The Optimization of Host-cell Protein Detection using data-independent SWATH-MS

Milla Neffling[1]; Randy Arnold[2]; Eric Johansen[2],;Remco Van Soest[3], Sahana Mollah[2]; and Christie L. Hunter[2]
[1] AB SCIEX, Warrington, UK; [2]AB SCIEX, USA, [2]Eksigent Technologies, USA

INTRODUCTION

Monoclonal antibodies (mAb) are major target-oriented biotherapeutics to treat an array of human diseases. mAbs are typically produced in biological systems such as Chinese hamster ovary (CHO) or other cell lines. Heterogeneity of IgG proteins due to post-translational modifications (PTMs), sequence variants, degradation products, and contaminants (such as host cell proteins) must be characterized completely to understand purity, stability and potency of the mAb product, and to avoid immunogenicity. Mass spectrometry is a superior method for the characterization of mAbs. Here we explore a data independent analysis approach that provides benefits over other MS strategies because it captures comprehensive quantitative MS/MS chromatograms of every fragment ion from a given sample that can be mined extensively post-acquisition.

Figure 1. Integrated Solution for Protein Characterization from Complex Proteomics Samples.

MATERIALS AND METHODS

Sample Preparation: 1 mg of monoclonal mouse IgG1 mAb was denatured with urea, reduced with TCEP, alkylated with MMTS and trypsin digested. A constant concentration of this digest was spiked with a commercial six-protein digest mixture of serum albumin, lactoperoxidase, carbonic anhydrase, glutamate dehydrogenase, alpha casein and lacto-globulin at a range of concentrations shown in Table 1.

Chromatography: Samples were analyzed using the Eksigent ekspert™ 425 System and a ChromXP column (0.5mm x 10cm C18-CL 3µm 120Å) Elution gradients of 5-35% B at 20 µL/min in 24 min were run with the column at 45 °C. Solvent A is 2% acetontrile; solvent B is 98% acetonitrile, both with 0.1% formic acid.

Mass Spectrometry: Unmodified and spiked mAb digests were analyzed using a TripleTOF® 5600 system. An information dependent acquisition (IDA) LCMS/MS method was used for initial peptide identification and to generate a peptide library for SWATH™ acquisition data interrogation. This IDA method consisted of a high resolution TOF MS survey scan followed by 20 MS/MS in a second with a minimum accumulation time of 50 msec. SWATH™ data-independent acquisitions were subsequently performed in triplicate on each sample, using a 20 Da Q1 window width, to obtain quantitative MS/MS chromatograms for every precursor ion between 400 and 1200 m/z.

Data Processing: IDA data were searched using ProteinPilot™ Software against a database containing the sequence of the antibody and the sequences of the model "host" cell proteome. Quantitative analysis was performed using the search results as a peptide library to inform SWATH™ peptide fragment ion chromatogram extraction using the SWATH™ Acquisition tool inside of PeakView® Software.

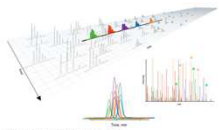
Level	Serum Albumin	Lacto-peroxidase	Carbonic Anhydrase	Glutamate Dehydrogenase	Alpha Casein	Lacto-globulin
1	415	485	182	350	148	114
2	207	242	91	175	74	57
3	104	121	46	88	37	29
4	52	61	23	44	18	14
5	26	30	11	22	9	7

Table 1: Model HCP levels in samples, expressed as parts per million.

MS/MS^{ALL} with SWATH™ Acquisition

What is it?

- **MS/MS^{ALL}**
  - A ***data-independent*** workflow enabled by TripleTOF® system technology that acquires high resolution quantifiable MS/MS data for all detectable analytes in a complex sample, in a single run



How does it work?

- SWATH™ Acquisition
  - Uses wide isolation windows stepped across a mass range, collecting high resolution composite MS/MS spectra in a chromatographic time scale

What does this enable?

- Data processing by generation of post-acquisition fragment ion XICs at high resolution for quantitation with confirmation of identity
- Quantitation and confirmation of everything in the sample
- Digital record of everything in your sample
- Single method for acquiring all your data

Single SWATH™ Acquisition Window

- Current strategy uses 25 Da window to cover the peptide mass range in a LC time frame
- 3 D data
- MS/MS on all precursors between 550 – 575 m/z

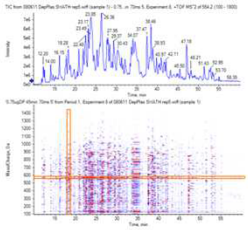


Figure 5. Example SWATH™ Data. Visual Depiction of a single SWATH™ acquisition data set. The chromatogram and heat map are from just the *m/z* 550-575 SWATH™. Inside this single acquisition there are 23 other 3-dimensional data sets just like the above. In the heat map, the X axis represents time, the Y axis is m/z, and intensity is represented by color. The horizontal red box indicates the q1 selection window, all ions outside that box are fragments. Each vertical stripe is an MSMS spectrum, for example, inside the vertical box.

PeakView™ SWATH™ Analysis Tool

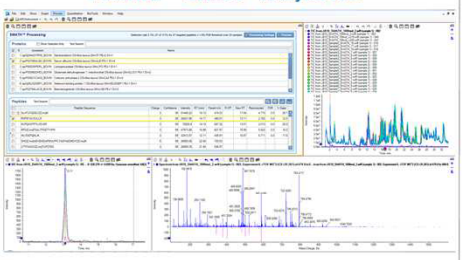


Figure 6. SWATH™ Analysis Software. Top Left: List of proteins and peptides from an ion library (in this case a Protein Pilot™ Group File from an IDA run of the measured sample. Top Right: TIC chromatograms of each SWATH™ data file. Bottom Left: XIC chromatograms of six fragment ions from the peptide selected in the Top Left pane. Bottom Right: Mirror plot showing the MSMS Spectrum collected at the top of the chromatogram in the bottom left pane (Blue) over the Spectrum from the ion library (Pink).

RESULTS

In this data independent workflow, the Q1 quadrupole is stepped at increments across the target mass range in 20 Da steps. Transmitted ions from the 20Da wide window are fragmented in the collision cell and fragments are analyzed in the TOF MS analyzer at high resolution. This is done in an LC cycle time, such that MS/MS spectra are acquired on every peptide in a sample. High resolution XICs are then generated post-acquisition for quantification. Profiling the heterogeneity of this mAb we were able to observe 100% sequence coverage for the Light Chain and 99.5% sequence coverage for the Heavy Chain. For each of the model MCPs, three or four peptides were used for SWATH™ quantitation.

Despite up to a 100,000 fold difference in abundance between the product and the contaminants, the quantitation of both HCPs and antibody was highly reproducible. At higher load levels, CVs were all between 3-7%; and even at the lowest level, where the HCP proteins were present at 7-30 ppm, the average CV was 7%, and the most variable protein had a CV of 12% (Figure 8). The data indicates that we can detect our "host cell protein" at levels far below 0.01% contamination. Using the quantitative SWATH™ methodology, in a single sample run we can acquire MS/MS data on every fragment ion from every precursor peptide ion between *m/z* of 400 and 1200. By examining the data retrospectively, we can quantify the extent of host cell protein contamination with MRM-like fidelity and sensitivity, without any up-front method development or foreknowledge of the PTM or contaminating protein.

SWATH™ Analysis: MarkerView™ Results - Protein Data

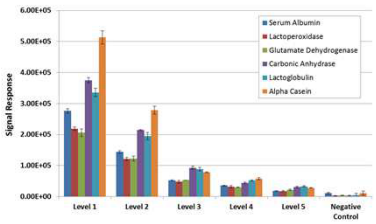


Figure 7. Protein level signal responses. Peptide level signals for each protein were summed, providing a direct measure of concentration. Values shown are the mean and standard deviations (error bars) of triplicate analyses.

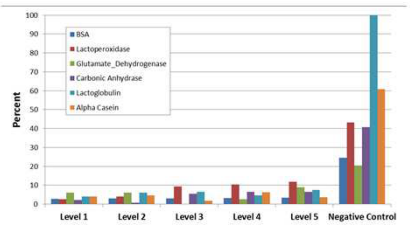


Figure 8. Protein level coefficients of variation across replicates. Consistently low CVs across the model HCP contamination levels indicate excellent quantitative accuracy down to the lowest tested level (protein present at 7-30 ppm). The very high CV values observed in the negative control confirm that the software could not find a reliable signal, as expected.

CONCLUSIONS

- **MS/MS^{ALL} with SWATH™ Acquisition** is a novel data-independent acquisition strategy that provides:
  - Comprehensive high resolution MS/MS data for all detectable ions
  - High quality quantitation similar to MRM with no method development
  - Easy and retrospective data interrogation
- SWATH™ data can be processed by PeakView™ Software and MarkerView™ Software or extracted for use with 3rd party informatics tools
- SWATH™ Acquisition is ideal for quantifying Protein Contaminants in Biologic protein products.
  - Quantitative sensitivity and fidelity rivaling ELISA without safety concerns of reagent preparation (not everything that produces a reaction in human produces a reaction in rabbit)
  - Captures a digital record of all fragments of all peptides in a protein product. This can be used to track changes over time and the data can serve as a digital archive of the current state of a sample at a given time. Can be retroactively mined for any protein contaminant concerns in the future.

REFERENCES

1. Nature Methods 6, 359-362 (2009).
2. J. Proteome Res. 2008, 7, (9), 3661-3667.
3. Electrophoresis, 20(18) 3551-67 (1999)

TRADEMARKS/LICENSING