## thermoscientific

# Characterization of Adeno-associated viral proteins and related proteoforms using topdown approach on a LC-Orbitrap Tribrid MS platform

## ABSTRACT

Purpose: Develop a top-down AAV viral protein mass analysis method for the phosphosite localization of VP2 and protein sequence confirmation of VP3 clip.

Methods: An AAV6 sample was analyzed with a targeted top-down approach on a Thermo Scientific<sup>™</sup> Orbitrap<sup>™</sup> Ascend Tribrid<sup>™</sup> mass spectrometer coupled with a Thermo Scientific<sup>™</sup> Vanguish<sup>™</sup> Horizon UHPLC system using multiple fragmentation techniques. The top-down data were processed using the Thermo Scientific<sup>™</sup> BioPharma Finder<sup>™</sup> 5.1 Software.

Results: The developed top-down method was able to localize two phosphorylation sites unambiguously by the multiple fragmentation techniques including EThcD, HCD and UVPD. The multiple fragmentation methods also allowed confident VP3 clip sequence confirmation with increased sequence coverage.

## INTRODUCTION

AAV capsid viral proteins (VPs) are important constituents of AAV products and play an important role in immunogenicity and tissue tropism in gene therapy. In-depth characterization of viral proteins, including their truncated forms and PTMs, is critical for developing and manufacturing AAV products to ensure their safety, quality, and efficacy. In our previous work, we applied multiple LC-MS/MS approaches (Intact, top-down and peptide mapping) for AAV6 viral protein characterization on a Thermo Scientific<sup>™</sup> Orbitrap Exploris<sup>™</sup> 480 MS<sup>1,2</sup> and identified multiple potential phosphorylation sites with peptide mapping approach.<sup>2</sup> However, it was difficult to localize the accurate site of the phosphorylation with the peptide mapping approach only because of the high sequence overlaps of the VP1 and VP2. Orbitrap Ascend Tribrid MS enables multiple fragmentation techniques with higher ion transfer efficiency for the high m/z ions and provides an ideal platform for the top-down analysis of intact proteins. In this study, we performed targeted top-down analyses using EThcD, HCD and UVPD on the HPLC-Orbitrap Tribrid MS platform for the phosphosite localization of VP2 and protein sequence confirmation of VP3 clip. For the targeted precursor ion selection, an additional intact protein analysis was carried out. The results are reported here.

### MATERIALS AND METHODS

### Sample Preparation

An AAV6 sample expressed in HEK293 cells was used. The AAV6 sample was buffer exchanged and concentrated into 80% H<sub>2</sub>O/20% acetonitrile containing 5 mM TCEP and 0.1% formic acid using 30K centrifugal filter (Amicon® Ultra, 0.5 mL). The estimated protein concentration of the AAV6 sample was 0.2 µg/µL after buffer exchange and concentration.

### **HPLC** conditions

For all experiments, chromatographic separations were carried out using a C4 stationary phase column (1.0 x 100 mm, 300 Å, 1.7 µm) on the Vanquish Horizon UHPLC system. 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. The Orbitrap Ascend Tribrid mass spectrometer equipped with the HMRn+, UVPD and ETD options was used for MS data collection. The ESI source settings were shown in the Table 2. The MS settings for the intact mass analysis are shown in Table 3. The MS settings for the top-down mass analysis are shown in Table 4.

### Data Analysis

Thermo Scientific<sup>™</sup> BioPharma Finder<sup>™</sup> 5.1 software was used for both intact protein and top-down data analysis.

#### Table 1. HPLC gradient condition

		0/5
Time	Flow ( ml/min)	%B
0	0.05	20
1	0.05	32
16	0.05	36
20	0.05	80
21.5	0.05	80
22	0.05	20
30	0.05	20

Sheath gas	25
Aux gas	5
Sweep gas	0
Spray voltage (+V)	3500
Capillary temp. (°C)	320
Vaporizer temp. (°C)	125

Table 3. MS settings for intact		
Scan range (m/z)	800-5000	
Application mode	Intact	
Pressure mode	Low	
Resolution	15000 at <i>m/z</i> 200	
RF lens (%)	60	
AGC target value	150	
Max inject time (MS)	50	
Microscans	10	
Source fragmentation (V)	15	

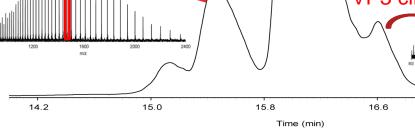


High mass range: 800-5	000 mz; Intact low pressure
150 % AGC target value	e; 5 micro scan; 15 v source f
Targeted precursor ion	
Isolation window	
	571.0
Fragment type	EThcD
	reaction time:8 ms; Max. r
Fragment parameters	inj.time:200 ms; Reagent f
	7E5; SA collision energy (S
	, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
Mass range	Normal
Scan range (m/z)	400-1400/1400-2000/400
Application mode	
Pressure mode	
Resolution (at m/z 200)	
RF lens <mark>(</mark> %)	
AGC target value	
Max inject time (MS)	
Microscans	

### Results

Precursor ion selection for the targeted top-down experiments An intact mass analysis was carried out first for the precursor ion selection. 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 charge 46 of VP2 clearly showed two phosphorylation sites (Figure 1, let insert). In order to generate the fragment ions directly linked to the two phospho-proteomeforms, m/z 1441.5 (the second phospho-proteoform) was selected as the precursor ion for all targeted top-down experiments. By using the isolation window of 6, the m/z 1439.65 (representing the first phospho-proteoform) was also co-selected. The selected precursor ion with isolation window of 6 also covers the charge 30 (m/z 1441.69) of VP3 clip and the acquired top-down data can be used not only for the phosphosite localization of VP2, but also for the protein sequence verification of VP3 clip.

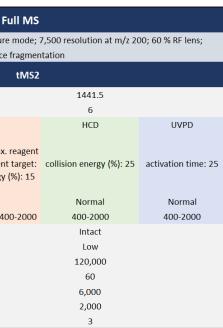
### 1437.86 46+ +1 phospho +2 phospho 1439.65 1441.50 M/M1435.0 1437.0 1439.0 1441.0 1443.0 1445.0 1380 m/z VP2 VP3 clip

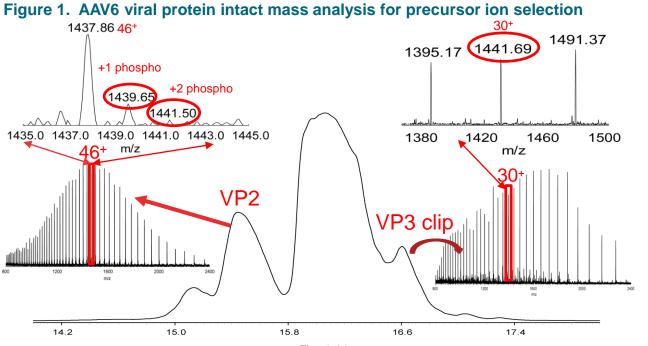


## Reiko Kiyonami<sup>1</sup>, Rafael Melani<sup>1</sup>, Roberto Gamez<sup>1</sup>, Kenneth Thompson<sup>2</sup>, Chao Yan Liu<sup>2</sup>, Min Du<sup>1</sup>, <sup>1</sup>Thermo Fisher Scientific, San Jose, CA, USA; <sup>2</sup>Thermo Fisher Scientific, Frederick, MD, USA

#### Table 2. ESI settings

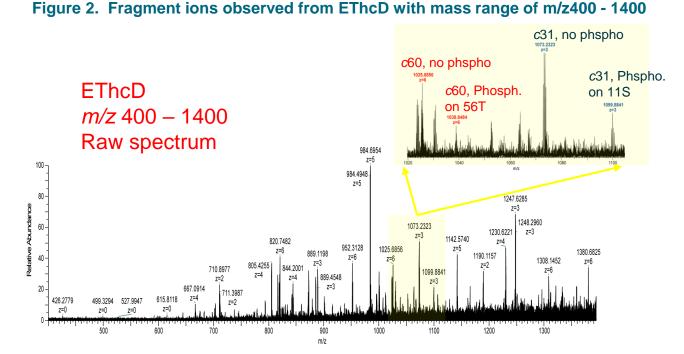
#### Table4. MS settings for top-down





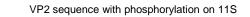
#### Confident localization of two phosphorylation sites from VP2 using multiple fragmentation methods

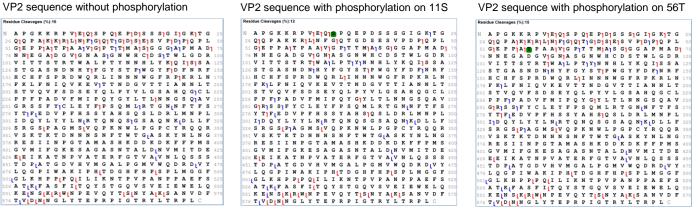
EthcD, HCD and UVPD were performed for the targeted top-down MS<sup>2</sup> experiments. A full MS scan was also performed in the same LC-MS run to ensure only the MS<sup>2</sup> data which are relevant to the VP2 or VP3 clip were used for data processing. Aimed to get more precursor ions into the c-Trap, high AGC value and up to 2000 ms ion injection time were used in all experiments. In order to get more fragments ions that are directly relevant to the phosphoproteome of VP2, we collected EThcD data using three different mass ranges. Figure 2 shows the fragment ions of EThcD collected with a mass range of m/z 400-1400. The *c*-type fragment pairs with and without phosphorylation in the specific amino acid locations were identified clearly (Figure 2, insert).



A total of five top-down raw files (Three EThcD, one HCD and one UVPD) were processed using the top-down workflow of Biopharma Finder 5.1 software. The fragment ion data across the eluted VP2 peak was averaged per raw file and the averaged spectra were searched against three VP2 proteoforms (VP2 proteoform without phosphorylation, VP2 phosphoproteoform with phosphorylation on the 11S and VP2 phospho-proteoform with phosphorylation on the 56T) using 15 ppm mass tolerance for the sequence coverage and phosphorylation localization confirmation. Figure 3 shows all three VP2 proteoforms were identified unambiguously with specific fragment ions using one of the EThcD (m/z 400 -1400) data. The HCD and UVPD data search results provided complementarity fragment ions for supporting the EThcD search results with increased sequence coverage. 26% sequence coverage was observed for the VP2 phospho-proteoform (11S) using UVPD (Figure 4). Overall, 50%, 45% and 47% sequence coverage were observed with the combined EThcD, HCD, UVPD data for the three VP2 proteoforms. Figure 5 shows the combined fragment map of the VP2 phospho-proteoform (56T) with the 47% sequence coverage.

#### Figure 3. Fragment maps observed from EThcD with mass range of m/z 400 - 1400





### REFERENCES

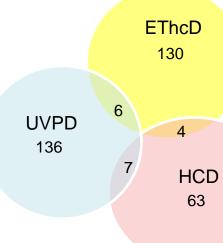
- 1. Thermo Scientific application note # 000253. Reiko Kiyonami, et al., Highly sensitive intact mass analysis of AAV capsid proteins using a UHPLC-FLD-HRAM MS platform.
- 2. Thermo Scientific application note # 000242. Reiko Kiyonami, et al., Adeno-associated virus capsid protein characterization and host cell protein profiling using micro-flow UHPLC-Orbitrap MS.

VP2 phospho-proteoform (11S)	Figure 5. Combined search result for VP2 phospho-proteoform (56T)	
Residue Cleavages (%):26	Residue Cleavages (%):47	
N A P G K K R P V E Q   P Q   E P D S S S G   G K T G 25 24 Q Q P A   K K K L N F G Q T G D S ] S   G Q A P A   K K K A N F G Q P T T M A S G G G A P M A D 75 51 G E P P A T P A A A Q P T T M A S G G G A P M A D 75 76 N N E G A D G V G N A S G P T T M A S G G G A P M A D 75 76 N N E G A D G V G N A S G P T T M A S G G G A P M A D 75 76 N N E G A D C V G N A S G P T T M A S G G G A P M A D 75 76 N N E G A D C V G N A S G P T T M A S G G G A P M A D 75 76 N N E G A D C V G N A S G P T T N A G C J S S A A D 75 76 N N E G A S N D N H Y F G Y S T P W G Y F D F N R F F 10 75 F K U F N I Q V K E V T T N D G V T T I A N N L T 20 76 F V F F S P R D W Q R L I I N N N W G F R P K R L N 175 76 T F E D V P F H I P Q Y G Y L T L N N G S Q A Y 250 76 Y T F E D V P F H S S Y A H S Q S L D R L M N P L 305 76 Y T F E D V P F H S S Y A H S Q S L D R L M N P L 305 76 Y T F E D V P F H S S Y A H S Q S L D R L M N P L 305 76 Y T F E D V P F H S S Y A H S Q S L D R L M N Y L 305 76 Y T F E D V P F H S S Y A H S Q S L D R L M N Y L 305 76 Y T F E D V P Y H Y L N R T Q N Q S G B A Q N K D L L F 325 76 Y T F E D V P Y H Y M Y A T E R F T W I G A S K Y N L N G 375 77 K E S I I N P G T A M A S N F T W T G A S K Y N L N G 375 75 V S K T K T D N N N S N F T W T G A S N Y A V L N L Q Q R 350 76 K E S I I N P G T A M A S N F T W T G A S N Y A V L Y S S A 100 76 K E S I I N P G T A M A S N F T W T G A Y N V Q G R D V Y Y 475 77 L Q G P I W A K [ I P H [ T D G H [ F H P [ S [ V I W N [ I T D E 425 76 L Q G P I W A K [ I P H [ T D G H [ F H P [ S [ L M G [ G F 506 76 L Q G P I W A K [ I P H [ T D G H [ F H P [ S ] L M [ G F 505 76 G L K H [ P [ P Q I L I K K N T [ P V ] A [ N [ P A [ E F S 525 77 K [ X ] N N F [ N K [ [ P K ] Y T S N Y A K S A N V D [ F 575 77 K [ X ] N N F F Y T E [ P R ] I G T R Y L T R P L C 77 [ T V ] N N M G L Y T E [ P R ] F I G T R Y L T R P L C	N A P G K K R P V E Q S P Q E P D S S S G G I G K T G S 20 Q Q P A K K R P V E Q S P Q E P D S S S G G I G K T G S 31 G P A P A K K R L N F G Q T G D S R S V P D P Q P L 1 32 G P A P A M P A D Y A V G P T T M A S G G G A P P P P G P A D T 40 N N E G A D G V G N A S G N W H C D S T W L G D R 1 10 V I T T S Y R T W A L P T Y N N M H L Y K Q I S S A 1 12 S H G A S N D N H Y F G Y S T P W G Y F D F N N R F T 13 H C H F S P R D W Q R L I I N N W G F R P K L N 1 16 F K L F N I Q V K E V T T N D G V T T I A N N L T T 2 20 S T V Q V F S D S E Y Q L P Y V L G S A A Q G C L 2 22 F P F P F A D V F M I P Q Y G Y L I L N N G V R F T F S 2 23 G R N S F Y Q C L E Y F P S Q M L R T G N N F T F S 2 34 G R M S L V T L N R M V Q C C L P Y V L G S S A Q N K D L L F 3 35 I J D Q Y L Y L M R T Q N Q S S A Q N K D L L F 3 36 I J D Q Y L Y Y L M R T Q N Q S G S A A O N K D L L F 3 37 R E S I I N P G T A M A S H K D D K V N L Q G Z S 36 I J D G Y L I Y K L N N S N F T W T I G A S K Y N L M G 37 R E S I I N P G T A M A S H K D D K V N L Q G S S 40 C Y M Y F G K E S A G A S N T A L D N V U D Q S S S 41 C D R W Y W A K I P H T D G Y P A N P L S S 52 G R R S S P A G A S N T A L D N V Y D R O C Y R C S S 36 I V Q Y F S I I N Y G T A M A S H K D D K V N L Q G S S 37 R E S I I N P G T A M A S H K D D K V N D D R D V Y A 40 C C W A I F G K E S A G A S N T A L D N V W D G R D Y Y C A 45 C L C P I W A K I P H T D G H F H P S S S 50 G L K H P P Q F I W A K I P H T D G H F H P S S S 50 G L K H P P Q F I W A K I P H T D G Y A N Y A K I S A N V D C S S 50 S C A T K F A S F I T T E R F G T Y A A K I S A Y N Y A K I S A N Y D S S S S S S S S S S S S S S S S S S	

## fragmentation methods

The fragment ion data across the eluted VP3 clip peak was averaged per raw file and searched against the proposed Vp3 clip sequence<sup>1</sup> using 15 ppm mass tolerance. Multiple fragmentation methods provided significant numbers of unique fragment ions for the confident confirmation of the VP3 clip sequence (Figure 6). The sequence coverage observed using the combined fragmentation fragment ions was 50% (Figure 7).

#### Figure 6. Venn diagram of fragment ions which matched the VP3 clip sequence



### CONCLUSIONS

- increased sequence coverage.

## TRADEMARKS/LICENSING

© 2023 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.

PO2023-46EN

### Increased sequence coverage for the truncated VP3 proteoform using multiple

Figure 7. Combined search result for VP3 clip A P G K K RIP VIEIQISIPIQIEIPIDIS S SIGIIG KITIG QIQIP AIKIKIRILINIFIGIQITIGIDISIEISIVIPIDIPIQIP LI GIEIPIPIAITIPIAIAIVIGIP TITIMIAISIGIGIGIAIPIM A D NINIEIG AIDIG VIGIN A SIGINIWIH CIDISIT W LIGID R V IITIS T RIT WIAILIPITIYIN N HIL YIKIO IISISIAI 6 S T G A S N D N HIYIF GIYIS T PIW GIY FIDIF NIRI HICHFSIPIRIDWQRLINNNIWIGIFIRPKRILIN 76 F K LIFIN I Q V KIEIVITIN D G V TIT I A NIN LIT STVQVFISIDSEYQLPIYVILIGSAHQGCI 6 PIPIF P A DIV F M IIPIOTYIGIYIL TILIN N G SIOIAIV ו GIR S S F YICIL EIY FIPISIQ MILIR T GININ F T F SI 76 Y T<sub>l</sub>FlE D V P FlH S S Y A H SlQ SlL D R L M N PlL I D Q YIL Y Y LIN R T Q N Q SIG S A Q NIKID LIL S R G S P A G M S V Q P K N W L P G P C Y R Q Q R IV SIKTKITDNINN SINFTWTIGIAS KYNLING <sup>°</sup> RESLIIN PGTAMLASH<sub>L</sub>KDDLKDLKFFPM S 1 G V M I F G K E S A G A S N T A L D L N V M I T D E 26 E L K A T N P V A T E R F G T V A V N L Q S S S IT DIPIAT G DIV H VM GAL P G M V W Q D R DIV LOGPIWAKIPHITDGHIFHPSPLMGG IG L K HIPIPIPIQII L I KIN TIP VIP AINIPIPIALE F S 26 Α Τ[K[F]A S[F][I]T Q[Y]S[T G[Q[V]S[V]E[I E]W[E]L[Q 51[K]E[N]S[K[R[W]N]P E[V]Q Y[T]S[N Y A]K[S Α N[V D[F 76[T]V]D[N]N[G]L[Y T E[P R[P]I G T R Y L T R P L C

A targeted top-down method that uses optimized mass spectrometer settings and multiple fragmentation techniques was developed to address the challenges of AAV viral protein PTM characterization and truncated AAV protein sequence verification.

The developed method was applied to direct fragmentation of low abundant VP2 phosphoproteofroms and VP3 proteoform with C-terminal truncation (VP3 clip).

Two phosphorylation sites were verified unambiguously from VP2.

The proposed VP3 clip sequence was confidently confirmed

The multiple fragmentation methods provided complementary fragment ions and enabled

