TD RéNaFoBiS 2018
SAXS

Dominique Housset

Aurélien Thureau

5/06/2018

Instructions

1) Using Primus, open the file:

`01_Mast/mast2_avg.dat`

- ☞ Is the protein folded ?
- ☞ What is the Rg value ?
- ☞ What is the MW of the sample?
- ☞ What is the Dmax value ?
- ☞ Generate a bead models.
- ☞ Is the NMR structure is compatible ?
- ☞ Can we do more ?

2) Using Primus, open the file:

`02_Pdz/p1p2_cut.dat`

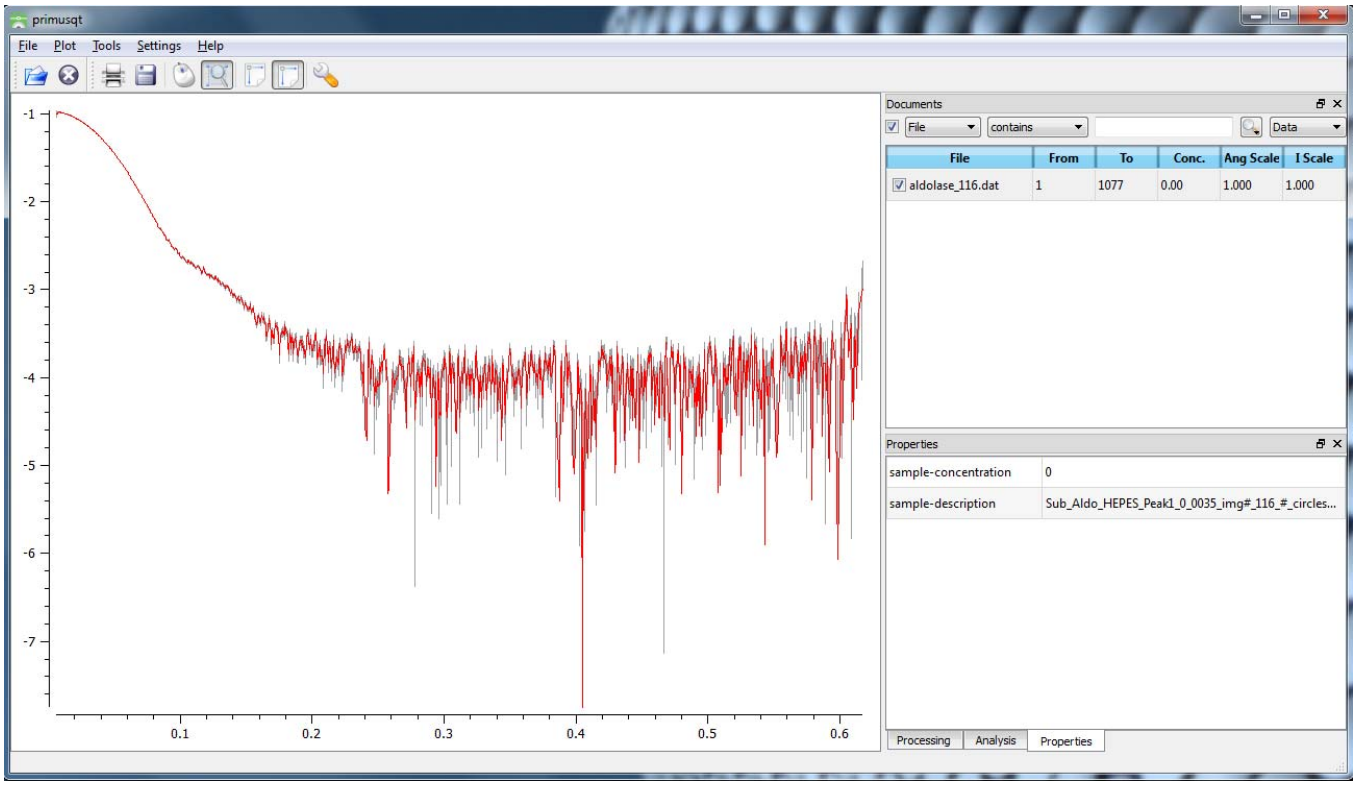
- ☞ What is the the MW of the sample?
- ☞ Is the Xray structure is compatible ?
- ☞ Look at the PDB file (pymol)
- ☞ Can we do more ?

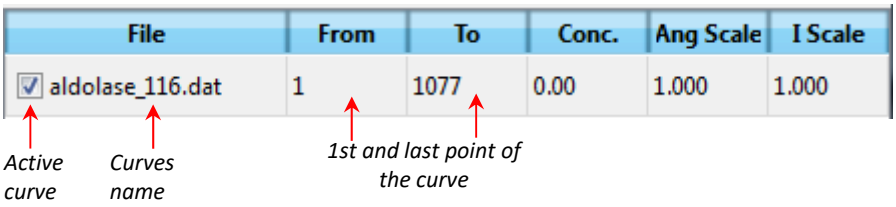
3) More difficult files:

`03_OthersExamples/`

Primus

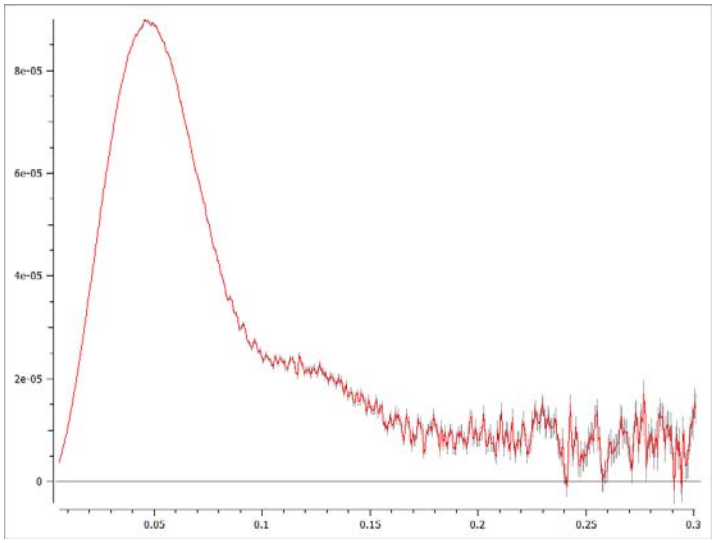
This program is used to display curves and apply some operation on the curves (scaling, substraction or merging…). It includes also a lot of script to estimate the Radius of Gyration, the Distance Distribution, the Porod Volume and the Molecular Weight.



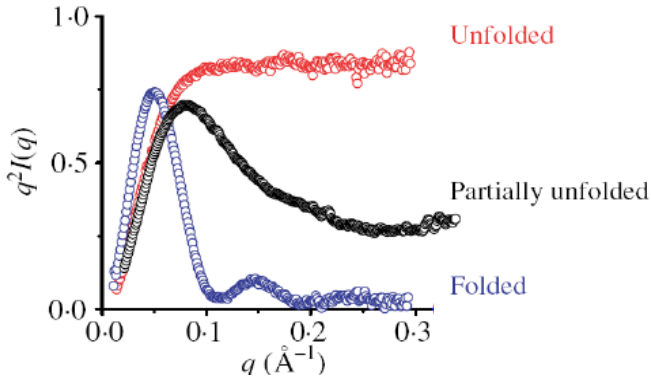


You can load the SAXS curves (click on “File” button). When the data are loaded in the table, the name appears in the “Documents” area.
With the buttons “From” and “To” you can remove respectively the first points or the last points of the curve. It’s useful to hide the noisy part of the curve (last points).

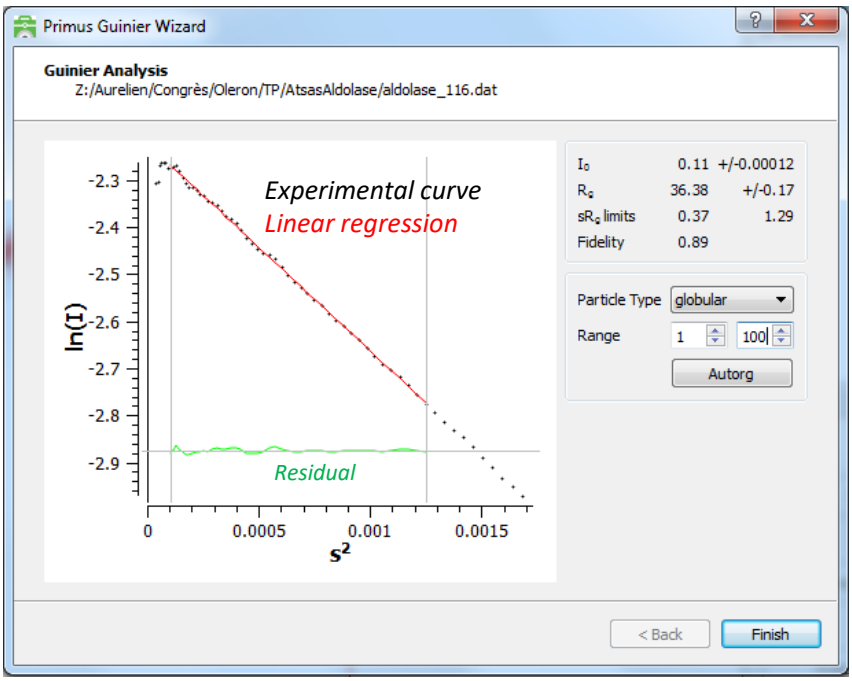
Primus



For Kratky representation, you can go to **Plot -> I*s² vs. s (Kratky plot).**
Only the first points are necessary (q between 0 and 0.3)



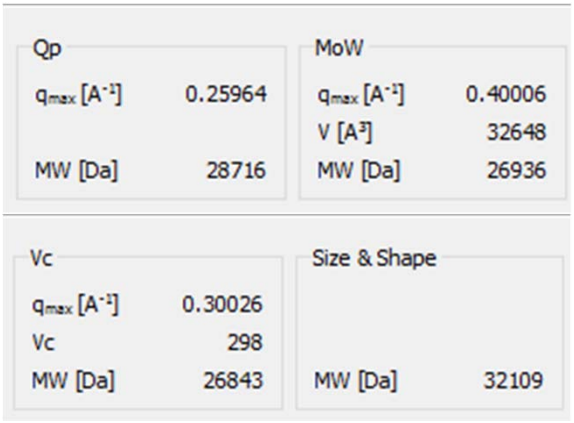
For Rg calculation, you can go to **Tools -> Analysis -> Radius of Gyration** and modify manually the parameters or use directly the program "AutoRg".



You can evaluate the quality of the fit with the residual plot which represents the distribution of the experimental points around the regression line. You can control also the limit s*Rg max which should be inferior to 1. This limit depends of the geometry of the object : For globular protein, the limit can be increased up to 1.3 whereas for elongated or unfolded protein, the Guinier region is more restricted (less than 1). The first points which correspond to the lowest values of q can be also removed due to large error measurement (data close to the beam stop).

Molecular Weight

For Molecular Weight estimation, you can go to **Tools -> Analysis -> Molecular Weight.**
A Guinier Analysis is necessary for the First Step.

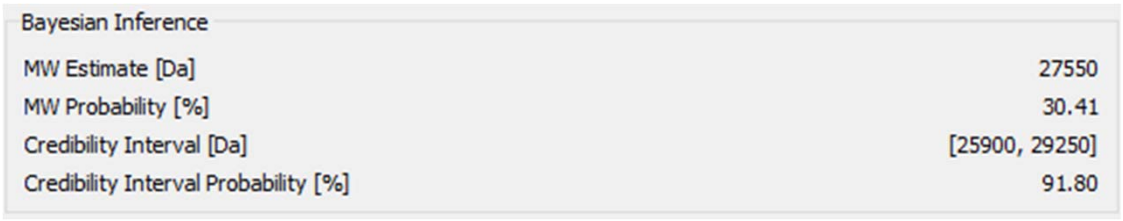


Qp : Based of Porod invariant (only folded protein)

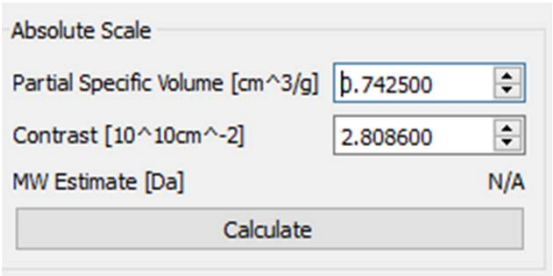
MoW : Integration of $I(q).q^2=f(q)$

Vc : Integration of $I(q).q=f(q)$ + Rg

Size&Shape : From Rg estimation



Statistic calculation to give an interval of the estimated MW

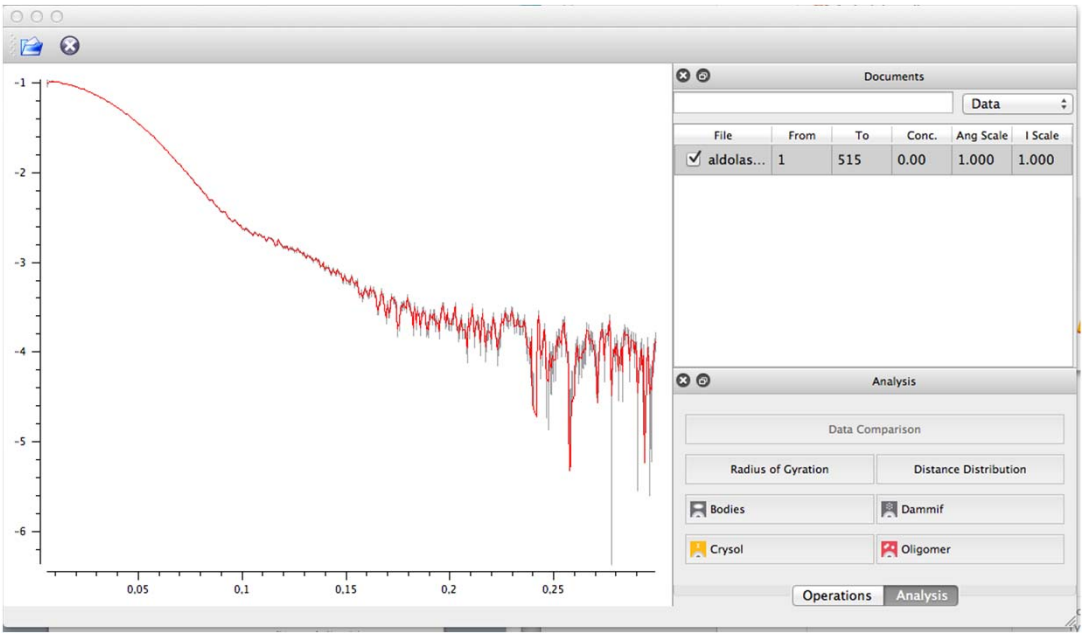


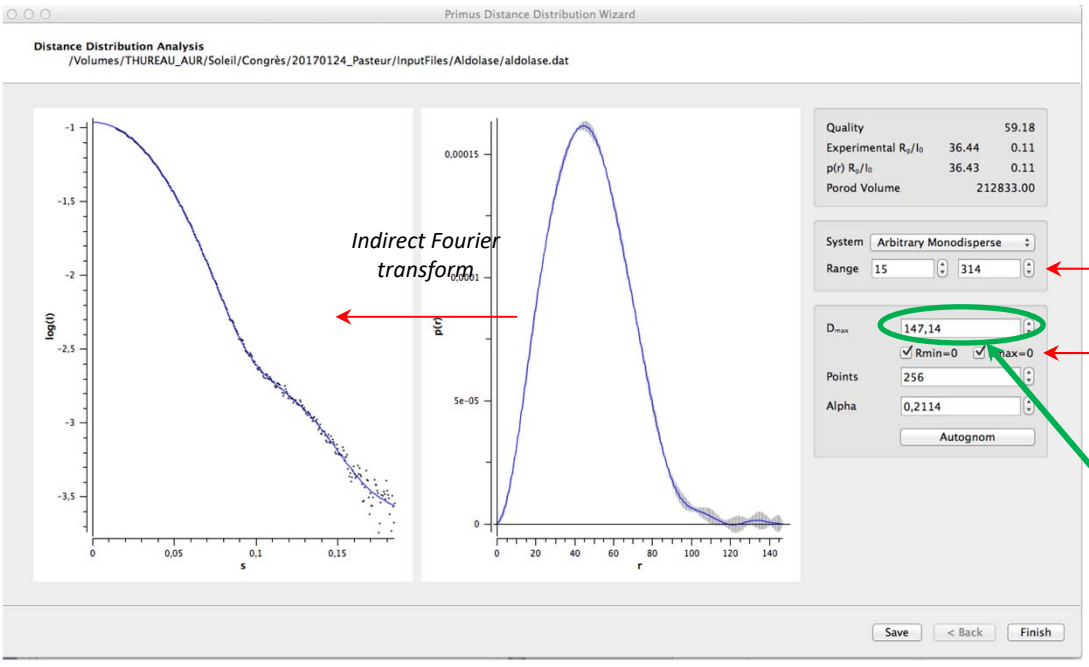
You need a normalized curve in absolute scale extrapolated to a concentration at 1mg/mL.

Partial Specific Volume and Contrast can be calculated from the 1D sequence

Distance Distribution

The program GNOM is used to determine the autocorrelation function p(r) from the SAXS data and to estimate the Dmax.
Using Primus, click on [Distance Distribution]





First and last points can be removed (noisy)

Limits condition (try to change)

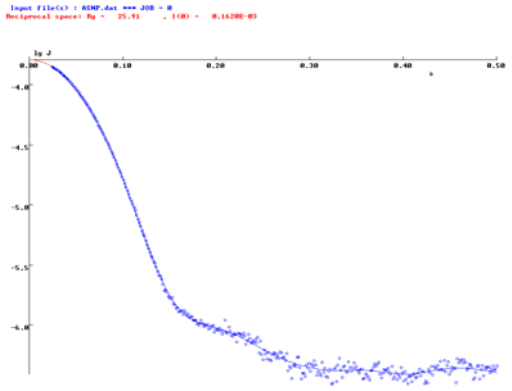
Dmax (to be evaluated) : start with 5*Rg

At the end, you will obtain a new file called <*.out> containing all informations about the p(r) determination such as parameters defined preliminary in GNOM (nBeg, nEnd, …), biophysical parameters (Rg, I(0) and Dmax). You will find also in ASCII format, the SAXS curve with the corresponding regularization curve and the p(r) function.
The file <*.out> will be used by ab initio program such as DAMMIF, DAMMIN or GASBOR.

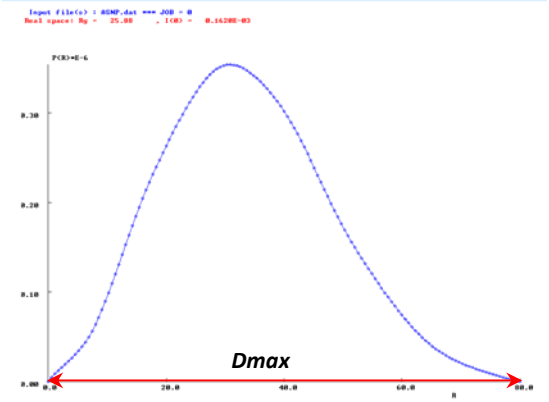
Distance Distribution

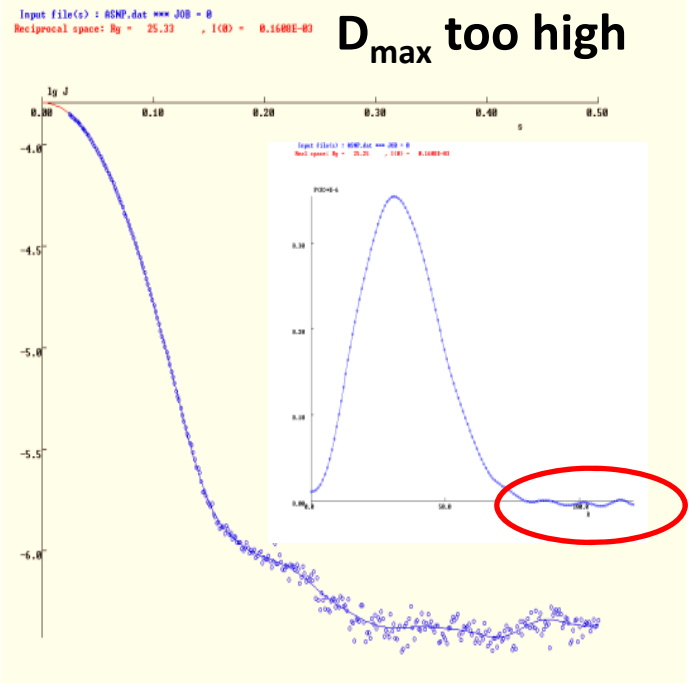
To determine a correct value of Dmax, we must proceed by trial and error to find a Rg calculated with GNOM similar to that found with the calculation of Guinier. We start in general with a value of Dmax equal to 4 or 5 times the value of Rg, and decrease gradually the value of the Dmax in order to obtain a smoothed p(r) that cuts the axis of the distance r.

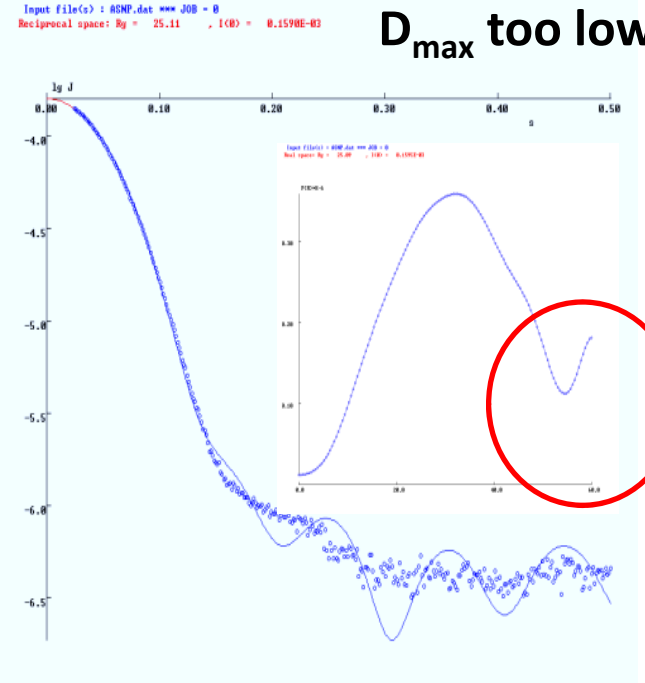
Regularized curve



Autocorrelation function p(r)

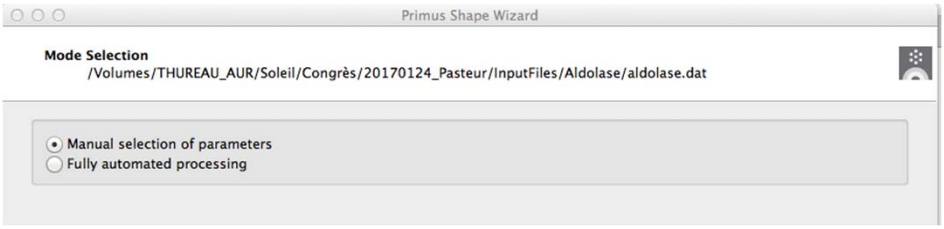




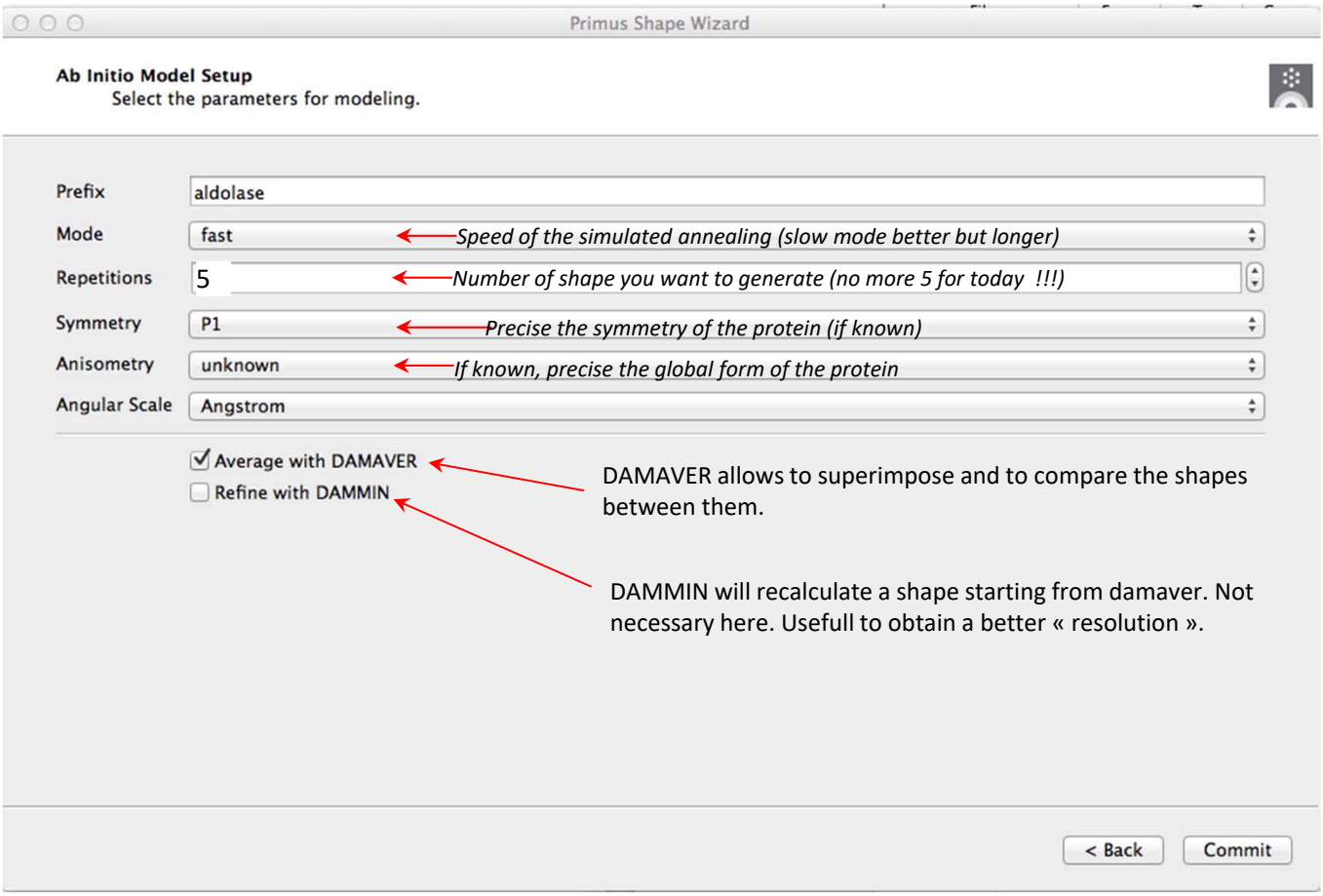


Dammif

The program DAMMIF is dedicated to low resolution shape modeling using a sphere containing beads (with a defined value of electronic density) as initial model.

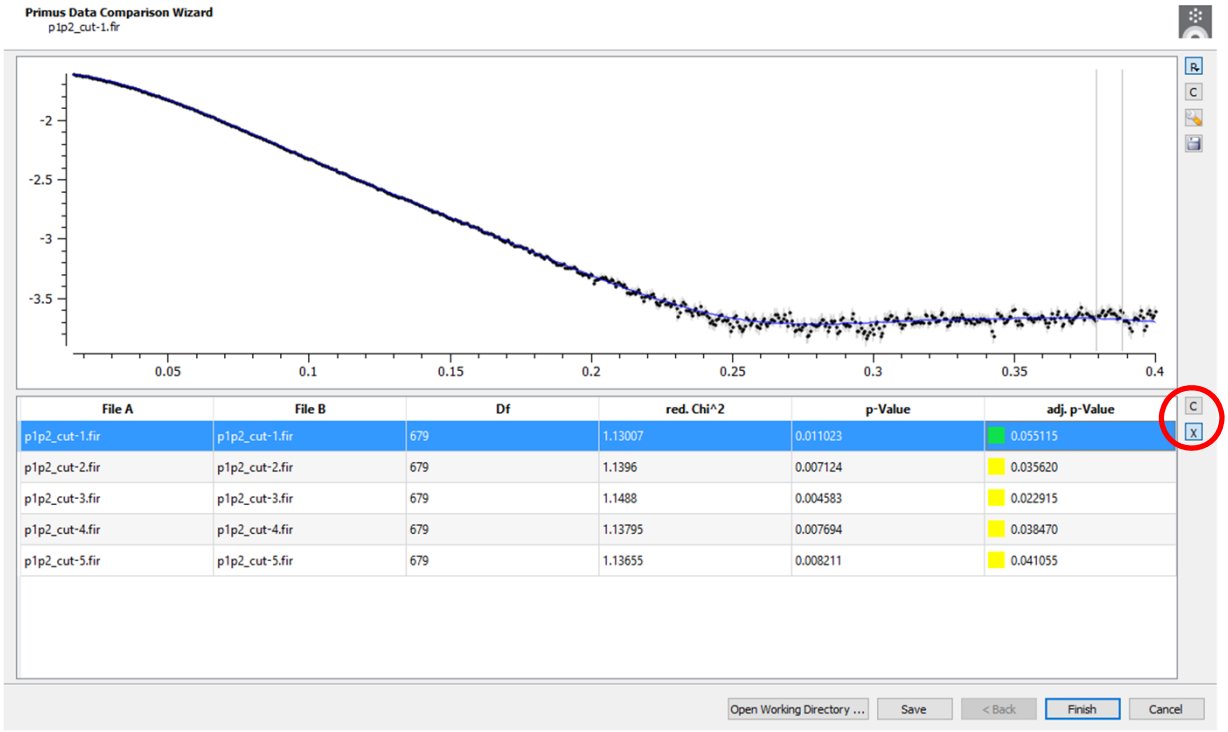


Rg and P(r) are necessary to generate a bead model.



Be patient, it will take few minutes/hours depending of your computer.
Let's have a break !!!

Dammif



Cormap Test
X^2 Test

For each repetition asked, a fit is proposed:

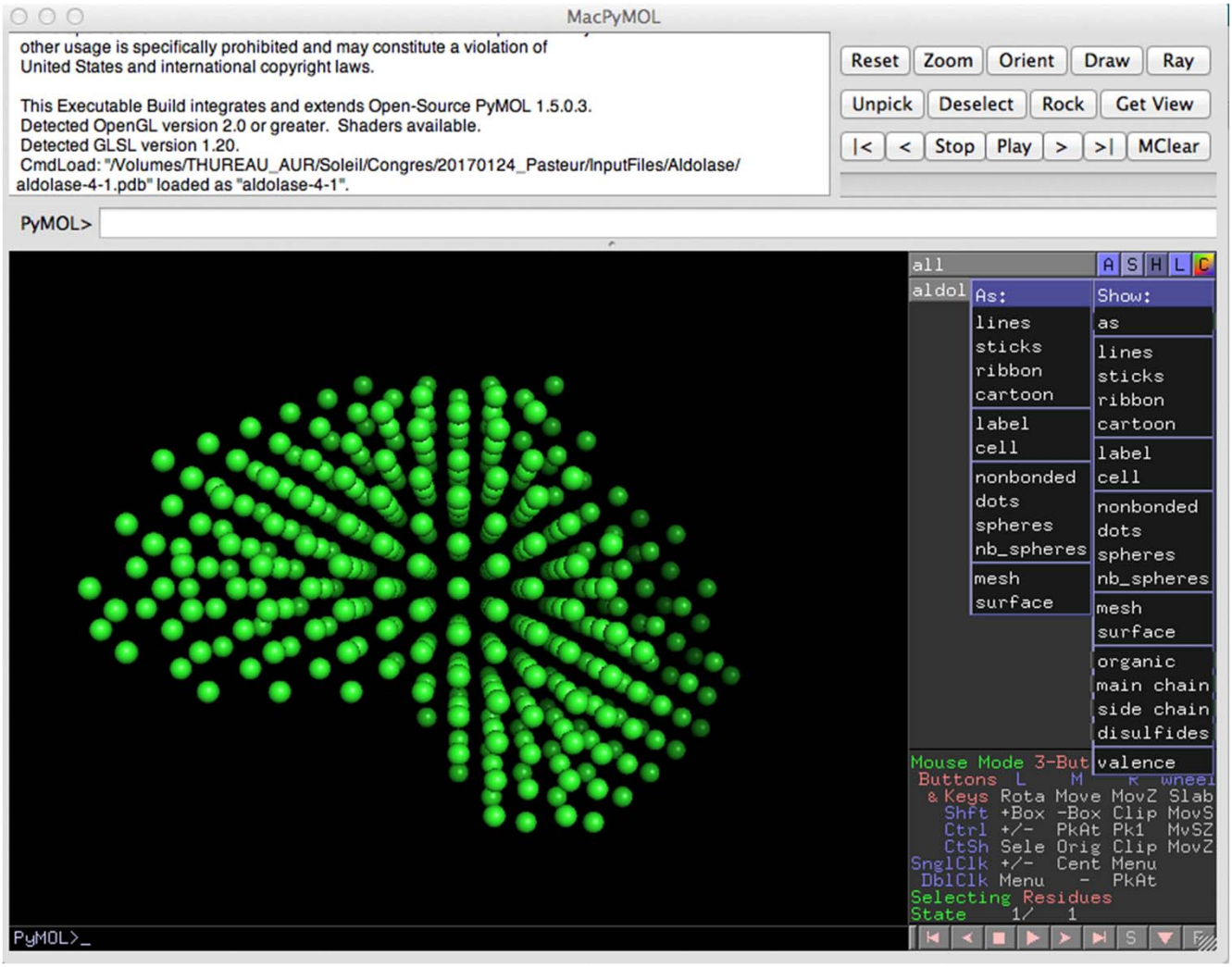
1) The statistics of Cormap Test have to be correct (green or yellow), otherwise it means that "*The hypothesis of similarity of experimental data and shape model could be rejected*"
2) The Chi square Test is also an other method to validate the models of shape.

By Saving the result:

1) You will obtain, a pdb file (-1.pdb) containing the shape composed of dummy residues, a fitting curve (smoothed curve) dam_xx.fit, a fitting curve corresponding to the experimental data dam_xx.fir and a log file with initial parameters and the process of minimization steps for each shape.
2) If you have perform an average with damaver, you will obtain also a table contained in the file damsel.log presents a matrix giving a non-deviation standard score (NSD) for each pair of shapes, and gives a classification of the shape. The most representative shape presents the lowest mean value of NSD. Be careful with the damfilt.pdb which represents a filtered shape corresponding to the common part off all shape. This model, in general case, does not fit the SAXS data !

3) You can open the pdb files (*-1.pdb) with pymol

Pymol



1) To have a nicer view of bead modelon pymol. Select Show -> As -> Spheres
2) Type the command line : `set sphere_scale, 2.5`

AND / OR

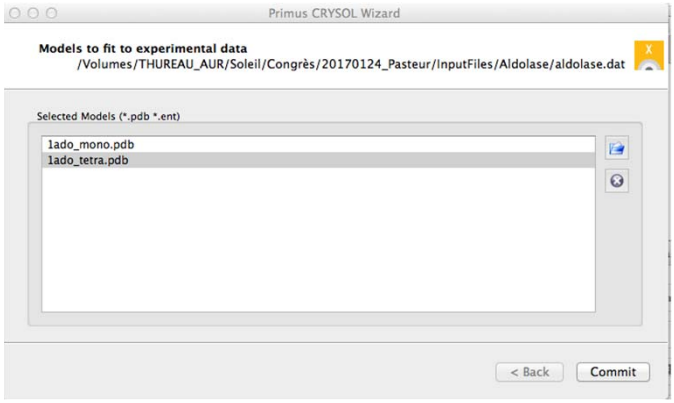
1) `alter <name of the object>,vdw=<value in the dammif log file at dummy atom radius>`
2) `set solvent_radius = 4.3`
3) Show -> As -> mesh

1) To represente the protein as cartoon: Show -> As -> Cartoon
2) To color some part of a protein (residue 6 to 107), type the command:

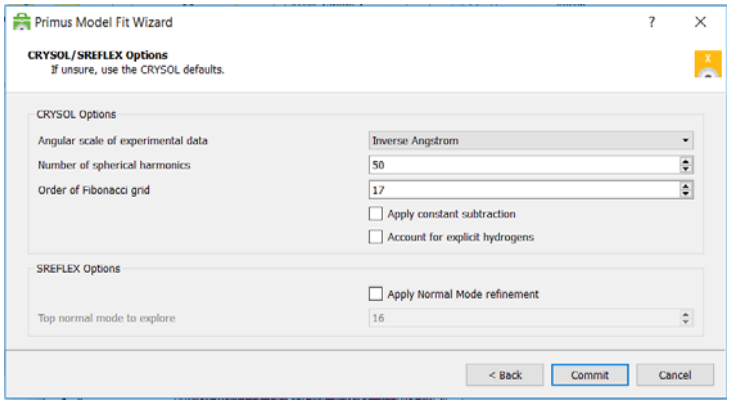
```
color red, resi 6-107
```

Crysol

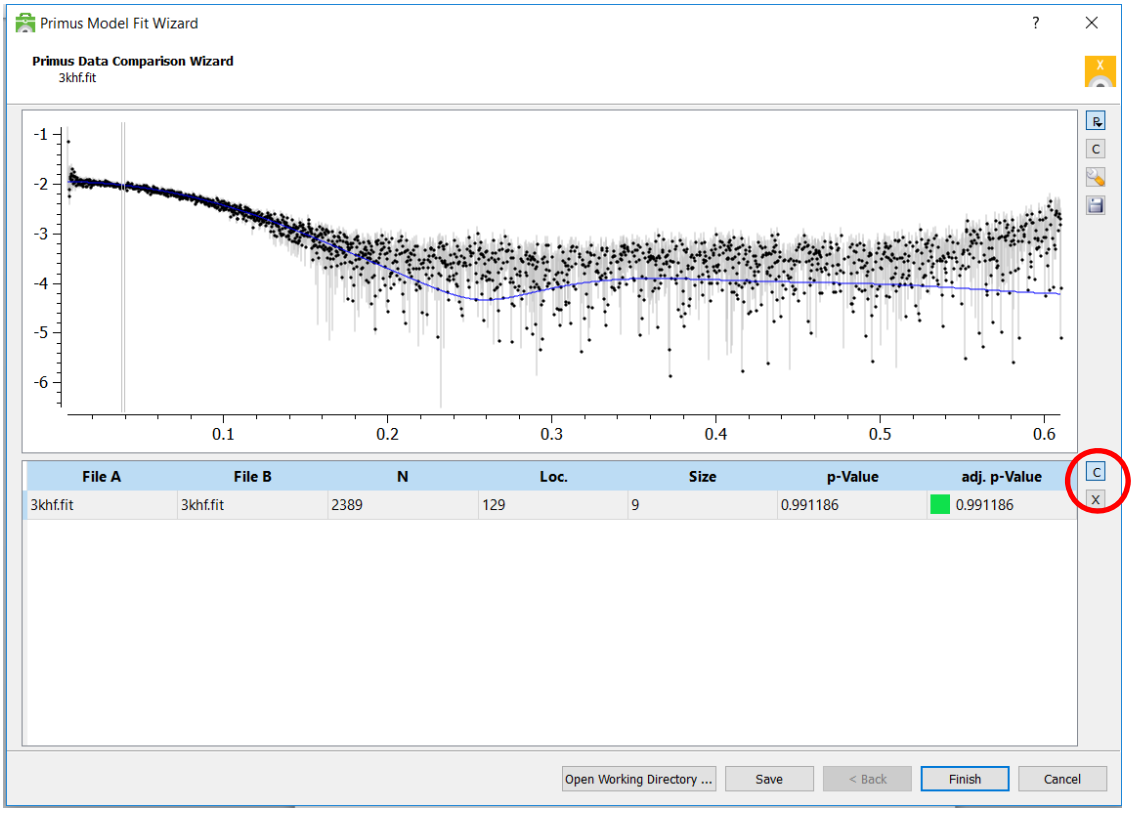
CRYSOL is a program for evaluating the solution scattering from macromolecules with known atomic structure and fitting it to experimental scattering curves from Small-Angle X-ray Scattering (SAXS). As an input one can use a PDB file with an X-ray or NMR structure of a protein or a protein-DNA(RNA) complex.



You can add as many pdb files as you want

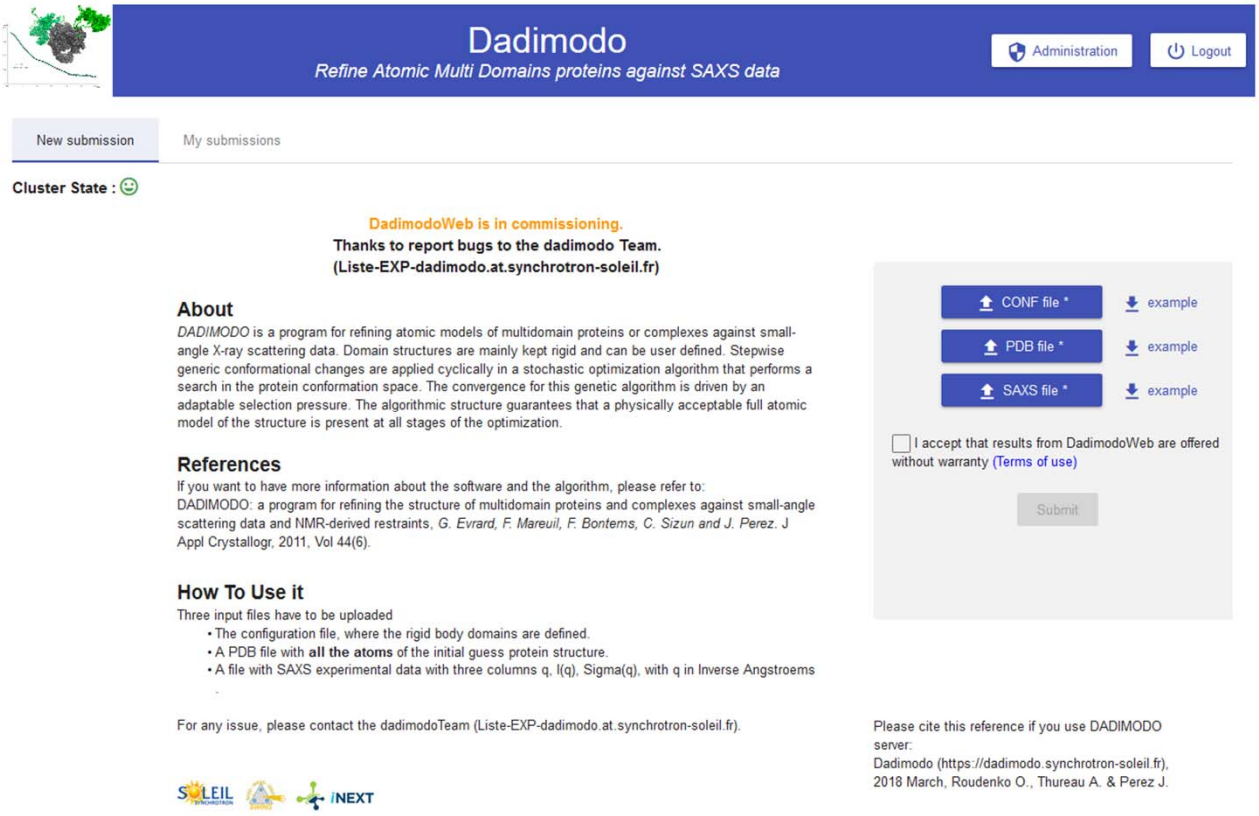


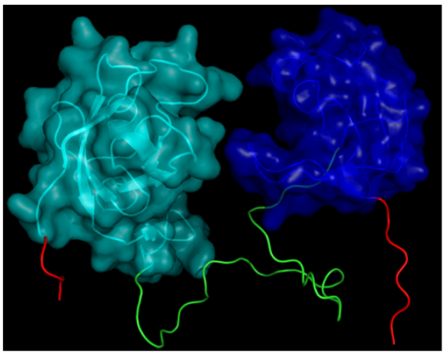
number of spherical harmonic determined by the relationship : ***Lmax = Qmax*Dmax/2 + 15***.
Usually, we set the value to 50 (max)



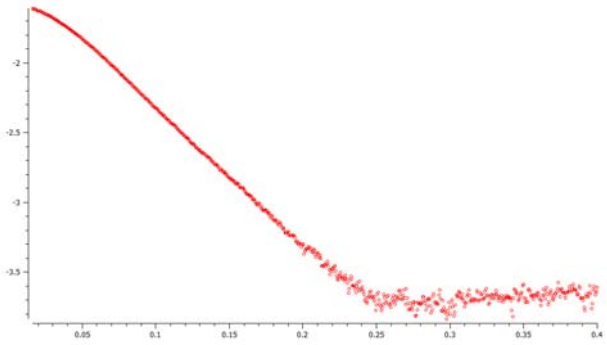
Cormap Test
X^2 Test

https://dadimodo.synchrotron-soleil.fr/





+



Pdb + rigid domains definition

Saxs curve