Feb 12, 2021

Fitting PDB files to SAXS data using FOXS web server

Chris Berndsen[1]
[1]James Madison University

1 Works for me This protocol is published without a DOI.

Chris Berndsen
James Madison University

SUBMIT TO PLOS ONE

ABSTRACT

FOXS is a server for fitting models to SAXS data. It can also predict the SAXS data for protein models in the absence of SAXS data.

Schneidman-Duhovny D, Hammel M, Tainer JA, and Sali A. Accurate SAXS profile computation and its assessment by contrast variation experiments. Biophysical Journal 2013. 105 (4), 962-974

Schneidman-Duhovny D, Hammel M, Tainer JA, and Sali A. FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles NAR 2016

PROTOCOL CITATION

Chris Berndsen 2021. Fitting PDB files to SAXS data using FOXS web server. **protocols.io** https://protocols.io/view/fitting-pdb-files-to-saxs-data-using-foxs-web-serv-bd3di8i6

CREATED

Mar 21, 2020

LAST MODIFIED

Feb 12, 2021

PROTOCOL INTEGER ID

34629

MATERIALS TEXT

Internet access
Known protein models
SAXS data in the form of Intensity vs. q (optional)

.dat file

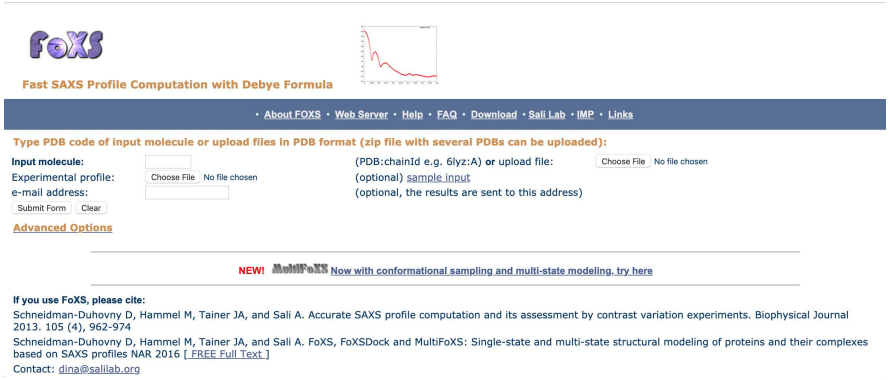
1 To perform this analysis you will need subtracted SAXS data in a .dat file that has the structure shown below:

```
# q intensity error
  0.203610E-01  0.326420E-01  0.141695E-02
  0.210620E-01  0.318600E-01  0.136604E-02
  0.217880E-01  0.320550E-01  0.134368E-02
  0.225380E-01  0.314110E-01  0.122690E-02
  0.233140E-01  0.305640E-01  0.943570E-03
  0.241170E-01  0.301890E-01  0.870996E-03
  0.249480E-01  0.299540E-01  0.969085E-03
  0.258070E-01  0.291470E-01  0.959823E-03
  0.266960E-01  0.291550E-01  0.953466E-03
  0.276160E-01  0.285550E-01  0.921233E-03
  0.285670E-01  0.279190E-01  0.751552E-03
  0.295510E-01  0.274570E-01  0.863553E-03
  0.305690E-01  0.269400E-01  0.749686E-03
  0.316220E-01  0.265030E-01  0.778591E-03
```

1.1 If you do not have subtracted SAXS data see the protocol on SCATTER or ATSAS to generate this file.

Loading the data into webserver

2 Navigate to the [FOXS webserver]



3 Provide the PDB file to be fit for the Input Molecule option.

This can be provided as a PDB code, a PDB file, or a zipped file of PDB. This latter option is appropriate for fitting structures from molecular dynamics simulations.

3.1 Provide the experimental data set as a .dat file in the Experimental Profile box.

3.2 Record the input files:

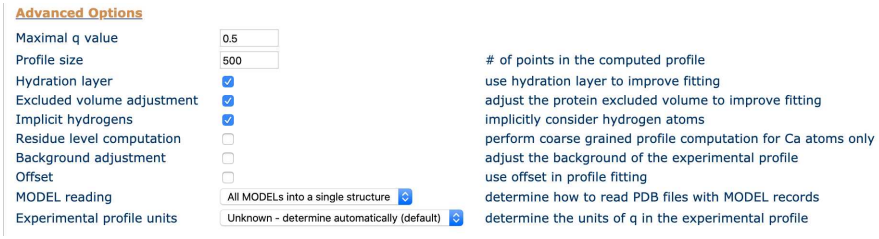
PDB file	
Experimental profile file	

3.3

Advanced options can also be set. These options are useful for fitting to energy minimized structures or ones from simulations.

Options to set:

- **Maximal q:** If data are bad at high q, reduce this value
- **Implicit hydrogens:** Uncheck if hydrogens are in the PDB file(s)
- **Offset**: Add a constant offset when fitting. This option should be used when fitting.



3.4 Press Submit Form and wait 1 to 30 minutes depending on the number of structures to be fitted.

Analyzing fitted profile

4 The results are shown in three regions:

1. Top left side shows a fit of the PDB to the data
2. Top right is visualization of the PDB
3. Bottom table showing statistics of the fit

4.1 The data table is ranked by X^2 values which range from infinity to 1, with 1 being the target value.

Ideally the X^2 value would be ~1 with c_1 and c_2 values ~1 and no more than 4. If either of the c_1 or c_2 is red, this suggests some issue with the fitting.

4.2 Record the numbers of the top 5 fits.

	X^2	c1	c2	Rg	file_name
1					
2					
3					
4					
5					

4.3

A useful visuaization of the data is plotting X^2 vs. Rg (radius of gyration). However, the table can be quite large and hard to record by hand. The code below runs in R and will "scrape" the table from the web into a .csv file.

Paste the URL into the spot listed as "web link from FOXS" and change the protein name and data set name in "proteinname_datasetname_FOXS.csv" as appropriate.

```
# Packages needed to scrape
library(rvest)
library(tidyverse)

# Assign web link to the variable "page"
page <- read_html("web link from FOXS")

# Pull the table nodes from the page
tbl <- page %>%
    html_nodes("table") %>%
# isolate the 7th table which is the data table
    .[7] %>%
# make it a table file
    html_table(fill = TRUE)

# convert the list into a data frame
tbl_df <- as.data.frame(tbl)

# Write a csv file to the working directory.

write.csv(tbl_df, "proteinname_datasetname_FOXS.csv")
```

Sometimes the table is the 6th element in the page, if the data table does not look correct, change the .[7] to .[6]

4.4 Record the name and location of the FOXS table as a note on this step.

5 Click on the fit file link for any structures that fit particularly well.

5.1 Record the location of the fit files as a note on this step.

multi-FOXS

6

If multiple PDB files were provided for fitting, FOXS will also run multi-FOXS which tries combinations of PDB files to produce a better fit. This can reveal the conformational ensemble of the molecule in solution.

These results can be access by clicking the multi-FOXS link.

6.1 Record any multi-FOXS results as a note on this step.