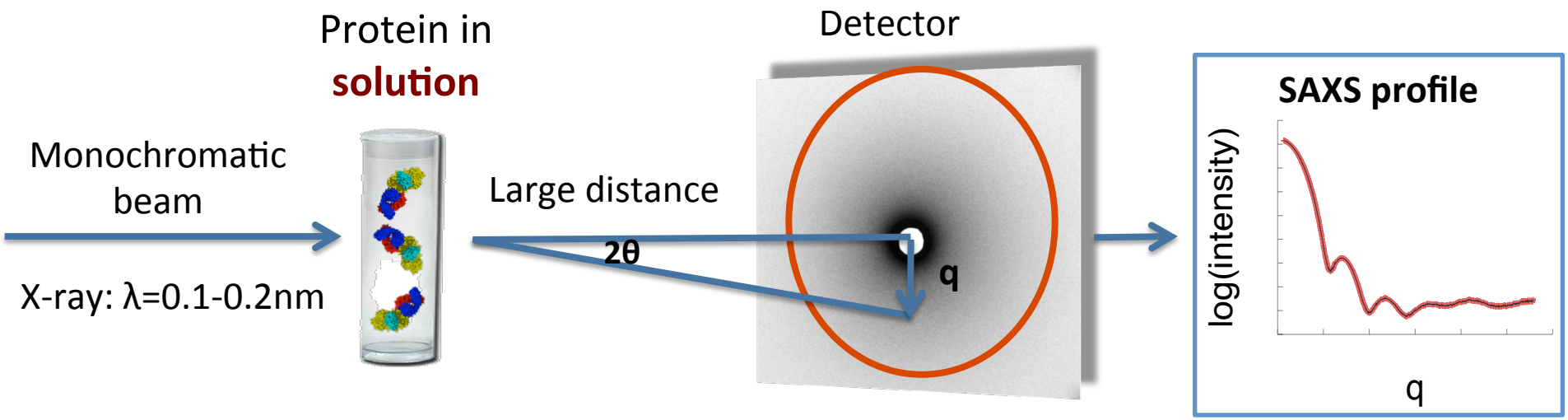
Small Angle X-ray Scattering (SAXS) software tools in Integrative Modeling Platform (IMP)

Dina Schneidman

3/6/2015

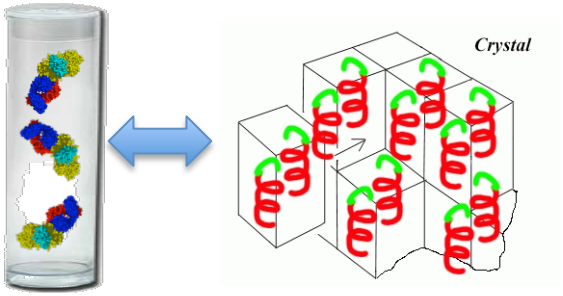
Small Angle X-Ray Scattering



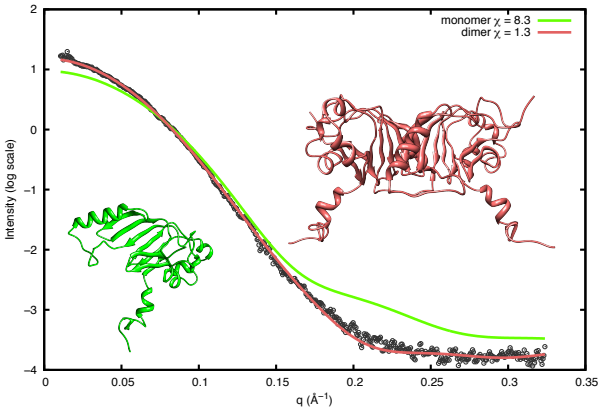
Data can be collected in several **minutes!**

Integrative Modeling and SAXS

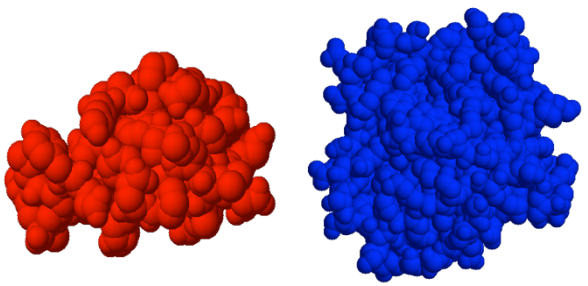
solution structure vs. crystal structure



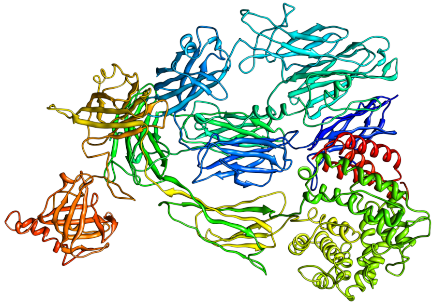
monomer, dimer or mixture?



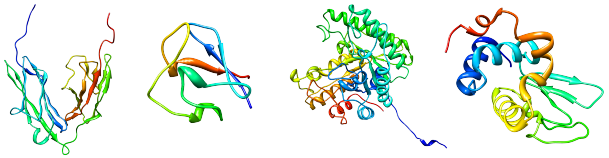
assembly of multi protein complexes



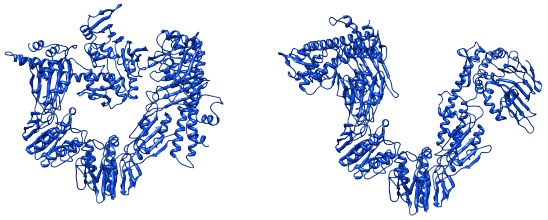
assembly of multi domain proteins



protein folding



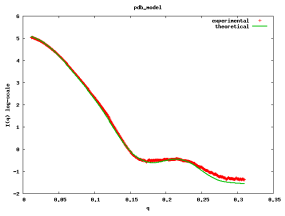
structural characterization of protein dynamics



Integrative Modeling and SAXS

Gathering information

standards for data validation



sparse data

sample heterogeneity

Designing model representation and evaluation

atomic representation

Debye formula
FoXS

Debye formula
MultiFoXS

Sampling good models

Docking, conformational sampling …

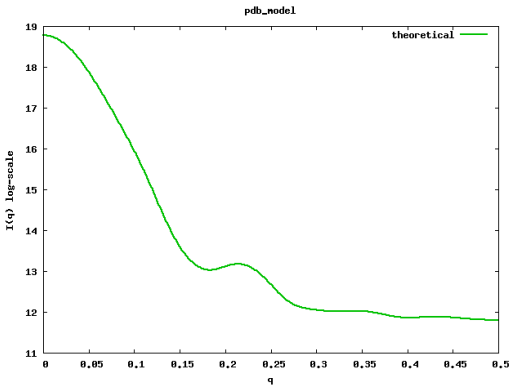
Analyzing models and information

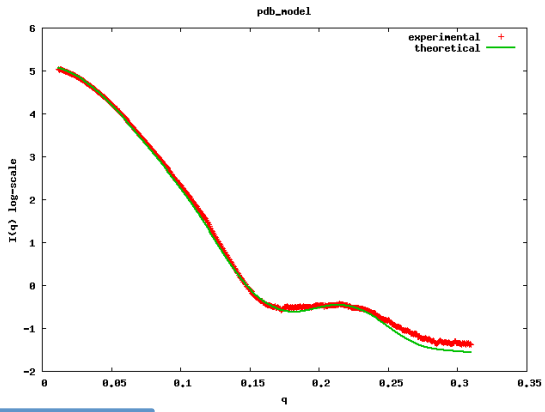
Fit to data, variance among top scoring models

Fast open-source X-ray Scattering

forward model

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





Compute theoretical profile

Fit experimental profile

Debye formula (1915)

$$I_m(q) = \sum_i \sum_j f_i(q) f_j(q) \frac{\sin(q d_{ij})}{q d_{ij}}$$

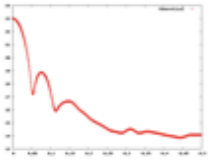
$$\chi^2 = \frac{1}{N-1} \sum_i \left[\frac{I_{\exp}(q_i) - c I_m(q_i)}{\sigma(q_i)} \right]^2$$

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

Schneidman-Duhovny D, Hammel M, Tainer J, Sali A. Biophys J 2013

FoXS

Fast SAXS Profile Computation with Debye Formula

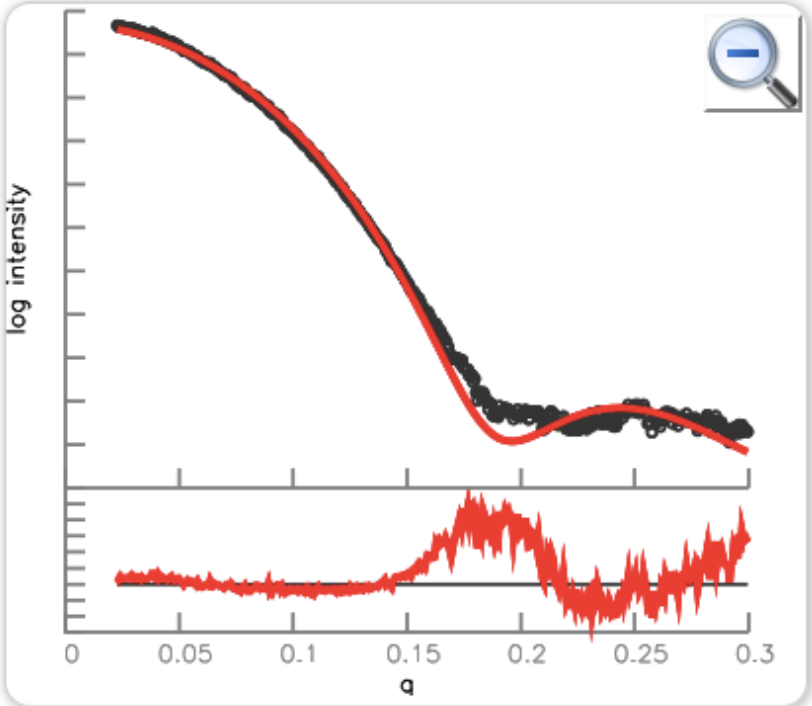


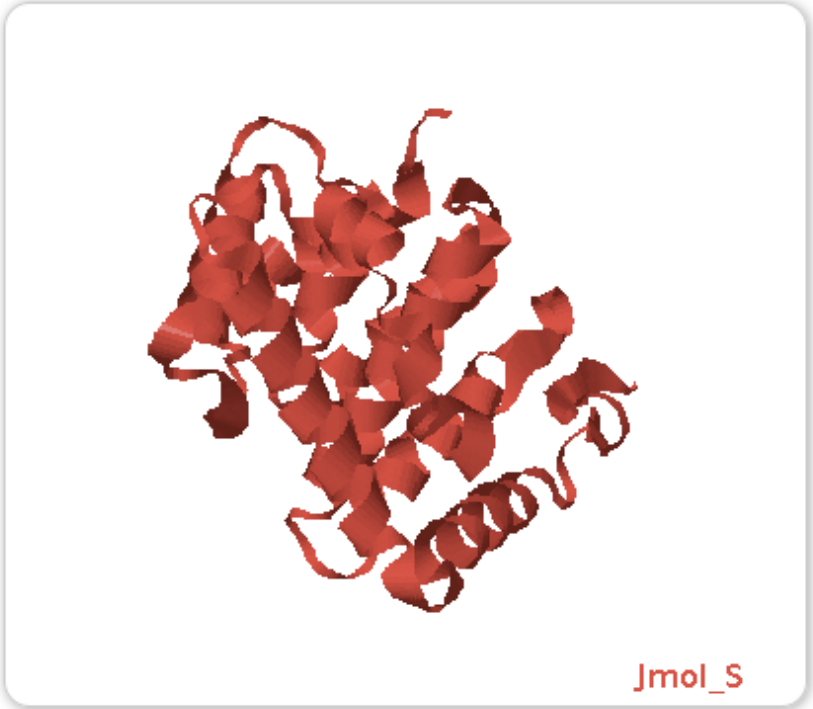
• About FOXS • Web Server • Help • FAQ • Download • Sali Lab • IMP • Links

PDB files
3KFO.pdb

Profile file
23922_merge.dat

Can't see interactive display? Use old interface

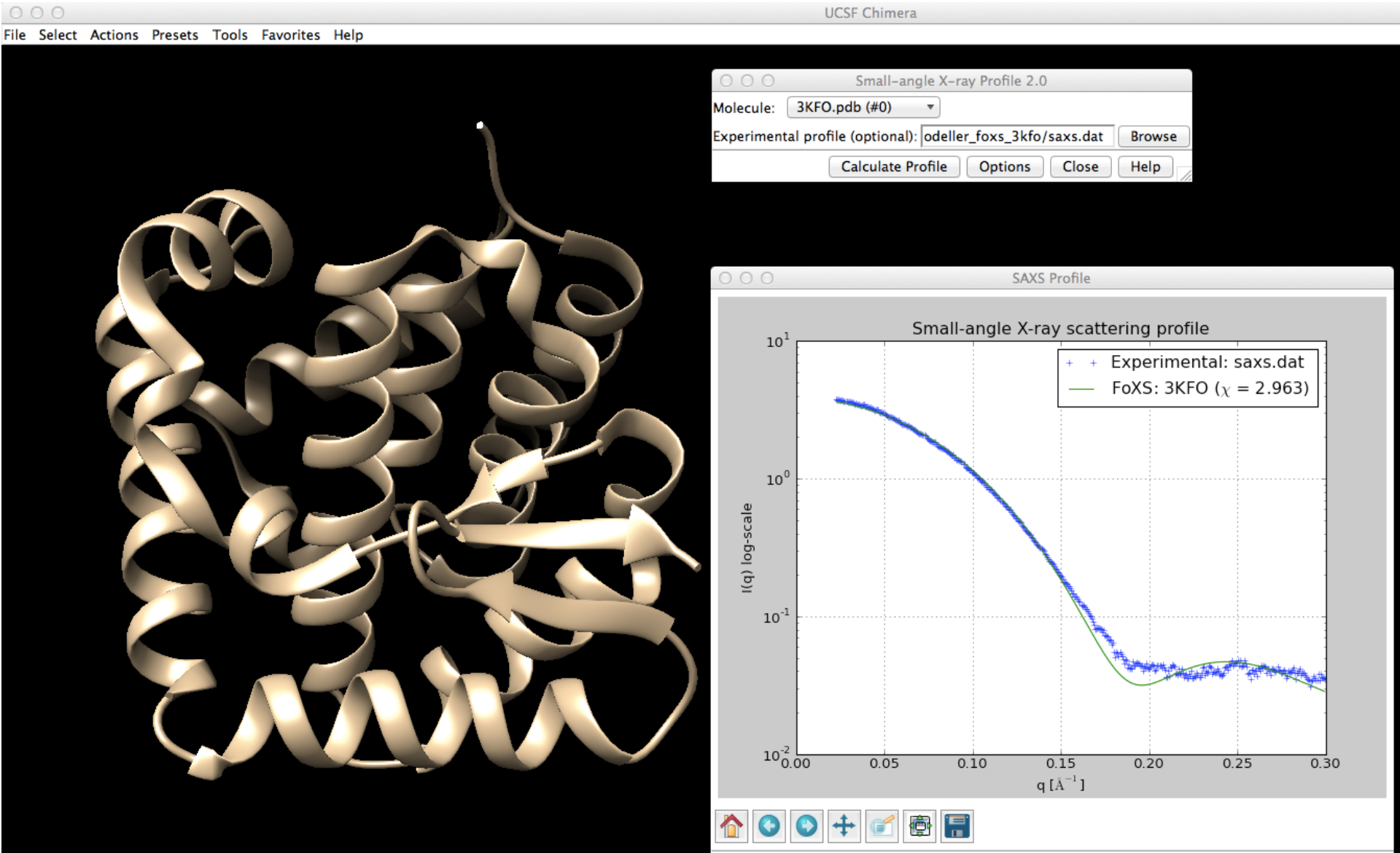




PDB file show/hide		χ	c_1	c_2	R_g	# atoms fit file
3KFO	☑	2.96	1.02	3.76	16.59	1669 3KFO_23922_merge.dat

Chimera-IMP-FoXS integration

- Tools->Higher Order Structure->SAXS



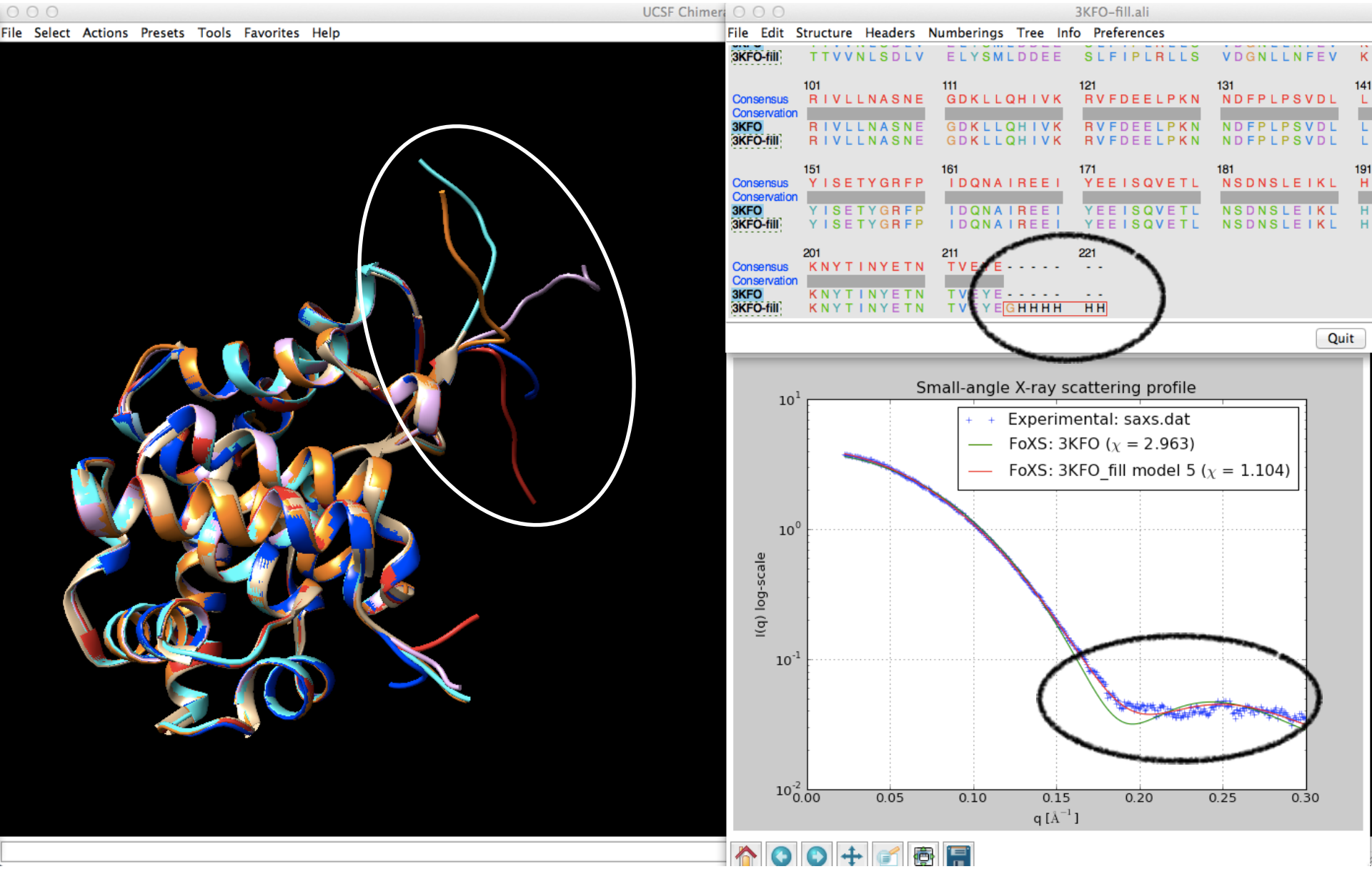
Chimera-Modeller-FoXS Modeling full length protein and fitting to SAXS profile

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

```
>P1;3KFO
structureX:3KFO: 946 :A:+213 :A:NUP133:XXX: 1.90: 0.19
--KIQYNLDTIDAEKNISNKLKKGEVQICKRFKNGSIREVFNILVEELKSTTVVNLSDLVELYSMLDDEESLFIPL
RLLSVDGNLLNFEVKKFLNALVWRRIVLLNASNEGDKLLQHIVKRVFDEELPKNNDFPLPSVDLLCDKSLLTPEYI
SETYGRFPIDQNAIREEIYEEISQVETLNSDNSLEIKLHSTIGSVAKEKNYTINYETNTVEYE-------*
>P1;3KFO-fill
sequence::::::::
LRKIQYNLDTIDAEKNISNKLKKGEVQICKRFKNGSIREVFNILVEELKSTTVVNLSDLVELYSMLDDEESLFIPL
RLLSVDGNLLNFEVKKFLNALVWRRIVLLNASNEGDKLLQHIVKRVFDEELPKNNDFPLPSVDLLCDKSLLTPEYI
SETYGRFPIDQNAIREEIYEEISQVETLNSDNSLEIKLHSTIGSVAKEKNYTINYETNTVEYEGHHHHHH*
```

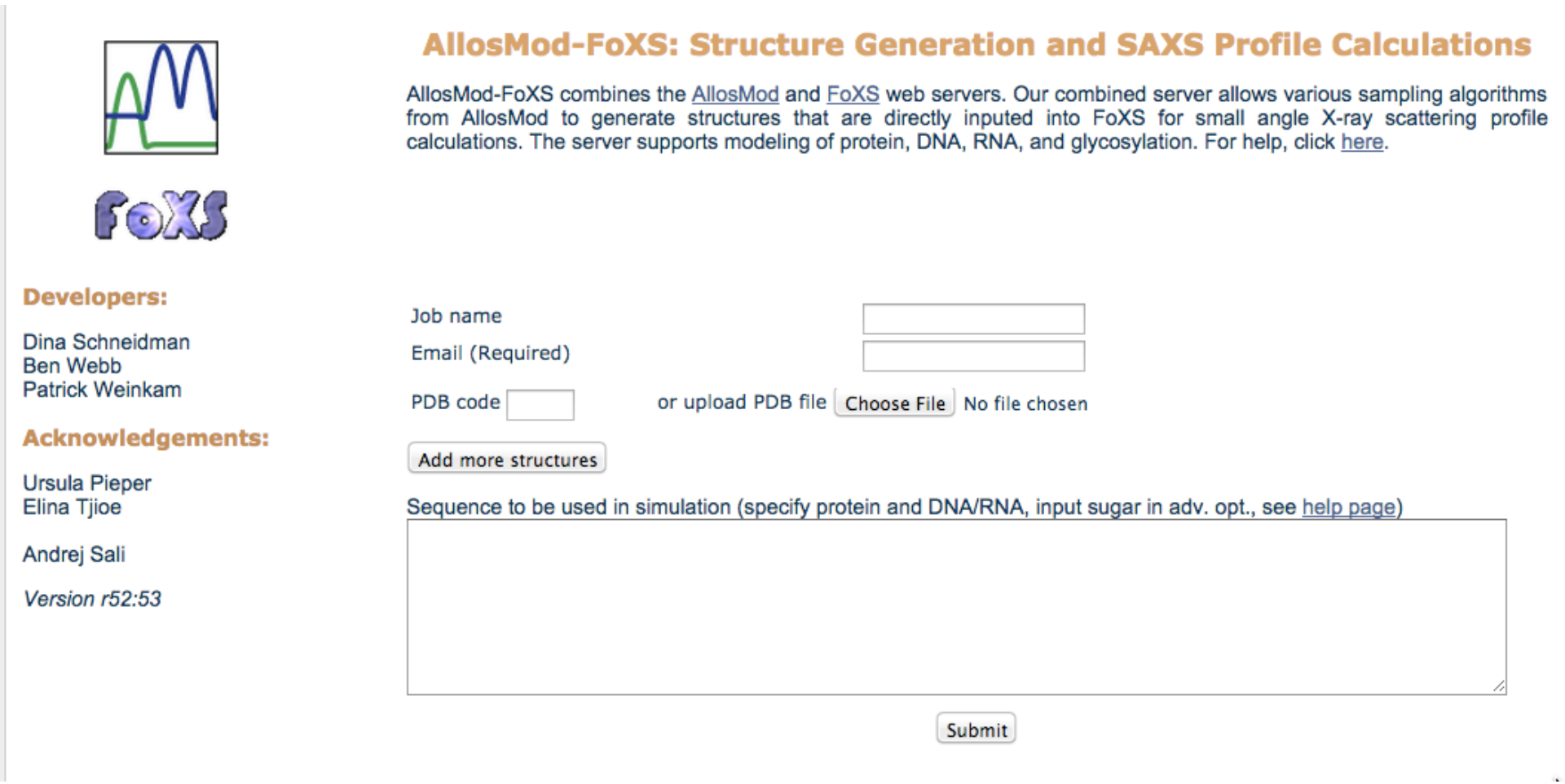
Chimera-Modeller-FoXS Modeling full length protein and fitting to SAXS profile

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


UCSF Chimera
File Select Actions Presets Tools Favorites Help
3KFO-fill.ali
File Edit Structure Headers Numberings Tree Info Preferences
Consensus
Conservation
3KFO
3KFO-fill
Quit
Small-angle X-ray scattering profile
Experimental: saxs.dat
FoXS: 3KFO (χ = 2.963)
FoXS: 3KFO_fill model 5 (χ = 1.104)
I(q) log-scale
q [Å^{-1}]

AllosMod-FoXS

- upload template PDB files and target sequence



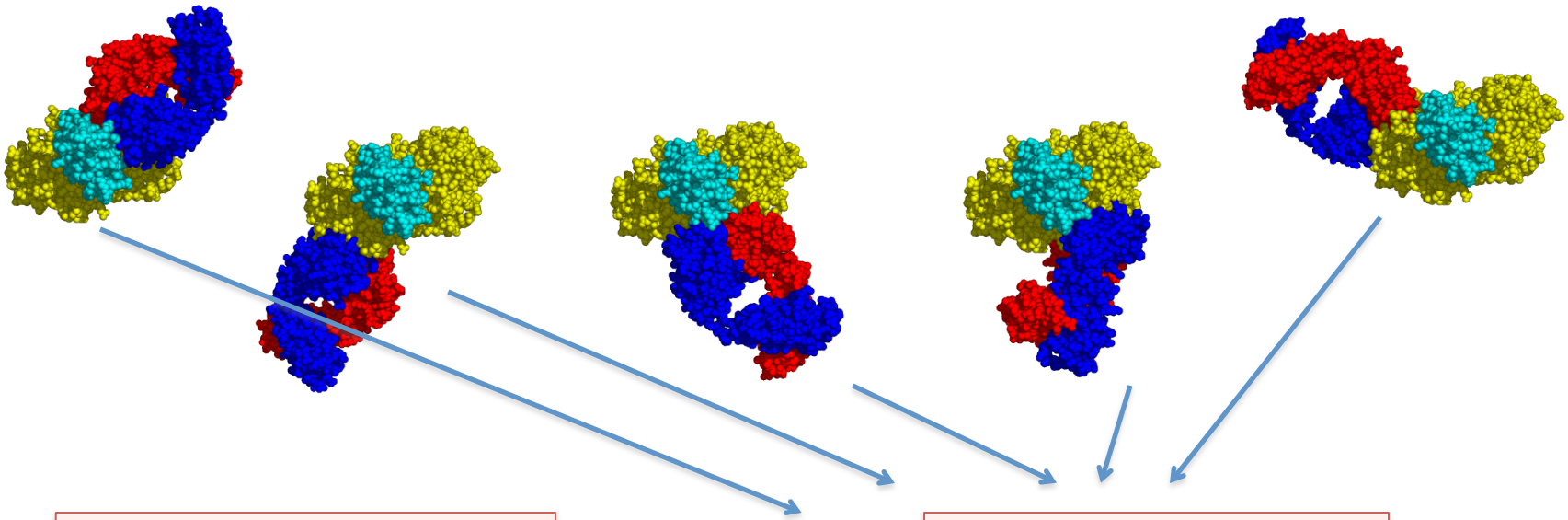
Running FoXS from IMP

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

> foxs *.pdb saxs.dat
foxs 3KFO-fill.B99990001.pdb 3KFO-fill.B99990002.pdb 3KFO-fill.B99990003.pdb 3KFO-fil
fill.B99990005.pdb 3KFO.pdb saxs.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
1817 atoms were read from PDB file 3KFO-fill.B99990005.pdb
1669 atoms were read from PDB file 3KFO.pdb
Profile read from file saxs.dat size = 456
Computing profile for 3KFO-fill.B99990001.pdb 1817 atoms
3KFO-fill.B99990001.pdb saxs.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 saxs.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 saxs.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 saxs.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 saxs.dat Chi = 1.20806 c1 = 1.02808 c2 = 0.994432 default chi
Computing profile for 3KFO.pdb 1669 atoms
3KFO.pdb saxs.dat Chi = 2.96235 c1 = 1.02261 c2 = 3.76 default chi = 9.8793
```

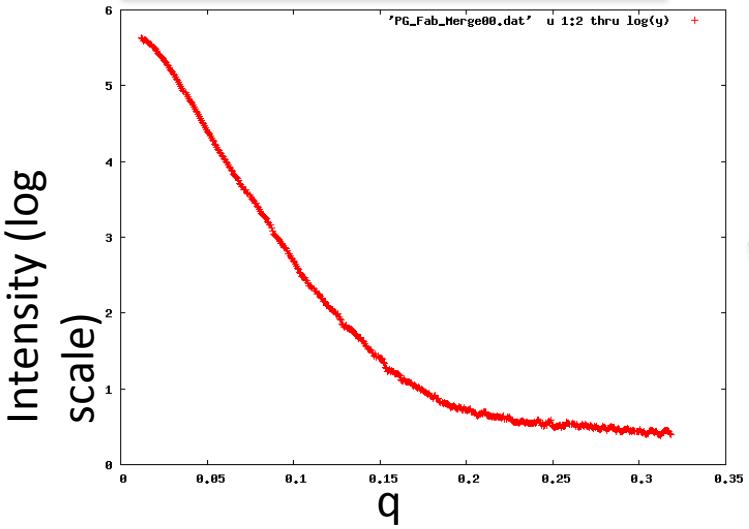
FoXS as a Restraint in IMP::saxs

```
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

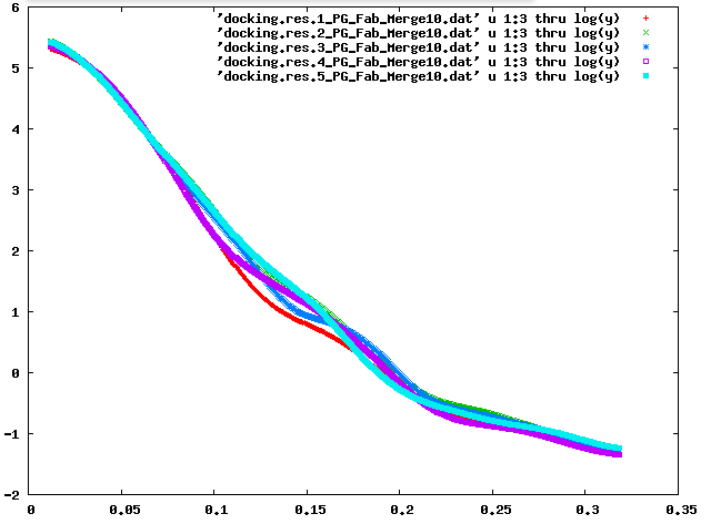


Experimental SAXS profile



compare

Computed SAXS profiles



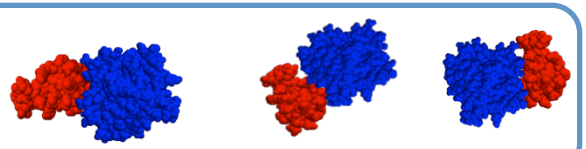
Integrative Docking Approach

Input structures 

~10^{11} models

Rigid Docking

PATCHDOCK



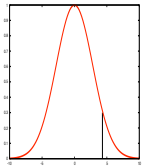
~10^{6} models

Data-guided Scoring

SAXS	EM2D	EM3D	NMR-RTC	CXMS	SOAP
FoXS					

~10^{3} models

Composite Scoring



10 models

Schneidman-Duhovny D et al. Bioinformatics 2012 salilab.org/idock

http://salilab.org/foxsdock

FoXS Dock

Macromolecular Docking with SAXS Profile

• About FOXSDock • Web Server • Help • FAQ • Download • FoXS • Sali Lab • IMP • Links

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

Receptor Molecule: (PDB:chainId e.g. 2kai:AB) **or upload file:** Browse...

Ligand Molecule: (PDB:chainId e.g. 2kai:I) **or upload file:** Browse...

Complex SAXS profile: Browse...

e-mail address: (the results are sent to this address)

Complex Type: Default

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

Submit Form Clear

Schneidman-Duhovny D, Hammel M, Sali A. Macromolecular docking restrained by a small angle X-ray scattering profile. J Struct Biol. 2010 [Abstract]

Contact: dina@salilab.org

FoXS Dock

omolecular Docking with SAXS Profile

Receptor	Ligand	SAXS Profile	Complex Type
1hl5F.pdb	1hl5E.pdb	hliq.dat	Default

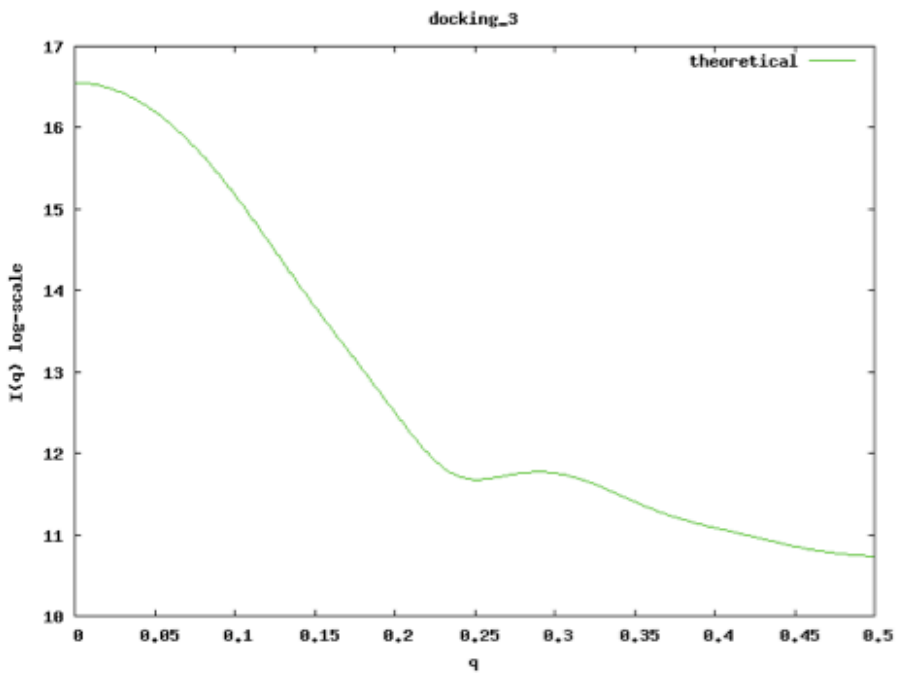
Model No	Z-Score	SAXS χ score	Energy score	Transformation	PDB file of the complex	
1	-4.510	2.429	-95.020	3.13 0.46 1.96 95.17 -65.06 8.25	result1.pdb	view
2	-3.540	2.354	-55.380	0.78 -0.09 0.19 2.86 -7.95 24.11	result2.pdb	view
3	-3.467	3.318	-72.980	0.19 0.15 -0.12 5.37 7.68 -1.93	result3.pdb	view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

»» show next 20

Download output file.

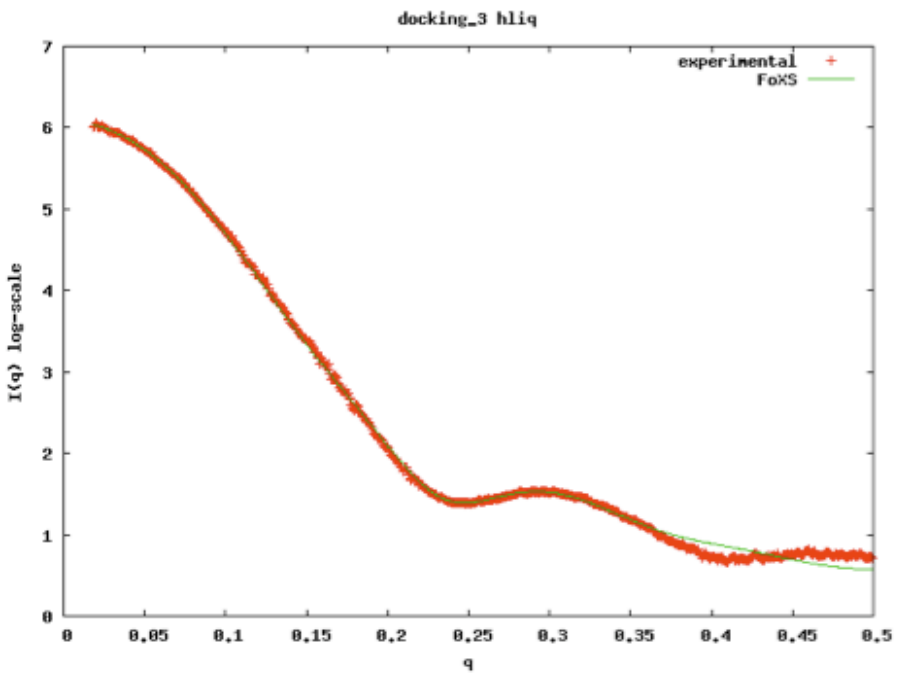
FoXSDock prediction - profile fit view

docking_3 Theoretical profile



Profile file

docking_3 Fit to experimental profile



Experimental profile fit file

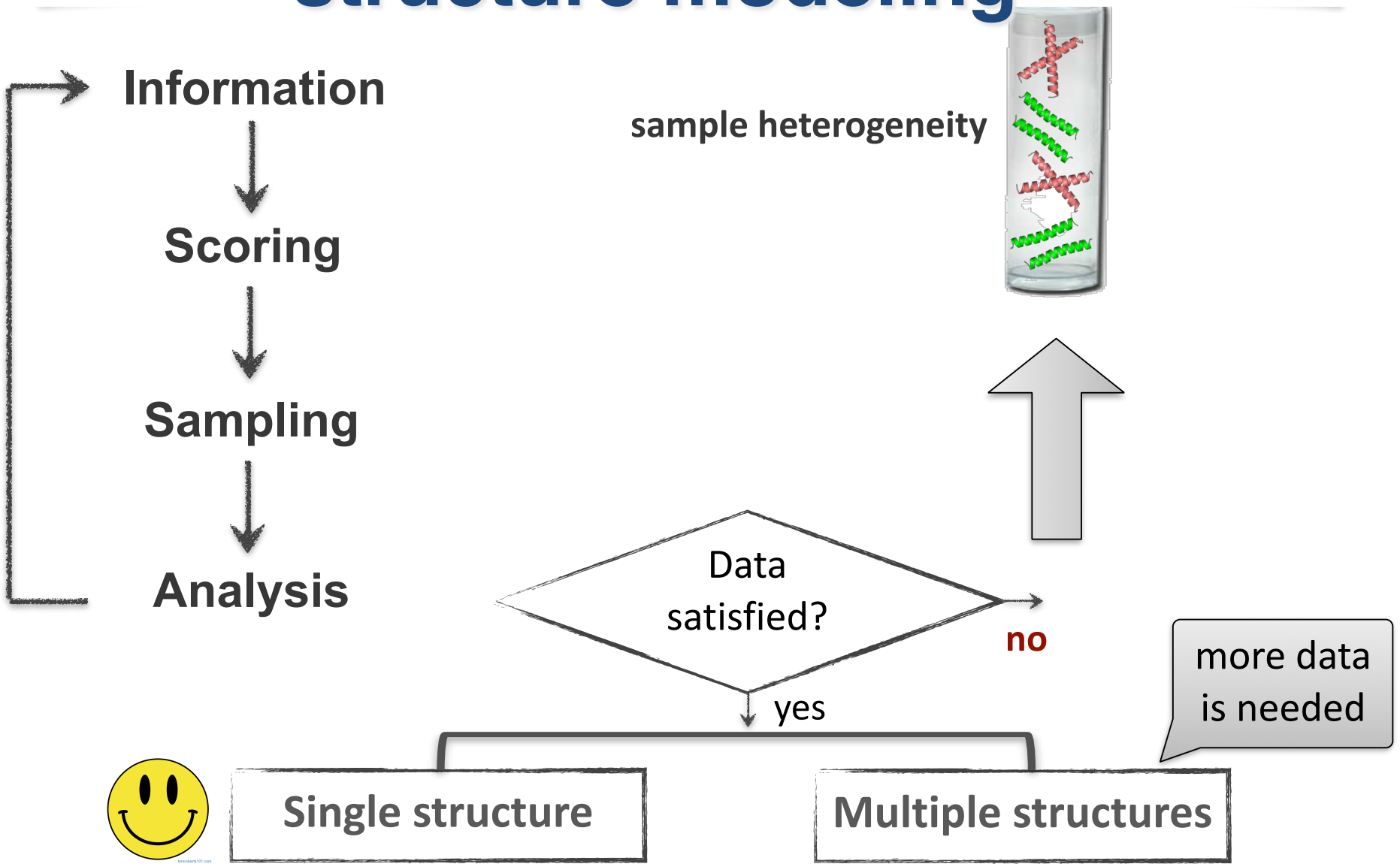
χ= 3.30192 c1 = 1.06 c2 = -0.200002

FoXSDock 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

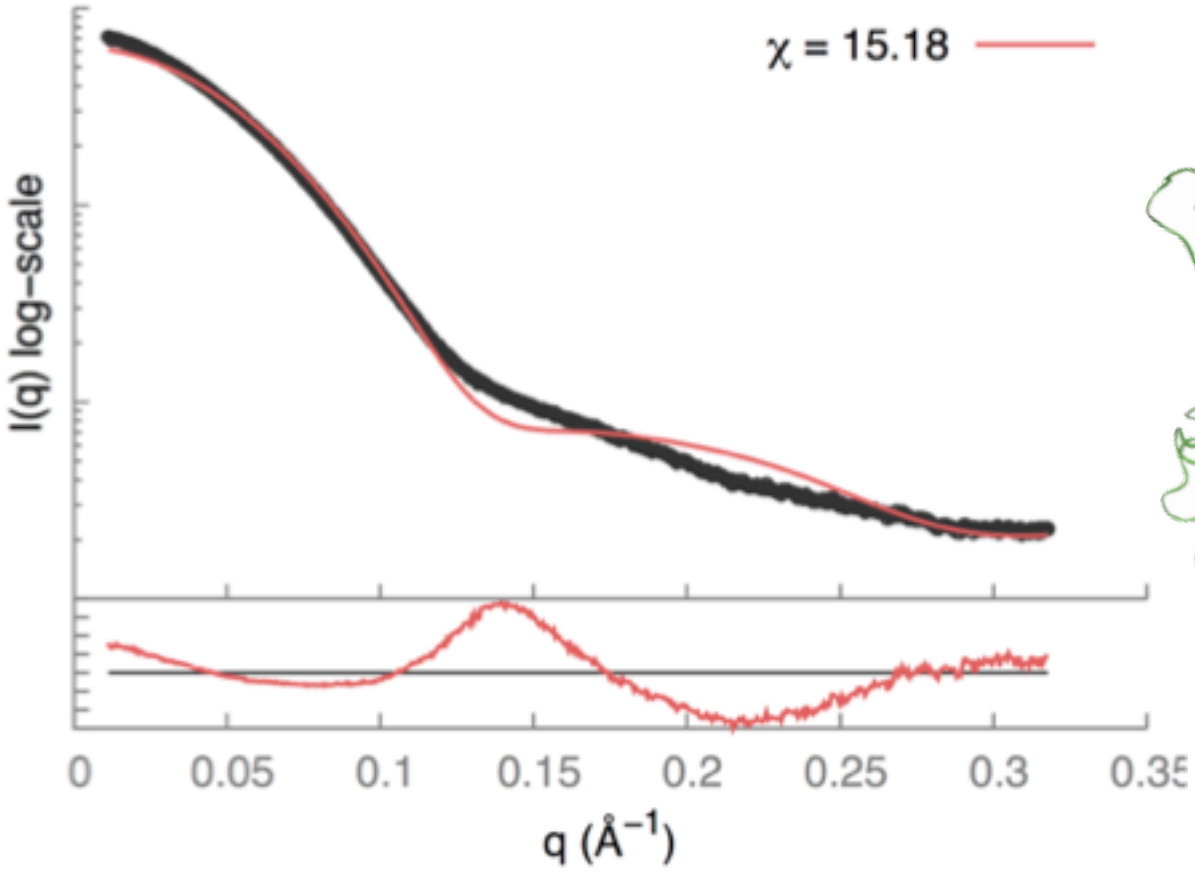
Uncertainty in integrative structure modeling

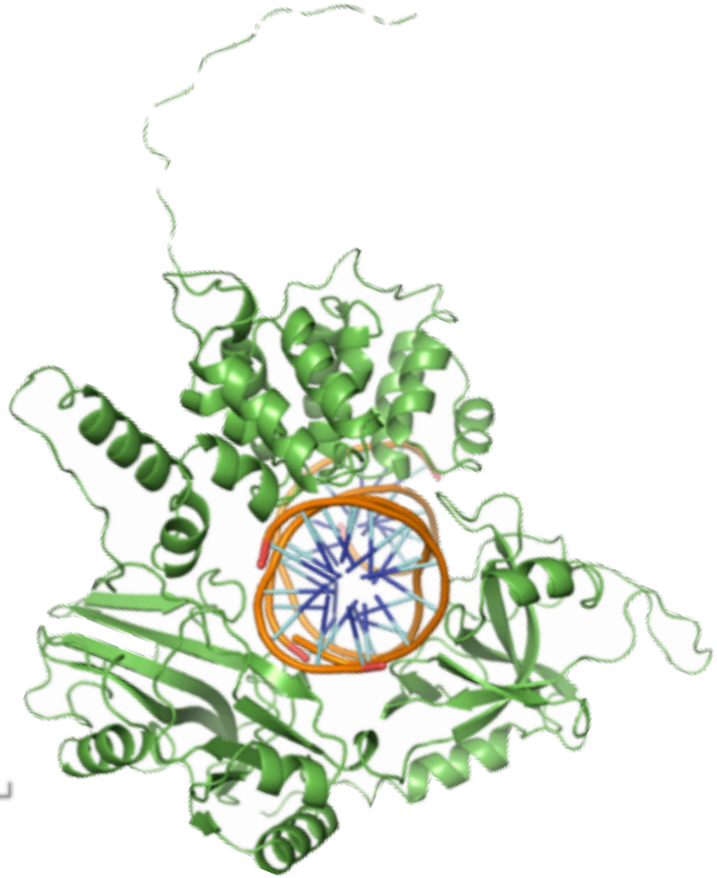


Schneidman-Duhovny, Pellarin, Sali. *COSB* 2014

DNA ligase with DNA

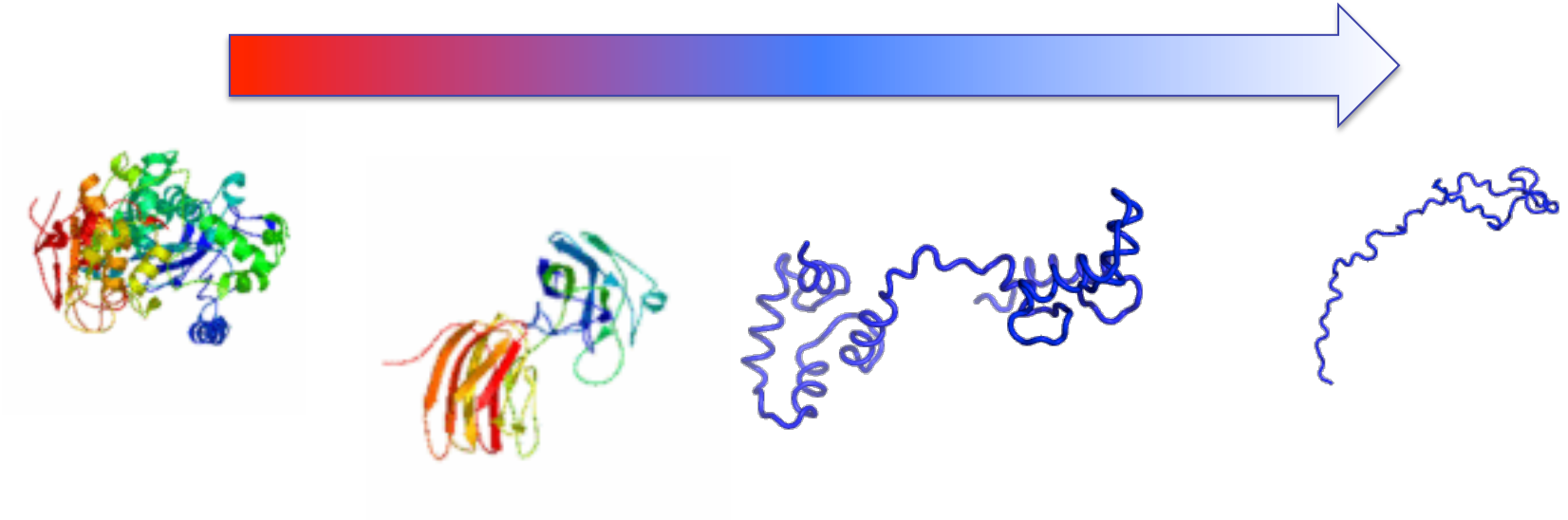
Essential functions in nuclear and mitochondrial DNA replication and repair





Disorder Comes in Flavors

"rigid" "flexible" "disordered"



	short fragments (≥10 and <30 residues)	long fragments (> 40 residues)
PDB	~ 40% of structures	~ 10% of structures
SwissProt		> 25% of sequences

(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

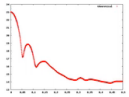
IMP::multi_state

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

FoXS



Fast SAXS Profile Computation with Debye Formula

• About FOXS • Web Server • Help • FAQ • Download • Sali Lab • IMP • Links

Type PDB code of input molecule or upload files in PDB format (zip file with several PDBs can be uploaded):

Input molecule: (PDB:chainId e.g. 6lyz:A) **or** upload file: Choose File ligase.zip

Experimental profile: Choose File iq.dat (optional) sample input

e-mail address: (optional, the results are sent to this address)

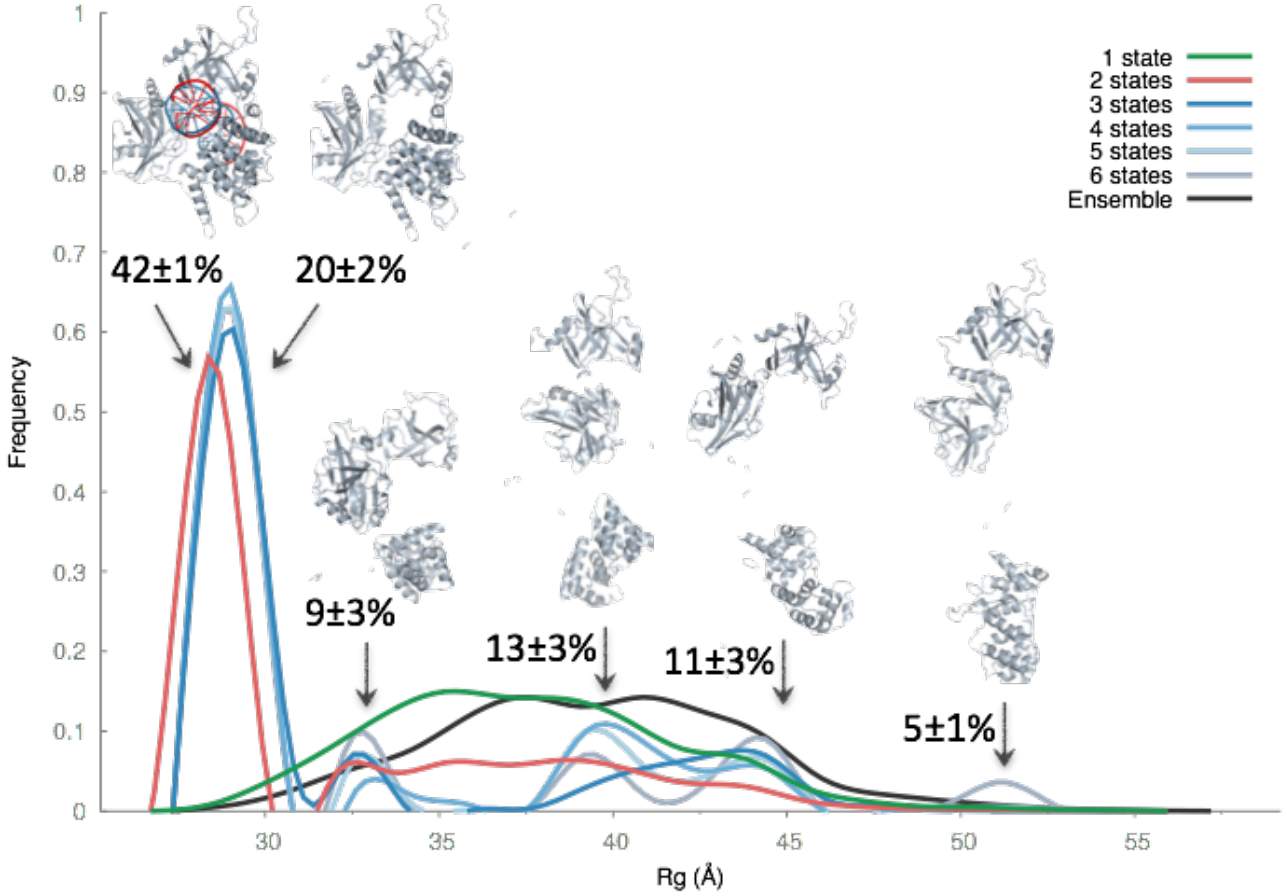
Submit Form Clear

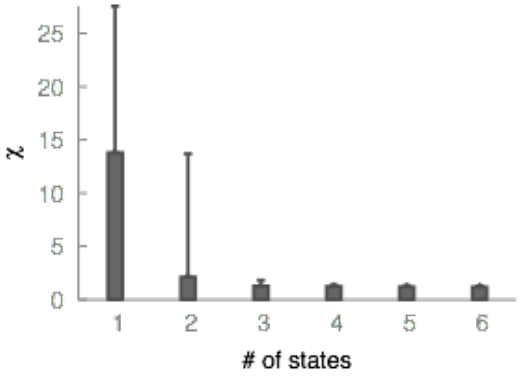
Advanced Options

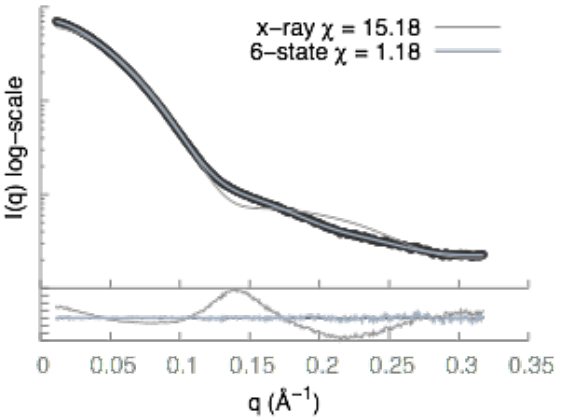
NEW! Search for Minimal Ensemble (MES) when you upload multiple PDB files.
NEW! Interactive display for profile plots and input structures.

If you use FoXS, please cite:
D. Schneidman-Duhovny, M. Hammel, JA. Tainer, and A. Sali. Accurate SAXS profile computation and its assessment by contrast variation experiments. Biophysical Journal 2013. 105 (4), 962-974
D. Schneidman-Duhovny, M. Hammel, and A. Sali. FoXS: A Web server for Rapid Computation and Fitting of SAXS Profiles. NAR 2010. 38 Suppl:W540-4 [FREE Full Text]
Contact: dina@salilab.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!

UCSF University of California San Francisco

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

BERKELEY LAB
LAWRENCE BERKELEY NATIONAL LABORATORY

Michal Hammel
Greg Hura
Rob Rambo
Susan Tsutakawa
John Tainer

SLAC

Lester Carter

UCSF University of California San Francisco

UCSF Chimera group
Yang Zheng
Elaine Meng
Thomas Goddard
Thomas Ferrin

Thanks to our users!