Application Note

Biotherapeutics and Biosimilars

Agilent

Trusted Answers

Host Cell Protein Analysis Using Agilent AssayMAP Bravo and 6545XT AdvanceBio LC/Q-TOF

Authors

Linfeng Wu, Shuai Wu and Te-Wei Chu

Agilent Technologies, Inc., Santa Clara, CA, USA

Introduction

Use of biopharmaceuticals such as monoclonal antibodies (mAbs) has been growing rapidly. Since biopharmaceuticals are generated from biological sources, some of the low-level host cell proteins (HCPs) could remain in the final products even after multiple purification steps. Due to their potential to affect product safety and efficacy, HCP level must be monitored and controlled in drug products according to regulatory requirements[1]. Traditionally, enzyme-linked immunosorbent assay (ELISA) is the standard method of choice for quantifying HCPs in protein therapeutics. However, ELISA lacks the specificity and coverage to identify and quantify individual HCPs. Therefore, LC/MS technologies have become a choice for HCP analysis. The main challenge during LC/MS analysis of HCPs exists in the coelution of low-abundance HCP peptides with the highly abundant peptides from the drug product. This requires better separation of the peptides and broad dynamic range of the LC/MS system.

This study demonstrates an HCP analysis workflow including the Agilent AssayMAP Bravo platform for automated sample preparation, the Agilent 6545XT AdvanceBio LC/Q-TOF for LC-MS/MS analysis, and the vendor-neutral software from Protein Metrics for data analysis (Figure 1). The AssayMAP Bravo platform was used for automated sample preparation, which included protein digestion, desalting, and on-cartridge high-pH reversed-phase (HPRP) fractionation. Both digested samples with or without HPRP fractionation were subjected to LC-MS/MS analysis to examine the platform performance. A new acquisition method, Iterative MS/MS, has been shown to improve protein identification. This standard-flow LC/Q-TOF system demonstrates broad dynamic range, excellent reproducibility, and semiquantitative capability. Our results also showed that all HCPs above 2 ppm were identified with high confidence by coupling on-cartridge sample fractionation with Iterative MS/MS.

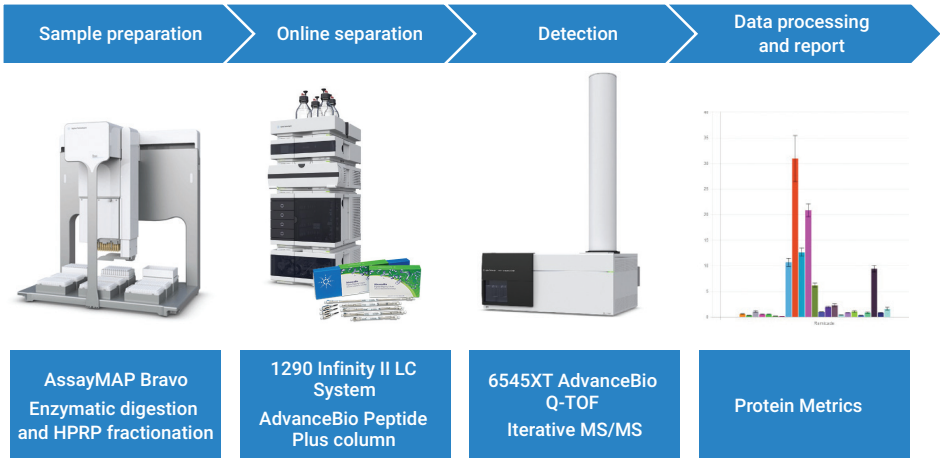

Figure 1. Overview of Agilent HCP Workflow.

Experimental

Materials

Human IgG1 mAb (an R&D product from a partner) was produced from Chinese Hamster Ovary (CHO) cells and purified with protein A. Proteomics Dynamic Range Standard Set (UPS2) was purchased from Sigma-Aldrich.

Instrumentation

- Agilent AssayMAP Bravo system
- Agilent 1290 Infinity II LC system including:
  - Agilent 1290 Infinity II High Speed Pump G7120A
  - Agilent 1290 Infinity II Multisampler G7167B
  - Agilent 1290 Infinity II Thermostatted Column Compartment G7116B
- Agilent 6545XT AdvanceBio LC/Q-TOF
- Agilent Dual Jet Stream ESI source

Sample preparation

UPS2 was spiked into the mAb at a 1:1,000 ratio. A second sample, comprising the mAb without UPS2 spiking, was prepared and analyzed in parallel as a negative control to ensure that the identification of UPS2 proteins was not due to the presence of sequence homologs in the control sample. Both samples were reduced, alkylated, trypsin-digested, and desalted using the AssayMAP Bravo system. The samples were then aliquoted; one aliquot was subjected directly to LC-MS/MS analysis using the Agilent 6545XT AdvanceBio LC/Q-TOF system, and the other aliquot was fractionated using HPRP method on the AssayMAP Bravo system, followed by LC-MS/MS analysis.

HPRP fractionation

The digested sample was fractionated into six fractions on an Agilent Reversed Phase (RP-S) cartridge using the preconfigured Protein Sample Prep WorkBench Fractionation application on the AssayMAP Bravo (Figure 2). RP-S cartridges were first primed with 70 % acetonitrile (ACN), 0.1 % trifluoroacetic acid (TFA) and equilibrated with 0.1 % TFA. Then, 150 µg of digested sample was loaded onto each cartridge, and eluted into six fractions using 10 mM ammonium formate buffer (pH 10) with a stepwise increase of ACN (10, 15, 20, 25, 30, and 90 %).

LC/MS analysis

LC separation was performed on an Agilent AdvanceBio Peptide Plus column (2.1 × 150 mm, 2.7 µm, p/n 675950-902) with 0.4 mL/min flow rate using a 60 minute LC method (Table 1). Each sample was analyzed on the 6545XT AdvanceBio LC/Q-TOF system using either a conventional Auto MS/MS method or an Iterative MS/MS method as indicated (Table 2).

Table 1. Liquid chromatography parameters.

LC Parameters	
Analytical column	AdvanceBio Peptide Plus, 2.1 × 150 mm, 2.7 µm (p/n 675950-902)
Mobile phase A	H_2O, 0.1 % formic acid
Mobile phase B	90 % ACN, 0.1 % formic acid
Flow rate	0.4 mL/min
Gradient	1.0 minutes → 3 %B 50.0 minutes → 21 %B 53.0 minutes → 90 %B 55.0 minutes → 90 %B 55.1 minutes → 3 %B
Stop time	60 minutes
Column temperature	60 °C

Table 2. MS parameters.

MS Parameters	
Drying gas temperature	290 °C
Drying gas flow	13 L/min
Nebulizer	35 psi
Sheath gas temperature	275 °C
Sheath gas flow	12 L/min
Isolation width	Narrow (~1.3 *m/z*)
Dynamic exclusion within run	1 spectrum release after 0.2 minutes
MS Mass range	250–1,700 *m/z*
MS Acquisition rate	10 spectra/second
MS/MS Mass range	50–1,700 *m/z*
MS/MS Acquisition rate	3 spectra/second
Acquisition mode	Iterative MS/MS or Auto MS/MS as indicated
Iterative MS/MS	Mass error tolerance: 15 ppm RT exclusion tolerance: −0.2 minutes to +0.4 minutes

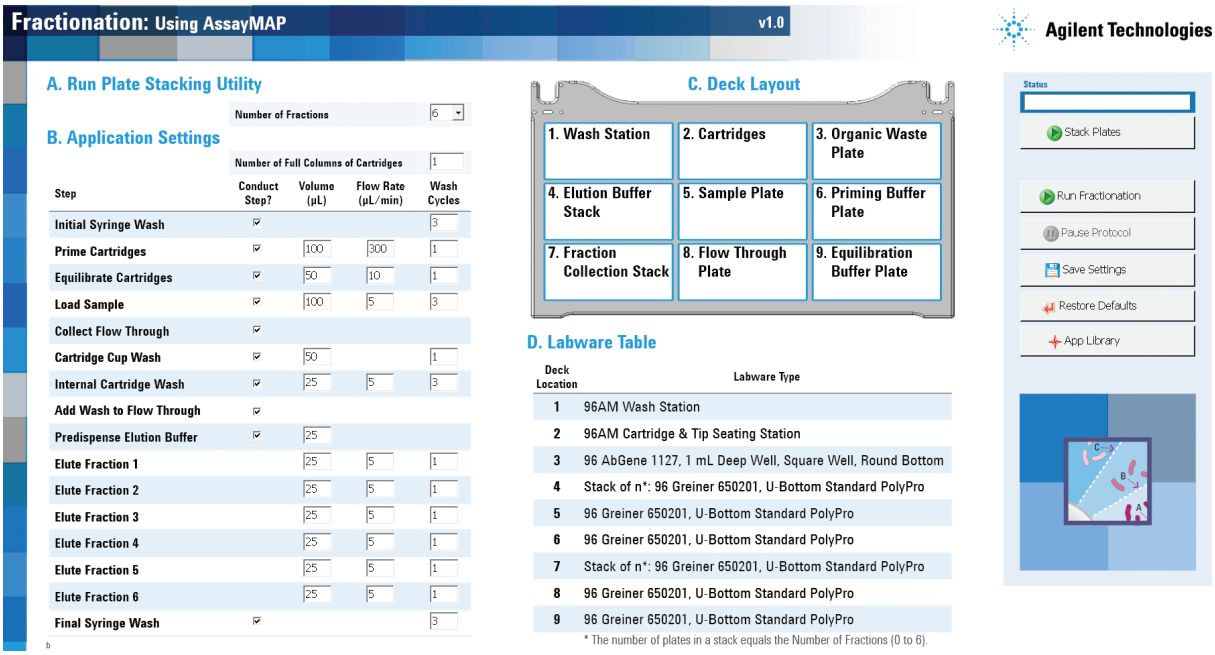

Figure 2. Fractionation application user interface with default settings.

Data processing

The raw data files were processed using software from Protein Metrics Inc. One raw data file was searched using Preview to generate Byonic parameters. All raw data files were searched against a Uniprot CHO K1 protein database that was concatenated with mAb and UPS2 protein sequences using Byonic. The search parameters included semispecific trypsin digestion, up to two missed cleavages, 20 ppm precursor mass tolerance, 30 ppm fragment mass tolerance, fixed cysteine (C) alkylation, variable methionine (M) oxidation, asparagine (N) and glutamine (Q) deamidation, and other variable modifications as suggested by the Preview analysis. The Byonic result files were imported into Byologic for further detailed analysis on a smaller set of protein sequences, and for result reporting. A minimum score of 150 for MS2 search was used to filter peptides.

Results and discussion

Comparison of Iterative MS/MS and Auto MS/MS

The Agilent 6545XT AdvanceBio Q-TOF system provides a new data acquisition method, Iterative MS/MS, which improves identification of low-abundance precursors (Figure 3). Using this method, the protein digest sample was subjected to multiple LC-MS/MS analyses. The first analysis was performed as a conventional Auto MS/MS run using data dependent acquisition. In the following Iterative LC-MS/MS analyses, precursors that were previously selected for MS/MS fragmentation were automatically excluded on a rolling basis with customizable mass error tolerance and retention time exclusion tolerance. As a result, more precursors were automatically interrogated by LC-MS/MS.

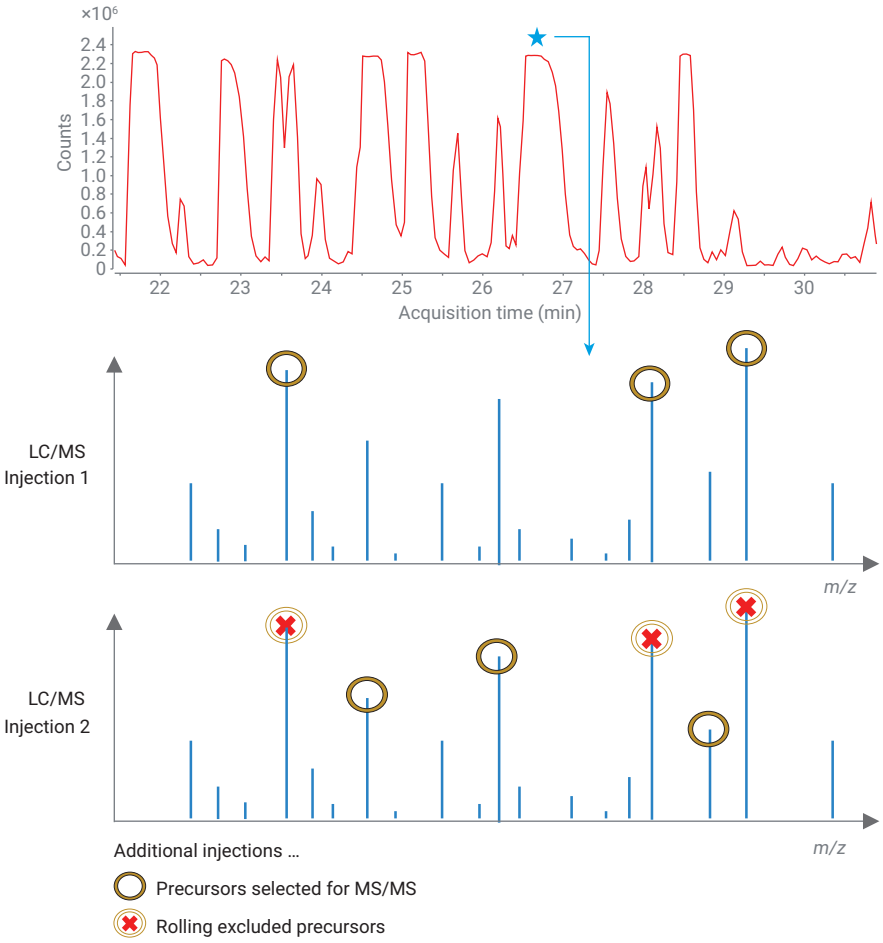

Figure 3. A diagram of an automated Iterative MS/MS acquisition method.

We first compared Iterative MS/MS to Auto MS/MS using the LC/Q-TOF system without any offline fractionation—that is, where the purified mAb with spiked-in UPS2 was analyzed in both methods. UPS2 is the proteomics dynamic range standard produced by Sigma-Aldrich, which is a complex protein mixture containing 48 human proteins at six concentrations ranging from 500 amoles to 50 pmoles. Spiking UPS2 into the mAb sample at a 1:1,000 ratio (w/w) before digestion resulted in protein levels ranging from 0.0004 to 313 ppm, among which 19 proteins were present above one ppm. This spiking sample mimics the wide dynamic range of the HCPs present in the therapeutic proteins. It also allows us to evaluate the sensitivity and dynamic range of different methods being investigated in this study.

Table 3 shows the total numbers of unique peptide sequences from mAb and UPS2 proteins that were identified by each acquisition method. Peptides with the same amino acid sequence but different modifications were counted as one unique peptide sequence. The sample loading amount for each injection was 32 µg, and three injections were carried out for each acquisition method. Overall, Iterative MS/MS identified more unique peptide sequences per protein across a wide dynamic range. All the spiked-in proteins above 8 ppm level were identified with high confidence. All additional results presented here were performed by Iterative MS/MS.

Table 3. Number of unique peptide sequences identified by Iterative MS/MS or Auto MS/MS acquisition methods.

Protein accession	Molecular weight (kDa)	Protein spiking level (ppm)	Three injections per method	
			Iterative MS/MS	Auto MS/MS
mAb_HC	49.7	NA	419	382
mAb_LC	24.0	NA	201	186
ALBU_HUMAN_spike	66.4	313.0	46	43
CAH2_HUMAN_spike	29.1	137.3	19	15
CAH1_HUMAN_spike	28.7	135.6	9	9
LEP__HUMAN_spike	16.2	76.2	4	1
HBB_HUMAN_spike	15.9	74.8	12	10
HBA_HUMAN_spike	15.1	71.3	7	6
UBIQ_HUMAN_spike	10.6	50.0	6	6
CO5_HUMAN_spike	8.6	40.4	4	4
CATA_HUMAN_spike	59.6	28.1	2	2
SUMO1_HUMAN_spike	38.8	18.3	3	1
NQO1_HUMAN_spike	30.7	14.5	2	0
PRDX1_HUMAN_spike	22.0	10.4	3	0
PPIA_HUMAN_spike	20.2	9.5	4	4
MYG_HUMAN_spike	17.1	8.0	2	1

Chromatographic reproducibility and dynamic range

We also assessed the chromatographic reproducibility and dynamic range using the above dataset acquired by three injections of Iterative MS/MS analyses. Figure 4 shows the overlaid chromatograms of the base peak chromatogram (BPC) and Extracted Ion Chromatogram of four coeluting precursor ions from three Iterative LC-MS/MS runs. Due to the high sample loading amount on the column, many of the signals from mAb peptides saturated the MS detector, which showed a plateau in the BPC. Details about those four coeluting peptide precursors are summarized in Table 4. Excellent chromatographic reproducibility was observed for the coeluting peptides, and quantitative reproducibility was excellent across an extremely broad dynamic range (peak intensities from 6.76 × 10^3 to 1.38 × 10^8). For the lowest-abundance myoglobin peptide, which was spiked in at 8 ppm, we observed an RSD value of 10.3 %. These data demonstrated the excellent in-spectrum dynamic range and reproducibility of the 6545XT AdvanceBio LC/Q-TOF system.

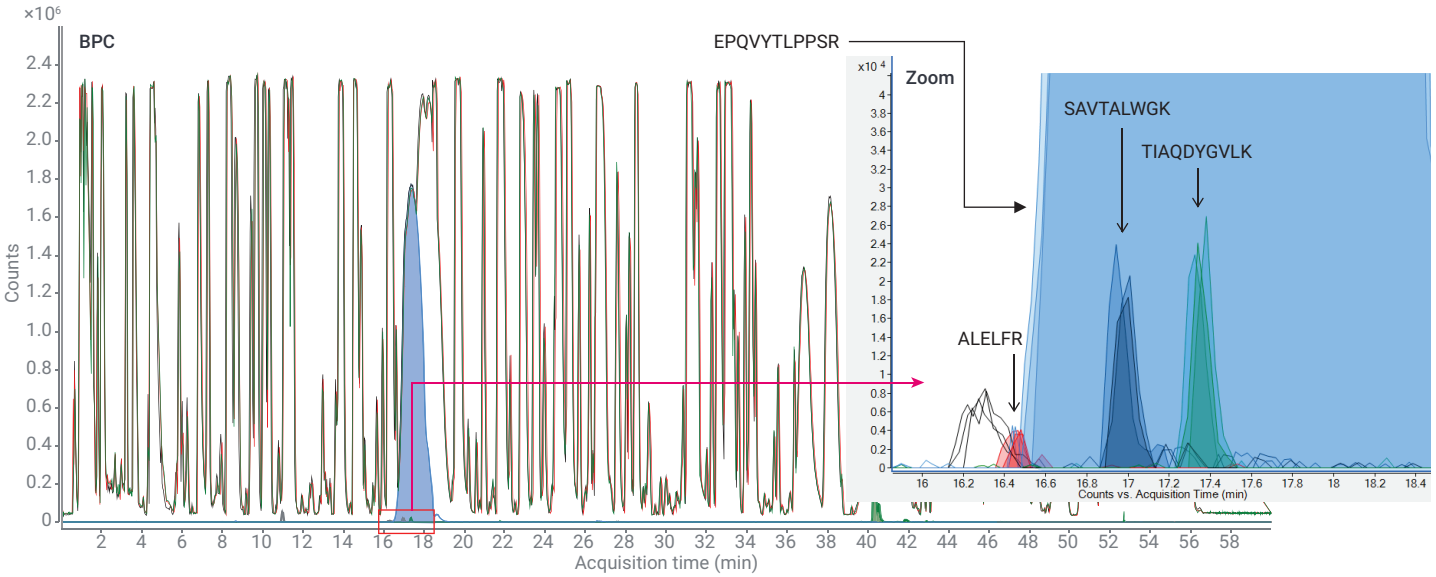

Figure 4. Overlaid chromatograms of three LC-MS/MS runs for BPC and extracted ion chromatograms of four selected peptide precursors.

Semiquantification

We also explored semiquantitative analytical capabilities of this workflow solution using the above data files acquired by three injections of Iterative MS/MS analyses. The total extracted ion chromatogram (XIC) for each identified UPS2 protein using all the measured peptides was normalized to the highest protein in the sample, which is mAb heavy chain in this study, and the values were exported directly from Byologic software. Figure 5 shows the normalized XIC per protein (Y-axis on the left) in the mAb control sample (blue) and the UPS2 spiking sample (orange) plotted with the actual UPS2 protein level in the spiking sample (green, Y-axis on the right). As expected, there are no XIC signals from UPS2 proteins in the mAb control sample, demonstrating the high specificity of this workflow solution. In the UPS2 spiking sample, the normalized XIC was correlated with the actual protein spiked-in level in ppm. These results indicate that the normalized XIC values reported from Byologic software offer a semiquantitative estimate of HCP abundance.

Table 4. Data of the four selected peptide precursors.

Peptide	Precursor ion (*m/z*)	Mass error (ppm)	Intensity	Intensity %RSD	Protein spiking level (ppm)	Protein name
ALELFR	374.7208	-1.1	6.76E+03	10.3 %	8	Myoglobin
TIAQDYGVLK	554.3049	-1.8	1.51E+05	6.2 %	10.4	Peroxiredoxin 1
SAVTALWGK	466.7659	4.8	1.36E+05	6.0 %	74.8	Hemoglobin subunit beta
EPQVYTLPPSR	643.844	1.0	1.38E+08	1.2 %	NA	mAb

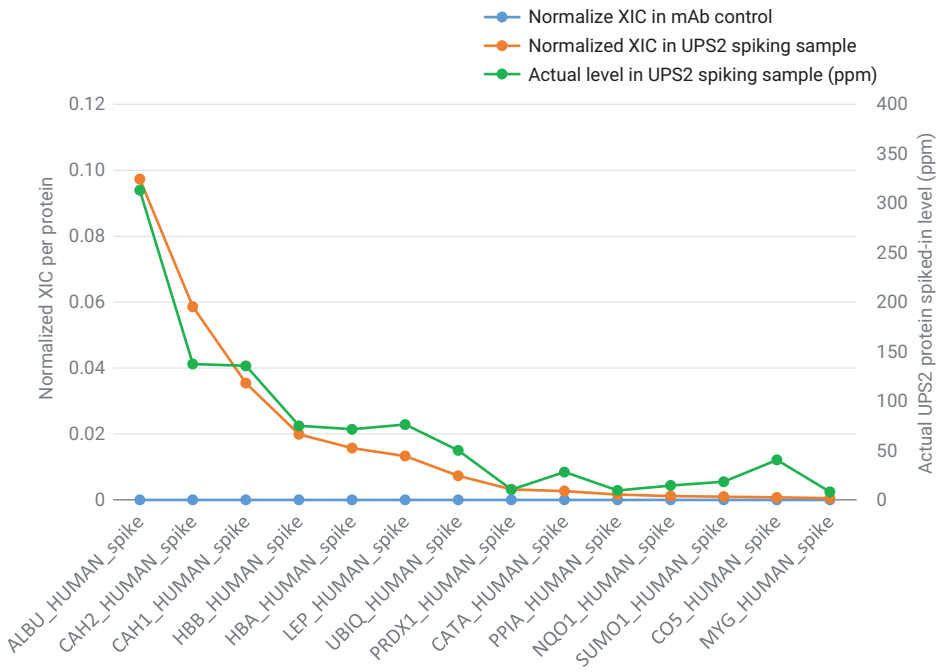

Figure 5. Comparison between normalized XIC per protein and actual UPS2 spike-in level. The XIC per protein was normalized to mAb heavy chain, and their values in mAb control sample (blue) and UPS2 spiking sample (orange) were plotted with the actual protein levels in the UPS2 spiking sample (green).

Improving identification by HPRP fractionation using AssayMAP Bravo

Several core protein sample preparation tasks have been automated using the Agilent AssayMAP Bravo sample preparation platform (for example, protein digestion, peptide cleanup, immuno-affinity purification, and so forth) with micro-chromatography cartridges and task-centric automation protocols[2–4]. With the same automation platform, we demonstrated improvements on identification sensitivity by HPRP fractionation using RP-S cartridges. One hundred fifty micrograms of digested sample were loaded on each RP-S cartridge and eluted into six fractions using an easy-to-use high-pH fractionation protocol. Each fraction was analyzed by two Iterative MS/MS runs. Figure 6 shows the TIC signal from the LC-MS/MS analysis of unfractionated sample compared to HPRP eluates, which demonstrated the chromatographic separation on the RP-S cartridge.

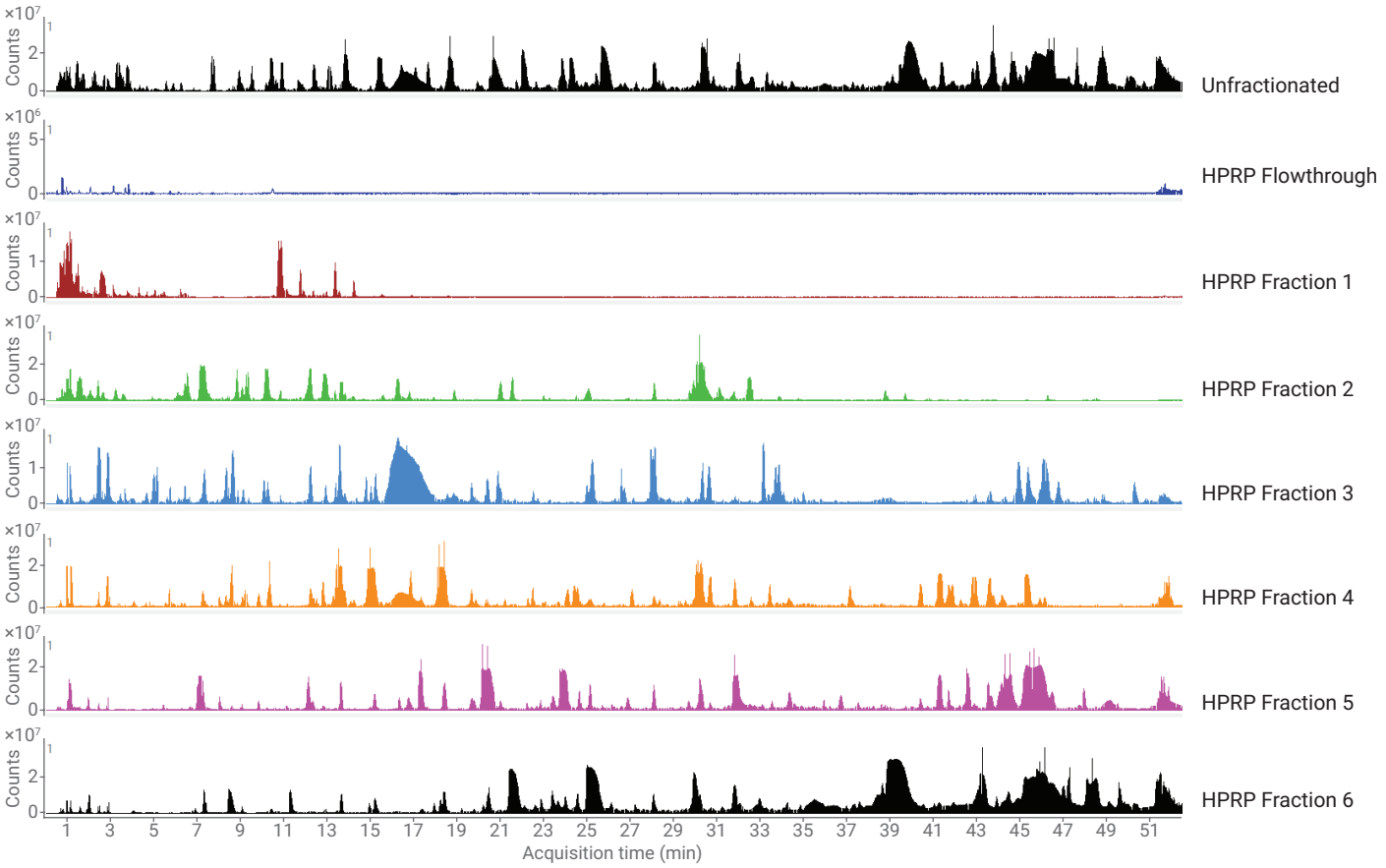

Figure 6. TIC signal from unfractionated and HPRP fractionated sample.

The MS/MS spectra of peptides from low-abundance spiked-in proteins were manually inspected to ensure high spectral quality and no homolog contamination from endogenous CHO HCPs. All the spiked-in proteins above 2 ppm were identified with high confidence, demonstrating an improvement in the identification sensitivity using HPRP fractionation (Table 5, compared to Table 3). Figure 7 shows an example MS/MS spectrum for peptide IEEIFK from the human KCRM protein, which was spiked in at 2 ppm.

Table 5. Identification of unique peptide sequences from the UPS2 spiking proteins using AssayMAP HPRP fractionation coupled with two Iterative MS/MS runs per fraction.

Protein accession	Molecular weight (kDa)	Protein spiking level (ppm)	No. unique peptide sequences
ALBU_HUMAN_spike	66.4	313.0	79
CAH2_HUMAN_spike	29.1	137.3	32
CAH1_HUMAN_spike	28.7	135.6	15
LEP_HUMAN_spike	16.2	76.2	6
HBB_HUMAN_spike	15.9	74.8	22
HBA_HUMAN_spike	15.1	71.3	14
UBIQ_HUMAN_spike	10.6	50.0	9
CO5_HUMAN_spike	8.6	40.4	6
CATA_HUMAN_spike	59.6	28.1	14
SUMO1_HUMAN_spike	38.8	18.3	11
NQO1_HUMAN_spike	30.7	14.5	8
PRDX1_HUMAN_spike	22.0	10.4	9
PPIA_HUMAN_spike	20.2	9.5	11
MYG_HUMAN_spike	17.1	8.0	2
CYB5_HUMAN_spike	16.0	7.6	2
EGR_HUMAN_spike	6.4	3.0	1
SYHC_HUMAN_spike	58.2	2.7	5
KCRM_HUMAN_spike	43.1	2.0	3

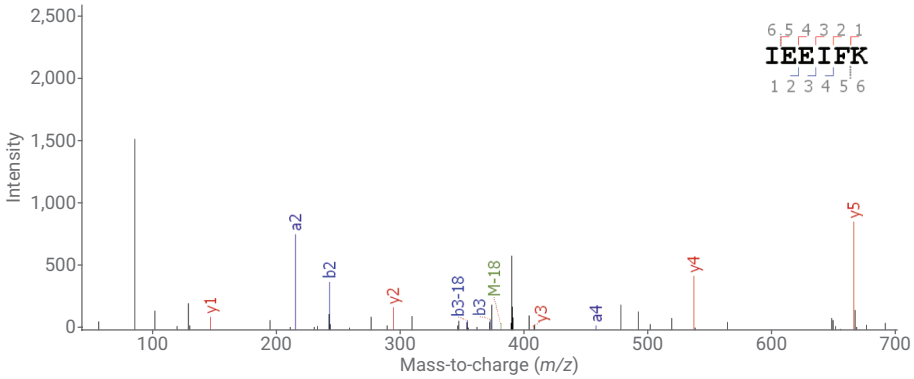

Figure 7. MS/MS spectrum of peptide IEEIFK originated from the KCRM protein spiked at 2 ppm.

The impact of HPRP fractionation on identifying endogenous CHO HCPs was also examined. Figure 8 shows the number of identified CHO proteins between unfractionated and HPRP fractionated samples at a 1 % protein false discovery rate. The unfractionated sample was analyzed by three Iterative MS/MS runs, and the HPRP fractionated samples were analyzed by two Iterative MS/MS runs per fraction. The results show a more than three-fold increase using HPRP fractionation coupled with Iterative MS/MS (138 versus 38).

All the above results demonstrate the advantage of on-cartridge HPRP fractionation in improving identification of HCPs.

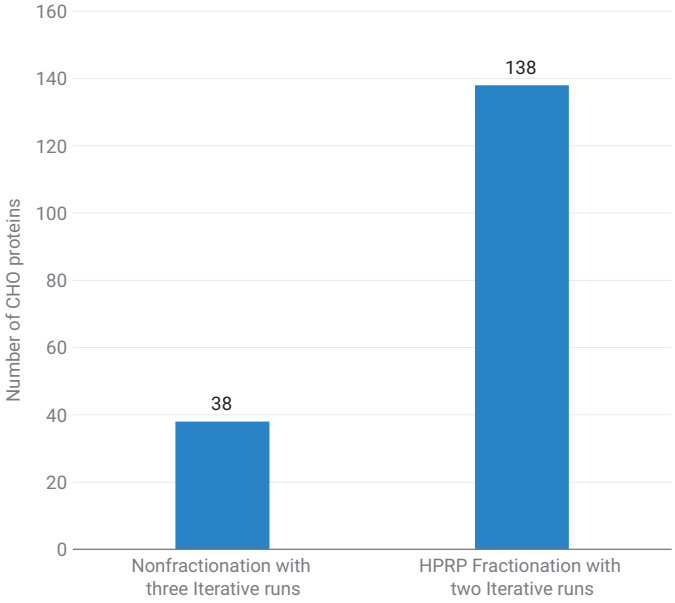

Figure 8. Comparison of number of identified endogenous CHO HCPs between unfractionated and HPRP fractionated samples coupled with Iterative MS/MS acquisition.

Conclusions

A host cell protein analysis workflow including the AssayMAP Bravo platform for automated sample preparation, the 6545XT AdvanceBio LC/Q-TOF for LC-MS/MS analysis, and the vendor-neutral software from Protein Metrics for data analysis, has been demonstrated.

- The AssayMAP Bravo platform using microchromatography cartridges and task-centric automation protocols has brought unprecedented reproducibility, scalability, flexibility, and ease-of-use to sample preparation automation.
- Iterative MS/MS improves protein identification coverage. Using LC-MS/MS coupled with Iterative MS/MS acquisition, all the spiked-in standard proteins above 8 ppm were identified with high confidence.
- Broad dynamic range and excellent reproducibility were demonstrated on this standard-flow LC/Q-TOF system.
- By adding HPRP fractionation using AssayMAP Bravo, all the spiked-in proteins above 2 ppm were identified with high confidence.
- Data files acquired by data-dependent acquisition allow simple and fast protein identification and semiquantification.

References

1. ICH Q6B Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products.
2. Automation for LC/MS Sample Preparation: High Throughput In-Solution Digestion and Peptide Cleanup Enabled by the Agilent AssayMAP Bravo Platform. *Agilent Technologies*, publication number 5991-2957EN.
3. Automation of Sample Preparation for Accurate and Scalable Quantification and Characterization of Biotherapeutic Proteins Using the Agilent AssayMAP Bravo Platform. *Agilent Technologies*, publication number 5991-4872EN.
4. Agilent AssayMAP Bravo Technology Enables Reproducible Automated Phosphopeptide Enrichment from Complex Mixtures Using High Capacity Fe(III)-NTA Cartridges. *Agilent Technologies*, publication number 5991-6073EN.

www.agilent.com/chem

For Research Use Only. Not for use in diagnostic procedures.

Printed in the USA, July 2, 2018
5991-9300EN

Agilent
Trusted Answers