

Atomic modeling of proteins via Small Angle X-ray Scattering (SAXS)

Susan Tsutakawa

**SIBYLS Beamline. Advanced Light Source
Synchrotron, Lawrence Berkeley Natl Lab**



Six Take Home Messages on What can SAXS do for you?

*X-ray Scattering by electrons provides **distances** between electrons.*

Small Angle X-ray Scattering measures all electron pair distances in a protein in solution

Atomic models can be quantitatively compared with SAXS data

SAXS can validate protein structure predictions

SAXS can reveal protein conformations occurring in solution at the atomic level

***Atomic models** are more powerful than shape because they can be tested.*

If I could ask for any scientific app, what would it be?

Accurate and reliable protein structure prediction



For proteins with no known orthologs, structure predictions currently are not reliable.



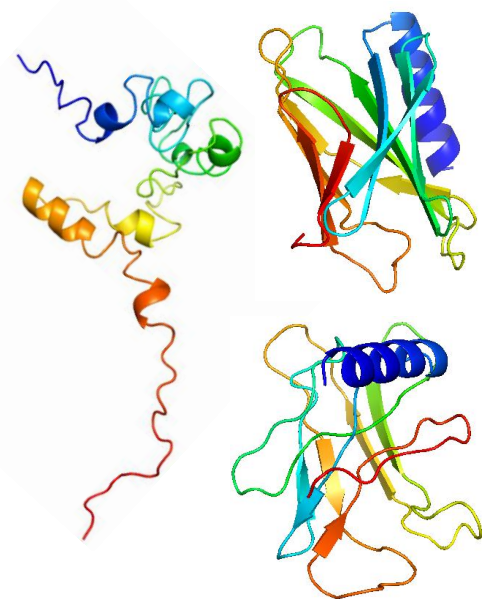
Amino Acid Sequence

>4ZBH
MGSSHHHHHSQDPNSNQAQ
ELNHELELEQLETKITVSSV
SLTGSTLNVVLENGSTNLY
DFQGFSVIVQYYANISNIST
FNLSLYNYTKNSNPSPYYWT
INTPLLAPGSQATLTIILPY
PPYPNTQATVVIVTNYGPSV
IWRGSL

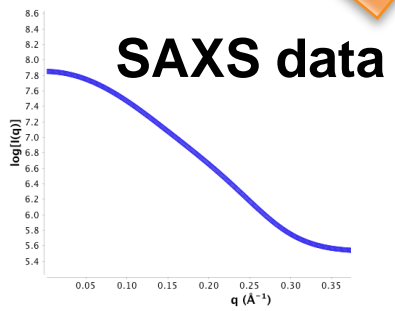
Protein Prediction Server



Structure Prediction



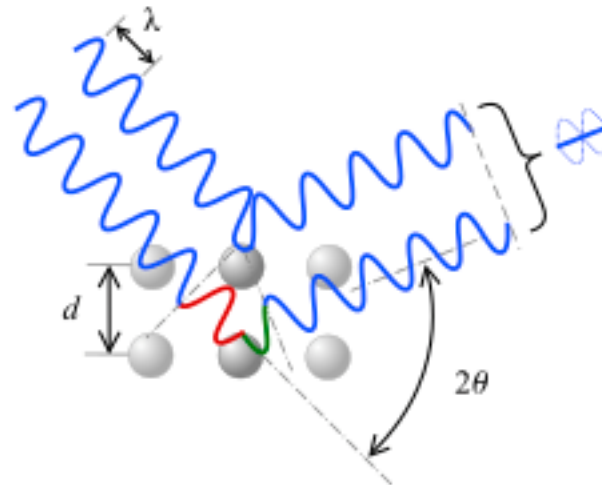
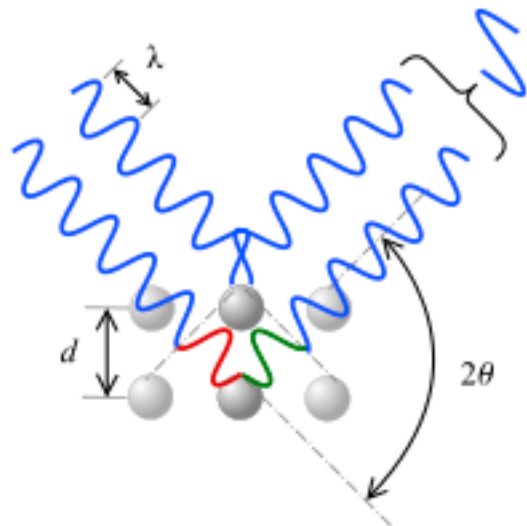
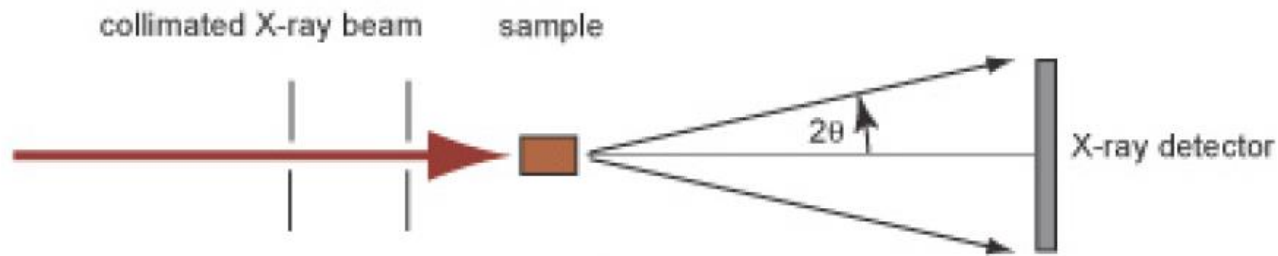
**SIBYLS
Beamline
12.3.1**



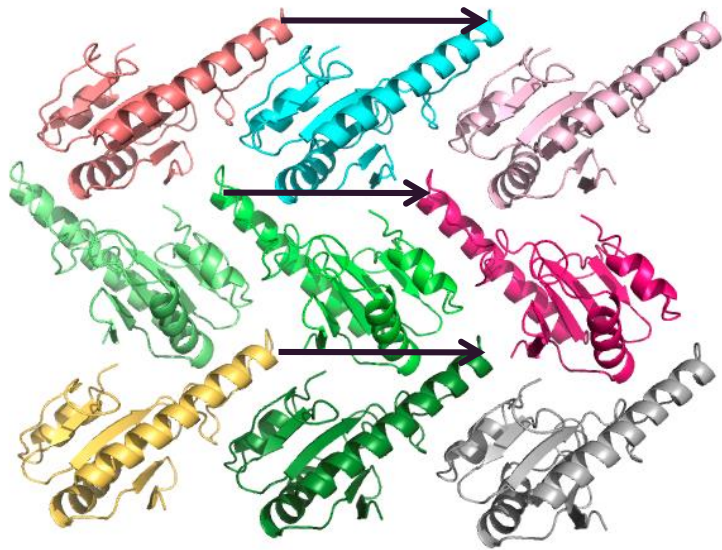
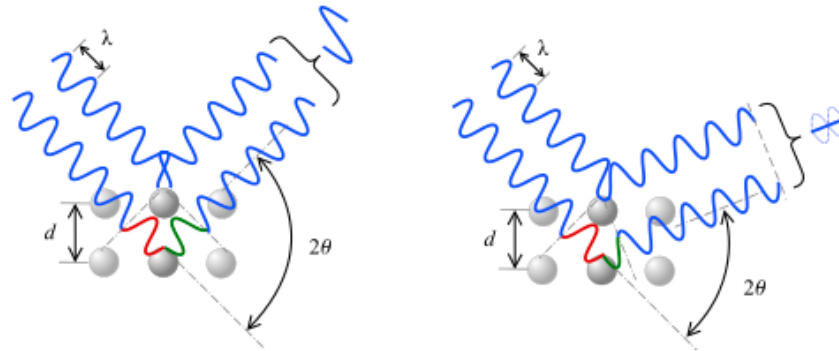
I propose that input of SAXS data can improve protein structure algorithms.

As in crystallography, SAXS uses elastic scattering of X-rays, where the X-rays are scattered by an electron without a change in energy. Scattered X-rays constructively or destructively combine with

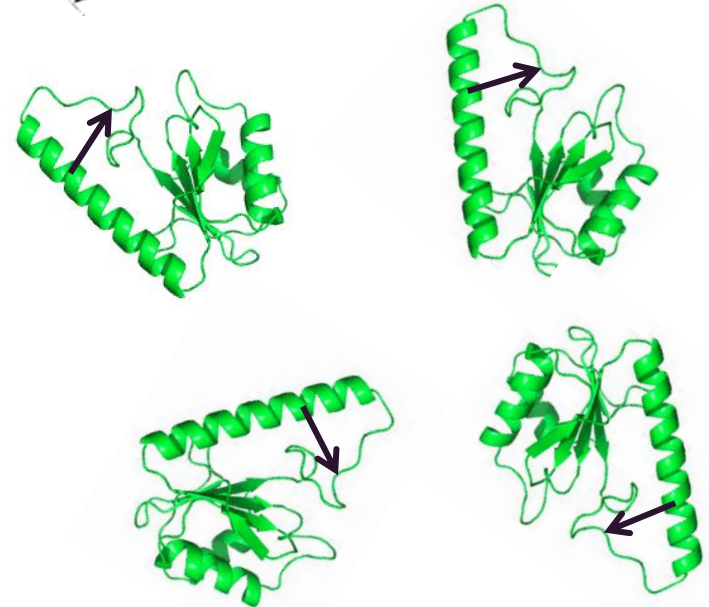
A



The X-ray scattering provides information on the distance between electrons.



In Crystallography, these electrons are related to each in crystallographic symmetry.

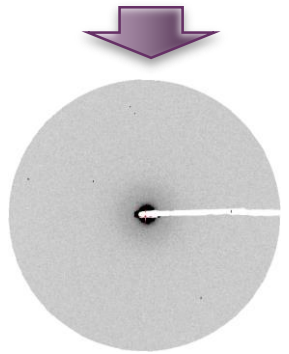


In SAXS, the intramolecular distances are constant; the scattering is coherent, and the amplitudes are added.

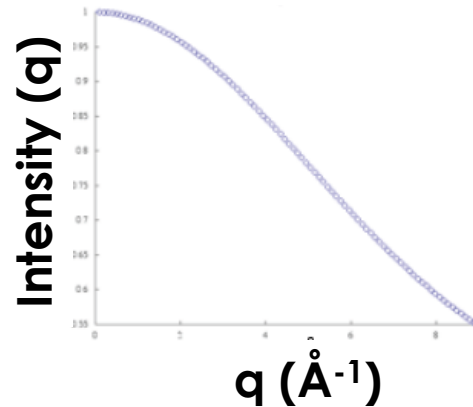
SAXS is a distance method, measuring all electron pair distances.



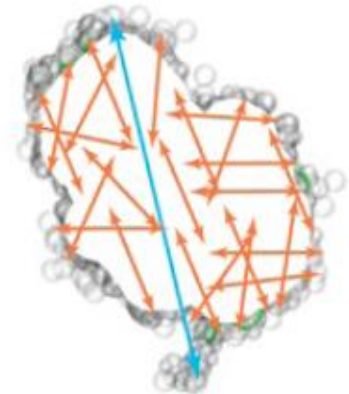
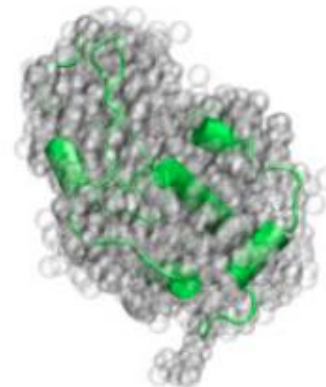
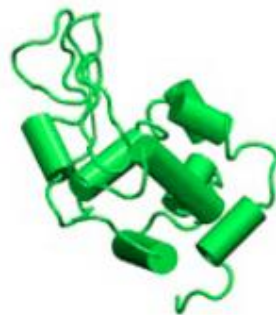
$$I(q) = 4\pi \int_0^{D_{\max}} P(r) \frac{\sin(qr)}{qr} dr$$



Scattering Curve

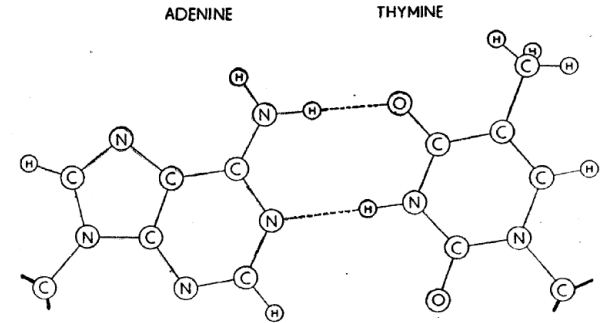
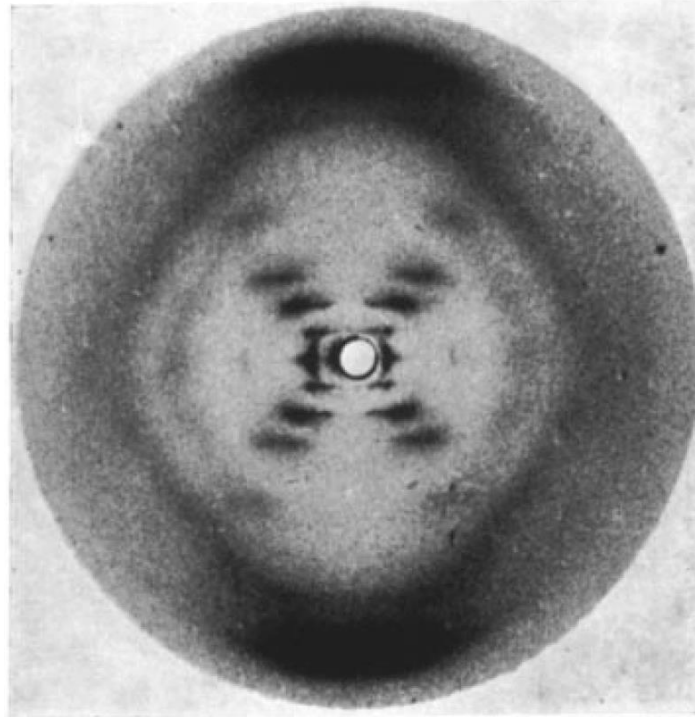
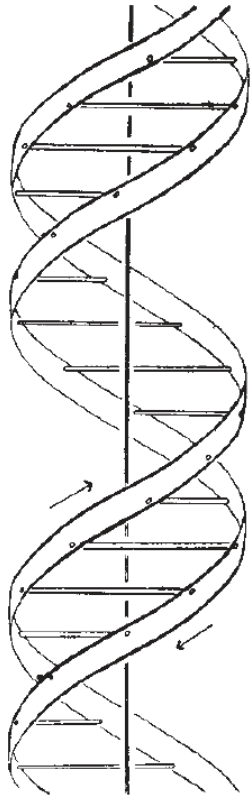


SAXS sample 30 ul
Protein 1-3 mg/ml
Exact Buffer



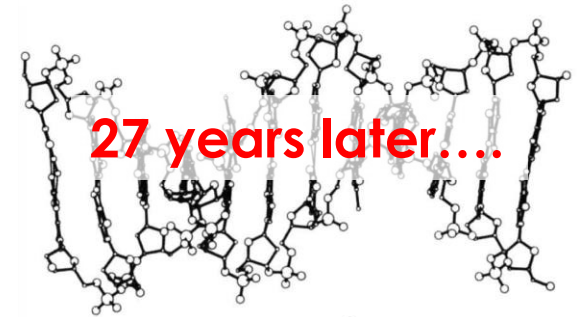
Distance information can validate an atomic model – as exemplified by validation of the model of DNA by fiber diffraction

May, 1953



Genetic Implications of the structure of Deoxyribonucleic Acid

Watson J.D. and Crick F.H.C.
Nature **171**, 964-967 (1953)



27 years later....

April, 1953

A Structure for Deoxyribose Nucleic Acid

Watson J.D. and Crick F.H.C.
Nature **171**, 737-738 (1953)

Molecular Configuration in Sodium Thymonucleate

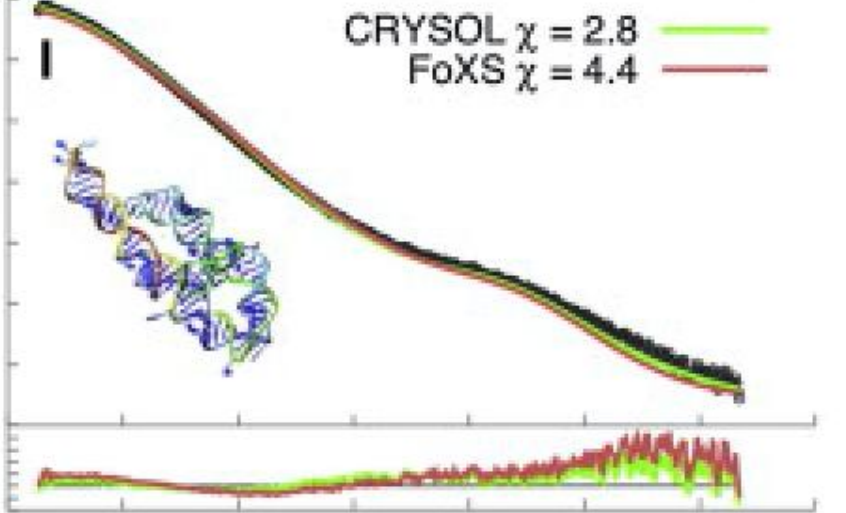
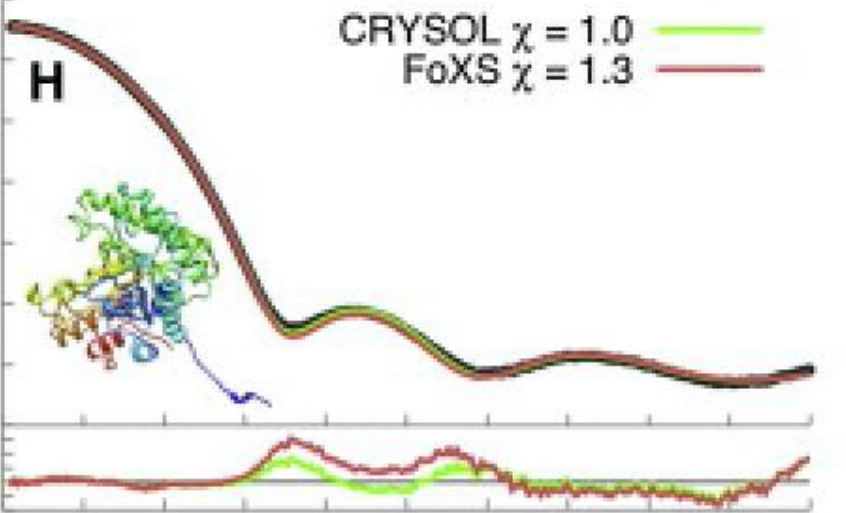
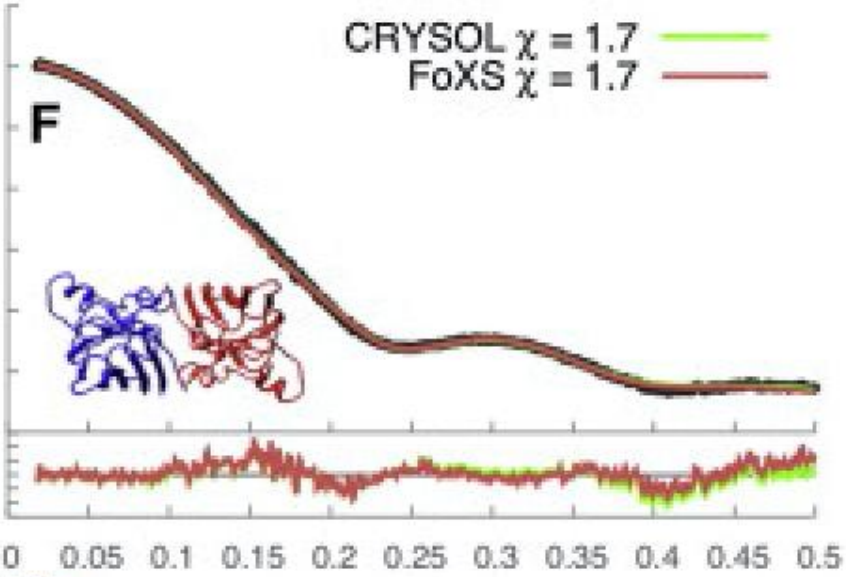
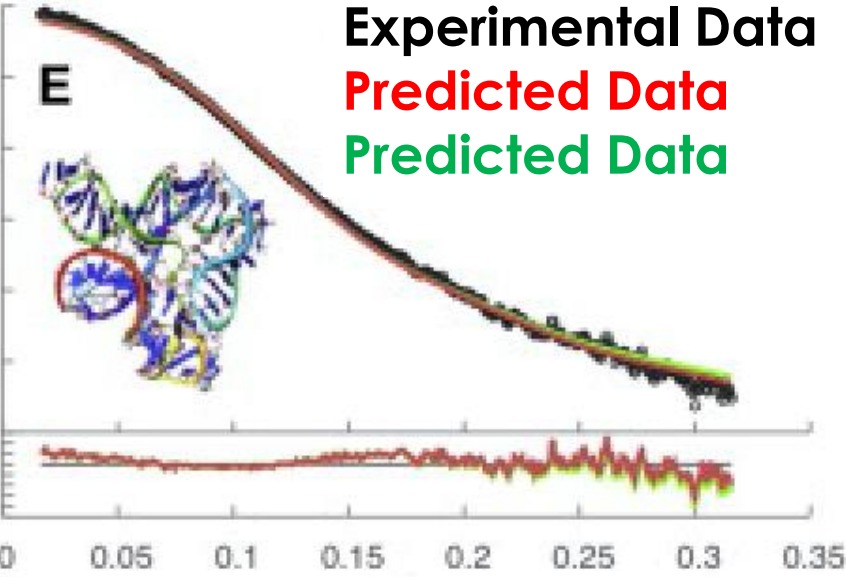
Franklin R. and Gosling R.G.
Nature **171**, 740-741 (1953)

Crystal structure analysis of a complete turn of B-DNA

Richard Wing*, Horace Drew, Tsunehiro Takano, Chris Broka, Shoji Tanaka, Keiichi Itakura† & Richard E. Dickerson
Nature **287**, 755-758 (1990).

BREAKTHROUGH for proteins:

Calculations of SAXS curves from atomic models matches the experimental data for well folded targets.





Can you use the ability to compare atomic structures to the target SAXS data for actual protein structure predictions?

Test if SAXS can differentiate models –
input an AA sequence to three easily
accessible servers.

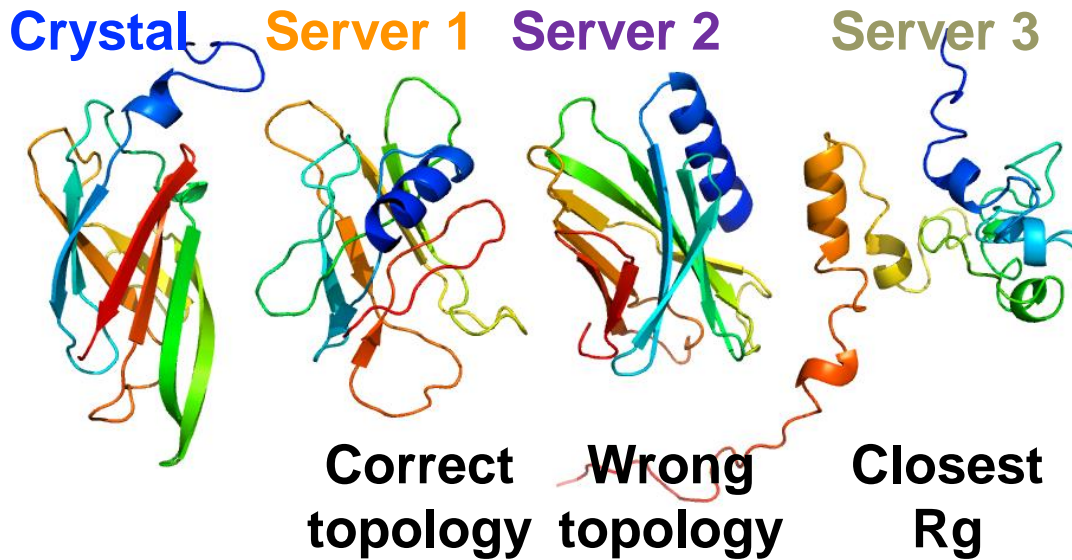
Crystal



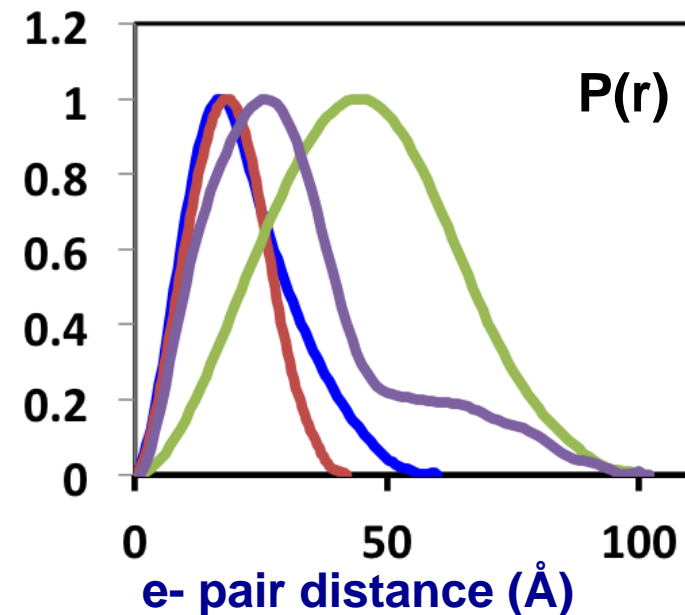
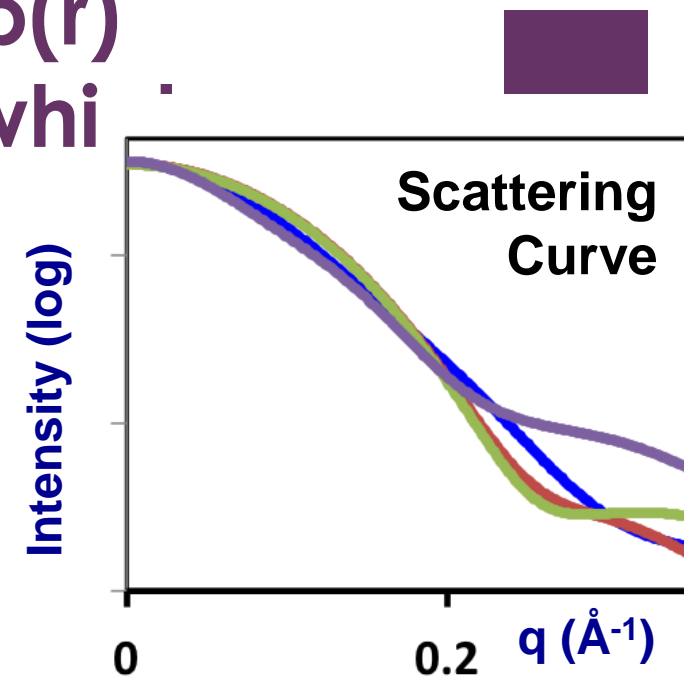
Chi-Lin Tsai

*None of the models were exact fits, but if you didn't have a
crystal structure to compare, how would you know?*

The Raw scattering curve or $p(r)$ function can help distinguish which model is better.

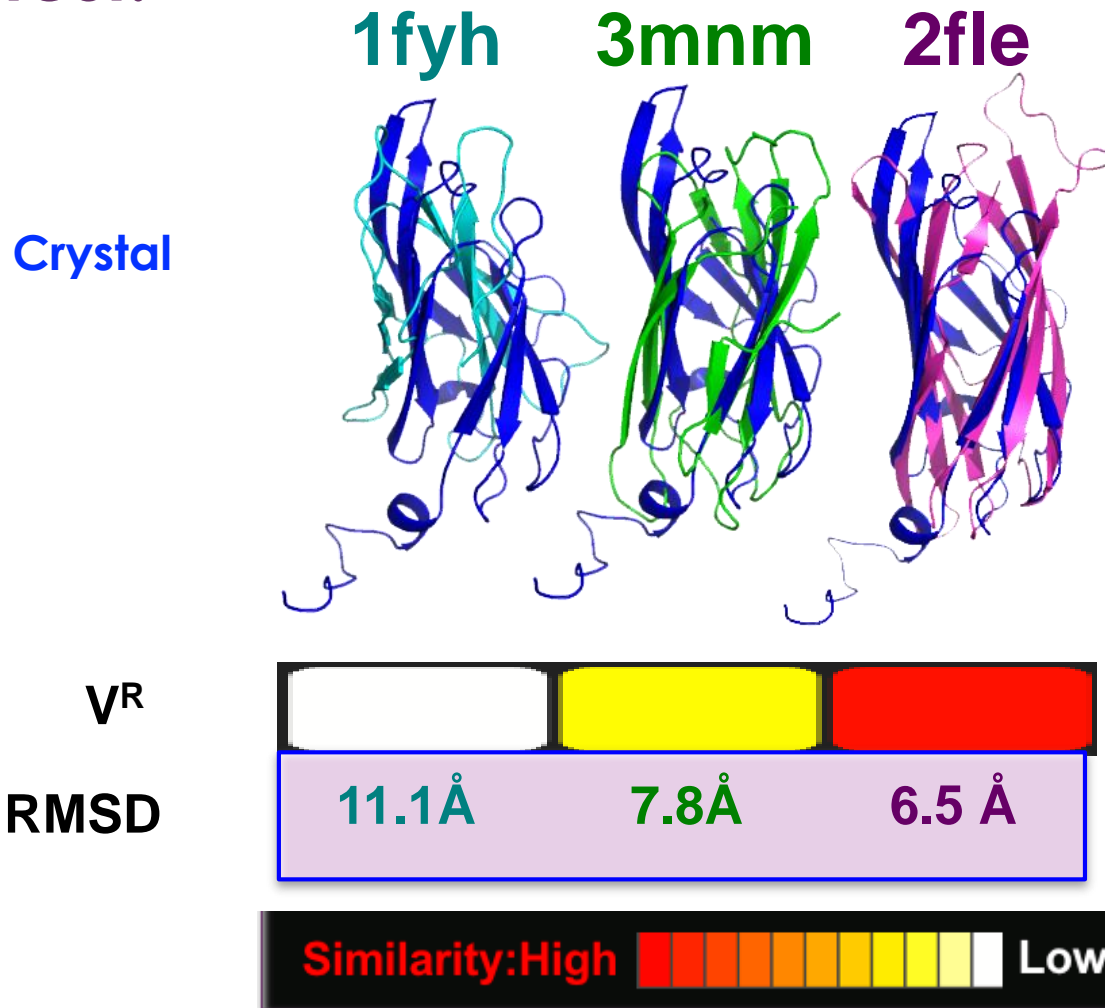


Hura et al, Nat. Methods, 2013



Can SAXS differentiate similarly-related folds?

The Dali test:



SAXS has the resolution to differentiate between different DALI models.

Premise: SAXS can validate or even promote accurate protein structure prediction.

Test: We tested this hypothesis with CASP

Parameters: Predictors were given

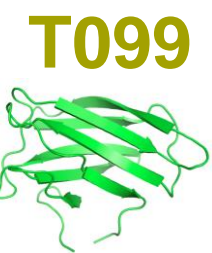
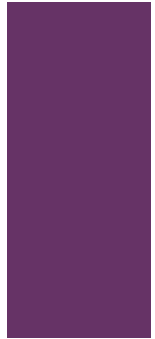
- 1) the AA sequence of a solved and not yet released crystal structure
- 2) the corresponding SAXS data (including shape, stoichiometry)
- 3) Two weeks

How did the predictors do?

Two ways that SAXS data helps.

1. Improve overall shape (density using gmfit tool)
2. Improve fold (GDT-TS score)

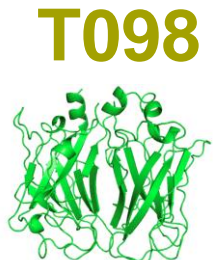
11 CASP13-SAXS targets include 4 monomers and 7 multimers, 14-340 kDa



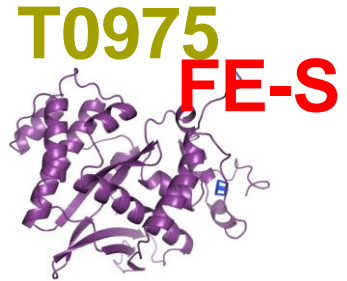
14



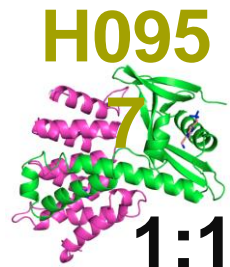
17k



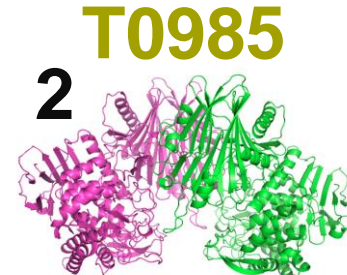
46
kDa



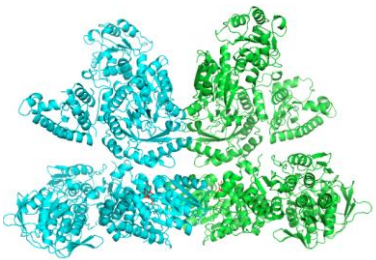
38 kD



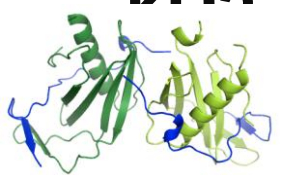
1:1
36
kD



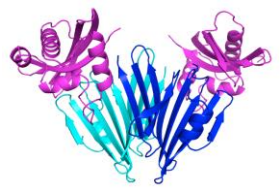
2
200 kDa



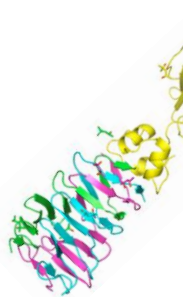
T099
homodim
340
kD



H098
2:2
40
kD



H0968
54
kDa
2

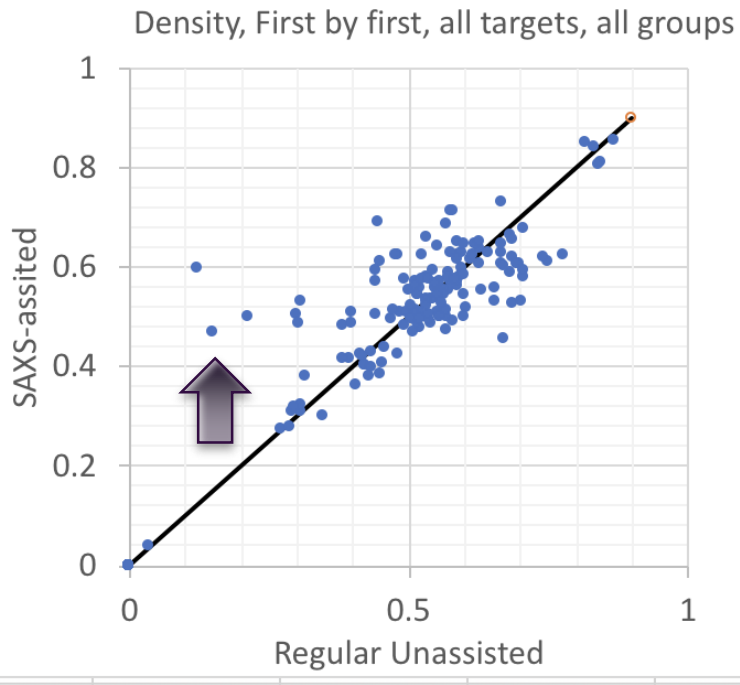
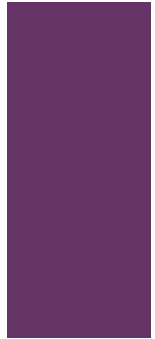


3:
1
48
kD

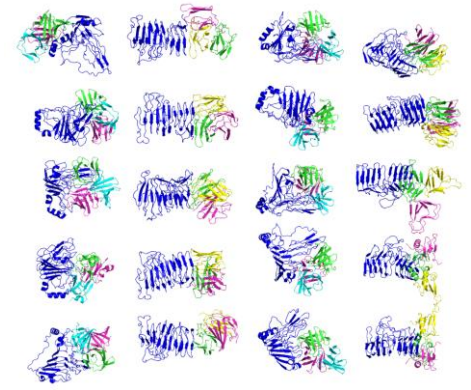


T098
homotrim
228
kD

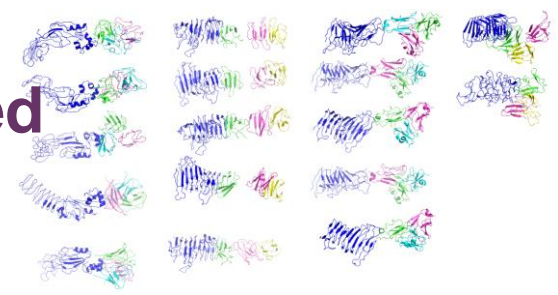
Shape – Predictors improved their overall shape with SAXS data.



Regular
(unassisted)



SAXS-assisted

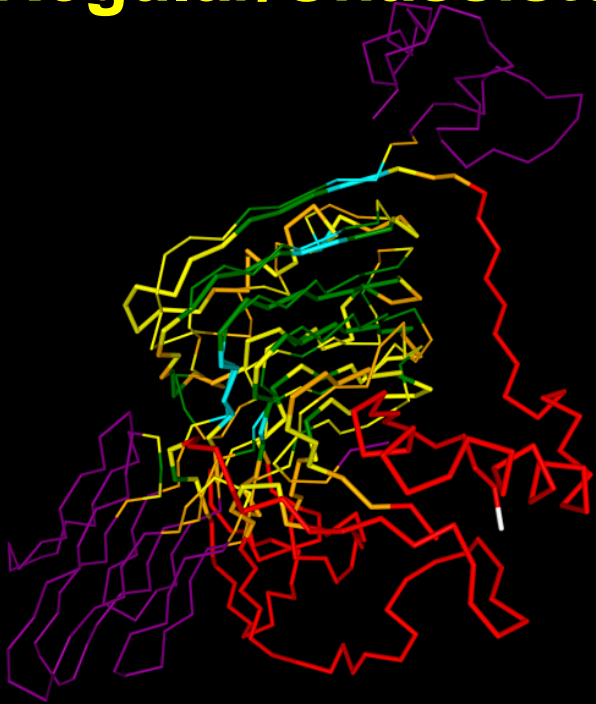


For the top scoring GDTTS model for H0953s2, can visibly see improved fold



196 Grudinin

Regular/Unassisted

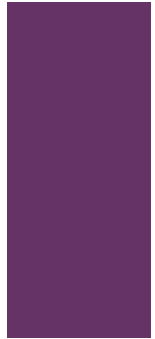


Purple=crystal

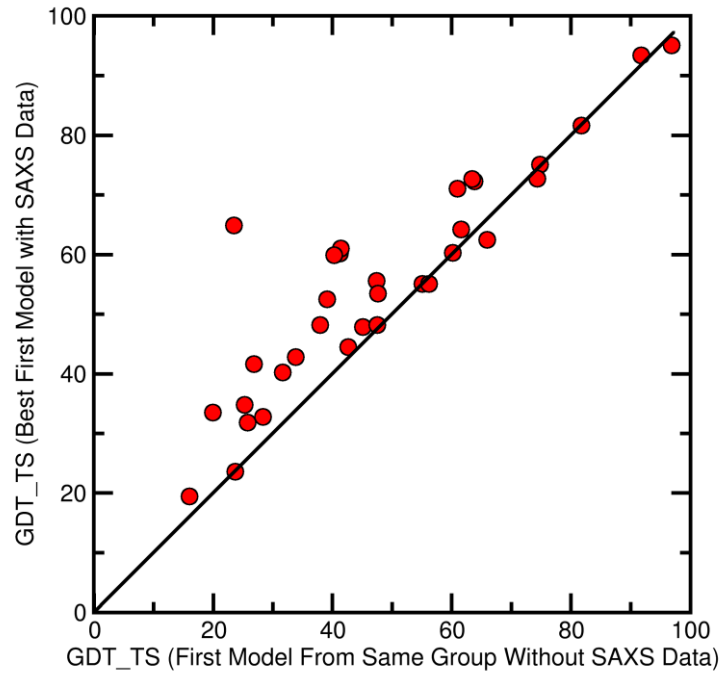
SAXS-assisted



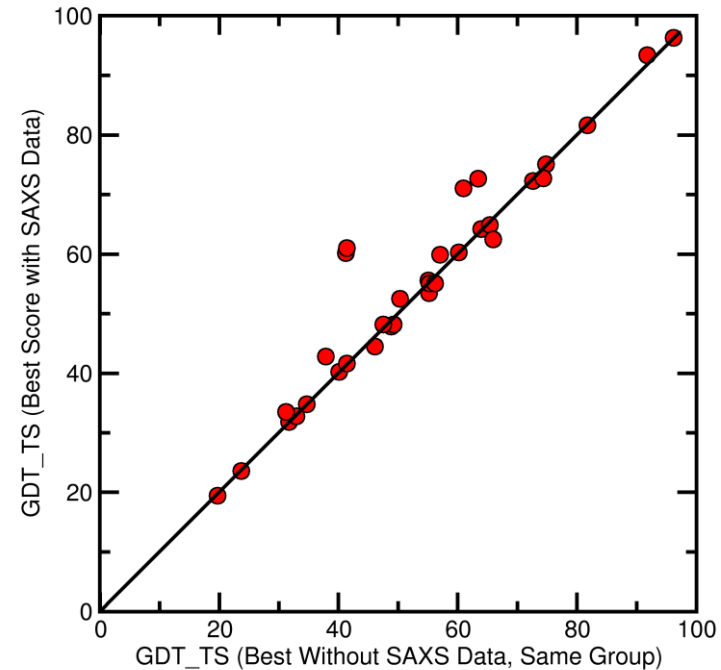
If you look at the top scoring model 1 for every target, compared to the regular entry from same group, see improved GDTTS. Do not see if compare best to best, suggesting improved ranking with SAXS.



SAXS Model 1 vs Reg Model 1



SAXS Best vs Best



For S09868 s1 which showed GDTTS improvement, see that whole protein fold improves.

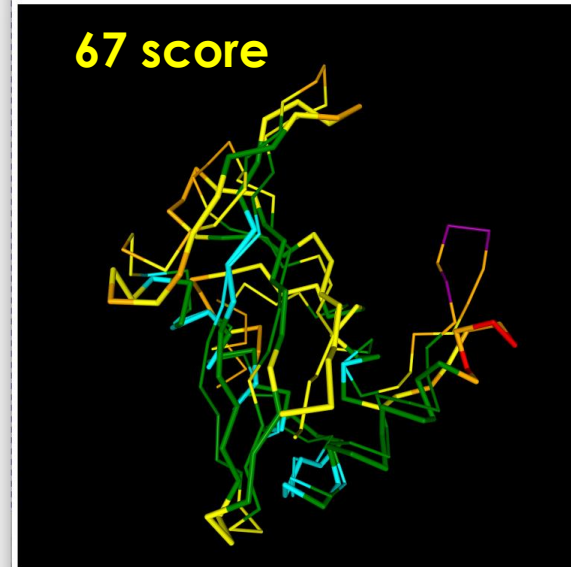
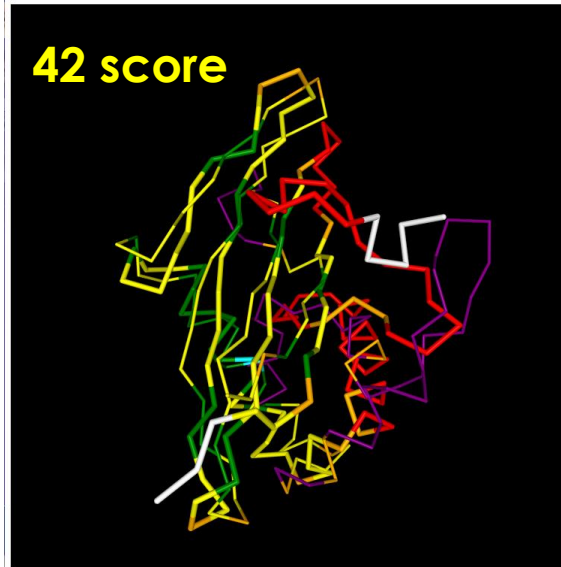


Regular

+SAXS

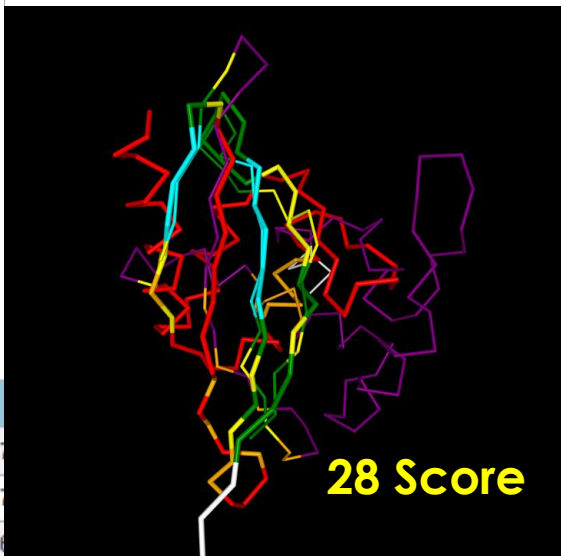
Group1

309/Seder1



Group2

329/D-Haven



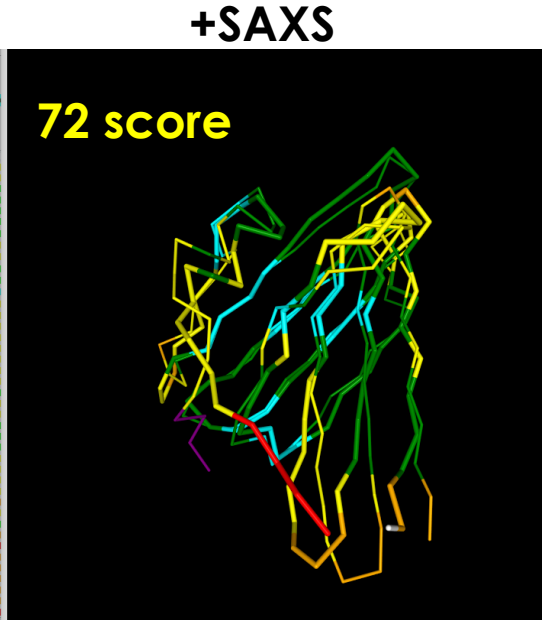
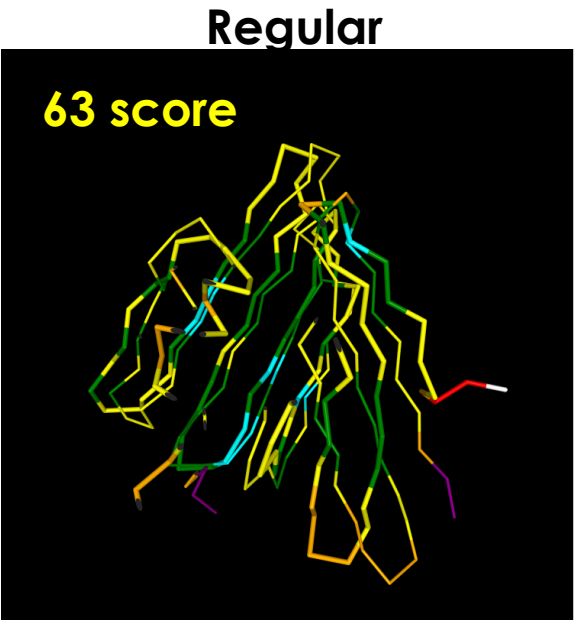
#	Model	(T0)	(S)
1.	_0968s1TS309_1	41.737	66.737
2.	_0968s1TS329_1	28.178	66.737
3.	_0968s1TS196_1	61.017	58.686

For S09868 s2 which showed GDTS improvement, see that edges improves. For biologists, edges are important – it's where active sites and interfaces are.



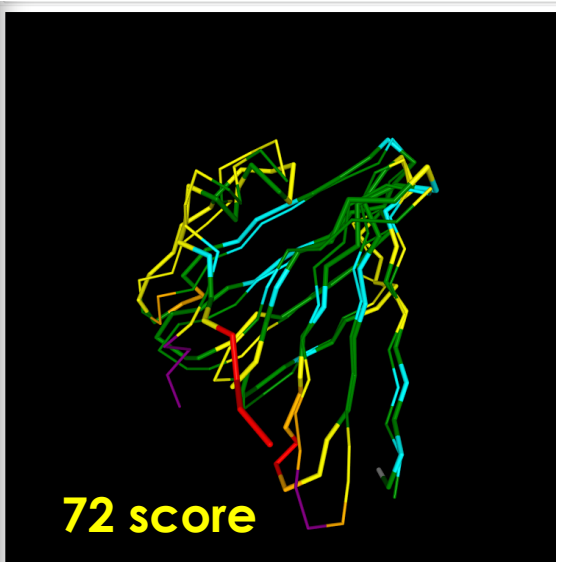
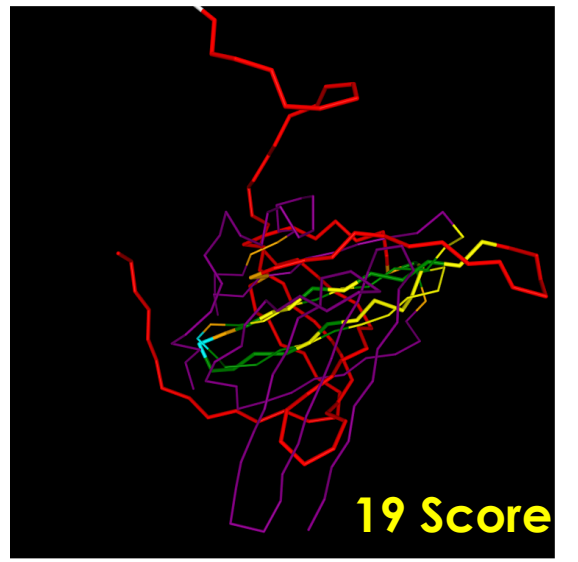
196/Grudinin
(#1 saxs)

Group1



329/D-Haven

Group2



For S0957 s2 which showed GDTS improvement, see that biologically-important edges improves.

Regular

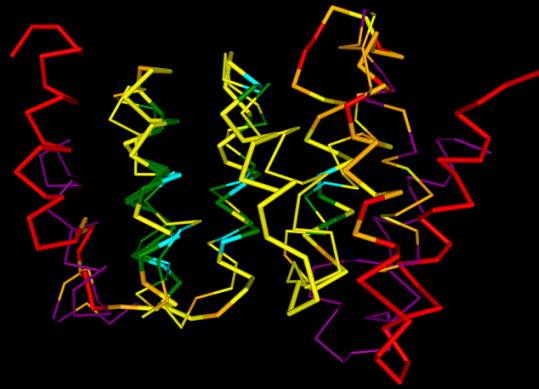
+SAXS

329/D-Haven

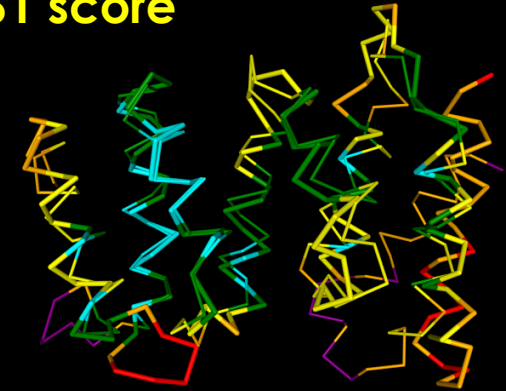
Group1

#	Model	(T0)	(S)	(S) Δ
1.	_0957s2TS329_1	41.452	60.968	19.516
2.	_0957s2TS196_1	40.000	56.935	16.935
3.	_0957s2TS135_1	40.000	49.355	9.355

41 score



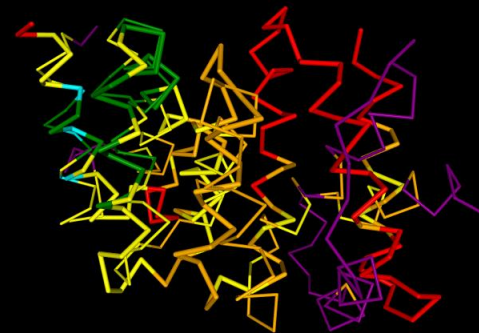
61 score



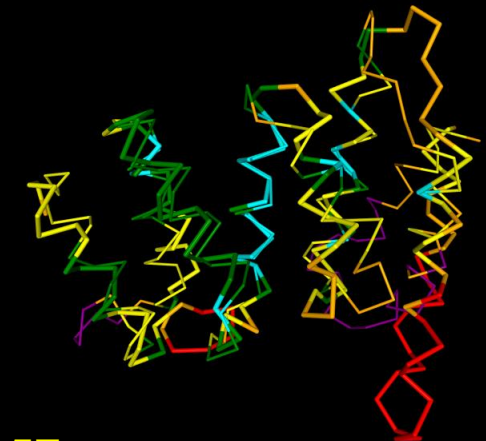
196/Grudunin

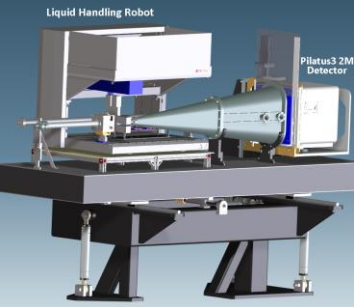
Group2

40 Score



57 score





Six Take Home Messages on What can SAXS do for you?

X-ray Scattering by electrons provides *distances* between electrons.

Small Angle X-ray Scattering measures all electron pair distances in a protein in solution

Atomic models can be quantitatively compared with SAXS data

SAXS can validate protein structure predictions

SAXS can reveal protein conformations occurring in solution at the atomic level

***Atomic models* are more powerful than shape because they can be tested.**

ANY QUESTIONS ON THIS SECTION?

SIBYLS Related Staff



John Tainer

Michal Hammel

Scott Classen



Jane Tanamachi



Curtis Hodge



David Shin



Kathryn Burnett



Susan Tsutakawa



Daniel Rosenberg

CASP Committee

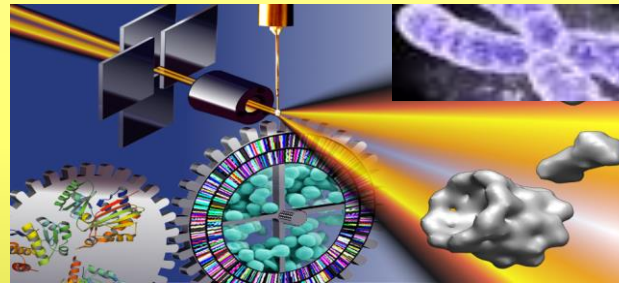
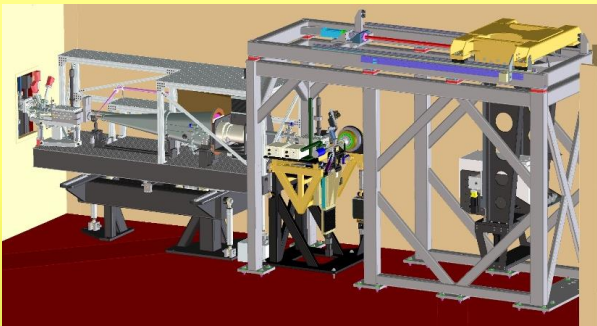


Andriy Kryshchak

Krzysztof Fidelis

John Mout

Dmytro Guzenko



DOE

CASP-SAXS Acknowledgements



CASP Commons

Vision Statement

Protein structures drive biology.

CASP Commons will engage the protein modeling community with the broader biological community to address important problems in biology and medicine.



CASP Commons Goals

Three key objectives:

- Structural models for biology
- Provide a bridge between the modeling and biological communities
- Drive methods for data-assisted modeling

CASP Commons

Implementation of the Vision

To engage the CASP scientific community in both 'regular' and 'data-assisted' protein structure modeling on a large number of biomedically-important proteins and complexes for which high-resolution experimental structures are not available.

Example CASP Commons Activities

- Targets broadly nominated by the biology community
- Targets involved in a particular disease (e.g. cancer), organism, or biological process.

CASP Commons

Targets and data are being generated by CASP Organizers

Proposed by high-impact biomedical research labs.

Range from 50 to 200 residues. May be monomers or oligomers.

No good templates can be identified for modeling.

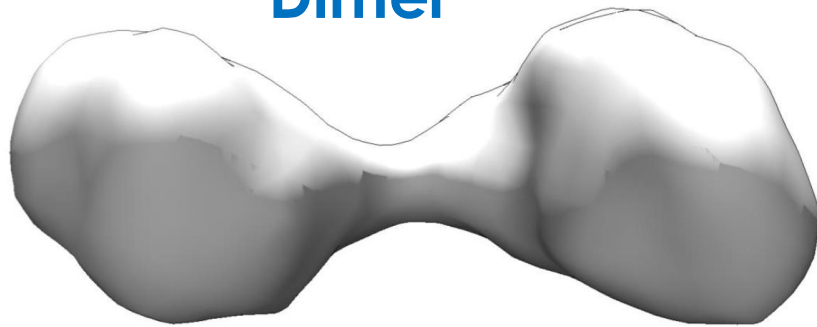
Shallow multiple sequence alignments ($N_{\text{eff}} / L < 2$).

Structures to eventually be determined by CASP Organizers – may not have 3D structures for assessment for some years.

Assessment will be an ongoing activity.

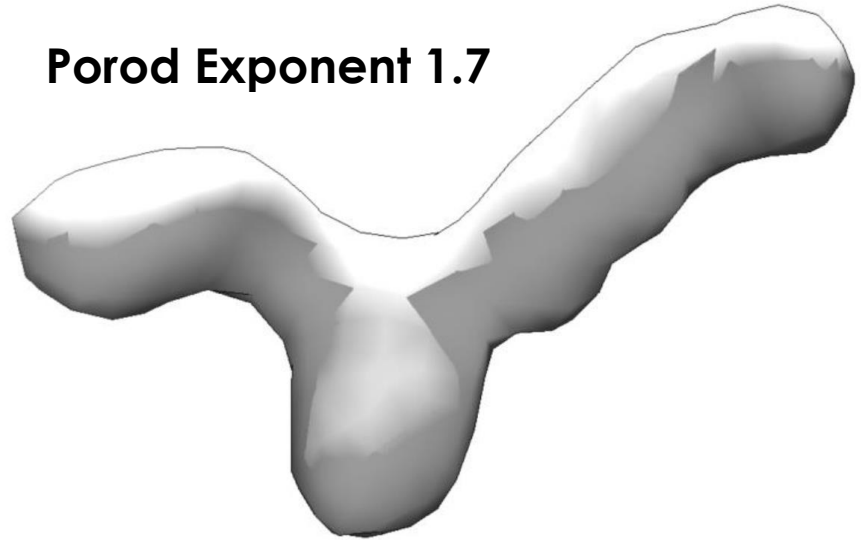
Of three CASP-Commons targets analyzed, one is a dimer in solution. One is globular. Two are very flexible.

Dimer

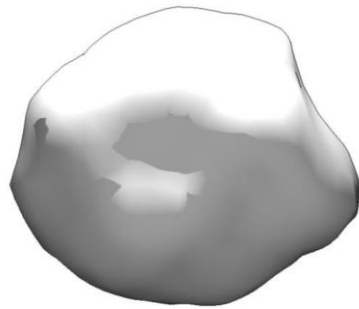


Porod Exponent 2.0

Porod Exponent 1.7



Well-folded



Porod Exponent 3.9

Porod Exponent
4 globular
3 Beads on a string
2 Unfolded

CASP Commons



Y. Ishida, N. Denissova, G. Liu,
G. V. T. Swapna, G. T. Montelione,

G. Hura, S. Tsutakawa, J. Tainer

J. Moult

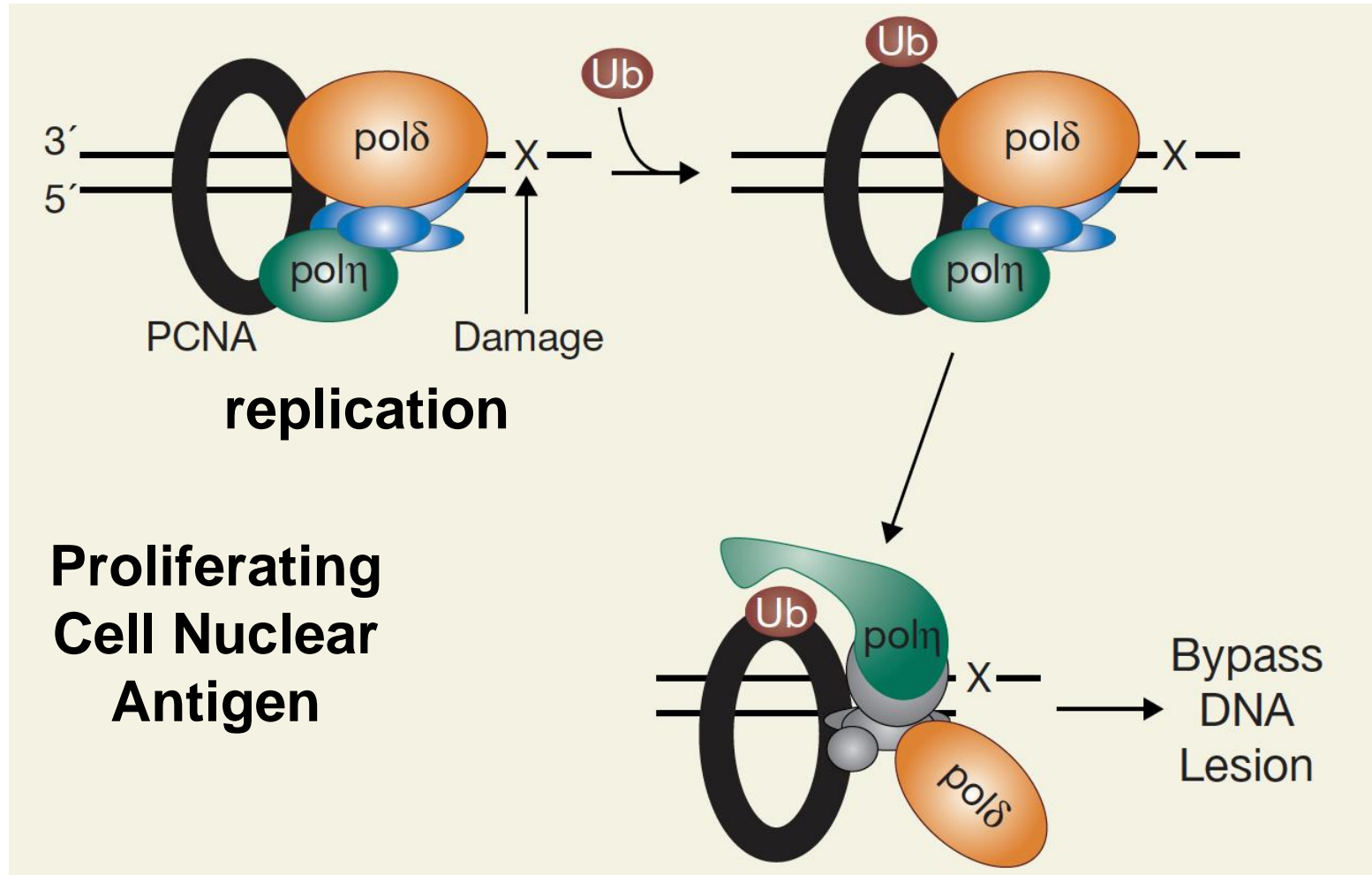
Chin-Hsien Emily Tai

K. Fidelis, A. Kryshfovych



SAXS is powerful for
characterizing proteins that
adopt multiple conformations

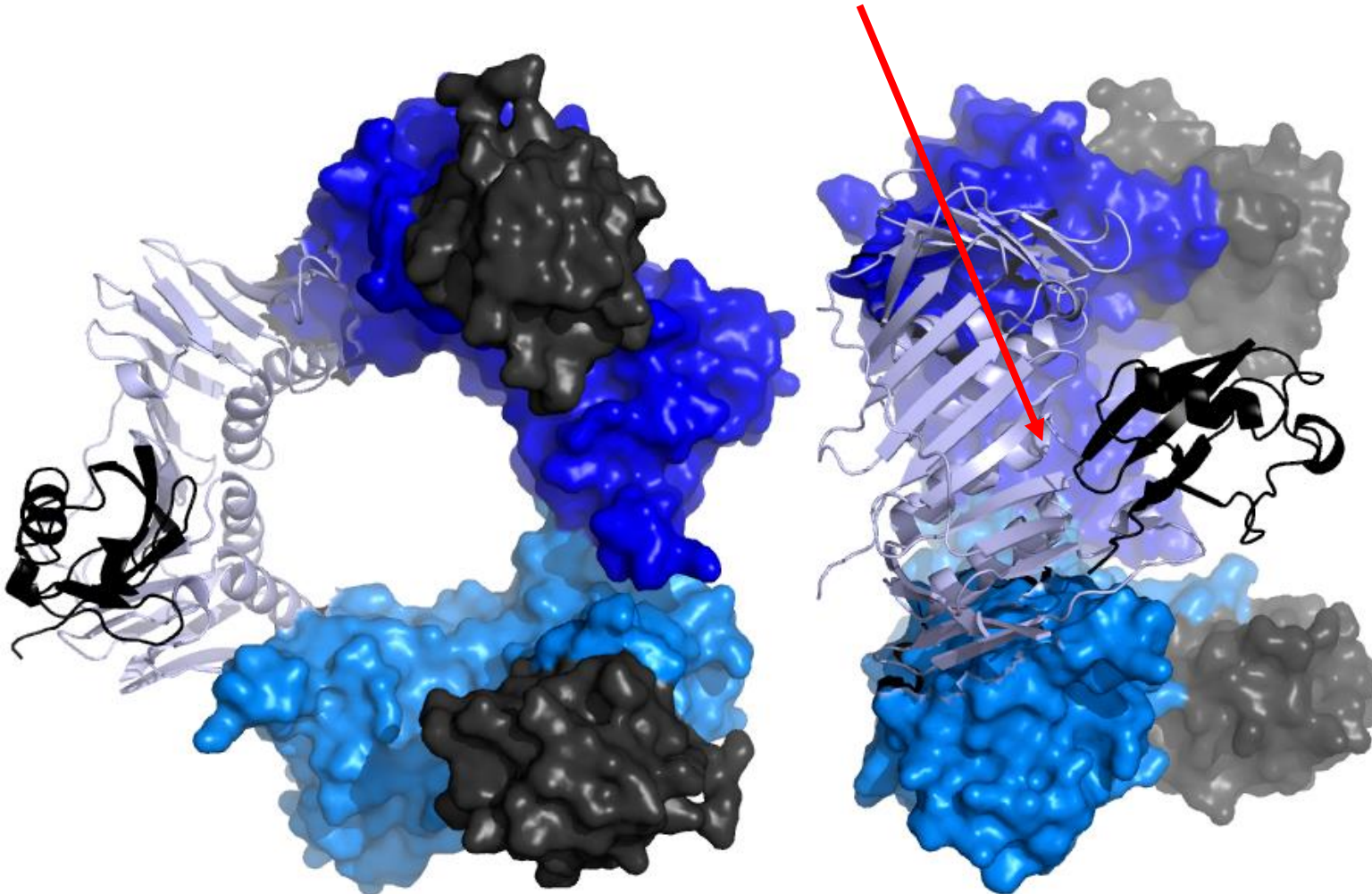
PCNA is a central protein in replication. It is modified when the replication complex encounters DNA damage.



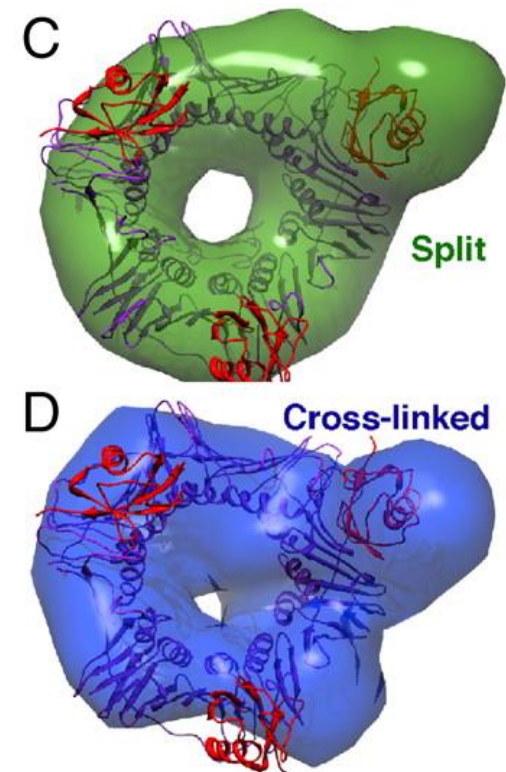
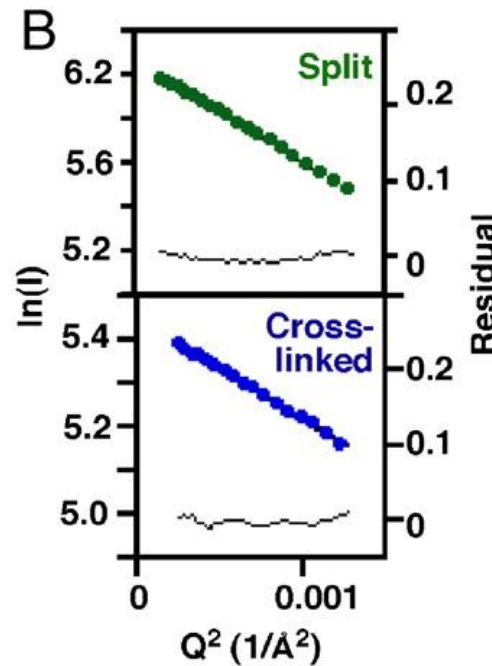
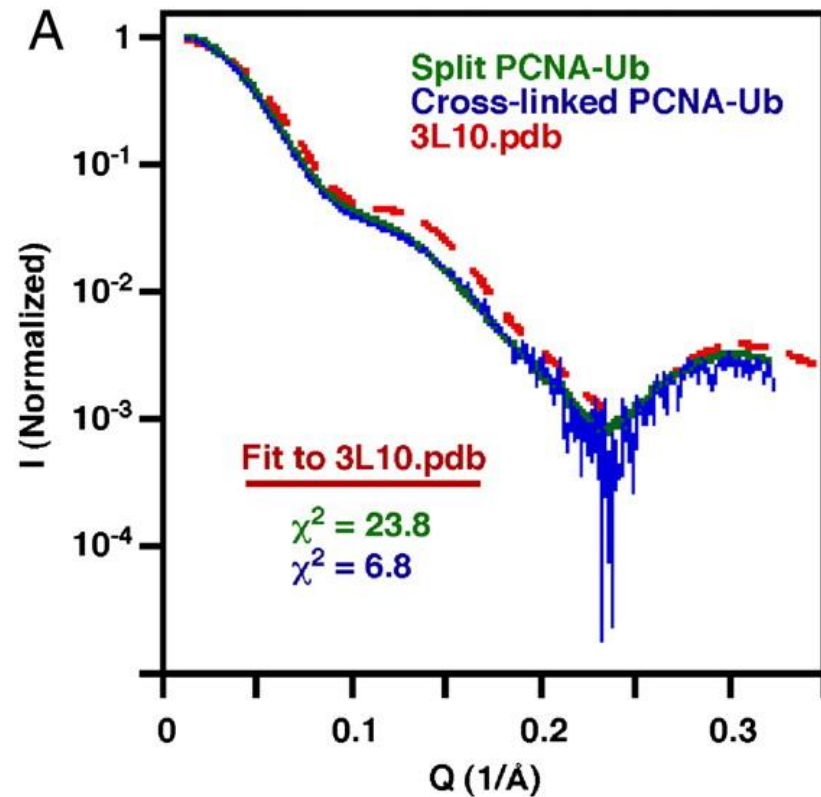
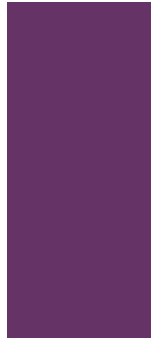
Swaminathan, Nature Cell Biology, 2004

The crystal structure of UbPCNA had the bypass interface buried. What was happening?

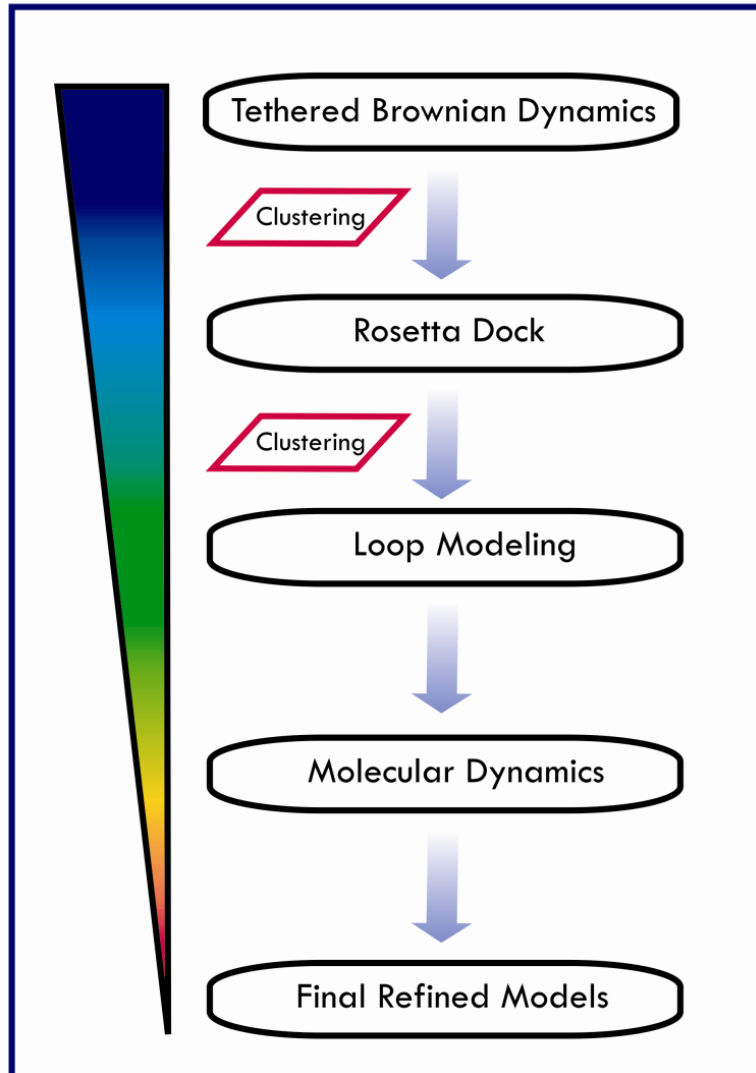
Pol Eta binding interface



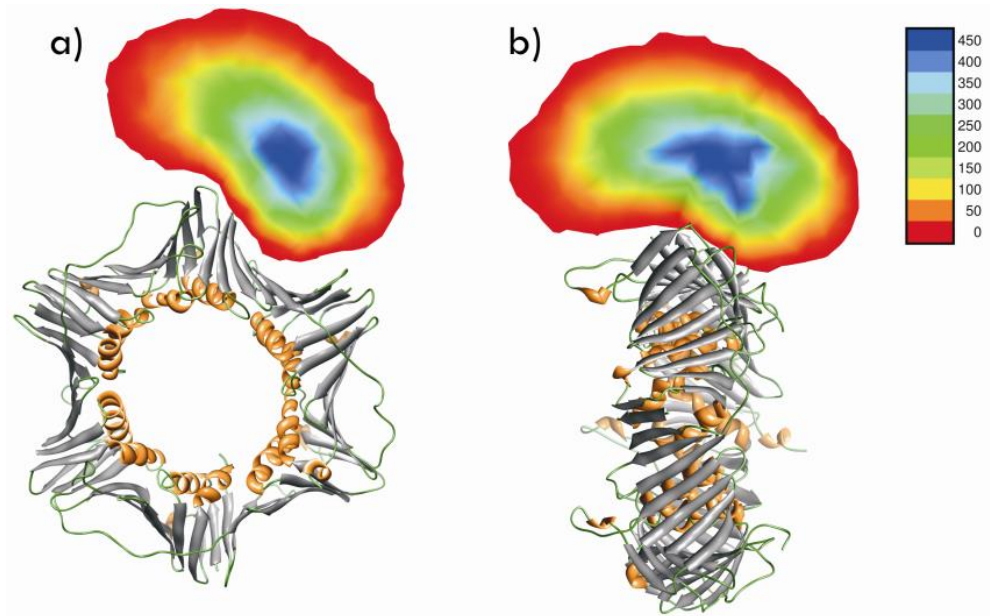
SAXS data on triubiquitinated PCNA (3 ubiquitin per homotrimer) was not consistent with three ubiquitin per homotrimer or flexible.



Multiscale Computational studies of PCNA-Ub conducted independently of SAXS study identified a novel position along the ring.



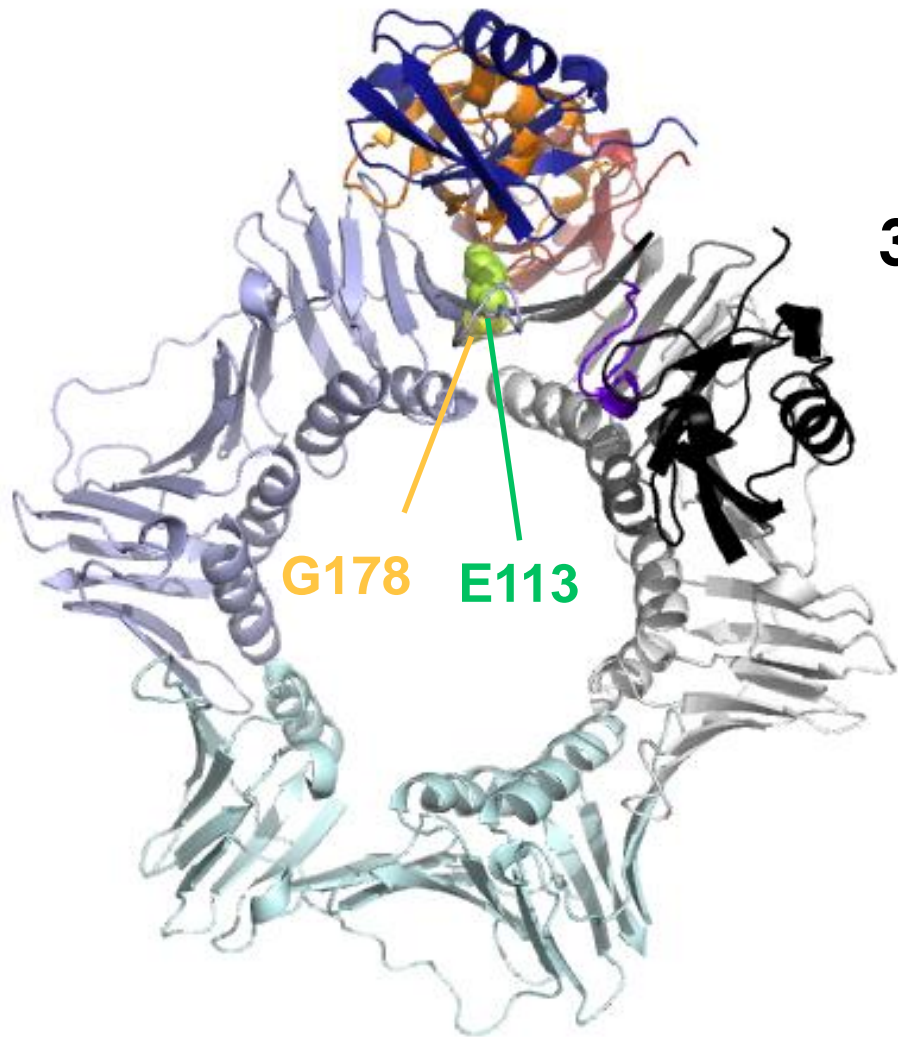
Tethered Brownian Dynamics



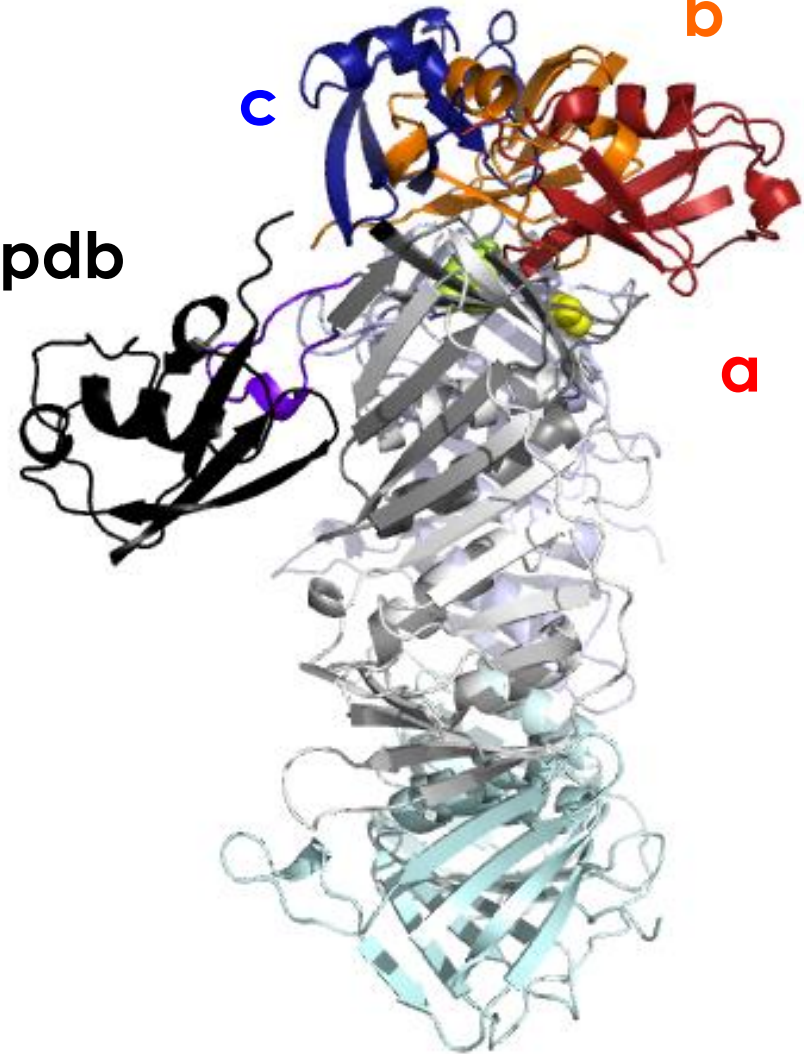
Ivaylo Ivanov, Andrew McCammon

Tsutakawa, PNAS, 2011

Multiscale Molecular Dynamics identified three positions – along ring that would be hidden in the SAXS envelope.



3l10.pdb



Tsutakawa, PNAS, 2011

An ensemble of atomic models revealed the Ub PCNA could adopt multiple conformations



GENERATE MODELS

MINIMAL ENSEMBLE SEARCH (MES)

MD positions a,b,c
Xtal position x
Flexible position f

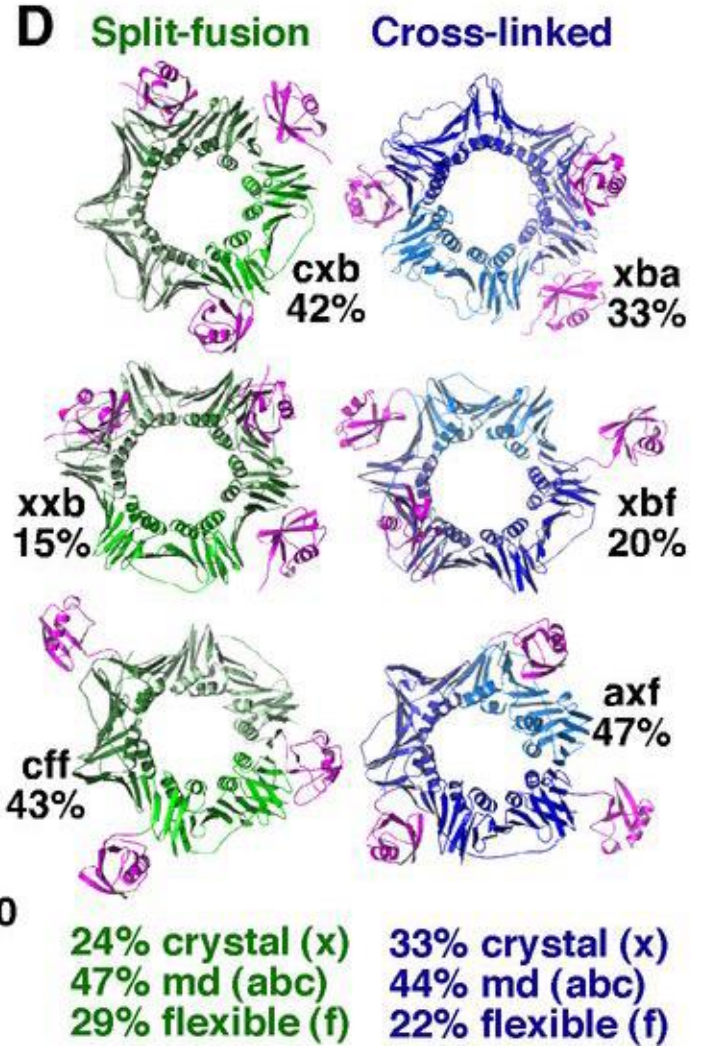
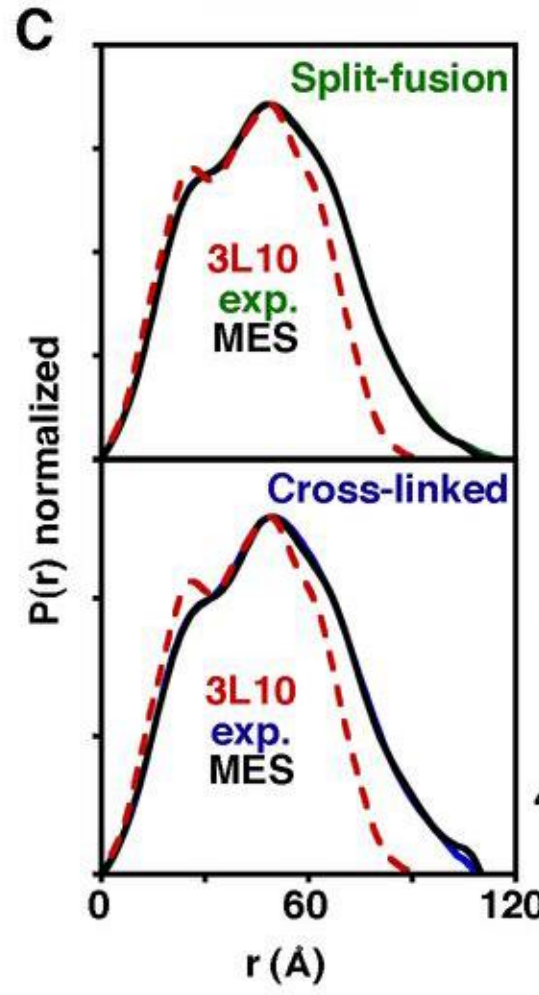
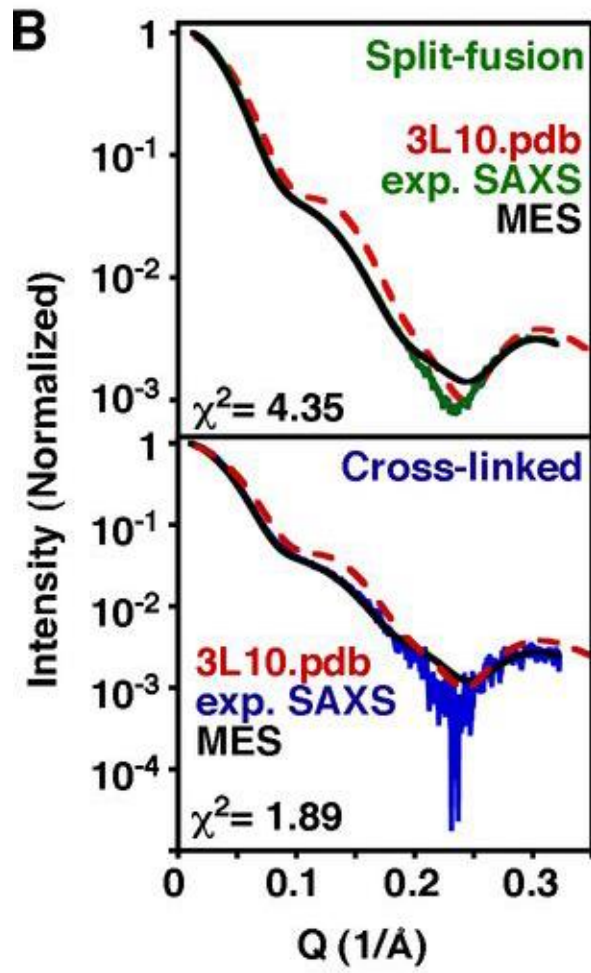


130 models
a,a,a a,a,x
a,a,b a,a,f
a,a,c a,a,x
etc



Ensemble of
3 models
that fit the
SAXS data

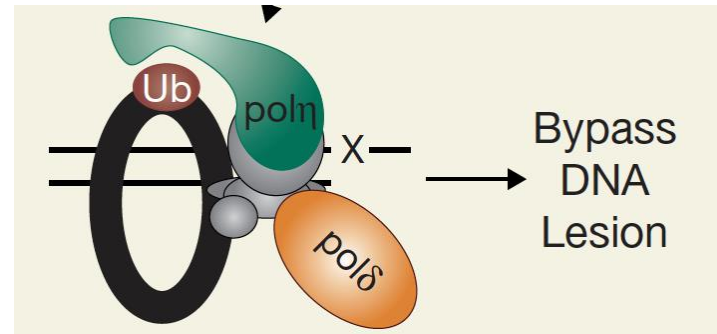
An ensemble of atomic models revealed the Ub PCNA could adopt multiple conformations



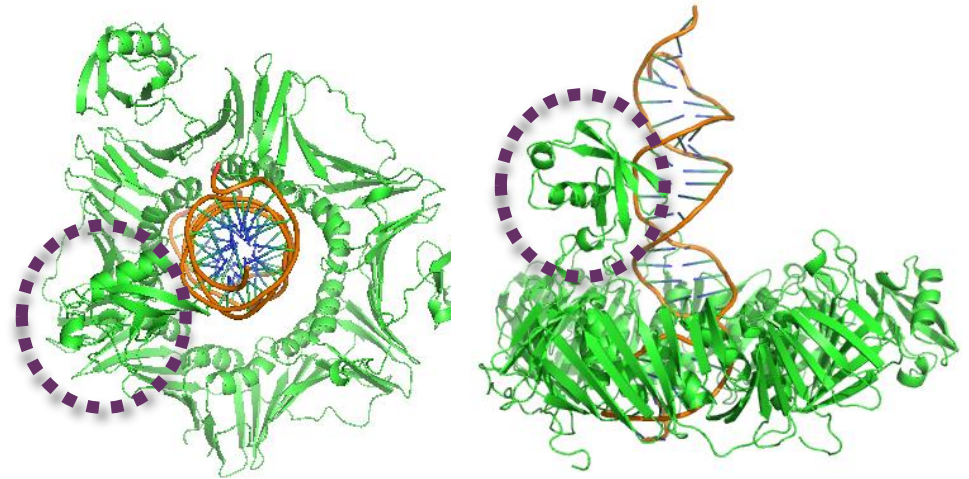
SAXS provides biologically relevant functional information.

Ubiquitinated PCNA adopts multiple conformations in solution.

In docked crystal position, ubiquitin could stabilize DNA conformation on PCNA



In flexible conformation, UbPCNA can bind to translesion polymerases.



MD simulation

- John Tainer
- Greg Hura
- Michal Hammel

- Todd Washington, U. of Iowa
 - Lynn Dieckman
 - Brett Freundenthal,
- Ivaylo Ivanov, Georgia State
 - Chunli Yan
- Zhihao Zhuang, U. of Maryland
 - Chris Weinacht

Acknowledgments

NIH