Fast and high-throughput LC-MS characterization, and peptide mapping of engineered AAV capsids using LC-MS/MS

Anh K. Lam,[1] Junping Zhang,[1] Dylan Frabutt,[1] Patrick L. Mulcrone,[1] Lei Li,[2] Lifan Zeng,[3] Roland W. Herzog,[1] and Weidong Xiao[1]

[1]Department of Pediatrics, Herman B Wells Center for Pediatric Research, Indiana University School of Medicine, Indianapolis, IN 46202, USA; [2]Department of Chemistry, Georgia State University, Atlanta, GA 30302, USA; [3]Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN 46202, USA

Adeno-associated virus (AAV) has emerged as a leading platform for gene therapy. With the skyrocketing rate of AAV research and the prevalence of many new engineered capsids being investigated in preclinical and clinical trials, capsid characterization plays a vital role in serotype confirmation and quality control. Further, peptide mapping the capsid proteins might inevitably be a future requirement by regulatory agencies since it is a critical step in good manufacturing practice (GMP) for biotherapeutic characterization. To overcome many challenges that traditional methods like SDS-PAGE and western blots carry, liquid chromatography and mass spectrometry (LC-MS) allows high resolution and sensitivity with great accuracy in characterizing the AAV capsid proteins. Our optimized LC-MS method provides quick sample preparation, a fast and high-throughput 4-min run, and high sensitivity, which allows for very efficient characterization of wild-type and engineered capsids. This study also reports the usage of LC-MS/MS peptide mapping of AAV capsid proteins to determine the most accessible lysine residues targeted by chemical modifications. Our detailed protocols are anticipated to promote the development and discovery of AAV variants with high accuracy and efficiency.

INTRODUCTION

Since first discovered accidentally in the 1960s, adeno-associated virus (AAV) has emerged as a leading platform for the gene therapy treatment of many monogenic diseases.[1] Aside from the two AAV-based treatments approved by the US FDA, there are currently more than 150 clinical trials worldwide, highlighting the immense potential of AAV vectors.[2]

Despite these successes, challenges such as high vector dose requirements, preexisting neutralizing antibodies against many AAV serotypes, decisions concerning administration route (intravenous [i.v.] infusion, intramuscular injection, aerosol dispersal), and broad biodistribution remain. Most of these problems are related to the capsids of the AAV vectors, which has led to the development of hundreds of new AAV serotypes and engineered capsids with improved characteristics.[3–6] However, the need for understanding the precise chemistry of each AAV capsid protein is unmet. Additionally, protein-based biologic drug products (monoclonal antibodies, antibody-drug conjugates, etc.) are often required to elucidate their primary amino acid sequences by several regulatory agencies. Since January of 2020, the US FDA recommends that all human gene therapy products be well characterized and evaluated to ensure critical quality attributes for potency, identity, and purity.[7] With the skyrocketing rate of AAV gene therapy research and the prevalence of engineered capsid in preclinical and clinical trials, peptide mapping the capsid proteins is a foreseeable future requirement since it is a critical step for biotherapeutic characterization. Traditional methods like SDS-PAGE and western blots can only estimate capsid serotypes with moderate confidence and are time consuming. To overcome these challenges, liquid chromatography and mass spectrometry (LC-MS) allows high sensitivity with great confidence in characterizing the AAV capsid proteins to the exact masses. LC-MS has proven to be a powerful analytical technique by offering a robust method to directly measure the protein molecular weight of the AAV capsid, thus providing both specificity and sensitivity for rapid confirmation of capsid serotypes. With our optimized method, intact LC-MS of AAV capsid can be performed in as little as 4 min, therefore allowing high-throughput sample screening of wild-type and engineered AAV capsids (more than 100 runs per 8 h). Peptide mapping methods for different AAV capsids (including chemically engineered capsids) were also optimized for high sequence coverage and precise amino acid structure identification using MS/MS.

Our methods offer the following advantages: quick sample preparation, a fast and high-throughput 4-min LC-MS run (while other studies required much longer runs ranging from 15 to 40 min[8–10]), high sensitivity from a modest amount of vector titers to characterize

A Intact LCMS characterization
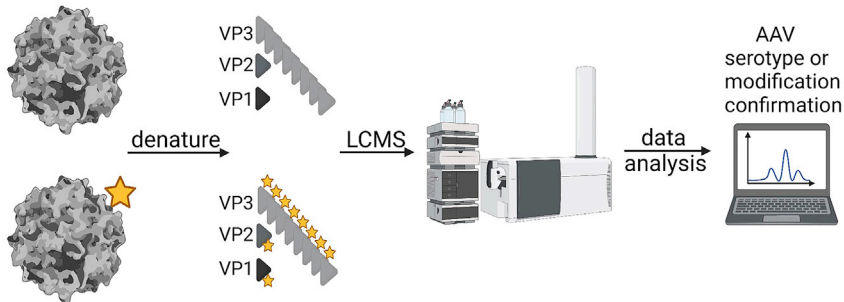
B Peptide mapping AAV VP proteins
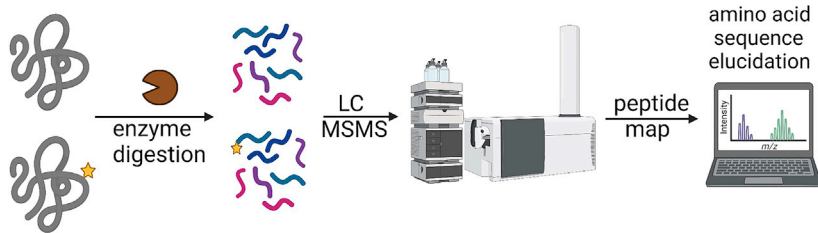
For example, Figure 2 shows an SDS-PAGE of multiple denatured AAV serotypes. The three VP proteins of each serotype were clearly separated in the gel, except for our AAV8, which had an extra band of VP1.5 as previously documented.[11] However, theoretical masses are not reflected in SDS-PAGE results as evidenced by the three AAV2 proteins (possessing the greatest theoretical masses) migrating further than the AAV5 and AAV8 VP proteins. The same principle applies for AAV5x, as the smallest VP3, VP2, and VP1 should migrate the furthest among tested AAVs, but this is not reflected in the predicted mass using SDS-PAGE. Clearly, small molecular weight differences cannot be measured reliably on SDS-PAGE. Western blot using specific antibodies might be possible to confirm each serotype; however, this is time consuming and often not accurate due to the high homology between capsids. Chemically engineered capsids (for example, AAV5x versus AAV5x-Dc) can be even more challenging to separate and characterize (especially with small molecule conjugations) using SDS-PAGE or western blots.

Analytical LC-MS and LC coupled tandem MS (LC-MS/MS) provides specificity, sensitivity, and speed in proteomics analysis and thus offers an ideal tool to distinguish AAV serotypes and map engineered capsid proteins. Here, AAV9 was used for intact LC-MS-related method development. Since AAV vectors are comprised of 60 monomers of VP proteins per capsid with VP1/2/3 in an approximately 1:1:10 stoichiometric ratio, the capsid should be denatured into individual VP proteins before LC-MS injection to ensure proper protein ionization (as demonstrated in Figure 1). A simple acidic condition of 0.1% formic acid (or 0.1% TFA) was added to the AAV samples 10 min before LC injection. Optimized protocols are shown in Tables 3 and 4.

chemically modified capsids, and peptide mapping of engineered AAV capsid proteins. Lys-targeted chemical modification of the AAV capsids, for the first time, has been characterized and peptide mapped in this study, therefore elucidating the most accessible lysine (Lys) residues of the AAV tested. Our detailed protocols (Figure 1 shows the schematic diagrams of the method workflows) are anticipated to promote the development and discovery of AAV variants with high efficiency and accuracy.

RESULTS
Intact AAV capsid characterization using LC-MS

The capsid of each AAV serotype consists of 3 VP proteins: VP1, VP2, and VP3, which share an overlapping open reading frame. VP1 is the largest, with an approximate size of ∼82 kDa, VP2 is around ∼66 kDa, and VP3 is the smallest and most abundant, with a size of ∼60 kDa. Table 1 shows the VPs' exact theoretical masses of different AAV serotypes: AAV2, AAV5x, AAV8, and AAV9. To demonstrate the power of LC-MS analysis, a biological engineered capsid—AAV5x—resulted from a single point mutagenesis of VP3 position #VP3-M277T methionine to threonine, plus several chemical-engineered capsids AAV5x-dansyl chloride (Dc) and AAV9-Dc were included in this study. Dc is a small molecule with a functional group of sulfonyl chloride, which is very reactive toward the amino group of exposed Lys residues on the AAV capsid. One molecule of Dc resulting from a covalent bond formation with an amino group causes a 233.1 Da mass gain of that VP protein.

The exact masses of the VP3 proteins alone can distinguish different AAV serotypes, which an SDS-PAGE cannot provide.

Table 1. Theoretical masses of AAV capsid proteins

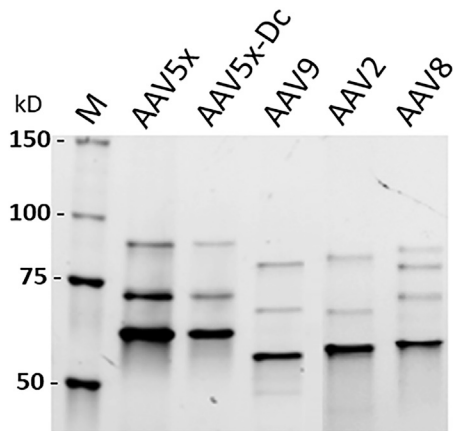|	Theoretical masses of AAV capsid proteins (Da)		
AAV serotype	VP1	VP2	VP3
AAV2	81,856	66,488	59,974
AAV5x	80,306	65,253	59,433
AAV8	81,667	66,519	59,805
AAV9	81,291	66,210	59,733

Figure 2. SDS-PAGE of multiple AAV serotypes
M, marker. 7.5% polyacrylamide gel; ~0.6E10 vg loaded.

For fast and high-throughput purposes, as a highlight in this study, a short column ZORBAX 300SB-C3 4.6 × 50 mm was used, as a single VP protein mass can sufficiently confirm an AAV capsid serotype variant (Table 1). The chromatographic profile of AAV9 is shown in Figure 3; a fast run of LC-MS comes at the expense of coelution of VP1-3 proteins at 2.8–3.0 min. The capsid proteins fully elute at ~3.0 min in each intact LC-MS run, indicating that each run can end at 4 min and still provide a high-throughput value for multiple batch screening and characterization of AAV serotypes. For the VP1/2/3 separation method, Zhang et al.'s protocol[9] is highly recommended. Wild-type AAV9's raw and deconvoluted mass spectra are also shown in Figure 3 (left side). For comparison, chemically engineered AAV9-Dc spectra are shown in Figure 3 (right side). The exact masses of each VP protein were detected as shown by the deconvoluted mass spectra, with VP3 (59,734 Da) displaying the most intense signal, reflecting its relative abundance compared with the other VP proteins, VP2 (66,211 Da) and VP1 (81,292 Da). We calculated the

Table 2. Liquid chromatography parameters for intact and peptide mapping analyses

Parameter	Intact analysis	Digested analysis
Column	Agilent ZORBAX 300SB-C3 4.6 × 50 mm, 3.5 μm	Agilent AdvanceBio Peptide Map 2.1 × 150 mm, 2.7 μm
Mobile phase A	water, 0.1% FA (formic acid)	water, 0.1% FA
Mobile phase B	acetonitrile, 0.1% FA	acetonitrile, 0.1% FA
Flow rate	0.7 mL/min	0.4 mL/min
Injection volume	10 μL (~2E10 vg)	20 μL
Column temperature	40°C	60°C
Gradient	0–0.2 min: 0%–20% B; 0.2–5 min: 20%–50% B; 5.1–6 min: 95%–95% B;	0–18 min: 5%–40% B; 18–28 min: 40%–90% B; 28–36 min: 90%–95% B; 36–40 min: 95% B
Post time	1 min	1 min

Table 3. Mass spectrometry Agilent 6545 LC-QTOF parameters

Parameter	Intact analysis	Digested analysis
Source	Dual Agilent Jet Stream ESI	Dual Agilent Jet Stream ESI
Acquisition mode	positive, 3,200 m/z mass range	positive, 3,200 m/z mass range
Gas temperature, °C	325	290
Gas flow, L/min	10	13
Nebulizer, psig	20	35
Sheath gas temperature, °C	400	275
Sheath gas flow, L/min	12	12
VCap, kV	4	4
Nozzle voltage, kV	2 kV	2
Fragmentor, V	220 V	175
Skimmer, V	45 V	65
Mass range	100–3,200 m/z	100–2,000 m/z (MS); 50–1,700 m/z (MS/MS)
Acquisition rate, spec/s	1.5	5/3 for MS/MSMS
Reference mass	121.0509 and 922.0098	922.0098
Mass tolerance, ppm	30	30
Isolation width	–	medium (~4 m/z)
Precursors/cycle	–	Top 10
Precursor threshold	–	200 counts + 0.01%
Collision energy	–	3.6*(m/z)/100–4.8
Sort precursors	–	by abundance only; +1, +2, +3, >+3
Isotope model	–	peptides
Purity	–	100% stringency and 30% cutoff

stoichiometric ratio of these VP1/2/3 to be 7:9:100 (Figure 3). From the AAV9-Dc deconvoluted mass spectrum, an additional mass gain of 233.1 Da was found as the second most intense peak right next to the VP3 peak. For better data interpretation, Figure 4 shows a close-up look around VP3 peaks of AAV9 and AAV9-Dc.

Figure 4 clearly shows that one *Dc* molecule was chemically conjugated to the VP3 of AAV9 capsid with an exact mass gain of 233.1 Da. This result indicates that regardless of multiple Lys residues (at least 8) predicted to be exposed on the serotype 9 capsid by the structural biology model,[12] only one Lys is the most accessible for chemical conjugation of Dc. Unlike western blots, where small-molecule chemical modifications can only be confirmed indirectly through the attachment of a reporter or epitope tag to the molecule being tested (as discussed in Lam et al.[3]), LC-MS directly measured and confirmed the covalent chemical addition in just a few minutes. The modification ratio is estimated to be ~45%–50%, as the original peak of the wild-type AAV9 VP3 has similar abundance to the modified peak. This is most likely due

Table 4. Mass accuracy of AAV capsid intact LC-MS

AAV serotype	AAV capsid proteins (Da)								
	VP1			VP2			VP3		
	Theoretical	Observed	Mass accuracy ppm	Theoretical	Observed	Mass accuracy ppm	Theoretical	Observed	Mass accuracy ppm
AAV2	81,856	81,855	12.22	66,488	66,490	30.08	59,974	59,975	16.67
AAV5x	80,306	80,309	37.36	65,253	65,255	30.65	59,433	59,434	16.83
AAV8	81,667	81,668	12.24	66,519	66,519	0.00	59,805	59,807	33.44
AAV9	81,291	81,292	12.30	66,210	66,211	15.10	59,733	59,734	16.74

to the sulfonyl chloride group of Dc molecules' instability in aqueous solution, since they hydrolyze quickly. Chemical modification is a common tool in protein engineering, especially to tag a protein with a dye marker or antibody. Many investigators have documented their results using chemical-engineered AAV capsids,[3,13–16] and countless studies investigating biological-engineered capsids have been performed (from directed evolution to structural-guided screening), but none of these capsid proteins were well characterized.[17–19] Our method was shown to provide an invaluable tool for quick and reliable AAV capsid characterization with minimum of ~2E10 vg per run.

Three additional AAV serotypes (AAV2, AAV5x, and AAV8) were analyzed using the same optimized method for intact LC-MS. Their deconvoluted mass spectra are shown in Figure 5. The focus of this study was the development of a LC-MS method for rapid VP3 identification. Still, coeluted VP2 and VP1 were found to greatly match their theoretical masses. These values were used to confirm the serotype of the analyzed AAV samples with excellent mass accuracy for most detected VP proteins, as shown in Table 2. Capsid AAV5x has one point mutation that resulted in a *Met* to *Thr* mutation, which is 30 Da less than the wild-type AAV5 VP3. The results here in Figure 5 confirm that mutation and the masses of VP1/2/3, since they all share the overlapping sequence of VP3. Our study used several different purification methods of AAV vectors including ultracentrifugation, affinity column, and ion exchange column (supplemental information). Although different purification methods were applied, the results from LC-MS/MS remained consistent. Using this technique, we have also tested several hundred engineered AAV capsids, and this method has proven to be productive and reproducible, as it allowed us to quickly screen for newly engineered AAV capsids. Our optimized intact LC-MS method can provide a robust, fast, and high-throughput workflow for both wild-type and engineered AAV capsids.

Peptide mapping analysis of AAV capsids using LC-MS/MS

Several regulatory agencies often require primary amino acid sequences of protein-based biologic drug products (monoclonal antibodies, antibody-drug conjugates, etc.). Because of the emergence of recombinant AAVs (rAAVs) in gene therapy, AAV capsid characterization might soon be a requirement in the near future. LC-MS/MS is one of the most valuable and advanced tools in analytical chemistry to elucidate confident and comprehensive peptide maps including post-translational modifications.

For LC-MS/MS method development, trypsin and chymotrypsin were used to digest AAV capsid proteins. Trypsin cleaves peptides on the C terminus of Lys and Arg, while chymotrypsin cleaves at the carboxyl side of aromatic amino acids Tyr, Phe, and Trp. Due to limited Lys and Arg residues, trypsin gave lower sequence coverage than chymotrypsin. Furthermore, with chemical modifications targeting Lys, cleavage by trypsin might be affected, which is the most likely reason chymotrypsin was superior in this case. Specifically, peptide mapping analysis of AAV9 VP3 resulted in a sequence coverage of 93% (trypsin) and 100% (chymotrypsin) as shown in Figure 6. Even though VP2 and VP1 are lower in abundance compared with VP3 monomers, their sequence coverage was also found to be 94% for VP2 and 93% for VP1 (Figures S1 and S2). An example of a peptide fragment spectrum with annotated b and y ions is shown in Figure 7. These *m/z* values of fragmented ions are unique fingerprint patterns for accurate peptide characterization.

Since chymotrypsin provided higher sequence coverage of AAV9 capsid proteins compared with trypsin, it was used to map the engineered capsid AAV9-Dc. As shown in Figure 8, the top panel shows the total compound chromatogram (TCC), the middle panel is the *m/z* of the ion amino acid position #308–333 detected, and the bottom panel is the fragmented product ions of the parent ion shown in the middle panel. As a result, the site-specific position of amino acid Lys at VP3#326 was found to be conjugated with one Dc molecule, hence the mass gain of 233.1 Da, while the corresponding wild-type sequence did not show the mass gain (which are indicated in the middle and bottom panels of Figure 8). This position was confirmed to be an exposed Lys on the intact capsid of AAV9 using the three-dimensional (3D) structure (*3ux1*)[12] from the Protein Data Bank.

Additionally, peptide mapping analyses were performed for AAV8, AAV5x, and AAV5x-Dc to assess the method's general applicability across AAV serotypes. Our peptide mapping method yielded sequence coverage from 92% (AAV5x VP3) to 100% (AAV8 VP3 and VP2) as shown in Figures S4 and S5. The most accessible amino acid modified with one Dc molecule was found to be Lys VP3#259 from AAV5x-Dc peptide mapping (Figure S5). Minimal percentages of post-translational modifications including phosphorylation, oxidation, acetylation, and deamidation were detected (Table 5), showcasing the power of the LC-MS/MS method for protein characterization. LC-MS was shown to be a superior methodology compared with gel-based assays in protein analyses of the AAV capsids, with great
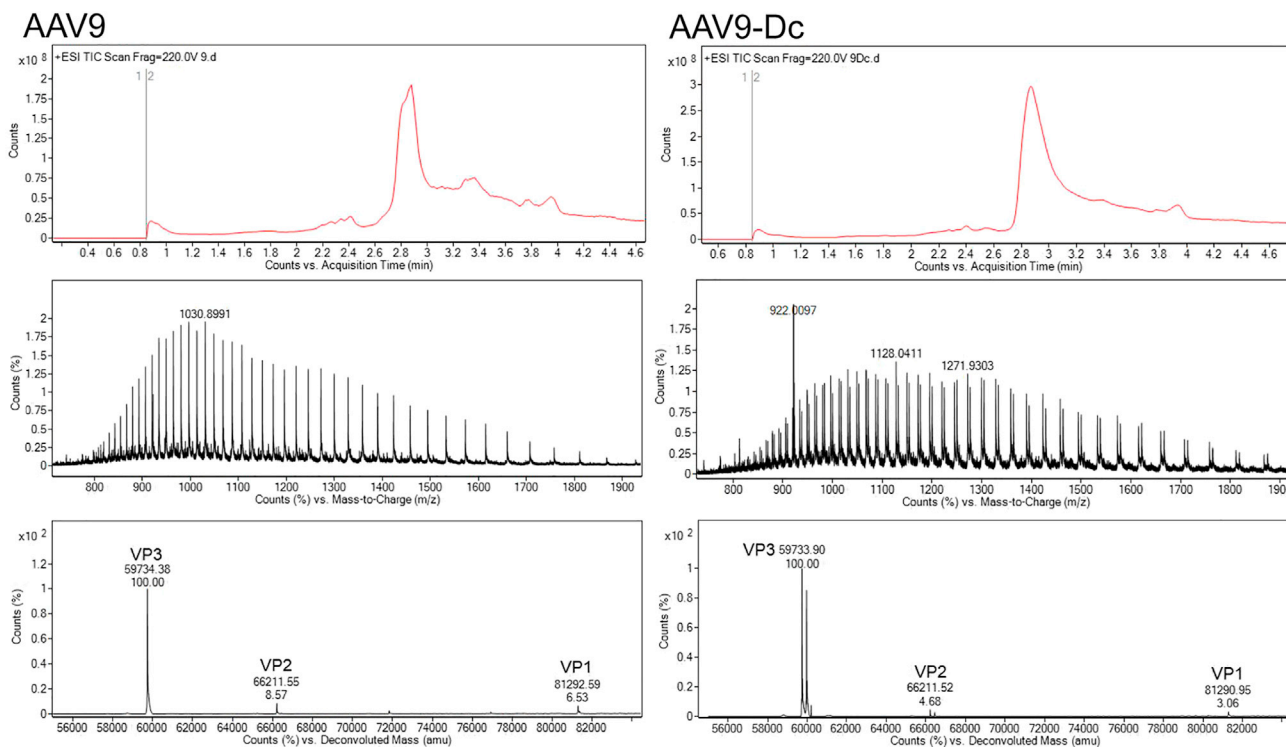
Figure 3. Intact LC-MS of AAV9 (left) and AAV9-Dc (right)

ESI-TIC spectra show the LC profile of capsid proteins (top panel). *m/z* spectra of the most abundant peak (~2.9 min) show the raw MS spectra of the VP3 protein (middle panel). Deconvoluted spectra show the VP3 protein exact mass (plus VP2 and VP1 also detected) and relative abundance percentage of each peak (bottom panel). Chemically engineered capsid AAV9-Dc was characterized with the exact mass of one Dc molecule added to the VP3 protein (see Figure 4 for a zoomed-in spectrum).

specificity and speed to ensure the quality and safety (contaminants in each AAV batch will be detected in LC) of the final vector product. Coupling MS/MS further provided an ideal tool to peptide map the precise amino acid sequence of each capsid protein including chemical conjugations, elucidating more insights about the chemistry of the AAV capsid proteins.

DISCUSSION

rAAV vectors are a safe and effective gene therapy delivery tool capable of treating a variety of diseases. A total of 17 different AAV capsids have been disclosed in clinical trials including wild-type and engineered capsids, and hundreds of new and engineered capsids are being investigated in preclinicals in search for the best vectors to benefit patients.[20] Capsid development and engineering efforts aim to improve AAV vectors to evade existing neutralizing antibodies (NAbs), enhance transduction efficiency, and/or redirect tissue tropism. Despite having the highest prevalence of NAbs (more than 70%),[21] the preferred vector is still AAV2 because it is the most characterized serotype. With more and more novel capsids becoming prevalent in preclinical and clinical trials, capsid characterization could be a tremendous benefit to improve quality control of capsid confirmation and amino acid sequence elucidation, as many of the AAV serotypes differ by a few amino acids. To avoid batch-to-batch contamination and human errors, it is also important to characterize every vector product in process

development as a good manufacturing practice. This requires sophisticated methods that were very time consuming a decade ago. The availability and instrumentation of LC-MS allowed us to develop and optimize a much faster and more reliable method for capsid protein characterization and a more in-depth LC-MS/MS workflow for peptide mapping the capsid proteins (Figure 1).

Traditional methods like SDS-PAGE (Figure 2) cannot distinguish AAV serotypes. Intact LC-MS (Figures 3, 4, and 5) with less than a 5-min run precisely confirmed the capsid serotype and chemical modifications with high mass accuracy (Table 2). This is a faster intact LC-MS protocol compared with previously published studies. For example, the Zhang et al. study[9] required at least ~10 min per run with a requirement of 1E11 vg, and Jin et al.[8] required 30–50 min per run with the use of 1E12 vg. Here, only 2E10 vg was needed for a 4-min run, which enables a more sensitive and higher-throughput outcome, as the number of engineered capsids in an investigational study can range up to the hundreds. Further, intact LC-MS determined the exact mass gain for capsid chemical modification (Figures 3 and 4), providing more understanding of the AAV capsid chemistry. To analyze the amino acid sequence in more details, peptide mapping using LC-MS/ MS for a longer run of 45 min successfully characterized the VP proteins of AAV9 (Figures 6 and 7), AAV5x, and AAV8
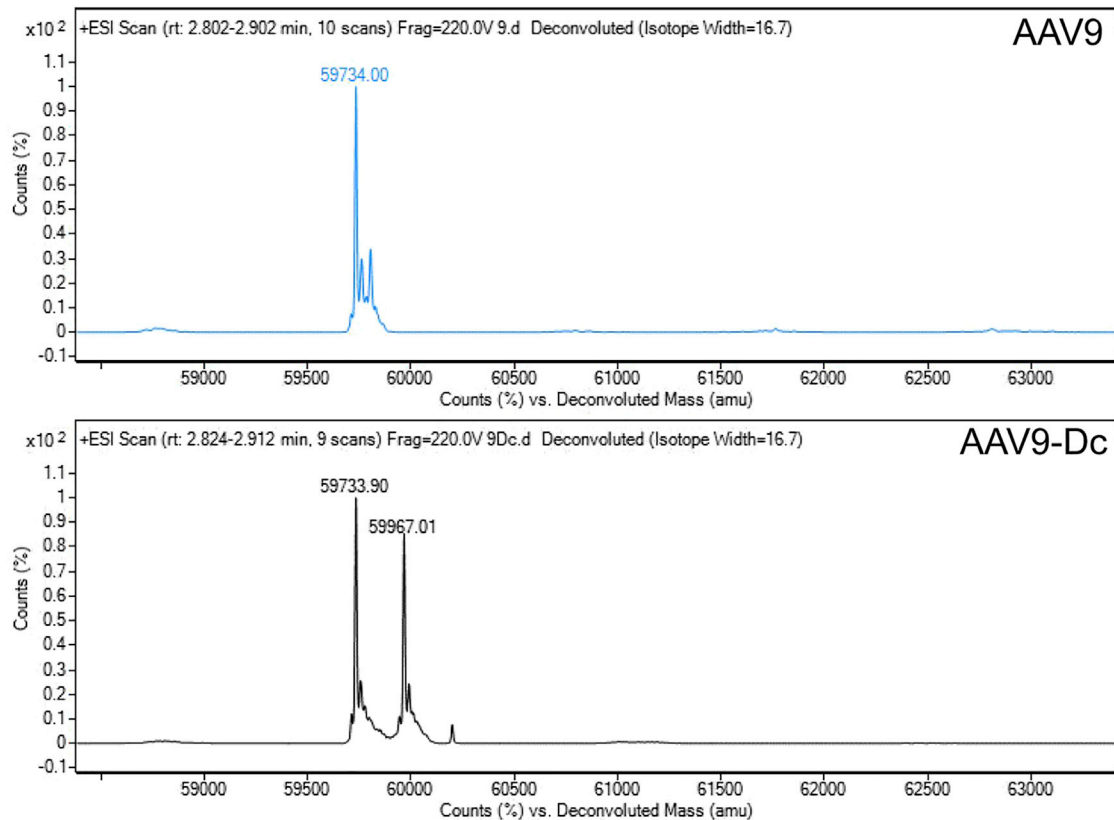
Figure 4. Comparison of LC-MS spectra of AAV9 and engineered-capsid AAV9-Dc

Deconvoluted mass spectra of capsid of AAV9 and capsid AAV9-Dc confirm the exact masses of serotype-9 VP proteins and the chemical-engineered capsid 9-Dcs, with the mass gain of 233.2 Da per one Dc molecule conjugation.

(Figures S4 and S5), with their sequence coverage ranging from 92% to 100%. The most accessible amino acid regarding the Dc-molecule covalent modification was found to be Lys-VP3#326 of AAV9 (Figure 8) and Lys VP3#259 (Figure S5) of AAV5x, highlighting the power and efficiency of LC-MS/MS. Additionally, post-translational modifications (PTMs) were detected in AAV9 VP3; the most common PTMs are listed in Table 5, with the total PTMs being less than 1%, as expected.

In summary, intact LC-MS characterization of AAV serotypes with our optimized method was demonstrated to be successful and very efficient. This method also shows significant improvements compared with other studies: fast and high throughput (4–5 min per run; more than 100 runs per 8 h), high mass accuracy, improved sensitivity due to minimal AAV titer (~2E10 vg) and sample preparation requirements (10 min), and consistent reproducibility. Peptide mapping analysis of the AAV capsid proteins (wild-type and engineered capsids) was shown to achieve high sequence coverage and precise amino acid structure elucidation, including the identification of covalent modifications. With the rapid growth of novel AAV serotype discovery and development, we envision that this LC-MS method will be of great assistance

to many laboratories that require fast and accurate confirmation and characterization of AAV serotypes. Since the precise chemistry of the AAV capsids is still less understood, our MS/MS method described herein hopefully will contribute to accelerating the development of this knowledge. Depending on the need of a particular experiment, our protocols can be modified to accommodate specific analyses, as a mass spectrum contains myriad information about capsid mass, stoichiometry, modifications, reactivity to certain compounds, and more.

MATERIALS AND METHODS

Cell lines
The HEK293 cell line was purchased from ATCC for AAV vector production. The cells were cultured in DMEM supplemented with 10% fetal bovine serum, 100 μg/mL penicillin, and 100 units/mL streptomycin (Invitrogen) and maintained in a humidified 37°C incubator with 5% CO_2.

AAV vector production
All rAAVs were produced by triple-plasmid transfection[22] of HEK293 cells using three plasmids: pRep-Cap (pH22, pH25×, pH28, or pH29), pVector (pdsAAV-CB-Gluc or pdsAAV-CB-EGFP), and pAd-Helper
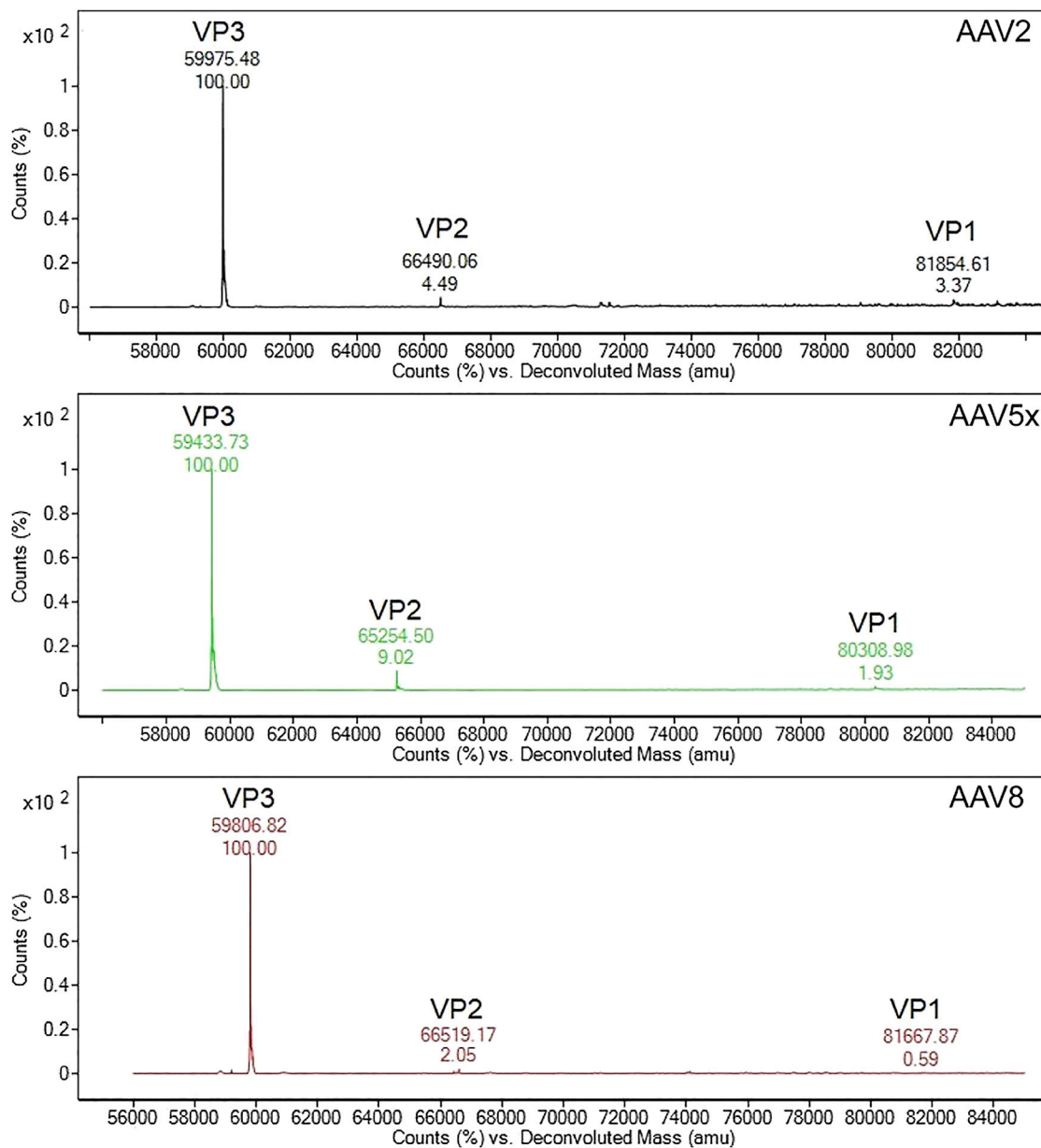
Figure 5. Intact LC-MS deconvoluted mass spectra of AAV2, AAV5x, and AAV8
Being the most abundant protein, VP3 was detected as 100% relative abundant peak in all serotypes.

(PfΔ6). The capsid of each rAAV may be chosen among natural sero-types (AAV2, AAV8, or AAV9) and may also be engineered by nonnatural methods such as mutagenesis (such as AAV5x) and chemical modification (AAV9-Dc and AAV5x-Dc). The three plasmids were delivered at the molar ratio of 1:1:1 into HEK293 cells (seeded in 2-L roller bottles) using PolyJet DNA In Vitro Transfection Reagent (SignaGen Laboratories). All rAAV vectors were collected after 72–96 h of transfection and were purified by LC (AKTA system; see supplemental information). The rAAV titers were determined by using ddPCR/qPCR with primers/probes targeting the transgene or pro-moter (see supplemental information for details).

SDS-PAGE
Mini-PROTEAN TGX stain-free precast gel (7.5%, a polyacrylamide gel containing a proprietary tri-halo compound to make proteins fluorescent directly in the gel, Bio-Rad, cat# 456-8026) was used. A sample of 6 μL rAAV (~1E12 vg/mL) was added to 2 μL 4× Laemmli sample buffer containing 10% β-mercaptoethanol and
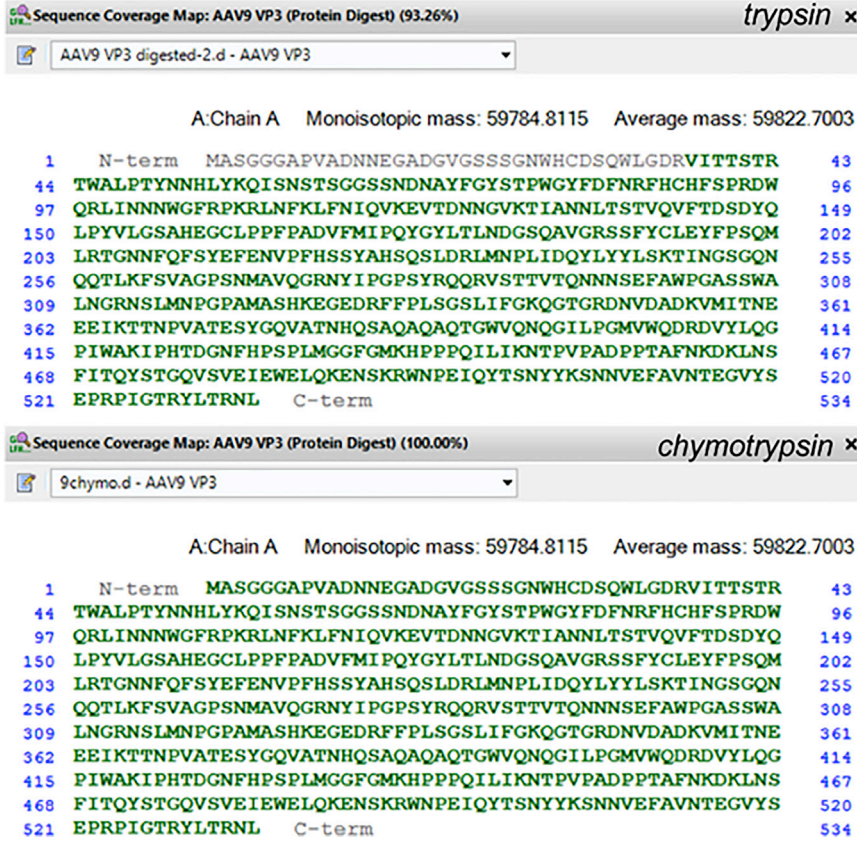
Figure 6. Screenshots of Agilent MassHunter BioConfirm 8 shows peptide map and sequence coverage of the AAV9 VP3 capsid protein

Top: trypsin digested; bottom: chymotrypsin digested. This protein has a 93.26% coverage using trypsin, and 100.00% coverage using chymotrypsin, as highlighted in green; the position of each amino acid of this protein is highlighted in blue.

was mixed well before being heated at 90°C for 5 min. The sample was cooled to room temperature and loaded into the gel lanes, together with a standard marker. Running buffer was 1× Tris/Glycine/SDS (Bio-Rad, cat# 1610772). The assembly was set and connected. Voltage was set to be constant, at 200 V. The gel was set to electrophorese for 30 min.

Instrumentation

All protocols documented were optimized from testing different parameters such as AAV titers, sample incubation in different acidic conditions of formic acid (FA) and TFA, times of sample incubation, varying the LC gradients for optimal elution of the AAV VP proteins at the desired of ∼3 min, column temperature, sample injection volume, and different serotypes, with reproducibility and robustness.

Agilent 1290 Infinity II LC: auto-sampler with 20-μL loop for both intact and digested analyses.

Agilent 6545 LC/QTOF: Dual-jet stream source. See Table 3 for LC setting details, and Table 4 for MS setting details.
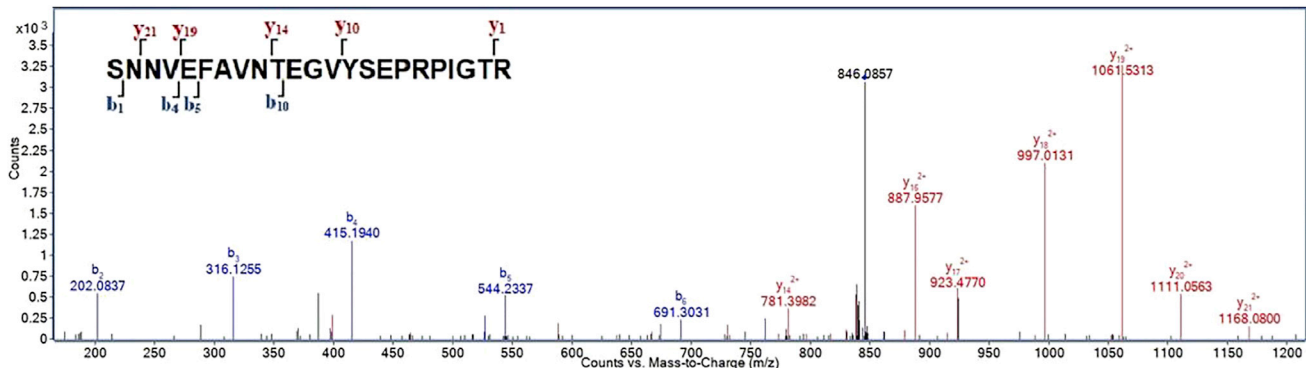

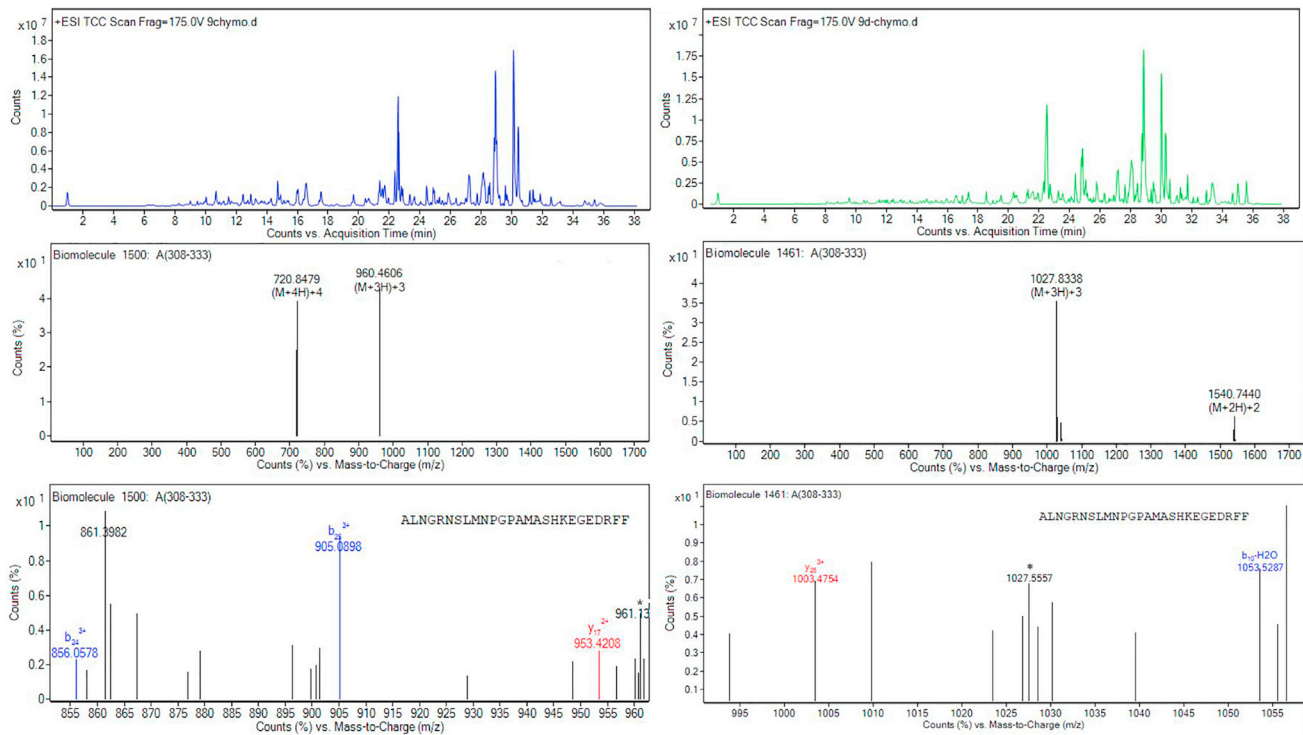**Figure 7. MS/MS spectrum of product ions of an AAV9 VP3 peptide sequence**

Figure 8. LC-MS/MS spectra of the digested capsid proteins of AAV9 (left) versus AAV9-Dc (right)

Peptide maps were elucidated with 100% VP3 sequence coverage for both serotypes. Top panel is total compound chromatogram (TCC); middle panel is the *m/z* of the ion amino acid position #308–333 detected; and bottom panel is the fragmented product ions of the one shown in the middle panel. The amino acid modified with one Dc molecule was identified as Lys at VP3#326, and the peptide sequence containing this position is shown here for both AAV9 (wild type) and AAV9-Dc (one Dc-molecule conjugation with the exact mass gain of 233.1 Da detected as shown in the middle and bottom panels of AAV9-Dc).

Sample preparation

For intact LC-MS analyses

1 μL FA (10% FA) was added to 9 μL AAVs (∼2E10 vg) and incubated at room temperature for 10 min before injection.

For peptide mapping LC-MS/MS

1E11 vg of AAVs was denatured in 6 M urea and 1 mM DTT at 90°C for 20 min and alkylated with iodoacetamide (15 mM) for 30 min at room temperature in the dark. Reduced and alkylated samples were cooled to room temperature and diluted with 3 equivalent volumes of buffer (50 mM Tris-HCl, 1 mM CaCl$_2$ [pH 7.5]) to ensure the urea concentration down to ≤2 M. The samples were then digested by 0.4 μg trypsin (Promega) or chymo-

trypsin (Thermo) overnight at 37°C. Digested samples were cleaned up using peptide-cleanup C18 spin columns (Agilent, cat# 5188-2750) right before LC-MS injection. A final 40-μL elution from each C18 cleanup column (following the manufacture's protocol) in elution buffer (70% ACN, 0.1% TFA) was collected and used directly for LC-MSMS analysis. The purpose of this step is to remove interfering contaminants and increase peptide concentration.

Data analyses

Agilent MassHunter BioConfirm B.08 software and Sequence Manager software were used for data analyses. The method report is attached as Data S1.

Table 5. Most detected post-translational modifications (PTMs) in AAV9 VP3 (chymotrypsin digested)

Sequence	PTM	Seq Mass diff (ppm)	Hits (BioConfirm)	Score (MFE)	Vol %
VLGSAHEGCLPPFPADVF	1*alkylation (iodoacetic acid) + 1*acetylation	6.29	9	100	0.35
VQNQGILPGMVW	1*deamidation + 1*acetylation	8.36	18	100	0.1
AHSQSLDRLMNPLIDQY	1*phosphorylation + 1*oxidation + 1*acetylation	1.31	47	100	0.05
RQQRVSTTVTQNNNSEF	1*deamidation + 1*acetylation	6.78	77	100	0.02
GQVATNHQSAQAQAQTGW	1*deamidation + 1*acetylation	4.72	30	100	0.01
AVNTEGVY	1*acetylation	0.47	2	97.1	0.01

DATA AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtm.2022.09.008.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

A.K.L. and W.X. designed the study. J.Z., D.F., and P.L.M. performed plasmid development and vector production for this study. A.K.L. performed purification. A.K.L. and L.Z. performed instrumentation of LC-MS/MS. All authors contributed to data analyses. A.K.L. wrote this manuscript with support and critical feedback from the others. All authors approved of the final version of the manuscript.

DECLARATION OF INTERESTS

A.K.L. and W.X. are filing a patent regarding the engineered AAV capsids studied in this work.

REFERENCES

1. Wang, D., Tai, P.W.L., and Gao, G. (2019). Adeno-associated virus vector as a platform for gene therapy delivery. Nat. Rev. Drug Discov. 18, 358–378. https://doi.org/10.1038/s41573-019-0012-9.

2. Kuzmin, D.A., Shutova, M.V., Johnston, N.R., Smith, O.P., Fedorin, V.V., Kukushkin, Y.S., van der Loo, J.C.M., and Johnstone, E.C. (2021). The clinical landscape for AAV gene therapies. Nat. Rev. Drug Discov. 20, 173–174. https://doi.org/10.1038/d41573-021-00017-7.

3. Lam, A.K., Frabutt, D., Li, L., and Xiao, W. (2021). Chemical modifications of the capsid for redirecting and improving the efficacy of adeno-associated virus vectors. Hum. Gene Ther. 32, 1433–1438. https://doi.org/10.1089/hum.2021.124.

4. Büning, H., and Srivastava, A. (2019). Capsid modifications for targeting and improving the efficacy of AAV vectors. Mol. Ther. Methods Clin. Dev. 12, 248–265. https://doi.org/10.1016/j.omtm.2019.01.008.

5. Yao, T., Zhou, X., Zhang, C., Yu, X., Tian, Z., Zhang, L., and Zhou, D. (2017). Site-specific PEGylated adeno-associated viruses with increased serum stability and reduced immunogenicity. Molecules 22, E1155. https://doi.org/10.3390/molecules22071155.

6. Pearce, H.A., Qian, H., Connell, T.U., Huang, D., Gottstein, C., Donnelly, P.S., Peter, K., Gregorevic, P., and Hagemeyer, C.E. (2019). Site-specific glycation and chemo-enzymatic antibody sortagging for the retargeting of rAAV6 to inflamed endothelium. Mol. Ther. Methods Clin. Dev. 14, 261–269. https://doi.org/10.1016/j.omtm.2019.07.003.

7. Human Gene Therapy for Rare Diseases - Guidance for Industry. (2020). FDA.gov.

8. Jin, X., Liu, L., Nass, S., O'Riordan, C., Pastor, E., and Zhang, X.K. (2017). Direct liquid chromatography/mass spectrometry analysis for complete characterization of recombinant adeno-associated virus capsid proteins. Hum. Gene Ther. Methods 28, 255–267.

9. Zhang, X., Jin, X., Liu, L., Zhang, Z., Koza, S., Yu, Y.Q., and Chen, W. (2021). Optimized reversed-phase liquid chromatography/mass spectrometry methods for intact protein analysis and peptide mapping of adeno-associated virus proteins. Hum. Gene Ther. 32, 1501–1511. https://doi.org/10.1089/hum.2021.046.

10. Zhang, X., Koza, S.M., Yu, Y.Q., and Chen, W. (2020). Optimizing adeno-associated virus (AAV) capsid protein analysis using UPLC and UPLC-MS. Waters Application Note.

11. Wang, Q., Dong, B., Pokiniewski, K.A., Firrman, J., Wu, Z., Chin, M.P.S., Chen, X., Liu, L., Xu, R., Diao, Y., and Xiao, W. (2017). Syngeneic AAV pseudo-particles potentiate gene transduction of AAV vectors. Mol. Ther. Methods Clin. Dev. 4, 149–158. https://doi.org/10.1016/j.omtm.2016.12.004.

12. DiMattia, M.A., Nam, H.-J., Van Vliet, K., Mitchell, M., Bennett, A., Gurda, B.L., McKenna, R., Olson, N.H., Sinkovits, R.S., Potter, M., et al. (2012). Structural insight into the unique properties of adeno-associated virus serotype 9. J. Virol. 86, 6947–6958.

13. Lee, G.K., Maheshri, N., Kaspar, B., and Schaffer, D.V. (2005). PEG conjugation moderately protects adeno-associated viral vectors against antibody neutralization. Biotechnol. Bioeng. 92, 24–34. https://doi.org/10.1002/bit.20562.

14. Kelemen, R.E., Mukherjee, R., Cao, X., Erickson, S.B., Zheng, Y., and Chatterjee, A. (2016). A precise chemical strategy to alter the receptor specificity of the adeno-associated virus. Angew Chem. Int. Ed. Engl. 55, 10645–10649. https://doi.org/10.1002/anie.201604067.

15. Katrekar, D., Moreno, A.M., Chen, G., Worlikar, A., and Mali, P. (2018). Oligonucleotide conjugated multi-functional adeno-associated viruses. Sci. Rep. 8, 3589. https://doi.org/10.1038/s41598-018-21742-x.

16. Reul, J., Muik, A., and Buchholz, C.J. (2019). Ligand coupling to the AAV capsid for cell-specific gene transfer. In Adeno-Associated Virus Vectors (Springer), pp. 35–50.

17. Marsic, D., Govindasamy, L., Currlin, S., Markusic, D.M., Tseng, Y.S., Herzog, R.W., Agbandje-McKenna, M., and Zolotukhin, S. (2014). Vector design Tour de Force: integrating combinatorial and rational approaches to derive novel adeno-associated virus variants. Mol. Ther. 22, 1900–1909. https://doi.org/10.1038/mt.2014.139.

18. Davidsson, M., Wang, G., Aldrin-Kirk, P., Cardoso, T., Nolbrant, S., Hartnor, M., Mudannayake, J., Parmar, M., and Björklund, T. (2019). A systematic capsid evolution approach performed in vivo for the design of AAV vectors with tailored properties and tropism. Proc. Natl. Acad. Sci. USA 116, 27053–27062. https://doi.org/10.1073/pnas.1910061116.

19. Tse, L.V., Klinc, K.A., Madigan, V.J., Castellanos Rivera, R.M., Wells, L.F., Havlik, L.P., Smith, J.K., Agbandje-McKenna, M., and Asokan, A. (2017). Structure-guided evolution of antigenically distinct adeno-associated virus variants for immune evasion. Proc. Natl. Acad. Sci. USA 114, E4812–E4821. https://doi.org/10.1073/pnas.1704766114.

20. Au, H.K.E., Isalan, M., and Mielcarek, M. (2021). Gene therapy advances: a meta-analysis of AAV usage in clinical settings. Front. Med. 8, 809118. https://doi.org/10.3389/fmed.2021.809118.

21. Kruzik, A., Fetahagic, D., Hartlieb, B., Dorn, S., Koppensteiner, H., Horling, F.M., Scheiflinger, F., Reipert, B.M., and de la Rosa, M. (2019). Prevalence of anti-adeno-associated virus immune responses in international cohorts of healthy donors. Mol. Ther. Methods Clin. Dev. 14, 126–133. https://doi.org/10.1016/j.omtm.2019.05.014.

22. Mulcrone, P.L., Zhang, J., Pride, P.M., Lam, A.K., Frabutt, D.A., Ball-Kell, S.M., Xiao, W., and Pan, D. (2022). Genomic designs of rAAVs contribute to pathological changes in the livers and spleens of mice. Advances in Cell and Gene Therapy 2022, 1–19. https://doi.org/10.1155/2022/6807904.