Enhanced Sample Preparation: Precision Peptide Mapping and Quantitation Using UHPLC and Mass Spectrometry

Michael Oliver PhD, Thermo Fisher Scientific

Biotherapeutic drugs have seen a significant growth in recent years compared to traditional synthetic small molecule pharmaceuticals. These drugs are highly complex in structure and due to this and the complexities observed during development and manufacture, thorough analysis and safety testing must be performed before any drug goes to market. Peptide mapping of protein biotherapeutics by proteolytic digestion and LC-MS for peak identification and sequence confirmation is an integral aspect of biotherapeutic characterisation. New methods in protein digestion are accelerating this characterisation, simplifying and strengthening safety and efficacy testing.

Introduction

Therapeutic proteins such as cytokines, growth factors, or hormones are engineered in the laboratory to treat cancer, anaemia, hepatitis, and other diseases [1]. These large molecule therapeutics have been successful as drug therapies due to their inherent specificity for a target and minimal risk for toxicity with most molecules [2]. Thus, protein biotherapeutics are increasing in development and use [1]. In addition to their increasing development and use, these compounds show a higher rate of success during clinical trials compared to synthetic small molecule drugs, resulting in a greater percentage of drugs coming to market more quickly.

While biotherapeutics have proven to be effective for a range of diseases, their structural complexity and lengthy production process highlights numerous features that require comprehensive analysis to ensure a safe and efficacious product [3]. Bioproduction and quality control can be monitored with robust procedures that follow strict guidelines set by regulatory organisations such as the European Medicines Agency (EMA) and the United States Food and Drug Administration (FDA) [4].

To ensure patient safety, it is vital that these biopharmaceuticals are comprehensively characterised to confirm that the product is properly defined. Confirmation of protein structure includes peptide maps that detail the entire protein and provide clear evidence of its molecular structure, as well as determining post-translational modifications and sequence variants. This process also provides an understanding of how the biotherapeutic interacts and behaves within the target biological system, as well as enabling identification of signature peptides for quantitation purposes [4].

The analytical techniques used during peptide mapping and quantitation have a reputation for being slow, laborious and subject to reproducibility issues. Current methods such as immuno affinity capture and in-solution digestion often exhibit poor sensitivity and introduce the potential for chemically induced modifications such as deamidation, oxidation, and carbamylation due to the variety of chemicals involved [5]. Even though these methods would benefit from automation to reduce manual errors that could affect data integrity and robustness, most are performed manually by multiple users, impacting regulatory compliance and limiting workflow efficiency and throughput. And since primary emphasis is typically put on actual quantitation, early steps such as sample preparation and digestion tend to see less updates in new technologies. In some instances, ultraviolet (UV) absorption can be used to simplify the characterisation of the biotherapeutic without confirming sequences by mass spectrometry (MS), placing an increased focus on clean sample preparation and precise separation techniques [12].

Complexity in peptide mapping procedures stems from the need to denature proteins and open their structure enough for optimal enzyme interaction. Multiple steps are required to produce a peptide map including reduction, alkylation, and digestion. Given that digestion time varies depending on the size and complexity of a protein, conventional methods could take up to 24 hours to complete [9]. Introduction of alternative methods that improve the quality of quantitative and qualitative analytical results can enhance workflow efficiency and reduce time spent on method development.

New technologies in sample preparation, automation, separation and detection are making the generation of robust and reproducible results easier and faster. Of particular promise for these applications are; peptide mapping and peptide quantitation workflows, with immoblized heat stable enzymes and immunoaffinity capture combines with immobilized heat stable enzymes. These are discussed below.

Experimental
Sample Preparation

Infliximab and rituximab drug products and lyophilised powder of three proteins, cytochrome c, carbonic anhydrase, and recombinant somatotropin, were dissolved in deionised water and adjusted to a final concentration of 10 mg/mL. Two approaches were employed for this evaluation: investigating the effect of using a manual approach to using the novel bead technology and also using an automated approach. The methodologies employed for both approaches are given below.

Manual Digestion

Somatotropin and rituximab were adjusted to 2 mg/mL with deionised water, using 100 µg protein per digestion reaction. The solutions were further diluted 1:4 with Thermo Scientific SMART Digest buffer. Digestion was conducted at 70°C and 1200 rpm with SMART Digest immobilized trypsin beads, incubating somatotropin for 15 minutes and rituximab for 45 minutes to ensure complete digestion of each protein in the shortest time. Immobilised resin was removed using filtration.

Automated Digestion

The Thermo Scientific KingFisher Duo Prime purification system was used to automate the protein digestion. The SMART Digest magnetic resin slurry was diluted and suspended in SMART Digest buffer to create a suspension of 15 µL original resin into 100 µL of buffer in each well of the dedicated 'resin lane' of a 96 deep well plate. 200 µL of 1:4 diluted buffer was prepared in each well of a separate row of the plate as the optional wash buffer. 50 µL each of infliximab, somatotropin, cytochrome c, and carbonic anhydrase were diluted into 150 µL of buffer in the dedicated "incubation lane" to allow for heating and cooling.

Digestion was performed at 70°C, incubating somatotropin for 15 minutes, cytochrome c and carbonic anhydrase for 20 minutes, and infliximab for 45 minutes to ensure complete digestion of each protein in the shortest time. Repeated insertion of the magnetic comb prevented sedimentation of the beads. Immediately after incubation, the magnetic beads were collected and removed from the reaction and the digest solution was actively cooled to 15°C.

UHPLC-UV separation conditions for cytochrome c, infliximab, rituximab, and carbonic anhydrase
Column: Thermo Scientific Hypersil Gold

UHPLC column 1.9 µm, 2.1 × 150 mm
Mobile phase A: Water + 0.05% trifluoroacetic acid
Mobile phase B: Water/acetonitrile/ trifluoroacetic acid (20:80:0.04 v/v/v)
Flow rate: 0.5 mL/min
Column temperature: 70°C (still air mode)
Injection volume: 5 µL
UV wavelength: 214 nm
Gradient: Table 1

Table 1. Mobile phase gradient

Time (min)	%A	%B	Flow (mL/min)
0.0	95	5	0.5
15.0	45	55	0.5
15.1	0	100	0.5
17.0	0	100	0.5
17.1	95	5	0.5
22	95	5	0.5

UHPLC-UV and UHPLC-MS separation conditions for somatotropin and infliximab
Columns: Hypersil GOLD 1.9 µm, 2.1 × 150 mm
Mobile phase A: Water + 0.1% formic acid
Mobile phase B: Acetonitrile + 0.1% formic acid
Flow rate: 0.3 mL/min
Column temperature: 70°C (still air mode)
Injection volume: 5 µL
UV wavelength: 214 nm
Gradient: Table 2

Table 2. Mobile phase gradient

Time (min)	%A	%B	Flow (mL/min)
0.0	96	4	0.3
30	25	75	0.3
30	0	100	0.3
35	0	100	0.3
35	96	4	0.3
45	96	4	0.3

MS Conditions

The Thermo Scientific Q Exactive Plus mass spectrometer prepared with a HESI-II probe was used for mass spectrometric detection using a full MS / dd-MS2 (Top5) experiment.

Ionisation: HESI Positive ion
Scan range: 140 to 2000 m/z
Source temperature: 350°C
Sheath gas pressure: 45 psi
Auxiliary gas flow: 10 arb
Spray voltage: 3.4 kV
Capillary temperature: 320°C
Resolution (Full MS) at m/z 200(FWHM): 70,000

Resolution (MS2) at m/z 200(FWHM): 17,500
Top-N MS2: 5
S-lens RF level: 60
Max inject time: 100 ms
Collision energy (CE): 27

Data Processing and Software

A Thermo Scientific Chromeleon chromatography data system software 7.2 SR4 was used for data processing, while Thermo Scientific Xcalibur software v 2.2 SP1.48 was used for MS data acquisition and BioPharma Finder 2.0 software for protein characterisation.

Results and Discussion
Increased Speed and Reproducibility of Sample Preparation

A fast and highly reproducible workflow has been developed that aids in complete digestion of proteins to all constituent peptides for characterisation and quantitation applications, as illustrated in Figures 1 and 2 where all peptides are shown from full digestion of the different therapeutics. As an alternative to in-solution trypsin digestion, using a heat-stable immobilised trypsin design offers a simplified process and reduces the time required for sample preparation. In-solution digestion requires denaturation, usually by heat or urea, followed by a reduction and then alkylation reaction, after which the reaction is quenched and diluted. Trypsin is then added, and digestion allowed to proceed overnight. As opposed to the more complex in-solution procedure, trypsin is immobilised onto beads and is inactive at room temperature. Once heated, the enzyme becomes active and proteins unfold, allowing cleavage of targeted sites for peptide fragmentation.

This method is thermally stable, and denaturation and digestion occur simultaneously, so there are fewer steps occurring at a faster pace. With this, there is no need for additional reduction or alkylation steps traditionally performed to ensure proteins are fully unravelled for improved digestion. Digestion time was optimised by taking time-points for analysis until no intact protein or large peptides persist. The speed of the method was validated by digesting carbonic anhydrase, where full digestion is achieved in five minutes (Figure 1). Improved sample
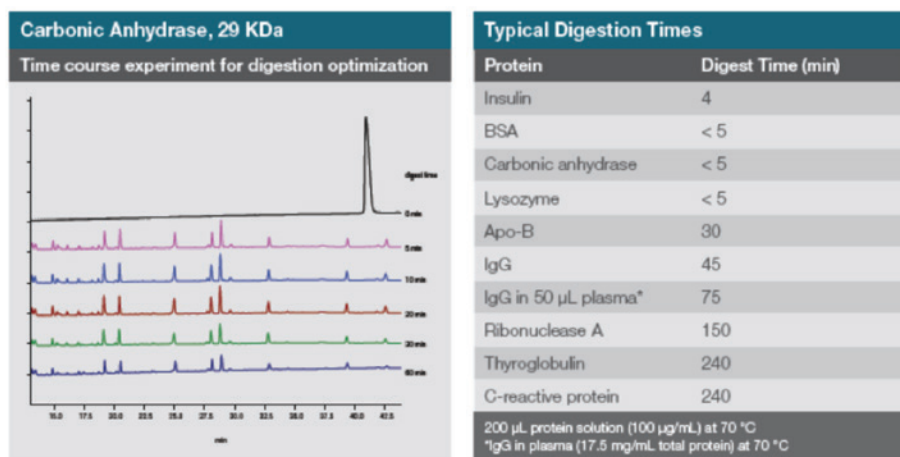
Carbonic Anhydrase, 29 KDa		Typical Digestion Times	
Time course experiment for digestion optimization		Protein	Digest Time (min)
		Insulin	4
		BSA	< 5
		Carbonic anhydrase	< 5
		Lysozyme	< 5
		Apo-B	30
		IgG	45
		IgG in 50 µL plasma*	75
		Ribonuclease A	150
		Thyroglobulin	240
		C-reactive protein	240

200 µL protein solution (100 µg/mL) at 70 °C
*IgG in plasma (17.5 mg/mL total protein) at 70 °C

Figure 1. Digestion optimisation of carbonic anhydrase using manual digestion.

Peak	1	2	3	4	5	6	7	8	9	10	11	12	13
%RSD (A_{rel})	2.75	1.87	2.45	0.71	1.27	1.90	3.60	2.09	2.35	3.92	1.11	0.72	2.42
%RSD (t_R)	0.12	0.03	0.05	0.04	0.03	0.02	0.03	0.03	0.03	0.02	0.01	0.01	0.01

Peak	1	2	3	4	5	6	7	8	9	10
%RSD (A_{rel})	2.12	1.46	1.53	3.03	1.71	1.26	1.51	2.60	1.14	1.64
%RSD (t_R)	0.32	0.1	0.03	0.04	0.02	0.02	0.01	0.01	0.01	0.01
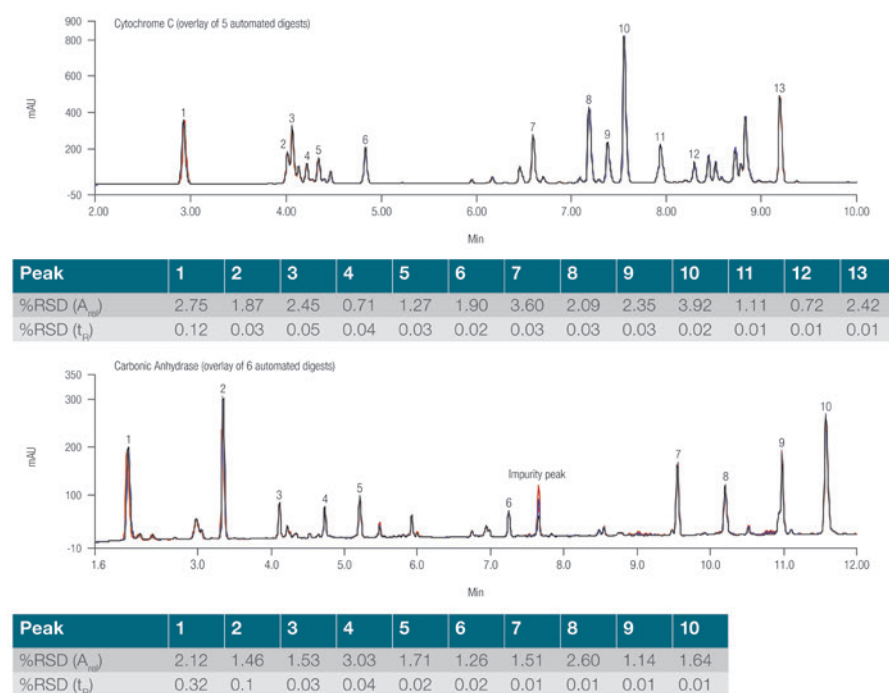
Figure 2. Automated digestion of cytochrome c and carbonic anhydrase. Displaying overlaid peptide maps of replicate digests of cytochrome c (top) and carbonic anhydrase (bottom).

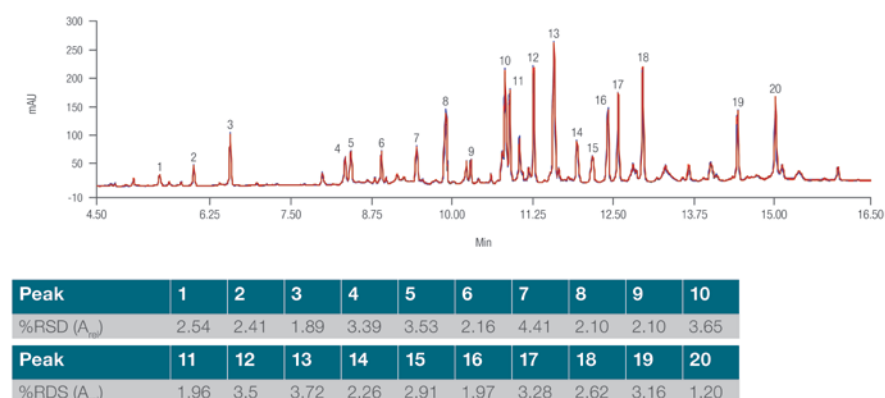| Peak	1	2	3	4	5	6	7	8	9	10
%RSD (A_{rel})	2.54	2.41	1.89	3.39	3.53	2.16	4.41	2.10	2.10	3.65
Peak	11	12	13	14	15	16	17	18	19	20
%RDS (A_{rel})	1.96	3.5	3.72	2.26	2.91	1.97	3.28	2.62	3.16	1.20

Figure 3. Manual digestion of rituximab performed by five different people. Overlaid peptide maps are displayed with percent RSD for each.

preparation with simpler and more efficient digestion can ease downstream separation and detection for peptide mapping processes.

Also fundamental to the success of peptide mapping is method reproducibility, which enables users to confidently assess sample data differences, easing peak integration and interpretation. Reproducibility

can be influenced by all steps of a process, including protein digestion, chromatographic separation, detector performance, and linearity and consistency in data handling [10]. Standardising various steps to create a comprehensive peptide mapping workflow improves reproducibility and increases analytical confidence.

Experimental reproducibility was demonstrated by both automated and manual digestion, using the bead-based methods described above. Automated replicate digests of cytochrome c and carbonic anhydrase generated peptides that were then separated and analysed by UHPLC-UV. Resulting chromatographs were overlaid and found to be nearly identical with average relative standard deviations (RSD) for relative peak areas of 2.08% and 1%, respectively (Figure 2). Supporting this data, potential influence from user variation was tested by manual digestion of rituximab by five different people, some of whom had never performed a protein digestion. Results show an average RSD for peak area of 2.74% across 20 chromatographic peaks, demonstrating the reproducibility and robustness of the protocol between users (Figure 3).

When adjusting sample digestion methodologies, it is important to evaluate the effects of the new method on potential modifications. Since traditional digestions include multiple chemicals and varying temperatures that can possibly cause chemically induced modifications, ensuring that digestion maintains or lessens these variations could affect data outcomes [7].

Modifications were characterised and compared between two methods, one using SMART Digest and one using an in-solution digestion. Rituximab, which consists of 1328 amino acid residues including 16 disulphide bonds, could contain potential modifications of amino acids, deamidation of asparagine or glutamine, and oxidation of methionine or tryptophan [6]. Figure 4 shows the relative abundance of all identified modifications, while Figures 5A and 5B show the extent of amino acid oxidation and deamidation. Summarising the data, Figure 6 displays the relative amount of total modifications measured.

Peptide Separation and Detection

The new digestion protocol combined with high resolution LC separation and high-resolution accurate mass-mass
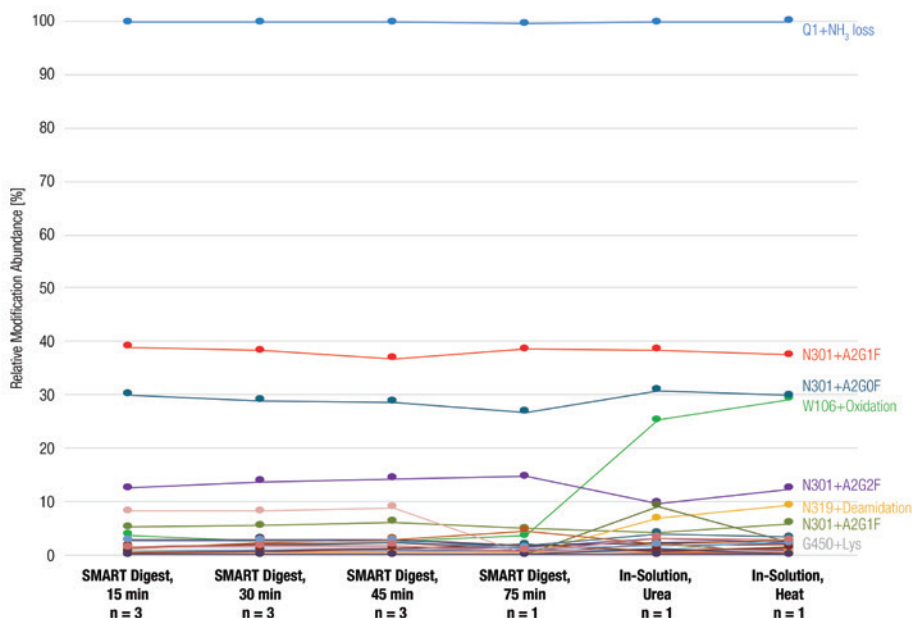
Figure 4. Relative abundance of 85 potential modifications including oxidation, double oxidation, glycation, glycosylation, NH3 loss, isomerisation, lysine truncation, methylation, dimethylation, and carbamylation.

spectrometry (HRAM-MS) was evaluated using the drug product infliximab. Complete digestion of the infliximab antibody and analysis by LC-MS confirmed 100% sequence coverage of the protein, shown in Figure 7 for both the heavy chain and light chain of the antibody. The combined protocol generated high quality data sets that could be effectively interrogated as shown in Figure 7. Additionally, the interpretation of results allows quantitation of modifications, which are important to how the molecule will interact within a biological system [6]. Peptide mapping provides identification, localisation, and quantification of various post translational and chemical modifications that might be present on the amino acid residues [8]. Results confirm that reproducible and complete digestion of infliximab was achieved, validating the procedure for the characterisation and quality control of
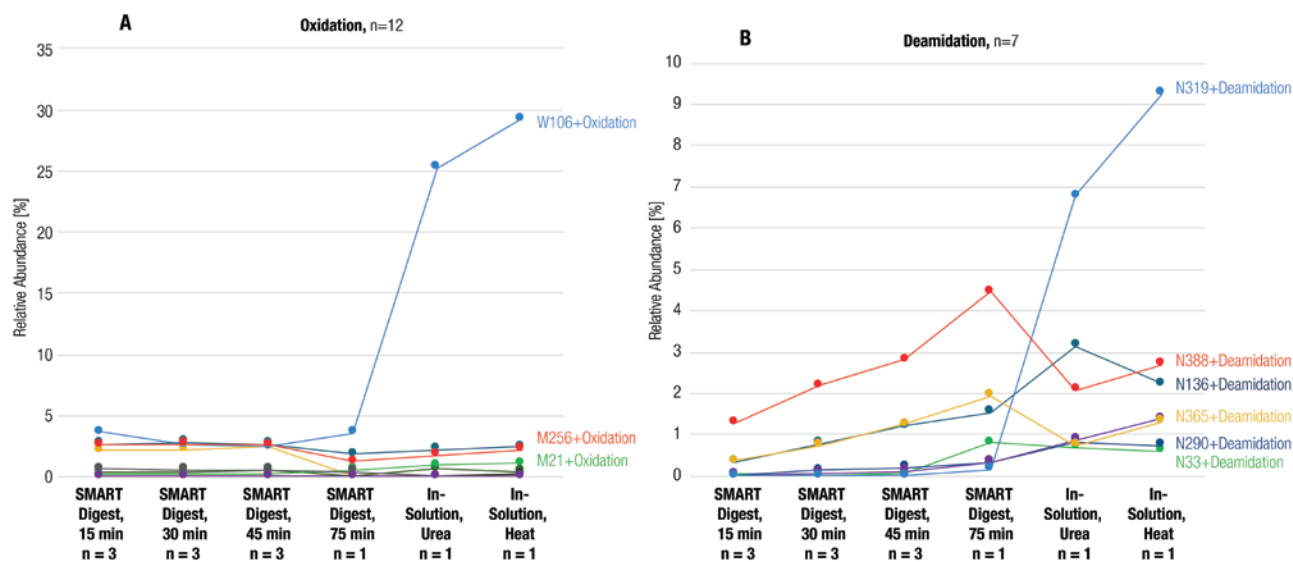
Figure 5. Relative abundance of 12 oxidations (A) and 5 deamidations (B) in different runs with various digestion methods.
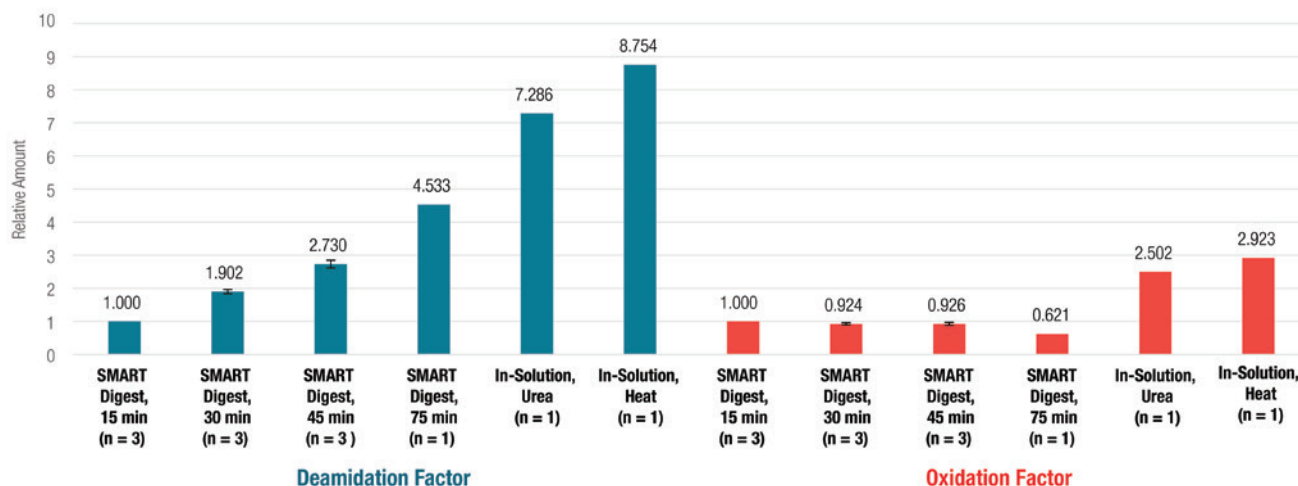

Figure 6. Relative amount of total deamidation and oxidation modifications measured for different digest conditions.
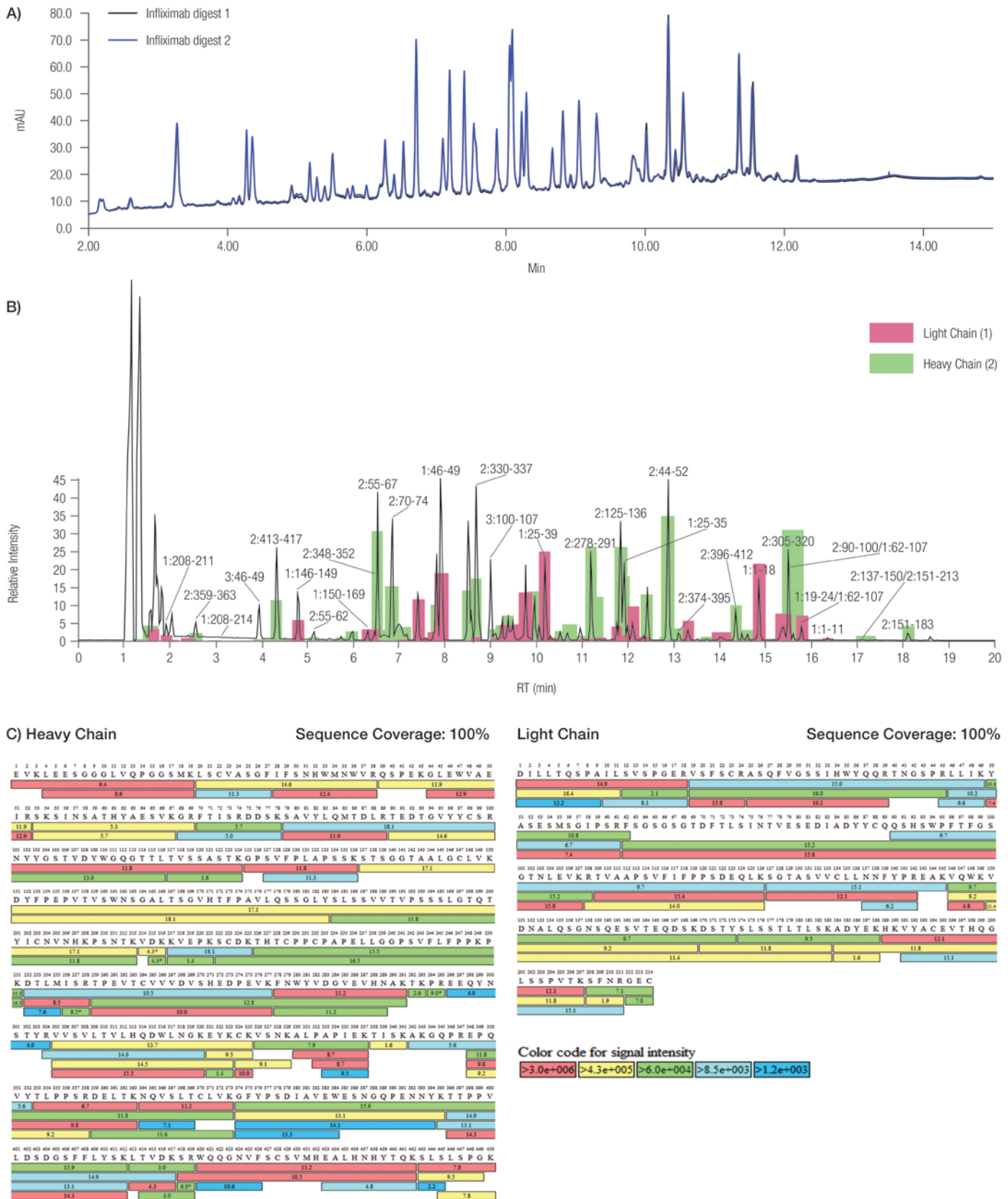
Figure 7. Automated digestion of infliximab using SMART Digest magnetic resin. (A) Overlaid peptide maps for two digests of the infliximab antibody. (B) Total ion chromatogram from infliximab indicating the peptide origin to light (1) and heavy chain (2). (C) Sequence coverage map of infliximab.

protein biopharmaceuticals.

With improved sample preparation techniques resulting in greater reproducibility and speed, MS data was also used to ensure that the digest conditions across both manual and automated digests were optimal and consistent. Correlating the relative peak areas for each digestion approach generates a linear regression curve with a slope of one. This indicates that results from the two digestion approaches are equivalent and reliable (Figure 8).

HRAM-MS detection supports characterisation of target analytes once proteins have been digested and peptides have been separated. HRAM Orbitrap LC-MS generated precise mass measurements for each peptide to generate a peptide map [10]. Compared to lower resolution MS systems that potentially incorrectly identify fragments and significantly impact
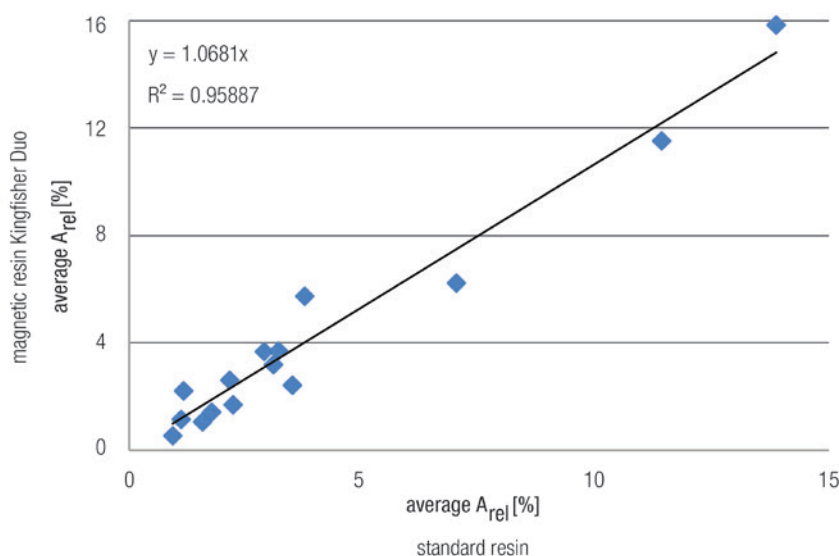
Figure 8. Linear regression curves comparing the peak area results from automated and manual digestion.
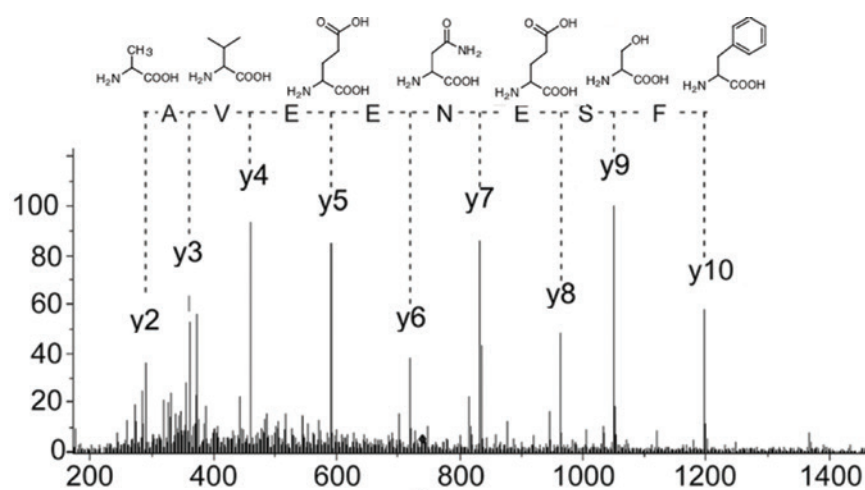

Figure 9. Example of peptide mass spectrum showing amino acid fragment ions

results, high resolution systems offer greater selectivity for robust identification.

Further confirmation could be obtained by fragmenting the individual peptides using LC-MS/MS to isolate and fragment peptides along their backbone producing signature fragment ions by higher-energy collisional dissociation (HCD) in the gas phase if needed. For example, MS parent ion spectra expose the accurate precursor ion masses of the peptides and MS/MS daughter ion spectra reveal the b- and y-ion amino acid fragments of each peptide, as shown in Figure 9.

In addition to mass accuracy and fragmentation of the ion of interest, MS instruments used in peptide mapping workflows must be able to match fast LC run times. Advanced UHPLC systems provided improved separations for these complex peptide samples combined with fast run times for higher throughput. Subsequently, peak widths are often <5 seconds, which then requires a complementary MS system

that can scan fast enough to ensure collection of high quality data. The scan speeds achievable using Orbitrap-based instruments have been demonstrated to exceed the demand for even the most rapid UHPLC separation protocols [11].

Automatic data processing provided quick and in-depth interpretation for a comprehensive view of all data sets, including amino acid sequence confirmation with mass tolerance, modification, identification, and retention time. Data could be examined in multiple displays to enable effective data mining such as sequence coverage maps and chromatographic peak shading that showed relative quantitation of co-eluting peptide peak contributions.

Conclusion

Sample preparation for biotherapeutic protein analysis is a key area where potential issues with sensitivity, reproducibility, and time constraints need to be controlled to

allow robust downstream data analysis [8]. Employing alternative techniques improves all aspects of therapeutics sample preparation, enabling confident results to be generated in less time. Downstream evaluation of LC-MS procedures confirms robust sample preparation and provides additional characterisation of peptides for a complete assessment of the total procedure.

Further method development in sample preparation before digestion combines an immunocapture and the digestion process into a single well. Quantitation studies benefit when immunoaffinity capture is typically used to increase sensitivity by purifying low level proteins from complex biological matrices. Co-immobilising the immunoaffinity reagents and heat-activated thermally stable trypsin onto a single bead allows accelerated digestion under protein denaturing conditions once the enzyme is activated after the binding of a capture reagent to the bead.

References

1. IFPMA – Regulations: https://www.ifpma.org/subtopics/biotherapeutic-medicines/

2. Al Deghaither D, Smaglo BG, Weiner LM. Beyond peptides and mAbs— current status and future perspectives for biotherapeutics with novel constructs. J Clin Pharm. 2015, 55: S3, S4-S20.

3. WHO – Biotherapeutic products: http://www.who.int/biologicals/vaccines/biotherapeutic-products/en/

4. WHO – Regulations: http://www.who.int/biologicals/biotherapeutics/rDNA_DB_final_19_Nov_2013.pdf

5. Jenkins N, Murphy L, and Tyther R. Post-translational Modifications of Recombinant Proteins: Significance for Biopharmaceuticals. Mol Biotechnol (2008) 39:113–118.

6. Nebija, D.; Kopelent-Frank, H.; Urban, E.; Noe, C. R.; Lachmann, B. Comparison of two-dimensional gel electrophoresis patterns and MALDI-TOF MS analysis of therapeutic recombinant monoclonal antibodies trastuzumab and rituximab. Journal of Pharmaceutical and Biomedical Analysis, 2011, 56, 684–91.

7. Ren, Da; Pipes, Gary D; Liu, Dingjiang; Shih, Liang-Yu; Nichols, Andrew C.; Treuheit, Michael J.; Brems, David N.; Bondarenko, Pavel V., An improved trypsin digestion method minimizes digestion-induced modifications on proteins. Analytical Biochemistry, 2009, 392, 12–21.

8. Dick, L.W.; Mahon, D.; Qiu, D.; Cheng, K.C., Peptide mapping of therapeutic monoclonal antibodies: Improvements for increased speed and fewer artifacts. Journal of Chromatography B, 2009, 877, 230-236.

9. Visser, J.; Feuerstein, I.; Stangler, T.; Schmiederer, T.; Fritsch, C.; Schiestl, M. Physicochemical and Functional Comparability Between the Proposed Biosimilar Rituximab GP2013 and Originator Rituximab. BioDrugs, 2013, 27, 495–507.

10. Glatter T., Ludwig C., Ahrne E., Aebersold R., Heck AJR, and Schmidt A. Large-Scale Quantitative Assessment of Different In-Solution Protein Digestion Protocols Reveals Superior Cleavage Efficiency of Tandem Lys-C/Trypsin Proteolysis over Trypsin Digestion. J. Proteome Res. 2012, 11, 5145–5156

11. Strader MB, Tabb DL, Hervey WJ, Pan C, and Hurst GB. Efficient and Specific Trypsin Digestion of Microgram to Nanogram Quantities of Proteins in Organic-Aqueous Solvent Systems. Anal. Chem. 2006, 78, 125-134.

12. Hong P, Koza S, Bouvier ESP. A Review: Size-exclusion chromatography for the analysis of protein biotherapeutics and their aggregates. Journal of Liquid Chromatography & Related Technologies. 2012, 35:20, 2923-2950

Acknowledgements