Small Angle X-ray Scattering (SAXS) software tools in Integrative Modeling Platform (IMP)

Dina Schneidman
3/6/2015
Small Angle X-Ray Scattering

Monochromatic beam
X-ray: $\lambda=0.1-0.2\text{nm}$

Protein in solution

Large distance $\theta$

Detector

$\log(\text{intensity}) = f(q)$

SAXS profile

Data can be collected in several minutes!
Integrative Modeling and SAXS

- Solution structure vs. crystal structure
- Assembly of multi domain proteins
- Monomer, dimer or mixture?
- Protein folding
- Structural characterization of protein dynamics

- Assembly of multi protein complexes
Integrative Modeling and SAXS

- Gathering information
- Designing model representation and evaluation
- Sampling good models
- Analyzing models and information

- Standards for data validation
- Sparse data
- Sample heterogeneity
- Atomic representation
- Debye formula
- FoXS
- Debye formula
- MultiFoXS
- Docking, conformational sampling ...
- Fit to data, variance among top scoring models
Fast open-source X-ray Scattering

**foXS**

**forward model**

A rapid method for computing a SAXS profile of a given structure and for matching of the computed and experimental profiles

Debye formula (1915)

\[ I_m(q) = \sum_i \sum_j f_i(q) f_j(q) \frac{\sin(qd_{ij})}{qd_{ij}} \]

\[ \chi^2 = \frac{1}{N-1} \sum_i \left[ \frac{I_{exp}(q_i) - cI_m(q_i)}{\sigma(q_i)} \right]^2 \]

Schneidman-Duhovny D, Hammel M, Sali A. NAR 2010

Schneidman-Duhovny D, Hammel M, Tainer J, Sali A. Biophys J 2013
Fast SAXS Profile Computation with Debye Formula

PDB files
3KFO.pdb

Profile file
23922_merge.dat

Can't see interactive display? Use old interface

<table>
<thead>
<tr>
<th>PDB file</th>
<th>show/hide</th>
<th>χ</th>
<th>c₁</th>
<th>c₂</th>
<th>Rg</th>
<th># atoms</th>
<th>fit file</th>
</tr>
</thead>
<tbody>
<tr>
<td>3KFO</td>
<td>✓</td>
<td>2.96</td>
<td>1.02</td>
<td>3.76</td>
<td>16.59</td>
<td>1669</td>
<td>3KFO_23922_merge.dat</td>
</tr>
</tbody>
</table>
Chimera-IMP-FoXS integration

- Tools->Higher Order Structure->SAXS
Chimera-Modeller-FoXS
Modeling full length protein and fitting to SAXS profile

- Chimera version 1.5 or higher
- Prepare alignment file in PIR format:

```
>P1;3KFO
--KIQYNLDTDAEKNIISNKLKKGEVQICKRFKNGSIREVFNILVEELKSTTVNVNLSDLVELYSMLDDEESLFIPL
RLLSVGDNLLNFEVKKFLNALVWRIRVLLNASNEGDLKLLQHIVKRVFDEELPNNDPLPESVLDCDKSLLTPEYI
SETYGRFPIDQNAIREEIYEEISQVETLSNDSLSEIKHLHSTIGSVAKEKNYNTINYNTE
```

```
>P1;3KFO-fill
sequence:::::::::
LRKIQYNLDTDAEKNIISNKLKKGEVQICKRFKNGSIREVFNILVEELKSTTVNVNLSDLVELYSMLDDEESLFIPL
RLLSVGDNLLNFEVKKFLNALVWRIRVLLNASNEGDLKLLQHIVKRVFDEELPNNDPLPESVLDCDKSLLTPEYI
SETYGRFPIDQNAIREEIYEEISQVETLSNDSLSEIKHLHSTIGSVAKEKNYNTINYNTE
```
Chimera-Modeller-FoXS
Modeling full length protein and fitting to SAXS profile

• In Chimera go to Tools-&gt;Sequence-&gt;Multalign Viewer and load the alignment file
• In the alignment viewer go to Structure-&gt;Modeller Tools to open Modeller window
• Select target sequence and template structure and run modeller (it will run in the background)
• Select Tools-&gt;Higher Order Structure-&gt;SAXS to run FoXS
• Select molecule and profile file
AllosMod-FoXS

- upload template PDB files and target sequence

AllosMod-FoXS: Structure Generation and SAXS Profile Calculations

AllosMod-FoXS combines the AllosMod and FoXS web servers. Our combined server allows various sampling algorithms from AllosMod to generate structures that are directly inputted into FoXS for small angle X-ray scattering profile calculations. The server supports modeling of protein, DNA, RNA, and glycosylation. For help, click here.

Developers:
Dina Schneidman
Ben Webb
Patrick Weinkam

Acknowledgements:
Ursula Pieper
Elina Tjoe
Andrei Sali

Version r52:53

---

Job name
Email (Required)

PDB code □ □ □ or upload PDB file Choose File No file chosen

Add more structures

Sequence to be used in simulation (specify protein and DNA/RNA, input sugar in adv. opt., see help page)
Running FoXS from IMP

> foxs
foxs
Usage: <pdb_file1> <pdb_file2> ... <profile_file1> <profile_file2> ...

> foxs *.pdb saxes.dat
1817 atoms were read from PDB file 3KFO-fill.B99990001.pdb
1817 atoms were read from PDB file 3KFO-fill.B99990002.pdb
1817 atoms were read from PDB file 3KFO-fill.B99990003.pdb
1817 atoms were read from PDB file 3KFO-fill.B99990004.pdb
1669 atoms were read from PDB file 3KFO.pdb
Profile read from file saxes.dat size = 456
Computing profile for 3KFO-fill.B99990001.pdb 1817 atoms
3KFO-fill.B99990001.pdb saxes.dat Chi = 2.2455 c1 = 1.0265 c2 = 0.51424 default chi =
Computing profile for 3KFO-fill.B99990002.pdb 1817 atoms
3KFO-fill.B99990002.pdb saxes.dat Chi = 1.16904 c1 = 1.0292 c2 = 0.438976 default chi =
Computing profile for 3KFO-fill.B99990003.pdb 1817 atoms
3KFO-fill.B99990003.pdb saxes.dat Chi = 1.13191 c1 = 1.02405 c2 = 1.43008 default chi =
Computing profile for 3KFO-fill.B99990004.pdb 1817 atoms
3KFO-fill.B99990004.pdb saxes.dat Chi = 1.38553 c1 = 1.02892 c2 = 0.815872 default chi =
Computing profile for 3KFO-fill.B99990005.pdb 1817 atoms
3KFO-fill.B99990005.pdb saxes.dat Chi = 1.20806 c1 = 1.02808 c2 = 0.994432 default chi =
Computing profile for 3KFO.pdb 1669 atoms
3KFO.pdb saxes.dat Chi = 2.96235 c1 = 1.02261 c2 = 3.76 default chi = 9.8793
FoXS as a Restraint in IMP::saxs

m = IMP.kernel.Model()

#! read PDB
mp = IMP.atom.read_pdb(self.get_input_file_name('6lyz.pdb'), m,
    IMP.atom.NonWaterNonHydrogenPDBSelector())

#! read experimental profile
exp_profile = IMP.saxs.Profile(self.get_input_file_name('lyzexp.dat'))

#! select particles from the model
particles = IMP.atom.get_by_type(mp, IMP.atom.ATOM_TYPE)

#! define restraint
saxs_restraint = IMP.saxs.Restraint(particles, exp_profile)
m.add_restraint(saxs_restraint)
score = saxs_restraint.evaluate(False)
Protein assembly application

Experimental SAXS profile

Computed SAXS profiles

Intensity (log scale)

q

compare
Integrative Docking Approach

Input structures

Rigid Docking

Data-guided Scoring

SAXS  EM2D  EM3D  NMR-RTC  CXMS  SOAP

Composite Scoring

~10^{11} models

~10^{6} models

~10^{3} models

10 models

Schneidman-Duhovny D et al. Bioinformatics 2012  salilab.org/idock
Macromolecular Docking with SAXS Profile

Type PDB codes of receptor and ligand molecules or upload files in PDB format

Receptor Molecule:  
(PDB:chainId e.g. 2kai:AB)

Ligand Molecule: 
(PDB:chainId e.g. 2kai:I)

Complex SAXS profile:  
Browse...

e-mail address:

Complex Type:  
Default

Submit Form  Clear

or upload file:  
Browse...

or upload file:  
Browse...

(the results are sent to this address)

Be sure to give receptor and ligand in the corresponding order!


Contact: dina@salilab.org
## Homolecular Docking with SAXS Profile

**Receptor**

- 1hl5E.pdb

**Ligand**

- 1hl5E.pdb

**SAXS Profile**

- hliq.dat

**Complex Type**

- Default

<table>
<thead>
<tr>
<th>Model No</th>
<th>Z-Score</th>
<th>SAXS χ score</th>
<th>Energy score</th>
<th>Transformation</th>
<th>PDB file of the complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-4.510</td>
<td>2.429</td>
<td>-95.020</td>
<td>3.13 0.46 1.96 95.17 -65.06 8.25</td>
<td>result1.pdb view</td>
</tr>
<tr>
<td>2</td>
<td>-3.540</td>
<td>2.354</td>
<td>-55.380</td>
<td>0.78 -0.09 0.19 2.86 -7.95 24.11</td>
<td>result2.pdb view</td>
</tr>
<tr>
<td>3</td>
<td>-3.467</td>
<td>3.318</td>
<td>-72.980</td>
<td>0.19 0.15 -0.12 5.37 7.68 -1.93</td>
<td>result3.pdb view</td>
</tr>
<tr>
<td>4</td>
<td>-3.464</td>
<td>2.559</td>
<td>-56.740</td>
<td>-1.93 0.26 3.13 95.15 -12.63 -30.80</td>
<td>result4.pdb view</td>
</tr>
<tr>
<td>5</td>
<td>-3.449</td>
<td>2.403</td>
<td>-52.850</td>
<td>2.96 0.04 1.93 88.25 -65.23 36.51</td>
<td>result5.pdb view</td>
</tr>
<tr>
<td>6</td>
<td>-3.400</td>
<td>1.843</td>
<td>-39.020</td>
<td>-1.77 0.18 -2.81 94.43 -8.13 -30.04</td>
<td>result6.pdb view</td>
</tr>
<tr>
<td>7</td>
<td>-3.323</td>
<td>1.665</td>
<td>-32.240</td>
<td>0.27 -0.15 -0.19 7.77 12.03 20.69</td>
<td>result7.pdb view</td>
</tr>
<tr>
<td>8</td>
<td>-3.320</td>
<td>2.088</td>
<td>-41.080</td>
<td>-0.12 0.63 2.92 118.57 -38.68 36.53</td>
<td>result8.pdb view</td>
</tr>
<tr>
<td>9</td>
<td>-3.180</td>
<td>4.136</td>
<td>-79.140</td>
<td>-2.91 0.76 1.81 90.65 -62.57 -14.80</td>
<td>result9.pdb view</td>
</tr>
<tr>
<td>10</td>
<td>-3.116</td>
<td>2.625</td>
<td>-44.520</td>
<td>3.14 -0.67 3.06 96.39 1.61 35.80</td>
<td>result10.pdb view</td>
</tr>
<tr>
<td>11</td>
<td>-3.087</td>
<td>2.893</td>
<td>-49.050</td>
<td>-3.13 -0.95 2.75 94.36 -6.74 42.67</td>
<td>result11.pdb view</td>
</tr>
<tr>
<td>12</td>
<td>-3.054</td>
<td>1.776</td>
<td>-24.050</td>
<td>0.60 0.95 -0.10 25.08 6.93 -30.32</td>
<td>result12.pdb view</td>
</tr>
<tr>
<td>13</td>
<td>-2.988</td>
<td>2.725</td>
<td>-41.610</td>
<td>-1.66 0.38 -2.15 60.19 -7.73 -41.67</td>
<td>result13.pdb view</td>
</tr>
<tr>
<td>14</td>
<td>-2.942</td>
<td>3.088</td>
<td>-47.540</td>
<td>-2.04 0.18 2.91 98.86 -19.38 -26.04</td>
<td>result14.pdb view</td>
</tr>
<tr>
<td>15</td>
<td>-2.928</td>
<td>2.968</td>
<td>-44.430</td>
<td>0.00 -0.10 -0.33 7.35 15.26 5.05</td>
<td>result15.pdb view</td>
</tr>
<tr>
<td>16</td>
<td>-2.886</td>
<td>2.334</td>
<td>-29.300</td>
<td>2.17 -0.26 -2.72 92.37 3.32 37.85</td>
<td>result16.pdb view</td>
</tr>
<tr>
<td>17</td>
<td>-2.852</td>
<td>2.919</td>
<td>-40.390</td>
<td>0.50 0.84 -0.02 26.10 4.82 -31.06</td>
<td>result17.pdb view</td>
</tr>
<tr>
<td>18</td>
<td>-2.832</td>
<td>2.904</td>
<td>-39.280</td>
<td>0.39 -0.06 0.30 6.43 -13.72 12.26</td>
<td>result18.pdb view</td>
</tr>
<tr>
<td>19</td>
<td>-2.828</td>
<td>3.668</td>
<td>-55.360</td>
<td>-2.74 -0.83 2.26 96.84 -36.42 51.00</td>
<td>result19.pdb view</td>
</tr>
<tr>
<td>20</td>
<td>-2.821</td>
<td>1.992</td>
<td>-19.470</td>
<td>0.30 -0.65 -2.78 122.17 -14.63 47.27</td>
<td>result20.pdb view</td>
</tr>
</tbody>
</table>

Download output file.
FoXS Dock prediction - profile fit view

Profile file

Experimental profile fit file
\( \chi = 3.30192 \ c_1 = 1.06 \ c_2 = -0.20002 \)
FoXS Dock in IMP::integrative_docking

> idock receptor.pdb ligand.pdb --saxs saxs.dat

• relies on PatchDock (bioinfo3d.cs.tau.ac.il) for docking
• relies on SOAP statistical potential for scoring
• other types of information supported:
  ✓ 3D electron microscopy density maps
  ✓ Electron microscopy 2D class averages
  ✓ Cross links
  ✓ Binding sites
  ✓ Distance constraints
Uncertainty in integrative structure modeling

Information → Scoring → Sampling → Analysis

Data satisfied?

- yes: Single structure
- no: Multiple structures

Sample heterogeneity

More data is needed
DNA ligase with DNA

Essential functions in nuclear and mitochondrial DNA replication and repair
Disorder Comes in Flavors

“rigid”  “flexible”  “disordered”

short fragments (≥10 and <30 residues)  long fragments (> 40 residues)

<table>
<thead>
<tr>
<th></th>
<th>PDB</th>
<th>SwissProt</th>
</tr>
</thead>
<tbody>
<tr>
<td>~ 40% of structures</td>
<td>~ 10% of structures</td>
<td>&gt; 25% of sequences</td>
</tr>
</tbody>
</table>

(Romero et al. 2001; Dunker et al 2000; Le Gall 2007)
Two stages in modeling

sampling of conformations

finding (all) multi-state models that fit SAXS data within noise

How to choose the sampling method?

Rapidly exploring Random Trees (RRT)

How to enumerate the relevant multi-state models?

Branch & bound enumeration

How to score them?

Non-negative least-square fitting
Multi-state modeling in IMP with SAXS

IMP::kinematics sampling of conformations

IMP::multi_state finding (all) multi-state models that fit SAXS data within noise

# generate 10000 conformations sampling torsion angles of residues in linker.txt
> rrt_sample ligase.pdb linker.txt -i 100000 -n 10000

# pre-calculate SAXS profiles (-p pre-calculate, -m reads each MODEL into separate structure), run in parallel for all PDBs
> foxs -p -m 2 nodes*.pdb

# generate a file with the name of all the pre-calculated SAXS profiles (ls *.pdb.dat > filenames)
> multi_foxs saxs.dat filenames
Multi-state modeling with SAXS online

NEW! Search for Minimal Ensemble (MES) when you upload multiple PDB files.
NEW! Interactive display for profile plots and input structures.

If you use FoXS, please cite:
Contact: dina@sallab.org
DNA ligase with DNA
Important SAXS Links

salilab.org/foxs  FoXS and MultiFoXS
salilab.org/foxsdock  FoXSDock
http://www.bioisis.net/  - datasets, tutorials, and software
http://sibyls.als.lbl.gov/  - Sibyls beamline
http://www.embl-hamburg.de/biosaxs/software.html  - ATSAS software
http://www.sasbdb.org/  - datasets

More in the eposter today
Thanks!

Seung Joong Kim
Riccardo Pellarin
Pat Weinkam
Barak Raveh
Yannick Spill
Guangqiang Dong
Natalia Khuri
Daniel Russel
Elina Tjioe
Ben Web
Ursulla Pieper
Andrej Sali

Michal Hammel
Greg Hura
Rob Rambo
Susan Tsutakawa
John Tainer
Lester Carter

UCSF Chimera group
Yang Zheng
Elaine Meng
Thomas Goddard
Thomas Ferrin
Thanks to our users!