Primary data inspection and reduction with **PRIMUS**: guided tutorial

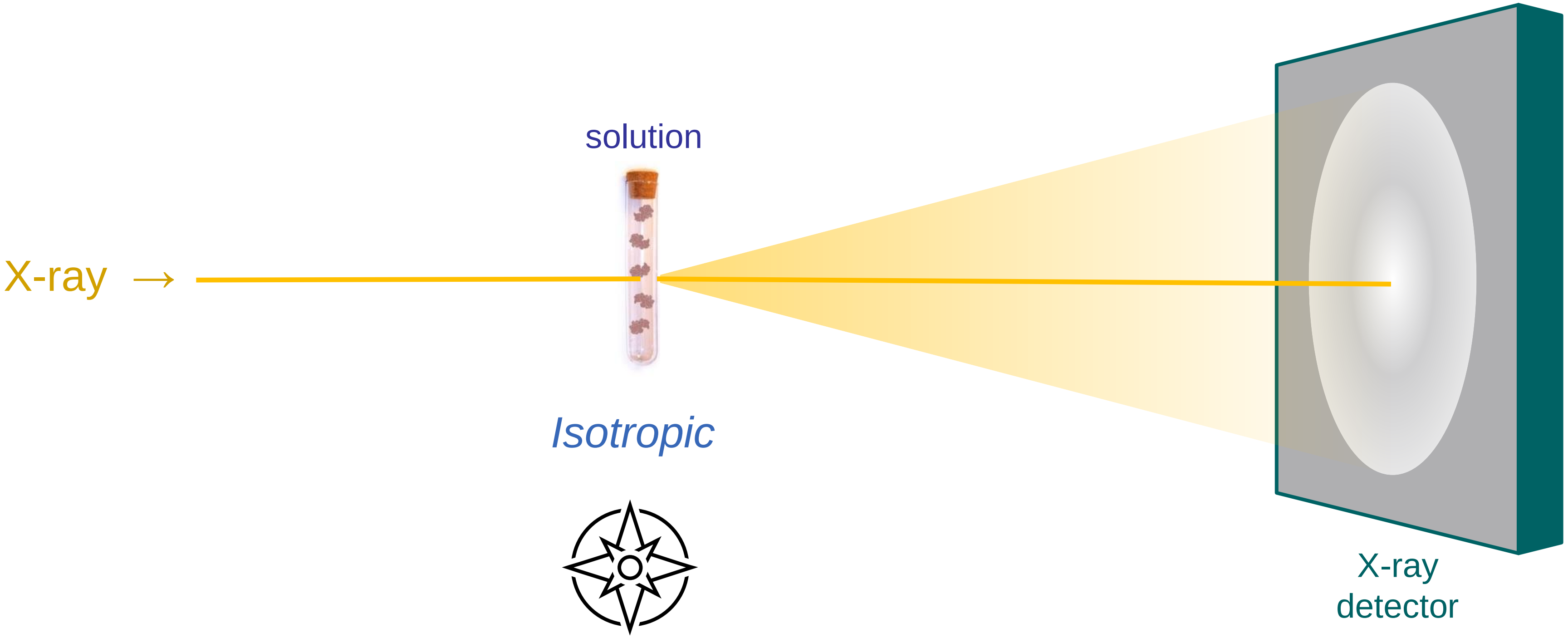
Stefano Da Vela, Melissa Gräwert, Al Kikhney, Cy Jeffries

BIOSAXS

The original solution

First, a brief summary...

SAXS data collection

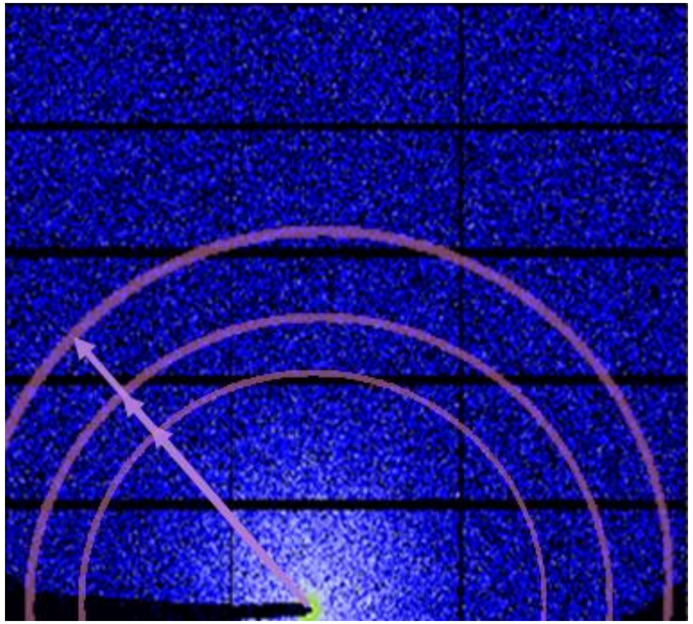


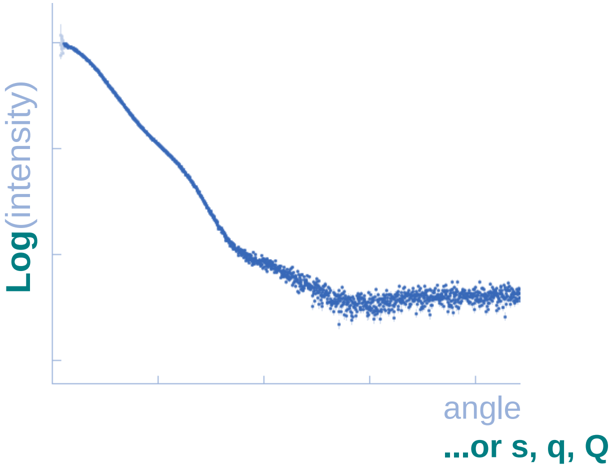
SANS: conceptually similar

SAXS data collection

Also Isotropic

SAXS data





SAXS data
(**background** subtracted)

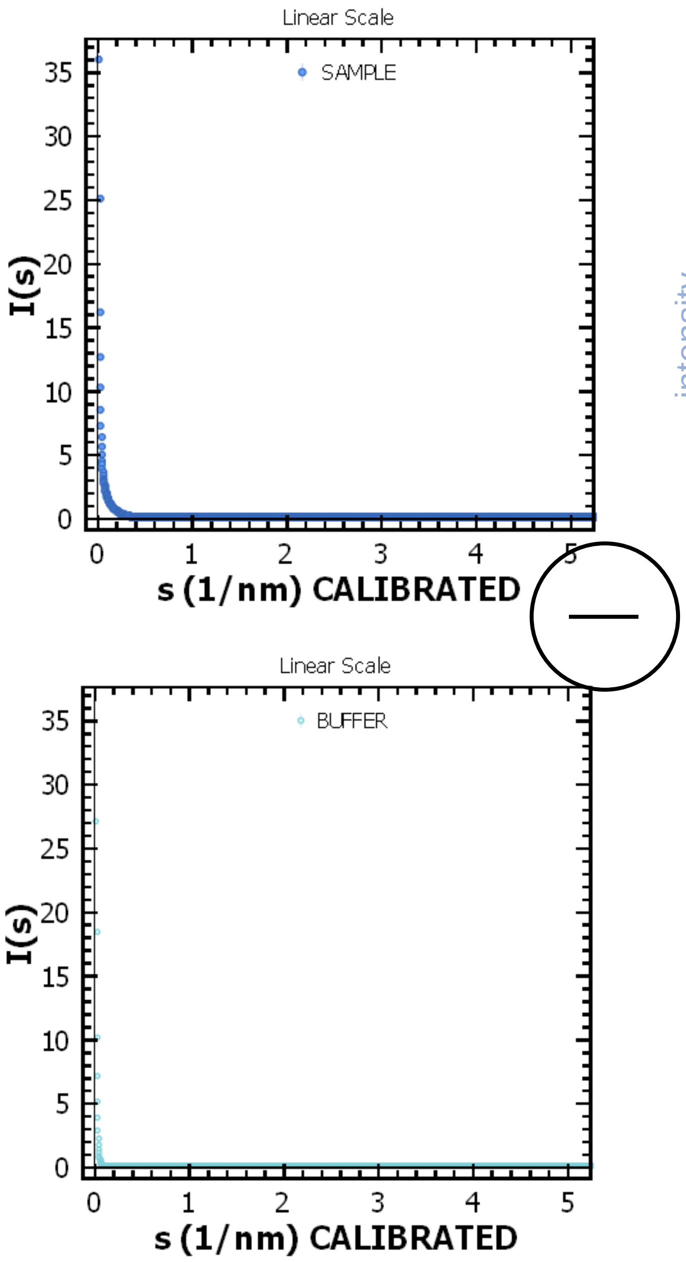
$$|s| = \frac{4\pi sin\vartheta}{\lambda}$$

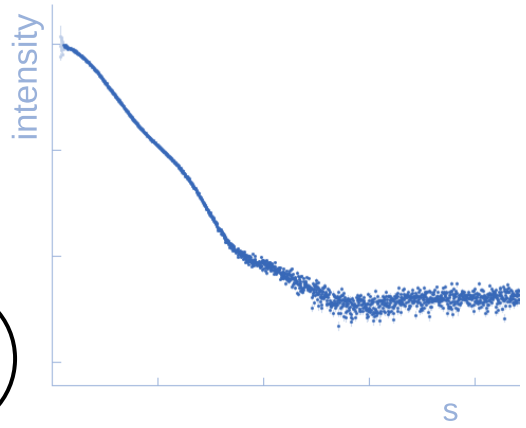
Units: reciprocal of a length
(commonly nm^{-1} or $Å^{-1} = 0.1 * nm^{-1}$)

1D SAXS
curve/profile

nm scale information

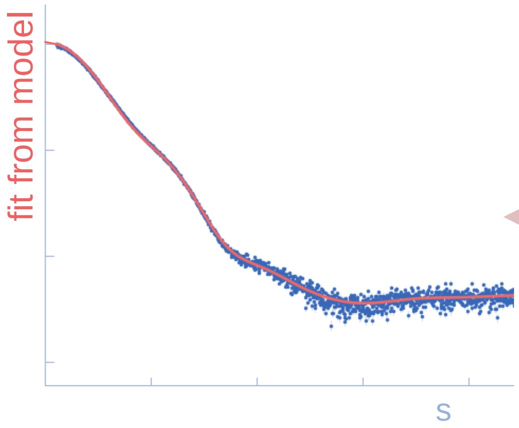
SAXS modelling

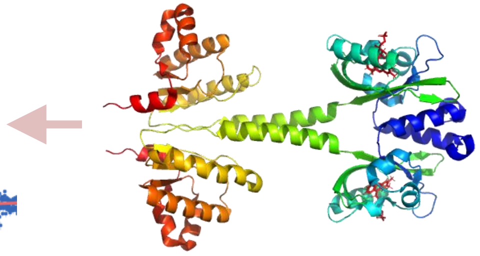




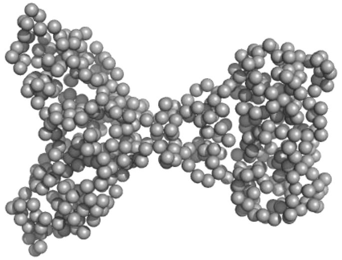
SAXS data

Size & shape:
R_g, D_{max}, p(r), MWs….





Models



Ab initio models

ATSAS software package

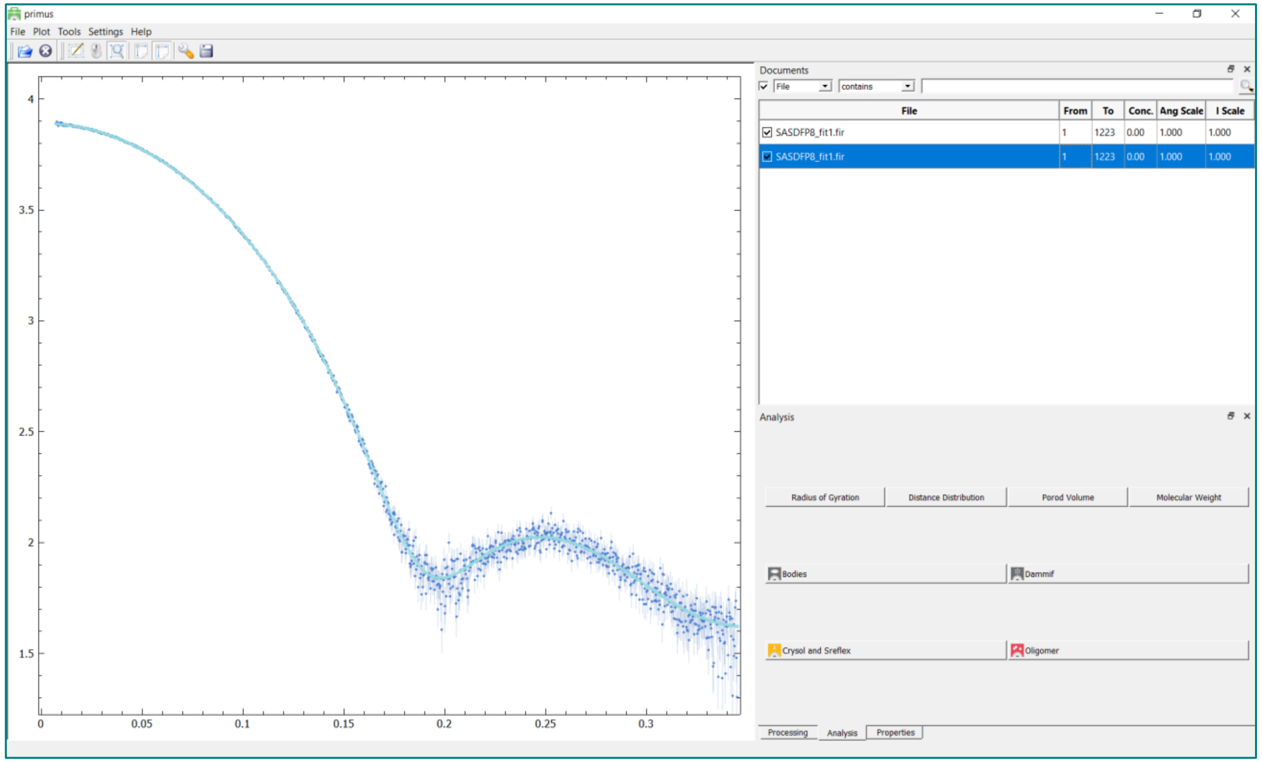
- >100 programs:
  - GUI
  - command-line
  - PyMOL plugin
- ATSAS online

- Commercial licenses: biosaxs.com/software
- Academic licenses: www.embl-hamburg.de/biosaxs/download.html

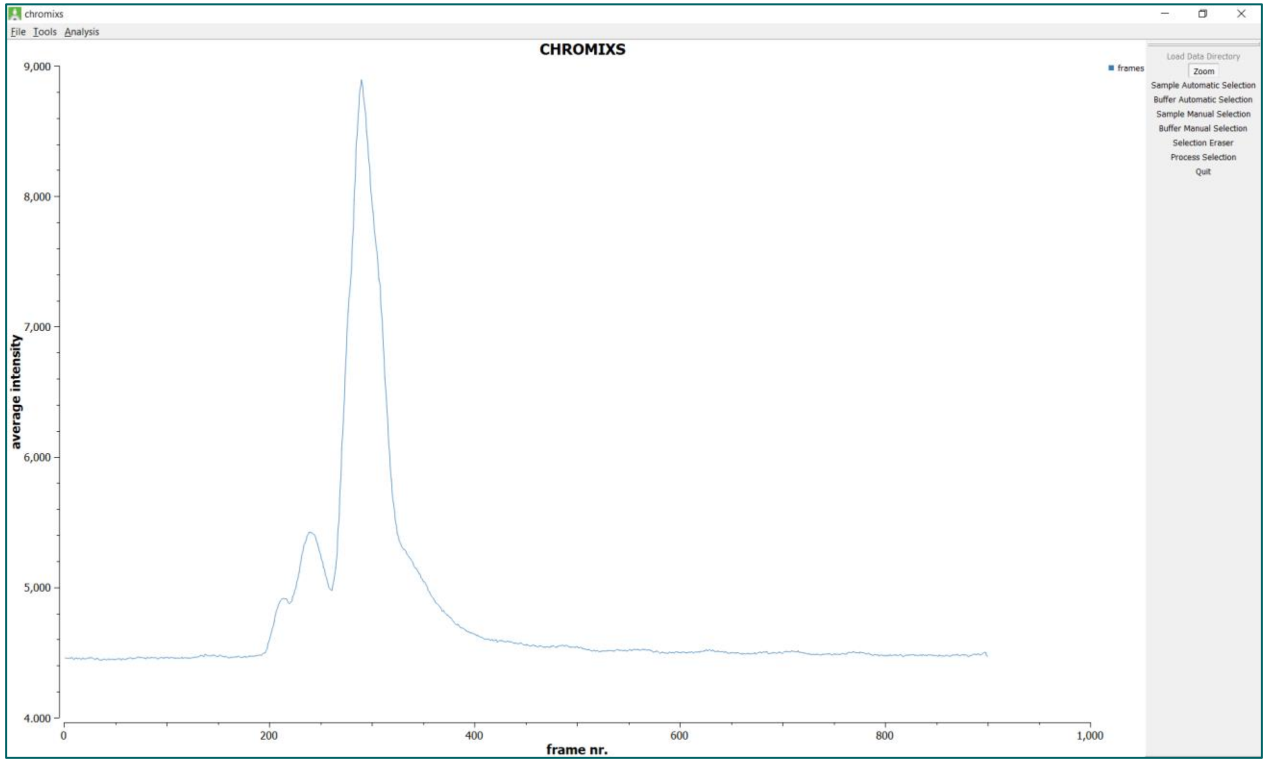
K. Manalastas-Cantos, P.V. Konarev, N.R. Hajizadeh, A.G. Kikhney, M.V. Petoukhov, D.S. Molodenskiy, A. Panjkovich, H.D.T. Mertens, A. Gruzinov, C. Borges, C.M. Jeffries, D.I. Svergun and D. Franke (2021)
ATSAS 3.0: Expanded functionality and new tools for small-angle scattering data analysis
J. Appl. Cryst. 54, 343-355

Focus: Primus, Chromixs

- PRIMUS: plotting SAXS curves and fits, elementary processing, overall parameters, modeling calling other programs of the ATSAS suite
- CHROMIXS: SEC-SAXS data reduction, inspection of chromatograms, overall parameters across SEC peak(s)



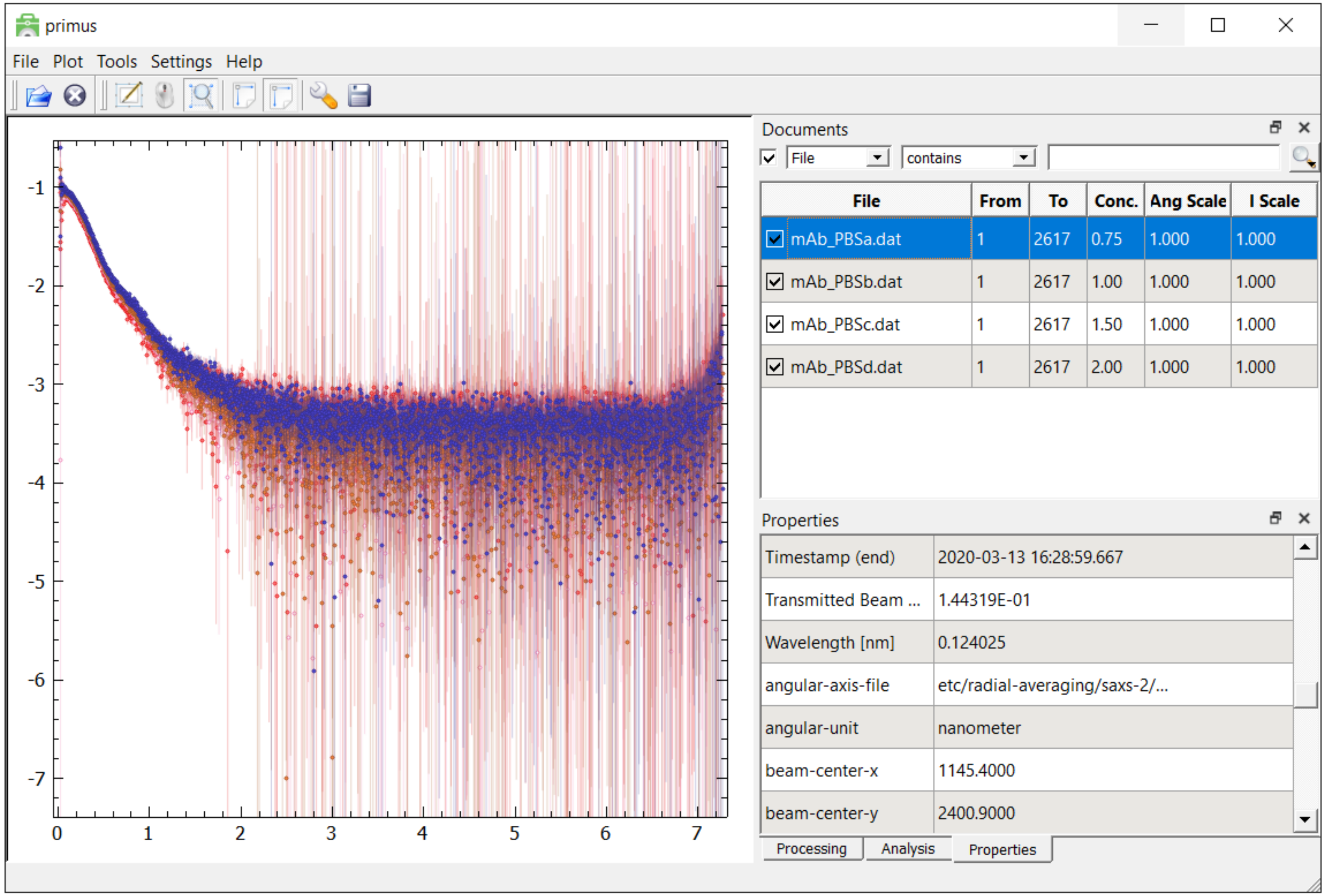
File	From	To	Conc.	Ang Scale	I Scale
☑ SASDFP8_fit1.fir	1	1223	0.00	1.000	1.000
☑ SASDFP8_fit1.fir	1	1223	0.00	1.000	1.000



PRIMUS: basics

PRIMUS: plotting the curves

- A concentration series



File	From	To	Conc.	Ang Scale	I Scale
☑ mAb_PBSa.dat	1	2617	0.75	1.000	1.000
☑ mAb_PBSb.dat	1	2617	1.00	1.000	1.000
☑ mAb_PBSc.dat	1	2617	1.50	1.000	1.000
☑ mAb_PBSd.dat	1	2617	2.00	1.000	1.000

Timestamp (end)	2020-03-13 16:28:59.667
Transmitted Beam ...	1.44319E-01
Wavelength [nm]	0.124025
angular-axis-file	etc/radial-averaging/saxs-2/...
angular-unit	nanometer
beam-center-x	1145.4000
beam-center-y	2400.9000

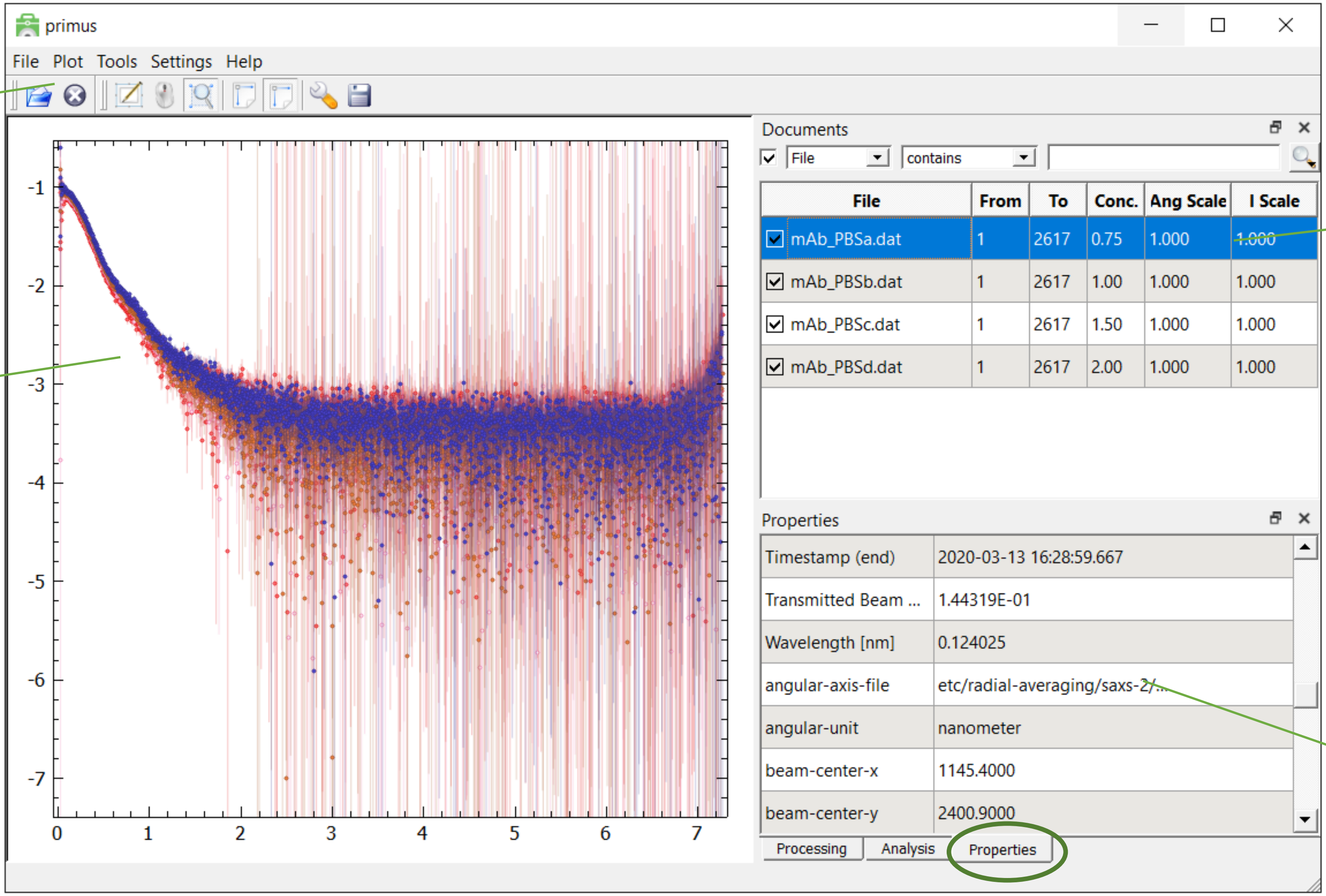
PRIMUS: plotting the curves

- A concentration series

Toolbar

- *Open/close*
- *Select range*
- *(auto)Zoom*
- *Plot settings, export image*

Log-lin plots (default) +error bars



File	From	To	Conc.	Ang Scale	I Scale
☑ mAb_PBSa.dat	1	2617	0.75	1.000	1.000
☑ mAb_PBSb.dat	1	2617	1.00	1.000	1.000
☑ mAb_PBSc.dat	1	2617	1.50	1.000	1.000
☑ mAb_PBSd.dat	1	2617	2.00	1.000	1.000

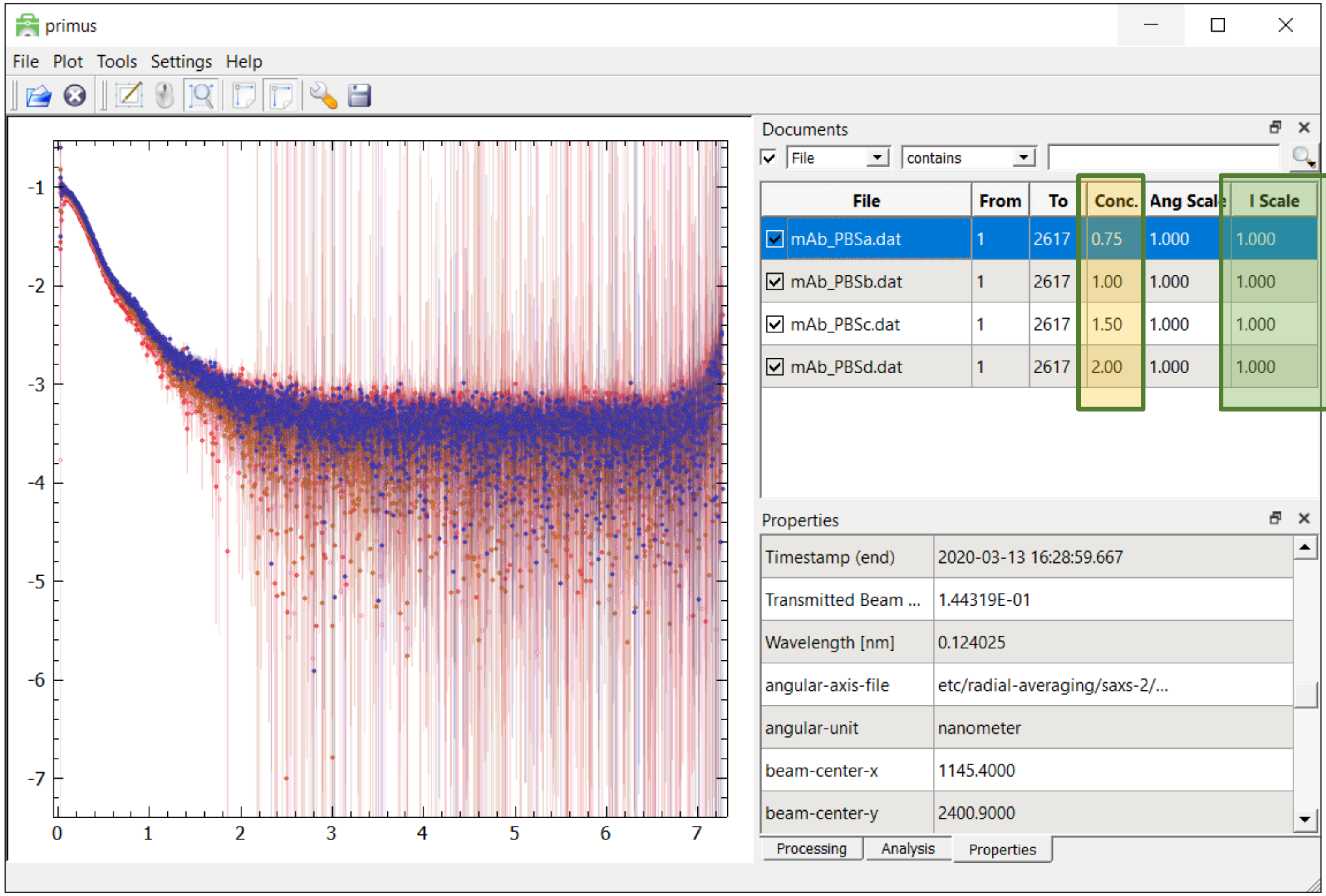
Timestamp (end)	2020-03-13 16:28:59.667
Transmitted Beam ...	1.44319E-01
Wavelength [nm]	0.124025
angular-axis-file	etc/radial-averaging/saxs-2/...
angular-unit	nanometer
beam-center-x	1145.4000
beam-center-y	2400.9000

Loaded data Selected=active

Metadata from footer

PRIMUS: plotting the curves

- A concentration series: I/c *vs* s



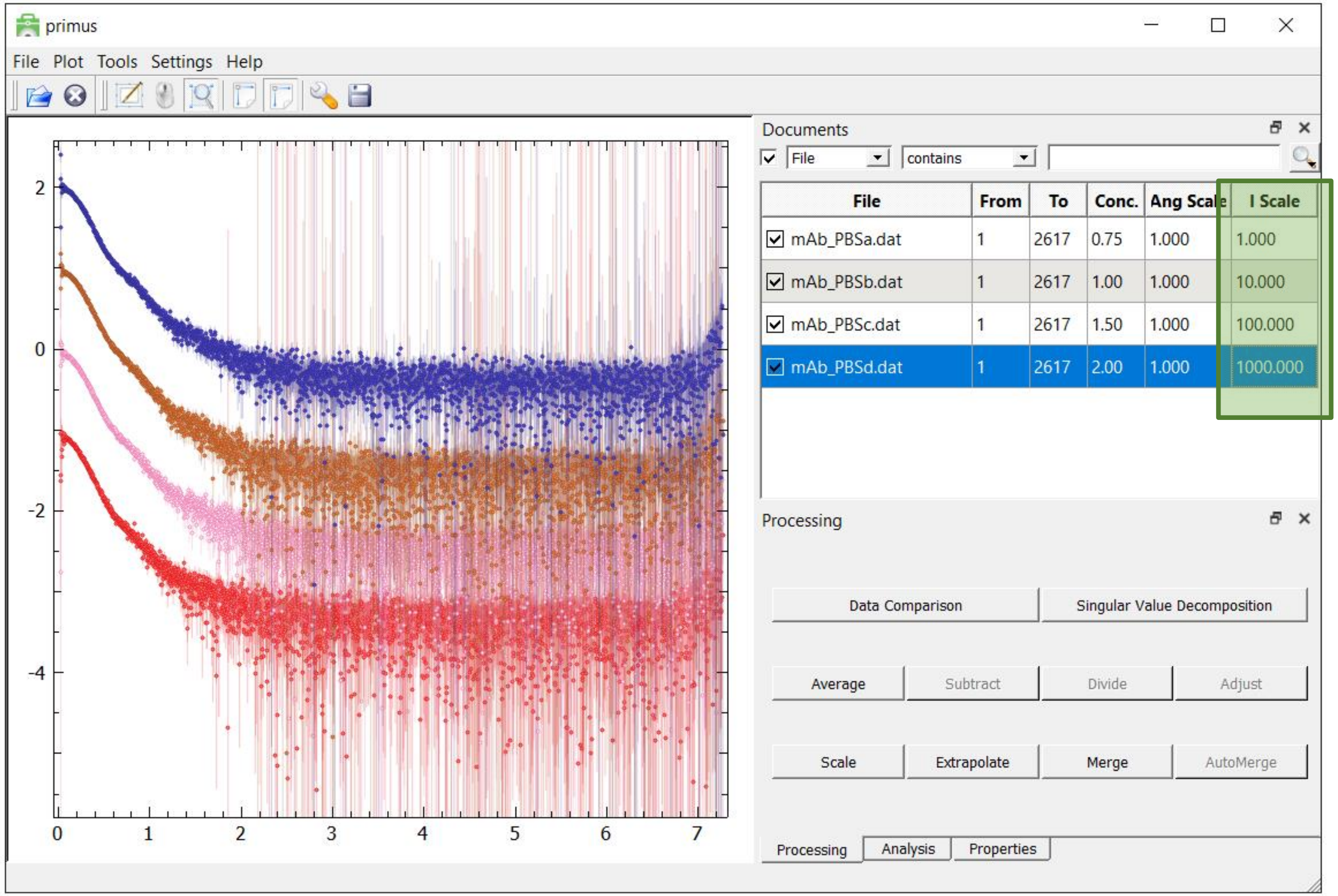
File	From	To	Conc.	Ang Scale	I Scale
☑ mAb_PBSa.dat	1	2617	0.75	1.000	1.000
☑ mAb_PBSb.dat	1	2617	1.00	1.000	1.000
☑ mAb_PBSc.dat	1	2617	1.50	1.000	1.000
☑ mAb_PBSd.dat	1	2617	2.00	1.000	1.000

Timestamp (end)	2020-03-13 16:28:59.667
Transmitted Beam ...	1.44319E-01
Wavelength [nm]	0.124025
angular-axis-file	etc/radial-averaging/saxs-2/...
angular-unit	nanometer
beam-center-x	1145.4000
beam-center-y	2400.9000

I Scale x (I/c)

PRIMUS: plotting the curves

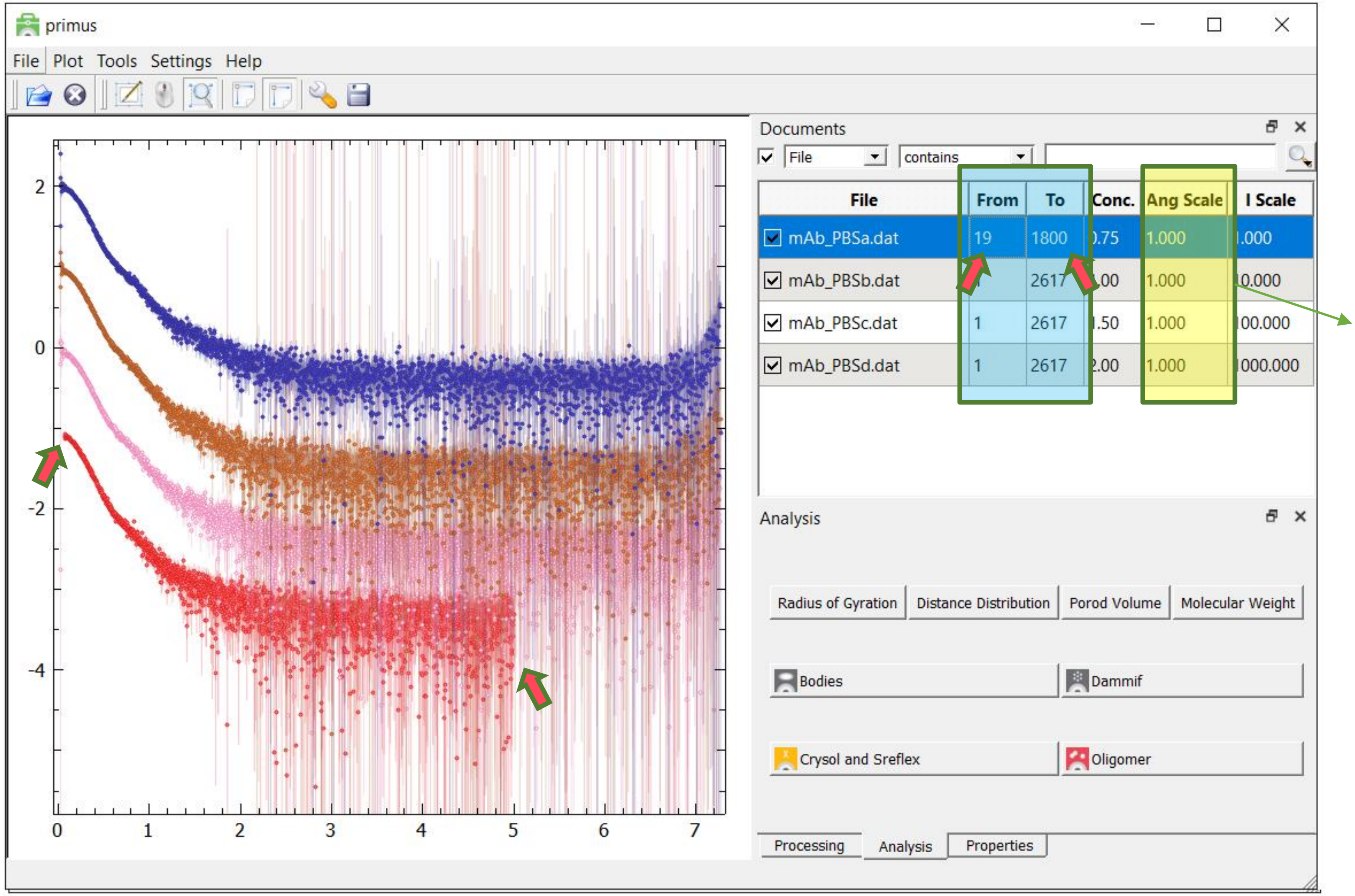
- A concentration series: I/c *vs* s *shifted for display*



File	From	To	Conc.	Ang Scale	I Scale
☑ mAb_PBSa.dat	1	2617	0.75	1.000	1.000
☑ mAb_PBSb.dat	1	2617	1.00	1.000	10.000
☑ mAb_PBSc.dat	1	2617	1.50	1.000	100.000
☑ mAb_PBSd.dat	1	2617	2.00	1.000	1000.000

PRIMUS: plotting the curves

- A concentration series: range and angular scale

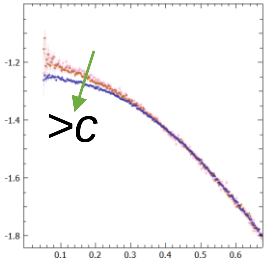


File	From	To	Conc.	Ang Scale	I Scale
☑ mAb_PBSa.dat	19	1800	.75	1.000	.000
☑ mAb_PBSb.dat	1	2617	.00	1.000	0.000
☑ mAb_PBSc.dat	1	2617	.50	1.000	00.000
☑ mAb_PBSd.dat	1	2617	.00	1.000	000.000

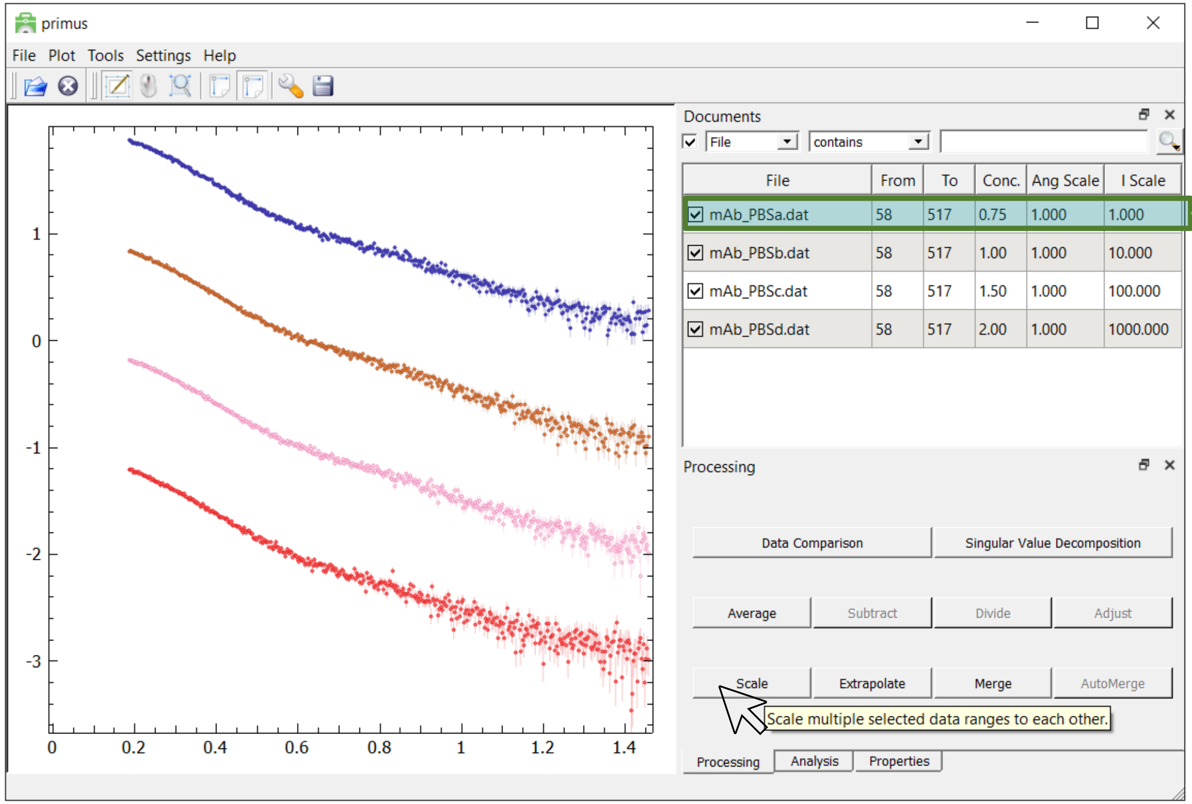
e.g. nm^{-1} to Å^{-1}

PRIMUS: plotting the curves

- A concentration series: scaling, comparing low angles

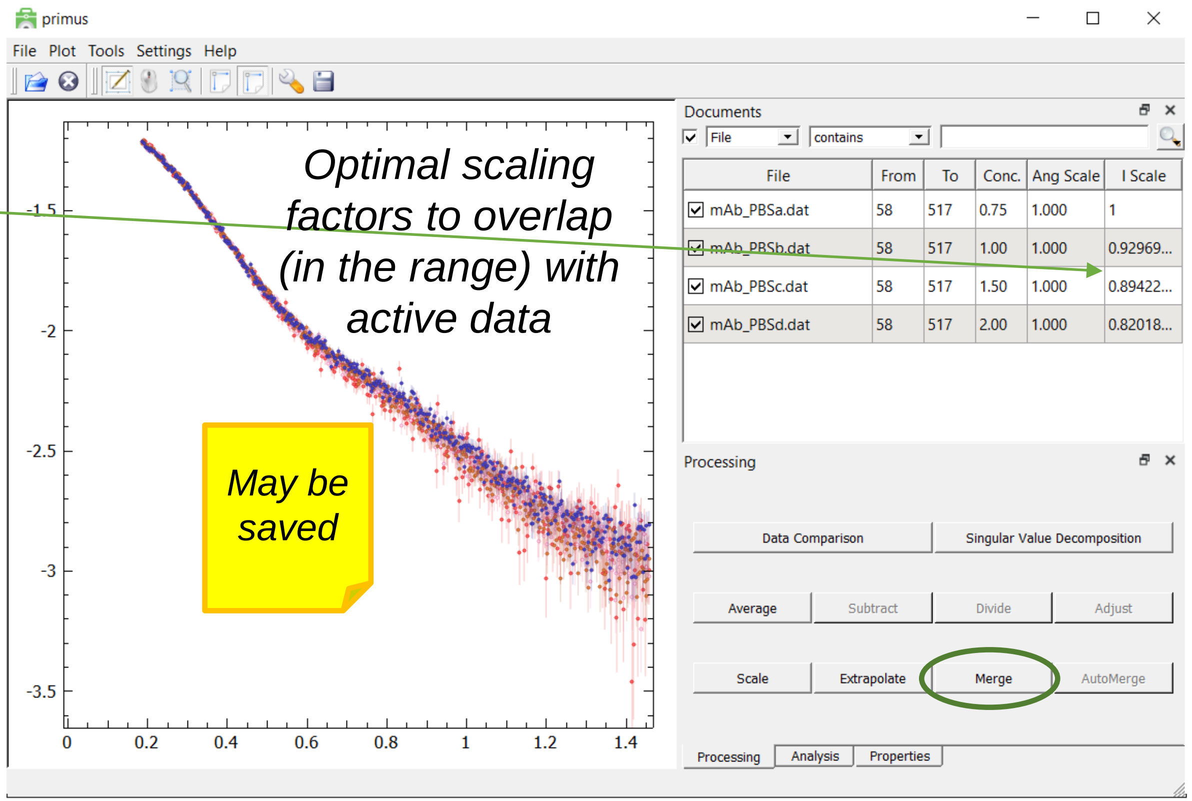


e.g. detection of structure factor or self-association



File	From	To	Conc.	Ang Scale	I Scale
☑ mAb_PBSa.dat	58	517	0.75	1.000	1.000
☑ mAb_PBSb.dat	58	517	1.00	1.000	10.000
☑ mAb_PBSc.dat	58	517	1.50	1.000	100.000
☑ mAb_PBSd.dat	58	517	2.00	1.000	1000.000

Scale multiple selected data ranges to each other.



File	From	To	Conc.	Ang Scale	I Scale
☑ mAb_PBSa.dat	58	517	0.75	1.000	1
☑ mAb_PBSb.dat	58	517	1.00	1.000	0.92969...
☑ mAb_PBSc.dat	58	517	1.50	1.000	0.89422...
☑ mAb_PBSd.dat	58	517	2.00	1.000	0.82018...

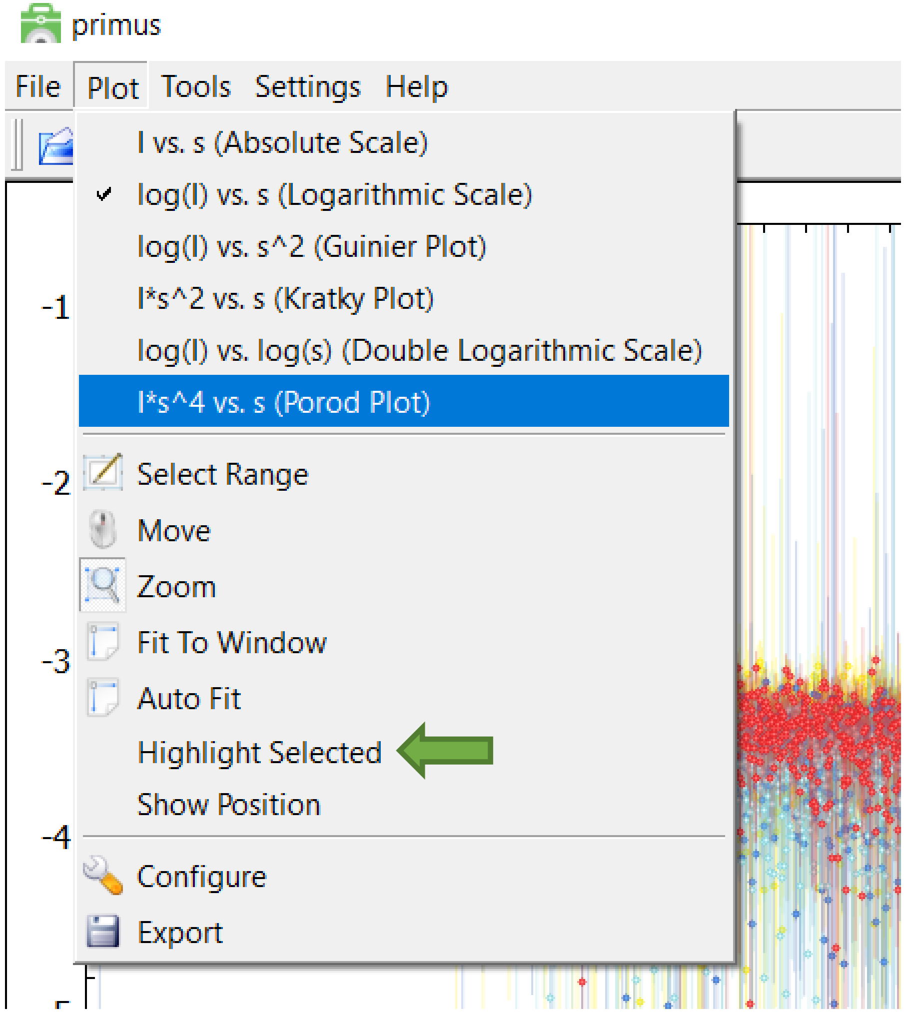
Optimal scaling factors to overlap (in the range) with active data

May be saved

Rule-of-thumb: choose highest concentration curve (higher signal-to-noise) with overlap to all the lowest concentration ones for parameters evaluation and modeling. Or: merge/extrapolate.

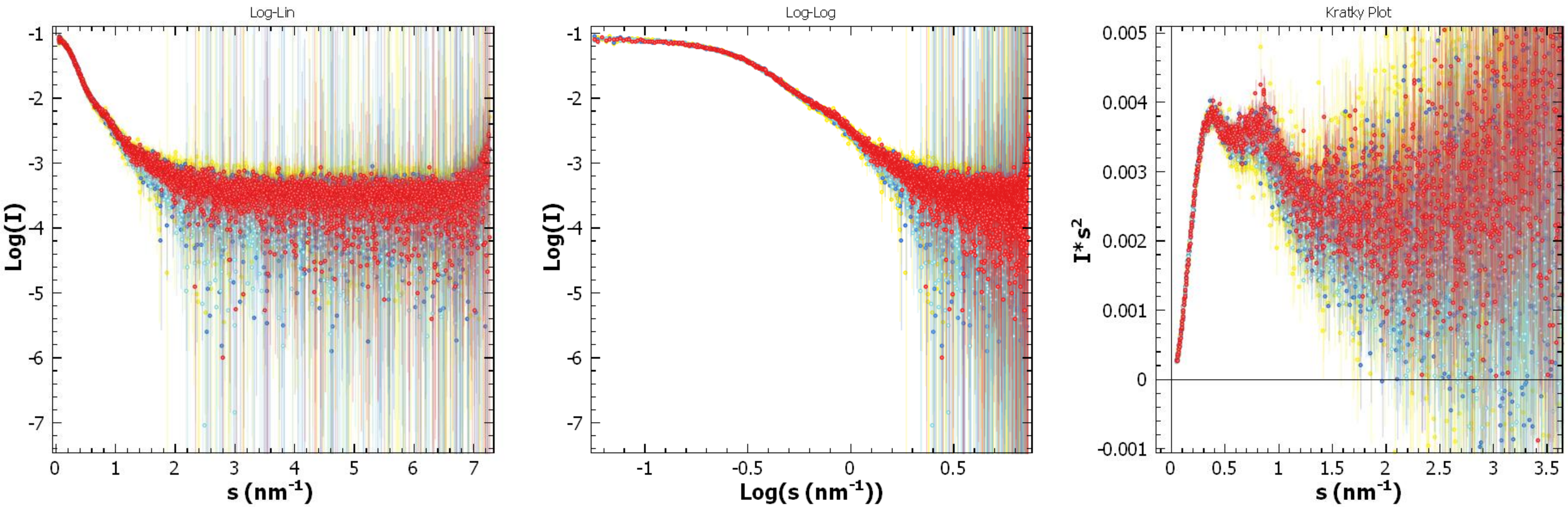
PRIMUS: plotting the curves

- Alternatives for plotting



PRIMUS: plotting the curves

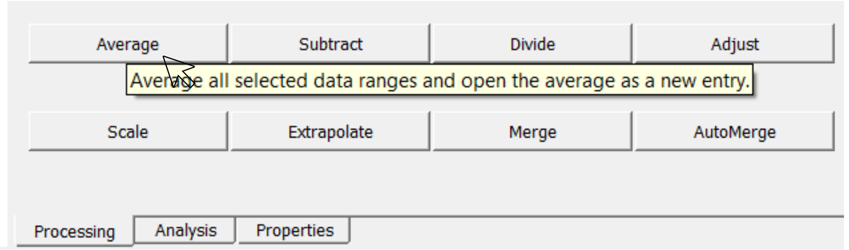
- Alternatives for plotting most commonly used for Bio-SAXS



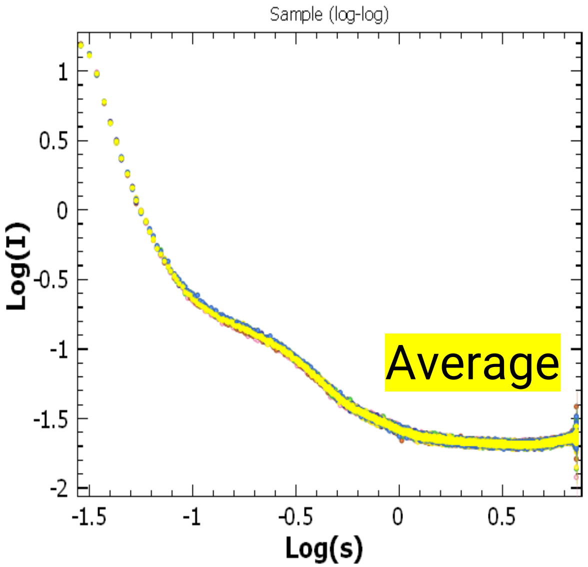
Same data in 3 plotting modes

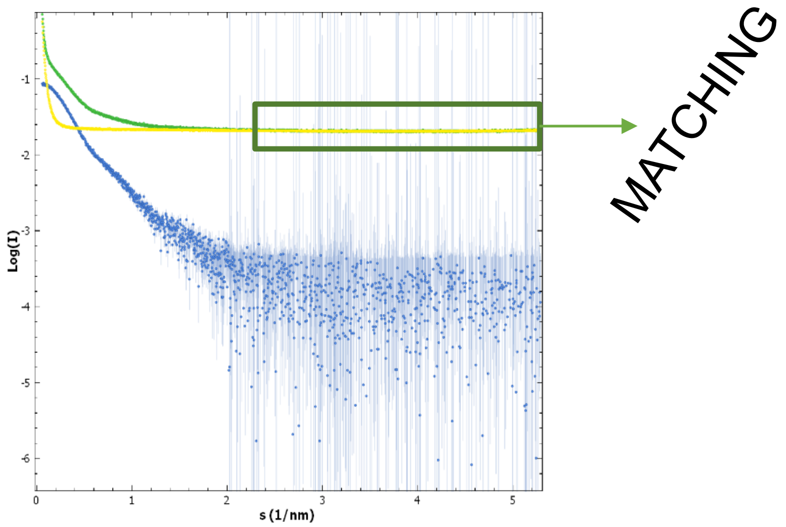
PRIMUS: plotting the curves

- More math with curves:



Average all selected data ranges and open the average as a new entry.



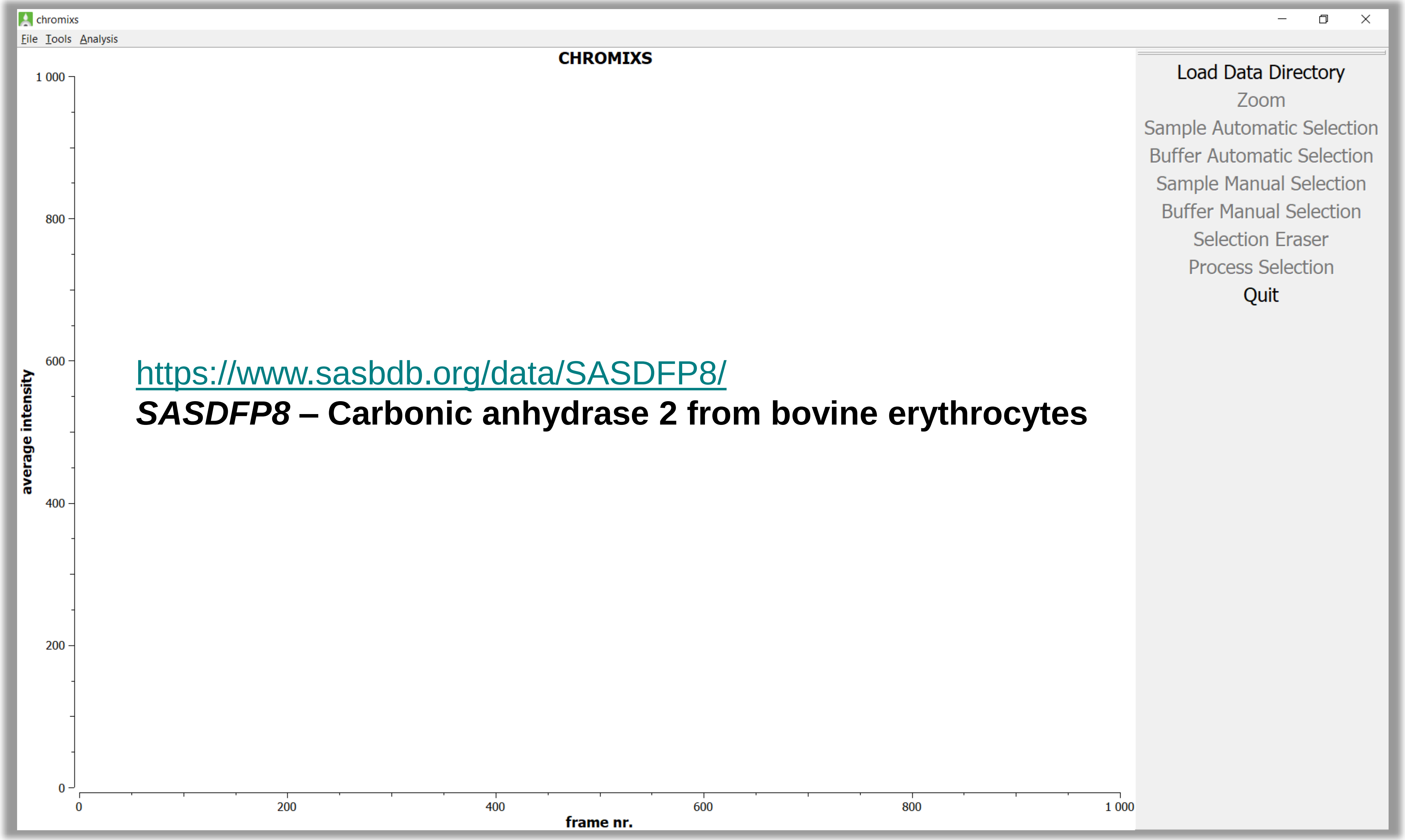


- Averaging exposures and checking for filling/damage
- Subtracting background manually
- Compensating inaccurate concentration normalizations
- Experimental structure factors
- ….

Mostly automated

CHROMIXS: SAXS curves from SEC-SAXS

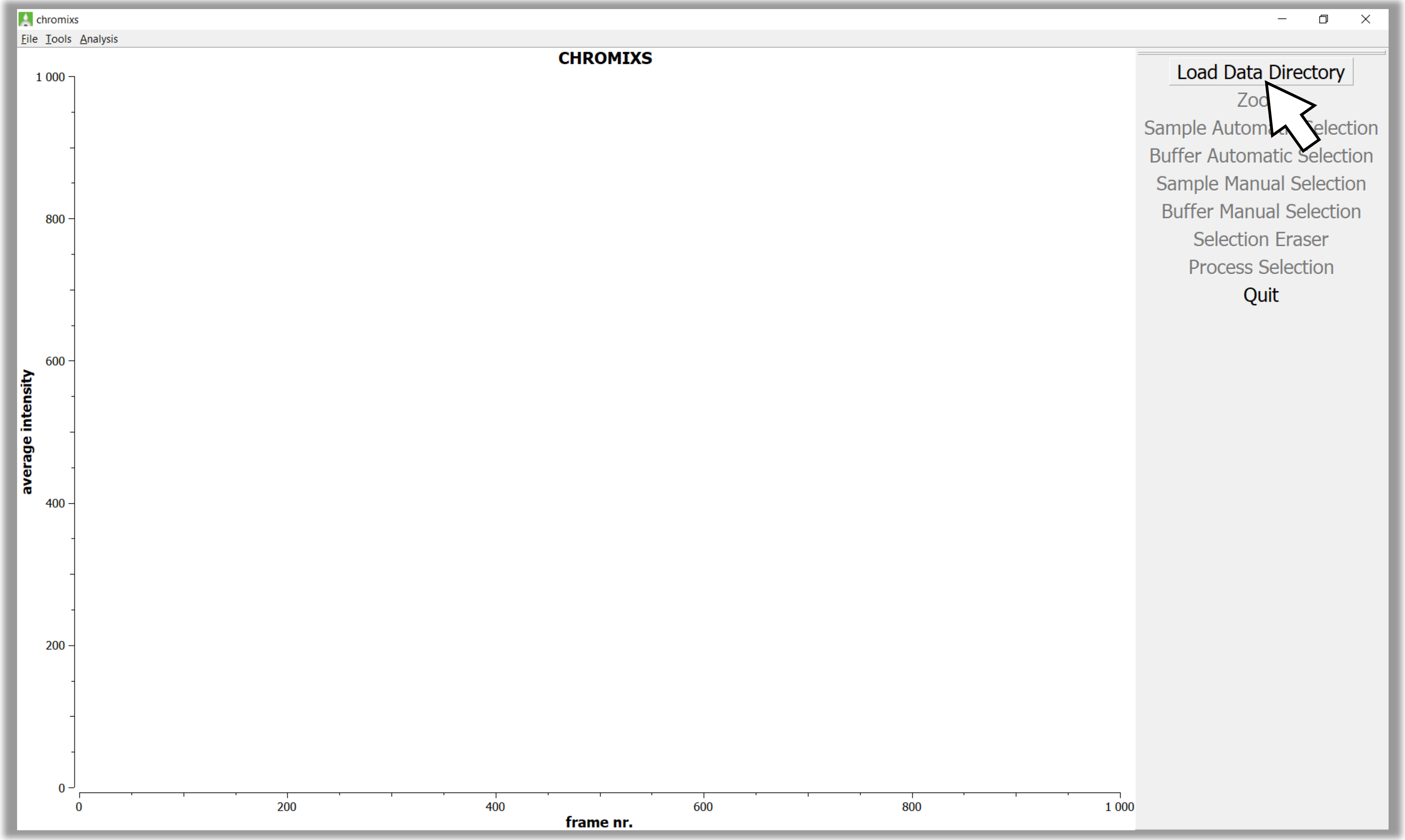
CHROMIXS: process SEC-SAXS data



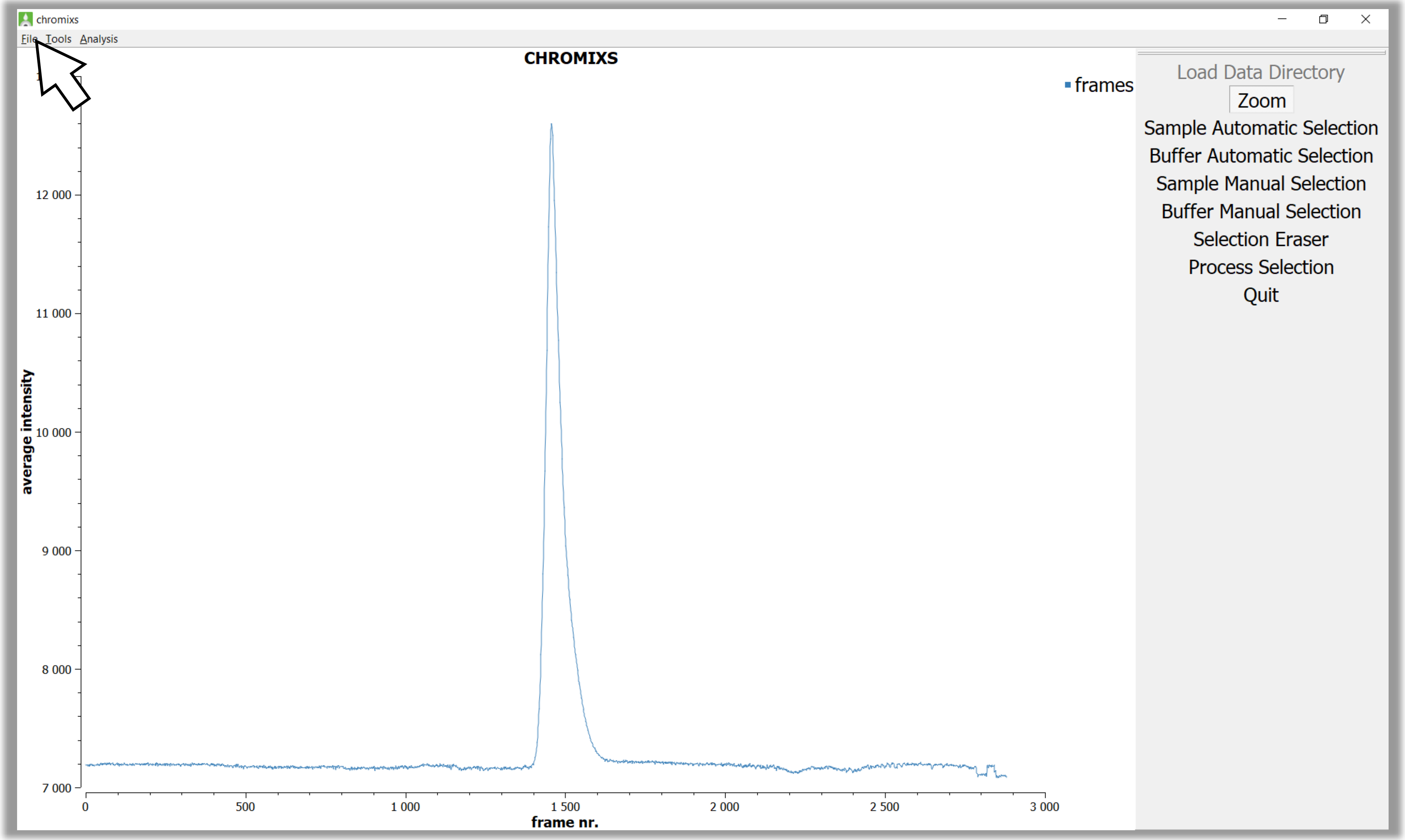
https://www.sasbdb.org/data/SASDFP8/

***SASDFP8* – Carbonic anhydrase 2 from bovine erythrocytes**

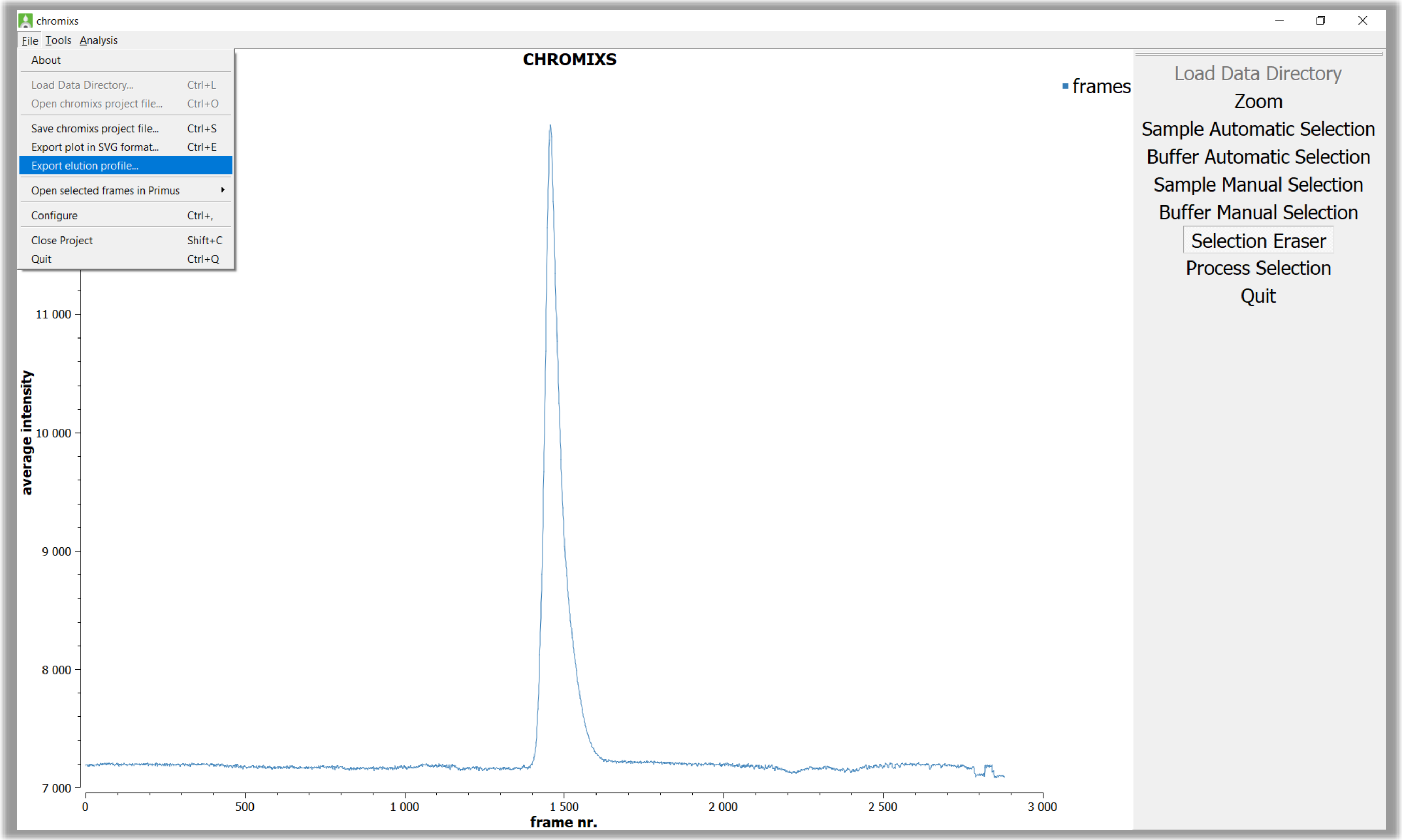
CHROMIXS: process SEC-SAXS data



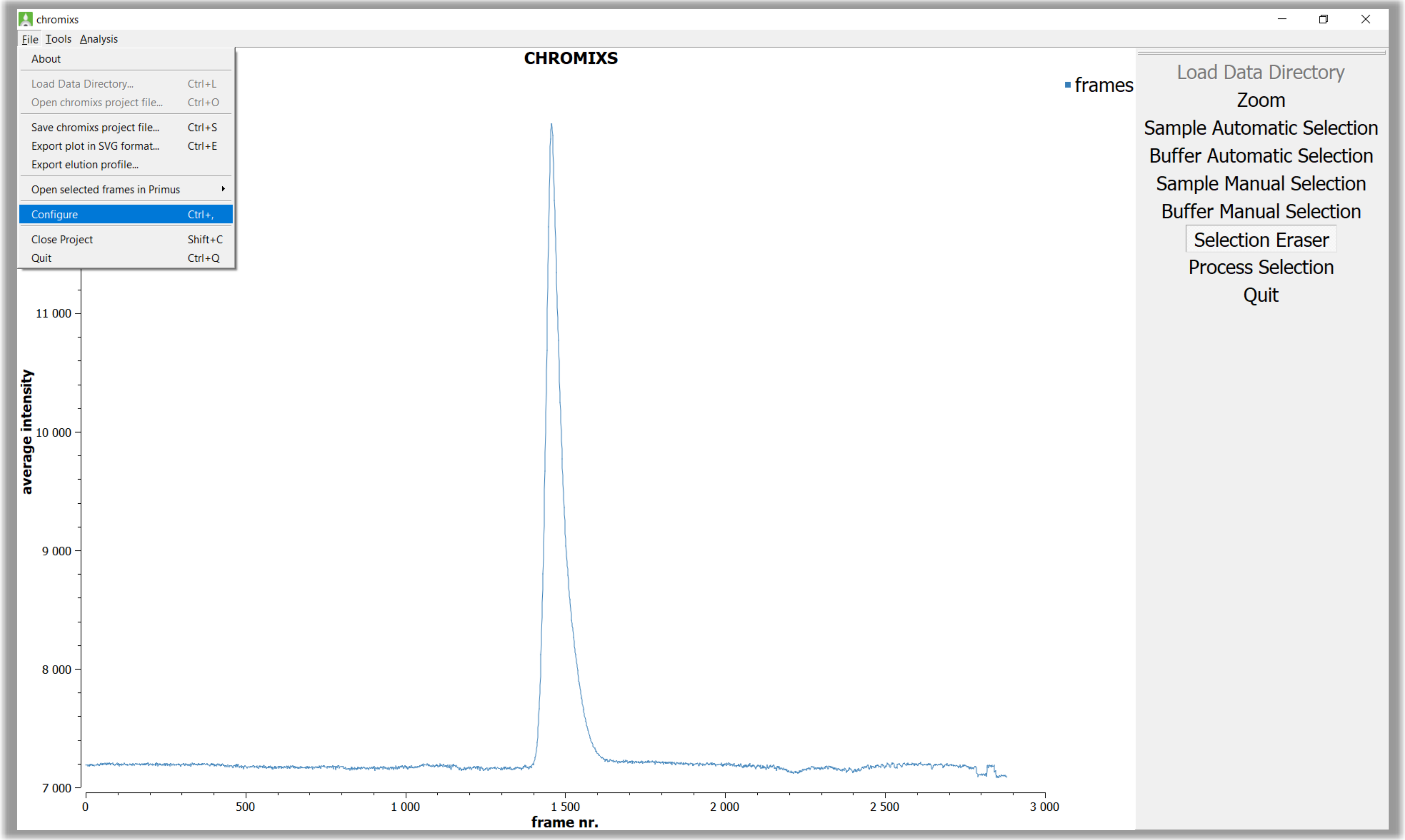
CHROMIXS: process SEC-SAXS data



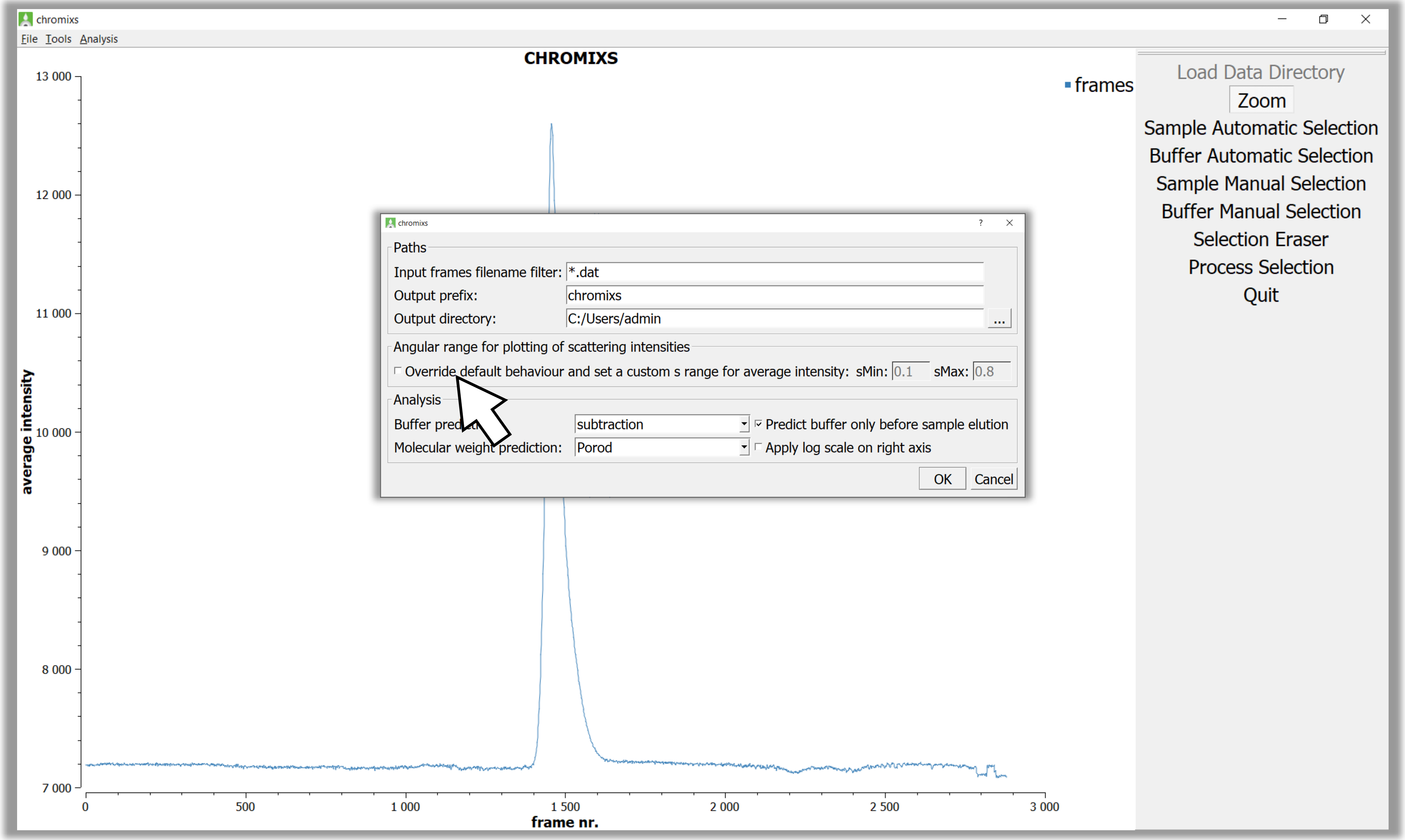
CHROMIXS: process SEC-SAXS data



CHROMIXS: process SEC-SAXS data



CHROMIXS: process SEC-SAXS data



Paths

Input frames filename filter: *.dat

Output prefix: chromixs

Output directory: C:/Users/admin

Angular range for plotting of scattering intensities

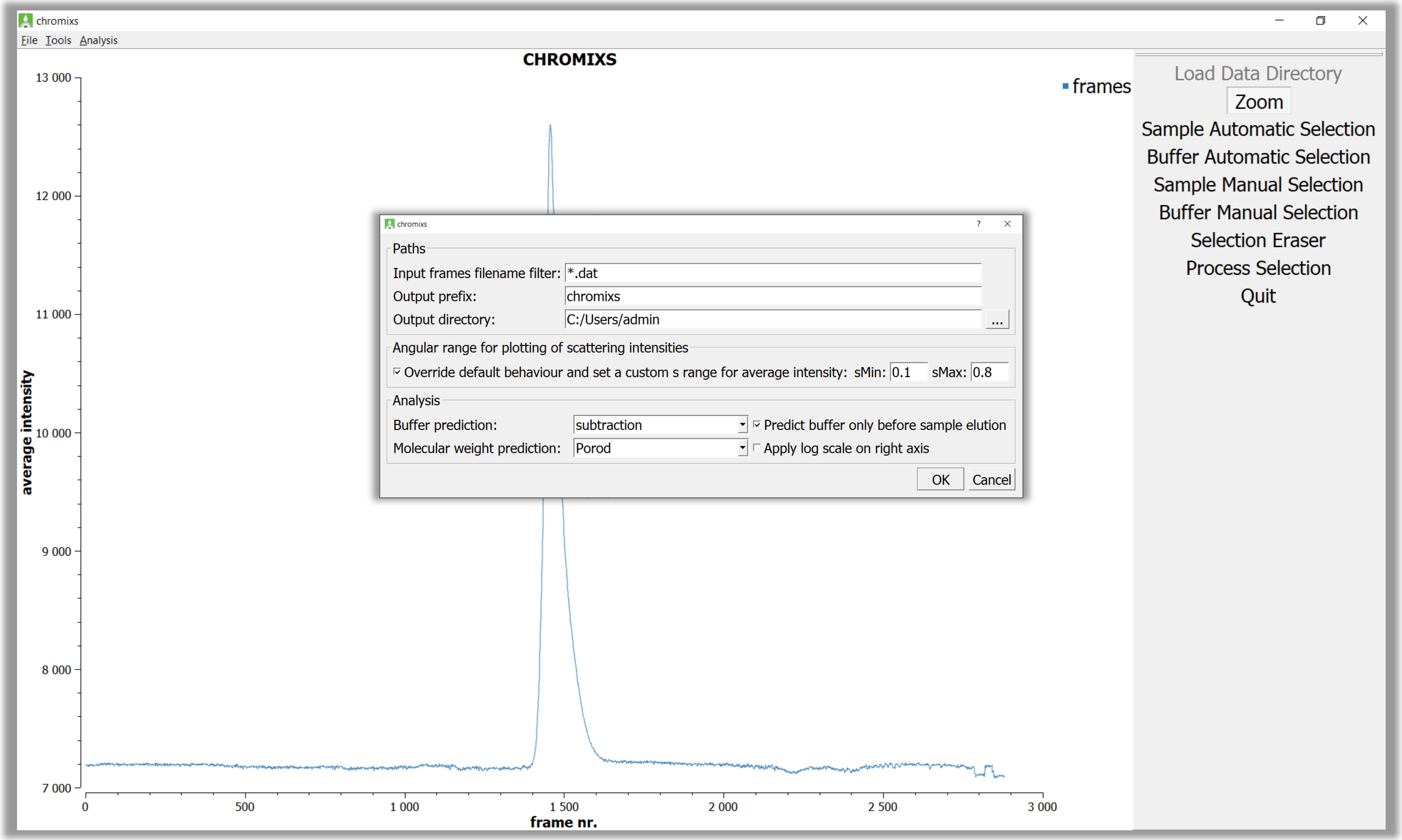
☐ Override default behaviour and set a custom s range for average intensity: sMin: 0.1 sMax: 0.8

Analysis

Buffer pred subtraction ☑ Predict buffer only before sample elution

Molecular weight prediction: Porod ☐ Apply log scale on right axis

CHROMIXS: process SEC-SAXS data



Paths

Input frames filename filter: *.dat

Output prefix: chromixs

Output directory: C:/Users/admin

Angular range for plotting of scattering intensities

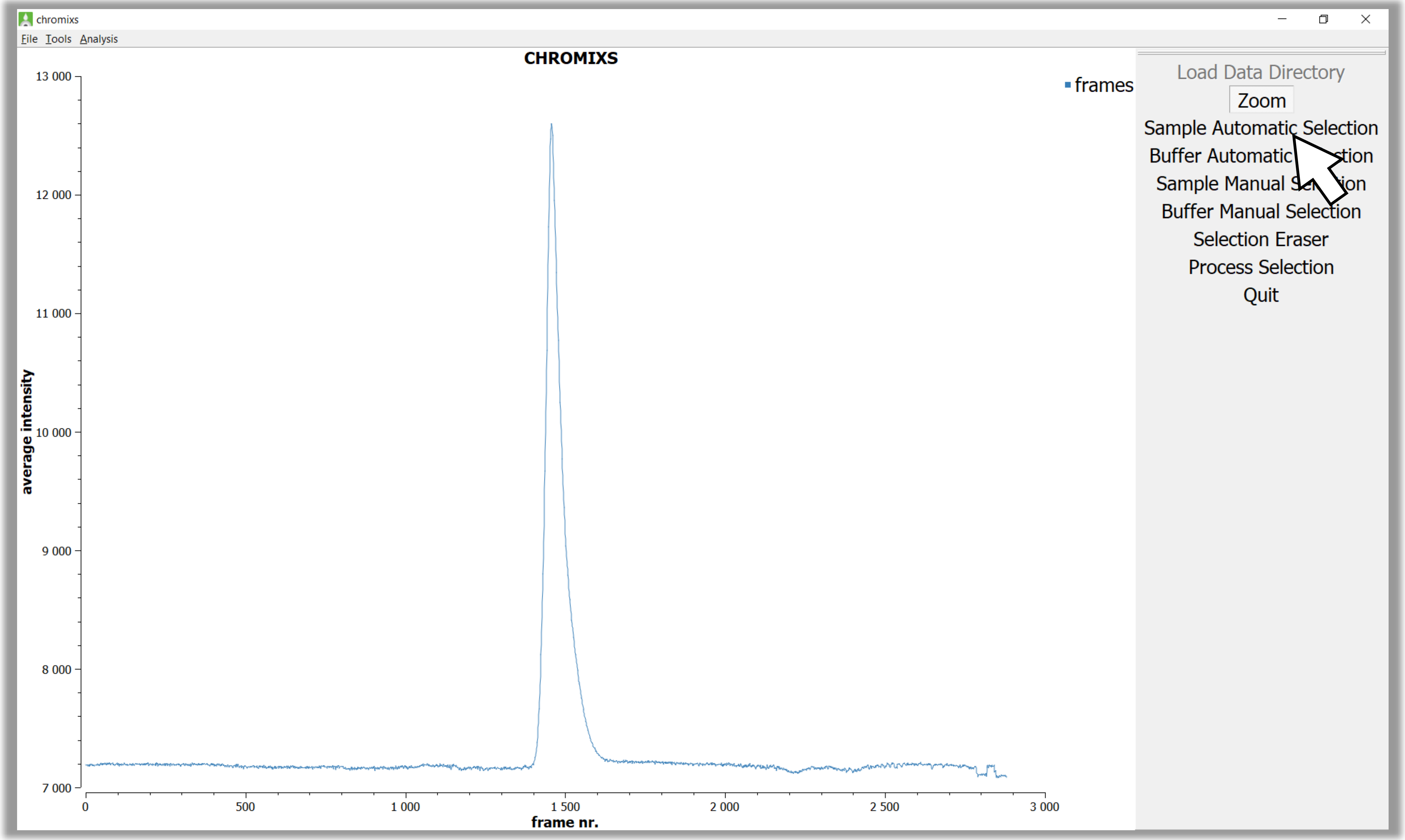
☑ Override default behaviour and set a custom s range for average intensity: sMin: 0.1 sMax: 0.8

Analysis

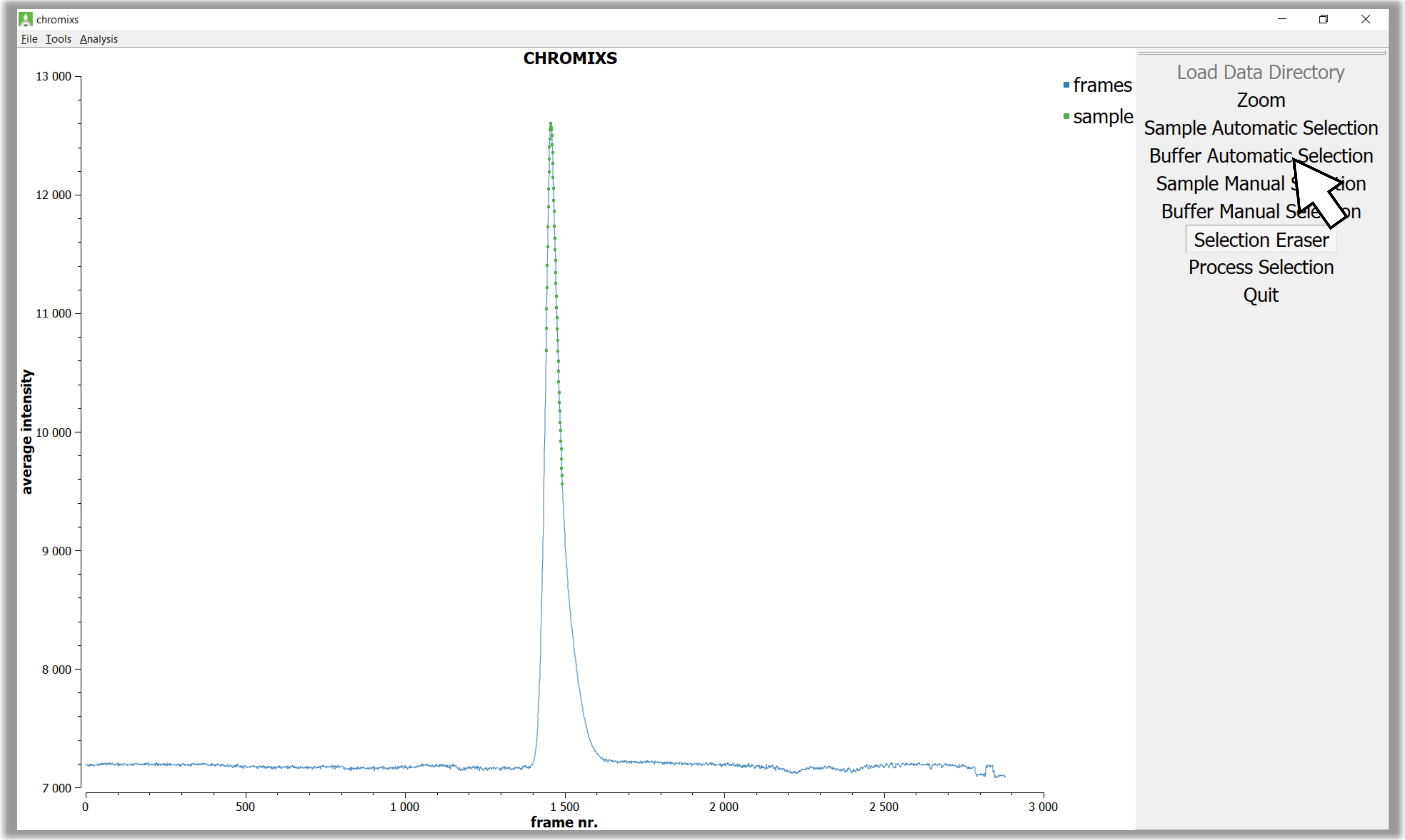
Buffer prediction: subtraction ☑ Predict buffer only before sample elution

Molecular weight prediction: Porod ☐ Apply log scale on right axis

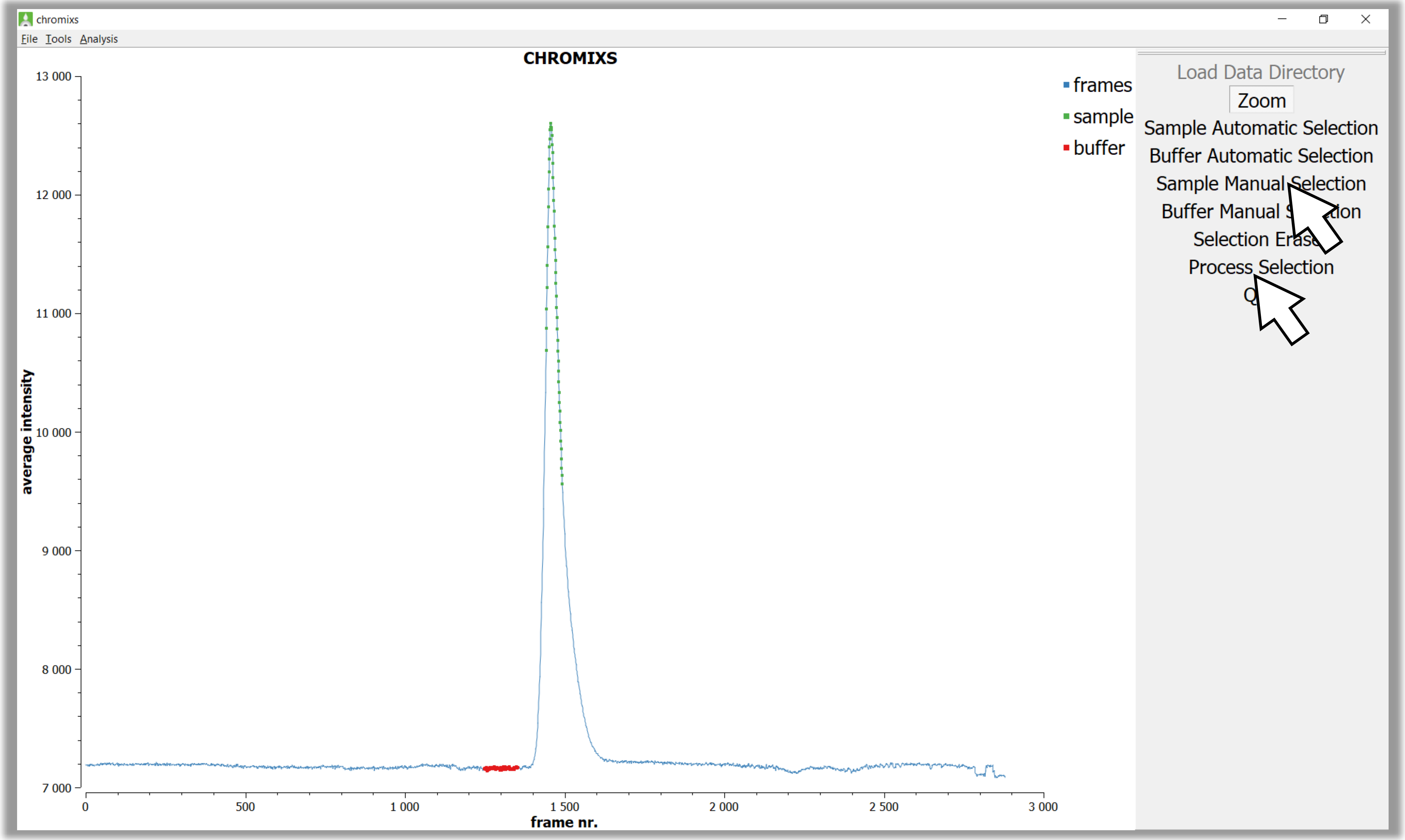
CHROMIXS: process SEC-SAXS data



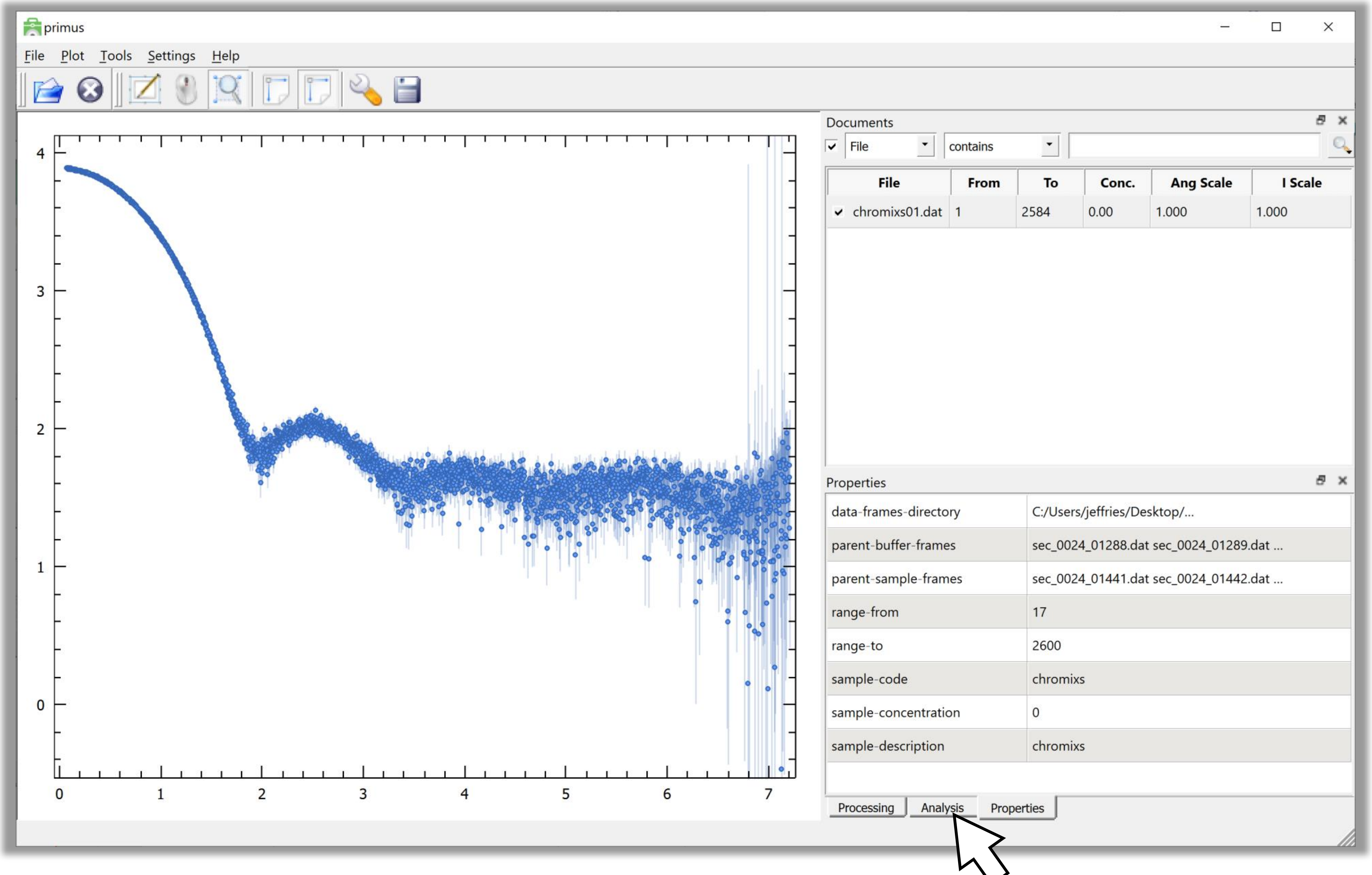
CHROMIXS: process SEC-SAXS data



CHROMIXS: process SEC-SAXS data



CHROMIXS opens PRIMUS

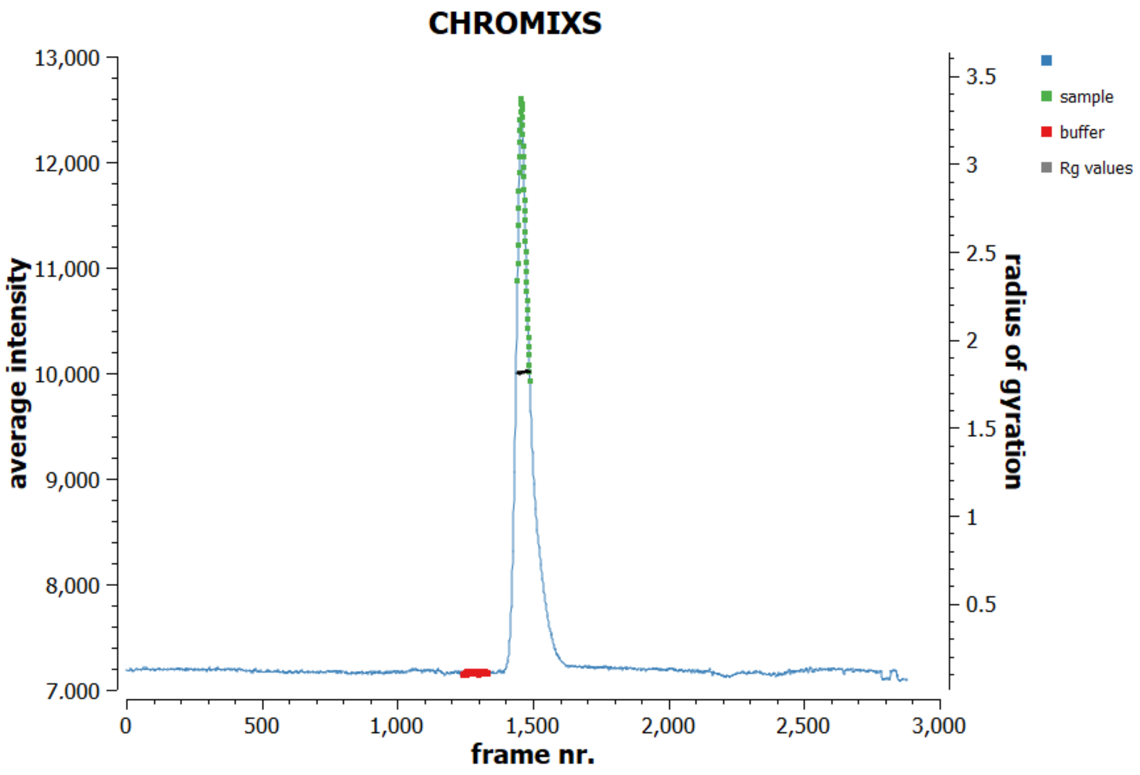


	File	From	To	Conc.	Ang Scale	I Scale
✓	chromixs01.dat	1	2584	0.00	1.000	1.000

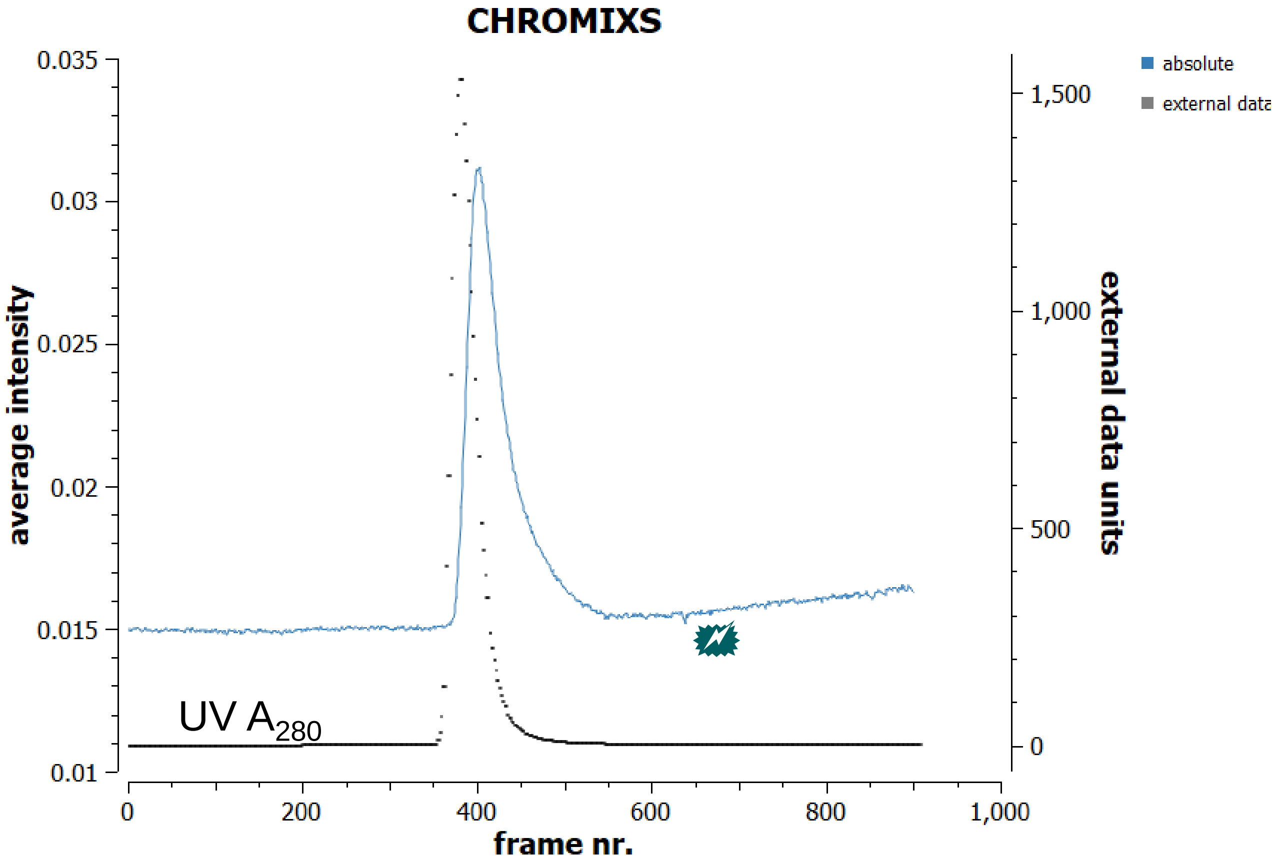
data-frames-directory	C:/Users/jeffries/Desktop/...
parent-buffer-frames	sec_0024_01288.dat sec_0024_01289.dat ...
parent-sample-frames	sec_0024_01441.dat sec_0024_01442.dat ...
range-from	17
range-to	2600
sample-code	chromixs
sample-concentration	0
sample-description	chromixs

Stable size underneath the peak, comparison with UV

- Correct background, monodispersion



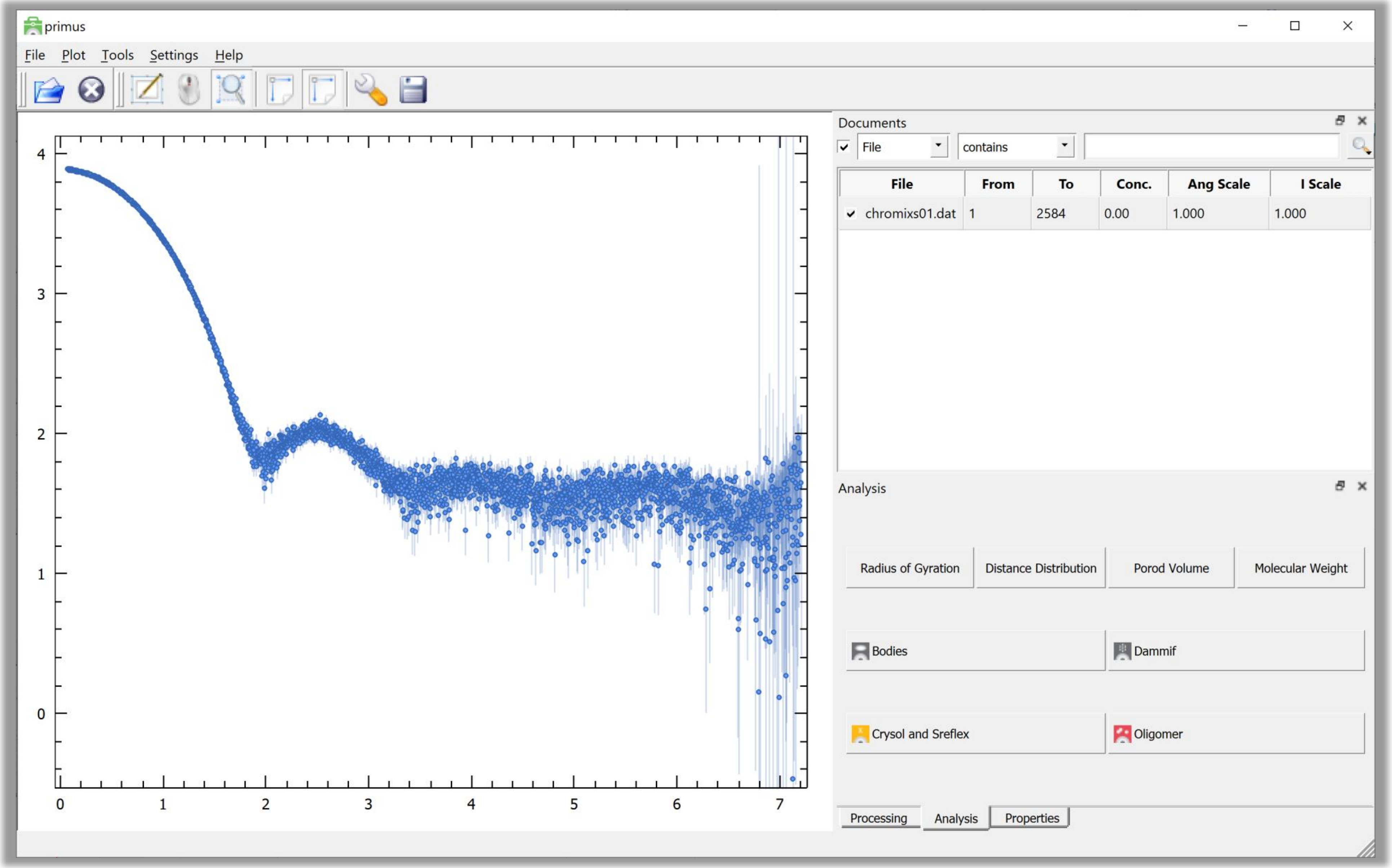
10/300



5/150

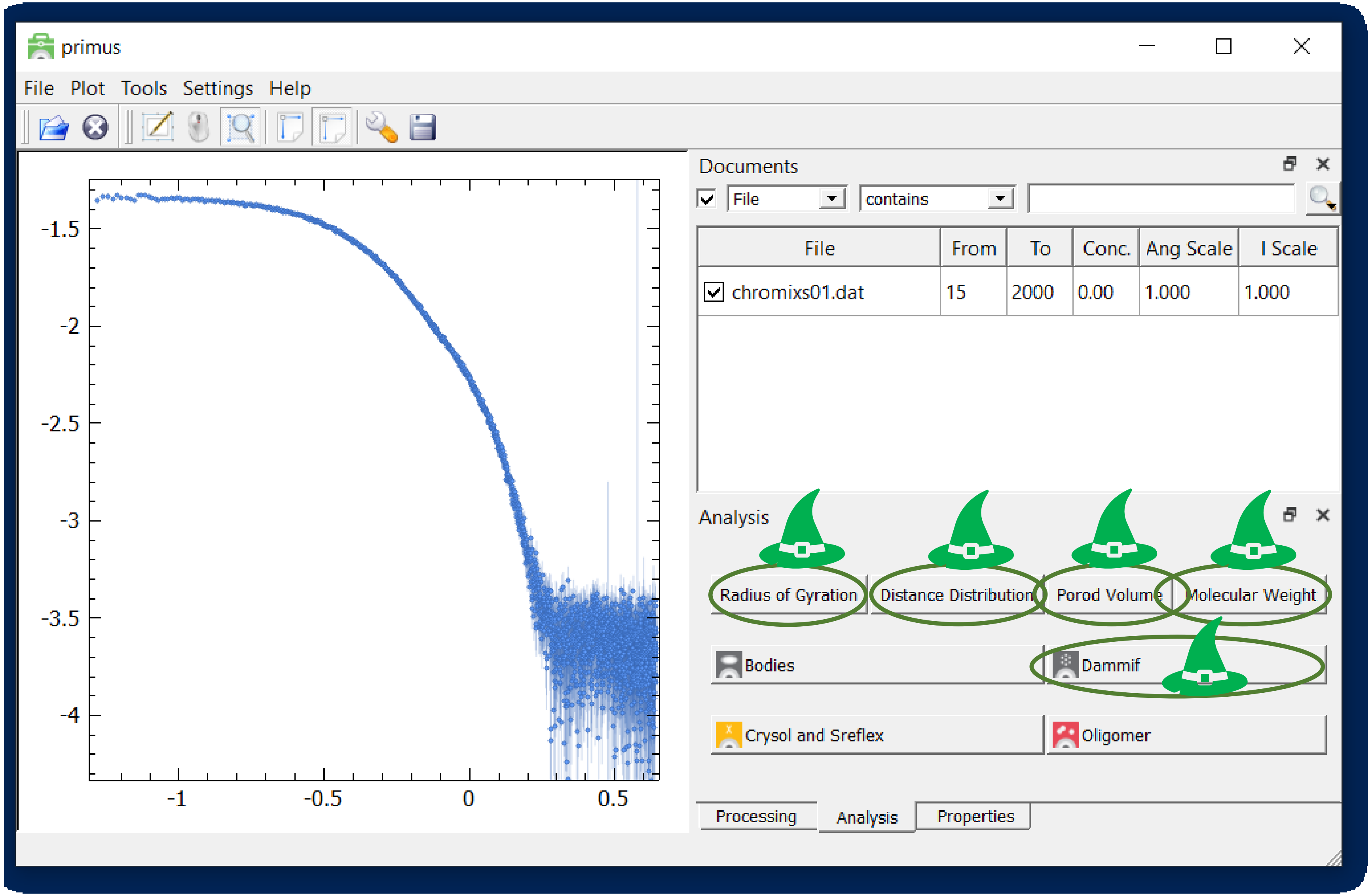
PRIMUS: parameters extraction and basic modeling

PRIMUS: processing our SEC-SAXS curve



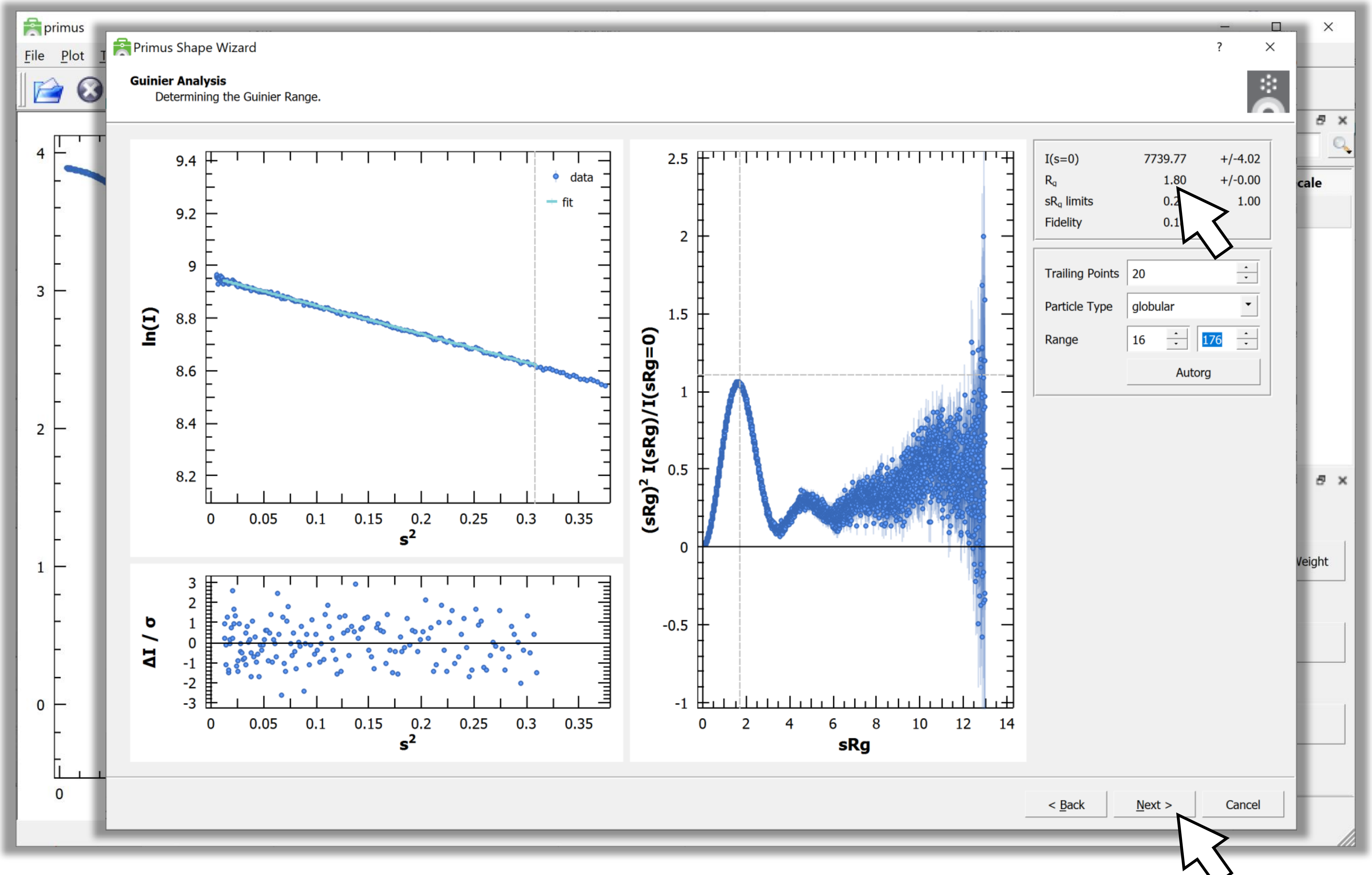
	File	From	To	Conc.	Ang Scale	I Scale
✓	chromixs01.dat	1	2584	0.00	1.000	1.000

Most Used (and Most Useful) Wizards



File	From	To	Conc.	Ang Scale	I Scale
☑ chromixs01.dat	15	2000	0.00	1.000	1.000

PRIMUS Guinier Analysis



Guinier Analysis

Determining the Guinier Range.

I(s=0)	7739.77	+/-4.02
R_g	1.80	+/-0.00
sR_g limits	0.2	1.00
Fidelity	0.1	

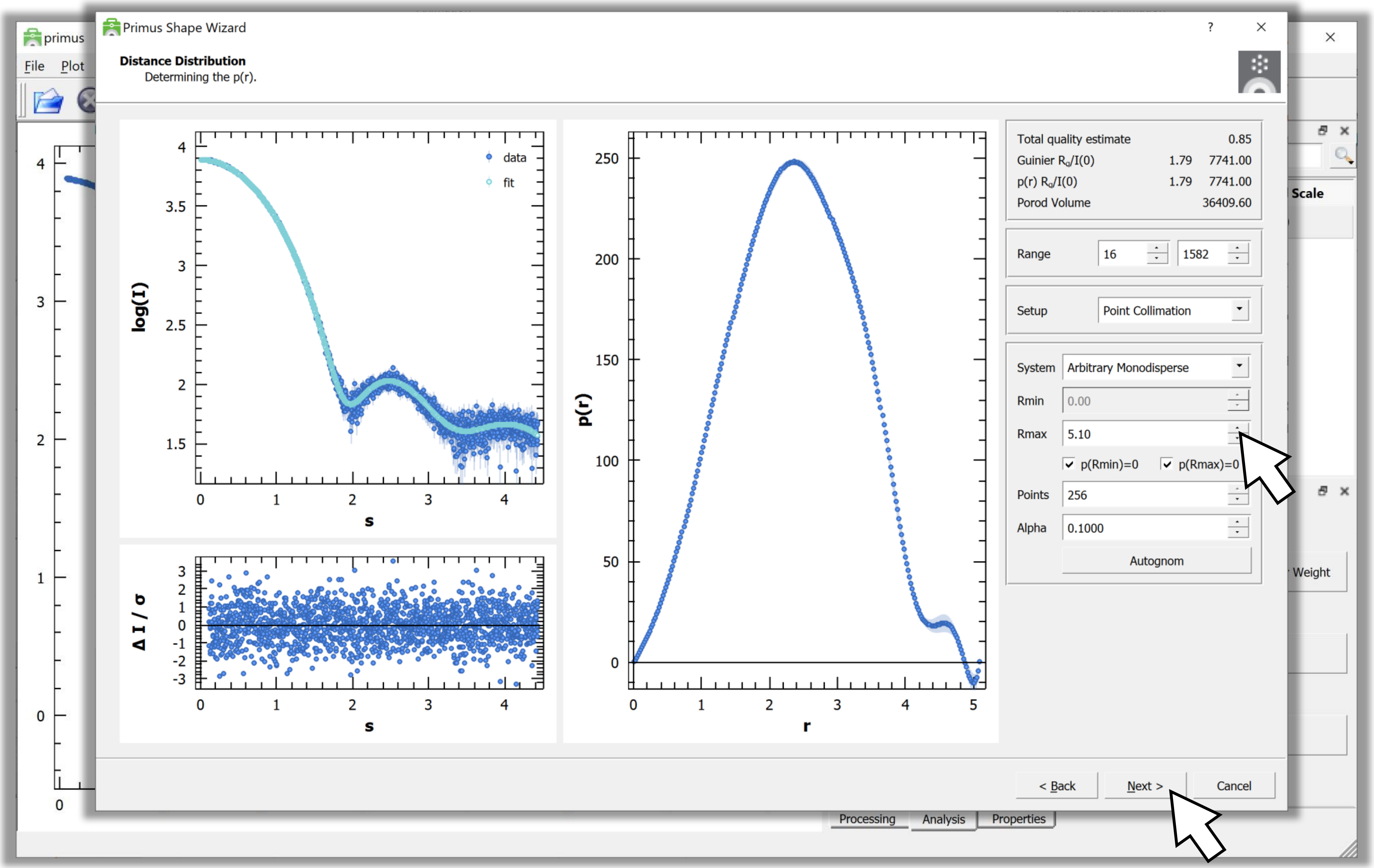
Trailing Points: 20

Particle Type: globular

Range: 16 176

Autorg

Distance Distribution (GNOM)



Primus Shape Wizard

Distance Distribution
Determining the p(r).

Total quality estimate		0.85
Guinier R_g/I(0)	1.79	7741.00
p(r) R_g/I(0)	1.79	7741.00
Porod Volume		36409.60

Range 16 1582

Setup Point Collimation

System Arbitrary Monodisperse

Rmin 0.00

Rmax 5.10

☑ p(Rmin)=0 ☑ p(Rmax)=0

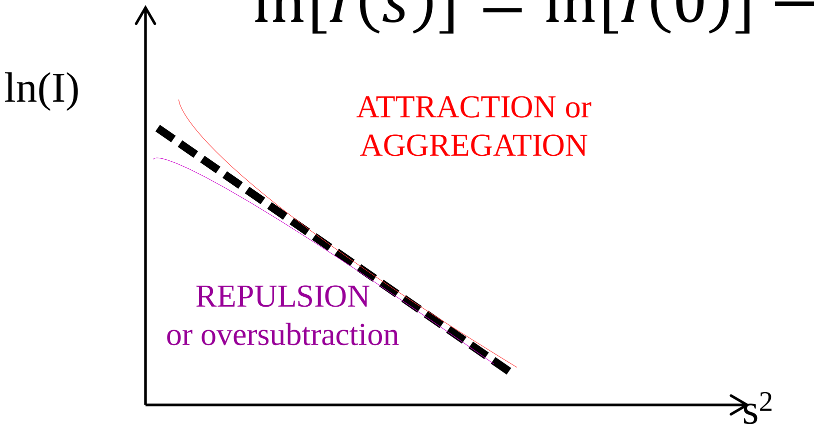
Points 256

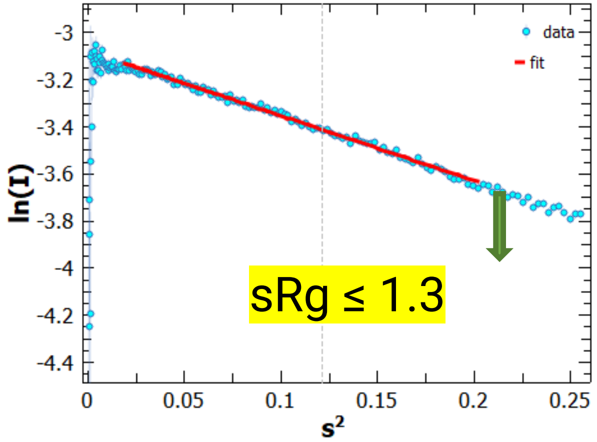
Alpha 0.1000

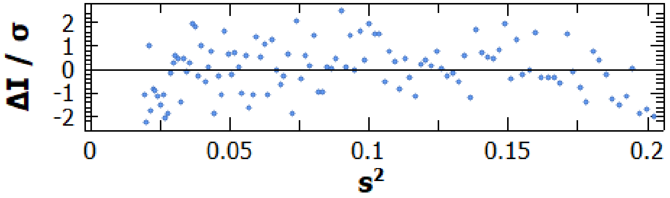
Rg from Guinier Plot *vs* from P(r)

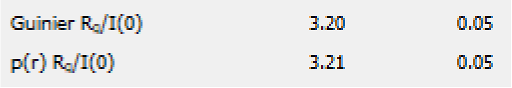
Guinier plot: check linearity !

$$\ln[I(s)] = \ln[I(0)] - \frac{1}{3}{R_g}^2 s^2$$



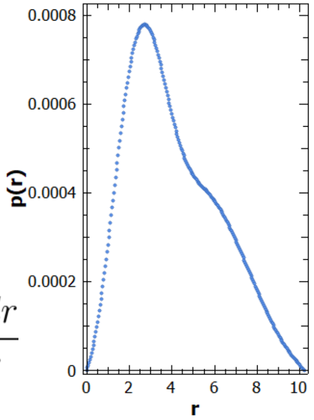


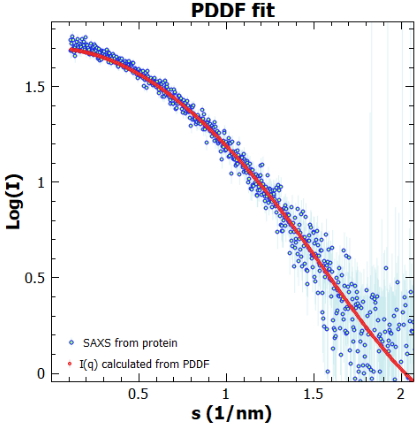




Guinier R_g/I(0)	3.20	0.05
p(r) R_g/I(0)	3.21	0.05

$$R_g^2 = \frac{1}{2}\frac{\int p(r) r^2 dr}{\int p(r) dr}$$

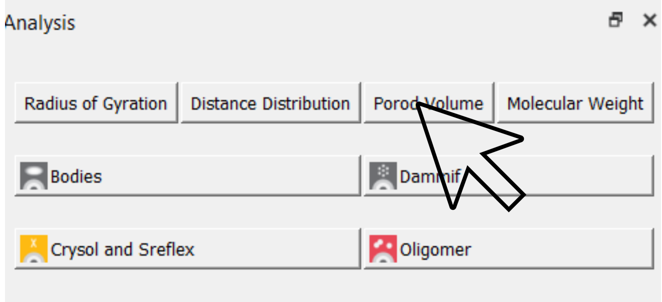


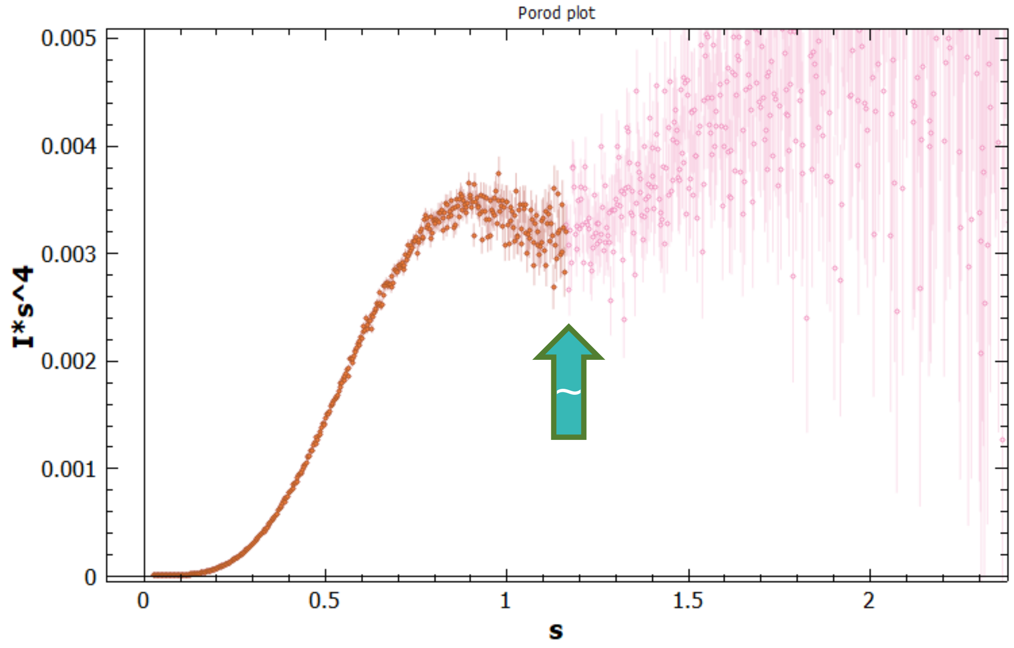


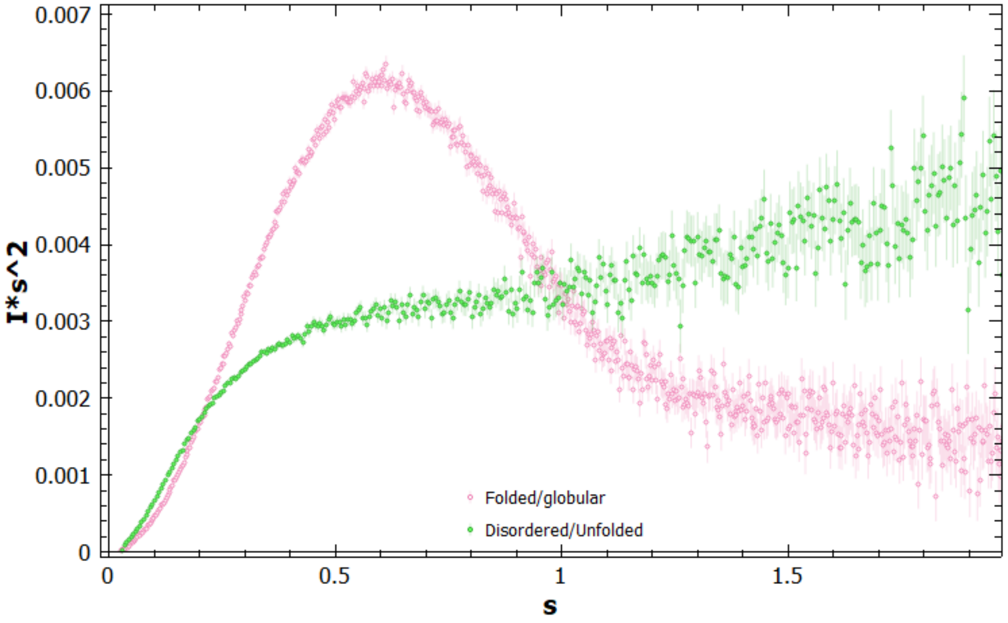
- ➢ Compare Rg (Guinier) *vs* Rg (PDDF): if Rg Guinier blatantly >…aggregates

Porod volume Vp

- Depends on a good background subtraction !
- For proteins MW~0.625*Vp NAs MW~Vp (check consistency with theoretical one !)
- May „look smaller“ for disodered molecules
- DATPOROD or GUI wizard
- Evtl.from „regularized“ curve from PDDF







$$V = 2\pi I(0)/\tilde{Q}$$

$$Q = \int_0^{\infty} s^2 I \, ds$$

Molecular mass (Mr, MW...)

- MW from I(0) [I(s) or PDDF-regularization]: a physical measurement
  - calibration: H_2O (T!), BSA
  - ~10% accuracy for „clean“ globular protein data
  - Accuracy crucially dependent on CONCENTRATION I(0)/c
  - n/a if data were manually scaled (e.g. Bkg correction, c inaccurate)
  - Sensitive to aggregation/structure factor

≠

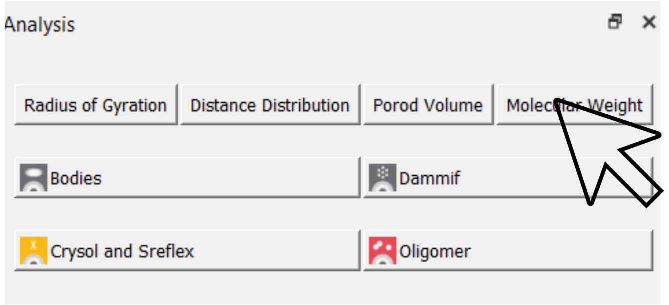
- SEC-SAXS MW from I(0): matching A280 (...) otherwise n/a

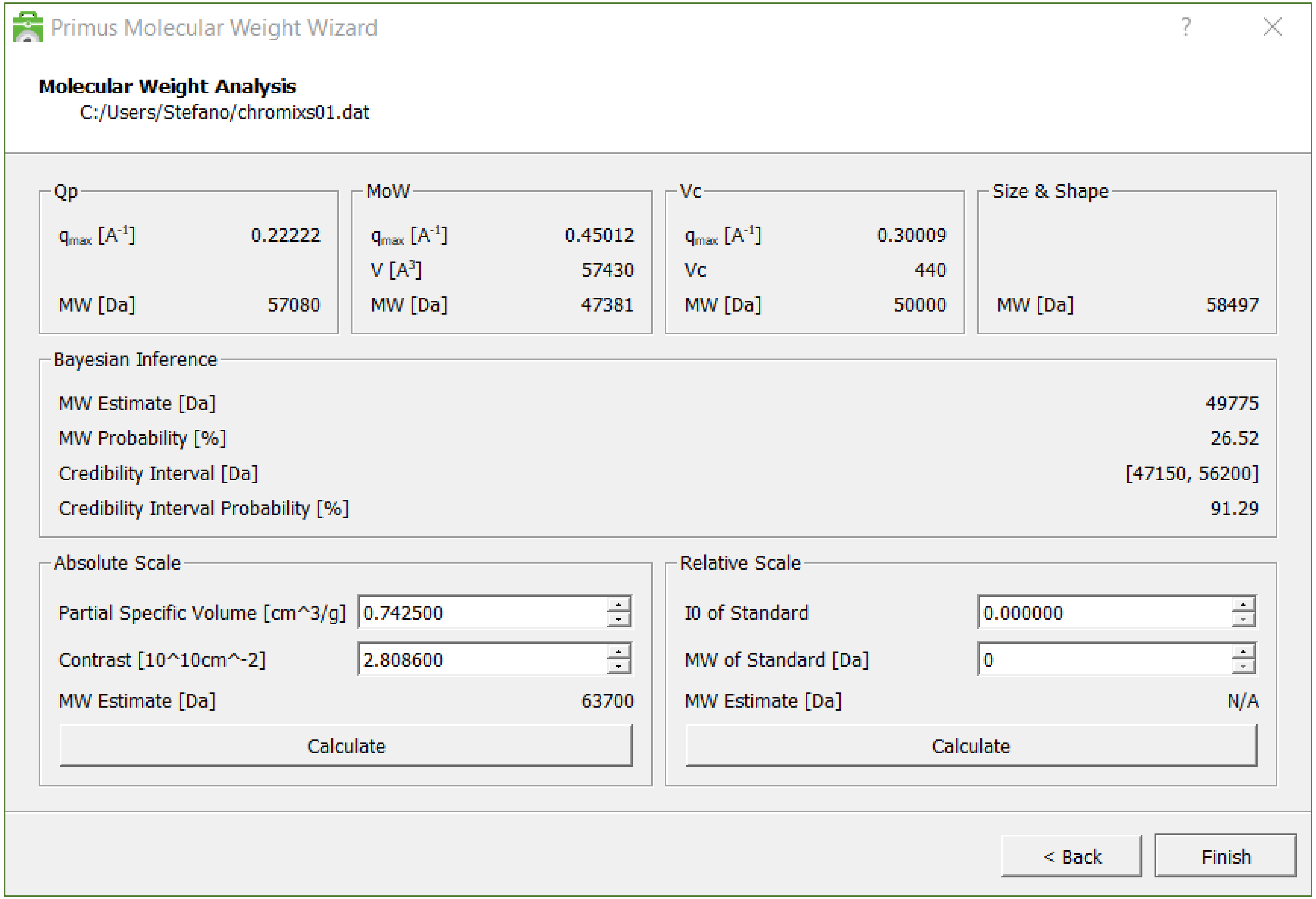
- Other MW metrics e.g. MoW, s&s, Vc, Qp, Vp...
  - *Bayesian* MW: e.g.

Bayesian Inference

MW Estimate [Da]	62350
MW Probability [%]	70.38
Credibility Interval [Da]	[60200, 66250]
Credibility Interval Probability [%]	92.44

PRIMUS MW wizard





Molecular Weight Analysis
C:/Users/Stefano/chromixs01.dat

Qp

q_{max} [A^{-1}]	0.22222
MW [Da]	57080

MoW

q_{max} [A^{-1}]	0.45012
V [A^3]	57430
MW [Da]	47381

Vc

q_{max} [A^{-1}]	0.30009
Vc	440
MW [Da]	50000

Size & Shape

MW [Da]	58497

Bayesian Inference

MW Estimate [Da]	49775
MW Probability [%]	26.52
Credibility Interval [Da]	[47150, 56200]
Credibility Interval Probability [%]	91.29

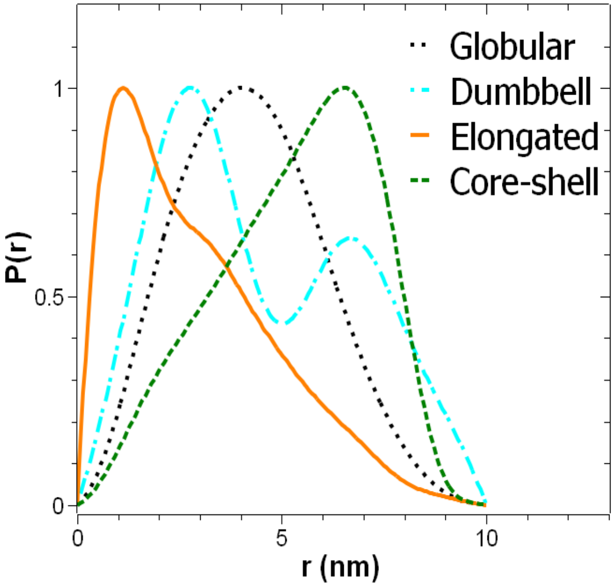
Absolute Scale

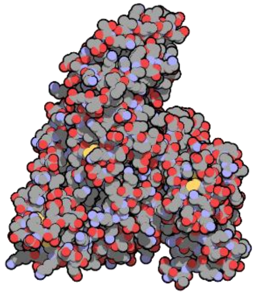
Partial Specific Volume [cm^3/g]	0.742500
Contrast [10^10cm^-2]	2.808600
MW Estimate [Da]	63700

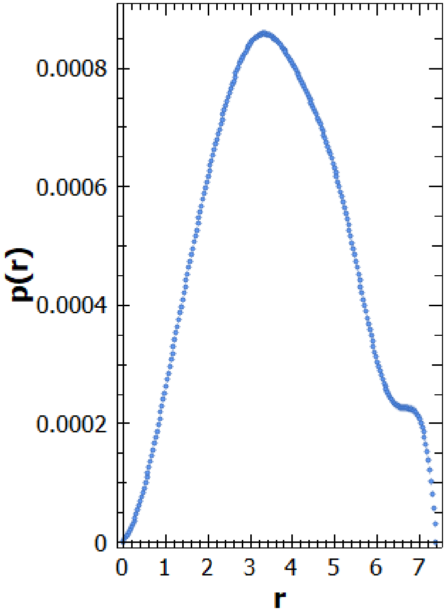
Relative Scale

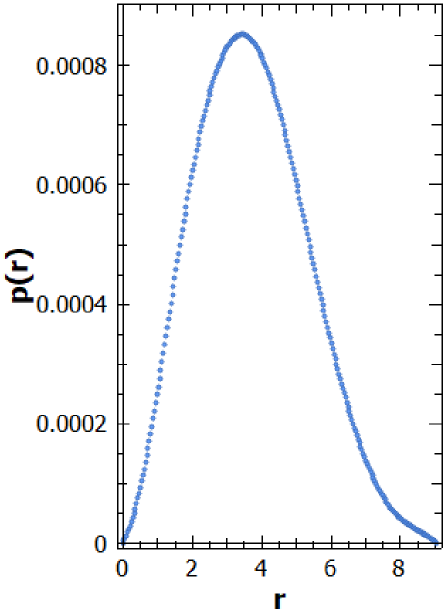
I0 of Standard	0.000000
MW of Standard [Da]	0
MW Estimate [Da]	N/A

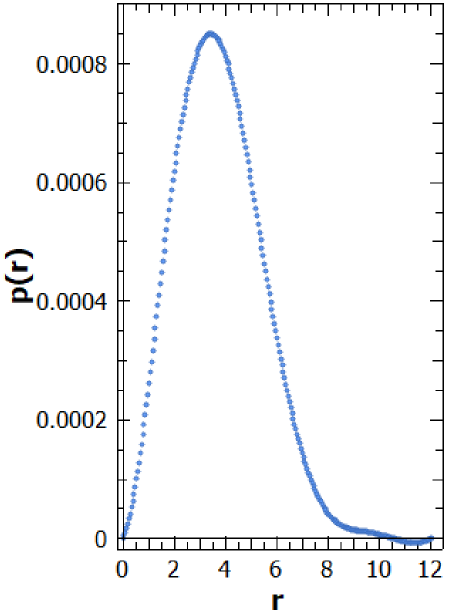
Dmax & PDDF

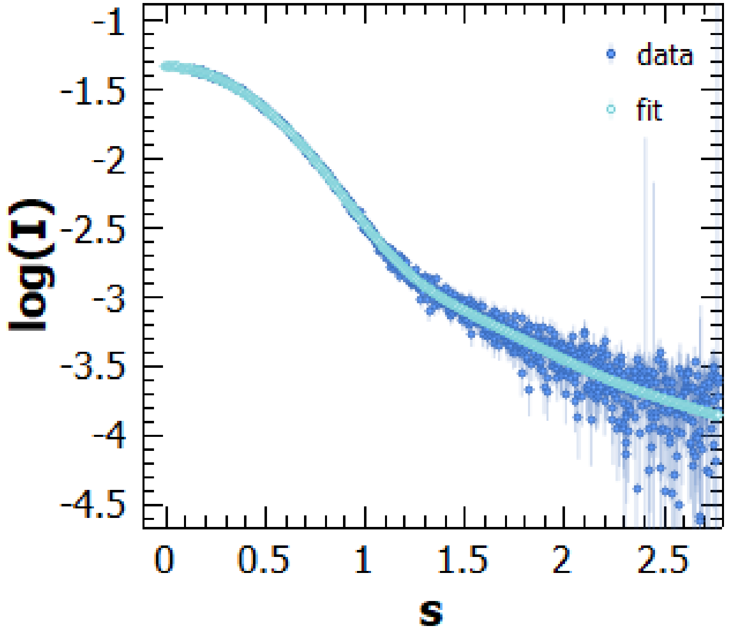

Globular
Dumbbell
Elongated
Core-shell
P(r)
0
0.5
1
0
5
10
r (nm)




p(r)
0
0.0002
0.0004
0.0006
0.0008
0 1 2 3 4 5 6 7
r


p(r)
0
0.0002
0.0004
0.0006
0.0008
0 2 4 6 8
r

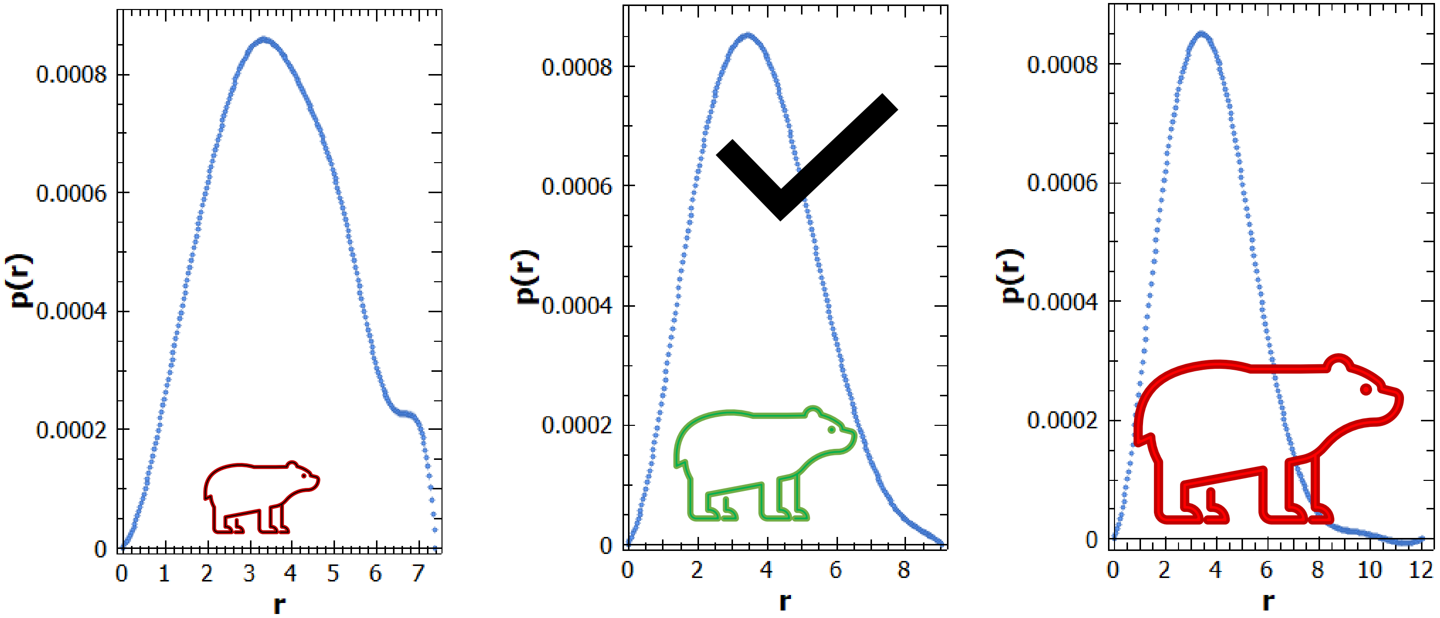

p(r)
0
0.0002
0.0004
0.0006
0.0008
0 2 4 6 8 10 12
r

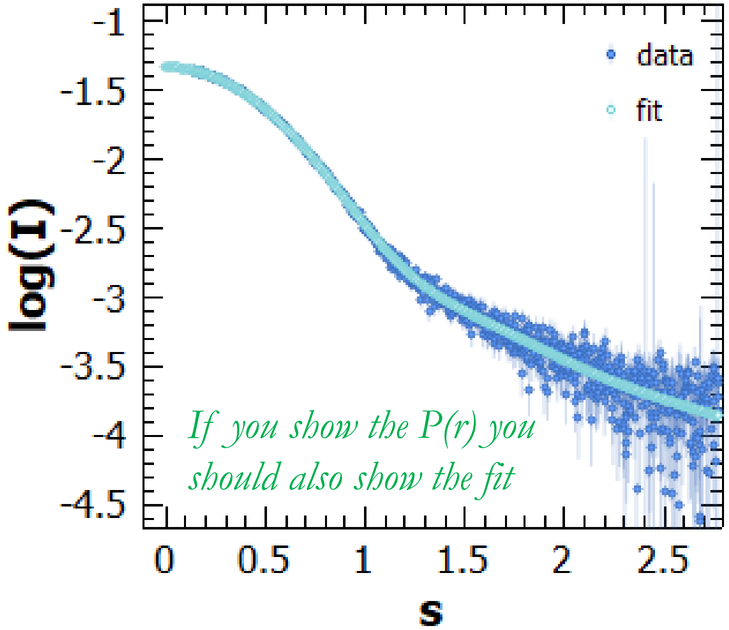

data
fit
log(I)
-1
-1.5
-2
-2.5
-3
-3.5
-4
-4.5
0 0.5 1 1.5 2 2.5
s

Dmax & PDDF

- P(r) for a compact folded globular protein
- Unless „extrapolated“ Dmax $\leq \pi/s_{min}$
- „Perceptual criteria“
- Unlikely to be more precise than 0.5 nm
- Dmax and fit only as good as the PDDF

Warning: D_{max}*s_{min} greater than PI.

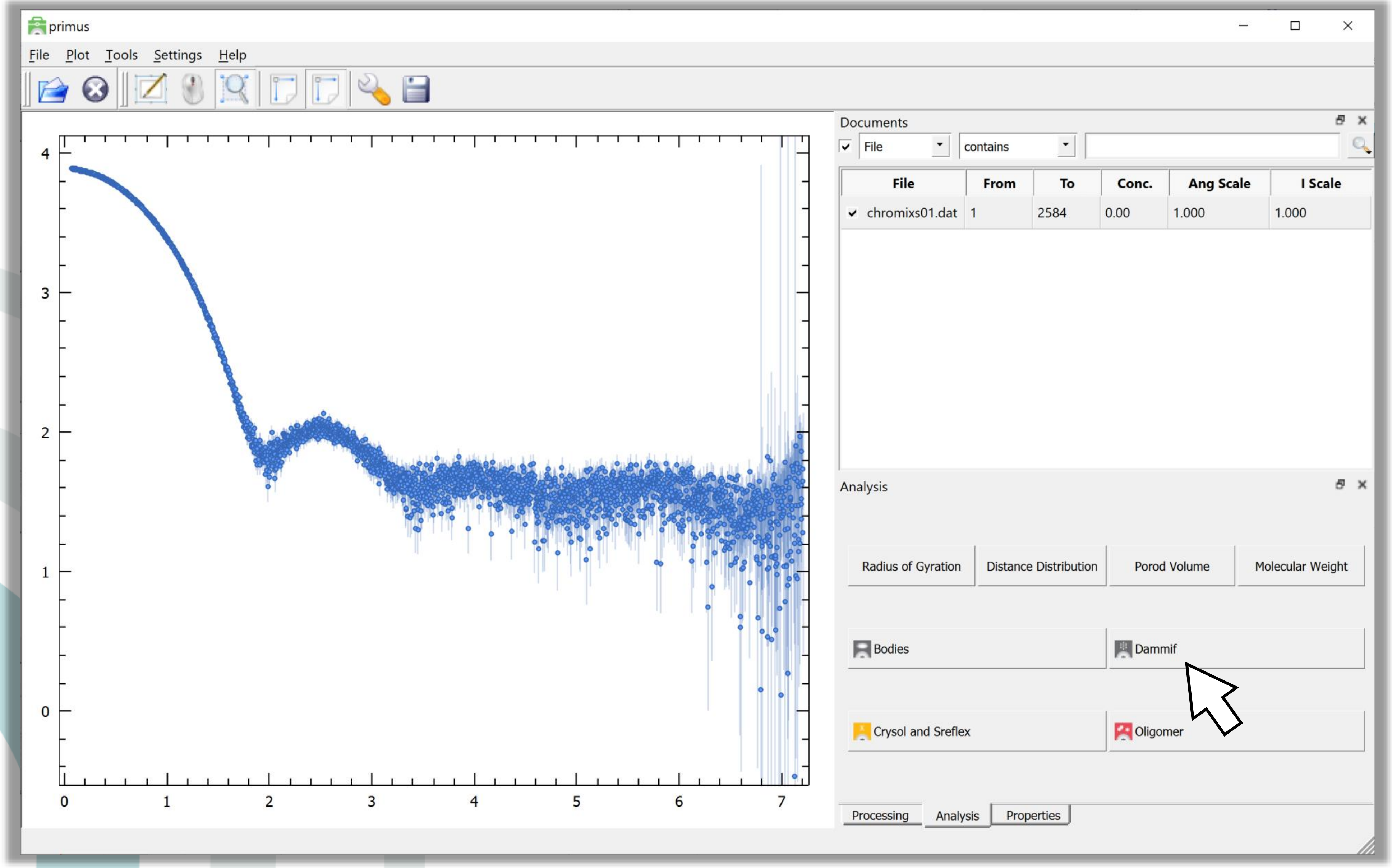




If you show the P(r) you should also show the fit

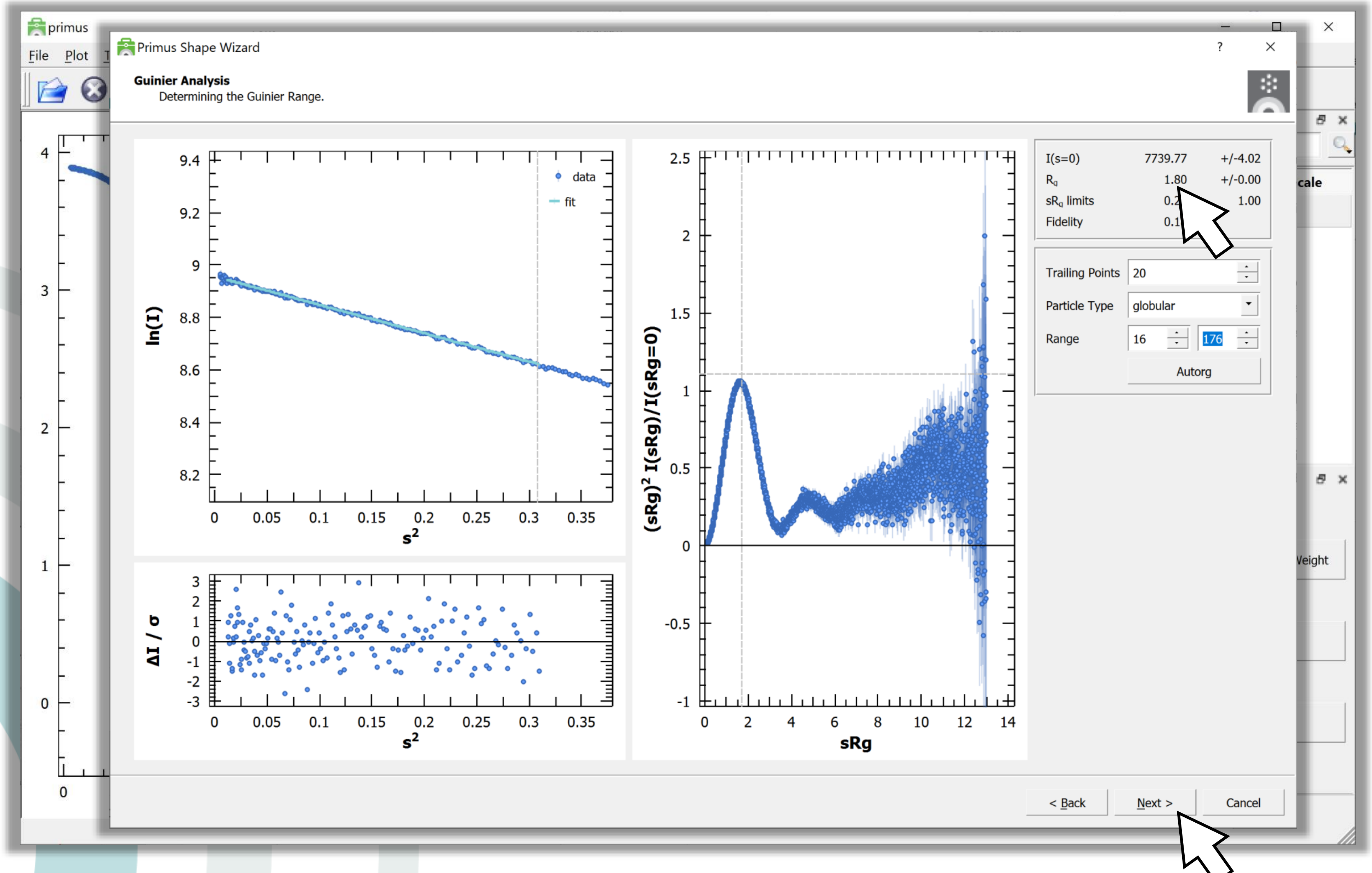
With a good P(r), we can start the modeling

PRIMUS: processing our SEC-SAXS curve



File	From	To	Conc.	Ang Scale	I Scale
✓ chromixs01.dat	1	2584	0.00	1.000	1.000

PRIMUS Guinier Analysis



Guinier Analysis

Determining the Guinier Range.

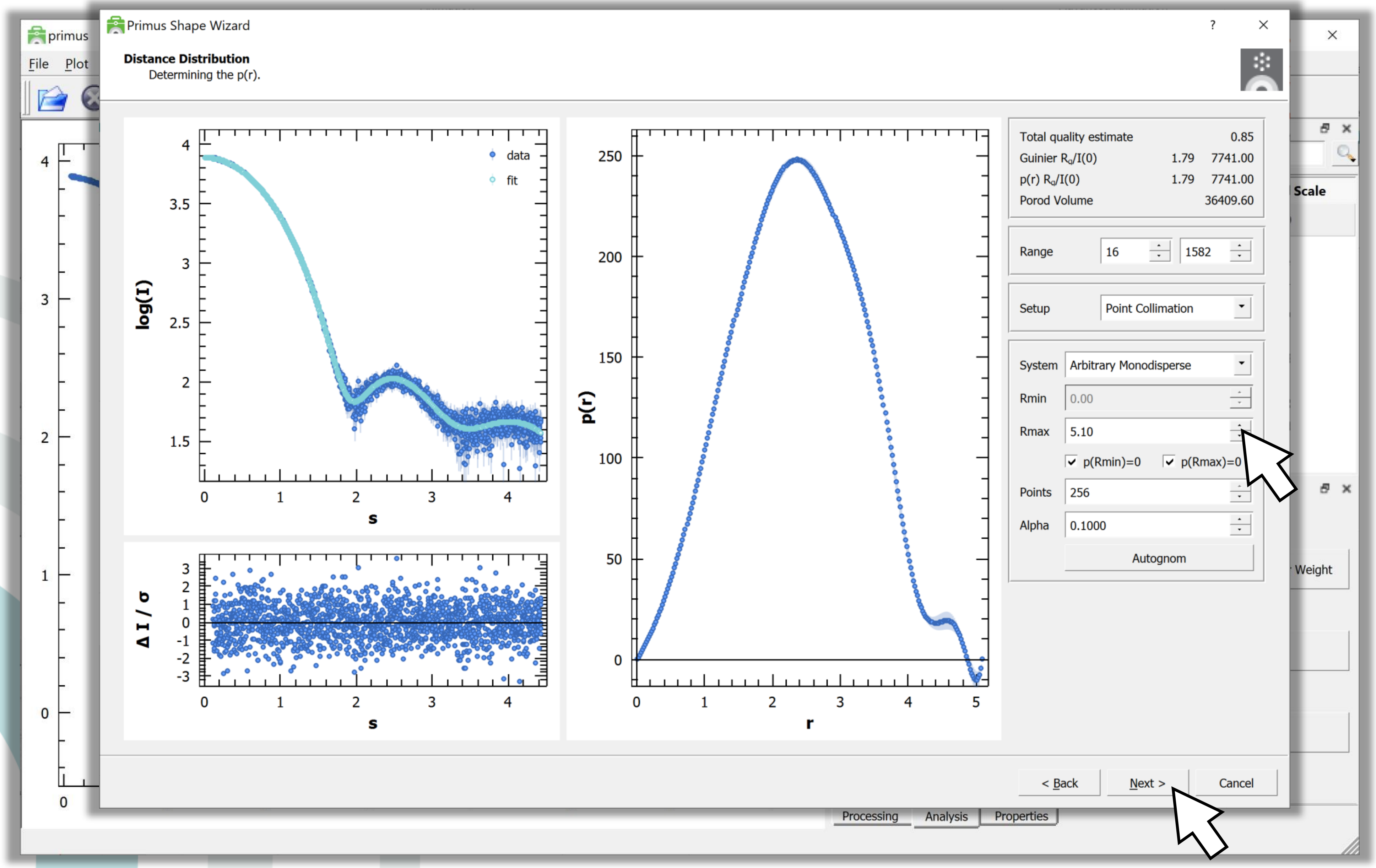
I(s=0)	7739.77	+/-4.02
R_g	1.80	+/-0.00
sR_g limits	0.2	1.00
Fidelity	0.1	

Trailing Points: 20

Particle Type: globular

Range: 16 176

Distance Distribution (GNOM)



Primus Shape Wizard

Distance Distribution

Determining the p(r).

Total quality estimate		0.85
Guinier R_g/I(0)	1.79	7741.00
p(r) R_g/I(0)	1.79	7741.00
Porod Volume		36409.60

Range 16 1582

Setup Point Collimation

System Arbitrary Monodisperse

Rmin 0.00

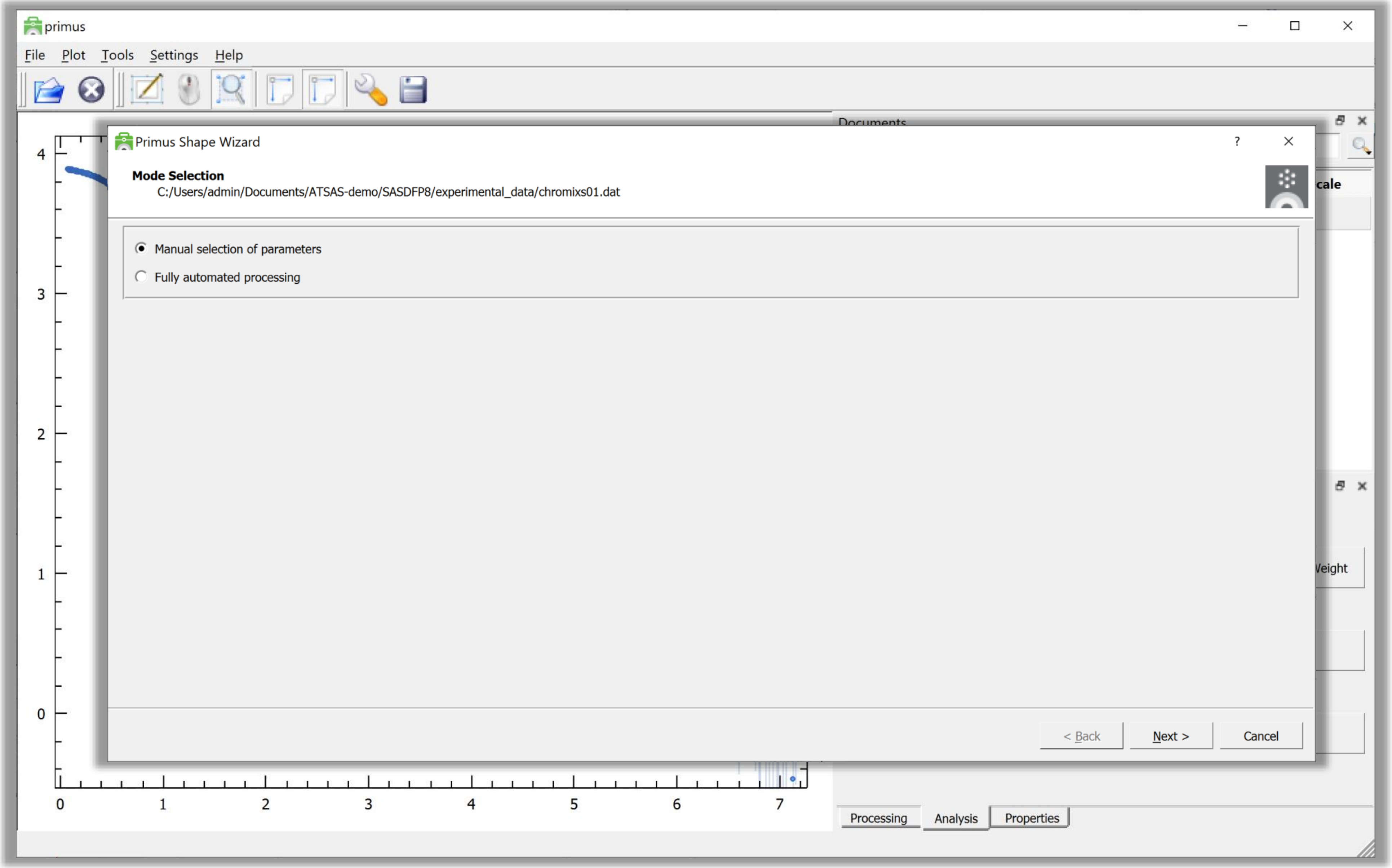
Rmax 5.10

☑ p(Rmin)=0 ☑ p(Rmax)=0

Points 256

Alpha 0.1000

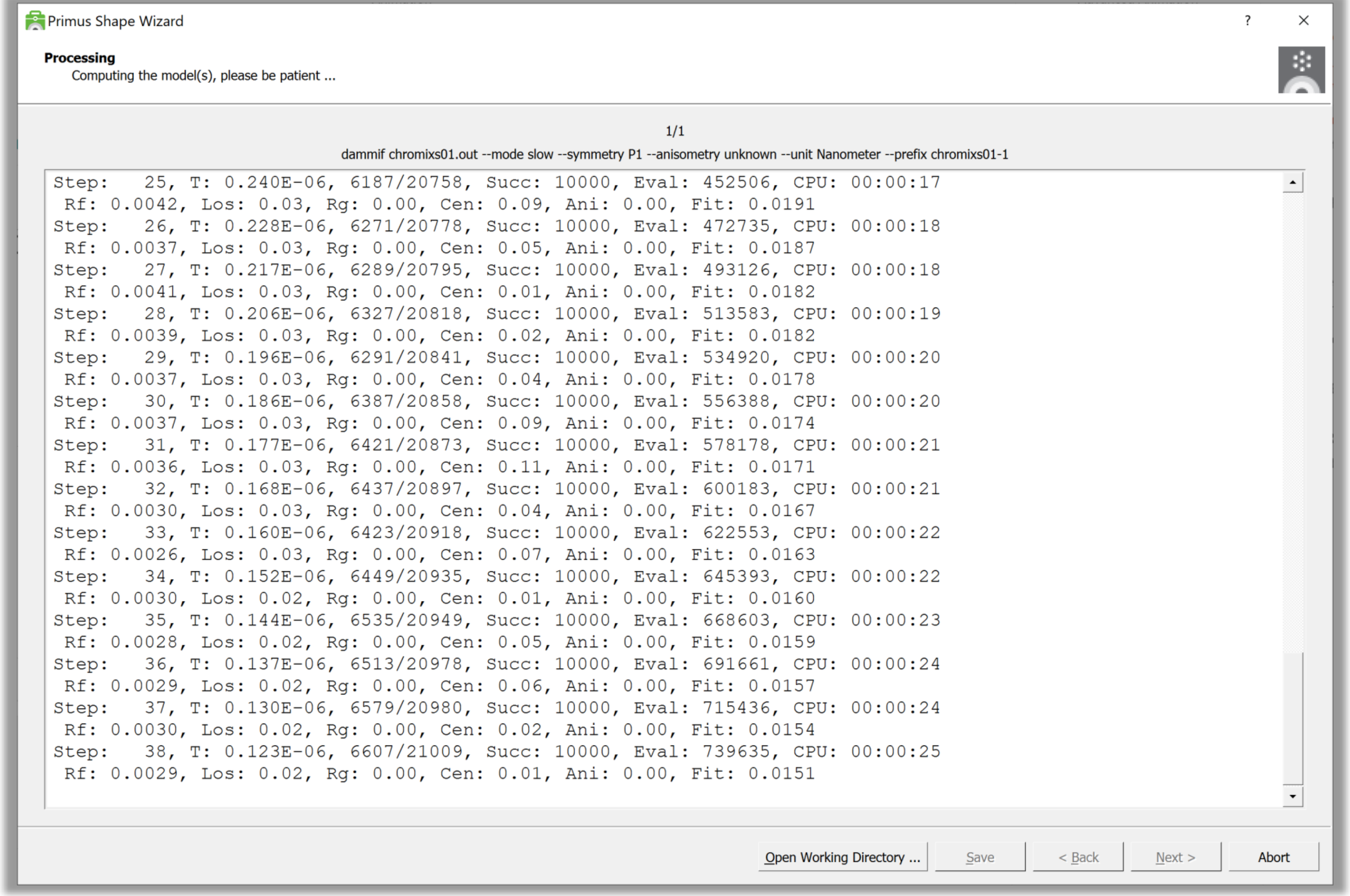
PRIMUS Shape Wizard



Mode Selection
C:/Users/admin/Documents/ATSAS-demo/SASDFP8/experimental_data/chromixs01.dat

● Manual selection of parameters
○ Fully automated processing

Running DAMMIF/DAMMIN



Processing

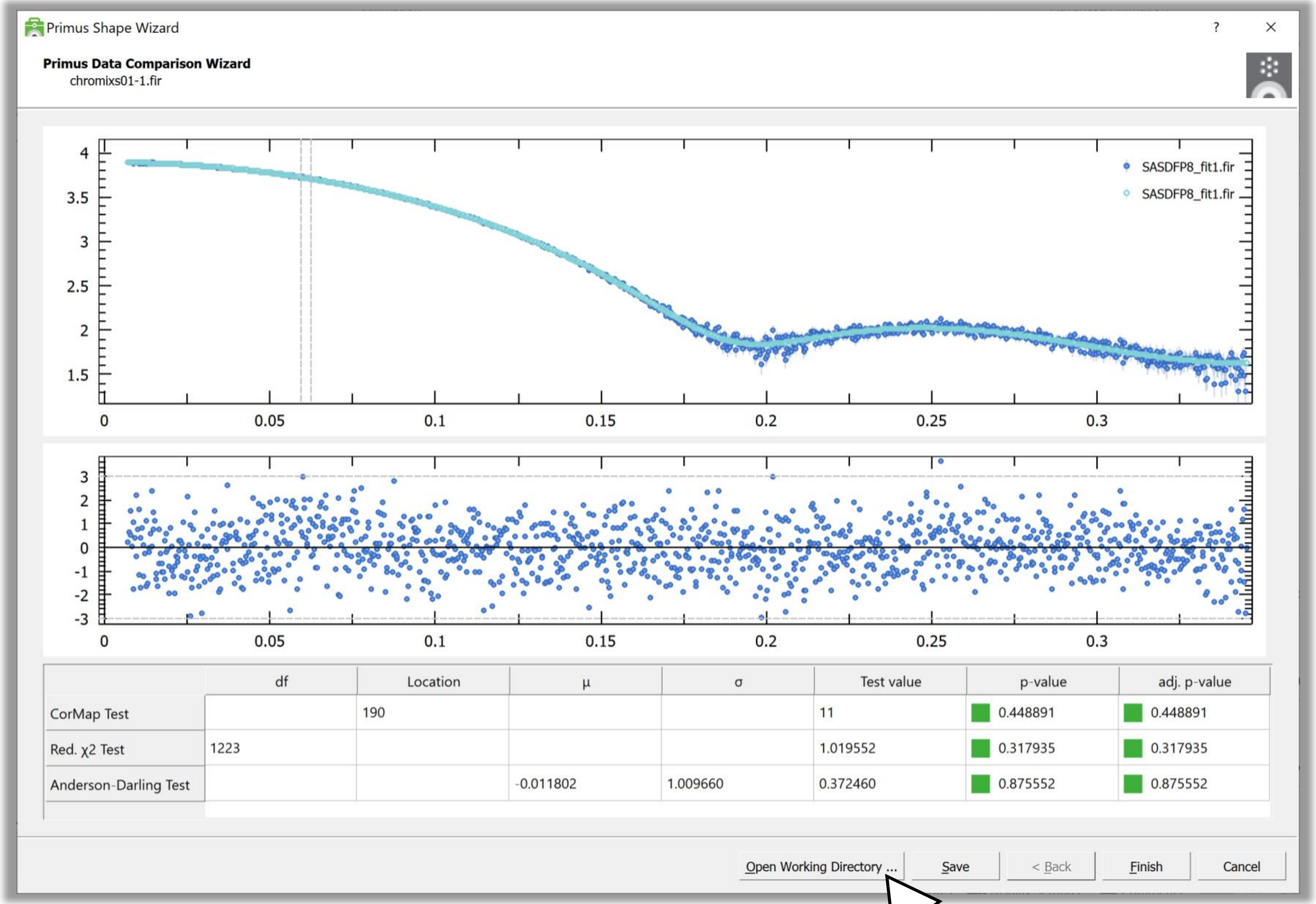
Computing the model(s), please be patient ...

1/1

```
dammif chromixs01.out --mode slow --symmetry P1 --anisometry unknown --unit Nanometer --prefix chromixs01-1
```

```
Step:   25, T: 0.240E-06, 6187/20758, Succ: 10000, Eval: 452506, CPU: 00:00:17
 Rf: 0.0042, Los: 0.03, Rg: 0.00, Cen: 0.09, Ani: 0.00, Fit: 0.0191
Step:   26, T: 0.228E-06, 6271/20778, Succ: 10000, Eval: 472735, CPU: 00:00:18
 Rf: 0.0037, Los: 0.03, Rg: 0.00, Cen: 0.05, Ani: 0.00, Fit: 0.0187
Step:   27, T: 0.217E-06, 6289/20795, Succ: 10000, Eval: 493126, CPU: 00:00:18
 Rf: 0.0041, Los: 0.03, Rg: 0.00, Cen: 0.01, Ani: 0.00, Fit: 0.0182
Step:   28, T: 0.206E-06, 6327/20818, Succ: 10000, Eval: 513583, CPU: 00:00:19
 Rf: 0.0039, Los: 0.03, Rg: 0.00, Cen: 0.02, Ani: 0.00, Fit: 0.0182
Step:   29, T: 0.196E-06, 6291/20841, Succ: 10000, Eval: 534920, CPU: 00:00:20
 Rf: 0.0037, Los: 0.03, Rg: 0.00, Cen: 0.04, Ani: 0.00, Fit: 0.0178
Step:   30, T: 0.186E-06, 6387/20858, Succ: 10000, Eval: 556388, CPU: 00:00:20
 Rf: 0.0037, Los: 0.03, Rg: 0.00, Cen: 0.09, Ani: 0.00, Fit: 0.0174
Step:   31, T: 0.177E-06, 6421/20873, Succ: 10000, Eval: 578178, CPU: 00:00:21
 Rf: 0.0036, Los: 0.03, Rg: 0.00, Cen: 0.11, Ani: 0.00, Fit: 0.0171
Step:   32, T: 0.168E-06, 6437/20897, Succ: 10000, Eval: 600183, CPU: 00:00:21
 Rf: 0.0030, Los: 0.03, Rg: 0.00, Cen: 0.04, Ani: 0.00, Fit: 0.0167
Step:   33, T: 0.160E-06, 6423/20918, Succ: 10000, Eval: 622553, CPU: 00:00:22
 Rf: 0.0026, Los: 0.03, Rg: 0.00, Cen: 0.07, Ani: 0.00, Fit: 0.0163
Step:   34, T: 0.152E-06, 6449/20935, Succ: 10000, Eval: 645393, CPU: 00:00:22
 Rf: 0.0030, Los: 0.02, Rg: 0.00, Cen: 0.01, Ani: 0.00, Fit: 0.0160
Step:   35, T: 0.144E-06, 6535/20949, Succ: 10000, Eval: 668603, CPU: 00:00:23
 Rf: 0.0028, Los: 0.02, Rg: 0.00, Cen: 0.05, Ani: 0.00, Fit: 0.0159
Step:   36, T: 0.137E-06, 6513/20978, Succ: 10000, Eval: 691661, CPU: 00:00:24
 Rf: 0.0029, Los: 0.02, Rg: 0.00, Cen: 0.06, Ani: 0.00, Fit: 0.0157
Step:   37, T: 0.130E-06, 6579/20980, Succ: 10000, Eval: 715436, CPU: 00:00:24
 Rf: 0.0030, Los: 0.02, Rg: 0.00, Cen: 0.02, Ani: 0.00, Fit: 0.0154
Step:   38, T: 0.123E-06, 6607/21009, Succ: 10000, Eval: 739635, CPU: 00:00:25
 Rf: 0.0029, Los: 0.02, Rg: 0.00, Cen: 0.01, Ani: 0.00, Fit: 0.0151
```

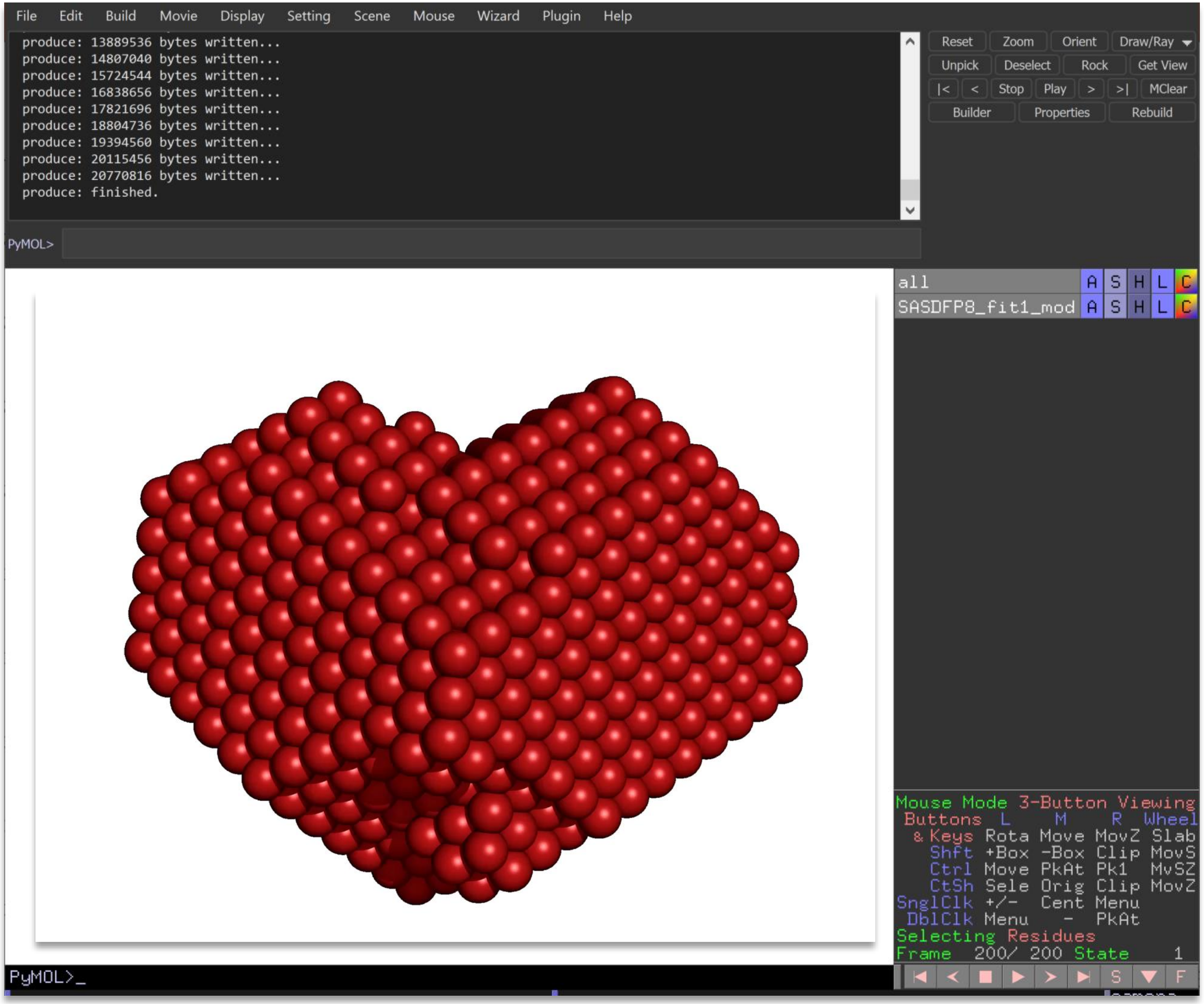
Running DAMMIF/DAMMIN



Primus Data Comparison Wizard
chromixs01-1.fir

	df	Location	µ	σ	Test value	p-value	adj. p-value
CorMap Test		190			11	🟩 0.448891	🟩 0.448891
Red. χ2 Test	1223				1.019552	🟩 0.317935	🟩 0.317935
Anderson-Darling Test			-0.011802	1.009660	0.372460	🟩 0.875552	🟩 0.875552

Let‘s now open one of the 10 models....

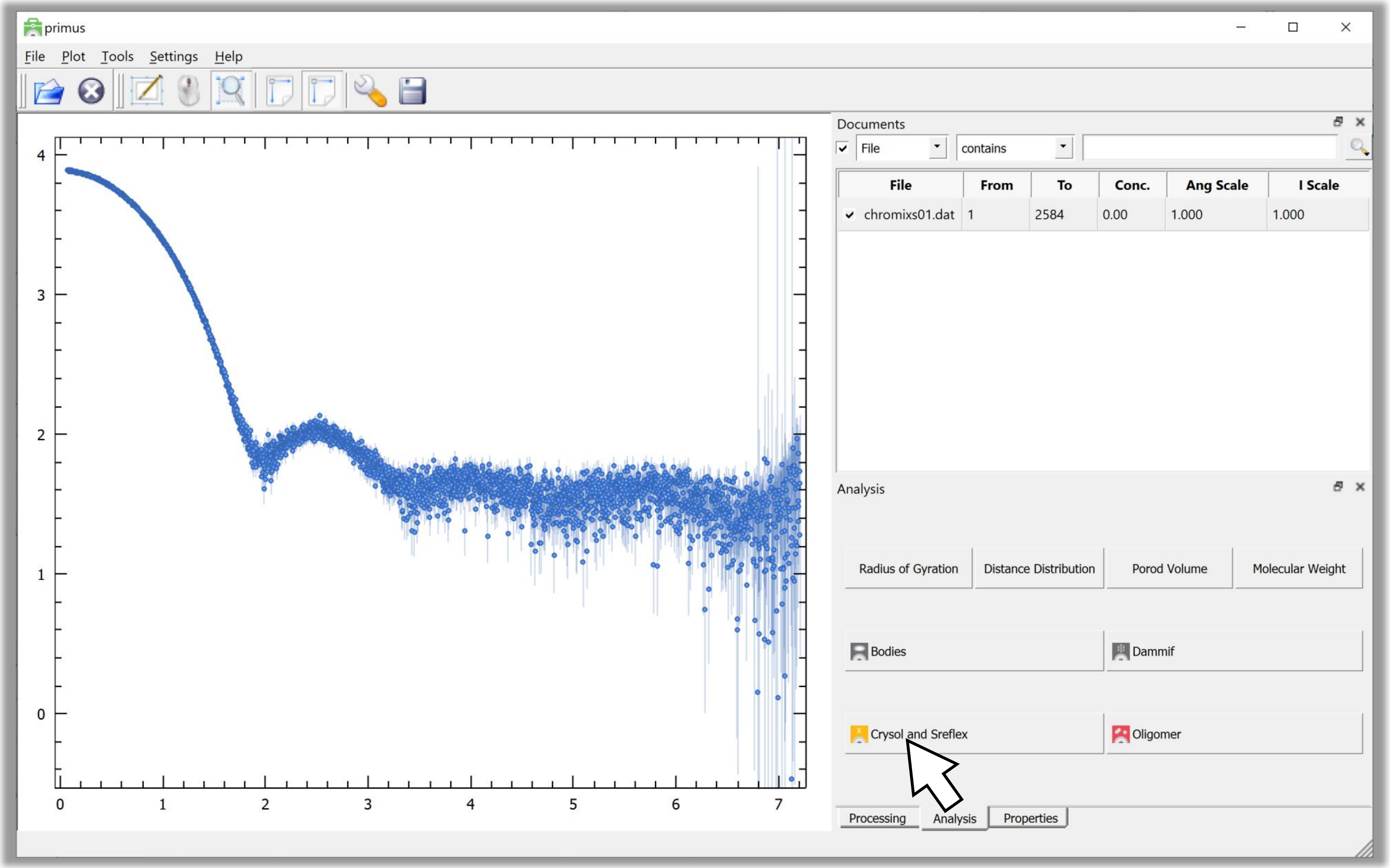


```
produce: 13889536 bytes written...
produce: 14807040 bytes written...
produce: 15724544 bytes written...
produce: 16838656 bytes written...
produce: 17821696 bytes written...
produce: 18804736 bytes written...
produce: 19394560 bytes written...
produce: 20115456 bytes written...
produce: 20770816 bytes written...
produce: finished.
```

```
PyMOL>
```

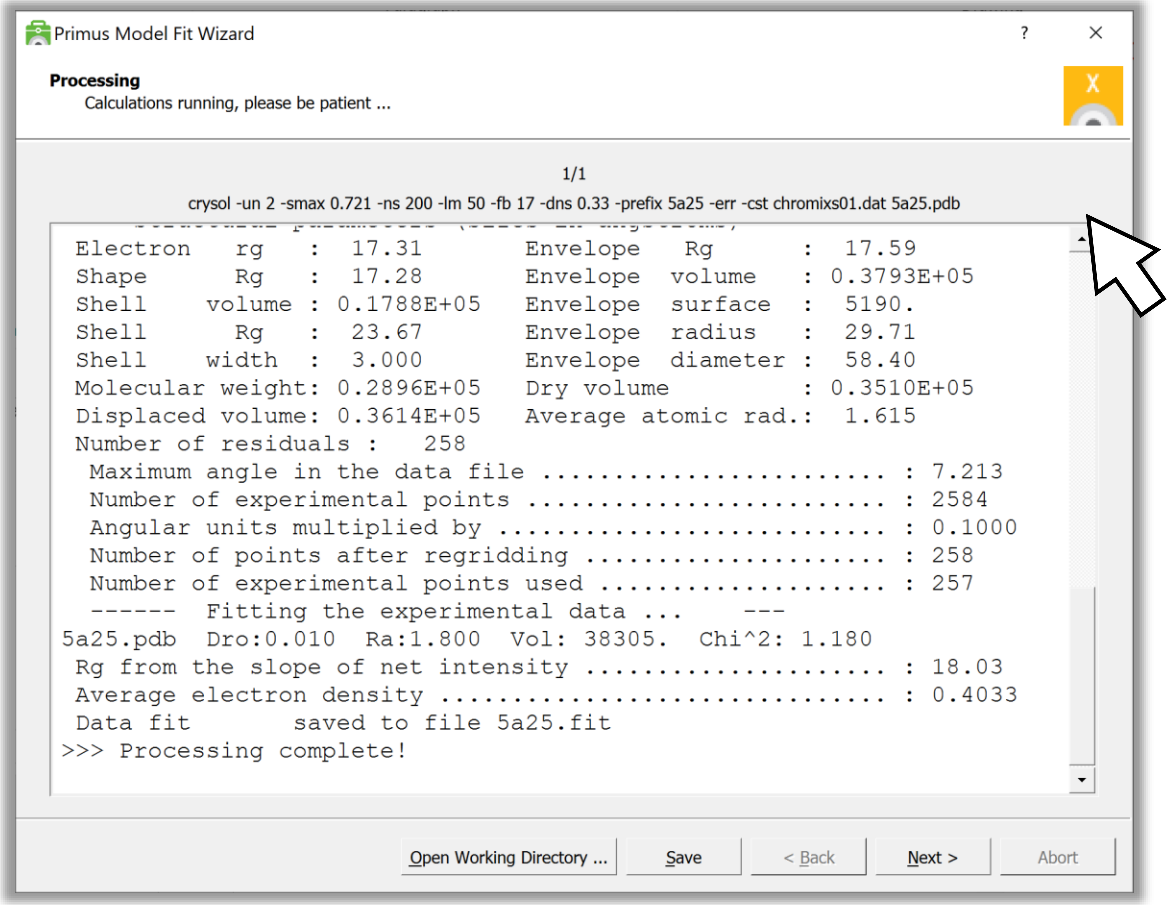
```
PyMOL>_
```

PRIMUS: CRYSOL



File	From	To	Conc.	Ang Scale	I Scale
✓ chromixs01.dat	1	2584	0.00	1.000	1.000

PRIMUS: CRYSOL



Processing

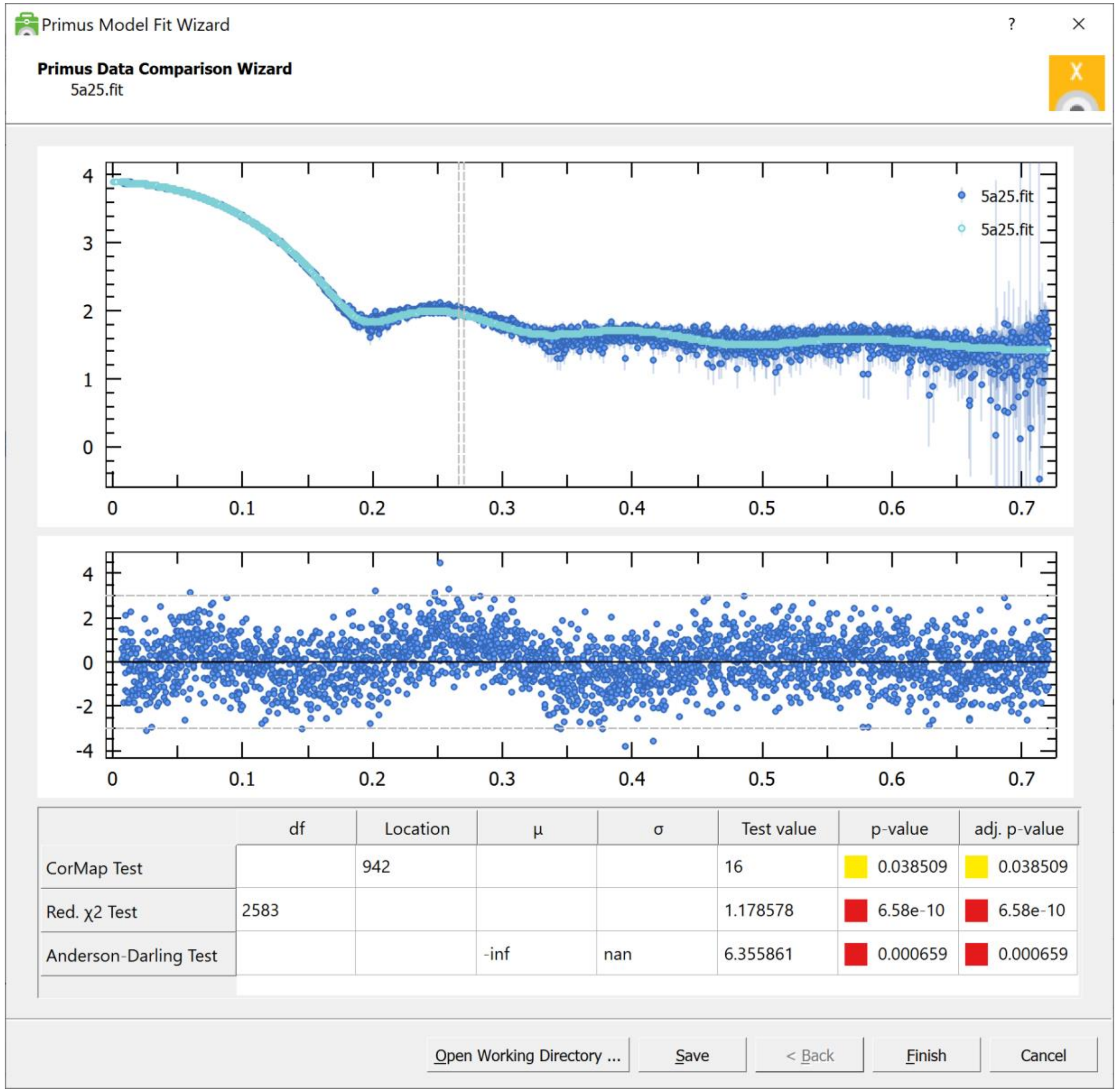
Calculations running, please be patient ...

1/1

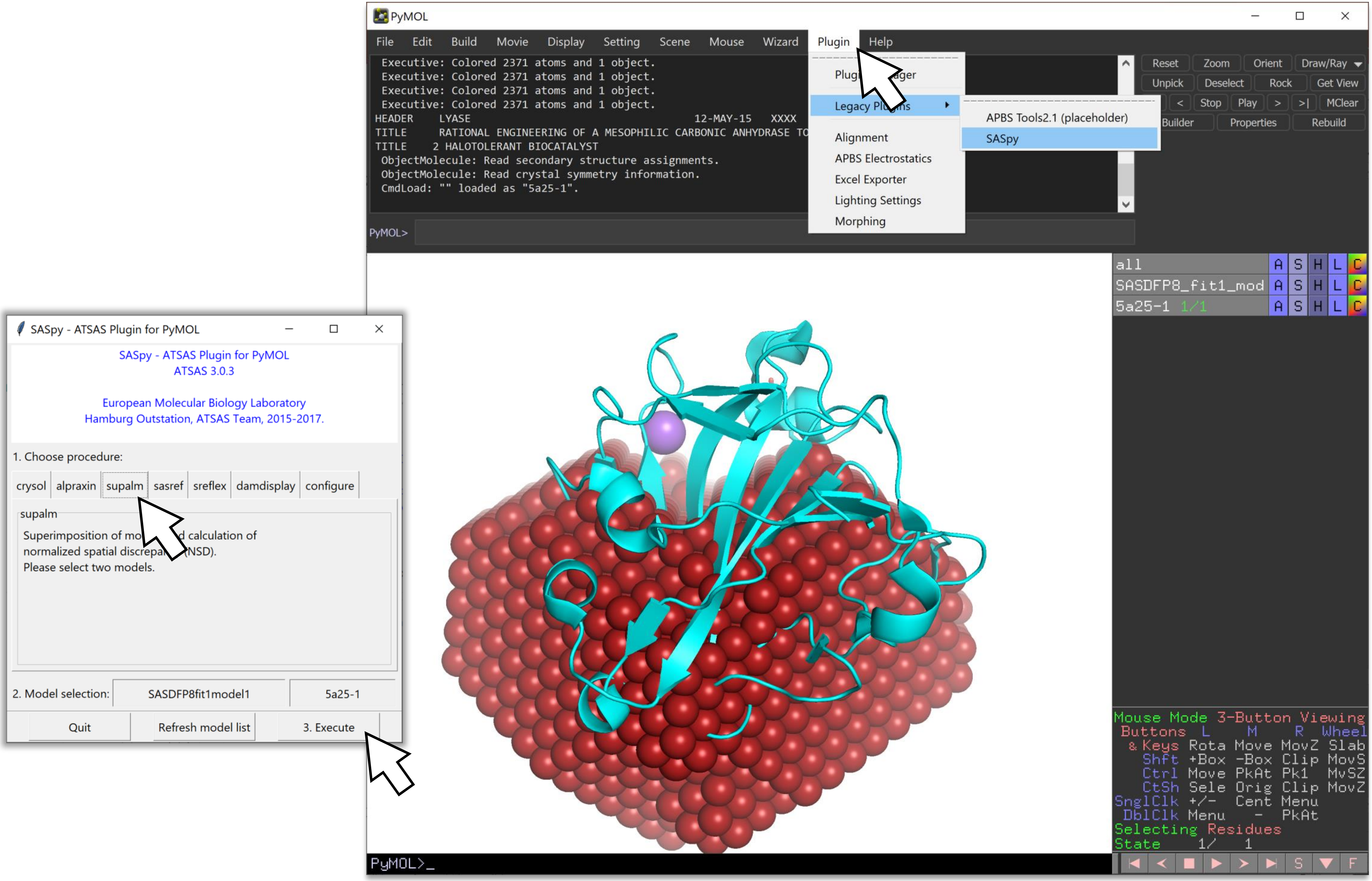
`crysol -un 2 -smax 0.721 -ns 200 -lm 50 -fb 17 -dns 0.33 -prefix 5a25 -err -cst chromixs01.dat 5a25.pdb`

```
 Electron   rg    :  17.31       Envelope   Rg       :  17.59
 Shell      Rg    :  17.28       Envelope  volume    : 0.3793E+05
 Shell     volume : 0.1788E+05   Envelope  surface   :  5190.
 Shell      Rg    :  23.67       Envelope  radius    :  29.71
 Shell     width  :  3.000       Envelope  diameter  :  58.40
 Molecular weight: 0.2896E+05   Dry volume           : 0.3510E+05
 Displaced volume: 0.3614E+05   Average atomic rad.:  1.615
 Number of residuals :   258
  Maximum angle in the data file ........................ : 7.213
  Number of experimental points ......................... : 2584
  Angular units multiplied by ........................... : 0.1000
  Number of points after regridding ..................... : 258
  Number of experimental points used .................... : 257
  ------  Fitting the experimental data ...    ---
5a25.pdb  Dro:0.010  Ra:1.800  Vol: 38305.  Chi^2: 1.180
 Rg from the slope of net intensity .................... : 18.03
 Average electron density .............................. : 0.4033
 Data fit       saved to file 5a25.fit
>>> Processing complete!
```

PRIMUS: CRYSOL



	df	Location	µ	σ	Test value	p-value	adj. p-value
CorMap Test		942			16	🟨 0.038509	🟨 0.038509
Red. χ2 Test	2583				1.178578	🟥 6.58e-10	🟥 6.58e-10
Anderson-Darling Test			-inf	nan	6.355861	🟥 0.000659	🟥 0.000659



```
 Executive: Colored 2371 atoms and 1 object.
 Executive: Colored 2371 atoms and 1 object.
 Executive: Colored 2371 atoms and 1 object.
 Executive: Colored 2371 atoms and 1 object.
HEADER    LYASE                                   12-MAY-15   XXXX
TITLE     RATIONAL ENGINEERING OF A MESOPHILIC CARBONIC ANHYDRASE TO
TITLE    2 HALOTOLERANT BIOCATALYST
 ObjectMolecule: Read secondary structure assignments.
 ObjectMolecule: Read crystal symmetry information.
 CmdLoad: "" loaded as "5a25-1".
PyMOL>
```

SASpy - ATSAS Plugin for PyMOL
ATSAS 3.0.3

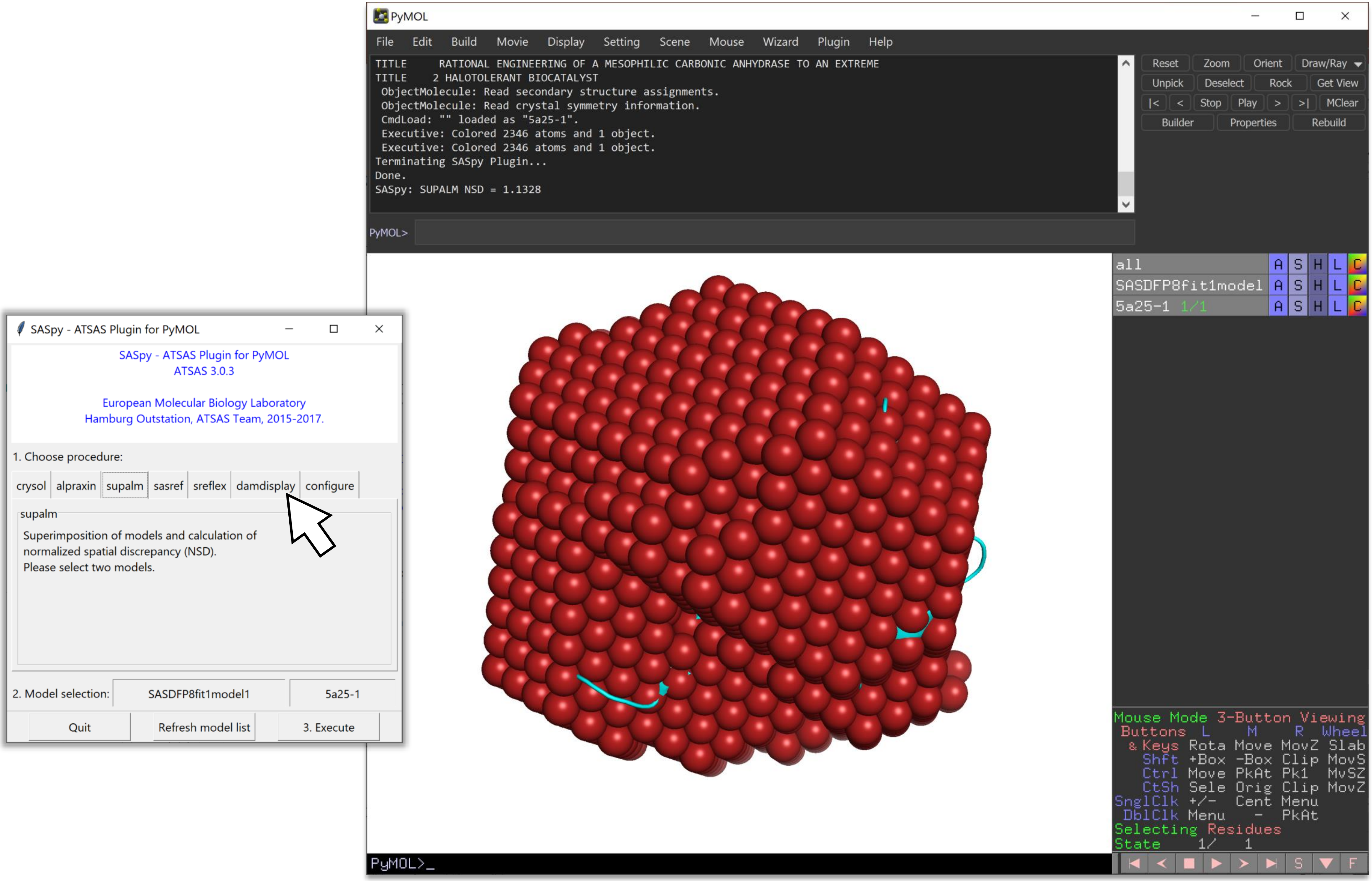
European Molecular Biology Laboratory
Hamburg Outstation, ATSAS Team, 2015-2017.

1. Choose procedure: supalm

supalm

Superimposition of models and calculation of normalized spatial discrepancy (NSD).
Please select two models.

2. Model selection: SASDFP8fit1model1 5a25-1



```
TITLE     RATIONAL ENGINEERING OF A MESOPHILIC CARBONIC ANHYDRASE TO AN EXTREME
TITLE    2 HALOTOLERANT BIOCATALYST
 ObjectMolecule: Read secondary structure assignments.
 ObjectMolecule: Read crystal symmetry information.
 CmdLoad: "" loaded as "5a25-1".
 Executive: Colored 2346 atoms and 1 object.
 Executive: Colored 2346 atoms and 1 object.
Terminating SASpy Plugin...
Done.
SASpy: SUPALM NSD = 1.1328
```

SASpy - ATSAS Plugin for PyMOL
ATSAS 3.0.3

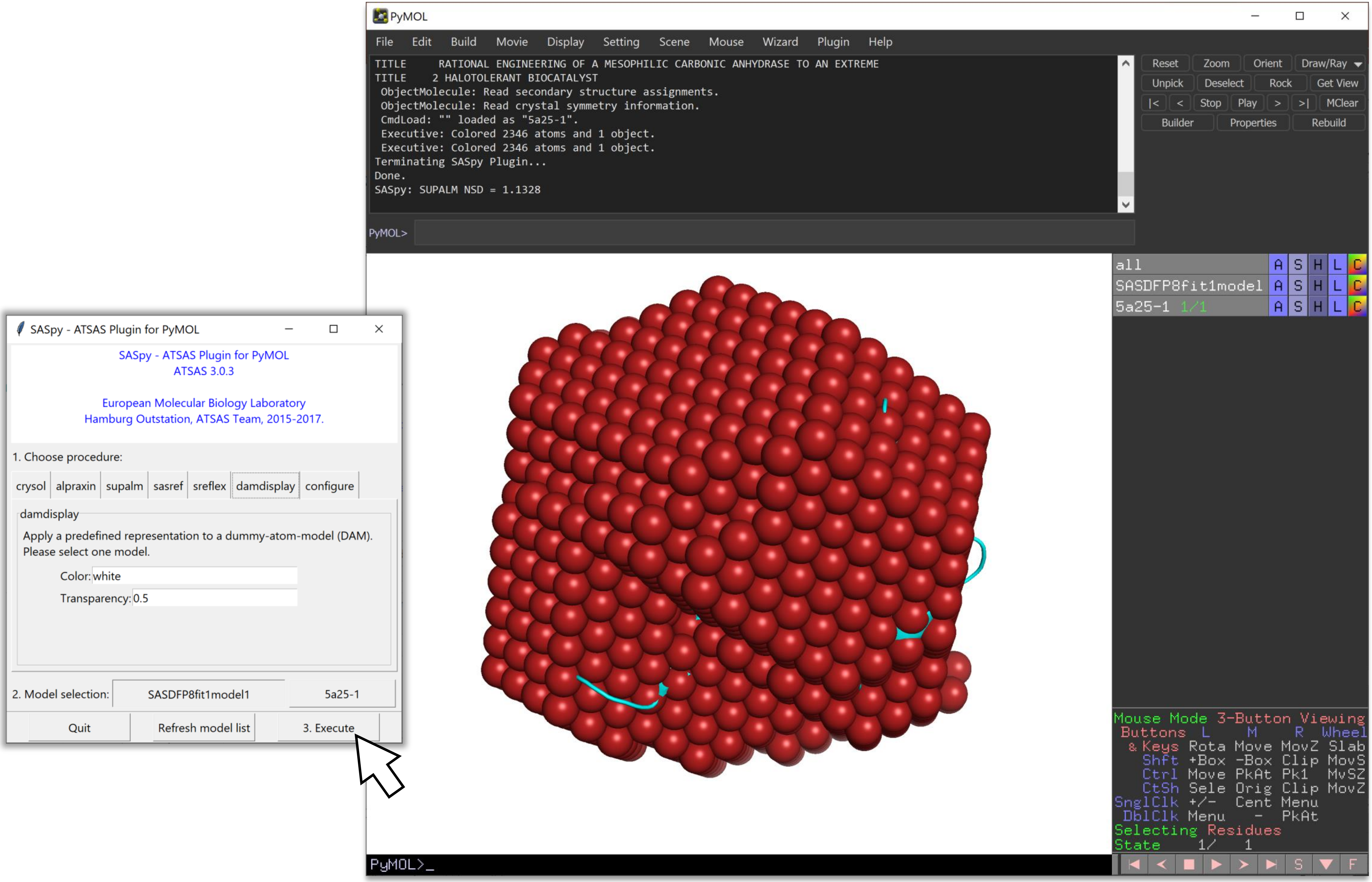
European Molecular Biology Laboratory
Hamburg Outstation, ATSAS Team, 2015-2017.

1. Choose procedure: supalm

supalm

Superimposition of models and calculation of
normalized spatial discrepancy (NSD).
Please select two models.

2. Model selection: SASDFP8fit1model1 5a25-1



```
TITLE     RATIONAL ENGINEERING OF A MESOPHILIC CARBONIC ANHYDRASE TO AN EXTREME
TITLE    2 HALOTOLERANT BIOCATALYST
 ObjectMolecule: Read secondary structure assignments.
 ObjectMolecule: Read crystal symmetry information.
 CmdLoad: "" loaded as "5a25-1".
 Executive: Colored 2346 atoms and 1 object.
 Executive: Colored 2346 atoms and 1 object.
Terminating SASpy Plugin...
Done.
SASpy: SUPALM NSD = 1.1328
```

SASpy - ATSAS Plugin for PyMOL

SASpy - ATSAS Plugin for PyMOL
ATSAS 3.0.3

European Molecular Biology Laboratory
Hamburg Outstation, ATSAS Team, 2015-2017.

1. Choose procedure:

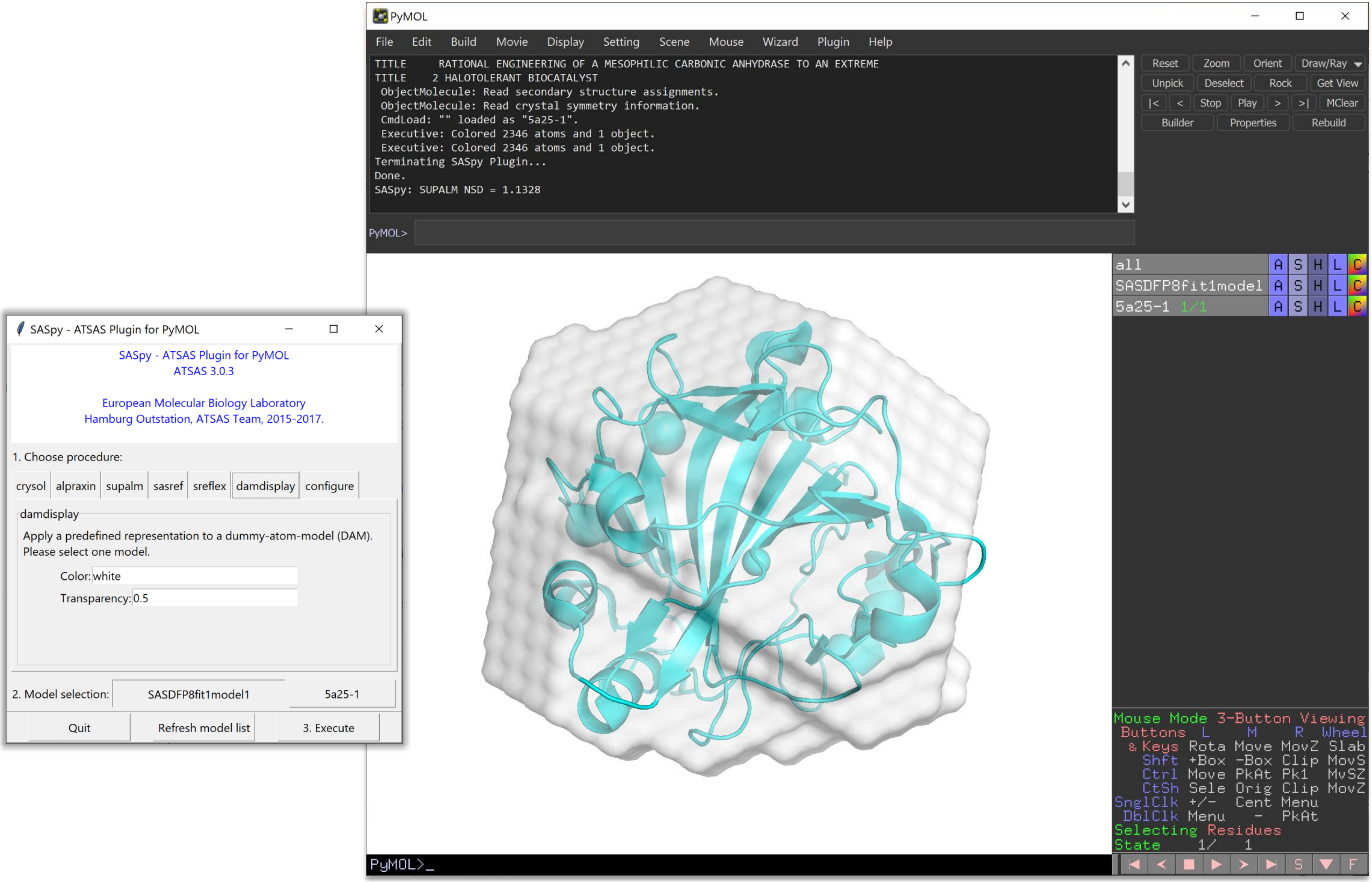
damdisplay

Apply a predefined representation to a dummy-atom-model (DAM).
Please select one model.

Color: white

Transparency: 0.5

2. Model selection: SASDFP8fit1model1



```
TITLE     RATIONAL ENGINEERING OF A MESOPHILIC CARBONIC ANHYDRASE TO AN EXTREME
TITLE    2 HALOTOLERANT BIOCATALYST
 ObjectMolecule: Read secondary structure assignments.
 ObjectMolecule: Read crystal symmetry information.
 CmdLoad: "" loaded as "5a25-1".
 Executive: Colored 2346 atoms and 1 object.
 Executive: Colored 2346 atoms and 1 object.
Terminating SASpy Plugin...
Done.
SASpy: SUPALM NSD = 1.1328
```

SASpy - ATSAS Plugin for PyMOL
ATSAS 3.0.3

European Molecular Biology Laboratory
Hamburg Outstation, ATSAS Team, 2015-2017.

1. Choose procedure:

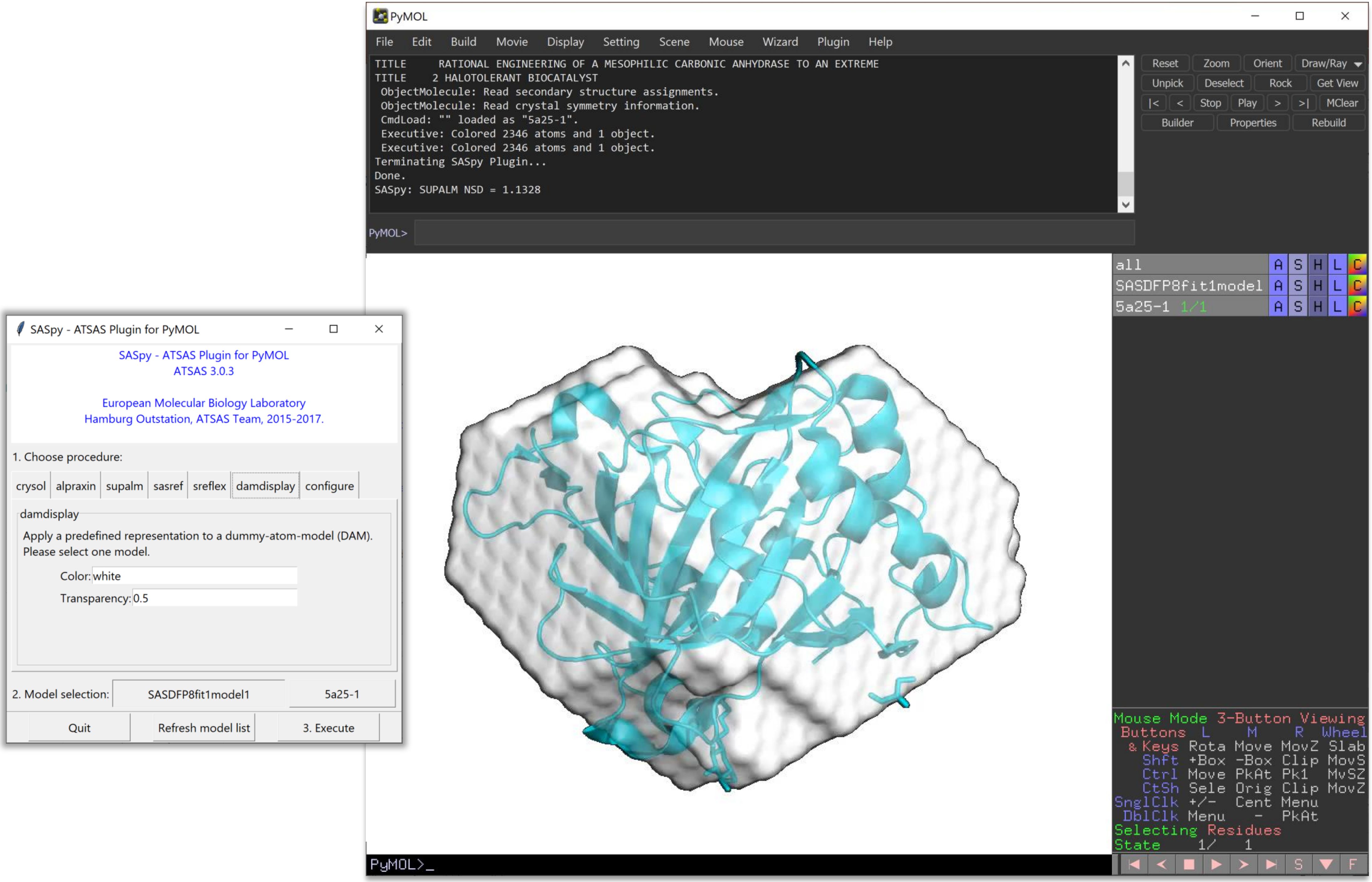
damdisplay

Apply a predefined representation to a dummy-atom-model (DAM).
Please select one model.

Color: white

Transparency: 0.5

2. Model selection:



PyMOL

```
TITLE     RATIONAL ENGINEERING OF A MESOPHILIC CARBONIC ANHYDRASE TO AN EXTREME
TITLE    2 HALOTOLERANT BIOCATALYST
 ObjectMolecule: Read secondary structure assignments.
 ObjectMolecule: Read crystal symmetry information.
 CmdLoad: "" loaded as "5a25-1".
 Executive: Colored 2346 atoms and 1 object.
 Executive: Colored 2346 atoms and 1 object.
Terminating SASpy Plugin...
Done.
SASpy: SUPALM NSD = 1.1328
```

SASpy - ATSAS Plugin for PyMOL

SASpy - ATSAS Plugin for PyMOL
ATSAS 3.0.3

European Molecular Biology Laboratory
Hamburg Outstation, ATSAS Team, 2015-2017.

1. Choose procedure:

damdisplay

Apply a predefined representation to a dummy-atom-model (DAM).
Please select one model.

Color: white

Transparency: 0.5

2. Model selection: SASDFP8fit1model1 | 5a25-1

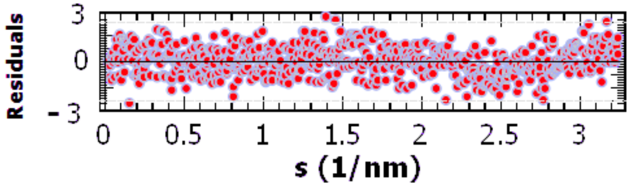
Metrics for fit quality

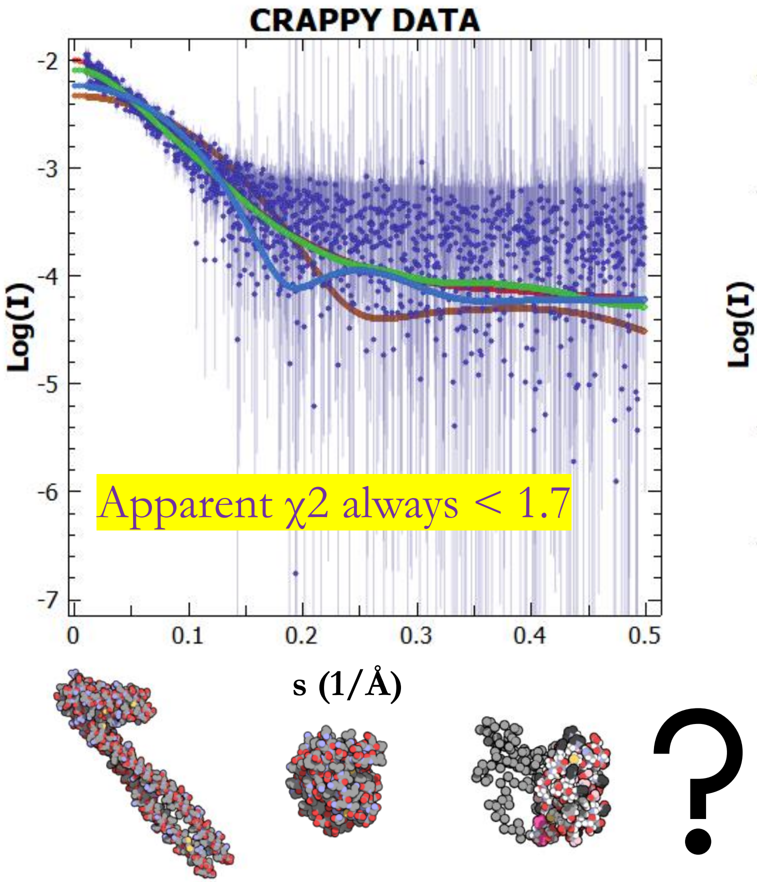
Most commonly used: χ2 (but…sensitive to I error estimate)
LOOK to the data (noise level/point dispersion, regions with larger disagreement/residuals)
Consider proper data range

$$\chi^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{I_{exp}(s_i) - I_{calc}(s_i)}{\sigma_i}\right)^2$$

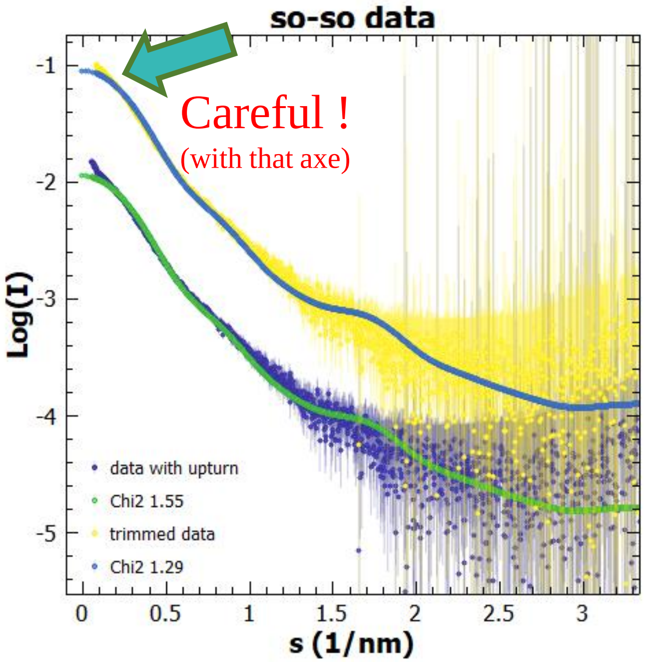
Cormap – p

Vr, χ^2_{free}





Apparent χ2 always < 1.7



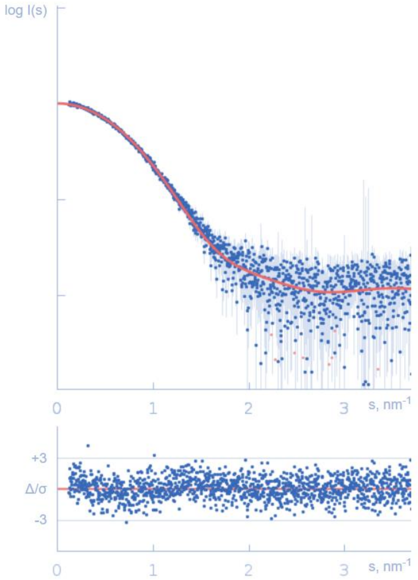
Careful !
(with that axe)

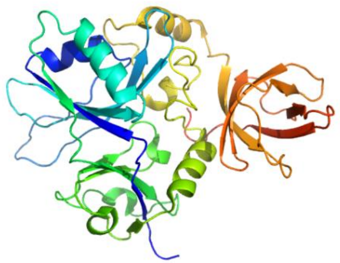
Smin: typically first Guinier point

Terminology: *ab initio models, fitting*

- Fit to the data
  - Agreement or discrepancy

    Between *ab initio or molecular model and the SAXS data*

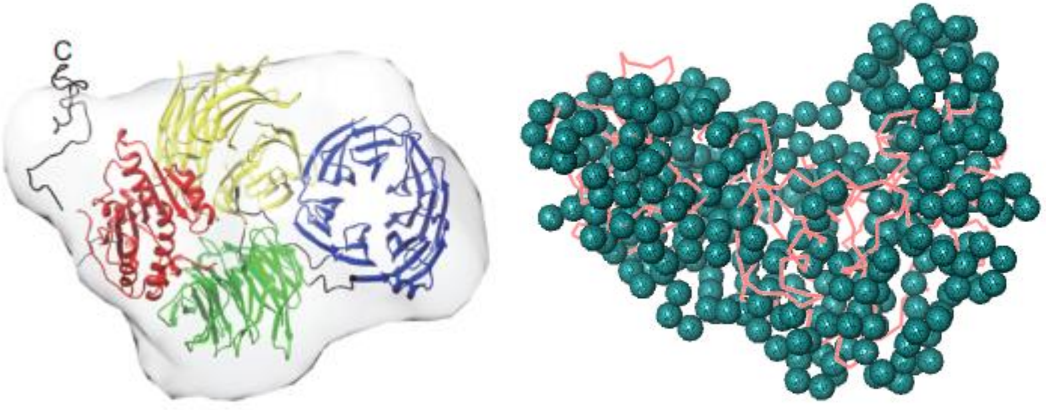




≠

- Overlap to the *ab initio* reconstruction
  - *Visual* comparison between models

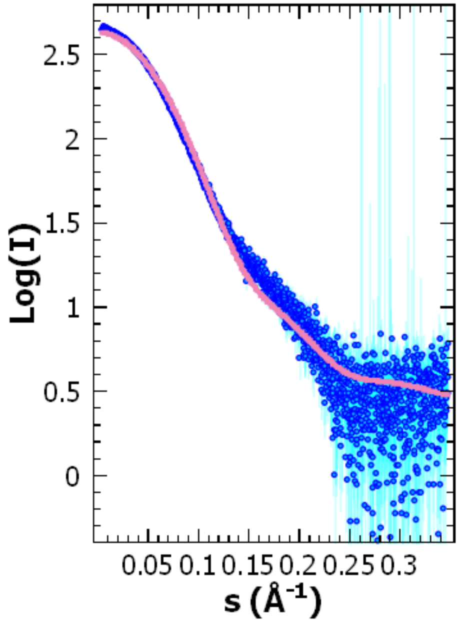
    maximizing the overlap (not a real fit)

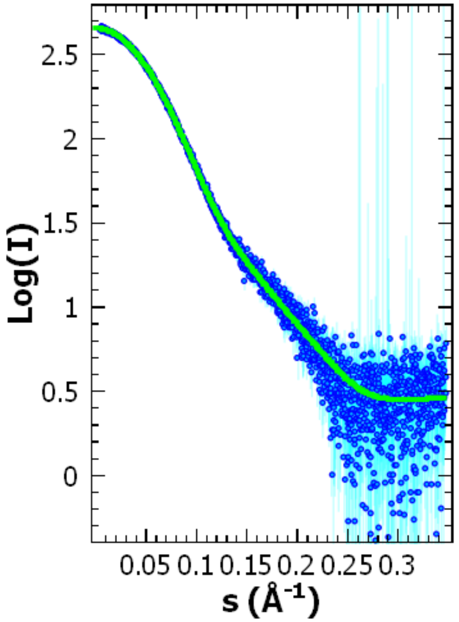


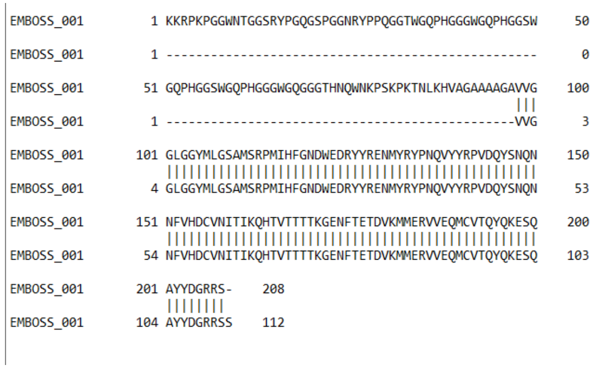
Not fitting ?

That’s good information too !

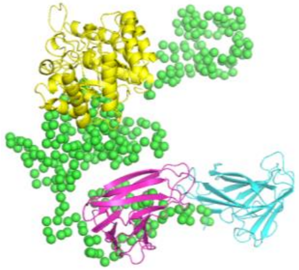
- Consider data (range, structure factor, subtraction)
- Consider structure completeness
- Consider association state(s)

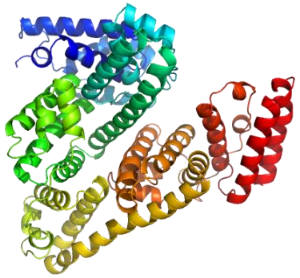


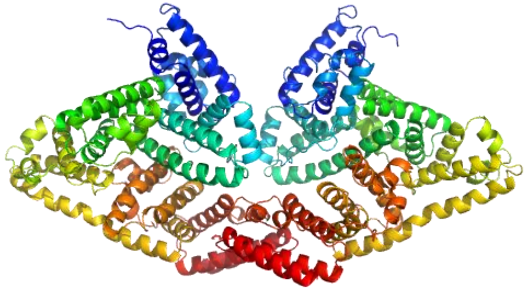




```
EMBOSS_001         1 KKRPKPGGWNTGGSRYPGQGSPGGNRYPPQGGTWGQPHGGGWGQPHGGSW     50

EMBOSS_001         1 --------------------------------------------------      0

EMBOSS_001        51 GQPHGGSWGQPHGGGWGQGGGTHNQWNKPSKPKTNLKHVAGAAAAGAVVG    100
                                                                    |||
EMBOSS_001         1 -----------------------------------------------VVG      3

EMBOSS_001       101 GLGGYMLGSAMSRPMIHFGNDWEDRYYRENMYRYPNQVYYRPVDQYSNQN    150
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
EMBOSS_001         4 GLGGYMLGSAMSRPMIHFGNDWEDRYYRENMYRYPNQVYYRPVDQYSNQN     53

EMBOSS_001       151 NFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERVVEQMCVTQYQKESQ    200
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
EMBOSS_001        54 NFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERVVEQMCVTQYQKESQ    103

EMBOSS_001       201 AYYDGRRS-    208
                     ||||||||
EMBOSS_001       104 AYYDGRRSS    112
```



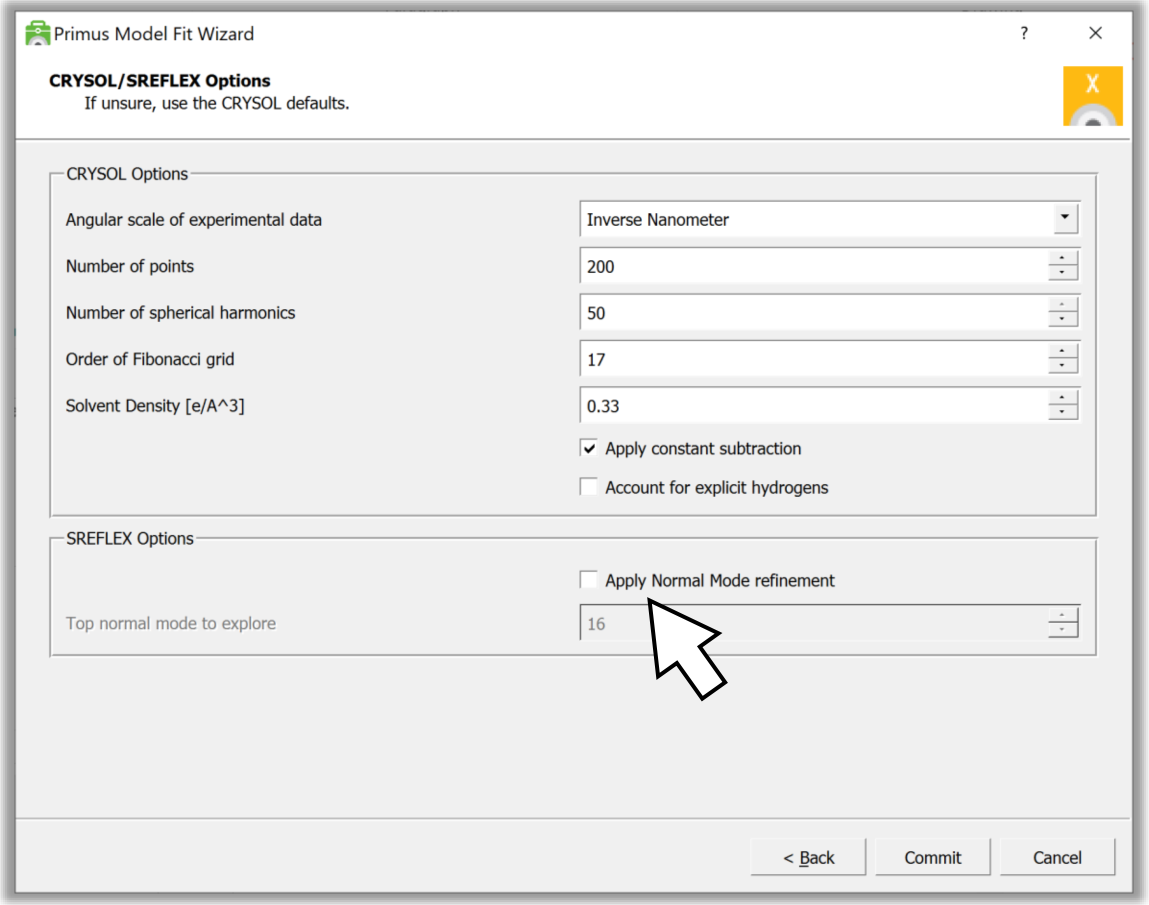




Not fitting and suspecting conformational change ?

- Crystal contacts
- Structure only in complex with interaction partner
- Etc.

SREFLEX: flexible refinement



CRYSOL/SREFLEX Options

If unsure, use the CRYSOL defaults.

CRYSOL Options

Angular scale of experimental data: Inverse Nanometer

Number of points: 200

Number of spherical harmonics: 50

Order of Fibonacci grid: 17

Solvent Density [e/A^3]: 0.33

☑ Apply constant subtraction

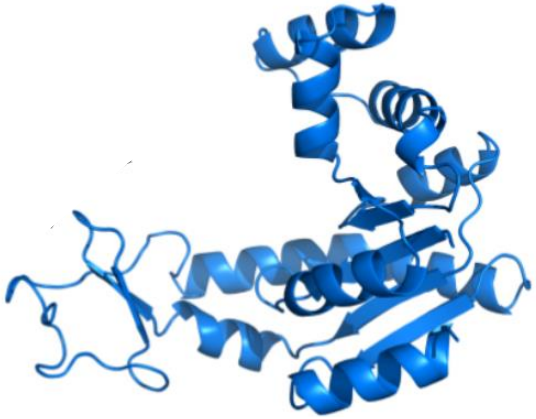
☐ Account for explicit hydrogens

SREFLEX Options

☐ Apply Normal Mode refinement

Top normal mode to explore: 16

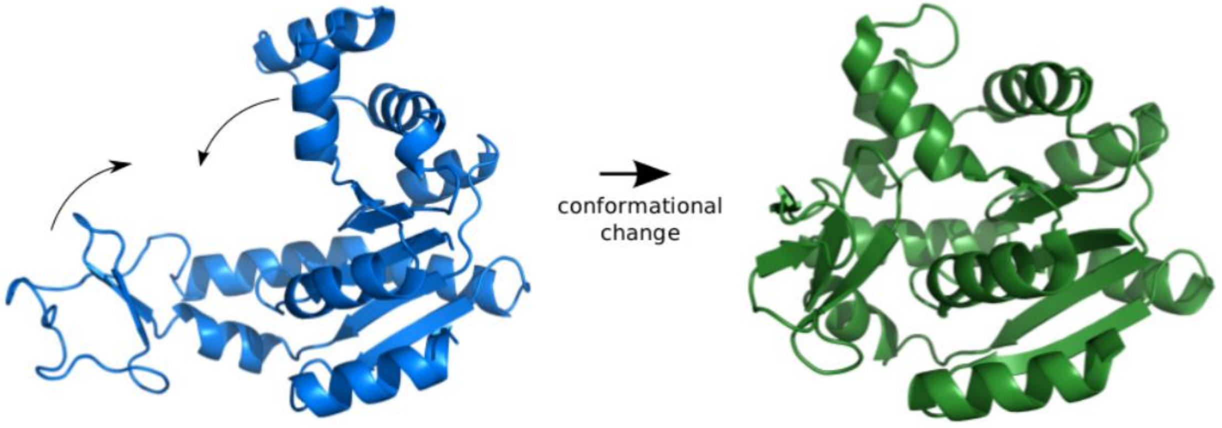
SREFLEX: refinement through flexibility



4ake

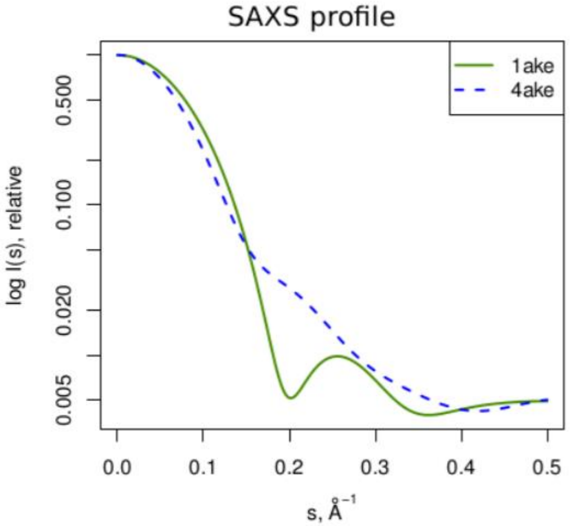
1ake

SREFLEX: refinement through flexibility



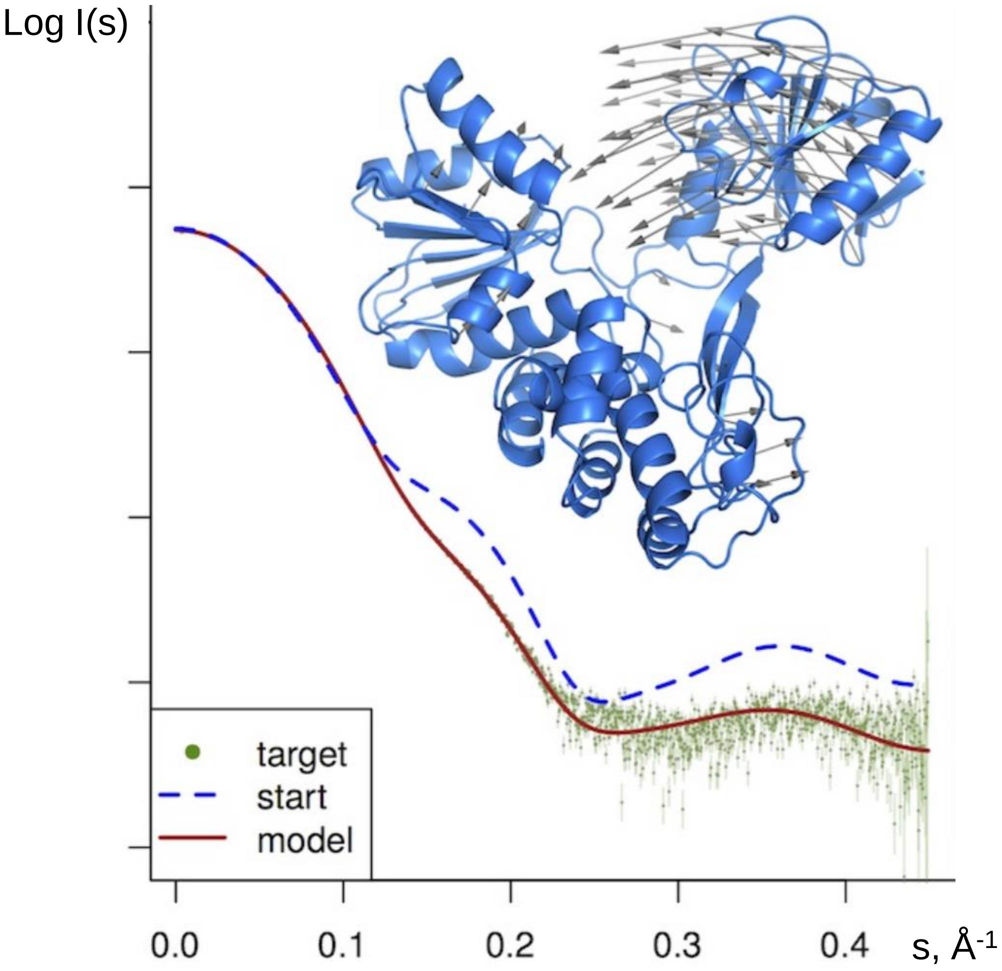
4ake

1ake



SREFLEX

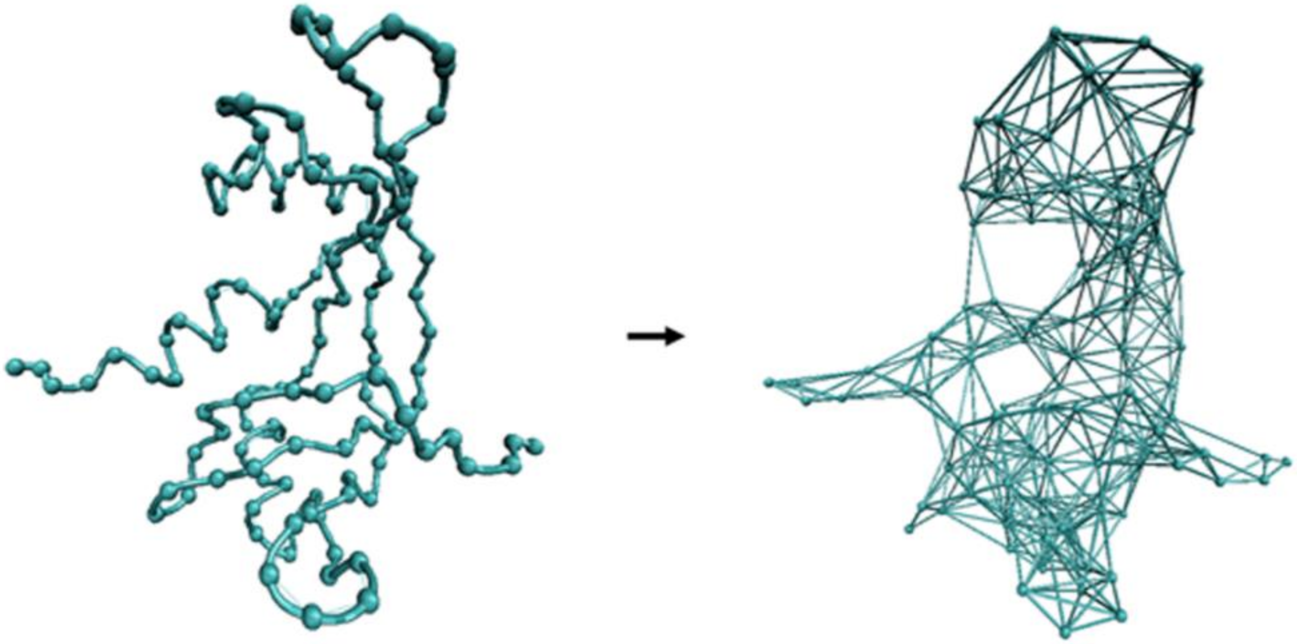
SAS REfinement through **FLEX**ibility based on normal mode analysis



SREFLEX: Panjkovich A. and Svergun D.I. (2016) *Phys. Chem. Chem. Phys.* 18, 5707-5719

ASREFLEX

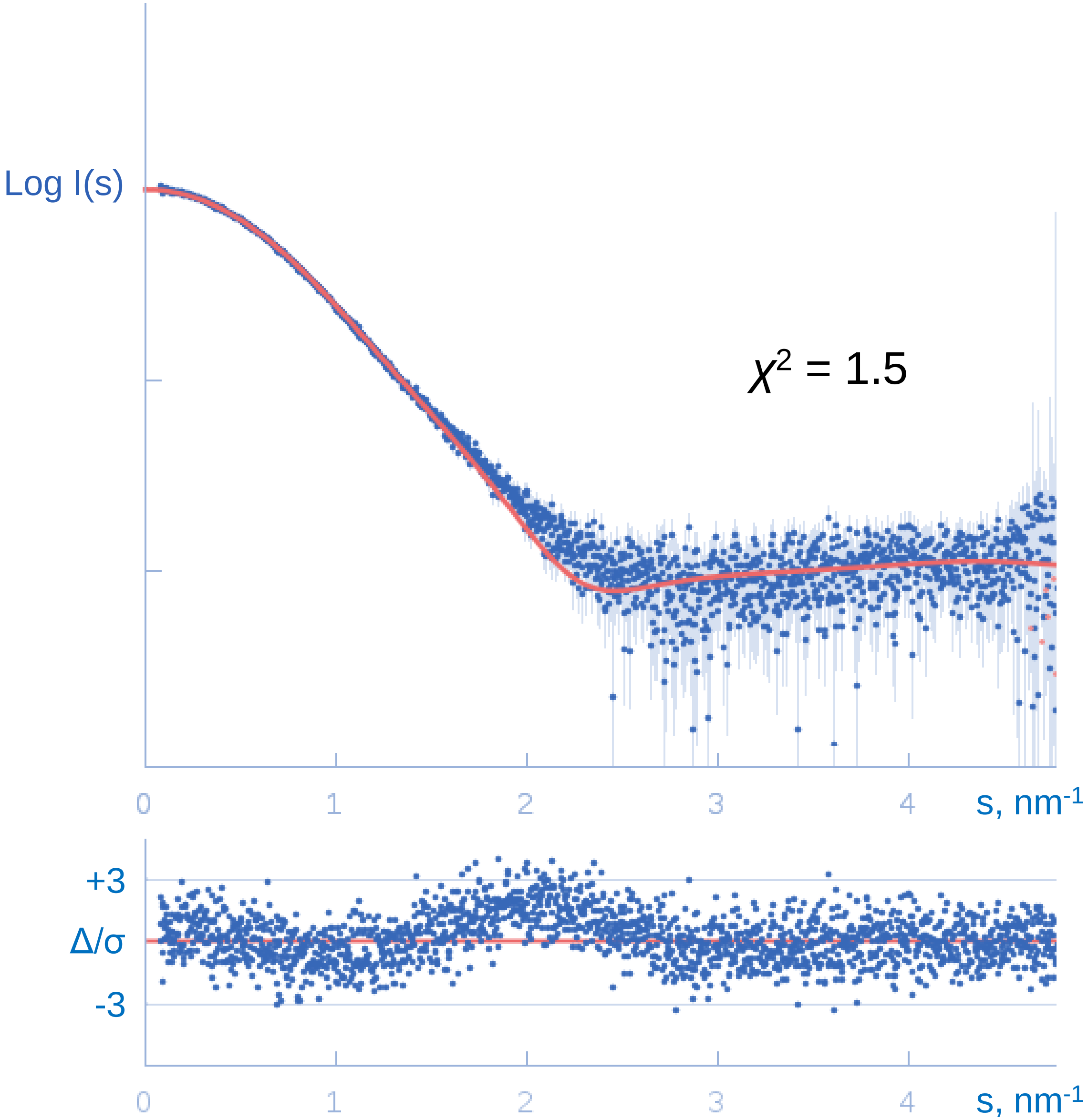
Estimating protein flexibility: normal mode analysis (NMA)

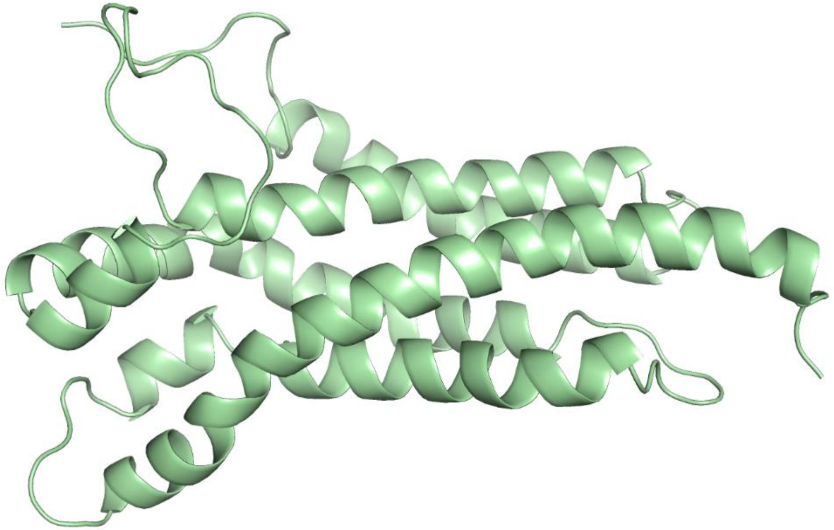


Cα trace elastic network model

Delarue & Sanejouand (2002) Simplified NMA of conformational transitions in DNA-dependent polymerases: the elastic network model. *J Mol Biol* 320:1011-1024

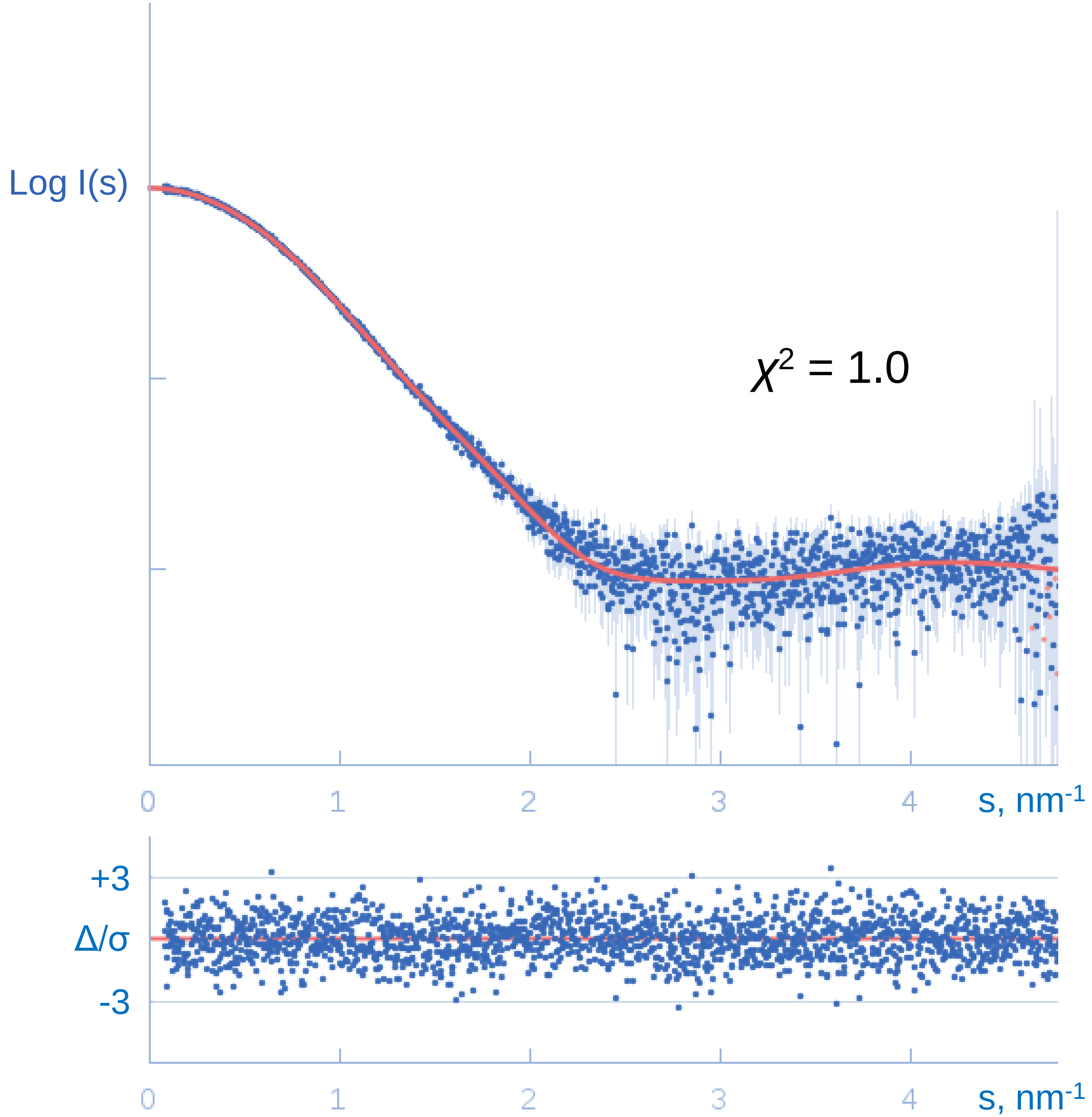
SREFLEX: refinement through flexibility

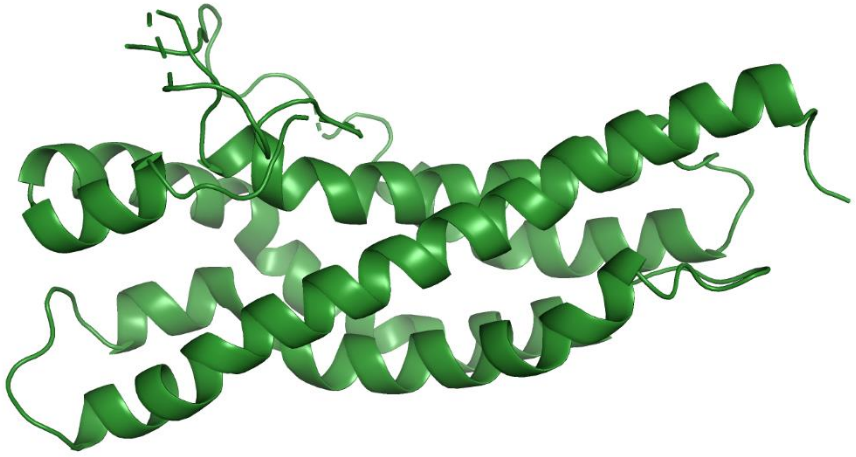




SASDC36 – Structural and functional dissection of the DH and PH domains of oncogenic Bcr-Abl tyrosine kinase

SREFLEX: refinement through flexibility





SASDC36 – Structural and functional dissection of the DH and PH domains of oncogenic Bcr-Abl tyrosine kinase

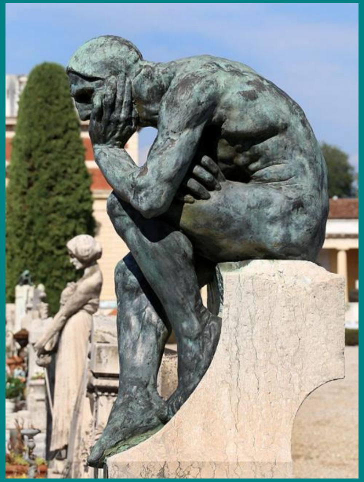
- HINTS:

  Click your way around by trial and error

  Look inside the files (open as text files .dat, .fir, .out etc.) !

  Learn to use CLI

  Look at fits visually and think (larger? smaller ? etc.)



BIOSAXS

The original solution

Thanks for your attention: keep exploring the programs….

sdavela@embl-hamburg.de

BIOSAXS

The original solution