Low ppm detection of Host Cell Proteins (HCPs) in biotherapeutics with optimized Orbitrap-based UHPLC HRAM MS

Amy J Claydon[1], Phil Widdowson[1], Tom Buchanan[1], Janusz Debski[1], Andrew Williamson[2]. [1]Thermo Fisher Scientific, Runcorn, UK. [2]Thermo Fisher Scientific, Hemel Hempstead, UK

ABSTRACT

Purpose: Proteins released from host cells during recombinant biotherapeutic production have unknown effects on drug safety and efficacy. Levels of residual co-purifying host cell proteins (HCPs) must be determined. Traditional immunoassay approaches can provide only limited information on immunologically reactive HCPs, whereas UHPLC-MS enables protein specific identification and quantitation.

Methods: A protein digest standard was spiked into an excess of a BSA protein digest to assess column performance and dynamic range capabilities of the Thermo Scientific™ Q Exactive™ Plus Hybrid Quadrupole-Orbitrap™ mass spectrometer with BioPharma option. Monoclonal antibody (mAb), also spiked with six-protein digest standard, was proteolyzed with magnetic resin immobilized trypsin. Peptides were separated by UHPLC over a 90 min linear gradient. Data were acquired by both DIA and DDA using Thermo Scientific™ Chromeleon™ Chromatography Data System (CDS) Software. Protein identification was performed using Thermo Scientific™ Proteome Discoverer™ software and Spectronaut™ X software (Biognosys AG, Schlieren, Switzerland).

Results: A multiplexed DIA experiment maximized protein coverage and extended MS dynamic range detection compared to DDA analysis of the same samples. When combined with Spectronaut X software, and its integrated search engine Pulsar, comprehensive data mining for identification of low abundance proteins within a biotherapeutic drug product was possible.

INTRODUCTION

The current regulatory gold-standard for assessing HCP levels in recombinant biotherapeutics is a sandwich immunoassay, or enzyme-linked immunosorbent assay (ELISA)[1]. This approach requires suitable anti-HCP antibody reagents to be generated, which can take considerable time and money to develop. Traditional ELISA approaches can also only provide limited information, in that they offer only bulk quantitation of HCPs that are immunologically reactive to the generated antibodies with no possibility of identifying individual proteins.

UHPLC separation combined with HRAM-MS, however, can quantify and identify HCPs on a protein-protein basis, enabling individual HCP monitoring throughout the purification stages[2,3]. The major challenge is to overcome the huge intrasample dynamic range to detect low ppm concentrations of residual HCPs within the concentrated biotherapeutic product. A novel sample preparation strategy, using non-denaturing digestion conditions, reduces the dynamic range of the sample and increases HCP identification when combined with Orbitrap-based MS detection[4].

In contrast with data-dependent acquisition (DDA), data-independent acquisition (DIA) fragments all peptides within comparatively large isolation widths (10 - 50 Da), therefore acquiring MSMS data for all m/z values within the specified mass range. This often leads to fragmentation of peptides that would typically be missed by a standard DDA ('Top N') approach and generates more signal for a single peptide increasing the reliability of relative quantification. The highly complex fragmentation spectra produced by DIA requires a high-resolution mass spectrometer and a spectral library, often generated by DDA analysis of an equivalent sample i.e. harvested culture medium with all potential HCPs at a high concentration.

MATERIALS AND METHODS

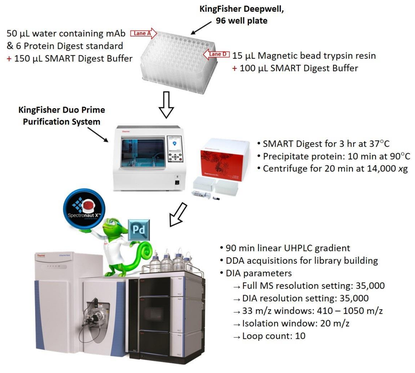
Sample Preparation

To assess the dynamic range capabilities of the Q Exactive Plus mass spectrometer, a vial containing 1 nmol Thermo Scientific™ Pierce™ BSA Protein Digest Standard was reconstituted with 200 µL 0.1% formic acid, and spiked with Thermo Scientific™ Pierce™ 6 Protein Digest Standard, to give an equivalent of 100 ppm and 50 ppm total spiked standard, to BSA.

The same 6 Protein Digest standard was spiked into 500 mg NISTmAb reference material at an equivalent level to 50 ppm of the total equimolar mix to intact NISTmAb. The mixture was added to 150 µL Thermo Scientific™ SMART Digest™ Buffer in a Thermo Scientific™ KingFisher™ Deep-well 96-well plate, and proteolyzed with 10 µl Thermo Scientific™ SMART Digest™ Trypsin, Magnetic Bulk Resin option, under non-denaturing conditions at 37°C for 3 hr, with a Thermo Scientific™ KingFisher™ Duo Prime Purification System (Figure 1). Any undigested protein after this time was precipitated by heating at 90°C for 10 min, then removed by centrifugation for 20 min at 14,000 xg.

A representative mAb, nivolumab (Opdivo®) was spiked with 6 Protein Digest standard (100 ppm to mAb), then proteolyzed using SMART Digest under denaturing conditions (70°C for 45 min), and non-denaturing conditions (37°C for 3 hr) following the same protocol for removal of undigested protein.

Figure 1. Overview schematic of analysis protocol



Test Method(s)

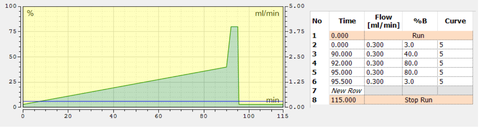
Peptides were separated by UHPLC using a Thermo Scientific™ Acclaim™ VANQUISH™ 250 x 2.1 mm (2.2 µm) C18 UHPLC column over a 90 min linear gradient of 3% - 40% B (mobile phase A: water, 0.1% formic acid, mobile phase B: acetonitrile, 0.1% formic acid) at a flow rate of 0.30 mL/min and a column temperature of 60°C.

An injection volume of 50 µL was used for BSA Protein Digest standard (on-column load of 300 pmol) to ensure spiked 6 Protein Digest standard at 100 ppm and 50 ppm are within the detectable range of the mass spectrometer; 5 fmol and 2.5 fmol respectively. An injection volume of 80 µL was used for NISTmAb reference material (on-column load of 2 mg mAb, if digestion was complete). Additional injections of 6 pmol on-column load 6 Protein Digest standard (25 µL) were analyzed for spectral library building. Data were acquired by both DDA and DIA using Chromeleon CDS software (v. 2.7.8) with a Q Exactive Plus mass spectrometer. DIA parameters are included in Figure 1.

Data Analysis

To enable protein identification using Spectronaut X software, raw data files were exported from Chromeleon CDS software. The DDA acquisitions of the high loading of 6 Protein Digest standard, plus DIA acquisitions of the same sample, were imported into Spectronaut X to generate a spectral library using the integrated search engine Pulsar. A .fasta file of proteins in the 6 Protein Digest standard, created from sequences present in the UniProt database, was assigned. The default search schema was modified to include a maximum peptide length of 75, and a minimum peptide length of 5. The raw files for the remaining DIA analyses were imported into Spectronaut X and searched using the created spectral library.

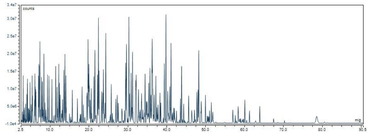
Figure 2. UHPLC gradient conditions



No	Time	Flow [ml/min]	%B	Curve
1	0.000	Run		
2	0.000	0.300	3.0	5
3	90.000	0.300	40.0	5
4	92.000	0.300	80.0	5
5	95.000	0.300	80.0	5
6	95.500	0.300	3.0	5
7	New Row			
8	115.000	Stop Run		

RESULTS

Figure 3. Base Peak Ion (BPI) chromatogram of 6 Protein Digest standard acquired in DDA for spectral library generation using Pulsar



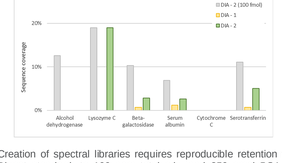
The 6 Protein Digest standard contains an equimolar amount of six proteins (see Table 1 for protein identifiers) that have been reduced and alkylated then proteolyzed with trypsin. The resulting peptide mixture is lyophilized. To assess DIA parameters, decreasing column loads of 6 Protein Digest standard were analyzed to enable comparison in terms of unique peptides detected and % sequence coverage (Table 1).

Table 1. Number of identified peptides using different DIA acquisition parameters with different column loads of 6 Protein Digest standard

DIA-1: m/z range 410 – 890 with fixed isolation window of 20 m/z; DIA-2: m/z range 410 – 1050 with fixed isolation window of 20 m/z

		25 fmol		50 fmol		100 fmol	
		DIA-1	DIA-2	DIA-1	DIA-2	DIA-1	DIA-2
P00330	Alcohol dehydrogenase		1	2	1	3	3
P00698	Lysozyme C		2		2	3	2
P00722	β-galactosidase	1	3	8	6	15	10
P02769	Serum albumin	1	4	3	3	5	4
P62894	Cytochrome C						
Q29443	Serotransferrin	1	4	4	2	7	7

Figure 4. % protein sequence coverage, if ≥ 2 unique peptides, using different DIA acquisition parameters for 25 fmol load of 6 Protein Digest standard (~4 fmol per individual protein)

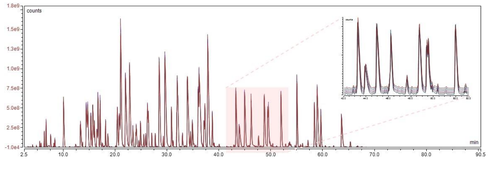


Peptides were more consistently detected at low loading using a wider mass range with fixed isolation windows (DIA-2, green bars). A higher loading of 6 Protein Digest standard (100 fmol) included for comparative purposes.

This also serves to emphasize the importance of instrument LOD when analyzing low abundance HCPs.

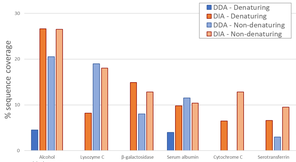
Creation of spectral libraries requires reproducible retention times. BSA was spiked with 6 Protein Digest standard, to 100 ppm total mix, and 250 pmol BSA was injected over multiple replicates. Detection level of 6 Protein Digest standard peptides was maintained, and the column continued to perform reproducibly (Figure 5).

Figure 5. Retention time reproducibility of BSA, despite high column loading



Nivolumab, as a representative mAb, spiked with 6 Protein Digest standard (100 ppm to pmol mAb), was proteolyzed under denaturing and non-denaturing conditions. Peptides were analyzed by DDA using Proteome Discoverer software, and DIA using Spectronaut X. Protein sequence coverage detected with each technique was compared (Figure 6). Although a vial of 6 Protein Digest standard contains 100 pmol protein, it cannot be guaranteed that each individual protein is present at (=100/6) 16.67 pmol, due to digestion variability. However, for the purpose of this analysis, each protein from the standard is assumed to be in equimolar amount at the peptide level, to enable calculation of the approximate mass of protein present for ppm (ng protein/mg mAb) calculations (Figure 6 & Table 2).

Figure 6. Protein coverage for DDA and DIA analyses following denaturing and non-denaturing digestion of nivolumab (left), plus approximate ppm (ng/mg) of each spiked protein (right)

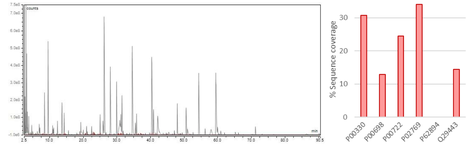


	Protein	MW	ng in digest	ppm
	Nivolumab (Hc & Lc)	146000	1600000	
P00330	Alcohol dehydrogenase	36849	6.76	4
P00698	Lysozyme C	16239	2.98	2
P00722	Beta-galactosidase	116483	21.36	13
P02769	Serum albumin	69293	12.70	8
P62894	Cytochrome C	11704	14.25	1
Q29443	Serotransferrin	77753	2.15	9

Digest conditions do impact sequence coverage of low abundance proteins with DDA. In a non-denaturing digest, much of the mAb remains intact[4], reducing the overall intensity of biotherapeutic peptides during LC-MS analysis. This allows for more 6 Protein Digest standard peptides to be detected, i.e. a higher sequence coverage, for DDA (light blue bars), compared to the same sample digested under denaturing conditions when also analyzed by DDA (dark blue bars).

Overall, higher sequence coverage was acquired with DIA (orange bars) compared to identical sample analyzed by DDA (blue bars). There was also less impact on sequence coverage from digestion conditions with DIA. However, reduction of the high concentration of mAb required to bring HCPs into the detectable range of the mass spectrometer favors the non-denaturing digestion conditions, in terms of amount of material injected onto the column and maintenance requirements of the mass spectrometer.

Figure 7. Overlaid BPI chromatogram of NISTmAb with 6 Protein Digest standard (left), with % coverage of proteins detected in NISTmAb after non-denaturing digestion



NISTmAb was spiked with 6 Protein Digest standard (50 ppm to pmol mAb) prior to non-denaturing proteolysis. Column loading was 11 fmol per individual standard protein. Only Cytochrome C, present at 0.7 ppm (Table 2), remained undetected (Figure 7, right).

Table 2. Approximate ppm (ng protein/mg mAb) of individual 6 Protein Digest standard proteins, spiked at 50 ppm of equimolar mix, in NISTmAb digest

		MW	pmol in digest	ng on column	ppm
	NISTmAb (heavy & light chains)	148,039	3,377	200,000	
P00330	Alcohol dehydrogenase	36,849	0.028	0.41	2.1
P00698	Lysozyme C	16,239	0.028	0.18	0.9
P00722	β-galactosidase	116,483	0.028	1.31	6.6
P02769	Serum albumin	69,293	0.028	0.78	3.9
P62894	Cytochrome C	77,753	0.028	0.13	0.7
Q29443	Serotransferrin	11,704	0.028	0.88	4.4

CONCLUSIONS

- Optimization of DIA parameters, including m/z range, isolation window and loop count, can have dramatic consequences on the number of detected peptides for low abundant proteins.
- Finding a compromise between the optimal DIA parameters and the transient time of the instrument is important for protein quantification if using peak area. Spectronaut X also permits quantification by apex peak height which helps to relax this requirement.
- Due to the complex nature of recombinant mAbs, non-denaturing digest conditions result in a largely intact biotherapeutic that can be precipitated with heat-treatment and removed from the sample prior to MS analysis .
- Incomplete digestion of the high abundance biotherapeutic dramatically reduces the dynamic range of the sample, which is beneficial for detection of low abundant species with DDA detection.
- Optimization of DDA parameters, such as Top N and dynamic exclusion, may increase protein coverage and needs to be assessed.
- Protein digestion, using the SMART Digest trypsin kit, with magnetic resin, in combination with the KingFisher Duo Prime purification system, produces highly reproducible peptide digests, under both denaturing and non-denaturing conditions.
- The automated approach also simplifies the procedure compared to standard in-solution protein digestion protocols.
- To bring any low abundant HCPs into the detectable range of the mass spectrometer, and thus enable detection of individual HCPs at the low ppm level in the final drug product, a column overloading amount of biotherapeutic must be injected. Even under these challenging conditions, the Acclaim VANQUISH C18 UHPLC column continued to deliver high efficiency and excellent separation.
- Combining the DIA data analysis algorithms in Spectronaut X and its integrated search engine, Pulsar, with HRAM Orbitrap data enables high protein coverage and comprehensive data mining for identification of low abundance proteins within a biotherapeutic drug product, using both a denaturing and non-denaturing digestion approach.
- Identification of HCPs in a biotherapeutic drug product requires a suitable spectral library to be generated. This requires samples of the drug product before the final polishing steps, such as the harvest cell culture fluid (HCCF) or post-Protein A fraction. Optimization of false discovery rates (FDR) statistics will also be essential.

REFERENCES

1. USP 41 Published General Chapter <1132> Residual Host Cell Protein Measurement in Biopharmaceuticals, May 2018
2. Houel *et al*. Improving the Dynamic Range of Host Cell Proteins Analysis Using an HRAM Orbitrap Mass Spectrometer. *ASMS Conference Poster,* 2018
3. Füssl *et al*. Comprehensive Characterisation of the Heterogeneity of Adalimumab via Charge Variant Analysis Hyphenated On-line to Native High Resolution Orbitrap Mass Spectrometry. *mAbs,* 2019, 11(1), 116-128
4. Huang, *et al*. A Novel Sample Preparation for Shotgun Proteomics Characterization of HCPs in Antibodies. *Analytical Chemistry*, 2017, 89(10), 5436-5444

ACKNOWLEDGEMENTS

We would like to thank our customers and colleagues in the BioPharma and Product Marketing teams for support. We would also like to thank to the support team at Biognosys for their help with the Spectronaut X software.

TRADEMARKS/LICENSING

PO65502-EN0519S

ThermoFisher SCIENTIFIC