Characterization of adeno-associated virus capsid proteins with peptide mapping

Easy to use solution for capsid protein characterization using the SCIEX X500B QTOF LC-MS/MS System with SCIEX OS Software and BPV Flex Software

Sean McCarthy, Kerstin Pohl, Esme Candish
SCIEX, Framingham, MA

Adeno-associated Viruses (AAVs) have become an attractive therapeutic vehicle for gene therapy development. The focus on this particular class of therapeutic is partially related to their being non-pathogenic, their ability to treat both dividing and non-dividing cells, and the ability to afford long term gene expression without incorporating into the host cell genome[1]. A number of different AAV serotypes have been identified and studied for their potential to provide treatment for a wide range of diseases. The selection of serotype aligns well with the disease target and affected tissue to ensure the greatest efficacy.

It is well understood that the development and production of AAV based therapeutics is a complex process with many challenges. From an analytical perspective, assessment of the ratio of full and empty capsids, the overall ratio of the capsid proteins (VP1, VP2, and VP3), and the assessment of both desired and non-desired post translational modifications (PTMs) continue to be challenging.

The characterization of AAV capsid proteins in particular is challenging for a number of reasons. First, the capsid proteins VP1, VP2, and VP3 are present at a ratio of 1:1:10 which can make the characterization of the lower abundance proteins VP1 and VP2 difficult. In addition, while typical yields of AAV therapies may be sufficient for use therapeutically, the material required for many analytical approaches can be limiting, which can be problematic when considering the cost of AAV production.

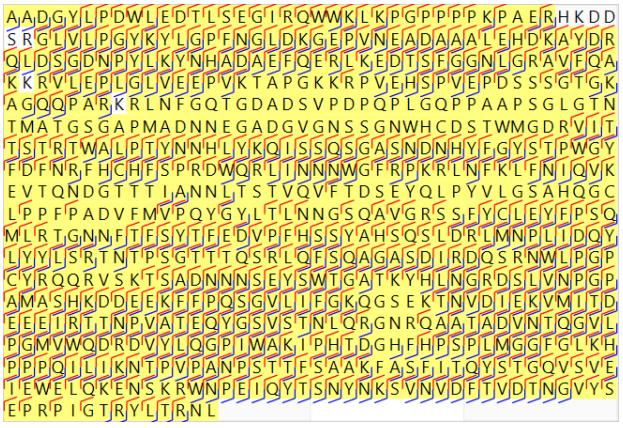

Figure 1. Sequence coverage map of AAV2 VP1 capsid protein. Highlighted sequence represents detected MS signal and red and blue lines denote observed MS/MS sequence confirmation. Overall sequence coverage is 98.9% from a tryptic digest.

SCIEX X500B QTOF System with SCIEX OS Software

Presented here is an approach for the characterization of AAV capsid proteins and associated PTMs using a bottom-up approach. The presented method limits sample loss from sample preparation using a simple one pot digestion strategy without the need for the use of dialysis or spin filters. Digested samples were analyzed using a SCIEX X500B QTOF System coupled to an ExionLC™ System which provided the required sensitivity to achieve nearly complete sequence coverage. Data acquisition was controlled using SCIEX OS Software and data processing using BPV Flex Software. Both software products have been designed for ease of use to enable both expert and non-expert users to deliver meaningful results to inform decision making.

Key features of the SCIEX solution for capsid characterization

- Ease of use and high quality MS and MS/MS data to ensure confidence in results for both expert and non-expert users
- Sensitivity to ensure detection of critical PTMs across all capsid proteins
- Intuitive and flexible interface of BPV Flex Software to facilitate data processing and curation

Methods

Sample preparation:

AAV-LacZ (AAV Serotype 2) was purchased from SignaGen Laboratories (PN SL100854) and used without any further purification. The sample contained 30 µL of total volume at a concentration of $2.25E^{13}$ gene copies/ml (GC/mL). This was converted into a total estimated protein concentration of 4.2 µg/ml. It is possible the actual protein concentration is greater, since the GC/mL value does not account for empty capsids. The digestion was executed within the vial provided by the vendor following the protocol included with the SCIEX Protein Preparation Kit. Digestion was accomplished using Trypsin/Lys-C mix from Promega (PN V5071) with incubation overnight at 37°C. The digestion was terminated by the addition of formic acid and the sample was analyzed directly.

Chromatography:

Chromatographic separation of the digested capsid samples was done using an ExionLC™ system. The column used was a Waters CSH C18, 2.1mm × 150mm.The column was set to a temperature of 40°C. Mobile phase A was water with 0.1% formic acid and mobile phase B was acetonitrile with 0.1% formic acid. The gradient conditions and flow rate used are shown in Table 1.

Table 1. LC Conditions for intact analysis.

Time (min	%A	%B	Flow Rate ml/min
Initial	*97*	*3*	*0.25*
0.5	*97*	*3*	*0.25*
50.0	*45*	*55*	*0.25*
50.1	*10*	*90*	*0.25*
55.0	*10*	*90*	*0.25*
55.1	*97*	*3*	*0.25*
75	*97*	*3*	*0.25*

Mass spectrometry:

Mass spectrometry detection was accomplished using a SCIEX X500B LC MS/MS System. The instrument was fitted with a Turbo V™ Ion Source with a TwinSprayer ESI probe. The instrument was controlled using SCIEX OS Software. Data were acquired using and IDA method with the TOF MS and IDA criteria as shown in Tables 2 and 3 respectively.

Table 2. MS parameters.

Parameter	Setting
Scan Mode	*Positive*
GS1	*50 psi*
GS2	*60 psi*
Curtain Gas	*30 psi*
Temperature	*550°C*
Ion Spray Voltage	*5500 V*
Time Bins to Sum	*6*
Accumulation Time	*0.25 sec*
TOF Start Mass (m/z)	*300*
TOF Stop Mass (m/z)	*2000*
Declustering Potential	*80 V*
Collision Energy	*10 V*

Table 3. IDA criteria.

Parameter	Setting
Maximum number of candidate ions	*15*
Intensity threshold	*100*
Dynamic Collision Energy	*True*
Exclude former candidate ions	*True*
Exclude former candidates for	*6 s*
Exclude candidates after	*2 occurrences*
Accumulation time	*0.06 sec*
Time bins to sum	*12*
TOF Start Mass (Da)	*100*
TOF Stop Mass (Da)	*1600*

Data processing:

Data was processing using BPV Flex Software 2.0. Processed data was manually curated and modifications were localized where possible to specific residues. VP1, VP2, and VP3 proteins were entered independently for processing.

Capsid protein sequence confirmation

As shown in Figure 1, nearly full sequence coverage was obtained for the capsid proteins as evidenced by MS and in many cases MS/MS confirmation. To confirm that each capsid protein was present, the N-terminal sequences of each of the capsid proteins was verified. By their nature, the entire sequence of VP3 is contained in VP2, and VP2 is contained in VP1. For this reason the C-terminal sequences of capsid proteins in AAVs are identical, however the N-terminal sequences of each protein following trypsin digestion are unique. As shown in Figure 2, the N-terminus of VP1 was identified and sequenced with 100% coverage by MS/MS. The N-terminus of VP1 was found to be acetylated consistent with previous reports. [2] Similarly, the N-terminal peptide of VP3 was also found to be acetylated as evidenced by the MS/MS data shown in Figure 3. The N-terminus of VP2 was also confirmed with MS and MS/MS data. In addition to confirmation of the N-terminal sequences, the C-terminal sequence which is common to VP1, VP2, and VP3 was confirmed. As shown in Figure 4, high sequence coverage was obtained which agrees well with previous reports. From these data it is possible to confirm the identity of each of the capsid proteins as well as confirm the presence of expected N-terminal post translational modifications on VP1 and VP3.. VP2 was not found to be modified agreeing with previous studies. [2]

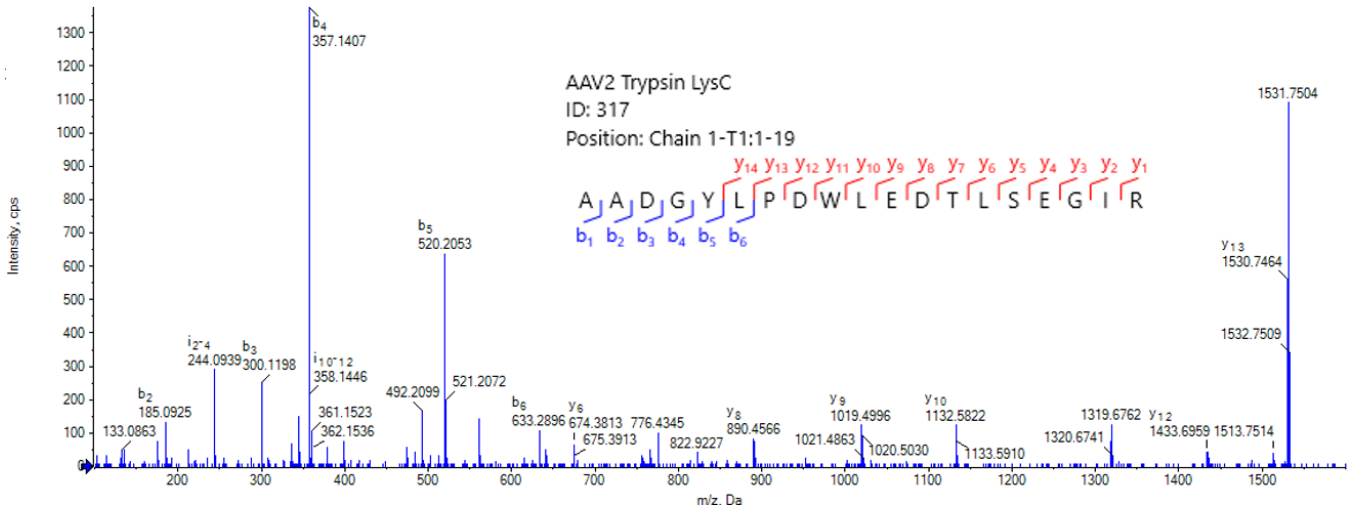

Figure 2. MS/MS sequence confirmation of N-terminal peptide of VP1. Comprehensive fragmentation of the peptide is observed supporting sequence confirmation and N-terminal acetylation.

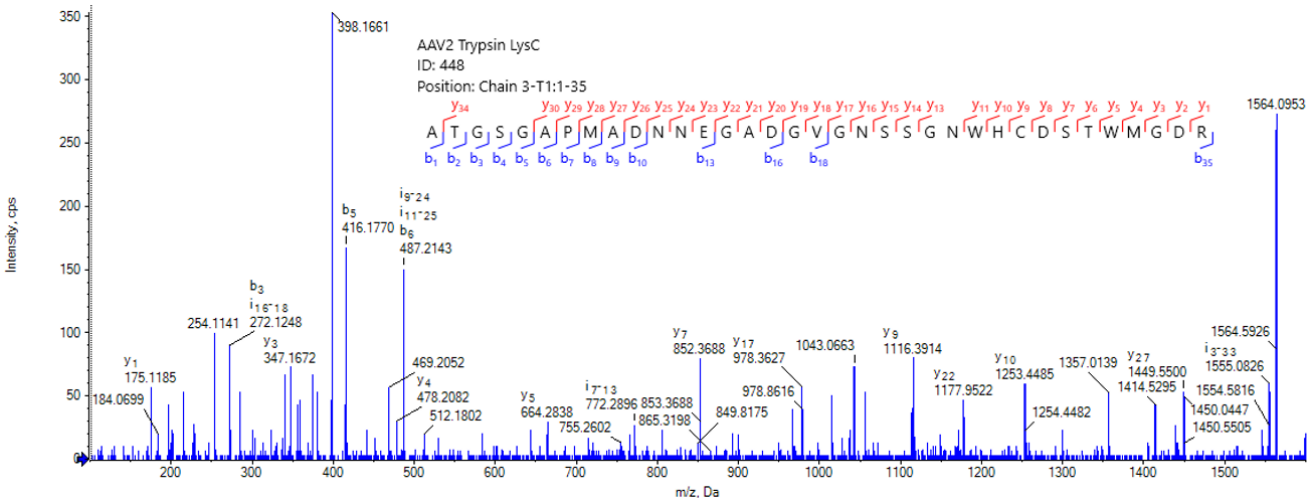

Figure 3. MS/MS sequence confirmation of N-terminal peptide of VP3. Nearly comprehensive fragmentation of the peptide is observed supporting sequence confirmation and N-terminal acetylation.

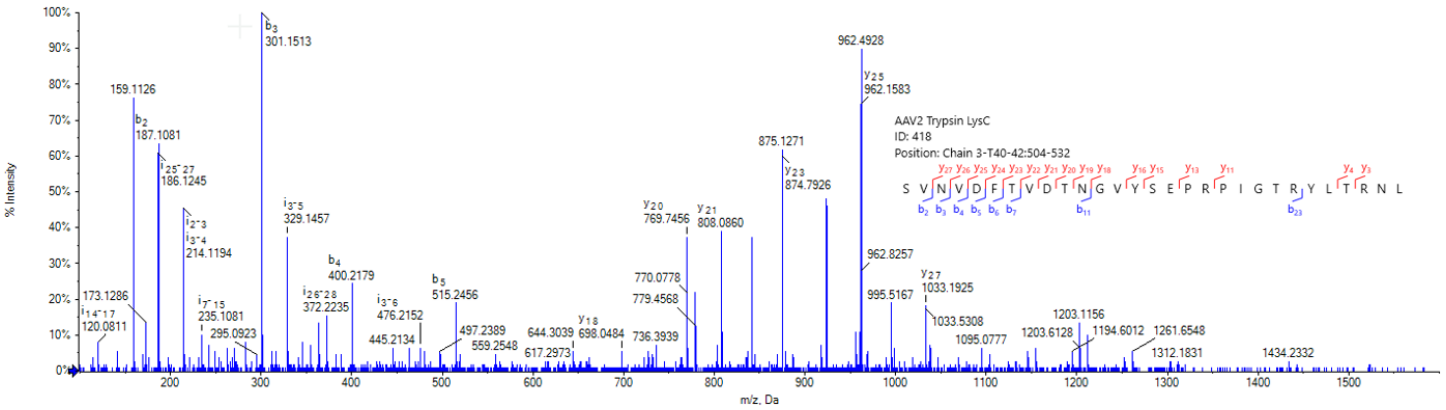

Figure 4. MS/MS sequence confirmation of C-terminal peptide common to VP1, VP2, and VP3. Descriptive fragmentation is observed supporting sequence confirmation and C-terminus.

Capsid protein post translational modifications

While the impact of post translational modifications on capsid proteins is still under investigation, early reports suggest that their presence or absence is attributable to altered vector function[2]. While the impact of specific modifications may be beneficial or deleterious, reports support that the presence of modifications are related to vector transduction and that their formation can be related to both the expression system used as well as the purification strategy used during production.

As part of this study, common post translational modifications such as oxidation and deamidation were included in the search criteria in BPV Flex Software. After processing, the data was reviewed in BPV Flex Software to verify the identification of the modifications. The position of modification is easily identified when reviewing the results using the interactive sequence viewer. Each identified peptide, and modified form if found, is designated by a green line. In addition, the location of the modification is indicated by a pink mark as shown in Figure 5.

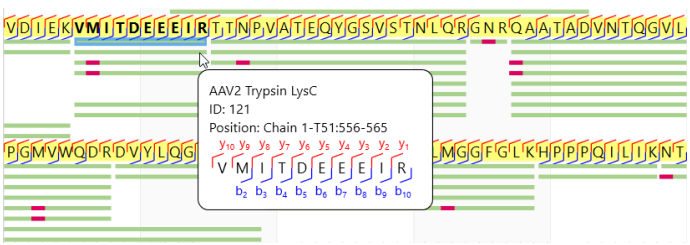

Figure 5. Interactive sequence viewer showing identified peptides and the location of modifications highlighted in red.

To more readily compare the modified and unmodified peptides, BPV Flex Software allows comparison of MS and MS/MS spectral data for multiple identified components simultaneously. A particularly useful application of this comparison is for the verification of PTMs as the observed fragment ion spectra should be distinctly different and potentially enable localization of modifications to specific residues.

As shown in Figure 6, there is a comparison of the peptide VMITDEEEIR in its unmodified and modified forms. This peptide is common across VP1, VP2, and VP3. As shown in Figure 5, this same peptide is identified in both its native and modified form, with suspected oxidation on the methionine residue. It is easy to confirm the presence of oxidation in the peptide by comparing the spectra shown in Figure 6. First the precursor mass indicates a clear shift which correlates to oxidation. In addition, the fragment ion series for both the singly charged b and y fragment ions are shifted by 16 Da according to the position of the oxidation of methionine. Finally, BPV Flex Software readily calculates the modification percent for each peptide with a positioned modification. In this case the extent of modification across all the capsid proteins was determined to be 3.6%.

Similarly, deamidation was identified on several residues within the sample. One example is shown in Figure 7 for a peptide unique to VP1, YLGPFNGLDK. Supporting the identification of deamidation is a corresponding shift in the observed precursor mass which is validated by MS/MS data that show a clear shift in mass for fragment ions which contain the deamidation. As with the oxidation discussed previously, the extent of modification is calculated automatically by BPV Flex Software. In this case the deamidation percent was found to be 33.6%.

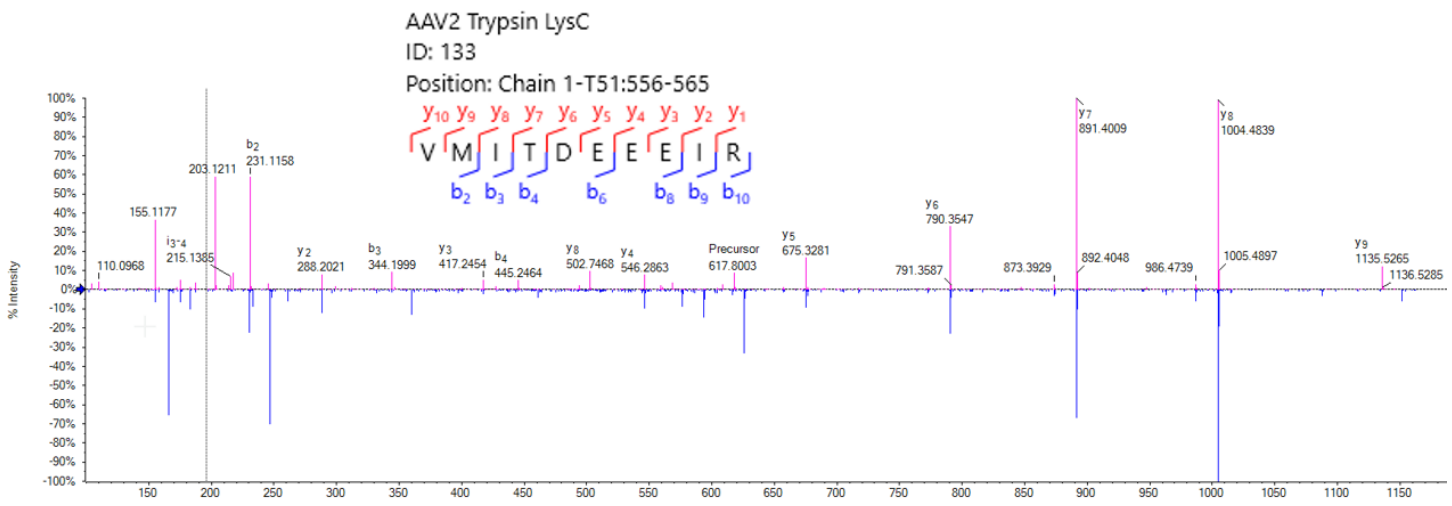

Figure 6. Comparison of unmodified (pink) and modified (blue) peptide VMITDEEEIR. Good correlation between fragment ions from y-ion series is observed with shift at y9 while b-ion series shows systematic shift of 16 Da due to oxidation of methionine **(M).**

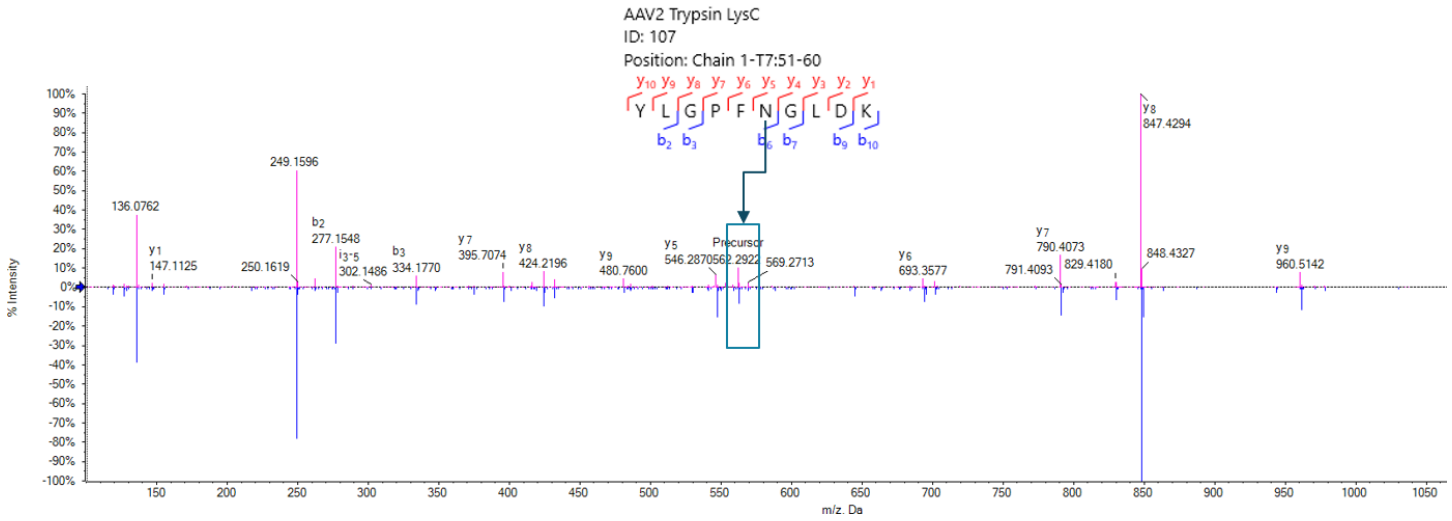

Figure 7. Comparison of unmodified (pink) and modified (blue) peptide YLGPFNGLDK from VP1. The observed mass shift for the precursor as well as fragment ions containing the deamidated residue (N) are shifted accordingly supporting the assignment.

With the sorting and filtering capabilities in BPV Flex Software it is easy to aggregate data for peptides to focus on the specific question of a study. As an example, to investigate the extent of deamidation across each of the sites in the capsid proteins, data can be grouped by peptide and then by modification percent. This results in a list of each peptide which is sub segmented based on each position which can potentially be modified. As all of the data for each peptide, and by extension modification position, is considered in its totality, the extent of modification at each position is easily understood as shown in Figure 8. While the example shown is for deamidation, other modifications or attributes can also be used for sorting and filtering as needed. As shown, the modification percent abundance position within the peptide, as well as the residue within the full protein are reported to expedite review and understanding of the results.

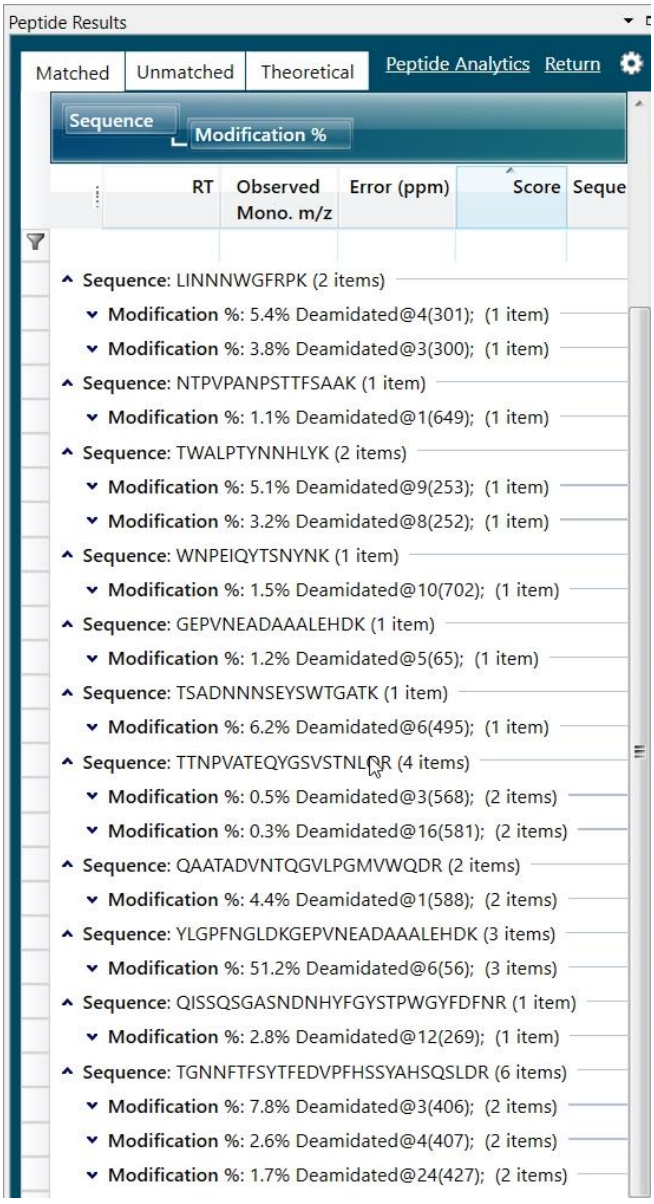

Figure 8. Use of BPV Flex Software grouping and sorting functionality to focus on attributes of interest.

Conclusions

- The SCIEX X500B QTOF System with BPV Flex software enables characterization of capsid proteins including confirmation of sequence of both C and N termini
- High quality MS/MS spectra enable unambiguous confirmation of peptide identification and localization of modifications
- Direct comparison of unmodified and modified peptides is easily done with BPV Flex Software using the comparative analysis tools including the ability to view spectra in mirror plots
- Identification and relative quantitation of PTMs is expedited using the flexible data processing, sorting and filtering capabilities of BPV Flex Software

References

1. Naso MF, Tomkowicz B, Perry WL 3rd, Strohl WR. Adeno-Associated Virus (AAV) as a Vector for Gene Therapy. BioDrugs. 2017;31(4):317–334. doi:10.1007/s40259-017-0234-5
2. Xiaoying Jin, Lin Liu, Shelley Nass, Catherine O'Riordan, Eric Pastor, and X. Kate Zhang.Human Gene Therapy Methods.Oct 2017.255-267.
3. Bertin Mary, Shubham Maurya, Sathyathithan Arumugam, Vikas Kumar, Giridhara R. Jayandharan. FEBS J. 2019 Dec;286(24):4964-4981

Headquarters
500 Old Connecticut Path | Framingham, MA 01701 USA
Phone 508-383-7700
sciex.com

International Sales
For our office locations please call the division headquarters or refer to our website at
sciex.com/offices