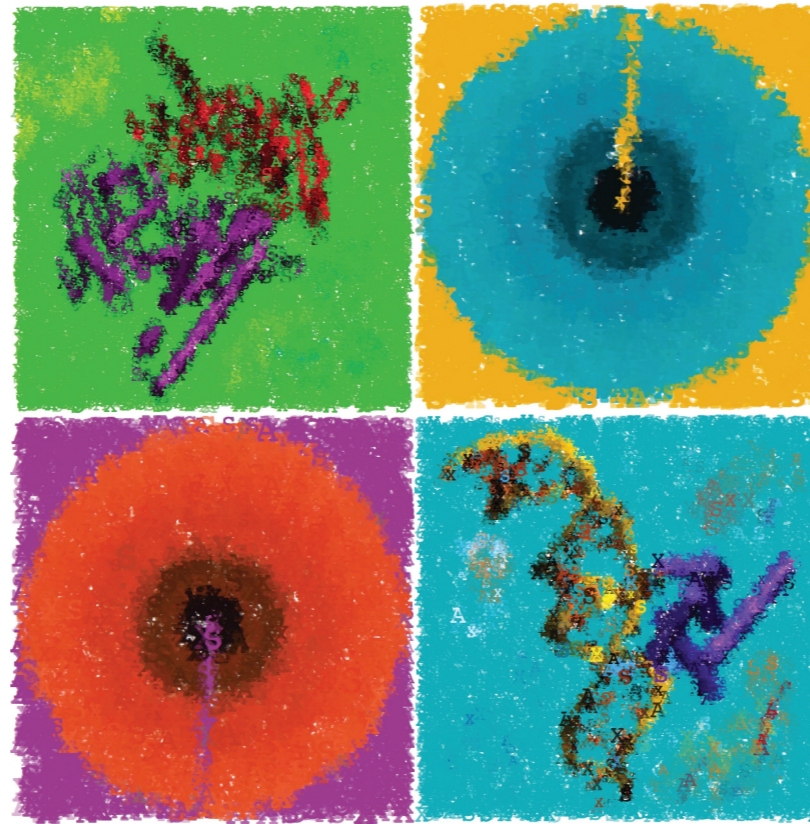


Integrative structural modeling using SAXS data



Dina Schneidman

האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM



Integrative Structure Modeling

Information

X-ray
crystallography

NMR
spectroscopy

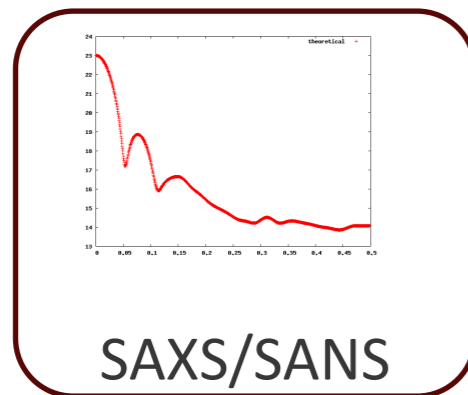
cross-linking
mass
spectrometry

affinity
purification

statistical
potential

Electron
microscopy

AlphaFold2
RosettaFold

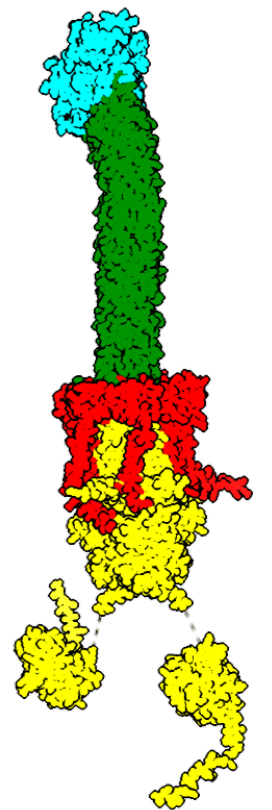


FRET site-directed
mutagenesis

Computational methods

```
while (no structure)  
  get_more_data();
```

Structural models



Ward A, Sali A, Wilson I. Integrative structural biology. Science 2013.

Rout M, Sali A. Principles for Integrative Structural Biology Studies. Cell 2019.

Thanks!

Lab

Edan Patt

Tomer Cohen

Ben Shor

Alon Aronson

Matan Halfon

Amir Weinfeld

Tanya Hochner

Adi Weshler

Gal Passi

Meitar Sela

Danielle Zaccai

Merav Braitbard

Jerome Tubiana

Shon Cohen



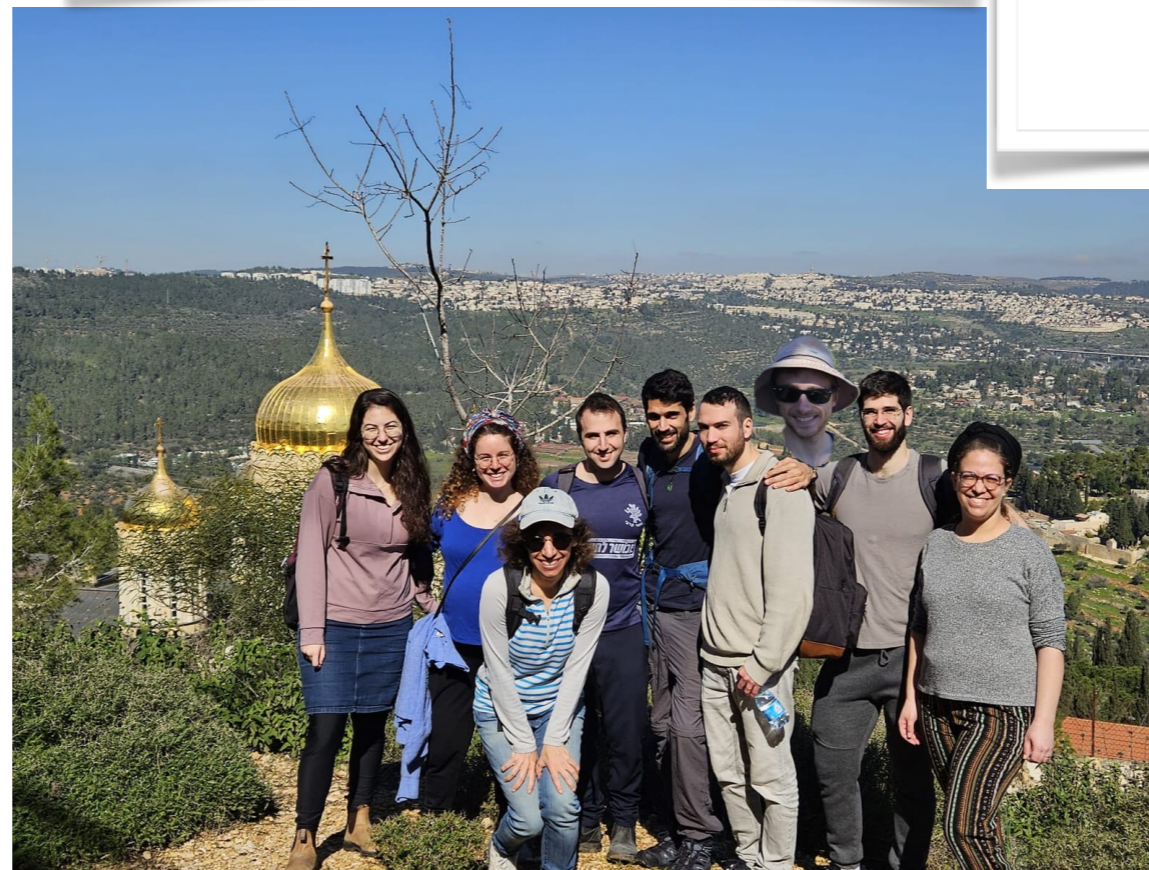
Michal Hammel

Greg Hura

Rob Rambo

Susan Tsutakawa

John Tainer



Nanobodies

Yufei Xiang

Zhe Sang

Yi Shi



Icahn
School of
Medicine at
Mount
Sinai

EGFR antibodies

Arvind Sivasubramanian



EIC
Pathfinder



ISRAEL
SCIENCE
FOUNDATION



MINERVA
STIFTUNG



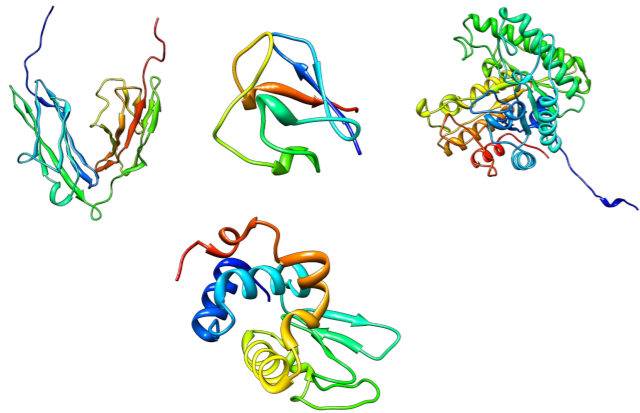
National Institute of
Allergy and
Infectious Diseases

משרד המדע
והטכנולוגיה
Ministry of Science & Technology

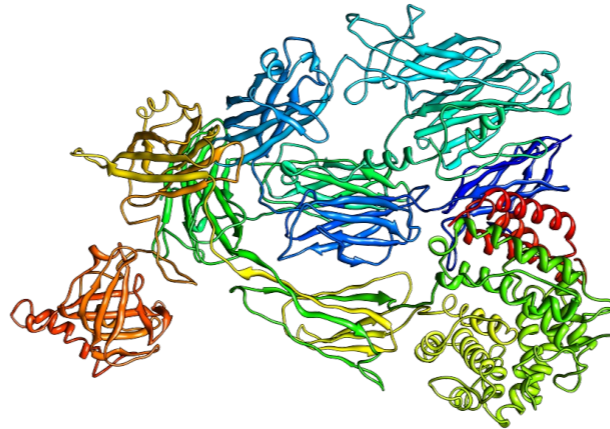


What are the modeling tasks to address with SAXS?

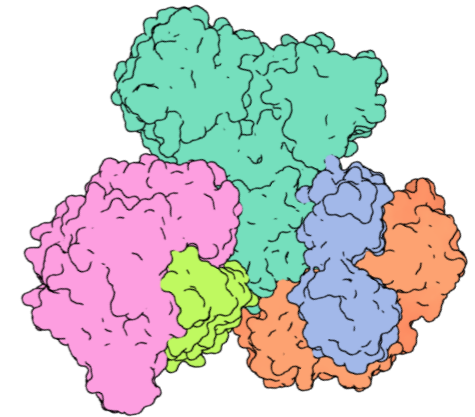
validation of protein structure prediction



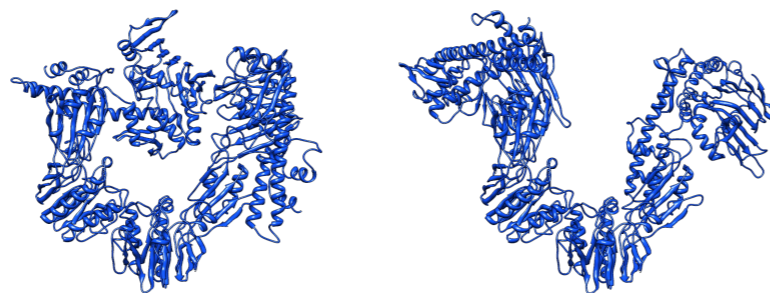
assembly of **multi-domain** proteins



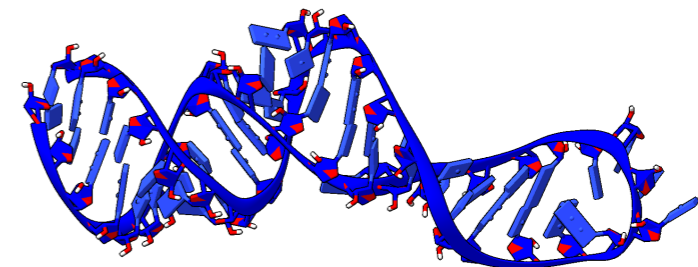
assembly of multi-protein **complexes**



structural characterization of **protein dynamics**



RNA

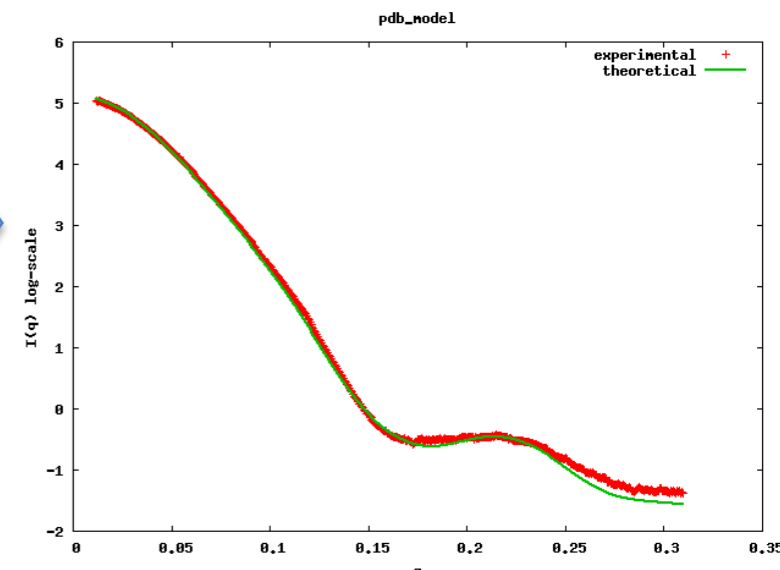
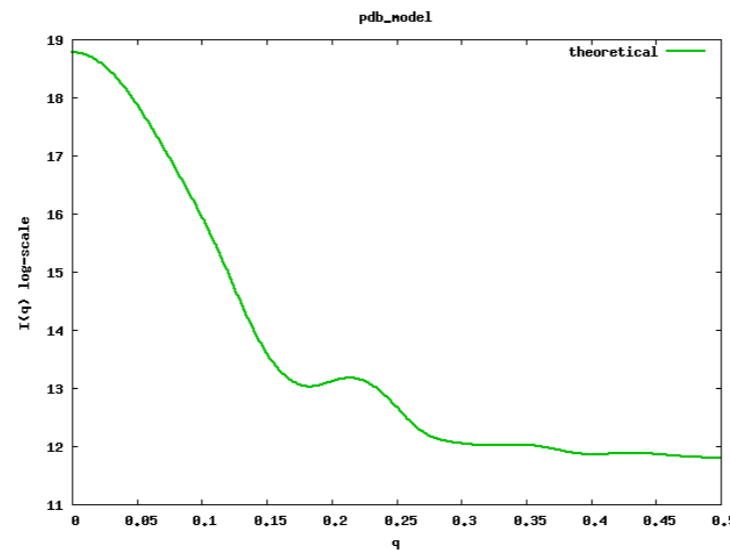
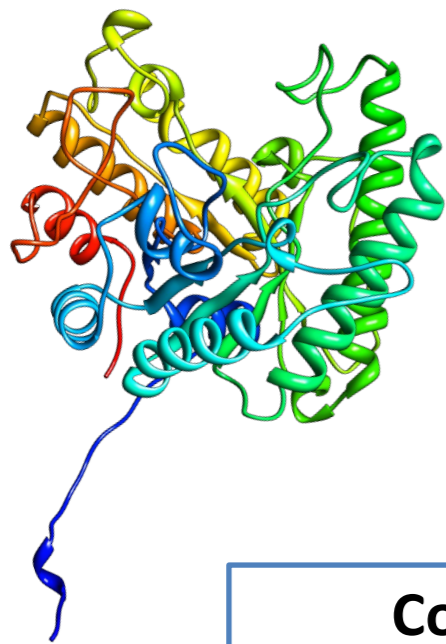


Scoring: Fast open-source X-ray Scattering

Forward modeling

foXS

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



Compute a theoretical profile

Fit to the experimental profile and compute the score

Debye formula (1915)

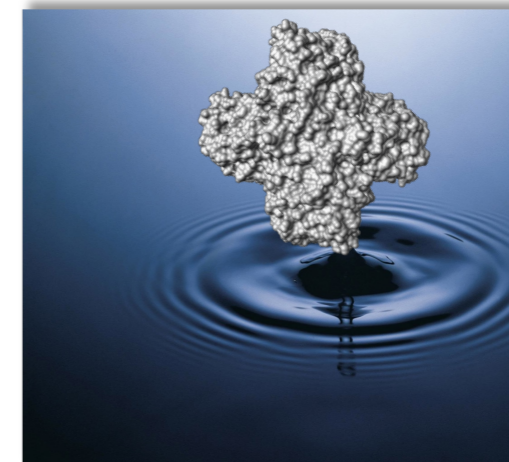
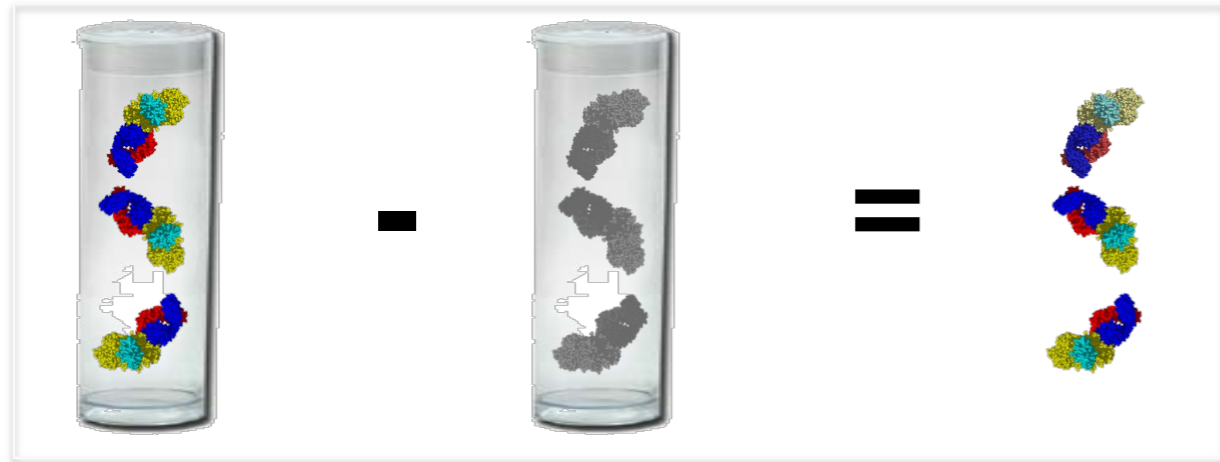
$$I(q) = \sum_{i=1}^N \sum_{j=1}^N f_i(q) f_j(q) \frac{\sin(qd_{ij})}{qd_{ij}}$$

$$\chi = \sqrt{\frac{1}{M} \sum_{i=1}^M \left(\frac{I_{exp}(q_i) - cI(q_i)}{\sigma(q_i)} \right)^2}$$

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

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

Scoring: Excluded Volume and Hydration Layer Density



vacuum

solvent excluded volume

hydration layer

$$f_i(q) = f_i^v(q) - C_1(q) f_i^s(q) + c_2 s_i f^w(q)$$

Increase/decrease atomic radii to obtain the best fit to the experimental profile

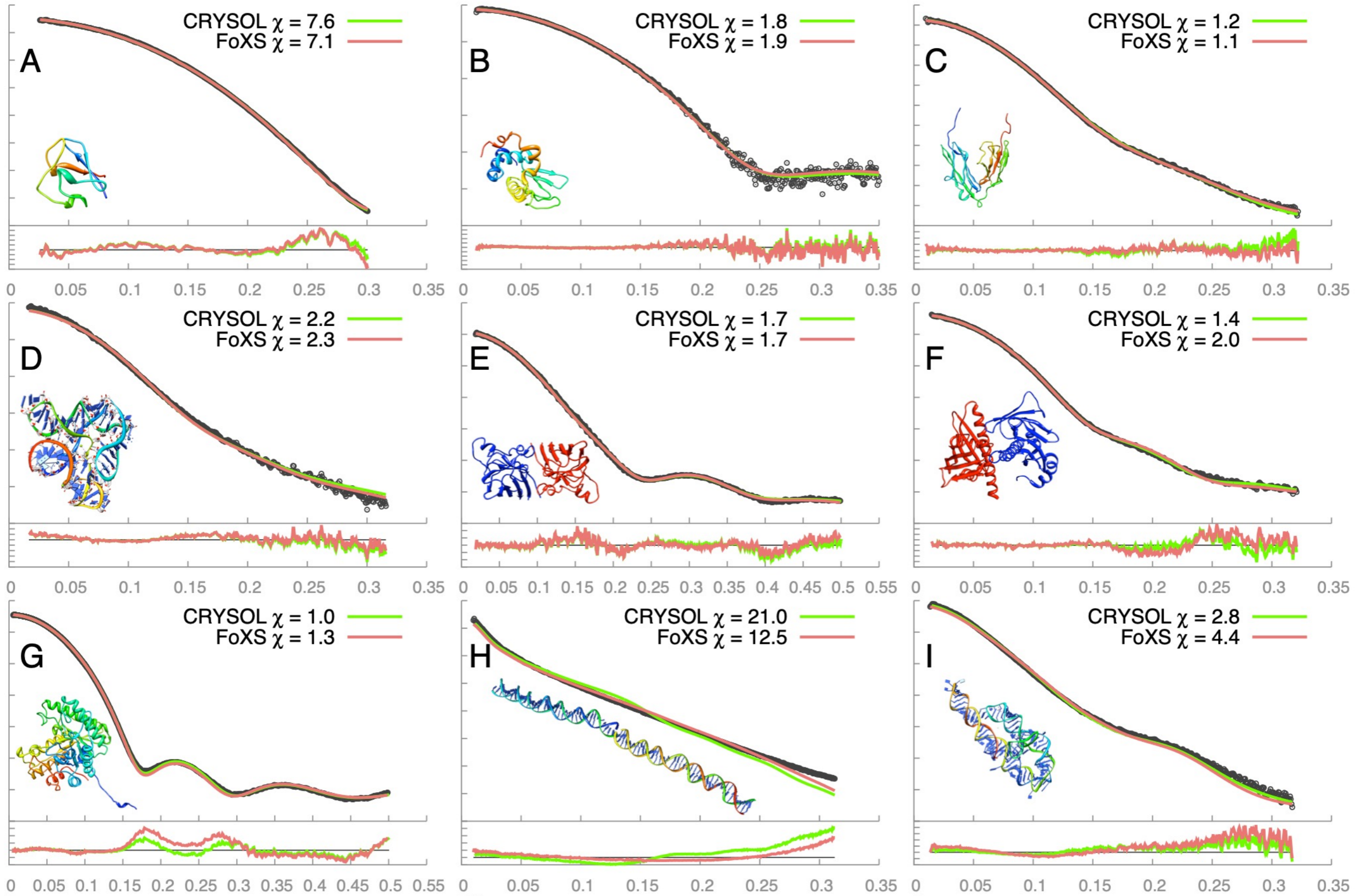
Add water form factor to solvent accessible atoms (s_i measures solvent accessibility [0-1])

5% variance in radius

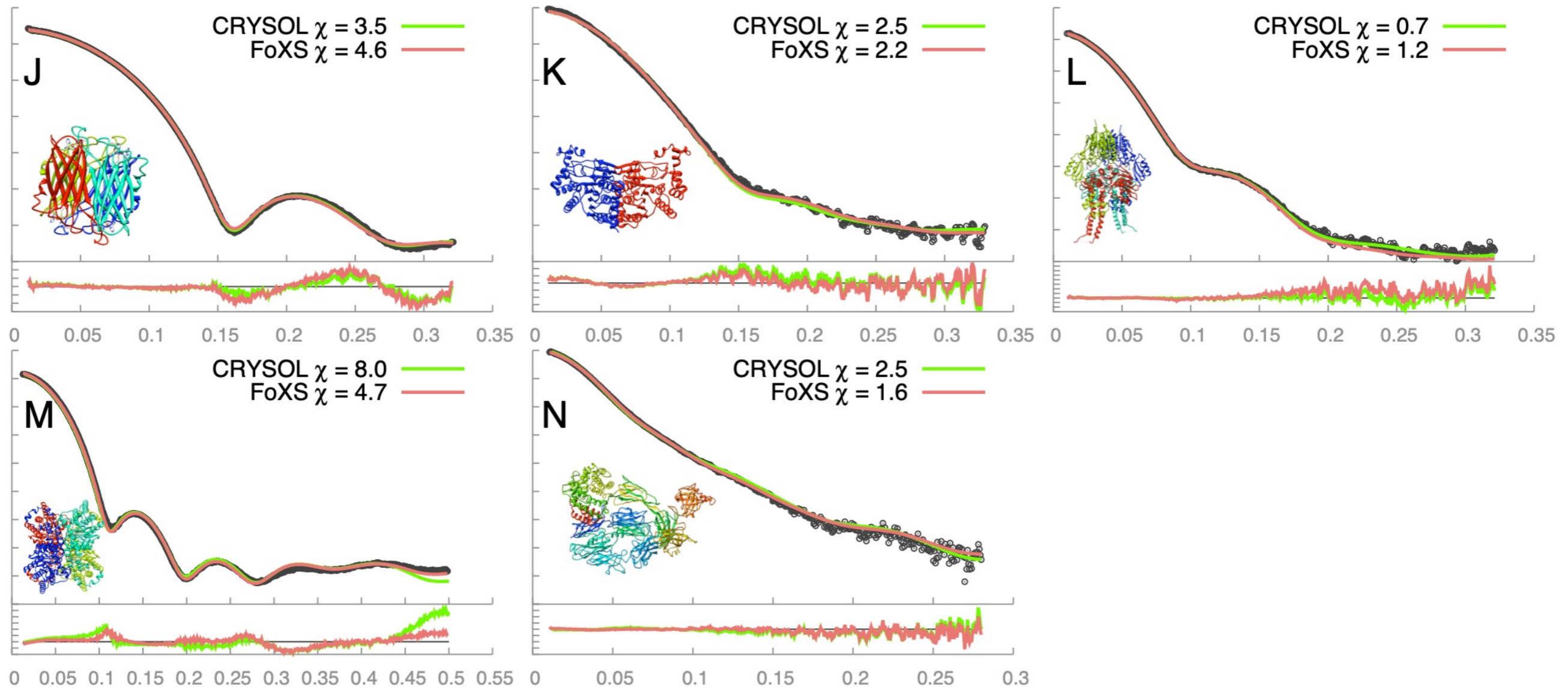
$$0.32 \text{ e}/\text{\AA}^3 \leq \rho \leq 0.38 \text{ e}/\text{\AA}^3$$

enumeration of 2 fitting parameters: c_1, c_2

X-ray structure vs. SAXS - good fits → publish

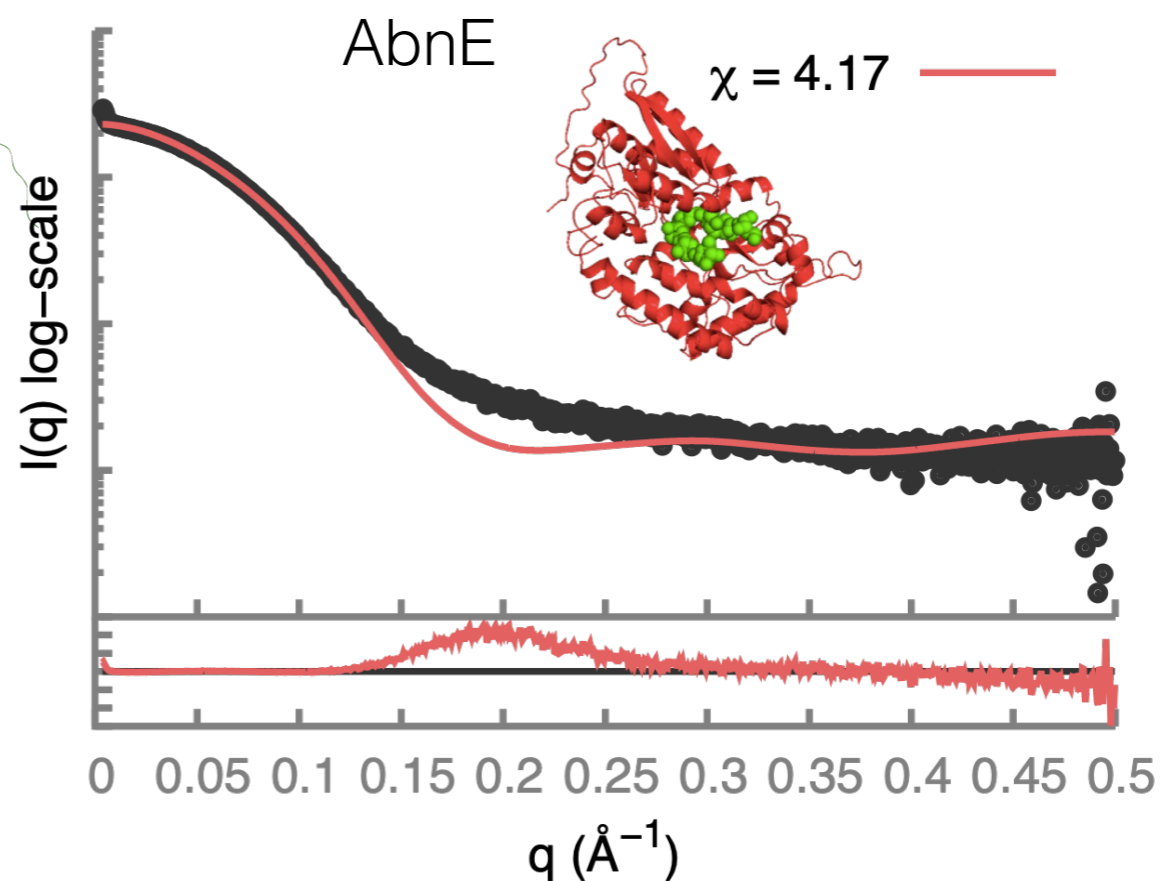
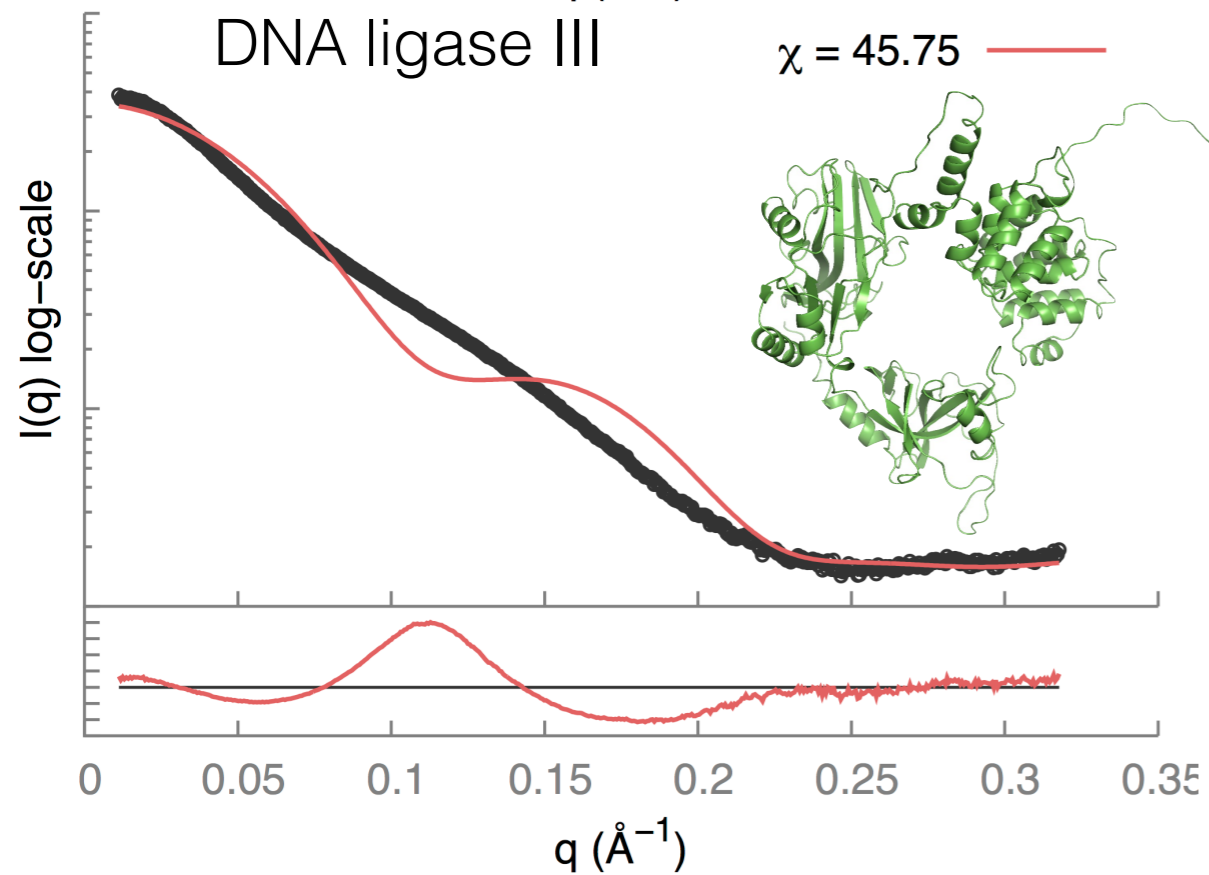
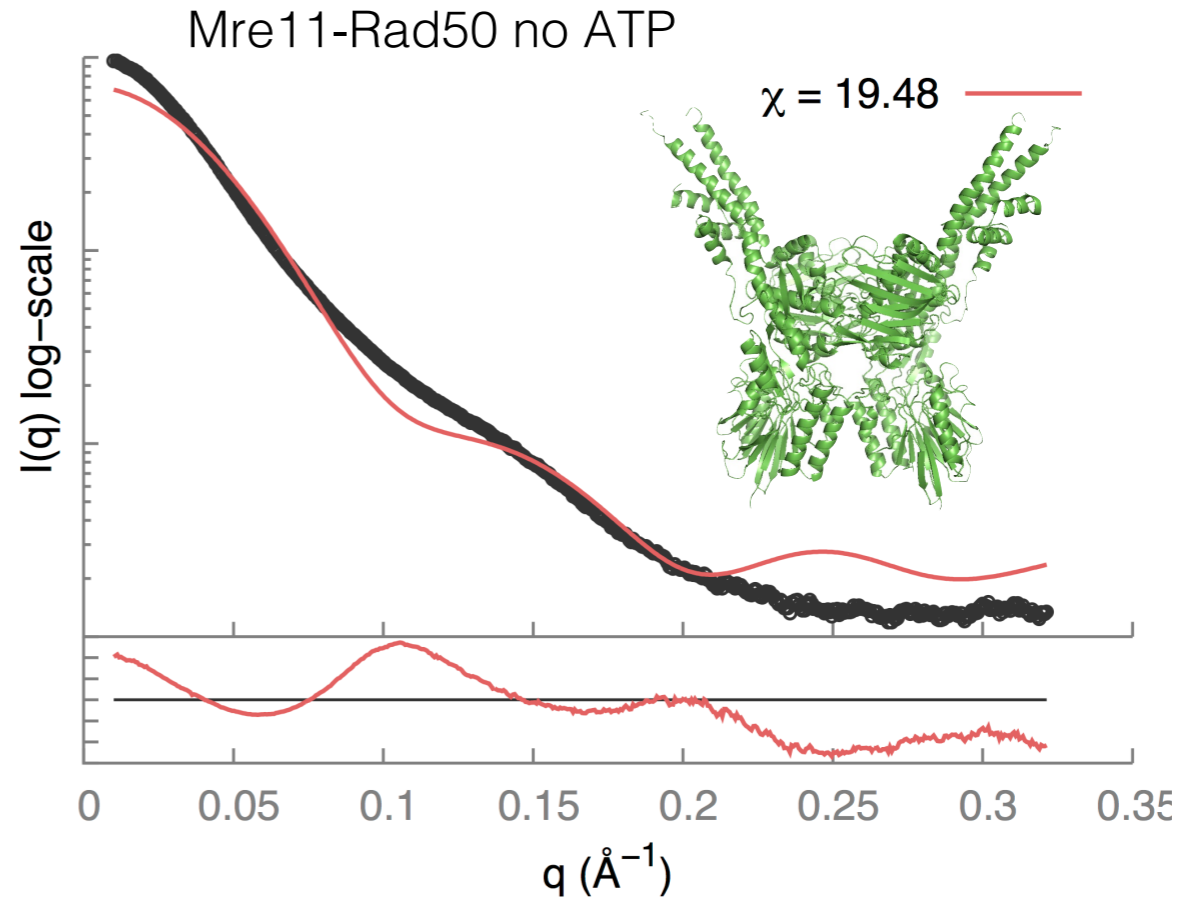
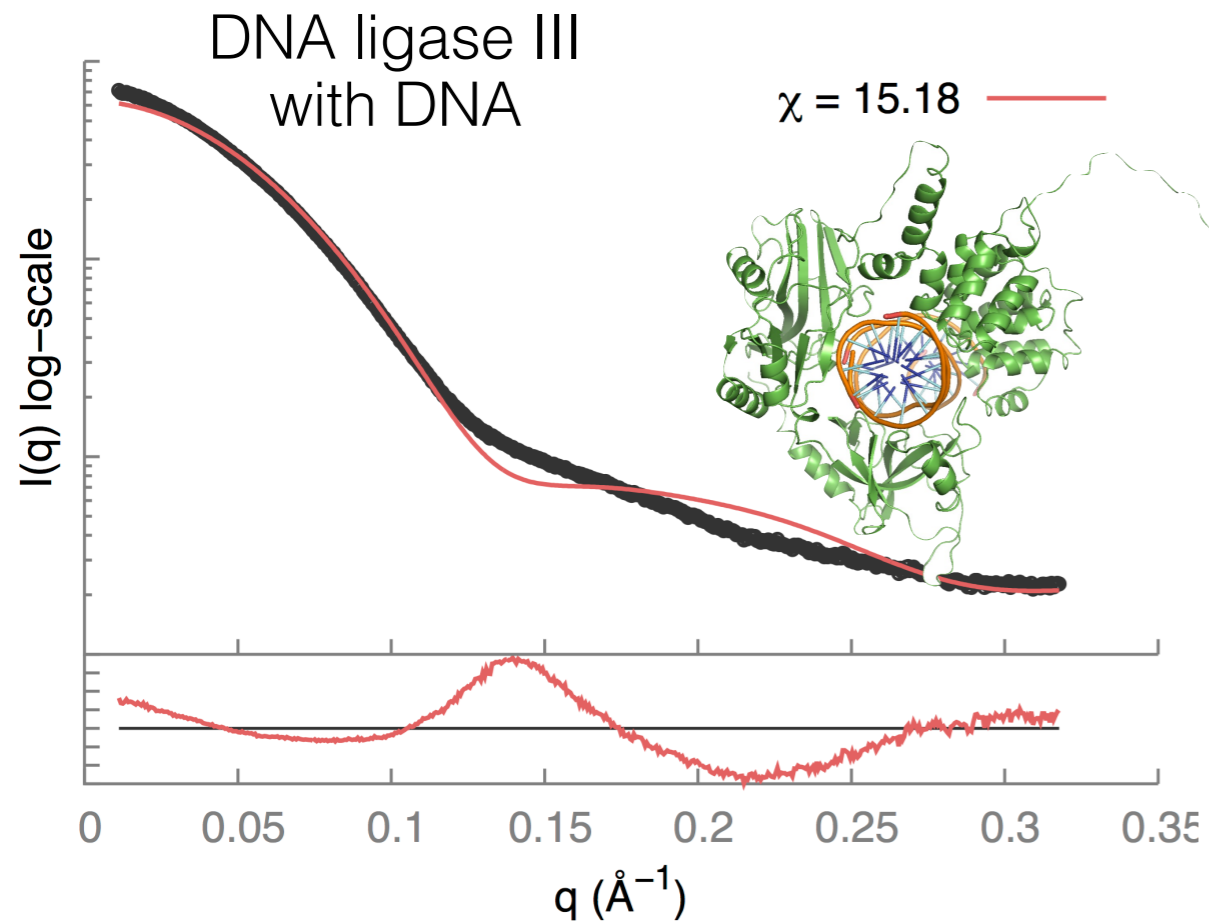


X-ray structure vs. SAXS - good fits → publish



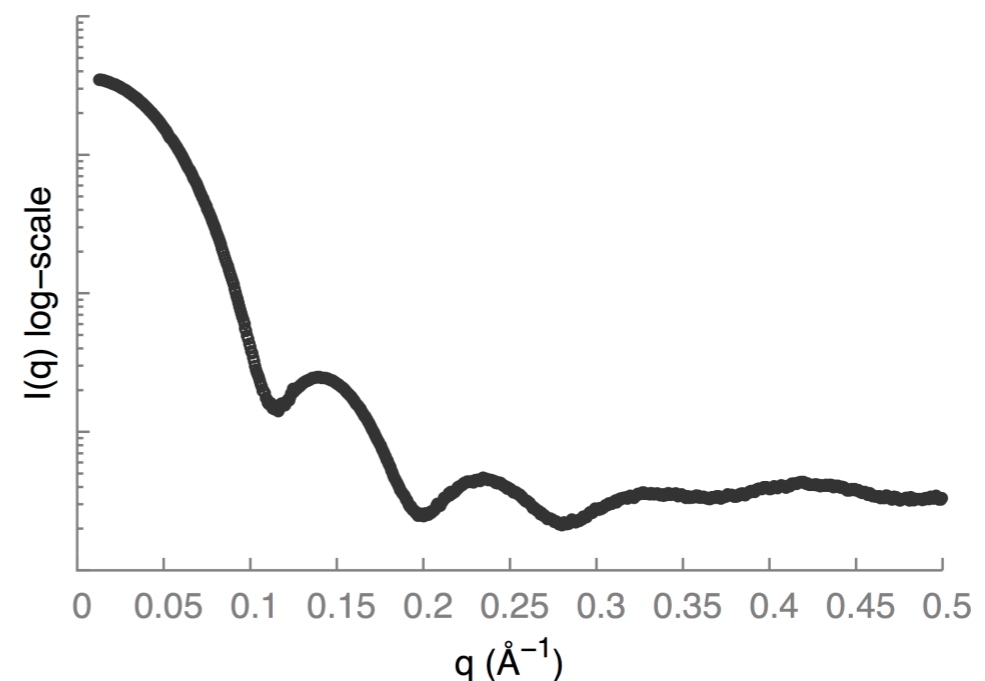
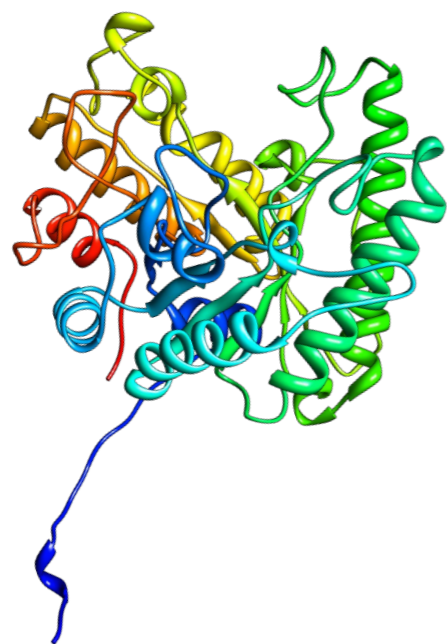
14 experimental datasets with x-ray structures

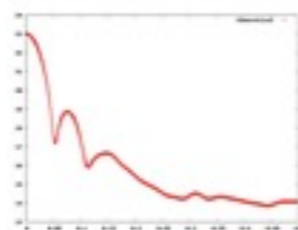
X-ray structure vs. SAXS - they don't fit!



X-ray structure vs. SAXS

- Data quality
- Missing residues/sugars
- Compositional heterogeneity
- Conformational heterogeneity
- both





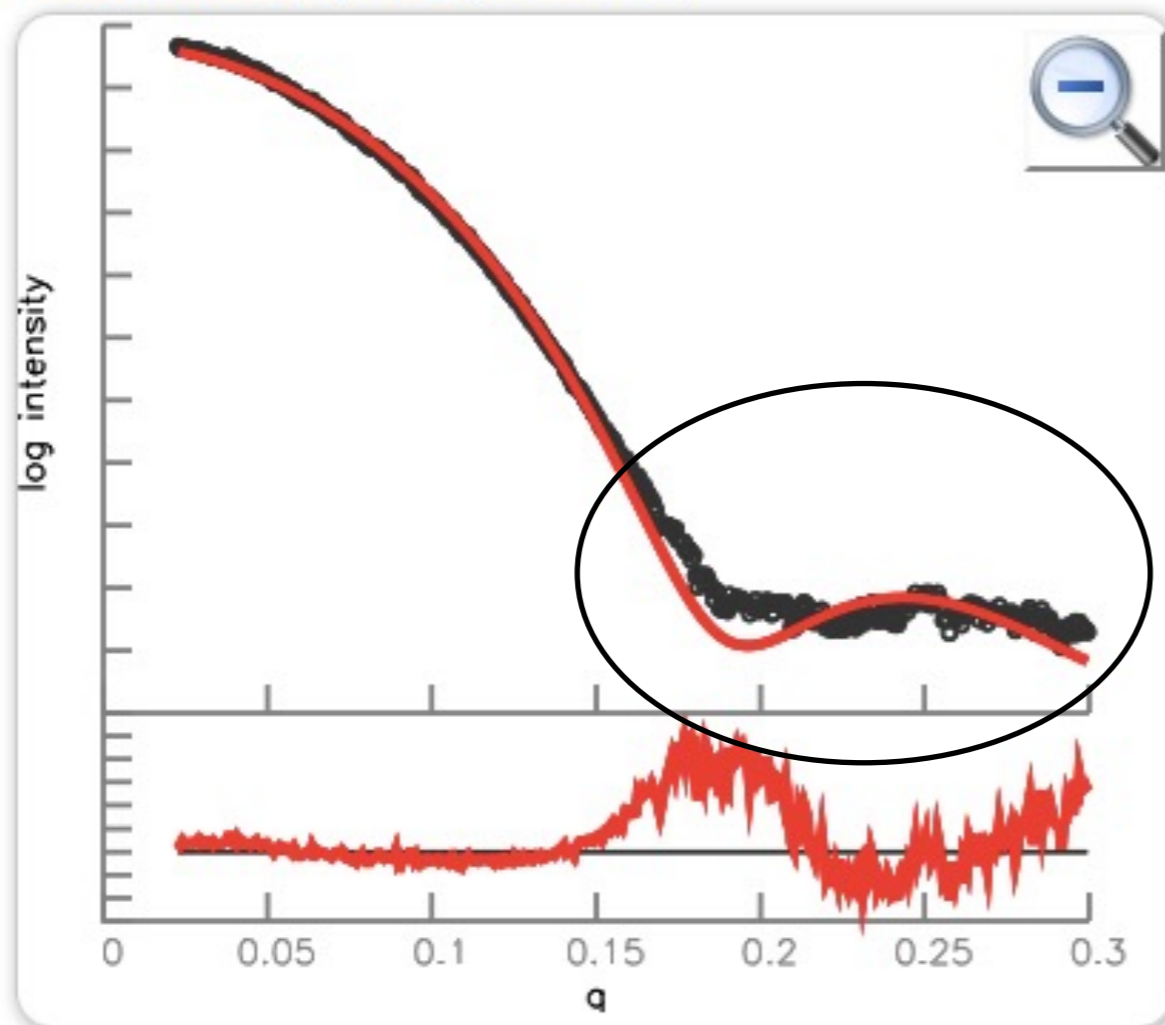
PDB files

[3KFO.pdb](#)

Profile file

[23922_merge.dat](#)

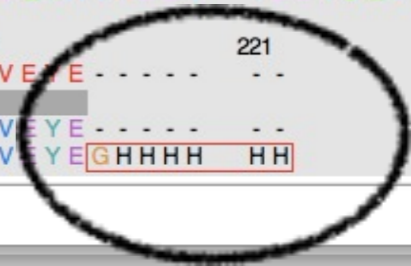
Can't see interactive display? Use [old interface](#)



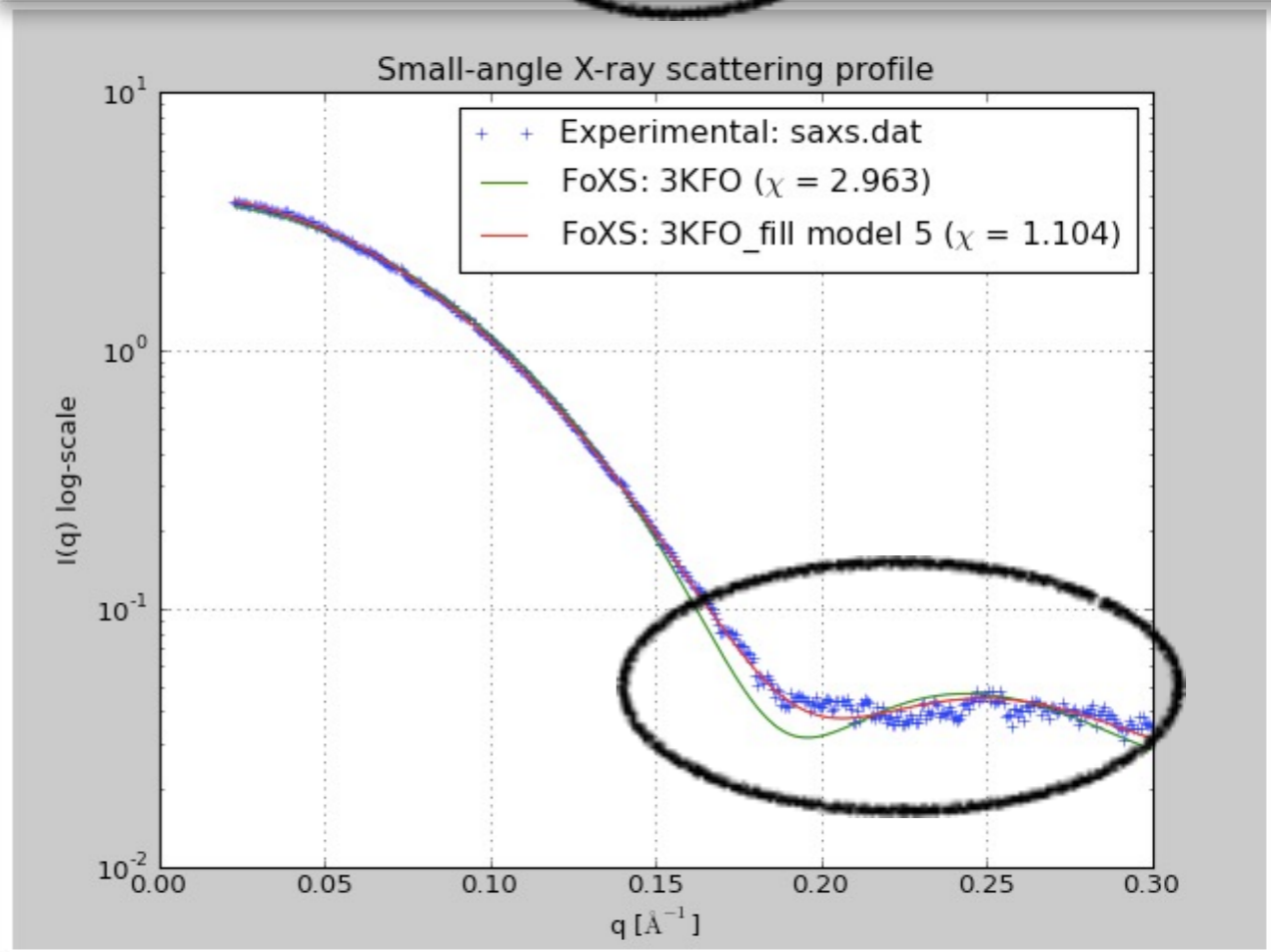
PDB file	show/hide	χ	c_1	c_2	R_g	# atoms	fit file
3KFO	<input checked="" type="checkbox"/>	2.96	1.02	3.76	16.59	1669	3KFO_23922_merge.dat

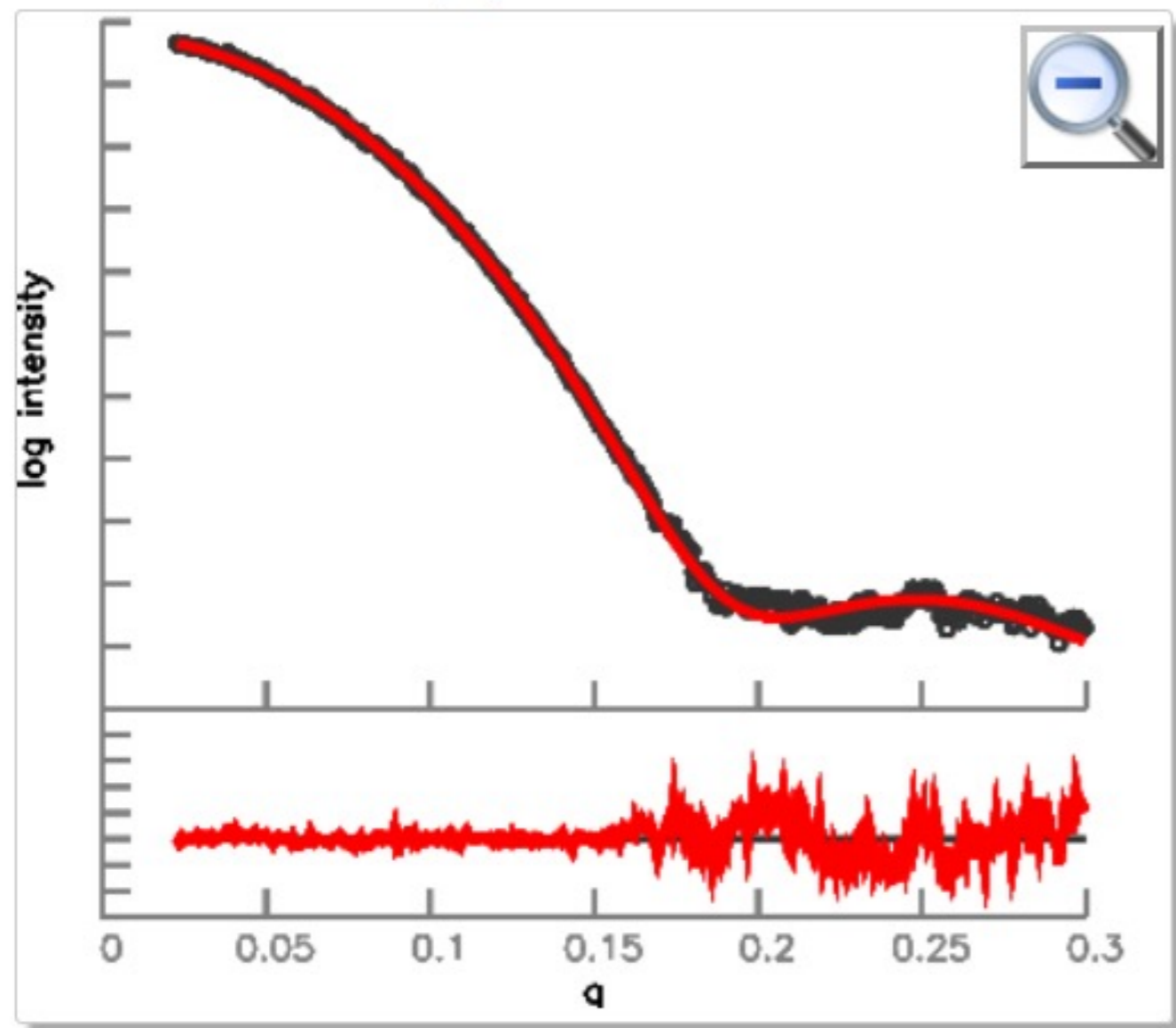


3KFO-fill:	TTVVNLSDLV	ELYSMLDDEE	SLFIPLRLLS	VDGNLLNFEV	K
101	RIVLLNASNE	GDKLLQHIVK	RVFDEELPKN	NDFPLPSVDL	141
Consensus	██████████	██████████	██████████	██████████	
Conservation	██████████	██████████	██████████	██████████	
3KFO	RIVLLNASNE	GDKLLQHIVK	RVFDEELPKN	NDFPLPSVDL	L
3KFO-fill:	RIVLLNASNE	GDKLLQHIVK	RVFDEELPKN	NDFPLPSVDL	L
151	YISETYGRFP	IDQNAIREEI	YEEISQVETL	NSDNSLEIKL	191
Consensus	██████████	██████████	██████████	██████████	
Conservation	██████████	██████████	██████████	██████████	
3KFO	YISETYGRFP	IDQNAIREEI	YEEISQVETL	NSDNSLEIKL	H
3KFO-fill:	YISETYGRFP	IDQNAIREEI	YEEISQVETL	NSDNSLEIKL	H
201	KNYTINYETN	TVEE	---	---	221
Consensus	██████████	██████████	██████████	██████████	
Conservation	██████████	██████████	██████████	██████████	
3KFO	KNYTINYETN	TVEYE	---	---	
3KFO-fill:	KNYTINYETN	TVEYE	GHHHH	HH	



Quit



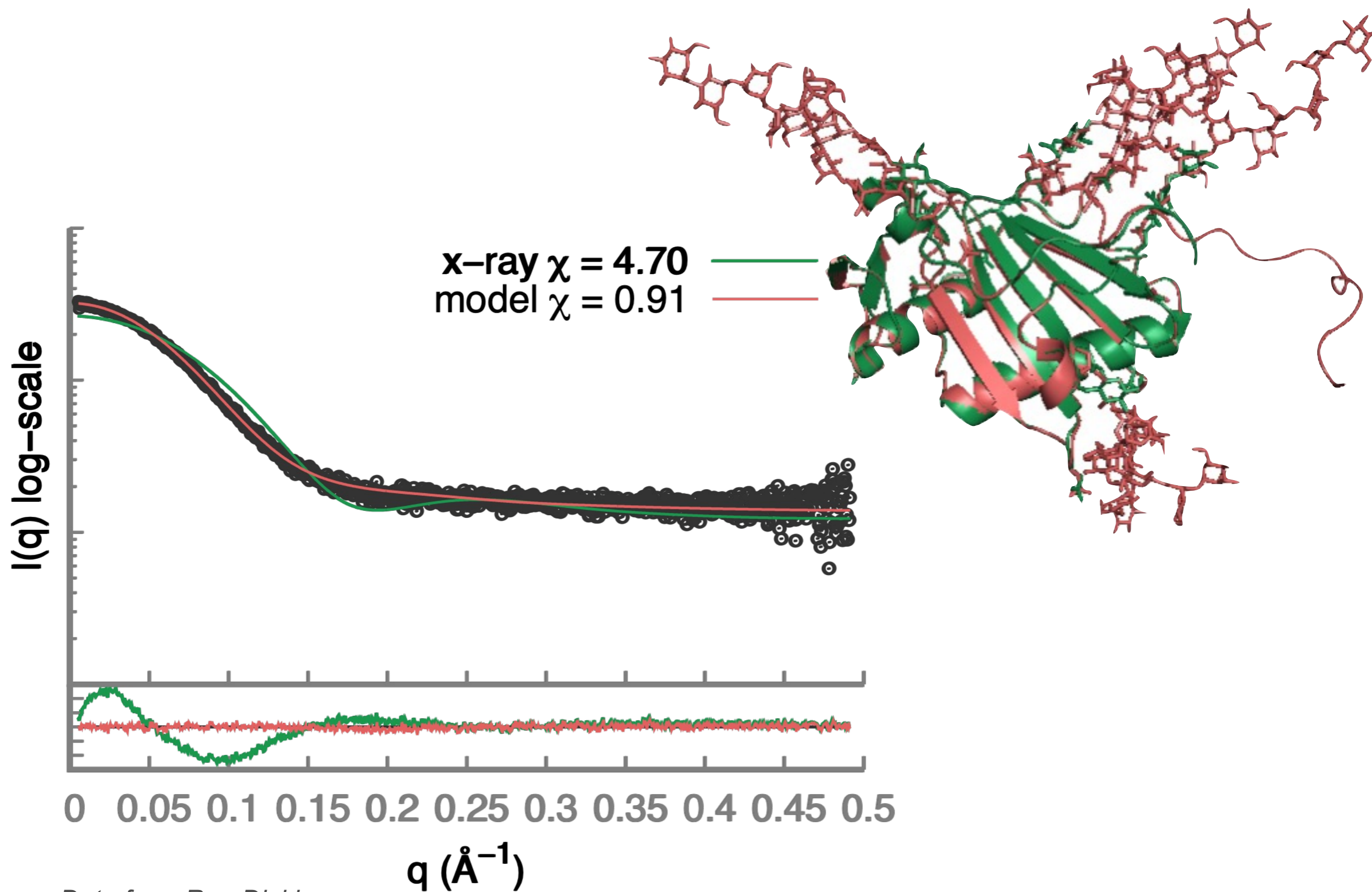


PDB file	<input type="checkbox"/> show all/hide all	χ	c_1	c_2	R_g	# atoms	fit file	png file
3KFOfill.B99990003	<input checked="" type="checkbox"/>	1.09	1.03	1.22	17.68	1817	3KFOfill.B99990003_saxs.dat	3KFOfill.B99990003_saxs.png
3KFOfill.B99990005	<input type="checkbox"/>	1.22	1.03	0.81	17.93	1817	3KFOfill.B99990005_saxs.dat	3KFOfill.B99990005_saxs.png
3KFOfill.B99990004	<input type="checkbox"/>	1.24	1.03	0.91	17.62	1817	3KFOfill.B99990004_saxs.dat	3KFOfill.B99990004_saxs.png
3KFOfill.B99990002	<input type="checkbox"/>	1.32	1.03	0.66	17.56	1817	3KFOfill.B99990002_saxs.dat	3KFOfill.B99990002_saxs.png
3KFOfill.B99990001	<input type="checkbox"/>	2.27	1.02	0.68	17.29	1817	3KFOfill.B99990001_saxs.dat	3KFOfill.B99990001_saxs.png
3KFO1	<input type="checkbox"/>	2.96	1.03	3.40	16.59	1669	3KFO1_saxs.dat	3KFO1_saxs.png

overfitting

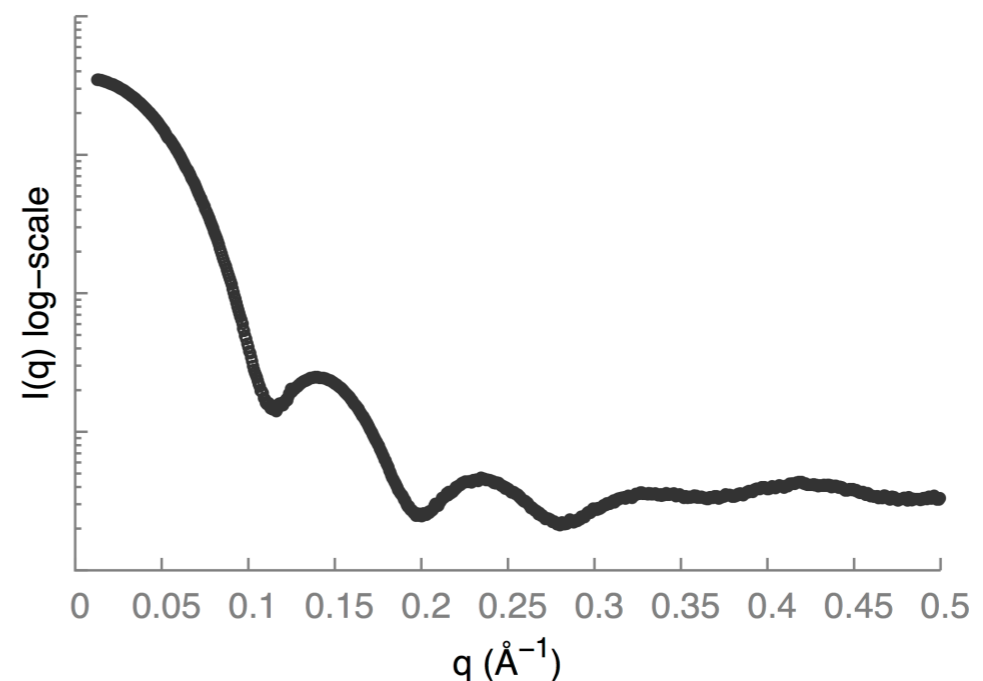
$0.32 \text{ e}/\text{\AA}^3 \leq \rho \leq 0.38 \text{ e}/\text{\AA}^3 \rightarrow c_2 < 4.0$

LASSA GP1



X-ray structure vs. SAXS

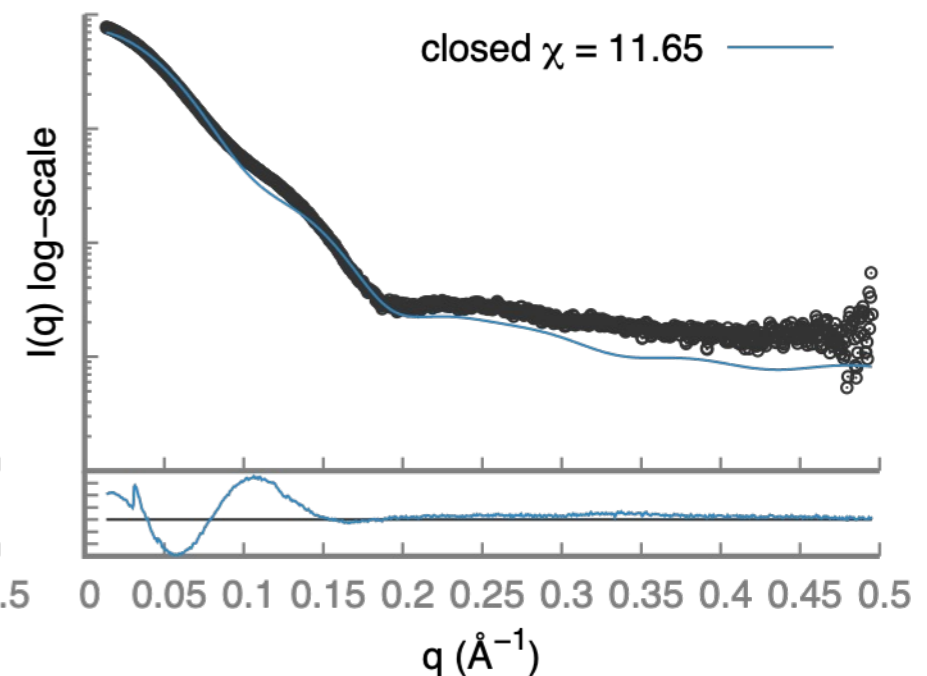
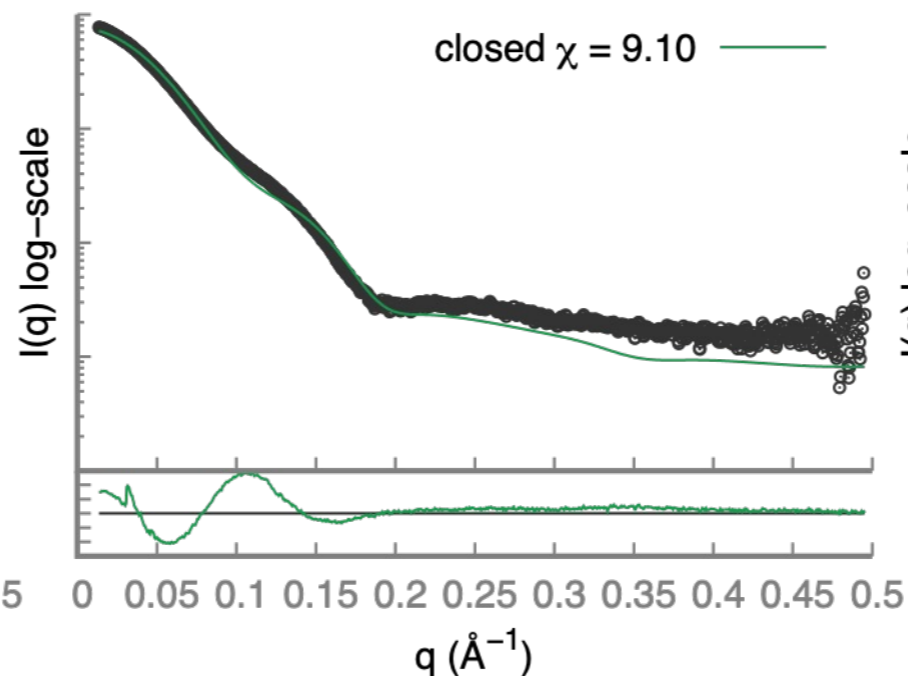
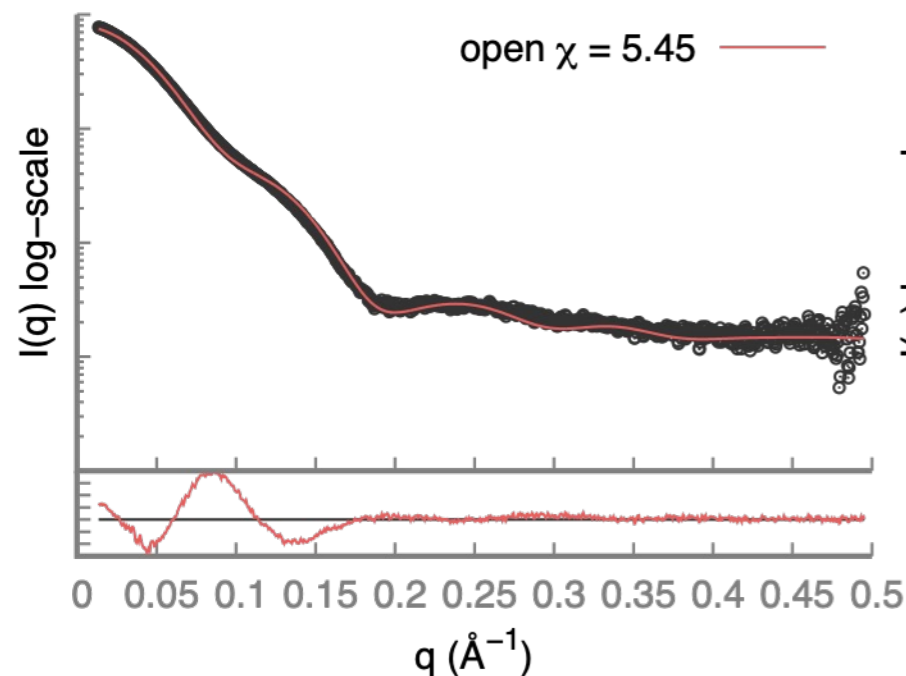
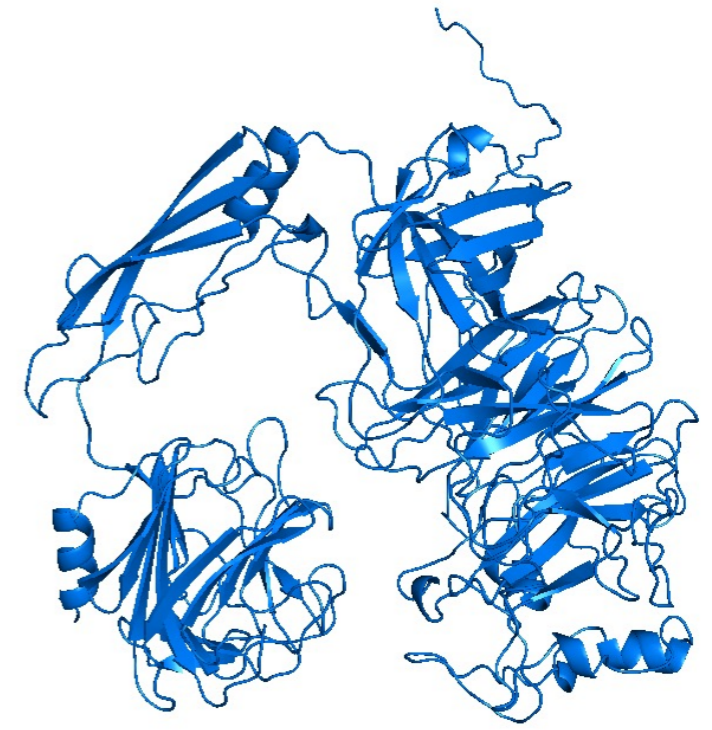
- Data quality
- Missing residues/sugars
- Compositional heterogeneity
- Conformational heterogeneity
- both



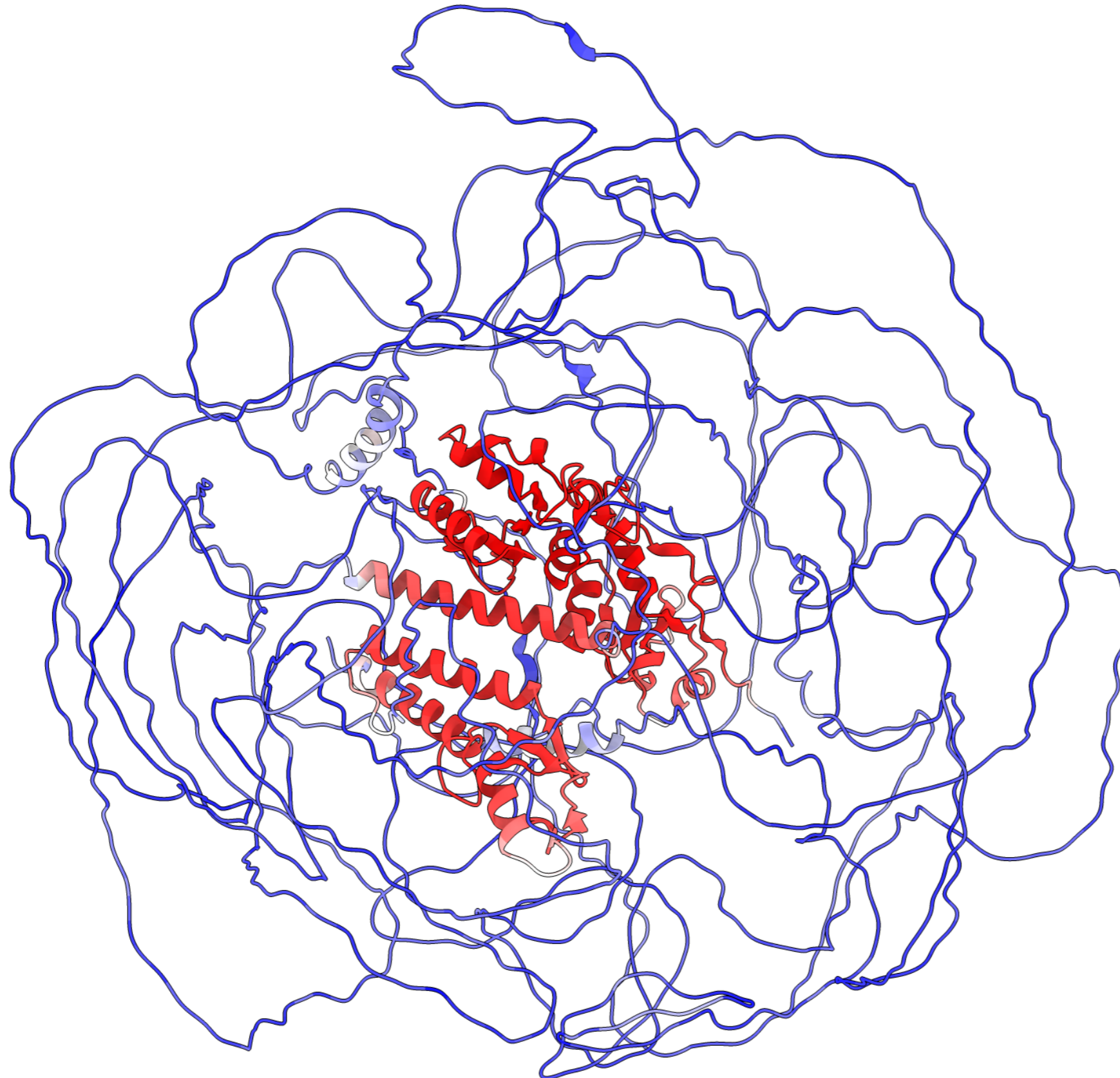
AbnA structures vs. SAXS



- 3 X-ray structures in different conformations do not fit the data



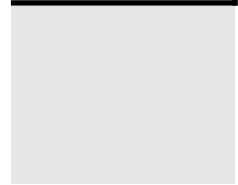
Dynamics Comes in Flavors and it is Common



ed”



nents



(all 2007)

PD

Sw

Dynamics and SAXS

- SAXS data can be easily collected for proteins that include disordered regions
- Data interpretation is challenging

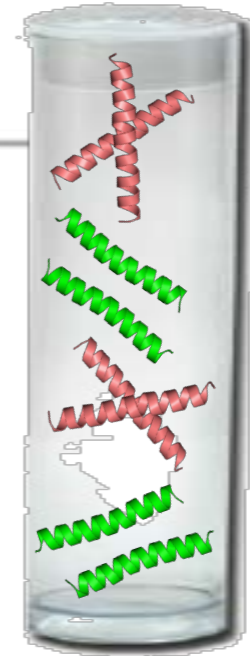


```
while(noSuccess)
{
    tryAgain();
    if(Dead)
        break;
}
```


Heterogeneous Sample Requires Multi-State Model

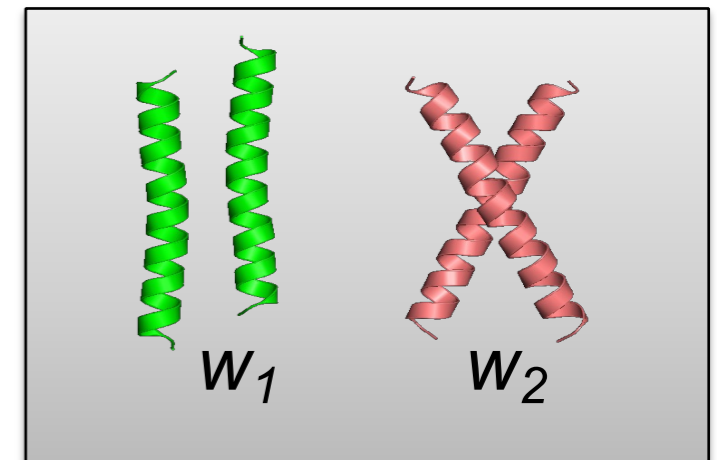
Heterogeneous sample

compositional or **conformational** heterogeneity in the sample used to generate the data



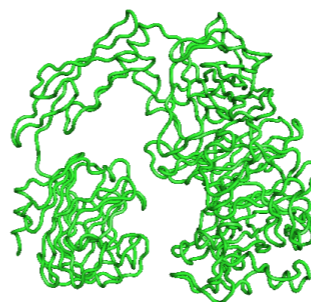
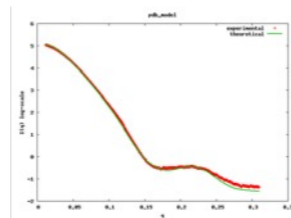
Multi-state model

a model that specifies two or more co-existing **structural states** and values for **any other parameter**



MultiFOXS

salilab.org/multifoxs

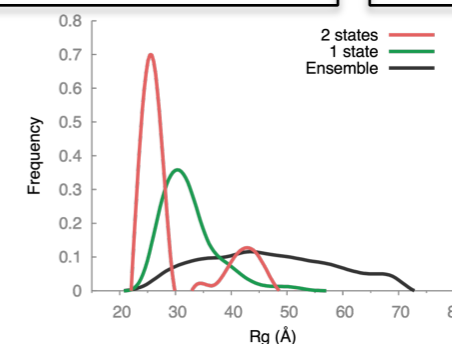
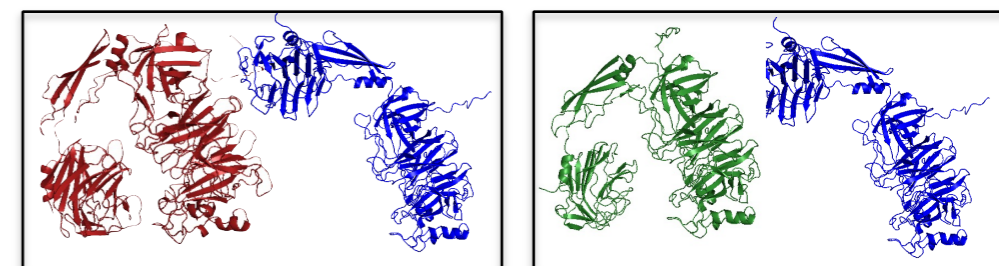
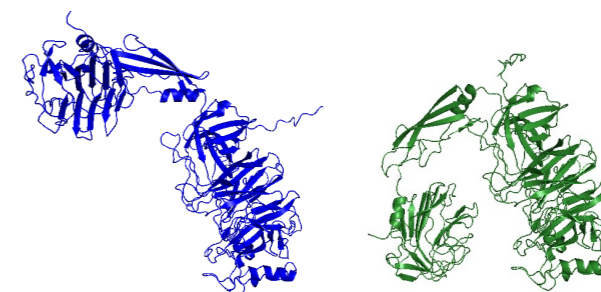
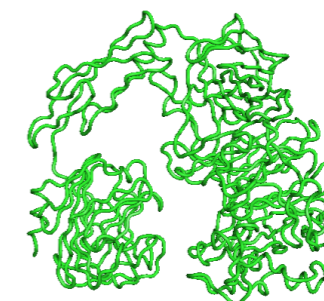


flexible residues:

35 A

36 A

...



Sampling

Rapidly exploring Random Trees (RRTs)



Scoring

Debye formula **foXS**
salilab.org/foxs



Enumeration

Enumeration of multi-state models that fit the data within noise



Analysis

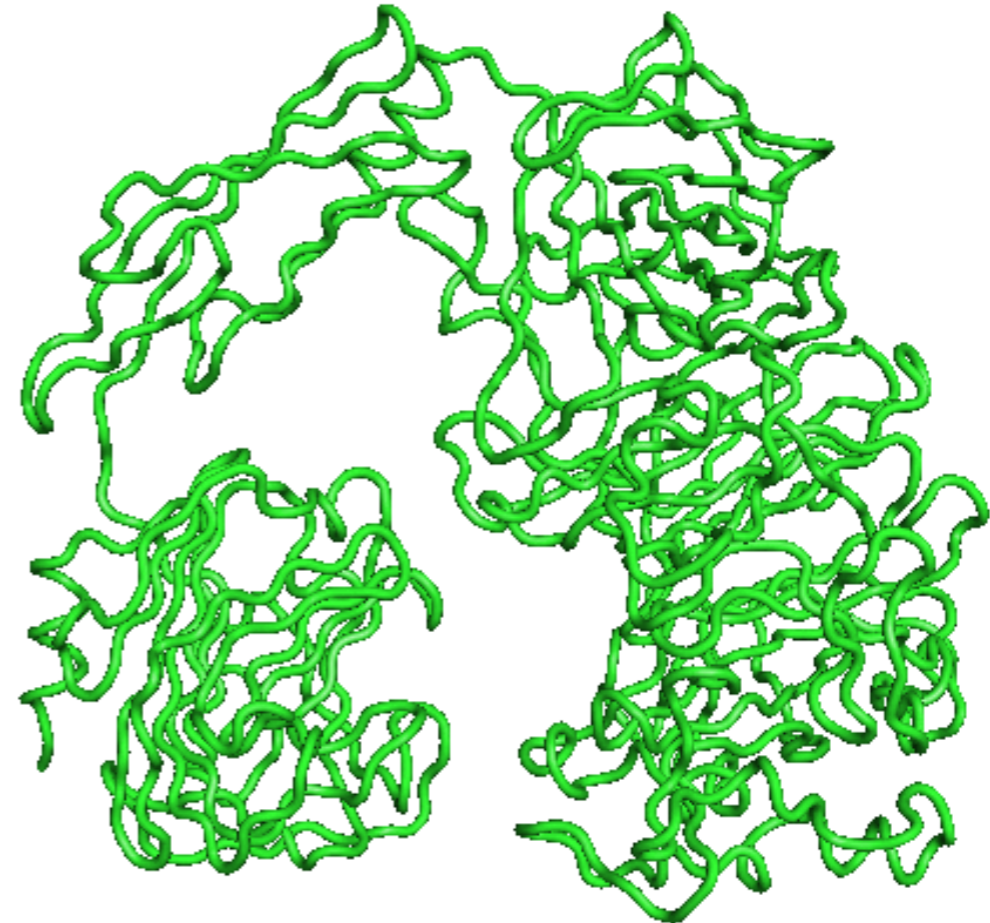
Quality of fit to data, Rg variance among top scoring models

Conformational sampling

Proteins and robots have similar degrees of freedom



Robotic arm

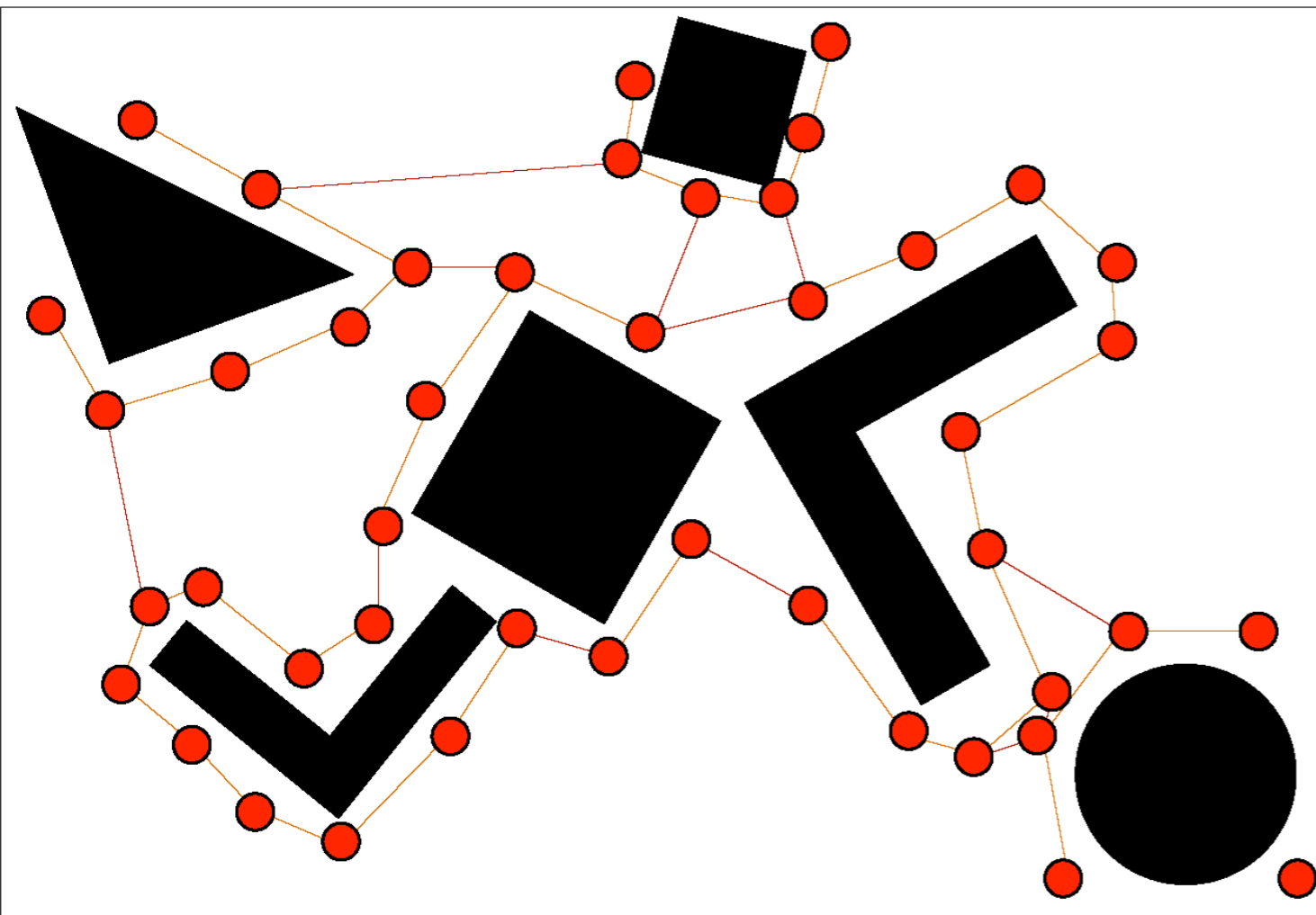


AbnA protein

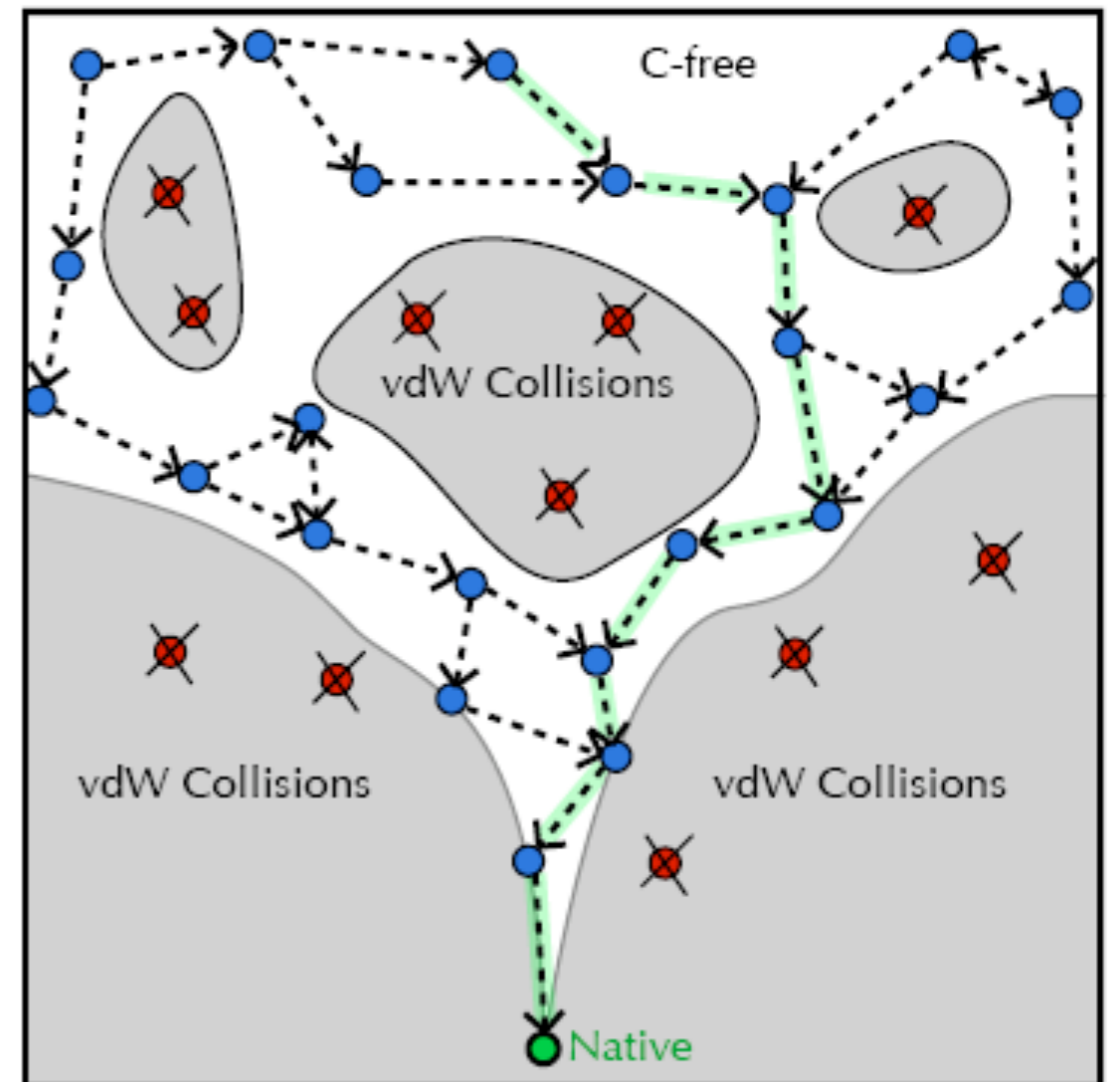
We rely on methods for Motion Planning developed in Robotics (*La Valle, Latomb, Kavraki, Cortes*)

Mapping collision free space with Rapidly exploring Random Tree (RRT)

Collision free space for robot



Collision free space for protein chain

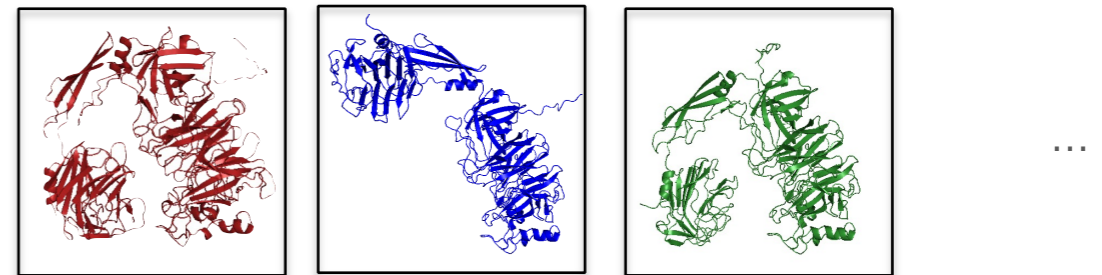


Enumeration of multi-state models

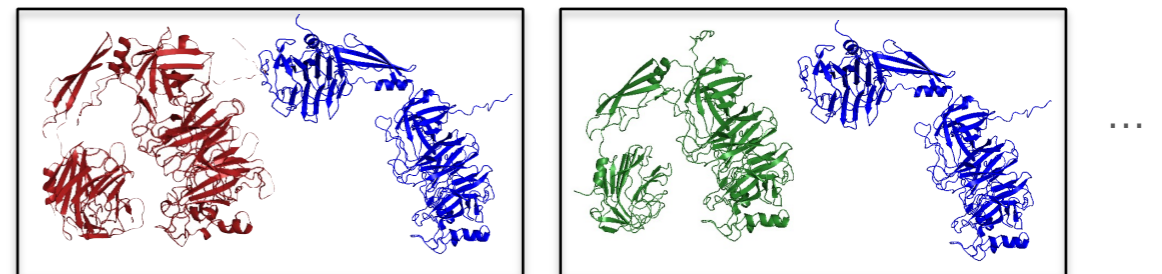
branch & bound deterministic algorithm

Multi-state models of size $i+1$ are generated by extending **best**
K (=10000) multi-state models of size i

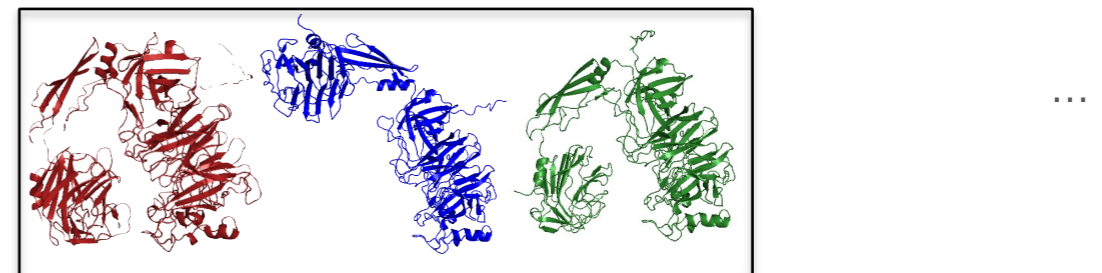
best K multi-state models of size 1:



best K multi-state models of size 2:

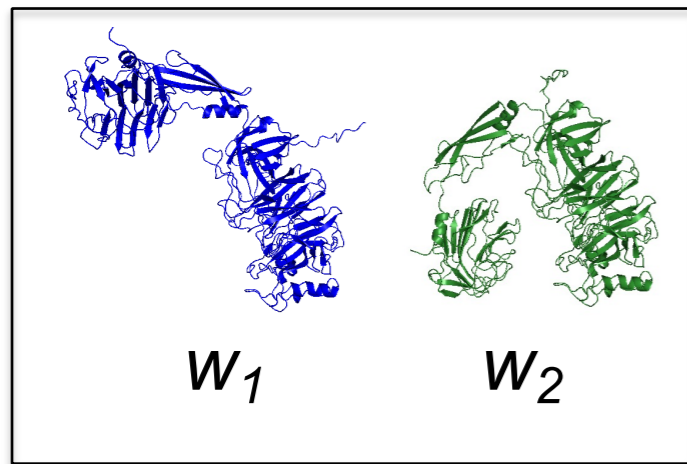


best K multi-state models of size 3:



...

Scoring of Multi-State Models

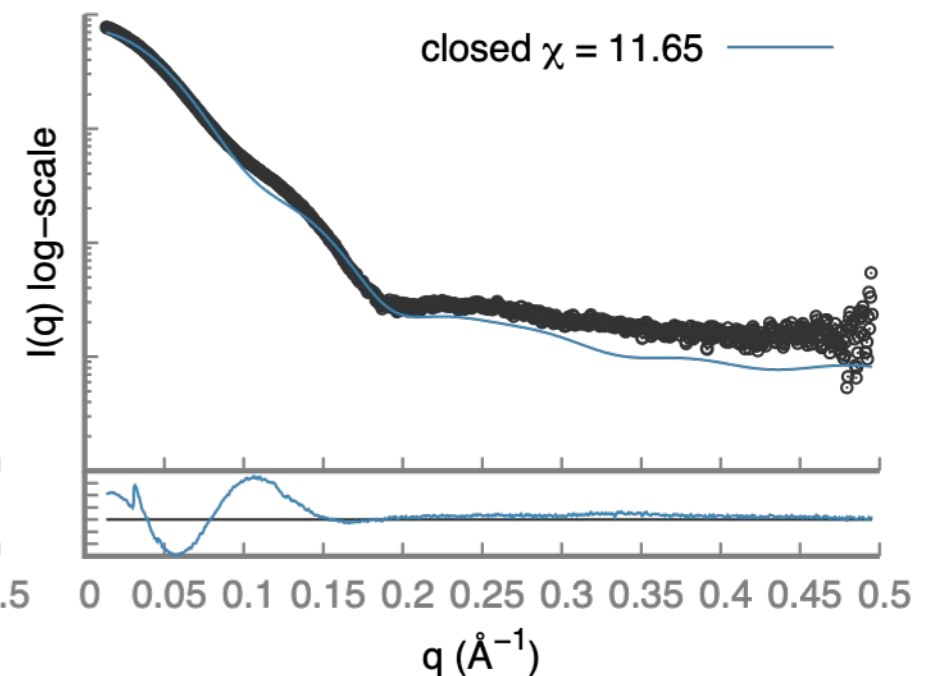
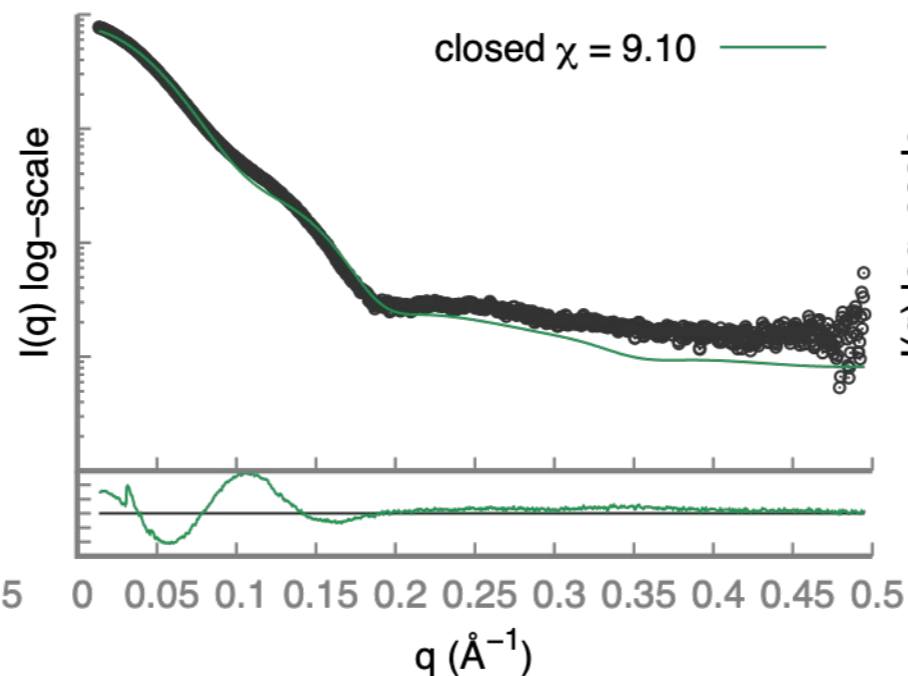
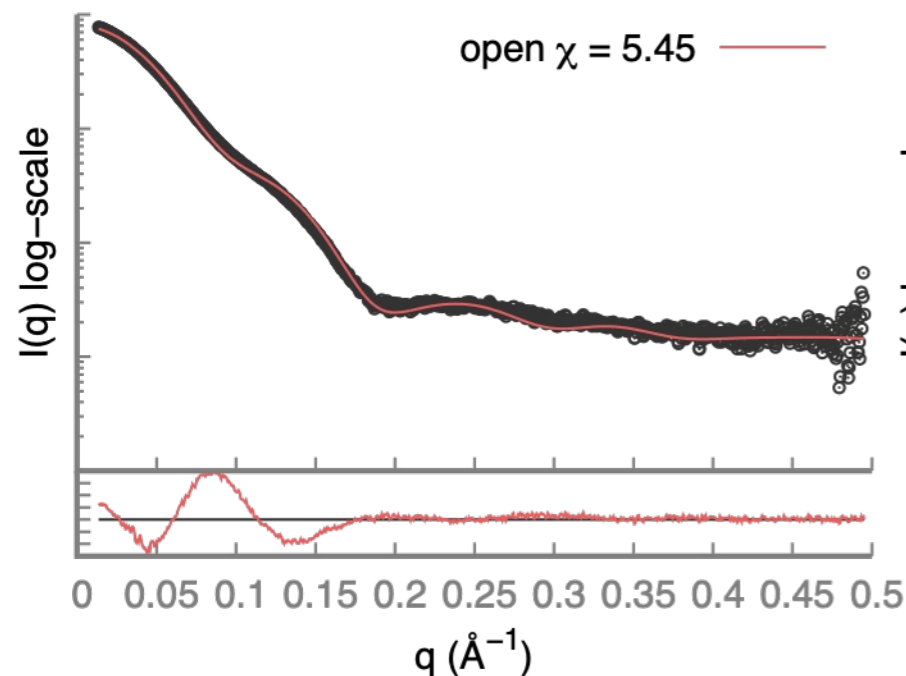
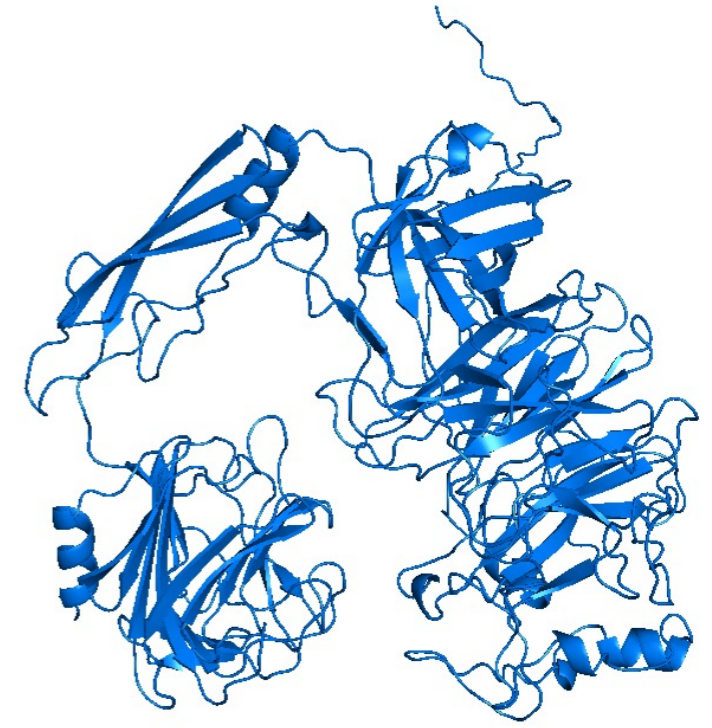


$$\chi = \sqrt{\frac{1}{S} \sum_{i=1}^S \left(\frac{I_{exp}(q_i) - c \sum_n w_n I_n(q_i, c_1, c_2)}{\sigma(q_i)} \right)^2}$$

- weights optimization is needed for each set of structural states
- Non-negative least square fitting (NNLS, Lawson & Hanson 1974)
- c_1 (excluded volume), and c_2 (hydration layer) are enumerated
- a single pair of c_1 and c_2 is used for all states in a multi-state model

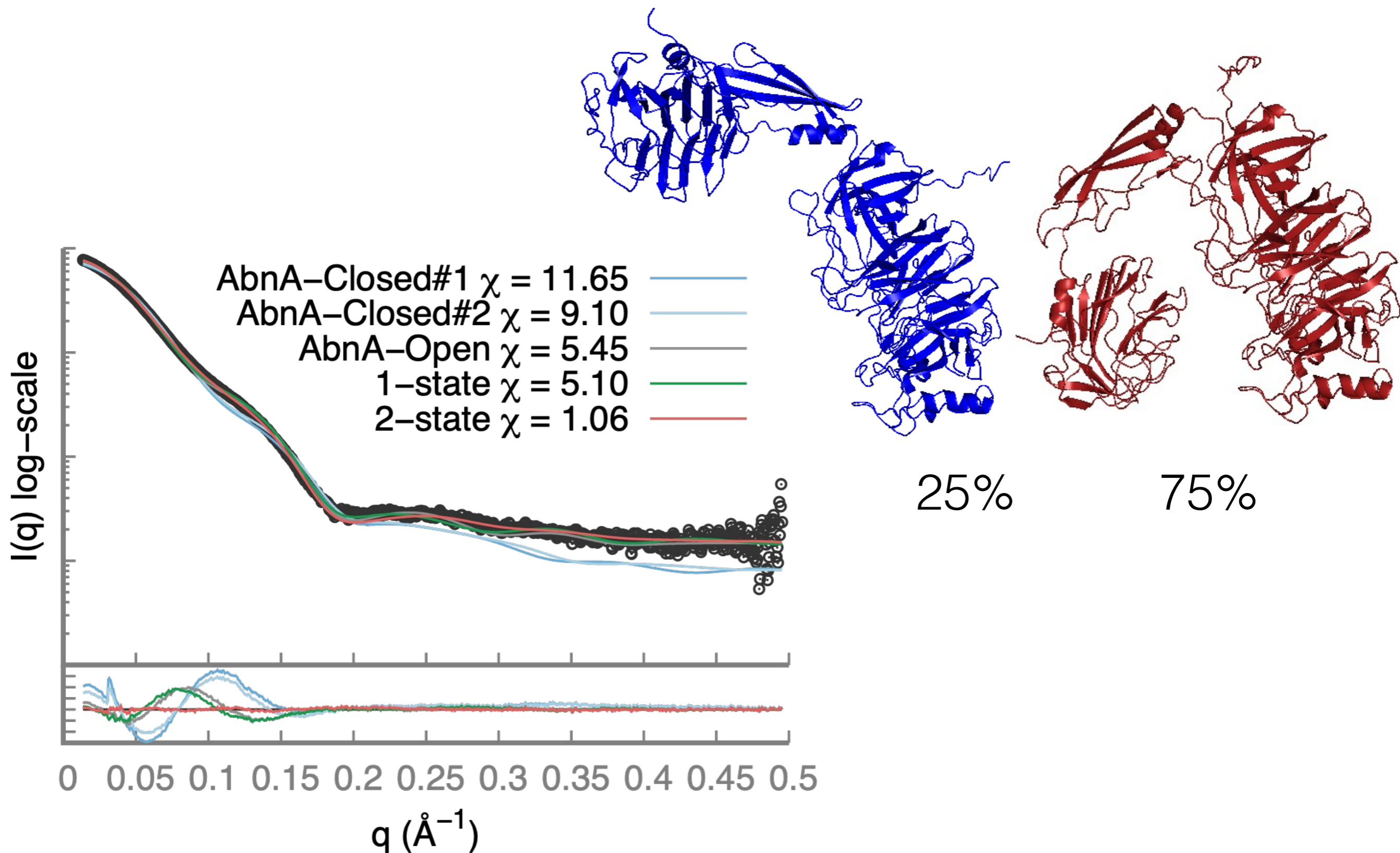
AbnA structures vs. SAXS

- 3 X-ray structures in different conformations do not fit the data



Multi-state Modeling

- Good fit to data obtained with open and closed conformations



Multi-state modeling with SAXS online

QuickTime Player File Edit View Window Help

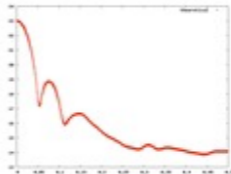
MultiFoXS Sali lab MultiFoXS service: ...

https://modbase.compbio.ucsf.edu/multifoxs/ Search

Most Visited Getting Started Google Google Scholar Google Maps Home - PubMed - ... RCSB Protein Dat... MultiFoXS

MultiFoXS

Multi-state modeling with SAXS profiles



[About MultiFoXS](#) • [Web Server](#) • [Help](#) • [FAQ](#) • [Download](#) • [FoXS](#) • [Sali Lab](#) • [IMP](#) • [Links](#)

Type PDB code for protein or upload file in PDB format

Input protein (PDB:chainId e.g. 2kai:AB) or upload file: input.pdb

Flexible residues hinges.dat

SAXS profile iq.dat

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

Advanced Parameters

Job name

Connect rigid bodies No file selected.

Number of conformations Use 100 to test your setup, 10,000 for final calculation

Contact: dina@salilab.org

BilboMD: high-temperature MD for linkers

SAXS Data Analysis with BILBOMD

BILBOMD Home

Check your jobs

Experiment Title: (letters, numbers, and spaces only)

Enter number of segments you have: (up to 10 PDB files. Each chainID and/or segID must be in a separate file.)

Select segment 1 (PDB name): XRCC4_1.pdb

Select segment 2 (PDB name): XRCC4_2.pdb

Select segment 3 (PDB name): LIGIV.pdb

Upload or create **const.inp** file: No file chosen

or

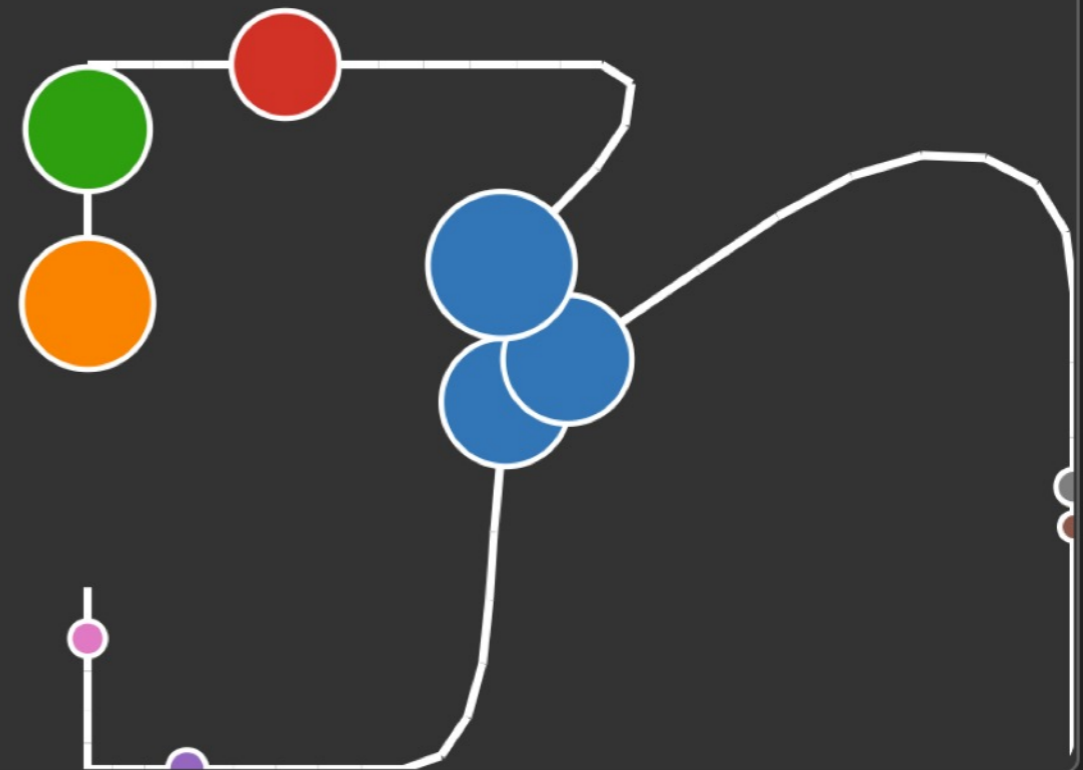
Experimental data: l4x4new.dat

Extent of Conformational Sampling:

R_g min (Å): (10-100)

R_g max (Å): (10-100)

Enter email to receive results:

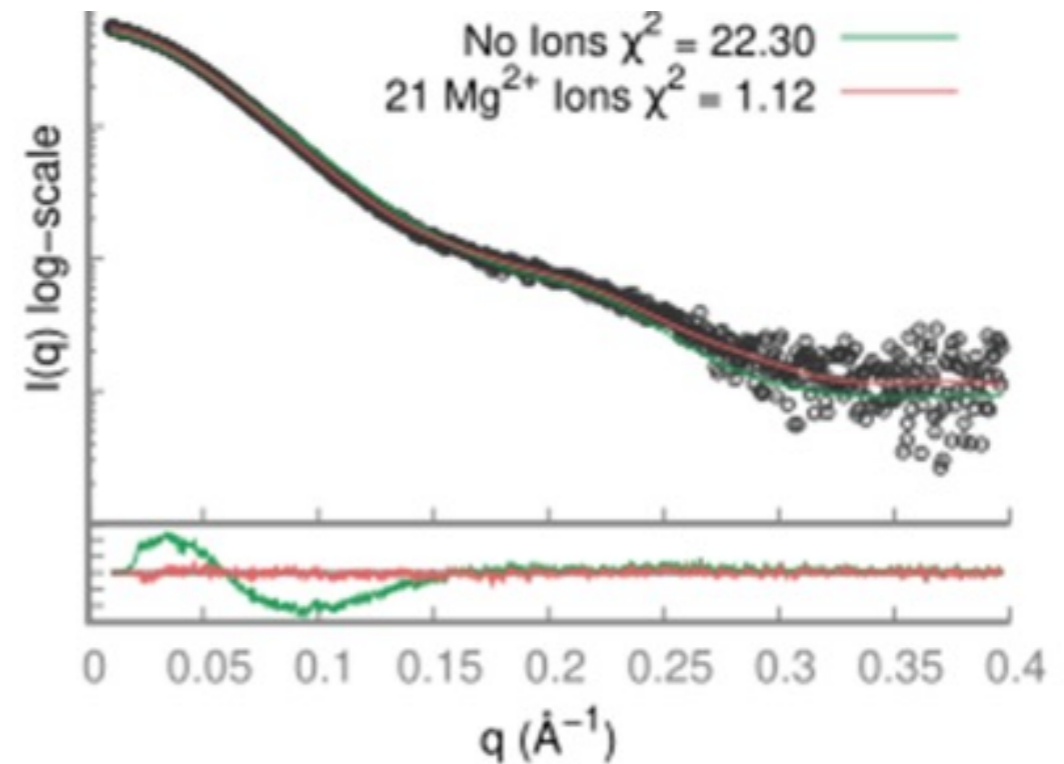
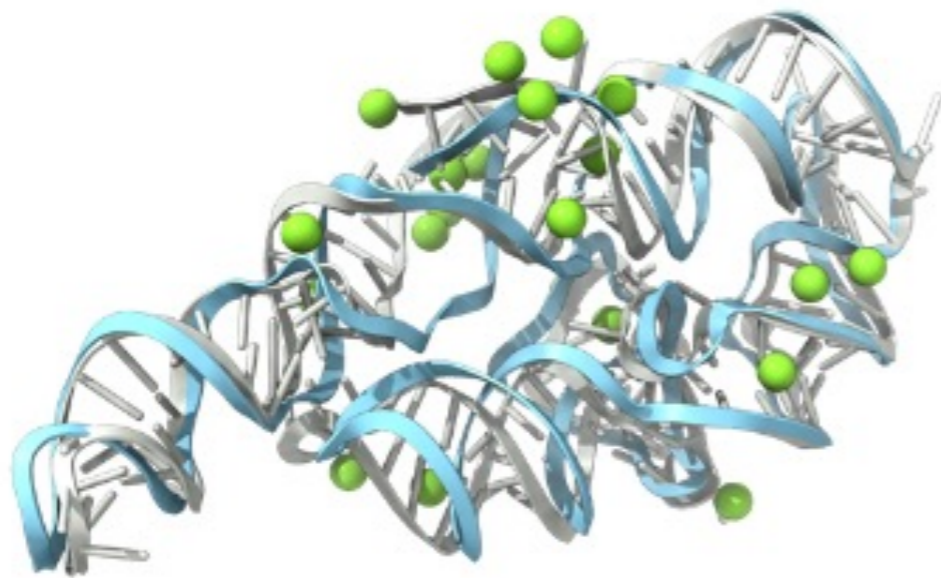


<https://bl1231.als.lbl.gov/bilbomd>

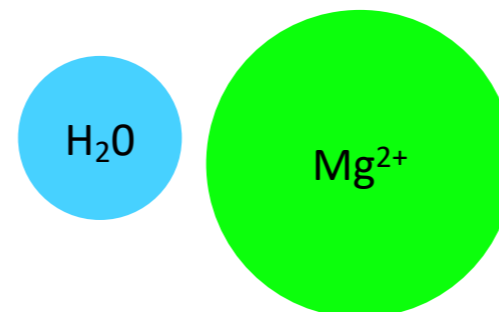
Pelikan M, Hura GL, Hammel M.2009

SAXS profile calculator for RNA structure validation

- RNA binds Mg^{2+} ions that are required for proper folding and charge neutralization



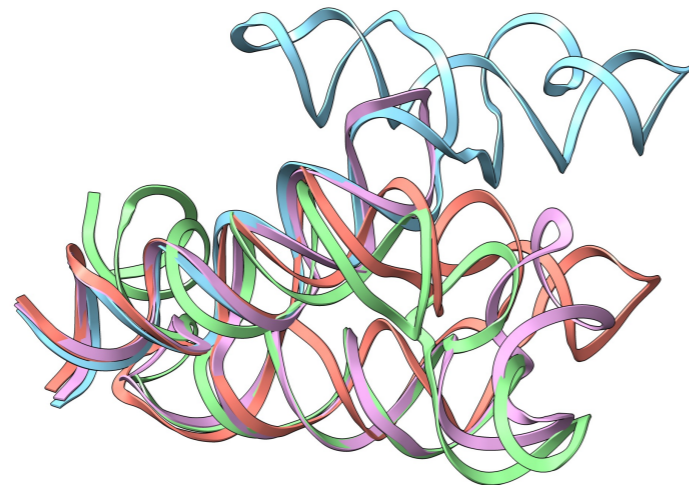
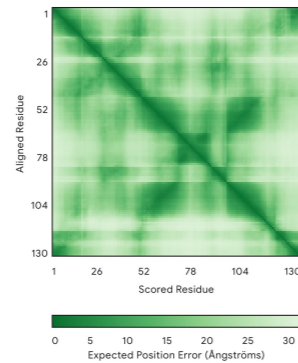
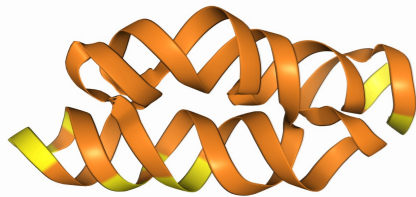
X-ray scattering Length density



SCOPER: Solution Conformation Predictor for RNA

Very high (pIDDT > 90) Confident (90 > pIDDT > 70) Low (70 > pIDDT > 50) Very low (pIDDT < 50)

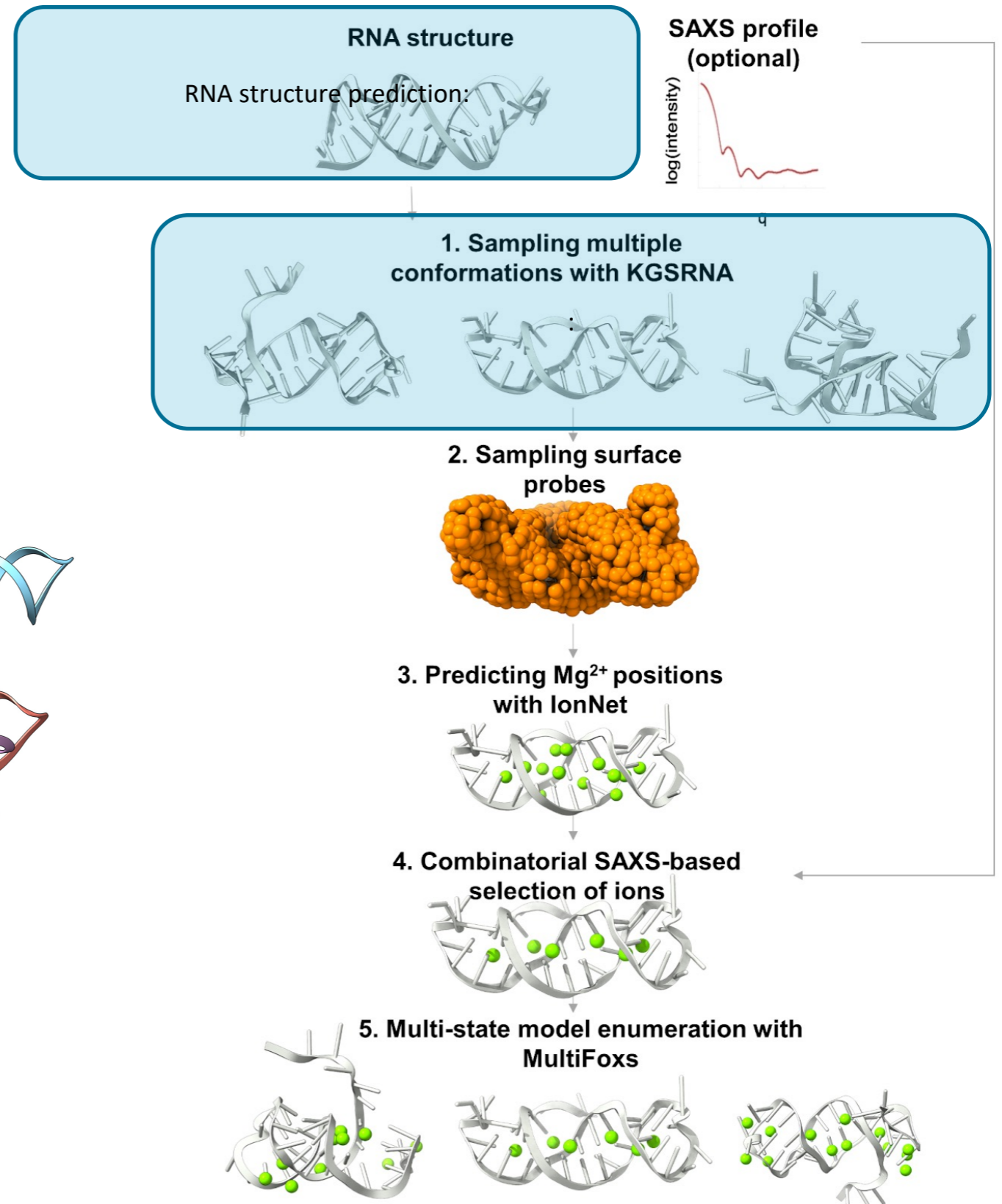
ipTM = - pTM = 0.27 [learn more](#)



Five AlphaFold3 models



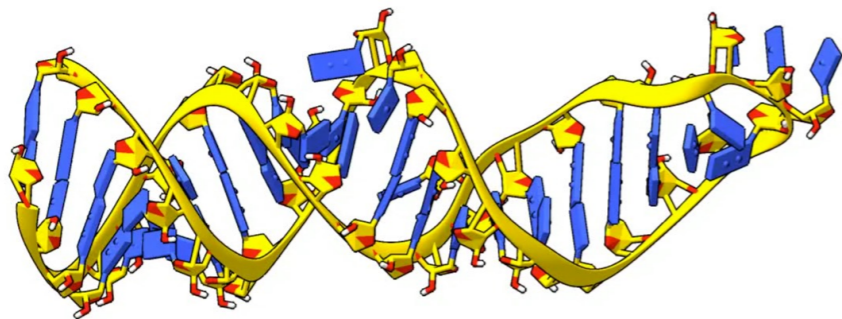
Edan Patt



Sampling RNA while preserving base pairing

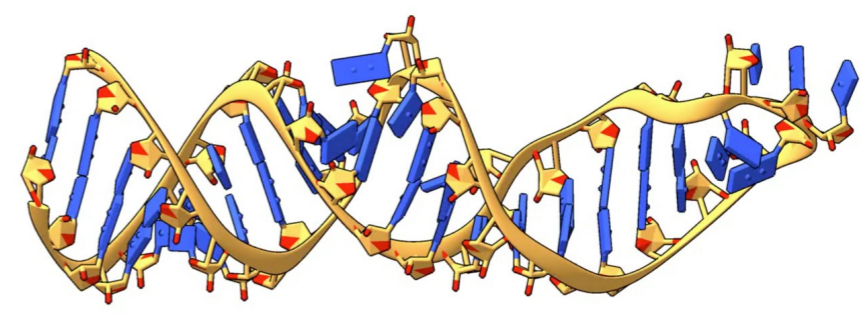
KGSRNA sampling

Kinematics-based approach to efficiently explore the native ensemble of RNA molecules

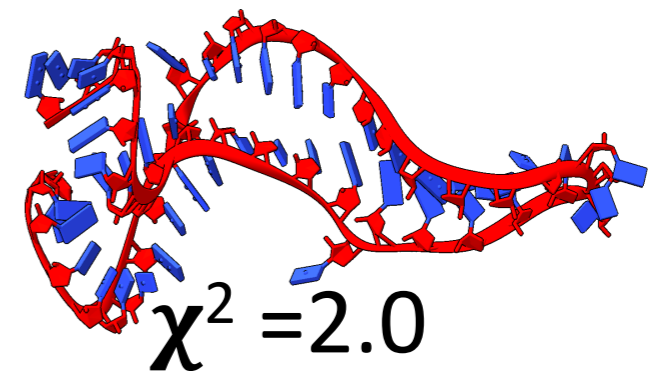
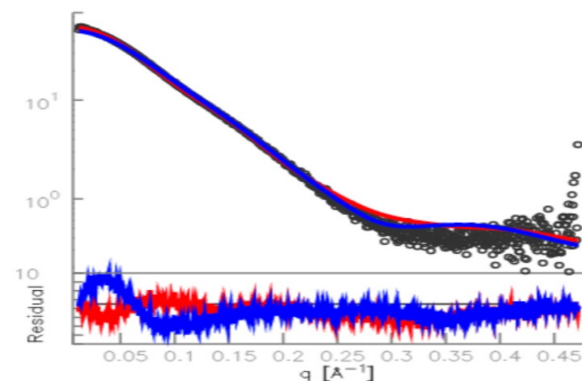
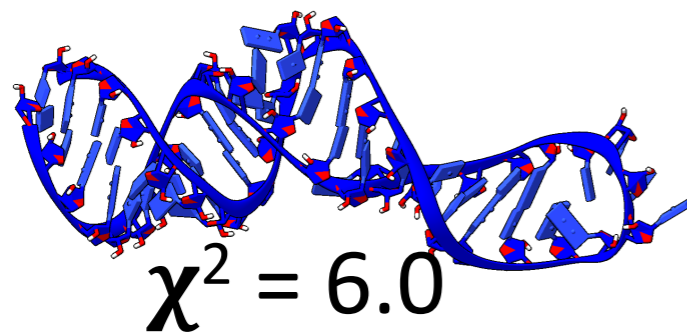


Normal mode sampling

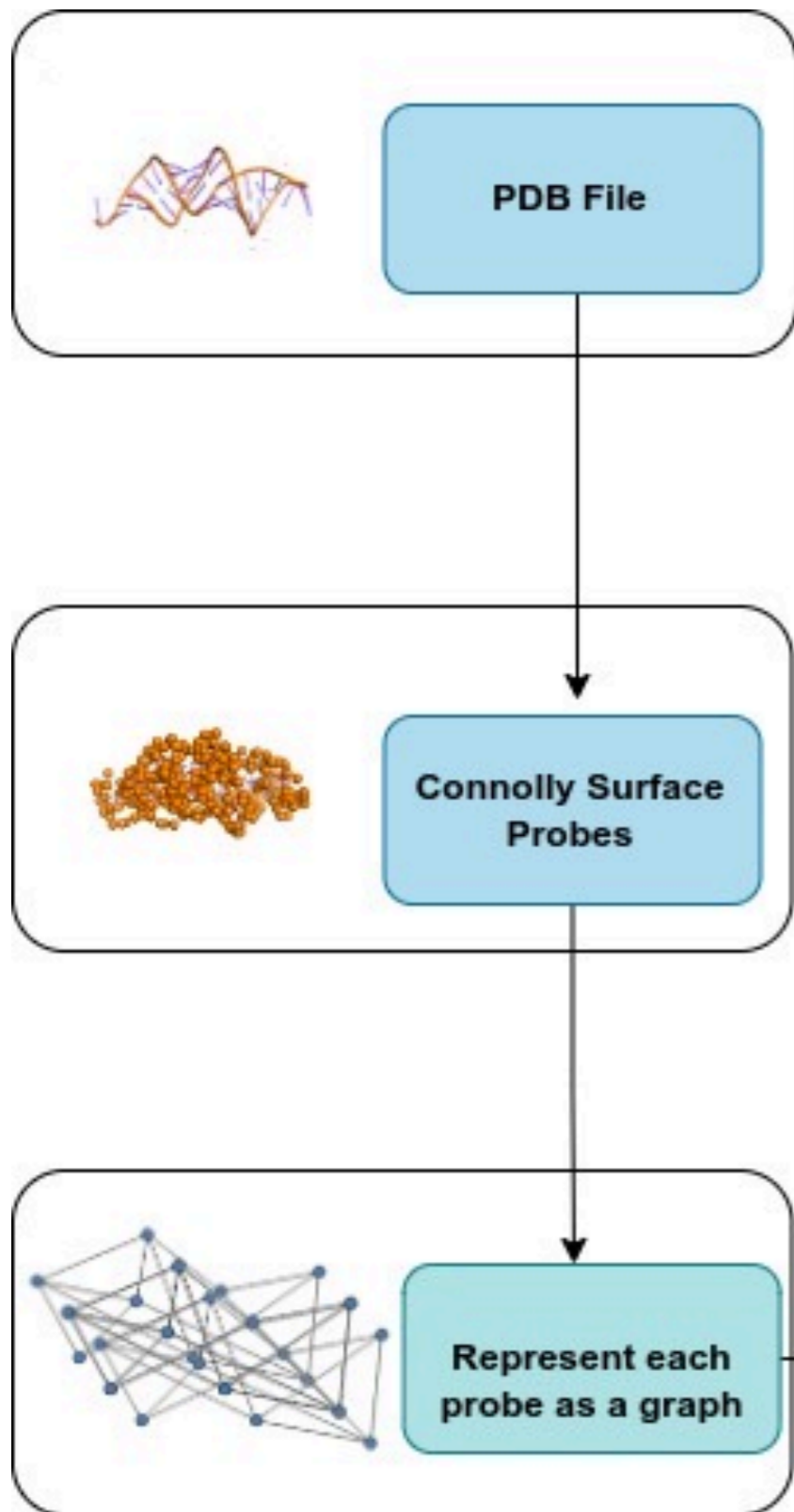
- The base pair interactions not preserved
- Oversampling
- Nonrealistic RNA structure



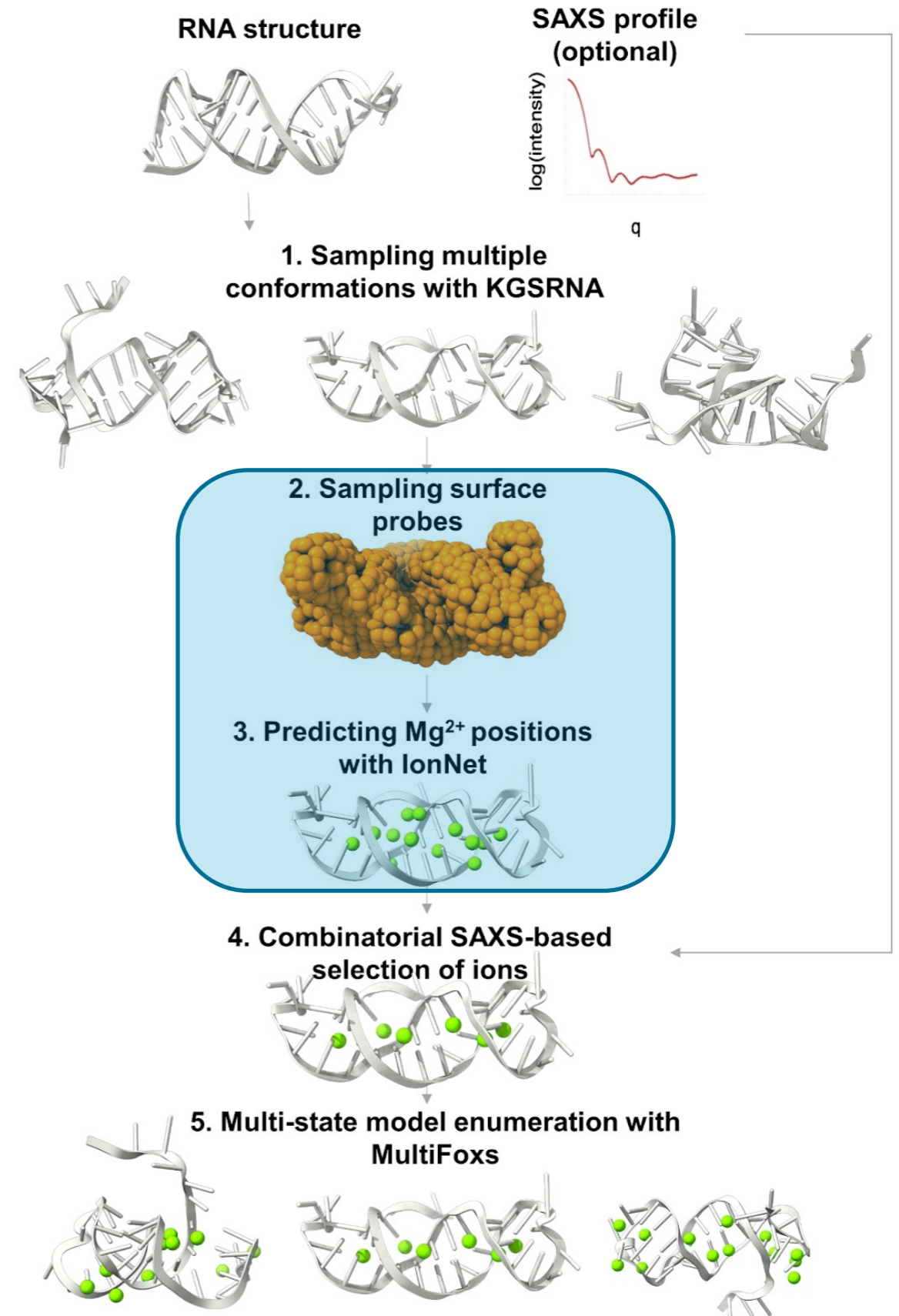
Normal Mode - best-fit conformer w/o Mg^{2+}
Overfit the SAXS data with nonrealistic conformers



IonNet: Mg²⁺ binding site predictor

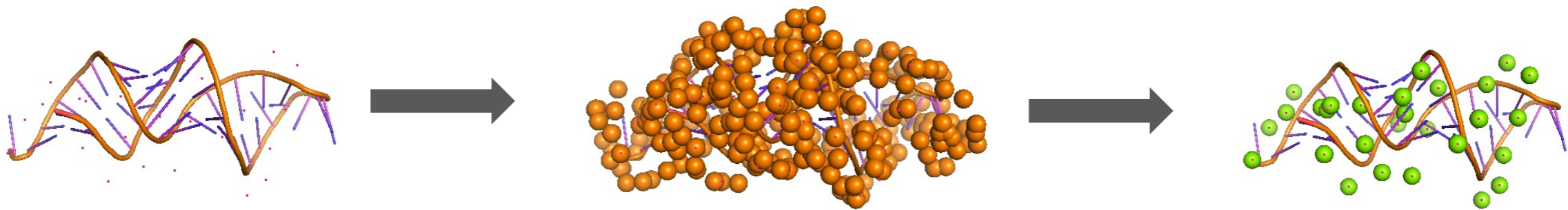
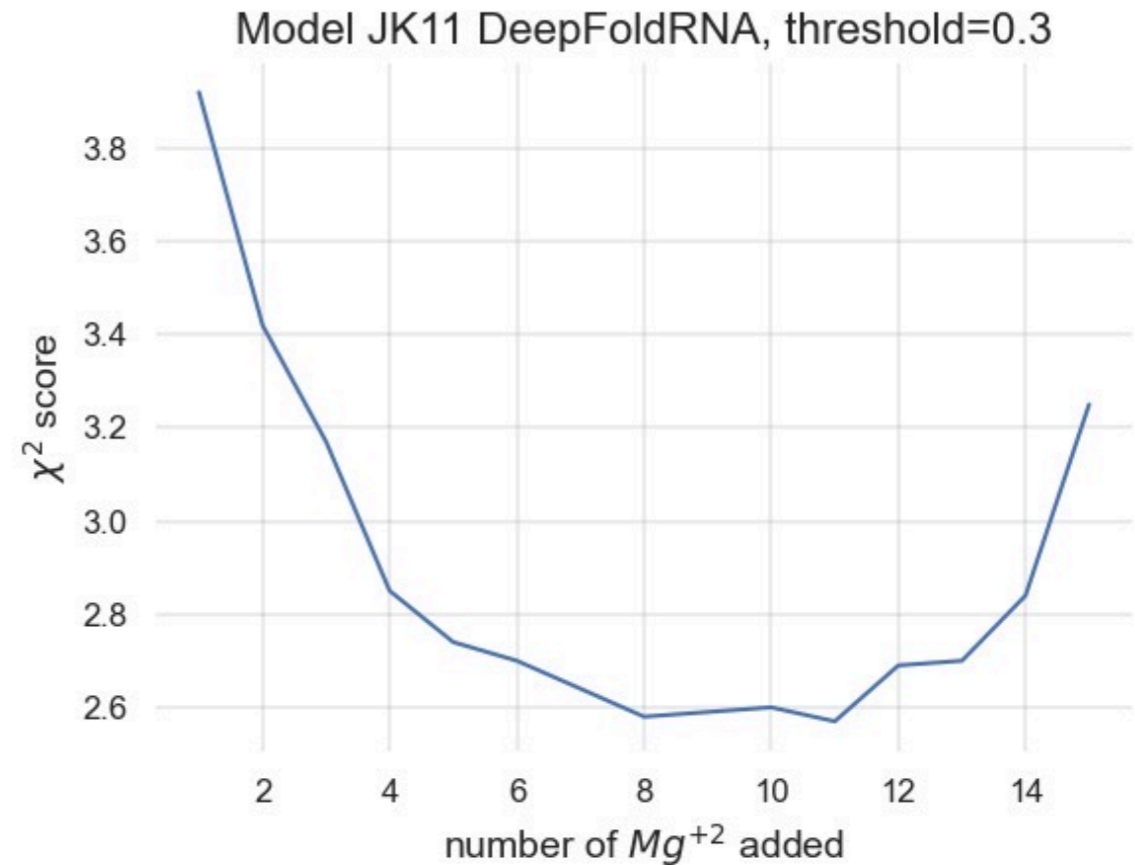


Graph Neural Networks (GNNs) to represent the geometry of the ion/water neighborhood



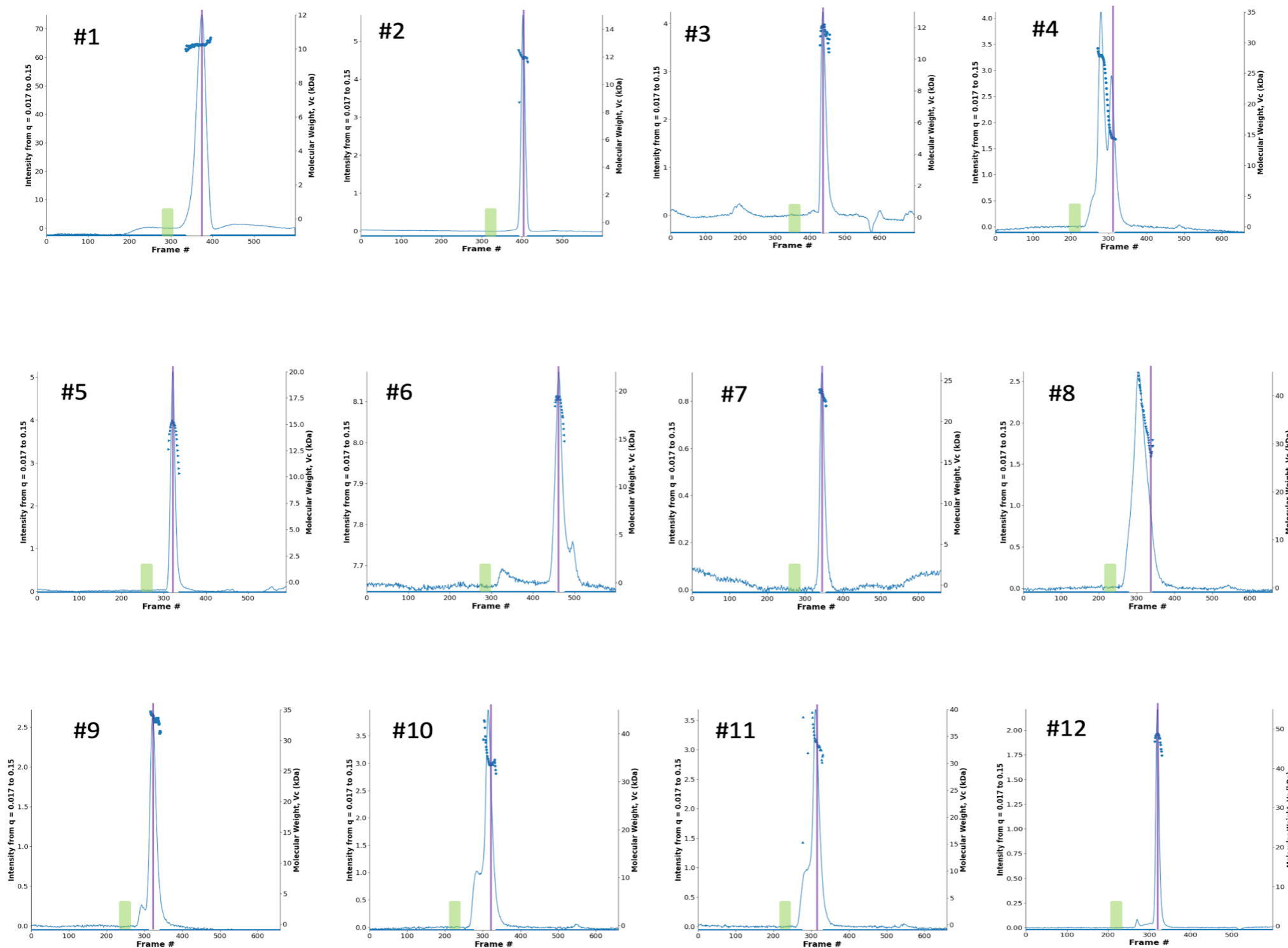
Using IonNet to predict Mg^{2+} positions for an RNA structure

- RNA 3D structure is covered with probes that are classified by the model
- The probes are added to the RNA, starting with the most likely one
- Fit to the experimental SAXS profile is used to select the optimal number of ions



High-quality benchmark dataset

Size exclusion coupled SAXS (SEC-SAXS) applied for RNAs benchmark of 12 RNA's

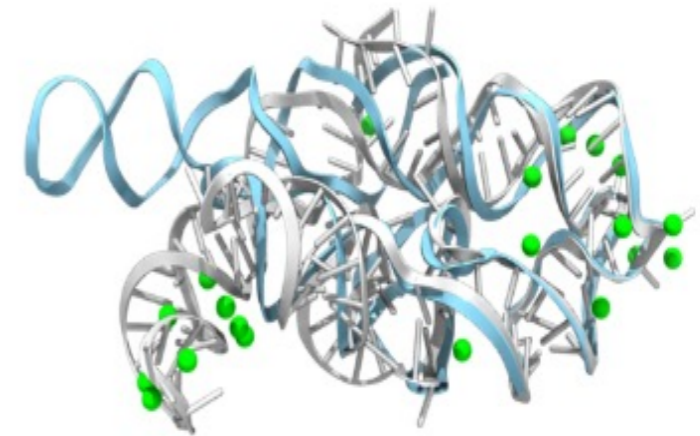
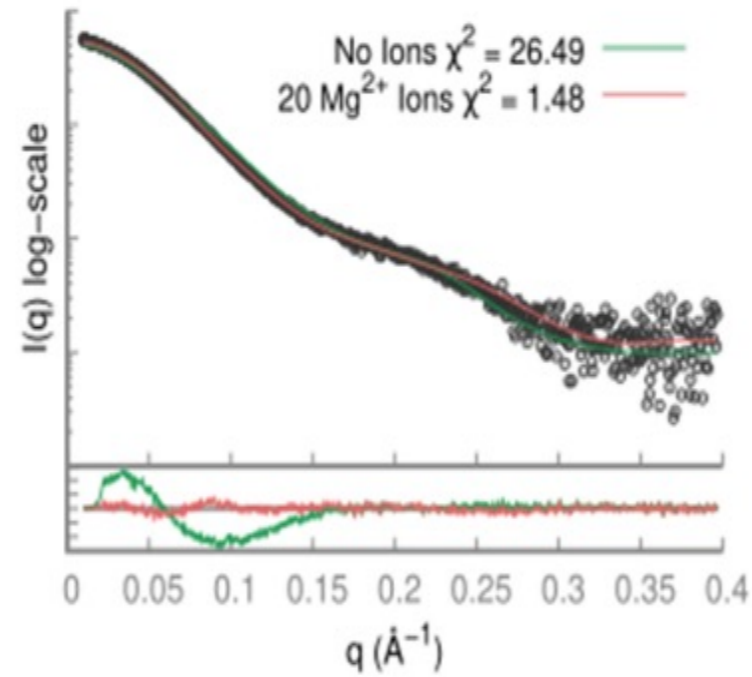
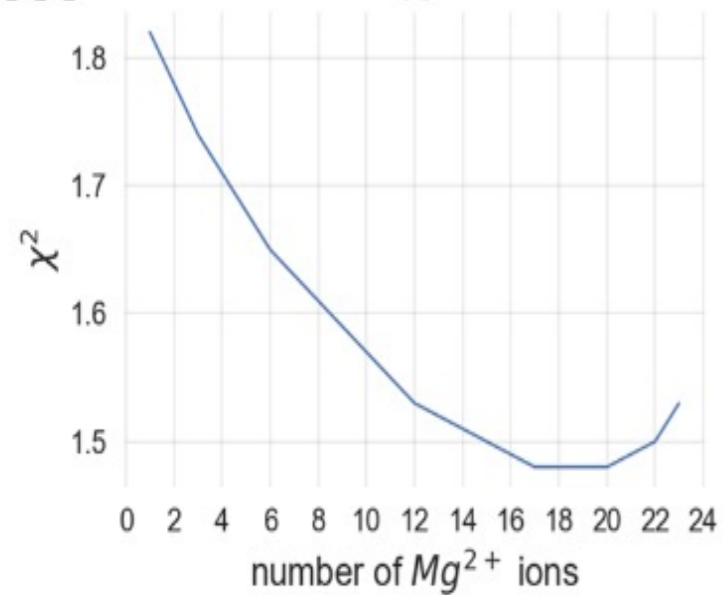


- SEC-SAXS chromatograms
- sample region
- buffer region
- Molecular weight

High-quality benchmark dataset

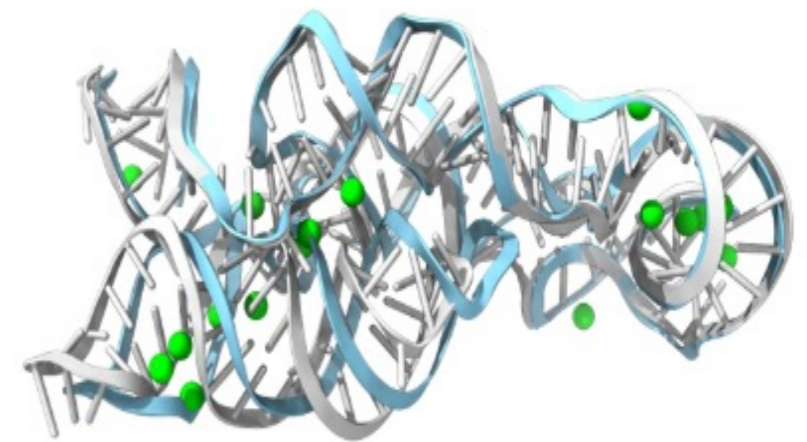
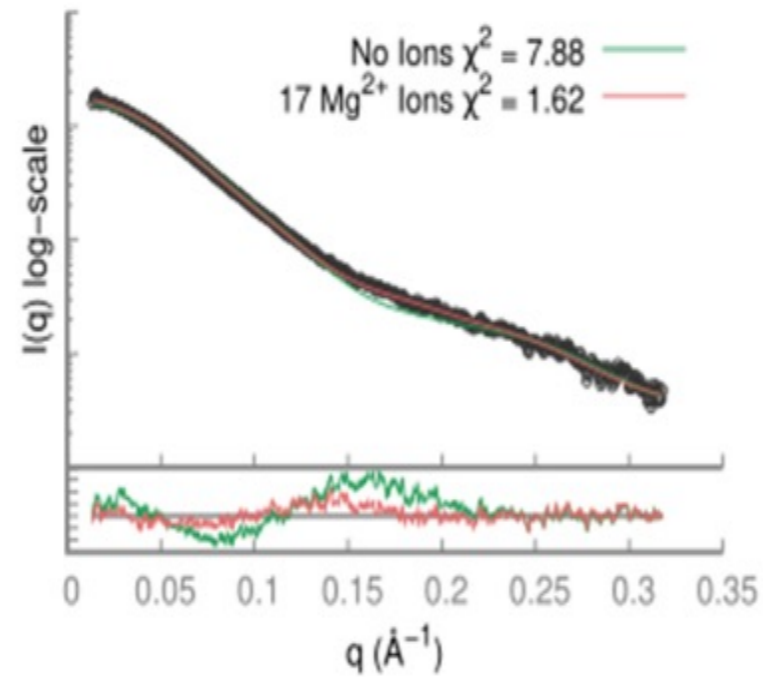
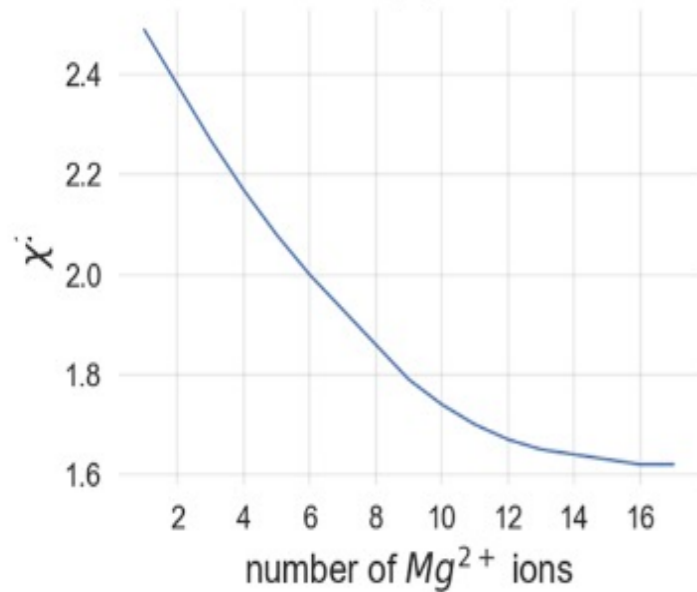
M

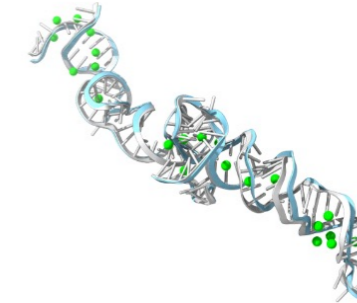
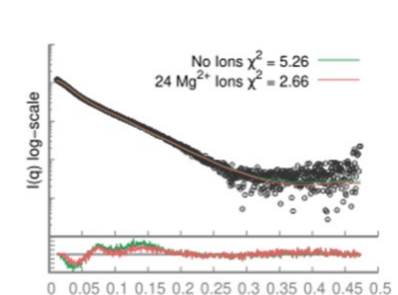
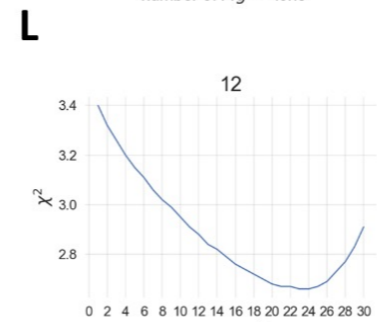
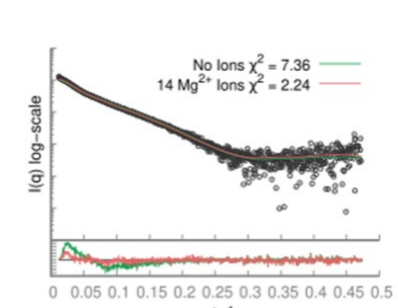
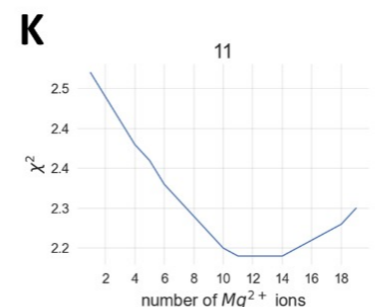
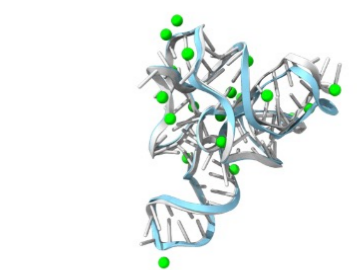
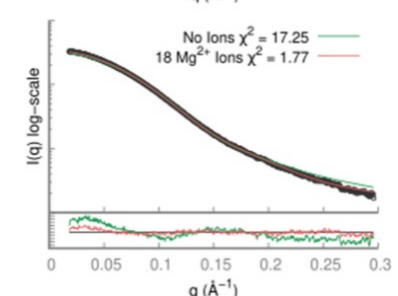
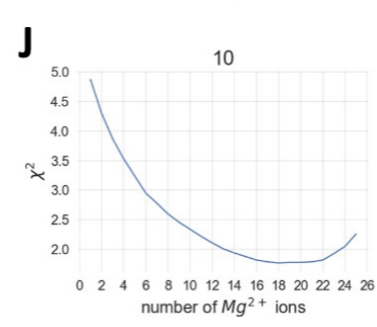
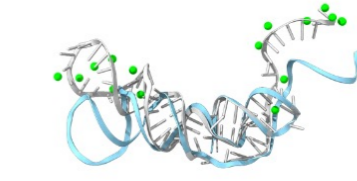
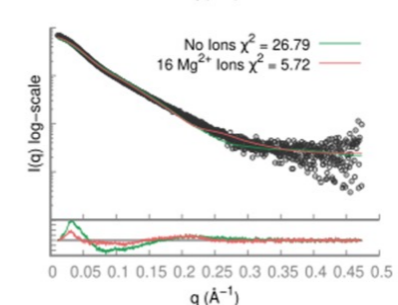
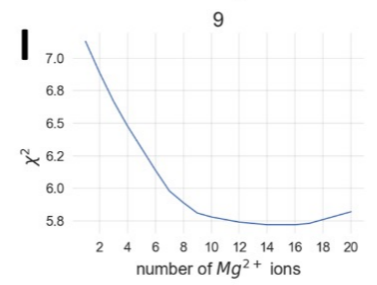
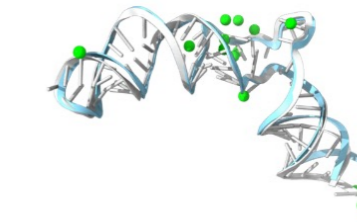
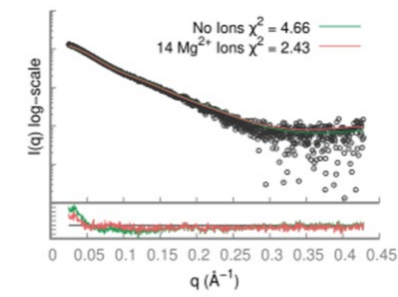
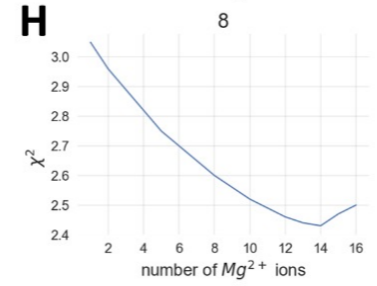
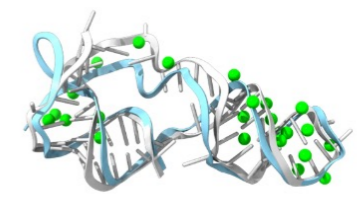
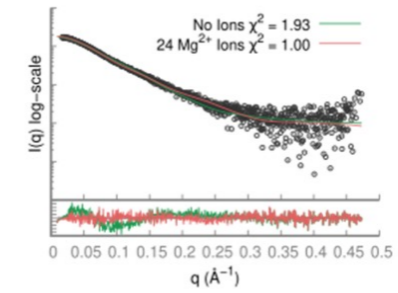
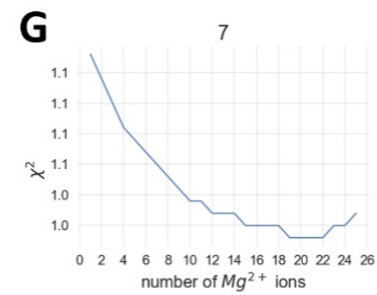
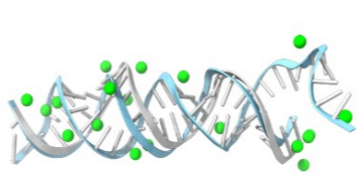
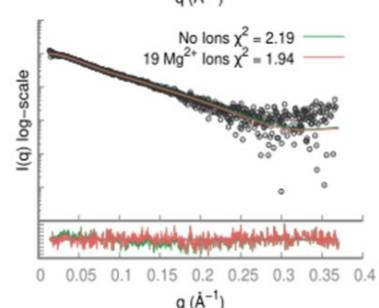
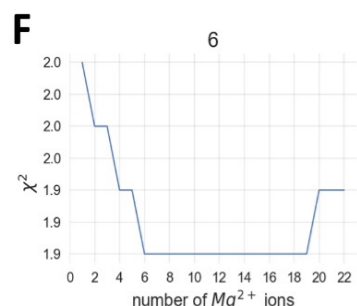
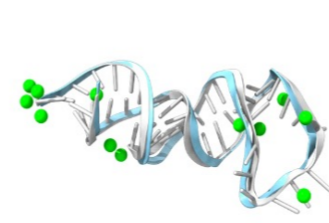
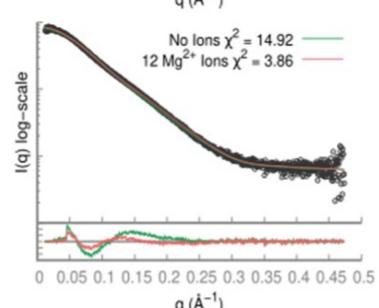
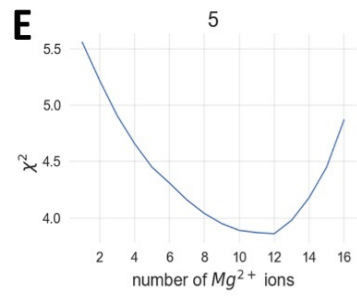
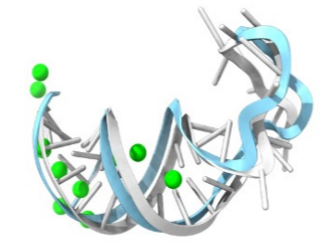
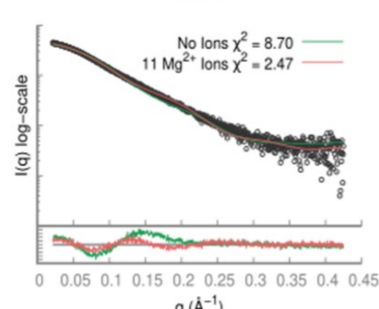
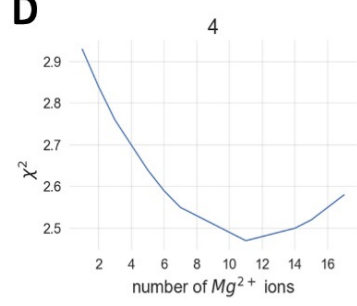
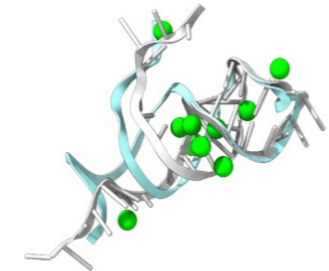
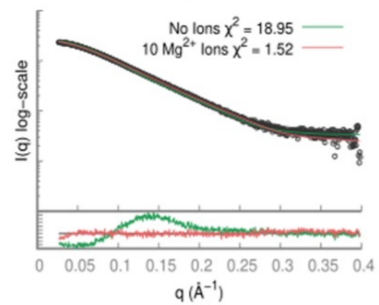
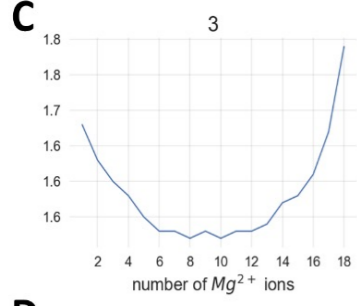
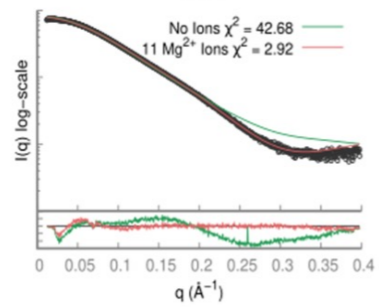
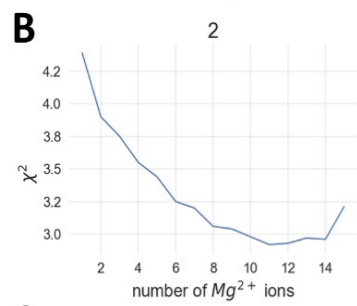
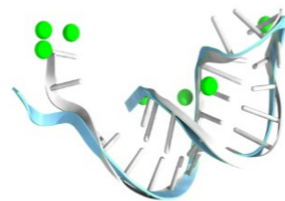
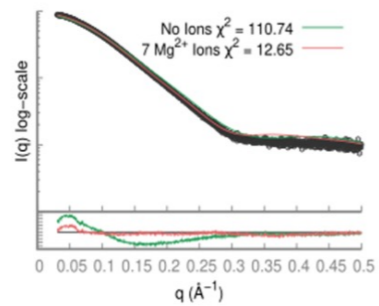
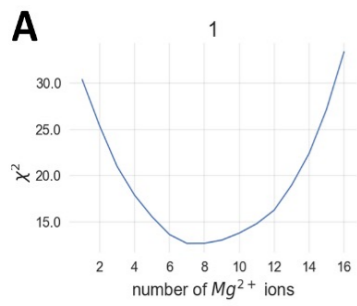
13



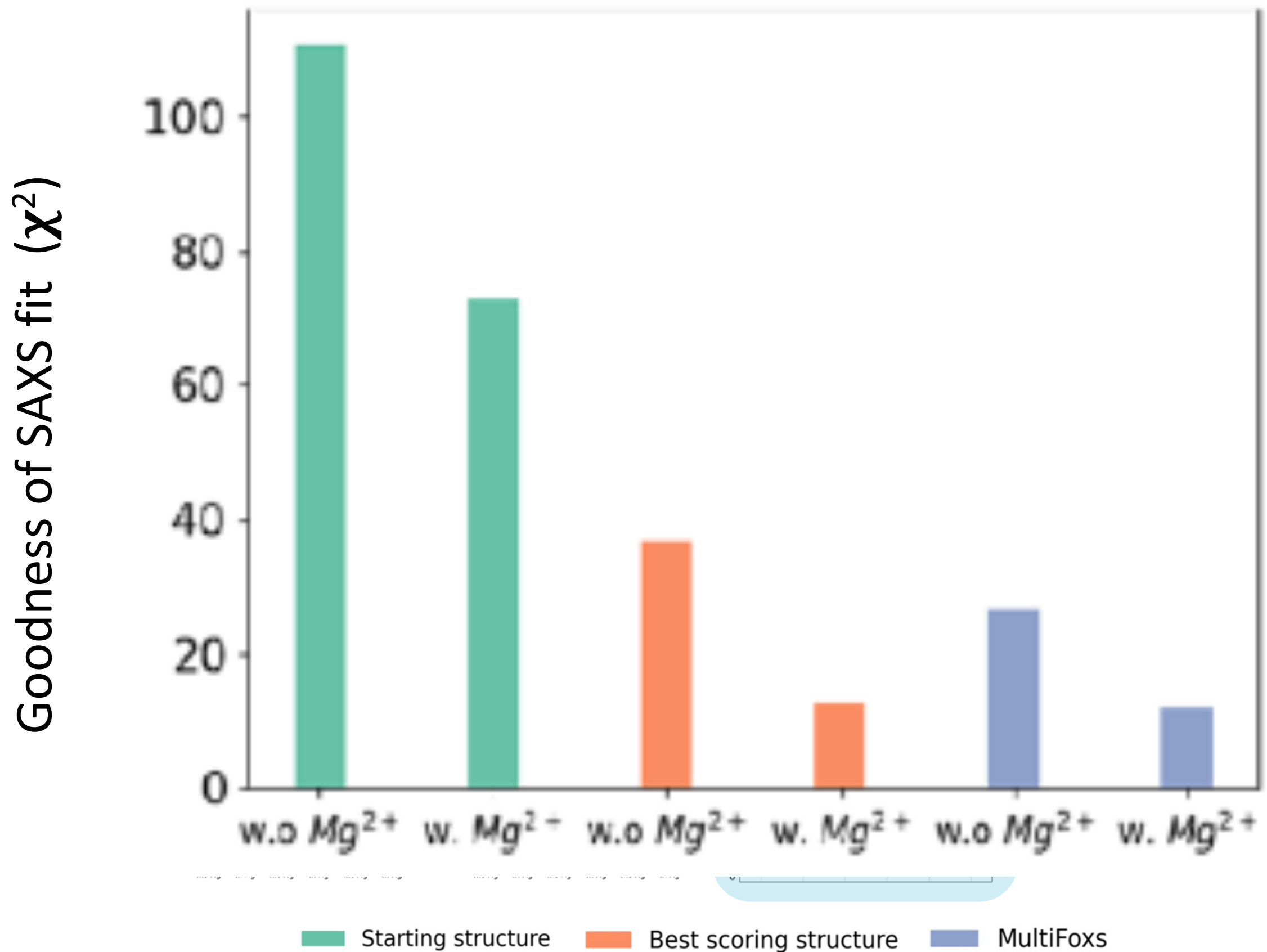
N

14

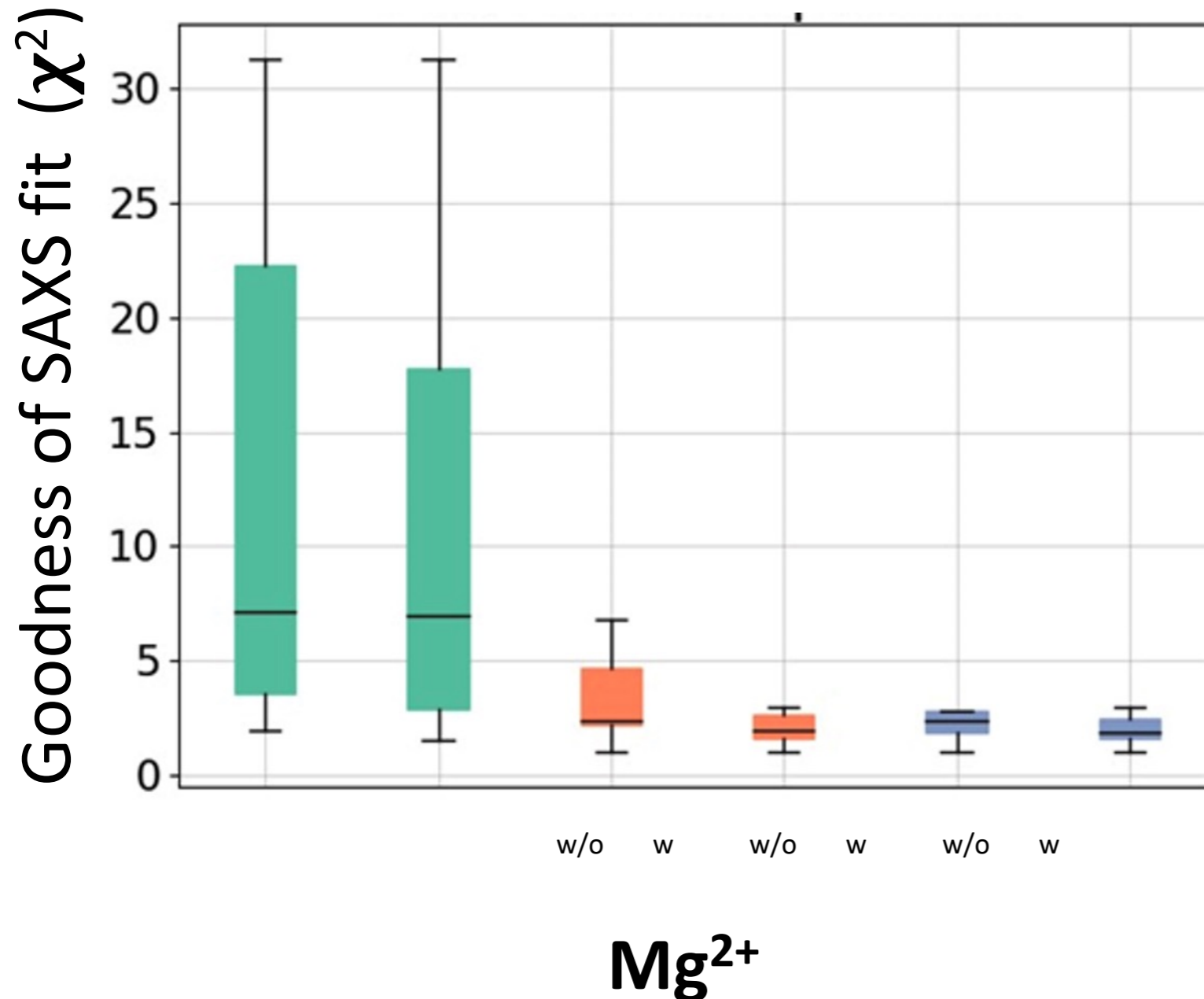




Flexibility, Mg^{2+} , and multiple states



Flexibility, Mg^{2+} , and multiple states



- RNA flexibility is responsible for poor SAXS fit
- The addition of Mg^{2+} ions improves the SAXS fit for the best scoring structure
- Multistate models have a minimal impact on the improvement of SAXS fit

SCOPER webserver

<https://bilbomd.bl1231.als.lbl.gov>

- Jobs
- BilboMD Classi
- BilboMD Auto
- Scoper
- inp Jiffy™
- PAE Jiffy™
- Users
- Admin

INSTRUCTIONS

SCOPER JOB FORM

Title

Upload your RNA *.pdb file

Select File

Upload your experimental SAXS data *.dat file

Select File

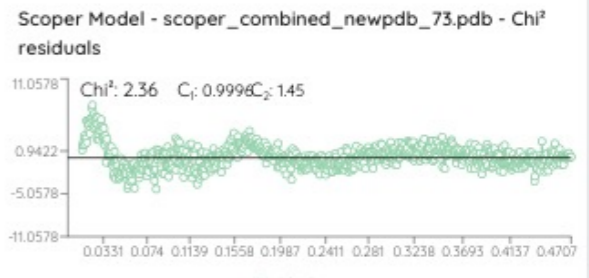
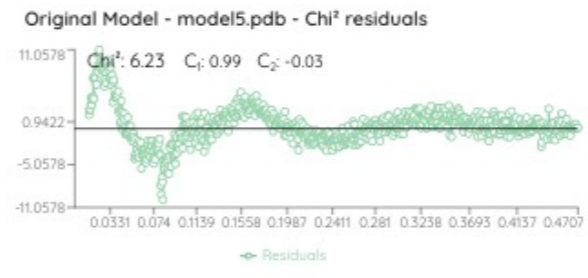
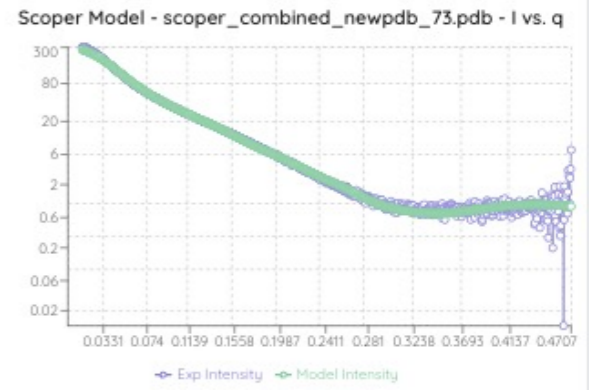
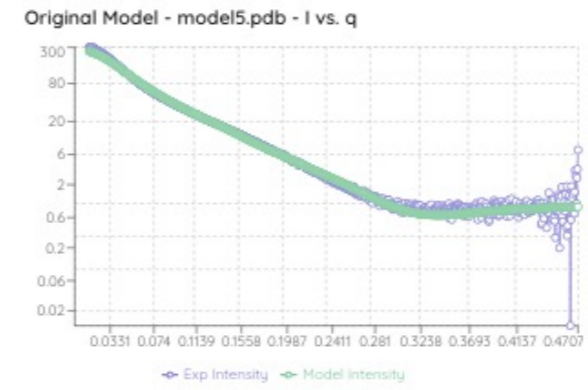
Submit ➤

SCOPER DETAILS

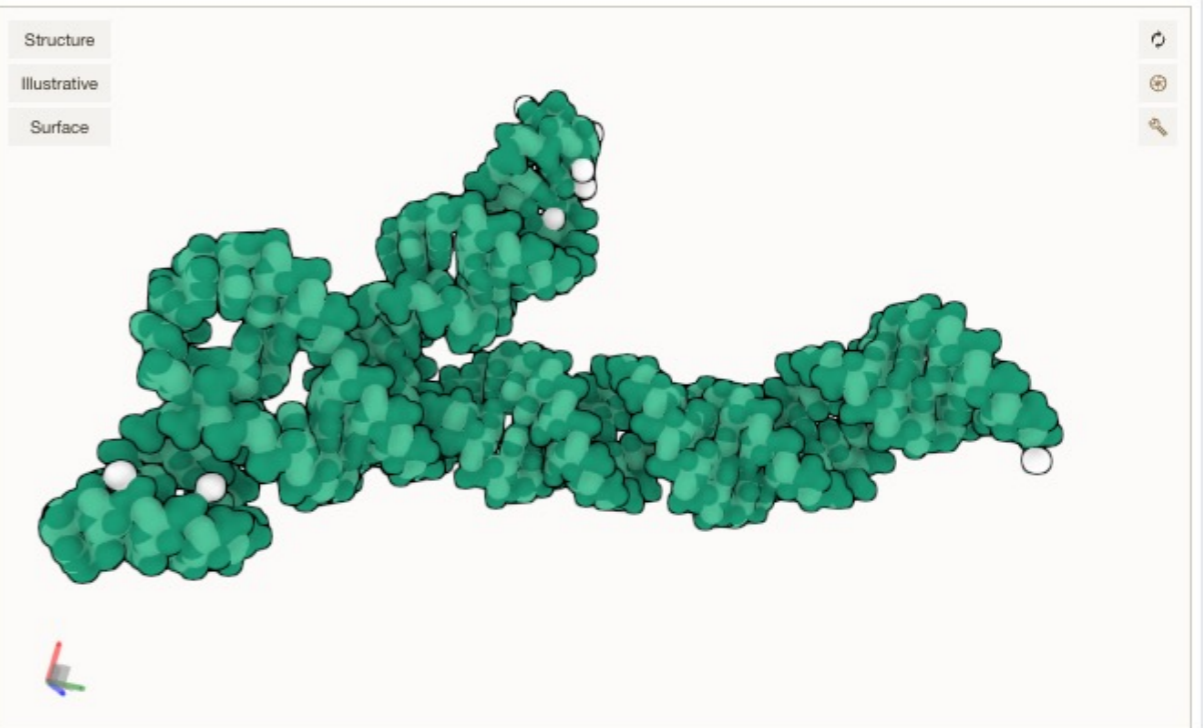
<input checked="" type="checkbox"/> reduce	<input checked="" type="checkbox"/> rnoview	<input checked="" type="checkbox"/> kgs	<input checked="" type="checkbox"/> foxs	<input checked="" type="checkbox"/> IonNet	<input checked="" type="checkbox"/> multifoys
KGS Number of Conformations to Generate	1000				
KGS Progress	1000				
Number of predicted Mg ions	15				

JOB DATABASE DETAILS

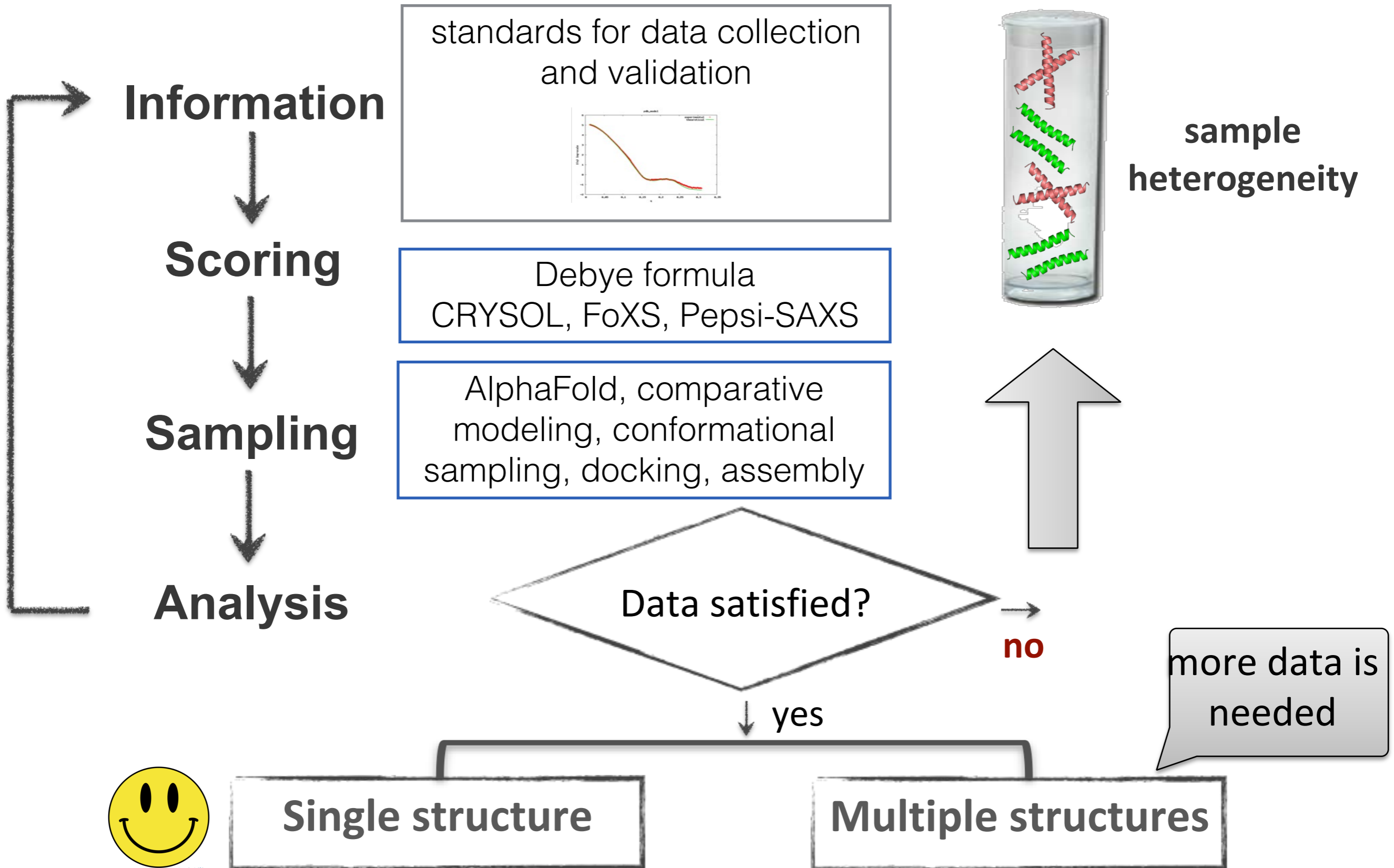
SCOPER FOXS ANALYSIS



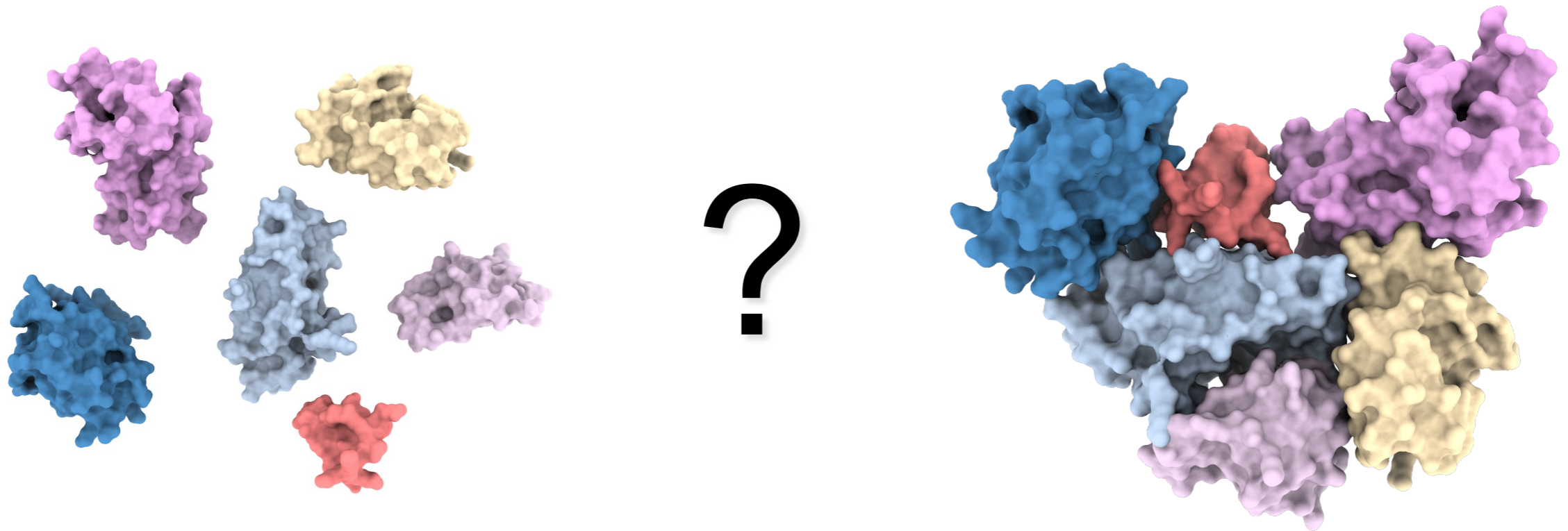
MOLSTAR VIEWER



Integrative Modeling and SAXS



Modeling protein interactions with or without SAXS



AlphaFold2 is effective in predicting complexes

- On typical benchmarks, **40-70%** of the complexes are correctly modeled vs. **20-30%** for docking algorithms
- Docking methods generate thousands of models including models that are close to the correct complex (PatchDock, ZDock, ClusPro...)
- Additional data, such as SAXS or crosslinks, helps to identify and validate correct models

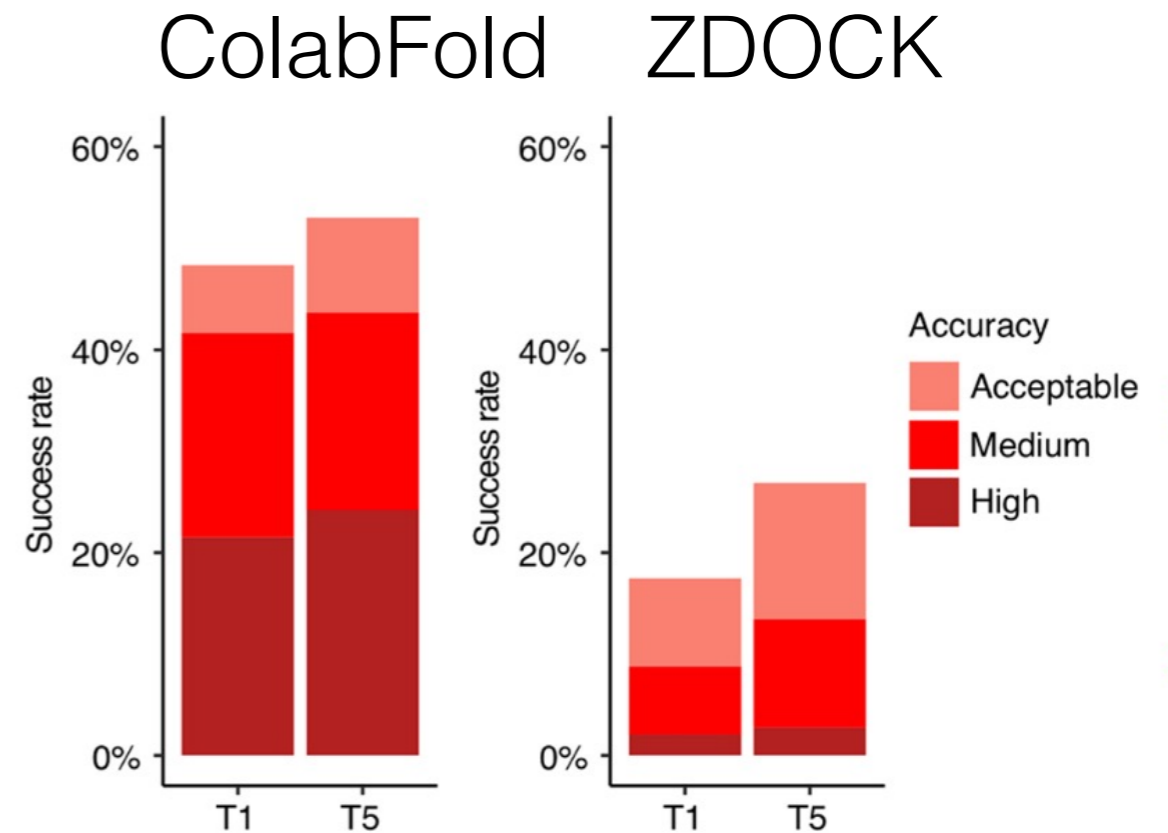


figure from Yin et al. Protein Science 2022

Docking



by Brian Jimenez

- success rate = # of benchmark complexes with acceptable or higher accuracy models, usually specified for topN predictions
- accuracy according to the CAPRI criteria (Acceptable, Medium, and High)

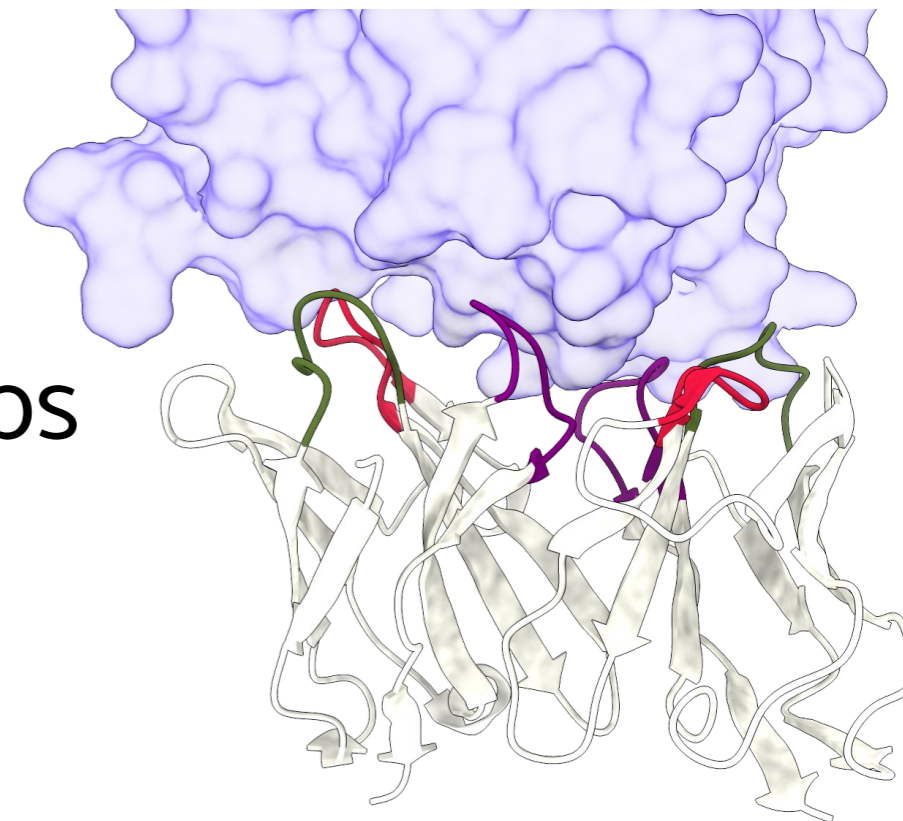
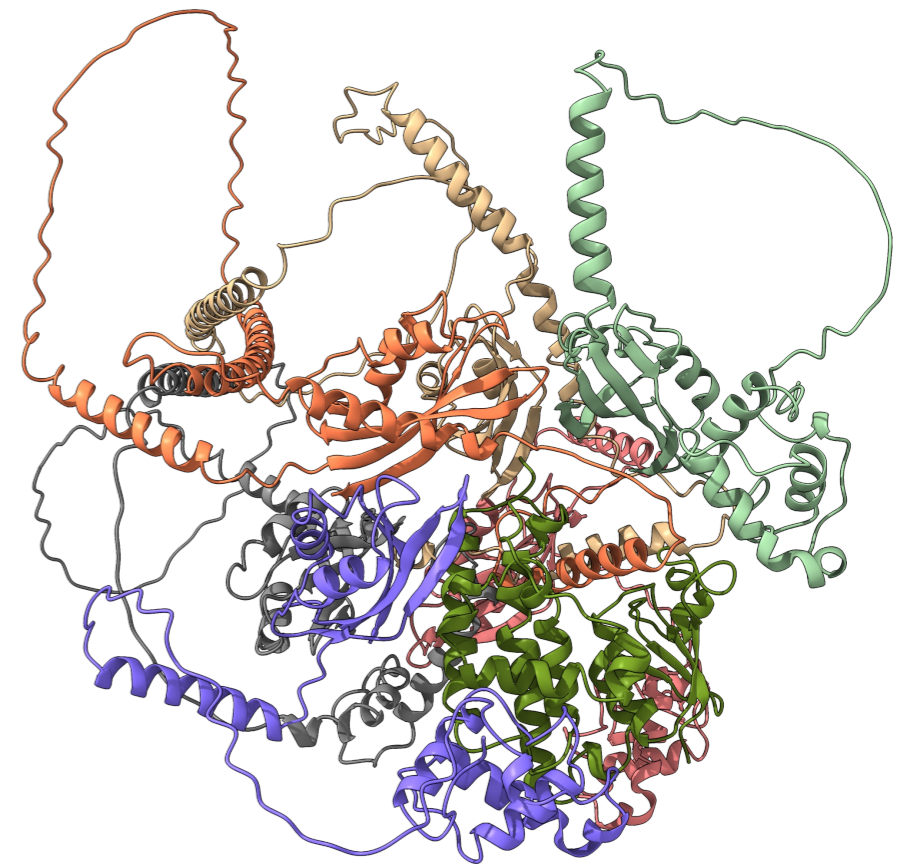
BUT...

Large assemblies are still difficult to model with AlphaFold-Multimer:

- GPU memory limitations
- sampling limitations
- out-of-domain inference
- converges to a single minima

Antibody-antigen systems

- interactions via highly variable loops



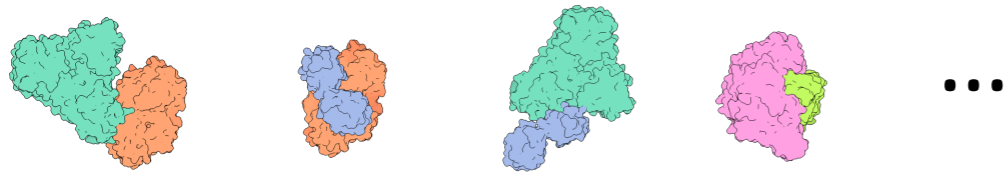
Combinatorial assembly based on AlphaFold2

subunit sequences:

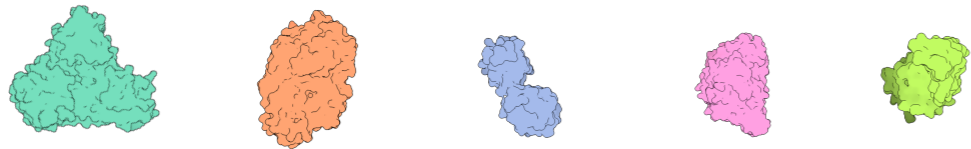


CombFold

Dock all pairs using AlphaFold (AF)

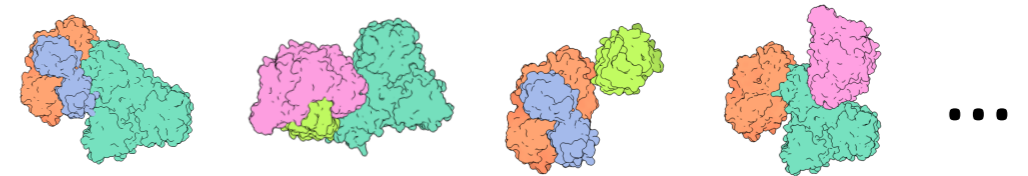


Select representative structure for each subunit

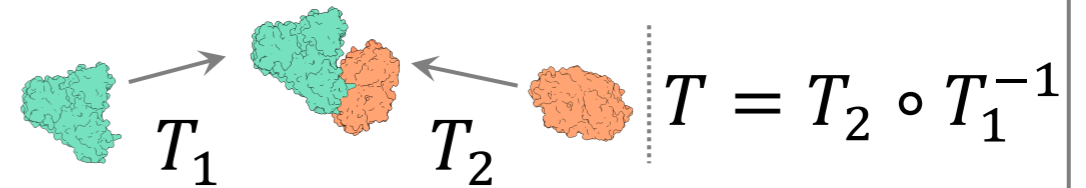


subunit structures

Use AF for larger subunit subsets

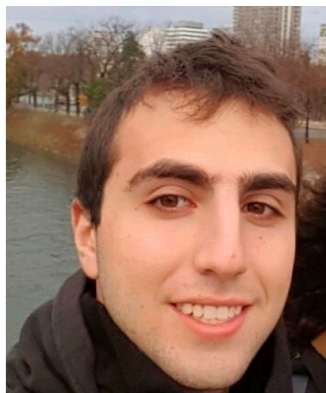
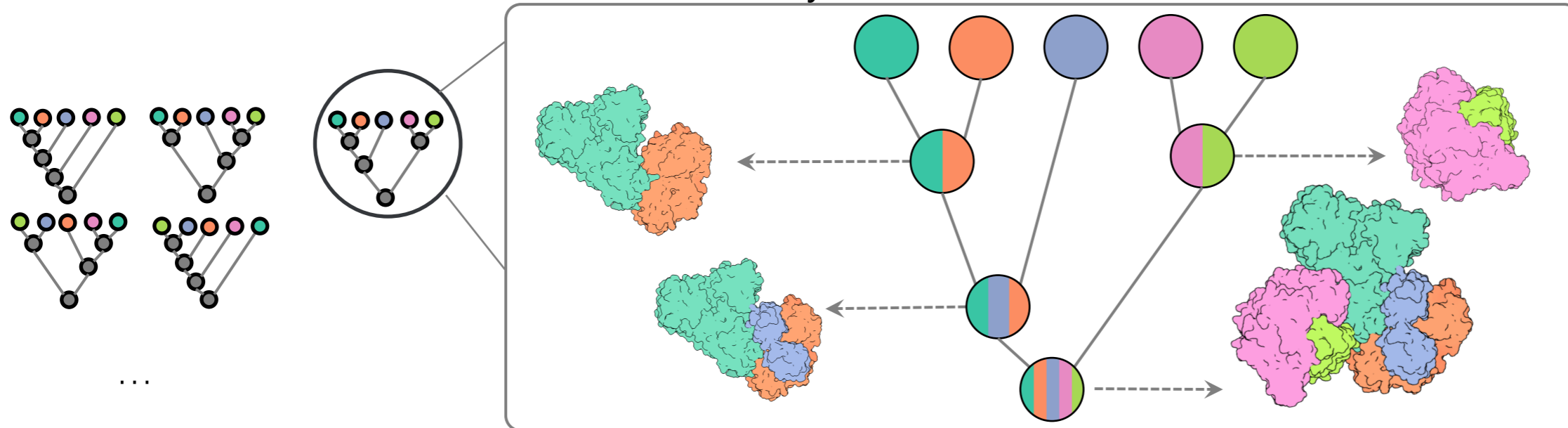


Convert AF models to transformations of representative subunit structures



pairwise transformations

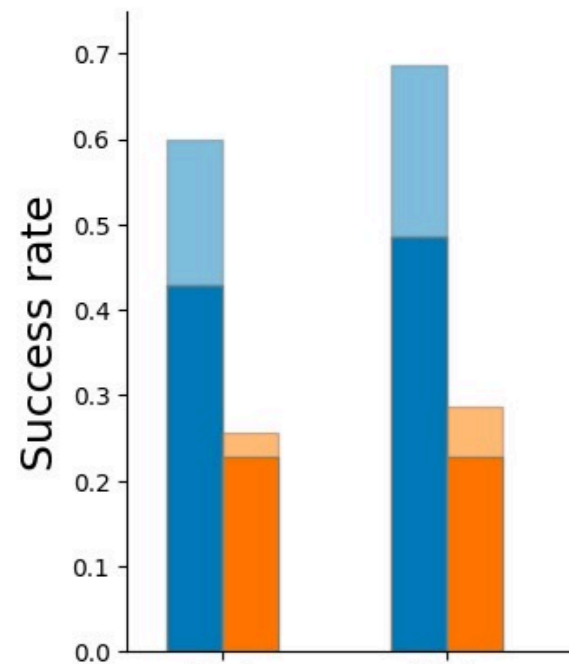
Combinatorial and hierarchical assembly



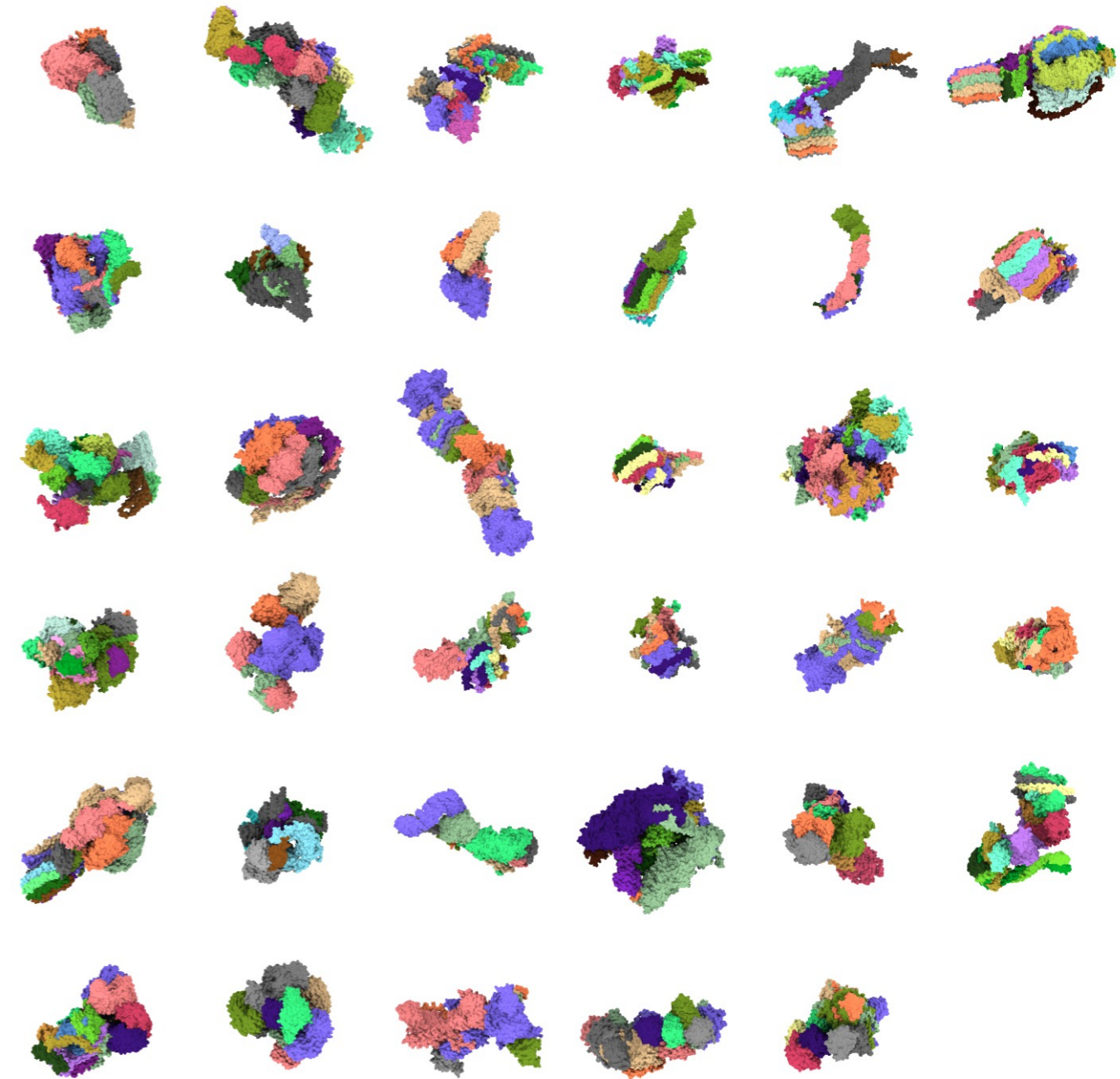
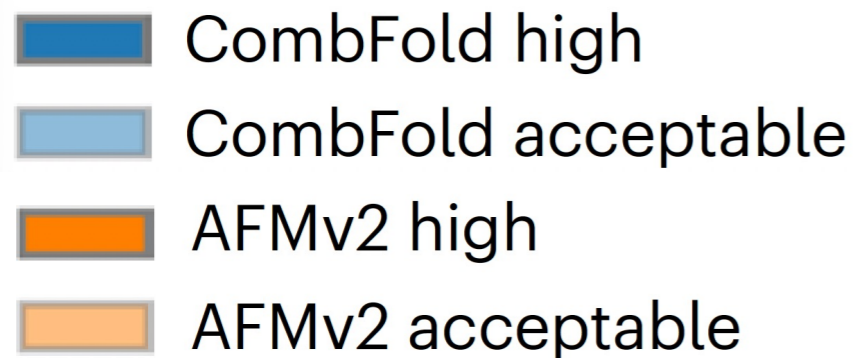
Ben Shor

Benchmarking heteromers

- 35 complexes (no overlap with the AFM training set)
- 5-20 chains
- 1,700-8,000 amino acids

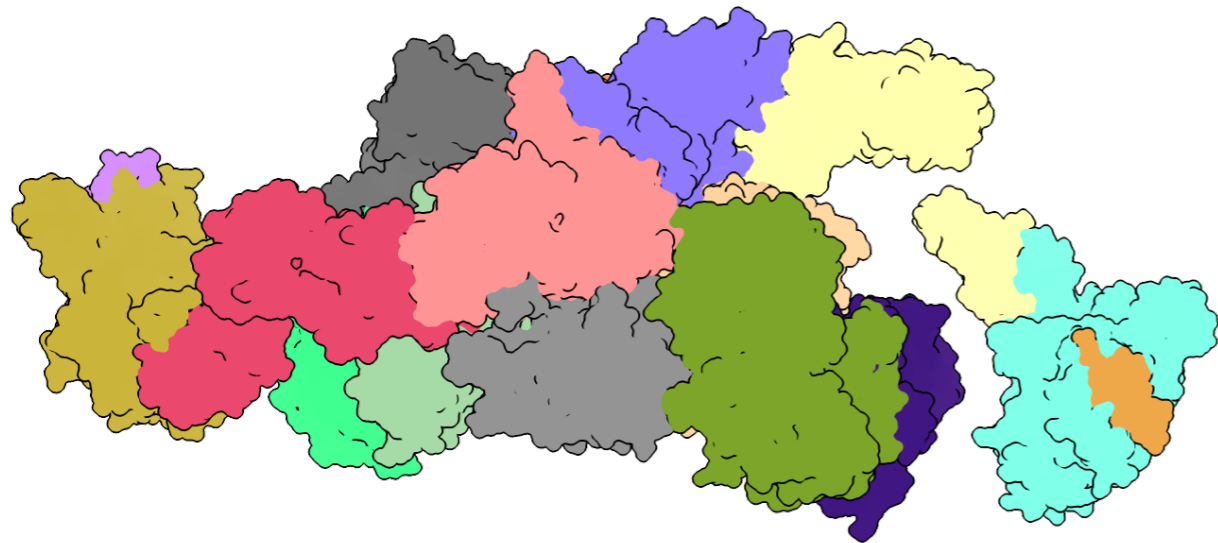


Top-1 Top-5

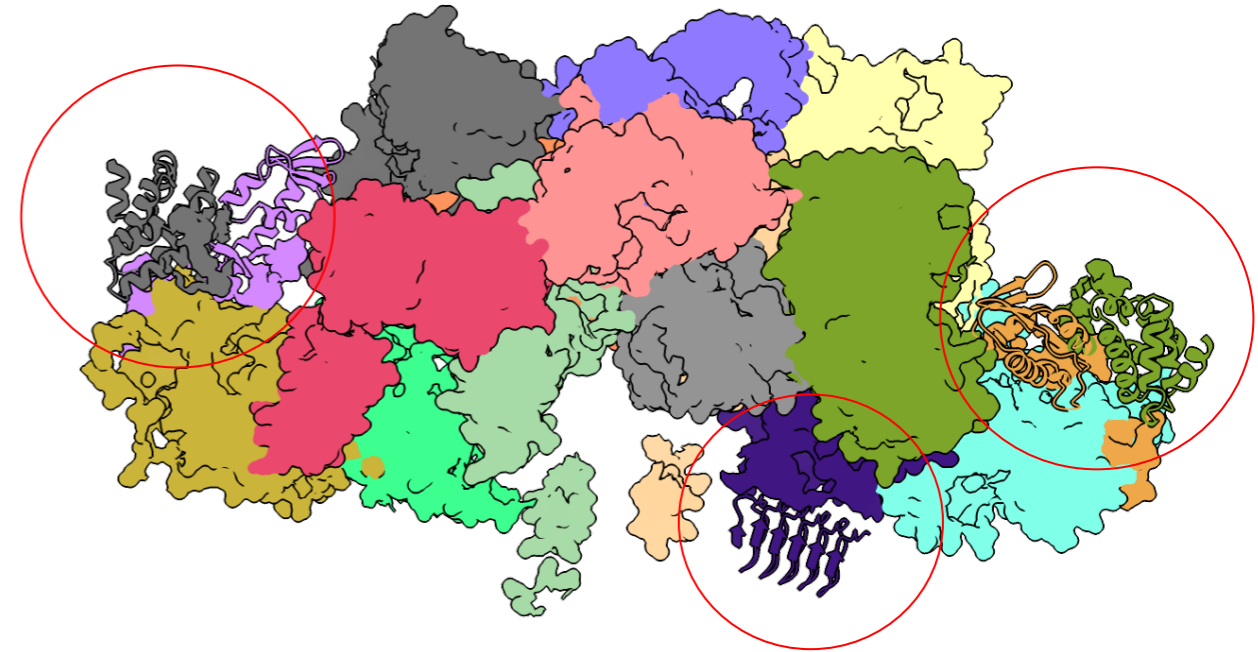


- High: TM-score > 0.8
- Acceptable: TM-score > 0.7

Modeling subunits that are missing in PDB structures



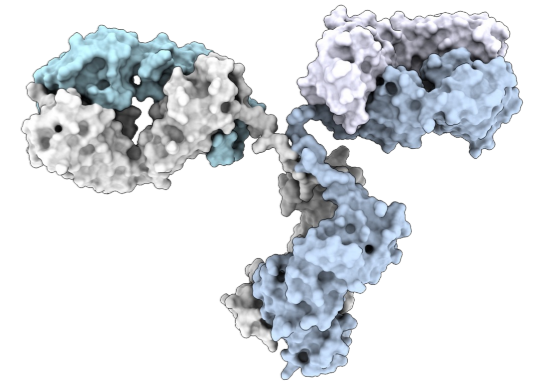
eIF2B:eIF2 complex (PDB 6I3M)
4,680 amino acids



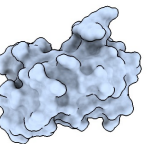
TM-Score 0.79
6,114 amino acids

- 20% increase in structural coverage compared to PDB entries in our Benchmark

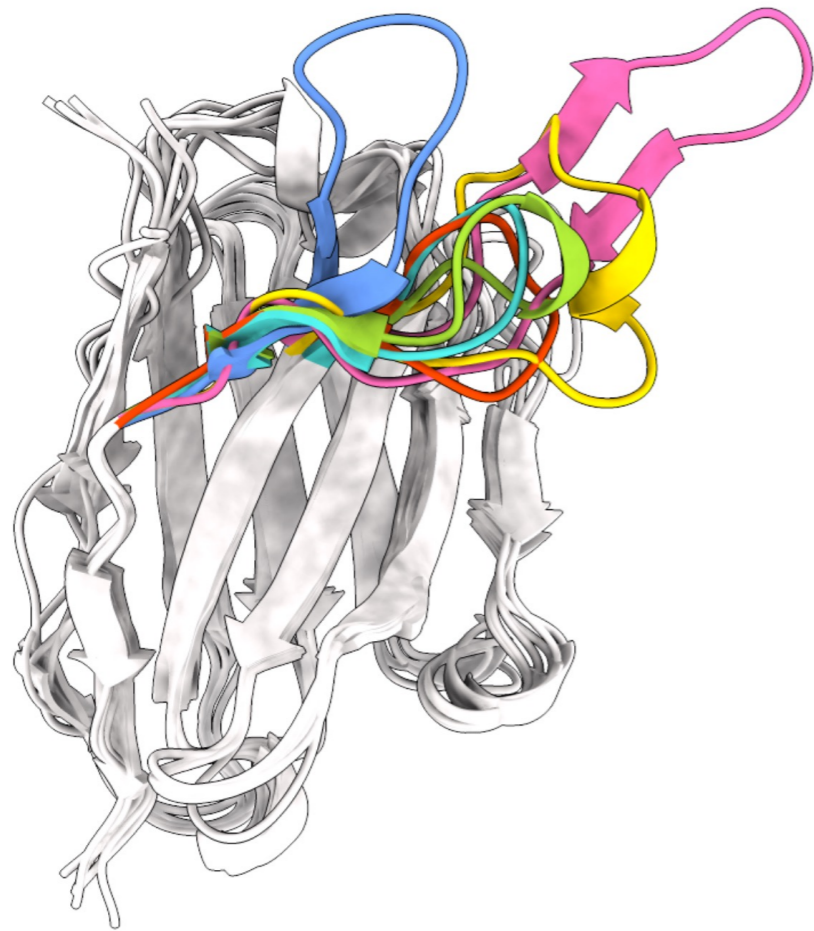
Why antibodies?



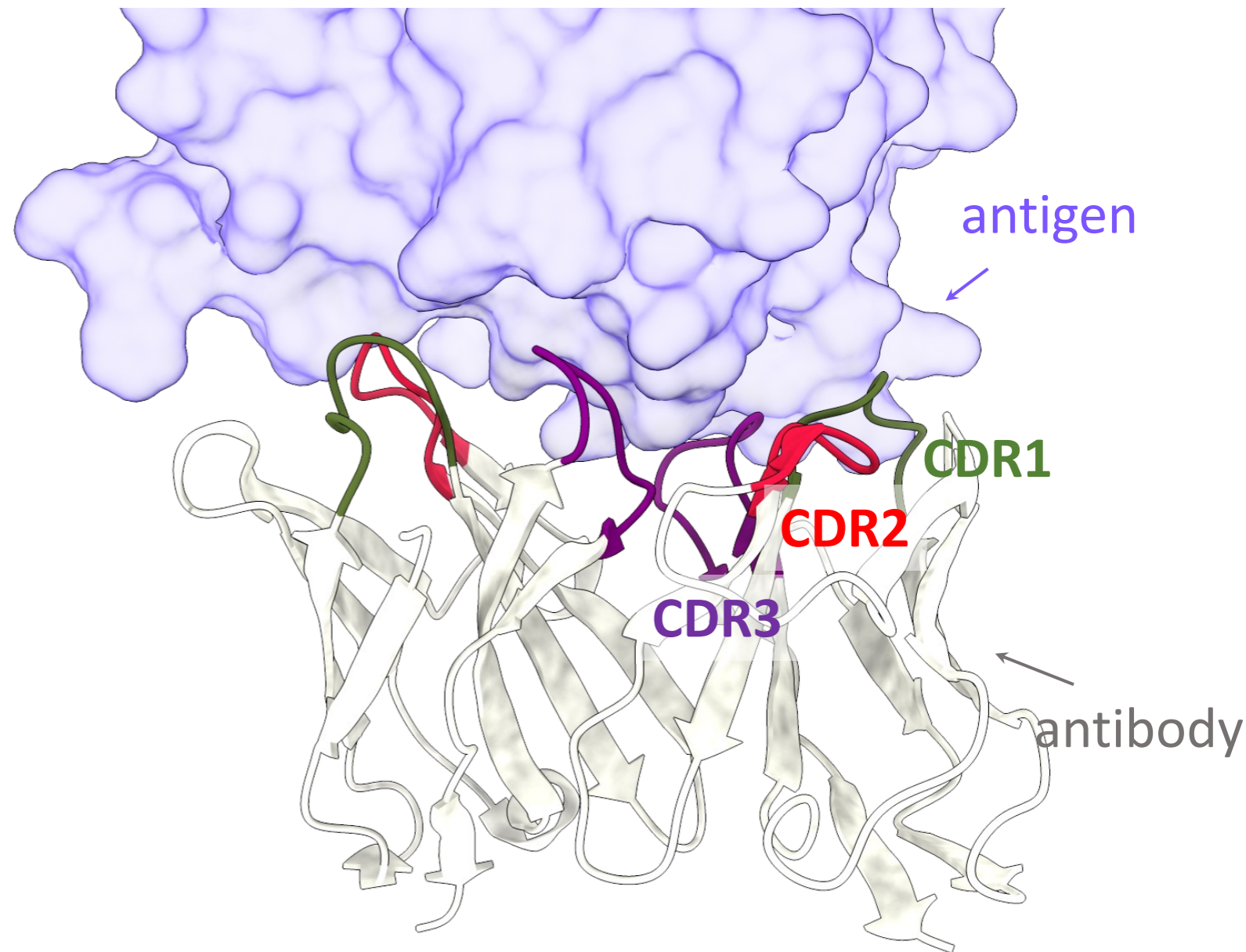
- A key component of the adaptive immune system
- A rapidly growing class of human therapeutics for a range of diseases, including **cancer, autoimmunity, inflammatory diseases, viral infections**
- There are over 100 approved antibody-based therapeutics and over 1,000 in clinical studies for a wide range of diseases.
- Nanobodies, heavy chain only antibodies - small, stable, highly similar to IgGs
- Accurate high-throughput computational methods have the potential to greatly accelerate the **discovery of new therapeutic antibodies**



Antibodies have a conserved frame region and variable loops



Alignment of frame regions (heavy chain)



Open challenges

Folding

Input: antibody sequence

Output: 3D structure



Docking/specificity

Input: structures (or sequences)

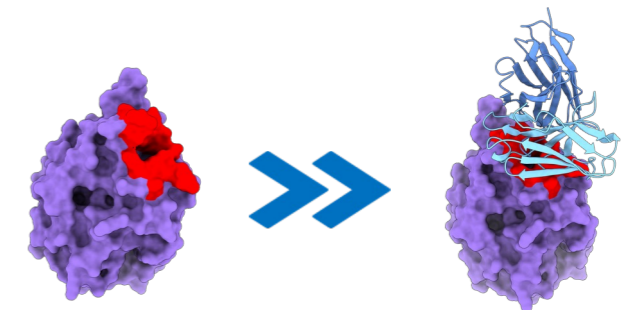
Output: antibody-antigen complex 3D structure



Design

Input: antigen structure and **epitope** (in red)

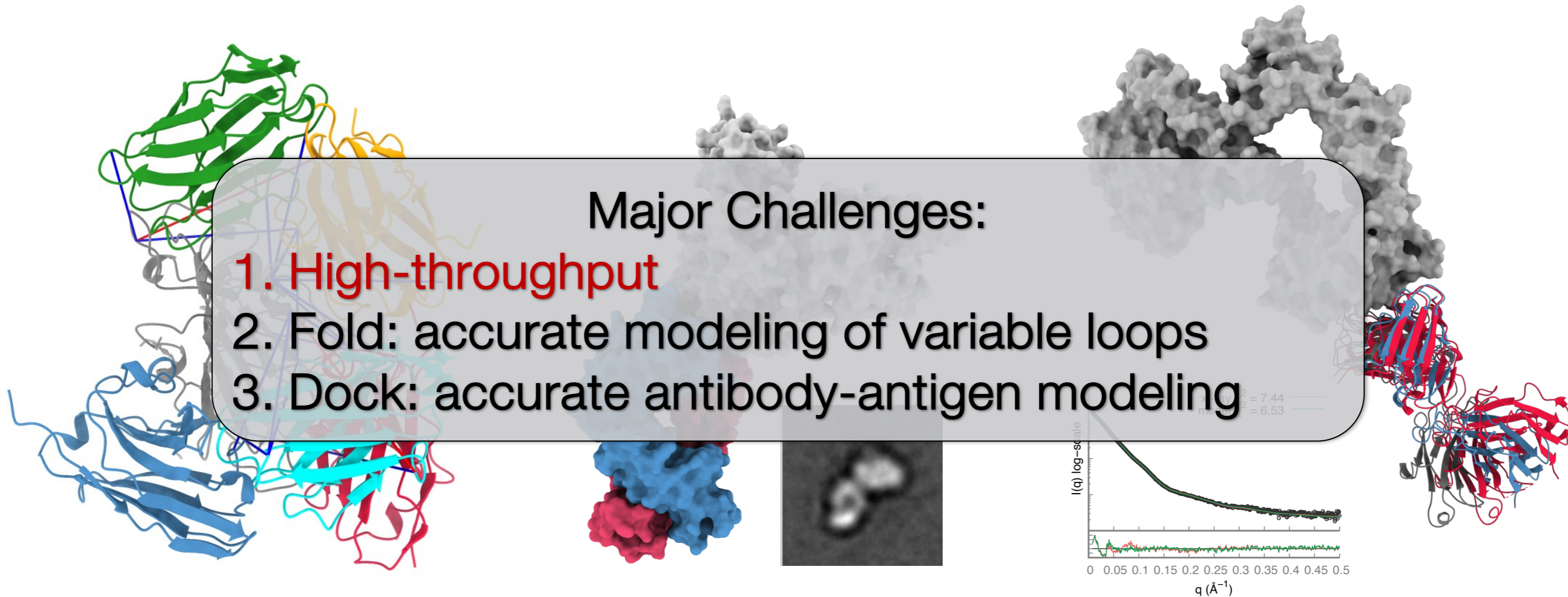
Output: antibody sequence/structure that binds to the given epitope



Integrative modeling of antibody-antigen complexes

Major Challenges:

1. High-throughput
2. Fold: accurate modeling of variable loops
3. Dock: accurate antibody-antigen modeling

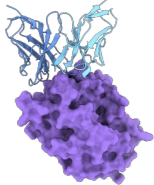
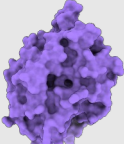
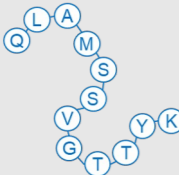
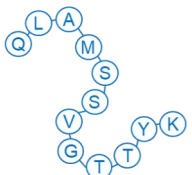


*SARS-CoV-2 nanobodies
with crosslinks
Xiang et al. Science 2020*

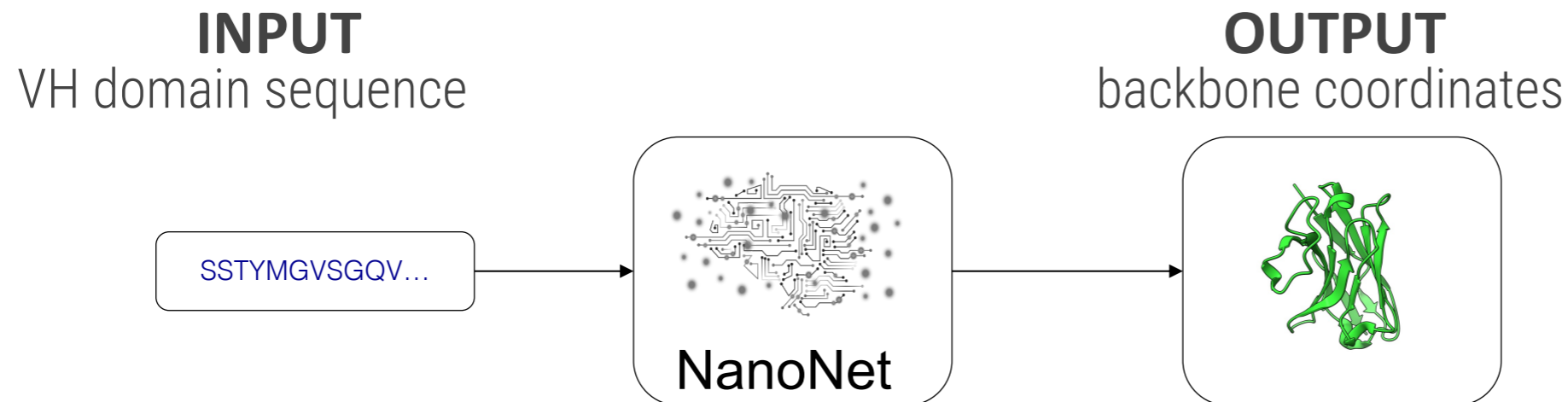
*PCSK9 antibody with 2D EM
Schneidman et al.
Bioinformatics 2012*

*EGFR antibodies with SAXS
profiles
Cohen et al. Meth. Enzymol
2023*

Data gaps for ML

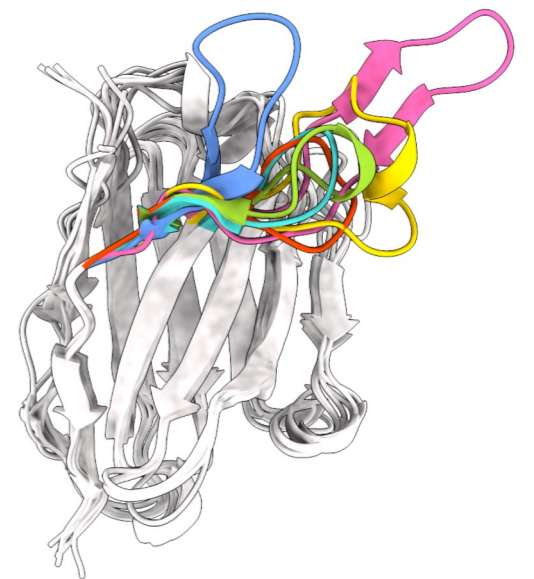
		Antibodies	TCRs
	complex structures	$\sim 10^3$	$\sim 10^2$
 	antigen-specific sequences	$\sim 10^5$	$\sim 10^5$
	TCR/antibody sequences	$\sim 10^9$	$\sim 10^9$

NanoNet: end-to-end antibody, nanobody, and TCR modeling without MSA

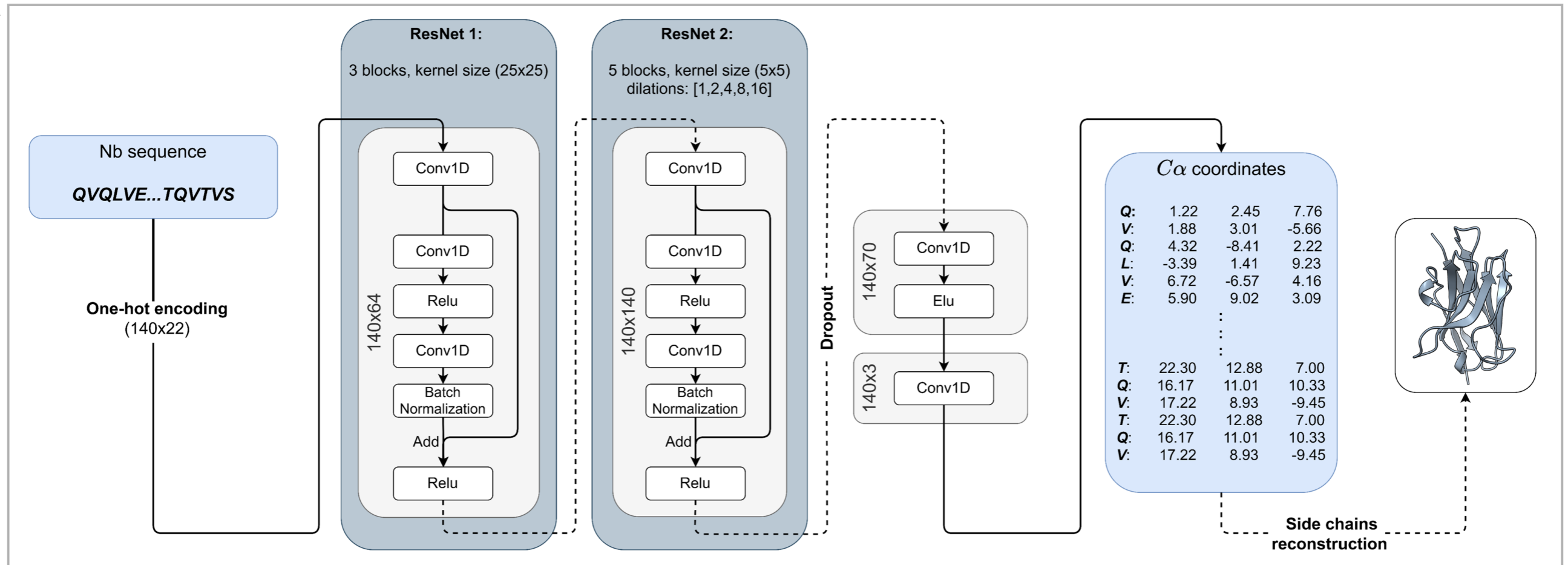


Tomer Cohen

- Invariance to rotations and translations can be achieved by frame region alignment
- All the structures of the training set were aligned on a randomly selected reference structure

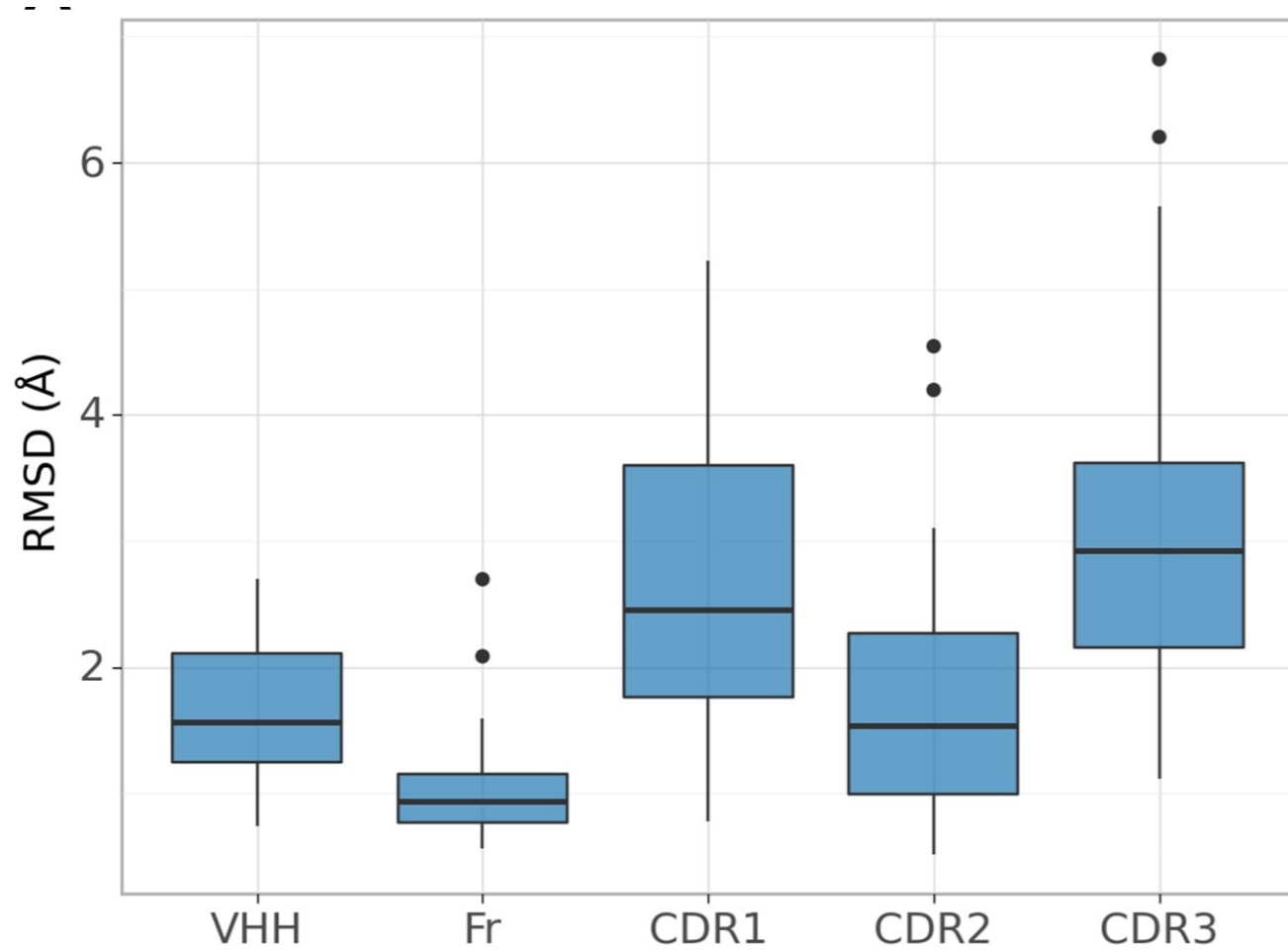


NanoNet architecture



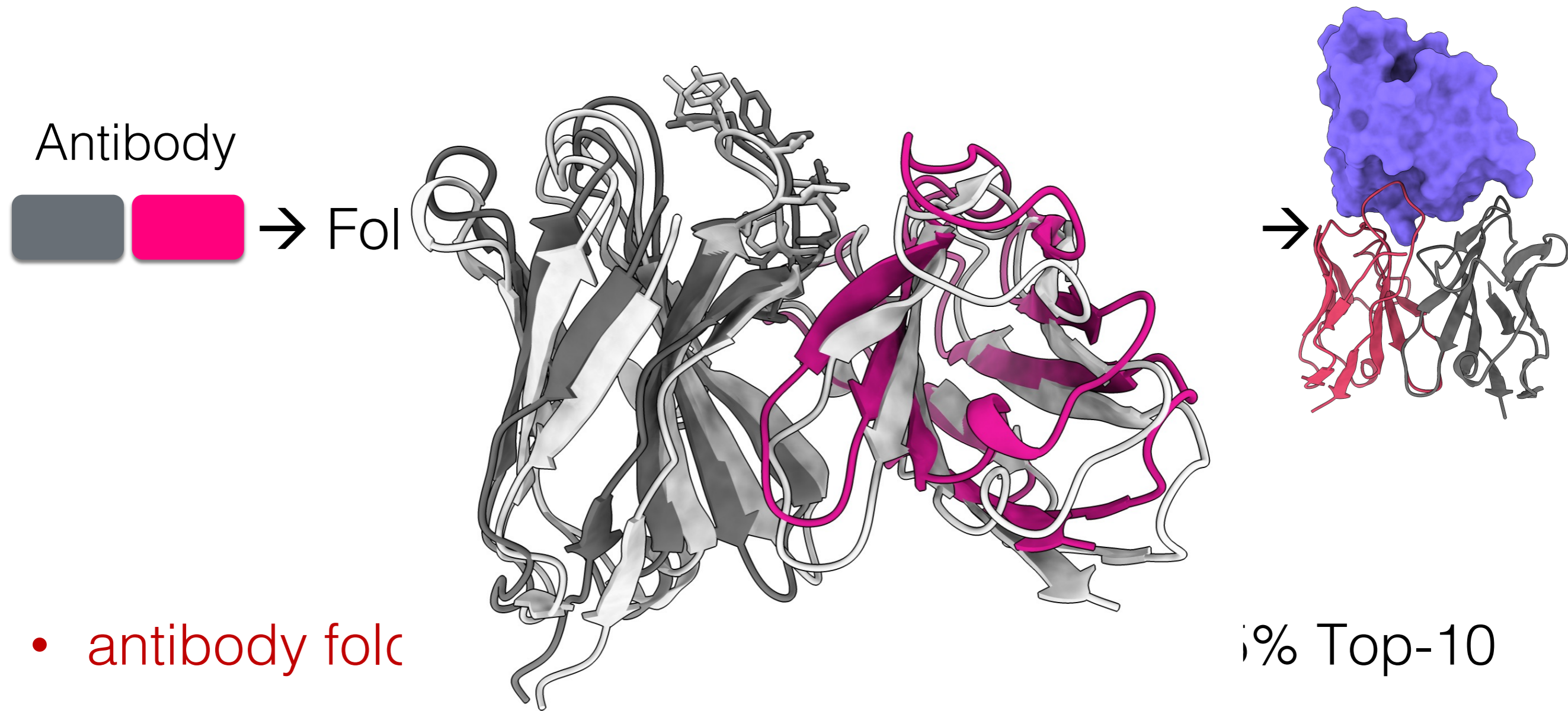
- Trained on ~1,800 antibody and nanobody structures
- Coordinates MSE as a loss function
- Structure prediction: ~6ms on GPU or ~20ms on a CPU
- \Rightarrow 1M structures in less than an hour on a CPU!
- <https://github.com/dina-lab3D/tutorials/tree/main/NanoNet>

NanoNet performance for nanobodies



- Accuracy comparable to AlphaFold2, IgFold, ABLooper...

Antibody folding and docking to antigen



- antibody fold success rate
 - the main problem is the accuracy of the antibody models
- % Top-10

Can we fold & dock simultaneously?

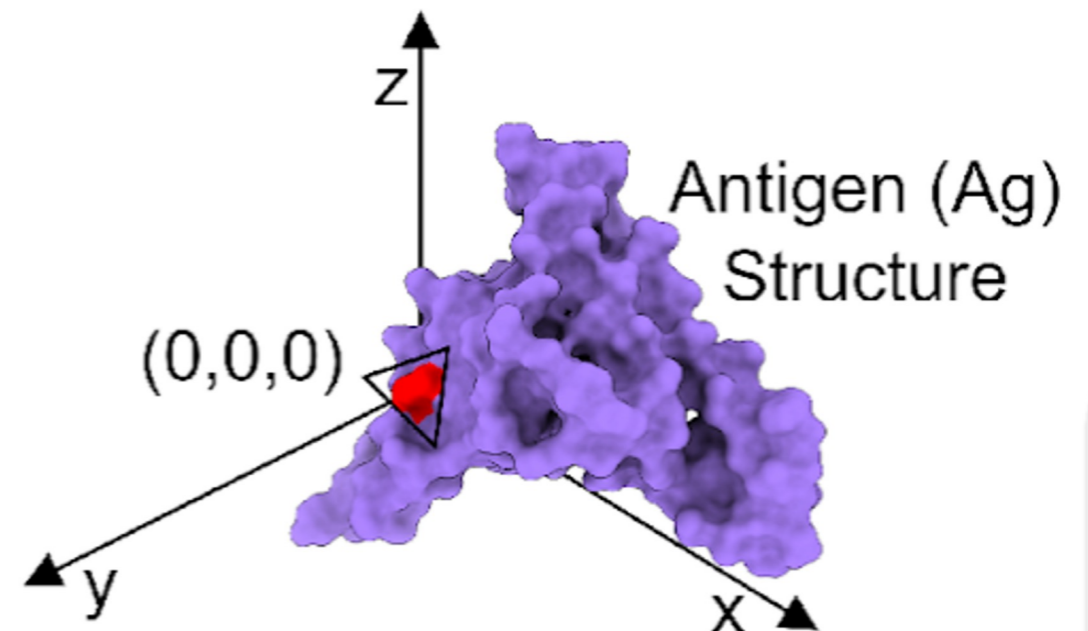
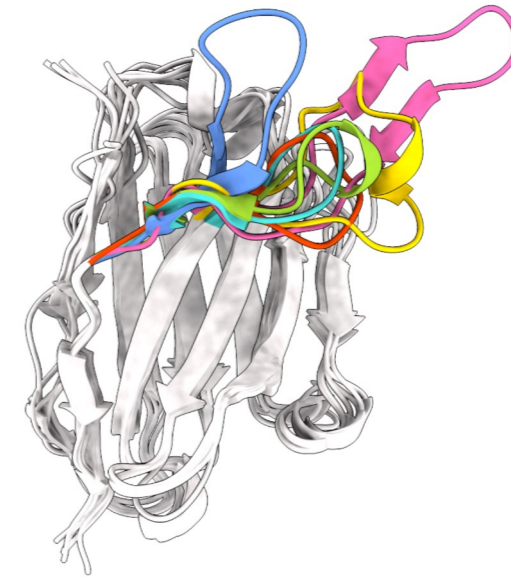


Input: antibody sequence + antigen structure

Output: complex structure

Transformational invariance

- **Antibody** – aligning the training set structures on a single representative structure for the heavy and the light chains
- **Antigen** – constructing an amino acid reference frame for the antigen (N-CA-C atoms) and transforming it to the global reference frame.

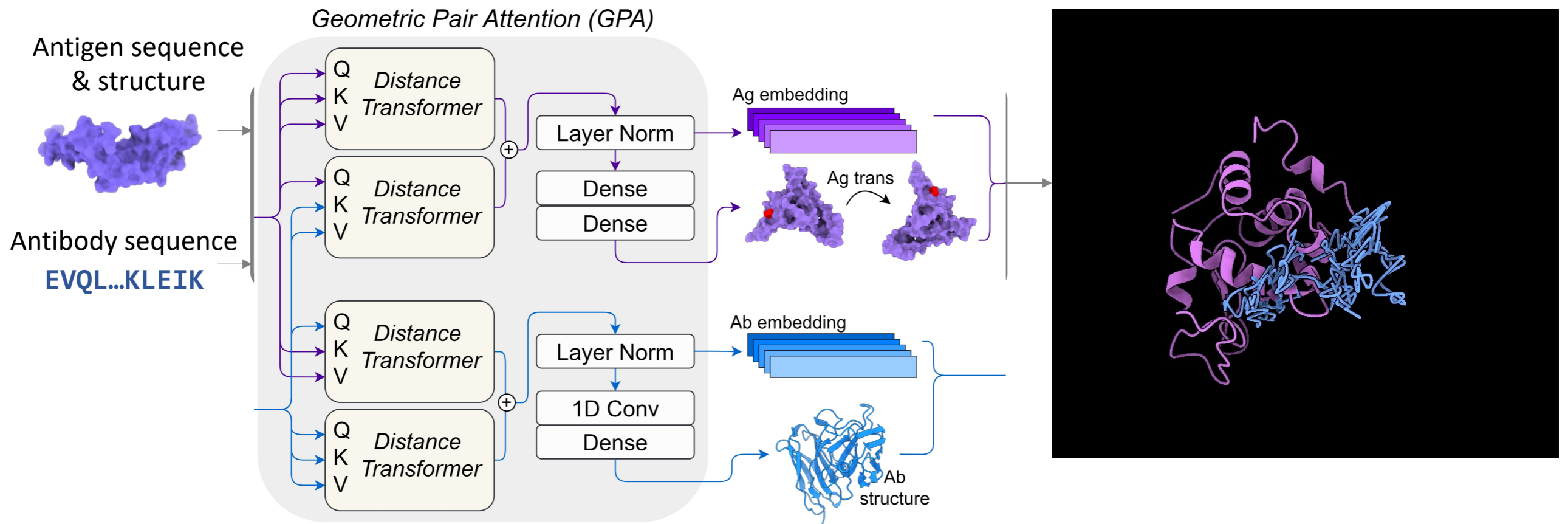


Fold & dock architecture

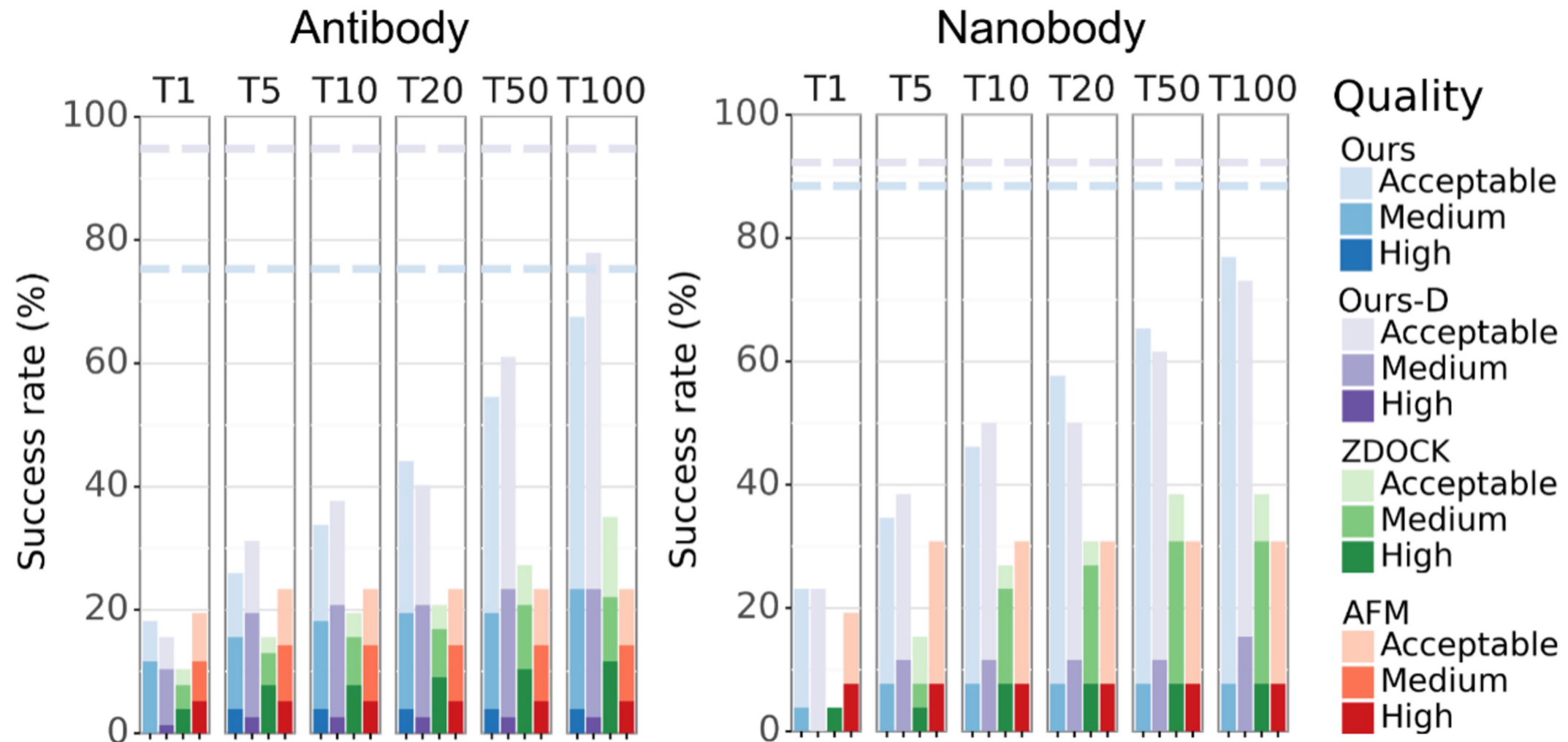
Designed to simulate the **biological antibody-antigen recognition**

Consists of several layers of **Geometric Pair Attention (GPA)** each containing **four** dedicated **Distance Transformer** modules

Each of the **four Distance Transformers** is responsible for a different aspect of the antibody-antigen interaction



Fold & Dock accuracy

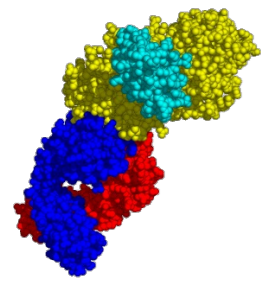


Success rate: fraction of test set complexes with Acceptable or higher quality models, usually specified for topN predictions

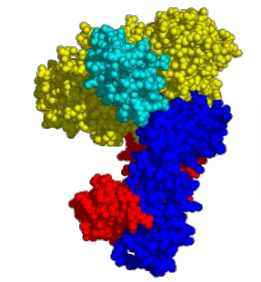
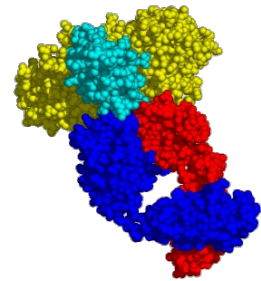
Quality: High, Medium, Acceptable, and Incorrect

Docking with SAXS profile of the complex

foXS Dock

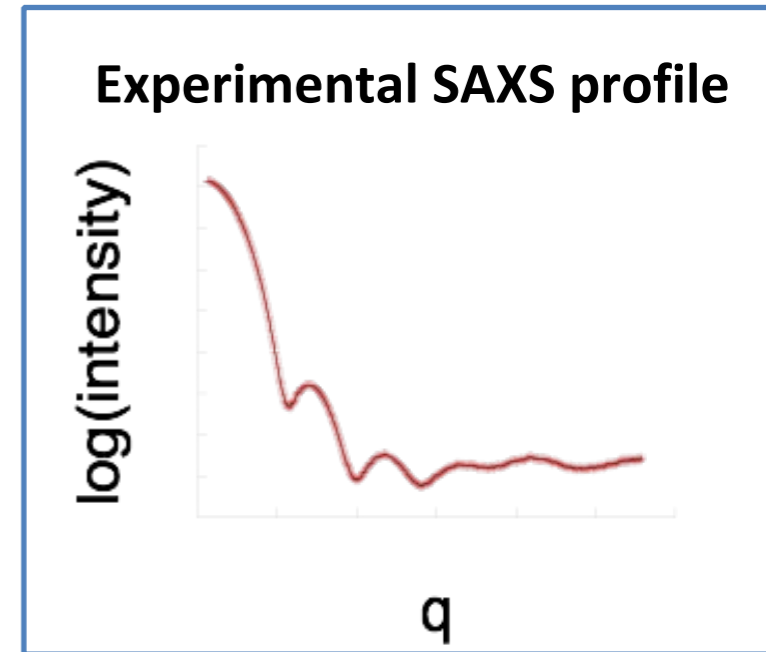
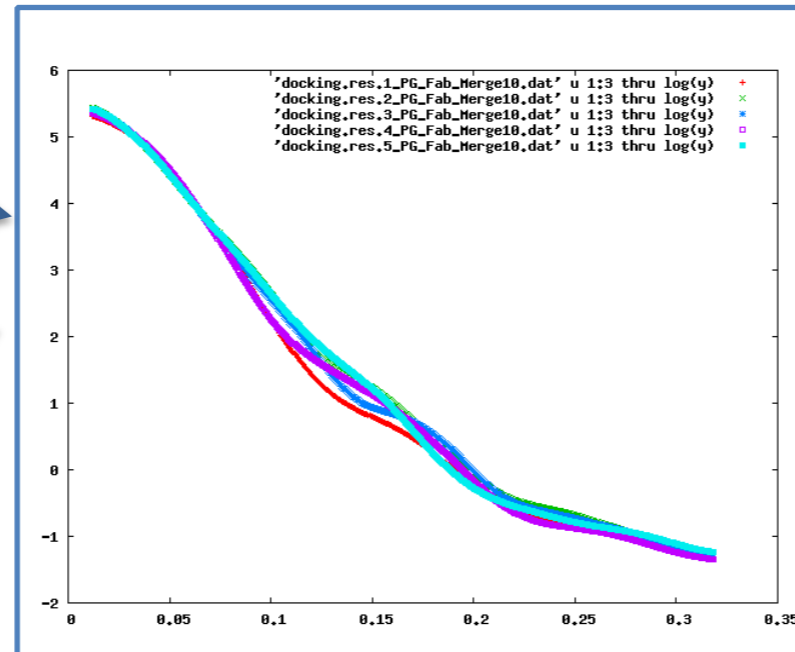


forward model



PATCHDOCK

FoldDock



foXS

Generate docking candidates

Compute theoretical SAXS profiles

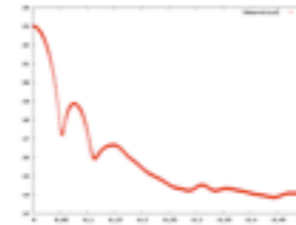
Fit experimental profile and compute the score

Debye formula

$$I(q) = \sum_{i=1}^N \sum_{j=1}^N f_i(q) f_j(q) \frac{\sin(qd_{ij})}{qd_{ij}}$$

$$\chi = \sqrt{\frac{1}{M-1} \sum_{i=1}^M \left(\frac{I_{exp}(q_i) - cI(q_i)}{\sigma(q_i)} \right)^2}$$

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:

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

or upload file:

Complex SAXS profile:

e-mail address:

(the results are sent to this address)

Complex Type:

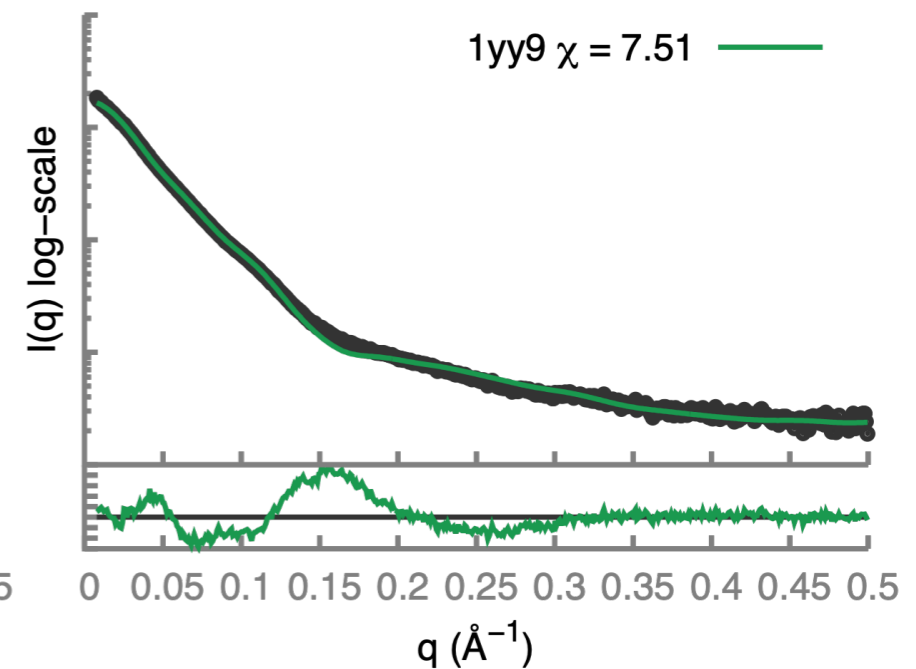
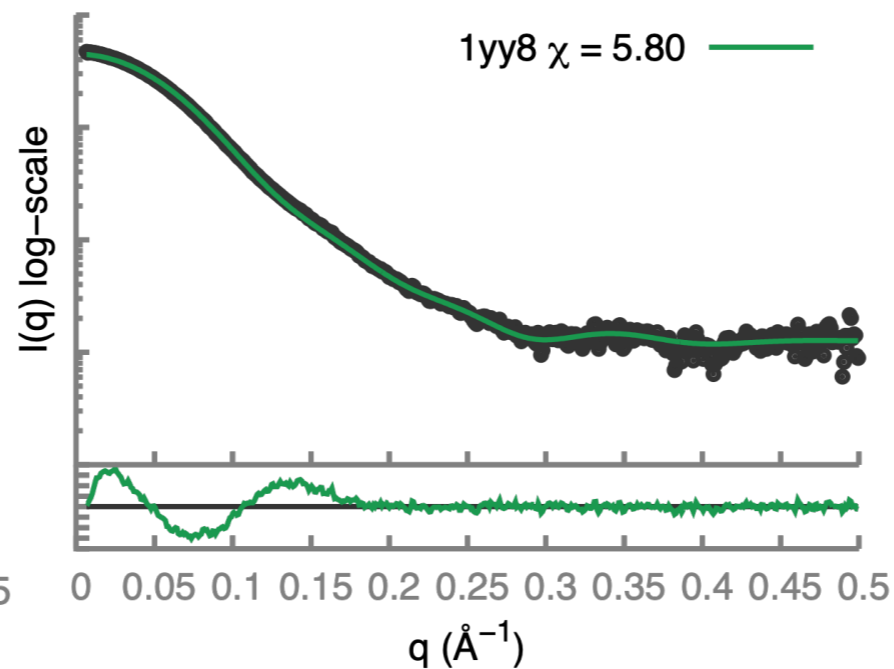
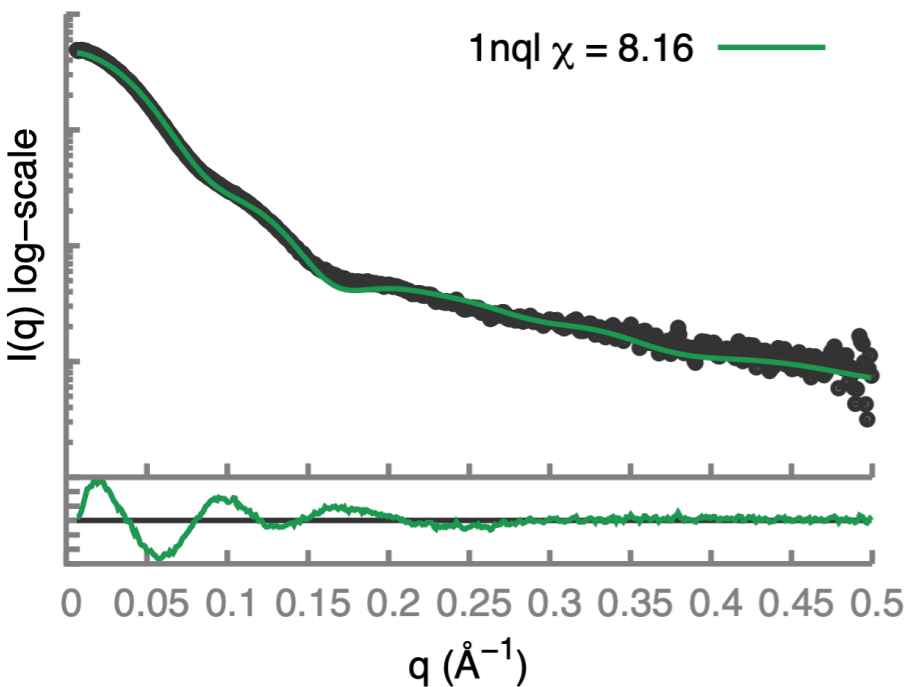
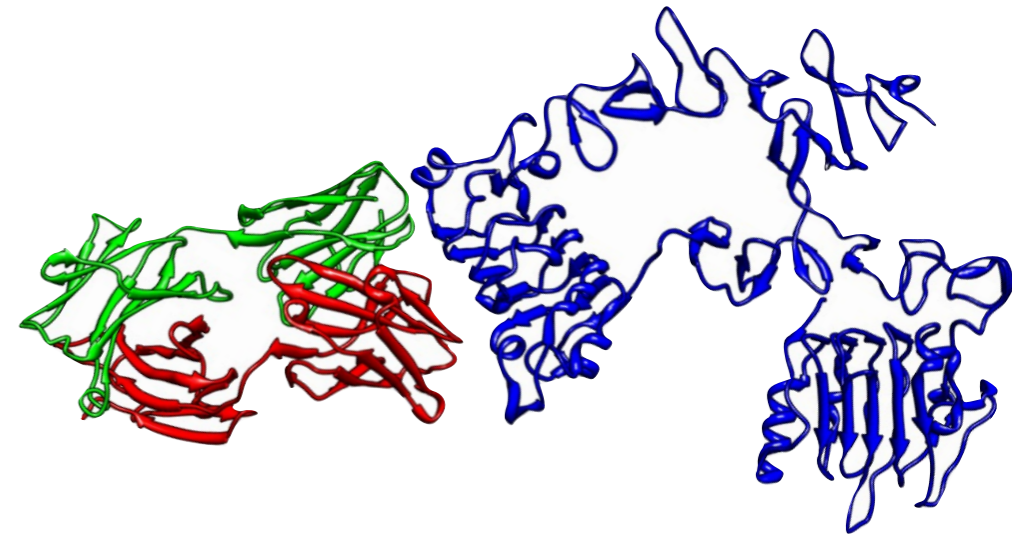
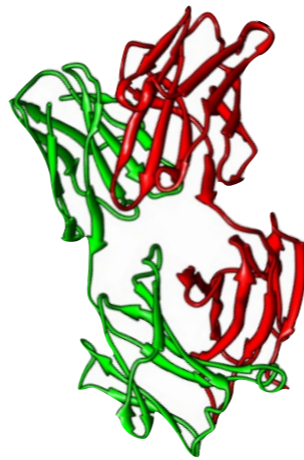
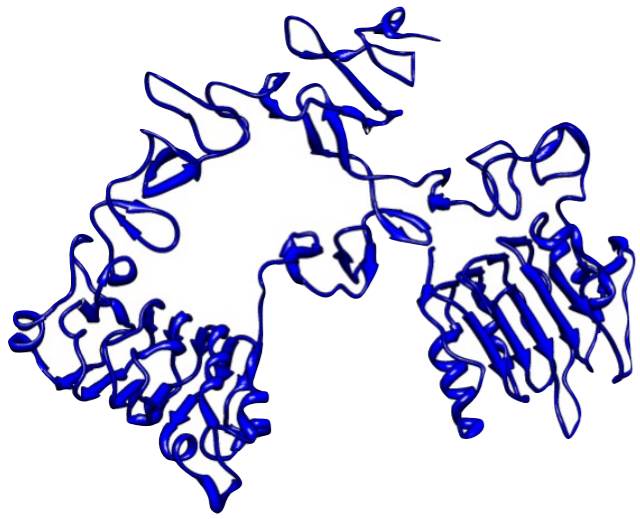
Be sure to give receptor and ligand in the corresponding order!

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

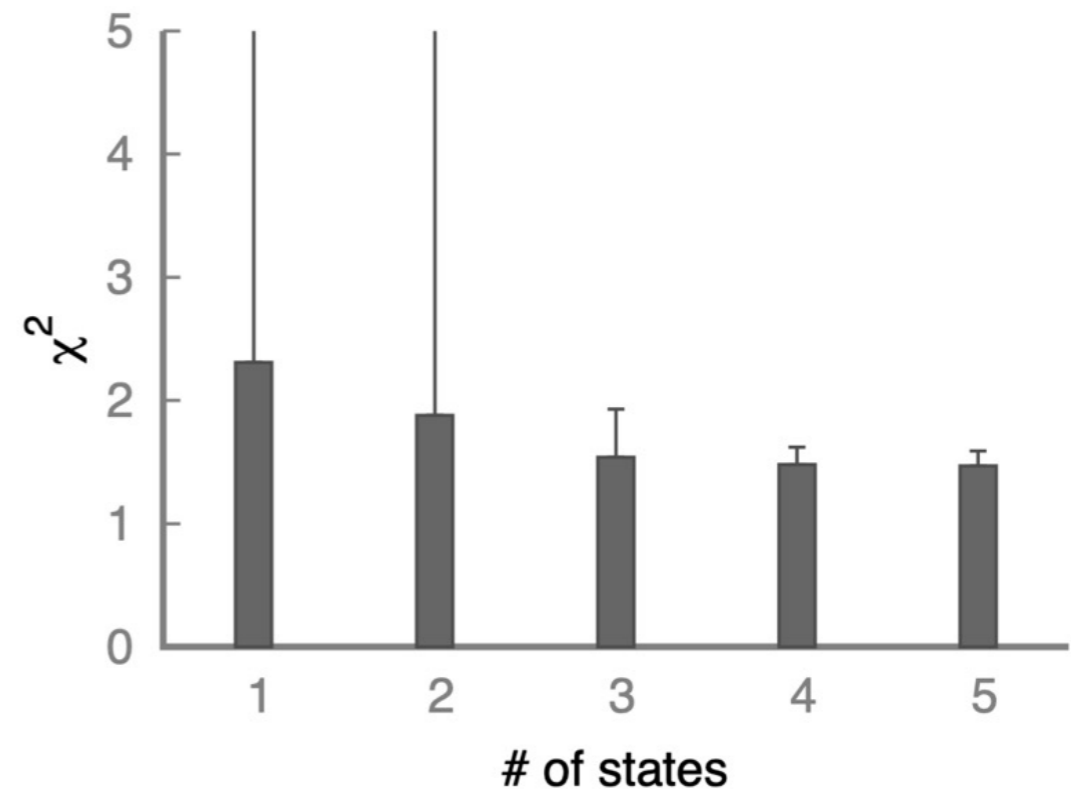
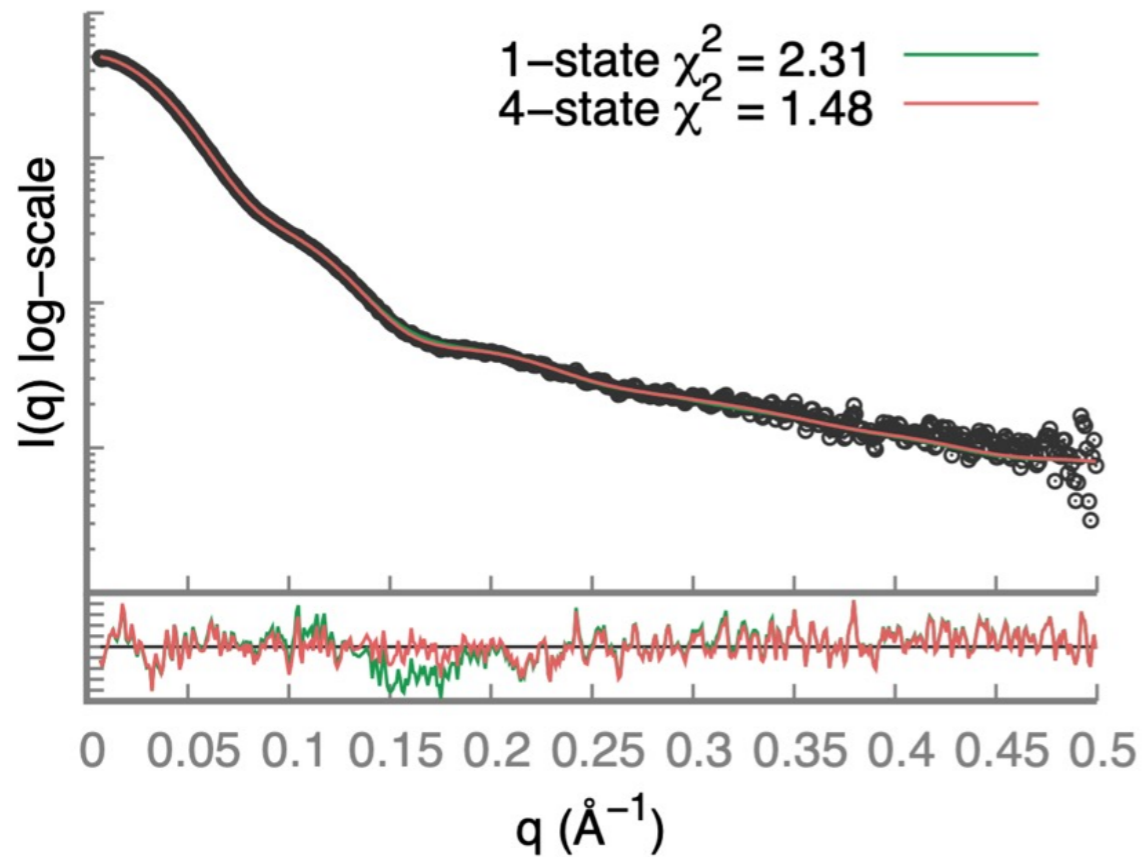
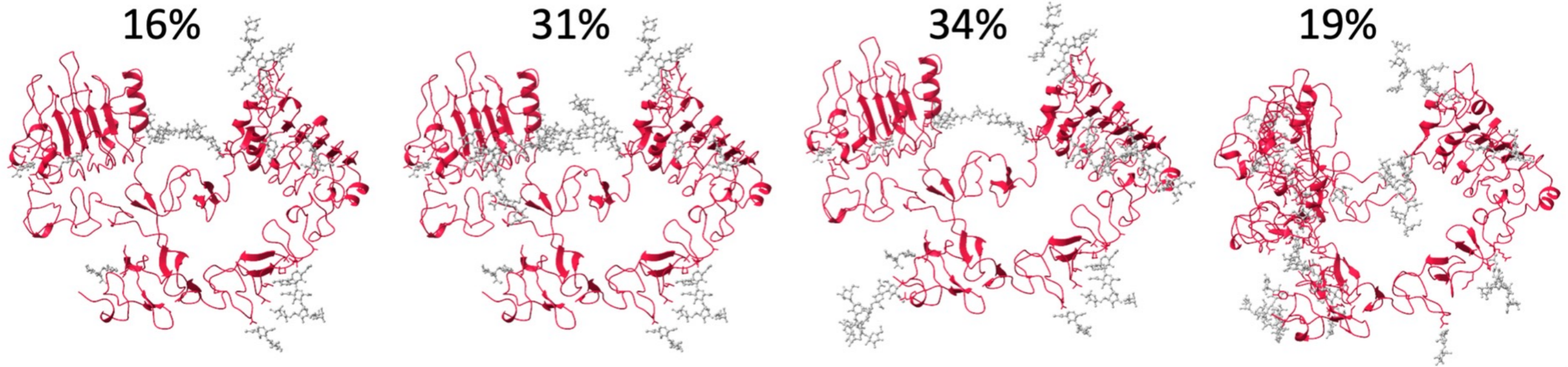
Contact: dina@salilab.org

EGFR-antibody complex with SAXS profiles

- SAXS profiles collected for EGFR, antibody, and their complex
- 4 antibodies

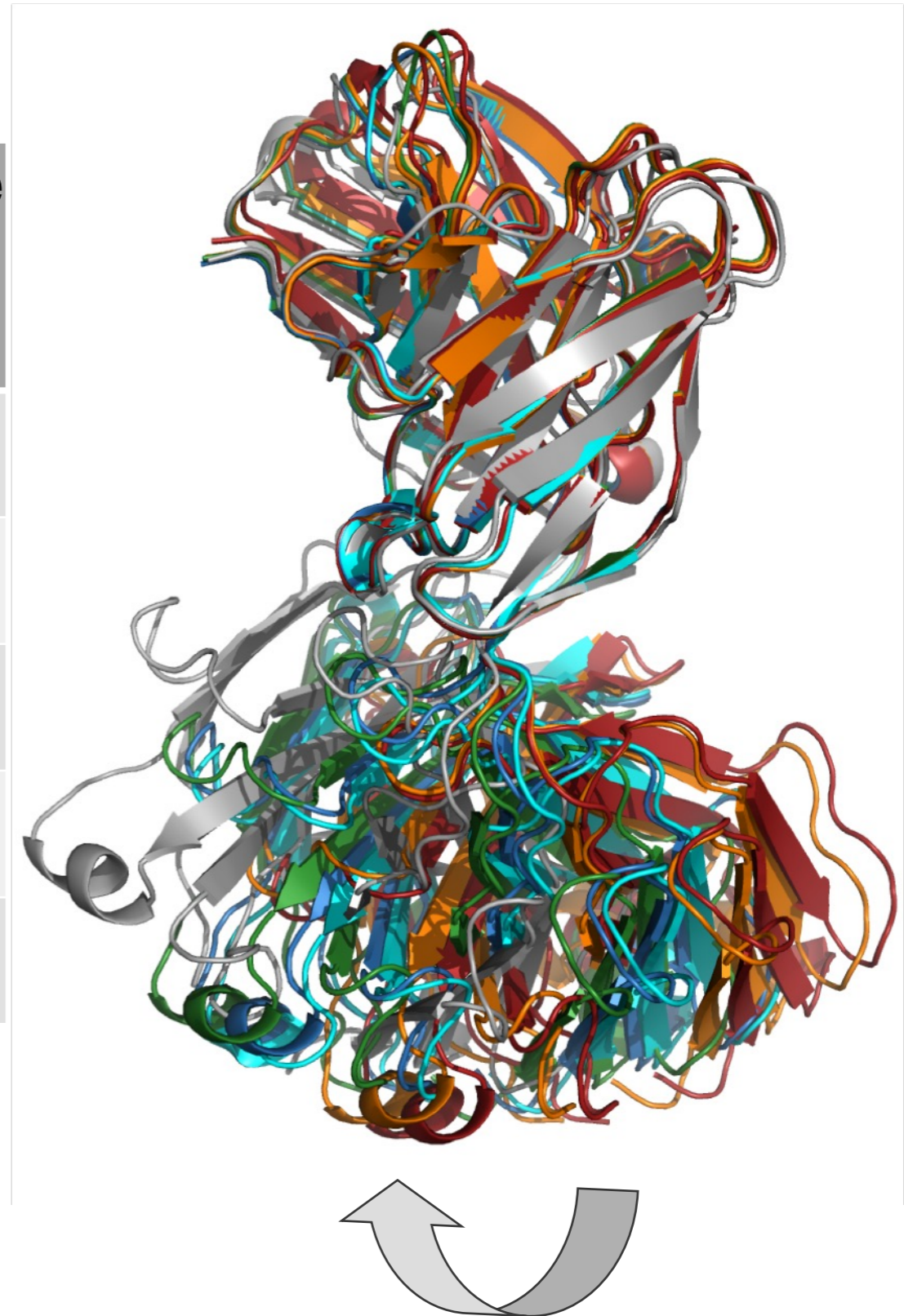


EGFR is flexible and glycosylated



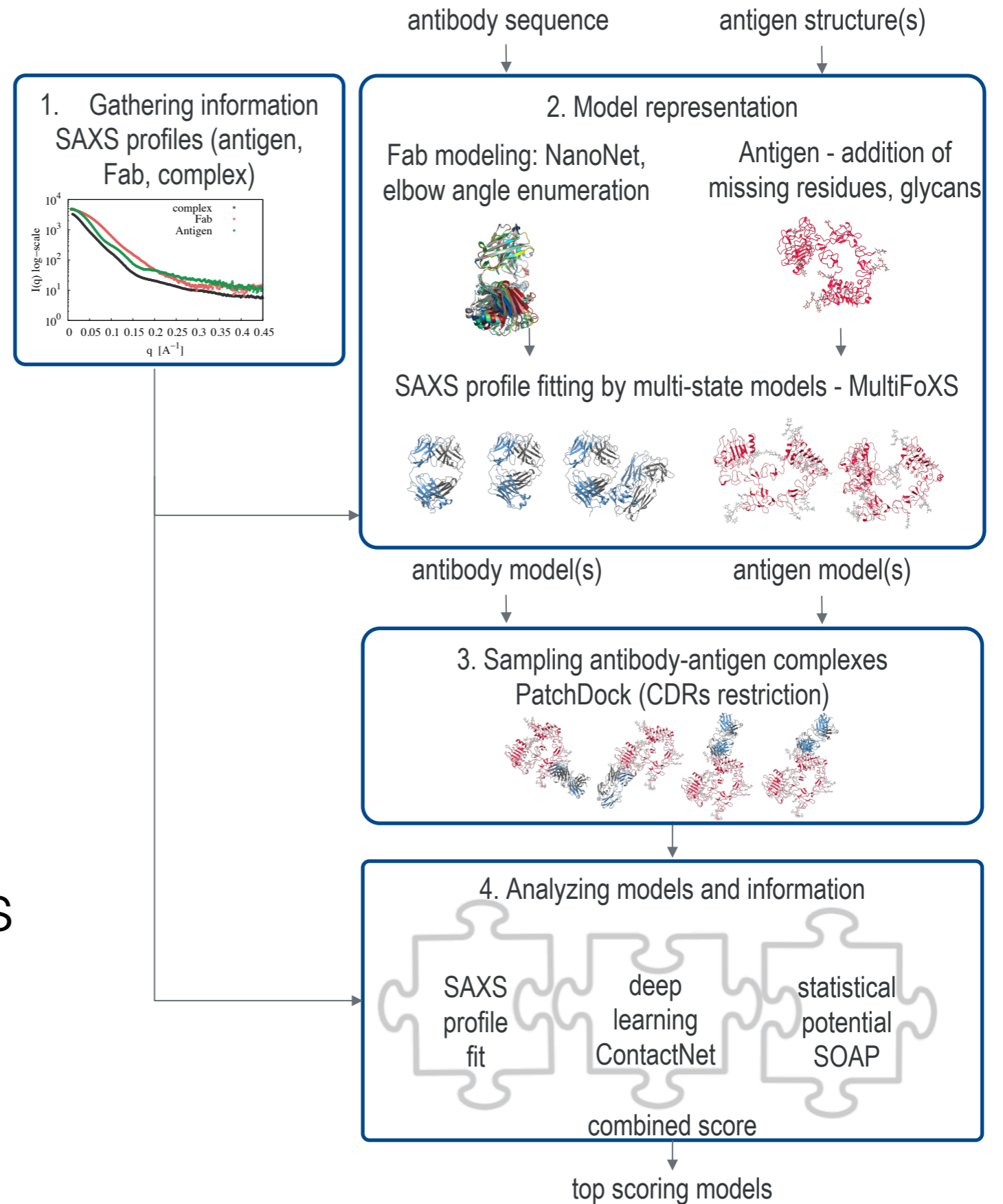
Fabs vary their elbow angle

PDB	χ^2 x-ray Fab	χ^2 single-state	χ^2 multiple elbow angles
1yy9	27.3	9.6	9.6
3b2u	20.4	11.3	11.3
3c09	19.0	4.9	4.9
3p0y	9.4	3.5	3.4
3sqo	69.0	6.5	3.6

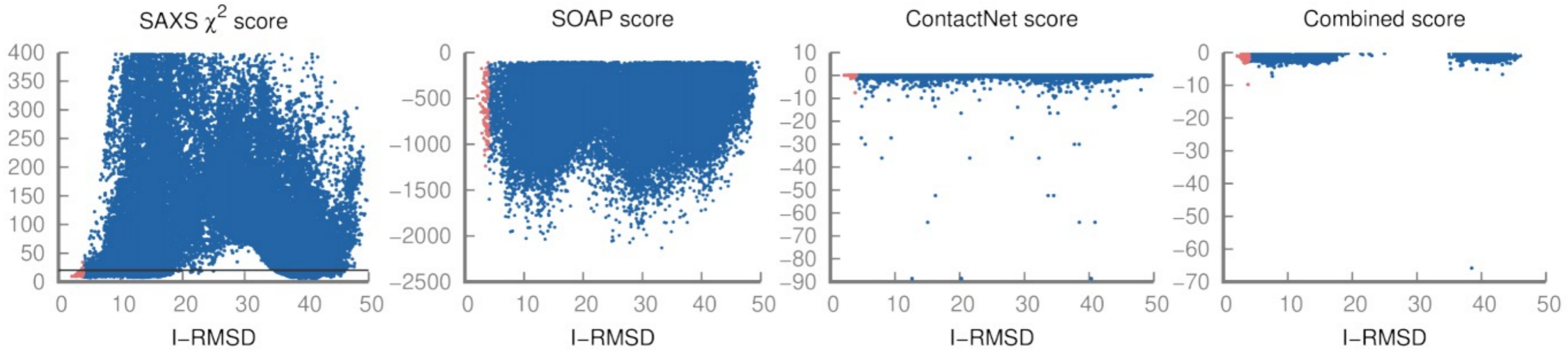


Let's dock!

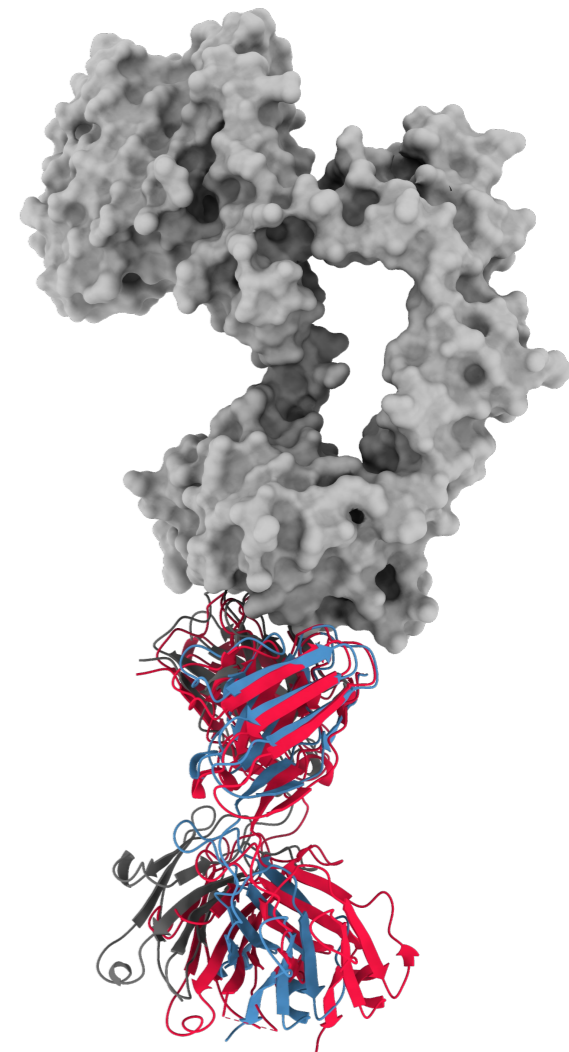
1. SAXS data collection
 - antibody, antigen, complex
2. Antibody and antigen modeling
 - single- or multi-state
3. Docking with all conformations
4. Scoring
 - SAXS – multiple states
 - interaction interface



Docking Results

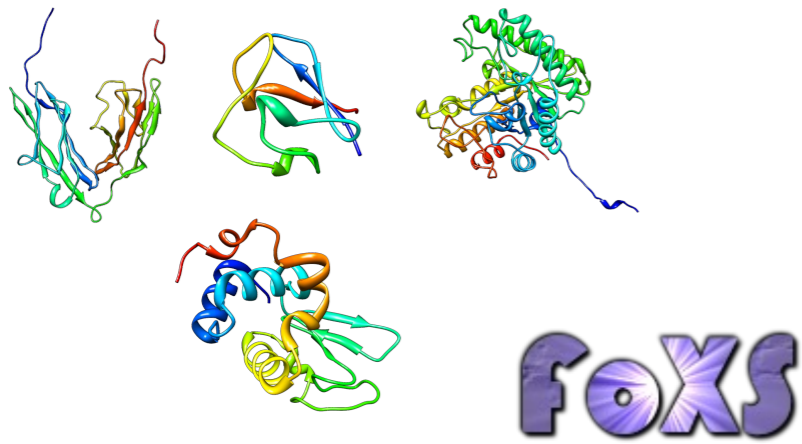


PDB	Rank by SAXS χ^2	Rank by SOAP	Rank (IRMSD) by ContactNet	Rank (IRMSD) by combined score
1yy9	1	30	1	1
3b2u	1	2	13	4
3c09	776	192	2122	178
3p0y	139	1228	45	2
3sqo	1022	38	1	1

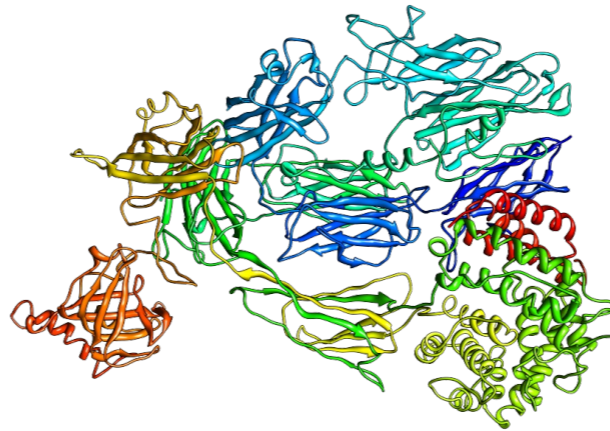


What are the modeling tasks to address with SAXS?

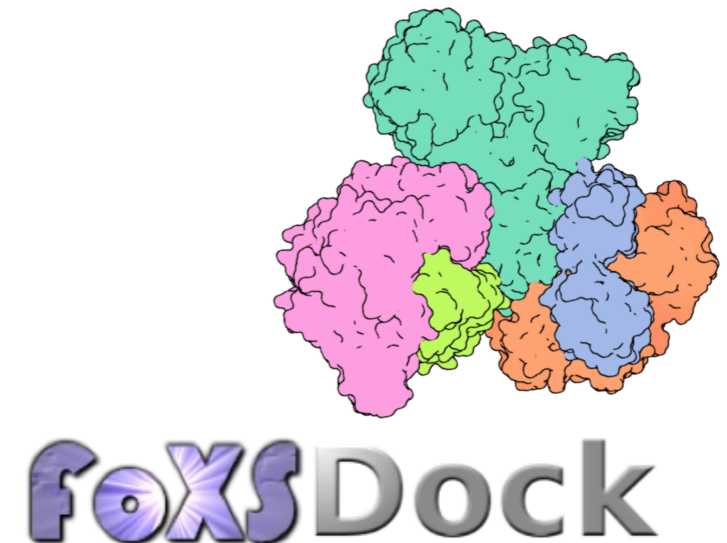
validation of protein structure prediction



assembly of **multi-domain** proteins

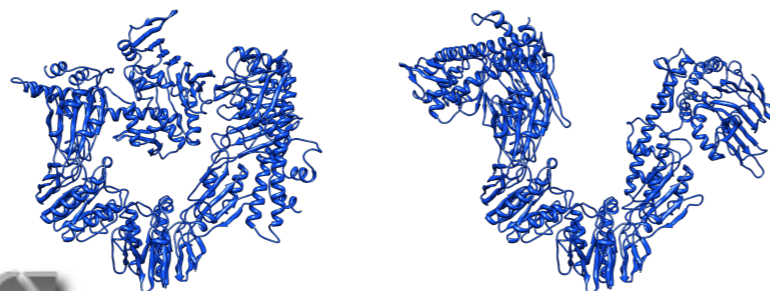


assembly of **multi-protein** complexes



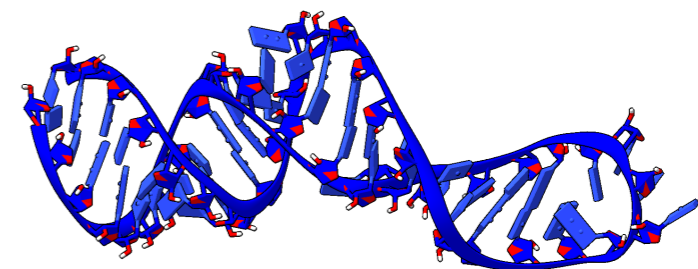
structural characterization of **protein dynamics**

BilboMD



MultiFoXS

RNA



SCOPER

Links

SAXS profile calculator

<https://modbase.compbio.ucsf.edu/foxs/>

foXS

Multi-state modeling

<https://modbase.compbio.ucsf.edu/multifoxs/>

MultiFoXS

BilboMD – MD + MultiFoXS

<https://bilbomd.bl1231.als.lbl.gov>

BilboMD

SCOPER

<https://bilbomd.bl1231.als.lbl.gov>

SCOPER

SAXS-based docking

<https://modbase.compbio.ucsf.edu/foxsdock/>

foXS Dock

Antibody-antigen structure prediction

<https://folddock.cs.huji.ac.il/>

Fold&Dock

AlphaFold-based assembly of complexes

<https://github.com/dina-lab3D/CombFold>

 **CombFold**