

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

M. Rocco<sup>a</sup>

J. Pérez<sup>b</sup>

P. Vachette<sup>c</sup>

E. H. Brookes<sup>d</sup>

<sup>a</sup>*Biopolimeri e Proteomica,*

*IRCCS Ospedale Policlinico San Martino, Genova, Italy*

<sup>b</sup>*Beamline SWING, Synchrotron SOLEIL, Gif-sur-Yvette, France*

<sup>c</sup>*I2BC, UMR 9198, Orsay, France*

<sup>d</sup>*Dept. of Biochemistry, U. of Texas Health Science Center at San  
Antonio, San Antonio, Texas*

<http://somo.aucnsolutions.com/>

# US-SOMO main window

SOMO Solution Modeler

Lookup Tables SOMO MD PDB Configuration File

**PDB Functions:**

Select Lookup Table C:\Program Files\UltraScan3\etc\somo.residue

Batch Mode/Cluster Operation

Load Single PDB File not selected

**Please select a PDB Structure:**

View/Edit PDB File PDB Editor

SAXS/SANS Functions

Run DMD

BD

**Bead Model Functions:**

**Bead Model Suffix:** A20R50hiOT / A10R30syOThyG5 / A20R50

Overwrite existing filenames  Add auto-generated suffix

Build SoMo Bead Model Build AtoB (Grid) Bead Model

Build SoMo Overlap Bead Model Build vdW Overlap Bead Model

View ASA Results Grid Existing Bead Model Visualize Bead Model

Batch Mode/Cluster Operation View Bead Model File

Load Single Bead Model File not selected

SAXS/SANS Functions  Automatically Calculate Hydrodynamics

**Hydrodynamic Calculations:**

Calculate RB Hydrodynamics SMI Calculate RB Hydrodynamics ZENO

Show Hydrodynamic Calculations Open Hydrodynamic Calculations File

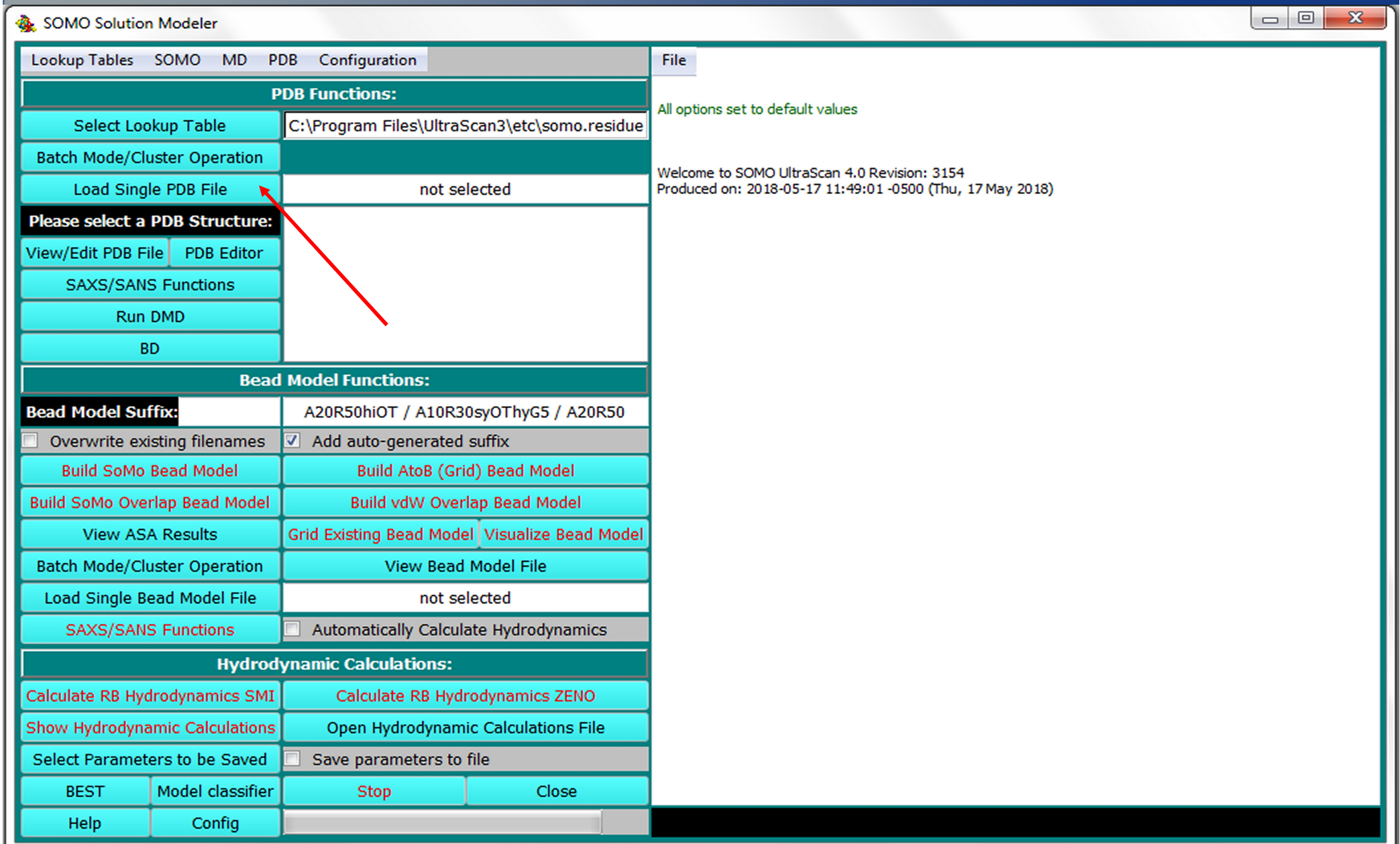
Select Parameters to be Saved  Save parameters to file

BEST Model classifier Stop Close

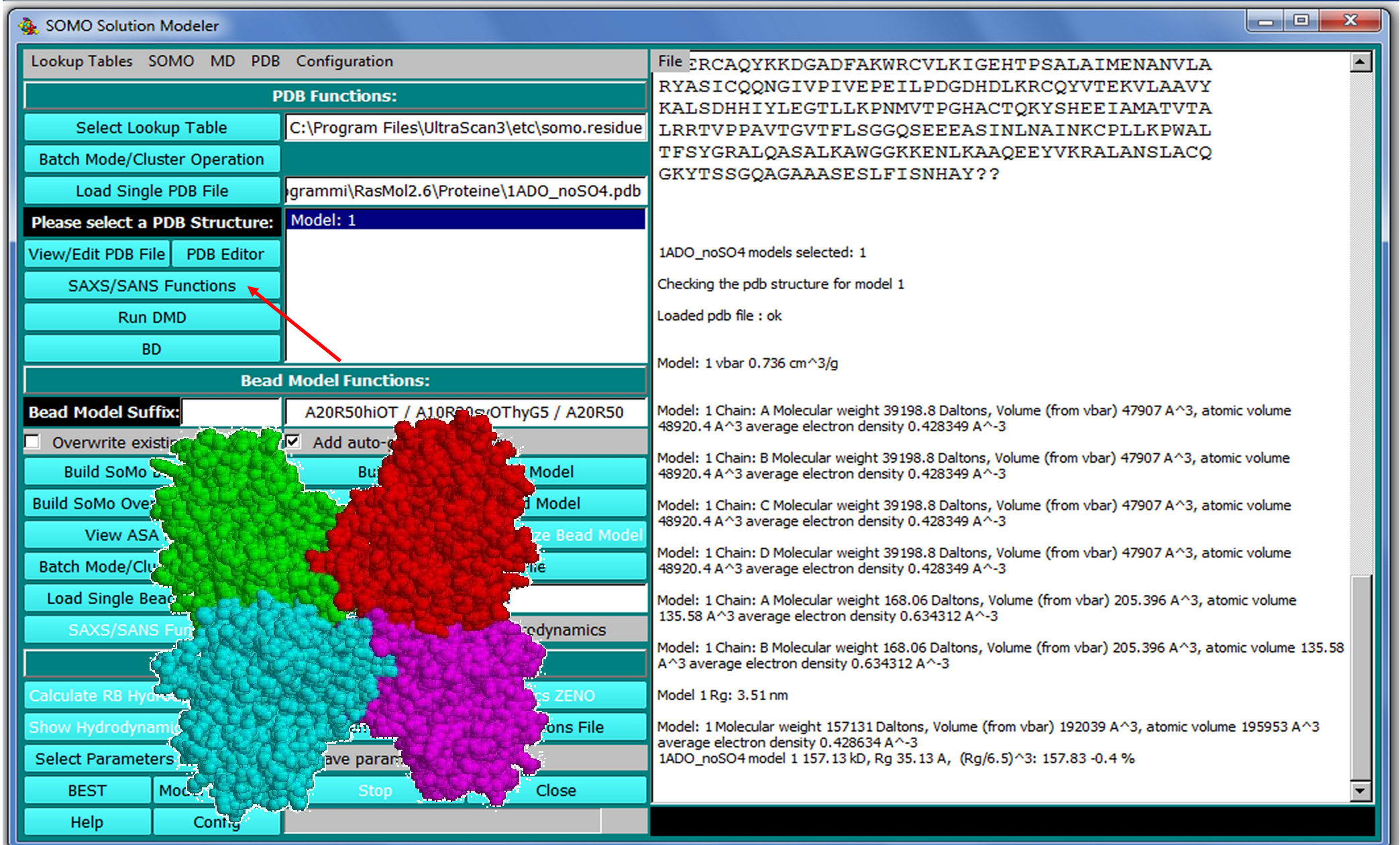
Help Config

All options set to default values

Welcome to SOMO UltraScan 4.0 Revision: 3154  
Produced on: 2018-05-17 11:49:01 -0500 (Thu, 17 May 2018)



# US-SOMO main window



The screenshot displays the US-SOMO main window, which is divided into several functional areas:

- Menu Bar:** Lookup Tables, SOMO, MD, PDB, Configuration.
- PDB Functions:** Select Lookup Table (C:\Program Files\UltraScan3\etc\somo.residue), Batch Mode/Cluster Operation, Load Single PDB File (grammi\RasMol2.6\Proteine\1ADO\_noSO4.pdb), Please select a PDB Structure: Model: 1, View/Edit PDB File, PDB Editor, SAXS/SANS Functions (highlighted with a red arrow), Run DMD, BD.
- Bead Model Functions:** Bead Model Suffix: A20R50hiOT / A10R50eOThyG5 / A20R50,  Overwrite existing,  Add auto-..., Build SoMo..., Build SoMo Over..., View ASA..., Batch Mode/Cluster..., Load Single Bead..., SAXS/SANS Functions..., Calculate RB Hydro..., Show Hydrodynamic..., Select Parameters..., BEST, Mod..., Stop, Close, Help, Config.
- 3D Model:** A central 3D molecular model is shown, composed of several subunits colored in green, red, cyan, and magenta.
- Log Window:** File: ERCAQYKKDGADFAKWRCVLKIGEHTPSALAIMENANVLA RYASICQQNGIVPIVEPEILPDGDHDLKRCQYVTEKVLAAVY KALSDHHIYLEGTLKPNMVT PGHACTQKYSHEEIAMATVTA LRRTVPPAVTGVTFSLGGQSEEEASINLNAINKCPLLPWAL TFSYGRALQASALKAWGGKKENLKAAQEEYVKRALANSLACQ GKYTSSGQAGAAASESLFISNHAY??  
1ADO\_noSO4 models selected: 1  
Checking the pdb structure for model 1  
Loaded pdb file : ok  
Model: 1 vbar 0.736 cm<sup>3</sup>/g  
Model: 1 Chain: A Molecular weight 39198.8 Daltons, Volume (from vbar) 47907 A<sup>3</sup>, atomic volume 48920.4 A<sup>3</sup> average electron density 0.428349 A<sup>-3</sup>  
Model: 1 Chain: B Molecular weight 39198.8 Daltons, Volume (from vbar) 47907 A<sup>3</sup>, atomic volume 48920.4 A<sup>3</sup> average electron density 0.428349 A<sup>-3</sup>  
Model: 1 Chain: C Molecular weight 39198.8 Daltons, Volume (from vbar) 47907 A<sup>3</sup>, atomic volume 48920.4 A<sup>3</sup> average electron density 0.428349 A<sup>-3</sup>  
Model: 1 Chain: D Molecular weight 39198.8 Daltons, Volume (from vbar) 47907 A<sup>3</sup>, atomic volume 48920.4 A<sup>3</sup> average electron density 0.428349 A<sup>-3</sup>  
Model: 1 Chain: A Molecular weight 168.06 Daltons, Volume (from vbar) 205.396 A<sup>3</sup>, atomic volume 135.58 A<sup>3</sup> average electron density 0.634312 A<sup>-3</sup>  
Model: 1 Chain: B Molecular weight 168.06 Daltons, Volume (from vbar) 205.396 A<sup>3</sup>, atomic volume 135.58 A<sup>3</sup> average electron density 0.634312 A<sup>-3</sup>  
Model 1 Rg: 3.51 nm  
Model: 1 Molecular weight 157131 Daltons, Volume (from vbar) 192039 A<sup>3</sup>, atomic volume 195953 A<sup>3</sup> average electron density 0.428634 A<sup>-3</sup>  
1ADO\_noSO4 model 1 157.13 kD, Rg 35.13 A, (Rg/6.5)<sup>3</sup>: 157.83 -0.4 %

# US-SOMO SAS MODULE

US-SOMO: SAS Functions

**PDB Filename:** 1ADO\_noS04

**Definition files:**

- Load Atom Definition File:
- Load Hybridization File:
- Load SAXS Coefficients File:

**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

Standard  Kratky plot

Create standard output files

SAXS  F-DB  SH-DB  Q-DB  Crysol

SANS  F-DB  SH-DB  Q-DB  Cryson

**File suffix:**

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):**

**Smoothing:**

Raw  SAXS  SANS  Normalize

Residue contrib. range (Angstrom):

Compute P(r) Distribution

File

### SAXS Curve

**P(r) Distribution Curve**

A red arrow points from the 'Compute P(r) Distribution' button in the control panel to the top-left corner of the 'P(r) Distribution Curve' plot area.

# US-SOMO SAS MODULE

US-SOMO: SAS Functions

**PDB Filename:** 1ADO\_noSO4

**Definition files:**

- Load Atom Definition File:
- Load Hybridization File:
- Load SAXS Coefficients File:

**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

Standard  Kratky plot

Create standard output files

SAXS  F-DB  SH-DB  Q-DB  Crysol

SANS  F-DB  SH-DB  Q-DB  Cryson

**File suffix:**

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):**

**Smoothing:**

Raw  SAXS  SANS  Normalize

Residue contrib. range (Angstrom):

Compute P(r) Distribution

**File**

Number of atoms 11048. Bin size 1.

P(r) curve file: C:\Users\mattia\ultrascan\somo\saxs\1ADO\_noSO4\_1b1.spr\_x created.

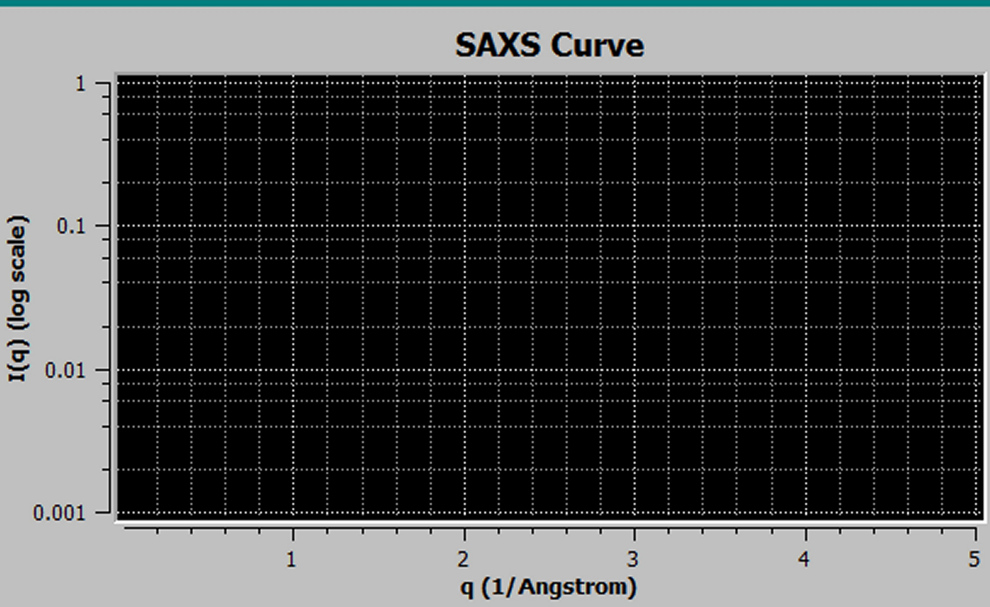
1ADO\_noSO4 Molecular weight 157131 (computed from pdb)

**P(r): Bin size: 1 "1ADO\_noSO4"**

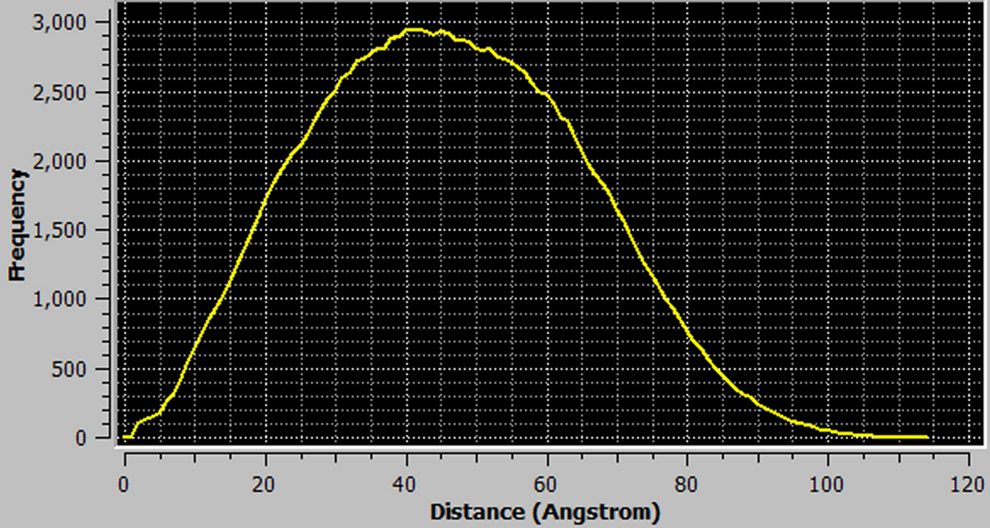
Stop | Open Options Panel

Help | Close

**SAXS Curve**



**P(r) Distribution Curve**



# US-SOMO SAS MODULE

US-SOMO: SAS Functions

**PDB Filename:** 1ADO\_noSO4

**Definition files:**

- Load Atom Definition File:
- Load Hybridization File:
- Load SAXS Coefficients File:

**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

Standard  Kratky plot

Create standard output files

SAXS  F-DB  SH-DB  Q-DB  Crysol

SANS  F-DB  SH-DB  Q-DB  Cryson

**File suffix:**

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):**

**Smoothing:**

Raw  SAXS  SANS  Normalize

Residue contrib. range (Angstrom):

Compute P(r) Distribution

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

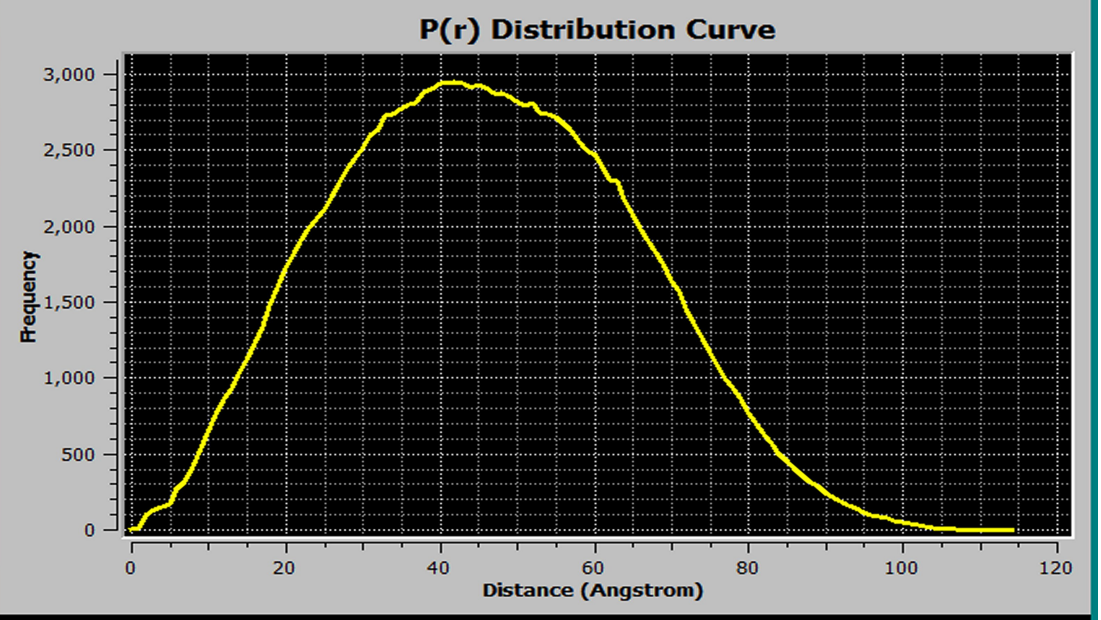
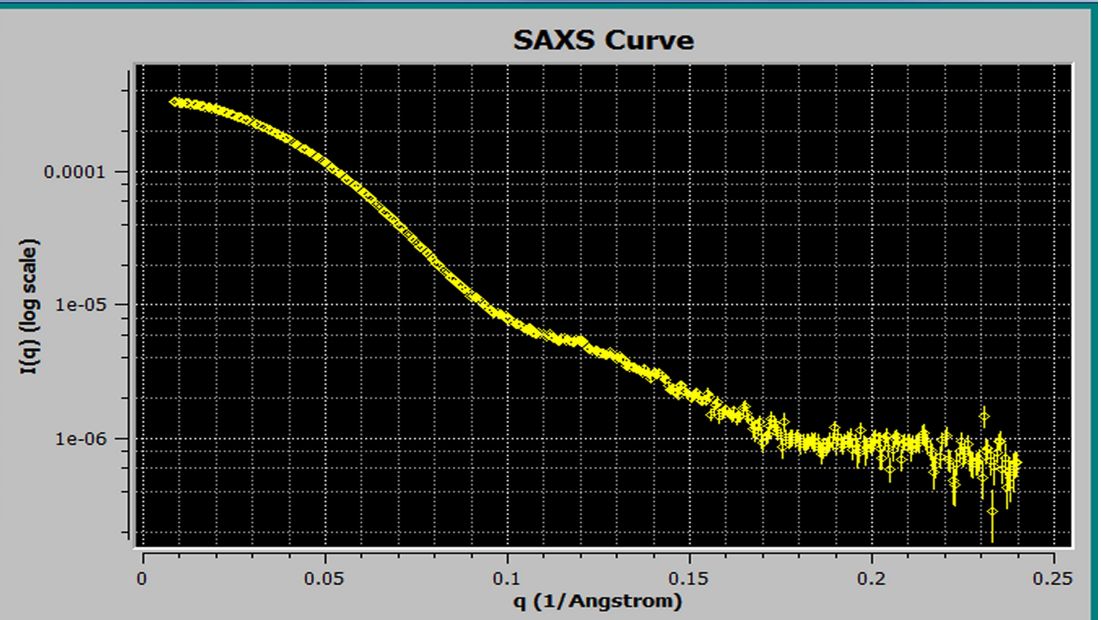
I(q) plot done

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.spr\_x created.

1ADO\_noSO4 Molecular weight 157131 (computed from pdb)



# US-SOMO SAS MODULE

US-SOMO: SAS Functions

**PDB Filename:** 1ADO\_noSO4

**Definition files:**

- Load Atom Definition File:
- Load Hybridization File:
- Load SAXS Coefficients File:

**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

Standard  Kratky plot

Create standard output files

SAXS  F-DB  SH-DB  Q-DB  Crysol

**SAXS Curve**

**P(r) Distribution Curve**

**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.spr\_x created.

1ADO\_noSO4 Molecular weight 157131 (computed from pdb)

Stop | Open Options Panel

Help | Close

IFT: Indirect Fourier Transform using Bayesian Analysis to generate  $P(r)$  vs.  $r$   
By Steen Hansen (see J. Appl. Cryst. (2014) 47, 1469-1471, and refs. therein)

# US-SOMO SAS MODULE

**US-SOMO: SAS Functions**

**PDB Filename:** 1ADO\_noSO4

**Definition files:**

Load Atom Definition File:

Load Hybridization File:

Load SAXS Coefficients File:

**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

Standard  Kratky plot

Create standard output files

SAXS  F-DB  SH-DB  Q-DB  Crystol

**SAXS Curve**

$I(q)$  (log scale)

q (1/Angstrom)

**US-SOMO SAXS: IFT**

**Indirect Fourier Transform**

**Qmin [Angstrom<sup>-1</sup>]:** 0.0062862

**Qmax [Angstrom<sup>-1</sup>]:** 0.200505

Fit background

**Maximum diameter [Angstrom]:**   Fix

**Starting value for the Lagrange multiplier (Alpha):**   Fix

**Desmearing constant:**

**Number of points in p(r):** 50

**Number of extra error calculations:**

**Transformation/Regularization:**

- Debye (default -> returning p(r) with positivity constraint)
- Negative (Debye transformation -> returning p(r) without positivity constraint)
- MaxEnt using an ellipsoid of revolution as prior (-> p(r) -positivity constraint)
- Bessel (for cylindrical scatterers -> cross section distribution)
- Cosine (lamellae -> thickness distribution)
- Size (using spheres only -> size distribution)

Non-dilute solution

Help Close Process

**(r) Distribution Curve**

Distance (Angstrom)

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 Functions

**PDB Filename:** IADO\_noS04

**Definition files:**

- Load Atom Definition File:
- Load Hybridization File:
- Load SAXS Coefficients File:

**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<sup>2</sup> 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:**

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):**

**Smoothing:**

Raw  SAXS  SANS  Normalize

Residue contrib. range (Angstrom):

Compute P(r) Distribution

**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

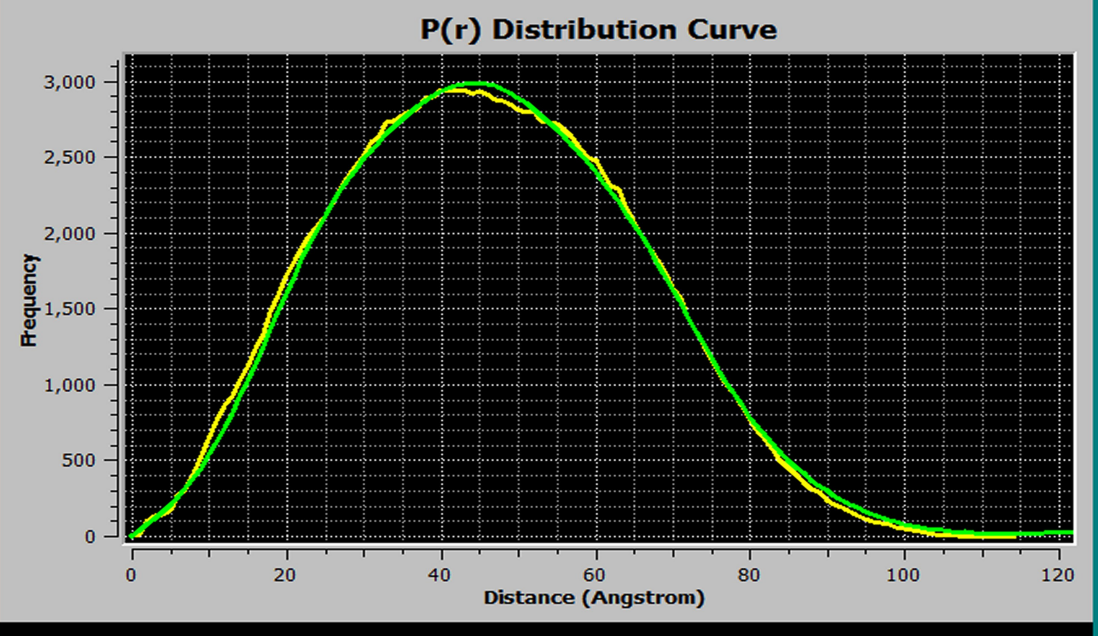
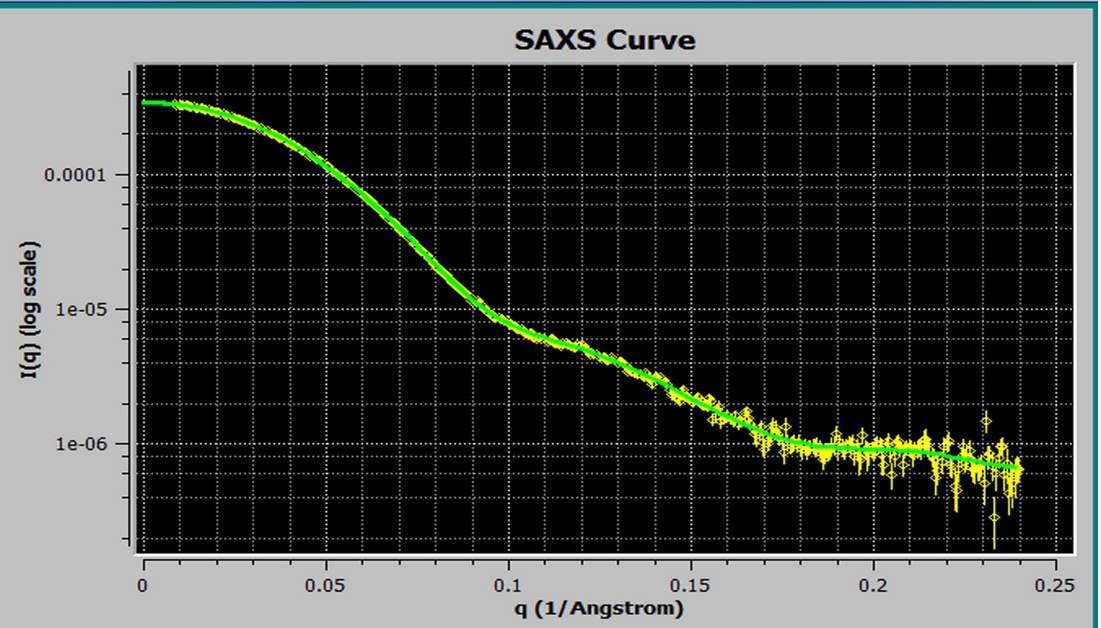
fitting range: 0.00857209 to 0.238167 with 402 points

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
    
```



# US-SOMO SAS MODULE

US-SOMO: SAS Functions

**PDB Filename:** 1ADO\_noS04

**Definition files:**

- Load Atom Definition File:
- Load Hybridization File:
- Load SAXS Coefficients File:

**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<sup>2</sup> 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:**

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):**

**Smoothing:**

Raw  SAXS  SANS  Normalize

Residue contrib. range (Angstrom):

Compute P(r) Distribution

**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

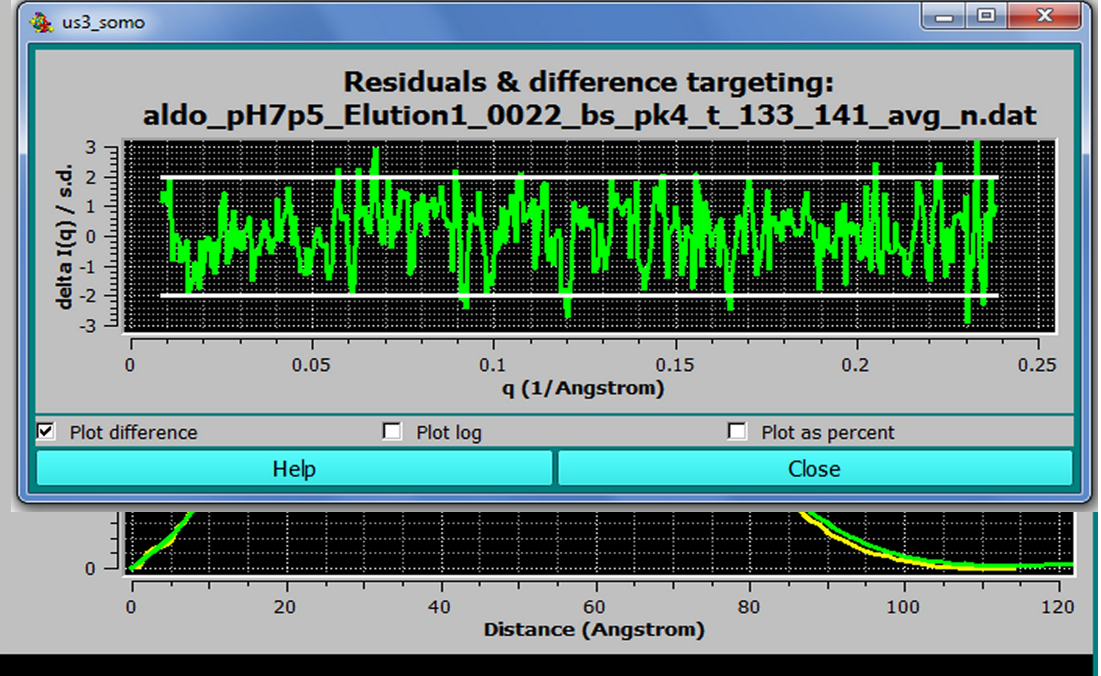
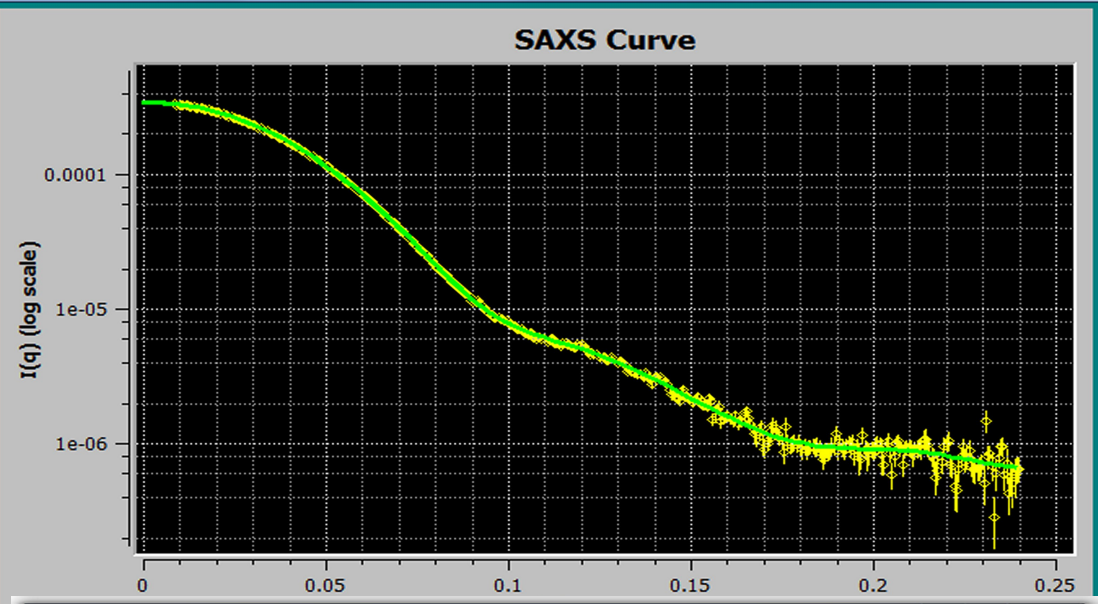
fitting range: 0.00857209 to 0.238167 with 402 points

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
    
```



# US-SOMO SAS MODULE

US-SOMO: SAS Functions

**PDB Filename:** 1ADO\_noSO4

**Definition files:**

- Load Atom Definition File:
- Load Hybridization File:
- Load SAXS Coefficients File:

**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

Standard  Kratky plot

Create standard output files

SAXS  F-DB  SH-DB  Q-DB  Crystol

SANS  F-DB  SH-DB  Q-DB  Cryston

**File suffix:**

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):**

**Smoothing:**

Raw  SAXS  SANS  Normalize

Residue contrib. range (Angstrom):

Compute P(r) Distribution

**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

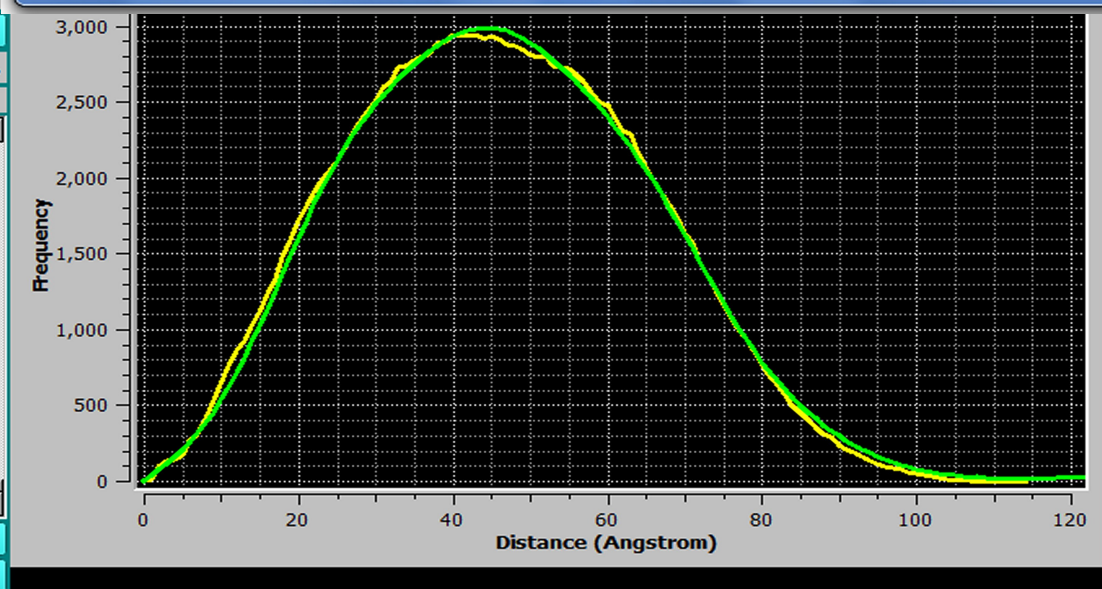
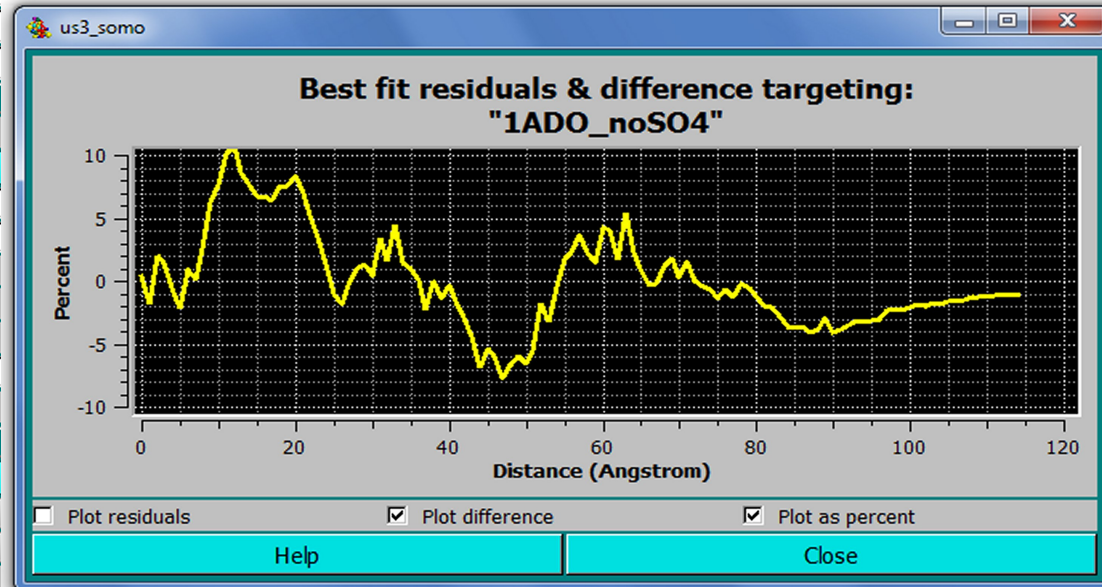
fitting range: 0.00857209 to 0.238167 with 402 points

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
    
```



# US-SOMO SAS NNLS utility

**US-SOMO: SAS Plotting Functions**

PDB Filename:

Definition files:

SAS I(q) Plotting Functions:

Load SAXS Curve | Load GNOM File | Load Plotted | Set Grid | Clear SAXS Curve

IFT | Search | Data | HPLC | Guinier | Legend

Guinier  CS  TV q<sup>2</sup> 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 Plotting Functions:

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

Bin size (Angstrom):  1

Smoothing:  0

Raw  SAXS  SANS  Normalize

Residue contrib. range (Angstrom):  Display

Compute P(r) Distribution

File

I(q) vs q plot legend:

results-1.csv "/root/andy/tgfb2tm\_bc\_bc\_open\_swap\_notails\_fixed.pdb  
Model: 1"

Chi<sup>2</sup> fitting  
fitting range: 0.00929495 to 0.310625 with 70 points  
Scaling factor: 0.999971 chi<sup>2</sup>=53.5353 df=69 nchi=0.880837  
sdf=0.0817236 nchi\*sdf=0.0719852  
results\_1.csv "/root/andy/tgfb2tm\_bc\_closed\_sub\_mm.pdb Model: 1"

Stop | Open Options Panel

Help | Close

**SAXS Curve**

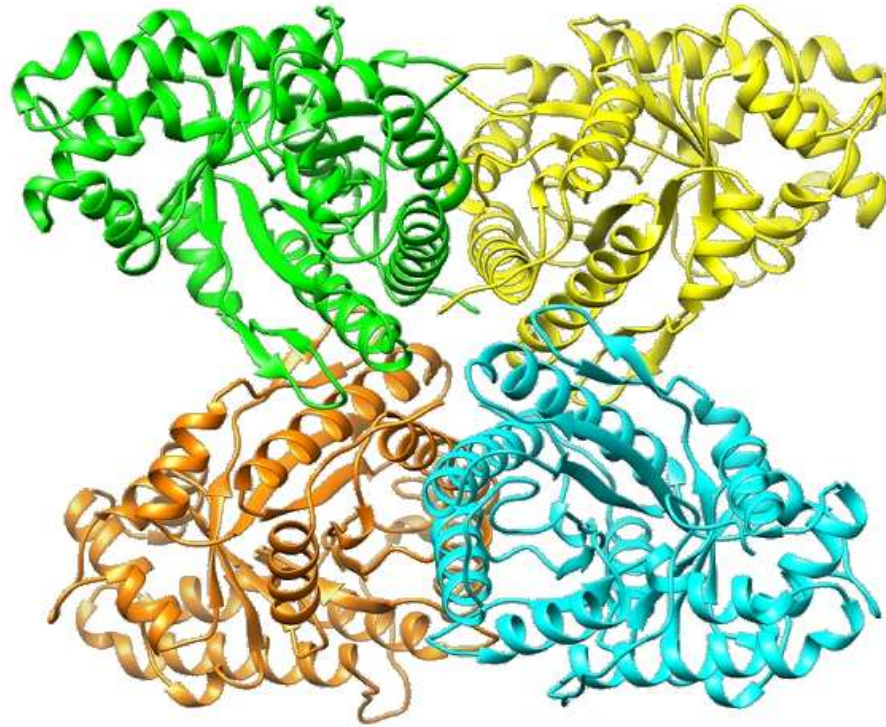
**us\_hydrodyn**

Residuals & difference targeting:  
c3sec.dat

Plot difference  Plot log  Plot as percent

Help | Close

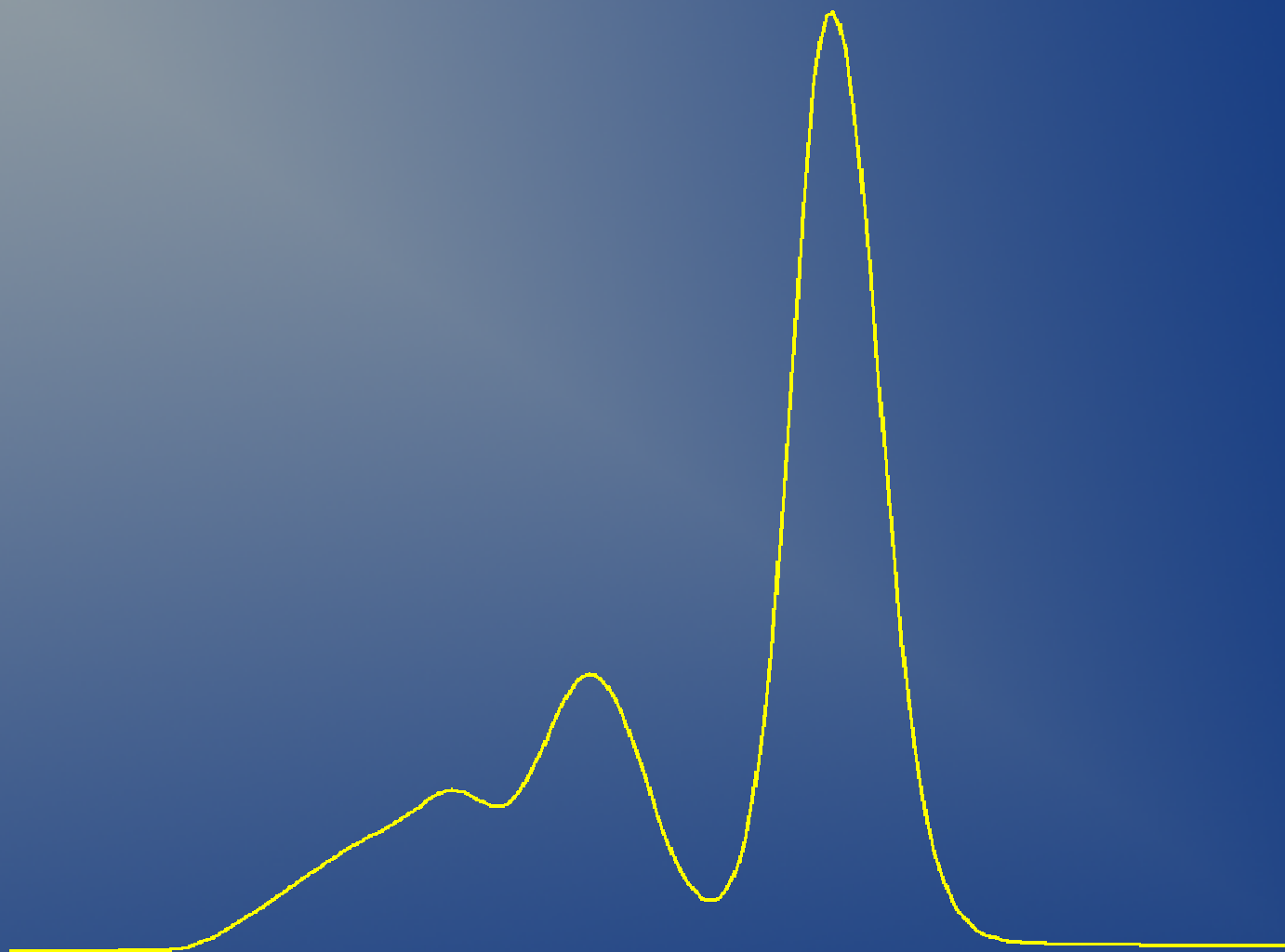
# The HPLC-SAXS module: SEC-SAXS of aldolase



A homotetramer whose crystal structure has been solved

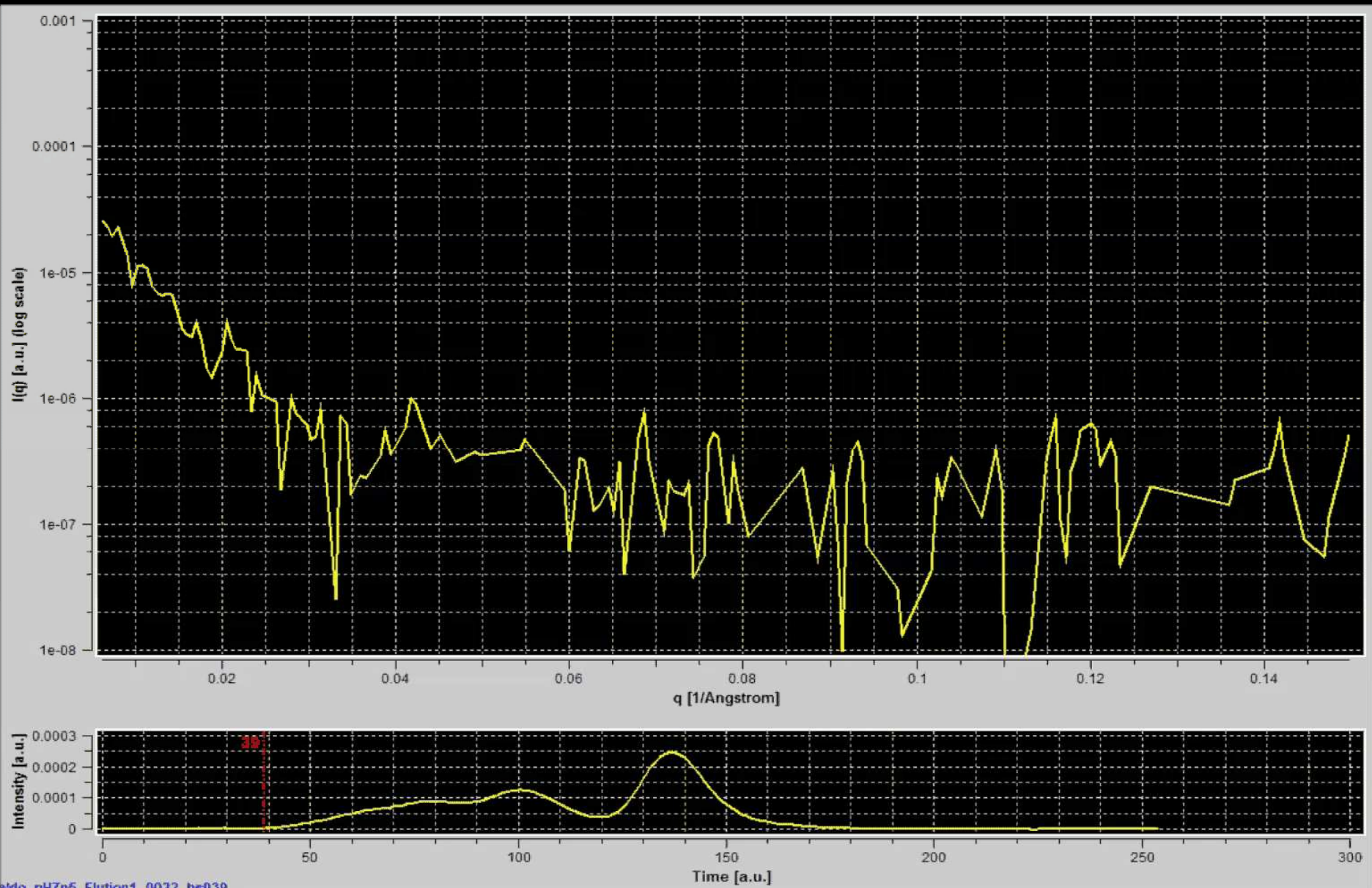
# SEC UV Trace of Aldolase

A<sup>280</sup>



Time

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



# US-SOMO HPLC-SAXS module

US-SOMO: SAXS HPLC

Developed by Emre Brookes, Javier Pérez, Patrice Vachette and Mattia Rocco (see *J. App. Cryst.* 46:1823-1833, 2013; *J. App. Cryst.* 49:1827-1841, 2016 )

**Data files**  
C:/Users/mattia/ultrascan/somo/saxs

Lock Add files Similar Concentrations Remove files

0 of 0 files selected

Sel. all Sel. Unsel. Adv. Sel. View Movie Log X Log Y  Err Rescale

Normalize Average To SOMO/SAS Width Color

Bin Smooth SVD Make I(t) Test I(t) Make I(q)

Concentration load Repeat Set Detector

**Produced Data**  
C:/Users/mattia/ultrascan/somo/saxs

0 of 0 files selected

Select all Invert Similar Remove Save CSV Save

Show Show only

**Messages**

File

3D Concentration reference Residuals CorMap Analysis Save Plots Cancel Keep

Blanks analysis Baseline Baseline apply Timeshift Timescale SD eval. SD apply

Gaussians Global Gaussians Scale Analysis Trial make I(q) Guinier

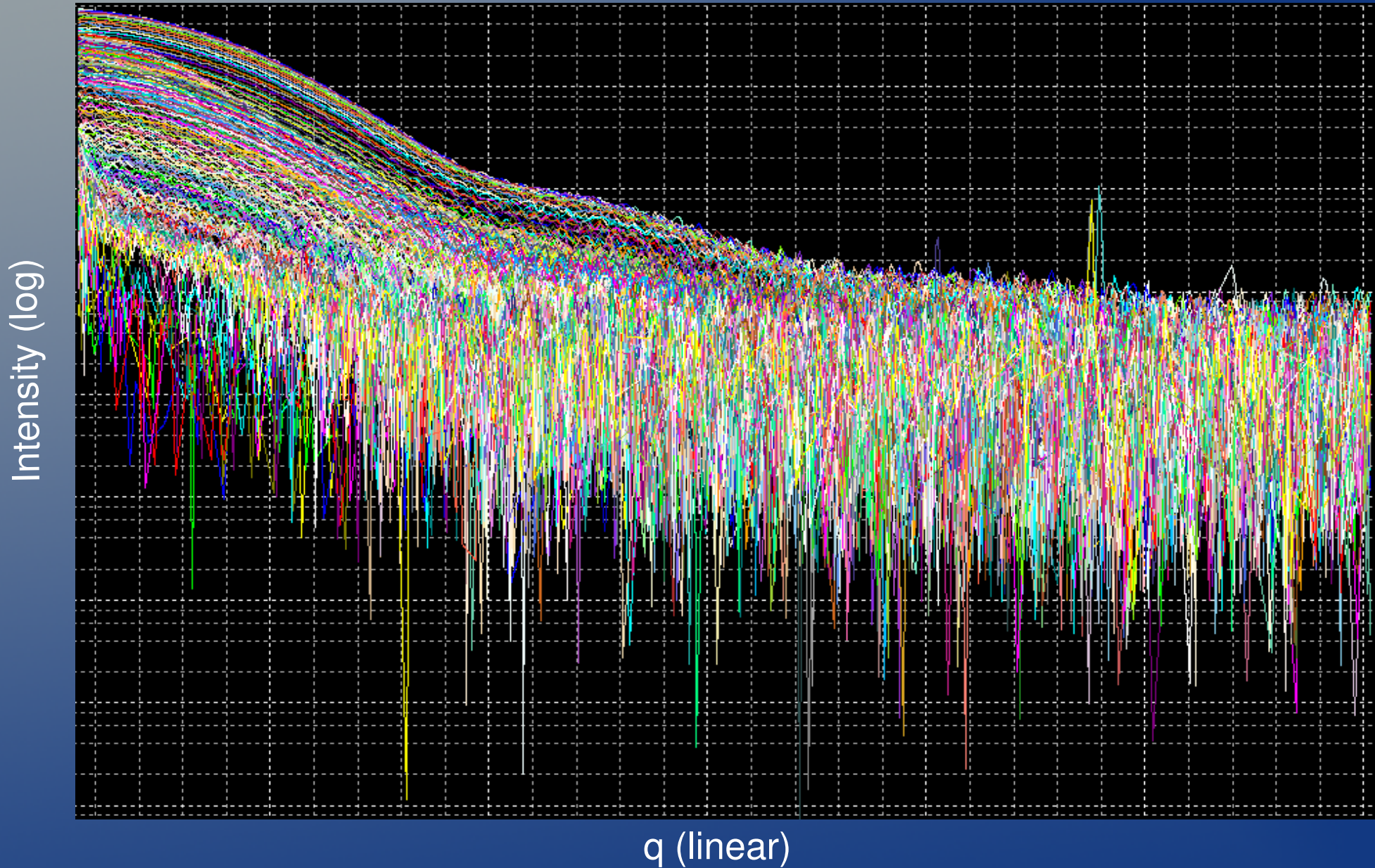
Select Visible Remove Vis Crop Common Crop Vis Crop Zeros Crop Left Undo Crop Right Legend

Help Options Close

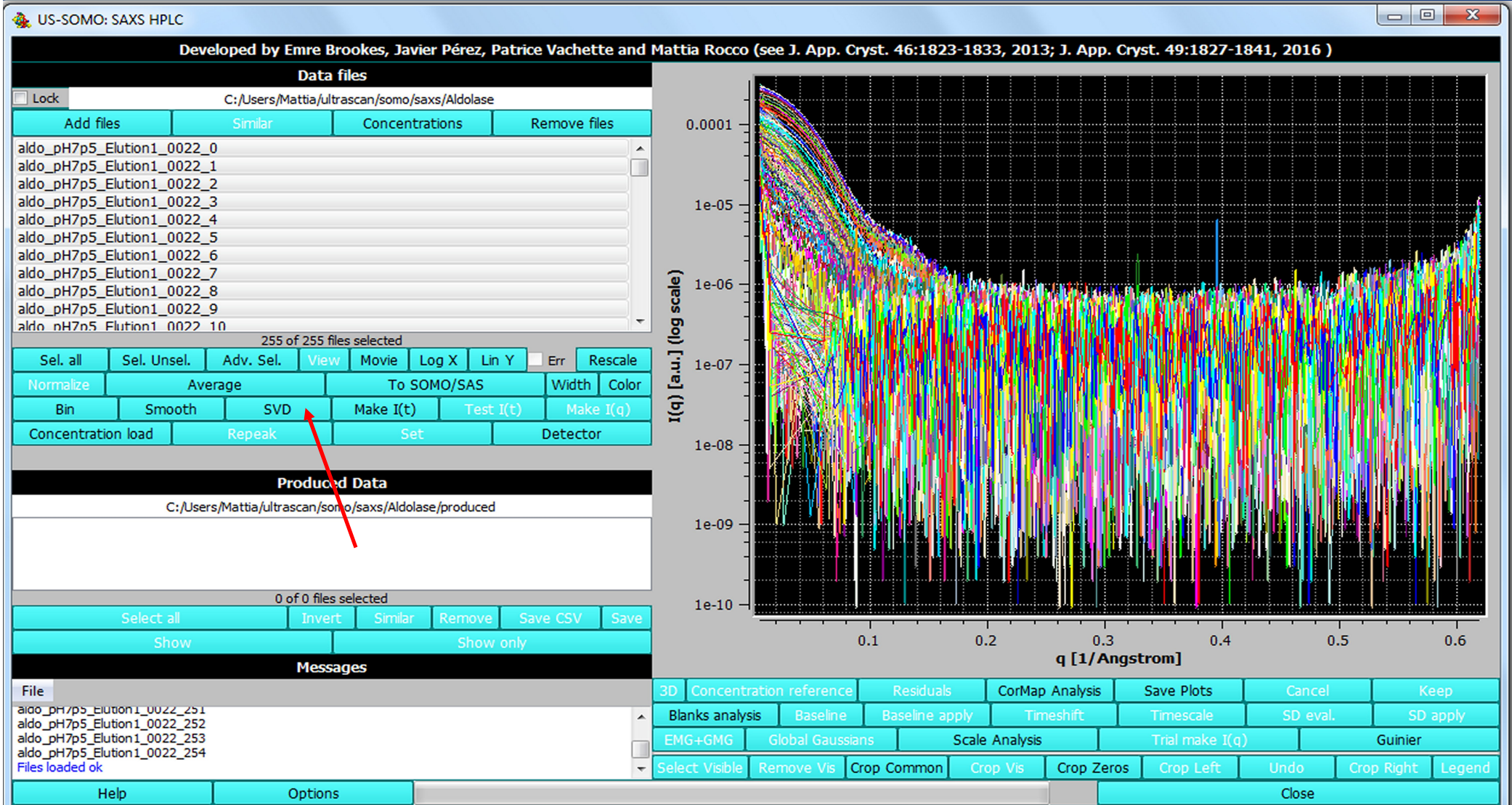




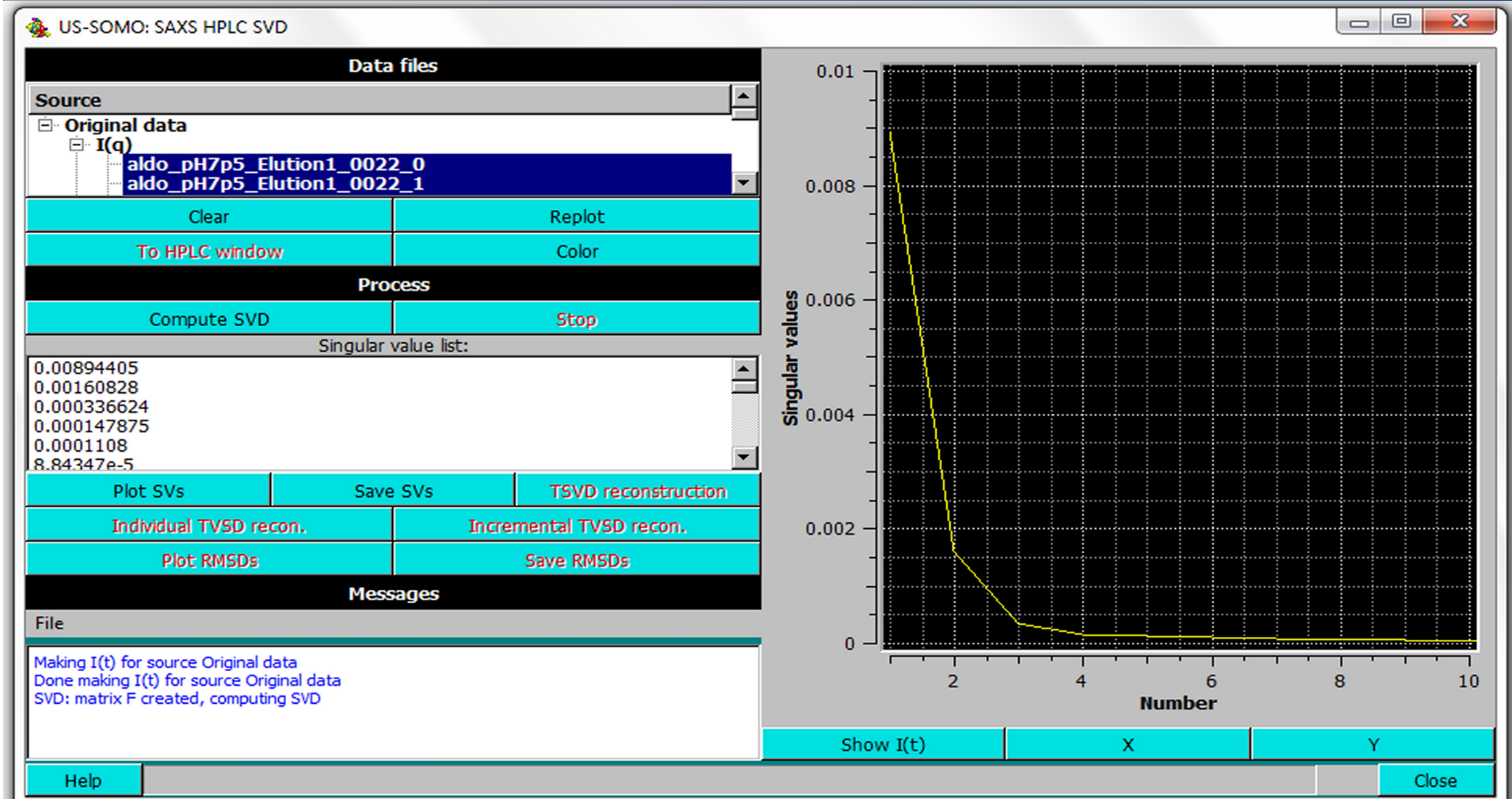
SEC-SAXS of aldolase: set of  $I(q)$  data  
each corresponding to a specific time or frame



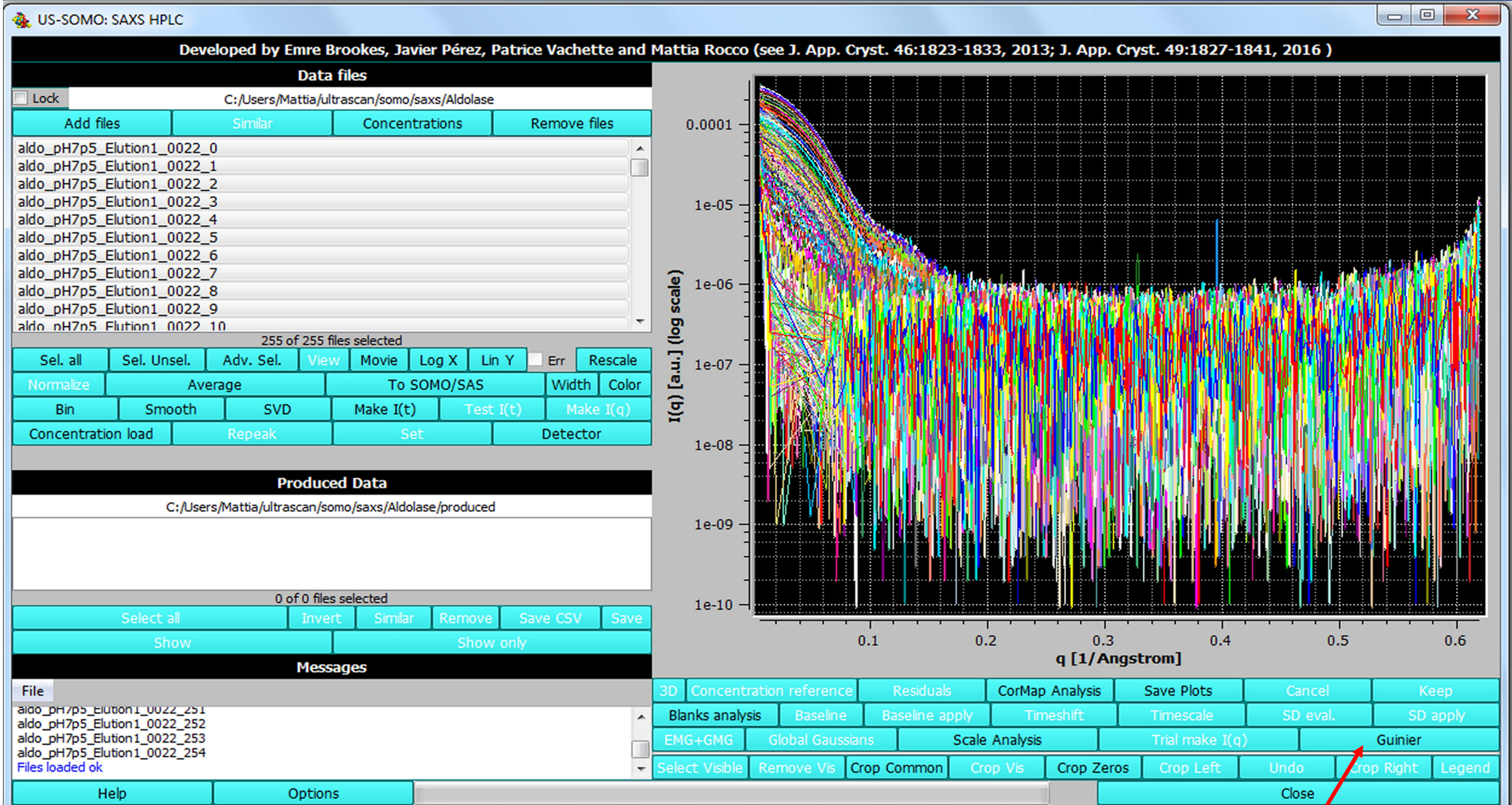
# HPLC-SAXS module with loaded I(q) vs. q data



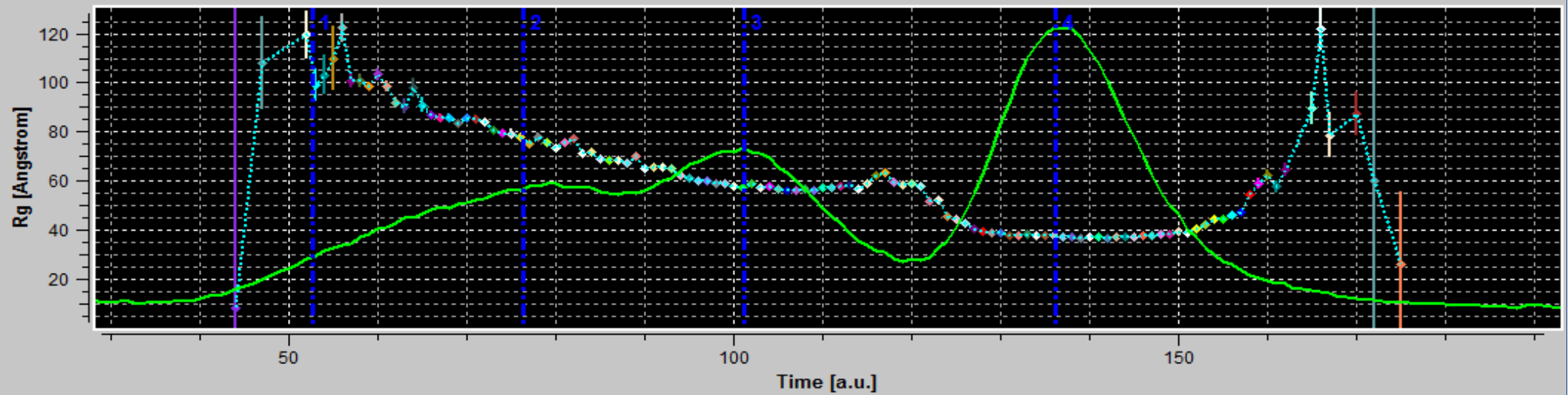
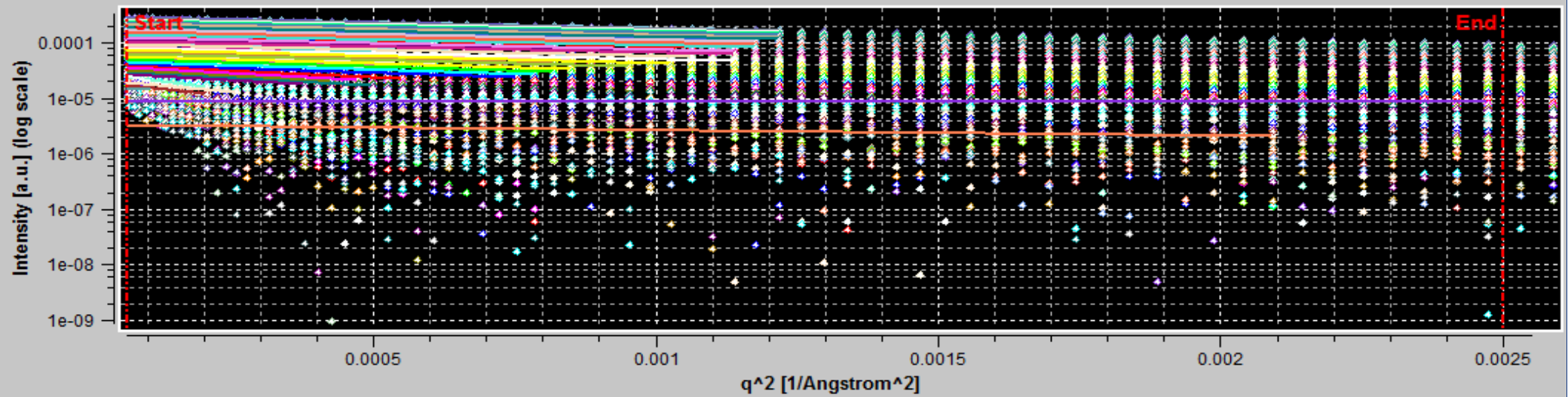
# HPLC-SAXS module with loaded $I(q)$ vs. $q$ data



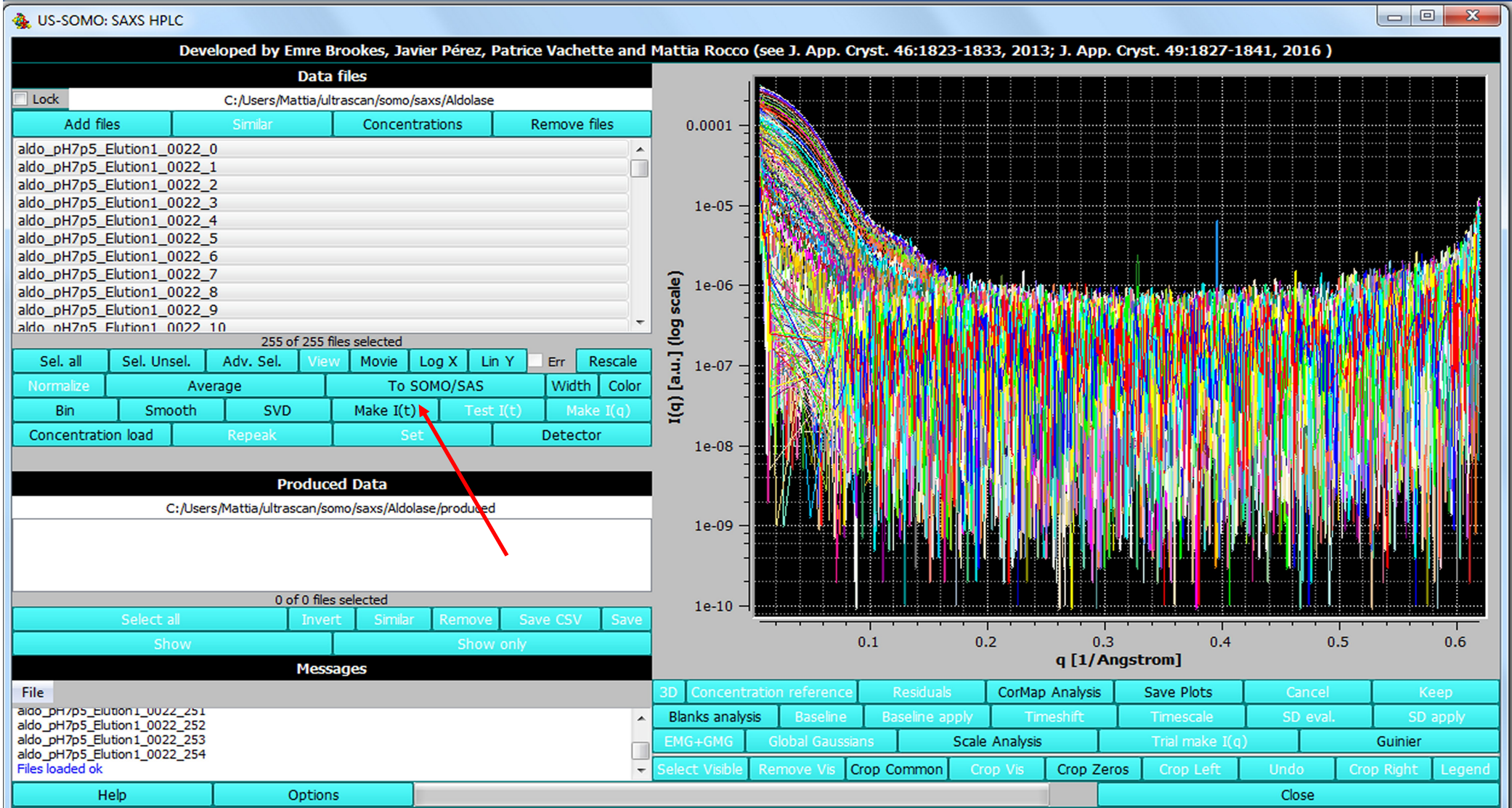
# HPLC-SAXS module with loaded I(q) vs. q data



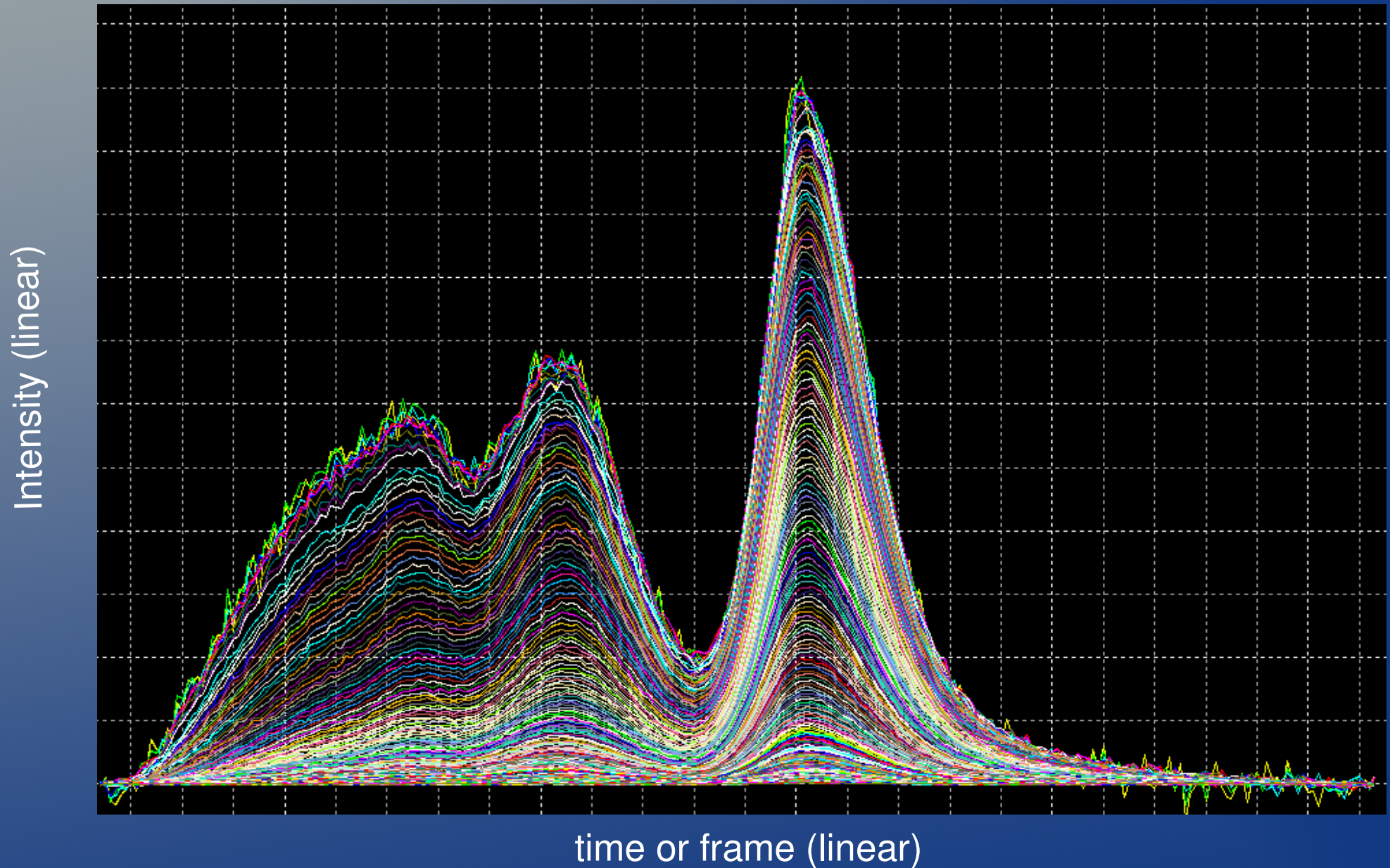
# Aldolase Rg – original data



# 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):

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

- Half-Gaussian modified Gaussian (GMG):

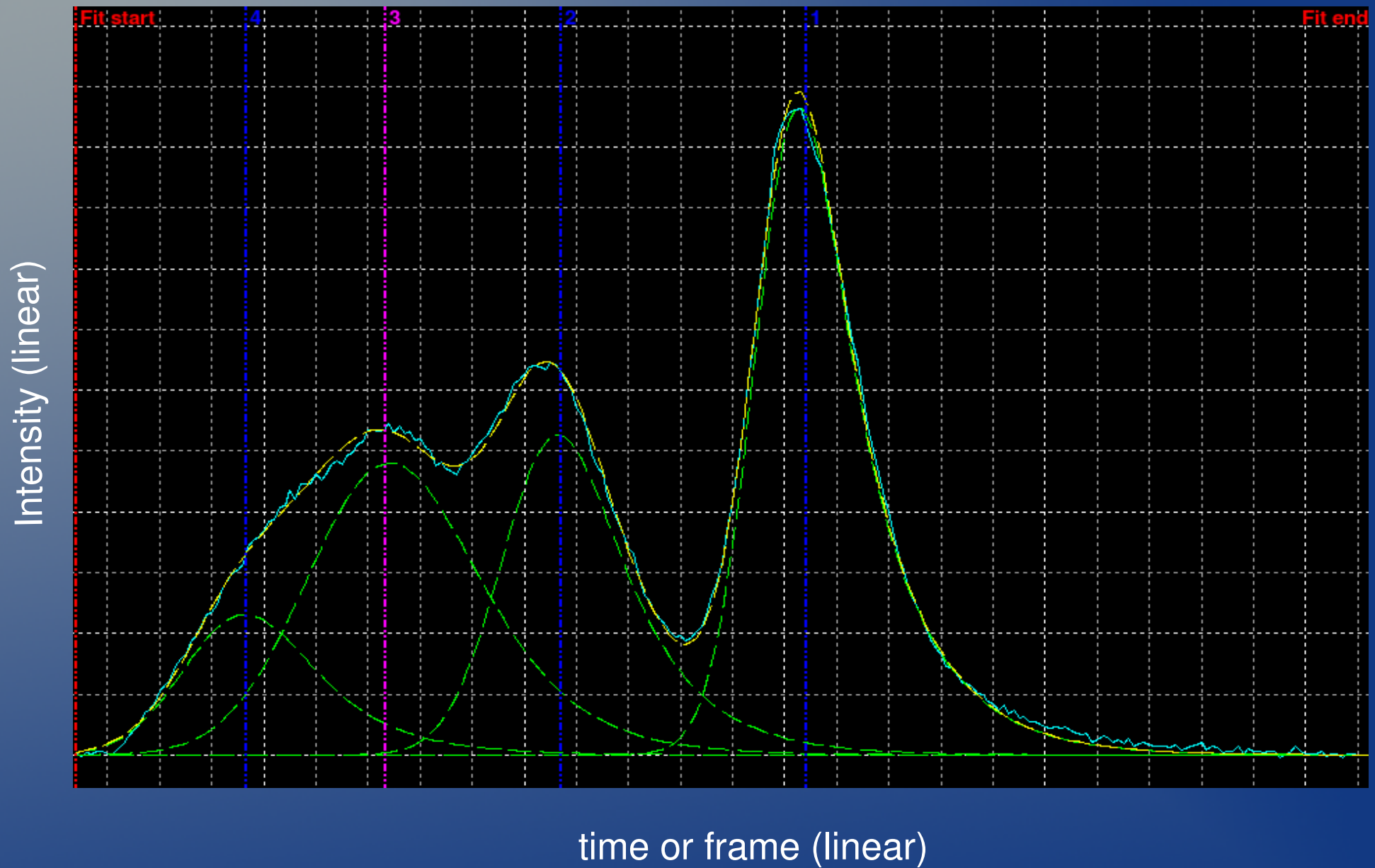
$$y = \frac{a_0 \exp\left(-\frac{1}{2}\frac{(x-a_1)^2}{a_3^2+a_2^2}\right) \left[1 + \operatorname{erf}\left(\frac{a_3(x-a_1)}{\sqrt{2}a_2\sqrt{a_3^2+a_2^2}}\right)\right]}{\sqrt{2\pi}\sqrt{a_3^2+a_2^2}}$$

- EMG+GMG:

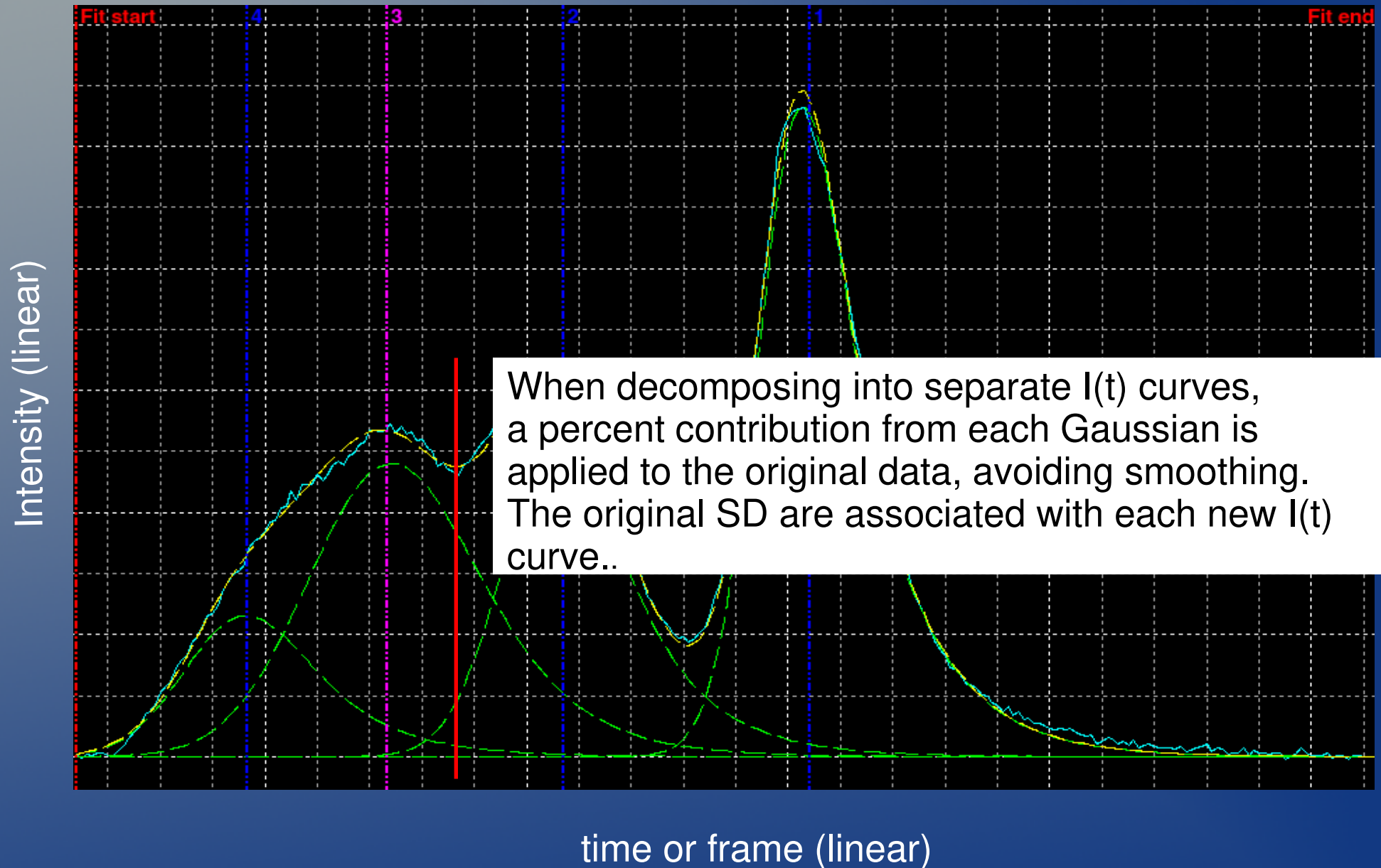
$$y = \frac{a_0}{4a_3} \exp\left(\frac{2a_1a_3 - 2a_3x + a_2^2}{a_3^2}\right) \operatorname{erfc}\left(\frac{a_1a_3 - a_3x + a_2^2}{\sqrt{2}a_2a_3}\right) + \frac{a_0}{2\sqrt{2\pi}\sqrt{a_2^2+a_4^2}} \exp\left(-\frac{1}{2}\frac{(a_1-x)^2}{a_2^2+a_4^2}\right) \operatorname{erfc}\left(\frac{a_2(a_1-x)}{\sqrt{2}a_2\sqrt{a_2^2+a_4^2}}\right)$$



# Aldolase EMG+GMG fit of one $I(t)$



# Aldolase EMG+GMG fit of one $I(t)$



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

Developed by Enrico Brocchini, Daniel Brocchini, Gianni Vascotto and Mattia Rossi (see <http://dx.doi.org/10.1002/eqm2.1216>)

**Data files**

Lock  C:/WORK\_EMRE/US\_SOMO/Aldolase

Add files	Similar	Concentrations	Remove files
aldo_pH7p5_Elution1_0022__It_q0_613126			
aldo_pH7p5_Elution1_0022__It_q0_613693			
aldo_pH7p5_Elution1_0022__It_q0_614823			
aldo_pH7p5_Elution1_0022__It_q0_615388			
aldo_pH7p5_Elution1_0022__It_q0_615953			
aldo_pH7p5_Elution1_0022__It_q0_616518			
aldo_pH7p5_Elution1_0022__It_q0_617083			
aldo_pH7p5_Elution1_0022__It_q0_617647			
aldo_pH7p5_Elution1_0022__It_q0_618212			
aldo_pH7p5_Elution1_0022__It_q0_618777			

154 of 1327 files selected

Set all	Set Unsel.	Adv. Sel.	View	Move	Log X	Log Y	Err	Rescale
Normalize	Average		To SOMO/SAS		Width	Color		
Bin	Smooth	SVD	Make I(t)	Test I(t)	Make I(q)			
Concentration load	Repeak	Set	Detector					

**Produced Data**

C:/WORK\_EMRE/US\_SOMO/Aldolase/produced

aldo_pH7p5_Elution1_0022__It_q0_615388
aldo_pH7p5_Elution1_0022__It_q0_615953
aldo_pH7p5_Elution1_0022__It_q0_616518
aldo_pH7p5_Elution1_0022__It_q0_617083
aldo_pH7p5_Elution1_0022__It_q0_617647
aldo_pH7p5_Elution1_0022__It_q0_618212
aldo_pH7p5_Elution1_0022__It_q0_618777

0 of 1072 files selected

Select all	Invert	Similar	Remove	Save CSV	Save
Show	Show only				

Reverse  Use standard deviations  By percent  Group

**Messages**

File

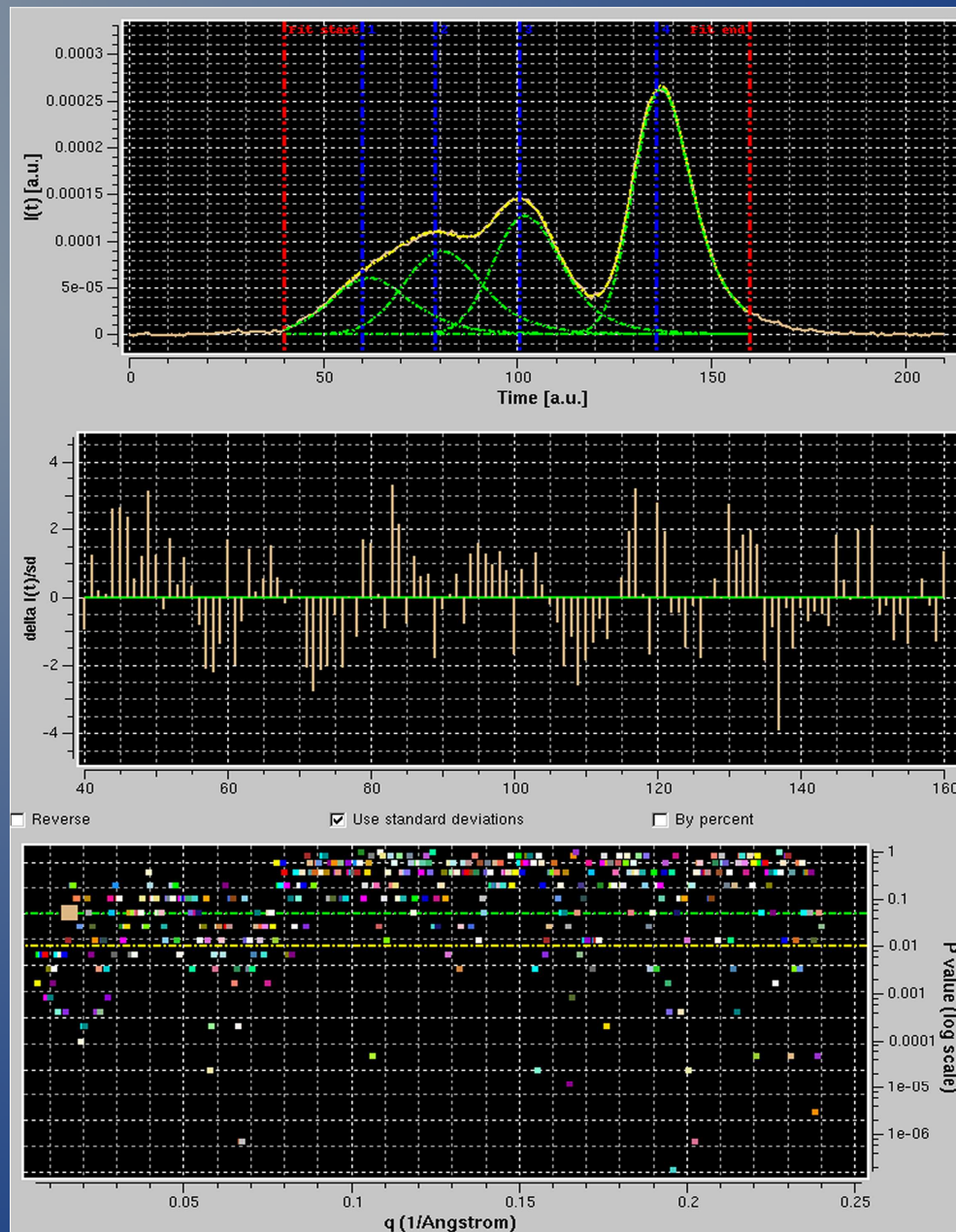
37.1% 0.01 > P pairs  
 P value analysis summary:  
 62.0% P >= 0.01 (42.6% P >= 0.05) + (19.4% 0.05 > P >= 0.01) pairs  
 38.0% 0.01 > P pairs  
 P value analysis summary:  
 70.8% P >= 0.01 (52.6% P >= 0.05) + (18.2% 0.05 > P >= 0.01) pairs  
 29.2% 0.01 > P pairs

Top plot:  $I(t)$  [a.u.] vs Time [a.u.]

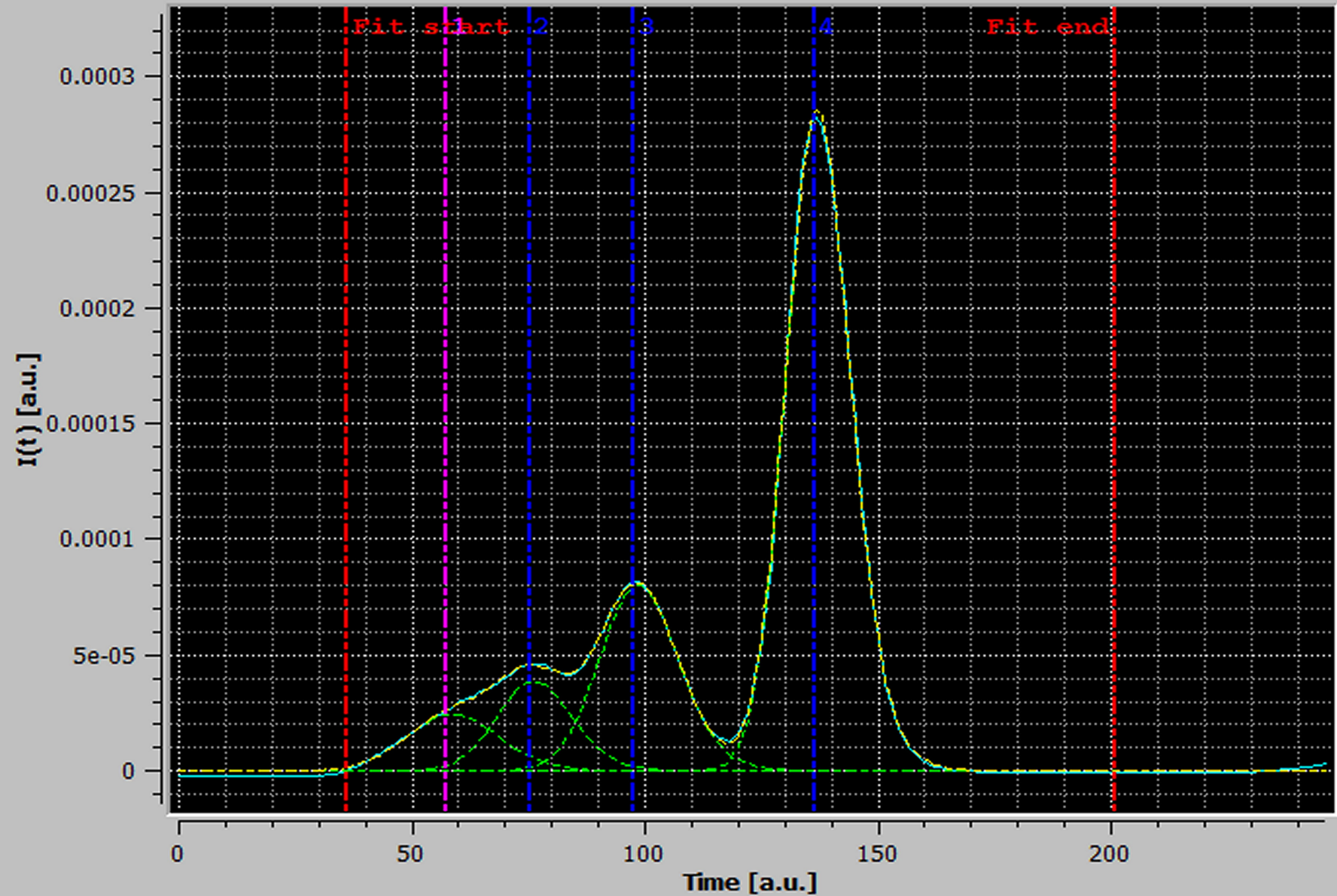
Bottom plot:  $\Delta I(t)/sd$  vs Time [a.u.]

3D	Concentration reference	Residuals	Show CorMap	Save Plots	Global fit by q	Cancel	Keep	
EMG+GMG	Global Gaussians	Scale Analysis	Trial make I(q)		Guinier			
<input type="checkbox"/> Scroll	<input checked="" type="checkbox"/> P >= 0.05	<input checked="" type="checkbox"/> 0.05 > P >= 0.01	<input checked="" type="checkbox"/> P < 0.01	Make result curves	To produced data			
< 4 of 4 >	136.155	6.26136	10.5099	-1.48966	Save			
<input checked="" type="checkbox"/> SD	<input checked="" type="checkbox"/> Eq width	<input checked="" type="checkbox"/> Eq dist1	<input checked="" type="checkbox"/> Eq dist2	Global Fit	Recompute nChi <sup>2</sup> 1.8211	33.1787	199.966	
Help							Options	Close

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



# HPLC-SAXS module: UV Gaussians fit



# HPLC-SAXS module: Make I(q) panel

US-SOMO: SAXS HPLC : Make I(q)

**US-SOMO: SAXS HPLC : Make I(q)**

Create sum of peaks curves

Add SD computed %-wise from the difference between the sum of Gaussians and the original I(q)

If zeros are produced when computing SDs:  Average adjacent SDs  Set to 0.1 % of peak's I(q)

Average and normalize resulting I(q) curves by Gaussian, using top % of max. intensity 5

Do you want to set the concentration file Gaussians centers, widths and skewness to the SAXS-optimized values, adjusting the amplitudes and keeping the areas constant?

This implies that all the species that were defined as Gaussians contributing to the SAXS signal also contribute to the concentration signal.  
**Be aware that this option will result in an apparent mass artificially approximately constant along each of the deconvoluted Gaussian peaks, reflecting just the oscillations in the original SAXS data.**  
However, the apparent average mass for each peak should be a closer approximation to the real value when significant band broadening occurs between the concentration and the SAXS detectors.

I0 standard experimental value (a.u.) :

**Concentrations will be computed and will be written along with PSVs to the output I(q) curves**

Gaussian	Extinction coefficient (ml mg <sup>-1</sup> cm <sup>-1</sup> )	Partial specific volume (ml/g)
1	0.877	0.736
2	0.877	0.736
3	0.877	0.736
4	0.877	0.736

Duplicate Gaussian 1 values globally

Help      Quit      Make I(q) without Gaussians      Continue

# HPLC-SAXS module: Make I(q) panel

US-SOMO: SAXS HPLC : Make I(q)

### US-SOMO: SAXS HPLC : Make I(q)

Create sum of peaks curves

Add SD computed %-wise from the difference between the sum of Gaussians and the original I(q)

**If zeros are produced when computing SDs:**  Average adjacent SDs  Set to 0.1 % of peak's I(q)

Average and normalize resulting I(q) curves by Gaussian, using top % of max. intensity

Do you want to set the concentration file Gaussians centers, widths and skewness to the SAXS-optimized values, adjusting the amplitudes and keeping the areas constant?

This implies that all the species that were defined as Gaussians contributing to the SAXS signal also contribute to the concentration signal. **Be aware that this option will result in an apparent mass artificially approximately constant along each of the deconvoluted Gaussian peaks, reflecting just the oscillations in the original SAXS data.**

However, the apparent average mass for each peak should be a closer approximation to the real value when significant band broadening occurs between the concentration and the SAXS detectors.

I0 standard experimental value (a.u.) :

**Concentrations will be computed and will be written along with PSVs to the output I(q) curves**

Gaussian	Extinction coefficient (ml mg <sup>-1</sup> cm <sup>-1</sup> )	Partial specific volume (ml/g)
1	0.877	0.736
2	0.877	0.736
3	0.877	0.736
4	0.877	0.736

Duplicate Gaussian 1 values globally

Help Quit Make I(q) without Gaussians Continue

# HPLC-SAXS module: Make I(q) panel

US-SOMO: SAXS HPLC : Make I(q)



## US-SOMO: SAXS HPLC : Make I(q)

Create sum of peaks curves

Add SD computed %-wise from the difference between the sum of Gaussians and the original I(q)

**If zeros are produced when computing SDs:**  Average adjacent SDs  Set to 0.1 % of peak's I(q)

Average and normalize resulting I(q) curves by Gaussian, using top % of max. intensity

Do you want to set the concentration file Gaussians centers, widths and skewness to the SAXS-optimized values, adjusting the amplitudes and keeping the areas constant?

This implies that all the species that were defined as Gaussians contributing to the SAXS signal also contribute to the concentration signal.

**Be aware that this option will result in an apparent mass artificially approximately constant along each of the deconvoluted Gaussian peaks, reflecting just the oscillations in the original SAXS data.**

However, the apparent average mass for each peak should be a closer approximation to the real value when significant band broadening occurs between the concentration and the SAXS detectors.

I0 standard experimental value (a.u.) :

### Concentrations will be computed and will be written along with PSVs to the output I(q) curves

Gaussian	Extinction coefficient (ml mg <sup>-1</sup> cm <sup>-1</sup> )	Partial specific volume (ml/g)
1	0.877	0.736
2	0.877	0.736
3	0.877	0.736
4	0.877	0.736

Duplicate Gaussian 1 values globally

Help

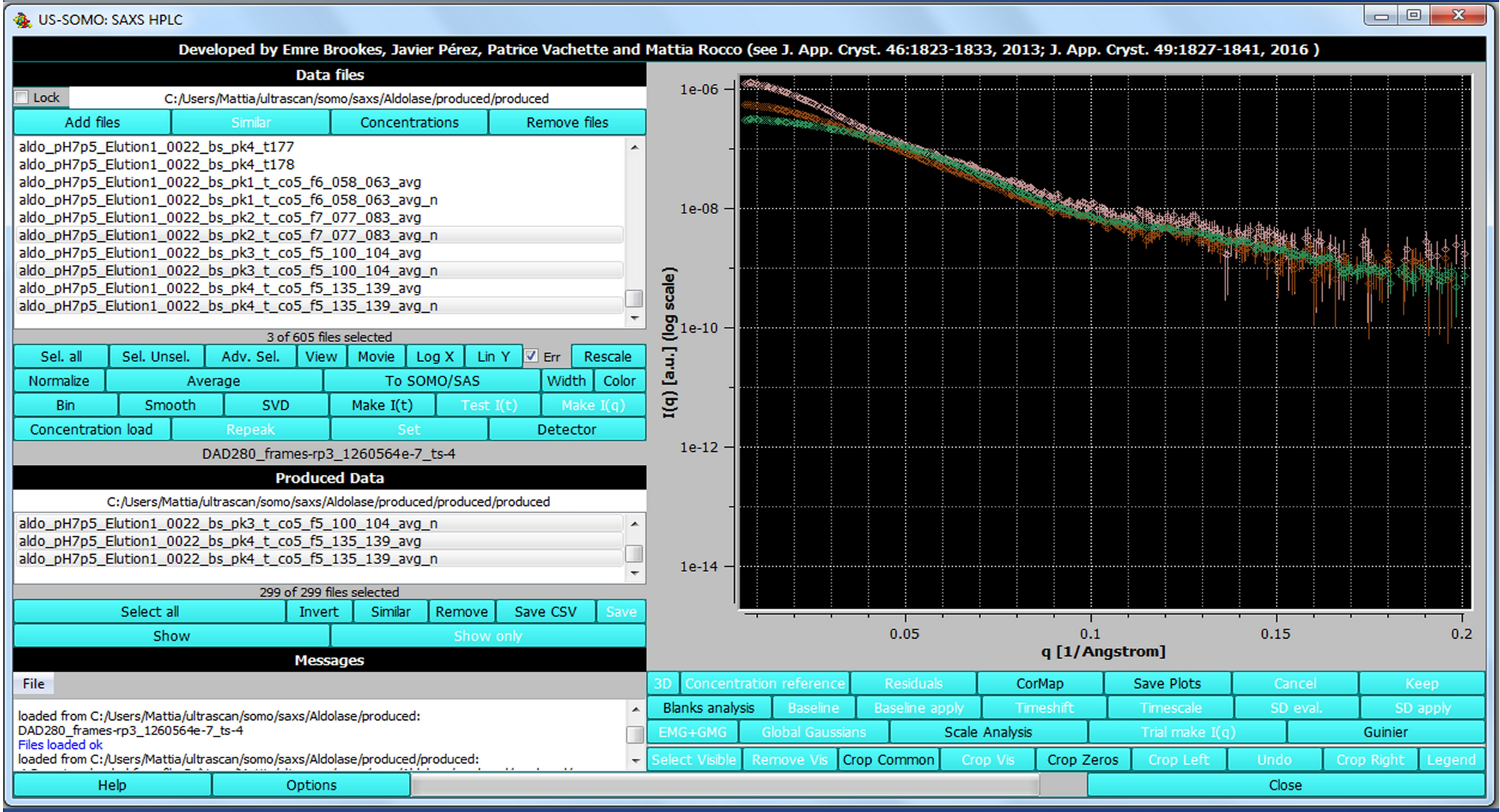
Quit

Make I(q) without Gaussians

Continue

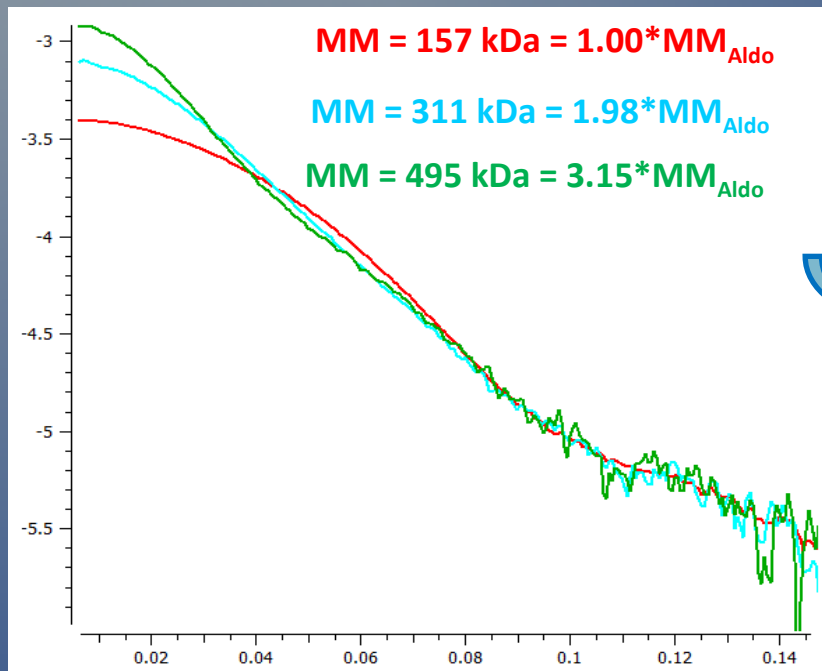


# HPLC-SAXS module: auto-averaged final decomposed $I(q)$ vs. $q$ datasets

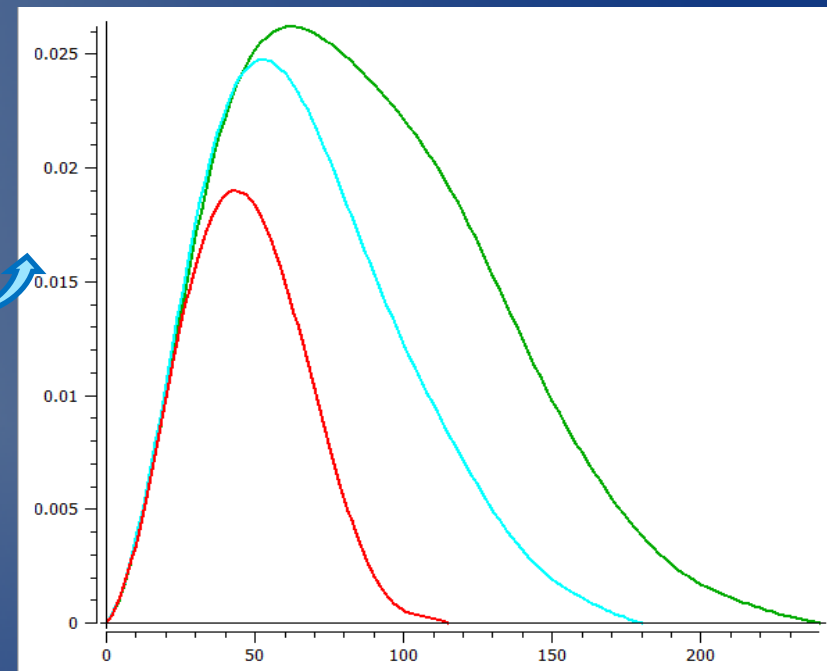


# SEC-SAXS : Molar Mass (MM) of aldolase components pK<sub>4</sub>, pK<sub>3</sub> & pK<sub>2</sub>

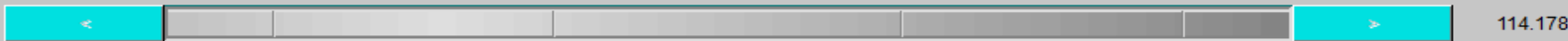
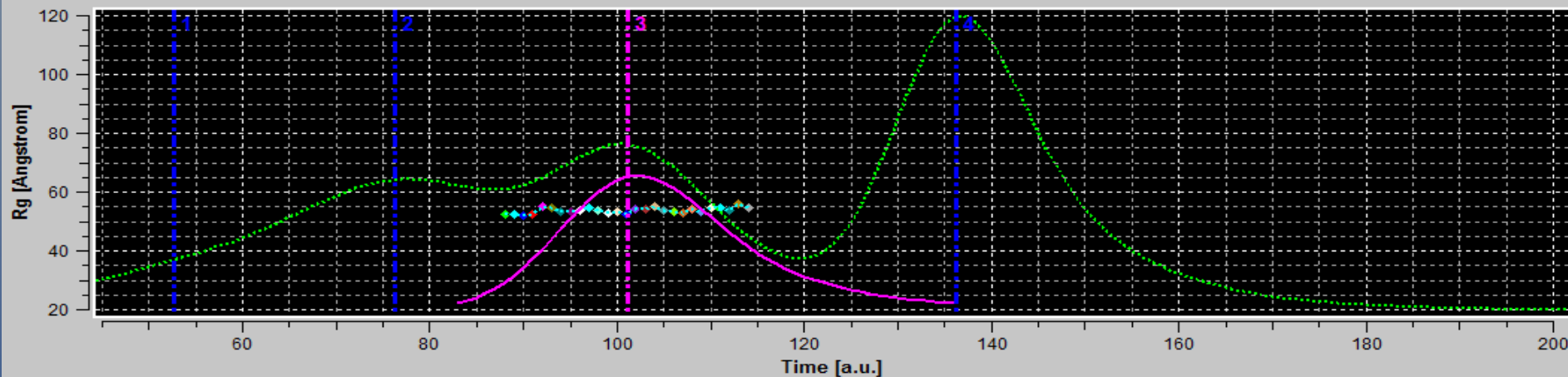
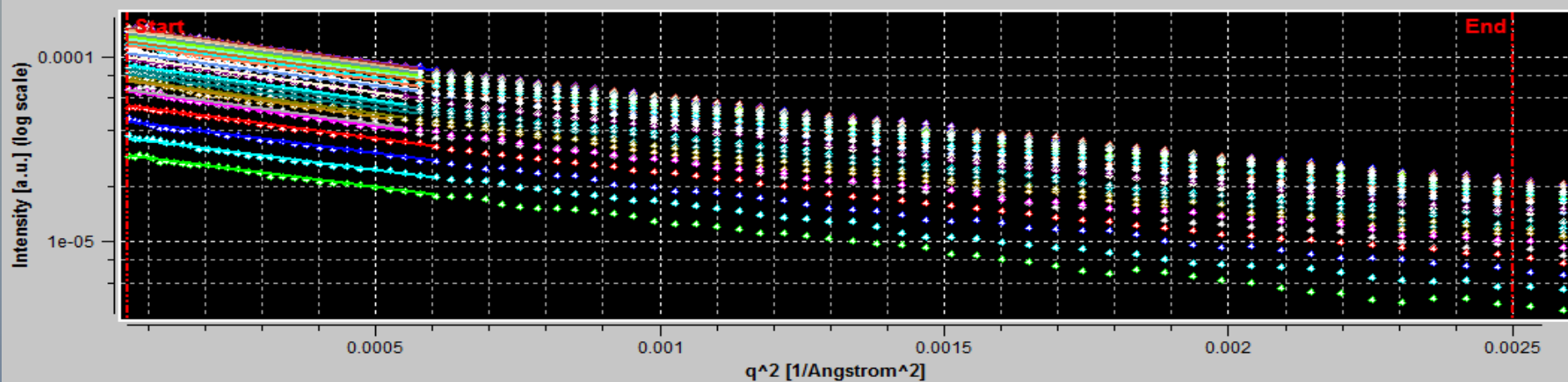
I(q)



p(r)



# US-SOMO HPLC-SAXS: aldolase Test I(q) 3<sup>rd</sup> peak



Guinier

3D Concentration reference Rg plot Approx. MW plot Residuals CorMap Save Plots Cancel Keep

EMG+GMG Global Gaussians Scale Analysis Trial make I(q) Guinier

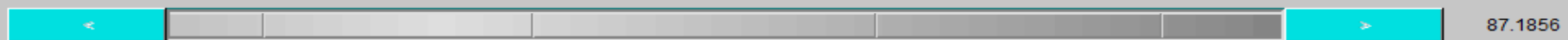
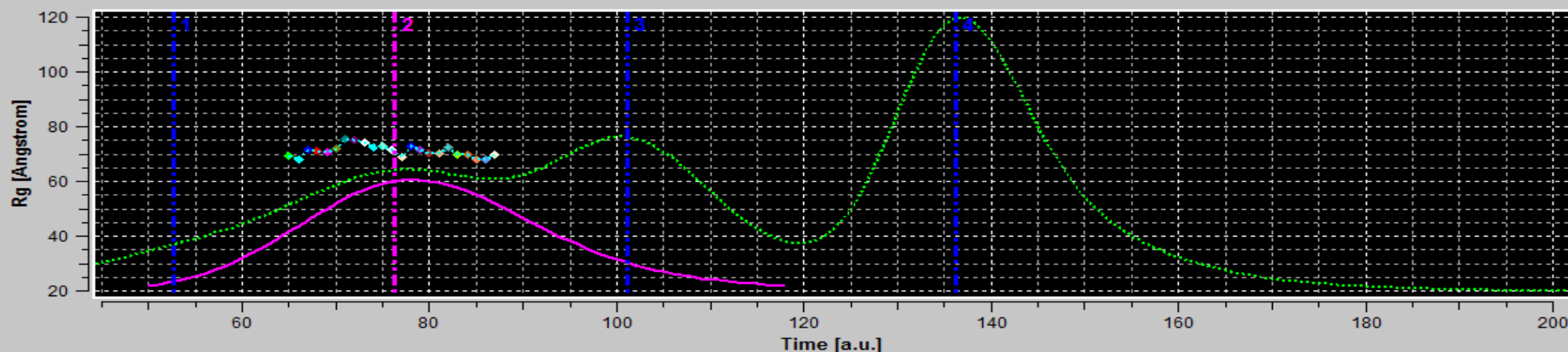
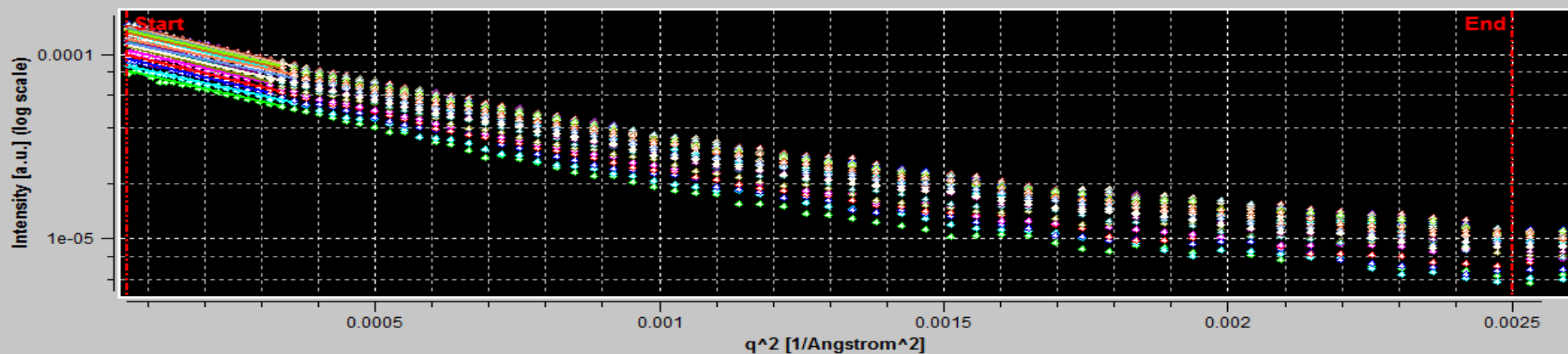
Time range for Rg plot: 45 200 Rg range: 20 120  Lock range Replot

Scroll q range: 0.00800062 0.05 plot extension: 5e-05 5e-05  SD  qmax\*Rg limit: 1.3

Avg. 27 curves qmax\*Rg 1.286 [1.271:1.299] Rg 53.7 (1.0) [52.1:56.0] I0 1.07e-04 (3.72e-05) [3.12e-05:1.53e-04]

MW[RT] Avg. 3.334e+05 (1.493e+04) [3.134e+05:3.774e+05]

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



Guinier

3D Concentration reference Rg plot Approx. MW plot Residuals CorMap Save Plots Cancel Keep

EMG+GMG Global Gaussians Scale Analysis Trial make I(q) Guinier

Time range for Rg plot: 45 200 Rg range: 20 120  Lock range Replot

Scroll q range: 0.00800062 0.05 plot extension: 5e-05 5e-05  SD  qmax\*Rg limit: 1.3

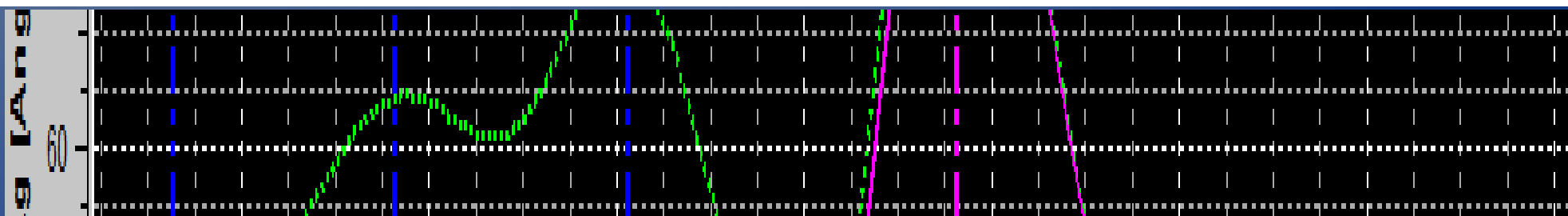
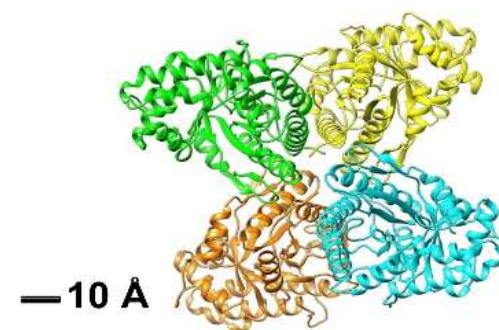
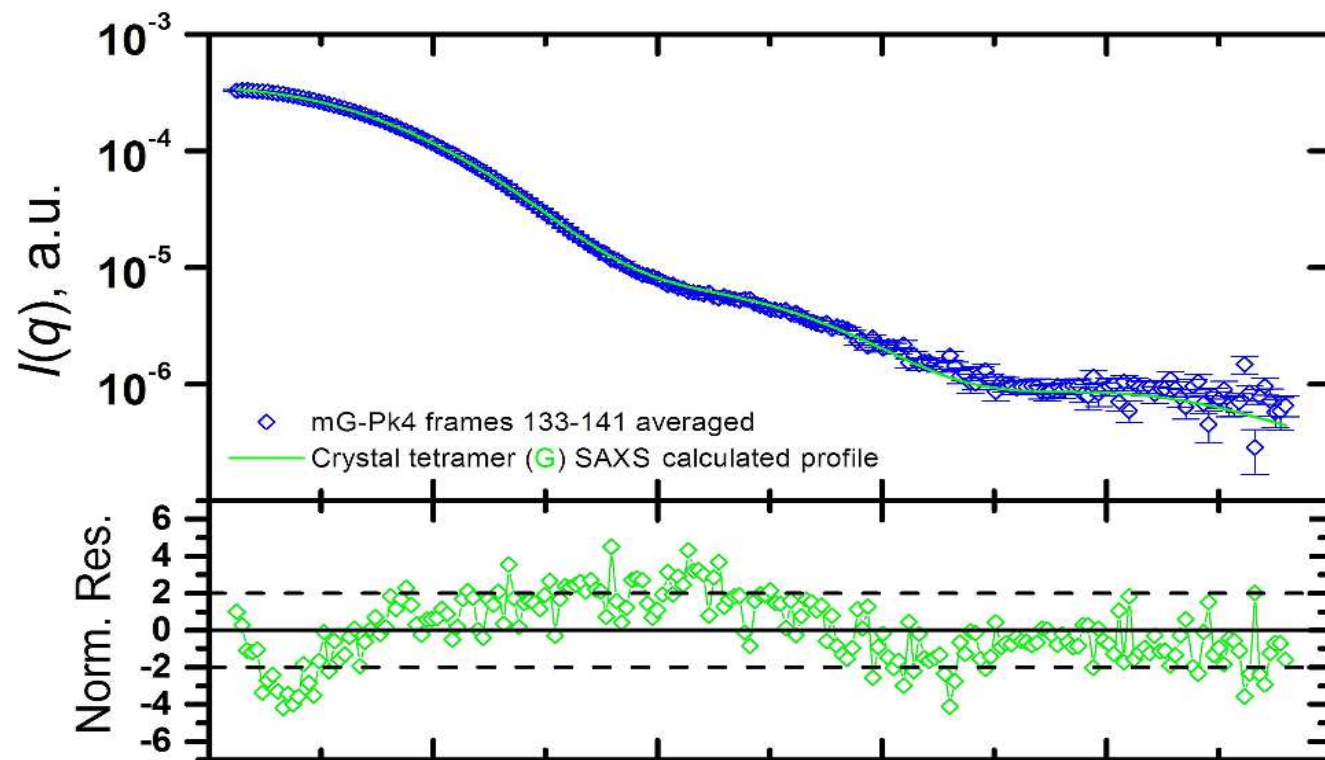
Avg. 23 curves qmax\*Rg 1.281 [1.251:1.299] Rg 71.2 (2.1) [68.1:75.4] I0 1.36e-04 (2.11e-05) [8.98e-05:1.60e-04]

MW[RT] Avg. 6.609e+05 (3.892e+04) [6.093e+05:7.409e+05]

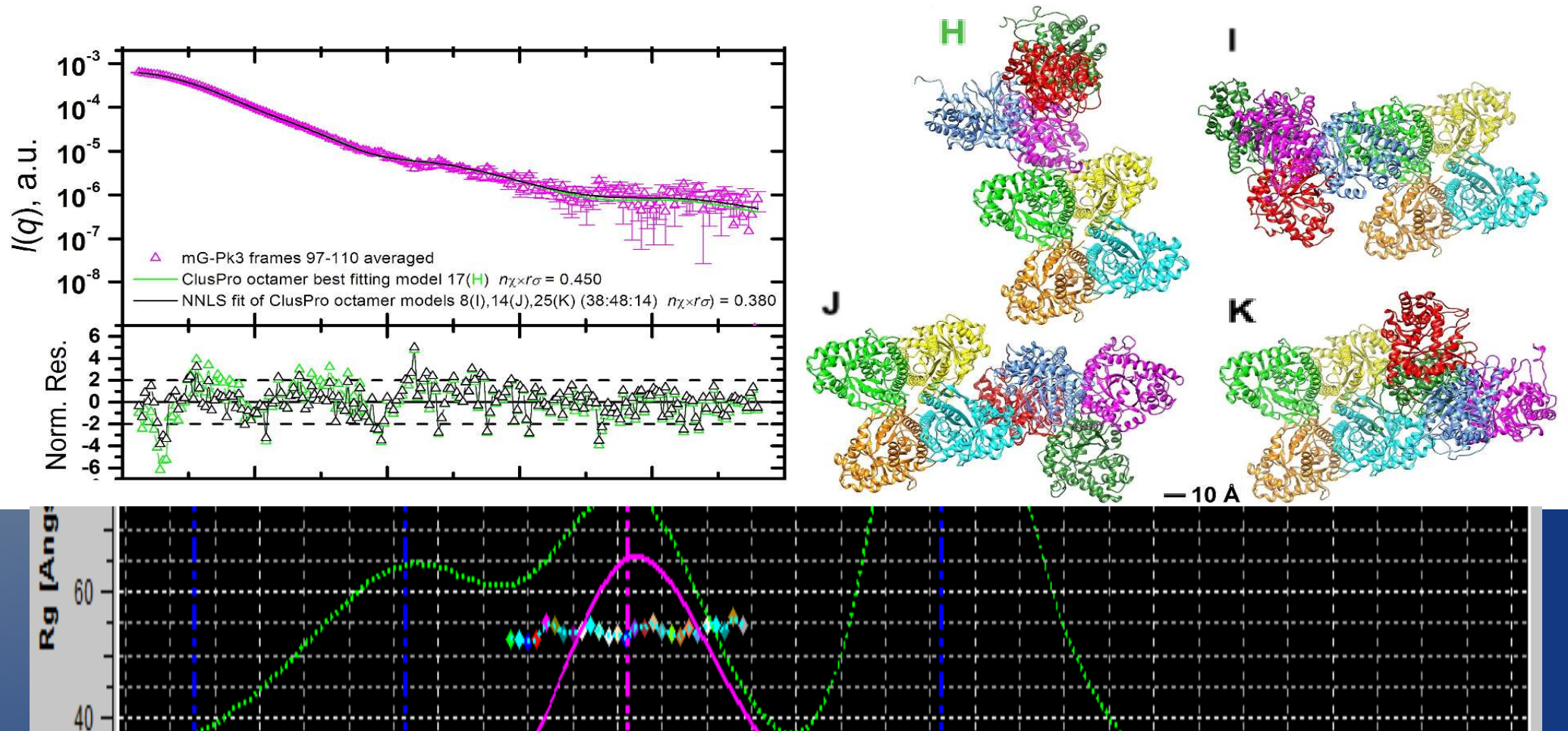
# Modeling

- “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

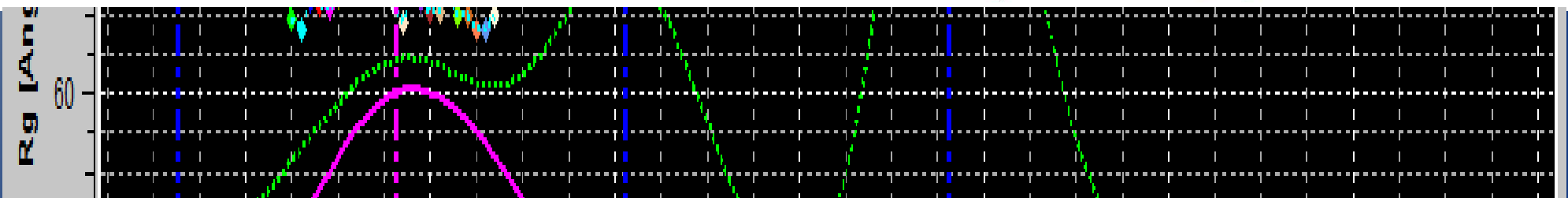
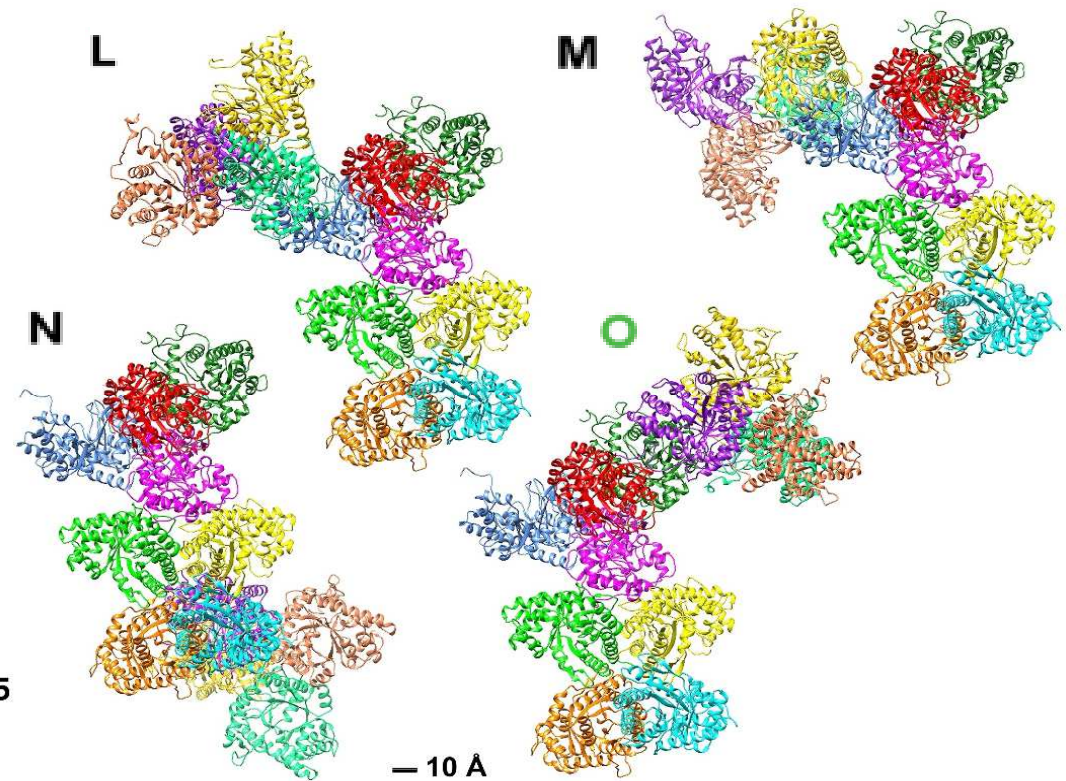
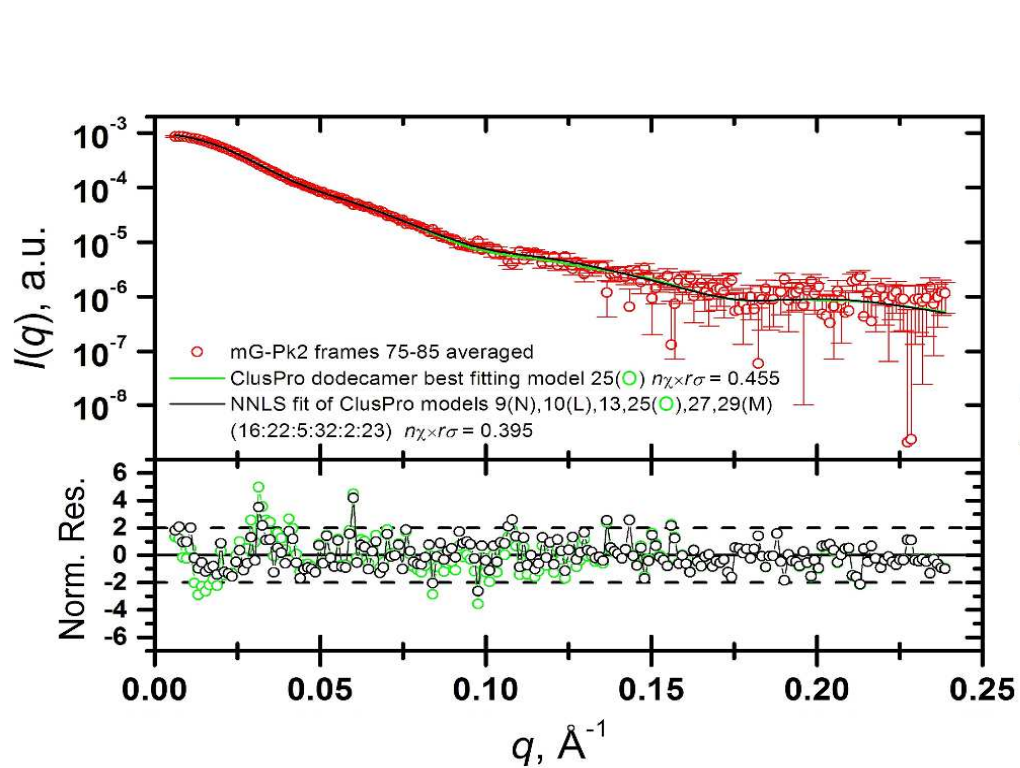
# Deconvoluted aldolase HPLC-SAXS data



# Deconvoluted aldolase HPLC-SAXS data



# Deconvoluted aldolase HPLC-SAXS data

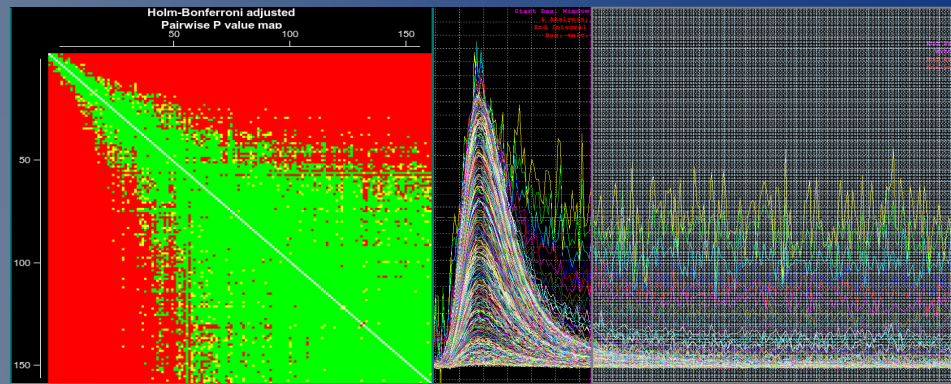




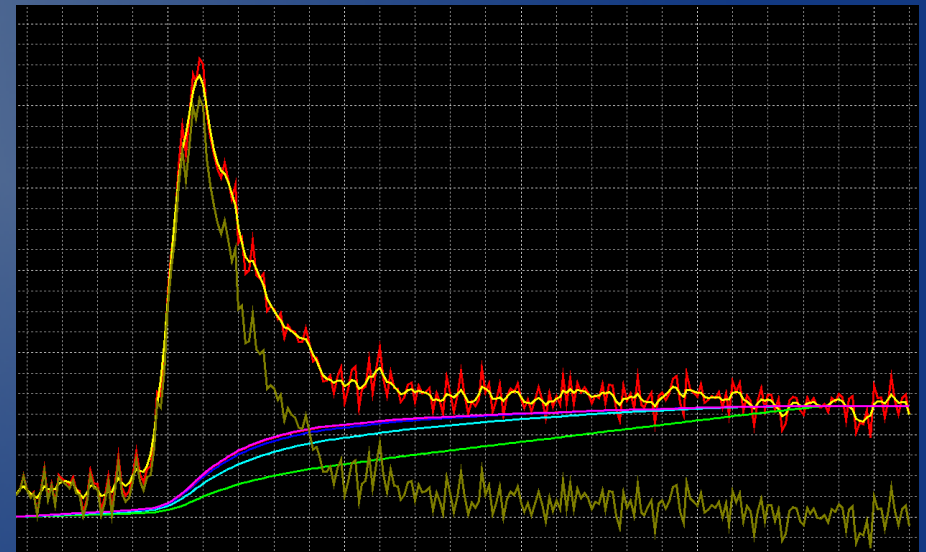
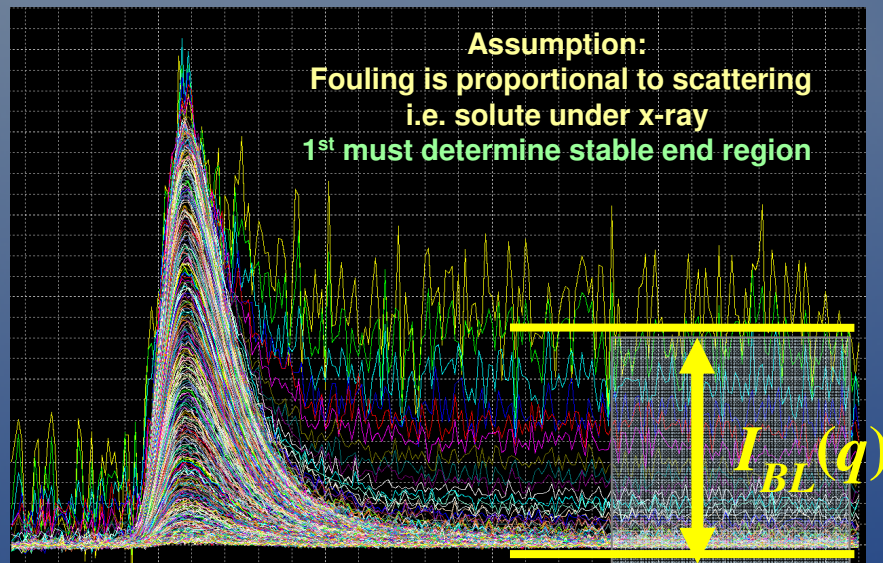
# Other SEC-SAXS tools

*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



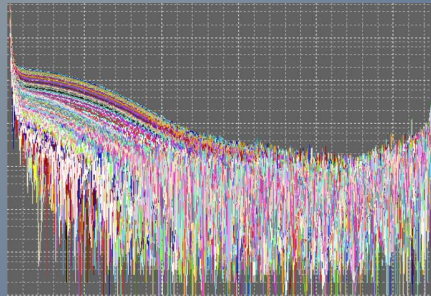
Integral baseline correction



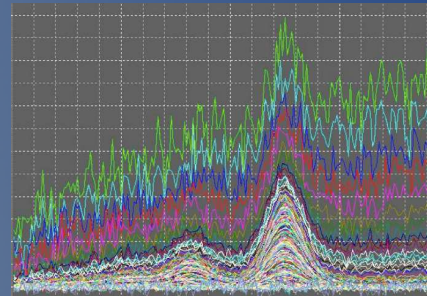
# US-SOMO: HPLC/SEC-SAXS

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

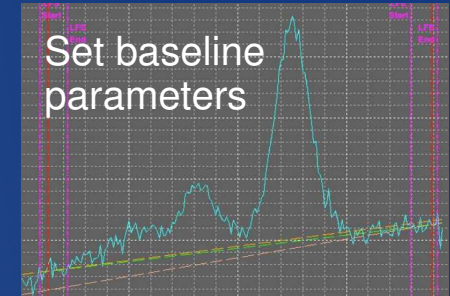
Collect  
HPLC-  
SAXS  
data



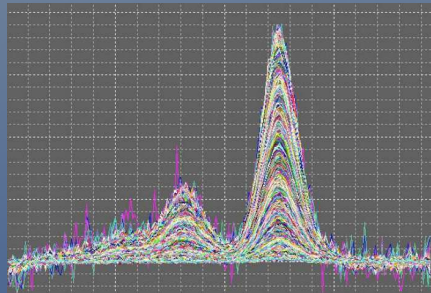
Make  $I(t)$



Select  
typical  
curve



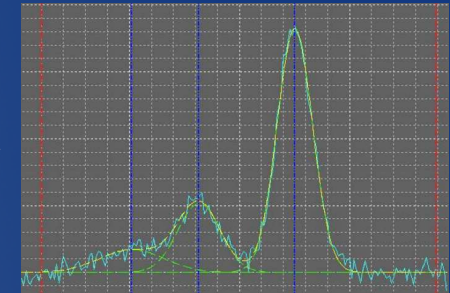
Apply b.l.  
params to  
all curves



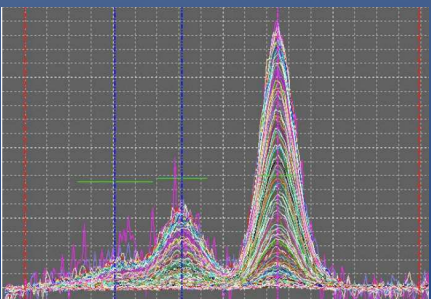
Select  
typical  
curve



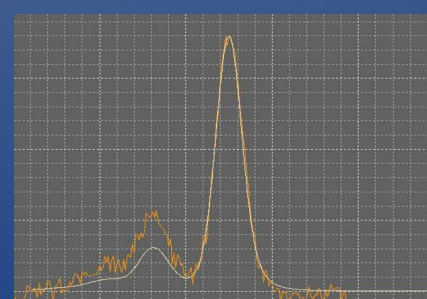
Gauss  
Fit



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



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



**EPSRC**  
Engineering and Physical Sciences  
Research Council