LC/MS Analysis of the Monoclonal Antibody Rituximab Using the Q Exactive Benchtop Orbitrap Mass Spectrometer

Application Note 591

Martin Samonig[1,2], Christian Huber[1,2] and Kai Scheffler[2,3]
[1]Division of Chemistry and Bioanalytics, University of Salzburg, Salzburg, Austria
[2]Christian Doppler Laboratory for Innovative Tools for Biosimilar Characterization, University of Salzburg, Salzburg, Austria
[3]Thermo Fisher Scientific, Dreieich, Germany

Key Words

Monoclonal antibody, intact protein mass measurement, sequence confirmation, protein deconvolution, top-down sequencing

Goal

Analysis and characterization of a monoclonal antibody using an optimized LC/MS workflow based on monolithic columns coupled online with the Thermo Scientific™ Q Exactive™ benchtop Orbitrap™ mass spectrometer.

Introduction

Monoclonal antibodies (mAbs) are one of the fastest growing classes of pharmaceutical products. They play a major role in the treatment of a variety of conditions such as cancer, infectious diseases, allergies, inflammation, and auto-immune diseases. Because mAbs can exhibit significant heterogeneity, extensive analytical characterization is required to obtain approval for a new mAb as a therapeutic product. Mass spectrometry has become an essential tool in the characterization of mAbs, providing molecular weight determinations of intact proteins as well as separated light and heavy chains, elucidation of glycosylation and glycan structures, confirmation of correct amino acid sequences, and identification of impurities such as host cell proteins (HCP) inherent to the production process.

Rituximab, which is known under the trade names Rituxan® (Biogen Idec/Genentech) in the United States and MabThera® (Roche) in Europe, is a recombinantly produced, monoclonal chimeric antibody against the protein CD20. It was one of the first new generation drugs in cancer immune therapy. Rituximab was approved by the U.S. Food and Drug Administration in 1997 and by the European Commission in 1998 for cancer therapy of malignant lymphomas. The variable domain of the antibody targets the cell surface molecule CD20, that can be found in some non-Hodgkin lymphomas.

In this application note, the capabilities and performance of the Q Exactive benchtop Orbitrap mass spectrometer in analyzing the intact and reduced forms of rituximab are demonstrated as well as sequence confirmation analyses using a combined top-down and bottom-up approach.

Furthermore, the sensitivity of two chromatographic setups using monolithic columns coupled online to the mass spectrometer is evaluated. The data obtained demonstrate superior resolution and mass accuracy of the Q Exactive mass spectrometer and present it as a high-confidence screening tool for accelerated and accurate biopharmaceutical product development and characterization.

Thermo
SCIENTIFIC

Experimental

Sample Preparation

The commercially available monoclonal antibody rituximab was used in all experiments. Rituximab is a sterile, clear, colorless, preservative-free, concentrated solution for intravenous infusion. It was supplied at a concentration of 10 mg/mL, formulated in 7.35 mg/mL sodium citrate buffer containing 0.7 mg/mL polysorbate 80, 9.0 mg/mL sodium chloride, and sterile water, and ready for injection. The pH was adjusted to 6.5 with sodium hydroxide or hydrochloric acid.

Prior to LC/MS analysis, rituximab was dialyzed due to polysorbate 80 in the sample. The dialysis was performed with a Thermo Scientific™ Slide-A-Lyzer™ dialysis cassette with a molecular weight cut off (MWCO) of 3.5 kDa. A 1 mL sample of rituximab was dialyzed for 48 h against 2 L of 20% aqueous acetonitrile (ACN) at 4 °C.

For analysis of the light and heavy chains of rituximab, disulfide bonds were reduced by incubation for 30 min at 60 °C with 5 mM tris(2 carboxyethyl)phosphine (TCEP).

For the bottom-up analysis of digested mAb, the sample was alkylated with 20 mM iodoacetamide (IAA) for 30 min at room temperature in the dark after the reduction step. The sample was purified with Thermo Scientific™ Pierce™ C18 tips dried in a Thermo Scientific™ SpeedVac™ concentrator and dissolved in 0.5 M triethylammonium bicarbonate buffer (TEAB). Sequencing grade modified trypsin (Promega) was added twice in a total ratio of 1:15 (w/w) at 0 h and 1.5 h and digestion was allowed to proceed for 2.5 h at 37 °C. The digest was stopped by addition of triflouroacetic acid (TFA) to approximately pH 3.

All samples were supplied in autosampler vials containing glass inserts (micro-inserts 0.1 ml, clear glass, VWR).

Liquid Chromatography

A monolithic 160 x 0.20 mm i.d. poly(styrene-divinylbenzene) copolymer (PS-DVB) capillary column, prepared according to a previously published protocol[1], and a Thermo Scientific™ PepSwift™ monolithic 250 x 0.20 mm i.d PS-DVB capillary column were used. Protein separations were performed with a Thermo Scientific™ Dionex™ UltiMate™ 3000 RSLCnano system that included a detector equipped with a 3 nL z-shaped capillary detection cell.

Separations were accomplished at 55 °C with a gradient of 20–60% acetonitrile (ACN) in 0.050% aqueous triflouroacetic acid (TFA) in 10 min at a flow rate of 1 µL/min. For the proteolytic digest with trypsin, the gradient was adapted to run at 0–50% B in 30 min. For the reduced antibody samples, a gradient from 35–45% B in 15 min was selected.

Protein separation in a higher scale was performed using a Thermo Scientific™ ProSwift™ RP-10R monolithic 50 mm x 1.0 mm i.d. column with an UltiMate 3000 RSLCnano system that included a 45 nL detection cell. The column was run with a flow rate of 60 µL/min and a column temperature set to 55 °C. The gradient used was 26–80% B in 20 min. For the reduced antibody, a gradient of 26–56% B in 20 min was chosen to separate the heavy and the light chain.

The recorded back pressure of the monolithic columns for the gradients described above was in the range of 190 to 260 bar for the PepSwift 250 mm x 0.2 mm i.d. column and 120 to 180 bar for the ProSwift RP-10R 50 mm x 1 mm i.d. column.

For all experiments, the solvents used were water with 0.05% TFA (A) and acetonitrile with 0.05% TFA (B). The LC gradients are described in Tables 1 and 2.

Table 1. LC gradients used for experiments with the PepSwift 250 mm x 0.2 mm i.d. column, at a flow rate of 1 µL/min

Time [min]	Intact mAb [%B]	Time [min]	Reduced mAb [%B]	Time [min]	mAb Digest [%B]
0.0	20	0.0	35	0.0	0
10.0	60	15.0	45	30.0	50
10.1	85	15.1	85	30.1	85
16.0	85	21.0	85	40.0	85
16.1	20	21.1	35	40.1	0
30.0	20	30.0	35	50.0	0

Table 2. LC gradient used for experiments with the ProSwift RP-10R 50 mm x 1 mm i.d. column, at a flow rate of 60 µL/min

Time [min]	Intact mAb [%B]	Time [min]	Reduced mAb [%B]
0.0	26	0.0	26
15.0	80	15.0	56
20.0	80	15.1	80
20.1	26	20.0	80
30.0	26	20.1	26
		30.0	26

Mass Spectrometry

The Q Exactive benchtop Orbitrap mass spectrometer was used for all experiments in this study. Experiments using the ProSwift RP-10R 50 mm x 1 mm i.d. column were performed using the Thermo Scientific™ IonMax™ source with the heated electrospray ionization (HESI) sprayer, applying 4 kV spray voltage and sheath gas and auxiliary gas flow rates of 15 and 5 units, respectively.

All other experiments were performed using the Thermo Scientific™ NanoFlex™ ion source equipped with 15 cm PicoTip® emitter (New Objective, Woburn, USA; 20 µm i.d., 360 µm o.d., 10 µm tip), running with a flow rate of 1 µL/min. A source voltage of 1.5 kV was applied.

Method details are provided in Table 3.

Table 3. Mass spectrometric parameters used for all experiments

	Intact Antibody	Reduced Antibody	Top Down AIF	5-plex MS/MS (Targeted MS^2)	Antibody Digest
Method type	Full MS	Full MS (2 segments)	Full MS-AIF	Targeted MS^2	Full MS-dd top 10 HCD
Total run time	30 min	0–15.8/15.8–30 min	25 min	25 min	40 min
Scan range *m/z*	1800–5000	800–3500/700–2500	300–2500	Fixed first mass 300	350–2000
Resolution (full MS/MS^2)	17,500/x	140,000/17,500	70,000	n.a./70,000	70,000/17,500
AGC Full MS	3.00 x 10^6	3.00 x 10^6	3.00 x 10^6	5.00 x 10^5	3.00 x 10^6 (MS)/1.00 x 10^5(MS^2)
Max inject time (Full MS/MS^2)	150 ms	150 ms/200 ms	150 ms	150 ms	100 ms/100 ms
Isolation window	n.a.	n.a.	n.a.	10 Th	2 Th
Microscans	10	5	5	5	1
Capillary temperature	275 °C	275 °C	275 °C	275 °C	275 °C
S-lens RF level	80	80	50	50	50
SID [eV]	80	0/60	n.a.	LC 0/HC 20	n.a.
NCE [%]	n.a.	n.a.	10 to 30	10 to 30	25

Source CID

The source CID (SID) parameter is a DC offset (0–100 eV) that is added to the source DC offset. The source DC offset consists of three voltages: capillary DC, S-lens DC, and S-lens exit lens. The application of this DC offset by setting the source CID parameter results in collisions of the analytes inside the injection flatapole with residual gas molecules present in the source region of the instrument.

All-Ion Fragmentation

All-ion fragmentation (AIF) is a fragmentation type in which all ions generated in the source are guided through the ion optics of the mass spectrometer, accumulated in the C-trap, and sent together to the higher-energy collisional dissociation (HCD) cell for fragmentation. In this case, the quadrupole is not set to select a particular precursor but operated in RF-only pass-through mode. For the analysis of intact proteins, this is a useful method since different charge states often show different fragmentation behavior and it is not easy to predict which one works best.

Data Analysis

Full MS spectra were deconvoluted using Thermo Scientific™ Protein Deconvolution™ software version 2.0. From the intact antibody and the intact heavy chain, the spectra acquired at a resolution setting of 17,500 were deconvoluted using the ReSpect™ algorithm. High resolution spectra from the intact light chain acquired at a resolution of 140,000 and top-down spectra acquired at 70,000 resolution were deconvoluted using the Xtract algorithm. To identify glycoforms of the intact antibody and the intact heavy chain obtained after reduction, the masses were compared to the expected masses with the various combinations of commonly found glycoforms.

The top-down HCD and AIF spectra were deconvoluted using the Xtract algorithm in the Thermo Scientific™ Qual Browser™ utility. Fragment ion assignment was performed using Thermo Scientific™ ProSightPC™ software version 3.0 in single protein mode with a fragment ion tolerance of 5 ppm.

The dataset obtained from the proteolytic digest was processed with Thermo Scientific™ Proteome Discoverer™ software version 1.4, using the SEQUEST® algorithm. A three-protein-entry database was used consisting of the light chain, the heavy chain in two variants carrying either Ala or Val at position 219, and trypsin. Mass tolerances were set to 10 ppm for the precursor and 20 mmu for the fragment ions. Four variable modifications were considered: carbamidomethylation (Cys), oxidation (Met), deamidation (N, Q), Gln to pyro-Glu conversion, and N,N-dimethylation (Lys) (relevant for identification of trypsin autolysis products only).

Results and Discussion

Rituximab is an IgG1 class chimeric monoclonal antibody against the protein CD20, which consists of two light chains with 213 amino acids and two heavy chains with 451 amino acids each in length. The light and heavy chains are connected via 12 intrachain and 4 interchain disulfide linkages (Figure 1). The antibody is decorated with glycan structures attached to residue Asn^{301} of each of the two heavy chains. The composition and length of the attached glycans is quite diverse, resulting in a microheterogeneity of the molecule. The variety and relative abundance of the different glycostructures is essential for the efficiency of the antibody as a biological drug. The nomenclature of common glycans attached to antibodies are listed in Figure 2.

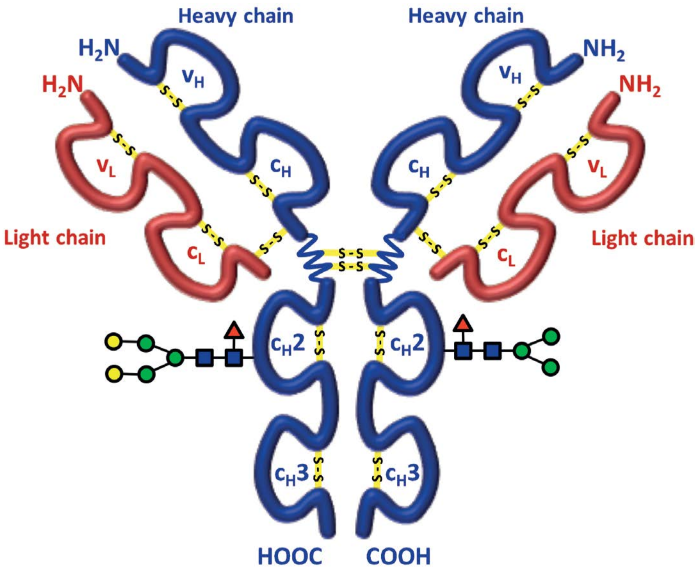


Figure 1. Schematic of molecular structure for the humanized IgG1 class monoclonal antibody rituximab

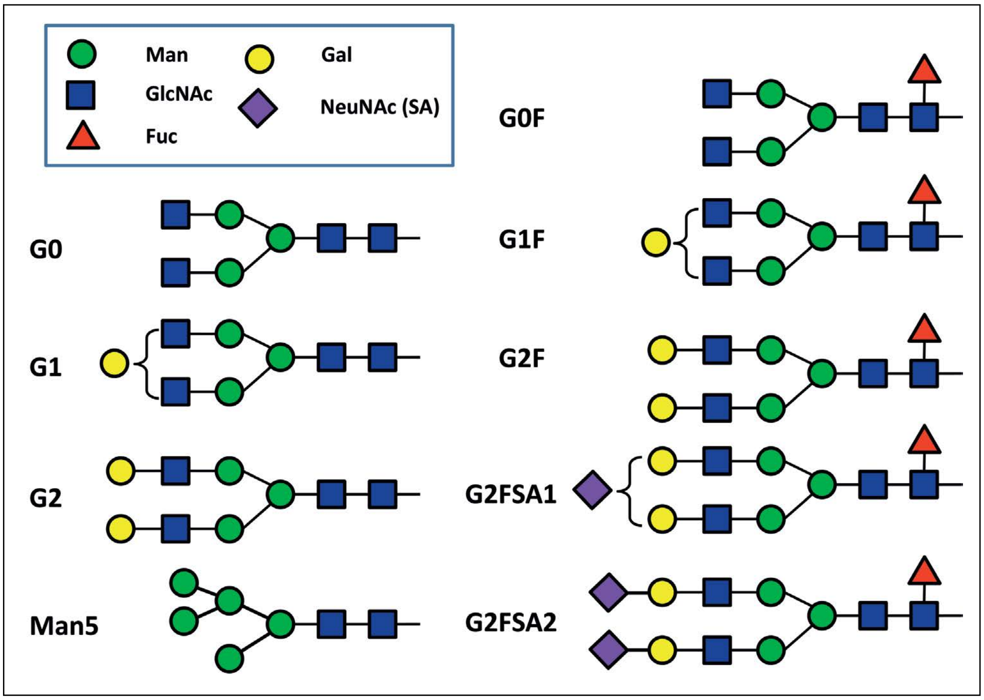


Figure 2. Nomenclature of carbohydrate structures commonly observed on antibodies

The full MS spectrum obtained from 20 ng rituximab applied to a 25 cm x 0.2 mm i.d. monolithic column is displayed in Figure 3. The mass spectrum, acquired over *m/z* 1800–5000 shows the typical charge distribution observed for large proteins. The most abundant charge state (z=+45) at *m/z* 3269, represented in the zoomed in insert, nicely pictures the four most abundant glycoforms of the intact antibody.

The intact mass of these four most abundant glycoforms and a series of less abundant glycoforms is obtained after the deconvolution of the full MS mass spectrum shown in Figure 4. The assignment of the peaks was based on the calculation of the proteins sequence, taking into account the various anticipated glycan structures shown in Figure 2.

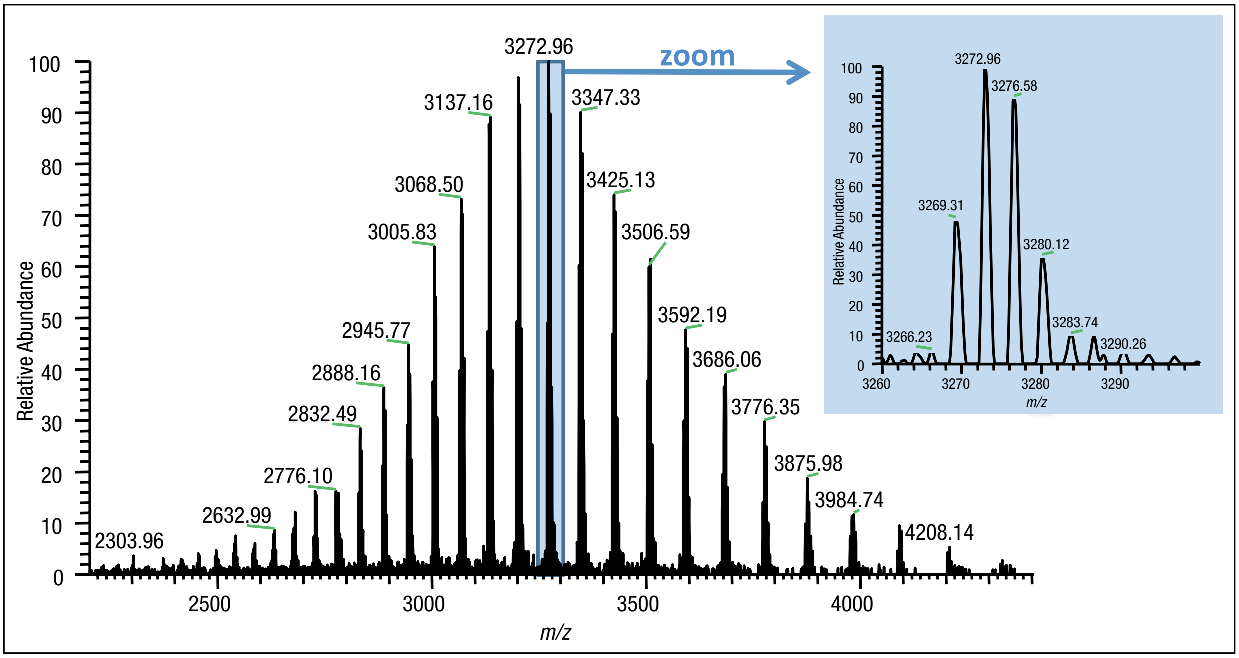
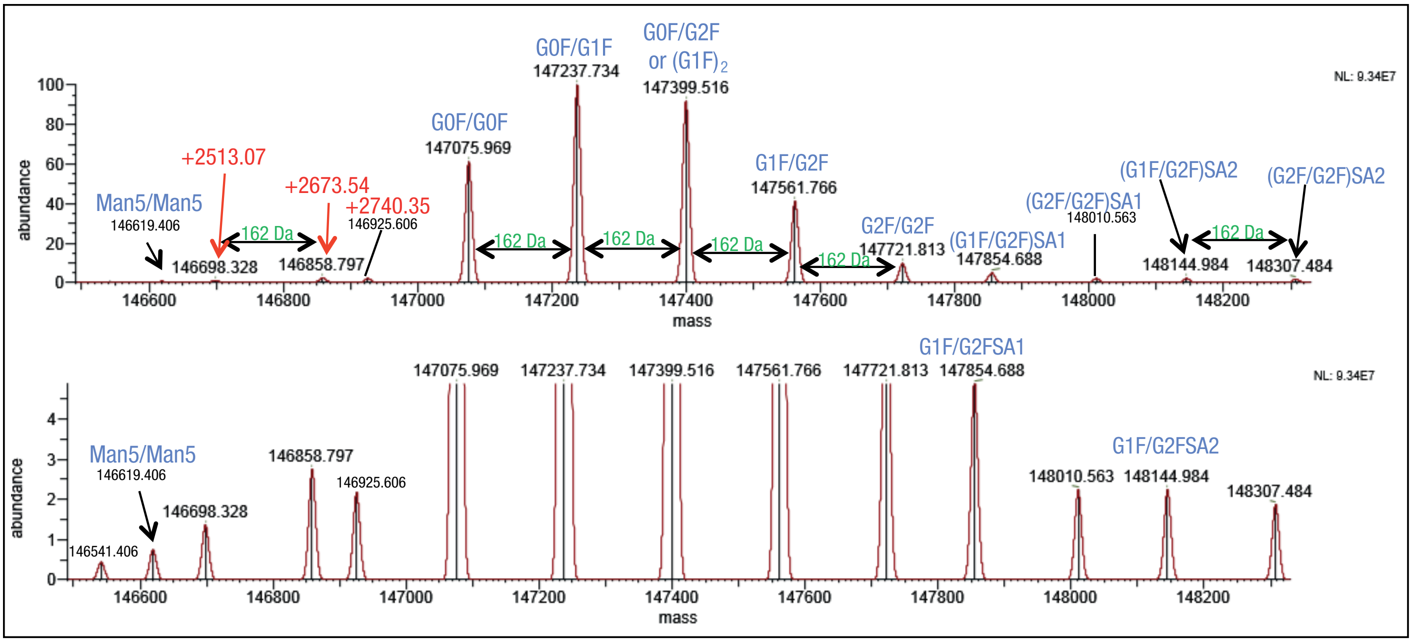


Figure 3. Single scan full MS spectrum (10 µscans) of rituximab, acquired from 10 ng sample loaded on a 250 x 0.2 mm i.d. column. The insert shows a zoomed in view of the most abundant charge state (z=+45). The observed peak pattern in the insert represents the different glycoforms of the molecule.



	G0F/G0F	G0F/G1F	G0F/G2F or (G1F)$_2$	G1F/G2F	G2F/G2F
Measured	147,075.969	147,237.734	147,399.516	147,561.766	147,721.813
Theoretical	147,074.985	147,237.126	147,399.267	147,561.408	147,723.549
Δ ppm	-6.7	-4.1	-1.7	-2.0	11.8

Figure 4. Deconvoluted mass spectrum of rituximab with annotated glycoforms (top) and comparison of theoretical and measured masses for the five most abundant glycoforms (table)

For the acquisition of the full MS spectrum of the intact antibody, the optimum setting of the source CID was evaluated (Figure 5). This setting was found to be crucial in obtaining a high quality spectrum. The application of 25–90% source CID is beneficial for most proteins. For this sample the optimum setting was 80% SID.

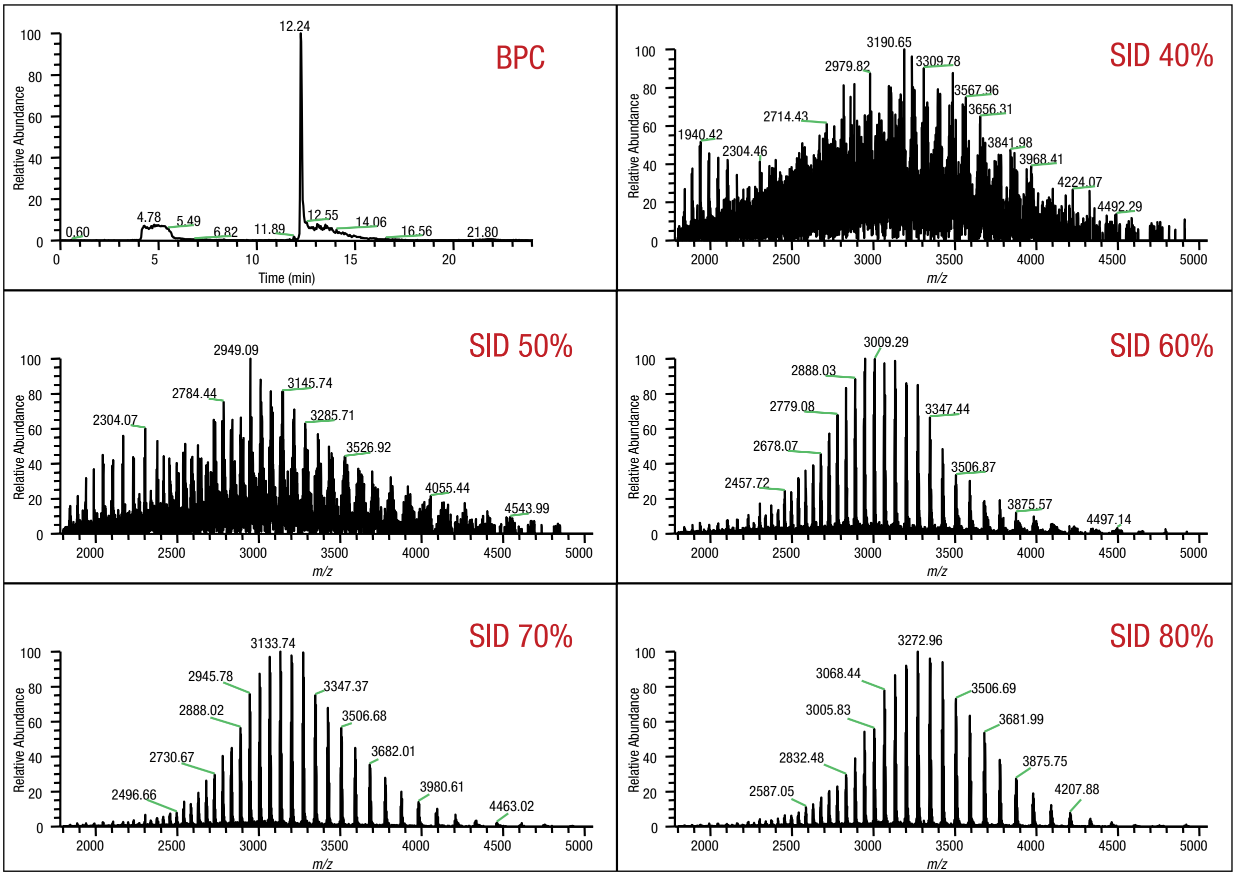


Figure 5. Full MS spectra acquired from 1 ng intact rituximab, applying increasing settings of source CID (SID)

The calculation of the masses for the light chain, unglycosylated heavy chain, and intact fully assembled antibody is presented in Table 4, showing the step-by-step calculation starting with the 213 respectively 451 amino acids of the light and heavy chain. Both protein sequences contain an N-terminal glutamine, which is anticipated to be modified to a pyro-glutamic acid, resulting in a deduction of mass of 17.0265 Da. Moreover, the C-terminal lysine present in the heavy chain is likely to be cleaved off, reducing the molecular weight by another 128.09497 Da. For assembling the intact antibody, a total of 16 disulfide linkages is considered by abstracting 32 protons. The glycan structures on each of the two heavy chains will add between 1217.1 and 2352.1 Da in mass. It has to be considered that the two chains can carry different glycans, resulting in a mixed composition, e.g. G01/G2F. Chemical composition and masses of individual carbohydrates are listed in Table 5. The monoisotopic and average atomic masses of the elements used to calculate molecular weights in Tables 4 and 5 are listed in Table 6.

Table 4. Chemical composition and step-by-step calculation of monoisotopic and average mass for the light and heavy chain, including their modifications as well as the intact antibody rituximab with various glycoforms. Detected masses shown in Figures 4, 6, and 7 are presented in the blue cells.

Elemental compositions	C	H	N	O	S	MW (monoisotopic)	MW (average)
Light chain (LC) full sequence aa 1-213	1016	1577	273	328	6	23,042.34369	23,056.5
N-terminal pyro Glutamic acid	1016	1574	272	328	6	23,025.31714	23,039.4
N-terminal pyro Glutamic acid, 2 intrachain S-S bonds	1016	1570	272	328	6	23,021.28584	23,035.4
2 x LC (N-term. pyroGlu)	2032	3148	544	656	12	46,050.63428	46,078.9
2 x LC (N-term. pyroGlu, 2 intrachain S-S bonds each)	2032	3140	544	656	12	46,042.57168	46,070.8
Heavy chain (HC) full sequence aa 1-451	2197	3389	577	676	16	49,183.40813	49,214.0
N-terminal pyro Glutamic acid	2197	3386	576	676	16	49,166.38158	49,197.0
minus C-term. K (aa 1-450)	2191	3374	574	675	16	49,038.28661	49,068.8
minus 4 intrachain S-S bonds	2191	3366	574	675	16	49,030.22401	49,060.8
HC-G0F (pyro-Glu, - K, fully reduced)	2247	3466	578	714	16	50,482.82048	50,514.2
HC-G1F (pyro-Glu, - K, fully reduced)	2253	3476	578	719	16	50,644.87330	50,676.3
HC-G2F (pyro-Glu, - K, fully reduced)	2259	3486	578	724	16	50,806.92613	50,838.5
HC minus 4 intrachain S-S bonds + G0F	2247	3458	578	714	16	50,474.75788	50,506.1
2 x HC (pyroGlu, - K)	4382	6748	1148	1350	32	98,076.57323	98,137.7
2 x HC (pyroGlu, - K, 4 intrachain S-S bonds each)	4382	6732	1148	1350	32	98,060.44803	98,121.6
Man5 (HexNAc)2 (Hex)5	46	76	2	35	0	1216.42286	1217.1
G0 (HexNAc)4 (Hex)3	50	82	4	35	0	1298.47596	1299.2
G0F (HexNAc)4 (Hex)3 Fuc	56	92	4	39	0	1444.53387	1445.3
G1 (HexNAc)4 (Hex)4	56	92	4	40	0	1460.52878	1461.3
G1F (HexNAc)4 (Hex)4 Fuc	62	102	4	44	0	1606.58669	1607.5
G2 (HexNAc)4 (Hex)5	62	102	4	45	0	1622.58161	1623.5
G2F (HexNAc)4 (Hex)5 Fuc	68	112	4	49	0	1768.63951	1769.6
G1FSA (HexNAc)4 (Hex)4 Fuc SA	73	119	5	52	0	1897.68211	1898.7
G1FSA2 (HexNAc)4 (Hex)4 Fuc (SA)2	84	136	6	60	0	2188.77752	2190.0
G2FSA (HexNAc)4 (Hex)5 Fuc SA	79	129	5	57	0	2059.73493	2060.9
G2FSA2 (HexNAc)4 (Hex)5 Fuc (SA)2	90	146	6	65	0	2350.83035	2352.1
Man5/Man5 (HexNAc)4 (Hex)10	92	152	4	70	0	2432.84572	2434.2
G0F/G0F (HexNAc)8 (Hex)6 (Fuc)2	112	184	8	78	0	2889.06774	2890.7
G0F/G1F (HexNAc)8 (Hex)7 (Fuc)2	118	194	8	83	0	3051.12056	3052.8
G1F/G1F (HexNAc)8 (Hex)8 (Fuc)2	124	204	8	88	0	3213.17338	3215.0
G1F/G2F (HexNAc)8 (Hex)9 (Fuc)2	130	214	8	93	0	3375.22621	3377.1
G2F/G2F (HexNAc)8 (Hex)10 (Fuc)2	136	224	8	98	0	3537.27903	3539.2
G1F/G2FSA (HexNAc)8 (Hex)9 (Fuc)2 SA	141	231	9	101	0	3666.32162	3668.3

Elemental compositions	C	H	N	O	S	MW (monoisotopic)	MW (average)
G1F/G2FSA2 (HexNAc)8 (Hex)9 (Fuc)2 (SA)2	152	248	10	109	0	3957.41704	3959.6
G2F/G2FSA (HexNAc)8 (Hex)10 (Fuc)2 SA	147	241	9	106	0	3828.37445	3830.5
G2F/G2FSA2 (HexNAc)8 (Hex)10 (Fuc)2 (SA)2	158	258	10	114	0	4119.46986	4121.7
Sum 2 x HC +2 x LC (4 x pyroGlu, -2K)	6414	9896	1692	2006	44	144,127.20750	144,216.6
minus 32 S-S bond protons	6414	9864	1692	2006	44	144,094.95710	144,184.3
2HC + 2LC - 16 S-S bonds + (Man5)2	6506	10016	1696	2076	44	146,527.80282	146,618.5
2HC + 2LC - 16 S-S bonds + (G0F)2	6526	10048	1700	2084	44	146,984.02484	147,075.0
2HC + 2LC - 16 S-S bonds + G0F/G1F	6532	10058	1700	2089	44	147,146.07766	147,237.1
2HC + 2LC - 16 S-S bonds + G0F/G2F or (G1F)2	6538	10068	1700	2094	44	147,308.13049	147,399.3
2HC + 2LC - 16 S-S bonds + G1F/G2F	6544	10078	1700	2099	44	147,470.18331	147,561.4
2HC + 2LC - 16 S-S bonds + G2F/G2F	6550	10088	1700	2104	44	147,632.23613	147,723.5
2HC + 2LC - 16 S-S bonds + G1F/G2F SA	6555	10095	1701	2107	44	147,761.27872	147,852.7
2HC + 2LC - 16 S-S bonds + G1F/G2F (SA)2	6566	10112	1702	2115	44	148,052.37414	148,143.9
2HC + 2LC - 16 S-S bonds + G2F/G2F SA	6561	10105	1701	2112	44	147,923.33155	148,014.8
2HC + 2LC - 16 S-S bonds + G2F/G2F (SA)2	6572	10122	1702	2120	44	148,214.42696	148,306.1

Table 5. Chemical composition and masses of monosaccharides

	Sum Formula	Monoisotopic Mass	Average Mass	C	O	N	H
Sialic Acid	$C_{11}O_8NH_{17}$	291.09542	291.3	11	8	1	17
Galactose	$C_6O_5H_{10}$	162.05282	162.1	6	5	0	10
N-Acetylglucosamine	$C_8O_5NH_{13}$	203.07937	203.2	8	5	1	13
Mannose	$C_6O_5H_{10}$	162.05282	162.1	6	5	0	10
Fucose	$C_6O_4H_{10}$	146.05791	146.1	6	4	0	10

Table 6. Monoisotopic and average atomic masses of the elements used to calculate the molecular masses in Tables 4 and 5

	Monoisotopic Mass	Average Mass
C	12.0000000	12.01074
H	1.00782503	1.00794
N	14.0030740	14.00674
O	15.9949146	15.99940
S	31.9720707	32.06608

The initial calculation based on the sequence published in the DrugBank database[2] resulted in a mass that did not match the masses obtained in our experiments. Comparison of the DrugBank sequence with a previously published sequence[3] revealed a difference in one amino acid at position 219, located in the conserved region of the heavy chain, Ala versus Val. The sequence containing the Ala at position 219 did match well with the results obtained from intact mass measurements as well as with previously reported results.[4] To further verify this, a series of additional experiments was performed.

After reducing the antibody (without alkylation), the analysis of separated light and heavy chain was performed applying different resolution settings to account for whether or not isotopic resolution can be achieved based on molecular weight. Due to the smaller molecular weight of the light chain, it is possible to obtain an isotopically resolved spectrum, whereas for the heavy chain this is not possible since it is about twice as large as the light chain. To apply different resolution settings, the method was set up in two segments (140k resolution for the scans acquired from 0 to 15.8 min, and 17.5k resolution from 15.8 to 30 min) and the gradient was optimized to achieve well-separated peaks of the light and heavy chain.

On both monolithic columns evaluated in this study (PepSwift 250 mm x 0.2 mm i.d. and ProSwift RP-10R 50 mm x 0.1 mm i.d.), the separation of the two peaks by more than 2 min was equally possible (Figure 6). The mass spectra obtained from the light chain and from the heavy chain (Figures 6b and 6c) were submitted for deconvolution. The isotopically resolved light chain spectrum was deconvoluted using the Xtract algorithm, resulting in a monoisotopic molecular weight of 23,025.3758 Da, which matches the calculated mass by 2.5 ppm. The heavy chain was deconvoluted using the ReSpect algorithm, resulting in three peaks, each of which represents one of the major glycoforms, G0F, G1F, and G2F (Figure 7).

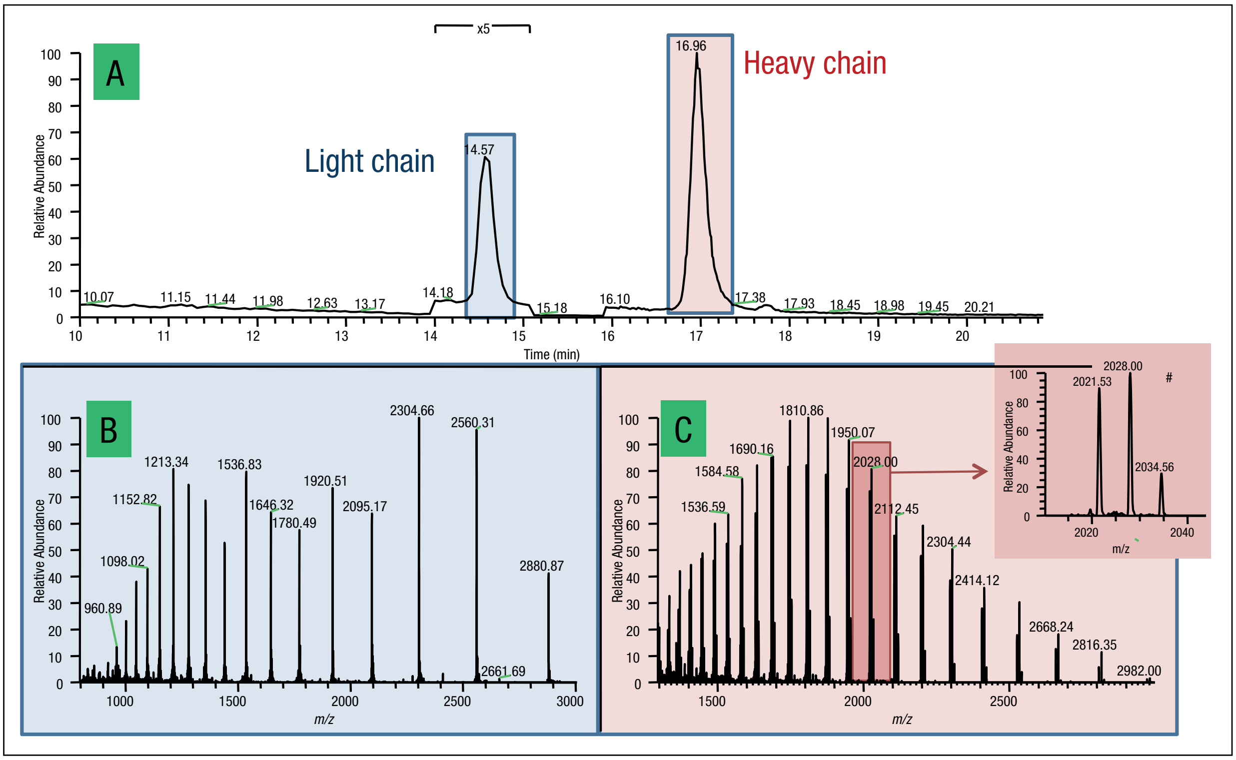
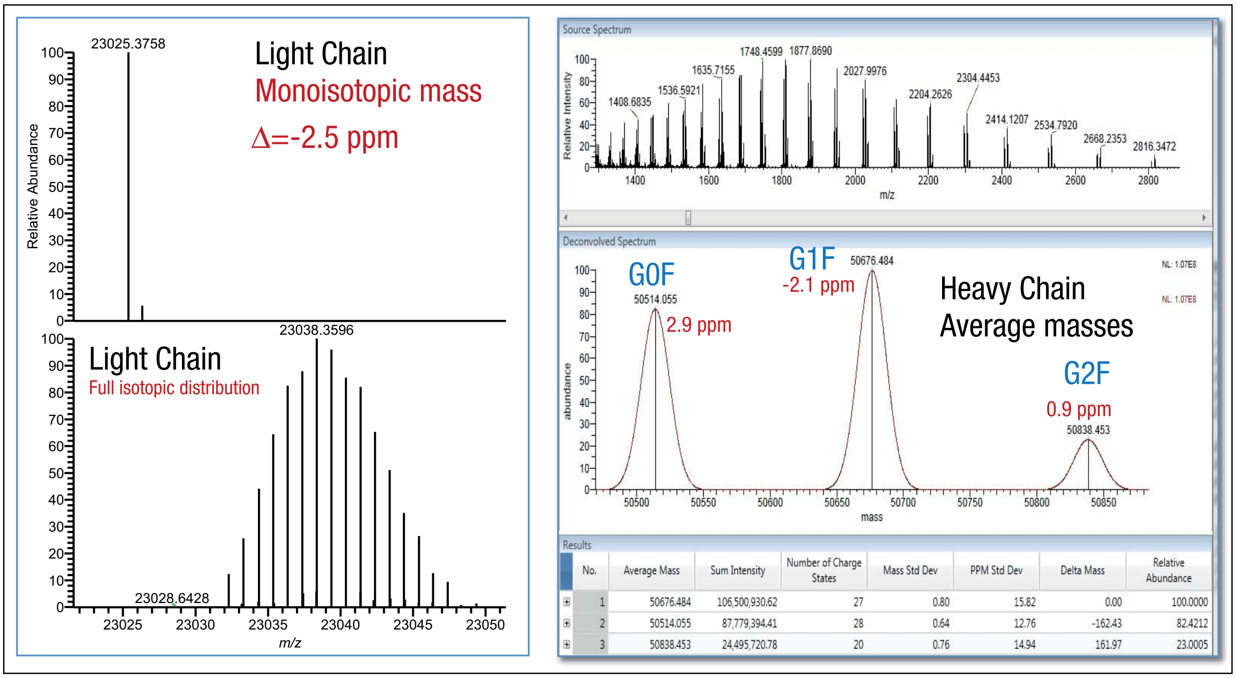


Figure 6. Chromatogram (**A**) and full MS spectra of light (**B**) and heavy chain (**C**) from reduced rituximab. The insert in panel C shows a zoomed in view of charge state z=+25, with the three peaks representing three different glycoforms.



No.	Average Mass	Sum Intensity	Number of Charge States	Mass Std Dev	PPM Std Dev	Delta Mass	Relative Abundance
1	50676.484	106,500,930.62	27	0.80	15.82	0.00	100.0000
2	50514.055	87,779,394.41	28	0.64	12.76	-162.43	82.4212
3	50838.453	24,495,720.78	20	0.76	14.94	161.97	23.0005

Figure 7. Deconvolution results of the light and heavy chain. The light chain, acquired at a resolution setting of 140,000 in full scan mode, was deconvoluted using the Xtract algorithm, obtaining an accurate monoisotopic mass as well as the full isotopic envelope (left). The heavy chain, detected at 17,500 resolution, was deconvoluted with the ReSpect algorithm providing average masses (right).

To assess the limit of detection of the instrument setup using the 250 x 0.2 mm monolithic PepSwift column, a series of LC/MS runs were acquired. Between 50 pg and 20 ng of the intact antibody was applied on column (Figure 8), starting with the lowest concentration. Two blanks were run before the sequence and between each sample to exclude carryover effects. With this setup, 500 pg was found to be the lowest amount that still achieved a good spectrum for deriving the most abundant glycoforms of the intact antibody. Here it is worth pointing out that for the lowest concentrations it was crucial to prepare the samples fresh without storing them for several hours in the autosampler prior to analysis.

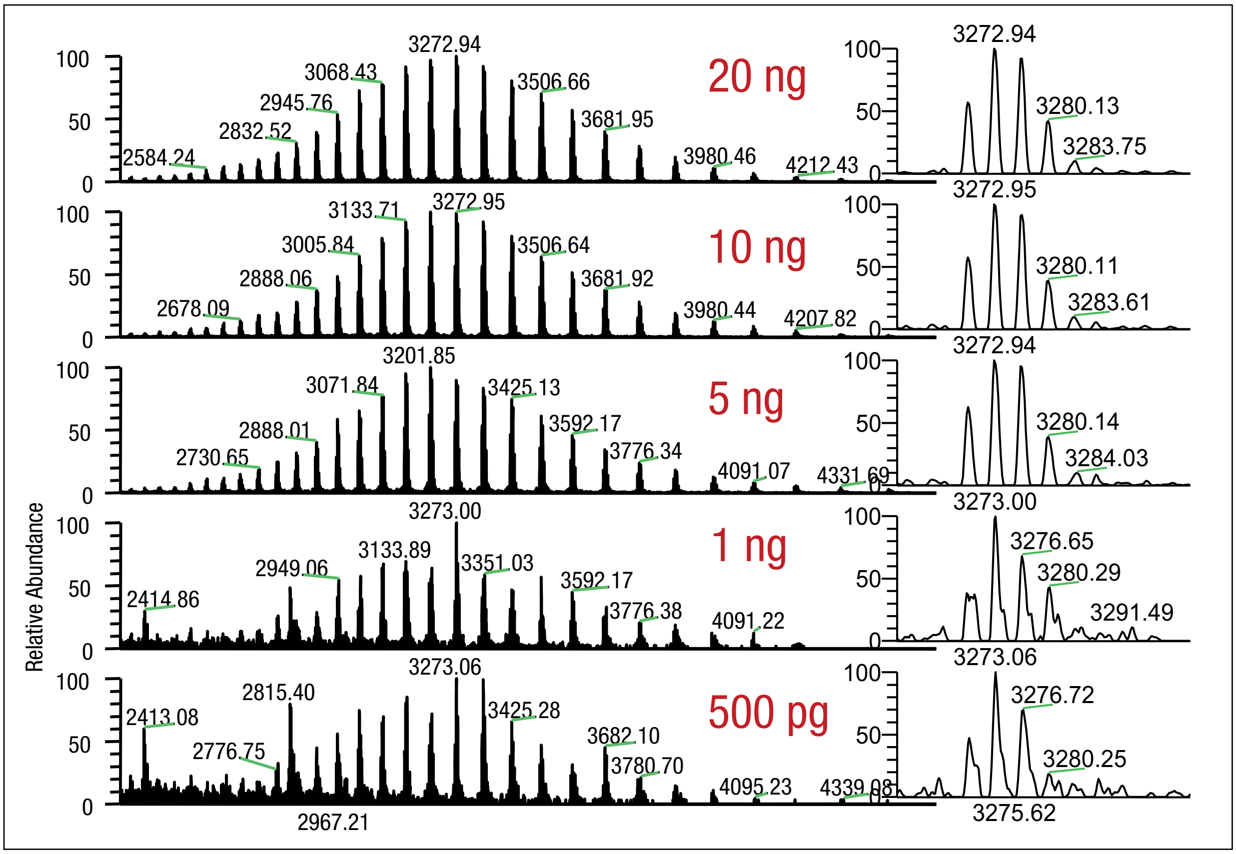


Figure 8. Full MS spectra from a dilution series of 20 ng to 500 pg of intact antibody, applied on a 250 mm x 0.2 mm i.d. monolithic PepSwift column

On the 50 mm x 1 mm i.d. monolithic ProSwift column, 30 ng and 150 ng of intact antibody were applied, both of which produced high quality spectra (Figure 9). Based on the 30 ng load it can be estimated that the lowest amount still yielding a sufficient spectrum quality to be between 5 and 10 ng.

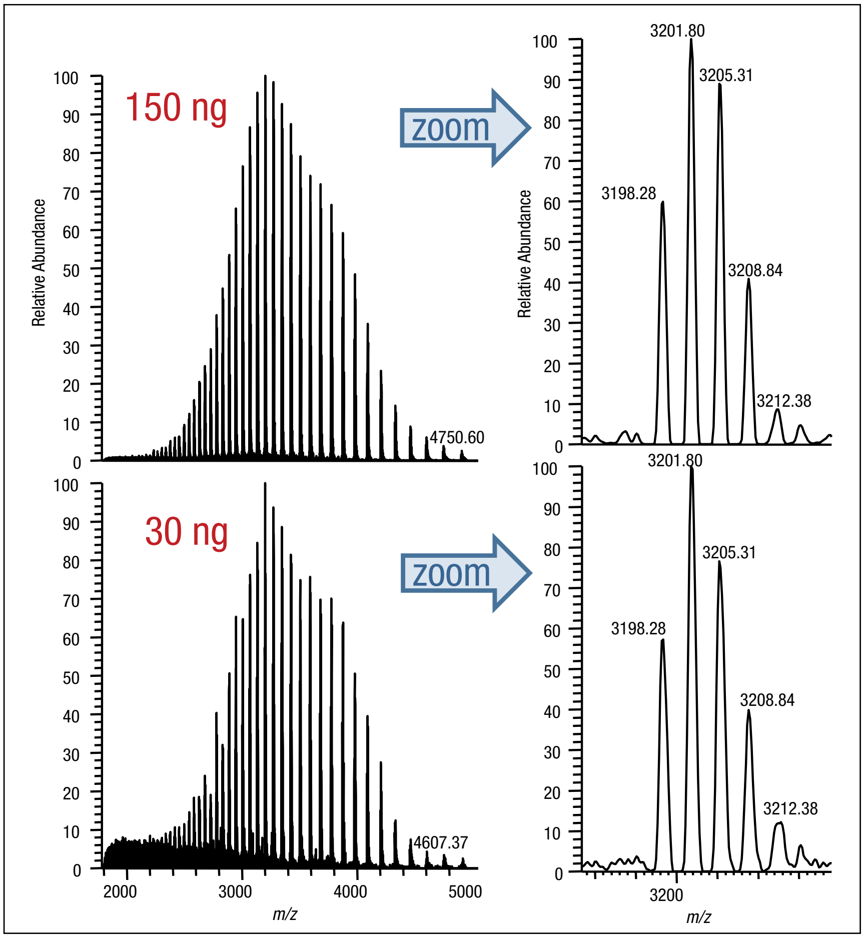


Figure 9. Full MS spectra (left) and zoomed in view of the highest abundant charge state (right) of 150 ng and 30 ng loads of intact rituximab applied on a 50 mm x 1 mm i.d. monolithic ProSwift column

In an attempt to further confirm the sequences of the light and heavy chains, two types of top-down experiments were performed: all-ion fragmentation (AIF) with fragmentation upon collision in the HCD cell and a multiplexed (5-plex), targeted MS^2 experiment on five selected charge states each of the light and heavy chains. All spectra were acquired at 70,000 resolution. For the targeted MS spectrum, a retention-time-dependent mass list was used, targeting first the earlier eluting light chain (RT 13.16 min: *m/z* 1536.96, 1646.6, 1773.3, 1920.7, 2095.4) and later the heavy chain (RT 16–20 min: *m/z* 1584.6, 1635.7, 1684.7, 1748.5, 1810.9). In this type of experiment, the first charge state listed on the inclusion list is selected and sent to the HCD cell for fragmentation. The product ions are stored in the HCD cell while the second charge state is isolated, sent to the HCD cell, fragmented, and stored in the cell until the fifth charge state has also been fragmented. All ions from the five individual isolation and fragmentation steps are sent together to the Orbitrap analyzer, resulting in one single fragment ion spectrum.

The fragment ion assignment for the light chain is displayed in Figure 10. There is good coverage of both the N- and C-terminal ends as well as some fragments in the center of the sequence, resulting in 28% coverage, respectively 15% of the theoretical fragments. For the heavy chain, fragmentation was less efficient with both methods and resulted in about 20 fragments, most of which represent the sequence termini.

To further confirm the sequences, a bottom-up approach was performed using a digest with trypsin following reduction and alkylation of the antibody. The chromatogram obtained from the digest is displayed in Figure 11. A database search against a four-entry database containing the light chain, both variants of the heavy chain, and trypsin revealed a sequence coverage of the light chain of 96% and for the heavy chain of 78.8% (Figure 12). The two short missing peptides from the light chain (LEIK and EAK) could be detected as intact masses only in the full MS spectra, whereas the peptide EAK was identified based on the accurate mass corresponding to the peptide containing a missed cleavage **EAK**VQWK. Taking into account the peptides identified based on MS/MS spectra and based on accurate masses of the small intact peptides, sequence coverage for the light chain is 100%.

Figure 10. Matched sequence coverage of the rituximab light chain based on fragment ions obtained from AIF experiments. Seventeen b- and 50 y-ions were assigned, corresponding to 15.7% of the theoretical number of fragments (67 of 426).

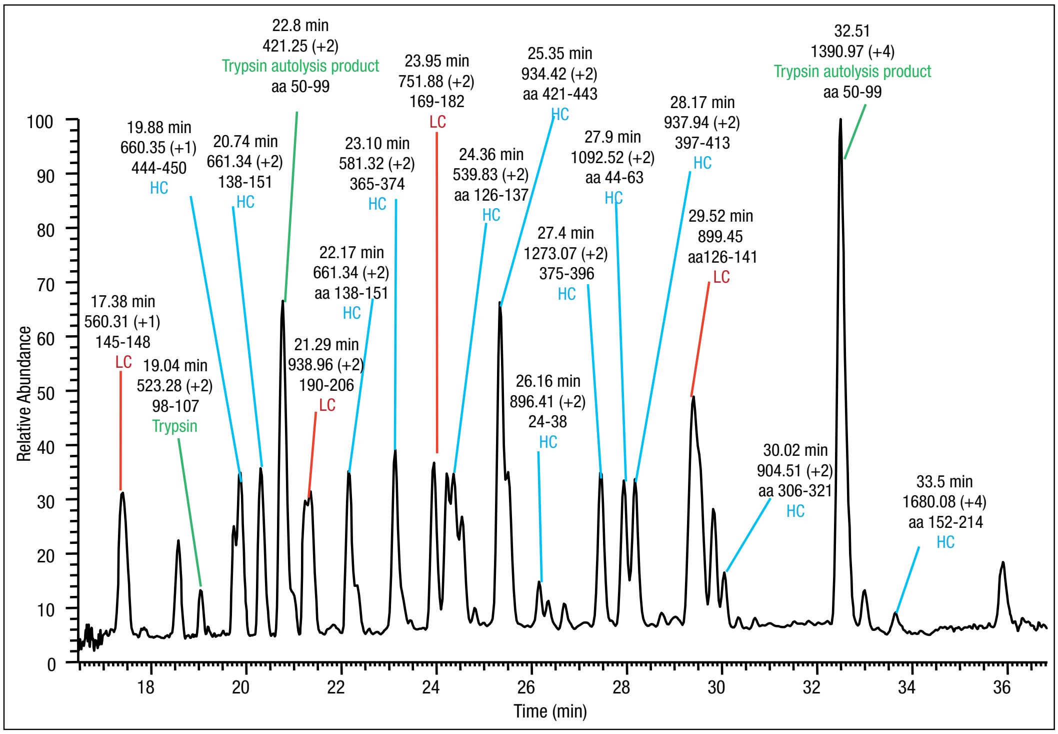
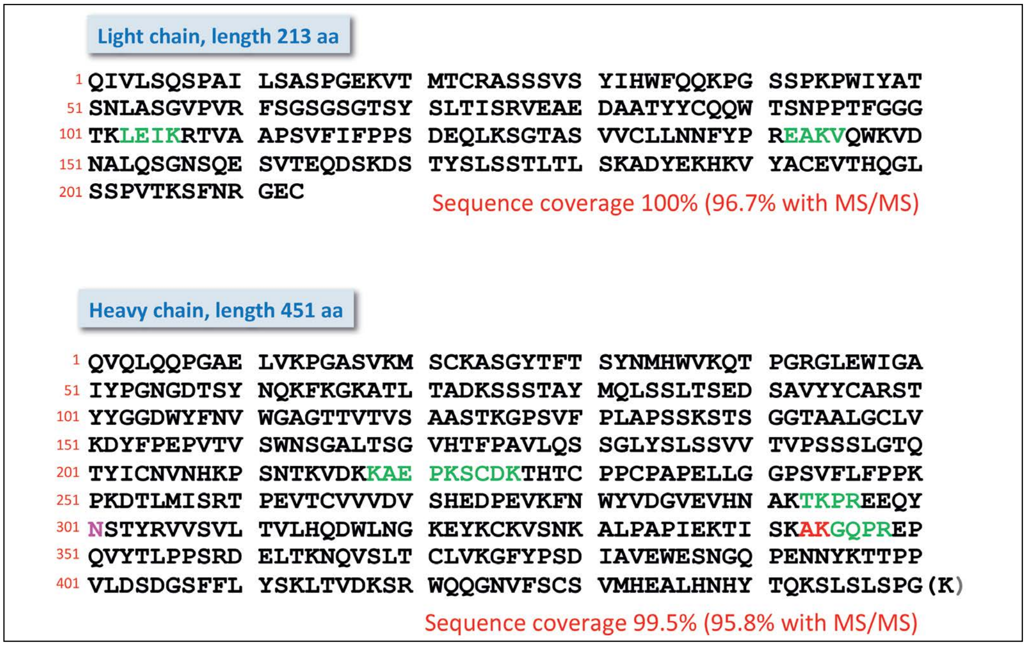


Figure 11. Base peak chromatogram of a digest using trypsin on the reduced and alkylated antibody rituximab



Light chain, length 213 aa

```
  1 QIVLSQSPAI LSASPGEKVT MTCRASSSVS YIHWFQQKPG SSPKPWIYAT
 51 SNLASGVPVR FSGSGSGTSY SLTISRVEAE DAATYYCQQW TSNPPTFGGG
101 TKLEIKRTVA APSVFIFPPS DEQLKSGTAS VVCLLNNFYP REAKVQWKVD
151 NALQSGNSQE SVTEQDSKDS TYSLSSTLTL SKADYEKHKV YACEVTHQGL
201 SSPVTKSFNR GEC
```

Sequence coverage 100% (96.7% with MS/MS)

Heavy chain, length 451 aa

```
  1 QVQLQQPGAE LVKPGASVKM SCKASGYTFT SYNMHWVKQT PGRGLEWIGA
 51 IYPGNGDTSY NQKFKGKATL TADKSSSTAY MQLSSLTSED SAVYYCARST
101 YYGGDWYFNV WGAGTTVTVS AASTKGPSVF PLAPSSKSTS GGTAALGCLV
151 KDYFPEPVTV SWNSGALTSG VHTFPAVLQS SGLYSLSSVV TVPSSSLGTQ
201 TYICNVNHKP SNTKVDKKAE PKSCDKTHTC PPCPAPELLG GPSVFLFPPK
251 PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY
301 NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP
351 QVYTLPPSRD ELTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP
401 VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG(K)
```

Sequence coverage 99.5% (95.8% with MS/MS)

Figure 12. Amino sequence of light and heavy chains from rituximab. Amino acids shown in black letters represent the parts identified based on MS/MS spectra. Sequences confirmed based on MS full scan data as intact peptides only are shown in green. The two amino acids shown in red (AK) as part of the heavy chain could neither be covered on the MS nor on the MS/MS level. Resulting sequence coverage for the light chain is 100% (96% with MS/MS) and 99.5% (98.8% with MS/MS) for the heavy chain. Asparagin[251] in the heavy chain represents the glycosylation site.

For the heavy chain, the peptide GQPR was also identified based on the accurate mass of the intact peptide. Lastly, the peptide containing the glycosylation site at position Asn[301] was not identified in its unglycosylated form based on an MS/MS spectrum. A database search including the expected glycans as modifications was successful. In addition, the glycopeptides can easily be detected in the full scan spectra in different glycosylated forms and in different charge states, and the MS/MS spectra can easily be spotted due to the presence of a characteristic peak pattern. The G0F-containing peptide is shown as an example in Figure 13, representing the intact precursor and the typical fragmentation pattern obtained from glycopeptides using HCD-type fragmentation: the two hexonium ions at mass 204 (HexNAc) and 366 (Hex-HexNAc) as well as the fragment ions nicely showing the sequence ladder of released hexose (*m/z* 162), N-acetylhexosamine (203), and Fucose (146). Considering all peptides on the MS full scan level and based on MS/MS spectra via database searches, the sequence coverage of the heavy chain is 99.5%, leaving only two amino acids not covered (aa 343-344).

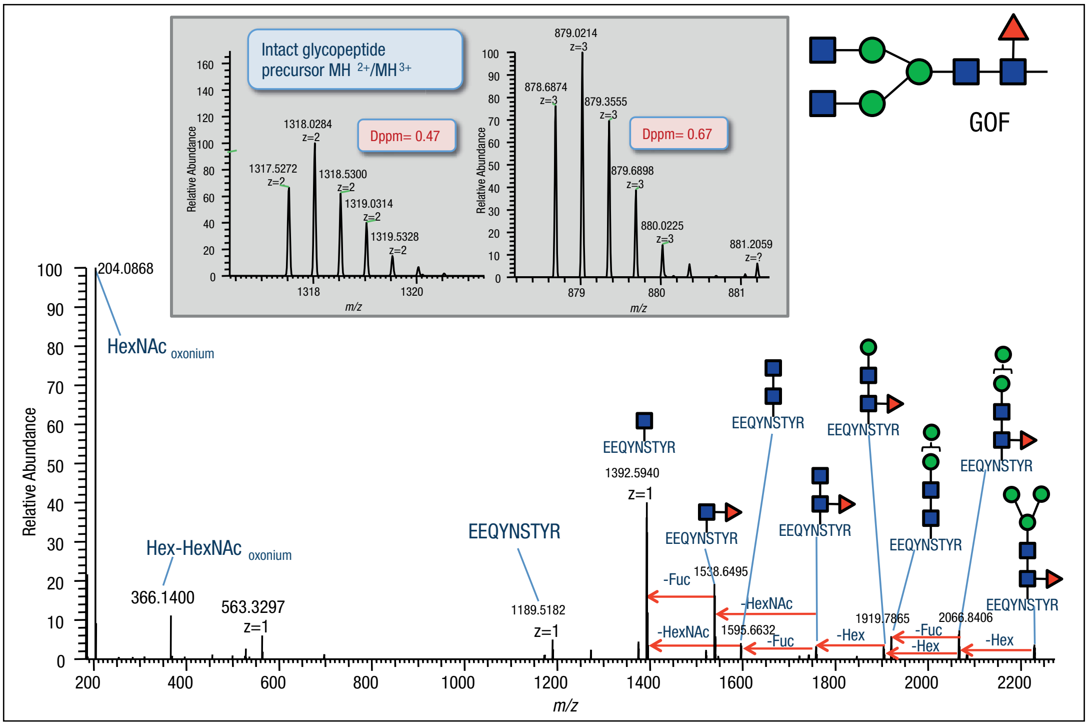


Figure 13. MS/MS spectrum of the glycopeptide aa 297-305 (EEQYN*STYR, *=G0F) obtained from the triply charged glycopeptide precursor. Inserts show the isotope patterns of doubly and triply charged intact precursors detected in the full scan spectrum.

Conclusion

In this study, a workflow is presented that combines fast chromatography, using two sizes of monolithic columns, and high resolution Orbitrap mass spectrometry of intact, as well as reduced, rituximab, sequence verification by AIF, and multiplexed HCD top-down fragmentation, supplemented by a bottom-up approach.

The data presented here also demonstrate the sensitivity of the applied LC-MS instrument setup, still obtaining a good quality MS spectrum from as low as 500 pg of the intact antibody loaded on column. Furthermore, for the analysis of the reduced mAb, a chromatographic separation of the light and heavy chains was achieved allowing for their detection at different resolution settings.

The data obtained from this workflow allow the determination of the molecular weight of the intact antibody, the confirmation/verification of the amino acid sequence of light and heavy chain, and the identification and evaluation of the relative abundance of various glycoforms of rituximab.

Acknowledgements

The authors would like to thank Daniel Pürstinger for his help in sample preparation and Remco Swart for providing the PepSwift and ProSwift monolithic columns used in this study. The financial support by the Austrian Federal Ministry of Economy, Family, and Youth and the National Foundation of Research, Technology, and Development is gratefully acknowledged.

References

1. Premstaller, A.; Oberacher, H. and Huber, C.G. High-Performance Liquid Chromatography-Electrospray Ionization Mass Spectrometry of Single- and Double Stranded Nucleic Acids Using Monolithic Capillary Columns. *Anal. Chem*. **2000**, *72*, 4386-4393.
2. http://www.drugbank.ca/drugs/DB00073
3. Nebija, D.; Kopelent-Frank, H.; Urban, E.; Noe, C. R. and Lachmann, B. Comparison of two-dimensional gel electrophoresis patterns and MALDI-TOF MS analysis of therapeutic recombinant monoclonal antibodies trastuzumab and rituximab. *Journal of Pharmaceutical and Biomedical Analysis*, **2011**, *56*, 684–91.
4. Kuribayashi, R.; Hashii, N.; Harazono, A. and Kawasaki, **N.** Rapid evaluation for heterogeneities in monoclonal antibodies by liquid chromatography/mass spectrometry with a column-switching system. *Journal of Pharmaceutical and Biomedical Analysis,* **2012**, *67-68*, 1–9.

www.thermofisher.com

Africa	+43 1 333 50 34 0
Australia	+61 3 9757 4300
Austria	+43 810 282 206
Belgium	+32 53 73 42 41
Canada	+1 800 530 8447
China	800 810 5118 (free call domestic) 400 650 5118
Denmark	+45 70 23 62 60
Europe-Other	+43 1 333 50 34 0
Finland	+358 9 3291 0200
France	+33 1 60 92 48 00
Germany	+49 6103 408 1014
India	+91 22 6742 9494
Italy	+39 02 950 591
Japan	+81 45 453 9100
Latin America	+1 561 688 8700
Middle East	+43 1 333 50 34 0
Netherlands	+31 76 579 55 55
New Zealand	+64 9 980 6700
Norway	+46 8 556 468 00
Russia/CIS	+43 1 333 50 34 0
Singapore	+65 6289 1190
Spain	+34 914 845 965
Sweden	+46 8 556 468 00
Switzerland	+41 61 716 77 00
UK	+44 1442 233555
USA	+1 800 532 4752

Thermo SCIENTIFIC

Part of Thermo Fisher Scientific