Automated proteolytic digestion for high precision peptide mapping and targeted bioanalysis

Alexander Schwahn[1], Ken Cook[2], Amy Farrell[3], Kai Scheffler[4], Martin Samonig[4], Jon Bardsley[5], Suraj Patel[2], and Jonathan Bones[3]
Thermo Fisher Scientific, [1]Reinach (Switzerland), [2]Hemel Hempstead (UK), [4]Germering (Germany), [5]Runcorn (UK); [3]NIBRT, Mount Merrion, Blackrock, Co Dublin.

ABSTRACT

We describe a robust, reproducible, high-precision, automated, digestion workflow that is appropriate for use in protein characterisation. Automation provides an ease of use as a general approach to peptide mapping with both LC-only and LC-MS processes. It overcomes the difficulty with reproducibility between users and different laboratories. We confirm the peptide map quality with multiple examples of proteins and show that high sequence coverage with outstanding reproducibility is possible using high-resolution, accurate-mass mass spectrometry.

INTRODUCTION

Peptide mapping has been used for many years in the characterisation of proteins. It has become an important analytical requirement for the testing of biotherapeutic proteins where it is used to measure several critical quality attributes (CQA) required to grantee the safety and efficacy of these drug products. The analysis is used to confirm that the correct sequence has been expressed for the protein and to check for post-translational and chemical modifications. Mass spectrometry (MS) is coupled to liquid chromatography (LC) for peak identification and confirmation of the sequence. However, many quality control (QC) methods use detection by ultraviolet (UV) absorption only after the peaks identities have been confirmed.[1] Although a widely used and accepted technique, in-solution trypsin digestion protocols required for sample preparation have had little standardisation, are labour intensive and prone to manual errors. These errors affect the quality of the analytical data compromising the ability to reproducibly characterize a protein to the required standard. The digestion must be reproducible and the chromatography extremely stable to allow unambiguous peptide identification based on chromatographic retention time [1].
Here we describe the use of an immobilised heat stable Trypsin. Immobilisation allows for automation and an excess of protease to be used. Heat stability allows the digestion to occur under denaturing conditions for the target protein. This work details the automated peptide mapping of several proteins to show a global applicability and reproducibility of the automated digestion protocol and subsequent analysis [2]. The combination of the Thermo Scientific™ SMART Digest™ magnetic beads and the Thermo Scientific™ KingFisher™ Duo Prime purification system was used to automate the digestion process to produce high quality, reproducible peptide mapping data.

MATERIALS AND METHODS

Equipment

Thermo Scientific™ Q Exactive™ Plus, Thermo Scientific™ Vanquish™ Horizon UHPLC System, KingFisher Duo Prime Purification System, SMART Digest Trypsin Kit, Magnetic Bulk Resin option, SMART Digest Trypsin Kit, with filter/collection plate , Thermo Scientific™ Hypersil GOLD™ column 3 µm, 2.1 × 150 mm.

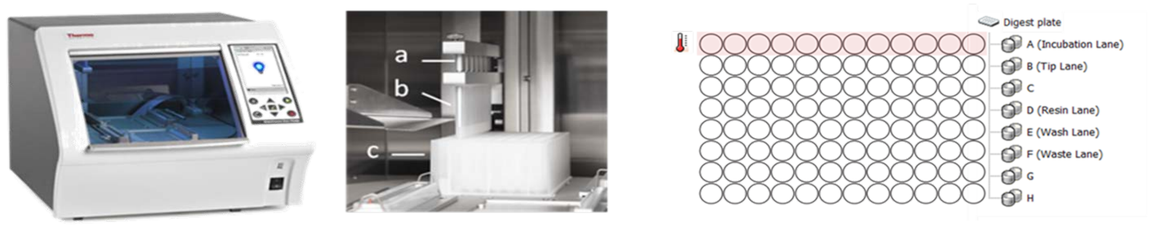
Data Analysis

Thermo Scientific™ BioPharma Finder 2.0, Thermo Scientific™ Chromeleon CDS 7.2, Thermo Scientific™ Xcalibur™ software v 2.2

RESULTS

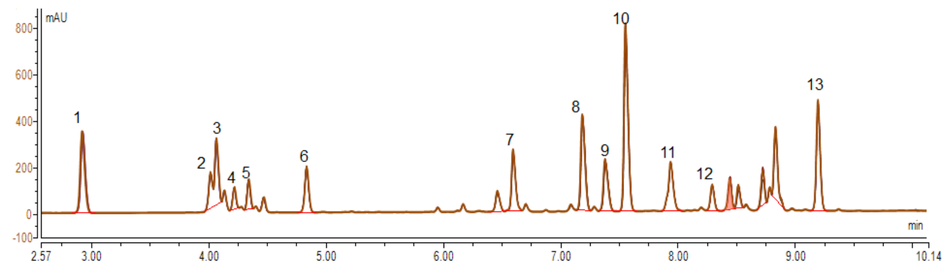
To verify the applicability of the automated trypsin digestion protocols we needed to show ease of use, reproducibility and complete digestion of several proteins using the robotic manipulation of heat stable Trypsin coated magnetic beads.

Figure 1 KingFisher Duo Prime Purification System with well layout for Trypsin Digestion

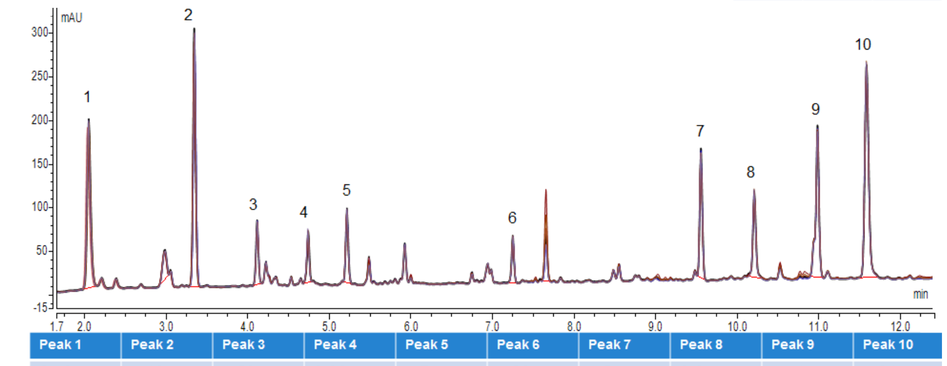


The reproducibility of the automated tryptic digestion with the KingFisher Duo Prime purification system was tested with the proteins Cytochrome C and carbonic anhydrase. Replicate digests [n=5] were conducted and the generated peptides were separated and analysed by UHPLC-UV. The corresponding peptide maps are shown as an overlay of every chromatogram in Figure 2. Both Cytochrome C and carbonic anhydrase were readily digested in 20 minutes at 70 C using the automated SMART Digest kit protocol, resulting in complete digestion of the proteins. An average RSD for relative peak area of 2.08% was achieved for the peaks annotated with Cytochrome C. Several of the peptide produced had peak area RSD values of 1% and below. Carbonic anhydrase gave similar highly reproducible results with an average peak area RSD value of 1.8.

Figure 2. Automated digestion of Cytochrome C and carbonic anhydrase using SMART Digest magnetic resin with the KingFisher Duo Prime system. Overlaid peptide maps of different digests of Cytochrome C (upper panel) and carbonic anhydrase (lower panel). Digest solutions were injected without further purification. %RSD values for relative peak area (upper) and retention time (lower) and are given for the peaks indicated.



Peak 1	Peak 2	Peak 3	Peak 4	Peak 5	Peak 6	Peak 7	Peak 8	Peak 9	Peak 10	Peak 11	Peak 12	Peak 13
2.75	1.87	2.45	0.71	1.27	1.9	3.6	2.09	2.35	3.92	1.11	0.72	2.42
0.12	0.03	0.05	0.04	0.03	0.02	0.02	0.03	0.03	0.02	0.01	0.01	0.01

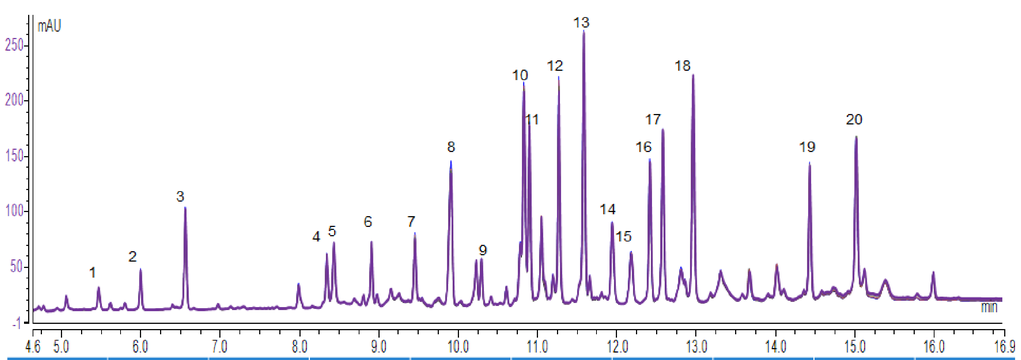


Peak 1	Peak 2	Peak 3	Peak 4	Peak 5	Peak 6	Peak 7	Peak 8	Peak 9	Peak 10
2.12	1.46	1.53	3.03	1.71	1.26	1.51	2.6	1.14	1.64
0.32	0.1	0.03	0.04	0.02	0.02	0.01	0.01	0.01	0.01

This level of reproducibility can be visualized by the high consistency of the Vanquish Horizon UHPLC system gradients and injection accuracy, which gives identical chromatography and makes integration and interpretation of the peaks easier.

This level of reproducibility in protein digestion has never been reported before so the degree of influence between different users was characterized. This is important with different users within the same laboratory and also in method transfer to different laboratories. To assess the robustness and ease of use of the SMART Digest kit protocol in general between different users, an experiment was performed during a protein chromatography workshop with five different people performing a manual digestion using the SMART Digest kit, some of the people involved had never performed a protein digestion before. To make this more realistic the protein chosen for the test was the monoclonal antibody [mAb] Rituximab which is a large protein that is more difficult to digest with the resulting peptide map being much more complicated. Figure 3 shows the results of the experiment.

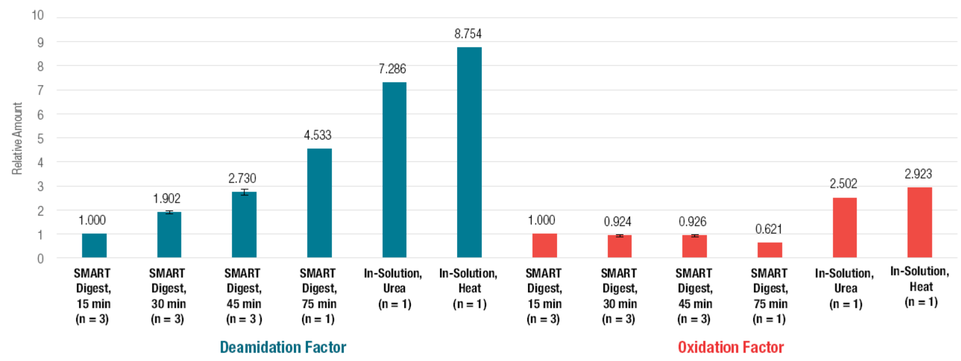
Figure 3. Manual digestion of Rituximab performed by 5 different people. Overlaid peptide maps of the monoclonal antibody Rituximab. Digest solutions were injected without further purification and peptides were separated using the Vanquish™ Horizon UHPLC System. Percentage relative standard deviation (%RSD) values for relative peak area are given for the peaks indicated.



Peak 1	Peak 2	Peak 3	Peak 4	Peak 5	Peak 6	Peak 7	Peak 8	Peak 9	Peak 10
2.54	2.41	1.89	3.39	3.53	2.16	4.41	2.10	2.10	3.65
Peak 11	**Peak 12**	**Peak 13**	**Peak 14**	**Peak 15**	**Peak 16**	**Peak 17**	**Peak 18**	**Peak 19**	**Peak 20**
1.96	3.51	3.72	2.26	2.91	1.97	3.28	2.62	3.16	1.20

The result from multiple users show an average RSD value for peak area of 2.74 over 20 different peaks in a complex chromatogram. Considering that this result was achieved from a protein digestion of a large monoclonal antibody performed by five people, the robustness of the protocol between different users is very apparent. The ease of use is also demonstrated in that some of the digestions were done by people who have no experience with protein digestion techniques.
The automated protocol was further evaluated by LC-MS, a 45 min incubation at 70 ° C enabled the complete digestion of all antibody samples tested and resulted in 100% sequence coverage for both the light and heavy chains.

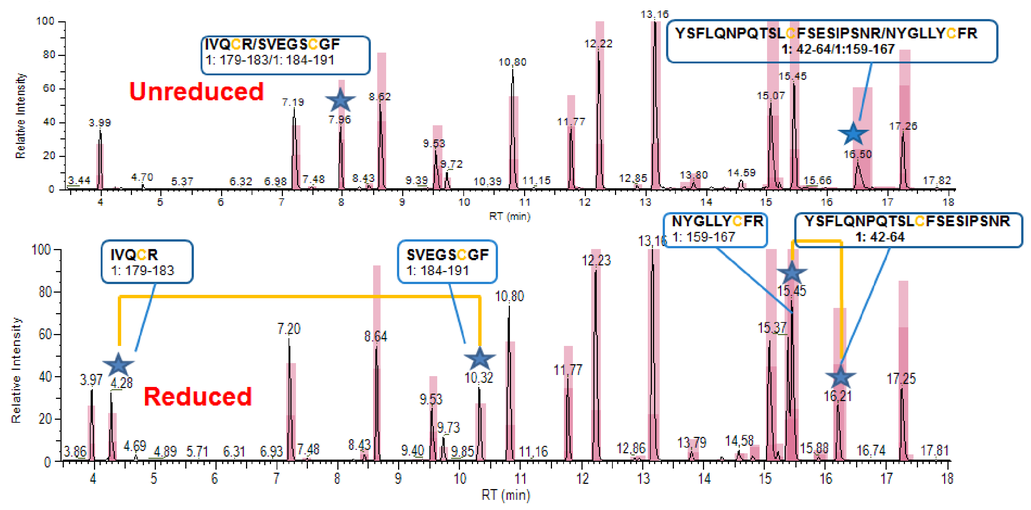
Figure 4. Chemical modifications in a mAb digest



The chemical modifications found in Rituximab SMART digests were analysed to determine the level of modifications that could be induced during the digestion protocol. Deamidation was found to increase over time but still remained at a lower level than compared to an overnight in solution trypsin digest. 45 minutes is the requirement for complete digestion of a mAb which leaves the induced modifications at a very low level. Oxidation does not increase with time and again is lower than compared to a typical overnight digest. Deamidation increases as a factor of time, temperature and pH. The SMART digest is buffered at a lower pH of 7.0 rather than the standard pH 8.0, this and the speed of digestion help to keep modifications low.

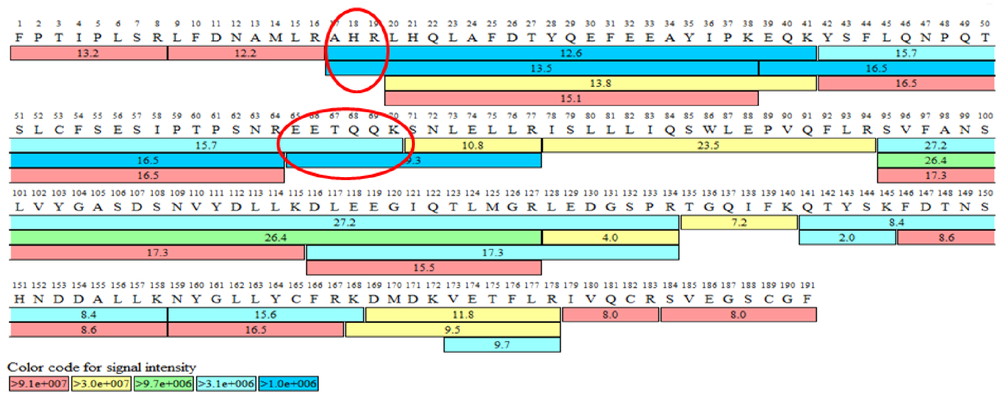
Somatotropin represents a good model system for peptide mapping and intact protein analysis with 18 peptides expected of varying size and two disulphide bond linkages. There are also several sites available for possible post-translational modifications by deamidation, isomerization, and oxidation.
Figure 4 shows the automated identification of the peptide bond linkages using Thermo Scientific™ BioPharma Finder 2.0 software. The digest is done under heat denaturing conditions and so does not require reduction and alkylation of the disulphide bonds beforehand. This allows unambiguous identification of the linkage positions in the protein.

Figure 5. Base peak chromatograms obtained for the unreduced (top) and reduced (bottom) somatotropin trypsin digest. Cysteine containing peptides are labelled with blue stars indicating four disconnected peptides in the reduced sample and two peptides in the unreduced sample bearing disulphide bonds.



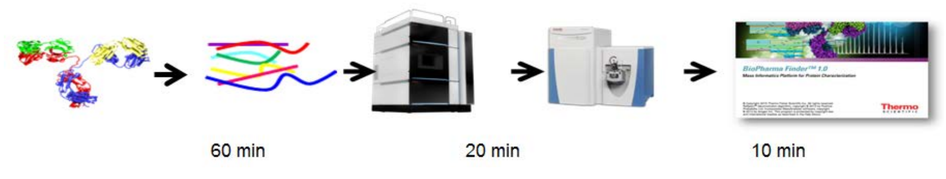
To achieve 100% sequence coverage of a protein in a routine manner every time requires a very precise digestion that is repeatable. Figure 6 shows the coverage Map of Somatotropin which achieves 100% sequence coverage based on the reliable appearance of a very small amount of some missed cleavage products. The areas in the sequence map marked with a red circle are the full cleavage products which are very small hydrophilic peptides eluting in the column void and making detection difficult.

Figure 6. Sequence coverage map for somatotropin showing 100% coverage



The timed schematics of the automated digestion protocol is shown in figure 7. Within 90 minutes the protocol is completed and the results will be available. With an in solution digestion the preparation of all the reagents required to start the digestion may still not be completed.

Figure 7 Outline of the automated Trypsin digestion protocol.



CONCLUSIONS

- **The SMART Digest kit is readily applicable for the characterization and quality control of modern bio-pharmaceuticals and holds the benefits of automation.**
- **Highest reported reproducibility for proteolytic digestion**
- **Proven ease of use with untrained people**
- **Precision of digestion possible with routine 100% sequence coverage**
- **Verification of the correct disulphide bond linkages**
- **Quantification of post-translational modifications**

REFERENCES

1. Samonig M., Schwahn A., Cook K., Oliver M., Swart R., SMART Digest compared to classic in-solution digestion of rituximab for in-depth peptide mapping characterization, Thermo Fisher Scientific Application Note 1159, 2016
2. SMART Digest Kit Technical Guide, Thermo Fisher Scientific, Runcorn UK, 2015

TRADEMARKS/LICENSING