BECKMAN COULTER
Life Sciences

Demystifying DGE-AUC Part 4: Fundamentals of Data Analysis

Akash Bhattacharya
Beckman Coulter Life Sciences, 4510 Byrd Dr, Loveland, CO 80538

Abstract

This application note is the fourth in a technical series covering the basics of density gradient equilibrium analytical ultracentrifugation or DGE-AUC. In it, we will cover the basics of DGE-AUC data analysis. A data analysis case study on Adeno Associated Virus (AAV) samples[1-4] using the Origin software package will be described.

Introduction

Background

In the previous installments of this series, we discussed the fundamentals of Analytical Ultracentrifugation (AUC) and its physical basis. We compared different AUC experimental techniques and touched upon a few common use cases – such as the quantification of loading fraction of gene therapy products such as AAV[5,6]. We then explored the physics behind DGE-AUC, its data acquisition, visualization and optimization. Then we discussed the workflow of a DGE-AUC experiment, paying special attention to the process of sample preparation. This workflow is shown in Figure 1.

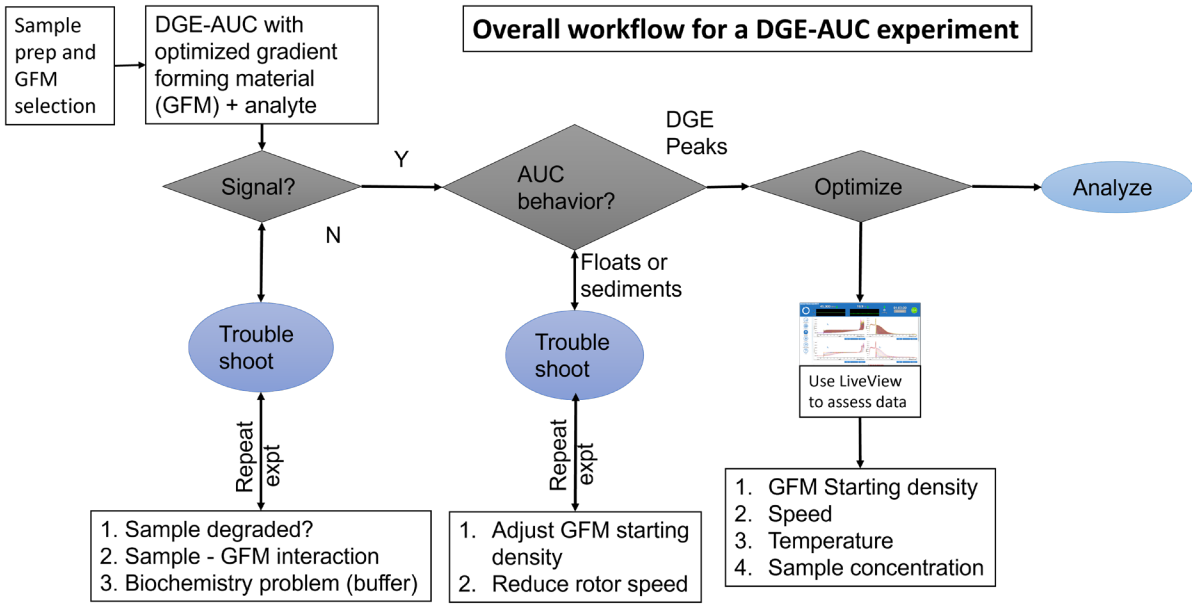

Figure 1: Workflow for a DGE-AUC experiment

Analysis and interpretation of DGE-AUC data:

The DGE-AUC analysis toolbox

One of the attractive features of DGE-AUC is that the data is easy and intuitive to interpret. Plotting the raw data enables quick visual interpretation. Quantitative analysis of the data yields much more detailed insight into the population distributions of different species. Such analysis can be done using common spreadsheet software. More complex operations such as peak integration can also be done using spreadsheets but are certainly easier using dedicated graphing software.

Finally, the most complex data analysis operations, which are used for challenging datasets, can be performed using full-featured programming languages and/or numerical computing environments. Thus, DGE-AUC analysis can be done using very basic and commonly used software tools, and the use of more sophisticated data science packages provides significant benefits.

Some of the tools available for DGE-AUC are shown in Table 1.

Analysis Type	What can you do here?	Examples	Software category	Scripting	Paid/Free
Essentials	1. Data import and visualization (both Abs and RIF) 2. Truncate data 3. Sector adjustment 4. Add and average scans 5. Peak integration	MS Excel	Spreadsheet	No	Paid
Optional	Essentials, and also 1. Baseline adjustment 2. Noise filtering	Origin, SigmaPlot, GraphPad, JMP, LabPlot, Grace, Gnuplot, SciDavis, etc.	Graphing software	Sometimes	Paid / Free
Advanced	Essentials, optional and also 1. MultiGauss fitting 2. Fully customizable noise filtering 3. Batch analysis and report generation	Python, R, C, C++, MATLAB, Octave	Full-featured programming language	Yes	Free (except MATLAB)

Table 1: Comparison of different software packages for DGE-AUC data analysis

Morphology of experimental DGE-AUC absorbance data for an analyte

Before we discuss DGE-AUC data analysis, it is desirable to understand the features of a DGE-AUC absorbance scan. In Figure 2, features of a DGE-AUC absorbance scan recorded at 230 nm for a sample of AAV9 in 1.35 g/mL CsCl in PBS are shown.

The air gap region is boxed in cyan, the meniscus region in orange, the density gradient region, which will be used for analysis, is boxed in grey, while the cell bottom region is boxed in maroon.

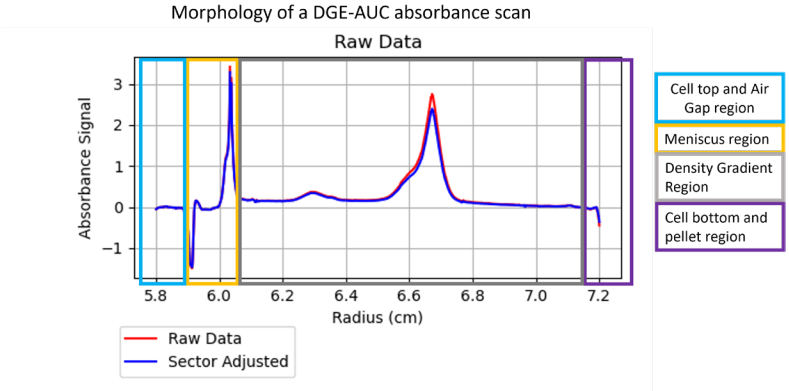

Figure 2: Morphology of an absorbance scan dataset for DGE-AUC of AAV in CsCl

Only the region boxed in grey – i.e., the density gradient region, will be used for data analysis. Thus, following the import of raw data from the Optima AUC instrument, the first step of actual analysis is to plot and visualize the data. Having done so, it is necessary to truncate the data to only the region of interest – which is "below" the menisci and "above" the cell bottom. The flowchart in Figure 3 shows the steps involved in processing a DGE-AUC dataset.

Workflow for analysis of DGE-AUC absorbance data of analytes.

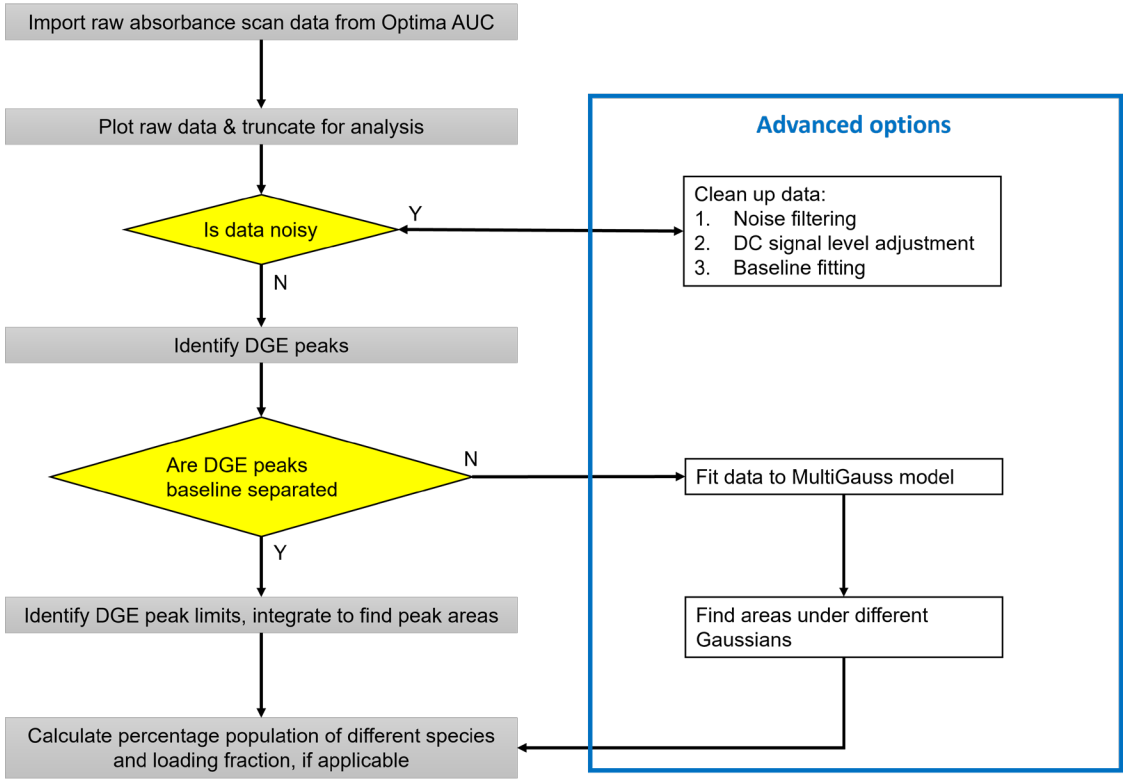

Figure 3: Workflow for analysis of DGE-AUC absorbance data of analytes

These steps are split into the essential steps on the left and the optional or advanced steps in the blue box on the right in Figure 3. The following steps are required for analysis of DGE-AUC absorbance data of analytes.

1. Download raw data from the Optima AUC instrument.

2. If analysis will be performed using MS Excel or a similar spreadsheet program, then export equilibrium absorbance scans to a spreadsheet. This will typically be the last 12 scans of a 200-scan experiment – corresponding to the last hour of a 16-hour experiment.
   or
   If analysis will be performed using Origin or a similar graphing package, import raw Optima AUC data into the software. This can be done using Import ASCII or similar functions.

3. Plot and visualize the absorbance scans.

4. Visual inspection: Determine whether the analyte is at thermodynamic equilibrium based on the above plot. If equilibrium is established, then select the last scan.

5. Visual inspection: Identify menisci and cell bottom region.

6. Data transform: Truncate data to retain only the density gradient region. This step is optional. The menisci usually have high positive/negative values of absorbance. Therefore, including the menisci will effectively compress the Y-axis and make visual inspection of small DGE-peaks in the raw data potentially difficult. However, data transforms and peak analysis can be performed while retaining the full data range.

7. Visual inspection: Verify if the data is noisy? Can DGE peaks be identified?

8. Identify DGE-AUC peaks.

9. Visual inspection: Are DGE-AUC peaks baseline separated? If yes, then identify left and right radial limits of different DGE-AUC peaks.

10. Data transform: Use Peak Wizard (for Origin) or similar tool to smooth baseline and replot.

11. Data transform: Use Peak Wizard (for Origin) or similar tool to select DGE-peak centers and integration windows. Be sure to not select menisci in lieu of DGE peaks.

12. Data transform: Use Peak Wizard (for Origin) or similar tool to integrate peaks. Record or export results to a spreadsheet for report.

DGE-AUC data analysis example with Origin

Sample Information

The following analysis steps were performed using data acquired for a sample of AAV9 in CsCl based buffer. The sample conditions were as follows: AAV9 in 1.35 g/mL CsCl with PBS buffer at pH 7.4. DGE-AUC experiment with multiple speed steps: 7, 14, 21, 28, 35, 42 krpm rotor speeds, 100 scans per stage, absorbance data at 260 nm (data was recorded for 230, 260 and 280 nm), 5 min intervals, 4°C. The last scan (# 100) from stage 6 (42 krpm) was used for analysis in Origin.

Step 1: Data Import

Origin allows raw data to be imported using the menu below

Main Menu > Data > Import From File > Multiple ASCII

following the usual conventions about column separators (space), etc. Upon import, the third data column will be filled with zeroes, as it shows the absorbance of the reference sector (which is set to 0). This column can be dropped. There will be ~ 1,400 rows of data with the radius incrementing by 0.001 cm in each row starting from 5.8 cm to a maximum of 7.2 cm. The data table looks like this:

Radius	Absorbance Signal
cm	AU 280 nm
5.801	-0.08805
5.8019	-0.06321
5.803	-0.04374
5.8041	-0.03164
5.805	-0.02865
5.806	-0.02118
5.8071	-0.01662
5.8079	-0.01907
5.8091	-0.01753
5.8101	-0.01372
5.811	-0.01153
5.812	-0.00884
5.813	-0.01047

Table 2: Optima AUC raw absorbance data imported into Origin

The raw data can be plotted as shown below. There are two peaks, one between 6.2 cm and 6.4 cm and the other between 6.5 and 6.8 cm. The menisci are located at radii less than 6.1 cm.

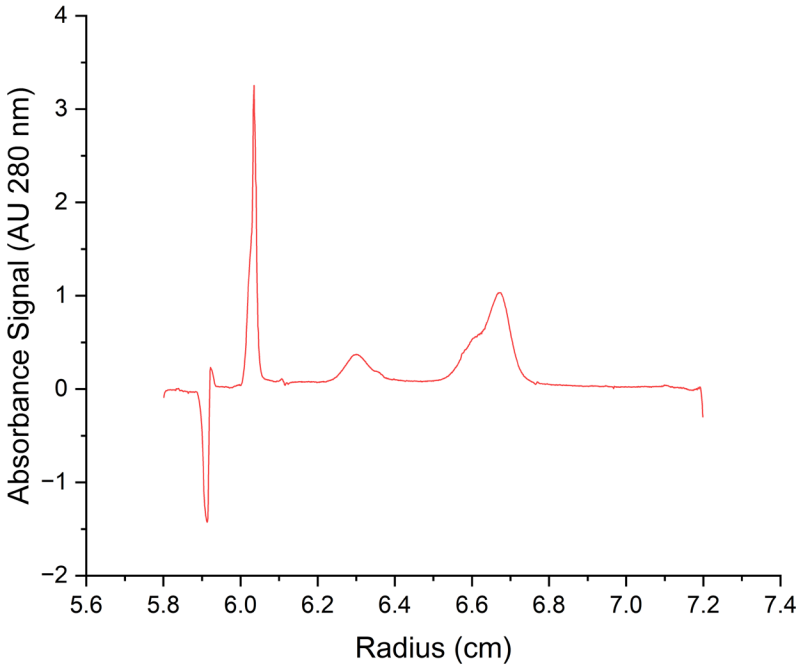

Figure 4: Raw AUC absorbance data for AAV9 in CsCl buffer

Step 2: Select baseline

Origin will automatically find points to adjust the baseline. There is also an option which allows the user to select new points or modify existing points. There are several options to calculate a baseline using polynomial and other functions. It is recommended to pick enough points to capture the entire baseline, but care must be taken to avoid actual DGE peaks. A minimum of six baseline points is recommended. The points picked are shown below in red in Figure 5.

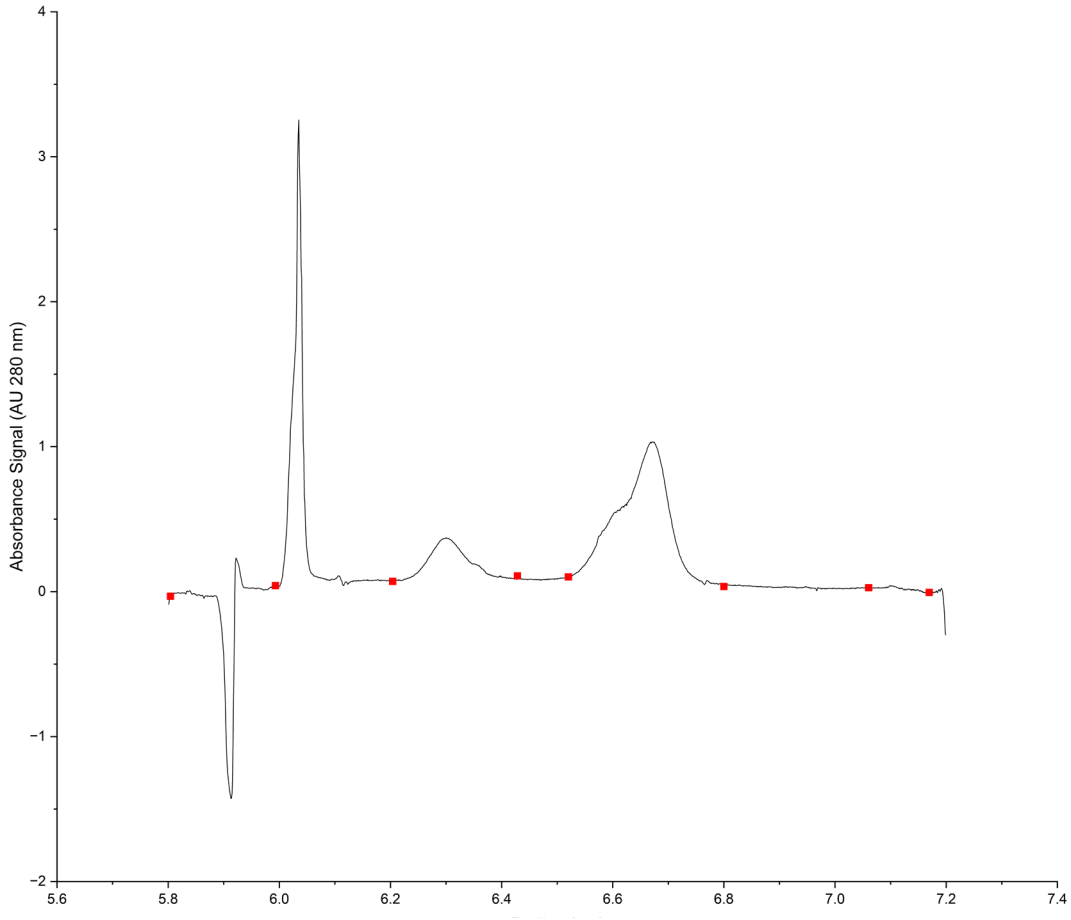

Figure 5: Baseline points on AUC absorbance data for AAV9 in CsCl buffer

Step 3: Calculate baseline

After selecting suitable points and picking a baseline calculation method, Origin will calculate a baseline and save it to the data table. This can be plotted as shown below. It must be ensured that the baseline does not incorporate DGE peaks or the menisci regions. If the baseline is distorted near these regions (peaks or menisci), this baseline should be rejected, and a new one should be calculated. Using different baseline selection points may help with finding a better baseline. Calculating a baseline by interpolation instead of fitting might also yield better results. The baseline is plotted below in Figure 6:

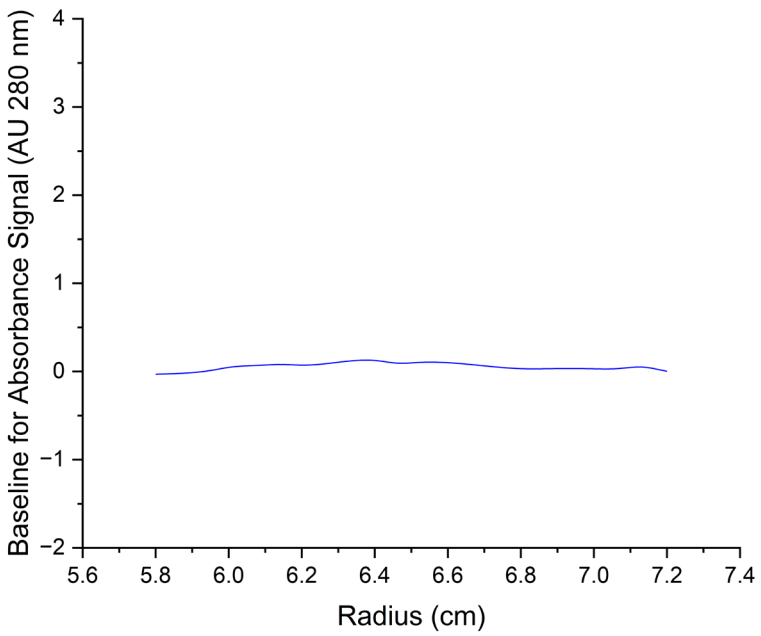

Figure 6: Baseline for AUC absorbance data for AAV9 in CsCl buffer

Step 4: Subtract baseline from raw data

Origin saves the raw data after baseline subtraction in the data table. This is plotted below in Figure 7:

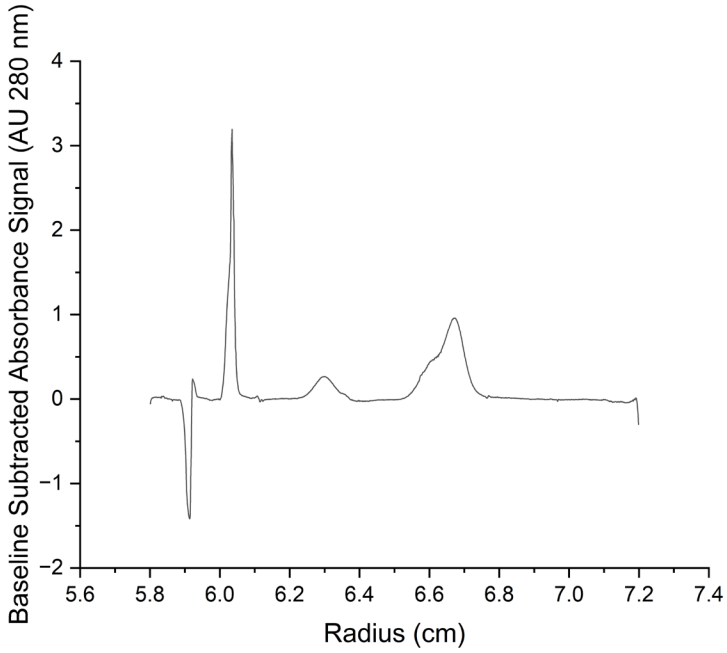

Figure 7: Adjusted AUC absorbance data after baseline subtraction for AAV9 in CsCl buffer

The adjusted (i.e., baseline subtracted data) should display a flat "ground level" of Y = 0. The DGE peaks and menisci should not be significantly distorted as compared to the raw data. At this juncture, it is useful to plot overlays of the raw data, the baseline and the adjusted data after baseline subtraction. This is shown below in Figure 8:

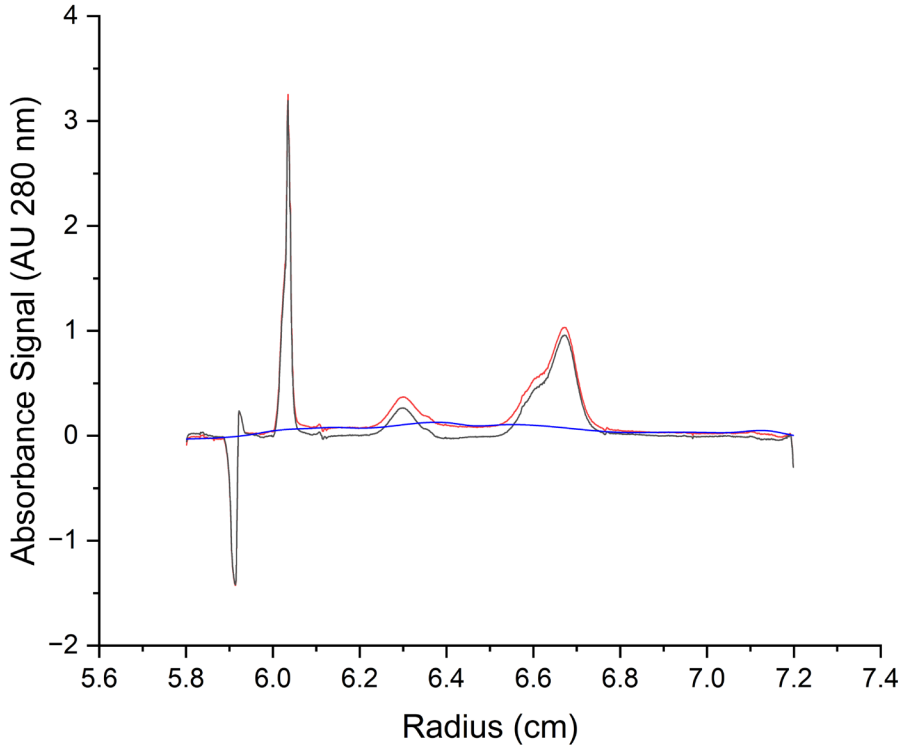

Figure 8: Raw AUC absorbance data (red trace), baseline (blue trace), adjusted data after baseline subtraction (black trace) for AAV9 in CsCl buffer

This is usually the final stage for visual inspection of data before peak analysis. If any inconsistencies are seen—such as a baseline which "captures" actual DGE peaks—then the baseline subtraction procedure must be repeated with different parameters.

Step 5: Peak Selection

The Peak Analysis wizard in Origin will automatically select peaks for integration. The center positions of the selected peaks can be moved by the user. Additional peaks can be added manually. Origin will also automatically suggest integration windows for all the selected peaks. These can also be adjusted manually. In this example, Origin selected several peaks, which included the menisci. Auto-selected peaks at the menisci were deleted. The two remaining peak markers in the plot below correspond to the two DGE peaks which were identified previously in Step 1 of this workflow.

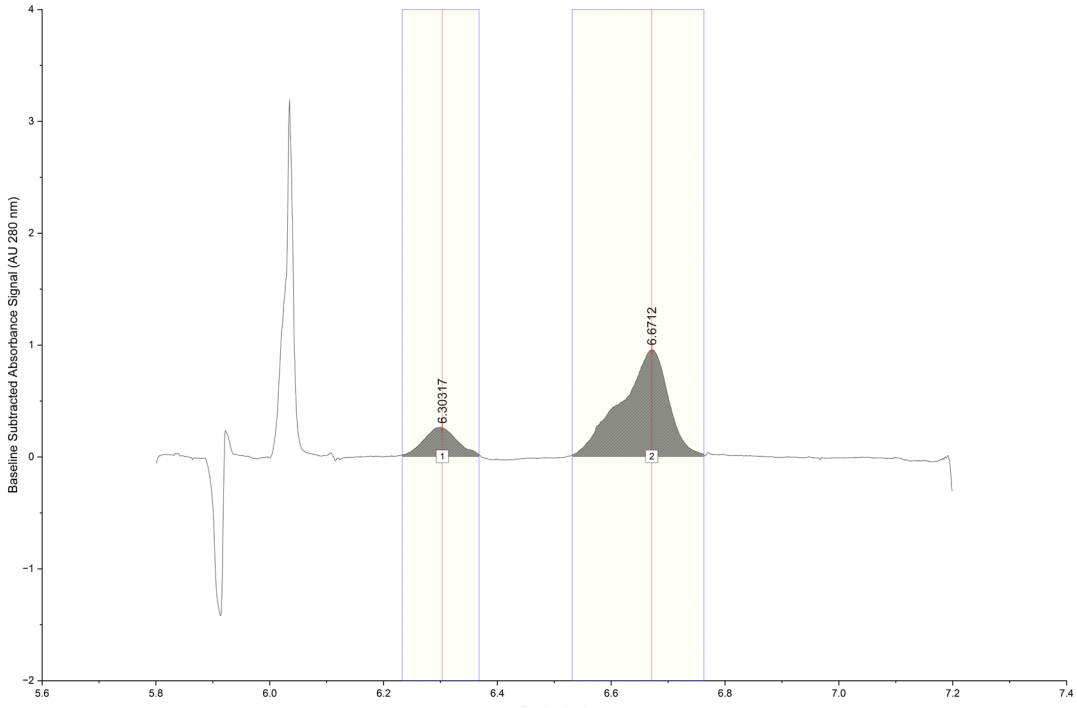

Figure 9: DGE peaks and integration windows for AAV9 in CsCl buffer

Step 6: Peak Integration

Index	Area	AreaIntgP (%)	Curve Area	Row Index	Beginning X	Ending X	FWHM	Left Half Width	Right Half Width	Center	Height	Centroid
1	0.01851	9.04432	0.20465	500	6.233	6.3681	0.06521	0.03331	0.0319	6.3	0.26674	6.30072
2	0.0933	45.59032	0.20465	874	6.532	6.7622	0.08813	0.05942	0.02871	6.6742	0.95823	6.65134

Table 3: Integration output from Origin Peak Wizard for AAV9 in CsCl buffer

The area-under-curve corresponding to each peak can be extracted from this table. Summing the total area under each curve allows us to extract the percentage area corresponding to each peak, and therefore each species. This is shown below:

Peak #	Area	FWHM	Center	Height	Summed Area	% Area
1 (left)	1.85E-02	6.52E-02	6.30E+00	2.67E-01	1.12E-01	16.55
2 (right)	9.33E-02	8.81E-02	6.67E+00	9.58E-01		83.45

Table 4: Integration output from Origin peak analysis wizard for AAV9 in CsCl buffer - peak percentage area calculation

Discussion and Conclusions

This is the fourth installment of a technical series covering the basics of DGE-AUC. In this application note, we discussed the different aspects of DGE-AUC data analysis, the software tools which are recommended for analysis, and the analysis workflow. We have demonstrated the use of Origin – a popular graphing and analysis software package to analyze a dataset acquired on a sample of AAV9 in PBS buffer with a CsCl gradient. This example demonstrates one of the core strengths of DGE-AUC –

the data can be easily interpreted like a chromatogram. The data is indeed WYSIWYG (what you see is what you get). While it is possible to perform sophisticated data transformation procedures to extract the valuable insights from challenging datasets (with high levels of noise, for instance), the basic data analysis is very simple and can be carried out using easy-to-use software tools.

More sophisticated analysis methods can be used, employing more sophisticated tools ranging from numerical computing packages like MATLAB/Octave all the way to full-featured programming languages like Python, C, C++, FORTRAN, R, etc. These tools offer almost infinite flexibility in analysis as well as report generation. This will be discussed in a subsequent installment of this series.

References

1. Balakrishnan, B. & Jayandharan, G. R. Basic biology of adeno-associated virus (AAV) vectors used in gene therapy. Curr Gene Ther 14, 86-100 (2014). **https://doi.org/10.2174/1566523214666140302193709**

2. Grimm, D. & Büning, H. Small But Increasingly Mighty: Latest Advances in AAV Vector Research, Design, and Evolution. Hum Gene Ther 28, 1075-1086 (2017). **https://doi.org/10.1089/hum.2017.172**

3. Naso, M. F., Tomkowicz, B., Perry, W. L., 3rd & Strohl, W. R. Adeno-Associated Virus (AAV) as a Vector for Gene Therapy. BioDrugs 31, 317-334 (2017). **https://doi.org/10.1007/s40259-017-0234-5**

4. Pupo, A. et al. AAV vectors: The Rubik's cube of human gene therapy. Mol Ther 30, 3515-3541 (2022). **https://doi.org/10.1016/j.ymthe.2022.09.015**

5. Bepperling, A. & Best, J. Comparison of three AUC techniques for the determination of the loading status and capsid titer of AAVs. Eur Biophys J 52, 401-413 (2023). **https://doi.org/10.1007/s00249-023-01661-0**

6. Richter, K. et al. Purity and DNA content of AAV capsids assessed by analytical ultracentrifugation and orthogonal biophysical techniques. Eur J Pharm Biopharm 189, 68-83 (2023). **https://doi.org/10.1016/j.ejpb.2023.05.011**

Legal Notices

2024-GBL-EN-106822-v1

danaher