



Universiteit
Leiden
The Netherlands

Boosting mass spectrometry-based analytics for biopharma

Gstöttner, C.J.

Citation

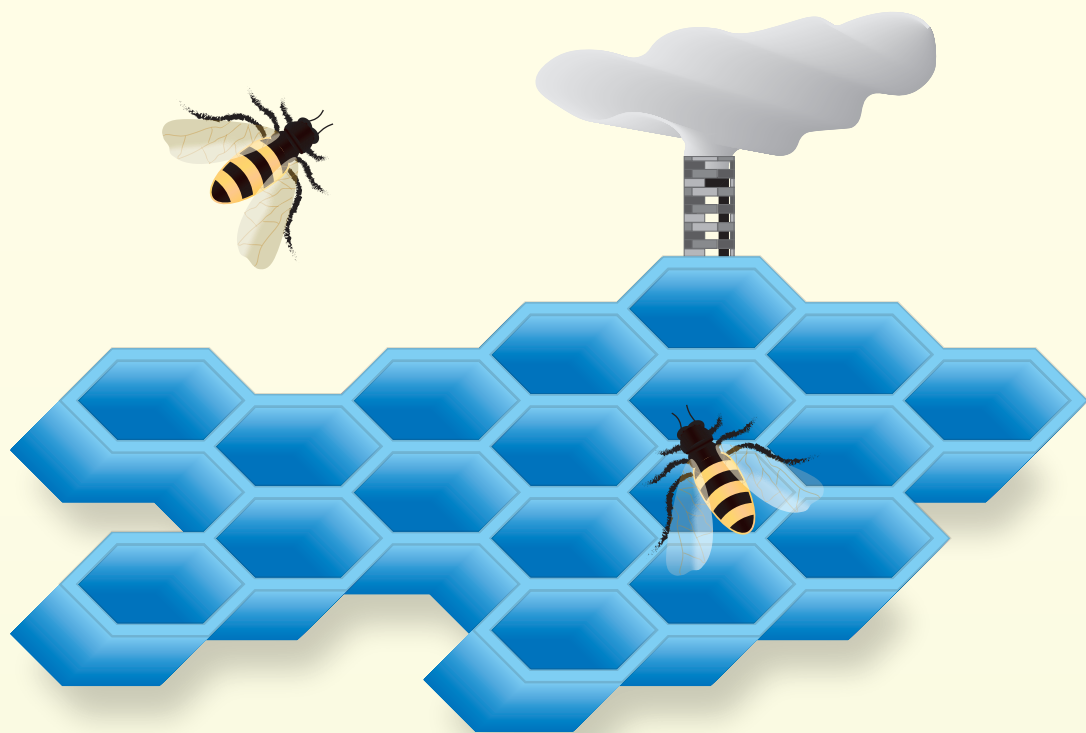
Gstöttner, C. J. (2021, November 30). *Boosting mass spectrometry-based analytics for biopharma*. Retrieved from <https://hdl.handle.net/1887/3245884>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3245884>

Note: To cite this publication please use the final published version (if applicable).



Boosting mass spectrometry-based analytics for biopharma

Christoph Johann Gstöttner



Boosting **mass spectrometry-based** **analytics for biopharma**

Christoph Johann Gstöttner

ISBN: 978-94-6419-333-6

© 2021 Christoph Johann Gstöttner. All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means without permission of the author or the journals holding the copyrights of the published manuscripts. All published material was reprinted with permission.

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

This work was supported by the European Commissions Horizon 2020 project “Analytics for Biologics” (Grant agreement ID 765502).

Cover Design and Layout: Christoph Johann Gstöttner

Printed by: Gildeprint BV, Enschede, The Netherlands

Financially supported by: Roche Diagnostics GmbH, Penzberg, Germany

Boosting mass spectrometry-based analytics for biopharma

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 30 november 2021
klokke 16.15 uur

door

Christoph Johann Gstöttner
Geboren te Zwiesel, Duitsland
op 25 januari 1993

Promotor: Prof. dr. M. Wuhrer

Co-promotor: Dr. E. Domínguez Vega

Leden promotiecommissie: Prof. dr. R.E.M Toes

Prof. dr. R. P.H. Bischoff,
*Department of Analytical
Biochemistry, Rijksuniversiteit Groningen,
The Netherlands*

Dr. I. Kohler,
*Division of BioAnalytical Chemistry,
Amsterdam Institute of Molecular and Life Sciences,
Vrije Universiteit Amsterdam, The Netherlands*

Dr. A. Gargano,
*Faculty of Science, Van 't Hoff Institute for Molecular
Sciences, university of Amsterdam, The Netherlands*

"If at first the idea is not absurd, then there is no hope for it."

Albert Einstein

Table of Contents

Chapter 1	Introduction	9
Chapter 2	Fast and automated characterization of antibody variants with 4D-HPLC/MS	35
Chapter 3	Fast analysis of antibody-derived therapeutics by automated multidimensional liquid chromatography – mass spectrometry	53
Chapter 4	Monitoring glycation levels of a bispecific monoclonal antibody at subunit level by ultrahigh resolution MALDI FT-ICR mass spectrometry	71
Chapter 5	Characterization and prediction of positional 4-hydroxyproline and sulfotyrosine, two post-translational modifications that can occur at substantial levels in CHO cells-expressed biotherapeutics	97
Chapter 6	Intact and subunit-specific analysis of bispecific antibodies by sheathless CE-MS	125
Chapter 7	Sheathless CE-MS as a tool for monitoring exchange efficiency and stability of bispecific antibodies	147
Chapter 8	Affinity capillary electrophoresis – mass spectrometry as tool to unravel proteoform-specific antibody-receptor interactions	159
Chapter 9	Structural and functional characterization of SARS-CoV-2 RBD domains produced in mammalian cells	183
Chapter 10	Discussion and future perspectives	207
	Nederlandse Samenvatting	222
	English Summary	225
	Curriculum Vitae	229
	List of Publications	230
	PhD Portfolio	233
	Acknowledgments	236



Chapter 1

Introduction



Biopharmaceutical antibodies and new antibody formats

The word biopharmaceuticals is very generic and comprises every pharmaceutical produced by living organisms. That can range from nucleic acids over proteins to complete cells or even blood products, which themselves contain a mixture of several different bioactive compounds. One large class of protein biopharmaceuticals are antibodies, which naturally are produced by immune cells (B cells) and are present in high amounts in the blood. These antibodies - also called immunoglobulins or in short Ig - can be further divided into five isotypes in humans, namely IgG, IgM, IgA, IgE and IgD. IgG is by far the most abundant isotype in blood plasma (more than 70%) and binds with very high specificity and affinity to its antigen on a pathogen (opsonization), activating an immune response and thereby allowing to fight and survive infections. Already in 1981 scientists successfully tested an intravenous immunoglobulin G (IVIg) therapy, based on captured IgGs from healthy donors, to treat children with idiopathic thrombocytopenic purpura¹.

To fight the wide variety of exogenous substances and pathogens, our body produces a plethora of different IgG antibodies - i.e. different clones, each specific for a particular antigen. To increase antigen specificity and treatment efficacy, scientists pursued therapies based on one single clone (monoclonal) antibody (mAb) instead of polyclonal IgG preparations. The development of hybridoma technology (where specific B cell clones are fused with immortalized myeloma tumor cells) revolutionized biopharmaceutical industry due to its capacity to recombinantly produce large amounts of mAbs and is currently the base for many modern mAbs². One of the first marketed mAbs was used to treat diffuse large cell non-Hodgkin's lymphoma. It was applied in combination with chemotherapeutics and showed a significantly elevated cure rate of cancer patients. Since then, around 50,000 patients with this disease are cured each year³.

Until now, biopharmaceutical companies have developed and marketed 100 mAbs⁴ for various types of cancer, autoimmune diseases, cardiovascular or neurologic diseases. Conventional mAbs are large proteins consisting of four subunits (two identical light chains (Lc) and two identical heavy chains (Hc)) with a molecular weight of around 150 kDa. However, pharma industry is continuously improving antibodies (*e.g.* higher stability, binding avidity or improved pharmacokinetics) and conventional mAbs are being slowly displaced by new antibody formats with new or enhanced properties. These recently emerging new antibody formats range from small nanobodies⁵ over Fc fusion proteins⁶ to bispecific antibodies (BsAbs) and elongated versions thereof⁷, allowing to bind two different antigens.

Bispecificity, in particular, opened a myriad of new possibilities of therapeutic applications. An example is the ability to block two different receptors and thereby prevent their activation, allowing to block cell proliferation or inflammatory processes⁸. This mechanism can also be used for cancer therapy to bring cancer cells and immune cells in close proximity⁸. Treatment

with the BsAb Emicizumab was shown to be effective in patients with haemophilia A by replacing the function of the missing factor VIII and mitigating the interaction between factor XI and X⁹. However, next to the benefit of targeting two molecules, bispecificity comes also with major biotechnological and analytical challenges due to the higher structural complexity of these molecules.

Functional BsAbs consist, instead of two, of four different chains: two different Lcs and Hcs (**Figure 1**). Proper assembly of these chains is not trivial and different antibody engineering strategies have been developed over the years. A well-known one is the “*knob-into-hole*” technology which guides the correct assembly of the Hcs^{10, 11} (**Figure 1**). To this end, a small amino acid in the CH3 domain of one Hc is exchanged by a large one and vice versa in the second Hc. This way, a knob and a hole are generated, allowing only one possibility to assemble the heavy chains. To further enhance the heterodimer formation of both heavy chains, additional amino acid exchanges, allowing ionic interactions and hydrogen bond formation, can be introduced¹². Regarding the correct assembly of the different Lcs, the so called “*CrossMab*” technology has been implemented¹³⁻¹⁶. The clue behind this approach is to swap either the complete Fab part on one Hc and its Lc or exchange only the variable part between one Lc and one Hc (**Figure 1**). Another simplistic and broadly established approach is the controlled Fab-arm exchange where BsAbs are generated by reassembly of antigen-binding arms of individually expressed monospecific parental antibodies containing two point mutations¹⁷. As bispecificity is generally enough to induce the desired therapeutic response, BsAbs often contain additional modifications in the CH2 domain of the Hc to abrogate the binding to Fc receptors and thereby abolish the immune response. Typical modifications are the exchange of the amino acids L234A, L235A, P329G, known in short as LALA-PG mutation¹⁸.

Despite all these biotechnological advances, production of BsAbs often comes with misassembled side products, such as light chain dimers, wrongly assembled heavy or light chains or entire monospecific antibodies (either covalently or non-covalently bound), which have to be carefully monitored, both qualitatively and quantitatively^{16, 19, 20}. Especially the latter ones are considered critical because of their increased immunogenicity, due to lower stability and, therefore, higher probability to form aggregates²¹.

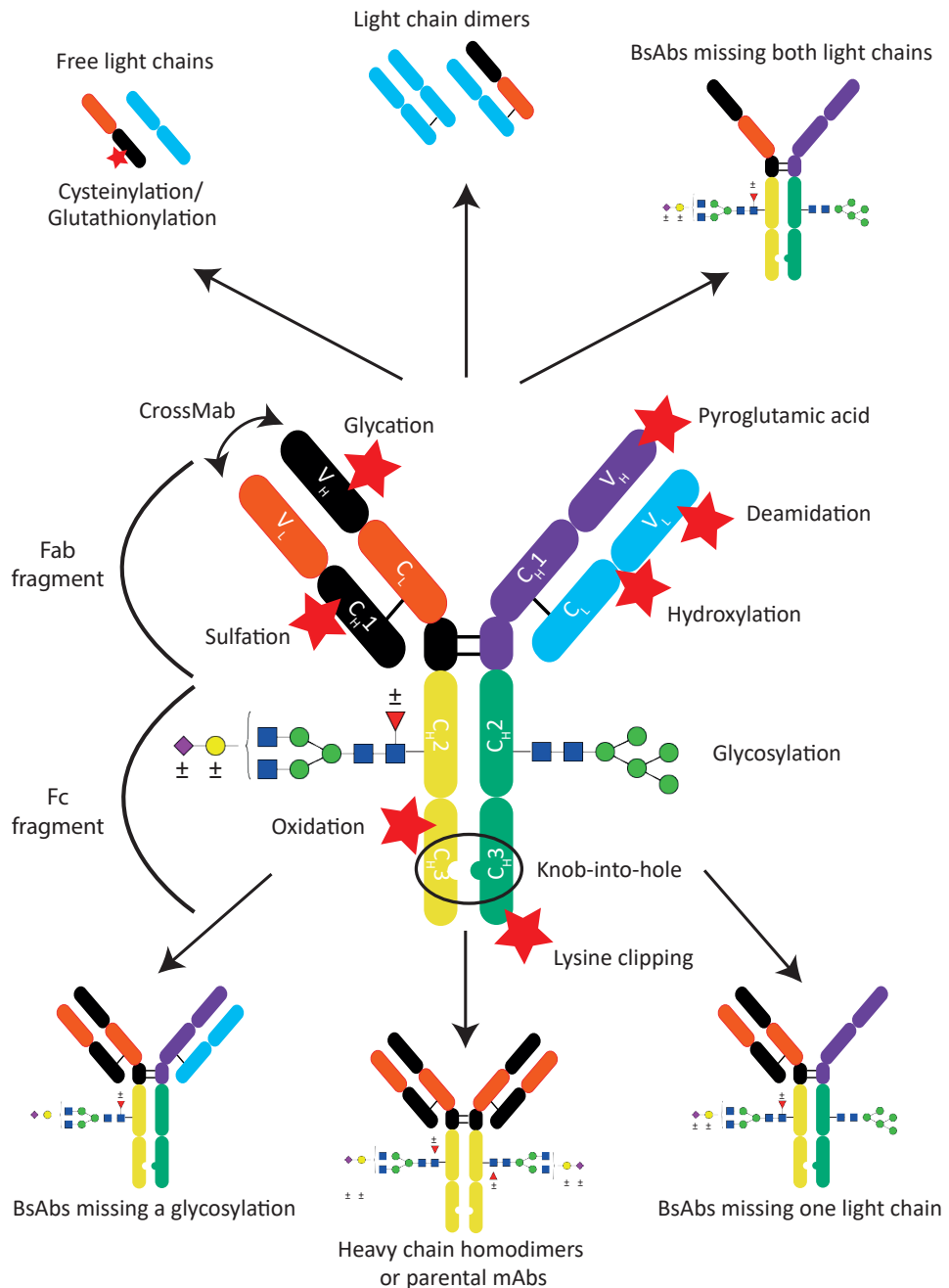


Figure 1. Common heterogeneity and modifications of BsAbs. BsAbs have a IgG-like blueprint containing two disulfide bridges connecting the two heavy chains and two additional connecting the corresponding heavy and light chain. The central BsAb illustrate typical PTMs and modifications from the engineering process. Misassembled products and antibodies missing a glycosylation site are depicted in the side figures.

1.1 Antibody proteoforms and their functional variability

Recombinantly (and naturally) produced antibodies often comprise several post-translational modifications (PTMs), resulting in a mixture of different “proteoforms” instead of a single molecular product. Most antibodies contain a conserved N-glycosylation site at Asn297 of the Hc. Therapeutic mAbs mainly exhibit complex type glycans (left glycan in **Figure 1**) and to a minor extent high mannose glycans (right glycan in **Figure 1**). Depending on the production system, the heterogeneity and abundance of certain glycan types can be different²². Additionally to glycosylation, other modifications can appear during the production or storage of the antibody. Often encountered antibody PTMs are C-terminal lysine clipping, sulfation, hydroxylation, N-terminal and lysine glycation, deamidation of asparagine residues, oxidation of methionine residues, N-terminal pyroglutamic acid formation and cysteinylolation or glutathionylation of cysteine residues²³⁻²⁵ (**Figure 1**). These modifications can influence the physicochemical properties of the antibody but also its functional characteristics and, therefore, should be carefully assessed.

Antibodies consist of a variable region specific for antigen binding and a constant region which determines the antibody half-life and activates effector cells (via binding to Fc receptors (FcRs)). Therefore, not only the nature but also the position of the modification can determine the functionality, stability and pharmacokinetic behavior of an antibody proteoform. An example is deamidation of asparagine residues, which was shown to occur with high probability in the complementary-determining region (CDR) of antibodies²⁶. This can result in reduced antigen-binding affinity lowering drug potency²⁷⁻³⁰. Another PTM which frequently occurs during antibody production or long term storage in formulation buffer, is glycation^{31, 32}. Similar to deamidation, it was shown that depending on the position glycation can impact antigen binding^{33, 34}.

PTMs occurring in the Fc region of antibodies do not directly influence the binding of the antigen but rather influence their pharmacokinetic behavior or effector functions. Oxidation of methionine residues mainly occurs on two conserved residues in the CH2 and CH3 domains, namely Met252 and Met428^{35, 36}. Oxidation of these residues alters the structure of the Fc region³⁷, resulting in a detrimental impact on the binding affinity to the neonatal Fc receptor (FcRn)³⁸⁻⁴⁰ and to a minor extent to other Fc gamma receptors (FcγR)^{38, 41}. As the FcRn is responsible for recycling antibodies, a lower FcRn binding affinity shortens the half-life of oxidized proteoforms³⁹. Also, glycosylation can influence the pharmacokinetics of antibodies⁴². Whereas aglycosylated antibodies do not show a significant difference in their half-life compared to the glycosylated antibody⁴³⁻⁴⁵, certain glycoforms can have a drastic impact. Some examples are the high mannose glycoforms^{46, 47}, which are removed by the mannose receptor⁴⁶⁻⁴⁸ or hypersialylated antibodies, which are cleared from circulation more slowly due to lower binding to asialoglycoprotein receptor⁴⁹. Controversially, other

studies show no difference between sialylated or non-sialylated antibodies^{50, 51}.

Glycosylation also plays a critical role in the pharmacodynamics of antibodies. These pharmacodynamic effects are depending on the interaction of antibodies with Fcγ receptors and C1q, activating thereby immune cells. This results in complement-dependent cytotoxicity (CDC), antibody-dependent cellular phagocytosis (ADCP) or antibody-dependent cell-mediated cytotoxicity (ADCC)⁵². For example, aglycosylated or hemi-glycosylated antibodies show reduced binding to Fcγ receptors⁵³⁻⁵⁵ or C1q⁵⁶, abrogating (or reducing in case of hemi-glycosylated antibodies) CDC and ADCC^{53, 55} activity. Within glycoforms afucosylation shows a 50 fold increase in binding affinity to the FcγRIIIa receptor^{57, 58} and subsequently an increase in the ADCC activity⁵⁹⁻⁶¹. Complex type glycans normally contain a core fucose, however, some antibodies are engineered to lack this fucose which increases their potency⁶². A similar but less pronounced effect was seen for high mannose glycoforms⁶³. Additionally, high mannose glycoforms decrease CDC⁶³ response and have a lower binding affinity to FcγRIIIa and FcγRIIb⁴⁶. These examples show that next to the structural characterization, the determination of binding affinities for each specific proteoform is of utmost importance.

1.2 Mass spectrometry for antibody characterization

The large diversity of antibody proteoforms, side products and engineering processes makes their analysis and characterization not trivial. Mass spectrometry (MS) is a powerful tool for the detailed characterization of antibody proteoforms. In biopharmaceutical industries MS is particularly used in early and late-stage development or for characterization of new occurring peaks during product release. Recently, first reports about validation and implementation of MS based QC methods are also being published^{64, 65}. Yet, currently, there is no universal method to assess all the heterogeneity and different approaches are available depending on the antibody and the question to answer. **Figure 2** illustrates the current approaches for MS antibody characterization. *Bottom-up* approaches rely on protein digestion using endoproteases and analysis of the obtained peptides by MS, usually combined with tandem MS (MS/MS) (**Figure 2A**). *Middle-up* approaches often consist of a limited cleavage of the mAb in the hinge region and analysis of the resulting antigen-binding fragment ((Fab)₂) and crystallizable fragment (Fc) (**Figure 2B**). To obtain information on the N- and C- termini, these subunits can be further fragmented by different fragmentation techniques (*middle-down*). mAbs can also be analyzed at the *intact level* without or with fragmentation (*top-down*) (**Figure 2C**). Next to structural, the *functional characterization* of mAbs can be performed using MS-based approaches, which can be considered the 4th level of antibody characterization (**Figure 2D**). The utility of these approaches will be discussed in the following sections.

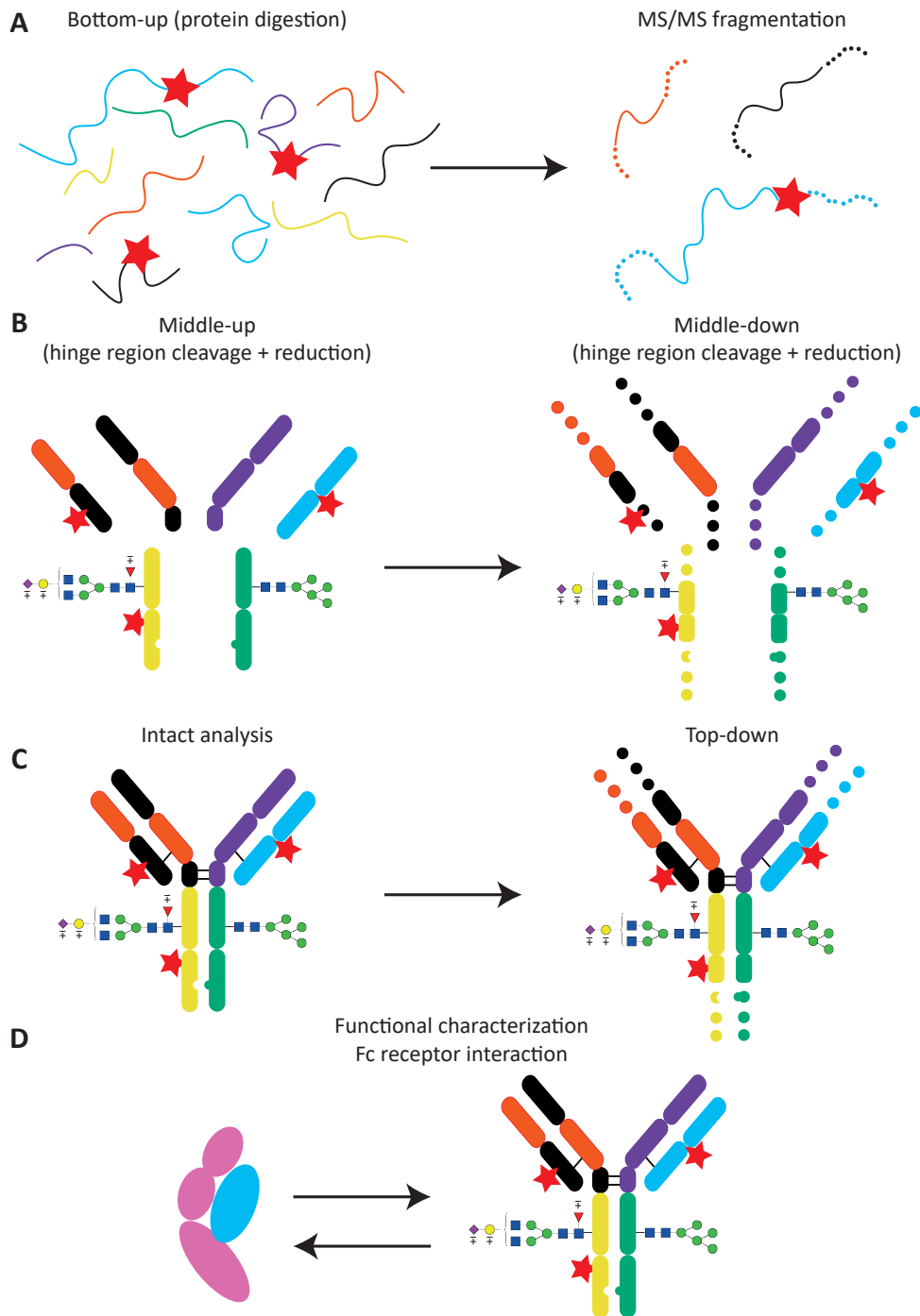


Figure 2: Different levels of mAb analysis. A) Bottom-up and MS/MS fragmentation, B) Middle-up/down, C) intact and Top-down D) functional interaction.

For the analysis of antibodies mainly two ionization techniques are used, electrospray ionization (ESI) and matrix assisted laser desorption/ionization (MALDI). In MALDI the sample is spotted together with a matrix on a MALDI target plate and allowed to crystallize. The crystallized samples are then transferred to the vacuum and targeted with a laser which leads to desorption and thereby ionization of the sample. Depending on the matrix used, MALDI is considered a soft ionization technique, which allows to detect labile proteoforms⁶⁶. Another characteristic of MALDI is, that it mainly generates singly charged ions, which simplify data analysis⁶⁷. However, to enhance the mass accuracy and dynamic range of the MS - opening the possibility to analyze larger proteins - the generation of multiply charged ions is sometimes favorable⁶⁸⁻⁷⁰. Fragmentation in MALDI can either be performed post-source or by in-source decay (MALDI-ISD). Whereas post-source fragmentation is not efficient for large proteins as mainly singly charged ions are generated in MALDI, MALDI-ISD results in efficient fragmentation and good sequence coverage⁷¹. Similar to the intact MALDI analysis the resulting fragment ions are mainly singly charged, simplifying data analysis. This fragmentation technique has shown to be very powerful for conventional antibody top-down or middle-down characterization⁷²⁻⁷⁴. A drawback of this technique is that no precursor ions can be selected, but rather many molecules present in the sample are fragmented simultaneously, challenging data analysis of complex mixtures.

In ESI, the molecule of interest (dissolved in a liquid) is dispersed into the ionization chamber via a sprayer needle onto which a high voltage is applied (often assisted by a coaxial gas flow and heat), resulting in charged droplets and via desolvation in ionized proteins in the gas phase^{75, 76}. In contrast to MALDI, where mainly singly charged ions are formed, ESI delivers multiply charged ions. Fragmentation of these ions can either be performed in the source, although these examples are rather rare for proteins⁷⁷, or more commonly after the ionization process. This often involves selection of precursor ions and, thereby, targeted fragmentation. Most fragmentation techniques for bottom-up are high-energy collisional dissociation (HCD) or collision-induced dissociation (CID), while for top/middle-down mainly ultraviolet photodissociation (UVPD), electron-capture dissociation or electron-transfer dissociation (ETD) are used⁷⁸. For detection of fragment ions and intact proteins, commonly time of flight (TOF)⁷⁹, Fourier transform ion cyclotron resonance (FTICR)⁸⁰ or Orbitrap⁸¹ mass analyzers are used.

Upfront separation of proteoforms with small mass differences is beneficial for confident structure assignments. Whereas MALDI-MS is mainly used as stand-alone technique, ESI-MS is commonly combined with an upfront separation technique^{82, 83}. Separation techniques that can be directly coupled with MS and that have shown to be helpful for antibody characterization - and to a lesser extent new antibody formats - are reversed phase liquid chromatography (RPLC), hydrophilic interaction chromatography (HILIC), ion exchange chromatography (IEC) and size exclusion chromatography (SEC)^{20, 21, 84-87}. Another technique that is gaining attention (especially for middle-up and intact antibody characterization) is

Capillary Electrophoresis (CE)-MS. Hyphenation of CE with MS has been traditionally done by using a sheath-liquid to close the electric circuit. However, this sheath-liquid interface leads to dilution of the sample providing limited sensitivity. Different solutions have been developed to overcome this problem, such as nanoflow sheath-liquid or sheathless interfacing, which have demonstrated similar performance but boosted sensitivity compared to conventional sheath-liquid CE-MS^{88, 89}.

1.3 Current analytical challenges in biopharmaceutical industry

While several MS-based analytical platforms have been developed and implemented for antibody characterization in the biopharmaceutical industry, the continuous development of novel antibody therapeutics opens a demand for new and improved analytical methods. Some of the current analytical challenges addressed in this thesis are described below.

1.3.1 Automation of bottom-up approaches

With the rising number of antibody applications and additional possibilities arising from new antibody formats, the pipelines of pharmaceutical companies are heavily filled. This also implies an increase in molecules that need to be characterized in various stages from early research until approval by the authorities, formulation development or stability studies using stress conditions (*e.g.* oxidative stress, high temperature or high levels of glucose)²⁹.

Bottom-up approaches combined with LC-MS are the gold standard for antibody PTM assessment^{29, 90, 91}. They are commonly employed in pharmaceutical companies for the initial characterization of mAbs to define CQAs (arising during elongated storage or forced degradation) and during QC release testing of marketed antibodies to characterize unknown peaks. Product variants and PTMs need to be confidentially assigned and a risk assessment performed to ensure patient safety and drug efficacy.

Typical sample preparation in bottom-up approaches require multiple steps (denaturation, disulfide bridge reduction, alkylation and enzymatic digestion) resulting in extensive hands-on and incubation times (~18 hours for trypsin) and increasing the risk of inducing artificial modifications⁹²⁻⁹⁴. For chromatographic peak characterization the process is even more complex as the peak needs to be tediously fractionated prior bottom-up characterization. Such workflows can easily take several days and often lack a confidential peak assignment due to overlapping with other proteoforms.

Robotic platforms are useful to save hands-on time on sample preparation (reduction alkylation and digestion) and contribute to standardization such as reducing sample handling errors. Such approaches can process up to 96 samples simultaneously and reduce digestion time to 4 hours⁹⁵. Still, samples require further LC-MS analysis which leads to additional

storage time. Despite the full pipelines of pharmaceutical industries, 96 samples are barely analyzed at once, but rather 10 to 20 samples, not ideal for a robotic platform which requires experienced users and time to program. In short, an intermediate automated solution for bottom-up sample preparation and analysis (and applicable for QC peak characterization) is still missing.

Using multidimensional LC, few steps have been made to perform an online digestion of protein, either by mixing the trypsin and the sample in a pressurized sample loop⁹⁶ or by developing immobilized trypsin reactors (IMERs). Such multidimensional systems with an on-line trypsin digestion step, followed by a RPLC separation and MS detection have been described to characterize apolipoproteins⁹⁷, protein mixtures and antibodies^{98, 99}. However, these systems are not user friendly and they are not able to fractionate peaks from a pre-separation. Furthermore, none of these systems permits protein reduction, resulting in peptides connected by disulfide bridges, thereby complicating data analysis.

1.3.2 Complexity of new antibody formats and unexpected modifications

Although pharmaceutical labs have established an armament of analytical techniques for characterization of conventional mAbs, new antibody formats comprise higher complexity bringing new analytical challenges.

In the particular case of BsAbs, additional side products are often encountered as consequence of the misassembling of different antibody chains – *i.e. macroheterogeneity*. Especially of interest are the homodimers of Hcs, which can show high immunogenicity²¹. To analyze this type of side products sample preparation should be minimized and the analysis needs to be performed at the intact level. Bottom-up approaches are not suitable for monitoring BsAb misassembling as high-order structures are lost during endoprotease digestion. MS has been extensively applied for the analysis of intact conventional mAbs. However, the macroheterogeneity can be compromised in stand-alone MS as aggregate artifacts or disassembling can happen during the ESI process²⁰. Hyphenation with a separation technique allows discrimination of pre-existing complexes from ESI artifacts due to the introduction of an orthogonal dimension. SEC-MS permits to monitor free chains, incomplete assemblies (*e.g.* BsAbs missing a Lc or Hc) and aggregates. However, different antibodies (*i.e.* homodimers or parental monospecific antibodies depending on the production mechanism) can not be resolved by SEC as they are in the same mass range²⁰. Although MS can provide mass distinction between antibodies, homo- and heterodimers often exhibit similar masses resulting in signal superposition and difficult assignment. Chromatographic techniques suitable for homodimer and parental antibody separations are hydrophobic interaction chromatography (HIC) and IEC, however, they normally require high amount of non-volatile salts in the mobile phases making hyphenation to MS not straightforward^{84, 85}. CE-MS has shown great potential in the characterization of intact and

hinge region cleaved conventional mAbs^{100, 101}, although the applicability for BsAbs has never been explored so far.

Next to this, new formats can undergo non-conventional or unexpected modifications due to different exposure of amino acids, presence of additional linker elements or amino acid exchanges compared to conventional antibodies¹⁰²⁻¹⁰⁴. For instance, an unexpected modification that can occur during antibody production is hydroxylation of amino acids such as lysine or proline. Commonly these modifications are observed in proteins such as collagen to enhance the water solubility¹⁰⁵. However, recently they have been found on antibodies and more frequently on new antibody formats^{103, 106} and seem to occur during fermentation processes through protein translation errors and misincorporation of hydroxyproline instead proline¹⁰⁷. Because the mass difference is exactly identical to that of oxidation, characterization of this modification is further complicated if there is an amino acid prone to oxidation nearby (*e.g.* methionine, tryptophan, cysteine or histidine).

Localization of PTMs is also an important aspect, as the position of the modification will determine the effect on functionality. Localization can be focused on antibody chains (Fab₂, Fc, Hc, Lc) or amino acid level. For BsAbs, chain localization is especially important as small variations introduced during BsAb engineering can result in different heterogeneity of counter chains (*e.g.* Hc and Hc'). Here, bottom-up workflows can, in some cases, result in equal peptides hampering localization to the specific chain. Middle-up/down approaches - where the BsAb is reduced and/or partially cleaved using hinge region endoproteases - can be a solution for these cases (**Figure 2B**). Techniques which have proven to be suitable for subunit analysis are HILIC- or RP-MS^{21, 86, 108}, CE-MS^{109, 110} and MALDI-MS^{73, 74}. To localize PTMs at the amino acid level top-down or middle-down approaches are employed^{72-74, 111}. Top-down applications in particular require minimal sample preparation (*i.e.* desalting or buffer exchange), thereby decreasing the risk of introducing unintended PTMs. More challenging is the localization of labile PTMs (*e.g.* sulfation) as they can be lost during standard "harsh" fragmentation, such as CID or HCD^{112, 113}. A way to localize sulfation can be to use a soft fragmentation techniques such as ETD or a combination between ETD and CID (ETcID)¹¹², which allows in some cases to keep the PTM on the protein backbone. However this technique requires enrichment of sulfated proteoforms to enable localization of the PTM, because still the sulfation is lost partially and additionally the fragmentation is in some cases not trivial. Even though there are several intact, top-down or middle-up/down applications for conventional antibodies the ability of these platforms and approaches for the characterization of the more complex BsAbs has barely been explored.

1.3.3 Binding assessment in a proteoform-specific manner

After monitoring different modifications and proteoforms the next question arising is how these modifications influence the characteristics and functionality of the therapeutic product. As discussed at the beginning of this section, these proteoforms can have a severe impact

on pharmacokinetics and pharmacodynamics and the assessment of the affinity for each single proteoform is therefore very important. Unfortunately, approaches for the functional characterization of individual proteoforms in mixtures are largely missing. Techniques used traditionally to determine binding affinities are surface plasmon resonance (SPR) or enzyme-linked immunosorbent assay (ELISA). These techniques, however, are not able to distinguish between different proteoforms and give an overall response for all proteoforms in a sample. To study the binding affinity of specific proteoforms, either isolation or proteoform specific production in an “as pure as possible” form is required^{39, 60, 114-116}. However, these processes are very tedious and can still induce confounding by the presence of small amounts of other (glyco)variants^{57, 116}. Affinity LC tried to overcome these drawback and columns for the assessment of FcRn and FcγRIIIa have been developed¹¹⁷⁻¹²⁰. This allowed to determine the relative binding of different mAb proteoforms in mixtures. However, columns for all Fc receptors or other receptors are not available and the determination of quantitative binding affinities is not possible and it requires pH stress for elution. All the above mentioned approaches have the immobilization of the receptor in common. How far this reflects the natural situation, which in some cases requires dimerization of the receptor before binding is questionable. ESI-MS has also been employed for monitoring protein-binding in a proteoforms specific manner^{121, 122}. However, gas phase binding often deviates from in solution as many complexes are gas-phase unstable¹²³. Therefore, immobilization-free in-solution binding techniques that provide affinity constants in a proteoform-selective manner are desired but still missing.

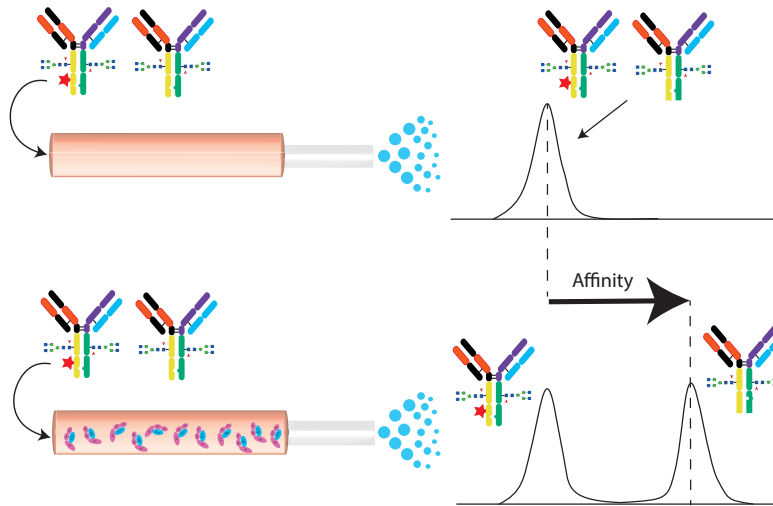


Figure 3: Illustration of the principle of mobility-shift affinity CE. Example of two mAb proteoforms without (upper panel) and with an interaction partner in the BGE (lower panel). The example shows a separation of a antibody proteoform which does not interact with the receptor (with red star), whereas the other antibody proteoform migrates later, due to its affinity to the receptor. This order of migration is only the case if the receptor has a lower electrophoretic mobility compared to the antibody.

One particular appealing mode of affinity CE (ACE) for proteoform-specific binding assessment is the mobility shift approach. In this approach, the capillary is filled with one interaction partner followed by an injection of the second one. After applying an electric field, the proteoforms that do not interact will maintain their electrophoretic mobility while proteoforms interacting with the receptor will change their electrophoretic mobility migrating in a different place (**Figure 3**). This technique has been used to study the interaction of proteins with metal ions^{124, 125} and demonstrated the ability to obtain quantitative binding information in equilibrium constants (K_d). Yet, these studies were conducted without MS detection, excluding the possibility to determine the affinity of overlapping proteoforms in a mixture simultaneously (*e.g.* oxidation or glycosylation). Protein-protein ACE-MS was for the first time shown in an exploratory study for the interaction between two small test proteins¹²⁶. The system was based on sheath-liquid CE-MS providing limited sensitivity and hampering the applicability to larger or more complex proteins such as antibodies and Fc receptors. Exploiting and developing ACE-MS for proteoform-specific antibody Fc-binding would overcome one of the major current challenges in biopharma characterization.

1.4 Scope of the thesis

The aim of the research described in this thesis is the development of new MS-based analytical platforms that permit to address current limitations and challenges in the characterization of antibody and antibody-derived therapeutics as described in **Chapter 1**.

For detailed structural characterization of therapeutic proteins including PTM assessment commonly bottom-up proteomics strategies are applied. The analytical methodologies are well-established for antibody samples and identification of unknown peaks for quality control (QC) purposes, however sample preparation requires extensive hands-on time - from several hours to days - and increases the risk of inducing unintended modifications. In this thesis this drawback is addressed by integrating all the steps in an automatized multidimensional LC-MS (mD-LC-MS) platform. In **Chapter 2**, a mD-LC-MS method was developed for the on-line bottom-up characterization of peaks separated on ion exchange chromatography. The approach drastically reduced the total analysis time and proved several benefits over conventional offline sample preparation. In **Chapter 3** the novel mD-LC-MS platform was further optimized to increase retention of highly polar peptides by adaptation of the flow rates of the different dimensions. The optimized method provided sequence coverages comparable to the standard offline workflow and allowed detection of hydrophilic PTMs missed prior to optimization. The mD-LC-MS approach was intended to directly analyze formulated mAb samples without pre-separation. The developed platforms are currently in use in various pharma laboratories.

In contrast to bottom-up approaches - requiring endoprotease digestion - top-down or middle-down MALDI-ISD approaches offer the possibility to sequence N- and C-termini of mAbs and subunits (Hc and Lc) with minor sample preparation. **Chapter 4** shows the ability of MALDI-(ISD)-FT-ICR-MS for the analysis of new antibody formats with a higher degree of complexity, such as BsAbs. Next to sequence information, PTMs close to the N- or C-terminus could be localized by MALDI-ISD. Due to the “softer” fragmentation process, very labile modifications, as shown for sulphation in **Chapter 5**, can be retained on the protein backbone and confidently assigned.

Next to classical PTMs, production of BsAbs also yields misassembled side products, which might lead to immunogenic side reactions. To monitor mispairing, the native structure of the antibodies have to be maintained during the analysis (*i.e.* at the intact level). CE-MS has shown to be a powerful technique for the characterization of intact proteins and was employed in **Chapter 6** to analyze two homologous BsAbs. Next to their intact analysis the BsAbs were analyzed at the middle-up level to localize PTMs to specific subunits (Lc1, Lc2, Fd'1, Fd'2, (Fc/2)1 or (Fc/2)2). Similarly, **Chapter 7** provides a sheathless CE-MS approach to monitor the exchange efficiency of in-house produced BsAbs. Here, the presence of highly heterogenic Fab-glycosylated antibodies required a separation prior to MS characterization.

The functional characterization of antibodies in a proteoform-selective way remains one of the major unaddressed challenges of pharmaceutical companies, which still relies on complex engineering and/or isolation strategies. **Chapter 8** presents an innovative approach based affinity CE-MS for the simultaneous assessment of structure and function of antibodies. The method permitted to determine affinities in a proteoform-specific manner and to study binding stoichiometry. The approach was employed to monitor the influence of oxidation on the interaction with the neonatal Fc receptor.

In the course of this PhD thesis, the COVID-19 pandemic appeared, imposing a drastic change of habits and priorities in our lives but also in our research. Recombinant SARS-CoV-2 proteins are essential instruments in the fight against COVID-19. In **Chapter 9**, the receptor binding domains (RBD) of the S protein produced in two different mammalian systems (HEK293 versus CHO) were characterized. Here, the lack of previous structural information instigated the necessity of a multilevel MS approach (intact, glycopeptide and released glycans). Also, the functional differences between both RBD samples, such as binding to anti RBD antibodies and to the ACE2 receptor, were evaluated.

In the last chapter, **Chapter 10**, the results are discussed with special emphasis on the aspects of automation, top/middle-down and native MS, CE and especially affinity CE in an industrial setting. In addition, the potential and benefit of collaborations between industry and academia are discussed.

1.5 References

1. Imbach, P., *et al.* High-dose intravenous gammaglobulin for idiopathic thrombocytopenic purpura in childhood. *Lancet* (London, England) 1981, 1228-31.
2. Köhler, G., Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 1975, 256, 495-497.
3. Grillo-López, A. J. The first antibody therapy for cancer: a personal experience. *Expert Rev Anticancer Ther* 2013, 13, 399-406.
4. Mullard, A. FDA approves 100th monoclonal antibody product. *Nat Rev Drug Discov* 2021.
5. Yang, E. Y., Shah, K. Nanobodies: Next Generation of Cancer Diagnostics and Therapeutics. *Front Oncol* 2020, 10.
6. Jafari, R., *et al.* Fc-fusion Proteins in Therapy: An Updated View. *Curr Med Chem* 2017, 24, 1228-1237.
7. Suurs, F. V., *et al.* A review of bispecific antibodies and antibody constructs in oncology and clinical challenges. *Pharmacol Ther* 2019, 201, 103-119.
8. Kontermann, R. E. Brinkmann, U. Bispecific antibodies. *Drug Discov Today* 2015, 20, 838-847.
9. Blair, H. A. Eficizumab: A Review in Haemophilia A. *Drugs* 2019, 79, 1697-1707.
10. Xu, Y., *et al.* Stafford, R. L.; Steiner, A. R.; Sato, A. K.; Hallam, T. J.; Yin, G., Production of bispecific antibodies in “knobs-into-holes” using a cell-free expression system. *MAbs* 2015, 7, 231-42.
11. Ridgway, J. B., Presta, L. G., Carter, P. ‘Knobs-into-holes’ engineering of antibody CH3 domains for heavy chain heterodimerization. *Protein Eng* 1996, 9, 617-21.
12. Moore, G. L., *et al.* A robust heterodimeric Fc platform engineered for efficient development of bispecific antibodies of multiple formats. *Methods* 2019, 154, 38-50.
13. Klein, C., Schaefer, W., Regula, J. T. The use of CrossMAb technology for the generation of bi- and multispecific antibodies. *MAbs* 2016, 8, 1010-20.
14. Schaefer, W., *et al.* Immunoglobulin domain crossover as a generic approach for the production of bispecific IgG antibodies. *Proc Natl Acad Sci USA* 2011, 108, 11187-92.
15. Klein, C., *et al.* Engineering therapeutic bispecific antibodies using CrossMab technology. *Methods* 2019, 154, 21-31.
16. Atwell, S., *et al.* Stable heterodimers from remodeling the domain interface of a homodimer using a phage display library. *J Mol Biol* 1997, 270, 26-35.
17. Labrijn, A. F., *et al.* Controlled Fab-arm exchange for the generation of stable bispecific IgG1. *Nat Protoc* 2014, 9 (10), 2450-63.
18. Lo, M., *et al.* Effector-attenuating Substitutions That Maintain Antibody Stability and Reduce Toxicity in Mice. *J Biol Chem* 2017, 292, 3900-3908.
19. Elliott, J. M., *et al.* Antiparallel conformation of knob and hole aglycosylated half-antibody homodimers is mediated by a CH2-CH3 hydrophobic interaction. *J Mol Biol*

- 2014, 426, 1947-57.
20. Habberger, M., *et al.* Rapid characterization of biotherapeutic proteins by size-exclusion chromatography coupled to native mass spectrometry. *MAbs* 2016, 8, 331-339.
 21. Woods, R. J., *et al.* LC-MS characterization and purity assessment of a prototype bispecific antibody. *MAbs* 2013, 5, 711-722.
 22. Werner, R. G., Kopp, K., Schlueter, M. Glycosylation of therapeutic proteins in different production systems. *Acta Paediatr* 2007, 96, 17-22.
 23. Jefferis, R. Posttranslational Modifications and the Immunogenicity of Biotherapeutics. *J Immunol Res* 2016, 2016, 5358272.
 24. Liu, H., *et al.* In vitro and in vivo modifications of recombinant and human IgG antibodies. *MAbs* 2014, 6, 1145-1154.
 25. Beck, A., Liu, H. Macro- and Micro-Heterogeneity of Natural and Recombinant IgG Antibodies. *Antibodies (Basel)* 2019, 8, 18.
 26. Lu, X., *et al.* Deamidation and isomerization liability analysis of 131 clinical-stage antibodies. *MAbs* 2019, 11, 45-57.
 27. Huang, L., *et al.* In vivo deamidation characterization of monoclonal antibody by LC/MS/MS. *Anal Chem* 2005, 77, 1432-9.
 28. Harris, R. J., *et al.* Identification of multiple sources of charge heterogeneity in a recombinant antibody. *J Chromatogr B Biomed Sci Appl* 2001, 752, 233-245.
 29. Habberger, M., *et al.* Assessment of chemical modifications of sites in the CDRs of recombinant antibodies: Susceptibility vs. functionality of critical quality attributes. *MAbs* 2014, 6, 327-39.
 30. Yan, B., *et al.* Succinimide formation at Asn 55 in the complementarity determining region of a recombinant monoclonal antibody IgG1 heavy chain. *J Pharm Sci* 2009, 98, 3509-3521.
 31. Fischer, S., Hoernschemeyer, J., Mahler, H. C. Glycation during storage and administration of monoclonal antibody formulations. *Eur J Pharm Biopharm* 2008, 70, 42-50.
 32. Wei, B., *et al.* Glycation of antibodies: Modification, methods and potential effects on biological functions. *MAbs* 2017, 9, 586-594.
 33. Kennedy, D. M., Skillbn, A. W., Self, C. H. Glycation of monoclonal antibodies impairs their ability to bind antigen. *Clin Exp Immunol* 1994, 98, 245-251.
 34. Mo, J., *et al.* Quantitative analysis of glycation and its impact on antigen binding. *MAbs* 2018, 10, 406-415.
 35. Chumsae, C., *et al.* Comparison of methionine oxidation in thermal stability and chemically stressed samples of a fully human monoclonal antibody. *J Chromatogr B* 2007, 850, 285-294.
 36. Lam, X. M., Yang, J. Y.; Cleland, J. L. Antioxidants for Prevention of Methionine Oxidation in Recombinant Monoclonal Antibody HER2. *J Pharm Sci* 1997, 86, 1250-1255.
 37. Liu, D., *et al.* Remmele, R. L., Structure and Stability Changes of Human IgG1 Fc as a Consequence of Methionine Oxidation. *Biochemistry* 2008, 47, 5088-5100.

38. Bertolotti-Ciarlet, A., *et al.* Impact of methionine oxidation on the binding of human IgG1 to Fc Rn and Fc gamma receptors. *Mol Immunol* 2009, 46, 1878-82.
39. Wang, W., *et al.* Impact of methionine oxidation in human IgG1 Fc on serum half-life of monoclonal antibodies. *Mol Immunol* 2011, 48, 860-866.
40. Gao, X., *et al.* Effect of Individual Fc Methionine Oxidation on FcRn Binding: Met252 Oxidation Impairs FcRn Binding More Profoundly than Met428 Oxidation. *J Pharm Sci* 2015, 104, 368-377.
41. Cymer, F., *et al.* Oxidation of M252 but not M428 in hu-IgG1 is responsible for decreased binding to and activation of hu-FcγRIIa (His131). *Biologicals* 2017, 50, 125-128.
42. Liu, L., Antibody glycosylation and its impact on the pharmacokinetics and pharmacodynamics of monoclonal antibodies and Fc-fusion proteins. *J Pharm Sci* 2015, 104, 1866-1884.
43. Jung, S. T., *et al.* Bypassing glycosylation: engineering aglycosylated full-length IgG antibodies for human therapy. *Curr Opin Biotechnol* 2011, 22, 858-867.
44. Liu, L., *et al.* Pharmacokinetics of IgG1 monoclonal antibodies produced in humanized *Pichia pastoris* with specific glycoforms: a comparative study with CHO produced materials. *Biologicals* 2011, 39, 205-10.
45. Clarke, S. J., *et al.* A phase I, pharmacokinetic (PK), and preliminary efficacy assessment of ALD518, a humanized anti-IL-6 antibody, in patients with advanced cancer. *J Clin Oncol* 2009, 27, 3025-3025.
46. Yu, M., *et al.* Production, characterization and pharmacokinetic properties of antibodies with N-linked Mannose-5 glycans. *MAbs* 2012, 4, 475-487.
47. Goetze, A. M., *et al.* High-mannose glycans on the Fc region of therapeutic IgG antibodies increase serum clearance in humans. *Glycobiology* 2011, 21, 949-59.
48. Alessandri, L., *et al.* Increased serum clearance of oligomannose species present on a human IgG1 molecule. *MAbs* 2012, 4, 509-520.
49. Bas, M., *et al.* Fc Sialylation Prolongs Serum Half-Life of Therapeutic Antibodies. *J Immunol* (Baltimore, Md. : 1950) 2019, 202, 1582-1594.
50. Naso, M. F., *et al.* Engineering host cell lines to reduce terminal sialylation of secreted antibodies. *MAbs* 2010, 2, 519-527.
51. Scallon, B. J., *et al.* Higher levels of sialylated Fc glycans in immunoglobulin G molecules can adversely impact functionality. *Mol Immunol* 2007, 44, 1524-1534.
52. Ravetch, J. V. Fc receptors. *Current opinion in immunology* 1997, 9, 121-5.
53. Jefferis, R. Recombinant antibody therapeutics: the impact of glycosylation on mechanisms of action. *Trends Pharmacol Sci* 2009, 30, 356-62.
54. Mimura, Y., *et al.* The influence of glycosylation on the thermal stability and effector function expression of human IgG1-Fc: properties of a series of truncated glycoforms. *Mol Immunol* 2000, 37, 697-706.
55. Jung, S. T., *et al.* Aglycosylated IgG variants expressed in bacteria that selectively bind FcγRI potentiate tumor cell killing by monocyte-dendritic cells. *Proc Natl Acad Sci*

- USA 2010, 107, 604-609.
56. Lund, J. *et al.* Multiple interactions of IgG with its core oligosaccharide can modulate recognition by complement and human Fc gamma receptor I and influence the synthesis of its oligosaccharide chains. *J Immunol* 1996, 157, 4963-9.
 57. Dekkers, G., *et al.* Decoding the Human Immunoglobulin G-Glycan Repertoire Reveals a Spectrum of Fc-Receptor- and Complement-Mediated-Effector Activities. *Front Immunol* 2017, 8.
 58. Shinkawa, T., *et al.* The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. *The Journal of biological chemistry* 2003, 278, 3466-73.
 59. Pereira, N. A., *et al.* The “less-is-more” in therapeutic antibodies: Afucosylated anti-cancer antibodies with enhanced antibody-dependent cellular cytotoxicity. *MAbs* 2018, 10, 693-711.
 60. Li, T., *et al.* Modulating IgG effector function by Fc glycan engineering. *Proc Natl Acad Sci USA* 2017, 114, 3485.
 61. Ferrara, C., *et al.* Modulation of therapeutic antibody effector functions by glycosylation engineering: influence of Golgi enzyme localization domain and co-expression of heterologous beta1,4-N-acetylglucosaminyltransferase III and Golgi alpha-mannosidase II. *Biotechnol Bioeng* 2006, 93, 851-61.
 62. Gerdes, C. A., *et al.* GA201 (RG7160): A Novel, Humanized, Glycoengineered Anti-EGFR Antibody with Enhanced ADCC and Superior *In Vivo* Efficacy Compared with Cetuximab. *Clin Cancer Res* 2013, 19, 1126.
 63. Kanda, Y., *et al.* Comparison of biological activity among nonfucosylated therapeutic IgG1 antibodies with three different N-linked Fc oligosaccharides: the high-mannose, hybrid, and complex types. *Glycobiology* 2007, 17, 104-18.
 64. Sokolowska, I., *et al.* 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* 2020, 92, 2369-2373.
 65. Schilling, M., *et al.* Development and validation of a platform reduced intact mass method for process monitoring of monoclonal antibody glycosylation during routine manufacturing. *Bioengineered* 2020, 11, 1301-1312.
 66. Lee, C., Ni, C.-K. Soft Matrix–Assisted Laser Desorption/Ionization for Labile Glycoconjugates. *J Am Soc Mass Spectrom* 2019, 30 (8), 1455-1463.
 67. Singhal, N., *et al.* MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Front Microbiol* 2015, 6, 791-791.
 68. Li, J., *et al.* Matrix Assisted Ionization: New Aromatic and Nonaromatic Matrix Compounds Producing Multiply Charged Lipid, Peptide, and Protein Ions in the Positive and Negative Mode Observed Directly from Surfaces. *J Am Soc Mass Spectrom* 2012, 23, 1625-1643.

69. Chan, K. W. S., Cook, K. D. Extended mass range by multiple charge: Sampling quadruply charged quasimolecular ions of poly(ethylene glycol) 4000. *Organic Mass Spectrom* 1983, 18, 423-425.
70. Choi, H., *et al.* Effects of Matrices and Additives on Multiple Charge Formation of Proteins in MALDI–MS Analysis. *J Am Soc Mass Spectrom* 2019, 30, 1174-1178.
71. Debois, D., *et al.* MALDI in-source decay, from sequencing to imaging. *Top Curr Chem* 2013, 331, 117-41.
72. Nicolardi, S., *et al.* Improved N- and C-Terminal Sequencing of Proteins by Combining Positive and Negative Ion MALDI In-Source Decay Mass Spectrometry. *Anal Chem* 2020, 92, 12429-12436.
73. van der Burgt, Y. E. M., *et al.* Structural Analysis of Monoclonal Antibodies by Ultrahigh Resolution MALDI In-Source Decay FT-ICR Mass Spectrometry. *Anal Chem* 2019, 91, 2079-2085.
74. Srzentić, K., *et al.* Interlaboratory Study for Characterizing Monoclonal Antibodies by Top-Down and Middle-Down Mass Spectrometry. *J Am Soc Mass Spectrom* 2020, 31, 1783-1802.
75. Bruins, A. P. Mechanistic aspects of electrospray ionization. *J Chromatogr A* 1998, 794, 345-357.
76. Ho, C. S., *et al.* Electrospray ionisation mass spectrometry: principles and clinical applications. *Clin Biochem Rev* 2003, 24, 3-12.
77. Ramos, A. A., *et al.* Tandem Parallel Fragmentation of Peptides for Mass Spectrometry. *Anal Chem* 2006, 78, 6391-6397.
78. Fornelli, L., *et al.* Accurate Sequence Analysis of a Monoclonal Antibody by Top-Down and Middle-Down Orbitrap Mass Spectrometry Applying Multiple Ion Activation Techniques. *Analytical chemistry* 2018, 90, 8421-8429.
79. Cameron, A. E., Eggers, D. F. An Ion “Velocitron”. *Rev Sci Instrum* 1948, 19, 605-607.
80. Nikolaev, E. N., Kostyukovich, Y. I., Vladimirov, G. N. Fourier transform ion cyclotron resonance (FT ICR) mass spectrometry: Theory and simulations. *Mass Spectrom Rev* 2016, 35 (2), 219-58.
81. Maple, H. J., *et al.* Application of the Exactive Plus EMR for automated protein-ligand screening by non-covalent mass spectrometry. *Rapid Commun Mass Spectrom* 2014, 28, 1561-8.
82. Reusch, D., *et al.* Comparison of methods for the analysis of therapeutic immunoglobulin G Fc-glycosylation profiles-Part 2: Mass spectrometric methods. *MAbs* 2015, 7, 732-742.
83. Bondarenko, P. V., *et al.* Mass Measurement and Top-Down HPLC/MS Analysis of Intact Monoclonal Antibodies on a Hybrid Linear Quadrupole Ion Trap–Orbitrap Mass Spectrometer. *J Am Soc Mass Spectrom* 2009, 20, 1415-1424.
84. Wang, C., *et al.* A systematic approach for analysis and characterization of mispairing in bispecific antibodies with asymmetric architecture. *MAbs* 2018, 10, 1226-1235.

85. Kluters, S., *et al.* Solvent modulated linear pH gradient elution for the purification of conventional and bispecific antibodies: Modeling and application. *J Chromatogr A* 2015, 1418, 119-129.
86. Faid, V., *et al.* Middle-up analysis of monoclonal antibodies after combined IgdE and IdeS hinge proteolysis: Investigation of free sulfhydryls. *J Pharm Biomed Anal* 2018, 149, 541-546.
87. Füssl, F., *et al.* Charge Variant Analysis of Monoclonal Antibodies Using Direct Coupled pH Gradient Cation Exchange Chromatography to High-Resolution Native Mass Spectrometry. *Anal Chem* 2018, 90, 4669-4676.
88. Höcker, O., Montealegre, C., Neusüß, C. Characterization of a nanoflow sheath liquid interface and comparison to a sheath liquid and a sheathless porous-tip interface for CE-ESI-MS in positive and negative ionization. *Anal Bioanal Chem* 2018, 410, 5265-5275.
89. Haselberg, R., *et al.* Performance of a sheathless porous tip sprayer for capillary electrophoresis–electrospray ionization-mass spectrometry of intact proteins. *J Chromatogr A* 2010, 1217, 7605-7611.
90. Bongers, J., *et al.* Validation of a peptide mapping method for a therapeutic monoclonal antibody: what could we possibly learn about a method we have run 100 times? *J Pharm Biomed Anal* 2000, 21, 1099-1128.
91. Roberts, G. D., *et al.* An Integrated Strategy for Structural Characterization of the Protein and Carbohydrate Components of Monoclonal Antibodies: Application to Anti-Respiratory Syncytial Virus MAb. *Anal Chem* 1995, 67, 3613-3625.
92. Goyon, A., *et al.* From proof of concept to the routine use of an automated and robust multi-dimensional liquid chromatography mass spectrometry workflow applied for the charge variant characterization of therapeutic antibodies. *J Chromatogr A* 2020, 1615, 460740.
93. Liu, S., *et al.* Mildly acidic conditions eliminate deamidation artifact during proteolysis: digestion with endoprotease Glu-C at pH 4.5. *Amino Acids* 2016, 48, 1059-1067.
94. Mouchahoir, T., Schiel, J. E. Development of an LC-MS/MS peptide mapping protocol for the NISTmAb. *Anal Bioanal Chem* 2018, 410, 2111-2126.
95. Bauer, L. G., *et al.* Monitoring modifications in biopharmaceuticals: Toolbox for a generic and robust high-throughput quantification method. *J Chromatogr B* 2020, 1148, 122134.
96. López-Ferrer, D., *et al.* On-line Digestion System for Protein Characterization and Proteome Analysis. *Anal Chemistry* 2008, 80, 8930-8936.
97. Toth, C. A., *et al.* On-column trypsin digestion coupled with LC-MS/MS for quantification of apolipoproteins. *J Proteomics* 2017, 150, 258-267.
98. Naldi, M., *et al.* Towards automation in protein digestion: Development of a monolithic trypsin immobilized reactor for highly efficient on-line digestion and analysis. *Talanta* 2017, 167, 143-157.
99. Hedström, M., *et al.* Miniaturized on-line digestion system for the sequential

- identification and characterization of protein analytes. *J Chromatogr A* 2007, 1146, 17-22.
100. Han, M., *et al.* Intact mass analysis of monoclonal antibodies by capillary electrophoresis—Mass spectrometry. *J Chromatogr B* 2016, 1011, 24-32.
 101. Giorgetti, J., *et al.* Intact monoclonal antibodies separation and analysis by sheathless capillary electrophoresis-mass spectrometry. *Eur J Mass Spectrom* 2019, 25, 324-332.
 102. Spahr, C., *et al.* Recombinant human lecithin-cholesterol acyltransferase Fc fusion: analysis of N- and O-linked glycans and identification and elimination of a xylose-based O-linked tetrasaccharide core in the linker region. *Protein Sci* 2013, 22, 1739-53.
 103. Spahr, C., *et al.* High-resolution mass spectrometry confirms the presence of a hydroxyproline (Hyp) post-translational modification in the GGGP linker of an Fc-fusion protein. *MAbs* 2017, 9, 812-819.
 104. Hou, Y., *et al.* Nutrient Optimization Reduces Phosphorylation and Hydroxylation Level on an Fc-Fusion Protein in a CHO Fed-Batch Process. *Biotechnol J* 2019, 14, e1700706.
 105. Raju T. S. Hydroxylation of Proteins. In *Co- and Post-Translational Modifications of Therapeutic Antibodies and Proteins*, 2019; pp 119-131.
 106. Xie, Q., Moore, B., Beardsley, R. L. Discovery and characterization of hydroxylysine in recombinant monoclonal antibodies. *MAbs* 2016, 8, 371-378.
 107. Boddapati, S., *et al.* Evidence for co-translational misincorporation of non-canonical amino acid hydroxyproline in recombinant antibodies produced in Chinese Hamster Ovary (CHO) cell lines. *PLOS ONE* 2020, 15, e0241250.
 108. Bobály, B., *et al.* Analysis of recombinant monoclonal antibodies in hydrophilic interaction chromatography: A generic method development approach. *J Pharm Biomed Anal* 2017, 145, 24-32.
 109. Haselberg, R., *et al.* Heterogeneity assessment of antibody-derived therapeutics at the intact and middle-up level by low-flow sheathless capillary electrophoresis-mass spectrometry. *Anal Chim Acta* 2018, 1044, 181-190.
 110. Giorgetti, J., *et al.* Combination of intact, middle-up and bottom-up levels to characterize 7 therapeutic monoclonal antibodies by capillary electrophoresis – Mass spectrometry. *J Pharm Biomed Anal* 2020, 182, 113107.
 111. Fornelli, L., *et al.* Middle-Down Analysis of Monoclonal Antibodies with Electron Transfer Dissociation Orbitrap Fourier Transform Mass Spectrometry. *Anal Chem* 2014, 86, 3005-3012.
 112. Chen, G., *et al.* Distinguishing Sulfotyrosine Containing Peptides from their Phosphotyrosine Counterparts Using Mass Spectrometry. *J Am Soc Mass Spectrom* 2018, 29, 455-462.
 113. Nemeth-Cawley, J. F., Karnik, S., Rouse, J. C. Analysis of sulfated peptides using positive electrospray ionization tandem mass spectrometry. *J Mass Spectrom* 2001, 36, 1301-11.
 114. Stracke, J., *et al.* A novel approach to investigate the effect of methionine oxidation on

- pharmacokinetic properties of therapeutic antibodies. *MAbs* 2014, 6, 1229-1242.
115. Hajduk, J., *et al.* Interaction analysis of glycoengineered antibodies with CD16a: a native mass spectrometry approach. *MAbs* 2020, 12, 1736975.
 116. Dekkers, G., *et al.* Multi-level glyco-engineering techniques to generate IgG with defined Fc-glycans. *Sci Rep* 2016, 6, 36964.
 117. Lippold, S., *et al.* Glycoform-resolved FcγRIIIa affinity chromatography-mass spectrometry. *MAbs* 2019, 11, 1191-1196.
 118. Schlothauer, T., *et al.* Analytical FcRn affinity chromatography for functional characterization of monoclonal antibodies. *MAbs* 2013, 5, 576-586.
 119. Gahoual, R., *et al.* Detailed Characterization of Monoclonal Antibody Receptor Interaction Using Affinity Liquid Chromatography Hyphenated to Native Mass Spectrometry. *Anal Chem* 2017, 89, 5404-5412.
 120. Thomann, M., *et al.* In vitro glycoengineering of IgG1 and its effect on Fc receptor binding and ADCC activity. *PloS one* 2015, 10, e0134949-e0134949.
 121. Hajduk, J., *et al.* Interaction analysis of glycoengineered antibodies with CD16a: a native mass spectrometry approach. *MAbs* 2020, 12, 1736975-1736975.
 122. Liu, J., Konermann, L. Protein–Protein Binding Affinities in Solution Determined by Electrospray Mass Spectrometry. *J Am Soc Mass Spectrom* 2011, 22, 408-417.
 123. Bich, C., *et al.* Probing the Hydrophobic Effect of Noncovalent Complexes by Mass Spectrometry. *J Am Soc Mass Spectrom* 2010, 21, 286-289.
 124. Alhazmi, H. A., *et al.* A comprehensive platform to investigate protein–metal ion interactions by affinity capillary electrophoresis. *J Pharm Biomed Anal* 2015, 107, 311-317.
 125. Hutanu, A., *et al.* Methionine oxidation of proteins analyzed by affinity capillary electrophoresis in presence of silver(I) and gold(III) ions. *Electrophoresis* 2021, 42, 1209-1216.
 126. Domínguez-Vega, E., *et al.* Simultaneous Assessment of Protein Heterogeneity and Affinity by Capillary Electrophoresis–Mass Spectrometry. *Anal Chem* 2015, 87, 8781-8788.





Chapter 2

Fast and automated characterization of antibody variants with 4D-HPLC/MS

2

**Christoph Gstöttner¹, Denis Klemm¹, Markus Habberger², Anja Bathke¹, Harald Wegele²,
Christian Bell¹, Robert Kopf¹**

¹F. Hoffmann-La Roche, Basel, Switzerland

²Roche Pharma Technical Development, Penzberg, Germany

Reprinted and adapted with permission from Analytical Chemistry, 2018, 90, 2119-2125, DOI:
10.1021/acs.analchem.7b04372

© 2018 American Chemical Society

2.1 Abstract

Characterization of unknown monoclonal antibody (mAb) variants is important in order to identify their potential impact on safety, potency and stability. Ion exchange chromatography (IEC) coupled with UV detection is frequently used to separate and quantify mAb variants in routine quality control (QC). However, characterization of the chromatographic peaks resulting from an IEC separation is an extremely time consuming process, involving many cumbersome steps. Presented here is an online four-dimensional high performance liquid chromatography-mass spectrometry (4D-HPLC/MS) approach, developed to circumvent these limitations. To achieve this a 2D-HPLC system was extended through the introduction of additional modules, hence enabling fully automated bioseparation of mAbs, fractionation of peaks, reduction, tryptic digestion and reversed-phase (RP) separation of resulting peptides followed by MS detection. The entire separation and analytical process for an unknown peak is performed in less than 1.5 hours, leading to a significant time saving, with comparable sequence coverage. To show the comparability with the traditional offline process, a proof of concept study with a previously characterized mAb1 is presented in this paper.

2.2 Introduction

The chemical amino acid modifications of therapeutic antibodies have been extensively investigated and well reviewed.² The most common degradations are oxidation of methionine and tryptophan, deamidation of asparagine, isomerization of aspartate and glycation of lysine residues³⁻¹⁰ Monitoring of these modifications is important, because they show altered stability as well as an impact on biological function, especially deamidation of asparagine residues located in the complementarity-determining regions (CDRs).¹¹⁻¹⁵ In addition to deamidation, methionine oxidation also requires monitoring, due to the fact that oxidized mAbs are known to influence in vivo product stability, receptor binding, structure and exhibit accelerated plasma clearance.¹⁶⁻¹⁸

Normally characterization of mAb modifications is a time-consuming and multi-step offline process, where the samples are fractionated, reduced and alkylated followed by tryptic digestion for several hours. More recently Native MS is receiving increased application,¹⁹ however the structural information obtained is only protein specific and pre-fractionation or post-digestion is still required to localize the detected modifications.

Tryptic digestion in solution presents some disadvantages including trypsin autodigestion, long incubation times and low efficiency.²⁰ Several methodologies have been reported to improve or automate the digestion step of proteins,^{21,22} however in many cases their applicability to routine analysis is limited due to the complexity of the system. López-Ferrer et al. described a system with a pressurized loop, which enabled online digestion of protein samples.²³ In addition, immobilization of trypsin on a monolithic column to generate a trypsin bioreactor, so called immobilized enzyme reactor (IMER) was reported.²⁴ Hedström et al. discussed a miniaturized system where trypsin was immobilized in a fused silica capillary directly followed by a reversed phase separation.²⁵ These systems all involve direct digestion of the introduced protein sample, without the possibility to separate variants up-front. The main disadvantage of such methodologies is that protein samples, in reality, often consist of a mixture of different protein variants, and subsequent results are complicated by this heterogeneity.

In 2011, Alvarez et al. presented a 2D LC-MS technique for online reduction, separation and detection of mAb variants. The method enabled detection and localization of larger modifications to the respective reduced protein chains, however the system configuration was rather complex.²⁶ In a recent publication we presented a fully automated online 2D system also enabling online reduction and separation.²⁷ These 'middle-down' approaches, whilst informative, share the limitation that the modifications measured are not localizable to the amino acid level (or at least the peptide) and nor are the methods sensitive enough to enable detection of smaller modifications such as deamidation (m/z delta 1 Da).

More recently, a 1D method applying ion exchange chromatography directly coupled with

a mass spectrometer for native ESI-MS analysis was reported. The system proved effective for determination of larger mass-shift modifications, however smaller modifications such as deamidation, required additional sample preparation steps.²⁸ Implementing digestion with IdeS prior to the IEC analysis enabled location of the modification to the Fc or Fab fragment, but for the exact position additional peptide mapping would be required.

Recently several different approaches have been presented in the field of mAb analysis using 2D-HPLC, for example the analysis of mAb breakdown products.²⁹ For a recent overview about different methods for antibody characterization please refer to in Sandra et al.³⁰. Different chromatographic support combinations have been tested, such as strong cation exchange (SCX) combined with reversed-phase (RP), RP combined with RP and hydrophilic interaction chromatography (HILIC) combined with RP for the analysis of digested mAbs.³¹ Benefits of this technology for mAb analysis have been illustrated by Sandra et al.³² The combination of cation exchange chromatography (CEX) with RP was also shown to be a powerful tool for separation of mAb fragments or peptides especially combined with MS detection.³³

Another method, developed by Tran et al., including an IEC separation, on-line reduction and on-line digestion with electrospray time-of-flight (ESI-TOF)-MS characterization has also been reported. The technique was however only applied to smaller proteins and required a customized capillary chromatographic system using a stop-flow technique.³⁴ Full automation and online digestion of large molecules, which is necessary for pharmaceutical drug product quality control, was not presented.

In this article we describe for the first time a fully automated 4D HPLC-MS system for online separation, reduction, digestion and peptide mapping which addresses the above mentioned disadvantages. The system enables automated fractionation of IEC-separated charged variants, online reduction and tryptic digestion, subsequent chromatographic resolution of the generated peptides and detailed characterization through online coupling to ESI-MS/MS. To achieve full automation of this process, a standard, commercially available, 2D-HPLC was extended by the introduction of further standard modules. A significant advantage of our system is that no adjustments to the IEC separation conditions or buffers is required for introduction to the MS, but rather the methods of commercial QC quality control can be directly used. The system is simple to configure and the resulting methods robust and user-friendly. As such the application of the method in research and development labs is highly feasible and offers tremendous improvements over current workflows. In this paper a proof of concept study of our system with a mAb, previously characterized by standard offline analytical methods¹ is presented.

2.3 Experimental Section

2.3.1 Production of stressed material

To generate stressed material for characterization of mAb variants, a recombinant IgG1 antibody expressed in Chinese hamster ovary cells, hereafter described as mAb2, was used (Roche Diagnostics GmbH, Penzberg, Germany). The hydrophobicity is comparable to a standard IgG1 and the isoelectric point is 9.1. The concentration of the formulation was 25 mg/mL in a 20 mM His-HCl (Sigma-Aldrich, Missouri, USA) buffer system at pH 5.5. The generation of antibody variants, including deamidated, oxidized and glycosylated versions of mAb2, was achieved using various stress conditions, namely high pH, elevated temperature, H_2O_2 (Merck KGaA, Darmstadt, Germany) and raised levels of glucose. Forced deamidation was achieved by diluting 800 μ L of mAb2 with 1.2 mL of 200 mM TRIS-HCl (Sigma-Aldrich, Missouri, USA) at pH 9.0 and incubating for 7 days at 37°C. Forced oxidation was achieved by adding 0.2% H_2O_2 to the formulated mAb2 and incubation for 24 h at 25 °C. For generating glycosylated mAb2, the formulated antibody was diluted with a stock solution of glucose (Merck KGaA, Darmstadt, Germany) to a final concentration of 1 M. This was followed by incubation for 7 days at 25°C.

2.3.2 4D-HPLC hardware configuration for analysis of charged variants

The 4D-HPLC system was completely configured using standard modules from Agilent Technologies (Waldbronn, Germany). The system is based on a standard 2D-HPLC with the following modules: 1260 Bio ALS autosampler (G5667A), 1290 thermostat (G1330B), 1290 binary pump (G4220A), 1260 quaternary pump (G5611A), 1290 column oven (G1316C) configured with a 6 port / 2 position valve (G4231B) and a 1260 UV detector (G1314F). Critical for the 2D application is the interface module consisting of a 1290 8 port / 2 position valve (G4236A) and two selector valve (decks) with 6 x 40 μ L loops each for multiple heart cutting (G4242A multiple heart cutting kit). This standard system was further extended with an additional binary pump (G4220A), a quaternary pump (G1311A), a second column oven (G1316A) configured with a 6 port / 2 position valve (G4231B) and a 6 port / 2 position external valve (G1158B). The system was controlled by two installations of the OpenLab software package, one for the two quaternary pumps, the two column ovens and the external valve, the second for the remaining modules. For full automation, an additional self-designed macro "valve event plugin" from ANGI (Gesellschaft für angewandte Informatik, Karlsruhe, Germany) is necessary. This macro generates a start signal for the modules in the second OpenLab software for each fraction. A start signal for the MS system is also simultaneously triggered and thus a single MS TIC for every fraction is produced. The 'remote' start signals are delivered by a contact closure board, which is integrated in the autosampler.

2.3.3 First dimension - IEC separation and fractionation

IEC separation was performed to separate charged variants from the main mAb species. For separation, a 4.0 x 250 mm ProPac WCX-10 analytical cation exchange column (Dionex, Thermo Scientific, Massachusetts, USA) was used. A step gradient from 5 to 100% B in 56 min with a flowrate of 1 mL/min of 20 mM MES pH 6.2 as solvent A and 20 mM MES (Merck KGaA, Darmstadt, Germany), 0.75 M NaCl (Sigma – Aldrich, Missouri, USA), pH 6.2 as solvent B was applied. The column temperature was set to 40°C and UV detection at 280 nm. For each injection 175 µg mAb2 was used.

2.3.4 Second dimension – RP trapping, reduction and elution

A 2.1 x 12.5 mm AdvancedBio RP cartridge column (Agilent Technologies, Waldbronn, Germany) was used for subsequent trapping and reduction. As mobile phase A and B, 0.1% FA (Fluka, Missouri, USA) in H₂O and in ACN (Merck KGaA, Darmstadt, Germany), respectively were used. Reduction was achieved using 20 mM DTT (Dithiothreitol) (mobile phase C) for 10 minutes, followed by isocratic elution of the cartridge with 60% B to the 3rd dimension column for 25 min. Longer reduction times were evaluated, however even after 10min incubation full reduction of the antibody was obtained. After reduction no intact material could be detected by MS. To prevent carryover, the cartridge was subsequently flushed with 100% mobile phase B for 5 mins at the end of each run. For gradient details see **Table 1**.

Table 1. Gradient and valve position information for the second dimension trapping, reduction and elution.

Time (min)	A (%)	B (%)	C (%)	Flow (mL/min)	Valve 3 switching (Position)
0.00	99.0	1.0	0.0	0.50	1-2
5.00	99.0	1.0	0.0	0.50	
5.01	0.0	0.0	100.0	0.10	
15.00	0.0	0.0	100.0	0.10	
15.01	99.0	1.0	0.0	0.50	
20.00	99.0	1.0	0.0	0.50	
20.01	40.0	60.0	0.0	0.06	1-6
45.00	40.0	60.0	0.0	0.06	1-2
45.01	0.0	100.0	0.0	0.10	
47.00	0.0	100.0	0.0	0.10	
47.00	0.0	100.0	0.0	0.50	
50.00	0.0	100.0	0.0	0.50	
51.00	99.0	1.0	0.0	0.50	

2.3.5 Third dimension - tryptic digestion

The reduced mAb chains eluted from the second dimension RP column (60% B, 60 μ L/min) were directed to a 2.1 x 100 mm StyrosZyme TPCK-Trypsin column, (OraChrom, Massachusetts, USA) in the third dimension. The column temperature was set to 40°C. Digestion buffer comprising of 50 mM TRIS, 10mM CaCl₂ (Merck KGaA, Darmstadt, Germany) at pH 8.0 was permanently introduced via a T-junction with a flow rate of 250 μ L/min. This leads to a final solvent composition of 11.6% ACN diluted in digestion buffer for elution of the peptides to the fourth dimension RP column. The effective digestion time or contact time on the trypsin column is approximately one minute, given the flow rate of the second dimension column and added digestion buffer. Subsequent cleaning of the column after digestion was achieved with 50 % ACN and 50 % digestion buffer at a flowrate of 0.5 mL/min for 5 min.

2.3.6 Fourth dimension peptide trapping and peptide mapping

Peptides were trapped on an UPLC ACQUITY BEH C18 column 2.1 x 50 mm, 1.7 μ m (Waters Corporation, Massachusetts, USA). Mobile phases of 0.1% FA in H₂O and 0.1% FA in ACN were used as A and B, respectively. A gradient of 1% to 100% B in 87 minutes was applied, however the starting conditions are, in reality, closer to 11.6% B, due to the relatively high composition of ACN in the elution solvent from the second dimension (60%), despite dilution with digestion buffer in the 3rd dimension. The trapped peptides were first washed 5 min with 1% ACN, and thereafter eluted to the MS system with the gradient shown in **Table 2**. The solvent composition is maintained at 1% mobile phase B for the first 50 min until introduction of the digested peptides from the third dimension column, following that the elution gradient begins. The MS analysis was performed with an Impact II ESI Q-TOF (Bruker, Bremen, Germany). The connection between mass spectrometer and 4D HPLC were fused silica capillaries. For the MS data acquisition, the HyStar™ software was used and for the evaluation of results the DataAnalysis software. The spectra were generated in positive ion mode over the *m/z* range 300 – 2500 and a frequency of 1 Hz. The endplate offset was set to 500 V and the capillary to 4500 V. The theoretical Peptide masses and modifications were obtained by in silico trypsin digestion with GPMaw (General Protein/Mass Analysis for Windows).

Table 2. Fourth dimension RP column elution gradient conditions (for peptide mapping).

Time (min)	A (%)	B (%)	Valve 2 switching (Position)	Valve 4 switching (Position)
0.00	99.0	1.0	1-2	1-2
20.00	99.0	1.0	1-6	
50.00	99.0	1.0		1-6
51.00	99.0	1.0		
58.00	85.0	15.0		
65.00	81.0	19.0		
82.00	65.0	35.0		
86.00	30.0	70.0		
87.00	0.0	100.0		
90.00	0.0	100.0	1-2	1-2
92.00	99.0	1.0		
96.00	99.0	1.0		

2.3.7 Process workflow

In the first dimension charged variants are separated by IEC separation. Fractions of interest can be stored in one of the 12 loops of deck A and B. After the first selected fraction is stored in one of the loops, valve one is switched and the fraction is flushed, by the second dimension quaternary pump, to the 2D RP trapping cartridge (AdvancedBio RP 2.1 x 12.5 mm). Reducing agent is subsequently delivered to the trapping cartridge for 10 min by the same pump enabling protein reduction. To prevent degradation of the enzymatic column by the reducing agent, the trapping cartridge is first washed with 1% mobile phase B prior to elution of the reduced mAb chains to the 3D trypsin cartridge. The fourth pump delivers digestion buffer via a T-junction. The mAb is digested by elution through the trypsin cartridge and the peptides are trapped on a second RP column (UPLC ACQUITY BEH C18 column; 2.1 x 50 mm, 1.7 μ m). The trapped peptides are washed to prevent contamination of the MS system by the high-salt containing (CaCl₂ and TRIS) digestion buffer. Subsequently, the peptides are separated using the third pump and analyzed by ESI-QTOF MS/MS. The entire process is depicted schematically in the flowchart shown in **Figure 1**.

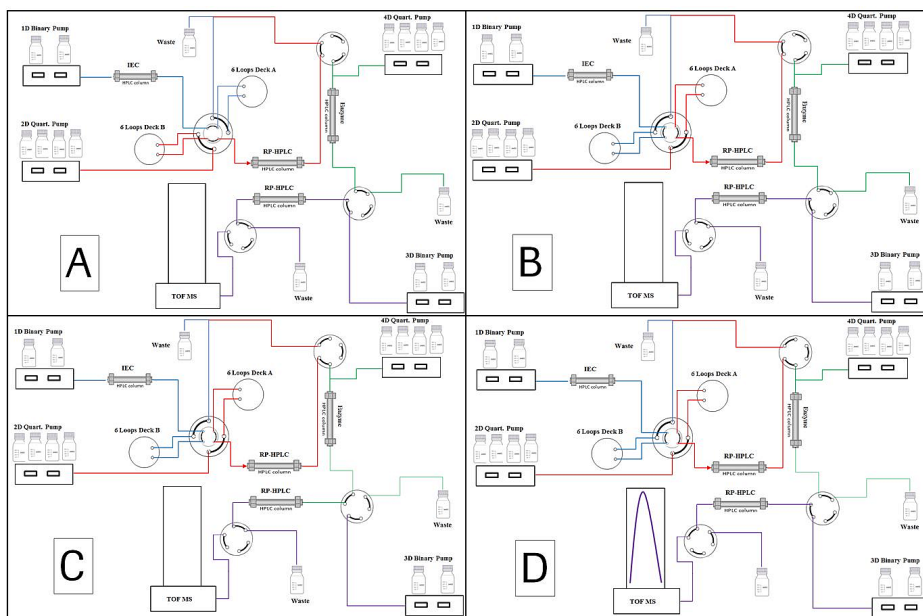


Figure 1. Schematic representation of the automated process workflow, enabling protein isolation, online reduction, digestion and detection. A) The IEC fractions of interest from the 1D separation are stored in Deck A or Deck B. B) Valve 1 is switched, enabling transfer of the isolated 1D fraction(s) to the 2D RP column. The sample is subsequently reduced and washed. C) Valves 2 and 3 are switched, hence eluting the reduced mAb chains from the 2D RP column to the 3D digestion column. The digested peptides are subsequently trapped on the online 4D RP column. D) Valves 2 and 4 are switched, enabling chromatographic elution and LC/MS detection of the trapped peptides from the 4D RP column.

2.4 Results and Discussion

2.4.1 Proof of concept study

In order to demonstrate the productive efficiency of the developed 4D-HPLC-MS system a proof of concept study was performed. A previously characterized therapeutic antibody (mAb2) was used as the benchmark. Results of this previous characterization work, which was performed using the traditional offline process, have been published by Habberger et al.¹ The goal was to detect and localize all reported modifications including deamidation, oxidation and glycation in stressed and non-stressed samples.

Samples exposed to pH stress conditions and reference material were analyzed. High pH stress was chosen, as compared to temperature stress, in order to impart higher levels of deamidation. Online fractionation was carried out for the acidic peak of both samples. Measurement of the reference material was also performed to determine sensitivity to low

levels of deamidation (see **Table 3** for expected levels). Additionally, the main IEC peak was also fractionated in order to enable determination of the efficiency of the IEC separation. Ideally no deamidation should be detected in the main peak fraction, as the loss of positive charge resulting from deamidation leads to a shift to the acidic region of the IEC separation (**Figure 2**).

Table 3. Quantitation results of deamidated peptides in reference material and pH stressed material. Values sourced from Ref¹.

Peptide	Relative amount of deamidated peptide (%)*	
	Reference material	pH stressed sample
LC-deamid-30,31	1.8	11.0
LC-deamid-50,52,53	0.5	0.9
LC-deamid-92	7.5	27.1
HC-deamid-384,389,390	1.6	26.0

* Values calculated relative to wild-type peptide (LC/MS peptide mapping)

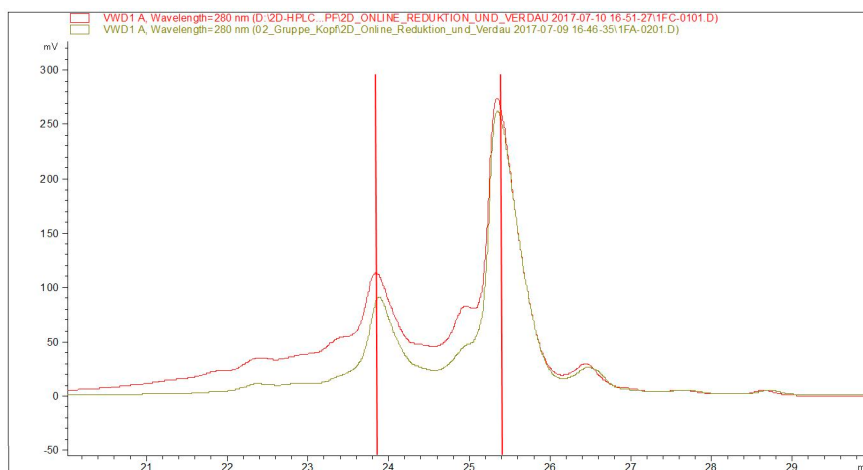


Figure 2. IEC-UV chromatogram of pH stressed material (red trace) and reference material (brown trace). Fractions (approximately 3.4 µg of AP2) are depicted by red bars.

All four deamidation sites could be detected in both the acidic peak fraction (AP2) of the pH stressed and the non-stressed reference material, as shown in **Figure 3**. The mass difference was the expected 1 Da. The high sensitivity of our system is demonstrated by our ability to detect the deamidation of the light chain at position 50,52,53, which is only abundant with an amount of 0.5% in the non-

stressed reference material. As can be seen in **Figure 3B**, this peptide could also be detected in the main peak fraction, while the other three deamidated peptides are only detectable in the acidic region. One possible explanation for this is that the asparagine residues on this peptide are buried in the hydrophobic core of the mAb and hence interaction with the IEC matrix is inhibited. This is further corroborated by the fact, that also during pH stress nearly no deamidation occurs at these three asparagine residues. The second signal also detected (**Figure 3B**) is attributed to a co-eluting peptide.

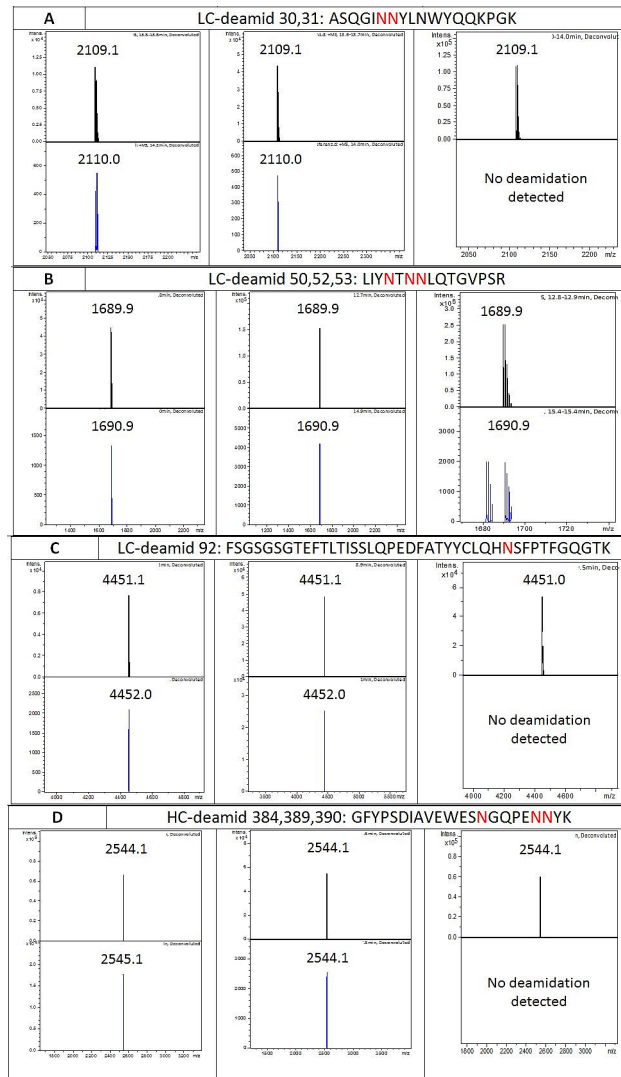


Figure 3. (A-D) Spectral results showing the deamidated peptides detected in the pH-stressed, reference material and fractionated main peak samples.

In addition to deamidation, glycation was also one of the modifications found in the molecule.¹ In order to determine if glycation at high and low levels can also be detected with our system, artificially glycated material and the reference material were analyzed. The main peak, in which no glycation should be detected, was also fractionated and analyzed to confirm the efficiency of the IEC separation. The IEC chromatogram, including the fractionation, is depicted in **Figure 4**.

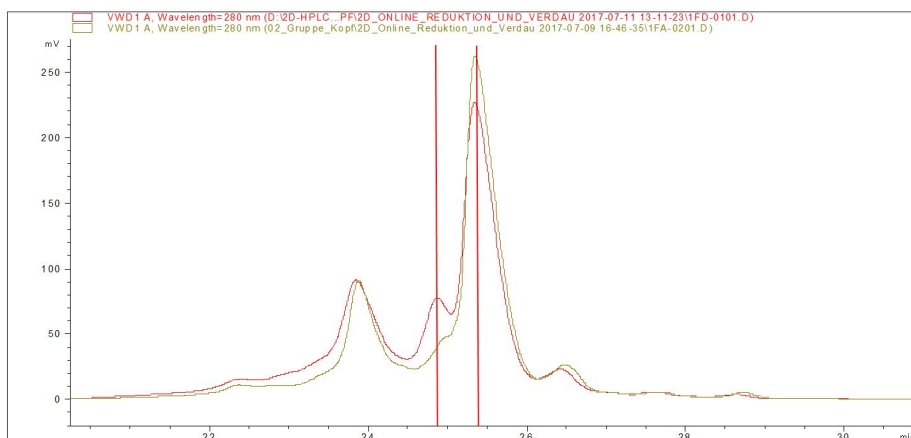


Figure 4. IEC chromatogram of material stressed with high amount of glucose (red trace) and reference material (brown trace). Fractions (approximately 1.68µg of AP1) are depicted by red bars.

The expected mass difference of 162 Da for a glucose molecule could be detected in both the stressed material and in the reference material. The glycation was confirmed at the same position (HC-Lys-33) as reported previously.¹ In the main peak fraction no glycation could be detected, see **Figure 5**.

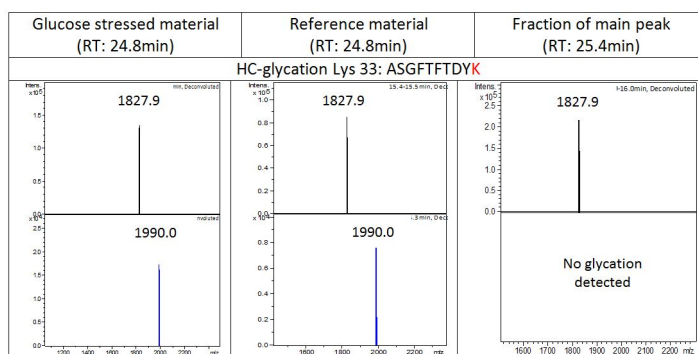


Figure 5. Spectral results showing the glycated peptide, as detected in the glucose-stressed, reference material and fractionated main peak samples.

Detection of the six oxidation sites, also reported for mAb2,¹ was also investigated with the 4D HPLC-MS system. For this purpose, the main IEC peak, as well as the basic fraction were

analyzed. The relative amounts of oxidized peptides detected in the H₂O₂-stressed sample and the reference material are provided in **Table 4**.¹

Table 4. Quantitation results of oxidized peptides in reference material and oxidation stressed material, as sourced from Ref¹.

Peptide	Relative amount of oxidized peptide (%)	
	Reference material	H2O2 stressed sample
LC-Met 4	3.3	47.9
HC-Met 48	1.8	1.8
HC-Met 80	1.9	1.9
HC-Met 100c	3.3	99.4
HC-Met 252	3.0	99.0
HC-Met 428	1.6	80.4

Five out of the six modifications were determined in the two fractions on our system. The peptide corresponding to the sixth modification (HC-Met 252) however could not be detected. This is a result of the fact that the acetonitrile content needed to elute the reduced chains from the 2D RP cartridge leads to a composition of approximately 12% ACN over the subsequent RP trapping column, despite maximal dilution with digestion buffer. Subsequently small, polar peptides (< 1.3kDa), including the HC-Met 252 peptide, are not retained in the trapping step and therefore not detected during peptide mapping.

The online reduction and digestion steps of the developed system have also been successfully applied to different antibody formats confirming high repeatability of this set-up (data not shown). After online reduction and digestion no intact material or intact chains were detected. Reported here is the proof of concept of the 4D online method carried out with the previously fully characterized mAb.¹ The goal to detect all previously published modifications was successful in terms of all four deamidation sites and the glycation at HC Lysine 33. These modifications could also be detected at the low levels present in the reference material, including the deamidation of the LC-Asn-50,52,53 peptide (total amount of 0.5% in the reference material), showing the sensitivity of the assay. In addition, no deamidation-modified peptides were detected in the fractionated IEC main peak demonstrating that the IEC separation performs efficiently.

In addition to deamidation and glycation, known oxidation modifications¹ were also investigated. Since oxidation does not result in a different charge pattern, separation of the oxidized from the native species by IEC is not expected. Therefore, the main peak and basic peaks from the 1D IEC were fractionated and analyzed by the 4D system. With our assay

we could determine five out of the six expected oxidation sites, however the HC-Met-252 oxidized peptide could not be detected. This is due to inability to trap the peptide resulting from the slightly higher amount of ACN compared to the common process. This inability to bind small, hydrophilic peptides is largely related to the dimensions of the peptide mapping column. This column that is used behind the trypsin column is rather short to achieve a backpressure below the limit of the trypsin column. Other alternatives potentially enhancing trapping of small hydrophilic peptides as well as providing additional resolution to separate aspartate isomerization, such as C30 or monolith columns, will be further investigated.

2.5 Conclusion

With the 4D HPLC setup presented here it is possible to characterize an unknown peak within a mAb formulation in a minimum timeframe. The characterization of five peaks of an IEC chromatogram, which is a typical number for IEC characterization, takes only approximately 9 h, as compared to 52 h using the common process. This time saving is achieved through full automation of the entire process from fractionation to peptide mapping followed by MS detection. Only the fractions of interest from the first dimension separation are chosen, as in normal 2D HPLC. Compared to our previous system ²⁷ (in press) which was limited to online reduction of mAb variants and subsequent separation of the reduced chains, this 4D HPLC system enables detection of low molecular weight modifications, including i.e. deamidation. Due to the coupling with online MS/MS, another benefit of this system is that the modifications can be localized to amino acid level. Sequence coverage of mAb2 (light chain 94%, heavy chain 86%) is comparable to the offline procedure (light chain 94%, heavy chain 94%). Especially advantageous is that, the presented 4D HPLC system also enables pre-fractionation by IEC so that separated, selected charged variants, rather than the whole sample, are reduced and digested, as in other systems.^{35, 36} Through this preselection of a charged variant of interest, higher overall signal intensity and sensitivity is achieved. With a single IEC separation run it is possible to fractionate up to nine peaks, and as such, only a minimum amount of material is required for complete IEC separation and isolation. The proposed system hence also offers enhanced possibilities for characterization of unknown peaks in early phase mAb development, where often only limited material is available. The system is configured such that, in addition to IEC separation, SEC could also be used as the separation mechanism of the first dimension. With this configuration size variants could hence also be characterized. The 4D system presented here was developed such that commercial quality control separation methods could directly be applied in the first dimension without any adaptation necessary. This offers significant benefits and time-saving in method transfer. Further experiments are ongoing and include optimization of peptide retention and development of methods using alternative digestion enzymes such as pepsin and LysC.

2.6 Supporting information

Supplementary information is available free of charge on the ACS website: DOI: 10.1021/acs.analchem.7b04372

2.7 Acknowledgments

We wish to thank the Roche Global Strategy Team (GST) for supporting the project, and especially Rico Schärer from Agilent Technologies for his excellent technical support. Many thanks also to Patrick Bulau and Lea Bonnington for reviewing this manuscript and Christof Finkler for providing the opportunity to develop this system.

2.8 References

1. Habegger, M., *et al.* Assessment of chemical modifications of sites in the CDRs of recombinant antibodies: Susceptibility vs. functionality of critical quality attributes. *MAbs* 2014, 6, 327-339.
2. Liu, H., *et al.* In vitro and in vivo modifications of recombinant and human IgG antibodies. *MAbs* 2014, 6, 1145-1154.
3. Yang, H., Zubarev, R. A. Mass spectrometric analysis of asparagine deamidation and aspartate isomerization in polypeptides. *Electrophoresis* 2010, 31, 1764-1772.
4. Harris, R. J. Heterogeneity of recombinant antibodies: linking structure to function. *Dev Biol (Basel)* 2005, 122, 117-127.
5. Hovorka, S.; Schoneich, C. *J Pharm Sci* 2001, 90, 253-269.
6. Ji, J. A., *et al.* Oxidative degradation of pharmaceuticals: theory, mechanisms and inhibition. *J Pharm Sci* 2009, 98, 4485-4500.
7. Quan, C., *et al.* A study in glycation of a therapeutic recombinant humanized monoclonal antibody: where it is, how it got there, and how it affects charge-based behavior. *Anal Biochem* 2008, 373, 179-191.
8. Fischer, S., Hoernschemeyer, J., Mahler, H. C. Glycation during storage and administration of monoclonal antibody formulations. *Eur J Pharm Biopharm* 2008, 70, 42-50.
9. Pace, A. L., *et al.* Asparagine deamidation dependence on buffer type, pH, and temperature. *J Pharm Sci* 2013, 102, 1712-1723.
10. Zhang, B., *et al.* Unveiling a glycation hot spot in a recombinant humanized monoclonal antibody. *Anal Chem* 2008, 80, 2379-2390.
11. Yan, B., *et al.* Succinimide formation at Asn 55 in the complementarity determining region of a recombinant monoclonal antibody IgG1 heavy chain. *J Pharm Sci* 2009, 98, 3509-3521.
12. Vlasak, J., *et al.* Identification and characterization of asparagine deamidation in the light chain CDR1 of a humanized IgG1 antibody. *Anal Biochem* 2009, 392, 145-154.

13. Diepold, K., *et al.* Simultaneous Assessment of Asp Isomerization and Asn Deamidation in Recombinant Antibodies by LC-MS following Incubation at Elevated Temperatures. *PLoS One* 2012, 7, e30295.
14. Harris, R. J., *et al.* Identification of multiple sources of charge heterogeneity in a recombinant antibody. *J Chromatogr B Biomed Sci Appl* 2001, 752, 233-245.
15. Huang, L., *et al.* In vivo deamidation characterization of monoclonal antibody by LC/MS/MS. *Anal Chem* 2005, 77, 1432-1439.
16. Schlothauer, T., *et al.* Analytical FcRn affinity chromatography for functional characterization of monoclonal antibodies. *MAbs* 2013, 5, 576-586.
17. Bertolotti-Ciarlet, A., *et al.* Impact of methionine oxidation on the binding of human IgG1 to Fc Rn and Fc gamma receptors. *Mol Immunol* 2009, 46, 1878-1882.
18. Stracke, J., *et al.* A novel approach to investigate the effect of methionine oxidation on pharmacokinetic properties of therapeutic antibodies. *MAbs* 2014, 6, 1229-1242.
19. Habberger, M., *et al.* Rapid characterization of biotherapeutic proteins by size-exclusion chromatography coupled to native mass spectrometry. *MAbs* 2016, 8, 331-339.
20. Ma, J., *et al.* Monolith-based immobilized enzyme reactors: Recent developments and applications for proteome analysis. *J Sep Sci* 2007, 30, 3050-3059.
21. Toth, C. A., *et al.* On-column trypsin digestion coupled with LC-MS/MS for quantification of apolipoproteins. *J Proteomics* 2017, 150, 258-267.
22. Naldi, M., *et al.* Towards automation in protein digestion: Development of a monolithic trypsin immobilized reactor for highly efficient on-line digestion and analysis. *Talanta* 2017, 167, 143-157.
23. Lopez-Ferrer, D., *et al.* On-line digestion system for protein characterization and proteome analysis. *Anal Chem* 2008, 80, 8930-8936.
24. Calleri, E., *et al.* Trypsin-Based Monolithic Bioreactor Coupled On-Line with LC/MS/MS System for Protein Digestion and Variant Identification in Standard Solutions and Serum Samples. *J Proteome Res* 2005, 4, 481-490.
25. Hedstrom, M., *et al.* Miniaturized on-line digestion system for the sequential identification and characterization of protein analytes. *J Chromatogr A* 2007, 1146, 17-22.
26. Alvarez, M., *et al.* On-line characterization of monoclonal antibody variants by liquid chromatography-mass spectrometry operating in a two-dimensional format *Anal Biochem* 2011, 419, 17-25.
27. Bathke A., *et al.* Rapid Online Characterization and Reduction of Protein Modifications Using Fully Automated Two-Dimensional High Performance Liquid Chromatography–Mass Spectrometry. *LCGC* 2018, 31, 10-21.
28. Leblanc, Y., *et al.* Charge variants characterization of a monoclonal antibody by ion exchange chromatography coupled on-line to native mass spectrometry: Case study after a long-term storage at +5°C. *J Chromatogr B Analyt Technol Biomed Life Sci* 2017, 1048, 130-139.
29. Mazur, M. T., *et al.* A Platform for Characterizing Therapeutic Monoclonal Antibody

- Breakdown Products by 2D Chromatography and Top-Down Mass Spectrometry. *AAPS J* 2012, 14, 530-541.
30. Sandra, K., *et al.* Characterizing Monoclonal Antibodies and Antibody–Drug Conjugates Using 2D-LC–MS. *Lc Gc Eur* 2017, 30, 149-157.
 31. Vanhoenacker, G., *et al.* Comprehensive two-dimensional liquid chromatography of therapeutic monoclonal antibody digests. *Anal Bioanal Chem* 2015, 407, 355-366.
 32. Sandra, K., Sandra, P. The opportunities of 2D-LC in the analysis of monoclonal antibodies. *Bioanalysis* 2015, 7, 2843-2847.
 33. Stoll, D. R., *et al.* Direct Identification of Rituximab Main Isoforms and Subunit Analysis by Online Selective Comprehensive Two-Dimensional Liquid Chromatography–Mass Spectrometry. *Anal Chem* 2015, 87, 8307–8315.
 34. Tran, B. Q., *et al.* On-Line multitasking analytical proteomics: How to separate, reduce, alkylate and digest whole proteins in an on-Line multidimensional chromatography system coupled to MS. *J Sep Sci* 2008, 31, 2913-2923.
 35. Zheng, Q., Zhang, H., Chen, H. Int Integration of online digestion and electrolytic reduction with mass spectrometry for rapid disulfide-containing protein structural analysis. *J Mass Spectrom* 2013, 353, 84-92.
 36. Ma, J., *et al.* Online integration of multiple sample pretreatment steps involving denaturation, reduction, and digestion with microflow reversed-phase liquid chromatography-electrospray ionization tandem mass spectrometry for high-throughput proteome profiling. *Anal Chem* 2009, 81, 6534-6540.



Chapter 3

Fast analysis of antibody-derived therapeutics by automated multidimensional liquid chromatography – mass spectrometry

3

**Sanne Pot^{1#}, Christoph Gstöttner^{1#}, Katrin Heinrich², Sina Hoelterhoff³, Ingrid Grunert²,
Michael Leiss², Anja Bathke³ and Elena Domínguez-Vega¹**

[#]These authors contributed equally to this work

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

² Roche Pharma Technical Development, Penzberg, Germany

³ F. Hoffmann-La Roche, Basel, Switzerland

Reprinted and adapted with permission from *Analytica Chimica Acta*, 2021, 1184, 339015, DOI:
10.1016/j.aca.2021.339015

3.1 Abstract

Characterization of post-translational modifications (PTMs) of therapeutic antibodies is commonly performed by bottom-up approaches, involving sample preparation and peptide analysis by liquid chromatography-mass spectrometry (LC-MS). Conventional sample preparation requires extensive hands-on time and can increase the risk of inducing artificial modifications as many off-line steps - denaturation, disulfide-reduction, alkylation and tryptic digestion - are performed. In this study, we developed an on-line multidimensional (mD)-LC-MS bottom-up approach for fast sample preparation and analysis of (formulated) monoclonal antibodies and antibody-derived therapeutics. This approach allows on-column reduction, tryptic digestion and subsequent peptide analysis by RP-MS. Optimization of the 1D -and 2D flow and temperature improved the trapping of small polar peptides during on-line peptide mapping analysis. These adaptations increased the sequence coverage (95-98% versus 86-94% for off-line approaches) and allowed identification of various PTMs (*i.e.* deamidation of asparagine, methionine oxidation and lysine glycation) within a single analysis. This workflow enables a fast (<2 hours) characterization of antibody heterogeneities within a single run and a low amount of protein (10 µg). Importantly, the new mD-LC-MS bottom-up method was able to detect the polar, fast-eluting peptides: Fc oxidation at Hc-Met-252 and the Fc N-glycosylation at Hc-Asn-297, which can be challenging using mD-LC-MS. Moreover, the method showed good comparability across the different measurements (RSD of retention time in the range of 0.2-1.8% for polar peptides). The LC system was controlled by a standard commercial software package which makes implementation for fast characterization of quality attributes relatively easy.

3.2 Introduction

Therapeutic monoclonal antibodies (mAbs) and new antibody formats are highly complex bio-molecules potentially affected by chemical and enzymatic modifications, also known as post-translational modifications (PTMs) [1, 2]. PTMs - including lysine glycation, deamidation of asparagine, aspartic acid isomerization and methionine oxidation - can occur during manufacturing or storage [3, 4]. As these modifications can affect the pharmacological properties (*i.e.* stability, safety and potency) of the final drug product, their assessment is mandatory to ensure the product quality, efficacy and consistency [5-7]. Structural characterization of mAbs at an amino acid level is commonly performed by bottom-up approaches, involving sample preparation followed by peptide analysis by liquid chromatography–mass spectrometry (LC-MS) [8, 9]. Sample preparation for conventional bottom-up approaches includes many off-line steps such as denaturation, disulfide bridge reduction, alkylation and enzymatic digestion. As a consequence, characterization of mAbs requires extensive hands-on time and can increase the risk of inducing artificial modifications [6, 9]. Furthermore, off-line digestion comes with a few drawbacks, such as long incubation time (several hours), low efficiency and trypsin autodigestion [5]. Hence, there is a high demand for automation of bottom-up approaches which can be used for routine characterization of quality attributes. Bauer et al. [6] proposed an automated sample preparation procedure for peptide mapping using a liquid handling robot, allowing denaturation, disulfide bridge reduction, alkylation and tryptic digestion in four hours. On the downside, these automated robotic platforms are only beneficial when analyzing a large number of samples. Multidimensional LC-MS bridges the automated sample preparation process and the subsequent LC-MS analysis within one set-up, having the opportunity to analyze various samples step-by-step. This technique has been widely adopted for mAb characterization due to the introduction of commercially available 2D-LC systems and immobilized enzyme reactors (IMERs) [7]. For instance, immobilized trypsin reactors have proven to be more efficient than off-line digestion as digestion can be achieved in minutes [5, 10, 11]. In recent years, several studies implemented on-column disulfide bridge reduction and immobilized trypsin reactors between two separation techniques, allowing a full automated bottom-up approach for a peak characterization [5, 7, 8, 11-13]. Gstöttner et al. [5] developed a 4D-LC set-up for the HPLC-peak characterization of mAb charge variants using a commercial 2D-LC system of Agilent which was extended with additional modules. Online fractions of the cation exchange (CEX) separation were taken, followed by reduction, tryptic digestion and peptide analysis by RP-MS. Similarly, Camperi et al. [12] characterized protein modifications of mAbs from cell culture supernatants using protein affinity chromatography (ProA) in the first dimension. Compared to conventional off-line bottom-up approaches, these new automated platforms enable faster identification and characterization of protein modifications (hours versus days) [5, 12, 14]. In addition, upon comparing the off-line and on-line approach, Goyon et al. [11] reported less artificial modifications within the online

approach. Altogether, these automated workflows rely on pre-separation of the antibody in the first dimension, followed by online fraction collection of the separated antibody variants. Hereafter, each collected fraction undergoes on-column reduction, tryptic digestion and RP-MS analysis. These online approaches are very powerful to selectively fractionate unknown peaks from quality control (QC) methods for further bottom-up characterization. However, one of the remaining challenges in on-line bottom-up approaches using mD-LC systems is trapping of small polar peptides in the RP-C18 column. This is a consequence of the high ACN content necessary to elute the reduced chains from the reduction-trap column [5]. Furthermore, many mD-LC platforms are based on non-commercial 2D-LC systems without integrated software. This requires a certain amount of expertise and experience making it very complex to implement mD-LC for routine characterization of quality attributes [15].

In the present study, we developed an on-line bottom-up strategy using a mD-LC-MS set-up for fast sample preparation and analysis of mAbs (reduction – tryptic digestion – RP-MS). The proposed mD-LC-MS approach relies on the direct injection of (formulated) mAb samples without any pre-separation and/or fractionation of antibody variants. The mD-LC set-up is based on a commercial Agilent 2D-LC system with easy-to-use hardware and software. The use of an immobilized trypsin column reduced the digestion time significantly and enabled the digestion in minutes instead of hours and without manual sample preparation time. Furthermore, various LC parameters such as 1D and 2D flow rate and column temperature were optimized providing better trapping of polar peptides resulting in higher sequence coverages.

3.3 Experimental section

3.3.1 Reagents and samples

Tris(hydroxymethyl) amino methane (TRIS) base ($\geq 99.8\%$), DL-Dithiothreitol (DTT) ($\geq 99\%$), hydrogen chloride (ACR reagent grade 37%), D-(+)-Glucose ($\geq 99.5\%$), L-Histidine ($\geq 99.5\%$), sucrose ($\geq 99\%$) were purchased from Sigma-Aldrich (St. Louis, MO). Acetonitrile (Ultra LC-MS grade) was provided by Actua-All Chemicals (Randmeer, Oss, The Netherlands). Formic acid ($\geq 99\%$) was purchased from VWR Chemicals (Radnor, PA). Calcium Chloride Dihydrate was obtained from J.T. Baker (99%) (Deventer, The Netherlands). Hydrogen peroxide (30%) and trifluoroacetic acid (TFA) were purchased from Merck (Darmstadt, Germany). Tween-20 was provided by Roche Diagnostics (Mannheim, Germany). Deionized water was obtained from a Purelab ultra (ELGA Labwater, Ede, The Netherlands). Formulated, conventional and bispecific, mAbs were provided by F. Hoffmann-La Roche (Basel, Switzerland) and Roche Diagnostics (Penzberg, Germany).

3.3.2 Sample preparation

Formulated mAb1 (119 mg/mL in 5 mM L-histidine, 60 mM α,α -trehalose, 0.01% polysorbate 20, pH 6.0) and BsAb1 (120 mg/mL in 20 mM His-acetate, 25 mM NaCl, 160 mM Sucrose, 7 mM L-methionine, 0.04% polysorbate 20, pH 5.5 ± 0.2) were simply diluted to 1 mg/mL before mD-LC-MS experiments with mobile phase A of the first dimension. mAb2 and BsAb2 (25.2 mg/mL) were buffer exchanged in a formulation buffer containing 240 mM sucrose, 0.02% polysorbate, 10 mM L-methionine in 20 mM His-HCl, pH 5.5. Prior mD-LC-MS experiments, the samples were diluted to 1 mg/mL with mobile phase A of the first dimension. Additionally, mAb2 was stressed under various stress conditions. Oxidative stress was performed by mixing mAb2 with hydrogen peroxide 0.2% (H_2O_2) (v/v). The mixture was incubated at room temperature for 24 hours. Forced deamidation was achieved by incubating mAb2 with a final concentration of 10 mg/mL with 200 mM Tris-HCl, pH 9 at 37°C for 3 and 7 days. Glycation was induced by incubating mAb2 with a final concentration of 12.5 mg/mL with 1M glucose at room temperature for 7 days. All stressed samples were kept at -20°C until analysis and diluted to 1 mg/mL with mobile phase A of the first dimension before mD-LC-MS experiments.

3.3.3 Instrumentation

mD-LC system Instrumentation

The system was completely configured using standard modules from the Agilent 1290 Infinity II 2D-LC system (Agilent Technologies, Waldbronn, Germany). The system consisted of the following modules: 1260 Bio Multisampler (G5668A), 1290 High Speed Binary Pump (G7120A), two 1260 Bio-Inert Quaternary Pump (G5654A), 1260 Multicolumn Thermostat (G7116A) configured with a 6 port/2 position valve (G1170A), two 1290 Multicolumn Thermostats (G7116B) configured with a 6 port/2 position valve (G5631) and a 6port/2position external valve (G1170A). The system was controlled by one OpenLab software package. An additional self-designed macro "valve event plugin" from ANGI (Gesellschaft für angewandte Informatik, Karlsruhe, Germany) was used to generate a start signal for the MS system. The mD-LC was directly connected to a QTOF mass spectrometer (Bruker Daltonics, Bremen, Germany), equipped with an electrospray ionization (ESI) ion source. A detailed overview of the mD-LC set-up (gradients, flows and valve positions) can be found in **Table S1**.

First dimension: Trapping and Reduction

An AdvancedBio RP cartridge (2.1 I.D. x 12.5 mm) (Agilent Technologies) was used for subsequent trapping and reduction of 10 μg therapeutic antibody samples. Mobile phases were composed of water (A) and ACN (B) containing both 0.1% FA and 0.05% TFA. Mobile phase C was composed of 20 mM DTT in 50 mM Tris, pH 7.5. The column temperature was

set to 60°C. The antibody is first trapped onto the RP cartridge using 1% B and subsequently reduced using mobile phase C, resulting in heavy chains (Hc) and light chains (Lc). After reduction, the column was flushed with 1% B to remove any remaining DTT followed by the elution of mAb chains at 60% B at a flow rate of 45 µL/min. Next, a cleaning and reconditioning step was performed using 100% B for 5 minutes followed by 1% B for 5 minutes.

Second dimension: Tryptic digestion

During the elution step the valves located after the first and second dimension were set in-line allowing the transfer of reduced mAb chains from the AdvancedBio RP cartridge to a StyroZyme TPCK-Trypsin column (2.1 I.D. x 100 mm) (OraChrom, Massachusetts, USA). The column temperature was set to 37°C. Mobile phase A consisted of 50 mM Tris and 10 mM CaCl₂, pH 8 and mobile phase B of ACN. Mobile phase A was continuously introduced via a T-junction. The trypsin column was set in-line with 1D-RP and 3D-RP for 27 minutes using mobile phase A at a flow rate of 450 µL/min followed by a cleaning and reconditioning step. The column was flushed with 50% B at 1 mL/min for 5 minutes followed by 100% A at 1 mL/min for 5 minutes, this procedure was performed two times.

Third dimension: Peptide mapping

Tryptic peptide mixtures were separated on a RP XSelect CSH C18 column (Waters CSH RP-C18, particle size 3.5 µm, 2.1 mm I.D. x 100 mm) (Waters, Mildford, MA). The mobile phases were composed of water (A) and ACN (B) containing both 0.1% FA. The trapped tryptic peptides were first desalted using 1% B for 5 minutes. After desalting, the RP-C18 column was set in-line with the MS and the peptides were subsequently separated using a gradient from 1% to 35% B in 45 minutes and switching to 80% B for 6 minutes. The flow rate was set at 300 µL/min and the column temperature was set to 55°C.

Mass spectrometry

HyStar software (Bruker Daltonics) was used for the MS data acquisition. The acquisition of peptides was performed in positive ion mode in a m/z range of 150-2500. The in-source collision energy was 7 eV. The source parameters were set as follows: 500 V end plate offset, 4500 V capillary voltage, 5.5 L/min dry gas flow rate, 180°C dry gas temperature and 2.2 bar nebulizer gas pressure. The MS/MS spectra were generated by using fragmentation mode collision-induced dissociation (CID). The isCID energy and collision energy were set to 5 eV and 70 eV. MS data processing (*i.e.* assignment and peak integration) was manually performed using DataAnalysis (Bruker Daltonics) and Skyline 20.2 software (MacCoss Lab, U Washington).

3.4 Results and discussion

3.4.1 mD-LC-MS method development

In the present study, we developed an automated on-line bottom-up strategy using a mD-LC-MS system as shown in **Figure 1**. This on-line bottom-up strategy enables on-column reduction, tryptic digestion and peptide analysis by RP-MS.

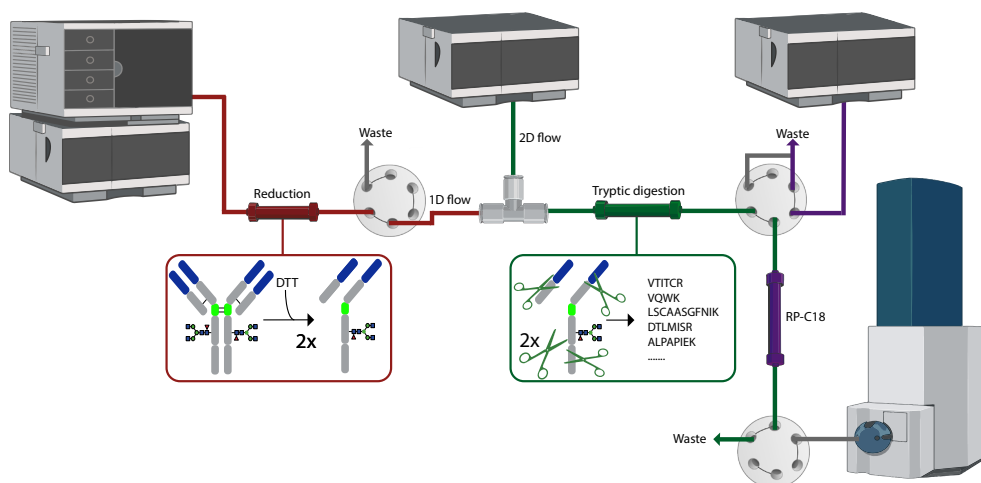


Figure 1. Schematic representation of the automated on-line bottom-up approach for mAbs using mD-LC-MS which enables on-line on-column reduction by RP-HPLC (1D), tryptic digestion (2D) and peptide mapping by RP-MS (3D).

One major challenge in bottom-up approaches using on-column reduction is the trapping of polar peptides on the RP-C18 column, since an increased concentration of acetonitrile (~60% ACN) is necessary for the elution of mAb chains from the reduction-trap column. As shown in earlier studies the amount of ACN was already lowered to ~11.6% but further dilution was limited due to the relatively low backpressure limit of the trypsin reactor (<200 bar) and high backpressure of the RP-C18 column (Waters BEH RP-C18, particle size 1.7 μm , 50 x 2.1 mm) [5]. Consequently, many polar peptides were not trapped on the RP-C18 column and therefore not detected during peptide mapping resulting in low sequence coverages (Light chain, Lc: 94% and Heavy chain, Hc: 86%). In this study, various parameters were optimized to improve trapping of small polar peptides. To address this challenge a column with optimized characteristics such as particle size, length and stationary phase was chosen. More precisely, the particle size was increased from 1.7 to 3.5 μm (Waters CSH RP-C18, particle size 3.5 μm , 2.1 mm I.D. x 100 mm) to get the possibility of reducing the amount of ACN even more while observing the pressure limit of the trypsin reactor. Besides, the column was longer (100mm) and had a CSH stationary phase to improve chromatographic

separation.

Furthermore, to improve chromatographic separation and trapping of polar peptides, the selected column was longer (50 mm versus 100 mm) and consisted of a different stationary phase (CSH versus BEH) which is more selective towards basic compounds [11]. This implementation increased the sequence coverage of mAb1 from 90% to 95% for the Hc, whereas the sequence coverage for the Lc remained at 96% (**Figure 2** and **Table S2**). The increased sequence coverage can be attributed to the improved trapping of polar peptides, including “IYPTNGYTR”, “ALPAPIEK”, as the applied column has a higher selectivity for basic compounds compared to the BEH RP-C18 column [11].

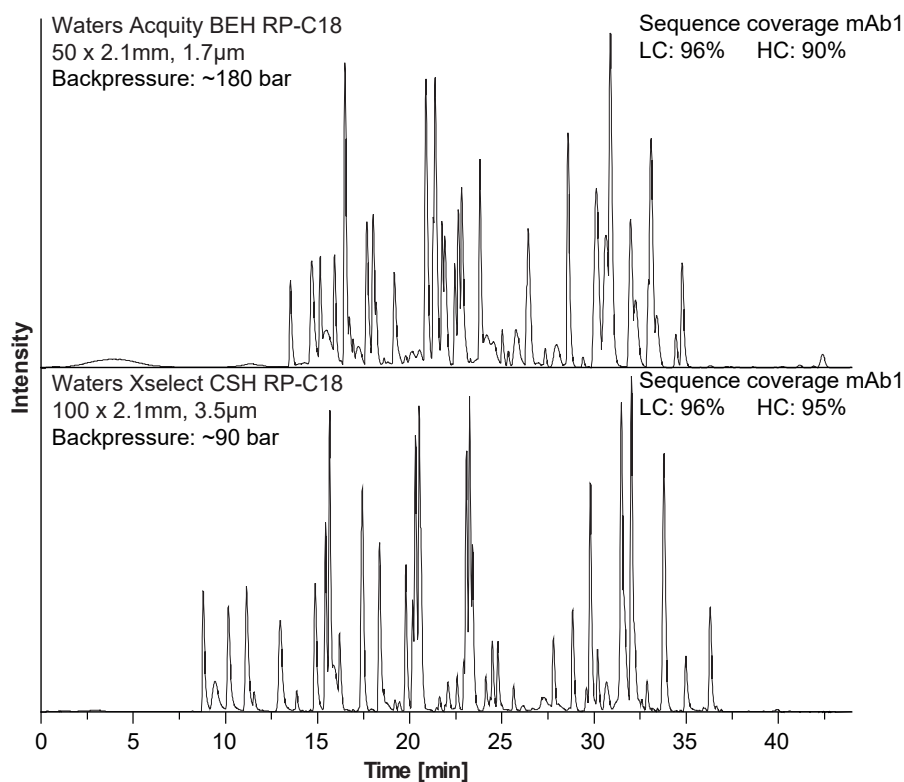


Figure 2. Comparison of two different RP-C18 columns in the mD-LC-MS set-up using mAb1, prior flow rate and temperature optimization. The Waters Acquity BEH RP-C18 column (particle size 1.7 μm, 50 x 2.1 mm) (Top) and the Waters XSelect CSH RP-C18 column (particle size 3.5 μm, 100 x 2.1 mm) (Bottom). A detailed overview of the identified peptides for both columns is shown in **Table S2**.

Moreover, the increased particle size of the CSH RP-C18 column led to a decreased backpressure (from ~180 bar to ~90 bar) when the column was set in-line with the first and second dimension. Lowering the backpressure of the RP-C18 column made the optimization of the ACN content possible. The dilution was optimized by varying the 1D- and 2D flow rates.

A decrease of the 1D flow rate from 60 $\mu\text{L}/\text{min}$ to 50 or 45 $\mu\text{L}/\text{min}$ resulted in a decrease of ACN from 11.6% to 10% or 9.2% (**Figure 3**). Dilution of the ACN content from 11.6% to 9.2% resulted in trapping of several polar peptides including the Fc oxidation at Hc-Met-252 (according to the Kabat numbering [16]). This oxidized peptide is one of the most polar peptides eluting early in RP-C18 and is particularly important because it decreases binding to the neonatal receptor (FcRn) reducing the half-life of the antibody [1]. To further dilute the amount of ACN, the 2D flow rate was increased from 250 to 350 or 450 $\mu\text{L}/\text{min}$ which led to increased trapping of polar peptides. A flow rate of 450 $\mu\text{L}/\text{min}$ lowered the ACN content from 9.2% to 5.5% and led to a sequence coverage of 97% and 96% for the Hc and Lc. Hence, this allowed the identification of the glycopeptide “TKPREEQYNSTR”, which plays an important role for complement-dependent cytotoxicity (CDC) and antibody-dependent cell-mediated cytotoxicity (ADCC) [17].

Additionally, the temperature dependence of polar peptides trapping in RP-C18 was investigated by performing the trapping and separation on the RP-C18 at 85°C, 75°C, 65°C and 55°C (**Figure 3**). As expected, an increase of the temperature led to faster elution rather than trapping of polar peptides, as the retention factor is inversely proportional to the temperature [18], whereas a low temperature improved trapping of polar peptides without affecting the overall separation. For instance, the glycopeptide “TKPREEQYNSTR” was better retained at 55°C compared to higher temperatures. Furthermore, lower temperatures enhance the column lifetime and decrease the risk of generating artificial PTMs during trapping. For this reason the temperature was set to 55°C.

In summary, the use of a different RP-C18 column (Waters CSH RP-C18, particle size 3.5 μm , 2.1 mm I.D. x 100 mm) together with reduction of ACN content (from 11.6% to 5.5%, respectively) and optimized column temperature (from 65 to 55°C) improved trapping of polar peptides on the RP-C18 column. Subsequently, the sequence coverage for the Hc increased from 90% to 97%, whereas the Lc remained at 96%. Besides, the optimized method allowed detection of polar Fc peptides, which play a critical role in effector functions and recycling. Only some very small polar peptides, containing less than six amino acid residues, on the Lc (*i.e.* GEC and FSGR) and Hc (*i.e.* QAPGK, VDK, VEPK, and K) were not detected. Note that these small peptides are often also missed in conventional offline bottom-up workflows.

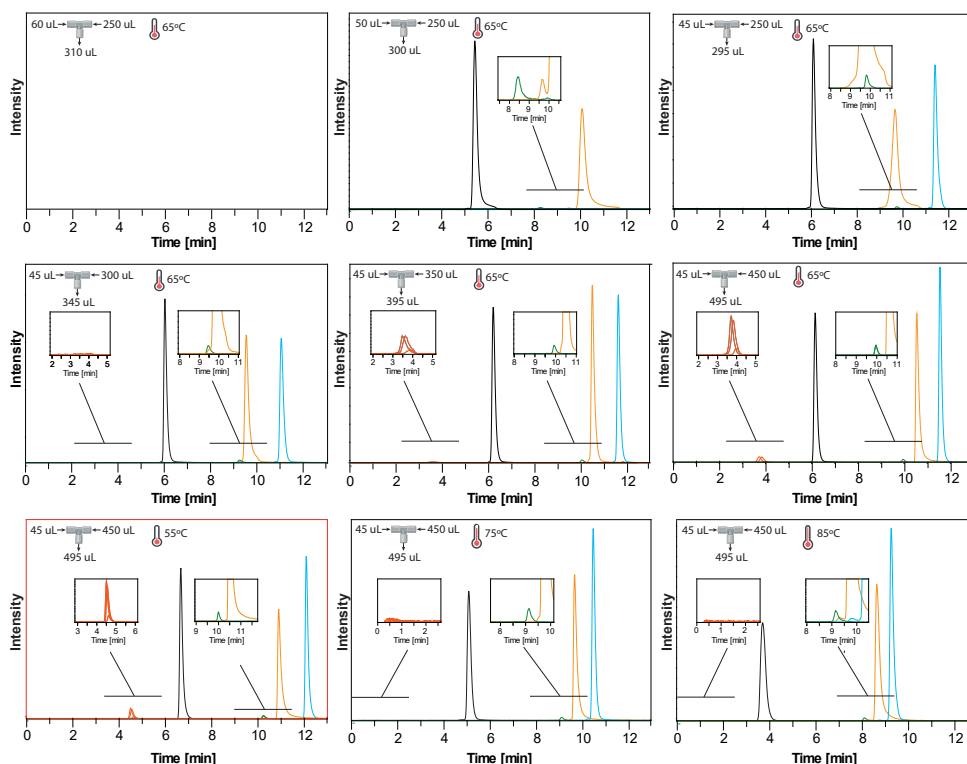


Figure 3. Optimization of the polar peptide trapping in the mD-LC-MS set-up using mAb1. The method was optimized by evaluating the 1D Flow (A) 2D Flow (B) and temperature in the 3D RP-C18 column (C). ● TKPREEQYNSTR glycosylation ● DTLMIISR oxidation ● VTITCR ● AEDTAVYYCSR ● VDNALQSGNSQESVTEQDSK. A detailed overview of all the optimization steps can be found in **Table S2**. Final conditions are highlighted in red.

3.4.2 Method evaluation for the direct analysis of formulated antibodies and new antibody-derived therapeutics

To demonstrate the performance of the mD-LC-MS set-up for the analysis of mAbs, both desalted and formulated mAb1 samples were analyzed. No differences in sequence coverage or separation (**Figure S1**) was observed between both samples indicating that formulation additives (*i.e.* polysorbate, saccharose) are not interfering with the analysis. Hence, injection of formulated mAbs can be performed directly without sample handling reducing total analysis time and the risk of inducing artificial modifications. To prove the generic applicability of the method for the analysis of diverse antibody formats, two conventional mAbs and two different bispecific mAb (BsAbs) were analyzed with the mD-LC-MS set-up. **Figure 4** shows the base peak chromatograms (BPCs) obtained for the four analyzed

antibodies. Comparable sequence coverages between 95-98% were obtained, highlighting the applicability of this approach for new antibody-derived therapeutics. Importantly, for all antibodies the Fc glycopeptides could be identified. The total analysis time, including on-column reduction, tryptic digestion and peptide analysis by RP-LC-MS was performed in less than 2 hours. Compared to conventional bottom-up strategies using manual sample preparation (offline approach), the total analysis time is considerably lower (<2 versus >12 hours), lower amount of protein is needed (10 µg versus 350 µg) and comparable sequence coverages are observed (≥95%) [8, 14, 19]. Furthermore, often mD-LC-MS bottom-up approaches are rather complex due to the use of non-commercial 2D-LC systems without integrated software [8, 11]. The proposed workflow is based on a commercial 2D-LC system and focuses on the analysis of the whole sample. Therefore, there is no need for online fraction collection of separated mAb variants which requires less modules (one pump and three valves less) and no plug-in. As a result, the mD-LC method consists of only three dimensions and is controlled by a single commercial software (OpenLAB CDS Software from Agilent) which makes implementation relatively easy and simplifies the work-around.

The method was further evaluated by assessing the repeatability (intra- and inter-day precision). To assess the intra-day precision, six replicates of formulated mAb1 were injected on one day using the same sample and mobile phases (**Figure S2**). The inter-day precision was assessed by the analysis of triplicates of formulated mAb1 on two different days using freshly prepared mobile phases and sample (**Figure S3**). For both intra- and inter-day precision, all measurements showed the same sequence coverage (Lc: 96% and Hc: 97%). The repeatability was further evaluated by the relative standard deviation (RSD) of the retention time of four polar peptides. Low intra- and inter-day variation was observed for the retention time (RSD 0.1-1.8%) of the four selected polar peptides (**Table S3**).

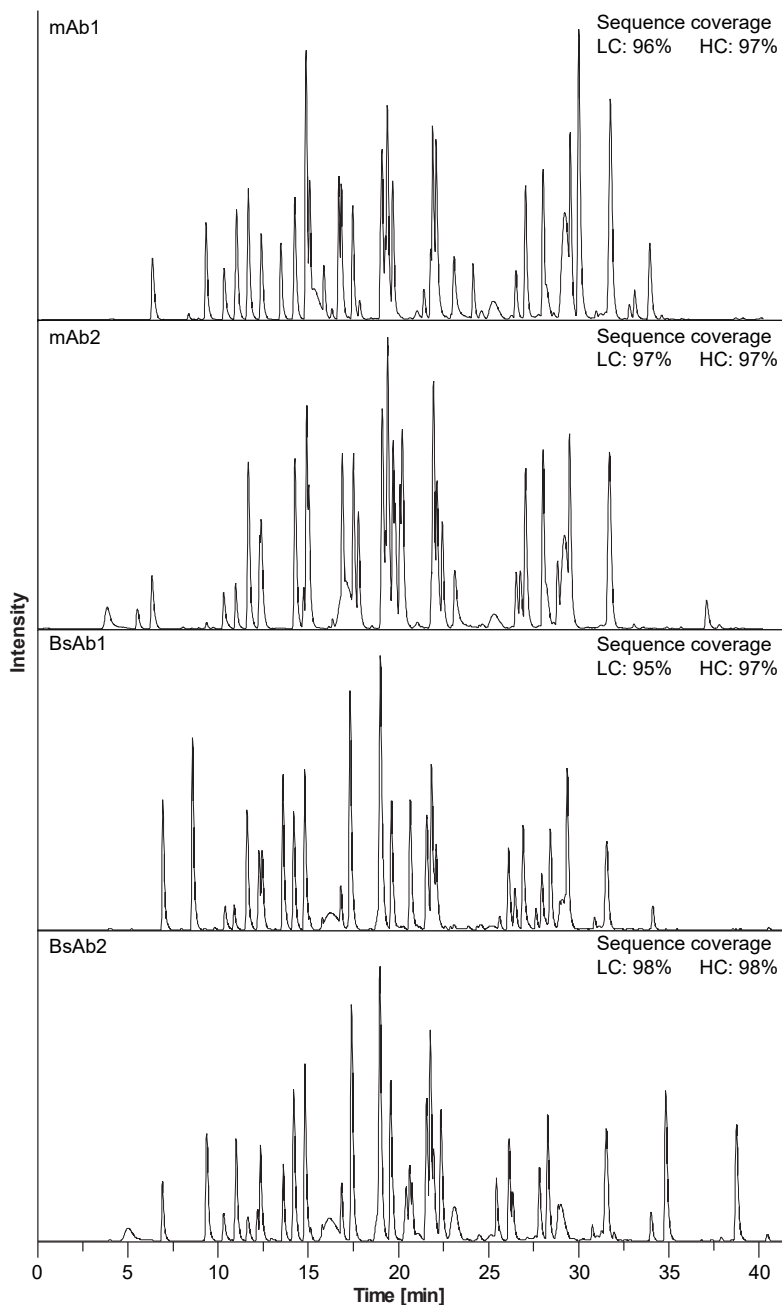


Figure 4. mD-LC-MS BPCs of formulated conventional mAbs (mAb1 and mAb2) and formulated bsAbs (BsAb1 and BsAb2) and their obtained sequence coverages, respectively. mAb1 (Lc: 96 % and Hc 97 %), mAb2 (Lc and Hc: 97 %), BsAb1 (Lc: 95% and Hc: 97%) and BsAb2 (Lc and Hc: 98%).

3.4.3 Analysis of stressed antibodies and PTM assessment

Next to sequence assessment, characterization of PTMs in therapeutic mAbs is highly important as these may alter the stability and/or affect the biological function of the protein [19]. PTMs such as deamidation of asparagine, methionine oxidation and lysine glycation are particularly important and can be considered as critical or non-critical quality attributes, which should be monitored carefully. To evaluate the capability of the established method to determine PTMs, a therapeutic antibody (mAb2) was exposed to various stress conditions (i.e. elevated pH, hydrogen peroxide, elevated glucose levels) [19] and analyzed with the proposed mD-LC-MS platform. The results were compared to a conventional off-line method as reported by Habberger et al. [19]. To induce deamidation, mAb2 was exposed to pH 9 at 37°C for 3 and 7 days. Analysis of the stressed samples revealed new signals eluting after their unmodified peptide with a mass increase of ~1 Da and were assigned as the deamidated variants and confirmed by MS/MS analysis. All four potential Asn sites were found deamidated (Lc-Asn-deamid-92, Lc-Asn-deamid-50, 52, 53, Lc-Asn-deamid-30, 31 and Hc-Asn-deamid-384, 389, 390) (**Figure S4**) with increasing intensities after 3 and 7 days (**Table S4**). Similarly as Habberger et al. [19], the most susceptible deamidation sites were found on Lc-Asn-92 and Hc-Asn-384, 389, 390, whereas Lc-Asn-50, 52, 53 and Lc-Asn-30,31 were less deamidation prone. Deamidation on Lc-Asn-92, which is particularly important as it is known to decrease the antigen binding site [19], increased from 15% to 36% after 7 days. Deamidation on Hc-Asn-384, 389, 390 and Lc-Asn-30, 31 resulted in multiple peaks as deamidation occurs on multiple amino acid positions in the respective tryptic peptide. Furthermore, methionine oxidation sites were assessed by the analysis of an H₂O₂ stressed mAb2 sample. The developed method was able to detect all six oxidation sites, including the oxidation at Hc-Met-252 (**Figure S5**). This peptide could not be detected with our previous mD-LC-MS method, due to the inability to trap polar peptides. However, Fc oxidation at Hc-Met-252 is of functional-relevance for the FcRn binding and therefore should be monitored carefully [20]. The most susceptible oxidation sites were found on Hc-Met-100c, Hc-Met-252 and Hc-Met-428 (from 1-2% to 97-98% after H₂O₂ stress), whereas Hc-Met-48 and Hc-Met-80 were less prone to oxidation and remained 1%, respectively (**Table S4**). In Addition, oxidation on Hc-Met-4, which is suspected to have an effect on the mAb2 target binding activity [19], increased from 1% to 58%, upon oxidative stress. Finally, lysine glycation was assessed by measuring a glucose stressed mAb2 sample. Depending on the mAb, glycation can have a detrimental effect on antigen binding as shown in several reports [21, 22]. Glycation on lysine residues affects the digestion pattern of trypsin, resulting in miss-cleaved peptides with an additional mass of +162 Da. As shown in **Figure S6**, both Lys-glycation sites (Hc-Lys-62, and Hc-Lys-33) were identified as miss-cleaved peptides as reported by Habberger et al. [19]. For Hc-Lys-33 the glycation level increased from 3% to 11%, whereas Hc-Lys-62 showed a minimal increase from 0.1% to 3%, respectively (**Table S4**). While these two glycation-sites show negligible to moderate effect on the binding activity of this mAb [19], it is still very important to be able to identify susceptible glycation-sites in the CDR region.

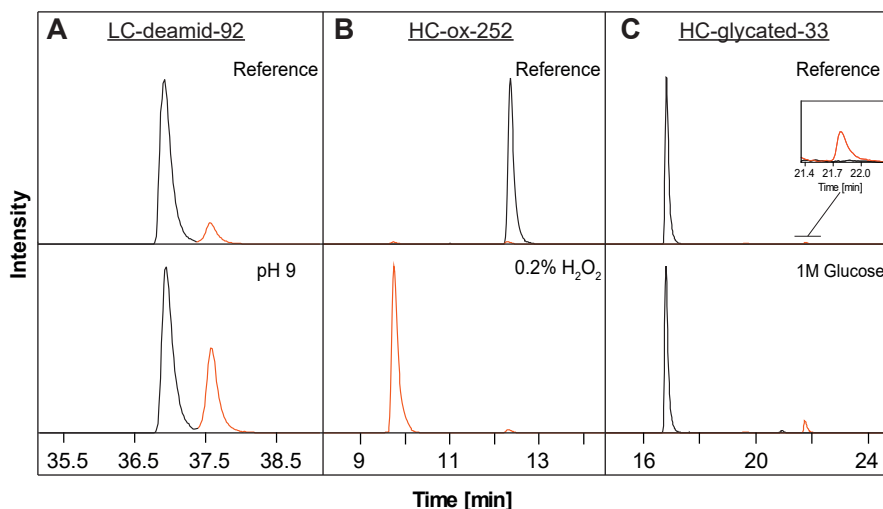


Figure 5. Extracted ion chromatograms (EIC) of A) Lc-Asn-92 (black) and deamidated Lc-Asn-92 (orange), B) Hc-Met-252 (black) and oxidized Hc-Met-252 (orange) and C) Hc-Lys-33 (black) and glycated Hc-Lys-33 (orange) of the reference material (upper panel) and stressed material (lower panel) of mAb2.

3.5 Conclusion

In this work we propose a new mD-LC-MS bottom-up strategy for the fast online characterization of (formulated) antibody-derived therapeutics (conventional and bispecific mAbs) without the need of any sample preparation. The method enables on-column reduction, tryptic digestion and peptide analysis by RP-MS. The easy-to-use hardware and software allow characterization of antibody-derived therapeutics in a straightforward manner. Compared to conventional off-line bottom-up approaches, the proposed workflow requires less hands-on time as no off-line steps need to be performed. Optimization of the method enabled trapping of polar peptides and improved the sequence coverage (95-98%) compared to previous mD-LC-MS methods (86-94%). Furthermore, the method permits to detect polar peptides, including Fc oxidation at Hc-Met-252 and Fc glycosylation at Hc-Asn-297, which could not be detected with previous mD-LC-MS methods. Furthermore, several PTMs (*i.e.* deamidation of asparagine, lysine glycation and methionine oxidation) can be detected and characterized within a single analysis. Therefore, the proposed mD-LC-MS bottom-up approach allows direct and fast characterization of antibody-derived therapeutics in less than 2 hours using only 10 µg of protein. Importantly, the proposed workflow is faster, requires lower amount of protein and obtains similar sequence coverages as conventional off-line bottom-up approaches. We believe that this method can considerably increase throughput for the characterization of antibody derived therapeutics in pharmaceutical labs.

3.6 Supporting information

Supplementary information is available free of charge on the Analytica Chimica Acta website:
DOI: 10.1016/j.aca.2021.339015

3.7 Acknowledgments

This work was supported by the Analytics for Biologics project (Grant agreement ID 765502) of the European Commission.

3.8 References

1. Gahoual R., *et al.* Detailed Characterization of Monoclonal Antibody Receptor Interaction Using Affinity Liquid Chromatography Hyphenated to Native Mass Spectrometry. *Anal Chem* 2017, 89, 5404-5412.
2. S. Fekete S., *et al.* Chromatographic, Electrophoretic, and Mass Spectrometric Methods for the Analytical Characterization of Protein Biopharmaceuticals. *Anal Chem* 2016, 88, 480-507.
3. Liu H., *et al.* In vitro and in vivo modifications of recombinant and human IgG antibodies. *MAbs* 2014, 6, 1145-1154.
4. Beyer B., *et al.* Microheterogeneity of Recombinant Antibodies: Analytics and Functional Impact. *Biotechnol J* 2018, 13.
5. Gstöttner C., *et al.* Fast and Automated Characterization of Antibody Variants with 4D HPLC/MS. *Anal Chem* 2018, 90, 2119-2125.
6. Bauer L. G., *et al.* Monitoring modifications in biopharmaceuticals: Toolbox for a generic and robust high-throughput quantification method. *J Chromatogr B* 2020, 1148, 122134.
7. Graf T., *et al.* Recent advances in LC-MS based characterization of protein-based bio-therapeutics - mastering analytical challenges posed by the increasing format complexity. *J Pharm Biomed Anal* 2020 186, 113251.
8. Camperi J., *et al.* Development of a 3D-LC/MS Workflow for Fast, Automated, and Effective Characterization of Glycosylation Patterns of Biotherapeutic Products. *Anal Chem* 2020, 92, 4357-4363.
9. Mouchahoir T., Schiel J. E. Development of an LC-MS/MS peptide mapping protocol for the NISTmAb. *Anal Bioanal Chem* 2018, 410, 2111-2126.
10. Kim J. H., *et al.* Integrated sample preparation methodology for proteomics: analysis of native proteins. *Anal Chem* 2013, 85, 8039-8045.
11. Goyon A., *et al.* From proof of concept to the routine use of an automated and robust multi-dimensional liquid chromatography mass spectrometry workflow applied for the charge variant characterization of therapeutic antibodies. *J Chromatogr A* 2020, 1615, 460740.

12. Camperi J., *et al.* Fast and Automated Characterization of Monoclonal Antibody Minor Variants from Cell Cultures by Combined Protein-A and Multidimensional LC/MS Methodologies. *Anal Chem* 2020, 92, 8506-8513.
13. Goyon A., *et al.* Streamlined Characterization of an Antibody–Drug Conjugate by Two-Dimensional and Four-Dimensional Liquid Chromatography/Mass Spectrometry. *Anal Chem* 2019, 91, 14896-14903.
14. Camperi J., *et al.* Multi-dimensional LC-MS: the next generation characterization of antibody-based therapeutics by unified online bottom-up, middle-up and intact approaches. *Analyst* 2021 146, 747-769.
15. Ehkirch A. C. A. The Current Status and Future of Two- and Multidimensional Liquid Chromatography in Pharmaceutical R&D and QC. *LCGC* 2020.
16. Johnson G., Wu T. T. Kabat Database and its applications: 30 years after the first variability plot. *Nucleic Acids Res* 2000, 28, 214-218.
17. Zheng K., Bantog C., Bayer R. The impact of glycosylation on monoclonal antibody conformation and stability. *MAbs* 2011, 3, 568-576.
18. Schöbinger M., Klein O. J., Mitulović G. Low-Temperature Mobile Phase for Peptide Trapping at Elevated Separation Temperature Prior to Nano RP-HPLC-MS/MS. *Separations* 2016, 3.
19. Habegger M., *et al.* Assessment of chemical modifications of sites in the CDRs of recombinant antibodies. *MAbs* 2014, 6, 327-339.
20. Wang W., *et al.* Impact of methionine oxidation in human IgG1 Fc on serum half-life of monoclonal antibodies. *Mol Immunol* 2011, 48, 860-866.
21. Mo J., *et al.* Quantitative analysis of glycation and its impact on antigen binding. *MAbs* 2018, 10, 406-415.
22. Wei B., *et al.* Glycation of antibodies: Modification, methods and potential effects on biological functions. *MAbs* 2017, 9, 586-594.





Chapter 4

Monitoring glycation levels of a bispecific monoclonal antibody at subunit level by ultrahigh resolution MALDI FT-ICR mass spectrometry

4

Christoph Gstöttner¹, Dietmar Reusch², Markus Habberger², Irina Dragan¹, Peter van Veelen¹, David P. A. Kilgour³, Yury O. Tsybin⁴, Yuri E.M. van der Burgt¹, Manfred Wuhrer¹, Simone Nicolardi¹

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

²Roche Pharma Technical Development, Penzberg, Germany

³Department of Chemistry, Nottingham Trent University, Nottingham, United Kingdom

⁴Spectroswiss, EPFL Innovation Park, Lausanne, Switzerland

"This is an Accepted Manuscript of an article published by Taylor & Francis in mAbs on 18 Dec 2019, available online at <https://doi.org/10.1080/19420862.2019.1682403>"

4.1 Abstract

Bispecific monoclonal antibodies (BsAbs) are engineered proteins with multiple functionalities and properties. The “bi-specificity” of these complex biopharmaceuticals is a key characteristic for the development of novel and more effective therapeutic strategies. The high structural complexity of BsAbs poses a challenge to the analytical methods needed for their characterization. Modifications of the BsAb structure, resulting from enzymatic and non-enzymatic processes, further complicate the analysis. An important example of the latter type of modification is glycation, which can occur in the manufacturing process, during storage in formulation or *in vivo* after application of the drug. Glycation affects structure, function and stability of monoclonal antibodies, and consequently, detailed analysis of glycation levels is required. Mass spectrometry (MS) plays a key role in the structural characterization of monoclonal antibodies and top-down, middle-up and middle-down MS approaches are increasingly used for the analysis of modifications. Here, we apply a novel middle-up strategy, based on IdeS digestion and matrix-assisted laser desorption ionization (MALDI) Fourier transform ion cyclotron resonance (FT-ICR) MS, to analyze all six different BsAb subunits in a single high-resolution mass spectrum, namely two light chains, two half fragment crystallizable regions and two Fd' regions, thus avoiding upfront chromatography. This method was used to monitor glycation changes during a 168h forced-glycation experiment. In addition, hot spot glycation sites were localized using top-down and middle-down MALDI-in-source decay FT-ICR MS, which provided complementary information compared to standard bottom-up MS.

4.2 Introduction

Bispecific monoclonal antibodies (BsAbs) are recombinant proteins designed to recognize two different epitopes on either one or two different antigens.¹ This characteristic is derived from the combination of specificities of two different monoclonal antibodies (mAbs) and has shown great potential for the development of novel therapeutic strategies.^{2,3} In immunoglobulin G (IgG)-like BsAbs, each antibody molecule contains two different light chains and two different heavy chains that are (inter)connected by disulfide bonds between cysteines to form the typical Y-shaped structure.⁴ In the simplest format, a BsAb is bivalent however, valency can be increased by adjusting the format. As is the case for mAbs, the structural complexity of BsAbs is increased by post-translational modifications (PTMs) such as glycosylation – an enzymatic process that leads to the attachment of glycan structures to specific amino acids (e.g. asparagine-297 of the conserved N-linked glycosylation site of the fragment crystallizable region (Fc)).^{5,6} Furthermore, the structure of BsAbs can change during or after production as a result of glycation, backbone truncation, oxidation or deamidation.⁷ Glycation is a non-enzymatic modification that can affect function and stability of pharmaceutical proteins.^{8,9} It results from the reaction of a primary amine of a protein and an aldehyde group of a reducing sugar and occurs primarily on lysine residues and to a lesser extent on arginine residues and N-terminal amino acids. Both mAbs and BsAbs can undergo glycation during the manufacturing process, storage in formulation or once administered to the patient (during *in vivo* circulation). Determination of the glycation state is important to ensure the structural consistency of the pharmaceutical product throughout the entire manufacturing process.¹⁰⁻¹² This analysis is also important during the development phases, when susceptibility tests for glycation are combined with activity tests to design stable antibodies. Factors affecting glycation of biopharmaceuticals include the level and type of sugars in the cell culture broth, the temperature, and the pH used in the cell culture process.¹³ Although each primary amine in a mAb sequence can potentially be glycated, only a limited number of sites are glycated during production or storage. These glycation hot spots are not determined by any specific consensus motif however, the proximity to catalytic carboxylic acid amino acyl residues seems to play an important role in glycation. Since glycation levels of mAbs are usually low, under normal production conditions, mAbs are often stressed by incubation with high concentrations of reducing sugars in order to promote glycation and facilitate the identification of glycation hot spots. Such stressed or forced conditions increase the level of glycation at the hot spot sites.

The characterization of antibodies with respect to their PTMs, including glycation, is complex and requires the use of various analytical methods, in which mass spectrometry (MS) plays a key role by providing the tools for a multi-level characterization.¹⁴⁻¹⁹

For the characterization of recombinant proteins such as mAbs and BsAbs, bottom-up MS analysis is often preferred to other MS-based strategies.²⁰ This method includes chemical

reduction of disulfide bonds and enzymatic digestion (*e.g.* with trypsin), followed by peptide analysis by liquid chromatography (LC)-MS/MS, leading to extensive amino acid sequence coverage.^{21, 22} Further in-depth MS-analysis allows characterization of PTMs and disulfide bond connectivities.

For bottom-up analysis, it is noted that sample treatment can unintendedly result in additional modifications (artefacts), and that the relationships between specific combinations of PTMs are difficult to assess.²³⁻²⁷ Intact protein analysis and top-down characterization overcomes these limitations since these strategies circumvent digestion and determine the mass of the intact recombinant proteins, including those that carry additional modifications.²⁸⁻³¹ However, such MS methods that provide so-called intact proteoform profiles are less sensitive and require the use of high-end instrumentation with specific setups. Therefore, an intermediate strategy has been optimized for mAbs that consists of a single proteolytic cleavage in the protein backbone, with specific enzymes such as KGP (GingisKHAN) and IdeS, to cleave the heavy chain in the hinge region, resulting in two parts, namely a fragment antigen-binding (Fab) and Fc.^{30, 32, 33} Fab and Fc can then either be mass analyzed in the intact form (*i.e.* middle-up) or sequenced after reduction of disulfide bonds (*i.e.* middle-down).^{16,}

30, 34-41

Previously, we have developed both top- and middle-down methods for the structural analysis of mAbs using matrix-assisted laser desorption ionization (MALDI) in combination with in-source decay (ISD) fragmentation on a Fourier transform ion cyclotron (FT-ICR) MS platform.^{42, 43} These methods result in ultrahigh resolution absorption mode mass spectra that allow relatively fast and simple mAb sequence analysis without the need for chromatography, desalting steps or offline reduction of disulfide bonds, and furthermore include information of the major glycoforms. The MALDI-ISD fragments provide complementary structural information compared to other fragmentation techniques such as electron-transfer dissociation (ETD) and collision-induced dissociation (CID) that are widely used in electrospray ionization (ESI)-based MS analysis.⁴² Recently, we have applied this approach for the characterization of a sulfated mAb and localized sulfation on a tyrosine of the Lc.⁴⁴

In the current study, we apply a middle-up MS strategy to analyze all six different subunits generated after IdeS digestion and chemical reduction of Ang-2-VEGF (A2V) BsAb (*i.e.* 2 different Lc, two different Fd and two different Fc/2) in a single mass spectrum. This allows monitoring of glycation levels during a forced-glycation experiment performed on intact A2V BsAb. Sequential ISD fragment analysis (middle-down) is used to localize hot spot protein backbone glycation sites proximal to N-terminus of A2V BsAb heavy chains.

4.3 Experimental Section

4.3.1 Reagents and samples

A stock solution of formulated Ang-2-VEGF BsAb at 25 mg/mL (in 20 mM L-histidine, 25 mM sodium chloride, 160 mM sucrose, 7 mM L-methionine, 0.04% (w/v) polysorbate 20, pH 5.5 adjusted with acetic acid) was provided by Roche Diagnostics GmbH and was stored at -80 °C until use. Trifluoroacetic acid (TFA) was purchased from Thermo Fisher Scientific. Formic acid (FA) and 2',4',6'-trihydroxyacetophenone (THAP) were purchased from Honeywell. Fabricator (IdeS) was purchased from Genovis. Diammonium hydrogen citrate (DAHC), ammonium bicarbonate, 1,5-DAN (handle with care, check material safety data sheet), acetonitrile (MS grade), horse myoglobin and dithiothreitol (DTT) were purchased from Sigma-Aldrich. Glucose was purchased from Merck. Ultrapure water (Milli-Q; Merck) was used throughout this study. For intact mass analysis, a THAP solution was prepared by dissolving 7.6 mg of THAP in 375 µL acetonitrile containing 0.1% TFA, and adding 125 µL of an 18 mg/mL aqueous solution of DAHC. The matrix was stored at -20 °C for up to 2 weeks. For ISD experiments, a saturated solution of 1,5-DAN was prepared in acetonitrile/water/FA, 50%:49.95%:0.05%.

4.3.2 Forced glycation

Glycation was achieved by incubation of formulated A2V BsAb with glucose. To this end, 30 µL of formulated BsAb were mixed with 30 µL of glucose 2 M and kept at 25 °C for up to 168 h.

4.3.3 IdeS digestion of A2V BsAb

Non-glycated and glycated A2V BsAb were diluted with 100 mM NH_4HCO_3 , pH 6.8 to a final concentration of 1 mg/mL. IdeS was used according to the protocol of the manufacturer, meaning 1 mg bispecific mAb was digested with 1000 units of IdeS for 1.5 h at 37 °C. The incubation was quenched using a final concentration of 0.05% TFA and the sample was used immediately.

4.3.4 Chemical reduction of disulfide bonds

Reduction of the disulfide bonds of A2V BsAb was performed using DTT. To this end, a 1 M DTT solution was prepared freshly and added to intact or IdeS-digested BsAb, to a final concentration of 20 mM. Reduction was thus performed for 1 h at 60 °C.

4.3.5 Desalting by C4-based solid-phase extraction

Prior to MS measurement, intact and IdeS-digested BsAb were desalted by solid-phase extraction (SPE), using a 10 µL ZipTip® (Merck Millipore) filled with 0.6 µL of C4 resin.

First, the SPE-tip was flushed 3 times with 15 μ L of a solution of acetonitrile:water:FA (50%:49.9%:0.1%) and equilibrated by flushing 3 times with 15 μ L of water. Then, 5 μ L of digested mAb was diluted with 5 μ L of water and allowed to bind to the SPE-tip by pipetting 20 times in and out. Finally, the loaded SPE-tip was flushed 3 times with 15 μ L of water and the sample was eluted in 3 μ L of acetonitrile:water:FA (50%:49.9%:0.1%).

4.3.6 MALDI spotting

For MALDI middle-up measurements, 1 μ L of each sample was spotted onto a polished steel MALDI target plate (Bruker Daltonics) and gently mixed with 1 μ L of THAP MALDI matrix. The droplet was allowed to dry at room temperature. For MALDI-ISD measurements, 1 μ L of each sample was spotted onto a polished steel MALDI target plate. Thereafter, 1 μ L of 1,5-DAN matrix was added and mixed gently until the onset of small crystal formation. After allowing the spot to dry at room temperature, the samples were measured immediately.

4.3.7 MALDI mass spectrometry measurements

MALDI-(ISD) FT-ICR measurements were performed as previously reported.⁴² Briefly, measurements were carried out on a Bruker 15T solarix XR FT-ICR mass spectrometer equipped with a CombiSource and a ParaCell (Bruker Daltonics). The MS system was controlled using ftmsControl software (Bruker Daltonics) and operated in MALDI-mode using a Smartbeam-II Laser System (Bruker Daltonics) at a frequency of 500 Hz with 200 laser shots per measurement. High trapping potentials (up to 9.5 V) and ParaCell DC biases (up to 9.3 V) were used to optimize mass measurement precision. MALDI FT-ICR MS spectra were acquired in serial mode (i.e. 10 single spectra per sample) in the m/z -range 3495-30000 using THAP as a MALDI matrix. MALDI-ISD FT-ICR MS spectra were obtained from a single spot using two acquisition methods for the m/z -ranges between m/z 1012-5000 and between m/z 1012-7000, using 1,5-DAN as a MALDI matrix. Each MALDI-ISD FT-ICR MS spectrum was generated from a variable number of summed spectra (i.e. 10 for middle-up analysis and up-to 200 for middle-down analysis). All MS spectra were acquired with 1,000,000 data points and all data files were recorded to include the transient data (fid) file. The measurement time depended on the number of acquired spectra, ranging from approximately 2 min for the acquisitions in the m/z -range 3,495-30,000 to approximately 8 min for the acquisitions in the m/z -range 1,012-5,000. External calibration of MALDI-ISD FT-ICR MS acquisition methods was performed using ISD *c-type* fragments of myoglobin while the acquisition method used for MALDI FT-ICR MS analysis was calibrated using the most intense isotopic peak of each non-glycated A2V BsAb subunit (i.e. Lc1, Lc2, Fd'1, Fd'2 and GOF glycoform for both (Fc/2)1 and (Fc/2)2).

4.3.8 Bottom-up mass spectrometry

Glycated (168 h) A2V BsAb was subjected to 4-12% PAGE (NuPAGE Bis-Tris Precast Gel,

Life Technologies). Bands were cut from the gel, and the proteins subjected to reduction with DTT, alkylation with iodoacetamide and in-gel trypsin digestion using Proteineer DP digestion robot (Bruker).

Tryptic peptides were extracted from the gel slices, lyophilized, dissolved in 95/3/0.1 v/v/v water/acetonitrile/FA and subsequently analyzed by LC-MS using a Thermo Easy nLC 1200 system and a Thermo Orbitrap LUMOS MS system. Trypsin digests were injected onto a homemade precolumn (100 μm $\times$ 15 mm; Reprosil-Pur C18-AQ 3 μm , Dr. Maisch, Ammerbuch, Germany) and eluted via a homemade analytical nano-HPLC column (25 cm $\times$ 75 μm ; Reprosil-Pur C18-AQ 3 μm). The gradient was run from 10% to 40% solvent B (20/80/0.1 water/acetonitrile/FA v/v/v) in 30 min. The nano-HPLC column was drawn to a tip of ~ 15 μm and acted as the electrospray needle of the MS source.

4.3.9 Data-dependent acquisition with HCD

The Orbitrap Lumos FTMS system (Thermo Fisher Scientific, Bremen, Germany) was operated in data-dependent acquisition (DDA, top-N) using HCD as a fragmentation technique. In the master scan (MS1) the resolving power was 120,000 at m/z 200, the m/z -range was 400-1500, and the AGC target was set to 400,000, with a maximum fill time of 50 ms. Dynamic exclusion was set to $n=1$ with exclusion duration of 10 s. Precursor ions with charge states from 2+ to 5+ were isolated using the quadrupole with an isolation window of m/z 1.2. HCD fragment ions were generated using a collision energy of 32 V. MS/MS spectra were acquired in the Orbitrap mass analyzer from m/z 110. The resolution was set to 30,000 at m/z 200 with an AGC target of 50,000 and a maximum fill time of 60 ms.

4.3.10 Data-dependent acquisition with EThcD

LC-MS/MS was also performed using DDA with EThcD (top-N). MS1 parameters were as for DDA analysis with HCD. The calibrated charge-dependent parameters for ion-ion reactions were used with additional supplemental activation at 15%. MS/MS spectra were acquired in the Orbitrap mass analyzer in the m/z -range 200-3000. The resolution was set to 60,000 at m/z 200 with an AGC target of 50,000 and a maximum fill time of 60 ms.

4.3.11 Selected ion monitoring with EThcD

In order to obtain high-quality EThcD spectra of specific glycosylated tryptic peptides, MS/MS analysis was limited to selected precursor ions. MS1 and MS/MS parameters were as described for DDA analysis with EThcD, with some differences. The MS/MS scan resolution was set to 120,000 at m/z 200 with an AGC target of 50,000 and a maximum fill time of 128 ms. Multiple MS/MS spectra of the same precursor ions were averaged to improve the S/N ratio. The selected precursor ions were the following: m/z 1056.5 ($[M+4H]^{4+}$ of glycosylated EVQLVESGGGLVQPGGSLRLSC(Carbamidomethyl)AASGYTFTNYGMNWWVR of Hc2);

m/z 1022.5 ($[M+2H]^{2+}$ of glycosylated EVQLVESGGGLVQPGGSLR of Hc2); m/z 1049.6 and 700.1 ($[M+2H]^{2+}$ and $[M+3H]^{3+}$ of glycosylated pyroQVQLVQSGAEVKKPGASVK of Hc1). The quadrupole was set to cycle through the selected precursor ions throughout the entire run. In a post-analysis process, raw data (reduced profile mass spectra) were converted to peak lists using Proteome Discoverer version 2.2 (Thermo Fisher Scientific) and investigated using Mascot v. 2.2.07 (www.matrixscience.com) and the amino acid sequence of A2V BsAb subunits. Mascot searches were with 10 ppm and 0.02 Da deviation for precursor and fragment mass, respectively, and trypsin as enzyme. Up to two missed cleavages were allowed, and carbamidomethyl on Cys was set as a fixed modification. Methionine oxidation, acetylation (N-terminus and lysine) and glycation (additional 162.0528 Da on N-terminus, lysine, and arginine) were set as variable modifications. The results from this search were compared using Scaffold software version 4.8.9 (www.proteomesoftware.com). A minimum peptide threshold of 80% was used for peptide identification. MS² spectra of modified peptides were also visually inspected.

4.3.12 Spectra processing

MALDI FT-ICR MS spectra were visualized using DataAnalysis 5.0 SR1 (Bruker Daltonics). For each data file, consisting of 10 acquired MALDI FT-ICR MS spectra, an average mass spectrum was obtained after an internal calibration using the most abundant isotopic peak of each non-glycosylated A2V BsAb subunit as calibrants. Theoretical m/z -values were obtained using the online tool MS-Isotope (<http://prospector.ucsf.edu/prospector/mshome.htm>) and theoretical elemental compositions calculated using an in-house developed tool in Excel. Absorption mode MALDI-TOF FT-ICR mass spectra were generated, directly from the transient data files, using AutoVectis software suite (Spectroswiss, Lausanne, Switzerland), as previously reported.⁴² aFT mass spectra were internally calibrated using theoretical *c*-type fragment ions of Hc2 and processed using AutoSquid v6 tool for the identification of *c*- and *z*+1-type fragment ions. Fragment ions were assigned with an error tolerance of 10, 20 ppm for the mass spectra produced between m/z 1012-5000 and m/z 1012-7000, respectively. The tolerance for the quality value for the match between the observed and theoretical isotopic distribution was 0.65. Calibrated aFT mass spectra were exported in the "comma-separated value" (csv) format, copied in a text file and visually inspected in mMass.⁵⁸ To this end, the elemental composition of ISD fragments was calculated using an in-house developed tool in Excel and theoretical isotopic abundance distributions were generated using the online tool envipat Web 2.2 (<http://www.envipat.eawag.ch/index.php>),⁵⁹ exported in the csv format and compiled in a text file. Relative quantification of glycosylated fragment ions was based on the comparison of the intensity of most abundant isotopic peaks of non-glycosylated and glycosylated fragments in aFT spectra. Plots were generated in Excel and exported in Adobe Illustrator CS6 to make the figures.

4.4 Results

4.4.1 Middle-up MS analysis of A2V BsAb subunits and their glycation levels

IdeS digestion followed by chemical reduction of A2V BsAb resulted in the generation of six different subunits (*i.e.* polypeptide chains) varying in size from approximately 22 to 28 kDa (**Figure 1** and **Table S1**). The magnitude mode (mFT) MALDI FT-ICR MS spectrum generated from the analysis of digested and reduced A2V BsAb is shown in Figure 2A. All six different subunits are clearly observed in a single spectrum. It is furthermore noted that all species are detected as doubly charged ions ($[M+2H]^{2+}$) and are isotopically resolved (resolving power 61000 at m/z 13837.63). Ultrahigh resolution allows confident identification of all subunits and is particularly beneficial for the Fc/2 subunits. The sequences of (Fc/2)1 and (Fc/2)2 are similar, with just five amino acid differences at positions 113, 118, 130, 132, 171 (**Table S1**). Moreover, both Fc/2 subunits are N-glycosylated, resulting in detection of multiple glycoforms (not to be confused with glycated forms): G0, G0F, G1F and G2F (**Table S1**).

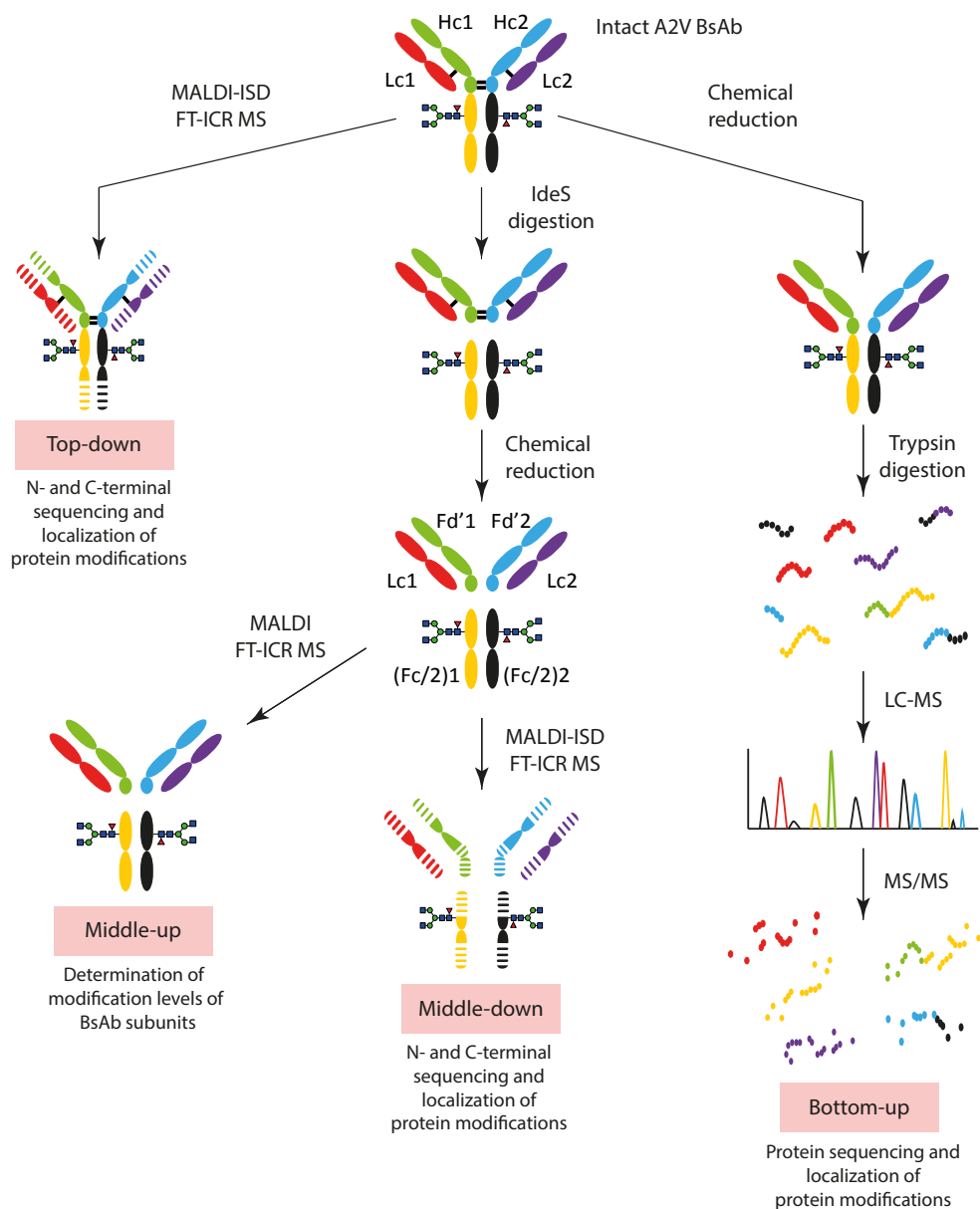


Figure 1. Workflow of analysis followed in this study. A2V BsAb was analyzed by top-down and middle-down MALDI-ISD FT-ICR MS, by middle-up MALDI FT-ICR MS and by bottom-up LC-MS/MS.

In order to evaluate the performance of our method to determine glycation levels in A2V BsAb, forced glycation of the intact BsAb was performed by a prolonged (i.e. up to 168 h) incubation with glucose. In **Figure 2**, panels B-D show mFT MALDI FT-ICR MS spectra of

glycated, IdeS-digested and chemically reduced A2V BsAb respectively. In all six polypeptide chains, glycation increased over time. After 168 h of forced glycation, one or two glucose residues (with m/z increases of 81.03 and 162.06, respectively) were detected on Lc1, Lc2, and Fd'2, while one additional glucose was detected on Fd'1. Glycation levels determined from relative intensities of (Fc/2)1 and (Fc/2)2 ions can be monitored over time, although it is noted that these overlap with some of the Fc/2 N-glycosylated forms, namely G1F and G2F.

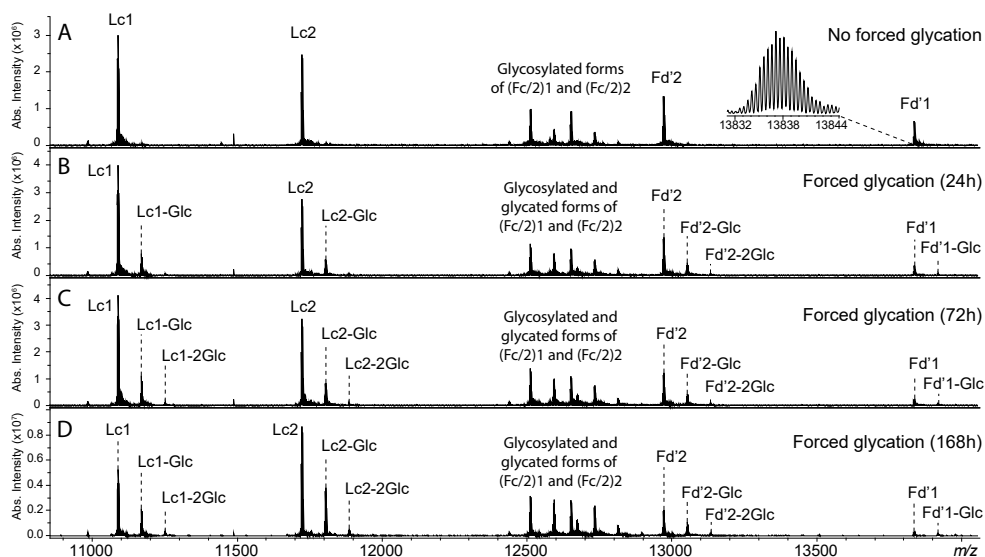


Figure 2. mFT MALDI FT-ICR MS spectra of IdeS-digested and chemically reduced A2V BsAb analyzed (A) prior to and (B-D) after forced glycation. All polypeptide chains, including glycosylated Fc/2 portions, were detected in a single spectrum. Enlargements of such spectra are shown in **Figures S1 to S5**.

For all subunits, it appears that mono-glycation is predominant compared to di-glycation. Enlargements of the spectra depicted in **Figure 2** are provided in **Figure S1-5**, to show the glycation level of each polypeptide chain. Mono-glycation levels were approximately 41%, 45%, 51%, 57%, 28% and 24% for Lc1, Lc2, Fd'1, Fd'2, G0F-(Fc/2)1 and G0F-(Fc/2)2 respectively. The contribution of the glycated G0 glycoform to the signal of the glycated G0F glycoform was not taken into account. The mass measurement error of each glycated species was lower than 10 ppm.

4.4.2 Sequencing of A2V BsAb by top-down and middle-down MS

Intact and IdeS-digested A2V BsAb were analyzed by mFT MALDI-MS/MS using 1,5-diaminonaphthalene (1,5-DAN) as a MALDI matrix. Fragment ions, mainly c-type and z+1-type, were generated from all polypeptide chains. This resulted in complex MS spectra

(Figure S6) that require ultrahigh resolving power for confident assignments. To further enhance spectral resolution these spectra were visualized in absorption mode (aFT; featuring higher spectral resolution) next to the mFT spectra. This resulted in a slightly higher number of assigned fragment ions. For example, from the analysis of IdeS-digested A2V BsAb 466 and 442 fragment ions were identified in aFT spectra and mFT, respectively. Differences between mFT and aFT spectra are exemplified in Figure S7. As described above, Hc1 and Hc2 exhibit 97.6% sequence identity in the Fc regions. Thus all c-type fragments generated from (Fc/2)1 and (Fc/2)2 with m/z lower than 7000 were identical and were assigned to both Hc1 and Hc2 sequences. The total sequence coverage resulting from the interpretation of aFT spectra was 38%, 41%, 32%, 35% for Lc1, Lc2, Hc1, and Hc2, respectively (Figure 3). Note that, the presence of a proline in the sequence prevents the formation of c - and $z+1$ -fragments at this residue. The signal-to-noise ratio (S/N) of identified fragment ions ranged from 5 to 5200 (Figure S8 and Figure S9).

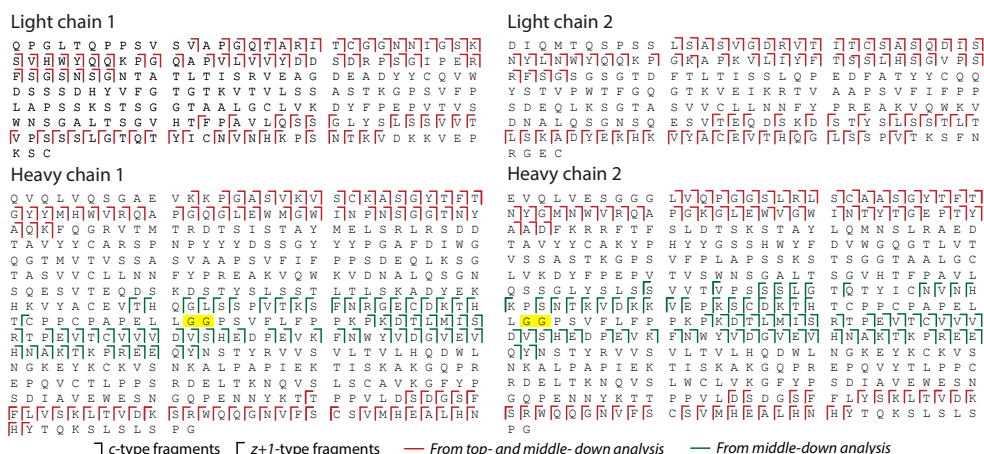


Figure 3. Total sequence coverage for Lc1, Lc2, Hc1, and Hc2. IdeS cleavage site is highlighted in yellow. N-terminal proQ was detected on Lc1 and Hc1.

4.4.3 Localization of hot spot glycation sites by top-down and middle-down MS

Glycated A2V BsAb samples were analyzed by aFT MALDI-MSD FT-ICR MS, as intact species and IdeS-digested subunits in the m/z -range 1000-7000 as described for non-glycated A2V BsAb. The appearance of glycated fragment ions in the mass spectra was investigated, assuming potential glycation of lysine or arginine and considering the heterogeneity of the sample with glycated BsAb subunits differing in the position of the glucose residue. Glycated MALDI-MSD fragments of Hc2 from c11 to c18 (i.e. Leu11 to Leu18) were not detected while, both non-glycated and glycated c19 fragments (i.e. m/z 1881.0194 and m/z 2043.0722, respectively) were detected with glycation increasing overtime during glycation experiments

(Figure 4 and Figure S10), indicating glycation on Arg19.

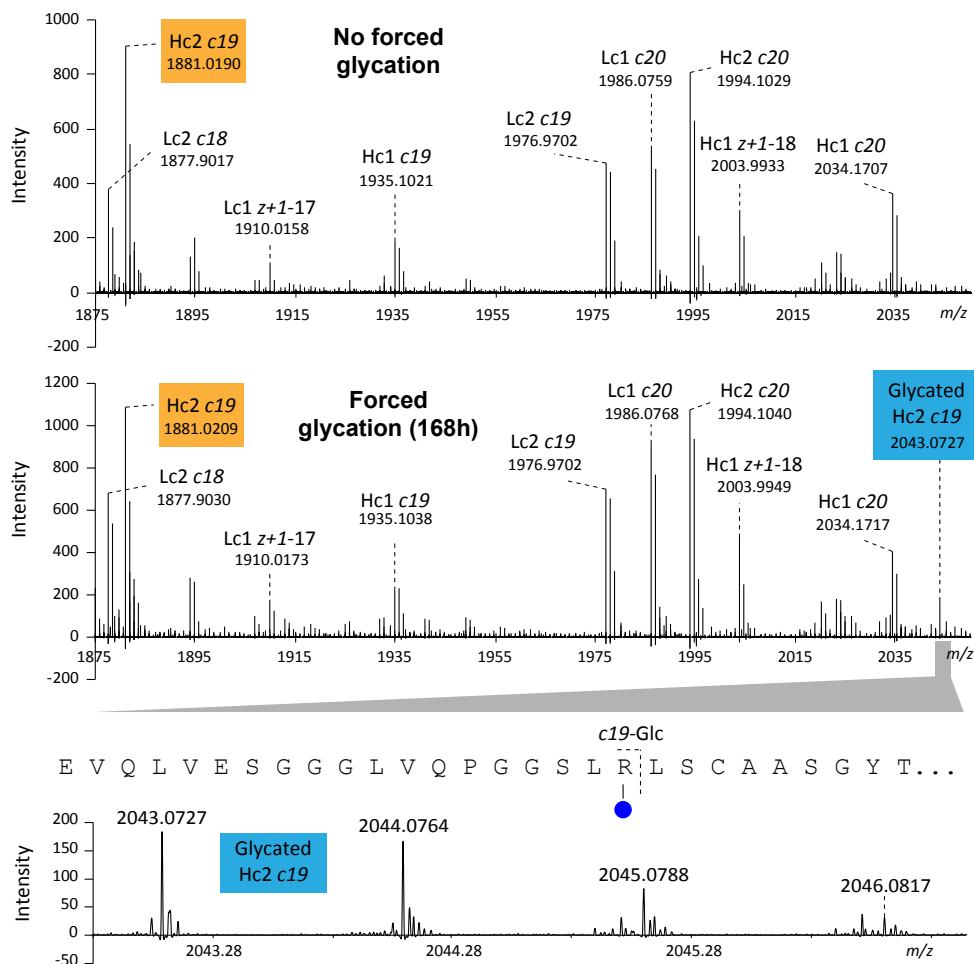


Figure 4. Enlargement of the aFT MALDI-MS/MS spectra obtained from non-glycated (top) and glycated (bottom) intact A2V BsAb. Both non-glycated and glycated c19 fragment ion from Hc2 were detected at m/z 1881.0209 and 2043.0727, respectively, after a 168 h incubation with glucose.

After 168 h, the non-glycated-to-glycated intensity ratio was approximately 8:1 (i.e. intensity c19/intensity glycated c19). Glycated c-type fragments of Hc2 larger than c19 were also detected, however, glycation of Arg19 affected the identification of other glycation sites. For example, glycated c43 fragment, which includes Arg19, Arg38 and Lys43, was detected but the presence of glycation on Arg38 or Lys43 could not be determined since glycation on Arg19 contributed to the signal. Di-glycated c-type fragments of Hc2 were not observed. The N-terminus of Hc1 could not be glycosylated since the primary amine was converted into a pyroglutamate. As a consequence, the first potential glycation site of Hc1, identifiable from

c-type fragments, was Lys12. Indeed, a glycated c12 fragment on Hc1 was detected at m/z 1429.7538 (**Figure 5**). After 168 h, the glycation of Lys12 of Hc1 (non-glycated-to-glycated intensity ratio of approximately 30:1) was reduced compared to glycation of Arg19 of Hc2 (non-glycated-to-glycated intensity ratio of approximately 8:1). The level of glycation on larger c-type fragments was similar and di-glycated c-type fragments of Hc1 were not observed.

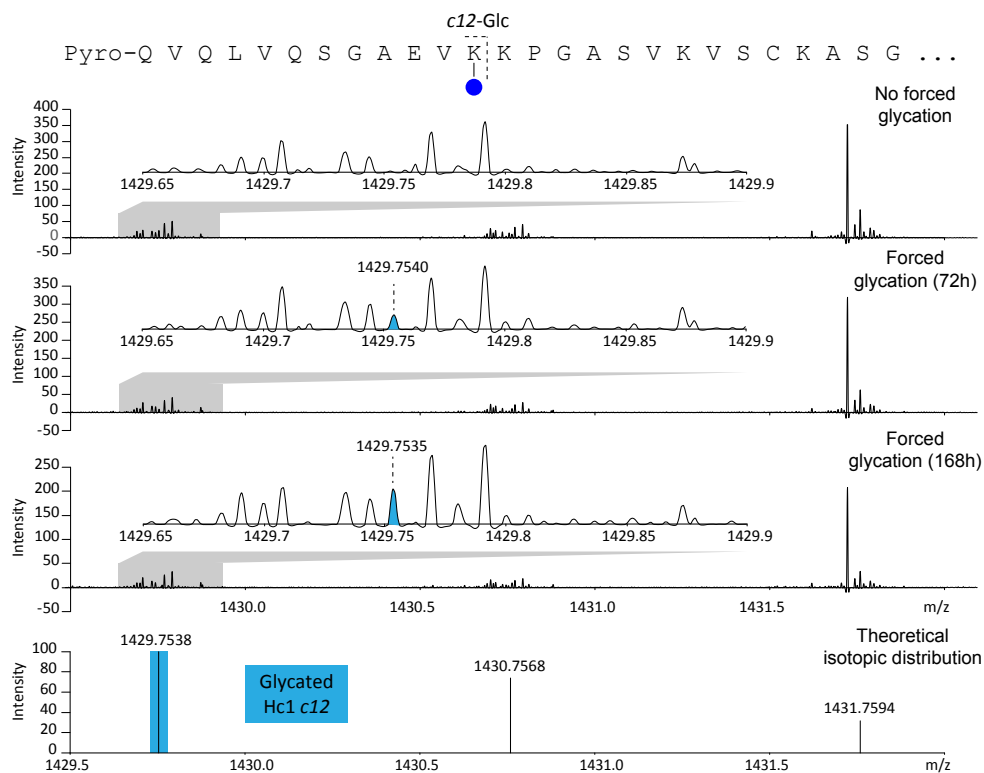


Figure 5. Enlargement of an aFT MALDI-MS/MS spectrum generated from the analysis of non-glycated and glycated intact A2V mAb. Detection of a glycated c12 fragment ion of Hc1 allowed the identification of a glycation site on Lys12.

The analysis of IdeS-digested A2V BsAb revealed a glycation site on (Fc/2)1 and/or (Fc/2)2. In fact, a glycated c11 fragment (i.e. glycation on Lys10; m/z 1346.7360) of (Fc/2)1 and/or (Fc/2)2 was observed (**Figure 6**). The intensity of this fragment ion increased upon forced glycation. After 168 h, the non-glycated c11 fragment (m/z 1184.6841) was approximately 7 times more intense than the glycated c11 fragment. As described for Hc2, glycation of one site affected the determination of glycation level of other sites on larger glycated fragment ions. For example, a glycated c12 fragment was observed but the presence of glycation on Lys12 could not be confirmed since the same c-fragment also contain Lys10.

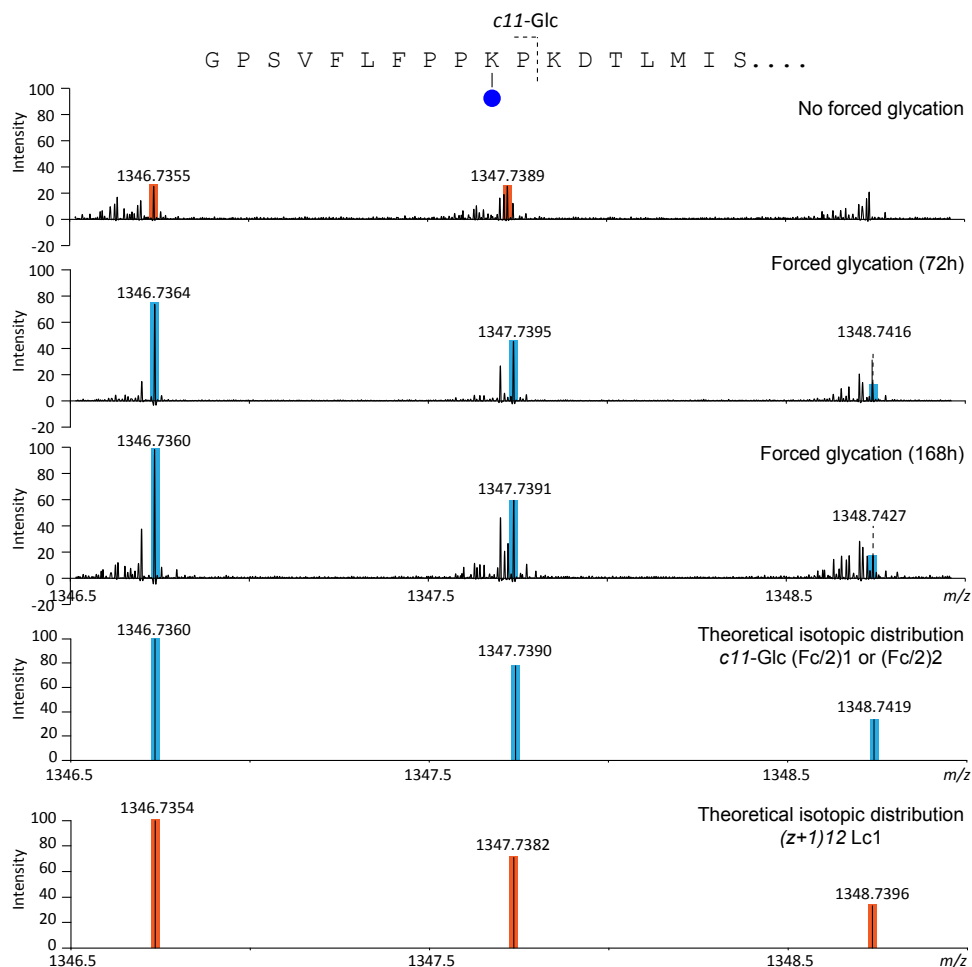


Figure 6. Enlargement of an aFT MALDI-MS/MS spectrum generated from the analysis of non-glycated and glycated IdeS-digested A2V mAb.

Differences in glycation levels between (Fc/2)1 and (Fc/2)2 could not be determined due to a 100% sequence identity of *c*-type fragment ions generated from these subunits in the *m/z*-range investigated. Finally, no signals were detected that corresponded to di-glycated *c*-type fragments. None of the fragment ions generated from Lc1 and Lc2 and detected in the *m/z*-range 1000-7000 contained an additional glucose residue. Similarly, glycated *z*+1-type fragments were not observed for any of the six subunits generated after IdeS digestion and reduction of A2V BsAb. A summary of identified glycation sites by top-down and middle-down MS is reported in **Table 1**.

Table 1. Glycation sites identified by top-down and middle-down MALDI-ISD FT-ICR MS, and standard bottom-up LC-MS/MS.

		Glycated sites
Top-down MS	Hc1	K12
	Hc2	R19
Middle-down MS	(Fc/2)1 and (Fc/2)2	K262 on (Fc/2)1 or K252 on (Fc/2)2
Bottom-up LC-MS/MS with HCD	Lc1	N-terminus Q1; K30
	Lc2	K45; K126; K149; K169; K190
	Hc1	K23; K171; K191; K212; K238; 342; K376
	Hc2	N-terminus E1; K43; K76; K228; K332
Bottom-up LC-MS/MS with ETD	Hc1	K63; K171; K205; K304; K333; K342; K350; K425
	Hc2	K43; K76; K139; K294; K323; K332; K340; K415

4.4.4 Localization of hot spot glycation sites by bottom-up MS

The glycated light and heavy chains of A2V BsAb (from a 168 h glycation experiment) were separated by gel electrophoresis and digested with trypsin, before being analyzed by LC-MS/MS, with higher-energy collisional dissociation (HCD) and in data-dependent acquisition (DDA) mode. Observed peptides were matched to the sequence of the different A2V subunits and results from this analysis are reported in Table S2. Sequence coverage from bottom-up analysis was 66%, 79%, 87% and 83% for Lc1, Lc2, Hc1 and Hc2, respectively (**Table S2** and **Figure S11**). The glycation sites identified are reported in **Table 1** and **S3**. As an example, the HCD spectrum from the glycated tryptic peptide EVQLVESGGGLVQPGGSLR of Hc2 is reported in **Figure S12**, revealing glycation of the N-terminus.

In addition, the tryptic digests were analyzed by LC-MS/MS with ETD in DDA mode and results are summarized in **Table S2**. The sequence coverage was 85% and 82% for Hc1 and Hc2, respectively (**Figure S11**). The glycation sites identified are reported in **Table 1** and **S3**. As an example, the ETD spectrum from the glycated tryptic peptide VVSVLTVLHQDWLNGKEYK of Hc1 is reported in **Figure S12**. The tryptic digest of heavy chains of A2V BsAb was also analyzed by LC-MS/MS with EThcD on selected precursor ions. The EThcD spectra obtained from the glycated tryptic peptide pyroQVQLVQSGAEVKKPGASVK of Hc1 and EVQLVESGGGLVQPGGSLR of Hc2 are reported in **Figure S13**.

4.5 Discussion

4.5.1 Simultaneous analysis of A2V BsAb subunits by middle-up MALDI FT-ICR MS

Measurement of the intact mass of proteins provides direct information on protein modifications.²⁸ For complex proteins, such as BsAbs, the determination of the mass of each protein subunit (*e.g.* Lc and Hc) is important to determine the structural integrity of the molecule. The cleavage of BsAbs into Fab and Fc regions, using specific enzymes, further increases the number of polypeptide chains in the sample. In fact, six different subunits can be obtained from a BsAb by combining specific hinge-region proteolytic cleavage with chemical reduction of disulfide bonds of Fab and Fc regions. In this study, we aimed to develop a fast MS-based method for the simultaneous analysis of all six subunits obtained from A2V BsAb. To this end, we used MALDI and ultrahigh resolution FT-ICR MS to measure two different Lc, two different Fd, and two different glycosylated Fc/2 regions, in a single spectrum. Previously, we have reported that MALDI FT-ICR MS analysis of mAbs after IdeS digestion and without chemical reduction allows the analysis of glycosylated Fc/2, solely. Intact Fab portions of the analyzed mAbs were about 100 kDa and therefore were not detectable under the used conditions.^{33, 42} In the current study, we reduced the disulfide bonds and we used THAP as a MALDI matrix to increase the intensity of doubly charged ions which were detected in the m/z -range 11000-14000. Theoretically, in the sample, all six polypeptide chains were present at the same concentration. Consequently, the observed differences in signal intensities may be attributed to differences in ionization efficiency and ion transmission. The use of a 15 T MALDI FT-ICR MS system allowed us to perform measurements at isotopic resolution, with high sensitivity at high m/z . In addition, compared to ESI-MS, the detection of the different BsAb subunits as doubly charged ion species by MALDI FT-ICR MS allowed us to increase the mass difference between species and avoid overlapping signals. Furthermore, MALDI is more tolerant to the presence of salts or other additives in the sample than ESI. These characteristics allowed middle-up analysis of BsAbs subunits by MALDI FT-ICR MS after minimal sample preparation and without the need for chromatographic separation.

4.5.2 Determination of glycation level of A2V BsAb subunits by middle-up MALDI FT-ICR MS

The MALDI FT-ICR MS method developed was then applied for the analysis of glycated A2V BsAb. In our study, we were able to monitor the glycation level of all polypeptide chains obtained from A2V BsAb in a forced glycation experiment, performed up to 168 h. Monoglycation was predominant for all chains and di-glycation was likewise observed for all chains except for Fd1. However, the detection of di-glycated Fd1 was probably confounded by the lower signal intensity for this chain, therefore Fd1 di-glycation cannot be excluded.

Glycation of glycosylated (Fc/2)1 and (Fc/2)2 resulted in isobaric glycoforms which affected the interpretation of the glycation levels. This level was, in fact, based on the most abundant glycoform (i.e. G0F), solely.

Glycation of A2V BsAb at the intact level was not investigated. Firstly, with the current technology, the analysis of intact BsAb (~150 kDa) ions by MALDI FT-ICR MS is not possible. Secondly, standard MS-based methods for the analysis of the complex mixture of isomeric structures which result from the glycation of an intact (and glycosylated) BsAb are not available. In fact, techniques such as boronate affinity chromatography, capillary isoelectric focusing and liquid chromatography, commonly used for the analysis of glycated mAbs at intact level, only allow the separation of non-glycated from glycated structures, without isomeric resolution.^{9, 11, 45, 46} Thus, although intact mass analysis provides information on the number of sugar molecules attached to a BsAb, the contribution of the different isomeric structures (i.e. combinations of glycated BsAb subunits) to the glycation level cannot be determined.^{10, 45}

In addition, the presence of glycosylated structures affects the quantification of glycation levels. This limitation can be overcome by removal of the glycans prior to MS analysis (e.g. using Peptide:N-glycosidase F (PNGase F)).^{47, 48} Similarly, our middle-up approach also overcomes this limitation for the non-glycosylated BsAb subunits, namely the light chains and the Fd' subunits.

4.5.3 Structural characterization of A2V BsAb by top-down and middle-down MALDI-ISD FT-ICR MS

In this study, we also evaluated the performance of our recently developed MALDI-ISD FT-ICR MS method for the characterization of mAbs, with regard to the analysis of BsAbs. As shown previously, the use of 1,5-DAN as a MALDI matrix predominantly provided N-terminal c-type fragment ions while C-terminal z+1-type ions were detected at a lower intensity. The hydrogen donor property of 1,5-DAN causes the reduction of disulfide bonds during ionization, allowing the analysis of intact and digested A2V BsAb without chemical reduction.⁴⁹

Fragment ions were measured in the m/z-range 1000-7000. At m/z lower than 1000, the presence of matrix cluster ions affected the sensitivity for ISD fragment ions while at m/z higher than 7000, assignment of fragment ions was compromised by a lower resolving power and a lower sensitivity. Thus, mass measurements in this m/z-range allowed us to characterize regions of amino acid sequences close to both the N- and C-terminus. Internal regions of Hc1 and Hc2 could be only characterized after IdeS digestion.

In a previous study, we showed that our method provides complementary amino acid sequence information, when compared with other fragmentation techniques such as ETD

and CID.⁴² This is due to the fact that a more continuous array of fragment ions is observed in the m/z -range analyzed compared to other fragmentation techniques. Using our approach, ISD fragment ions from all six different subunits were detected in one single spectrum, with (Fc/2)1 and (Fc/2)2 having a 100% identity of *c*-type fragment ions in the analyzed m/z -range.

Ultrahigh resolution measurements and aFT visualization simplified the interpretation of MALDI-ISD FT-ICR MS spectra, allowing us to manage their complexity and increasing the confidence in the identification. However, although resolving power in aFT spectra was approximately double that in the comparable mFT spectra, the number of assigned ISD fragments increased only slightly. This result was associated with the already high quality of mFT spectra where the number of overlapping ion species was minor compared to the total number of observed fragment ions.

ESI-based MS methods allow the use of a variety of fragmentation techniques, which in general provide complementary sequence information. These techniques rely on the isolation of a precursor ion prior to MS/MS analysis. Typical isolation windows for top-down MS analysis of proteins are larger than 3 m/z units. Thus, the isolation of a single ion species from a dense MS spectrum is not trivial and is usually difficult to achieve in automatic precursor ion selection mode. The overlapping of multiply charged fragment ions from more than one ion species complicates MS/MS spectra, affecting sequence coverage and identification scores. This problem can be tackled by online protein separation techniques (e.g. LC and CE) and specific gas-phase ion/ion reactions.^{50, 51}

In our approach, precursor ion selection is avoided but the complexity of fragmentation spectra is reduced by both the generation of singly charged fragment ions (which ensure a maximal m/z difference between ion species) and the use of ultrahigh resolution. Improving the quality of MALDI-ISD FT-ICR MS spectra acquired at m/z higher than 7000 would lead to a higher sequence coverage. Thus, further research focused on novel acquisition strategies (e.g. using new data acquisition systems) and MALDI-ISD conditions (e.g. different ISD matrices) is warranted.^{52, 53}

4.5.4 Localization of hot spot glycation sites by middle-down MALDI-ISD FT-ICR MS

MALDI-ISD FT-ICR MS was used to monitor glycation levels of A2V BsAb by relative quantification, based on the comparison between the intensity of non-glycated and glycated fragment ions. Furthermore, MALDI-ISD pinpointed hot spots for glycation close to both N- and C-terminal portions of the BsAb. Hot spot glycation sites were identified on Lys12 of Hc1, Arg19 of Hc2, and Lys10 of (Fc/2)1 and/or (Fc/2)2 (or Lys262 of Hc1 and Lys252 of Hc2), from the observation of glycated *c*-type fragments. Standard bottom-up MS analysis of trypsin-digested light and heavy chains of A2V mAbs complemented these results.

N-terminal glycation of Hc2 was identified in a bottom-up analysis, while MALDI-ISD determined glycation of Arg19 in Hc2. In the bottom-up approach, each identified glycosylated lysine was found in a tryptic peptide that contained a missed cleavage at the glycosylated lysine. Such a missed cleavage at glycosylated Arg19 would yield EVQLVESGGGLVQPGGSLR(Glc) LSCAASGYTFTNYGMNWVR with a carbamidomethylated cysteine. Bottom-up analysis is not optimal for large peptides in general and indeed this peptide was not detected in our MS scans.

With regard to glycation on Lys43 of Hc2, results differed between bottom-up and MALDI-ISD. The bottom-up approach determined a glycosylated Lys43. The MALDI-ISD method identified a glycosylated c43 fragment from Hc2, however, this was not sufficient to pinpoint to glycation on Lys43 due to the presence of (an additional) glycation on Arg19, coming before Lys-43 in the sequence. Hence, a limitation of MALDI-ISD MS analysis is that in the case of multiple glycosylations within a fragment ion it may be hard to pinpoint all the glycosylations to specific sites. This limitation is common to other top-down MS methods. In fact, although the more commonly used ESI-based top-down approaches typically provide enough sequence information for identification, localization, and quantification of a few positional isomers of modified proteins⁵⁴⁻⁵⁶, the analysis of complex mixtures of isomers is more difficult.⁵⁷

The sensitivity of bottom-up MS analysis was improved by selected ion monitoring. For example, glycation on Lys12 of Hc1 was detected by MALDI-ISD MS but not by bottom-up MS analysis in DDA mode. However, the glycosylated peptide pyroQVQLVQSGAEVKKGASVK of Hc1 was identified by visual inspection of the LC-MS/MS data. Analysis of such a peptide by ETD using selected ion monitoring allowed us to improve the quality of the MS/MS spectra and to localize the glycation site on Lys12 with high confidence. It is furthermore noted that the application of alternative proteases such as endoproteinase GluC or chymotrypsin potentially increases the number of identified glycation sites.

In conclusion, in this study, a fast method for monitoring of A2V BsAb modifications at the subunit level, based on ultrahigh resolution MALDI FT-ICR MS, was presented. This method allowed the analysis of glycosylated A2V BsAb, in a forced glycation experiment, without the need for chromatographic separation of BsAb subunits providing complementary information compared to standard bottom-up MS analysis. In addition, we have shown that our recently developed MALDI-ISD FT-ICR MS method, developed for the structural characterization of mAbs, can be also applied for the analysis of even more complex biopharmaceuticals such as A2V BsAb. MALDI-ISD localized hot spot glycation sites proximal to N-termini of A2V subunits and provided structure information complementary to standard bottom-up MS analysis with HCD and ETD.

4.6 Supporting information

Supplementary information is available free of charge on the mAbs website: DOI: 10.1080/19420862.2019.1682403

4.7 Acknowledgments

This work was supported by the Analytics for Biologics project (Grant agreement ID 765502) of the European Commission and by an Investment Grant NWO Medium with project number 91116004 to P.A.v.V. which is (partially) financed by ZonMw.

Notes:

The authors declare the following competing financial interest(s): Dr. Tsybin is an employee of Spectroswiss, which commercializes AutoVectis software.

4.8 References

1. Brinkmann U., Kontermann R. E. The making of bispecific antibodies. *MAbs* 2017, 9, 182-212.
2. Sedykh S.E., *et al.* Bispecific antibodies: design, therapy, perspectives. *Drug Des Dev Ther* 2018, 12, 195-208.
3. Krishnamurthy A., Jimeno A. Bispecific antibodies for cancer therapy: A review. *Pharmacol Ther* 2018, 185, 122-34.
4. Liu H., May K. Disulfide bond structures of IgG molecules. *mAbs* 2012, 4, 17-23.
5. Zheng K., Bantog C., Bayer R. The impact of glycosylation on monoclonal antibody conformation and stability. *MAbs* 2011, 3, 568-76.
6. Higel F., *et al.* N-glycosylation heterogeneity and the influence on structure, function and pharmacokinetics of monoclonal antibodies and Fc fusion proteins. *Eur J Pharm Biopharm* 2016, 100, 94-100.
7. Liu H., *et al.* Heterogeneity of monoclonal antibodies. *J Pharm Sci* 2008, 97, 2426-47.
8. Kennedy D. M., Skillen A. W., Self C. H. Glycation of monoclonal antibodies impairs their ability to bind antigen. *Clin Exp Immunol* 1994, 98, 245-51.
9. Wei B., *et al.* Glycation of antibodies: Modification, methods and potential effects on biological functions. *mAbs* 2017, 9, 586-94.
10. Saleem R. A., *et al.* A chemical and computational approach to comprehensive glycation characterization on antibodies. *MAbs* 2015, 7, 719-31.
11. Quan C., *et al.* A study in glycation of a therapeutic recombinant humanized monoclonal antibody: where it is, how it got there, and how it affects charge-based behavior. *Anal Biochem* 2008, 373, 179-91.
12. Brady L. J., Martinez T., Balland A. Characterization of nonenzymatic glycation on a

- monoclonal antibody. *Anal Chem* 2007, 79, 9403-13.
13. Yuk I. H., *et al.* Controlling glycation of recombinant antibody in fed-batch cell cultures. *Biotechnol Bioeng* 2011, 108, 2600-10.
 14. Zhang Z., Pan H., Chen X. Mass spectrometry for structural characterization of therapeutic antibodies. *Mass Spectrom Rev* 2009, 28, 147-76.
 15. Beck A., Sanglier-Cianfèrani S., Van Dorsselaer A. Biosimilar, Biobetter, and Next Generation Antibody Characterization by Mass Spectrometry. *Anal Chem* 2012, 84, 4637-46.
 16. Tran B. Q., *et al.* Glycosylation characterization of therapeutic mAbs by top- and middle-down mass spectrometry. *Data in brief* 2016, 6, 68-76.
 17. Miller A. K., *et al.* Characterization of site-specific glycation during process development of a human therapeutic monoclonal antibody. *J Pharm Sci* 2011, 100, 2543-50.
 18. Zhang J., *et al.* Rapid identification of low level glycation sites in recombinant antibodies by isotopic labeling with ¹³C6-reducing sugars. *Anal Chem* 2012, 84, 2313-20.
 19. Chelius D., *et al.* Structural and functional characterization of the trifunctional antibody catumaxomab. *MAbs* 2010, 2, 309-19.
 20. Zhang Y., *et al.* Protein Analysis by Shotgun/Bottom-up Proteomics. *Chem Rev* 2013, 113, 2343-94.
 21. Wang L., *et al.* Structural Characterization of a Recombinant Monoclonal Antibody by Electrospray Time-of-Flight Mass Spectrometry. *Pharm Res* 2005, 22, 1338-49.
 22. Srzentić K., *et al.* Advantages of Extended Bottom-Up Proteomics Using Sap9 for Analysis of Monoclonal Antibodies. *Anal Chem* 2014, 86, 9945-53.
 23. Stadlmann J., *et al.* Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides. *Proteomics* 2008, 8, 2858-71.
 24. Wang Y., *et al.* Characterization and Comparison of Disulfide Linkages and Scrambling Patterns in Therapeutic Monoclonal Antibodies: Using LC-MS with Electron Transfer Dissociation. *Anal Chem* 2011, 83, 3133-40.
 25. Marx V. Making sure PTMs are not lost after translation. *Nat Methods* 2013, 10, 201.
 26. Venne A. S., Kollipara L., Zahedi R. P. The next level of complexity: Crosstalk of posttranslational modifications. *Proteomics* 2014, 14, 513-24.
 27. Ren D., *et al.* An improved trypsin digestion method minimizes digestion-induced modifications on proteins. *Anal Biochem* 2009, 392, 12-21.
 28. Tran J. C., *et al.* Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature* 2011, 480, 254-8.
 29. Tvardovskiy A., *et al.* Top-down and Middle-down Protein Analysis Reveals that Intact and Clipped Human Histones Differ in Post-translational Modification Patterns. *Mol Cell Proteomics* 2015, 14, 3142-53.
 30. Lermyte F., *et al.* Top or Middle? Up or Down? Toward a Standard Lexicon for Protein Top-Down and Allied Mass Spectrometry Approaches. *J Am Soc Mass Spectrom* 2019.
 31. Donnelly DP, Rawlins CM, DeHart CJ, Fornelli L, Schachner LF, Lin Z, *et al.* Best practices

- and benchmarks for intact protein analysis for top-down mass spectrometry. *Nat Methods* 2019, 16, 587-94.
32. Vincents B., *et al.* Cleavage of IgG1 and IgG3 by gingipain K from *Porphyromonas gingivalis* may compromise host defense in progressive periodontitis. *FASEB J* 2011, 25, 3741-50.
 33. Sjögren J., Olsson F., Beck A. Rapid and improved characterization of therapeutic antibodies and antibody related products using IdeS digestion and subunit analysis. *Analyst* 2016, 141, 3114-25.
 34. D'Atri V., *et al.* Hydrophilic Interaction Chromatography Hyphenated with Mass Spectrometry: A Powerful Analytical Tool for the Comparison of Originator and Biosimilar Therapeutic Monoclonal Antibodies at the Middle-up Level of Analysis. *Anal Chem* 2017, 89, 2086-92.
 35. Sorensen M., *et al.* Comparison of originator and biosimilar therapeutic monoclonal antibodies using comprehensive two-dimensional liquid chromatography coupled with time-of-flight mass spectrometry. *mAbs* 2016, 8, 1224-34.
 36. Resemann A., *et al.* Full validation of therapeutic antibody sequences by middle-up mass measurements and middle-down protein sequencing. *mAbs* 2016, 8, 318-30.
 37. Ayoub D., *et al.* 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* 2013, 5, 699-710.
 38. Srzentic K., *et al.* Multiplexed Middle-Down Mass Spectrometry as a Method for Revealing Light and Heavy Chain Connectivity in a Monoclonal Antibody. *Anal Chem* 2018, 90, 12527-35.
 39. Haselberg R., *et al.* Heterogeneity assessment of antibody-derived therapeutics at the intact and middle-up level by low-flow sheathless capillary electrophoresis-mass spectrometry. *Anal Chim Acta* 2018, 1044, 181-90.
 40. Leblanc Y., *et al.* Charge variants characterization of a monoclonal antibody by ion exchange chromatography coupled on-line to native mass spectrometry: Case study after a long-term storage at +5 degrees C. *Journal of chromatography B, Analytical technologies in the biomedical and life sciences* 2017, 1048, 130-9.
 41. Faid V., *et al.* Middle-up analysis of monoclonal antibodies after combined IgdE and IdeS hinge proteolysis: Investigation of free sulfhydryls. *Journal of Pharmaceutical and Biomedical Analysis* 2018, 149, 541-6.
 42. van der Burgt Y. E. M., *et al.* Structural Analysis of Monoclonal Antibodies by Ultrahigh Resolution MALDI In-Source Decay FT-ICR Mass Spectrometry. *Anal Chem* 2019, 91, 2079-85.
 43. Kilgour D. P. A., *et al.* Autophaser: An Algorithm for Automated Generation of Absorption Mode Spectra for FT-ICR MS. *Anal Chem* 2013, 85, 3903-11.
 44. Tyshchuk O., *et al.* Characterization and prediction of positional 4-hydroxyproline and sulfotyrosine, two post-translational modifications that can occur at substantial levels

- in CHO cells-expressed biotherapeutics. *mAbs* 2019, 11, 1219-32.
45. Mo J., *et al.* Quantitative analysis of glycation and its impact on antigen binding. *MAbs* 2018, 10, 406-15.
 46. Fischer S., Hoernschemeyer J., Mahler H. C. Glycation during storage and administration of monoclonal antibody formulations. *Eur J Pharm Biopharm* 2008, 70, 42-50.
 47. Goetze A. M., *et al.* Rates and impact of human antibody glycation in vivo. *Glycobiology* 2012, 22, 221-34.
 48. Zhang B., *et al.* Unveiling a glycation hot spot in a recombinant humanized monoclonal antibody. *Anal Chem* 2008, 80, 2379-90.
 49. Fukuyama Y., Iwamoto S., Tanaka K. Rapid sequencing and disulfide mapping of peptides containing disulfide bonds by using 1,5-diaminonaphthalene as a reductive matrix. *J Mass Spectrom* 2006, 41, 191-201.
 50. Huang T. Y., McLuckey S. A. Top-down protein characterization facilitated by ion/ion reactions on a quadrupole/time of flight platform. *Proteomics* 2010, 10, 3577-88.
 51. Bashyal A., *et al.* Top-Down Analysis of Proteins in Low Charge States. *J Am Soc Mass Spectrom* 2019, 30, 704-17.
 52. Kooijman P. C., *et al.* Increased throughput and ultra-high mass resolution in DESI FT-ICR MS imaging through new-generation external data acquisition system and advanced data processing approaches. *Sci Rep* 2019, 9, 8.
 53. Fukuyama Y., Izumi S., Tanaka K. 3-Hydroxy-2-Nitrobenzoic Acid as a MALDI Matrix for In-Source Decay and Evaluation of the Isomers. *J Am Soc Mass Spectrom* 2018, 29, 2227-36.
 54. Zhang H., Ge Y. Comprehensive analysis of protein modifications by top-down mass spectrometry. *Circ Cardiovasc Genet* 2011, 4, 711.
 55. Lin Z., *et al.* Simultaneous Quantification of Protein Expression and Modifications by Top-down Targeted Proteomics: A Case of the Sarcomeric Subproteome. *Mol Cell Proteomics* 2019, 18, 594-605.
 56. Pesavento J. J., Mizzen C. A., Kelleher N. L. Quantitative Analysis of Modified Proteins and Their Positional Isomers by Tandem Mass Spectrometry: Human Histone H4. *Anal Chem* 2006, 78, 4271-80.
 57. Woodling K. A., *et al.* Identification of single and double sites of phosphorylation by ECD FT-ICR/MS in peptides related to the phosphorylation site domain of the myristoylated alanine-rich c kinase protein. *J Am Soc Mass Spectrom* 2007, 18, 2137-45.
 58. Strohm M., *et al.* mMass data miner: an open source alternative for mass spectrometric data analysis. *Rapid Commun Mass Spectrom* 2008, 22, 905-8.
 59. Loos M., *et al.* Accelerated Isotope Fine Structure Calculation Using Pruned Transition Trees. *Anal Chem* 2015, 87, 5738-44.





Chapter 5

Characterization and prediction of positional 4-hydroxyproline and sulfotyrosine, two post-translational modifications that can occur at substantial levels in CHO cells-expressed biotherapeutics

Oksana Tyshchuk^{1#}, Christoph Gstöttner^{2,#}, Dennis Funk¹, Simone Nicolardi², Stefan Frost¹, Stefan Klostermann³, Tim Becker⁴, Elena Jolkver⁴, Felix Schumacher¹, Claudia Ferrara Koller⁵, Hans Rainer Völger¹, Manfred Wuhrer², Patrick Bulau⁶, and Michael Mølhøj¹

[#]These authors contributed equally to this work

¹Roche Pharma Research and Early Development, Large Molecule Research, Roche Innovation Center Munich, Germany

²Leiden University Medical Center, Center for Proteomics and Metabolomics, Leiden, The Netherlands

³Roche Pharma Research and Early Development Informatics, Roche Innovation Center Munich, Germany

⁴xValue GmbH, Willich, Germany

⁵Roche Pharma Research and Early Development, Large Molecule Research, Roche Innovation Center Zurich, Switzerland

⁶Roche Pharma Technical Development Penzberg, Germany

Reprinted and adapted with permission from mAbs, 2019, 11, 1219-1232, DOI: 10.1080/19420862.2019.1635865

© 2019 The Authors. Published with license by Taylor & Francis Group, LLC.



5.1 Abstract

Biotherapeutics may contain a multitude of different PTMs that need to be assessed and possibly monitored and controlled to ensure reproducible product quality. During early development of biotherapeutics, unexpected PTMs might be prevented by in silico identification and characterization together with further molecular engineering. Mass determinations of a hIgG1 (mAb1) and a bispecific IgG-ligand fusion protein (BsAbA) demonstrated the presence of unusual PTMs resulting in major +80 Da, and +16/+32 Da chain variants, respectively. For mAb1, analytical cation exchange chromatography demonstrated the presence of an acidic peak accounting for 20%. A +79.957 Da modification was localized within the light chain CDR-2 and identified as a sulfation based on accurate mass, isotopic distribution, and a complete neutral loss reaction upon CID. Finally, top-down ultrahigh resolution MALDI-ISD FT-ICR MS of modified and unmodified Fabs allowed the allocation of the sulfation to a specific Tyr residue. An aspartate in amino-terminal position -3 relative to the affected Tyr was found to play a key role in determining the sulfation. For BsAbA, a +15.995 Da modification was observed and localized to three specific Pro residues explaining the +16 Da chain A, and +16 Da and +32 Da chain B variants. The BsAbA modifications were verified as 4-hydroxyproline and not 3-hydroxyproline in a tryptic peptide map via co-chromatography with synthetic peptides containing the two isomeric forms. Finally, our approach for an alert system based on in house in silico predictors is presented which aims at preventing these PTMs by molecular design and engineering during early biotherapeutic development.

5.2 Introduction

The development of biotherapeutics requires an extensive characterization of the molecular purity, homogeneity, integrity, assessment of potential critical quality attributes (pCQAs), including the analysis of heterogeneities from post-translational modifications. PTMs occur at distinct amino acid side chains or peptide linkages, and well known chemical antibody modifications include deamidation, isomerisation, oxidation, reduction, glycation, cysteinylolation, and trisulfides. Common heterogeneities due to enzyme-catalyzed PTMs of therapeutic antibodies include Fc domain N-glycosylation, heavy chain C-terminal lysine cleavage and proline amidation, as well as incomplete removal of signal peptides.¹ Some modifications are added shortly after translation is completed, others occur after folding during passage through the Golgi apparatus, upon bioprocessing, storage, stress (e.g. elevated temperatures, oxidation, or light) or after administration.² Enzyme-catalyzed PTMs account for a huge number of different modifications, and involve enzymes like kinases, phosphatases, hydroxylases and transferases, which add or remove functional groups, proteins, peptides, lipids or oligosaccharides to amino acid side chains. Occasionally, unexpected enzyme-catalyzed PTMs have been reported in constant domains or frame-work regions of IgGs like hydroxylysine, glutamine-linked glycosylation, and non-consensus Asn-glycosylation.³⁻⁵ Also the complementarity-determining regions (CDRs) with their inherent variability and solvent exposure have been identified with unusual enzyme-catalyzed PTMs like O-fucosylation and sulfotyrosine.^{6,7} Further uncommon enzyme-catalyzed PTMs (based on the experience and knowledge from IgGs and mAbs) have been reported in non-IgG domains and linkers of antibody-fusion proteins.⁸⁻¹²

In therapeutic proteins PTMs are especially critical if they negatively influence drug potency or drug safety. Even if PTMs just increase the product heterogeneity of the desired molecules, when identified they need to be either diminished or removed by manufacturing process optimizations or require further protein engineering, or at least have to be monitored and controlled to demonstrate batch consistency or comparability of manufactured clinical material. While N-glycosylation sites are usually predictable due to the consensus sequence NXS/T, N-glycosylation of the antigen-binding (Fab) domain is present in the CDRs of some therapeutic antibodies and increase the product heterogeneity further.¹³ The same is true for cysteinylations in CDRs.¹⁴ Although PTMs observed in natural IgGs might pose a lower safety risk for monoclonal antibodies, failure to detect low level PTMs of <1% abundance may represent a safety risk due to potential direct or indirect effects on immunogenicity.^{2,15} In these cases, PTMs are considered relevant pCQAs. Another aspect is the risk to not fulfil predefined release criteria due to medium level PTMs of >1% abundance influencing e.g. the charge heterogeneity, which is widely analyzed during batch release and/or for the direct comparison of product batches to ensure drug quality and bio-process consistency.

Technically, the main challenges to study PTMs are the development of specific detection,

characterization and purification methods. Fortunately, these technical obstacles are being overcome with a variety of reliable characterization methods. E.g. suitable and non-inducing tryptic digestion methods for their analytical detection and characterization have been reported for the assessment of deamidations and succinimide formation^{16, 17}, trisulfide quantitation¹⁸, disulfide bond analysis¹⁹, or the availability of electron-transfer/higher-energy collision dissociation (ET_hcD) to give more complete fragmentation and more reliable phosphorylation site localization compared to electron transfer dissociation (ETD) or higher-energy collision dissociation (HCD) alone.²⁰ Recently, matrix-assisted laser desorption ionization (MALDI) was combined with in-source decay (ISD) fragmentation and ultrahigh resolution FT-ICR MS for a fast characterization of mAbs providing complementary sequence information compared to other MS-based techniques.²¹ Other approaches have been focusing on the development of structure based prediction tools for the identification, e.g. of deamidation, isomerisation, and oxidation hot spots in monoclonal antibodies²²⁻²⁴ or other in silico PTM prediction tools.²⁵⁻²⁷

Herein, we describe the integration of innovative PTM analysis and in silico identification tool for the detection and verification of 4-hydroxyproline (4Hyp) and sulfotyrosine (sTyr) in antibody-based therapeutics, two PTMs that are technically challenging to verify and with the potential to be present at substantial levels in biotherapeutics produced in Chinese hamster ovary (CHO) cell. 4Hyp is difficult to distinguish from its isomeric form 3-hydroxyproline (3Hyp), and sTyr is challenging to localize, as the modified amino acid is very labile using common MS/MS technologies. Here, we demonstrate the identification of 4Hyp based on chromatographic separation of peptides containing the hydroxyproline isomers, and the use of top-down ultrahigh resolution MALDI-ISD FT-ICR MS to allocate the sulfation. We present our strategy how to identify and monitor these uncommon PTMs during early development of biotherapeutics.

5.3 Experimental Section

5.3.1 Enzymes, peptides and proteins

Endoproteases and alkaline phosphatase were purchased from Promega. Synthetic peptides (95-98% HPLC-purity) were synthesized by Biosyntan GmbH. PNGase F was obtained from Custom Biotech, Roche Diagnostics GmbH. BsAbA was transiently expressed in CHO cells. mAb1 was stably expressed in CHO cells and purified from 250 L platform fed batch fermentations.

5.3.2 Intact and reduced mass analysis by ESI-QTOF-MS

100 µg BsAbA or mAb1 were Fc-deglycosylated by adding 45 U PNGase F, and 100 mM sodium phosphate buffer to a final volume of 230 µL, followed by incubation at 37 °C for 16

h. Reduction was done by adding 115 μL 100 mM TCEP in 4 M guanidine hydrochloride to 115 μL (50 μg) Fc-deglycosylated protein followed by incubation at 37 °C for 30 min. The samples were desalted by HPLC on a self-packed Sephadex G25 (Amersham Biosciences) 5 $\times$ 250 mm column at room temperature using an 8 min isocratic gradient with 40% acetonitrile with 2% formic acid (v/v) at 1 mL/min. By monitoring the UV absorption at 280 nm the protein peaks were collected using a fraction collector. Approx. 50 μg protein was injected. Total masses were determined by ESI-QTOF-MS on a maXis 4G QTOF mass spectrometer (Bruker Daltonik) equipped with a TriVersa NanoMate source (Advion). Calibration was performed with sodium iodide (Tof G2-Sample Kit 2; Waters). For the deglycosylated and deglycosylated/reduced proteins, data acquisition was performed at m/z 1000–4000 (ISCID: 130.0 eV), and m/z 600–2000 (ISCID: 0.0 eV), respectively. Data acquisition with Fabs was done at m/z 900–3000 (ISCID: 0.0 eV). The raw mass spectra were evaluated and transformed into individual relative molar masses using an in-house developed software tool. The quantitative evaluation of the mass spectra was performed by summing up contributions of m/z ion intensities of all charge states forming the dominant part (larger than 20%) of the charge state envelope as observed for the most abundant individual product mass. Then all peak contributions (fitted as Gaussians) of all signals in these charge states were used to calculate the relative contents of the individual species.

5.3.3 Analytical cation exchange chromatography of mAb1

Cation exchange chromatography was performed on a 4 $\times$ 250 mm ProPac WCX-10 column (Thermo Fisher Scientific) using 20 mM MES, pH 6.0 (eluent A) and 20 mM MES, 500 mM NaCl, pH 6.0 (eluent B), 1 mL/min flow rate, 25°C column temperature, and the following gradient: 75% eluent A [0–5 min], 25% to 56% eluent B [5–57 min], 56% to 100% eluent B [57–58 min], 100% eluent B [58–63 min], 0% to 75% eluent A [63–64 min], 75% eluent A [64–70 min]. 50 μg of mAb1 was injected and detected by 280 nm absorbance.

5.3.4 Isolation of modified and unmodified mAb1 Fab fragments

160 mg of mAb1 were diluted to 1 mg/mL in 0.2 $\times$ PBS, 150 mM NH_4Ac , pH 7.0 and incubated 10 min at 37°C. 3.3 mg Lys-C was added and the digest incubated at 37°C for 1.5 h. Following, the cleaved Fabs were isolated by a combination of two connected affinity chromatography columns under standard conditions. The mixture of Lys-C cleaved antibody, Fc-domain and Fab fragments was passed through both chromatography columns where the remaining non-cleaved antibody plus the cleaved Fc-domain was retained on MabSelect™ SuRe™ (GE Healthcare) whereas all cleaved Fabs were captured with CaptureSelect™ IgG-CH1 (Thermo Fisher Scientific). After acidic elution of the Fabs with 150 mM acetic acid, the pH value was adjusted with 1 M Tris to pH 5.5 and a size exclusion polishing step with Superdex™ 200 (GE Healthcare) in 20 mM His/HisHCl, 140 mM NaCl at pH 6.0 applied. To separate the Fab

mixture containing acidic and main peak Fabs, the sample was concentrated to >40 mg/mL with centrifugal spin columns and buffer exchanged to 10 mM sodium phosphate, pH 7.0. Multiple runs using a CIEX ProPac™ WCX-10 analytical column (Thermo Fisher Scientific) with 10 mM sodium phosphate, pH 7.0 (eluent A) and 10 mM sodium phosphate, 500 mM NaCl, pH 7.0 (eluent B) and a gradient of 0-38% eluent B within 9.5 column volumes at a flow rate of 1 mL/min led to a base line separation of both Fab species. With this method 50 µg purified Lys-C cleaved Fabs were injected multiple times and detected by 280 nm absorbance. Separated peaks were pooled, concentrated, aliquoted and stored at -80°C. The concentrations of the Fabs were determined by UV280 nm and using the molar extinction coefficients calculated on the basis of the amino acid sequences. The purified Fabs were reanalyzed by analytic CIEX and ESI-QTOF-MS.

5.3.5 LC-MS/MS peptide mapping

The antibodies were denatured and reduced in 0.3 M Tris-HCl pH 8, 6 M guanidine-HCl and 20 mM dithiothreitol (DTT) at 37° C for 1 hour, and alkylated by adding 40 mM iodoacetic acid (C13: 99%) (Sigma-Aldrich) at room temperature in the dark for 15 min. Excess iodoacetic acid was inactivated by adding DTT to 40 mM. The alkylated fusion protein was buffer exchanged using NAP5 gel filtration columns and a proteolytic digestion with trypsin was performed in 50 mM Tris-HCl, pH 7.5 at 37° C for 16 hours. The reaction was stopped by adding formic acid to 0.4% (v/v). Thermolysin digests were performed as described by Tyshchuk et al.¹¹ Digested samples were stored at -80°C and analyzed by UPLC-MS/MS using a nanoAcquity UPLC (Waters) and an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific). About 2.5 µg digested fusion protein was injected in 5 µL. Chromatographic separation was performed by reversed-phase on a BEH300 C18 column (1x 150 mm, 1.7 µm) or a CSH130 C18 column, 1x 150 mm, 1.7 µm (Waters) using mobile phase A and B containing 0.1% (v/v) formic acid in UPLC grade water and acetonitrile, respectively, 60 µL/min flow rate, 50°C column temperature, and the following gradient: 1% mobile phase B [0-3 min], 1% to 40% mobile phase B [3-93 min], 40% to 99% mobile phase B [93-94 min], 99% mobile phase B [94-96 min], 99% to 1% mobile phase B [96-97 min], and 1% mobile phase B [97-105 min]. Two injections of mobile phase A were performed between sample injections using a similar 50 min gradient up to 99% mobile phase B to prevent carry-over between samples. Synthetic peptides were spiked into to digests at different levels.

High-resolution MS spectra were acquired with the Orbitrap mass analyzer, and detection of CID, HCD, or ETD MS/MS fragment ion spectra in the ion trap with dynamic exclusion enabled (repeat count of 1, exclusion duration of 15 s (± 10 ppm)). The Orbitrap Fusion Lumos was used in the data-dependent mode. CID essential MS settings were: full MS (AGC: 2×10^5 , resolution: 1.2×10^5 , m/z range: 300-2000, maximum injection time: 100 ms), MS/MS (AGC: 5.0×10^3 , maximum injection time: 100 ms, isolation window (m/z): 2). HCD essential MS settings were: full MS (AGC: 4×10^5 , resolution: 1.2×10^5 , m/z range: 300-

2000, maximum injection time: 50 ms), MS/MS OT (AGC: 5.0×10^5 , maximum injection time: 500 ms, isolation window: 2), Orbitrap resolution was 15×10^3 , MS/MS IT (AGC: 1.0×10^4 , maximum injection time: 100 ms, isolation window: 2). ETD essential MS settings were: full MS (AGC: 2.0×10^5 , resolution: 6.0×10^5 , m/z range: 300-2000, maximum injection time: 100 ms). MS/MS IT (reaction time: 60-80 ms, reagent target: 1.0×10^6 , maximum reagent injection time: 200 ms, AGC: 1.0×10^5 , maximum injection time: 150 ms, isolation window: 2, supplemental activation was used without and with 10-35% collision energy). Orbitrap resolution was 15×10^3 or 30×10^3 . MS/MS IT (AGC: 5.0×10^4 , maximum injection time: 100-150 ms, isolation window: 2). Normalized collision energy was set to CID: 35% (activation q: 0.25); HCD: 28%; ETD: 15-35%.

A complementary ETHcD method based on HCD and ETD as data dependent fragmentation techniques involved full scan MS acquired with the Orbitrap mass analyzer, and parallel detection of ETD and HCD fragment ion spectra in the ion trap and Orbitrap mass analyzer, respectively. A fixed cycle time was set for the full scan with as many as possible data dependent MS/MS scans. Full MS: same setting as for HCD. For HCD, the MS/MS was detected in the ion trap and the setting were as follows: MS/MS (AGC: 1.0×10^4 , maximum injection time: 100 ms, isolation window: 2, charge state 1-3). Normalized collision energy: 28%. ETD: reaction time was set to 60 ms, 1×10^5 reagent target, reagent injection time: 200 ms, charge state 3-8. Supplemental activation collision energy was set to 30%. The AGC target was set to 5.0×10^4 , Orbitrap resolution was 1.5×10^4 , the precursor isolation window was 2 and the maximum injection time was set to 500 ms.

High-resolution MS spectra were acquired with the Orbitrap mass analyzer in negative mode, and parallel detection of CID or HCD MS/MS fragment ion spectra in the ion trap and Orbitrap with dynamic exclusion enabled as targeted mass experiment. For negative ion mode analysis mobile phases containing 5-10 mM ammonium acetate, 0.1% (v/v) triethylamine, or 5 mM ammonium formate in UPLC grade water and acetonitrile were used. CID essential MS settings were: full MS (AGC: 2×10^5 , resolution was used 6.0×10^5 , 1.2×10^5 , 2.4×10^5 , m/z range: 300-2000, maximum injection time: 100 ms), MS/MS IT (AGC: 1.35×10^5 , maximum injection time: 100-200ms, isolation window: 2). Normalized collision energy was set to 5-45%, activation q: 0.25, isolation window: 2. HCD essential MS settings were: full MS (AGC: 4×10^5 , resolution: 1.2×10^5 , m/z range: 300-2000, maximum injection time: 50 ms), MS/MS OT (AGC: 1.0×10^5 , maximum injection time: 500 ms, isolation window: 2). Normalized collision energy was set to 10-35% additional with stepped collision energy. Orbitrap resolution was 15×10^3 . MS/MS IT (AGC: 1.0×10^4 , maximum injection time: 100 ms, isolation window: 2). Normalized collision energy was set to 10-35%.

5.3.6 MS/MS data evaluation

Analysis of the Orbitrap MS/MS data and the PTM identification was performed using the PEAKS studio 8.0, 8.5 and X software (Bioinformatics Solutions Inc.). Manual data

interpretation and quantification was performed using the Xcalibur Qual Browser v.4.0 (Thermo Fisher Scientific). The Protein Calculator (Thermo Fisher Scientific) was used to calculate theoretical masses, and the XICs were generated with the most intense isotope mass using a mass tolerance of 5 ppm.

5.3.7 Top-down MALDI-ISD FT-ICR mass spectrometry of isolated mAb1 Fabs

MALDI-ISD FT-ICR MS experiments were performed on a 15 T solarix XR system equipped with a CombiSource and a ParaCell (Bruker Daltonics). MS measurements were performed as previously reported with some modification.²¹ Briefly, 1 μ L of each sample was spotted onto a ground steel MALDI target plate together with 1 μ L of 1,5-diaminonaphthalene (saturated in 50% ACN, 49.9% H₂O, 0.1% formic acid). For each spot, an average spectrum was obtained from the acquisition of 200 spectra in the m/z-range 2023-30,000 with 512 k data points. Absorption mode spectra were obtained using AutoVectis software suite (Spectroswiss) and evaluated using mMass data miner.⁵⁰

5.3.8 Development and testing of sTyr and 4Hyp prediction tools

In silico prediction was attempted with two machine learning approaches, the k-nearest neighbor algorithm (kNN) and the random forest (RF) algorithm. For that purpose, public data from Swiss-Prot was used as training reference with non-mammalian and completely redundant sequences removed. Residues for which PTMs were annotated were treated as PTM-positive, residues in the same records with no annotation or explicit exclusion of PTM (rarely reported) were treated as PTM-negative. The dataset used was trimmed to 21 amino acid long peptides, with the target amino acid (P or Y, respectively) located at position 11 with the adjacent N- and C-terminal 10 amino acid positions. The dataset for sTyr contained 6304 peptides (out of which 588 contained sTyr at position 11), the dataset for 4Hyp contained 4055 peptides (with 299 4Hyp examples). Features, used for modelling included the peptide sequences and in the case of sTyr also physicochemical and biochemical properties of various selected amino acids as provided at <ftp://ftp.genome.jp/pub/db/community/aaindex/aaindex1>.⁵¹ In addition, we created for both PTM-types an undersampled dataset with an equal modified to unmodified sequence ratio. For training, 75% of peptides were used (with the ratio of modified to unmodified sequences left constant), and the remaining 25% were used for testing. After assessment of model quality as given in **Table SI**, the models were used for predicting the 4Hyp and sTyr in BsAbA and mAb1, respectively, which were previously unseen by the prediction tools.

5.4 Results

5.4.1 mAb1 and BsAbA contain major mass variants

The molecular masses of mAb1 (a monoclonal human IgG1) and BsAbA (a bispecific IgG-fusion protein) expressed in CHO cells were analyzed by electrospray ionization quadrupole time-of-flight mass spectrometry (ESI-QTOF-MS). In addition to the expected mass of mAb1 (142735 Da), an unknown +80 Da variant (142815 Da) was present at 21% in the deconvoluted mass spectrum of the N-deglycosylated antibody (**Figure 1A**). Upon further reduction, the +80 Da modification could be allocated to the light chain (Lc) at 10% (**Figure 1B**). Assuming a binominal distribution, 10% modified Lc at reduced level would result in 18% of mAb1 having one +80 Da modification and 1% with both Lc's carrying the +80 Da modification at intact antibody level (**Figure S1**). The 21% +80 Da variant for the intact N-deglycosylated antibody (**Figure 1A**) indicates the modifications on the two Lc's to be independent.

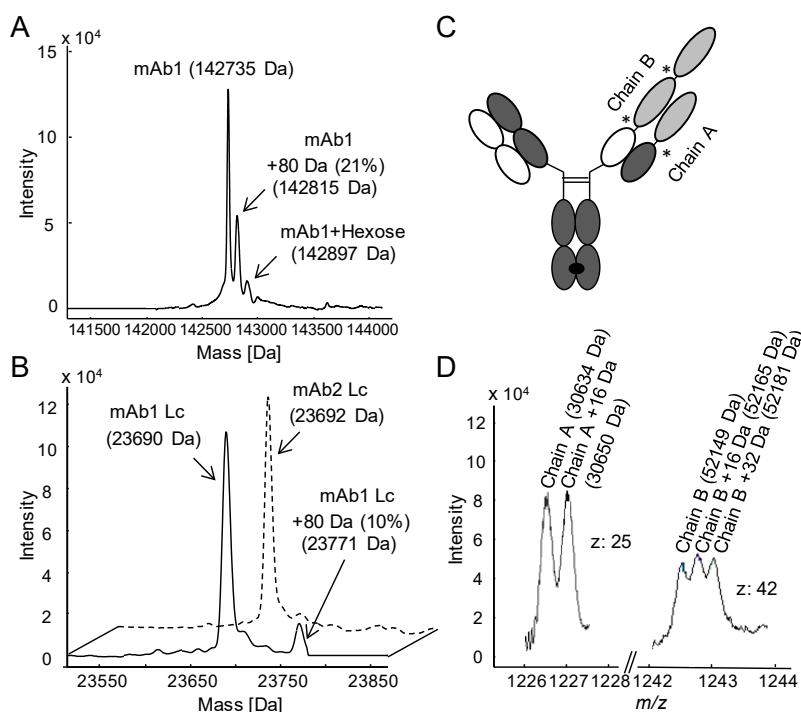


Figure 1. Intact and reduced mass analysis of human IgG1s mAb1 and mAb2, and bispecific IgG-fusion protein BsAbA expressed in Chinese hamster ovary cells. Deconvoluted mass spectra of (a) the N-deglycosylated mAb1, and (b) the reduced light chain (Lc) of mAb1 (solid line) and mAb1-related antibody mAb2 (dotted line). (c) Schematic illustration of BsAbA consisting of four different chains and based on the CrossMabCH1-CL format including three 4-1BB ligand domains (light grey) allocated to chain A and B. * denotes glycine-serine linkers. (d) Mass spectrum of the N-deglycosylated and reduced BsAbA (chain A, charge state z: 25; chain B, charge state z: 42).

BsAbA consist of four different chains and is based on the CrossMab^{CH1-CL} format^{28,29} including three 4-1BB ligand domains allocated to two molecular chains A and B (**Figure 1C**) comparable to the molecules reported by Claus et al.³⁰ In the mass spectrum of the N-deglycosylated and reduced BsAbA, two chains with the expected mass of 30634 Da (chain A) and 52149 Da (chain B) were present as well as an unknown +16 Da variant of chain A (30650 Da), and +16 Da (52165 Da) and +32 Da (52181 Da) variants of chain B (**Figure 1D**). The +16 Da variant of chain A was quantified to around 50% relative abundance, and the +16 Da and +32 Da variants of chain B were quantified to each around 33%.

5.4.2 The +80 Da modification makes mAb1 acidic

We speculated, if the +80 Da mAb1 variant could be explained by a phosphorylation or a sulfation (both +80 Da), an amino acid sequence variant, or a more complex combination of various events. Analyzed by cation exchange chromatography (CIEX), mAb1 demonstrated the presence of a significant acidic variant peak which was quantified to 20% by peak integration (**Figure 2A**) matching the 21% +80 Da variant of the intact mAb1 detected by ESI-MS (**Figure 1A**). To determine whether the +80 Da modification of the mAb1 Lc was in fact causing the presence of the acidic peak, the antibody was cleaved into antigen-binding (Fab) domains by a limited endoprotease Lys-C digestion and purified by a combination of MabSelect SuRe affinity chromatography to remove the Fc domain and CH1 affinity chromatography. A main Fab peak and a second more acidic Fab peak were separated and isolated by CIEX (**Figure 2B**). Subsequently, the isolated fragments were re-analyzed by CIEX (**Figure 2C**) and ESI-MS (**Figure 2D**). The mass analysis of the isolated Fabs confirmed that only the acidic eluting Fab contained the +80 Da modification. The Fabs were also isolated to obtain close to 100% non-sulfated and sulfated portions to determine if there was any functional impact of the modification, and to allow a differential analysis by MS. A surface plasmon resonance-based target binding assay did not demonstrate any functional impact of the modification (data not shown).

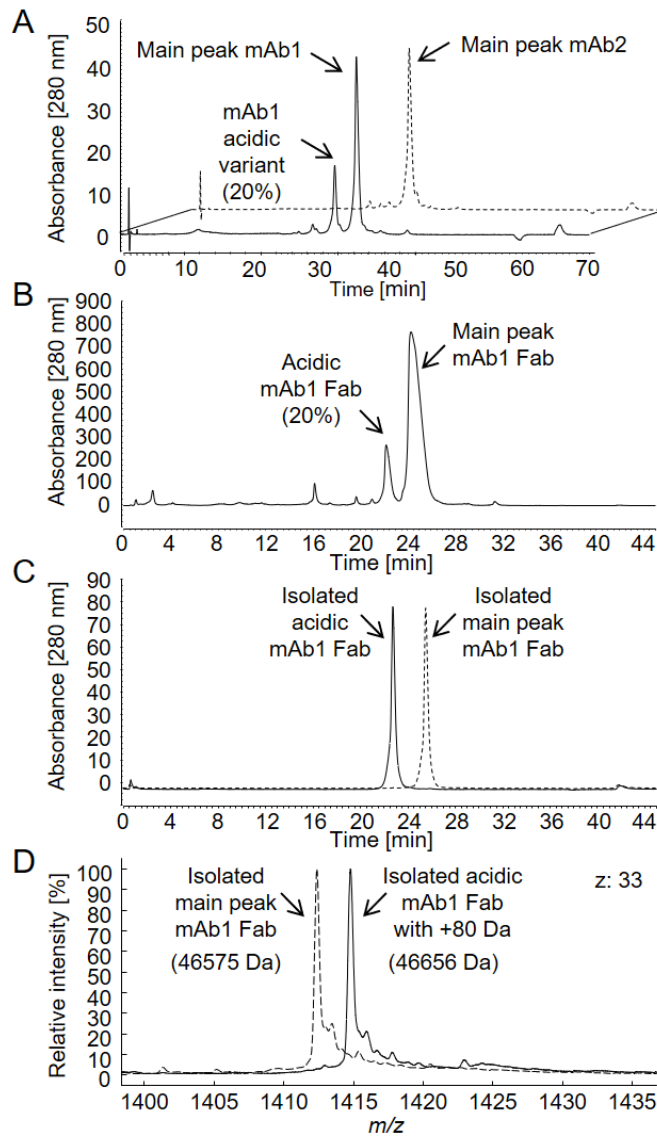


Figure 2. Cation exchange chromatography of (a) the intact mAb1 (solid line) and mAb2 (dotted line), (b) CH1 affinity purified antigen-binding (Fab) domains succeeding a limited endoproteinase Lys-C digest of mAb1 and removal of fragments with an Fc domain by MabSelect SuRe affinity chromatography, and (c) isolated acidic and main peak mAb1 Fabs. (d) Mass spectra (charge state $z: 33$) of the isolated Fabs with (acidic Fab, solid lines; 46656 Da) and without (main peak Fab, dotted lines; 46575 Da) the +80 Da modification.

5.5 LC-MS/MS peptide mapping of mAb1

In addition, mAb1 was analyzed by LC-MS/MS peptide mapping to assess and interpret the +80 Da variant. Evaluation of the LC-MS data from a chromatographic separation of a tryptic digest on a reverse-phase C18 column proved the presence of a mAb1 peptide (LIYSASDL DYGVPSR) with a mass corresponding to a +79.957 Da modification. On a C18 column, the modified peptide co-eluted with the unmodified peptide (data not shown), and a relative quantification by integration of the EIC chromatograms indicated only 4.6% modified peptide, which did not match the 10% relative abundance determined by ESI-QTOF-MS for the N-deglycosylated/reduced mAb1 (**Figure 1B**). We speculate whether co-elution on the reverse-phase C18 column caused reduced ionization efficiency (ion suppression) of the modified peptide. Notably, when analyzed using a mixed-mode hybrid reverse-phase C18 column carrying a low level surface charge, the modified and unmodified tryptic peptides were chromatographically separated with the modified peptide (41.3 min versus 38.3 min) being more retained on the charged surface hybrid column (**Figure 3A**). A relative quantitative evaluation (charge states z : 2 and 3) of the unmodified and the modified peptide determined the modification to be present at 9.4% relative abundance which is in accordance with the 10% modified Lc as determined by ESI-MS (**Figure 1B**).

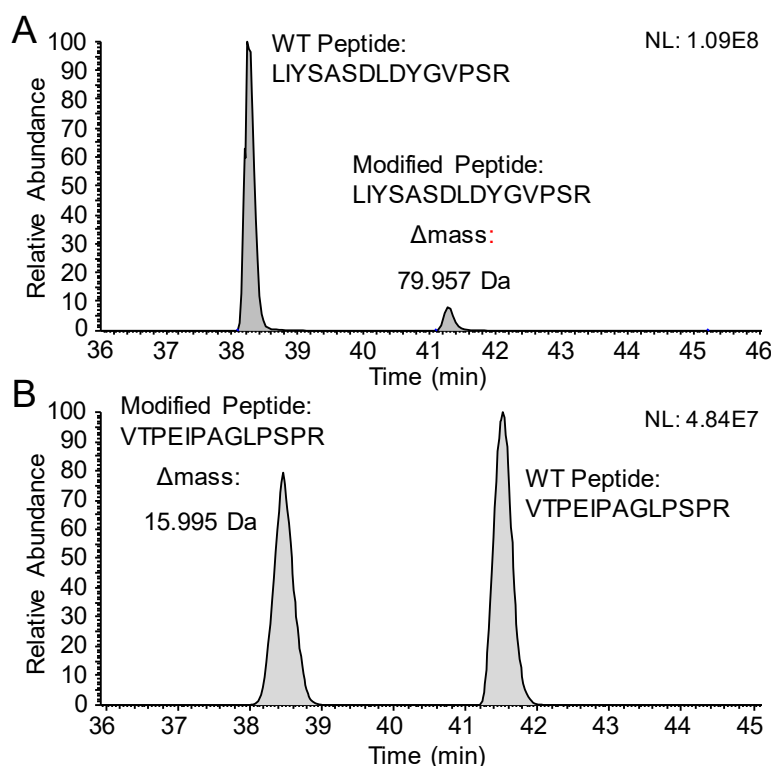


Figure 3. (a) Extracted ion current (EIC) chromatogram of the unmodified (elution time 38.3 min, z :

2 and 3 and m/z window 559.6096–552.6184 + 828.4097–828.4229; MA: 1121769959 (90.6%)) and the +79.957 Da modified (elution time 41.3 min, z: 2 and 3 and m/z window 579.2621–579.2713 + 868.3879–868.4017; MA: 116324914 (9.4%)) tryptic peptides (LIYSASDL DYGVPSR) of mAb1 separated on a mixed-mode hybrid CSH130 C18 column. (b) EIC chromatogram of the +15.995 Da modified (elution time 38.4 min, charge states z: 2 and 3 and m/z = 450.5772–450.5932 + 675.3647–675.3807; manually integrated peak area (MA): 728231181 (46%)) and the unmodified (elution time 41.5 min, z: 2 and 3 and m/z = 445.2459–445.2619 + 667.3666–677.3826; MA: 869604154 (54%)) tryptic peptides (VTPEIPAGLPSPR) of BsAbA separated on a BEH300 C18 reverse-phase column. A relative comparison of the integrated EIC chromatograms quantified the +79.957 Da and +15.995 Da modifications to 9.4% and 46%, respectively. NL, normalized intensity level.

Detailed evaluation of the CID-MS/MS data of the identified modified and unmodified mAb1 tryptic peptides demonstrated close to identical fragment ion profiles for both peptides featuring complete neutral loss of the +79.957 Da modification (data not shown). However, based on the accurate mass, the +79.957 Da modification suggested the presence of a sulfation (SO_3 , monoisotopic mass: +79.9568 Da) rather than a phosphorylation (HPO_3 , monoisotopic mass: +79.9663 Da). Also the isotopic distribution with relative abundance of the 2nd, 3rd, and 4th isotopic peak of 91.4%, 50.0% and 20.4% peak area relative to the monoisotopic peak was found to be in agreement with sulfation (theoretical distribution: 2nd, 91.197%; 3rd, 51.295%; and 4th isotopes, 21.463%) rather than phosphorylation (theoretical distribution: 2nd, 90.422%; 3rd, 46.164%; and 4th isotopes, 17.100%). A treatment with alkaline phosphatase as described in Tyshchuk et al.¹¹ did not remove the +79.957 Da modification from the tryptic peptide (data not shown). In addition, it is known that the sulfoester bond is labile when fragmented by CID causing a complete neutral loss of 80 Da from the precursor^{31,32} which is in line with the above-mentioned experimental results. Without having the position identified, all these observations together with the acidic elution of the modified Fab (**Figure 2**) pointed towards sulfation. A sulfation adds an acidic group to either a tyrosine, a serine or a threonine residue, with sTyr being most commonly observed.³³⁻³⁵ Of note, the modified tryptic peptide contained two tyrosine and three serine residues (LIYSASDL DYGVPSR) as potential sulfation sites, and further experiments were performed to identify the modified residue including several alternative digests. However, no single or combined set of digests allowed us to unambiguously determine the site of the mAb1 sulfation.

Various fragmentations techniques including ETD, HCD, ETHCD in positive ion mode and CID and HCD in negative ion mode were tested but did not provide diagnostic sulfated fragment ions that would allow identification of the sulfation site (data not shown). Also, spiking experiments with synthetic peptides containing sTyr (sY) in the positions 3 (LIYSASDL DYGVPSR) and 10 (LIYSASDL DYGVPSR) and LC-MS using the charged surface hybrid column did not exclude any of the tyrosine residues as the site of sulfation. Both synthetic peptides co-eluted with the modified tryptic peptide (data not shown).



5.5.1 Identification of the sulfation site of mAb1 by MALDI-ISD FT-ICR mass spectrometry

To unambiguously determine the site of the mAb1 sulfation, several fragmentation techniques were tested including MALDI-PSD-TOF MS in negative ion mode with different matrices as well as derivatization of the carboxyl-groups close to the sulfation by dimethylamidation or ethyl esterification (hypothesis: protons provided by carboxyl groups close to sulfotyrosine might enhance the sulfate loss), and ESI-FT-ICR-ETD MS/MS in positive ion mode of the synthetic peptides. However, all methods suffered from complete neutral loss of the SO₃ unit and/or low signal intensity. The site of sulfation was finally identified by the characterization of the non-sulfated and sulfated intact Fab fragments of mAb1 (**Figure 2C** and **D**) by ultrahigh resolution MALDI-ISD coupled to FT-ICR MS. The interpretation of MALDI-ISD FT-ICR MS spectra was based on the assignment of c-type and z+1-type fragment ions, solely. These fragment types were the most abundant ions in the spectra and since are generated via a radical-mediated mechanism, as for ETD, were considered the most valuable ions for the characterization of a labile PTM such as sulfation. The results from this analysis are summarized in **Figure 4**. While the non-sulfated c50 fragment ion was detected in the spectra of both non-sulfated and sulfated Fab at m/z 5544.907, the sulfated c50 fragment ion at m/z 5624.864, was not detected (**Figure 4A-C**). Similarly, sulfation was also not detected on Ser51 and Ser53 (LIYSA₅₁DLDYGVPSR) (**Figure S2**). The non-sulfated c57 fragment ion was detected in the spectrum of the non-sulfated Fab at m/z 6296.210 (**Figure 4D**). However, the sulfated c57 fragment ion of m/z 6376.167 was clearly verified in the ISD spectrum of the sulfated Fab (**Figure 4E** and **Figure S3**). In combination, these data allowed the localization of sulfation to Tyr57 (LIYSASDL₅₇LDYGVPSR). In addition, the intense c-type fragments c65*-c67* (covering the framework region C-terminal to the CDR2, asterisks denoting sulfated fragment ions) of the sulfated Fab light chain were in line with sulfation at Tyr57 (**Figure S4**).

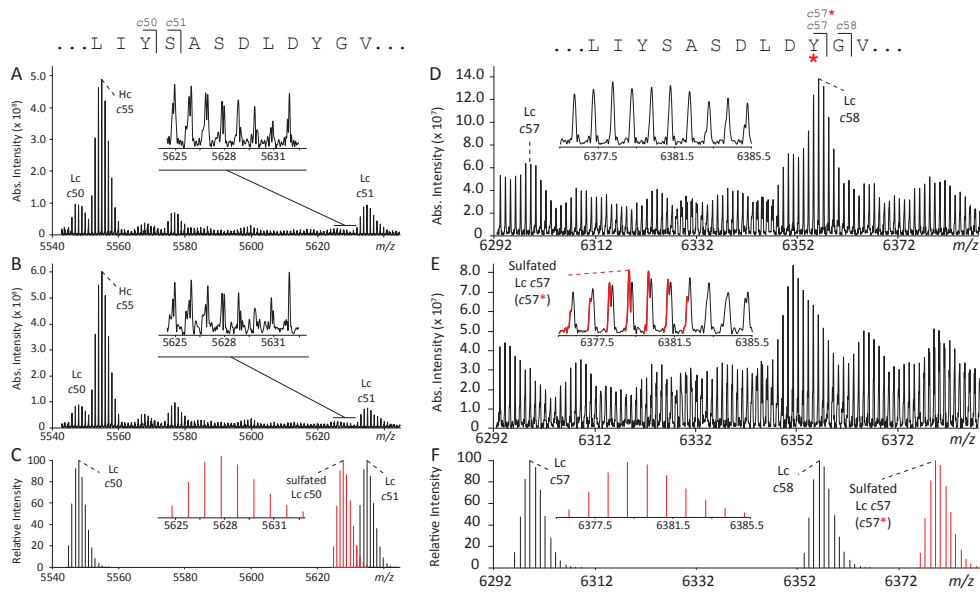


Figure 4. Identification of the sulfation site by top-down MALDI-MS/MS of antigen-binding (Fab) domains. A comparison was made between the fragmentation spectra generated from the mAb1 Fab without (a and d) and with (b and e) sulfation. In the m/z-range 5540–5640, detected fragment ions (a and b) matched the theoretical isotopic distributions of c-type fragment ions (c) of the light chain (Lc) without sulfation at Tyr50. In the m/z-range 6292–6390, clear differences were observed between the MS spectra generated from Fab portion (d) without and (e) with sulfation. The presence of the sulfation on Tyr57 of the Lc was confirmed by the detection of the fragment c57 with sulfation (i.e., c57*) at m/z 6376.14 (inset in e) which was in good agreement with the calculated theoretical isotopic distribution (f). Increased confidence in the identification of sulfated c57 fragment was obtained from absorption mode visualization (Figure S3). Larger sulfated c-type fragments were detected at higher m/z (Figure S4).

5.5.2 Aspartate in amino-terminal position -3 is involved in determining the sulfation of mAb1

In search for structural motifs that would instruct Tyr57 sulfation, we analyzed a second antibody, mAb2, which shows high sequence homology to mAb1. Notably, MS analysis after reduction did not indicate any +80 Da variant of the light chain (Figure 1B), and when analyzed by CIEX, mAb2 did not contain a distinct acidic variant as compared to mAb1 (Figure 2A). The amino acid sequences of the light and heavy chains of mAb1 and mAb2 are 98.6% and 99.8% identical, respectively, with the following few amino acid differences between the Lc's; Asn31 to Thr31, Lys33 to Arg33 (both CDR-1), and Asp54 to Thr54 (CDR-2) (mAb1 to mAb2, respectively). The modified mAb1 tryptic peptide thereby contains an aspartate (Asp54) in position 7 (LIYSASDLDYGVPSR), whereas the corresponding mAb2 tryptic

y2+ (modified peptide: m/z 272.1717) and y3+ (modified peptide: m/z 359.2037), and b2+ (modified peptide: m/z 201.1239) to b9+ (modified peptide: m/z 878.4987) fragment ions were detected for both peptides, locating the +15.995 Da modification of the BsAbA tryptic peptide to the proline residue in position 10 (VTPEIPAGLPSPR). The observed mass difference between the y4+ and y3+ ions of 113.0477 Da was found in perfect agreement with the presence of a hydroxyproline (theoretical mass increment of 113.0477 Da) in position 10 and virtually ruled out alternative explanations such as proline to leucine and/or isoleucine (113.0841 Da) sequence variants. Also, identical misincorporations in both chain A and B of BsAbA were unlikely unless proline limitation during fermentation was playing a role. Of note, the +15.995 Da modified peptide appeared to be less hydrophobic than the unmodified tryptic peptide as evidenced by less retention in RP-LC (**Figure 3B**). In contrast, both leucine and isoleucine misincorporations would be expected to increase the hydrophobicity of the modified peptide.

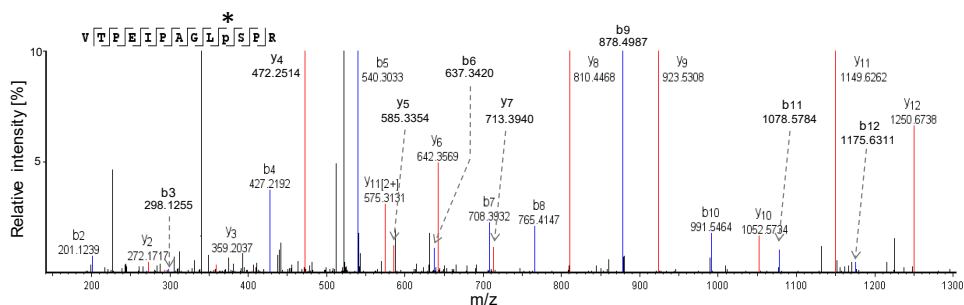


Figure 6. Orbitrap MS/MS data obtained by collision-induced dissociation of the doubly protonated +15.995 Da modified tryptic peptide (VTPEIPAGLPSPR) of BsAbA. Mass differences between the b10+ (m/z 991.5464) to b12+ (m/z 1175.6311), and y4+ (m/z 472.2514) to y12+ (m/z 1250.6738) fragments of the modified peptide and the corresponding theoretical ions of the unmodified peptide were identified. In addition, no modified b2+ (m/z 201.1239) to b9+ (m/z 878.4987), y2+ (m/z 272.1717), and y3+ (m/z 359.2037) fragment ions for the modified peptide were detected, implying the +15.995 Da modification to be localized to the proline residue in position 10 (*).

The modified peptide is part of the C-terminus of the three ligand domains of BsAbA (**Figure 1C**). The mass variants observed by intact ESI-QTOF-MS agrees with the fact that the identified peptide is present once in chain A containing one of the three identical ligand domains and in duplicate in chain B containing two of the ligand domains of BsAbA. The +32 Da variant of chain B (**Figure 1D**) thereby corresponds to both peptides/proline residues being modified. Also the 46% relative quantification at peptide level (**Figure 3B**) correlates with the signal intensities of the +16 and +32 Da variants of the reduced and deglycosylated BsAbA (**Figure 1D**).

5.5.4 Spiking of synthetic peptides prove the presence of 4Hyp in BsAbA and excludes 3Hyp, Leu, and Ile

Hydroxyproline is known to exist in two isomeric forms, trans-3- (3Hyp) and trans-4-hydroxyproline (4Hyp)^{37,38} with the same nominal mass of 113 Da as leucine and isoleucine. Leucine and isoleucine are successfully distinguished in MS³ experiments (e.g. ETD-HCD) by secondary fragmentation of z-ions to generate C_g-C_γ bond cleavages and to produce w-ions for the specific side chains.^{39,40} A similar approach has been applied to differentiate 3Hyp and 4Hyp using high energy CID mass spectrometry.³⁷ However, when performing different MS³ fragmentation experiments using an Orbitrap Fusion Lumos, we were not able to generate the w-ions for the specific Hyp side chains (data not shown). We therefore tested, whether we could determine the identity of the +15.995 Da modification by spiking of synthetic peptides containing 4Hyp (VTPEIPAGL4HypSPR), 3Hyp, (VTPEIPAGL3HypSPR), leucine (VTPEIPAGLLSPR), and isoleucine (VTPEIPAGLI_SSPR) to the BsAbA tryptic digest with subsequent analysis by LC-MS. EICs of the unmodified and modified tryptic peptides with and without spiking of the synthetic peptides are shown in **Figure 7** (charge states z: 2 and 3). In a spiking experiment with the synthetic 4Hyp-containing peptide, the area of the EIC peak eluting at 38.3-38.5 min and representing the modified peptide increased 8-fold (**Figure 7E**), demonstrating the co-elution of the +15.995 Da modified tryptic peptide of BsAbA and the synthetic 4Hyp peptide. With elution times of 38.3-38.5 min (**Figure 7A-E**) and 39.8 min (**Figure 7D**) of the 4Hyp and 3Hyp-containing peptides, respectively, the spiking experiment demonstrate that both synthetic hydroxyproline-containing peptides are less hydrophobic on the reverse-phase column relative to the unmodified peptide (41.3-41.5 min) as expected (**Figure 7**). Both the synthetic isoleucine (46.5 min) and leucine (48.3 min) containing peptides eluted later than the unmodified peptide (**Figure 7B** and **C**). In conclusion, the spiking experiments demonstrated a chromatographic separation of the two peptides containing the isomeric hydroxyprolines and clearly allowed identification of the BsAbA modification as 4Hyp.

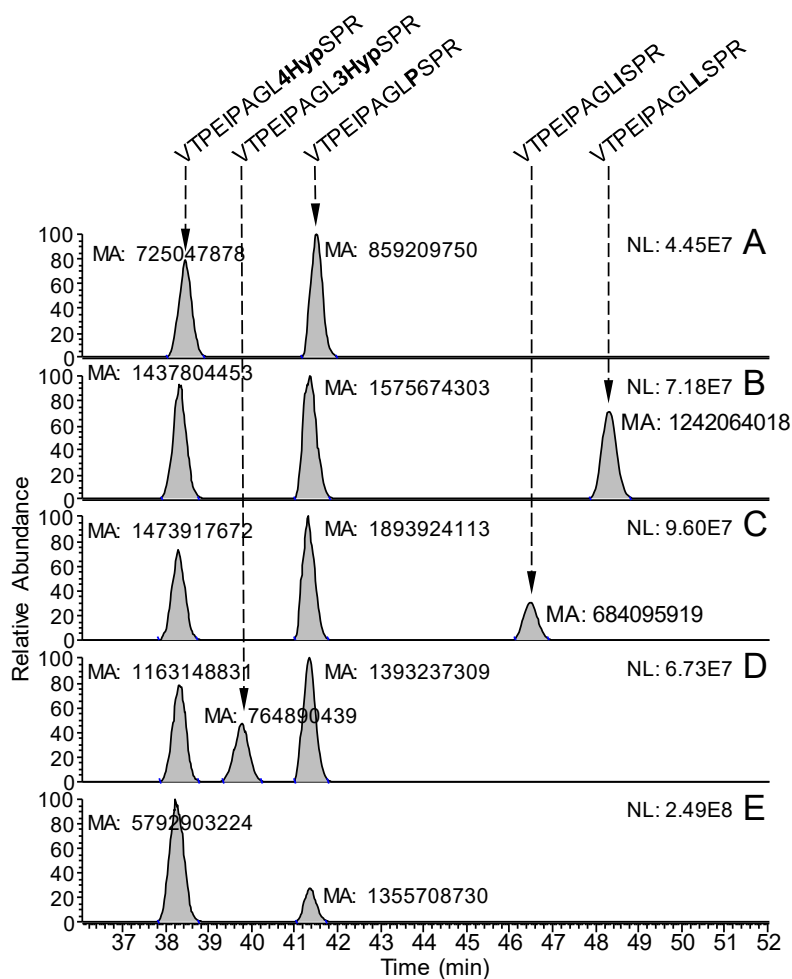


Figure 7. Identification of the modified proline residue as 4-hydroxyproline by spiking of synthetic peptides to the tryptic digest of BsAbA. Extracted ion current chromatograms (charge states z : 2 and 3) of (a) the +15.995 Da modified (38.4 min) and unmodified (41.5 min) tryptic peptides (VTPEIPAGLSPR) without spiking, and spiked with synthetic peptides containing (b) leucine (VTPEIPAGLLSPR, 48.3 min), (c) isoleucine (VTPEIPAGLISPR, 46.5 min), (d) 3-hydroxyproline (VTPEIPAGL3HypSPR, 39.8 min), and (e) 4-hydroxyproline (VTPEIPAGL4HypSPR, 38.3 min). The peak area of the +15.995 Da modified tryptic peptide increased when the tryptic digest was spiked with the synthetic peptide containing 4-hydroxyproline. NL, normalized intensity level; MA, manually integrated peak area.

5.5.5 In silico identification to hint at the potential presence of sTyr and 4Hyp at early development stage

Early in silico identification of the 4Hyp and sTyr formations could have prevented the heterogeneities of BsAbA and mAb1. On this basis, we developed in-house in silico models based on peptide sequences extracted from the Swiss-Prot database and, in the case of sTyr, also the physicochemical and biochemical properties of various selected amino acids (see Material and Methods). As the focus of our research is the development of therapeutic proteins for human applications, all non-mammalian sequences were excluded in this development, especially as they don't show the typical mammalian pattern around PTM prone positions. For both sTyr and 4Hyp, we obtained very good model qualities (**Table S1**) with respect to all considered performance measures. In particular, we outperformed the sTyr in silico identification performance as reported by Yang⁴² in terms of sensitivity while keeping specificity at highest level, especially with undersampling. First, we applied our random forest (RF) and k-nearest neighbor (kNN) models to the sequence RLIYSASDLDYGVPSRFGSGG comprising the mAb1 CDR-2, and all other Tyr residues found in the protein sequence of mAb1. In total, 18 Tyr residues were present. None of the Tyr residues were predicted as modified with a probability of $\geq 50\%$, but the position RLIYSASDLDYGVPSRFGSGG had the highest probability for modification among all 18 sites tested, both for RF (pPTM=0.18) and kNN (pPTM=0.33), while the similar position RLIYSASTLDYGVPSRFGSGG of mAb2 ranked in the middle for RF (pPTM=0.11), and lowest for kNN (pPTM=0). Thus, given our prior suspicion on the presence of a sulfation, the algorithms gave the correct information about the location of the PTM, despite the very difficult setting (modified and unmodified sequence differ only at one position). In comparison, the computational tyrosine sulfation site predictor SulfoSite⁴³ predicts the modified tyrosine residue mAb1 with a probability of 95% and the identical but unmodified position in mAb2 with a probability of 75%. In the case of the 4Hyp modification in BsAbA, the kNN had a 50% probability for the affected residue, which was the highest probability observable for all sites and methods in the chain A sequence. Thus, the in silico tool again identified the correct location. Sequence logos illustrating the most conserved amino acids around sTyr and 4-Hyp residues based on the mammalian peptide sequences we extracted from the Swiss-Prot database are shown in **Figure 8**.

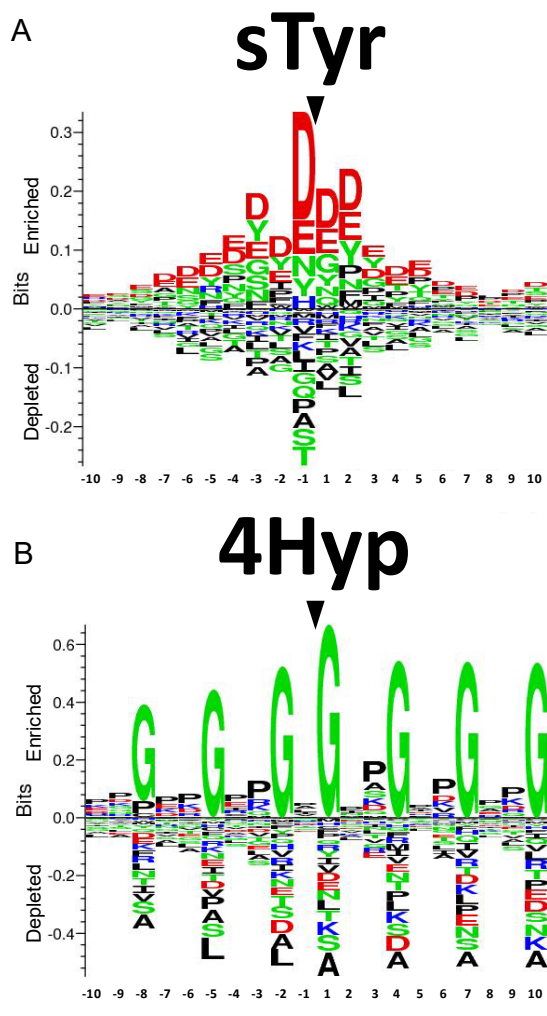


Figure 8. Sequence logos showing the most conserved amino acids around (a) sulfotyrosine (sTyr) and (b) 4-hydroxyproline (4Hyp) residues based on mammalian peptide sequences extracted from the Swiss-Prot database. For scaling reasons, the positions of the central sTyr and 4Hyp residues were marked after creating the logos using Seq2Logo.

5.6 Discussion

Several enzyme-catalyzed PTMs are well known to antibodies and other biotherapeutics, others are more unexpected. Some enzyme-catalyzed PTMs have been reported at significant levels in therapeutic proteins transiently expressed in human embryonic kidney (HEK) cells during early development. However, once the proteins are stably expressed in CHO cells, the very same PTMs have been found to be present at low or very low levels.

E.g. the O-xylosylations of glycine-serine linkers have been reported to be 10-fold higher in fusion proteins expressed in HEK compared to CHO cells.¹⁰ This observation matches with our experience with several different fusion proteins in development. We also reported 11% phosphoserine in a glycine-serine linker of a fusion protein transiently expressed in HEK cells compared to <1% when stably expressed in CHO cells.¹¹ Evaluation of further batches of the fusion protein from CHO cells has demonstrated levels <0.1% (unpublished data). Other than cultivation and feeding conditions, the root cause for these quantitative differences between the two expression systems may include differences in drug productivity, and in the expression level, activity, or localization of enzymes involved, and the availability of important co-factors.

However, unexpected PTMs are not always at low levels in CHO produced proteins. Here, we report the detection of substantial levels of positional 4Hyp and sTyr in two proteins expressed in CHO cells resulting in significant mass and for sTyr charge variant heterogeneities. If a charge variant is influenced by critical process parameters and not kept at a certain level, it might cause out-of-specification events revealing lack of process control. The accumulation of PTMs for bi- or multivalent molecules is non-linear as pictured in the supplementary **Figure S1**. Consequently, the expected influence on an ion exchange chromatography profile is also non-linear.

Both 4Hyp and sTyr were technically challenging to characterize and verify. The identification of the 4Hyp was finally accomplished by chromatographic separation of peptides containing the two isomeric forms of hydroxyproline. To our knowledge, the differentiation of the two isomeric forms of hydroxyproline by spiking of synthetic peptides has not previously been reported. In the case of the sTyr, all fragmentation methods tested including ETD, ETHcD, and negative ion mode CID did not allow to determine the position of the sulfation of mAb1. Finally, a novel top-down MALDI-ISD FT-ICR MS method²¹ recently developed for the characterization of the amino acid sequence of mAbs, was successfully applied to intact Fab fragments proving the applicability of this powerful ultrahigh resolution method for the analysis of sulfated proteins and therapeutic proteins in general. To our knowledge, this is the first application of MALDI-ISD FT-ICR MS for the characterization of sulfated mAbs. Antibody CDR sulfation was first reported by Zhao et al.⁷ using positive ion mode ETD. Often, however, sulfations have been reported to be undetectable using conventional positive ion mode fragmentation techniques due to poor ionization efficiency and sulfate loss preceding peptide backbone fragmentation.^{35,44}

Recently, hydroxyproline was reported at a similar high level (40%) in a partially rigid G₄P linker as part of an Fc-growth factor fusion protein stably expressed in CHO cells.⁹ Also the sTyr recently reported in the Lc CDR-1 of a CHO-derived antibody was determined to a relatively high (around 20%) at the peptide level (corresponding to ~40% at intact level).⁷ Together, these observations and the data presented here demonstrate that CHO cells are

generally capable of introducing high levels of hydroxyproline and sulfotyrosine in therapeutic proteins which thereby significantly increase the heterogeneity of the molecules. In a very recent study by Hou et al.¹², the impact of specific nutrients on phosphoserine (~20%) and hydroxylysine (~25%) of a CHO platform fed batch produced Fc-fusion protein was reported. Although the exact influence of the nutrients is unknown, increased vitamin C, ferric citrate, and niacinamide feeding rates and a decreased cysteine feeding rate reduced the phosphorylation level to ~3%. An increase in the niacinamide and cysteine feeding rates reduced the hydroxylation level to around 10%. Whether similar reductions are possible with 4Hyp and sTyr is unknown. However, as the pharma industry is often using standardized fed-batch cultivation and feeding strategies, molecule-specific major adaptations are mostly undesired.

In the recent years, several pharma companies have reported the development of structure based tools for the prediction of chemical PTMs like deamidation, isomerisation, and oxidation hot spots in therapeutic antibodies²²⁻²⁴. Also several open-source web-based predictors have been developed to suggest positions of enzyme-catalyzed PTMs based on consensus sequences/logos (see e.g. www.expasy.org/proteomics, www.cbs.dtu.dk/databases/PTMpredictions, www.modpred.org/). Some of these predictors are available as downloads, however, often they are not optimal for therapeutic proteins, as they are not exclusively based on human or mammalian protein sequences. Nevertheless, in the pharma industry it is not customary to upload amino acid sequence information of early development biotherapeutics to non-regulated external web-sites. With the detection and verification of the substantial levels of the enzyme-catalyzed PTMs in mAb1 and BsAbA, we therefore decided to pursue a prevention strategy by developing in silico tools for sTyr and 4Hyp and several other enzyme-catalyzed PTMs found in mammalian proteins. The purpose is to obtain an early note to where enzyme-catalyzed PTMs may occur in drug candidates, prevent that PTMs at low relative abundance may stay undiscovered or are uncovered in later phases of drug development, where it may be too late to alter the molecule. Furthermore, suggested positions for PTMs will help guide a positional analytical approach and minimize time-consuming analysis which may delay drug development.

As with several other PTMs, the activity of prolyl hydroxylases and tyrosyl sulfotransferases are governed by the adjacent amino acids relative to the affected residues. Examination of modified tyrosines has revealed that sulfotyrosines are characterized by acidic amino acids in the immediate environment, especially on the amino-terminal side of the affected tyrosine with the majority possessing a negative residue in the -1 position being the most important single determinant.⁴⁵⁻⁴⁷ Also the sequence logo we generated for sTyr in mammalian proteins illustrates that acidic amino acids are surrounding the sTyr (**Figure 8A**). Bundgaard et al.⁴⁷ suggested that acidic residues in the positions -1 and possibly -3 enhance sulfation, although they are not required to obtain a partial sulfation. When comparing the position of the sTyr in mAb1 with the sequence logo, the aspartates in the positions -1 and -3 relative

to the modified Tyr (RLIYSASDLDYGVPSRFSGSG) are likely both involved in determining the sTyr in mAb1. As the Asp in position -1 alone did not cause tyrosine sulfation in mAb2, the Asp in position -3 in mAb1 is likely essential for the affinity of tyrosyl sulfotransferase for this PTM to occur in mAb1. Moreover, the Gly residue in position +1 of mAb1 is enriched in the sTyr sequence logo.

The sequence logo we generated for 4Hyp in mammalian proteins implies key roles for Gly residues in the positions -8, -5, -2, +1, +4, +7, and +10 relative to the 4Hyp, and several enriched positions for Pro (**Figure 8B**). Similar observations have been reported for predictors based not exclusively on mammalian proteins.^{48,49} When comparing the position of the 4Hyp in the 4-BBL ligand domains of BsAbA with the sequence logo we generated for 4Hyp in mammalian proteins (**Figure 8B**), the Gly in position -2 relative to the modified proline (VTPEIPAGL4HypSPR) is likely critical for the proline hydroxylation of BsAbA. Also the Pro residues in -7, -4 and +2 are enriched in the 4Hyp sequence logo. To our knowledge, 4Hyp in human 4-1BBL has not been reported. As BsAbA involves 4-1BBL domains with carboxy-terminal glycine-serine linkers (VTPEIPAGL4HypSPRSEGGGSGGGGS; (see also **Figure 1C**) the linker Gly residue in position +7 might be involved in determining the 4Hyp of BsAbA. This could indicate that the specific sequence of BsAbA determines the presence of the 4Hyp.

5.7 Supporting information

Supplementary information is available free of charge on the mAbs website: DOI: 10.1080/19420862.2019.1635865

5.8 Acknowledgments

We thank Eva Greiter and Björn Mautz for experimental support, the Roche Penzberg management for continuous support, and Yury O. Tsybin (Spectroswiss, Lausanne, Switzerland) and David P. A. Kilgour (Nottingham Trent University, Nottingham, U.K.) for providing AutoVectis software and training support.

5.9 References

1. Beck A., Liu H. Macro- and micro-heterogeneity of natural and recombinant IgG antibodies. *Antibodies*. 2019, 8, 18.
2. Kuriakose A., Chirmule N., Nair P. Immunogenicity of biotherapeutics: causes and association with posttranslational modifications. *J Immunol Res*. 2016, 1298473.
3. Xie Q., Moore B., Beardsley R. L. Discovery and characterization of hydroxylysine in recombinant monoclonal antibodies. *MAbs*. 2016, 8, 371–78.
4. Valliere-Douglass J. F., *et al.* Asparagine-linked oligosaccharides present on a non-

- consensus amino acid sequence in the CH1 domain of human antibodies. *J Biol Chem.* 2009, 284, 32493–506.
5. Valliere-Douglass J. F., *et al.* Glutamine-linked and non-consensus asparagine-linked oligosaccharides present in human recombinant antibodies define novel protein glycosylation motifs. *J Biol Chem.* 2010, 285, 16012–22.
 6. Valliere-Douglass J. F., *et al.* O-fucosylation of an antibody light chain: characterization of a modification occurring on an IgG1 molecule. *Glycobiology.* 2009, 19, 144–52.
 7. Zhao J., *et al.* Characterization of a novel modification of a CHO-produced mAb: evidence for the presence of tyrosine sulfation. *MAbs.* 2017, 9, 985–95.
 8. Spahr C., *et al.* Recombinant human lecithin-cholesterol acyltransferase Fc fusion: analysis of N- and O-linked glycans and identification and elimination of a xylose-based O-linked tetrasaccharide core in the linker region. *Protein Sci.* 2013, 22, 1739–53.
 9. Spahr C., *et al.* High-resolution mass spectrometry confirms the presence of a hydroxyproline (Hyp) post-translational modification in the GGGGP linker of an Fc-fusion protein. *Mabs.* 2017, 9, 812–19.
 10. Spencer D., *et al.* O-xylosylation in a recombinant protein is directed at a common motif on glycine-serine linkers. *J Pharm Sci.* 2013, 102, 3920–24.
 11. Tyshchuk O., *et al.* Detection of a phosphorylated glycine-serine linker in an IgG-based fusion protein. *MAbs.* 2017, 9, 94–103.
 12. Hou Y., *et al.* Nutrient optimization reduces phosphorylation and hydroxylation level on an Fc-fusion protein in a CHO fed-batch process. *Biotechnol J.* 2018, 3, e1700706.
 13. Huang L., *et al.* Impact of variable domain glycosylation on antibody clearance: an LC/MS characterization. *Anal Biochem.* 2006, 349, 197–207.
 14. McSherry T., *et al.* Cysteinylation of a monoclonal antibody leads to its inactivation. *MAbs.* 2016, 8, 718–25.
 15. Jefferis R. Posttranslational modifications and the immunogenicity of biotherapeutics. *J Immunol Res.* 2016, 2016, 5358272.
 16. Huang H. Z., Nichols A., Liu D. Direct identification and quantification of aspartyl succinimide in an IgG2 mAb by RapiGest assisted digestion. *Anal Chem.* 2009, 81, 1686–92.
 17. Diepold K., *et al.* Simultaneous assessment of Asp isomerization and Asn deamidation in recombinant antibodies by LC-MS following incubation at elevated temperatures. *PLoS One.* 2012, 7, :e30295.
 18. Cornell C., *et al.* A high-throughput hydrophilic interaction liquid chromatography coupled with a charged aerosol detector method to assess trisulfides in IgG1 monoclonal antibodies using tris(2-carboxyethyl)phosphine reaction products: tris(2-carboxyethyl) phosphine-oxide and tris(2-carboxyethyl)phosphine-sulfide. *J Chromatogr A.* 2016, 1457, 107–15.
 19. Zhang W., *et al.* Complete disulfide bond assignment of a recombinant immunoglobulin G4 monoclonal antibody. *Anal Biochem.* 2002, 311, 1–9.



20. Frese C. K., *et al.* Unambiguous phosphosite localization using electrontransfer/higher-energy collision dissociation (ETHcD). *J Proteome Res.* 2013, 12, 1520–25.
21. Van der Burgt Y. E. M., *et al.* Structural analysis of monoclonal antibodies by ultra-high resolution MALDI in-source decay FT-ICR mass spectrometry. *Anal Chem.* 2019, 91, 2079–85.
22. Sydow J. F., *et al.* Structure-based prediction of asparagine and aspartate degradation sites in antibody variable regions. *PLoS One.* 2014, 9, e100736.
23. Yan Q., *et al.* Structure based prediction of asparagine deamidation propensity in monoclonal antibodies. *MAbs.* 2018, 10, 901–12.
24. Sankar K., *et al.* Prediction of methionine oxidation risk in monoclonal antibodies using a machine learning method. *MAbs.* 2018, 10, 1281–90.
25. Eisenhaber B., Eisenhaber F. Prediction of posttranslational modification of proteins from their amino acid sequence. *Methods Mol Biol.* 2010, 609, 365–84.
26. Gianazza E., *et al.* In silico prediction and characterization of protein post-translational modifications. *J Proteomics.* 2016, 134, 65–75.
27. Mohabatkari H., Rabiei P., Alamdarani M. New achievements in bioinformatics prediction of post translational modification of proteins. *Curr Top Med Chem.* 2017, 17, 2381–92.
28. Klein C., Schaefer W., Regula J. T. The use of CrossMAb technology for the generation of bi- and multispecific antibodies. *MAbs.* 2016, 8, 1010–20.
29. Regula J. T., *et al.* Variable heavy-variable light domain and Fab-arm CrossMabs with charged residue exchanges to enforce correct light chain assembly. *Protein Eng Des Sel.* 2018, 31, 289–99.
30. Claus C., *et al.* Tumor-targeted 4-1BB agonists for combination with T-cell bispecific antibodies as off-the-shelf therapy. *Sci Transl Med.* 2019, 11, pii: eaav5989.
31. Shi X., *et al.* Tandem mass spectrometry of heparan sulfate negative ions: sulfate loss patterns and chemical modification methods for improvement of product ion profiles. *J Am Soc Mass Spectrom.* 2012, 23, 1498–511.
32. Chen G., *et al.* Distinguishing sulfotyrosine containing peptides from their phosphotyrosine counterparts using mass spectrometry. *J Am Soc Mass Spectrom.* 2018, 29, 455–62.
33. Niehrs C., Beisswanger R., Huttner W. B. Protein tyrosine sulfation, 1993—an update. *Chem Biol Interact.* 1994, 92, 257–71.
34. Medzihradszky K. F., *et al.* O-sulfonation of serine and threonine: mass spectrometric detection and characterization of a new posttranslational modification in diverse proteins throughout the eukaryotes. *Mol Cell Proteomics.* 2004, 3, 429–40.
35. Stone M. J., *et al.* Tyrosine sulfation: an increasingly recognised post-translational modification of secreted proteins. *N Biotechnol.* 2009, 25, 299–317.
36. Lin W. H., *et al.* Recognition of substrates by tyrosylprotein sulfotransferase. Determination of affinity by acidic amino acids near the target sites. *J Biol Chem.* 1992, 267, 2876–79.

37. Kabat E. A., *et al.* Sequences of proteins of immunological interest. Bethesda (MD): U.S. department of health and human service; 1991.
38. Kassel D. B., Biemann K. Differentiation of hydroxyproline isomers and isobars in peptides by tandem mass spectrometry. *Anal Chem.* 1990, 62, 1691–95.
39. Srivastava A. K., *et al.* Hydroxyproline: A potential biochemical marker and its role in the pathogenesis of different diseases. *Curr Protein Pept Sci.* 2016, 17, 596–602.
40. Johnson R. S., *et al.* Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Anal Chem.* 1987, 59, 2621–25.
41. Lebedev A. T., *et al.* Discrimination of leucine and isoleucine in peptides sequencing with Orbitrap fusion mass spectrometer. *Anal Chem.* 2014, 86, 7017–22.
42. Smeeton N. C. Early history of the kappa statistic. *Biometrics.* 1985, 41, 795.
43. Yang Z. R. Predicting sulfotyrosine sites using the random forest algorithm with significantly improved prediction accuracy. *BMC Bioinformatics.* 2009, 10, 361.
44. Chang W. C., *et al.* Incorporating support vector machine for identifying protein tyrosine sulfation sites. *J Comput Chem.* 2009, 30, 2526–37.
45. Monigatti F., Hekking B., Steen H. Protein sulfation analysis – a primer. *Biochim Biophys Acta.* 2006, 1764, 1904–13.
46. Hortin G., *et al.* Characterization of sites of tyrosine sulfation in proteins and criteria for predicting their occurrence. *Biochem Biophys Res Commun.* 1986, 141, 326–33.
47. Bundgaard J. R., Vuust J., Rehfeld J. F. New consensus features for tyrosine O-sulfation determined by mutational analysis. *J Biol Chem.* 1997, 272, 21700–05.
48. Shi S. P., *et al.* PredHydroxy: computational prediction of protein hydroxylation site locations based on the primary structure. *Mol Biosyst.* 2015, 11, 819–25.
49. Li S., *et al.* HydPred: a novel method for the identification of protein hydroxylation sites that reveals new insights into human inherited disease. *Mol Biosyst.* 2016, 12, 490–98.
50. Strohm M., *et al.* mMass data miner: an open source alternative for mass spectrometric data analysis. *Rapid Commun Mass Spectrom.* 2008, 22, 905–08.
51. Kawashima S., Kanehisa M. AAindex: amino acid index database. *Nucleic Acids Res.* 2000, 28, 374.





Chapter 6

Intact and subunit-specific analysis of bispecific antibodies by sheathless CE-MS

Christoph Gstöttner^a, Simone Nicolardi^a, Markus Habberger^b, Dietmar Reusch^b, Manfred Wuhrer^a, Elena Domínguez-Vega^a

^aLeiden University Medical Center, Center for Proteomics and Metabolomics, Leiden, The Netherlands

^bRoche Pharma Technical Development, Penzberg, Germany

Reprinted and adapted with permission from *Analytica Chimica Acta*, 2020, 1134, 18-27, DOI: 10.1016/j.aca.2020.07.069

6.1 Abstract

Bispecific antibodies (BsAb) are next-generation, antibody-based pharmaceuticals which come with a great functional versatility and often a vast structural heterogeneity. Although engineering of the primary sequence of BsAbs guides the proper pairing of the different chains, several side products can often be observed contributing to the macroheterogeneity of these products. Furthermore, changes in the amino acid sequence can result in different protein modifications which can affect the properties of the antibody and further increase the structural complexity. A multi-methods approach can be used for the characterization of their heterogeneity but new analytical strategies are needed for a more accurate and in-depth analysis.

Here, we present a combination of intact antibody and subunit-specific mass measurements using sheathless capillary electrophoresis-mass spectrometry for assessing the macro- and microheterogeneity of BsAbs. Two homologous BsAbs with the same bispecificity but slightly different amino acid sequences were analyzed. Intact measurements were performed using a positively coated capillary and a background electrolyte (BGE) consisting of 3% acetic acid. For intact BsAbs, the separation permitted the characterization of free light chains, homo- and heterodimers as well as incomplete assemblies. For subunit-specific measurements, BsAbs were hinge region cleaved using two different enzymes (SpeB and IdeS) followed by disulfide-bond reduction. The six different subunits (Lc1, Lc2, Fd'1, Fd'2, (Fc/2)1 and (Fc/2)2) were separated using the same positively-coated capillary and a BGE consisting of 20% acetic acid and 10% methanol. Mass measurements of hinge region cleaved antibodies were performed at isotopic resolution (resolving power 140000 at m/z 1100) for a more confident analysis of low abundance proteoforms. For both BsAbs several proteoforms with e.g. pyroglutamic acid (Pyro-Glu) or glycation which could not be properly assigned at the intact level, were accurately determined in the subunits showing the complementarity of both approaches.

6.2 Introduction

The last decade has been marked by biotechnological innovations that have permitted the exploitation of novel ways of antibody production and biopharmaceuticals development. Traditional monoclonal antibodies (mAbs) are being slowly displaced by increasingly complex antibody formats which can tackle some limitations of mAbs (e.g. higher binding avidity, lower resistance or higher cytotoxicity¹) opening new ways for therapeutic intervention². Examples of these formats are Fc-fusion proteins³ or bispecific monoclonal antibodies (BsAbs)⁴. BsAbs, in particular, are taking a predominant position in biopharma as they can bind to two different epitopes. This makes BsAbs especially attractive for cancer immunotherapy as they can bring two target cells in close proximity (e.g. a cancer cell and a cytotoxic T cell). Another example is the treatment of haemophilia A by the recently approved BsAb emicizumab which binds the blood coagulation factors IX and X⁵.

Production of bispecific antibodies requires proper assembly of two different light chains (Lc) and two different heavy chains (Hc). Recent strategies in antibody engineering such as knob-into-hole (for Hc-Hc pairing) or crossmab (for Lc-Hc assembly) have enabled efficient pairing of chains within BsAbs (**Figure 1A**)^{6,7}. Still, side products such as free chains, incomplete antibodies or aggregates are commonly observed during production which can induce unwanted immunological responses⁸. This so-called macroheterogeneity should be carefully monitored to ensure patient safety. In addition, antibodies are inherently heterogeneous molecules containing different glycoforms and many other post-translational modifications (PTMs) which can impact the functionality of the antibody^{9,10}. Importantly, the location of these modifications can influence their effect on the functionality of the protein. For instance, oxidation will affect antigen binding or reduce the half-life of the antibody depending on whether it is positioned in the antigen binding fragment (Fab) or in the Fc region^{10,11}. Therefore, next to the identification also the assignment to antibody domains is important.

confidence to the assignment of low mass differences, but also allocating the modifications to domains and subunits. The combination of intact antibody and subunit analysis provides a reliable assessment of both macro- and microheterogeneity. However, due to the large differences in mass their analysis often required different analytical strategies. For the macroheterogeneity analysis, SEC-MS is generally used^{13, 18-20} while for subunit analysis RPLC-MS²¹⁻²³ or HILIC-MS²⁴ are more suitable.

Capillary electrophoresis (CE)-MS has demonstrated to be a powerful technique to analyze intact proteins and their proteoforms. Due to the characteristics of the separation capillary (open tube), CE can be used to separate proteins of a wide range of sizes. This makes CE-MS a very attractive technique for the analysis of both, intact mAbs and subunits without the need for alternative platforms²⁵⁻²⁸. In particular, when hyphenated via the sheathless interface, high ionization efficiencies and sensitivity can be obtained which might be especially useful for the detection of low abundant PTMs. Sheathless CE-MS has already demonstrated its ability to separate either intact²⁷ or hinge region-cleaved^{25, 26, 29} monospecific mAbs and their proteoforms. BsAbs, however, present larger macroheterogeneity due to the more complex production process. Furthermore, after hinge region cleavage and reduction, engineered BsAbs consist generally of six individual subunits (**Figure 1B and C**) instead of the three commonly present in traditional mAbs. Therefore, in this work, we explore the use of sheathless CE-MS for intact and subunit-specific characterization of BsAbs. Two homologous BsAbs obtained by different engineering processes were included in the study. Analysis of the intact BsAbs allowed monitoring incomplete assemblies and free chains. Using two alternative hinge region-specific enzymes, the six individual subunits were obtained and consequently characterized using sheathless CE-MS. This permitted the determination of PTMs such as glycation and assignment to specific portions of the molecule.

6.3 Materials and Methods

6.3.1 Reagents

Sodium hydroxide ($\geq 99.8\%$), di-thiothreitol (DTT) ($\geq 98\%$), tris(2-carboxyethyl)phosphine (TCEP) ($\geq 98\%$), hydrogen chloride (ACR reagent grade 37%), sodium chloride ($\geq 99.8\%$), methanol anhydrous ($\geq 99.8\%$) and Tris-HCl ($\geq 99.8\%$) were obtained from Sigma-Aldrich (St. Louis, MO). Methanol (Ultra LC-MS grade) was provided by Actua-All Chemicals (Randmeer, Oss, The Netherlands). Water (ULC/MS - CC/SFC grade) was purchased from Biosolve Chimie SARL (Dieuze, France). NaHCO_3 ($\geq 99.7\%$), was purchased from Merck KGaA (Darmstadt, Germany). Acetic acid ($\geq 99.7\%$) was purchased from VWR Chemicals (Radnor, PA). 8 M aqueous solution Guanidinium chloride (GdnCl) was obtained from Thermo Scientific (Waltham, MA). Ammonium acetate stock solution 7.5 M (for molecular biology) was purchased from Sigma-Aldrich. 50% tri-methoxysilylpropyl modified polyethylenimine (PEI) solution in isopropanol was obtained from Gelest (Morrisville, NC). IdeS (Fabricator) and

SpeB (Fabulous) endoproteinases were purchased from Genovis (Lund, Sweden). Formulated BsAb1 (pI 7.6) and BsAb2 (pI 7.4) were provided by Roche Diagnostics (Penzberg, Germany). The protein test mixture was obtained from Sciex (Brea, CA).

6.3.2 Sample preparation for intact BsAbs measurement

For intact measurements, BsAb1 and BsAb2 were three times desalted with water using 10 kDa Vivaspin MWCO filter tubes (Sartorius, Göttingen, Germany) at 10000xg at 4 °C. The samples were afterward diluted to a final concentration of 1 µg/µL and either directly measured or stored for a short period at 4 °C.

6.3.3 Hinge region cleavage of BsAbs

Two enzymes were used for the hinge region cleavage of the bispecific antibodies (IdeS and SpeB for BsAb1 and BsAb2, respectively). For IdeS digestion, the BsAb1 was diluted in 100 mM NaHCO₃ pH 6.8 (adjusted with acetic acid) to a final concentration of 1 µg/µL. For digestion, one Unit of IdeS per µg of mAb was added and incubated for 90 min at 37 °C. For SpeB digestion, the BsAb2 was diluted in 100 mM Tris-HCl pH 8.0 to a final concentration of 1 µg/µL and incubated with SpeB (one unit per µg mAb) at 37 °C for 60 min.

6.3.4 Reduction of hinge region cleaved BsAbs

After hinge region cleavage, BsAbs were denatured by diluting the BsAbs 1:1 (v:v) with 8 M GdnCl in 100 mM Tris pH 8.0 solution. After dilution, 1 M DTT was added to a final concentration of 20 mM. The samples were incubated at 60 °C for 30 min. After reduction, the samples were three times desalted with either water (BsAb1) or 100 mM ammonium acetate pH 3.0 (BsAb2) using 10 kDa Vivaspin MWCO filter tubes at 10000xg at 4 °C. The samples were directly measured or stored for a short period at 4 °C.

For TCEP reduction, used during reduction optimization, a 1 M TCEP solution in 100 mM NaHCO₃ at pH 6.8 was prepared and added to the sample to a final concentration of 100 mM TCEP. Samples were desalted as described above.

6.3.5 Sheathless CE-MS

Sheathless CE-MS experiments were performed on a Sciex CESI 8000 instrument. Bare fused silica capillaries with a porous tip were obtained from Sciex (separation capillary: 91 cm x 30 µm i.d., capillary volume 643 nL; conductive line: 70 cm x 50 µm i.d., capillary volume 1374 nL) and were in-house coated with PEI following the protocol described by Sciex³⁰. Briefly, the capillaries were pre-conditioned by flushing the separation capillary for 20 min with 0.1 M NaOH (75 psi, forward pressure, 21 capillary volumes), 10 min (75 psi, forward pressure, 10.5 capillary volumes) with 0.1 M HCl, 20 min with H₂O (75 psi, forward pressure, 21 capillary volumes), 20 min with MeOH (75 psi, forward pressure, 21 capillary

volumes) and the conductive line 5 min with MeOH (50 psi, reverse pressure). Subsequently, the separation capillary was flushed with PEI coating solution for 30 min (75 psi, forward pressure, 31.5 capillary volumes) and incubated overnight. The next day the capillary was cleaned with MeOH for 10 min (75 psi, forward pressure, 10.5 capillary volumes), 30 min from a second vial with MeOH (75 psi, forward pressure 31.5 capillary volumes) and 5 min the conductive line with MeOH (100 psi, reverse pressure). Hereafter, the separation capillary was conditioned by flushing it for 3 min with H₂O (100 psi, forward pressure, 4.2 capillary volumes), 2 min with 1 M NaCl (100 psi, forward pressure, 2.8 capillary volumes), followed by 3 min with H₂O (100 psi, forward pressure 4.2 capillary volumes) and the conductive line for 3 min with H₂O (100 psi, reverse pressure). For the performance check the separation capillary and conductive line were filled with 100 mM ammonium acetate pH 3.0 (8 min at 100 psi forward pressure, 11.2 capillary volumes, and 4 min at 100 psi reverse pressure). The performance of the capillary was checked using the protein test mixture. Intact measurements were performed using a BGE of 3% acetic acid. For hinge region cleaved BsAbs 20% acetic acid containing 10% methanol was used as BGE. Before each run, the separation capillary was flushed for 4 min with BGE (100 psi, forward pressure, 4.2 capillary volumes) and the conductive line for 2 min with BGE (75 psi, reverse pressure). The samples were injected by applying 2.5 psi for 15 s (5.6 nL, 0.87% of the total capillary volume) followed by a plug of BGE (0.5 psi for 25 s, 1.9 nL). The separations were performed at 20 °C by applying 20 kV at reversed polarity for 45 min. After each run, the capillary was ramped down to 1 kV in 5 min.

The capillary, containing a grounded sheath metal, was connected via a nano-electrospray ionization source, which could be adjusted via a XYZ stage either to an Impact qTOF-MS or a solarix 12 T FT-ICR-MS equipped with a ParaCell (Bruker Daltonics, Bremen, Germany). The system operated in positive mode using a capillary voltage of 1200 V, a dry gas flow of 1.8 L/min and temperature of 180 °C. For the Qtof-MS, the quadrupole ion energy and collision cell voltage were set at 5.0 and 20.0 eV, respectively. Transfer and pre-pulse storage times were 120.0 and 25.0 μ s, respectively. For in-source collision-induced dissociation (isCID) 100 eV between funnel 1 and 2, was used. The monitored m/z range was 500–6000. During FT-ICR-MS analysis the trapping potentials were set to 5.5 V and the ParaCell DC biases up to 1.5 V. The Q1 mass was set to m/z 800 while the time-of-flight to the ICR cell was set to 1.2 ms. The mass spectra were acquired in an m/z -range of either 202.70-3000, 405.41-3000 or 589.68-3000 with 1 M data points (with transient times either 0.144, 1.25 or 1.78 s). The accumulation time was 0.1 s. Mass spectra deconvolution was performed using the maximum entropy algorithm in the DataAnalysis software from Bruker Daltonics (Bremen, Germany). Monoisotopic masses were determined using the SNAP option from de DataAnalysis software after deconvolution.

6.4 Results and Discussion

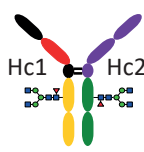
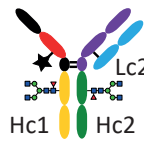
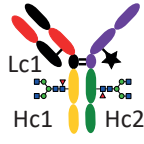
We analyzed two antibodies, namely BsAb1 and BsAb2, with the same bi-specificity and affinity but slightly different amino acid sequences due to the different engineering approaches. BsAb1 has a CH1-CL exchange in the binding site 1, a knob-into-hole in the CH3 but no additional amino acid substitutions in the CH2 portion (**Figure 1A**). BsAb2 has a CH1-CL exchange in the binding site 2, a knob-into-hole in the CH3 (with an additional exchange of 2H for 2A in the CH2 and CH3 subunits) and an L250A L251A P345G mutation in the CH2 portion introduced to “silence” the interaction with FcRs (**Figure 1A**)³¹.

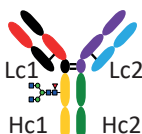
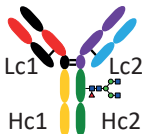
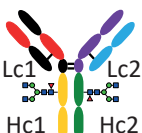
6.4.1 Intact mass measurements of BsAbs

BsAbs consist of four different chains that need to be correctly assembled during production often resulting in larger macroheterogeneity than traditional mAbs. In this study, we aimed to explore the potential of sheathless CE-MS to monitor the proper assembly of BsAbs.

Sheathless CE-MS analysis of BsAbs was performed using a positive coated capillary to avoid adsorption of the positively-charged proteins to the capillary wall. PEI was selected as it provides stable covalently-bound positive surfaces with a pH-independent anodic electroosmotic flow. This allows us to obtain efficient protein separations and provides good nanospray stability without the need of additional pressure during separation. Different BGEs were evaluated for intact BsAbs separations (**Figure S1**). Using 25 mM acetic acid, the effect of the pH on the separation and detection of BsAbs and their assemblies was evaluated. Increasing the pH from 3.0 to 6.5 did not result in the detection of additional assemblies (i.e. non-covalent assemblies) while it had a negative influence on the sensitivity due to the presence of ammonium ions which suppress the ionization. Therefore, analyses of intact BsAbs were performed at low pH. Increasing the concentration of acetic acid from 25 to 500 mM (3%) resulted in resolution of several peaks corresponding to different protein subunits and assemblies (**Figure 2A**). Good sensitivity and quality of the mass spectra were observed using these conditions (**Figure S2**). The assignments with the mass error of the different detected signals are shown in **Table 1**.

Table 1. Assignment of the species detected after sheathless CE-MS analysis of intact BsAb1.

Peak	Proteoform		Theoretical average mass (Da)	Observed average mass (Da)	Error (Da)	
5		Lc1	22172.4	22173.0	0.6	
		Lc1 (+Cys)	22292.5	22291.9	0.6	
		Lc1 (+GSH)	22478.7	22478.2	0.5	
4		Lc2	23445.8	23445.3	0.5	
		Lc2 (+Cys)	23565.9	23565.1	0.8	
		Lc2 (+GSH)	23752.1	23751.3	0.8	
		Lc1-Lc2	45618.1	45617.6	0.5	
	Lc2-Lc2	46891.5	46890.9	0.6		
3		Hc1-Hc2	(G0F/G0)*	103742.1	103747.7	3.6
			(G0F/G0F)	103888.2	103887.9	2.5
			(G0F/G1F)*	104050.4	104054.1	0.1
			(G1F/G1F) #	104212.5	104212.6	3.7
			(G1F/G2F)*	104374.7	104377.2	0.3
			(G2F/G2F)	104536.8	104533.2	5.6
2		BsAb (-Lc1) (+Cys)	(G0F-G0)*	127308.0	127307.8	0.2
			(G0F-G0F)	127454.2	127454.0	0.1
			(G0F-G1F)*	127616.3	127621.6	5.3
		BsAb (-Lc1) (+GSH)	(G0F/G0F)	127640.3	127638.0	2.3
			(G0F-G1F)*	127802.5	127796.3	6.2
			(G1F-G1F)	127616.3	127960.9	3.7
		BsAb (-Lc2) (+Cys)	(G0F/G0)*	126034.6	126029.7	4.9
			(G0F/G0F)	126180.7	126176.5	4.2
BsAb (-Lc2) (+GSH)	(G0F/G0F)	126366.9	126365.2	1.7		
	(G1F/G0F)*	126529.1	126527.0	2.1		

1		BsAb (singly glycosylated)	G0	147914.9	147918.7	3.9	
	Or		G0F	148061.0	148062.3	1.3	
			G1F	148223.1	148222.9	0.3	
			G2F	148385.3	148388.0	2.7	
		BsAb	(G0/G0)	149214.1	149216.7	2.6	
			(G0F/G0)*	149360.2	149361.2	1.0	
			(G0F/G0F)	149506.3	149505.7	0.6	
			(G0F/G1F)*	149668.5	149666.6	1.9	
			(G1F/G1F)#	149830.6	149829.1	1.6	
			(G1F/G2F)*	149992.8	149992.1	0.7	
			(G2F/G2F)	150154.9	150150.2	4.7	

The main peak (peak 1) corresponded to the intact antibody with different glycoforms, ranging from G0/G0 to G2F/G2F (Inset **Figure 2A**). The intact antibody was detected with lysine clipping (Kclip) on both heavy chains as well as with a mass difference of ~36 Da lower which suggests two pyroglutamic acid formations either on Lc1 and Hc1 coming from glutamine or on Hc2 coming from glutamic acid. However, at the intact level, we could not assign which of these three chains were modified. In peak 1, a low abundance signal (below 2% compared to the main form) was observed which corresponded to the BsAb missing one glycosylation in one of the Hc. Migrating in front of the main peak (peak 2) we detected the BsAb missing either Lc1 or Lc2 and their corresponding glycoforms. In both cases, the cysteine of the non-paired Hc was observed cysteinylated (+120.2 Da) or glutathionylated (+306.3 Da). In peak 3 we observed the correctly assembled Hc dimer missing both Lcs (Mw ~ 100 kDa). The Hc dimer was preceded by two species of approximately 45 kDa which corresponded to the Lc hetero- (Lc1-Lc2) and homodimers (Lc2-Lc2) (peak 4). Finally, in peak 5 of the electropherogram, free Lc1 and Lc2 either non-modified, cysteinylated, or glutathionylated were observed. These results illustrate the larger macroheterogeneity of BsAbs compared to standard monospecific mAbs where fewer possibilities for misassembling such as heterodimers can occur.

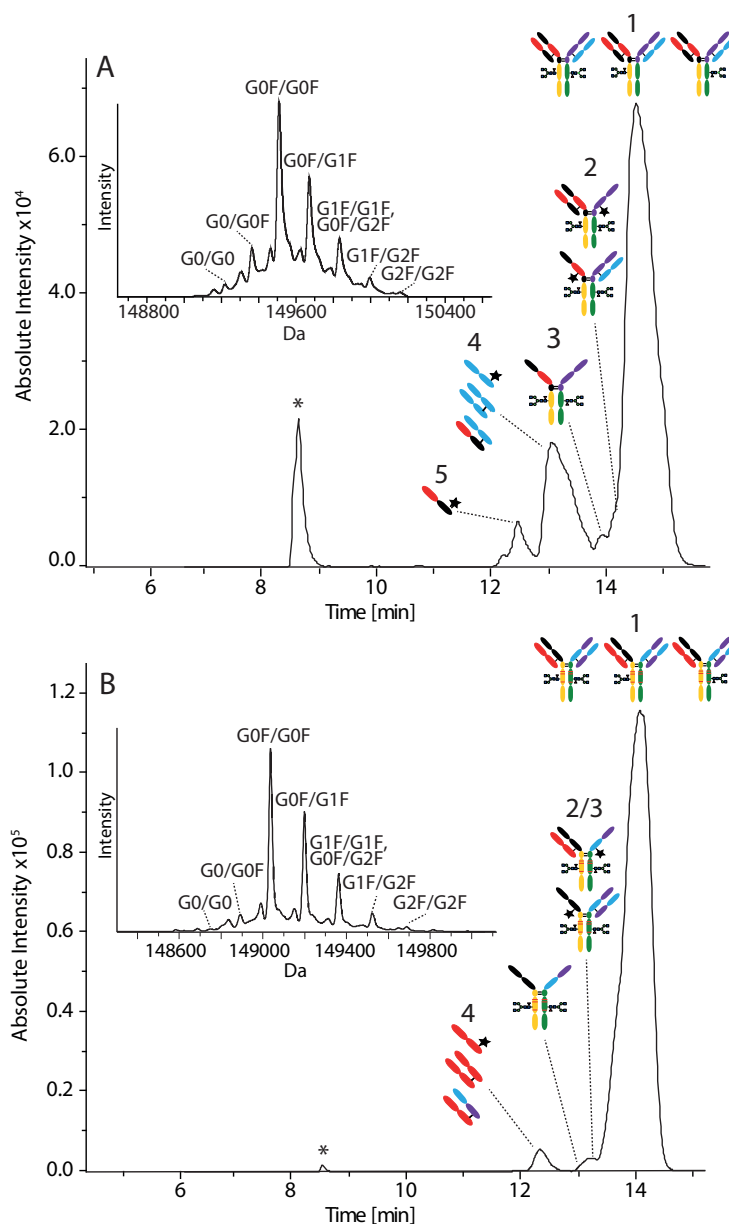


Figure 2. Sheathless CE-qTOF-MS of intact BsAbs. A) Base peak electropherogram (BPE) of BsAb1 (labeled from 1 to 5, see Table 1) and deconvoluted spectrum of peak 1 (inset). B) Base peak electropherogram (BPE) of BsAb2 (labeled from 1 to 4, see Table S1) and deconvoluted spectrum of peak 1 (inset). BGE: 10% Acetic acid. (*) non-protein peak.

Similar patterns were observed for the BsAb2 (Figure 2B and Table S1). The antibody was also detected with both C-terminal lysines clipped but only with one pyroglutamic

acid formation. In this case, the Lc1 does not contain terminal glutamine and, therefore, only the Hc1 and Hc2 can be modified. Regarding the different assemblies, although the antibody missing the crosslinked Lc (Lc2) was detected, no evidence was observed of free Lc2 in the MS. Also, the levels of Lc homo- and heterodimers observed for BsAb2 in peak 4 were lower compared to BsAb1. These results show that despite the large homology on the primary structure of these two antibodies, the different engineering processes resulted in an overall lower macroheterogeneity for BsAb2. Of note, for neither of the BsAbs evidence of mispairing (e.g. Hc1-Lc2) within chains was observed.

6.4.2 Hinge region digestion and reduction of differently engineered bispecific antibodies

Due to the presence of two binding sites and different engineering of the Fc region (knob-into-hole), BsAbs comprise six different subunits, namely Lc1 and Lc2, Fd'1 and Fd'2 and (Fc/2)1 and (Fc/2)2. These subunits, originally linked to form the typical Y-shape structure of immunoglobulin-G (IgG) mAbs, can be separated by proteolytic hinge region cleavage and reduction of disulfide bonds (**Figure 1B** and **C**). Compared to conventional IgG1-mAbs, the hinge region of BsAb1 is not altered. The enzyme IdeS exhibits a high specificity to cleave IgG1 antibodies just below the hinge region (PELLG/GPS in CH2 domain) and has been extensively used for subunit-specific characterization of conventional mAbs [26, 32-34]. Thus, IdeS was selected for the digestion of the BsAb1. Using IdeS no intact antibody was detected after a 90 min digestion indicating efficient proteolytic cleavage (**Figure S3**). BsAb2 contains an L250A L251A P345G mutation in the CH2 portion (**Figure 1A**). These amino acid substitutions are close to the cleavage site of the IdeS (PEAAG/GPS) hampering the cleavage of the antibody (no enzymatic cleavage achieved, data not shown). Therefore, a second enzyme, SpeB, was evaluated. In contrast to IdeS, SpeB cleaves above the hinge region (SCDKT/HPC) (**Figure 1C**) efficiently allowing the digestion of BsAb2. Using SpeB and digestion time of 60 min complete cleavage of BsAb2 was achieved (**Figure S3**).

After hinge region cleavage, the sample was reduced to generate the six individual subunits. reduction was performed using 100 mM TCEP or 20 mM DTT containing GdnCl as a chaotropic agent. Using TCEP the reduction of the BsAb subunits was not complete resulting in a complex mixture of different chains with different reduction levels. For Lc1 the reduced Lc1 (22164.18 Da) and the Lc1 with one internal disulfide bridge closed (22162.26 Da) were observed (theoretical mass of fully reduced Lc1 22163.90 Da) (**Figure S4**). (Fc/2)1 and (Fc/2)2 were detected either completely reduced or with one or two disulfide bridges closed while both Fds were completely reduced under these conditions. Using 20 mM DTT with 4M GdnCl as a chaotropic agent full reduction was achieved (**Figure 3**) and only completely reduced chains and their proteoforms were detected. Therefore 20 mM DTT with GdnCl was used to reduce the BsAbs after the hinge region cleavage.

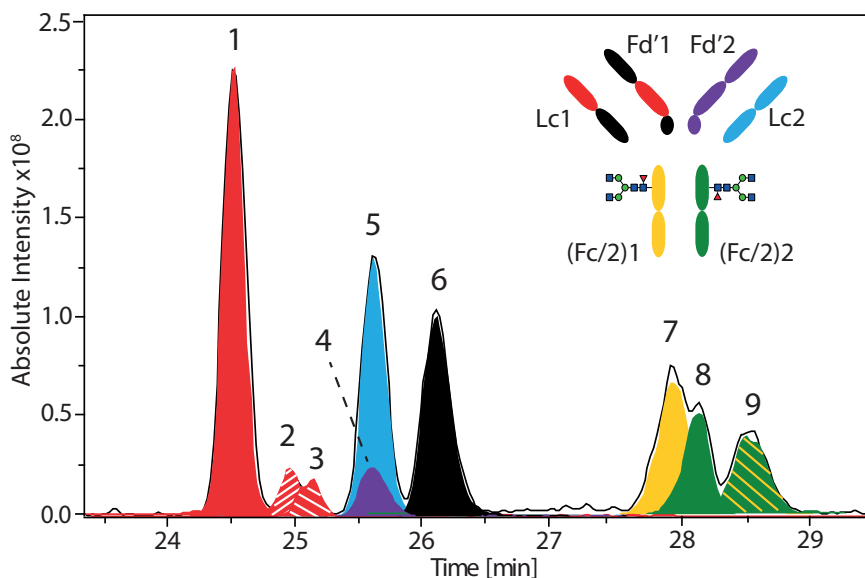


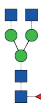
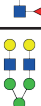
Figure 3. BPE (black trace) and EIEs (color fill) for the six chains and proteoforms of BsAb1 obtained by sheathless CE-FT-ICR-MS after IdeS cleavage and DTT reduction (labelled from 1-8, compare **Table 2**). Peak number 9 contains both Fc/2 fragments as a complex. BGE: 20% Acetic acid containing 10% MeOH. EIEs: Peak 1 ($1009.10+1057.06+1109.96\pm 0.1$ m/z); Peak 2 ($1009.78+1057.91+1110.70\pm 0.1$ m/z); Peak 3 ($999.50+1047.14+1099.50\pm 0.1$ m/z); Peak 4 ($1038.80+1082.04+1129.09\pm 0.1$ m/z); Peak 5 ($1066.94+1117.75+1173.49\pm 0.1$ m/z); Peak 6 ($1025.95+1065.41+1107.94\pm 0.1$ m/z); Peak 7 ($927.83+963.49+1001.95\pm 0.1$ m/z); Peak 8 ($938.26+974.26+1013.15\pm 0.1$ m/z); Peak 9 ($1398.96+1438.98+1481.22\pm 0.1$ m/z).

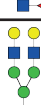
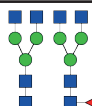
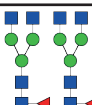
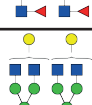
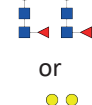
6.4.3 BsAb subunit-specific characterization by sheathless CE-FT-ICR-MS

To address with high confidence the microheterogeneity of BsAbs a subunit-specific characterization is preferred. In this work, we also study the performance of sheathless CE-MS for the characterization of BsAb subunits. The separation of the subunits was performed applying the same PEI-coated capillary than for the intact mass measurements. For the separation of the subunits, different acidic BGEs were evaluated. Using 10% acetic acid most of the chains were separated, however, the comigration of some proteoforms was observed (**Figure S5**). The addition of an organic modifier increased the viscosity of the BGE and reduced the EOF often resulting in a better separation of proteins in positively-charged coated capillaries²⁵. The addition of 10% MeOH to the BGE increased the separation window and allowed to resolve more species (**Figure S5**). A similar effect was observed with the increase of the acetic acid content from 10 to 20%, likely a consequence of the higher ionic strength. Further increase of acetic acid or MeOH content did not improve separation.

Using 20% acetic acid containing 10% MeOH all subunits were efficiently separated with the exception of Fd'2 and Lc2 which comigrated (**Figure 3**). These two subunits have very similar charge and hydrodynamic radius and, therefore, were not electrophoretically resolved (**Table S2**). For the rest of the subunits, although their hydrodynamic radius were very similar, their different number of charges allowed their CE separation (less positive proteins migrates earlier in positively-coated capillaries using reverse polarity) (**Table S2**). The formation of a Pyro-Glu on the N-termini of the Lc1 resulted in the loss of the primary amine and, therefore, in a lower electrophoretic mobility compared to the non-modified Lc. Glycosylation of the Fc/2, on the other hand, did not modify the charge of the protein and the different glycoforms comigrated in the same peak. Using 20% acetic acid containing 10% MeOH the repeatability of the method was assessed by the analysis of BsAb1 subunits after reduction and hinge-region digestion in three different days. The method showed good intraday (n=6) and interday (n=9) repeatability with RSD values of migration time between 0.5% and 1.7% (**Figure S6**).

Table 2. Assignment of the species detected by sheathless CE-MS for IdeS digested and reduced BsAb1.

Peak	Subunit	Sequence	PTM	Theoretical monoisotopic mass (Da)	Observed monoisotopic mass (Da)
1	Lc1	1-213	Pyro-Glu	22163.90	22164.04
2	Lc1	1-213	-	22180.92	22180.90
3	Lc1	3-213	truncation	21955.81	21955.79
4	Fd'2	1-242	-	25929.66	25930.01
5	Lc2	1-214	-	23436.43	23436.52
6	Fd'1	1-252	Pyro-Glu	27656.20	27656.37
7	(Fc/2)1	215-462 -Kclip		24862.32	24862.46
				25008.38	25008.51
				25170.43	25170.58
				25332.49	25332.61

8	(Fc/2)2	243-452 -Kclip		25143.44	25143.56
				25289.50	25289.62
				25451.55	25451.67
				25613.61	25613.78
9	(Fc/2)1 + (Fc/2)2	215-462 -Kclip + 243-452 -Kclip		50151.82	50151.12
				50297.88	50297.16
				50460.93	50458.23
			 or 	50621.98	50620.25

A complete subunit-specific characterization of the BsAb1 using the optimized conditions is presented in **Table 2**. To accurately assign all present proteoforms, we employed FT-ICR-MS. For the measurements, a transient time of 1.7476 s and a low m/z of 589.68 were used. **Figure S7** and **S8** show the 20+ charge state of Lc1 ($m/z \sim 1110$) with a resolving power of 145000. The used transient time permitted to isotopically resolve all the antibody subunits while providing enough spectra per peak. The intensity observed for the different subunits reflects their different ionization efficiency, with the Lc having the higher and the glycosylated Fc/2 portions the lower ionization efficiency. **Figure S9** shows the observed mass spectra for the six different subunits using CE-FR-ICR-MS, as well as the isotopic resolution obtained

for Lc1. This allowed the determination of their monoisotopic mass with an error between 2 and 13 ppm (**Table 2**). The deconvoluted mass spectra of all six subunits are shown in **Figure 4**. For both Fc/2 portions, several N-glycosylation forms (G0, G0F, G1F, and G2F) were observed being G0F the most abundant glycoform. In addition to glycosylation, other PTMs were detected within the BsAb1 subunits. Both Lc1 and Fd'1 were N-terminal pyroglutamic acid modified. Fd'2, which contains a terminal glutamic acid did not show this modification. For Lc1, both, unmodified and Pyro-Glu modified proteoforms could be detected while for Fd'1 only the Pyro-Glu form was observed. The middle-up analysis of BsAbs allowed the assignment of these modifications to each particular chain which was not possible at the intact level. However, due to the reduction step, the cysteinylated or glutathionylated free chains were not observed in the subunit-specific analysis. Peak 3 was assigned to the Lc1 with an N-terminal truncation (missing a proline and a glutamine). In addition to the different subunits and proteoforms, fragments with masses between 50151.12 and 50620.25 Da were observed (peak 9, **Figure 3**). These masses correlated well with the calculated mass for the two reduced Fc/2 subunits (Fc1/2+Fc2/2) (theoretical mass G0/G0F, 50151.82 Da; observed mass, 50151.12 Da). IdeS cleaves below the hinge region of the antibody and, therefore, these subunits may only be associated by non-covalent interactions. This type of associations has been previously described in CE-MS for IdeS-digested mAbs when acetic acid BGEs with concentrations below 10% were employed²⁶. Higher concentrations or addition of organic solvents provided full dissociation of the Fc complex in the studied mAbs, resulting in only Fc/2 fragments in the electropherogram²⁶. As mentioned earlier, the Fc portions of BsAb1 are modified (i.e. knob-into-hole) to ensure the correct assembly of the two different heavy chains. We hypothesize that the strong interaction between the introduced amino acids (knob-into-hole) kept part of the two engineered Fc/2 subunits together (or led to a fast reassembly) even in presence of BGEs with high acetic acid and MeOH content.

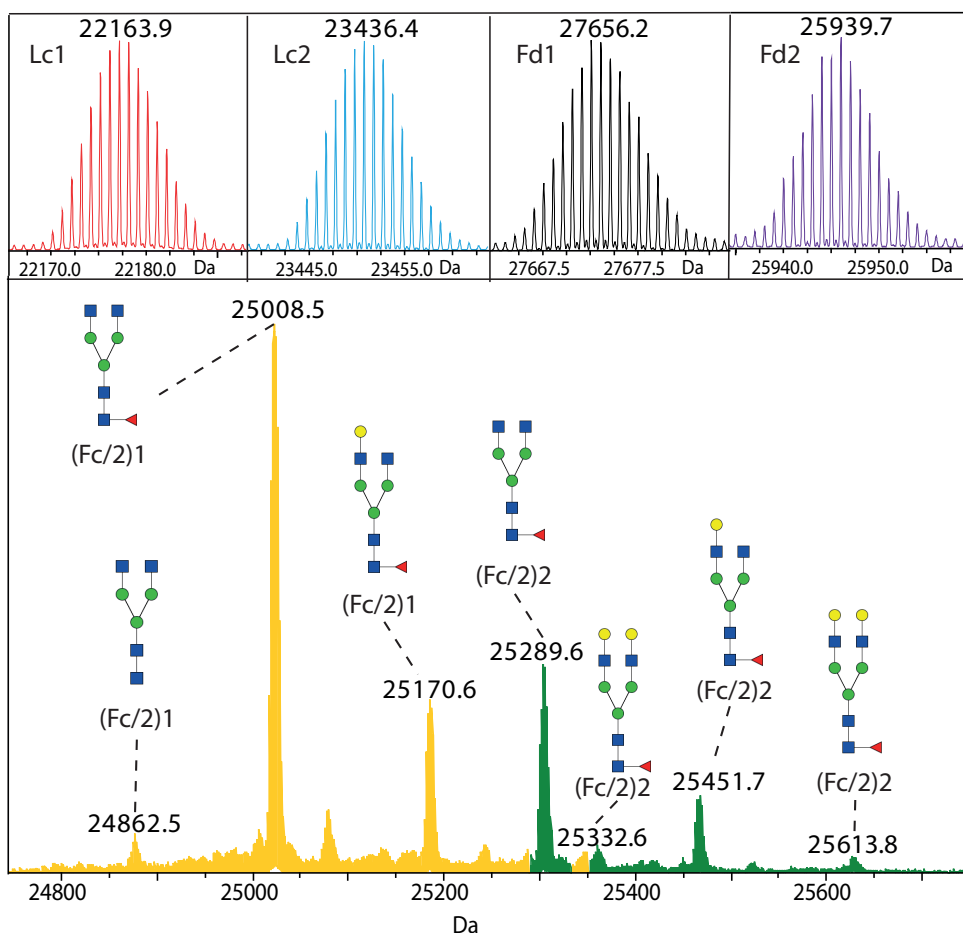


Figure 4. Deconvoluted mass spectra of the Lc1 (red), Lc2 (blue), Fd'1 (black) and Fd'2 (purple) obtained after sheathless CE-FT-ICR-MS of mAb1. Additional Fc/2 fragments ((Fc/2)1 in yellow and (Fc/2)2 in green) are shown with different glycoforms (G0, G0F, G1F, G2F). BGE: 20% Acetic acid containing 10% MeOH.

The same conditions were applied for the analysis of the bispecific BsAb2 obtained after SpeB digestion and reduction (**Figure 5**). Interestingly, the separation profile of these two antibodies was quite different. The migration window of the six subunits of BsAb2 was shorter, only 3.5 min compared to 5 min for BsAb1, and the migration pattern changed. In this case, comigration of Fd'2 and Lc2 was not observed. The alteration in the migration pattern can be explained by the different cleavage site of SpeB together with the exchange of certain amino acids of BsAb2 with respect BsAb1. This change in the primary structure was reflected in a different charge of the analyzed subunits and, therefore, in their electrophoretic mobility (**Tables S2 and S3**).

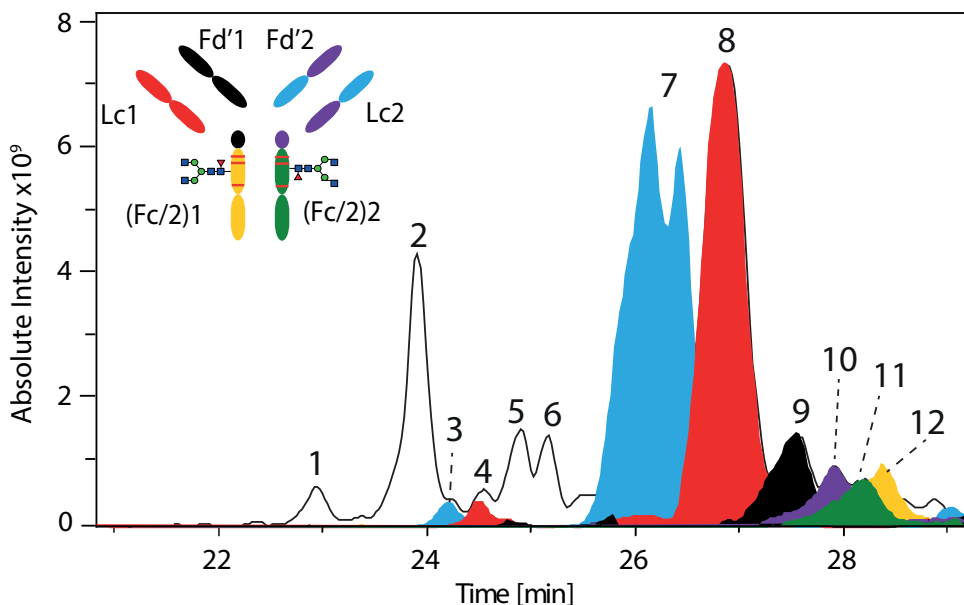


Figure 5. BPE (black trace) and EIEs (color fill) for all the six subunits (7-12, compare Table S4) of BsAb2 obtained by sheathless CE-FT-ICR-MS after SpeB cleavage and DTT reduction. Truncated versions of BsAb2 are numbered from 1-6 (compare Table S5). BGE: 20% Acetic acid containing 10% MeOH. EIEs: Peak 7 (1012.15+1060.11+1113.12±0.1 m/z); Peak 8 (1165.59+1116.86+1171.95±0.1 m/z); Peak 9 (1040.36+998.21+084.97±0.1 m/z); Peak 10 (917.61+950.83+986.08±0.1 m/z); Peak 11 (922.47+956.14+993.47±0.1 m/z); Peak 12 (1004.28+1044.41+967.16±0.1 m/z).

A total of 36 proteoforms including truncated forms could be detected (**Table S4 and S5**). In this case, Fd'2 was detected with a Pyro-Glu instead of Fd'1. The rest of the subunits did not carry N-terminal glutamine or glutamic acid. For both Lcs a signal with additional +162.05 Da was observed which was assigned to their glycosylated forms. This modification was not observed at the intact level as the mass shift overlapped with different galactosylation levels of the glycoforms. This illustrates once more the complementarity of both approaches. Furthermore, the Fd's were detected without C-terminal Thr. For Fd'2 also a form without His and Thr was found. These specific truncations have been previously reported and assigned to unspecific cleavages of the enzyme in the hinge region with SpeB³⁵. In addition, several truncated versions of Lc1, Lc2 and Fd'1 with masses between 5569.72 Da and 24828.13 Da were detected. These masses were not observed when the intact BsAb2 was analyzed suggesting that they may be the result of unspecific cleavages by endoprotease SpeB. Still, reliable subunit-specific characterization of BsAb2 was possible permitting to assign modifications such as pyro-Glu to specific chains and to determine levels of glycation in the antibody.

6.5 Conclusion

This manuscript shows, for the first time, the potential of sheathless CE-MS for the characterization of BsAbs. A combination of intact and subunit-specific analyses permitted the assessment of their macro- and microheterogeneity. The developed method was successfully applied to characterize two highly homologous BsAbs produced using common engineering techniques. By analyzing the intact antibodies several incomplete assemblies and free chains were detected. The levels of these assemblies were different between the BsAb illustrating that different engineering processes can result in different macroheterogeneity. By performing a subunit-specific analysis, particular PTMs such as glycation could be determined and assigned to each specific subunit. For BsAb1 with an unaltered hinge (compared to conventional mAbs) the well-accepted enzyme IdeS provided effective digestion while for BsAb2 containing the common LALA PG mutation an alternative enzyme, SpeB, permitted the characterization of the antibody subunits. Both, intact BsAbs and BsAb-subunits could be analyzed using a PEI-coated capillary and an acidic BGE showing large applicability. The followed strategy has potential for the characterization of new BsAbs formats of different sizes such as Fab-elongated BsAbs or other Fc fusion protein.

6.6 Supporting information

Supplementary information is available free of charge on the Analytica Chimica Acta website:
DOI: 10.1016/j.aca.2020.07.069

6.7 Acknowledgments

This work was supported by the Analytics for Biologics project (Grant agreement ID 765502) of the European Commission and the Netherlands Organization for Scientific Research NWO (SATIN project, Grant No. 731.017.202).

6.8 References

1. Schmid A. S., Neri D. Advances in antibody engineering for rheumatic diseases. *Nat Rev Rheumatol* 2019, 15, 197-207.
2. Tiller K. E., Tessier P. M., Advances in Antibody Design. *Annu Rev Biomed Eng* 2015, 17, 191-216.
3. Czajkowsky D. M., *et al.* Fc-fusion proteins: new developments and future perspectives. *EMBO Mol Med* 2012, 4, 1015-1028.
4. Sedykh S. E., *et al.* Bispecific antibodies: design, therapy, perspectives. *Drug Des Dev Ther* 2018, 12, 195-208.
5. Knight T., Callaghan M. U. The role of emicizumab, a bispecific factor IXa- and factor

- X-directed antibody, for the prevention of bleeding episodes in patients with hemophilia A. *Ther Adv Hematol* 2018, 9, 319-334.
6. Brinkmann U., Kontermann R. E. The making of bispecific antibodies. *MAbs* 2017, 9, 182-212.
 7. Klein C., Schaefer W., Regula J. T. The use of CrossMAb technology for the generation of bi- and multispecific antibodies. *MAbs* 2016, 8, 1010-1020.
 8. Ignjatovic J., *et al.* Aggregation of Recombinant Monoclonal Antibodies and Its Role in Potential Immunogenicity. *Curr Pharm Biotechnol* 2018, 19, 343-356.
 9. Abes R., Teillaud J. L. Impact of Glycosylation on Effector Functions of Therapeutic IgG. *Pharmaceuticals (Basel)*, 2010, 3, 146-157.
 10. Wang W. R., *et al.* Impact of methionine oxidation in human IgG1 Fc on serum half-life of monoclonal antibodies. *Mol Immunol* 2011, 48, 860-866.
 11. Wei Z. P., *et al.* Identification of a single tryptophan residue as critical for binding activity in a humanized monoclonal antibody against respiratory syncytial virus. *Anal Chem* 2007, 79, 2797-2805.
 12. Zhang H., Cui W. D., Gross M. L. Mass spectrometry for the biophysical characterization of therapeutic monoclonal antibodies. *Febs Lett* 2014, 588, 308-317.
 13. Haberman M., *et al.* Rapid characterization of biotherapeutic proteins by size-exclusion chromatography coupled to native mass spectrometry. *MAbs* 2016, 8, 331-339.
 14. Duivelshof B. L., *et al.* A generic workflow for the characterization of therapeutic monoclonal antibodies—application to daratumumab. *Anal and Bioanal Chem* 2019, 411, 4615-4627.
 15. Tassi M., *et al.* Advances in native high-performance liquid chromatography and intact mass spectrometry for the characterization of biopharmaceutical products. *J Sep Sci* 2018, 41, 125-144.
 16. Camperi J., Pichon V., Delaunay N. Separation methods hyphenated to mass spectrometry for the characterization of the protein glycosylation at the intact level. *J Pharm Biomed Anal* 2020, 178, 112921.
 17. Resemann A., *et al.* Full validation of therapeutic antibody sequences by middle-up mass measurements and middle-down protein sequencing. *MAbs* 2016, 8, 318-330.
 18. Turner A., *et al.* Development of orthogonal NISTmAb size heterogeneity control methods. *Anal and Bioanal Chem* 2018, 410, 2095-2110.
 19. Woods R. J., *et al.* LC-MS characterization and purity assessment of a prototype bispecific antibody. *MAbs* 2013, 5, 711-722.
 20. Liu H., Gaza-Bulseco G., Chumsae C. Analysis of Reduced Monoclonal Antibodies Using Size Exclusion Chromatography Coupled with Mass Spectrometry. *J Am Soc Mass Spectrom* 2009, 20, 2258-2264.
 21. Ding W., *et al.* Improving Mass Spectral Quality of Monoclonal Antibody Middle-Up LC-MS Analysis by Shifting the Protein Charge State Distribution. *Anal Chem* 2018, 90, 1560-1565.

22. Sokolowska I., *et al.* Subunit mass analysis for monitoring antibody oxidation. *MABs* 2017, 9, 498-505.
23. Wang C., *et al.* A systematic approach for analysis and characterization of mispairing in bispecific antibodies with asymmetric architecture. *MABs* 2018, 10, 1226-1235.
24. D'Atri V., *et al.* Hydrophilic Interaction Chromatography Hyphenated with Mass Spectrometry: A Powerful Analytical Tool for the Comparison of Originator and Biosimilar Therapeutic Monoclonal Antibodies at the Middle-up Level of Analysis. *Anal Chem* 2017, 89, 2086-2092.
25. Belov A. M., *et al.* Complementary middle-down and intact monoclonal antibody proteoform characterization by capillary zone electrophoresis - mass spectrometry. *Electrophoresis* 2018, 39, 2069-2082.
26. Haselberg R., *et al.* Heterogeneity assessment of antibody-derived therapeutics at the intact and middle-up level by low-flow sheathless capillary electrophoresis-mass spectrometry. *Anal Chim Acta* 2018, 1044, 181-190.
27. Giorgetti J., *et al.* Intact monoclonal antibodies separation and analysis by sheathless capillary electrophoresis-mass spectrometry. *Eur J Mass Spectrom* 2019, 25, 324-332.
28. Chen C. H., *et al.* Intact NIST monoclonal antibody characterization-Proteoforms, glycoforms-Using CE-MS and CE-LIF. *Cogent Chem*, 2018, 4, 1480455.
29. Giorgetti J., *et al.* Combination of intact, middle-up and bottom-up levels to characterize 7 therapeutic monoclonal antibodies by capillary electrophoresis - Mass spectrometry. *J Pharm Biomed Anal*, 2020, 182, 113107.
30. Marcia R Santos C. K. R., Fonslow B., Guttman A. A covalent, cationic polymer coating method for the CESI-MS analysis of intact proteins and polypeptides. 2015.
31. Schlothauer T., *et al.* Novel human IgG1 and IgG4 Fc-engineered antibodies with completely abolished immune effector functions. *Protein Eng Des Sel* 2016, 29, 457-466.
32. Michalikova K., *et al.* Middle-Up Characterization of the Monoclonal Antibody Infliximab by Capillary Zone Electrophoresis-Mass Spectrometry. *Lc Gc Europe* 2019, 32, 130-137.
33. Faid V., *et al.* Middle-up analysis of monoclonal antibodies after combined IgdE and IdeS hinge proteolysis: Investigation of free sulfhydryls. *J Pharm Biomed Anal* 2018, 149, 541-546.
34. Gstottner C., *et al.* Monitoring glycation levels of a bispecific monoclonal antibody at subunit level by ultrahigh-resolution MALDI FT-ICR mass spectrometry. *MABs* 2020, 12, 1682403.
35. Lippold S., *et al.* Proteoform-Resolved FcγRIIIa Binding Assay for Fab Glycosylated Monoclonal Antibodies Achieved by Affinity Chromatography Mass Spectrometry of Fc Moieties. *Front Chem* 2019, 7.



Chapter 7

Sheathless CE-MS as a tool for monitoring exchange efficiency and stability of bispecific antibodies

Christoph Gstöttner¹, Dana L.E.Vergoossen², Manfred Wuhrer¹, Maartje G.M.Huijbers^{2,3}
and Elena Domínguez-Vega¹

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

²Leiden University Medical Center, Department of Human Genetics, Leiden, The Netherlands

³Leiden University Medical Center, Department of Neurology, Leiden, The Netherlands

Reprinted and adapted with permission from Electrophoresis, 2021, 42, 171-176, DOI:
10.1002/elps.202000166

© 2020 Wiley-VCH GmbH



7.1 Abstract

Bispecific monoclonal antibodies (BsAbs) are receiving great attention due to their extensive benefits as biopharmaceuticals and their involvement in IgG4 mediated autoimmune diseases. While the production of BsAbs is getting more accessible, their analytical characterization remains challenging. We explored the potential of sheathless CE-MS for monitoring exchange efficiency and stability of in-house produced bispecific antibodies. Two IgG4 bispecific antibodies with different molecular characteristics were prepared using controlled Fab (Fragment antigen binding) -arm exchange. Separation of BsAbs from their parent monospecific antibodies was achieved using a PEI-coated capillary and acidic background electrolytes permitting reliable assessment of the exchange efficiency. This was especially valuable for a Fab-glycosylated BsAb where the high glycan heterogeneity resulted in an overlap of masses with the monospecific parent antibody, hindering their discrimination by MS only. The method showed also good capabilities to monitor the stability of the generated BsAbs under different storage conditions. The levels of degradation products were different for the studied antibodies indicating pronounced differences in stability. Overall, the proposed method represents a useful analytical tool for exchange efficiency and stability studies of bispecific antibodies.

7.2 Main

Current advances in antibody engineering have provided effective tools for the production of bispecific antibodies (BsAbs) [1]. The dual-targeting capacity of BsAbs has been especially beneficial for therapeutic purposes permitting *e.g.* redirecting of immune effector cells to tumor cells, targeting tumor angiogenesis or blocking signalling pathways [2]. Next to the therapeutic intervention, the use of BsAbs is highly relevant in many other clinical areas [3]. For instance BsAbs have been exploited for diagnostic purposes due to their capacity to bind simultaneously to a specific antigen and a detection moiety [3]. BsAbs also occur *in vivo* as human and cyno IgG4 antibodies are uniquely capable of exchanging half-antibodies with other IgG4 molecules through a stochastic process called Fab (Fragment antigen binding) -arm exchange [4]. Due to the large heterogeneity in circulating IgG4, it is estimated that ~99% of IgG4 molecules are bispecific in humans [5]. The formation of bispecific antibodies alters their effector functions as they for example cannot crosslink antigens nor form immune complexes. Bispecific functionally monovalent antibodies have furthermore been associated with (IgG4-mediated) autoimmune diseases and in some cases may increase the pathogenicity of autoantibodies [5-7]. Therefore, well-characterized BsAbs are important tools to facilitate their application as therapeutics, diagnostic tools or to study their role in health and (autoimmune) diseases.

Production of BsAbs has considerably improved over the last years and several ways of production, including quadroma technology, chemical conjugation and genetic approaches have been developed [1,2]. In particular, recombination of antigen-binding arms between individually expressed monospecific antibodies containing two matching point mutations (known as controlled Fab-arm exchange), has permitted generation of BsAbs not only in specialized production plants but also in different academic research labs [8].

The successful production of BsAbs has to be monitored using proper analytical techniques to confirm efficient heterodimer formation and assess occurring modifications and potential degradation products. In particular, presence of some monospecific, homodimeric antibody versions in BsAb preparations may confound *in vivo* and *in vitro* experiments and represents a risk factor for eventual therapeutic applications. Current methods for assessing exchange efficiency in the formation of BsAbs mainly focus on separation techniques such as cation-exchange chromatography (CEX) and hydrophobic interaction chromatography (HIC) or the use of mass spectrometry. However, when parent monospecific antibodies feature similar molecular properties (*e.g.* mass, charge variant distribution or hydrophobicity) the assessment of the levels of heterodimer formation versus homodimer becomes challenging. Capillary electrophoresis is a separation technique that has proven high potential for the characterization of mAbs, providing resolution based on charge, mass and also shape of the protein [9-11]. Furthermore, in contrast to conventional CEX and HIC using salt buffers, CE permits direct hyphenation with MS providing confident confirmation of the identity



of the separated species. Especially, when combined with nanoelectrospray ionization via sheathless CE-MS higher ionization efficiency and sensitivity can be obtained compared to sheath-liquid approaches, which is fundamental to determine low abundant species [12].

In this study we explored for the first time the capabilities of CE-MS to assess the heterodimer formation versus homodimer as well as degradation products of BsAbs. Two IgG4 BsAbs with different molecular properties were included in the study. To this end, monospecific IgG4 antibodies mAb1 and mAb2 were independently combined with a control antibody (mAb-Ct) resulting in bispecific functionally monovalent antibodies BsAb1 and BsAb2. The BsAbs were prepared in-house based on the protocol described by Labrijn et al. [8], but adapted for IgG4 antibodies by Vergoossen et al. [7]. Mutated monospecific IgG4 antibodies (S228P) were expressed in HEK293 cells and CHO cells. mAb-Ct contained, in addition to the S228P mutation, two matching point mutations at the CH3 domain (F405L and R409K) and was added in 30% excess compared with mAb1 and mAb2 to obtain a high exchange efficiency of the BsAb1 and BsAb2 and minimize contamination with the monospecific parental clone.

BsAb1 as well as the corresponding monospecific versions were desalted in Milli-Q water using Vivaspin filters (GE Healthcare, Munich, Germany) and analysed under acidic conditions using a BGE consisting of 10% acetic acid (pH 2.1) (acetic acid from VWR Chemicals, Radnor, PA). CE separations were performed in a CESI 8000 system (Sciex, Brea, CA). Bare fused silica capillaries with a porous tip (91 cm x 30 µm i.d) were employed (OptiMS, Sciex, Brea, CA). To avoid adsorption of the positively charged antibodies to the wall of the capillary, the surface was coated in-house with polyethylenimine (PEI) (Gelest, Morrisville, PA). The PEI coating was performed and conditioned following the protocol described by the manufacturer [13]. At the beginning of each run the separation capillary was flushed for 4 min with BGE (100 psi forward pressure) and the reverse line for 2 min (75 psi reverse pressure). The antibodies (1 µg/µL) were injected hydrodynamically by applying 2.5 psi for 15 s. Separation was performed using a reversed polarity of 20 kV at 20°C for 45 min. After each run the capillary was ramped down from 20 kV to 1 kV with 50 psi forward and reverse pressure in 5 min. To obtain good ionization efficiency and sensitivity for low abundance species CE was hyphenated with nano-ESI-MS via a sheathless interface (Sciex) to a Qtof-MS (Bruker Daltonics, Bremen, Germany). Detailed detection parameters are described elsewhere [14]. **Figure 1** shows the electropherograms obtained for the BsAb1 as well as the parent antibodies mAb1 and mAb-Ct. Both monospecific antibodies were baseline separated as consequence of their different overall charge (pI 6.5 for mAb1 and 8.5 for mAb-Ct calculated via Protipi.ch). The BsAb1 heterodimer formed of ½ mAb1 and ½ mAb-Ct migrated between them reflecting its hybrid nature. BsAb preparations contained around 30% excess of mAb-Ct which was clearly detected in the electropherogram (H3N4F1/H3N4F1 mAb detected at 150236.8 Da; calculated mass 150235.8 Da). Regarding the parent mAb1 only a minimal signal (~0.8%; H3N4F1/H3N4F1 mAb detected at 146137.7 Da; calculated mass 146139.9 Da) was detected indicating a high exchange efficiency of BsAb1 formation. The deconvoluted

mass spectra of the main peak revealed a mass of 148188.9 Da which is consistent with the theoretical mass calculated for the H3N4F1/H3N4F1 glycoform of BsAb1 (148187.9 Da) excluding the formation of artefacts in the protein during the exchange procedure (**Figure 1**).

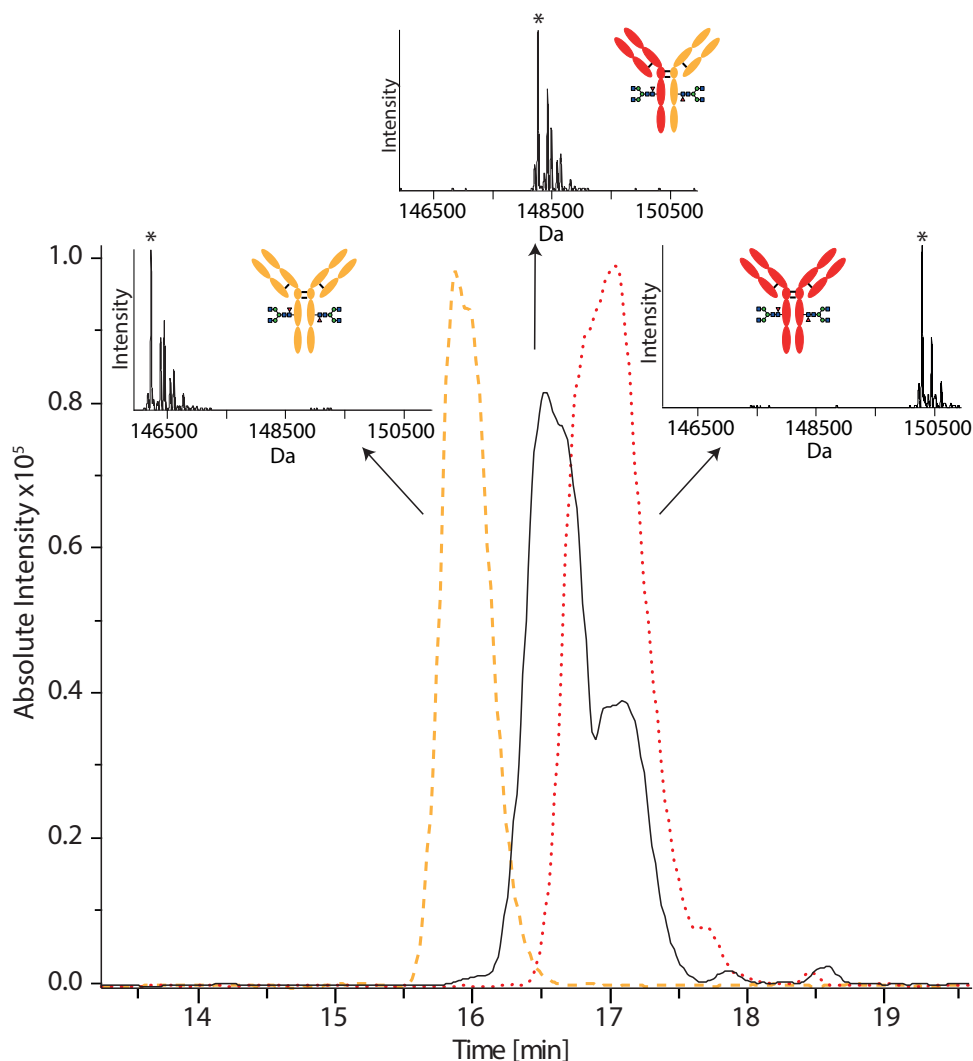


Figure 1. Base peak electropherogram (BPE) obtained after analysis of BsAb1 (solid line), mAb1 (dashed line) and mAb-Ct (dotted line). For each peak the deconvoluted mass spectra is depicted. Signals marked with * represent the H3N4F1/H3N4F1 Fc glycoform. Separation conditions: PEI-coated capillary; BGE, 10% AA; -20 kV, 20 °C.

BsAb2 was generated using a Fab glycosylated antibody mAb2 and mAb-Ct. In this case, some of the glycoforms of the parental and the bispecific antibody have very similar molecular masses. The Fab glycosylation of mAb2 introduced additional heterogeneity

resulting in a plethora of glycoforms. This complexity results in an overlap of signals in the deconvoluted mass spectrum when acquired by stand-alone MS hampering interpretation (**Figure S1**). For instance, for mAb2 a signal at 150606.2 Da was observed which corresponds to a glycan composition of H3N4F1/H3N4F1 for the Fc portion and H3N4/H3N5F1 (or H3N5/H3N4F1) for the Fab portion (calculated mass 150603.1 Da). This mass, however overlaps with another glycoform from BsAb2 with the glycan composition of H3N4F1/H3N4F1 for the Fc portion and H4N5F1 for the Fab detected at 150611.7 Da (calculated mass 150611.3 Da) not permitting to assess the presence of the mAb2 in the bispecific preparation. In these cases, a good separation of all the species proved to be crucial for assessing exchange efficiency. Using a BGE consisting of 10% acetic acid, overlap of the different species was observed in sheathless CE-MS (**Figure S2**). This is most probably a consequence of the highly similar calculated pI for mAb-Ct (8.5) and mAb2 (7.9). To increase resolution between the antibodies, different BGEs and separation temperatures were evaluated. The addition of MeOH (Actu-All Chemicals, Oss, The Netherlands) to the BGE (10 and 20%) resulted in a partial separation of the different species due to the increase on viscosity and consequent decrease on the EOF. Increasing the concentration of acetic acid (10-30%) further improved the resolution between peaks. Finally, using a BGE consisting of 30% acetic acid and 10% methanol, the decrease in the separation temperature from 20 to 15 °C allowed the baseline separation of both parental antibodies (**Figure 2**). This permitted to determine the identity of the separated antibodies with high confidence. Due to the Fab glycosylation, mAb2 showed a relatively lower ionization efficiency in the mass spectrometric measurement than mAb-Ct. This resulted in a similar signal for mAb-Ct (~30%) compared to BsAb2 (~70%) (**Figure 2**). Therefore, the amount of remaining mAb2 (~1 %) was estimated by spiking BsAb2 with mAb2. The addition of 0.85% of mAb2 was still detectable in the MS indicating that we can detect mAb2 amounts below 0.5%.

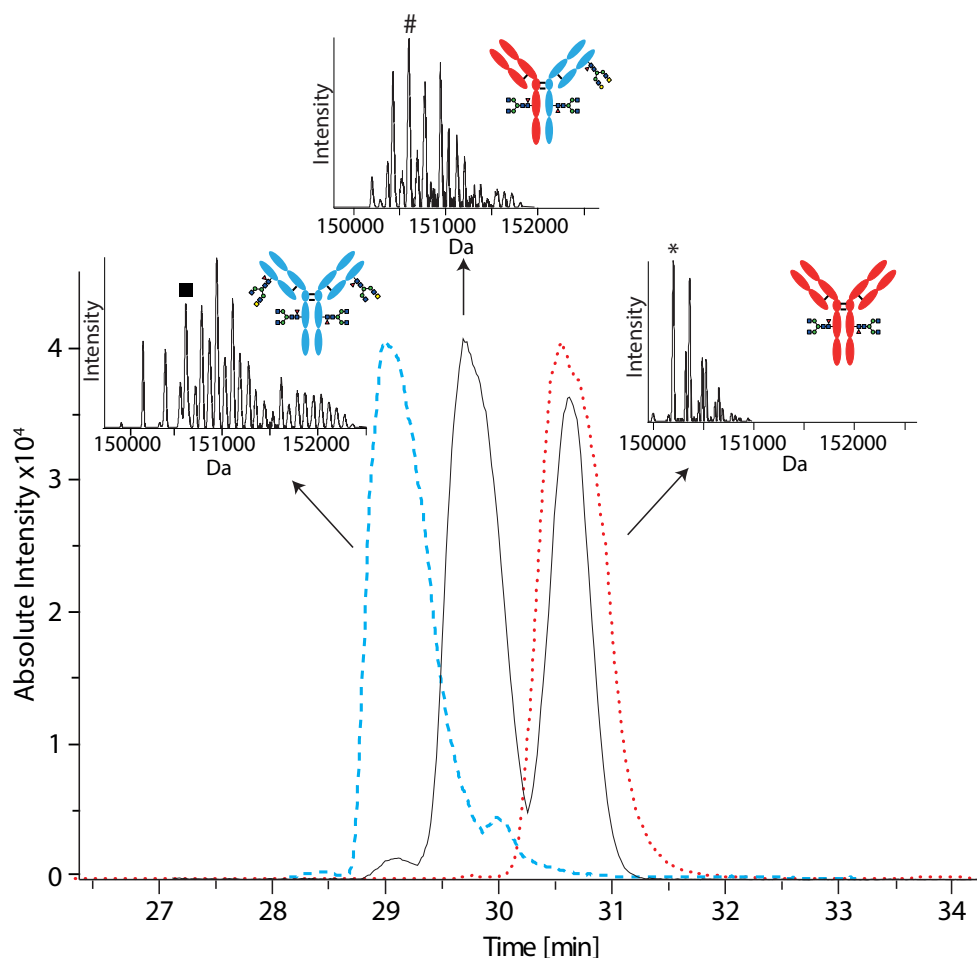


Figure 2. Base peak electropherogram (BPE) obtained after analysis of BsAb2 (solid line), mAb2 (dashed lines) and mAb-Ct (dotted line). For each peak the deconvoluted mass spectra is depicted. Signal marked with * represent the H3N4F1/H3N4F1 Fc glycoform without Fab glycosylation. Signal marked with # carry next to the Fc glycosylation one H3N5F1 glycan on one Lc. Signal with carry next to the Fc glycosylation additional H3N5F1 glycans on both Lcs. Separation conditions: PEI-coated capillary; BGE, 30% AA + 10% MeOH; -20 kV, 15 °C.

Stability of these molecules is another key aspect in BsAb characterization. Degradation products such as antibodies missing a light chain, half antibodies or light chain dimers are often reported for these type of molecules [15]. Therefore, we also explored the ability of sheathless CE-MS to monitor the stability of BsAbs. To this end, antibodies were stored before and after exchange for different periods of time (0-6 months) at 4 °C or at -20 °C in 1 x PBS solution. Before analysis the samples were desalted in Milli-Q water using Vivaspinn filters. For BsAb1, the analysis of the reference antibody already revealed low abundance

signals, at a migration time of 20 min. The masses of these signals were around 50 kDa and were assigned to various Fab fragments of mAb-Ct of different length (**Table S1**). No evidence of degradation products coming from mAb1 nor BsAb1 were observed in the freshly prepared samples. Storage of the sample for 2 months resulted in an increase in the levels of mAb-Ct-Fab fragments. The increase was more notable in the sample stored at 4 °C than in the one stored at -20 °C and got more pronounced after 6 months of storage (**Figure S3**). For BsAb2 the same type of Fab fragments coming from mAb-Ct were observed in the reference sample (not stored) (**Table S1**). In this case, the use of a different BGE (30% acetic acid + 10% MeOH) resulted in comigration of Fab fragments with the intact mAb-Ct. Still, it was possible to separately detect these species in the mass spectra and to determine their relative amount (**Figure S4**). In contrast to BsAb1, the amount of Fab fragments did not seem to increase during the storage of the sample. The same experiments were performed with the monospecific versions of the antibodies. For mAb-Ct, the same Fab fragments were observed in the reference sample. In line with the observations made for BsAb2 the storage conditions did not result in a noticeable increase of the amount of Fab fragments (**Figure S4**). For the parent molecules mAb1 and mAb2 we could not detect any degradation products in the reference standard, nor in the stored samples (data not shown). We hypothesize that the Fab fragments in the BsAb2 sample come from the remaining mAb-Ct (~30%) whereas in the BsAb1 they could also proceed from the degradation of the BsAb, which could indicate a lower stability of this BsAb. However, we could not find antibodies missing the Fab portion to confirm this theory. For BsAb1 and BsAb2 no evidence of further degradation products (*i.e.* additional PTMs or fragments) were observed suggesting that both BsAbs are stable over time. In vivo stability was also addressed by injecting the BsAb in mice and analyzing the circulating antibodies after 11 days by CE-MS (data available in [7]). No signs of additional species or degradation products were observed nor rearrangements of antibody arms indicating stability of the BsAbs in vivo.

To summarize, in this work we studied the potential of sheathless CE-MS to monitor exchange efficiency and stability of in-house produced BsAbs. The use of a positively coated capillary in combination with an acidic BGE provided adequate peak efficiency and quality of the mass spectra for the intact antibodies. BsAbs were separated from the parents monospecific antibody versions permitting reliable determination of the exchange efficiency even for BsAb2 where masses of the different species overlapped. The methods also allowed to monitor degradation products after prolonged storage and in vivo experiments. The method may exhibit limitations for the assessment of the exchange efficiency of BsAb where the parent mAbs exhibit similar calculated pIs but could be still valuable for their stability assessment. Although here this method has been applied to IgG4 monovalent BsAbs, it should be translatable to a variety of BsAb including bivalent BsAbs and BsAb from other subclasses often employed with therapeutic purposes (*i.e.* IgG1).

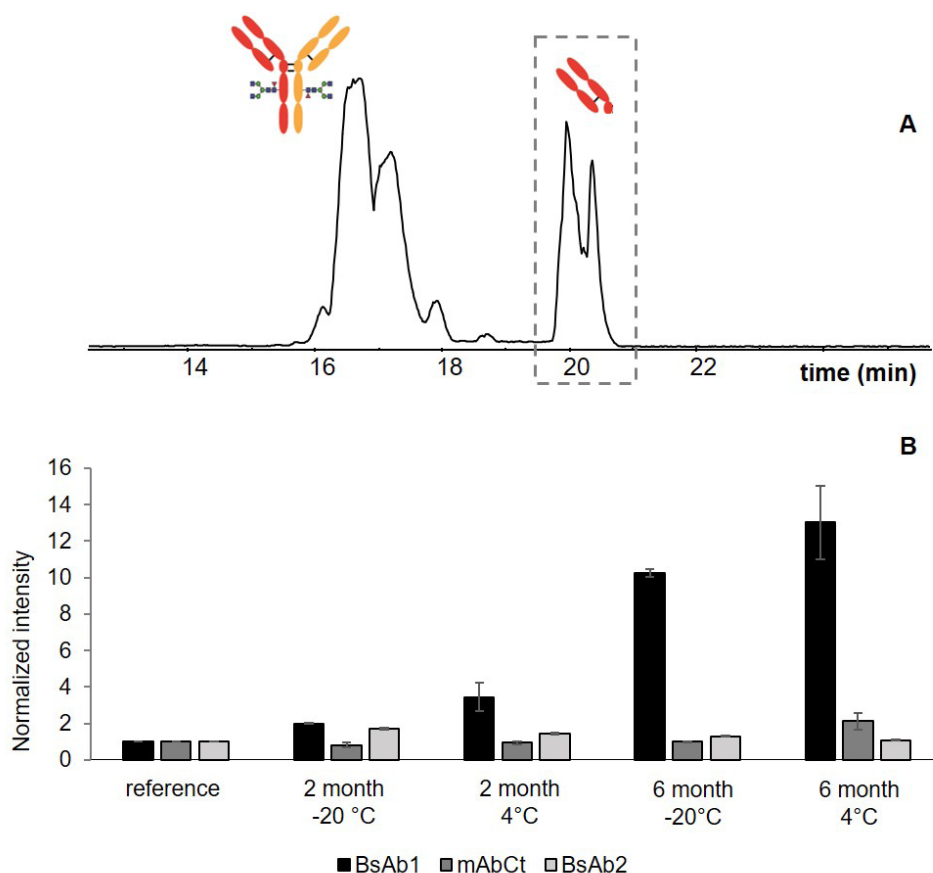


Figure 3. Base peak electropherogram (BPE) obtained after analysis of BsAb1 stored for 6 months at 4 °C (A). Normalized intensity obtained for the extraction of mAb-Ct-Fab (LC+Fd' (1-240)) stored for different time and temperature (B).

7.3 Supporting information

Supplementary information is available free of charge on the Electrophoresis website: DOI: 10.1002/elps.202000166

7.4 Acknowledgments

This work was supported by the Analytics for Biologics project (Grant agreement ID 765502) of the European Commission. MH and DV are members of the European Reference Network for Rare Neuromuscular Diseases [ERN EURO-NMD]. MH receives financial support from the LUMC (OIO, 2017), Health Holland/TKI Target to B consortium, Prinses Beatrix Spierfonds (W.OR-17.13 and W. OR-19.13) and the Dutch Science Organization (VENI 0915016181

0040). MH receive royalties from licensed patent applications on MuSK-related research.

7.5 References

1. Labrijn A. F., *et al.* Bispecific antibodies: a mechanistic review of the pipeline. *Nat Rev Drug Discov* 2019, 18, 585-608.
2. Fan G., *et al.* Bispecific antibodies and their applications. *J. Hematol. & Oncol.* 2015, 8, 130.
3. Byrne H., *et al.* A tale of two specificities: bispecific antibodies for therapeutic and diagnostic applications. *Trends Biotechnol* 2013, 31, 621-632.
4. Rispen T., *et al.* Mechanism of Immunoglobulin G4 Fab-arm Exchange. *J Am Chem Soc* 2011, 133, 10302-10311.
5. Konecny I., *et al.* IgG4 autoantibodies against muscle-specific kinase undergo Fab-arm exchange in myasthenia gravis patients. *J Autoimmun* 2017, 77, 104-115.
6. Huijbers M. G., *et al.* MuSK myasthenia gravis monoclonal antibodies: Valency dictates pathogenicity. *Neurol - Neuroimmunol* 2019, 6, e547.
7. Vergoossen D. L. E., *et al.* Functional monovalency amplifies the pathogenicity of anti-MuSK IgG4 in myasthenia gravis. *Proc Natl Acad Sci USA* 2021, 118, e2020635118.
8. Labrijn A. F., *et al.* Controlled Fab-arm exchange for the generation of stable bispecific IgG1. *Nat Protoc* 2014, 9, 2450-2463.
9. Domínguez-Vega E., *et al.* Capillary Electrophoresis of Proteins and Peptides. *Methods and Protocols*, Springer New York 2016, 25-41.
10. Haselberg R., *et al.* Heterogeneity assessment of antibody-derived therapeutics at the intact and middle-up level by low-flow sheathless capillary electrophoresis-mass spectrometry. *Anal Chim Acta* 2018, 1044, 181-190.
11. Michalikova K., *et al.* Middle-Up Characterization of the Monoclonal Antibody Infliximab by Capillary Zone Electrophoresis–Mass Spectrometry. *Lc Gc Eur* 2019, 32, 130-137.
12. Haselberg R., de Jong G. J., Somsen G. W. Low-flow sheathless capillary electrophoresis-mass spectrometry for sensitive glycoform profiling of intact pharmaceutical proteins. *Anal Chem* 2013, 85, 2289-2296.
13. Marcia R Santos C. K. R., Fonslow B., Guttman A. A covalent, cationic polymer coating method for the CESI-MS analysis of intact proteins and polypeptides. 2015.
14. Gstöttner C., *et al.* Intact and subunit-specific analysis of bispecific antibodies by sheathless CE-MS. *Anal Chim Acta* 2020, 1134, 18-27.
15. Habberger M., *et al.* Rapid characterization of biotherapeutic proteins by size-exclusion chromatography coupled to native mass spectrometry. *MAbs* 2016, 8, 331-339.





Chapter 8

Affinity capillary electrophoresis – mass spectrometry as tool to unravel proteoform-specific antibody-receptor interactions

Christoph Gstöttner^a, Michaela Hook^b, Tony Christopeit^c, Alexander Knaupp^c, Tilman Schlothauer^c, Dietmar Reusch^b, Markus Habberger^b, Manfred Wuhrer^a, Elena Domínguez-Vega^{a*}

^aLeiden University Medical Center, Center for Proteomics and Metabolomics, The Netherlands.

^bRoche Pharma Technical Development, Penzberg, Germany

^cPharma Research and Early Development, Roche Innovation Center Munich, Germany

Manuscript under review

8.1 Abstract

Monoclonal antibody (mAb) pharmaceuticals consist of a plethora of different proteoforms with different functional characteristics, including pharmacokinetics and pharmacodynamics, requiring their individual assessment. Current binding techniques do not distinguish between coexisting proteoforms requiring tedious production of enriched proteoforms. Here, we have developed an approach based on mobility shift-affinity capillary electrophoresis - mass spectrometry (ACE-MS), permitting to determine binding of co-existing mAb proteoforms to Fc receptors (FcRs). For high-sensitivity MS analysis we used a sheathless interface providing adequate mAb sensitivity allowing functional characterization of mAbs with high sensitivity and dynamic range. As a model system, we focused on the interaction with the neonatal FcR (FcRn), which determines the half-life of mAbs. Depending on the oxidation status, proteoforms exhibited different electrophoretic mobility shifts in presence of FcRn which could be used to determine their affinity. We confirmed the decrease of FcRn affinity with antibody oxidation and observed a minor glycosylation effect, with higher affinities for galactosylated glycoforms. Next to relative binding, the approach permits determination of individual K_d values in-solution resulting in values of 422 nM and 139 nM for oxidized and non-oxidized variants. Hyphenation with native MS provides unique capabilities for simultaneous heterogeneity assessment from mAbs, FcRn and complexes formed. The latter provides information on binding stoichiometry revealing 1:1 and 1:2 for antibody:FcRn complexes. The use of differently engineered Fc-only constructs allowed to distinguish between symmetric and asymmetric binding. The approach opens unique possibilities for proteoform-resolved antibody binding studies to FcRn and can be extended to other FcRs and protein interactions.

8.2 Introduction

Monoclonal antibodies (mAbs) have demonstrated to be beneficial for treating various diseases ranging from oncology over cardiovascular to infectious diseases [1]. Most therapeutic strategies using mAbs require binding to Fc receptors (FcRs) to trigger an immune response and for efficient recycling [2, 3]. The neonatal Fc receptor (FcRn), in particular, is involved in antibody recycling and transport through polarized membranes [3]. Recycling of mAbs from circulation determines the half-life of antibodies in serum (i.e. pharmacokinetics), whereas transcytosis is crucial for neonates as they are not able to produce enough IgGs during their first months [4]. This cellular transport and recycling mechanism is based on a pH-dependent binding of the antibody to the FcRn.

Antibody therapeutics contain not only one defined molecule but several different proteoforms of the same antibody. These proteoforms range from different glycoforms attached to the conserved N-glycosylation site at Asn297 [5,6] to additional posttranslational modifications (PTMs) on the protein backbone such as oxidation or deamidation. Antibody binding to FcRn is proteoform-dependent, and proteoforms that are not able to bind to the receptor (e.g. due to a modification close to the binding site) are directed to the lysosome for degradation [7]. Binding to FcRn occurs via the CH2-CH3 domain of human IgG [8, 9]. In this region two Met are present in human IgG1 (Met252 and Met428) which are crucial for the binding. Oxidation of Met252 (and to a minor extent Met428) drastically affects the binding to FcRn reducing its serum half-life [10-12]. The role of Fc glycosylation in FcRn binding has been more controversial. While some studies showed distinct FcRn binding between different glycoforms (e.g. deglycosylated, galactosylated or sialylated) [13-15], others reported no binding difference [16-18]. Recently, glycosylation of the Fab region has also been shown to influence the binding with the FcRn due to a decreased complex association [19]. However, most of these reports focus on specific features and do not consider potential confounding by other proteoforms (e.g. influence of oxidation during glycosylation studies). This highlights the importance of proteoform-selective binding assessment to draw reliable conclusions.

For studying the interaction between mAbs and FcRn enzyme-linked immunosorbent assays (ELISA) [20] and surface plasmon resonance (SPR) [21, 22] are the most widely employed techniques. In SPR, quantitative information on the antibody and Fc receptor affinity can be obtained (K_d , K_{on} , K_{off}) [23]. With ELISA an affinity value (EC_{50}) can be obtained, which describes the concentration necessary to obtain half of the maximum binding effect [24]. However, the major drawback of these techniques is that no distinction between different proteoforms can be made and an overall K_d or EC_{50} value for all the proteoforms of an antibody sample is obtained. Therefore, binding information on single proteoforms requires tedious proteoform isolation or engineering procedures [24,25] and yet faces the problem of potential confounding by modifications that go unnoticed. In an effort to

overcome this challenge, the use of affinity liquid chromatography hyphenated with mass spectrometry (affinity LC-MS) using FcRn columns has been proposed [26]. In this approach the antibody is injected at the binding pH and eluted by gradually increasing the pH to induce dissociation. Proteoforms with lower affinity will elute earlier than higher affinity ones, permitting to study differences in binding in a straightforward way. However, no K_d or EC_{50} values can be determined and, therefore, only relative differences in retention time under stress conditions can be studied [27]. Furthermore, immobilization of the FcRn receptor to a stationary phase does not permit further structural characterization of the interaction regarding e.g. stoichiometry and limit their application to FcRn, as columns with other immobilized FcRs are not commercially available. In short, currently there is not an optimal solution for proteoform-resolved study of antibody binding to FcRn and other FcRs.

Capillary electrophoresis (CE) separates analytes in an open-tube capillary without stationary phase. As separations occur in-solution, non-covalent (protein) interactions can be maintained during the analysis opening great possibilities to study protein binding under physiologically relevant conditions. One approach that has shown immense potential to study the affinity of individual proteoforms is the mobility-shift approach. In this approach, the capillary is filled with one (free) interaction partner (e.g. receptor), whereas the second one is injected in the system (e.g. mixture of proteoforms). During separation, proteoforms interact with the receptor influencing their effective electrophoretic mobility which can be exploited to determine K_d s [28, 29]. Traditionally, affinity CE is performed with UV detection, bringing significant limitations as many proteoforms often comigrate in one peak. Hyphenation with MS overcome this issue as structural heterogeneity is assessed and resolved. Mobility shift ACE-MS has so far only been employed to study the interaction of very small peptides with antibiotics [30-32] or stromal cell-derived factor-1 with different sulfated oligosaccharides [33]. Recently, we have for the first time shown the capabilities of mobility-shift ACE-MS for monitoring protein-protein interactions using two small model proteins, i.e., trypsinogen (24 kDa) and aprotinin (6.5 kDa) [29]. Hyphenation with MS was done using a sheath-liquid interface resulting in a limited sensitivity and hampering the possibility to apply the approach to antibodies and FcRs (approx. 150 and 50 kDa, respectively).

Sheathless interfaces for CE-MS operate with nano-ESI and have demonstrated to provide better ionization efficiency and suffer less from ionization suppression opening possibilities to application to larger proteins such antibodies [34]. In this work, we developed a new approach based on mobility-shift ACE-MS using sheathless interfacing for the proteoform-specific assessment of the interaction between mAb and the FcRn receptor. Different engineered antibodies and oxidized samples were employed to develop and demonstrate the capabilities of the approach. We show that we can simultaneously monitor the heterogeneity of the antibody and the FcRn and specifically assess their binding in-solution. This allowed us to determine specific K_d values for the different proteoforms in a mixture

without the need of isolation or proteoform engineering. Furthermore we determined the antibody:FcRn complex stoichiometry 1:2 with symmetrical FcRn binding.

8.3 Experimental section

8.3.1 Chemicals

All reagents employed were of analytical grade. Ammonium hydroxide, dithiothreitol (DTT), 7.5 M ammonium acetate (AmAc) solution, hydrogen chloride, guanidinium hydrochloride (Gua-HCl), tris-hydroxymethyl aminomethane (Tris), calcium chloride and H_2O_2 were obtained from Sigma Aldrich (Steinheim, Germany). Formic acid (FA), acetonitrile (ACN) and iodoacetic acid (IAA) were purchased from Thermo Scientific (Pittsburgh, PA). NAP-5 columns were obtained from GE Healthcare (Chicago, IL). Deionized water was obtained from a Milli-Q purification system (EMD Millipore, Burlington, MA). Bio-Spin 6 columns were purchased from Bio-Rad (Hercules, CA). 10 kDa VivaSpin MWCO filters were provided by Sartorius (Göttingen, Germany). Trypsin proteomic-grade was obtained from Roche Diagnostics (Mannheim, Germany).

8.3.2 Samples

Standard and engineered mAbs, Fc-only constructs and FcRn were provided by Roche Diagnostics (Penzberg, Germany). The FcRn consisted of the human beta-2-microglobulin with a (G4S)₄ linker [35], followed by the extracellular domain of the human IgG receptor FcRn large subunit p51, a His 10-tag and an Avi-tag. Oxidation of mAb samples (1 $\mu\text{g}/\mu\text{L}$) was achieved by adding different amounts of a 1% H_2O_2 solution resulting in a final H_2O_2 concentration ranging from 0 to 0.2% and incubation for 24 h at 25°C. mAbs and Fc-only samples were buffer exchanged to 50 mM AmAc pH 6.0 using NAP-5 columns. The columns were equilibrated with a complete buffer fill four times. Next, 300 μL (1 mg) of mAb samples were loaded onto the column and allowed to drip through. After adding an additional 350 μL of 50 mM AmAc pH 6.0, the sample was eluted with 500 μL 50 mM AmAc pH 6.0. The concentration of the samples was determined using a NanoPhotometer (Implen, Munich, Germany) and diluted to obtain a concentration of 1 $\mu\text{g}/\mu\text{L}$ for mAb samples or mixtures of oxidized and non-oxidized mAbs and 0.75 $\mu\text{g}/\mu\text{L}$ for Fc-only samples prior injection. Lysozyme from chicken egg (Sigma Aldrich) was desalted and buffer exchanged to 50 mM AmAc pH 6.0 using 10 kDa VivaSpin columns. Lysozyme was diluted to a final concentration of 2 $\mu\text{g}/\mu\text{L}$. Similarly, the FcRn receptor (4.9 $\mu\text{g}/\mu\text{L}$) was desalted and buffer exchanged to 50 mM AmAc pH 6.0 using 10 kDa VivaSpin columns and diluted to the desired concentration (0.5 to 12 μM).

8.3.3 Affinity CE-MS

Affinity CE experiments were carried out on a Sciex (Framingham, MA) CESI 8000 instrument

using OptiMS neutrally-coated capillaries (Sciex) with a porous tip (length, 91 cm; 30 μ m i.d.; 150 μ m o.d.). The capillaries were prepared following the instructions of the manufacturer. The capillary was flushed for 5 min (100 psi, forward) with 0.1 M HCl, 10 min (100 psi, forward) with 50 mM AmAc (pH 3.0) and 30 min (100 psi, forward) with MilliQ, followed by an equilibration for 16-18 h. This procedure was performed only prior first injection. Before each run, the capillary was flushed for 2 min with 0.1 M HCl (100 psi, forward pressure), 2 min with MilliQ (100 psi, forward pressure), 2 min with 50 mM AmAc pH 6.0 (100 psi, reverse pressure) and 2 min with 50 mM AmAc pH 6.0 (100 psi, forward pressure). Following the capillary was filled for 2 min (100 psi, forward pressure) with 50 mM AmAc pH 6.0 containing different concentrations of FcRn receptor or without receptor. Afterwards, a marker protein (lysozyme) was injected (1.5 psi, 15 s), followed by the antibody sample (2.5 psi, 15 s) and a plug of BGE with or without receptor (1 psi, 25 s). The separation was carried out for 45 min with 20 kV at 25° C. After the separation was complete, the voltage was ramped down to 1 kV in 5 min.

The outlet of the separation capillary was placed in a nano-electrospray ionization source at the entrance of an Orbitrap Exactive Plus Extended Mass Range mass spectrometer (Thermo, Waltham, MA). The mass spectrometer was operated in positive ionization mode using the following parameters: capillary voltage 1.8-2.2 kV; ion injection time 200 ms; HCD energy 100 eV; skimmer voltage 15 V and S-Lens voltage 15 V. Mass spectra were recorded in a mass range of 1,000 to 15,000 m/z with a resolution of 17,500 at m/z 200. MS control and data acquisition were performed using the Xcalibur (Thermo Fisher Scientific) software. For data analysis, the Intact Mass software from Protein Metrics (Cupertino, CA) was used.

For calculation of the K_d s, the electropherograms were aligned using the marker protein (lysozyme) and the electrophoretic mobility shift compared to the measurement without FcRn was calculated for each proteoform. Following data were plotted using the software GraphPadPrism using the one side specific binding model.

8.3.4 Assessment of mAb oxidation levels by LC-MS/MS

To determine the oxidation levels of the mAbs and Fc-only samples, 50 μ g (50 μ L) of sample were denatured with an equal amount of denaturation buffer (8 M Gua-HCl, 0.4 M Tris/HCl, pH 8.5). Afterwards, the samples were reduced by adding DTT to a final concentration of 20 mM and incubated for 60 min at 50°C. Following, the samples were alkylated with IAA in a final concentration of 50 mM and kept for 30 min in the dark. For buffer exchange, prior to trypsin digestion, the samples were loaded on Bio-Spin 6 columns. Beforehand the columns were 2 min centrifuged at 1000xg to remove the storage liquid and three times equilibrated with 500 μ L digestion buffer (50 mM Tris/HCl, 2 mM CaCl₂, pH 7.5) and centrifuged for 2 min at 1000xg in between. The reduced and alkylated samples were loaded onto the column and centrifuged for 4 min at 1000xg collecting thereby the eluate. 25 μ g trypsin were dissolved in 100 μ L 10 mM HCl solution. 3 μ L of trypsin solution were added to each sample and

incubated for 18 hours at 37°C. To stop the digestion, 17 μL of 10% TFA solution were added and the samples were diluted 1:1 (v/v) with MilliQ prior to LC-MS measurement. 10 μL sample were injected into an Acquity UPLC (Waters, Milford, MA, USA) equipped with a CSH C18, 1.7 μm , 130 $\text{\AA}$, 2.1x150 mm column (Waters). The separation was performed using 0.1% FA in H₂O as mobile phase A and 0.1% FA in ACN as mobile phase B at a column temperature of 65°C. A linear gradient from 1% B to 35% B in 45 min was used for the separation of peptides, followed by a cleaning of 3 min at 8% B and a re-equilibration for 4 min at 1% B. The LC was coupled to an Orbitrap Velos mass spectrometer operated in positive ionization mode. The isolation width for CID fragmentation was 1 m/z, the fragmentation energy was set to 35 eV, the activation q was 0.25 and the activation time 10 ms. The resolution was set to 30000, and an m/z range was 200 to 2000 in full scan mode was monitored. The amount of oxidation for the sites Met252 and Met428 were calculated by determining the area under the curve of the non-oxidized and oxidized peptide extracted ion chromatograms and determining the relative ratio of oxidized peptide in each sample.

8.4 Results and Discussion

8.4.1 Development of a mobility-shift ACE-MS approach for proteoform-resolved antibody - FcRn binding

The binding of antibodies to FcRn takes place in the endosome at a pH between 5.5 and 6.0. Mobility shift affinity approaches rely on the different mobility of the protein and receptor and in consequence, the protein-receptor complex. To study which conditions show the largest different in mobility between FcRn and mAb1, a mixture of both proteins was injected (1 $\mu\text{g}/\mu\text{L}$ each) and analyzed using a background electrolyte consisting of 50 mM AmAc at pH 5.5 or 6.0. At these pHs, both FcRn receptor (calculated pI 6.24) and mAb1 (calculated pI 8.47) are positively charged. Therefore, separations were performed using a neutrally-coated capillary to avoid adsorption to the capillary wall. 50 mM AmAc at pH 5.5 resulted only in a minor difference in mobility between FcRn and mAb1 (**Figure S1**). Using AmAc at 6.0, FcRn exhibited significantly lower mobilities than mAb1 (**Figure S1**), indicating that a shift in the mobility could be observed for the FcRn-antibody complex.

For affinity CE experiments, FcRn was added to the BGE, and after filling the capillary with the receptor, the mixture of antibody proteoforms was injected. Antibody proteoforms with high affinity towards FcRn would interact with the receptor during the separation and exhibit lower effective mobility, while proteoforms with no interaction should not change their effective mobility compared to their analysis in absence of receptor. To correct for changes in the ionic strength and viscosity of the BGE with addition of different amounts of FcRn, a marker protein showing no interaction with FcRn (lysozyme, molecular weight 14.3 kDa, calculated pI 10.36) was employed.

Initial experiments were performed with an engineered antibody which bears no glycans in the Fc domain (NGmAb). Oxidation of Met252 is known to influence the binding between the antibody and the FcRn receptor. Therefore, to fully explore the capabilities of the approach, analyses were performed with NGmAb material which was intentionally oxidized as confirmed by bottom-up proteomics. For NGmAb oxidation levels for Met 252 were 41%, comprising a mixture of antibodies containing 0, 1 or 2 oxidations at this site. First, the mixture of oxidized and non-oxidized species were analyzed in absence of FcRn receptor using 50 mM AmAc at pH 6.0.

Upon oxidation, NGmAb showed unchanged electrophoretic mobility under standard CE conditions (**Figure 1A**) which was in line with expectations as oxidation comes with only a relatively small increase in intact mass while not affecting mAb charge.

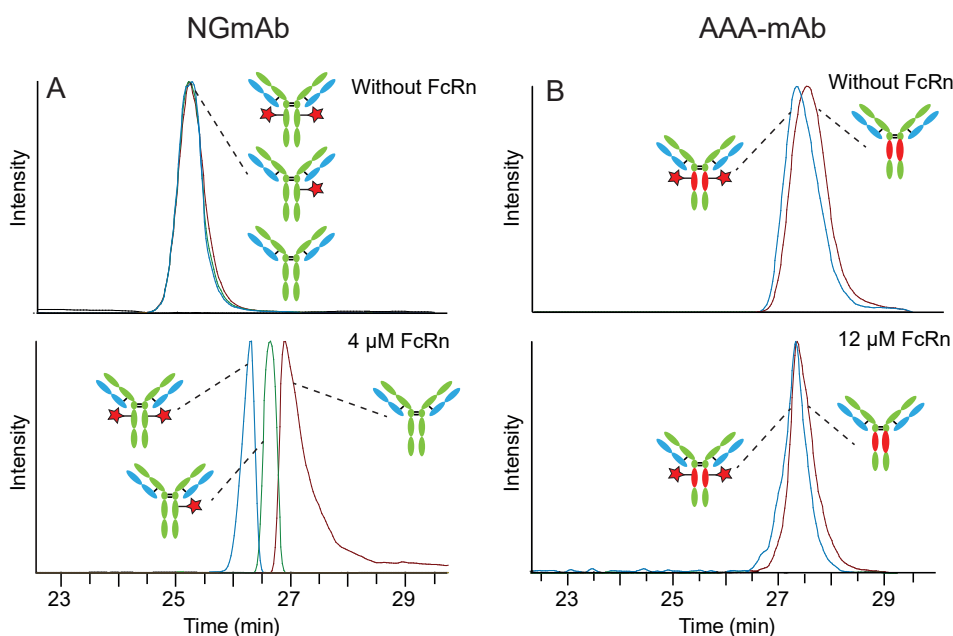


Figure 1. Sheathless CE-MS separation obtained for a H_2O_2 stressed A) NGmAb or B) AAA-mAb sample using a BGE without FcRn (upper panel) and containing FcRn (lower panel). The blue trace corresponds to the extracted ion electropherograms (EIEs) of the double-oxidized (two red stars), the green trace corresponds to the EIEs of mono-oxidized (one red star) and the brown trace shows the EIEs of the non-oxidized mAbs present in the mixture. Signal intensities are normalized.

Reported K_d s for antibodies and FcRn are between 3.2–284.7 nM [36]. Therefore, for affinity experiments FcRn was added to the BGE at concentrations in the range of nM to low μ M. Analysis of the oxidized NGmAb sample using 4 μ M FcRn in the BGE resulted in a clear shift of the effective electrophoretic mobility for all the species indicating that all of them interact with FcRn (**Figure 1A**). More interestingly, the different species start to separate indicating

different binding affinities. The double oxidized form of the NGmAb migrated in the first place indicating that it has the lowest binding affinity while the unmodified NGmAb which migrates the latest has the highest binding affinity. The mono-oxidized NGmAb showed an intermediate shift between the double and non-oxidized antibody indicating a decrease on binding affinity. However, is important to note that the observed decrease corresponds to the inaccessibility of one of the Fc chains for binding to FcRn in solution, but does not imply in-vivo half-life decrease. In addition, even if a decrease on binding affinity was observed for the oxidized forms, still binding to FcRn was observed for all the species.

To confirm that these observations correspond to the specific binding of the antibody to the FcRn and are not the result of non-specific interactions or any other artifact, we analyzed an antibody which has been silenced for binding with FcRn (AAA-mAb or triple-A mutant). The antibody was intentionally oxidized prior analysis resulting in oxidation levels of 99.6% for Met252 (i.e. double oxidized antibodies). Analysis of a 1:1 mixture between oxidized and non-oxidized AAA-mAb in absence of FcRn resulted in two overlapping signals at 27.5 min resulting from the double-oxidized and the unmodified triple-A antibody (**Figure 1B**). Of note, different antibodies (e.g. triple-A and NGmAb) exhibit different electrophoretic mobilities and therefore, the obtained migration times can not be taken as absolute values of affinity but relative with their analysis in absence of receptor (i.e. mobility shift). Addition of up to 12 μM FcRn receptor to the BGE did not induce any shift on the mobility or additional separation indicating that the AAA-mAb is not binding to FcRn (**Figure 1B**). These results confirmed that the shift on the mobility observed for the NGmAb antibody, indeed is induced by the specific binding of the antibody with the FcRn receptor.

To demonstrate the applicability of the method for simultaneous binding assessment of multiple proteoforms, a standard glycosylated mAb was employed. As in previous cases, a 1:1 mixture of oxidized (97.7% for Met252) and non-oxidized mAb1 was analyzed, resulting in co-migration in absence of FcRn (**Figure 2A**). Addition of 0.25 and 1 μM of receptor in the BGE induced a shift on the mobility for both species with a partial separation between the double-oxidized and the non-oxidized mAb for the latest (**Figure 2A**). Further increase on the concentration (up to 6 μM) did not result in a higher mobility shift and/or further separation suggesting that a plateau is reached above 1 μM (**Figure 2A**). One characteristic of mobility shift affinity CE approaches is their capability to provide quantitative information about binding. Next to monitor differences in relative binding for different proteoforms we also wanted evaluate the capabilities of the method to calculate their absolute binding – i.e. determination of affinity constants in a proteoforms-selective manner. After determining and plotting the shift on the effective electrophoretic mobility for the measured concentrations of FcRn a binding curve was obtained. By fitting the curve using nonlinear regression a K_d value of 139 ($\pm$ 67) nM for the non-oxidized and 422 ($\pm$ 157) nM for the oxidized mAb was obtained (**Figure S2**). Reported K_d values are in the range of 3.2–284.7 nM [36] depending in the specific structural features of the antibody (Fc and Fab portion) which are in line with the

obtained values. However, the errors associated to our K_d determinations were significantly elevated indicating that more points and further correction of electrophoretic mobilities would be needed for accurate K_d determinations but this initial exploration shows that the current approach could be used to determine K_d s of different antibody proteoforms.

In addition to oxidation different glycoforms can potentially affect the binding with FcRn. mAb1 comprise different glycoforms dominated by complex-type glycans containing different core fucosylation and galactosylation levels. The influence of these glycans on FcRn binding was investigating by looking at their shift on their electrophoretic mobility (**Figure 2B**). While different glycoforms do not influence electrophoretic mobility (**Figure 2B, upper panel**), shifted profiles were observed in presence of FcRn (**Figure 2B, lower panel**). The shift difference between glycoforms was significantly lower than the observed for oxidoforms but indicate some degree of influence on FcRn binding. In line with recent publications we did not observe any effect on fucosylation in binding while higher galactosylation slightly increased binding affinity [13-15].

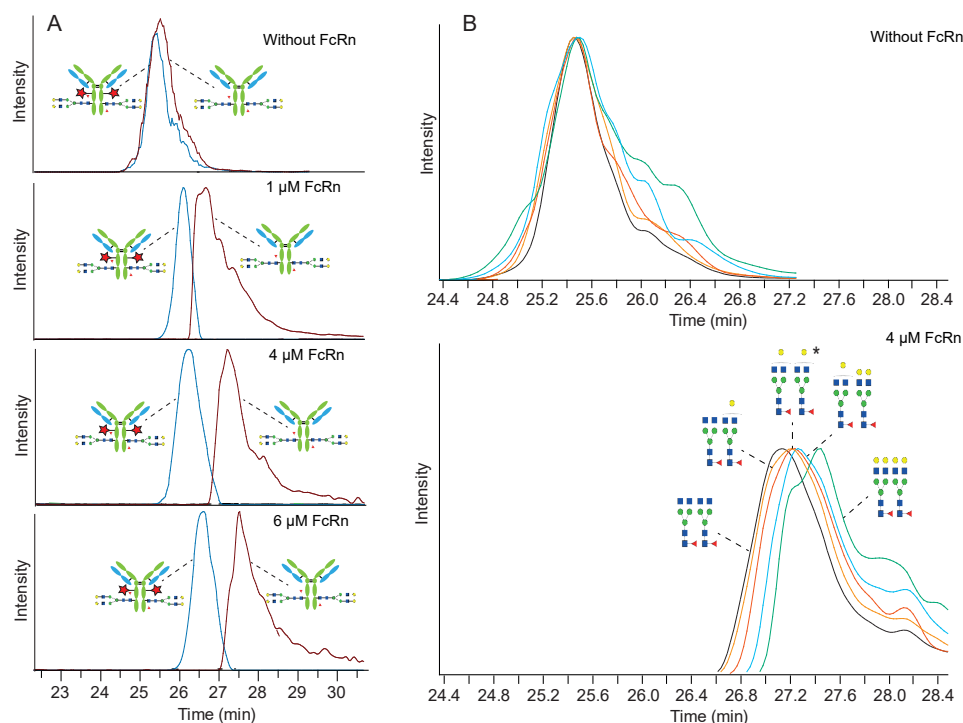


Figure 2. Sheathless CE-MS separation obtained for a 1:1 mixture of a H₂O₂ stressed and reference mAb1 using a BGE without or with different amounts of FcRn. A) Illustration of the effect of oxidation. The blue trace corresponds to the EIEs of the double-oxidized (two red stars) and the brown trace shows the EIEs of the non-oxidized mAbs. In both cases the antibody with G0F and G1F glycoforms were extracted. B) Illustration of the effect of glycosylation. In different colors EIEs of different glycoforms of the non-oxidized antibody. Signal intensities are normalized. *Either G1F/G1F or G2F/G0F.

8.4.2 Structural assessment by on-line native MS

Affinity separations were performed with online native MS detection permitting, next to functional assessment, structural characterization. The use of 50 mM AmAc as BGE provided good signal intensities for the antibody proteoforms. For the NGmAb, which is non-glycosylated, next to main form, two additional signals corresponding to the antibody with one (+162 Da) and two times glycation (+324 Da) were observed. After addition of the receptor to the BGE, additional signals appeared in the mass spectra corresponding to the free FcRn and in complex with the antibody, which were more perceptible with increasing amounts of FcRn in the BGE. The constant presence of FcRn in the MS also caused a slight decrease of antibody signal intensity. The plateau was reached at around 4 μ M of FcRn depending on the antibody. Still, at higher concentrations of FcRn (up to 12 μ M was tested) the ionization suppression was moderate and good sensitivity was obtained for the detection of the antibody proteoforms and their complexes. The constant FcRn signal, on the other hand, permitted the simultaneous characterization of the receptor heterogeneity during the affinity experiments. FcRn was detected as a complex pattern of signals at around 49 kDa, corresponding to different glycoforms at the N-glycosylation site at position 220 on our FcRn construct (**Figure 3** and **Figure S3**). The glycoforms were assigned based on their average mass and on previously reported glycopeptide data [37] (**Figure S3**). Next to the antibody and the receptor, NGmAb in complex with FcRn was detected in a stoichiometry of 1:1 or 1:2 (**Figure 3**). At concentrations of FcRn of 2 μ M only NGmAb in complex with 1 FcRn molecule (**Figure S4A**) was observed while at 4 μ M NGmAb in complex with 2 FcRn molecules was also noticeable (**Figure S4B**). Further increase of FcRn concentration up to 12 μ M higher resulted in an increase of the signals in particular for the corresponding to the NGmAb-FcRn 1:2 complex (**Figure S4C, D**). **Figure 3** shows the zoomed spectra of the complex of the NGmAb with one FcRn molecules. Several signals were observed corresponding to the antibody in complex with different FcRn glycoforms (**Figure 3**) which can be clearly represented after simulation of the glycoprofile of the NGmAb-FcRn for the most abundant FcRn glycoforms (**Figure 3**, red traces). The NGmAb in complex with 2 FcRn was also observed at around 243 kDa, as a complex pattern of signals where the main glycoforms were detected (**Figure 3**).

Using the glycosylated standard mAb1 resulted in a more complex pattern of signals for the complex formed between the mAb1 and FcRn. Still, we were able to detect the antibody in complex with one FcRn (**Figure S5**). The deconvoluted mass spectrum of this complex is dominated by the glycosylation of the antibody (G0/G0 to G2F/G2F) (**Figure S5**) which still contains different glycoforms of the receptor resulting in broader signals. The complex (**Figure S5** red lines) was simulated using the two mayor glycoforms of the FcRn receptor resulting in a good match between the theoretical and the observed spectra. The complex of mAb1 with two FcRn molecules could not be detected, most probably due to the high complexity arising from the glycosylation of the antibody and the two FcRn molecules resulting in a higher spread of the signals and, therefore, reduced signal intensity.

Monitoring the heterogeneity of all the species including the antibody, receptor and the complex results multiple benefits. Next to establishment of the stoichiometry, it permit to study if the variability of the FcRn receptor (e.g. different glycoforms) have an influence in the binding with the antibody -i.e. it permits to study the interaction in a receptor proteoform specific manner. For our particular example on FcRn we could not observe any clear preference of certain FcRn glycoform in binding with NGmAb based on the obtained data, suggesting no influence of the FcRn glycosylation on the binding with the antibody. The observed glycation on the NGmAb (**Figure 3**) did also not show any shift in mobility as expected.

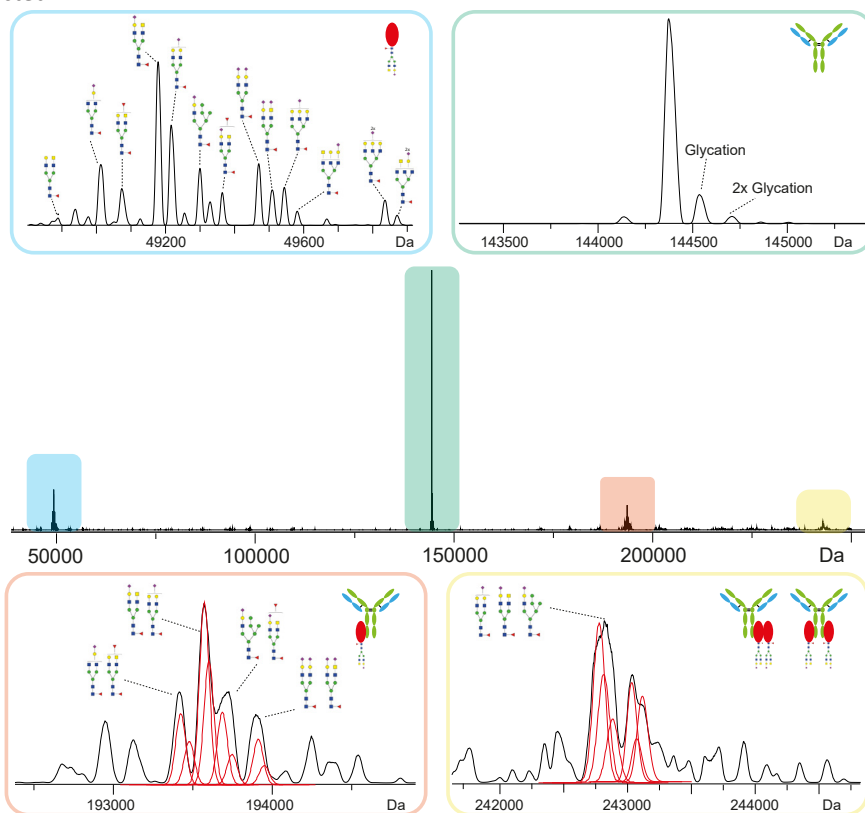


Figure 3. Deconvoluted mass spectrum of the non-stressed NGmAb using a BGE containing 12 μ M FcRn. Blue, zoom of FcRn containing several different glycoforms. Green, zoom of NGmAb containing 0, 1 or 2 glycations. Vermillion, zoom of NGmAb in complex with 1 FcRn. Yellow, zoom of NGmAb with 2 FcRn molecules. Complexes heterogeneity comes from different glycoforms of the FcRn (most abundant ones are simulated in red).

8.4.3 Evaluation of antibody:FcRn binding stoichiometry and symmetry using Fc-only constructs

Unlike other Fc receptors, the binding of antibodies to FcRn occurs in a 1:2 stoichiometry. This has been demonstrated by crystal structures of rat FcRn and rat IgG2a [38]. However, whereas some reports suggest a dimerization of the rat FcRn and binding only to one side of the Fc portion [38, 39] other studies using column binding assays [40] or using mutated Fcs [41] suggested binding of one FcRn per Fc chain. For human FcRn similar opposite observations have been made. While no FcRn dimers could be observed in crystallization experiments [42] recent publications show strong evidences on self FcRn interaction in a pH dependent manner [43]. Recently, cryo EM analysis of mAbs:FcRn showed 1:1 and 1:2 stoichiometry with one FcRn molecule binding to each side of the mAb-Fc [44].

Our MS results show 1:2 binding for NGmAb but still this does not answer the question on binding symmetry. To test whether we can also study if we have an asymmetric or symmetric binding of the FcRn with our mobility-shift CE-MS approach, we analyzed two different engineered Fc constructs [45]. These constructs consist on the Fc portion (two Fc/2 chains) connected by disulfide bridges. One of the constructs comprised two unmodified Fc/2 chains (wt/wt) and should allow for FcRn binding on both sides. The second construct carried a triple-A mutation in one of the Fc/2 chains (wt/AAA) permitting only binding of the FcRn to one of the two Fc/2 chains. Injection of a mixture of both constructs without FcRn led to a separation of both Fc constructs due to their different pI (and therefore, different electrophoretic mobility) with the wt/wt migrating earlier than the wt/AAA (**Figure 4A**). After filling of the capillary with 6 μ M FcRn receptor a clear shift on the mobility for both constructs was observed (**Figure 4B**). For wt/AAA the mobility shift was minor while for the wt/wt constructs a larger shift was observed.

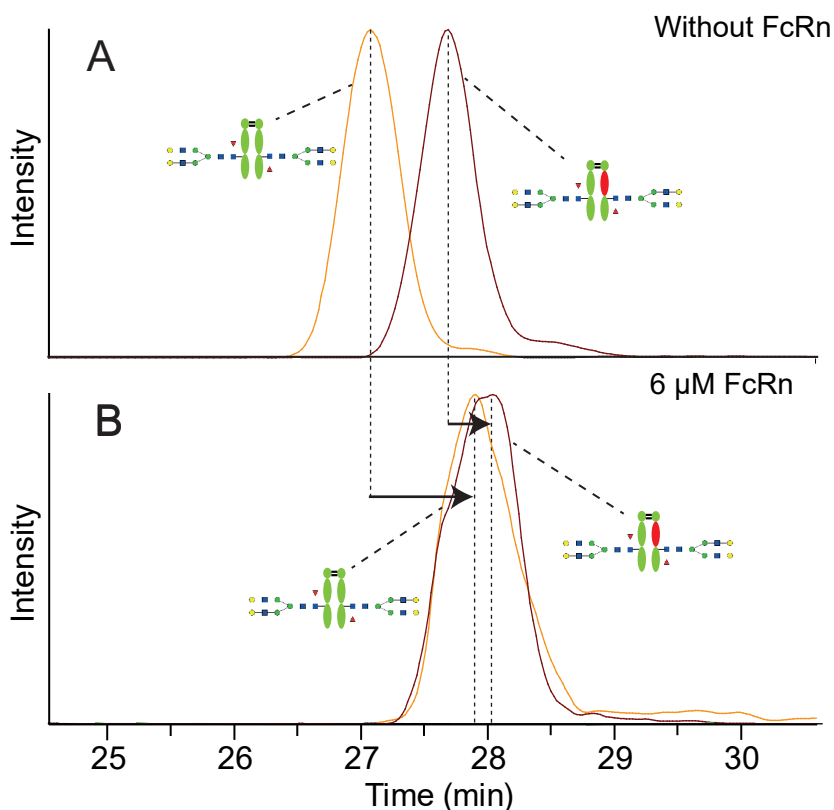


Figure 4. Sheathless CE-MS separation obtained for a 1:1 mixture Fc constructs (wt/wt (green) or wt/AAA (green and red)) using a BGE A) without FcRn or B) containing 6 μ M FcRn. The orange trace represent the EIE of the wt/wt Fc-construct and brown trace shows the EIE of the wt/AAA Fc-construct. Signal intensities are normalized.

These results indicates, that the construct which allowed a two side (symmetrical) binding showed a larger shift than the one where only one FcRn molecule can bind. These results suggest that under these conditions the observed 1:2 complex correspond most likely to the one FcRn molecule bound to each side of the mAb-Fc as reported recently [44]. Looking at the mass spectra no dimers of the FcRn were observed supporting this observation. However, dissociation during ionization can not be excluded. Regarding the complexes we could only observe the complex between the Fc constructs (wt/wt and wt/AAA) and one FcRn (**Figure S6**). As for mAb1, the complex with two FcRn receptor for the wt/wt construct could not be detected due to the large heterogeneity coming from the glycosylation of the Fc construct and the two glycosylated FcRn molecules. Furthermore, from this experiment it can be concluded that the missing Fab portion do not abrogate binding of the Fc to FcRn and still a mobility shift can be observed. Although the absolute affinity of this molecules would be different from their complete counterparts containing the Fab domain, the proposed

approach can be applied to monitor the effect of Fc modifications in FcRn binding opening possibilities to application to polyclonal samples.

8.5 Conclusion

We developed a unique approach based on mobility shift ACE-MS to study the affinity of individual antibody proteoforms in a sample to FcRn. This approach represents the first affinity ACE-MS for functional studies on mAbs and overcome most of limitations faced with current binding techniques. We have focus in FcRn antibody interaction, which is of special interest for antibody recycling and thereby important for antibody therapeutic pharmacokinetics. We demonstrated that proteoforms with different binding affinities (i.e. oxidoforms) exhibit different shift in the electrophoretic mobility in presence of FcRn. This permits to determine their individual binding affinity without necessity of enrichment or purification of different forms. We also showed that K_d values could be individually determined for the oxidized and non-oxidized mAb1 in mixture, providing lower K_d for the non-oxidized mAb1. The mass spectrometric detection allows not only to characterize the antibody and its proteoforms, but also the heterogeneity of the receptor and the complexes formed. Another advantage of MS detection is the possibility to also characterize and study binding affinity of overlapping species. Analyzing a standard glycosylated mAb showed slight difference in binding of Fc-glycoforms to FcRn with galactosylated variants showing higher binding as recently suggested in few reports. Similar observations were made through MS monitoring of the complex between mAb and FcRn receptor where no preferred binding of FcRn or antibody glycoforms could be monitored. In the particular case of the FcRn receptor heterogeneity our data did not show specific FcRn glycoforms in higher abundance in the complex compared to the background FcRn signal but more work would be needed to confirm this due to the complexity of the spectra. Reversing the approach (i.e. adding the mAb in the BGE and inject the receptor) could provide information on proteoform receptor binding in a more accurate way, and will be further studied. Furthermore we investigated the binding stoichiometry between mAbs and the FcRn. Our MS spectra revealed 1:1 and 1:2 mAb:FcRn complexes which under the applied conditions seems to be symmetrically bound. We used two different Fc-only constructs which allowed one or two side binding of the FcRn. We found binding of both Fc-only construct to FcRn with the one allowing both sides binding providing a larger mobility shift compared to the construct allowing only one side binding.

The flexibility and simplicity of switching receptors (just changing the BGE) next to the low amount of receptor necessary compared to classical affinity LC, makes this approach very attractive for effector function antibody monitoring in a proteoform selective manner. Due to the binding similarity of FcRn to other FcRs we believe that this approach could be extended to study their binding. More studies are warranted on the interaction of mAbs with other Fc receptors.

8.6 Supporting Information

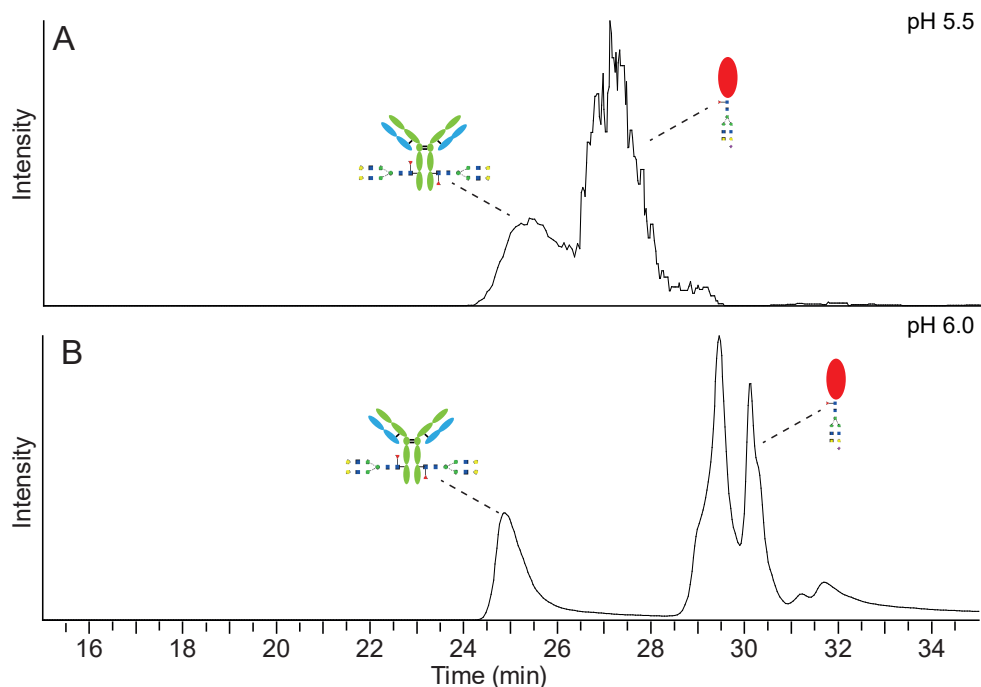


Figure S1. Sheathless CE-MS separation obtained for mAb1 and FcRn using A) 50 mM AmAc pH 5.5 and B) 50 mM AmAc pH 6.0.

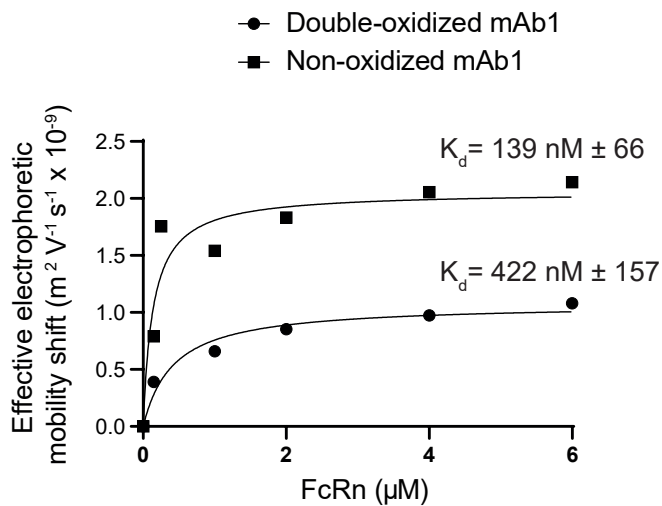


Figure S2. Binding curves obtained by plotting the mobility shift for non-oxidized and oxidized mAb at different concentrations of FcRn.

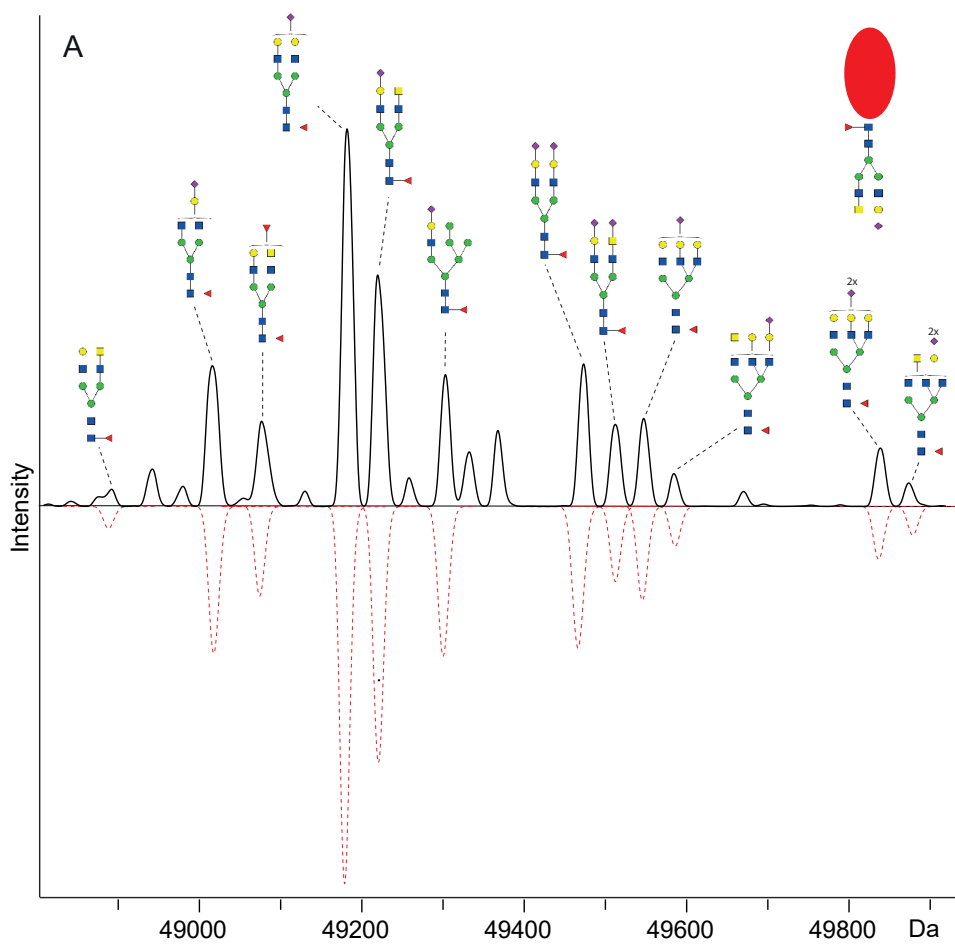


Figure S3. Glycosylation of FcRn receptor (black line) and simulated glycoforms in dashed red line.

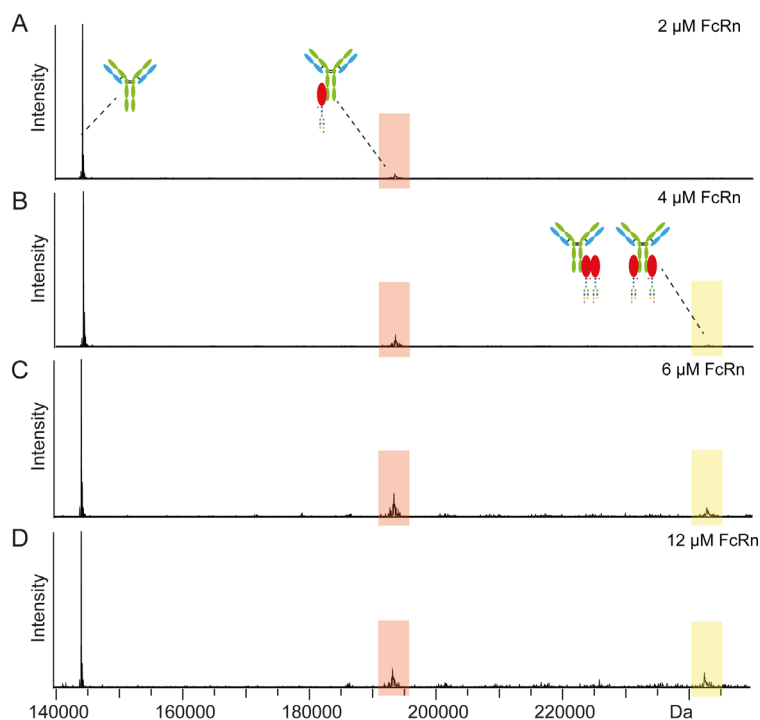


Figure S4. Deconvoluted mass spectra after ACE-MS separation of NGmAb using different amount of FcRn receptor in the BGE. A) 2 μ M FcRn receptor, B) 4 μ M FcRn receptor, C) 6 μ M FcRn receptor and D) 12 μ M FcRn receptor. Complex between NGmAb and 1 FcRn receptor is highlighted in vermilion and NGmAb in complex with 2 FcRn receptors highlighted in yellow.

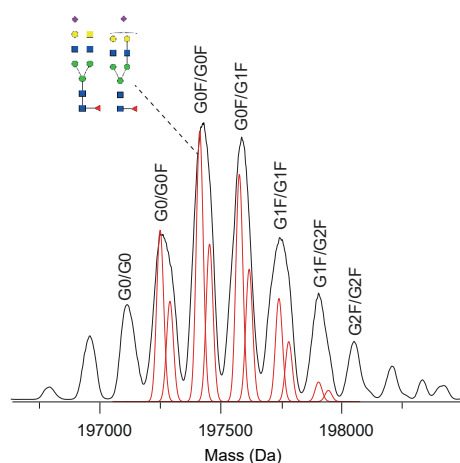


Figure S5. Deconvoluted mass spectra of the non-oxidized peak highlighting the area where the complex between mAb:FcRn appear and simulation of the mass spectra for the complex with the two major glycoforms of the FcRn receptor (red lines).

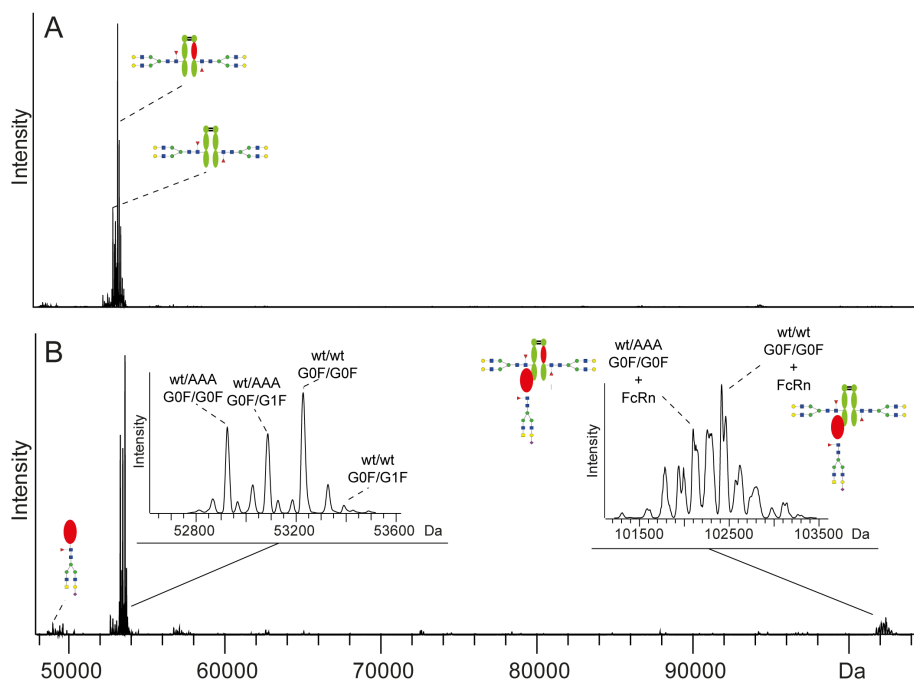


Figure S6. Deconvoluted mass spectra after ACE-MS separation of a 1:1 mixture Fc constructs (wt/wt (green) or wt/AAA (green and red)) using a BGE containing different amounts of FcRn. A) No FcRn and B) 6 μ M FcRn. Shown are the free FcRn, the Fc constructs and the complex of the Fc constructs with one FcRn.

8.7 Acknowledgments

We thank Protein Metrics for the temporal software license of their Intact mass software for data analysis. This work was supported by the Analytics for Biologics project (Grant agreement ID 765502) of the European Commission and the LUMC Fellowship 2020 to E.D.V.

8.8 References

1. Lu R. M., *et al.* Development of therapeutic antibodies for the treatment of diseases. *J Biomed Sci* 2020, 27, 1.
2. Koenderman L. Inside-Out Control of Fc-Receptors. *Front Immunol* 2019, 10, 544.
3. Pyzik, M., *et al.* The Neonatal Fc Receptor (FcRn): A Misnomer? *Front Immunol* 2019, 10, 1540.
4. Borghi S., *et al.* FcRn, but not Fc gamma Rs, drives maternal-fetal transplacental transport of human IgG antibodies. *P Natl Acad Sci* 2020, 117, 12943-12951.

5. Chen C. H., *et al.* Intact NIST monoclonal antibody characterization—Proteoforms, glycoforms—Using CE-MS and CE-LIF. *Cogent Chem* 2018, 4, 1480455.
6. Kabat E. A. National Institutes of Health, Columbia University. Sequences of proteins of immunological interest. U.S. Dept. of Health and Human Services, Public Health Service, National Institutes of Health: Bethesda, MD, 1991.
7. Ghetie V., Ward E. S. Multiple Roles for the Major Histocompatibility Complex Class I–Related Receptor FcRn. *Annu Rev Immunol* 2000, 18, 739-766.
8. Martin W. L., *et al.* Crystal Structure at 2.8 Å of an FcRn/Heterodimeric Fc Complex: Mechanism of pH-Dependent Binding. *Mol Cell* 2001, 7, 867-877.
9. Kim J. K., F, *et al.* Mapping the site on human IgG for binding of the MHC class I-related receptor, FcRn. *Eur J Immunol* 1999, 29 (9), 2819-25.
10. Bertolotti-Ciarlet A., *et al.* Impact of methionine oxidation on the binding of human IgG1 to Fc Rn and Fc gamma receptors. *Mol Immunol* 2009, 46 (8-9), 1878-82.
11. Mo J., *et al.* Understanding the Impact of Methionine Oxidation on the Biological Functions of IgG1 Antibodies Using Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal Chem* 2016, 88 (19), 9495-9502.
12. Wang W., *et al.* Impact of methionine oxidation in human IgG1 Fc on serum half-life of monoclonal antibodies. *Mol Immunol* 2011, 48, 860-866.
13. Wada R., Matsui M., Kawasaki N. Influence of N-glycosylation on effector functions and thermal stability of glycoengineered IgG1 monoclonal antibody with homogeneous glycoforms. *MAbs* 2019, 11, 350-372.
14. Dashivets T., *et al.* Multi-Angle Effector Function Analysis of Human Monoclonal IgG Glycovariants. *PLOS ONE* 2015, 10, e0143520.
15. Jennewein M. F., *et al.* Fc Glycan-Mediated Regulation of Placental Antibody Transfer. *Cell* 2019, 178, 202-215.
16. Simmons L. C., *et al.* Expression of full-length immunoglobulins in *Escherichia coli*: rapid and efficient production of aglycosylated antibodies. *J Immunol Methods* 2002, 263, 133-147.
17. Pawlowski J. W., *et al.* Influence of glycan modification on IgG1 biochemical and biophysical properties. *J Pharm Biomed Anal* 2018, 151, 133-144.
18. Jefferis R. Isotype and glycoform selection for antibody therapeutics. *Arch Biochem Biophys* 2012, 526, 159-166.
19. Volkov M., *et al.* ACPA Illustrating the impact of IgG Fab-glycosylation on transplacental transfer of. *Ann Rheum Dis* 2021, 80, 1044.
20. Lu Y., *et al.* Identification of IgG1 variants with increased affinity to FcγRIIIa and unaltered affinity to FcγRI and FcRn: Comparison of soluble receptor-based and cell-based binding assays. *J Immunol Methods* 2011, 365, 132-141.
21. Bertolotti-Ciarlet A., *et al.* Impact of methionine oxidation on the binding of human IgG1 to FcRn and Fcγ receptors. *Mol Immunol* 2009, 46, 1878-1882.

22. Magistrelli G., *et al.* Robust recombinant FcRn production in mammalian cells enabling oriented immobilization for IgG binding studies. *J Immunol Methods* 2012, 375, 20-29.
23. Neuber T., *et al.* Characterization and screening of IgG binding to the neonatal Fc receptor. *MAbs* 2014, 6, 928-942.
24. Li T., *et al.* Modulating IgG effector function by Fc glycan engineering. *Proc Natl Acad Sci* 2017, 114, 3485.
25. Stracke J., *et al.* A novel approach to investigate the effect of methionine oxidation on pharmacokinetic properties of therapeutic antibodies. *MAbs* 2014, 6, 1229-1242.
26. Gahoual R., *et al.* Detailed Characterization of Monoclonal Antibody Receptor Interaction Using Affinity Liquid Chromatography Hyphenated to Native Mass Spectrometry. *Anal Chem* 2017, 89, 5404-5412.
27. Cymer F., *et al.* Evaluation of an FcRn affinity chromatographic method for IgG1-type antibodies and evaluation of IgG variants. *Bioanalysis* 2017, 9, 1305-1317.
28. Hutanu A., *et al.* Methionine oxidation of proteins analyzed by affinity capillary electrophoresis in presence of silver(I) and gold(III) ions. *Electrophoresis* 2021, 42, 1209-1216.
29. Domínguez-Vega E., *et al.* Simultaneous Assessment of Protein Heterogeneity and Affinity by Capillary Electrophoresis–Mass Spectrometry. *Anal Chem* 2015, 87, 8781-8788.
30. Lynen F., *et al.* Considerations concerning interaction characterization of oligopeptide mixtures with vancomycin using affinity capillary electrophoresis-electrospray mass spectrometry. *Electrophoresis* 1999, 20, 2462-2474.
31. Dunayevskiy Y. M., *et al.* Simultaneous Measurement of Nineteen Binding Constants of Peptides to Vancomycin Using Affinity Capillary Electrophoresis–Mass Spectrometry. *J Med Chem* 1998, 41, 1201-1204.
32. Machour N., *et al.* Analysis of virtual two-dimensional gels based upon affinity capillary electrophoresis hyphenated to ion trap-mass spectrometry. *Electrophoresis* 2005, 26, 1466-1475.
33. Fermas S., *et al.* Sulfated oligosaccharides (heparin and fucoidan) binding and dimerization of stromal cell-derived factor-1 (SDF-1/CXCL 12) are coupled as evidenced by affinity CE-MS analysis. *Glycobiology* 2008, 18, 1054-1064.
34. Haselberg R., *et al.* Performance of a sheathless porous tip sprayer for capillary electrophoresis-electrospray ionization-mass spectrometry of intact proteins. *J Chromatogr A* 2010, 1217, 7605-11.
35. Chen X., Zaro J. L., Shen W. C. Fusion protein linkers: property, design and functionality. *Adv Drug Deliv Rev* 2013, 65, 1357-1369.
36. Piche-Nicholas N. M., *et al.* Changes in complementarity-determining regions significantly alter IgG binding to the neonatal Fc receptor (FcRn) and pharmacokinetics. *MAbs* 2018, 10, 81-94.
37. Jensen P. F., *et al.* Investigating the interaction between the neonatal Fc receptor and

- monoclonal antibody variants by hydrogen/deuterium exchange mass spectrometry. *Mol Cell Proteom* 2015, 14, 148-161.
38. Burmeister W. P., *et al.* Crystal structure at 2.2 Å resolution of the MHC-related neonatal Fc receptor. *Nature* 1994, 372, 336-343.
 39. Vaughn D. E., *et al.* Identification of critical IgG binding epitopes on the neonatal Fc receptor. *J Mol Bio* 1997, 274, 597-607.
 40. Huber A. H., *et al.* Crystallization and Stoichiometry of Binding of a Complex between a Rat Intestinal Fc Receptor and Fc. *J Mol Bio* 1993, 230, 1077-1083.
 41. Martin W. L., Bjorkman P. J. Characterization of the 2:1 Complex between the Class I MHC-Related Fc Receptor and Its Fc Ligand in Solution. *Biochemistry* 1999, 38, 12639-12647.
 42. West A. P., Bjorkman P. J. Crystal Structure and Immunoglobulin G Binding Properties of the Human Major Histocompatibility Complex-Related Fc Receptor. *Biochemistry* 2000, 39, 9698-9708.
 43. Soltermann F., *et al.* Quantifying Protein-Protein Interactions by Molecular Counting with Mass Photometry. *Angew Chem* 2020, 59, 10774–10779.
 44. Sun Y., *et al.* Antigen physiochemical properties allosterically effect the IgG Fc-region and Fc neonatal receptor affinity. *MAbs* 2020, 12(1):1802135
 45. Skolaut A., Schlothauer T. Fc-Region variants with modified FcRn- and maintained protein A-binding properties Patent WO WO2015107026A1, July 23, 2017.





Chapter 9

Structural and functional characterization of SARS-CoV-2 RBD domains produced in mammalian cells

**Christoph Gstöttner¹, Tao Zhang¹, Anja Resemann², Sophia Ruben³, Stuart Pengelley²,
Detlev Suckau², Tim Welsink³, Manfred Wuhrer¹,
Elena Domínguez-Vega¹**

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

²Bruker Daltonik GmbH, Bremen, Germany

³InVivo BioTech Services GmbH, Hennigsdorf, Germany

Reprinted with permission from *Analytical Chemistry*, 2021, 93, 17, 6839–6847, DOI: 10.1021/acs.analchem.1c00893

Copyright © 2021 The Authors. Published by American Chemical Society

9.1 Abstract

As the SARS-CoV-2 pandemic is still ongoing and dramatically influences our life, the need for recombinant viral proteins for diagnostics, vaccine development, and research is very high. The spike (S) protein, and particularly its receptor binding domain (RBD), mediates the interaction with the ACE2 receptor on host cells and may be modulated by its structural features. Therefore, well characterized recombinant RBDs are essential. We have performed an in-depth structural and functional characterization of RBDs expressed in Chinese hamster ovary (CHO) and human embryonic kidney (HEK293) cells. To structurally characterize the native RBDs (comprising N- and O-glycans and additional post-translational modifications) a multilevel mass spectrometric approach was employed. Released glycan and glycopeptide analysis were integrated with intact mass analysis, glycan-enzymatic dissection and top-down sequencing for comprehensive annotation of RBD proteoforms. The data showed distinct glycosylation for CHO- and HEK293-RBD with the latter exhibiting antenna fucosylation, higher level of sialylation and a combination of core 1 and core 2 type O-glycans. Additionally, using an alternative approach based on N-terminal cleavage of the O-glycosylation, the previously unknown O-glycosylation site was localized at T323. For both RBDs, the binding to SARS-CoV-2 antibodies of positive patients and affinity to ACE2 receptor was addressed showing comparable results. This work not only offers insights into RBD structural and functional features but also provides an analytical workflow for characterization of new RBDs and batch-to-batch comparison.

9.2 Introduction

Since the outbreak of COVID-19, the SARS-CoV-2 virus has infected more than 100 million individuals and influences our daily lives. The coronavirus is an enveloped RNA virus containing three different structural proteins in the membrane, the envelop (E) protein, the membrane (M) protein and the spike (S) glycoprotein.¹ The S protein is heavily glycosylated and forms a trimer on the SARS-CoV-2 surface. Each S protein carries 22 *N*-glycosylation sites² and consists of an S1 and an S2 subunit. Whereas the S2 subunit is necessary for membrane fusion, the S1 subunit directly interacts with the angiotensin converting enzyme 2 (ACE2) receptor in the human respiratory tract and facilitates the entry into the host cell.³ In particular, the receptor binding domain (RBD) of the S1 subunit mediates the interaction with the ACE2 receptor.⁴ This domain carries two *N*-linked glycans at positions N331 and N343 and, depending on the source, one or two *O*-linked glycosylation sites (T323/S325) are occupied.^{2, 5, 6} The predicted *O*-glycosylation site at T323 is not present in the S protein of SARS-CoV-1 and has been hypothesized to have an influence on the interaction with the ACE2 receptor⁷ and may be critical for conformational changes of the RBD.⁸ Further studies have shown the role of glycosylation on ACE2-RBD binding. Crystallographic structures have confirmed the interaction between RBD and ACE2 receptor and highlighted the importance of ACE2 glycosylation in the interaction.⁹ However, the RBD used was expressed in insect cells, and the role of RBD glycosylation was not addressed.⁹ In a more recent publication the interaction of the *N*-glycan at position N343 of the RBD with the ACE2 receptor has been suggested by modeling the interaction using atomistic molecular dynamic simulations.¹⁰ Another study revealed that the infectivity decreased by 1200 times (decrease of relative infectivity to 0.083%) when both *N*-glycosylation sites (N331 and N343) of the spike protein were silenced compared to the wild type versions. These findings suggest that glycosylation is involved in the binding to the receptor – either directly or by providing conformational stabilization.¹¹ In line with its role in virus-ACE2 interaction, the RBD is the primary target of neutralizing antibodies.¹² Interestingly, an antibody named S309, pulled from serum from a SARS-CoV-1 recovered patient, bound an epitope containing also glycosylation.¹³ This binding site is also conserved in SARS-CoV-2 and was predicted as an epitope for neutralizing antibodies.¹⁴ The neutralizing S309 antibody was shown to bind to a protein epitope (331-344) in combination with the glycan at N343. The antibody showed especially strong interaction with the core fucose and to a minor extent with the remaining glycan structure.¹³ All these findings emphasize that the assessment of RBD glycosylation is highly important.

Recombinantly produced S proteins are essential tools in the fight against SARS-CoV-2, contributing to a further understanding of the interaction mechanism, providing efficient components for diagnostic purposes, and helping in vaccine development.¹⁵ However, it is important to understand that glycosylation and other structural characteristics may differ considerably between different biotechnologically produced proteins and their natural forms. Considering the relevance of RBD glycosylation on ACE2 binding and recognition by

neutralizing antibodies, the use of well-characterized S proteins is essential. The S protein, in particular, has been produced as full-length protein as well as in a short version containing the RBD.¹⁶ Site-specific glycosylation analysis of the 22 *N*-glycosylation sites of the recombinant S glycoprotein expressed in human embryonic kidney 293 (HEK293) cells showed mainly complex type glycans but also, at certain glycosylation sites, high mannose structures and in lower amounts some hybrid structures.^{2, 5} In particular, the two *N*-glycosylation sites (N331 and N343), which are located in the RBD were found to carry mainly complex type glycans as well as spurious amounts of high mannose glycans. Similarly, for the S1 subunit recombinantly produced in HEK293 cells, mainly complex type glycans were found at both sites.⁵ As predicted by Uslupehlivan et al.⁷ an *O*-glycosylation at the position T323 and/or at S325 was found. Whereas one report, analyzing the whole S protein recombinantly produced, showed only trace levels of *O*-glycans² another study detected high levels of *O*-glycosylation.⁶ Still, in these studies the localization of the *O*-glycosylation site at T323 and/or S325 was not possible. *O*-glycosylation at these two positions is also thought to be important to stabilize the conformation of the RBD or to introduce conformational changes⁸. All these studies have been performed using recombinant versions of the spike protein or subunits thereof. Interestingly, upon expression in HEK293 cells the amount of sialic acids varied from low to high depending on whether the complete S protein or only the S1 subunit was produced⁵, suggesting that the *N*- and *O*-glycosylation is dependent on the context of the protein (S, S1 or RBD only). So far only terminal glycan epitopes of HEK293 produced RBD have been studied by NMR, however an assessment of the entire *N*-glycan structure, composition and the relative quantification of the *N*-glycans could not be achieved.¹⁷ Furthermore, the *O*-glycosylation was neglected completely. Also, a characterization of the intact RBD is still missing, providing information on the combination of the glycans as well as on additional protein backbone modifications.

Here we present an in-depth structural and functional characterization of two commercially available SARS-CoV-2 RBDs produced in two different expression systems, HEK293 and Chinese hamster ovary (CHO) cells. To achieve comprehensive structural information, a multilevel characterization was performed (**Figure 1**). Both RBD samples were initially analyzed at the intact level by top-down sequencing using MALDI in-source-decay (MALDI-ISD) mass spectrometry (MS)^{18, 19} and by sheathless capillary electrophoresis (CE)-MS after full *N*- and *O*-deglycosylation to establish their protein sequences and putative non-glycan modifications. With the established sequences and the help of sequential glycosidase treatment, the glycoforms on both *N*- and the *O*-glycosylation sites were assigned. Our findings were confirmed by glycopeptide analysis and the analysis of released *O*-glycans by porous graphitized carbon (PGC) nano-LC-ESI-MS/MS. To assess functional differences between the two RBD samples, we determined their binding characteristics to ACE2 and sera of patients who recovered from a previous SARS-CoV-2 infection.

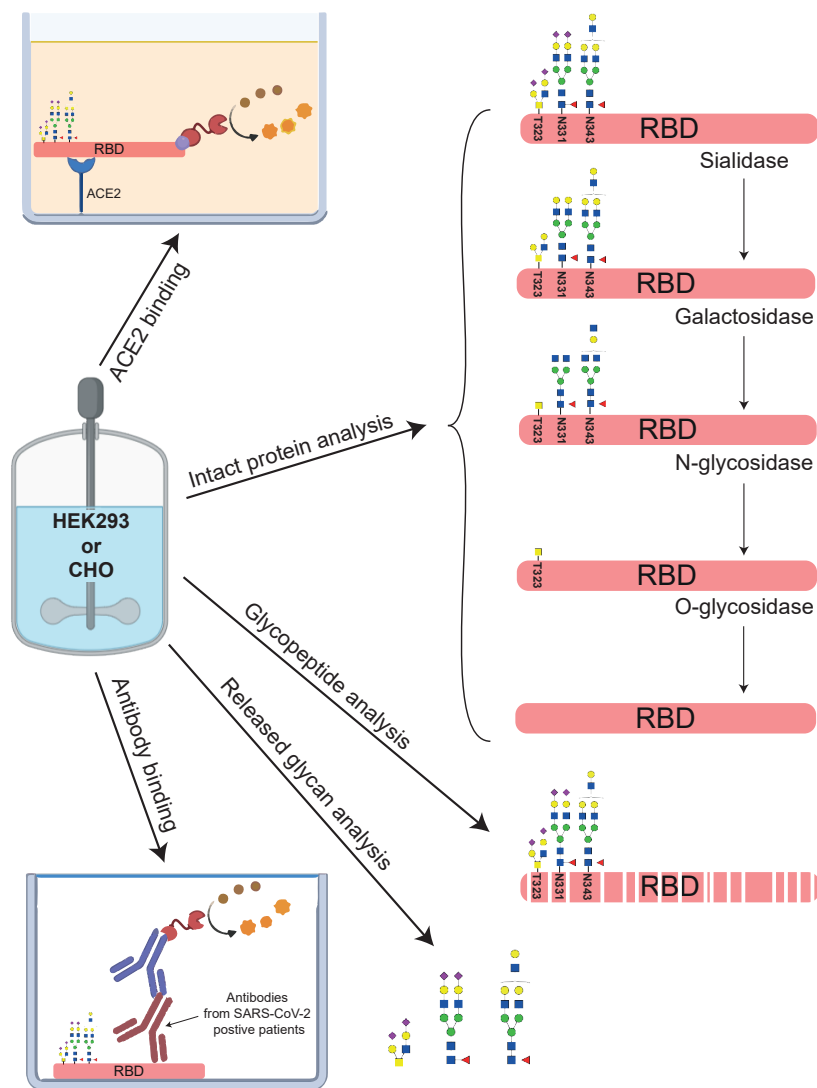


Figure 1. Multilevel characterization of CHO- and HEK293-RBD. Next to intact protein analysis using CE-MS and MALDI-MSD-MS the samples are structurally characterized by glycan dissection, glycopeptide analysis and released glycan analysis. Their functional characterization was performed by measuring their binding characteristics to ACE2 and anti-SARS-CoV-2 antibodies.

9.3 Experimental Section

9.3.1 Reagents and samples

Reagents used for this study were at least of analytical grade, for more details consult

Information S1. Recombinant RBDs (Wuhan-Hu-1-isolate (MN908947)), either transiently expressed in HEK293 or stably expressed in CHO cells, were used (InVivo Biotech Services, Henningsdorf, Germany). The constructs contained the amino acid sequence 319 to 541 with a C-terminal 6xHis-Tag. Recombinant RBDs were purified using immobilized metal affinity chromatography and a size exclusion polishing step. The samples were stored in 20 mM sodium phosphate, 300 mM NaCl, pH 7.2. Glycosidases SialEXO (sialidases α 2-3, α 2-6 & α 2-8), GalactEXO (galactosidases β 1-3 and β 1-4), OglyZOR (endo- α -N-acetylgalactosaminidase), OpeRATOR (*O*-protease), α 1-2 fucosidase and α 1-3,4 fucosidase were obtained from Genovis (Lund, Sweden). Peptide *N*-glycosidase F (PNGaseF) was purchased from Roche Diagnostics (Mannheim, Germany).

9.3.2 Intact RBD analysis by sheathless CE-MS

Samples for intact protein analysis were buffer exchanged with Tris pH 6.8 using 10 kDa Vivaspin MWCO filters Sartorius (Göttingen, Germany) to a final concentration of 1 μ g/ μ L. Subsequently the enzymes SialEXO, GalactEXO, OglyZOR or OpeRATOR were added according to the producers specifications and incubated overnight at 37°C. PNGaseF was added in a ratio 1:5 (v/v) and incubated overnight at 37°C. For removing antenna fucosylation 20 μ g of HEK293-RBD sample was incubated with 4.4 μ g fucosidase α 1-2 or 8.8 μ g fucosidase α 1-3,4 overnight at 37°C. Afterwards all samples were buffer exchanged with 100 mM ammonium acetate pH 3.0 and subsequently analyzed by sheathless CE-MS. For sheathless CE measurements a Sciex CESI 8000 instrument combined with a Bruker Impact qTOF-MS was used. Bare-fused silica CE capillaries were obtained from Sciex (Framingham, MA) and subsequently coated with polyethylenimine (PEI).²⁰ Afterwards the sample was injected (15 s, 5 psi) and analyzed by applying 20 kV (reversed polarity) at 20°C for 45 min. For detailed CE and MS settings see **Information S2**. Intact RBD data were deconvoluted using the DataAnalysis 5.3 (Bruker Daltonics) maximum entropy algorithm. Deconvoluted mass spectra were baseline subtracted and smoothed using the Gaussian smoothing function with a width of 0.5 Da and one cycle. Glycoforms were assigned based on the average masses observed and observed mass shifts with enzyme treatments. Reconstructed intact spectra from bottom-up data were simulated using an algorithm in R developed by Yang et al. which is publicly available and can be found at <https://github.com/Yang0014/glycoNativeMS>.²¹ For our data, instead of simulating the charge state distribution of the protein we simulated the deconvoluted mass spectra, so no correction for the relative abundance of different charge states needed to be performed.

9.3.3 MALDI-ISD top-down protein sequence analysis

Fully deglycosylated RBD samples were reduced for 30 min at 50°C using DTT. 2 μ L of the reduced sample were spotted on a hydrophilic anchor of an MTP BigAnchor sample plate and incubated. After 2 min, the remaining droplet was removed and the spot was

washed using 0.1% TFA in water. Subsequently, 1 μ L of sDHB matrix solution (25 μ g/ μ L in 50% acetonitrile/49.9% water/ 0.1% TFA) was deposited on the dried sample spot. MALDI-MS spectra were acquired with a rapifleX MALDI-TOF MS instrument (Bruker). Detailed MS settings and spectra assignment can be found in **Information S3**.

9.3.4 Glycopeptide analysis by RPLC-MS/MS

For glycopeptide generation a double digestion with trypsin followed by elastase was performed. In short, samples were reduced (45 mM DTT), alkylated (100 mM IAA) and the reaction stopped (45 mM DTT). After overnight tryptic digestion the reaction was stopped (90°C) and elastase was added. After overnight incubation with elastase the reaction was stopped (formic acid). A more detailed description can be found in **Information S4**. The digested samples were separated using a nanoElute (Bruker) nanoflow UHPLC equipped with an aurora 25 cm x 75 μ m C18 column with a particle size of 1.6 μ m (IonOpticks, Parkville, Victoria, Canada) and analyzed on a timsTOF PRO (Bruker). Information about the LC and MS parameters can be found in **Information S5**. MS/MS data of N-linked glycopeptides were generated following Hinneburg et al.²¹ and processed using DataAnalysis 5.3 and MGF peak lists were imported into BioPharma Compass 2021 (Bruker) and further analyzed using the glycopeptide analysis workflow. This workflow includes MS/MS spectra classification using typical fragmentation patterns to determine the peptide mass and the glycan mass. The peptide sequences were then identified using the theoretical digest feature of the software. In a second search, the classified MS/MS spectra were analyzed using the GlycoQuest search engine within BioPharma Compass 2021 software. For quantification the most intense isotope of each charge state ($[M+H]^+$, $[M+2H]^{2+}$, or $[M+3H]^{3+}$) was extracted and the results normalized to the total peak intensity of all glycopeptides within one sample to 100%.

9.3.5 Released O-glycan analysis by PGC nano-LC-ESI-MS/MS

Released O-glycan alditols from RBD samples were prepared using a 96-well plate sample preparation method as previously described.²³ In brief, after N-glycan removal (PNGaseF, 2 U, overnight incubation at 37 °C), the O-glycans were released via reductive β -elimination. The analysis of O-glycan alditols was performed on an Ultimate 3000 UHPLC system (Dionex/Thermo) equipped with an in house-packed PGC trap column (5 μ m Hypercarb, 320 μ m x 30 mm) and an in-house-packed PGC nano-column (Grace Discovery Sciences, Columbia, MD, USA) (3 μ m Hypercarb 100 μ m x 150 mm) coupled to an amaZon ETD speed ion trap (Bruker) following a method described previously.²³ More detailed information can be found in the supplementary information (**Information S6**). Structures of detected glycans were confirmed by MS/MS in negative mode²⁴. Glycan structures were assigned on the basis of the known MS/MS fragmentation patterns in negative-ion mode²⁵⁻²⁷, elution order, and general glycobiological knowledge, with help of Glycoworkbench²⁸ and Glycomod

software.²⁹ Relative quantification of individual glycans was performed by normalizing the total peak area of all glycans within one sample to 100%.

9.3.6 SARS-CoV-2-IgG ELISA with RBD antigens

The antigens (intact or glycosidase treated HEK293-RBD or CHO-RBD) were used to coat immunoassay plates at a concentration of 2 µg/mL. Subsequently, blocking was performed with blocking buffer containing 1% bovine serum albumin (BSA). Serum collected >10 weeks after onset of first symptoms from 12 SARS-CoV-2 PCR positive tested donors was applied to the RBD-coated wells at a dilution of 1:101. As a negative control, a 1:101 diluted serum pool containing 10 sera taken from healthy individuals in the year 2012 was used. The wells were three times washed with 250 µl PBS containing 0.1% Tween 20. Subsequently, bound IgG from sera was detected by anti-human IgG-horseradish peroxidase (HRP) and developed with tetramethylbenzidine. The reaction was stopped using sulfuric acid. The absorbance was measured at 450 nm using a microplate reader. In order to determine the linear correlation between the absorption values, the Pearson correlation was calculated using GraphPad Prism 9. Linear regression was employed for the visualization of the correlation. For this analysis, the negative control value was excluded.

9.3.7 ACE2 receptor binding assay

Intact and glycosidase treated RBD, S1 subunit and the S protein (InVivo Biotech) were biotinylated using 10 molar excess of NHS-LC-Biotin (Thermo Scientific). Immunoassay plates were coated with 2.5 µg/mL recombinant angiotensin converting enzyme-2 (ACE2, Sigma Aldrich). Subsequently, the assay plate wells were blocked using blocking buffer containing 1% BSA. Dilutions of the biotinylated antigens ranging from 1 µg/mL to 0.001 µg/mL were applied to the ACE2-coated assay plate wells in duplicates. The wells were three times washed with 250 µL PBS containing 0.1% Tween 20. Bound biotinylated antigens were detected using streptavidin peroxidase conjugate (Roche) and developed with tetramethylbenzidine. The reaction was stopped using sulfuric acid. The absorbance was measured at 450 nm using a microplate reader. GraphPad Prism 9 was used to plot log(dose) response curves (variable slope, four parameters) and to compute nonlinear fits which were utilized to calculate the half-maximal concentrations (EC_{50}).

9.4 Results and Discussion

9.4.1 Structural characterization of CHO- and HEK293-RBD

We characterized the RBD domains (amino acids 319-541 containing a C-terminal His-tag) produced in CHO and HEK293 cells. This domain contains two *N*-glycosylation sites at positions N331 and N343 as well as two potential *O*-glycosylation sites at T323 and S325. Analysis of the intact RBDs revealed a complex pattern of signals comprising different *N*-

and *O*-glycans and additional protein backbone modifications (see section assessment of RBD N-glycosylation). Therefore, to unravel this heterogeneity and achieve comprehensive structural characterization, we used a multilevel approach (**Figure 1**). Next to classical released glycan and glycopeptide approaches, we applied a step-by-step dissection of glycans at the intact level similar to the work of Wohlschläger et al. for the biopharmaceutical etanercept.³⁰

9.4.2 Characterization of the protein backbone after *N*- and *O*-glycan removal

To get information on the integrity of the protein backbone, the RBD proteins were treated with PNGaseF to remove the *N*-glycans and with an endo- α -*N*-acetylgalactosaminidase in combination with a mix of sialidases to remove the *O*-glycans. After deglycosylation, the RBDs were analyzed by CE-MS for intact mass and MALDI-MS for top-down sequencing. Analysis of the samples by CE-MS resulted in one main peak in the base peak electropherogram (BPE) corresponding to the RBD. For the RBD produced in CHO cells, the observed averaged mass (26033.9 Da) was 119.0 Da higher than the theoretical mass calculated solely based on the amino acid sequence (25914.9 Da) (**Figure 2A**). As the RBD contains a free cysteine at position C538³¹, it was presumably cysteinylated (+119.1 Da). This modification is often observed for free light chains during antibody production using CHO cells.³² After reducing the disulfide bonds with DTT, the mass of the deglycosylated and completely reduced protein was 25923.0 Da, which is consistent with the expected theoretical mass (25923.0 Da) confirming the presence of cysteinylation (data not shown). For the RBD expressed in HEK293 cells, a deconvoluted mass of 26144.9 Da was observed (**Figure 2B**). Considering a cysteinylation of the free cysteine also in the HEK293-RBD, an additional mass difference of 110.9 Da was observed compared to the theoretical mass indicating additional modifications.

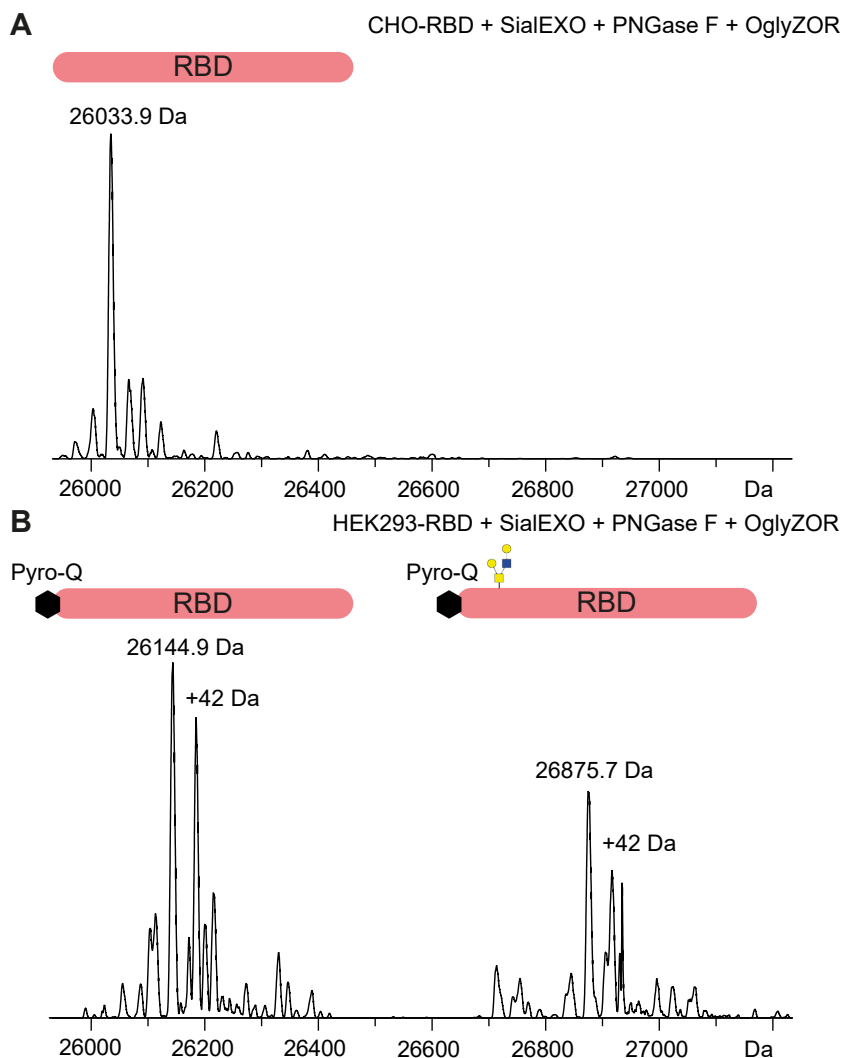


Figure 2. Deconvoluted ESI mass spectra of the main peak observed after sheathless CE-MS for A) CHO-RBD and B) HEK293-RBD. HEK293-RBD carries an additional N-terminal pyroglutamate. Blue square, *N*-acetylglucosamine; yellow square, *N*-acetylgalactosamine; yellow circle, galactose.

Both RBD samples were analyzed at the intact level by top-down sequencing using MALDI-MS/MS after full reduction and *N*- and *O*-deglycosylation. MALDI-MS/MS provides large terminal sequence tags and detects terminal (and internal) PTMs from intact proteins. For both RBDs we detected a C-terminal sequence tag of 73 amino acids (RMS error of 0.028 Da) verifying the expected C-terminus of the sequence. The N-terminal sequence tag, however, was not consistent between both RBDs. While the CHO-RBD sequence was in full agreement with the MALDI-MS/MS spectrum (**Figure S1A**), the HEK293-RBD sequence was found to be N-terminally extended by pyro-Glu (mass shift of 111.03 Da). 52% of the sequence were confirmed this

way including an N-terminal sequence tag of 46 residues for the non-glycosylated and 29 residues for the *O*-glycosylated form (**Figure S1B**). Of note, the HEK293-RBD came with a slightly different sequence compared to the CHO-RBD, resulting in an additional glutamine at the *N*-terminus after cleavage of the signaling peptide (**Figure S1B**). Taking pyroglutamic acid formation into consideration, the measured mass perfectly fits the theoretical mass (26145.1 Da). The N-terminal pyroglutamic acid formation was further confirmed by bottom-up analysis (data not shown). In addition, we found a species migrating before the main signal in the BPE with a +42.3 Da mass difference (26187.4 Da). This difference in mass might correspond to acetylation, however, it could not be confirmed by bottom-up analysis after trypsin digestion. This species was only observed in HEK293-RBD and not in CHO-RBD.

9.4.3 RBD *O*-glycan characterization

The RBD has two potential *O*-glycosylation sites at position T323 and S325. To characterize the *O*-glycans, the de-*N*-glycosylated RBD (after treatment with PNGaseF) were analyzed by CE-MS (**Figure S2**) and the released *O*-glycans by PGC nano-LC-ESI-MS/MS (**Figure 3**), respectively. **Table S1** summarizes the obtained results. In CHO-RBD mainly core 1 structures with 2 sialic acids (H1N1S2) were observed. HEK293-RBD showed a much more diverse *O*-glycosylation pattern with a core 1 structure with 2 sialic acids as the main signal. Additionally, several core 2 structures, with and without fucose as well as with sulfation were detected. These data are in accordance with the data in **Figure 2B** for the sample treated with an endo- α -*N*-acetylgalactosaminidase, where an additional signal of 26875.7 Da was observed. This mass corresponds to an H2N2 modification (theoretical mass 26875.8 Da) which is presumably a core 2 *O*-glycan which cannot be cleaved by the endo- α -*N*-acetylgalactosaminidase. Similar glycoforms were not detected in CHO-RBD. The position of fucoses, either to the terminal galactose or the *N*-acetylglucosamine were confirmed by MS fragmentation and treatment with different fucosidases (**Figure S3**).

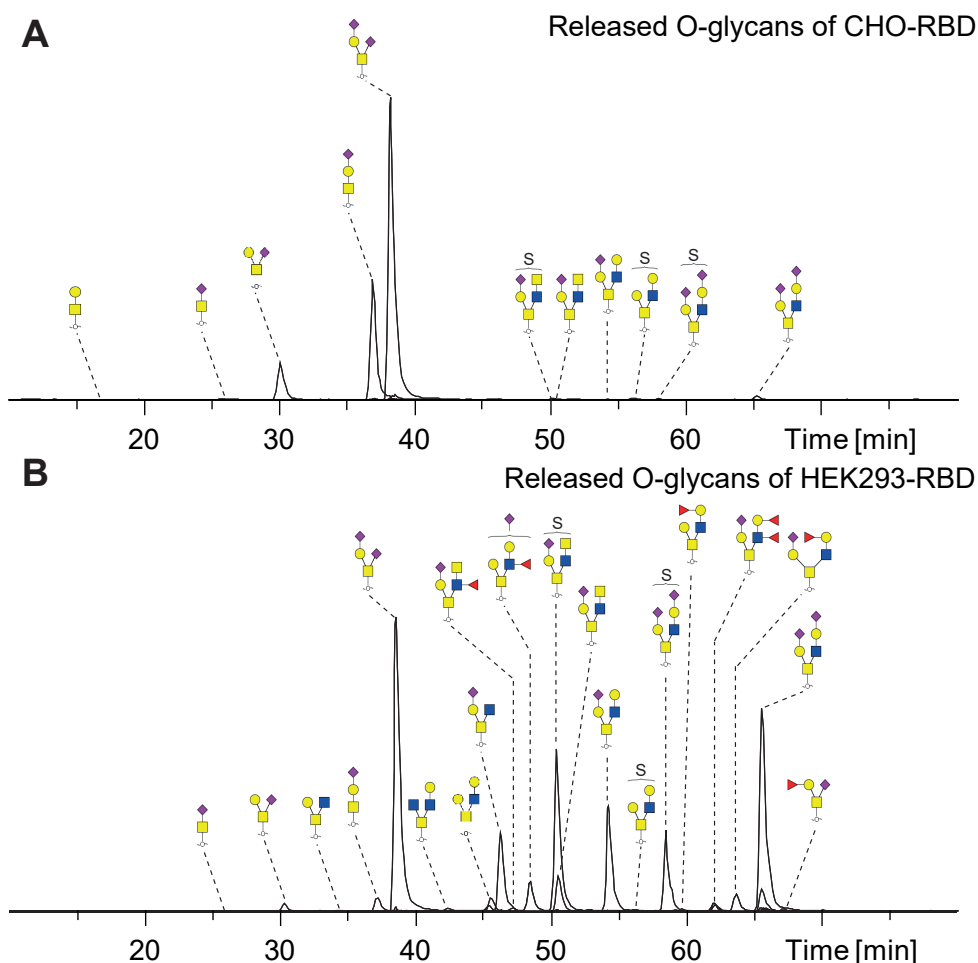


Figure 3. PGC nano-LC-ESI-MS/MS of released O-glycans from A) CHO-RBD and B) HEK293-RBD. Yellow square, *N*-acetylgalactosamine; yellow circle, galactose; blue square, *N*-acetylglucosamine; red triangle, fucose; purple diamond, *N*-acetylneuraminic acid (sialic acid), S, sulfate.

Hitherto published S protein bottom-up studies have not revealed the *O*-glycan attachment site, leaving both T323 and S325 as valid options.^{2, 6} To resolve this, we follow a different strategy based on *N*-glycan removal (with PNGaseF) and *O*-protease cleavage using the enzyme OperATOR in combination with a mix of sialidases. The *O*-protease cleaves the protein at the N-terminal site of an *O*-glycosylation site (**Figure S3**). This would result in a loss of the amino acids 319-322 from the RBD if the *O*-glycosylation is on T323 or 319-324 in case S325 carries the *O*-glycan. As shown in **Figure S4A**, only one signal with a mass of 25919.0 Da was observed for the CHO material, which correlates to the RBD cleaved at T323 with a core 1 *O*-glycan H1N1 (theoretical mass 25918.8 Da). No RBD cleaved at S325 was observed. In HEK293-RBD, the same signal with a core 1 glycan H1N1 was observed also indicating that

the glycosylation is located at T323 (**Figure S4B**). Next to this signal, an amount of uncleaved RBD with presumably core 2 *O*-glycan structures was detected. This is in line with a recent article that shows that the used *O*-protease can cleave N-terminal to core 1 but not core 2 glycan structures, similar to endo- α -*N*-acetylgalactosaminidase.³³ To confirm that the core 2 structures are located at T323 we performed MALDI-MS analysis using super-DHB as a matrix. In contrast to CID fragmentation in bottom-up analysis, MALDI-MS is known to result in mainly singly charged c- and z+2 ions with labile modifications remaining intact, which allows to localize *O*-linked glycosylation sites. MALDI-MS of CHO-RBD after *N*-glycan and sialic acid removal showed N-terminal fragments (c-ions) with core 1 structure (H1N1) for the site T323 and no additional glycosylated fragments at S325 (+365.1 Da) could be detected confirming the presence of the *O*-glycosylation at T323 (**Figure S5A**). No fragments of T323 without glycosylation were observed indicating full site occupancy. The analysis of HEK293-RBD comprising only core 2 structure *O*-glycans (after *N*-glycan and core 1 *O*-glycan removal) clearly showed that the core 2 structures are also located on T323 (**Figure S5B**). In conclusion, the combination of intact analysis of *O*-protease treated RBDs together with MALDI-MS allowed location of the *O*-glycans to the T323 with high confidence.

9.4.4 Assessment of RBD *N*-Glycosylation

We studied the two *N*-glycosylation sites N331 and N343 using a bottom-up glycopeptide approach combined with a step-by-step dissection of glycans at the intact level.³⁰ Whereas the glycopeptide data yielded the glycan composition per site, we studied the combination of these glycans at the intact level.²¹

The intact RBDs were first incubated with a mixture of sialidases and galactosidases. These enzymes removed the sialic acids and the terminal galactoses on the *N*- as well as *O*-glycans, resulting in considerably simplified deconvoluted mass spectra as shown in **Figure S6** and **Figure S7**. Overall, mainly fucosylated complex type glycans were observed for both RBDs (>90.0%). This is in line with previous studies on HEK293 produced intact S protein or S1 subunit.^{2, 5} The dissection of the terminal galactoses also permits direct distinction between *N*-acetyllactosamine (LacNAc) repeats and additional antenna. CHO-RBD showed a higher relative abundance of LacNAc repeats than HEK293-RBD. Additionally, CHO-RBD showed a higher antennarity, with two triantennary structures as the most abundant glycoforms, contrary to HEK293 material with di- and triantennary glycans as major signals. This was confirmed by the analysis of the glycopeptides (**Figure S8**, **Table S2** and **S3**). In general, more di- and triantennary structures were observed for HEK293-RBD (N331: 78.4% and N343: 84.5%) compared to the CHO-RBD (N331: 48.1% and N343: 72.1%). Therefore, CHO-RBD showed a larger contribution of tetraantennary structures and LacNAc repeats (N331: 42.7% and N343: 24.4%) compared to HEK293-RBD (N331: 15.7% and N343: 9.9%). Between both glycosylation sites a difference in the number of LacNAc repeats and high-antennary structures was observed with minor amounts on N343 compared to N331.

In addition, we found spurious amounts of hybrid type glycans only at N343 and only in HEK293-RBD (0.8%), similar to previous publications.^{2,5} Regarding high mannose glycans, low amounts were detected for both RBDs. In the case of HEK293-RBD, Man5 and Man6 glycans with and without phosphorylation were observed, whereas in the case of CHO-RBD Man5 and Man6 glycans carrying an additional phosphate were detected. For HEK293-RBD the high mannose and phosphorylated high mannose structures show combined abundances of 3.1% (N331) and 1.3% (N343) in line with the findings of Watanabe et al.² for full-length S protein. For CHO-RBD the distribution was more skewed with 7.7% (N331) and 0.9% (N343).

Besides, in the deconvoluted mass spectrum of the HEK293-RBD treated with sialidase and galactosidase a pattern of signals with a mass shift of +308.1 Da was observed. A combination of one galactose and one fucose could explain this mass shift. It was shown in literature that β -galactosidases are not able to remove the terminal galactose if an antenna fucose, either linked to the galactose itself or to the *N*-acetylglucosamine, is present.³⁴ Therefore, we incubated the HEK293-RBD either with α 1-2 fucosidase and in parallel with α 1-3,4 fucosidase to remove fucoses linked to the galactose and *N*-acetylglucosamine, respectively. Incubation with these fucosidases allowed the β -galactosidase to remove the terminal galactose. As shown in **Figure S3**, after removing the differently linked antenna fucoses, the +308.1 signals disappear completely confirming antenna fucosylation and providing information on the linkage of the antenna fucose with a predominant linkage to the *N*-acetylglucosamine. Additionally, antenna fucosylation in HEK293-RBD was confirmed by glycopeptide analysis (26.7% and 33.8% of antenna fucosylation on N331 and N343, respectively) (**Figure S9, Table S2 and S3**). No antenna fucosylation was found for CHO-RBD with any of the approaches.

The analysis of the intact RBDs without any previous enzymatic treatment resulted in a very complex mass spectrum (**Figure 4 and S10**).

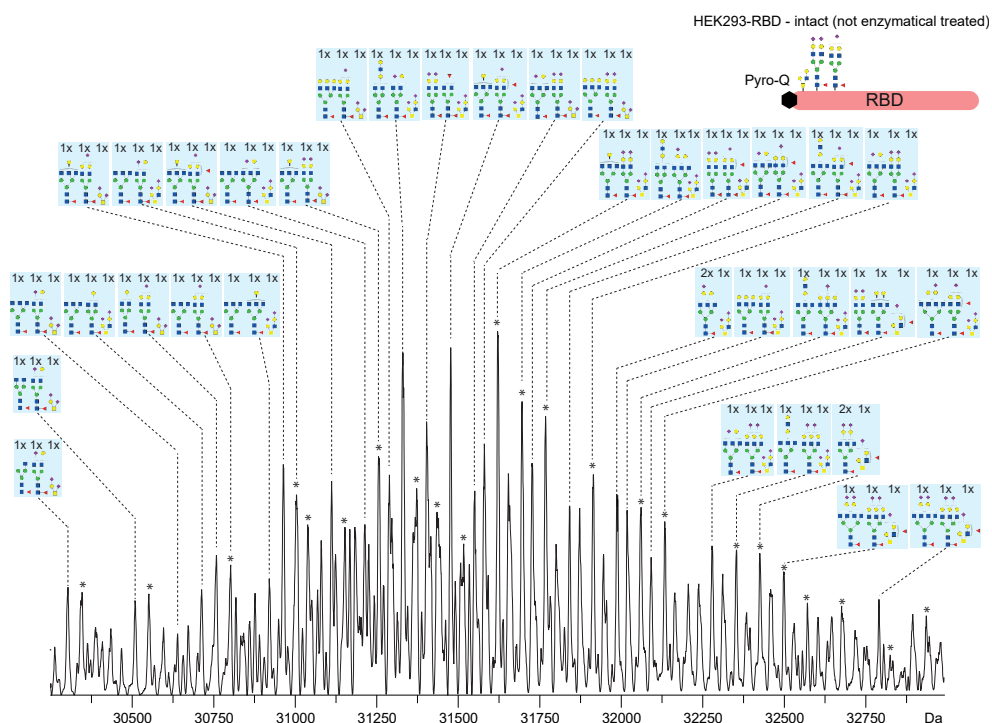


Figure 4. Deconvoluted mass spectra of the intact RBD (not enzymatically treated) produced by HEK293 cells. The assignments were based on previous enzyme treatments, mass and glycopeptide as well as released *O*-glycan data. Peaks marked with an asterisk * are presumably the acetylated variant of RBD. Yellow square, *N*-acetylgalactosamine; yellow circle, galactose; blue square, *N*-acetylglucosamine; red triangle, fucose; purple diamond, *N*-acetylneuraminic acid (sialic acid).

Based on the information obtained for the released *O*-glycans, glycopeptide and enzymatically-treated intact RBDs the spectra were confidently assigned. In particular, the HEK293-RBD exhibited a large heterogeneity. In addition to the variability resulting from the acetylation and antenna fucosylation, which were not observed in CHO-RBD, also a higher degree of sialylation was observed for HEK293-RBD. This was also supported by the glycopeptide data in which the CHO-RBD showed only 2.5% or 0.8% sialylated glycans, whereas the HEK293-RBD contained 56.4% or 36.3% on N331 or N343, respectively (**Figure S11**, **Table S2** and **S3**). Furthermore, in CHO-RBD only monosialylated species were observed, while in the HEK293-RBD mono-, di- or trisialylated species were detected. Interestingly, these high sialylation levels were not observed by Watanabe et al. who reported only 22% sialylation on N331 and 4% sialylation on N343 for the full length S protein.² This might be due to the different constructs with RBD expression versus S1 subunit or S-protein. A similar effect of lower sialylation on the complete S protein has also been previously reported by comparing the entire S protein expressed in HEK293 cells with the S1 subunit.⁵ Whereas S protein carried either ~ 40% or 10% sialylation on N331 or N343, the S1 subunit carried

80% or 50% on N331 or N343, respectively. These findings highlight that sialylation, which influences the isoelectric point of a protein, can change with the length of protein expressed (S, S1 or RBD) and must be taken into account.

Finally, to support our assignments, the intact mass spectra were reconstructed from the glycopeptide data as described.²¹ **Figure 5A**, shows a very strong correlation for the CHO-RBD confirming the assignments. For HEK293-RBD reconstructed mass spectrum showed a shift towards lower masses compared to the intact profile (**Figure 5B**). Proteins with higher sialylation levels often show a discrepancy between the intact and the glycopeptide-reconstructed mass spectra²¹ which could be attributable to a non-random combination or to a biased ionization efficiency of sialylated glycopeptides.

9.4.5 Functional characterization of CHO- or HEK293-RBD

Next to the structural characterization of the RBD also a functional characterization was performed. To assess the RBD functionality, an ACE2 receptor binding assay and a binding assay for anti-RBD antibodies from COVID-19 patient sera were performed.

9.4.6 SARS-CoV-2 antibody binding assay

The binding of anti-SARS-CoV-2-IgG antibodies in sera taken from 12 Covid-19 patients more than 10 weeks after the onset of the first symptoms and RBD was determined using an ELISA assay. For the assay, the intact CHO- or HEK293-RBD, as well as deglycosylated RBD samples, were used. As a negative control, a pool of 10 sera collected in 2012 was used. As expected, higher absorption values were observed for the 12 COVID-19 patient sera compared to the negative control indicating binding of specific antibodies to the RBDs (**Figure S12**). Both intact RBD samples showed similar absorption values, while absorption values were slightly elevated when the deglycosylated RBDs were employed. Also, for the negative controls, the deglycosylated RBD showed elevated absorption values compared to the intact RBD. This is most probably the result of unspecific binding which may be increased after deglycosylation. This was also reflected in the correlation of the absorption values of the patient sera. The correlation of these values gained using intact RBDs and their deglycosylated version (CHO: R2 0.9681, HEK293: 0.9755) was slightly lower than the correlation obtained using intact RBD produced by CHO and HEK293 cells (R2 0.9894) or deglycosylated CHO- versus deglycosylated HEK293-RBD (R2 0.9905) (**Figure S13**).

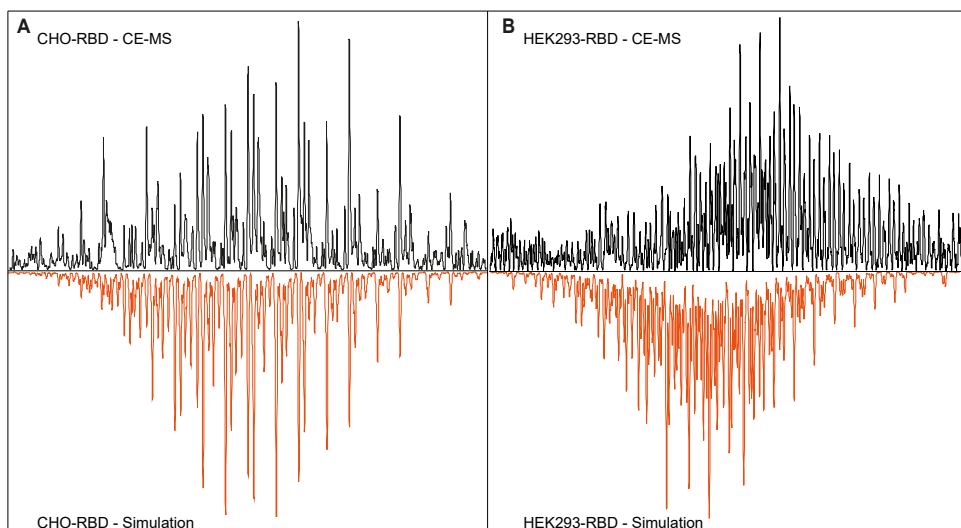


Figure 5. Deconvoluted mass spectra of the RBD intact (non enzymatical treated) (black trace) and in silico simulated mass spectrum (vermillion trace) of RBD produced in A) CHO cells or B) HEK293 cells.

9.4.7 ACE2 receptor binding assay

Using a plate-based ACE2 receptor binding assay, a dose-dependent binding of both CHO- and HEK293-RBD was observed (**Figure 6A**). Comparing the RBDs expressed in both production systems, similar binding properties were observed with a trend towards a lower EC_{50} value for CHO-RBD. As glycosylation has been hypothesized to have a role in the interaction with ACE2^{10, 11}, we additionally tested the RBDs after deglycosylation. Deglycosylation was found to reduce the binding of the RBDs as reflected in an approximately 2 times increase of the EC_{50} values. Of note, Qianqian et al. reported that the infectivity is reduced to only 0.08% after removal of the *N*-glycosylation sites.¹¹ Our results, although they show a slight variation between the glycosylated and non-glycosylated versions, do not explain this drastic difference in infectivity. Furthermore, the presence of endoglycosidases could also have affected the biotinylation rate resulting in the overall reduced binding values. Supporting the hypothesis of Qianqian, these glycans may be crucial for stabilizing the trimeric spike protein rather than influencing the binding affinity.¹¹

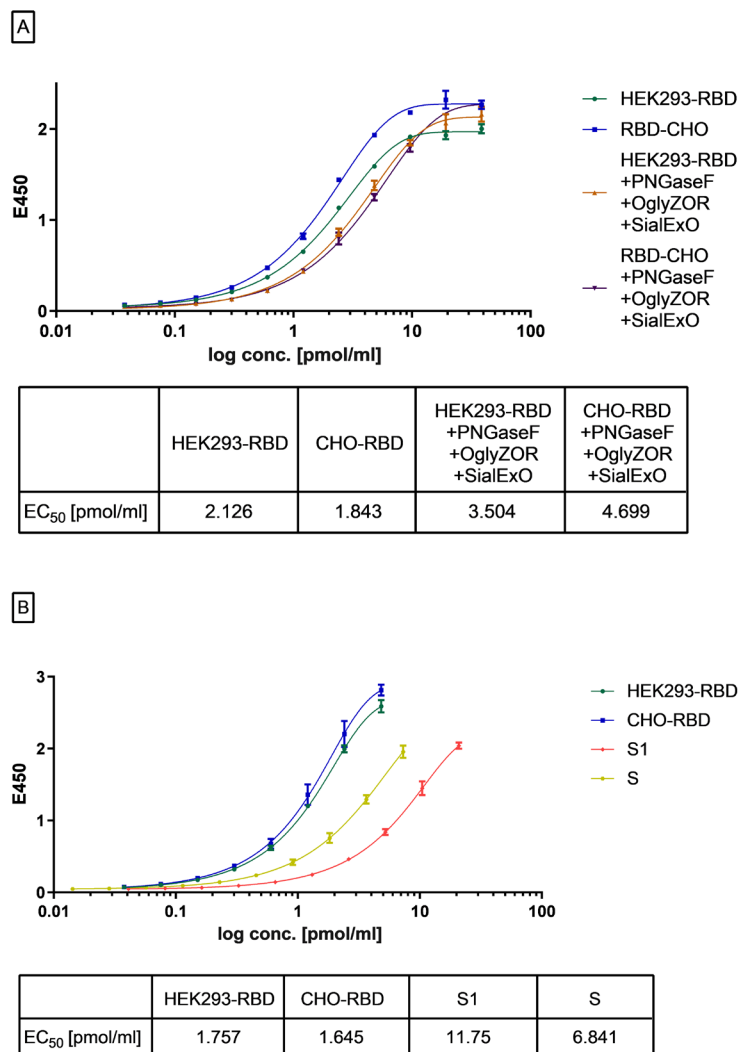


Figure 6. ACE2 binding assay. Dose-response curves of CHO- or HEK293-RBD binding to the ACE2 receptor. Comparison to A) deglycosylated version of CHO- or HEK293-RBD and B) S or S1 subunit (produced in HEK293 cells).

Additionally, we compared the ACE2 receptor binding affinity of both RBDs to the S1 subunit, as well as the intact S protein. The binding affinity of the ACE2 receptor to the RBDs was significantly increased compared to the S1 subunit and the S protein (**Figure 6B**). This might be due to the higher accessibility to ACE2 of the RBD only compared to the S1 subunit or even the S protein. As shown by Casalino et al., in the trimeric S protein the RBD can be in an up or down conformation and, therefore, more or less accessible to the ACE2 binding.⁸ By using the S protein or the S1 subunit, some parts of the RBD might not be accessible or less accessible for

the ACE2 receptor which may explain the reduced binding affinity. These results highlight the importance of a proper selection of recombinant proteins.

9.5 Conclusion

RBD proteoforms were comprehensively characterized by combining intact protein, glycopeptide and released glycan analysis with enzymatic glycan dissection and top-down sequencing. The combination of multiple MS workflows was fundamental for assigning the intact RBD proteoforms. In particular, glycan dissection of the intact protein using sequentially different glycosidases has shown to be very powerful to annotate complex intact RBD spectra. The glycopeptide data, next to providing site-specific information, were used to simulate an *in silico* intact spectrum and corroborate our assignments. This approach was applied to RBD samples from both CHO and HEK293 cells. In case of the low sialylated CHO-RBD a very strong correlation was obtained. However, for the HEK293-RBD the simulated spectrum showed a clear shift to a lower mass associated with the loss of sialic acids or an ionization bias of the glycopeptides. This stresses the importance of assessing intact proteoforms to avoid skewing of the data in any direction. The observed differences in *N*-glycosylation, with higher sialylation levels and antenna fucosylation for HEK293-RBD and low sialylation levels but high-antennary and LacNAc repeat structures for CHO-RBD are typical for the two expression systems. For the *O*-glycans, CHO-RBD showed mainly core 1 type glycans while HEK293-RBD presented a combination of core 1 and core 2 type *O*-glycans. Furthermore, using alternative approaches, such as N-terminal cleavage at the *O*-glycosylation site and MALDI-MS we localized the *O*-glycosylation site to T323, previously unknown. Further steps will focus on validation of the method for batch-to-batch comparison of different RBD batches and for release testing in companies producing RBD based vaccines. Full structural characterization of the S protein instead of RBD would be challenging due to the higher molecular mass and heterogeneity (22 *N*-glycosylation sites). Still, some of the strategies applied in this approach, as complete deglycosylation and analysis of protein backbone or N-terminal cleavage at the *O*-glycosylation sites could provide additional information to previous published works. From a functional point of view, both RBDs showed similar binding to antibodies from COVID-19 patient sera as well as to the ACE2 receptor. The ACE2 binding was increased compared to S protein or S1 subunit, likely due to higher accessibility. After deglycosylation, the binding of RBDs to anti-spike antibodies remained unaffected. For ACE2 a minor decrease in the EC_{50} values was observed, which does not fully explain the infectivity decrease with aglycosylation observed in previous studies. These findings suggest that glycosylation of the RBD plays a role in conformational stabilization rather than affecting binding affinity between ACE2 and RBD. Further studies are warranted on the influence of RBD glycosylation on S conformation, ACE2 binding as well as virus infectivity and biology.

9.6 Supporting information

Supplementary information is available free of charge on the ACS website: DOI: 10.1021/acs.analchem.1c00893

9.7 Acknowledgments

We thank Eckhard Belau and Waltraud Evers for preparing the samples and analysis. This work was supported by the Analytics for Biologics project (Grant agreement ID 765502) of the European Commission.

Notes:

The authors declare the following competing financial interest(s): InVivo BioTech Services is a biotechnology company producing antibodies and proteins, including SARS-CoV-2 antigens.

9.8 References

1. Li H., *et al.* Coronavirus disease 2019 (COVID-19): current status and future perspectives. *Int. J. Antimicrob. Agents* 2020, 55, 105951.
2. Watanabe Y., *et al.* Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* 2020, 369, 330.
3. Hoffmann M., *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 2020, 181, 271-280.e8.
4. Wong S. K., *et al.* A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. *J. Biol. Chem.* 2004, 279 (5), 3197-201.
5. Wang D., *et al.* Comprehensive Analysis of the Glycan Complement of SARS-CoV-2 Spike Proteins Using Signature Ions-Triggered Electron-Transfer/Higher-Energy Collisional Dissociation (ETHcD) Mass Spectrometry. *Anal. Chem.* 2020, 92, 14730-14739.
6. Sanda M., Morrison L., Goldman R. N and O glycosylation of the SARS-CoV-2 spike protein. *bioRxiv* 2020, 2020.07.05.187344.
7. Uslupehlivan M., Şener E. Glycoinformatics approach for identifying target positions to inhibit initial binding of SARS-CoV-2 S1 protein to the host cell. *bioRxiv* 2020, 2020.03.25.007898.
8. Casalino L., *et al.* Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein. *ACS Central Sci.* 2020, 6, 1722-1734.
9. Lan J., *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* 2020, 581, 215-220.
10. Mehdipour A. R., Hummer, G. Dual nature of human ACE2 glycosylation in binding to SARS-CoV-2 spike. *bioRxiv* 2020, 2020.07.09.193680.
11. Li Q., *et al.* The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and

- Antigenicity. *Cell* 2020, 182, 1284-1294.e9.
12. Premkumar L., *et al.* The receptor-binding domain of the viral spike protein is an immunodominant and highly specific target of antibodies in SARS-CoV-2 patients. *Sci. Immunol.* 2020, 5, eabc8413.
 13. Pinto D., *et al.* Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* 2020, 583, 290-295.
 14. Zhou D., *et al.* Identification of 22 N-glycosites on spike glycoprotein of SARS-CoV-2 and accessible surface glycopeptide motifs: Implications for vaccination and antibody therapeutics. *Glycobiology* 2020, 31, 69-80.
 15. Muhuri M., Gao G. Is smaller better? Vaccine targeting recombinant receptor-binding domain might hold the key for mass production of effective prophylactics to fight the COVID-19 pandemic. *Signal Transduct. Target. Ther.* 2020, 5, 222.
 16. He Y., *et al.* Identification of immunodominant sites on the spike protein of severe acute respiratory syndrome (SARS) coronavirus: implication for developing SARS diagnostics and vaccines. *J. Immunol. (Baltimore, Md. : 1950)* 2004, 173, 4050-7.
 17. Lenza M. P., *et al.* Structural Characterization of N-Linked Glycans in the Receptor Binding Domain of the SARS-CoV-2 Spike Protein and their Interactions with Human Lectins. *Angew. Chem.-Int. Edit.* 2020, 59, 23763-23771.
 18. Resemann A., *et al.* Full validation of therapeutic antibody sequences by middle-up mass measurements and middle-down protein sequencing. *mAbs* 2016, 8, 318-30.
 19. Gstöttner C., *et al.* Monitoring glycation levels of a bispecific monoclonal antibody at subunit level by ultrahigh-resolution MALDI FT-ICR mass spectrometry. *mAbs* 2020, 12, 1682403.
 20. Marcia R Santos C. K. R., Fonslow B., Guttman A. A covalent, cationic polymer coating method for the CESI-MS analysis of intact proteins and polypeptides. 2015.
 21. Yang Y., *et al.* Hybrid mass spectrometry approaches in glycoprotein analysis and their usage in scoring biosimilarity. *Nat. Commun.* 2016, 7, 13397.
 22. Hinneburg H., *et al.* The Art of Destruction: Optimizing Collision Energies in Quadrupole-Time of Flight (Q-TOF) Instruments for Glycopeptide-Based Glycoproteomics. *J. Am. Soc. Mass Spectrom.* 2016, 27, 507-19.
 23. Zhang T., *et al.* Development of a 96-well plate sample preparation method for integrated N- and O-glycomics using porous graphitized carbon liquid chromatography-mass spectrometry. *Mol. Omics* 2020, 16, 355-363.
 24. Madunić K., *et al.* Colorectal cancer cell lines show striking diversity of their O-glycome reflecting the cellular differentiation phenotype. *Cell. Mol. Life Sci.* 2020, 78, 337-350.
 25. Karlsson N. G., *et al.* Negative ion graphitised carbon nano-liquid chromatography/mass spectrometry increases sensitivity for glycoprotein oligosaccharide analysis. *Rapid Commun. Mass Spectrom.* 2004, 18, 2282-92.
 26. Karlsson N. G., Schulz B. L., Packer N. H. Structural determination of neutral O-linked oligosaccharide alditols by negative ion LC-electrospray-MSn. *J. Am. Soc. Mass*

- Spectrom. 2004, 15, 659-72.
27. Anugraham M., *et al.* A platform for the structural characterization of glycans enzymatically released from glycosphingolipids extracted from tissue and cells. *Rapid Commun. Mass Spectrom.* 2015, 29, 545-561.
 28. Ceroni A., *et al.* GlycoWorkbench: A Tool for the Computer-Assisted Annotation of Mass Spectra of Glycans. *J. Proteome Res.* 2008, 7, 1650-1659.
 29. Cooper C. A., Gasteiger E., Packer N. H. GlycoMod--a software tool for determining glycosylation compositions from mass spectrometric data. *Proteomics* 2001, 1, 340-9.
 30. Wohlschlager T., *et al.* Native mass spectrometry combined with enzymatic dissection unravels glycoform heterogeneity of biopharmaceuticals. *Nat. Commun.* 2018, 9, 1713.
 31. Hati S., Bhattacharyya S. Impact of Thiol-Disulfide Balance on the Binding of Covid-19 Spike Protein with Angiotensin-Converting Enzyme 2 Receptor. *ACS Omega* 2020, 5, 16292-16298.
 32. Gstöttner C., *et al.* Intact and subunit-specific analysis of bispecific antibodies by sheathless CE-MS. *Anal. Chim. Acta* 2020, 1134, 18-27.
 33. Trastoy B., *et al.* Structural basis of mammalian mucin processing by the human gut O-glycopeptidase OgpA from *Akkermansia muciniphila*. *Nat. Commun.* 2020, 11, 4844.
 34. Guile G. R., *et al.* Identification of highly fucosylated N-linked oligosaccharides from the human parotid gland. *Eur. J. Biochem.* 1998, 258, 623-56.





Chapter 10

Discussion and future perspectives



Production of biopharmaceuticals requires several levels of characterization to guarantee an effective and safe drug. Similar to other sectors, new products with novel properties are regularly being introduced by the biopharmaceutical industry. This includes optimized conventional antibodies (e.g. improved pharmacodynamics or pharmacokinetics) and new antibody formats with new or enhanced properties. However, analytical technologies are not evolving with the same speed, resulting in a higher load/pressure on the analytical laboratories, increasing development cost and eventually product prices. One example are suitable sized platforms to tackle integrated sample preparation and analysis. Another, major analytical hurdle is the investigation of Fc functions of specific mAb proteoforms, enabling to improve conventional antibody therapeutics (efficacy or pharmacokinetics), (re)-define CQAs or produce new antibody formats. Boosting biopharma analytics is therefore crucial and benefits from close collaboration between industry and academia instead of independent development. The mentioned points will be discussed in the following sections of this thesis.

10.1 Automated sample preparation and MS analysis

MS characterization of antibodies, necessary in several stages of mAb development, stability and formulation assessment, is traditionally performed by bottom-up approaches. Recently, the possibility of performing multiple attribute monitoring (MAM)¹ for CQA analysis is gaining more interest. MAM approaches are based on protein digestion (bottom-up) and analysis by LC-MS/MS. In MAM, information on several CQAs can be obtained from one measurement instead of the traditional single CQA assessment using different platforms. Whereas the concept of getting maximum information out of one measurement and remove several analytical platforms increases efficiency, still it is based on a tedious sample preparation (reduction, alkylation, digestion) prior LC-MS/MS analysis. In biopharma laboratories, this sample preparation is either performed in a manual manner² or following an off-line automated sample preparation by a pipetting robot^{3,4}. Not only the manual approach feels a bit outdated for a highly innovative sector as the biopharmaceutical industry, but it also has the risk to induce unintended modifications⁵ which might stay unnoticed especially when MAM approaches are employed. On the other hand, also robotic platforms for off-line automated sample preparation come with hurdles. A robotic platform needs to be purchased and maintained, trained personal is required and a large amount of samples is needed to make it cost-efficient. Still, following robotic sample preparation, the analysis of the samples by LC-MS needs to be performed, which leads to extended storage times and thereby also increases the risk of unintended modifications. Because in the biopharmaceutical industry samples are often not in the range of hundreds at one time point but rather an almost continuous flow of samples with often only a handful of samples at any given time, robotic sample preparation platforms often feel like using a sledge-hammer to crack a nut.

Chapter 3 presents an elegant solution to perform bottom-up sample preparation and analysis in the same platform, closing the gap for continuous sample analysis and reducing the risk of unintended modifications. The presented concept is based on an commercial 2D-LC system, which is nowadays very common in analytical laboratories. These systems can be easily extended with additional commercial modules and perfectly integrated into the existing software (**Chapter 3**). Also the need for special trained personal is not necessary as most laboratory workers are able to operate an LC system. The samples can be injected directly from formulation without the need for any desalting. All relevant sample preparation and analytical steps from reduction, tryptic digestion over separation of peptides and MS analysis are performed automatically. Whereas in the platform presented in **Chapter 2** some hydrophilic peptides were lost, **Chapter 3** shows improved conditions which allow to obtain very similar results to the standard manual procedure but with lower hands-on time. Also, there is no storage time after sample preparation as in robotic platforms and samples are directly processed online, which avoids the risk of inducing unintended modifications. Next to automation, the use of immobilized enzyme reactors (IMERs) results in more efficient and, therefore, fast digestion. The reduced digestion time is consequence of a higher local trypsin concentration due to immobilization and due to the faster mass transfer between antibody and trypsin⁶. Similarly, the reduction takes only a few minutes instead of 30 min commonly used in the off-line procedure. The method is also applicable to new formats such as BsAbs providing high sequence coverage, comparable to the standard approach.

While MS approaches are mainly used in product characterization as well as formulation and stability testing, release testing in quality control (QC) is largely based on separation methods with UV detection. Still, when an unexpected peak appears, an in-depth MS characterization of these signals is required. Direct hyphenation of QC methods to MS is in most cases not doable due to the use of high amounts of non-volatile salts in the mobiles phases. The standard procedure in the pharmaceutical industry is collection of the peaks by several off-line fractionation steps to obtain enough material and characterization by bottom-up LC-MS. In the cases where the new species are not baseline separated, other proteoforms can be cofractionated complicating a confident assignment of the unexpected species. Even if the peak of interest is well resolved from other mAb proteoforms, this process consumes extensive hands-on time for the fractionation and reanalysis. **Chapter 2** describes an automated system where QC separations can be directly implemented, the peak of interest can be fractionated and individually characterized in an on-line bottom up approach. In contrast to the approximately 52 hours necessary in a regular off-line fractionation, reanalysis and characterization by a bottom-up approach of 5 peaks, this platform can perform this process in around 9 hours in an automated way (**Chapter 2**). The developed platforms (**Chapter 2 and 3**) have been very well accepted and are currently established and used on a daily basis in various pharmaceutical laboratories. Furthermore, there has been a high potential and broad interest by academic and different pharmaceutical laboratories represented by the number of articles based on the method presented in

Chapter 2⁷⁻¹⁰. Despite the full automation at the hardware level, in **Chapter 2** two controlling software tools are currently still necessary, which is not a problem itself, but would benefit from a complete software integration by the vendor in future, allowing even further to enhance the implementation of these approaches in biopharma.

The developed platforms focus on trypsin digestion (which is the standard protocol) for generating the peptides. However, some biopharmaceuticals require alternative endoproteases to provide efficient cleavage (*e.g.* LysC or GluC). Until now, next to trypsin, only papain, IdeS or pepsin columns are commercially available, which either cleave only in the hinge region of antibodies (papain, IdeS) or provides non-specific digestion of the Fc subunit (pepsin), limiting the applicability of the approach. To overcome this bottleneck on-line in-solution digestion – *i.e.* without immobilization of the endoprotease in a column – would be beneficial. This would open the possibility to use other endoproteases commonly used in industry, such as LysC or GluC or endoprotease mixtures, to ensure high sequence coverage and detection of all relevant CQAs. However, as mentioned before the digestion using an IMER has the benefit of a relatively high local endoprotease concentration and a very fast mass transfer between endoprotease and antibody. Mixing both enzyme and antibody on-line might result in a lower local endoprotease concentration and autodigestion therefore, decreasing digestion efficiency. A solution could be the use of proper mixers or to non-covalently trap the endoprotease on a column, which could allow higher concentrations of the enzyme and prevent autodigestion partially.

Besides biopharmaceutical characterization, this platform would be very useful to study the pharmacokinetic behaviour of drugs in serum or disease-related antibodies where often a large number of samples have to be analyzed to draw reliable conclusions. However, sample amounts needed for this analytical scale platform are in the range of μg and therefore, miniaturization would be necessary to reduce sample consumption and further boost the MS sensitivity (*i.e.* nano-ESI). Unfortunately, commercial multidimensional systems in a nano or capillary scale are not available yet, and only in-house developed systems have been reported which lack robustness and high throughput applicability. Therefore, technological developments towards miniaturized multidimensional LC systems by LC vendors are necessary to further push their applicability in pharma and clinical laboratories.

10.2 Intact, middle-up/down and native approaches in biopharma

Bottom-up approaches are very useful for the localization of modifications on the protein backbone and determination of the mAb sequence. However, after endoprotease digestion, some information is lost, leaving only pieces of the puzzle instead of the complete picture. Even when the complete puzzle can be constructed from the pieces, the obtained picture can misrepresent the reality. A clear example of this situation is illustrated in **Chapter 9**.

Comparing the results of intact protein analysis and a simulation of the intact spectrum from bottom-up results showed quite some discrepancies. Whereas the results were comparable in the case of non-charged glycoforms, sialylated glycans were underestimated in the bottom-up approach resulting in an overall lower mass. Similar observations were made in another report, where erythropoietin was analyzed¹¹. The most likely explanation is that the ionization bias on the sialylated species (bearing a negative charge) is higher at the peptide level, due to the lower total number of charges compared to the intact protein which is highly charged. Also the ion transmission energies might not be optimal for each glycopeptide group (*e.g.* sialylated vs non-sialylated glycopeptides), affecting the relative intensity between species. Additional PTMs on the protein backbone would further complicate the situation as co-occurrence of modifications can have an interplay. This example nicely demonstrates the importance to analyze complex glycoproteins on the intact level to get a more real picture of the actual proteoforms. For antibody samples where most PTMs are known, such intact solutions are often enough to provide reliable characterization. Yet, for complex glycoproteins that are first-time characterized, as shown in **Chapter 9** for RBDs, the use of complementary platforms is inevitable. Here using a combination of glycan release, bottom-up and intact MS approaches was necessary to assign all proteoforms of a new protein with very high confidence. After full assignment, however, a comparability analysis of *e.g.* RBD samples could be performed at the intact level to ensure batch-to-batch similarity in a fast and straightforward way with limited sample manipulation.

Next to biases and loss of information on co-occurring modifications – *i.e.* proteoform identity – also some crucial aspects for biopharma characterization cannot be attained by bottom-up approaches. One clear example is the misassembling side products of new antibody formats such as BsAbs. In those cases, the analysis at the intact level is mandatory. In **Chapter 6** we developed a methodology based on CE-MS for the assessment of misassembled side products in BsAb therapeutics. We observed that the engineering process can alter the heterogeneity obtained, highlighting the necessity of these type of approaches. Another benefit of the intact analysis is the possibility to assess the exchange efficiency of bispecific antibodies generated by controlled Fab-arm exchange (**Chapter 7**), which is not possible with stand-alone MS due to the mass similarity of the parental mAbs. Yet, although intact characterization is necessary for macroheterogeneity assessment, it fails on PTM localization. **Chapter 6** shows the benefits of complementary intact and middle-up characterization for BsAbs. Whereas intact analysis allows to get the information on these misassembled side products, the middle-up level allows to localize the proteoforms to a specific subunit and increase certainty in the assignment of proteoforms with small mass differences (higher accuracy). The versatility of CE-MS allowed to assess BsAbs on both levels and determine macro- and microheterogeneity by simply changing the BGE. Some new antibody formats have a modified hinge region (PG LALA mutation), abrogating the use of the classical enzyme IdeS. To overcome this, an alternative enzyme SpeB is reported in **Chapter 6**, able to cleave hinge-region modified antibodies allowing thereby their middle-up

characterization.

Recently top-down and middle-down approaches are gaining more and more attention due to new developments in fragmentation techniques, such as ECD and ETD or UVPD, but also MALDI-MSD, allowing to obtain very good sequence coverage¹². Whereas middle-down approaches require a hinge-region cleavage and/or reduction of disulfide bridges, top-down approaches require only a desalting step or no prior sample preparation. In both cases localization of the PTMs is possible, however middle-down permits accessibility to the hinge region and thereby boosts sequence coverage. It was shown that middle down sequencing can provide a sequence coverage of around 60-89% depending on the subunit¹², whereas top-down of conventional mAbs reaches only around 30% sequence coverage¹³. Similarly, in **Chapter 4** we reported higher sequence coverage for IdeS cleaved bispecific antibodies compared to top-down fragmentation due to the presence of fragments from the hinge region of the Hc.

Another expanding field, which greatly benefited from the recent technological advances in MS and is clearly gaining momentum is native mass spectrometry. Native MS maintains the higher order structure of the protein during analysis permitting to study protein complexes. The implementation of Orbitrap mass analyzers has revolutionized the field of native MS allowing to analyze native proteins with high resolution. Especially instruments with an extended mass range permit to detect with good resolution antibody and new antibody formats in their native state ($\sim m/z$ 6000-8000) and their complexes (*e.g.* dimers, mAb-FcR complexes $\sim m/z$ 7000-9000)¹⁴. Another benefit of native MS is the higher sensitivity due to the lower number of charge states (and therefore lower spread of the signal) compared to denaturing MS. However, many mass spectrometers so far are not able to allow good declustering or desolvation of native proteins to benefit from this increased sensitivity. Orbitrap mass analyzers provide more efficient desolvation compared to FT-ICR or Q-TOF mass analyzers, due to the possibility of storing the ions in an HCD cell with a higher gas pressure¹⁵. Overall, all these advances significantly boosted native MS in the last few years and first applications of native MS in biopharmaceutical analysis have been demonstrated¹⁶⁻¹⁸ and applied by pharma companies and undoubtedly will be expanded in the next years. In this lines, **Chapter 8** demonstrated the benefits of native MS to study protein-protein interactions and determination of the binding stoichiometry by detection of the formed FcRn-antibody complexes.

10.3 Affinity CE for the selective functional characterization of antibody proteoforms

Antibody biopharmaceuticals consist of a plethora of proteoforms, each with a potentially different function. Well-known examples are afucosylation and its increased affinity to the FcγRIIIa¹⁹ or the decreased binding of oxidated antibodies to the FcRn²⁰. These are extreme

examples that show a severe biological effect. However, many more proteoforms could/ have not been addressed with existing binding techniques (*e.g.* SPR or ELISA) because of their inability to distinguish between different species such as different proteoforms. These techniques determine an overall affinity for the mixture or require the production/isolation of enriched antibody proteoforms, which often do not result in a 100% pure form and therefore do not provide confident results²¹.

The challenge of proteoform-specific binding affinity assessment can be clearly illustrated through the example of glycosylation. Let's assume that we have two samples containing a high and low amount of core fucose and they are tested for their FcγRIIIa binding. Obviously, the mentioned techniques will determine a difference between both samples and link the increased binding affinity to antibody afucosylation. However, still the question of the impact of other glycan features such as galactosylation is not solved. To test for this, glycoengineered high and low galactosylated samples have to be analyzed in the context of high and low core fucose. However, alteration of galactosylation will also slightly change fucosylation levels, biasing the results. Furthermore, highly galactosylated antibodies may comprise 2 to 4 galactoses per antibody and low galactosylated ones from 0 to 2 making it hard to point out which galactoses influence the binding²¹. Additionally, galactosylated samples often exhibit sialic acids, which can also reduce binding to the FcγRIIIa receptor in context of low levels of fucose but high amounts of bisection¹⁹. In addition to the challenge of producing pure glycoforms, additional proteoforms often appear during production and storage, such as deamidation or oxidation, influencing the binding. On top of this, the situation gets even more complex if we consider the heterogeneity of the receptors (*e.g.* different glycoforms), which can also influence binding affinity and should be taken into account²². This example shows the urgent need for methods able to determine the binding affinities in a proteoform specific manner.

Affinity LC-MS (AC-MS) has recently demonstrated its ability to study binding of proteoforms with different glycosylation or oxidation status to FcγRIIIa and FcRn^{23, 24}. AC-MS offers information in relative binding (dissociation) of different proteoforms but in contrast to SPR or ELISA do not provide quantitative information (*i.e.* affinity constants). Although AC-MS opens possibilities for proteoform-selective binding studies, it comes still with certain limitations and drawbacks. After binding to the receptor, immobilized in the stationary phase, the proteoforms are eluted by a pH stress to disrupt the interaction. For instance, in FcγRIIIa elution is achieved using a pH gradient from pH 5 to 3 which is not a physiological pH. These conditions are far from a biological representation and the results should be carefully interpreted. Furthermore, in AC-MS (and most binding techniques) the receptor is immobilized to the stationary phase. This obviously does not represent the native situation where the receptors are located in the cell membrane and allowed to form in some instances higher-order structures. In addition, the stationary material can cause secondary interactions between the antibody and the stationary phase which can not be distinguished from actual

binding. These secondary interactions are dependent on the protein sequence resulting in different elution times for different antibodies even if they exhibit the same binding affinity. To exclude unspecific binding, negative controls with no binding to FcRs (*e.g.* same antibody with a LALA-PG mutation to silence the binding) are necessary for each particular antibody, hampering the possibility of comparing binding affinities of different antibody samples or application to antibody mixtures²⁵. Commercial availability of Fc receptor columns in AC-MS is another major challenge as now it is limited to FcRn and FcγRIIIa. Receptor immobilization and self-packing of the columns are possible, but it requires high knowledge, plenty of the receptor and can lead to high lot to lot variations. As mentioned before, the heterogeneity of the Fc receptor (*e.g.* glycosylation) can also influence the binding to the antibody²². Therefore, glycosylation changes of the receptor would require new packing of affinity LC columns which is a very tedious process.

Mobility-shift ACE approaches showed the ability to determine affinity constants between proteins or peptides and small molecules^{26, 27}, and recently between small proteins²⁸. In **Chapter 8** we demonstrated that ACE-MS can also be exploited to study antibody-FcR binding, clearly overcoming the drawbacks from the previously mentioned techniques. This is due to the fact that the receptor is free in the capillary and that the approach monitors differences in equilibrium (not only dissociation as compared to AC-MS). This allows to study binding affinity under physiological conditions without any pH stress. As the receptor is free in the BGE, it can form higher-order structures in the CE capillary. Furthermore, the receptor can be titrated by simply adding different concentrations to the BGE, permitting to obtain quantitative information on the binding (*i.e.* affinity constants, K_d s). Especially when using MS, the affinity of each proteoform can be determined in a mixture (even if overlapping) without the need of producing highly pure glycoengineered proteoforms. Hyphenation with MS also permits to simultaneously assess receptor heterogeneity (which is detected as a constant signal in the MS) and to monitor the complexes formed between antibodies and the receptors. The latter provides valuable information on the binding stoichiometry as shown in **Chapter 8**, where antibodies in complex with one (1:1 binding) or two FcRn molecules (1:2 binding) were observed. Furthermore, depending on the complexity of the receptor, monitoring the complexes formed can potentially deliver some information on the influence of receptor heterogeneity in the binding.

While the approach can be applied in a straightforward way to monitor relative binding of proteoforms, K_d determinations still suffer from lack of robustness. This is mainly due to the fact that electrophoretic mobilities can be affected by subtle changes in buffer characteristics (*e.g.* receptor concentration, viscosity) or in the capillary wall (*e.g.* adsorption). To correct for these changes we used a marker protein which does not have any interaction with the FcRn receptor. However, as the marker protein migrates before the antibody proteoforms, changes occurring after can not be corrected. Including a second marker protein (one migrating before and the other after the antibody) would allow for better correction on

differences in migration time and more accurate electrophoretic mobility determination. Furthermore, recent software developments permit easy conversion of electropherograms to electrophoretic mobilities using two markers, permitting easy implementation of this strategy^{29, 30}. This would permit to monitor very small differences in affinity with high confidence and to determine K_d values more accurately.

Adding the receptor in the BGE comes with high flexibility. A receptor can be easily exchanged for another one permitting multiple applications with the same capillary (*e.g.* various FcRs). Furthermore, ACE-MS can be employed to study the influence of receptor heterogeneity on the binding by simply reversing the approach, avoiding the necessity of receptor engineering. For the reverse approach, instead of the receptor, the antibody can be added to the BGE and the receptor injected, allowing to separate receptor proteoforms based on their different affinity to the antibody. In contrast to AC, where milligrams of receptor are necessary, in ACE only a few micrograms are required for a study as the capillary is filled with nanograms of the receptor. This permits applications to receptors that are difficult to produce and where limited amounts are available. Next to FcRs, the high mannose receptor is of high interest in the biopharma industry due to its involvement in the clearance of high mannose antibody proteoforms. The high mannose receptor (and many other C-type lectin family receptors) bring however additional challenges as they often show calcium dependent binding. The addition of calcium to the BGE would result in a drastic decrease of the MS signal due to adduct formation between the antibody and calcium ions. To overcome this drawback alternative strategies, such as a partial filling of the capillary with the receptor, would be necessary to avoid MS signal suppression, yet such approaches still have to be explored.

In summary, mobility-shift ACE-MS is a highly flexible tool, able to determine affinity constants in a proteoform-specific manner, is closer to native conditions than AC-MS approaches, provides very rich structural information and requires very low amounts of antibody and receptor. Overall for future, this line has the possibility to help optimizing current antibody drugs and can also find clinical relevance to study different binding behavior of antibodies in disease.

10.4 Industry and academy: A symbiosis?

Symbiosis happens very often in nature. One concrete example are honeybees and flowers. Honeybees get nectar and pollen from multiple flowers and thereby transport pollen from flower to flower fertilizing them. Everyone has a benefit. The honeybees have food and the population can grow. The plants will produce seeds allowing also to increase their population, which is again positively influencing the honeybees. Industry and academia could also make a perfect symbiosis due to their different interest and complementarity. When both work closely together, the industry which does not have time to test and exploit new technologies or develop new methods can give these questions to academia, that does have time but

struggles to get access to interesting problems and/or the molecules. The industrial partner does get access to people with a lot of know-how and the time to develop new methods that can be transferred, whereas the academia gets funding from the industrial partner and can promote and exploit their research.

Another principle in some bees is nectar robbing. Here bees bite a hole in the flower to get the nectar bypassing therefore, the pollination. This way the bees can get their nectar, but flowers can not reproduce and the population can not grow, which is not beneficial for both of them in the long term. To minimize the threats, plants often develop protective mechanisms to hinder them from the robber bees. Although this situation is rarely seen in nature due to obvious unsustainability reasons, this relationship can be seen often between academia and industry. The industry is constantly worried that academia could “steal” something from them and therefore developed mechanisms, which make collaborations very difficult. In many cases, instead of collaborating, pharmaceutical companies try to do everything on their own. In others, the collaborations are not very successful due to too many hurdles and agreements that need to be signed (*e.g.* IP rights, MDAs, MTAs), publication approvals and other confidentiality aspects because pharmaceutical industries want to protect their property or distrust the academic groups.

However, this is company dependent and there are companies that have successful collaborations for years, whereas others never or barely collaborate. This PhD thesis has been achieved in the frame of a very successful and fruitful academic-industry collaboration. A key aspect for this success was a previous internship performed at the industrial partner and the two additional short secondments performed during the project. This allowed the establishment of a relation of trust and widened the vision on the real interest and challenges faced by industry. Unfortunately, not all researchers have these opportunities and experiences, as either the secondment in industry requires a lot of paperwork and initial training for the student (often secondments below three months are not even considered by industry as the student need a 2-3 weeks training before entering to the lab) or the industry does not have the technology that is being exploited in the project. Big consortia such as the Marie Curie Horizon 2020 consortium Analytics for Biologics (A4B) help to establish initial contacts but still suffer from these problematics. As a matter of fact, from 15 PhD students in the A4B consortium only few spent some time at one of the industrial partners. To encourage these internships industrial partners should facilitate secondments to students and grant some exceptions regarding training regulations (as in most of the cases the student will not work on business-critical projects) and the academic supervisors should understand that even if there is no rocket science happening during the secondment period, this would be highly beneficial for the student. Internships not only expand the capabilities and experience of the students but also allow to strengthen the industry-academia relationship. In fact, the exchanges should not be limited to students but rather be open to other personal to tear this wall between industry and academia down. This would also benefit researchers for

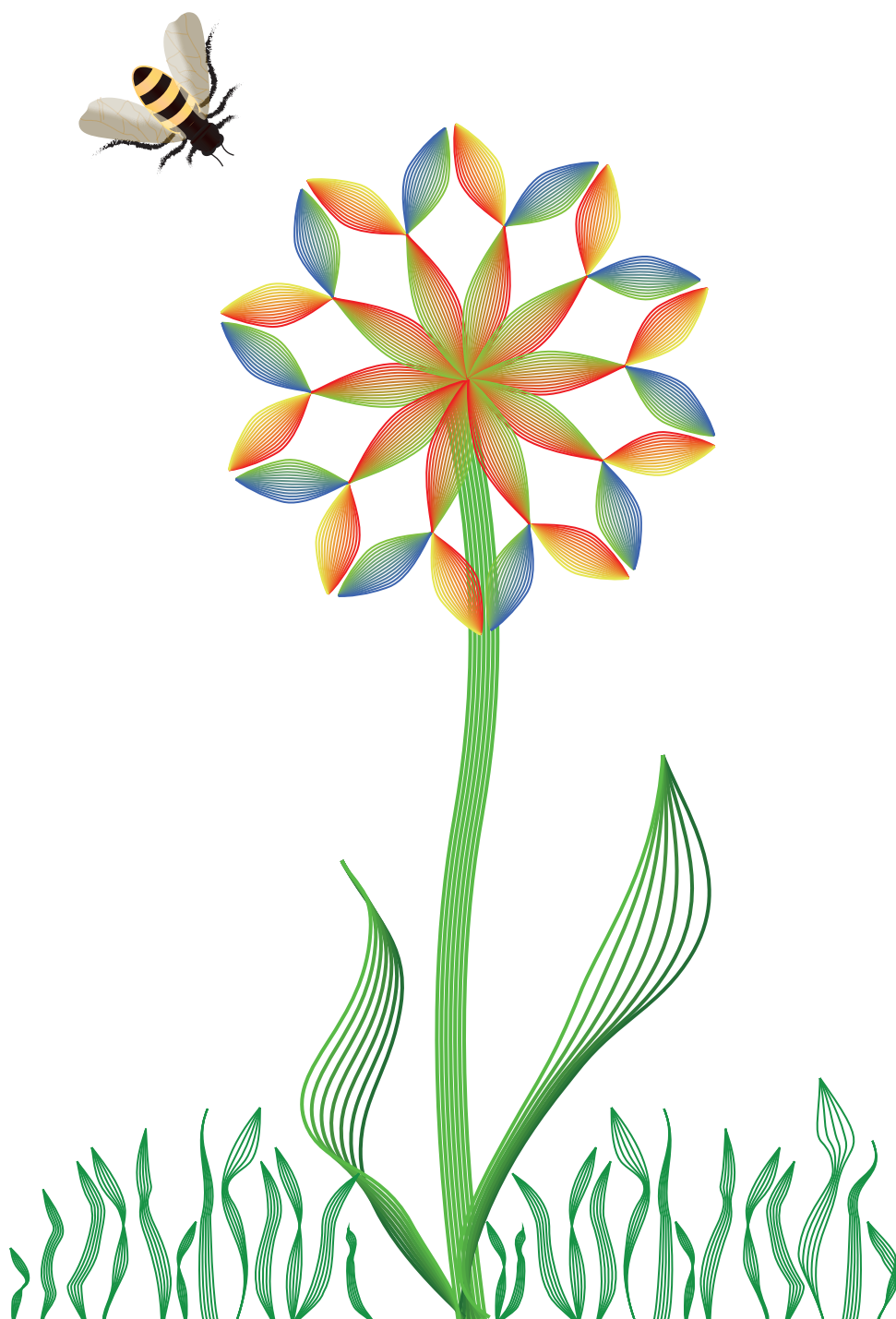
possible late changes in their career paths. Shared positions could also have a tremendously positive influence but yet, are rarely supported from both sides.

Funding by governmental institutions should be more supportive and should invest more in these collaborations. It is often thought that pharmaceutical industries should sponsor most of the collaboration costs instead of being supported by public funds. In many cases, this can cause the industrial partners to be hesitant regarding collaborations as they feel academy is only interested in their money instead of seeing the real benefit of the collaboration. However, the COVID-19 pandemic showed that also the societal benefit is dependent on a strong pharmaceutical industry. Without tests, a vaccination or a treatment of the COVID-19, lockdowns would have been longer and the economy would have suffered even more. Many companies based their vaccines on academic developments and academic groups saw how their discoveries could be applied and save lives. Besides all the negative sides, maybe the corona pandemic has a positive influence on the global vision of industry-academia relations. An example is the development of a pandemic institute financed by the Fraunhofer society and the Bavarian government in collaboration with the Roche diagnostic division aimed to get a better understanding on infectious diseases and built directly next to Roche (Penzberg) to promote the collaboration³¹.

10.5 References

1. Rogers R. S., et al. Development of a quantitative mass spectrometry multi-attribute method for characterization, quality control testing and disposition of biologics. *MABs* 2015, 7, 881-890.
2. Mouchahoir T., Schiel, J. E. Development of an LC-MS/MS peptide mapping protocol for the NISTmAb. *Anal Bioanal Chem* 2018, 410, 2111-2126.
3. Bauer L. G., et al. Monitoring modifications in biopharmaceuticals: Toolbox for a generic and robust high-throughput quantification method. *J Chromatogr B* 2020, 1148, 122134.
4. Song Y. E., et al. Automated mass spectrometry multi-attribute method analyses for process development and characterization of mAbs. *J Chromatogr B* 2021, 1166, 122540.
5. Liu S., et al. Mildly acidic conditions eliminate deamidation artifact during proteolysis: digestion with endoprotease Glu-C at pH 4.5. *Amino Acids* 2016, 48, 1059-1067.
6. Wouters B., et al. On-line microfluidic immobilized-enzyme reactors: A new tool for characterizing synthetic polymers. *Anal Chim Acta* 2019, 1053, 62-69.
7. Goyon A., et al. From proof of concept to the routine use of an automated and robust multi-dimensional liquid chromatography mass spectrometry workflow applied for the charge variant characterization of therapeutic antibodies. *J Chromatogr A* 2020, 1615, 460740.
8. Camperi J., et al. Fast and Automated Characterization of Monoclonal Antibody Minor Variants from Cell Cultures by Combined Protein-A and Multidimensional LC/MS Methodologies. *Anal Chem* 2020, 92, 8506-8513.
9. Goyon A., et al. Streamlined Characterization of an Antibody-Drug Conjugate by Two-Dimensional and Four-Dimensional Liquid Chromatography/Mass Spectrometry. *Anal Chem* 2019, 91, 14896-14903.
10. Camperi J., Guilleme D., Stella C. Targeted Bottom-up Characterization of Recombinant Monoclonal Antibodies by Multidimensional LC/MS. *Anal Chem* 2020, 92, 13420-13426.
11. Yang Y., et al. Hybrid mass spectrometry approaches in glycoprotein analysis and their usage in scoring biosimilarity. *Nat Commun* 2016, 7, 13397-13397.
12. Srzentić K., et al. Interlaboratory Study for Characterizing Monoclonal Antibodies by Top-Down and Middle-Down Mass Spectrometry. *J Am Soc Mass Spectr* 2020, 31, 1783-1802.
13. Fornelli L., et al. Top-down analysis of immunoglobulin G isotypes 1 and 2 with electron transfer dissociation on a high-field Orbitrap mass spectrometer. *J Proteom* 2017, 159, 67-76.
14. Leney A. C., Heck A. J. R. Native Mass Spectrometry: What is in the Name? *J Am Soc Mass Spectr* 2017, 28, 5-13.
15. Rose R. J., et al. High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nat Methods* 2012, 9, 1084-1086.

16. Habegger M., et al. Rapid characterization of biotherapeutic proteins by size-exclusion chromatography coupled to native mass spectrometry. *MAbs* 2016, 8, 331-339.
17. Füssl F., et al. Comprehensive characterisation of the heterogeneity of adalimumab via charge variant analysis hyphenated on-line to native high resolution Orbitrap mass spectrometry. *MAbs* 2019, 11, 116-128.
18. Füssl F., et al. Charge Variant Analysis of Monoclonal Antibodies Using Direct Coupled pH Gradient Cation Exchange Chromatography to High-Resolution Native Mass Spectrometry. *Anal Chem* 2018, 90, 4669-4676.
19. Dekkers G., et al. Decoding the Human Immunoglobulin G-Glycan Repertoire Reveals a Spectrum of Fc-Receptor- and Complement-Mediated-Effector Activities. *Front Immunol* 2017, 8.
20. Bertolotti-Ciarlet A., et al. Impact of methionine oxidation on the binding of human IgG1 to Fc Rn and Fc gamma receptors. *Mol Immunol* 2009, 46, 1878-82.
21. Dekkers G., et al. Multi-level glyco-engineering techniques to generate IgG with defined Fc-glycans. *Sci Rep* 2016, 6, 36964.
22. Hayes J. M., et al. Identification of Fc Gamma Receptor Glycoforms That Produce Differential Binding Kinetics for Rituximab. *Mol Cell Proteom* 2017, 16, 1770-1788.
23. Lippold S., et al. Glycoform-resolved FcγRIIIa affinity chromatography-mass spectrometry. *MAbs* 2019, 11, 1191-1196.
24. Gahoual R., et al. Detailed Characterization of Monoclonal Antibody Receptor Interaction Using Affinity Liquid Chromatography Hyphenated to Native Mass Spectrometry. *Anal Chem* 2017, 89, 5404-5412.
25. Wessels U., et al. Detection of antidrug antibodies against human therapeutic antibodies lacking Fc-effector functions by usage of soluble Fcγ receptor I. *Bioanalysis* 2016, 8, 2135-2145.
26. Dunayevskiy Y. M., et al. Simultaneous Measurement of Nineteen Binding Constants of Peptides to Vancomycin Using Affinity Capillary Electrophoresis–Mass Spectrometry. *J Med Chem* 1998, 41, 1201-1204.
27. Hutanu A., et al. Methionine oxidation of proteins analyzed by affinity capillary electrophoresis in presence of silver(I) and gold(III) ions. *Electrophoresis* 2021, 42, 1209-1216.
28. Domínguez-Vega E., et al. Simultaneous assessment of protein heterogeneity and affinity by capillary electrophoresis-mass spectrometry. *Anal Chem* 2015, 87, 8781-8.
29. Codesido S D. N., et al. A Mature ROMANCE: A Matter of Quantity and How Two Can Be Better than One. *ChemRxiv* 2020.
30. González-Ruiz V., et al. ROMANCE: A new software tool to improve data robustness and feature identification in CE-MS metabolomics. *Electrophoresis* 2018, 39, 1222-1232.
31. <https://www.merkur.de/lokales/weilheim/penzberg-ort29272/penzberg-erhaelt-pandemie-forschungsinstitut-90045621.html>.



Appendices

Nederlandse samenvatting

Antilichamen worden veelvuldig ingezet als medicijn bij de behandeling van verschillende ziektes. Aanvankelijk werden deze therapeutische antilichamen uit bloed van donoren gezuiverd, echter tegenwoordig worden deze veelal recombinant geproduceerd. Het kan hierbij zowel om een conventioneel antilichaam gaan als een antilichaam met geoptimaliseerde functionaliteiten. Evenals endogene antilichamen bevatten recombinant geproduceerde antilichamen verschillende post-translationele modificaties (PTMs), die potentieel de functie kunnen beïnvloeden. Het is van groot belang de effecten van deze PTMs in kaart te brengen teneinde de veiligheid en effectiviteit van de medicijnen te garanderen (**Hoofdstuk 1**). Ondanks dat er verschillende analytische platforms succesvol zijn ontwikkeld voor de karakterisering van monoklonale antilichamen (mAbs), is het belangrijk dat technologie verder wordt ontwikkeld en geoptimaliseerd voor nauwkeurige analyse van nieuwe therapeutische antilichamen. De volgende aspecten zijn hierbij van belang: automatisering, de detectie van onverwachte PTMs en proteovorm-specifieke functionele karakterisering.

In **Hoofdstuk 2** en **3** wordt automatisering van zogenaamde *bottom-up* methoden voor de snelle karakterisering van antilichamen beschreven. Door automatisering kunnen er meer monsters in een kortere tijd geanalyseerd worden. De doorlooptijd voor analyse wordt verkort en het risico op introductie van onbedoelde modificaties wordt kleiner. In **Hoofdstuk 2** wordt de integratie van een *bottom-up* methode met een multidimensionaal vloeistofchromatografie systeem beschreven. De antilichamen worden hierbij gescheiden op basis van lading, gevolgd door een *bottom-up* karakterisering. Deze methode maakt gedetailleerde en geautomatiseerde karakterisering van antilichaamvarianten met een verschil in lading mogelijk zonder dat er fracties geïsoleerd worden. Dit is een voordeel aangezien zulke fracties vaak een mengsel van verschillende varianten bevatten als gevolg van niet-perfekte chromatografische scheiding. Daarnaast kan met deze aanpak *offline* monstervoorbewerking omzeild worden. Hierdoor worden de volgende stappen uitgevoerd zonder dat er handen aan te pas komen: reductie, digestie met trypsine en meting met LC-MS(/MS). In **Hoofdstuk 3** wordt deze nieuwe aanpak met een online digestie met trypsine verder ontwikkeld door middel van optimalisatie van de kolom, *flow rate* en temperatuur. Op deze manier kan de aminozuursequentie (primaire structuur) van de antilichamen met eenzelfde kwaliteit worden vastgesteld als bij de conventionele *offline* procedures. Dit hoofdstuk laat daarnaast zien dat het mogelijk is geformuleerde mAbs en bispecifieke antilichamen (BsAbs) direct en snel te analyseren zonder extra monstervoorbehandeling. Hiermee wordt een alternatief geboden voor eerdergenoemde geautomatiseerde platforms.

Bottom-up onderzoeksstrategieën zijn geschikt om monospecifieke antilichamen in detail te karakteriseren, echter voor het lokaliseren van modificaties op nieuwe antilichaam

producten zijn additionele technieken noodzakelijk. Zo is bijvoorbeeld van een proteolytisch glycopeptide niet vast te stellen bij welke zware keten van het BsAbs deze hoort. Daarom zijn in **Hoofdstuk 4** *middle-up/down* procedures ontwikkeld gebaseerd op *matrix-assisted laser desorption ionization* (MALDI)-MS voor de lokalisering van PTMs van BsAbs op het niveau van de ketens en de aminozuren. Het voordeel van deze methodes wordt gedemonstreerd aan de hand van geglyceerde BsAb monsters, waar meerdere glycaties op de ketens zijn gedetecteerd. De exacte plaats van de glycatie kon worden achterhaald wanneer deze zich dichtbij de N- of C-terminus bevond door fragmentatie met *in-source decay* (ISD). In **Hoofdstuk 5** wordt MALDI-ISD toegepast voor de detectie van een labiele modificatie (sulfatering) op een antilichaam. Deze sulfatering kon met ISD toegekend worden aan een tyrosine residu, iets wat niet mogelijk is met andere fragmentatietechnieken. Naast sulfatering wordt in **Hoofdstuk 5** ook een 4-hydroxyproline modificatie beschreven op een nieuw antilichaam product. Tot slot is een *in silico* model ontwikkeld om modificaties al in een vroeg stadium (tijdens het ontwerpproces van mAbs en BsAbs) te kunnen voorspellen. Hierdoor kan de sequentie voorafgaand aan de expressie aangepast worden teneinde te voorkomen dat ongewenste modificaties in het eindproduct aanwezig zijn.

Hoofdstuk 6 beschrijft een alternatieve methode gebaseerd op capillaire elektroforese (CE)-MS om BsAbs te bestuderen op het *middle-up* en intacte niveau. Hiervoor worden twee zuster BsAbs gebruikt die binden aan hetzelfde deel van een antigeen maar verschillend geproduceerd zijn. In dit hoofdstuk wordt de toepassing van een nieuwe hinge-regio endoprotease (*SpeB*) onderzocht voor antilichamen met een aangepaste hinge regio (vergeleken met IgG1) als alternatief voor het veelgebruikte *IdeS* enzym. Met de ontwikkelde CE-MS methode voor de bestudering van de intacte structuur van een BsAbs is het mogelijk om naast het antilichaam ook nevenproducten te karakteriseren, zoals ontbrekende BsAbs-ketens, vrije lichte ketens met verschillende PTMs of BsAbs die één N-glycosylering missen. Minimale monstervoorbehandeling is van groot belang teneinde verlies van de quaternaire structuur te voorkomen tijdens het onderzoek naar macroheterogeniteit. Deze CE-MS methode is verder toegepast om de correcte samenstelling van BsAbs te controleren. **Hoofdstuk 7** demonstreert de toepasbaarheid van CE-MS op het niveau van intacte antilichamen om de uitwisselingsefficiëntie te monitoren tijdens bispecifieke antilichaamproductie met behulp van een gecontroleerde Fab-armuitwisseling. De methode is geschikt voor antilichamen met en zonder Fab-glycosylering. Daarnaast maakt deze methode het mogelijk om afbraakproducten die optreden tijdens langdurige opslag te monitoren.

De aanwezigheid van PTMs op antilichamen kan vragen oproepen over hun impact op de functionaliteit. Vaak worden deze functies gemeten met behulp van bindingstechnieken, zoals ELISA of SPR. Echter, deze technieken leveren slechts een gemiddelde waarde van de functie van een compleet mAb monster. Om functie van proteovormen vast te kunnen stellen aan de hand van deze technieken moeten er fracties opgevangen worden of

specifieke varianten geproduceerd worden, wat in beide gevallen veel tijd kost. **Hoofdstuk 8** behandelt deze problematiek en draagt mobiliteitsverschuivingsaffiniteit CE-MS aan als alternatieve techniek om antilichaaminteracties op een proteovorm-specifieke manier te bestuderen zonder additionele prefractionering. Met de ontwikkelde methodiek wordt de interactie bestudeerd tussen verschillende antilichaam-proteovormen en de FcRn-receptor, welke verantwoordelijk is voor de halfwaardetijd van antilichamen. De informatie over de affiniteit wordt verkregen door de verschuiving in mobiliteit na toevoeging van de receptor. De koppeling met MS brengt de benodigde structurele informatie. Verder laat dit hoofdstuk zien dat door het gebruik van verschillende hoeveelheden FcRn-receptor in de achtergrondelektrolyt, K_d -waarden voor elke proteovorm tegelijkertijd kunnen worden bepaald.

Het laatste deel van dit proefschrift (**Hoofdstuk 9**) richt zich op de karakterisering van het recombinant geproduceerde SARS-CoV-2 receptorbindend domein (RBD). Om een totaalbeeld van het RBD te krijgen ondanks de immense complexiteit en beperkte voorkennis, is het noodzakelijk om *multilevel* benadering toe te passen, namelijk O-glycaan analyse, *bottom-up* N-glycopeptide analyse, MALDI *top-down* fragmentatie en intacte CE-MS metingen. Daarnaast worden een endo- en exoglycosidase enzym ingezet teneinde de complexiteit van RBD te reduceren. Op deze manier kunnen alle glycovormen en PTMs toegekend worden. Twee RBD-monsters geproduceerd in verschillende celtypen (CHO en HEK293) worden gekarakteriseerd en heterogeniteit wordt in kaart gebracht. Verder wordt in dit hoofdstuk voor het eerst aangetoond waar de O-glycosylering in het RBD is gelokaliseerd. Hiervoor is een combinatie van intacte en *top-down* MS analyses nodig. Na structurele karakterisering worden bindingsassays met antilichamen van SARS-CoV-2-patiënten en de ACE2 receptor uitgevoerd om de binding (en mogelijke verschillen) tussen de twee RBD-monsters te onderzoeken.

Het laatste hoofdstuk (**Hoofdstuk 10**) bevat een algemene discussie van de onderwerpen die zijn besproken in dit proefschrift. Dit hoofdstuk beschrijft de noodzaak voor geautomatiseerde platforms voor monsteranalyse in de farmaceutische industrie en hoe deze verder verbeterd kunnen worden. Verder wordt de relevantie van intacte en *middle-up/down* native MS bediscussieerd. Hierbij wordt specifiek aandacht gegeven aan de in **Hoofdstuk 9** geconstateerde discrepanties tussen de *bottom-up* en intacte MS-data. Daarnaast wordt er een sectie gewijd aan de functionele karakterisering van PTMs op mAbs. Hierin wordt het gebruik en verdere ontwikkeling van nieuwe technieken besproken. Ook wordt er een perspectief geboden op de invoering van deze technieken in de kliniek om de veranderingen van antilichaam-proteovormen in ziektes te kunnen detecteren. Tenslotte zijn er afsluitende gedachten met betrekking tot de (af en toe) gelimiteerde relatie tussen de industrie en academie en worden er mogelijke strategieën genoemd om deze relaties te versterken.

English Summary

Antibody-based therapeutics are widely used to treat various diseases. Initially, antibodies were enriched from blood specimens before administration to patients but nowadays most are produced recombinantly, either as conventional antibodies or with optimized functions. These recombinantly produced antibody therapeutics commonly carry a plethora of post-translational modifications (PTMs) that potentially alter their function and therefore need to be monitored to ensure high safety and efficacy (**Chapter 1**). Although several analytical platforms have been established for the characterization of conventional monoclonal antibodies (mAbs), aspects such as automation, the appearance of unexpected PTMs in novel antibody formats or proteoform-specific functional characterization still require further analytical developments.

Chapter 2 and **3** address automation of bottom-up approaches for fast antibody characterization. Automation allows analysis at higher sample throughput thereby reducing hands-on time and minimizing the risk of introducing unintended modifications. **Chapter 2** is focused on integrating bottom-up characterization after a charge variant separation in a multidimensional liquid chromatography (LC) platform. The described method allows in-depth characterization of separated variants in an automated way avoiding tedious collection of chromatographic fractions. Furthermore, the approach provides a solution to overcome offline sample preparation, including reduction, overnight tryptic digestion and LC-mass spectrometry (MS)(/MS) analysis. In **Chapter 3** this new online tryptic digestion approach was further developed. Antibody sequence coverages similar to the ones obtained by offline procedures were achieved by optimization of parameters such as column, flow rates and temperature. In addition, this chapter provides a new platform for direct analysis of formulated mAb and bispecific antibody (BsAb) samples without the need of any sample pre-treatment allowing to process several samples in a short timeframe and thereby providing an alternative to robotic platforms.

Although bottom-up approaches allow in-depth characterization of monospecific antibodies, the newer antibody formats often require additional solutions to localize a modification with regard to their quaternary structure. For example, in BsAbs, many heavy chain proteolytic peptides are identical even though the chains are partially different, hampering the possibility to assign PTMs to specific subunits. **Chapter 4** describes middle-up/down and top-down approaches based on matrix assisted laser desorption ionization (MALDI)-MS for localization of PTMs of BsAbs at the subunit or amino acid level. The applicability of the methods was demonstrated by the analysis of glycosylated BsAb samples. Several glycosylations were detected at the subunit level, for which the modified site could be determined when close to the N- or C-terminus by applying in source decay (ISD) fragmentation. In **Chapter 5** the MALDI-ISD approach was additionally applied to an antibody presenting a very labile modification, *i.e.* sulfation. Since ISD is a soft fragmentation technique it allowed to assign the sulfation to

a tyrosine residue, which was not possible with other fragmentation techniques. Next to sulfation **Chapter 5** also describes the assignment of a 4-hydroxyproline which occurred on new antibody formats. Finally, an in-house *in silico* tool was developed allowing to predict the appearance of modifications already during mAb or BsAb design and thereby allowing to modify the sequence before expression to avoid them.

Chapter 6 describes an alternative approach to study BsAbs at the middle-up and intact level based on capillary electrophoresis (CE)-MS. Two sister BsAbs, with same antigen target but different engineering process were employed. In this chapter, a novel hinge-region endoprotease (SpeB) was explored as alternative to the commonly used IdeS for antibodies with a modified hinge region compared to IgG1. Application of the CE-MS method for the analysis of BsAbs at the intact level allowed to characterize side products, such as BsAbs missing chains, free light chains carrying different PTMs or BsAbs missing one N-glycosylation. This type of macroheterogeneity required the analysis with limited sample treatment to avoid loss of information on their quaternary structure. This intact CE-MS approach was in addition applied to monitor correct assembly of BsAbs. **Chapter 7** therefore demonstrates the applicability of intact CE-MS to monitor the exchange efficiency during bispecific antibody production using a controlled Fab-arm exchange. The method was suitable for antibodies with and without Fab glycosylation and to assess degradation products occurring during elongated storage.

Equally important next to characterization of antibody PTMs is to evaluate the impact of these modifications on antibody functionality. Functional characteristics are often addressed using binding techniques, such as ELISA or SPR. However, these techniques provide an overall response for a mAb sample requiring tedious fractionation or production of single variants. In **Chapter 8** these difficulties are addressed and a technique is described to study antibody interactions in a proteoform specific-manner based on mobility shift affinity CE-MS without the need for any prefractionation. The developed approach focused on studying the interaction of different antibody proteoforms with the FcRn receptor, responsible of antibody half-life. Affinity information was obtained by means of their different mobility shifts upon addition of the receptor while hyphenation to MS provided direct structural information. Furthermore, it is shown that by using different amounts of FcRn in the background electrolyte K_D values can be simultaneously determined for each proteoform.

The last part of this thesis (**Chapter 9**) is focused on the characterization of recombinantly produced SARS-CoV-2 receptor binding domain (RBD). Due to the complexity and limited knowledge on the RBD, a multilevel approach was necessary to obtain an entire picture. To this end, released O-glycan analysis, bottom-up N-glycopeptide analysis, MALDI top-down fragmentation and CE-MS intact analysis were combined with different endo- and exoglycosidase treatments to obtain a complete assignment of glycoforms and PTMs. Two RBD samples produced in different cell types (CHO and HEK293) were characterized

revealing specific heterogeneity. Furthermore, the chapter describes the localization of the O-glycosylation site in the RBD for the first time. This was achieved by a combination of intact and top-down MS analysis. After structural characterization, binding assays with antibodies from SARS-CoV-2 patients and ACE2 were performed to study potential differences in binding between the two RBD samples.

The final chapter (**Chapter 10**) comprises a general discussion of this thesis. The chapter discusses the need of automated sample platforms in the pharmaceutical industry and how they could be further improved. Furthermore, the relevance for intact and middle-up/down native MS is discussed giving special attention to the discrepancies between bottom-up and intact data based on the observations reported in **Chapter 9**. Another discussion point is the functional characterization of mAb PTMs with various techniques, further developments and a view on potential clinical use to study the implications of antibody proteoforms in diseases. Finally, some personal thoughts on the (sometimes) limited relation between industry and academia and some possible strategies to strengthen this relationship are shared.

Curriculum Vitae

Christoph Johann Gstöttner was born on the 25th of January 1993 in Zwiesel, Germany. At the age of six he went to primary school Bodenmais, followed by the secondary school in Zwiesel where he graduated in 2011. Following this Christoph studied Biology in Regensburg, which he finished in 2014 with a bachelor work in the functional characterization of various RNA Polymerase I mutants. Hereafter he continued with a master in biology focused in the area of biochemistry. During his master he became more and more acquainted with proteins, either measuring their activity and stability, producing whole protein complexes recombinantly or analyzing these by mass spectrometry. The latter he performed during an internship which he conducted 2016 at Roche in Basel, where he also performed his practical work for his master thesis on the automation of protein analysis by using multidimensional LC hyphenated to mass spectrometry. He graduated in 2017 as Master of Science with an excellent grade. Because of his fascination for biopharmaceutical characterization by mass spectrometry he got enrolled in a Marie Curie Horizon 2020 PhD consortium, called Analytics for Biologics (A4B). He performed his PhD research at the Leiden University Medical Center in the Center for Proteomics and Metabolomics under the supervision of Prof. Dr. Manfred Wührer and Dr. Elena Domínguez Vega. During this time he developed in close collaboration with industrial partners methods for the structural characterization of antibodies and SARS-CoV-2 RBDs and developed an approach for the functional characterization of antibodies and Fc receptors. After his PhD he will continue this line as a postdoctoral researcher at the CPM in collaboration with Roche Diagnostics in Penzberg.

List of publications

1. **Gstöttner C**, Klemm D, Habegger M, Bathke A, Wegele H, Bell C, Kopf R. *Fast and Automated Characterization of Antibody Variants with 4D HPLC/MS*. Analytical Chemistry 2018, 90, 2119–2125 (**Chapter 2**).
2. Pot S*, **Gstöttner C***, Heinrich K, Hoelterhoff S, Grunert I, Leiss M, Bathke A, Domínguez-Vega A. *Fast analysis of formulated antibody-derived therapeutics by automated multidimensional liquid chromatography – mass spectrometry*. Analytica Chimica Acta 2021, 1184, 339015 (**Chapter 3**).

*equal contribution

3. **Gstöttner C**, Reusch D, Habegger M, Dragan I, van Veelen P, Kilgour DPA, Tsybin YO, van der Burgt YEM, Wuhrer M, Nicolardi S. *Monitoring glycation levels of a bispecific monoclonal antibody at subunit level by ultrahigh-resolution MALDI FT-ICR mass spectrometry*. mAbs 2020, 12(1), 1682403 (**Chapter 4**).
4. Tyshchuk O*, **Gstöttner C***, Funk D, Nicolardi S, Frost S, Klostermann S, Becker T, Jolkver E, Schumacher F, Koller CF, Völger HR, Wuhrer M, Bulau P, Mølhøj M. *Characterization and prediction of positional 4-hydroxyproline and sulfotyrosine, two post-translational modifications that can occur at substantial levels in CHO cells-expressed biotherapeutics*. mAbs 2019, 11(7), 1219–1232 (**Chapter 5**).

*equal contribution

5. **Gstöttner C**, Nicolardi S, Habegger M, Reusch D, Wuhrer M, Domínguez-Vega E. *Intact and subunit-specific analysis of bispecific antibodies by sheathless CE-MS*. Analytica Chimica Acta 2020, 1134, 18-27 (**Chapter 6**).
6. **Gstöttner C**, Vergoossen DLE, Wuhrer M, Huijbers MGM, Domínguez-Vega E. *Sheathless CE-MS as a tool for monitoring exchange efficiency and stability of bispecific antibodies*. Electrophoresis 2021, 42, 171–176 (**Chapter 7**).
7. **Gstöttner C**, Hook M, Christopeit T, Knaupp A, Schlothauer T, Reusch D, Habegger M, Wuhrer M, Domínguez-Vega E. *Affinity capillary electrophoresis – mass spectrometry as tool to unravel proteoform-specific antibody-receptor interactions*. 2021, manuscript under review (**Chapter 8**).
8. **Gstöttner C**, Zhang T, Resemann A, Ruben S, Pengelley S, Suckau D, Welsink T, Wuhrer M, Domínguez-Vega E. *Structural and Functional Characterization of SARS-CoV-2 RBD Domains Produced in Mammalian Cells*. Analytical Chemistry 2021, 93, 6839–6847 (**Chapter 9**).

9. Van Schaick G*, **Gstöttner C***, Büttner A, Reusch D, Wuhrer M, Domínguez-Vega E. *Anion exchange chromatography–Mass spectrometry for monitoring multiple quality attributes of erythropoietin biopharmaceuticals*. Analytica Chimica Acta 2021, 1143, 166-172.

***equal contribution**

10. **Gstöttner C**, Zhang T, Resemann A, Pengelley S, Suckau D, Asperger A, Wuhrer M, Domínguez-Vega E. *Recombinant SARS-CoV-2 Receptor Binding Domain: Comprehensive Top-Down Sequence Confirmation, Curation and O-Glycosylation Site Determination*. Bruker (Technical Note)
11. **Gstöttner C**, Zhang T, Resemann A, Pengelley S, Suckau D, Asperger A, Wuhrer M, Domínguez-Vega E. *Proteoform characterization of recombinant SARS-CoV-2 Receptor Binding Domains*. Bruker (Technical Note) (In preparation).
12. Bathke A, Bell C, **Gstöttner C**, Kopf R, Klemm D. *Rapid Online Characterization and Reduction of Protein Modifications Using Fully Automated Two-Dimensional High Performance Liquid Chromatography–Mass Spectrometry*. LCGC Europe 2018, Volume 31(1), 10–21.
13. Vergoossen DLE, Plomp JJ, **Gstöttner C**, Fillié-Grijpma YE, Augustinus R, Verpalen R, Wuhrer M, Parren PWHI, Dominguez-Vega E, van der Maarel SM, Verschuuren JJ, Huijbers MG. *Functional monovalency amplifies the pathogenicity of anti-MuSK IgG4 in myasthenia gravis*. PNAS 2021, 118(13), e2020635118.
14. Lock S, **Gstöttner C**, Wuhrer M, Dominguez-Vega E. *Monitoring exchange efficiency and stability of bispecific antibodies (BsAb) by Capillary Electrophoresis-Mass Spectrometry (CE-MS)*. Sciex 2021, Document number: RUO-MKT-02-12642-A (Technical Note)
15. Komuczki D, Dutra G, **Gstöttner C**, Dominguez-Vega E, Jungbauer A, Satzer P. *Media on-demand: Continuous reconstitution of a chemically defined media directly from solids*. Biotechnology and Bioengineering 2021, 1-13.
16. Van Schaick G, Domínguez-Vega E, **Gstöttner C**, van den Berg-Verleg JH, Akeroyd M, Olsthoorn MMA, Wuhrer M, Heck AJR, Abello N, Franc V. *Native structural and functional proteoform characterization of the prolyl-alanyl-specific endoprotease EndoPro from Aspergillus niger*. Journal of Proteome Research 2021, 20, 4875-4885.

17. Van de Bovenkamp FS, Dijkstra DJ, Borggreven NV, Pool J, Zuijderduijn R, Abendstein L, **Gstöttner C**, Nilsson SC, Holmdahl R, Gelderman KA, Blom AM, Domínguez-Vega E, Parren PWHI, Sharp TH, and Trouw LA. *Protecting cells from complement-mediated cytotoxicity using engineered bispecific antibodies*. (In preparation)

Bookchapters:

18. **Gstöttner C**, Kaur H, Wuhrer M. *Glycosylation analysis*. Chapter in: Monoclonal Antibodies Physicochemical Analysis. Elsevier 2021, 1st Edition.
19. **Gstöttner C**, Haselberg R, Wuhrer M, Somsen GW, Domínguez-Vega E. *Assessment of macro- and microheterogeneity of monoclonal antibodies using Capillary Electrophoresis hyphenated with Mass Spectrometry*. Chapter in: Capillary Electrophoresis-Mass Spectrometry: Methods and Protocols, Springer Methods in Molecular Biology in press.

PhD Portfolio

PhD student:	Christoph Johann Gstöttner
PhD period:	January 2018 – November 2021
Promotors:	Prof. Dr. Manfred Wuhrer
Co-Promotor:	Dr. Elena Domínguez Vega
Department:	Center for Proteomics and Metabolomics

Training Courses

Soft skills:

- A4B Summer School
- Personal effectiveness workshop and talent dynamics
- Leadership and Career Development
- Scientific Presentation and writing
- Legal requirements, industrial application and entrepreneurship
- Grant writing workshop
- Job application strategies

General training:

- BROK
- Basic Methods and Reasoning in Biostatistics
- EFTMS Course (Basic training in FT-ICR)
- Advanced FT-ICR User School
- Protein Chromatography – Engineering Fundamentals and Measurements for Process Development and Scale-up
- Top-Down and Middle-Down Characterization of Proteins

Awards and Grants:

- Innovation grant from Roche Diagnostics (Germany)
Project title: “Automatized workflows for gene therapy product characterization”
Co-applicant: 96.035 Euro
- Global Strategy Team (GST) grant from F. Hoffmann-La Roche (Switzerland)
Project title: “Rapid antibody analysis using multidimensional separations and mass spectrometry”
Co-applicant: 66.613 Euro
- CASSS Student Travel Grant
- Separations Travel Award
- Best Poster Award (EFTMS Conference)

Secondments:

- 2018 - Roche Diagnostics (Department Analytical Biochemistry Technical Development Europe & GLP Analytics (PTDEA), Penzberg, Germany)
- 2019 - Roche Diagnostics (Department Analytical Biochemistry Technical Development Europe & GLP Analytics (PTDEA), Penzberg, Germany)

Supervision

- Parnia Shams Ghahfarokhi (MSc)
- Sanne Pot (MSc)
- Sacha A. Boon (MSc)
- Annika van der Zon (MSc)

Scientific contributions in conferences:

Poster presentations:

- Coupling of capillary electrophoresis with FTICR-MS for enhanced characterization of intact proteins. **13th European Fourier Transform Mass Spectrometry (EFTMS)**, 2018, Freising, Germany.
- Sheathless CE-MS for the characterization of new therapeutic antibody formats. **Global CESI-MS Symposium**, 2018, Leiden, The Netherlands.
- How many chains can we see? A novel middle-down approach based on MALDI-FTICR MS for fast and easy characterization of modified bispecific antibodies **Chemistry as innovative Science (CHAINS)**, 2018 Chemistry, Veldhoven, The Netherlands.
- Characterization of monoclonal antibodies and their modifications by MALDI-(in-source decay)-FT-ICR MS. **1st European Top-Down Proteomics Symposium (TDP)**, 2019, Paris, France.
- MALDI-(in-source decay)-FT-ICR MS for the fast characterization of new therapeutic antibodies. **Analytical Technologies Europe (ATEurope)**, 2019, Dublin, Ireland.
- Studying the structure function relationship of new antibody-derived therapeutics using sheathless CE-MS. **Analytical Technologies Europe (ATEurope)**, 2020.

Oral Presentations:

- Structural and functional characterization of monoclonal antibodies and new antibody formats by mass spectrometry. **Chemistry as innovative science (CHAINS)**, 2019, Veldhoven, The Netherlands.

- Affinity Sheathless CE-MS for the selective study of FcRn - Antibody interaction. **36th International symposium on microscale separations and bioanalysis (MSB)**, 2020.
- Structural and functional characterization of Spike-RBD produced in different organisms. **eNVMS meeting on Covid-19 Research**, 2021.
- Affinity sheathless-CE-MS as a new tool for the functional assessment of antibody - Fc receptor interactions. **40th International Symposium on the Separation of Proteins, Peptides & Polynucleotides (ISPPP)**, 2021.

Round tables:

- Phase-relevant Protein Characterization, **Analytical Technologies Europe (ATEurope)**, 2020 Virtual symposium, November 2020.

Acknowledgments

First of all, I would like to thank everyone who is able to attend my defence during this difficult corona situation. But also, I want to thank you all for the last 4 years where I learned a lot and had a lot of fun.

A great thanks goes to my supervisor Manfred, who enabled me to perform a PhD in his laboratory. You always supported me during our meetings, especially those with Roche where we all could brainstorm and come up with new ideas and projects. Thanks for critically questioning my ideas, but also for giving me the support necessary to realize them. Thanks also for giving me feedback and thereby allow me to strengthen my weaknesses.

A special thanks also to my co-promotor Elena who always supported me during the 4 years. You always trusted me in the lab and really gave me freedom to develop my own ideas, especially with the mD-LC. I learned a lot from you about MS but muuuuuch more about CE. I remember still the moment when we tried for the first time affinity CE-MS. I was really amazed and surprised that it worked out so nice. You were a great supervisor, not only in respect to my PhD, but even more importantly you prepared me perfectly for my future career.

Thank you, Simone, for welcoming and supervising me in the first period of my PhD. You taught me a lot about MALDI and especially FTICR. I learned (and still learn) plenty about this impressive and very complex instrument. It was nice to work with you on the first publications of my thesis. Thanks also to Yuri for reviewing my PhD thesis and give it the last shape.

I would like to thank my amazing paranymphs, Guusje and Steffen. Especially Guusje, not only for guiding me through my defence but also because you taught me a lot about Dutch culture and beautiful places (Achterhoek, Zeeland (Kokkel and Oyster collecting), Amsterdam...). You were a great officemate and friend and I really enjoyed the time we had together. Also, I would like to thank my unofficial paranymphs Willem and Elena, which supported Guusje making possible that I defend together with Steffen.

A special thanks also to Wei and Di, who explained me so much about China. I enjoyed cooking together Hotpot or going to the beach for a swim with the inflatable crocodile. Here again a great thanks to the local of Katwijk, Willem, which knows the best hidden beach spots, where 18 degrees feel like a hot summer day. Also, you were always a good friend and we had several 'hikes' in the dunes together, which I enjoyed a lot!

Thanks also to all the students (and nearly PhD student) which I supervised in the last 4 years, Annika and Sacha, Parnia and Sanne. It was very nice working with all of you. I also learned guiding people and planning projects. I enjoyed working with all of you and wish you good luck on your future way.

Thanks to my consortium and to all my A4B colleagues for the nice training events and meetings we had in the course of the almost 4 years.

Ein herzliches Dankeschön an German #2, Büronachbar, A4B Partner, Paranymphe und Freund, Steffen. Unsere Geschichte reicht schon weit zurück, als wir beide noch Roche Praktikanten waren und zudem im selben Haus lebten. In Basel aber auch hier in den Niederlanden hatten wir eine super Zeit. Wir sind viel gereist und übernachteten in den verrücktesten Unterkünften, entweder mit dem Gefühl, dass sich der Flughafen oder Bahnhof in unserem Hotelzimmer befindet, auf einem Boot mit Sauna oder einem Schloss. Alles Gute in SSF.

Ein riesiges Dankeschön für die super Zusammenarbeit an die Kollegen von Roche. Vor allem Dietmar und Markus Habberger vielen Dank für die Hilfe in diversen Projekten. Es war super schön, wenn ihr in Leiden wart, bzw. wir in Penzberg. Vielen Dank auch an Michaela für die unzähligen Proben die du nach Leiden geschickt hast. Des Weiteren Dankeschön an Robert, Anja, Ingrid, Sina, Katrin, Raphael, Michi und Markus Haindl für die gute Zusammenarbeit. Vor allem vielen Dank an Dietmar, Michi und Christian, für eure sehr hilfreichen Ratschläge vor aber auch während meinem PhD!

Ein herzliches Dankeschön auch an meine Familie und Freunde, die mich immer unterstützt haben und mich zu Hause immer wieder zur ein oder anderen Wanderung herzlich empfangen haben. Vielen Dank Papa, Mama und Reinhold, dass ihr mir auf meinem Weg immer geholfen habt, egal wohin die Reise ging, entweder nach Regensburg, Basel, Penzberg, Stuttgart oder nach Leiden. Die letzte Zeit war ein bisschen schwierig, da wir uns wegen Corona nicht bzw. nur online sehen konnten, aber es kommen sicher wieder andere Zeiten und mal sehen, wo die Reise als nächstes hingeht.

