

Post-Translational Modifications of Biopharmaceuticals – A Challenge for Analytical Characterisation

By Dr Fiona Greer and Dr Andrew Reason of M-Scan Ltd (UK) and Dr Mark Rogers of M-Scan Inc (US)

The list of biopharmaceuticals produced by recombinant DNA technology, transgenics and synthetic manufacture continues to grow to encompass therapeutic and diagnostic agents ranging from antibodies to vaccines. Regardless of the method of production, all these products present the biotechnology community with complex challenges in manufacture and, ultimately, analytical characterisation.

Dr Fiona Greer, Dr Andrew Reason and Dr Mark Rogers are principals of M-Scan (Mass Spectrometry Consultants and Analysts). Together, they have over 45 years' experience in the analysis of proteins and carbohydrates using mass spectrometry.

Following a PhD in Protein Biochemistry from Aberdeen University, Fiona 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.

Andrew received his PhD in Biochemistry from Imperial College, University of London in 1992 and joined M-Scan in 1994. He is Managing Director of M-Scan Ltd, Wokingham, UK.

Mark received his PhD in Biochemistry from the University of Westminster, London and has been with M-Scan for 13 years. He is currently Vice President of M-Scan Inc, Pennsylvania, US.

Recognising the importance of this topic, a recent international conference was held on 'The Impact of Post-Translational and Chemical Modifications on Protein Therapeutics' (1). At this meeting of leading scientists from academia and industry, a keynote address was given by Dr Keith Weber, Deputy Director, Division of Monoclonal Antibodies, Office of Therapeutics Research and Review, Center for Biologics Evaluation and Research (CBER), FDA. Dr Weber presented the current CBER perspective on post-translational and chemical modifications, emphasising above all that CBER's primary concerns are "safety, purity, potency and strength of bioproducts".

WHY THE CONCERN?

In comparison with small molecule drugs, protein pharmaceuticals are extremely complex entities. As well as microheterogeneities formed during fermentation and downstream processing, the expressed product may differ from the DNA base sequence due to errors of translation, mutation, insertion and deletion, together with

non-protein post-translational modifications such as glycosylation.

The concern is that these modifications will affect the immunogenicity, stability, pharmacokinetics and, most importantly, efficacy of the therapeutic. In addition to these clinical and regulatory worries, there are also important commercial considerations. Detecting and understanding post-translational product changes may lead to modification of the fermentation/purification procedures which would not only eliminate or minimise undesirable changes, but also lead to improved process yields.

TYPES OF MODIFICATION

Almost all protein therapeutics will contain some type of modification – either intentional or unintentional. Intentional changes can be of the *in vitro* type, for example, conjugation (radiolabelling, PEGylation and so on); reduction of thiols; or other treatments which are required following production to make an active therapeutic. As part of the manufacturing process, these changes must be well-controlled and monitored.

The list of unintentional modifications is frighteningly long! These may be inherent in the biological system used (such as glycosylation) or may occur during or after purification (for example proteolytic degradation). During cell culture and processing, exposure to glucose can cause glycation and exposure to urea can result in carbamylation. Table 1 lists just some of the more common post-translational, co-translational, mis-translational (mis-incorporation) and synthetic processing modifications cited in the literature. More comprehensive information may be obtained from websites such as the Association of Biomolecular Resource Facilities (ABRF) (2) and the National Biomedical Research Foundation

Table 1: Common Post-Translational and Other Processing Modifications of Proteins

Acetylation
Acylation
Addition of lipid (palmitoylation, prenylation)
Amidation (deamidation)
Carbamylation
Carboxylation
Formylation
Gla [gamma carboxyglutamic acid]
Glycosylation [N-linked, O-linked]
Glycation
Methylation
Norleucine
Phosphorylation [de-phosphorylation]
Sulphation
Proteolysis
Methionine oxidation
Di-sulphide bond formation

Protein Information Resource (PIR-RESID) (3) databases, and in reference (4).

THE ANALYTICAL CHALLENGE

Dr Weber concluded his presentation by saying that it was important to “know your product as best you can” and that “making it well-characterised requires thorough analysis and good data”. It is clear that an assessment of all post-translational and other changes is important for full characterisation of any protein therapeutic. The dilemma now facing analytical chemists is what kinds of characterisation techniques are appropriate and informative and will provide reliable data which can be validated for regulatory purposes.

Over the past 20 years or so, various mass spectrometric techniques have been applied to the structural characterisation of biomolecules, and as experience has shown, these can provide versatile and sensitive methods of firstly detecting, then analysing, and even quantifying changes in anticipated product primary structure. Unfortunately, there is no one ideal or universal mass spectrometer for this application; all have their advantages and disadvantages (usually cost!), and it is often necessary to utilise more than one type of instrument during a complex biopharmaceutical study.

The most commonly utilised ionisation techniques include electrospray mass spectrometry (ES-MS), online ES-MS (where the MS is coupled to an HPLC, or less commonly CE), matrix assisted laser desorption ionisation mass spectrometry (MALDI-MS) and for derivatised carbohydrates, fast atom bombardment mass spectrometry (FAB-MS) and gas chromatography mass spectrometry (GC-MS).

ES-MS ionisation allows the mass range of ionisable molecules to encompass intact proteins up to approximately 150kDA. This technique is widely applied using small or benchtop quadrupole instruments, often triple-quadrupoles, capable of MS/MS experiments to produce sequence or fragment data as well as molecular weight information. One advantage of ES-MS is that it can be directly coupled to a liquid online sample introduction method such as HPLC, allowing separation and simultaneous detection of the components of complex mixtures.

MALDI ionisation, usually coupled with a time-of-flight (TOF) analyser, and delayed extraction (DE) technology is also capable of intact mass measurement on extremely large molecules

up to approximately 500kDA, together with sequencing applications.

A new generation of instruments has emerged, comprising a quadrupole analyser followed by a collision cell with an orthogonal injection of ions into a TOF analyser. These instruments (for example, Q-TOF and Q-STAR) are currently the most powerful tandem MS/MS sequencing instruments available (5).

Further information on the different types of mass spectrometry ionisation and instrumentation suitable for the analysis of biopolymers can be found in a recent publication of the *Methods in Molecular Biology* series (4).

MASS MAPPING

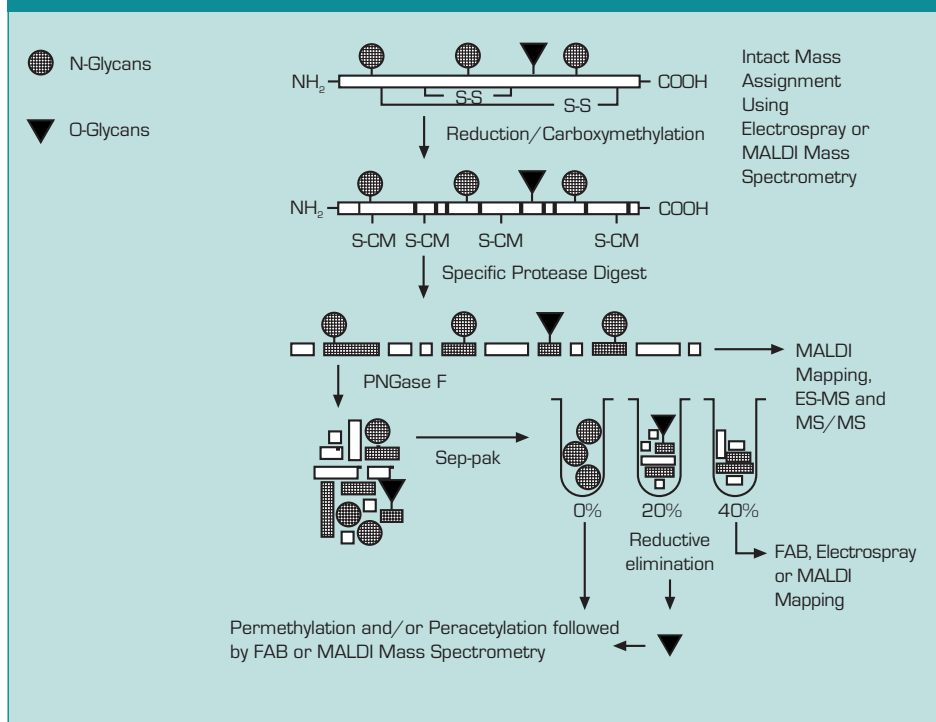
The first step in determining whether a biopharmaceutical product has the correct anticipated structure, for instance that the gene sequence has been correctly translated and that there have been no errors, insertions, deletions or mutations, is usually a simple molecular weight measurement. Depending on the size of the molecule, this is usually performed by MALDI-MS or ES-MS. This measurement would ‘flag’ any discrepancy between the theoretical mass and the actual mass, and depending on the mass range and resolution of the technique, may provide a clue to the type of modification(s).

However, in order to take a closer look at any potential modifications, MS-MAPPING procedures must be carried out. Analogous to LC peptide mapping, the molecule is initially digested into smaller parts using enzymic or chemical means, and then the mixture of peptides produced is analysed using ES, MALDI or FAB-MS. If the mixture is too complex, it can be analysed using online LC-MS, bringing the additional dimension of molecular weight to the peptides separated in the UV profile. In this experiment, differences between the measured masses and the theoretical masses of the anticipated peptides can be quickly observed and even the corresponding peptides isolated and collected for further study. An additional benefit of an MS approach is that the technique relies on measuring mass changes, so that non-protein modifications such as sulphation, phosphorylation or the addition of lipid or carbohydrate, can also be detected.

GLYCOSYLATION

According to Dr Weber, glycosylation is the most common post-translational modification

Figure 1: Strategies for Structural Elucidation Using Mass Spectrometry



encountered in submissions to CBER. It is also a highly complex issue which can be significantly affected by even simple changes in manufacturing processes, such as a pH change of growth media.

Glycosylation often performs a role in the activity of the glycoprotein. The carbohydrate portion of a glycoprotein may target the molecule to a specific location. For example, glycans lacking sialic acid are cleared from the blood and targeted to liver hepatocytes, and a phosphorylated mannose on lysosomal enzymes targets the glycoprotein to a receptor, initiating sorting of the enzymes to the lysosome. Even where the carbohydrate simply performs a structural role (such as holding the active site of an enzyme in the 'correct' 3-D shape), loss or alteration of the saccharide(s) present can have a significant effect on the activity of the glycoprotein. Therefore, structural analysis of the carbohydrate portion of a recombinant glycoprotein is essential to assess the types of glycosylation present on a protein; allow a comparison of the glycosylation on the recombinant product with the natural glycoprotein (if available); and to assess the antennae of the oligosaccharides present for incomplete or antigenic motifs.

Various chromatographic and electrophoretic techniques exist for glycan analysis. However, the same mass spectrometers which have been used to analyse the protein portion of the

molecule can also be used to provide information on the carbohydrate.

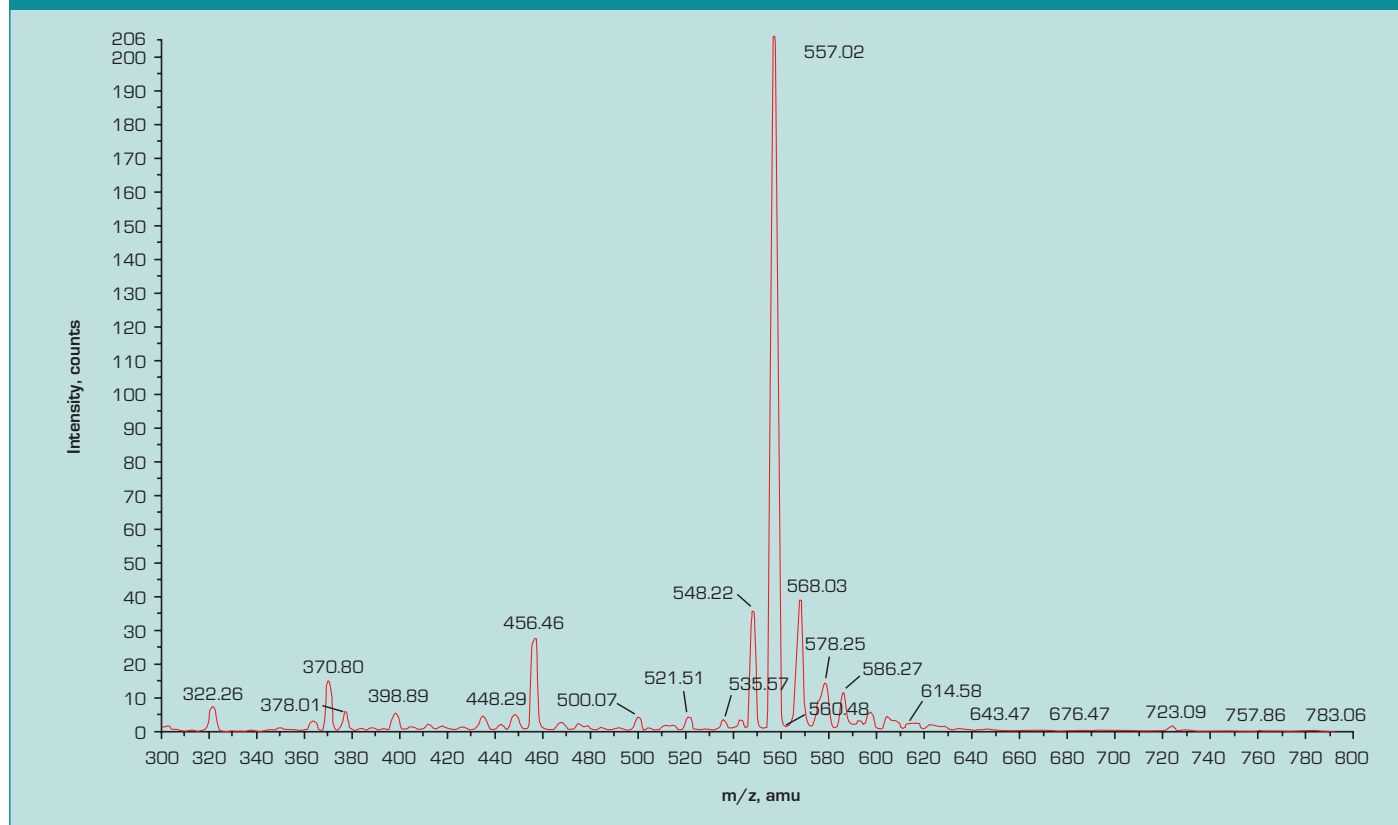
Powerful strategies can be used to analyse both free (un-derivatised) and derivatised samples to determine sites of glycosylation of both N- and O-linked structures; determine the identity of terminal non-reducing ends (potentially the most antigenic structures); and to identify the type of oligosaccharide present (6).

Figure 1 shows a typical strategy which can be used to elucidate the structure of both the protein and carbohydrate portions of a glycoprotein. Following intact mass determination using either ES-MS or MALDI-MS, the glycoprotein can be reduced/alkylated (to cleave disulphide bridges and protect the Cys residues), subjected to proteolysis, then specifically digested with N-glycanase (PNGase F) which will cleave N-linked

glycans from (mammalian) glycopeptides. Sequential elution of fractions from a column will result in a pool of N-Glycan carbohydrate containing molecules, a pool of O-linked glycopeptides and a pool of the remaining peptides. A reductive β -elimination step will remove the O-linked glycans from the peptides.

Both the O-linked and the N-linked glycans can be subsequently analysed using FAB or MALDI-MS following appropriate derivatisation (usually permethylation or peracetylation). The 'mapping' of the peptide portion allows confirmation of the cDNA derived amino acid sequence. Analysis of the peptide/glycopeptide mixture by MS and MS/MS and comparison with the products of deglycosylation allows elucidation of the structure of the oligosaccharides, including location of glycosylation sites as well as the population of oligosaccharides at each individual site if required. To complete the picture, monosaccharide composition analysis can confirm the presence of the carbohydrate, quantify the amounts of monosaccharides present, and broadly assign the class of carbohydrate, such as high mannose oligosaccharide, polysaccharide and so on. Methylation analysis by GC-MS plus the use of specific enzymes can then assign the way in which the monosaccharides are linked together. Enzymatic confirmation of linkage configurations may then allow the full structure of the glycoprotein to be determined.

Figure 2: Parent Ion Scan Detecting a Phosphopeptide in a Tryptic Digest



Carbohydrate analysis is undergoing a period of rapid expansion. This expansion is due to a number of factors including improved instrumentation, the availability of analytical kits which aid the non-specialist, plus a growing awareness of the importance of glycosylation. With the aid of the MS instruments mentioned here, more and more challenging problems and complex strategies for analysis of the increasing number of recombinant biopharmaceuticals can be tackled.

PHOSPHORYLATION

The detection of protein phosphorylation and the identification of the type of modified amino acid residue (serine, threonine or tyrosine) has in the past been a complex undertaking, usually involving some type of radiolabelling (for example ^{32}P). A technique of selective detection of fragment ions produced by collision-induced dissociation during MS/MS experiments allows a specific 'marker' for the presence of this modification.

Alternatively, parent ion scanning from a specific fragment can be used to pinpoint phosphopeptides in complex mixtures which are often obtained during proteolytic 'mapping' studies. Figure 2 illustrates a spectrum of a parent ion scan of a peptide produced from a tryptic digest looking for the parents of m/z 79 – the

characteristic phosphate fragment. This spectrum clearly demonstrates the detection of the $(\text{M}-2\text{H})^2$ at m/z 557 for the expected phosphopeptide. These types of MS investigation may be performed directly or combined with online separation techniques (LC-MS or GC-MS) and provide a means of either rapid screening of multiple samples or structural investigation of a single protein with more than one phosphorylation site.

CONCLUSION

It has been recognised that therapeutic products produced using biotechnological methods are rarely single pure polypeptides or proteins, but a mixture of post-translationally modified molecules. It is important that reliable and sensitive methods are available to detect and identify these modifications. In addition, cost-effective quality control of batch materials is required if the product is to be brought to the market successfully. A variety of mass spectrometric strategies have been developed to enable characterisation of both proteins and glycoproteins and these techniques are routinely used as validated assays to allow characterisation of 'well-specified' molecules. ♦

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

References

- (1) **The Impact of Post-Translational and Chemical Modifications on Protein Therapeutics**, May 2002, San Diego, organised by IBC USA Conferences, Inc
- (2) **ABRF Delta Mass site: A database of protein PT modifications** (www.abrf.org/index.cfm/dm.home)
- (3) **PIR-RESID database of PT modifications** (www.pir.nbrf.georgetown.edu/pirwww/dbinfo/resid.html)
- (4) **Smith BJ (Ed.) Protein Sequencing Protocols, Methods in Molecular Biology, 2002**
- (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**
- (6) **Dell A and Morris HR, Glycoprotein Structure Determination by Mass Spectrometry, Science, 291: pp2,351-2,356, 2001**