Application Note

Biopharma/Pharma

Agilent

Trusted Answers

Adeno-Associated Virus Characterization with Agilent 6545XT AdvanceBio LC/Q-TOF and Protein Metrics Byos Software

Authors

Wendi A. Hale and
Christopher M. Colangelo
Agilent Technologies Inc.
Lexington, MA, USA

Michelle English
Protein Metrics
Cupertino, CA, USA

Abstract

This application note describes a workflow for the characterization of some critical quality attributes (CQAs) of adeno-associated viruses (AAVs), such as post-translational modification (PTM) identification, host cell protein (HCP) discovery, and amino acid misincorporation. The workflow was composed of an Agilent 1290 Infinity II LC coupled to an Agilent 6545XT AdvanceBio LC/Q-TOF, and was analyzed with Protein Metrics Byos software.

Introduction

AAVs are a crucial tool for the development and delivery of cutting-edge gene therapies and have been successful in treating inherited retinal diseases and spinal muscular atrophy. AAVs are composed of an icosahedral protein shell with a single-stranded genome of approximately 4.7 kb. There are three capsid proteins called VP1, VP2, and VP3. The intact AAVs act as vehicles to protect and deliver oligonucleotide therapeutics.

As AAVs grow in popularity in the form of therapeutic delivery platforms, it is essential that all CQAs of the therapeutic product are maintained. CQAs such as PTMs and truncations are routinely monitored by mass spectrometry. Other CQAs such as HCP analysis and sequence variant analysis (SVA), also known as amino acid misincorporation analysis, have traditionally been monitored with techniques such as ELISA and next-generation sequencing tools. However, identification and relative quantitation of these CQAs are shifting to being monitored by mass spectrometry as hardware and software become more powerful and improve in sensitivity and speed. In biopharmaceutical industry labs, speed is essential to improve product quality. This application note shows developed LC/MS methods for AAV characterization for HCP and SVA. In addition, orthogonal methods of subunit and peptide mapping were compared directly to capture discrepancies and to determine PTMs.

Experimental

Instrumentation

Agilent 1290 Infinity II LC system, including:

- Agilent 1290 Infinity II high-speed pump (G7120A)
- Agilent 1290 Infinity II multisampler (G7167B) fitted with 20 µL loop for intact analysis, 40 µL loop for peptide-mapping analysis
- Agilent 1290 Infinity II multicolumn thermostat (G7116B)
- Agilent 6545XT AdvanceBio LC/Q TOF

Materials

AAV8 was produced by Lake Pharma (now Curia; Worcester, MA, USA). Molecular weight cutoff filters and tris(2-carboxyethyl)phosphine (TCEP) were purchased from MilliporeSigma. Trypsin and rAsp-N were purchased from Promega Corporation.

Sample preparation

For intact analysis, AAVs underwent a buffer exchange three times at 10,000 g with a 10 kDa molecular weight filter. The buffer contained 5 mM TCEP, 80% H_2O, and 20% acetonitrile with 0.1% formic acid (v/v). After the sample was collected, it was incubated at room temperature before injection.

For PTM, HCP, and SV analyses, the AAVs underwent denaturation, reduction, alkylation with iodoacetamide, and digestion. The enzymes used in this experiment were trypsin and rAsp-N.

LC/MS analysis

LC/MS analysis was performed on a 1290 Infinity II LC coupled to a 6545XT AdvanceBio LC/Q-TOF with a dual Agilent Jet Stream source. Agilent MassHunter Acquisition (B.09.00) software with the large-molecule SWARM autotune feature was used for intact analysis. The instrument was further calibrated and operated in standard mass mode. The iterative MS/MS feature was used for the other workflows. Each dataset was processed with Byos software from Protein Metrics. The SVA validator algorithm was applied to the sequence variant workflow. LC parameters are displayed in Table 1 and LC/Q-TOF parameters are displayed in Table 2.

Table 1. LC parameters.

Parameter	Value
Column	Agilent AdvanceBio peptide mapping column, 2.1 × 150 mm
Column Temperature	60 °C
Mobile Phase A	Water, 0.1% formic acid
Mobile Phase B	Acetonitrile, 0.1% formic acid
Flow Rate	0.4 mL/min
Gradient	Time (min) %B 0 to 3 3 3 to 50 3 to 35 50 to 60 35 to 97 60 to 62 97 62 to 62.5 97 to 3 62.5 to 65 3
Post Time	5 min
Approximate Concentration Injected	1.8 e13 capsids/mL

Table 2. Agilent 6545XT AdvanceBio LC/Q-TOF parameters.

Parameter	Value
Source	Dual Agilent Jet Stream
Gas Temperature and Flow	325 °C and 13 L/min
Nebulizer	35 psig
Sheath Gas Temperature and Flow	275 °C and 12 L/min
VCap	4,000 V
Nozzle	0 V
Fragmentor	170 V
Acquisition Rate	5/3 spec/s for MS and MS/MS
Reference Masses	121.0509, 922.0098

Results and discussion

Adeno-associated virus (AAV) analysis considerations

For LC/MS analysis, AAV assays have some challenges that are not present in other therapeutics such as monoclonal antibodies. Protein yield is on the order of magnitudes lower, so sample amount is greatly limited. In addition, some portions of the amino acid sequence are riddled with ragged ends or frequent lysine or arginine residues, whereas other portions of the sequence are devoid of lysine and arginine. Therefore, the use of a second enzyme, in addition to trypsin, is highly recommended. Instead of developing separate methods for PTM, HCP, and SV analyses, it makes the most sense to develop one LC/MS method that can encompass the need for all three types of analyses and combine it with iterative MS/MS to ensure that low-level peptides are characterized.

Host cell protein (HCP) analysis

HCP analysis is commonly performed with ELISA. ELISA, while easy to use, is not specific and cannot identify individual proteins. HCP identification by LC/MS/MS is becoming more common with traditional biotherapeutic analysis, but it is not frequently used as an HCP assay with AAVs. Detecting and characterizing low-abundance HCPs is an ideal situation for using the iterative MS/MS feature in MassHunter Acquisition software. In the iterative MS/MS workflow, the first injection acts as a normal auto MS/MS function, where the top N precursor ions are selected and isolated for MS/MS fragmentation. Subsequent iterative runs automatically add any precursors that were selected for MS/MS to a rolling exclusion list, so that lower abundant precursors are characterized with MS/MS. In addition, iterative data results can be combined in Byos software.

Due to limited sample amount, two iterative runs were acquired with the tryptic digest and two iterative runs were acquired with the Asp-N digest. A contaminant database in Byos software was added to the FASTA file to monitor relative abundance of proteins such as trypsin and rAsp-N compared to the host cell proteins. The VP1 protein sequence was added to the FASTA file as well. All three capsid proteins have overlapping amino acid sequences and VP1 is the largest protein that encompasses the sequences of VP2 and VP3. As expected, most proteins were identified as the AAV8 VP1 capsid protein (Figure 1, pink). Relative quantitation was achieved with the default setting of adding the area of the three most abundant extracted ion chromatograms (EICs). Of the top five most abundant contaminant proteins, three were from the contaminant database: trypsin (Figure 1, dark blue), Asp-N (Figure 1, orange), and keratin (Figure 1, red). A protein annotated as a heat shock protein was fairly abundant (Figure 2, blue), but the remaining HCPs were all low in abundance. A more quantitative approach would need to be performed to determine the exact amount of HCP.

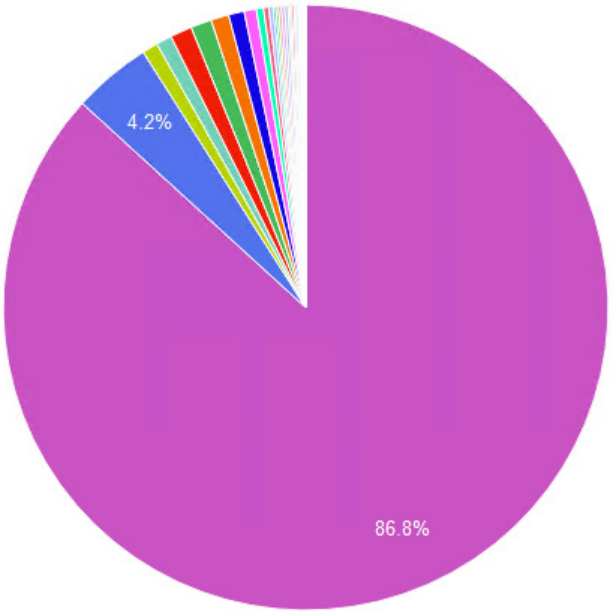

Figure 1. The top three EICs of each protein were added together for relative quantitation. As expected, AAV8 (dark pink) is overwhelmingly abundant compared to protein impurities.

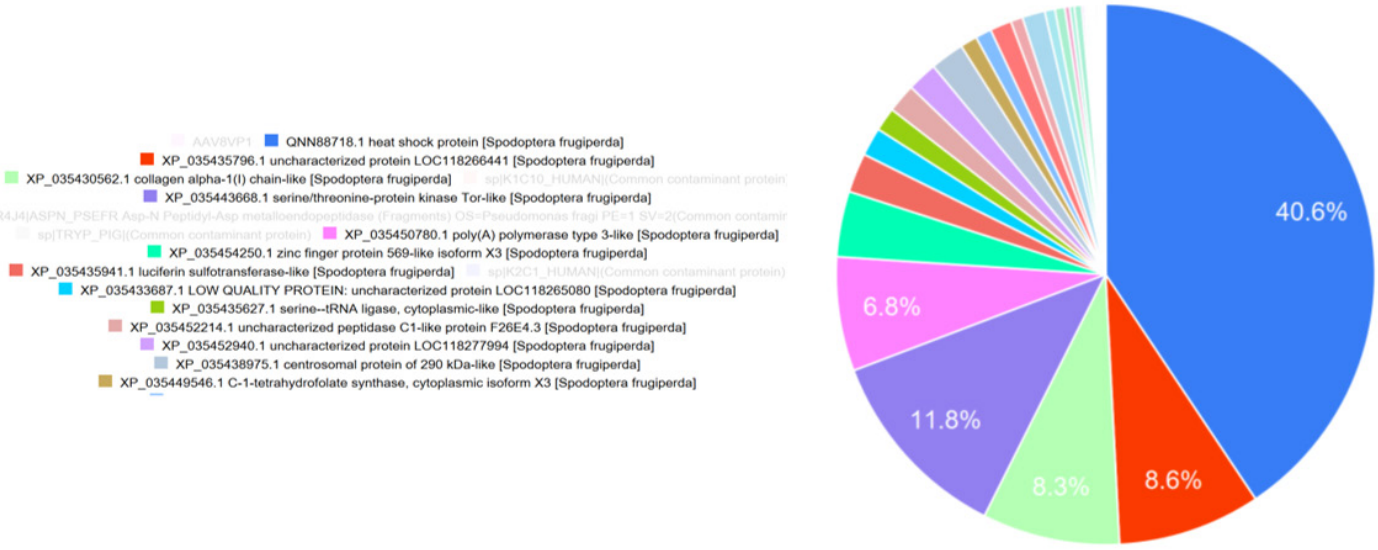

Figure 2. Top three EICs of all HCPs (AAV8 VP1 and contaminants removed). Only one protein is more abundant than artefact proteins. The remaining proteins include a few uncharacterized proteins and several enzymes.

Sequence variant analysis (SVA)

Like HCP analysis, SVA by LC/MS/MS is becoming routine for traditional biotherapeutics, but less common for AAV analysis. Next-generation sequencing is the gold standard, although it can take weeks to obtain results. Sequence variants can occur at any step of the protein expression process, from mutations during DNA replication to misincorporations during protein translation events. While the exact allowed percentage is still being debated for antibodies, approximately 1% is considered acceptable, although limits as strict as 0.1% are being proposed.

While an LC/MS/MS assay for SVA does not take as long as multiple weeks to collect and analyze the data, the data analysis is still extensive. Multiple features in Byos make this process easier. The sequence variant workflow automatically includes the full set of single amino acid substitutions as modifications. In addition, the SVA validator will label sequence variants with "Critical", "Warning" or "Clear" based on criteria such as lack of wild type peptide, peptide score, percent EIC ratio, and differences in retention time. While it is still recommended to inspect the MS/MS data to verify sequence variants, the validator quickly eliminates false positives and narrows down the list for true variants.

For this set of experiments, the default parameters for the validator were used, including the percent EIC ratio threshold set to 1%. Ratios up to 10% were manually inspected for multiple MS/MS fragments that included the variant.

Four of six sequence variants had an EIC ratio under 1% (Table 3). Each of the variants' MS/MS spectra were manually validated, noting multiple b/y ions that confirmed the variant (Figure 3). The two variants above 1% were serine to asparagine, which is a common misincorporation due to a G/U mismatch during translation.

Table 3. List of verified sequence variants and their relative abundance.

Variant	EIC Ratio%	Variant Amino Acid Position in VP1
Gly → Asp	0.41	545
Ile → Met	0.03	648
Met → Xle	0.57	404
Met → Xle	0.48	560
Ser → Asn	4.47	557
Ser → Asn	2.60	670

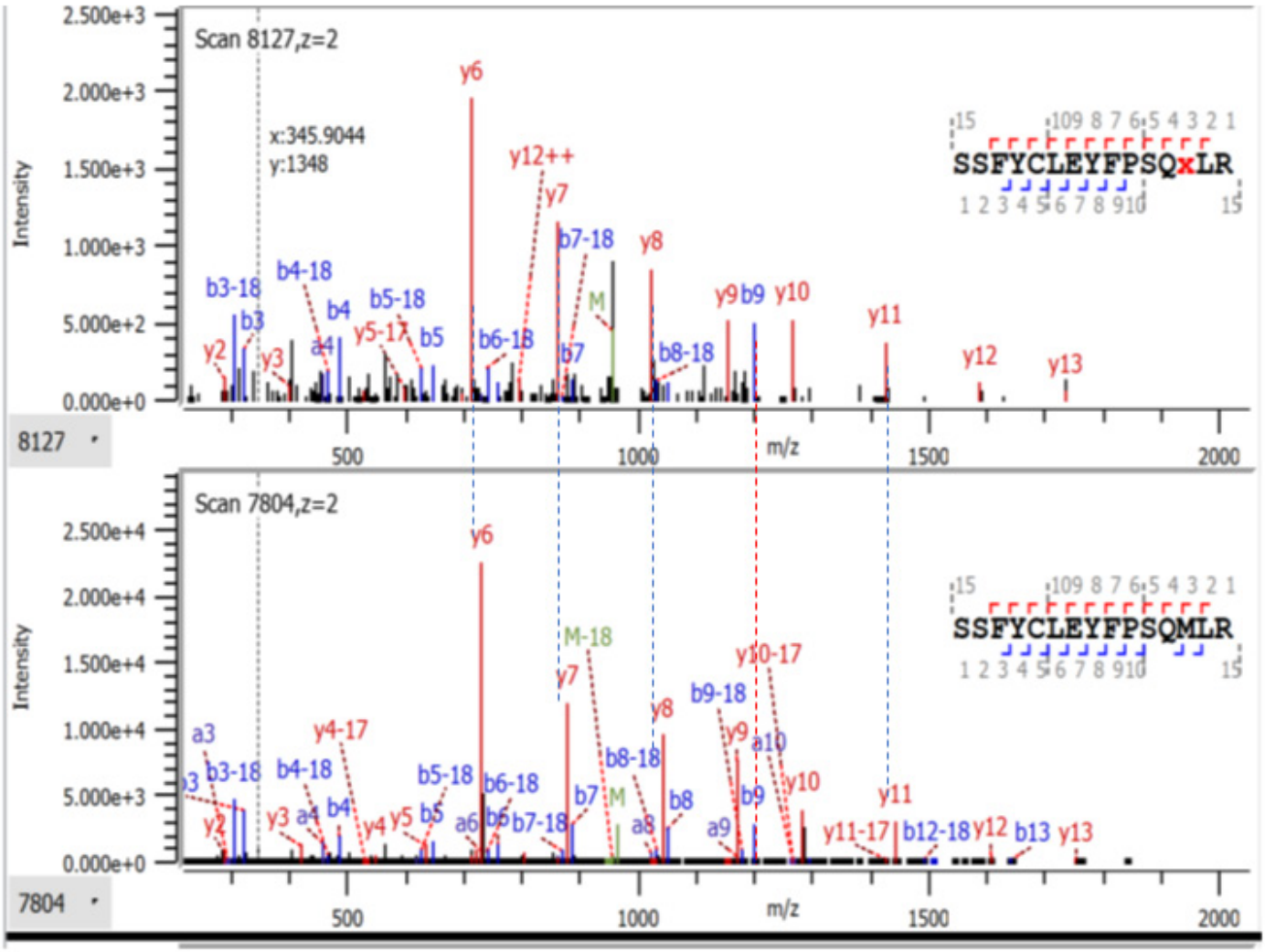

Figure 3. MS/MS confirmation of M to I/L variant. Dotted blue lines were added to clarify differences in y-ions, while the red dotted line for b9 ions confirms they are at the same *m/z* value.

Post-translational modification (PTM) analysis and reconstruction

Reconciling and comparing orthogonal methods such as intact analysis and peptide mapping has its challenges. There are major differences in experimental procedures, sample preparation, and chromatography and mass spectrometry source conditions, as well as ionization differences between the two types of methods. To mitigate this challenge, Byos software has a feature that takes the peptide-mapping data, including PTM information, and reconstructs it into a theoretical, intact mass spectrum. The theoretical, intact spectrum is created by aggregating the contributions from all the peptide ions. The theoretical spectrum is compared with an acquired mass spectrum and the two are overlaid. A second feature is the reconstruction table, which provides the data for the peptide ions that make up the intact spectrum (Figure 4). The reconstruction table can be adjusted and edited to search for any discrepancies in the two data sets. The reconstruction feature and table make it easier for those developing biotherapeutics to determine challenges and discrepancies earlier in the development process.

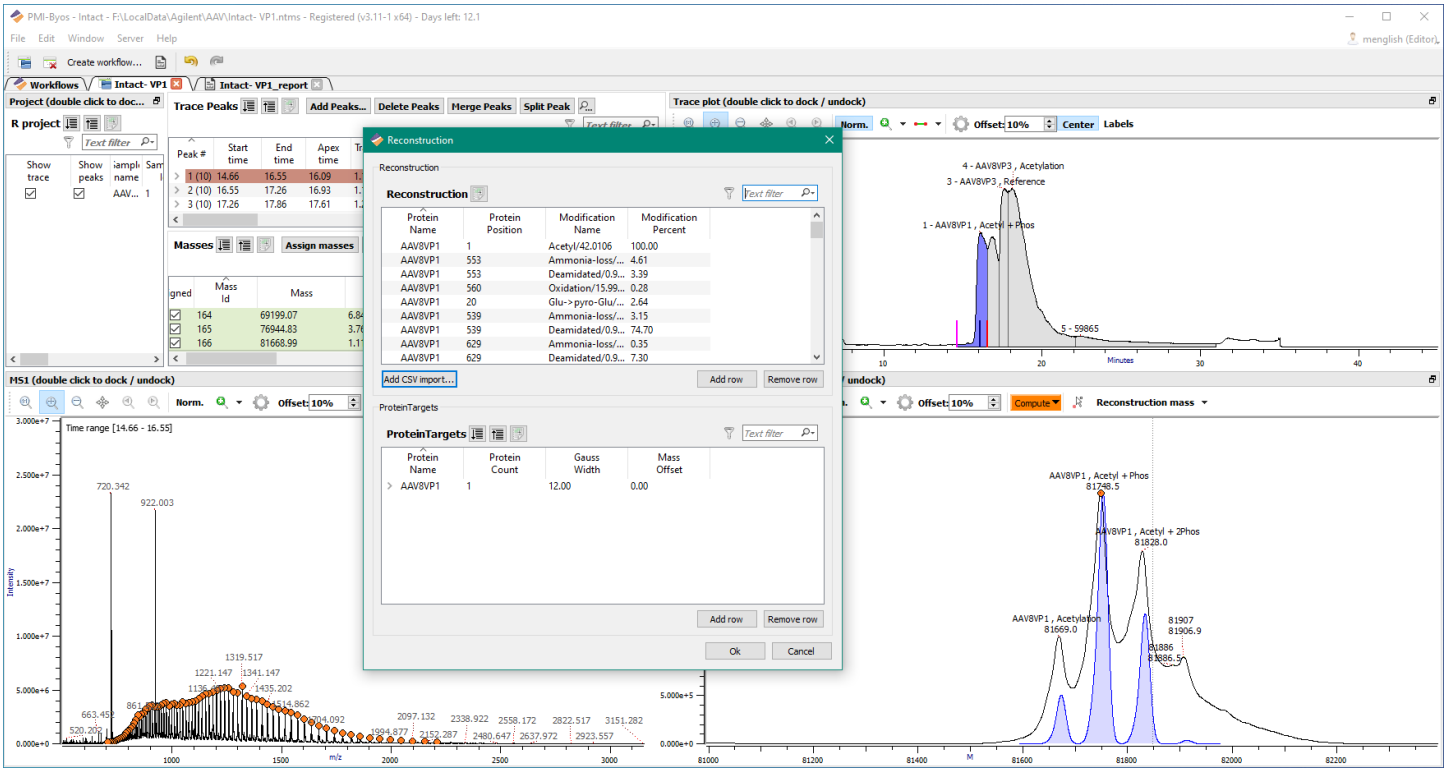

Figure 4. The reconstruction table for VP1 in Protein Metrics Byos software. Modifications can be curated in the table and are exported from the peptide mapping modifications report.

In Byos software, the intact AAV data had shoulder peaks that were to the left of the main peak. When the data were originally analyzed in Agilent MassHunter BioConfirm software, these were resolved peaks, but the mass difference from the main peak did not add up to a known PTM or mass loss. The reconstruction was able to confirm that the intact shoulder peak was an amalgam of multiple PTMs and provided greater clarity to the composition of the species that were present (Figure 5). Generally, the intact and peptide-level spectra were comparable, although there were some species in the peptide map that were not accounted for and they were evident at the intact level. One example of this is the third phosphorylation in VP1 and VP2, as well as the unphosphorylated VP2 (Figures 5 and 6). In the VP3 reconstruction data, the unacetylated peak is overestimated in the peptide-mapping data (Figure 7). This is likely due to a clip that was present in the intact data, and the clip data are included in the reconstruction.

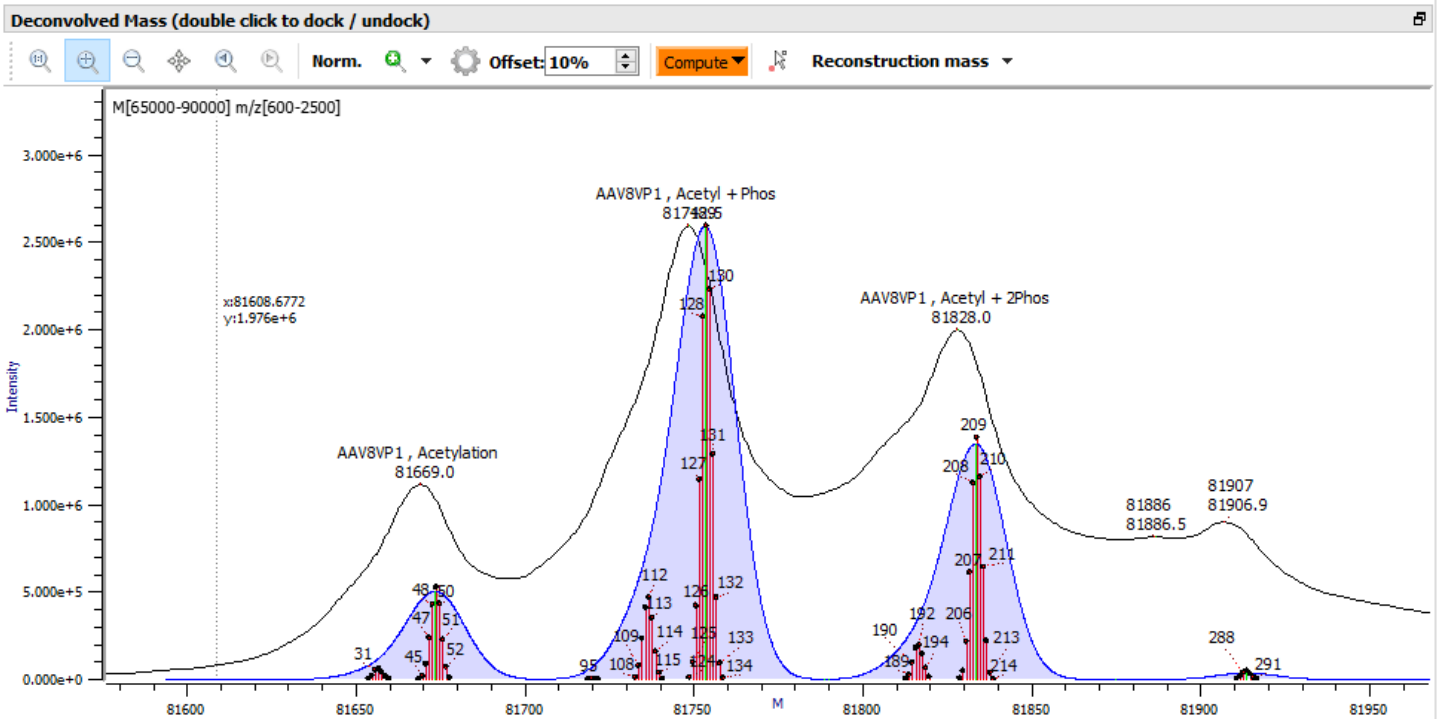

Figure 5. Reconstruction data of VP1. Red lines represent relative quantitation of modifications on specific amino acids. Red lines are used to draw the theoretical blue intact spectrum. Shoulder peaks in the experimental spectrum (outlined in black) are a result of multiple low-abundant and -mass PTMs, which is revealed in the reconstruction data.

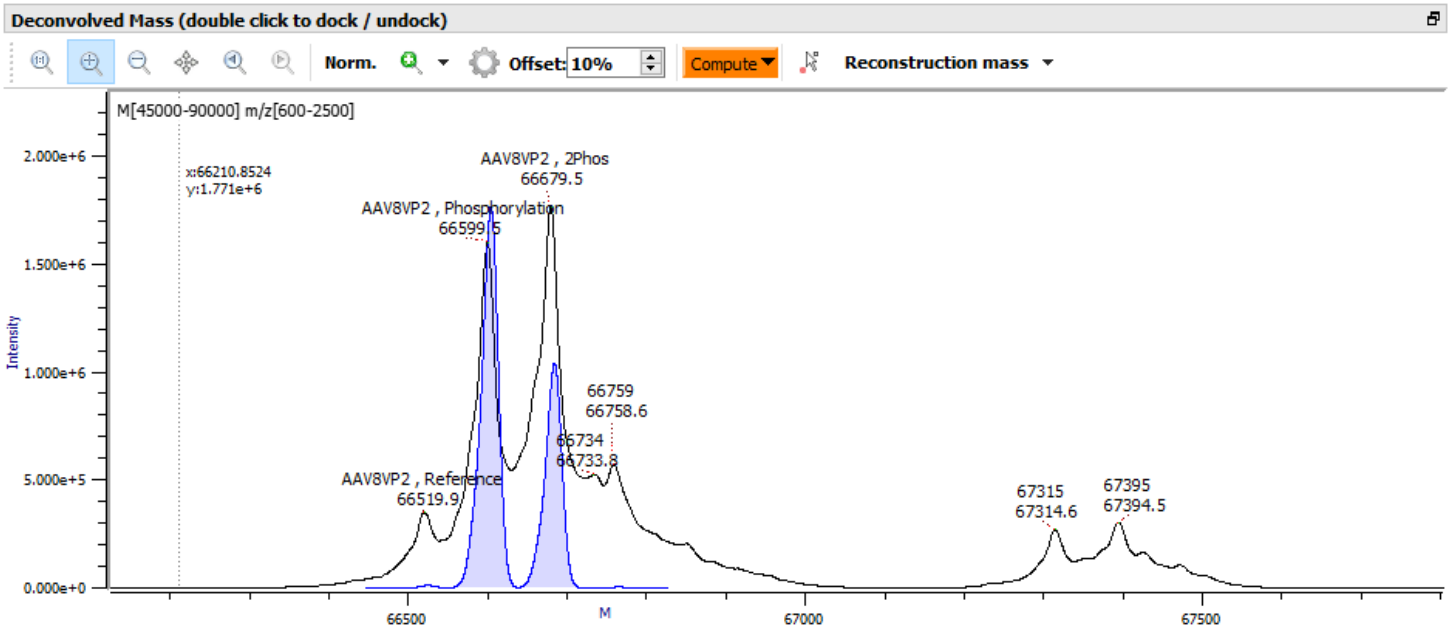

Figure 6. Reconstruction data of VP2. The reconstruction data underestimate the second and third phosphorylation in VP1 (Figure 2) and VP2.

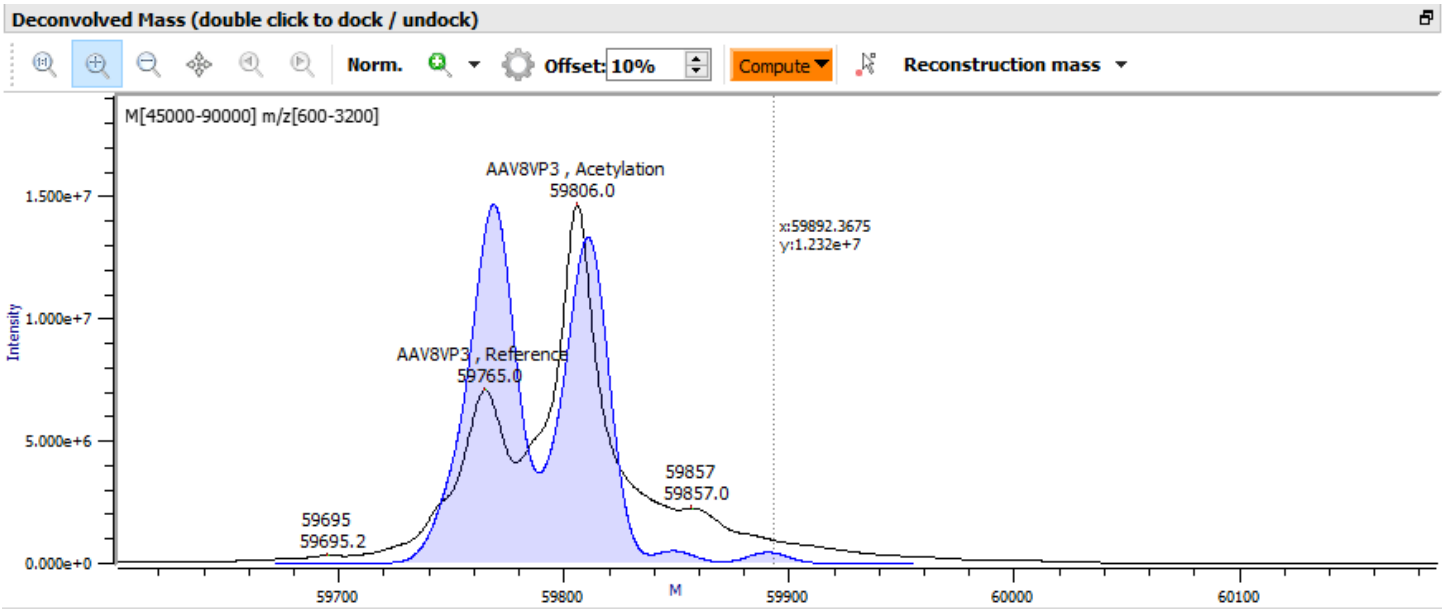

Figure 7. Reconstruction data of VP3. The nonacetylated data is overstated in the peptide mapping data. This may be due to a clip, as the clip mapping data were included in the reconstruction.

Conclusion

This study demonstrates a comprehensive workflow for AAV capsid proteins with an Agilent 1290 Infinity II LC coupled to an Agilent 6545XT AdvanceBio LC/Q-TOF. Host cell proteins and sequence variants were discovered and relatively quantified with the same acquisition LC/MS/MS runs. In addition, post-translational modifications are confirmed at the peptide-mapping level and compared to the intact analysis data with the reconstruction tool using Protein Metrics Byos software. This robust and easy-to-implement workflow is ideal for sample-limited applications like AAV capsid analysis.

References

1. Hale, W. A.; Colangelo, C. M. Characterization of Viral Vector Particles Using the Agilent 6545XT AdvanceBio LC/Q-TOF. *Agilent Technologies application note*, publication number 5994-1980EN, **2020**.
2. Liau, B. LC/MS of Intact Adeno-Associated Virus Capsid Proteins for Rapid Confirmation of Product Identity. *Agilent Technologies application note*, publication number 5994-2434EN, **2020**.
3. Huang, Y. *et al.* Toward Unbiased Identification and Comparative Quantification of Host Cell Protein Impurities by Automated Iterative LC–MS/MS(HCP-AIMS) for Therapeutic Protein Development. *J. Pharm. Biomed.* Anal. **2021**, *200*, 114069–114079.
4. Valliere-Douglass, J. *et al.* Biopharmaceutical Industry Practices for Sequence Variant Analyses of Recombinant Protein Therapeutics. *PDA J. Pharm. Sci. and Tech.* **2019**, *73*, 622–634.
5. Zhang, A. *et al.* A General Evidence-Based Sequence Variant Control Limit for Recombinant Therapeutic Protein Development. *mAbs*. **2020**, *12(1)*, 1791399–1791409.
6. Zhang, Z. *et al.* G/U and Certain Wobble Position Mismatches as Possible Main Causes of Amino Acid Misincorporations. *Biochem*. **2013**, *52*, 8165–8176.

www.agilent.com

DE15843892

This information is subject to change without notice.

Printed in the USA, July 14, 2022
5994-5110EN

Agilent
Trusted Answers