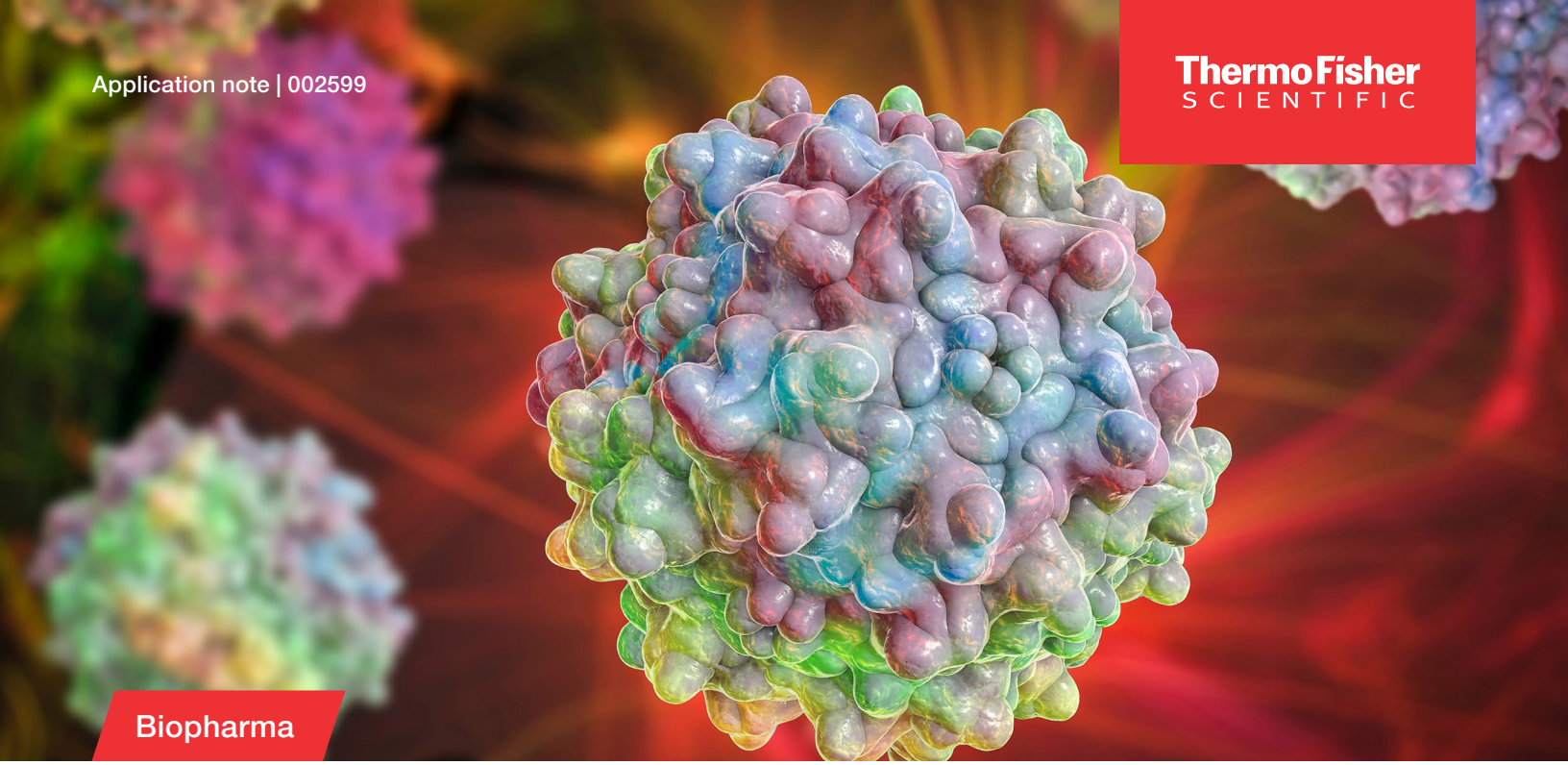

Characterization of adeno-associated viral proteins proteoforms using a top-down MS approach on a UHPLC-Orbitrap Ascend Tribrid mass spectrometer

Authors

Reiko Kiyonami[1], Rafael Melani[1], Roberto Gamez[1], Kenneth Thompson[2], Chao Yan Liu[2], Min Du[3]

[1]Thermo Fisher Scientific, San Jose, CA, USA

[2]Thermo Fisher Scientific, Frederick, MD, USA

[3]Thermo Fisher Scientific, Cambridge, MA, USA

Keywords

AAV viral proteins, posttranslational modification (PTM), top-down MS, Orbitrap Ascend Tribrid mass spectrometer, multiple dissociation techniques, EThcD, HCD, UVPD, BioPharma Finder 5.1 software

Application benefits

- Confident determination of PTM localization on adeno-associated virus (AAV) viral proteins and sequence confirmation of truncated AAV viral proteoforms using a top-down MS approach
- Improve sequence coverage with a top-down approach using multiple complementary fragmentation techniques

Goal

Demonstrate the benefit of using multiple dissociation techniques on a Thermo Scientific™ Orbitrap™ Ascend Tribrid™ mass spectrometer to improve LC-MS based top-down characterization of AAV viral proteins

Introduction

Recombinant AAV vectors have emerged as the leading gene delivery vehicles for gene therapy due to their high-efficiency transduction and safety.[1-3] A wild-type AAV particle contains a single-stranded DNA genome and a capsid comprised of three structural viral proteins (VP1, VP2, and VP3) that assemble the viral particle at an approximate ratio of 1:1:10.[1] Each viral protein (VP) has an overlapping sequence at the C-terminus and

differs only at its N-terminal region.[4,5] The surface composition of the AAV capsid is an essential component involved in cell binding, internalization, and trafficking within the targeted cell.[5]

Complete characterization of viral proteins is required to ensure the safety, quality, and efficacy of AAV products.[6] Recent advances in LC-MS are being adopted as fast analytical methods for AAV viral protein characterization.[7,8] Unlike therapeutic proteins, the AAV is more complex in structure and limited in sample volume with low titer due to small batch sizes, yielding a very low amount of capsid proteins for subsequent LC-MS analysis. The desired LC-MS platform for the characterization of the AAV VPs needs to offer high sensitivity for detecting low abundant viral proteins and provide accurate intact mass measurements.[9] Furthermore, each VP has a highly overlapping sequence at the C-terminus. When digesting VPs to the peptide mixtures, the identified peptides with PTM could belong to either VP1, VP2 or VP3, making assigning the PTM sites at the intact VP level challenging.[10] Top-down MS analysis can help to localize the PTM site at the intact VP level and to confirm the amino acid sequence of truncated VP proteoforms by generating each VP proteoform-specific fragment ion information. The challenge for top-down MS analysis is to get comprehensive sequence coverage of larger proteoforms.

The new generation Orbitrap Ascend Tribrid mass spectrometer offers increased sensitivity, faster MS/MS scan rate and multiple dissociation techniques for improved fragmentation sequence coverage while performing top-down MS experiments.[11,12] As proof of concept, we performed targeted top-down MS experiments using EThcD, HCD, and UVPD on an HPLC-Orbitrap Ascend Tribrid mass spectrometer. We analyzed the AAV6 serotype by top-down MS using different ion activation methods to confirm Ser-11 (S11) and Thr-56 (T56) phosphorylation in VP2 and the sequence of VP3 clip described in previous studies.[9,13] The additional fragment ion information offered by the EThcD and UVPD fragmentation methods clearly detected the *c* and *b* type of fragment ion pairs with and without phosphorylation sites. It also significantly improved the sequence coverages for both VP2 (with and without phosphorylations) and VP3 clip.

Experimental

Sample preparation

An AAV6 sample expressed in HEK293 cells was buffer exchanged and concentrated into 80% H_2O/20% acetonitrile containing 5 mM TCEP and 0.1% formic acid using a 30K centrifugal filter (Amicon™ Ultra, 0.5 mL). The estimated protein concentration was 0.2 µg/µL after buffer exchange.

Chromatography

The Thermo Scientific™ Vanquish™ Horizon UHPLC system consisting of the following modules was used for the protein separation. A Thermo Scientific™ nanoViper™ fingertight fitting (50 µm × 350 mm, P/N 6041.5540) was used for connecting the outlet of the column to the mass spectrometer.

- Thermo Scientific™ Vanquish™ System Base Horizon/Flex (P/N VF-S01-A-02)
- Thermo Scientific™ Vanquish™ Binary Pump H (P/N VH-P10-A)
- Thermo Scientific™ Vanquish™ Split Sampler FT (P/N VH-A10-A)
- Thermo Scientific™ Vanquish™ Column Compartment H (P/N VH-C10-A-03)

For all experiments, chromatographic separations were carried out using a C4 stationary phase column (1.0 × 100 mm, 300 Å, 1.7 µm). The solvent A was water with 0.1% difluoroacetic acid, and the solvent B was acetonitrile with 0.1% difluoroacetic acid. The column temperature was set to 80 °C. The flow rate was 50 µL/min. The gradient condition used is listed in Table 1.

Table 1. HPLC gradient condition

Time (min)	%B
0	20
1	32
16	36
20	80
21.5	80
22	20
30	20

Mass spectrometry

The Orbitrap Ascend Tribrid mass spectrometer equipped with the HMRn+, UVPD, and ETD options was used for MS data collection. The instrument was operated with Thermo Scientific™ Xcalibur™ 4.6 software and controlled by Orbitrap Tribrid Series Instrument Control Software 4.0 SP1. The ESI source settings are shown in Table 2. The MS settings for the intact mass analysis are shown in Table 3. The MS settings for top-down MS analysis are shown in Table 4.

Data processing

Thermo Scientific™ BioPharma Finder™ 5.1 software was used for intact protein and top-down MS data analysis.

Table 2. ESI settings

Parameter	Value
Sheath gas	25
Aux gas	5
Sweep gas	0
Spray voltage	3,500 V
Capillary temp.	320 °C
Vaporizer temp.	125 °C

Table 3. MS settings for intact analysis

Parameter	Value
Scan range	800–5,000 *m/z*
Application mode	Intact
Pressure mode	Low
Resolution	15,000 at *m/z* 200
RF lens	60%
AGC target value	150%
Max inject time	50 ms
Microscans	10
Source fragmentation	15 V

Table 4. MS settings for top-down analysis

Parameter	Value		
Full MS			
Mass range	High		
Scan range	800–5,000 *m/z*		
Application mode	Intact		
Pressure mode	Low		
Resolution	7,500 (at *m/z* 200)		
RF lens	60%		
AGC target value	150%		
Microscans	5		
Source frgamentation	15 V		
tMS²			
Targeted precursor ion	1,441.50		
Isolation window	6		
Fragment type	**EThcD**	**HCD**	**UVPD**
Fragment parameters	Reaction time: 8 ms Max. reagent inj. time: 200 ms Reagent target: 7E5 SA collision energy: 15%	Collision energy: 25%	Activation time: 25
Mass range	Normal	Normal	Normal
Scan range	400–1,400 *m/z*	400–2,000 *m/z*	400–2,000 *m/z*
Application mode	Intact		
Pressure mode	Low		
Resolution	120,000 (at *m/z* 200)		
RF lens	60%		
AGC target value	6,000%		
Max inject time	2,000 ms		
Microscans	3		

Results and discussion

Precursor ion selection for the targeted top-down MS experiments

Intact mass analysis was carried out first for selecting the representative precursor ion which can provide specific fragment ions related with VP2 phospho-proteoforms and VP3 clip for the top-down experiments. Figure 1 shows the TIC of the detected viral proteins and MS spectra of VP2 and VP3 clip. The enlarged *m/z* mass range for the 46+ charge state of VP2 clearly shows two phosphorylated proteoforms (Figure 1, left insert). To generate the fragment ions directly linked to the two phospho-proteoforms, *m/z* 1,441.5 (the second phospho-proteoform) was selected as tzhe precursor ion for all targeted top-down experiments. By using the isolation window of 6 amu, the *m/z* 1,439.65 (representing the first phospho-proteoform) was also co-isolated. The selected precursor ion with an isolation window of 6 amu also covers the 30+ charge state (*m/z* 1,441.68) of VP3 clip.

The acquired top-down MS data during the VP2 elution time were used for the phosphosite localization of VP2. The acquired top-down MS data during the VP3 clip peak elution time were used for the protein sequence verification of VP3 clip.

Confident localization of two phosphorylation sites from VP2 using multiple dissociation methods

Targeted top-down LC-MS experiments were performed using EThcD, HCD, and UVPD. A full MS scan experiment was also performed in the same LC-MS run to ensure only the MS^2 data that are relevant to the VP2 or VP3 clip were used for data processing. To get more precursor ions into the C-trap, a high AGC value and up to 2,000 ms ion injection time were used in all experiments. Figure 2 shows the fragment ions of EThcD collected with a mass range of *m/z* 400–1,400. The c-type fragment ion pairs with and without phosphorylation in the specific amino acid locations (11S and 56T) were identified clearly (Figure 2, insert).

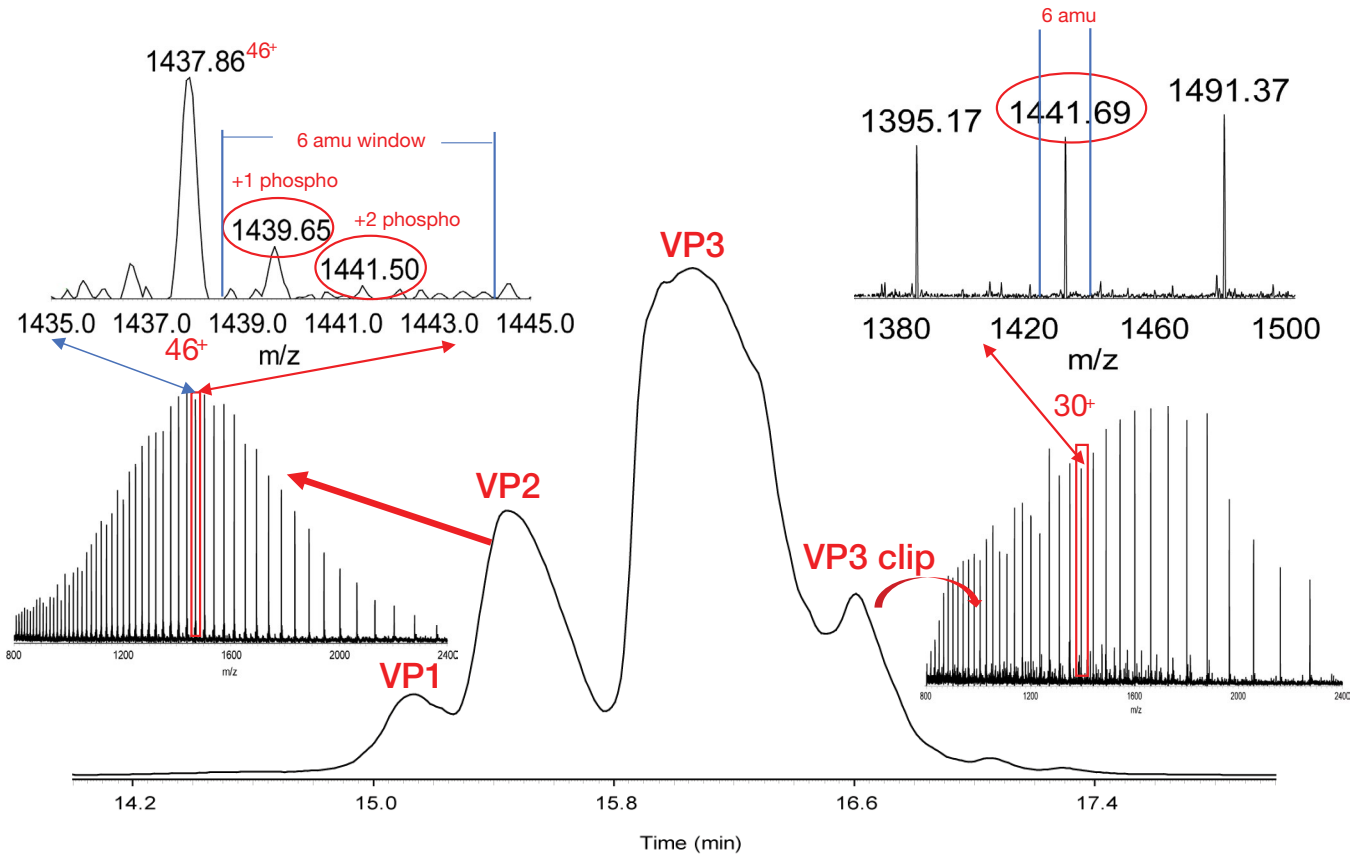

Figure 1. Targeted precursor ion selection based on AAV viral protein intact mass analysis data

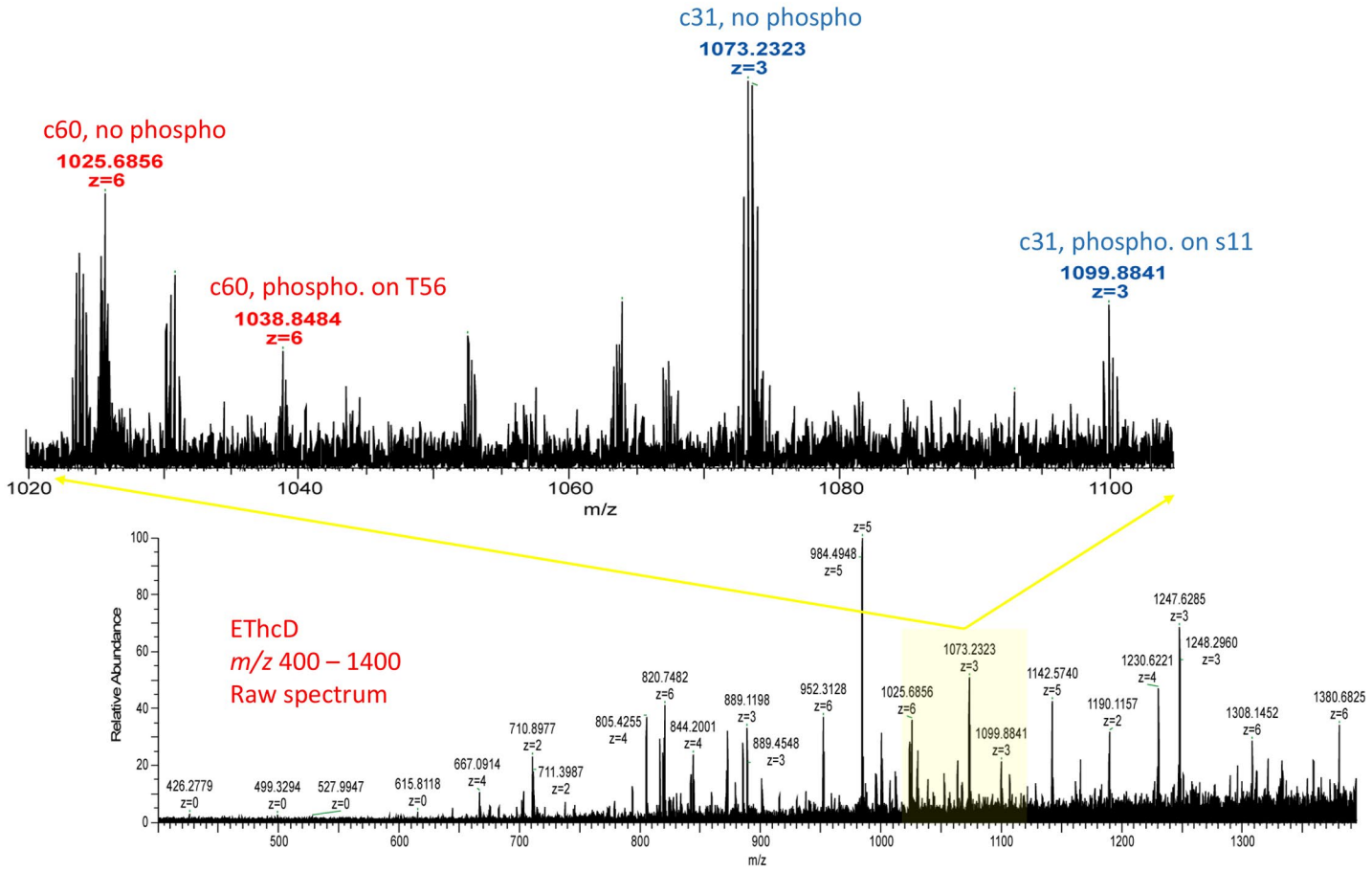

Figure 2. Fragment ions observed from EThcD with mass range of *m/z* 400–1,400

A total of three top-down raw files (EThcD, HCD, and UVPD) were processed using the top-down workflow in Biopharma Finder 5.1 software. The fragment ion data across the eluted VP2 peak were averaged per raw file and the averaged spectra were searched against four VP2 proteoforms (VP2 proteoform without phosphorylation, VP2 phospho-proteoform with phosphorylation on the S11, VP2 phospho-proteoform with phosphorylation on the T56, and VP2 phospho-proteoform with double phosphorylations on the S11 and T56) using 15 ppm mass tolerance for the sequence coverage and phosphorylation localization confirmation. The presence of S11 and T56 phosphorylations in VP2 was confirmed by fragment ions detected in the spectra obtained from all three dissociation methods. Figure 3, Figure 4, and Figure 5 show the sequence coverages for the VP2 phospho-proteoforms with phosphorylation on the S11 and phosphorylation on the T56 and the combination of S11 and T56 for the double phosphorylated proteoform from the EThcD, HCD, and UVPD, respectively. The phosphorylated serine and threonine residues are highlighted in green. HCD and UVPD provided complementary fragment ions for supporting the EThcD search results with increased sequence coverage. The observed residue cleavages using each fragmentation method and the combined residue cleavages are summarized in Table 5 and Figure 6.

Increased sequence coverage for the truncated VP3 proteoform using multiple dissociation methods

It is important to characterize not only major VP proteoforms, but also all associated processed proteoforms to ensure the safety and quality of AAV products. In our previous AAV6 study, we proposed and confirmed the sequence of the detected AAV3 clip peak using the top-down approach with the HCD dissociation method and observed 13% sequence coverage.[9] In this study, in addition to HCD, the EThcD and UVPD top-down data were also used for confirming the proposed VP3 clip sequence. The fragment ion spectra collected across the eluted VP3 clip peak were averaged per top-down experiment raw file and searched against the proposed VP3 clip sequence using 15 ppm mass tolerance. The multiple dissociation methods provided significant numbers of unique fragment ions (Figure 7). Significantly increased sequence coverage of 65% was observed using the combined fragmentation data (Figure 8).

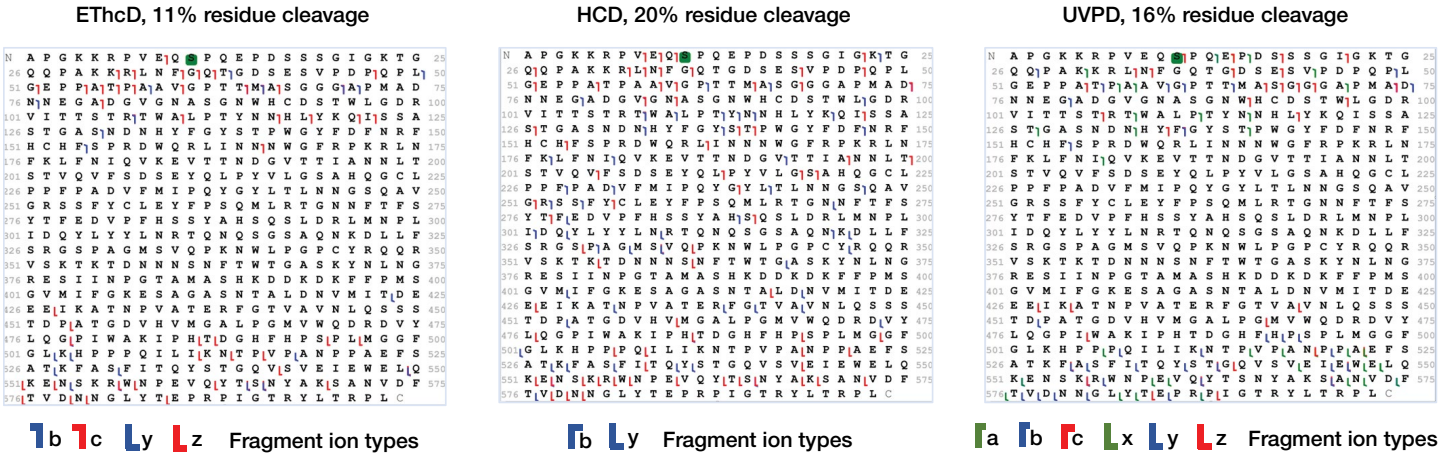

Figure 3. The sequence coverages for the VP2 with the phosphorylation on S11 observed with EThcD (*m/z* 400–1,400), HCD, and UVPD

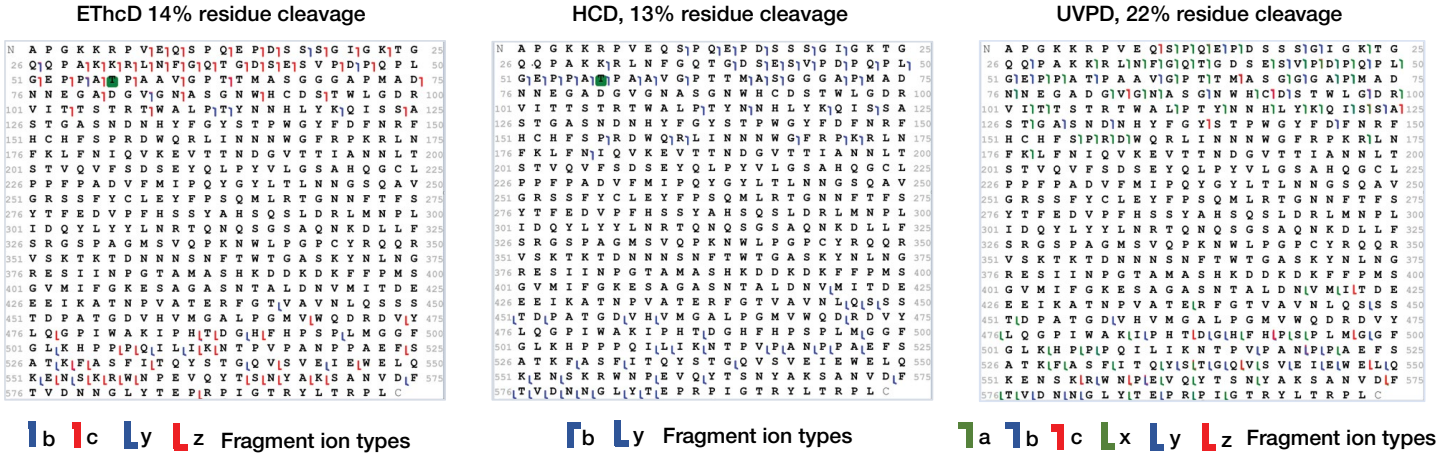

Figure 4. The sequence coverages for the VP2 with the phosphorylation on T56 observed with EThcD (*m/z* 400–1,400), HCD, and UVPD

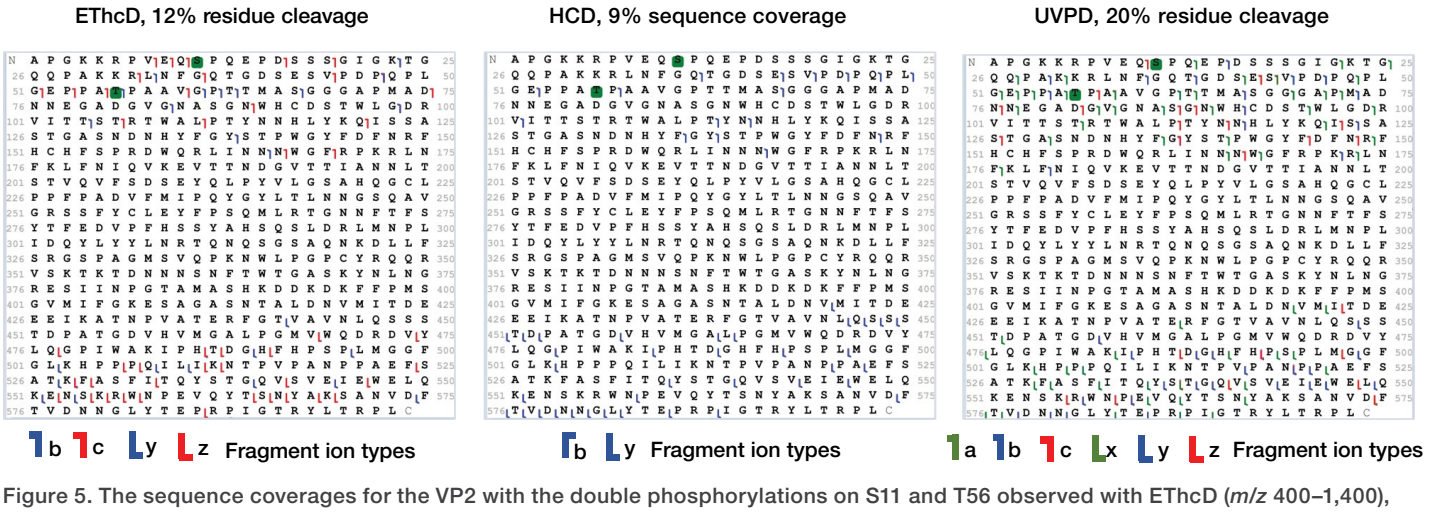

Figure 5. The sequence coverages for the VP2 with the double phosphorylations on S11 and T56 observed with EThcD (*m/z* 400–1,400), HCD, and UVPD

Table 5. The residue cleavages for the VP2 and its phospho-proteoforms observed with each individual EThcD (*m/z* 400–1,400), HCD UVPD, combined three EThcD and combined three EThcD, HCD, and UVPD

	Residue cleavages (%)			
	Non-phosphorylated VP2	VP2 phospho-proteoform, (S11)	VP2 phospho-proteoform, (T56)	VP2 combined phospho-proteoform, (S11,T56)
EThcD (*m/z* 400–1,400)	15	11	14	12
HCD (*m/z* 400–2,000)	14	11	13	9
UVPD (*m/z* 400–2,000)	17	16	17	20
Combining three EThcD, HCD, and UVPD	50	45	47	35

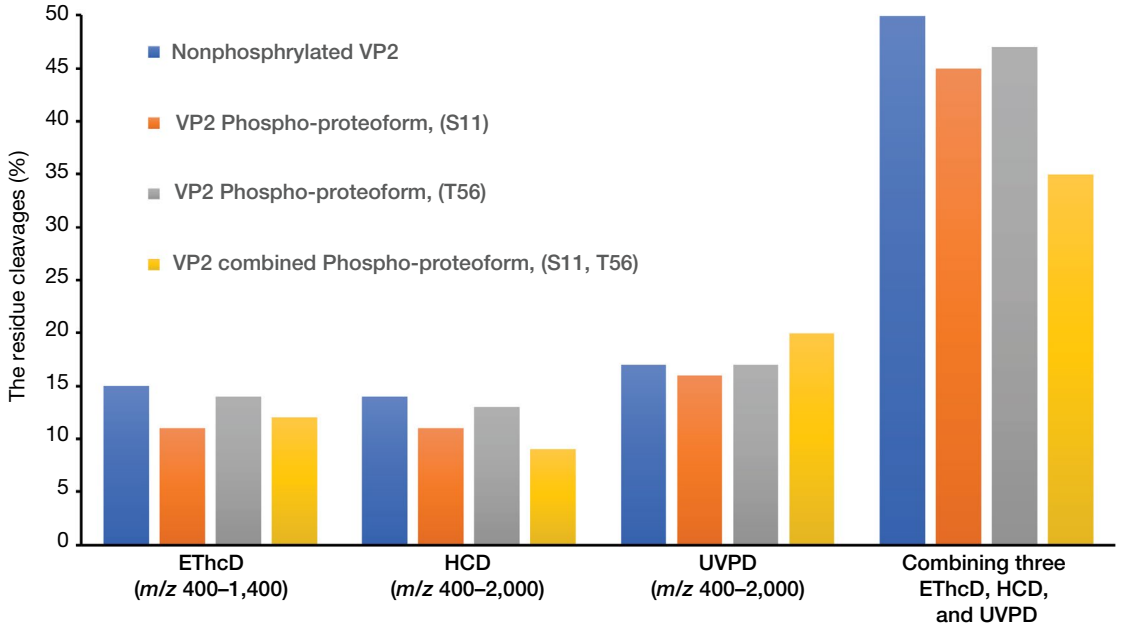

Figure 6. The residue cleavages for the VP2 and its phospho-proteoforms observed with each individual EThcD (*m/z* 400–1,400), HCD UVPD, combined three EThcD, and combined three EThcD, HCD, and UVPD. Significantly increased sequence coverages are observed by combining the complementary fragment ions from multiple dissociation methods.

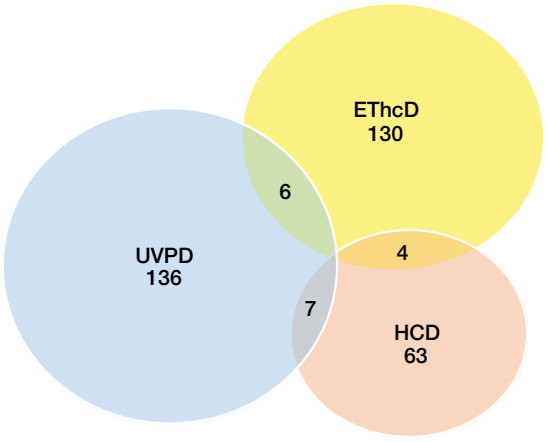

Figure 7. Venn diagram of fragment ions that matched the VP3 clip sequence

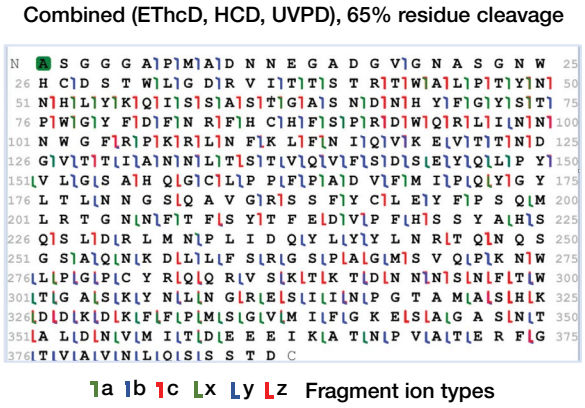

Figure 8. The combined sequence coverages for the VP3 clip by combining the complementary fragment ions from three EThcD, HCD, and UVPD

Conclusions

- A targeted top-down MS method that uses optimized mass spectrometer settings and multiple dissociation techniques was developed to address the challenges of characterizing low-level site-specific proteoforms.
- The developed method was applied to direct fragmentation of low abundant VP2 phospho-proteofroms and VP3 proteoform with C-terminal truncation (VP3 clip).
- Two phosphorylation sites on VP2 were verified unambiguously.
- The proposed VP3 clip sequence was confidently confirmed.
- The multiple dissociation techniques provided complementary fragment ions and resulting in more than doubled sequence coverage.

References

1. Samulski, R.J.; Muzyczka, N. AAV-Mediated Gene Therapy for Research and Therapeutic Purposes, *Annu. Rev. Virol.* **2014,** *1*(1), 427–451.
2. Naso, M.F.; Tomkowicz, B.; Perry, W.L.; Stroh, W.R. Adeno-Associated Virus (AAV) as a Vector for Gene Therapy, *BioDrugs* **2017,** *31,* 317–334.
3. Dong, B.; Duan, X.; Chow, H.Y.; Chen, L.; Lu, H.; Wu, W.; Hauck, B. Proteomics Analysis of Co-Purifying Cellular Proteins Associated with rAAV Vectors, *PLOS ONE,* **2014,** *9*(2).
4. Popa-Wagner, R.; Porwal, M.; Kann, M.; Reuss, M.; Weimer, M.; Florin, L.; Kleinschmidt; J.A. Impact of VP1-Specific Protein Sequence Motifs on Adeno-Associated Virus Type 2 Intracellular Trafficking and Nuclear Entry. *Journal of Virology,* **2012**, *86*(17), 9163–9174.
5. Van Vliet, K.M.; Blouin, V.; Brument, N.; Agbandje-McKenna, M.; Snyder, R.O. The Role of the Adeno-Associated Virus Capsid in Gene Transfer. Methods *Mol Biol.* **2008,** *437,* 51–91.
6. U.S. FDA, Chemistry, Manufacturing, and Control Information for Human Gene Therapy Investigational New Drug Applications; Guidance for Industry.
7. Jin, X.; Liu, L.; Nass, S.; O'Riordan, C.; Pastor, E.; Zhang, X.K. Direct Liquid Chromatography/Mass Spectrometry Analysis for Complete Characterization of Recombinant Adeno-Associated Virus Capsid Proteins. *Human Gene Therapy Methods* **2017,** *28*(5), 255–267.
8. Liu, A.P.; Patel, S.K.; Xing, T.; Yan, Y.; Wang, S.; Li, N. Characterization of Adeno-Associated Virus Capsid Proteins Using Hydrophilic Interaction Chromatography Coupled with Mass Spectrometry. *Journal of Pharmaceutical and Biomedical Analysis* **2020,** *189,* 113481.
9. Kiyonami, R.; Srzentić, K.; Viner, R.; Liu, W.; Du, M.; Irvin, K.; Thompson, K.; Liu, C.Y.; Zmuda, J. Thermo Fisher Scientific, Application Note 000253: Highly sensitive intact mass analysis of AAV capsid proteins using a UHPLC-FLD-HRAM MS platform.
10. Kiyonami, R.; Sutton, J.; Irvin, K.; Cook, K.; Toole, E.; Du, M. Thermo Fisher Scientific, Application Note 000242: Adeno-associated virus capsid protein characterization and host cell protein profiling using micro-flow UHPLC-Orbitrap MS.
11. Thermo Fisher Scientific, Orbitrap Ascend Tribrid mass spectrometer.
12. Thermo Fisher Scientific, Gene Therapy Analysis.
13. Srzentić, K.; Damoc, E.; Günder. M. Thermo Fisher Scientific, ASMS 2021 Poster 66125: Characterization of Adeno-Associated Viral assemblies on an Ultra-High Mass Range Hybrid Quadrupole-Orbitrap Mass Spectrometer.

Learn more at **thermofisher.com/genetherapyanalysis**