Instructions

The programs of ATSAS suite are available form http://www.embl-hamburg.de/biosaxs/download.html

Before downloading the package, you need to create an account which will be used to process later jobs on the cluster of EMBL.

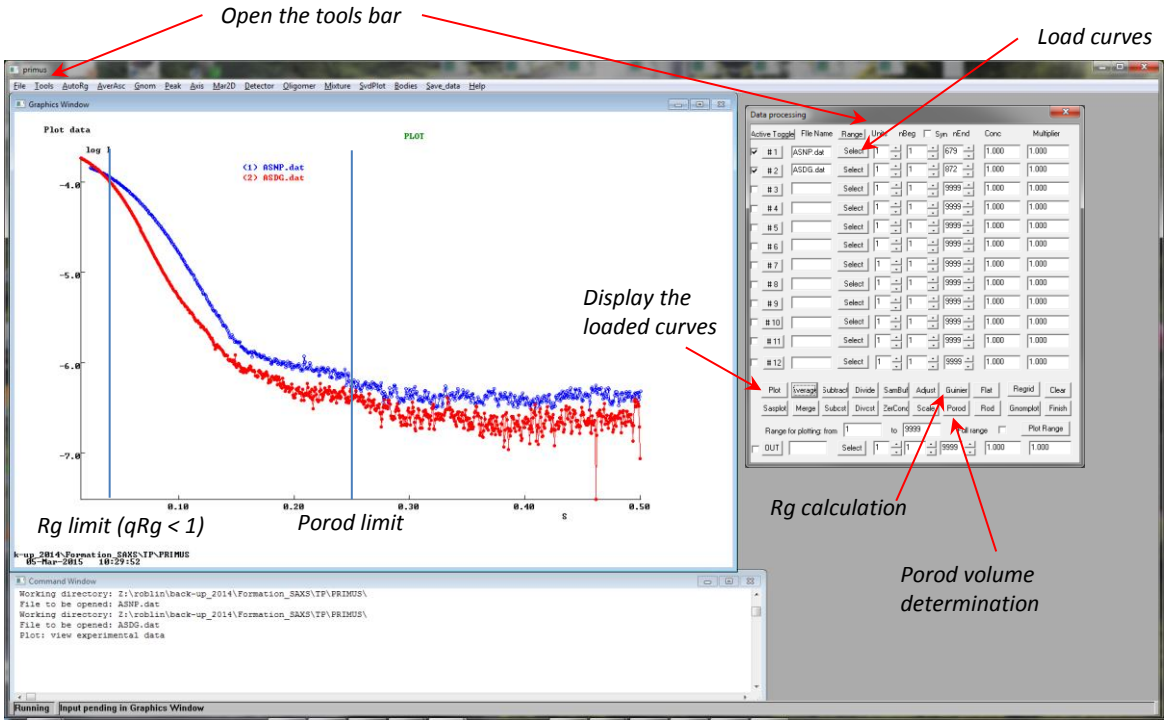
For the practical course, the file are located in the the folder “Input_file”. In the second folder “solution” you will find the final models or results generated from the different program.

- Open the PRIMUS program and load the data files of ASNP and ASDG
- Calculate the Rg value for ASNP and ASDG
- Determine the Porod volume of the two proteins
- Conclusion
- With GNOM program, determine an approximation value of Dmax to obtain a p(r) function for each protein
- Compare the Rg values obtained previously with Guinier extrapolation
- With the gnom.out file, launch a calculation with DAMMIF for ASNP and ASDG on your computer. In parallel, you can connect to ATSAS online to start 10 runs simultaneously with DAMMIF and the DAMAVER suite.
- Determine the most representative shape and compare the shape of ASNP and ASDG to the structure of ASNP (see in the folder Solution/DAMMIF-ASNP or ASDG)
- With the SASREF program, generate a model of ASDG with and without given constrains
- With CRYSOL, calculate the SAXS theorical curve with the structure of ASNP and compare to the experimental data
- Perform the same procedure with the data of ASDG
- Calculate and compare the model of ASDG (run1.pdb and run2.pdb) obtained with SASREF to the experimental curves of ASDG
- To finish, try to generate a model of ASDG with a new conformation of the N-terminal part with BUNCH

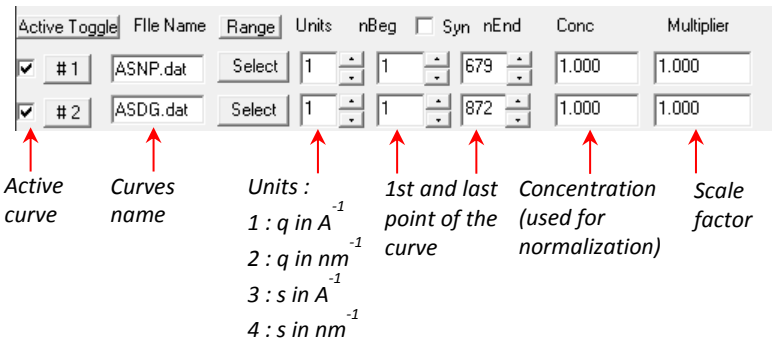
Conclusion

PRIMUS

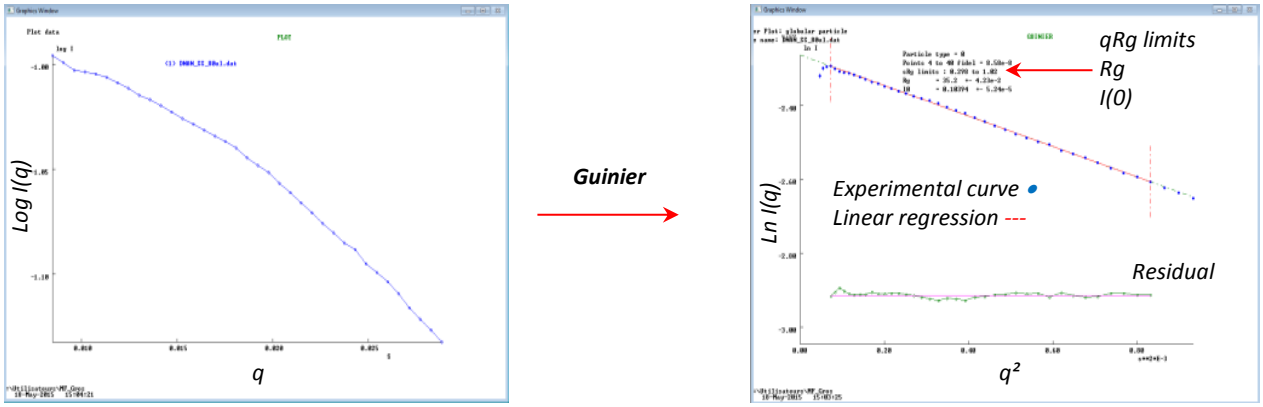
This program is used to display curves and apply some operation on the curves (scaling, substraction or merging…) and to determine the gyration radius Rg, the Porod volume.



By clicking on the "Tools" menu, you will display a table where you can load the different SAXS curves (click on "Select" button). When the data are loaded in the table, the name appears in the "File name" area. To display a curve, select the curve and click on the "plot" button. With the buttons "NBeg" and "nEnd" you can remove respectively the first points or the last points of the curve. If you want to compare many curves each other by superimposition, you select the curve and click on the "scale" button. The program applies a scale factor whose value is reported in box "Multiplier" and can be changed manually to modify the scaling.

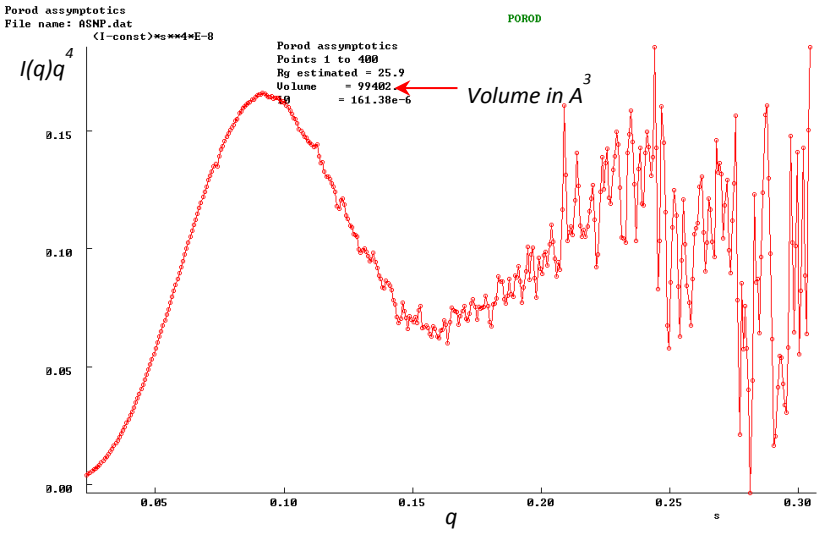


For Rg calculation, you can use the function "Guinier" and modify manually the parameters or use directly the program "AutoRg" present on the general toolbar. Before to click on the "Guinier" button, you must restrict the number of points by taking into account the portion small angles as qRg is less than 1 (try nEnd = 50 to begin). You can click on the "Plot" button to display just the small angles area and after use the "Guinier" function.



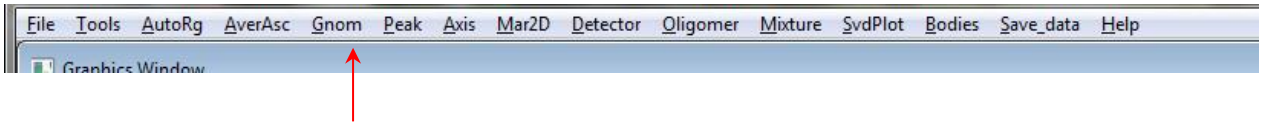
You can evaluate the quality of the fit with the residual plot which represents the distribution of the experimental points around the regression line. You can control also the limit q*Rg max which should be inferior to 1, but for this limit depends of the geometry of the object. For globular protein, the limit can be increased up to 1,4 whereas for elongated or unfolded protein, the Guinier region is more restricted (less than 0,8). 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).

For Porod volume calculation, you can use the function "Porod" and modify manually the parameters (nBeg or nEnd). This calculation is correct for globular object and is not appropriate for elongated or unstructured object. There is no precise angular limit to apply this law, but in general the data is cut from 0.2 to 0.25 to $Å^{-1}$. As Guinier calculation, you select the correct area and click on "Porod" button and after on "Porod" button.

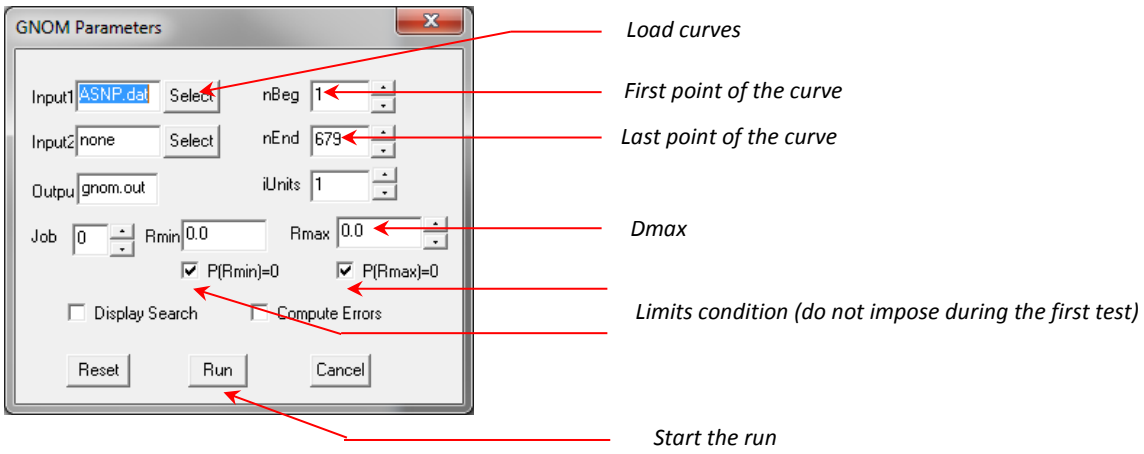


GNOM

The program GNOM is used to determine the autocorrelation function p(r) from the SAXS data. GNOM can be launch directly from PRIMUS toolbar or from ATSAS folder with gnomqw.exe.

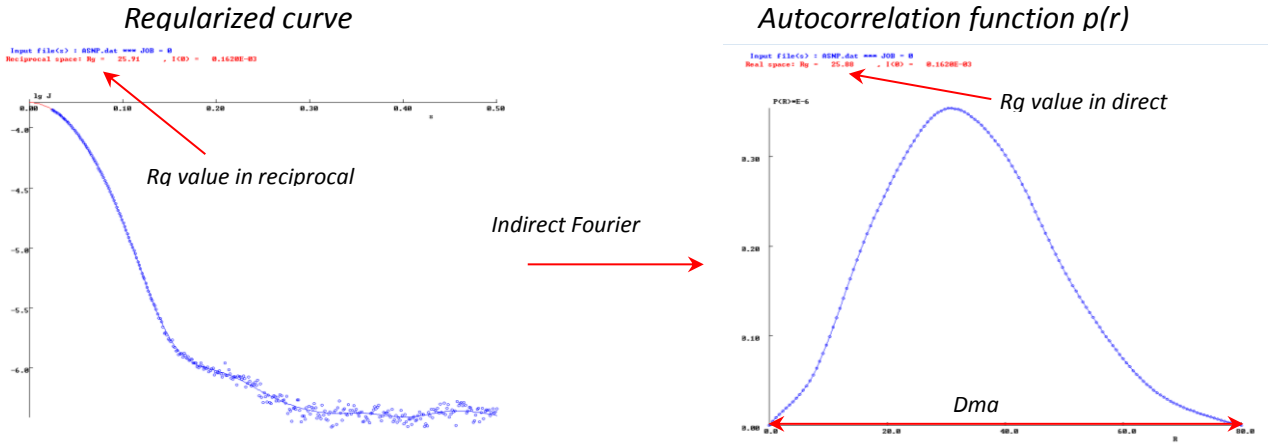


From PRIMUS, a new window appears where some parameters can be modified to perform a p(r) calculation.

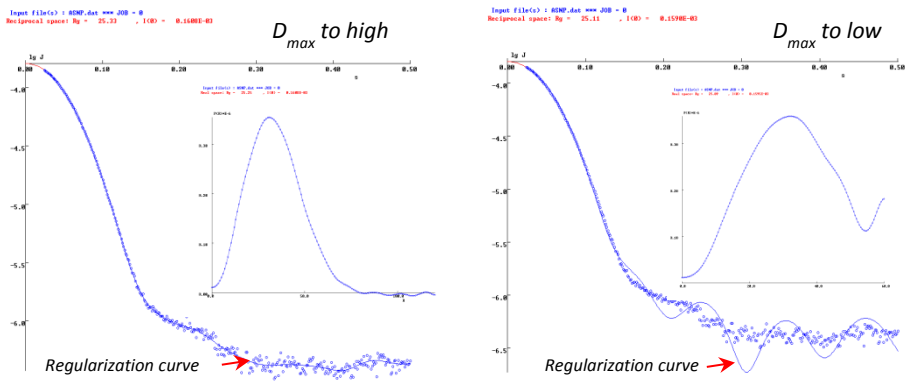


In the box input1, you can load the experimental data, and precise the number of points at the beginning and the end of the curve which will be omitted. The value of Rmin remains equal to 0 and the Rmax correspond to the internal maximum distance of the particle. For the first tests, the conditions where $P(r_{min})$ and $P(r_{max}) = 0$ must be imposed, and can be removed during the last process. When all the parameters are defined, the calculation can be launched by pressing the "Run" button.

A new window appears with the regularization fit of the SAXS curve, and when the fit is validated, the corresponding p(r) is showed on a second window.

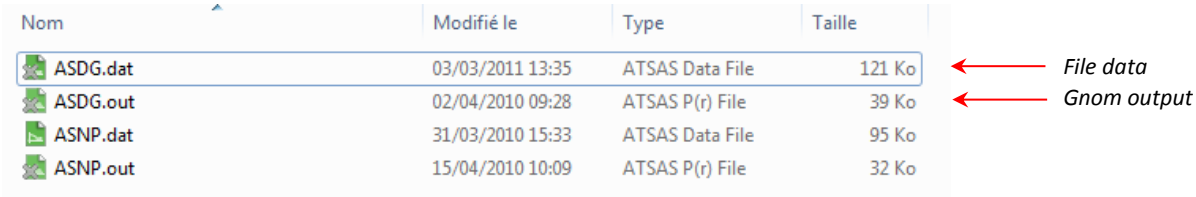


To determine a correct value of Dmax, we must proceed by trial and error to find a Rg calculated with GNOM similar to that found with the calculation of Guinier. We start in general with a value of Dmax equal to 4 or 5 times the value of Rg, and decrease gradually the value of the Dmax in order to obtain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



During the process, you will press return to validate each step of the p(r) calculation and the end, the program offers a summary table grouping parameters to appreciate the quality of the fit (for more, read the paper Svergun D.I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495-503).

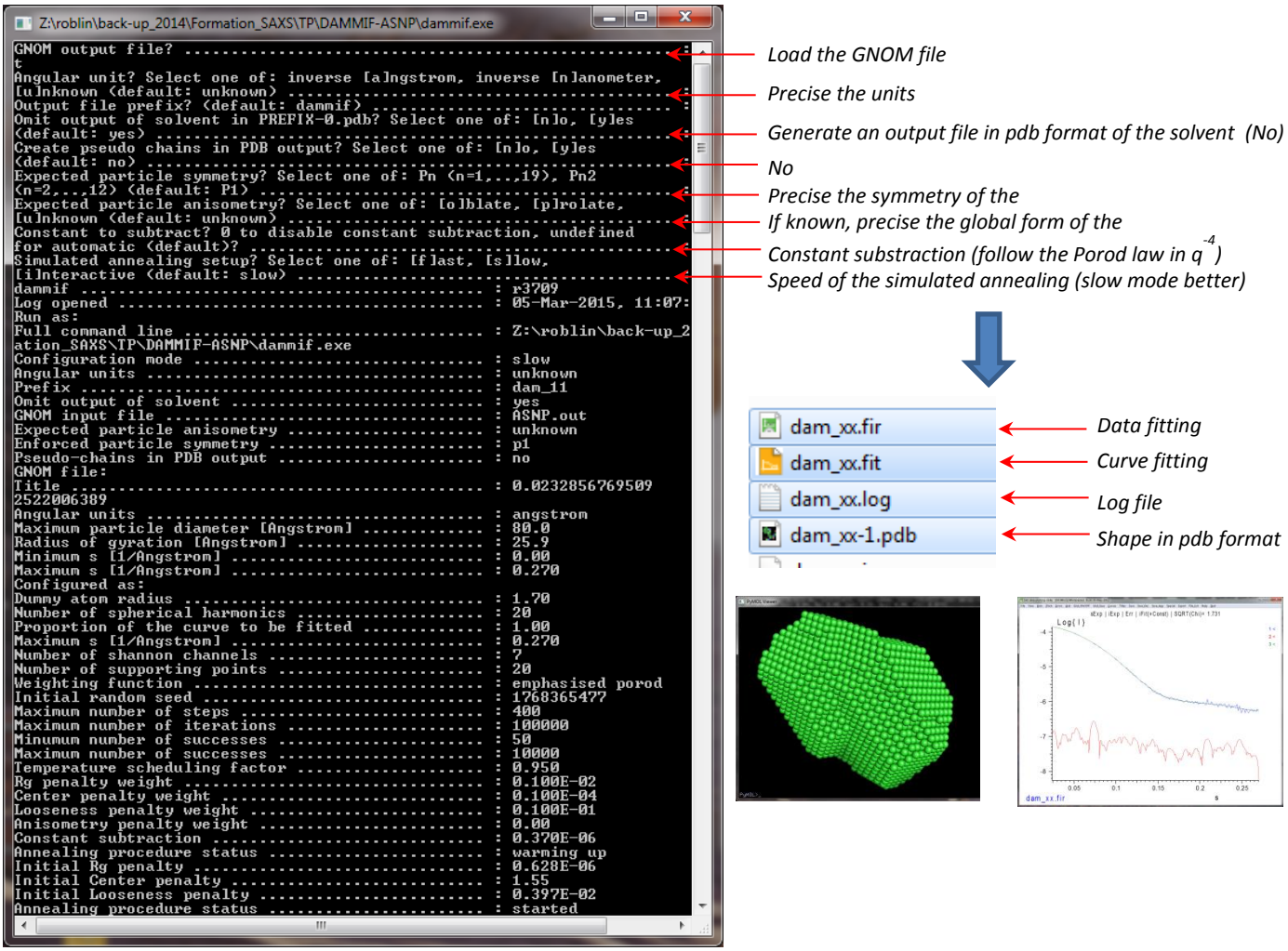
At the end, you will obtain a new file called .out containing all informations about the p(r) determination such as parameters defined preliminary in GNOM (nBeg, nEnd, …), biophysical parameters (Rg, I(0) and Dmax). You will find also in ASCII format, the SAXS curve with the corresponding regularization curve and the p(r) function.



The file.out will be used by ab initio program such as DAMMI, DAMMIN or GASBOR.

DAMMIF

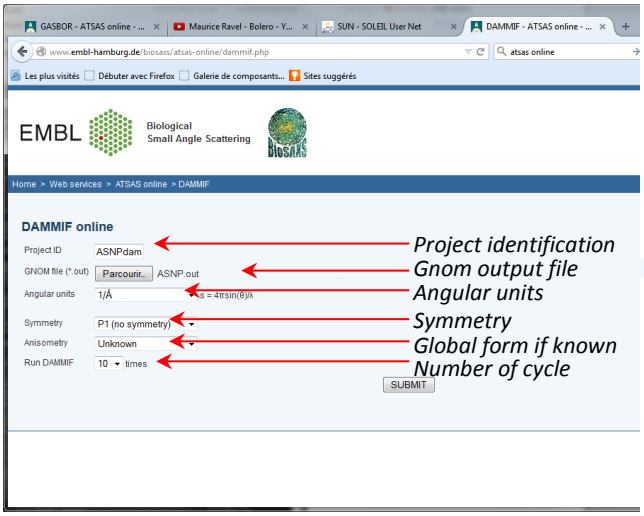
The program DAMMIF is dedicated to low resolution shape modeling using a sphere containing beads (with a defined value of electronic density) as initial model. By clicking on dammif.exe, a command window appears:

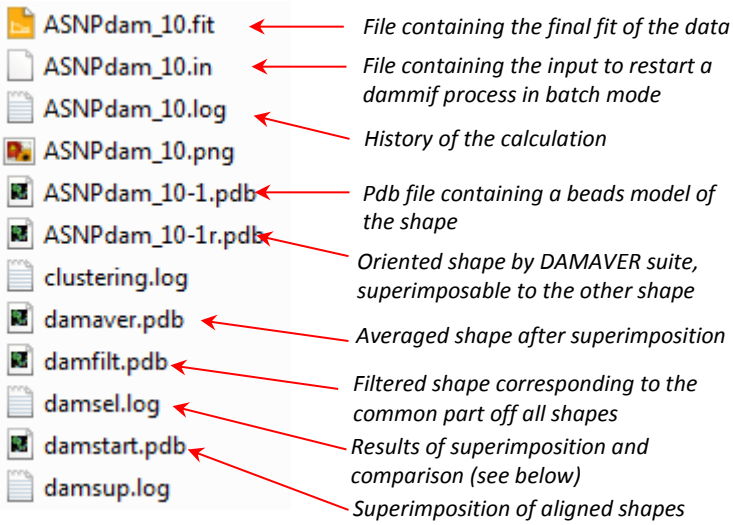


By following the instruction described below, you will obtain at the end, a pdb file containing the shape composed of dummy residues, a fitting curve (smoothed curve) dam_xx.fit, a fitting curve corresponding to the experimental data dam_xx.fir and a log file with initial parameters and the process of minimization steps.

DAMMIF on atsas online (low resolution shape, superimposition, comparison)

http://www.embl-hamburg.de/biosaxs/atsas-online/





With DAMMIF on ATSAS online, you can launch many runs simultaneously in order to generate a set of envelope that can be compared with each other with the DAMAVER package suite. This package allows to superimpose and to compare the shapes between them. 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 whch represents a filtered shape corresponding to the common part off all shape. This model, in general case, does not fit the SAXS data !

Matrix giving the NSD score between the different shape

Criteria of selection

Most representative shape which have the lowest mean value of NSD

```
  ---  Created by DAMSEL     Thu Apr 15 18:00:07 2010 ---
 List file name ..................................... :
damsel.inp
    ------------------------------------------
     Cross-correlation NSD table by SUPCOMB
    ------------------------------------------
 File  Aver  0-1  2-1  3-1  4-1  5-1  6-1  7-1  8-1  9-1  emp
dam_10 0.52  0.00 0.51 0.55 0.44 0.52 0.56 0.58 0.51 0.52 0.52
dam_2- 0.52  0.51 0.00 0.52 0.52 0.53 0.48 0.56 0.51 0.52 0.50
dam_3- 0.54  0.55 0.52 0.00 0.56 0.54 0.51 0.54 0.55 0.54 0.56
dam_4- 0.52  0.44 0.52 0.56 0.00 0.51 0.53 0.56 0.53 0.50 0.55
dam_5- 0.53  0.52 0.53 0.54 0.51 0.00 0.53 0.56 0.52 0.51 0.52
dam_6- 0.52  0.56 0.48 0.51 0.53 0.53 0.00 0.55 0.54 0.43 0.52
dam_7- 0.55  0.58 0.56 0.54 0.56 0.56 0.55 0.00 0.55 0.57 0.52
dam_8- 0.53  0.51 0.51 0.55 0.53 0.52 0.54 0.55 0.00 0.52 0.55
dam_9- 0.51  0.52 0.52 0.54 0.50 0.51 0.43 0.57 0.52 0.00 0.51
temp.p 0.53  0.52 0.50 0.56 0.55 0.52 0.52 0.52 0.55 0.51 0.00
  Aver 0.53  0.52 0.52 0.54 0.52 0.53 0.52 0.55 0.53 0.51 0.53

 Mean value of  NSD =      0.527
 Variation  of  NSD =      0.012
 Recommend to discard files fith NSD > Mean + 2*Variation

        dam_9-1.pdb  --- Reference // Aver NSD =  0.514
        dam_6-1.pdb  --- Include   // Aver NSD =  0.516
        dam_2-1.pdb  --- Include   // Aver NSD =  0.518
       dam_10-1.pdb  --- Include   // Aver NSD =  0.522
        dam_4-1.pdb  --- Include   // Aver NSD =  0.523
        dam_5-1.pdb  --- Include   // Aver NSD =  0.525
        dam_8-1.pdb  --- Include   // Aver NSD =  0.529
           temp.pdb  --- Include   // Aver NSD =  0.529
        dam_3-1.pdb  --- Include   // Aver NSD =  0.540
        dam_7-1.pdb  --- Discard   // Aver NSD =  0.554
```

SASREF

The program SASREF is used to perform molecular modeling in rigid bodies against SAXS data. To start correctly a run with SASREF, you will need the SAXS data file, the pdb file and the .alm file of each part corresponding respectively to the atomic structure of the part and the corresponding amplitude. The .alm file is generated with CRYSOL starting from the pdb file. To generate a correct amplitude file, use 15 spherical harmonics and a q_{max} value given by the relation $q_{max} = 2*(L_{max} - 5)/D_{max}$ (where L_{max} is the number of spherical harmonics and D_{max} the maximum size of the particle). The number of point used to generate the theoretical curve does not excess 51 points. If possible, use a condition contacts file to restrain the possibilities because many conformations can fit the SAXS curve.

```
  ====== SASREF06  started on          05-Mar-2015   12:14:36

Computation mode (User or Expert) ..... <         User >: e
Log file name ......................... <         .log >: run3
Project identificator ................................ : run3
Enter project description ............. : modelling of ASDG in P2
Random sequence initialized from ...................... : 121512
Input total number of subunits ........ <            1 >:
Symmetry: P1...19 or Pn2 (n=1,..,12) ... <           P1 >: P2
Input total number of scattering curves  <            1 >:
Input first & last subunits in 1-st construct <            1,

        ***  Enter file name, 1-st experimental data  ***

 Working directory: Z:\roblin\back-up_2014\Formation_SAXS\TP\SASREF\
 File to be opened: ASDG.dat
Number of experimental points found ................... : 871
Angular units in the input file :
4*pi*sin(theta)/lambda [1/angstrom] (1)
4*pi*sin(theta)/lambda [1/nm      ] (2)
2*   sin(theta)/lambda [1/angstrom] (3)
2*   sin(theta)/lambda [1/nm      ] (4)  <            1 >:
Fitting range in fractions of Smax ..... <        1.000 >: 0.5
Number of points reduced to ........................... : 421
Experimental radius of gyration ....................... : 35.83
Number of points in the Guinier Plot .................. : 35

        ***  Amplitudes, 1-st subunit  ***

 Working directory: Z:\roblin\back-up_2014\Formation_SAXS\TP\SASREF\
 File to be opened: ASDG_monomer00.alm
Maximum order of harmonics ............................ : 15
Number of points in partial ampliudes ................. : 51
 SASREF --W- Lm reduced to compute cross term
Current subunit:   5117 atoms read, center at    20.72   30.55  -22.85
Initial rotation by alpha ............. <          0.0 >:
Initial rotation by beta .............. <          0.0 >:
Initial rotation by gamma ............. <          0.0 >:
Initial shift along X ................. <        20.72 >: 0
Initial shift along Y ................. <        30.55 >: 0
Initial shift along Z ................. <       -22.85 >: 0
Fix the subunit at this position? [ Y / N ] <           No >:

Spatial step in angstroems ............ <        5.000 >:
Angular step in degrees ............... <       20.00 >:
ALMGRZ --- :   91800 summation coefficients used
Cross value ........................................... : 34.23
Discontiguity value ................................... : 0.0
Cross penalty weight .................. <       10.00 >:
Disconnectivity penalty weight ........ <       10.00 >:

        ***  PLEASE SELECT THE CONTACTS CONDITIONS FILE  ***

  Working directory: Z:\roblin\back-up_2014\Formation_SAXS\TP\SASREF\
  File to be opened: Contacts.cnd
Condition #  1: Distance   4.800
   Between subunit #  1, Residues from  CA  PRO A 566  to  CA  PRO A 566
       and subunit #  2, Residues from  CA  ALA A 347  to  CA  ALA A 347
Condition #  2: Distance   6.800
   Between subunit #  1, Residues from  CA  ALA A 383  to  CA  ALA A 383
       and subunit #  2, Residues from  CA  SER A 571  to  CA  SER A 571
Condition #  3: Distance   8.500
   Between subunit #  1, Residues from  CA  HIS A 348  to  CA  HIS A 348
       and subunit #  2, Residues from  CA  ASP A 564  to  CA  ASP A 564
Condition #  4: Distance   9.000
   Between subunit #  1, Residues from  CA  ARG A  27  to  CA  ARG A  27
       and subunit #  2, Residues from  CA  ALA A  75  to  CA  ALA A  75
Condition #  5: Distance  13.000
   Between subunit #  1, Residues from  CA  VAL A 617  to  CA  VAL A 617
       and subunit #  2, Residues from  CA  ALA A 383  to  CA  ALA A 383
 Contacts conditions penalty ........................... : 42.28
 Contacts penalty weight ............... <       10.00 >: 200
 Expected particle shape: <P>rolate, <O>blate,
  or <U>nknown ......................... <      Unknown >:
 Shift penalty is normalized by ........................ : 26.63
 Shift penalty ......................................... : 0.0
 Shift penalty weight .................. <       1.000 >:
 Total penalty ......................................... : 8799.
 1-st curve:
 Theoretical  points from           3  to          37  used
 The best chi values:93.15830
 Initial fVal .......................................... : 1.748e+4
 Initial annealing temperature ......... <       10.00 >:
 Annealing schedule factor ............. <      0.9000 >:
 Max # of iterations at each T ......... <        5000 >:
 Max # of successes at each T .......... <         500 >:
 Min # of successes to continue ........ <          50 >:
 Max # of annealing steps .............. <         100 >:
  ====  Simulated annealing procedure started  ====
```

Select expert mode

Give a name for the run

Add a comment to describe the run

Precise the symmetry of the particle

Load the data

Precise the angular units

The portion of the curve fitted. Be careful, SASREF use 15 spherical harmonics, and the q range depends also of the size of particle.

Load the amplitude of each subunits calculated previously with CRYSOL (L = 15 and nb of points 51)

Put the center of the mass to the origin

Don't fix the subunits

The penalty can be modified to avoid steric clashes or discontinuous model.

You can add a contact file which contains a list of fixed distance between residus. Not necessary to work with SASREF but recommended to limit the number of possibility.

The contacts conditions penalty must be increased to force the program to respect the contact conditions and the default value is not sufficient

Keep the default parameters for the simulated annealing

At the end of the run, you will obtain a set of file with the pdb file of the final model, the fitting curve .fit and the log file. To verify the validity of the model, you can perform a CRYSOL calculation against SAXS data with more spherical harmonics (L = 50) and take the totality of the curve for the fit.

CRYSOL

The program CRYSOL is employed to calculate a theoretical SAXS curve starting from a pdb file and to compare the result to the experimental SAXS data.

```
------------------------------------------------
                Program options :
0 - evaluate scattering amplitudes and envelope
1 - evaluate only envelope and Flms
2 - read CRYSOL information from a .sav file
------------------------------------------------

Enter your option ..................... <           0 >:

           ***  PLEASE SELECT THE PDB FILE NAME  ***

Working directory: Z:\roblin\back-up_2014\Formation_SAXS\TP\CRYSOL-ASDG\
File to be opened: ASDG-dimere-PISA.pdb
 2 (AB) chains were found in ASDG-dimere-PISA.pdb
Process chain (0: all chains) .........................: 0
 ------------------------------------------------
        Following file names will be used:
ASDG-dimere-PISA01.log -- CRYSOL log-file          (ASCII)
ASDG-dimere-PISA01.sav -- save CRYSOL information  (binary)
ASDG-dimere-PISA01.flm -- multipole coefficients   (ASCII)
ASDG-dimere-PISA01.int -- scattering intensities   (ASCII)
ASDG-dimere-PISA01.fit -- fit to experimental data (ASCII)
ASDG-dimere-PISA01.alm -- net partial amplitudes   (binary)
 ------------------------------------------------
 Maximum order of  harmonics ........... <          15 >: 50
 Order of Fibonacci grid ............... <          18 >:
 ----------  Reciprocal space grid  -------------
  ( in s = 4*pi*sin(theta)/lambda [1/angstrom] )
Maximum s value ....................... <       1.000 >: 0.5
Number of points ...................... <          51 >: 256
Account for explicit hydrogens? [ Y / N ] <          No >:
Read atoms and evaluate geometrical center ...
 Number of atoms read ................................ : 10234
Percent processed      10  20  30  40  50  60  70  80  90 100
Processing atoms   :>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
Center of the excess electron density:  0.000  0.075  0.000
Processing envelope:>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
Coefficients   saved to file ASDG-dimere-PISA01.flm
CRYSOL data    saved to file ASDG-dimere-PISA01.sav
--- Structural parameters (sizes in angstroms) ---
Electron   rg   :  33.04       Envelope   Rg      :  32.80
Shape      Rg   :  33.08       Envelope  volume   : 0.2160E+06
Shell    volume : 0.5276E+05   Envelope  surface  : 0.1625E+05
Shell      Rg   :  42.11       Envelope  radius   :  58.85
Shell    width  :  3.000       Envelope  diameter :  115.5
Molecular weight: 0.1445E+06   Dry volume         : 0.1752E+06
Displaced volume: 0.1804E+06   Average atomic rad.:  1.614
Number of residuals :  1302
 Fit the experimental curve [ Y / N ] .. <         Yes >:

           ***  PLEASE SELECT THE DATA FILE NAME  ***

Working directory: Z:\roblin\back-up_2014\Formation_SAXS\TP\CRYSOL-ASDG\
File to be opened: ASDG.dat
Subtract constant ..................... <          no >: y
Title: ? Data from ASDG.dat
 Maximum angle in the data file ........................ : 0.5004
 Number of experimental points ......................... : 872
Angular units in the input file:
4*pi*sin(theta)/lambda [1/angstrom] (1)
4*pi*sin(theta)/lambda [1/nm]       (2)
2 *  sin(theta)/lambda [1/angstrom] (3)
2 *  sin(theta)/lambda [1/nm]       (4)  <           1 >:
 Angular units multiplied by ........................... : 1.000
 Number of points after regriding ...................... : 201
Electron density of the solvent, e/A**3  <      0.3340 >:
 Number of experimental points used .................... : 200
 ------  Fitting the experimental data ...    ---
Plot the fit [ Y / N ] ................. <         Yes >:
Another set of parameters [ Y / N ] .... <          No >:
Rg from the slope of net intensity ..................... : 35.07
Average electron density ............................... : 0.4624
Data fit       saved to file ASDG-dimere-PISA01.fit
Intensities    saved to file ASDG-dimere-PISA01.int
Net amplitudes saved to file ASDG-dimere-PISA01.alm
 Press CR to terminate the program
```

Select option 0 (default)

Select the pdb file

Select the totality of the pdb contains or just some chain

*Enter the number of spherical harmonic determined by the relationship : **Lmax – 15 = Qmax*Dmax / 2***

For the Fibonacci grid, less the default value

Give the Qmax value for the fit

Enter number of point for the fitting curve (default to 256)

Less the default proposition

After the processing of atoms and envelope, CRYSOL generate two files, the .flm file containing the enveloppe structure and function parameters , the .sav file containing all information to restart CRYSOL with already made calculation

CRYSOL give also biophysical parameters such as the Rg of enveloppe or the volume (in A^3)

Select the experiment data file

You can add a supplementary parameter of the fitting process to offset the backgroud (linked to the substraction of solvent in major case)

Put the correct value of the angular units

For the classical buffer without excess of salt of organic molecule, the default value is correct

You can display the experimental curve fitting

A the end CRYSOL generate the .alm file containing the amplitudes which are used by SASREF or BUNCH, the .fit file in ascii format usable with another spreadsheet, and the .int file containing the intensity of the different component (electron, border shape and excluded volume)