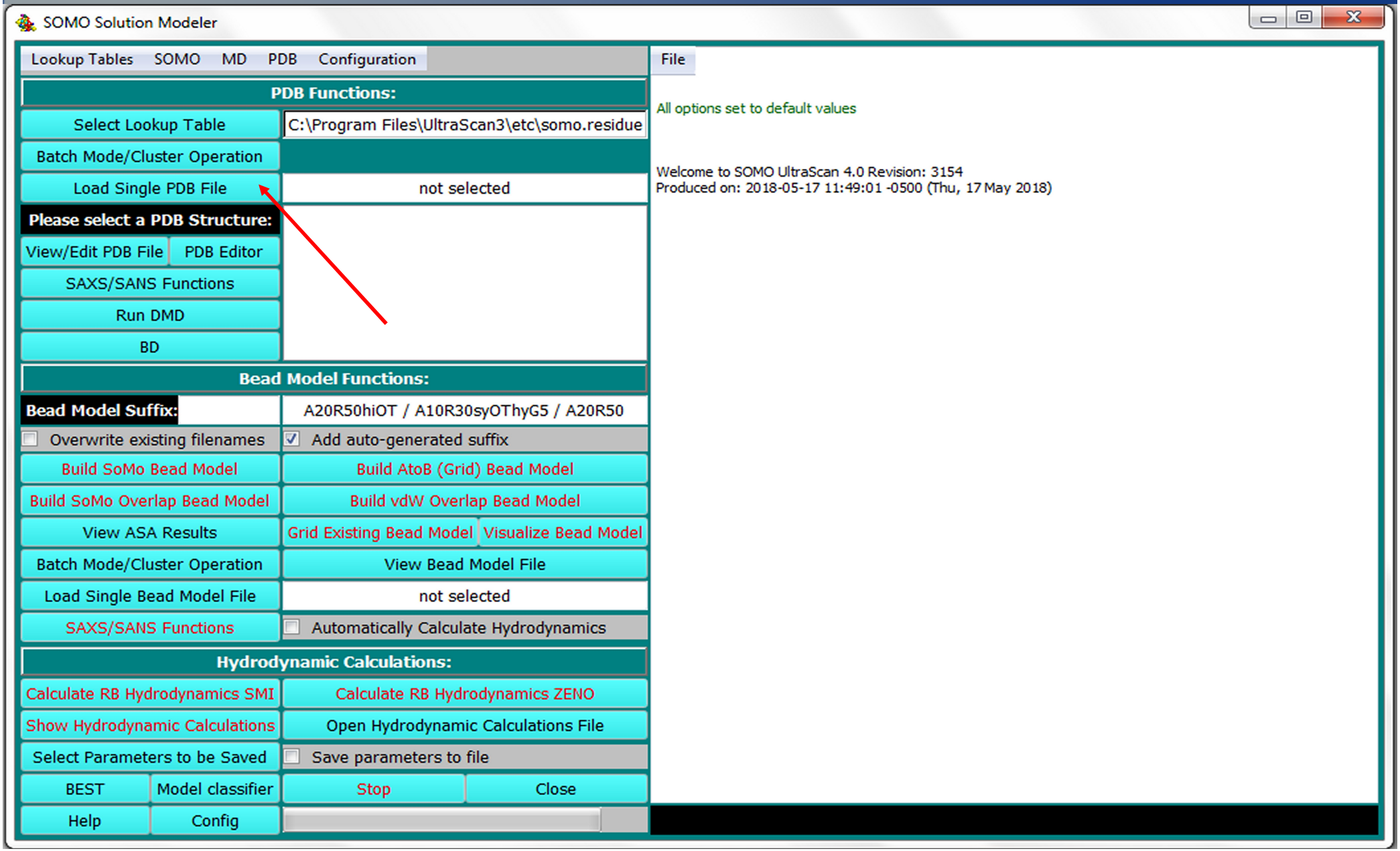
The Small-Angle Scattering and HPLC-SAXS modules of the US-SOMO software suite

M. Rocco[a]
J. Pérez[b]
P. Vachette[c]
E. H. Brookes[d]

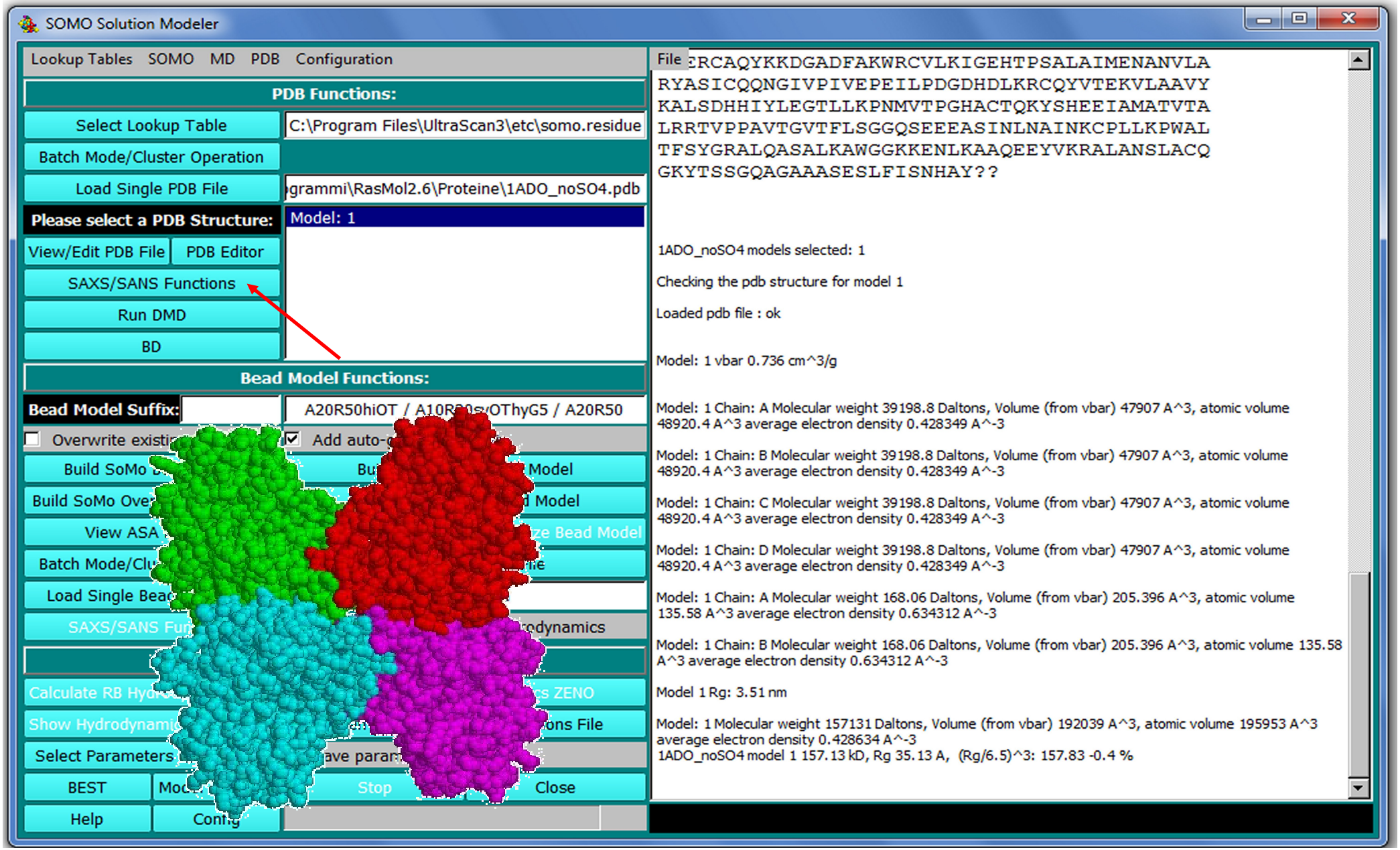
[a]*Biopolimeri e Proteomica, IRCCS Ospedale Policlinico San Martino, Genova, Italy*
[b]*Beamline SWING, Synchrotron SOLEIL, Gif-sur-Yvette, France*
[c]*I2BC, UMR 9198, Orsay, France*
[d]*Dept. of Biochemistry, U. of Texas Health Science Center at San Antonio, San Antonio, Texas*

http://somo.aucsolutions.com/

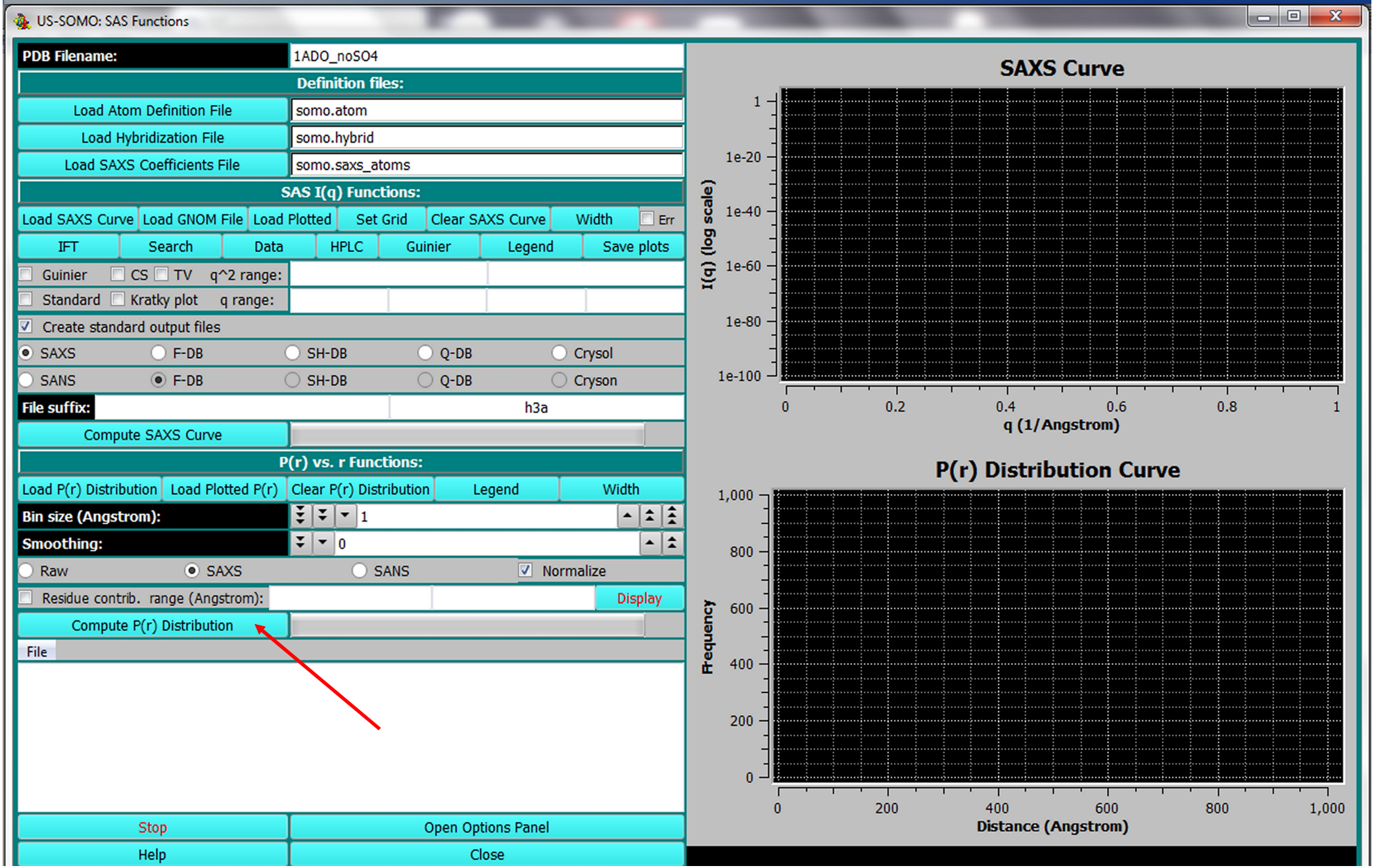
US-SOMO main window



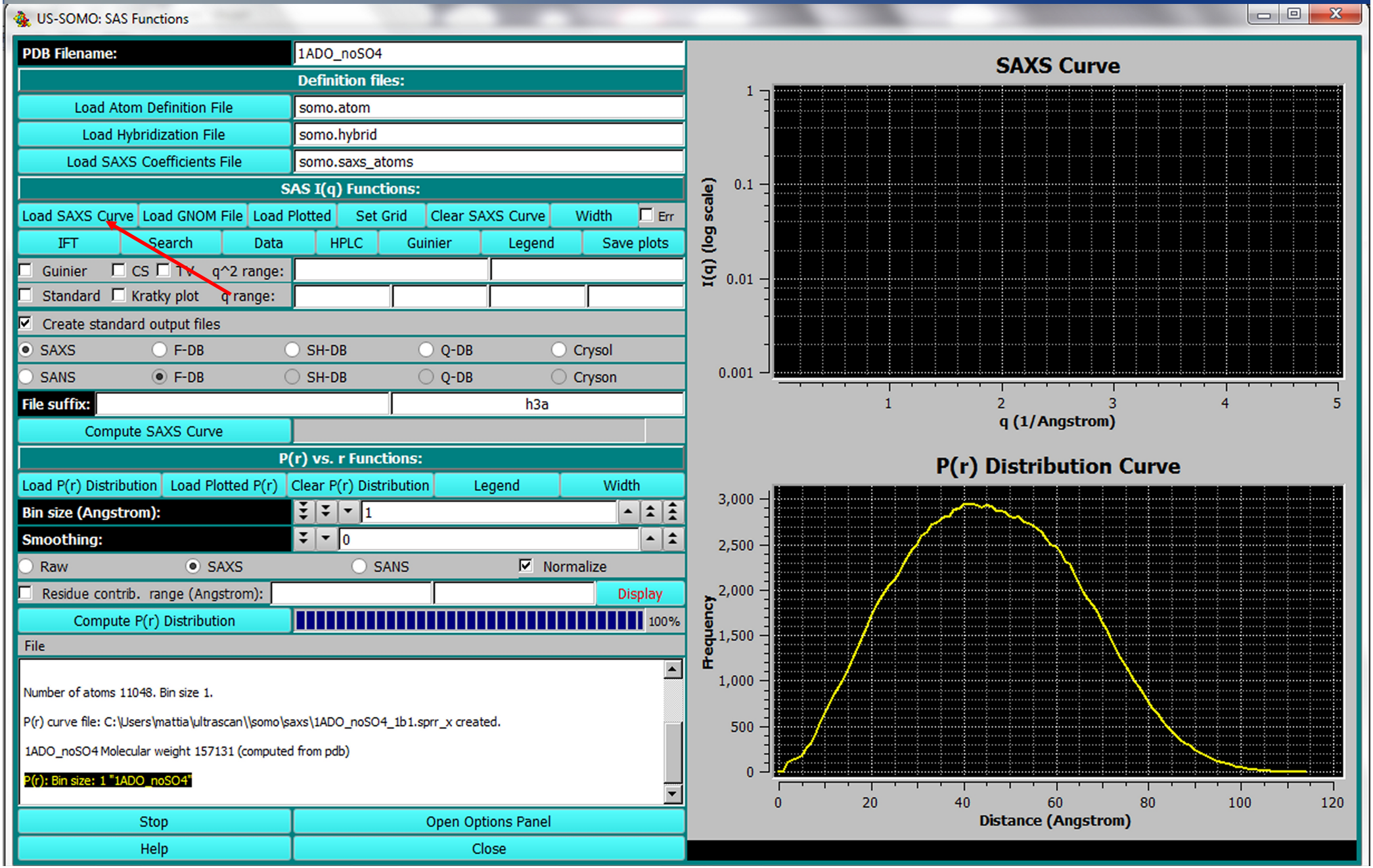
US-SOMO main window



US-SOMO SAS MODULE



US-SOMO SAS MODULE



US-SOMO SAS MODULE

US-SOMO: SAS Functions

PDB Filename: 1ADO_noSO4

Definition files:

Load Atom Definition File: somo.atom
Load Hybridization File: somo.hybrid
Load SAXS Coefficients File: somo.saxs_atoms

SAS I(q) Functions:

Load SAXS Curve | Load GNOM File | Load Plotted | Set Grid | Clear SAXS Curve | Width | Err
IFT | Search | Data | HPLC | Guinier | Legend | Save plots

Guinier CS TV q^2 range:
Standard Kratky plot q range:
Create standard output files

SAXS F-DB SH-DB Q-DB Crysol
SANS F-DB SH-DB Q-DB Cryson

File suffix: h3a

Compute SAXS Curve

P(r) vs. r Functions:

Load P(r) Distribution | Load Plotted P(r) | Clear P(r) Distribution | Legend | Width

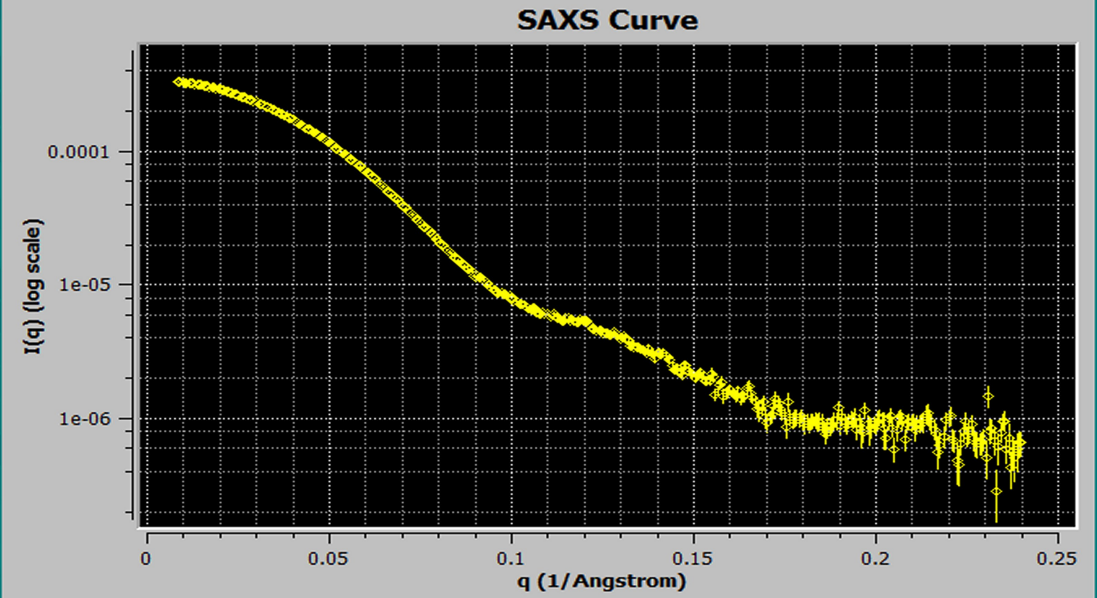
Bin size (Angstrom): 1
Smoothing: 0

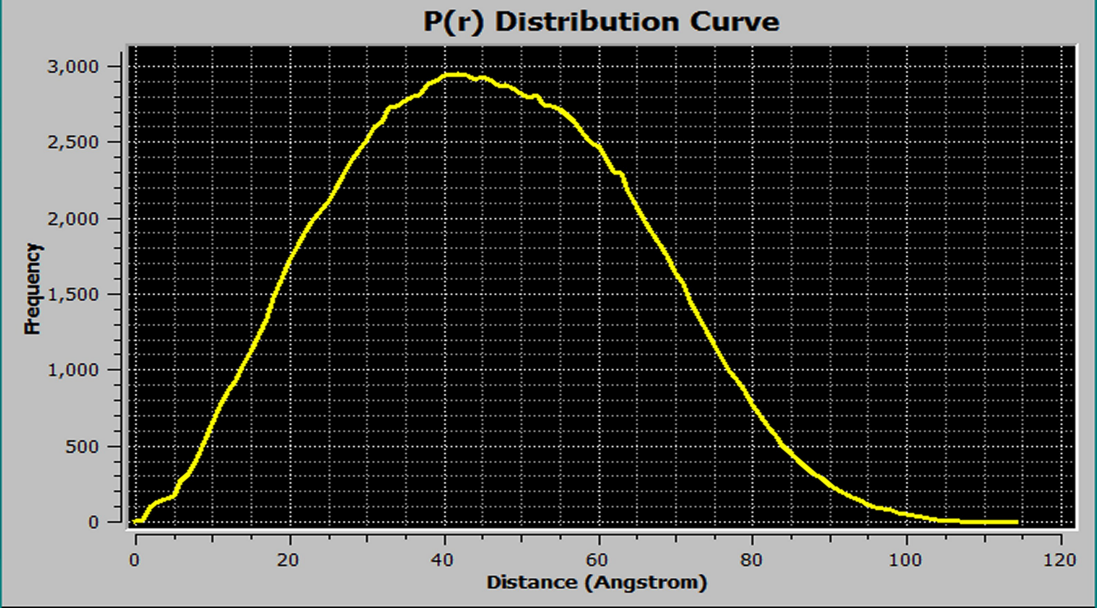
Raw SAXS SANS Normalize
Residue contrib. range (Angstrom): Display
Compute P(r) Distribution 100%

File

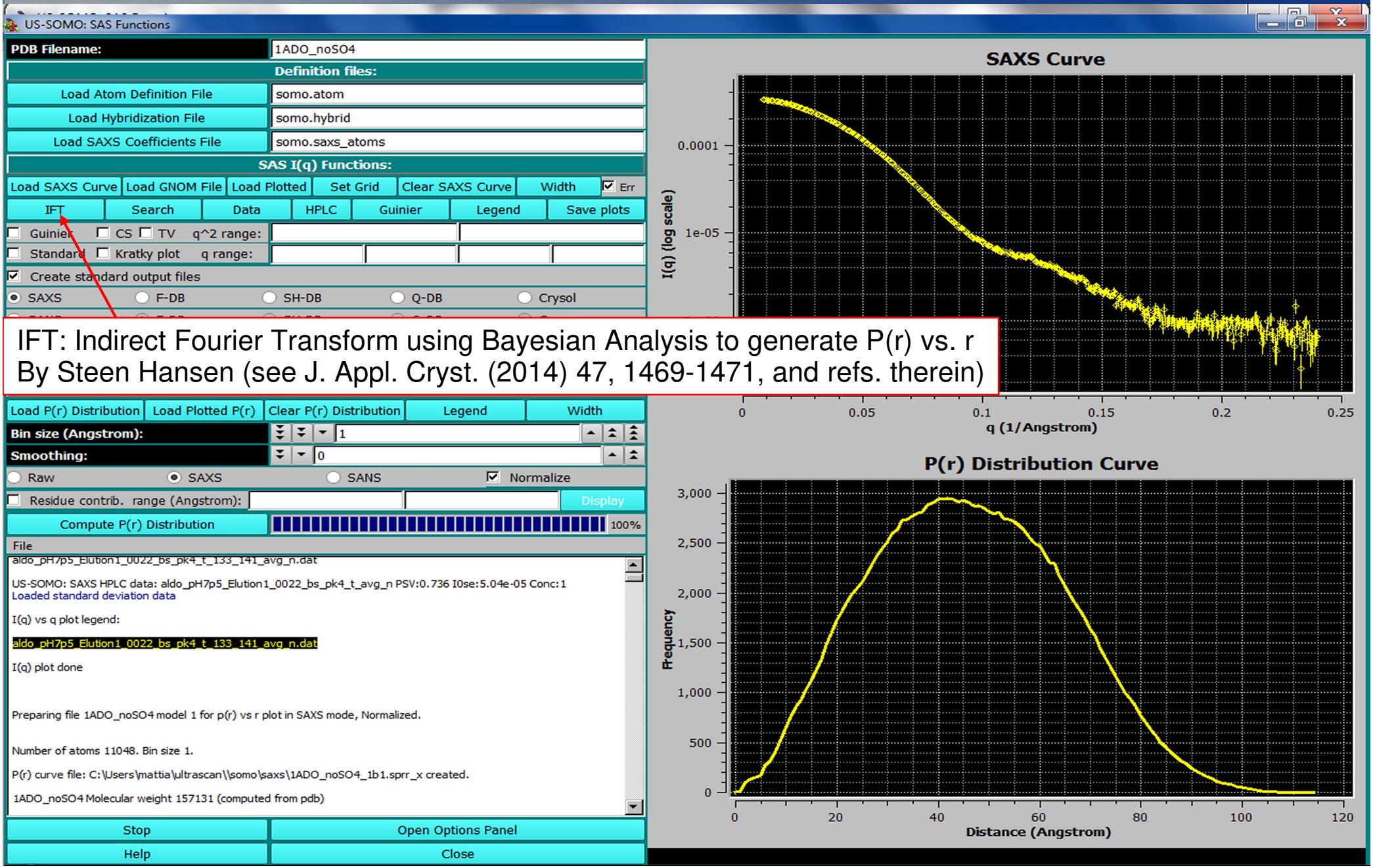
```
aldo_pH7p5_Elution1_0022_bs_pk4_t_133_141_avg_n.dat

US-SOMO: SAXS HPLC data: aldo_pH7p5_Elution1_0022_bs_pk4_t_avg_n PSV:0.736 I0se:5.04e-05 Conc:1
Loaded standard deviation data

I(q) vs q plot legend:

aldo_pH7p5_Elution1_0022_bs_pk4_t_133_141_avg_n.dat

I(q) plot done


Preparing file 1ADO_noSO4 model 1 for p(r) vs r plot in SAXS mode, Normalized.

Number of atoms 11048. Bin size 1.

P(r) curve file: C:\Users\mattia\ultrascan\somo\saxs\1ADO_noSO4_1b1.sprr_x created.

1ADO_noSO4 Molecular weight 157131 (computed from pdb)
```

Stop | Open Options Panel
Help | Close

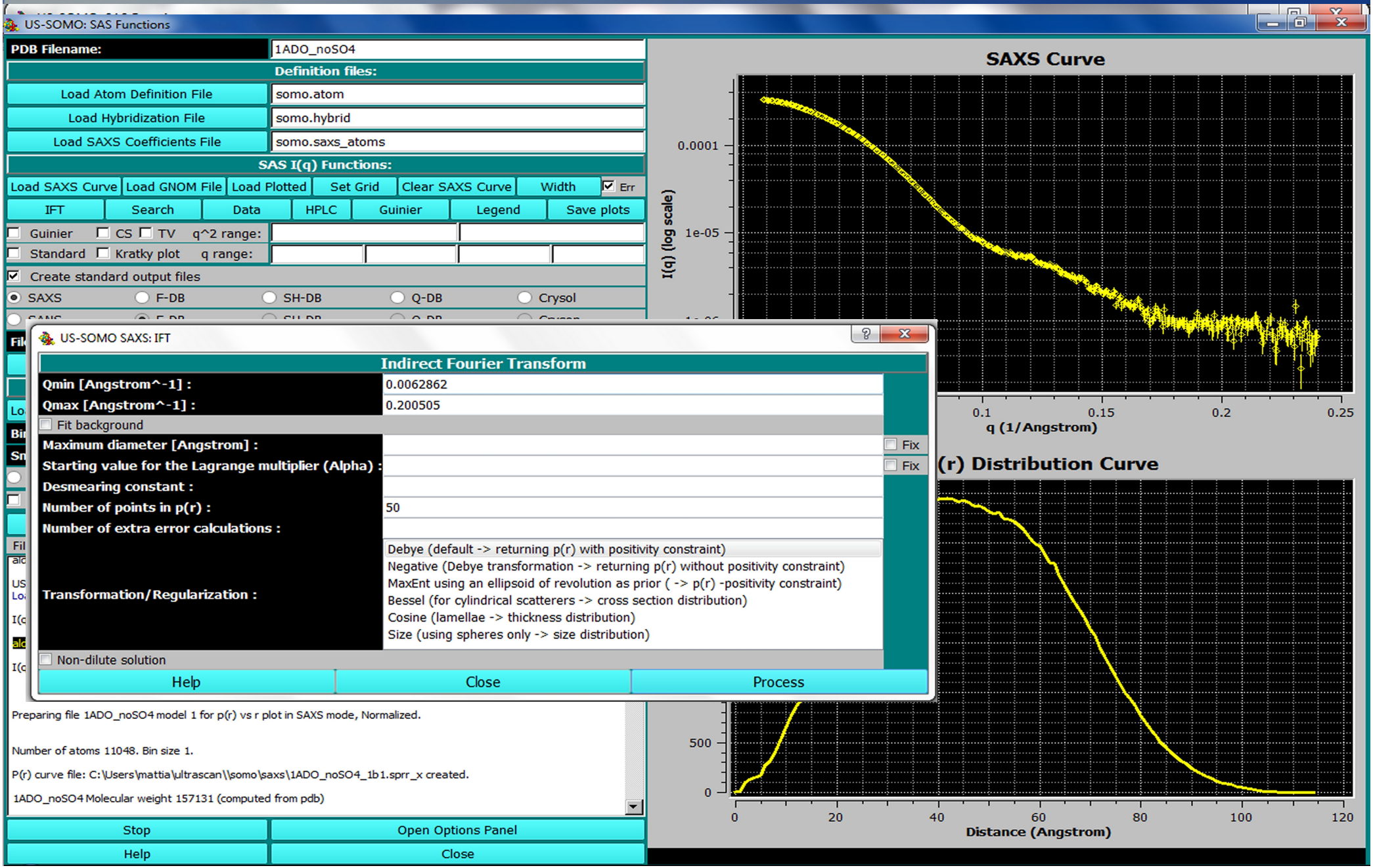




US-SOMO SAS MODULE



US-SOMO SAS MODULE



US-SOMO SAS MODULE

US-SOMO: SAS Functions

PDB Filename: 1ADO_noSO4

Definition files:

Load Atom Definition File | somo.atom
Load Hybridization File | somo.hybrid
Load SAXS Coefficients File | somo.saxs_atoms

SAS I(q) Functions:

Load SAXS Curve | Load GNOM File | Load Plotted | Set Grid | Clear SAXS Curve | Width | Err
IFT | Search | Data | HPLC | Guinier | Legend | Save plots

Guinier | CS | TV | q^2 range:
Standard | Kratky plot | q range:

Create standard output files

SAXS | F-DB | SH-DB | Q-DB | Crysol
SANS | F-DB | SH-DB | Q-DB | Cryson

File suffix: h3a

Compute SAXS Curve

P(r) vs. r Functions:

Load P(r) Distribution | Load Plotted P(r) | Clear P(r) Distribution | Legend | Width

Bin size (Angstrom): 1
Smoothing: 0

Raw | SAXS | SANS | Normalize

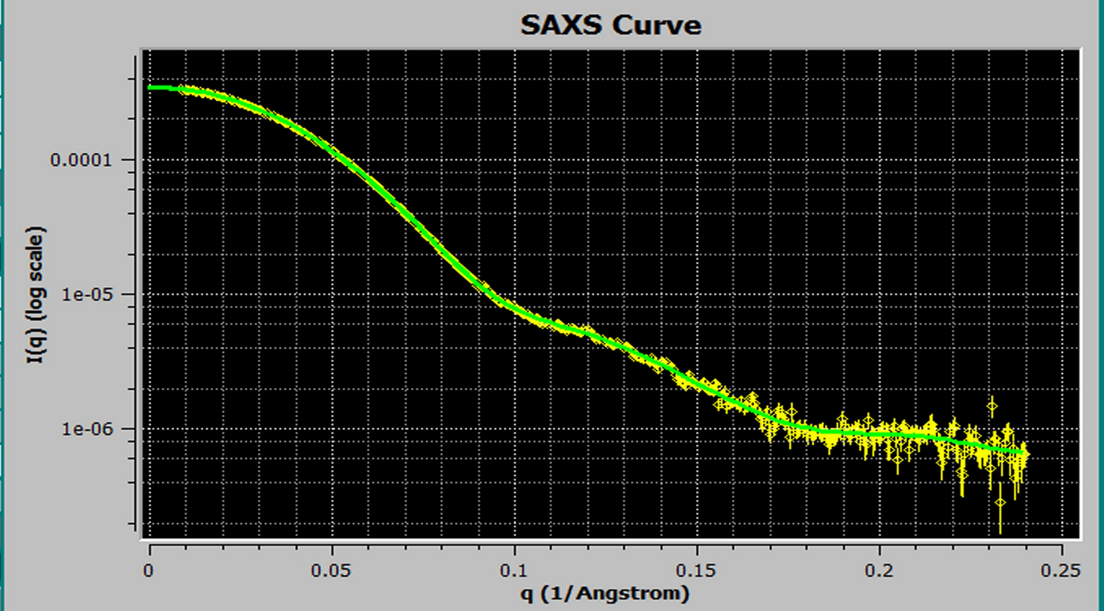
Residue contrib. range (Angstrom): | Display

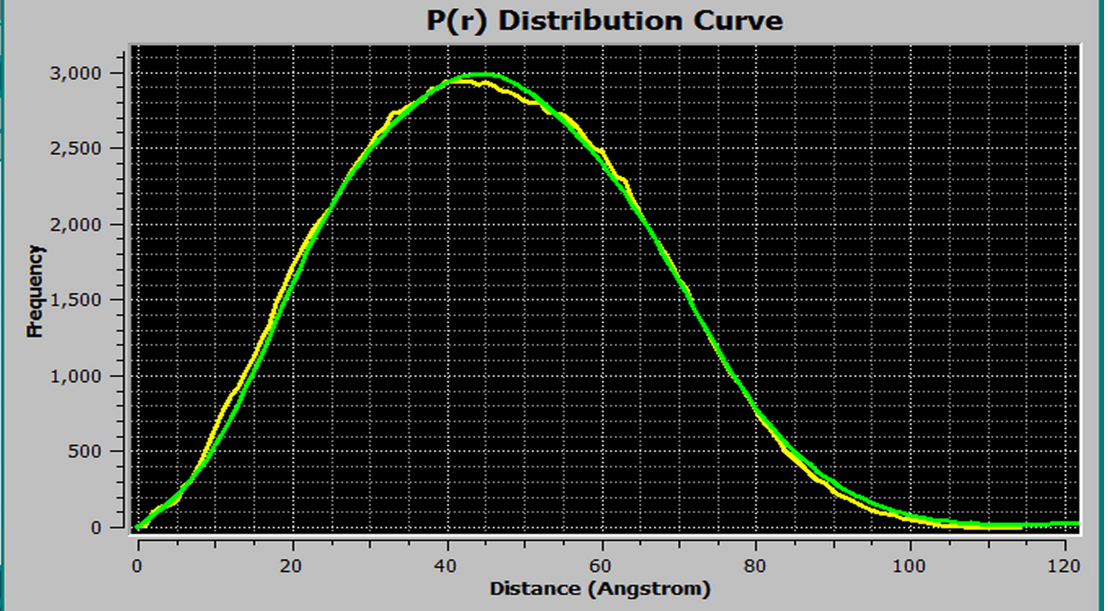
Compute P(r) Distribution | 100%

File

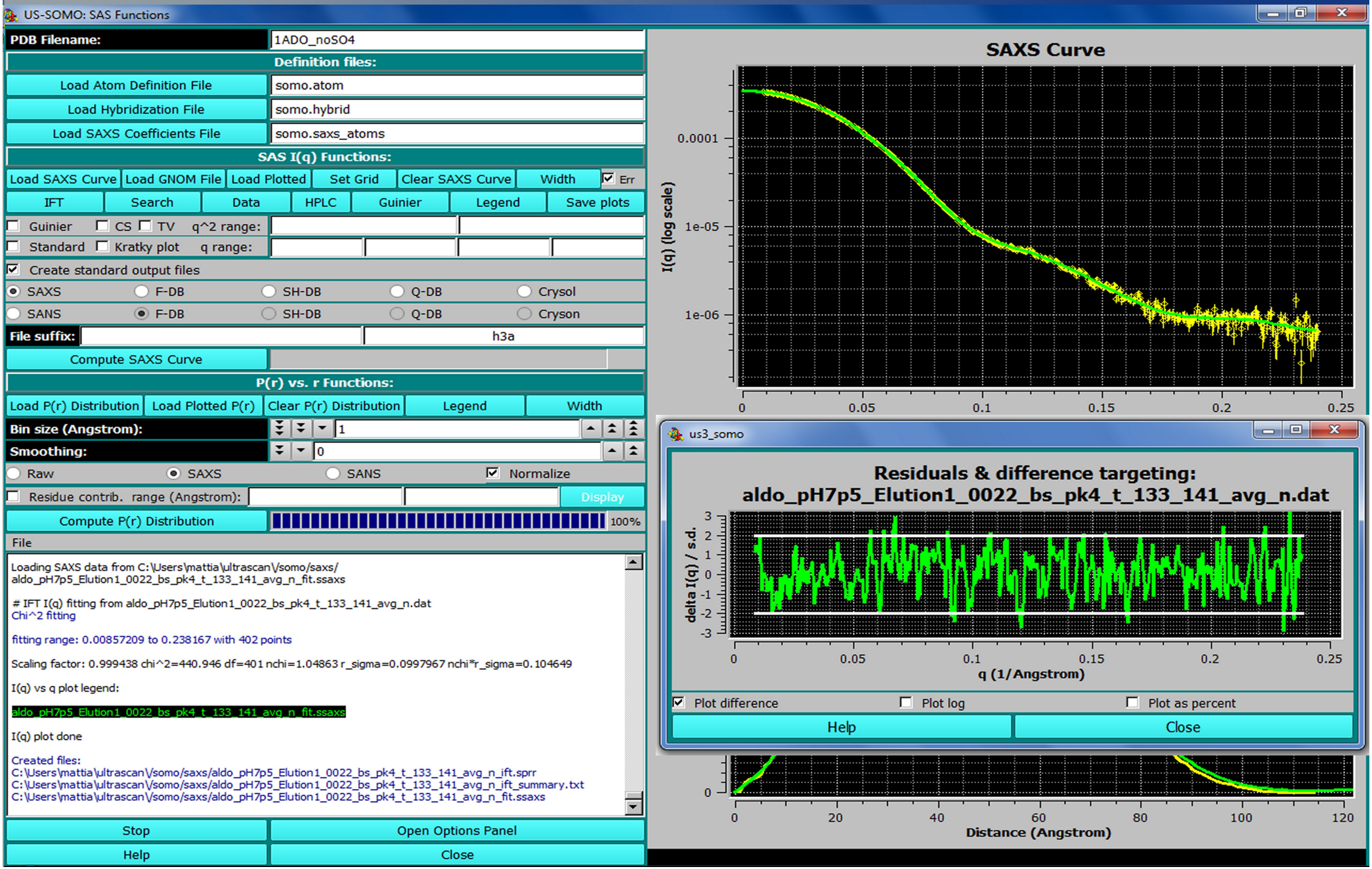
```
Loading SAXS data from C:\Users\mattia\ultrascan\somo/saxs/
aldo_pH7p5_Elution1_0022_bs_pk4_t_133_141_avg_n_fit.ssaxs

# IFT I(q) fitting from aldo_pH7p5_Elution1_0022_bs_pk4_t_133_141_avg_n.dat
Chi^2 fitting

fitting range: 0.00857209 to 0.238167 with 402 points

Scaling factor: 0.999438 chi^2=440.946 df=401 nchi=1.04863 r_sigma=0.0997967 nchi*r_sigma=0.104649

I(q) vs q plot legend:

aldo_pH7p5_Elution1_0022_bs_pk4_t_133_141_avg_n_fit.ssaxs

I(q) plot done

Created files:
C:\Users\mattia\ultrascan\somo/saxs/aldo_pH7p5_Elution1_0022_bs_pk4_t_133_141_avg_n_ift.sprr
C:\Users\mattia\ultrascan\somo/saxs/aldo_pH7p5_Elution1_0022_bs_pk4_t_133_141_avg_n_ift_summary.txt
C:\Users\mattia\ultrascan\somo/saxs/aldo_pH7p5_Elution1_0022_bs_pk4_t_133_141_avg_n_fit.ssaxs
```

Stop | Open Options Panel
Help | Close

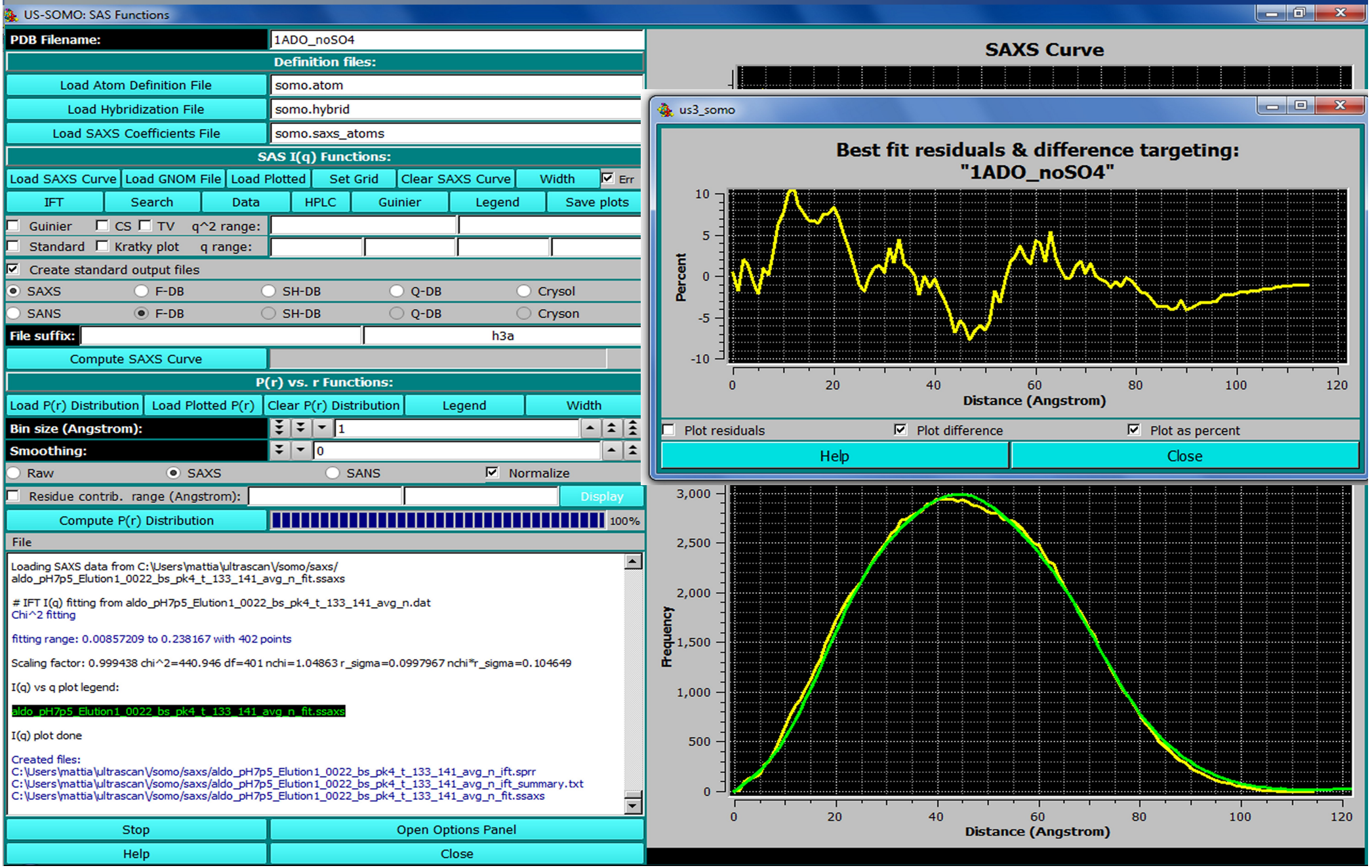




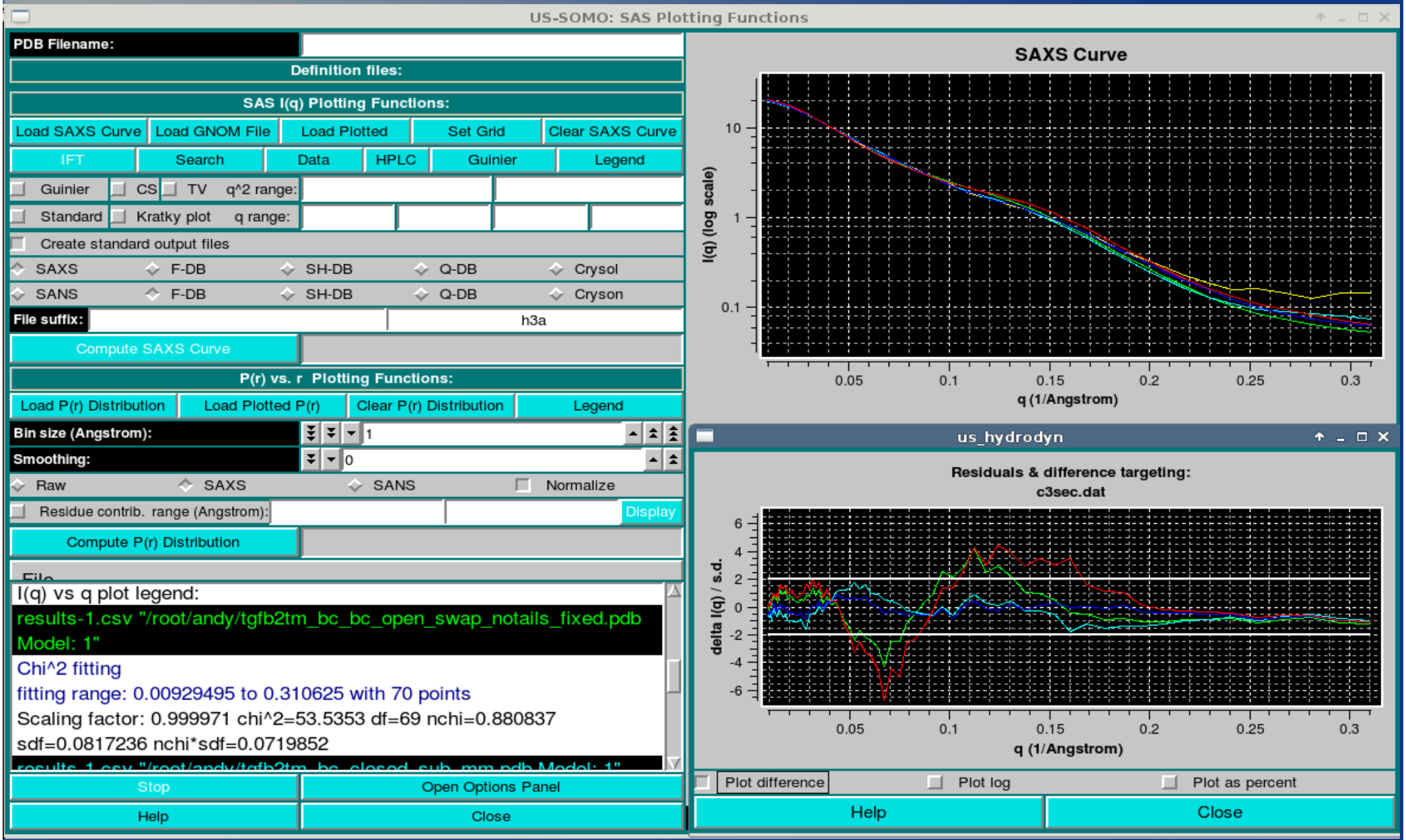
US-SOMO SAS MODULE



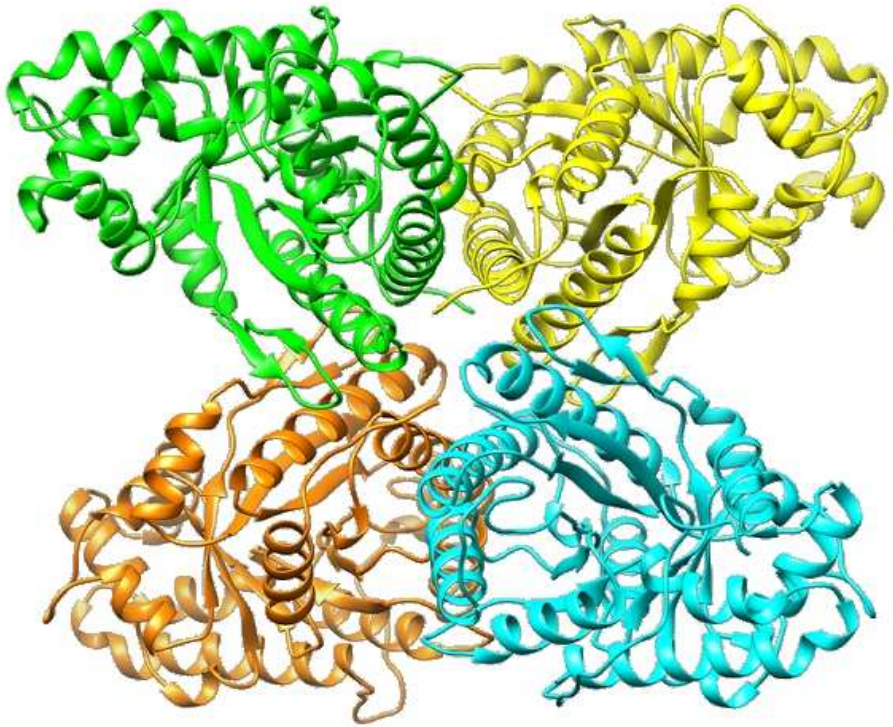
US-SOMO SAS MODULE



US-SOMO SAS NNLS utility

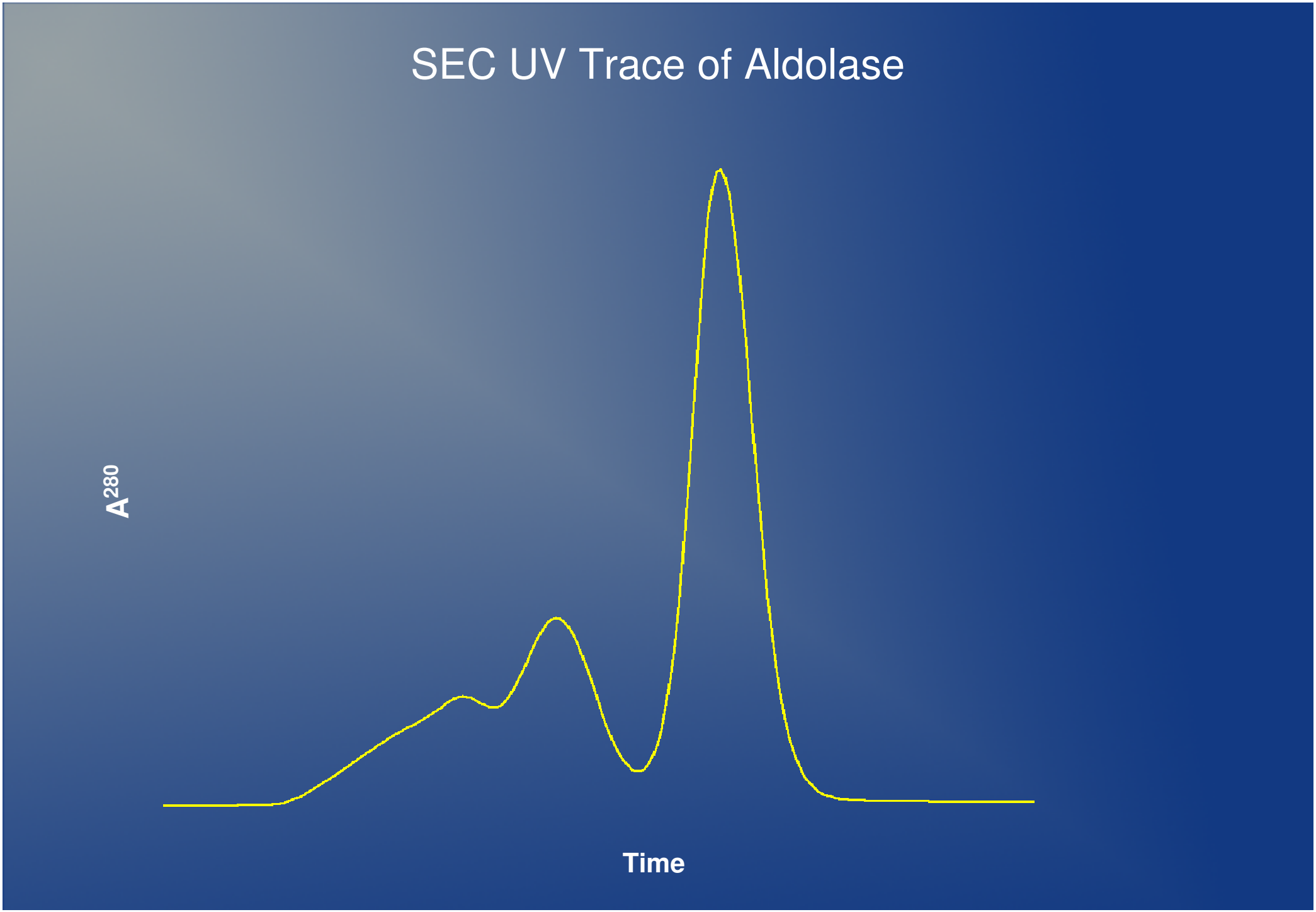


The HPLC-SAXS module: SEC-SAXS of aldolase

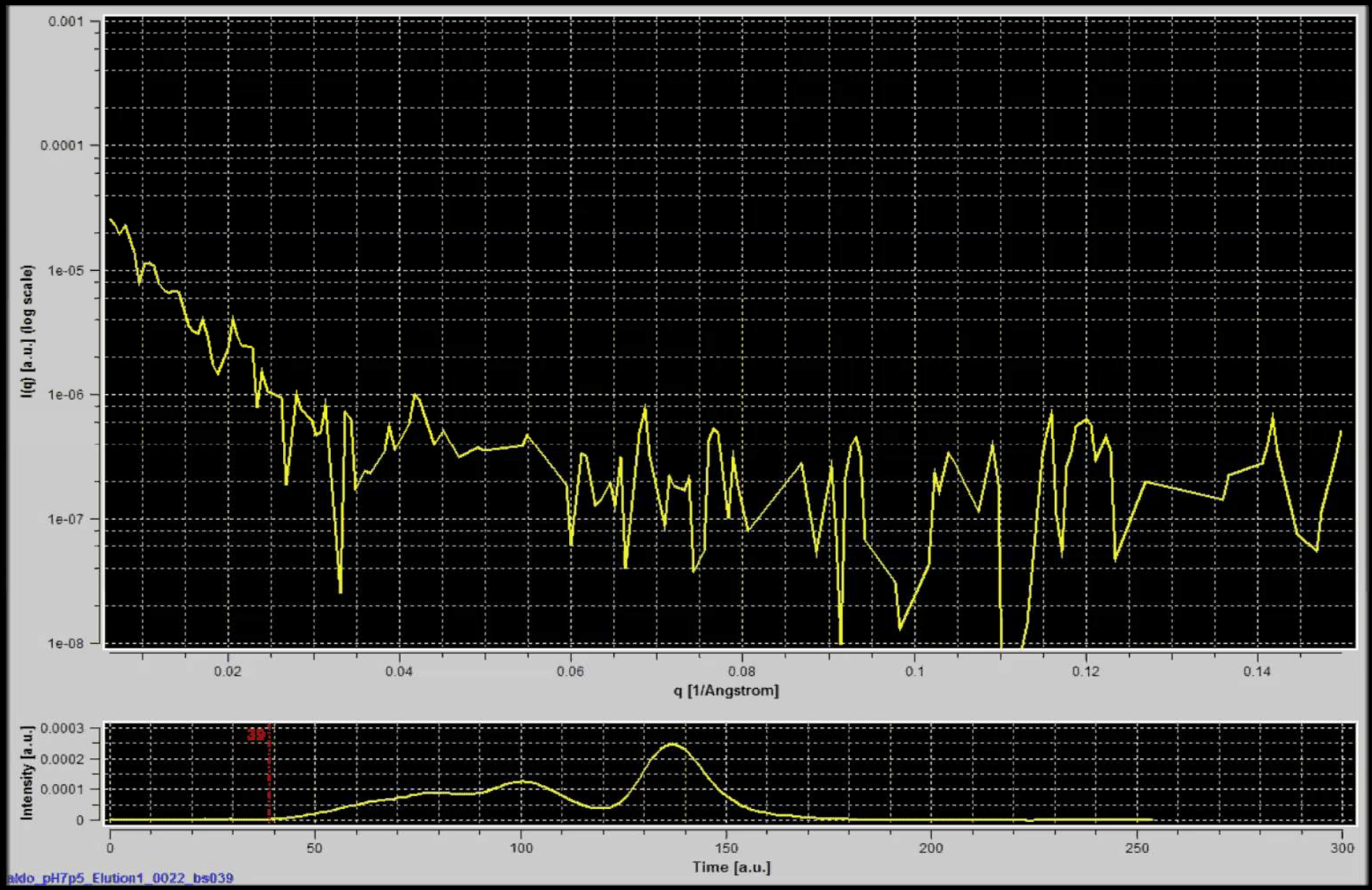


A homotetramer whose crystal structure has been solved

SEC UV Trace of Aldolase

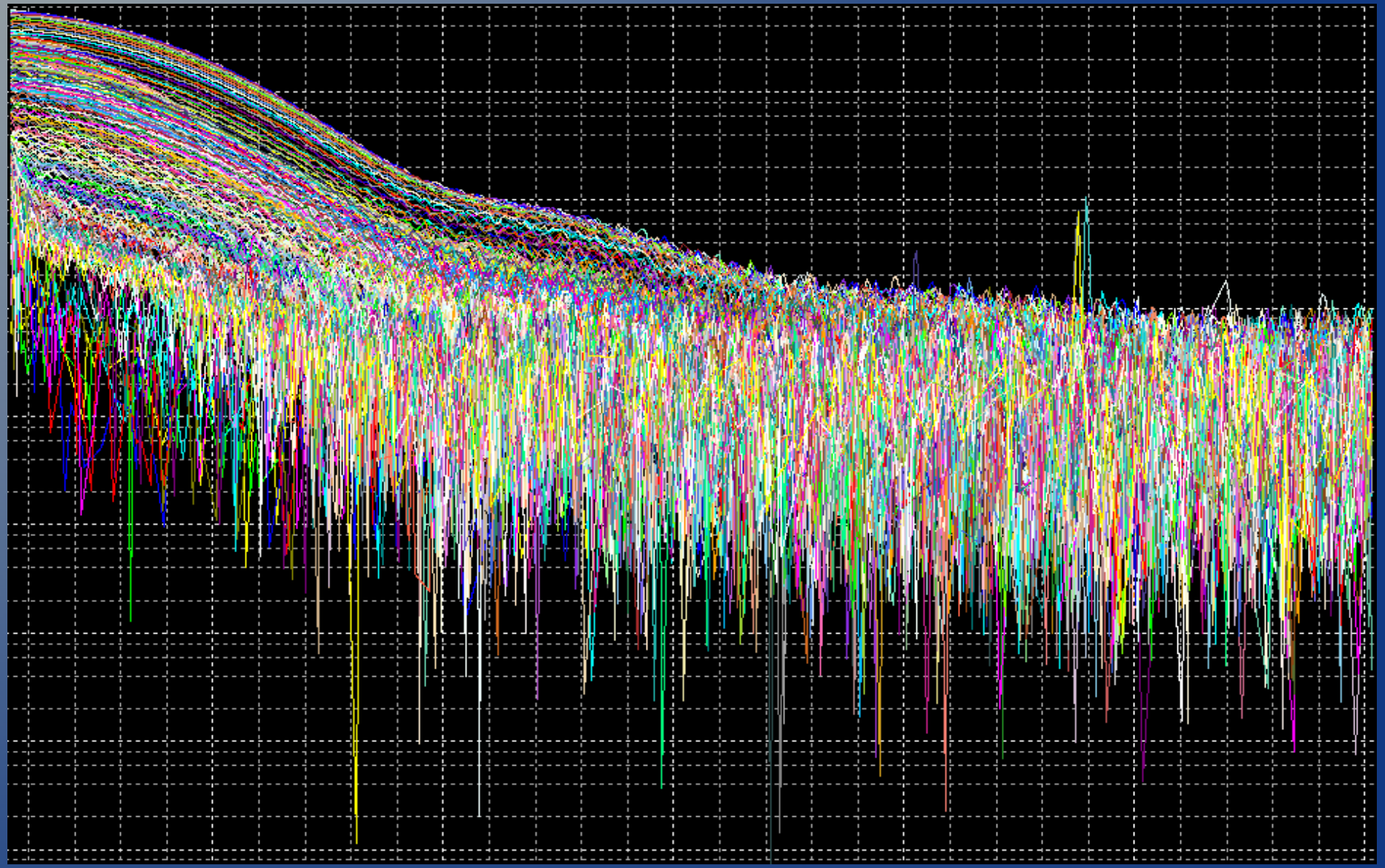


SEC-SAXS : aldolase I(q) vs. q elution profile

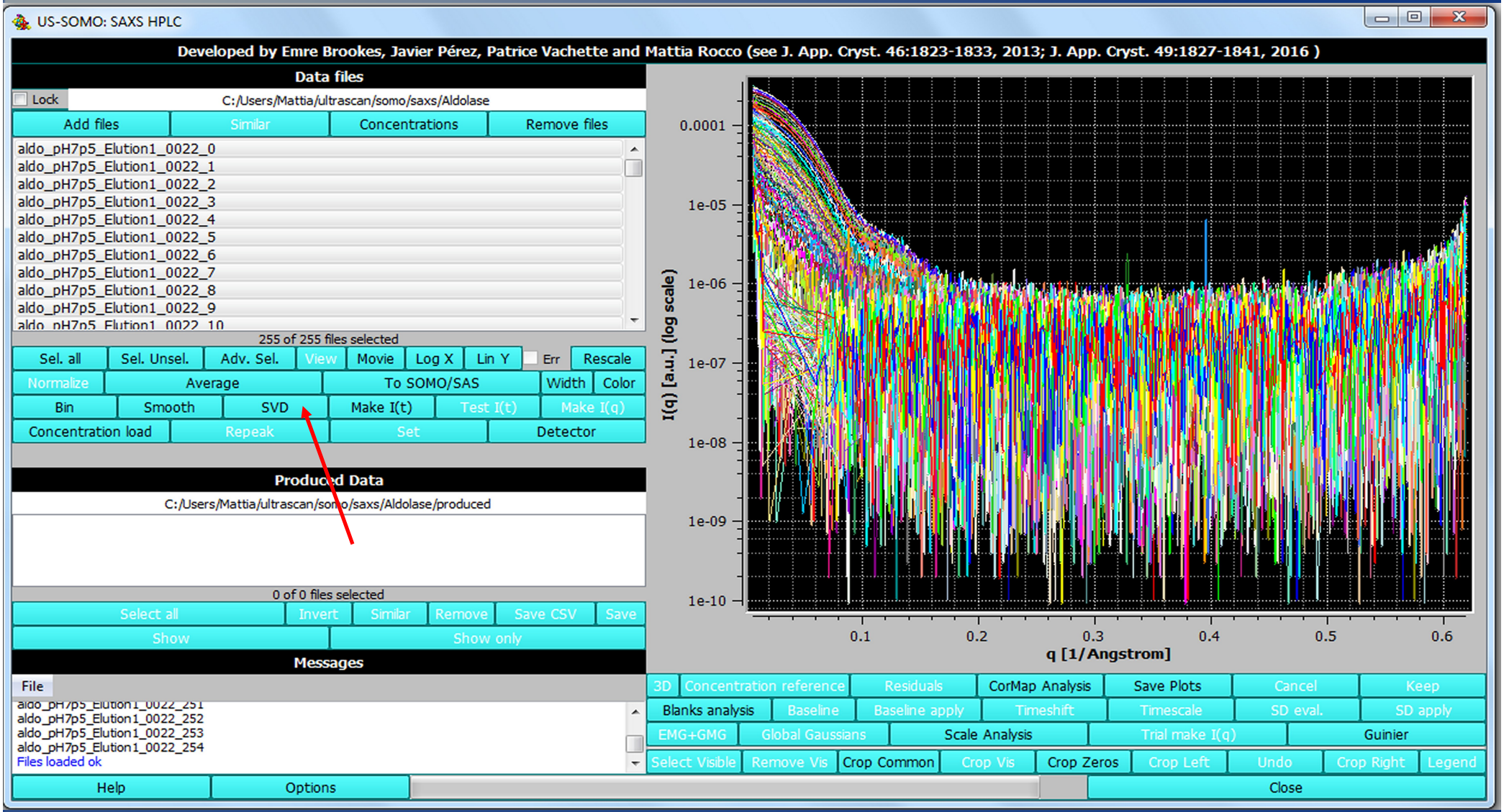


US-SOMO HPLC-SAXS module

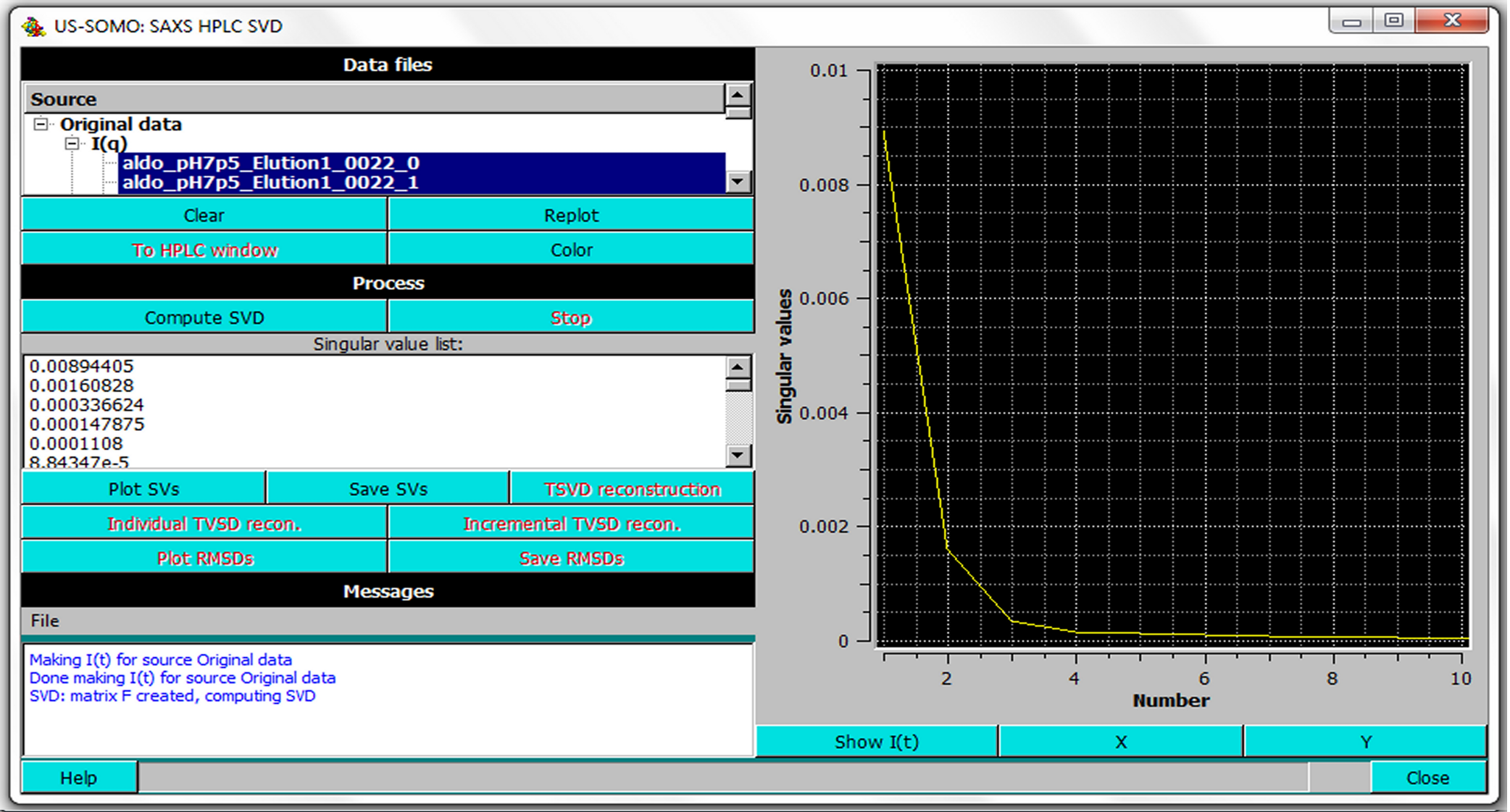
SEC-SAXS of aldolase: set of I(q) data each corresponding to a specific time or frame



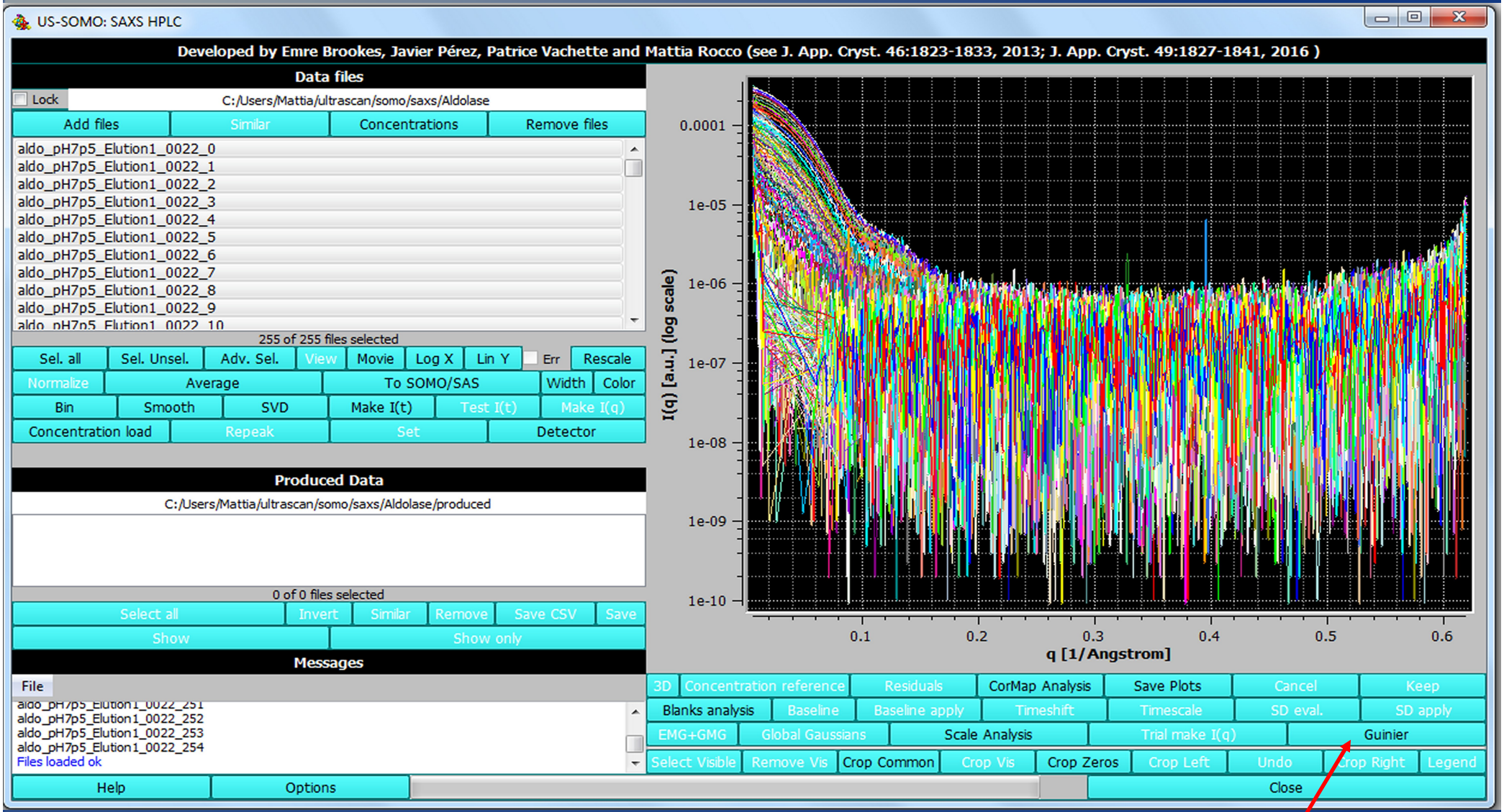
HPLC-SAXS module with loaded I(q) vs. q data



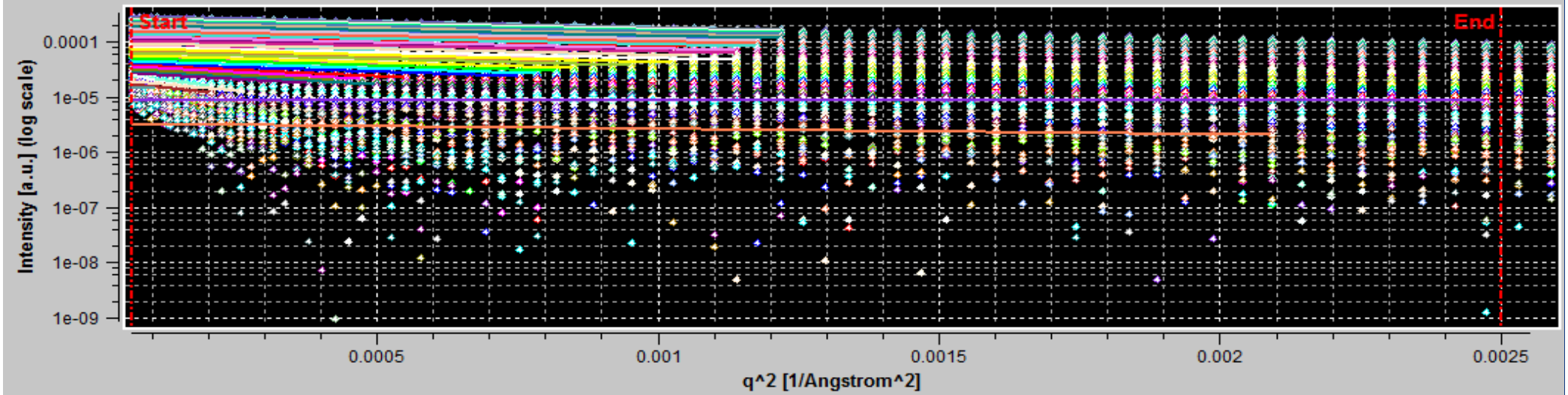
HPLC-SAXS module with loaded I(q) vs. q data

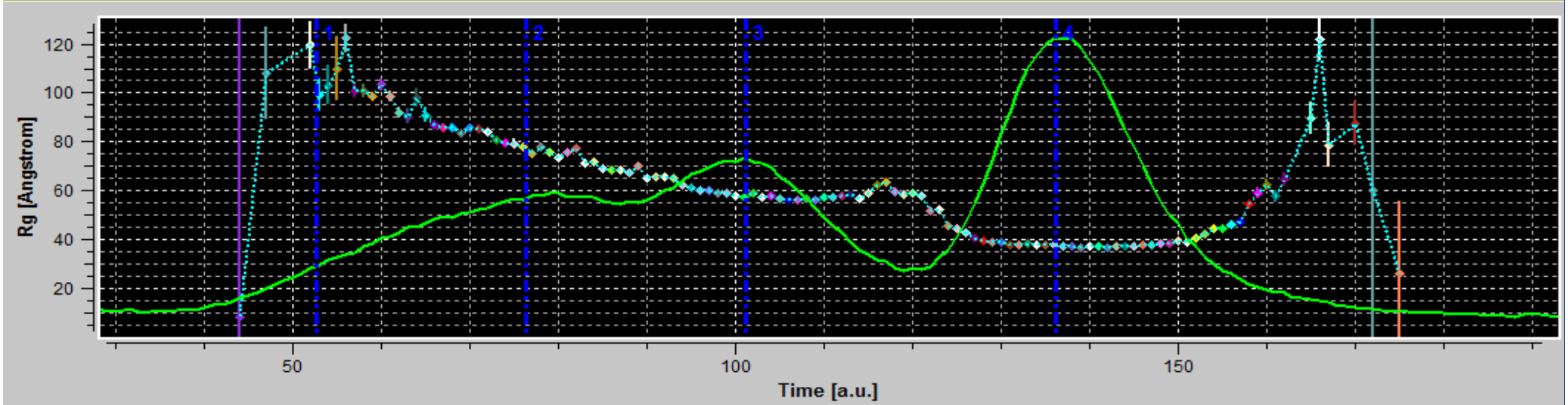


HPLC-SAXS module with loaded I(q) vs. q data

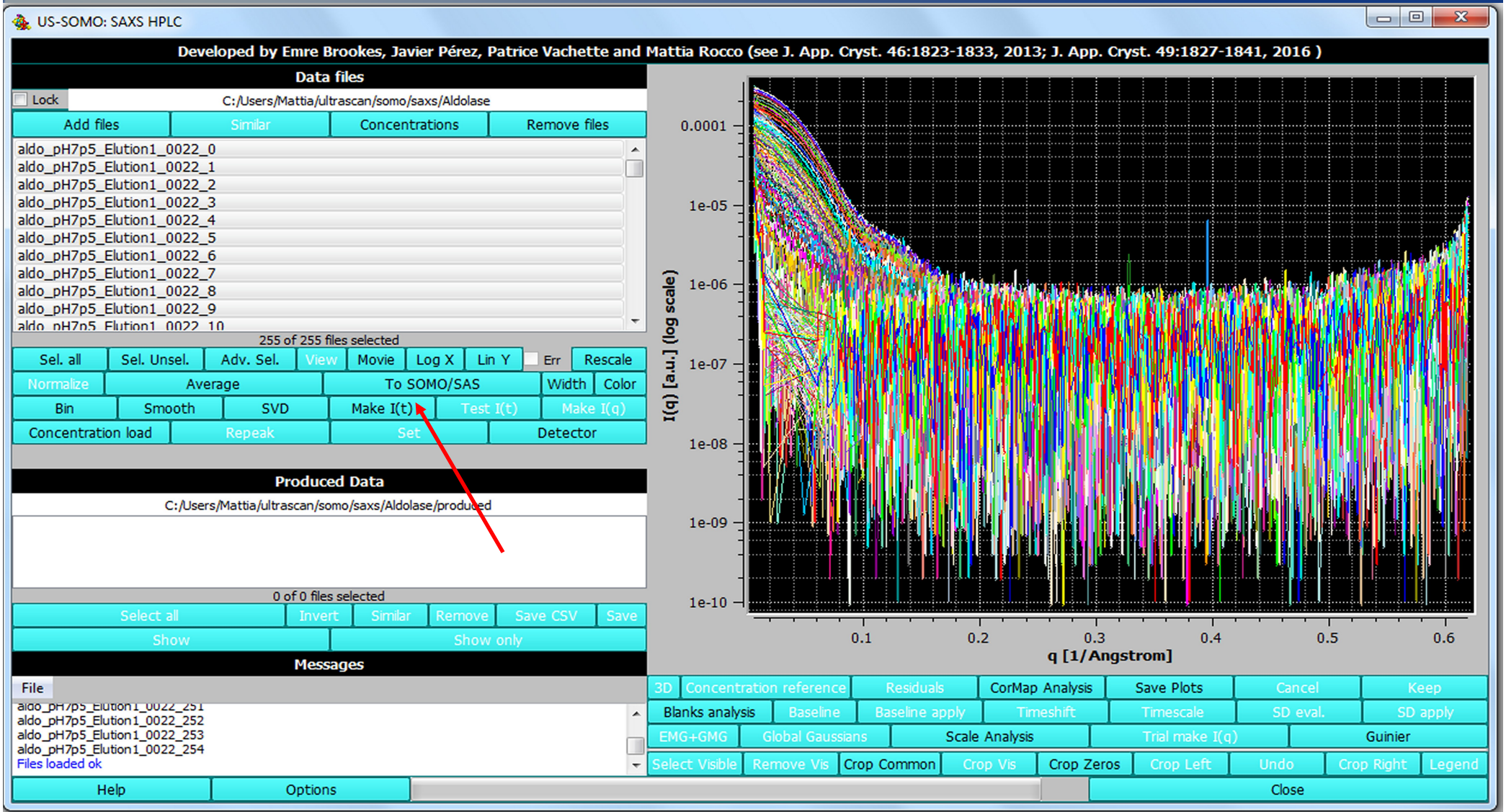


Aldolase Rg – original data

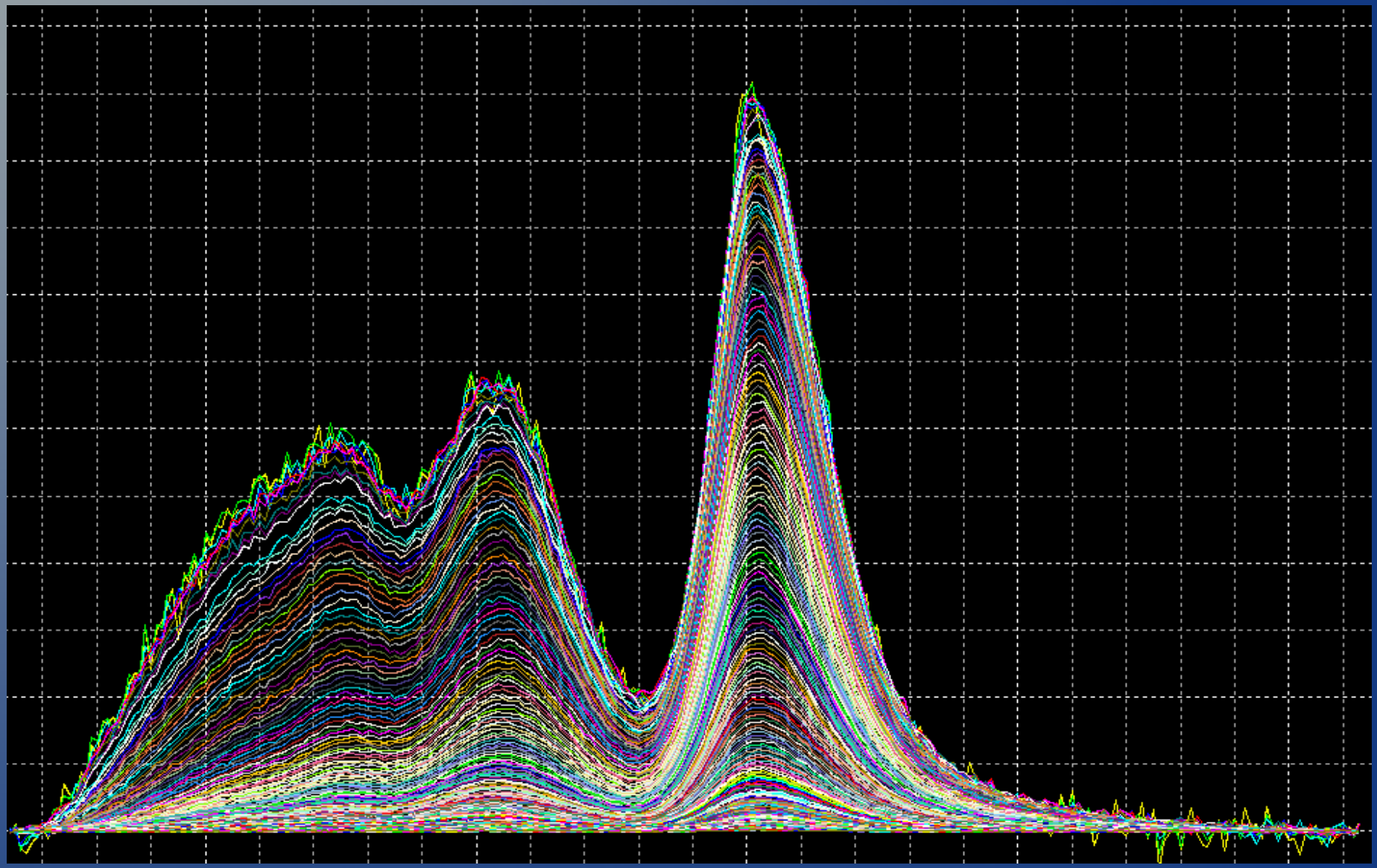




HPLC-SAXS module with loaded I(q) vs. q data



Transposition of the aldolase SEC-SAXS data: set of I(t), each corresponding to a specific *q* value



Gaussian functions for SEC-SAXS deconvolution

- Normal Gaussian:

$$y = \frac{a_0}{\sqrt{2\pi}a_2}\exp\left[-\frac{1}{2}\left(\frac{x - a_1}{a_2}\right)^2\right]$$

- Exponentially modified Gaussian (EMG):

$$y = \frac{a_0}{2a_3}\exp\left(\frac{a_2^2}{2a_3^2} + \frac{a_1 - x}{a_3}\right)\left[\mathrm{erf}\left(\frac{x - a_1}{\sqrt{2}a_2} - \frac{a_2}{\sqrt{2}a_3}\right) + \frac{a_3}{|a_3|}\right]$$

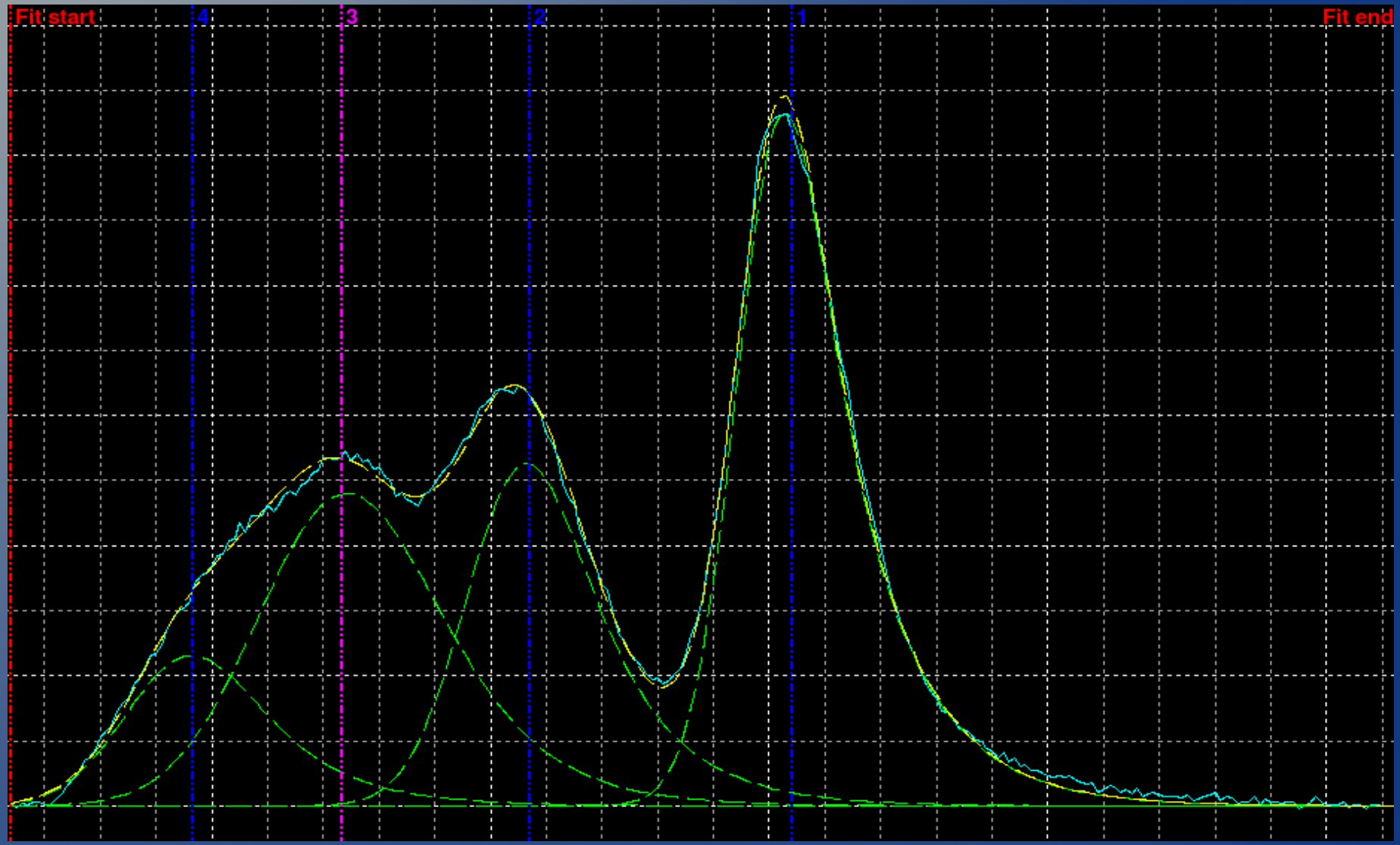
- Half-Gaussian modified Gaussian (GMG):

$$y = \frac{a_0\exp\left(-\frac{1}{2}\frac{(x - a_1)^2}{a_3^2 + a_2^2}\right)\left[1 + \mathrm{erf}\left(\frac{a_3(x - a_1)}{\sqrt{2}a_2\sqrt{a_3^2 + a_2^2}}\right)\right]}{\sqrt{2\pi}\sqrt{a_3^2 + a_2^2}}$$

- EMG+GMG:

$$y = \frac{a_0}{4a_3}\exp\left(\frac{2a_1a_3 - 2a_3x + a_2^2}{a_3^2}\right)\mathrm{erfc}\left(\frac{a_1a_3 - a_3x + a_2^2}{\sqrt{2}a_2a_3}\right) + \frac{a_0}{2\sqrt{2\pi}\sqrt{a_2^2 + a_4^2}}\exp\left(-\frac{1}{2}\frac{(a_1 - x)^2}{a_2^2 + a_4^2}\right)\mathrm{erfc}\left(\frac{a_4(a_1 - x)}{\sqrt{2}a_2\sqrt{a_2^2 + a_4^2}}\right)$$

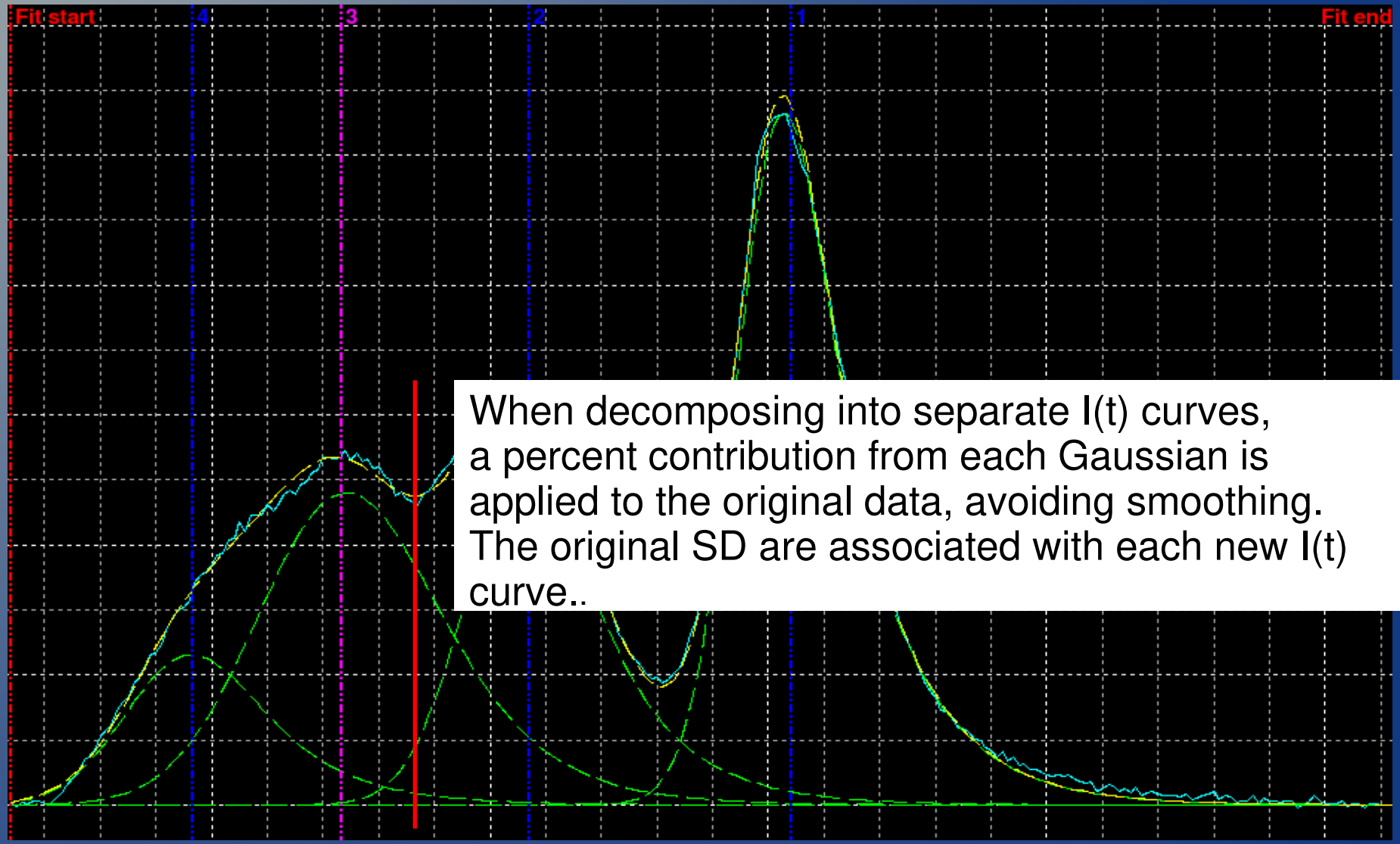
Aldolase EMG+GMG fit of one I(t)



Intensity (linear)

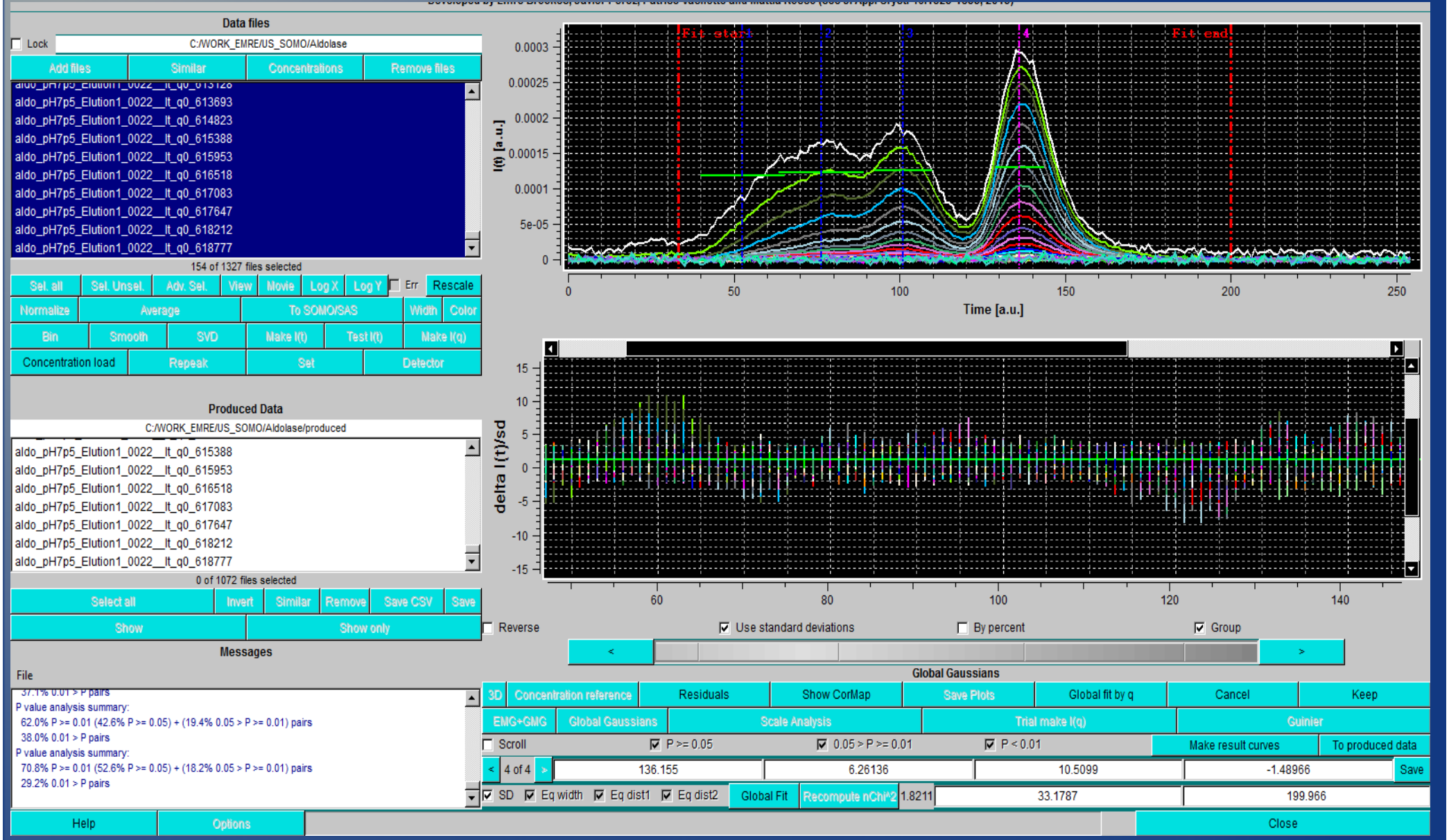
time or frame (linear)

Aldolase EMG+GMG fit of one I(t)

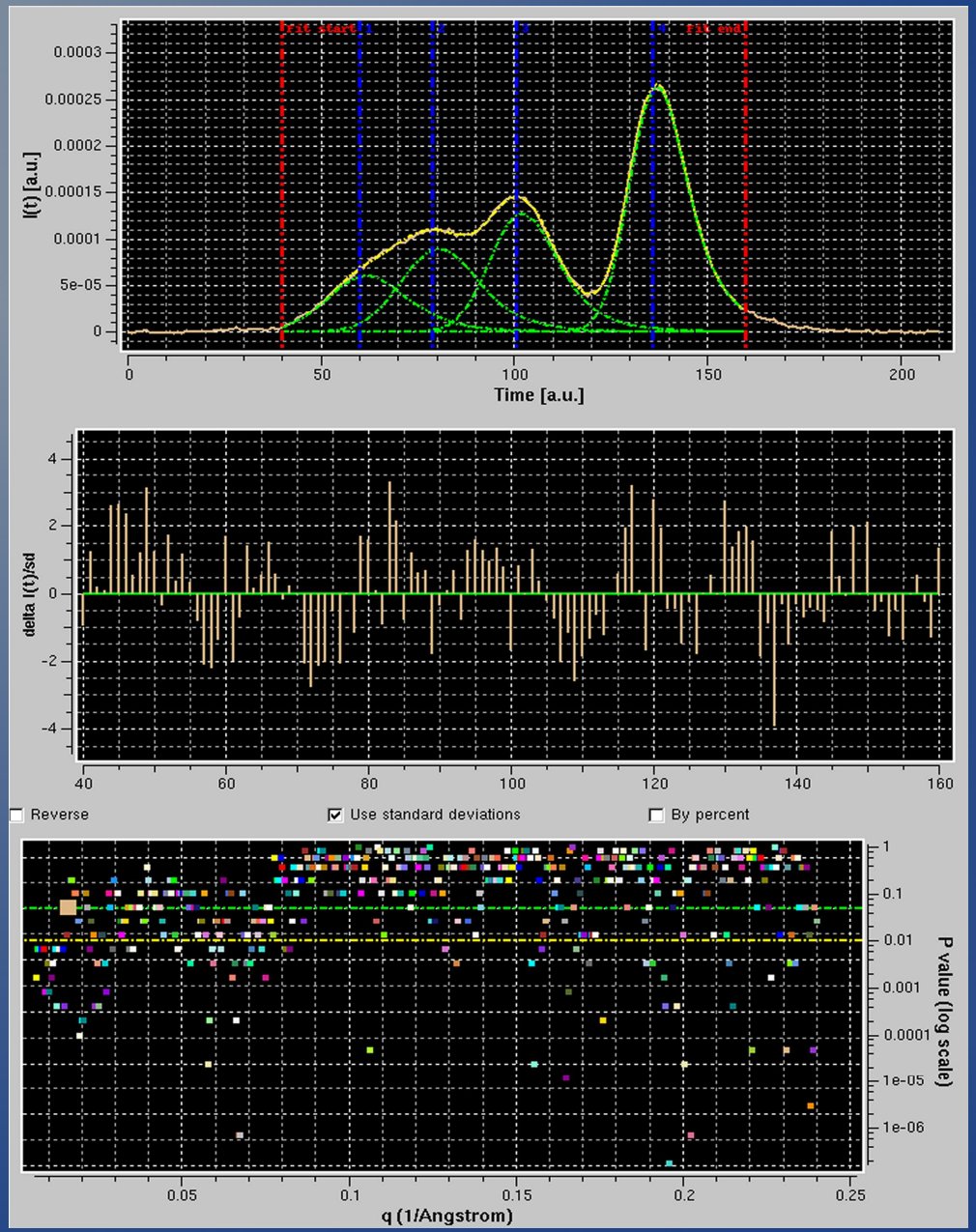


When decomposing into separate I(t) curves, a percent contribution from each Gaussian is applied to the original data, avoiding smoothing. The original SD are associated with each new I(t) curve..

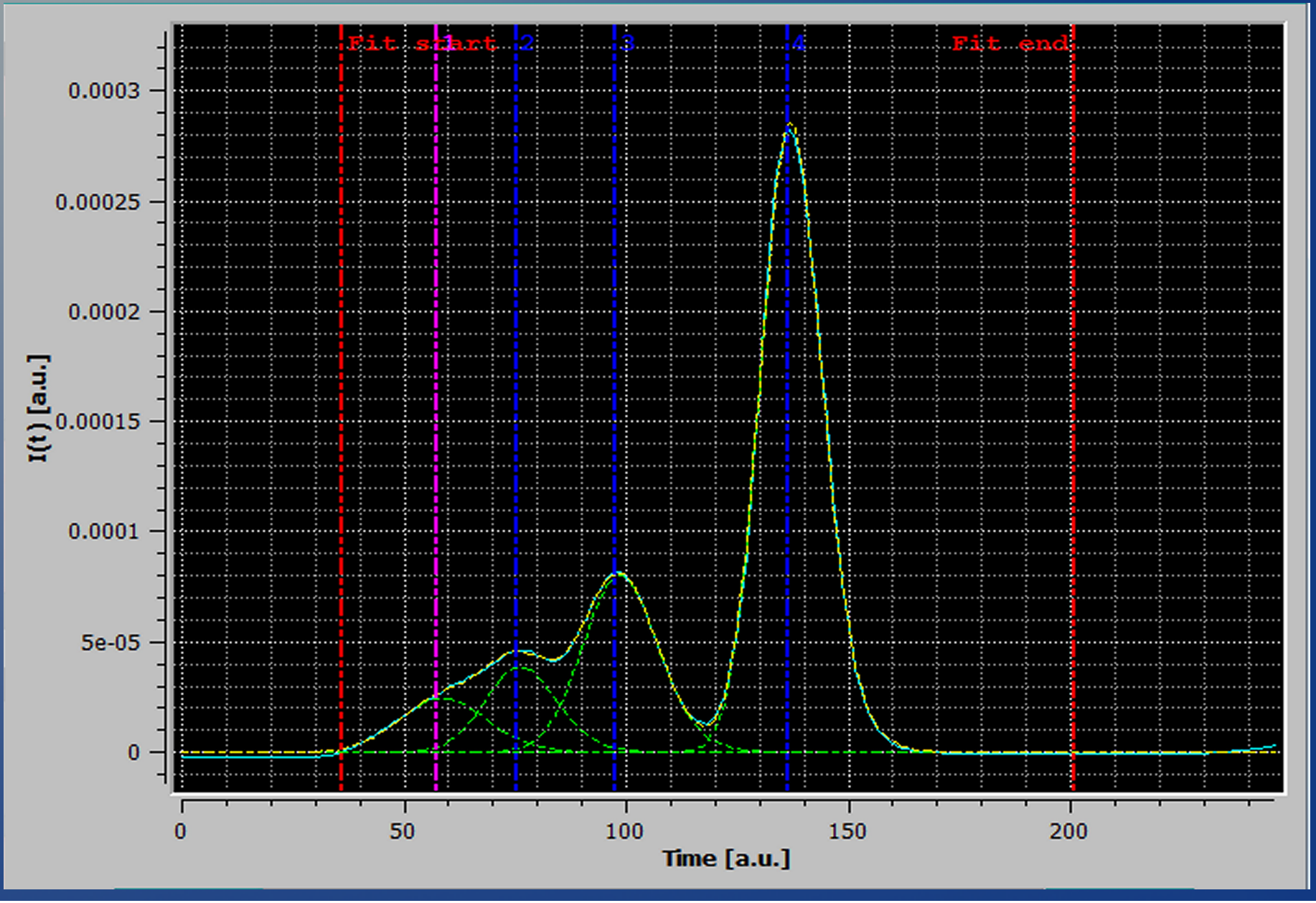
US-SOMO HPLC-SAXS: aldolase global EMG+GMG fitting



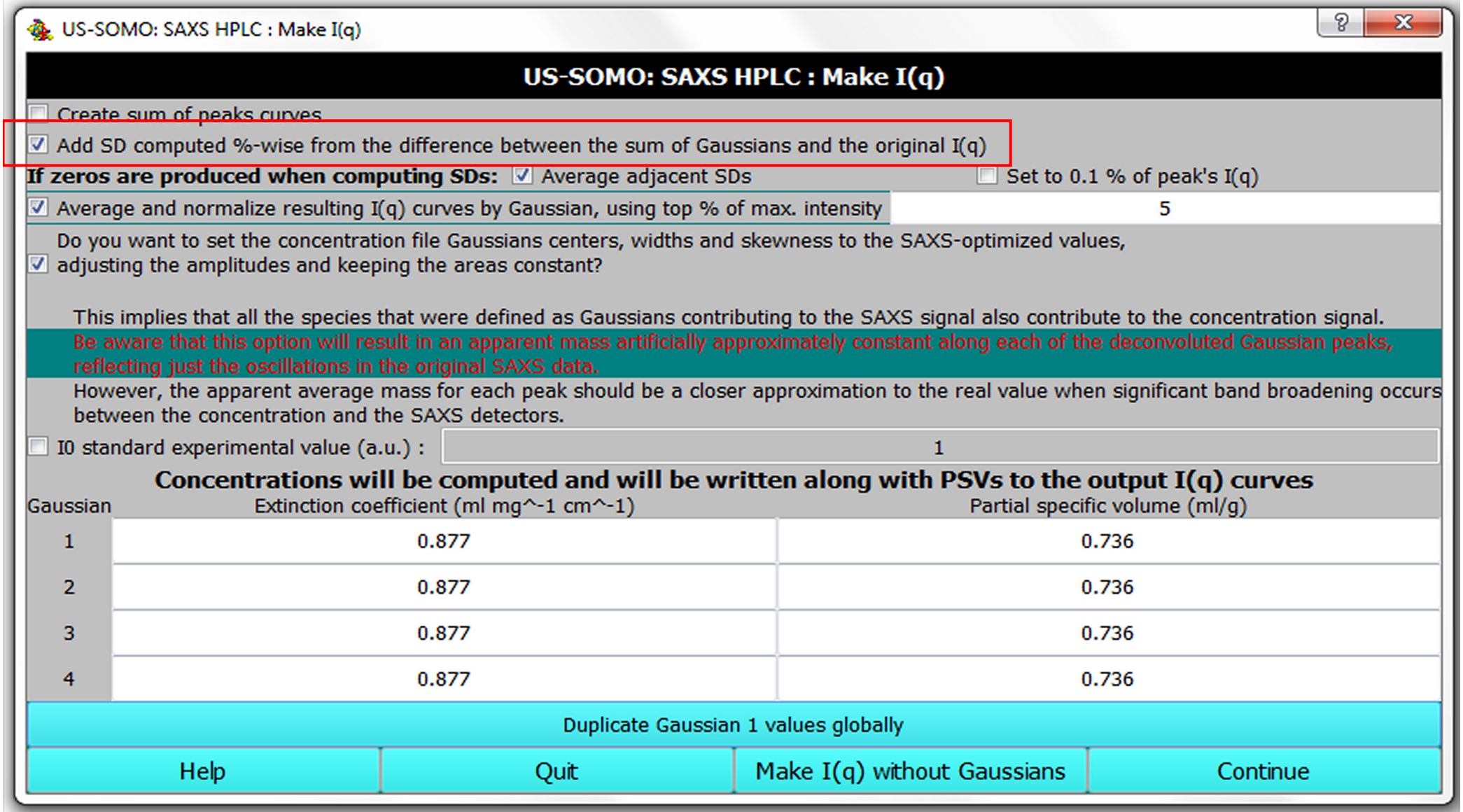
Global Gaussians in “scroll” mode, with residuals and P-values



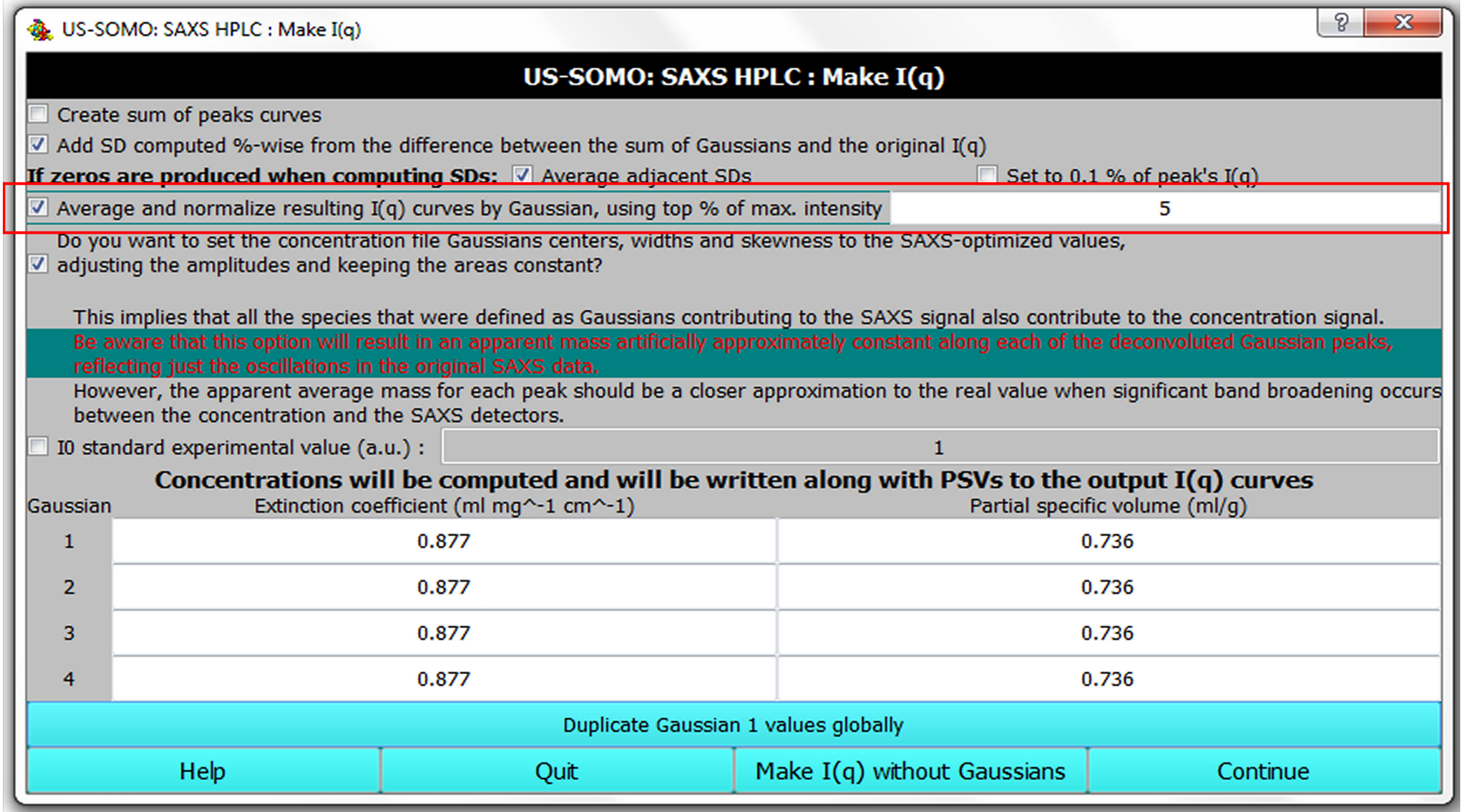
HPLC-SAXS module: UV Gaussians fit



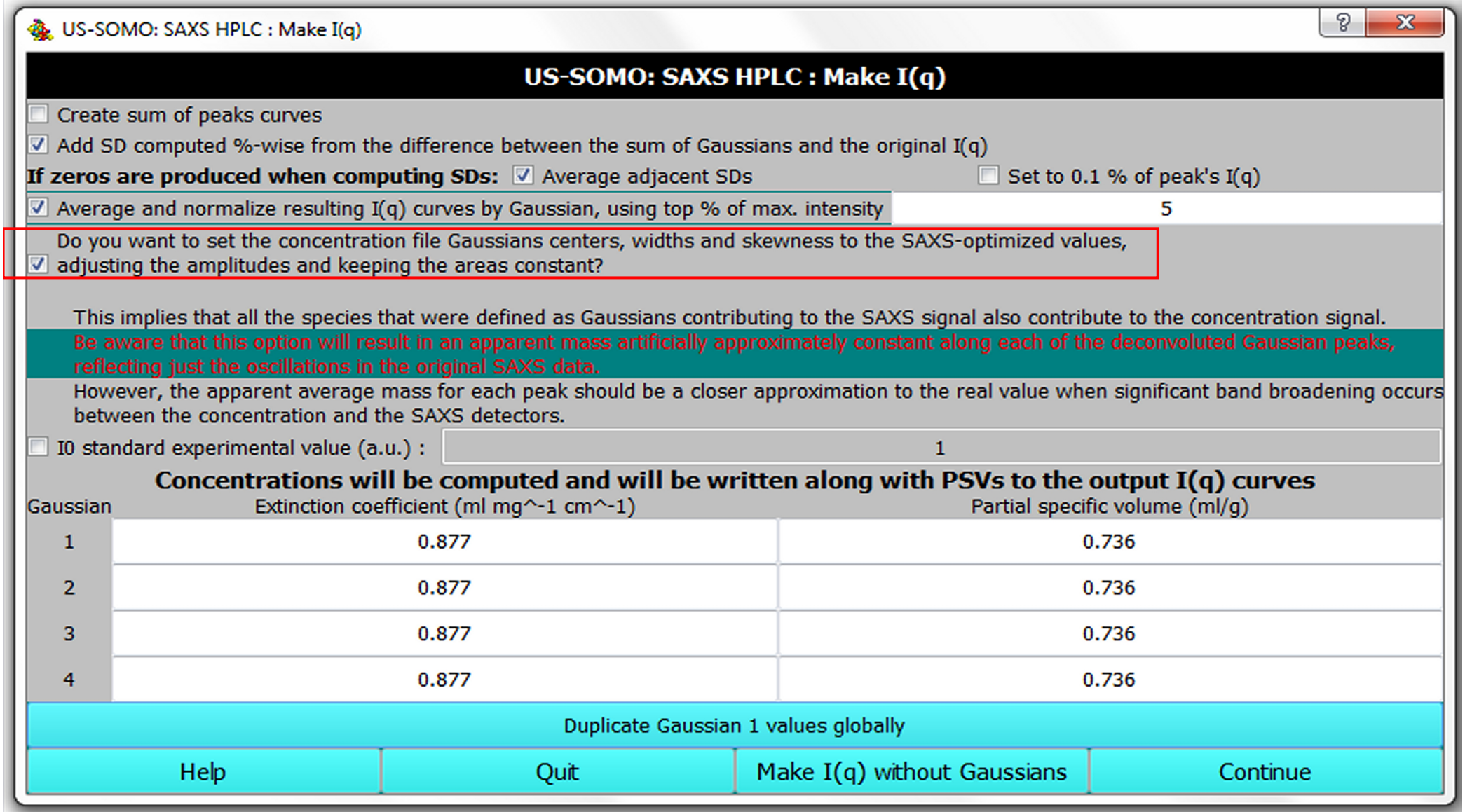
HPLC-SAXS module: Make I(q) panel



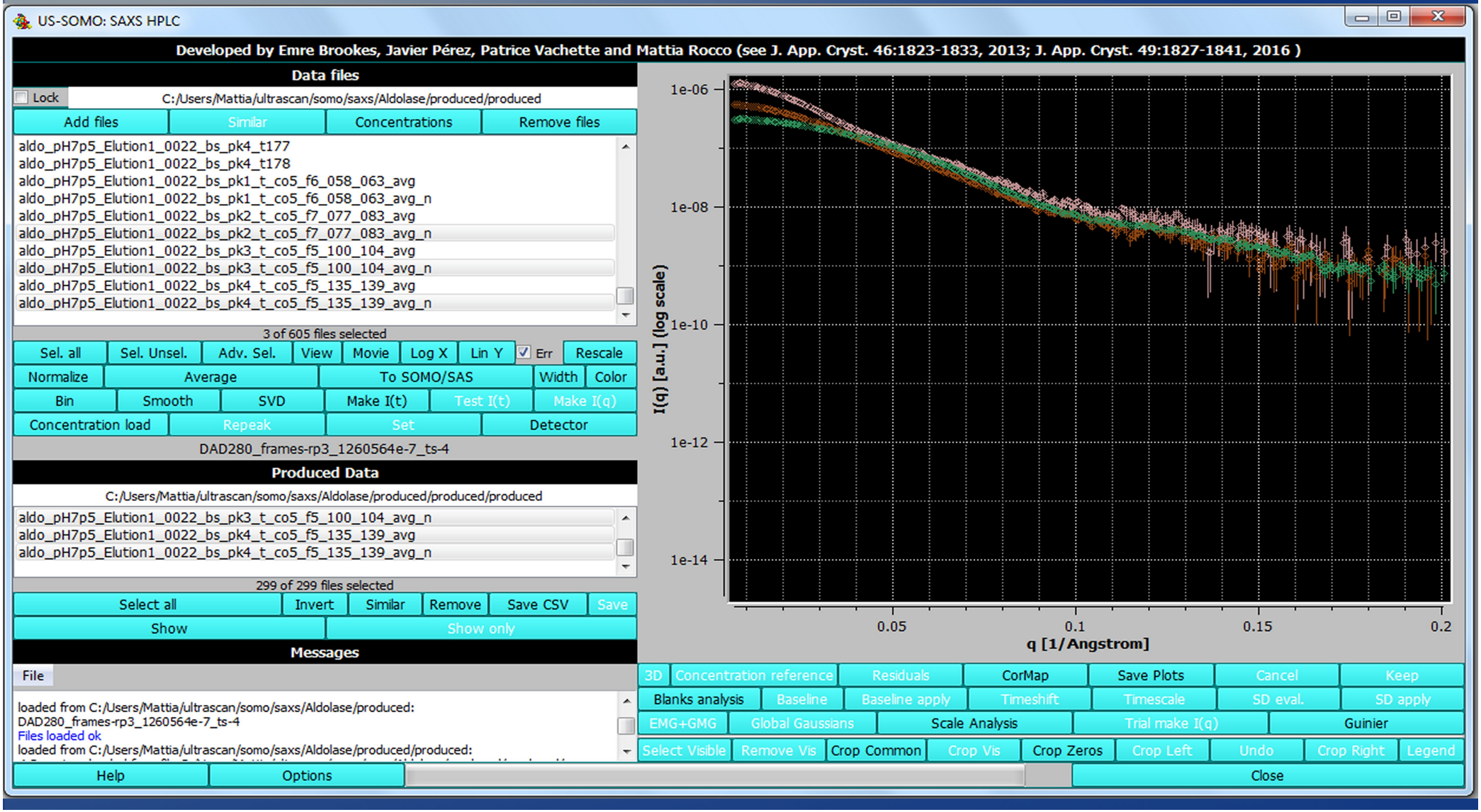
HPLC-SAXS module: Make I(q) panel



HPLC-SAXS module: Make I(q) panel

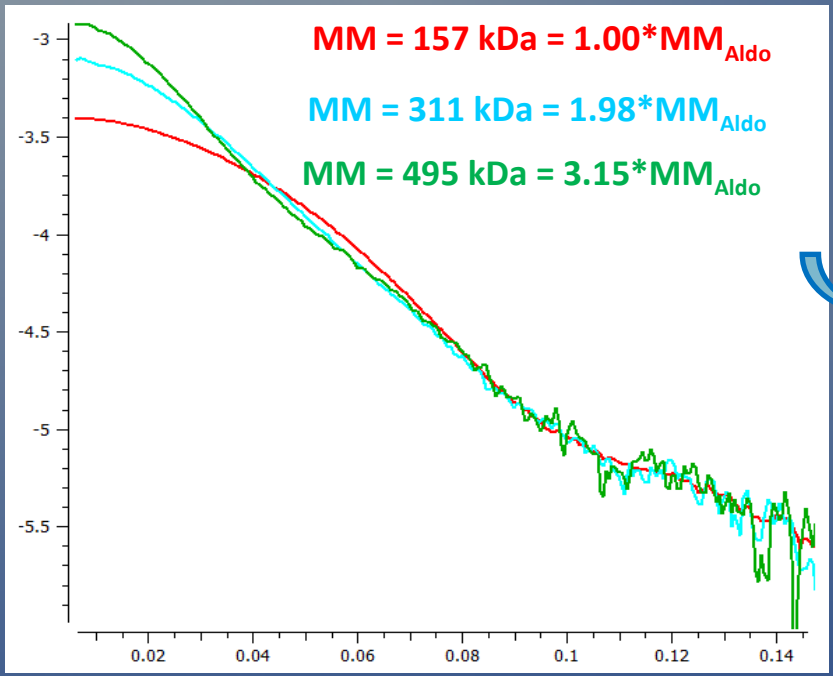


HPLC-SAXS module: auto-averaged final decomposed I(q) vs. q datasets

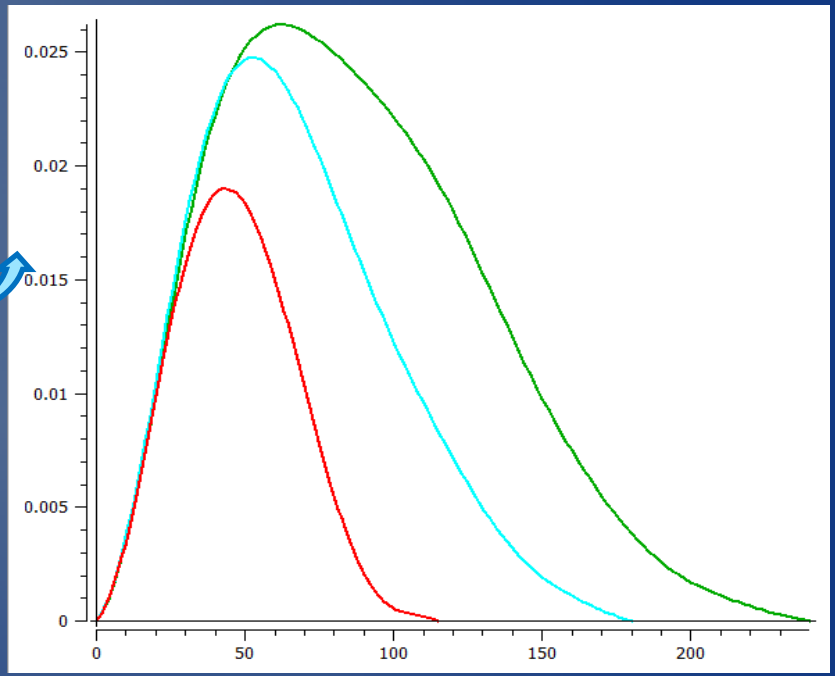


SEC-SAXS : Molar Mass (MM) of aldolase components pK_4, pK_3 & pK_2

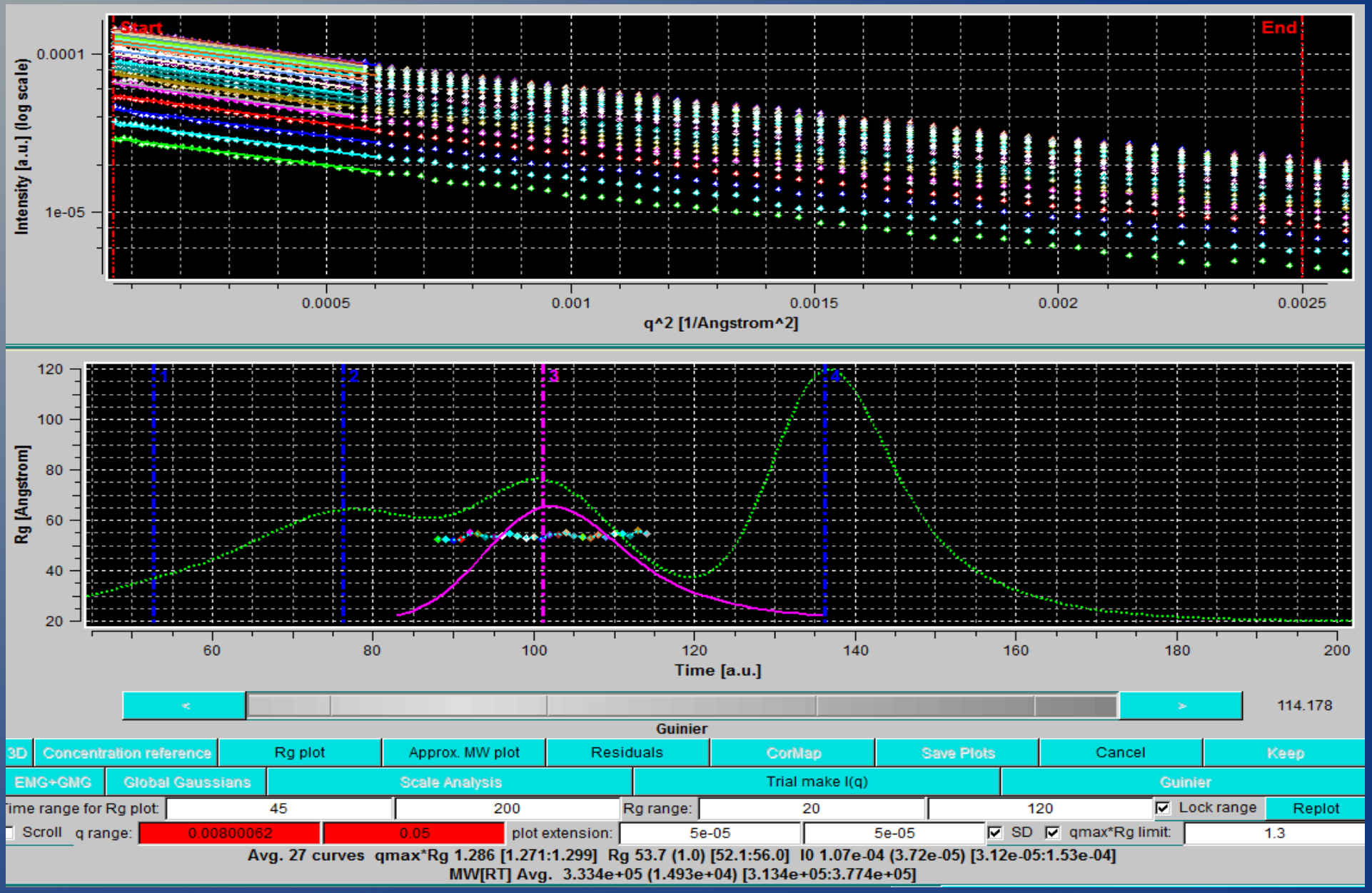
I(q)



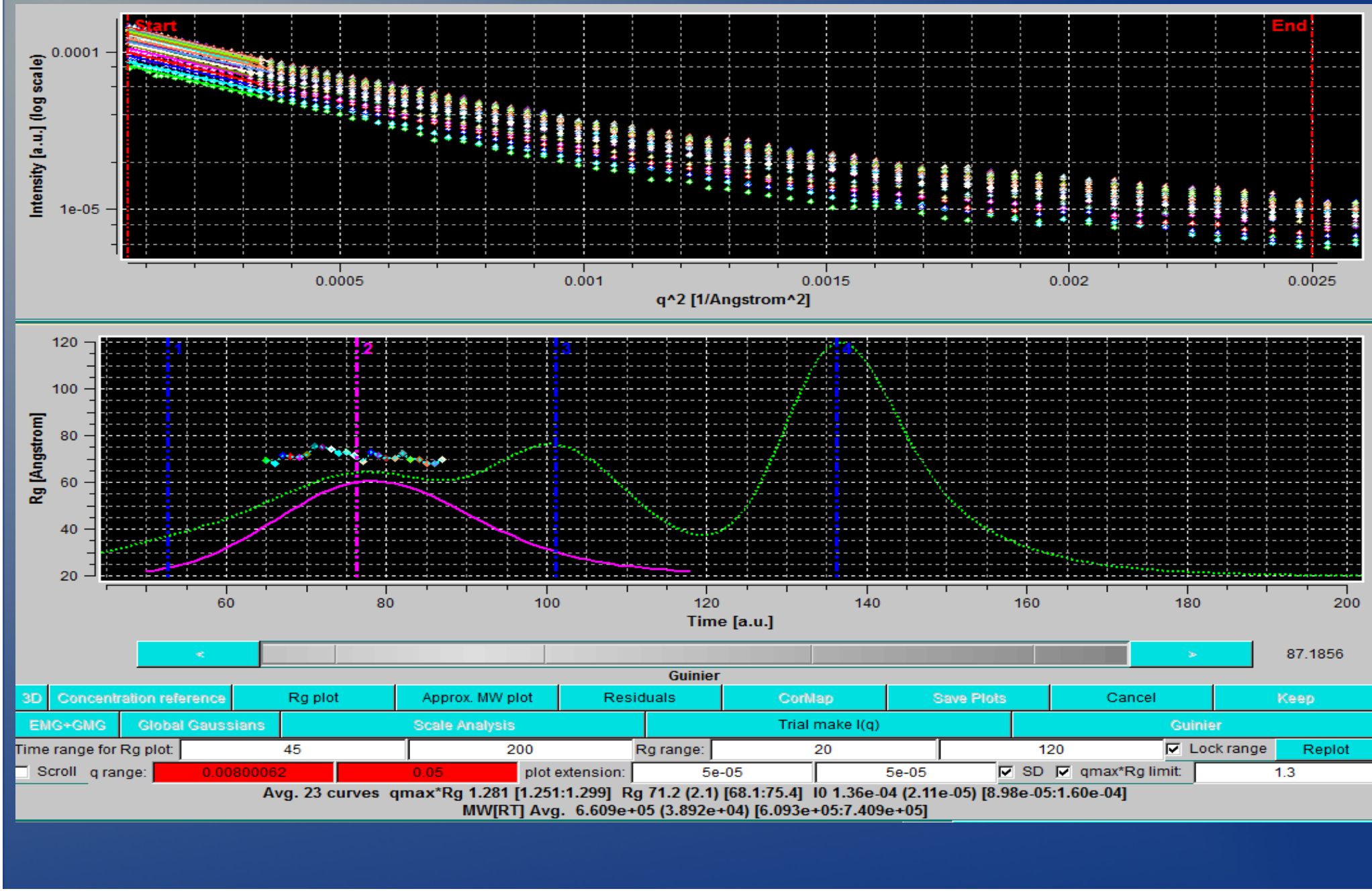
p(r)



US-SOMO HPLC-SAXS: aldolase Test I(q) 3rd peak



US-SOMO HPLC-SAXS: aldolase Test I(q) 2nd peak

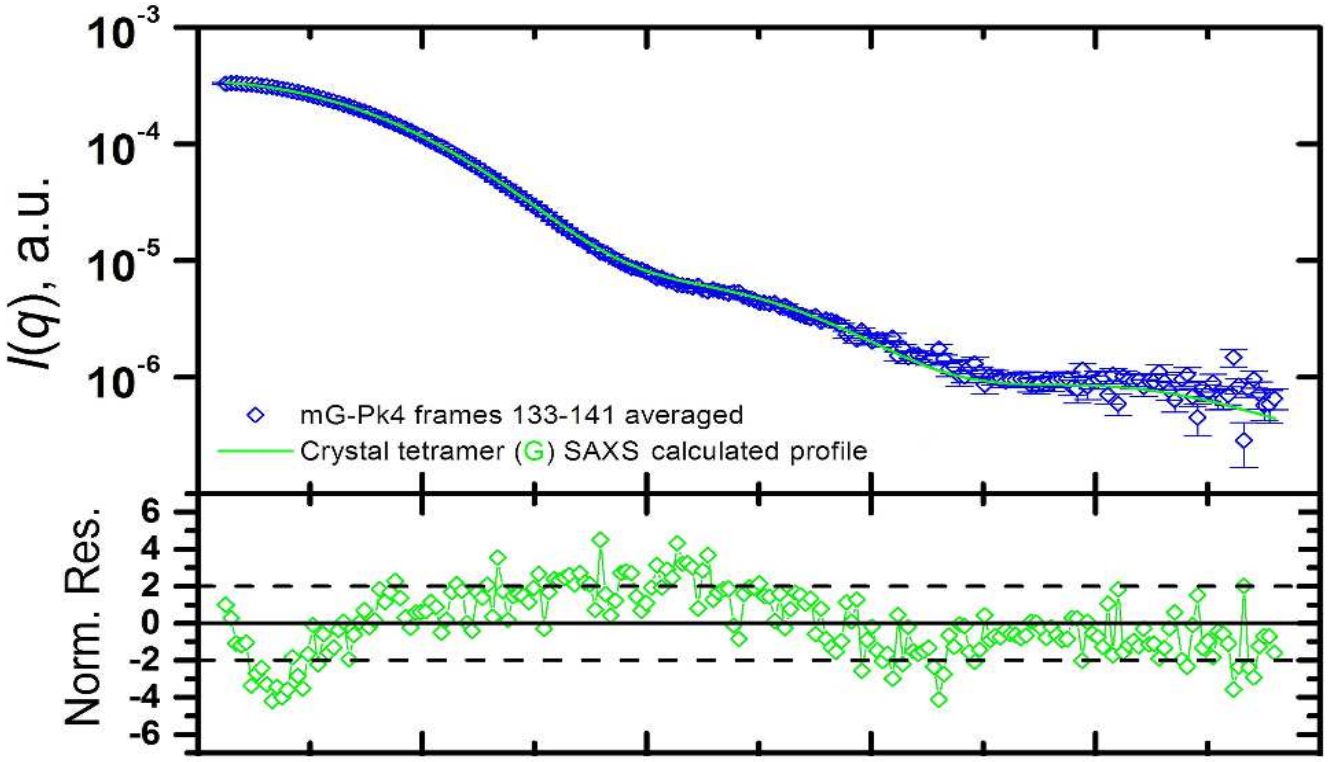


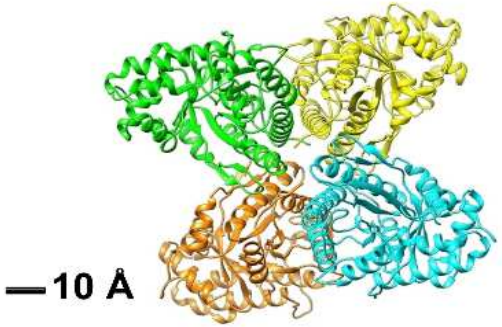
Modeling

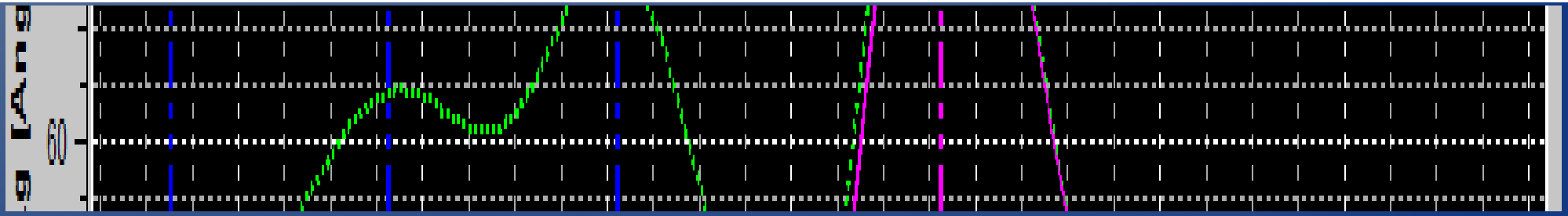
• “Balanced” supramolecular complexes of the aldolase tetramer were generated by the ClusPro server (https://cluspro.bu.edu/publications.php), filtered against the averaged SAXS curves derived from each decomposed peak

• The final SAXS curves were computed by the WAXSiS server (http://waxsis.uni-goettingen.de/) and compared with the averaged SAXS curves using the SAS NNLS module of US-SOMO

Deconvoluted aldolase HPLC-SAXS data

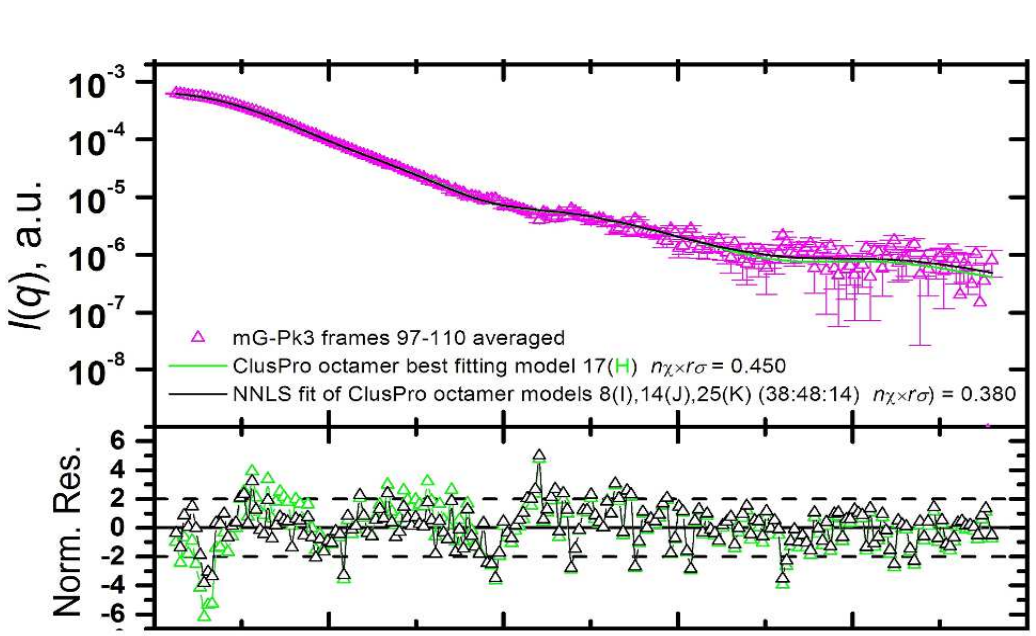


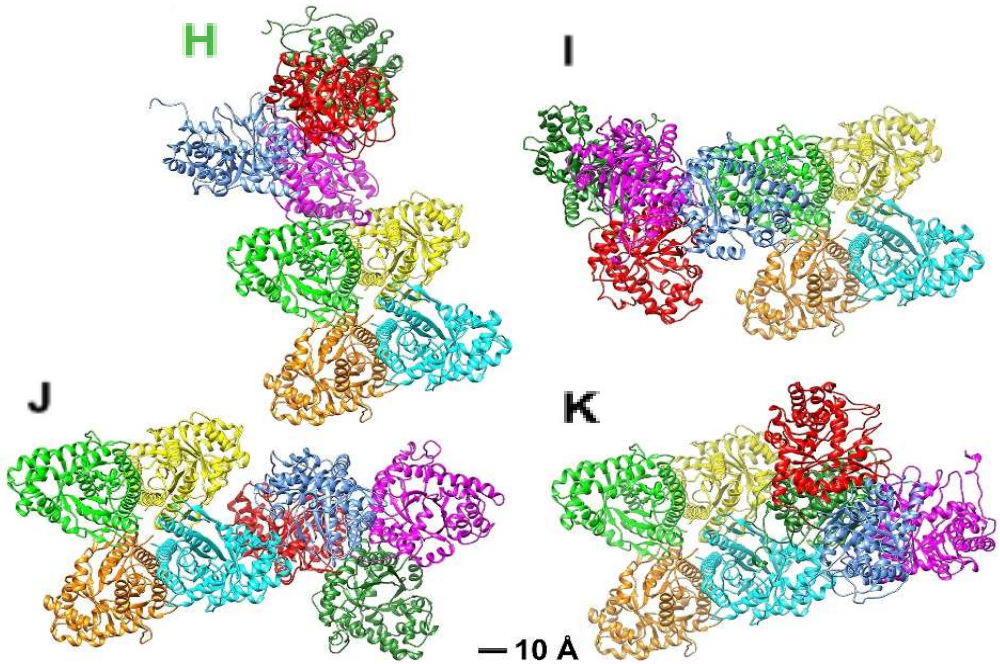


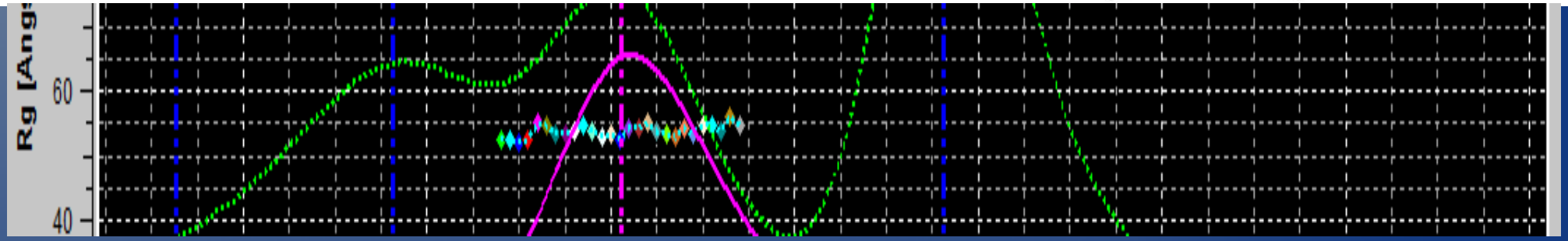


Brookes, E. et. al. [2016]
J. Appl. Cryst.

Deconvoluted aldolase HPLC-SAXS data

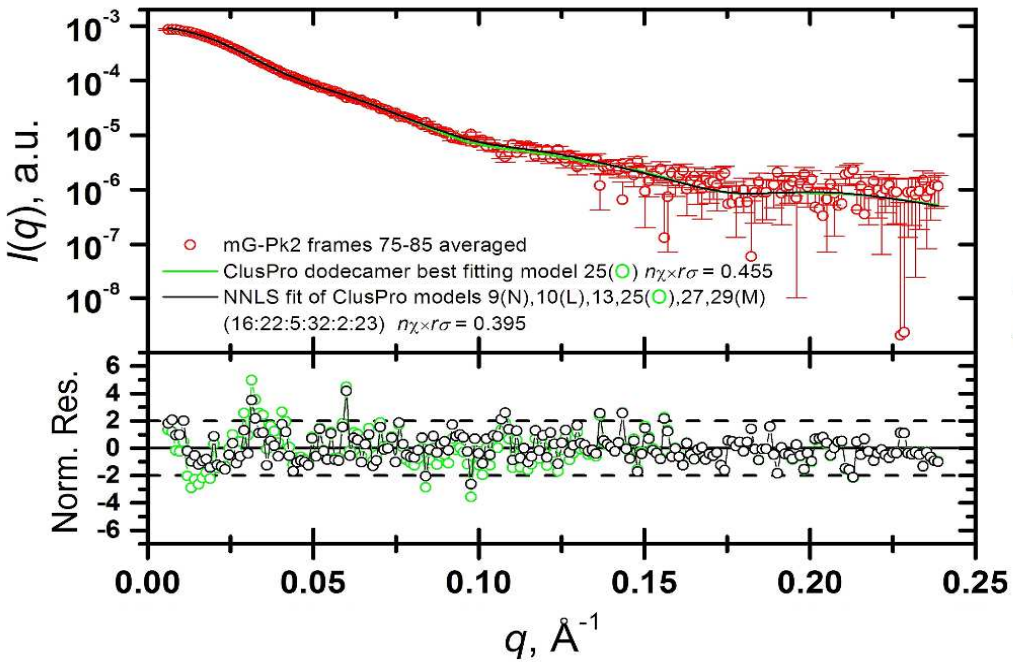


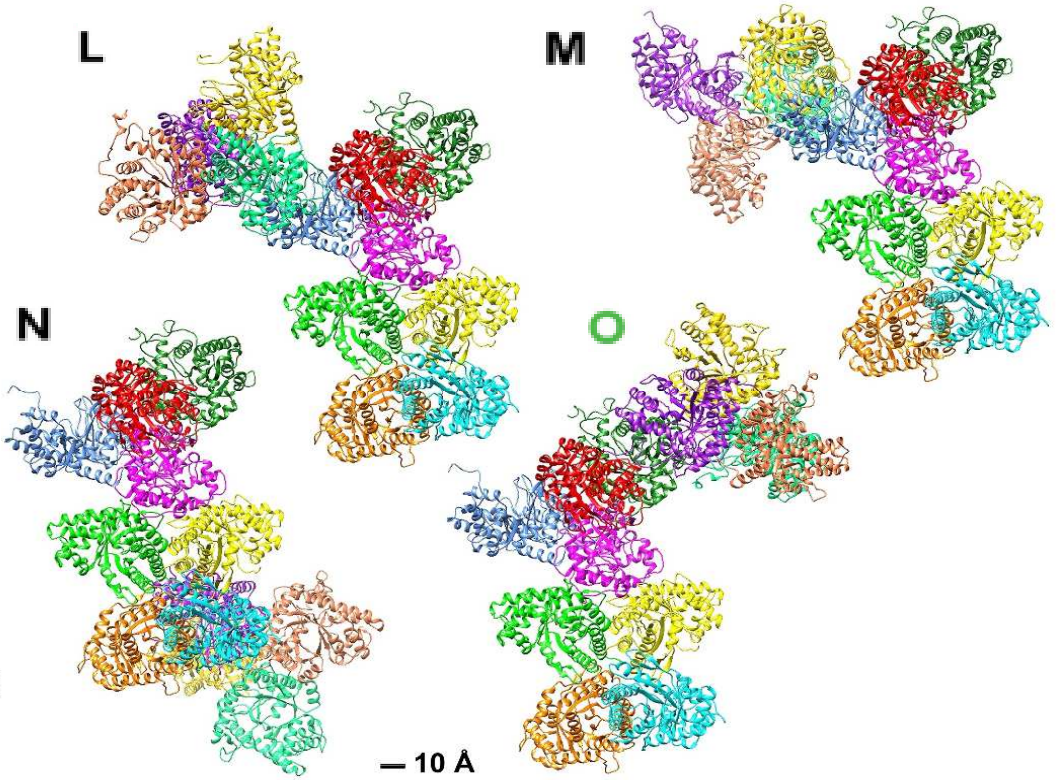


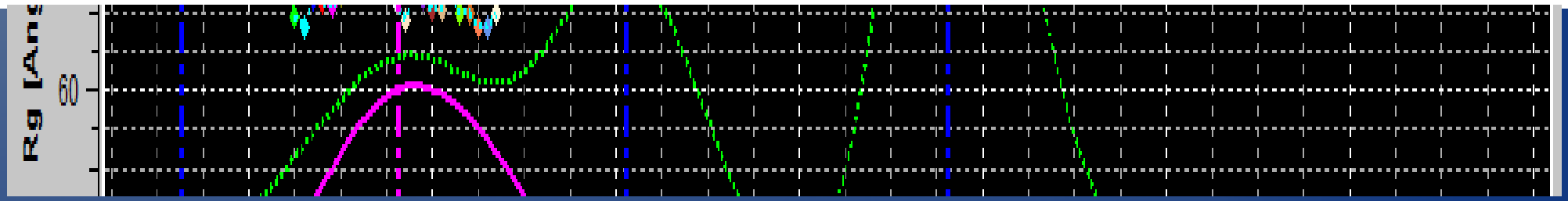


Brookes, E. et. al. [2016]
J. Appl. Cryst.

Deconvoluted aldolase HPLC-SAXS data





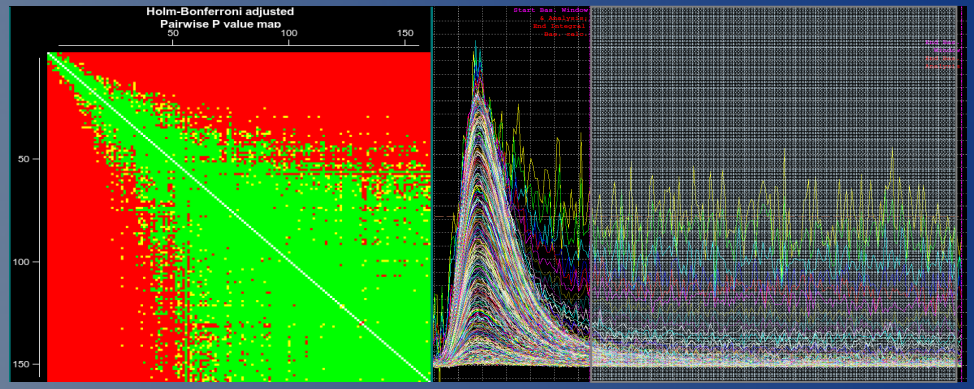


Brookes, E. et. al. [2016]
J. Appl. Cryst.

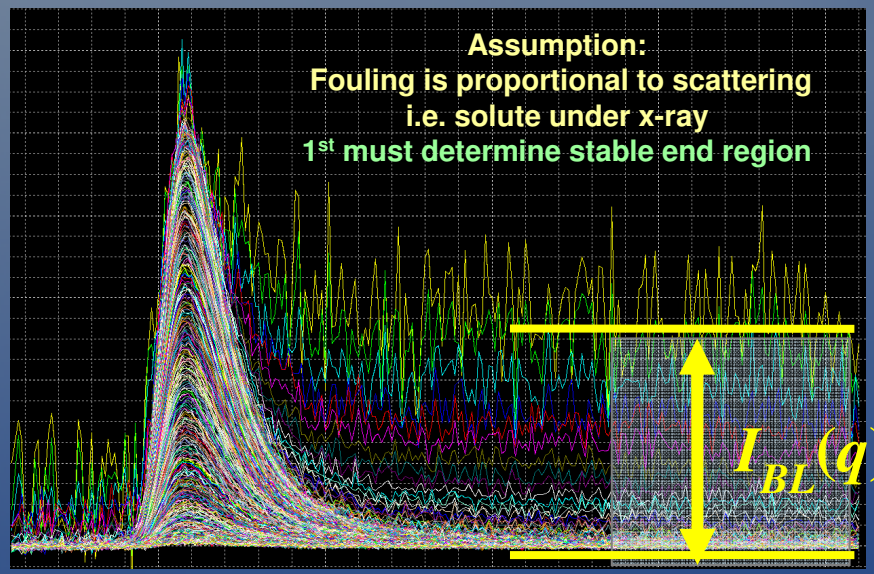
Other SEC-SAXS tools

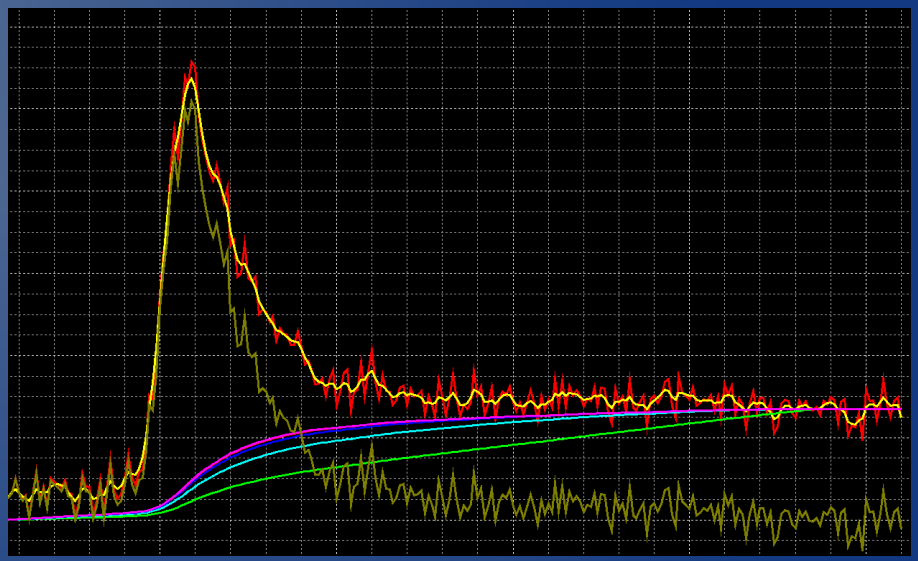
Brookes, E., Vachette, P., Rocco, M. & Perez, J. [2016] J. Appl. Cryst. 49(5).

Pairwise P value map for visual representation of similarity tests of pairs of frames



Integral baseline correction

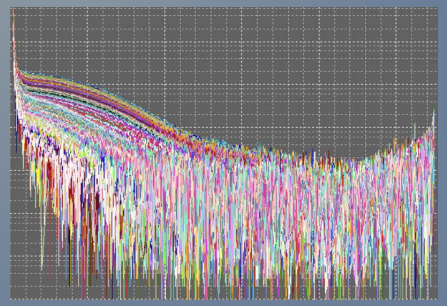




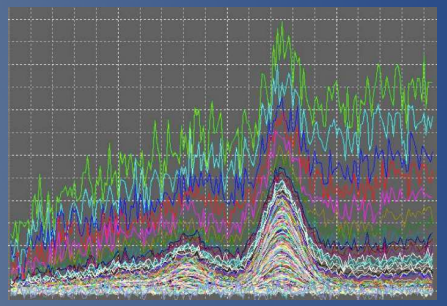
US-SOMO: HPLC/SEC-SAXS

Brookes E., Perez J, Cardinali B, Profumo A., Vachette P & Rocco M. (2013)
J. Appl. Cryst. 46, 1823-1833

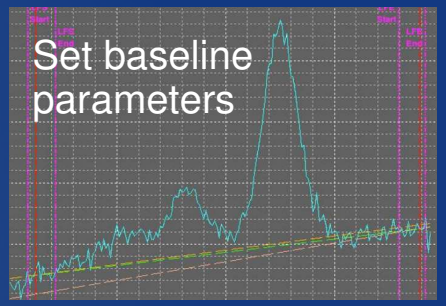
Collect HPLC-SAXS data →



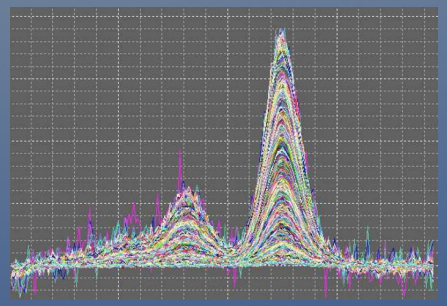
Make I(t) →



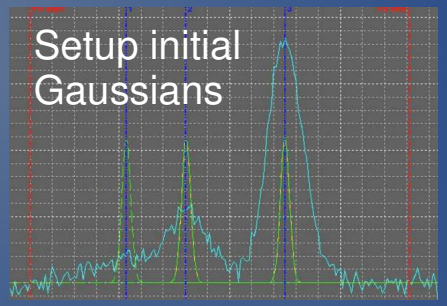
Select typical curve →



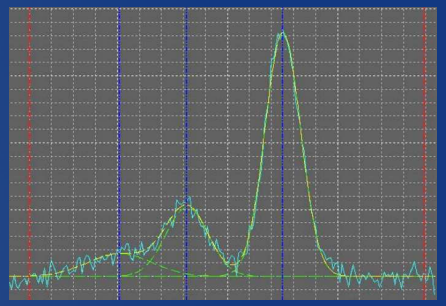
Apply b.l. params to all curves →



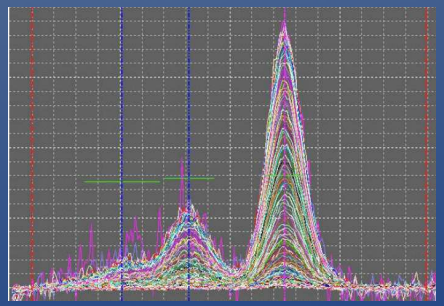
Select typical curve →



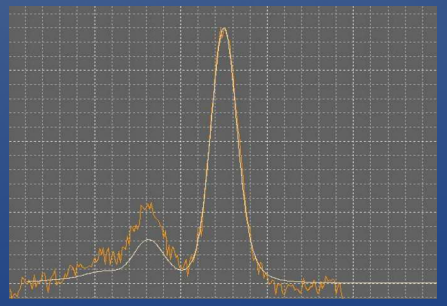
Gauss Fit →



Global Gauss Fit →



Opt. apply to conc. curve →



Make I(q) →

Set of I(q) curves for each Gaussian peak
Automatic frame averaging and normalization

Thanks for listening

email us at
emre@biochem.uthscsa.edu mattia.rocco@hsanmartino.it
or visit http://somo.aucsolutions.com/

Acknowledgments

- NIH K25GM090154 / PI Brookes
- NSF CHE-1265817 / PI Brookes
- NSF XSEDE MCB140255 / PI Brookes
- NIH R01 GM120600 / PI Demeler

SCHOOL OF MEDICINE
UT HEALTH SCIENCE CENTER®
DEPARTMENT OF BIOCHEMISTRY

Funded by a grant from the National Institute of General Medical Sciences of the National Institutes of Health

NSF

EPSRC
Engineering and Physical Sciences Research Council