Characterization and Quantification of Adeno-associated Virus Fill States Using Analytical Ultracentrifugation

Nikki Machalek, Amber Raines and N. Karl Maluf; KBI Biopharma, Louisville, CO USA

Introduction

Recombinant adeno-associated viruses (AAV) are used extensively as a vector for gene therapy. AAV products are composed of a proteinaceous capsid that encapsulates the single stranded DNA genome. Preparations of purified AAVs typically contain species ranging from empty capsids, various partially-packaged capsids that contain less than the full complement of intended DNA, fully-packaged capsids which represent the drug product, along with viral/capsid aggregates and fragments. Characterization and quantification of these species is necessary to ensure safety and efficacy of gene therapy treatments.

Recently, Maruno et al.[1] published an elegant and comprehensive analysis using analytical ultracentrifugation to characterize AAV size distributions using multiple detection wavelengths. This approach was successful due to the improved absorbance optical detection capabilities of the Beckman Optima Analytical Ultracentrifuge, which allowed the authors to characterize and quantify the protein and nucleic components of various species observed in those size distributions. In this poster, we discuss implementation of these methods in our lab to provide a more comprehensive description of AAV size distributions for our clients. We show how this analysis can be used to identify empty, partially-packaged and full capsids, calculate the ratio of protein to DNA present in well resolved species, and compare the multi-wavelength results with expected values that arise from a consideration of the shape properties of AAVs. Furthermore, we show how this approach can help eliminate false peak identification. We also conducted a statistical analysis of the measurement precision and have established acceptance criteria (for this commercially available AAV-5 model system) for fundamental measures such as main peak sedimentation coefficient, peak percentages and capsid fill state.

Experimental

Commercial AAV serotype 5 was purchased from Virovek. The expected genome size for the product is 2.5 kb, while the wild type genome is 4.7 kb. The stock was provided at approximately 2×10^{13} vg/mL, and its concentration was determined using a Nanodrop OneC. For sedimentation velocity experiments, the stock was diluted to a target OD at 280 nm of about 0.25 using Gibco DPBS. The diluted sample was loaded into one chamber of a 2 sector, charcoal-Epon centerpiece (1.2 cm pathlength) while the reference chamber was loaded with DPBS. The experiments were conducted using a Beckman Optima Analytical Ultracentrifuge, at 20 °C, using a rotor speed of 16,000 RPM. Scans were simultaneously collected on each sample at five absorbance wavelengths: 230 nm, 255 nm, 260 nm, 270 nm and 280 nm.

Theory

Two methods for calculating the fill state of an AAV species were evaluated in this work. The first method relies on the fact that the proteinaceous AAV capsids and the ssDNA genome show different absorbance spectra; i.e. Capsids and ssDNA have differing extinction coefficients over a wide range of λ.

Consider a solution containing an AAV species. The total absorbance of the solution, A_T, at a particular wavelength, λ, is given by:

$$A_T = \varepsilon_D D_T l + \varepsilon_C C_T l \qquad [1]$$

where ε_D and ε_C are the extinction coefficients for DNA and Capsids, respectively (and are functions of λ), expressed in units of $(\text{mg/mL})^{-1}\text{cm}^{-1}$; D_T and C_T are the total concentrations of DNA and Capsids (expressed in units of mg/mL); and l is the pathlength (1.2 cm). Equation 1 can be linearized as follows:

$$\frac{A_T}{\varepsilon_D l} = DT + \frac{\varepsilon_C}{\varepsilon_D} C_T \qquad [2]$$

where A_T and $\varepsilon_D/\varepsilon_C$ are functions of λ. Thus, a plot of $A_T/(\varepsilon_D l)$ vs. $\varepsilon_C/\varepsilon_D$ will yield a straight line with a slope of C_T and an intercept of D_T. This implies that the relative concentrations of DNA and Capsid proteins can be measured for a species that is well resolved by sedimentation velocity analysis over a range of analysis wavelengths. Further, the DNA mass fraction can be calculated from the measured slope and intercept as $D_T/(D_T+C_T)$.

The second method follows from the insight that the frictional properties of empty, partial and full capsids are hydrodynamically indistinguishable. This implies that ratios of sedimentation coefficients of various capsid species will be independent of hydrodynamic friction, and thus can be determined from knowledge of just the buoyant molar mass properties of the species. To illustrate the method, consider the Svedberg equation for a particular species, i:

$$s_i = \frac{M_i(1 - \bar{\upsilon}_i \rho)}{N_A f_i} \qquad [3]$$

where M_i is the molar mass (g/mol) of species i, $\bar{\upsilon}_i$ is the partial specific volume (mL/g) of species i, ρ is the solution density (g/mL), N_A is Avogadro's number and f_i is the frictional coefficient of species i. Let EC represent empty capsids and let species i correspond to a particular fill state (i.e. a certain proportion of Capsid and DNA). Taking the ratio of species i to EC yields (after some algebra):

$$\frac{s_i}{s_{EC}} = \frac{M_{EC}(1 - \bar{\upsilon}_{EC}\rho) + M_D(1 - \bar{\upsilon}_D\rho)}{M_{EC}(1 - \bar{\upsilon}_{EC}\rho)} \qquad [4]$$

Notice that since f_i is independent of fill state, each species f_i cancels upon taking the ratio. Published values for M_{EC}, $\bar{\upsilon}_{EC}$, $\bar{\upsilon}_D$ are available[1], and s_i/s_{EC} can be measured experimentally. Thus, Equation 4 can be solved for M_D, and the DNA mass fraction can be calculated as $M_D/(M_D+M_{EC})$.

Results

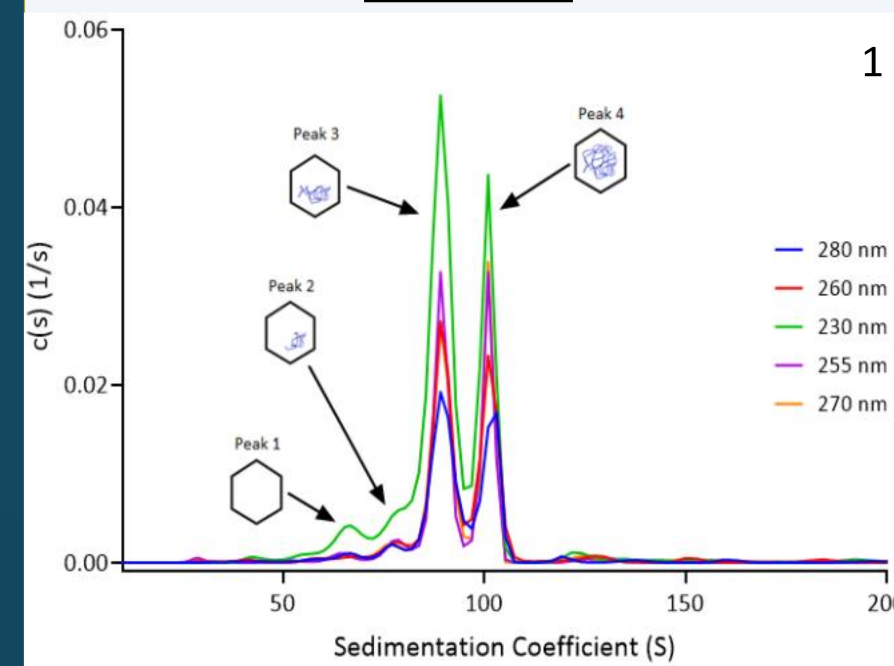

Figure 1 (above) shows the c(s) size distributions calculated for a single replicate analyzed at five analysis wavelengths. This sample shows a small fraction of empty capsids (at about 64 S), along with at least three additional filled species. Note that peak 3 (at about 90 S) corresponds to the most abundant species, but an additional species is observed that sediments faster (peak 4; at about 103 S). Note the relative proportions of the observed peaks vary with analysis wavelength due to differences in the extinction coefficients of the ssDNA and Capsid proteins. This sample also shows small amounts of fragments and viral aggregates.

While it is tempting to conclude peak 4 corresponds to the full AAV product, while peaks 2 and 3 correspond to partially packaged AAV, analysis of the DNA mass fraction of these data (discussed below) suggests a different interpretation.

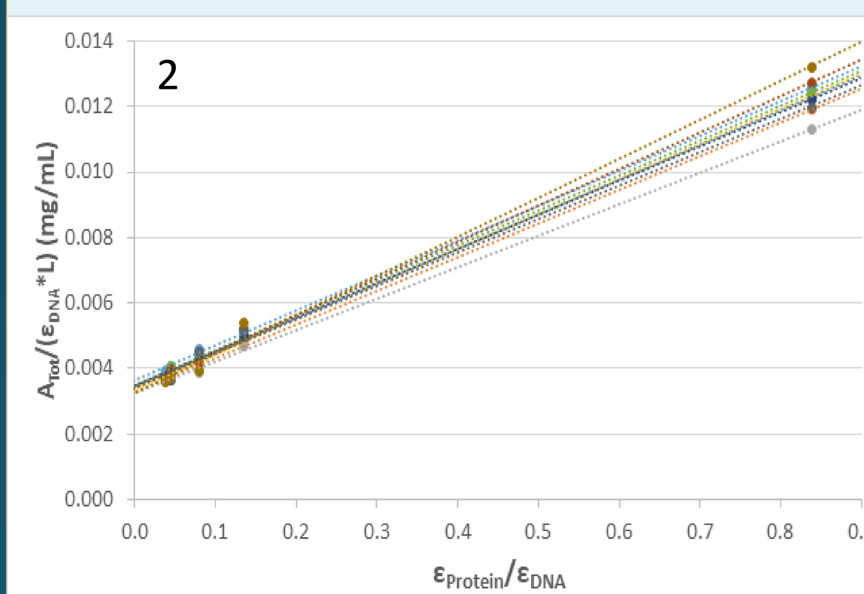

Figure 2 (above) shows the linear transformation of the integrated peak intensities (shown is peak 4) collected over 10 replicates. The data from each replicate were fitted to a linear function. From the slope and intercept, the average DNA mass fraction for this peak was calculated as 24.4%. Note that the expected mass fraction for an AAV with a 2.5 kb genome is 17.2%. Furthermore, from the sedimentation coefficient ratio analysis, the calculated mass fraction for peak 4 is 26.8%, which agrees reasonably well with the multi-wavelength analysis, and also indicates peak 4 does not correspond to the full virus (a single packaged 2.5 kbp genome). Rather, peak 4 likely represents an over packaged virus.

Table 1 (below) shows the calculated DNA mass fractions for the four peaks shown in Figure 1, calculated from the multi-wavelength approach and the sedimentation coefficient ratio approach. The two approaches agree reasonably well with each other, with the sedimentation coefficient ratio approach consistently estimating a slightly higher DNA mass fraction. Note that the measured DNA mass fraction suggests that peak 2 corresponds to a partially packaged virus (i.e. it contains less than the full genome), while peak 3 corresponds to the full virus (i.e. it contains the full, 2.5 kb genome). As previously mentioned, peak 4 likely corresponds to an over-packaged AAV.

Peak	Method	Mass fraction of DNA
Peak 4	Multi-wavelength	0.244
	Sedimentation coefficient ratio	0.268
Peak 3	Multi-wavelength	0.178
	Sedimentation coefficient ratio	0.195
Peak 2	Multi-wavelength	0.085
	Sedimentation coefficient ratio	0.110
Peak 1	Multi-wavelength	0.036
	Sedimentation coefficient ratio	0.012 *

Table 2 (below) shows the results of a REML statistical model of the measurement precision.

Parameter	Peak	Average value	Detection threshold, 3σ
Sedimentation coefficient, s* (S)	Peak 1	64.25	±2.22
	Peak 2	77.23	±5.70
	Peak 3	89.51	±1.17
	Peak 4	103.00	±4.53
Peak percentage (%)	Peak 1	5.8	±2.4
	Peak 2	7.8	±7.2
	Peak 3	50.4	±6.9
	Peak 4	30.0	±3.9
	DNA Containing	88.2	±3.6
	Total LMWS	1.2	±0.9
	Total HMWS	5.2	±4.2
A260/A280	Peak 1	0.96	±2.19
	Peak 2	1.44	±1.26
	Peak 3	1.23	±0.21
	Peak 4	1.30	±0.21
	DNA Containing	1.26	±0.09
Mass fraction of DNA using multi-wavelength method	Peak 1	0.244	±0.036
	Peak 2	0.178	±0.030
	Peak 3	0.085	[0, 0.350] *
	Peak 4	0.036	[0, 0.150] *

Takeaways

- Multiwavelength SV-AUC analysis can identify the fill state of AAV species.
- Fill state identification agrees well between multi-wavelength analysis and sedimentation coefficient ratio approach.
- Fastest sedimenting peak may not reflect full virus. Rather, it may reflect an over-packaged virus.

References

1) Maruno T., Usami K., Ishii K., Torisu T., and Uchiyama S. (2021) Comprehensive Size Distribution and Composition Analysis of Adeno-Associated Virus Vector by Multiwavelength Sedimentation Velocity Analytical Ultracentrifugation. Journal of Pharmaceutical Sciences, 110:3375-3384.

Copyright © 2014 Alliance Protein Laboratories, Inc.