BioXTAS RAW Documentation

Release 1.6.2

Jesse B. Hopkins

Oct 28, 2019

Contents

1 What is BioXTAS RAW? 1

2 Features 2

3 History and Usage 3

4 Licensing 4

4.1 Installation 4
4.2 Getting Help 15
4.3 Tutorial 15
4.4 Videos 116
4.5 Cite 116
4.6 Manual 118
4.7 Changes 191

Index 222

CHAPTER 1

What is BioXTAS RAW?

BioXTAS RAW is a GUI based, free, open-source Python program for reduction and analysis of small-angle X-ray solution scattering (SAXS) data. The software is designed for biological SAXS data. It is available on windows, macOS (and OS X), and linux. It provides an alternative to closed source programs such as Primus and Scatter for primary data analysis. Because it can calibrate, mask, and integrate images it also provides an alternative to synchrotron beamline pipelines that scientists can install on their own computers and use both at home and at the beamline.

Find out how to get it!

CHAPTER 2

Features

- Analysis of radius of gyration (R_g) and I(0) via Guinier fit.
- Analysis of molecular weight via I(0) comparison to standards, absolute calibration, correlation volume (V_c) and corrected Porod volume (V_p) methods.
- Calculation of inverse Fourier transforms (IFTs) via GNOM and a Bayesian indirect Fourier transform (BIFT).
- Calculation of envelopes (dummy atom models) using DAMMIF, DAMMIN, DAMAVER, and DAMCLUST.
- Calculation of electron density using the DENSS algorithm
- Easy processing of in-line chromatography coupled SAXS data, including size-exclusions coupled SAXS (SEC-SAXS) data.
- Deconvolution of SEC-SAXS data using singular value decomposition (SVD) and evolving factor analysis (EFA).
- Standard data operations such as averaging, subtraction, merging, and rebinning.
- Creation and plotting of 1D scattering profiles from 2D detector images, including Pilatus, CBF, Eiger, and more than 20 other types of images.

CHAPTER 3

History and Usage

RAW was first developed in 2008 by Soren Skou as part of the biological x-ray total analysis system (BioXTAS) project. Since then it has been extensively developed, with recent work being done by Jesse Hopkins.

RAW is actively used as the primary analysis software at the MacCHESS G1 BioSAXS Beamline. and the BioCAT (18-ID) beamline at the APS. It is also used at various other beamlines, including:

- BL19U2 (SSRF)
- I911-4 (aka I911-SAXS) beamline at MAX-LAB

SAXSLAB distributes RAW with some of its homesources, and RAW is used at various other homesources around the world.

Do you use RAW? *Let us know!*

CHAPTER 4

Licensing

RAW source code is released under a GPLv3 license. Both the source code and the prebuilt versions of RAW are free for anyone to download and use.

4.1 Installation

RAW can be downloaded from sourceforge. There are prebuilt installers for windows (.msi) and mac (.dmg). Installation on Linux requires building from source.

RAW is also available through SBGrid.

Some features in RAW require a separate installation of the ATSAS package.

4.1.1 Detailed installation instructions

RAW Install Guide for Microsoft Windows

There are two ways to install RAW on Windows. The easiest, and recommended, way is to use a prebuilt installer (`.msi` file). If you want to install a version for which there is no prebuilt installer, you will need to install RAW from the source code.

Here you can find information on:

Using a prebuilt installer

The recommend way to install RAW on Windows is using a prebuilt installer. To install from a prebuilt installer, simply download the `RAW-x.y.z-win32.msi` (where `x.y.z` is the version number) file from sourceforge (http://sourceforge.net/projects/bioxtasraw), and double click it to run the installer.

Important Notes:

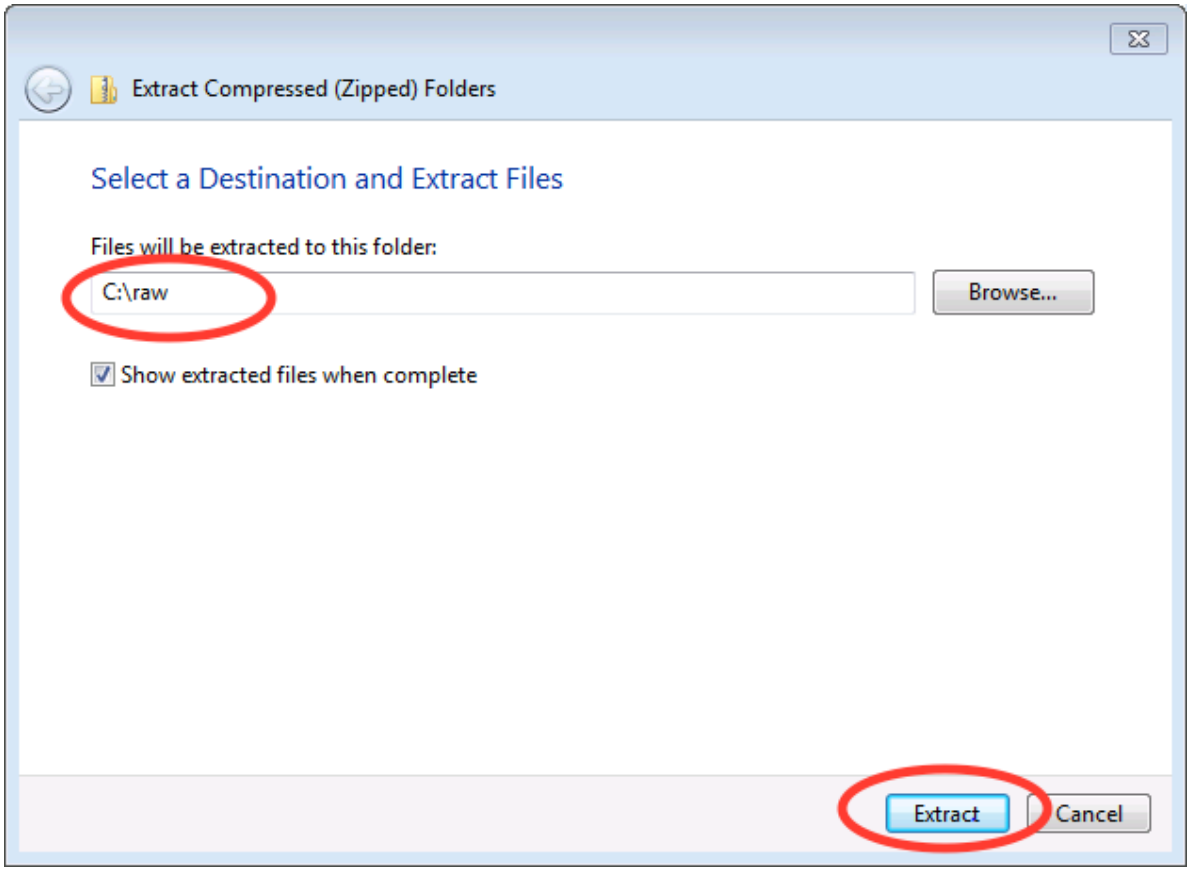
- Because the RAW team is an 'unknown' developer, you will probably see some security warnings when you install RAW. When you do, just give the installer and program permission to run on your computer.
- RAW is used by (relatively) few people, which means many virus scanners have not seen the RAW software before. Occasionally virus scanners will mark a file (typically `RAW.exe`) as a threat (it will usually be in the 'general malware' category). If this happens, please do the following:
  - Upload the file to https://www.virustotal.com/ and see if any other virus scanners identify it as a virus (it is always possible someone hijacked the installer somehow!).
  - If most or all of the virus scanners on virustotal.com clear the file, make an exception for it in your virus scanner.
  - Contact the RAW developers, so we can report the false identification to the virus scanner company and get the file whitelisted in future definitions files.

General instructions for installing from source (advanced users)

1. Install Microsoft Visual C++ 2008 Redistributable and Visual C++ Compiler for Python.
2. Install Python 2.7 (if it isn't already installed) and add it to your system path.
3. Install the following Python packages (version indicated if less than most recent):
   - numpy
   - scipy
   - matplotlib
   - pillow
   - wxpython
   - lxml
   - h5py
   - cython
   - fabio
   - pyFAI
   - hdf5plugin
   - numba
4. Download the RAW source file (`RAW-x.y.z-Source` where `x.y.z` is the version number) from sourceforge (http://sourceforge.net/projects/bioxtasraw)
5. Extract RAW to a directory of your choice.
6. In the top level RAW directory run `python setup.py build_ext --inplace` to build the extensions.
7. In the `bioxtasraw` subdirectory run `RAW.py` using python.
8. Enjoy!
   - If you have problems, please consult the detailed installation guides and the *solutions to common problems* below. If that doesn't help, please contact the developers.

Windows 7, 8.1, and 10 install from source instructions

1. RAW on windows can be installed using 64 bit (x64) or 32 bit (x86) python. Unless you know you need a 32 bit build, you should install the 64 bit version.
2. Download and install the Microsoft Visual C++ 2008 Redistributable Package.
   - Download the 64 bit version from: https://www.microsoft.com/en-us/download/details.aspx?id=15336
   - Download the 32 bit version from: https://www.microsoft.com/en-us/download/details.aspx?id=29
   - Run the installer with the default options.
3. Download and install the Microsoft Visual C++ Compiler for Python.
   - Download from: https://www.microsoft.com/en-us/download/details.aspx?id=44266
   - Run the installer with the default options.
4. Install Miniconda python distribution
   - Download the free miniconda python 2.7 installer from: https://conda.io/miniconda.html
     - Make sure you get the python 2.7 version!
     - Pick the appropriate 64 bit/32 bit version that matches the redistributable you downloaded in the previous step.
   - Run the installer with the default options.
   - More detailed install instructions are available here: https://conda.io/docs/user-guide/install/windows.html
5. Install the necessary python packages
   - Open an anaconda prompt by clicking on the start menu -> All Programs -> Anaconda2 -> Anaconda Prompt
   - Run the following commands in the anaconda prompt:
   - `conda upgrade conda pip wheel setuptools`
   - `conda install numpy scipy matplotlib pillow wxpython numba h5py lxml cython numexpr`
   - `pip install hdf5plugin silx fabio pyfai`
6. Download RAW from sourceforge (http://sourceforge.net/projects/bioxtasraw)
   - Go to the Files tab on the linked website and download the `RAW-x.y.z-Source.zip` file, where *x.y.z* is the version number (for example, 1.0.0).
7. Expand the downloaded zip file into the downloads folder
   - Right click on the download and select *Extract All*
   - Accept the default location for files to be extracted.



8. In Windows Explorer, confirm that the file named `setup.py` is in your top level expanded raw directory. If it isn't, it's likely that when you expanded the RAW download, you ended up with unnecessary layers of directories. Find the directory with `setup.py` in it, and make that the top level folder.
9. Build the extensions
   - Open an anaconda prompt as in Step 4 of these instructions.
   - Type `cd C:\raw`
   - Hit enter
   - Type `python setup.py build_ext --inplace`
   - Hit enter
10. Run `RAW.py` from the command line
    - Open an anaconda prompt as in Step 4 of these instructions.
    - Type `cd C:\raw\bioxtasraw`
    - Hit enter
    - Type `python RAW.py`
    - Hit enter
11. Enjoy!
    - If you have trouble with the installation, please see the *solutions to common problems* section below.

Common problems/troubleshooting

Prebuilt installer:

- Because the RAW team is an 'unknown' developer, you will probably see some security warnings when you install RAW. When you do, just give the installer and program permission to run on your computer.
- RAW is used by (relatively) few people, which means many antivirus programs have not seen the RAW software before. Occasionally virus scanners will mark a file (typically `RAW.exe`) as a threat (it will usually be in the 'general malware' category). If this happens, please do the following:
  - Upload the file to https://www.virustotal.com/ and see if any other antivirus programs identify it as a problem (it is always possible someone hijacked the installer somehow!).
  - If most or all of the antivirus programs on virustotal.com clear the file, make an exception for it in your virus scanner.
  - Contact the RAW developers, so we can report the false identification to the virus scanner company and get the file whitelisted in future definitions files.

From source:

- If you are updating your RAW installation, you should completely delete the old RAW source files, and then replace them with the new ones.

RAW Install Guide for macOS and OS X

There are two ways to install RAW on Mac. The easiest, and recommended, way is to use a prebuilt app package (`.dmg` file). If you want to install a version for which there is no prebuilt installer, you will need to install RAW from the source code.

Here you can find information on:

Using a prebuilt app package

The recommended way to install RAW on Mac is using a prebuilt app package. To install from a prebuilt app package simply download the `RAW-x.y.z-mac.dmg` (where `x.y.z` is the version number) file from sourceforge (http://sourceforge.net/projects/bioxtasraw), double click it to open the dmg, and drag the `RAW.app` file to your Applications folder (or wherever you want to install RAW).

Important Notes:

- Because the RAW team is an unidentified developer, you may get a warning message the first time you run the program. If that happens, right click on RAW and select *Open* from the right click menu, and then click the *Open* button in the window that appears.
  - *Note:* This requires administrator permissions.

General instructions for installing from source (advanced users)

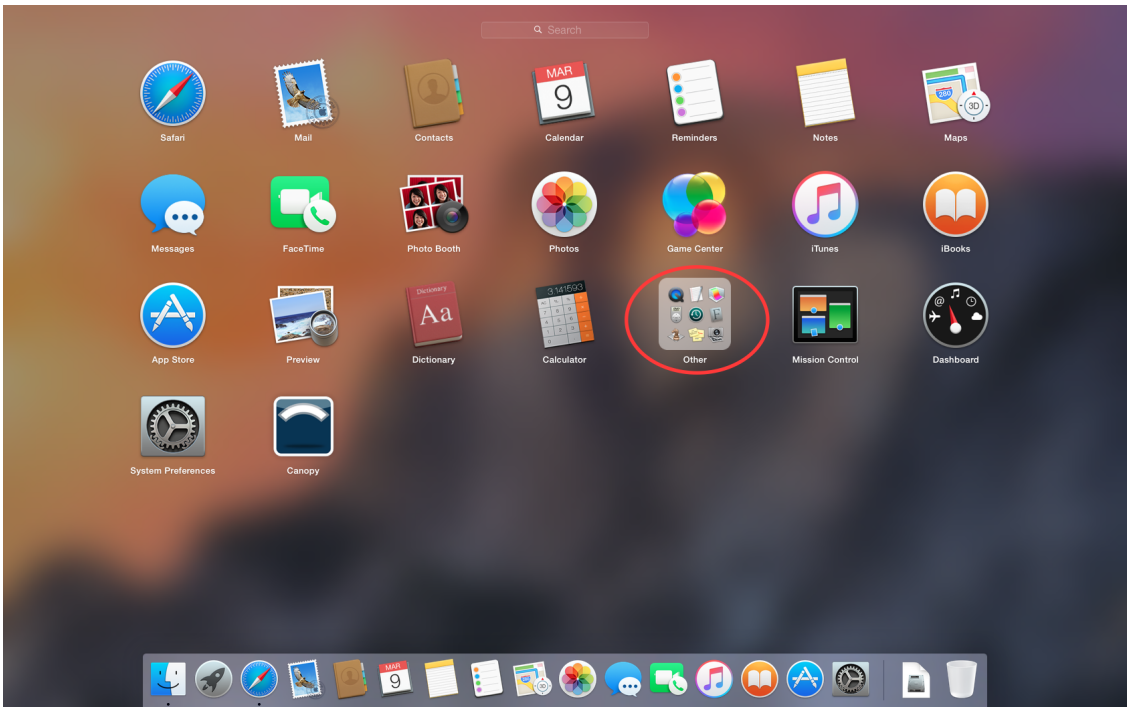
1. Install a standalone version of python 2.7 (recommended, not required).
2. Install the following python packages (most recent version of each recommended):
   - numpy
   - scipy

- matplotlib
- pillow
- wxpython
- lxml
- h5py
- cython
- fabio
- pyFAI
- hdf5plugin
- numba

3. Download the latest RAW sourcecode from sourceforge (http://sourceforge.net/projects/bioxtasraw)
4. Extract RAW to a directory of your choice.
5. In the top level RAW directory run `python setup.py build_ext --inplace` to build the extensions.
6. In the `bioxtasraw` subdirectory run `RAW.py` using python.
7. Enjoy!
   - If you have problems, please consult the detailed installation guide and the *solutions to common problems* below. If that doesn't help, please contact the developers.

OS X and macOS install from source instructions

1. Install Miniconda python distribution
   - Download the free miniconda python 2.7 installer from: https://conda.io/miniconda.html
     - Make sure you get the python 2.7 version!
   - Open a Terminal window by starting the Terminal app.
     - Using the Launcher, it is located in the Other section.



  - Browsing the Applications folder in Finder, Terminal is in the Utilities folder.
- In the terminal window type `cd ~\Downloads` and hit enter.
- In the terminal window type `bash Miniconda2-latest-MacOSX-x86_64.sh` and hit enter.
- Agree to all of the prompts.
- More detailed install instructions are available here: https://conda.io/docs/user-guide/install/macos.html
- Close the terminal window.

2. Install numpy scipy matplotlib pillow wxpython numba h5py lxml cython numexpr using conda.

- Open a new terminal window as in the previous step
- Type `conda upgrade conda pip wheel setuptools` and hit enter. Agree to all the prompts.
- Type `conda install numpy scipy matplotlib pillow wxpython numba h5py lxml cython numexpr` and hit enter. Agree to all the prompts.

3. Install the hdf5plugin, silx, fabio, and pyFAI packages using pip
   - In the terminal window type `pip install hdf5plugin silx fabio pyfai`
   - Hit enter
   - Once the installation finishes, close the terminal window.
4. Download RAW from sourceforge
   - http://sourceforge.net/projects/bioxtasraw
   - Navigate to the *Files* tab and download the latest source code, `RAW-x.y.z-Source.zip`. Or download the latest development version from the git by navigating to the *Code* tab.
5. Expand the downloaded zip file in the Downloads folder by double clicking on it.
   - This step may not be necessary, some browsers may automatically expand zip files.
6. In the terminal or in the graphical file manager, confirm that the file named `setup.py` is in your expanded raw directory. If it isn't, it's likely that when you expanded the RAW download, you ended up with unnecessary layers of directories. Find the directory with `setup.py` in it, and make that the top level folder.
7. Move the RAW files to Applications folder
   - Move the folder that contains all of the RAW files to the `Applications` folder. As above, this would be the folder with `setup.py` in it.
   - Rename the folder that you just moved to `raw`.
8. In a terminal, change directory into the top level RAW folder
   - If you used the suggested path of `Applications/raw` type: `cd /Applications/raw`
9. Build the extensions.
   - `python setup.py build_ext --inplace`
10. Navigate to the `bioxtasraw` subfolder
    - From the top level RAW folder it should be `cd ./bioxtasraw`
11. Run RAW
    - `pythonw RAW.py`
12. Enjoy!
    - In the future, you can start RAW as in the previous step.
    - If RAW doesn't work, check out the *solutions to common problems*

Common problems/troubleshooting

Installing the prebuilt app package:

- Because the RAW team is an unidentified developer, you may get a warning message the first time you run the program. If that happens, right click on RAW and select *Open* from the right click menu, and then click the *Open* button in the window that appears.

  - This requires administrator privileges

- If the above doesn't work, you can run the RAW app from the command line. Navigate to `RAW.app/Contents/MacOS` and run the RAW unix executable file (`./RAW`) in that directory.

Installing from source:

- If you fail to build the extensions before running RAW, RAW will crash at some point. Be sure to run the `python setup.py build_ext --inplace` before starting RAW.
- Using the RAW autorg function requires a relatively recent version of numba. If you get an error running this function update your numba version to the most recent.

RAW Install Guide for Linux

Here you can find information on:

General instructions for installing from source (advanced users)

1. Install python 2.7 (if it isn't already installed).
2. Install the following python packages (version indicated if less than most recent):
   - numpy
   - scipy
   - matplotlib
   - pillow
   - wxpython
   - lxml
   - h5py
   - cython
   - fabio
   - pyFAI < 0.18
   - hdf5plugin
   - numba
3. Download RAW from sourceforge (http://sourceforge.net/projects/bioxtasraw)
4. Extract RAW to a directory of your choice.
5. In the top level RAW directory run `python setup.py build_ext --inplace` to build the extensions.
6. In the `bioxtasraw` subdirectory run `RAW.py` using python.
7. Enjoy!
   - If you have problems, please consult the detailed installation guides below and check out the *solutions to common problems*.
   - If you want, see the section on *making a desktop shortcut for RAW*.

Linux install from source instructions

Tested on: Ubuntu 16.04, Linux Mint 18.3, Debian 9.4, Scientific Linux 6.8 7.4 (which should be equivalent to Redhat and CentOS), and OpenSUSE Leap 42.3.

1. Open a new terminal window (in many distros you can right click on the desktop and select *New Terminal* or *Open in terminal*).
2. Download Miniconda Python (Python 2.7) distribution and install.
   - http://conda.pydata.org/miniconda.html
   - Make sure you chose the python 2.7 installer.
   - Save to the downloads folder
   - Open a terminal and run the following commands:
     - `cd ~/Downloads`
     - `bash ./Miniconda2-version.sh` (where `-version` which will depend on the version you download)
   - Accept the default installation location.
   - At the end, say "yes" to have the conda python install put in your system path.
   - Close the terminal window.
3. Install python packages. Open a new terminal window and run the following commands (in many distros you can right click on the desktop and select *New Terminal* or *Open in terminal*).
   - `conda upgrade conda pip wheel setuptools`
   - `conda install numpy scipy matplotlib pillow wxpython numba h5py lxml cython numexpr`
   - `pip install silx fabio "pyFAI<0.18" hdf5plugin`
4. Download RAW from sourceforge
   - http://sourceforge.net/projects/bioxtasraw/
   - The download button on the main page should default to the right download for your OS.
5. Expand the RAW download to your location of choice.
   - We suggest `~/raw`
   - Make sure there are no spaces in the file path (you can check by navigating to the raw directory in a terminal window and using `pwd`).
   - In the terminal or in the graphical file manager, confirm that the file named `setup.py` is in your raw directory. If it isn't, it's likely that when you expanded the RAW download, you ended up with unnecessary layers of directories. Find the directory with `setup.py` in it, and make that the top level folder.
6. In a terminal, change directory into the top level RAW folder
   - If you used the suggested path of `~/raw` type: `cd ~/raw`
7. Build the extensions.
   - `python setup.py build_ext --inplace`
8. Navigate to the `bioxtasraw` subfolder
   - From the top level RAW folder it should be `cd ./bioxtasraw`

9. Run RAW
   - `python RAW.py`
10. RAW is now installed. Enjoy!
    - If you want, see the section on *making a desktop shortcut for RAW*.
    - If RAW doesn't work, check out the *solutions to common problems*.

Instructions for setting up a RAW desktop shortcut

All files referred to are initially located in the RAW `LinuxLib` folder.

1. Add the `start_raw` file to your path:
   - `sudo cp ~/raw/LinuxLib/start_raw /usr/local/bin`
2. Make the `start_raw` file executable:
   - `sudo chmod +x /usr/local/bin/start_raw`
3. Copy the `RAW.desktop` file to the desktop:
   - `cp ~/raw/LinuxLib/RAW.desktop ~/Desktop/`
4. Right click on the `RAW` file on the desktop, and select *Properties*
5. Click on the Permissions tab, and make sure *Allow executing file as program* is checked.
6. Note: depending your distribution/shell, you have to edit the `start_raw` file to use a different shell. By default it uses bash.

Common problems/troubleshooting

- Sometimes, compilers can have trouble if there are spaces in the filepath. Try installing RAW so that there are no spaces in the file path (navigate the folder in the terminal, type `pwd` and see what the result is).
- If you have installed a standalone python distribution (such as Enthought Canopy or miniconda/anaconda), it is possible that it isn't set to default, so when you run `python RAW.py`, you are using the wrong python.
  - You can verify which python you are using the command `which python` in the terminal.
  - You can set the correct python to default by modifying your appropriate profile file (such as the .bash_profile), or setting the `$PATH` environmental variable.
  - You can also specify the full path to the version of python you want to use in the command, such as `~/miniconda2/bin/python`
- Note that when you change environmental variables in one terminal window, you need to restart other windows for this to take effect. If you aren't using the right python (or compiler, etc), trying closing all of your terminal windows and opening a new one.
- If you fail to build the extensions before running RAW, RAW will crash at some point. Be sure to run the `python setup.py build_ext --inplace` before starting RAW.
- Using the RAW autorg function requires a relatively recent version of numba. If you get an error running this function update your numba version to the most recent.
- On Scientific Linux 6, and thus probably on Red Hat 6 (untested), RAW completely fails to work with wxpython 3.0 and certain python distributions (namely the Enthought python).

  - If RAW completely doesn't start, check and make sure you have wxpython 2.8 installed. This requires that you have matplotlib<=1.4.
  - In general, RAW should work with wxpython 3 and 4.
- As for the last time we tested it (2017) the Enthought Canopy python package DOESN'T WORK on Ubuntu or Linux Mint with wxpython.

4.2 Getting Help

4.2.1 General Questions/Comments

For questions on how to use the software, first check out the documentation on this site.

If that doesn't answer your question, or for other comments, please contact the RAW mailing list. If that doesn't work for you, you can contact Soren and Jesse directly: soren (dot) skou (at) saxslab (dot) com and hopkins (dot) jesse (at) gmail (dot) com.

4.2.2 Reporting A Bug

We appreciate any and all bug reports, it helps us improve the software! If you have a bug to report, the best way is to contact the RAW mailing list. If that doesn't work for you, you can contact Soren and Jesse directly: soren (dot) skou (at) saxslab (dot) com and hopkins (dot) jesse (at) gmail (dot) com.

If you have a bug report, please include: operating system, RAW version number, installation type (package or from source), what you were doing when the bug occurred, the appropriate error message(s) and sample data so we can recreate the bug. That will enable us to find and solve the problem as quickly as possible.

4.3 Tutorial

4.3.1 Introduction

Overview

This tutorial covers SAXS data processing with RAW. You will learn how 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 and normalized Kratky analysis
- Test the similarity of scattering profiles
- Load and process SEC-SAXS data
- Carry out singular value decomposition (SVD) and evolving factor analysis (EFA) to evaluate and analyze SEC-SAXS data
- Do baseline correction on SEC-SAXS data
- Merge SAXS/WAXS data from two detectors

- Carry out Pair-distance distribution analysis (BIFT and GNOM)
- Evaluate ambiguity of 3D shape reconstructions (AMBIMETER)
- 3D reconstructions with bead models (DAMMIF/N and DAMAVER)
- 3D reconstructions with electron density (DENSS)
- Calibrate RAW for integrating images
- Mask images for integration
- Set up normalization and save processing settings
- Set absolute scaling in RAW using water and glassy carbon
- Set a molecular weight standard in RAW

Section 1 covers basic processing with RAW, and *Section 2* covers advanced processing with RAW. *Section 3* covers how to set up RAW for integrating images for those who do not already have a configuration file.

Requirements

- BioXTAS RAW >= v1.6.0 (most recent is best).
  - *Install instructions*
- Tutorial data.
  - Available from: https://sourceforge.net/projects/bioxtasraw/files/?source=navbar
- ATSAS programs, >= v2.8.0 (for parts of *Section 2* of the tutorial only).
  - Download and installation 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.

Other useful materials

1. An overview and tutorial of RAW produced by Jesse Hopkins for SBGrid, which can be viewed here: https://youtu.be/XGnJDs3N2MI
2. There are RAW tutorial videos produced by Richard Gillilan, which can be viewed here: http://bit.ly/bioxtast. Data for these tutorial videos is available here: http://bit.ly/bioxtasd.
3. ATSAS resources:
   - Manuals: http://www.embl-hamburg.de/biosaxs/manuals/
   - User forum: http://www.saxier.org/forum/
4. Electron density (DENSS) resources available at DENSS.org
   - Particularly useful is the section on visualizing the results and aligning with known structures.

Notes

If you are only interested in using RAW to process data, and are not interested in how to set up RAW to calibrate your data, you do not need to look at *Section 3*.

RAW depends on user feedback to get better. If you have questions, find bugs, or think a part of this tutorial is unclear, please let the developers know. The best way to do this is via the RAW google group: http://bit.ly/rawhelp

You can find additional developer contact information on the RAW website: https://sourceforge.net/projects/bioxtasraw/

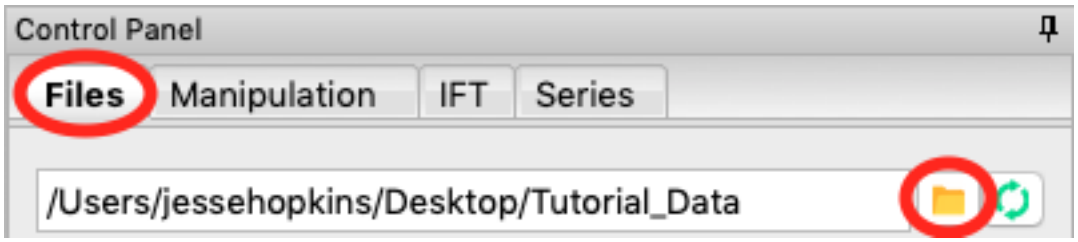
4.3.2 Basic processing

This section will guide you through basic processing of files in RAW. It includes: integrating images into scattering profiles, averaging and subtracting scattering profiles, doing Guinier fits, MW analysis, processing SEC-SAXS data, processing data from 2 detectors (SAXS and WAXS data), and a few other things. It refers to the *RAW tutorial data*.

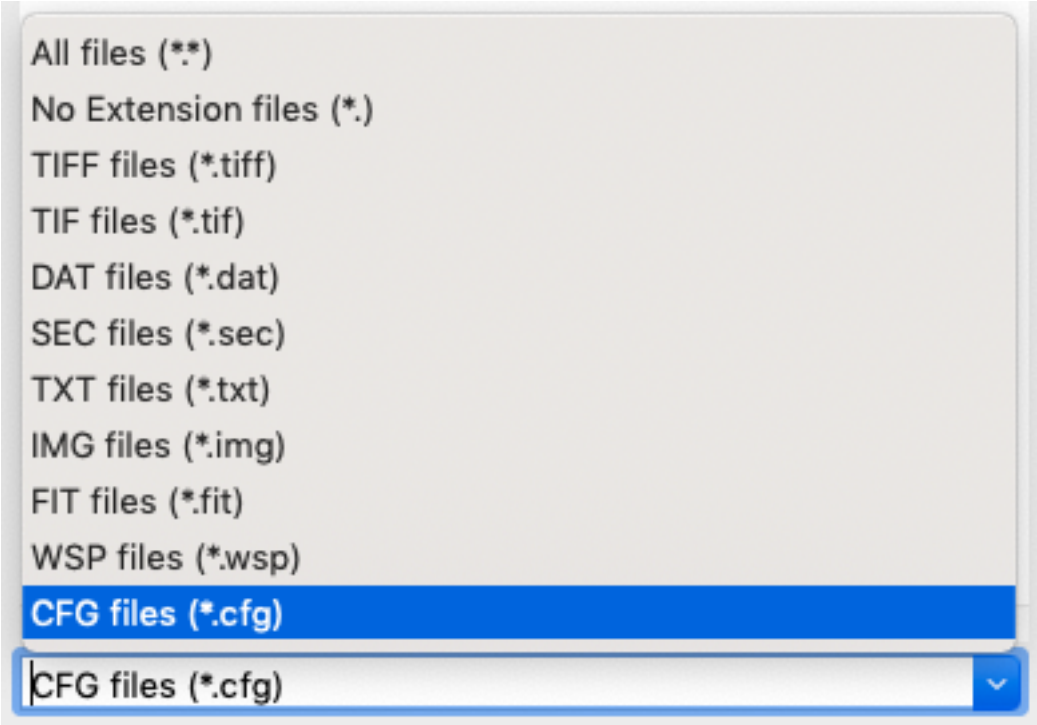
Select a section below to view the tutorial, or use the next and back buttons at the bottom of the page to navigate through it in order.

Loading configuration files and images, creating subtracted scattering profiles, saving profiles

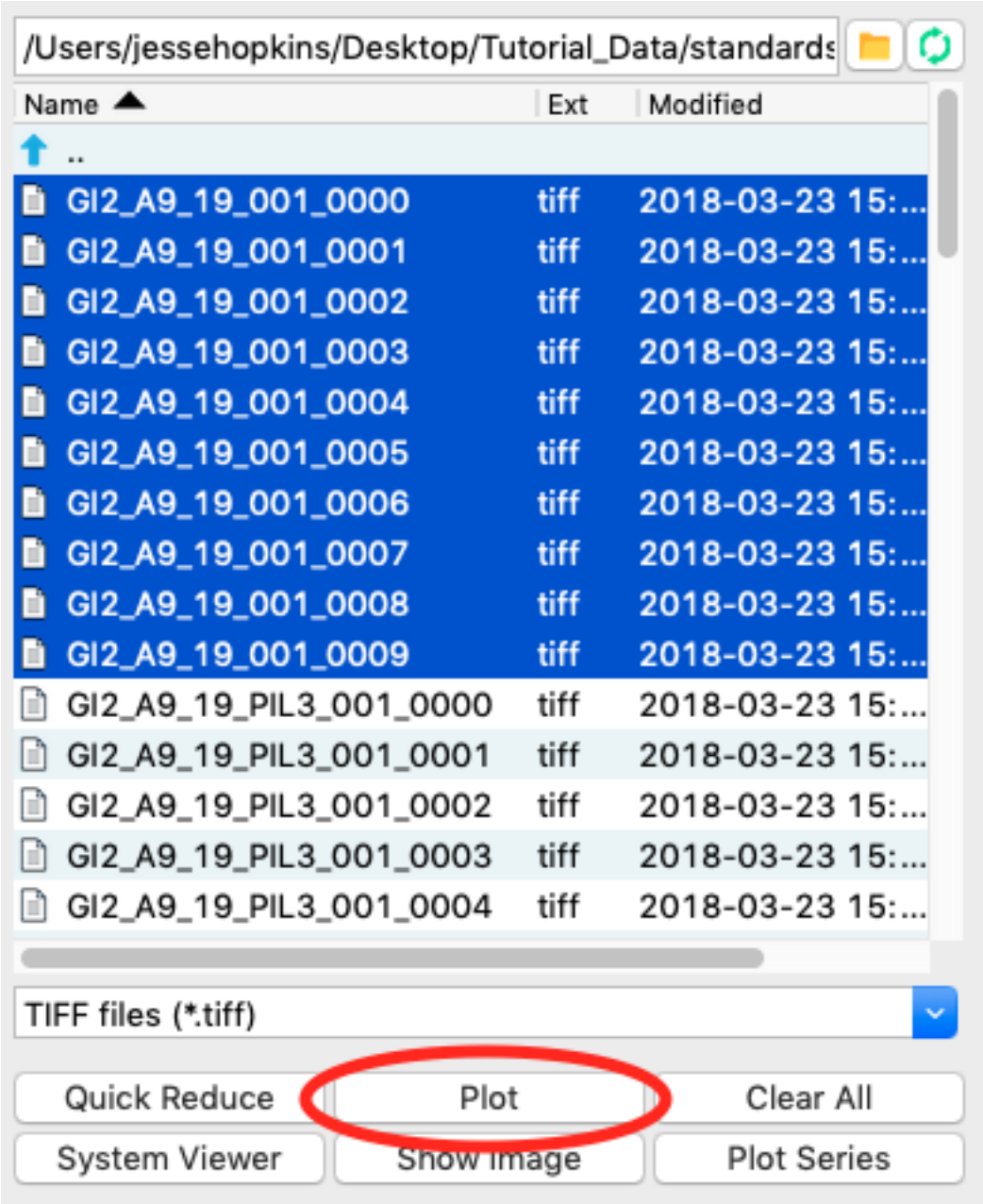
1. Open RAW. The install instructions contain information on installing and running RAW.
2. In the files tab, click on the folder button and navigate to the **Tutorial_Data/standards_data** folder. Click the open button to show that folder in the RAW file browser.



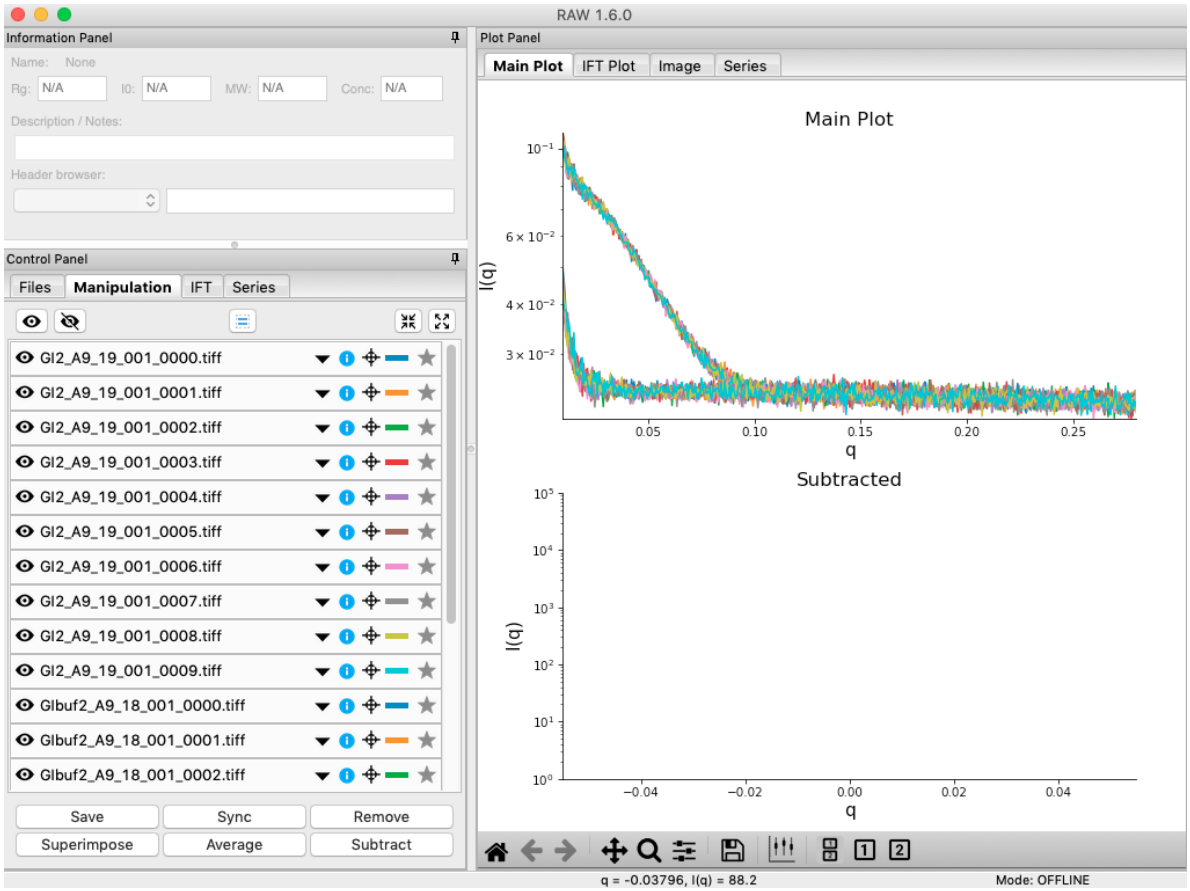
3. At the bottom of the File tab in the Control Panel, use the dropdown menu to set the file type filter to "CFG files (*.cfg)".



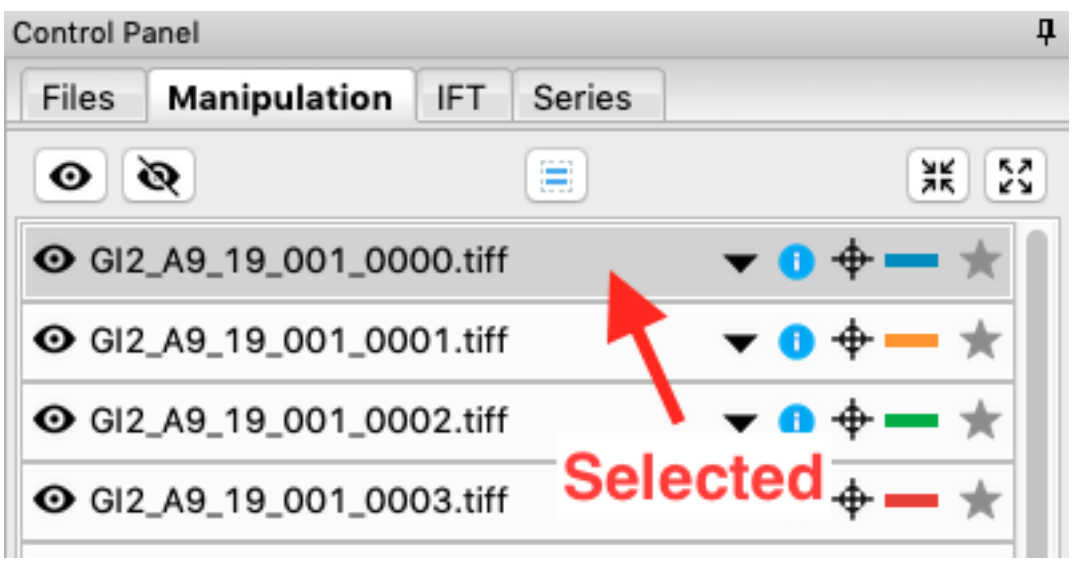
4. Double click on the **SAXS.cfg** file to load the SAXS configuration. This loads the beamline configuration into the program.
   - *Note:* Any time you are going to process images, you need to load the appropriate configuration!
5. Change the file type filter to "TIFF files (*.tiff)".
6. At the CHESS G1 station, typically ~10 images are collected from a given sample. To load in the 10 images for the glucose isomerase (GI) sample, start 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 process separately later.
7. Click the plot button to integrate all of the images and plot the integrated scattering profiles on the Main 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.



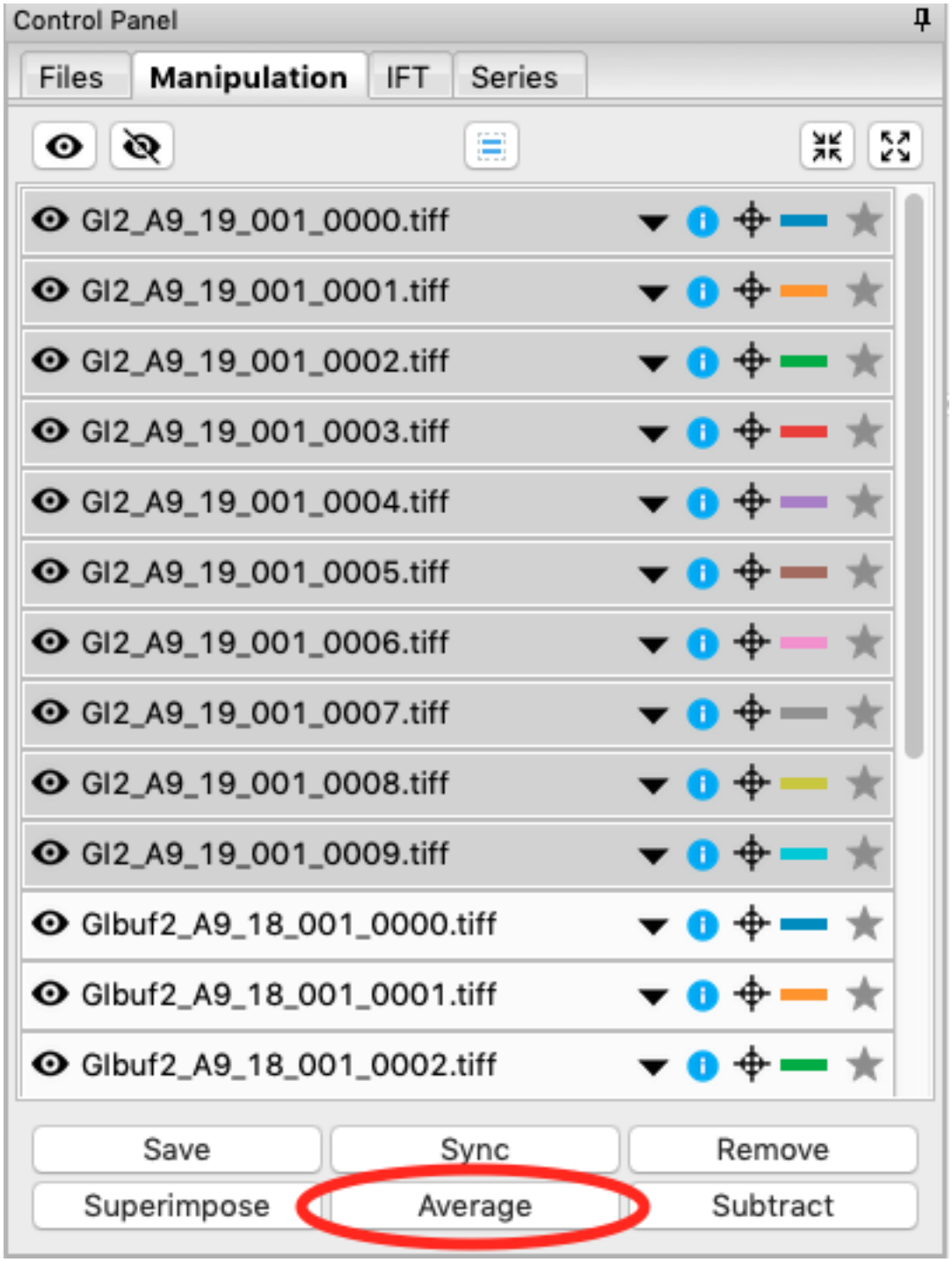
8. Plot the **GIbuf2** scattering profiles from the images. These are measured scattering from the matching buffer, without any protein, for the GI sample.
9. Click on the Manipulation tab. This is where you can see what scattering profiles are loaded into RAW, and manipulate/analyze them.
   - *Checkpoint:* If you've successfully loaded the images given, you should see twenty scattering profiles in the manipulation 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 gray, indicating it is selected.



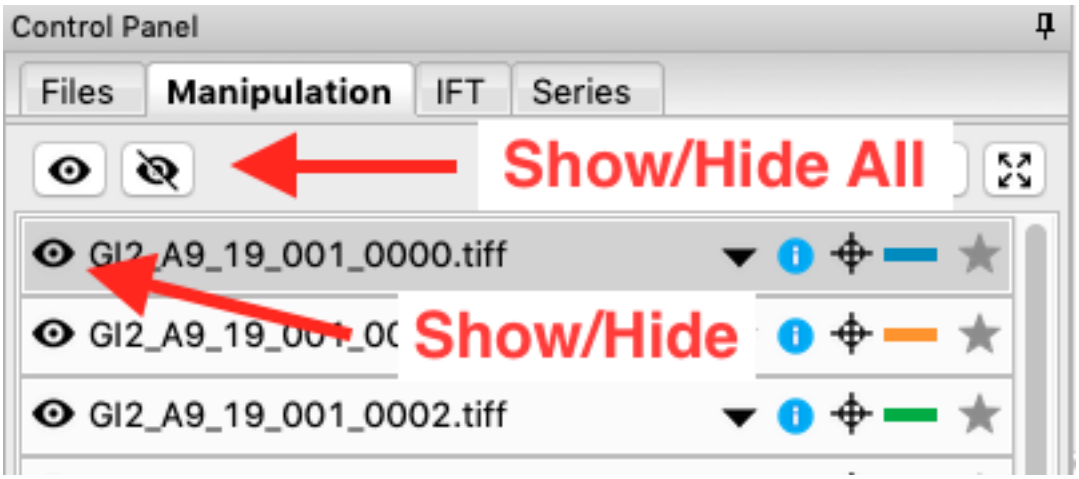
11. Select all of the GI scattering profiles
    - *Tip:* Again, the ctrl(/apple) key or the shift key can be used to select multiple scattering profiles.
    - *Warning:* Select only the GI profiles, not the GI buffer profiles.



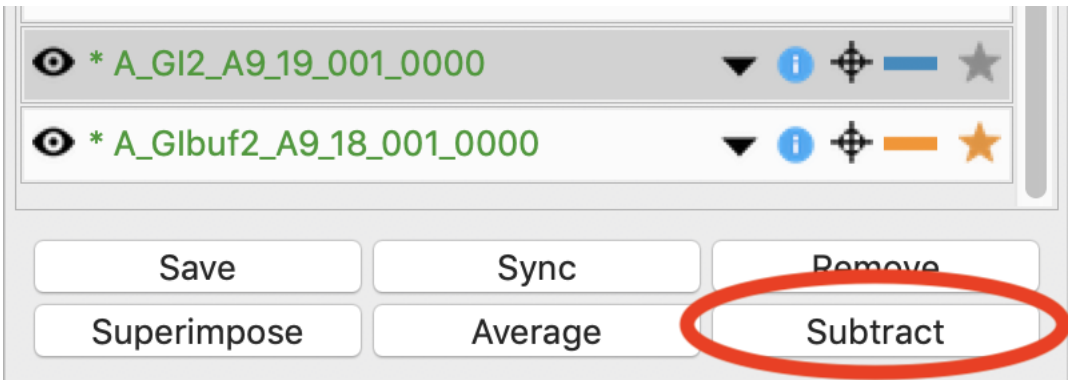
12. Use the average button to average all of the scattering profiles collected into a single curve.
    - *Checkpoint:* The averaged scattering profile should appear at the bottom of the manipulation list. You may have to scroll down to see it. The filename will be in green, and will start with **A_**, 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.

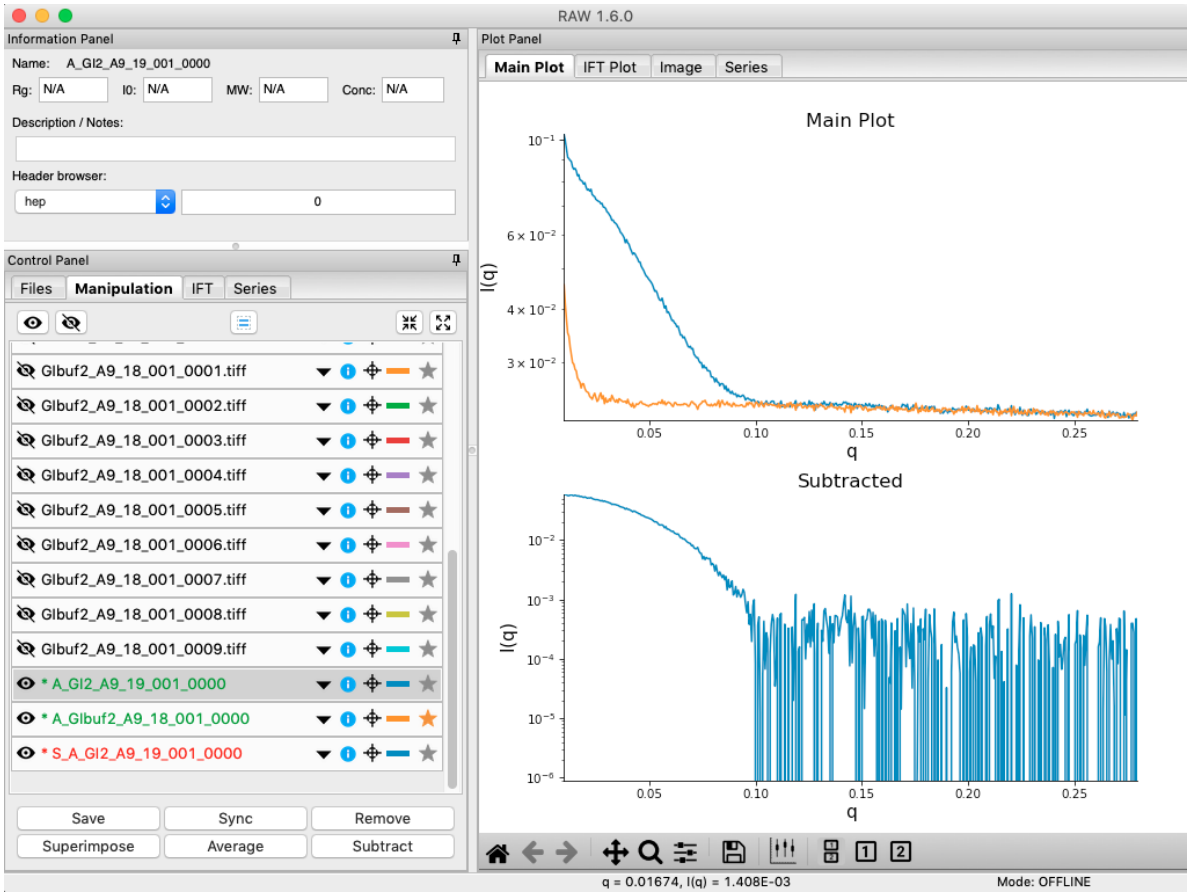
- *Tip:* The eye and eye with the line through it at the top of the manipulation panel can be used to show/hide sets of loaded profiles at once. If no profiles are selected, these buttons show/hide all loaded profiles. If some profiles are selected, these buttons show/hide just the selected profiles. 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 manipulation item should be shown in the manipulation menu, with the name in red and a **S_** prefix indicating it is a subtracted file.

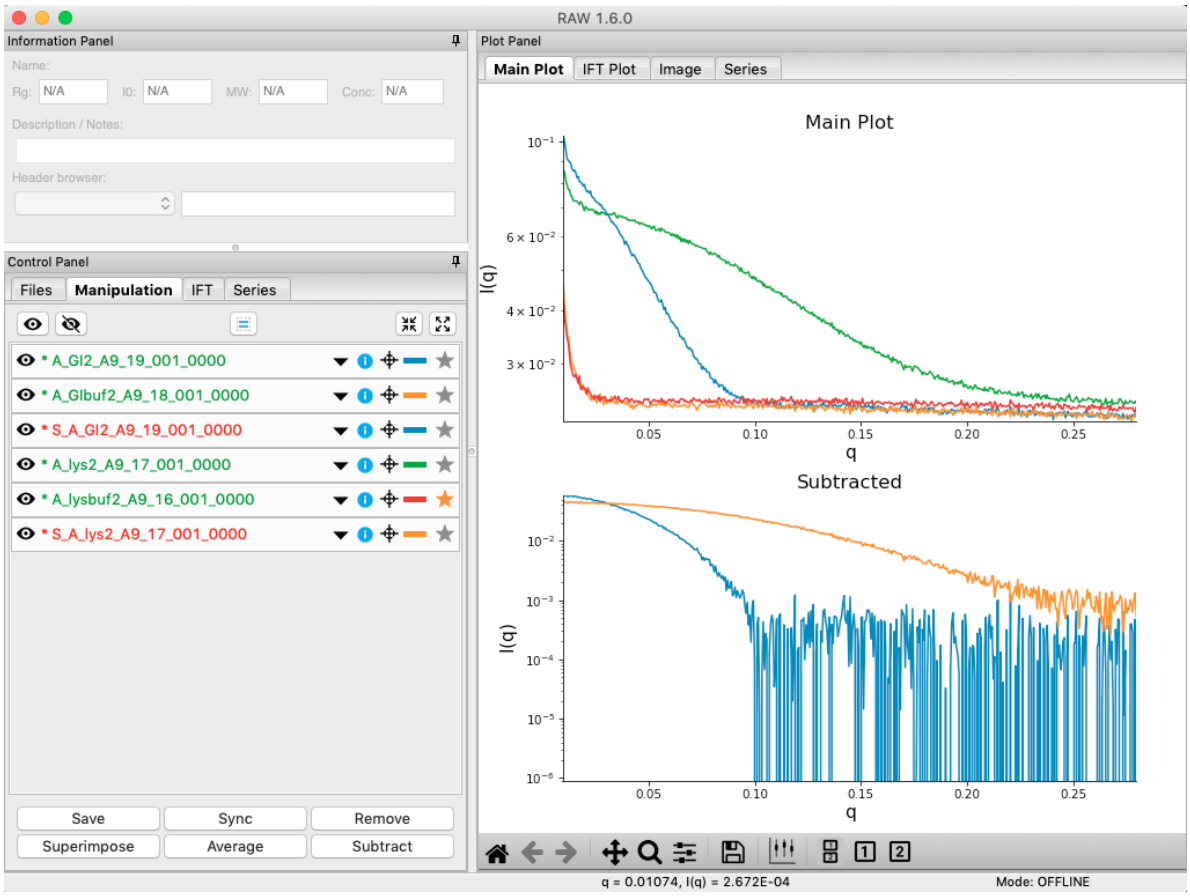


16. You don't need the individual image scattering profiles any more. Select all of those (but not your averaged or subtracted profiles!) and click remove.

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

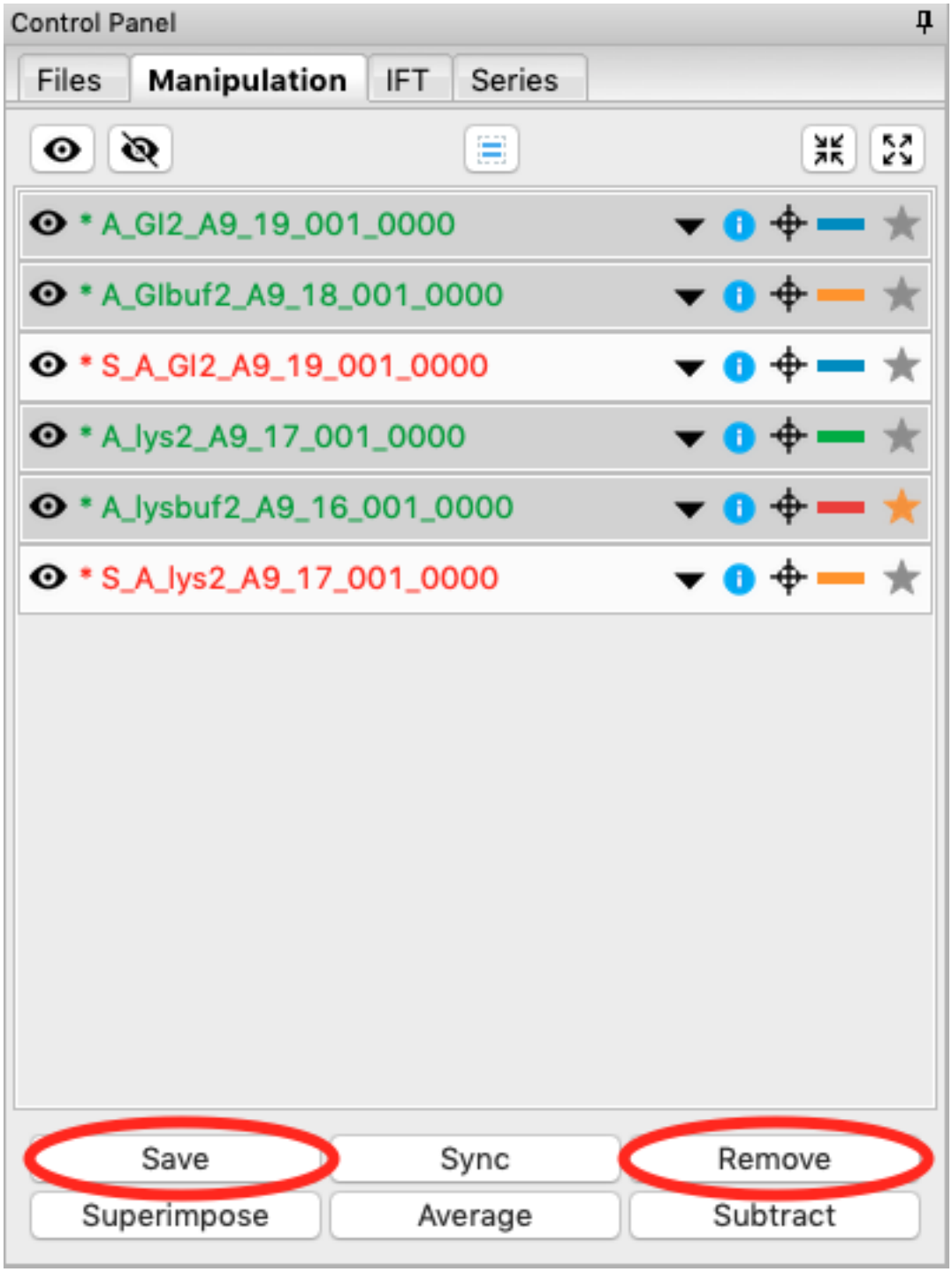
17. Load in the **lys2** and **lysbuf2** files, average, and subtract to create a subtracted lysozyme scattering profile. The concentration of this sample was 4.27 mg/ml. Remove all of the profiles that are not averaged or subtracted profiles.

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



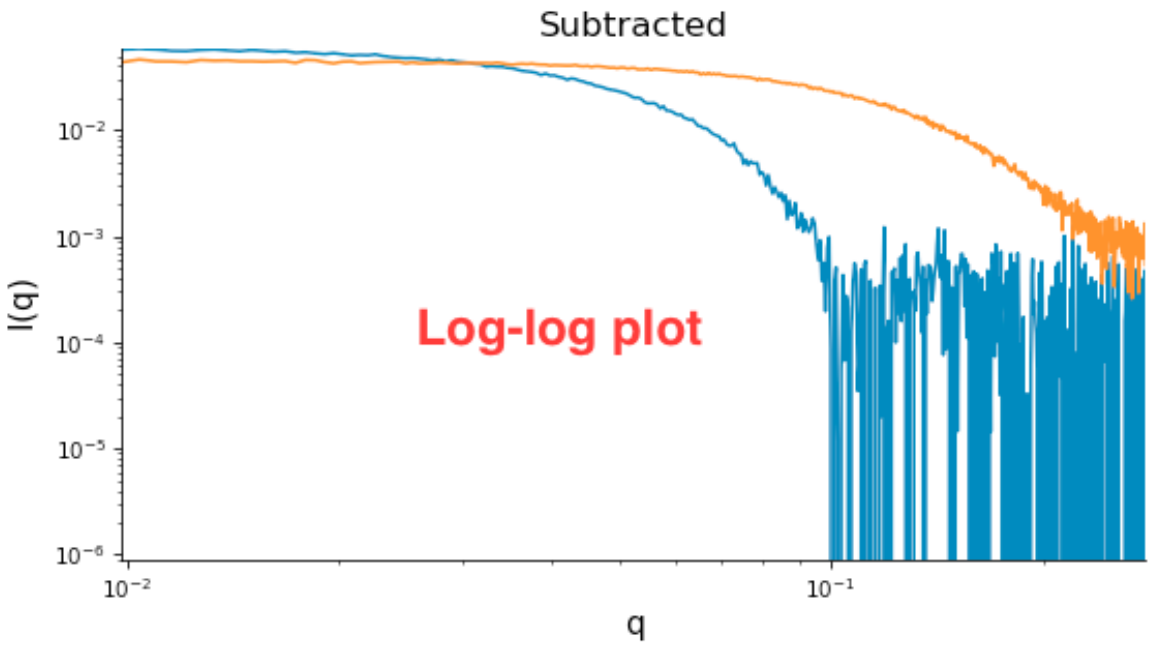
18. We're done with the averaged profiles. Select all of the averaged profiles and click the "Save" button to save them in the **standards_data** folder. Note that in the filename in the manipulation list, the * at the front goes away. This indicates there are no unsaved changes to those scattering profiles.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

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



19. Right click on the subtracted plot, move the cursor over 'Axes' and select the Log-Log option. 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 (Log-Lin, default option in RAW) or double-log plot (depending on the features of interest).

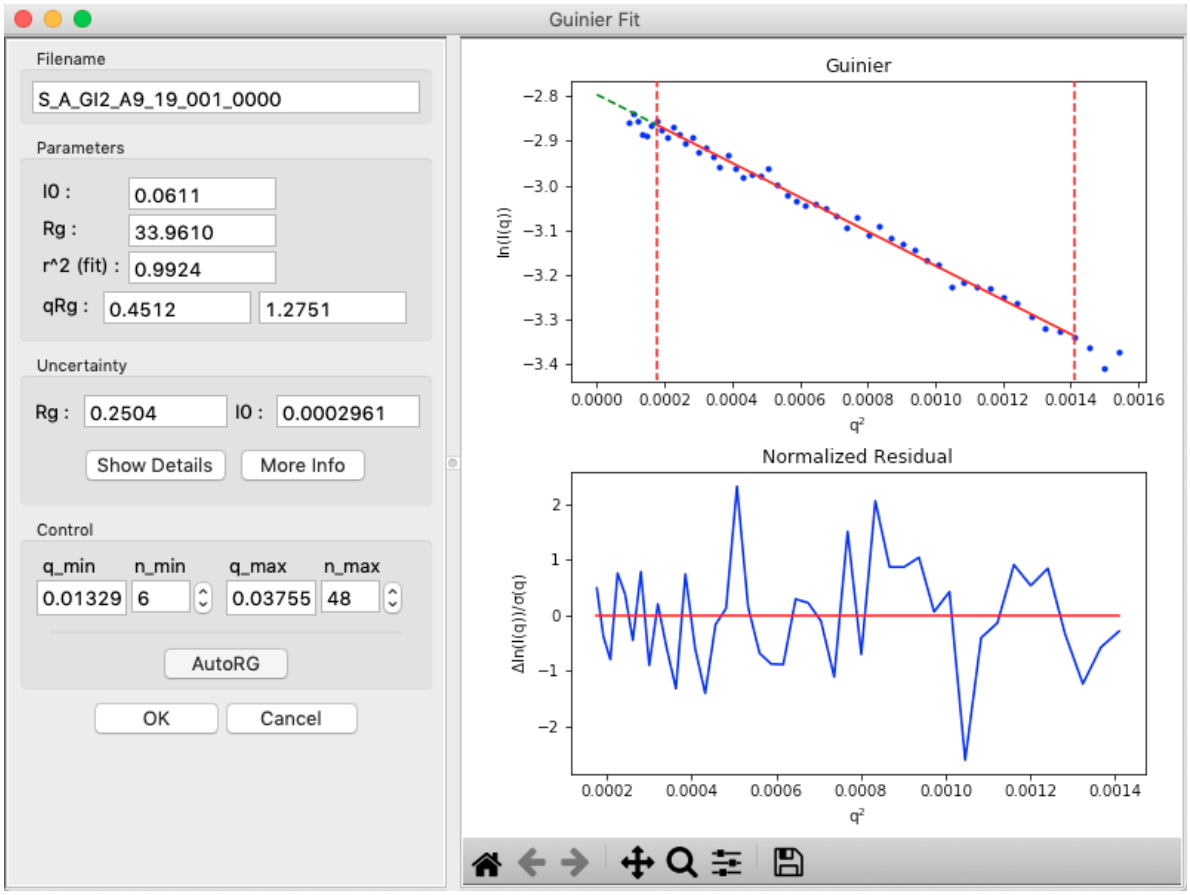


Guinier analysis

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 (ctrl click on macs without a right mouse button) on the subtracted GI scattering profile in the Manipulation list and select "Guinier fit". In the plots on the right, 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.
   - *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), while the two qR_g boxes give, left to right, $q_{min}R_g$ and $q_{max}R_g$ respectively.



2. In the "Control" panel, you'll see that n_min is now 6. This means RAW has cut off the first six points of the scattering profile in the fit. Use the arrow buttons next to the n_min box to adjust that to zero. Check whether the R_g changes.
3. In the "Parameters" panel, note that $q_{max}R_g$ is only ~1.27. Recall that for globular proteins like GI, it is typical to have $q_{max}R_g$ ~1.3. Adjust n_max until that is the case, watching what happens to the R_g and the residual.
   - *Question:* The literature radius of gyration for GI is 32.7 Å. How does yours compare?
4. 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 autorg function or an estimated uncertainty in R_g and I(0) based on variation of the selected interval start and end points.
5. Click the "OK" button to keep the results.
   - *Checkpoint:* If you now select the GI scattering profile, in the information panel at the top you should see the R_g and I(0) that you just found.
   - *Note:* Clicking the "Cancel" button will discard the results.
6. Repeat the Guinier analysis for lysozyme.
   - *Try:* Increase q_{min} and/or decrease q_{max} 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.

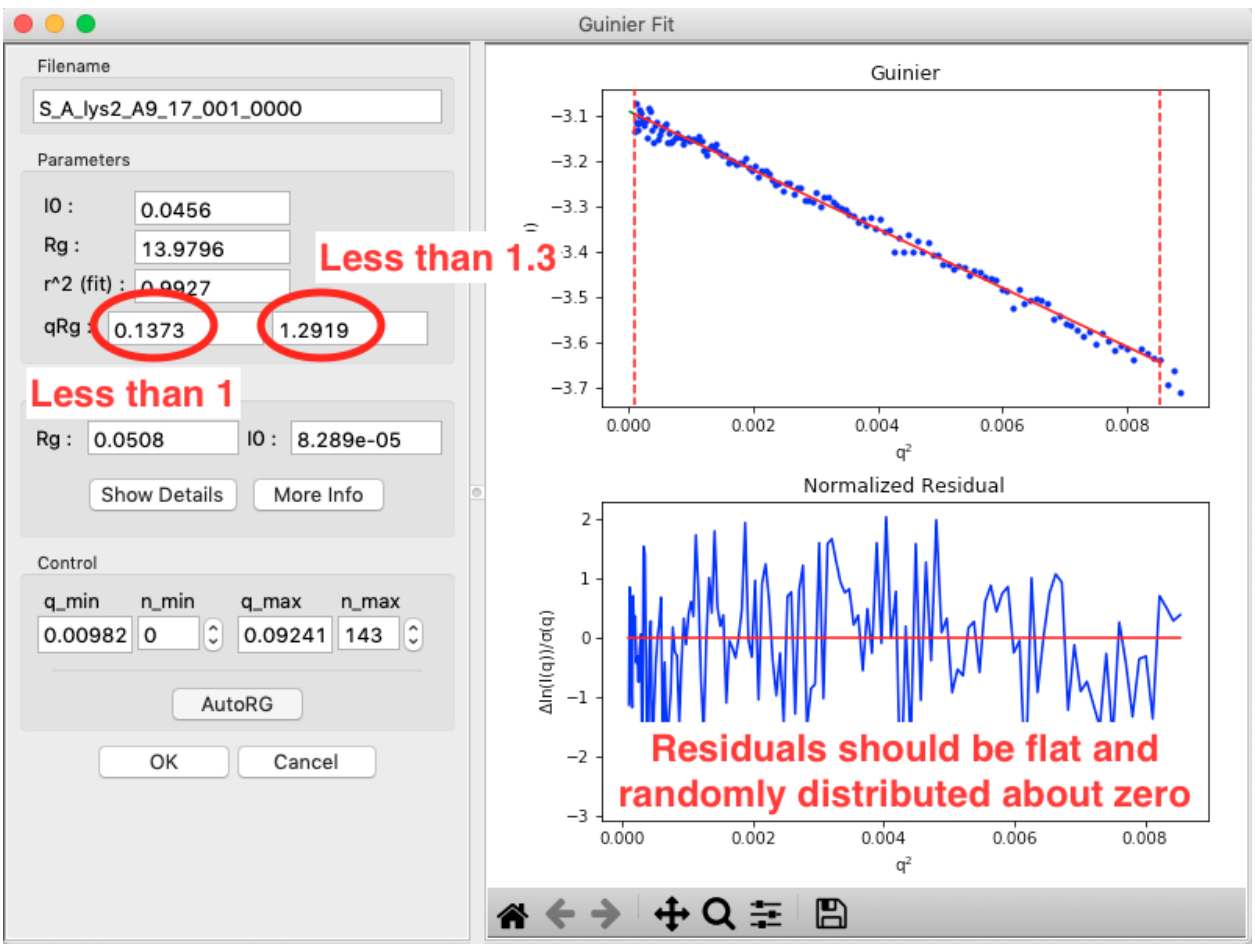
Aside: Criteria for a good Guinier region

For a globular protein, you are looking for three essential components in your Guinier fit:

- $q_{min}R_g < 1.0$. This states that the minimum q of your fit, q_{min}, times the R_g of your fit should be less than 1.0. This criteria ensures you have enough q range to properly estimate the R_g and I(0) values.
- $q_{max}R_g < 1.3$. This states that the maximum q of your fit, q_{max}, times the R_g of your fit should be less than 1.3. This criteria ensures you remain in the linear range of the Guinier approximation for the 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.

Additionally, you shouldn't have to excluded very many points at the start of the fit. A few is generally fine, as the points nearest the beamstop can be noisy (depending on the exact details of the measurement). 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 to understands you observed these effects and decided to proceed with analysis despite the non-ideality.

Note that for non-globular systems, such as rod-like shapes, the fitting range shifts as the linear region of the Guinier approximation shifts.



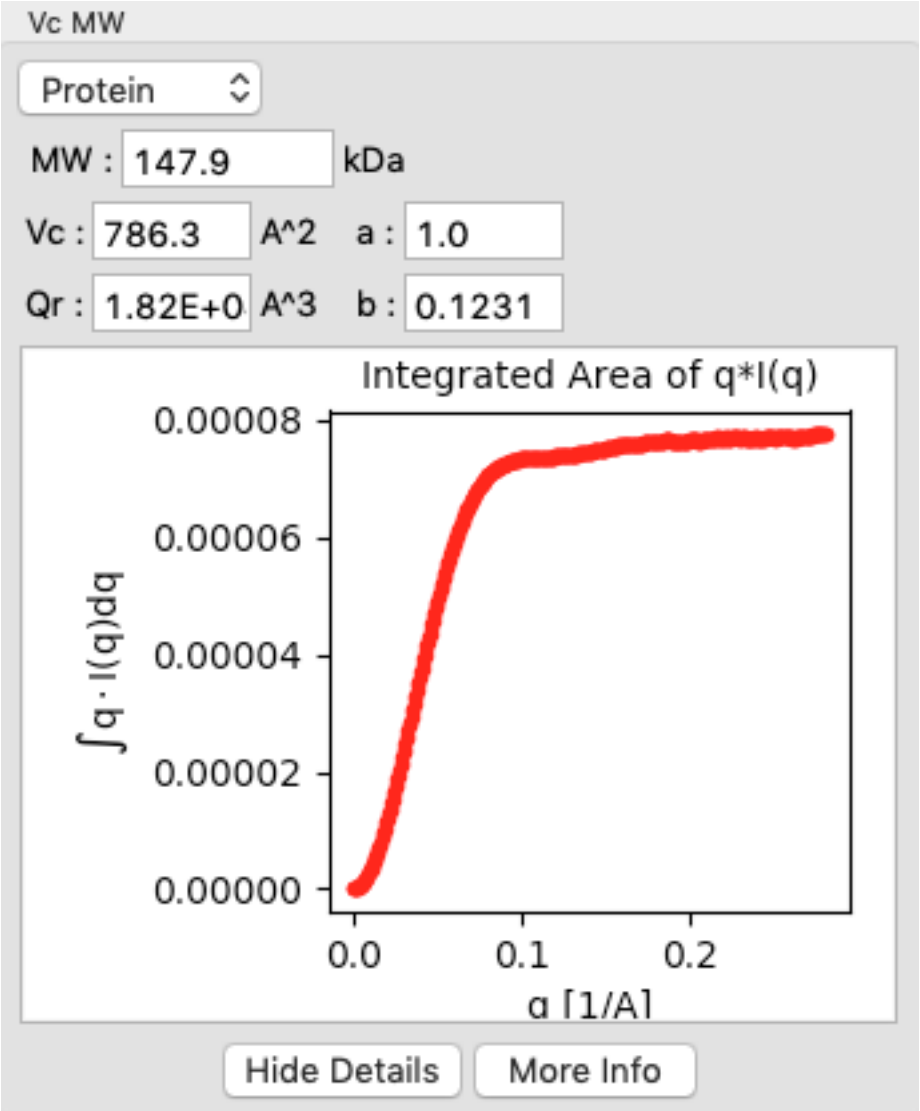
Molecular weight analysis

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.

1. In RAW, right click on the subtracted GI scattering profile in the manipulation panel and select "Molecular weight." At the top of the panel is a description of the methods used, and 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.

2. 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?
3. Click on the "Show Details" button for the Vc MW panel. You should see a graph, which shows the integrated area of $qI(q)$ vs. q. For this method to be accurate, this value needs to converge at high q.



4. Click the "OK" button to save your analysis.
   - *Note:* The "Cancel" button discards the analysis.
5. 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 method work for the lysozyme data?

Aside: Discussion of the four methods:

Calculation of MW from reference to a known protein 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).

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.

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.

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.

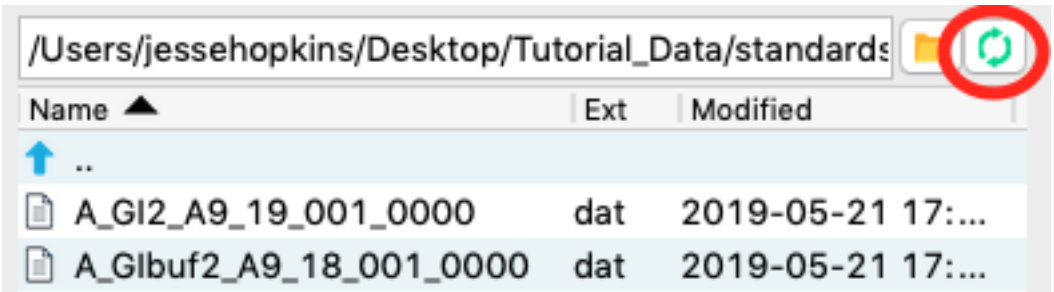
This method can yield inaccurate results if:

- 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).
- Your molecule is not a protein (e.g. RNA/DNA or a protein-nucleic acid complex).

Saving analysis information

1. Save your subtracted scattering profiles in the **standards_data** folder.
2. 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.

3. Remove the subtracted scattering profiles from RAW by selecting both of them and clicking the "Remove" button.
4. Load the saved subtracted scattering profiles back into RAW. Note that if you select one in the Manipulation 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.



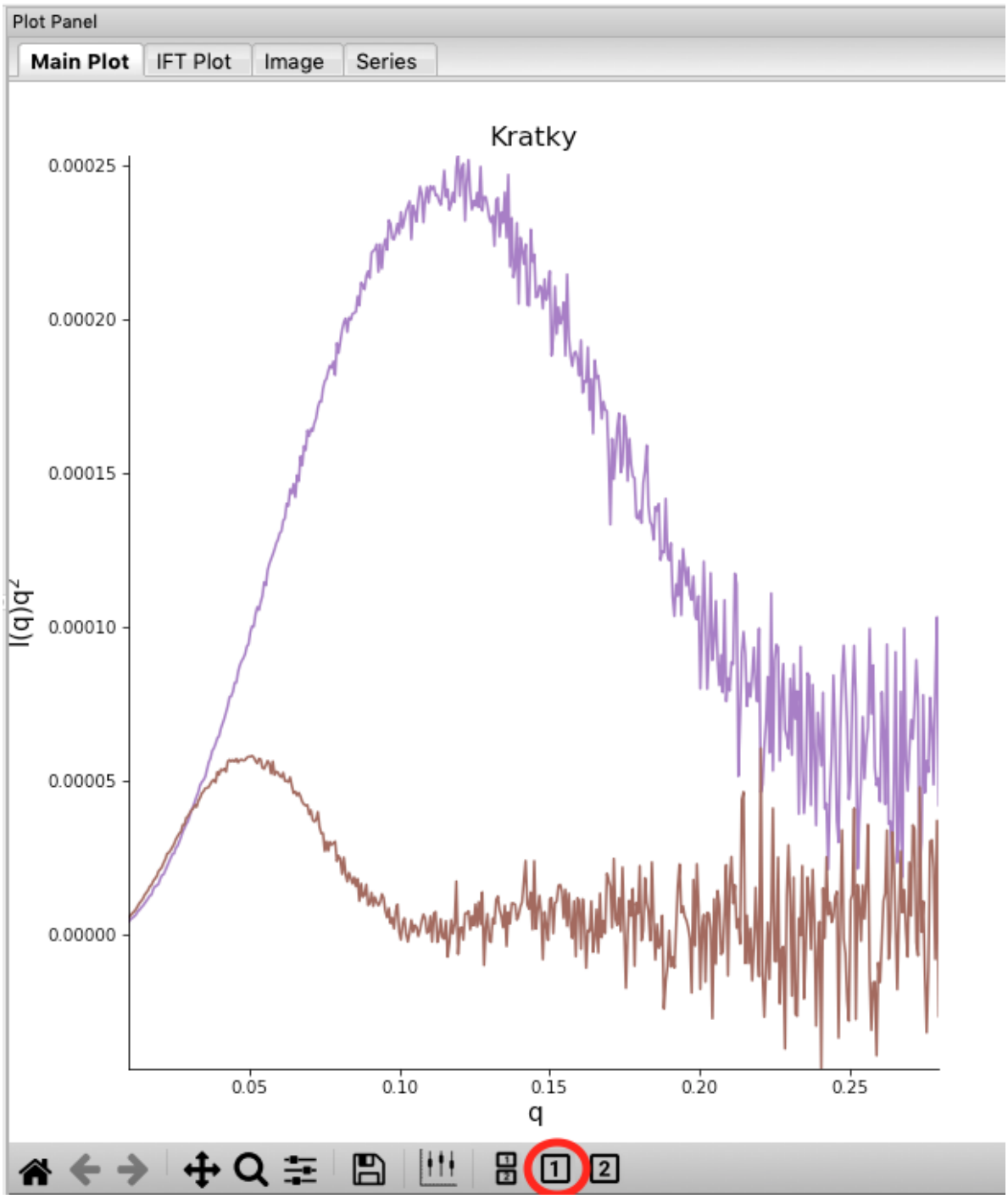
   - *Try:* Open the saved subtracted scattering profile **S_A_GI2_A9_19_001_0000.dat** in a text editor such as Notepad (windows) or TextEdit (mac). You should see all of the data in three columns, followed by header information. If you scroll down far enough, the header information contains all of the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

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_gor $(q^2V_c)I(q)/I(0)$ vs. $q(V_c)^{1/2}$. These dimensionless plots can provide semi-quantitative analysis of flexibility and disorder. More information about can be found here and references therein: http://www.bioisis.net/tutorial/21.

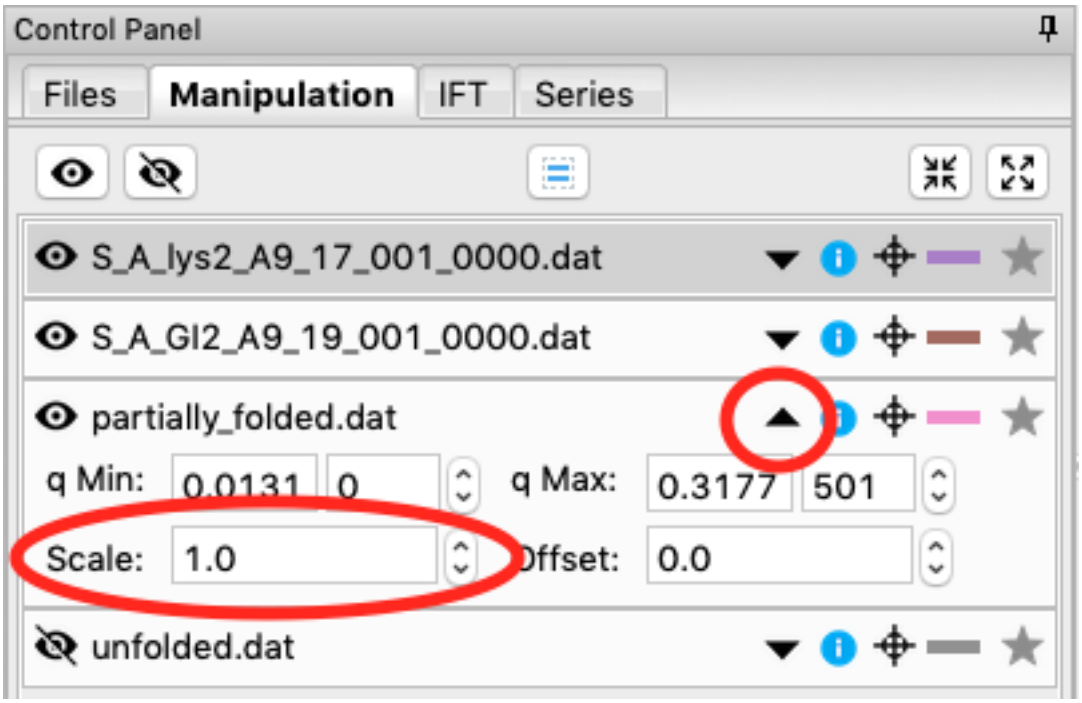
1. If you haven't already, save the loaded scattering profiles, remove all loaded profiles from RAW, and reload your saved GI and Lysozyme profiles, as described in *the previous section*.
2. Put the top plot on Kratky axes.
   - *Tip:* Right click on the plot to change the plot type.
3. Show only the top plot by clicking on the 1 in the plot control bar below the plots.



4. 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.
5. Load the two scattering profiles in the **Tutorial_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, and has the BIOISIS ids of 2LYSRR and 3LYSRR.

6. 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 manipulation 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.

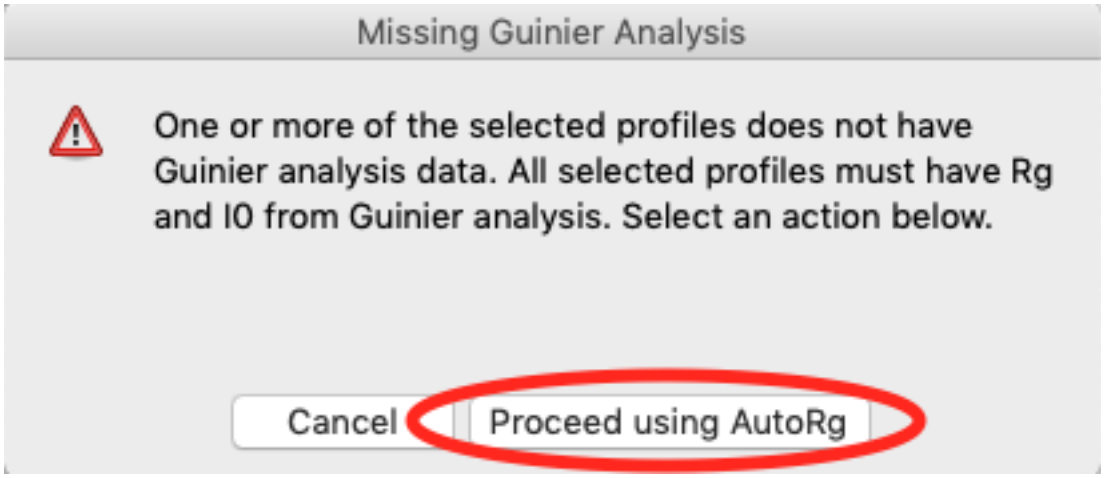
   - *Tip:* The up and down arrows will only adjust the last digit of the scale factor.



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).

8. Select all four loaded scattering profiles, right click, and select the Normalized Kratky Plot option.

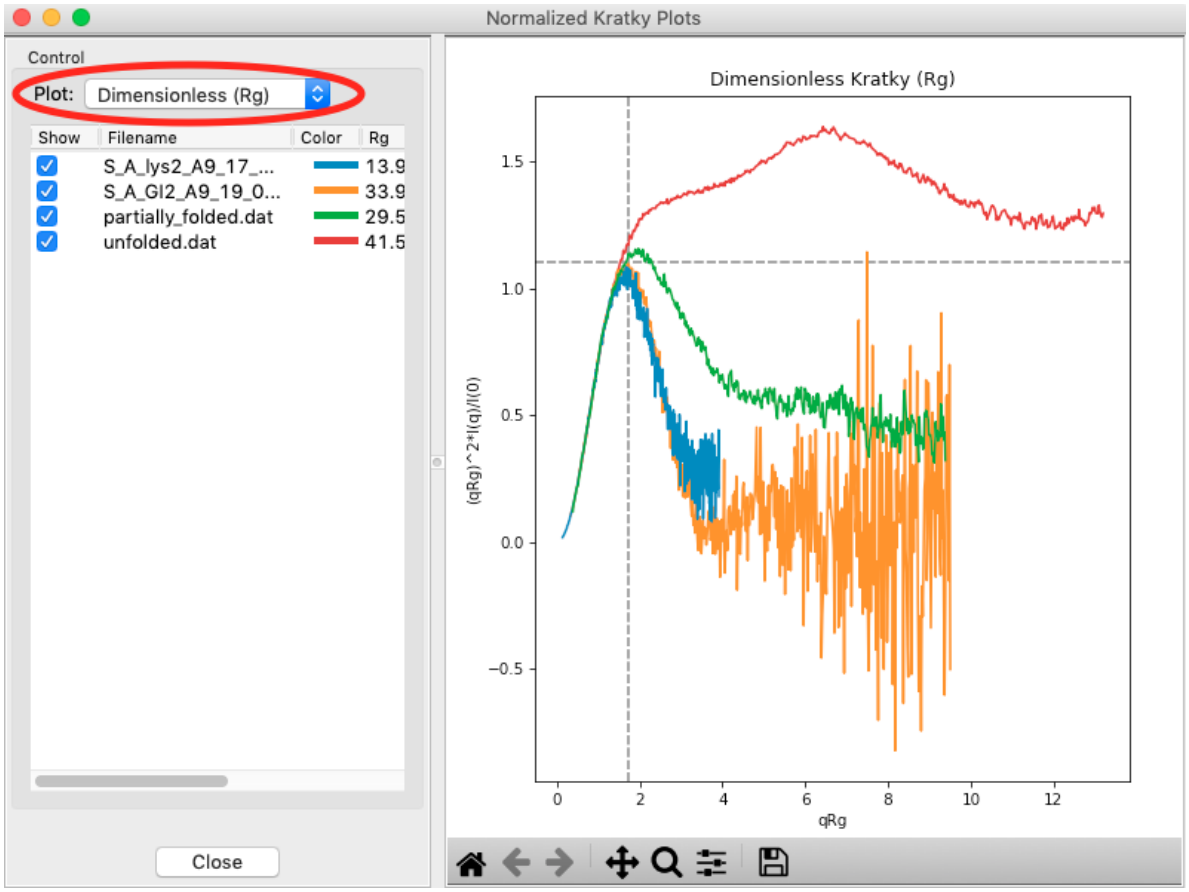
9. Normalized Kratky plots require Guinier analysis to be done. If one or more profiles are missing this information, RAW will show the following window. You can either cancel, and do the fits manually, or you can proceed with RAW's automatic determination.



10. Click the Proceed using AutoRg button to proceed to the Normalized Kratky Plot window using RAW's automatic fitting for R_g.

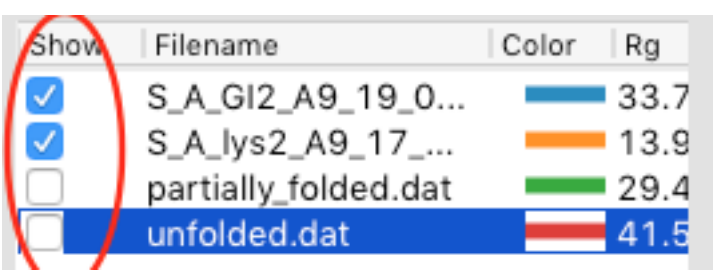
11. By default, the plot is the Dimensionless R_g plot. Use the dropdown "Plot" menu at the top to select the Normalized (by I(0)) and Dimensionless Vc plots.

    - *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.



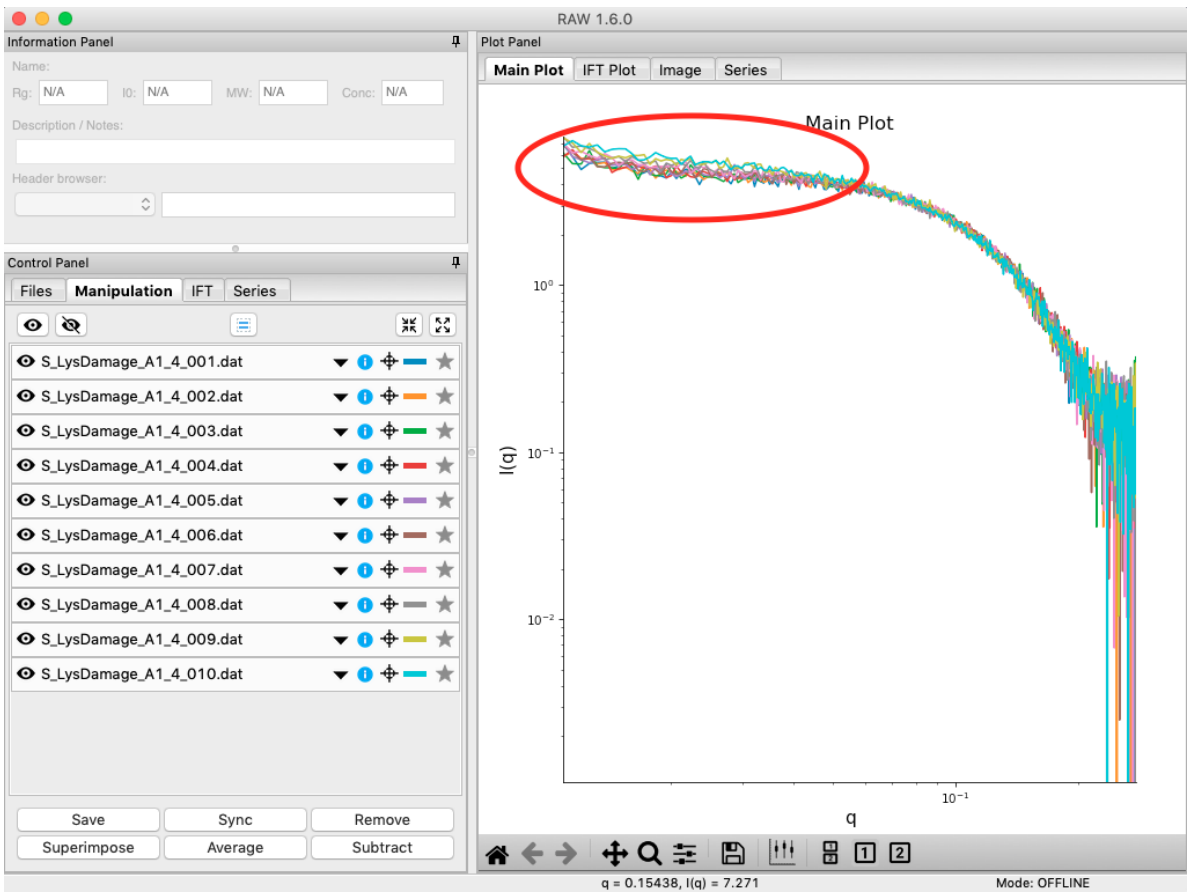
13. Right click on the plot and select "Export Data As CSV" to save the dimensionless data for further processing or plotting with another program.

14. Click the Close button to close the Normalized Kratky Plot window.

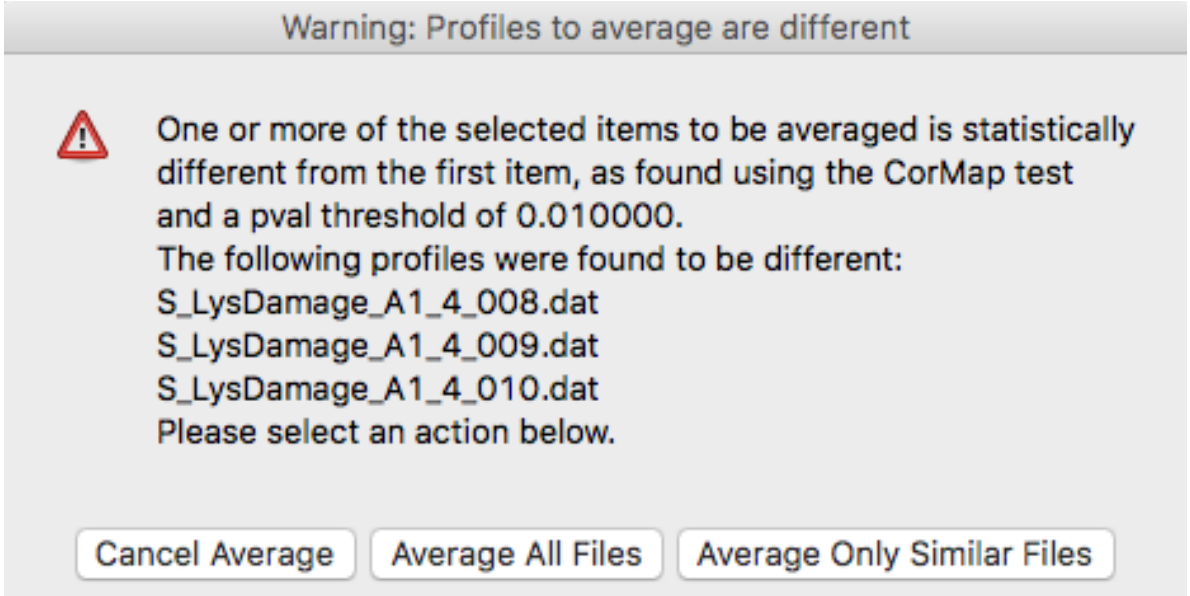
Similarity Testing

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.

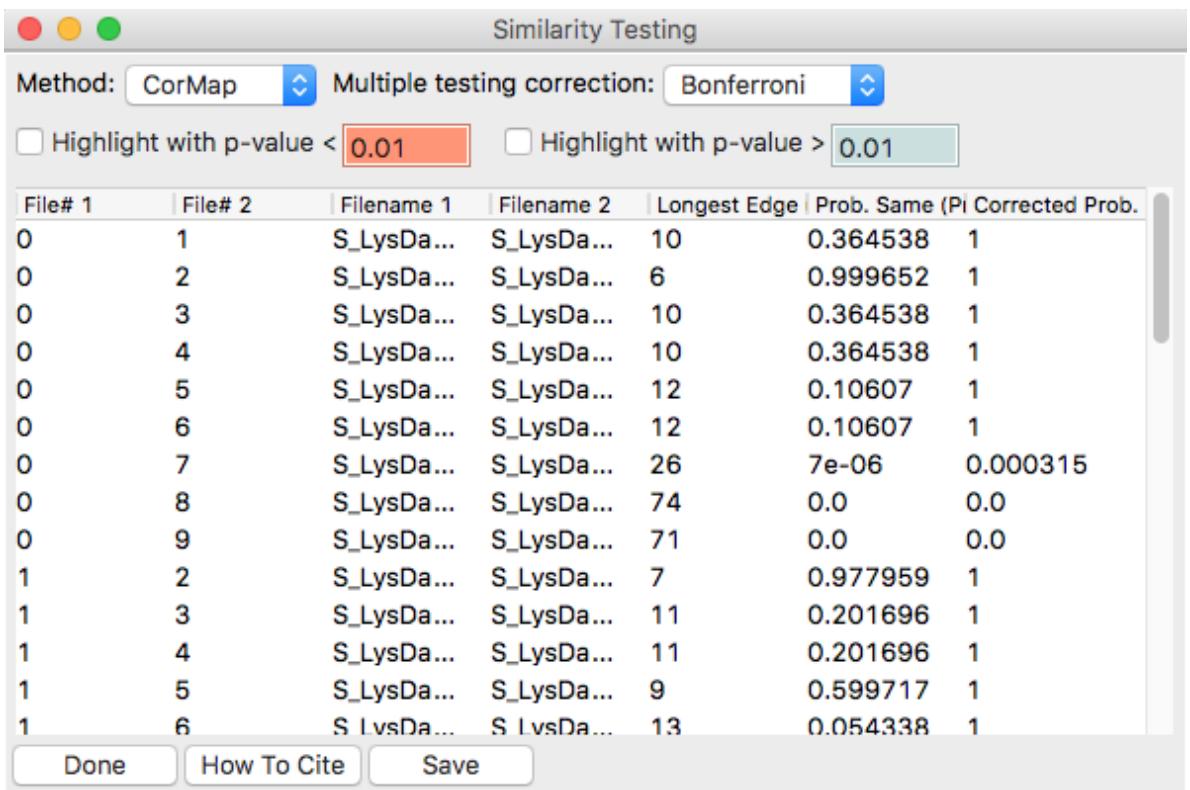
1. Clear any data loaded into RAW. Load all of the profiles in the **Tutorial_Data/damage_data** folder into the main plot. Show only the top plot.
   - *Tip:* In the Files tab, click the "Clear All" button.
2. Put the plot on a log-log scale. You should see that the profiles are different at low q.
   - *Note:* These data are showing what radiation damage 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.



3. Select all of the 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.



4. 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.
5. Select all of the profiles except the new averaged one, and right click and select "Similarity Test".

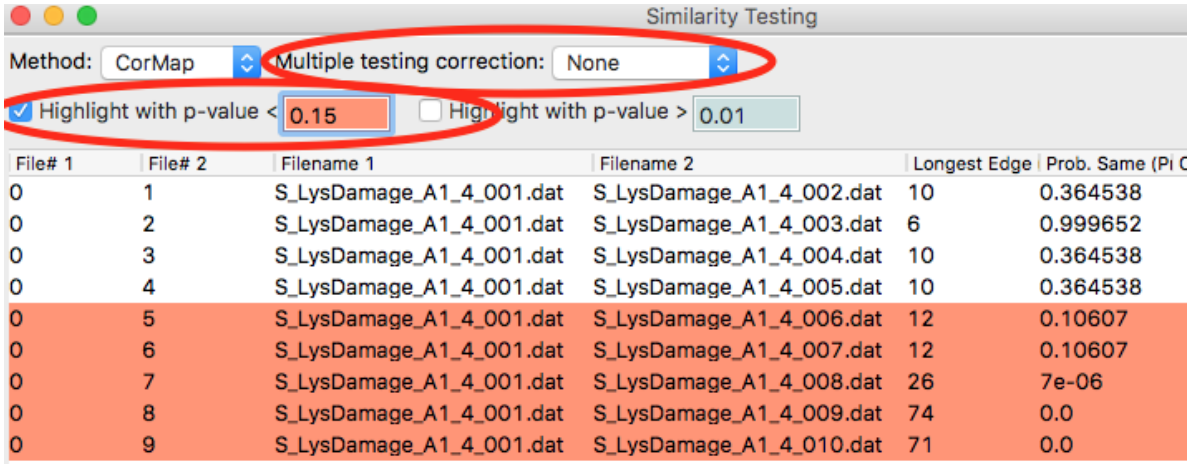


File# 1	File# 2	Filename 1	Filename 2	Longest Edge	Prob. Same (Pi	Corrected Prob.
0	1	S_LysDa...	S_LysDa...	10	0.364538	1
0	2	S_LysDa...	S_LysDa...	6	0.999652	1
0	3	S_LysDa...	S_LysDa...	10	0.364538	1
0	4	S_LysDa...	S_LysDa...	10	0.364538	1
0	5	S_LysDa...	S_LysDa...	12	0.10607	1
0	6	S_LysDa...	S_LysDa...	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

6. 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

7. 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 C
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

8. 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, it is reasonable to suspect that we should discard profiles 6-10, not just 8-10 as in the automated version.
9. Save the similarity test data as a **.csv** by clicking the "Save" button.
10. Close the similarity testing window by clicking the "Done" button.
11. Average profiles 1-5.
12. Hide all of the profiles except the two averaged profiles on the plot.
    - *Question:* Is there a difference between the two? What about if you do a Guinier fit?
    - *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.

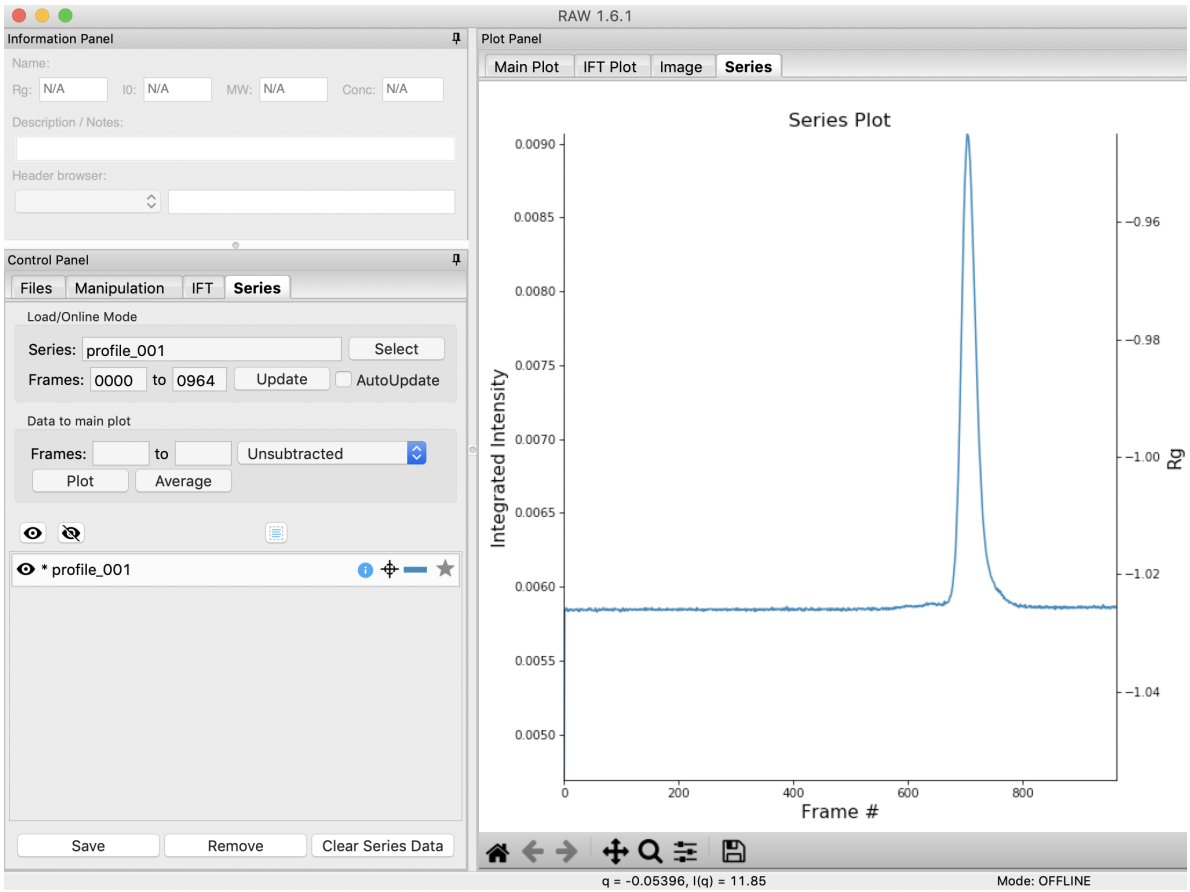
Basic SEC-SAXS processing

In a typical SEC-SAXS run, images are continuously collected while the eluate (outflow) of a size exclusion column flows through the SAXS sample cell. As proteins scatter more strongly than buffer, a plot of total scattered intensity vs. time, the so-called SAXS chromatograph (or scattergram), will show a set of peaks similar to what is seen by UV absorption measurement of the SEC system. RAW includes the capability to do routine processing of SEC-SAXS data. This includes creating the SAXS chromatograph from the data, plotting R_g, MW, and I(0) across the peaks, and extracting specific frames for further analysis.

Note: In RAW, this is called Series analysis, as the same tools can be used for other sequentially sampled data sets.

1. Clear any data loaded into RAW. Click on the Series tab in the control panel. Click the "Select" button. Navigate to the **Tutorial_Data/sec_data/sec_sample_1** folder and select any of the **.dat** files in the folder.
    - *Tip:* In the Files tab, click the "Clear All" button to clear all data in RAW.
    - *Troubleshooting:* If you get an error message, it means you don't have a configuration file loaded. Load the SAXS.cfg file referenced *earlier*.

2. The SEC run will automatically load. RAW should automatically show you the Series plot panel. If not, click on the Series tab in the plot panel.

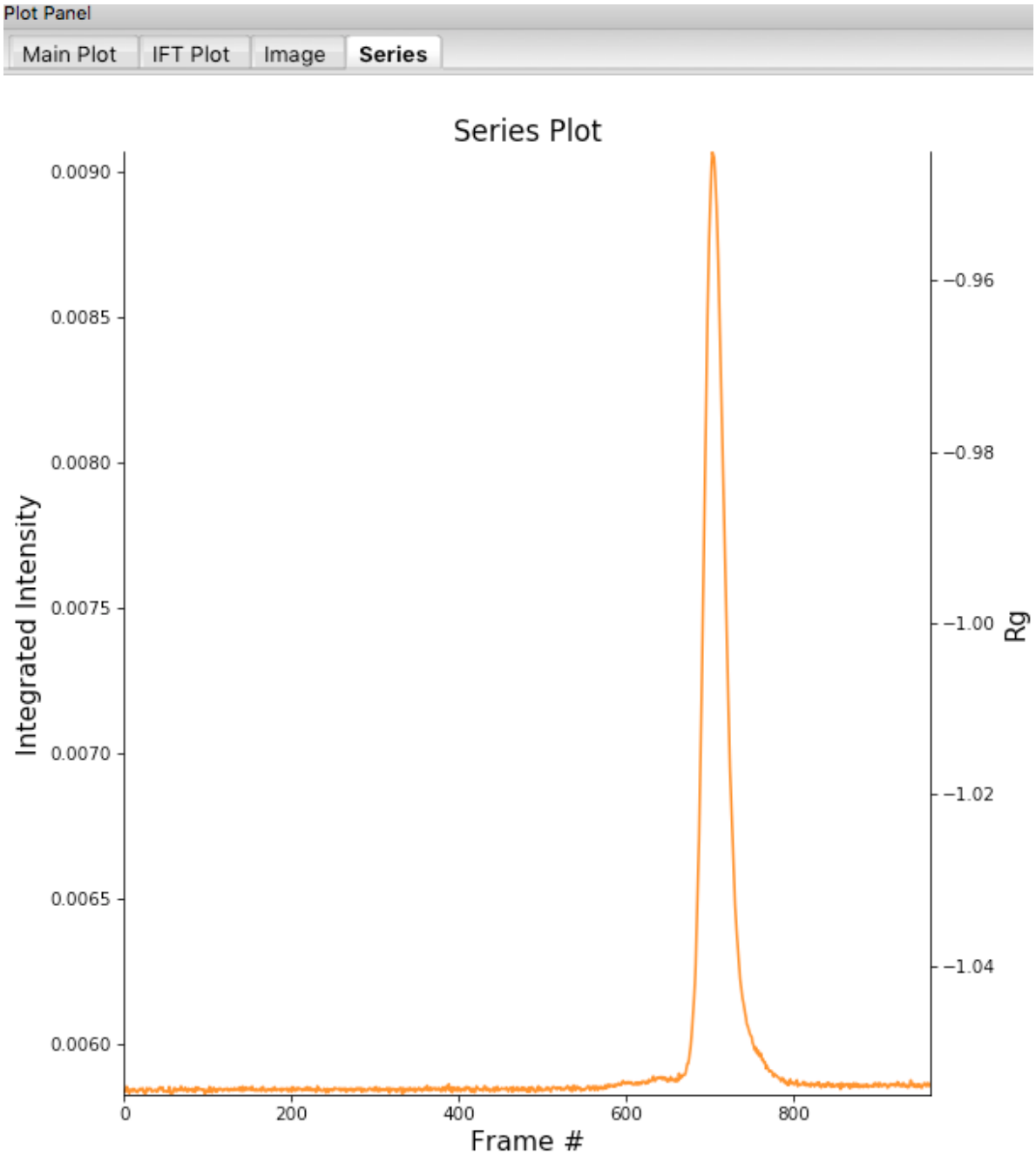


- *Try:* Each point on this curve is the integrated intensity of a scattering profile. You can figure out which one by right clicking on the filename in the Series list and selecting 'Show data'. This will show you the frame number and integrated intensity displayed on the plot, and the filename corresponding to the displayed frame number.

3. Drag the plot so that you can clearly the see the first frame. You'll notice it has a significantly lower intensity than the rest of the frames. This happens occasionally at the MacCHESS G1 beamline (where the data was taken). It can make it harder to tell what the data is doing.

- *Tip:* Select the crossed arrows in the plot control bar, and then click and drag on the plot to move the curve around on the screen.

4. Go to the Files control tab and navigate to the **sec_sample_1** data directory. Click on the second data file, **profile_001_0001.dat**. Scroll down to the bottom of the file list, and shift click on the last file, **profile_001_0964.dat**. This should highlight all of the files in between, as well as the two you clicked on.
5. Click on the "Plot Series" button. You will see the same curve plotted as before, but without the very first scattering profile. Remove the other loaded data set. Now you should have a curve where the baseline is very close to the bottom of the plot.

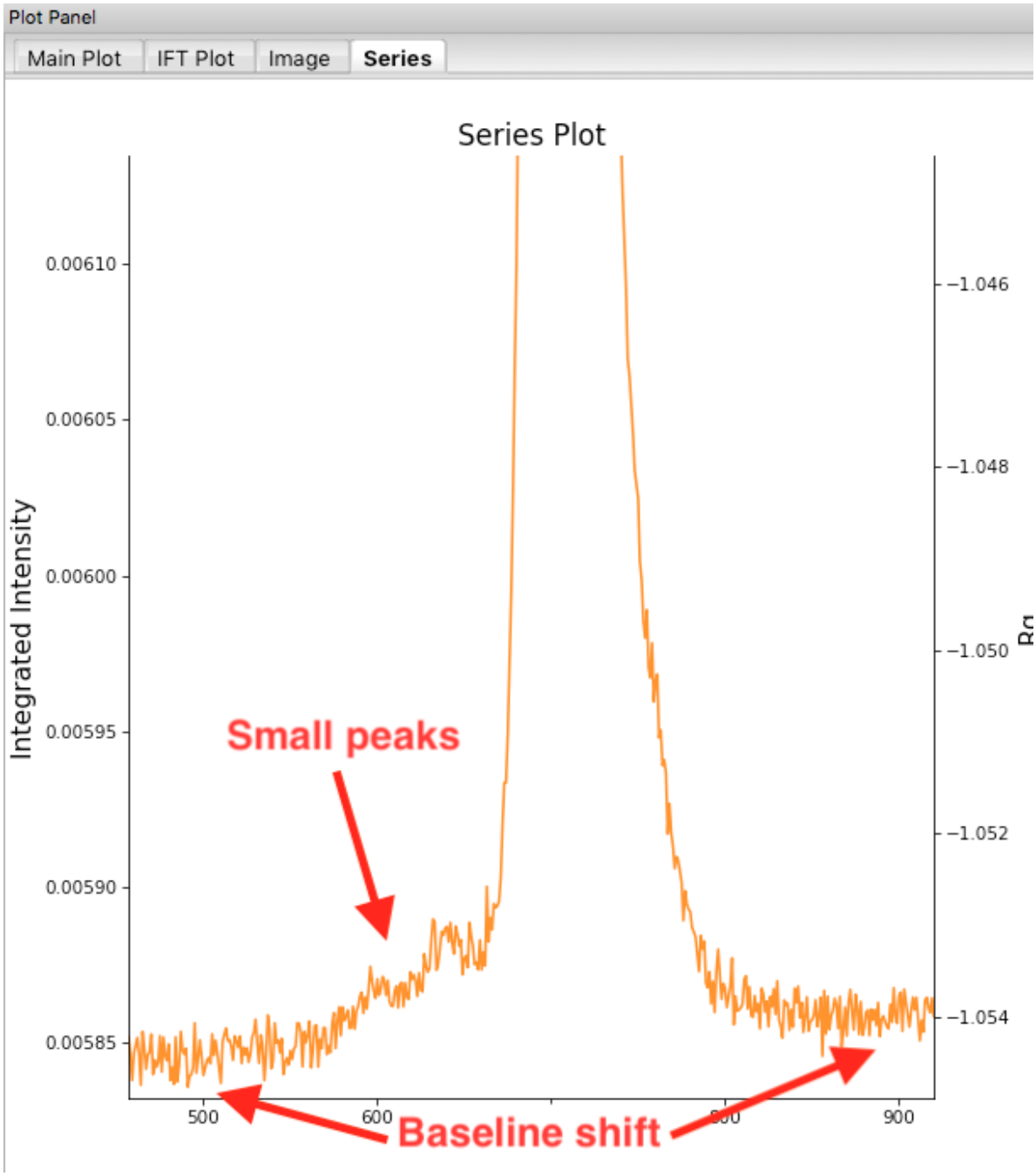


6. In some cases it is more useful to look at the mean intensity, the intensity at a specific q value, or the intensity in a range of q values than the total intensity. Right click on the plot and select mean intensity for the left axis y data. Then try the intensity at q=*0.02* and in the q range 0.01-0.03.
   - *Note:* You need to have the drag button in the plot control bar unselected to get a right click menu.

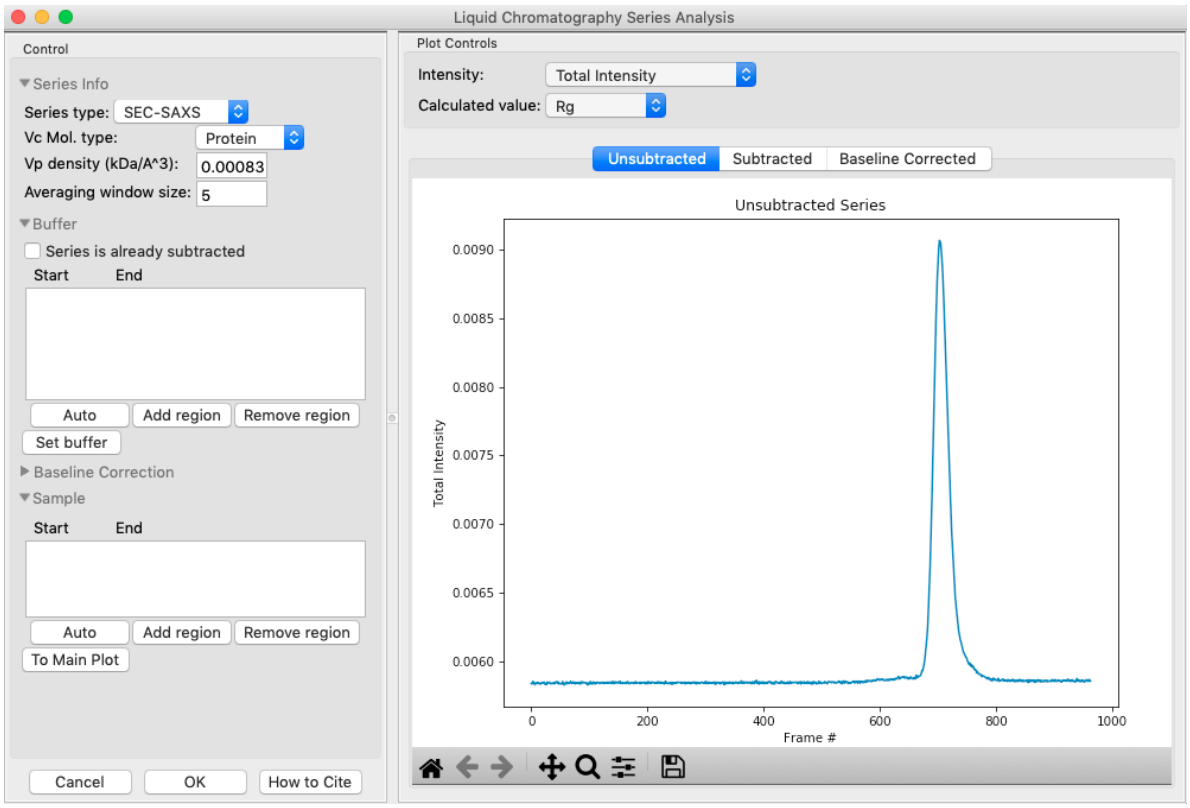
- *Tip:* CHROMIXS, in the ATSAS software displays the average intensity over the q range 0.01-0.08 $^{-1}$. To achieve a similar display in RAW, set your q range for the intensity to that. Note that RAW will display the sum of the intensity over the range while CHROMIXS displays mean intensity for the range, so the results won't be exactly the same.

7. Return to plotting the integrated intensity. Zoom in near the base of the peak. Notice that there are two smaller peaks on the left, likely corresponding to higher order oligomers that we don't have the signal to properly resolve. Also notice that the baseline after the peak is not the same as the baseline before the peak. This can happen for several reasons, such as damaged protein sticking to the sample cell windows.

   - *Tip:* Click on the magnifying glass at the bottom of the plot, then click and drag on the plot to select a region to zoom in on.

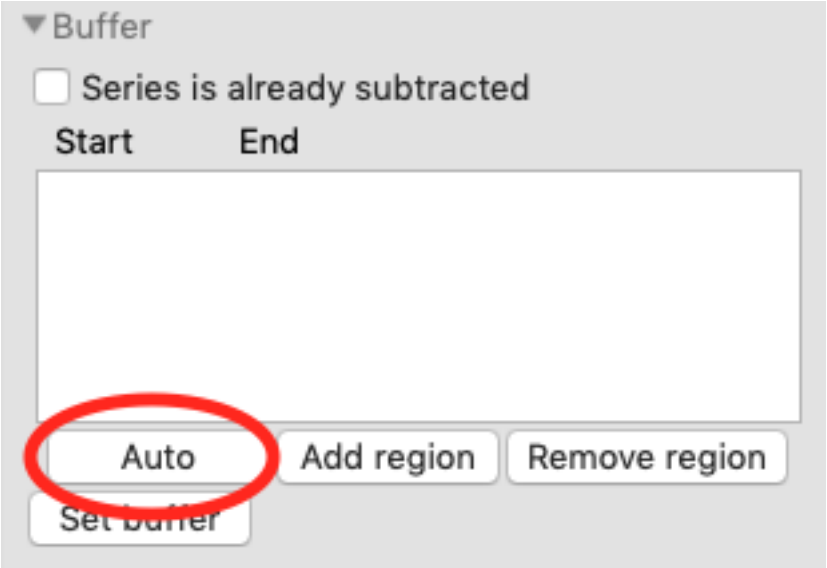


8. Zoom back out to the full plot.
   - *Tip:* Click the Home (house) button at the bottom of the plot.
9. In order to determine if we really have a single species across the peak, we will calculate the R_g and MW as a function of frame number. Right click on the filename in the Control Panel and select LC Series analysis to open the LC Series analysis panel.
   - *Note:* At the top of the control panel in this window, in the 'Series info' section you'll see several settings. If you had RNA instead of protein, you would use the Vc Mol. type menu to select that option. This affects the calculation of the molecular weight. You could also change the Vp density away from the default value, or change the averaging window (discussed below).

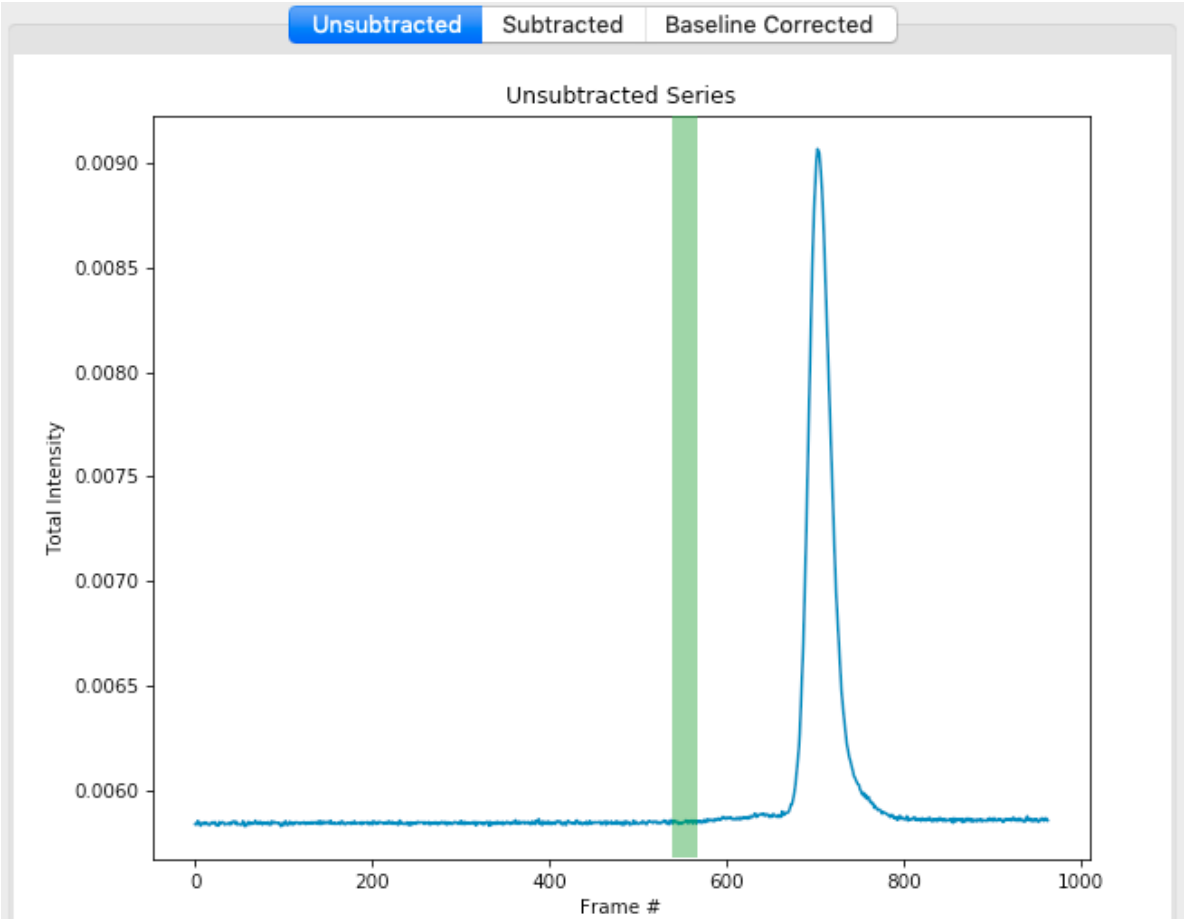


10. The LC Series analysis panel provides basic and advanced analysis tools for liquid chromatography experiments. Here we will show how to select buffer and sample regions, and send final processed data to the main plot. The advanced baseline correction features are discussed later.

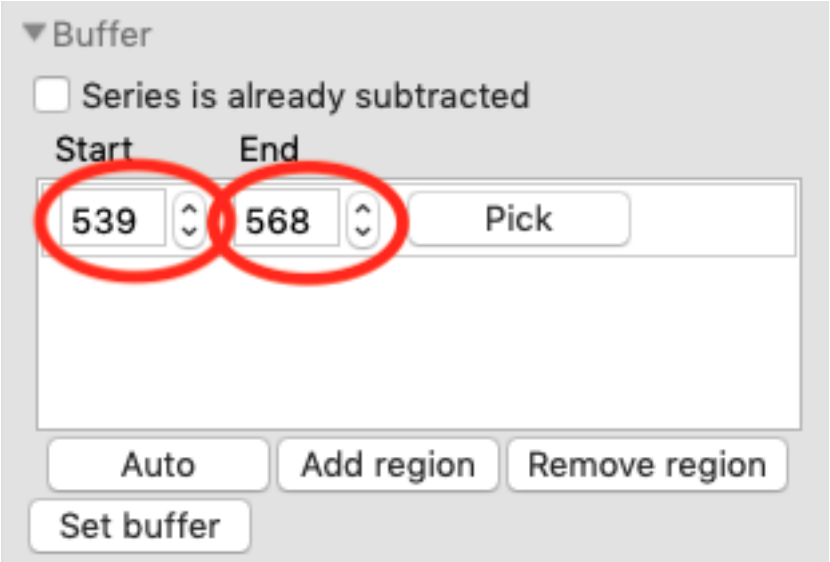
    In order to calculate R_g and other parameters as a function of elution time (Frame #), we need to define a buffer region. RAW can do this automatically. In the 'Buffer' section click the 'Auto' button.



- *Checkpoint:* You should see a buffer range show up in the buffer list, with defined start and end values. The region will be shown in green on the Unsubtracted plot.

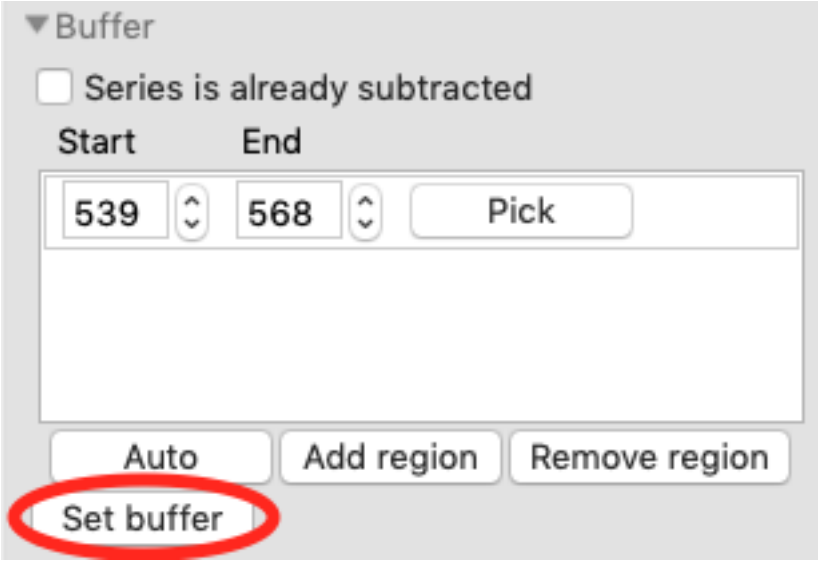


11. You can make fine manual adjustments to the buffer range if necessary. Zoom in on the baseline with the buffer region. Use the up/down arrows for the Start and End points to adjust the buffer region a little bit. You will see the region on the plot update as you make the changes.



- *Warning:* The automatic buffer determination can be wrong! Always be sure to manually inspect the region it picked. In particular, large flat leading edge shoulders next to the main peak can look like a baseline region to the algorithm, and will often mistakenly be picked.
- *Tip:* If the SAXS data isn't clear (noisy, low signal, etc.), it can be useful to inspect the UV trace associated with the SEC elution to see where there are minor elution components that you should exclude from your buffer selection.

12. Zoom back out on the plot. Reset the buffer range to 539 to 568 by typing those values in the Start/End range boxes and hitting enter.
13. To set the buffer region, create a set of subtracted profiles, and calculate structural parameters as a function of elution time, click the 'Set buffer' button. This may take a while to calculate.

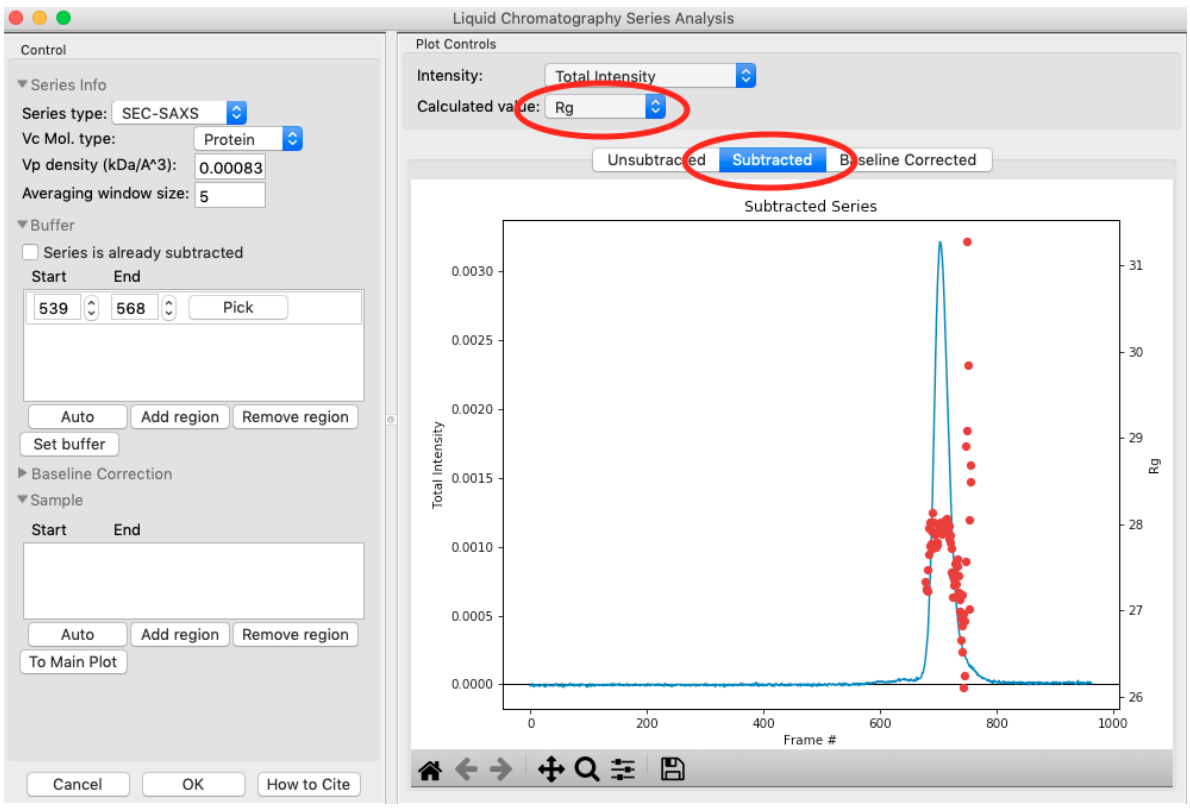


- *Note:* All of the files in the given buffer range will be averaged and used as a buffer. A sliding average window (size defined by the 'Averaging window size' in the 'Series Info' section) is then moved across the SEC curve. So for a window of size five, the profiles corresponding to frames 0-4, 1-5, 2-6, etc will be averaged. From each of these averaged set of curves, the average buffer will be subtracted, and RAW will attempt to calculate the R_g, MW, and I(0). These values are then plotted as a function of frame number.

- *Warning:* It is important that the buffer range actually be buffer! In this case, we need to make sure to not include the small peaks before the main peak.

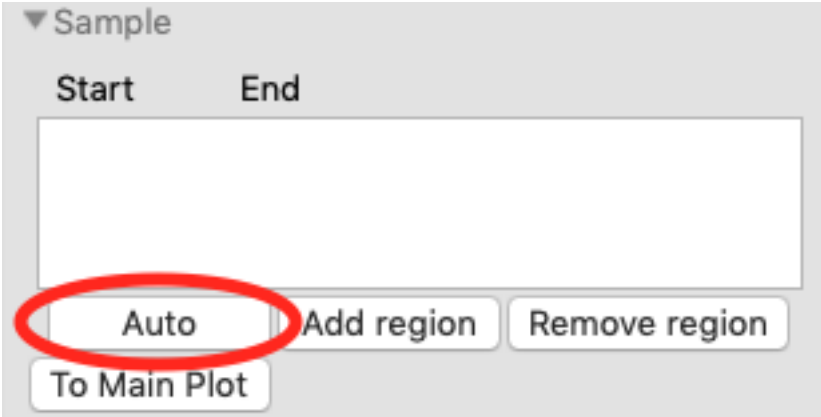
14. Once the calculation is finished, the window should automatically display the Subtracted plot. If it doesn't, click on the 'Subtracted' tab in the plot. On this plot there is a new Intensity vs. Frame # curve, representing the subtracted data. There is also a set of markers, showing one of the calculated parameters. By default the R_g displayed.. The calculated parameters are plotted on the right Y axis. You can show R_g, I(0), and MW calculated by the volume of correlation (Vc) and adjusted Porod volume (Vp) methods. Click on the 'Calculated value' menu to switch between the different displays.

    - *Try:* Show the R_g, MW (Vc), and MW (Vp). Notice that the MW estimate varies between the two different methods.
    - *Note:* You'll notice a region of roughly constant R_g across the peak. To either side there are regions with higher or lower R_g values. Some of these variations, particularly on the right side, are from scattering profiles near the edge of the peak with lower concentrations of sample, leading to more noise in determining the R_g values. There may also be some effects from the small peaks on the leading (left) side of the peak, and from the baseline mismatch between left and right sides of the peak.

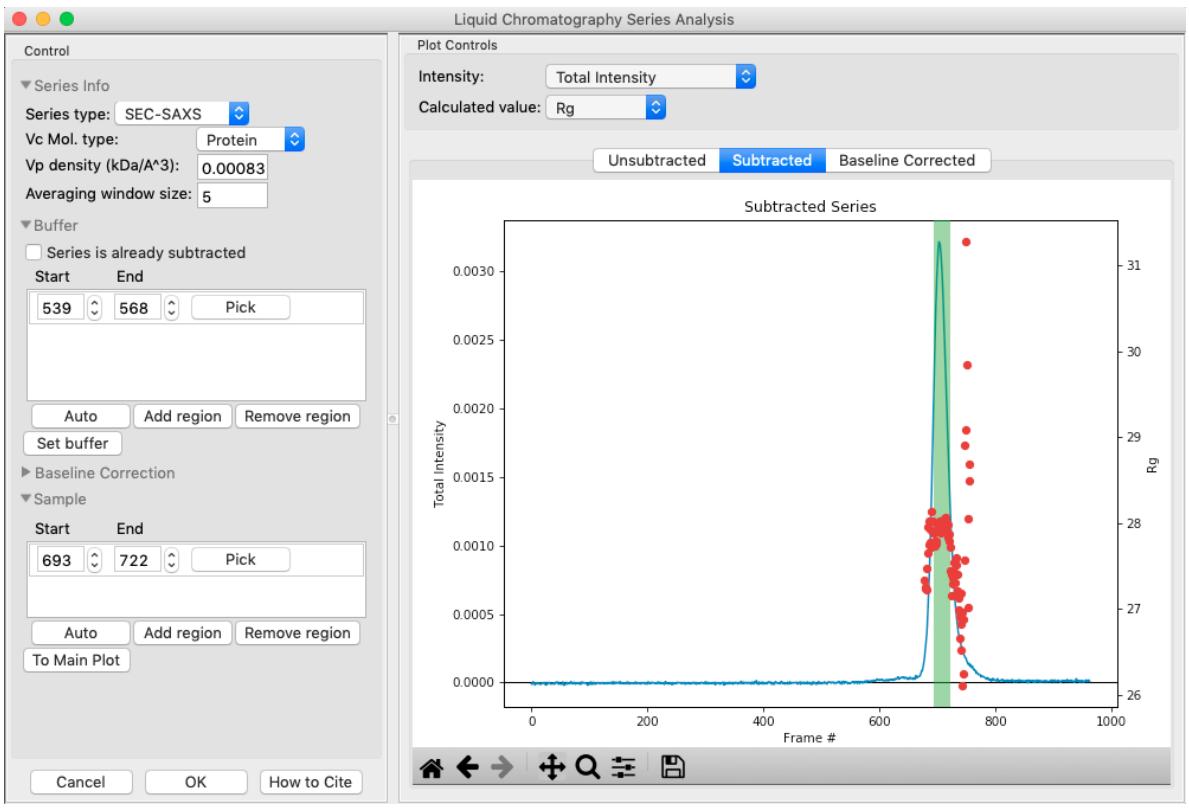


15. A monodisperse peak should display a region of flat R_g and MW near the center. Note that some spread on either edge can come from small shoulders of other components, bad buffer selection, or just the low signal to noise in the tails of the peak. Zoom in on the R_g and MW values across the peak to verify that these show a significant flat region.

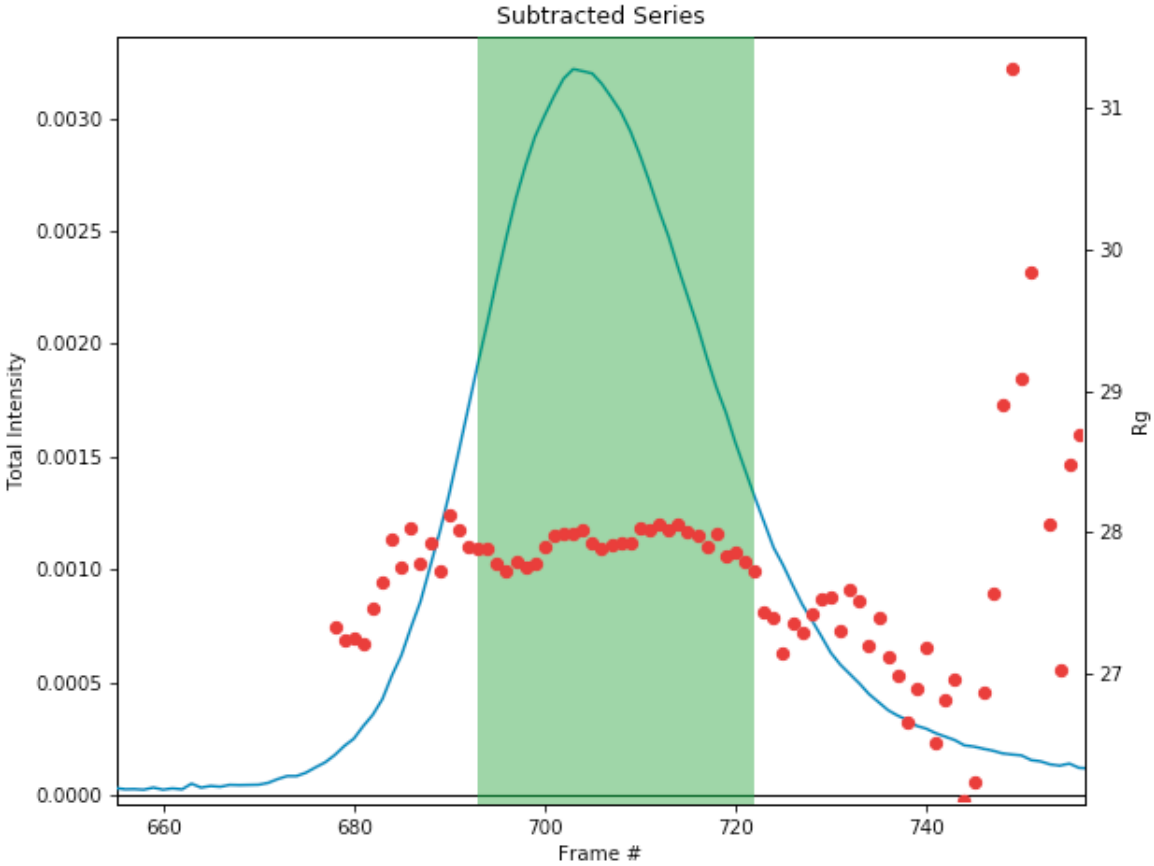
    RAW can automatically determine a good sample region (good being defined as monodisperse and excluding low signal to noise data). To do this, click the 'Auto' button in the Sample region.



- *Checkpoint:* You should see a sample range show up in the sample list, with defined start and end values. The region will be shown in green on the Subtracted plot.



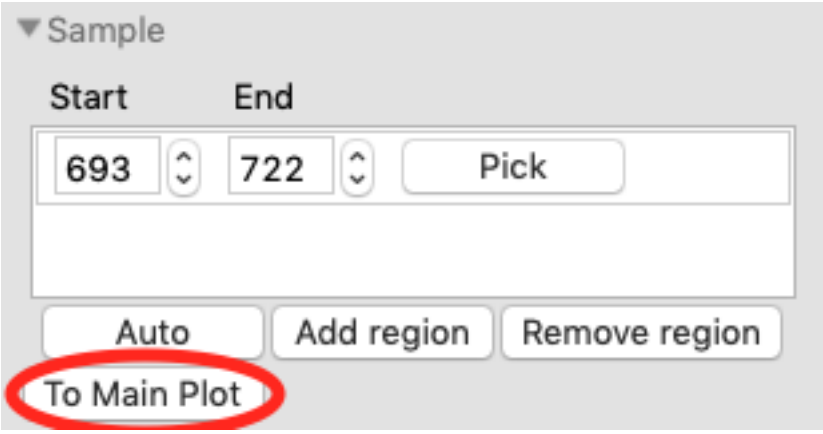
16. In the plot, zoom in on the peak region and verify that the R_g and MW seem flat in the selected sample range.



- *Tip:* You can manually adjust the sample region range in the same way as the buffer range, using the controls in the Start/End boxes.

17. Once you are satisfied with the region picked (should be 693-722), click the 'To Main Plot' button. This averages the selected region and sends the resulting average to RAW's Main Plot.

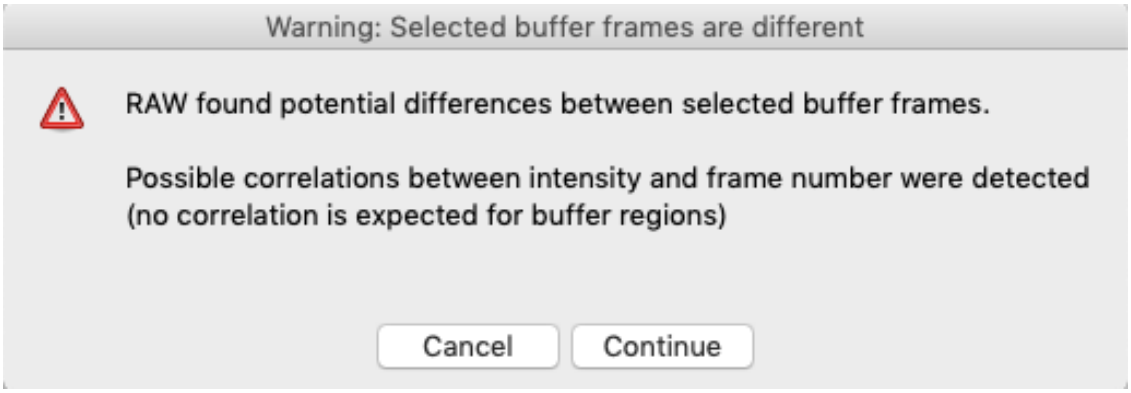
    - *Note:* RAW first averages the selected sample and buffer regions in the unsubtracted data, then subtracts. This avoids the possibility of correlated noise that would arise from averaging the subtracted files.



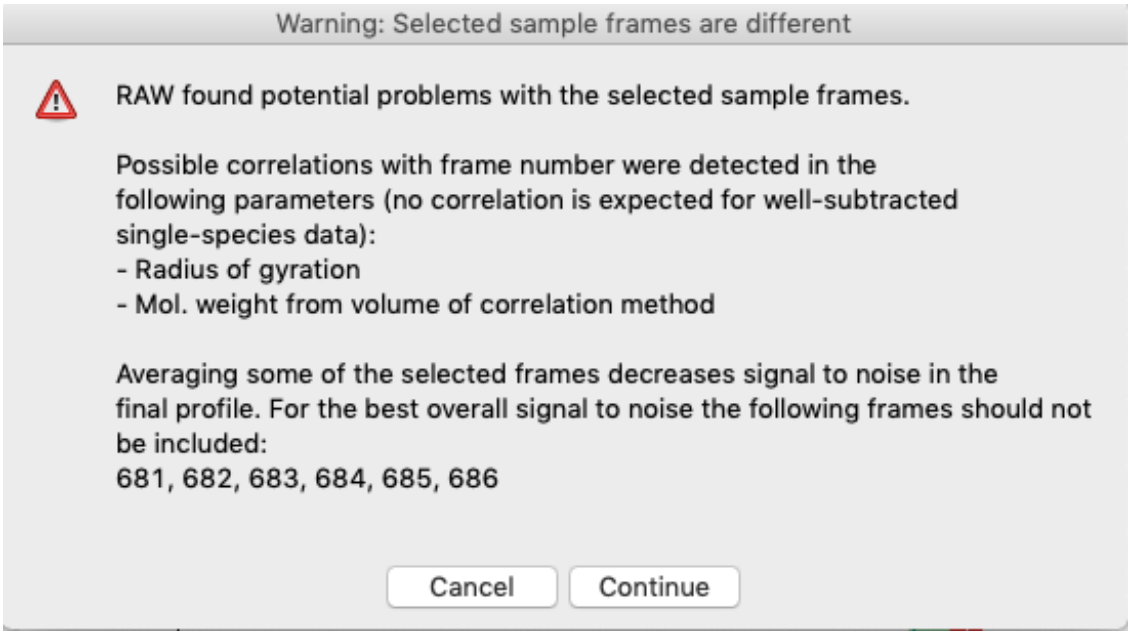
18. If you adjust the sample or buffer region in a way that could be problematic, RAW will warn you. Try this.

    - Adjust the Buffer end to include some of the leading edge peaks, such as ending at 600. You will want to click on the 'Unsubtracted' plot to see the buffer range. Then click 'Set Buffer'. You will see a warning

window telling you want might be wrong with the selected region. Click 'Cancel'.



- Adjust the Sample start to include some of the non-flat region, such as starting at 681. Then click 'To Main Plot'. You will see a warning window telling you want might be wrong with the selected region. Click 'Cancel'.



- *Note:* For buffer regions, RAW checks frame-wise similarity across the whole q range and at low and high q, correlations in intensity, and whether there are multiple singular values in the selected region.

  For sample regions, RAW checks frame-wise similarity across the whole q range and at low and high q, correlations in calculated values, whether there are multiple singular values in the selected region, and if some of the selected frames decrease the signal to noise of the average.

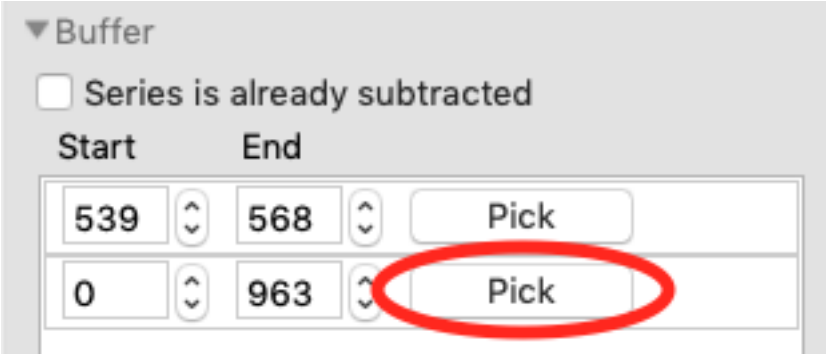
19. Click 'OK' to close the window and save your analysis results. If you reopen the LC analysis window you will see the buffer and sample regions you selected are remembered.

20. Click on the Main Plot tab and the Manipulation tab. You should see one scattering profile, the buffer subtracted data set you sent to the main plot. Carry out Guinier and MW analysis.

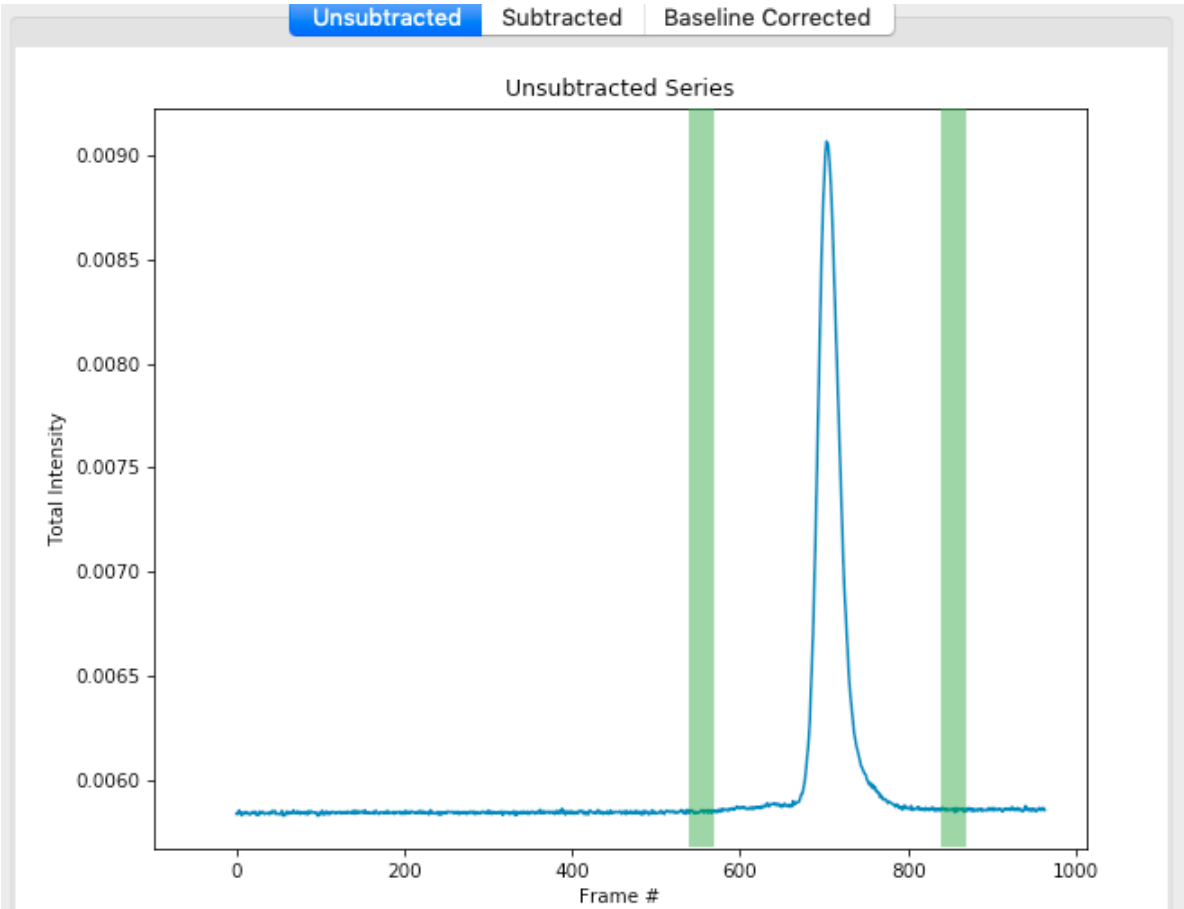
    - *Note:* The I(0) reference and absolute calibration will not be accurate for SEC-SAXS data, as the concentration is not accurately known.
    - *Question:* How does the R_g and MW you get from the averaged curve compare to what RAW found automatically for the peak?

- *Tip:* Make sure your plot axes are Log-Lin or Log-Log. Make sure that both plots are shown by clicking the 1/2 button at the bottom of the plot window.

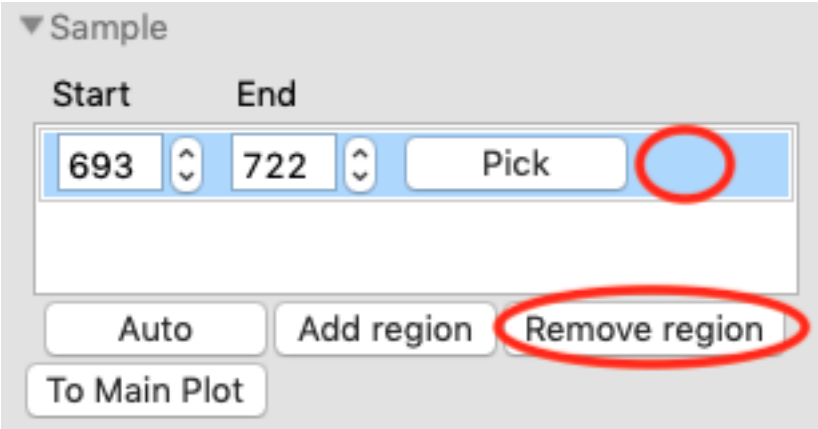
21. This particular dataset shows a small difference between initial and final buffer scattering profiles. A better scattering profile might be obtained by using buffer from both sides of the peak. To do so, start by reopening the LC Series Analysis panel.
22. Switch to showing the unsubtracted profiles by clicking on the 'Unsubtracted' plot tab.
23. Add a second buffer region by clicking the 'Add region' button.
24. For the second region, click the 'Pick' button.



25. Move your mouse across the plot. You will see a vertical green line moving with the mouse cursor. This represents the start of the buffer region. Click once to fix the start point where you click. Move the mouse further to the right and click again to fix the end point of the buffer region.



26. Once you are happy with the second buffer region, click 'Set buffer'. A range like ~840-870 is reasonable.
27. A warning window will pop up. In this case, we have purposefully chosen two buffer regions because they are different, so ignore the warning and click 'Continue'.
28. Remove the old sample region by clicking in the empty space to the right of the 'Pick' button to highlight it, and then clicking the 'Remove region' button.

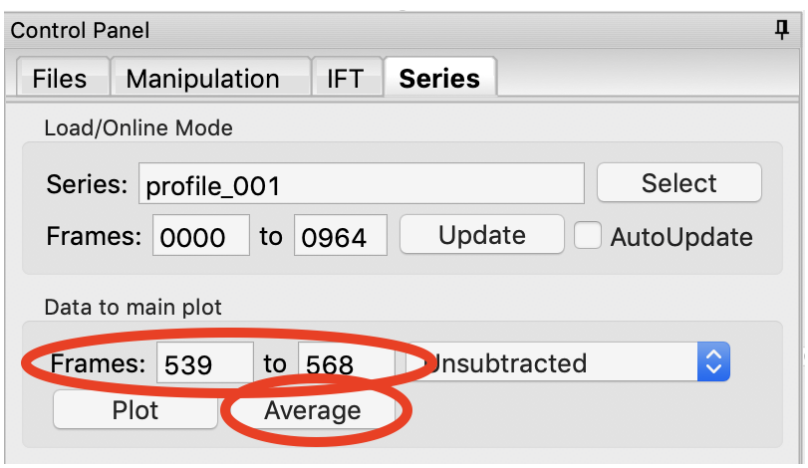


29. Click the 'Auto' button to automatically find a new sample region. Click the 'To Main Plot' button to send that new region to the main plot.
    - *Try:* You can see what the data subtracted by just the second buffer region looks like by removing the first buffer region, setting the buffer again, finding a new good sample region, and sending new region to the

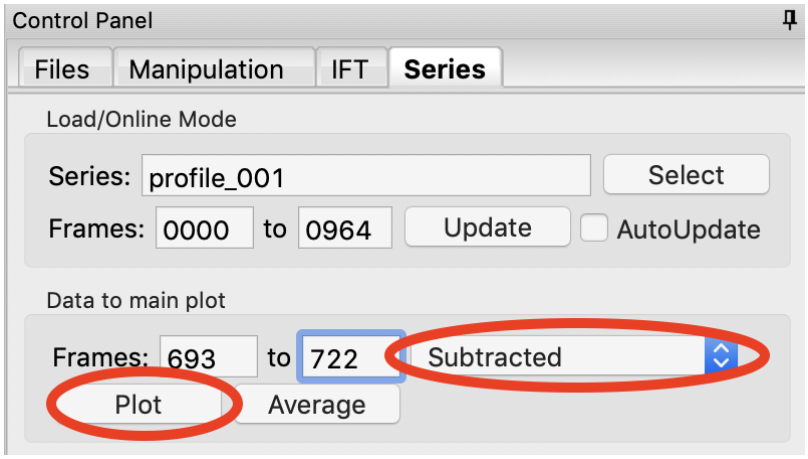
main plot.

30. Cancel out of the LC Series analysis window. This will not save the changes you made to the buffer and sample regions.
31. Carry out the R_g and MW analysis on the new curve. How does the scattering profile compare to the one that you generated using only buffer from before the peak?
    - *Tip:* You should see subtle but noticeable differences in the Guinier fit.
    - *Note:* An alternative approach to using several buffer regions is to use a single buffer region and apply a baseline correction. Both approaches have advantages and disadvantages. If you want to do EFA deconvolution, it is best to not use a baseline correction, however in other cases it will be more accurate as it doesn't assume a single average buffer across the peak.
32. Return to the Series control and plot panels.
33. If you want to look at either individual profiles or the average of a range of profiles you can send profiles to the main plot. To select which series curve to send profiles from, star the series curve of interest.

34. In the 'Data to main plot' section enter the frame range of interest. For this dataset, try the buffer range you selected: 539 to 568. Then click the 'Average' button. That will send the average buffer to the main plot.
    - *Try:* Send the average of the sample range you selected to the main plot (693 to 722), carry out the subtraction, and verify it's the same as the curve produced by the 'To Main Plot' button in the LC Series Analysis panel.
    - *Question:* When you send the sample average to the main plot you will get a warning that the profiles are different. Why?

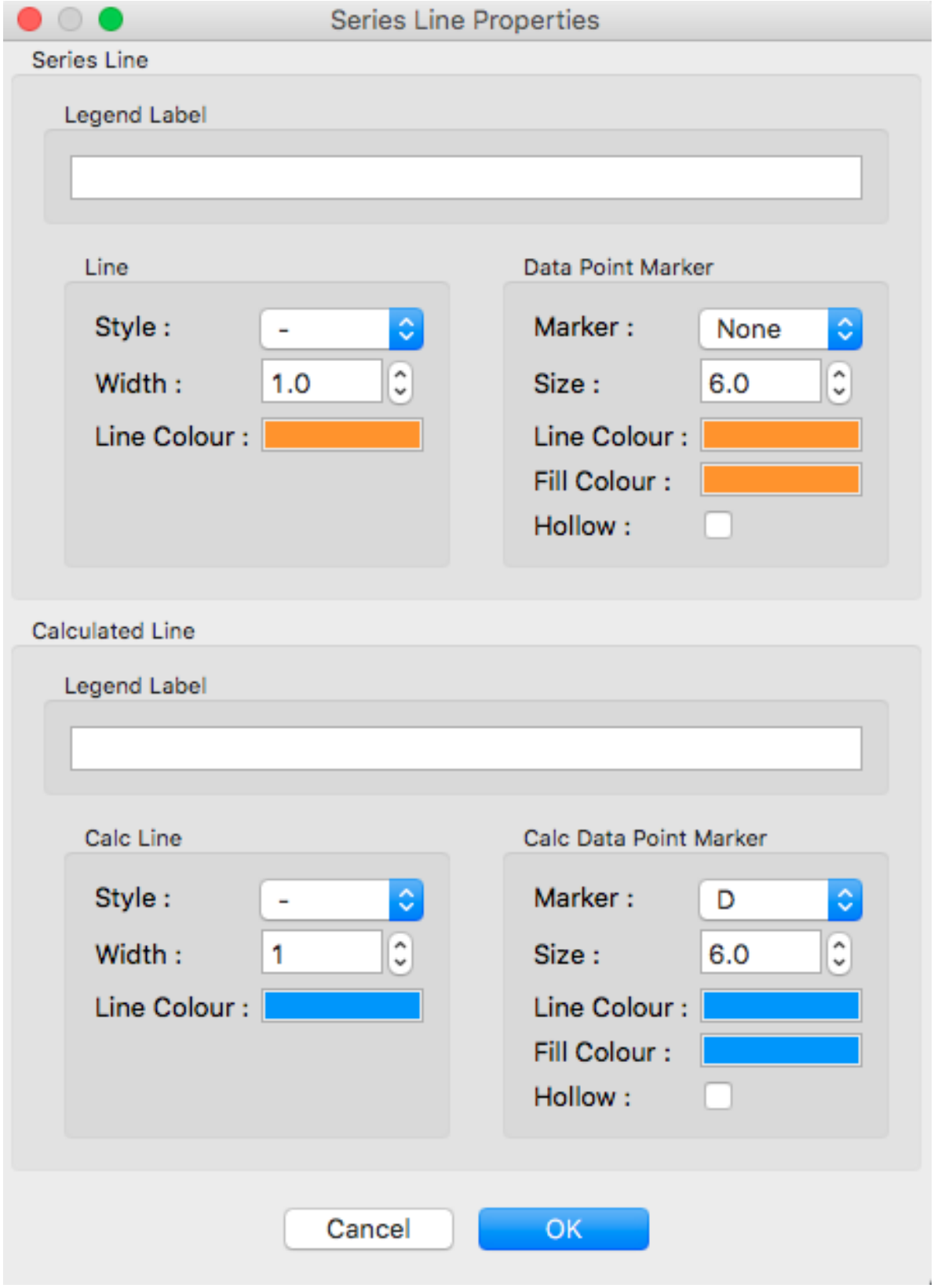


35. You can also send subtracted (or baseline corrected data) to the main plot. For the selected sample range, select the 'Subtracted' frames and send each individual profile to the plot using the 'Plot' button.
    - *Try:* Average these profiles and verify they match the subtracted profiles for this data set generated previously.



36. Click on the colored line next to the star in the Series control panel. In the line properties control panel this brings up, change the Calc Marker color to something different. Add a line to the Calc Markers by selecting line style '-' (solid), and adjust the line color to your liking.

    - *Tip:* You can do the same thing to adjust the colors of the scattering profiles in the Manipulation and IFT control tabs.



37. Load the Bovine Serum Albumin (BSA) SEC-SAXS data contained in the **sec_sample_2** data folder. Hide the first SEC-SAXS chromatograph.

38. Select a good buffer region, and calculate the R_g and MW across the peak for the BSA.
    - *Warning:* Don't forget to star the curve you want to set/update parameters for!
    - *Tip:* If you hover your mouse cursor over the info icon, you will see the buffer range and window size used to calculate the parameters.
    - *Question:* Is the BSA peak one species?
39. Find the useful region of the peak (constant R_g/MW), and send the buffer and sample data to the main plot. Carry out the standard R_g and MW analysis on the subtracted scattering profile. For BSA, we expect R_g ~28 Å and MW ~66 kDa.
40. In the Series control tab, right click on the name of BSA curve in the list. Select export data and save it in an appropriate location. This will save a CSV file with the frame number, integrated intensity, radius of gyration, molecular weight, filename for each frame number, and a few other items. This allows you to plot that data for publications, align it with the UV trace, or whatever else you want to do with it.
    - *Try:* Open the **.csv** file you just saved in Excel or Libre/Open Office Calc.
41. Select both items in the Series control panel list, and save them in the **sec_data** folder. This saves the Series plot data in a form that can be quickly loaded by RAW.
    - *Try:* Clear the Series data and then open one of your saved files from the Files tab using either the "Plot" or "Plot Series" button.

WAXS processing and merging

Several SAXS beamlines use two (or more) detectors to collect different q regions. For example, the MacCHESS G1 beamline uses dual Pilatus detectors to measure SAXS and WAXS from q ~0.008 – 0.75 $Å^{-1}$. The SAXS detector has q ~< 0.25 $Å^{-1}$ and the wide-angle scattering (WAXS) data has q >~ 0.25 $Å^{-1}$. All of the data that you have been working with so far has been SAXS data. Some experiments can make use of the WAXS data. In this part of the tutorial you will learn the basics of processing it.

1. Clear any data in RAW.
2. Navigate to the **standards_data** and load the **WAXS.cfg** file.
3. Plot the **lysbuff2 PIL3** and **lys2 PIL3** files. These are the images from the WAXS detector. Average these files and create a subtracted WAXS scattering profile.
    - *Tip:* Filenames should be **lysbuff2_<stuff>_PIL3_<stuff>.tiff** and **lys2_<stuff>_PIL3_<stuff>.tiff**.
4. Load the saved subtracted SAXS scattering profile for the lysozyme standards data.
    - *Note:* You should have saved it in the **standards_data** folder, and it is likely named **S_A_lys2_A9_17_001_0000.dat**.
5. Move the SAXS scattering profile you just loaded to the bottom plot by right clicking on it in the Manipulation list and selecting "Move to bottom plot."
6. The WAXS data is not on the same scale as the SAXS data. For this data, the known scale factor to apply is 0.000014 to the WAXS data.
    - *Note:* The scale factor can be calculated as the ratio of solid angles subtended by the pixels on the SAXS and WAXS detectors, plus any scale factor for absolute calibration and normalization included for one curve but not the other.
7. Star the WAXS data. Right click on the SAXS data and select Other Operations -> Merge. This will create a new merged scattering profile. The new file will have the prefix **M_** to indicate it is a merged file.

- *Tip:* If you can't see it, that's probably because it appeared on the upper plot, and is hidden by the very large intensities of the averaged WAXS files. Either try hiding those, or move the Merged curve to the lower plot.

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 main plot, the corresponding manipulation 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 go to Options -> Advanced Options -> Molecular weight, you can change the default type of molecule used in the MW estimation from the volume of correlation. This affects the default option selected in the MW window.
4. 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.
5. 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).
6. You can rename a curve by right clicking on the appropriate entry in the list and choosing rename.
7. 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.
8. 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.

4.3.3 Advanced processing

This section will guide you through advanced processing of files in RAW. It includes: pair-distance distribution analysis using GNOM and BIFT methods, ambiguity assessment for shape reconstructions, 3D reconstructions with bead models and electron density, SEC-SAXS data deconvolution using singular value decomposition (SVD) and evolving factor analysis (EFA), and baseline corrections for SEC-SAXS data. It refers to the *RAW tutorial data*.

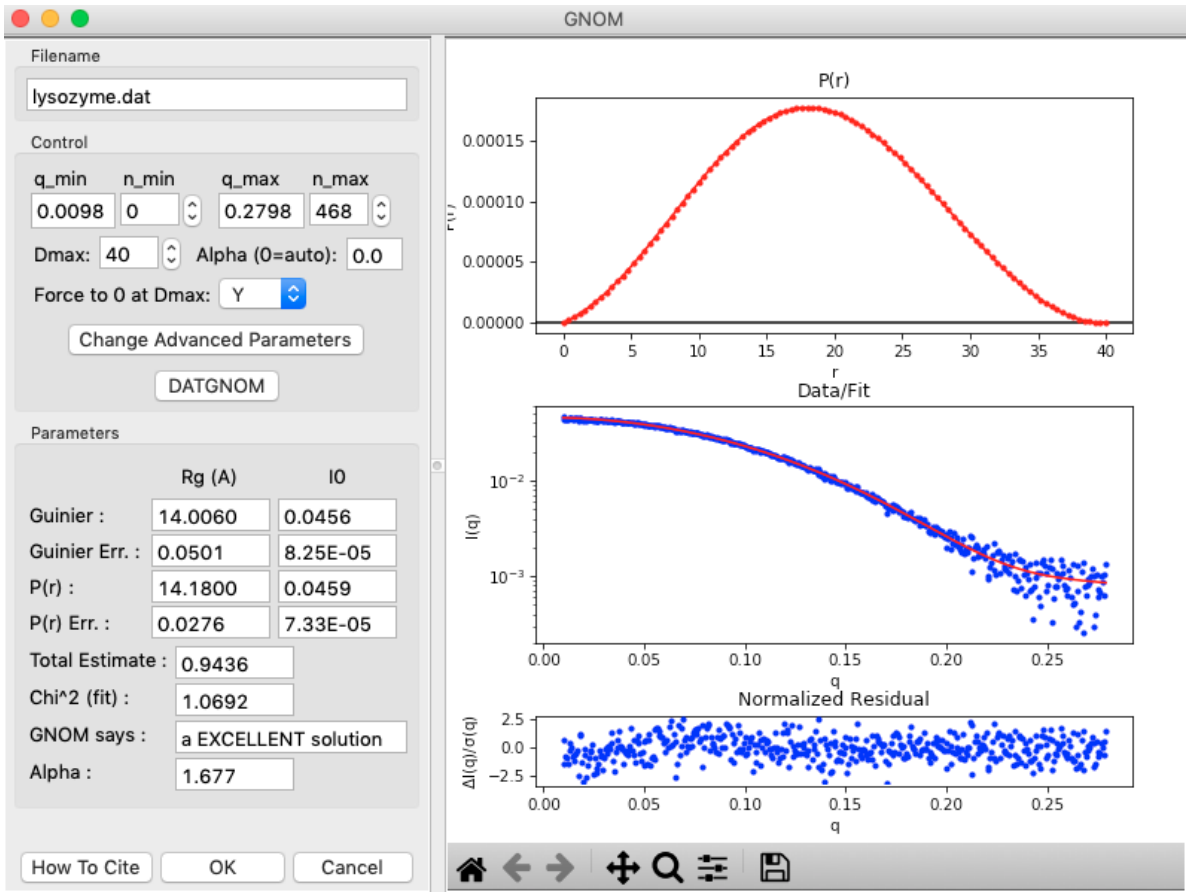
Select a section below to view the tutorial, or use the next and back buttons at the bottom of the page to navigate through it in order.

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 you need *ATSAS installed* to do this part of the tutorial.

1. Open RAW. The install instructions contain information on installing and running RAW. If RAW is already open, clear all loaded data in RAW.
2. Open the **lysozyme.dat** file in the **Tutorial_Data/reconstruction_data** folder.

3. Right click on the lysozyme profile in the Manipulation list and select "IFT (GNOM)".
   - *Note:* RAW will automatically try to find an appropriate maximum dimension (D_{max}) by running the DATGNOM program from the ATSAS software package.
   - *Troubleshooting:* If you do not have the GNOM option in the right click menu, RAW does not know where your ATSAS programs are 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.



4. 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).
   - *Note:* The fit line is the Fourier transform of the P(r) function, and is also called the regularized intensity.
5. 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.
   - *Tip:* The Guinier and P(r) R_g and I(0) values should agree well. 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.
6. Vary the D_{max} value up and down in the range of 30-50 in steps of 1. Observe what happens to the P(r) and the quality of the solution.
   - *Note:* D_{max} is in units of Å.
   - *Tip:* Recall that we want the following qualities in a P(R) function:

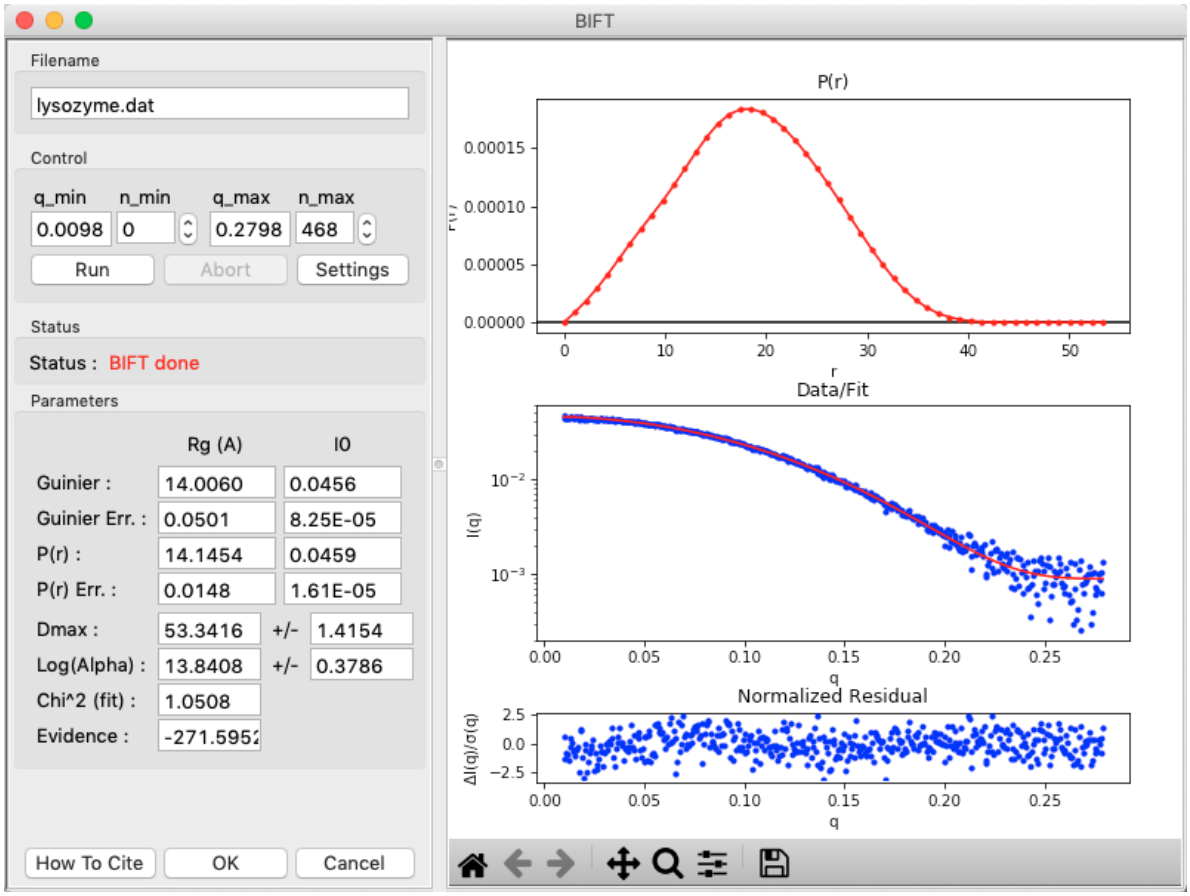
    1. No negative values.
    2. A gentle approach to D_{max} (not suddenly forced down).
    3. Minimal oscillation.

7. Return the D_{max} value to that found by DATGNOM by clicking the "DATGNOM" button. D_{max} should be 40. 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".
    - *Try:* Vary D_{max} with this option turned off.
8. Reset it so that the P(r) function is again being forced to zero at D_{max}.
9. Set the D_{max} back to 40, 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 **lysozyme.out** item in the **reconstruction_data** folder.
    - *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.

Pair-distance distribution analysis – BIFT in RAW

RAW also has a built in method for determining the P(r) function using a Bayesian IFT method. This has the advantage of only have one possible solution. More information on this method can be found in the RAW paper and manual and references therein.

1. Right click on the lysozyme profile in the Manipulation list you loaded *previously*. Select "IFT (BIFT)" from the resulting menu.



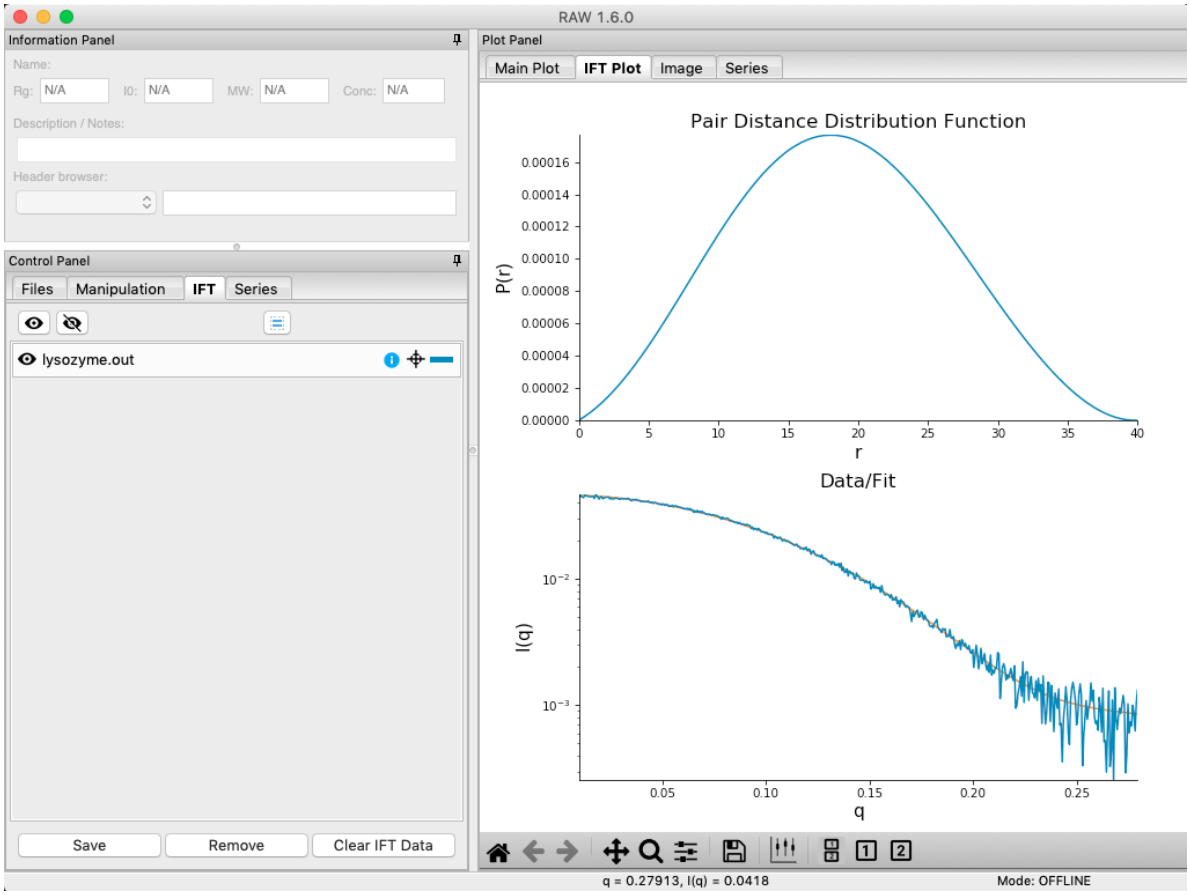
2. The BIFT panel has plots on the right. These show the P(r) function (top panel), the data (middle panel, blue points) and the fit line (middle panel, red line), and the fit residual (bottom panel).
3. On the left of the BIFT panel are the controls and the resulting parameters. Note that in this case you do not control the D_{max} value, the BIFT method finds that for you automatically. Because BIFT can take some time to run, if you change the *q* range for the data set you have to click the 'Run' button to run BIFT again.
4. Note that for this dataset, BIFT has overestimated the D_{max} value, however by looking at the P(r) function you can quickly see that the true D_{max} is somewhere near 40, in good agreement with what we got from GNOM.
5. Click OK to exit the BIFT window. This saves the results into the RAW IFT panel.
6. Click on the IFT Control and Plot tabs. This will display the BIFT output you just generated. Save the **lysozyme.ift** item in the **reconstruction_data** folder.

Note: As of now, BIFT output from RAW is not compatible with DAMMIF or other ATSAS programs. However, it is compatible with electron density determination via DENSS.

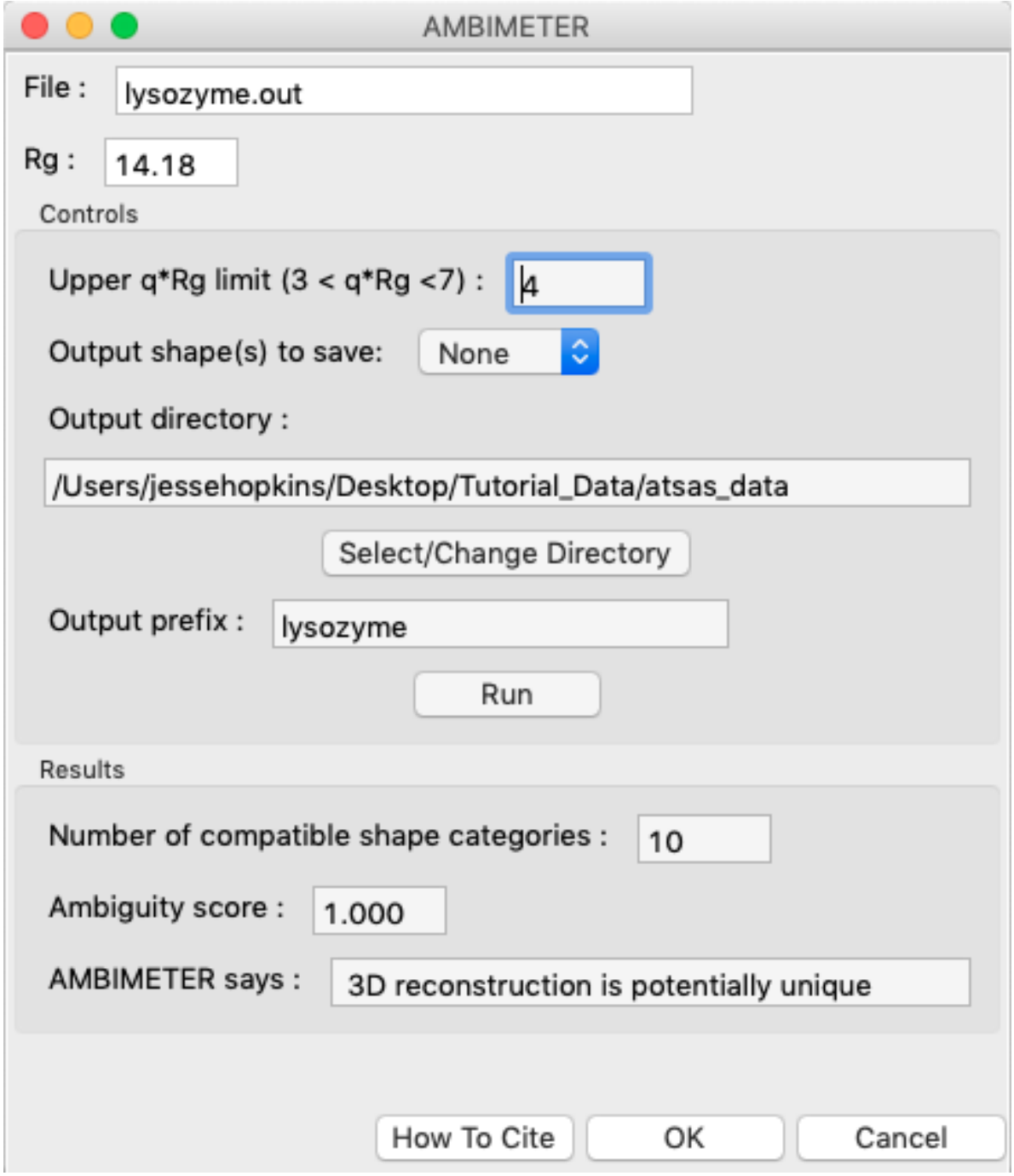
Assessing ambiguity of 3D shape information - AMBIMETER in RAW

It is impossible to determine a provably unique three-dimensional shape from a scattering profile. This makes it important to determine what degree of ambiguity might be expected in our reconstructions. The program AMBIMETER from the ATSAS package does this by comparing the measured scattering profile to a library of scattering profiles from relatively simple shapes. The more possible shapes that could have generated the scattering profile, the greater ambiguity there will be in the reconstruction. We will use RAW to run AMBIMETER. Note that you need *ATSAS installed* to do this part of the tutorial.

1. Clear all of the data in RAW. Load the **lysozyme.out** file that you saved in the **reconstruction_data** folder in a previous part of the tutorial.
   - *Note:* If you haven't done the previous part of the tutorial, or forgot to save the results, you can find the **lysozyme.out** file in the **reconstruction_data/lysozyme_complete** folder.



2. Right click on the **lysozyme.out** item in the IFT list. Select the "AMBIMETER" option.
3. The new window will show the results of AMBIMETER. It includes the number of shape categories that are compatible with the scattering profile, the ambiguity score (a-score) (log base 10 of the number of shape categories), and the AMBIMETER interpretation of whether or not you can obtain a unique 3D reconstruction.
   - According to the original paper, "an a-score below 1.5 practically guarantees a unique ab initio shape determination, whereas when the a-score is in the range 1.5–2.5 care should be taken, perhaps involving cluster analysis, and for a-scores exceeding 2.5 unambiguous reconstruction without restrictions (for example, on symmetry and/or anisometry) is highly unlikely."
   - *Note:* AMBIMETER can also save the compatible shapes (either all or just the best fit). You can do that by selecting the output shapes to save, giving it a save directory, and clicking run. We won't be using those shapes in this tutorial.



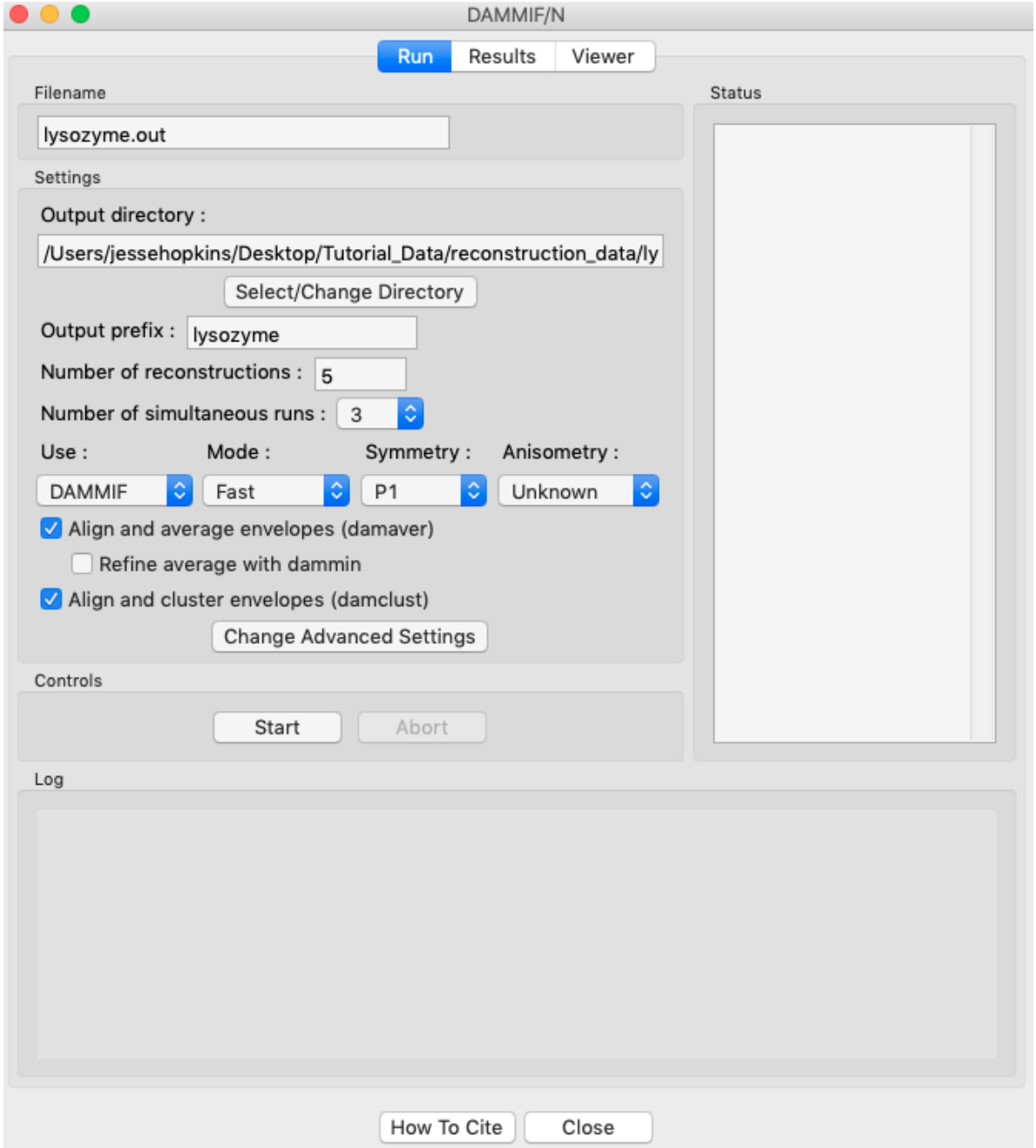
4. Click "OK" to exit the AMBIMETER window.

3D reconstruction with bead models – DAMMIF/N and DAMAVER in RAW

Shape reconstruction in SAXS is typically done using bead models (also called dummy atom models, or DAMs). The most common program used to generate these shapes is DAMMIF (and, to a lesser degree, DAMMIN) from the ATSAS package. We will use RAW to run DAMMIF/N. Because the shape reconstruction is not unique, a number of distinct reconstructions are generated, and then a consensus shape is made from the average of these reconstructions. The program DAMAVER from the ATSAS package is the most commonly used program for building consensus

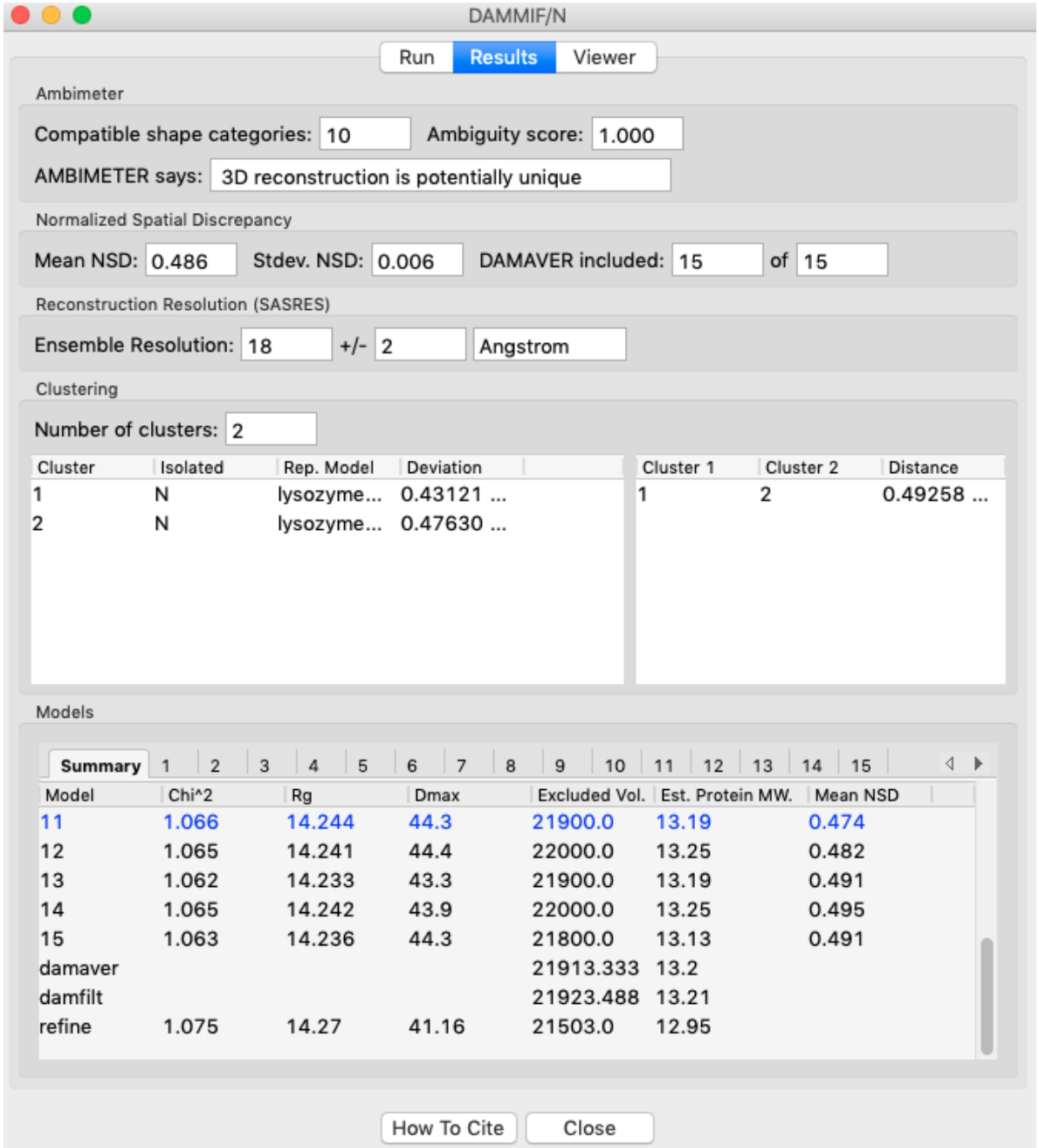
shapes. Note that you need *ATSAS installed* to do this part of the tutorial.

1. Clear all of the data in RAW. Load the **lysozyme.out** file that you saved in the **reconstruction_data** folder in a previous part of the tutorial.
   - *Note:* If you haven't done the previous part of the tutorial, or forgot to save the results, you can find the **lysozyme.out** file in the **reconstruction_data/lysozyme_complete** folder.
2. Right click on the **lysozyme.out** item in the IFT list. Select the "Bead Model (DAMMIF/N)" option.
3. Running DAMMIF generates a lot of files. Click the "Select/Change Directory" button, make a new folder in the **reconstruction_data** directory called **lysozyme_dammif** and select that folder.
4. Change the number of reconstructions to 5.
   - *Note:* It is generally recommended that you do 15-20 reconstructions. However, for the purposes of this tutorial, 5 are enough.
   - *Note:* For final reconstructions for a paper, DAMMIF should be run in Slow mode. For this tutorial, or for obtaining an initial quick look at results, Fast mode is fine.
5. Uncheck the "Refine average with dammin" checkbox.
   - *Note:* For final reconstructions for a paper, DAMMIN refinement should be done. However, it is quite slow, so for the purposes of this tutorial we won't do it.
6. If it's not already checked, check the "Align and cluster envelopes (damclust)" checkbox.



7. Click the "Start" button.
   - *Note:* The status panel will show you the overall status of the reconstructions. You can look at the detailed status of each run by clicking the appropriate tab in the log panel.
8. Note that by default the envelopes are aligned and averaged using DAMAVER, and then the aligned and averaged profile is refined using DAMMIN. Clustering analysis is also done by default.
   - Some settings are accessible in the panel, and all settings can be changed in the advanced settings panel.
9. Wait for all of the DAMMIF runs and DAMAVER to finish. Depending on the speed of your computer this could take a bit.
10. Once the reconstructions are finished, the window should automatically switch to the results tab. If it doesn't,

click on the results tab.

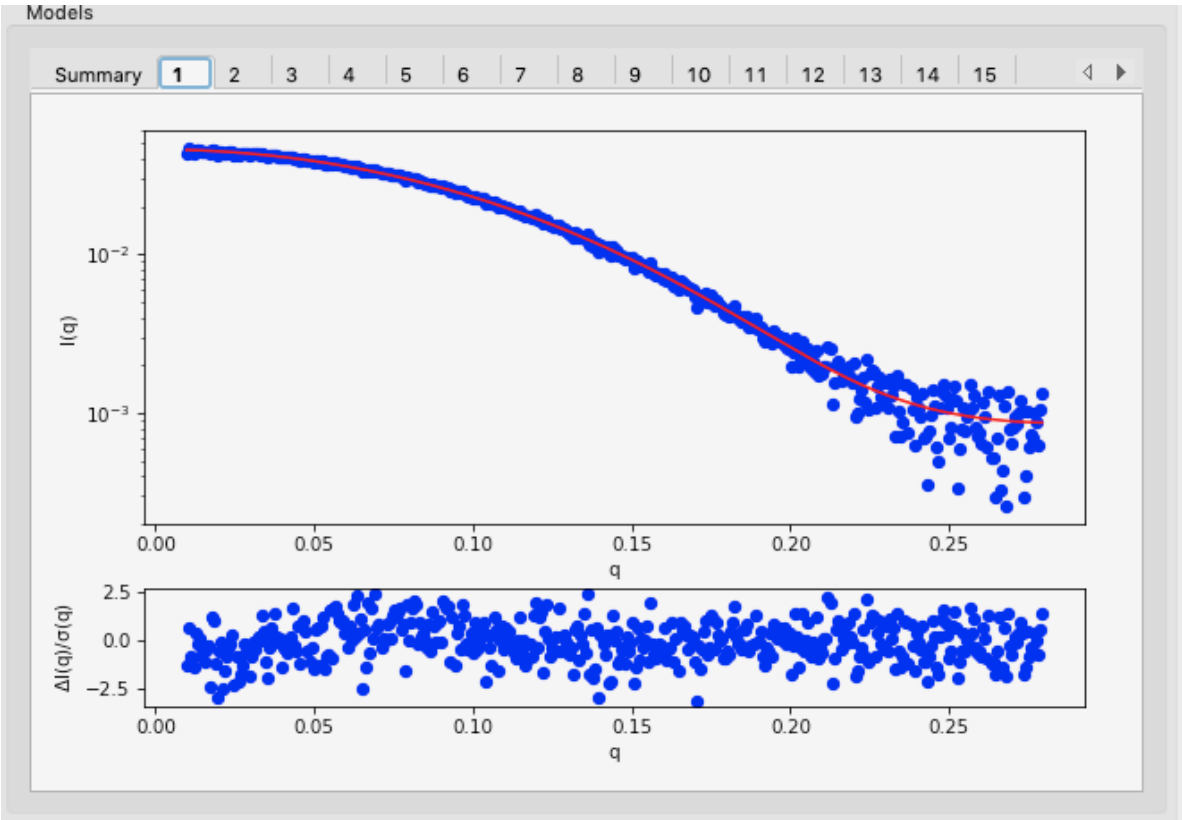


11. The results panel summarizes the results of the reconstruction run. At the top of the panel there is the ambimeter evaluation of how ambiguous the reconstructions might be (see previous tutorial section). If DAMAVER was run, there are results from the normalized spatial discrepancy (NSD), showing the mean and standard deviation of the NSD, as well as how many of the reconstructions were included in the average. If DAMAVER was run on 3 or more reconstructions, and ATSAS >=2.8.0 is installed, there will be the output of SASRES which provides information on the resolution of the reconstruction. If DAMCLUST was run the number of clusters, information on each cluster, and the distance between each cluster is shown.

12. Information on each individual model is shown at the bottom. The summary tab gives the model chi squared, R_g, D_{max}, excluded volume, molecular weight estimated from the excluded volume, and, if appropriate, mean NSD of the model.

- Any models rejected from the average by DAMAVER will be shown in red in the summary tab list.
- *Tip:* The model highlighted in blue in the summary tab is the 'most probable' model, this can be used as your final bead model instead of doing a dammin refinement.
- *Note:* Above we show the results for 15 runs instead of the 5 in the tutorial.

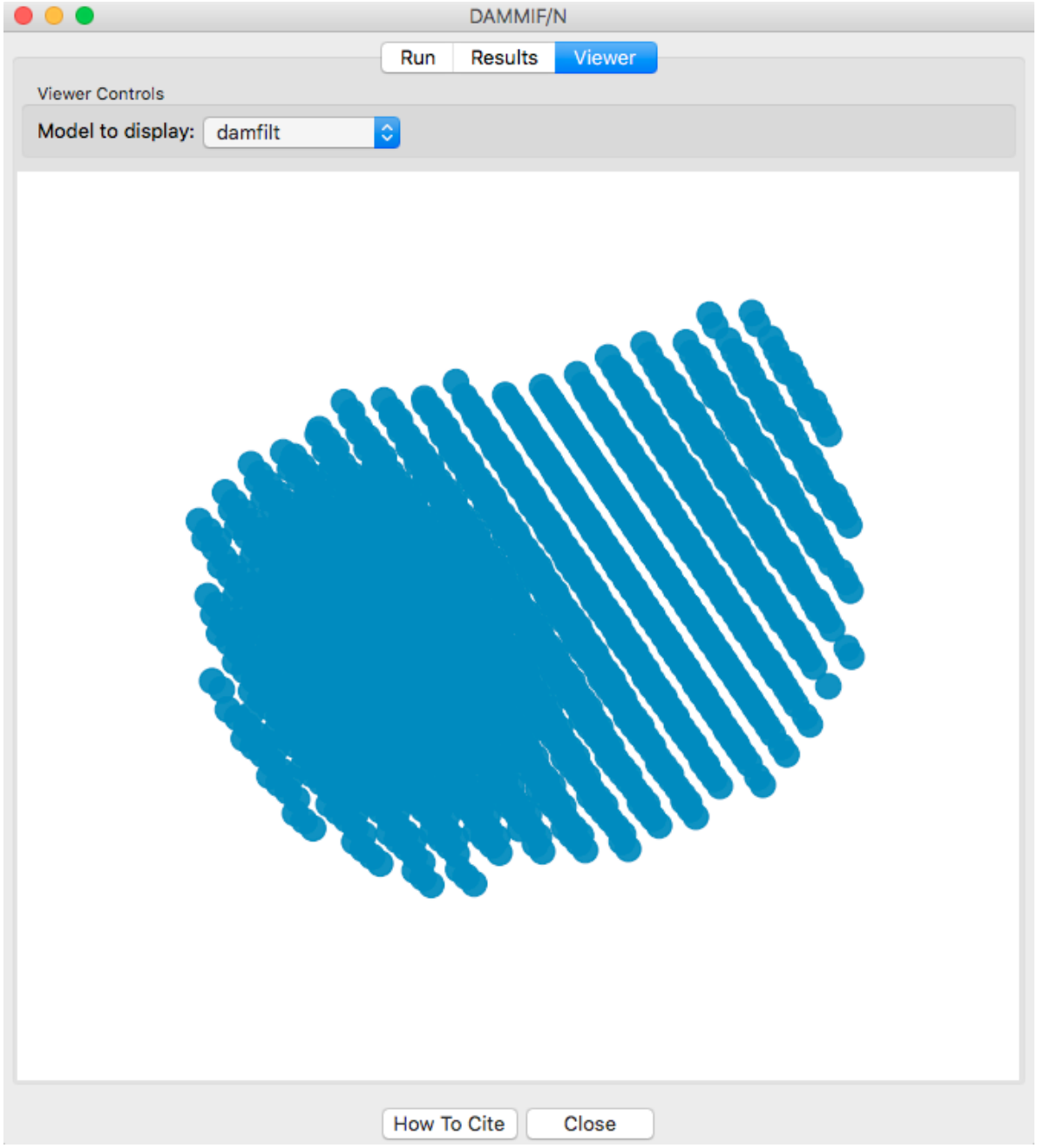
13. Also, each individual model has a tab which shows the data, the model fit, and the residuals.



14. The results summary shown in Summary tab is automatically saved as a **<prefix>_dammif_results.csv** csv file, e.g. for this data as **lysozyme_dammif_results.csv**. All the plots shown on the individual model tabs are automatically saved as a multi-page pdf file with the same name.

15. Click on the Viewer tab to open the model viewer.

    - *Note:* The model viewer is intended for a fast first look at the results. It is not currently up to the standards of a program like pyMOL.



16. Click and drag the model to spin it.
    - Note: For lysozyme, it should look more or less like a flattened sphere.
17. Right click and drag the model to zoom in and out.
18. Use the "Model to display" menu in the Viewer Controls box to change which reconstruction is displayed.
19. Click the "Close" button when you are finished looking at the results and reconstructions.
20. The results from individual DAMMIF runs are saved in the selected output folder with the name **<prefix>_xx**, where *xx* is the run number: 01, 02, etc. For this tutorial, that would be **lysozyme_01**, **lysozyme_02**, and so on. The different files produced are described in the DAMMIF manual.
    - *Note:* Generally, the file of interest is the **-1.pdb** file, in this case **lysozyme_01-1.pdb**, **lysozyme_02-**

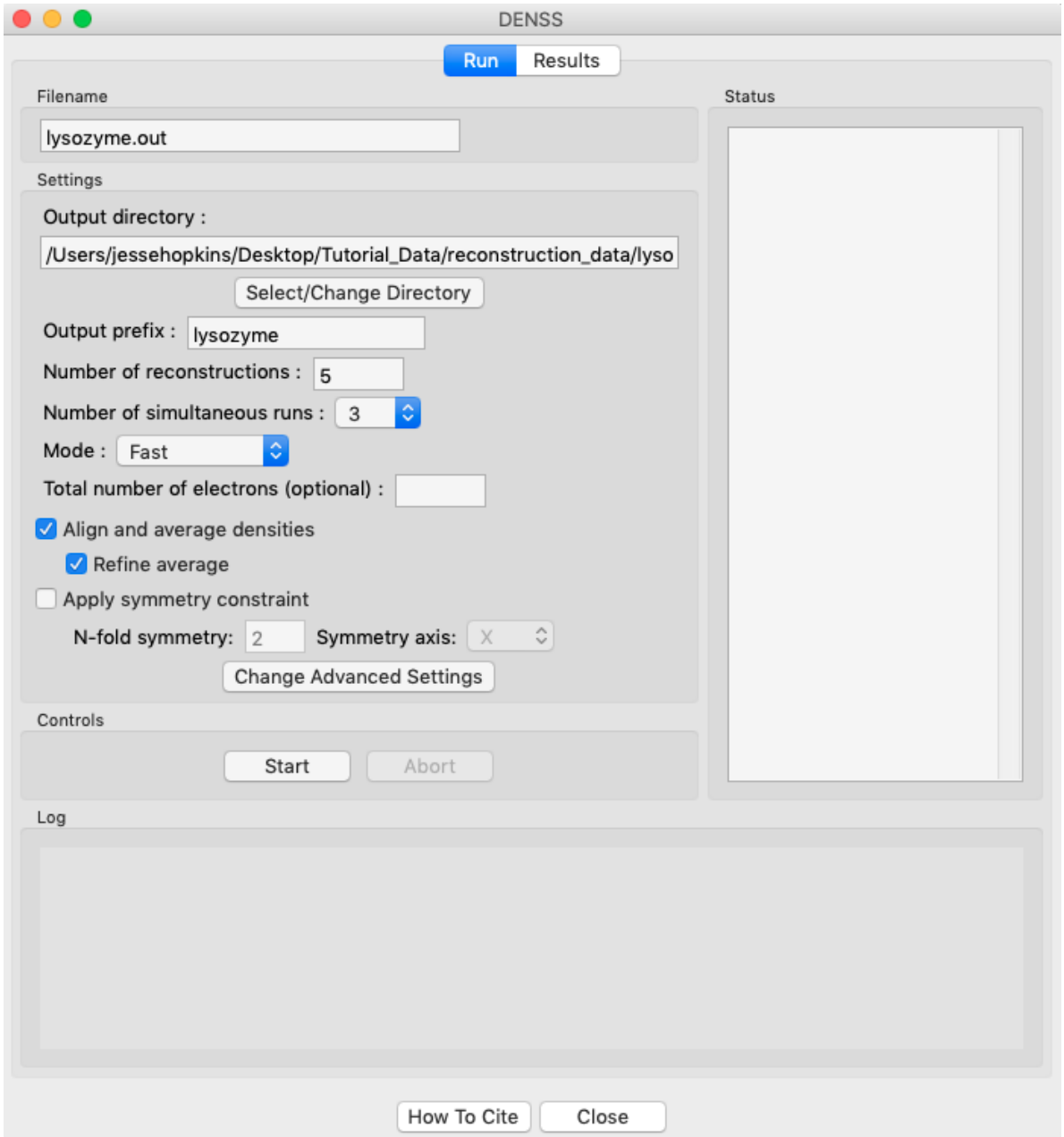
1.pdb, etc.

21. If averaging was done with DAMAVER, the results are saved in the selected output folder with the given prefix, in this case **lysozyme**. The output files generated are described in the DAMAVER manual.
    - *Note:* Generally, the files of interest are the generated pdbs: **<prefix>_damaver.pdb** and **<prefix>_damfilt.pdb**. For this tutorial, those would be **lysozyme_damaver.pdb** and **lysozyme_damfilt.pdb**.
22. If clustering was done with DAMCLUST, the results are saved in the selected output folder with the given prefix (for this tutorial, **lysozyme**). The files generated are described in the DAMCLUST manual.
23. If refinement was done with DAMMIN, the results are saved in the selected output folder as **refine_<prefix>**, e.g. for this tutorial **refine_lysozyme**. The files generated are described in the DAMMIN manual.
    - *Note:* Generally, the file of interest is the **-1.pdb** file, in this case **refine_lysozyme-1.pdb**.

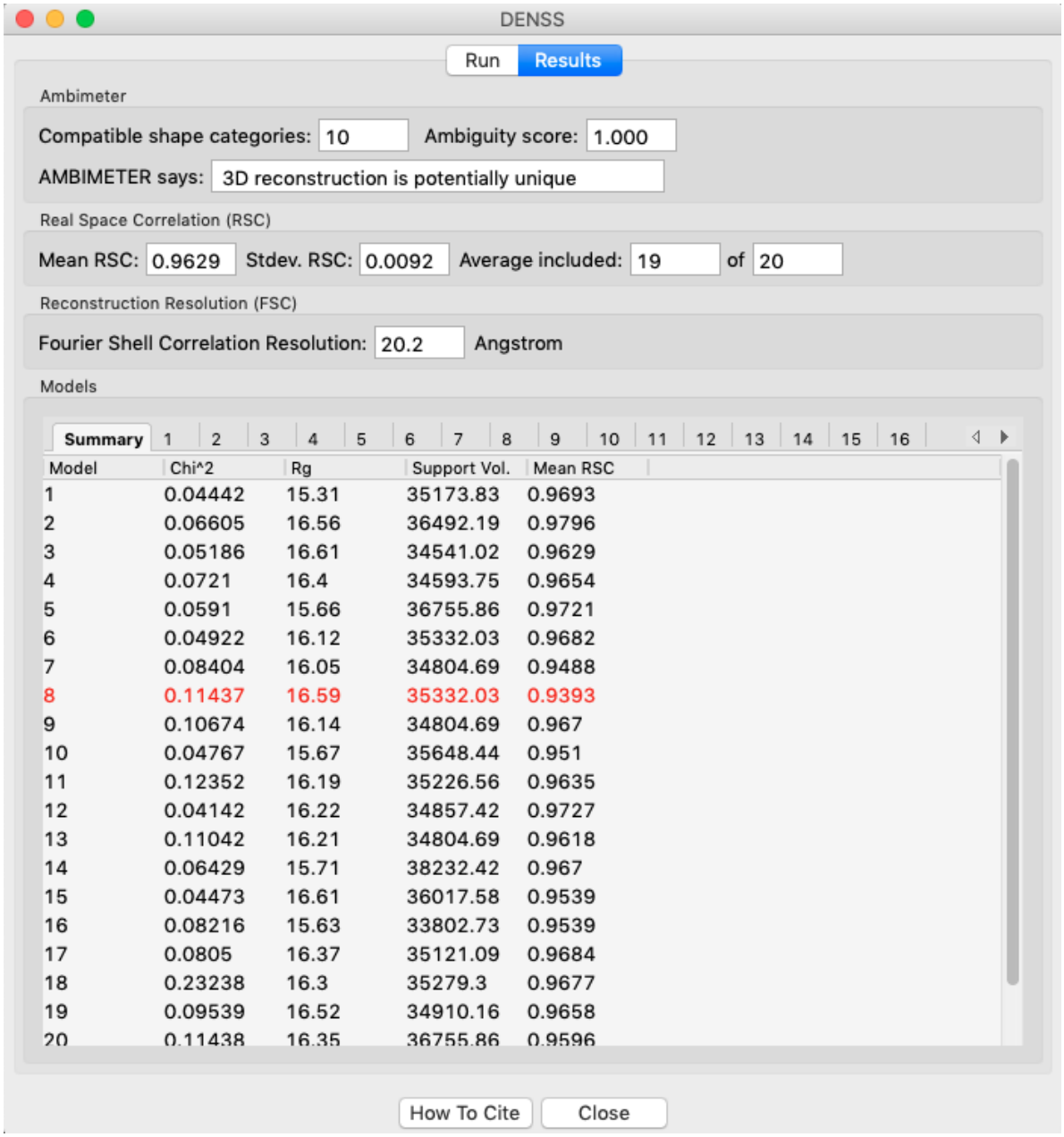
3D reconstruction with electron density – DENSS in RAW

A new, exciting method for doing 3D shape reconstructions in SAXS yields actual electron density, rather than bead models. There are many potential advantages to this, but one significant one is easy handling of systems like RNA-Protein complexes or membrane proteins surrounded by lipids or detergents, which have more than one electron density. Bead models typically only have two (molecule and solvent) or three bead densities, and so typically fail to reconstruct these complex objects. DENSS has been fully implemented in RAW and will be used to reconstruct these electron densities.

1. Clear all of the data in RAW. Load the **lysozyme.out** file that you saved in the **reconstruction_data** folder in a previous part of the tutorial.
    - *Note:* If you haven't done the previous part of the tutorial, or forgot to save the results, you can find the **lysozyme.out** file in the **reconstruction_data/lysozyme_complete** folder.
2. Right click on the **lysozyme.out** item in the IFT list. Select the "Electron Density (DENSS)" option.
3. Running DENSS generates a lot of files. Click the "Select/Change Directory" button, make a new folder in the **reconstruction_data** directory called **lysozyme_denss** and select that folder.
4. Change the number of reconstructions to 5 and the mode to Fast.
    - *Note:* It is generally recommended that you do at least 20 reconstructions. However, for the purposes of this tutorial, 5 are enough.
    - *Note:* For final reconstructions for a paper, DENSS should be run in Slow mode. For this tutorial, or for obtaining an initial quick look at results, Fast mode is fine.



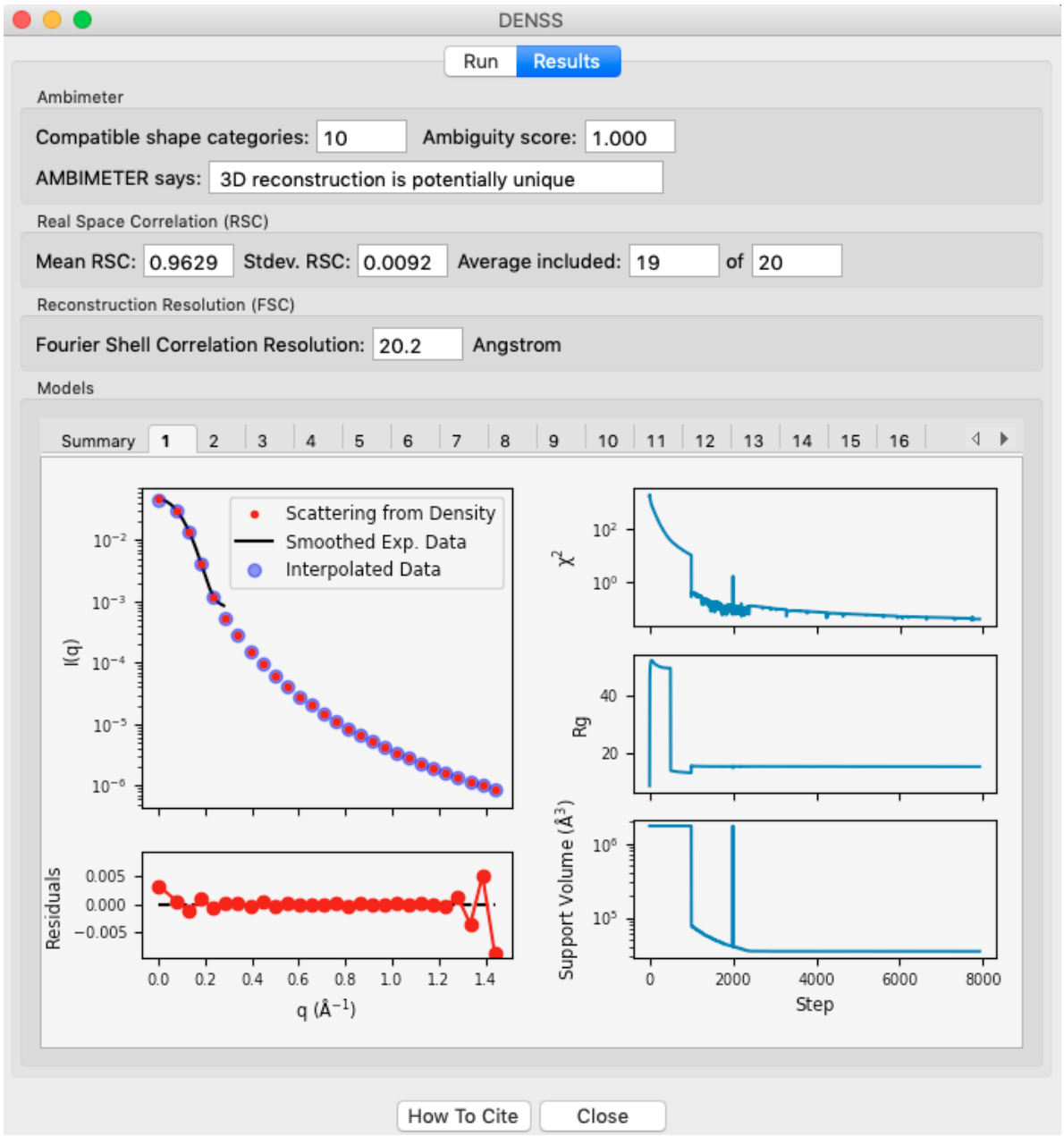
5. Click the "Start" button.
   - *Note:* The status panel will show you the overall status of the reconstructions. You can look at the detailed status of each run by clicking the appropriate tab in the log panel.
6. Note that by default the densities are aligned and averaged, including enantiomer filtering, and a refined density is created from the average.
7. Wait for all of the DENSS runs and averaging to finish. Depending on the speed of your computer this could take a bit.
8. Once the reconstructions are finished, the window should automatically switch to the results tab. If it doesn't, click on the results tab.



Model	Chi^2	Rg	Support Vol.	Mean RSC
1	0.04442	15.31	35173.83	0.9693
2	0.06605	16.56	36492.19	0.9796
3	0.05186	16.61	34541.02	0.9629
4	0.0721	16.4	34593.75	0.9654
5	0.0591	15.66	36755.86	0.9721
6	0.04922	16.12	35332.03	0.9682
7	0.08404	16.05	34804.69	0.9488
8	0.11437	16.59	35332.03	0.9393
9	0.10674	16.14	34804.69	0.967
10	0.04767	15.67	35648.44	0.951
11	0.12352	16.19	35226.56	0.9635
12	0.04142	16.22	34857.42	0.9727
13	0.11042	16.21	34804.69	0.9618
14	0.06429	15.71	38232.42	0.967
15	0.04473	16.61	36017.58	0.9539
16	0.08216	15.63	33802.73	0.9539
17	0.0805	16.37	35121.09	0.9684
18	0.23238	16.3	35279.3	0.9677
19	0.09539	16.52	34910.16	0.9658
20	0.11438	16.35	36755.86	0.9596

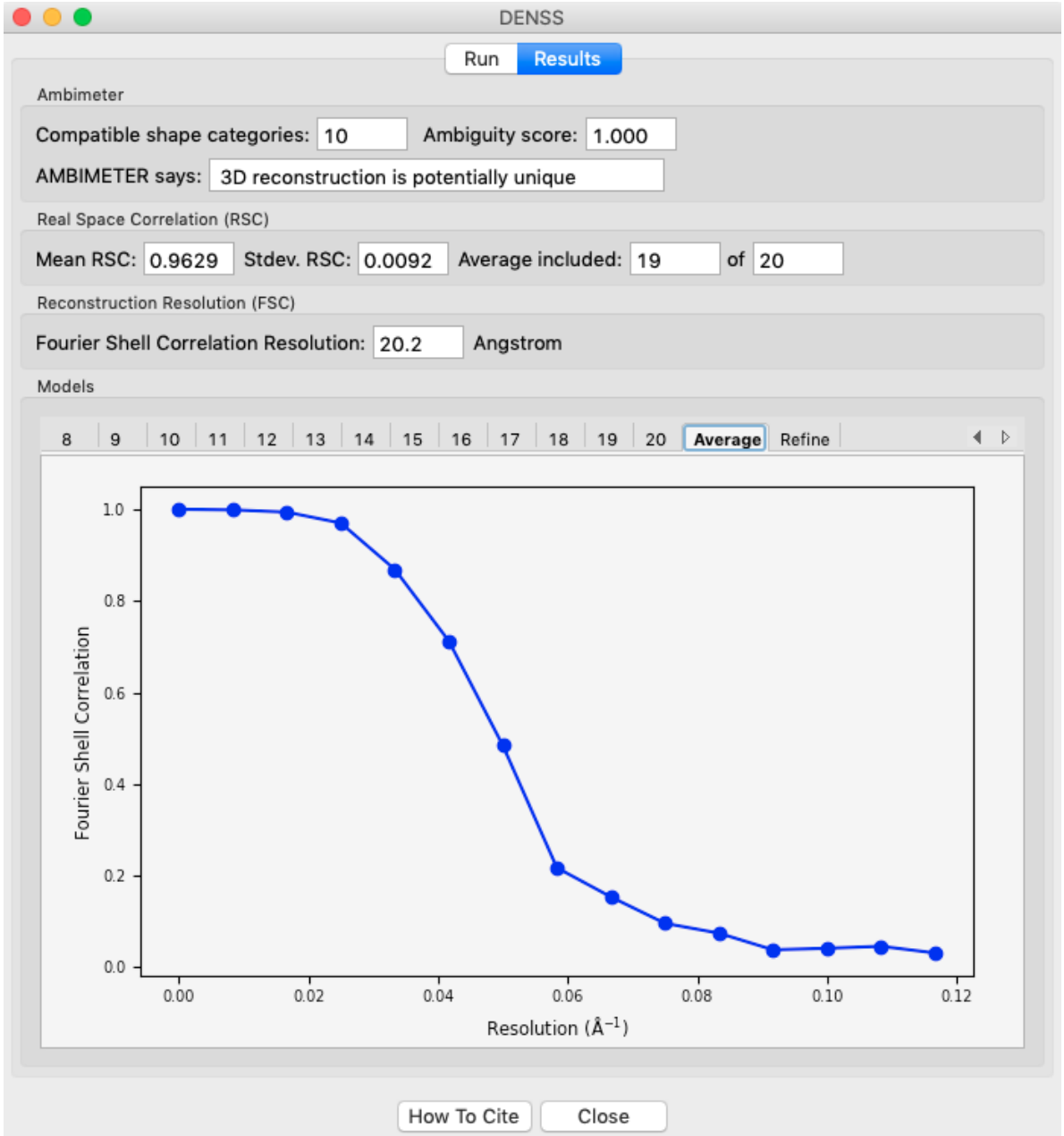
9. The results panel summarizes the results of the reconstruction runs. At the top of the panel there is the ambimeter evaluation of how ambiguous the reconstructions might be (see earlier tutorial section). If averaging was run there is an estimate of the reconstruction resolution based on the Fourier shell correlation. In the models section there are several tabs. The summary tab shows the chi^2, R_g, support volume, and RSC to the reference model. If any model was not included in the averaging it is highlighted in red.
   - Verify that the Rg is close to the expected value, and that the chi^2 and support volumes are relatively consistent between models.
   - *Note:* Above we show the results for 20 runs instead of the 5 in the tutorial.
10. Individual model results are displayed in the numbered tabs. For each individual model there are plots of: the original data and the model data (scattering from density); the residual between the original data and the model data; and chi squared, R_g and support volume vs. refinement step.
   - Verify that the residual between the actual data and the model data is small.

- Check that the chi squared, R_g, and support volume have all plateaued (converged) by the final steps.



11. If the densities were averaged, the average tab will display the Fourier shell correlation vs. resolution.

    - *Note:* The reconstruction resolution is taken as the resolution in angstroms where the correlation first crosses 0.5.



12. The results summary shown in Summary tab is automatically saved as a **<prefix>_denss_results.csv** csv file, e.g. for this data as **lysozyme_denss_results.csv**. All the plots shown on the individual model tabs are automatically saved as a multi-page pdf file with the same name.

13. Click the "Close" button when you are finished looking at the results and reconstructions.

14. The results from the individual DENSS runs are saved in the selected output folder as **<prefix>_xx.mrc** where *xx* corresponds to the run number: 01, 02, etc. For this tutorial that would be **lysozyme_01.mrc**, **lysozyme_02.mrc**, etc.

15. If averaging was done, final average density is saved in the selected output folder as **<prefix>_aver.mrc**. For this tutorial, that would be **lysozyme_aver.mrc**.

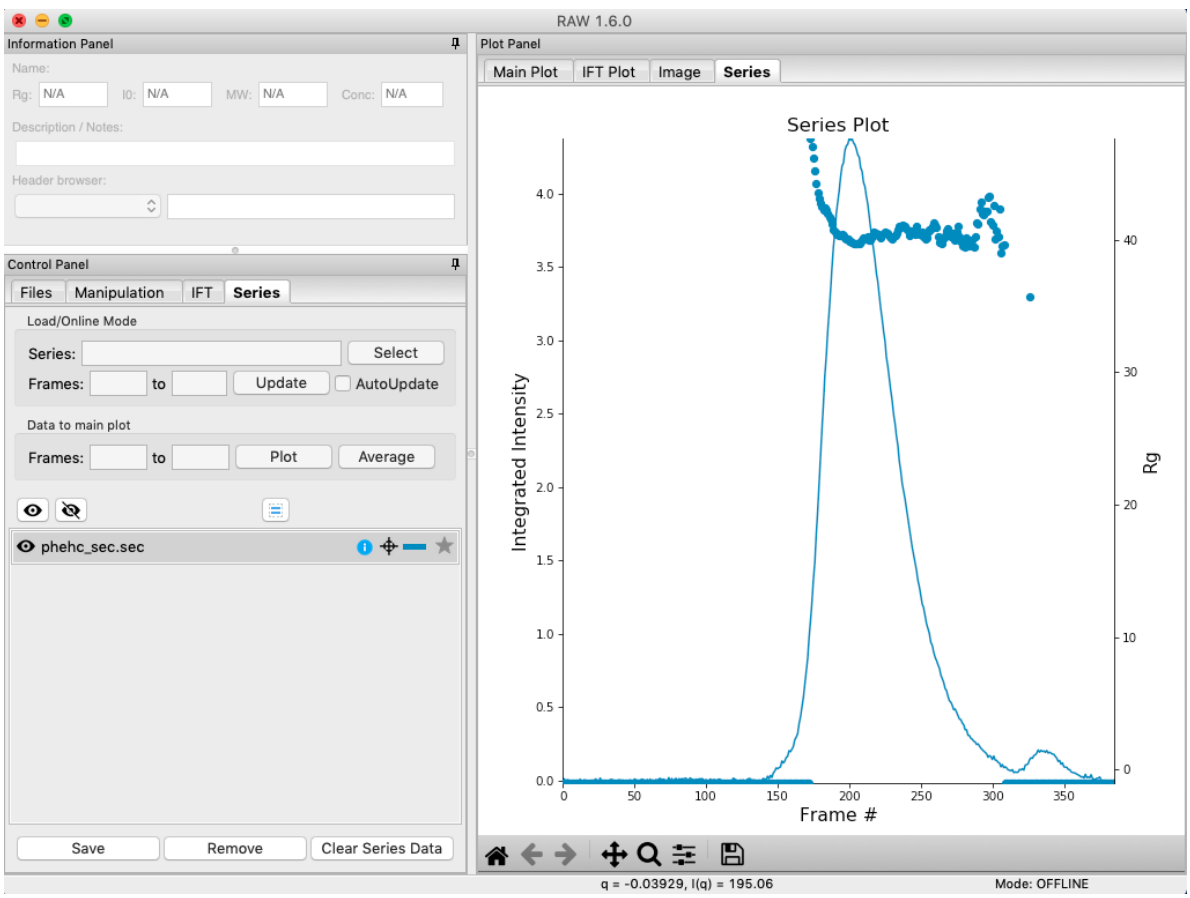
16. If refinement was done, the final refined density is saved in the selected output folder as **<prefix>_refine.mrc**. For this tutorial that would be **lysozyme_refine.mrc**.

Note: **.mrc** files can be opened in Chimera and pyMOL. For tips about how to visualize the density and align it with known structures see the appropriate sections here: http://www.tdgrant.com/denss/tips/.

Advanced SEC-SAXS processing – Singular value decomposition (SVD) and evolving factor analysis (EFA)

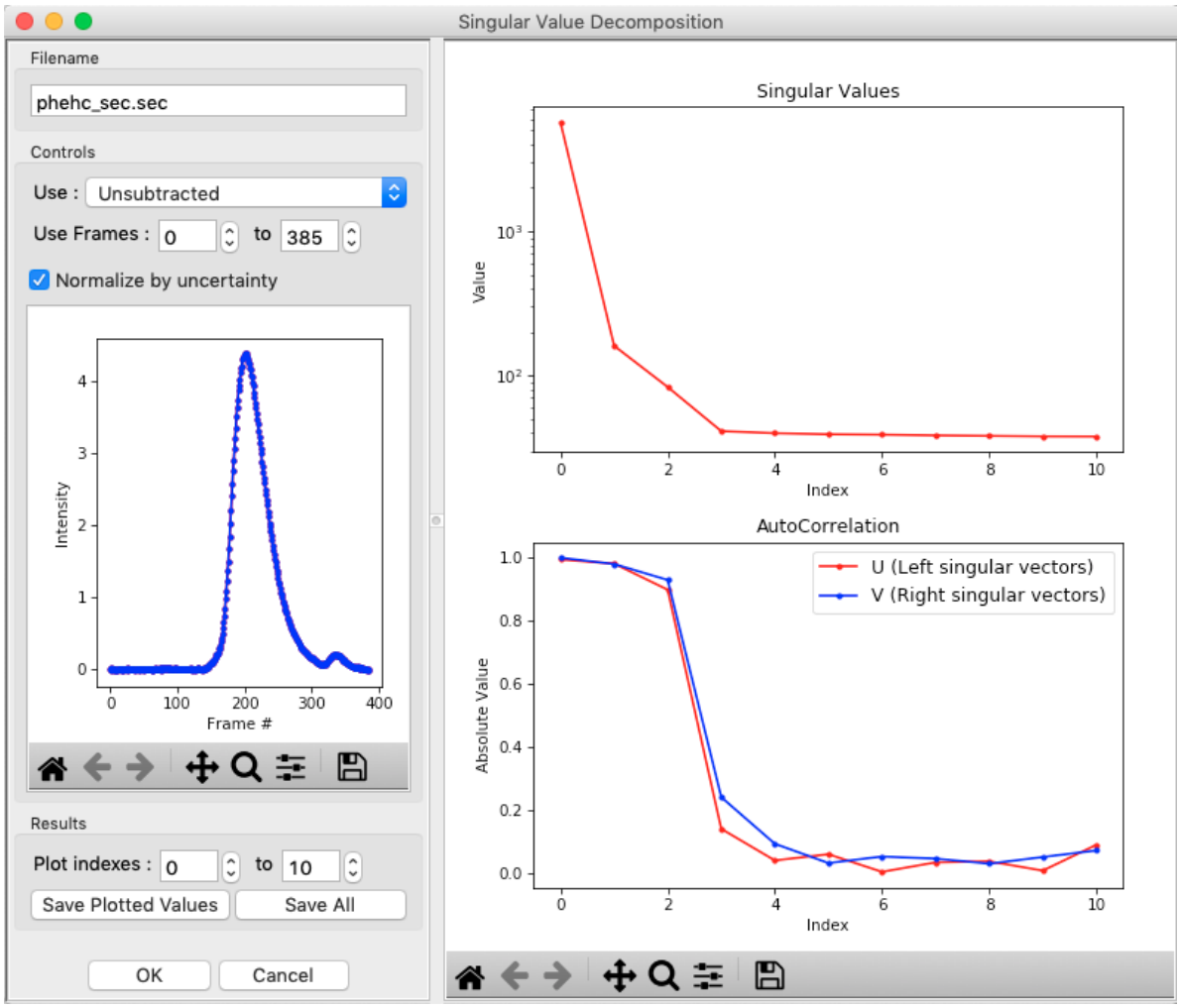
Sometimes SEC fails to fully separate out different species, and you end up with overlapping peaks in your SEC-SAXS curve. It is possible to apply more advanced mathematical techniques to determine if there are multiple species of macromolecule in a SEC-SAXS peak, and to attempt to extract out scattering profiles for each component in an overlapping peak. Singular value decomposition (SVD) can be used to help determine how many distinct scatterers are in a SEC-SAXS peak. Evolving factor analysis (EFA) is an extension of SVD that can extract individual components from overlapping SEC-SAXS peaks.

1. Clear all of the data in RAW. Load the **phehc_sec.sec** file in the **sec_data** folder.
   - *Note:* The data were provided by the Ando group at Cornell University and is some of the data used in the paper: *Domain Movements upon Activation of Phenylalanine Hydroxylase Characterized by Crystallography and Chromatography-Coupled Small-Angle X-ray Scattering*. Steve P. Meisburger, Alexander B. Taylor, Crystal A. Khan, Shengnan Zhang, Paul F. Fitzpatrick, and Nozomi Ando. Journal of the American Chemical Society 2016 138 (20), 6506-6516. DOI: 10.1021/jacs.6b01563

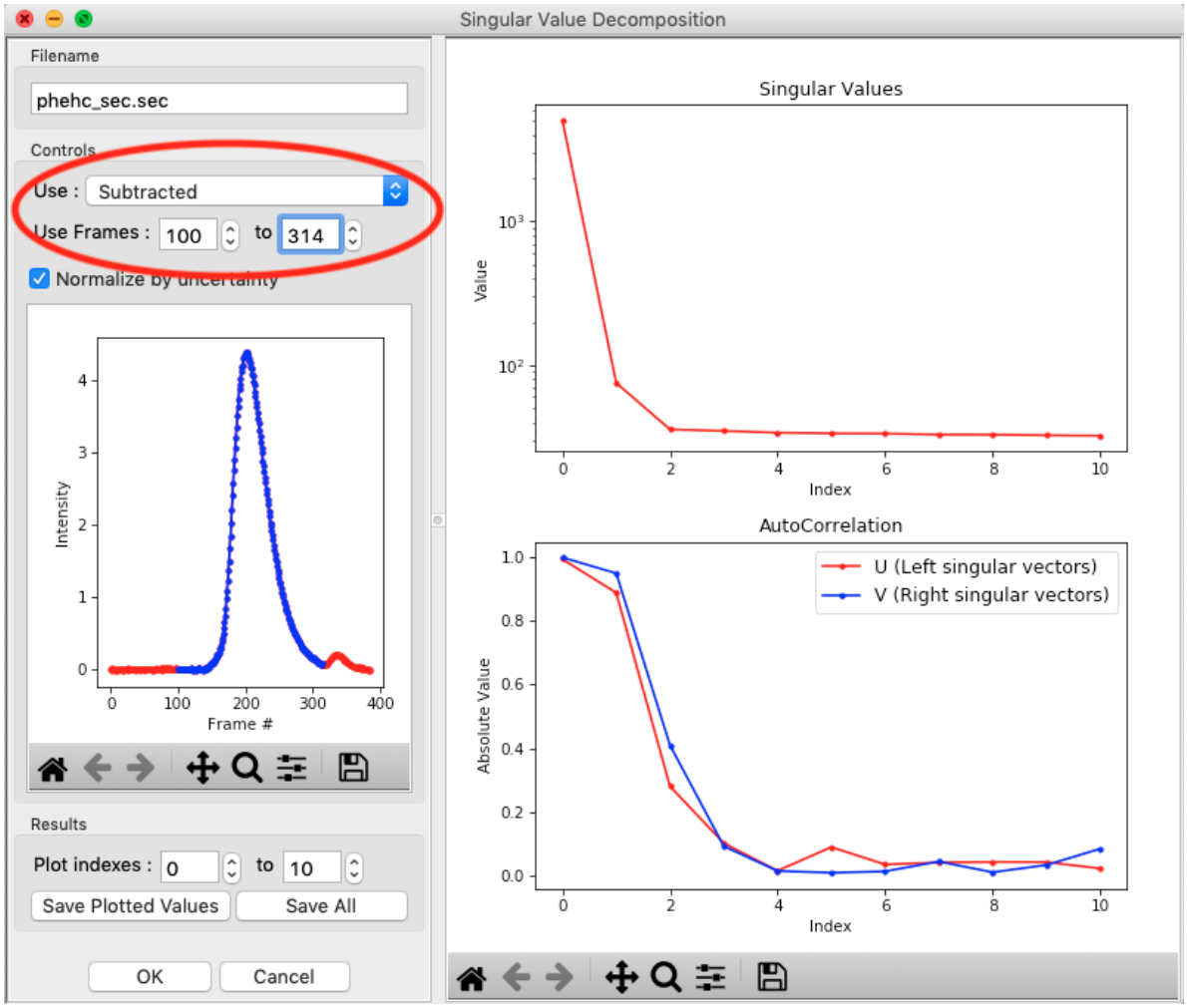


2. Right click on the **phehc_sec.sec** item in the Series list. Select the "SVD" option.
3. The SVD window will be displayed. On the left are controls, on the right are plots of the value of the singular values and the first autocorrelation of the left and right singular vectors.

- *Note:* Large singular values indicate significant components. What matters is the relative magnitude, that is, whether the value is large relative to the mostly flat/unchanging value of high index singular values.
- *Note:* A large autocorrelation indicates that the singular vector is varying smoothly, while a low autocorrelation indicates the vector is very noisy. Vectors corresponding to significant components will tend to have autocorrelations near 1 (roughly, >0.6-0.7) and vectors corresponding to insignificant components will tend to have autocorrelations near 0.

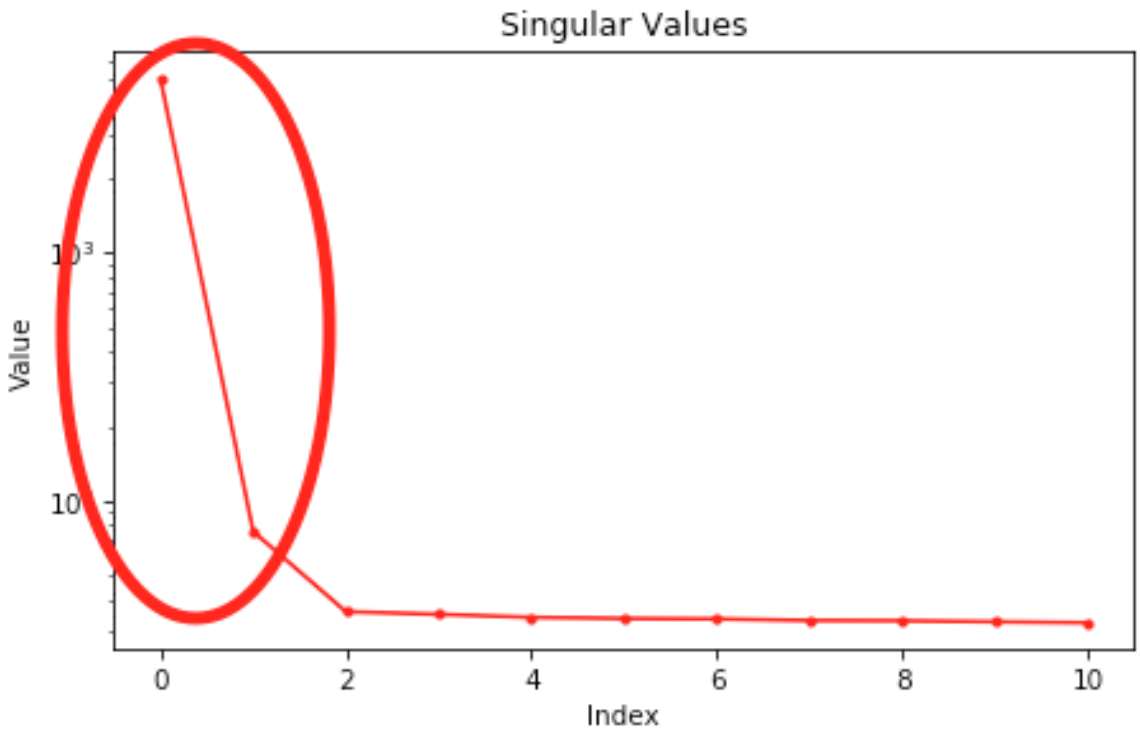


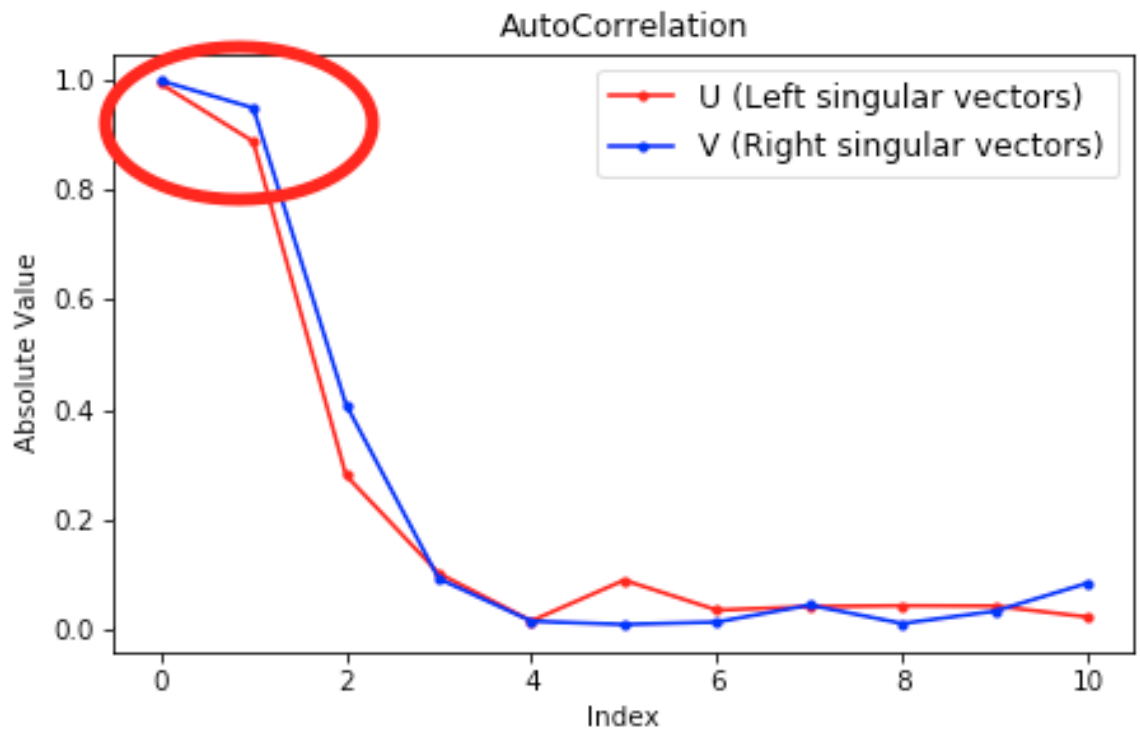
4. Adjust the starting frame number to 100, the ending frame number to near 300, and switch to using Subtracted data.
   - *Note:* The blue points are in the plot on the left are the region being used for SVD, while the red points shows the rest of the SEC-SAXS curve.



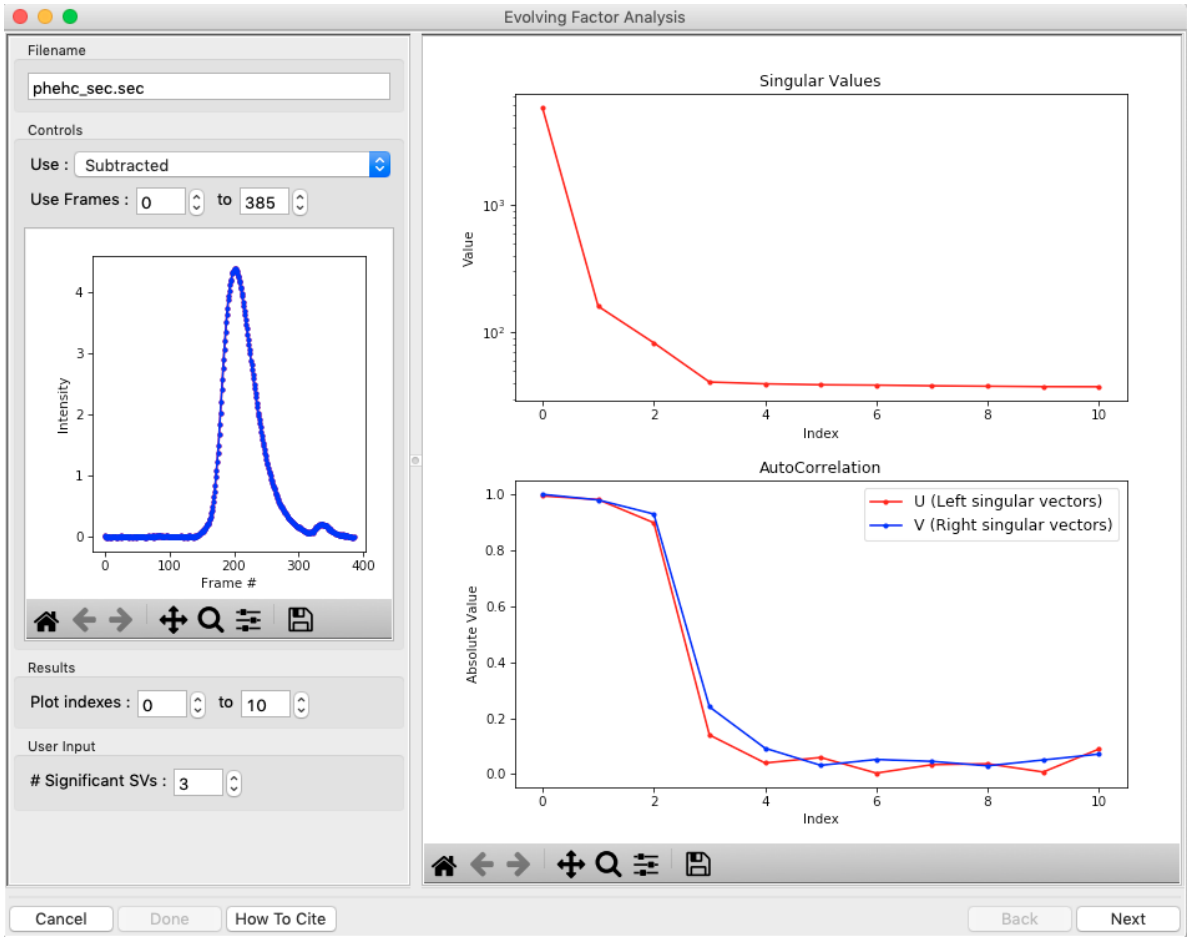
5. We have now isolated the peak. Looking at the top plot, we see there are two singular values significantly above the baseline level, and from the autocorrelation we see two values with both left and right singular vectors autocorrelations near 1. This indicates that there are two scattering components in the peak, even though there are no obvious shoulders in the region we selected

   - *Try:* Adjust the starting and ending values and seeing how that changes the SVD results. Is there a region of the peak you can isolate that has just one significant component?
   - *Note:* Normally, changing between Unsubtracted and Subtracted SEC-SAXS profiles should remove one significant singular value component, corresponding to the buffer scattering. In this data, you will see no difference, as the profiles used to produce the SEC-SAXS curve were already background subtracted.
   - *Note:* You can save the SVD plots by clicking the Save button, as with the plots in the main RAW window. You can save the SVD results, either just the plotted values or all of the values, using the two Save buttons in the SVD panel.





6. Close the SVD window by clicking the OK button.
7. We will now use EFA to attempt to extract out the two scattering components in the main peak in this data. Right click on the **phehc_sec.sec** item in the Series list. Select the "EFA" option.



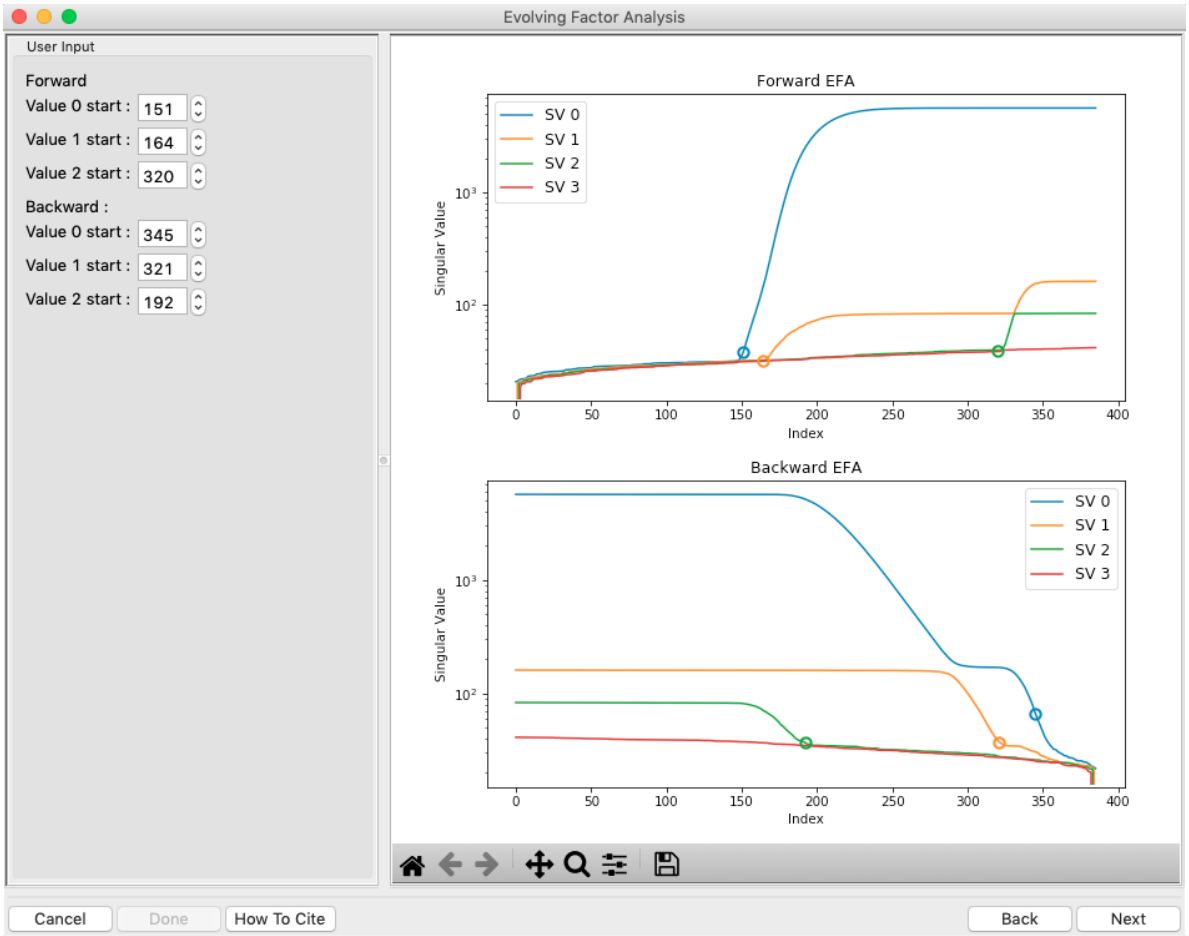
8. For successful EFA, you want to use Subtracted data, and you typically want to have a long buffer region before and after the sample. For this data set, using the entire frame range (from 0 to 385) is appropriate. With other data sets, you may need to change the frame range to, for example, remove other, well separated, peaks from the analysis.

9. RAW attempts to automatically determine how many significant singular values (SVs) there are in the selected range. At the bottom of the control panel, you should see that RAW thinks there are three significant SVs in our data. For this data set, that is accurate.

   - *Note:* You should convince yourself of this by looking at the SVD results in the plots on this page, using the same approach as in Steps 3-5 above.
   - *Note:* There is a hint of a fourth component. You can rerun this exercise using four components and see if that changes the results.

10. Click the "Next" button in the lower right-hand corner of the window to advance to the second stage of the EFA analysis.

    - *Note:* It may take some time to compute the necessary values for this next step, so be patient.



11. This step shows you the "Forward EFA" and "Backward EFA" plots. These plots represent the value of the singular values as a function of frame.

    - *Note:* There is one more singular value displayed on each plot than available in the controls. This is so that in the following Steps you can determine where each component deviates from the baseline.

12. In the User Input panel, tweak the "Forward" value start frames so that the frame number, as indicated by the open circle on the plot, aligns with where the singular value first starts to increase quickly. This should be around 147, 164, and 324.

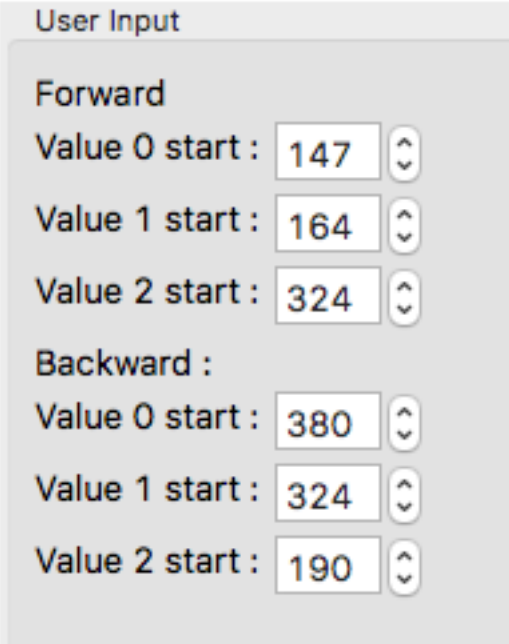
    - *Note:* For the Forward EFA plot, SVD is run on just the first two frames, then the first three, and so on, until all frames in the range are included. As more frames are added, the singular values change, as shown on the plot. When a singular values starts increasingly sharply, it indicates that there is a new scattering component in the scattering profile measured at that point. So, for the first ~150 frames, there are no new scattering components (i.e. just buffer scattering). At frame ~147, we see the first singular value (the singular value with index 0, labeled SV 0 on the plot) start to strongly increase, showing that we have gained a scattering component. We see SV 1 start to increase at ~164, indicating another scattering component starting to be present in the scattering profile.

13. In the User Input panel, tweak the "Backward" value start frames so that the frame number, as indicated by the open circle on the plot, aligns with where the singular value first starts to increase quickly, reading the plot left to right (i.e. where it drops back to near the baseline). This should be around 380, 324, and 190.

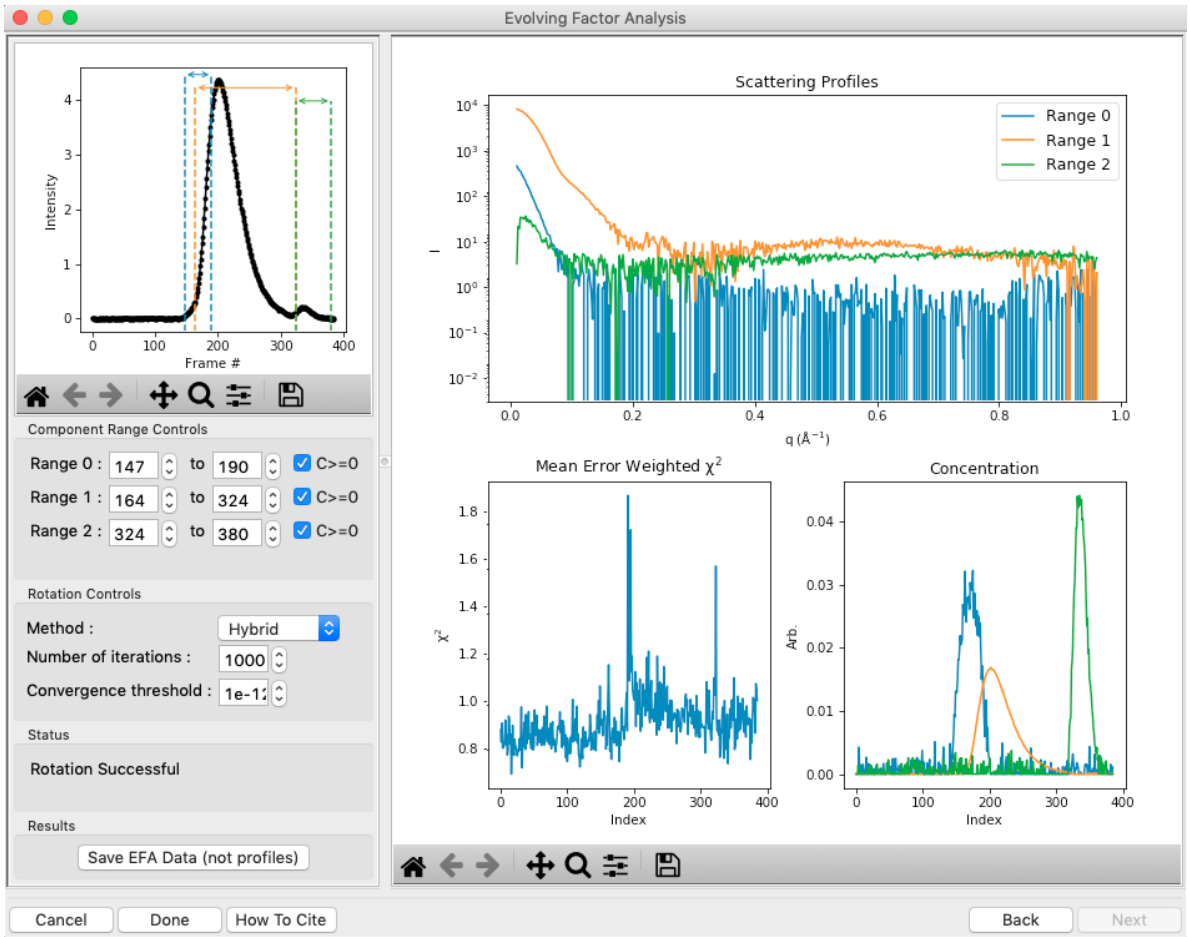
    - *Note:* For the Backward EFA plot, SVD is run on just the last two frames, then the last three, and so on, until all frames in the range are included. As more frames are added, the singular values change, as shown on the plot. When a singular values starts increasingly sharply (as seen from right to left), it indicates that

there is a new scattering component in the scattering profile measured at that point.

- *Note:* The algorithm for determining the start and end points is not particularly advanced. For some datasets you may need to do significantly more adjustment of these values



14. Click the "Next" button in the bottom right corner to move to the last stage of the EFA analysis.



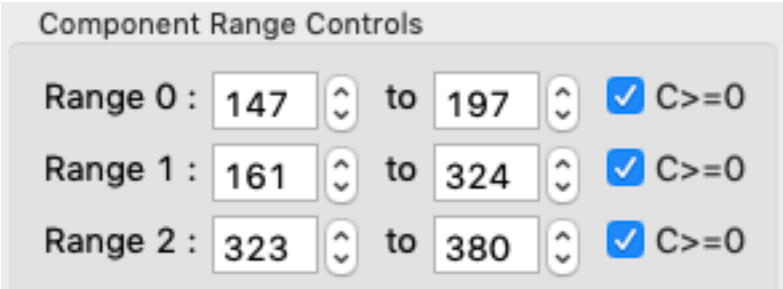
15. This window shows controls on the left and results on the right. In the controls area, at the top is a plot showing the SEC-SAXS curve, along with the ranges occupied by each scattering component, as determined from the input on the Forward and Backward EFA curves in stage 2 of the analysis. The colors of the ranges correspond to the colors labeled in the Scattering Profiles plot on the top right and the Concentration plot in the lower right. This panel takes the SVD vectors and rotates them back into scattering vectors corresponding to real components.

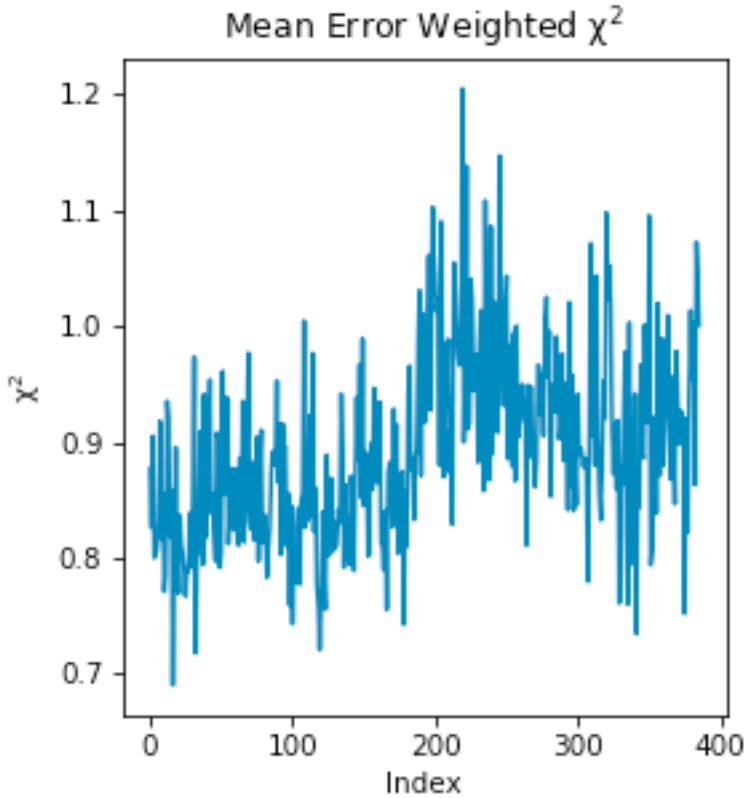
    - *Note:* This rotation is not guaranteed to be successful, or to give you valid scattering vectors. Any data obtained via this method should be supported in other ways, either using other methods of deconvolving the peak, other biophysical or biochemical data, or both!

16. Fine tune the ranges using the controls in the "Component Range Controls" box. Adjust the start of Range 2 down until it overlaps with Range 1.

    - *Question:* What is the effect on the chi-squared plot?

17. Adjust the starts and ends of Range 0 and the start of Range 1 by a few points until the spikes in the chi-squared plot go away. After these adjustments, Range 0 should be about 147 to 197, Range 1 from 161 to 324, and Range 2 from 323 to 380.

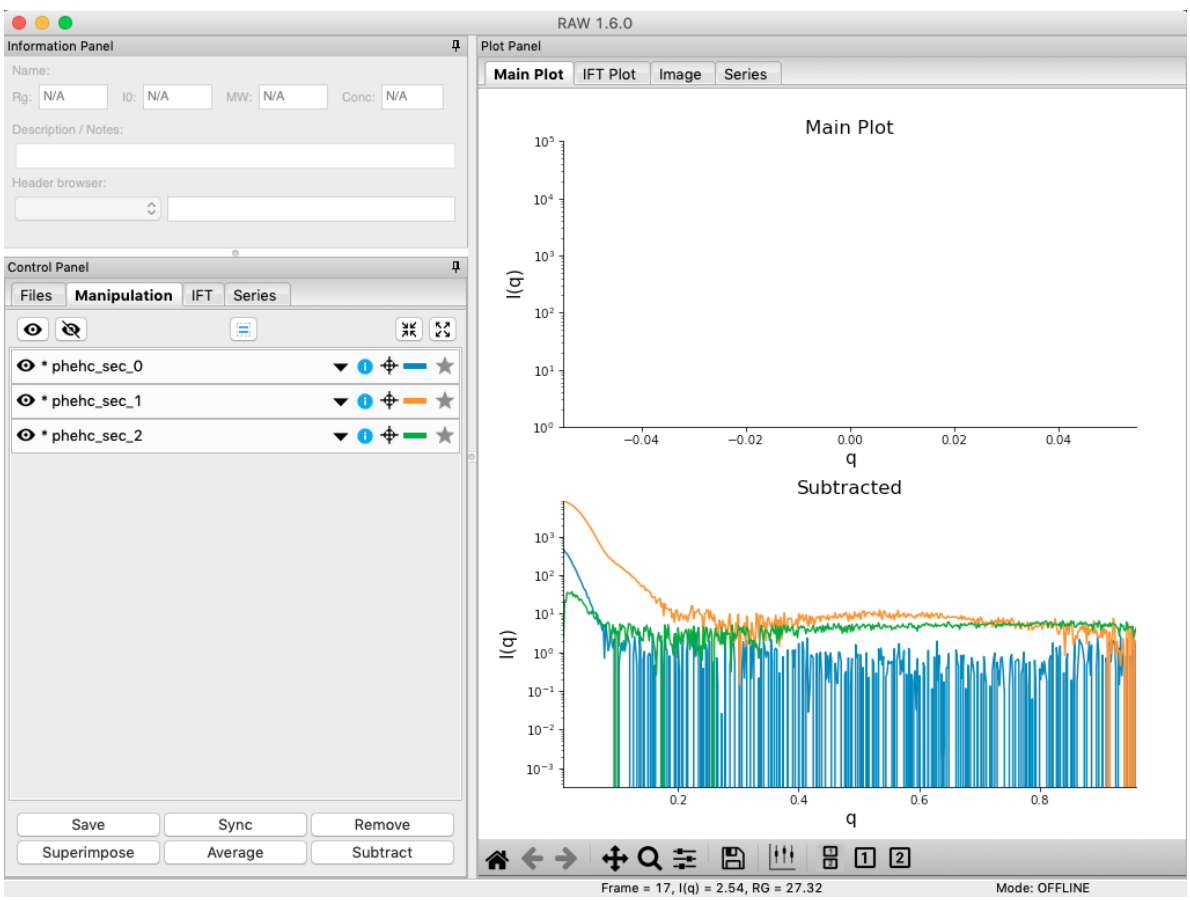




18. To see these changes on the Forward and Backward EFA plots, click the "Back" button at the bottom right of the page. Verify that all of your start and end values are close to where the components become significant, as discussed in Steps 12 and 13.
19. Click the "Next" button to return to the final stage of the EFA analysis.
20. In the Rotation Controls box, you can set the method, the number of iterations, and the convergence threshold. As you can see in the Status window, the rotation was successful for this data. If it was not, you could try changing methods or adjusting the number of iterations or threshold.
21. Examine the chi-squared plot. It should be uniformly close to 1 for good EFA. For this data, it is.
22. Examine the concentration plot. You'll see three peaks, corresponding to the concentrations for the three components. In the Range Controls, uncheck the Range 0 C>=0 box. That removes the constraint that the concentration must be positive. If this results in a significant change in the peak, your EFA analysis is likely poor, and you should not trust your results.
    - *Note:* The height of the concentration peaks is arbitrary, all peaks are normalized to have an area of 1.
23. Uncheck all of the C>=0 controls.

- *Question:* Do you observe any significant changes in the scattering profiles, chi-squared, or concentration when you do this? How about if you uncheck one and leave the others checked?

24. Recheck all of the C>=0 controls. You have now verified, as much as you can, that the EFA analysis is giving you reasonable results.
    - *Reminder:* Here are the verification steps we have carried out, and you should carry out every time you do EFA:
        1. Confirm that your selected ranges correspond to the start points of the Forward and Backward EFA values (Steps 12-13).
        2. Confirm that your chi-squared plot is close to 1, without any major spikes (Step 21).
        3. Confirm that your concentrations are not significantly altered by constraining the concentration to be positive (Steps 22-23).
25. Click the "Save EFA Data (not profiles)" to save the EFA data, including the SVD, the Forward and Backward EFA data, the chi-squared, and the concentration, along with information about the selected ranges and the rotation method used.
26. Click the "Done" button to send the scattering profiles to the Main Plot.
27. In the main RAW window, go to the Manipulation control tab and the Main plot. If it is not already, put the Main plot on a semi-Log or Log-Log scale.



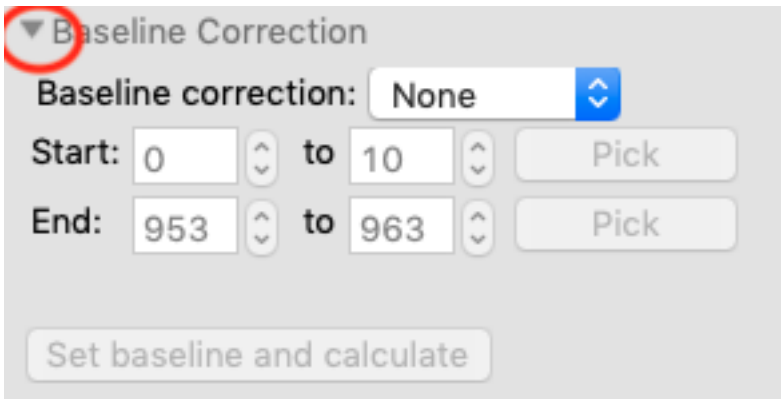
28. The three scattering profiles from EFA are in the manipulation list. The labels _0, _1, and _2 correspond to the 0, 1, and 2 components/ranges.
    - *Note:* Regardless of whether you use subtracted or unsubtracted data, these scattering profiles will be

buffer subtracted, as the buffer represents a scattering component itself, and so (in theory) even if it is present will be separated out by successful EFA.

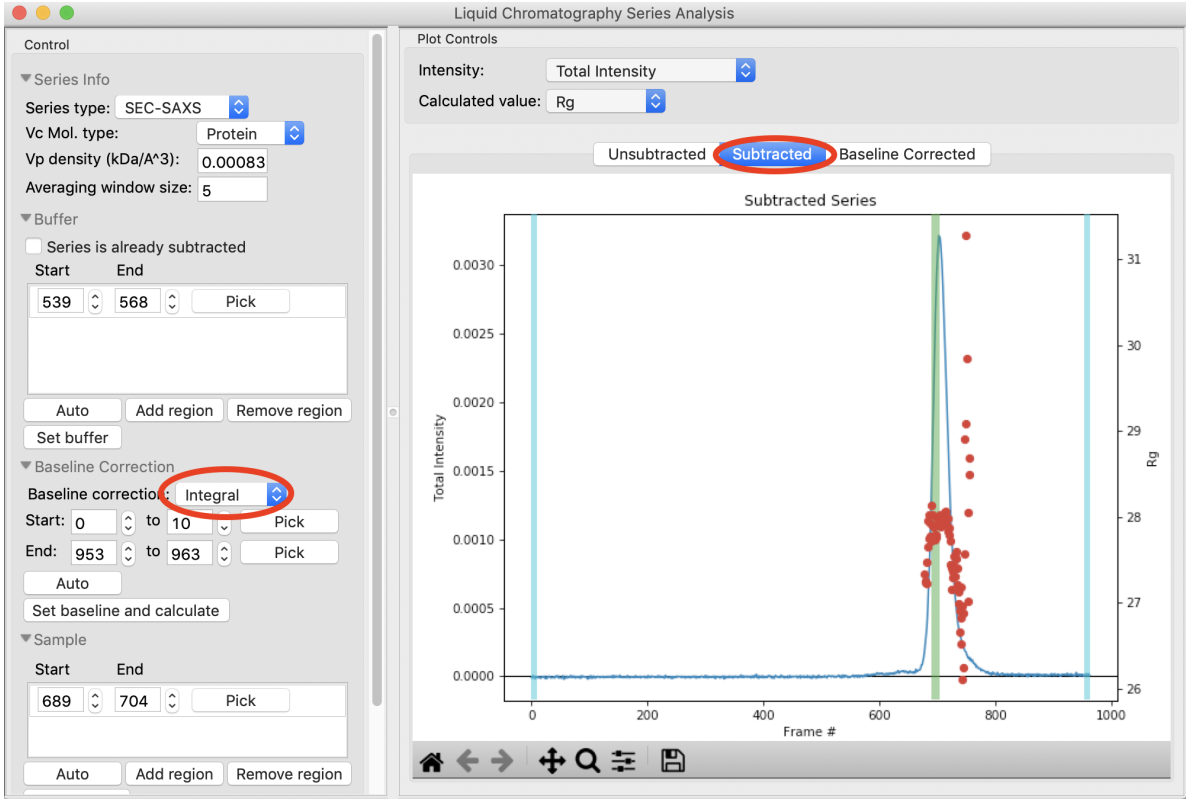
Advanced SEC-SAXS processing – Baseline correction

Sometimes SEC data shows a baseline drift. This can be due either to instrumental changes (such as beam drift), or changes in the measured system, such as capillary fouling. RAW provides the ability to correct for these forms of baseline drift using either a linear or integral baseline method. The linear baseline method is best for instrumental drifts, while the integral baseline method is best for capillary fouling. Both baseline methods apply a distinct correction for each q value.

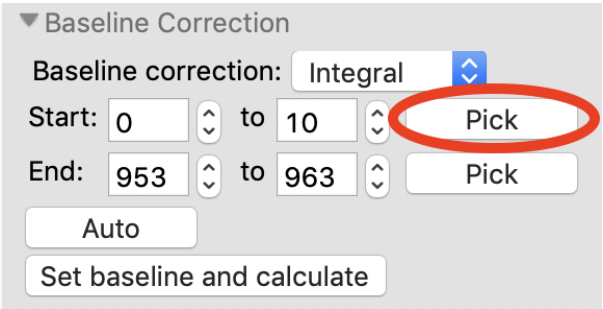
1. Clear all of the data in RAW. Load the **baseline.sec** SEC data in the **sec_data** folder.
   - *Note:* This is the same as what you *previously saved* in an earlier part of the tutorial.
2. Open the LC Series analysis panel.
3. To baseline correct data, you should only have buffer regions selected before the peak. Select the 'Unsubtracted' plot tab and verify that your buffer regions are before the peak of interest.
4. Use the triangle to expand the Baseline Correction section.



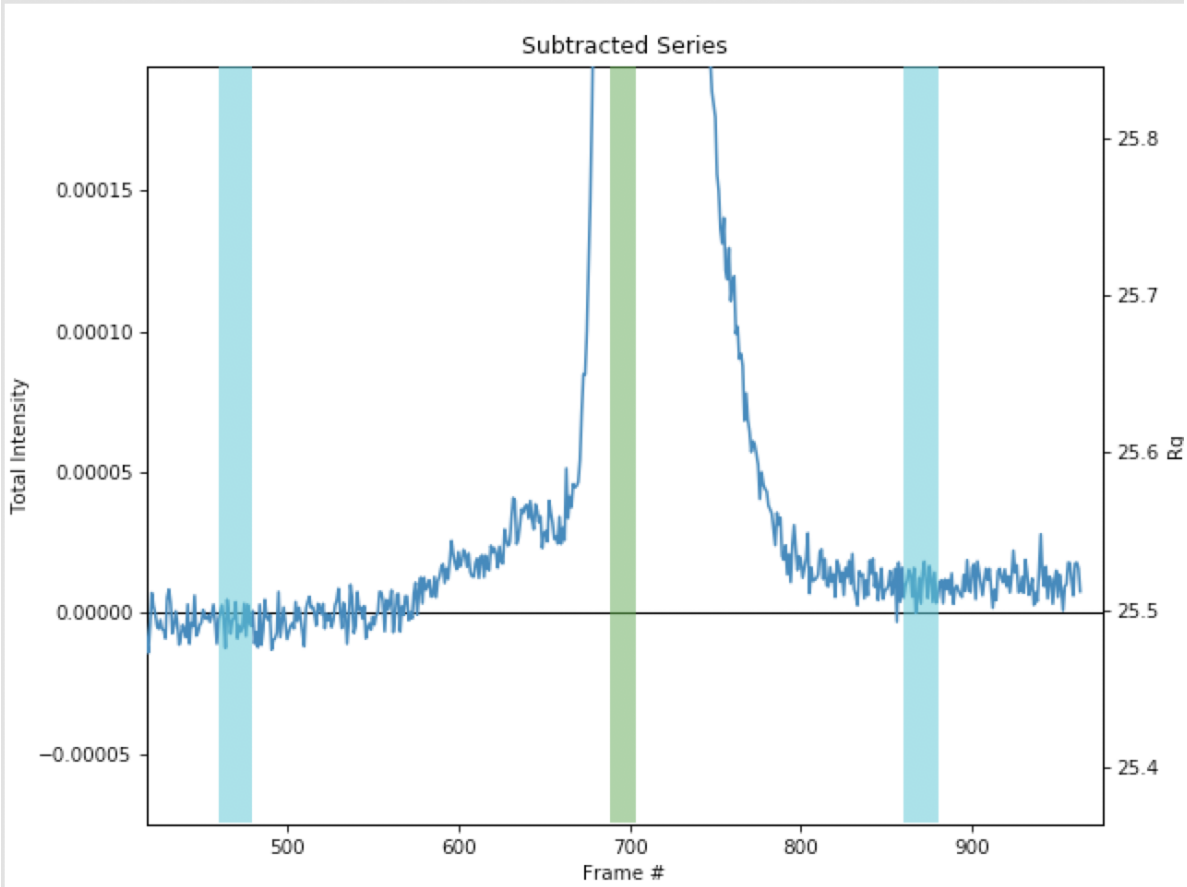
5. Set baseline correction to 'Integral'. You should see the start and end controls become active, and two blue regions representing your selected start and end regions appear on the 'Subtracted' plot.
   - *Note:* This should automatically show the 'Subtracted' plot tab. If it doesn't, select the 'Subtracted' plot tab.



6. Zoom in near the base of the subtracted peak. Click the 'Pick' button for the baseline correction start region and select a start region in the flat baseline area just before the start of the peaks.

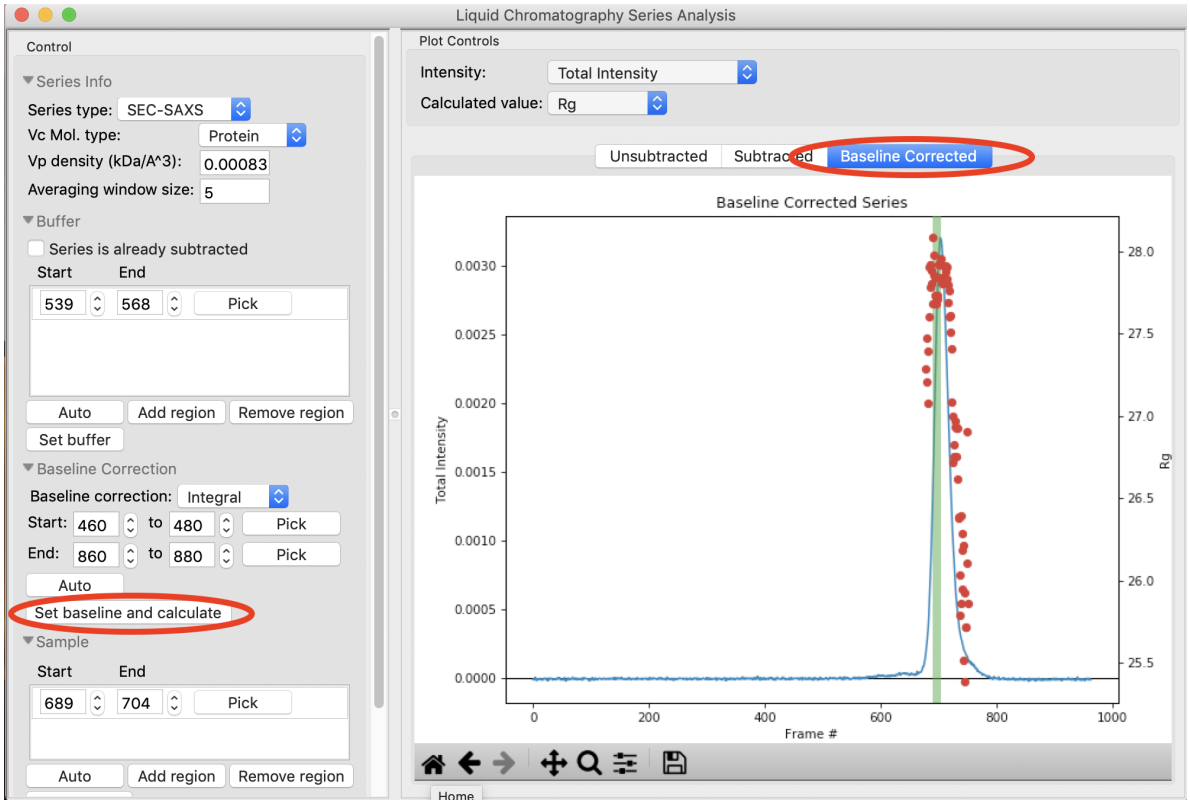


7. Pick an end region in the flat region just after the end of the peaks.
   - *Note:* You should end up with regions ~460-480 and 860-880
   - *Try:* You can use the 'Auto' button to automatically find start and end regions. For this dataset, it ends up a little close to the peaks, so some manual adjustment is necessary.



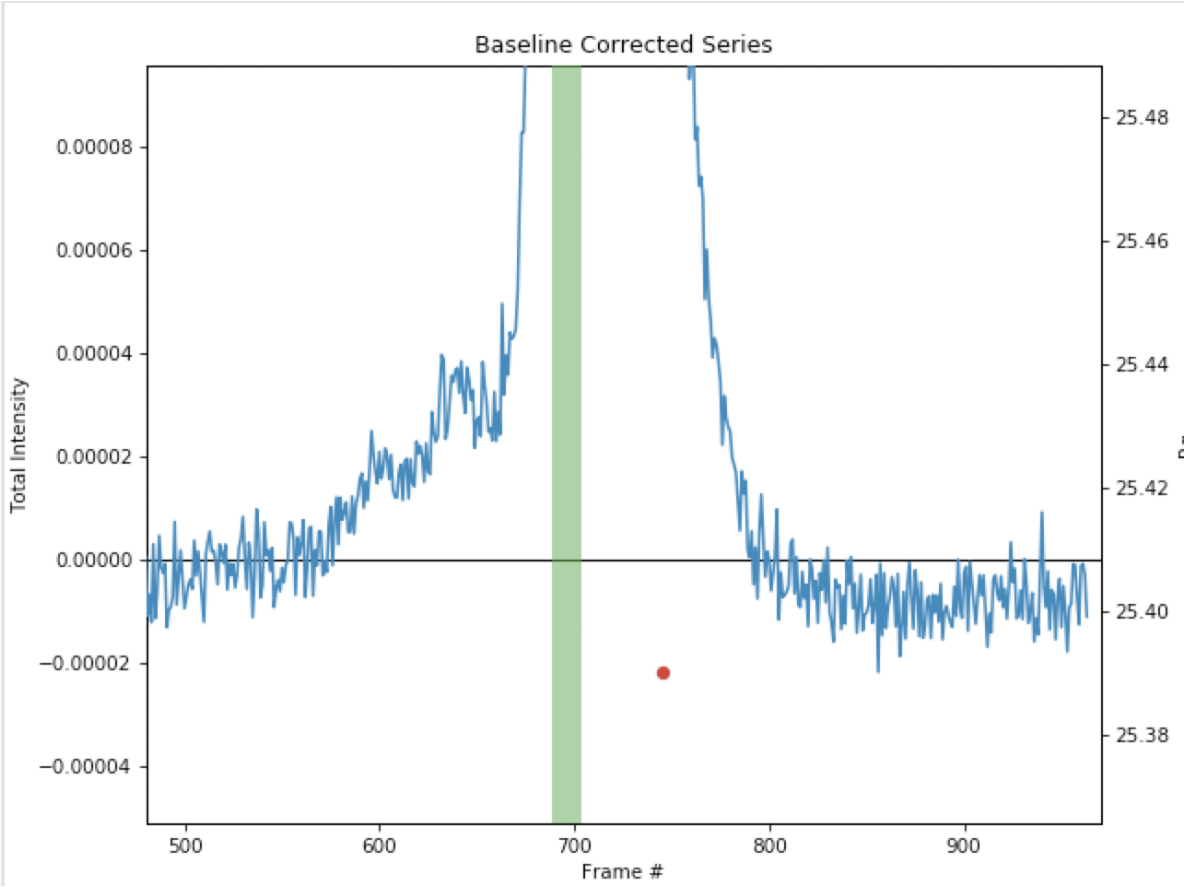
8. Click the 'Set baseline and calculate' button. The 'Baseline Corrected' plot should automatically show after you set the baseline region and calculate. If not, change to that plot tab.

    - *Note:* The start and end points should be set in regions with no change in the baseline. If they aren't, RAW will give a warning. Try changing the end region to ~800-820 to see such a warning.

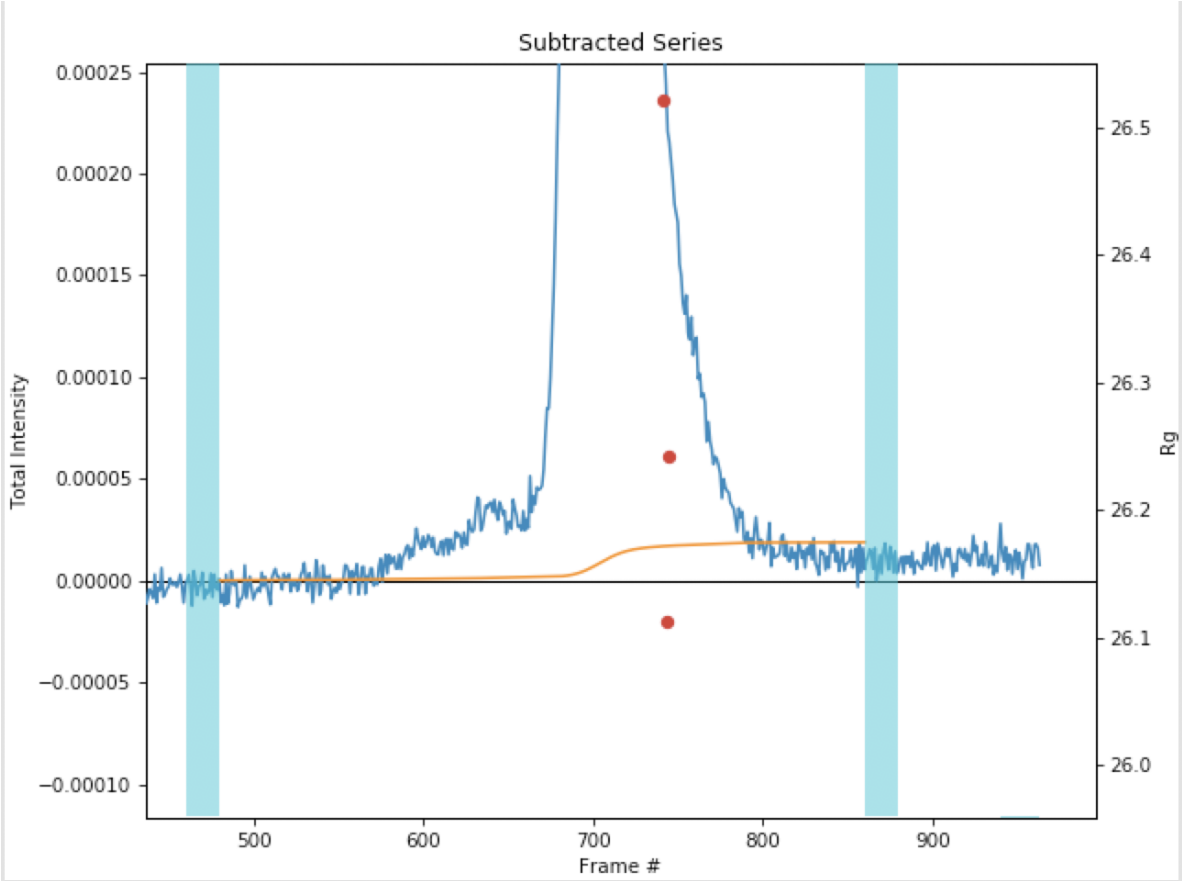


9. Zoom in on the base of the peak in the baseline corrected dataset. You should see that the baseline is actually a little overcorrected. This is because the integral baseline correction only allows for positive or no change in the baseline, so if some q values need a negative correction the total baseline ends up overcorrected, as the positive values are brought down but the negative values are not brought up.

   - *Note:* You can change the intensity display to individual q values or a q range and look at different points in q to try figure out which q values are causing the issue.



10. Switch back to the subtracted plot and zoom in on the base of the peak. You'll see the calculated baseline shown in orange.



11. Change the intensity display to 'Intensity in q range' and try several different q ranges. This will allow you to see what q values are getting the baseline overcompensated.
    - *Try:* Recommended regions to try for this dataset are 0.01-0.02, 0.05-0.06, 0.1-0.2, 0.2-0.27.
    - *Note:* You should find that it is the high q ranges that are being overcorrected. This may imply that the profiles are mostly noise in that range. If you example the profiles and determine that is the case, you could truncate the profiles to lower q before doing the baseline correction.
12. Switch back to displaying the total intensity and the Baseline Corrected plot.
13. Remove any existing sample region and find a new sample region using the 'Auto' button. Send that region to the main plot.
    - *Try:* Compare with what you found for this dataset *earlier in the tutorial*.

Note: The linear baseline correction is done in much the same way as the integral correction, so it is not covered in this tutorial. The only differences are that there is no 'Auto' determination of start and end regions, and that you should pick start and end regions that are linearly changing, not flat.

4.3.4 Creating a configuration file

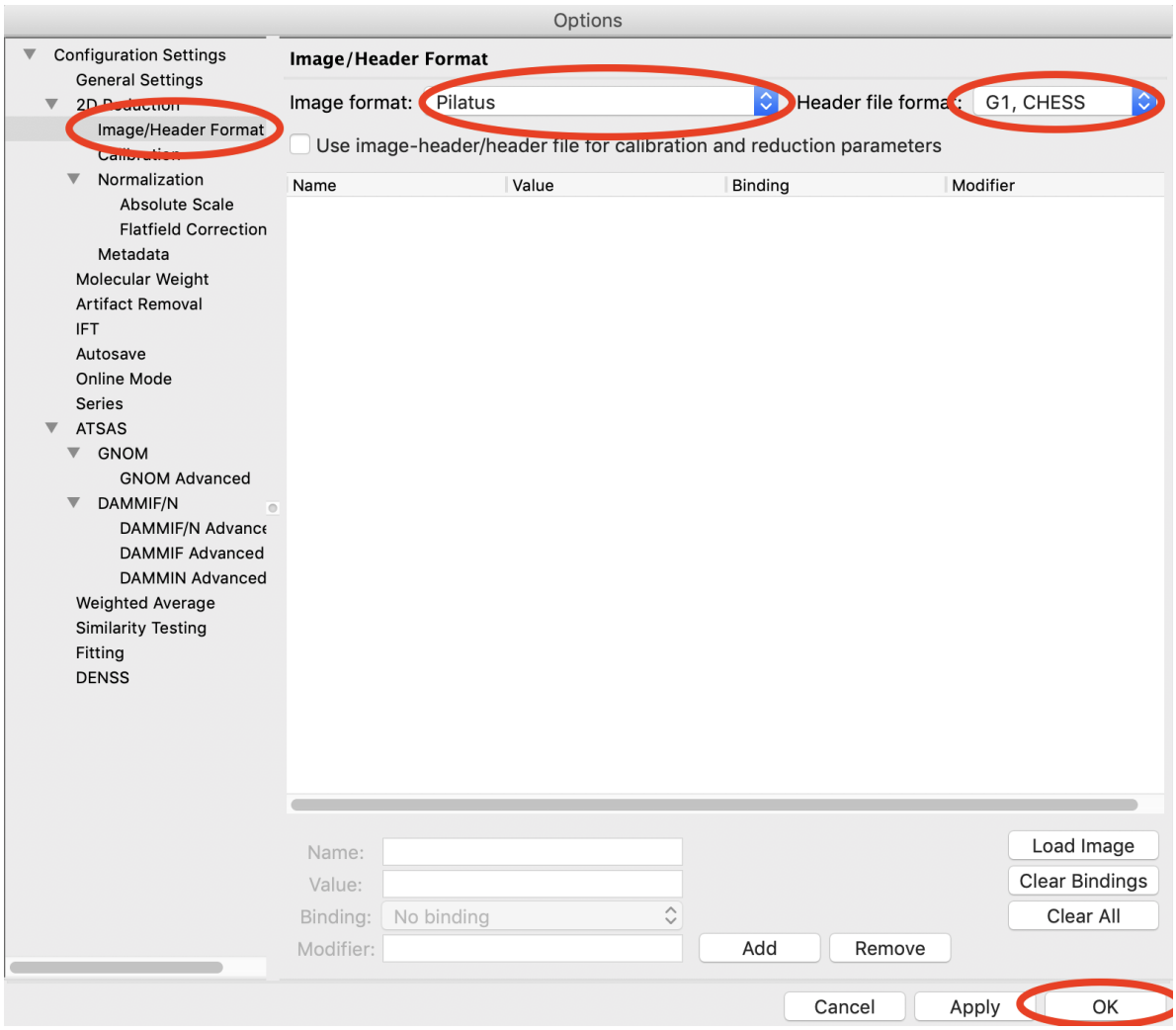
This section will guide you through creating a configuration file for RAW that allows you to integrate 2D images into 1D scattering profiles. It refers to the *RAW tutorial data*.

Select a section below to view the tutorial, or use the next and back buttons at the bottom of the page to navigate through it in order.

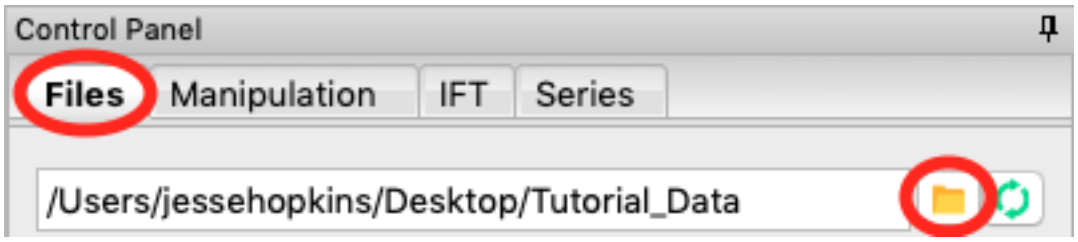
Centering and calibration – Automated method

The first step is to set the beam center, x-ray wavelength, and sample to detector distance. Before this can be done, you have to set the image and file header type in the Options window. The best way to find the beam center and sample to detector distance is using the automated method in RAW.

1. Open RAW. The install instructions contain information on installing and running RAW.
2. Open the Options window by going to the Options menu (top of the RAW window or in the system bar, depending on your OS) and selecting "Advanced Options"
3. In the options window, select the Image/Header Format section on the left.
4. In the area on the right, set the Image format dropdown menu to "Pilatus" and the Header file format to "G1, CHESS".

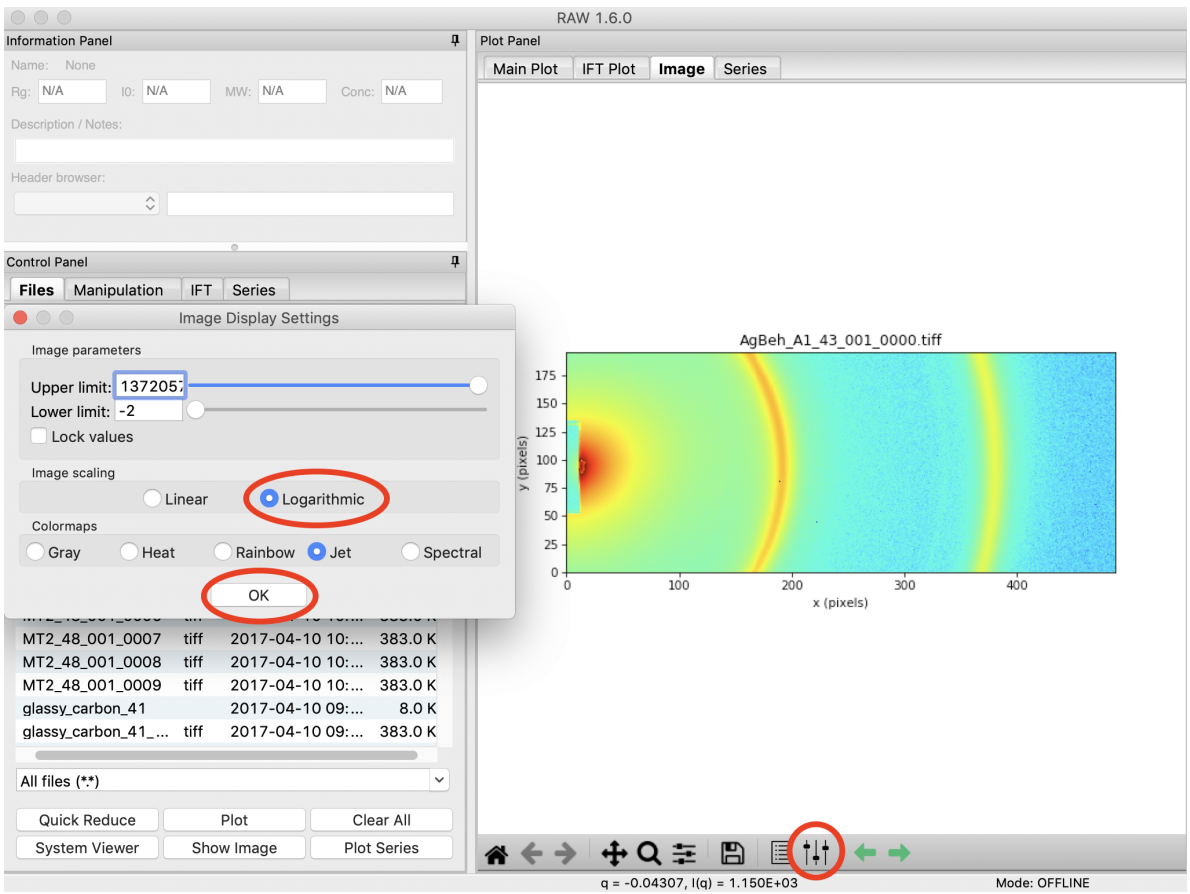


5. Click the OK button to close the window and save the changes to the settings.
6. In the files tab, click on the folder button and navigate to the **Tutorial_Data/calibration_data** folder.

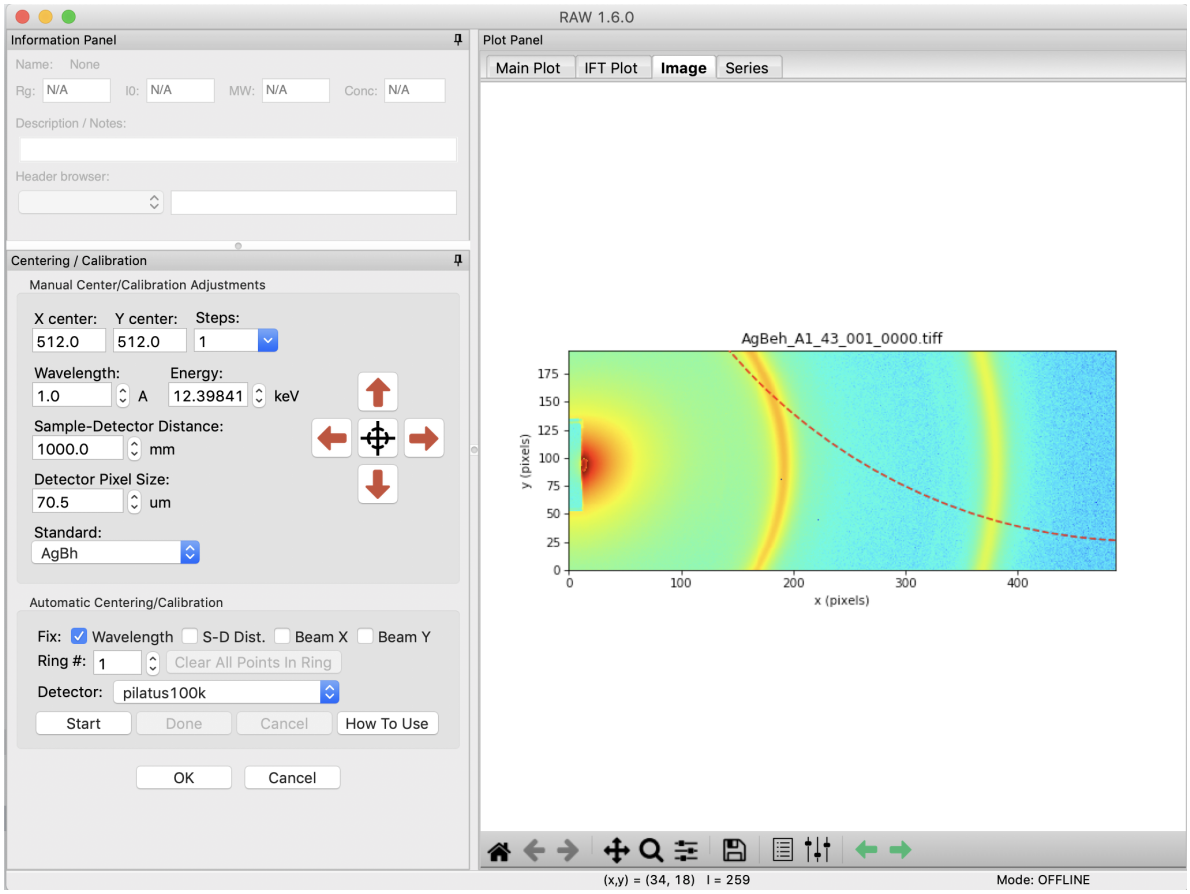


7. Select the **AgBeh_A1_43_001_0000.tiff** file by clicking on it once and click the show image button at the bottom of the screen.

8. In the Image Plot Panel that is now showing, click on the Image Display Settings button (looks like slider bars) at the bottom of the screen.
9. In the window that appears, set the scale to logarithmic and and click "OK".

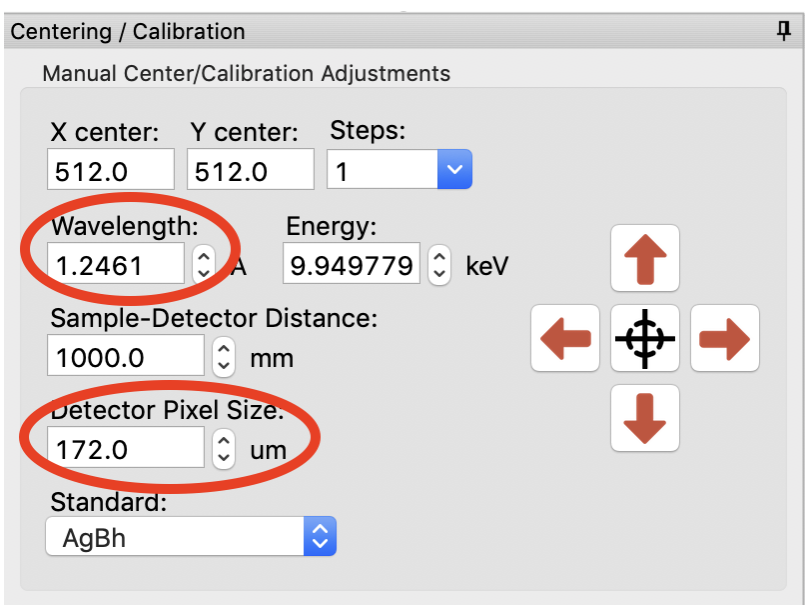


10. Open the Centering/Calibration panel by going to the Tools menu and selecting "Centering/Calibration".

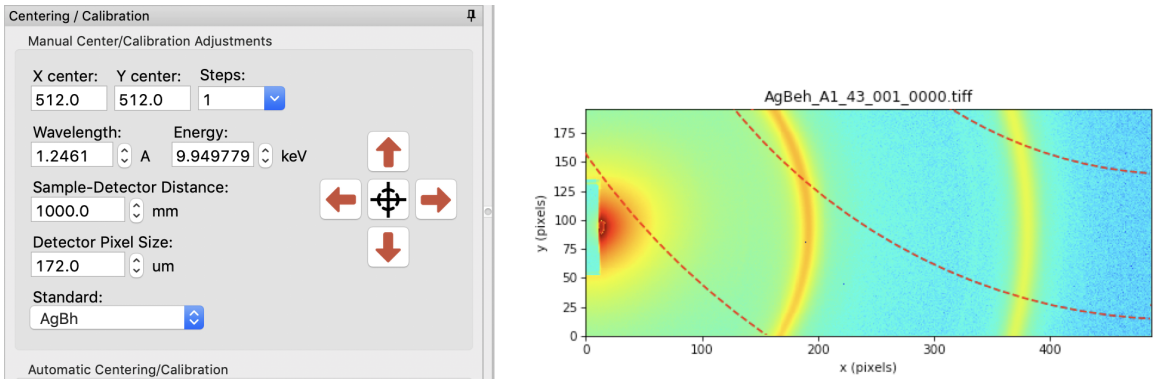


11. In the Centering/Calibration panel set the wavelength to 1.2461 Å. Set the detector pixel size to 172.0 μm.

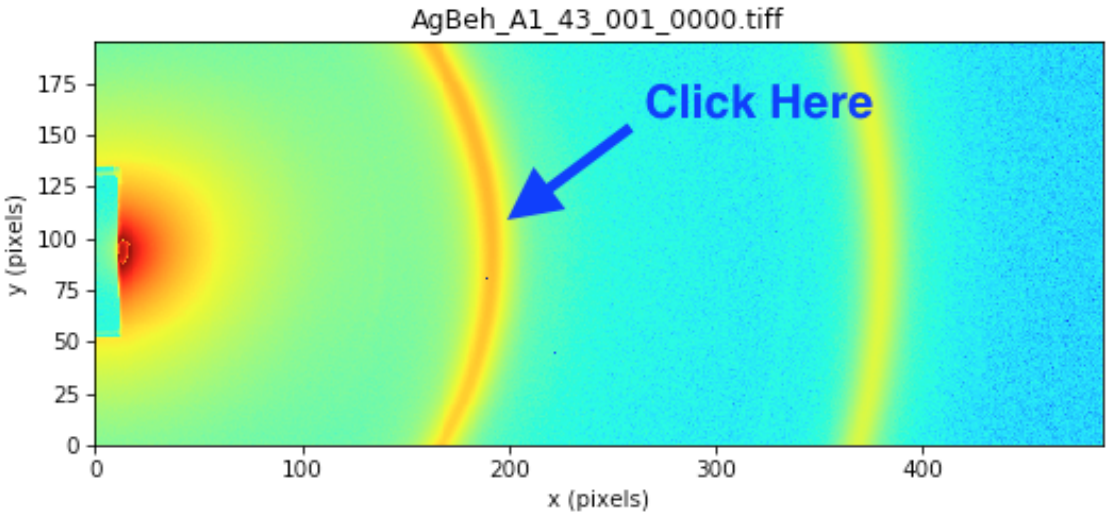
    - *Note:* The x-ray energy/wavelength and detector pixel size are previously known values, and are not found in RAW.
    - *Tip:* You can set the value in two ways. Either using the up/down arrows next to the box (spin controls) or directly typing the value into the box.



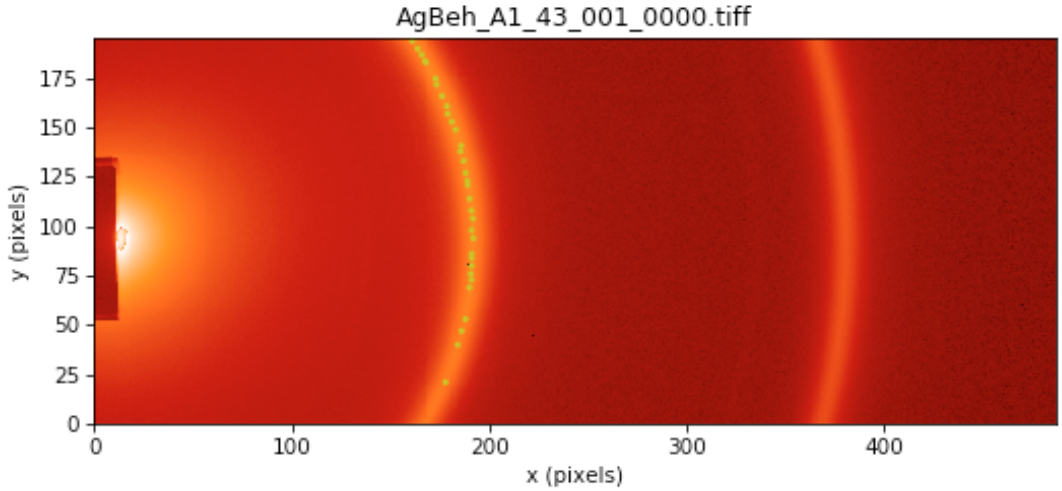
12. The goal of centering and calibration is to find a beam center position and sample to detector distance that causes the displayed Silver-Behenate ring pattern to match up with the rings on the image.
    - *Note:* The beamstop is the dark blue patch extending out from the center of the left edge of the detector.
13. *Checkpoint:* You should currently have a screen that looks like the one below.



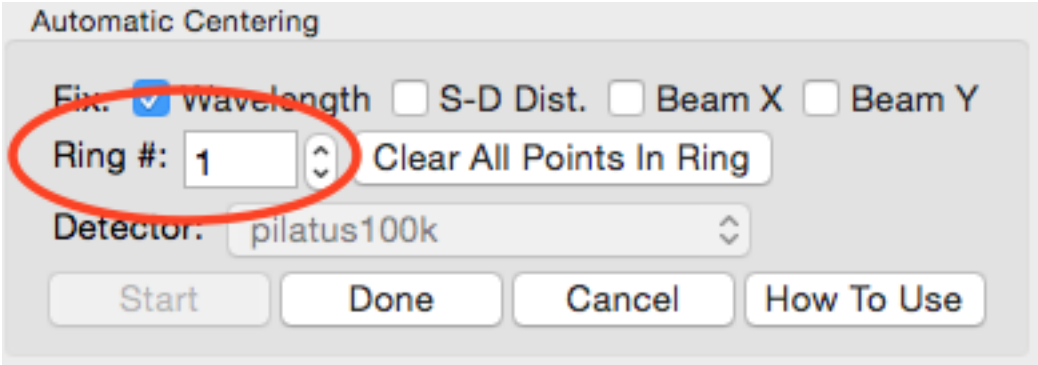
14. In the "Manual Centering/Calibration Adjustments" panel, make sure the correct standard is selected, in this case, AgBh (silver behenate).
15. In the Automatic Centering panel, make sure the correct detector is selected, in this case "pilatus100k".
    - *Note:* If you cannot find your detector in that list, select "detector", and make sure that you have entered the correct detector pixel size in the manual centering panel.
16. Click the "Start" button in the "Automatic Centering/Calibration" panel.
17. Make sure the "Ring #" is set to 0. Click on a point with strong intensity in the Silver Behenate ring nearest the beamstop (left most ring, in this case, near x = 200).
    - *Note:* For some experimental setups, one or more of the largest d-spacing rings may not be visible on the detector. In this case, you need to figure out what the first visible ring on the detector is, and set the ring number to that. So, if the third ring was the first one on the detector, the Ring # would be set to 2 (the ring number is zero index, so 0 corresponds to the first ring, 1 to the second ring, and so on).
    - *Try:* If you set the Manual settings to approximately right, hover the mouse above the image and use the scroll wheel to zoom out. This will let you see all of the centering rings, and figure out which ring is the first one visible on the image. Once you're done, hit the Home button to return to the zoom of the entire image.



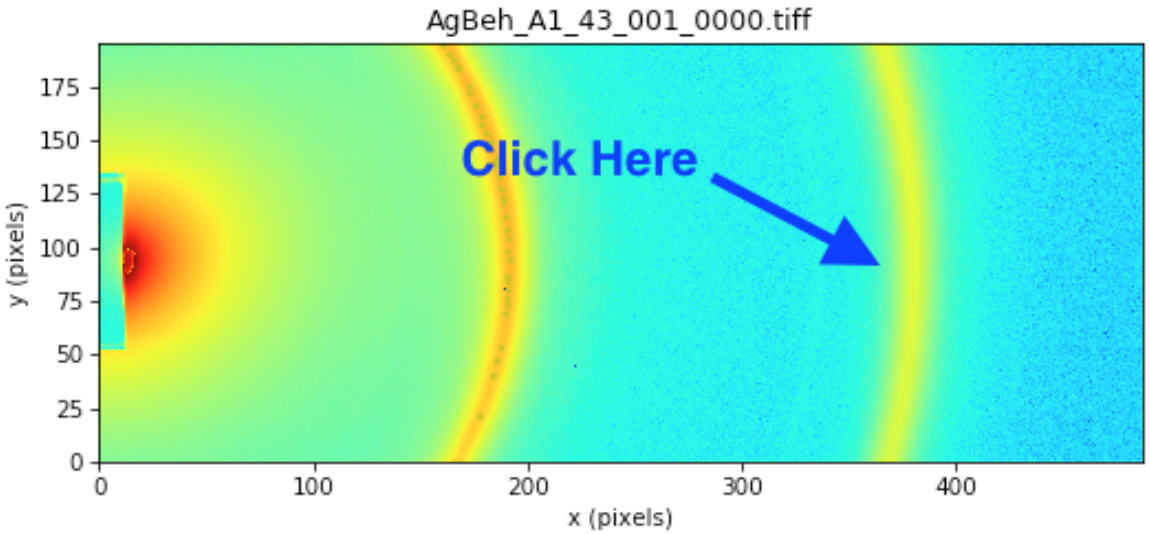
18. The peak intensity points in that ring will be automatically found, and labeled with yellow-green dots.
    - *Note:* If it didn't find very many points, try clicking again on another part of the ring, and it will add more points to your selection.
    - *Note:* If you have the same ring separated by a gap (due to detector module gaps, beamstop shadow, or geometry, click on the separated parts of the rings to add points from all sections. The autofind algorithm will only find peaks in contiguous regions.
    - *Tip:* Due to the color map selected, the points may be hard to see. Try changing to the heat map to see the selected points, like in the image below.



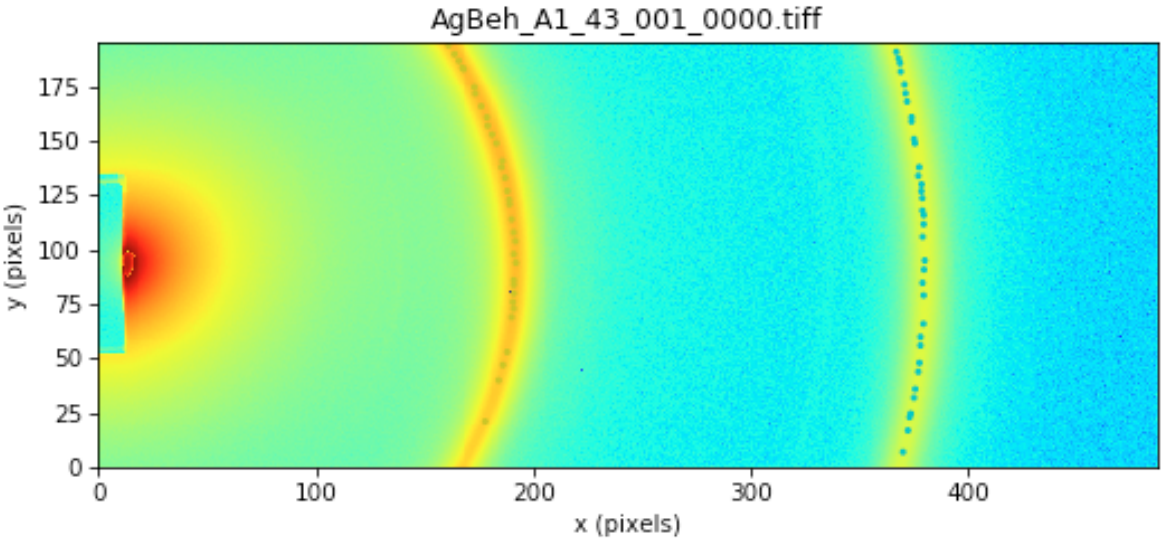
19. Change the "Ring #" to 1.



20. Click on a peak intensity point of the second visible ring.



21. The peak intensity points in that ring will be automatically found, and labeled with blue dots.

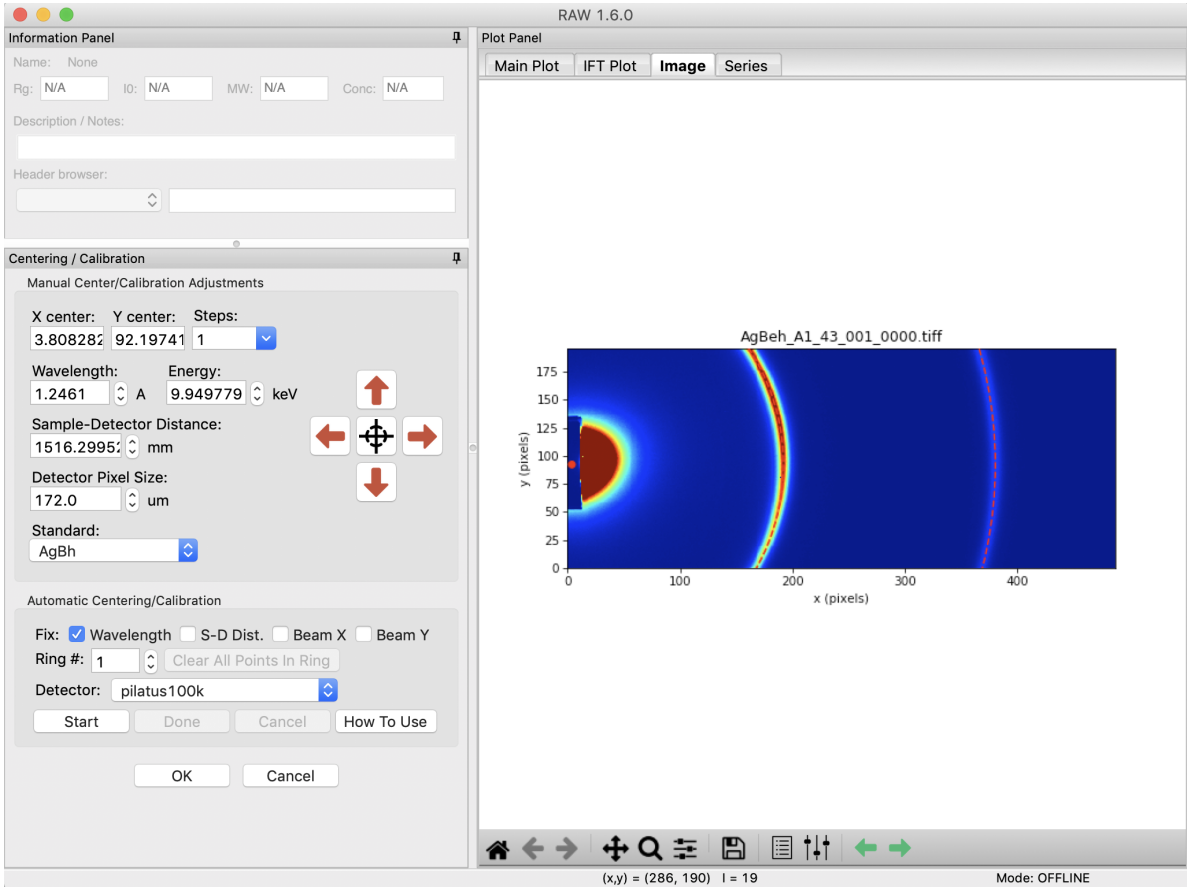


22. Click the "Done" button in the "Automatic Centering/Calibration" panel and beam position, sample to detector

distance, and detector pixel size will all be automatically filled in.

- *Note:* If the automatic centering fails, carry out Steps 5-7 of Part 2. Giving the system starting points that are approximately in the right place can help it refine to the precise location.

23. In the Image Display Settings, set the color scale back to Linear, and the Upper limit to 9000. You should now be able to easily see the centering rings and beam center on the image.

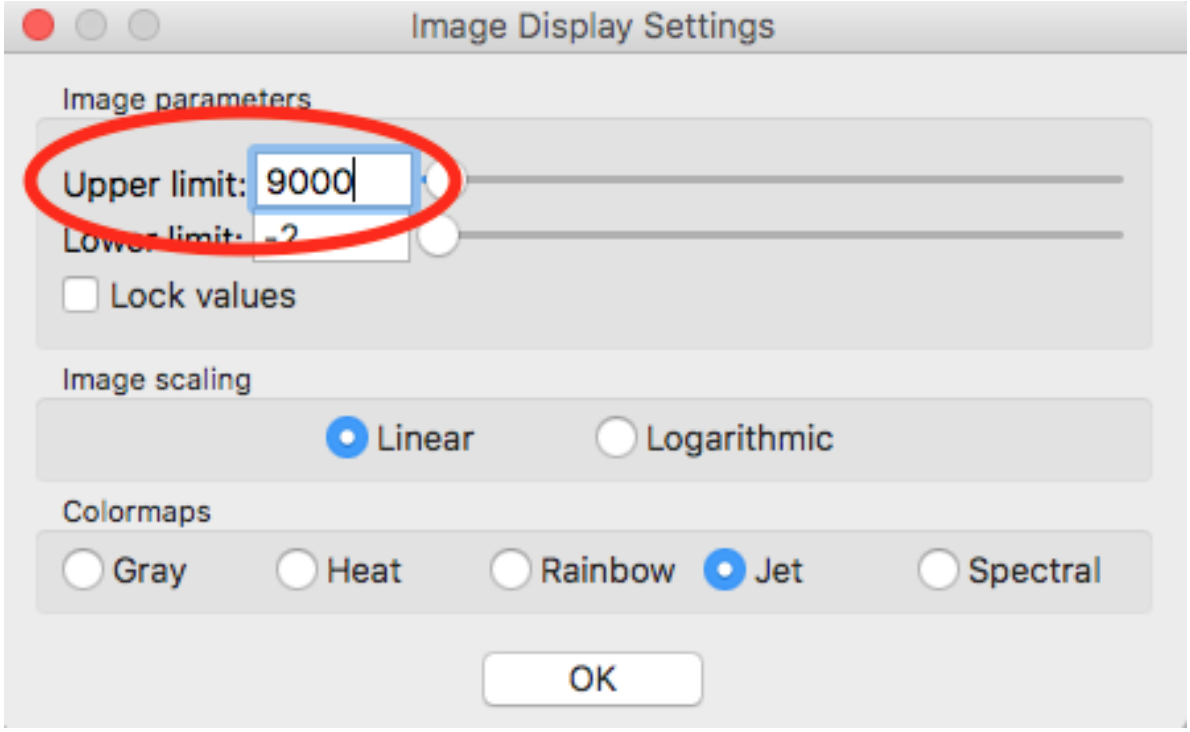


24. Click the OK button in the Centering/Calibration panel to save your settings and exit the panel.

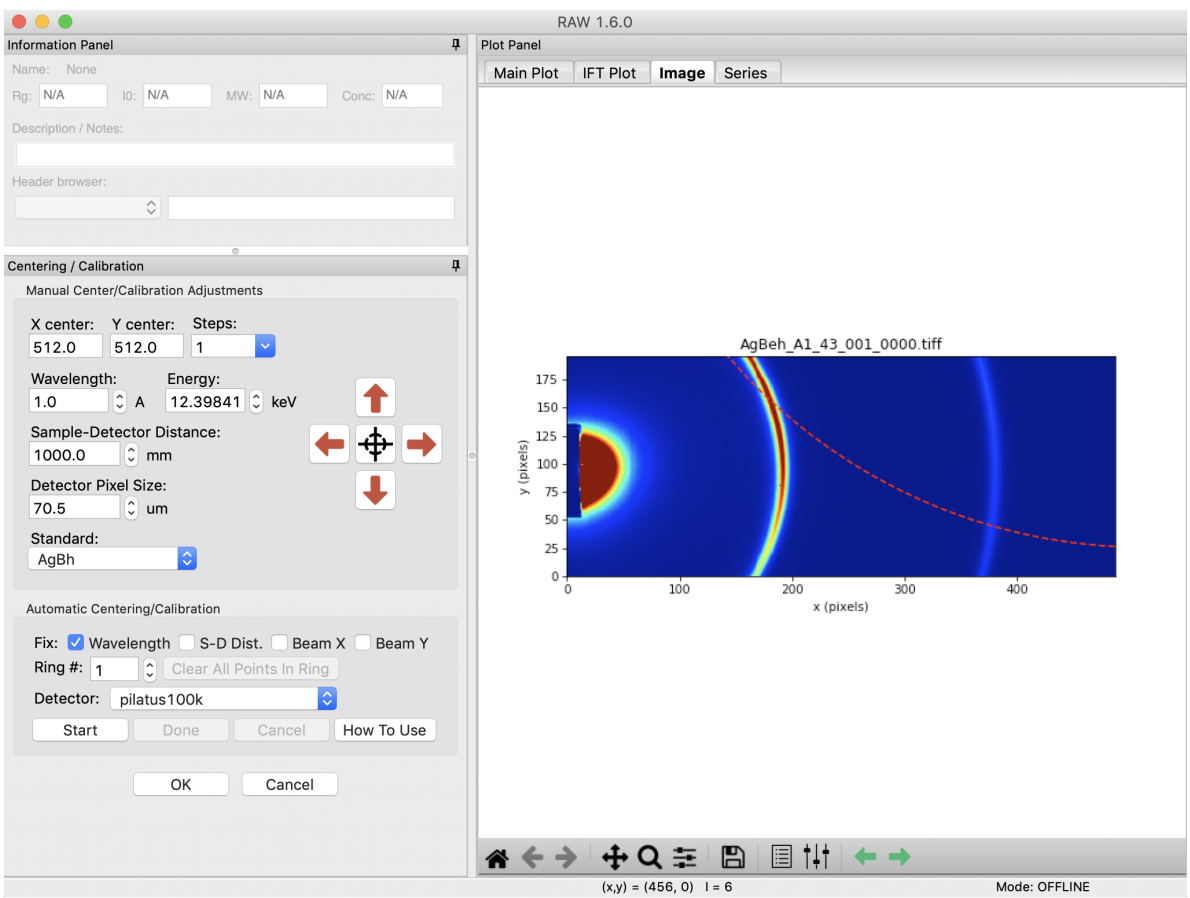
Centering and calibration – Manual method

If the automated method fails, you can also carry out centering and calibration

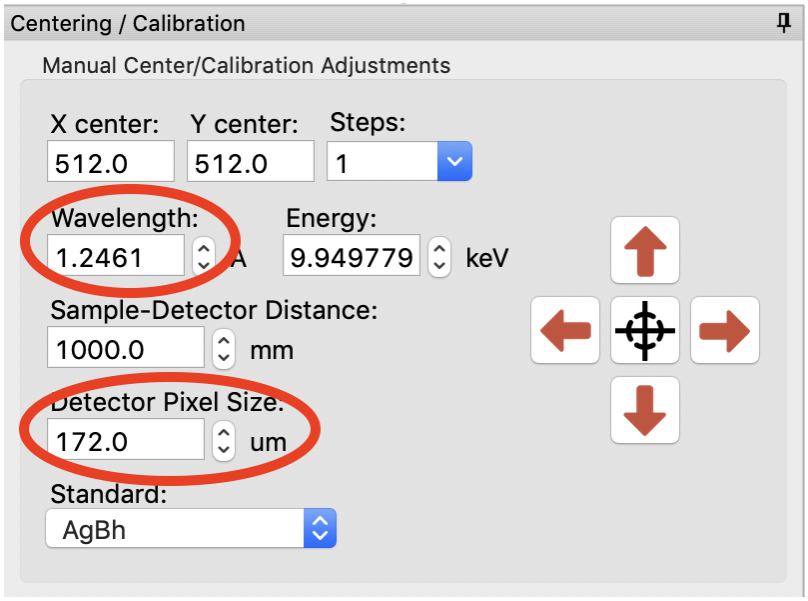
1. If you haven't already, carry out Steps 1-9 of Part 1, so that the **AgBeh_A1_43_001_0000.tiff** file is shown in the Image Plot Panel with an upper limit of 90000 set for the image.
2. In the Image Display settings (see Part 1, Step 8), set the upper limit to 9000 and click "OK" to close the settings window.



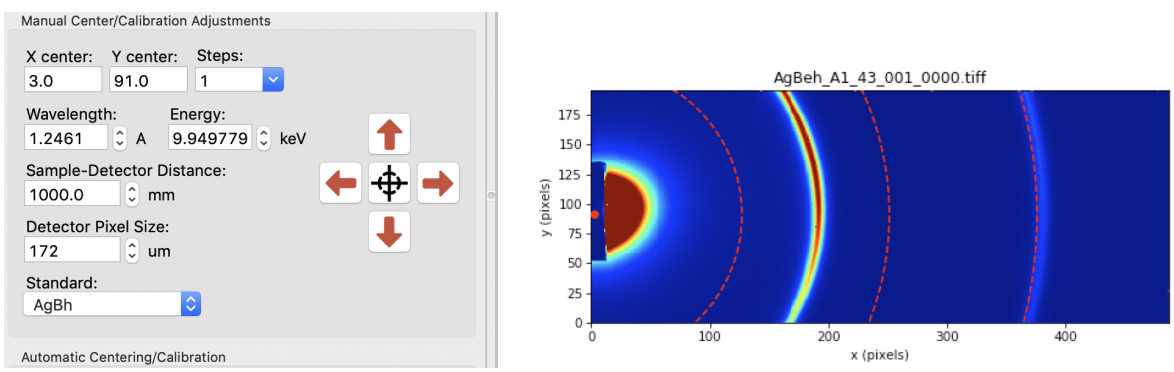
3. Open the Centering/Calibration panel by going to the Tools menu and selecting "Centering/Calibration".



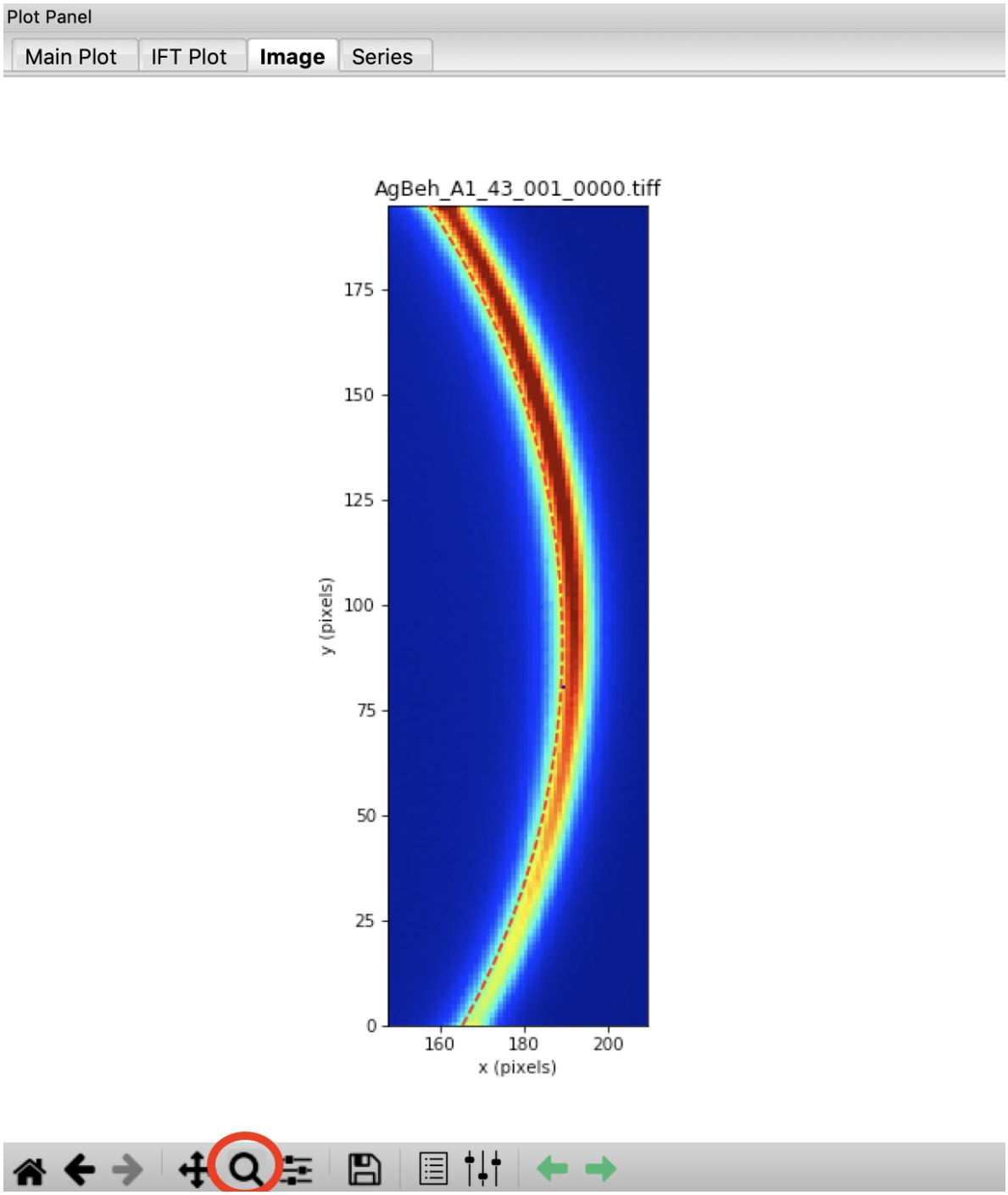
4. In the Centering/Calibration panel set the wavelength to 1.2461 Å. Set the detector pixel size to 172.0 μm.
   - *Note:* The x-ray energy/wavelength is a previously known value, and is not found in RAW.



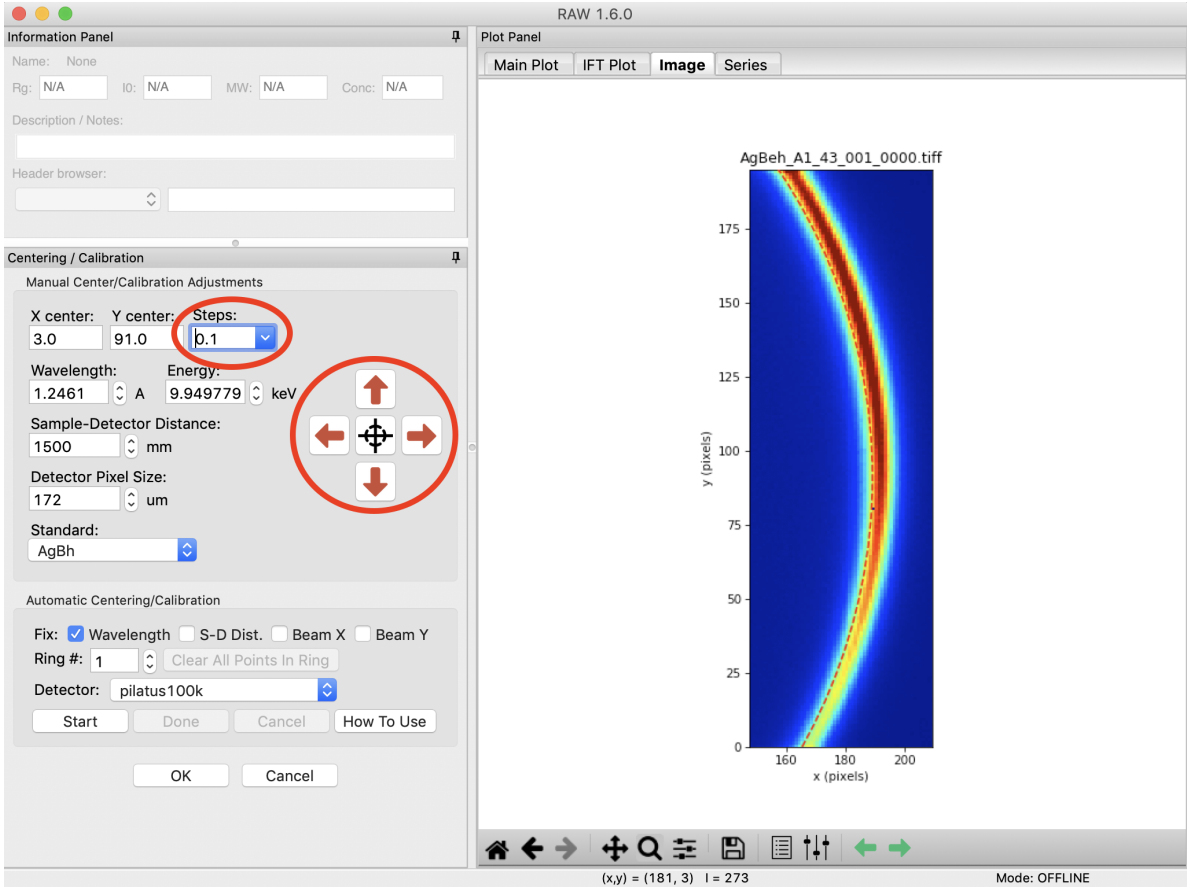
5. We know that the beam is probably near the center of the beamstop. Click on the crosshairs button and click in the center of the beamstop.
   - *Note:* The beamstop is the dark blue patch extending out from the center of the left edge of the detector.
6. *Checkpoint:* Once you have done that, three rings should be displayed on the image, and the current beam center is shown as a red dot on the image. This is shown below.



7. Set the Sample-Detector Distance to 1500 mm.
   - *Note:* The rough sample to detector distance is known from direct measurement of the experimental setup. Calibration can be done without this knowledge, but will take longer.
8. Using the zoom tool at the bottom of the Plot window, zoom in on the first ring shown on the detector.



9. Using the red arrow buttons, move your beam center until the dashed red line (first silver behenate ring) matches with the hottest (most red) pixels of the first ring shown on the detector.
   - *Tip:* Change the step size up or down to adjust how far the arrows move the rings.



10. Use the home button to zoom out to the whole extent of the detector image.

11. Open the image display settings and set the upper limit to 2000. Zoom in on the second silver behenate ring on the image (near x=380). Use the spin controls (up/down arrows) on the sample detector distance to adjust the distance until the red dashed ring matches perfectly with the hottest pixels on the detector image.

12. Zoom back out to the full extent of the image. Set the image display upper limit back to 9000. Zoom in on the first ring, and adjust the beam center position until that ring perfectly matches the hottest pixels on the detector.

    - *Tip:* If you had the ring properly centered in y the first time, you should only need to adjust the x position (left/right).

13. Iterate steps 10-13 until both rings are well aligned. You should find an X center of ~4, a Y center of ~ 92.5, and a sample-detector distance of ~1517 mm.

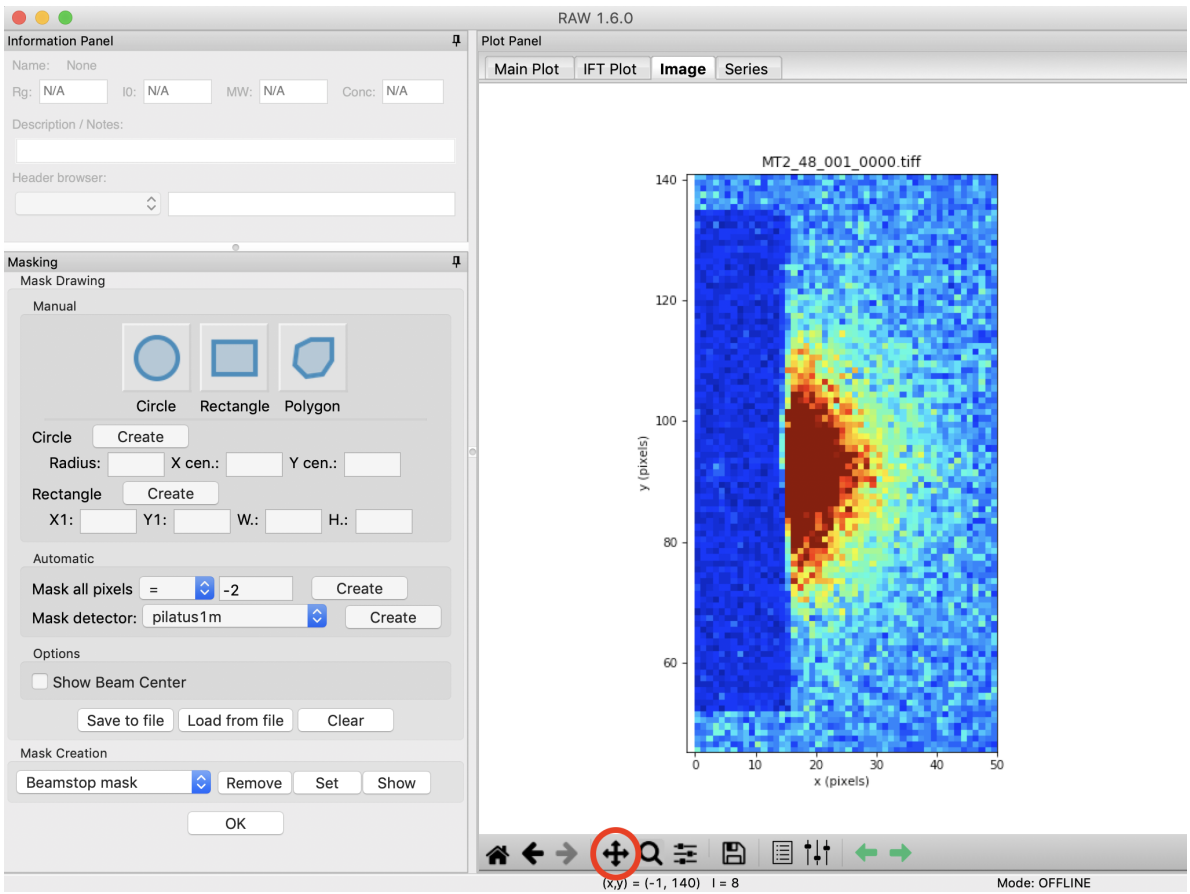
    - *Tip:* You can adjust the step size for moving the beam center using the "Steps" menu. You can either pick a value from the dropdown menu, or enter your own value in the box.
    - *Tip:* The spin controls adjust the last digit of the value in the box. So if you want to adjust the sample-detector distance more finely than 1 mm, set the distance to (for example) 1500.0 and then the spin controls will change the distance in steps of 0.1 mm.

14. Compare the manually found center and sample detector distance with those you found automatically in Part 1 (if you did that). If you did a careful job in this part, they should compare well.

Masking

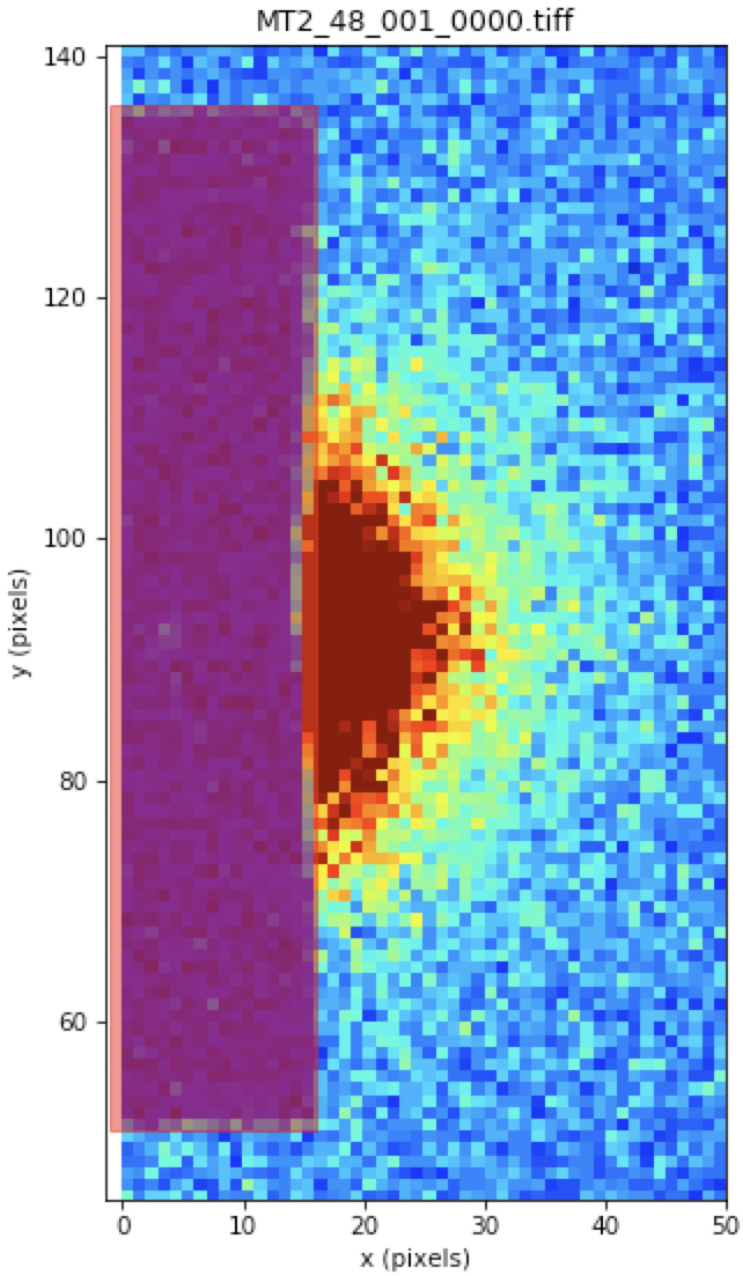
This section teaches you how to mask out unwanted portions of your image, such as the beamstop and bad detector pixels. It assumes you have just done *Part 1* or *2*. If not, open RAW as in Step 1 and set your data folder as in Step 6 of *Part 1*.

1. In the Files tab, select the **MT2_48_001_0000.tiff** file and click the show image button.
2. Set the image upper limit to 50.
3. Open the masking panel by clicking "Masking" in the Tools menu.
4. Zoom in around the beamstop.
   - *Note:* The beamstop is the blue rectangular area on the left edge of the detector
5. Select the Pan tool and left click and drag the image to the right until you can see a blank (white) canvas to the left of the beamstop.



6. Click on the rectangle mask drawing button and click on a point in the white part of the canvas just above the edge of the beamstop.
7. Move the cursor until the red outline of the rectangle covers the beamstop from top to bottom, and out to the right edge of the middle of the beamstop. Click again to make the rectangle mask.
   - *Note:* There will still be a bit of the beamstop at the bottom right edge that is not covered by this rectangle.
   - *Tip:* If you mess up drawing the mask, click on the masked region (shaded red) and click your backspace/delete key to remove it.
   - *Tip:* You can also resize a rectangle (or circle) mask my right clicking on it and selecting resize.

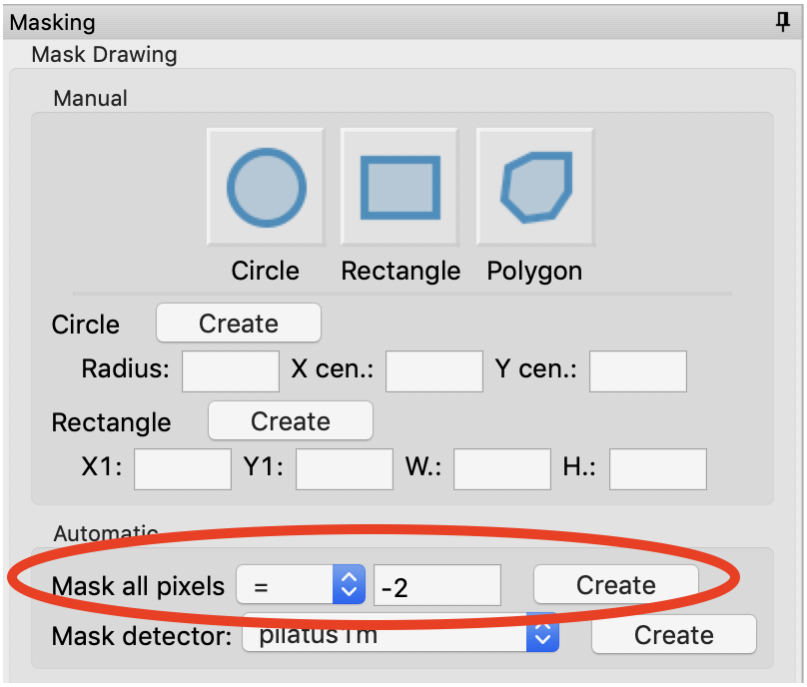
- *Tip:* You can create a rectangle of defined position and size by entering the x and y coordiantes of the bottom left corner and the width and height of the rectangle then clicking 'Create'.



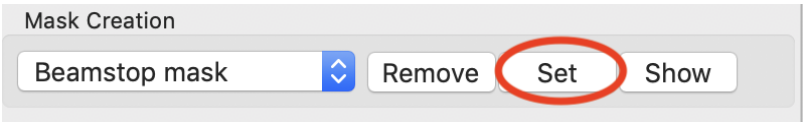
8. This beamstop is quite square, so a simple square mask works. A circle is also easy to draw:
   - Click on the Circle mask drawing button.
   - Click at the center
   - Move the mouse out to make it the size you want, and click again to finish the circle.
   - *Tip:* Like a rectangle, you can also draw a circle of defined position and size by entering the radius, x center, and y center, and click the 'Create' button.
9. If you need to draw another (non-square or circle) shape, you would do the following:

- Click on the Polygon mask drawing button. Left click to place the first vertex.
- Continue left clicking to place more vertices to draw the desired shape.
- Right click to connect the last point you put down to the first point, and finish drawing the polygon.

10. Zoom back out to the full extent of the image.
11. Next you need to mask out the bad pixels on the detector. On a Pilatus detector, bad pixels usually have a value of -2.
12. In the automatic section of the Mask Drawing controls, set Mask All Pixels = -2 and click 'Create'.
    - This control allows you to mask pixels at, above, or below a given threshold. It is particularly useful for things like bad pixels, where the value is known.



13. In the masking panel, make sure that "Beamstop mask" is selected in the Mask Creation dropdown menu. Click the set button to set the mask you just made as the beamstop mask.



14. Click the OK button to exit the masking panel.

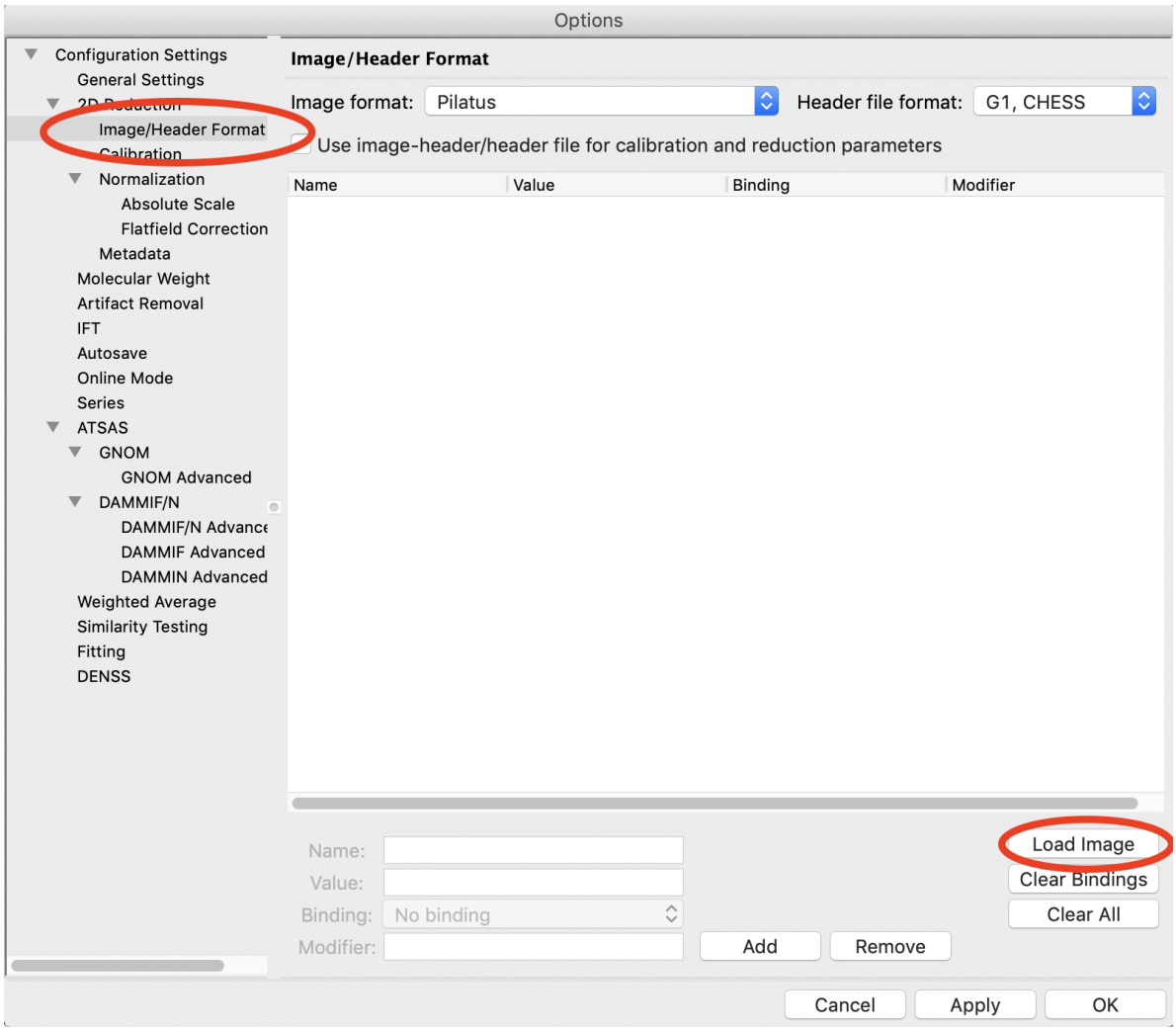
Additional Tips

1. If you have a larger detector with panel gaps, the Mask detector function in the Automatic section of the Mask Drawing tools will generate a gap mask for many different detectors.
2. You can resize rectangle and circle masks by right clicking on them and selecting resize.
3. You can 'invert' a mask, which for a beamstop mask means it will only include pixels in the mask. This can be useful if, for example, you have a circular shadow from your flight tube on your detector.
4. If you know the mask parameters, you can put in coordinates for a circle or rectangle and create a mask with a known position and size.

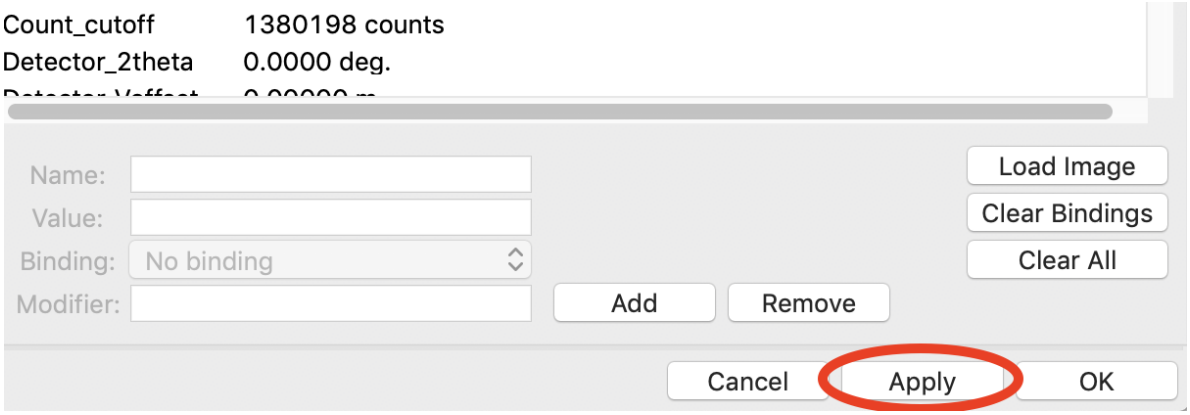
Setting normalization and other options

This section teaches you how to set up normalization by a beamstop counter, and other options. It assumes you have completed *Parts 1* (or *2*) and *3*.

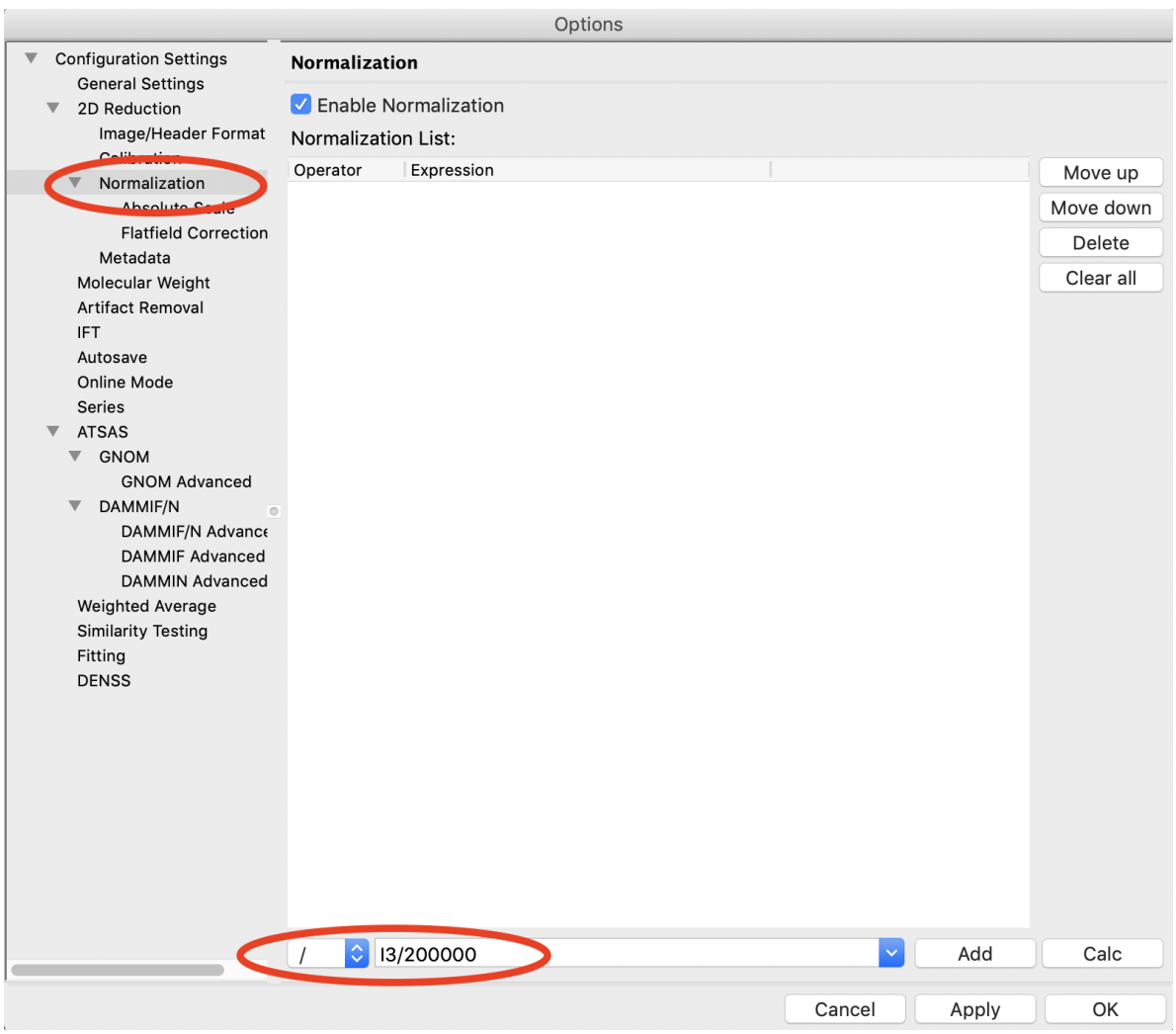
1. Open the Options window by selecting "Advanced Options" in the Options menu.
2. In the window that shows up select the Image/Header Format section on the left. In the area on the right click the Load Image button.



3. In the window that pops up, select the **AgBeh_A1_43_001_0000.tiff** file. Click the Open button.
   - *Note:* You can select any image of the appropriate type, not just the behenate.
4. In the Image/Header Format window you should now see header values loaded into the list. Click the Apply button at the bottom of the screen.



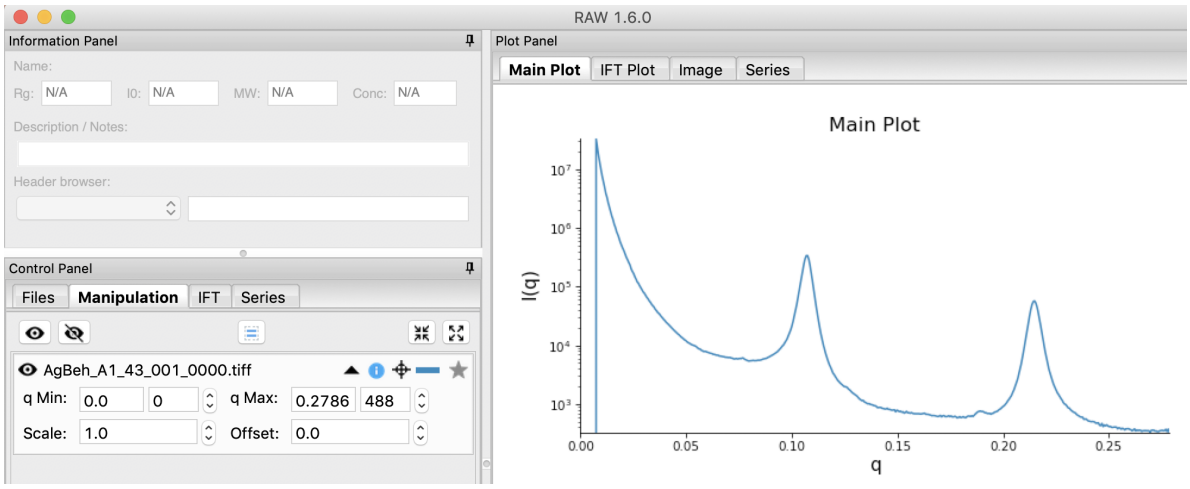
5. Click on the Normalization section in the options list on the left.
6. In the fields at the bottom of the Normalization panel, make sure "/" is selected in the left dropdown menu, and enter I3/200000 in the large field.
   - *Note:* It is typical in SAXS to normalize by the transmitted intensity. At the CHESS G1 beamline, the beamstop counter is name I3, which is why we are using that name in the normalization expression.



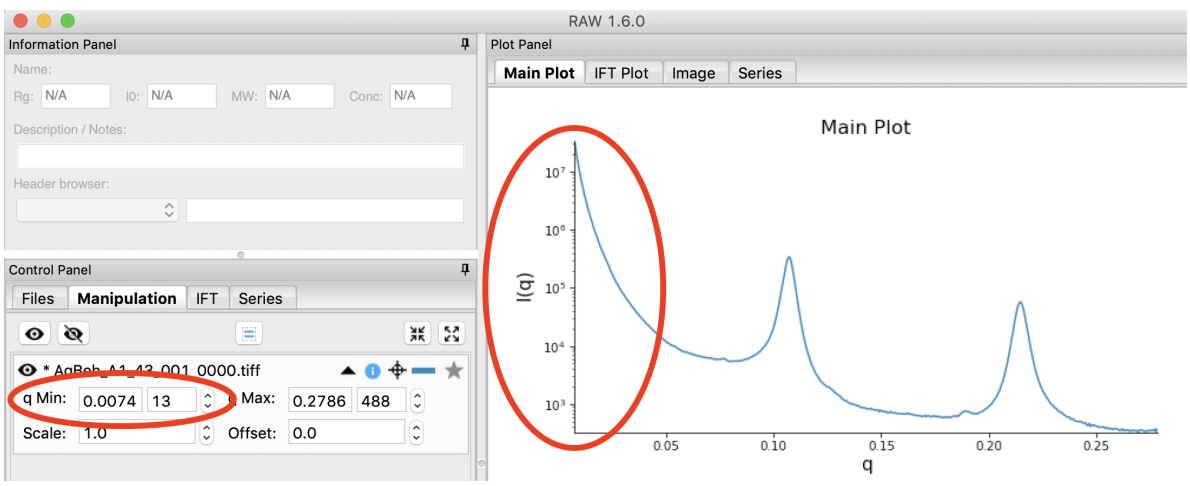
7. Click the Calc button to evaluate the expression for the counter values loaded in the Image/Header Format tab.

You should get a value of 0.02404.

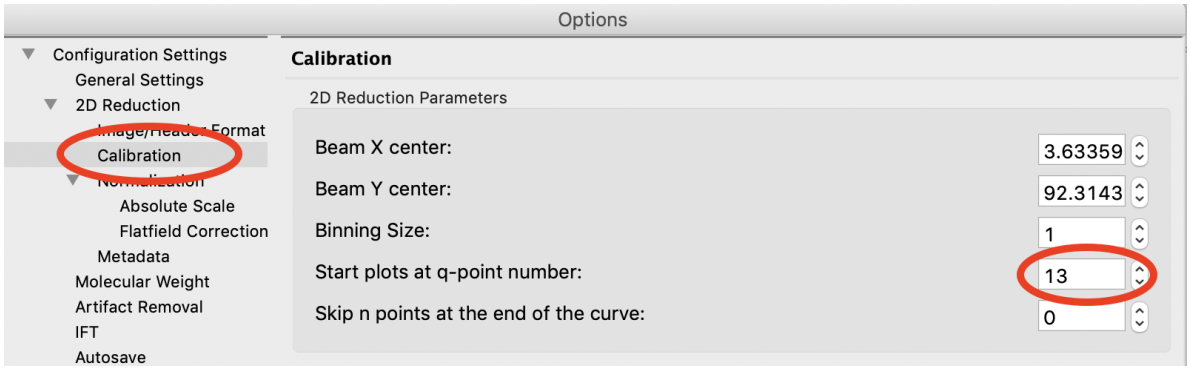
8. Click the Add button to add the expression to the normalization list.
9. Make sure the "Enable Normalization" checkbox at the top of the page is checked.
10. Click OK to exit the options window.
11. In the file list, select the **AgBeh_A1_43_001_0000.tiff** file and click the Plot button. You will see a curve get plotted in the top panel of the Main Plot.
12. Click on the manipulation tab. You will see a data item loaded in the manipulation list.



13. Adjust the start point for q Min to remove the points with zero value at the start of the curve (these are q points entirely in the mask). Set q Min so that the first point is the peak of the curve on the main plot. This should be around point 13 (depending on your mask).

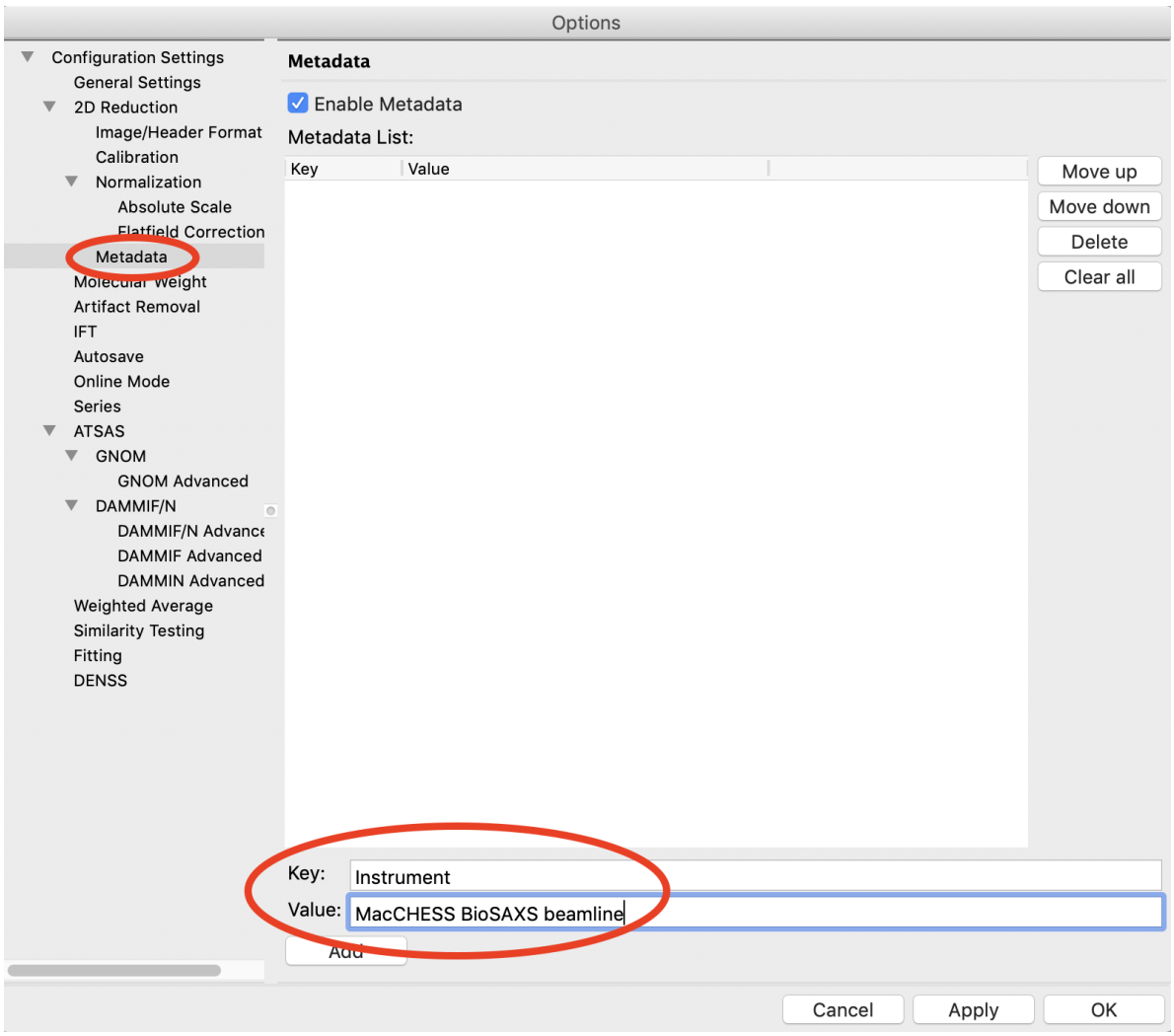


14. Open the Options window as in Step 1.
15. Click on the Calibration section in the options list on the left. Set "Start plots at q-point number" to the number you just found in Step 13.
    - *Note:* This makes it so that every curve loaded from now on will by default not display the first n points, which are covered by the beamstop.

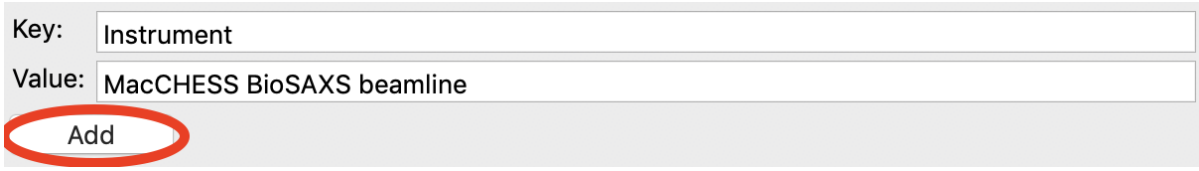


16. RAW also allows you to add arbitrary metadata to your radially integrated files. One use case is to provide metadata keys for data deposition (e.g. in the SASBDB). To start, click on the Metadata section in the options list on the left of the Options window.

17. In the 'Key' field enter 'Instrument'. In the 'Value' field enter 'MacCHESS BioSAXS beamline'.



18. Click the Add button.

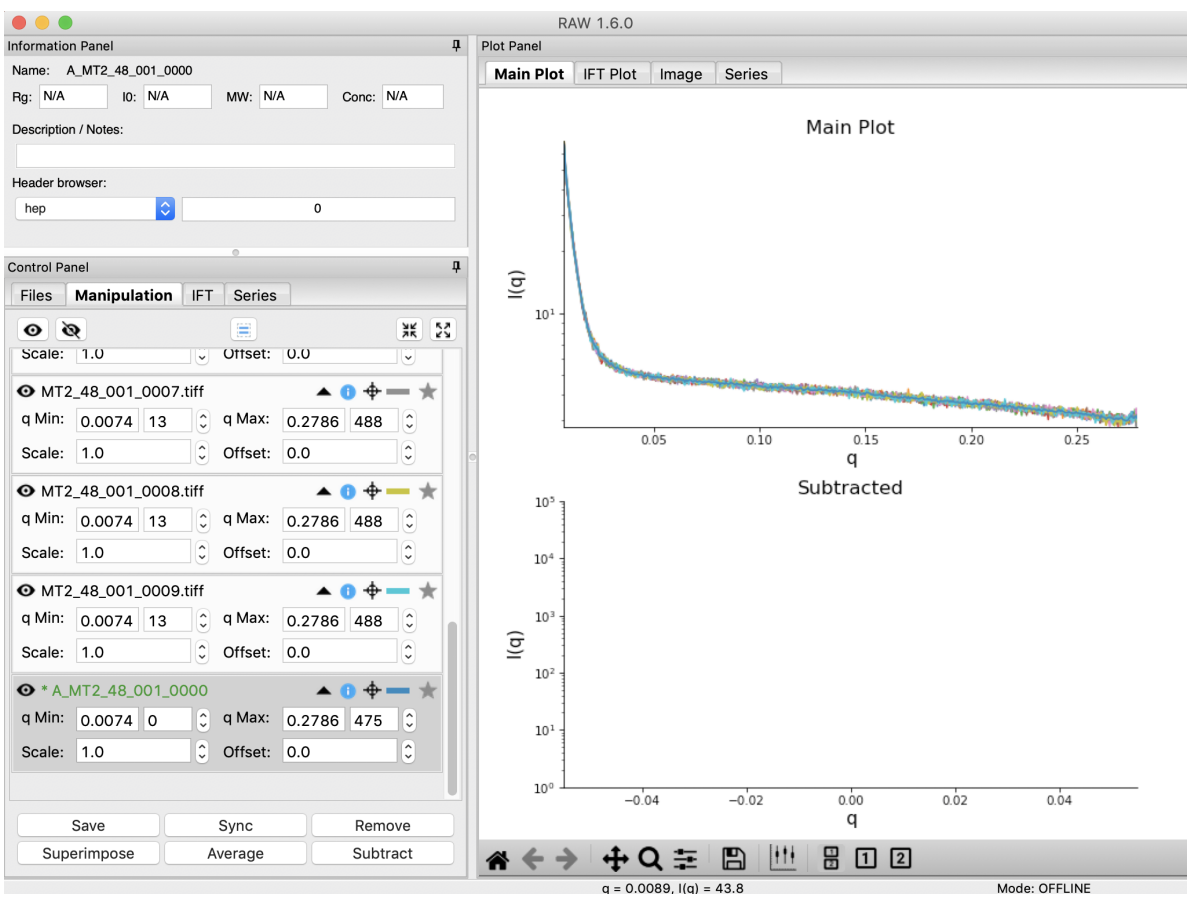


19. Click the OK button to exit the options window and save your changes.
20. You have configured everything necessary, and are now ready to save your settings. Go to the File menu and select "Save Settings".
21. Save the settings as **SAXS.cfg**.
22. These settings can now be used to process images, and can be reloaded when you open RAW by selecting "Load Settings" from the File menu.

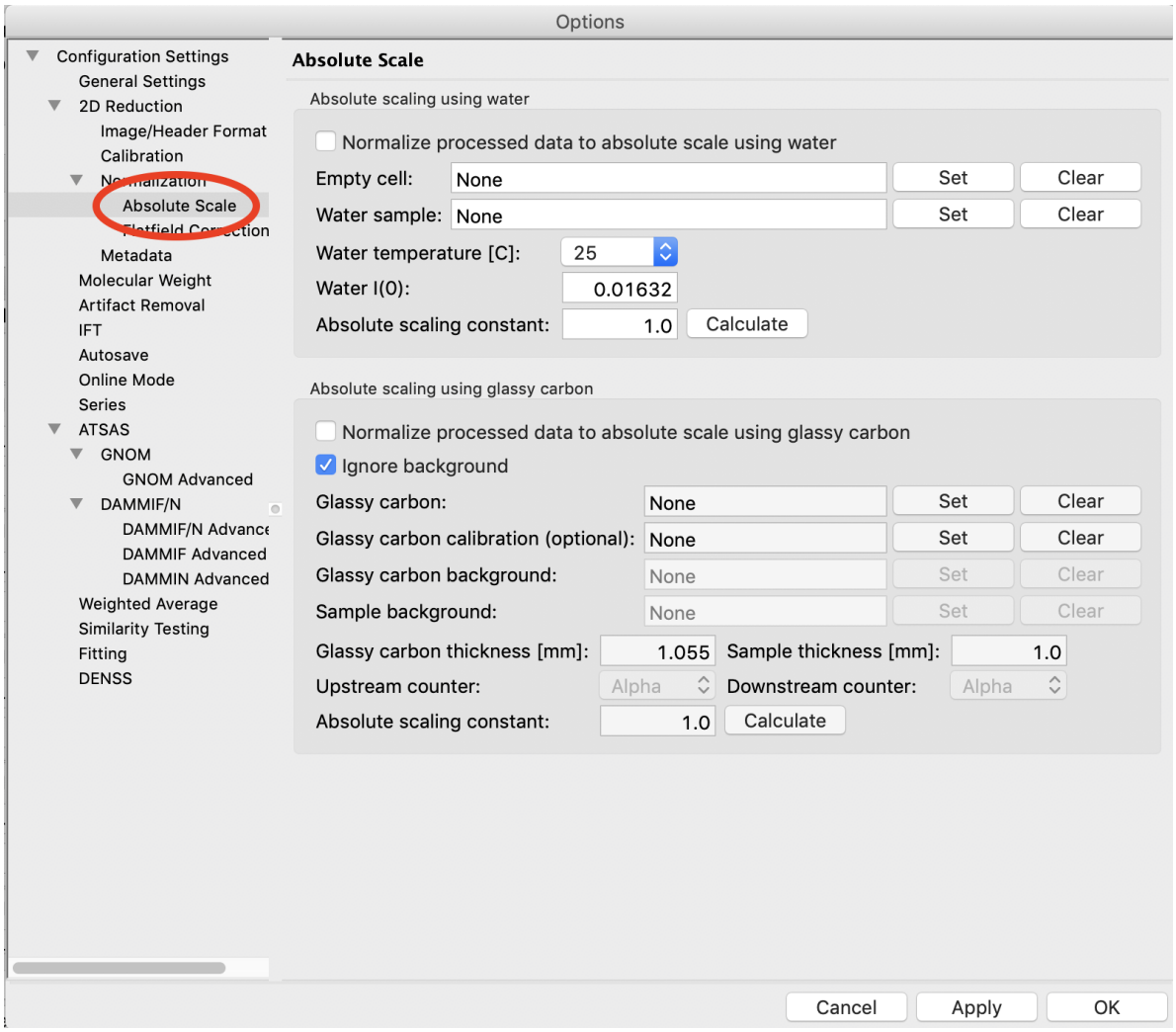
Setting absolute scale with water

This section teaches you how to set up absolute scale using water as a reference. It assumes you have completed *Parts 1* (or *2*), *3* and *4*. Note that you can use water or glassy carbon (*Part 6*) for absolute scale calibration in RAW.

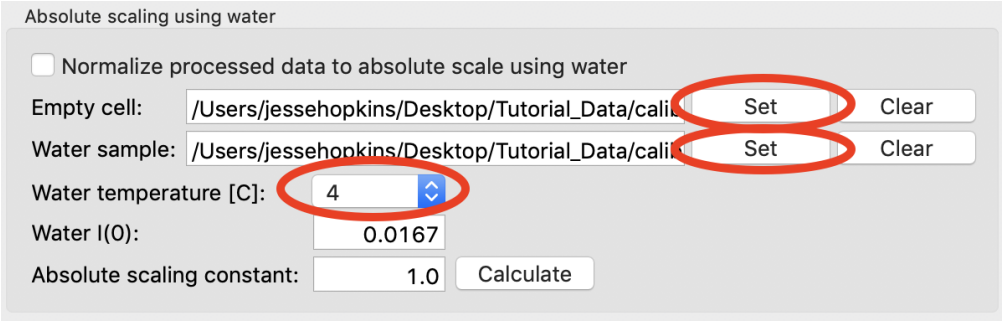
1. Using the settings from the previous parts of the tutorial, plot all of the **MT2_48_001_000x.tiff** files, where x is 0-9, on the main plot.
   - *Tip: Section 1 Part 1* of this tutorial document teaches you how to do this.
2. Average the **MT2** files you just loaded. Save the average in the **calibration_data** folder.



3. Repeat steps 1 and 2, plotting, averaging and saving, for the **water2_49_001_000x.tiff** files.
4. Open the Options window by selecting "Advanced Options" in the Options menu.
5. Click on the Absolute Scale section in the options list on the left.

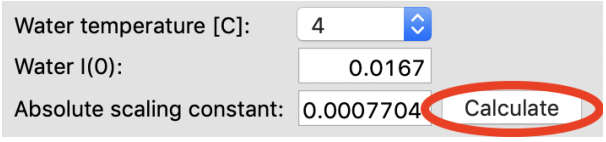


6. Click on the Empty cell "Set" button and select the **A_MT2_48_001_0000.dat** file.
7. Click on the Water sample "Set" button and select the **A_water2_49_001_0000.dat** file.
8. Set the Water temperature to 4 C.

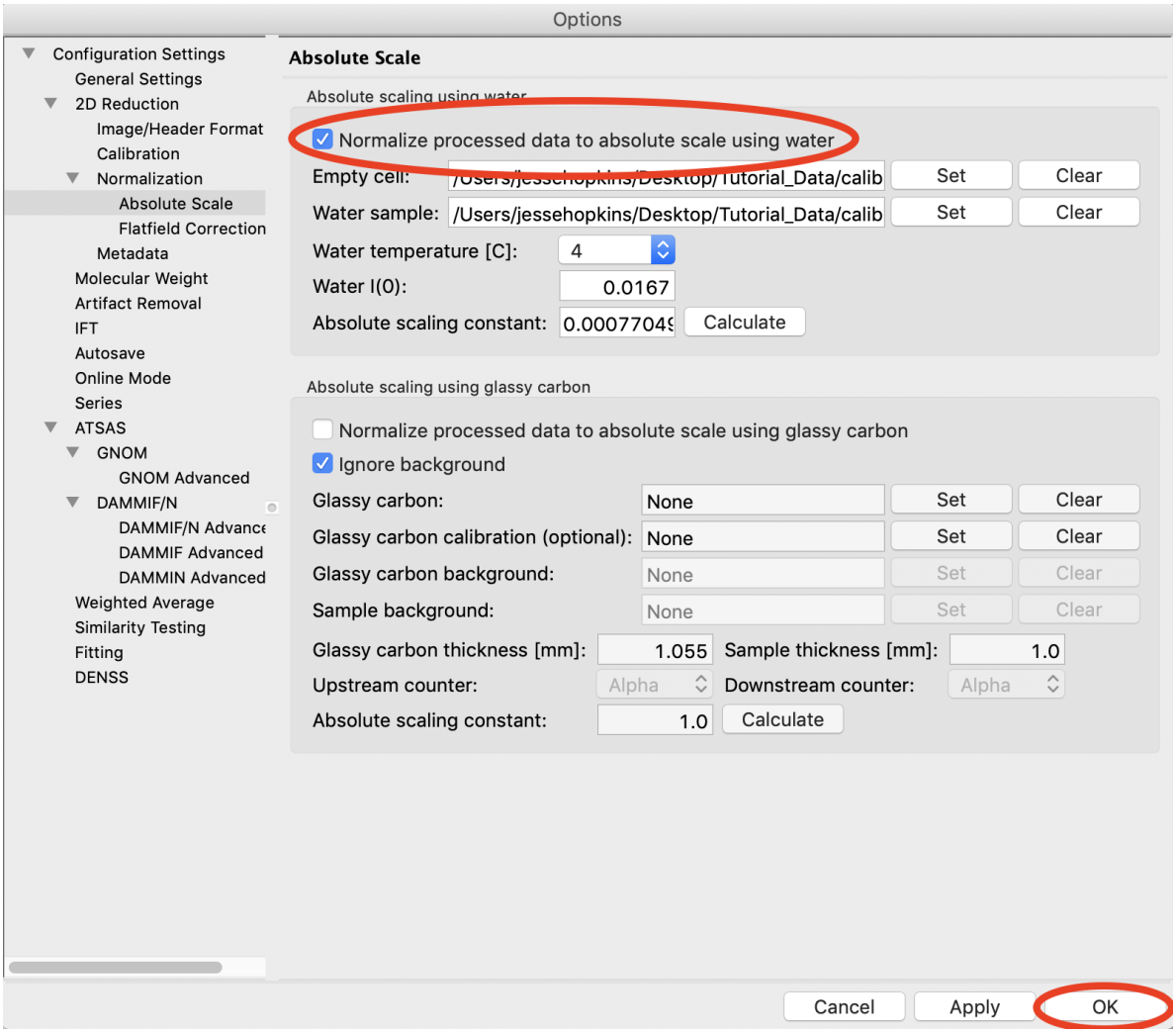


9. Click the Calculate button to calculate the Absolute Scaling Constant. You should get a value near 0.00077.

- *Tip:* You can also use images to set the absolute scale. This may give worse results, as the signal to noise of the averaged file should be better than for a single image.
- *Note:* It is important that you not change your normalization settings once you have set the absolute scaling constant. If you do, you will have to recalculate the absolute scaling constant. Also, make sure absolute scale is turned off before you calculate the scale constant, otherwise you will get a bad scaling constant (see the manual for details).



10. Check the "Normalize processed data to absolute scale" checkbox. Click "OK" to exit the advanced options window and save the changes.



11. Save the settings for later use.

Setting absolute scale with glassy carbon

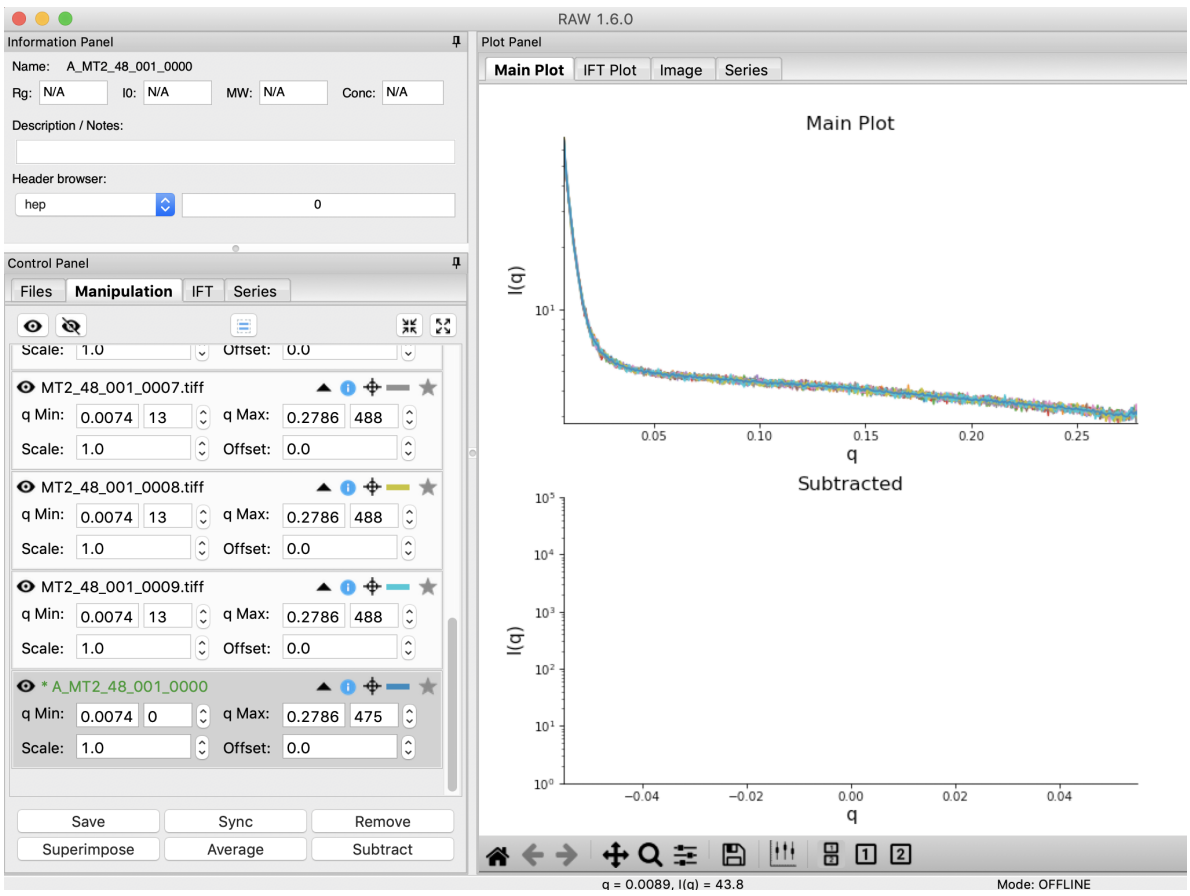
This section teaches you how to set up absolute scale using glassy carbon (NIST SRM 3600) as a reference. It assumes you have completed *Parts 1* (or *2*), *3* and *4*. Note that you can use water (*Part 5*) or glassy carbon for absolute scale

calibration in RAW.

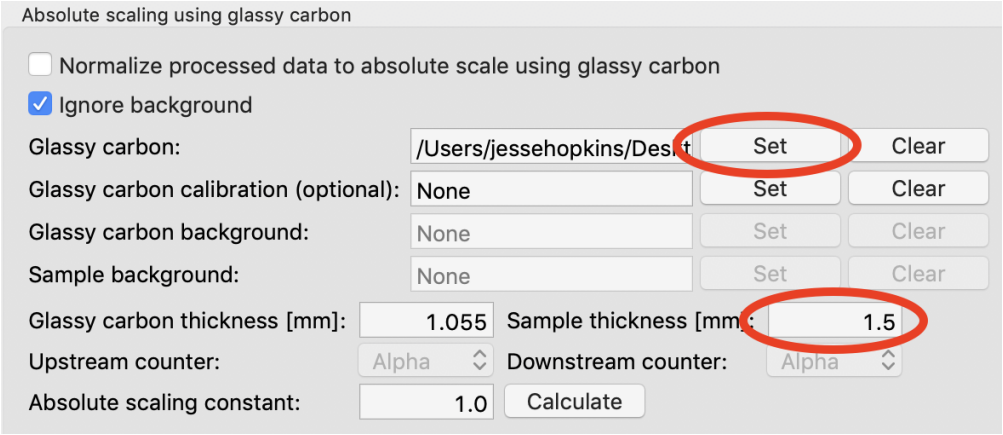
There are two ways to use glassy carbon as a standard in RAW. One way follows the NIST protocol, and will deliver the most accurate results. However, this method depends on all measurements having reliable flux measurements upstream and downstream of the sample. It also requires accurate measurements of the background of the glassy carbon measurement and the sample measurements. The second way is more similar to that used by water, in that it essentially ignores the background (assumes it to be small). This approach only requires regular normalization and a single measurement of the background for the glassy carbon sample.

The simple approach, "ignoring" background:

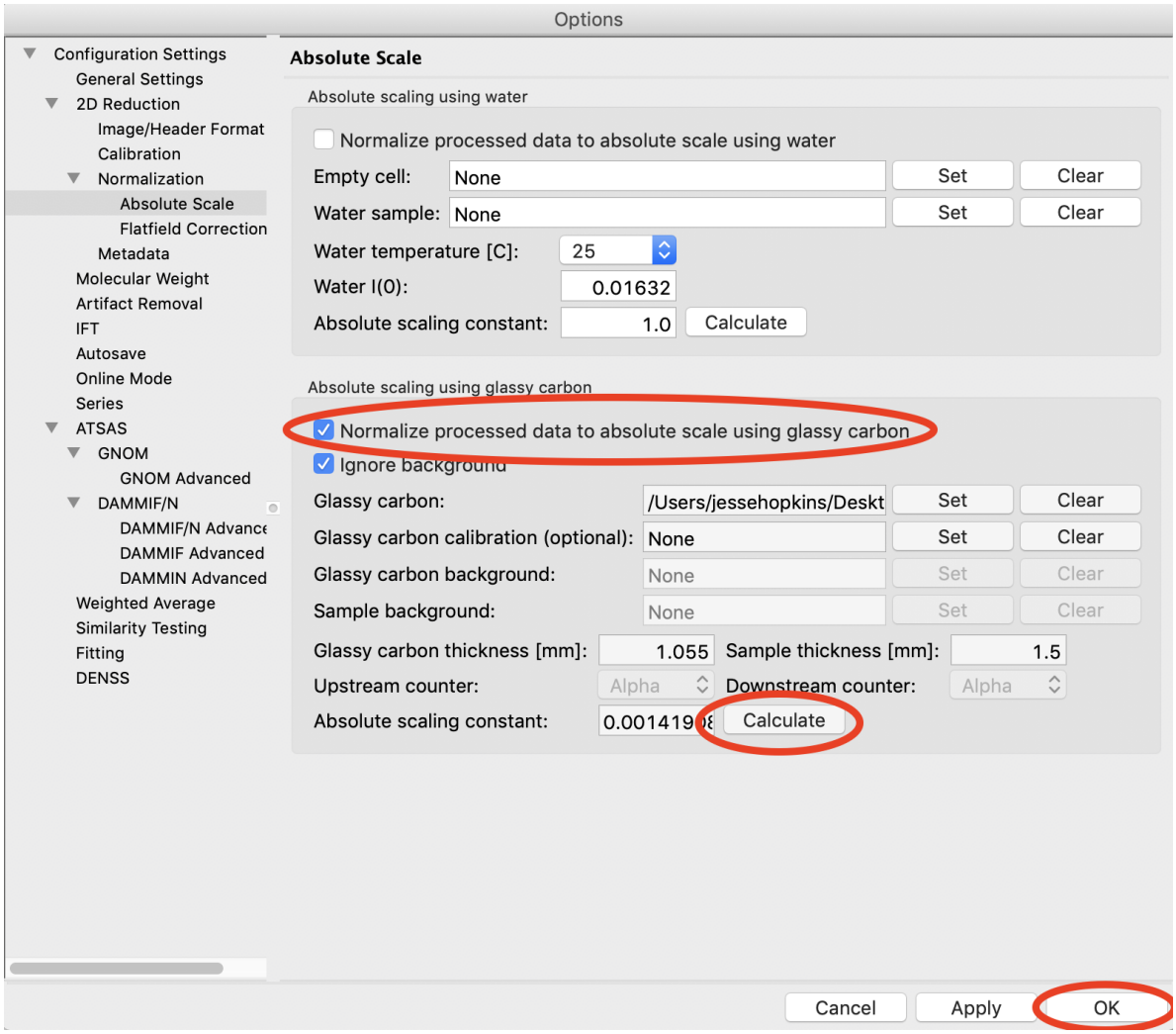
1. Load/use the settings from part 4 (without absolute scale set from water, part 5).
2. Plot all of the **glassy_carbon_41_001_000x.tiff** files, where x is 0-9, on the main plot.
   - *Tip: Section 1 Part 1* of this tutorial document teaches you how to do this.
3. Average the **glassy_carbon** files you just loaded. Save the average in the **calibration_data** folder.
4. Open the Options window by selecting "Advanced Options" in the Options menu.
5. Click on the Absolute Scale section in the options list on the left.



6. Click on the Glassy carbon "Set" button and select the **A_glassy_carbon_41_001_0000.dat** file.
7. Set the Sample thickness to 1.5 mm.



8. Click "Calculate" button. You should get something near 0.0014.
   - *Note:* It is important that you not change your normalization settings once you have set the absolute scaling constant. If you do, you will have to recalculate the absolute scaling constant. Also, make sure absolute scale is turned off before you calculate the scale constant, otherwise you will get a bad scaling constant (see the manual for details).
9. Check the "Normalize processed data to absolute scale using glassy carbon" checkbox.
10. Click "OK" to exit the advanced options panel, saving the changes.

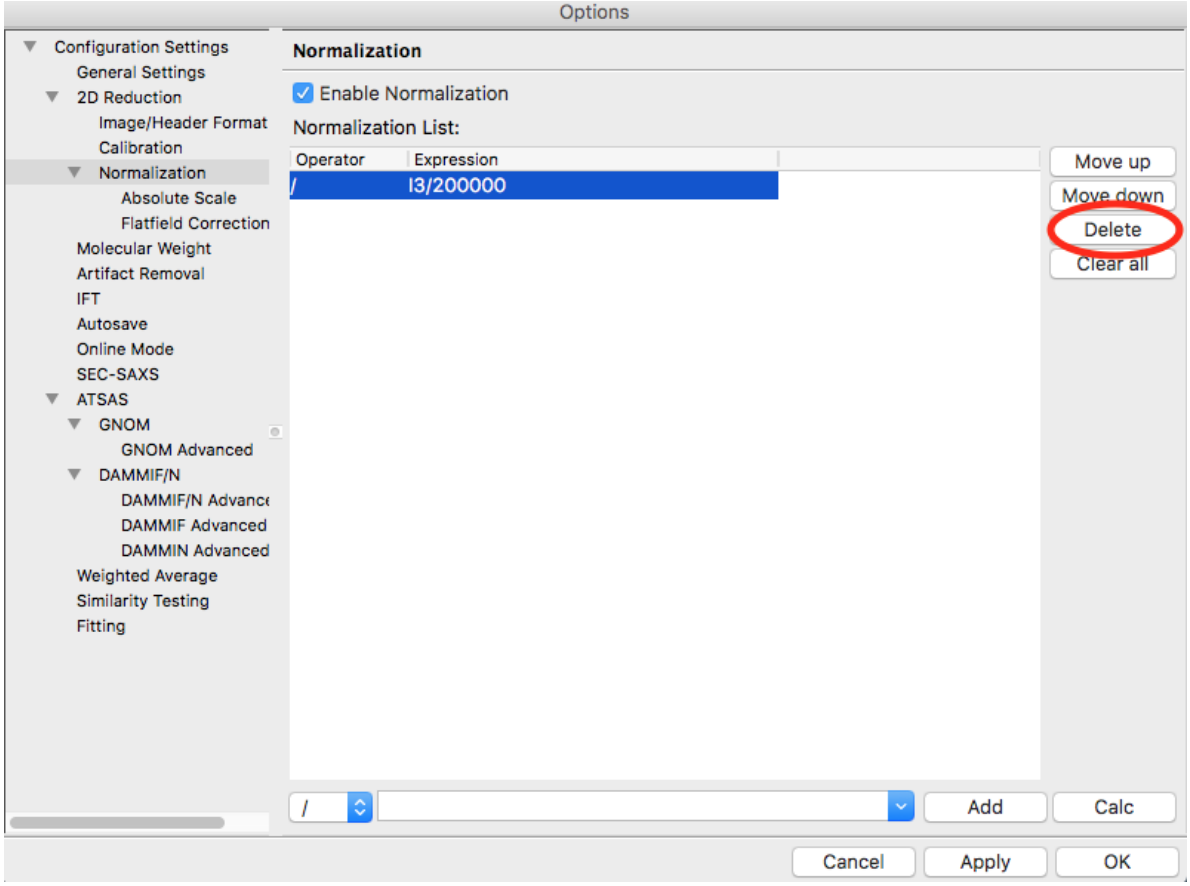


11. Save the settings for future use.

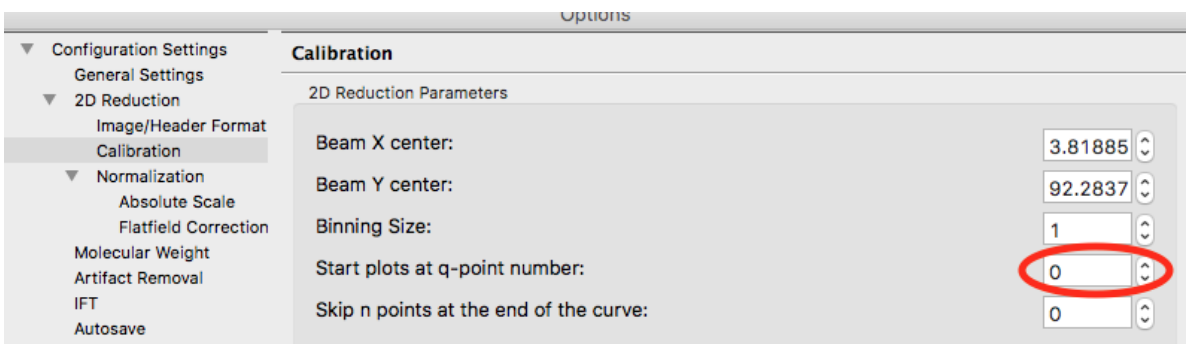
The full, NIST recommended approach:

Important note: All of the normalization (including flux, transmission, etc) happens through the absolute scale panel. You shouldn't have anything set in the Normalization panel (unless you are doing something like subtracting off a constant pedestal from the image).

1. Load/use the settings from part 4 (without absolute scale set from water, part 5).
2. Open the Options window by selecting "Advanced Options" in the Options menu.
3. Click on the Normalization section in the options list on the left.
4. Remove any/all items in the Normalization List by highlighting them in the list and clicking the "Delete" button.

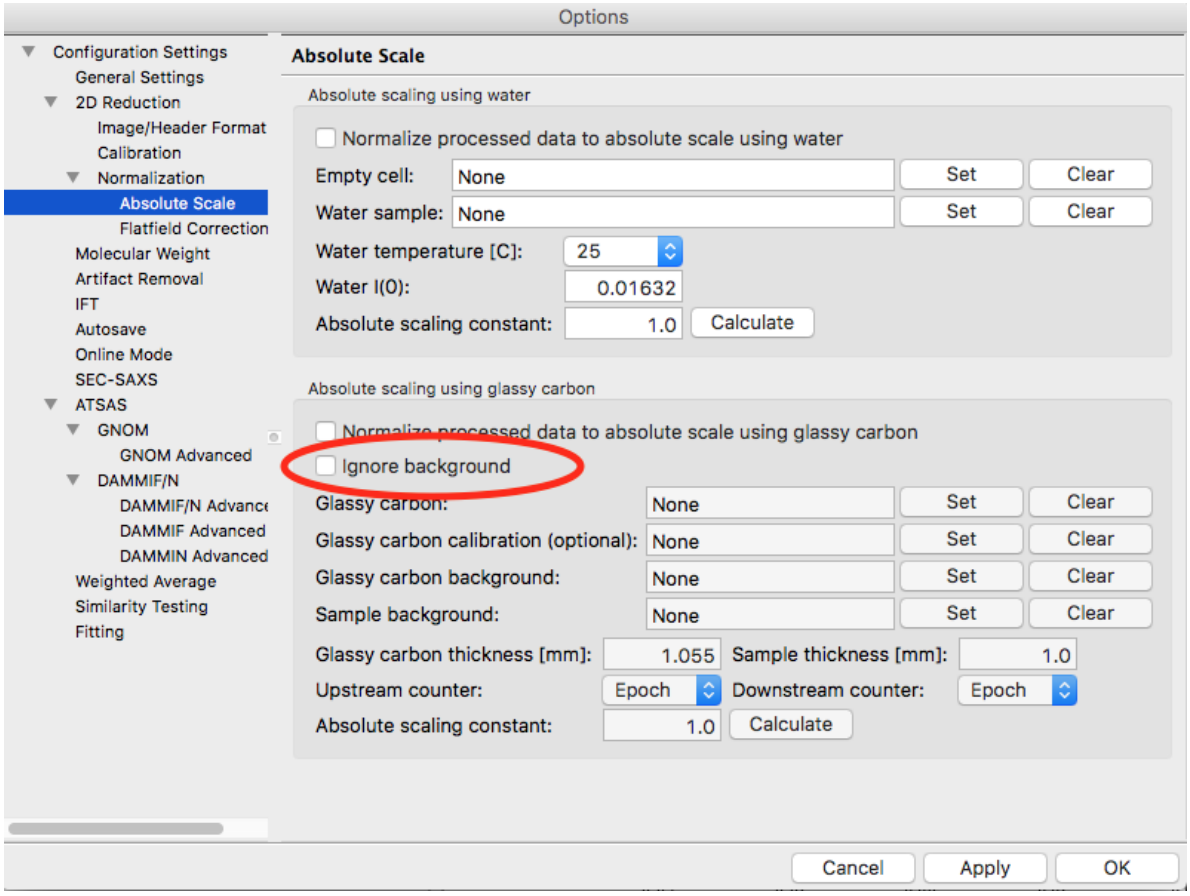


5. Turn off any absolute scaling already in place.
6. Click on the Calibration section in the options list on the left.
7. Change the "Start plots at q-point number" to 0.

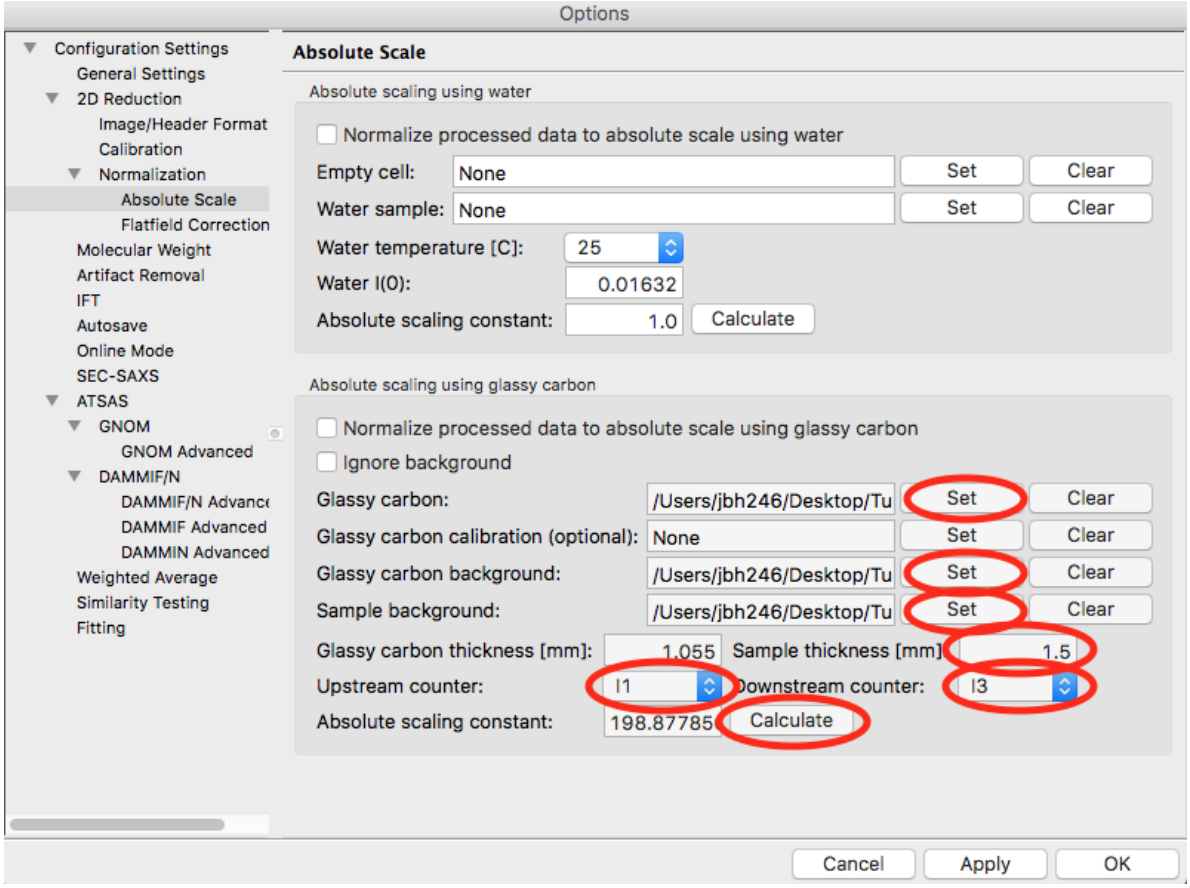


8. Click "OK" to exit the advanced options window and save the changes.
9. Plot the **glassy_carbon_41_001_0000.tiff** file.
   - *Tip:* *Section 1 Part 1* of this tutorial document teaches you how to do this.
10. Save the **glassy_carbon** profile in the **calibration_data** folder.
11. Plot, average, and save the **vac_37_001_000x.tiff** and ** MT2_48_001_000x.tiff ** files, where x is 0-9.
    - *Tip:* Because you aren't normalizing by beam intensity, these averages may have profiles that are not similar (see *Section 1 Part 6*). In that case, average just the similar profiles.

12. Open the Options window and select the Absolute Scale section.
13. Uncheck the Ignore background checkbox.



14. Click the Glassy carbon "Set" button and select the **glassy_carbon_41_001_0000.dat** file.
15. Click the Glassy carbon background "Set" button and select the **A_vac_37_001_0000.dat** file.
16. Click the Sample background "Set" button and select the **A_MT2_48_001_0000.tiff** file.
17. Set the Sample thickness to 1.5 mm.
18. Set the Upstream counter to I1.
19. Set the Downstream counter to I3.
20. Click the "Calculate" button. You should get an absolute scaling constant near 198.
    - *Note:* This approach will only work if the .dat files you select for the glassy carbon, glassy carbon background, and sample background contain the upstream and downstream counter values. This happens automatically with RAW. Otherwise, you should use images, which will have more noise, but should allow RAW to find all of the appropriate counter values.
    - *Note:* It is important that you not change your normalization settings once you have set the absolute scaling constant. If you do, you will have to recalculate the absolute scaling constant. Also, make sure absolute scale is turned off before you calculate the scale constant, otherwise you will get a bad scaling constant (see the manual for details).



21. Check the "Normalize processed data to absolute scale using glassy carbon" checkbox.
22. Click on the Calibration section in the options list on the left.
23. Change the "Start plots at q-point number" to 13.
24. Click "OK" to exit the advanced options panel, saving the changes.
25. Save the settings for future use.

Comparison note:

We find that for the example data given here, the two methods of glassy carbon calibration agree within ~1.5%. The best approach depends on how strong your background scattering is relative to the rest of the scattering in the system.

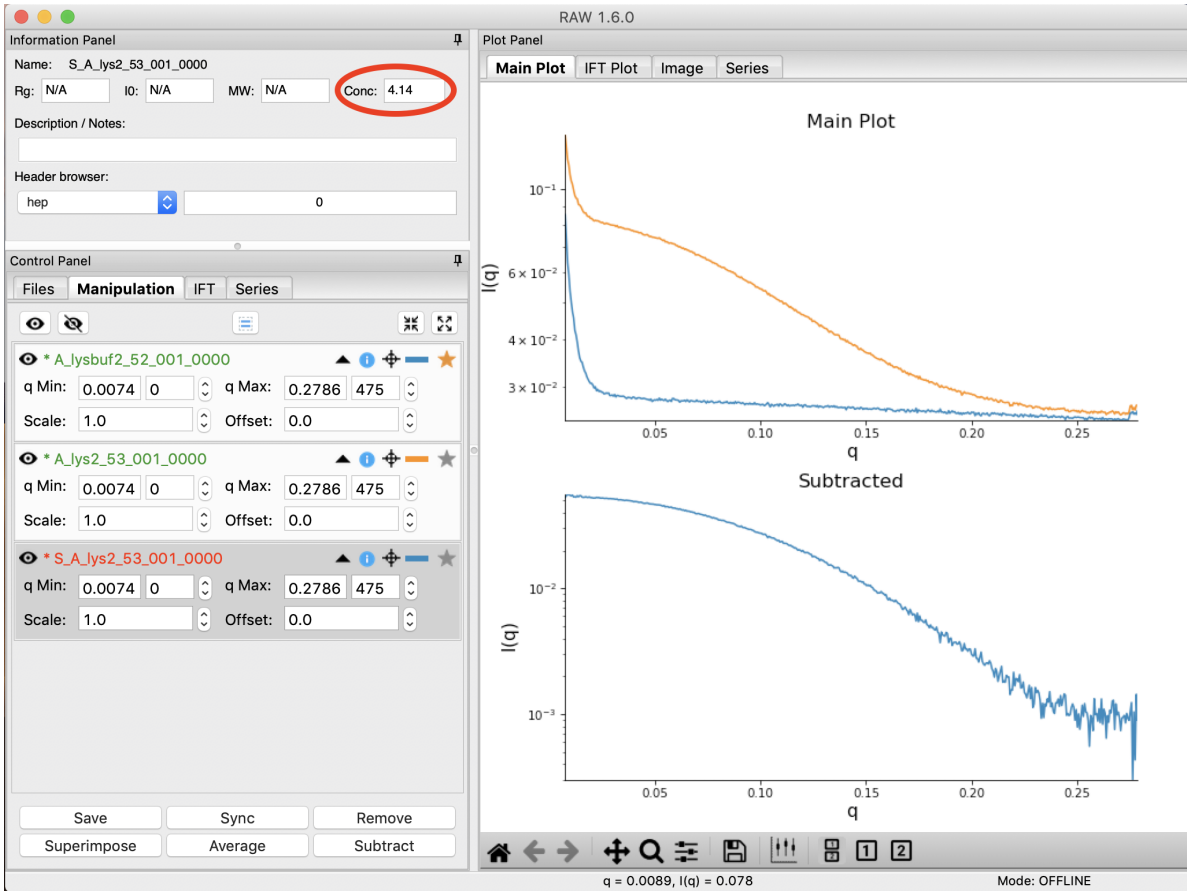
Setting a molecular weight standard

One method for determining molecular weight from a scattering profile is comparison to a known scattering profile with known molecular weight. This part will teach you how to set that known standard in RAW.

1. Load/use the settings from *Parts 4*, *5*, or *6*.
2. Plot all of the **lysbuf2_52_001_000x.tiff** files, where x is 0-9, on the main plot.
   - *Tip: Section 1 Part 1* of this tutorial document teaches you how to do this.
3. Average the **lysbuf2** files you just loaded. Save the average in the **calibration_data** folder.
4. Repeat steps 2-3 for the **lys2_52_001_000x.tiff** files.
5. Subtract the averaged buffer profile (**lysbuf2**) from the averaged sample profile (**lys2**).

- *Tip:* *Section 1 Part 1* of this tutorial document teaches you how to do this.

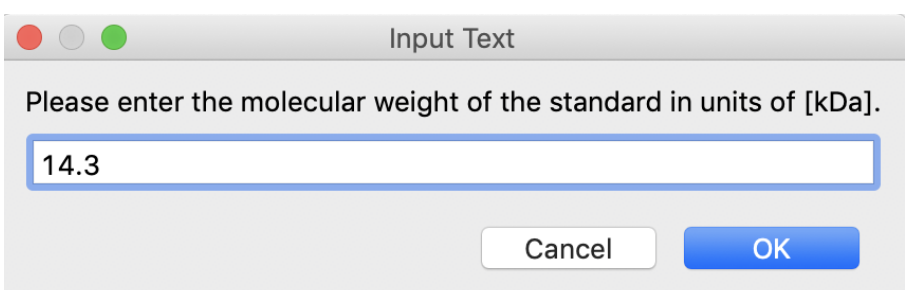
6. Select the subtracted profile by clicking on it. In the information panel, set the concentration in the Conc box to 4.14 (this is concentration in mg/ml).



7. Perform a Guinier fit on the subtracted profile.

- *Tip:* *Section 1 Part 2* of this tutorial document teaches you how to do this.

8. Right click on the subtracted profile and select the "Other Operations->Use as MW Standard" option.
9. Enter the molecular weight of the standard in kDa in the box that appears. For this lysozyme sample, the molecular weight is 14.3 kDa.



10. Click "OK" to save the molecular weight standard.
11. Save the settings for future use.

4.4 Videos

We have made various demo and instructional videos for RAW. These are linked below.

4.4.1 RAW demo for SBGrid

4.4.2 RAW tutorial videos from MacCHESS

For these videos, tutorial data is available from http://bit.ly/bioxtasd

Loading and plotting profiles

Collecting and processing data

Creating a config file

Combining SAXS and WAXS data

Radiation damage

Aggregation

Concentration Effects

4.5 Cite

4.5.1 RAW

If you use RAW in your research, please cite the newest RAW paper:

BioXTAS RAW: improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. J. B. Hopkins, R. E. Gillilan, and S. Skou. Journal of Applied Crystallography (2017). 50, 1545-1553.

DOI: 10.1107/S1600576717011438

You can also cite the previous RAW paper if you like:

BioXTAS RAW, a software program for high-throughput automated small-angle X-ray scattering data reduction and preliminary analysis. S. S. Nielsen, K. Noergaard Toft, D. Snakenborg, M. G. Jeppesen, J. K. Jacobsen, B. Vestergaard, J. P. Kutteraand L. Arleth. Journal of Applied Crystallography (2009). 42, 959-964.

DOI: 10.1107/S0021889809023863

4.5.2 ATSAS

If you use RAW to control any of the ATSAS programs (DAMMIF, DAMMIN, DAMAVER, DAMCLUST, AMBIMETER, SASRES), in addition to the RAW paper, please cite the appropriate paper given on their documentation pages.

4.5.3 Evolving Factor Analysis (EFA)

If you used the EFA function in RAW to deconvolve overlapping chromatography data, in addition to the RAW paper please cite the following paper:

Domain Movements upon Activation of Phenylalanine Hydroxylase Characterized by Crystallography and Chromatography-Coupled Small-Angle X-ray Scattering. S. P. Meisburger, A. B. Taylor, C. A. Khan, S. Zhang, P. F. Fitzpatrick, N. Ando. Journal of the American Chemical Society (2016). 138(20), 6506-6516.

DOI: 10.1021/jacs.6b01563

4.5.4 Electron Density (DENSS)

DENSS

If you used the electron density (DENSS) function in RAW to calculate electron density, in addition to the RAW paper please cite the following paper:

Ab initio electron density determination directly from solution scattering data. T. D. Grant. Nature Methods (2018) 15, 191–193.

DOI: 10.1038/nmeth.4581

4.5.5 Molecular Weight

Volume of Correlation

If you used the volume of correlation method to determine molecular weight in RAW, in addition to the RAW paper please cite the following paper:

Accurate assessment of mass, models and resolution by small-angle scattering. Rambo, R.P. & Tainer, J.A. Nature (2013). 496, 477-481

DOI: 10.1038/nature12070

Corrected Porod Volume

If you used the corrected Porod volume method to determine molecular weight in RAW, in addition to the RAW paper please cite the following paper:

SAXSMoW 2.0: Online calculator of the molecular weight of proteins in dilute solution from experimental SAXS data measured on a relative scale. V. Piiadov, E. Ares de Araujo, M. Oliveira Neto, A. F. Craievich, and I. Polikarpov. Protein Science (2019). 28(2), 454-473.

DOI: 10.1002/pro.3528

4.5.6 Baseline Correction

If you used the integral baseline correction method in RAW, in addition to the RAW paper please cite the following paper:

US-SOMO HPLC-SAXS module: dealing with capillary fouling and extraction of pure component patterns from poorly resolved SEC-SAXS data. E. Brookes, P. Vachette, M. Rocco, and J. Pérez. Journal of Applied Crystallography (2016). 49, 1827-1841.

DOI:10.1107/S1600576716011201

4.6 Manual

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

4.6.1 Introduction to RAW and this documentation

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

RAW

BioXTAS RAW is a program for analysis of Small-Angle X-ray Scattering (SAXS) data. The software enables: creation of 1D scattering profiles from 2D detector images, standard data operations such as averaging and subtraction, analysis of radius of gyration (Rg) and molecular weight, and advanced processing using GNOM, DAMMIF, and AMBIMETER (requires ATSAS installation). It also allows easy processing of inline SEC-SAXS data.

RAW is written in python (mostly) and C++ (a few small bits for speed). It is open source and free for anyone to use. If you do use RAW, we ask that you cite the following paper if you publish or present your results:

BioXTAS RAW, a software program for high-throughput automated small-angle X-ray scattering data reduction and preliminary analysis, J. Appl. Cryst. (2009). 42, 959-964.

Some of the features of RAW include:

- Calibrate, mask, radially integrate, and normalize 2D images to make 1D scattering profiles
- Average, subtract, merge, rebin, and interpolate scattering profiles
- Easily process in-line SEC-SAXS data
- Calculate radius of gyration (Rg) and I(0) via Guinier fit
- Calculate the molecular weight via I(0) comparison to standards, absolute calibration, correlation volume, and corrected Porod volume
- Run GNOM, DAMMIF, and AMBIMETER (requires ATSAS installation)
- Run singular value decomposition (SVD) and evolving factor analysis (EFA) on datasets
- Calculate P(r) functions using a Bayesian indirect Fourier transform (BIFT)
- Can read the following image formats:
  - Pilatus Tiff
  - CBF
  - SAXSLab300
  - ADSC Quantum
  - Bruker
  - Gatan Digital Micrograph
  - EDNA-XML
  - Eiger

- ESRF EDF
- FReLoN
- Nonius KappaCCD
- Fit2D spreadsheet
- FLICAM
- General Electric
- Hamamatsu CCD
- HDF5
- ILL SANS D11
- MarCCD 165
- Mar345
- Medoptics
- MPA (multiwire)
- Numpy 2D Array
- Oxford Diffraction
- Pixi
- Portable aNy Map
- Rigaku SAXS format
- 16 bit TIF
- 32 bit TIF

RAW is currently being developed by Jesse Hopkins and Soren Nielsen.

More information on RAW is available from the RAW website:

https://sourceforge.net/projects/bioxtasraw/

RAW needs your help! If you find bugs in the program, or errors in the documentation or tutorial, please let us know. Your input is the way we get better!

This documentation

The purpose of this documentation is to clearly and completely document the functions of RAW. This document is not intended to act as a tutorial to RAW, for a tutorial please refer to the tutorial document and videos available from the RAW website. Instead, the goal is to document what features are available in each section of RAW, how to use them, and what they do.

This document is laid out in chapters discussing each general area of RAW (for example, the Files tab in the Control Panel is a single chapter). When appropriate, a chapter will cover two related areas, such as the Manipulation tab of the Control Panel and the Main Plot Panel. Some areas will be physical panels or windows of RAW, as the previous examples, while some will be features, such as the Online Mode chapter. The exceptions to this are the second chapter, which covers how to install RAW, and the third chapter, which provides a brief introduction to the different windows of RAW and how they relate.

Typically, the algorithms used will not be fully documented here, as they are available in the source code or appropriate citations (and are subject to change as we improve and expand RAW).

4.6.2 Installing RAW

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

This chapter provides an overview of how to install RAW on Windows, Linux and Macintosh. Detailed installation guides for each system are available separately. It covers:

- Licensing
- System requirements
- Installing from source code.
- Installing from pre-compiled binaries.
- Obtaining the newest version of RAW

Licensing

RAW is open source software, which means the source code is freely available and can be modified/extended as the experienced users sees fit. RAW is available under the GPL V3 license which means that if a modified version of RAW, or any software that uses any of the code available in RAW, is made available to users, then the full source code must also be made available. For a full description of the GPL licensing terms, see the file gpl-3.0.txt included with the RAW source code, or visit http://www.gnu.org/licenses/gpl.html

System requirements

The faster the machine the better the experience.

Minimum Requirements

1.6 GHz 32 bit (x86) Computer

1 GB of memory

1024x768 resolution display

Windows 7 or newer, Mac OSX 10.9 or newer, Linux

Note: The newest version of RAW may run on Windows XP or Vista and OS X 10.4+, but it has not been tested by the developers on these older systems.

Optimal Requirements

2.5+ GHz 32/64 bit (x86) Dual/Quad Core Computer

2+ GB of memory

1280x1024+ resolution display

Windows 7 or newer, Mac OSX 10.9 or newer, Linux

Installing RAW from pre-compiled binaries

A prebuilt installer is available for RAW for Windows and Mac. We recommend that most users install the prebuilt versions. The windows installer has been tested on Windows 7, 8.1, and 10. It may also work for older versions. The Mac installer has been tested on macOS/OS X 10.9 - 10.12. It may also work for older versions.

Instructions for prebuilt installers are available here:

- *Windows*
- *Mac*

Installing RAW from source code

Installing RAW from source code takes more work but has the advantage that the newest version can always be downloaded and quickly installed once the necessary tools have been installed. Installing a pre-compiled version of RAW is easier, but compiled versions are not updated as often as the source code due to the time consuming process of making a compiled executable. Compiled versions may also only be available for certain platforms.

Instructions for install RAW from source are available for

- *Windows*
- *Mac*
- *Linux*

NOTE: As of writing, the RAW is **only** compatible with Python version 2.7 and not the newer 3.x versions.

Obtaining the newest version of RAW

The newest version of RAW can always be found on Sourceforge. The released source code and compiled binaries can be downloaded at:

https://sourceforge.net/projects/bioxtasraw/files/

To obtain the very latest unreleased version (that might or might not be stable) go to: https://sourceforge.net/p/bioxtasraw/code/HEAD/tree/trunk/src/

And click "Download Snapshot" near the top of the screen. Unpack the content into a folder and run RAW.py as usual.

NOTE: When installing new source code on top of old it is important to delete the old source code, including the compiled C libraries with the extensions .so, .pyc, and .pyd, otherwise RAW could be loading old code and become unstable.

Integrating ATSAS with RAW

RAW allows you to do analysis with some of the programs from the ATSAS package directly from RAW. Currently, you can use GNOM, DAMMIF, and AMBIMETER in RAW. This requires a separate ATSAS installation, as the RAW developers are not allowed to distribute the ATSAS package with RAW.

Installing the ATSAS package

The ATSAS package is available from EMBL, and can be downloaded here:

http://www.embl-hamburg.de/biosaxs/download.html

Installation instructions are available here:

http://www.embl-hamburg.de/biosaxs/manuals/install.html

We recommend installing the packages in the default installation location.

To use all of the programs through RAW, you need ATSAS version 2.7.1 or greater. GNOM and DAMMIF may work for earlier versions of the ATSAS package, but the RAW developers have not tested this.

Locating the ATSAS package for RAW

RAW will attempt to automatically locate the ATSAS package when you start up RAW (and when you load a configuration file). It may fail to do this, in which case you will need to set the location of the ATSAS programs manually. To do this:

- Open the "ATSAS" section of the Options window.
- Uncheck the "Automatically find the ATSAS bin location"
- Either by typing the path or using the "Select Directory" button, select the "bin" folder inside the main ATSAS folder. This folder should have a dammif executable inside of it.

Running without compiled extensions

RAW compiles certain extensions that are written in C++ in order to maximize the speed of the program. These extensions are involved in the following tasks: Making polygon masks, integrating images into scattering profiles, and carrying out the BIFT analysis. All of these extensions are also available in native python code, but run much more slowly. If RAW is unable to compile these extensions, a warning message will display when the program is started.

While RAW is able to run without the extensions compiled, it will significantly impact performance of the listed tasks. We recommend troubleshooting the RAW installation, or reinstalling RAW to get these to compile. The RAW installation guides contain detailed install instructions and some solutions to common problems with the installation. Please refer to those for more details.

4.6.3 Overview of RAW

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

The main screen of RAW shows three distinct panels: The Information Panel (top left), the Control Panel (bottom left), and the Plot Panel (right). Additional windows that can be shown are the Options window, and various Analysis windows. In addition, there is a menu bar at the top of RAW (either in the RAW window or in the system menu bar, depending on your operating system).

Any of the three panels docked in the main RAW window can be moved relative to the other panels. To do this, click on the title bar of the panel, and drag the panel to the desired location (you should see a blue rectangle indicating where the panel will be put). Additionally, panels can be undocked from the main window by clicking on the pin icon on the right side of the title bar of the panel. Undocked panels can be re-docked with the main window using the same method as rearranging the panels, drag the title bar until you see a blue rectangle appear indicating where the panel will go.

The portion of the total area available for the left (information and control) and right (plot) sides of RAW is controlled by the separator bar. You can click and drag on this bar to change this ratio.

The main RAW window can be resized by clicking and dragging on a corner or edge.

The Control panel

The control panel is where you manipulate files and items loaded into RAW. It has four tabs: Files, Manipulation, IFT, and SEC. The Files tab is for viewing files on the system disk, and loading the files into RAW. The Manipulation tab is for working with individual scattering profiles, the IFT tab is for working with inverse Fourier transforms, and the SEC tab is for working with SEC-SAXS data. You change tabs just by clicking on them. The order of the tabs in the control panel can be rearranged by clicking and dragging the tabs.

The Plot panel

The plot panel contains four tabs: the Main Plot, IFT Plot, Image plot, and SEC plot tabs. The Main Plot tab is for viewing individual scattering profiles. The IFT Plot tab is for viewing inverse Fourier transforms, the Image plot tab is for viewing detector images, and the SEC plot tab is for viewing SEC-SAXS data. The order of the tabs can be changed by clicking and dragging.

Multiple different plot tabs can be displayed at the same time. Click and drag the tab to any edge of the Plot Panel (up, down, left, or right) and you should see a blue box appear indicating where the tab will go. Additional splits after the first can be accomplished by dragging other tabs, and tabs can be moved between different tab bars (in this case, you should see the bar light up blue, rather than a full blue box for the panel). The amount of screen occupied by the different plots can be set by clicking and dragging the separator bars. To recombine, drag all of the plot tabs into a single bar.

Navigation/control bar

Each Plot tab has a navigation/control bar on the bottom. Most navigation bars have some unique buttons, but they all share the following buttons:

Home. Zoom the to fit all the displayed data.

Move back and fourth in the zoom history.

Pan/Zoom. Click and drag the left mouse button to pan and the right mouse button to zoom.

Area Zoom. Click and drag to zoom to the selected area.

White-space settings. Ability to adjust the white-space around the plots.

Image save. Ability to save the displayed plot(s) in formats such as .png, .svg, .eps, .pdf and more.

Cursor readout

Every plot has a cursor readout. If the mouse is hovering over a position on the plot, the coordinates (x-axis and y-axis values) will read out in the bottom bar of RAW. For the image plot, the intensity at those coordinates is also displayed. For the SEC plot, the value of the secondary y-axis is also displayed.

Plot settings

For each of the Main Plot, IFT Plot, and SEC plot, one of the options in the right-click menu is "Plot Options". Selecting this will open a window that allows you to set:

- Plot title, x-axis label, and y-axis label text, font size, and bold/italicization
- Legend title, font size, and bold/italication
- Whether the legend is visible (default: no)
- Legend font size, transparency, border, and shadow
- Turn auto limits on/off for the Axes and manually set the axes limits if desired
- Turn off the left/right/top/bottom plot borders
- Turn on a line at y = 0 (the Zero Line)
- Adjust the axes tick label font sizes

The axis labels and title will accept some LaTeX commands if they are placed between $ symbols, for example: $\int I(q)$ will display as $\int I(q)$. The bold and italic options will not change LaTeX text.

The settings will be changed for the plot/subplot that was clicked on.

The Information panel

The information panel displays information about a selected data item in the Manipulation panel. It will show, if available, the Rg, I(0), and MW. It provides a place to enter the sample concentration, description/notes about the sample, and quickly look at the different header values of the data item.

If the current Control panel is changed from the Manipulation panel to the IFT or SEC panel, the Information panel is cleared. If the panel is changed back to the Manipulation panel, the data in the Information panel are restored.

The Options window

The options panel is opened by clicking on the Options->Advanced Options menu item. This opens a separate window where many of the RAW settings can be viewed/changed. The panel has two pieces. On the left is the set/category of options, and on the right the actual options panel. For example, if you click on the General Settings section in the left panel, the general settings will be displayed on the right.

Note that in the options tree on the left, the triangles can be used to expand/collapse more options for many of the categories.

Once you have changed the settings, you simply need to click "OK" and they will be saved in memory. In some cases you may want to set the options without exciting the panel. To do this, click "Apply". To exit without saving any changes click "Cancel".

Analysis windows

Various analysis windows can be opened in RAW. These will be discussed in detail *later*.

Menus

The top menu bar of RAW contains the File, Options, View, Tools, and Help menu. The View and Tools menu are simply another way to access options found elsewhere, while the File, Options, and Help menu have items that cannot otherwise be accessed. These will be discussed *later*.

4.6.4 RAW Settings and the Options window

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

This section covers the RAW settings, in particular focusing on the Options window and what all of the settings there do. It also covers how to save and load settings in RAW.

Saving and loading settings

To save the current RAW settings to disk (a .cfg file):

1. Go to the "File" menu bar and select "Save Settings"
2. Select a filename and a location for the configuration (config) file and click "Save"

To load RAW settings from disk (a .cfg file):

1. Go to the "File" menu bar and select "Load Settings"
2. Select the appropriate file and click "Open"

Note: Saving a configuration file saves all of the settings in RAW that are set in the Options window. It also saves any masks that have been created.

Changing settings

Settings in RAW are generally changed in the Options window. Below we describe how to open the options window, how to change all of the settings, and what the settings do.

Opening the Options window

To open the options window:

1. Go to the "Options" menu
2. Select "Advanced Options"

There are two parts to the Options window. There is an options tree on the left that determines which options panel is displayed on the right. The part on the right displays the options associated with the selected option in the option tree.

Closing the Options window

To save the changes made to the settings in the Options window, close the window by clicking the "OK" button. To exit without saving any of the changes in the Options window, use the "Cancel" button or the system close button (an "x" in the upper left or right corner of the panel).

General settings panel

The general settings are all check boxes and can be set by checking and unchecking the boxes. They control the following items:

Hide controls on manipulation items for new plots

If this option is selected, new Manipulation data items start out *collapsed*, if not selected they start out expanded. Defaults to off.

Write header on top of dat files

The *.dat* and *.ift* file "header" information can be saved at the top or bottom of the file. If this is selected, it is written at the top, if not it is written at the bottom. Defaults to off.

Use header for mask creation (SAXSLAB instruments)

If this option is set, if a SAXSLab instrument is being used, the image header will be used to make the beamstop mask. Defaults to off.

Detector is rotated 90 degrees (SAXSLAB instruments)

If this option is set, it indicates to RAW that the SAXSLAB detector is rotated 90 degrees. This affects mask creation from the image header. Defaults to off.

Start online mode on startup

If this option is selected and an *Online mode startup directory* has been picked, then when RAW starts, it will automatically turn on online mode, with the selected directory as the target online directory.

2D Reduction

In the 2D reduction panel, there are reduction options.

Correct of the change in the solid angle of the pixels

If this option is selected, the scattering intensity is corrected for the change in solid angle of pixels as you move along the detector. This is implemented as the standard $\cos^3(2\theta)$ where 2θ is the scattering angle.

Image/Header Format panel

This panel in the Options window allows you to set the image format, the header file format (if any), load the image header into RAW to set up normalizations, and set up bindings to use information in the image header or header file as calibration and reduction parameters.

Image format

The image format can be set using the drop down menu. A full description of supported image types is available in the *file types section*.

Header file format

The header file is a file that accompanies the detector image on some beamlines. This file often contains additional information such as, diode values, ion chamber readings, exposure time and date. Currently RAW supports the following header files: I711, MAXLab; I911-4, MAXLab; F2, CHESS; G1, CHESS; G1 WAXS, CHESS; G1 Eiger, CHESS; BL19U2, SSRF; and BioCAT, APS.

Note: If you wish to have a new header file format added to RAW, please contact the developers.

Loading/Viewing header information

If you wish to view header information from either the header file or the image header, click the "Load Image" button and select the image of interest.

If you wish to use the header information to normalize the image, load the image using the "Load Image" button and then click the "Apply" button at the bottom of the screen. This will save the counter values in such a way that RAW can set up the normalization appropriately.

Using image header information for calibration and reduction (turning on and setting bindings)

RAW has the ability to use header information for calibration and reduction settings. The method for doing this is to set a "binding" between the counter value and the calibration value. The calibration values that can be obtained from the image header or header file are: Beam X Center, Beam Y Center, Detector Pixel Size, Sample Detector Distance, and Wavelength.

To create a binding:

1. Check the "Use image-header/header file for calibration and reduction parameters" box.
2. Load the image and header file values into the list as described *above*.
3. In the list of the image header and header file names and values, click on the name of the header parameter that you want to use as one of the calibration values. This will fill in the Name and Value in the appropriate fields in the lower left hand portion of the panel.
4. Using the "Binding" menu, select what calibration parameter should use this header value. In the binding column of the header list, you will see this calibration parameter displayed.

Note: These values overwrite the same values set elsewhere in the settings. So if you bind the Beam X Center to use a value from the header, no matter what you set it to in the Calibration panel of the Options window it will use the value from the header.

Note 2: Make sure that your header file values match the expected units for the calibration parameter. The beam center values should be in pixels on the detector, the detector pixel size should be in microns, the sample detector distance in mm, and the wavelength in angstroms.

Adding a modifier to a binding

Once a binding is set, it is possible to add a modifier to the binding, which affects the value obtained from the header. This might be used in a case where the header value is not in the appropriate units.

To set a modifier:

1. In the list of the image header and header file names and values, click on the name of the bound header parameter that you want to set a modifier for.
2. In the Modifier field at the bottom of the panel, type in a mathematical expression. This expression may contain any of the header values (including but not limited to the header value selected for the binding). It may contain "+" "-" "*" and "/" for addition, subtraction, multiplication, and division. The following strings are restricted, and apply specific mathematical functions: *acos, asin, atan, atan2, ceil, cos, cosh, degrees, exp, fabs, floor, fmod, frexp, hypot, ldexp, log, log10, modf, pow, radians, sin, sinh, sqrt, tan, tanh*, all of which correspond to the functions of the same name in the python math library (https://docs.python.org/2/library/math.html#module-math).

3. Click the "Add" button. You should get a popup window that evaluates the expression for the current loaded header values. Once you close that window, the modifier should be listed in the Modifer column of the header list.

Changing or removing a modifier to a binding

To change a modifier to a binding, do the steps to add a modifier, *above*. When you click on the header item in step 1, the modifier will be shown in the Modifier field at the bottom of the panel, and you can make changes as appropriate in step 2.

To remove a modifier to a binding:

1. In the list of the image header and header file names and values, click on the name of the bound header parameter that you want to remove a modifier to.
2. Click the "Remove" button (next to the Modifier field at the bottom of the panel).

Removing bindings

To remove a single binding:

1. In the list of the image header and header file names and values, click on the name of the bound header parameter that you want to remove a binding to.
2. In the "Binding" menu at the bottom of the panel, select "No binding".

To remove all bindings, click the "Clear Bindings" button.

Disabling bindings for calibration and reduction

To disable the use of bindings for calibration and reduction, either *remove* all bindings or uncheck the "Use image-header/header file for calibration and reduction parameters" checkbox.

Clear All

The "Clear All" button clears all bindings, and removes the current loaded header/header file values from the panel.

Calibration panel

The calibration panel allows you to set the beam center, binning size, number of points skipped at the start and end of a scattering profile, the sample to detector distance, wavelength, detector pixel size, and whether or not the Q range is calibrated.

Setting calibration parameters

The calibration paramters are: Beam center (x and y), sample-detector distance, wavelength, and detector pixel size. These can all be set by entering a value in the appropriate field on this panel or using the spin controls. However, it is more natural to set these values from the *Calibration/Centering panel*.

Note: Changing the settings in the calibration/centering panel will change the values in this panel, and vice versa.

Note 2: All of these calibration values are overridden by the bindings described *above*, if a binding for the particular calibration parameter is set.

Start and end points

The "Start plots at q point number" value sets the first q point shown on the plot when a scattering profile is integrated. It is zero indexed (first point is zero). So if it is set to 5, the plot will start with the 6th q point in the q-vector. This is typically used to get rid of the beamstop shadow from the integrated profiles.

The "Skip n points at the end of the curve" value sets the last point shown on the plot when a scattering profile is integrated. If it is set to zero, all points are shown. So if it is set to 5, the last point shown will be the 5th to last point in the q-vector. This is typically used to remove end points if something, for example the downstream flight tube window, is shadowing a high q region of the detector.

Binning

The default binning for integrated scattering profiles can be set using the "Binning Size" option. It accepts integer values. A binning size of one corresponds to q bins that are one pixel wide. A binning size of 2 corresponds to q bins that are 2 pixels wide, and so on.

Note: The q size of a bin of a given pixel size will depend on the calibration parameters.

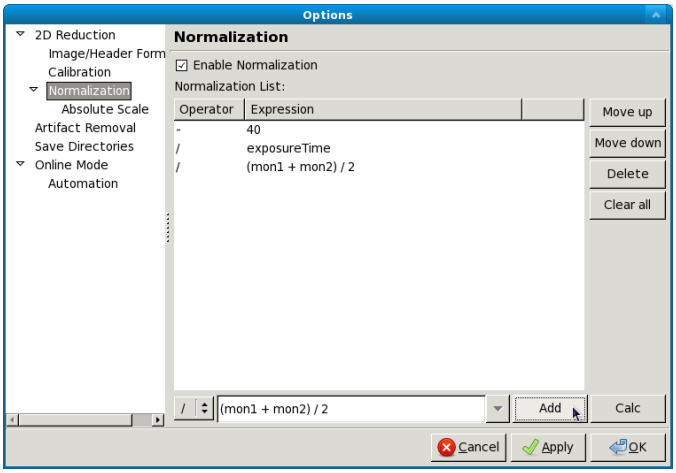
Calibrating the q-range

If you do not wish to calibrate the q-range of integrated scattering profiles, uncheck the "Calibrate Q-range" box. The scattering profile will then be displayed as intensity vs. bin number. This option is checked by default.

Normalization panel

The normalization panel allows you to normalize integrated scattering profiles by some value. Typically a counter value is used that is proportional to the beam intensity transmitted through the sample (such as a beamstop counter from an active beamstop).

Enabling and disabling normalization



To enable normalization for integrated scattering profiles, check the "Enable Normalization" checkbox (checked by default). To disable, uncheck the "Enable Normalization" checkbox.

Setting up normalization operations

To add a new operation to the normalization procedure:

1. Select the operator to be used (/, *, -, + corresponding to division, multiplication, subtraction, and addition respectively)
2. Enter the desired expression in the expression box.
3. Click the "Calc" button to view the result of the entered expression.
4. Click the "Add" button.

Note: This expression may contain any of the header values (including but not limited to the header value selected for the binding). It may contain "+" "-" "*" and "/" for addition, subtraction, multiplication, and division. The following strings are restricted, and apply specific mathematical functions: *acos, asin, atan, atan2, ceil, cos, cosh, degrees, exp, fabs, floor, fmod, frexp, hypot, ldexp, log, log10, modf, pow, radians, sin, sinh, sqrt, tan, tanh*, all of which correspond to the functions of the same name in the python math library (https://docs.python.org/2/library/math.html#module-math).

Reordering and removing normalization operations

The order in which the operations are carried out can be changed by selecting the operation in the normalization list and using the Move Up and Move Down buttons. Operations can be removed by selecting the operation in the list and clicking the Delete button.

Normalizing by a header value

It is often desired to normalize the data by exposure time or incoming / transmitted beam intensity, and/or remove offsets on the detector.

To do so:

1. Load a header file into RAW as *described*. Be sure to hit the "Apply" button after loading!
2. Return to the Normalization panel.
3. Add a normalization value as in steps, in the expression box enter the name of the header value you wish to normalize by along with any other mathematical operations.

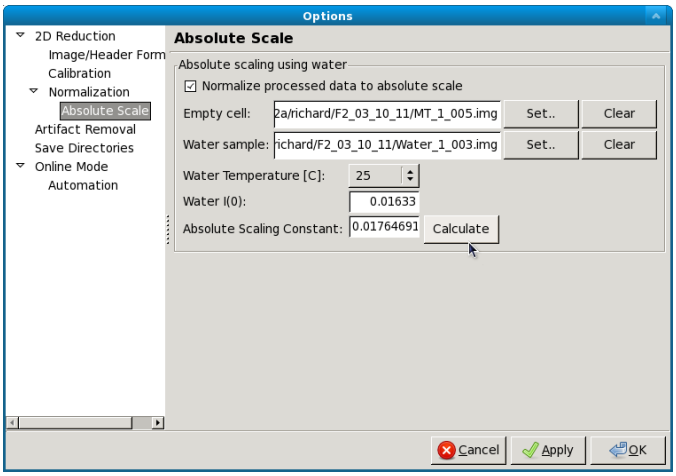
Normalizing by a region of interest (ROI)

RAW has the ability to normalize by a region of interest on the image. Every pixel in the region of interest is summed, and that can be used to normalize in the same way as a header value.

To normalize by an ROI:

1. Set an *ROI mask*.
2. Add an operation to the *normalization list*, but use "roi_counter" (without quotes) as the header value. For example, to divide by the roi value, select the "/" operator and enter roi_counter in the expression box, then add that to the list.

Absolute scale panel



RAW is able to scale loaded image data to absolute scale using water as a standard. Water has a known, temperature dependent absolute scale value at the forward scattering I(0). Water has a relatively flat scattering profile, which makes it possible to estimate the forward scattering, I(0), from an average of the intensity. To obtain the pure water signal, the water sample obtained in a sample cell must have the empty cell subtracted from it.

To set up Absolute scale:

1. Click the "Set" button for the empty cell. Select either an image or text (such as .dat) file of the empty cell scattering.
2. Click the "Set" button for the water sample. Select either an image or text (such as .dat) file of the water scattering.
3. Select the water temperature in degrees centigrade.
4. Click the "Calculate" button. An absolute scaling constant should appear.
5. Enable absolute scale normalization by checking "Normalize processed data to absolute scale" check box.

The algorithm uses the middle third part of the water scattering curve to estimate I(0) by the average intensity.

Note: The selected files must have been normalized in exactly the same way as the rest of the data that is to go on absolute scale. If loading an image, that means not changing the normalization parameters after calculating the absolute scale. If normalization parameters are changed the absolute scale constant will have to be re-calculated. It is particularly important that the images or profiles used to calculate absolute scale not have been saved with absolute scale already on (for example, from a previous calibration).

Turning off absolute scale

To turn off absolute scale, uncheck the "Normalize processed data to absolute scale" checkbox in the Absolute scale panel.

Flatfield correction panel

If a flatfield file is available, RAW can do a flatfield correction of the data. To do so, click the "Set" button, and select the flatfield image. Then check the "Enable flatfield correction" box.

When RAW applies a flatfield correction, it divides every image it processes by the flatfield image, on a per-pixel basis. The assumption is that every pixel in the flatfield image should have gotten the same intensity, so any variation comes from variation in the detector.

Turning off flatfield correction

To turn off flatfield correction, uncheck the "Enable flatfield correction" checkbox in the Flatfield correction panel.

Molecular weight panel

The molecular weight panel of the Options windows allows control of the parameters used to calculate molecular weight in the *molecular weight window* and the *SEC calculated parameters*. All four methods are described in more detail *elsewhere*.

Molecular Weight Estimation Using a Standard

This subpanel corresponds to parameters for the MW estimation by comparison to a known standard. While all of the parameters of the standard can be set/changed in this box, the standard MW (in kDa), the standard I(0), the standard concentration (in mg/ml), and the standard filename (only for reference), it is more natural to change these settings by loading the standard scattering profile into RAW and using the *Use as MW Standard* option.

Molecular Weight Estimation From Volume of Correlation

This subpanel corresponds to the parameters used for the volume of correlation method of estimating molecular weight. This method is the method used for calculating MW in the *SEC panel*). The protein and RNA coefficients correspond to the *A and B coefficients*. The default type selection selects if the MW calculation defaults to Protein or RNA. Changing this option will change whether the MW calculated in the SEC panel is for protein or RNA.

Molecular Weight Estimation From Corrected Porod Volume

This subpanel corresponds to the parameters for the MW calculation by corrected Porod volume. For this method, the only parmater that can be changed is the protein density in kDa/Å:sup:*3*.

Molecular Weight Estimation From Absolute Intensity Calibration

These parameters correspond to the parameters necessary for calculating the molecular weight when a scattering profile is on an absolute scale.

Reset MW Parameters To Defaults

If you have customized the MW parameters for a particular sample, you can restore the parameters to the RAW defaults (which are the defaults from the relevant papers for each method). There are no default settings for the estimation using a standard.

Artifact removal panel

Zingers are pixel values on the detector that are unusually high due to either cosmic radiation or readout errors. RAW includes three methods that can be used for zinger removal.

Zinger removal by smoothing

A window of "Window Length" data points can be run across the data and discard values that are more than "Std" standard deviations away from the average of the points in the window. A starting index is given to specify where on the data curve the window should start.

Zinger removal when averaging

If three or more exposures of the same sample are available, then these can be used to eliminate zingers by comparing the intensity values of each data set to the others. An intensity value in a data-set that is larger than x standard deviations (Sensitivity) from third quintile of all related data points in the rest of the data sets is removed and replaced by the average of the third quintile.

Zinger removal after radial averaging

This method is the most effective method for removing zingers. Pixel intensities in the image for the same q value are compared and should be fairly constant. Values that are more than "Sensitivity" standard deviations away from the median are discarded.

IFT panel

RAW currently supports one built-in method for determining the inverse Fourier transform (IFT) of a scattering profile, the Bayesian IFT (BIFT) method. In the future we anticipate supporting a python based implementation of the GNOM algorithm called pyGNOM, but currently that is not available.

BIFT

The BIFT panel allows you to set the BIFT Grid-Search parameters. These define the large grid that the BIFT algorithm searches over before doing the fine search near the best value on the grid.

Dmax Upper Bound

Sets the largest maximum dimension (Dmax) value that will be used in the coarse grid search, in Å.

Dmax Lower Bound

Sets the smallest Dmax value that will be used in the coarse grid search, in Å.

Dmax Search Points

Total number of Dmax values in the coarse grid search that. These are evenly distributed between the lower and upper bounds.

Alpha Upper Bound

Sets the largest alpha value that will be used in the coarse grid search.

Alpha Lower Bound

Sets the smallest alpha value that will be used in the coarse grid search.

Alpha search points

Sets the total number of alpha values in the coarse grid search. These are distributed logarithmically between the lower and upper bound.

P(r) Points

Sets the number of points in the calculated P(r) curve.

Save Directories panel

This panel controls the settings for *automated saving of data*.

Auto Save

In this subpanel, the checkboxes control whether or not RAW automatically saves Processed image files, Averaged data files, and Subtracted data files. When the boxes are checked, that file type will be automatically saved, when they are unchecked, it will not.

Save Directories

This panel allows you to selected the directories into which files will be saved for each of the automated saving file types (Processed, Averaged, Subtracted). To pick a directory, click the "Set" button and use the window that opens to select a folder. Click "Open" once the appropriate folder is selected. To clear a directory click the " Clear" button.

Note: A save directory must be selected before an Auto Save checkbox can be enabled.

Online Mode panel

This panel includes settings for the *online mode*. This lets you enable online filtering and set up the filter list.

Enabling/disabling online filtering

Online filtering filters files by filename, so that you can control which files are loaded into RAW automatically. To enable this mode, check the "Enable Online Filtering" checkbox. To disable this mode, uncheck that checkbox.

Adding a filter item to the online filter list

A filter item consists of three parts. First, there is the Ignore/Open operator. This allows you specify whether you want RAW to ignore files with the given filter string in their name, or to only open files that have the given filter string in their name. To set this option, use the dropdown selector box at the bottom of the panel and select either "Ignore" or "Open only with".

The next part of the filter is the filter string. This is the string that RAW looks for in the filename. To set this, enter a string into the filter string box at the bottom of the panel.

The final part of the filter is the location of the filter string. This sets where in the filename RAW looks for the given filter string. This can be set to: "At start", which means RAW only applies the filter Ignore/Open action to files with the filter string at the start of the file name; "Anywhere", which means RAW applies the filter Ignore/Open action to files with the filter string anywhere in the file name; and "At end", which means RAW only applies the filter Ignore/Open action to files with the filter string at the end of the file name. To set this, use the dropdown selector box at the bottom of the panel and select one of those three options.

Once you have set the Ignore/Open option, entered a filter string, and selected the location of the filter string, click the "Add" button to add the filter item to the Online Filter List.

Reordering and removing filter items

The order in which the filtering is carried out can be changed by selecting the item in the filter list and using the Move Up and Move Down buttons. Items can be removed by selecting the operation in the list and clicking the Delete button. All filter items can be removed using the "Clear all" button.

SEC-SAXS panel

The SEC-SAXS panel controls settings related to the SEC-SAXS data processing.

Intensity ratio (to background) threshold for calculationg Rg, MW, I0

In order to speed up the *calculation of Rg, MW, and I0 as a function of frame* for SEC-SAXS data, a ratio of the frame intensity to the background intensity is taken. If that value is less than the threshold set here, the frame is skipped. To attempt to calculate structural parameters for all frames, set this threshold to -1.

ATSAS panel

The top level ATSAS panel allows you to control where the ATSAS bin location is (the folder with all of the ATSAS programs in it). By default, RAW will attempt to automatically find the ATSAS installation. If you wish to set the location yourself, uncheck the "Automatically find the ATSAS bin location" checkbox, and either type the location into the ATSAS bin location field or use the "Select Directory" button to select the appropriate directory.

Note: If you uncheck and then check the Automatically find checkbox, RAW will attempt to find the ATSAS directory again. This can be useful if, for example, you install ATSAS and want RAW to find the new installation without restarting RAW.

GNOM panel

The top level GNOM panel allows you to set the commonly used advanced settings for the ATSAS software GNOM, which is run from the *GNOM window*. All of these options correspond in name and allowable values to those of GNOM as described in the GNOM manual: http://www.embl-hamburg.de/biosaxs/manuals/gnom.html

Settings can be rest to their defaults (which correspond to the GNOM defaults) by clicking the "Reset to default" button. This resets the settings in this panel and in the GNOM Advanced panel.

GNOM Advanced panel

This GNOM panel allows setting GNOM settings which are not as commonly used in GNOM. Again, all of the options correspond in name and allowable values to those of GNOM as described in the GNOM manual: http://www.embl-hamburg.de/biosaxs/manuals/gnom.html

Settings can be rest to their defaults (which correspond to the GNOM defaults) by clicking the "Reset to default" button in the GNOM panel. This resets the settings in this panel and in the GNOM panel.

DAMMIF panel

The top level DAMMIF panel allows setting two things: First, the default settings for DAMMIF that are set when the *DAMMIF window* is opened. Second, standard settings that are available in Fast and Slow mode can be set in the "Standard Settings" subpanel. All of the settings correspond in name and allowable values to those in the DAMMIF manual: http://www.embl-hamburg.de/biosaxs/manuals/dammif.html

Settings can be rest to their defaults (which generally correspond to the DAMMIF defaults) by clicking the "Reset to default" button. This resets the settings in this panel and in the DAMMIF Advanced panel.

DAMMIF Advanced panel

The settings in the DAMMIF advanced panel are only used when the "Custom" mode is selected in the DAMMIF panel. This is equivalent to the interactive DAMMIF mode at the command line. Unless otherwise noted, a value of -1 for a field indicates that it will use the default setting. The settings correspond in name and allowable values to those in the DAMMIF manual: http://www.embl-hamburg.de/biosaxs/manuals/dammif.html

Settings can be rest to their defaults (which generally correspond to the DAMMIF defaults) by clicking the "Reset to default" button in the DAMMIF panel. This resets the settings in this panel and in the DAMMIF panel.

4.6.5 The Centering/Calibration panel

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

The Centering/Calibration panel allows you to set the beam center, x-ray wavelength or energy, sample to detector distance, and detector pixel size. This allows RAW to calibrate the q position for each pixel, and integrate the image into a one dimensional scattering profile.

Opening the Centering/Calibration panel

Before opening the Centering/Calibration panel, you should make sure that an appropriate image for calibration (such as the scattering from silver behenate) is *open* in the *Image panel*. To open the Centering/Calibration panel, go to the Tools menu and click on "Centering/Calibration". The Centering/Calibration panel will be placed (temporarily) over the control panel.

Setting calibration parameters

The Wavelength, Energy, Sample-Detector Distance, and Detector Pixel Size can all be set in the Manual Centering/Calibration Adjustments panel of the Centering/Calibration panel. For each, they can be set in one of two ways. First, a value can be directly typed into the appropriate field. Second, the spin controls can be used to adjust the value in the field. The spin controls only change the last digit of the value, so if the sample detector distance is set at 1000, the up spin control will change it to 1001, whereas if it is set at 1000.0, the up spin control will change it to 1000.1.

If a centering pattern is displayed on the image, changing these calibration will change the pattern. In order for the pattern to update when a value is typed into a field, you must either hit the enter key, or click out of the field.

Note: Setting the wavelength will automatically set the appropriate energy, and vice versa. You only need to set one of these values.

Manually setting the beam center

The beam center is displayed in x and y coordinates in the Manual Center/Calibration Adjustments panel. These coordinates correspond with the x and y coordinates shown on the image. If a centering pattern is selected in the Pattern drop down menu, the eam center is also displayed on the image as a red dot 6 pixels in diameter.

The location of the beam center can be changed in three ways. First, values can be directly entered into the X center and Y center fields. As with the calibration parameters, in order to update the pattern display after entering a value in these fields, you must either click out of the field or hit the enter key.

Second, the step size of the beam center in pixels can be set using the drop-down "Steps" menu. The Steps menu is also a field you can type into, so you can set it to any numerical value if the options available in the list are insufficient. Once you have set this, the large red centering arrows on the right side of the Manual Center/Calibration Adjustments box can be used to move the beam center <steps> pixels in the direction indicated by the arrow, where <steps> is the

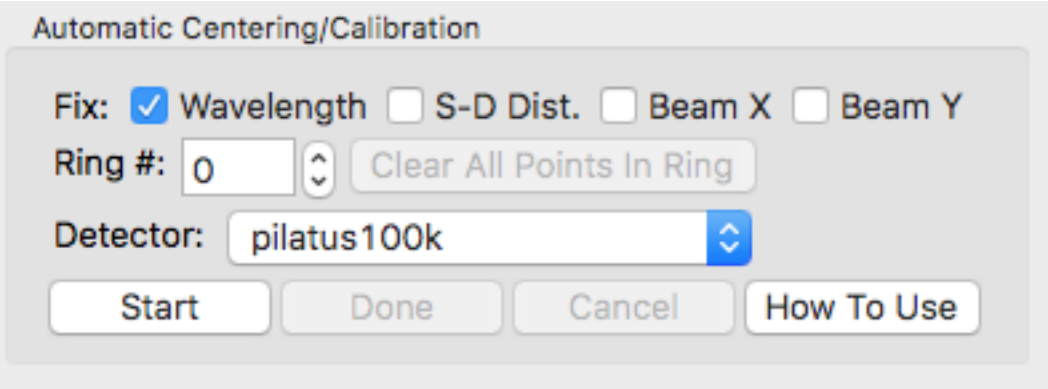
value of the "Steps" field/menu. The arrows can be held down for repeated motion in the same direction with the same step size.

Third, the target button in the center of the centering arrows can be clicked. You then click on the image, and the beam center will be set to that position on the image, to the nearest pixel.

Automatic centering and distance calibration

There are two different automated centering methods available in RAW. Which one is available in your version depends on whether or not you have the *pyFAI* library installed. With either of these methods, you will have the best chance to obtain accurate results if you manually calibrate the center and distance first, so that the automated centering can refine on that starting position.

With pyFAI (recommended)



In the Manual Centering/Calibration Adjustments panel:

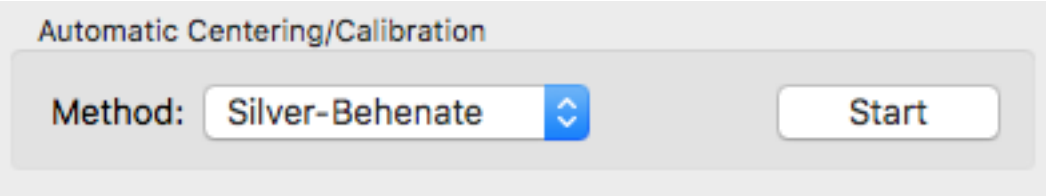
1. Select the appropriate Standard.

In the Automatic Centering/Calibration panel:

1. Select the parameters to hold constant by checking boxes in the Fix section.
2. Select the detector type to be used. If your detector is not in the list, select Other.
3. Set the Ring # to the index of the first ring visible on the detector. The ring index starts at zero for the largest d-spacing ring (nearest the beam) and increments by one for each ring thereafter. *IMPORTANT:* The first ring visible on your detector image may not be ring 0!
4. Click the Start button. Then click on the first ring in the image. Points in that ring will be automatically selected. Click on other parts of the ring as necessary to fill in points.
5. Increment the ring number as appropriate for the next ring visible on the image (usually increment by 1, for example from 0 to 1), and click on the next ring on the image to select points there. Repeat for all visible rings.
6. (If necessary) To remove points in a ring, set the Ring # to that ring, and click the Clear All Points In Ring button.
7. Click the Done button once you have selected points in all of the visible standard rings. At this point, automatic centering and calibration will be carried out.

Note: Ring points cannot be selected if the Pan or Zoom tool are selected.

Without pyFAI (not recommended)



This automated centering function attempts to position the center and find the sample-detector distance based on the innermost ring of a Silver Behenate sample:

1. Click the start button in the "Centering/Calibration Panel"
2. Select at least 3 points just outside the innermost ring in the Silver Behenate image.
3. Hit the "Done" button.

This function attempts to find the center of the circle that all of the points defined by your clicks make. If that is found, it then uses the known wavelength, pixel size, and q value of the first silver behenate ring to calculate the sample-detector distance.

Centering patterns

With pyFAI (recommended)

If the pyFAI library is installed, more than twenty different calibrants are available, all of those supported by the pyFAI package (https://github.com/silx-kit/pyFAI/tree/master/pyFAI/resources/calibration), including Silver-Behenate (AgBh). Rings are displayed as red dashed lines at all appropriate q values.

Without pyFAI (not recommended)

Only one centering pattern is available in RAW. This is the Silver-Behenate pattern, which displays red dashed rings on the detector corresponding to the expected rings from Silver-Behenate. RAW displays the first 5 rings, which at q = 0.1076, 0.2152, 0.3229, 0.4305, and 0.538 Å:sup:-*1*. The q values are labeled in RAW, but only show at the bottom of the rings, and so you typically have to zoom out of the image to show the label.

4.6.6 The Masking panel

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

The masking panel allows you to create various types of masks that will be used when the image is integrated into a one dimensional scattering profiles.

Types of masks

Beamstop mask

Unwanted regions such as the beamstop, bad pixels, or gaps between detector panels need to be excluded from the detector image before processing the image. Creating a beamstop mask allows you to do so. For a beamstop mask, when a normal mask is drawn (shown as red on the image), the area inside the mask is excluded from integration. When an inverted mask is used (green on the image), only the area within the mask is included during integration. These types of masks can be combined: normal masked regions will be excluded from integration, even if some or all of the same region is included in the inverse mask region. This allows you to create an inverse mask, and mask out undesired features within that inverse mask.

ROI counter mask

RAW allows you to define a region of interest (ROI) mask. If you have an ROI mask defined, every time an image is integrated to a scattering profile, all of the pixel intensities in the ROI mask will be summed, and the value will be available as a counter. In this case, the normal mask (red color) is an include only mask, while the inverted mask (green color) is an exclusion mask. Thus, any areas with a normal mask will be counted in the ROI, while any areas with an inverted mask will be ignored by the ROI.

Typically this is used for normalization purposes, for example to normalize by the transmitted beam through a semi-transparent beamstop. The counter name that this value goes to is *roi_counter*.

SAXSLAB Beamstop mask

The SAXSLAB home source instruments embed the beamstop mask within the image header. Selecting this beamstop mask option will allow you to show this mask on the image. This mask cannot be changed within RAW.

Note: Both a beamstop mask and a SAXSLAB beamstop mask can be simultaneously defined in RAW. In that case, they are both used to mask the image.

Readout-dark mask

The area that is selected by a readout-dark mask is averaged for each image, and that average value is subtracted off of every pixel in the image. The standard deviation in the pixel counts is added in quadrature with the noise in the scattering profile calculated during integration (the standard deviation of the counts in each q bin). In this case, the normal mask (red color) is an include only mask, while the inverted mask (green color) is an exclusion mask. Thus, any areas with a normal mask will be averaged to get the readout-dark counts, while any areas with an inverted mask will be ignored by the average.

Setting a readout-dark mask is intended to allow RAW to compensate for dark counts on the detector (most common with CCD detectors). Typically it is used when a fixed area of the detector is permanently masked, such as by lead tape applied to the detector face.

Creating a mask

In order to creating a mask, an appropriate image must first be l:ref:*open <showimage>* in the *Image panel*. Mask creation is then done as follows:

1. Select the appropriate mask type from the drop down masking menu in the Mask Creation panel.
2. Use the mask drawing tools to draw a mask on the image.
3. For the circle and rectangle tools, left click once to start drawing, move the mouse until you have the desired shape, and click again to stop drawing.
4. For a polygon, left click once to start drawing, and continue left clicking to add vertices. To finish drawing and connect the final point to the initial point, right click.
5. To invert a mask section, right click on the section and select "Inverted mask".
6. To remove a mask section, left click on the section to select it and click the delete key.
7. You may draw as many mask sections as you want.
8. Once done, click the "Set" button to save the mask.

Note: Clicking "Set" overwrites the existing mask in memory. To add new regions to an existing mask, see below.

Viewing and modifying a mask

At times you may wish to view an existing mask, and possibly modify it. To do so:

1. Select the mask type in the Mask Creation panel drop down menu.
2. Click "Show" to show the existing mask.
3. Draw new masked regions and/or remove old masked regions as above.
4. Once you have finished modifying the mask, click "Set" to save your changes.

Removing a mask

To completely remove a mask, select the appropriate type of mask in the drop down menu in the Mask Creation panel, and click the "Remove" button.

Saving/Loading a mask to/from disk

A mask may be saved to disk by itself (saving the settings saves the mask as well), using the "Save" button in the mask drawing panel, and loaded into RAW using the "Load" button. Once a mask is loaded, select the appropriate mask type from the drop down menu in the Mask Creation panel and click "Set" to set the loaded mask as that mask type.

Note: The "Save" button does not save the mask in program memory! You must use the "Set" button to set a mask as the current mask for use. Save and load are exclusively for saving/loading to/from disk.

Clearing a mask

The "Clear" button clears any mask from the image. However, to save these changes, you must click "Set" for the mask type of interest. To remove a mask from use in RAW, see *above*.

Mask drawing options

The check box "Show Beam Center" puts a red circle of diameter 6 pixels on the image where the beam center is (as set in the Centering/Calibration panel).

4.6.7 The Files tab in the Control Panel

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

Changing file directories

There are three ways to change the directory whose contents are displayed in the Files tab:

- Click on the folder icon next to the bar to open a directory window. This allows you to select the directory you want to display.
- Type a path into the bar at the top of the Files tab and hit enter to go to a directory.
- Double click the up arrow in the File list (blue arrow with a filename of '..') to navigate up a directory level. Double click on any displayed directory to open it and show the contents in the Files tab.

Note: The Files tab simply displays the files in your system. None of these files are loaded into RAW until you take some action on them (for example, *plotting them*)

Updating the file list

The file list in the Files tab does not automatically update when new files are placed in the folder. In order to update the file list to show new files, you must click the Refresh button (two green arrows in a circle) to the right of the Folder button near the top right of the Files tab.

Plotting files

Plotting is done from the Files tab in the Control Panel. It can be done in two ways:

1. Simply double-click an image, compatible ASCII file, or RAW sec file to plot it.
2. Select one or more files to be plotted at the same time and click the "Plot" button

The Calibration, Normalization, Masking, and other Advanced Options will be used to convert images into one dimensional scattering profiles when an image is selected and plotted. The Normalization settings, including absolute scale factor, counter norms, and solid angle correction status, the configuration file used, and the image load path are saved as part of the manipulation history for a scattering profile in RAW created from an image.

Note 1: Selecting files is as simple as left clicking on them. Multiple files can be selected by shift or ctrl (command on OS X) left clicking on the files.

Note 2: Some image files can only be loaded if the image type in the Advanced Options is set appropriately.

Filtering and searching the file list

Below the file list is a filter/search bar. You can either select filters from a list or create your own using wildcard characters. The filter bar accepts '*' as a wildcard matching any number of characters and '?' as a wildcard matching a single character.

Examples:

- If you were looking for all files with 'BSA' anywhere in the filename you would write: *BSA* in the filter/search bar. The two * wildcard matches any characters before and after the 'BSA'.
- If you wanted all files with 'BSA' at the start of the filename you would write BSA* in the filter/search bar. The single * wildcard will match any characters after the 'BSA'.
- If you wanted all of the files with 'BSA' at the start of the filename and '.tiff' at the end of the filename you would write **BSA*.tiff**. The single * wildcard will match any characters between BSA and .tiff in the filename.
- If you wanted all of the files with 'BSA_000x.tiff' where x is a single character, you would write **BSA_000?.tiff**. The ? wildcard will match any single alphanumeric character in that position.

Sorting the file list

The columns at the top of the file list, Name, Ext, Modified and Size, show the file name, the file extension, the file modified time, and the file size respectively. By clicking on the column heading, you can sort the file list by those attributes. By default, the file list is sorted by name, ascending (low to high, A to Z). This is indicated by the upward pointing green arrow in the Name column heading. Clicking on the Name column heading will switch the order of the sort to descending (high to low, Z to A) sort. Clicking again will return it to the ascending sort. Clicking on another column heading will sort the file list by that column, and again can be switched between ascending and descending.

Files tab buttons

Quick reduce

Processes selected images and save them to ASCII files without plotting the scattering profiles using the integration parameters (calibration, masking, etc) current set in RAW.

System Viewer

Attempts to open/run the selected file using whatever software is set up in the operating system to handle the file format. You can use the 'E' keyboard key as a shortcut.

Plot

This plots the selected files in RAW, assuming the formats can be read by RAW. If the files are images, they will be integrated into one dimensional scattering profiles using the integration parameters (calibration, masking, etc) set in RAW. Files are shown in the following locations:

- Images (which RAW integrates) and 1D scattering profiles (typically ".dat" files) opened by RAW display in the Main Plot tab of the Plot panel. The Manipulation panel shows the items in the Main Plot and lets change and process the curves individually.
- IFT files (.out and .ift extensions) are loaded into the IFT Plot Panel of RAW. The IFT panel in the Control Panel shows the items in the IFT Plot and lets you change and process the curves individually.
- SEC files (.sec extensions) are loaded into the SEC Plot Panel of RAW. The SEC panel in the Control Panel shows the items in the SEC plot and lets you change and process the curves individually.

Note: If a file has previously been saved from RAW, all of the saved information (see later sections) will be loaded when the file is reloaded. For example, if a subtracted scattering profile had a Guinier analysis performed on it and was saved as a ".dat" file, when that saved file is loaded back into RAW, the Rg and I(0) values will be loaded as well.

Show image

Shows the image selected in the file list in RAW's image plot. You can use the 'S' keyboard key as a shortcut.

Note: If multiple files are selected, the image shown is from the first file in the list.

Clear all

Resets and clears all plots and also clears the item lists in the Manipulation, IFT, and SEC tab.

Plot SEC

The "Plot SEC" button plots the selected images and/or 1D scattering profiles as a *SEC curve*. The images are integrated into 1D scattering profiles.

Note: This button can also be used to plot files with the .sec extension that have previously been saved by RAW.

Manipulating files and folders

Files and folders can be manipulated in the Files tab by right clicking on a filename in the file list. The pop-up menu has the options to create a new folder, rename a file, copy/cut/paste file(s) or delete file(s).

Note: These options change the files on disk, not just in RAW! So if you delete a file here, it will be deleted from your disk. If you want to work with the files you have loaded into RAW, see the sections on the *Manipulation*, *IFT*, and *SEC Control Panels*.

4.6.8 The Manipulation Panel and Main Plot Window

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

When individual scattering profiles are loaded, they are placed in the Manipulation tab of the Control Panel, and displayed in the Main Plot window. This section will cover the features of the Manipulation panel and the Main Plot Panel.

The Manipulation Panel

The Manipulation Panel refers to the panel that is shown in the Control Panel when the Manipulation Tab is selected.

The panel consists of two parts. The top part is where individual scattering profiles that are loaded into RAW are shown. The bottom part consists of buttons that allow you to manipulation these scattering profiles. Further manipulation of the scattering profiles can be done from the right click (context) menu and from the Tools menu.

An individual item in the Manipulation Panel list is called a manipulation data item.

All items loaded into the manipulation panel are initially displayed in the top plot of the Main Plot.

Manipulation Data Items

When a new item is plotted, a data item is added to the manipulation list in the manipulation tab. This allows you to control the properties of the data item, and perform manipulations and analysis on it.

The name of the data item is displayed at the top of each item. If an item is given a different name for the plot legend, this legend name is displayed in [square brackets] next to the item name. On the same line as the item name, on the right side of the data item, there are a *series of buttons* that can be used for further manipulation of the data item.

Note: If there is a * to the left of the item name (between the checkbox and the item name), it indicates there are unsaved changes to the item. This can occur if you change the scale, offset, or q range of the item, if the item is a newly created scattering profile (such as integrated from an image, or an averaged or subtracted item), if there is new analysis information associated with the item (such as Rg and I0), and from binning and interpolating the scattering profile.

Offsetting, scaling and adjusting the q-range of a data item

Each data item lets the user interactively adjust the scaling, offset and q-range by using the spin-buttons or entering values in the associated text field and hitting enter. These changes will be updated on the scattering profile displayed on the main plot.

Offset and Scale

The spin-buttons in the scaling and offset boxes will adjust the last digit of the displayed number by 1. So if the scale is 1.0, using the up spin button will adjust it 1.1, while the down spin button will adjust it to 0.9. If the scale was set to 1.15, the up button would adjust it to 1.16, while the down button would adjust it to 1.14.

The scale can never be set to zero by the spin buttons, so if it is adjusted down in a way that would do so, it starts adjusting the next significant figure. For example, if you use the down button on a scale of 0.1, the new scale will be 0.09. A negative scale factor is treated like a positive scale factor (the absolute value of the scale factor is applied), thought it can cause unexpected behavior from the spin buttons.

Q-range

Adjusting q Min and q Max can be done in two ways. Each q limit has two boxes associated with it. The left box for each displays the actual q value, and a q value can be typed into those boxes. As there are a finite set of points in q space, RAW will automatically determine the nearest q point, and set the q value to this nearest point. The box on the right is the (0 indexed) point in the q vector. That is, if for q Min the value reads 15, then the scattering profile is being displayed starting at point number 15 in the q vector (note that this is the 16th point of the q vector, because it is zero indexed). The spin buttons adjust these point numbers, so if you use the up spinner on q Min, it will start at the next point in the q vector.

Showing/Hiding data items on the Main plot

To show/hide the scattering profile associated with a data item on the plot, click the checkbox next to the sample filename. If the checkbox is checked, the item is currently shown on the Main plot, if the checkbox is unchecked, the item is currently hidden on the main plot.

Data item buttons

On the right of the top line of the data item (the same line as the item name and show/hide checkbox) are a series of buttons for controlling the data item. In order of left to right, these buttons have the following properties.

Collapse/Expand item controls

In order to see more data items in the list, or to prevent accidental changes to the q Min, q Max, Scale, or Offset values, the controls for a data item can be hidden using the triangle button in the upper right of the data item. Note that the direction of the triangle switches when it is toggled. If the triangle is pointing up, the controls are expanded, if the triangle is pointing down the controls are collapsed.

Extended Info

By hovering the mouse over the i in the blue circle button, a tooltip will appear which shows more information about the item, including Rg and MW.

Locate Line

The target button is used to highlight the scattering profile on the graph that is associated with the data item. When the target is pressed, it 'bolds' the line of the scattering profile (increases the line width by several points). When the target is pressed again, the line width is set back to normal. You can tell if a line is currently bolded, as the target will be orange instead of grey.

Line Properties

The colored line button has two purposes. First, the color matches the current color of the scattering profile in the Main Plot. Second, when pressed it opens a *line properties dialog* which allows you to set the legend label; the line style, width, and color; the data point marker style, size, line color, and fill color; and the error bar line style, width, and color.

Mark

The star button marks an item. This is used when doing operations such as subtraction, syncing, superimposing or merging. In those cases, the marked (also referred to as starred) item has a special significance.

Selecting data items

A single data item can be selected by clicking on the item name in the Manipulation list (similar to how you would select files in your system file browser). When an item is selected, the color of the item background changes from white to gray. If the item is currently selected, clicking on it will cause it to be unselected. Note that for a regular click, all other selected items will be unselected when a new item is selected.

Multiple items may be selected in two ways. If the Control key (Command key on Macs) is held down while clicking on items, each item that is clicked on will be added to the set of selected items. If a single item is first selected and then the Shift key is held down and another item is selected, all of the items in the list between the two items will be selected (including the second item that is clicked on).

All of the items in the list can be selected in two ways. The first is using the *select all button*, the second is pressing Ctrl-A (Cmd-A), the Control (Command) key and the A key at the same time when you are in the Manipulation panel. All items can be unselected by clicking in a empty spot of the Manipulation list (but not above or below the list), or by clicking on an already selected item.

Note: If you have a set of selected items and wish to remove some, holding down the Control (Command) key and clicking on selected items will deselect them without affecting the other selected items.

The top buttons of the Manipulation Panel

The Manipulation Panel has a set of five buttons at the top of the panel. These buttons have the following effects, listed from left to right.

Show All

Clicking on the button that looks like an eye will show all scattering profiles. This is the same as if you manually set all of the show/hide checkboxes in the data items to on.

Hide All

Clicking on the button that looks like an eye with a red x through it will hide all scattering profiles. This is the same as if you manually set all of the show/hide checkboxes in the data items to off.

Select All

Clicking on the button that looks like a spreadsheet with selected cells will select all of the Manipulation items.

Collapse All

Clicking on the button that looks like an upward pointing arrow with a box under it will collapse all of the data item controls. This is the same as if you manually toggled the Expand/Collapse button for each data item to the collapse position.

Expand All

Clicking on the button that looks like a downward pointing arrow with a box over it will expand all of the data item controls. This is the same as if you manually toggled the Expand/Collapse button for each data item to the expand position.

Synchronizing settings for several data items

Settings for several data items can be easily synchronized.

1. Click the star icon on the data item you want the other items synchronized to.
2. *Select* one or more data items, you wish to synchronize with the starred item.
3. Click the "Sync" button
4. Select what parameters you want synchronized
5. Hit the "OK" button

Parameters that can be synchronized in this way are: q min, q max, n min, n max, scale, offset, line style, line width, and line marker.

Renaming a data item

Data items can be renamed by selecting the data item of interest and selecting "Rename" in the right click popup menu.

Note: While no characters are expressly forbidden in the filename, RAW does not sanitize file names before saving, and thus special characters such as '/' and '\' are likely to cause problems when the file is saved.

Saving data items

All scattering profiles can be saved to a standard 3 column ASCII format with a header. To save:

1. Select the item(s) to be saved.
2. Click the "Save" button or select "Save selected file(s)" from the right click menu.
3. In the window that pops up, navigate to the directory in which you want to save the files.
4. If you are saving a single item, the window will give you an opportunity to rename your file if desired. Click "Save" when ready.
5. If you are saving multiple items, you simply need to select the folder for the items to be saved in, and click "Open". The items will be saved with the same names displayed in the Manipulation Panel, in the folder that you chose.

Items are saved in a standard format with three columns of data. These columns have listed headers, and are Q, I, and Error in I respectively. Depending on your settings, at the bottom (default) or top of the file there will be a "Header". This header contains any information loaded as 'Header' information from the image or separate header file, the normalization factor, any notes you made in the Notes section of the information box, any analysis results (such as Rg, MW), and the manipulation history saved by the averaging, subtracting, merging, rebinning, and interpolating functions (see following sections).

The files are saved with a ".dat" extension, which is the standard for scattering profiles and can be read in by most standard SAXS processing software (including Scatter, Primus, and programs in the ATSAS package). The files are simply text files, and can be opened and viewed in any standard text editor. The header is saved in the JSON format.

Removing data items from the manipulation list

To remove one or more data items, select them and do one of the following:

1. Press the "Delete" key on the keyboard
2. Click the "Remove" button
3. Select "Remove" from the right click menu

Averaging data

Averaging data is an important part of processing SAXS data. At the moment in RAW, data can only be averaged if the q vectors match exactly. To average data:

1. *Select* two or more data items that should be averaged.
2. Either click the "Average" button or right-click on a selected item and choose "Average" from the popup menu.

After averaging, a new data item will be displayed in the manipulation panel with the filename in green and an 'A_' prefix. All averaged items are initially displayed on the top plot of the Main plot, regardless of where the items being averaged are displayed.

A record of which profiles were averaged is saved in the metadata for the averaged profile. This can be viewed within RAW by right clicking on the item and selecting "Show history", or externally if the scattering profile is saved as a .dat file and opened by an external text editor. This history contains all of the history of the individual files, so if you average a set of averaged (or subtracted) files, you can dig down through the metadata to see what files were used at each step.

Subtracting data

Subtracting data is an important part of processing SAXS data. RAW can handle subtracting data that has different q vectors, as long as there is a region of overlap between the q vectors. Data subtraction is done in the following way:

1. Click the star icon on the data item to mark the scattering profile that should be used for subtraction (typically the background or buffer scattering profile).
2. *Select* another data item, or multiple items.
3. Either click the "Subtract" button or right-click and select "Subtract" from the popup menu. Note: this subtracts the marked item (step 1) from the selected items (step 2).

After subtraction, a new data item will be displayed in the Manipulation panel with the filename in red and a 'S_' prefix. All subtracted items are initially displayed on the bottom plot of the main plot, regardless of where the items being subtracted are displayed.

A record of which profiles were involved in the subtraction is saved in the metadata for the subtracted profile This can be viewed within RAW by right clicking on the item and selecting "Show history", or externally if the scattering profile is saved as a .dat file and opened by an external text editor. This history contains all of the history of the individual files, so if you subtract a set of averaged (or subtracted) files, you can dig down through the metadata to see what files were used at each processing step.

Data with different q vectors is handled as follows. First, the range of overlap is established. Second, RAW checks whether these overlapping regions have matching q vectors, if so, RAW carries out the subtraction. If not, RAW bins both curves so that the q vectors in the overlapping region match, and then carries out the subtraction.

Superimposing data

Scattering profiles can be superimposed on each other automatically. This function tries to calculate a scale and an offset that minimizes difference between two or more curves using a least-squares method. To superimpose:

1. Click the star icon on the data item you want the other items superimposed to.
2. *Select* one or more data items, that you wish to superimpose onto the starred item.
3. Click the "Superimpose" button

Merging data

Merging scattering profiles is typically done if you have measured multiple q ranges using different detectors. RAW can merge these profiles into one single profile. To merge profiles in RAW:

1. Click the star icon on one of the curves. This items name will be used for the name of the merged item, otherwise it does not matter.
2. *Select* one or more data items, that you wish to merge together with the starred item
3. Click the "Merge" button, select "Merge" from the right click menu, or select "Merge" from the Tools->Operations menu.

After merging, a new data item will be displayed in the Manipulation panel with a filename with a 'M_' prefix. All merged items are initially displayed on the top plot of the Main plot, regardless of where the items being merged are displayed.

Merging will average together the overlapping regions, if any exist. It can be done on an arbitrary number of data items, and is done in serial. So first the two items with the two lowest q vectors are merged, then that item is merged with the data item with the next lowest q vector, and so on. For example, if you had selected items with a minimum q vector of 0.01, 0.25, and 0.75, it would first merge the 0.01 and 0.25 items, and then merge that resulting item with the 0.75 item.

A record of which profiles were merged is saved in the metadata for the averaged profile. This can be viewed within RAW by right clicking on the item and selecting "Show history", or externally if the scattering profile is saved as a .dat file and opened by an external text editor.

Note: If items are merged without an overlap region, there will simply be no data in those intervening points.

Rebinning data

Rebinning scattering profiles is typically done to improve signal to noise of individual points. As most scattering profiles are significantly oversampled (compared to the Shannon sampling frequency), there is no loss of information from doing this. To rebin curves in RAW:

1. *Select* one or more data items, that you wish to rebin with the same binning factor.
2. Select "Rebin" from the right click menu, or select "Rebin" from the Tools->Operations menu.
3. Select the bin reduction factor, and if you want logarithmic binning check the logarithmic box. Note: the log binning is log base 10.
4. Click "OK".

After rebinning, a new data item will be displayed in the Manipulation panel with a filename with an 'R_' prefix. All rebinned items are initially displayed on the top plot of the Main plot, regardless of where the items being rebinned are displayed.

A record of which profiles were averaged is saved in the metadata for the rebinned profile. This can be viewed within RAW by right clicking on the item and selecting "Show history", or externally if the scattering profile is saved as a .dat file and opened by an external text editor.

Note: The bin reduction factor can be thought of as how many q points will go into a single bin in the new profile. So a bin reduction factor of 2 will result in half as many q points in the new profile, with each q point being the average of two q points from the previous profile (note that this may not quite be true for q points at the start and end of the profile). Rebinning will slightly reduce the q range of the scattering profile, as the new q value will be less than the extreme values in each bin.

Interpolating data

RAW can interpolate a scattering profile to find values at all points in a reference q vector. This can be useful for creating scattering profiles with perfectly matched q-vectors. To interpolate curves in RAW:

1. Click the star icon on the scattering profile you wish to interpolate to. The q vector for this profile will be used as the interpolation points for the other profile(s).
2. *Select* one or more data items, that you wish to interpolate to match the q vector of the starred item
3. Select "Interpolate" from the right click menu, or select "Interpolate" from the Tools->Operations menu.

After interpolating, a new data item will be displayed in the Manipulation panel with a filename with an 'I_' prefix. All interpolated items are initially displayed on the top plot of the Main plot, regardless of where the items being merged are displayed.

Interpolation works as follows: The overlapping q region is found for the two curves. Within this region, for every point in the q vector of the reference (starred) scattering profile an intensity value is found from the item being interpolated via linear interpolation (using the scipy.interpolate.interp1d function). A new scattering profile is then created from these interpolated values and the reference q vector.

A record of which profiles were averaged is saved in the metadata for the averaged profile. This can be viewed within RAW by right clicking on the item and selecting "Show history", or externally if the scattering profile is saved as a .dat file and opened by an external text editor.

Converting q-scale

Q-values may sometimes be in inverse angstroms (1/Å), while others use inverse nanometers (1/nm). The q vector can be multiplied or divided by 10 to accommodate for this difference when comparing two curves on different q-scales.

1. Select the data item(s) to be scaled and right-click on them.
2. Select the "Convert q-scale" option and choose whether to adjust the q-scale upwards or downwards by a factor of 10.

Note: All of the processing in RAW assumes that the scattering profiles are in inverse angstroms.

Normalizing by concentration

The overall intensity of the scattering profile should be proportional to the concentration, so it can be useful to normalize scattering profiles by the measured concentration value. To do so:

1. Set the concentration of the item in the *Information panel* or the *MW window*.
2. Select the item to be normalized.
3. Right click on the item and select "Normalize by conc"

Note: This function simply calculates 1/conc and sets that value as the scale factor. This is equivalent to doing the same calculation and manually setting the scale factor of the scattering profile.

Saving analysis results

After analyzing the data using the Guinier, MW, BIFT, or GNOM panels, the resulting parameters such as Rg, I(0), MW, Dmax, and others can be saved to a comma separated format that easily loads into a spreadsheet program such as Excel. To do so:

1. Select the data items of interest.
2. Right-click on a selected data item and select "Save all analysis info"
3. Choose a name and a location for the .csv file
4. Load the .csv file into a spreadsheet program.

This saves a spreadsheet where each row is an individual scattering profile, and each column an analysis parameter.

Saving select data item information

Information about data items including any analysis parameters, header file information, image header information, scale factor, offset, concentration, and notes can be selectively saved to a comma separated format that easily loads into a spreadsheet program such as Excel. To do so:

1. Select the data items of interest
2. Right-click on a selected data item and select "Save item info"
3. A window will pop up with two panels. The panel on the left is all of the information available about the selected items, sorted into appropriate categories (such as Image Header and Guinier Analysis).
4. In the window, select the information of interest in the left hand panel by clicking (including Ctrl and Shift clicking) on it.
5. Once all of the variables of interest are selected (blue), click the "->" button to put them in the list of data that will be saved. They should appear in the right hand panel.
6. If you accidentally selected an item, select the item and click the "<-" button to remove it from the list of data to be saved.
7. You can add additional items to the list by repeating steps 4 and 5 until everything you want to save is selected.
8. Click the "OK" button.
9. Choose a name and a location for the .csv file
10. Load the .csv file into a spreadsheet program.

This saves a spreadsheet where each row is an individual scattering profile, and each column one of the selected variables.

Note: This differs from the save function described *above* in two ways. First, it allows you select which information you want to save. Second, it can save header information and other parameters about the scattering profile beyond the analysis information.

Data point browsing

The q, I, and I error values for each individual point in a scattering profile on the data curve can be inspected using the data browser. To do so:

1. Right-click on the data item of interest.
2. Select "Show data" in the popup menu.

Displaying the detector image

If the plotted data was loaded from an image, the image can be shown in the Image tab of RAW. To do so, right-click on the data item of interest and click "Show image" in the popup menu.

Displaying the header information

If the plotted data has header information, either from an image header or separate header file, this can be displayed in RAW. To do so, right-click on the data item of interest and click "Show header" in the popup menu.

Displaying the history information

If the plotted data has history information, either from integration of an image or from manipulation inside RAW such averaging or subtraction, this can be displayed in RAW. To do so, right-click on the data item of interest and click "Show history" in the popup menu.

The Manipulation data item right click menu options

When you right click on a data item, a popup menu is shown. This section describes what each item on the menu does.

Subtract

Carries out *subtraction*.

Average

Carries out *averaging*.

Merge

Carries out *merging*.

Rebin

Carries out *rebinning*.

Interpolate

Carries out *interpolation*.

Use as MW Standard

Sets the selected item to be the MW standard used in the *I(0) reference MW* calculation method. To do this the concentration must be set, for example in the Information Panel (it can also be set in the MW panel). A Guinier fit must also have been done, as this requires a known I(0) value for the reference scattering profile. In the popup window that opens, enter the known protein MW in units of kDa.

Normalize by conc

Normalizes the profile by the concentration.

Remove

Removes the item.

Guinier fit

Opens the Guinier fit analysis panel for the selected scattering profile.

Molecular Weight

Opens the molecular weight analysis panel for the selected scattering profile.

BIFT

Opens the Bayesian indirect Fourier transform (BIFT) analysis panel for the selected scattering profile.

GNOM (ATSAS)

Opens the GNOM analysis panel for the selected scattering profile. Note: this option is only available if RAW has detected a valid ATSAS package installation.

SVD

Opens the singular value decomposition (SVD) analysis panel for the selected scattering profiles (must have at least 2 profiles selected).

EFA

Opens the evolving factor analysis (EFA) panel for the selected scattering profiles (must have at least 2 profiles selected).

Convert q-scale

Has a submenu with two options: **q * 10** which multiples the scattering profile q vector by a factor of 10, and **q / 10** which divides the scattering profile q vector by a factor of 10.

Show image

If the scattering profile was integrated from an image, and the image is still in the same location as when the profile was first integrated, it displays the image in the Image Plot Panel.

Show data

Shows the q, I(q) and I_error(q) data.

Show header

Shows the header information.

Show history

This shows the history of the scattering profile.

Move to top plot

Moves the scattering profile to the top subplot of the Main plot (no effect if the scattering profile is already plotted on the top plot).

Move to bottom plot

Moves the scattering profile to the bottom subplot of the Main plot (no effect if the scattering profile is already plotted on the top plot).

Save all analysis info

Saves all of the analysis info of the selected data item(s).

Save item info

Saves select item info.

Save selected file(s)

Saves the selected data item(s).

The Manipulation panel bottom buttons

There are six buttons at the bottom of the Manipulation panel. They are:

Save

This button saves the selected data item(s).

Sync

This button syncs the settings of the selected data item(s) to the starred data item.

Remove

This button removes the selected data item(s) from the Manipulation panel.

Superimpose

This button superimposes the selected data item(s) onto the starred data item.

Average

This button averages the selected data item(s).

Subtract

This button subtracts the selected data item(s).

The Main Plot window

The Main Plot window displays individual scattering profiles associated with items in the Manipulation panel. It has two subplots, which by default are titled "Main Plot" (top) and "Subtracted" (bottom). Any item that is initially loaded is plotted on the top plot.

The features that are general between all of the plots are described *elsewhere*. This section will describe features unique to this plot.

Changing axes and plot types

Right-click in the plot window to view a pop-up menu with different axis settings.

The available plot modes are:

- Lin-Lin
- Log-Lin
- Log-Log
- Lin-Log
- Guinier plot (ln(I(q)) vs. q^2)
- Porod plot (q^4I(q) vs. q)
- Kratky plot (q^2I(q) vs. q)

The Main plot toolbar

In addition to the plot toolbar buttons *shared by all of the plots*, the Main plot has the following buttons:

Toggle errorbars. Shows the errorbars on the plotted curves.

Top/Bottom plot. Shows both the top and the bottom plot.

Top plot. Shows only the top plot.

Bottom plot. Shows only the bottom plot.

Moving profiles between plots

Scattering profiles can be moved between the top and bottom plots. To do so:

1. Select the data item(s) of interest in the Manipulation panel.
2. Right-click on a selected data item and select "Move to top plot" or "Move to bottom plot" from the popup menu.

Setting the legend label

Instead of the filename showing in the legend, a special label can be created for that data item.

1. Click on the *Line Properties* button of the data item.
2. In the Line Properties window that pops up, enter the desired legend label in the Legend Label box.
3. Click "OK".

The new legend label will be shown in [square brackets] next to the data item name. To revert to t he standard label (the filename), erase the custom label.

Locating data items

The data item associated with a plotted scattering profile can be quickly found by clicking plotted curve. When a curve is clicked on, it will flash 'bold' (increased line weight), and the associated data item will be selected in the Manipulation panel.

Note: If there are other selected items in the Manipulation panel, they will be unselected.

4.6.9 The IFT Control Panel and IFT Plot Panel

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

When inverse Fourier transforms (P(r) functions) are loaded into RAW or generated by BIFT or GNOM in RAW, they are placed in the IFT tab of the Control Panel, and displayed in the IFT Plot window. This section will cover the features of the IFT control panel and IFT plot panel.

The IFT control panel

The IFT control panel refers to the panel that is shown in the Control panel when the IFT tab is selected.

The panel consists of two parts. The top part is where individual IFT items loaded into RAW are shown. The bottom part consists of buttons that allow you to save or remove those IFT items. Further manipulation of the IFT items can be done from the right click (context) menu and from the Tools menu.

An individual item in the IFT list is called an IFT data item. It is associated with a P(r) function, the scattering profile used to generate that P(r) function, and the scattering profile generated from the P(r) function.

All items loaded into the IFT panel have a P(r) function displayed in the top plot of the IFT Plot window, and an experimental scattering profile and a scattering profile from the P(r) function displayed in the bottom plot of the IFT Plot window.

IFT Data Items

When a new item is plotted, a data item is added to the IFT list in the IFT tab. This allows you to control the properties of the data item, and perform analysis on it.

The name of the data item is displayed for each item. If an item is given a different name for the plot legend, this legend name is displayed in [square brackets] next to the item name. On the same line as the item name, on the right side of the data item, there are several buttons that can be used for *further manipulation* of the data item.

Note: If there is a * to the left of the item name (between the checkbox and the item name), it indicates there are unsaved changes to the item. This can occur if the item is newly created IFT data (such as from the BIFT or GNOM panels), or if item properties such as the name have been changed.

Showing/Hiding data items on the IFT plot

To show/hide all of the curves (P(r), experimental data, scattering profile from P(r) curve) associated with a data item on the plot, click the checkbox next to the filename. If the checkbox is checked, the item is currently shown on the IFT plot, if the checkbox is unchecked, the item is currently hidden on the IFT plot.

Data item buttons

On the right of the data are a series of buttons for controlling the data item. In order of left to right, these buttons have the following properties.

Locate Line

The target button is used to highlight the scattering profile on the graph that is associated with the data item. When the target is pressed, it 'bolds' the line each plotted curve associated with the data item (increases the line width by several points). When the target is pressed again, the line width is set back to normal. You can tell if a line is currently bolded, as the target will be orange instead of grey.

Line Properties

The colored line button has two purposes. First, the color matches the current color of the P(r) function in the IFT Plot. Second, when pressed it opens a *line properties dialog* which allows you to set the legend label; the line style, width, and color; the data point marker style, size, line color, and fill color; and the error bar line style, width, and color for each line associated with the IFT data item.

Selecting data items

A single data item can be selected by clicking on the item name in the IFT list (similar to how you would select files in your system file browser). When an item is selected, the color of the item background changes from white to gray. If the item is currently selected, clicking on it will cause it to be unselected. Note that for a regular click, all other selected items will be unselected when a new item is selected.

Multiple items may be selected in two ways. If the Control key (Command key on Macs) is held down while clicking on items, each item that is clicked on will be added to the set of selected items. If a single item is first selected and then the Shift key is held down and another item is selected, all of the items in the list between the two items will be selected (including the second item that is clicked on).

All of the items in the list can be selected in two ways. The first is using the *select all* button, the second is pressing Ctrl-A (Cmd-A), the Control (Command) key and the A key at the same time when you are in the IFT panel. All items can be unselected by clicking in an empty spot of the IFT list (but not above or below the list), or by clicking on an already selected item.

Note: If you have a set of selected items and wish to remove some, holding down the Control (Command) key and clicking on selected items will deselect them without affecting the other selected items.

The top buttons of the IFT Panel

The IFT Panel has a set of three buttons at the top of the panel. These buttons have the following effects, listed from left to right.

Show All

Clicking on the button that looks like an eye will show all IFT items. This is the same as if you manually set all of the show/hide checkboxes in the data items to on.

Hide All

Clicking on the button that looks like an eye with a red x through it will hide all IFT items. This is the same as if you manually set all of the show/hide checkboxes in the data items to off.

Select All

Clicking on the button that looks like a spreadsheet with selected cells will select all of the IFT data items.

Renaming a data item

Data items can be renamed by selecting the data item of interest and selecting "Rename" in the right click popup menu.

Note: While no characters are expressly forbidden in the filename, RAW does not sanitize file names before saving, and thus special characters such as '/' and '\' are likely to cause problems when the file is saved.

Saving data items

IFT items are saved in two different formats, depending on whether the item was generated by BIFT (".ift") or GNOM (".out"). The procedure to save either is the same. To save:

1. Select the item(s) to be saved.
2. Click the "Save" button or select "Save selected file(s)" from the right click menu.
3. In the window that pops up, navigate to the directory in which you want to save the files.

4. If you are saving a single item, the window will give you an opportunity to rename your file if desired. Click "Save" when ready.
5. If you are saving multiple items, you simply need to select the folder for the items to be saved in, and click "Open". The items will be saved with the same names displayed in the IFT Panel, in the folder that you chose.

BIFT items are saved as ".ift" files, which is a text file with RAW specific formatting of the text. The first two lines are "BIFT" and the "Filename: <filename>". After that, the P(r) function is saved as 3 column data. The first column is "R", the second column is "P(R)" and the third column is "Error", these headers are included as the third line of the file. After the P(R) function there are two blank lines, followed by the scattering data. This is saved in four columns, the first three are "Q", "I(q)" and "Error" which correspond to the experimental data, while the fourth column is "Fit" which is the scattering profile from the P(r) function. After this data is written there is a "header" written, which consists of the Chisquared, algorithm used, I(0) value, log base 10 of the alpha value, Dmax, Rg, and the filename, all saved in JSON format. The files are simply text files, and can be opened and viewed in any standard text editor.

GNOM items are saved in the standard ".out" format. This is described in the ATSAS manual for GNOM. These files can be directly input into any ATSAS (or other) program that requires a GNOM .out file as input.

Removing data items from the IFT list

To remove one or more data items, select them and do one of the following:

1. Press the "Delete" key on the keyboard
2. Click the "Remove" button
3. Select "Remove" from the right click menu

Sending data to the main plot

If you wish to examine the experimental scattering profile and fit to this profile from the P(r) function more closely, the data can be sent to the main plot. To do this, right click on the IFT data item and select "To Main Plot". This will plot items on the Main plot and add them to the Manipulation list.

For a ".ift" file generated by BIFT, the following items will be added to the Manipulation list and Main plot. In all cases, <filename> corresponds to the filename of the IFT data item without the extension (so "my_ift".ift would have a filename of "my_ift").

<filename>_data – This is the experimental data that was used to generate the P(r) curve.

<filename>_fit – This is the scattering profile generated from the P(r) curve.

For a ".out" file generated by GNOM, the same two curves as for a BIFT item (above) are sent to the main plot, there is also one additional file.

<filename>_extrap – This is the scattering profile generated from the P(r) curve, extrapolated to q=0.

Running DAMMIF on the P(r) function

RAW allows you to run DAMMIF on a P(r) function from within RAW. Currently, this can only be done on P(r) items generated by GNOM (these can be generated in RAW, or loaded in after being generated outside of RAW). To run DAMMIF, select an appropriate IFT data item (a ".out" item), and either right click and select the "Run DAMMIF" option or from the Tools->ATSAS menu select "DAMMIF". This opens the *DAMMIF window*.

Running AMBIMETER on the P(r) function

RAW allows you to run AMBIMETER on a P(r) function from within RAW. Currently, this can only be done on P(r) items generated by GNOM (these can be generated in RAW, or loaded in after being generated outside of RAW). To run AMBIMETER, select an appropriate IFT data item (a ".out" item), and either right click and select the "Run AMBIMETER" option or from the Tools->ATSAS menu select "AMBIMETER". This opens the *AMBIMETER window*.

Data point browsing

Each individual point of the r; P(r); error in P(r); experimental q, I(q), and error; I(q) from the P(r) function; and, for GNOM generated IFT data items, the q and I(q) values extrapolated to q=0 vectors; can be inspected using the data browser. To do so:

1. Right-click on the data item of interest.
2. Select "Show data" in the popup menu.

The IFT data item right click menu options

When you right click on a data item, a popup menu is shown. This section describes what each item on the menu does.

Remove

Removes the item.

To Main Plot

This *sends the scattering profile data to the main plot*.

Run DAMMIF

This item is only available for ".out" files generated by GNOM. It opens the *DAMMIF window*.

Run AMBIMETER

This item is only available for ".out" files generated by GNOM. It opens the *AMBIMETER window*.

SVD

Opens the singular value decomposition (SVD) analysis panel for the selected scattering profiles (must have at least 2 items selected).

EFA

Opens the evolving factor analysis (EFA) panel for the selected scattering profiles (must have at least 2 items selected).

Show data

Shows the *individual data points*.

Rename

Renames the data item,

Save selected file(s)

Saves the selected data item(s).

The IFT panel bottom buttons

There are three buttons at the bottom of the IFT control panel. They are:

Save

This button *saves the selected data item(s)*.

Remove

This button *removes the selected data item(s)* from the IFT panel.

Clear IFT Data

This button clears all loaded IFT data. It works the same as if you had selected all of the IFT data items and then removed them.

The IFT Plot window

The IFT Plot window displays P(r) data (top plot), the scattering profiles generated from the P(r) data (bottom plot), and the experimental scattering profile used to generate the P(r) data (bottom plot). Each set of three curves is associated with a single IFT data item in the IFT control panel. The two subplots are the Pair Distance Distribution Function (top) and Data/Fit (bottom) plots.

The features that are general between all of the plots are described *elsewhere*. This section will describe features unique to this plot.

Changing axes and plot types

Right-click in the Data/Fit (top) plot to view a pop-up menu with different axis settings.

The available plot modes are:

- Lin-Lin
- Log-Lin
- Log-Log
- Lin-Log
- Guinier plot (ln(I(q)) vs. q^2)
- Porod plot (q^4I(q) vs. q)
- Kratky plot (q^2I(q) vs. q)

The axes cannot be changed for the P(r) (top) plot.

The IFT plot toolbar

In addition to the plot toolbar buttons *shared by all of the plots*, the IFT plot has the following buttons:

Toggle errorbars. Shows the error bars on the plotted curves.

Top/Bottom plot. Shows both the top and the bottom plot.

Top plot. Shows only the top plot.

Bottom plot. Shows only the bottom plot.

4.6.10 The SEC Control Panel and SEC Plot Panel

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

When SEC-SAXS data is loaded into RAW, it is placed in the SEC tab of the Control Panel, and displayed in the SEC Plot window. This section will cover the features of the SEC control panel and the SEC plot panel.

SEC-SAXS data is assumed to come from continuous collection of images while sample is being eluted from an FPLC column. Depending on the frame rate of the detector and the total elution time, this can generate thousands of images. The basic approach RAW uses to deal with this data is to load every scattering profile (or integrate every image into a scattering profile), and associate it with a single SEC data item. For each SEC item, the total (or average) intensity in each scattering profile is plotted vs. frame number, where frame number is simply what position the item was loaded in (so if 10 items are loaded, the frame number for the first item loaded would be 0, for the next item loaded 1, etc). For standard SEC-SAXS data, this intensity vs. frame number should look similar to the UV chromatograph of the sample after it passes through the column. Some analysis can be carried out directly on this SEC data item, and the scattering profiles of interest can be extracted for further analysis.

The SEC control panel

The SEC control panel refers to the panel that is shown in the Control panel when the SEC tab is selected.

The panel consists of three parts. The top part is a set of controls for loading SEC items, controlling the SEC online mode, and analyzing SEC items. The middle part is where individual SEC items loaded into RAW are shown. The bottom part consists of buttons that allow you to save or remove the SEC items. Further manipulation of the SEC items can be done from the right click (context) menu.

An individual item in the SEC list is the called a SEC data item. It is associated with two curves on the SEC plot, one plotted on the right y axis and one on the left y axis.

SEC Data Items

When a new item is plotted, a data item is added to the SEC list in the SEC control panel. This allows you to control the properties of the data item, and perform analysis on it.

The name of the data item is displayed for each item. If an item is given a different name for the plot legend, this legend name is displayed in [square brackets] next to the item name. On the same line as the item name, on the right side of the data item, there are several buttons that can be used for further manipulation of the data item.

SEC data items are different from Manipulation and IFT data items as they have many (often hundreds or thousands) of scattering profiles associated with a single data item, and they can have new profiles appended to them after they are loaded.

Note: If there is a * to the left of the item name (between the checkbox and the item name), it indicates there are unsaved changes to the item. This can occur if the item is newly created IFT data (such as from the BIFT or GNOM panels), or if item properties such as the name have been changed.

Showing/Hiding data items on the SEC plot

To show/hide the curves on the plot associated with a data item, click the checkbox next to the filename. If the checkbox is checked, the item is currently shown on the SEC plot, if the checkbox is unchecked, the item is currently hidden on the SEC plot.

Data item buttons

On the right of the data item are several buttons for controlling the data item. In order of left to right, these buttons have the following properties.

Extended Info

By hovering the mouse over the i in the blue circle button, a tooltip will appear which shows more information about the item. It shows the buffer range and window size used (if any) to calculate the structural parameters of the SEC data item as a function of frame number.

Locate Line

The target button is used to highlight the scattering profile on the graph that is associated with the data item. When the target is pressed, it 'bolds' the line each plotted curve associated with the data item (increases the line width by several points). When the target is pressed again, the line width is set back to normal. You can tell if a line is currently bolded, as the target will be orange instead of grey.

Line Properties

The colored line button has two purposes. First, the color matches the current color of the Intensity curve in the SEC Plot. Second, when pressed it opens a *line properties dialog* which allows you to set the legend label; the line style, width, and color; the data point marker style, size, line color, and fill color; and the error bar line style, width, and color for each line associated with the SEC data item.

Mark

The star button marks an item. This is used when doing operations such as sending data frames to plot or calculating structural parameters. In those cases, the marked (also referred to as starred) item has a special significance.

Loading data items

SEC data items can be loaded from the *files tab*. There is an additional way to load files available in the SEC control panel, this method is intended to be used in conjunction with the SEC online mode. This method is described in the rest of this section.

At the top of the SEC control panel, there is a section for loading files. It has a button "Select file in SEC run", and then descriptive information about the file: the image prefix, the run number, and the frame numbers of the file. Note that none of these descriptive fields can be edited. Below the descriptive fields is an "Update" button and an "Automatically Update" checkbox.

To load a SEC curve using this method, click the "Select file in SEC run" button. In the file browser that pops up, select any file in the SEC data set. The data set will automatically be loaded into RAW as a SEC item, and the descriptive

information will be displayed in the appropriate boxes. This loaded file can then be used with the SEC automatic update mode described *below*.

Note: Due to how files are loaded from the SEC control panel, it only works for certain header file types. RAW must be able to automatically determine what files are associated with the SEC run, from any file in the run. At the moment, this can only be done with the G1 and G1 WAXS header file formats. If you want to have a particular beamline's file format added to this loading (and thus able to use the online mode) please contact the RAW developers.

Updating a SEC data item

If a SEC data item is loaded via the SEC control panel, it can be updated if additional files are added to the folder that are part of the data collection. This can only be done for the most recent file loaded via the SEC control panel, for which the descriptive information is shown at the top of the SEC control panel.

To do this update, hit the "Update" button on the SEC control panel. RAW will automatically determine all of the files associated with the SEC run, based on the file you selected with the "Select file in SEC run" button, determine if any of those files have not yet been loaded, and if so, load the files and add them to the SEC item as appropriate.

This is useful when working at the beamline while data is being collected, as you may want to start analysis of the SEC curve before all of the images are taken.

SEC automatic update (online mode)

The SEC control panel can be used in an online mode, where a SEC curve is automatically updated as data comes in. All of the conditions for manually updating a curve, described above, must be met. Instead of using the "Update" button as in that section, check the "Automatically Update" box in the SEC control panel.

Note: The automatic update will stay on even if you load a new file into the SEC panel using the "Select file in SEC run" button. It will switch to updating this newly loaded file. As with the "Update" button described above, the automatic update only applies to the SEC data item most recently loaded in the SEC panel.

Selecting data items

A single data item can be selected by clicking on the item name in the SEC list (similar to how you would select files in your system file browser). When an item is selected, the color of the item background changes from white to gray. If the item is currently selected, clicking on it will cause it to be unselected. Note that for a regular click, all other selected items will be unselected when a new item is selected.

Multiple items may be selected in two ways. If the Control key (Command key on Macs) is held down while clicking on items, each item that is clicked on will be added to the set of selected items. If a single item is first selected and then the Shift key is held down and another item is selected, all of the items in the list between the two items will be selected (including the second item that is clicked on).

All of the items in the list can be selected in two ways. The first is using the *select all* button, the second is pressing Ctrl-A (Cmd-A), the Control (Command) key and the A key at the same time when you are in the SEC panel. All items can be unselected by clicking in an empty spot of the SEC list (but not above or below the list), or by clicking on an already selected item.

Note: If you have a set of selected items and wish to remove some, holding down the Control (Command) key and clicking on selected items will deselect them without affecting the other selected items.

The top buttons of the SEC item list

The SEC item list has a set of three buttons at the top of the panel. These buttons have the following effects, listed from left to right.

Show All

Clicking on the button that looks like an eye will show all SEC items. This is the same as if you manually set all of the show/hide checkboxes in the data items to on.

Hide All

Clicking on the button that looks like an eye with a red x through it will hide all SEC items. This is the same as if you manually set all of the show/hide checkboxes in the data items to off.

Select All

Clicking on the button that looks like a spreadsheet with selected cells will select all of the SEC data items.

Renaming a data item

Data items can be renamed by selecting the data item of interest and selecting "Rename" in the right click popup menu.

Note: While no characters are expressly forbidden in the filename, RAW does not sanitize file names before saving, and thus special characters such as '/' and '\' are likely to cause problems when the file is saved.

Saving data items

SEC items are saved as ".sec" files, and is the only data that RAW does not save in a human readable format. To save:

1. Select the item(s) to be saved.
2. Click the "Save" button or select "Save selected file(s)" from the right click menu.
3. In the window that pops up, navigate to the directory in which you want to save the files.
4. If you are saving a single item, the window will give you an opportunity to rename your file if desired. Click "Save" when ready.
5. If you are saving multiple items, you simply need to select the folder for the items to be saved in, and click "Open". The items will be saved with the same names displayed in the SEC Panel, in the folder that you chose.

SEC items often contain hundreds or thousands of scattering profiles, so they are not saved in a human readable format. The ".sec" files that RAW saves can only be read by RAW.

Removing data items from the SEC list

To remove one or more data items, select them and do one of the following:

1. Press the "Delete" key on the keyboard
2. Click the "Remove" button
3. Select "Remove" from the right click menu

Exporting SEC data

The following data can be exported in a spreadsheet ready format: frame number, integrated intensity, mean intensity, Rg, Rg error, I(0), I(0) error, MW, filename, and, if available intensity a q=<#> where <#> is a user selected value, for each individual point.

To do so:

1. Select the item(s) to be saved.
2. Select "Export data" from the right click menu.
3. In the window that pops up, navigate to the directory in which you want to save the file(s).
4. If you are saving a single item, the window will give you an opportunity to rename your file if desired. Click "Save" when ready.
5. If you are saving multiple items, you simply need to select the folder for the items to be saved in, and click "Open". The items will be saved with the same names displayed in the SEC Panel, in the folder that you chose.

The data is saved as a comma separated value (".csv") file. This can be opened directly into most spreadsheet programs, such as Excel.

Saving all SEC scattering profiles

If you want to save every individual scattering profile loaded into a SEC data item, you can do so by:

1. Select the item(s) to be saved.
2. Select "Save all profiles as .dats" from the right click menu.
3. In the window that pops up, navigate to the directory in which you want to save the file(s).
4. Select the folder for the items to be saved in, and click "Open". The items will be saved with the same filenames displayed in the *data browser*.

Sending data to the main plot

Individual scattering profiles can be sent to the main plot for further analysis. This utilizes the middle section of the controls at the top of the SEC Control panel.

To do so:

1. Star the SEC data item containing the scattering profiles of interest (note: if only one SEC data item is loaded, it does not have to be starred).
2. Enter the data frames of interest in the "Select Data Frames:" boxes. The box on the left is the first frame of interest, the box on the right is the last frame of interest. All of the frames between those two endpoints (inclusively) are selected.
3. Either click the "Frames To Main Plot" button, which will send each individual frame selected in part 2 to the Main plot, or click "Average To Main Plot" which will send the average of the selected frames to the Main plot.
4. Click on the Manipulation panel and Main plot panel to view the scattering profiles.

Calculate structural parameters

You can calculate the Rg, MW, and I(0) as a function of frame number for a SEC profile. This is done using the "Calculate/Plot Structural Parameters" section of the SEC control panel (the bottom section of the controls at the top of the panel).

It is important to have a big picture for what is happening when this is done. First you set a buffer range. All of the scattering profiles in that range will be averaged, and then subtracted from every loaded scattering profile. Next you set a window size. This window will be slid across the curve and all of the frames within it averaged (note: this average is of the buffer subtracted scattering profiles). This Rg, MW, and I(0) for this averaged profile is then calculated. Those values are assigned to the center frame of the window. The window is then slid down the curve one frame, and the process is repeated until the window reaches the end of the SEC data. For example, if the window size is 5, the first 5 frames, frames 0, 1, 2, 3, and 4, are averaged, and have the Rg, MW, and I(0) calculated. Then window is moved over by one, and frames 1, 2, 3, 4, and 5 are averaged and have the Rg, MW, and I(0) calculated, and so on.

To do this:

1. Star the SEC data item containing the scattering profiles of interest (note: if only one SEC data item is loaded, it does not have to be starred).
2. Enter the buffer range in the "Buffer Range" boxes. The box on the left is the starting buffer frame and the box on the right is the final buffer frame.
3. Enter the window size.
4. Select the appropriate molecule type (Protein or RNA) for the molecular weight calculation.
5. Click the "Set/Update Parameters" button.
6. The structural parameters will be plotted as a function of frame number in the SEC plot. If RAW was unable to determine the parameters for a particular window, then all parameters in that window are set to -1.

The Rg and I(0) calculations are done using RAW's *automatic rg function*. The molecular weight calculation is done using the *volume of correlation method*.

Note: See *below* for how to show different structural parameters on the SEC plot.

Data point browsing

The Frame number, integrated intensity, mean intensity, Rg, Rg error, I(0), I(0) error, MW, filename, and, if available intensity a q=<#> where <#> is a user selected value, for each individual point can be inspected using the data browser. To do so:

1. Right-click on the data item of interest.
2. Select "Show data" in the popup menu.

The SEC data item right click menu options

When you right click on a data item, a popup menu is shown. This section describes what each item on the menu does.

Remove

Removes the item.

Export data

Exports the SEC data item to a spreadsheet.

Save all profiles as .dats

Saves all of the scattering profiles loaded into the SEC item as individual .dat files.

Save

Saves the selected data item(s).

SVD

Opens the singular value decomposition (SVD) analysis panel for the selected SEC item.

EFA

Opens the evolving factor analysis (EFA) panel for the selected SEC item.

Show data

Shows the individual data points.

Rename

Renames the data item.

The SEC panel bottom buttons

There are three buttons at the bottom of the SEC control panel. They are:

Save

This button saves the selected data item(s).

Remove

This button removes the selected data item(s) from the SEC panel.

Clear SEC Data

This button clears all loaded SEC data. It works the same as if you had selected all of the SEC data items and then removed them.

The SEC Plot window

The SEC plot window has only one plot, however data can be plotted on both the left and right y axes. The left axis plots intensity from scattering profiles, while the right axis plots structural parameters (Rg, MW, I(0)). The items associated with the plotted curves are shown in the SEC control panel list.

The features that are general between all of the plots are described *elsewhere*. This section will describe features unique to this plot.

Changing axes and plot types

Right-click in the plot to view a pop-up menu with different axis settings. In this plot, you can change what data is plotted on each axis.

Y Data (Left Axis)

This allows you to change intensity plotted on the left y axis. The methods are:

- Integrated intensity (default), which is

$$I_{tot} = \int_{q_{min}}^{q_{max}} I(q)dq$$

- Mean intensity, which is the average intensity across the whole scattering profile
- Intensity at q=..., which allows the user to specify a q value, and displays the intensity at the nearest q point to that specified value.

Y Data (Right Axis)

This allows you to change which calculated structural parameter is plotted on the right y axis.

1. RG (default) which plots the Rg
2. MW which plots the molecular weight
3. I(0) which plots the I(0)
4. None, which turns off the right y axis.

X Data

This allows you to change what is plotted on the x axis.

- Frame Number, which plots the intensity and structural parameter as a function of frame number.
- Time, which puts the x axis in terms of experimental time.

Note: The time display on the x axis is only available for certain types of header files. Currently only G1 and G1 WAXS header files will have associated time values.

4.6.11 The Image plot panel

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

The Image plot panel displays any detector image that RAW is able to read. The image name is shown at the top of the plot, while the x and y axes are in dimensions of pixels. There is no right click menu for this plot, but scrolling with a mouse scroll wheel will zoom in and out. It is possible to display regions where there is no image data, these will show up as a white background (for example, if you zoom out and show negative pixel ranges.

In addition to the plot toolbar buttons *shared by all of the plots*, the Image plot has the following buttons:

Displays the image header. This does not display any information from separate header files, only what is in the image itself.

Opens the image display settings dialog described in below.

Shows the previous/next image in the same directory as the current image (as sorted by filename, A-Z).

The Image Display Settings dialog

The Image Display Settings dialog lets you change how the image is displayed on the screen. It has three sections, Image parameters, Image scaling, and Colormaps.

Image parameters

There are three slider bars here. The Upper and Lower limit sliders let you set what pixel values in the image correspond to the most extreme values of the color scale chosen. The Brightness color bar does not work at the moment. By default, the upper and lower limit values are automatically set each time an image is loaded into the Image plot panel. However, checking the Lock values checkbox will prevent the current set values from changing. This allows you to scroll through images maintaining constant upper and lower limits.

Image scaling

Image scaling allows you to use a linear or logarithmic (base 10) color scale. The logarithmic scale uses the matplotlib.colors.SymLogNorm function (more available here: http://matplotlib.org/api/colors_api.html), with a linear threshold set at 1. So the image is logarithmic in color display above 1, and linear in color display below 1 (this is so 0 and negative value pixels can be handled).

Colormaps

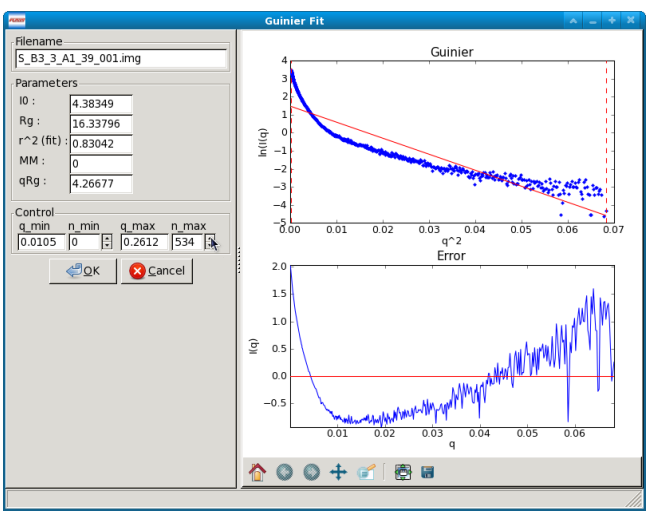
The colormaps option allows you to change the color scale used to display intensity values in the Image plot. The color maps are those available through matplotlib (http://matplotlib.org/users/colormaps.html), a limited selection of which are available in RAW: Gray, Heat, Rainbow, Jet (default), Spectral.

Note: Currently there is also a brightness setting, but it does not work properly.

4.6.12 Analysis windows

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

There are a variety of analysis windows in RAW: the Guinier Fit window, the Molecular Weight window, the BIFT window, the GNOM window, the AMBIMETER window, the DAMMIF window, the SVD window, and the EFA window. In this chapter we will cover how to open these windows and carry out the analysis.



Guinier Fit window

Data that has been properly background subtracted can be analyzed in the standard fashion via a Guinier Fit to yield the radius of gyration and the scattering intensity at zero angle. The Guinier Fit window allows you to interactively do the fit.

Opening the Guinier fit window

The Guinier fit window can be opened by selecting a data item in the Manipulation Panel and either right clicking and selecting "Guinier fit" or selecting "Guinier fit" from the "Tools" menu. Whichever data item is selected will be the one that is being analyzed. If multiple items are selected, the first (top item in the list) will be sent to the Guinier Fit panel.

Once the window opens you will see two panels. The left panel has the filename of the data loaded into the Guinier Fit panel, the fit parameters, and the fit controls. The right panel has two plots that show the fit and residual. If no previous Guinier fit analysis has been saved for the data item, the fit range defaults to the whole scattering profile. If a previous Guinier fit analysis was saved for the data item, the fit range will be the range of that analysis.

Once you have completed a satisfactory Guinier fit, clicking the "OK" button will save the fit for the data item selected. That will allow you to save the Guinier analysis parameters when *saving all results*, view the results in the information window, and save it with the scattering profile if that is saved as a ".dat" file. Clicking the "Cancel" button (or the close window button from the operating system, typically in the upper left or right corner) will exit the window without saving the analysis results.

Note 1: Only one Guinier Fit window can be open at once. If a window is already open and you try to open a new one, it will close the current window and open a new one.

Note 2: RAW will automatically attempt to automatically determine the Rg when the Guinier window is opened, if no previous Guinier fit has been done. If a previous Guinier fit has been done, RAW will set the q min and q max to the previous values.

Guinier Fit Controls

The Guinier Fit Controls allow you to manually set the q_min and q_max values over which the Guinier fit is done. You can do this in two ways. First, if you type a value into the q_min or q_max box and hit enter, RAW will find the nearest q point, and set the q value to that nearest point. Second, you can adjust the n_min and n_max values either by typing an integer value in the appropriate box or using the spin controls. The n_min and n_max controls change the minimum and maximum points of the q vector used (zero indexed), and the q_min and q_max are updated to match. For example if n_min is set to 5 and n_max is set to 130, the 5th indexed point through the 130th indexed point in the q vector (so points 6-131, because of the zero indexing) will be used in the Guinier fit.

The AutoRG button calls a function that attempts to automatically determine the best range of data to use for the Guinier fit. It considers the following criteria: q_min*Rg as small as possible, q_max*Rg as close to 1.3 as possible, smallest fit error in the Rg, smallest fit error in the I(0) possible, r^2 of the fit as close to one as possible, and the largest range of q values used as possible.

Guinier Fit Plots

There are two plots in the Guinier fit window. The top plot is the standard Guinier plot, ln(I(q)) vs. q^2, shown in the range of n_min -20 to n_max +3 (if possible) as set in the Guinier Control panel. The scattering profile is marked by the blue points, the fit is shown as the solid straight red line, the fit extrapolated to q=0 is shown by the green dashed line, and the range of the fit, q_min and q_max, are shown by vertical red dashed lines.

The bottom plot has the residual between each point in the scattering profile and the fit, plotted as a Δln(I(q)) vs. q^2. The red line is a line at zero. Note that the x range of this plot may be slightly reduced compared to the top plot, as it only displays the fit, while the top plot can display points beyond the limits of the fit.

Note: There are no controls available on right click for these plots, but there is a standard *plot control bar*.

Guinier Fit Parameters

The fit parameters available in the Guinier Fit window are:

- I(0), the extrapolated scattering intensity at zero angle
- Rg, the radius of gyration
- r^2 (fit), the r^2 value of the fit
- qRg, which is q_min*Rg and q_max*Rg for the left and right boxes respectively.

Molecular weight

Data that has been background subtracted and had a Guinier fit carried out can be analyzed to find the molecular weight of the macromolecule. RAW provides four different ways of calculating the molecular weight, which are described below.

Opening the Molecular Weight Window

The Molecular Weight window can be opened by selecting a data item in the Manipulation Panel and either right clicking and selecting "Molecular Weight" or selecting "Molecular Weight" from the "Tools" menu. Whichever data item is selected will be the one that is being analyzed. If multiple items are selected, the first (top item in the list) will be the one analyzed.

The window has two parts. The top part is a description of the MW methods and a panel with the Guinier fit parameters. The bottom part is four panels providing the calculated MW from each method, and the ability to see more details and learn more about each method. All of the MW methods require a Guinier fit to have been done, as they rely on the I(0) value. Two of the methods require knowing the sample concentration, and the same two methods also depend on (different) calibration of the scattering profile.

Once you are satisfied with the molecular weight analysis, clicking the "OK" button will save the analysis for the data item selected. That will allow you to save the molecular weight analysis parameters when *saving all results*, view the results in the information window, and save the results with the scattering profile if that is saved as a ".dat" file. Clicking the "Cancel" button (or the close window button from the operating system, typically in the upper left or right corner) will exit the window without saving the analysis results.

Note: Only one molecular weight window can be open at once. If a window is already open and you try to open a new one, it will close the current window and open a new one.

The molecular weight controls

There are several controls available in the molecular weight panel. In the top panel, the Guinier Fit button can be used to open the Guinier panel and (re)analyze the scattering profile using that method, so that an I(0) value is available for the molecular weight panel.

Each of the four lower panels has a "Show Details" and "More Info" button. The "More Info" button simply provides an extended description of the method used, and, when relevant, a citation. The "Show Details" button provides extended information on the parameters used to calculate the MW for a given method. None of these parameters are editable in the molecular weight panel. When the "Show Details" button is clicked, it becomes a "Hide Details" button, which hides the extra parameters.

The concentration box in both the first and fourth panels can have the sample concentration entered, if it has not already been (for example, in the Information panel). These are linked, so that any change to one also changes the other. Sample concentration should be in mg/ml.

The volume of correlation method uses different parameters for Proteins and RNA, and the drop down menu can be used to toggle between those two sets of parameters. The default setting can be changed in the Options window.

If RAW is set to normalize scattering profiles to an absolute scale, the "Intensity on Absolute Scale" box will be checked in the fourth panel. If RAW is not normalizing to an absolute scale, it will be unchecked. If that is set incorrectly for a particular scattering profile for some reason, it can be manually toggled to the correct position.

At the bottom of the window there are three buttons. The "OK" and "Cancel" buttons have been described above. The "Change Advanced Parameters" button opens the Options window, with the Molecular Weight section shown. This allows you to change the parameters used to calculate the molecular weight for each method of calculation.

The molecular weight parameters

There are a number of parameters listed in the molecular weight panel. In the top panel there are:

Filename

Gives the filename of the data item being analyzed.

Guinier parameters

Gives the I(0) and Rg from the Guinier fit.

In the I(0) Ref. MW panel there are:

Concentration

The sample concentration in mg/ml, this field accepts input.

MW

The sample MW calculated by this method, in kDa.

Ref. I(0)

The I(0) of the molecular weight standard set for RAW. This field corresponds to the I(0) field in the "Molecular Weight Estimation Using a Standard" box in the Molecular Weight section of the Options window, and can be changed there or by *setting the MW standard*.

Ref. MW

The molecular weight of the molecular weight standard set for RAW. This field corresponds to the MW field in the "Molecular Weight Estimation Using a Standard" box in the Molecular Weight section of the Options window, and can be changed there or by *setting the MW standard*.

Ref. Concentration

The concentration of the molecular weight standard set for RAW. This field corresponds to the Conc. field in the "Molecular Weight Estimation Using a Standard" box in the Molecular Weight section of the Options window, and can be changed there or by *setting the MW standard*.

File – The data item name of the molecular weight standard set for RAW. This field corresponds to the Filename field in the "Molecular Weight Estimation Using a Standard" box in the Molecular Weight section of the Options window, and can be changed there or by *setting the MW standard*.

In the Vc MW panel there are:

MW

The sample molecular weight calculated by this method, in kDa.

Vc

The volume of correlation calculated by this method, in $Å^2$.

Qr

The Qr parameter calculated by this method, in Å^3.

a

The macromolecular type (protein/RNA) dependent "a" parameter used for the calculation. This field corresponds to the "Protein (RNA) Coef. A" field in the "Molecular Weight Estimation From Volume of Correlation" box in the Molecular Weight section of the Options window, and can be changed there. The value depends on whether Protein or RNA is selected in the drop down menu at the top of this panel.

b

The macromolecular type (protein/RNA) dependent "b" parameter used for the calculation. This field corresponds to the "Protein (RNA) Coef. B" field in the "Molecular Weight Estimation From Volume of Correlation" box in the Molecular Weight section of the Options window, and can be changed there. The value depends on whether Protein or RNA is selected in the drop down menu at the top of this panel.

This panel also has a plot which shows $\int qI(q)dq$ vs. q, over the q-range of the scattering profile. For this method to be accurate, the integral value needs to have converged at high q (the graph needs to be flat at high q).

In the Vp MW panel there are:

MW

The sample molecular weight calculated by this method, in kDa.

Vp

The Porod volume calculated by direct integration of the scattering profile, in Å^3.

Corrected Vp

The corrected Porod volume based on the method described *below*, in Å^3.

Macromolecule Density

The density of the macromolecule, used to calculate the molecular weight. This field corresponds to the "Density" field in the "Molecular Weight Estimation from Corrected Porod Volume" box in the Molecular Weight section of the Options window, and can be changed there.

In the Abs. MW panel there are:

Concentration

The sample concentration in mg/ml, this field accepts input.

MW

The sample MW calculated by this method, in kDa.

 # electrons per mass dry macromolecule

The dry mass number density of electrons for the macromolecule, in e-/g. This field corresponds to the "Electrons per dry mass of macromolecule" field in the "Molecular Weight Estimation from Absolute Intensity Calibration" box in the Molecular Weight section of the Options window, and can be changed there.

 # electrons per volume of buffer

The number density of electrons for the protein buffer/solvent, in e-/cm^3. This field corresponds to the "Electrons per volume of aqueous solvent" field in the "Molecular Weight Estimation from Absolute Intensity Calibration" box in the Molecular Weight section of the Options window, and can be changed there.

Protein partial specific volume

The partial specific volume of the macromolecule, in cm^3/g. This field corresponds to the "Partial specific volume of the macromolecule" field in the "Molecular Weight Estimation from Absolute Intensity Calibration" box in the Molecular Weight section of the Options window, and can be changed there.

Scattering length of an electron

The scattering length of an electron in cm. This field corresponds to the "Scattering length of an electron" field in the "Molecular Weight Estimation from Absolute Intensity Calibration" box in the Molecular Weight section of the Options window, and can be changed there.

Scattering contrast per mass

The calculated scattering contrast per mass. This is calculated from the other parameters as $r_0(\rho_{Mmac} - \rho_{solv}\bar{\nu})$ where r_0 is the scattering length of an electron, ρ_{Mmac} is the electrons per dry mass of macromolecule, ρ_{solv} is the electrons per volume of aqueous solvent, and $\bar{\nu}$ is the partial specific volume of the protein.

The molecular weight methods

Four different methods are used to calculate the molecular weight of the macromolecule from the background subtracted scattering profile.

I(0) Referenced molecular weight calculation (I(0) Ref. MW panel)

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:

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

where MW is the molecular weight, c is the concentration, and the m and st subscripts denote 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, I(0) is not well estimated from the Guinier fit, or the contrast between the macromolecule and buffer is significantly different between the reference and sample.

Volume of correlation based molecular weight calculation (Vc MW panel)

This method uses the approach described in: Rambo, R. P. & Tainer, J. A. (2013). Nature. 496, 477-481. First, the volume of correlation, V_c, is calculated as

$$V_c = \frac{I(0)}{\int qI(q)dq}$$

Unlike the Porod volume, V_c is expected to converge for both compact and flexible macromolecules. Physically, V_c can be interpreted as the particle volume per self-correlation length, and has units of $Å^2$. V_c and the radius of gyration, Rg, are then used to calculate a parameter $Q_r = V_c^2/R_g$. The molecular weight is then calculated as:

$$MW = \left(\frac{Q_r}{b}\right)^a$$

where *a* and *b* are empirically determined constants that depend upon the type of macromolecule. More details on the calculation are in the reference. The authors claim the error in MW determination is ~5-10%.

This method can yield inaccurate results if the integral $\int qI(q)dq$ doesn't converge, which may indicate the scattering profile is not measured to high enough q or that there is a bad buffer match. It also requires accurate determination of I(0) and Rg. It doesn't work for protein-nucleic acid complexes.

Corrected Porod Volume method (Vp MW panel)

This method uses the approach described in: Fischer, H., de Oliveira Neto, M., Napolitano, H. B., Polikarpov, I., & Craievich, A. F. (2009). J. Appl. Crystallogr. 43, 101-109. First, the Porod volume, V_p, is determined. True determination of the Porod volume requires the scattering profile measured to infinite q. A correction is applied to V_p to account for the limited range of the measurement. The authors report a maximum of 10% uncertainty for calculated molecular weight from globular proteins.

This method can yield inaccurate results if the molecule is not globular. It requires accurate determination of I(0). It also requires an accurate protein density. It only works for proteins.

Note: To do the integration, RAW extrapolates the scattering profile to I(0) using the Guinier fit. The authors of the original paper used smoothed and extrapolated scattering profiles generated by GNOM. This may cause discrepancy. To use this method on GNOM profiles, use the online SAXS MoW calculator located at: http://www.if.sc.usp.br/~saxs/

Absolute calibrated intensity method (Abs. MW panel)

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:

$$MW = \frac{N_A I(0)/c}{\Delta\rho_M^2}$$

where N_A is the Avagadro number, c is the concentration, and $\Delta\rho_M^2$ is the scattering contrast per mass described *above*. 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 off, or if the partial specific volume of the macromolecule in solution is incorrect. I(0) and the concentration in solution must be well determined. Unless the scattering contrast is adjusted, this method will only work for proteins.

BIFT

The BIFT window allows you to run a Bayesian Indirect Fourier Transform (BIFT) method on background subtracted scattering profiles to find the P(r) function. The advantage to this method over the method implemented by GNOM is that once the search parameters are set, there is no subjective input required from the user, a single "best" solution is provided by the algorithm. The BIFT algorithm being used to find the P(r) is that of: Hansen, S. (2000). J. Appl. Crystallogr. 33, 1415-1421.

Opening the BIFT Window

The BIFT window can be opened by selecting a data item in the Manipulation Panel and either right clicking and selecting "BIFT" or selecting "BIFT" from the "Tools" menu. Whichever data item is selected will be the one that is being analyzed. If multiple items are selected, the first (top item in the list) will be the one analyzed.

The window has two parts. The left part shows the file being analyzed, and the controls, parameter outputs, and status for the BIFT. The right is two plots, the top showing the P(r) function found by the BIFT and the bottom showing the experimental data and the scattering profile generated from the P(r) function. The space allotted to each side can be adjusted by clicking and dragging the separator bar. The whole window can be resized by clicking and dragging an edge or corner.

When the window is first opened, it runs a BIFT analysis to find the P(r) function of the data, based on the current settings. These settings can be altered from the *BIFT panel* in the Options window.

Once you are satisfied with the BIFT results, clicking the "OK" button will save the Dmax, real space Rg, real space I(0), the χ^2 for the fit, and $\log_{10}\alpha$ for the data item selected. That will allow you to save the BIFT analysis parameters

when *saving all results* and save the results with the scattering profile if that is saved as a ".dat" file. Additionally, a new IFT data item will be created, which will be shown in the *IFT Control and Plot panels*. Clicking the "Cancel" button (or the close window button from the operating system, typically in the upper left or right corner) will exit the window without saving the analysis results or new IFT item. IFT items created by BIFT will have an extension ".ift".

Note: Only one BIFT window can be open at once. If a window is already open and you try to open a new one, it will close the current window and open the new one.

The BIFT Controls

A BIFT analysis of the scattering profile is automatically run when the window is opened. BIFT has very few controls available to the user, though settings can be customized. The controls consist of three buttons:

Run

Reruns the BIFT analysis. Needs to be done if the settings are changed after the BIFT panel is opened.

Abort

Aborts the BIFT analysis if it is currently running.

Settings

Opens the Options window and shows the settings for BIFT. If settings are changed, the Run button must be used to generate a new P(r) function with the changed settings.

The BIFT parameters

The BIFT panel displays the following parameters:

Dmax

The maximum dimension of the P(r) function found by the BIFT algorithm. This is in units of 1/q, which RAW assumes to be Å.

Log(Alpha)

The log base 10 of the alpha value found as optimal by the BIFT search.

Rg (A)

The radius of gyration in Å (assumed, actual units of 1/q). This is shown from the Guinier fit (if available) and the P(r) function. The value from the P(r) function is the value calculated in real space by

$$R_g = \frac{\int_0^{D_{max}} r^2 P(r) dr}{2 \int_0^{D_{max}} P(r) dr}$$

I(0)

The scattering at zero angle. This is shown from the Guinier fit (if available) and the P(r) function. The value from the P(r) function is the value calculated in real space by

$$I(0) = 4\pi \int_0^{D_{max}} P(r) dr$$

chi^2 (fit)

The χ^2 value of the scattering profile from the P(r) function to the experimental data.

The BIFT status

The status box for the BIFT search shows parameters that update as the BIFT search is performed. Once the search is over, they show the parameters of the final solution. The status items displayed are:

Status

An overall status, which can be: Performing search grid, Performing Fine Search, BIFT done, or BIFT canceled.

Evidence

The evidence value for a given search point.

Chi

The χ^2 value of a given search point.

Alpha

The log base 10 of the alpha value of a given search point.

Dmax

The maximum dimension of the current search point.

Current Search Point

The current search point (numbered along the search grid, essentially arbitrary).

Total Search Points

The total number of search points, equal to the number of Dmax search points multiplied by the number of alpha search points.

Note: The status window does not update the evidence, chi, alpha, or dmax values during the fine search, only at the end of the fine search.

The BIFT plots

There are two plots in the BIFT window. The top plot shows the P(r) function in red. The units for the r (bottom) axis of this plot are 1/q, which RAW assumes to be Å. A black line is displayed at zero on the plot for reference. The bottom plot shows the measured scattering profile data as blue points, and the scattering profile generated from the P(r) function in red.

Note: There are no controls available on right click for these plots, but there is a standard *plot control bar*.

The BIFT algorithm

The algorithm used is described in Hansen, S. (2000). J. Appl. Crystallogr. 33, 1415-1421. In RAW, a coarse grid is used for an initial search, and then a fine optimization is performed from the best point in that search space. The limits of the coarse grid and the number of points in the coarse grid can be set in the Options window.

GNOM (ATSAS)

RAW allows you to run certain analyses using the ATSAS software package from within RAW. One of the programs that can be run from RAW is GNOM, which performs an IFT to find the P(r) function. Using the ATSAS package programs requires a *separate installation* and (possibly) some additional configuration of RAW.

Opening the GNOM Window

The GNOM window can be opened by selecting a data item in the Manipulation Panel and either right clicking and selecting "GNOM (ATSAS)" or selecting "GNOM" from the "Tools"->"ATSAS" menu. Whichever data item is selected will be the one that is being analyzed. If multiple items are selected, the first (top item in the list) will be the one analyzed.

The window has two parts. The left part shows the filename being analyzed, and the controls and parameter outputs for GNOM. The right part has two plots, the top showing the P(r) function found by the GNOM, and the bottom showing the experimental data and the scattering profile generated from the P(r) function. The space allotted to each side can be adjusted by clicking and dragging the separator bar. The whole window can be resized by clicking and dragging an edge or corner.

When the window is first opened, if no previous GNOM analysis is available for the data item, RAW runs DATGNOM from the ATSAS package analysis to find a P(r) function of the data. Generally, better results are obtained from DATGNOM when an Rg value is available from the Guinier fit. If GNOM analysis has previously be done on the data item, RAW will display the P(r) function corresponding to the Dmax value found by that analysis.

Once you are satisfied with the GNOM results, clicking the "OK" button will save the Dmax, Total Estimate, real space Rg, real space I(0), and the starting and ending q values for the data item selected. That will allow you to save the GNOM analysis parameters when *saving all results* and save the results with the scattering profile if that is saved as a ".dat" file. Additionally, a new IFT data item will be created, which will be shown in the *IFT Control and Plot panels*. Clicking the "Cancel" button (or the close window button from the operating system, typically in the upper left or right corner) will exit the window without saving the analysis results or new IFT item. IFT items created by GNOM will have an extension ".out".

Note 1: Only one GNOM window can be open at once. If a window is already open and you try to open a new one, it will close the current window and open the new one.

Note 2: If the GNOM option is unavailable in the right click menu for a data item, it indicates that RAW is unable to find the *ATSAS programs* on your computer.

The GNOM Controls

The GNOM Controls allow you to manually set the q_min and q_max values GNOM uses. You can do this in two ways. First, if you type a value into the q_min or q_max box and hit enter, RAW will find the nearest q point, and set the q value to that nearest point. Second, you can adjust the n_min and n_max values either by typing an integer value in the appropriate box or using the spin controls. The n_min and n_max controls change the minimum and maximum points of the q vector used (zero indexed), and the q_min and q_max are updated to match. For example if n_min is set to 5 and n_max is set to 130, the 5th indexed point through the 130th indexed point in the q vector (so points 6-131, because of the zero indexing) will be used by GNOM. The Dmax value can be either by typing an integer value in the appropriate box or using the spin controls. Changing any of these values (q_min, q_max, and Dmax) automatically updates the P(r) function.

The "Change Advanced Parameters" button opens the Options panel and shows the options for GNOM. This allows you to change advanced parameters for your GNOM analysis.

The "DATGNOM" button runs the DATGNOM program from the ATSAS software package. The Dmax it finds is rounded to the nearest integer, and GNOM is run with that Dmax value to generate the P(r) function.

The GNOM parameters

The GNOM panel displays the following parameters from the GNOM fit:

Rg (A)

The radius of gyration in Å (assumed, actual units of 1/q). This is shown from the Guinier fit (if available) and the P(r) function. The value from the P(r) function is the value calculated in real space by

$$R_g = \frac{\int_0^{D_{max}} r^2 P(r) dr}{2 \int_0^{D_{max}} P(r) dr}$$

I(0)

The scattering at zero angle. This is shown from the Guinier fit (if available) and the P(r) function. The value from the P(r) function is the value calculated in real space by

$$I(0) = 4\pi \int_0^{D_{max}} P(r) dr$$

Total Estimate

The "Total Estimate" produced by GNOM. A value close to 1 is good.

chi^2 (fit)

The χ^2 value of the scattering profile from the P(r) function to the experimental data.

GNOM says

The subject interpretation of the quality of the P(r) function provided by GNOM.

The GNOM plots

There are two plots in the GNOM window. The top plot shows the P(r) function in red. The units for the r (bottom) axis of this plot are 1/q, which RAW assumes to be Å. A black line is displayed at zero on the plot for reference. The bottom plot shows the measured scattering profile data as blue points, and the scattering profile generated from the P(r) function in red.

Note: There are no controls available on right click for these plots, but there is a standard *plot control bar*.

AMBIMETER (ATSAS)

RAW allows you to run certain analyses using the ATSAS software package from within RAW. One of the programs that can be run from RAW is AMBIMETER, which provides an estimate of the ambiguity a 3D shape reconstruction will have, based on the scattering profile generated from the P(r) function. Using the ATSAS package programs requires equires a *separate installation* and (possibly) some additional configuration of RAW.

Opening the AMBIMETER Window

The AMBIMETER window can be opened by selecting a data item in the IFT Panel and either right clicking and selecting "Run AMBIMETER" or selecting "AMBIMETER" from the "Tools" -> "ATSAS" menu. Whichever data item is selected will be the one that is being analyzed. If multiple items are selected, the first (top item in the list) will be the one analyzed. Currently, AMBIMETER only works on IFT items generated by GNOM (".out" files in the IFT panel). The AMBIMETER window shows the file it is being run on, the Rg (real space form the P(r) function), and controls and results.

When the window is first opened, AMBIMETER is run on the data. Once you are satisfied with the GNOM results, clicking the "OK" or "Cancel" will close the window. Because of the strict save format required for ".out" files to be used by the ATSAS package, the AMBIMETER results are not saved anywhere, and must be manually saved by the user (such as writing it down).

Note 1: Only one AMBIMETER window can be open at once. If a window is already open and you try to open a new one, it will close the current window and open the new one.

Note 2: If the AMBIMETER option is unavailable in the right click menu for a data item, it indicates that RAW is unable to find the *ATSAS programs* on your computer or that you do not have a recent enough version of the ATSAS package installed (version 2.7.1 or greater required for AMBIMETER).

The AMBIMETER Controls

The AMBIMETER controls allow you to adjust the maximum q value used by AMBIMETER, by adjusting the upper q*Rg limit between 3 and 7. Note that if the maximum q value of the scattering profile times the Rg is less than the limit set, the whole curve is used.

The AMBIMETER program can also save output shapes. For more information about this, see the AMBIMETER manual available with the ATSAS software. In the window, you can select which shapes to save, None (default), Best (one shape, the best fit), or All (all of the shapes that fit). If you are saving shapes, you should then select the output directory to save them in, and provide an output prefix. The shapes will with the prefix value provided in the Output prefix box, as described in the AMBIMETER manual. Clicking the Run button is necessary to rerun AMBIMETER after any settings have been changed.

The AMBIMETER Results

The results section shows the output from AMBIMETER. It reports:

Number of compatible shape categories

The number of compatible shape categories, as described in the AMBIMETER manual.

Ambiguity score

Log base 10 of the number of compatible shape categories.

AMBIMETER says

The subjective interpretation of the ambiguity score provided by AMBIMETER.

DAMMIF (ATSAS)

RAW allows you to run certain analyses using the ATSAS software package from within RAW. One of the programs that can be run from RAW is DAMMIF and the accompanying programs DAMAVER and DAMCLUST, which carry out 3D shape reconstructions based on the P(r) function and scattering profile. Using the ATSAS programs requires a *separate installation* and (possibly) some additional configuration of RAW.

Opening the DAMMIF Window

The DAMMIF window can be opened by selecting a data item in the IFT Panel and either right clicking and selecting "Run DAMMIF" or selecting "DAMMIF" from the "Tools"->"ATSAS" menu. Whichever data item is selected will be the one that is being analyzed. If multiple items are selected, the first (top item in the list) will be the one analyzed. Currently, DAMMIF only works on IFT items generated by GNOM (".out" files in the IFT panel).

The DAMMIF window shows current DAMMIF settings, controls, the log output for each separate DAMMIF and DAMAVER run, and the overall status of the processing. The "Close" button closes the window. If this is done before the DAMMIF processing is finished, it will abort the processing.

Note 1: Only one DAMMIF window can be open at once. If a window is already open and you try to open a new one, it will close the current window and open the new one.

Note 2: If the DAMMIF option is unavailable in the right click menu for a data item, it indicates that RAW is unable to find the *ATSAS programs* on your computer.

Note 3: The DAMMIF processing can be run in the background while further data processing is done in RAW.

The DAMMIF settings

The DAMMIF window settings section allows you to change the most commonly used DAMMIF settings. The "Change Advanced Settings" button allows you to change all of the advanced settings of DAMMIF.

Output directory

This sets the output directory for DAMMIF and DAMAVER results, and can be set either by typing a directory into the box and hitting enter, or using the Select/Change Directory button. The directory defaults to the directory showing the Files tab.

Output prefix

The output prefix will be appended to each DAMMIF and DAMAVER file. It defaults to the filename being analyzed. This should contain no spaces.

Number of reconstructions

The total number of DAMMIF models to construct.

Number of simultaneous runs

The number of DAMMIF models to reconstruct simultaneously. DAMMIF runs on a single core, and typically takes 100% of the resources of that core. On multicore machines, the number of simultaneous runs can be set between 1 and the total number of cores. The default is the number of cores minus one.

Mode

Sets the Mode used by DAMMIF for the reconstructions. The "Fast" and "Slow" modes are as described in the DAMMIF manual. The "Custom" mode is equivalent to the "Interactive" mode for DAMMIF, and allows the use of the more advanced settings, as set in the Options panel in the "DAMMIF Advanced" section.

Symmetry

Allows the symmetry to be set, if known, as described in the DAMMIF manual.

Anisometry

Allows the anisometry to be set, if known, as described in the DAMMIF manual.

Change Advanced Settings

This opens the Options panel and shows the options for DAMMIF. This allows you to change advanced parameters for your DAMMIF analysis.

Align and average envelopes (damaver)

If this is selected and two or more reconstructions are generated using DAMMIF, then once all reconstructions are finished DAMAVER is automatically run on the DAMMIF reconstructions. This runs in a mode equivalent to "damaver –a" at the command line.

Align and cluster envelopes (damclust)

If this is selected and two or more reconstructions are generated using DAMMIF, then once all reconstructions are finished DAMCLUST is automatically run on the DAMMIF reconstructions.

Note: The damaver and damclust options are mutually exclusive, you can select DAMAVER or DAMCLUST but not both.

The DAMMIF controls

There are only two control buttons, "Start", which starts the DAMMIF reconstructions, and "Abort", which aborts the DAMMIF reconstructions. Start is only available if DAMMIF reconstructions are not currently running. Abort is only available if DAMMIF reconstructions are running.

Note: DAMMIF requires that the P(r) function be written to disk as a ".out" file. RAW will check whether there is an existing ".out" file with the same name that will be overwritten before running. It will also check whether the DAMMIF files generated will overwrite any existing files. In either case, it will provide a warning to let you know that is happening.

Note 2: DAMMIF gets the starting random seed value from the computer clock time in seconds. In order to produce different reconstructions, the start of each reconstruction must occur at least 1s after the previous one. This introduces a noticeable delay when starting up a large number simultaneous reconstructions.

The DAMMIF status

The status panel provides an overview of the current status of the DAMMIF runs. It updates with the following status:

Starting Processing

Indicates that the initial processing has started.

Starting DAMMIF run <#>

Here the <#> corresponds to the run number (1 up to the total number of reconstructions), and this corresponds to the numbers in the Log panel. This indicates that the given DAMMIF run has started.

Finished DAMMIF run <#>

Here the <#> corresponds to the run number (1 up to the total number of reconstructions), and this corresponds to the numbers in the Log panel. This indicates that the given DAMMIF run has finished.

Starting DAMVER

Indicates that the reconstructions are now being aligned and averaged by DAMAVER.

Finished DAMAVER

Indicates that the reconstructions have finished being aligned and averaged by DAMAVER.

Finished Processing

Indicates that all processing has finished.

Processing Aborted!

Indicates that the processing was aborted before everything finished.

If necessary, the status window is scrollable.

The DAMMIF log

The log area provides details of the current and finished DAMMIF and DAMAVER processing. This is the output that would be displayed in the console window if DAMMIF or DAMAVER were run from the command line.

When the DAMMIF window is first opened, the log window will be empty. Once DAMMIF processing is started, a set of different panels accessible via the tabs at the top will be opened. Any tab with a number corresponds with a DAMMIF run (output from DAMMIF run 1 will display in tab "1", and so on). The runs are simply numbered sequentially starting with 1 and ending with the total number of reconstructions. If DAMAVER will also be run, the last tab is "Damaver", which shows the output of that processing.

Before a given run is started, the window associated with the tab will be empty. Once it starts, it will be updated with the output from DAMMIF or DAMAVER that would normally display in the console. Once it finishes, it will no longer update, but the log can still be viewed. The long windows are scrollable. You can change which output you are viewing by clicking on the tabs.

Note: The output in the log window is not saved, however by default DAMMIF writes this output to a log file, and the DAMAVER output is available in a set of output files. See the ATSAS manuals for each program for more details.

SVD

Singular value decomposition (SVD) is a mathematical technique that is a model independent approach that provides information on the number of unique elements in a data set. Formally, singular value decomposition of a m x n matrix M is a factorization of into three matrices such that

$$M = U\Sigma V^*$$

where U is an m x m unitary matrix, called the left singular values; Σ is a diagonal m x n matrix, where the diagonal values are the singular values, and V^* is the conjugate transpose of an n x n unitary matrix V, the right singular vectors. A typical interpretation of singular value decompositions is that the number of singular values significantly above the baseline level represents the number of significant distinct components in the data set.

RAW allows the user to use either scattering profiles or P(r) functions as the data set. This is typically applied to scattering profiles in a SEC-SAXS data set, and the number of significant singular values corresponds to the number of distinct scatterers in the data set. For SVD done across a single well-separated peak from the chromatograph, there would be two significant components: one from the buffer and one from the protein. For SVD done on a poorly separated monomer-dimer peak, there would be three significant components: buffer, monomer, and dimer. RAW allows users to select a range of scattering profiles for SVD, and displays the singular values σ_i and the autocorrelation of the left and right singular vectors R_i for each ith singular value defined as

$$R_i = \Sigma_n X_{i,n} X_{i,n+1} \tag{4.1}$$

where X is the U or V singular vector matrix.

Opening the SVD Window

The SVD window can be opened by selecting a single data item in the SEC Panel or two or more data items in the Manipulation or IFT Panels and either right clicking and selecting "SVD" or selecting "SVD" from the "Tools" menu.

The SVD window shows current SVD settings, controls, and results. The "Okay" button closes the window and saves the parameters used, the "Cancel" button closes the window and saves nothing.

Note: Only one SVD window can be open at once. If a window is already open and you try to open a new one, it will close the current window and open the new one.

The SVD Controls

The SVD controls give the user (up to) three items to control.

Use

This option allows the user to select whether they use Unsubtracted or Subtracted data. Note that this choice is only available for SEC curves, where both unsubtracted and subtracted data can exist within the same data item. Subtracted data can only be selected if the SEC curve has had *structural parameters* calculated for it.

Use Frames

This option changes the range of frames used for the SVD. If a SEC data item is selected, the frame number corresponds to the frame number in the SEC Plot. If Manipulation or IFT data items are selected in, the frame number is the same as how it would be displayed if the data were plotted on the *SEC plot*. The plot in the controls window shows the SEC curve of the selected. The red points correspond to all of the data in the data set, the blue points correspond to the data being used for SVD.

Normalize by uncertainty

If this option is selected, the intensity at a given q (SEC and Manipulation data) or r (P(r) data) value is divided by the average uncertainty across all frames at that q or r value. When doing EFA analysis, the data is normalized in this way, so this option allows corresponding SVD analysis.

The SVD Results

The SVD results are plotted in the two plots on the right side of the SVD panel. The top plot shows the singular values as a function of index (by default singular values are ordered from largest to smallest). The bottom plot shows the autocorrelation, defined in equation (4.1), of the left (red) and right (blue) singular vectors as a function of index.

The results control panel allows you to control which singular value indices are plotted. It has two further controls:

Save Plotted Values

This saves the singular values and autocorrelation values as a function of index that are plotted in the plots on the right. The data is saved as a comma separated value (.csv) file.

Save All

This saves all of the singular value information, which is to say, the full U, Σ, and V^* matrices. It also saves the autocorrelation values for all indices. The data is saved as a comma separated value (.csv) file.

EFA

Evolving factor analysis (EFA) is a model independent approach that extends SVD to allow separation of scattering profiles from mixed solutions, particularly overlapping chromatographic peaks from different species. This method was recently applied to SEC-SAXS data (see: Meisburger, S. P., Taylor, A. B., Khan, C. A., Zhang, S., Fitzpatrick, P. F., & Ando, N. (2016). J. Am. Chem. Soc. jacs.6b01563). An improved version of the method described by Meisburger et al. has been implemented in RAW. EFA in RAW starts with SVD, then proceeds by finding the component start and end points in the EFA plot and then rotating the significant singular value vectors into real scattering profiles. RAW implements two new methods for rotation of the singular vectors besides the iterative approached described by Meisburger et al. The first new method is the explicit calculation method described in: Maeder, M. (1987). Anal. Chem. 59, 530–533, while the second is a hybrid method that uses the explicit calculation as a seed for the iterative approach, allowing much faster convergence of the rotation.

Because EFA analysis is relative complex, it consists of three distinct screens in RAW, which are reached by clicking the "Next" and "Back" buttons in the EFA window. These screens step you through the EFA analysis. The window can be closed at any time with the "Cancel" button. If the EFA analysis succeeds, the window can be closed with the "Done" button, which sends the extracted scattering profiles to the Main plot, and saves information about the parameters used in the SEC item (if a SEC item was selected).

Opening the EFA Window

The EFA window can be opened by selecting a single data item in the SEC Panel or two or more data items in the Manipulation or IFT Panels and either right clicking and selecting "EFA" or selecting "EFA" from the "Tools" menu.

The EFA window shows current EFA settings, controls, and results. The "Okay" button closes the window and saves the parameters used, the "Cancel" button closes the window and saves nothing.

Note: Only one EFA window can be open at once. If a window is already open and you try to open a new one, it will close the current window and open the new one.

EFA Page 1 – SVD

The first EFA page essentially reproduces the SVD window described *above*. The difference is that there is a panel for User Input, "# Significant SVs" where the user inputs the number of significant singular values/vectors in the data set. RAW will automatically attempt to determine that, but the user can adjust it. If the user changes any of the SVD controls, RAW will not refine the guess for the number of significant singular values.

Once the user is happy with the data range and has determined number of significant singular vectors in the data, they click the "Next" button to move to Page 2.

EFA Page 2 – Evolving Factors

The second EFA page presents the results of the Evolving Factor Analysis. The details are described in the above referenced papers. In short, for Forward EFA, SVD is done on pieces of the data set, starting with just the first two frames, then the first three frames, and so on, until the entire data set is used. The singular values are then plotted as a function of final frame index used in that particular SVD, the top plot in this second page. This lets the user determine where certain components start in the data set, when there is a strong increase in a singular value is when a component starts in the data set. For Backward EFA, the same thing is done, except using the last two frames, then the last three frames, etc. The Backward EFA plot, the bottom plot on this page, shows the user where components exit the data set. Both the Forward and Backward EFA plots show one more value than the user marked as significant in Page 1, so that the user can judge where the value diverges from the baseline.

The User Input panel on the left side of the page consists of "Forward" and "Backward" sections each with a number of values equal to the user input of significant singular values from Page 1. These controls allow the user to set where each value diverges from the baseline.

Once the starting points for the values are set for both Forward and Backward EFA, clicking the "Next" button will take the user to Page 3.

Clicking the "Back" button allows the user to go back to Page 1 and adjust the results there.

EFA Page 3 – Rotation

The third EFA page allows the user to control the rotation of the singular vectors into scattering profiles and view the result of that rotation. There are several different control boxes:

Component Range Controls

These allow the user to adjust the ranges of the components in the scattering profile. There are a number of controls equal to the number of significant singular vectors selected on Page 1. Each control allows you to set the start and end point of the range, and each control has a "C>=0" checkbox. When checked, that forces the concentration profile for that component to be never negative (a physical constraint that can help with rotation by may hide mathematical errors). When the ranges are adjusted, the plot above this control is updated. The arrows and dashed lines of this plot

shown the component ranges, with colors that correspond to the colors in the legend of the Scattering Profiles plot in the top right.

Rotation Controls

Method

This controls the rotation method. There are three choices. Iterative applies the approach of Meisburger et al. (2016). Explicit applies the approach of Maeder (1987). Hybrid uses the explicit approach to seed the iterative approach.

Number of iterations

This can only be set for the Hybrid and Iterative approaches, and sets the maximum number of iterations the algorithm will attempt. If convergence is not reached by the end of these iterations, the method will fail.

Convergence Threshold

This sets the threshold at which the iterative and hybrid solutions are said to be converged. This threshold is defined in Meisburger et al. (2016), and for the nth rotation is the absolute difference between the concentration profiles of the n-1 and n solutions.

Status

The status box displays the status of the rotation. If the rotation has succeeded it will say "Rotation Successful." If the rotation is in progress it will say "Rotation in progress." If the rotation has failed it will provide a message with some on the failure, that starts with "Rotation failed."

Results

Save EFA Data (not profiles)

This button saves the EFA data, including the SVD data, the number of significant values, the ranges Forward and Backward EFA data, the ranges and concentration constraints for each component, the rotation method and other parameters used, and the mean error weighted χ^2 and concentration data. This is all saved as a .csv file.

In addition to the control boxes, there are also three plots. The top plot shows the scattering profiles obtained via the rotation (if the rotation is successful). The bottom plots show: Left – the mean error weighted χ^2 as a function of frame, which is a measure difference between the scattering profiles in a given frame as measured and as produced from the rotated scattering profiles and concentration profiles; Right – the concentration profiles for each component, which are color as the scattering profile (top) plot. The concentration profiles are normalized to an area of 1, and so are on an arbitrary scale.

4.6.13 Menus

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

RAW has 5 program wide menus: File, Options, View, Tools, and Help.

The File menu

The File menu has five choices:

Load Settings

This loads RAW settings from a .cfg file. This can also be done by double clicking on the .cfg file in the Files tab of the Control Panel.

Save Settings

This saves all of the RAW settings, including image calibration and the image mask, to a .cfg file so they can be loaded again later.

Load Workspace

This loads a RAW *workspace*.

Save Workspace

This saves a RAW *workspace*.

Quit

This quits the RAW program. If you have unsaved data or are running DAMMIF reconstructions you will be asked to confirm that you wish to quite.

The Options menu

The options menu has one choice and one submenu. The "Advanced Options" choice opens the *Options window*. The submenu is labeled "Online mode" and the three options available there are described *here*.

The View menu

The view menu has three top level options and five submenus. The three top level options are: "Show image", "Show data", and "Show header". All of these require that a single manipulation item be selected, and otherwise behave as the options of the same name in the Manipulation data item *right click menu*.

The view menus named "Top Main Plot Axis" and "Bottom Main Plot Axis" change the plot axes/scale of the top and bottom main plots respective. Each option in the two menus corresponds to the options in the main plot *right click menu*.

The view menus named "SEC Plot Left Y Axis", "SEC Plot Right Y Axis", and "SEC Plot X Axis" change the data plotted on the SEC Plot axes as the options in the SEC plot *right click menu*.

The Tools menu

The Tools menu is split into three parts. In all of the items in the top part of the menu apply to Manipulation data items and all of them require one or more Manipulation data items to be selected. The first submenu is "Operations". This contains a set of operations that can be run on Manipulation data items, and have identical effects to items of the same name in the manipulation *right click menu*. The second submenu is "Convert q-scale" and the options in the menu have identical effects to items of the same name in the Manipulation data items, and have identical effects to items of the same name in the manipulation *right click menu*. The third item is "Use as MW Standard" and has an identical effect as the item of the same name in the manipulation *right click menu*.

The middle part of the Tools menu consists of analysis tools. The "Guinier Fit", "Molecular Weight" and "BIFT" options, and the "ATSAS" submenu option "GNOM" all require a single Manipulation data item to be selected. These all open the corresponding *analysis windows*. The "DAMMIF" and "AMBIMETER" options require a single IFT data item to be selected, and open the corresponding *analysis windows*. The "SVD" and "EFA" require either that multiple Manipulation or IFT data items be selected for that a single SEC data item be selected, and open the corresponding *analysis windows*.

The lower part of the Tools menu consists of two options. The "Centering/Calibration" option requires that an appropriate image for centering/calibration be loaded in the Image plot panel. It then opens the *Centering/Calibration panel*. The "Masking" option requires that an appropriate image for masking be loaded in the Image plot panel. It then opens the *Masking panel*.

The Help menu

The Help menu has two options on it. The "Help!" option shows a window describing how to find help for RAW (including a reference to this document). The "About" provides a very brief description of RAW, includes the RAW citation, provides the license agreement, and lists the developers.

4.6.14 Line properties dialog

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

The Manipulation, IFT, and SEC data items can all open a line properties dialog. This dialog will let you control the following properties for each line associated with the object:

Line: Style (solid, dashed, dotted), width, color

Error bars: Style (solid, dashed, dotted), width, color

Data point marker: Shape (the marker shapes are those used by matplotlib, which are defined here: http://matplotlib.org/api/markers_api.html), size, edge line color, fill color, and whether or not the markers are hollow.

The line properties dialog also lets you set the legend label for the object, if any is needed.

If the dialog is exited with "OK", all changes take effect. If the dialog is exited with "Cancel", the old settings are restored.

4.6.15 The Information panel

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

The information panel displays information about a selected data item in the Manipulation panel. It shows the name of the item at the top. It also shows the following analysis, if available, the Rg in Å, I(0), and MW in kDa (from reference to a set MW standard). There is a box into which sample concentration can be entered in mg/mL. Notes or a description of the sample can also be entered, and will be saved as part of the "header" when a .dat file is saved. Finally, the drop down list in the header browser can be used to display and quickly view any header value available for the data item.

If the current Control panel is changed from the Manipulation panel to the IFT or SEC panel, the Information panel is cleared. If the panel is changed back to the Manipulation panel, the data in the Information panel are restored.

4.6.16 Workspaces

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

RAW allows you to save and load sets of files that you are analyzing, called "workspaces". To save a workspace, go the File menu and select "Save Workspace". To load a workspace, go to the File menu and select "Load Workspace".

When you save a workspace, all of the items in the Manipulation, IFT, and SEC panel are saved. When you load a workspace, all of those saved items are loaded back up, with all of the analysis information and line settings they had previously. Any files already open in RAW remain open when a workspace is loaded.

Note: The settings are not saved when a workspace is saved.

4.6.17 Online mode

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

RAW has an online mode that will watch a folder, and automatically load any files placed in the folder that RAW can read. This is typically used to automatically load in data as it is being collected at the beamline. This is distinct from the *automatic update feature* of the SEC Control panel, and both features can be active at once.

Turning on/off online mode

In order to turn on online mode, go to the Options->Online Mode menu and click the "Online" option. When you go online, a file browser will open and you will pick a folder. This folder (and only this folder, online mode does not recurse into subfolders) will be monitored by RAW. Every time a file is added (or updated) in the folder, RAW checks to see whether it can load the file. If it can, it will be loaded into RAW (for example, detector images will be integrated and loaded into the Manipulation panel and Main plot).

To disable online mode, go to the Options->Online Mode menu and click the "Offline" option.

Note: The online status of RAW is indicated in the status bar at the bottom of the main RAW window. On the right side, it says "Mode: <status>" where <status> is OFFLINE or ONLINE. This is only toggled by this online mode, not the automatic update feature of the SEC panel.

Changing directories in online mode

If you wish to change the directory that RAW is monitoring in online mode, go to the Options->Online Mode menu and click the "Change Directory" option. A file browser will appear, and you will use that to select the new folder for the Online mode.

Online mode filter

In some cases, you may not want every file that enters the directory (that can be read by RAW) to be loaded into RAW. In this case, it is possible to set a filter so that only certainly files can be loaded into RAW. The filter can consist of any number of individual items, which are set as follows. First, chose whether an item is include (open only if) or exclude (don't open if). Then provide a string for RAW to match in the file name. Then specify where in the filename RAW should match the string (start, end, anywhere). This filter can be set in the *options window*.

4.6.18 File types

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

Output file types

RAW outputs the following file types:

.cfg – These are files containing all of RAW's settings (configuration files). This file type is not human readable.

.dat – These are files containing *three column scattering profile data* with a "header" at the start or end of the file for additional information. This file type is human readable.

.ift – These are files containing *three column BIFT data* followed by four column scattering profile data, with a "header" at the start or end of the file for additional information. This file type is human readable.

.msk – These are files containing a RAW mask. This file type is not human readable.

.out – These are files containing GNOM P(r) data and are written in the standard format described in the ATSAS manual for GNOM. This file is human readable.

.sec – These are files containing *SEC-SAXS curves*, which can be loaded back into RAW and contain all of the relevant scattering profiles and structural parameters. This file type is not human readable.

.wsp – These are files for *saved workspaces*. This file type is not human readable.

Input file types

In addition to the output file types described above, all of which can be read in by RAW, RAW can load the following types of input files.

ASCII files:

.csv – If data is in a 3 column csv format, it is loaded assuming the columns are q, I, Error.

.dat (2 column) – If a .dat file has exactly 2 columns separated by whitespace, it is loaded as q I.

.dat (4 column) – Produced by FoXs when fitting a theoretical curve to a scattering profile. Loads the experimental and simulated (_FIT) scattering profiles.

.fir – Output by various ATSAS programs such as DAMMIF. Loads the experimental and simulated (_FIT) scattering profiles.

.fit – Output by various ATSAS programs such as DAMMIF. Loads the experimental (often smoothed) and simulated (_FIT) scattering profiles.

.int – Output by CRYSOL if no fitting is used. Loads the scattering intensity in solution.

.rad – Depreciated file format for RAW saved scattering profiles.

.txt - 2 or 3 column data can be read in as if it was a .dat file.

Image files:

Image files are often distinguished not just by their extension, but by their header format. For example, many detectors produce tif files, but require different methods to read in the appropriate header data. RAW uses the fabIO python module to load images. The complete list o f images and extensions supported can be found here: http://pythonhosted.org/fabio/getting_started.html#list-of-file-formats-that-fabio-can-read-and-write

We reproduce that here with the detector/format followed by the file extension in parenthesis: * ADSC Quantum (.img)

- Bruker (.sfrm)
- CIF binary files (.cbf)
- EDNA-XML (.xml)
- ESRF EDF (.edf)
- Eiger (.h5)
- FReLoN
- Fit2D spreadsheet (spr)
- Gatan Digital Micrograph (.dm3)
- General Electric (.No?)
- HDF5 (Hierarchical Data Format 5) (.h5)

- Hamamatsu CCD (.tif)
- MarCCD/Mar165 (.mccd)
- Mar345 image plate (.mar3450)
- Nonius KappaCCD (.kccd)
- Numpy 2D array (.npy)
- Oxford Diffraction (.img)
- Pixi (.?)
- Pilatus Tiff (.tif or .tiff)
- Portable aNy Map (.pnm)
- Rigaku SAXS format (.img)
- Tagged Image File Format (.tif)

In addition, we have written custom methods to support the following detector/file formats:

- 32 bit TIF (.tif)
- FLICAM (.tif)
- ILL SANS D11
- Medoptics (.tif)
- Multiwire (.mpa)
- SAXSLab300

4.6.19 Automated saving

WARNING: The manual is current several versions out of date. While it may still be useful for some users, please refer to the tutorial for the most up-to-date information.

RAW can automatically saving the following types of files:

- Processed image files (after integration)
- Averaged data files
- Subtracted data files

To turn on any of these capabilities:

1. Open the *Save Directories* panel of the Options window.
2. For the appropriate file type (Processed, Averaged, or Subtracted), select a save directory by clicking the "Set" button.
3. Check the appropriate Auto Save checkbox.
4. Click "OK" to exit the Options window and keep the changes to the settings.

4.7 Changes

4.7.1 1.6.2

Release date: 2019-10-28

Overview

The RAW team is pleased to announce the release of RAW version 1.6.2. This version fixes several critical bugs that could prevent DENSS from running. There are no other changes.

4.7.2 1.6.1

Release date: 2019-10-21

Overview

The RAW team is pleased to announce the release of RAW version 1.6.1. This version contains numerous small bug fixes and improvements, particularly for the new series analysis methods released in version 1.6.0.

We anticipate this will be the last release before RAW version 2.0, which will be our first python 3 compatible release. We're aiming to release version 2.0 sometime around the end of the year.

Note: as part of this release we did test with ATSAS 3.0 (pre-release version), and RAW seems to be compatible with it.

All changes:

- Fixed a bug where opening the Ambimeter panel could fail if ATSAS was installed in a directory with a space in the path name.
- Fixed a bug where if RAW didn't load a settings file when opened it wouldn't automatically find the ATSAS directory on startup.
- Fixed a possible bug where using the LC Series Analysis panel on series data being loaded in online mode could fail.
- Added intensity type selector for the series panel when sending data to the main plot.
- Fixed a bug where automatic integral baseline start/end region determination could set the wrong control limits.
- Fixed a bug where zero lines on various plots (like the GNOM P(r)) plot weren't getting redrawn when necessary.
- Fixed a bug with autoscaling in the LC Series Analysis plots when changing the data type displayed.
- Fixed a bug where sample and buffer region comparison in the LC Series Analysis panel could return the wrong indices for dissimilar profiles.
- Fixed a bug where profile comparison in LC Series Analysis would skip the first profile.
- Fixed a bug where a very short series (<22 frames) could cause errors when opening the LC Series Analysis panel.
- Improved automatic buffer search in the absence of major peaks.
- Improved automatic buffer search to search on the right side of the main peak if it doesn't find a good buffer region on the left side.

- Removed a bias to the left side of the peak in the automated sample region search.
- Removed actual baseline correction values from being saved in file history, as history could get very long (>100000 lines).
- Added a cutoff for header length, at which point RAW will stop saving file history. This avoids saving extremely large text files.
- Added compatibility for pyFAI 0.18 (note: on linux and python 2 pyFAI 0.18 seems to be broken, stick with 0.17).
- Added a new way of loading HDF5 files with definitions done in external files.
- Added ability to load HDF5 files from LiX.
- Fixed a bug where identical selected regions in the LC Series Analysis window didn't count as overlapping.
- Fixed a couple of typos in messages.
- Fixed a bug where Ambimeter and GNOM couldn't run if the current working directory was read only.
- Improved how ATSAS programs are called, and added use of temporary file names and directories.
- Fixed a bug where when dragging the image plot with the masking showing you could sometimes get an unexpected error.
- Added the name of the series to the LC Series Analysis panel.
- Fixed a bug where baseline subtracted profiles were improperly being skipped when calculating Rg etc in the LC Series Analysis panel.
- Fixed a bug where if a series was loaded with a baseline already calculated, changing buffer range or other parameters wouldn't properly recalculate baseline corrected values.
- Fixed a bug where if you set a baseline, then set it back to none, when exiting the LC Series Analysis window the baseline calculated values would be saved instead of the regular subtracted values.
- Fixed a bug where if you loaded in a series curve with baseline correction, then turned off baseline correction, it wouldn't have any calculated values.
- Fixed a bug where the LC Series Analysis panel would resize itself when any of the collapsible panes were collapsed or expanded.
- Fixed a bug in the LC Series Analysis panel where if you had a range that was one frame long, when you closed and reopened the analysis window you couldn't adjust the range.
- Fixed a bug in the LC Series Analysis panel where if you had a one frame long range you could get a reported correlation in the range.
- Fixed some issues with window height where windows weren't opening large enough for all of their contents.
- Fixed some issues with window size where windows could open up bigger than the screen.
- Fixed a bug where in certain circumstances opening the SVD and EFA windows could fail.
- Added compatibility with numba >= 0.44

4.7.3 1.6.0

Release date: 2019-06-07

Overview

The RAW team is pleased to announce the release of RAW version 1.6.0. This version contains several major changes:

- Completely new and improved SEC-SAXS processing, including new automated buffer and sample region selection and baseline correction. There are also significant speed improvements for SEC-SAXS processing, in addition to the new features.
- Completely redone BIFT, which fixes several bugs (both minor and major), and adds Monte Carlo error estimation and extrapolation fit of data to I(0).
- RAW now preserves matching metadata across processes like averaging and subtraction. Metadata is now saved with keywords compatible with the SASBDB to make uploading there easier for users. Users can now also provide arbitrary metadata during data reduction.
- All new icons which are compatible with retina displays, including changing out the check mark for showing/hiding data on plots for an eye, which we hope will be more intuitive.
- RAW now loads the last used config, rather than the last saved config, when it starts up.
- Any analysis window (Guinier, MW, GNOM, etc) can now be opened more than once, allowing easy comparison or side-by-side processing of data sets.

You can see the full set of changes below.

We also want to note that we're not anticipating any other major feature releases this year. With the upcoming end of life for python 2 at the end of 2019, we need to focus on making RAW work with python 3. Once that is done we will start doing major feature development again.

All changes:

- Updated DENSS to have the latest features, including refining averaged structures, symmetry constraints, and the 'Membrane' protein mode.
- Completely redid BIFT code from the ground up. This fixes several bugs, and now includes Monte Carlo estimation of errors.
- Completely redid series analysis for SEC-SAXS data. Now includes automated buffer and sample region determination and baseline correction.
- Added residual plots to GNOM, BIFT, and DAMMIF/N results.
- Fixed several bugs related to setting error bar line styles.
- Added ability to add arbitrary metadata to a file header when an image is processed by RAW
- Updated the adjusted Porod volume MW method to match the newly published MoW2 approach.
- Fixed a bug where info panel data could get improperly set
- All appropriate fields in MW panel now editable.
- You can now open any analysis window more than once (previously only one instance of each window was allowed).
- Fixed a bug where in the GNOM window changing q_min or q_max didn't update the IFT results.
- RAW now loads the last used config (saved or loaded) by default rather than the last saved.
- RAW now preserves all shared header values when averaging, subtracting, or merging datasets.
- Added visual guidelines to the dimensionless Kratky plot.
- Added option to display normalized residuals, now on by default.

- Added Rigaku HiPix to known images (requires Fabio 0.9.0)
- Guinier panel can now export Guinier fit data so users can make the Guinier plots in their plotting software of choice.
- RAW's file list no longer displays hidden files.
- Can now read in time of each data point for BioCAT data.
- Fixed a bug where closing the BIFT window with BIFT running would crash RAW.
- Better formatting for numbers displayed in the status bar.
- Fixed a bug where windows could be too large on low resolution displays.
- Fixed a bug where series plot calculated data were not highlighted by the locater button.
- Fixed a bug where markers were not highlighted by the locater button for any plot.
- Fixed a bug where when selecting a line by clicking on it the plot markers were not highlighted.
- Fixed a bug where selecting a line on the IFT plot didn't work.
- Can now display unsubtracted, subtracted, or baseline corrected intensity in the main series plot.
- Fixed a bug where series data could be truncated when exporting.
- Fixed a bug where the SVD window wasn't doing the SVD on non-error-normalized curves.
- Moved cormap to cython for speed, increased by at least 5x.
- Modified layout of the repository to standardize.
- Autorg now uses numba for just-in-time compiling. Speed increase of 2 orders of magnitude.
- Fixed bugs that would occur when quick reduce, plot, plot series, or show images were used on folders, '..', or with no files selected.
- Added ability to plot intensity over a q range for series plots.
- All-new icons that work with retina displays, including a new 'eye' for show/hide instead of a check box (hopefully more intuitive).
- Fixed a few bugs in the DAMMIF/N GUI.

4.7.4 1.5.2

Release date: 2019-04-04

Overview

The RAW team is pleased to announce the release of RAW version 1.5.2. The only change is this version is a modification to how BioCAT header files are read in, to accommodate a new header file format at that beamline.

4.7.5 1.5.1

Release date: 2018-11-01

Overview

The RAW team is pleased to announce the release of RAW version 1.5.1. This version contains several small bug fixes. Normally we might wait to release these until more major changes had happened, but there is a workshop using RAW soon and we wanted these bugs fixed before then. Significant issues that have been eliminated:

- An issue where the electron density output from DENSS could fail to load into pymol correctly because the default scaling was too small (still loaded into Chimera fine). Strictly speaking I think this is a workaround for a bug in pymol ...
- Several bugs with running GNOM, including using data with minimal sampling (<100 points).
- Fixed a bug where the .app package for Mac wasn't displaying natively on retina displays, so the text was fuzzy.

You can see the full set of changes below.

All changes:

- Fixed a bug where automatic loading of BioCAT SEC data wouldn't work if there was more than one underscore in the filename.
- Fixed a bug where automatic loading of BioCAT SEC data wouldn't work if there existed another file with the same name but different extension as one of the image files.
- Added parsing of damsup.log file for bead models, which allows highlighting of the representative model in the dammif summary.
- Dammif results summary now saved by default.
- DENSS results summary now saved by default.
- Fixed several bugs in the GNOM interface that could cause it to fail.
- Fixed a bug that prevented some .fit files from being loaded.
- Fixed a bug where the Rg for BIFT was being calculated incorrectly.
- Fixed a bug where temporary files (with a .tmp prefix) would mess up SEC autoupdates.
- Fixed an issue where you couldn't run DENSS twice without closing the panel between each run.
- Fixed an issue where the default scaling for DENSS was too small, and caused issues loading the electron densities into pymol.
- Fixed a bug text in some items and list controls would display 'fuzzy' on high dpi monitors. This is still an issue for the plot labels.
- Added the ability to run damaver and damclust on the same set of reconstructions.
- Fixed a bug where the .app package for Mac wasn't displaying natively on retina displays, so the text was fuzzy. Note that in order to fix this, even after you install the new version you may have to do the following:
  1. Enter the following commands in your terminal:

     ```
     /System/Library/Frameworks/CoreServices.framework/Frameworks/LaunchServices.
     ↪framework/Support/lsregister -f /Applications/RAW.app
     /System/Library/Frameworks/CoreServices.framework/Frameworks/LaunchServices.
     ↪framework/Support/lsregister -kill -seed
     /System/Library/Frameworks/CoreServices.framework/Frameworks/LaunchServices.
     ↪framework/Support/lsregister -f /Applications -all local,user
     ```

  2. You may then have to right click on RAW.app, select 'Get Info' and uncheck the box 'Open in low resolution mode'

4.7.6 1.5.0

Release date: 2018-08-23

Overview

The RAW team is pleased to announce the release of RAW version 1.5.0. This version focused on several significant updates that will be invisible to most users. Namely:

- RAW is now compatible with wxpython4
- RAW no longer uses weave, which has been essentially unsupported for years, to compile code. It now uses the numba just-in-time compiler.

This will make it much easier for us to support RAW, and should make it easier for users to install RAW from source on any platform. It also prepares us for the inevitable transition to Python 3 that has to happen in the next several years.

In addition to a range of bug fixes and small enhancements detailed below, RAW also now incorporates the new DENSS alignment code. This is all done in python, in RAW. This removes the dependency on EMAN2, and means that all parts of density reconstructions work on Windows!

Finally, RAW is now saving configuration files in JSON format. This is human readable, and makes the RAW configuration files more open and accessible for other programs to use. However, this does mean that earlier versions of RAW will not be able to open configuration files created with version 1.5.0 or later. However, configuration files created in earlier versions of RAW ARE compatible with version 1.5.0.

All changes:

- Fixed a bug where if atsas is in the path but not installed RAW will still find the directory from the path.
- ATSAS filepaths and filenames should be able to deal with spaces.
- Fixed various strange threadsafe bugs on debian 8.
- Weighted average now checks for similarity
- Fixed a bug where the dammif results window wouldn't work when you did only one dammif run and had damaver checked.
- Fixed a bug where dammin in normal mode wouldn't work on windows.
- Fixed a bug where dammif/n wouldn't abort on windows.
- Added in new expected shape parameter for dammif in custom mode.
- Fabio, hdf5plugin, and pyfai are now required dependencies, rather than optional dependencies
- Mode all previously compiled code into using the numba just-in-time compiler. This is important because the previous code was compiled with weave, which has been unsupported for years.
- Fixed a bug where users could give dammif/n file prefixes that were too long for damaver.
- Fixed a bug where canceling out of the color change dialog didn't cancel the color change.
- Made the plot options box resizable (important for computers with large font size).
- Fixed a bug where the sec plot right axis framestyle wouldn't properly restore if you canceled out of the plot options dialog.
- Significant code restructuring and cleanup.

- EFA calculations are now in a thread, so it might not freeze the whole GUI.
- Circle and rectangle masks are now resizable.
- Added ability to automatically mask pixels at/above/below a given threshold.
- Added ability to automatically mask images based on known detector panel gaps.
- Added ability to create predefined size/location circle and rectangle masks.
- Added ability to control detector image left-right flip and up-down flip.
- Fixed a bug where RAW could crash under certain conditions when exporting analysis info.
- Fixed a bug where the Guinier window would give an error under certain circumstances.
- GNOM and BIFT windows now show scattering profiles on log-lin axes.
- RAW is now wxpython4 compatible.
- Added alpha as an available setting in the GNOM window.
- Fixed several bugs in the GNOM window that caused RAW to unnecessarily calculate the P(r) function, slowing down the program.
- Added drag and drop file loading for both the plot and control panels.
- Settings are now saved in JSON format, which is human readable, to increase compatibility and ease of use by other programs. This means that settings saved from RAW 1.5.0 are not compatible with previous versions of RAW. Settings saved from previous version of RAW ARE compatible with RAW 1.5.0.
- DENSS now uses custom python code for aligning and averaging density. This removes the requirement on EMAN2, which means all parts of DENNS will work on Windows.
- The image plot now maintains the same zoom when you change images. Previously it would zoom back out to the full image whenever you showed a new image.
- Fixed a bug where the SVD would sometimes not open correctly.
- Fixed a bug where if there was one pixel in the q bin during integration the error would be set to 0 instead of the square root of the value
- Fixed a bug where nans or infinities in the SVD matrix would break SVD/EFA without an appropriate error message.

4.7.7 1.4.0

Release date: 2018-03-20

Overview

The RAW team is pleased to announce the release of RAW version 1.4.0. This is a major feature release for us! The big new feature is that RAW can now use the DENSS method to calculate electron density from SAXS scattering! You can read more about this at http://denss.org/.

To fully use this new feature (for density averaging and enantiomer filtering) you have to install EMAN2 (http://blake.bcm.edu/emanwiki/EMAN2/Install) which, sadly, doesn't work on Windows. Windows users can still generate densities, but they won't be able to average them. A new tutorial on DENSS in RAW is now available in the documentation (https://bioxtas-raw.readthedocs.io/en/latest/tutorial/s2_denss.html).

The other feature many folks will be interested in is the new error calculation for Guinier fits, which is a much requested feature. This is now available whenever you open the Guinier panel, and saves and exports with the rest of the analysis information as expected.

We've also done the usual set of bug fixes and tweaks. You can find a full list of changes below.

All changes:

- Added DENSS method for calculating electron density from SAXS profiles
- Added support for EMAN2 averaging and enantiomer testing of DENSS results
- Fixed a bug where the GNOM window could fail to exit and save the .out file to the IFT tab
- Changed the default DAMMIF mode to slow.
- Changed when the 'please wait' message appears when loading SEC-SAXS files in autoupdate mode. Now it only shows up if more than 5 files are loaded at once.
- Fixed a bug where advanced options for GNOM and DAMMIF couldn't be set while the respective analysis windows were open.
- Fixed a bug where the spectral color map couldn't be displayed, breaking the image control panel.
- Fixed a bug where ambimeter would try to run in the DAMMIF window even if ambimeter wasn't available.
- Fixed a bug where if files were averaged or subtracted and had analysis history, that analysis would get transfered to the new file.
- Fixed a bug where Guinier fit limits would be improperly displayed on the plot when the Guinier window was first opened.
- Fixed a bug where calls to set up the DAMMIF results window could be non thread safe.
- Added estimate of the parameter (Rg and I0) errors for a Guinier fit.
- Reformatted the MW display to make it more compact.
- Changed how numbers are displayed in all of the analysis windows, to better handle very large or very small values.
- GNOM, Ambimeter, DAMMIF windows now open much faster.
- Added support for BioCAT header files (new style).
- Added support for autoloading of BioCAT Series curves.
- Added GNOM P(r) parameters (Rg, I0) errors to the GNOM window, and the estimated Guinier errors.
- Guinier parameter errors and GNOM P(r) parameter errors are now saved with profiles, and with analysis info spreadsheets.
- Fixed bugs where spin controls could raises errors if a user entered a blank value.
- Values from analysis windows are now saved with more precision.
- Rearranged the manipulation item right click menu to make it more compact, put some less-used items on sub-menus.
- Changed 'SEC' labels to 'Series' labels.
- Fixed an off by one error in SEC autoupdate that could occur for certain file names.
- Renamed and rearranged some menu items in the IFT item right click menu.
- Added universal newline support when loading in scattering data.

- Fixed a bug where averaging could fail if all the averaged files were different form the first file.
- Fixed a bug where similarity testing could fail with an overflow error if there were too many points in the scattering profile.
- Minor improvements to plotting speed with large numbers of files.
- Fixed a bug where having no positive values in a curve displayed on a log-y axis would cause an error.
- Updated the documentation to include a DENSS tutorial. Updated various other parts of the documentation, including the images, to reflect other new features.
- Updated all of the installation documentation.
- Removed the RAW-Windows-Source-Install-Essentials file from the downloads.

4.7.8 1.3.1

Release date: 2017-11-01

Overview

The RAW team is pleased to announce the release of RAW version 1.3.1. This is a very minor release. Several small bugs have been fixed, and we have updated the citations in the program to reflect the release of the new RAW paper. Most of the major work in this release went into updating the documentation, which we have already released on the new website: http://bioxtas-raw.readthedocs.io/

All changes:

- Made RAW compatible with pyFAI 0.14 (not back compatible with 0.13)
- Improved the multiwire loading function
- Updated some citations and error messages in the program
- Revamped and updated all of the documentation and tutorials. It is now in sphinx format, in the RAW SVN for better tracking.
- Updated the RAW citation to reflect the newly released RAW paper.
- Updated the .app build on mac.

4.7.9 1.3.0

Release date: 2017-08-19

Overview

The RAW team is pleased to announce the release of RAW version 1.3.0. This release is a major feature release, and we're very excited that you get to use all of the fun new stuff we've added in! The major new features are:

- Similarity testing for scattering profiles using the CorMap test. This allows statistical testing of whether or not profiles are similar. This is done automatically when averaging profiles or picking a buffer region of a SEC curve, and is also available in the right click menu for profiles, IFTs, and sec files. In the automatic check, if it detects files that may be different, you'll see a message asking you how you want to proceed.
- Normalized Kratky plots can now be made, and are accessible through the right click menu.

- We've added a results summary panel for dammif/n reconstructions that shows the NSD, resolution (if SASRES is installed), and statistics about the individual reconstructions including chi squared, rg, dmax, excluded volume, and molecular weight. There is also a new dammif results viewer panel that lets you get a basic look at the reconstructions (this panel is still very simple).
- Absolute scaling can now be done using the NIST glassy carbon standard SRM 3600.

In addition to all of these major changes, we've made the usual range of small tweaks, bug fixes, and enhancements. See the full list of changes below.

Finally, we're happy to announce that we're also releasing a new tutorial, that has been updated to include tutorials for all of the new features mentioned above!

All changes:

- Fixed a bug where switching between linear and log scale in the image display could change the overall scaling of the image without changing the displayed limits in the dialog.
- Added a new dammif/n results summary panel.
- Added a new dammif/n results viewer panel.
- Added a new normalized kratky plot panel
- Changed how multiple images in a single file are deal with when loaded as a sec curve (now each is loaded as an individual point on the curve).
- Added a new check for statistical similarity between profiles (or IFTs or SEC curves).
- Now on average, RAW automatically checks whether the profiles are statistically similar.
- Fixed a bug where the first image loaded from a file with multiple images in t was flipped left-to-right relative to the rest.
- Fixed a bug where if a configuration file is loaded and doesn't contain certain setting keys (a configuration made with a previous version where those settings don't exist, for example), those settings are now set to default, rather than left as whatever is loaded in RAW.
- Added ability to view all images in a single file if the file contains more than one image.
- Added ability to use glassy carbon (NIST SRM 3600) to calibrate absolute scale.
- Fixed a bug in subtraction that could result in the q and i vectors being rounded.
- Fixed a bug where if the beam center was in the masked region of the image it could be assigned a non-zero value.
- Fixed a bug where a RAW setting for a choice type with default value of None could cause an error when trying to set the field in the Advanced Options window.
- Added a check for syncing items to make sure that an item is starred and an item is selected.
- Added ability to reset all settings to default values from the advanced options panel.
- Marker face, marker edge, and error bar colors are now saved when you save a workspace.
- Error bars now show up correctly for Guinier, Kratky, and Porod axes in the Main Plots.
- Added ability to use error weighting in fits, and ability for user to toggle that on and off in the advanced options panel. Fitting is now by default done with error weighting.
- RAW can now load .txt files.
- Fixed a bug where on a single core machine there would be no default selection for the number of simultaneous runs in the dammif/n window.

- Font list now includes matplotlib fonts
- Changed LaTeX symbols to default to regular instead of italics.
- Fixed a bug where line size on a plot would change when opening/closing the line properties window without making any changes to the line size in the window.
- Added ability to use fractional line sizes.
- Fixed a typo in the readme
- Removed a message asking if you're sure you want to load the workspace.
- RAW now checks whether or not you're saving something when it quits. If it is saving something, it warns you that you might now want to quit.
- Legend labels are now saved with a workspace.
- Fixed a bug where the legend label for IFT items would get changed from the default when you opened the line properties window.
- Fixed a bug where the calculated markers for a SEC item would show when loading a workspace even if the item wasn't supposed to be visible.
- Added sync and superimpose to the right click menu, tools menu.
- Added the program version to integrated dat files history.
- Added integration method and calibration parameters to the integrated dat files history.
- Fixed a bug where a dammin refine would try to run even if damaver didn't run.
- Fixed a bug where superimpose could break for different q vectors.
- Fixed a bug where the slider and custom color boxes in the color dialogs didn't change line/marker colors.
- Fixed a bug where in autoupdate mode the SEC plot could fail to switch between rg, mw, i0 on the right axis.
- Fixed a bug where you couldn't resize custom question dialogs.
- Fixed a bug where SVD/EFA wouldn't work with some sec data loaded in autoupdate mode.
- Fixed a bug where when updating the SEC data in autoupdate mode, an improper q value could be used when getting the intensity at a given q.
- Fixed a bug where if improper values were entered in the buffer range or window size and the set/update parameter button was pressed, if autoupdate mode was on it would stop.
- Removed the error printing on startup that backup.ini file could not be found.
- Fixed a bug where carrying out EFA to panel 3, then going back to panel 1 and changing the frame range used, then carrying out EFA again could cause an error in the rotation.
- Fixed a bug where for unsubtracted profiles from images, EFA would use the full profile rather than the appropriately truncated profile.
- Fixed a bug where the options panel couldn't be opened twice in windows.
- Added a check to prevent errors with missing lines when changing plot type in the main plots.
- Added a check to prevent index errors when setting the q range of a sasm.
- Fixed a bug where online mode would show an error if the directory being watched was removed.
- Added a choice in the GNOM panel to force dmax to zero or not.
- Added ability to use superimpose to find scale, offset, or scale and offset.
- Fixed a bug where EFA results wouldn't export due to getting the wrong q values from the scattering profiles.

- Fixed a bug with new versions of numpy not integrating images correctly. (actually fixed in 1.2.3 rerelease)
- Changed the generic error message. (actually fixed in 1.2.3 rerelease)
- Fixed a bug where temporary files that vanish in the online directory could raise an error. (actually fixed in 1.2.3 re-release)
- Fixed a bug that could cause intensity integration to fail in the sec plot. (actually fixed in 1.2.3 re-release)
- Fixed a bug where calculating the scale constant of water could cause the main thread to lock up if it had an error.
- Verified compatibility with ATSAS 2.8.2.
- Fixed a bug where in the prebuilt windows version any plots not in the main window (for example, Guinier plots) couldn't be saved.
- Fixed a bug where line colors didn't reset properly when canceling out of any of the line properties dialogs.
- Fixed a bug where the SVD window could have no default selection for type of profile to use.
- Fixed a bug where the advanced options window didn't open properly centered on the parent window.
- Minor speed improvements from code streamlining.

4.7.10 1.2.3

Release date: 2017-05-08

Overview

The RAW team is pleased to announce the release of RAW version 1.2.3. The release again mostly focuses on bug fixes, speed improvements, and other small improvements to the user experience. There is one bit of exciting news: we are releasing a prebuilt version for Mac! Users can now download a .dmg with a RAW.app in it. This can be installed via drag-and-drop, like other app files, and run just like any other app. We hope this will make installation much easier for mac users. If you want to try this, the download is available in the usual area, and the mac install instructions have been updated.

In addition to the new prebuilt version, we've also made errors more obvious, now if there is an unhandled error in the program, rather than failing silently it will pop up a dialog box to let you know. We're hoping this is seen as an improvement!

All changes:

- Made numerous changes to fix strange behavior in frozen version on mac
- Created instructions for building a frozen version on mac
- RAW icon now shows up in the dock instead of the top bar on mac
- Fixed how RAWWorkDir is used in the program, and how it gets set. It now gets set appropriately for each type of OS
- Switched to using an embedded version of the BioXTAS logo, for easier packaging
- Changed the default directory for RAW if there is no previous directory. It now uses the documents directory
- Fixed a bug in the Porod volume calculation that in some cases could extrapolate to q<0
- Added Guinier extrapolation to the volume of correlation molecular weight calculation

- Changed how Guinier extrapolation is done for the adjusted porod volume method
- Updated some of the text in the More Info buttons of the MW panel
- Fixed a bug where changing the q vector of a scattering profile would print an error in the console
- Updated the A and B coefficients for the adjusted porod volume method to perfectly match those used in the paper
- Updated GNOM and BIFT windows to both report reduced chi squared values
- Fixed a bug where having an ROI mask set could prevent loading image headers in the calibration section of the advanced settings
- Tweaked the MW, GNOM, and BIFT GUIs
- Fixed a bug where GNOM wouldn't run on SL6 with ATSAS 2.7.2
- Fixed a bug where rescaling profiles wouldn't work on a kratky plot
- Fixed a bug where IFT data plot could display the wrong scale for Guinier and Porod plots
- Changed how Guinier plots are displayed from I vs. q2 on a loglin scale to log(I) vs. q2 on a linlin scale, to match with labels shown on the plot
- Fixed a bug that could cause autorg to crash
- Changed the circle masking tool to draw more quickly/smoothly
- Improved responsiveness of dragging masks on an image
- Fixed a bug where the beam center wouldn't turn off if the masking panel was closed
- Improved responsiveness of updating positions of calibrant rings and beam center when working in the centering panel
- Fixed some bugs that could happen when switching between calibration and masking windows without hitting the okay or cancel buttons first
- Fixed a bug where VC integration was highly unstable in some cases, required switching from simpsons method to trapezoid method for numerical integration
- Attempted to fix a not reproducible bug where clicking the clear all button could cause a segfault on linux
- Fixed a bug so that the info panel is only cleared if the user actually decides to clear all items when clicking the clear all button
- Fixed a bug where loading FoXS files with fits would not load the fit
- Fixed a bug where PIL.Image couldn't load files (prevented loading of SAXSLAB300 images)
- Fixed a bug where if an image load returned no header, RAW could crash
- Did some futureproofing in the code
- Fixed a possible memory leak when loading certain image types
- Attempted to fix an irreproducible bug where masking would fail because pixel positions were floats instead of ints
- Error bars, if shown, now move properly with the line on scale and offset
- Fixed a bug where the Guinier window didn't respect the q limits set on the manipulation panel
- Set the default plot type to log-lin instead of lin-lin
- Fixed a bug where using the next/previous image buttons would cause the image to flicker if a fixed range were set for the color scale

- Fixed a bug where scaling q didn't mark the item as modified
- Fixed a bug where online mode loading more than one image didn't update the image plot
- Changed the green for the average file name text from green to forest green, which may be easier to read
- Fixed a bug where the centering panel being displayed without an image loaded could cause an error.
- Fixed a bug where the ATSAS 2.8.0 GNOM wouldn't run if an Rg for the profile had not been calculated.
- Fixed a bug where DAMCLUST wouldn't run.
- Added a global error handler to pop up a dialog for unhandled errors.
- Attempted to fix a bug where the program could run out of control ids on mac, causing a crash.
- Fixed a bug where damclust and dammin refine could both be selected in the advanced options window.
- Fixed a bug where dammin refine could be selected without damaver being selected in the advanced options window
- Fixed a bug where setting a flatfield image could fail if there wasn't an absolute scale normalization factor set
- Fixed a bug where GNOM and BIFT autosaving could be turned on without directories selected.
- Fixed a bug where switching from linear to log scale or vice versa with limits locked in the image display would set the slider bar maximum value incorrectly.
- Removed tifffile.py (no longer used).
- Fixed a bug where automated centering wouldn't work with the newest pyFAI
- Fixed a bug where typing an incomplete LaTeX expression in the plot label could cause an error.
- Added some error checking to running GNOM/DATGNOM in case it fails for some reason.

4.7.11 1.2.2

Release date: 2017-03-10

Overview

The RAW team is pleased to announce the release of RAW version 1.2.2. This release mostly focuses on bug fixes, speed improvements, and other small improvements to the user experience. However, there are several changes/new features we think many of our users may want to know about:

1. RAW now has the ability to do weighted averages of scattering profiles (accessible by the right click menu in the main control panel)
2. RAW is now compatible with ATSAS version 2.8.0.
3. You can now run DAMMIN from RAW (previously on DAMMIF was available). This includes using DAMMIN to refine the damstart.pdb file output from DAMAVER, which is now the default option.
4. RAW can now handle files with multiple images in them, such as Eiger hdf5 files. This is an ongoing project, so some features, such as image viewing and SEC plotting do not yet handle these types of files perfectly.
5. We have changed how the show/hide and collapse/expand buttons work. Previously they affected all items. Now if no items are selected they affect all items, otherwise they affect the selected items. We hope that once users are accustomed to this change they will find it useful.
6. RAW has a new header type available, P12 Eiger header files.

Additionally, RAW users should be aware that we have added an additional dependency, the weave package (to replace scipy.weave, which was removed in scipy version 0.19), and that RAW is not yet compatible with matplotlib version 2.0 (released January 2017). We are working on updated install instructions to reflect these changes, and those will be available (hopefully) next week.

As always, we appreciate user feedback, as that is how we improve the program. If you have questions, need help, or want to report a bug, please contact us!

All changes:

- Added ability to do a weighted average in RAW, using either error based weighting per q point or weighted by a counter value.
- Many small changes to the code to streamline how plotting works, which should results in modest speed improvements, particularly when working with large numbers of plotted files.
- If autoscaling is on for plots, plots should now autoscale in all appropriate instances (previously they didn't autoscale when moving items between plots, rescaling the q range, and a few other instances)
- Trimmed out many dead functions to make the code easier to maintain.
- Changed how the visibility check box for control panel items works, which improved show/hide speed for a single item when lots of files were loaded by a factor of 2.
- Improved speed for certain actions that resulted in marking lots of items as modified.
- Fixed a bug in autorg where error for the rg value could fail to be calculated
- Fixed a bug in running GNOM for ATSAS <2.8 where certain advanced settings couldn't be used.
- Fixed a bug where flatfielding would fail when using pyFAI to integrate images (not yet publicly available)
- Fixed a bug where using the roi_counter would fail when using pyFAI to integrate images (not yet publicly available)
- Fixed a bug where dezingering would fail using python for integration (instead of the compiled c++ modules)
- Removed the SASIft.py file that was unused.
- Fixed a bug where having nothing entered for limits in the plot options panel (such as when typing a new limit) would print an error message in the console.
- Fixed a bug where loading a roi_counter header value with no image header would give an error.
- Fixed a bug where legend position wasn't maintained when all items were removed or hidden on a plot.
- Fixed a bug where the legend wouldn't go away if all items on the SEC plot were hidden and there had previously been a legend.
- Updated how legend settings are handled in plot options to improve speed and maintainability.
- Fixed a bug where plot titles and axes labels didn't reset appropriately when using the clear all button.
- Fixed a bug where the plot options font selector boxes didn't work.
- Fixed a bug where not all settings were restored to previous values when canceling out of the plot options dialog.
- Fixed a bug where the Porod volume calculation was not getting properly interpolated to q=0.
- Fixed a bug where hitting the next/previous image buttons in the RAW Image plot would throw an error and crash RAW if the image currently displayed wasn't in the current working directory of the Files panel.
- Fixed a bug where saving items wasn't threadsafe on scientific linux 6.
- Fixed a bug in how the error bars for log(I) were calculated in the autorg function.

- Switched the autorg to calculate the Guinier fit without error weighting, to match how it is done in the Guinier panel.
- Fixed a bug where the how to cite button in the dammif frame wasn't getting properly placed in wxpython < 3.0.
- Addeed the ability to run dammin from the DAMMIF (now DAMMIF/N) window.
- Added the ability to use dammin to refine damstart files from dammin/f in the DAMMIF window.
- Fixed a bug where autoMW, autoRG did not respect the limits set for the scattering profile in the manipulation controls.
- Changed how the show/hide and collapse/expand buttons work. Previously they affected all items. Now if no items are selected they affect all items, otherwise they affect the selected items.
- Added compatibility for gnom5 from ATSAS 2.8.
- Counters available for normalization now show up in the combo box in the normalization list panel.
- Made some progress fixing a windows specific bug having to do with hitting enter after clicking a button in another panel.
- Fixed some bugs on windows where the mouse would get captured and not released by txtctrl boxes.
- Fixed a bug where the rename option wasn't working in the file overwrite dialog.
- Moved the version number into the RAWGlobals.py file.
- Improved speed of saving items from RAW, by a factor of ~160x for a large number of files on my test machine.
- Fixed a bug in the Guinier panel where the maximum point shown in the plot and used for the fit was one less than the maximum point shown in the spin control.
- Tweaked the autorg function to allow some intervals with qmaxRg > 1.3 (up to 1.35) to improve fitting.
- Fixed a bug where interpolate did not work on multiple selected scattering profiles.
- Fixed a bug where interpolate was giving the interpolated file the wrong name.
- Fixed a bug where writing the header could cause RAW to crash due to improper json serialization.
- Changed how normalization deals with zero values. Instead of raising an error it prints a warning.
- Added the GNU disclaimer at the top of all .py files that didn't have it.
- Added a header type for P12 Eiger, Petra III
- Updated image loading and all associated functions to handle multiple images in a single file, for example eiger files.
- Added filtering of headers so that () and [] characters are removed, as header names with these characters could not be used for normalization.
- Fixed a bug where image and other headers were getting filtered differently.
- Added some new file types to the TestData folder.
- Added error catching for json formatting of file headers upon save. If the header can't be serialized properly, the files saves without a header (used to cause a crash).
- Fixed a bug where ambimeter could fail if there were spaces in the filename.
- Fixed a bug where with older versions of wxpython and matplotlib, failure to find points in the autocentering mode could cause RAW to freeze.
- Fixed a bug where quick reduce would crash if it couldn't find the header.

- Replaced the dependency on scipy.weave with the weave package (which is a fork of scipy.weave), as scipy.weave is removed in scipy 0.19.

4.7.12 1.2.1

Release date: 2016-12-02

Overview

The RAW team is happy to announce the release of RAW version 1.2.1. This version focuses on bug fixes and small improvements to the user experience. There were a few significant changes:

1. In addition to numerous bug fixes, the EFA technique can now be used with explicit, iterative, or hybrid methods for computing the concentration profiles of the components. Previously, only the iterative approach was available.
2. We added a new automated centering and calibration routine using the pyFAI library, for better determination of beam center and sample-detector distance.

In addition to a new version of RAW, we have also released new installation instructions for all platforms.

As always, we appreciate user feedback, as that is how we improve the program. If you have questions, need help, or want to report a bug, please contact us!

All changes:

- Updated online mode so RAW only plots files if there are files to plot. This prevents some flickering when files enter the directory but are not plotted for any reason (such as not being suitable images).
- Updated online mode so that the "Processing incoming file…" status doesn't linger forever after an image is processed, but goes away suitably quickly.
- Fixed a bug that prevented EFA from running on scattering profiles that don't use the full range of their q vector.
- Fixed a bug where concentration wasn't saved when the 'save all analysis info' option was used.
- Fixed a bug where changing SEC plot axes while SEC live update is going could cause a crash
- Fixed a bug where Normalization information got saved in the scattering profile processing parameters twice, once with a capital N, once with a lowercase n.
- Fixed a bug where the wrong upper limit was getting set for the end of range controls in the third EFA control panel.
- Fixed a bug where if no normalizations were set in the normalization list, the solid angle correction would not be saved in the normalization history list for the scattering profile.
- Made a change where if EFA has a converged solution, if the ranges are changed it uses that solution as a starting point. This leads to faster convergence to the new solution.
- Added ability to display calibration rings from any calibrant in the pyFAI library.
- Fixed a bug where plotting certain scattering profiles on a Kratky plot would cause RAW to crash
- Fixed a bug where having the SEC plot set to display the intensity at a particular q value would prevent structural parameters from being calculated, and in some cases could prevent new SEC items from being plotted.
- Fixed a bug where the plot legend wasn't updated if the plot was turned on, then off, and then items were removed from the plot.

- Added an energy box in the centering and calibration window, so that if energy is entered, wavelength is automatically calibrated, and vice versa.
- Fixed a bug where changing centering values with no centering values selected could crash RAW.
- Added ability to explicit calculation of concentrations for EFA, as opposed to currently iterative method.
- Added ability to use a hybrid method for calculation EFA, using the explicit calculation as a starting point, then refining iteratively.
- Added ability to chose rotation method for EFA in the third EFA control panel.
- Fixed a bug where the range plot in the third EFA panel was not refreshing properly when the number of significant values was changed.
- Fixed a bug where the info panel was not updated when a scattering profile was selected by clicking on it on the main plot.
- Updated build commands for making a windows installer, including adding some explicit hooks for pyFAI and pyinstaller.
- Added the optional use of the hdf5plugin to RAW to support eiger images.
- Fixed a bug in the image display where the dialog box could fail to open because the maximum value in the image was greater than 2^31-1 (the maximum value a wx slider can handle).
- Added a feature for automatic centering and fitting of the beam center and sample to detector distance. Requires pyFAI to be installed.
- Added a header reader for g1 eiger files, which have the spec header file one level up from the image files.
- Fixed a bug where the RAW ROI could not consistently be used for normalization.

4.7.13 1.2.0

Release date: 2016-10-25

Overview

The RAW team is very pleased to announce the release of version 1.2.0. We've added two major new features, the first of which is the ability to perform SVD on a set of scattering profiles, IFTs, or a SEC-SAXS curve. We've also implemented the exciting new evolving factor analysis (EFA)[1] method for deconvolving overlapping data. This is primarily intended to be applied to SEC-SAXS data, but it is implemented so that it can be applied to any set of scattering profiles or IFTs. We want to note that while EFA is an exciting new technique, it is still in ongoing development. We intend continuing development on the stability and utility of the algorithm.

We will release an updated tutorial document and dataset which includes examples of doing SVD and EFA soon.

As always, we appreciate feedback from users, either positive or negative.

The RAW Team

[1] Steve P. Meisburger, Alexander B. Taylor, Crystal A. Khan, Shengnan Zhang, Paul F. Fitzpatrick, and Nozomi Ando. Journal of the American Chemical Society 2016 138 (20), 6506-6516.

All changes:

- Added the solid angle correction to the normalization parameters in the sasm history, so that if it is used, that use is recorded.

- Fixed a bug where SAXSLAB images could not be loaded when using version 3.0 or newer of the pillow library.
- Added in the ability to use a RAW defined beamstop mask in addition to a SAXSLAB beamstop mask for SAXSLAB data.
- Fixed a bug (on OSX, wxpython 3.0) where clicking the OK button in the Masking Panel was returning the plot window to the IFT panel instead of the Main Panel.
- Added in some dialog boxes letting users know they can't modify the SAXSLAB header mask in RAW. Previously, the Remove and Set buttons in the masking panel appeared to work for the SAXSLAB header beamstop mask, but in reality did nothing. Now they still do nothing, but pop up a dialog letting the user know that nothing has happened (and no longer appear to do anything).
- Added a molecule type choice to the SEC calculate parameters panel, so that the user no longer has to change the default molecule type in the mol weight options panel.
- Fixed a bug where the Clear All button was not properly clearing some fields in the SEC control panel.
- Added SVD capability.
- Fixed a bug which prevented some .sec curves from being loaded.
- Added overwrite checking to the .sec saving function.
- Fixed a bug where the SEC item filename didn't change when the item was saved with a different name.
- Made how SEC names are deal with consistent with how scattering profile names are dealt with.
- Added overwrite checking to the Export data option for SEC curves.
- The parameters on a SEC plot now default to markers, not lines.
- Fixed a bug where in a 3 column data file with no non-data first line (empty or otherwise), the first data point would get cut off.
- Added evolving factor analysis (EFA) capability
- Added 'How To Cite' buttons for the RAW functions that incorporate other people's work, so that they can correctly cite the methods.
- Added in backwards compatibility for loading .sec files from previous versions of RAW, and workspaces with saved .sec files from previous versions of RAW.
- Saving/Loading a workspace now preserves the file order in the workspace.
- Fixed a bug where selecting log axes would crash RAW if you tried to do so before loading any data.
- Fixed a bug where the legend label for ift and sec items got set when it didn't need to be.

4.7.14 1.1.0

Release date: 2016-08-22

Overview

The RAW team is happy to announce the release of version 1.1.0. While there are several significant new features, the major milestone that pushed us into version 1.1 is the integration (after almost a year) of the RAW code that has been available on this website and the RAW code improvements made by Soren Skou for use with the SAXSLAB home-source machines. All of RAW is now unified, and we intend to have only one development trunk for the foreseeable future (though we may have temporary branches for major feature development).

We have also added in a solid angle correction for integrating images into scattering profiles. This correction accounts for the change in solid angle of a pixel as you change q. We have tested it against the solid angle correction implemented in pyFAI, and found that the results are identical. This effect will get stronger at higher q, and cause an overall increase in intensity of integrated profiles. On a Pilatus detector, the solid angle correction has a ~0.5% effect on integrated intensity at q=0.25 A^-1 and ~4% effect at q=0.75 A^-1.

Major new features include:

- The solid angle correction mentioned above
- Improved speed when calculating Rg, MW, and I(0) for SEC-SAXS curves (up to a factor of 7 faster in our limited testing)
- Ability to read in multiwire (.mpa) files
- Ability to read in headers from SAXS beamline BL19U2 at the Shanghai Synchrotron Radiation Facility
- Merging, rebinning, and interpolating now all save history information like averaging and subtracting have
- Scattering profile history (either: averaging, subtracting, merging, rebinning, and interpolating, or information about loading in and normalization) can now be viewed within RAW by right clicking and selecting 'Show history'
- RAW is now (mostly) compatible with wxpython 3.0 on Linux

Beyond these changes, there are numerous small improvements, visual tweaks, and bug fixes. You can find a full list of those below.

Simultaneous with this release we are also releasing updated installation guides for all platforms. We are happy to say that we are confidant enough in our ability to produce prebuilt windows installers that we now recommend that windows users install from the .msi files unless they know that they need to compile from source.

As always, we appreciate any feedback (positive, or, especially, negative), bug reports, and suggestions for new features!

All changes:

- Fixed a bug that prevented BIFT from running in uncomplied mode
- AutoRG now runs automatically when the Guinier window opens, assuming there is no previous Guinier analysis
- Fixed a bug where BIFT failing to find a solution caused RAW to crash
- If autosave is active, and a the folder vanishes, autosave now detects that, and is disabled, instead of crashing RAW
- When RAW settings are loaded, all folders and files in the settings (autosave directory, online directory, flatfield file) are checked. If they cannot be found, these settings are disabled, and the user is notified.
- Visual improvements of the BIFT window, DAMMIF window, and some options windows
- Fixed a bug where analysis windows would show up behind the main window, where you could move them by dragging the title bar without losing focus on the analysis windows, and where you could bring them to the front without first clicking on the main window
- Changed the layout in the SEC tab to be more descriptive, and to save space
- Changed welcome dialog info
- Fixed display problems of the Guinier and GNOM windows under wxpython 2.8 on Ubuntu
- Added the ability to start online mode at startup with a predefined directory

- Added the option of automatic saving of BIFT and GNOM results
- Updated save functions in RAW so that files that RAW saves are not automatically loaded back into RAW
- Added in option (on by default) to apply a solid angle correction to the integrated data to account for change in solid angle of the pixels with q
- Fixed several small bugs with the online mode: crashing when the online mode directory ceased to exist, online mode being able to start without selecting an online directory
- All counters and image header parameters now automatically have any spaces in the file name replaced with underscores, so that they do not crash the normalization
- DAMCLUST is now available as an alternative to DAMAVER after running DAMMIF
- Merging, rebinning, and interpolation now add to the file history in the same way that subtracting and averaging have
- Added a new feature to view the file manipulation history or load history within RAW (right click on a scattering profile in the manipulation list and select 'Show history')
- Added a sorting function to the .dat file saving so that file parameters should always appear in the same order in the saved file
- Fixed a bug where a tiff file with the wrong header getting read in as a Pilatus tiff file would cause RAW to hang up
- Added extra error catching to the file header load function
- Sped up calculation of SEC-SAXS Rg, MW, and I0 by adding a threshold function. The threshold checks the ratio of integrated sample intensity (or whatever intensity is being used for the SEC plot) of the average buffer to the average sample files. If the intensity is not above the set threshold (1.02 by default), it does not try to calculate the parameters. This means all of the buffer curves are automatically skipped, and calculation is much faster. It depends on the threshold and the data, but I saw speed increases of up to ~7x. This can be set by the user in the new SEC-SAXS panel in the Advanced Options window.
- Changed how normalization information is saved when a .dat file is saved. Now, normalization information is only saved when it is applied. The absolute scale factor applied is also now saved
- Added more files to the list of files that can be loaded in online mode
- Updated sync function so that files are only marked as modified when something is changed during the sync
- Modified how the centering arrows work to catch faster clicks, and to (mostly) prevent two moves with one click (noticed on a mac)
- Masks with zero area are no longer saved as masks
- Added the ability to load some multiwire detector files (.mpa files)
- Added the ability to read in the header for files from BL19U2 at the Shanghai Synchrotron Radiation Facility
- If the image or beamline header contains a concentration key word, that is now set as the sample concentration in RAW when the image is loaded
- Fixed a problem where ambimeter in the ATSAS 2.7.2 package could not be run
- Fixed numerous small and large visual problems with running RAW on linux with wxpython 3.0. I now believe that RAW can be considered compatible with wxpython 3.0 on all platforms, but there are still occasional sizing issues on Linux that it does not handle perfectly
- Fixed a bug where damaver and damclust would not run if the directory path contained a space

4.7.15 1.0.3

Release date: 2016-07-20

Overview

We're releasing the latest version of RAW, 1.0.3 today. This includes several minor bug fixes. The timing of the release is done so that the version being demoed at the ACA meeting (http://www.amercrystalassn.org/2016-scientific-program#SAXS) will be identical to the latest release.

All changes:

- Fixed a bug where saving a mask without an image loaded would cause an error.
- Fixed a bug where attempting to show a SAXSLAB BS Mask without a SAXSLAB image loaded would cause an error.
- Fixed a bug where autosaving for files (processed image files, averaged files, subtracted files) could be turned on without a valid save directory selected.
- Added a feature so that when an autosave directory is cleared, autosave for that file type is turned off.
- Fixed a bug where the final lines of the damaver output were not being shown in the dammif window.
- Added some extra information to the two most common error messages we get contacted about: inability to load an image type, and inability to load a header file.
- Fixed an error where if an image header contained non-unicode characters, when a scattering profile generated from that image header was saved it would crash RAW. Fixed the same error if the header was shown.
- Removed some unused settings values.
- Removed the brightness bar in the image settings pop up window, as it was currently disabled. This may be re-enabled in the future.
- Set the image settings pop up window to have the default upper value be the max pixel value, rather than 65535.
- Fixed a bug where starting two dammif runs in the same window (running it again after either aborting or letting the current runs finish) did not clear the old log tabs.
- Fixed a bug where entering a wavelength longer than ~115 A resulted in an error. Now a window pops up informing you of the error and you have to re-enter the wavelength value.
- Fixed a bug where the quick reduce dialog was not displaying, and thus quick reduce could not be used.
- Profiles reduced using quick reduce will now have a q range corresponding to the start/end skip points in RAW, consistent with items loaded into RAW and saved from there.
- Fixed a bug where certain .fit files and FoXS .dat files with 4 columns would not plot properly.
- Fixed a bug where the x and y axis values of the Guinier plot were not updating when the data range was changed
- Relabeled the residual plot in the Guinier window with the correct axis labels.
- Updated how GNOM, BIFT, an Guinier plots are refreshed for improved speed and to remove certain display glitches.
- Changed the header display in the image panel to be read only (since changes there were not saved).
- Removed the automation and SANS options panels, as they had no effect. These may be reenabled in the future.
- Changed the default bin size in RAW for q spacing from 2 to 1.

- Removed some extraneous print statements.
- Cleaned up RAWAnalysis.py code and some code in SASFileIO.py
- Added ability to load .fir files.
- Fixed a bug where most of the new image types added in RAW 1.0.2 were not being recognized by RAW.

4.7.16 1.0.2

Release date: 2016-06-22

Overview

We're happy to announce that we're releasing RAW 1.0.2. This is another version focusing on small bug fixes and speed improvements, to try to increase the stability and usability of the software. As always, please report any bugs you find to us, so we can fix them!

The one major change is the inclusion of the fabIO package (https://pypi.python.org/pypi/fabio) for opening images. This has allowed us to support a number of new image types. RAW now supports images in the following formats:

- Pilatus TIff
- CBF
- SAXSLab300
- ADSC Quantum
- Bruker
- Gatan Digital Micrograph
- EDNA-XML
- ESRF EDF
- FReLoN
- Nonius KappCCD
- Fit2D spreadsheet
- FLICAM
- General Eelctric
- Hamamatsu CCD
- HDF5
- ILL SANS D11
- MarCCD 165
- Mar345
- Medoptics
- Numpy 2D Array
- Oxford Diffraction
- Pixi
- Portable aNy Map

- Rigaku SAXS format
- 16 bit TIF
- 32 bit TIF

All changes:

- Removed tifffile warnings upon opening RAW
- Improved the SEC-SAXS online mode based on user feedback to make it easier to work with.
- Fixed an issue where active masks could be removed from memory when saving config files.
- Fixed an issue where no warning was being displayed when config files failed to save properly.
- Improved the speed of selecting large numbers of manipulation, IFT, and SEC items by at least 3 orders of magnitude.
- Updated how loading and plotting works to improve speed by a factor of ~2.5 for both loading and subtracting large numbers of items.
- Updated the Plot Sec button to improve the speed of file loading in certain cases.
- Fixed a bug where FLICAM images could no longer be loaded due to changes in how tiffs are loaded in pillow >=3.0
- Removed some possible issues with loading items where files were not getting closed correctly.
- Fixed an error where rebinning an item under certain conditions could crash RAW.
- Added a warning if a users tries to update or send frames from a hidden SEC curve (assumes that they forgot to change their selection)
- Fixed a big where sending the same frames twice to the main plot from a SEC curve would cause various problems with RAW.
- IFT items are now marked as modified when they are renamed.
- Fixed an error caused by clicking on the top item of the advanced options configuration tree
- Fixed an error in the Image tab where selecting the pan/zoom buttons wouldn't always properly toggle the button in the toolbar.
- Fixed a bug where the popup menu for inverting the mask couldn't show.
- Fixed a bug where panning or zooming when centering would turn off the silver behenate centering rings
- Fixed a bug in OS X where holding down the centering arrows didn't continuously move the beam center position
- Fixed a bug where the centering arrows wouldn't move the beam center in smaller than integer steps (when holding them down).
- Updated the sync function to greatly increase speed when used with lots of items.
- Updated the superimpose function to greatly increase the speed when used with lots of items.
- The file panel now automatically refreshes when you switch to the file tab.
- Added the ability to use the common keyboard shortcut ctrl-A to select all items in the manipulation, IFT, and SEC lists.
- Fixed an issue with the beam center indicator in the masking panel vanishing when it should not.
- Fixed a bug where error bar color was not maintained when moving a line between different plots.

- Fixed a bug where the error bar color selector for the manipulation and IFT line properties displayed the wrong color in the line properties box.
- Added the ability to change the calculated line name in the SEC line properties box.
- Fixed an issue where, if the legend position had been changed, it reset to the default position when the legend was updated.
- Fixed an issue where the legend shadow went away when legend was updated.
- Added ability to load many more image types using the fabIO library.
- Fixed a bug where the wrong legend label would sometimes be used for SEC curves in windows.

4.7.17 1.0.1

Release date: 2016-05-23

Overview

We're very happy to announce that we are releasing RAW 1.0.1. This is a minor release, concentrating on bug fixes and small changes to the user interface.

There is one very exciting piece of news, which is that this release comes with a prebuilt windows installer (.msi file)! This should make it much easier for those on windows to install the program. We're currently working on a similar thing for OS X.

We are also happy to announce that, to the best of our testing, RAW is compatible with wxpython 3.0 on OS X and Windows (Linux is still a work in progress).

All changes:

- Fixed a bug where online mode without an online filter would load files twice.
- Fixed a bug which caused dammif to crash when run in a directory where the path contained a space.
- Masking panel now defaults to the beamstop mask, not the ROI mask.
- Fixed a bug where if OS X preview files became visible on another system, loading them would crash RAW.
- Fixed an intermittent bug where in scientific linux 6 and wxpython 2.8, occasional calls to the File List would crash RAW.
- Added in error catching, so attempting to load bad .cfg files (either corrupted, or non-RAW files with the same extension) doesn't crash RAW.
- Added in automatic verification of saved .cfg files, to check they can be loaded back into RAW.
- Scrolling with the third mouse button in the Image plot panel, but outside of an image, no longer produces errors in the console.
- Moving manipulation items between plots now respects visibility of the manipulation items.
- The plot axes now automatically refresh when the scale or offset of an item is changed if the axes are set to autoscale.
- Tool tips now work in wxpython 3.0 on OS X
- Selecting the "remove" option in a right click context menu in the Manipulation, IFT, or SEC control tabs no longer causes a seg fault in wxpython 3.0 and OS X.

- Removed MM and conc from Guinier panel, to unify GUi so that MW information is only in the MW panel.
- Added ability to change online mode directory without going offline and back online.
- Added a sort to the online mode, so that files should load in order if multiple files are detected in a given online mode load sequence.
- Added a size check to the online mode load, so that if a file fails to load because it hadn't finished writing/copying, it should load when it is finished.
- Removed the Load button in the SEC control panel .SEC items are now loaded automatically once the file is selected.
- Added an online mode for SEC-SAXS
- Fixed a bug in how SEC-SAXS data was updated when no parameters were being calculated.
- Added a feature so that RAW's online mode will not load in files that RAW saves in the online directory.
- Fixed a bug occasionally preventing the ATSAS directory from being automatically detected.
- Changing control tabs now automatically clears/loads the info window as appropriate.
- Fixed a bug with running datgnom from inside RAW that caused it to fail in certain circumstances.

4.7.18 1.0.0

Release date: 2016-05-06

Overview

Very exciting news, we're moving the project out of beta! That doesn't mean there aren't still bugs, or that we're done adding features. But it does mean that we're happy with the current build (and that we ran out of numbers to increment in beta).

The major new features in this release:

- Added support for running GNOM from RAW
- Added support for running DAMMIF from RAW
- Added support for running DAMAVER from RAW
- Added support for running AMBIMETER from RAW
- Major overhaul of the IFT panel, so it actually works, which involved changing how BIFT is run.

All other changes:

- Added support for reading in FoXS .dat files that have both experimental and model intensities in them
- Fixed a bug where after using the Clear SEC data button RAW could still think there were unsaved changes in the SEC panel
- After removing an item from a plot, the plot axes will automatically resize (unless automatic axes size is turned off in plot options)
- Added a README file in the RAW directory with information on installation and getting help
- Fixed an issue with the porod volume MW calculator crashing if the scattering profile extended to q greater than 0.45 A^-1

- Fixed a bug where MW for RNA was not properly calculated in the SEC plot
- Added ability to save all integrated scattering profiles from a SEC curve as dats
- Fixed an issue where header for save analysis csv files was not using the correct delimiter
- Fixed an issue where beam center did not initially show up correctly in the centering/calibration panel
- Fixed a bug where changing font size for the plot title and axis labels had no effect
- Fixed an issue where the home button in the sec plot didn't work if the calc data existed but was not shown
- Added complied windows 8 exentions, updated compiled windows 7 extensions.
- Various other small bug fixes.

4.7.19 1.0.0b

Release date: 2016-03-24

Overview

We are proud to announce that RAW version 1.0.0b has been released for download! This version includes a huge number of new features and bug fixes.

Our favorite new features are:

- Easy processing of in-line SEC-SAXS data
- New molecular weight panel for calculating mol. weight from the volume of correlation, adjusted porod volume, and absolute scaling.
- AutoRG now available.
- Uncompiled running, which allows RAW to run as long as the appropriate python packages are installed, even if the extension files cannot be compiled.
- Files saved as .dats now automatically save all analysis information in the header, and reload it into RAW when loaded again.

We have also made significant improvements to speed and responsiveness:

- Sped up loading and plotting for large numbers of files on a test machine by a factor of ~30
- Sped up subtraction of large numbers of scattering profiles by a factor of ~4
- Improved responsiveness when large numbers of scattering profiles are plotted.

Also, there are new, up-to-date install guides available for Windows, Mac, and Linux. Check them out in the files tab.

Finally, we have cleaned up both the code repository and the files area.

If you have questions, or feedback, please contact us!

All changes:

SEC-SAXS data processing:

Added capability to process SEC-SAXS data. This included adding a new SEC tab in the control panel, a new SEC plot, and a new SECM data class.

SEC-SAXS data is collected by continuous framing of the detector while sample is being pumped through a column. The output of that column is connected to the SAXS cell. The new RAW addition allows users to load all of the

detector images collected during column elution into a new data type, the SECM. The overall frame intensity is plotted vs. frame number or time, and this should look very similar to the UV-chromatograph that an FPLC produces. The users can then select a range of frames from this curve, and send them to the main plot for processing as normal.

Additionally, the users can select a specified buffer range, and an average window size. The window is then slid across the curve, and the scattering profiles within the window are averaged. The averaged buffer is subtracted from the curves in the window, and radius of gyration, molecular weight, and I(0) are automatically calculated. These are then plotted on the same plot as the 'SAXS chromatograph' (intensity vs. frame #), allowing users to quickly get a feel for what is in each peak they measured.

Major code additions:

- There is now a SEC Panel, SEC Item panel, and SEC Control panel class, based on the Manipulation panel and Item Panel in RAW.py.
- There is now a new plot class in RAWPlot.py, the SECPlot, which allows for multiple axes on the same plot, and handles the various plotting options.
- There is a SECM class in SASM.py, which is the data structure for this new thing.
- There is a new SASCalc.py file, which contains the autorg and automw functions. The autorg is pure python, based on the ATSAS package autorg function. It could probably use some tuning of the various parameters. The automw is also purely python, and based on the Rambo & Tainer correlation volume method for determining molecular weight.
- There is a new save/load format, extension .sec, for saving SEC objects.
- The SEC data is saved when the workspace is saved.
- Various bits and pieces everywhere have been adjusted to accommodate these new panels.

Online mode filtering:

- Added an online mode section in the advanced option panel. This allows you to turn on online filtering, and give a set of strings that allow you to ignore certain files when they enter the watched folder. You can either set a list of strings in the file name to ignore, or a list of strings that must be in the file name, or some combination. You can also set the location where these strings must occur: at the start, end, or anywhere.

MW Panel:

- Added a new analysis panel for finding MW. It has methods for MW by I(0) ratio (also in Guinier plot), MW by the Rambo & Tainer method of the volume of correlation, MW by the Porod volume (corrected by the method of Fisher), and MW by absolute intensity.
- Users can modify default calculation values for the MW in the advanced options MW panel.

Speed improvements:

- Changed the loadAndPlot function so that it only updates the curves on the plot every 20 curves loaded (and at the end), and only updates the legend after all the curves are loaded. On my machine, for ~400 data files (pilatus 100K tiffs) this sped up loading and plotting by ~30x (~40 s vs. 20 minutes & 15 s).
- Changed the subtractItems function so that it only updates the curves on the plot every 20 curves loaded (and at the end), and only updates the legend after all the curves are loaded, as with the _loadAndPlot function. On my machine, this sped up subtraction by ~4x (1 min 7 s vs. 4 min 5 s for ~400 manipulation items).
- Updated online mode to take advantage of the faster plotting, by passing all of the files to be plotted to 'loadAndPlot' at once, rather than one at a time (will only matter if files are coming in faster than the online mode update timer)
- Changed the legend to be off by default (since it significantly hinders performance). Changed the update legend and the legend plot options dialog functions so that this all still works. This seemed to improve load in performance for ~400 data files by ~15% (35 s compared to 40 s).

Uncompiled running:

- Removed all attempts to compile unused extensions.
- Added in try/except cycles for importing and compiling compiled extensions.
- Rewrote compiled extensions scipy.weave code (essentially c code) as pure python.
- Set it so that if RAW is unable to compile extensions, it displays a warning message to users on startup, and then runs with the pure python versions.
- Compilation is particularly an issue on windows, so hopefully this will make deploying to windows much easier (even though the program will run slower). Particularly for versions where a windows installer is not available.
- This required the inclusion of a RAWGlobals file, which contains a variable that notes whether or not the compiled extensions were successfully imported.

Minor changes:

- Switched from PIL to pillow. PIL is not longer under active development, pillow is a fork of PIL that is still supported. Also, pillow is included in the default enthought python installation, while PIL no longer is.
- Fixed an issue where integrated scattering profiles could end up with different numbers of points. This was simply disabling the zero trim command in the integration routine.
- Added in an option to skip points at the end of a scattering profile (identical to the skip at the beginning). This was needed after the removal of the zero trim command when you have entire range of high q masked out (such as to eliminate shadowing from the beam pipe). This setting is accessible in the advanced options calibration dialog.
- Added in a parse function and header profile for log files from the BioCAT beamline.
- Removed the requirement that the beam position be an integer.
- Added in the ability to add a 'zero line' to the main plots (a horizontal line at y=0), in the plot options dialog.
- Fixed the plot options dialog so that it can be opened when no items are loaded in the plot.
- Fixed how the plot options dialog handles legend settings, so it doesn't break if there are no curves already plotted.
- Fixed plot options so that setting x limits and y limits when auto limits is not checked actually affects the graph. Also fixed the limits so that they properly acquire the current axis limits when plot options is opened.
- Made it so that turning auto fitting axes back on forces the plot to autofit the axes when the plot option dialog is closed with the okay button.
- Fixed a bug where the legend would turn on when an item was hid/shown in the manipulation panel, even if the legend was previously turned off.
- Fixed a bug where error bars didn't turn off when an item was hidden in the manipulation panel with error bars turned on.
- Made it so that the borders check boxes in the plot options window actually cause the borders (and tick marks) to turn on and off in the plot.
- Changed the Guinier plot panel so that it automatically updates the MW when the concentration is entered (instead of needing one of the up/down arrows to be hit in the spin control)
- Fixed a bug where the MW of a SASM object wasn't updated when the SASM object was set as a MW standard.
- Fixed a bug in the menu creation of the file browser pane where the right click menu wouldn't open on a mac (wxpython >=2.9.2.4)
- Fixed a bug where the concentration of a sasm object was getting improperly set when the clearinfo function in the information panel was called.

- Made the info panel Rg, MW, and I(0) boxes read only, since user modified values in those boxes aren't saved
- Made the info panel conc box update whenever it gets text, so that if you update the concentration and click on another sasm it still saves the concentration.
- Fixed the options window not opening at the right size.
- Switched to using json to save/load sasm parameter dictionary contents in .dat files. This allows easy saving and loading of dictionaries in human readable format. So now all parameters (header, counters, analysis, etc) are loaded. NOTE: THIS IS NOT STRICTLY BACKWARD COMPATIBLE. RAW can still load old .dat files (and primus .dat files), but it cannot load analysis information out of the old files. This doesn't really affect anything, as for the old files the analysis information didn't load anyways.
- Modified how saving of averaged files history is done. Added in saving of subtracted files history. Now all of the averaging and subtracting manipulation history of a file is saved in the history entry of the parameters dictionary. This works even when you average or subtract files that are already averaged or subtracted. It is mostly human readable in the saved .dat file (though as you get more layers deep in averaging or subtracting it gets hard to tell what is what).
- Fixed a bug where the correct qmin and qmax weren't loading in the Guinier window when a previous Guinier analysis had been done.
- Changed it so that when guinier or mol. weight analysis is done, if the results are different from previous analysis, the scattering profile is marked as modified to denote that the results are not saved.
- Fixed a bug where plot axes didn't auto resize when curves were moved from the top main plot to the bottom main plot and vice versa.
- Fixed a bug where selecting 'Help!' in the help menu crashed RAW. No in-program help is yet available, but a message dialog now tells the user to look for help on the raw project homepage.
- Set 'okay' button to be selected by default in the welcome window.
- Fixed a bug where on mac, last saved settings wouldn't load from the dialog on startup (this may have also been affecting other OSes).
- Enabled normalization by ROI counter using an 'ROI counter mask' (formerly called a transparent beamstop mask).
- Fixed a bug where minor tick marks weren't turning off for log axes that weren't shown (such as top and right) (I believe this was introduced by an updated version of matplotlib, I don't remember seeing it before).
- Fixed the logarithmic image scale display in the image panel. It works now, and is enabled.
- Disabled nexus support to remove error on starting raw (can be easily re-enabled, it is simply commented out in a couple places in SASFileIO).
- Updated the manipulation and IFT item saves so that it offers the choice to rename the file when saving a single file, and so that there are more explicit instructions when saving multiple files.
- Fixed a bug in the rebin function, where it wasn't setting the qrange according to the original sasm.
- Fixed a bug where comparison of q vectors to test for subtraction was done by length rather than elementwise by q.
- When scattering profiles with different q vectors are subtracted, choosing to force the subtraction now actually carries out the subtraction (with appropriate matching/rebinning of the q vectors).
- Fixed a bug so that the average function now tests the q vectors point wise, rather than by length, to make sure they actually match.
- Added a feature to export all analysis information from sasm objects as an alternative to selecting which analysis features you want to save.

- Update the old save analysis feature to be called 'save item info' in the menu, since it can save things that aren't analysis. Updated the layout of that window a little bit, and added ability to save the new MW analysis info into the item.
- Added scattering profile manipulation into the tools menu: average, subtract, merge, rebin, interpolate, normalize by concentration, change q scale, set as MW standard.
- Upon quitting, RAW now checks whether there are unsaved changes to manipulation or SEC items, and asks for confirmation of quitting if there are.
- Added show image, show data, show header options to the view menu.
- File list maintains sort order upon refresh.
- Doing a Guinier fit on a scattering profile that is all zeros no long crashes RAW.
- Subtraction can handle mismatched q vectors.
- Autosave for averaged and subtracted files now available.
- Features supporting SAXSLab300 image format now available.

Index

Symbols

$PATH, 14

E

environment variable
    $PATH, 14