White Paper

Considerations When Performing an Intact Mass Analysis of a Monoclonal Antibody by LC-MS

Nicholas Michael
Protein Technical Specialist

Pat Easton
Biomolecular Analysis Manager

rssl
science with service

Considerations When Performing an Intact Mass Analysis of a Monoclonal Antibody by LC-MS

Introduction

Monoclonal antibodies are one of the most common types of biopharmaceuticals. Knowledge of an antibody's molecular weight is essential; if this differs from predicted weight then this indicates that there is a change to the structure. One way of screening for these changes is through intact mass analysis by liquid chromatography-mass spectrometry (LC-MS). Although this is a useful screening technique, data still has to be interpreted with care. A case study will be used in this paper to highlight a few issues that need to be considered when interpreting this data.

IgG antibodies are large, complex macromolecules of around 150kDa comprised of two light and two heavy chains complexed via several disulphide bridges that have to be aligned correctly to ensure correct folding.

Additional modifications may be either co-translational e.g. N-linked glycosylation on the heavy chains or post-translational e.g. further processing of the N-linked glycosylation or oxidation of methionines. It is common, however to refer to co-translational modifications as post-translational modifications (PTMs). Incorrect glycosylation's can cause a biotherapeutic protein to be immunogenic or have a reduced half-life (1).

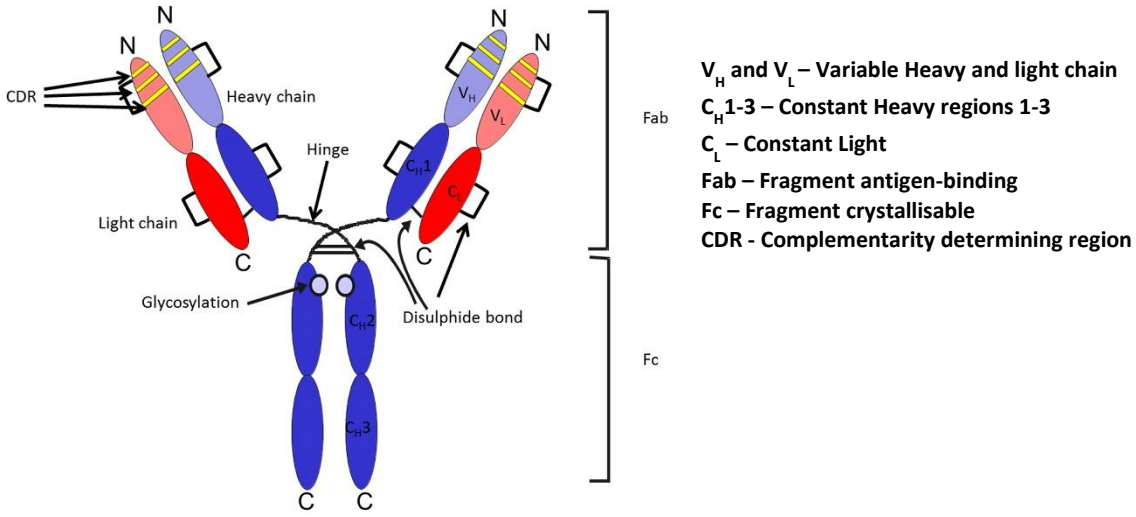

Figure 1. A generic, IgG1 monoclonal antibody structure

The type and extent of PTM formation is very sensitive to environmental factors during production and subsequent processing (2), as a result it is important that the degree of allowable heterogeneity is established when determining the critical quality attributes during development. Regulators demand analytical data to confirm that such known variabilities are within defined acceptance criteria.

Experimental

The antibody used as a model protein in this study was a murine IgG monoclonal antibody purified from Mouse Myeloma NS-1 cell culture media, with a known amino acid sequence and the following known modifications:

Glutamine to pyro-glutamic acid on Heavy chain N-termini
Glycosylation of Asparagine residue 292
Clipping of Heavy chain C-terminal Lysine
17 Disulphide bridges (2 intra light, 4 intra heavy, 5 interchain)

The accurate, average molecular weights were determined both theoretically and experimentally for the intact and disulphide bond-reduced molecule. The accurate mass of the deglycosylated, non-reduced and the deglycosylated, reduced molecules were also determined. For reduction of disulphide bonds, dithiothreitol (DTT) was used. For deglycosylation Peptide-N-Glycosidase F (PNGase F) was used, which selectively cleaves between asparagine residues and the innermost GlcNAc residue of the N-linked oligosaccharides that are found in mammalian glycoproteins.

The intact mass experiments were carried out using a Thermo QExactive LC-MS. The QExactive was run in an MS1 only mode in which molecular ion information only was acquired.

The samples were loaded onto a C3 reverse-phase HPLC column and eluted off with an increasing organic solvent gradient with a constant level of formic acid. UV absorbance at 280 nm was acquired immediately prior to entering the mass spectrometer. In brief, the ions entered the mass spectrometer from the electrospray ionisation (ESI) source and encountered the quadrupoles which were set to allow passage of all ions within the chosen m/z range. These ions accumulated in what is known as the C-trap. After a predetermined number of charges were detected in the C-trap the ions were injected into the Orbitrap, where they orbited, creating a detectable image current. Using a Fourier Transform function this was converted into a mass spectrum.

Nominal, Monoisotopic and Average Molecular Weights

When looking at large molecules such as large peptides or proteins what is determined is the average molecular weight. In mass spectrometry it is important to appreciate the difference in terminology when it comes to discussing mass.

The nominal mass is the sum of the masses of the most abundant isotope e.g. H1, C12, N14, O16 etc. and not accounting for mass defects (loss of mass due to binding energies).

The monoisotopic mass is the sum of the masses of the most abundant isotope, using the exact mass accounting for mass defects e.g. ^{1}H = 1.007825, ^{12}C =12.000000, ^{14}N = 14.003074, ^{16}O = 15.994915.

The average mass is the sum of the masses of the average mass of each element taking into account the isotopes and the isotope distribution e.g. H = 1.00794, C = 12.011, N = 14.00674, O = 15.9994.

The impact of the isotopic distribution on average masses is much more pronounced with large molecules such as proteins. This can be seen using carbon and its ^{13}C isotope as an example. The proportion of ^{13}C on Earth is 1.11% with the bulk of the remainder as ^{12}C. Therefore for a small molecule such as aspirin with only nine carbons, the ^{13}C isotope is low in comparison to the ^{12}C isotope. As a molecule increases in mass and in particular carbon content, the contribution from the ^{13}C isotope increases. As a result, for peptides above 1000 Daltons the ^{13}C isotope becomes larger than the ^{12}C isotope (3).

Predicted protein average molecular weights therefore make the assumption that the protein has the same isotope distribution as the reference used in the calculator, a reference file that describes the isotope distribution of each element in nature. The 'natural' distribution however is not consistent within living systems and natural variability in isotope distribution can have noticeable effects when looking at large molecules such has antibodies (2).

This is illustrated in Table 1 where the predicted average molecular weight of the murine model antibody was calculated based on its elemental composition using Chemcalc (4), an online mass calculator. It can be seen that the value obtained for the molecular weight varies depending on the year in which the reference file was released. Also shown are calculations using a commercially available protein calculator, named ProteinCalc.

Molecular formula: C_{6584} H_{10124} N_{1706} O_{2086} S_{52}

Reference File	Average Molecular Weight	Monoisotopic Mass
1995	148222.10716	148128.404874
2012	148220.77614	148128.4044560844
2013	148220.640631	148128.4044654614
ProteinCalc	148220.3920	148128.4049

Table 1: Nominal masses of the model antibody

The cause of the differences seen in the reference files, whether or not due to new equipment and/or new sources of raw elements, is beyond the scope of this paper. When looking at large proteins such as antibodies, which are circa 150 kDa, such effects are more pronounced than for smaller molecules.

Therefore, although it is important to ensure that the most up-to-date reference file is chosen, the largest discrepancy is generally less than 1.8 Daltons.

Analysis of the Intact and Deglycosylated Antibody

In addition to obtaining accurate (average molecular) masses of the intact non-reduced and reduced molecules, the accurate mass of the deglycosylated non-reduced molecule and the accurate mass of the deglycosylated reduced molecules were also determined, both theoretically and experimentally.

The image of the intact antibody acquisition is shown in Figure 2 as an example of the typical data obtained. The total ion chromatogram (TIC) is at the top, below that is the UV chromatogram and below that is the summed mass spectrum of the peak. It can be seen that the mass spectrum has what is known as a charge envelope and has a typical bell-shaped normal distribution. Each of the peaks represents the same molecule but with an increasing number of protons as one moves down the mass/charge (m/z) axis.

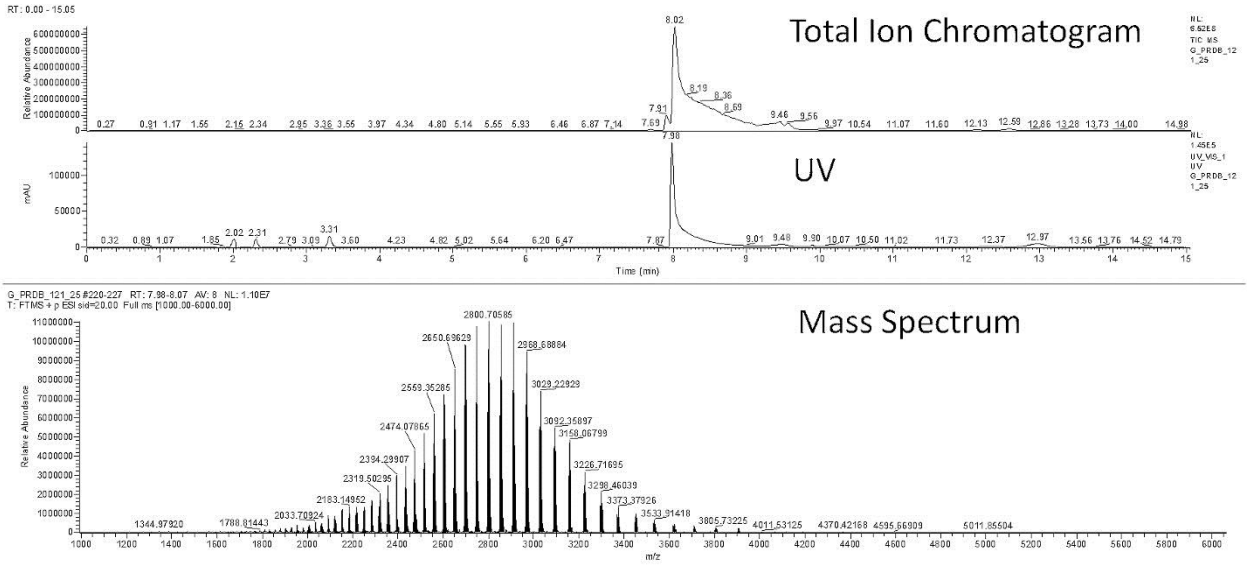

Figure 2: Mass spectrum of the intact, unmodified antibody

When the spectrum in Figure 2 is examined in more detail it can be seen that each peak in the mass spectrum is in fact a family of ions comprising of at least five significant members (see Figure 3). These members represent the different glycoforms of the antibody.

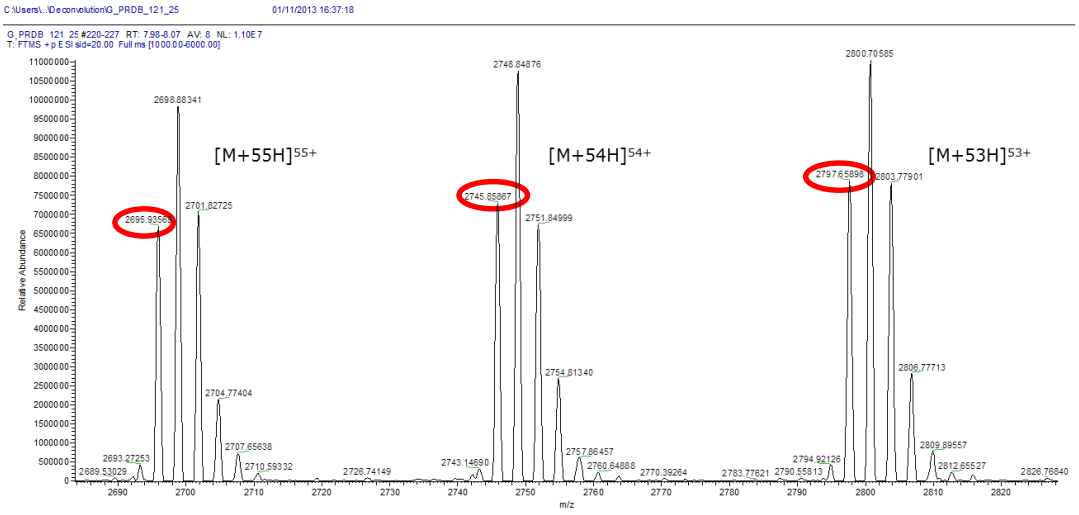

Figure 3: Zoom-in of the mass spectrum of the intact, unmodified antibody

The process used to determine the mass of the neutral molecule from the charge envelope is known as deconvolution and is performed using commercially available software. In simple terms, the software attempts to match an ion to a neutral mass trying different charge states. It requires adjacent ion families across the charge envelope to broadly agree and then an average neutral mass is generated. Using the spectrum in Figure 3 as an example, the mass of the antibody can be obtained by basic arithmetic using the first significant member of each family (shown ringed in Figure 3) as follows:

(2695.93569 x 55) – 55 = 148,221.5 Da.
(2745.85867 x 54) – 54 = 148,222.4 Da.
(2797.65898 x 53) – 53 = 148,222.9 Da.

The value generated by the deconvolution software for the same member was 148,221.9 Da. as shown in Figure 4. This is in close agreement of the predicted value of 148,220.4 Da (see Table 3). The other isoforms differ by units of approximately either 146 Da. or 162 Da. and therefore are attributable to a differing number of units of dHex or Hex sugars, respectively. Two glycan structures are illustrated for each peak because the intact antibody has two heavy chains and there will therefore contain one glycan structure on each chain.

It is noteworthy that there appears to a small amount of the glycosylation, G0, as this is not described on the datasheet provided by the antibody supplier. However it may be that a fucose unit fell off whilst on its way into the mass spectrometer. In-source fragmentation is a known phenomenon and should therefore be considered when investigating unexpected results.

It is important to realise that the glycosylation assignments are based purely on mass and expected glycosylations. Therefore any isobaric (same molecular weight) structures will not be differentiated. This underlines the importance of care when interpreting data, particularly for glycosylated proteins when alternative approaches should be used for definitive assignments of glycan structures (5).

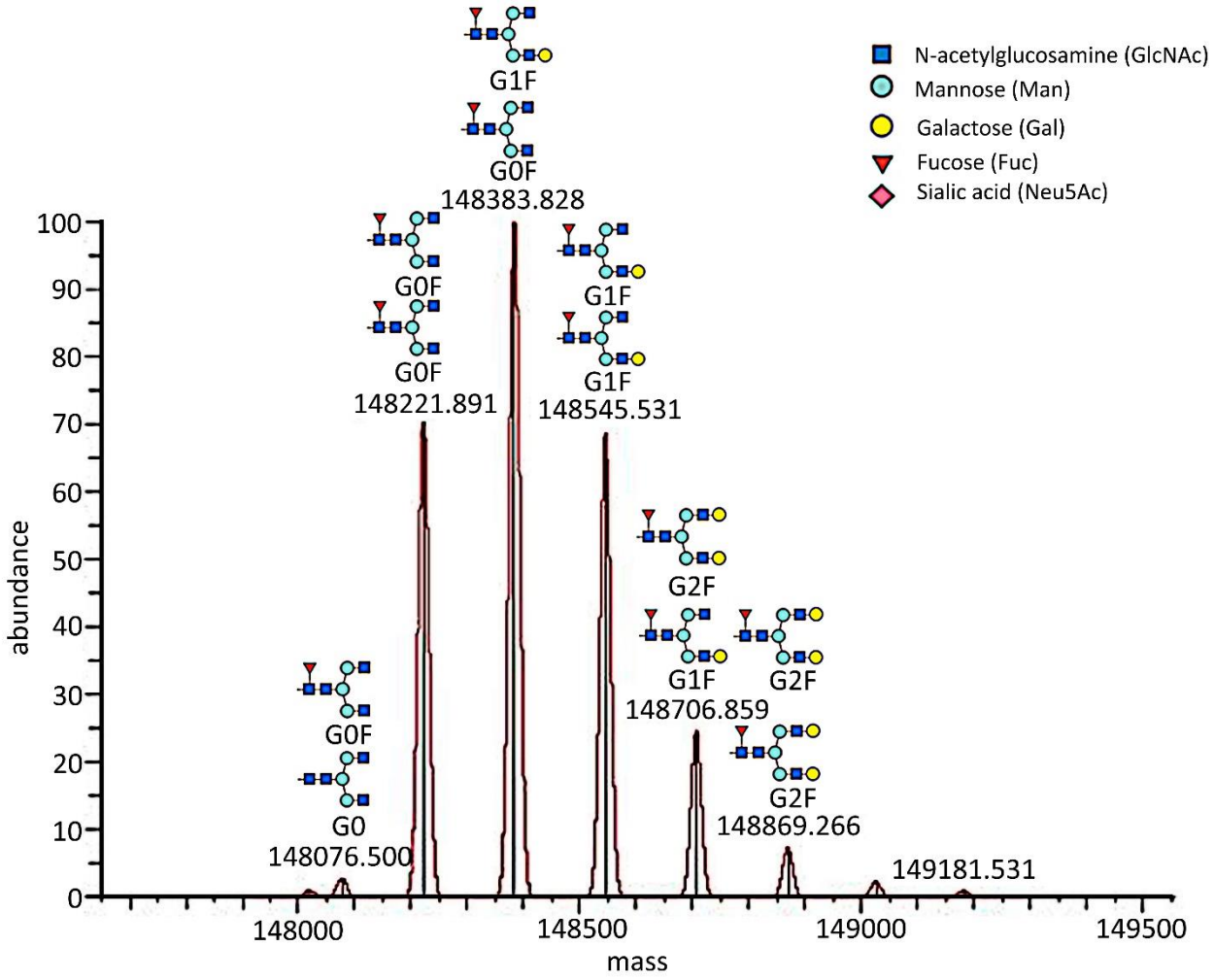

Figure 4: Deconvoluted mass spectrum of the intact, unmodified antibody showing the main glycan structures

Figure 5 shows a section of the mass spectrum of the PNGase F-treated and therefore deglycosylated antibody. It is apparent that the glycofamily profile seen in in Figure 3 has now disappeared. This therefore confirms that the glycosylation illustrated in Figure 4 was N-linked.

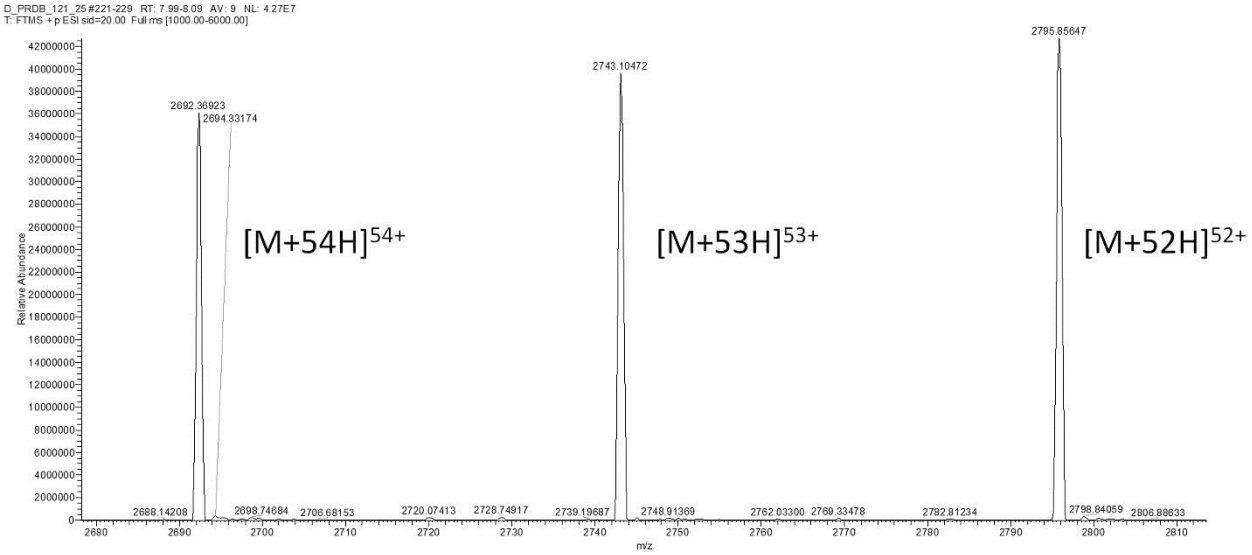

Figure 5: Mass spectrum of the deglycosylated antibody

When the spectrum in Figure 5 is deconvoluted, this derives a mass of 145,332.4 Da. (data not shown) for the deglycosylated antibody. The difference between this and the G0F/G0F mass (148,221.9 Da.) shown in Figure 4 is 2,889.5 Da. This equates to a loss of 1,444.75 Da. from each heavy chain which is in agreement with the loss of a G0F moiety.

Analysis of the Reduced Antibody

The next treatment analysed was the reduced antibody. In Figure 6, it can be seen that there are two peaks, one for the light chain and one for the heavy chain, each with their own charge envelope. As expected, the heavy chain spectrum shows a similar glycoprofile as was observed in the spectrum of the intact molecule.

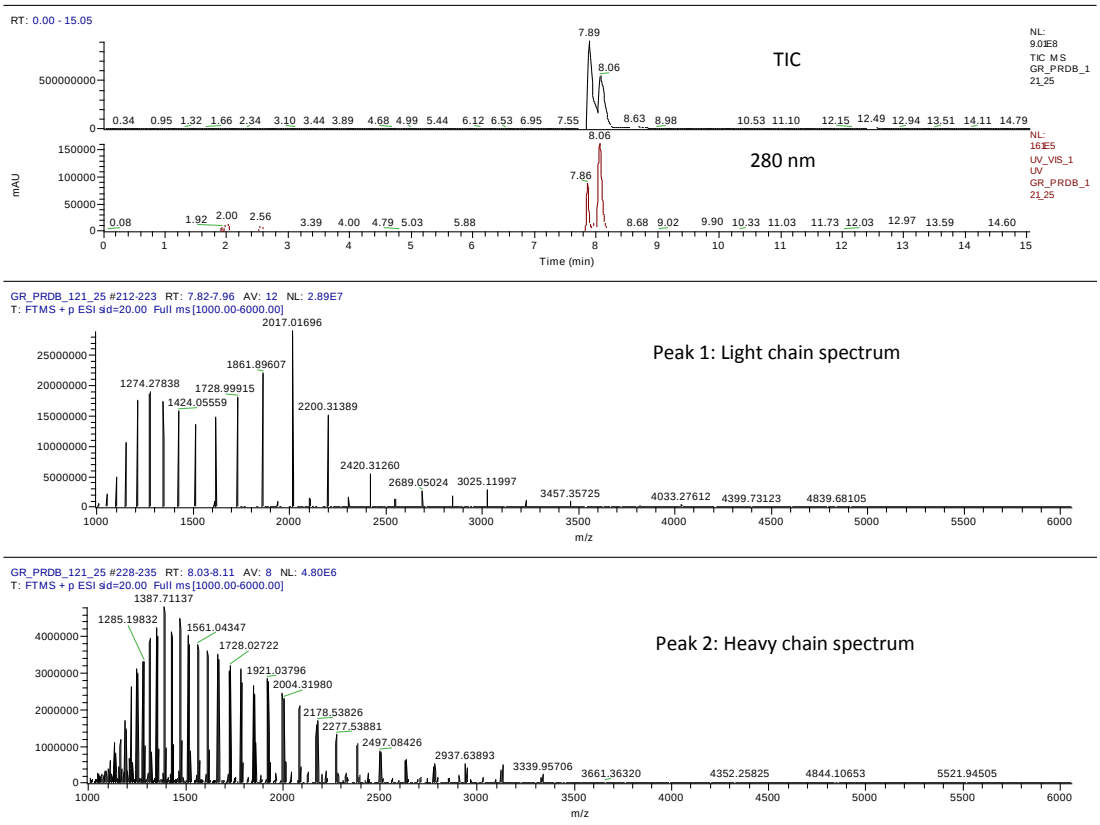

Figure 6: Mass spectrum of the reduced antibody

There are some interesting features to note in the mass spectrum of the light chain. It actually has three charge envelopes which partially overlap. This is shown in more detail in Figure 7 where the apex for each charge envelope is at 1274.3 m/z, 2017.0 m/z and 2420.3 m/z respectively. Each charge envelope has a normal distribution.

Having more than one charge envelope for the same molecule is suggestive of different tertiary structures within the gas phase (6). Proteins in an unfolded state acquire more protons than in a more compact state resulting in a lower m/z ratio, therefore the first charge envelope centred around 1274.3 m/z would be associated with a more unfolded light chain. Caution needs to be exercised when interpreting this data though as instrument settings and mobile phase conditions can also affect tertiary structure resulting in artefactual changes (7). The use of standards such as such as reference samples prepared alongside the test samples can be used to indicate if the observed charge envelope distributions are experimentally-derived or are indicative of true differences in tertiary structure.

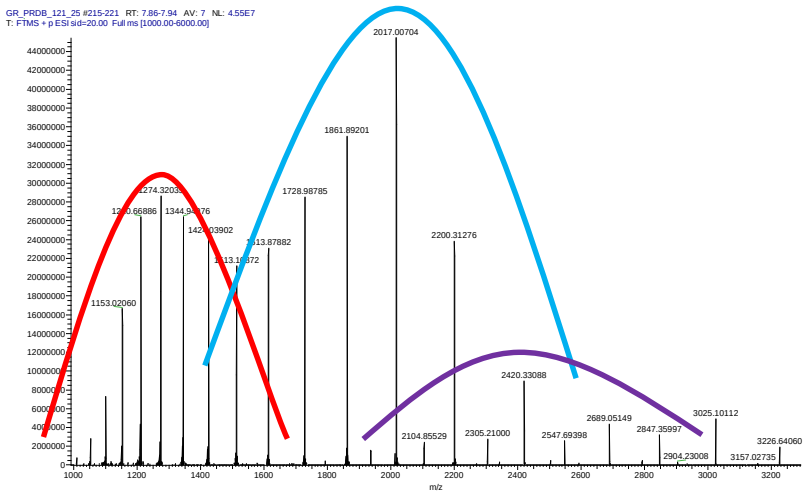

Figure 7: Zoom-in of the mass spectrum of the light chain indicating the three charge envelopes

When the spectrum in Figure 7 was deconvoluted, two major peaks were observed with masses of 24,192.3 Da and 48,387.3 Da, as shown in Figure 8.

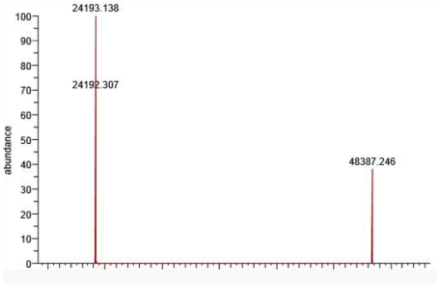

Figure 8: Deconvoluted mass spectrum of the light chain

It is no coincidence that the higher value is double the lower value. It is known that deconvolution software can often make such results that are sometimes spurious. However in this instance, in addition to the light chain mass, there is indeed a molecule that is twice the mass of the light which almost certainly is two light chains. Due to the reduction step in sample preparation it is likely that the interaction is a non-covalent association, something that is well-documented (8).

The proof that this dimer is likely to have existed at the time it entered the mass spectrometer can be obtained from a comparison of the theoretical m/z for each potential charge state for a both a light chain (24,193.6 Da.) and two light chains (48,387.2 Da.). As would be expected, every monomer light chain's m/zs are identical to every other homodimer light chain's m/zs, therefore half the homodimer light chain's m/zs are unique to a homodimer light chain. The ions 2305.2 m/z and 2547.7 m/z in the spectrum can only match a homodimer light chain.

Analysis of the deconvoluted monomeric heavy chain highlighted that there were no masses that correlated to light-heavy chain dimers. It also confirmed the intact mass data showing that the most abundant glycoform is G0F, followed closely by G1F. The abundance of G2F is significantly lower than these and glycoform G0 is only just detectable.

The final condition was deglycosylated and reduced. The results from this analysis were as expected (see summary in Table 3) with the light chain peak giving identical results as the non-deglycosylated reduced sample. This is as would be expected since the light chains are not glycosylated. The spectrum of the heavy chain was also as expected with the glycosylation profile disappearing as it had for the deglycosylated and non-reduced sample. Again the suggestive homodimer profile was present at 48,478.4 Da. and 96,957.3 Da (data not shown).

A comparison of the experimentally derived masses with the theoretically predicted masses is shown in Table 3. All masses were within 1.5 Da. of the predicted, so increasing the confidence that there are no major unexpected truncations, extensions or modifications. However, small amounts of modification would not necessarily be detected by an intact mass experiment. Also there is no precise sequence information derived from this analysis. If for example the light chain's residues were in a different order but still contained the same numbers of each amino acid the intact mass would remain the same.

Molecule	Condition	Predicted Average Mass (Da.)	Experimental Average Mass (Da.)
Light chain	*Partially reduced	24,193.6	24,192.3
	Fully reduced	24,197.7	-
Heavy chain	*Partially reduced (G0F)	49,921.6	49,922.3
	Fully reduced (G0F)	49,929.7	-
	*Partially reduced & deglycosylated	48,477.3	48,478.4
	Fully reduced & deglycosylated	48,485.3	-
Whole molecule	Intact (G0F + G0F)	148,220.4	148,221.9
	Deglycosylated	145,331.7	145,332.4

Table 3
*In the sample preparation conditions used, inter-chain disulphides are far more susceptible to reduction than intra-chain disulphides

Conclusions

Mass spectrometry can give very precise and accurate results. Intact mass analysis can be used as a screen prior to detailed peptide mapping analysis for characterisation and monitoring of post-translational modifications. It has the advantages of minimal sample preparation resulting in fewer sample-handling artefacts and shorter acquisition times. Although data interpretation is relatively straightforward care still needs to be taken as discussed above.

In addition, because experimental error can be up to 2 Daltons when looking at such large molecules, intact mass analysis is not suitable for detecting very small differences such as a single deamidation (+0.98 Da. mass shift) or a single disulphide bridge reduction/formation (+/- 2.02 Da. mass shift). It is however very good for detecting larger mass differences such as C-terminal lysine clipping of the heavy chain.

References

1. Ha S, Ou Y, Vlasak J, Li Y, Wang S, Vo K, Du Y, Mach A, Fang Y, Zhang N. Isolation and characterization of IgG1 with asymmetrical Fc glycosylation. Glycobiology. 2011 Aug;21(8):1087-96. doi: 10.1093/glycob/cwr047. Epub 2011 Apr 5. PubMed PMID: 21470983.
2. Zhang Z, Pan H, Chen X. Mass spectrometry for structural characterization of therapeutic antibodies. Mass Spectrom Rev. 2009 Jan-Feb;28(1):147-76. doi: 10.1002/mas.20190. Review. PubMed PMID: 18720354.
3. Strupat K. Molecular weight determination of peptides and proteins by ESI and MALDI. Methods Enzymol. 2005;405:1-36. Review. PubMed PMID: 16413308.
4. Chemcalc.org – www.chemcalc.org
5. Michael N, Neville D, Easton P, Liew LP, Royle L. Orthogonal Approaches for the Analysis of Protein Sequence and Post Translational Modifications of a Monoclonal Antibody. RSSL white paper
6. Konermann L, Douglas DJ. Unfolding of proteins monitored by electrospray ionization mass spectrometry: a comparison of positive and negative ion modes. J Am Soc Mass Spectrom. 1998 Dec;9(12):1248-54. PubMed PMID: 9835071.
7. Zamani L, Lindholm J, Ilag LL, Jacobsson SP. Discrimination among IgG1-kappa monoclonal antibodies produced by two cell lines using charge state distributions in nanoESI-TOF mass spectra. J Am Soc Mass Spectrom. 2009 Jun;20(6):1030-6. doi: 10.1016/j.jasms.2009.01.008. Epub 2009 Feb 27. PubMed PMID: 19251439.
8. Leitzgen K, Knittler MR, Haas IG. Assembly of immunoglobulin light chains as a prerequisite for secretion. A model for oligomerization-dependent subunit folding. J Biol Chem. 1997 Jan 31;272(5):3117-23. PubMed PMID: 9006964.

Nicholas Michael – Protein Technical Specialist

Nicholas Michael is a Protein Technical Specialist. He studied with the Department of Pharmacy at the University of Brighton and School of Crystallography at Birkbeck, London. Nicholas has significant experience identifying unknown proteins and pathways, and has also worked in the private sector, where he has used his LC-MS skills to characterise manufactured monoclonal antibodies to current Good Manufacturing Practice.

Pat Easton – Biomolecular Analysis Manager

Pat Easton is Biomolecular Analysis Manager at RSSL. Following her PhD involving the analysis of leukotrienes by mass spectrometry, Pat has seventeen years' experience within a biotechnology environment at Amersham International (later GE Healthcare) focusing on protein and genomic arrays, protein labelling and detection. She then moved into a Quality Control Manager role at GE Healthcare where she stayed for six years. Following this, Pat joined RSSL as the Pharmaceutical Chemistry Laboratory Manager. Her combination of protein chemistry and cGMP experience puts her in a good position to lead and manage the services that RSSL offers.

For further information, please contact Customer Services on enquiries@rssl.com