Characterization of a multi-specific antibody therapeutic with peptide mapping

Peptide mapping analysis of a trispecific antibody therapeutic with the SCIEX X500B QTOF System and Protein Metrics Inc. software

Armelle Martelet[1], Severine Clavier[1], Bruno Genet[1], Zoe Zhang[2], Kerstin Pohl[2]
[1]SANOFI, Vitry, FRANCE; [2]SCIEX, USA

Here, the peptide mapping analysis of a trispecific monoclonal antibody (mAb) including sequence coverage assessment as well as the identification and localization of various post-translational modifications (PTMs) is shown.

Antibody-based derivatives have received increased attention as therapeutics in biopharmaceutical development because of their unique capacity to engage multiple molecular targets[1], engage immune effector cells[2] or even penetrate tissues previously inaccessible to conventional mAbs[3]. In order to address complex disease pathways more precisely (by involving multiple targets, for example), new modalities such as multi-specific antibodies have become a promising therapeutic platform. These molecules can be more complex and their development requires increased attention to ensure safety and efficacy. Liquid chromatography coupled to mass spectrometry (LC-MS) is frequently used as a characterization tool. However, it can be challenged by the additional structural complexity of multi-specific antibodies, due to different domains and PTMs such as glycosylation. To address this analytical challenge, intact and subunit mass analysis of a trispecific mAb were performed previously[4]. Here, the characterization is complemented by a comprehensive peptide mapping analysis. This method is suitable as a general assay for the characterization of complex biotherapeutics as well as further attribute monitoring in both, upstream and downstream development.

Key features of characterization of a multi-specific antibody

- Excellent data quality for increased confidence in the results for a wide range of biopharmaceutical applications with a benchtop, high resolution TOF-MS
- Easy to use, yet powerful tools are provided for expert and non-MS trained scientists by a newly designed MS system
- Direct compatibility of raw data files with state-of-the-art data processing tools from Protein Metrics Inc., especially suited for complex biotherapeutics characterization
- A generic platform-based method suitable to characterize biopharmaceuticals from standard proteins to complex new modalities

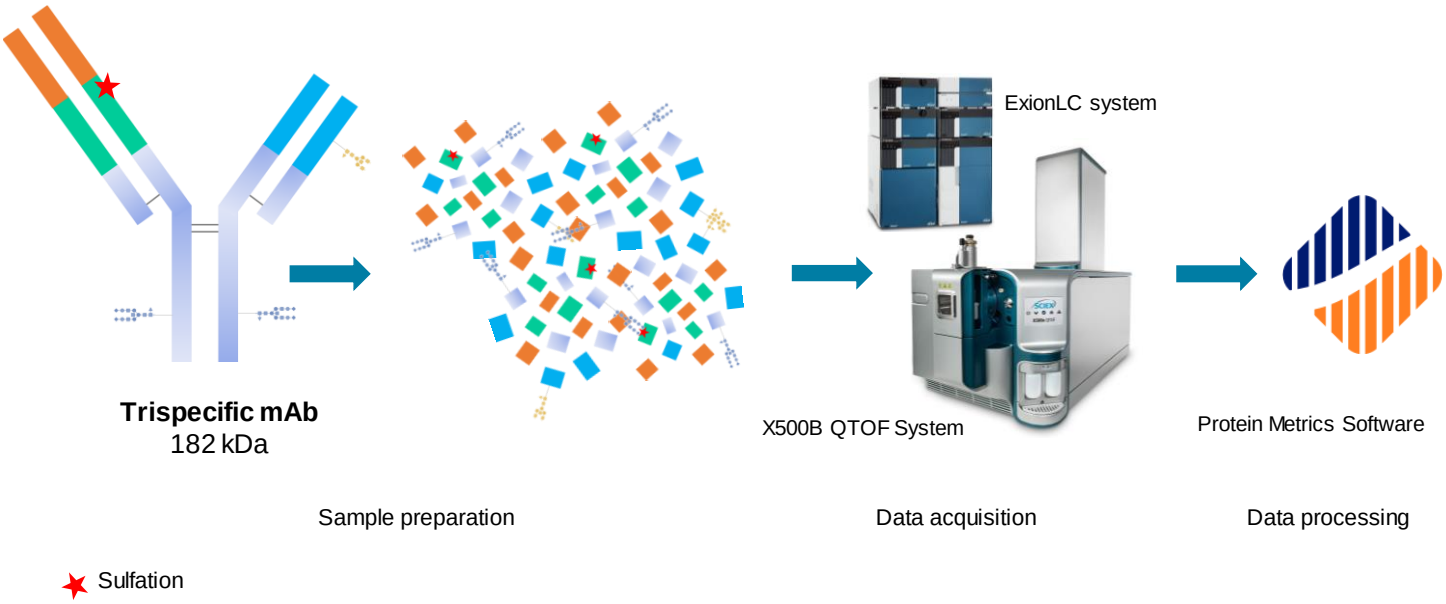

Figure 1. Overview of workflow for characterization of the multi-specific mAb. The sample was reduced and enzymatically digested into peptides. The digest was analyzed using an ExionLC System coupled to an X500B QTOF System. Raw data were analyzed using Byos software from Protein Metrics Inc.

Methods

Sample preparation: A multi-specific antibody (100 µg) was denatured with 7.0 M guanidine-HCl, 50 mM Tris, pH 7.5. The denatured protein was subjected to reduction with 10 mM DTT at room temperature for 30 minutes, followed by alkylation with iodoacetamide at 20 mM for 20 minutes, in the dark, at room temperature. The alkylation was quenched with 4 µL of 50 mM DTT. The guanidine-HCl was then diluted to a final concentration of 1 M. Half of the sample was digested overnight with trypsin/LysC (Promega, sequence grade) at a ratio of 1:10 (protein/enzyme) (*w:w*) and the other half was digested overnight with Asp-N (Promega, sequence grade) at a ratio of 1:25 (protein/enzyme) (*w:w*) at 37 °C.

Chromatography: Separation of peptides was accomplished using an ExionLC system fitted with a Waters CSH C18 column (150×2.1 mm, 1.7 µm, 130 Å) at 50°C using the gradient shown in Table 1. Mobile phase A consisted of 0.1% formic acid in water while mobile phase B consisted of 0.1% formic acid in acetonitrile. 2 µg of tryptic sample and 4 µg of AspN-digested sample were injected.

Table 1. Gradient information.

Time (min)	Flow rate (mL/min)	%A	%B
Initial	*0.3*	*99*	*1.0*
5.0	*0.3*	*99*	*1.0*
6.0	*0.3*	*90*	*10*
25.0	*0.3*	*75*	*25*
65.0	*0.3*	*60*	*40*
70.0	*0.3*	*40*	*60*
70.5	*0.3*	*10*	*90*
74.0	*0.3*	*10*	*90*
74.1	*0.3*	*99*	*1.0*
75.0	*0.3*	*99*	*1.0*
75.1	*0.3*	*10*	*90*
79.0	*0.3*	*10*	*90*
79.1	*0.3*	*99*	*1.0*
83.0	*0.3*	*99*	*1.0*

Mass spectrometry : The sample was analyzed with a SCIEX X500B QTOF System fitted with a Turbo V™ Ion Source with Twin Sprayer probe using information-dependent acquisition (IDA) for the 8 most abundant ions (Table 3).

Table 2. Mass spectrometry parameters.

Parameter	MS	MS/MS
Scan mode	*TOFMS*	*IDA dependent*
Gas 1	*60 psi*	
Gas 2	*60 psi*	
Curtain gas	*55 psi*	
Source temperature	*350°C*/425°C***	
Ion spray voltage	*5500 V*	
Declustering potential	*20 V*/80 V***	
Collision energy	*7 V*	*dynamic*
CAD gas	*6*/8***	
Maximum candidate ion	*8*	
Intensity threshold	*100 cps*	
Charge states	*1 to 6*	
Exclusion time	*6 s after 2 occurrences*	
Start mass	*200 m/z*	*100 m/z*
Stop mass	*1,800 m/z*	*1,500 m/z*
Accumulation time	*0.25 s*	*0.04 s*
Time bins to sum	*4*	*6*

** tryptic digest*
*** AspN digest*

Data processing: Data were processed using the Protein Metrics Inc. Byos software version 3.9.

Verifying the sequence

The multi-specific antibody contains four chains, with each chain consisting of a different primary amino acid (aa) sequence (Figure 3). Furthermore, the complex mAb showed three *N*-linked glycosylation sites in HC1, HC2 and LC2 and a modification site (indicated by the asterisk in Figure 3), which could either be linked to a phosphorylation (+79.9663 Da) or a sulfation (+79.9568 Da) based on previous subunit results.[4]

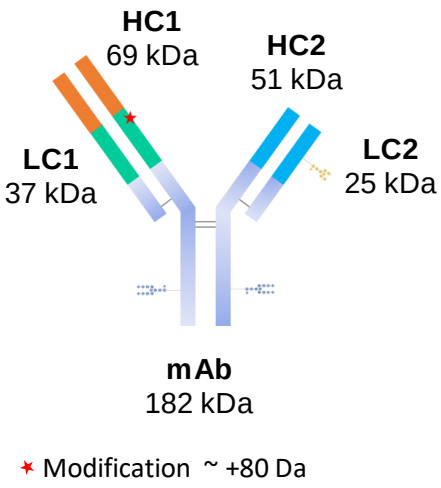

Figure 3. Overview of trispecific mAb.

To fully evaluate molecular integrity and composition of the multi-specific antibody, a comprehensive peptide mapping analysis was performed. The confirmation of the sequence is of particular importance, as part of the characterization throughout the development to ensure the safety and efficacy of the molecule. The multi-specific antibody was digested with trypsin/Lys-C and Asp-N, respectively. With each preparation alone a high sequence coverage > 86 % was obtained, however, only the combination of results led to a full confirmation of all sequences by MS/MS data (Table 4). Since tryptic digestion alone can result in very short peptides, that may not reliably retain using reversed phase conditions or may not provide information-rich fragmentation, a combination of enzymes is usually key for a full characterization and full confidence in the correct sequence.

Table 4. Sequence coverage* for trispecific mAb.

Chain	Trypsin/Lys-C	AspN	Combined
LC1	*(327 of 338) 96.75%*	*(332 of 338) 98.22%*	*(338 of 338) 100%*
LC2	*(203 of 210) 96.67%*	*(210 of 210) 100%*	*(210 of 210) 100%*
HC1	*(531 of 611) 86.91%*	*(587 of 611) 96.07%*	*(611 of 611) 100%*
HC2	*(419 of 450) 93.11%*	*(426 of 450) 94.67%*	*(450 of 450) 100%*

**based on MS/MS information*

Protein termini and their modifications

To evaluate the results from this peptide mapping platform assay, all N-terminal peptides derived from the four chains were used as examples (Figure 4). For the four peptides, great MS/MS sequence coverages (70-92 %) were obtained without further optimization of the MS parameters using a data dependent acquisition approach and collision-induced fragmentation (CID). For instance, the MS/MS spectrum shown in Figure 4D provided a 92% aa coverage of the HC2 N-terminal peptide with the confirmation of N-terminal pyroglutamate modification verified by a high intensity b-ion series (blue hashmarks). The N-terminal peptide of the HC1 derived from a tryptic/Lys-C digest only consists of three residues, which was not retained well with reversed-phase conditions and is therefore prone to be missed for unambiguous identification and for obtaining MS/MS information. In contrast, the N-terminal peptide from the Asp-N sample preparation provides a reasonable length for analysis and resulted in good MS/MS coverage (Figure 4C).

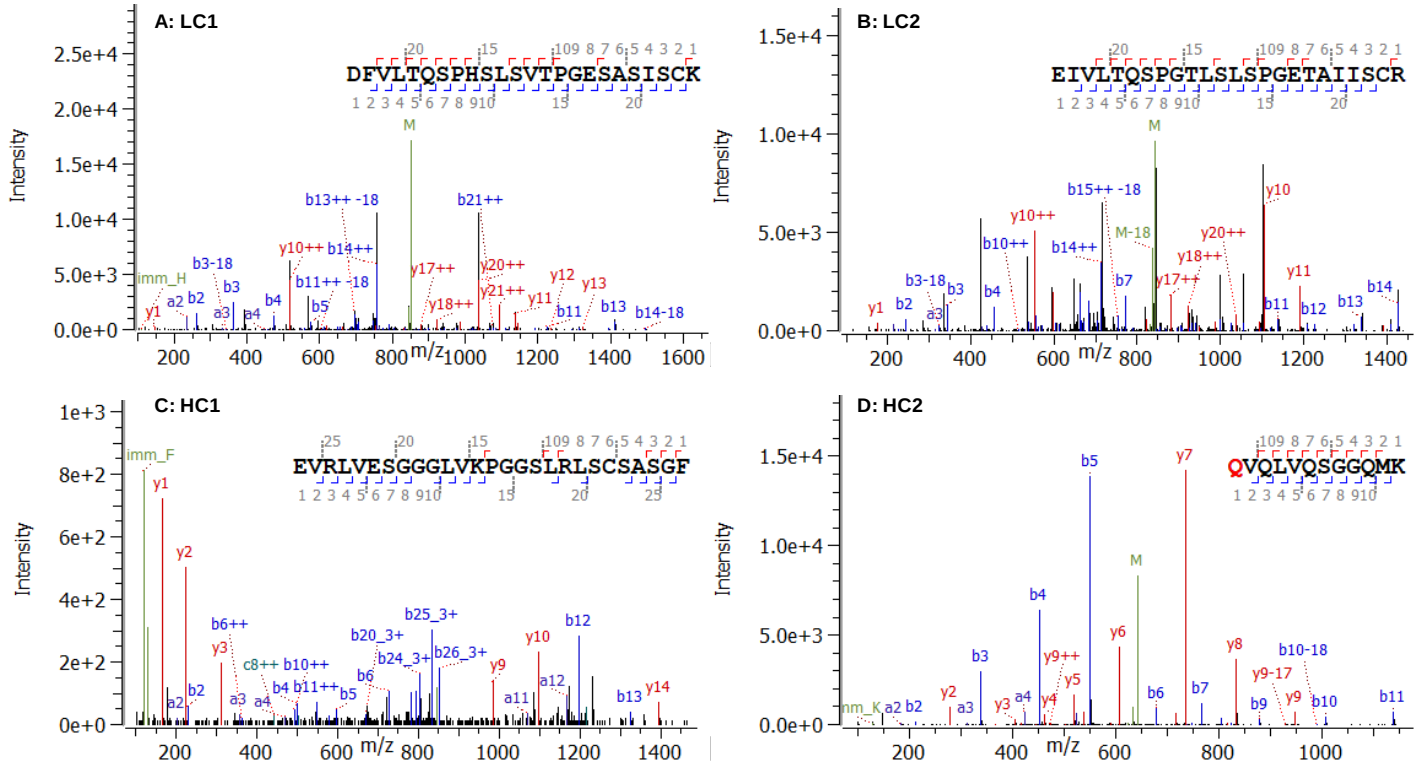

Figure 4: MS/MS spectra of four N-terminal peptides. Precursor and precursor derived ions are indicated in green, *y*-ions in red and *b*-ions in blue. Great MS/MS sequence coverages were obtained for all termini: 95.7% for LC1, 95.7 % for LC2, 69.2 % for HC1and 90.9 % for HC2. A, B, D derived from trypsin/Lys-C digest; C from Asp-N digest.

Understanding deamidations

Besides confirming the sequence, the identification and localization of different PTMs which can impact the biomolecule's efficacy or safety are additional objectives during its development, by leveraging peptide mapping analysis. One example is shown in Figure 5, which represents the extracted ion chromatograms (XIC) along with MS/MS spectra of the modified and the unmodified versions of a peptide. Three distinct chromatographic peaks were detected which were linked to the unmodified peptide (main peak at Rt = 28.2 min, Figure 5A) and two modifications (peaks at Rt = 28.7, Figure 5B and 29 min, Figure 5C). Since the sequence of the peptide with a total length of 22 AA contains four potential deamidation sites (three asparagine N and one glutamine Q), MS/MS information is required to understand the exact position of the modification. The correct assignment of deamidations quite often poses a challenge since chromatographic baseline separation of modified and unmodified peptides and peptides with different sites of modification can be difficult to achieve. The small mass shift caused by the modification (+ 0.9840 Da) leads to a partial overlap of the isotopic patterns of modified and unmodified peptides complicating the assignment already on MS1 level. In addition, chimeric MS/MS spectra of modified peptides can be a result of partially co-eluting modified peptides. Furthermore, the usually low abundance of modified peptides might lead to missing MS/MS information.

However, here, excellent quality MS/MS data were acquired with an automated IDA approach, enabling the localization of modifications via data interpretation by Protein Metrics Inc. software: The main peak could be assigned to the unmodified peptide proven by assigning all abundant fragment ions and achieving a great MS/MS sequence coverage (Figure 5A). The subsequently eluting low abundant peaks were linked to deamidation events within the same peptide. Both low abundant peaks could be assigned to a deamidation event at N14. The AA motif NG poses a hotspot for deamidation of the asparagine and is known to appear as two peaks, which can be differentiated by retention time.[6] The later eluting peak can be assigned to the deamidated species of asparagine resulting in an aspartic acid (Asp, Figure 5C). The earlier eluting modified peptide can be linked to isomerization of aspartic acid (isoAsp, Figure 5B).

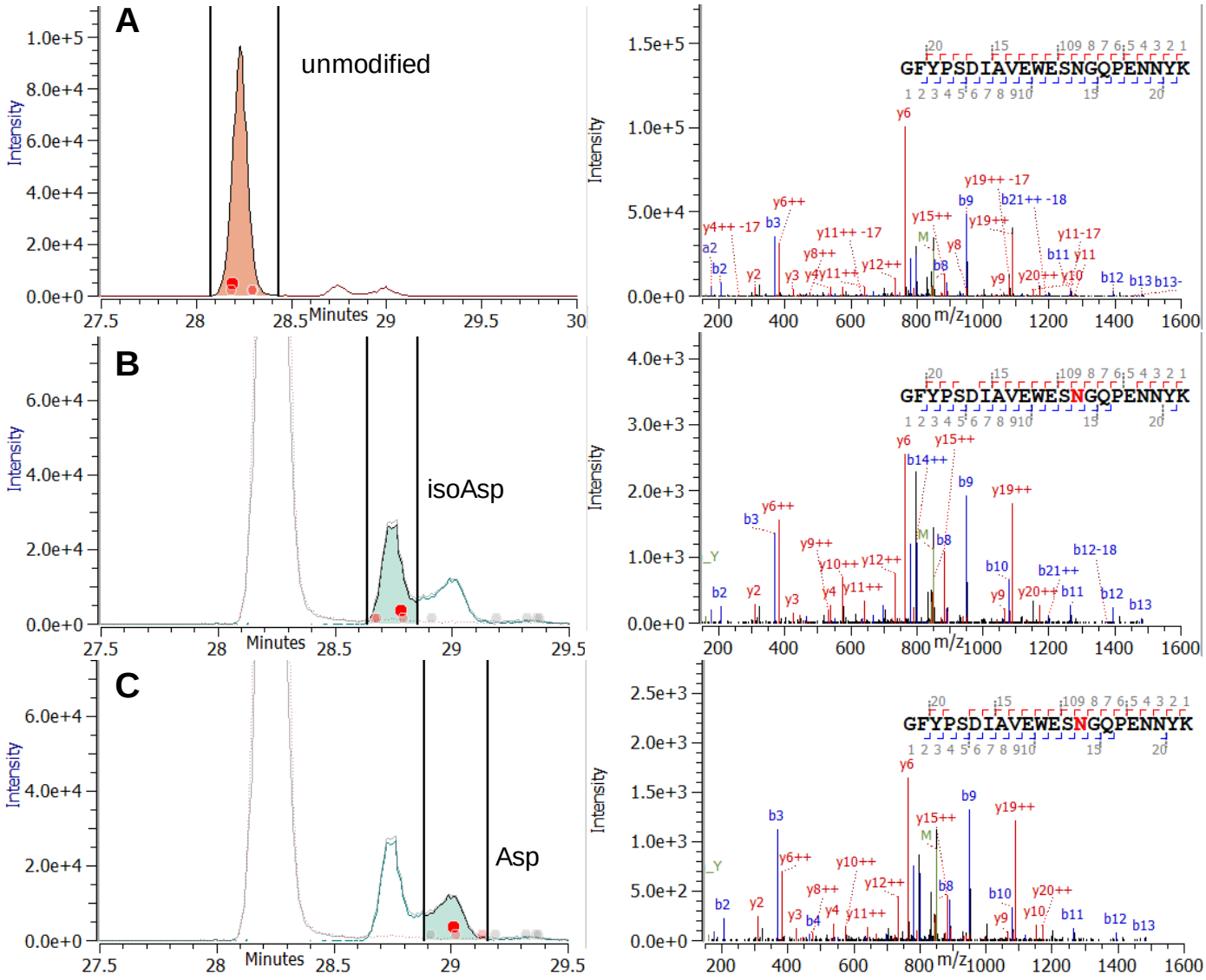

Figure 5. Data of the unmodified and deamidated peptide GFYPSDIAVEWESNGQPENNYK. XICs for the respective peptide with red dots highlighting the time points of MS/MS data acquisition (left hand side) and MS/MS spectra with sequence coverage for peptides respectively (right hand side) are shown. The spectra provide high sequence coverage as well as the localization of the modification from four potential sites within the peptide. A:unmodified peptide, B: deamidation on N14 with isoAsp formation, C: deamidation on N14.

Phosphorylation or sulfation?

A particularly interesting case was found on the HC1 of the trispecific mAb. The HC1 as well as the intact mAb showed a mass shift of ~ 80 Da in previous studies.[4] Intact or subunit analysis alone cannot reveal if this mass shift is linked to phosphorylation (+79.9663 Da) or a sulfation (+79.9568 Da) since the mass difference between both types of modification is too small compared to the overall mass observed using an intact or subunit approach. Applying peptide mapping analysis, the peptide containing the modification was detected as a singly-charged peptide with five residues. The data were processed allowing for phosphorylation and sulfation as potential modifications. The software identified that the given peptide contained a sulfation (Figure 6). The nature of the modification is justified by the precursor mass accuracy, the availability of fragment ions and the ability to determine how well alternative assignments could be made. The best fit assignment is therefore determined and the spectrum annotated accordingly, resulting in the assignment of a sulfation, while ruling out phosphorylation by the software. Alternative matches, however, can still be accessed by a user (Figure 7). Compared to a sulfation, a

phosphorylation showed a significantly higher mass error and lower score (Figure 7).

Sulfations are known to affect the binding of proteins[6] and are therefore likely to be classified as critical quality attributes to be monitored during development and throughout the production pipeline.

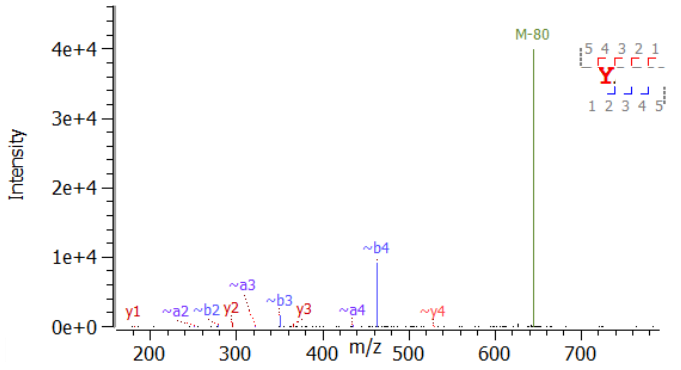

Figure 6. MS/MS spectra of modified peptide. Processing of the peptide mapping data was performed allowing for phosphorylation and sulfation as a potential modification for tyrosine (Y). The singly-charged peptide was identified with a sulfation. Proprietary sequence information was blanked out.

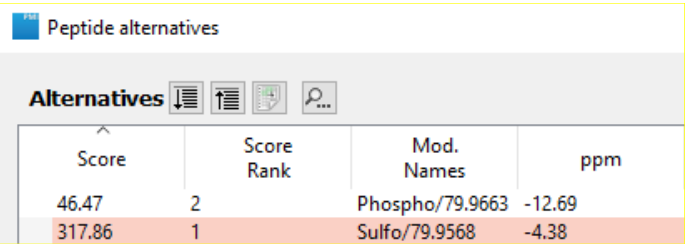

Figure 7. Alternative peptide matches for the modified peptide. Processing of the peptide mapping data was performed allowing for phosphorylation and sulfation as a potential modification for tyrosine (Y). The singly-charged peptide was identified with a sulfation showing a significantly higher score and lower mass error compared to phosphorylation.

N-linked glycosylation

Most biopharmaceuticals are proteins which contain glycosylations. The identification, relative quantification and in some cases modulation of a biotherapeutic's glycosylation profile is an important part during therapeutic development in order to enable a stable, but not immunogenic drug. With the market driving towards new modalities with higher complexity, increased glycosylation microheterogeneity can be expected, which, in turn, increases the analytical burden at the same time.

For the trispecific mAb sample, three glycosylation sites were identified in the LC2, HC1 and HC2.[4] The most abundant glycoforms were identified as G0F, G1F and G2F for both heavy chains.[4] In addition, trace levels of G0 and G0F-GlcNAc were detected in HC2.[4] These results are well in alignment with the studies on peptide level (Figure 8 and 9). However, because the two heavy chains share the same sequence with regards to the consensus peptide containing the glycosylation, peptide mapping analysis cannot differentiate the glycosylation between HC1 and HC2. As a result the relative percentage calculation for the different glycans is an average for both chains.

For accurate information on the glycan distribution of each chain individually, a subunit approach needs to be performed.[4] Hence, there is a need for combining information derived from different approaches (intact, subunit and peptide mapping analysis) when it comes to achieving a comprehensive characterization of complex multi-specific antibody.

Via peptide mapping, G0, G0F and G1F were identified as the dominant glycan species for the HC (Figure 8). The third *N*-linked glycosylation site on the multi-specific antibody is located on the LC2. The most abundant glycoforms observed in subunit analysis were G1S, G2S, G2S2, G2F+2NeuAc as well as trace levels of tri-antennary glycans.[4] Although some glycopeptides of the LC2 were identified from the tryptic digest, more glycans were identified in AspN digest, including G2S, G2S2, G2+NeuAc+NeuGc, G2F+2NeuAc (Figure 9), which is consistent with previous observations.[4] Further low abundant glycopeptides could be identified based on MS1 information. For a customized assay, these low intense forms can be part of an inclusion list in order to obtain MS/MS data despite their low abundance in order to increase confidence in the assignment via diagnostic fragment ions.

				MS Id ←	3
				MS Alias name ←	20191030 sanofi multi specific trypsin lysc_2 (%)
GlycanShortName ↑	Mod. Names ↑	Protein name ↑	Mod. AAs ↑		
G0	NGlycan/1298.4760	mAb|Sanofi_HC1	N		6.36
G0F	NGlycan/1444.5339	mAb|Sanofi_HC1	N		81
G1F	NGlycan/1606.5867	mAb|Sanofi_HC1	N		12.7

Figure 8. Identified glycopeptides on HC1 from the tryptic digest. All identified glycopeptides include MS/MS confirmation. The relative percentages for each glycoform are automatically calculated based on MS1 area information and color-shaded based on abundance for G0, G0F and G1F.

					MS Id ←	1
					MS Alias name ←	20191030 sanofi multi specific aspn _4ug (%)
GlycanShortName ↑	Mod. Names ↑	Protein name ↑	Mod. AAs ↑	Var. Pos. Protein ↑		
	NGlycan/1913.6770	mAb|Sanofi_LC2	N	70		18.7
	NGlycan/2204.7724	mAb|Sanofi_LC2	N	70		65.6
	NGlycan/2220.7674	mAb|Sanofi_LC2	N	70		5.91
G0	NGlycan/1298.4760	mAb|Sanofi_LC2	N	70		5.31
G2FS2	NGlycan/2350.8304	mAb|Sanofi_LC2	N	70		4.49

Figure 9. Identified glycopeptides on LC2 from the AspN digest. All identified glycopeptides include MS/MS confirmation. The relative percentages for each glycoform are automatically calculated based on MS1 area information and color-shaded based on abundance 1298.4760 was automatically linked to G0 and 2350.8304 to G2F+2NeuAc (G2FS2); 1913.6770 can be linked to G2S, 2204.7724 to G2S2, 2220.7674 to G2+NeuAc+NeuGc.

Overall, this assay provides an additional layer of information beyond previous intact and subunit analysis. Not only does it offer high confidence in the sequence confirmation but also in PTM identification as well as localization, even for very challenging cases based on excellent data quality combined with state-of-the-art software processing via Protein Metrics Inc. software. The assay proved to be a platform, suitable for challenging proteins, but still leaves room for specific adjustments if needed.

Conclusions

- A complete workflow to characterize complex new modalities in a streamlined manner is offered by the X500B QTOF System with Protein Metrics Inc. software solutions
- Ease-of-use combined with excellent data quality on the X500B System provides a template workflow platform for the characterization of complex next generation biotherapeutics as well as standard proteins
- Direct compatibility with Protein Metrics Inc. software workflows facilitates the comprehensive and time-efficient analysis of these complex therapeutics

Reference

1. Egan TJ *et al.* (2017) Novel multispecific heterodimeric antibody format allowing modular assembly of variable domain fragments. *MAbs* **9(1),** 68-84.
2. Lum LG and Thakur A (2011) Targeting T cells with bispecific antibodies for cancer therapy. *BioDrugs*, **25(6)**, 365-379.
3. Karaoglu Hanzatian D *et al.* (2018) Brain uptake of multivalent and multi-specific DVD-Ig proteins after systemic administration. *MAbs* **10(5)**, 765-777.
4. Comprehensive characterization of a multi-specific antibody therapeutics. SCIEX technical note RUO-MKT-02-12221-A.
5. Geiger T *et al* (1987) Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. *J Biological Chemistry*, **262**(2):785-797.
6. Zhu JZ *et al.* (2011) Tyrosine sulfation influences the chemokine binding selectivity of peptides derived from chemokine receptor CCR3. *Biochemistry* **50(9)**, 1524-1534.

SCIEX

Headquarters
500 Old Connecticut Path | Framingham, MA 01701 USA
Phone 508-383-7700
sciex.com

International Sales
For our office locations please call the division headquarters or refer to our website at
sciex.com/offices