Tutorial 1: Basic SAXS data analysis with RAW

Introduction:

This tutorial covers basic SAXS data processing. While the details are program specific, the general techniques should transfer between programs. You will use BioXTAS RAW and learn to:

- Process images into scattering profiles
- Average, subtract and save scattering profiles
- Find R_g and I(0) by Guinier analysis
- Find molecular weight by four different methods
- Do Kratky analysis
- Find a P(r) function

You will also learn how to perform basic data validation on measured scattering profiles. You will learn to identify:

- Concentration effects (inter-particle interaction)
- Aggregation
- Radiation damage

Requirements:

1. BioXTAS RAW, version 2.1.0 (newest).
    - Install instructions are available from: https://bioxtas-raw.readthedocs.io/en/latest/install.html
    - Install instructions as well as the RAW program are also available in the course materials.
2. ATSAS programs, version >=3.0.1.
    - Download and install instructions are available from: http://www.embl-hamburg.de/biosaxs/download.html
    - Requires a free registration for academic users. Industrial users must pay to use.
3. Tutorial data.

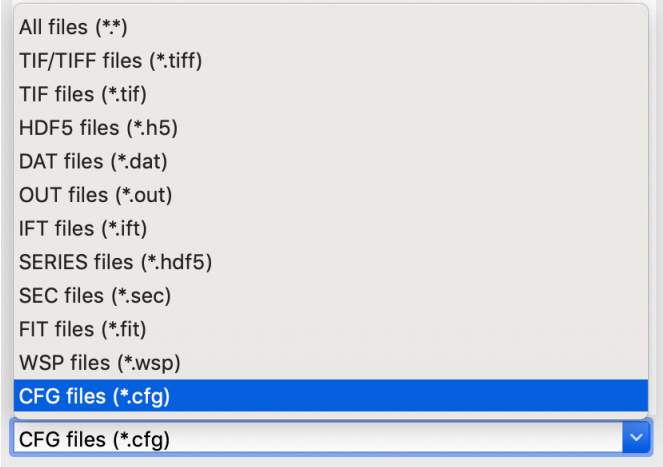
Other useful materials:

1. There are RAW tutorial videos, which can be viewed here: https://bioxtas-raw.readthedocs.io/en/latest/videos.html
2. ATSAS
    - Manuals: http://www.embl-hamburg.de/biosaxs/manuals/
    - User forum: http://www.saxier.org/forum/

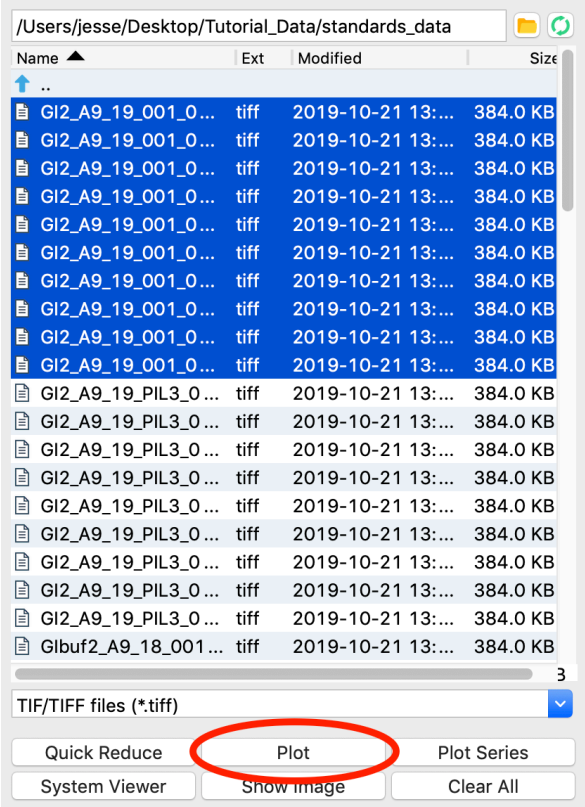
Part 1. Loading configuration files and images, creating subtracted scattering profiles, saving profiles

1. Open RAW. The install instructions contain information on installing and running RAW.
2. In the files tab, click on the folder button and navigate to the **Example_Data/standards_data** folder.

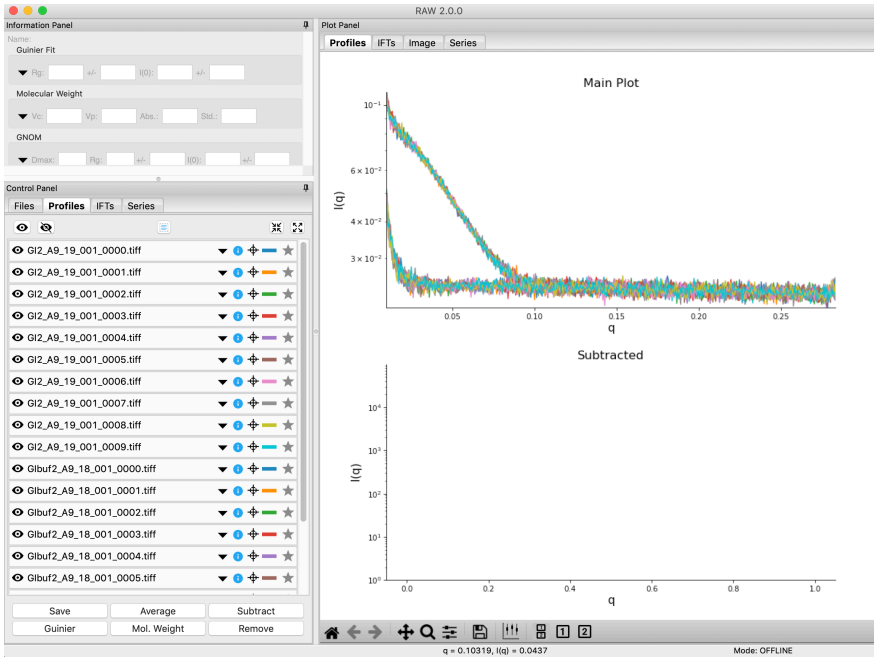
3. Set the file type filter to "CFG files (*.cfg)".



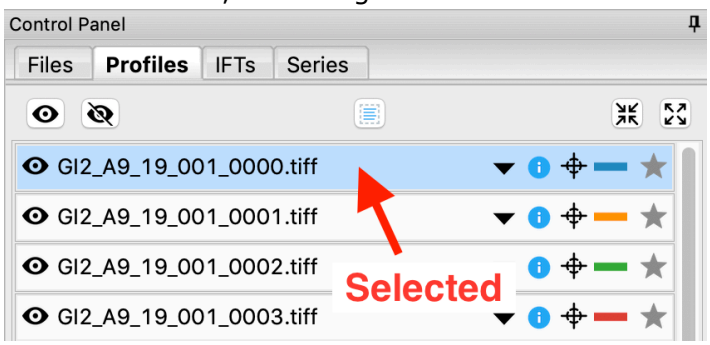
4. Double click on the **SAXS.cfg** file to load the SAXS configuration. This loads the beamline configuration into the program.
    - *Note*: Any time you are going to process images, you need to load the appropriate configuration!
5. Change the file type filter to "TIFF files (*.tiff)".
6. For batch mode samples, typically only a handful of images are collected from a given sample, in this case 10 per sample. Load in the 10 images for the glucose isomerase (GI) sample by selecting the files **GI2_A9_19_001_xxxx.tiff**, where **xxxx** will range from **0000** to **0009**. These files are measured scattering from 0.47 mg/ml GI.
    - *Tip:* you can hold down the ctrl key (apple key on macs) while clicking to select multiple files individually. You can also click on a file, and then shift click on another file to select those files and everything between them.
    - *Warning:* Don't load the files with **PIL3** in their name. Those are the wide-angle scattering (WAXS) data, which we will not use.
7. Click the plot button to integrate all of the images and plot the integrated scattering profiles on the Profiles plot.
    - *Note:* Typically, once the images are integrated we work only with the scattering profiles. However, it is useful to keep the images around in case you want to reprocess the data.



8. Plot the **GIbuf2** scattering profiles from the images. These are measured scattering from the matching buffer, without any protein, for the GI sample.
9. Click on the Profiles tab. This is where you can see what scattering profiles are loaded into RAW, and manipulate/analyze them.
    - *Checkpoint:* If you've successfully loaded the images given, you should see twenty scattering profiles in the profiles list, with names like **GI2_A9_19_001_0000.tiff** or **GIbuf2_A9_18_001_0000.tif**.

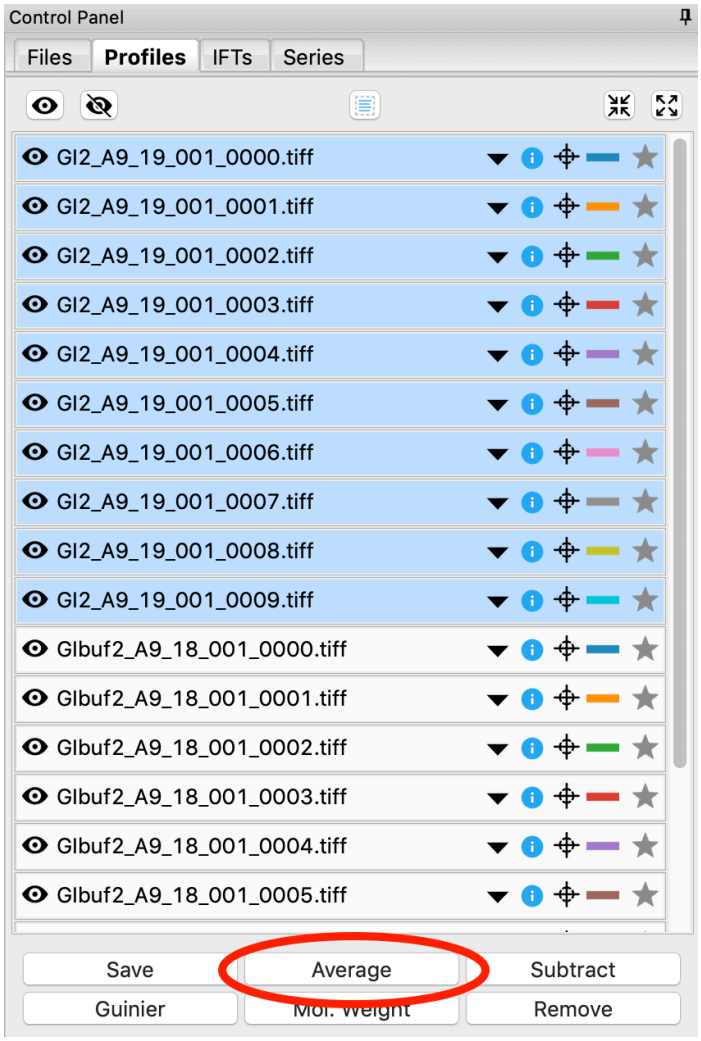


10. Click on a filename to select the scattering profile. The background should turn blue, indicating it is selected.



11. Select all of the GI scattering profiles

- *Tip:* Again, the ctrl(/apple) key or the shift key can be used to select multiple scattering profiles.
- *Warning:* Select only the GI profiles, not the GI buffer profiles.



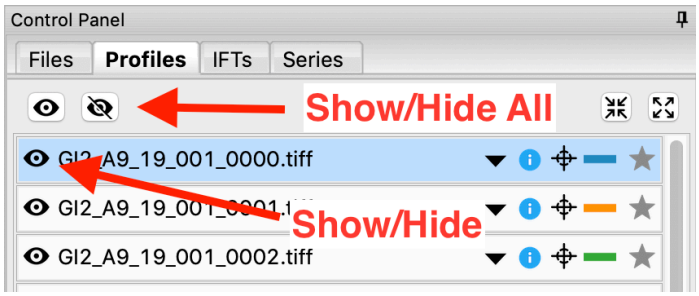
12. Use the average button to average all of the scattering profiles collected into a single curve.

- *Checkpoint:* The averaged scattering profile should appear at the bottom of the profiles list. You may have to scroll down to see it. The filename will be in green, and will start with **A_**, indicating it is an average scattering profile.

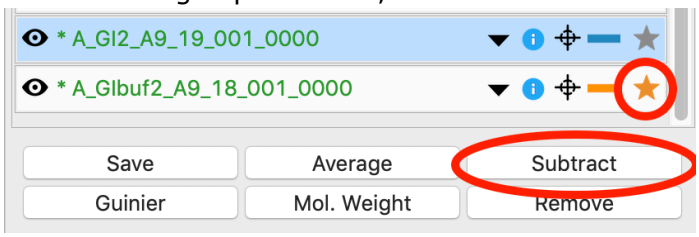
13. Average all of the GI buffer scattering profiles.

14. In order to see the averaged scattering profiles, you will need to hide the individual profiles from the plot. Clicking on the eye to the left of the filename will show/hide a scattering profile. When the eye is shown, the profile is shown on the plot, when the eye has a line through it, the profile is hidden. Hide all of the profiles except the two averaged curves.

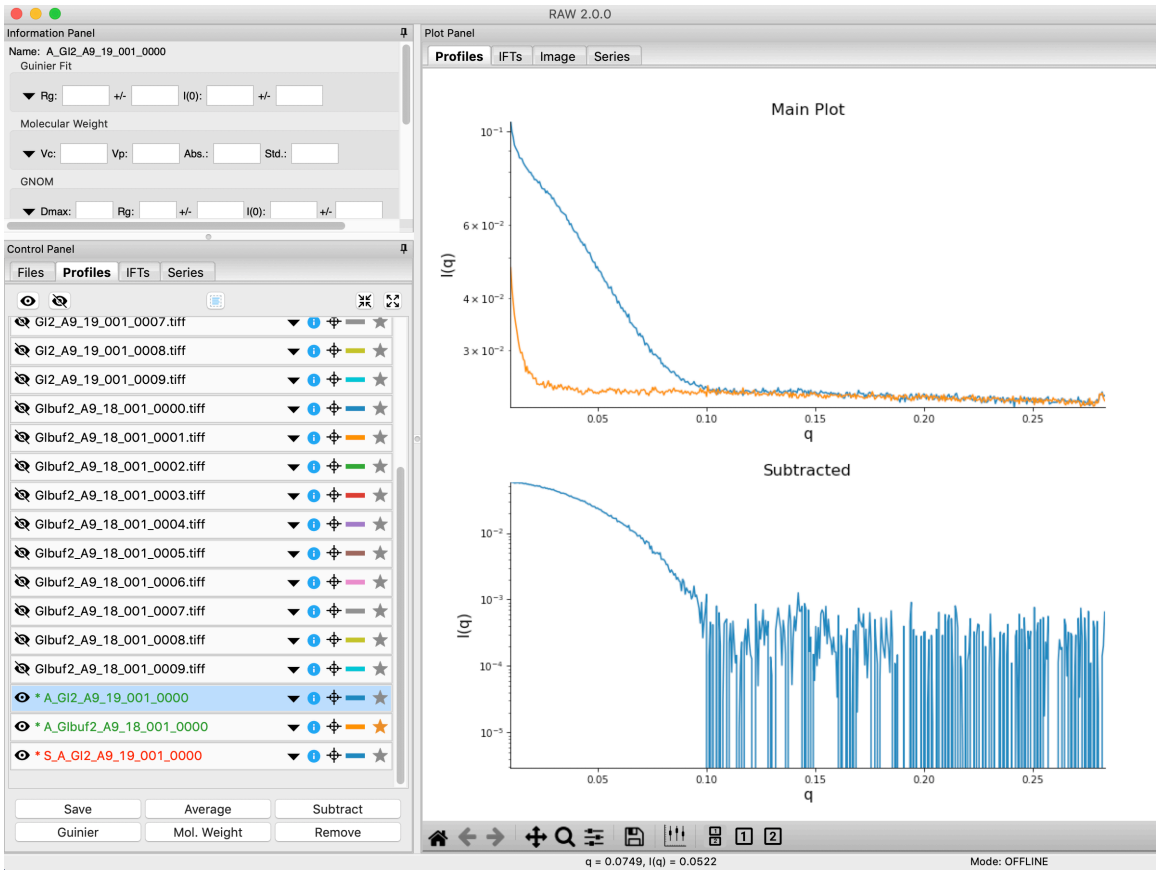
- *Tip:* The eye and eye with the line through it at the top of the profiles panel can be used to show/hide sets of loaded profiles at once. If no profiles are selected, these buttons show/hide all loaded profiles. If some profiles are selected, these buttons show/hide just the selected profiles. Try selecting all but the averaged files and using the show/hide all buttons.



15. Next you need to subtract the buffer scattering profile from the measured protein scattering (which is really the scattering of the protein plus the scattering of the buffer). Star the averaged buffer file, and select the averaged protein file, then click the subtract button.



- *Checkpoint:* The subtracted scattering profile should be shown in the lower plot. A new profiles item should be shown in the Profiles list, with the name in red and a **S_** prefix indicating it is a subtracted file.



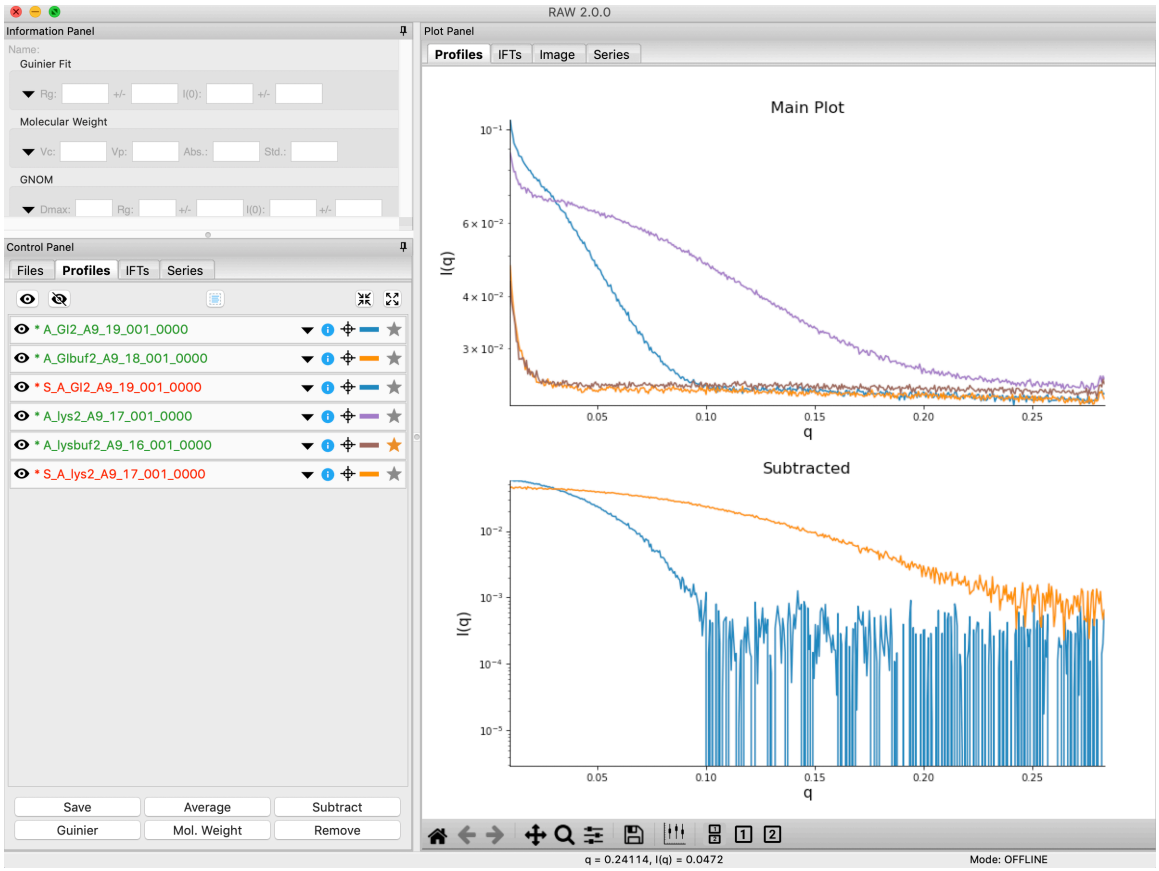
16. You don't need the individual image scattering profiles any more. Select all of those (but not your averaged or subtracted profiles!) and click remove.

- *Note:* This only removes the scattering profiles from RAW. The images on your hard drive are unaffected.

17. You can also load files into RAW by dragging and dropping files onto the RAW window. Load in the **lys2** images by selecting them in your file browser, then dragging them onto the top plot.

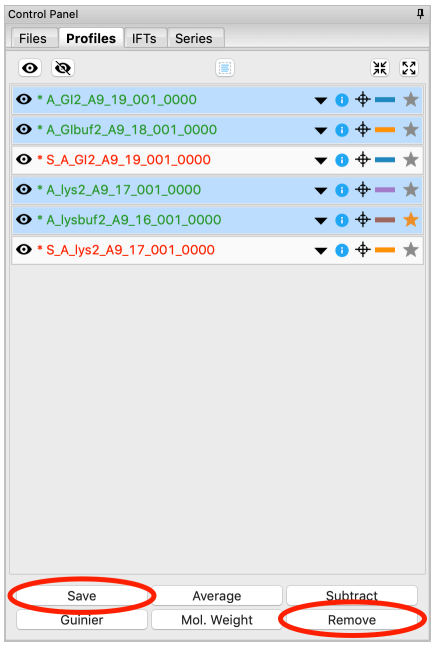
18. Load in the **lysbuf2** files. Average both the lysozyme and buffer data, and subtract to create a subtracted lysozyme scattering profile. The concentration of this sample was 4.27 mg/ml. Remove all of the profiles that are not averaged or subtracted profiles.

- *Tip:* In order to tell which curve is which in a plot, click on the target icon (right) in the profiles list. This will bold that curve in the plot. Click the target icon again to return the curve to normal.



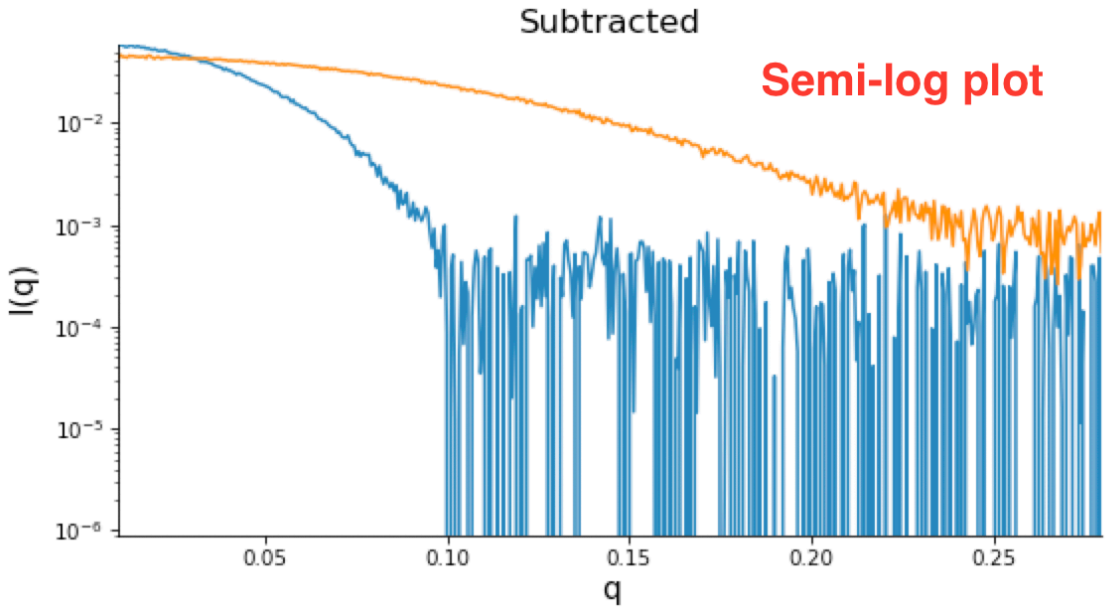
19. We're done with the averaged profiles. Select all of the averaged profiles and click the "Save" button to save them in the **standards_data** folder. Note that in the filename in the profiles list, the * at the front goes away. This indicates there are no unsaved changes to those scattering profiles.

- *Note*: This saves them with a **.dat** extension. This is the standard format for SAXS scattering profiles, and is also human readable.



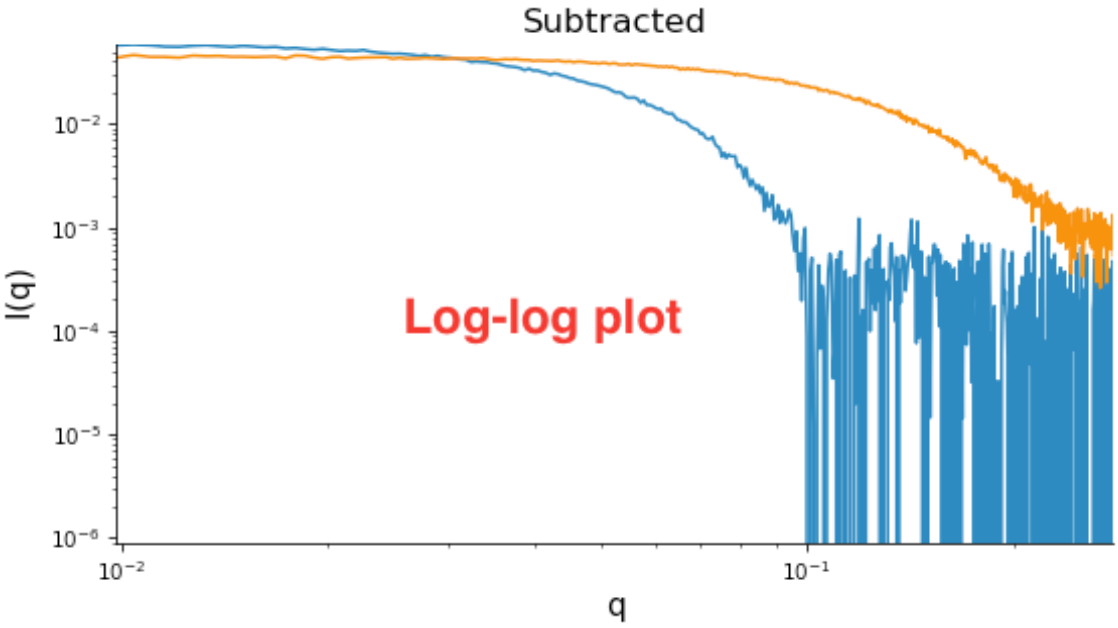
20. Right click on the subtracted plot, move the cursor over 'Axes'. Notice that the currently selected plot type is Log-Lin (a semi-log plot).

- *Question:* You should see a (noisy) 'bounce' in the GI scattering profile. Why does this happen?



21. Change the subtracted plot style to Log-Log (double log) plot. Well-behaved globular proteins will intersect the intensity axis roughly perpendicularly.

- *Note:* It is best practice to display SAXS data, particularly in publications, on either a semi-log or double-log plot (depending on the features of interest).

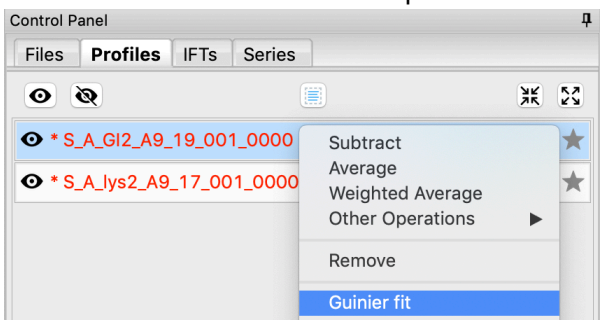


Part 2. Guinier analysis

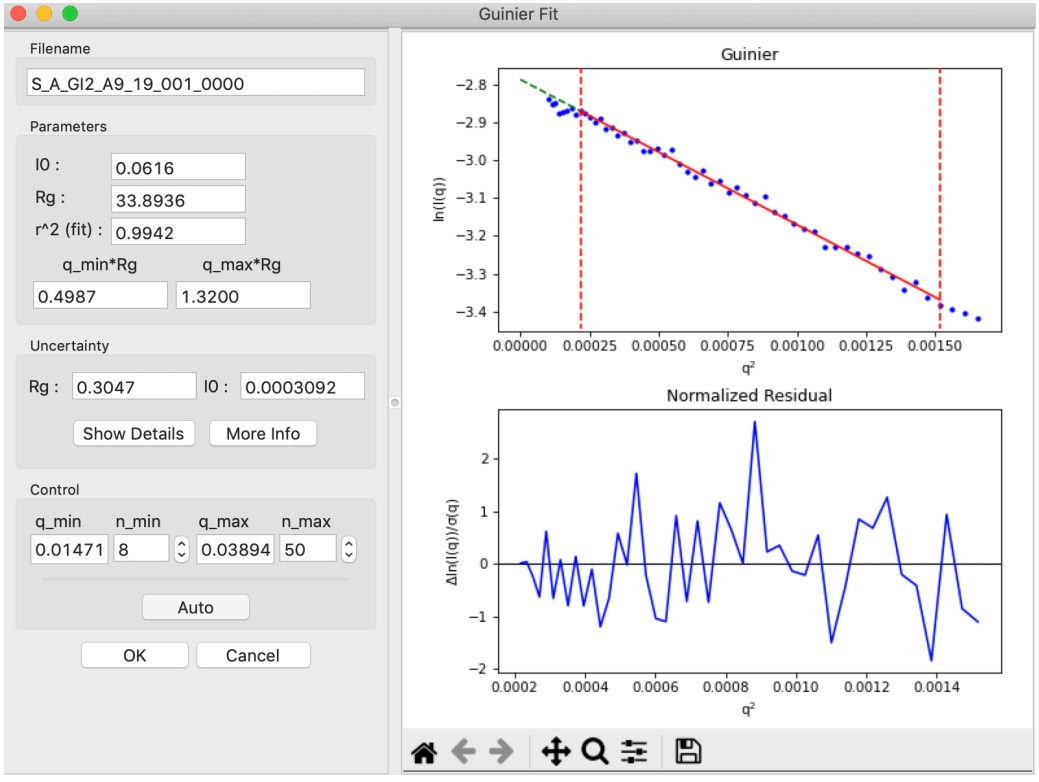
Recall Guinier's approximation at low-q: $I(q) \approx I(0)\exp(-R_g^2 q^2 / 3)$.

R_g and I(0) can be determined by performing a linear fit in the Guinier plot (a plot of $ln(I)$ vs. q^2). The fitting region should normally have $q_{max}R_g < 1.3$ for globular proteins. This fitting region is called the "Guinier region."

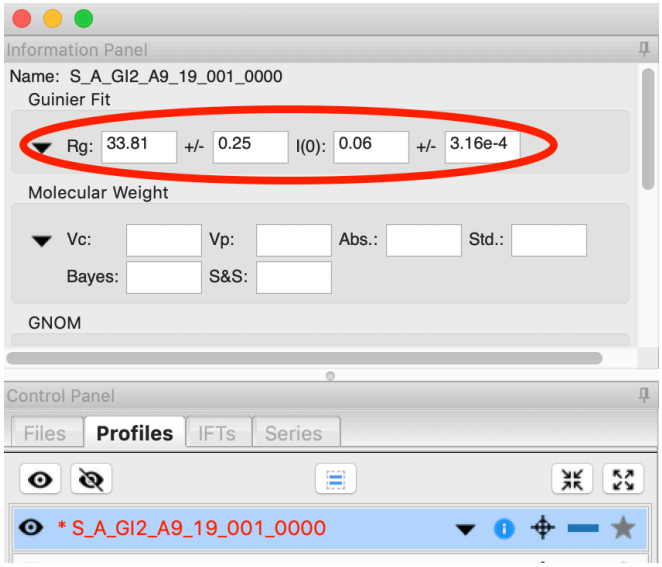
1. In RAW, right click on the subtracted GI scattering profile in the Profiles list and select "Guinier fit". The Guinier fit window will open.
    - *Note:* You can also click the 'Guinier' button at the bottom of the Profiles control panel.



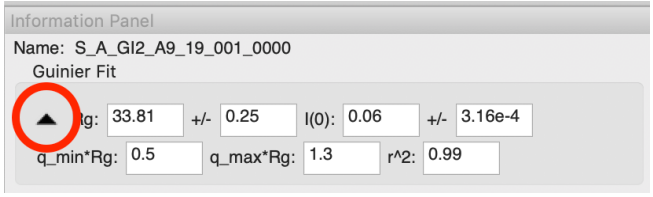
2. In the Guinier window, the top plot shows you the Guinier plot and the fit, while the bottom plot shows you the residual of the fit.
    - *Note:* RAW automatically tries to find the best Guinier region for you when the Guinier window is opened for the first time. Clicking the "Auto" button will redo this automatic determination.
    - *Note*: The R_g value is in units of $1/q$ (e.g. if q is in $Å^{-1}$ then R_g is in Å).



3. In the "Control" panel, you'll see that n_min is 8. This means RAW has skipped the first few low q points for the Guinier fit. You can see a little dip in the lowest q values, which may be why it was skipped. Use the arrow buttons next to the n_min box to adjust it down several points to remove that dip and check whether the R_g changes. Once you're done return n_min to 8.
4. In the "Parameters" panel, note that $q_{max}R_g$ is ~1.32. Recall that for globular proteins like GI, it is typical to have $q_{max}R_g$ ~1.3. Adjust n_max down slightly until that is the case, watching what happens to the R_g and the residual.
    - *Try*: increase n_max until you are truly out of the linear region. How long does that take?
    - *Question*: the literature radius of gyration for GI is 32.7 Å. How does yours compare?
5. RAW also provides an estimate of the uncertainty in both the R_g and I(0) values for the Guinier fit, shown in the Uncertainty section.
    - *Note:* This is the largest of the uncertainties from the fit (standard deviation of fit values calculated from the covariance matrix), and either the standard deviation of R_g and I(0) across all acceptable intervals found by the automatic R_g function or an estimated uncertainty in R_g and I(0) based on variation of the selected interval start and end points.
6. Click the "OK" button to keep the results,
    - *Note:* Clicking the "Cancel" button will discard the results.
7. If you now select the GI scattering profile, in the information panel above the control panel you should see the R_g and I(0) that you just found.



- *Tip:* Click on the triangle to expand the Guinier info section and see more details on the fit.



8. Repeat the Guinier analysis for lysozyme.
    - *Try*: Increase qmin and/or decrease qmax to verify that the R_g does not change significantly in the Guinier region.
    - *Tip*: If you hover your mouse cursor over the info icon (just left of the target icon) for a given scattering profile it should show you the R_g and I(0) of your Guinier analysis.

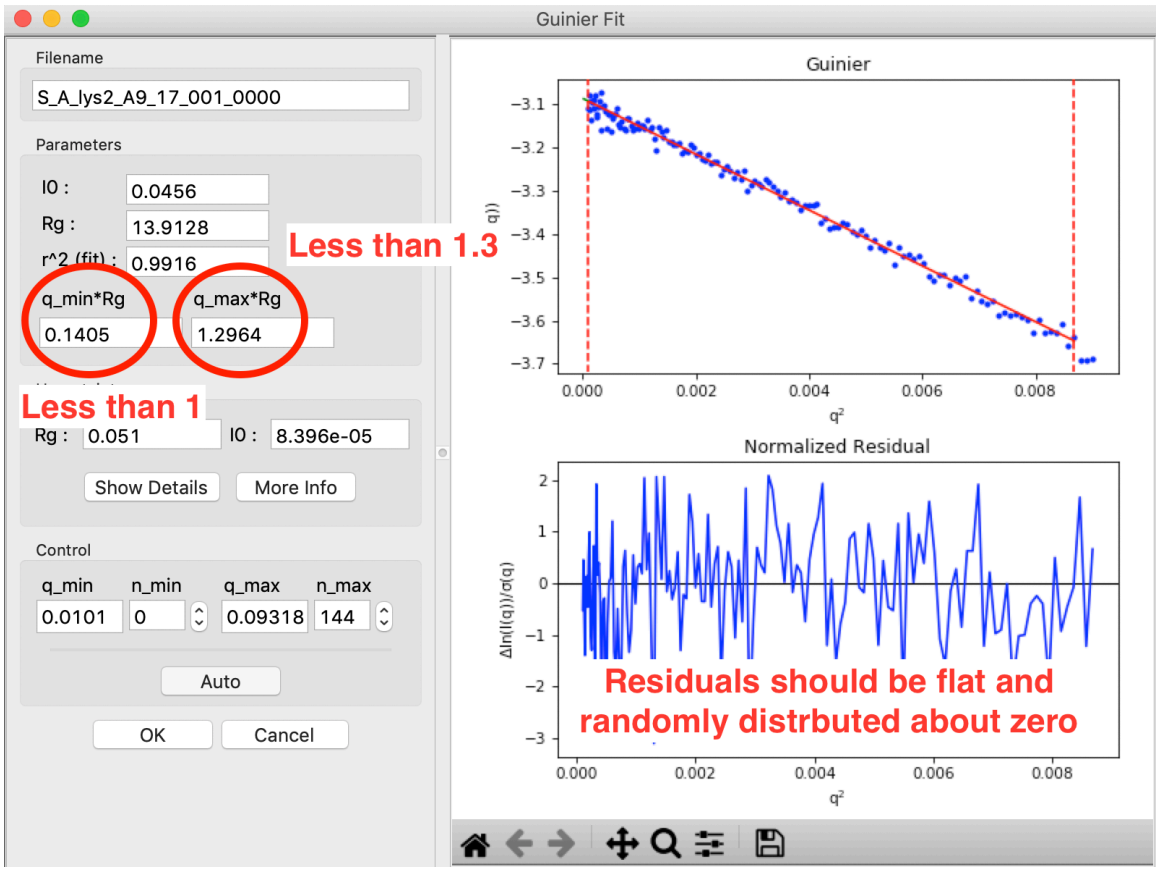
Aside: Criteria for a good Guinier region

You are looking for four essential components in your Guinier fit:

- **$q_{min}R_g < 0.65$.**
    - The minimum q of your fit, q_{min}, times the R_g of your fit should be less than 0.65. This criteria ensures you have enough *q* range to properly estimate the R_g and I(0) values. For globular particles (sphere- or disk-like), you can get away with $q_{min}R_g < 1.0$.
- **$q_{max}R_g < 1.3$ (globular) or $q_{max}R_g < 1.0$ (extended).**
    - The maximum q of your fit, q_{max}, times the R_g of your fit should be less than 1.3 for globular (sphere- and disk-like) particles and less than 1.0 for extended (rod-like) particles. This ensures you remain in the linear range of the Guinier approximation for the fit.
- **The Guinier fit residuals should be flat and randomly distributed about zero.**
    - If your residuals have a 'smile' (above zero near start and end of fit, below in the middle), or a 'frown' (below zero near start and end of fit, above in the middle), it indicates you have non-ideal data. The 'smile' is characteristic of aggregation, the 'frown' characteristic of interparticle repulsion.
- **The fit extends to the lowest available q point.**
    - You shouldn't have to exclude very many points at the start of the fit. A few is generally fine, as points nearest the beamstop can be noisy (depending on the exact details of the measurement). Having to exclude more than ~3-5 points at the low q may indicate a problem with your data.

If you have a small amount of aggregation or repulsion it may manifest as a small upturn or downturn at low *q* that, once excluded, doesn't seem to affect the fit residual (i.e. no 'smile' or 'frown'). In these cases, you may proceed, but exercise caution as your data may be subtly affected. Also, be sure whoever you present the data so that your audience understands you observed these effects and decided to proceed with analysis despite the non-ideality.

Having a good Guinier fit is a major quality check, and a good sign that your data is from a monodisperse sample with no interparticle interactions.


Guinier Fit
Filename
S_A_lys2_A9_17_001_0000
Parameters
I0 : 0.0456
Rg : 13.9128
r^2 (fit) : 0.9916
q_min*Rg
0.1405
q_max*Rg
1.2964
Less than 1.3
Less than 1
Rg : 0.051
I0 : 8.396e-05
Show Details
More Info
Control
q_min
n_min
q_max
n_max
0.0101
0
0.09318
144
Auto
OK
Cancel
Guinier
ln(I(q))
q²
Normalized Residual
Δln(I(q))/σ(q)
Residuals should be flat and
randomly distrbuted about zero

Part 3. Molecular weight analysis

RAW provides four forms of molecular weight analysis:

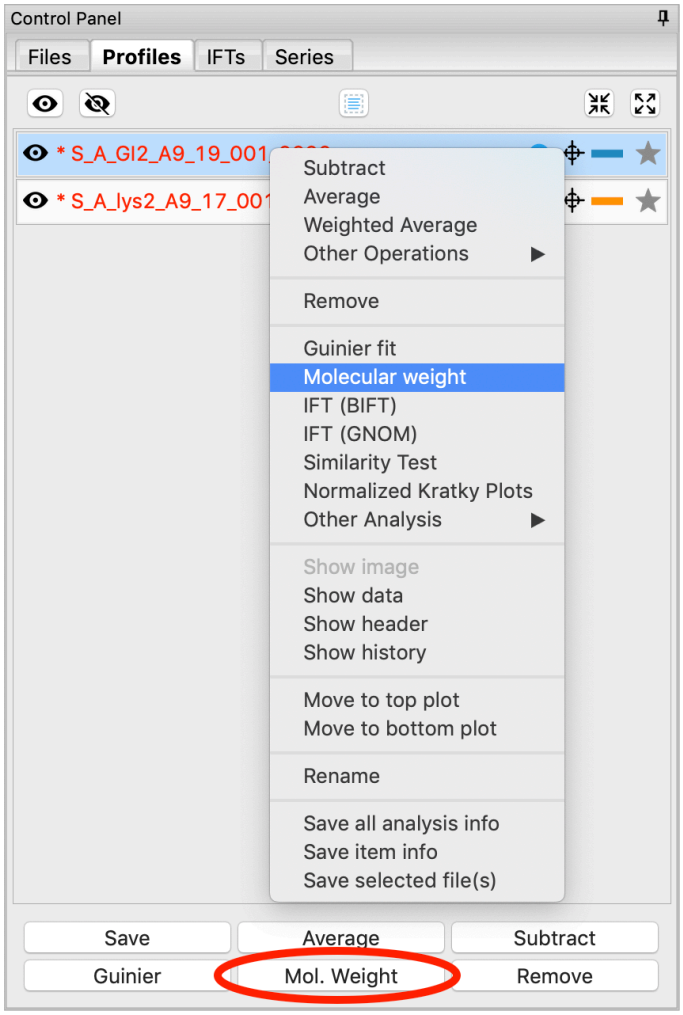
- Referencing I(0) to that of a known standard
- From the volume of correlation using the method of Rambo and Tainer
- From the adjusted Porod volume using the method of Fisher et al.
- From the value of I(0) on an absolute scale.

If ATSAS is installed, RAW also provides an addition two methods of molecular weight calculation, using the ATSAS tools, for a total of six different methods:

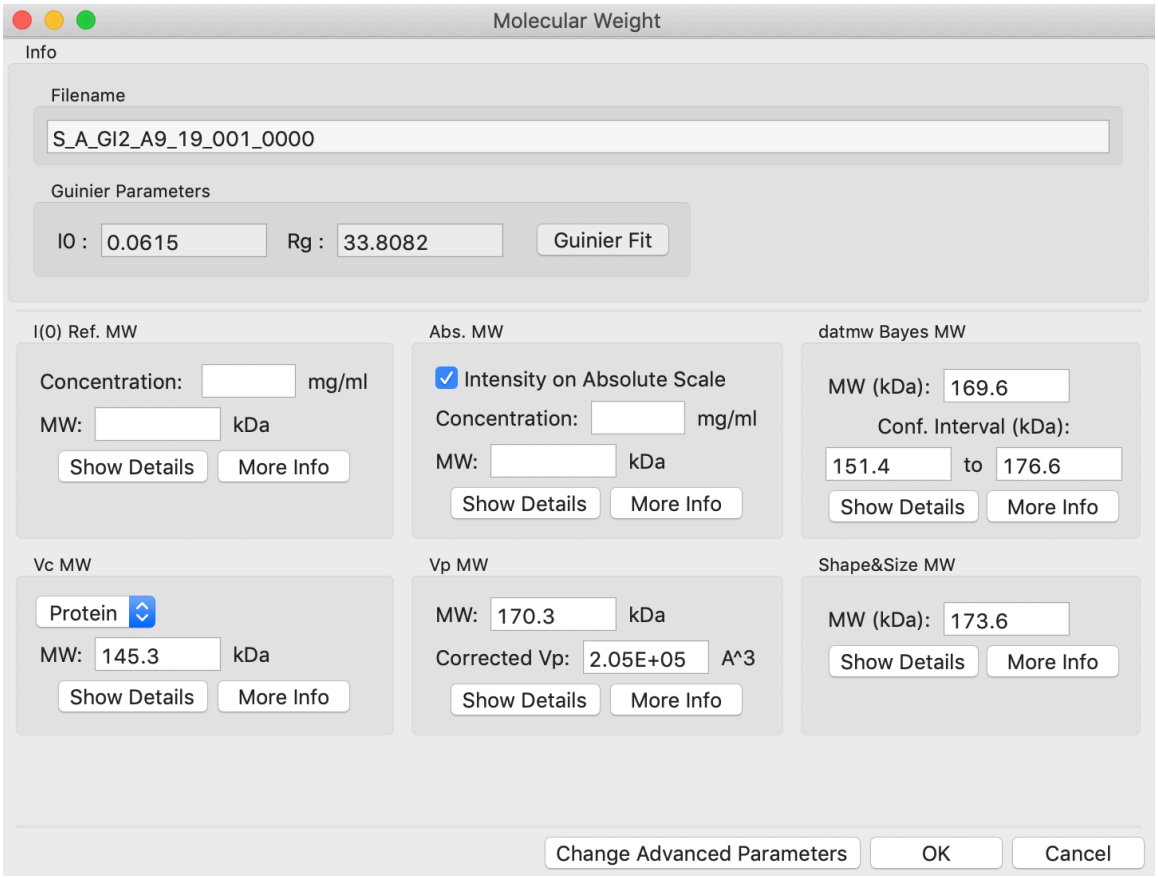
- From classification by machine learning (ATSAS datclass/Shape&Size)
- From a Bayesian estimation based on concentration independent methods (ATSAS datmw bayes)

Each of these has its own advantages and failure modes. A brief discussion of these can be found in RAW by clicking the more info button in the Molecular Weight panel.

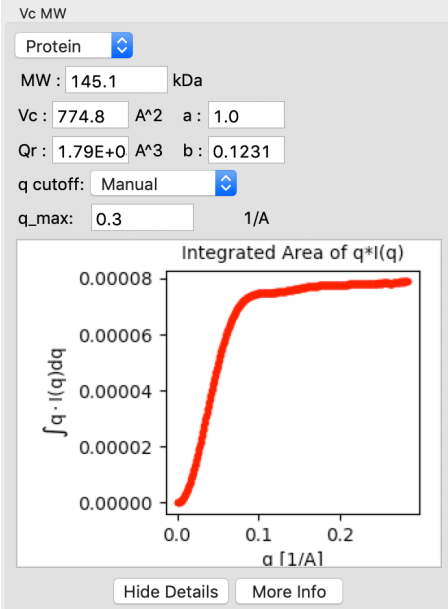
1. In RAW, right click on the subtracted GI scattering profile in the Profiles panel and select "Molecular weight." Alternative click on the "Mol. Weight" button at the bottom of the profiles panel.



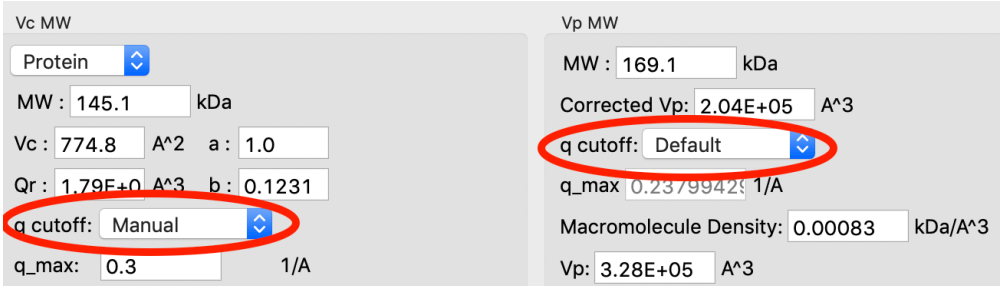
2. At the top of the panel are the results of your Guinier fit. All four methods require a good Guinier fit, so you can use that button to redo the fit if necessary. In the lower part of the panel, the results of the four estimates for MW are shown.
    - *Note:* Neither the I(0) Ref. MW panel nor the Abs. MW panel should be reporting a MW.
    - *Note:* If you have ATSAS installed and accessible to RAW, you should see six panels with MW, as in the image below. If you don't have ATSAS installed you will see just four panels, the rightmost panel in each row will be missing.
    - *Tip:* To learn more about any of the methods click on the "More Info" button.



3. In either concentration box, enter the sample concentration of 0.47 mg/ml. Notice that you now get results from all four methods of MW calculation.
    - *Question:* The expected MW value for GI is 172 kDa. How do your results compare?
4. Click on the "Show Details" button for the Vc MW panel. You should see a graph, which show the integrated area of $q*I(q)$ vs. q. For this method to be accurate, this value needs to converge at high q.
    - *Question:* Why should this value become constant as q increases?



5. Click on the "Show Details" for the Vp MW panel. You'll notice that both Vp MW and Vc MW have a "cutoff" selection. At higher q you can start to get scattering from flexibility or intra-molecular features that may reduce the reliability of the MW estimate. RAW automatically cuts off the scattering at usually good values, but you can change this manually if you need to.



6. Click the "OK" button to save your analysis.
   - *Note:* the "Cancel" button discards the analysis.
7. Repeat the MW analysis for the lysozyme sample, which had a concentration of 4.27 mg/ml. The expected MW of lysozyme is 14.3 kDa.
   - *Question:* Does the Vc MW method work for your lysozyme data?
   - *Note:* Even when data is measured out to sufficient q that the integral has converged, my experience is that the Vc MW method is inaccurate for molecules smaller than ~20 kDa.

Aside: Discussion of the molecular weight methods:

Calculation of MW from reference to a known reference standard

The scattering at zero angle, I(0) is proportional to the molecular weight of the macromolecule, and the concentration and contrast of the macromolecule in solution. If a reference sample of known molecular weight and concentration is measured, it can be used to calibrate the molecular weight of any other scattering profile with known concentration (assuming constant contrast between reference and sample, and a monodisperse sample). Molecular weight is calculated as:

$$MW_m = \left(\frac{I(0)_m}{c_m}\right)\left(\frac{MM_{st}}{\left(\frac{I(0)_{st}}{c_{st}}\right)}\right)$$

where MW is the molecular weight, c is the concentration, and the subscripts 'm' and 'st' designate quantities from the macromolecule of interest and the standard respectively. For a reference see, among many, Mylonas, E. & Svergun, D. I. (2007). J. Appl. Crystallogr. 40, s245-s249.

This method can yield inaccurate results if:
- The reference is not properly calibrated (concentration, I(0) measurement).
- I(0) is poorly determined.
- Sample concentration is poorly determined.
- The contrast between the macromolecule and buffer is significantly different between the reference and sample.
- The standard is a significantly different shape than the sample.

Calculation of MW from absolute scale
This uses the absolute calibration of the scattering profile to determine the molecular weight, as described in Orthaber, D., Bergmann, A., & Glatter, O. (2000). J. Appl. Crystallogr. 33, 218-225. By determining the absolute scattering at I(0), if the sample concentration is also known, the molecular weight is calculated as:

$$MW = \left(\frac{N_a I(0)}{c\Delta\rho_M^2}\right)$$

where N_A is the Avogadro number, c is the concentration, and $\Delta\rho_M$ is the scattering contrast per mass. The accuracy of this method was assessed in Mylonas, E. & Svergun, D. I. (2007). J. Appl. Crystallogr. 40, s245-s249, and for most proteins is <~10%.

This method can yield inaccurate results if:
- The absolute calibration is not accurate.
- I(0) is poorly determined.
- Sample concentration is poorly determined.
- Scattering contrast per unit mass is wrong. This depends on the buffer, macromolecule type (protein vs. nucleic acid), and the macromolecule partial specific volume (which can depend on shape/flexibility). The defaults are for a buffer with the electron density of water and compact globular proteins.

Volume of Correlation method
This method uses the approach described in: Rambo, R. P. & Tainer, J. A. (2013). Nature. 496, 477-481. This method should work for both compact and flexible macromolecules. The authors claim the error in MW determination is ~5-10%.

This method can yield inaccurate results if:
- The integral of q*I(q) doesn't converge (click 'Show Details' to see), which can indicate the scattering profile is not measured to high enough q or that there is a bad buffer match.
- I(0) and/or Rg are poorly determined.
- You have a protein-nucleic acid complex.
- Your molecule is less than ~15-20 kDa.

Adjusted Porod Volume method
This method uses the approach described in: V. Piiadov, E. Ares de Araujo, M. Oliveira Neto, A. F. Craievich, and I. Polikarpov. Protein Science (2019). 28(2), 454-473. It applies a correction to the Porod volume for the finite length of the measurement. The authors report a median of 12% uncertainty for calculated molecular weight from globular proteins.

This method can yield inaccurate results if:
- The molecule is not globular (i.e. is flexible or extended).
- I(0) is poorly determined.
- The protein density used is inaccurate (can be changed in RAW).
- Your molecule is not a protein (e.g. RNA/DNA or a protein-nucleic acid complex).

Comparison to known structures
In Franke, D., Jeffries, C. M. & Svergun, D. I. (2018). Biophys. J. 114, 2485–2492. DOI: 10.1016/j.bpj.2018.04.018 they describe a machine learning method that categories SAXS data into shape categories based on comparison with a catalog of known structures from the PDB. By finding the nearest structures in shape and size (also the name of the method: Shape&Size), they can obtain estimates for the molecular weight of the sample.

The authors found that, for the theoretical scattering profiles used for testing, the method calculated molecular weights within 10% of the expected value for 90% test data. Another paper found that for the test dataset the median molecular weight was correct and the median absolute deviation was 4%. Again, it seems reasonable to say that the uncertainty in molecular weight from this method is ~10% for most systems, though there are outliers.

This method can yield inaccurate results if:
- The system is flexible.
- Your molecule is not a protein.

From Bayesian inference
In Hajizadeh, N. R., Franke, D., Jeffries, C. M. & Svergun, D. I. (2018). Sci. Rep. 8, 7204. DOI: 10.1038/s41598-018-25355-2 they describe a method for calculating a molecular weight using Bayesian inference with the molecular weight calculations from the Porod volume, volume of correlation, and comparison to known structures methods as the evidence. Essentially, it takes a large test dataset of theoretical

scattering profiles, calculates the molecular weight for each using each method, then creates a probability distribution for each method that describes the probability of obtaining a particular calculated molecular weight given the true molecular weight. These probabilities are combined across all the methods, and the most likely molecular weight is thus estimated.

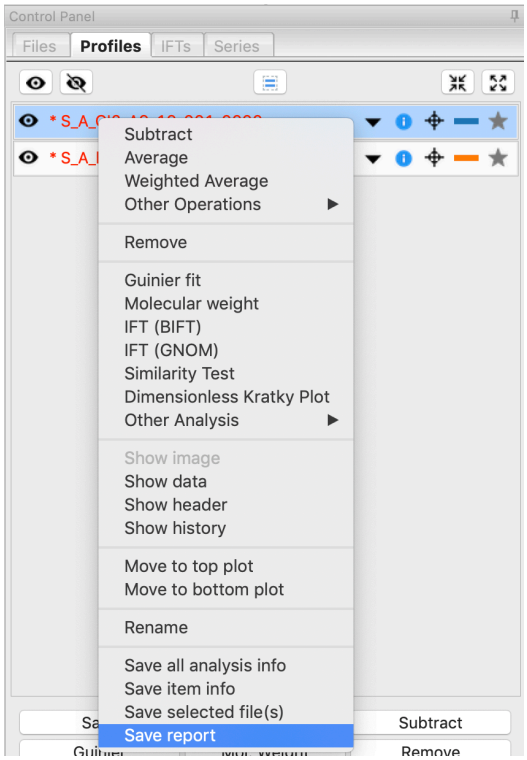
They found that for the theoretical scattering profiles used, the median molecular weight from this method was accurate and the median absolute deviation was 4%. Overall, they reported that it was more accurate than any individual method. It may be that the uncertainty in this method is usually closer to ~5% than 10% for the other methods.

This method could yield inaccurate results if:
- Your molecule is not a protein
- There are significant data quality issues.

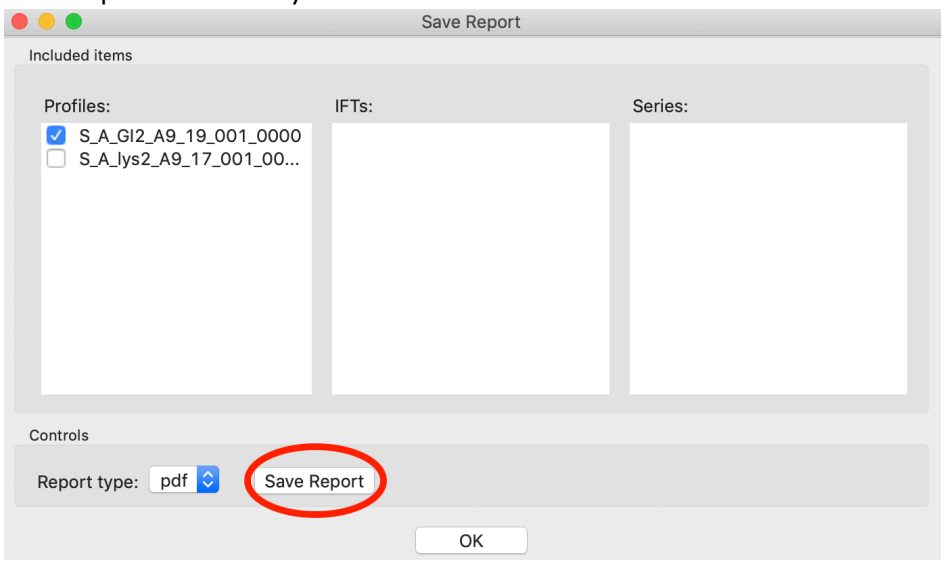
Part 4. Saving analysis information

In RAW there are several ways to save analysis results. First, we will save a summary pdf file of the results, then a .csv file with all of the results.

1. Select the subtracted GI scattering profile in the Profiles panel and right click on it and select "Save report".

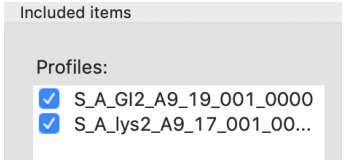


2. In the panel that opens, click the "Save Report" button and save the report as **gi_report.pdf** in the **standards_data** folder. This will save a PDF report with plots and analysis for the GI data.

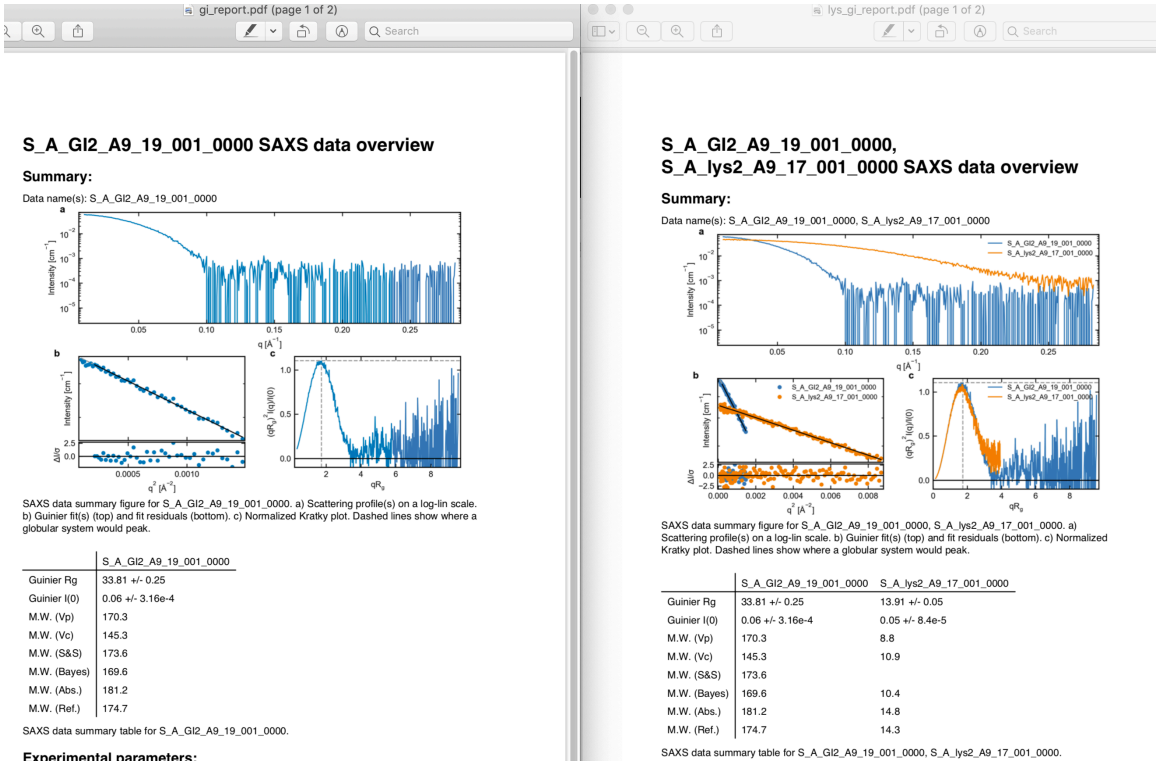


3. You can also save a report for multiple datasets at once. Select both the lysozyme and GI datasets by checking both of the checkboxes. Then click

"Save Report" to save a report with both datasets as **lys_gi_report.pdf** in the **standards_data** folder.



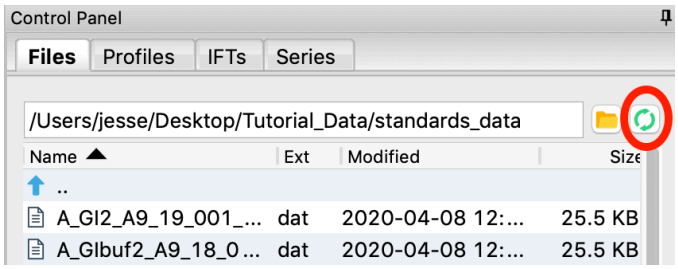
4. Open the two reports in a pdf viewer. You should see just the GI data in the first, whereas the second pdf should have plots and columns in the table for both the lysozyme and GI data.



5. Once you're done saving reports, click the "OK" button to close the window.
6. Next we will save the profiles with the analysis, and save all the analysis as a .csv file.
7. Save your subtracted scattering profiles in the **standards_data** folder.
8. Select both subtracted profiles, right click on one of them, and select 'Save all analysis info.' Give it an appropriate name and save it in the **standards_data** folder.
    - *Note:* This saves a **.csv** file with all of the analysis information for the selected scattering profiles.
    - *Try:* Open the **.csv** file in Microsoft Excel or Libre/Open Office Calc. You should see all of the analysis that you just did.
9. Remove all the scattering profiles from RAW by selecting all of them and clicking the "Remove" button.
10. Load the saved subtracted scattering profiles back into RAW. Note that if you select one in the Profiles list, the information panel in the upper left corner of RAW populates with analysis information. The analysis information is saved

with the scattering profile, so if you forget to save it in a **.csv**, you can load in the profiles later and do it then.

- *Note:* To get new files to show up in the file tab, you may have to click the refresh button. Also, make sure to that your file type filter is either All files or DAT files.



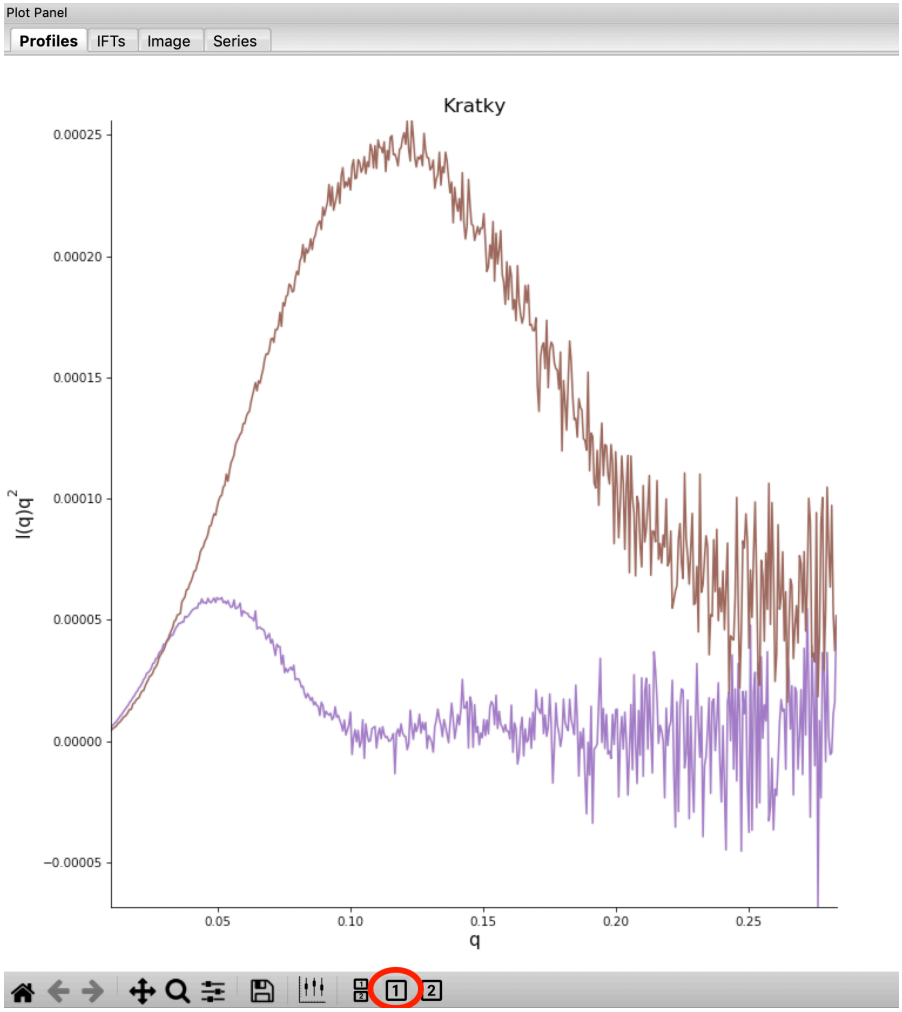
- *Try:* Open the saved subtracted scattering profile **S_A_GI2_A9_19_001_0000.dat** in a text editor such as Notepad (windows) or TextEdit (mac). You should see all of the data in three columns, followed by header information. If you scroll down far enough, the header information contains all of the analysis information, as well as the files that were averaged and subtracted to make the scattering profile.

Part 5. Kratky analysis

A Kratky plot is a plot of $q^2I(q)$ vs. q. Kratky plots can qualitatively assess the flexibility and/or degree of unfolding in samples. Unfolded (highly flexible) proteins should have a plateau in the Kratky plot at high q, while compact, globular proteins will have a bell-shaped (Gaussian) peak. A partially unfolded (flexible) protein may have a combination of the bell-shape and plateau, or a plateau that slowly decays to zero.

Normalized Kratky plots are plots of $q^2I(q)/I(0)$ vs. q. This normalizes scattering profiles by mass and concentration. Dimensionless Kratky plots are presented as either $(qR_g)^2I(q)/I(0)$ vs. qR_g or $(q^2V_c)I(q)/I(0)$ vs. $q(V_c)^{1/2}$. These dimensionless plots can provide semi-quantitative analysis of flexibility and disorder.

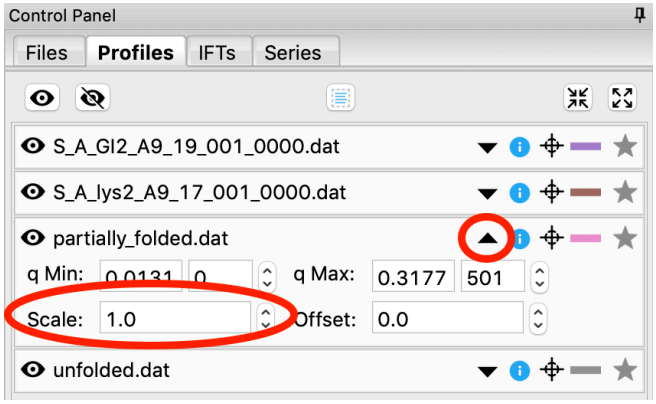
1. Put the top plot on Kratky axes.
    - *Tip:* Right click on the plot to change the plot type.
2. Show only the top plot by clicking on the 1 in the plot control bar below the plots.



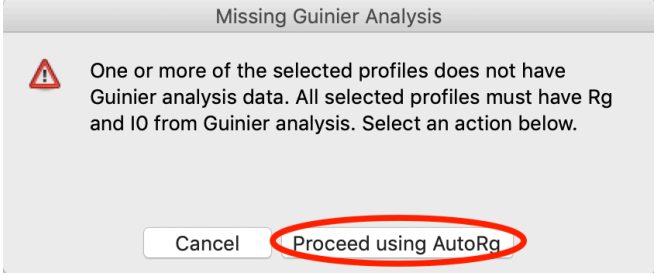
3. Both GI and lysozyme show the classic bell shape, indicating they are completely folded.

- *Warning:* bad buffer subtraction can also result in a Kratky plot that appears to show some degree of flexibility. Excellent buffer subtraction is required for accurately analysis with this technique.

4. Load the two scattering profiles in the **Example_Data/flexibility_data** folder.
    - *Note:* The **unfolded.dat** file is the scattering profile of an unfolded lysine riboswitch. The **partially_folded.dat** file is same lysine riboswitch, but in the biologically functional configuration. The data were downloaded from the BIOISIS database (bioisis.net), and have the BIOISIS ids of 2LYSRR and 3LYSRR
5. SAXS data can be presented on an arbitrary scale, which is why these two profiles have intensity that is much larger than the lysozyme and GI data (which is on an absolute scale). Use the triangle button for each item in the profiles menu to show more options. Hide one of the newly loaded data sets, and adjust the scale factor on the other until you can comfortably see it and your lysozyme and GI data. Repeat the scale adjustment for the other data set.
    - *Tip:* The up and down arrows will only adjust the last digit of the scale factor.

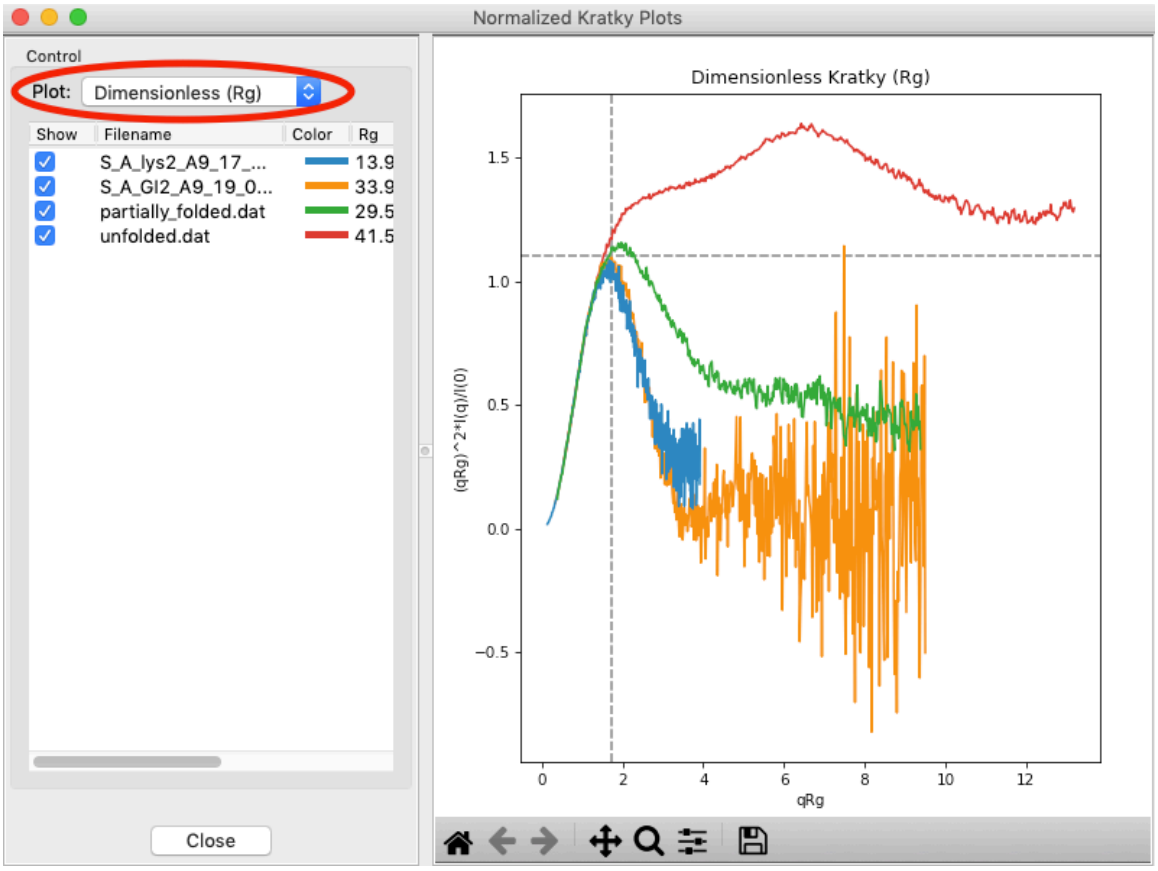


6. Hide the lysozyme and GI data.
    - *Question:* What differences do you observe in the two riboswitch data sets?
7. Kratky analysis can also be done on normalized or dimensionless data. RAW supports normalization by I(0), and non-dimensionalization by R_g and Vc (the volume of correlation). This allows semi-quantitative analysis of flexibility.
8. Select all four loaded scattering profiles, right click, and select the Dimensionless Kratky Plot option.
9. Normalized and dimensionless Kratky plots require Guinier analysis to be done. If one or more profiles are missing this information, RAW will show the following window. You can either cancel, and do the fits manually, or you can proceed with RAW's automatic determination.
10. Click the Proceed using AutoRg button to proceed to the Dimensionless Kratky Plot window using RAW's automatic fitting for R_g.



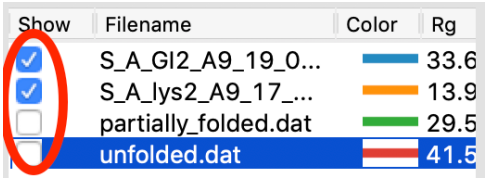
11. By default, the plot is the Dimensionless R_g plot. Use the dropdown "Plot" menu at the top to select the Normalized (by I(0)) and Dimensionless Vc plots.

- Tip: The dashed gray lines on the Dimensionless (Rg) plot are guidelines for a globular protein. For a globular protein the peak position should be at $qR_g = \sqrt{3} \approx 1.73$, while peak height should be $3/e \approx 1.1$.



12. Return to the Dimensionless R_g plot. Use the check boxes to hide the partially_folded and unfolded data sets on the plot. Note that both the lysozyme and GI data look very similar on this plot, showing they have similar shapes and (lack of) flexibility.

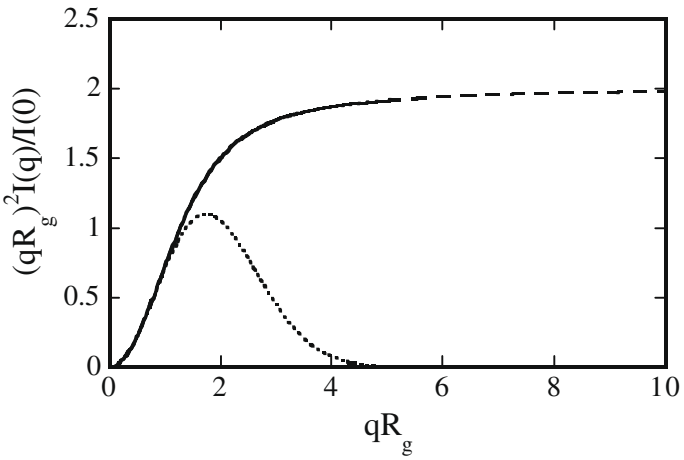
- *Tip*: You can click on the colored line in the Color column to change the color of an item on the plot.



13. Right click on the plot and select "Export Data as CSV" to save the dimensionless data for further processing or plotting with another program.
14. Turn the partially_unfolded and unfolded data sets back on. What does the dimensionless Kratky plot tell you about the samples?
15. Click the 'Close' button to close the Dimensionless Kratky plot window.

Aside: Dimensionless Kratky plots

Globular compact particles that obey Guinier's approximation will show a maxima of 1.1 at $qR_g=\sqrt{3} \approx 1.73$. A curve rising to a plateau at a value of 2 is expected for an ideal random chain, which may be approximated by a completely unfolded protein. A fully extended chain will continue to slope upwards without a plateau.



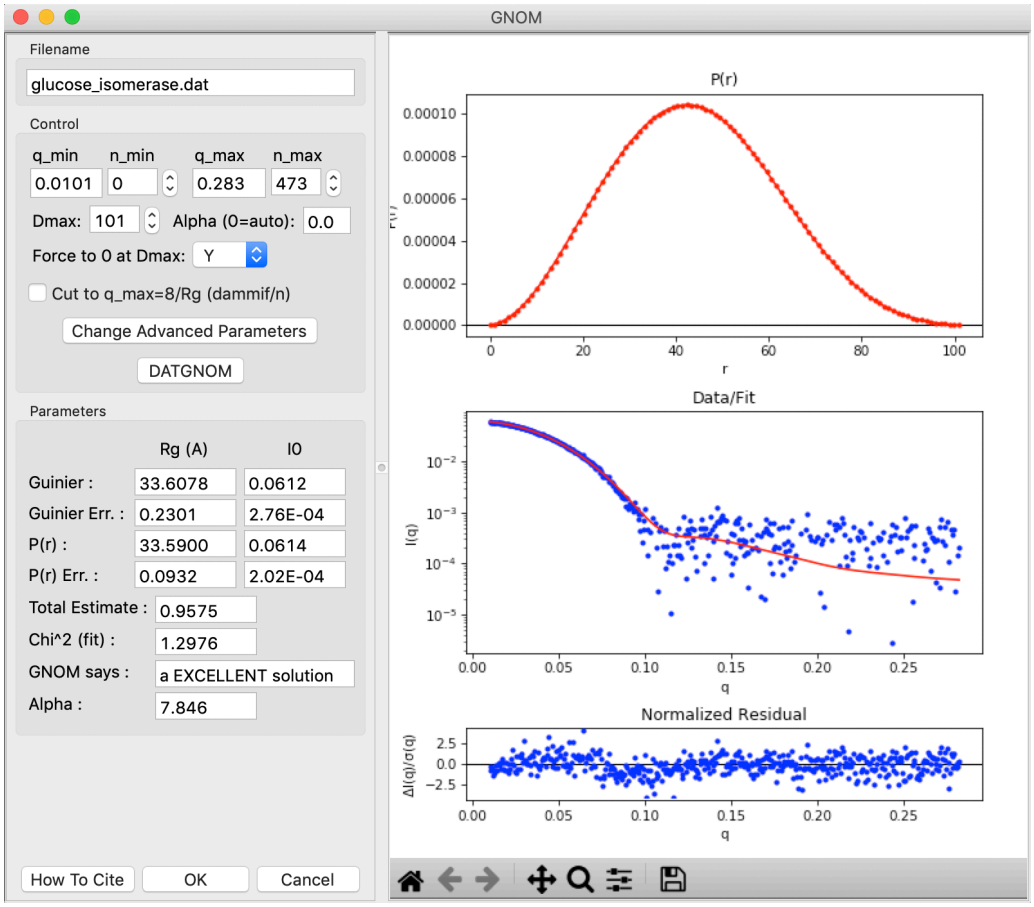
Theoretical dimensionless Kratky profiles for a globular (dotted) and completely unfolded (solid/dashed) protein. Figure from *Durand, D., et al. (2010). J. Struct. Biol. 169, 45–53*.

Because the non-dimensionalization puts the scattering profiles on the same scale, the shape of the profile can be directly compared. A shift of the peak to the right, or a plateau, indicates greater flexibility of the system.

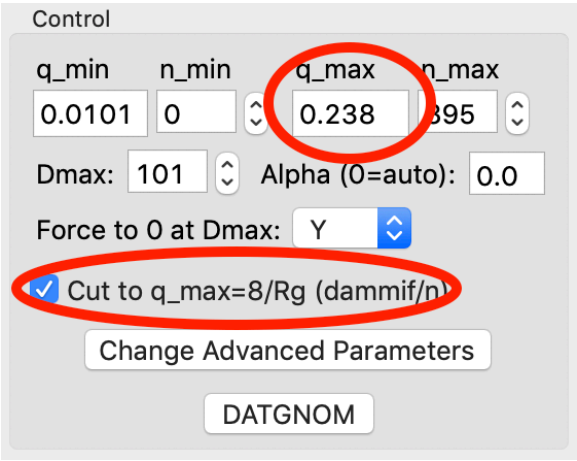
Part 6. Pair-distance distribution analysis – GNOM in RAW

The first step in most advanced data processing is to calculate the P(r) function, the inverse Fourier transform of I(q). This cannot be calculated directly from the scattering profile, so indirect Fourier transform methods are typically used. The most common such method is implemented in the GNOM program from the ATSAS package. We will use RAW to run GNOM. Note that RAW also supports a Bayesian method for generating the P(r) function, BIFT.

1. Open the **glucose_isomerase.dat** file in the **Example_Data/reconstruction_data** folder.
2. Right click on the glucose_isomerase profile in the Manipulation list and select "IFT (GNOM)".
    a. *Note:* RAW will automatically try to find an appropriate maximum dimension (D_{max}) by running the DATGNOM program from the ATSAS software package.
    b. *Troubleshooting:* If you do not have the GNOM option in the right click menu, RAW does not know where your ATSAS version is installed. If you installed the ATSAS programs after starting RAW, restart RAW and it will attempt to automatically find them. If that has failed, go to the Options->Advanced Options menu and choose the ATSAS settings ("ATSAS"). Uncheck the option to "Automatically find the ATSAS bin location" and specify the location manually.



3. The GNOM panel has plots on the right. These show the P(r) function (top panel), the data (middle panel, blue points) and the fit line (middle panel, red line), and the fit residual (bottom panel).
    a. *Note:* The fit line is the Fourier transform of the P(r) function and is also called the regularized intensity.
4. On the left of the GNOM panel are the controls and the resulting parameters. You can alter the data range used, the D_{max} value, and the alpha value used.
    a. *Tip:* The Guinier and P(r) R_g and I(0) values should agree well for mostly rigid particles. The total estimate varies from 0 to 1, with 1 being ideal. GNOM also provides an estimate of the quality of the solution. You want it to be at least a "REASONABLE" solution.
5. Vary the D_{max} value up and down in the range of 80-110. Observe what happens to the P(r) and the quality of the solution.
    a. *Note:* D_{max} is in units of Å.
    b. *Tip:* See the aside at the end of the section for details of what we look for in a good P(r) function. Generally speaking, we want:
        i. A smooth approach to zero at D_{max}.
        ii. A good fit to the measured scattering profile.
        iii. Goes to zero at r=0 and r=D_{max}.
        iv. The Rg and I(0) from the Guinier fit and P(r) function agree well.
        v. The P(r) function is positive everywhere.
    c. *Question:* What D_{max} values seem reasonable to you? What does GNOM have to say about them?
6. Return the D_{max} value to that found by DATGNOM by clicking the "DATGNOM" button. D_{max} should be 101. By default, GNOM forces the P(r) function to zero at D_{max}. For a high-quality data set and a good choice of D_{max}, P(r) should go to zero naturally. Change the "Force to 0 at Dmax" option to "N".
    a. *Try:* Vary D_{max} with this option turned off.
7. Reset it so that the P(r) function is again being forced to zero at D_{max}.
8. RAW makes it easy to truncate your data to a q_{max} of $8/R_g$, which is the recommended maximum q value for bead model reconstructions with DAMMIF. Check the 'Cut to q_max=8/Rg' box to truncate the data.



    a. *Note:* The qmax goes from 0.283 to 0.238 when you check the box.

   b. *Tip:* For electron density reconstruction with DENSS use the full available q range.

9. Make sure the D_{max} is 101 and click OK. This saves the results into the RAW IFT panel.
10. Click on the IFT Control and Plot tabs. This will display the GNOM output you just generated. Save the **glucose_isomerase.out** item in the **reconstruction_data** folder.
    a. *Note:* This saved file is all of the GNOM output, in the GNOM format. It can be used as input for any program that needs a GNOM **.out** file.
    b. *Tip:* When you select the IFT item in the IFT control panel, data from the IFT shows up in the Information panel.
11. Repeat the analysis with the lysozyme data from earlier in the tutorial. Try the following method of finding Dmax, which is what I usually do:
    1. Open the GNOM interface. It defaults to what GNOM thinks is a reasonable D_{max}.
    2. If necessary, set the starting q value for the P(r) function to match that of the Guinier fit (RAW should do this automatically for datasets where the Guinier fit was done in RAW).
    3. Set the D_{max} value to 2-3 times larger than the initial value.
    4. Look for where the D_{max} value drops to 0 naturally. Set the D_{max} value to this point.
    5. Turn off the force to zero at D_{max} condition.
    6. Tweak D_{max} up and down until it naturally goes to zero with the force to zero off.
    7. Turn the force to zero at D_{max} condition back on.
    8. If needed, truncate the P(r) function to a maximum of $8/R_g$ or q~0.3 $Å^{-1}$ for bead model reconstructions with DAMMIF/N.

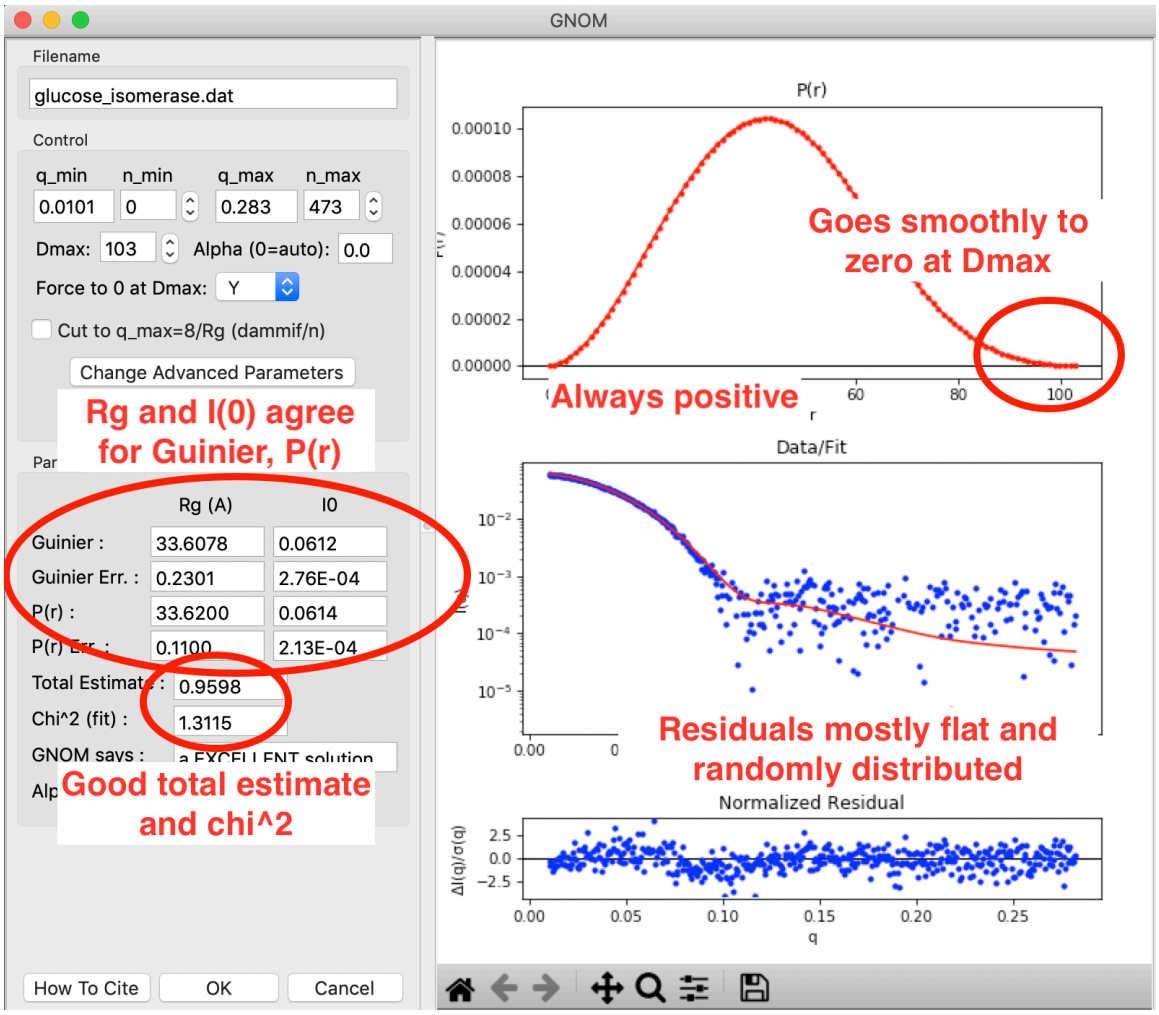
Aside: Criteria for a good P(r) function

We employ the following criteria to determine if a P(r) function is good:

1. The P(r) function falls gradually to zero at Dmax.
2. The P(r) function fits the measured scattering profile
3. The P(r) function goes to zero at r=0 and r=D_{max}.

Additionally, the following criteria usually apply:

1. The R_g I(0) from the Guinier fit and the P(r) function agree well.
2. The P(r) function is always positive.



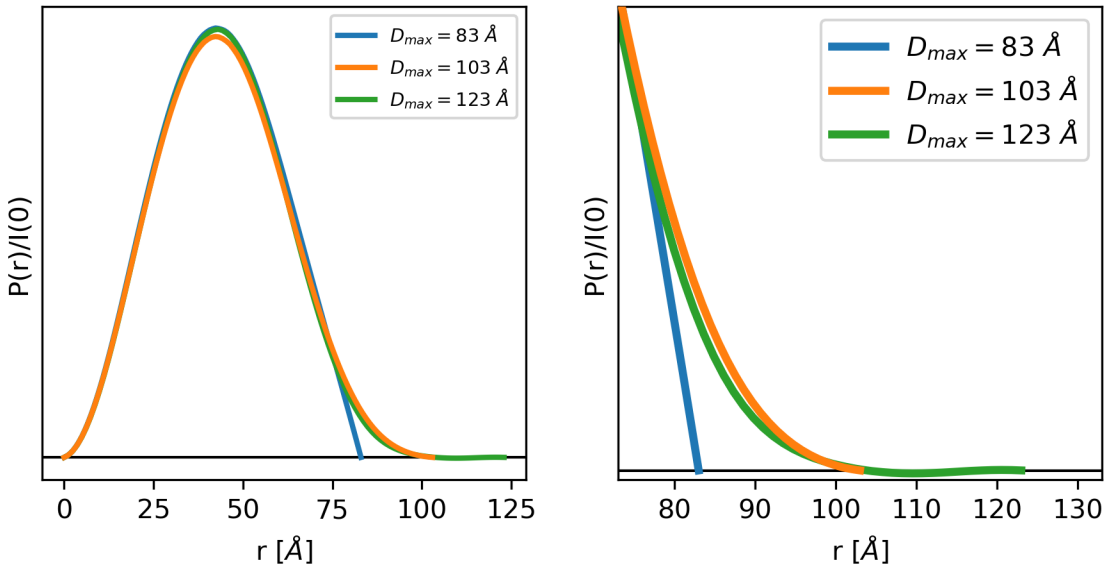
The figure above shows a P(r) function done in RAW using GNOM for glucose isomerase. This shows what a good P(r) function looks like. The function goes smoothly to zero at D_{max}, it is always positive, there is good agreement between the Guinier and P(r) R_g and I(0) values, and the residuals are mostly flat and randomly distributed. You can see a small systematic deviation in the residuals below q~0.125. This could be smoothed out by reducing the alpha value, which may be slightly over weighting the regularization parameters vs. the actual fit to the data.

A more thorough discussion of each criterion is given below.

The P(r) function falls gradually to zero at D_{max}

This is perhaps the most important and most subjective criterion for determining whether you have a good P(r) function. The idea is straightforward: macromolecules do not have perfectly sharp boundaries. Because they have side chains that stick out, and have some amount of flexibility in solution (even if limited to solvent exposed side chains), there is no distance at which you go from many electron pairs to no electron pairs within the macromolecule. As such, the P(r) function should gradually approach zero at the maximum dimension, rather than being cut off.

Essentially, if this criterion is met you have picked an appropriate D_{max} for the system. If you underestimate the D_{max}, then the P(r) function has an abrupt descent to zero, while an overestimated D_{max} usually shows an oscillation about zero. This is shown in the figure below.



The left and right plots show three different P(r) functions for the same protein (glucose isomerase). The difference between the three is the D_{max}, which is either 83 (blue), 103 (orange) or 123 (green) Angstrom. The left plot shows the full P(r) function. The different D_{max} values yield similar P(r) functions, so much so that they end up plotted on top of each other for most of their r values. The right plot is the same functions showing just the end, as P(r) approaches zero at D_{max}.

In the plot above, we can clearly see that for a D_{max} of 83, the P(r) function is forced abruptly down. For a D_{max} of 103, the function has a smooth approach to zero. For a D_{max} of 123 the function reaches zero and then oscillates about it. From this we can conclude that 103 is a good value for D_{max}, whereas 83 is underestimated and 123 is overestimated.

The P(r) function fits the measured scattering profile

This criterion is straightforward. The transformation of the P(r) function to I(q) should fit the measured scattering profile. This can be evaluated both through the χ^2 value of the fit, which should be close to 1, and the normalized residuals between the fit and the data, which should be flat and randomly distributed about zero.

The P(r) function goes to zero at r=0 and r= D_{max}

The reason for this criterion is straightforward. The P(r) function should go to zero at r=0 because it is the r^2 weighted number of electron pairs in the macromolecule. As r goes to zero, so does r^2, and thus so must P(r). The P(r) function should go to zero r= D_{max} because D_{max} is the maximum dimension of the particle. Beyond that distance there should be no electron pairs in the particle. This criterion is usually enforced by conditions in the IFT calculation.

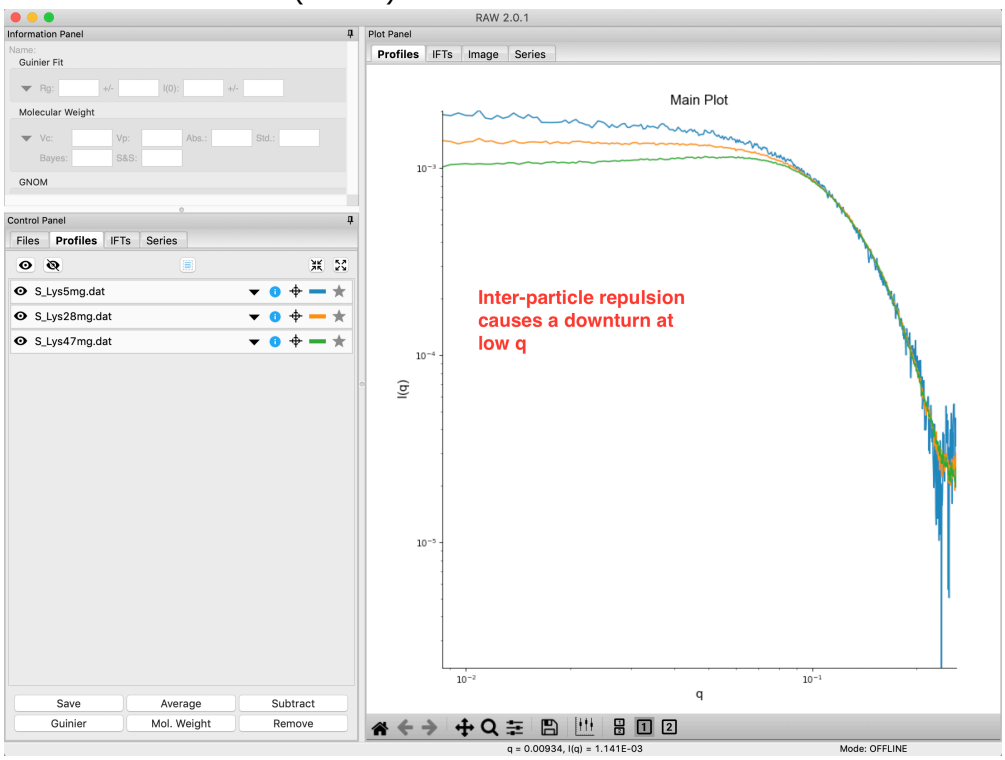
The R_g and I(0) from the Guinier fit and the P(r) function agree well
The R_g and I(0) values can be determined directly from the P(r) function. This provides a complementary approach to the Guinier fit. For well-behaved rigid systems, R_g and I(0) should agree well between both methods. If they do not, it may suggest a problem in either the Guinier fit or the P(r) function. However, for flexible and disordered systems, it has been observed that the P(r) R_g and I(0) values are characteristically larger, and more reliable, than the Guinier R_g and I(0) values.

The P(r) function is always positive
This criterion usually applies, as for most macromolecules the presence of a negative number of electron pairs has no meaning. However, when dealing with membrane proteins that are encapsulated in lipids or detergents, this criterion is no longer valid. In those cases, the lipid/detergent may have a lower electron density than the buffer. As scattering is measured relative to the solvent, this lower density will appear as negative electron pairs in the P(r) function. For example, proteins embedded in lipid nanodiscs have a characteristic dip in the P(r) function that can go negative.

Part 7. Concentration effects

Sample concentration can change the measured scattering profile. Typically, increasing the concentration leads to inter-particle interactions. These usually manifest as repulsion, and leads to a downturn at low *q*. This is why it is important to measure multiple concentrations of a given sample using SAXS, to make sure your data are unaffected by the concentration.

1. Clear any data you have loaded in RAW. You can do this by clicking the "Clear All" button in the Files tab. Load the three scattering profiles in the **Example_Data/concentration_data** folder. These are scattering profiles of lysozyme at 5, 28, and 47 mg/ml.
2. Examine the scattering profiles on linear, semi-log, and double log plots.
    - *Question:* What appears to be the effect of the concentration on the scattering profile? Over what *q* range does it alter the scattering profile?
    - *Question:* Does changing the concentration affect the folded-ness of this protein? How can you tell?
    - *Question:* Why does the lowest concentration profile have the most noise (they were all collected using the same total exposure time)?
    - *Try:* You can click on the magnifying glass in the plot control bar and click and drag it on a plot to zoom in a region. By clicking on the crossed arrows, you can click and drag on the plot to move scattering profiles relative to the plot. Right clicking and dragging with the crossed arrows selected will zoom you in or out. You can hit the home button (house) to return to the default view.



3. Find the R_g and I(0) using Guinier analysis from the three profiles.
    - *Question:* How does the concentration affect the R_g and I(0)?
    - *Try:* Adjust q_{min} and/or q_{max} of the Guinier region. Does the R_g remain constant over the Guinier region for any of the curves? What does that tell you about the reliability of the R_g and I(0) numbers from each curve?
4. Find the molecular weight for the three samples.
    - *Warning:* Because you haven't loaded an appropriate **SAXS.cfg** file, the molecular weight from I(0) reference will not give reliable results. As these curves are not on an absolute scale, the absolute scale calibrated molecular weight will also fail.
    - *Question:* How does the concentration affect the calculated MW?
    - *Question:* Does the MW depend on the *q* range you selected for the Guinier fit?
5. Find the P(r) function for the three samples.
    - *Question:* How does the repulsion change the Dmax? How about the shape of the P(r) function?
    - *Try:* Extend the Dmax past what you found to be the optimal value. What happens to the P(r) function?

Aside: Concentration effects

If you increase the concentration of the protein in solution enough, the macromolecules may start to repel each other due to electrostatic effects (this can also happen if you don't have enough salt in your buffer). This leads to the characteristic downturn at low q that you see in the lysozyme profiles above. This has the effect of artificially decreasing the R_g, I(0), MW, and D_{max} relative to the expected values with no inter-particle interactions.

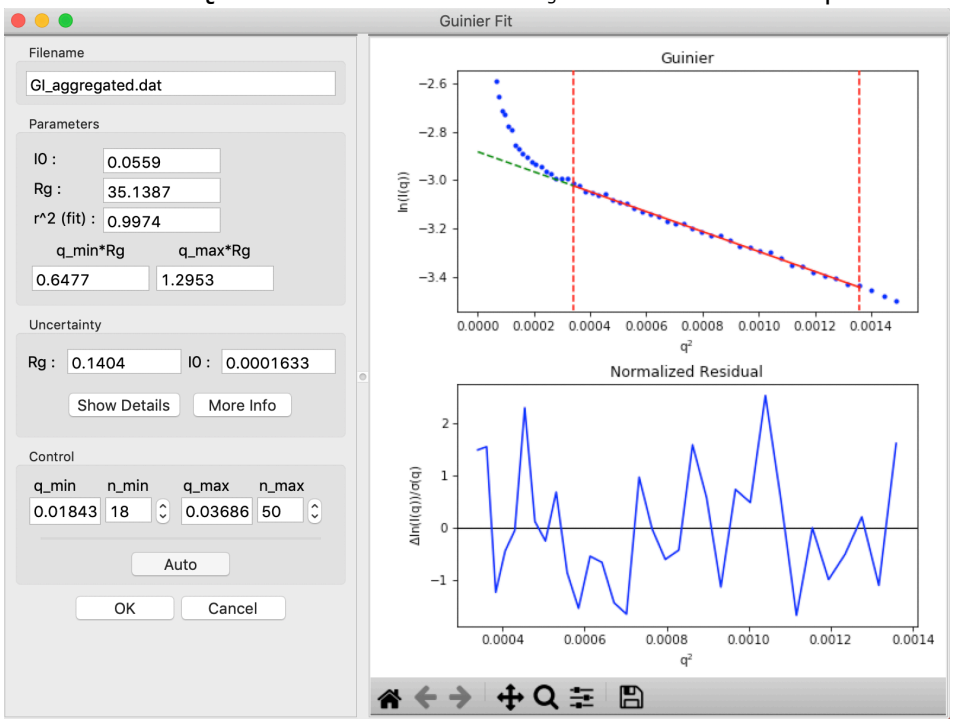
Often when collecting SAXS data you don't know what the expected values of these parameters is, so it can be hard to tell if they've affected. Learning the effect of this on the Guinier fit and P(r) function is thus particularly important. Repulsive interparticle interactions result in a structure factor that causes a downturn in the scattering profile at low q. These effects are clearly seen in the Guinier fit residual, where it shows up as a 'frown', with the residual below zero near the start and end of the fit and above zero in the middle.

For the P(r) function, the easiest way to see repulsion is to extend the D_{max} out past what you found to be a 'good' D_{max}. If the P(r) function goes and stays negative, then you have a repulsive interaction. If it stays near zero (possibly with some small oscillation about zero), then you are not seeing a repulsive interaction.

Part 8. Aggregation

A common problem in SAXS is the presence of a small amount of large aggregates in solution. These will alter the low-*q* scattering that is measured. Sometimes aggressive centrifuging can mitigate this. It is almost always a good idea to spin down your samples before measurement.

1. Clear any data that is open in RAW. Load the **S_Before_Centrifuge.dat** and **S_After_Centrifuge.dat** files from the **Example_Data/aggregation_data** folder. These are the same sample before and after aggressive centrifugation.
2. Look at the data on a variety of different plots.
   - *Question:* How did the scattering profile change with centrifugation?
   - *Question:* Which data set is more monodisperse? How can you tell?
3. Calculate the R_g and I(0) of each profile using a Guinier fit.
   - *Question:* How does the presence of aggregates affect the Guinier fit?
4. You should have found it essentially impossible to find a good Guinier region for the sample before it was centrifuged. Remember the requirements for a good Guinier fit: Evenly/randomly distributed residuals; $q_{min}R_g < 1.0$, $q_{max}R_g < 1.3$
   - *Try:* If you thought you found a good region, try again, keeping these requirements in mind.
5. Load the **GI_aggregated.dat** and **GI_unaggregated.dat** curves in the **aggregation_data** folder.
   - *Question:* How do these two GI curves differ?
6. Carry out Guinier analysis on both GI scattering profiles.
   - *Question:* For the aggregated curve, are you able to find a good Guinier region for the fit?
   - *Question:* How does the R_g for each curve compare?



7. Calculate the M.W. for each GI profile.
    - *Question:* How do they compare? Does it agree with what you would expect for aggregated vs. unaggregated samples?
8. Calculate the P(r) function for both profiles.
    - *Question:* How do they compare? Is there a difference in D_{max}?
9. In some cases, as with the aggregated GI scattering profile, you can work around small amounts of aggregation to obtain seemingly good Guinier regions. However, it is still best to retake your data and obtain aggregate free scattering profiles. Note that you will have obtained an R_g that was quite a bit larger (~35.3 Å compared to ~33.5 Å) from the GI scattering profile with aggregates. The M.W. and D_{max} from the aggregated profile are also larger than the unaggregated result. This indicates that despite there being some amount of a seemingly good Guinier region, the scattering profile is significantly influenced by the aggregation.

Aside: Aggregation

Aggregation causes a characteristic upturn at low q. This can either be caused by aggregates initially present in your sample, or by radiation induced aggregation (radiation damage). The effect of aggregation can also be clearly seen in the Guinier fit residual, where it shows up as a 'smile', with the residual above zero near the start and end of the fit and below zero in the middle.

For the P(r) function, the presence of these larger aggregate particles causes the D_{max} to be hard to determine. Typically this manifests as there being a significantly extended tail on the P(r) distribution, which does not fall to zero naturally regardless of the chosen D_{max}. The P(r) function calculated Rg and I(0) will also be larger than they would be for the monodisperse sample.

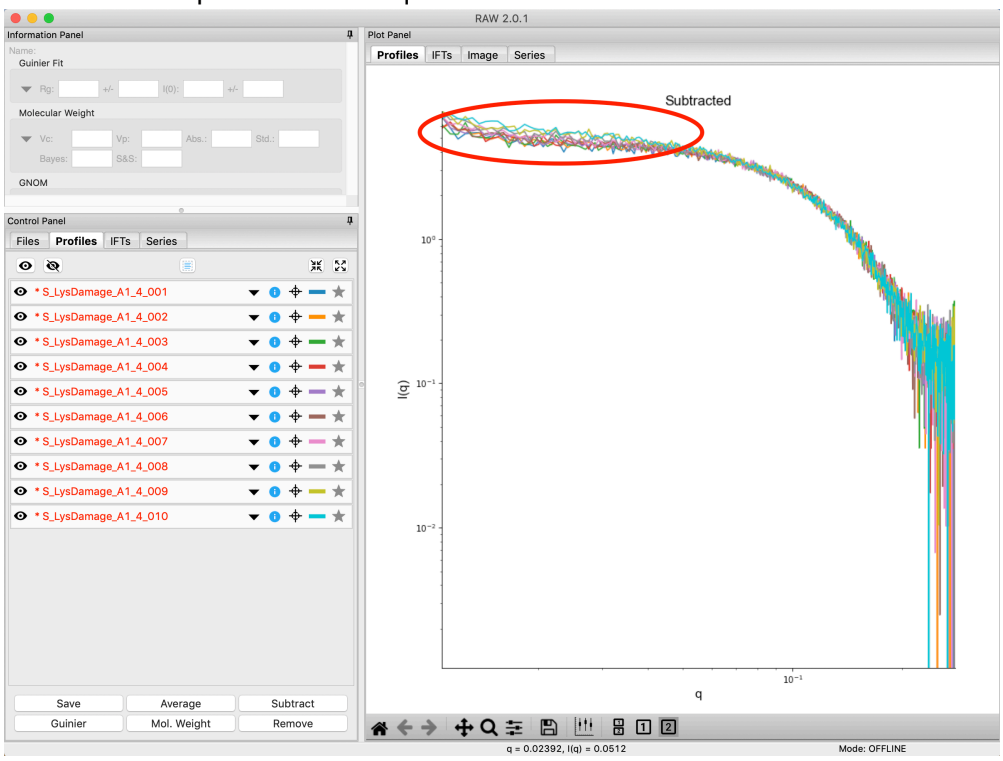
Small amounts of aggregation can look similar to the P(r) function for a flexible system, so other methods should be used to determine the true state of the system. Guinier analysis will usually reveal the presence of aggregates, and Kratky plot will show if the system is flexible. If you are unsure whether your P(r) function is showing aggregation or flexibility, check with these other techniques.

Part 9. Radiation damage

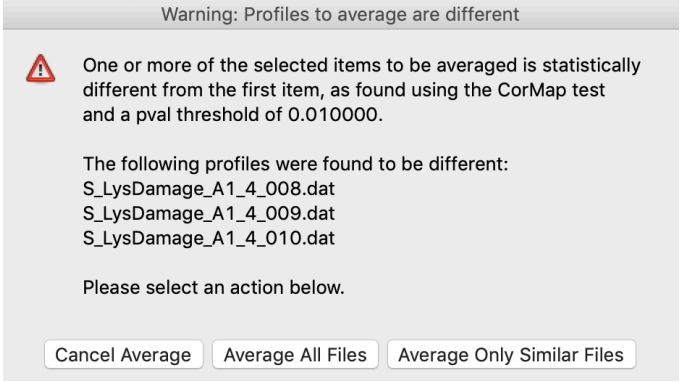
SAXS samples are sensitive to radiation damage from the X-rays. Radiation damage typically manifests as aggregates in solution, leading to the characteristic problems shown in Part 2. Controlling for radiation damage in your experiments is important to make sure that your data is not corrupted.

RAW has the ability to test scattering profiles for statistical similarity. Currently, only one test is available: The Correlation Map test. This can be done manually and is also done automatically when scattering profiles are averaged. This can be useful when you're dealing with data that may show changes in scattering from radiation damage or other possible sources.

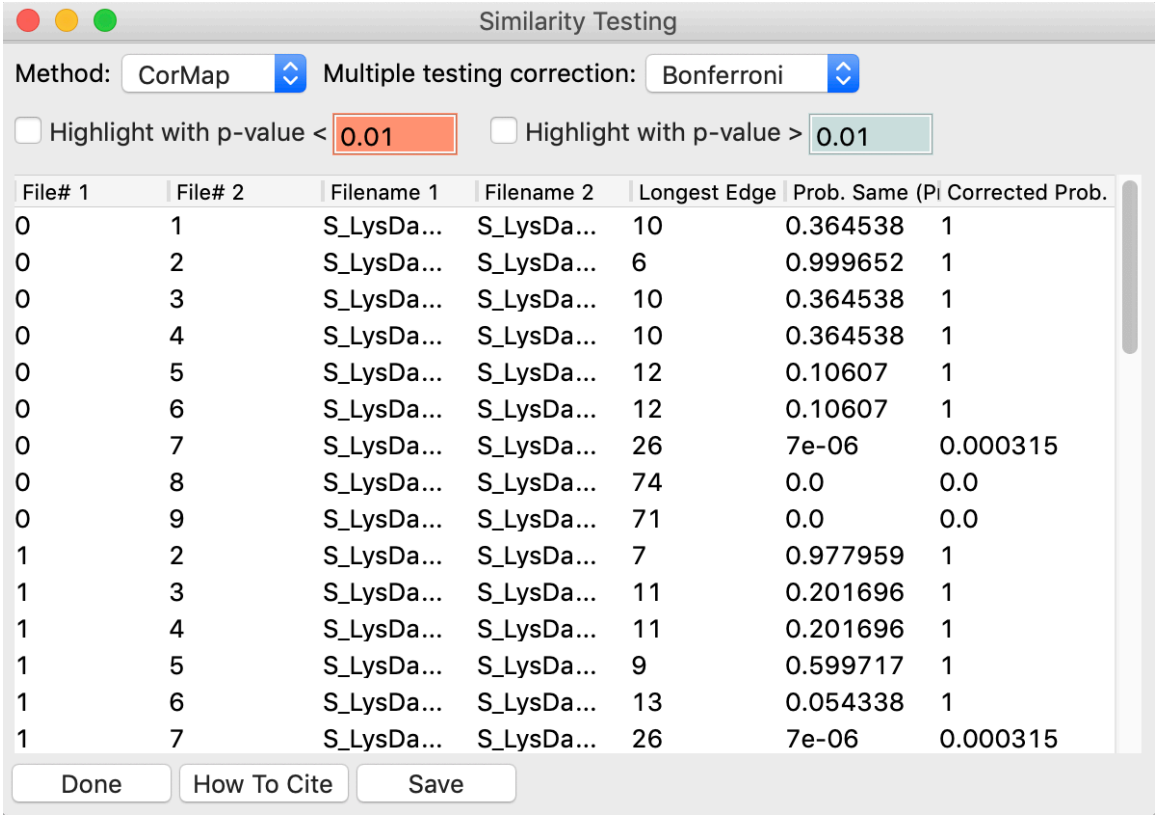
1. Clear any data that is open in RAW. Load all of the scattering profiles in the **Example_Data/damage_data** folder.
    - *Note:* This contains both unsubtracted lysozyme scattering profiles and lysozyme buffer scattering profiles.
2. Subtract the average buffer from each individual lysozyme scattering profile. Remove all the profiles except the subtracted profiles.
3. Put the Subtracted plot on a log-log scale. You should see that the profiles are different at low q.
    - *Note:* These data are showing what radiation damage typically looks like in a data set. They are consecutive profiles from the same sample, and as total exposure of the sample increase (frame number increases), the sample damages. In this case, the damage is manifesting as aggregation, which shows up as an uptick in the profiles at low q .



4. Select all of the subtracted profiles and average them. You will get a warning message informing you that not all the files are statistically the same.
    - *Note:* This is only as good as the statistical test being used, and the cutoff threshold selected. In the advanced options panel, you can select the test, whether or not it is corrected for multiple testing, and the threshold used.



5. Click the "Average Only Similar Files" button.
    - *Note:* This averages only those profiles found to the same as the first file, for the given statistical test.
    - *Tip:* Averages always end up on the top plot. You can move them to the bottom plot by right clicking on the profiles item and selection "Move to bottom plot".
6. Select all of the profiles except the new averaged one, and right click and select "Similarity Test".

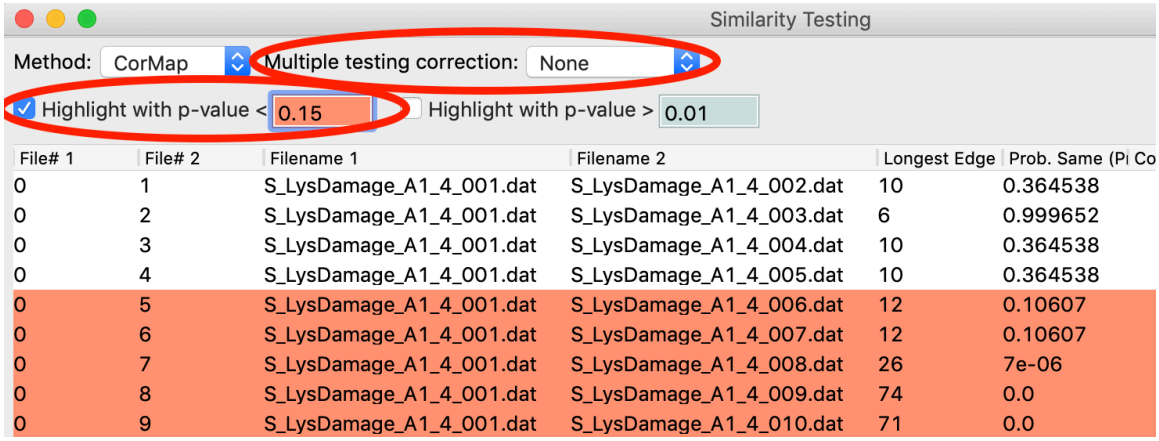


File# 1	File# 2	Filename 1	Filename 2	Longest Edge	Prob. Same (P	Corrected Prob.
0	1	S_LysDa...	S_LysDa...	10	0.364538	1
0	2	S_LysDa...	S_LysDa...	6	0.999652	1
0	3	S_LysDa...	S_LysDa...	10	0.364538	1
0	4	S_LysDa...	S_LysDa...	10	0.364538	1
0	5	S_LysDa...	S_LysDa...	12	0.10607	1
0	6	S_LysDa...	S_LysDa...	12	0.10607	1
0	7	S_LysDa...	S_LysDa...	26	7e-06	0.000315
0	8	S_LysDa...	S_LysDa...	74	0.0	0.0
0	9	S_LysDa...	S_LysDa...	71	0.0	0.0
1	2	S_LysDa...	S_LysDa...	7	0.977959	1
1	3	S_LysDa...	S_LysDa...	11	0.201696	1
1	4	S_LysDa...	S_LysDa...	11	0.201696	1
1	5	S_LysDa...	S_LysDa...	9	0.599717	1
1	6	S_LysDa...	S_LysDa...	13	0.054338	1
1	7	S_LysDa...	S_LysDa...	26	7e-06	0.000315

7. The similarity testing window (above) shows the results of the pairwise tests done using the CorMap method. Expand the window and the Filename columns to allow you to see the full filenames along with the probabilities.

File# 1	File# 2	Filename 1	Filename 2	Longest Edge	Prob. Same (Pi	Corrected Prob. Same
0	1	S_LysDamage_A1_4_001.dat	S_LysDamage_A1_4_002.dat	10	0.364538	1
0	2	S_LysDamage_A1_4_001.dat	S_LysDamage_A1_4_003.dat	6	0.999652	1
0	3	S_LysDamage_A1_4_001.dat	S_LysDamage_A1_4_004.dat	10	0.364538	1

8. Using the menu at the top, turn off multiple testing correction. Change the highlight less than value to 0.15, and highlight those pairs.



File# 1	File# 2	Filename 1	Filename 2	Longest Edge	Prob. Same (Pi	Co
0	1	S_LysDamage_A1_4_001.dat	S_LysDamage_A1_4_002.dat	10	0.364538	
0	2	S_LysDamage_A1_4_001.dat	S_LysDamage_A1_4_003.dat	6	0.999652	
0	3	S_LysDamage_A1_4_001.dat	S_LysDamage_A1_4_004.dat	10	0.364538	
0	4	S_LysDamage_A1_4_001.dat	S_LysDamage_A1_4_005.dat	10	0.364538	
0	5	S_LysDamage_A1_4_001.dat	S_LysDamage_A1_4_006.dat	12	0.10607	
0	6	S_LysDamage_A1_4_001.dat	S_LysDamage_A1_4_007.dat	12	0.10607	
0	7	S_LysDamage_A1_4_001.dat	S_LysDamage_A1_4_008.dat	26	7e-06	
0	8	S_LysDamage_A1_4_001.dat	S_LysDamage_A1_4_009.dat	74	0.0	
0	9	S_LysDamage_A1_4_001.dat	S_LysDamage_A1_4_010.dat	71	0.0	

9. Without multiple testing correction, and using a less stringent threshold for similarity, we see that more profiles are selected here (profiles 6-10) than were excluded from the average using the automatic test. Because we know radiation damage increases with dose (exposure), it is reasonable to suspect that we should discard profiles 6-10, not just 8-10 as in the automated version.
10. Save the similarity test data as a **.csv** by clicking the "Save" button.
11. Close the similarity testing window by clicking the "Done" button.
12. Average profiles 1-5, as determined by the similarity testing done above.
13. Using visual inspection of the subtracted scattering profiles, try to determine in what profile you first start to see radiation damage. Create an average of the undamaged curves.
    - *Tip:* As you will have observed earlier in the exercise, radiation damage typically manifests as an increase in the scattering profile at low *q*. It can help to hide everything but the first measured scattering profile, and then compare each subsequent profile to the first by itself. Also, it can help to look at different plot types.
    - *Tip:* Another useful method for looking for radiation damage is to subtract the first scattering profile from all the others and examine changes in these 'difference' profiles.
14. Find the R_g and I(0) by Guinier analysis of the first profile (**LysDamage_A1_4_001**), and the three averaged subtracted profiles that you have made.
15. *Question:* What do you observe about the R_g? How about the *q* range used in the Guinier fit? About the residual curves?

- *Note:* In this case, the differences are subtle, a ~1-2% increase in R_g. So the automated determination did a reasonable job. However, it is generally good to double check your set of profiles both visually and using the Similarity Test panel when the automated test warns you of outlier profiles.

Part 10. A Few Additional Tricks

Here are some additional tricks that may make your life easier while using RAW:

1. If you click on a scattering profile in the Profiles plot, the corresponding profiles list item will be highlighted.
2. You can save the workspace by going to File->Save Workspace. This will save all of the scattering profiles, IFT curves, and Series curves. These will all be loaded again when you load the workspace.
    - *Note:* This does not save the settings!
3. If you have the crossed arrows selected in the plot control bar to drag a plot, right clicking and dragging allows you to zoom a plot.
4. You can turn error bars on and off for scattering profiles using the error bar button in the plot control (to the right of the save button).
5. You can rename a curve by right clicking on the appropriate entry in the list and choosing rename.
6. You can view the history of a scattering profile by right clicking on it and selecting Show History. For a curve that has been processed from an image, this will show you processing parameters such as normalization and any corrections applied to the scattering intensity. For a curve that is processed (such as an averaged of subtracted curve) it will show you the steps used to make that curve. For example, for an averaged curve, it will show you all of the files that were averaged.
7. All of the main RAW plots read out the mouse coordinates at the bottom of the RAW window, just below all of the tool buttons.