

Characterisation of Template Driven and Non-Template Driven Features of Biotherapeutics

by Esme Candish, PhD, Biopharma Application Scientist, Zoe Zhang, PhD, Biopharma Development Scientist,
Todd Stawicki, Biopharma Global Marketing Manager
Sciex, Framingham, MA, USA

Biotherapeutics pose many challenges to pharmaceutical researchers tasked with their development and characterisation. The complexity of the biologic as well as the specific demands of the development pipeline require analytical techniques that can provide the accuracy, precision, and speed necessary for analysing an array of biologic features. These features can be thought of as template driven (DNA-directed), non-template driven (variable), or mixed features. Many of the more powerful techniques for biotherapeutic characterisation have traditionally been found in research labs and have been slower, more complex, and less robust. Here, new advances in liquid chromatography high resolution mass spectrometry, capillary isoelectric focusing, and capillary electrophoresis have been tailored to the needs of the biologics researcher in downstream biopharmaceutical environments and greatly simplify and speed biologics characterisation.

Biotherapeutics, or biologics, represent a rapidly expanding segment of the drug therapy market, with virtually every major pharmaceutical manufacturer now actively pursuing the production of therapeutic materials using biological methods. Antibody based drugs represent the largest and most rapidly expanding class of biotherapeutics [1]. Along with this growth, traditional small molecule based pharmaceutical companies have had to embrace new processes and technologies that radically deviate from those used within the typical small molecule development pipeline.

In contrast to small molecule drugs, protein-based biotherapeutics, particularly monoclonal antibodies (mAbs), are large and complex, often consisting of heterogeneous mixtures. Whereas small molecule drugs are synthesised, biotherapeutics are created in cell culture by transferring a gene for a desired biotherapeutic to a living host organism. Although the basic biotherapeutic molecule that is produced is predominantly the result of translating the newly added DNA sequence, numerous, and often unwanted modifications can be incorporated into the biomolecule during development and storage.

When characterising biotherapeutics, the features of the product preparation and individual molecules can be thought of as 'template driven' or 'non-template driven' (Figure 1). Template driven features are those specified at the DNA level; specifically, the primary amino acid sequence of the

protein. Non-template driven features are not encoded at the DNA level and may be non-specific, unpredictable, and influenced by many factors. In between these two extremes are those features that can be considered to arise from a mixture of template and non-template driven processes where the feature may result from a post-translational modification event but can only occur if a specific amino acid or sequence of amino acids is present in the protein. These amino acids can be considered as liabilities. If these amino acids

are modified in a negative way, they become a liability to product quality or safety.

The extent of any one feature incorporated into a biotherapeutic can be difficult to predict. Adding to the difficulty is that these features can occur in tandem with one another, each with a different probability, creating a complex mixture for the biotherapeutic formulation. Empirical observation through detailed analysis is therefore the only alternative for fully characterising a biotherapeutic.

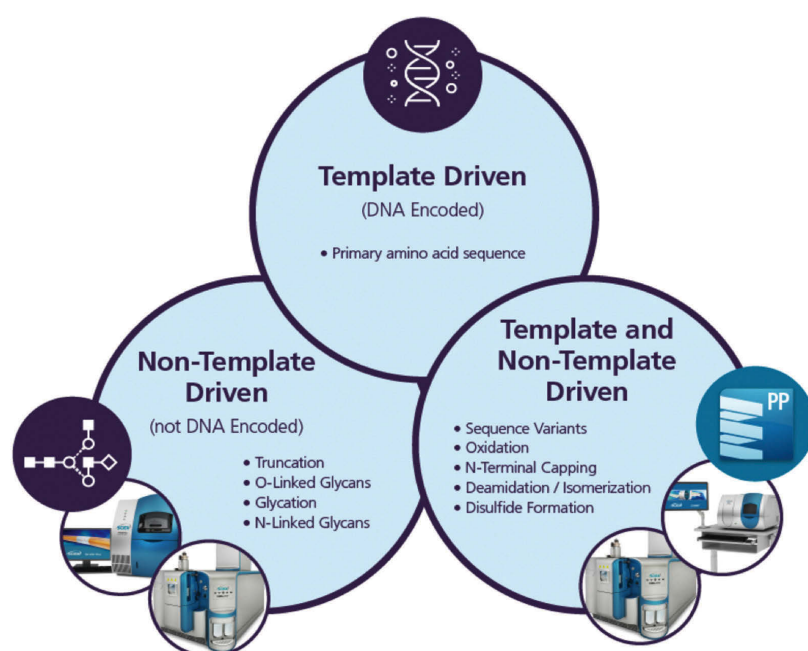


Figure 1: Categorisation of biotherapeutic features

The primary amino acid sequence, a template-driven feature, is typically one of the first characterisation studies to be performed as the fidelity of the protein sequence is of utmost importance. However, under certain conditions, sequence variants can arise. For example, DNA level mutations may be introduced by incorrect nucleotide incorporations during DNA replications. Misincorporations of the wrong amino acid can result if the cell culture medium is lacking in another, required amino acid. These mutations occur as a function of both the primary sequence as well as the local conditions for preparation and can propagate if undetected. Truncation is another primary sequence modification but is non-template-driven. Many c-terminal exopeptidases can cleave the c-terminal ends of recombinant proteins - notably mAbs, in addition to non-specific truncation occurring.

Glycosylation is a non-template driven feature and can be categorised into three types: O-linked glycosylation, N-linked glycosylation, and glycation. O-linked glycans are formed through a post-translational enzymatic process in which sugar molecules are attached to an oxygen atom of an amino acid residue, typically serine or threonine. N-linked glycans are formed when sugar molecules are attached to the nitrogen atom of an asparagine residue. Although there is usually a template-driven consensus sequence for the specific site of N-linked glycosylation, the composition of the glycan itself is entirely non-template driven. N-linked glycans are synthesised by a multitude of enzymes and are attached post translationally. While O-linked and N-linked glycosylation are typically key components of the biotherapeutic that are essential for overall efficacy, glycation is an entirely different thing. Glycation is the non-enzymatic attachment of a sugar molecule to a protein and can occur under certain storage or production conditions. Because enzymes are not involved in the process, glycation is highly unpredictable and impairs the functioning of the biotherapeutic.

Oxidation, deamidation, isomerisation, N-terminal modification, and disulphide bond formation are post-translational features that are controlled by both local amino acid liabilities and the conditions used for preparation and storage. Oxidation occurs primarily on methionine residues by reactive oxygen species within the environment, but can also occur on cysteine, histidine, tryptophan, and tyrosine. Asparagine and glutamine residues have

the potential for deamidation to become aspartate or glutamate, respectively. Modifications to the N-terminus can manifest as methylation or the cyclization of glutamic acid to pyroglutamate. Disulphide bonds can break and reform with incorrect pairings resulting in different protein conformations and scrambling of disulphide bridges.

Thus, biotherapeutic formulations typically consist of a variety of protein isoforms resulting from different combinations of these template- and non-template driven features. These features will all affect biophysical characteristics such as isoelectric point (pI) long-term stability, safety, and potency in various ways. Because there is very little a priori control over the majority of these features within the manufacturing process, full characterisation and close monitoring of the biologic is imperative throughout the development pipeline.

Charge heterogeneity analysis using capillary isoelectric focusing (cIEF) is one way to assess the pattern of microheterogeneity created by these isoforms. cIEF is a high-resolution technique that separates protein variants based on their pI. In order to ensure consistent product quality, this pattern is closely monitored during development and production. Charge heterogeneity analysis provides a view of all the isoforms within the biotherapeutic formulation at a macro level but does not provide a direct measurement of these modifications.

In order to observe these modifications more directly, other techniques are required. For example, capillary electrophoresis is an excellent analytical technique for analysing released glycans and is used extensively for the characterisation of glycans during clinical and commercial phases of development. However due to the large sample sets encountered, long sample preparation times, long separation times typically required, and the complexity of the data analysis, CE is not as amenable to a high throughput scenario as would be encountered during clone selection, cell culture optimisation, and at-line process control within target discovery/validation phases.

Intact molecular weight and subunit analysis are used as a means to confirm the primary sequence of the biotherapeutic. Often chromatography or a form of gel analysis is employed for these studies but liquid chromatography mass spectrometry (LC-MS) can provide much higher accuracy for the molecular weight determination and much higher resolution of components within mixtures. In the case of sequence variants, and post-translational modifications, LC-MS

peptide mapping and MS/MS techniques can provide both the identification of the variant amino acid as well as the location within the primary sequence. However, because these workflows demand high quality data with complex peak finding and data matching algorithms, the instrument expense and expertise required for these applications have historically delegated it to the research arm of biopharma. Thus, the challenge for instrument manufacturers has been to develop a small and affordable high mass accuracy high resolution instrument with powerful, easy to learn and use, automated biologics characterisation software for use in downstream biopharmaceutical laboratory environments.

In this study cIEF is demonstrated for charge heterogeneity analysis with a resolution that is difficult to achieve using other analytical platforms. Additionally, a novel capillary electrophoresis solution based on prepared kits for sample preparation and instrumentation and software designed for speed and automation is demonstrated for high throughput analysis of released N-glycans. And finally, the use of a high resolution accurate mass LC-MS/MS system specifically designed for biologics analysis is demonstrated for characterisation of biologic features.

Experimental

Samples

For the intact molecular weight analysis and oxidation study, NISTmAb, a well characterised humanised IgG monoclonal antibody standard, was obtained from the National Institute of Standards (#RM8671). For the charge heterogeneity analysis USP IgG was used. For the released N-glycan study, antibodies were MAK33-commercial purified antibody and IgG 1 k from cell culture sample (protein A captured).

Intact molecular weight analysis:

NISTmAb was diluted in 0.1% formic acid in water.

Released N-glycan analysis:

N-linked glycans were released from the antibodies by digestion with PNGase F and subsequently prepared using the CH100HT Glycan Screening Kit (SCIEX, Concord, ON). With the kit reagents the released glycans were captured by magnetic beads, labeled with 1-aminopyrene-3,6,8-trisulfonate (APTS), and then cleaned up by magnetic bead capture.

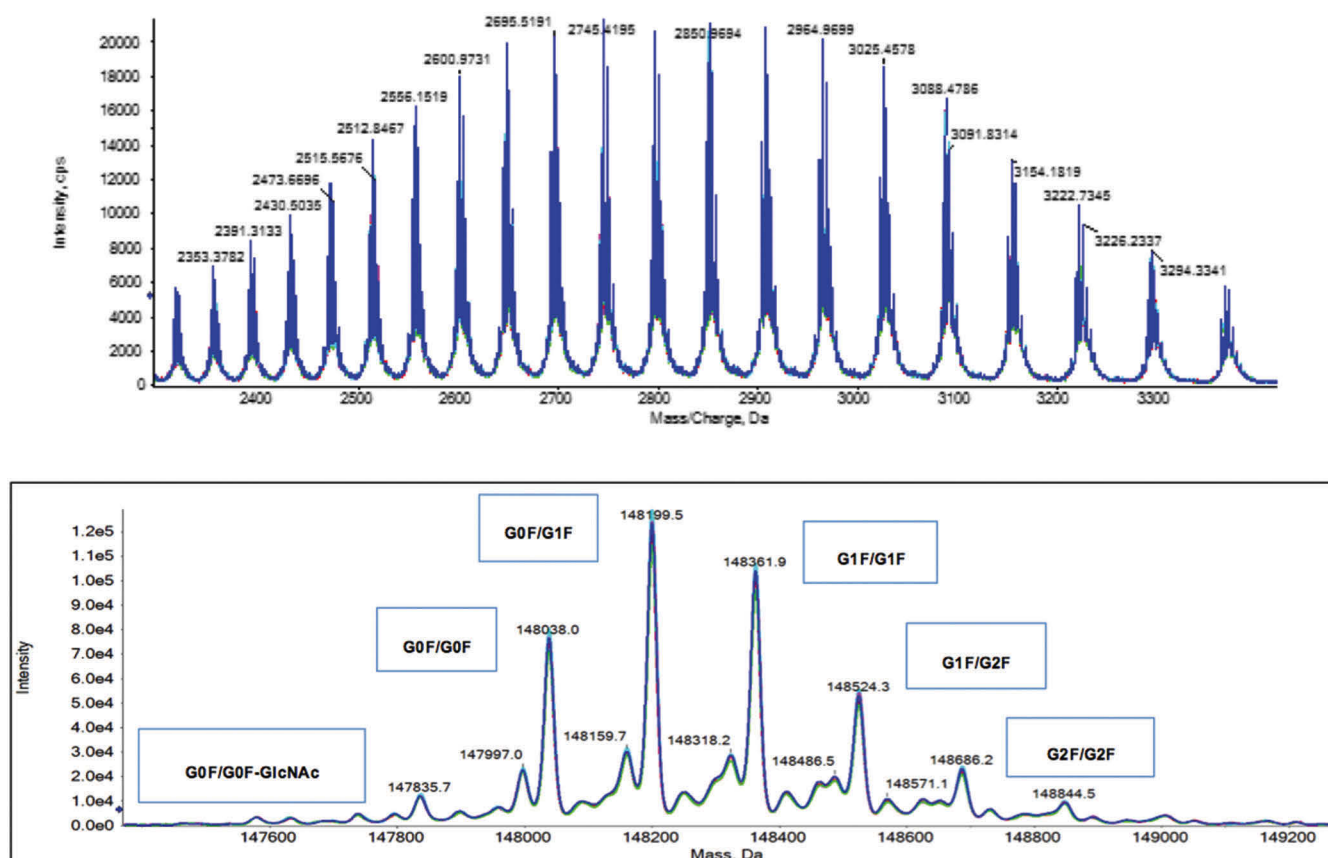


Figure 2: Overlay of four replicate LC-MS analyses of intact NISTmAb. Top: Multiply-charged raw data. Bottom: Reconstructed mass data. Different glycoforms are clearly detected with mass error <10ppm.

Charge heterogeneity analysis:

USP IgG was obtained in a master mix that included urea in cIEF gel, iminodiacetic acid (IDA), Pharmalyte 3-10, arginine, and peptide pI markers. The final IgG concentration was 0.24% m/v. This was reconstituted in DI water to obtain a concentration of 5 mg/mL.

Oxidation study:

NISTmAb was incubated with 0.003125% hydrogen peroxide (H_2O_2). Methionine oxidation was subsequently quenched by adding L-methionine. The treated sample was then denatured, reduced and alkylated using DL-dithiothreitol and 2-iodoacetamide (Sigma Aldrich). Trypsin (Promega) was added in a ratio of 1:30 (w:w; Trypsin:mAb) followed by an incubation at 37°C overnight. Digestion was stopped by adding formic acid and supernatant was subsequently measured using LC-MS/MS.

Liquid Chromatography Mass Spectrometry

All measurements were carried out in replicates on a X500B QTOF System (SCIEX, Concord, ON) coupled to a Turbo V™ ion

source. Chromatography was performed using an ExionLC™ (SCIEX, Concord, ON).

Intact molecular weight analysis:

LC Conditions: Phenomenex Aeris Widepore C4, 200Å, 3.6µm, 2.1x50mm, 80°C, 0.2mL/min, 200 ng injection.

Mobile phase A = 0.1% formic acid in water. Mobile phase B = 0.1% formic acid in acetonitrile.

MS Conditions: TOF-MS scan m/z 400-4000.

Oxidation study:

LC Conditions: 100x2.1, 1.7µm, C18 column, 40°C, 0.3mL/min, 6 µL injection volume.

Mobile phase A = 0.1% formic acid in water. Mobile phase B = 0.1% formic acid in acetonitrile.

MS Conditions: Data dependent acquisition (DDA). High resolution MS/MS data of ten candidate ions per cycle with a total cycle time of 1s were acquired.

Capillary Electrophoresis

Released N-glycan analysis:

The C100HT system (SCIEX, Concord, ON) was used.

The labelled released glycan samples were loaded on the ready-to-use 12-capillary GlycanSeparation Cartridge. The glycans were automatically separated with the automated method.

Charge heterogeneity analysis:

The PA 800 Plus Pharmaceutical Analysis System, (Sciex, Concord, ON) was used.

Each sample was diluted into a mixture of Urea-cIEF Gel, Pharmalyte 3-10 carrier ampholytes, cathodic stabiliser, anodic stabiliser and pI markers A, B, C.

The sample mixture was introduced to the neutral coated capillary, 50µm i.d. x 45 cm, at 15 psi for 150 seconds.

The cIEF separation was achieved using 30 KV over 25 minutes.

Data Analysis

Data from the released N-glycan and charge heterogeneity studies were processed using the CT100HT and PA800 software, respectively. Intact molecular weight data and oxidation data were processed using BioPharmaView™ software (SCIEX, Concord ON). For the oxidation studies, the

experimental peptide data were matched to the in-silico generated list of peptide masses with oxidation as a variable modification. The maximum error tolerance of 5 ppm was used for peptide matching.

Results

Intact Molecular Weight Analysis

Because the primary sequence of the biotherapeutic can be corroborated through a measurement of its molecular weight, intact mass analysis is frequently pursued as one of the first characterisation steps. Figure 2 shows the LC-MS analysis of a well characterised monoclonal antibody standard ('NISTmAb') simply diluted in 0.1% formic acid in water. The different N-linked glycoforms of the protein are clearly detected. At the top, the raw data show the multiply-charged envelope that is typical of large protein analysis. At the bottom, the reconstructed spectral data are shown with the different glycoforms labelled. This figure results from the overlay of four 200 ng injections of intact NISTmAb. The high correlation of the overlaid spectra demonstrates the very high precision and reproducibility of the analyses while the extremely small mass errors (less than 10 ppm) from the calculated theoretical mass demonstrate the high mass accuracy. Because there is no need for extensive sample preparation, sample preparation

artifacts are minimised. This is the highest level (tertiary analysis) of the complete biotherapeutic.

Released N-Glycan Analysis

Following the analysis of intact molecular weight, the relative levels of the different N-linked glycan glycoforms can be correlated with released N-linked glycan analysis methods for orthogonal assay confirmation. Figure 3 shows the analysis of a 96-well sample plate of N-linked glycans released from 96 mAb samples taken from a clone selection process. On the left, a visualisation of the plate samples is shown in which green dots indicate the sample 'passed' while red dots indicate the sample 'failed' with respect to pre-set glycan acceptance criteria. On the right the identities and relative abundances of N-linked glycans released from the mAb sample in well F3 are shown. The analysis takes only 8 minutes and peaks are automatically identified and labelled by the software once separation is complete. All 96 samples can be prepared automatically in ~1.5 hours and analysed in about 2.5 hours. At this speed, the analysis is high throughput and done at a scale and speed previously unattainable with other platforms. This makes the workflow amenable to the bioprocessing segment which can generate hundreds of samples during clone selection and cell culture optimisation.

Charge Heterogeneity Analysis:

Figure 4 shows the high resolution cIEF charge heterogeneity analysis of the NISTmAb sample. Multiple charge variants are observed even on this well-characterised and controlled monoclonal antibody. The charge forms are the physiochemical manifestation of deamidations, truncations, or unidentified post-translational modifications. This level of clear electrophoretic separation, with this level of high resolution of the different charge isoforms, is difficult, if not impossible, to achieve on other analytical platforms.

Oxidation Study:

In this study, oxidation was induced by incubating the NISTmAb with oxidising agent prior to digestion. Figure 5 shows the LC-MS and LC-MS/MS evaluation of the peptide DIQMTQSPSTLSASVGDR from this sample. Any oxidised forms are identified at three levels: at the chromatographic level, at the precursor ion level, and at the fragment ion level. In this case, an oxidised form of this peptide at only 1.5% relative abundance was easily detected. Mirror plots clearly show the difference in intact molecular weight in addition to changes in the MS/MS fragmentation pattern between the oxidised and non-oxidised forms. The data show that the workflow can detect, positionally confirm, and quantitate even very low level non-template driven modifications in an automated fashion.

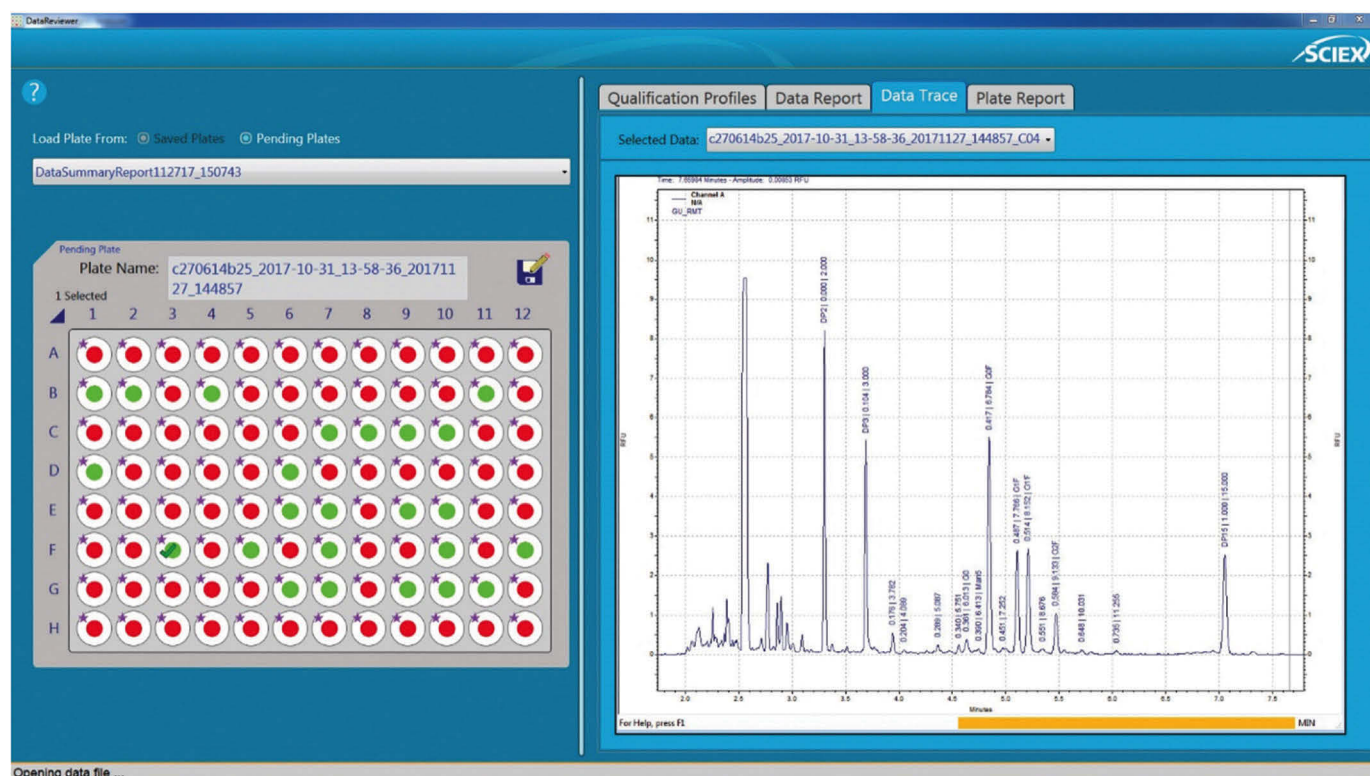


Figure 3: Fast N-glycan analysis. Left: Qualification profile and dynamic pass/fail display of the wells on 96 well plate. Green samples 'pass' while red samples 'fail'. Rows A and H: 9 glycan panel. Rows B, D, and F: MAK33-commercial purified antibody. Rows C, E, and G: IgG 1 k from cell culture sample (protein A captured). Right: Released N-linked glycans from well F3.

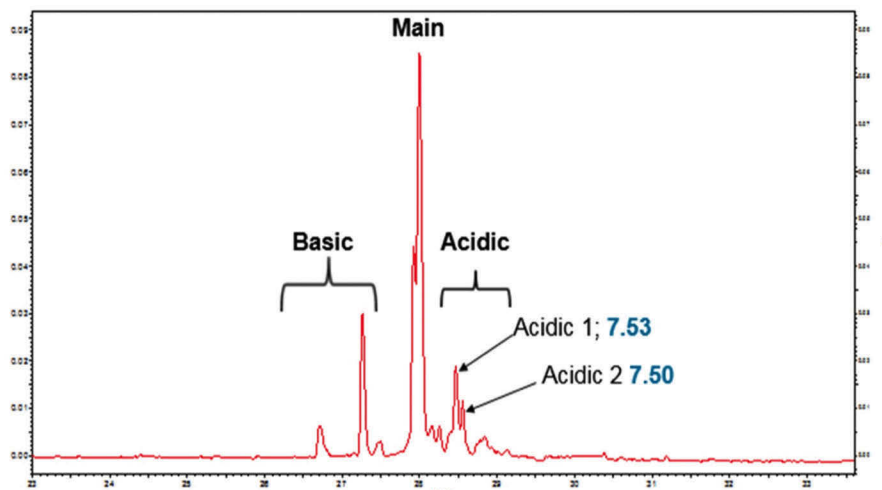


Figure 4: cIEF separation profile of USP IgG with 3-10 ampholytes. *pI* values of USP IgG were in the pH 7.0–7.6 range. The high resolution shows clear separation of multiple charge isoforms.

Conclusion

Analysis and characterisation of template driven and non-template driven features of biopharmaceuticals requires a range of analytical techniques that can meet the challenges arising from both the complexity inherent in the biologic formulation as well as the demands of the pipeline environment. High performance LC-MS/MS can be used for intact molecular weight analysis as well

as characterisation of modifications. Intact molecular weights clearly showing glycan isoforms can be detected with accuracies less than 10 ppm and modifications can be detected at very low levels, for example 1.5% oxidation as shown here. The cIEF charge heterogeneity analysis used here to characterise the various isoforms within biologic formulations shows a resolution typically not available on other platforms. And finally, the speed of the analysis of

released N-linked glycans enables the use of this technique in high throughput scenarios such as bioprocessing. Because same-day results can be obtained from hundreds of samples, real-time corrections can be enabled saving valuable time and resources.

References

1. Monoclonal Antibody Therapeutics Market by Application (Cancer, Autoimmune Diseases, Infection, Hematological Diseases, and Others), By Source and By End Users: Global Industry Perspective, Comprehensive Analysis, and Forecast, 2017 - 2023, Zion Market Research (2018) <https://www.zionmarketresearch.com/requestbrochure/monoclonal-antibody-therapeutics-market>
2. Intercompany Study to Evaluate the Robustness of Capillary Isoelectric Focusing Technology for the Analysis of Monoclonal Antibodies, Salas-Solano, O., Babu, K., Park, S.S. et al. *Chromatographia* 73 (2011) 1137. <https://doi.org/10.1007/s10337-011-2017-3>

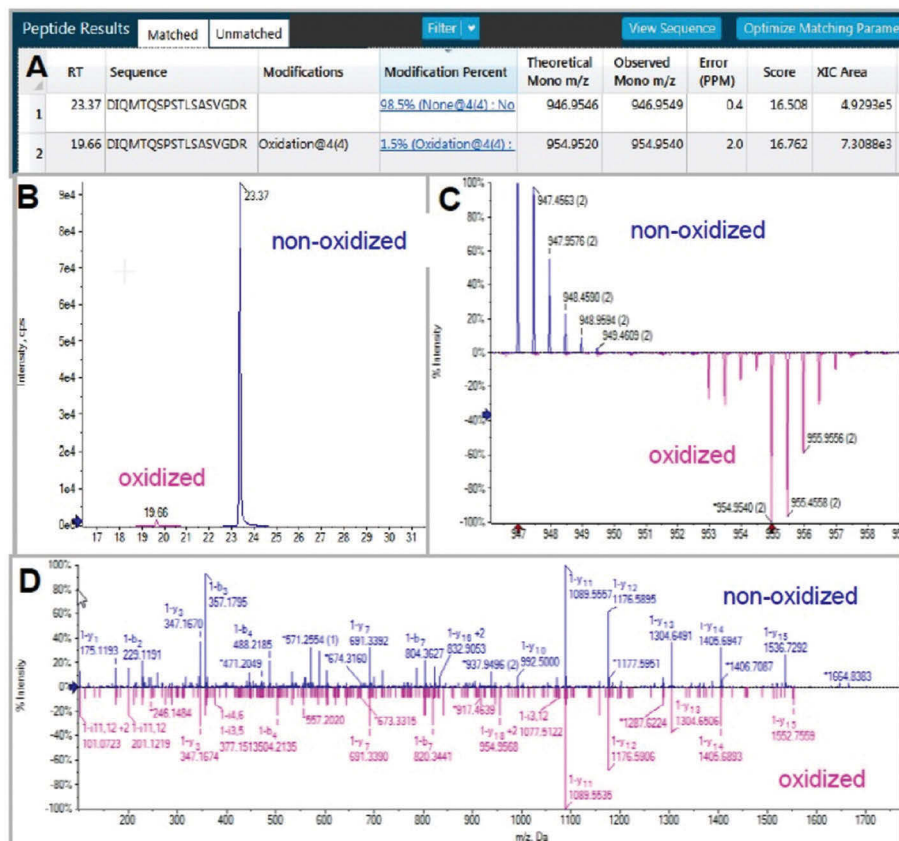


Figure 5: LC-MS and LC-MS/MS evaluation of the peptide DIQMTQSPSTLSASVGDR from NISTmAb sample. A: Table providing results from automated calculation of percentage modification by BioPharmaView software. B: Extracted ion chromatograms (XIC). C: Mirror plot of TOF-MS data. D: Mirror plot of MS/MS data.