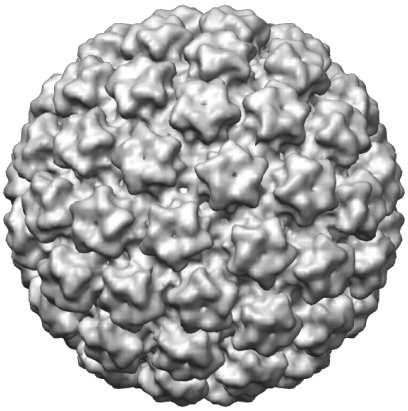
Scipion Tutorial Series

National Center for Biotechnology

Biocomputing Unit

Initial Volume Validation by SAXS



Scipion Team January 2018

Intended audience

This tutorial provides a guide to the validation of the initial volume estimations using SAXS curves in Scipion. Some basic knowledge about 3DEM image processing and basic computer skills are assumed.

We'd like to hear from you

We have tested and verified the different steps described in this tutorial to the best of our knowledge, but since our programs are in continuous development you may find inaccuracies and errors in this text. Please let us know about any errors, as well as your suggestions for future editions, by writing to scipion@cnb.csic.es.

Contents

1 General Introduction

1.1 Software Installation

To follow this tutorial you will need to download and install Scipion and the program Crysol available at https://www.embl-hamburg.de/biosaxs/crysol.html (Svergun et al., 1995).

For the full documentation please refer to the Scipion installation page.

1.2 SAXS and EM Initial Volume Estimation

Small angle scattering of X-rays (SAXS) is a technique to study biological structures that provides low resolution information (Koch et al., 2003). SAXS experiments give a one-dimensional scattering curve, where the scattered intensity recorded by a detector is represented as a function of the scattering angle.

The SAXS scattering curves can be used to complement the information obtained by EM. Specifically, in this tutorial we will take advantage of the low-resolution structural information revealed by SAXS to automatically pruned the initial 3D volumes obtained by the typically used methods in EM, helping the following refinement process to obtain more accurate 3D structures.

To this end, we approximate the initial volume using a pseudo-atoms model and simulate the SAXS curve expected for this pseudo-atomic structure. The match between the simulated and the experimental SAXS curves will be a good sign of the accuracy of the EM initial volume.

2 Initial Volume Validation by SAXS

2.1 Getting Started

In this tutorial we work with the Bos taurus catalase structure which PDB model is available at EMDB (6314). Moreover, the experimental SAXS curve for this structure is also available in SAXS database SASBDB (SASDA92). The SAXS curve must be

downloaded from SASBDB to measure the match between it and the simulated curve obtained with the proposed workflow (explained in the following sections).

For this structure, we generate two initial volume estimations to work with that can be downloaded using the following command:

```
scipion testdata --download SAXS_tutorial
```

This command will download (if this was not done earlier) the test data in directory

`$SCIPION_HOME/data/tests/SAXS_tutorial`

After downloading, you must launch the main GUI by typing: `scipion` Then, create a new project by clicking Create project button, type a *Project name* and click OK. The main project window will appear.

In the project window, the left panel shows a list with processing tasks (protocols) that can be used. You can also search protocols by Ctrl + F. The right panel displays the tree of protocols executed (runs) by user and their state: saved, running, finished or aborted. The special Analyze Results button is used to visualize outputs and plot results.

2.2 Basic Workflow

To generate the simulated SAXS curves we will follow the workflow shown in Figure 1.

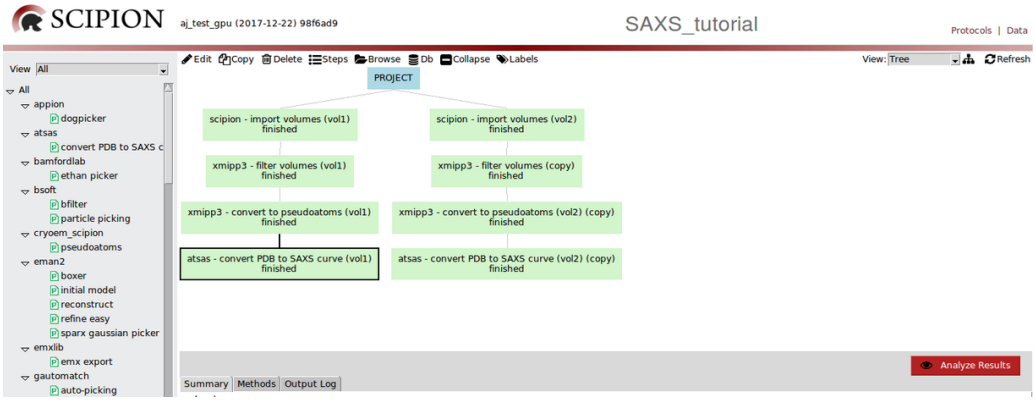


Figure 1: Project window with the target workflow.

First, the set of estimated initial volumes are read and filtered. Then, a pseudo-atoms model is applied to the filtered volumes. Finally, we estimate the SAXS curves from the previous pseudoatomic models. The simulated and the experimental SAXS curves will be compared using the χ^2 distance to select the best initial volumes.

In the next subsections we will explain every step of the workflow in detail.

2.2.1 Import volumes

The first step is simply to import the volumes available in the data for this tutorial. To this end we use the Scipion Import Volumes protocol available in Imports ... . In the configuration we only need to indicate the files directory with the complete path to the downloaded volumes. The remaining parameters in the form can stay with the default values.

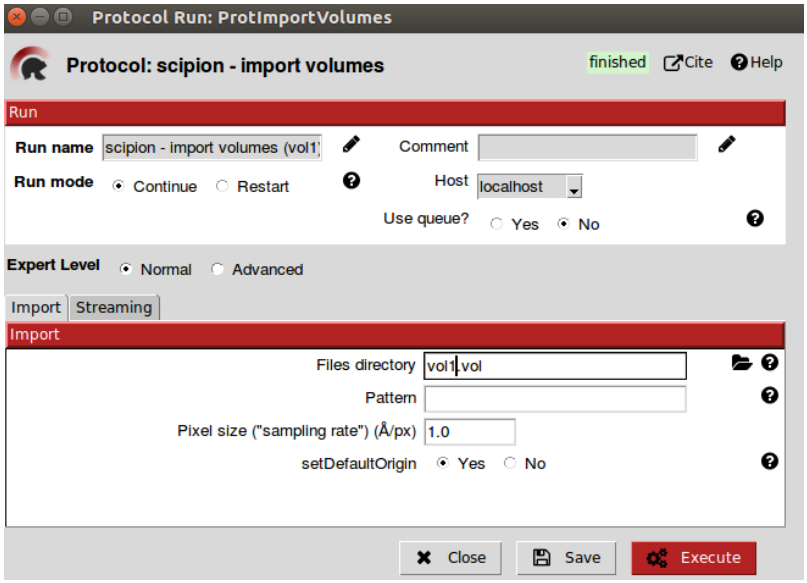


Figure 2: **Import volumes** Protocol GUI.

2.2.2 Filtering the volumes

We apply a low-pass filter to the initial volumes to discard any high resolution information. To this end, we use the protocol Filter Volumes in Xmipp, with a cut-off

resolution of 15 Å(some common resolutions could be in the range from 10 up to 20 Å). This protocol can be found at 3D » preprocess » ....

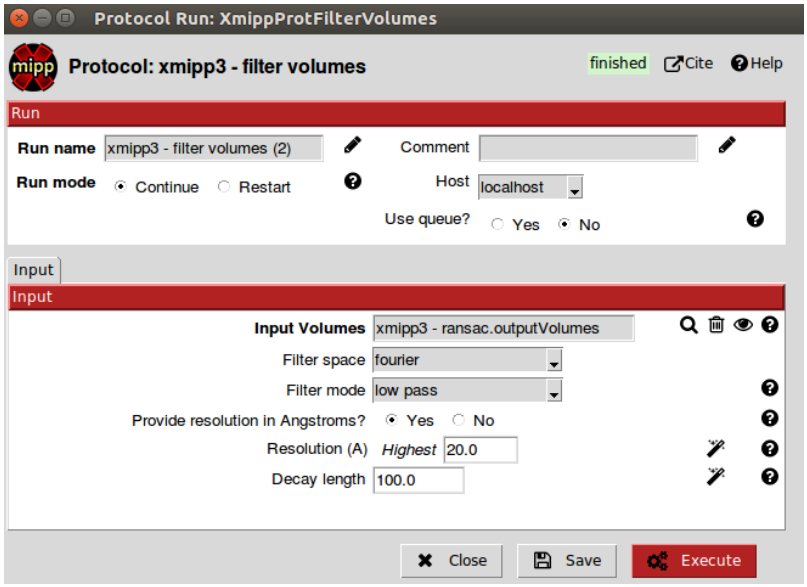


Figure 3: **Filter volumes** Protocol GUI.

2.2.3 Pseudo-atoms model

The filtered volumes are converted to a pseudo-atoms model with a Gaussian sigma between 1 and 2 voxels. We use the protocol *Convert to Pseudoatoms* in Xmipp that can be found at 3D » analysis » more » ... (see Figure 4).

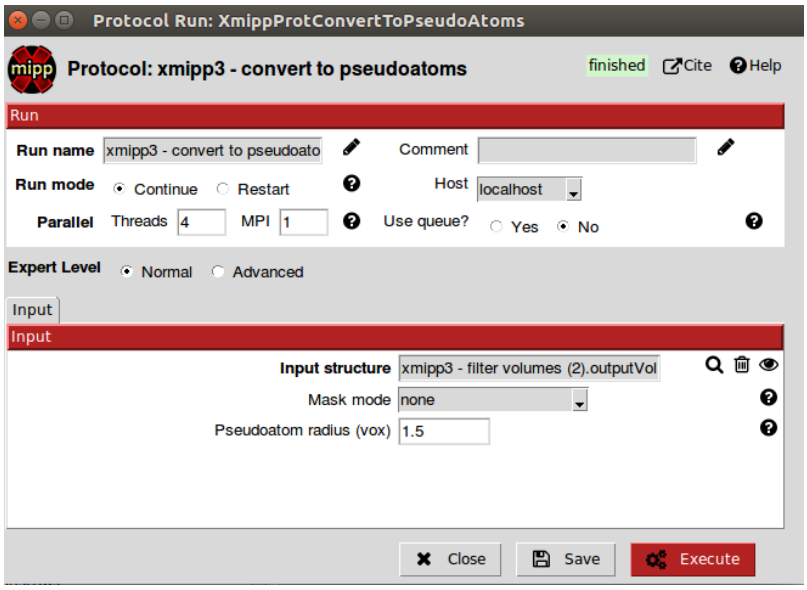


Figure 4: **Convert to pseudoatoms** Protocol GUI.

2.2.4 Comparison of SAXS curves

From the previous pseudo-atomic model, the simulated SAXS curves can be obtained using the protocol Convert PDB to SAXS Curve in Atsas. Moreover, this protocol allows us to compare these curves with the experimental ones available in SASBDB (previously download from the web). As it can be seen in Figure 5, we only need to specify the structure obtained in the previous step as Input structure, and the experimental SAXS curve from SASBDB. This protocol can be found in Scipion All > Atsas > ....

Crysol program will be in charge of calculating the comparison between both curves. If we press the **Analyze Results** button we will see the obtained curves. Moreover, in the **Summary** window we can see the χ^2 measure revealing the goodness of fit between both curves (if you don't find the χ^2 value in the **Summary**, you can find it clicking in **Browse**, **extra** folder, **crysol_summary.txt** file).

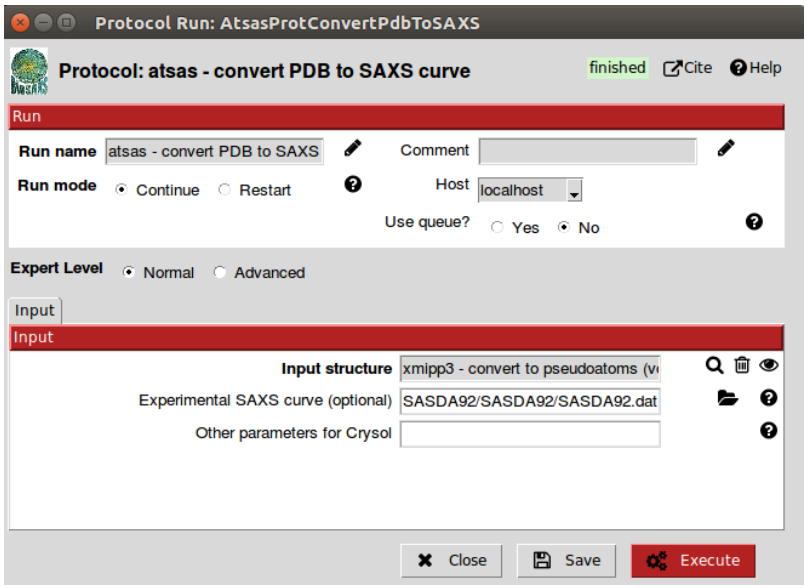
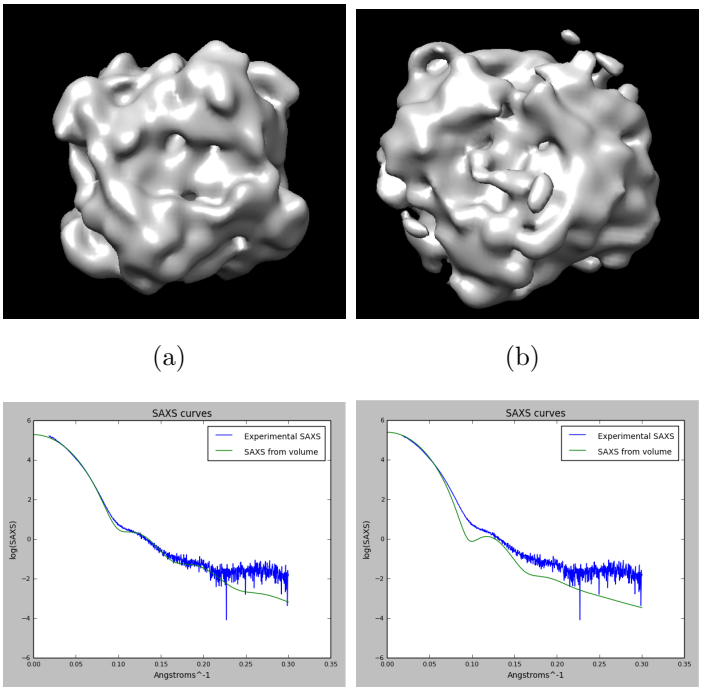


Figure 5: **Convert PDB to SAXS Curve** Protocol GUI.

In Figure 6 we show the obtained SAXS curves for both initial volumes and the χ^2 results achieved by the comparison between the experimental and the simulated SAXS curves. As it can be seen, with a good initial volume estimation the curves tend to maintain overlap until very high frequencies; on the other hand, with an incorrect initial estimation, the curves clearly diverge. The χ^2 values confirm these conclusions, as for a correct initial estimation the χ^2 obtained is lower than for an incorrect one.



(a) (b)

(c) (d)

Figure 6: Results. (a) Correct initial volume. (b) Incorrect initial volume. (c) Comparison between SAXS curves for the volume in (a) with $\chi^2 = 21.82$. (d) Comparison between SAXS curves for the volume in (b) with $\chi^2 = 143.91$.

References

Koch, M. H. J., Vachette, P., and Svergun, D. I. (2003). Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Quarterly Reviews of Biophysics*, 36(2):147–227.

Svergun, D., Barberato, C., and Koch, M. H. J. (1995). *CRYSOL* – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Applied Crystallography*, 28(6):768–773.