**Submission ID #: 61578**
**Scriptwriter Name: Bridget Colvin**
**Project Page Link:** https://www.jove.com/account/file-uploader?src=18781338

# Title: Analysis of SEC-SAXS data via EFA deconvolution and Scatter

**Authors and Affiliations: Mark D. Tully[1,*], Nicolas Tarbouriech[2], Robert P. Rambo[3], and Stephanie Hutin[4,*]**
*These authors contributed equally to the work

[1]European Synchrotron Radiation Facility, Structural Biology Group
[2]Institut de Biologie Structurale, University Grenoble Alpes, CEA, CNRS
[3]Diamond Light Source
[4]Laboratoire de Physiologie Cellulaire and Végétale, Université Grenoble Alpes/CNRS/CEA/INRA/BIG

**Corresponding Author:**
Stephanie Hutin
stephanie.hutin@cea.fr

**Co-authors:**
mark.tully@esrf.fr
nicolas.tarbouriech@ibs.fr
robert.rambo@diamond.ac.uk

 

# Author Questionnaire

**1. Microscopy**: Does your protocol demonstrate the use of a dissecting or stereomicroscope for performing a complex dissection, microinjection technique, or similar? **N**

**2. Software:** Does the part of your protocol being filmed demonstrate software usage? **Y**

**3. Filming location:** Will the filming need to take place in multiple locations (greater than walking distance)? **N**

**Protocol Length**
Number of Shots: **39**

**0.1 wide shot of the synchrotron ESRF beamlines**

# Introduction

**1. Introductory Interview Statements**

**REQUIRED:**

1.1. **Stephanie Hutin**: Bio-Small Angle X-ray Scattering provides a structural measurement of macromolecules or complexes under relevant biological conditions. Ideally, the sample to be measured should be monodisperse **[1]**.

1.1.1. INTERVIEW: Named talent says the statement above in an interview-style shot, looking slightly off-camera

**REQUIRED:**

1.2. **Mark D. Tully**: Although in some cases Size-Exclusion-Chromatography-SAXS is not sufficient to produce monodispersity, a software-based deconvolution step of the obtained SAXS data can be performed to obtain an idealized SAXS curve **[1]**.

1.2.1. INTERVIEW: Named talent says the statement above in an interview-style shot, looking slightly off-camera

**OPTIONAL:**

1.3. **Nicolas Tarbouriech**: Following this protocol, a deconvolution program and a user-friendly Scatter program can be used to analyze the Vaccinia DNA polymerase exo minus mutant **[1]**.

1.3.1. INTERVIEW: Named talent says the statement above in an interview-style shot, looking slightly off-camera

# Protocol

## 2. Primary Data Analysis

2.1. To perform a background subtraction of the size exclusion chromatography data, open the Java-based **Scatter four** program **[1]** and open the **SEC** (seck) tab **[2]**.

2.1.1. WIDE: Talent opening program and SEC tab, with monitor visible in frame

~~2.1.2. SCREEN: To be provided by Authors: SEC tab being opened~~

2.2. Drag and drop the reduced data files into the **Drop Data below** window **[1]** and click ~~the~~ **Output Directory** ~~button~~ and **Open** to set the output directory in which the data will be saved **[2]**.

2.2.1. WIDE: File(s) being dragged and dropped From Authors: probably 3.1.1. from the videographer, there was a mix up

2.2.2. SCREEN: 2_2_2: 00:01-00:10

2.3. Enter the sample name in the **Save as** box and click **TRACE [1]**.

2.3.1. SCREEN: 2_3_1: 00:05-00:14

2.4. Click the **Edit Details** button to edit the experimental details and fill out the appropriate fields ~~as possible~~ **[1]**.

2.4.1. SCREEN: 2_4_1: 00:02-00:30 *Video Editor: please speed up*

2.5. To select the buffer frames manually, click **Set Buffers [1]**.

2.5.1. SCREEN: 2_5_1: 00:01-00:04

2.6. Left click and drag to reselect the buffer as a flat on the trace curve before the void volume of the size exclusion chromatography column of approximately 100 frames and click **Set buffer** and **Update** to recalculate the **.sec* file **[1]**.

2.6.1. SCREEN: 2_6_1: 00:01-00:19 *Video Editor: please speed up*

2.7. Left click and drag to select a region of interest on the signal plot and left click and drag the crosshairs in the heat map plot to select and zoom the subset of frames that will be used for merging **[1]**.

2.7.1. SCREEN: 2_7_1: 00:01-00:17 *Video Editor: please speed up*

2.8. With another left click, highlight the frames in the heat map corresponding to the area to the bottom right of the crosshairs. Ideally, the frames should highlight a region with a predominately cyan color and a stable radius of gyration **[1]**.

2.8.1. SCREEN: 2_8_1

2.9. When satisfied with the selected frames, click **Merge** to merge the subtracted frames and click the **Analysis** tab to view the data **[1]**.

2.9.1. SCREEN: 2_9_1: 00:01-00:22 *Video Editor: please speed up*

**3. Data Deconvolution**

3.1. For deconvolution of the data, load the data set in the deconvolution program **[1]**

3.1.1. SCREEN: 3_1_1: 00:03-00:16 *Video Editor: please speed up*

3.2. In the **Control Panel** under **Files**, use the folder symbol to locate the data and highlight all of the **.dat* files **[1]**.

3.2.1. SCREEN: 3_2_1: 00:02-00:06

3.3. Click **Plot Series**. A plot of integrated intensity versus frame number will be drawn in the **Series Plot [1]**.

3.3.1. SCREEN: 3_3_1: 00:03-00:12 *Video Editor: please speed up*

3.4. Under the **Series** tab, click to highlight the curve and to open the **LC Analysis** pop-up window **[1]**.

3.4.1. SCREEN: 3_4_1: 00:05-00:22 *Video Editor: please speed up*

3.5. To select a suitable buffer region, click Add region to select an area before the peak of the chromatrogram and one after the solvent front and click **Set buffer**. Pop-up windows will indicate that frames have not been selected for the background **[1-TXT]**.

3.5.1. SCREEN: 3_5_1: 00:08-00:57 *Video Editor: please speed up* **TEXT: Alternative: If stable background, click Auto to select buffer region**

3.6. To start the EFA (E-F-A), ~~select and~~ right-click on the highlighted file and select **EFA** from the menu **[1-TXT]**.

3.6.1. SCREEN: 3_6_1: 00:02-00:15 **TEXT: EFA: evolving factor analysis**

3.7. In the pop-up window, the single value decomposition of the data set should be observed. In the **Controls** box, check the **Use Frames** box values to confirm ~~so~~ that the whole peak area to deconvolute is covered in the intensity plot **[1]**.

3.7.1. SCREEN: 3_7_1: 00:02-00:12 *Video Editor: please emphasize graph on left of screen with "single value decomposition of the data set"*

3.8. The **Singular Values** plot will show the intensity of the singular values above the baseline **[1]**.

3.8.1. SCREEN: 3_8_1 *Video Editor: please emphasize Singular Values plot*

3.9. If the left and right singular vectors do not match ~~as demonstrated~~, change the **Signifcant Singular Vector** number to 2 and move the frames until the left and right singular vectors are similar **[1]**.

3.9.1. SCREEN: 3_9_1: 00:05-00:26 *Video Editor: please speed up*

3.10. The EFA will be calculated, generating plots in the forward and backward directions for each vector and indicating when the components start and exit the solution profile for the selected size exclusion chromatography-SAXS (sacks) data **[1-TXT]**.

3.10.1. SCREEN: 3_10_1: 00:02 *Video Editor: please emphasize forward and backward plots when mentioned and emphasize data on left of popup window with "indicating … data"* **TEXT: SAXS: small angle x-ray scattering**

3.11. RAW will attempt to identify the ranges. If necessary, use the arrows to change the ranges so that each circle is at the start of an inflection point rising from or falling to baseline and click **Next [1]**.

3.11.1. SCREEN: 3_11_: 00:04-00:32 *Video Editor: please speed up*

3.12. To reduce or eliminate spikes, identify approximately which frame corresponds to the spike **[1]**, using the **Range Control** arrows to adjust the **Component Range Controls [2]**.

 

3.12.1. SCREEN: 3_12_1: 00:03-01:35 *Video Editor: please speed up*

~~3.12.2. SCREEN: To be provided by Authors: Component containing frame being identified/spikes decreasing~~

3.13. When a minimum chi-square has been achieved, click **Back** to perform a validation check. If the original EFA plots still look valid, click **Next** and **Save EFA Data** to save the plots. Then click **Done** to close the EFA window **[1]**.

3.13.1. SCREEN: 3_13_1: 00:05-00:36 *Video Editor: please speed up*

3.14. Then, in the RAW window, open **Profiles** in the plot panel to view the curves. In the **Profiles** tab of the control panel, right-click to save the curves as **.dat* files **[1-TXT]**.

3.14.1. SCREEN: 3_14_1: 00:04-00:26 *Video Editor: please speed up* **TEXT: Additional deconvolution and EFA BioXTAS RAW at https://bioxtas-raw.readthedocs.io/en/latest/**

## 4. Small Angle X-Ray Scattering (SAXS) Property Assessment

4.1. For SAXS determination, open the Scatter ANALYSIS tab **[1-TXT]** and click **G** to select the manual Guinier analysis tool **[2]**.

4.1.1. WIDE: Talent opening tab, with monitor visible in frame **TEXT: See Bioisis.net for in-depth SAXS determination tutorial**

4.1.2. Talent pressing G button

4.2. In the plot, add or remove points such that the residuals do not have a “smile” or “frown” feature. The selected data in the Guinier fit should not exceed the maximum *q* multiplied by radius of gyration limit of 1.3 **[1]**.

4.2.1. SCREEN: 4_2_1: 00:09-00:30 *Video Editor: please speed up*

4.3. Click **Normalized Kratky**. The resulting plot provides an assessment of the structural state of the macromolecule, globular, cylindrical, disordered, normalized for mass and concentration **[1]**.

4.3.1. SCREEN: 4_3_1: 00:07-00:16

4.4. Click **Volume-of-correlation**. The total scattered intensity and an integrated area of the total scattered intensity as a function of *q* plots will appear as a quick reference for validating the quality of the scattering curve **[1]**.

4.4.1. SCREEN: 4_4_1: 00:04-00:11 *Video Editor: please emphasize total scattered intensity and integrated area of total scattered intensity as a function of q plots when mentioned*

4.5. To start the flexibility analysis, click **Flexibility**. Each panel in the pop-up window will show a plot exploiting a power-law relationship that exists between the compact and elongated-flexible biopolymers. Use the slider at the bottom of the box to change the view of the data until a plateau in one of the plots is reached **[1]**.

4.5.1. SCREEN: 4_5_1: 00:05-00:35 *Video Editor: please speed up*

4.6. Immediately after performing the flexibility analysis, click **Volume**. A pop up of three graphs will be generated. The Porod-Debye plot ~~remembers~~ tracks where the slider from the flexibility plot was left and shows the plateaued area data **[1]**.

4.6.1. SCREEN: 4_6_1: 00:05-00:16 *Video Editor: please emphasize Porod-Debye plot when mentioned*

4.7. To calculate the volume of the particle, move the start and endpoints until the blue line on the plot fits the plateaued region. For an unbiased result, the residuals in the Porod-Debye exponent power-law fit should show no pattern **[1]**.

4.7.1. SCREEN: 4_7_1: 00:08-00:40 *Video Editor: please emphasize Porod Exponent Power Law Fit with "residuals … pattern"*

4.8. Under the **P of r** tab, the real-space distribution and the scattering curve for the sample can be observed. Ideally, the distribution curve will be smooth with no waves ~~present~~ and should just gently touch the x-axis **[1]**.

4.8.1. SCREEN: 4_8_1: 00:10-00:20 *Video Editor: please emphasize real-space distribution and scattering curve plots when mentioned, then emphasize distribution curve and contact with x-axis*

4.9. ~~Select the sample,~~ Right click on the sample name and click **Find DMAX** (D-max) to open a new window. The limits for the maximum dimension are pre-set with the suggested ~~$q_{max}$~~ maximum q-range, the maximal data points to be used for the calculation, the lower and upper maximum dimension limits and a lower and upper alpha score **[1]**.

4.9.1. SCREEN: 4_9_1: 00:04-00:18 *Video Editor: please speed up and indicate limits shown in pop-up window when mentioned*

4.10. Click **Start**. A composite distribution will be created, along with the suggested maximum dimension and alpha level. If these data are acceptable, close the window ~~and~~ to return to the P of r tab where the reciprocal space plot will now be cropped to match the suggested ~~$q_{max}$~~ maximum q-range **[1]**.

4.10.1. SCREEN: 4_10_1: 00:06-00:25 *Video Editor: please emphasize dmax and alpha level when mentioned*

4.11. Select the **Moore** model and click **Background** to set the alpha level and maximum dimension to the suggested values from the pop-up box **[1]**.

4.11.1. SCREEN: 4_11_1: 00:04-00:20 *Video Editor: please speed up*

4.12. Click **Refine**. A cross-validation plot will open, showing whether any points had to be rejected as indicated in red. If there are only a few rejected points and the distribution looks good, then the model is good **[1]**.

4.12.1. SCREEN: 4_12_1: 00:04-00:13 *Video Editor: please emphasize red points when mentioned as necessary*

4.13. To print a report, in the **Analysis** tab, left click to highlight the sample and right click on the sample name to select **Create report from single data set**. A text box will open to allow comments and a PDF document will be produced showing all of the figures and values generated **[1]**.

4.13.1. SCREEN: 4_13_1: 00:02-00:34 *Video Editor: please speed up*

# Protocol Script Questions

**A.** Which steps from the protocol are the most important for viewers to see?
n/a

**B.** What is the single most difficult aspect of this procedure and what do you do to ensure success?
2.8., 3.10.

# Results

**5. Results: Representative EFA Deconvolution and Scatter**

5.1. In this representative analysis, E9 DNA polymerase exonuclease minus mutant was bound to DNA and run using size exclusion chromatography-small angle X-ray scattering **[1]**. Two peaks were observed **[2]**.

5.1.1. LAB MEDIA: Figure 1
5.1.2. LAB MEDIA: Figure 1 *Video Editor: please emphasize peaks in blue data line*

5.2. The first, large peak represents the E9 DNA polymerase exonuclease minus mutant-DNA complex **[1]** and the second indicates the unbound state **[2]**.

5.2.1. LAB MEDIA: Figure 1 *Video Editor: please emphasize first peak*
5.2.2. LAB MEDIA: Figure 1 *Video Editor: please emphasize second peak*

5.3. While the classical approach of selecting frames provides a stable radius of gyration of the complex in the first peak **[1]**, the second peak is clearly merged and the radius of gyration across the plot shows that the second peak of interest does not have a stable radius of gyration due to cross-peak contamination **[2]**.

5.3.1. LAB MEDIA: Supplementary Figure 1 *Video Editor: please emphasize red line over peak 1*
5.3.2. LAB MEDIA: Supplementary Figure 1 *Video Editor: please emphasize red line over peak 2*
5.3.3. LAB MEDIA: Supplementary Figure 1 *Video Editor: please emphasize cyan region in bottom heat plot*

5.4. In this analysis, only 5 frames could be used that showed a semi-stable radius of gyration **[1]**. When subtracted, they gave a radius of gyration of 36.3 angstroms **[2]**.

5.4.1. LAB MEDIA: Figure 1 *Video Editor: please emphasize cyan region in bottom heat plot*
5.4.2. LAB MEDIA: Figure 2 *Video Editor: please emphasize green data*

5.5. When the peaks were deconvoluted using EFA, the corresponding curve for the second peak was overlaid with the original **[1]**, showing a clear decrease in signal to the noise and a lower radius of gyration of 34.1 angstroms **[2]**.

5.5.1. LAB MEDIA: Figure 2 *Video Editor: please emphasize blue data*

5.5.2. LAB MEDIA: Figure 2 *Video Editor: please emphasize green data*

5.6. A Kratky plot of the data reveals that the complex with the deconvoluted peak is more globular **[1]** as confirmed by the P-r curve, which gives a maximum dimension of 108.5 angstroms for the deconvoluted curve **[2]**.

5.6.1. LAB MEDIA: Figure 3 *Video Editor: please emphasize blue data*
5.6.2. LAB MEDIA: Figure 4 *Video Editor: please emphasize blue data line*

5.7. The non-deconvoluted data for this analysis is more elongated, with a maximum dimension of 120 angstroms **[1]**, most likely due to the heterogeneity arising from the unbound E9 polymerase minus exonuclease mutant **[2]**.

5.7.1. LAB MEDIA: Figure 4 *Video Editor: please emphasize green data line*
5.7.2. LAB MEDIA: Figure 4

# Conclusion

**6. Conclusion Interview Statements**

6.1. **Mark D. Tully**: The most critical steps are selecting the number of the singular values and the range of data used, as these values will greatly affect the accuracy of the deconvolution **[1]**.

6.1.1. INTERVIEW: Named talent says the statement above in an interview-style shot, looking slightly off-camera

6.2. **Nicolas Tarbouriech**: The results should not be taken on their own, but further assessed using additional techniques, such as analytical ultracentrifugation or multi-angle-laser-light-scattering, to allow their biological interpretation **[1]**.

6.2.1. INTERVIEW: Named talent says the statement above in an interview-style shot, looking slightly off-camera

6.3. **Stephanie Hutin**: Inline column coupled SAXS in combination with deconvolution and the user-friendly interface of Scatter provides a powerful package to obtain meaningful structural data even out of intrinsically difficult systems **[1]**.

6.3.1. INTERVIEW: Named talent says the statement above in an interview-style shot, looking slightly off-camera