

Mass spectrometric analysis of protein species of biologics

The global market for therapeutic proteins has been growing fast in the past decade. In comparison with conventional small-molecule drugs therapeutic proteins have a very complex chemical composition. These proteins are typically composed of dozens to hundreds of amino acids which are often post-translationally modified. Recombinant therapeutic proteins are produced by living organisms resulting in a heterogeneous mixture of recombinant protein species (isoforms, proteoforms). The pharmaceutical function and biological and toxicological properties of biologics and follow-on biosimilars are critically related to their exact chemical composition. Therefore, analytical tools capable of identifying and quantifying therapeutic protein species and their accompanying species, often differing in a few chemical moieties only, are necessary to guarantee efficacy and safety of therapeutic proteins. In this review we focus on the application of mass spectrometry for analyzing the exact chemical composition of therapeutic protein species, highlighting top-down mass spectrometry.

In this review we highlight the analysis of therapeutic proteins with mass spectrometric methods. Special focus is given to the analysis of therapeutic proteins with **top-down mass spectrometry** (TD-MS) or **middle-down MS** (MD-MS), implicating that intact proteins (TD) or large fragments of intact proteins (MD) are infused into the mass spectrometer and analyzed before and after fragmentation.

Therapeutic proteins are obtained from biological sources [1–4] in contrast to small molecule drugs, which are usually produced by chemical synthesis. Therapeutic proteins are manufactured either by recombinant technologies [1,2,4] or purified from sources such as human blood plasma [5]. As a result of their biological origin, large size and the requirement for complex processing steps, therapeutic proteins are not homogenous but are composed of many different species which differ in their exact chemical composition [1,3,6]. Heterogeneity of the chemical composition of recombinant therapeutic **protein species** is initially induced in host cells dur-

ing transcription, translation and protein synthesis. Additional heterogeneity can be introduced during downstream processing as well as during storage. In addition to changes in the amino acid sequence, glycosylation, deamidation, oxidation, racemization, condensation and reduction reactions can also occur, affecting the efficacy and safety of the therapeutic proteins.

Ideally a therapeutic protein species with a defined therapeutic action is characterized by its exact chemical composition including its full amino acid sequence (100% sequence coverage), all post-translational modifications (PTMs) including the positions of the modifications on the amino acids, the exact composition and position of glycans, position of the cysteines forming disulfide bonds and non-covalently bound cofactors. The therapeutic efficacy of a therapeutic protein species critically depends on its exact chemical composition [7]. Small changes in the chemical composition of a therapeutic protein species, forming a new species [8–10],

Verena Richter¹, Marcel Kwiatkowski¹, Maryam Omid¹, Azam Omid², Wesley D Robertson³ & Hartmut Schlüter*¹

¹University Medical Centre Hamburg-Eppendorf, Institute of Clinical Chemistry, Mass Spectrometric Proteomics, Hamburg, Germany

²Berlin Institute of Technology, Department of Ecological Impact Research & Ecotoxicology, Berlin, Germany

³CFEL - Atomically Resolved Dynamics Division, Max Planck Research Department for Structural Dynamics at the University of Hamburg, Hamburg, Germany

*Author for correspondence:
E-mail: h.schluefer@uke.de

Key Terms

Top-down mass spectrometry: Mass spectrometry analysis of intact proteins without previous proteolytical or chemical processing. The intact proteins are fragmented within the mass spectrometry to obtain sequence information.

Middle-down mass spectrometry: Proteins are cleaved in a few large fragments and analyzed by mass spectrometry. To obtain sequence information the large protein fragments are dissociated within the mass spectrometry.

Protein species: In this review the term 'protein species' instead of the term 'isoform' is used. The term was introduced by Jungblut *et al.* 1996 [20] and defined in detail with respect to the proteome by Jungblut *et al.* 2008 [8]. In 2012 Neil Kelleher introduced the term 'proteoform' [145]. Both terms can be used for the description of proteins which are coded by one single gene but differ in their exact chemical composition. The term 'proteoform' is gene-centric whereas the term 'protein species' is chemo-centric, thus it also includes man-made modifications introduced by chemical derivatization or biotechnological processes, as we describe here.

Biologics: According to the US FDA, biologics include human blood and plasma and their purified components, bacterial and viral vaccines, proteins such as growth factors, monoclonal antibodies, somatic cells and gene therapy products.

can even result in life-threatening side effects [11] as was observed with Cetuximab [12]. Here a Cetuximab species containing a galactose-alpha-1,3-galactose group was responsible for inducing anaphylaxis in patients having an IgE specific for galactose-alpha-1,3-galactose.

To ensure the efficacy of therapeutic proteins and the safety of patients, analytical methods for an unambiguous species identification as defined by their exact chemical composition as well as quantification are absolutely mandatory.

Monoclonal antibodies (mAbs) are one of the most important types of biopharmaceuticals. They comprise nearly 63% of biologics on the market [13]. Antibodies are glycoproteins composed of two heavy chains (HC) and two light chains (LD) and typically have a molecular mass of approximately 150 kD (Figure 1). Immunoglobulin G (IgG) and its derivatives are the most commonly used therapeutic mAbs [14]. The manufacturing and purification process for biologics is complex [3,15,16]. Minor changes in the individual bioprocessing steps either in up- or down-stream processing, can result in changes to the exact chemical composition of the therapeutic protein increasing the number of therapeutic protein species. As a consequence, the newly formed species may have a decreased efficacy or a total loss of activity [7,17-19]. In the worst case, these species can induce

severe immunotoxic side effects [20], as mentioned above. Many factors can be responsible for changes in the exact chemical composition of the therapeutic protein species. In the host cell, splicing of transcripts can result in the generation of different amino acid sequences [9,21]. Changes in culturing conditions can change the activities of the host cell's enzymes, which can add PTMs to the desired therapeutic protein [22]. Glycosylation steps of the therapeutic protein are often affected by this problem, resulting in variations in the glycan structure [1,23]. Host cell peptidases and proteases can hydrolyze the recombinant therapeutic protein producing truncated therapeutic protein species [23]. Therefore, strict manufacturing control is

mandatory to produce the therapeutic protein species with the desired chemical composition. In order to receive approval for commercial production of therapeutic proteins by authorities such as the US FDA and European Medicines Agency, the qualitative and quantitative composition of the therapeutic protein species must be determined by various analytical methods and their efficacy and safety must be further investigated by biological, toxicological and clinical tests [1,18]. These experiments provide a correlation between the exact chemical composition of the various therapeutic protein species and their efficacy and immunogenicity.

Bioanalytical tools used for the comprehensive analysis of therapeutic protein species are also very important for monitoring the protein stability. Unstable proteins may undergo spontaneous changes to their chemical composition that affect their biomedical functions [23]. Instability can result in isomerization, as observed at aspartyl and asparaginyl residues, which can undergo autocyclization to form isoaspartate, lead to the oxidation of cysteine residues and result in aggregation or denaturation [7].

Currently, the first patents for therapeutic proteins are resulting in out-of-patent follow-on therapeutic proteins, also termed biosimilars, which are being increasingly released into the market. Thus a biosimilar is a copy of a defined therapeutic protein with a defined therapeutic application [1,3,15,24]. Since the exact chemical composition is critically dependent on many up- and downstream-processing parameters, biosimilars are usually not completely identical to the originator protein [1,3,15,24].

In order to receive FDA approval, the similarities of a new biosimilar and its originator therapeutic protein must be determined to ensure efficacy and safety of the biosimilar [7,24]. The comprehensive characterization by bioanalytical tools of the exact chemical composition and its comparison with that of the originator species is one of the major requirements for its approval by the authorities.

Relevant techniques for the characterization of therapeutic proteins

As described above, the production process results in the production of diverse, closely related species along with the target therapeutic protein species. Therefore, it is important to apply a fractionation method which guarantees the separation of all of these species and which can provide a reliable approximation of the relative quantities of the different species. Although modern mass spectrometers offer high resolution power, there are limitations concerning the differentiation of species by MS. This is especially true for large proteins such as mAbs whose species differ only in small modi-

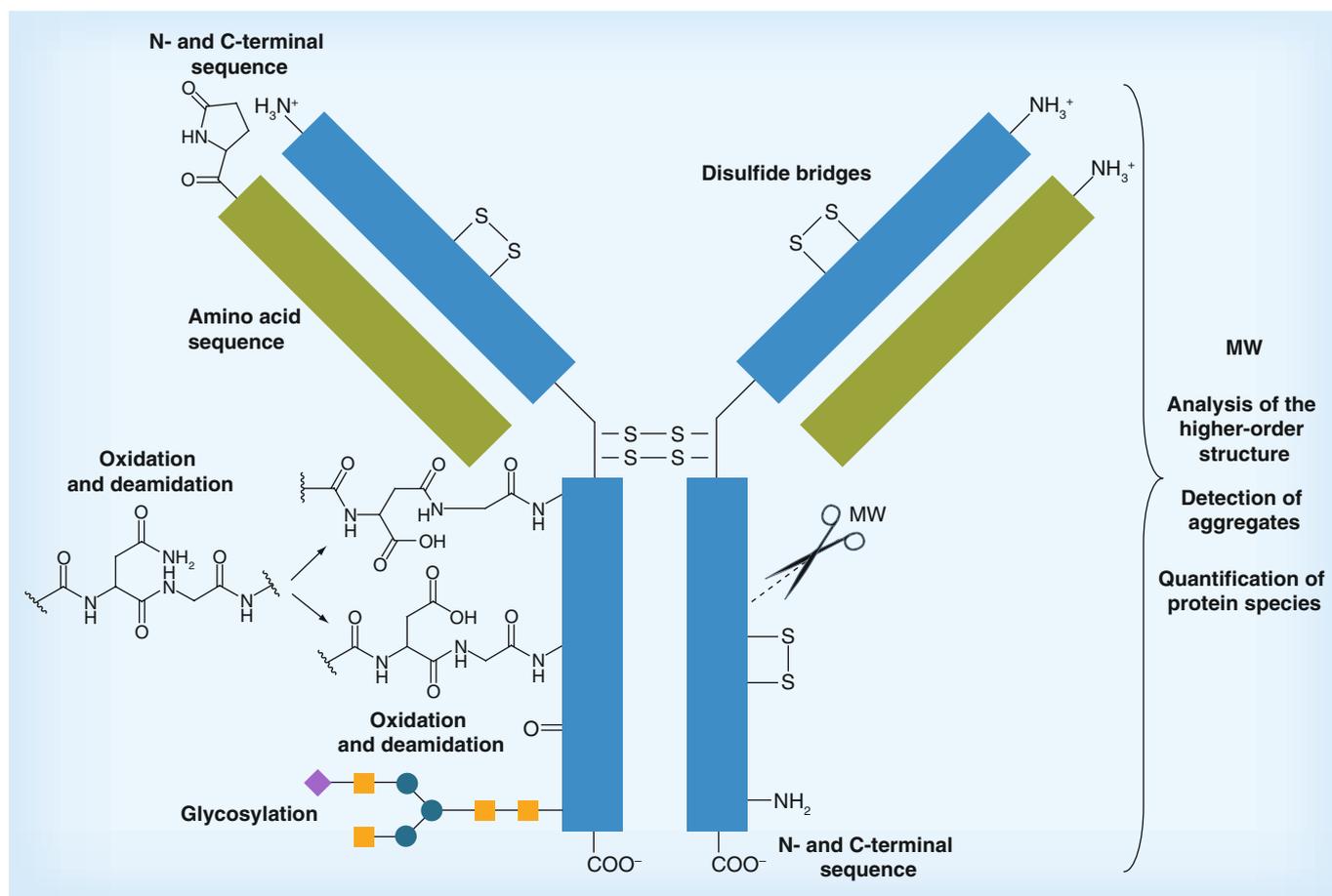


Figure 1. IgG antibody. The labels can also be found in the text and in Table 1 and refer to the analytics applicable to the illustrated region of the antibody. A comprehensive overview of main modifications including chemical structures of modified residues is given by Bischoff and Schlüter [10].

fications, such as deamidation. In order to simplify the following discussions, we assume that protein species are separated to near homogeneity prior to their mass spectrometric analysis.

The active therapeutic protein species with high efficacy is defined by its exact chemical composition which is comprised of the complete amino acid sequence, all PTMs including its glycan composition, as well as disulfide bonds. During the development of a new therapeutic protein the exact chemical compositions of the most potent species, as well as less active accompanying species, are determined. These data serve as a reference for the identification of the product and its byproducts during the production process. Figure 1 presents the important features of the chemical composition for therapeutic proteins exemplified by a scheme of an antibody. The labels in Figure 1 correspond to the subheadings in the text below and in Table 1, which lists structural features and the corresponding analytical methods for detecting these features. Under the heading 'MW', and the following sections, the features nec-

essary for describing the exact chemical composition of species are briefly highlighted.

Introduction of PTMs usually results in a change in the MW and/or isoelectric point (pI) of the species. Thus isoelectric focusing (IEF) is well suited for the detection of species with different charges [25]. Techniques based on capillary electrophoresis (CE) are especially useful because they offer the highest resolution currently obtainable, including capillary IEF (cIEF). However, CE and cIEF are restricted to the separation of very small amounts, because of the small volume of the capillaries. Therefore, CE and cIEF are not suited for preparative separation of species which yields purified species that require further characterization by other analytical or biological assays. For this aim, preparative IEF or ion exchange chromatography are usually applied. Cation-exchange chromatography was introduced to measure species with different pI, in particular lysine variants, by Harris in 1995 [26] and is now a popular and widely accepted technique in pharmaceutical analysis [25].

Table 1. Overview of relevant techniques for the characterization of therapeutic proteins.		
Section	Characterization	Techniques
MW	MW	LC-ESI MS, MS/MS [103,202] MALDI MS [202] Native MS [202] SEC in combination with light scattering [32] 1DE [203], SDS-PAGE, CE-SDS [100,203]
Amino acid sequence	Amino acid sequence	LC-MS/MS of proteolytic peptides resulting from enzymatic digestion of the protein species (BU) [36,73] N/C-terminal sequencing by MALDI-TOF-TOF using ISD and T ³ sequencing [184,185] TD-MS [144-148] MD-MS [148,199]
N- and C-terminal sequence	N-terminal sequence	Edman sequencing [38] MALDI-TOF-TOF using ISD [39] and T ³ sequencing [184,185] TD-MS [144-148] MD-MS [148,199]
N- and C-terminal sequence	Pyro-glutamination	BU [44,47,48] TD-MS [148] MD-MS [200]
N- and C-terminal sequence	C-terminal sequence	MALDI-TOF-TOF using ISD [39] and T ³ sequencing [184,185] TD-MS [144-148] MD-MS [148,199]
Instrumentation	C-terminal lysine	BU [25,44,47]
Oxidation and deamidation	Oxidation	BU [40-43] TD-MS [147] MD-MS [199,200]
Oxidation and deamidation	Deamidation	BU [44-46]
Less frequent PTMs of therapeutic proteins	γ -carboxylation [53]	BU
Less frequent PTMs of therapeutic proteins	β -hydroxylation [53]	BU
Less frequent PTMs of therapeutic proteins	Tyrosine-O-sulfatation [53]	BU
Less frequent PTMs of therapeutic proteins	Amidation [53]	BU
Less frequent PTMs of therapeutic proteins	Phosphorylation [53]	BU TD-MS [204,205]
Less frequent PTMs of therapeutic proteins	Succinylation [206]	BU
Less frequent PTMs of therapeutic proteins	Acetylation	BU
Less frequent PTMs of therapeutic proteins	ADP ribosylation	BU
Less frequent PTMs of therapeutic proteins	Isomerization: aspartate [46]	RP-HPLC for separation of isomers followed by BU [207] ISOQUANT detection kit (Promega, WI, USA) [208-210]

The different sections in the table correspond to the subheadings found in text and in [Figure 1](#).
 1DE: 1D native gel electrophoresis; AF4: Asymmetric flow field flow fractionation; AUC: Analytical ultracentrifugation; BU: Bottom-up; CD: Circular dichroism; CE: capillary electrophoresis; DLS: Dynamic light scattering; ELISA: Enzyme-linked immunosorbent assay; ESI: Electrospray ionization; HDX: Hydrogen/deuterium exchange; HPLC: High-performance liquid chromatography; IM: Ion-mobility; ISD: In-source decay; LC: Liquid chromatography; MALDI: Matrix-assisted laser desorption/ionization; MD: Middle-down; MS: Mass spectrometry; SDS: Sodium dodecyl sulphate; PTM: Post-translational modification; RP: Reversed phase; SDS-PAGE: Sodium dodecyl sulfate-polyacrylamide gel electrophoresis; SEC: Size exclusion chromatography; SRM: Selected reaction monitoring; TD: Top-down; TOF: Time-of-flight.

Table 1. Overview of relevant techniques for the characterization of therapeutic proteins (cont.).		
Section	Characterization	Techniques
Glycosylation	N-glycans: identification of the amino acid positions derivatized with glycans	Deglycosylation with PNGase F followed by BU [211] or HPLC [301,302]
	O-glycans: identification of the amino acid positions derivatized with glycans	Deglycosylation by hydrazinolysis followed by BU review [59] or HPLC [61]
Disulfide bonds	Disulfide bridges	BU in combination with derivatization [74–76] MALDI-ISD [212] IM-MS [213]
Analysis of the higher-order structure	Higher order structure	MS: HDX-MS [79,214,215], native MS [79], IM-MS [83] Chemical cross-linking [86,87] Spectroscopy: NMR [81,82], CD [80] x-ray crystallography, x-ray scattering
Detection of aggregates	Aggregates	HDX-MS [104,105] Native MS [103] Chemical-cross linking [89,106] Oxidative footprinting [107–109] DLS [97], AUC [93,95,97], AF4 [99] SEC [90] AUC [90] 1DE [32], SDS-PAGE [96] CE-SDS [96]
Quantification of protein species	Quantification of species	SRM [112,115,116], LC-MS and -MS/MS with internal standards [110] Quantitative amino acid analysis [216] Immunoassay: ELISA [111]

The different sections in the table correspond to the subheadings found in text and in [Figure 1](#).
 1DE: 1D native gel electrophoresis; AF4: Asymmetric flow field flow fractionation; AUC: Analytical ultracentrifugation; BU: Bottom-up; CD: Circular dichroism; CE: capillary electrophoresis; DLS: Dynamic light scattering; ELISA: Enzyme-linked immunosorbent assay; ESI: Electrospray ionization; HDX: Hydrogen/deuterium exchange; HPLC: High-performance liquid chromatography; IM: Ion-mobility; ISD: In-source decay; LC: Liquid chromatography; MALDI: Matrix-assisted laser desorption/ionization; MD: Middle-down; MS: Mass spectrometry; SDS: Sodium dodecyl sulphate; PTM: Post-translational modification; RP: Reversed phase; SDS-PAGE: Sodium dodecyl sulfate-polyacrylamide gel electrophoresis; SEC: Size exclusion chromatography; SRM: Selected reaction monitoring; TD: Top-down; TOF: Time-of-flight.

Regulatory quality demand guidelines from regulatory agencies, such as the FDA, European Medicines Agency or collectively the International Conference on Harmonization (ICH), require the characterization of biopharmaceuticals by using state-of-the-art analytics [27–29]. At the time of submission, the manufacturer is required to supply the in-depth characterization of the exact chemical composition of the most efficient therapeutic protein species (reference species), as well as the composition, efficacy and toxicology of other species which are generated during manufacturing [29]. Reference species standards are important for method development and process control, validation, assessment of purity, definition of relative potency, and safety. The primary reference standard is commonly derived from the manufacturing process established prior to the start of clinical trials. The primary reference standard should be available in sufficient quantities for usage over the next 10–20 years. It is also used to qualify new

working standards. Working standards are regularly produced for lot release and stability testing [30].

» MW

The MW is an important parameter toward the characterization of the exact chemical composition of a protein species. Classical electrophoretic and chromatographic techniques are typically applied, such size exclusion chromatography (SEC), often in combination with MS. Mazur *et al.* reported the use of a 2D chromatography, combining SEC and reversed-phase LC-MS (RP-LC-MS), for the characterization of mAb breakdown products by TD-MS [31].

For a fast qualitative analysis of homogeneity and the MW, 1D native gel electrophoresis and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) are often used [32]. Although both methods are rather robust, they provide rather approximate information concerning homogeneity, MW and quantity of species. Although more labori-

Key Term

Bottom-up mass spectrometry: mass spectrometry analysis of proteins on the peptide level. Proteins are proteolytically cleaved and the resulting peptides are fragmented within the mass spectrometry to obtain sequence information.

ous and time consuming, 2D electrophoresis provides significantly more insight into the heterogeneity of a therapeutic protein fraction including the MWs of the diverse species, their pI and their relative quantities.

CE methods as mentioned above, offer more detailed information useful for the characterization of a variety of therapeutic protein species [33–35].

The most accurate determinations of the MWs of species are obtained by modern MS-based methods. Generally, for the analysis of intact proteins, matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) are applied as desorption and ionization techniques. MS-based analytical methods will be discussed below in detail.

» Amino acid sequence

For recombinant therapeutic protein species the theoretical amino acid sequence is given by the nucleotide sequence of the gene which is cloned into the vector. However, during gene expression alternative splicing events may occur or host cell peptidases and proteases may truncate the protein that has been synthesized at the ribosome, leading to deviations from the theoretical amino acid sequence. Digestion of homogenous therapeutic protein species with trypsin as well as other proteases, such as endoproteinases Lys-C, Lys-N Asp-N, Glu-C and chymotrypsin, yields complementary information that is required to achieve 100% sequence coverage (bottom-up MS approach; BU) [36]. The comparison of proteolytic peptide profiles of the reference protein species with purified protein species from protein fractions collected at different steps of the up- and down-stream processing steps provides information concerning the quality of the therapeutic protein and the presence and quantity of undesired species characterized by changes in the exact chemical composition.

If a peptide fragment spectrum measured in a BU-approach cannot be assigned to the reference protein, often because of unusual modifications of the peptide side chains, de-novo sequencing, multiple fragmentation (MSⁿ) by ion-trap MS and/or Edman sequencing should yield the necessary data for deciphering the chemical composition of the proteolytic peptide.

» N- & C-terminal sequence

The amino acids glutamate or glutamine, if located at the N-terminus of a protein, can spontaneously or enzymatically be catalyzed to convert to the cyclic pyrrolidone (pGlu). Many mAbs contain pGlu

on the N-termini of the LD and HC [37]. The loss of a primary amine residue results in a more acidic protein.

N-terminal sequencing is a requirement according to the ICH Q6B guideline. Typically, Edman degradation [38] is used for the determination of 5–15 amino acid residues [39].

Truncation of C-terminal lysine is commonly observed in therapeutic mAbs and is a critical quality attribute. Therefore it must be monitored closely for product consistency. Together with lysine, arginine residues at the C-terminus of the HC are often absent in the final protein products of mammalian cells [26]. It is believed that carboxypeptidase activity and differences in expression and protein synthesis, following changes in the production process, are responsible for the lysine variants [25].

C-terminal sequencing is possible today with MALDI MS/MS instruments using in-source decay (ISD) [39] and T3-sequencing as well as TD-MS (explained in detail in ‘Analysis of therapeutic proteins by TD-MS’).

» Characterization of further PTMs

Analysis of peptide fragments of purified therapeutic protein species obtained by a combination of proteases by LC–MS/MS methods (BU-approach) can contribute valuable information about PTMs. This is especially true if the modified peptides are already known and characterized or the modifying group is available from a search engine such as MASCOT. The comparison of the LC–MS profiles of the reference protein species with that of the analyzed protein fraction can discern if a peptide from the therapeutic protein is modified with respect to the reference peptide.

BU is an important tool for PTM studies and permits the detection of oxidation [40–43], deamidation [44–46], and truncated protein species [47]. Furthermore, this method can be generally used for the detection of mAb variants such as N-terminal cyclization [44,47,48], C-terminal lysine processing [25,44,47] and glycosylation [49,50].

» Oxidation & deamidation

Analyzing methionine oxidation and deamidation is a key requirement of the ICH Q6B guidelines. These two modifications frequently occur due to thermal stress during purification and storage, for example.

Oxidation of proteins is usually observed for methionine and cysteine residues [10,51] and less frequently in tryptophan or tyrosine residues. The addition of oxygen atoms results in a change of hydrophobicity and mass (+15.9949 Da/+31.9898 Da).

The amino acids asparagine and to a smaller extent glutamine are prone to deamidation, a spontaneous non-enzymatic reaction [52]. Deamidation

(+0.9840 Da) leads to a change in the pI and therefore the presence of charge heterogeneity. For further details of the chemical reactions of side chains of amino acids see Bischoff and Schlüter [10].

» Less frequent PTMs of therapeutic proteins

Proteins can exhibit many more PTMs than the ones described above, but only a subset of PTMs are associated with therapeutic proteins. For example, γ -carboxylation and β -hydroxylation are characteristic modifications for only a few native and recombinant proteins, for example, blood factor VIIa and IX and activated protein C [53]. Other PTMs which are characteristic for only a limited number of therapeutic proteins are tyrosine-O-sulfatation, amidation or phosphorylation [53]. Acetylation is another modification that affects protein stability by the incorporation of a COCH₃ group [54]. As with phosphorylation and acetylation, ADP-ribosylation of proteins is involved in several intracellular processes and is rarely associated with therapeutic proteins. The attachment of one or more ADP and ribosome moieties to proteins is arranged by ADP-ribosyltransferase enzymes and can effect cell signaling and control. Zhang *et al.* reported the identification of lysine succinylation of the protein isocitrate dehydrogenase in *Escherichia coli* [55] and others have demonstrated the broad role of this PTM in cellular metabolism [56,57]. Isomerization of the amino acid asparagine 102 has also been described as a PTM of recombinant antibodies resulting in a low potency species [46]. Most of these PTMs can be detected by BU if high-resolution mass spectrometers are applied.

» Glycosylation

Glycosylation is an important structural feature in many biopharmaceuticals [53]. The ICH Q6B guideline requests the analysis of the carbohydrate content, carbohydrate chain structure, oligosaccharide pattern, and glycosylation sites. The heterogeneity of glycans, the similarity in their monosaccharide subunits and the variety of feasible linkages between the units, usually necessitate the application of more than one analytical method for carbohydrate characterization [58].

Strategies for glycoprotein analysis using HPLC

A general protocol for glycosylation analysis includes the release of glycoprotein-bound glycans, fluorescent labeling of the liberated glycans, and high-performance liquid chromatography (HPLC) purification and separation with online-fluorescence detection.

A common strategy for determining sialic acid residues is to first release sialic acid residues from the glycoprotein by acid hydrolysis, followed by fluorescent

labeling of released sialic acids with 1,2-diamino-4,5-methylenoxybenzene and subsequent HPLC with relative quantification [59,60]. N-linked glycan profiles are obtained by enzymatic liberation of the glycan by the enzyme PNGase F. O-linked glycans are commonly released by mild hydrazinolysis prior to derivatization with 2-aminobenzamide, and separation in normal phase- or RP-HPLC mode. For more information, about O-glycosylation the review by Zauner *et al.* is recommended [61]. For the determination of the composition and the extent of glycosylation, monosaccharide analysis is routinely applied. Monosaccharide residues are cleaved from the protein under acidic conditions, labeled with 2-aminobenzoic acid, and separated by RP-HPLC [301,302]. The obtained sample glycosylation profiles are compared with the reference standard profiles and typically relative quantification is performed. Alternatively, high-performance anion-exchange chromatography with pulsed amperometric detection can be performed where no prior derivatization step is necessary [62,63].

Beside anion-exchange chromatography-, normal phase- and RP-HPLC, hydrophilic interaction liquid chromatography and hydrophobic interaction chromatography are powerful chromatographic techniques to separate glycoproteins as well as liberated glycans [64,65].

MS-based strategies for glycoprotein analysis

MALDI- and ESI-MS are helpful tools in the characterization of glycosylated proteins [66,67] and many publications can be found, for example, for the glycosylation analysis of IgG antibodies [68].

MS-based glycosylation analyses commonly comprise three approaches involving intact glycoprotein analysis, the characterization of generated glycopeptides, and structural analysis of chemically or enzymatically released glycans [68–70]. Carbohydrates can be studied by analyzing the intact protein with MALDI or ESI due to the relatively large mass shift between various glycoforms (mass increment of one hexose unit is 162.0528 Da). MALDI-TOF MS is often used as a primary screening step in order to obtain an overview about the nature and diversity of glycans present [70]. LC-ESI MS/MS with low-energy collision-induced dissociation (CID) is frequently applied for the investigation of the sequence and linkage of glycans [70].

MS analysis at the glycopeptide level after proteolytic digestion of the therapeutic protein or analysis of glycans released from the protein are currently the methods of choice for obtaining sensitive and comprehensive glycosylation information [71,72]. Glycopeptides are often separated by RP [69] or hydrophilic interaction LC [71] prior to analysis by MS.

» Disulfide bonds

Localization of disulfide bonds is a key part in the ICH Q6B guideline. 1D gel-electrophoresis under reducing and non-reducing conditions can be performed to confirm the existence of disulfide-linked subunits of a protein species.

LC-MS/MS analysis of proteolytic peptides (BU-approach) under both reducing and non-reducing conditions yields the positions of the cysteines forming the disulfide bridge [73]. For example to determine disulfide linkage within a mAb requires a proteolytical digest under denaturing conditions including alkylation to prevent artificial disulfide formation [73–76]. After splitting the digest, one part is further reduced using dithiothreitol or tris(2-carboxyethyl)phosphine. The resulting peptide map is compared with the peptide map of the non-reduced part resulting in identification of the disulfide bonds.

Electron capture dissociation (ECD) and electron transfer dissociation (ETD) have been introduced as valuable alternatives to CID [77,78] (see also 'Fragmentation' section).

» Analysis of the higher-order structure

Information gained from analytics of the primary structure form the basis for the characterization of the higher-order structure. A mutation of an amino acid or the introduction of a PTM can have a significant effect on the higher-order structure of the protein. The ICH Q6B guideline recommends the analysis of higher-order structures, including the evaluation of any changes of the 3D structure. However, if extensive physicochemical information is obtained, but the higher-order structure cannot be confirmed, the guideline accepts relevant biological assays as an alternative if combined with specific quantitative measurements. This demonstrates the challenge one is facing when characterizing the 3D shape of complex biomolecules. However, due to the progressive development of specialized analytical tools, one can choose from a variety of techniques [79].

Spectroscopic techniques such as circular dichroism and Fourier transform infrared are commonly used to assess the secondary structure of proteins [80]. Raman spectroscopy is occasionally applied for this purpose. UV/Vis absorption, fluorescence and again circular dichroism are powerful techniques for monitoring alterations in tertiary structure [80].

The additional use of orthogonal methods is essential to characterize the complex biological molecules in more detail. x-ray crystallography, NMR, and small-angle x-ray scattering, are such orthogonal techniques. However, they have inherent limitations that might make their application to the analyses of biopharma-

ceutical products impractical. For example, for x-ray crystallography the protein needs to be amenable for crystallization, which is not always the case due to structural reasons, for example, because of the presence of PTMs. NMR has demonstrated its power especially in the studies of the higher-order structure of small biomolecules. The power of NMR was demonstrated for comparability studies of three non-glycosylated proteins spanning a MW range of 6–67 kDa [81], and in defining the bioactive conformational identity of recombinant human granulocyte macrophage-colony stimulation factor (MW ~14.5 kDa) [82]. Nevertheless, for the analysis of large biopharmaceuticals such as mAbs, these techniques still have limitations, that is, in terms of the size of mAbs, the existence of flexible/disordered regions within mAbs, the amount of material necessary, and the time to perform analysis [37]. Therefore, these techniques are not used in routine analytics, but primarily in formulation development [80].

Recently, MS-based technologies have been developed that can successfully complement the classical biophysical techniques and can partially circumvent their current limitations, that is, native MS, ion-mobility MS (IM-MS), and hydrogen/deuterium exchange MS (HDX-MS) [79,83]. In HDX-MS, exchangeable amide hydrogen atoms are replaced by deuterium, resulting in one additional measurable mass unit per amino acid. The localization and degree of H/D exchanges are determined by LC-MS/MS [84]. The H/D exchange is a function of protein structure and dynamics. HDX-MS is also popular for the characterization of interactions and epitope mapping [85].

Chemical cross-linking followed by enzymatic proteolysis is another method that has been combined with BU MS-based proteomics yielding structural information on the distance constraints within a molecule [86–88]. Unlike NMR or x-ray crystallography, it is applicable to analytes at low concentrations and it does not require protein crystals [89].

» Detection of aggregates

SEC is most commonly performed for a rough estimation of the MW and, more importantly, for detection and quantification of soluble aggregates such as dimers [90]. However, large aggregates may not enter the column and a certain degree of incomplete product recovery has been observed due to non-specific adsorption of the proteins to the chromatographic material. Therefore, aggregates might be overlooked with SEC [90,91]. Because of these limitations, the FDA usually requests verification of results obtained with SEC using orthogonal analytical techniques [29,92], as described as follows.

Analytical ultracentrifugation (AUC) has been proposed as a suitable alternative for SEC [93–95]. So far,

AUC is not used in routine product characterization because of the time-consuming measurement and the relatively high costs for the initial equipment as well as for operating the system [91]. At present, AUC can rather serve as a valuable tool for qualitative cross-check of data obtained by SEC as claimed by regulatory agencies [96].

In addition to AUC, field-flow fractionation and light scattering techniques belong to the column-free techniques which are increasingly applied for the detection and quantification of aggregates. Light scattering methods are highly sensitive to detect small amounts of large aggregates, as the intensity of the scattered light is often directly proportional to the MW [97].

Features of dynamic light scattering (DLS) are rapid and automated measurements, no perturbation, or dilution effects, thus circumventing potential dissociation of reversible aggregates [98]. Although the sensitivity of DLS for the detection of large molecules is excellent, quantification is not possible [96].

Asymmetric flow field flow fractionation (AF4) is a third matrix-free technique that is increasingly applied for aggregation analysis [99]. Similar to AUC or DLS, AF4 is usually used as a support method to assure that the data generated by SEC are accurate. The application of AF4 to biomolecules is rather new and often requires in-depth method development in order to obtain acceptable separation and robustness [96].

In the presence of SDS, non-covalent aggregates are disrupted enabling the detection of covalent aggregates. Under reducing conditions, SDS-PAGE can discriminate between proteins aggregated by means of disulfide bonds and those held together by other covalent bonds [96]. However, with native and SDS-PAGE, it is not possible to detect small amounts of aggregates, depending on resolution achieved and on sensitivity of the staining method [97]. Yet, SDS-PAGE is increasingly being replaced by its modern equivalent CE in combination with SDS (CE-SDS) or coupled to mass spectrometers. CE-SDS can be automated; it offers great reproducibility and online-detection capability and is better suited for robust quantification by UV light absorption rather than dye-binding. A combination of SEC and SDS-PAGE or CE-SDS is usually part of analytics in the quality control of pharmaceutical companies [96,100].

MS has been applied for non-covalent structures of proteins [101,102]. An important difference between the MS-based techniques and the conventional biophysical techniques described above is the ability of MS-based techniques to provide structural details by using fragmentation such as proteolytic digestion [73].

ESI-time-of-flight (TOF) MS is capable of accurate determination of the molecular mass of proteins and

aggregates [103]. Specifically, native ESI-TOF MS has proven its utility in structural characterization of intact protein aggregates owing to the ability of preserving the quaternary structure of a protein and maintaining non-covalent protein interaction [103].

Specifically for non-covalent changes in protein structure, HDX-MS [104,105], chemical-cross-linking [89,106] and oxidative foot-printing [107–109] have been successfully performed.

To sum up, applying SEC as the standard method for the characterization of protein size in combination with orthogonal methods, such as electrophoresis, AUC, DLS and/or AF4, reduces the event of missing important aggregation and truncation forms and is in accordance with the demand of regulatory agencies.

» Quantification of protein species

It must be emphasized that most of the assays for quantification of therapeutic protein species require the protein to be purified to near homogeneity prior to their analysis in order to prevent false results.

Amino acid analysis after total hydrolysis of the therapeutic protein is the gold standard for accurate quantification of proteins and peptides. It is applicable to complex as well as purified samples. Amino acid analysis is a key requirement of the ICH Q6B guideline. When the molar extinction coefficient of the therapeutic protein is known, the protein concentration can be approximately determined by spectrophotometry. The molar extinction coefficient can be assessed by UV/Vis absorbance measurements in combination with amino acid analysis which is in agreement with the ICH Q6B guideline. Total nitrogen determination, such as the Kjeldahl nitrogen assay, is another method for determining the protein concentration needed for definition of the molar extinction factor. However, the molar extinction coefficient is usually determined with the reference standard and all spectrometric measurements should be referenced to the standard. Colorimetric protein assays, such as Biuret or Lowry, can be applied for the determination of the total protein concentration.

With MS, individual protein species can be relatively quantified in the presence of other species [110], however the quality of the MS-based quantification is dependent on the degree of the difference of the chemical structure and the resulting difference in the MW. Protein species differing in one moiety regarding deamidation can not be separated by current mass spectrometers and therefore cannot be quantified sufficiently. Isotopically labeled protein species allow absolute quantification of species if a proteolytic peptide unique for the species is detectable. Thus MS quantification of some species in the presence of other

species is possible but it is recommended to separate the diverse species prior to quantification.

Besides the precise quantification of the therapeutic protein species, process- and product-related impurities as well as contaminants need to be detected, identified and quantified. Process-related impurities are host cell proteins (HCPs), host cell DNA and reagents used in cell culture. Degradation products of the therapeutic protein species occurring during manufacturing or storage are termed product-related impurities. Usually these species formed by degradation have a significantly decreased efficacy and may even cause side effects and therefore need to be monitored and removed.

The amount of contaminant HCP is usually determined using enzyme-linked immunosorbent assays (ELISA) [111]. Antibodies developed for immunochemical detection require prior knowledge about HCPs. Therefore, the use of immunochemical assays can be limited due to the inability to develop specific antibodies. Generic ELISA kits commercially available for the detection of HCPs are less specific than a specific immunoassay, and might not detect all the HCPs present in a sample [112]. However, MS-based assays for the detection of host cells have been developed recently by applying high-resolution mass spectrometers such as orbitraps [113].

LC-MS/MS in selected reaction monitoring (SRM; also called multiple-reaction monitoring) mode offers an alternative to immunochemical quantification methods [114–116]. With SRM, proteotypic peptides representative for specific proteins can be quantified in complex mixtures with high sensitivity, reproducibility and specificity. A big advantage of SRM is its multiplexity. Doneanu *et al.* developed a SRM assay for the quantification of HCP over five orders of magnitude in concentration [112].

A MS host cell assay based on data independent acquisition mode can circumvent the drawbacks of SRM [117].

Analysis of therapeutic proteins by TD-MS

Many of the established protocols for characterization of therapeutic proteins by MS as described above are based on the BU approach, implicating the proteolytic digestion of the therapeutic proteins by trypsin and other proteases prior to their mass spectrometric analysis. In the past few years the mass spectrometric characterization of intact proteins has become an additional important tool termed TD-MS analysis [118]. TD-MS of proteins is aiming for the comprehensive analysis of protein species, ideally including not only the MW determination but also yielding 100% sequence coverage and all PTM data. This aim until now has only

been achieved for some smaller proteins. However, TD-MS is especially useful, if data regarding the exact chemical composition of a defined therapeutic protein and its diverse species have already been collected by BU protocols. TD-MS yields not only data regarding the molecular mass of the diverse species but also additional information about their composition, including partial sequence information as well as some data about posttranslational modifications. Furthermore in combination with LC, TD-MS can provide information about the relative quantitative composition of a mixture of protein species within short analysis times. Thus TD-MS is a helpful tool for accompanying the production process of therapeutic proteins and for quality control of the purified product.

This part of the review will describe the function of mass analyzers, fragmentation techniques and data analysis software programs used for TD-MS as well as applications of TD-MS for characterizing therapeutic proteins.

» Instrumentation

Measuring only the mass of large therapeutic proteins, such as mAbs, without further fragmentation state-of-the-art mass spectrometers and mass resolutions higher than 10,000 provide little additional information and are sometimes counterproductive due to loss in sensitivity [37,73]. Significantly more information can be attained by using TD-MS. Fragmentation of intact proteins in the mass spectrometer yields a large amount of highly charged and sometimes overlapping fragment-ions. High mass resolution is mandatory for resolving the fragment-ions. In order to detect large intact proteins, instruments with higher mass resolving power are required. In this section three different mass analyzers well suited for TD-MS are briefly described.

TOF mass analyzer

In a TOF mass analyzer mass-to-charge ratio (m/z) is determined by measuring the time for ions to move through a field-free region between a source and a detector. Before entering the field-free region, ions with mass (m) and charge $q = z \times e$ are accelerated by an electric field with a potential V . Thereby the potential energy (E_{el}) of the ions in the field is converted into kinetic energy (E_{kin}):

$$E_{kin} = \frac{1}{2}m \times v^2 = q \times V = z \times e \times V = E_{el}$$

Equation 1

After acceleration, ions travel through the field free flight tube with a constant velocity until they reach the detector, a distance (d) away. By knowing the distance

(d), the time it takes ions to reach the detector can be correlated with the corresponding m/z (Equations 2 & 3).

$$t^2 = m/z \left(\frac{d^2}{2 \times e \times V} \right)$$

Equation 2

$$t = d \times \sqrt{\frac{1}{2 \times e \times V}} \times \sqrt{m/z}$$

Equation 3

The smaller ions arrive at the detector first, followed by larger ones. The mass resolving power of a TOF mass analyzer depends on ion flight path length. Typical TOF instruments have a flight path length of approximately 2 m. For further details concerning the function and principles of the TOF mass analyzers, we recommend the reviews of Cotter [119] and Guilhaus [120].

Modern orthogonal acceleration quadrupole TOF (QTOF) mass spectrometers can achieve a mass resolution up to 60,000 on both MS and MS/MS level with a mass accuracy of approximately 1 ppm at a spectral acquisition rate of 20 Hz [121].

Fourier transform ion cyclotron mass analyzer

The function and principles of a fourier transform ion cyclotron resonance (FTICR)-MS are well explained in a number of books and reviews [122–127]. Here a short description is provided about the function of the ICR cell. Ions are trapped in the ICR cell and move in a static magnetic field B , with a specific cyclotron frequency rotation ν_c (Hz):

$$\nu_c = \frac{e \times z \times B}{2 \times \pi \times m}$$

Equation 4

As shown in Equation 4, ν_c is independent of ion energy and amplitude but proportional to the charge and inversely proportional to the mass of an ion. Ion cyclotron motion by itself is not detectable because the ion cyclotron orbit is centered along the z -axis of the cell and does not generate a measurable signal. To obtain a detectable signal, ions are resonantly excited with a rotating or oscillating electric field, yielding a spatially coherent ion packet for a given m/z . Due to the excitation, ions are accelerated and their orbits increase in size, allowing ion cyclotron frequency detection. Ion cyclotron frequencies are measured non-destructively by passing two opposed electrodes on either side of the ion's flight path acquiring a time-domain image current transient of the ions. The time-domain signal is converted by Fourier transformation (FT) into a frequency-domain signal. As shown in Equation 4, the frequency-domain

signal can be correlated to a specific m/z . FTICR mass spectrometers have been shown to provide the highest mass accuracy (<1 ppm) [128] and mass resolution (up to 24,000,000) to date [129]. In a 9.4 T FTICR mass spectrometer (ICR orbital radius at 298K $r = 0.01$ mm), ions with $m/z = 1000$ and $\nu_c = 144.346$ kHz travel at a velocity of 9070 m/s [123]. Given that FTICR resolving power is equal to the number of cyclotron orbits acquired, data acquisition for 3 s of the same ion (m/z 1000) provides a mass resolution of approximately 400,000 [123]. This example illustrates a disadvantage of FTICR instruments: the acquisition time increases with the increase of resolution. However, Nikolaev developed a FTICR cell [129,130] that enables an outstanding resolving power ($R > 20,000,000$) [129] and mass accuracy [131] even with the use of a 7 T magnet. This ParaCell™ is implemented in Bruker solarix XR FTICR mass spectrometer and provides a mass resolution of up to 700,000 with a 1 s scan [132]. Thus the acquisition times for high resolution measurements were significantly reduced, making FTICR MS suitable for coupling with LC.

Fourier transform orbitrap mass analyzer

The orbitrap mass analyzer consists of a spindle-like inner electrode, which is surrounded by a second outer barrel-like, half splitted electrode [133]. An electrostatic quadro-logarithmic potential field is created between these two electrodes [134]. Ions are tangentially injected into the electric field with a spatially coherent motion, so no further excitation is needed for their detection, as in the ICR cell. The electrostatic interaction between ions and the inner electrode is balanced by centrifugal forces resulting in an elliptical orbit rotation. In addition, trapped ions of a given m/z move in oscillating rings along the axis of the inner electrode. The axial oscillating frequency ν_z (Hz) is independent of ion energy and amplitude. The frequency depends on the field curvature k and is inversely proportional to the square root of m/z :

$$\nu_z = \frac{\left(e \times z \times \frac{k}{m} \right)^2}{2 \times \pi}$$

Equation 5

Signal detection is achieved by means similar to an ICR cell, by measuring the image current. The outer electrode is equipped with a differential amplifier to measure the axial oscillation which is acquired as a time-domain transient that is converted by FT into a frequency-domain signal. According to Equation 5, ν_z is correlated to m/z of a specific ion. The image current can be amplified for the orbitrap in the same way as for FTICR resulting in comparable sensitivity and signal-

to-noise ratio (S/N). The state-of-the-art orbitrap instruments can achieve mass resolution up to approximately 240,000 [135] and a mass accuracy <1 ppm for internal and <3 ppm for external calibration. Compared with FTICR instrument, they provide larger trapping volumes and increased space charge capacity for higher masses. The orbitrap provides a higher mass resolution for ions with a m/z above 800, as compared with a 7 Tesla FTICR [136,137], due to the ions in the orbitrap having an ion cyclotron frequency that is proportional to $(e \times z/m)$ and an axial oscillating frequency that is proportional to $(e \times z/m)^{1/2}$, as shown in Equations 4 & 5. However, FTICR instruments provide higher mass resolving power for ions at lower m/z .

» Fragmentation

Since information about intact protein mass as well as sequence information is obtained from TD-MS, a suitable instrument must be able to fragment intact proteins inside the MS and to determine the m/z -values of the fragment ions. Current TD-MS fragmentation techniques can be divided into CID and electron triggered dissociation methods.

CID is based on the collision of ions with an inert reagent gas (e.g., He, Ar) within a collision cell. During CID, kinetic energy of the ions and collision molecules is converted into internal energy resulting in vibrational excitation, where the energy is redistributed over the entire ion with an equal probability before dissociation occurs [138,139]. The redistribution of the energy across the proteins results in non-conservative cleavage sites including sequence specific hot-spots for fragmentation during CID. CID mainly produces b- and y-ions (Figure 2) and can be used in combination with nearly every type of mass spectrometer (QTOF-, FTICR- and orbitrap instruments). Orbitraps provide a special form of collision-based fragmentation termed higher-energy collisional dissociation (HCD, formerly high-energy C-trap dissociation). HCD in comparison with CID is based on higher activation energy and shorter activation time. The HCD method produces mainly b- and y-ions which can be further fragmented to a-ions or even smaller fragments during HCD [140].

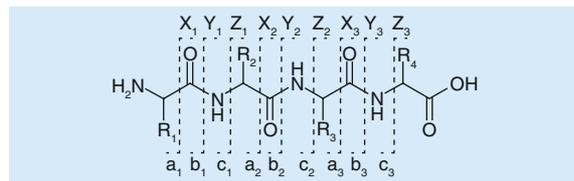


Figure 2. Fragments that occur during low energy dissociation of peptide and proteins according to the nomenclature of Roepstorff and Fohlman [217].

In contrast to CID and HCD, electron-triggered fragmentation methods are based on electron capture (ECD) [77] or the transfer of an electron from a donor molecule to the analyte ion (ETD) [141]. ECD produces a radical ion $[M+nH]^{n+\bullet}$. Neutralization of one elementary charge provides a free energy of 5–7 eV that is converted into internal energy and results in a spontaneous fragmentation without further energy redistribution over the whole protein. Therefore fragmentation is more conservative compared with CID with fewer structurally specific hot-spot sides undergoing fragmentation and provides a more widespread backbone dissociation yielding c- and z[•]-ions. At present, ECD is limited to ICR cells due to difficulties in storage of thermal electrons and cations in a quadrupole or an ion trap.

In order to utilize electron-induced fragmentation for TD-MS in other instruments other than FTICRs, ETD was developed and coupled to QTOF and hybrid instruments equipped with an ion trap (orbitrap). The principles of ETD are similar to those of ECD yielding mainly c- and z[•]-ions as well as the widespread fragmentation of the protein backbone. As the dissociation process in ECD and ETD is almost exclusively located to the protein backbone, side-chain modifications (e.g., glycosylation, phosphorylation) are virtually not affected. Additionally, labile structure elements of a protein such as phosphorylation sites and disulfide bonds that can undergo rearrangement reactions during CID are preserved in electron-triggered dissociation [142,143]. These features, together with the distinct backbone fragmentation, make ECD and ETD the method of choice for intact protein fragmentation.

One challenge in TD-MS is the relatively low signal intensity of fragment ions compared with fragment ions yielded by BU-approaches. A peptide with a MW between 1000 and 2000 Da produces a relatively small number of fragments. In comparison, dissociation of a protein will result in an extremely high number of fragments. Since the intensity of the precursor is distributed over the number of fragment ions, TD-MS suffers from a relatively low sensitivity in fragment spectra. This challenge is enhanced during ESI due to the protein concentration being distributed over a number of different charge states. To overcome this problem and to enhance sequence coverage in TD-MS, a number of charge states are isolated for fragmentation using an expanded isolation window [144–146] or by the use of ISD followed by an additional fragmentation step [147,148]. Both an expanded isolation window and ISD in combination with another fragmentation process results in a higher sequence coverage. However both methods are susceptible to impurities or complex samples with a number of different proteins or pro-

tein species. Another major challenge in TD-MS is that proteins larger than 500 residues are difficult to dissociate due to their gaseous tertiary structure [149]. To address this problem McLafferty and colleagues developed a technique termed prefolding dissociation using low energy collisions in the 1 Torr region in front of the skimmer to disrupt non-covalent interaction followed by high-energy collisions to fragment the backbone of the protein in the high-pressure region after the skimmer [149]. This technique has been successfully implemented by the group of McLafferty to dissociate proteins with a MW larger than 200 kDa [149]. Another well established technique developed to overcome this limitation is activated ion ECD. Using ECD, proteins larger than 20 kDa can lead to the formation of undissociated charge-reduced ions, where the protein is cleaved but fragments are held together via noncovalent hydrogen bonding or hydrophobic interactions [150]. To disrupt these noncovalent interactions and enhance the efficiency of ECD, slow heating activations of protein ions are used to destroy the non-covalent interactions. This activation can be achieved by collisions with background gases [151–153], IR irradiation [152,154,155] or increased temperature of the ICR cell [153,154]. Thus activated ion ECD generates more *c*- and *z*[•]-ions and increase significantly the sequence coverage [151,155–158]. Ledvina, Coon and colleagues implemented activated ion-ETD to a quadrupole linear ion trap [159] and recently in a modified collision cell on a hybrid LTQ-orbitrap [160].

Bruker developed a special fragmentation technique called T³-sequencing for TD-MS using a MALDI-TOF/TOF mass spectrometer [161]. This technique is a pseudo MS³. First, ions are dissociated by MALDI-ISD [162] and the resulting fragments are selected for a further MS/MS analysis. In MALDI-ISD mainly *c*-, *y*- and (*z*+2)-ions are generated [163]. The fragmentation process is still not completely understood but it has been shown that the MALDI matrix plays an important role in the formation of ISD ions. Due to excitation by photon absorption, hydrogen radicals are generated photochemically, originating from the matrix. An intermolecular transfer to the peptide backbone then takes place resulting in the formation of peptide radicals and thus generating ISD ions [164–166]. Using the T³-sequencing approach to analyze intact proteins, MALDI-ISD produces *c*- and *y*-ions of both the C- and N-termini of the protein. These ISD-ions are further analyzed by MS/MS yielding *b*- and *y*-ions, that confirm the N- and C-terminal sequence of the protein.

» Data analysis software

ESI, commonly used for TD-MS, produces multiple charged ions of intact proteins resulting in complex

spectra, especially if there are different proteins or protein species with overlapping isotopic envelopes. A major challenge is the annotation of the monoisotopic peak, which provides information concerning the correct mass of the corresponding molecule. For large proteins the monoisotopic peak is often not experimentally observed because of its relative low intensity and S/N. In order to avoid tedious manual spectra interpretation, deconvolution algorithms were developed to reduce spectra complexity and to determine the monoisotopic mass of a protein [167–172]. In 2000 Horn *et al.* developed the first fully automated algorithm for charge state determination, deconvolution and peak picking on both precursor ion and product ion level [172]. This algorithm is still today a benchmark for data processing in TD-MS [146] and implemented in a numerous data analysis and identification tools, including ProSight [173].

Next to raw data processing algorithms, three different search engines for TD-MS are currently available (MascotTB [174], PIITA [175] and ProSight [173]). The standard Mascot [176] search engine is limited to a precursor MW of 16 kDa. MascotTB was developed for intact protein identification supporting a precursor MW up to 110 kDa. MascotTB enables identification of different protein species including PTMs as well as amino acid sequence alterations. Unfortunately the supported mass range is still beneath the MW of intact antibodies which represent the major form of industrially produced biologics today.

PIITA (precursor ion independent TD algorithm) uses fragment ion information to identify a protein from a predicted gene. Following a gene match, the experimental precursor mass is compared with the MW information provided by the gene. Differences between observed and predicted masses are used to detect possible modifications or sequence alterations.

The third search engine, ProSight, developed by Meng *et al.* in 2001 [173] was the first search engine for TD-MS and is still the most flexible one today. ProSight uses the THRASH algorithm for raw data processing and searches with both precursor ion and product ion information against a database [177]. With ProSight it is possible to identify expected and unexpected PTMs as well as sequence alterations and truncations of the measured protein.

During the last years, MS companies have developed vendor-specific software solutions for TD-MS tailored for biopharmaceuticals. For example, Bruker Daltonics offers BioPharmaCompass, a fully automated solution for characterization and quality control of pharmaceutical proteins including a data analysis software package [178,179]. Bruker is applying a maximum entropy algorithm [180] to determine

charge and deconvolute spectra in combination with a custom algorithm, that accurately determines and extracts monoisotopic mass information from high-resolution MS [144].

Thermo has developed Protein Deconvolution 2.0 [181], a raw data processing tool that includes two algorithms for deconvolution and peak picking. Xtract [135,182] is used to manipulate isotopically resolved proteins and ReSpect [181,183] has been designed for the manipulation of isotopically unresolved proteins. Additionally, Thermo implemented ProSight (ProSightPC 2.0 Software) [135,182] into their data analysis pipeline for protein identification and included THRASH as a third raw data processing algorithm.

» Analysis of therapeutic proteins smaller than 100 kDa by TD-MS

Suckau and Resemann showed that MALDI-TD-MS using T³-sequencing can be used to characterize both the N- and C-termini of proteins covering a range from 6 to 67 kDa yielding up to 70 N-terminal and up to 50 C-terminal residues with an analysis time less than 1 h/protein [184,185]. Furthermore they were able to sequence an entire camelid single HC antibody (124 residues, 13.6 kDa) *de novo* using the T³-approach [186]. The success of MALDI-TD-MS is highly dependent on sample quality. If the protein species of interest was purified and effectively desorbed and ionized under MALDI conditions, a protein amount between 10 and 50 pmol per sample preparation should be sufficient to characterize the N- and C-termini, yielding identification and localization of possible modification sites. This approach is well suited for rapid routine quality checks during both the development and production process of therapeutic proteins, and is capable of yielding results in less than 1 h and in some cases within minutes.

Using TD-MS in combination with ESI, Moreno-Gordaliza *et al.* investigated the interaction between the antitumor drug cisplatin and insulin-identifying Pt binding sites [187]. Analysis was performed on a LTQ XL linear ion trap in zoom-scan mode in which the instruments resolving power was sufficient to resolve insulin molecules. TD-MS in combination with FTICRs and orbitraps provide a reliable method for the routine analysis of proteins with a MW less than 25 kDa [124,186,188–191] and has the potential to be routinely applied to proteins up to 60 kDa in the near future [177,192]. This method will provide a fast and effective method for analyzing small therapeutic proteins such as insulin or erythropoietin.

» Analysis of mAbs by TD-MS

In recent years TD-MS was used to analyze mAbs. Zhang *et al.* from Amgen reported the implementa-

tion of TD-MS towards characterizing the variable regions of mAbs. In 2007 Zhang *et al.* used a LTQ-orbitrap to characterize the variable region of IgG1 and IgG2 antibodies [147]. They used ISD to fragment all charge states of a protein simultaneously which greatly improved the sensitivity of the fragment spectra. Furthermore they reported an enhancement of fragmentation efficiency due to protein dissociation before gaseous tertiary structure formation took place. ISD of mAbs generated a series of b- and y-ions corresponding to the LD and HC, with an abundance of a series of b-ions corresponding to the entire variable region of both chains. For further characterization of the variable regions, these b-ions were isolated and fragmented by using CID. These studies yielded a large number of product ions (~15% sequence coverage), and clearly identified the location of oxidized methionines. Bondarenko *et al.* used the combination of ISD and CID with a LTQ-orbitrap to characterize disulfide species of an IgG2 antibody that had been separated by online reversed-phase chromatography [148]. Using ISD, 12 kDa fragments containing the N-Terminus were generated and further analyzed by CID allowing the differentiation between N-terminal glutamine and pGlu variants of the HC. Bondarenko, Ren and colleagues also successfully applied a LC-ESI-QTOF (LCT Premier, Micromass, Waters) to TD N-terminal sequencing of IgG LD and HCs [193]. They employed an ion guide voltage of 100 V in order to achieve a CID-like fragmentation. This resulted in the production of b-ion series which permitted N-terminal sequencing of the first seven residues. A group from Amgen highlighted the use of TD-MS by performing a rapid characterization of mAbs variable regions. Compared with classical BU-approaches, sample preparation is minimized thus reducing artificial modifications. Furthermore the required time for analysis is significantly decreased in comparison to BU. Tsybin and colleagues investigated the use of ETD for antibody characterization using both a state of the art QTOF [144] and a LTQ-orbitrap velo [145]. Using the QTOF and a m/z-isolation window of 100 Da, covering three charge states (49⁺, 50⁺, 51⁺), a sequence coverage of 21% was achieved. An orbitrap was used to achieve an overall sequence coverage of 32.7% by combining the fragmentation results of a narrow and a wide isolation window. The narrow isolation window included charge state 53⁺, 54⁺ and 55⁺ and generated 162 unique backbone cleavages covering 214 residues and 451 residues of the LD and the HCs, respectively. This corresponded to a sequence-coverage of approximately 24%. Although most of the backbone cleavages occurred in the disulfide-bond-free region, in total 32 cleavage sites were

assigned in the disulfide-protected region of the antibody. Using a wide isolation window (600 m/z units, covering charge states 47⁺–57⁺) 180 unique cleavage sites were generated yielding a sequence coverage of approximately 27%. By combining the results of the narrow and wide isolation window the total number of unique cleavage sites was increased to 217. This corresponds to a sequence-coverage of 32.7%. One of the main limitations in LC-based TD-MS is the insufficient S/N. This is due to rapid protein elution and results in a limited number of acquired scans available to be averaged. To overcome this problem and enhance S/N for product ions, Tsybin and colleagues summed up the orbitrap MS/MS transients coming from a number of distinct LC-MS/MS-runs before performing time-to-frequency conversion by FT [145]. Recently the Marshall group investigated the performance of ECD for analyzing intact antibodies using a 9.4 Tesla FTICR [146]. By isolating a single charge state, 162 unique cleavage sites were generated (~25% sequence coverage). Using a simultaneous fragmentation of all charge states yielded 228 unique cleavage sites pushing the sequence coverage forward to 34.4% compared with the work of Tsybin and colleagues [145]. These results demonstrate the superior fragmentation efficiency of ECD and ETD, as compared with CID, which results in higher sequence coverage in the application of TD-MS to therapeutic antibodies.

Furthermore the group of Gross recently (2013) used a combination of IM-MS and ECD-FTICR for analyzing differences in antibody conformation based on disulfide alterations [143]. From IM-MS, information about conformational changes, based on different gas phase collision cross-sections, were derived, showing different isomers of the wild type antibody. From the ECD-FTICR-MS more site-specific information concerning the conformational changes were derived.

TD-MS is an attractive tool to quantify the relative abundances of the diverse therapeutic protein species along with the main species. The influence of groups modifying the physicochemical properties of protein species originated by the same gene are significantly lower as compared with peptides [194–196]. With the protein-ion-ratio method developed by the Kelleher group, the relative amount of the diverse protein species can be calculated [196]. Together with information provided by BU-approaches, TD-MS can be used to quickly identify and quantify modified species of therapeutic proteins during both the development process and the industrial processing.

Further improvements to TD-MS for the analysis of therapeutic proteins was developed by Marshall and

Nikolaev. The Marshall group achieved a unit mass baseline resolution for an intact therapeutic mAb with a MW of 148 kDa by using a 9.4 Tesla FTICR, which utilized a ICR-transient time of 20 s and a mass resolving power of approximately 420,000 [197]. Nikolaev was able to resolve the isotopic patterns from a number of large proteins using a 7 Tesla FTICR equipped with his ICR cell [198]. The 66 kDa protein, bovine serum albumin, was isotopically resolved with a mass resolving power of 1,500,000 (ICR-transient: 24 s) as well as the 93 kDa yeast enolase dimer ($R = 800,000$; ICR-transient: 40 s). Furthermore Nikolaev was able to resolve the isotopic patterns of both an alcohol dehydrogenase tetramer (147 kDa; $R = 500,000$; ICR-transient: 30 s) and IgG1 (147 kDa, $R = 500,000$; ICR-transient: 28 s). These results, together with the high resolving power achieved for large proteins up to 148 kDa, have the potential to revolutionize the use of TD-MS for analyzing large therapeutic proteins such as mAbs. Up to now, due to the large unresolved isotopic envelope of mAbs (~25 Da), it has not been possible to detect modifications smaller than 20 Da such as oxidation (+15.9949 Da), pGlu formation from glutamine (-17.0265 Da) or deamidation (+0.9840 Da) on intact protein level. Recent developments suggest that these measurements could be possible in the near future.

» Analysis of large therapeutic proteins by MD-MS

Characterization of mAbs by TD-MS is almost entirely limited to the variable regions. To obtain further information about the primary structure and to enhance the sequence coverage, mAbs can be analyzed by a MD approach. In this approach mAbs are cleaved to a few larger fragments by a limited LysC-digest, the reduction of disulfide-bonds or by using both which yields three fragments of about 25 kDa, namely Fab HC, single chain Fc and LD [73,148,199]. The resulting sizes of the fragments, being much smaller than that of the intact therapeutic protein species, are well suited for LC-MS/MS-analysis using state-of-the-art QTOF, orbitrap or FTICR mass spectrometers given that the resolving power of these instruments is high enough to distinguish the isotopic patterns of both precursor and fragment ions. The reduction of disulfide bonds leads to a higher sequence coverage compared with TD-MS-analysis of intact mAbs because cleavage sites are obtained during fragmentation that were otherwise protected by the disulfide bond region [148,149,200]. Compared to TD-MS more sample preparation steps are necessary but it requires less time and protein than BU-approaches.

Conclusion

TD-MS is an attractive tool for analyzing and monitoring therapeutic proteins during their development and industrial processing, especially if data about the chemical composition of a defined therapeutic protein species and its accompanying species have already been collected by BU-approaches. For example, it would be advantageous to have a rapid method to characterize the variable regions of mAbs, because these regions determine their activity and even a small modification can have an impact on it. As described above, TD-MS is capable of sequencing the variable regions of mAbs thus identifying and localizing possible modification sites. This quality control measurement can be made in less than one hour utilizing minimal sample preparation steps minimizing artificial modifications. Even by using multiple LC-MS/MS runs as described above [145] the analysis results are obtained within a few hours, which is still less time than normal BU procedures require [37].

TD-MS of large proteins such as mAbs still presents a challenge. It is difficult for the most advanced instruments to resolve the isotopic pattern of the intact protein precursor ion and to obtain sufficient signal intensities on the fragment ion level. Here the MD approach presents a powerful alternative presently implemented in many biopharmaceutical laboratories. Nevertheless, recent and ongoing developments in instrumentation, fragmentation methods and proce-

dures as well as algorithms for data processing make TD-MS of large proteins increasingly attractive for the analysis of therapeutic proteins.

Future perspective

In the past few years there have been impressive improvements according MS technologies. This progress has enabled TD-MS of therapeutic protein species. However, today TD-MS is still difficult regarding larger proteins like antibodies. In 5–10 years limitations of TD-MS presumably will be overcome because it can be assumed that the progress in the development of mass spectrometers will continue regarding increases in resolution and mass accuracy, improved orthogonal fragmentation techniques yielding complementary fragment ion series and because of progress in the field of data processing and interpretation of MS data; thus, making time-consuming BU approaches needless, at least for monitoring up- and down-stream-processing.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Executive summary

Background

- » During manufacturing of therapeutic proteins not only the desired therapeutic protein species but many other closely related species are generated which differ more or less in their chemical composition.
- » Minor changes in the chemical composition of therapeutic protein species can have a great impact on the efficiency, stability and immunogeneity.
- » Thus a comprehensive qualitative and quantitative analysis of the therapeutic protein species is necessary.

Relevant techniques for the characterization of therapeutic proteins

- » Today many different methods for the analysis are allowing the determination of the exact chemical composition of closely related species. Many of these methods are based on the MS analysis of proteolytic peptides obtained from therapeutic protein species by incubation with proteases such as trypsin. A combination of this bottom-up approach usually yields all necessary information.

Analysis of therapeutic proteins by top-down-MS

- » Applying a combination of bottom-up approaches is time consuming. Here top-down MS (TD-MS) – the MS analysis of intact therapeutic protein species – is an alternative which requires significantly less time per sample but offers information about important features of the chemical composition of a species.

Future perspective

- » In the past few years there have been impressive improvements in MS technologies. This progress has enabled TD-MS analysis of therapeutic protein species. However, today TD-MS still encounters difficulties with regards to larger proteins such as antibodies. It is expected that in 5–10 years, limitations of TD-MS will be presumably overcome because it can be assumed that the progress in the development of MS will continue; regarding increases in resolution and mass accuracy, improved orthogonal fragmentation techniques yielding complementary fragment ion-series and because of progress in the field of data processing and interpretation of MS data; thus making time-consuming bottom-up approaches needless, at least for monitoring upstream- and downstream-processing.

References

Papers of special note have been highlighted as:

■ of interest

■ ■ of considerable interest

- 1 European Medicines Agency. Questions and answers on biosimilar medicines (similar biological medicinal products). European Medicines Agency, London, UK (2012).
- 2 Richmann DA. FDA Basics For Biotech Drugs, Biologics and Devices. King and Spalding, GA, USA (2001).
- 3 Amgen. Biologics and biosimilars: an overview. Amgen Inc., CA, USA (2012).
- 4 APA. The biosimilar pathway: where will it lead us? *Pharm. Today* 17, 67–76 (2011).
- 5 Farrugia A, Robert P. Plasma protein therapies: current and future perspectives. *Best Pract. Res. Clin. Haematol.* 19(1), 243–258 (2006).
- 6 Smith LM, Kelleher NL. Proteoform: a single term describing protein complexity. *Nat. Methods* 10(3), 186–187 (2013).
- 7 Dimitrov DS. Therapeutic proteins. *Methods Mol. Biol.* 899, 1–26 (2012).
- 8 Jungblut PR, Holzhtuter HG, Apweiler R, Schlüter H. The speciation of the proteome. *Chem. Cent. J.* 2, 16 (2008).
- 9 Schlüter H, Apweiler R, Holzhtuter HG, Jungblut PR. Finding one's way in proteomics: a protein species nomenclature. *Chem. Cent. J.* 3, 11 (2009).
- 10 Bischoff R, Schlüter H. Amino acids: chemistry, functionality and selected non-enzymatic post-translational modifications. *J. Proteomics* 75(8), 2275–2296 (2012).
- 11 Descotes J, Gouraud A. Clinical immunotoxicity of therapeutic proteins. *Expert Opin. Drug Metab. Toxicol.* 4(12), 1537–1549 (2008).
- 12 Chung Ch, Mirakhor B, Chan E *et al.* Cetuximab-induced anaphylaxis and IgE specific for galactose- α -1,3-galactose. *N. Engl. J. Med.* 358(11), 1109–1117 (2008).
- 13 Willow D. The regulation of biologic medicine: innovators' rights and access to healthcare. *Chi.-Kent J. Intell. Prop.* 6, 32–42 (2006).
- 14 Vermasvuori R. Production of recombinant proteins and monoclonal antibodies – techno-economical evaluation of the production methods. Thesis. Helsinki university of technology, Helsinki, Finland (2009).
- 15 Amgen. Biologics and biosimilars, a new complexity in the world of medicines. Amgen Inc., CA, USA (2012).
- 16 Calo-Fern'andez B. M1-Hjl: biosimilars: company strategies to capture value from the biologics market. *Pharmaceuticals* 5, 1393–1408 (2012).
- 17 Accenture. Biosimilars: emergence of a third market dynamic between original products and generics. Current situation, strategic options and recommendations. Accenture, Dublin, Ireland (2009).
- 18 Berkowitz SA, Engen JR, Mazzeo JR, Jones GB. Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars. *Nat. Rev. Drug Discov.* 11(7), 527–540 (2012).
- **Besides analytical tools, regulatory aspects are well elucidated.**
- 19 Nandania J, Akalkotkar A. Analysis of biopharmaceuticals and mass spectroscopy. *Int. J. Res. Pharm. Sci.* 2(4), 571–578 (2011).
- 20 Tam JWY. Biologics revolution: the intersection of biotechnology, patent law, and pharmaceutical regulation. *Georgetown Law J.* 98, 535–565 (2010).
- 21 Jensen ON. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr. Opin. Chem. Biol.* 8(1), 33–41 (2004).
- 22 Jenkins N. Modifications of therapeutic proteins: challenges and prospects. *Cytotechnology* 53(1–3), 121–125 (2007).
- 23 Manning MC, Chou DK, Murphy BM, Payne RW, Katayama DS. Stability of protein pharmaceuticals: an update. *Pharm. Res.* 27(4), 544–575 (2010).
- 24 Weise M, Bielsky MC, De Smet K *et al.* Biosimilars – why terminology matters. *Nat. Biotechnol.* 29(8), 690–693 (2011).
- 25 Dick LW Jr, Qiu D, Mahon D, Adamo M, Cheng KC. C-terminal lysine variants in fully human monoclonal antibodies: investigation of test methods and possible causes. *Biotechnol. Bioeng.* 100(6), 1132–1143 (2008).
- 26 Harris RJ. Processing of C-terminal lysine and arginine residues of proteins isolated from mammalian cell culture. *J. Chromatogr. A* 705(1), 129–134 (1995).
- 27 US FDA. Center for Biologics Evaluation and Research: strategic plan for regulatory science and research fy 2012–2016. US FDA, MD, USA (2012).
- 28 European Medicines Agency. Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: quality issues (revision 1). European Medicines Agency, London, UK (2012).
- 29 ICHQ6B. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use – Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products (1999).
- 30 Zölls S. Reference standards for therapeutic proteins: their relevance, development, qualification and replacement. Presented at: *The International Association for Biologicals*. MD, USA, 20–21 September 2011.
- 31 Mazur MT, Seipert RS, Mahon D, Zhou Q, Liu T. A platform for characterizing therapeutic monoclonal antibody breakdown products by 2D chromatography and top-down mass spectrometry. *AAPS J.* 14(3), 530–541 (2012).
- 32 Arakawa T, Philo JS, Ejima D, Tsumoto K, Arisaka F. Aggregation analysis of therapeutic proteins, Part 1. *BioProcess Int.* 4(10), 32–42 (2006).
- 33 Ma S, Nashabeh W. Analysis of protein therapeutics by capillary electrophoresis. *Dev. Biol.* 122, 49–68 (2001).
- 34 Hunt G, Nashabeh W. Capillary electrophoresis sodium dodecyl sulfate nongel sieving analysis of a therapeutic

- recombinant monoclonal antibody: a biotechnology perspective. *Anal. Chem.* 71(13), 2390–2397 (1999).
- 35 Lee HG. High-performance sodium dodecyl sulfate-capillary gel electrophoresis of antibodies and antibody fragments. *J. Immunol. Methods* 234(1–2), 71–81 (2000).
- 36 Srebalus Barnes CA, Lim A. Applications of mass spectrometry for the structural characterization of recombinant protein pharmaceuticals. *Mass Spectrom. Rev.* 26(3), 370–388 (2007).
- 37 Beck A, Wagner-Rousset E, Ayoub D, Van Dorsselaer A, Sanglier-Cianferani S. Characterization of therapeutic antibodies and related products. *Anal. Chem.* 85(2), 715–736 (2013).
- Nicely structured review giving very good insights about current state-of-the-art analytical methods used to characterize monoclonal antibodies (mAbs).
- 38 Edman P, Begg G. A protein sequenator. *Eur. J. Biochem.* 1(1), 80–91 (1967).
- 39 Mortz E, Nguyen TH, Crawford J. N- and C-terminal protein sequencing using in-source-decay (ISD) MALDI MS. Alphalyse Inc., CA, USA (2009).
- 40 Griffiths SW, Cooney CL. Development of a peptide mapping procedure to identify and quantify methionine oxidation in recombinant human alpha1-antitrypsin. *J. Chromatogr. A* 942(1–2), 133–143 (2002).
- 41 Chumsae C, Gaza-Bulseco G, Sun J, Liu H. Comparison of methionine oxidation in thermal stability and chemically stressed samples of a fully human monoclonal antibody. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 850(1–2), 285–294 (2007).
- 42 Kroon DJ, Baldwin-Ferro A, Lalan P. Identification of sites of degradation in a therapeutic monoclonal antibody by peptide mapping. *Pharm. Res.* 9(11), 1386–1393 (1992).
- 43 Yang J, Wang S, Liu J, Raghani A. Determination of tryptophan oxidation of monoclonal antibody by reversed phase high performance liquid chromatography. *J. Chromatogr. A* 1156(1–2), 174–182 (2007).
- 44 Wang L, Amphlett G, Lambert JM, Blattler W, Zhang W. Structural characterization of a recombinant monoclonal antibody by electrospray time-of-flight mass spectrometry. *Pharm. Res.* 22(8), 1338–1349 (2005).
- 45 Liu H, Gaza-Bulseco G, Sun J. Characterization of the stability of a fully human monoclonal IgG after prolonged incubation at elevated temperature. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 837(1–2), 35–43 (2006).
- 46 Harris RJ, Kabakoff B, Macchi Fd *et al.* Identification of multiple sources of charge heterogeneity in a recombinant antibody. *J. Chromatogr. B Biomed. Sci. Appl.* 752(2), 233–245 (2001).
- 47 Lewis DA, Guzzetta AW, Hancock WS, Costello M. Characterization of humanized anti-TAC, an antibody directed against the interleukin 2 receptor, using electrospray ionization mass spectrometry by direct infusion, LC/MS, and MS/MS. *Anal. Chem.* 66(5), 585–595 (1994).
- 48 Dick Lw Jr, Kim C, Qiu D, Cheng KC. Determination of the origin of the N-terminal pyro-glutamate variation in monoclonal antibodies using model peptides. *Biotechnol. Bioeng.* 97(3), 544–553 (2007).
- 49 Wagner-Rousset E, Bednarczyk A, Bussat MC *et al.* The way forward, enhanced characterization of therapeutic antibody glycosylation: comparison of three level mass spectrometry-based strategies. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 872(1–2), 23–37 (2008).
- 50 Hirayama K, Yuji R, Yamada N, Kato K, Arata Y, Shimada I. Complete and rapid peptide and glycopeptide mapping of mouse monoclonal antibody by LC/MS/MS using ion trap mass spectrometry. *Anal. Chem.* 70(13), 2718–2725 (1998).
- 51 Hermanson GT. *Bioconjugate Techniques, Second Edition.* Academic Press, London, UK (2008).
- 52 Geiger T, Clarke S. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *J. Biol. Chem.* 262(2), 785–794 (1987).
- 53 Walsh G, Jefferis R. Post-translational modifications in the context of therapeutic proteins. *Nat. Biotechnol.* 24(10), 1241–1252 (2006).
- 54 Mann M, Jensen ON. Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* 21(3), 255–261 (2003).
- 55 Zhang T, Bourret J, Cano T. Isolation and characterization of therapeutic antibody charge variants using cation exchange displacement chromatography. *J. Chromatogr. A* 1218(31), 5079–5086 (2011).
- 56 Tan M, Chen Y, Xie Z, Lu Z, Zhao Y. Proteome-wide screening of lysine succinylation in *Escherichia coli* reveals its broad roles in cellular metabolism. Presented at: *61st ASMS Conference on Mass Spectrometry and Allied Topics.* Minneapolis, MN, USA, 9–13 June 2013.
- 57 Chen Y, Peng C, Tan M, Dai L, Xie Z, Zhao Y. The first mammalian succinylome analysis reveals Sirt5-mediated lysine desuccinylation and its roles in diverse cellular pathways. Presented at: *61st ASMS Conference on Mass Spectrometry and Allied Topics.* Minneapolis, MN, USA, 9–13 June 2013.
- 58 Brooks SA. Strategies for analysis of the glycosylation of proteins: current status and future perspectives. *Mol. Biotechnol.* 43(1), 76–88 (2009).
- 59 Hara S, Yamaguchi M, Takemori Y, Nakamura M, Ohkura Y. Highly sensitive determination of *N*-acetyl- and *N*-glycolylneuraminic acids in human serum and urine and rat serum by reversed-phase liquid chromatography with fluorescence detection. *J. Chromatogr.* 377, 111–119 (1986).
- 60 Hara S, Yamaguchi M, Takemori Y, Furuhashi K, Ogura H, Nakamura M. Determination of mono-*O*-acetylated *N*-acetylneuraminic acids in human and rat sera by fluorometric high-performance liquid chromatography. *Anal. Biochem.* 179(1), 162–166 (1989).
- 61 Zauner G, Kozak RP, Gardner RA, Fernandes DL, Deelder AM, Wuhler M. Protein *O*-glycosylation analysis. *Biol. Chem.* 393(8), 687–708 (2012).
- 62 Rohrer JS. Analyzing sialic acids using high-performance anion-exchange chromatography with pulsed amperometric detection. *Anal. Biochem.* 283(1), 3–9 (2000).
- 63 Varki A, Cummings R, Esko J, Freeze H, Hart G, Marth J. Principles of Structural Analysis and Sequencing of Glycans.

- In: *Essentials of Glycobiology First Edition*. Cold Spring Harbor Laboratory Press, NY, USA (1999).
- 64 Zauner G, Deelder AM, Wuhrer M. Recent advances in hydrophilic interaction liquid chromatography (HILIC) for structural glycomics. *Electrophoresis* 32(24), 3456–3466 (2011).
- 65 El Rassi Z. Recent progress in reversed-phase and hydrophobic interaction chromatography of carbohydrate species. *J. Chromatogr. A* 720(1–2), 93–118 (1996).
- 66 Zaia J. Mass spectrometry of oligosaccharides. *Mass Spectrom. Rev.* 23(3), 161–227 (2004).
- 67 Morelle W, Canis K, Chirat F, Faid V, Michalski JC. The use of mass spectrometry for the proteomic analysis of glycosylation. *Proteomics* 6(14), 3993–4015 (2006).
- 68 Huhn C, Selman MH, Ruhaak LR, Deelder AM, Wuhrer M. IgG glycosylation analysis. *Proteomics* 9(4), 882–913 (2009).
- 69 Dalpathado DS, Desaire H. Glycopeptide analysis by mass spectrometry. *Analyst* 133(6), 731–738 (2008).
- 70 Morelle W, Michalski JC. Analysis of protein glycosylation by mass spectrometry. *Nat. Protoc.* 2(7), 1585–1602 (2007).
- 71 Kolarich D, Jensen PH, Altmann F, Packer NH. Determination of site-specific glycan heterogeneity on glycoproteins. *Nat. Protoc.* 7(7), 1285–1298 (2012).
- 72 Zauner G, Selman MH, Bondt A *et al.* Glycoproteomic analysis of antibodies. *Mol. Cell Proteomics* 12(4), 856–865 (2013).
- 73 Zhang Z, Pan H, Chen X. Mass spectrometry for structural characterization of therapeutic antibodies. *Mass Spectrom. Rev.* 28(1), 147–176 (2009).
- 74 Bloom JW, Madanat MS, Marriott D, Wong T, Chan SY. Intrachain disulfide bond in the core hinge region of human IgG4. *Protein Sci.* 6(2), 407–415 (1997).
- 75 Akashi S, Kato K, Torizawa T *et al.* Structural characterization of mouse monoclonal antibody 13–11 against a porphyrin derivative: identification of a disulfide bond in CDR-H3 of Mab 13–11. *Biochem. Biophys. Res. Commun.* 240(3), 566–572 (1997).
- 76 Zhang W, Marzilli LA, Rouse JC, Czupryn MJ. Complete disulfide bond assignment of a recombinant immunoglobulin G4 monoclonal antibody. *Anal. Biochem.* 311(1), 1–9 (2002).
- 77 Zubarev RA, Kelleher NL, McLafferty FW. Electron capture dissociation of multiply charged protein cations: a nonergodic process. *J. Am. Chem. Soc.* 120, 3265–3266 (1998).
- 78 Wu SL, Jiang H, Lu Q, Dai S, Hancock WS, Karger BL. Mass spectrometric determination of disulfide linkages in recombinant therapeutic proteins using online LC–MS with electron-transfer dissociation. *Anal. Chem.* 81(1), 112–122 (2009).
- 79 Kaltashov IA, Bobst CE, Abzalimov RR, Wang G, Baykal B, Wang S. Advances and challenges in analytical characterization of biotechnology products: mass spectrometry-based approaches to study properties and behavior of protein therapeutics. *Biotechnol. Adv.* 30(1), 210–222 (2012).
- 80 Joshi SB, Bhambhani A, Zeng Y, Middaugh CR. An empirical phase diagram–high-throughput screening approach to the characterization and formulation of biopharmaceuticals. In: *Formulation and Process Development Strategies for Manufacturing Biopharmaceuticals*, Jameel F, Hershenson S (Eds.). John Wiley and Sons, Inc., NJ, USA, 173–205 (2010).
- 81 Amezcua CA, Szabo CM. Assessment of higher order structure comparability in therapeutic proteins using nuclear magnetic resonance spectroscopy. *J. Pharm. Sci.* 102(6), 1724–1733 (2013).
- 82 Aubin Y, Gingras G, Sauve S. Assessment of the three-dimensional structure of recombinant protein therapeutics by NMR fingerprinting: demonstration on recombinant human granulocyte macrophage-colony stimulation factor. *Anal. Chem.* 80(7), 2623–2627 (2008).
- 83 Chen G, Warrack BM, Goodenough AK, Wei H, Wang-Iverson DB, Tymiak AA. Characterization of protein therapeutics by mass spectrometry: recent developments and future directions. *Drug Discov. Today* 16(1–2), 58–64 (2011).
- 84 Konermann L, Pan J, Liu YH. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem. Soc. Rev.* 40(3), 1224–1234 (2011).
- 85 Tu T, Drăgușanu M, Petre BA, Rempel DI, Przybylski M, Gross ML. Protein-peptide affinity determination using an H/D exchange dilution strategy: application to antigen-antibody interactions. *J. Am. Soc. Mass Spectrom.* 21(10), 1660–1667 (2010).
- 86 Young MM, Tang N, Hempel JC *et al.* High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proc. Natl Acad. Sci. USA* 97(11), 5802–5806 (2000).
- 87 Pearson KM, Pannell LK, Fales HM. Intramolecular cross-linking experiments on cytochrome c and ribonuclease A using an isotope multiplet method. *Rapid Commun. Mass Spectrom.* 16(3), 149–159 (2002).
- 88 Leitner A, Walzthoeni T, Kahraman A *et al.* Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell Proteomics* 9(8), 1634–1649 (2010).
- 89 Sinz A. Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. *J. Mass Spectrom.* 38(12), 1225–1237 (2003).
- 90 Krayukhina E, Uchiyama S, Nojima K, Okada Y, Hamaguchi I, Fukui K. Aggregation analysis of pharmaceutical human immunoglobulin preparations using size-exclusion chromatography and analytical ultracentrifugation sedimentation velocity. *J. Biosci. Bioeng.* 115(1), 104–110 (2013).
- 91 Jenkins N, Murphy L, Tyther R. Post-translational modifications of recombinant proteins: significance for biopharmaceuticals. *Mol. Biotechnol.* 39(2), 113–118 (2008).
- 92 US FDA. Guidance for Industry Immunogenicity Assessment for Therapeutic Protein Products. US FDA, MD, USA (2013).
- 93 Svedberg T, Pedersen KO. *The Ultracentrifuge, First Edition*. Oxford University Press, Oxford, UK (1940).

- 94 Schachman HK. *Is there a future for the ultracentrifuge? First Edition*. Harding SE, Rowe AJ, Horton JC (Eds). Royal Society of Chemistry, Cambridge, UK (1992).
- 95 Berkowitz SA. Role of analytical ultracentrifugation in assessing the aggregation of protein biopharmaceuticals. *AAPS J.* 8(3), E590–E605 (2006).
- 96 Den Engelsman J, Garidel P, Smulders R *et al.* Strategies for the assessment of protein aggregates in pharmaceutical biotech product development. *Pharm. Res.* 28(4), 920–933 (2011).
- This publication elaborately discusses the analysis of aggregates with a special focus on the development of therapeutic proteins.
- 97 Arakawa T, Philo JS, Ejima D, Sato H, Tsumoto K. Aggregation analysis of therapeutic proteins, part 3. *BioProcess Int.* 5(10), 52–70 (2007).
- 98 Kuebler S. Characterizing stable protein formulations. Detecting and quantifying aggregation using dynamic light scattering. *Genet. Eng. News* 27(20) (2007).
- 99 Arakawa T, Philo JS, Ejima D, Tsumoto K, Arisaka F. Aggregation analysis of therapeutic proteins, part 2. *BioProcess Int.* 5(4), 36–47 (2007).
- 100 Felten C, Solano OS. Capillary electrophoresis in quality control: part II: CE-SDS: method development and robustness. Beckman Coulter, Inc., CA, USA (2011).
- 101 Kaltashov IA, Eyles SJ. Studies of biomolecular conformations and conformational dynamics by mass spectrometry. *Mass Spectrom. Rev.* 21(1), 37–71 (2002).
- 102 Smith DI, Zhang Z. Probing noncovalent structural features of proteins by mass spectrometry. *Mass Spectrom. Rev.* 13(5–6), 411–429 (1994).
- 103 Kukrer B, Filipe V, Van Duijn E *et al.* Mass spectrometric analysis of intact human monoclonal antibody aggregates fractionated by size-exclusion chromatography. *Pharm. Res.* 27(10), 2197–2204 (2010).
- 104 Tobler SA, Fernandez EJ. Structural features of interferon-gamma aggregation revealed by hydrogen exchange. *Protein Sci.* 11(6), 1340–1352 (2002).
- 105 Zhang Z, Smith DL. Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci.* 2(4), 522–531 (1993).
- 106 Novak P, Giannakopoulos AE. Chemical cross-linking and mass spectrometry as structure determination tools. *Eur. J. Mass Spectrom.* 13(2), 105–113 (2007).
- 107 Maleknia SD, Ralston CY, Brenowitz MD, Downard KM, Chance MR. Determination of macromolecular folding and structure by synchrotron x-ray radiolysis techniques. *Anal. Biochem.* 289(2), 103–115 (2001).
- 108 Kiselar JG, Maleknia SD, Sullivan M, Downard KM, Chance MR. Hydroxyl radical probe of protein surfaces using synchrotron x-ray radiolysis and mass spectrometry. *Int. J. Radiat. Biol.* 78(2), 101–114 (2002).
- 109 Guan JQ, Chance MR. Structural proteomics of macromolecular assemblies using oxidative footprinting and mass spectrometry. *Trends Biochem. Sci.* 30(10), 583–592 (2005).
- 110 Hale JE. Advantageous uses of mass spectrometry for the quantification of proteins. *Int. J. Proteomics* 2013, 219452 (2013).
- 111 Wang X, Schomogy T, Wells K, Mozier NM. Improved HCP quantitation by minimizing antibody cross-reactivity to target proteins *BioProcess Int.* 8(1), 18–24 (2010).
- 112 Doneanu CE, Xenopoulos A, Fadgen K *et al.* Analysis of host-cell proteins in biotherapeutic proteins by comprehensive online two-dimensional liquid chromatography/mass spectrometry. *mAbs* 4(1), 24–44 (2012).
- 113 Bell A, Rogers R, Kowski T, Robert B. One-dimensional liquid chromatography analysis of host cell proteins in therapeutic antibodies using an orbitrap velos. *61st ASMS Conference on Mass Spectrometry and Allied Topics*. MN, USA, 9–13 June 2013.
- 114 Potter O, Yang Y, Yin H, Killen K. Determination of contaminant bacterial host cell proteins in recombinant proteins expressed in *E. coli* by LC-QQQ. *61st ASMS Conference on Mass Spectrometry and Allied Topics*. MN, USA, 9–13 June 2013.
- 115 Cox DM, Zhong F, Du M, Duchoslav E, Sakuma T, McDermott JC. Multiple reaction monitoring as a method for identifying protein posttranslational modifications. *J. Biomol. Tech.* 16(2), 83–90 (2005).
- 116 Maiolica A, Junger MA, Ezkurdia I, Aebersold R. Targeted proteome investigation via selected reaction monitoring mass spectrometry. *J. Proteomics* 75(12), 3495–3513 (2012).
- 117 Johansen E, Jonakin K, Hunter C. Simultaneous quantitative peptide mapping and host cell protein detection in an IgG1 monoclonal antibody preparation using data-independent acquisition. *61st ASMS Conference on Mass Spectrometry and Allied Topics*. MN, USA, 9–13 June 2013.
- 118 Chait BT. Chemistry. Mass spectrometry: bottom-up or top-down? *Science* 314(5796), 65–66 (2006).
- 119 Cotter RJ. Time-of-flight mass spectrometry: an increasing role in the life sciences. *Biomed. Environ. Mass Spectrom.* 18(8), 513–532 (1989).
- 120 Guilhaus M. Principles and instrumentation in time-of-flight mass spectrometry. *J. Mass Spectrom.* 30(11), 1519–1532 (1995).
- 121 Hartmer RG, Kaplan DA, Stoermer C, Lubeck M, Park MA. Data-dependent electron transfer dissociation of large peptides and medium size proteins in a QTOF instrument on a liquid chromatography timescale. *Rapid Commun. Mass Spectrom.* 23(15), 2273–2282 (2009).
- 122 Marshall AG, Hendrickson CL, Jackson GS. Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass Spectrom. Rev.* 17(1), 1–35 (1998).
- 123 Marshall AG, Hendrickson CL. High-resolution mass spectrometers. *Annu. Rev. Anal. Chem.* 1, 579–599 (2008).
- 124 Schaub TM, Hendrickson CL, Horning S, Quinn JP, Senko MW, Marshall AG. High-performance mass spectrometry: Fourier transform ion cyclotron resonance at 14.5 Tesla. *Anal. Chem.* 80(11), 3985–3990 (2008).
- 125 Madeira PJA, Alves PA, Borges CM. High resolution mass spectrometry using FTICR and orbitrap instruments.

- In: *Fourier Transform – Materials Analysis*. Salih Salih (Ed.). InTech, NY, USA (2012).
- 126 Marshall AG, Verdun FR. *Fourier Transforms in NMR, Optical and Mass Spectrometry: A Users Handbook*. Elsevier, Amsterdam, The Netherlands (1990).
- 127 Asamoto B, Dunbar RC. *Analytical Applications of Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*. VCH, NY, USA (1991).
- 128 Savory JJ, Kaiser NK, McKenna AM *et al.* Parts-per-billion Fourier transform ion cyclotron resonance mass measurement accuracy with a 'walking' calibration equation. *Anal. Chem.* 83(5), 1732–1736 (2011).
- 129 Nikolaev EN, Boldin IA, Jertz R, Baykut G. Initial experimental characterization of a new ultra-high resolution FTICR cell with dynamic harmonization. *J. Am. Soc. Mass Spectrom.* 22(7), 1125–1133 (2011).
- 130 Boldin IA, Nikolaev EN. Harmonization of Electric Field in FT ICR Cell. The New Approaches. In: *Proceedings of the 58th ASMS Conference on Mass Spectrometry Allied Topics*. Salt Lake City, UT, USA, 23–27 May 2010.
- 131 Nikolaev EN, Jertz R, Grigoryev A, Baykut G. Fine structure in isotopic peak distributions measured using a dynamically harmonized Fourier transform ion cyclotron resonance cell at 7 T. *Anal. Chem.* 84(5), 2275–2283 (2012).
- 132 Solarix XR: A New Era in Mass Spectrometry (1816259). *Bruker Daltonic brochure*. Bruker Daltonics, MA, USA (2013).
- 133 Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Graham Cooks R. The Orbitrap: a new mass spectrometer. *J. Mass Spectrom.* 40(4), 430–443 (2005).
- 134 Makarov A. Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal. Chem.* 72(6), 1156–1162 (2000).
- 135 Michalski A, Damoc E, Lange O *et al.* Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. *Mol. Cell Proteomics* 11(3), O111.013698 (2012).
- 136 Scigelova M, Makarov A. Orbitrap mass analyzer – overview and applications in proteomics. *Proteomics* 6 (Suppl. 2), 16–21 (2006).
- 137 Perry RH, Cooks RG, Noll RJ. Orbitrap mass spectrometry: instrumentation, ion motion and applications. *Mass Spectrom. Rev.* 27(6), 661–699 (2008).
- 138 Sleno L, Volmer DA. Ion activation methods for tandem mass spectrometry. *J. Mass Spectrom.* 39(10), 1091–1112 (2004).
- 139 Wells JM, McLuckey SA. Collision-induced dissociation (CID) of peptides and proteins. *Methods Enzymol.* 402, 148–185 (2005).
- 140 Olsen JV, Macek B, Lange O, Makarov A, Horning S, Mann M. Higher-energy C-trap dissociation for peptide modification analysis. *Nat. Methods* 4(9), 709–712 (2007).
- 141 Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl Acad. Sci. USA* 101(26), 9528–9533 (2004).
- 142 Brgles M, Kurtovic T, Halassy B, Allmaier G, Marchetti-Deschmann M. Studying disulfide bond rearrangement by MALDI-RTOF PSD and MALDI-TOF/RTOF high-energy CID (20 keV) experiments of peptides derived from ammodytotoxins. *J. Mass Spectrom.* 46(2), 153–162 (2011).
- 143 Jones LM, Zhang H, Cui W *et al.* Complementary MS methods assist conformational characterization of antibodies with altered S-S bonding networks. *J. Am. Soc. Mass Spectrom.* 24(6), 835–845 (2013).
- 144 Tsybin YO, Fornelli L, Stoermer C *et al.* Structural analysis of intact monoclonal antibodies by electron transfer dissociation mass spectrometry. *Anal. Chem.* 83(23), 8919–8927 (2011).
- 145 Fornelli L, Damoc E, Thomas PM *et al.* Analysis of intact monoclonal antibody IgG1 by electron transfer dissociation Orbitrap FTMS. *Mol. Cell Proteomics* 11(12), 1758–1767 (2012).
- 146 Mao Y, Valeja SG, Rouse JC, Hendrickson CL, Marshall AG. Top-down structural analysis of an intact monoclonal antibody by electron capture dissociation–fourier transform ion cyclotron resonance–mass spectrometry. *Anal. Chem.* 85(9), 4239–4246 (2013).
- For a 148 kDa mAb, a sequence coverage of 34.4% was achieved by the fragmentation of all charge states at once using electron capture dissociation top-down mass spectrometry (MS).
- 147 Zhang Z, Shah B. Characterization of variable regions of monoclonal antibodies by top-down mass spectrometry. *Anal. Chem.* 79(15), 5723–5729 (2007).
- 148 Bondarenko PV, Second TP, Zabrouskov V, Makarov AA, Zhang Z. 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.* 20(8), 1415–1424 (2009).
- Describes the use of top-down and middle-down MS to characterize different species of therapeutic mAbs.
- 149 Han X, Jin M, Breuker K, McLafferty FW. Extending top-down mass spectrometry to proteins with masses greater than 200 kilodaltons. *Science* 314(5796), 109–112 (2006).
- 150 Mikhailov VA, Cooper HJ. Activated ion electron capture dissociation (AI ECD) of proteins: synchronization of infrared and electron irradiation with ion magnetron motion. *J. Am. Soc. Mass Spectrom.* 20(5), 763–771 (2009).
- 151 Horn DM, Ge Y, McLafferty FW. Activated ion electron capture dissociation for mass spectral sequencing of larger (42 kDa) proteins. *Anal. Chem.* 72(20), 4778–4784 (2000).
- 152 Tsybin YO, Witt M, Baykut G, Kjeldsen F, Hakansson P. Combined infrared multiphoton dissociation and electron capture dissociation with a hollow electron beam in Fourier transform ion cyclotron resonance mass spectrometry. *Rapid Commun. Mass Spectrom.* 17(15), 1759–1768 (2003).
- 153 Oh HB, McLafferty FW. Variety of activation methods employed in 'activated-ion' electron capture dissociation mass spectrometry: a test against bovine ubiquitin 7⁺ ions. *Bull. Korean. Chem. Soc.* 27, 389–394 (2006).
- 154 Horn DM, Breuker K, Frank AJ, McLafferty FW. Kinetic intermediates in the folding of gaseous protein ions characterized by electron capture dissociation mass

- spectrometry. *J. Am. Chem. Soc.* 123(40), 9792–9799 (2001).
- 155 Hakansson K, Chalmers MJ, Quinn JP, McFarland MA, Hendrickson CL, Marshall AG. Combined electron capture and infrared multiphoton dissociation for multistage MS/MS in a Fourier transform ion cyclotron resonance mass spectrometer. *Anal. Chem.* 75(13), 3256–3262 (2003).
- 156 Tsybin YO, He H, Emmett MR, Hendrickson CL, Marshall AG. Ion activation in electron capture dissociation to distinguish between N-terminal and C-terminal product ions. *Anal. Chem.* 79(20), 7596–7602 (2007).
- 157 Ge Y, Lawhorn BG, Elnaggar M *et al.* Top down characterization of larger proteins (45 kDa) by electron capture dissociation mass spectrometry. *J. Am. Chem. Soc.* 124(4), 672–678 (2002).
- 158 Zabrouskov V, Whitelegge JP. Increased coverage in the transmembrane domain with activated-ion electron capture dissociation for top-down Fourier-transform mass spectrometry of integral membrane proteins. *J. Proteome Res.* 6(6), 2205–2210 (2007).
- 159 Ledvina AR, Beauchene NA, McAlister GC *et al.* Activated-ion electron transfer dissociation improves the ability of electron transfer dissociation to identify peptides in a complex mixture. *Anal. Chem.* 82(24), 10068–10074 (2010).
- 160 Ledvina AR, Rose CM, McAlister GC *et al.* Activated ion ETD performed in a modified collision cell on a hybrid QLT-orbitrap mass spectrometer. *J. Am. Soc. Mass Spectrom.* doi:10.1007/s13361-013-0621-1 (2013) (Epub ahead of print).
- 161 Suckau D, Resemann A, Schuerenberg M, Hufnagel P, Franzen J, Holle A. A novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics. *Anal. Bioanal. Chem.* 376(7), 952–965 (2003).
- 162 Brown RS, Lennon JJ. Sequence-specific fragmentation of matrix-assisted laser-desorbed protein/peptide ions. *Anal. Chem.* 67(21), 3990–3999 (1995).
- 163 Hardouin J. Protein sequence information by matrix-assisted laser desorption/ionization in-source decay mass spectrometry. *Mass Spectrom. Rev.* 26(5), 672–682 (2007).
- 164 Takayama M. N-Ca bond cleavages of the peptide backbone via hydrogen abstraction. *J. Am. Soc. Mass Spectrom.* 12, 1044–1049 (2001).
- 165 Takayama M. Susceptible region of myoglobins to in-source decay using matrix-assisted laser desorption/ionization coupled with delayed extraction reflectron time-of-flight mass spectrometer. *J. Mass Spectrom. Soc. Japon* 50, 304–310 (2002).
- 166 Kocher T, Engstrom A, Zubarev RA. Fragmentation of peptides in MALDI in-source decay mediated by hydrogen radicals. *Anal. Chem.* 77(1), 172–177 (2005).
- 167 Mann M, Meng CK, Fenn JB. Interpreting mass spectra of multiply charged ions. *Anal. Chem.* 61(15), 1702–1708 (1989).
- 168 Senko MW, Beu SC, McLafferty FW. Automated assignment of charge states from resolved isotopic peaks for multiply charged ions. *J. Am. Chem. Soc.* 6(1), 52–56 (1995).
- 169 Senko MW, Beu SC, McLafferty FW. Determination of monoisotopic masses and ion populations for large biomolecules from resolved isotopic distributions. *J. Am. Chem. Soc.* 6(4), 229–233 (1995).
- 170 Zhang Z, Guan S, Marshall AG. Enhancement of the effective resolution of mass spectra of high-mass biomolecules by maximum entropy-based deconvolution to eliminate the isotopic natural abundance distribution. *J. Am. Chem. Soc.* 8(6), 659–670 (1997).
- 171 Zhang Z, Marshall AG. A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra. *J. Am. Soc. Mass Spectrom.* 9(3), 225–233 (1998).
- 172 Horn DM, Zubarev RA, McLafferty FW. Automated reduction and interpretation of high resolution electrospray mass spectra of large molecules. *J. Am. Soc. Mass Spectrom.* 11(4), 320–332 (2000).
- 173 Meng F, Cargile BJ, Miller LM, Forbes AJ, Johnson JR, Kelleher NL. Informatics and multiplexing of intact protein identification in bacteria and the archaea. *Nat. Biotechnol.* 19(10), 952–957 (2001).
- 174 Karabacak NM, Li L, Tiwari A *et al.* Sensitive and specific identification of wild type and variant proteins from 8 to 669 kDa using top-down mass spectrometry. *Mol. Cell Proteomics* 8(4), 846–856 (2009).
- 175 Tsai YS, Scherl A, Shaw JL *et al.* Precursor ion independent algorithm for top-down shotgun proteomics. *J. Am. Soc. Mass Spectrom.* 20(11), 2154–2166 (2009).
- 176 Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20(18), 3551–3567 (1999).
- 177 Kellie JF, Tran JC, Lee JE *et al.* The emerging process of top down mass spectrometry for protein analysis: biomarkers, protein-therapeutics, and achieving high throughput. *Mol. Biosyst.* 6(9), 1532–1539 (2010).
- 178 Tremintin G, Albers C, Main LC, Baessmann C, Jabs W. Development of a methodology for rapid characterization of biological therapeutics, ThP 35. Presented at: *59th ASMS Conference on Mass Spectrometry and Allied Topics*. Denver, CO, USA. 5–9 June 2011.
- 179 Lubeck M, Albers C, Main LC, Baessmann C, Jabs W. A fully automated solution for quality control of biological therapeutics, P1183. Presented at: *10th HUPO World Congress*. Geneva, Switzerland, 3–7 September 2011.
- 180 Ferrige AG, Seddon MJ. Maximum entropy deconvolution in electrospray mass spectrometry. *Rapid Commun. Mass Spectrom.* 5, 374–379 (1991).
- 181 Lin S, Hao Z, Decrop W *et al.* Automated mAb workflow – from harvest cell culture to intact mass analysis of variants. *Column* 9(9), 1318 (2013).
- 182 Ahlf DR, Compton PD, Tran JC, Early BP, Thomas PM, Kelleher NL. Evaluation of the compact high-field orbitrap for top-down proteomics of human cells. *J. Proteome Res.* 11(8), 4308–4314 (2012).
- 183 Thermo Scientific™ Solutions For Intact-Protein Analysis – Better, Faster Decisions For Biotherapeutic Development. *Thermo Scientific brochure*. Thermo Fisher Scientific, CA, USA (2013).

- 184 Suckau D, Resemann A. T³-sequencing: targeted characterization of the N- and C-termini of undigested proteins by mass spectrometry. *Anal. Chem.* 75(21), 5817–5824 (2003).
- 185 Suckau D, Resemann A. MALDI top-down sequencing: calling N- and C-terminal sequence with high confidence and speed. *J. Biomol. Tech.* 20(5), 258–262 (2009).
- 186 Resemann A, Wunderlich D, Rothbauer U *et al.* Top-down *de novo* protein sequencing of a 13.6 kDa camelid single heavy chain antibody by matrix-assisted laser desorption/ionization-time-of-flight/time-of-flight mass spectrometry. *Anal. Chem.* 82(8), 3283–3292 (2010).
- **Complete *de novo* sequencing of a 13.6 kDa camelid single heavy chain antibody by matrix-assisted laser desorption/ionization-top-down-MS using T³-sequencing.**
- 187 Moreno-Gordaliza E, Canas B, Palacios MA, Gomez-Gomez MM. Top-down mass spectrometric approach for the full characterization of insulin-cisplatin adducts. *Anal. Chem.* 81(9), 3507–3516 (2009).
- 188 Lee SW, Berger SJ, Martinovic S *et al.* Direct mass spectrometric analysis of intact proteins of the yeast large ribosomal subunit using capillary LC/FTICR. *Proc. Natl Acad. Sci. USA* 99(9), 5942–5947 (2002).
- 189 Zabrouskov V, Senko MW, Du Y, Leduc RD, Kelleher NL. New and automated MSn approaches for top-down identification of modified proteins. *J. Am. Soc. Mass Spectrom.* 16(12), 2027–2038 (2005).
- 190 Wynne C, Edwards NJ, Fenselau C. Phyloproteomic classification of unsequenced organisms by top-down identification of bacterial proteins using capLC-MS/MS on an Orbitrap. *Proteomics* 10(20), 3631–3643 (2010).
- 191 Lee JE, Kellie JF, Tran JC *et al.* A robust two-dimensional separation for top-down tandem mass spectrometry of the low-mass proteome. *J. Am. Soc. Mass Spectrom.* 20(12), 2183–2191 (2009).
- 192 Tipton JD, Tran JC, Catherman AD *et al.* Nano-LC FTICR tandem mass spectrometry for top-down proteomics: routine baseline unit mass resolution of whole cell lysate proteins up to 72 kDa. *Anal. Chem.* 84(5), 2111–2117 (2012).
- 193 Ren D, Pipes GD, Hambly D, Bondarenko PV, Treuheit MJ, Gadgil HS. Top-down N-terminal sequencing of Immunoglobulin subunits with electrospray ionization time of flight mass spectrometry. *Anal. Biochem.* 384(1), 42–48 (2009).
- 194 Zabrouskov V, Han X, Welker E *et al.* Stepwise deamidation of ribonuclease A at five sites determined by top down mass spectrometry. *Biochemistry* 45(3), 987–992 (2006).
- 195 Zabrouskov V, Ge Y, Schwartz J, Walker JW. Unraveling molecular complexity of phosphorylated human cardiac troponin I by top down electron capture dissociation/electron transfer dissociation mass spectrometry. *Mol. Cell Proteomics* 7(10), 1838–1849 (2008).
- 196 Pesavento JJ, Mizzen CA, Kelleher NL. Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: human histone H4. *Anal. Chem.* 78(13), 4271–4280 (2006).
- 197 Valeja SG, Kaiser NK, Xian F, Hendrickson CL, Rouse JC, Marshall AG. Unit mass baseline resolution for an intact 148 kDa therapeutic monoclonal antibody by Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* 83(22), 8391–8395 (2011).
- **A unit mass baseline resolution of a 148 kDa therapeutical mAb was achieved using a Fourier transform ion cyclotron mass spectrometer equipped with a 9 Tesla magnet.**
- 198 Nikolaev EN, Boldin IA, Jertz R, Fuchser J, Kriete C, Baykut G. Isotopically resolved FT-ICR mass spectrometry of proteins above 100 kDa at 7 Tesla. P1588. Presented at: *10th HUPO World Congress*. Geneva, Switzerland, 3–7 September 2011.
- **With a new developed ion cyclotron resonance cell it was possible to achieve a unit mass baseline resolution of different proteins between 66 and 147 kDa including a mAb with a 7 Tesla Fourier transform ion cyclotron.**
- 199 Zhang J, Liu H, Katta V. Structural characterization of intact antibodies by high-resolution LTQ Orbitrap mass spectrometry. *J. Mass Spectrom.* 45(1), 112–120 (2010).
- 200 Wang B, Gucinski AC, Keire DA, Buhse LF, Boyne MT 2nd. Structural comparison of two anti-CD20 monoclonal antibody drug products using middle-down mass spectrometry. *Analyst* 138(10), 3058–3065 (2013).
- 201 Jungblut P, Thiede B, Zimny-Arndt U *et al.* Resolution power of two-dimensional electrophoresis and identification of proteins from gels. *Electrophoresis* 17(5), 839–847 (1996).
- 202 Carini M, Regazzoni L, Aldini G. Mass spectrometric strategies and their applications for molecular mass determination of recombinant therapeutic proteins. *Curr. Pharm. Biotechnol.* 12(10), 1548–1557 (2011).
- 203 Felten C, Solano OS. Capillary electrophoresis in quality control: part I: application for therapeutic proteins. Beckman Coulter, Inc., CA, USA (2010).
- 204 Ge Y, Rybakova IN, Xu Q, Moss RL. Top-down high-resolution mass spectrometry of cardiac myosin binding protein C revealed that truncation alters protein phosphorylation state. *Proc. Natl Acad. Sci. USA* 106(31), 12658–12663 (2009).
- 205 Ayaz-Guner S, Zhang J, Li L, Walker JW, Ge Y. *In vivo* phosphorylation site mapping in mouse cardiac troponin I by high resolution top-down electron capture dissociation mass spectrometry: Ser22/23 are the only sites basally phosphorylated. *Biochemistry* 48(34), 8161–8170 (2009).
- 206 Zhang Z, Tan M, Xie Z, Dai L, Chen Y, Zhao Y. Identification of lysine succinylation as a new post-translational modification. *Nat. Chem. Biol.* 7(1), 58–63 (2011).
- 207 Yang H, Zubarev RA. Mass spectrometric analysis of asparagine deamidation and aspartate isomerization in polypeptides. *Electrophoresis* 31(11), 1764–1772 (2010).
- 208 Johnson BA, Aswad DW. Optimal conditions for the use of protein L-isoaspartyl methyltransferase in assessing the isoaspartate content of peptides and proteins. *Anal. Biochem.* 192(2), 384–391 (1991).
- 209 Kharbanda KK, Mailliard ME, Baldwin CR, Sorrell MF, Tuma DJ. Accumulation of proteins bearing atypical isoaspartyl residues in livers of alcohol-fed rats is prevented by betaine administration: effects on protein-L-isoaspartyl

- methyltransferase activity. *J. Hepatol.* 46(6), 1119–1125 (2007).
- 210 Schurter BT, Aswad DW. Analysis of isoaspartate in peptides and proteins without the use of radioisotopes. *Anal. Biochem.* 282(2), 227–231 (2000).
- 211 Pan S, Chen R, Aebersold R, Brentnall TA. Mass spectrometry based glycoproteomics – from a proteomics perspective. *Mol. Cell Proteomics* 10(1), R110.003251 (2011).
- 212 Schnaible V, Wefing S, Resemann A *et al.* Screening for disulfide bonds in proteins by MALDI in-source decay and LIFT-TOF/TOF-MS. *Anal. Chem.* 74(19), 4980–4988 (2002).
- 213 Bagal D, Valliere-Douglass JF, Balland A, Schnier PD. Resolving disulfide structural isoforms of IgG2 monoclonal antibodies by ion mobility mass spectrometry. *Anal. Chem.* 82(16), 6751–6755 (2010).
- 214 Kaltashov IA, Bobst CE, Abzalimov RR, Berkowitz SA, Houde D. Conformation and dynamics of biopharmaceuticals: transition of mass spectrometry-based tools from academe to industry. *J. Am. Soc. Mass Spectrom.* 21(3), 323–337 (2010).
- 215 Bobst CE, Abzalimov RR, Houde D *et al.* Detection and characterization of altered conformations of protein pharmaceuticals using complementary mass spectrometry-based approaches. *Anal. Chem.* 80(19), 7473–7481 (2008).
- 216 Rutherford SM, Dunn BM. Quantitative amino acid analysis. *Curr. Protoc. Protein Sci.* Chapter 3, Unit 3.2 (2011).
- 217 Roepstorff P, Fohlman J. Letter to the Editors. *Biol. Mass Spectrom.* 11(11), 601 (1984).

» Websites

- 301 EGC by Procognia. Glycosylation analysis. Procognia Ltd, Ashdod, Israel (2012).
www.procognia.com/doc/EGC_brochure.pdf
- 302 Ludger. Ludger Ltd, Culham Science Centre, Oxfordshire: Guide to Glycosylation Analysis (2013).
www.ludger.com/support/guide.php