

Biopharmaceutical Characterisation – Considering Key Questions

By Dr Fiona M Greer, Director, Biochemical Services at M-Scan Ltd

Following a PhD in Protein Biochemistry from Aberdeen University, Fiona Greer joined M-Scan in 1984 to develop analytical services using the then new techniques of biopolymer mass spectrometry. She is Director of Biochemistry and VP of M-Scan Ltd, M-Scan SA and M-Scan Inc, and is also responsible for QA at all three laboratories.

In the race to bring new biotech derived products to market, researchers face many obstacles along the road from discovery to commercial success. So, any new analytical techniques which can help smooth the path, particularly through the regulatory approval stages, are always welcome. One such breakthrough has been the use of mass spectrometry (MS) to provide information about the primary protein and carbohydrate structure of biomolecules. MS techniques are now being applied routinely to the characterisation of products ranging in size and complexity from small synthetic peptides to conjugated antibody molecules. However, with the bewildering array of modern analytical instruments available, it is sometimes difficult to choose the most suitable strategies to use for a particular characterisation problem, or indeed decide if or when these studies should be performed. This article intends to address some of the frequently asked questions about biopharmaceutical characterisation using mass spectrometry.

QUESTION: WHY ARE RELIABLE ANALYTICAL CHARACTERISATION DATA REQUIRED THROUGHOUT ALL STAGES OF THE BIOTECH DRUG DEVELOPMENT PIPELINE?

Answer: A comprehensive understanding of the chemical structure, physical and biological properties, impurity profile and degradation pathways is an absolute requirement prior to the marketing of any drug substance. In contrast to small molecule drugs, biopharmaceuticals are large, complex entities, often comprising of multiple disulphide-bridged proteins. In general, they are heterogeneous as a consequence of both the biosynthetic processes used and the subsequent manufacturing and storage of the products. It is these microheterogeneities, together with the co- and post-translational modifications to the expected gene-derived protein sequence which require detection, identification and monitoring. A previous article has examined the use of MS in the analysis of post-translational modifications (PTMs) and lists some of the more common types of modification (1).

Perhaps the most important, and certainly the most complex of the PTMs mentioned is glycosylation. Glycoproteins are mixtures of 'glycoforms', that is the same polypeptide backbone but with different populations of sugar chains (glycans) attached. This glycosylation of a protein therapeutic can have an influence on pharmacokinetics and clearance *in vivo*, immunogenicity, solubility and protease resistance amongst other things. It has been shown to be important for biological activity – for instance in human IgG antibodies where changes in activity can be related to different levels of certain monosaccharides, particularly galactose and sialic acids. It is accepted that some degree of heterogeneity exists in most natural as well as engineered glycoproteins, however from a regulatory standpoint this must be shown to be consistent, within specifications and between production batches. The study of glycosylation (glycobiology) is too large an area to more than touch upon here. An overview was given in the previous article (1), but readers are also guided towards a recent discussion of MS glycoprotein strategies (2).

Knowledge regarding the structure of the biomolecule is important not just from a regulatory viewpoint, but also from a commercial one. Small differences between expressed versions of the same recombinant protein or glycoprotein may raise issues of patentability.

QUESTION: WHAT TYPES OF ANALYTICAL DATA ARE REQUIRED?

Answer: A general guide to the requirements for setting specifications to support new marketing applications is given in the document *ICH Topic Q6B, Specifications: Test procedures and Acceptance Criteria for Biotechnological/Biological Products* (3). Although this document does not recommend specific procedures, it lists in an appendix “examples of technical approaches which might be considered for structural characterisation”. The list of these requirements are summarised in Table 1.

Mass spectrometry is mentioned as a technique specifically under the headings of “peptide map”, “sulfhydryl group(s) and disulfide bridges” and “molecular weight or size”. However, an MS-based approach is also of value when determining protein sequence, including “terminal amino acid sequence”, and of course “carbohydrate structure” of glycoproteins.

QUESTION: WHEN SHOULD THESE ANALYTICAL TESTS BE PERFORMED?

Answer: The more analytical knowledge a company has about its molecule, the earlier in the product life cycle, the better to ensure product development and fulfil the regulatory guideline requirements. Extensive analytical testing has now become the route by which manufacturers comply with the United States’ FDA definition of a ‘specified biological product’. This was originally the concept of the ‘well characterized biological product’, introduced by the FDA at a meeting in Washington in 1995, to move away from the approach to traditional biologics in which products were defined by the manufacturing process.

As time evolves in a product life cycle and changes are made to cell lines, processes and manufacturing sites, reliable data will still be required to demonstrate that the product is comparable, with regards to safety and efficacy, to the product before the changes were made. Extensive physicochemical analytical data may be able to demonstrate comparability without the need to perform additional exhaustive new clinical studies, even after the product has been marketed. At the moment, an ICH group is working towards producing a harmonised guidance document (ICH Q5E) on ‘comparability’ later this year.

QUESTION: HOW CAN I USE MASS SPECTROMETRY TO PROVIDE ANALYTICAL CHARACTERISATION DATA?

Answer: Today there are a variety of modern MS techniques which can be applied to the structural characterisation of

Table 1: ICH Topic Q6 B (Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products)	
6.1	Appendix for physicochemical characterisation
6.1.1	Structural characterisation and confirmation a) Amino acid sequence b) Amino acid composition c) Terminal amino acid sequence d) Peptide map e) Sulfhydryl groups and disulfide bridges f) Carbohydrate structure
6.1.2	Physicochemical properties a) Molecular weight or size b) Isoform pattern c) Extinction co-efficient d) Electrophoretic patterns e) Liquid chromatographic patterns f) Spectroscopic profiles

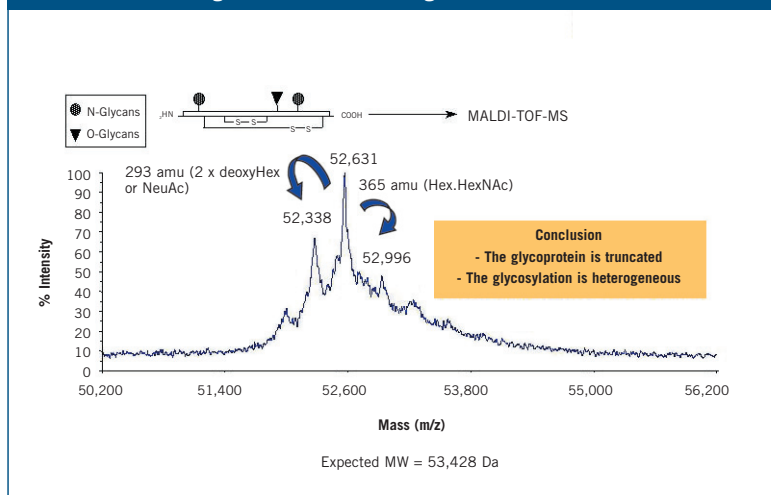
biomolecules. These versatile and sensitive methods can detect, then identify and even quantify changes in the primary protein structure of a product. For biopharmaceutical characterisation studies, the most commonly utilised ionisation techniques include electrospray (ES-MS) and matrix assisted laser desorption ionisation (MALDI). ES-MS can be used in a direct-injection mode but is particularly powerful when coupled to an HPLC as online LC-MS, allowing separation and simultaneous detection of the components of complex mixtures. MALDI is usually performed on an instrument with a time-of-flight (TOF) analyser and delayed extraction (DE) technology. In addition, for glycoprotein studies, where characterisation of the glycan portion of the molecule is required, then fast atom bombardment (FAB) and gas chromatography mass spectrometry GC-MS are also utilised.

However, although it is good to have a choice, experience is required to decide which is the most appropriate MS technique for which application. Consideration has to be given to the limitations of mass range, sensitivity and resolution (which affects mass accuracy) of each different instrument. Equally important are the protein chemical and enzymic treatments used to prepare the biopolymers for analyses. A biopharmaceutical structural characterisation strategy has to be planned with care.

First Step

As both ES-MS and MALDI-MS are capable of ionising intact proteins/glycoproteins up to approximately 150kDa and 500kDa respectively, they are useful for ‘molecular weight or size’ analysis as suggested by ICHQ6B. Thus a simple measurement can be the first step in confirming that a biotech product has the correct anticipated structure, that is that the protein sequence has been correctly translated from the gene with no errors, insertions, deletions or mutations. Any difference between the theoretical mass and the observed mass would point to a change in structure and depending on the resolution of the technique, the mass difference may provide a clue to the type of modification(s) present. An example is given

Figure 1: Molecular Weight Determination



in Figure 1. This MS test of molecular weight can be validated and used in a QC environment as an identity test.

Peptide Map

The next stage would be to take a closer look at the structure of the biomolecule, including any potential modifications and disulphide bridges and so on. For this, the intact molecule needs to be broken down into smaller parts. This is

the basis of the mass-mapping strategy (4). As with routine LC peptide mapping, the molecule is initially fragmented selectively using enzymic or chemical means. Then, however, the unseparated mixture of discrete peptides produced is analysed using ES-MS or more favourably, MALDI-MS to produce a mass fingerprint. If the mixture is too complex, online LC-MS is used, bringing the additional dimension of molecular weight to the peptides separated in the UV chromatogram. Differences between the measured masses and the theoretical masses of the anticipated peptides can be detected. If an ‘unknown mass’ appears in the map, the corresponding peptide can be isolated and collected for further study. As the MS technique relies on measuring mass changes, non-

protein modifications such as sulphation, phosphorylation or addition of lipid or carbohydrate, can be easily detected. And using the same types of instruments, the carbohydrate portions of glycoproteins can also be characterised.

This online LC-MS peptide mapping ‘fingerprinting’ approach to the characterisation of biomolecules has been well established. Again it can become a validated

specification test if required. However, it must be remembered that this strategy aims to confirm the structure of the protein against the theoretical masses of the anticipated peptides – it does not confirm the actual sequence or order of the individual amino-acid residues. It is possible to use certain types of mass spectrometers to produce sequence ion data as well as molecular weight information. Sequence analysis is carried out using either ES-MS/MS performed on triple-quadrupole instruments, or on a new generation of instruments of novel design – a quadrupole analyser followed by a collision cell with an orthogonal injection of ions into a TOF analyser. These instruments (for example, the Q-TOF and Q-STAR) provide the most powerful tandem MS/MS sequencing available (5).

Figure 2a: Deamidation

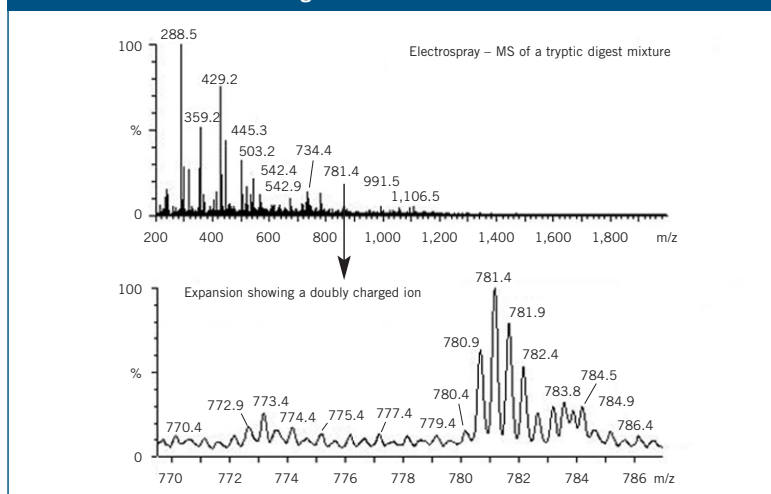
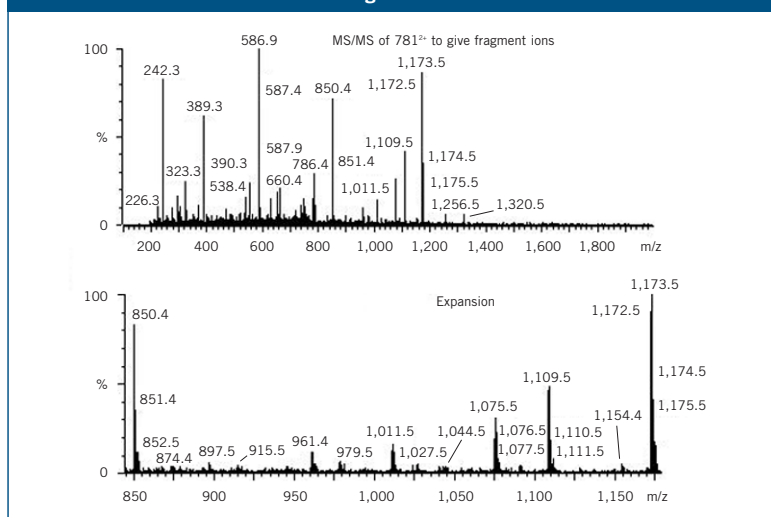


Figure 2b

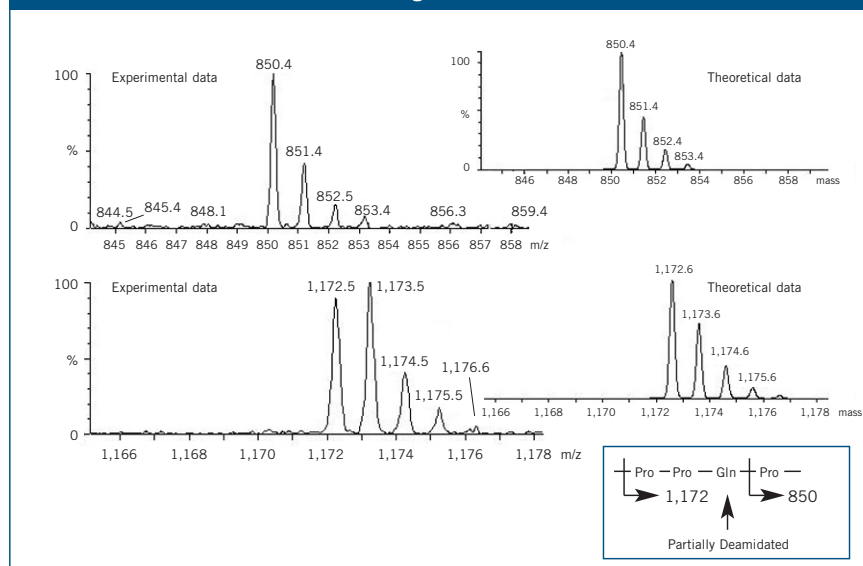


In contrast to the N-terminal region, where the automated Edman Degradation technique is used in gas phase sequencers, there is still no fully reliable analogous method for instrumental sequencing from the C-terminus. Confirmation of the C-terminal sequence is usually obtained using peptide mass mapping strategies combined with carboxypeptidase digestions and MS/MS sequencing.

Peptide Mapping Approach to Impurity ID

Appendix 6.2.2 of the ICH Q6B document, “Product-related impurities including degradation products”, describes three types of “molecular variants of the desired product”. These are listed below:

Figure 2c



- ◆ Truncated forms
- ◆ Other modified forms: deamidated, isomerised, mismatched S-S linked, oxidised or altered conjugated forms (for example glycosylation, phosphorylation)
- ◆ Aggregates

It notes with simple understatement that “such variants may need considerable effort in isolation and characterisation in order to identify the type of modification(s)”.

Using strategies similar to the experiments described above, MS can be used to examine other regions of the molecule where differences may have been initially detected in the peptide mass map. Indeed, with the sensitivity of the new instruments, screening for low abundance modifications, which are not immediately apparent in the mixture map, is also possible.

Deamidation can occur at glutamine (Gln) or asparagine (Asn) amino acid residues. It is characterised by a 1 mass unit shift to higher mass corresponding to the amide being converted into an acid (glutamic acid [Glu] or aspartic acid [Asp]). It can be identified mass spectrometrically by examining the isotopic composition of the relevant signal, once the contribution of the natural ^{13}C isotope is taken into account. Figure 2a shows an example from a tryptic digest including a signal at m/z 781.4. The doubly-charged ion m/z 781 is fragmented in Figure 2b. An expansion of two of the peptide fragment ions, 850.4 and 1172.5 are shown in Figure 2c, together with the computer calculated theoretical isotope ratios for these ions. The observed data for m/z 850 matches the theoretical closely. However, the signal at m/z 1173.5, the ^{13}C isotope of m/z 1172.5 has increased in intensity, indicating that the Gln within the sequence Pro-Pro-Gln-Pro has been partially deamidated.

In a similar way, other modifications can be detected by mass changes. Oxidation of methionine residues can form methionine sulfoxide (+16 mass units), or methionine sulphone (+32 mass

units). Phosphorylation of tyrosine, serine or threonine can be picked up by screening the tryptic digest mixture of peptides, looking, for example, for parent ions which are produced from m/z 79 – the characteristic phosphate fragment.

CONCLUSION

Of course, mass spectrometry is not the only tool in the biochemical analyst's toolbox. And it often requires the use of more than one type of mass spectrometer to generate characterisation data for complex glycoproteins. However, its ability to detect both protein and non-protein modifications, from carbohydrates and lipids to oxidation, sulphation and

phosphorylation ensures that it will be utilised to its maximum capacity. MS data have become a routine component of characterisation packages presented to regulatory authorities and MS is now regarded as a ‘suitable’ technique in many guidelines. ◆

The author can be contacted at services.ltd@m-scan.com

References

1. Greer F, Reason A and Rogers M, Post-translational modifications of biopharmaceuticals – a challenge for analytical characterisation, *European BioPharmaceutical Review*, pp106-111, Summer 2002
2. Dell A and Morris HR, Glycoprotein Structure Determination by Mass Spectrometry, *Science* 291, pp2,351-2,356, 2001
3. ICH Harmonised Tripartite Guideline, Topic Q6B. Specifications: Test Procedures and Acceptance Criteria for Biotechnological/ Biological products. Step 4, Consensus Guideline, March 1999 (CPMP/ICH/365/96)
4. Morris HR, Panico M and Taylor GW, FAB-MAPPING of recombinant DNA derived protein products, *Biochem. biophys. Res. Comm.* 117, pp299-305, 1983
5. Morris HR, Paxton T, Dell A, Langhorne J, Berg M, Bordoli RS, Hoyes J and Bateman RH, High Sensitivity Collisionally-Activated Decomposition Tandem Mass Spectrometry on a Novel Quadrupole Orthogonal Acceleration Time-of-Flight Mass Spectrometer, *Rapid Communications in Mass Spectrometry* 10, pp889-896, 1996