APPLICATION NOTE 21919

Confident monoclonal antibody sequence verification by complementary LC-MS techniques

Authors
Amy Farrell[1], Sara Carillo[1], Jonathan Bones[1], Kai Scheffler[2] and Ken Cook[3]
[1]National Institute for Bioprocessing Research and Training (NIBRT), Dublin, Ireland; [2]Thermo Fisher Scientific, Germering, Germany; [3]Thermo Fisher Scientific, Hemel Hempstead, UK

Keywords
NIBRT, biopharmaceutical, biotherapeutic, monoclonal antibody (mAb), IgG, peptide mapping, middle-up, intact protein analysis, sequence coverage, MAbPac RP columns, Vanquish Flex Binary UHPLC system, Q Exactive HF Hybrid Quadrupole-Orbitrap mass spectrometer, Q Exactive Plus Hybrid Quadrupole-Orbitrap mass spectrometer, BioPharma Finder

Application benefits

- Increased information from using multiple characterization methods for therapeutic protein primary sequence verification
- Using Thermo Scientific™ MAbPac™ RP columns for intact and middle-up applications to achievie very good separation with minimal carryover
- The Thermo Scientific™ Q Exactive™ Plus and Q Exactive™ HF Hybrid Quadrupole-Orbitrap™ mass spectrometers enable high-resolution accurate mass MS analysis of monoclonal antibodies on the intact, sub-unit, and peptide level with high confidence.

Goal

To highlight the possibility of errors originating from using only a single technique for primary sequence identification and to show the benefits of the application of multiple, orthogonal techniques to address this problem. Discussing the importance of investigation of protein primary sequence at different domains together with a high level of certainty on the data generated thanks to high-resolution accurate mass MS techniques.

Introduction

Recombinant monoclonal antibodies (mAbs) are well-established pharmacological therapeutics featuring long serum half-life in humans,

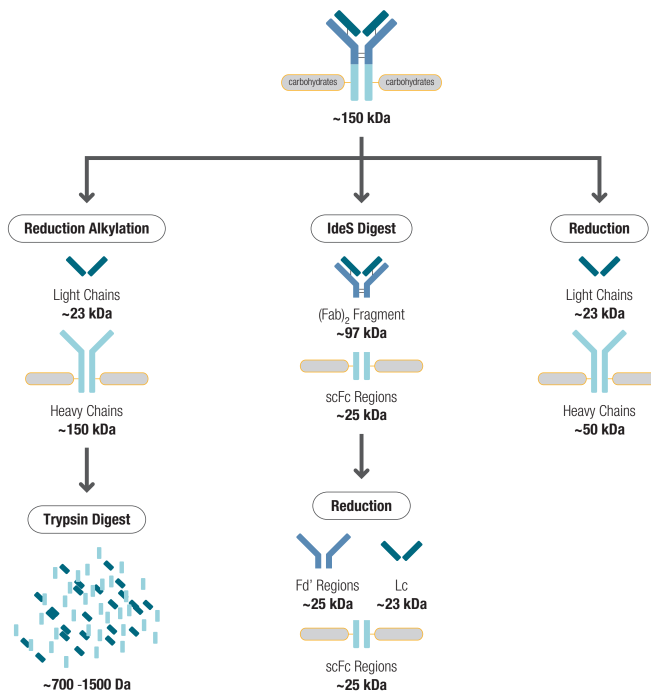
nibrt National Institute for Bioprocessing Research and Training

ThermoFisher SCIENTIFIC

high specificity to target antigens, and capabilities for the treatment of a wide range of ailments, such as cancer and inflammatory diseases. Therapeutic proteins, such as mAbs, are heterogeneous in nature, existing as a mixture of iso forms due to their tendency to undergo post-translational modifications (PTMs), potentially resulting in changes to their functional activity and structure. Due to the inherent complexity of therapeutic mAb samples, regulatory agencies require comprehensive characterization of mAb features to ensure product quality, safety, and efficacy.

Primary sequence verification is an important characterization step for therapeutic proteins and is most frequently performed using a peptide mapping experiment, wherein a protein is treated with a protease (e.g. trypsin) to produce a series of peptides, which are separated and detected, most commonly analyzed by liquid chromatography – tandem mass spectrometry (LC-MS/MS) and then interpreted using bioinformatics software. Peptide mapping also enables the identification and relative quantitation of PTMs including deamidation, oxidation, and glycosylation. Confirmation of primary structure is also critical for the characterization of biosimilars, as biosimilars must have an amino acid sequence that is identical to an innovator drug product in order to achieve regulatory approval in Europe and the USA.

A



B

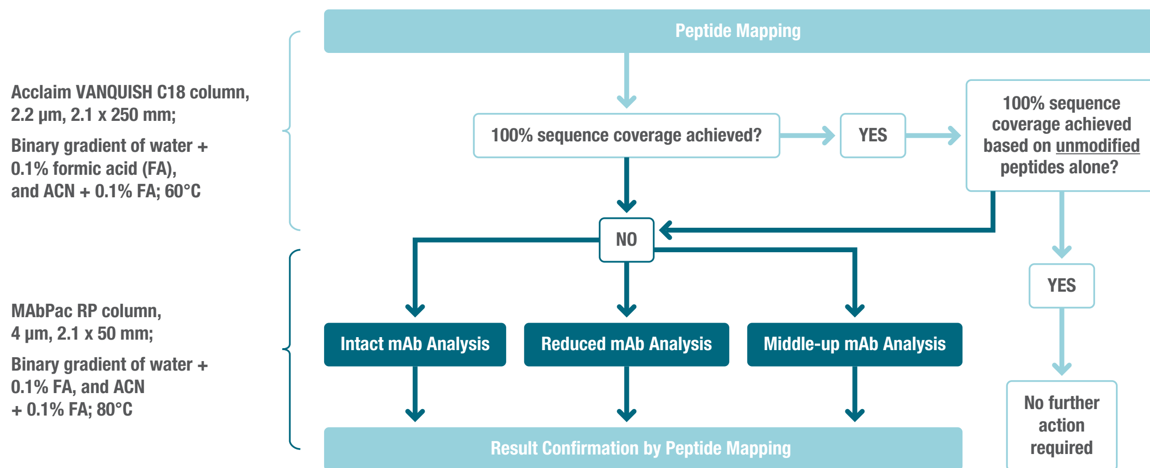


Figure 1. (A) Sample preparation workflow for bottom-up, middle-up, and reduced mAb analysis; (B) Flow of experiments to investigate inconclusive sequence verification results by peptide mapping analysis

Although a peptide mapping experiment may be useful for sequence verification and PTM identification, complementary LC-MS techniques may be advantageous when investigating erroneous amino acid substitutions in the primary structure of proteins and to perform comparative assessments of candidate biosimilars and innovator drug products. Using multiple complementary analytical techniques, such as intact protein and middle-up LC-MS analyses, variations in the primary structure of innovators and biosimilars have been observed, showing the benefits of using a multifaceted approach to primary sequence confirmation.[1,2]

Herein, we describe a solid framework of mass spectrometry-based methods for amino acid sequence assessment performed on a high-resolution analytical platform consisting of a Thermo Scientific™ Vanquish™ Flex Binary UHPLC and Thermo Scientific™ Q Exactive™ Plus and Thermo Scientific™ Q Exactive™ HF Hybrid Quadrupole-Orbitrap™ mass spectrometers. A combination of LC-MS techniques was applied to confidently elucidate the primary structure of a humanized IgG1 produced in-house. An amino acid sequence is available for the IgG1 under evaluation; however, natural variants incorporating some amino acids substitutions are known for the mAb heavy chain. Due to possible variations in the IgG1 primary structure, the proposed sequence was evaluated by intact molecular weight measurements and middle-up techniques (i.e. molecular weight determination on the domain level), as well as bottom-up peptide mapping experiments. The information gained from these experiments was used to assess the amino acid sequence, to localize sequence variants, and to enable the confident determination of the primary sequence (Figure 1).

Experimental

Recommended consumables

- Deionized, 18.2 MΩ•cm resistivity
- Water, Optima™ LC/MS grade (Fisher Chemical) (P/N 10505904)
- Water with 0.1% formic acid (v/v), Optima™ LC/MS grade (Fisher Chemical) (P/N 10188164)
- Acetonitrile with 0.1% formic acid (v/v), Optima™ LC/MS grade (Fisher Chemical) (P/N 10118464)
- Dithiothreitol (DTT) (Fisher Bioreagents) (P/N 10386833)
- Iodoacetic acid (IAA) (Acros Organics™) (P/N 10235940)
- Thermo Scientific™ Pierce™ MS Grade Trypsin Protease (P/N 13464189)
- Thermo Scientific™ Pierce™ Tris(2-carboxyethyl) phosphine hydrochloride (TCEP) (P/N 20490)
- Ammonium hydrogen carbonate (Acros Organics) (P/N 393212500)
- IdeS enzyme: FabRICATOR® (Genovis) (P/N A0-FR1-020)
- Urea, 99% (Acros Organics) (P/N 424580025)
- Thermo Scientific™ Pierce™ Protein Concentrators, 10 KDa MWCO, 0.5 mL (P/N 88513)
- Thermo Scientific™ Pierce™ 8 M Guanidine-HCl (P/N 10167783)
- Tris-HCl (Fisher Chemical) (P/N 10142400)
- Thermo Scientific™ Acclaim™ VANQUISH™ C18 column, 2.2 µm, 2.1 × 250 mm (P/N 074812-V)
- MAbPac RP column, 4 µm, 2.1 × 50 mm (P/N 088648)
- Thermo Scientific™ Virtuoso™ vial, clear 2 mL kit with septa and cap (P/N 60180-VT405)
- Thermo Scientific™ Virtuoso™ vial identification system (P/N 60180-VT100)

Sample handling equipment

Vanquish Flex Binary UHPLC system including:

- Binary Pump F (P/N VF-P10-A)
- Column Compartment H (P/N VH-C10-A)
- Split Sampler FT (P/N VF-A10-A)
- System Base Vanquish Horizon (P/N VH-S01-A)

- Q Exactive HF Hybrid Quadrupole-Orbitrap mass spectrometer (P/N IQLAAEGAAPFALGMBFZ) equipped with the Thermo Scientific™ Ion Max source with HESI II probe
- Q Exactive Plus Hybrid Quadrupole-Orbitrap mass spectrometer (P/N IQLAAEGAAPFALGMBDK) equipped with the Ion Max source with HESI II probe

Sample preparation

Bottom-up analysis of monoclonal antibody tryptic peptides

50 µg of mAb were diluted in 8 M urea in 100 mM Tris-HCl and added to a pre-rinsed 10 kDa MWCO filter. The sample was reduced by addition of DTT to a final concentration of 5 mM DTT and incubated at 65 °C for 30 minutes followed by alkylation with a final concentration of 15 mM IAA at 20 °C for 30 minutes in darkness. Samples were buffer exchanged into 50 mM ammonium bicarbonate. A ratio of 1:50 (w/w, enzyme/protein) was used for digestion with trypsin at 37 °C for 16 hours. Generated peptides were eluted from the filters by centrifugation and subsequently dried via vacuum centrifugation.

Intact protein analysis in reduced condition

10 µL of 8 M guanidine hydrochloride were added to 10 µL of IgG1 (2 µg/µL) and vortexed rigorously for 10 s before 2 µL of 500 mM tris-(2-carboxyethyl)phosphine hydrochloride (TCEP) were added. The mixture was then incubated for 45 minutes at room temperature.

Middle up analysis of IdeS-digested mAb

100 µg of IgG1 in phosphate buffered saline (PBS) were combined with 1 µL of the IdeS enzymatic digestion solution (67 units IdeS/µL in Optima grade water) and incubated at 37 °C for 2 hours with shaking at 500 rpm. Disulfide bonds were reduced by first adding guanidine hydrochloride solution to a final concentration of 4 M followed by rigorous vortexing for 10 s and then addition of TCEP to a final concentration of 50 mM. Following incubation for 45 minutes at 56 °C, samples were reduced to dryness via vacuum centrifugation.

LC conditions

Peptide mapping

Mobile phase A:	Water with 0.1% formic acid (v/v)
Mobile phase B:	Acetonitrile with 0.1% formic acid (v/v)
Flow rate:	0.25 mL/min
Column:	Acclaim VANQUISH C18, 2.2 µm, 2.1 × 250 mm
Column temperature:	60 °C (Still air mode)
Autosampler temp.:	5 °C
Injection volume:	10 µL
Injection wash solvent:	Methanol/water, 10:90
Needle wash:	Enabled pre-injection
Gradient:	See Table 1 for details

mAb intact analysis

Mobile phase A:	Water with 0.1% formic acid (v/v)
Mobile phase B:	Acetonitrile with 0.1% formic acid (v/v)
Flow rate:	0.25 mL/min
Column:	MAbPac RP, 4 µm, 2.1 × 50 mm
Column temperature:	80 °C (Still air mode)
Autosampler temp.:	5 °C
Injection volume:	1 µL
Injection wash solvent:	Methanol/water, 10:90
Needle wash:	Enabled pre-injection
Gradient:	See Table 2 for details

Intact protein analysis in reduced condition

Mobile phase A:	Water with 0.1% formic acid (v/v)
Mobile phase B:	Acetonitrile with 0.1% formic acid (v/v)
Flow rate:	0.25 mL/min
Column:	MAbPac RP, 4 µm, 2.1 × 50 mm
Column temperature:	80 °C (Still air mode)
Autosampler temp.:	5 °C
Injection volume:	2 µL
Injection wash solvent:	Methanol/water, 10:90
Needle wash:	Enabled pre-injection
Gradient:	See Table 3 for details

Middle-up

Mobile phase A:	Water with 0.1% formic acid (v/v)
Mobile phase B:	Acetonitrile with 0.1% formic acid (v/v)
Flow rate:	0.25 mL/min
Column:	MAbPac RP, 4 µm, 2.1 × 50 mm
Column temperature:	80 °C
Autosampler temp.:	5 °C
Injection volume:	1 µL
Injection wash solvent:	Methanol/water, 10:90
Needle wash:	Enabled pre-injection
Gradient:	See Table 4 for details

Table 1. Mobile phase gradient for bottom-up analysis

Time (min)	%A	%B	Curve
0.0	98	2	5
40.0	70	30	5
43.0	20	80	5
46.5	20	80	5
47.0	98	2	5
65.0	98	2	5

Table 2. Mobile phase gradient for intact protein analysis

Time (min)	%A	%B	Curve
0.0	75	25	5
1.0	75	25	5
9.0	65	35	5
10.0	20	80	5
11.0	20	80	5
12.0	75	25	5
19.0	75	25	5

Table 3. Mobile phase gradient for intact protein analysis in reduced condition

Time (min)	%A	%B	Curve
0.0	75	25	5
1.0	75	25	5
13.0	68	32	5
14.0	20	80	5
16.0	20	80	5
16.5	75	25	5
25.0	75	25	5

Table 4. Mobile phase gradient for middle-up analysis

Time (min)	%A	%B	Curve
0.0	75	25	5
1.0	75	25	5
16.0	68	32	5
17.0	20	80	5
18.0	20	80	5
18.5	75	25	5
28.0	75	25	5

MS conditions

Detailed tune and parameter settings are listed in Tables 5, 6, and 7.

Table 5. Summary of tune parameters

MS ion source parameters	Settings for bottom-up analysis	Settings for middle-up analysis	Settings for intact protein analysis	Settings for intact reduced protein analysis
Sheath gas pressure	40 psi	25 psi	20 psi	25 psi
Auxiliary gas flow (arbitrary units)	10	10	10	10
Probe heater temperature	150 °C	150 °C	150 °C	150 °C
Source voltage	3.8 kV	3.8 kV	3.6 kV	3.6 kV
Capillary temperature	320 °C	320 °C	320 °C	320 °C
S-lens RF level	60	60	100	80

Table 6. Summary of MS parameters

MS[1] method parameters	Settings for bottom-up analysis	Settings for middle-up analysis	Settings for intact protein analysis	Settings for intact reduced protein analysis
Run time	0 to 65 min	0 to 28 min	0 to 20 min	25 min with two segments (S1, S2): S1: 0–7.0, S2: 7.0–25.0 min
Polarity	Positive	Positive	Positive	Positive
Full MS mass range (*m/z*)	200–2000	600–2400	2200–5000	S1: 600–2500, S2: 600–3000
Resolution setting	70,000	240,000	35,000	S1: 140,000, S2: 17,500
AGC target value	3.0×10^6	3.0×10^6	3.0×10^6	3.0×10^6
Max injection time	100 ms	200 ms	200 ms	200 ms
In-source CID	0 eV	0 eV	80 eV	S1: 0, S2: 20 eV
Microscans	1	5	10	S1: 5, S2: 10

Table 7. Detail of MS^2 parameters for bottom-up analysis

MS^2 method parameters	Setting
Resolution setting	17,500
AGC target value	1.0×10^5
Isolation width	2.0 *m/z*
Signal threshold	1.0×10^4
Normalized collision energy (NCE)	28
TopN MS^2	5
Max injection time	200 ms
Fixed first mass	—
Dynamic exclusion	7.0 s
Loop count	5

Table 8. Thermo Scientific™ BioPharma Finder™ software parameter settings for analysis of middle-up and intact protein analysis

Middle-up	Setting
Xtract algorithm	Charge: 5 to 50
	Mass range: 600 to 2000 *m/z*
Intact protein	**Setting**
Intact: ReSpect™ algorithm	Charge: 10 to 100
	Mass range: 2200 to 5000 *m/z*
Reduced: ReSpect algorithm	Charge: 10 to 100
	Mass range: 600 to 3000 *m/z*

MS data processing

Detailed parameter settings are shown in Tables 8 and 9.

Table 9. BioPharma Finder software parameter settings for analysis of peptide mapping data

Component detection	Setting
Absolute MS signal threshold	1.60×10^5 counts
Identification	Setting
Mass accuracy	5 ppm
Minimum confidence	0.8
Maximum number of modifications for a peptide	1
Unspecified modification	-58 to +162 Da
N-glycosylation	CHO
Protease specificity	High
Variable modifications	Setting
N terminal	Glu → Pyro-Glu
C terminal	Loss of lysine
Side chain	Deamidation (NQ) Oxidation (MW) Carboxymethylation

Results and discussion

In this study, LC-MS/MS was used in several complementary analytical approaches for elucidation of the correct amino acid sequence of a humanized IgG1. Natural variants in the amino acid sequence in the Fc region of the heavy chain of the mAb are known, namely K38 to R38, D361 to E361, and L363 to M363 (Figure 2). In order to evaluate which of the natural variants are present in the amino acid sequence of the mAb under investigation, samples were prepared and analyzed by bottom-up, middle-up, and intact LC-MS analysis, as described in the experimental section. All data obtained were used for primary sequence evaluation.

Bottom-up analysis of an IgG1 tryptic digest using a proposed amino acid sequence

Sequence coverage maps of the mAb heavy and light chains were generated following LC-MS/MS analysis of IgG1 tryptic peptides. Data were searched against the proposed IgG1 amino acid sequence, yielding 100% sequence coverage for both the mAb heavy and light chains; this would typically suggest a complete match of the data with the proposed sequence. However, upon closer inspection of the data it was observed that two areas of the mAb HC were identified based on assignment of modified peptides alone (Figure 2). Potential peptide modifications were identified based on parameter settings input into BioPharma Finder software for analysis of peptide mapping data (Table 8). For the region of the mAb HC containing the sequence DELTK, peptides were identified based on full MS spectra with no supporting MS^2 data, or based on sequences containing multiple missed cleavages and a few identified b- and y-type fragment ions (Figure 3). The region of the mAb HC containing the amino acid sequence "MHWVK" was identified based on the closely matching precursor ions (~6 ppm) and a number of matching fragment ions, yet both relating to the peptide containing methionine-34 in oxidized form (Figure 4). However, the corresponding experimental and predicted MS spectra showed a low similarity value between the spectra (6.86×10^{-2}). In addition, some of the most prominent ions in the experimental spectra remained unassigend (e.g., 472.2747 *m/z)*. The results shown in Figures 3 and 4 highlight the importance of setting acceptance criteria for peptide identification, e.g. a maximum number of missed cleavages allowed per peptide, a minimum threshold for similarity between predicted and experimental spectra for peptide identification and a requirement for MS^2 data to confirm amino acid sequence identification. Furthermore, a combination of multiple sample preparation techniques, e.g. performance of additional mAb digests with endoproteinases such as Glu-C or chymotrypsin for peptide mapping experiments, or the use of complementary LC-MS analysis, such as intact protein measurements or middle-up protein analysis, may provide additional confidence in the obtained results. For the IgG1 investigated in this study, the sections of the mAb HC that were not identified based on unmodified peptides have previously been reported to have natural variants. We suspected that the absence of unmodified peptides might be a consequence of a variation in the mAb primary structure. To confirm this observation, further investigations were performed using additional experiments as outlined in Figure 1B.

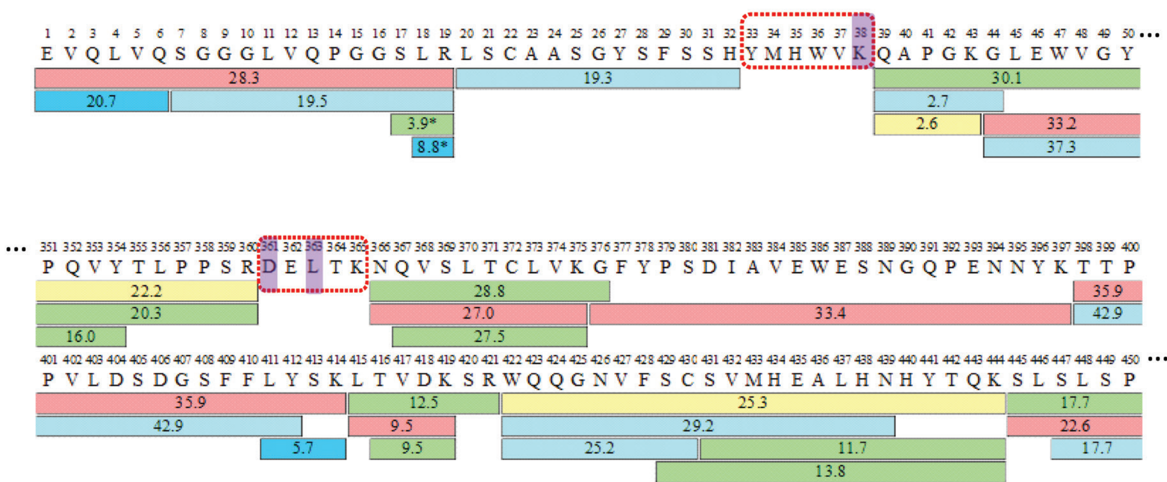


Figure 2. IgG1 sequence coverage obtained from BioPharma Finder software data analysis. Snapshots of the sequence containing the regions where the natural variants in amino acid sequence are likely to occur are highlighted in purple. The colored bars represent identified unmodified peptides with the numbers indicating the retention times.

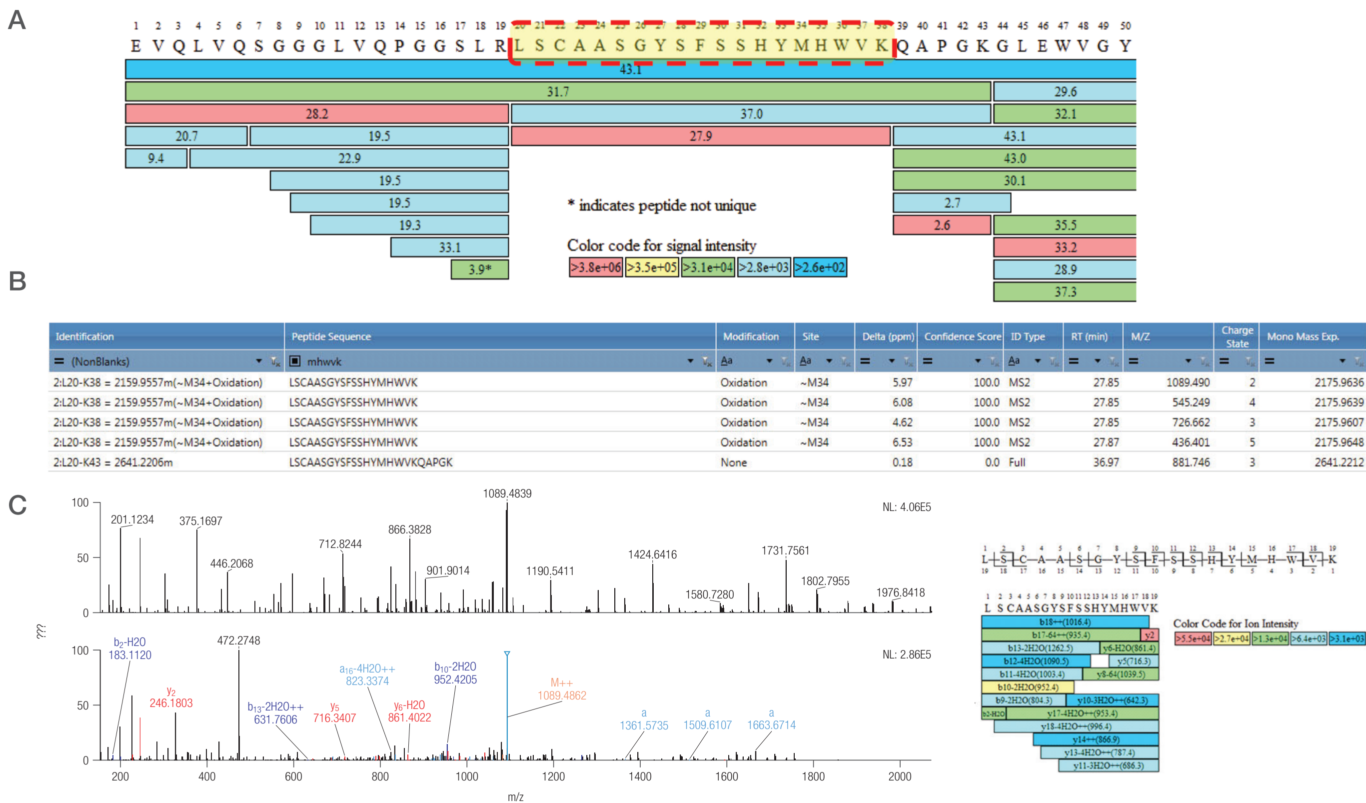


Figure 3. Identification of the peptide LSCAASGYSFSSHYMHWVK, based on the presence of oxidized Methionine-34. (A) Sequence coverage map obtained from BioPharma Finder software data analysis for a portion of the mAb heavy chain; (B) Table showing peptides containing the amino acid sequence LSCAASGYSFSSHYMHWVK, identified based on MS[1] and MS[2] data; (C) Predicted (upper panel) and experimental spectra (lower panel) for the tryptic peptide highlighted in (A). Experimental b- and y- type fragment ions used to identify the peptide LSCAASGYSFSSHYMHWVK are labeled.

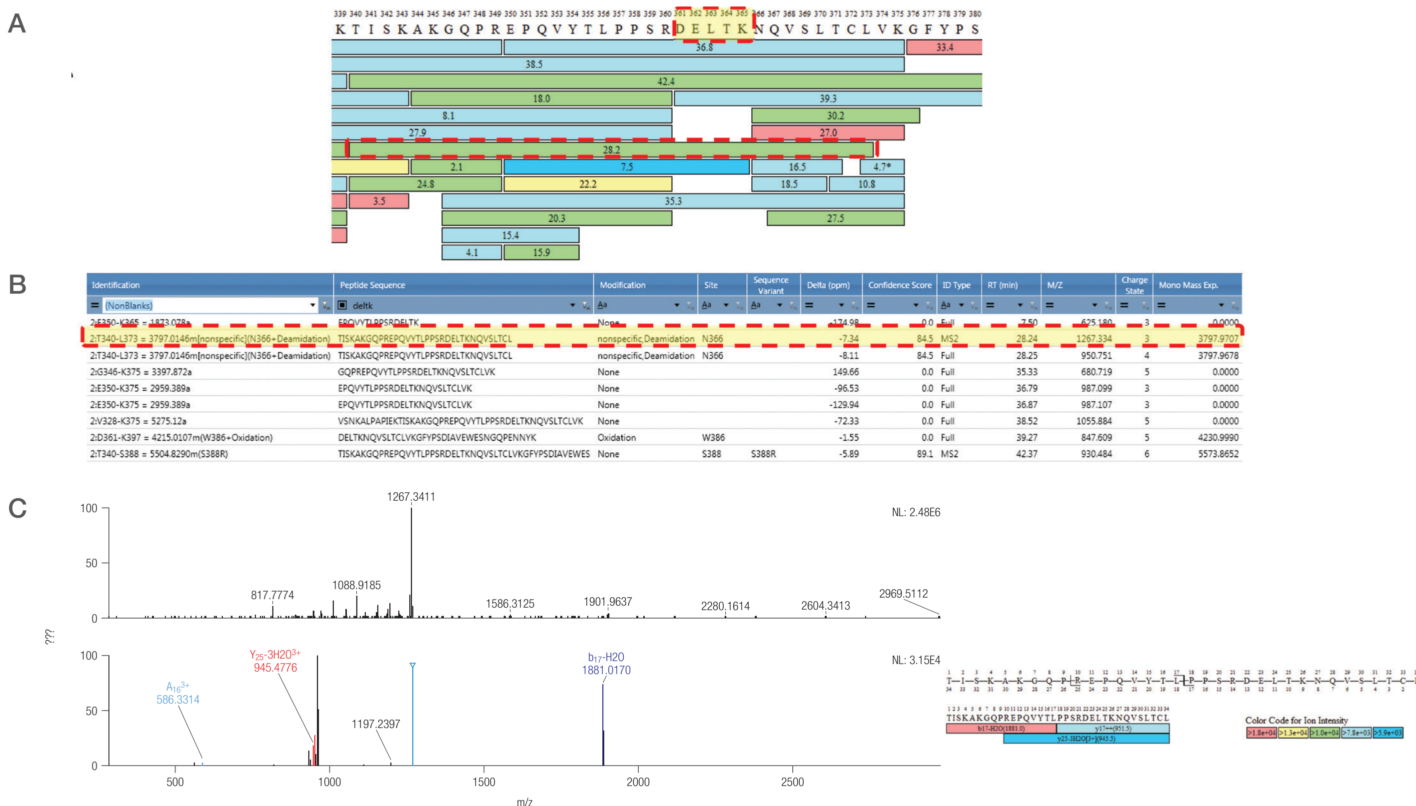


Figure 4. Identification of the amino acid sequence DELTK. (A) Sequence coverage map obtained from BioPharma Finder software data analysis for a portion of the mAb heavy chain; (B) The table is showing all peptides containing the amino acid sequence DELTK identified based on MS[1] and MS[2] data; (C) Predicted (upper spectrum) and experimental fragment ion spectra (lower panel) for the peptide DELTK highlighted in (A). Experimental b- and y- type fragment ions used to identify the peptide TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCL are labeled.

Calculation of theoretical masses for the IgG1 under evaluation

To correlate the experimental masses and interpret the LC-MS data generated, theoretical monoisotopic and average masses of the expected mAb isoforms were calculated. The theoretical mass values were corrected according to the incorporation of 16 disulfide bonds (-32 Da), loss of C-terminal lysine (-128 Da per HC), and the addition of the most commonly observed *N*-glycan species i.e. G0, G0F, G1F, and G2F contributing 1299.2, 1445.3, 1607.5, and 1769.6 Da, respectively, to the average mass of the biomolecules. Based on the peptide mapping results, pyroglutamic acid formation at the N-terminus of the heavy chain was not considered as a modification since the level of conversion of N-terminal glutamic acid to pyroglutamic acid was found to be just 0.89% (data not shown).

Mass measurement of intact IgG1

LC-MS analysis of the intact IgG1 under evaluation and subsequent data processing using BioPharma Finder software yielded high quality MS spectra and deconvoluted spectra, illustrating the high-resolution and mass accurate capabilities of the applied mass spectrometers. The analysis revealed six main glycoforms of the mAb (Figure 5). As shown in Table 10, the experimental mass of the most prominent IgG1 glycoforms in the samples analyzed were found to be inconsistent with corresponding theoretical masses (801.3–810.9 ppm mass difference). The large mass difference suggests differences in the amino acid sequence of the mAb analyzed in this study and the sequence used to calculate the theoretical mass.

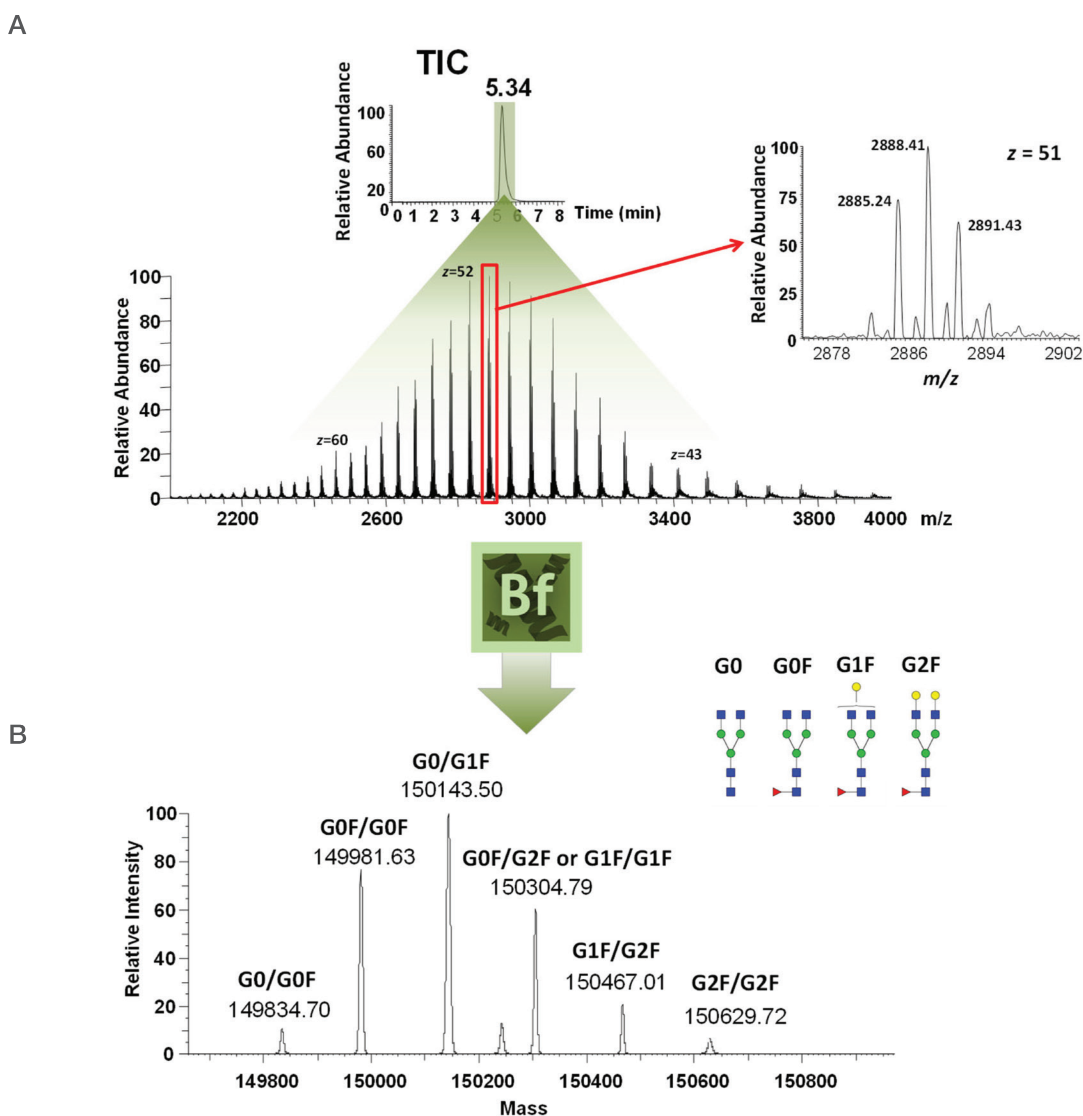


Figure 5. Analysis of the intact mAb. (A) TIC chromatogram and Full MS spectrum with a zoomed view of the most abundant charge state (z=+51); (B) Deconvolution result. Peak annotation corresponds to the experimental mass values and mAb glycoforms identified.

Table 10. Theoretical masses of the IgG1 under evaluation (proposed sequence) and the primary sequence incorporating known natural variants (verified sequence). The experimental masses for the different intact mAb glycoforms and mAb sub-units along with the difference between the experimental and theoretical mass (ppm) are also shown. Mass differences were calculated against experimental average masses determined using a resolution setting of ≤35,000 at *m/z* 200 and against monoisotopic mass for experimental masses determined using resolution settings ≥140,000 at *m/z* 200. For calculation of theoretical masses, the element masses were used as published previously.[4]

	Chain	Modifications	Experimental mass	Proposed sequence		Verified sequence	
				Theoretical average mass	Δ ppm (Da)	Theoretical average mass	Δ ppm
Intact mAb analysis	Intact	No C-term K, with 16 S-S bonds, G0/G0F	149834.70	149713.87	806.42 (120.83)	149834.03	4.47
Intact mAb analysis	Intact	No C-term K, with 16 S-S bonds, G0F/G0F	149981.63	149860.01	810.90 (121.92)	149980.17	9.73
Intact mAb analysis	Intact	No C-term K, with 16 S-S bonds, G0F/G1F	150143.50	150022.15	808.23 (121.35)	150142.31	7.93
Intact mAb analysis	Intact	No C-term K, with 16 S-S bonds, G1F/G1F	150304.79	150184.30	801.64 (120.49)	150304.45	2.26
Intact mAb analysis	Intact	No C-term K, with 16 S-S bonds, G1F/G2F	150467.01	150346.44	801.31 (120.57)	150466.60	2.72
Intact mAb analysis	Intact	No C-term K, with 16 S-S bonds, G2F/G2F	150629.72	150508.58	804.22 (121.14)	150628.74	6.51
Reduced mAb analysis	HC	No C-term K, no S-S bond, G0F	51049.55	50989.62	1173.96 (59.93)	51049.70	2.94
Reduced mAb analysis	HC	No C-term K, no S-S bond, G1F	51211.64	51151.77	1169.07 (59.87)	51211.84	3.91
Reduced mAb analysis	HC	No C-term K, no S-S bond, G2F	51373.66	51313.91	1163.05 (59.75)	51373.99	6.42
	Chain	**Modifications**	**Experimental mass**	**Theoretical mono-isotopic mass**	**Δ ppm (Da)**	**Theoretical mono-isotopic mass**	**Δ ppm**
Reduced mAb analysis	Lc	No S-S bonds	23941.9101	23941.8626	1.98	23941.8626	1.98
Middle-up analysis	Lc	No S-S bonds	23941.8932	23941.8626	1.28	23941.8626	1.28
Middle-up analysis	Fd'	No S-S bonds	25815.6493	25787.5992	1084.86 (28.0062)	25815.6054	1.7
Middle-up analysis	scFc	No C-term K, G0F	25220.4967	25188.4913	1269.02 (32.0054)	25220.4634	1.32
Middle-up analysis	scFc	No C-term K, G1F	25382.5079	25350.5441	1259.28 (31.9638)	25382.5162	0.33
Middle-up analysis	scFc	No C-term K, G2F	25544.5808	25512.5970	1252.08 (31.9838)	25544.5690	0.46

Mass measurement of reduced IgG1

To investigate whether the incongruence of mass was present in the Lc or HC of the IgG1 under evaluation, a treatment with guanidine hydrochloride and TCEP was performed to allow cleavage of all inter- and intra-chain disulfide bonds. The importance of proper sample preparation has been described previously.[3] For mass measurement of the reduced IgG1 sub-units, a method consisting of two segments was applied that switches the resolution setting at RT 7.0 min from 140,000 to 17,500. The goal of this method is the acquisition of the earlier eluting isotopically resolved light chain while

the later eluting HC, acquired at the lower resolution setting, results in isotopically unresolved spectra. The monoisotopic mass obtained upon deconvolution of the Lc spectra was determined to be 23,941.9101 Da (mass deviation 2.0 ppm). The average masses obtained upon deconvolution of the isotopically unresolved spectra of the HC, with the three most prominent *N*-glycans G0F, G1F, and G2F taken into account, were found to have large mass errors of 1173.96, 1169.07, and 1163.05 ppm. Considering the accurate mass that the applied instrument provides irrespective of the applied resolution setting and whether monoisotopic or average masses are obtained after deconvolution, the high mass deviation of approximately 60 Da obtained for the HC clearly indicates that a variation in the proposed heavy chain sequence exists.

Middle-up analysis of IdeS digested IgG1

To confirm that the change in the sequence is indeed present in the heavy chain, a middle-up approach was applied. The IgG1 under evaluation was digested with IdeS enzyme (FabRICATOR, Genovis), a cysteine protease with high specificity for the hinge region of IgG1. Following disulfide bond reduction, three mAb sub-unit populations are formed (Figure 1A): the intact light chain, a HC fragment containing the asparagine-293 N-glycosylation site and the C-terminus (scFc), and the remaining HC portion containing the N-terminus (Fd’). Following LC-MS analysis applying a high resolution setting of 240,000, data were processed using the Xtract algorithm in BioPharma Finder software. Three species corresponding to the different sub-units were detected in the total ion chromatogram (TIC).

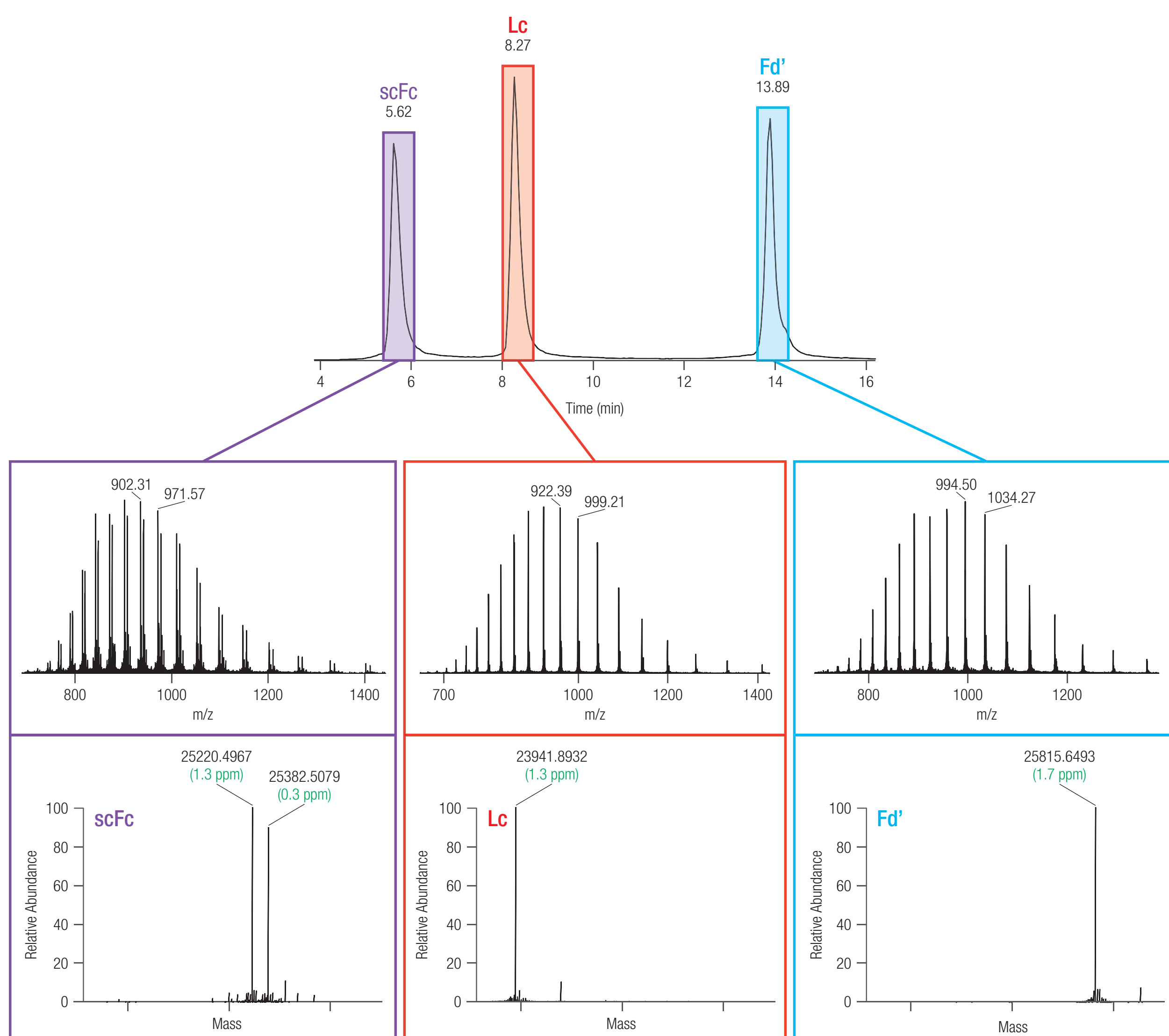


Figure 6. LC-MS analysis of IgG1 sub-units obtained upon IdeS-digest and disulfide bond reduction. (A) Total ion chromatogram; (B) Full MS spectra of the three species scFc, Lc, and Fd’; (C) Deconvoluted spectra annotated with experimental masses for each sub-unit and ppm error values calculated between the experimental and theoretical mass of the corrected mAb sequence.

The TIC, charge envelope profiles, and deconvoluted mass spectra for each peak in the TIC trace are shown in Figure 6. Experimental values for the three subunits were compared with theoretical monoisotopic masses for Lc, Fd', and the scFc, with the most prominent glycoforms G0F, G1F, and G2F taken into account (Table 10). In agreement with data generated after reduction of the intact IgG1, the mass difference between experimental and expected mass for Lc was 1.3 ppm, confirming the proposed sequence for this sub-unit. Conversely, following middle-up analysis of the sub-units relating to the heavy chain, a mass difference of ~32 Da and ~28 Da was observed for the scFc and Fd' subunits, respectively. These data suggested either an error in the proposed sequence of the heavy chain or unidentified modifications in both of the scFc and Fd' sub-units.

Correct sequence proposal

Following up on the assumption of of a possible error in the proposed sequence for the IgG1 under investigation, a correct sequence proposal was attempted. As discussed in the previous section, an error was found in two portions of the heavy chain from both middle-up and bottom-up analysis and the mass differences were found to be ~32 Da for scFc and ~28 Da for Fd' subunits of the mAb HC, respectively. On the basis of natural variants present for this sequence, some amino acids substitutions were speculated and verified. In the Fd' subunit, the peptide YMHWVK was identified exclusively as a modified peptide. A natural variant exists for the peptide with the sequence YMHWV**R**. The difference in these two peptides would account for a mass difference of ~28 Da observed in this region following middle-up LC-MS analysis. Similarly, in the scFc subunit, considering the mass difference of ~32 Da, the amino acid sequence **E**E**M**TK was evaluated in place of DELTK. To verify the modified sequence, bottom-up data analysis was repeated using the amended amino acid sequence. As a result, full sequence coverage for the heavy chain was achieved based on unmodified peptides. For the peptides assigned to the amended sequence sections the theoretical and experimental MS/MS spectra matched very well with fragment ions assigned to cover the entire peptide sequences (Figure 7).

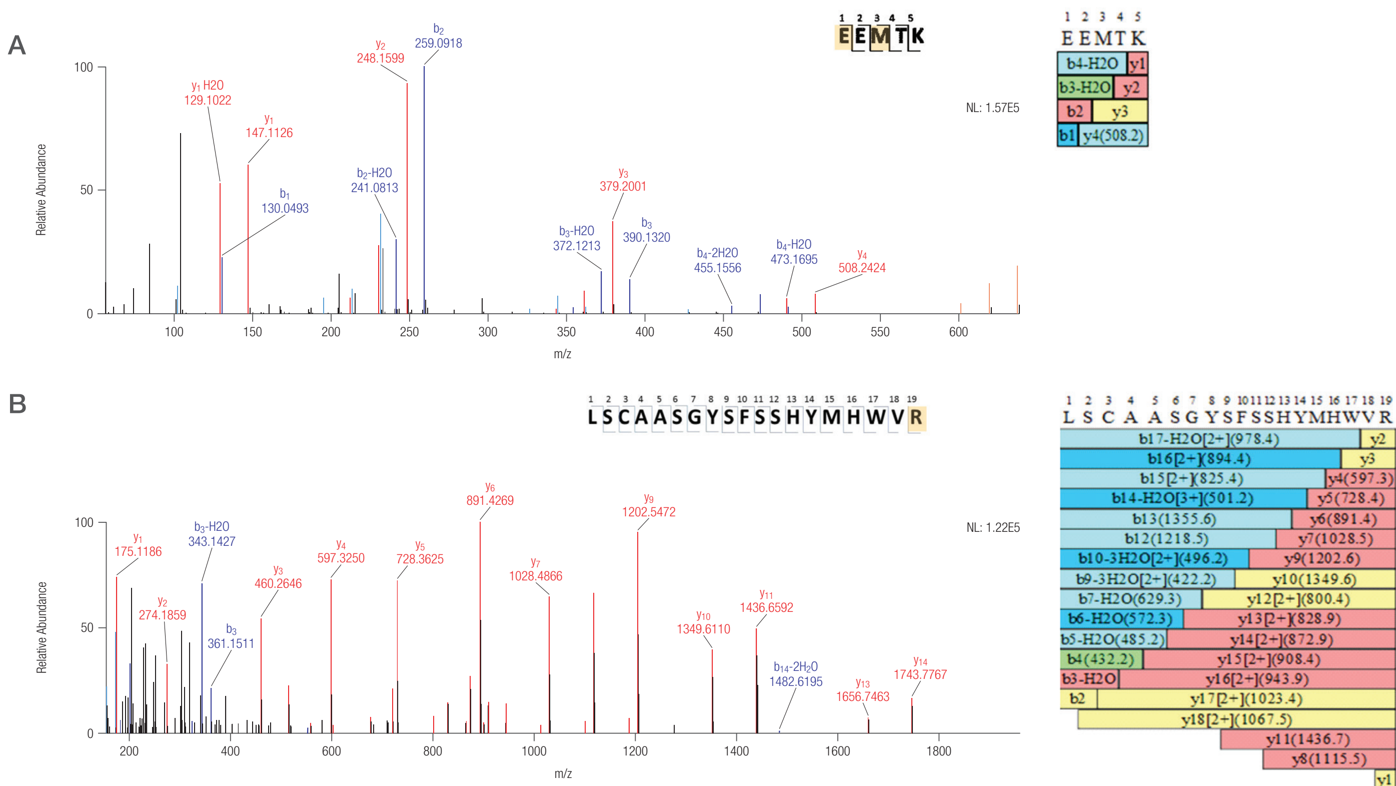


Figure 7. (A) MS/MS fragment ion spectrum and assignment of corrected sequence for peptide EEMTK; (B) MS/MS fragment ion spectrum and assignment of corrected sequence for peptide LSCAASGYSFSSHYMHWVR

Conclusions

- In this study we demonstrated the risk of confirming an incorrect primary sequence when only a single analytical approach is applied. Several cases were reported in the past on amendment of therapeutic protein sequences after a combination of orthogonal MS techniques was applied.

- An experimental workflow on a single LC-MS platform applying a peptide mapping experiment succeeded by a combination of intact protein and middle-up analysis enabled confident assessment of the mAb primary sequence.

- The high-resolution and accurate mass capabilities provided by the Q Exactive mass spectrometer platform significantly supports the high level of confidence required for primary sequence determination and confirmation of therapeutic biomolecules.

- A platform constituting MAbPac RP and Acclaim VANQUISH C18 columns together with a Vanquish Flex Binary UHPLC system hyphenated with Q Exactive Quadrupole-Orbitrap mass spectrometers has been successfully employed for biotherapeutics primary sequence assessment.

References

1. Ayoub, D.; Jabs, W.; Resemann, A.; Evers, W.; Evans, C.; Main, L. et al., Correct primary structure assessment and extensive glyco-profiling of cetuximab by combination of intact, middle-up, middle-down and bottom-up ESI and MALDI mass spectrometry techniques. *MAbs* **2013**, *5*, 699-710.
2. Tan, Q.; Guo, Q.; Fang, C.; Wang, C.; Li, B.; Wang, H. et al. Characterization and comparison of commercially available TNF receptor 2-Fc fusion protein products. *MAbs* **2012**, *4*, 761-74.
3. Scheffler, K.; Damoc, E. Antibody subunit analysis workflow on a quadrupole-Orbitrap mass spectrometer: from optimized sample preparation to data analysis. Thermo Fisher Scientific Application Note 72854 (2018). https://assets.thermofisher.com/TFS-Assets/CMD/Application-Notes/an-72854-lc-ms-antibody-subunit-analysis-an72854-en.pdf
4. Farrell, A.; Carillo, S.; Scheffler, K.; Cook, K.; Bones, J. Monoclonal antibody sequence assessment using a hybrid quadrupole-Orbitrap mass spectrometer. *Analytical Methods* **2018**, *10*, 3100-3109.

For Research Use Only. Not for use in diagnostic procedures.

More information at **thermofisher.com/NIBRT**

ThermoFisher SCIENTIFIC

©2019 Thermo Fisher Scientific Inc. All rights reserved. FabRICATOR is a registered trademark of Genovis AB. ReSpect is a trademark of Positive Probability Ltd. All other trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. This information is presented as an example of the capabilities of Thermo Fisher Scientific products. It is not intended to encourage use of these products in any manners that might infringe the intellectual property rights of others. Specifications, terms and pricing are subject to change. Not all products are available in all countries. Please consult your local sales representatives for details. **AN21919-EN 0419S**