



# Combination of hydrophilic interaction liquid chromatography and top-down mass spectrometry for characterisation of adeno-associated virus capsid proteins

Corentin Beaumal<sup>1</sup> · Felipe Guapo<sup>1</sup> · Josh Smith<sup>1</sup> · Sara Carillo<sup>1</sup> · Jonathan Bones<sup>1,2</sup>

Received: 24 January 2025 / Revised: 28 March 2025 / Accepted: 1 April 2025 / Published online: 21 April 2025  
© The Author(s) 2025

## Abstract

Adeno-associated virus (AAV) viral vector-based gene therapy is advancing rapidly, offering potential treatments for rare and severe diseases. The AAV capsid consists of a combination of three viral proteins (VPs), VP1, VP2, and VP3, ranging from 59 to 81 kDa and present at a theoretical bulk ratio of 1:1:10. This study employed hydrophilic interaction liquid chromatography (HILIC) and mass spectrometry (MS) to achieve robust separation and detailed characterisation of AAV9 capsid proteins. Advanced top-down MS approaches combining multiple fragmentation techniques (HCD, ETD, EThcD, and UVPD) were successfully applied, increasing the sequence coverage up to 40% for VP3 and confirming N-terminal acetylation on VP1 and VP3. The workflow demonstrated high reproducibility between injection duplicates and was subsequently applied to the characterisation of in-house produced biological replicates of AAV9 samples from HEK293 cells, showing consistent results across them. Analysis of AAV9 derived from Sf9 insect cells, a more complex sample due to higher levels of modification of the capsid VPs, further evidenced method versatility. Overall, this study highlights the potential of HILIC-MS and advanced top-down MS approaches for detailed characterisation of AAV capsid proteins.

**Keywords** Adeno-associated virus (AAV) · Top-down mass spectrometry (TD-MS) · Electron transfer dissociation (ETD) · Hydrophilic interaction liquid chromatography (HILIC) · Ultraviolet photodissociation (UVPD) · Gene therapy characterisation

## Introduction

Viral vector-based gene therapy is a rapidly advancing field with promising potential for treating rare and severe diseases [1–3]. This approach utilises selected vectors to deliver target genes to patients, either *in vivo* or *ex vivo*, with the goal of modifying or manipulating gene expression, or altering the biological properties of living cells for therapeutic use, often through the application of only a single dose to achieve long-lasting therapeutic effects [4, 5].

Among the commonly used vectors, adeno-associated virus (AAV) has gained considerable attention due to its favourable safety profile. As a non-pathogenic virus with primarily episomal expression and low integration frequency, AAV exhibits reduced immunogenicity compared to other viral vectors [6]. Wild-type AAV consists of a protein capsid encapsulating a therapeutic single-stranded DNA (ssDNA) genome (4.7 kb), flanked by two inverted terminal repeats (ITR) regions essential for genome replication and packaging. The open reading frames encoded within the capsid regulate replication (*rep*) and capsid protein synthesis and assembly (*cap*) [7–9]. Recombinant AAV (rAAV) vectors retain their ITRs, but *rep* and *cap* genes are replaced with the therapeutic gene of interest in either the ssDNA or in a double-stranded-self complementary (scDNA) conformation [7]. There are at least 12 wild-type serotypes of AAV identified in humans or primates, which differ in their capsid protein structure, thermal stability, and tissue tropism [10–12]. While the underlying mechanism for tissue tropism remains unknown, the AAV's lifecycle has been explored

✉ Jonathan Bones  
jonathan.bones@nibrt.ie

<sup>1</sup> Characterisation and Comparability Laboratory, NIBRT – National Institute for Bioprocessing Research and Training, Foster Avenue, Belfield, Blackrock, Dublin A94 X099, Ireland

<sup>2</sup> School of Chemical and Bioprocess Engineering, University College Dublin, Belfield, Dublin D04 V1 W8, Ireland

**Fig. 1** AAV9 capsid protein sequences. The blue arrow signifies the start of VP1, the orange arrow the start of VP2, and the purple arrow the start of VP3. The grey numbers convey the position of the amino acid residue above them in the VP1 protein sequence. Here, the amino acid residue considered to be the start of VP1 is the alanine residue with the grey 1 under it as the N-terminal methionine (grey M) is cleaved off in the cell

1	M	A	A	D	G	Y	L	P	D	W	L	E	D	N	L	S	E	G	I	R	E	W	W	A	L	K	P	G	A	P	Q	P	K	A	N	Q	Q	H	Q	D	N	A	R	G	L	V	L	P	G	Y	49
50	K	Y	L	G	P	G	N	G	L	D	K	G	E	P	V	N	A	A	D	A	A	A	L	E	H	D	K	A	Y	D	Q	Q	L	K	A	G	D	N	P	Y	L	K	Y	N	H	A	D	A	E	F	99
100	Q	E	R	L	K	E	D	T	S	F	G	G	N	L	G	R	A	V	F	Q	A	K	K	R	L	L	E	P	L	G	L	V	E	E	A	A	K	T	A	P	G	K	K	R	P	V	E	Q	S	P	149
150	Q	E	P	D	S	S	A	G	I	G	K	S	G	A	Q	P	A	K	K	R	L	N	F	G	Q	T	G	D	T	E	S	V	P	D	P	Q	P	I	G	E	P	P	A	A	P	S	G	V	G	S	199
200	L	T	M	A	S	G	G	G	A	P	V	A	D	N	N	E	G	A	D	G	V	G	S	S	S	G	N	W	H	C	D	S	Q	W	L	G	D	R	V	I	T	T	S	T	R	T	W	A	L	P	249
250	T	Y	N	N	H	L	Y	K	Q	I	S	N	S	T	S	G	G	S	S	N	D	N	A	Y	F	G	Y	S	T	P	W	G	Y	F	D	F	N	R	F	H	C	H	F	S	P	R	D	W	Q	R	299
300	L	I	N	N	N	W	G	F	R	P	K	R	L	N	F	K	L	F	N	I	Q	V	K	E	V	T	D	N	N	G	V	K	T	I	A	N	N	L	T	S	T	V	Q	V	F	T	D	S	D	Y	349
350	Q	L	P	Y	V	L	G	S	A	H	E	G	C	L	P	P	F	P	A	D	V	F	M	I	P	Q	Y	G	Y	L	T	L	N	D	G	S	Q	A	V	G	R	S	S	F	Y	C	L	E	Y	F	399
400	P	S	Q	M	L	R	T	G	N	N	F	Q	F	S	Y	E	F	E	N	V	P	F	H	S	S	Y	A	H	S	Q	S	L	D	R	L	M	N	P	L	I	D	Q	Y	L	Y	L	S	K	T	449	
450	I	N	G	S	G	Q	N	Q	T	L	K	F	S	V	A	G	P	S	N	M	A	V	Q	G	R	N	Y	I	P	G	P	S	Y	R	Q	R	V	S	T	T	V	T	Q	N	N	N	S	E	499		
500	F	A	W	P	G	A	S	S	W	A	L	N	G	R	N	S	L	M	N	P	G	P	A	M	A	S	H	K	E	G	E	D	R	F	F	P	L	S	G	S	L	I	F	G	K	Q	G	T	G	R	549
550	D	N	V	D	A	D	K	V	M	I	T	N	E	E	E	I	K	T	T	N	P	V	A	T	E	S	Y	G	Q	V	A	T	N	H	Q	S	A	Q	A	Q	A	Q	T	G	W	V	Q	N	Q	G	599
600	I	L	P	G	M	V	W	Q	D	R	D	V	Y	L	Q	G	P	I	W	A	K	I	P	H	T	D	G	N	F	H	P	S	P	L	M	G	G	F	G	M	K	H	P	P	P	Q	I	L	I	K	649
650	N	T	P	V	P	A	D	P	P	T	A	F	N	K	D	K	L	N	S	F	I	T	Q	Y	S	T	G	Q	V	S	V	E	I	E	W	E	L	Q	K	E	N	S	K	R	W	N	P	E	I	Q	699
700	Y	T	S	N	Y	K	S	N	N	V	E	F	A	V	N	T	E	G	V	Y	S	E	P	R	P	I	G	T	R	Y	L	T	R	N	L														735		

considerably more [8, 13], implicating structural differences, the transgene, and capsid genetic content as potential causes for functional variability observed among serotypes and production systems [14–17].

The icosahedral capsid of AAV is built by a combination of approximately 60 copies of three viral proteins (VPs), VP1, VP2, and VP3, at an approximate bulk ratio of 1:1:10, respectively [6]. However, evidence of a divergent capsid assembly [18] and variability in VP incorporation across expression systems underscore the structural complexity of AAV capsids [14, 19, 20]. Such heterogeneity, along with their large capsid size, makes the characterisation of AAV-based products challenging. To address this, accurate analysis of bulk VP ratios, as well as the location and abundance of post-translational modifications (PTMs), is of utmost importance to aid understanding of the viral activity and to ensure high product safety and efficacy during rAAV manufacturing [21, 22].

Liquid chromatography (LC) coupled to mass spectrometry (MS) has emerged as a powerful tool for the analytical characterisation of biotherapeutics [21, 23–25], and its high sensitivity makes it well suited for in-depth characterisation of samples of limited amount, like AAV-based gene therapy samples [23, 26, 27]. Most of the protein sequences, including the protein C-terminal, are shared between each VP (Fig. 1). Although bottom-up peptide mapping is an efficient way to fully characterise AAV capsid protein sequences, serotype identity, and identify PTMs [20, 28, 29], it fails to localise modifications located in each specific capsid protein when their location falls within their shared C terminal sequence. Alternative splicing gives VP2 an extra N-terminal domain compared to VP3, while VP1 possesses a further N-terminal domain, making it the largest capsid protein with a mass ranging from 80,336 to 82,106 Da depending on serotype [20, 30, 31].

Intact analysis of capsid protein by LC-MS, either under denaturing conditions [19] or followed by a denaturing protocol [20, 32], effectively addresses the challenge of separating VP1, VP2, and VP3. In most cases, this method enables comprehensive characterisation and determination of VP bulk ratios and PTM abundance [14, 19, 20, 30, 32]. When coupled with fluorescence-based detection, this enhances sensitivity further, making it suitable for analysing low-titre samples [27]. Following individual intact VP separation in denaturing conditions, top-down (TD) MS can be employed to fragment each intact proteoform, providing an alternative solution to the challenges encountered with peptide mapping. TD-MS approaches have already demonstrated their advantage over peptide mapping for characterisation of other biotherapeutics, such as determining the location of the payload in antibody-drug conjugates [33–36]. In addition, numerous studies point out the advantages of using multiple fragmentation techniques to increase sequence coverage and get complementary information [37–39]. While higher-energy collisional dissociation (HCD) was traditionally used, various fragmentation techniques, including electron transfer dissociation (ETD), electron transfer/higher-energy collisional dissociation (ET<sub>h</sub>CD), or ultraviolet photodissociation (UVPD), are now fully implemented on commercially available instruments, facilitating combination of the different fragmentation techniques to obtain complementary specific information such as sequence variants or PTM location.

However, the large mass of VPs (from 59 to 81 kDa) complicates dissociation and full elucidation of their primary amino acid sequence when precursor ion fragmentation is achieved by HCD and ETD techniques [40]. By exploring current advancements in ion activation, sequence coverage analysis through top-down MS can be improved, facilitating deep interrogation of the primary structure of proteins to precisely localise mass discrepancies generally related

to degradation events or PTMs within proteins [40]. Using multiple ion dissociation methods and manipulating product ions *via* gas-phase reactions, the feasibility of this application for AAV VP protein characterisation has been recently demonstrated, where the combination of multiple fragmentation methods along with proton transfer charge reduction achieved approximately 40% sequence coverage characterisation for the VP3 protein of AAV2 [41].

In the present study, we explore various fragmentation techniques available on the Thermo Scientific™ Orbitrap Eclipse™ Tribrid mass spectrometer to investigate top-down fragmentation of AAV9 VPs on the chromatographic timescale to determine sequence coverage and identification of PTMs. More specifically, we take advantage of highly informative fragmentation spectra to evaluate fragmentation repeatability across injection replicates and assess the benefits of combining multiple fragmentation techniques to build a solid foundation for using top-down for AAVs. Finally, the optimised method was applied to biological replicates from in-house produced AAV9 samples generated using the HEK293 cell line and then subsequently for the analysis of commercial AAV9 samples derived from the Sf9/baculovirus production system to explore potential differences between production systems.

## Materials and methods

### Chemicals and reagents

All reagents and solvents used were ACS reagent grade or better. Difluoroacetic acid 98% (DFA) was purchased from Arcos Organics (Fisher Scientific, Dublin, Ireland). Optima LC–MS grade acetonitrile (ACN), Thermo Scientific UHPLC-MS-grade water, Gibco™ Viral Production Medium (VPM), Gibco™ GlutaMAX™ Supplement (GlutaMAX™), Gibco™ Viral-Plex™ Complexation Buffer, the Gibco™ AAV-MAX Transfection Kit, Gibco™ AAV-MAX Lysis Buffer (Lysis Buffer), and Gibco™ Pluronic™ F-68 Non-ionic Surfactant (100X) (PF-68) were all sourced from Fisher Scientific (Dublin, Ireland). The Gibco™ AAV-MAX Transfection Kit contains the Gibco™ AAV-MAX Transfection Reagent, Gibco™ AAV-MAX Transfection Booster, and Gibco™ AAV-MAX Enhancer used for transfection. All other chemicals or solvents were obtained from Merck Sigma (Wicklow, Ireland).

### AAV9 production and purification

AAV9 samples from two different production systems were used. AAV9 from a HEK293 cell line production system was produced and purified in-house as described below. AAV9 (empty and full mix) from a Sf9 cell line production system

was purchased from Virovek (Hayward, CA, USA) at a concentration of  $2 \times 10^{13}$  viral particles per mL (vp/mL).

### AAV production and harvest

Gibco™ Viral Production Cells 2.0 (VPCs 2.0; Thermo Fisher Scientific Cat. No. A49784), a clonal derivative of the HEK293 F cell line, were maintained in suspension in VPM supplemented with 4 mM GlutaMAX™ for four passages. The incubation conditions for the cells were as follows: temperature, 37 °C; CO<sub>2</sub>, 8%; humidity, 80%; shake speed (25-mm orbital diameter), 120 rpm. On the final passage before transfection, cells were seeded in a 1 L shake flask at a viable cell density (vcd) of approximately  $0.6 \times 10^6$  cells/mL with a final culture volume of 240 mL. Three days after seeding, triple transfection was performed. The plasmids used for transfection were pAAV-GFP, pAAV2/9, and pAdDeltaF6. pAAV-GFP was a gift from John T. Gray (Addgene plasmid # 32395) while pAAV2/9 and pAdDeltaF6 were gifts from James M. Wilson (Addgene plasmid # 112867 and 112865, respectively) [42].

Triple transfection was performed following the procedure described in the AAV-MAX Helper-Free AAV Production System User Guide. Briefly, on the day of transfection, the VPCs 2.0 were diluted with fresh VPM supplemented with 4 mM GlutaMAX™ to a final vcd of  $3 \times 10^6$  cells/mL and used to generate three 120 mL cultures in 500 mL shake flasks, each with a vcd of  $3 \times 10^6$  cells/mL. 1.2 mL of AAV-MAX Enhancer (1:10 ratio Enhancer/culture volume) was added to each flask. The flasks were then incubated under the conditions described below, while the DNA/transfection complexation process was completed. DNA/transfection complexation was performed using 1.5 µg of total plasmid DNA/mL of culture. A plasmid molar ratio of 1:3:1 pAAV-GFP:pAAV2/9:pAdDeltaF6 was used. For each flask, the required amount of plasmid DNA was combined in a 15 mL tube and then diluted to a final volume of 12 mL (10% of the culture volume) with Gibco™ Viral-Plex™ Complexation Buffer. The solution was then mixed by gently inverting the tube 3–5 times and then stored on ice. Next, 360 µL of AAV-MAX Transfection Booster (3 µL/mL of culture volume) was transferred to a clean 1.5 mL Eppendorf tube where 720 µL of AAV-MAX Transfection Reagent (6 µL/mL of culture volume) was added. The solutions were mixed by gently pipetting up and down 3–5 times. The AAV-MAX Transfection Booster/AAV-MAX Transfection Reagent mixture was next added to the tube containing the diluted plasmid DNA. This complexation was then mixed by gentle inversion of the 15 mL falcon tube 3–5 times and incubated on ice for 20 min. After incubation, the transfection complexation was added to one of the incubating 120 mL cultures. This process was repeated for each of the 3 flasks transfected.

Seventy to seventy-two hours post-transfection, the AAVs from each flask were harvested. To harvest the produced

AAVs, AAV-MAX Lysis Buffer was added to each flask at a 1:10 dilution. This was immediately followed by the addition of  $\text{MgCl}_2$  and Benzonase (Merck, Darmstadt, Germany) to final concentrations of 2 mM and 90 U/mL respectively. Cultures were then incubated for an additional 2 h under the previously described incubation conditions. After incubation, the contents of each flask were centrifuged at  $16,000 \times g$  for 20 min to remove the cell debris. The supernatants were then collected and stored at  $-80^\circ\text{C}$  until ready for further downstream analysis.

### AAV purification

Harvested material was affinity purified as described by Guapo et al. [14]. In short, samples were sequentially filtered through 0.45  $\mu\text{m}$  and 0.2  $\mu\text{m}$  PES membranes (Fisher Scientific, Dublin, Ireland) before purification to remove any remaining cell debris. Purification was performed using an ÄKTA Avant 150 FPLC system using 5 mL POROS<sup>TM</sup> GoPure<sup>TM</sup> AAVX pre-packed columns (0.8  $\times$  10 cm) from Thermo Fisher Scientific (Paisley, UK). Loading was carried out on a flow rate of 1 mL/min, followed by two wash steps using 50 mM Tris, 1 M NaCl, 2 mM  $\text{MgCl}_2$ , and 0.01% PF-68 at pH 7.5 and 50 mM Tris, 0.2 M NaCl, 2 mM  $\text{MgCl}_2$ , and 0.01% PF-68 at pH 4.0 for 10 CVs each, respectively. Following equilibration with 50 mM Tris, 0.2 M NaCl, 2 mM  $\text{MgCl}_2$ , and 0.01% PF-68 at pH 7.5, elution of purified capsids was achieved with 80% 100 mM glycine, 2 mM  $\text{MgCl}_2$ , and 0.01% PF-68 at pH 2.0 for 10 CVs. The eluate was subsequently neutralised with 1.0 M Tris and then buffer-exchanged and concentrated in 0.01% PF-68 in  $1 \times$  PBS solution using 100 kDa Amicon<sup>®</sup> Ultra-0.5 filter devices (Merck, Darmstadt, Germany).

### Titre determination by SEC-LC-FLR

Size exclusion chromatography was carried out on a Thermo Scientific<sup>TM</sup> MABPac<sup>TM</sup> SEC-1 column (5  $\mu\text{m}$ , 2.1  $\times$  150 mm, 300 Å pore size, Thermo Fisher Scientific, Sunnyvale, CA, USA), with a flow rate of 75  $\mu\text{L}/\text{min}$  under isocratic conditions for 8 min with a buffer constituted of 50 mM sodium phosphate dibasic heptahydrate and 300 mM sodium chloride, pH 6.8. AAV titre determination was carried out based on triplicate measurements of peak areas quantified by fluorescence detection (excitation wavelength ( $\lambda_{\text{ex}}$ ) set to 280 nm, emission wavelength ( $\lambda_{\text{em}}$ ) set to 348 nm, and detector sensitivity set to 4) based on the assumption that capsids present elute at the exclusion limit for the SEC phase used. A standard curve was constructed by injecting different amounts ( $5 \times 10^9$ ,  $1 \times 10^{10}$ ,  $2.5 \times 10^{10}$ ,  $5 \times 10^{10}$ ,  $1 \times 10^{11}$  vp/column, respectively) in triplicate of a reference sample of AAV of known concentration. Peak areas were integrated using a standard processing method on Thermo Scientific<sup>TM</sup>

Chromeleon<sup>TM</sup> Chromatography Data System software version 7.2.10 ES. Using the calibration curve, the estimated capsid on a column of samples of unknown concentration was calculated following the fitting of a linear trendline with a correlation coefficient above 0.999. Capsid titre (vp/mL) could then be inferred using the quantitative experiment data and the volume injected.

### Instrumentation

Full-length AAV capsid protein separation was performed on a Thermo Scientific<sup>TM</sup> Vanquish<sup>TM</sup> Flex ultrahigh pressure liquid chromatography (UHPLC) instrument (Thermo Fisher Scientific, Germering, DE) consisting of a Quaternary Pump F (VF-P20-A), Split Sampler FT with 25  $\mu\text{L}$  autosampler loop (VF-A10-A), Column Compartment H (VH-C10-A), and Fluorescence Detector F (VF-D50-A), coupled to an Orbitrap Eclipse<sup>TM</sup> Tribrid mass spectrometer through a standard flow Thermo Scientific<sup>TM</sup> Ion Max NG<sup>TM</sup> ion source containing a heated electrospray ionisation (H-ESI) probe (Thermo Fisher Scientific, San Jose, CA, USA). UHPLC instrument module settings were as follows: autosampler temperature set to  $5^\circ\text{C}$ ; column compartment temperature set to  $25^\circ\text{C}$ ; fluorescence detector excitation and emission wavelengths set to 280 nm and 348 nm, respectively, and detector sensitivity set to 4. Instrument modules were controlled with Thermo Scientific<sup>TM</sup> Xcalibur<sup>TM</sup> software version 4.7.

### Hydrophilic interaction chromatography (HILIC)MS and TD-MS experiments

VP separation of AAV9 was performed using an Acquity UPLC Glycoprotein BEH Amide Column, 300 Å, 1.7  $\mu\text{m}$ , 2.1  $\times$  150 mm (Waters Corporation, Milford, MA, USA) kept at  $25^\circ\text{C}$  following a modified HILIC gradient [19]. Mobile phase A was UHPLC-MS grade water containing 0.1% (v/v) DFA, and mobile phase B was Optima LC-MS grade ACN containing 0.1% (v/v) DFA. Briefly, separation was performed at a flow rate of 0.1 mL/min, starting with an isocratic hold at 85% B for 3.5 min, followed by a decrease to 65% B over 0.1 min, and an isocratic hold at 65% B for 4.4 min. Then, a linear decrease from 65 to 58% B was performed over 21 min. The wash step was performed by simultaneously increasing the flow rate to 0.26 mL/min and decreasing mobile phase B to 0%, followed by a hold at 0% B for 3.5 min. Finally, column re-equilibration consisted of an increase to 85% B over 0.5 min, followed by an isocratic hold at 85% B and 0.26 mL/min for 5 min, and a decrease to 0.1 mL/min for 1 min. A volume of 1  $\mu\text{L}$  and 1.5  $\mu\text{L}$  of  $1.5 \times 10^{14}$  vp/mL (determined by SEC-FLR as mentioned in “Instrumentation” Section) AAV9 HEK293 (7  $\mu\text{L}$  and 10  $\mu\text{L}$  for AAV9 from Sf9) was injected for intact MS and top-down experiments, respectively. MS



acquisition was performed using an Orbitrap Eclipse™ Tribrid mass spectrometer instrument equipped with UVPD and ETD/PTCR options. All the measurements were performed using intact protein mode at low pressure (1 mTorr in ion-routing multipole) and using the following source parameters: positive ion spray voltage was set to 3000 V, sheath and auxiliary gas were set to 25 and 8 (arbitrary units), respectively, RF lens was set at 125%, ion transfer tube temperature was 320 °C, and vaporiser temperature 150 °C. Full scan MS spectra were collected in profile mode at 15,000 resolution (at 200  $m/z$ ) using 5 microscans and a 700 to 2500  $m/z$  mass range. AGC target set to  $2 \times 10^5$  and a maximum injection time to 100 ms. In-source CID dissociation energy was set to 35 V to assist with desolvation. Top-down MS<sup>2</sup> experiments were performed in a targeted design by isolating a single charge state precursor ion in the quadrupole using a 5  $m/z$  isolation window. All details of retention time windows, precursors charge states, and  $m/z$  are provided in Supplementary Table 1. Precursor ions were submitted to HCD, ETD, EThcD, and UVPD fragmentation. Resulting fragmentation spectra were recorded in profile mode without noise removal, using “full profile mode”. HCD used 20% and 30% normalised collision energies (NCE) based on the precursor charge state. ETD used 3 and 6 ms reaction time, and EThcD used the same reaction time with a supplemental activation energy of 15% NCE. The ETD reagent target was set at  $7 \times 10^5$  charges and the maximum ETD reagent injection time was 200 ms. UVPD used 15 and 30 ms activation time. All MS<sup>2</sup> spectra were recorded at a resolution of 120,000 (at 200  $m/z$ ) using 5 microscans, with an AGC target set to  $7.5 \times 10^4$  and maximum injection time to 400 ms.

## Data processing and analysis

Full length AAV9 intact viral protein identification was performed in Thermo Scientific™ BioPharma Finder™ software version 5.2 (Thermo Fisher Scientific, San Jose, CA, USA). Briefly, raw data were deconvoluted using the ReSpect™ deconvolution algorithm with the sliding windows feature, applying a mass tolerance of 20 ppm for VP identification. Details about the deconvolution parameters can be found in Supplementary Table 2.

Processing of top-down data from various fragmentation methods was performed using BioPharma Finder (version 5.2) using the multiconsensus option. MS/MS spectra were deconvoluted using the Xtract™ algorithm included in BioPharma Finder software by applying a signal-to-noise ratio of 7, a fit factor of 80%, and considering only charge states 1+ to 25+. For fragment assignment, a mass tolerance of 10 ppm was applied. Fragmentation maps were obtained from ProSightBP (Proteinaceous, Evanston, IL, USA).

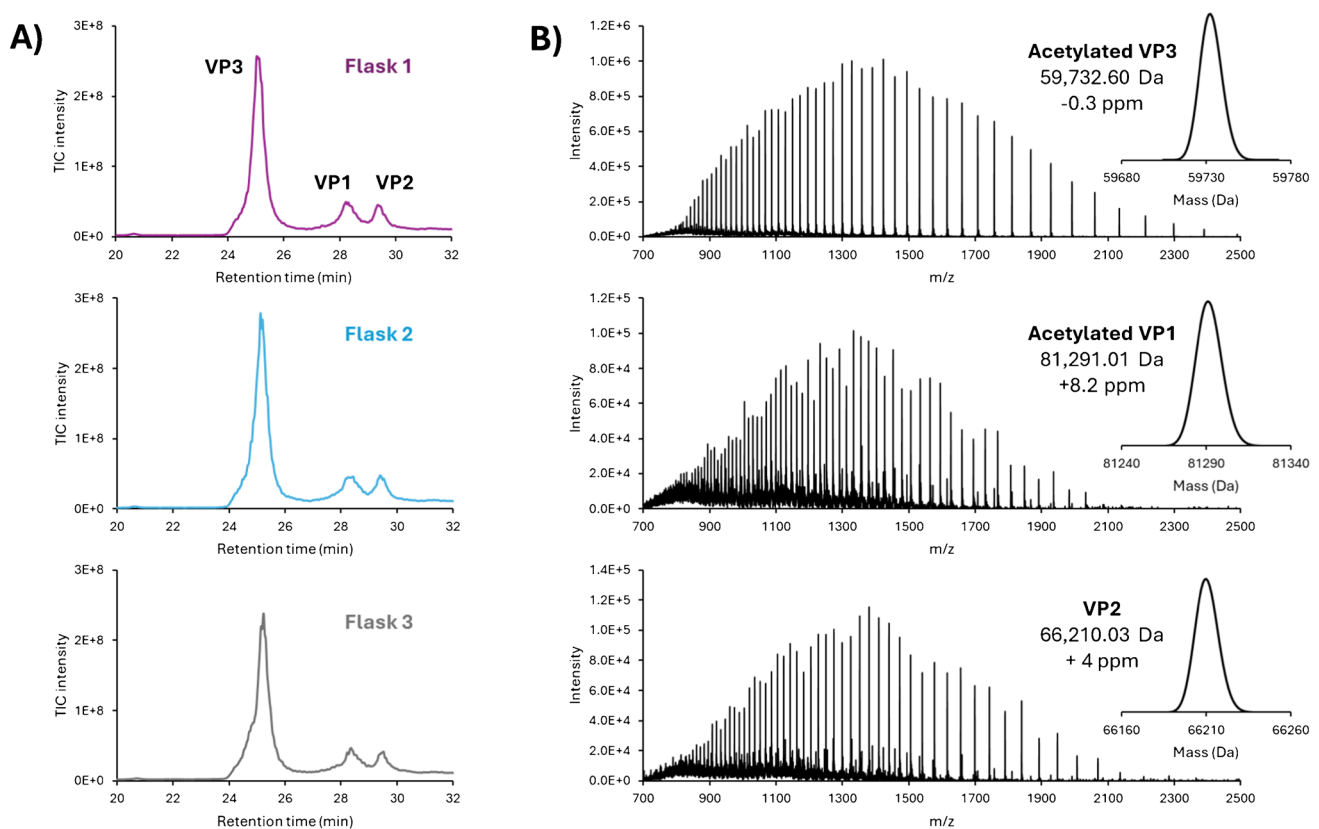
## Results and discussion

### AAV9 VP separation and intact characterisation

Achieving effective separation of AAV VPs is required for determining the VP ratio and identifying potential PTMs. Multiple recent studies have demonstrated the advantages of using HILIC, with DFA as a mass spectrometry compatible mobile phase additive for comprehensive AAV VP separation [19, 20, 32]. In this study, we used a 30-min gradient to separate AAV9 VPs. As shown in Fig. 2A, the complete separation of VP3 was observed, while VP1 and VP2 were partially separated, with high-quality MS spectra obtained for each VP. VP ratio calculations were performed using two methods: FLR-based and MS-based. For the FLR-based approach, VP stoichiometry was calculated using the integrated area under the elution peak of each VP. For the MS-based approach, ratios between VPs were calculated using each VP's area in the extracted ion chromatogram (XIC). In both cases, an approximate VP bulk ratio of 1:1:10 between VP1, VP2, and VP3 was found (Supplementary Table 3). This finding aligns with the expected theoretical ratio between capsid proteins for functioning AAV9 as documented in the literature [43]. Each VP exhibited a high-quality signal, which is crucial for accurate deconvolution. Notably, charge states 38+ to 95+ for VP1, 30+ to 77+ for VP2, and 24+ to 74+ for VP3 were observed and used for deconvolution. The results indicated excellent mass accuracy for the deconvoluted masses of each VP, with a mass error lower than 10 ppm for each (Fig. 2B). Furthermore, while VP2 was identified in its intact form, VP3 and VP1 were found to be acetylated (+ 42 Da). This observation is consistent with findings reported in other studies, whether for the AAV9 serotype, or for other serotypes [19, 30, 44].

### Top-down characterisation

Due to the high sequence homology among viral capsid proteins, their characterisation using a standard peptide mapping approach proves to be challenging [28, 45, 46]. To overcome this limitation, the combination of intact VP separation with TD-MS has emerged as a promising strategy. This approach allows for the characterisation of each individual VP and the identification of their potential PTMs [41]. Using the high reproducibility of chromatographic retention times, we developed a targeted method for each VP. A single charge state precursor was selected and then fragmented using various fragmentation techniques available on the Orbitrap Eclipse™ Tribrid instrument, including HCD, ETD, EThcD, and UVPD. We utilised two different activation energy/reaction times per fragmentation technique to induce fragmentation in different



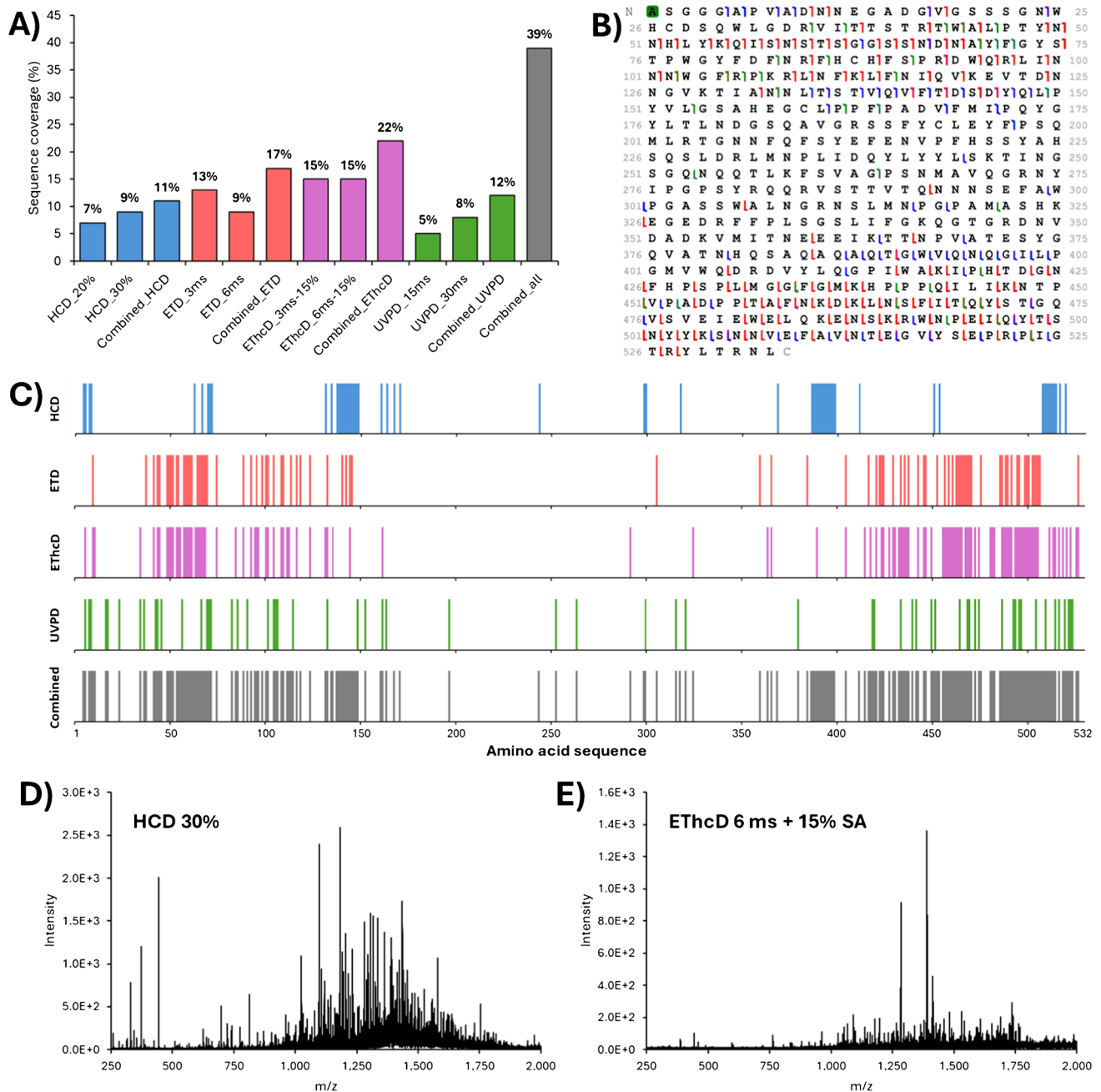
**Fig. 2** A) Total ion chromatogram (TIC) of AAV9 VP HILIC separation for the three biological replicates grown in parallel flasks 1–3. B) MS spectra of acetylated VP3 (top), acetylated VP1 (middle), and

VP2 (bottom) from the biological replicate 1. Insets represent the deconvoluted spectrum, the experimental average mass, and the deviation from the theoretical average mass

regions of the protein. As mentioned previously, the abundance and size of the VPs vary significantly, which can impact fragmentation efficiency. It is well established that the fragmentation of larger proteins is complex and often results in poor sequence coverage [47]. Here, VP3 (59.7 kDa) is smaller than VP2 (66.2 kDa) and VP1 (81.3 kDa) and has an intensity ten times greater than that of the other VPs, which enhanced fragment detection. The sequence coverage for each VP obtained from each run is illustrated in Fig. 3A and Supplementary Figure 1. Additionally, results from combining the same fragmentation technique (e.g., the two HCD runs) and of all the runs together (i.e., a total of 8 runs) are presented. The data indicates that fragmentation efficiency differs both between fragmentation techniques and among the different VPs. Specifically, VP3 exhibits the highest sequence coverage across all the techniques, achieving a sequence coverage of 39% when combining all the fragmentation techniques ( $n = 8$  runs, Fig. 3B). In comparison, considering the same number of runs, sequence coverage for VP1 and VP2 is 5% and 10%, respectively. These observations align with the expected higher intensity of the VP3 precursor and its smaller size. Similarly, although the precursor abundances for VP1 and VP2 are comparable, the smaller size of VP2 facilitates

fragmentation and leads to higher sequence coverage than observed for VP1.

Next, we examined the differences among the fragmentation techniques. For VP3, electron-based fragmentation (i.e., ETD and ETHcD) yielded the best results, providing sequence coverage of 17% and 22% for ETD and ETHcD, respectively, when combining results from the two runs for each technique. In addition, combining 2 HCD runs and 2 UVPD runs provided a sequence coverage of 11% and 12%, respectively. When considering all fragmentation techniques together for a single charge state ( $n = 8$  runs), a total sequence coverage of 39% was achieved. These results are consistent with a recently published study on empty AAV2 samples using proton transfer charge reduction after fragmentation [41]. In contrast, for VP1 and VP2, HCD proved to be the most efficient fragmentation technique, yielding to sequence coverages of 3% and 5%, respectively (Supplementary Figure 1). The low sequence coverage observed for VP1 and VP2, particularly for electron-based fragmentations, can mainly be attributed to the low abundance of the precursor ion for these two proteins. This results in low-intensity fragment ions that are less likely to pass the S/N filter during data processing and annotation.



**Fig. 3** Sequence coverage of AAV9 VP3. **A)** Bar plot representing the sequence coverage obtained for each fragmentation, as well as the combination of the two runs from each technique ( $n = 2$ ), and a combination of all the runs ( $n = 8$ ). **B)** Combined fragmentation map of AAV9 VP3. **C)** Bar code representations of fragment location from

HCD (blue), ETD (red), ETHcD (pink), and UVPD (green) runs ( $n = 2$  for each) and for the combination of all the runs ( $n = 8$ ) in the amino acid sequence. Fragmentation spectra of VP3 using **D)** HCD 30% NCE and **E)** ETHcD 6 ms + 15% SA

Another important aspect of this study is the benefit of using multiple fragmentation techniques and energies to improve the sequence coverage and gather information from the entire sequence of the VPs. Figure 3C represents the bar codes for each fragmentation technique and their combination for VP3. For HCD, only certain regions of the sequence are effectively fragmented (e.g., amino acids 135–150, 387–400,

and 508–517), primarily due to the fragmentation pathway of this technique [48, 49]. Conversely, electron-based techniques (i.e., ETD and ETHcD) exhibit a more random fragmentation pattern, with cleavage sites spread between amino acids 40 to 160 and 420 to 510. This difference between HCD and ETHcD fragmentation is directly observed on the fragmentation spectra, as shown for HCD 30% NCE and ETHcD

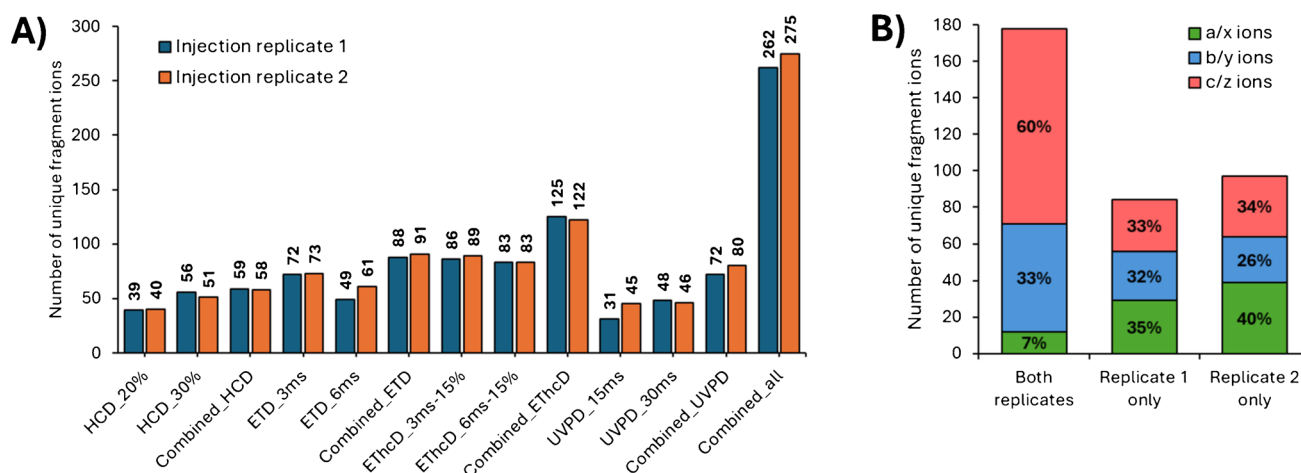
6 ms + 15% supplemental activation energy runs in Fig. 3D and E, respectively. Finally, UVPD generates fragment ions across the entire sequence without any fragmentation preference. In addition, the comparison of the fragment ions using a Venn diagram (Supplementary Figure 2) also highlights the need for multiple fragmentation techniques. Indeed, only 17 fragment ions (8% of all the identified sites) are shared by three or four techniques, while more than 50% of the fragment ions (111 fragment ions) have been only identified by a single fragmentation (36 by HCD, 14 by ETD, 33 by EThcD and 28 by UVPD). It can also be noted that 57 fragment ions are shared by ETD and EThcD, as expected regarding the similar pathway of these fragmentation techniques. The combination of each fragmentation techniques is therefore necessary to bring complementary information that enhances the overall sequence coverage of VPs, as evidenced by the combined bar code shown in Fig. 3C.

To assess the repeatability of our approach, duplicate injections were performed for each fragmentation technique. The number of unique fragment ions and sequence coverage obtained from both injection replicates were compared, with results presented in Fig. 4. For most of the injection duplicates, the number of unique fragment ions identified was very similar, with less than 10% variation across duplicates. Only ETD 6 ms (20%) and UVPD 15 ms (31%) showed a higher variation. This higher variation in these two runs can be attributed to the smaller number of unique fragment ions identified (i.e., 49 and 61 for ETD 6 ms and 31 and 45 for UVPD 15 ms for replicates 1 and 2, respectively). When examining the combined results for each technique, it was found that the number of unique fragment ions remained close between the injection duplicates. Overall, high levels of repeatability were achieved using the method, with 262 unique fragment ions being identified in replicate 1 and 275 in replicate 2. Of these, 178 were shared

between the two replicates, representing 68% and 65% of replicates 1 and 2, respectively (Fig. 4B). Among the ions found in both duplicates, 60% were *c* and *z* ions (mostly from ETD and EThcD), 33% were *b* and *y* ions (mainly from HCD), and only 7% were *a* and *x* (from UVPD), as depicted in Fig. 4B. Contrarily, unique fragment ions from UVPD (i.e., *a* and *x* ions) represent 35% and 40% of the unique fragment ions found only in injection replicates 1 and 2, respectively.

### Analysis of replicate in-house produced AAV9 using triple transfection in HEK293 cells

Three biological replicates of AAV9 samples were independently produced in separate flasks and purified in-house. These three biological replicates were analysed using LC-MS and top-down approaches, as mentioned previously. The separation of the three VPs demonstrated a high reproducibility, with less than 1% retention time variation between the biological replicates (Fig. 2). The VP ratio for each of the biological replicates was compared using the FLR- and MS-based methodologies described above. For all replicates, the observed experimental ratio was approximately 1:1:10 for VP1:VP2:VP3, which aligns with expectations from previous literature and highlights the reproducibility of the AAV9 production method used (Supplementary Table 3). Subsequently, the three biological replicates were analysed in duplicates using the top-down approach described earlier. The results are presented in Table 1, Fig. 5, and Supplementary Figure 3. Specifically, for each flask, we observed an overall sequence coverage of 4–5% for VP1, 7–10% for VP2, and 36–40% for VP3. Moreover, sequence coverage between the injection duplicates of a single biological replicate and among the biological replicates themselves was comparable for all VPs (Table 1).



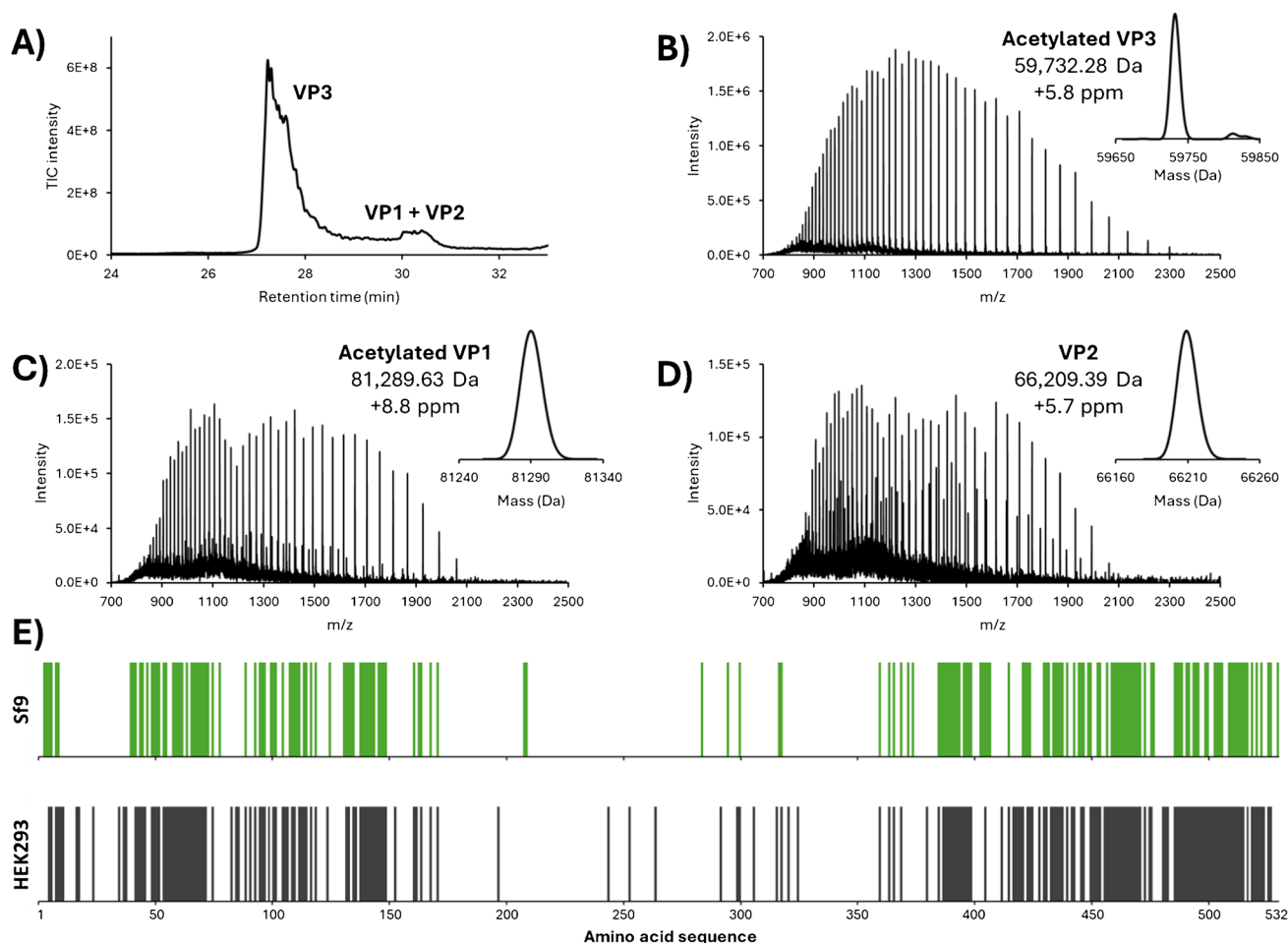
**Fig. 4** **A)** Number of unique fragments ions identified for each run, after combination of the runs from the same fragmentation technique ( $n = 2$ ), and after combination of all the runs ( $n = 8$ ) for injection

replicates 1 and 2. **B)** Number of *a/x* (green), *b/y* (blue), and *c/z* (red) fragment ions found in both injection replicates, or only in one or the other injection replicates after combination of all the runs



**Table 1** Combined sequence coverage of AAV9 VP1, VP2, and VP3 for each biological and injection replicate

Combined sequence coverage	Biological replicate 1		Biological replicate 2		Biological replicate 3	
	Injection 1	Injection 2	Injection 1	Injection 2	Injection 1	Injection 2
VP1	5%	5%	4%	4%	4%	4%
VP2	10%	9%	7%	7%	7%	8%
VP3	39%	40%	36%	39%	38%	39%

**Fig. 5** A) Total ion chromatogram (TIC) of AAV9 from Sf9 VP HILIC separation. MS spectra of B) acetylated VP3, C) acetylated VP1, and D) VP2. Insets represent the deconvoluted spectrum, the experimental average mass, and the deviation to the theoretical aver-

age mass. E) Bar code representations of fragments location for the combination of all the runs (grey,  $n = 8$ ) for AAV9 samples from Sf9 and HEK293 production systems

Additionally, the detailed results for VP3 underline the consistency of our approach across different samples. The coefficient of variation is below 15% for all the individual fragmentations and the combined results regarding sequence coverage and number of unique fragments ions identified (Supplementary Figure 4. A and 4.B, respectively) across all replicates (i.e., including both biological and injection replicates). This variation is particularly low for HCD-based techniques, as fragmentation pathways appeared less random.

### Application to AAV derived from Sf9 insect cells

AAV production methods typically involve either the transient transfection of HEK293 cells or the recombinant baculovirus infection of the *Spodoptera frugiperda* (Sf9) cells. The HEK293 cell line is the most widely used system for AAV production but remains a high-cost approach and suffers from limited scalability [9, 50], while Sf9 production systems often lead to high yields [50, 51] and achieve high full/empty ratios at lower cost. Nevertheless, for Sf9, optimisations are required to achieve the

desired 1:1:10 ratio of VP1:VP2:VP3 [9], and it has been noted that VPs generally have a greater number of PTMs and exhibit more truncated VP species [52, 53]. To explore the differences between the two expression systems, the method developed for the HEK293 cell line was applied to the AAV9 from the Sf9 production system. AAV9 empty reference material produced using baculovirus infections of Sf9 cells was purchased from Virovek (Hayward, CA, USA). We first performed an intact mass analysis of the AAV9 VPs. The results show a good separation of VP3 from other VPs, while VP1 and VP2 are poorly separated, making ratio calculation difficult (Fig. 5). For Sf9 samples, both VP1 and VP3 were found to be acetylated, while VP2 was found in its intact form, as observed for samples from HEK293. The mass error for all the VPs, when comparing the deconvoluted mass to the theoretical mass, was less than 10 ppm (Fig. 5). Several low-abundant species were also identified, corresponding to VP-bearing PTMs such as deamidation, phosphorylation, or oxidation. Although these species are present at low abundance, their existence may account for the low separation efficiency of the VPs and the broad peaks observed in the analysis. Fragmentation of the VPs using the same approach used for samples from HEK293 was also performed. For VP3, the sequence coverage from different fragmentation methods were as follows: 12% for HCD, 22% for ETD, 11% for EThcD, and 2% for UVPD, resulting in a combined sequence coverage of 33%, which is close to the 39% observed previously for HEK293 samples. Comparison of the bar code representations of both production systems confirmed the similarity of the fragmentation pattern, thereby strengthening confidence in our methodology (Supplementary Table 4 and Fig. 5E). The slightly lower combined sequence coverage observed for Sf9 compared to HEK293 might be explained by the possible co-isolation of other low abundant VP3 species precursors (e.g., phosphorylated VP3). This co-isolation could create signal interference, ultimately decreasing the overall signal-to-noise ratio. However, pinpointing phosphorylation on the sequence is extremely challenging due to the very low intensity of the phosphorylated VP3 fragment ions compared to the predominant VP3 species. To do so, a better separation of the different proteoforms would be required. In conclusion, this study on AAV samples from Sf9 demonstrates that our methodology is effective for analysing AAV samples from different production systems. Unfortunately, this approach remains limited for VP1 and VP2, with combined sequence coverage of 2% and 3%, respectively, likely due to their coelution and low abundance.

## Conclusions

This study successfully demonstrates a comprehensive approach to analysing of AAV9 VPs using HILIC-MS-based methods. The methodology enabled the effective separation

of VP1, VP2, and VP3, confirming the theoretical 1:1:10 ratio for the capsid proteins. The combination of multiple fragmentation techniques (HCD, ETD, EThcD, and UVPD) on a single charge state proved to be beneficial and necessary for enhancing sequence coverage on the LC time scale and enabled characterising N-terminal acetylation on VP1 and VP3. Moreover, injection duplicates further validated the repeatability of the methodology, demonstrating minimal variability. It confirms the versatility of the approach and serves as solid foundation for future work on AAVs using top-down approaches. This work also highlights the consistency and the reproducibility of our TD-MS approach across biological replicates, as well as for AAV9 samples from different production systems. The study establishes a robust, versatile, and highly repeatable methodology that offers detailed insights into capsid protein characterisation. Our findings align with recently published literature on AAV2 [41], and additionally highlight the reproducibility achieved when considering biological replicates and production systems, making our approach available to serve as a solid foundation for in-depth VP analysis across other AAV serotypes. Notwithstanding, further optimisation of the LC method and implementation of additional ion-ion reaction techniques (e.g., proton transfer reaction) may be necessary to improve the sequence coverage and PTM location, particularly for low-abundance critical PTMs such as deamidation [54]. Currently, VP1 and VP2 detailed characterisation remains challenging, and single proteoforms characterisation remain for now a limitation due to their low abundance.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00216-025-05874-4>.

**Acknowledgements** The authors gratefully acknowledge Nora Crushell for her help in sample purification. The authors also thank Thermo Fisher Scientific for the provision of funds and instrument access to support this research through a collaborative research agreement.

**Author contribution** Corentin Beaumal: conceptualisation; methodology; validation; formal analysis; investigation; writing—original draft; visualisation. Felipe Guapo: conceptualisation; methodology; investigation; writing—review and editing. Josh Smith: resources; methodology; writing—review and editing. Sara Carillo: writing—review and editing; project administration; supervision. Jonathan Bones: conceptualisation; supervision; project administration; writing—review and editing; funding acquisition. All authors have given approval to the final version of the manuscript.

**Funding** Open Access funding provided by the IReL Consortium.

**Data availability** The data generated and analysed in this study are available from the authors by request.

## Declarations

**Competing interests** The authors declare the following competing financial interest(s): J. Bones received funding to support undertaking this study as part of a funded collaboration between NIBRT and Thermo Fisher Scientific. C. Beaumal, F. Guapo, J. Smith, and S. Carillo are employed on this collaborative project. Beyond this, the authors are not aware of any affiliations, memberships, funding or

financial holdings that might perceive as affecting the objectivity of this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Kessler PD, Podsakoff GM, Chen X, McQuiston SA, Colosi PC, Matelis LA, Kurtzman GJ, Byrne BJ. Gene delivery to skeletal muscle results in sustained expression and systemic delivery of a therapeutic protein. *Proc Natl Acad Sci U S A*. 1996;93(24):14082–7. <https://doi.org/10.1073/pnas.93.24.14082>.
- Samulski RJ, Muzyczka N. AAV-mediated gene therapy for research and therapeutic purposes. *Annu Rev Virol*. 2014;1(1):427–51. <https://doi.org/10.1146/annurev-virol-031413-085355>.
- Au HKE, Isalan M, Mielcarek M. Gene therapy advances: a meta-analysis of AAV usage in clinical settings. *Front Med (Lausanne)*. 2021;8:809118. <https://doi.org/10.3389/fmed.2021.809118>.
- Muhuri M, Levy DI, Schulz M, McCarty D, Gao G. Durability of transgene expression after rAAV gene therapy. *Mol Ther*. 2022;30(4):1364–80. <https://doi.org/10.1016/j.ymthe.2022.03.004>.
- Landhuis E. The definition of gene therapy has changed. *Nature*. 2021. <https://doi.org/10.1038/d41586-021-02736-8>.
- Bulcha JT, Wang Y, Ma H, Tai PWL, Gao G. Viral vector platforms within the gene therapy landscape. *Signal Transduct Target Ther*. 2021;6(1):53. <https://doi.org/10.1038/s41392-021-00487-6>.
- Li C, Samulski RJ. Engineering adeno-associated virus vectors for gene therapy. *Nat Rev Genet*. 2020;21(4):255–72. <https://doi.org/10.1038/s41576-019-0205-4>.
- Wang D, Tai PWL, Gao G. Adeno-associated virus vector as a platform for gene therapy delivery. *Nat Rev Drug Discov*. 2019;18(5):358–78. <https://doi.org/10.1038/s41573-019-0012-9>.
- Wang JH, Gessler DJ, Zhan W, Gallagher TL, Gao G. Adeno-associated virus as a delivery vector for gene therapy of human diseases. *Signal Transduct Target Ther*. 2024;9(1):78. <https://doi.org/10.1038/s41392-024-01780-w>.
- Bennett A, Patel S, Mietzsch M, Jose A, Lins-Austin B, Yu JC, Bothner B, McKenna R, Agbandje-McKenna M. Thermal stability as a determinant of AAV serotype identity. *Mol Ther Methods Clin Dev*. 2017;6:171–82. <https://doi.org/10.1016/j.omtm.2017.07.003>.
- Gigout L, Rebollo P, Clement N, Warrington KH Jr, Muzyczka N, Linden RM, Weber T. Altering AAV tropism with mosaic viral capsids. *Mol Ther*. 2005;11(6):856–65. <https://doi.org/10.1016/j.ymthe.2005.03.005>.
- Gao G, Vandenberghe LH, Alvira MR, Lu Y, Calcedo R, Zhou X, Wilson JM. Clades of adeno-associated viruses are widely disseminated in human tissues. *J Virol*. 2004;78(12):6381–8. <https://doi.org/10.1128/jvi.78.12.6381-6388.2004>.
- Balakrishnan B, Jayandharan GR. Basic biology of adeno-associated virus (AAV) vectors used in gene therapy. *Curr Gene Ther*. 2014;14(2):86–100. <https://doi.org/10.2174/1566523214666140302193709>.
- Guapo F, Donohue N, Strasser L, Boi S, Füssl F, Rainbow-Fletcher A, Getty P, Anderson I, Barron N, Bones J. A direct comparison of rAAV5 variants derived from the baculovirus expression system using LC-MS workflows demonstrates key differences in overall production yield, product quality and vector efficiency. *Int J Mol Sci*. 2024;25(5):2785. <https://doi.org/10.3390/ijms25052785>.
- Heldt CL, Skinner MA, Anand GS. Structural changes likely cause chemical differences between empty and full AAV capsids. *Biomedicines*. 2024;12(9):2128. <https://doi.org/10.3390/biomedicines12092128>.
- Meyer NL, Chapman MS. Adeno-associated virus (AAV) cell entry: structural insights. *Trends Microbiol*. 2022;30(5):432–51. <https://doi.org/10.1016/j.tim.2021.09.005>.
- Zhang R, Xu G, Cao L, Sun Z, He Y, Cui M, Sun Y, Li S, Li H, Qin L, Hu M, Yuan Z, Rao Z, Ding W, Rao Z, Lou Z. Divergent engagements between adeno-associated viruses with their cellular receptor AAVR. *Nat Commun*. 2019;10(1):3760. <https://doi.org/10.1038/s41467-019-11668-x>.
- Wörner TP, Bennett A, Habka S, Snijder J, Friese O, Powers T, Agbandje-McKenna M, Heck AJR. Adeno-associated virus capsid assembly is divergent and stochastic. *Nat Commun*. 2021;12(1):1642. <https://doi.org/10.1038/s41467-021-21935-5>.
- Smith J, Guapo F, Strasser L, Millán-Martín S, Milian SG, Snyder RO, Bones J. Development of a rapid adeno-associated virus (AAV) identity testing platform through comprehensive intact mass analysis of full-length AAV capsid proteins. *J Proteome Res*. 2023;23(1):161–74. <https://doi.org/10.1021/acs.jproteome.3c00513>.
- Zhang X, Jin X, Liu L, Zhang Z, Koza S, Yu YQ, Chen W. Optimized reversed-phase liquid chromatography/mass spectrometry methods for intact protein analysis and peptide mapping of adeno-associated virus proteins. *Hum Gene Ther*. 2021;32(23–24):1501–11. <https://doi.org/10.1089/hum.2021.046>.
- Gimpel AL, Katsikis G, Sha S, Maloney AJ, Hong MS, Nguyen TNT, Wolfrum J, Springs SL, Sinskey AJ, Manalis SR, Barone PW, Braatz RD. Analytical methods for process and product characterization of recombinant adeno-associated virus-based gene therapies. *Mol Ther Methods Clin Dev*. 2021;20:740–54. <https://doi.org/10.1016/j.omtm.2021.02.010>.
- Kontogiannis T, Braybrook J, McElroy C, Foy C, Whale AS, Quaglia M, Smales CM. Characterization of AAV vectors: a review of analytical techniques and critical quality attributes. *Mol Ther Methods Clin Dev*. 2024;32(3):101309. <https://doi.org/10.1016/j.omtm.2024.101309>.
- Serrano MAC, Furman R, Chen G, Tao L. Mass spectrometry in gene therapy: challenges and opportunities for AAV analysis. *Drug Discov Today*. 2023;28(1):103442. <https://doi.org/10.1016/j.drudis.2022.103442>.
- Tassi M, De Vos J, Chatterjee S, Sobott F, Bones J, Eeltink S. Advances in native high-performance liquid chromatography and intact mass spectrometry for the characterization of biopharmaceutical products. *J Sep Sci*. 2018;41(1):125–44. <https://doi.org/10.1002/jssc.201700988>.
- Wu Z, Wang H, Tustian A, Qiu H, Li N. Development of a two-dimensional liquid chromatography-mass spectrometry platform for simultaneous multi-attribute characterization of

- adeno-associated viruses. *Anal Chem.* 2022;94(7):3219–26. <https://doi.org/10.1021/acs.analchem.1c04873>.
26. Fekete S, Aebischer MK, Imiolek M, Graf T, Ruppert R, Lauber M, D'Atri V, Guillaume D. Chromatographic strategies for the analytical characterization of adeno-associated virus vector-based gene therapy products. *TrAC Trends Anal Chem.* 2023;164:117088. <https://doi.org/10.1016/j.trac.2023.117088>.
  27. Green EA, Lee KH. Analytical methods to characterize recombinant adeno-associated virus vectors and the benefit of standardization and reference materials. *Curr Opin Biotechnol.* 2021;71:65–76. <https://doi.org/10.1016/j.copbio.2021.06.025>.
  28. Guapo F, Strasser L, Millán-Martín S, Anderson I, Bones J. Fast and efficient digestion of adeno associated virus (AAV) capsid proteins for liquid chromatography mass spectrometry (LC-MS) based peptide mapping and post translational modification analysis (PTMs). *J Pharm Biomed Anal.* 2022;207:114427. <https://doi.org/10.1016/j.jpba.2021.114427>.
  29. Lam AK, Zhang J, Frabutt D, Mulcrone PL, Li L, Zeng L, Herzog RW, Xiao W. Fast and high-throughput LC-MS characterization, and peptide mapping of engineered AAV capsids using LC-MS/MS. *Mol Ther Methods Clin Dev.* 2022;27:185–94. <https://doi.org/10.1016/j.omtm.2022.09.008>.
  30. Jin X, Liu L, Nass S, O'Riordan C, Pastor E, Zhang XK. Direct liquid chromatography/mass spectrometry analysis for complete characterization of recombinant adeno-associated virus capsid proteins. *Hum Gene Ther Methods.* 2017;28(5):255–67. <https://doi.org/10.1089/hgtb.2016.178>.
  31. Srivastava A, Lusby EW, Berns KI. Nucleotide sequence and organization of the adeno-associated virus 2 genome. *J Virol.* 1983;45(2):555–64. <https://doi.org/10.1128/jvi.45.2.555-564.1983>.
  32. Liu AP, Patel SK, Xing T, Yan Y, Wang S, Li N. Characterization of adeno-associated virus capsid proteins using hydrophilic interaction chromatography coupled with mass spectrometry. *J Pharm Biomed Anal.* 2020;189:113481. <https://doi.org/10.1016/j.jpba.2020.113481>.
  33. Beaumal C, Deslignière E, Diemer H, Carapito C, Cianféroni S, Hernandez-Alba O. Improved characterization of trastuzumab deruxtecans with PTMR and internal fragments implemented in middle-down MS workflows. *Anal Bioanal Chem.* 2024;416(2):519–32. <https://doi.org/10.1007/s00216-023-05059-x>.
  34. Lieu LB, Nagy C, Huang J, Mullen C, McAlister GC, Zabrouskov V, Srzentić K, Durbin KR, Melani RD, Fornelli L. Enhanced payload localization in antibody–drug conjugates using a middle-down mass spectrometry approach with proton transfer charge reduction. *Anal Chem.* 2024. <https://doi.org/10.1021/acs.analchem.4c03872>.
  35. Watts E, Williams JD, Miesbauer LJ, Bruncko M, Brodbelt JS. Comprehensive middle-down mass spectrometry characterization of an antibody–drug conjugate by combined ion activation methods. *Anal Chem.* 2020;92(14):9790–8. <https://doi.org/10.1021/acs.analchem.0c01232>.
  36. Hernandez-Alba O, Houel S, Hessmann S, Erb S, Rabuka D, Huguet R, Josephs J, Beck A, Drake PM, Cianferani S. A case study to identify the drug conjugation site of a site-specific antibody–drug–conjugate using middle-down mass spectrometry. *J Am Soc Mass Spectrom.* 2019;30(11):2419–29. <https://doi.org/10.1007/s13361-019-02296-2>.
  37. Fornelli L, Srzentić K, Huguet R, Mullen C, Sharma S, Zabrouskov V, Fellers RT, Durbin KR, Compton PD, Kelleher NL. Accurate sequence analysis of a monoclonal antibody by top-down and middle-down orbitrap mass spectrometry applying multiple ion activation techniques. *Anal Chem.* 2018;90(14):8421–9. <https://doi.org/10.1021/acs.analchem.8b00984>.
  38. Oates RN, Lieu LB, Srzentić K, Damoc E, Fornelli L. Characterization of a monoclonal antibody by native and denaturing top-down mass spectrometry. *J Am Soc Mass Spectrom.* 2024;35(9):2197–208. <https://doi.org/10.1021/jasms.4c00224>.
  39. Dhenin J, Dupre M, Druart K, Krick A, Mauriac C, Chamot-Rooke J. A multiparameter optimization in middle-down analysis of monoclonal antibodies by LC-MS/MS. *J Mass Spectrom.* 2023;58(3):e4909. <https://doi.org/10.1002/jms.4909>.
  40. Toby TK, Fornelli L, Kelleher NL. Progress in top-down proteomics and the analysis of proteoforms. *Annu Rev Anal Chem (Palo Alto Calif).* 2016;9(1):499–519. <https://doi.org/10.1146/annurev-anchem-071015-041550>.
  41. Kline JT, Huang J, Lieu LB, Srzentić K, Bergen D, Mullen C, McAlister GC, Durbin KR, Melani RD, Fornelli L. Top-down mass spectrometry analysis of capsid proteins of recombinant adeno-associated virus using multiple ion activations and proton transfer charge reduction. *Proteomics.* 2024;25:e2400223. <https://doi.org/10.1002/pmic.202400223>.
  42. Gray JT, Zolotukhin S. Design and construction of functional AAV vectors. *Methods Mol Biol.* 2011;807:25–46. [https://doi.org/10.1007/978-1-61779-370-7\\_2](https://doi.org/10.1007/978-1-61779-370-7_2).
  43. DiMattia MA, Nam HJ, Van Vliet K, Mitchell M, Bennett A, Gurda BL, McKenna R, Olson NH, Sinkovits RS, Potter M, Byrne BJ, Aslanidi G, Zolotukhin S, Muzyczka N, Baker TS, Agbandje-McKenna M. Structural insight into the unique properties of adeno-associated virus serotype 9. *J Virol.* 2012;86(12):6947–58. <https://doi.org/10.1128/jvi.07232-11>.
  44. Zhou Y, Wang Y. Direct deamidation analysis of intact adeno-associated virus serotype 9 capsid proteins using reversed-phase liquid chromatography. *Anal Biochem.* 2023;668:115099. <https://doi.org/10.1016/j.ab.2023.115099>.
  45. Xing T, Li S, Tang S, Huang Y, Liu G, Yan Y, Liu D, Wang S, Zhi L, Shameem M, Li N. Distinct chemical degradation pathways of AAV1 and AAV8 under thermal stress conditions revealed by analytical anion exchange chromatography and LC-MS-based peptide mapping. *J Pharm Biomed Anal.* 2024;251:116452. <https://doi.org/10.1016/j.jpba.2024.116452>.
  46. Qin X, Li X, Chen L, Gao T, Luo J, Guo L, Mollah S, Zhang Z, Zhou Y, Chen HX. Characterization of adeno-associated virus capsid proteins by microflow liquid chromatography coupled with mass spectrometry. *Appl Biochem Biotechnol.* 2024;196(3):1623–35. <https://doi.org/10.1007/s12010-023-04656-x>.
  47. Fornelli L, Parra J, Hartmer R, Stoermer C, Lubeck M, Tsybin YO. Top-down analysis of 30–80 kDa proteins by electron transfer dissociation time-of-flight mass spectrometry. *Anal Bioanal Chem.* 2013;405(26):8505–14. <https://doi.org/10.1007/s00216-013-7267-5>.
  48. Dongré AR, Jones JL, Somogyi Á, Wysocki VH. Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: evidence for the mobile proton model. *J Am Chem Soc.* 1996;118(35):8365–74. <https://doi.org/10.1021/ja9542193>.
  49. Wysocki VH, Tsaprailis G, Smith LL, Brei LA. Mobile and localized protons: a framework for understanding peptide dissociation. *J Mass Spectrom.* 2000;35(12):1399–406. [https://doi.org/10.1002/1096-9888\(200012\)35:12%3c1399::Aid-jms86%3e3.0.Co;2-R](https://doi.org/10.1002/1096-9888(200012)35:12%3c1399::Aid-jms86%3e3.0.Co;2-R).
  50. Liu S, Li J, Peraramelli S, Luo N, Chen A, Dai M, Liu F, Yu Y, Leib RD, Li Y, Lin K, Huynh D, Li S, Ou L. Systematic comparison of rAAV vectors manufactured using large-scale suspension cultures of Sf9 and HEK293 cells. *Mol Ther.* 2024;32(1):74–83. <https://doi.org/10.1016/j.ymthe.2023.11.022>.



51. Wang A, Wang Y, Wu S, Zuo W, Guo C, Hong W, Zhu S. Efficient production of an avian adeno-associated virus vector using insect cell/baculovirus expression system. *J Virol Methods*. 2017;240:26–31. <https://doi.org/10.1016/j.jviromet.2016.11.005>.
52. Rumachik NG, Malaker SA, Poweleit N, Maynard LH, Adams CM, Leib RD, Cirolia G, Thomas D, Stamnes S, Holt K, Sinn P, May AP, Paulk NK. Methods matter: standard production platforms for recombinant AAV produce chemically and functionally distinct vectors. *Mol Ther Methods Clin Dev*. 2020;18:98–118. <https://doi.org/10.1016/j.omtm.2020.05.018>.
53. Galibert L, Savy A, Dickx Y, Bonnin D, Bertin B, Mushimiyimana I, van Oers MM, Merten OW. Origins of truncated supplementary capsid proteins in rAAV8 vectors produced with the baculovirus system. *PLoS One*. 2018;13(11):e0207414. <https://doi.org/10.1371/journal.pone.0207414>.
54. Giles AR, Sims JJ, Turner KB, Govindasamy L, Alvira MR, Lock M, Wilson JM. Deamidation of amino acids on the surface of adeno-associated virus capsids leads to charge heterogeneity and altered vector function. *Mol Ther*. 2018;26(12):2848–62. <https://doi.org/10.1016/j.ymthe.2018.09.013>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.