BioSAXS data processing and interpretation

Edward H. Snell

Hauptman Woodward Medical Research Institute, Buffalo, NY 14203, USA.

~~BioSAXS data processing~~ and interpretation

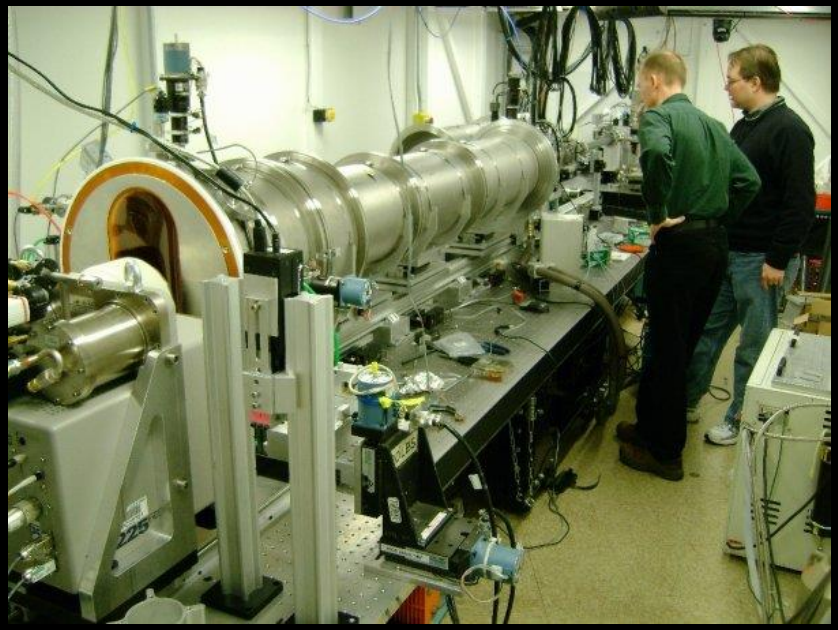
Saved for the practical sessions

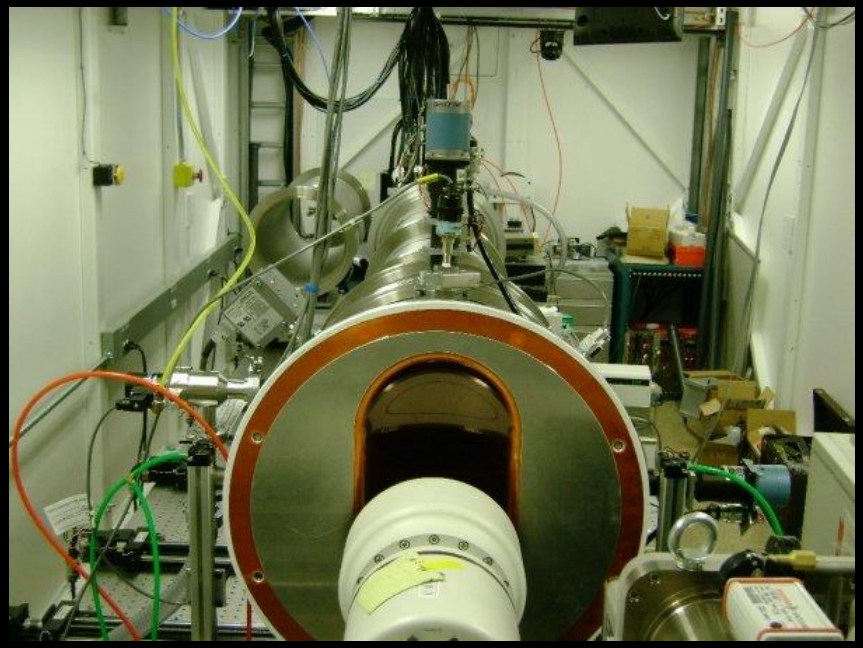
And how you can use it with other techniques

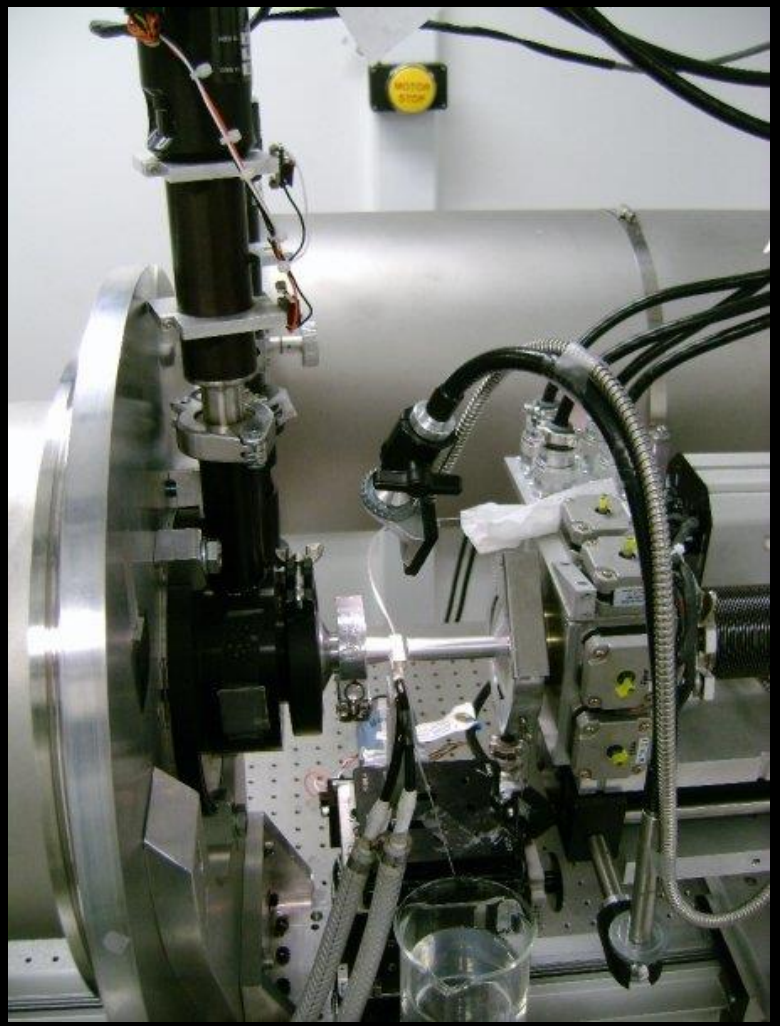
Edward H. Snell

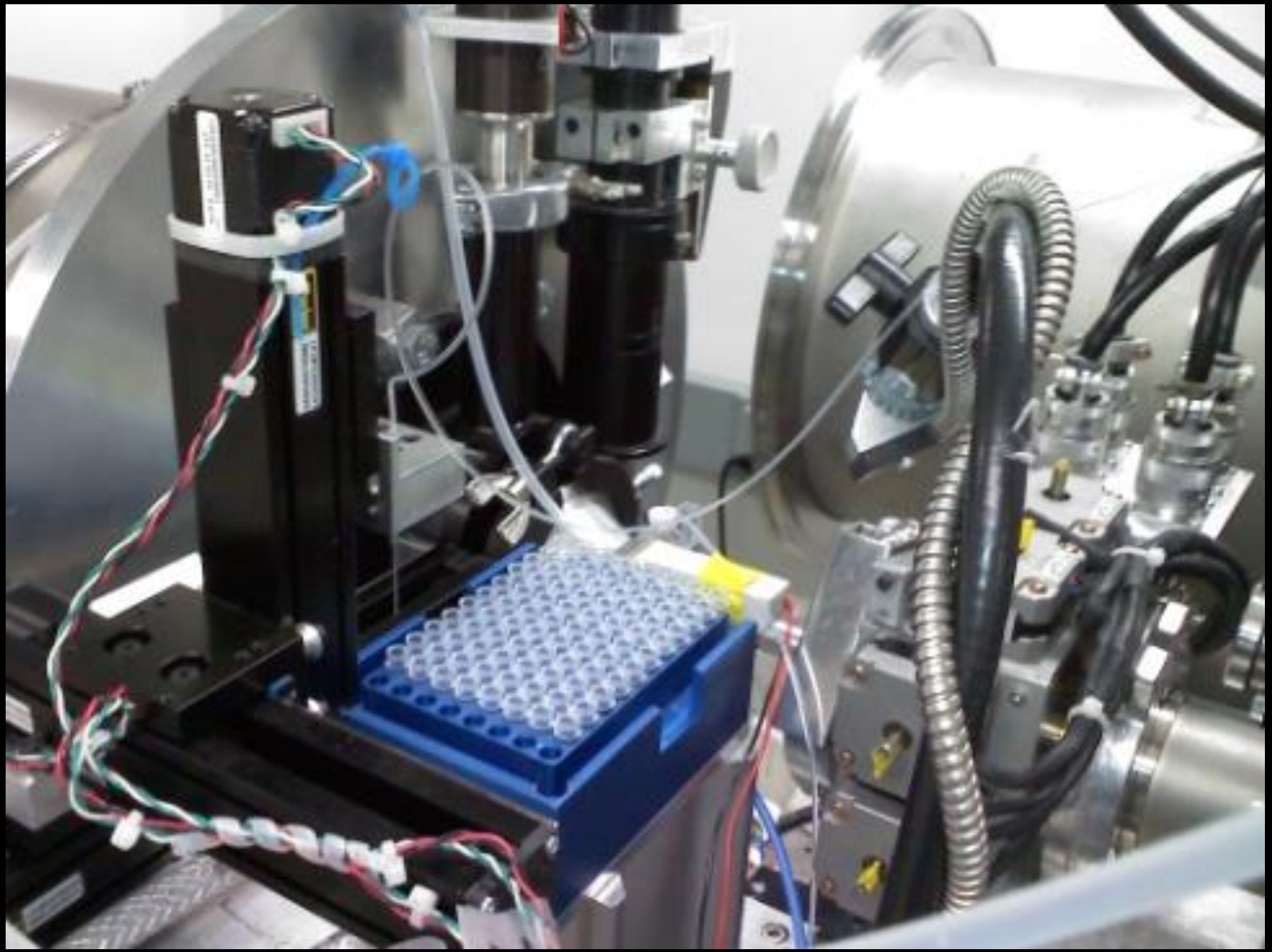
Hauptman Woodward Medical Research Institute, Buffalo, NY 14203, USA.

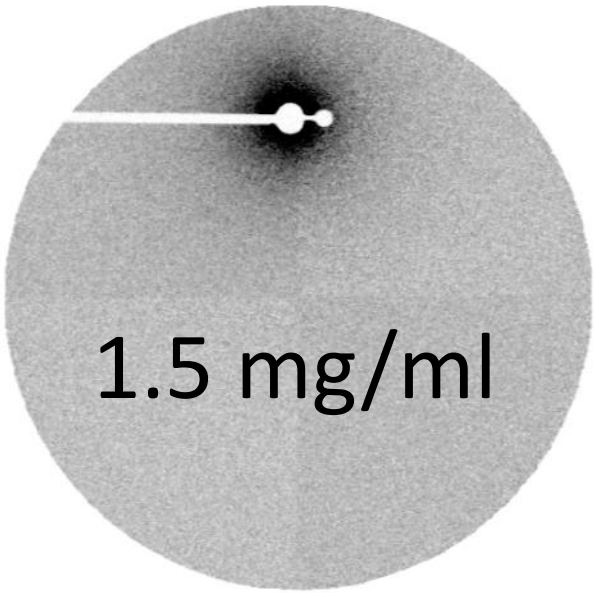
SAXS is a simple experiment

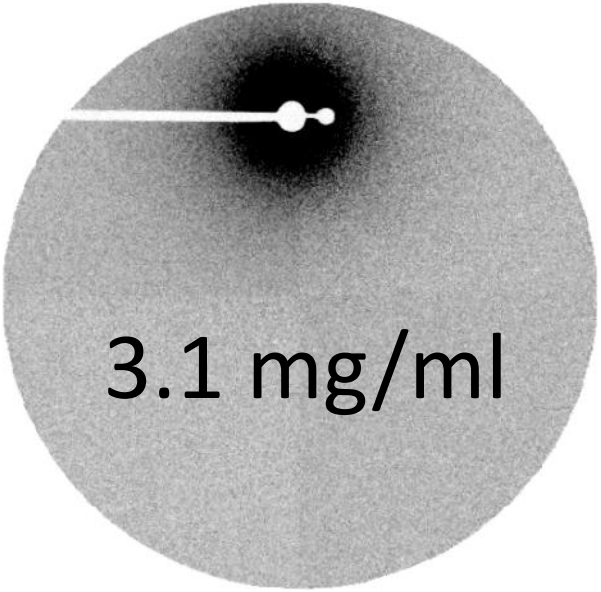


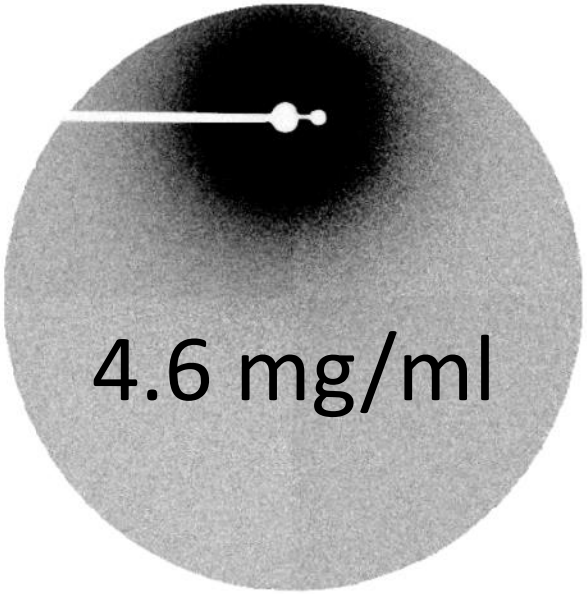


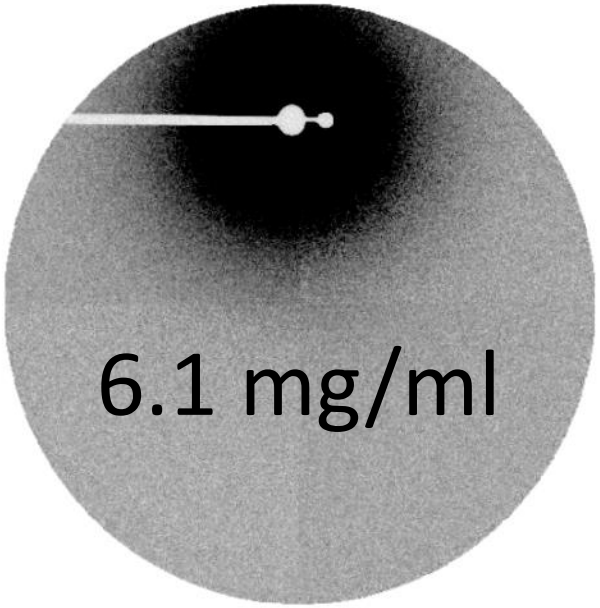


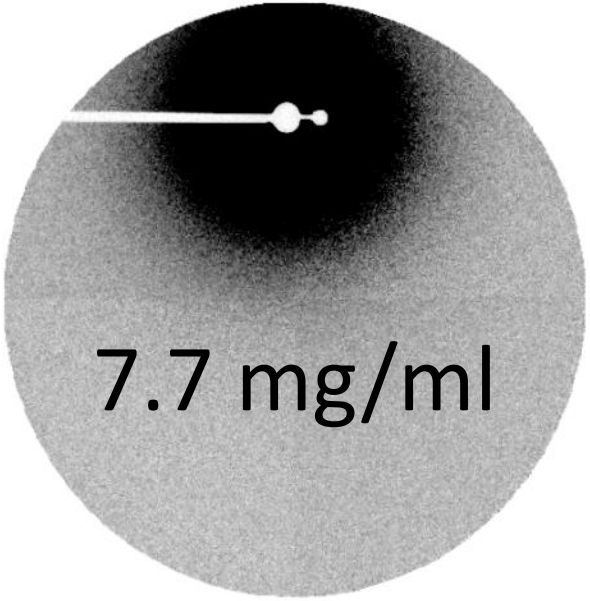



1.5 mg/ml

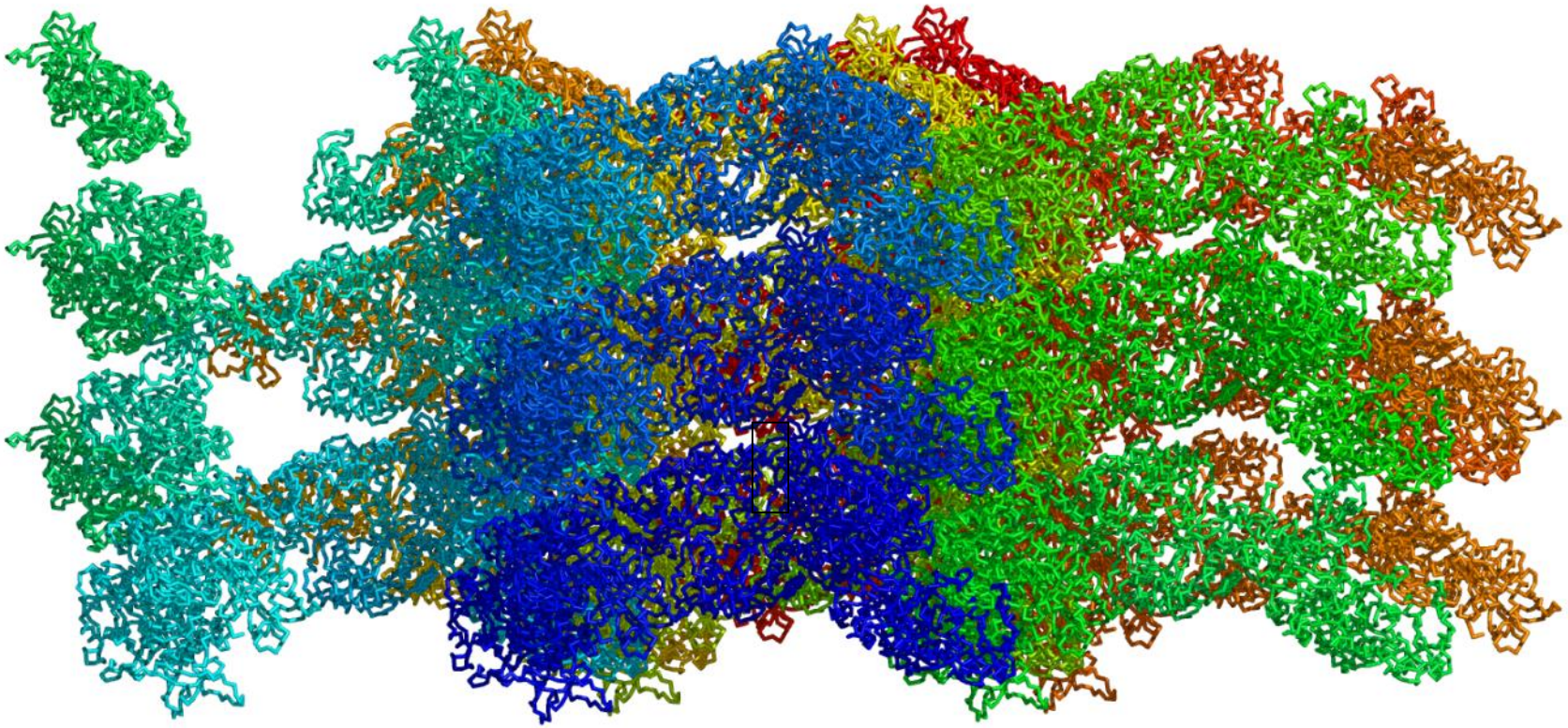

3.1 mg/ml


4.6 mg/ml

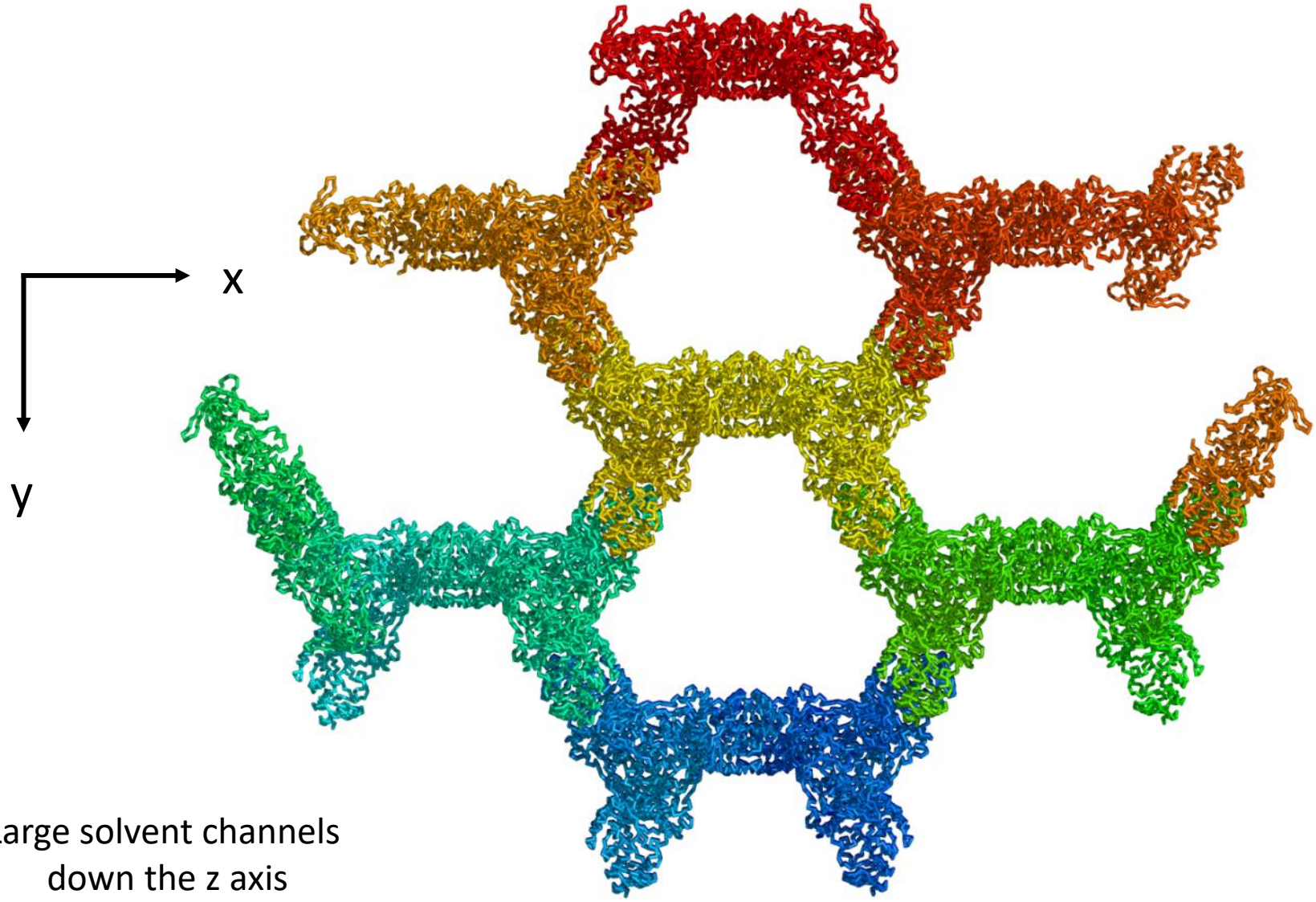

6.1 mg/ml


7.7 mg/ml

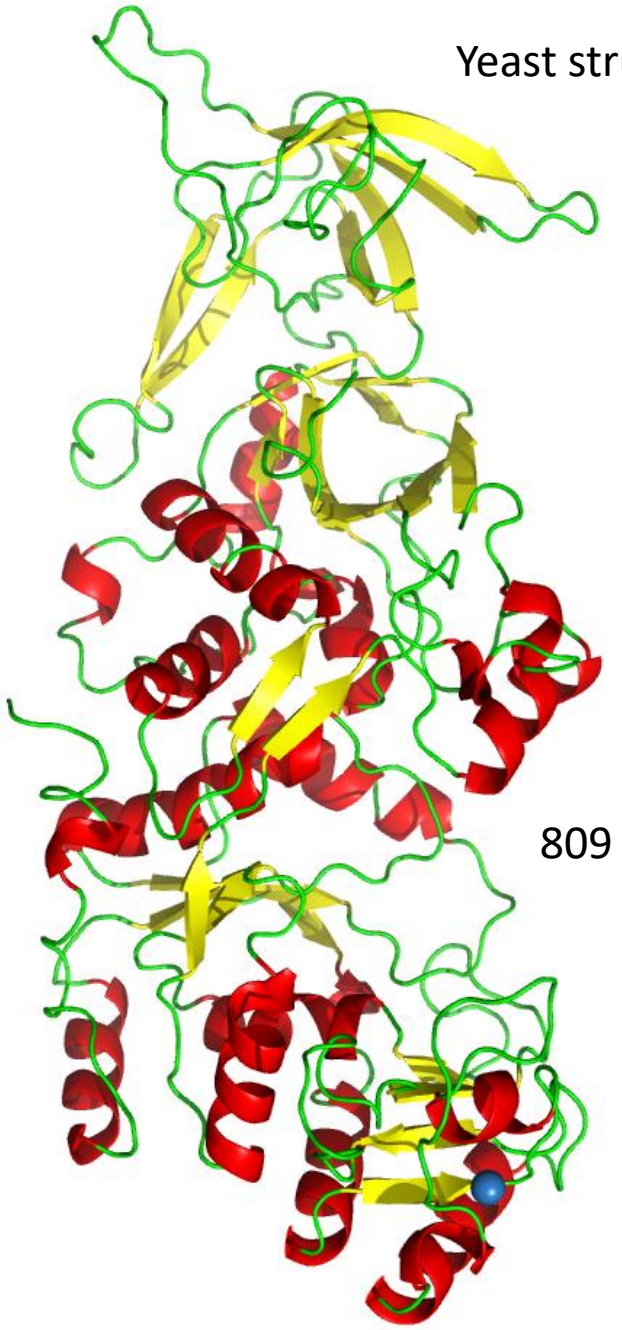
SAXS is a simple experiment
but a powerful one



Tight packing in z and y



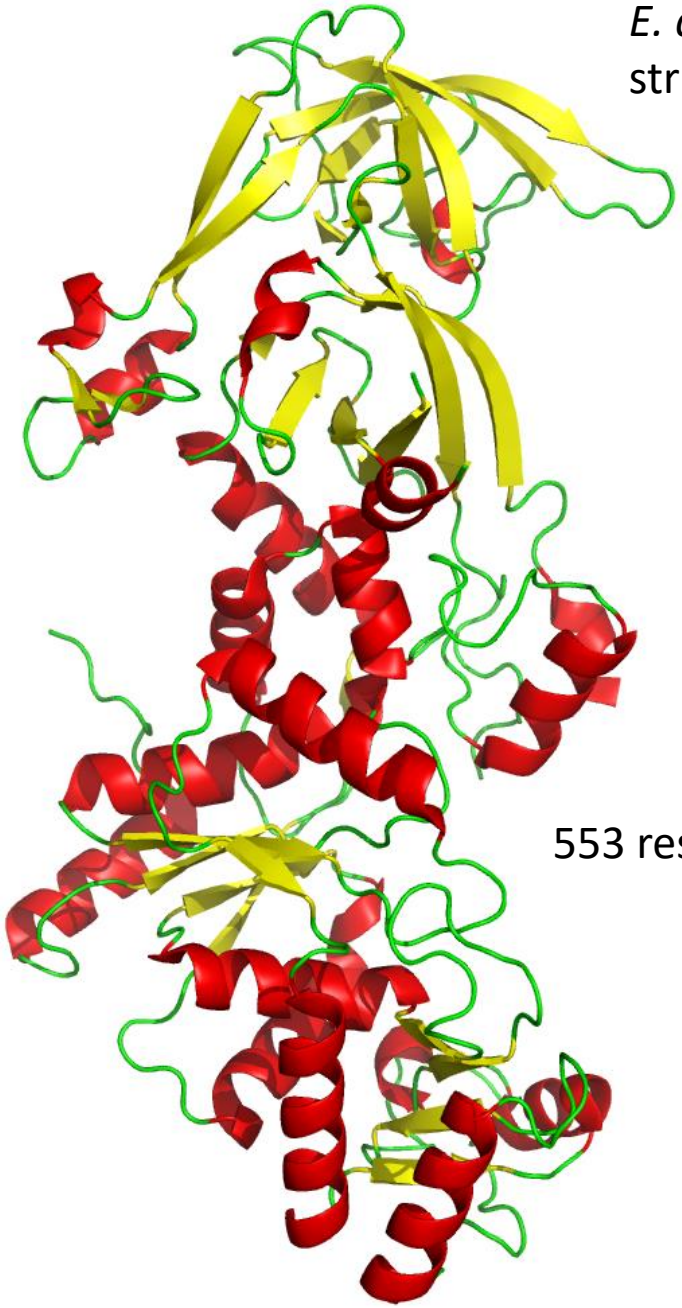
Large solvent channels
down the z axis



Yeast structure

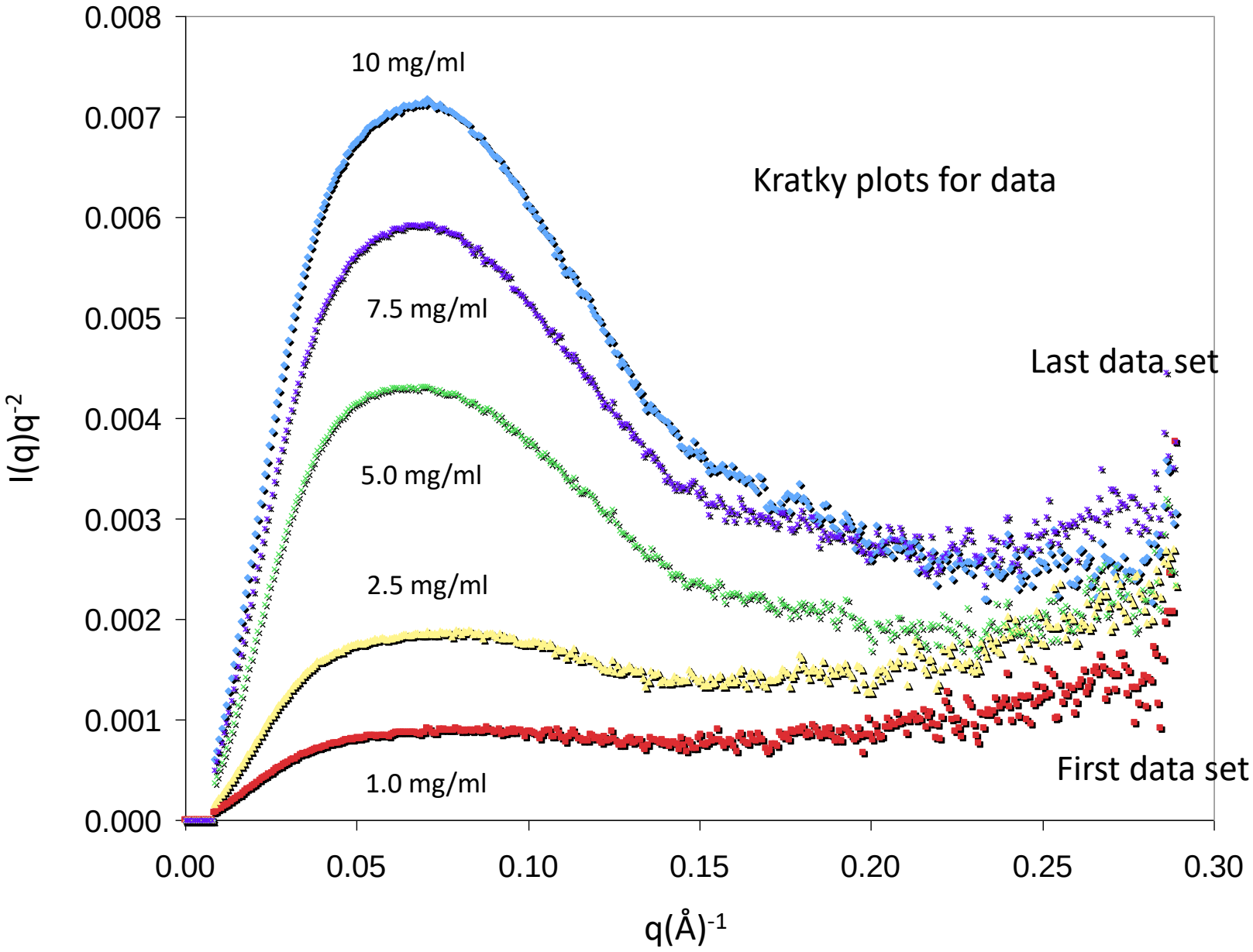
?

809 residues



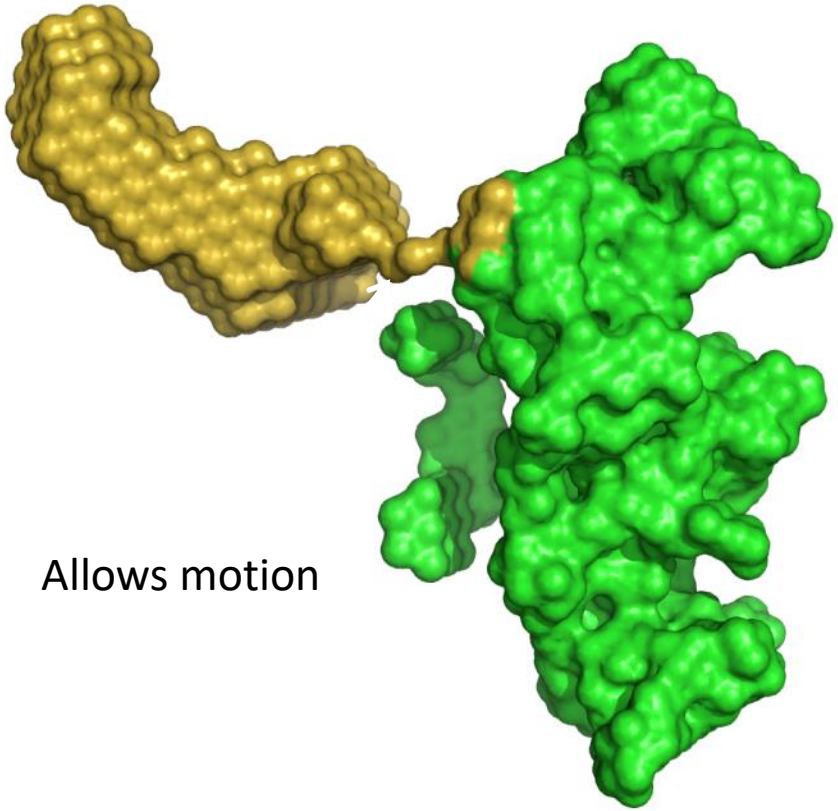
E. coli. structure

553 residues


Kratky plots for data
10 mg/ml
7.5 mg/ml
5.0 mg/ml
2.5 mg/ml
1.0 mg/ml
Last data set
First data set
I(q)q-2
q(Å)-1
0.000
0.001
0.002
0.003
0.004
0.005
0.006
0.007
0.008
0.00
0.05
0.10
0.15
0.20
0.25
0.30

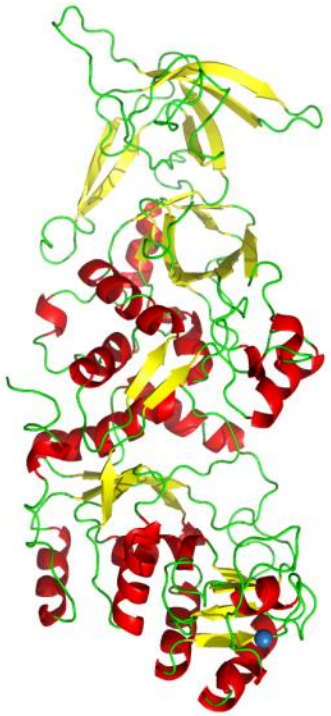
Envelope reconstruction using the crystallographic structure

N terminal domain



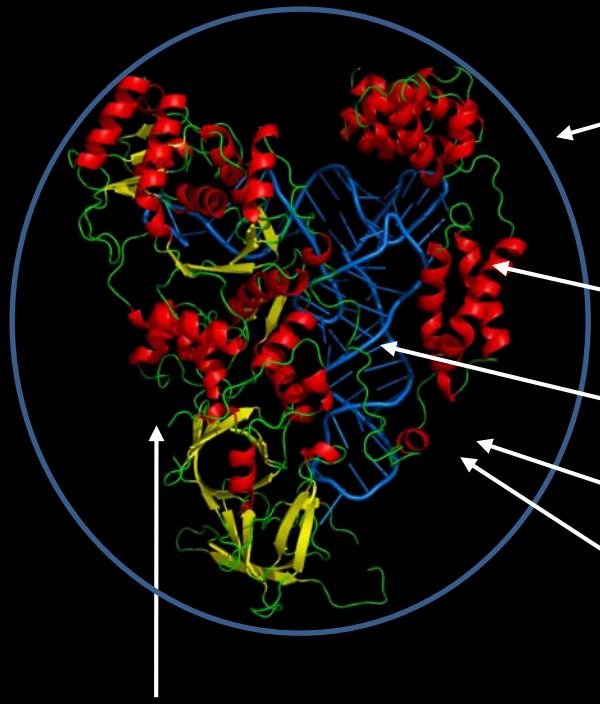
Allows motion

C terminal domain



The crystal structure (which shows only the C-domain)

Eukaryotic Gln tRNA synthetase



SAXS data indicating a larger but well folded system in solution

A Sherlock analysis indicated a preferential pH

The truncated terminal was crystallized

It was extracted directly from the screening plate and X-rayed to give the structure.

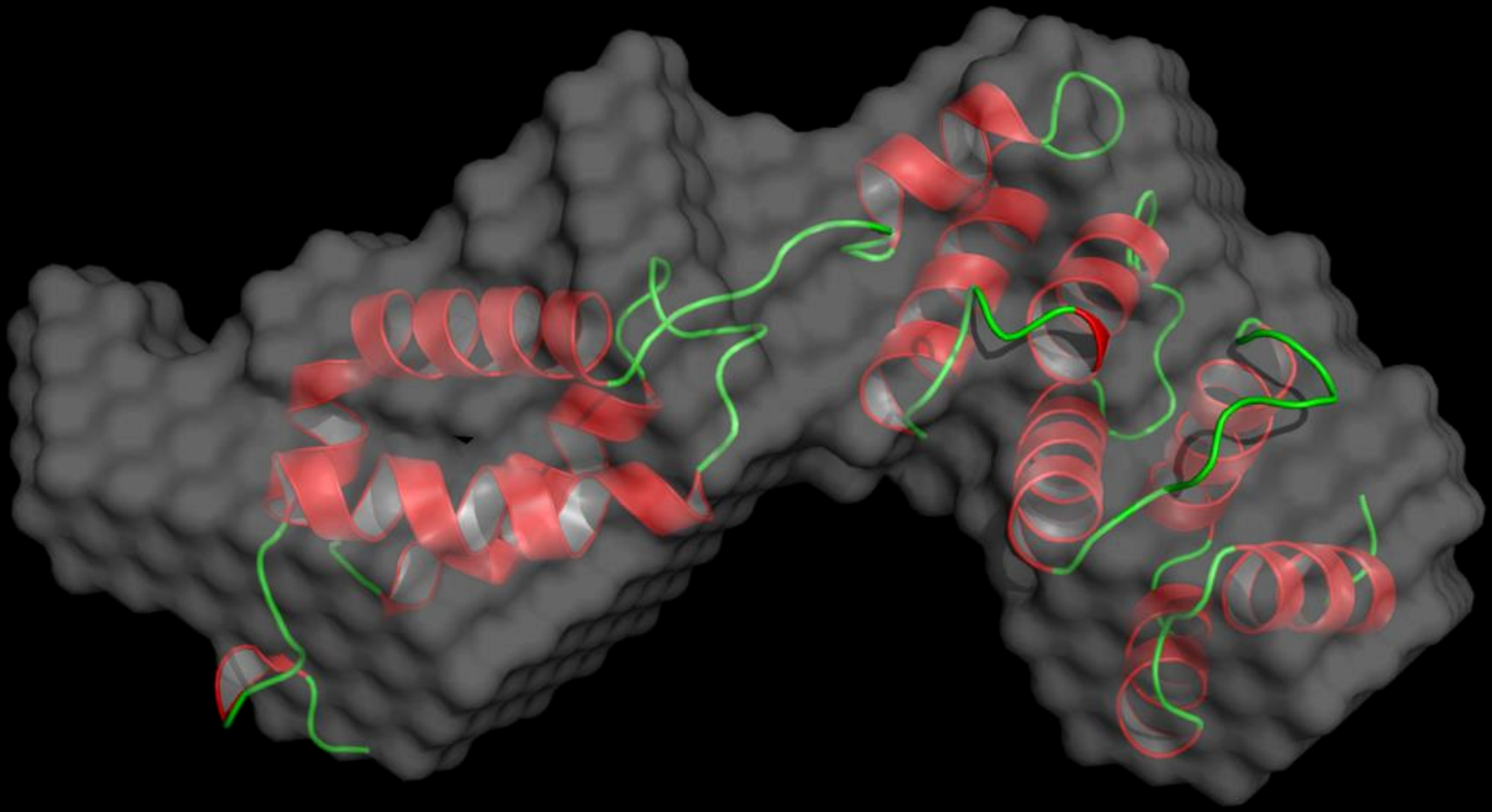
tRNA was docked in

SAXS aided by sequence analysis identified a flexible region

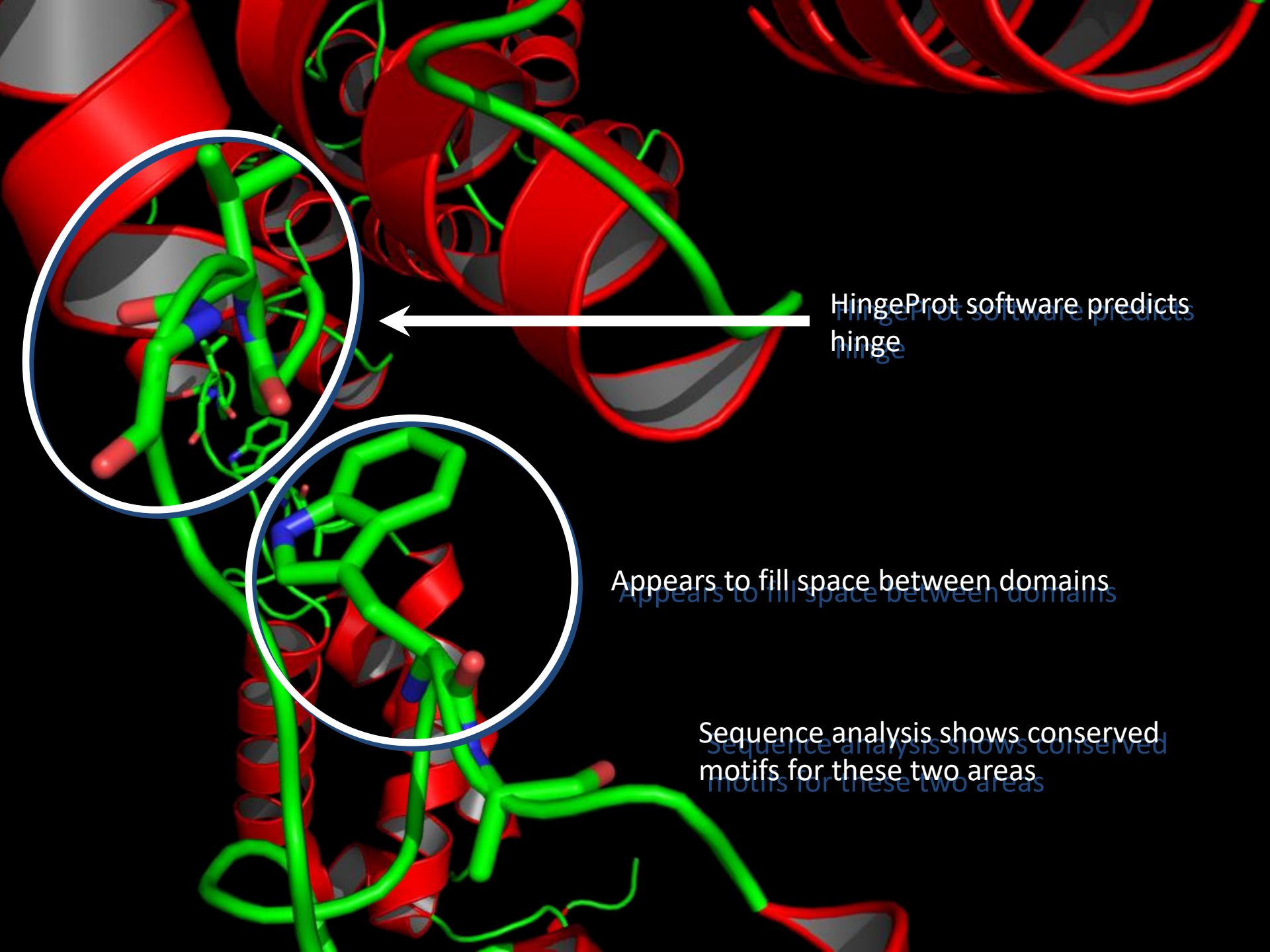
Homology modeling (FREAD) gave the flexible region

Crystallized the C-terminal in the standard screen, conditions chosen that were already known to be good cryo-conditions.

A combination of crystallography, SAXS, homology modeling and computational modeling was used to give the complete structure and tested by biochemical analysis.

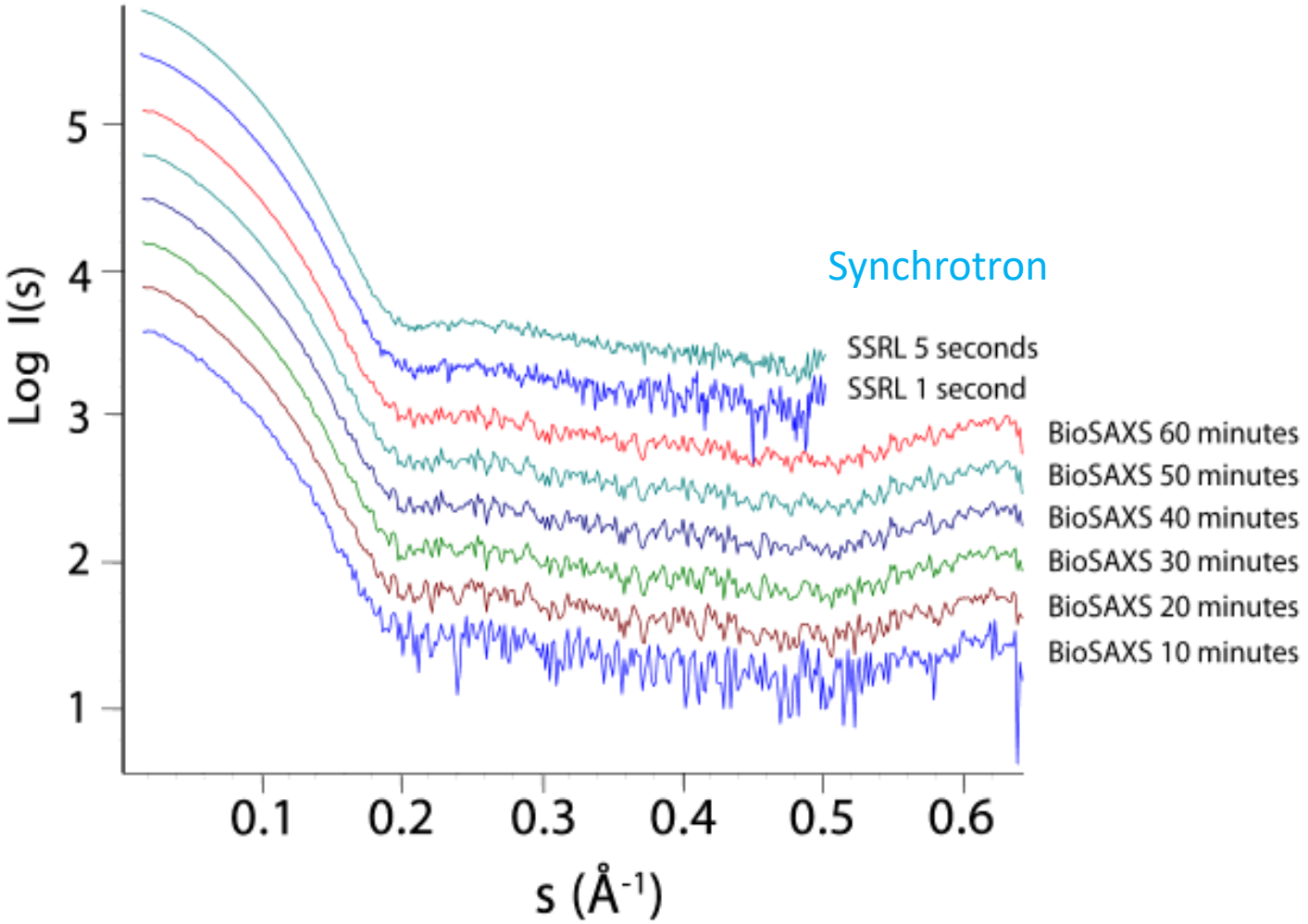


N-terminal domain SAXS and crystallographic structure

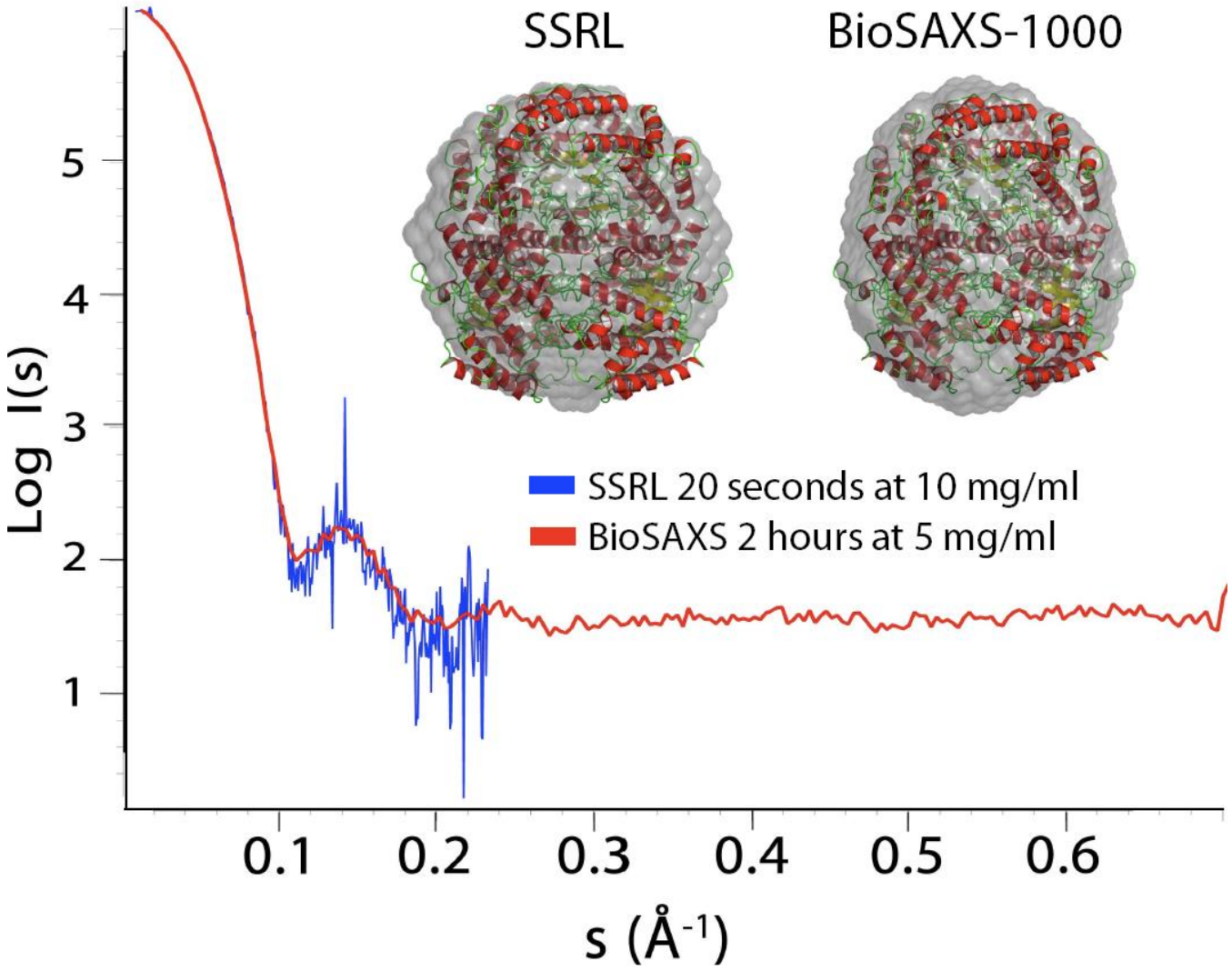

HingeProt software predicts hinge
Appears to fill space between domains
Sequence analysis shows conserved motifs for these two areas

SAXS analysis depends on shape of the curve, not intensity.

Information comes from shape and not intensity

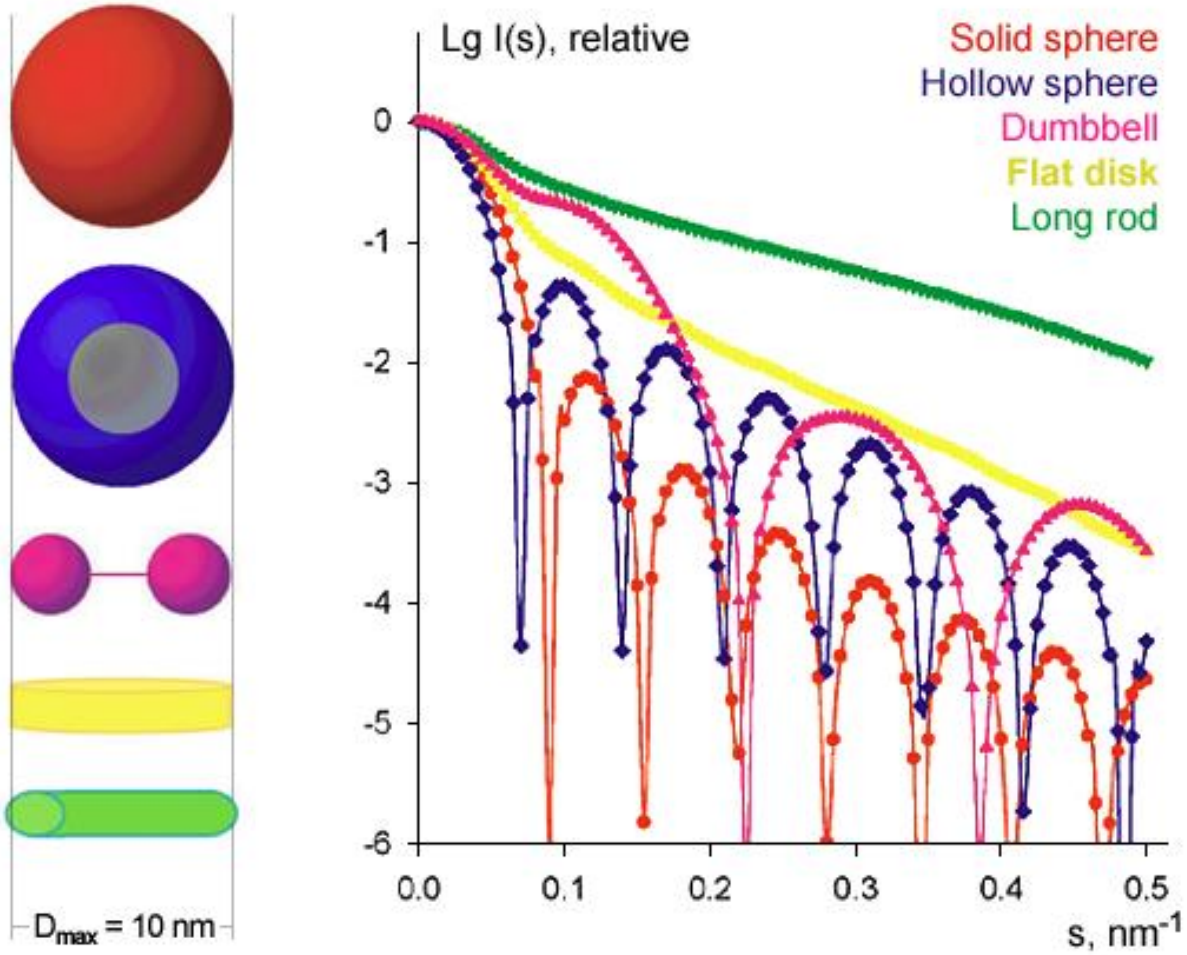


Laboratory data scaled to synchrotron



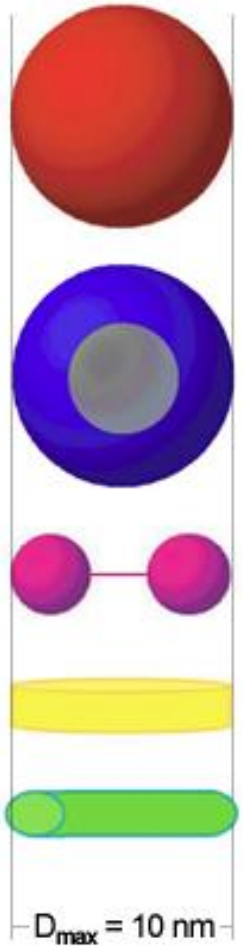
How do we interpret SAXS data?

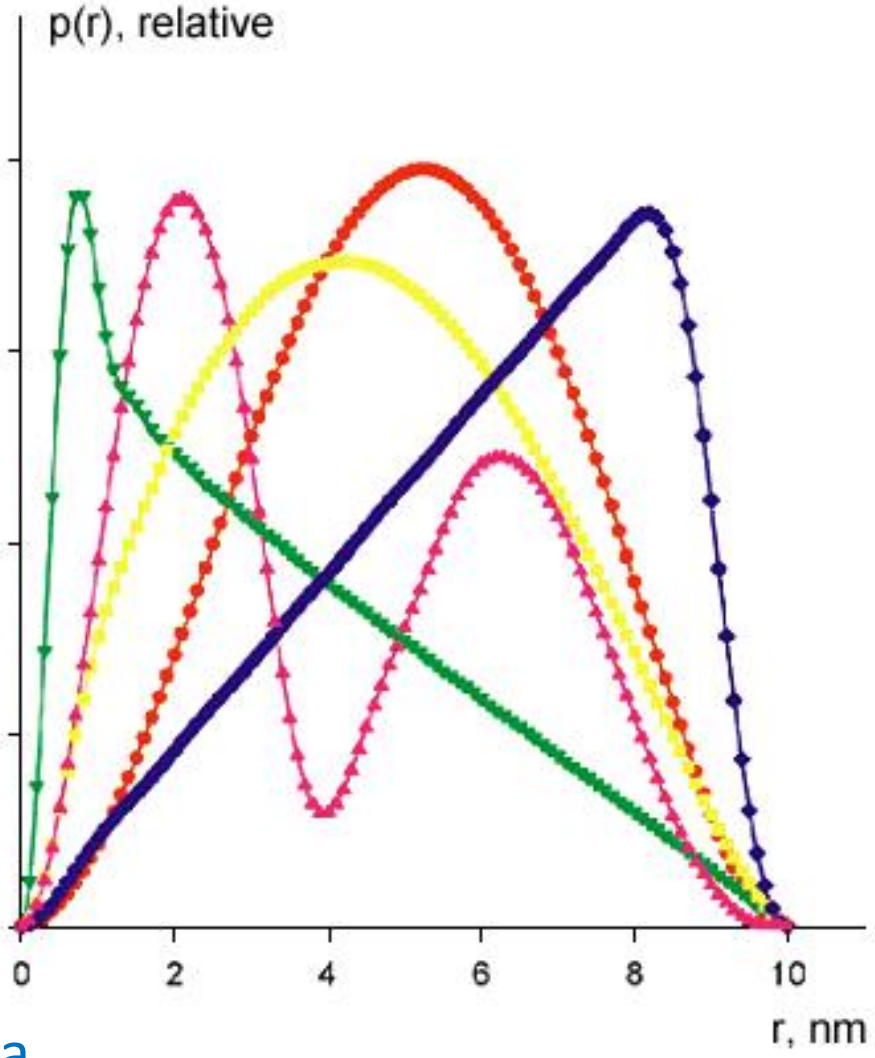
Data



From: Small-angle scattering studies of biological macromolecules in solution, Svergun and Koch, Rep. Prog. Phys., 1735-1782 (2003)

Pair distribution function

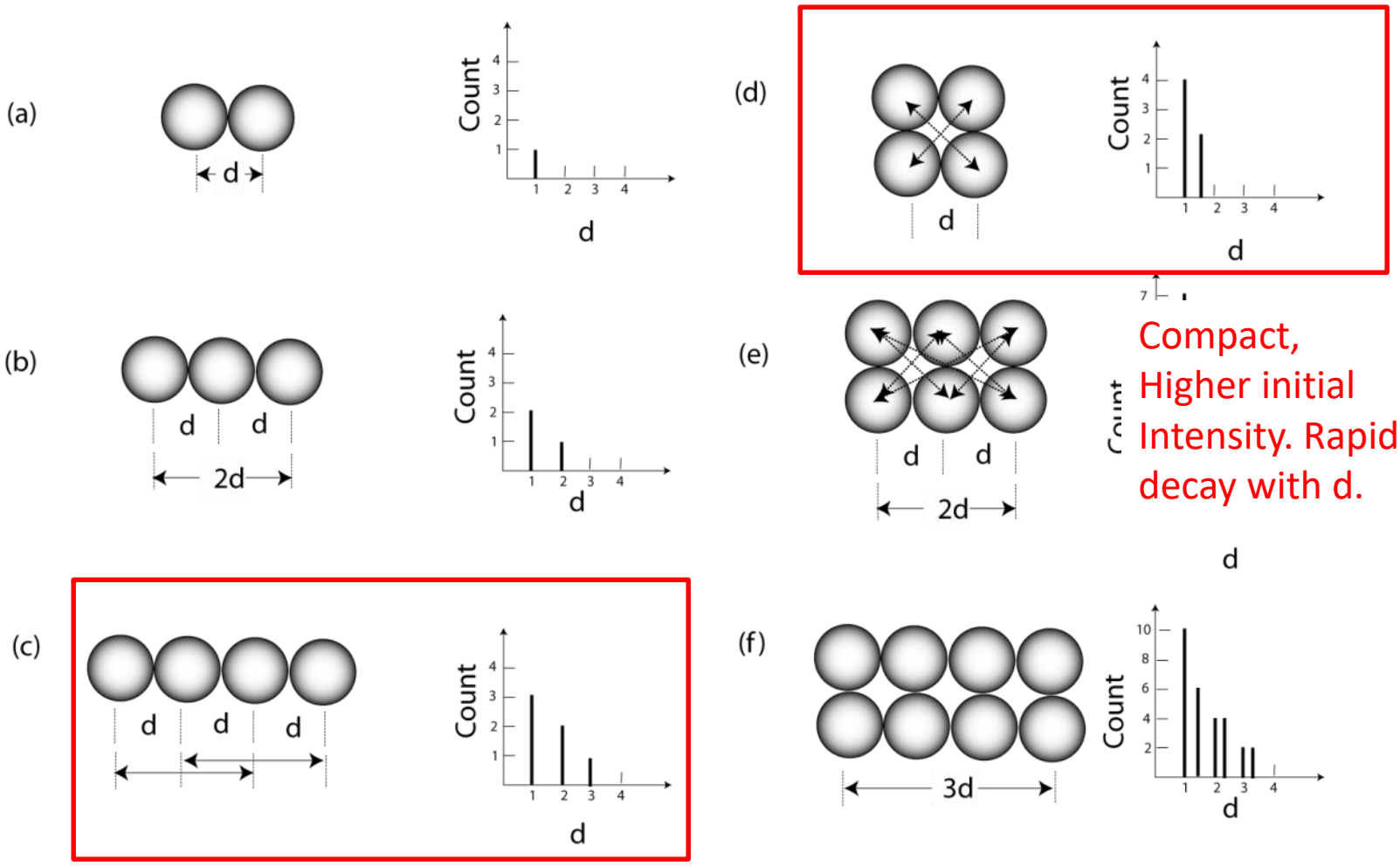




Fourier transform of data.

From: Small-angle scattering studies of biological macromolecules in solution, Svergun and Koch, Rep. Prog. Phys., 1735-1782 (2003)

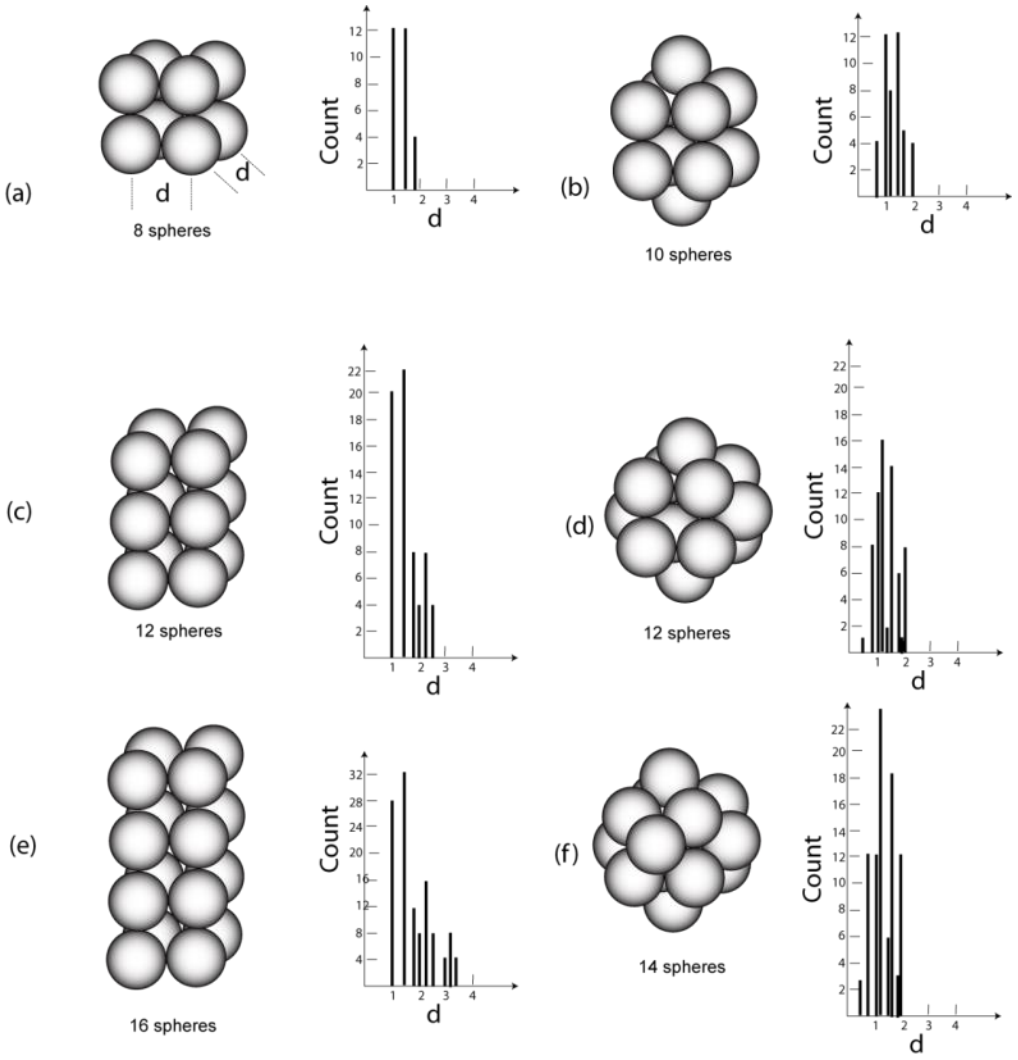
What does solution scattering give you?



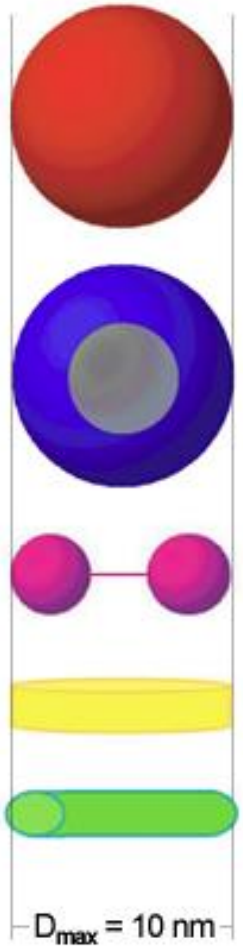
Compact, Higher initial Intensity. Rapid decay with d.

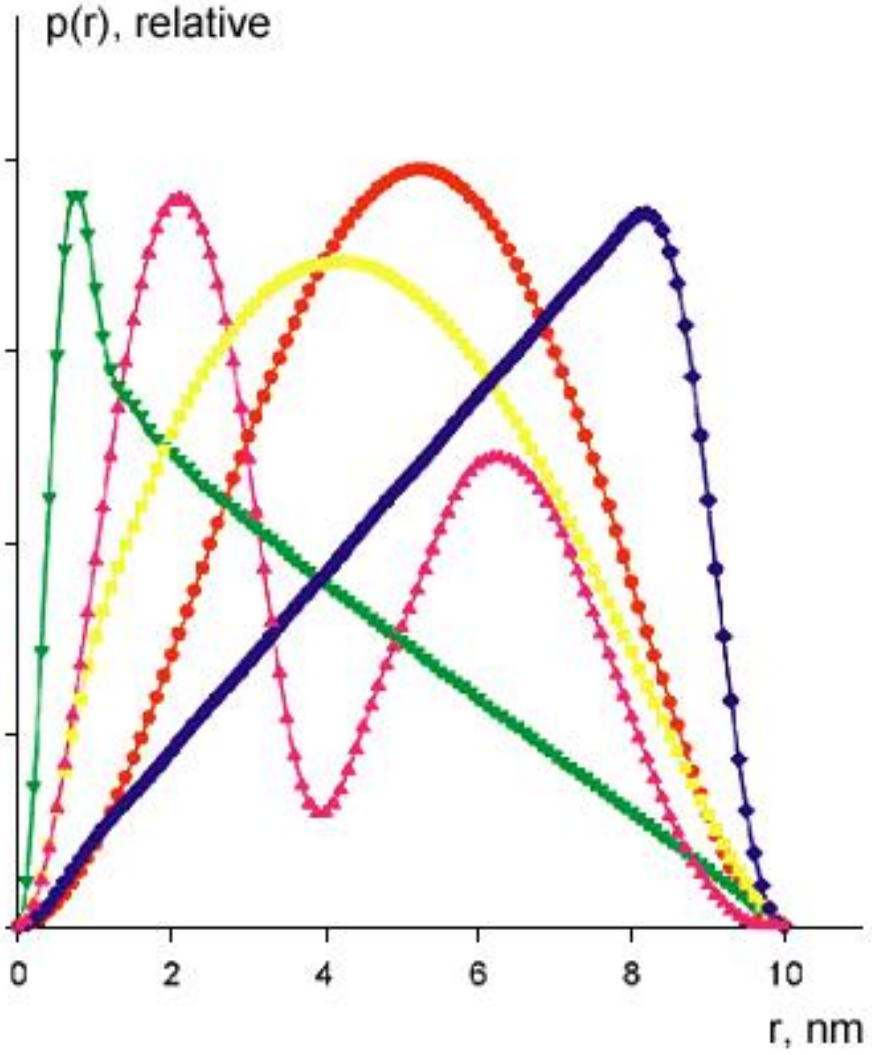
Long, slow decay with d.

Extended to three dimensions



Pair distribution function

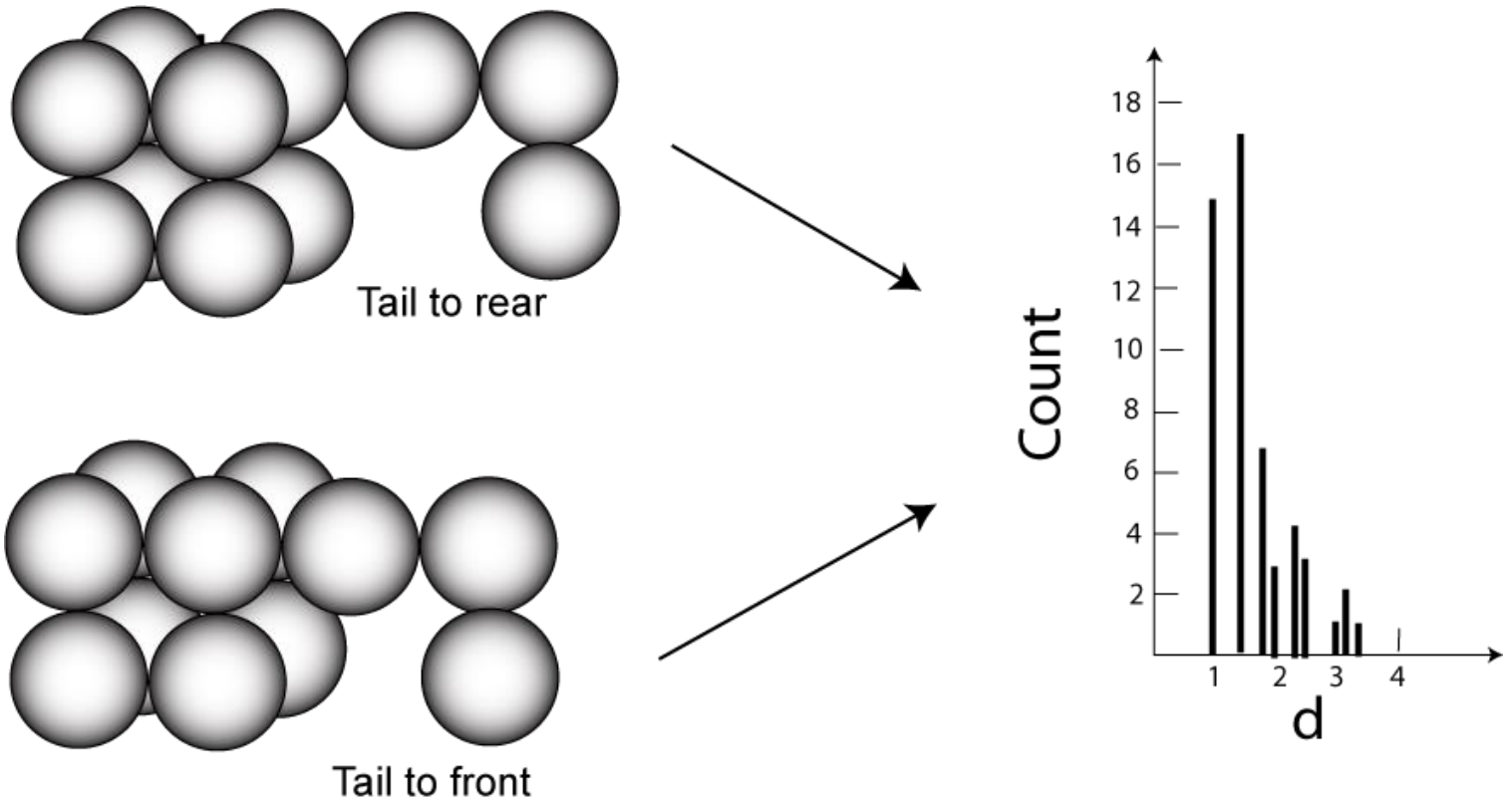




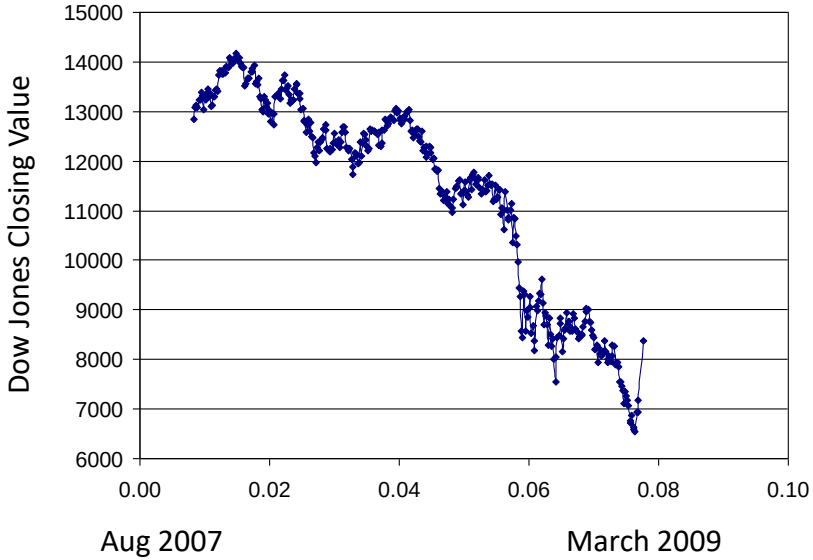
From: Small-angle scattering studies of biological macromolecules in solution, Svergun and Koch, Rep. Prog. Phys., 1735-1782 (2003)

SAXS is underdetermined

And then the problems …



Can we use X-ray solution scattering?

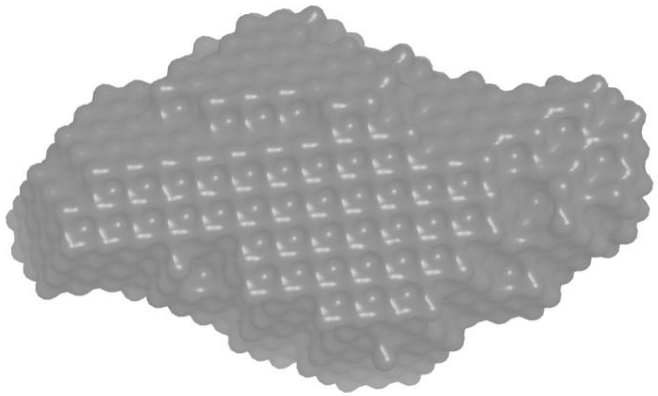


It's possible to fit multiple envelopes to the data.

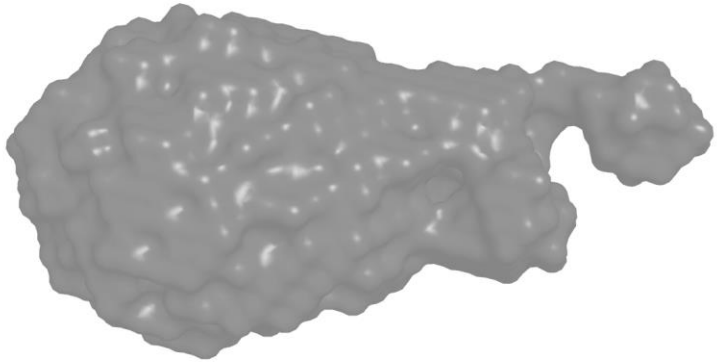
You will always get an envelope despite the data!

SAXS is complementary to crystallography

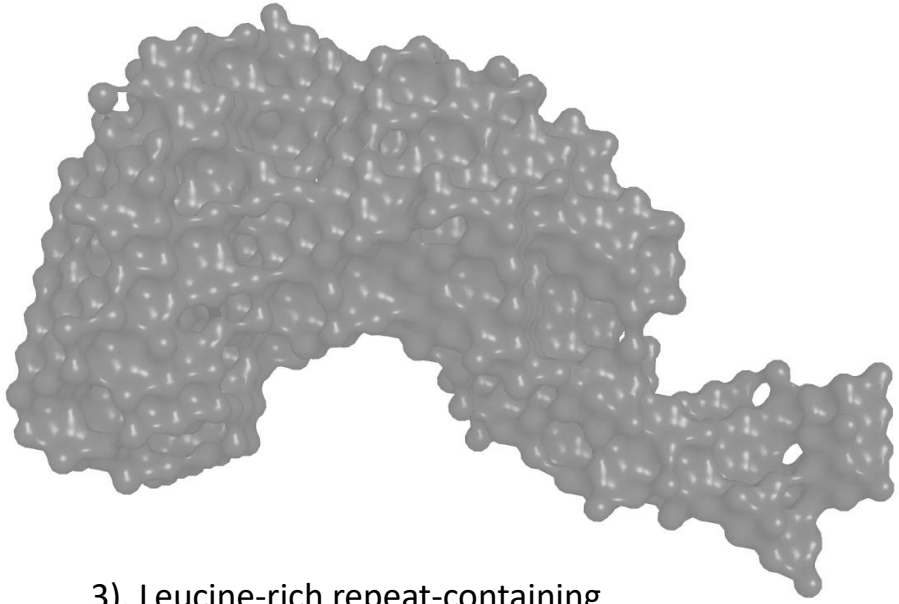
Ab intio envelopes



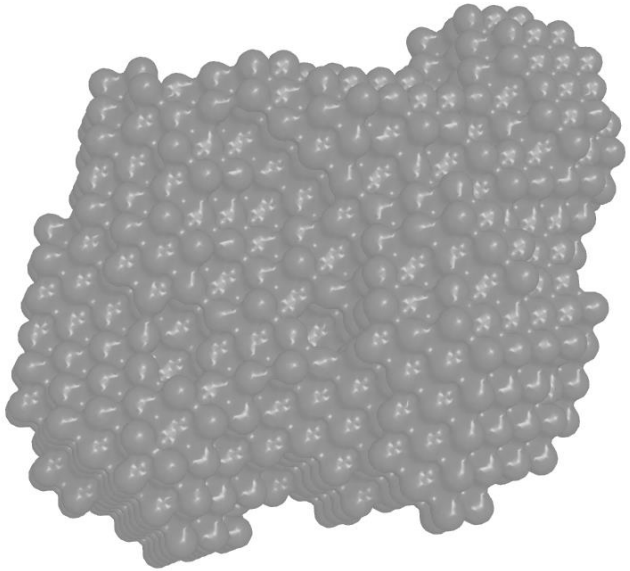
1). alr0221 protein from Nostoc (18.6 kDa)



2). C-terminal domain of a chitobiase (17.9 kDa)

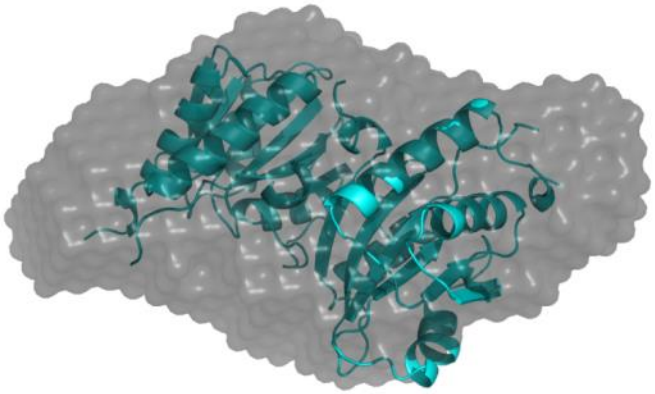


3). Leucine-rich repeat-containing protein LegL7 (39 kDa)

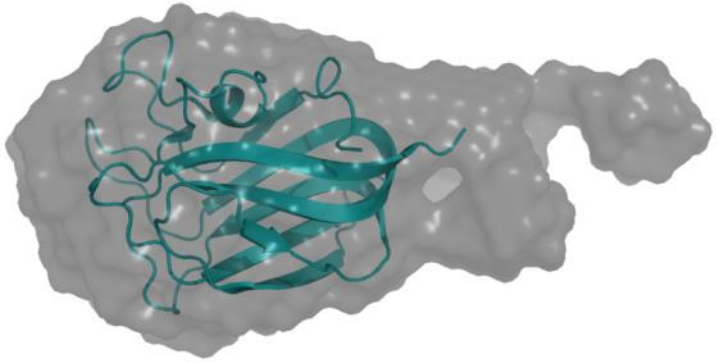


4). *E. Coli.* Cystine desulfurase activator complex (170 kDa)

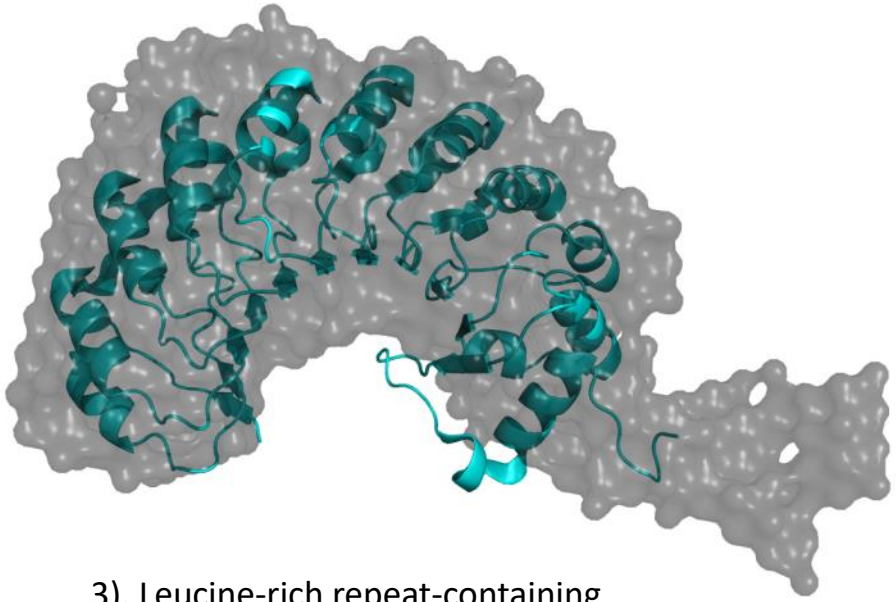
Overlaid with subsequent X-ray structures



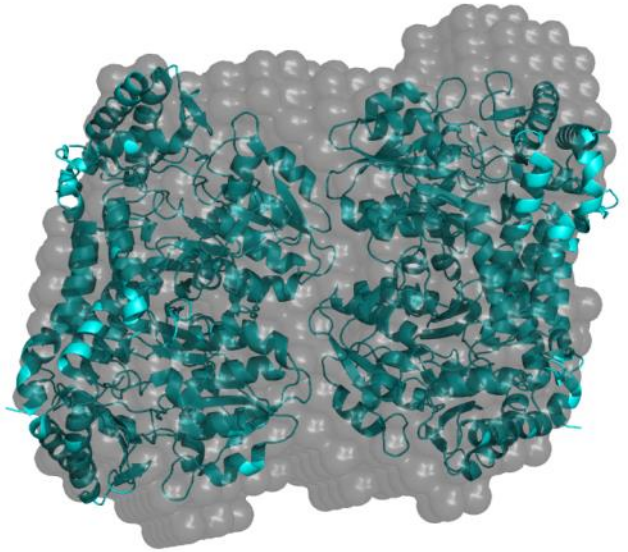
1). alr0221 protein from Nostoc (18.6 kDa)



2). C-terminal domain of a chitobiase (17.9 kDa)

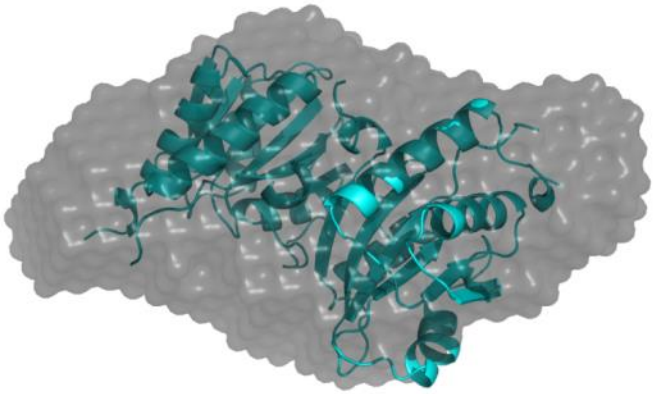


3). Leucine-rich repeat-containing protein LegL7 (39 kDa)

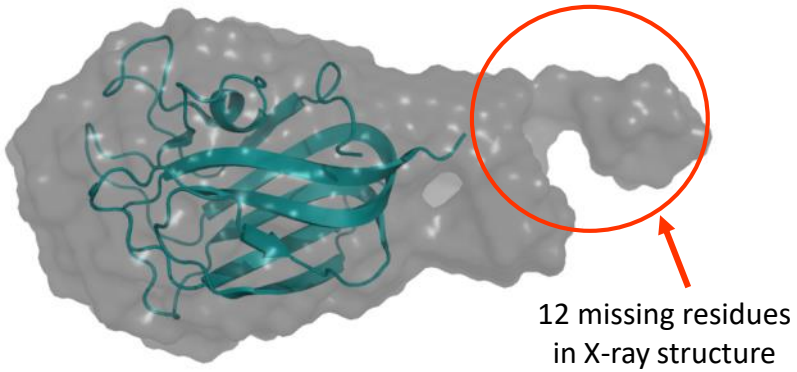


4). *E. Coli.* Cystine desulfurase activator complex (170 kDa)

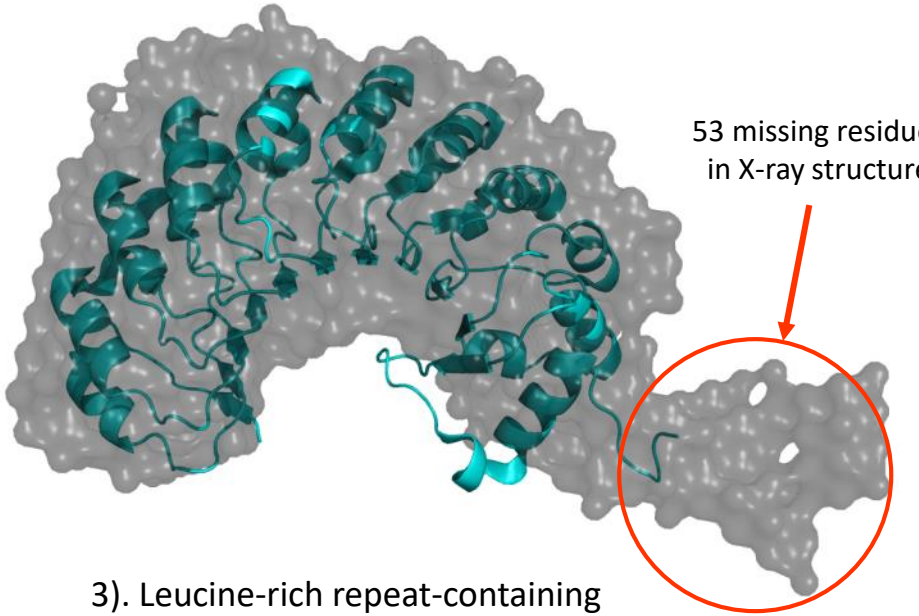
And data on what was missing …



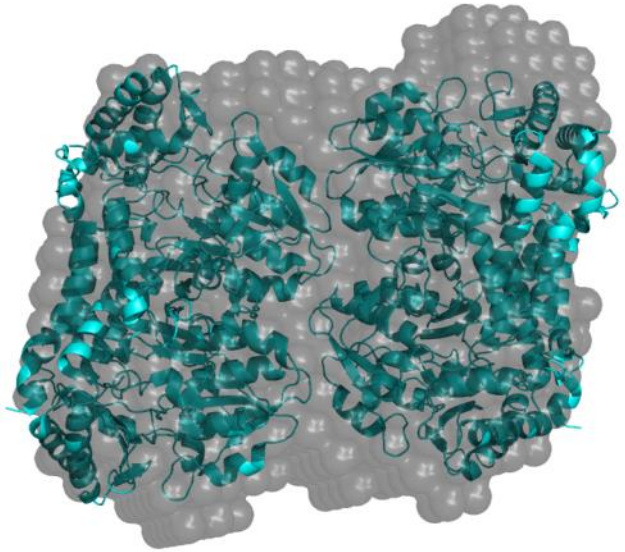
1). alr0221 protein from Nostoc (18.6 kDa)



2). C-terminal domain of a chitobiase (17.9 kDa)



3). Leucine-rich repeat-containing protein LegL7 (39 kDa)



4). *E. Coli.* Cystine desulfurase activator complex (170 kDa)

SAXS is a simple experiment
but a powerful one

It is easily interpreted but it has limitations

It is sensitive to all conformations of the molecule in solution and to residues missing in the crystal structure

#	Name	NESG ID	PDB	Ref	State	Conc	MW	Res
Samples where crystallographic structures were available								
1	Domain of Unknown Function	DhR2A	3HZ7	16	M	6.9	9523	87
2	Diguanylate cyclase with PAS/PAC sensor	MqR66C	3H9W	17	D	8.2	13611	210
3	Nmul_A1745 protein from *Nitrosospira multiformis*	NmR72	3LMF	18	T	6.9	14069	484
4	Domain of Unknown Functiion	DhR85C	3MJQ	19	D	10.7	14609	252
5	Sensory box/GGDEF family protein	SoR288B	3MFX	20	D	9.1	14779	258
6	MucBP domain of the adhesion protein PEPE_0118	PtR41A	3LYY	21	M	9.5	14300	131
7	Sensory box/GGDEF domain protein	CsR222B	3LYX	22	D	12.7	15341	248
8	HIT family hydrolase	VfR176	3I24	23	D	11.0	17089	298
9	EAL/GGDEF domain protein	McR174C	3ICL	24	M	5.0	18738	171
10	Diguanylate cyclase	MqR89A	3IGN	25	M	7.5	20256	177
11	Putative NADPH-quinone reductase	PtR24A	3HA2	26	D	9.5	20509	354
12	MmoQ (Response regulator)	McR175G	3LJX	27	M	8.8	32032	288
13	Putative uncharacterized protein	DhR18	3HXL	28	M	9.6	48519	446
Samples where multiple constructs and crystallographic structures were available								
14	Putative hydrogenase	PfR246A (78-226)	3LRX	29	D	11.4	17701	316
15		PfR246A (83-218)	3LYU	30	D	8.4	16321	284
16	Alr3790 protein	NsR437I	3HIX	31	M	5.3	11760	105
17		NsR437H	3HIX	31	M	6.5	15700	141
Samples where NMR structures were available								
18	MKL/myocardin-like protein 1	HR4547E	2KW9 (NMR)	32	D	10.4	8276	75
19	MKL/myocardin-like protein 1	HR4547E	2KVU (NMR)	33	D	10.4	8276	75
20	Putative peptidoglycan bound protein (LPXTG motif)	LmR64B	2KVZ (NMR)	34	M	5.0	9712	85
21	E3 ubiquitin-protein ligase Praja1	HR4710B	2L0B (NMR)	35	M/D	5.6	10297	91
22	Transcription factor NF-E2 45 kDa subunit	HR4653B	2KZ5 (NMR)	36	M	10.0	10623	91
23	YlbL protein	GtR34C	2KL1 (NMR)	37	M	11.0	10661	94
24	Cell surface protein	MvR254A	2L0D (NMR)	38	Tri	5.9	12385	114
25	Domain of Unknown Function	MaR143A	2KZW (NMR)	39	M	6.6	16312	145
26	N-terminal domain of protein PG_0361 from *P.gingivalis*	PgR37A	2KW7 (NMR)	40	M	12.9	17485	157
Samples where both crystallographic and NMR structures were available								
27	GTP pyrophosphokinase	CtR148A	2KO1 (NMR)	41	D	8.0	10042	176
			3IBW	42	T	8.0	10042	176
28	Lin0431 protein	LkR112	2KPP (NMR)	43	M/Hep	6.3	12747	114
			3LD7	44	M	6.3	12747	100

We have structural data for a large number of the 600 samples that we have SAXS data for (~100 structures)

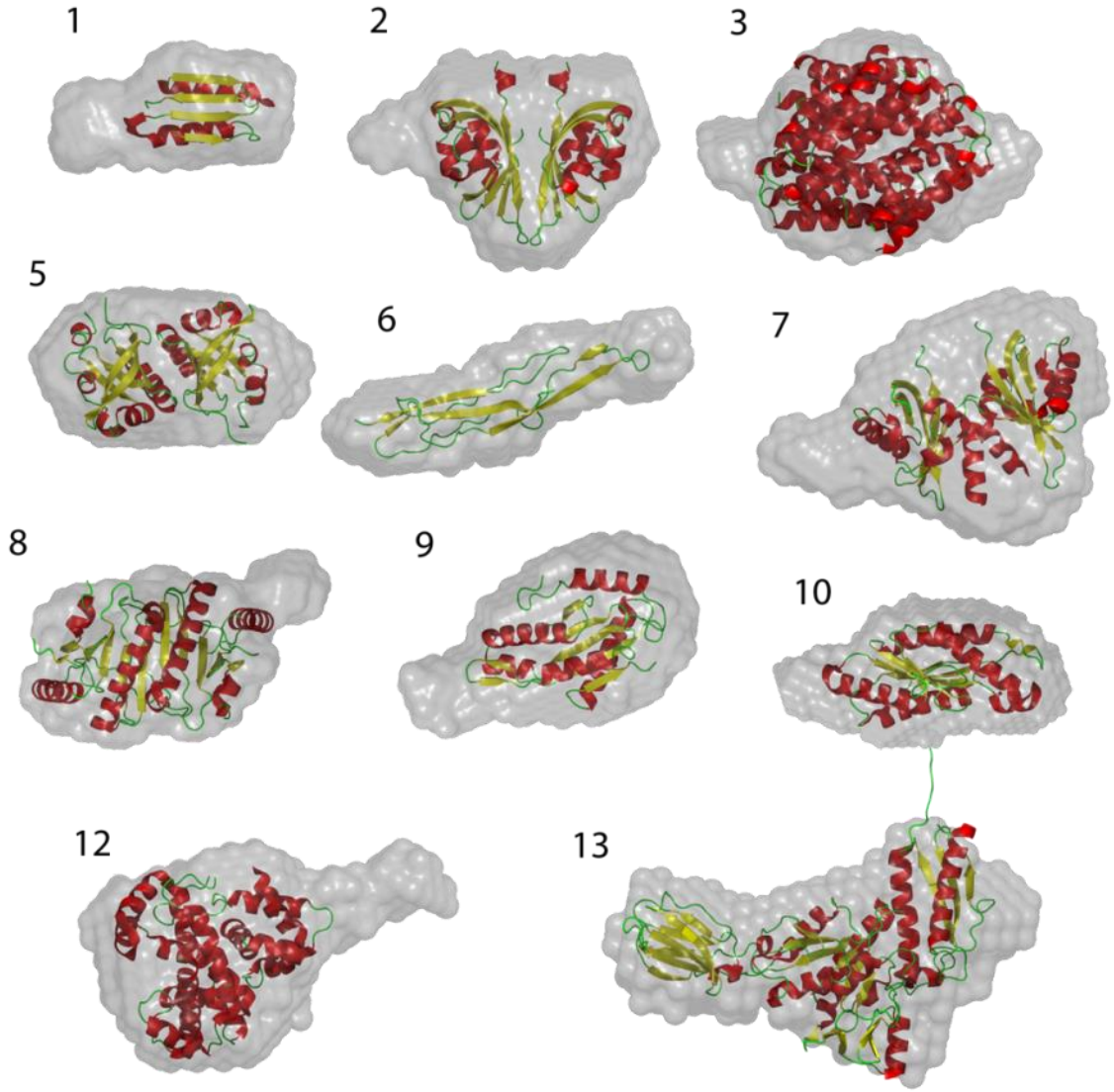
In an initial study with a subset of the 600 SAXS data sets we looked at 28 structures

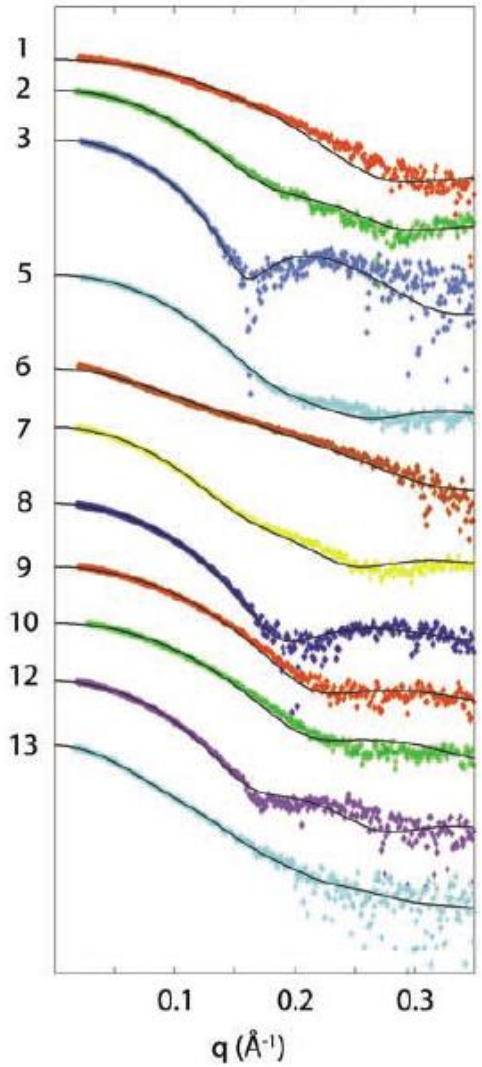
Table 1. Samples used for the SAXS analysis are divided into four sets. The first set (1-13) contains 13 proteins, each having crystallographic structures. The second set (14-17) contains 2 proteins with two different constructs of the first having two crystallographic structures and the second a single structure. The third set (18-26) contains 9 proteins, each having an NMR structure. The fourth set (27-28) contains two proteins where both NMR and crystallographic structures are available. The sample name, ID, PDB identifier, reference, the oligomeric solution state characterized on preparation by light scattering and gel filtration, initial concentration (mg/ml), molecular weight (Da) and number of residues are listed. The oligomeric solution s defined in the table as M (monomer), D (dimer), Tri (trimer), T (tetramer), Hep (Heptamer) or a combination. While all the samples have structures deposited in the PDB the majority are as yet unpublished. We are grateful to the authors in the references for the ability to use this structural data at this early stage.

#	Residues observed	# Res missing	Rg structure	Dmax structure	Rg SAXS	ΔRg	Dmax SAXS	Δ dmax	Porod MW	MW Ratio	SAXS oligomer[1]	Oligomer Assign.	SAXS fit (χ)
Samples where crystallographic structures were available													
1	74	13	13.7	42.0	14.9	1.2	53.2	11.2	7827	0.8	M		4.2
2	198	12	16.6	67.0	19.8	3.2	67.4	0.4	24555	1.8	D	sym	2.6
3	436	48	22.4	62.3	23.2	0.8	75.3	13.0	50064	3.6	T	sym	1.6
4	214	38	23.3	81.2	23.6	0.3	82.7	1.5	37348	2.6	D/T*	PDB	2.6
5	224	34	19.9	57.6	19.8	-0.1	64.2	6.6	28828	2.0	D	PDB	2.2
6	107	24	19.6	76.3	21.5	1.9	82.0	5.7	11085	0.8	M		6.1
7	236	12	21.4	64.7	22.2	0.8	76.8	12.1	31410	2.0	D	PDB	3.8
8	286	12	20.5	63.1	21.1	0.6	71.4	8.3	34786	2.0	D	PDB	2.0
9	162	9	17.6	54.0	18.7	1.1	65.5	11.5	20468	1.1	M		3.7
10	165	12	17.5	58.0	18.5	1.0	65.8	7.8	19069	0.9	M		4.2
11	336	18	26.1	80.8	26.0	-0.1	89.7	8.9	59937	2.9	D/T*	PDB/sym	1.4
12	252	36	21.3	61.5	22.5	1.2	81.9	20.4	37254	1.2	M		2.9
13	416	30	28.5	95.0	27.6	-0.9	98.5	3.5	40027	0.8	M		1.4
Samples where multiple constructs and crystallographic structures were available													
14	272	44	20.8	59.6	21.1	0.3	69.2	9.6	30670	1.9	D	PDB	1.9
15	258	26	21.1	61.8	22.0	0.9	79.7	17.9	32657	2.0	D	PDB	1.8
16	93	12	18.0	59.5	18.2	0.2	64.7	5.2	15875	1.3	D2	PDB	1.7
17	93	48	20.4	75.0	20.8	0.4	73.0	-2.0	15920	1.0	D1	PDB	2.5
Samples where NMR structures were available													
18	75	0	22.5	122.4	16.8	-0.9	58.4	-64.0	6771	0.8	M		4.7
19	75	0	17.7	94.4	16.5	-1.2	58.4	-36.0	6771	0.8	M		1.4
20	85	0	19.0	80.8	18.7	-0.3	68.0	-12.8	9724	1.0	M		1.7
21	91	0	16.4	71.0	15.9	-0.5	59.6	-11.4	7862	0.8	M		1.5
22	91	0	22.3	123.1	19.6	-2.7	68.0	-55.1	10762	1.0	M		1.6
23	87	7	14.3	55.8	14.5	0.2	49.7	-6.1	8479	0.8	M		1.4
24	114	0	16.5	67.8	19.6	3.1	66.6	-1.2	12609	1.0	M		5.9
25	145	0	49.0	325.5	26.6	-22.4	94.7	-230.8	15386	0.9	M		7.4
26	157	0	19.8	67.5	17.5	-2.3	60.6	-6.9	15238	0.9	M		2.1
Samples where both crystallographic and NMR structures were available													
27*									22589	2.2	D		
	158	18	18.1	52.5	19.0	0.9	68.3	15.8				PDB	2.4
28*	114	0	18.5	104.4	18.5	0.0	68.2	-36.2	10721	0.8	M		2.3
	87	13	14.8	44.1	18.4	3.6	68.2	24.1					7.4

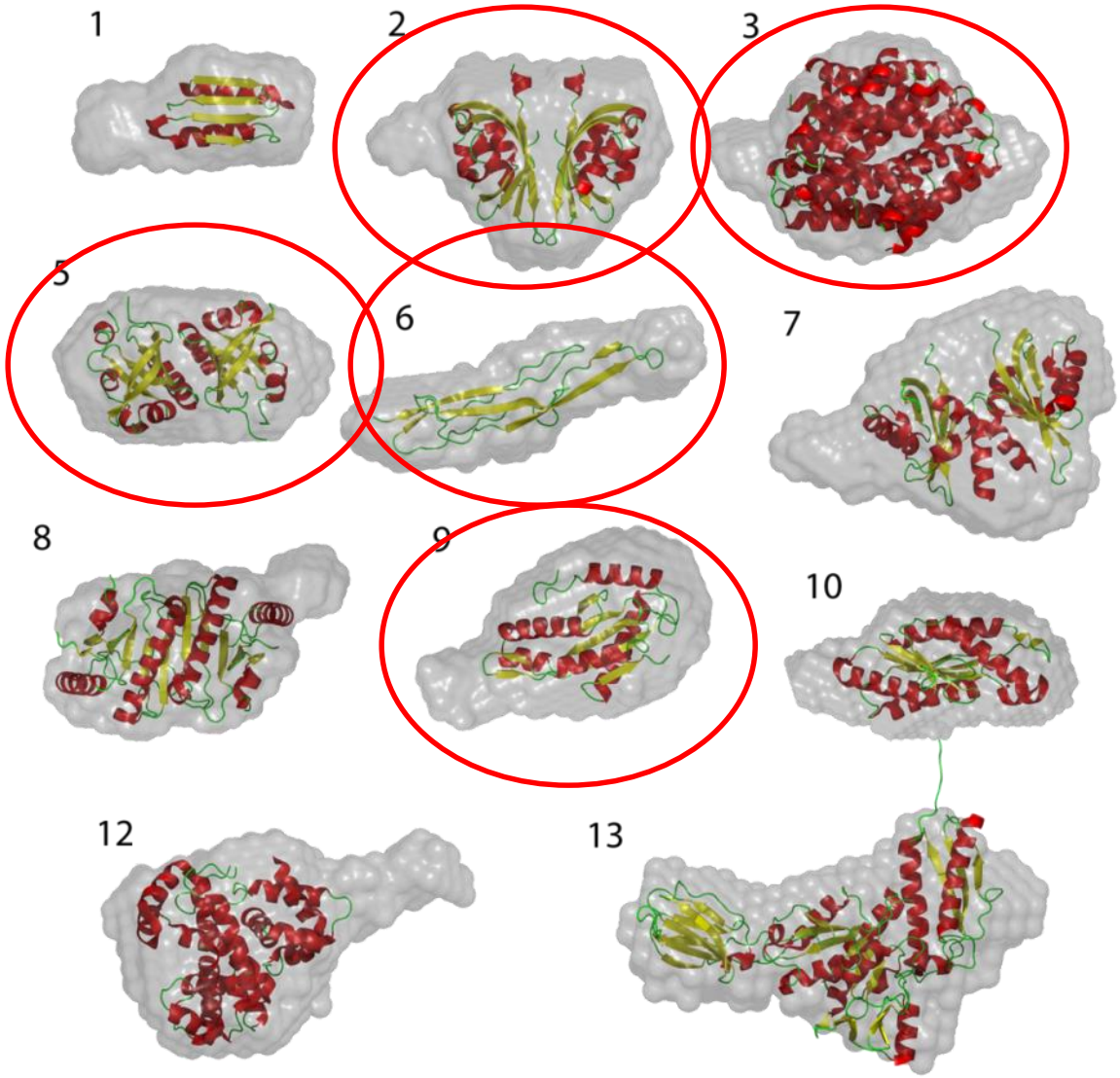
Table 2. A summary of structural (crystallography and NMR) and SAXS results. The sample # refers to the identical number in Table 1. The number of unresolved residues in the structure (mainly crystallographic) is listed together with the R_g and D_{max} (in Å) determined from the available structure. The R_g and D_{max} from the SAXS data are shown together with the difference from the available structural information. The molecular weight (in Da) calculated from a Porod analysis is listed along with the ratio of this weight with that derived initially from mass spectrometry in table 1. Finally the SAXS determined oligomer, (Monomer, Dimer or Tetramer), the relationship to the available structure and the χ of the fit are listed. A special case is described below for samples 16 and 17. Further details are given in the text.

Comparing X-ray structures





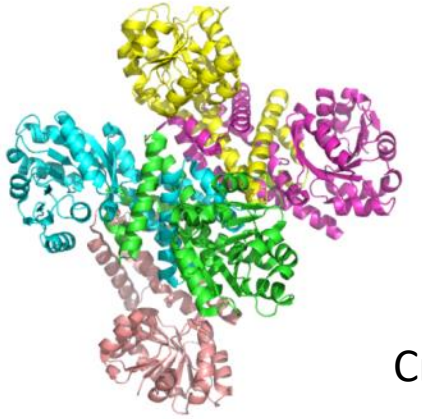
Comparing X-ray structures



Solution oligomer different than that suggested by biological unit in the PDB.

SAXS has added to the structural knowledge.

What is biologically correct, crystal or solution?

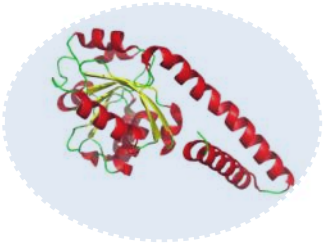


2izz from the PDB
(5 chains in PDB)

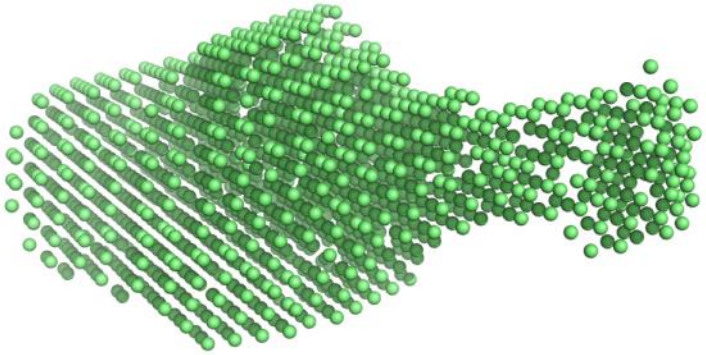
Crystal packing artifact

Another story

3gt0 from the PDB

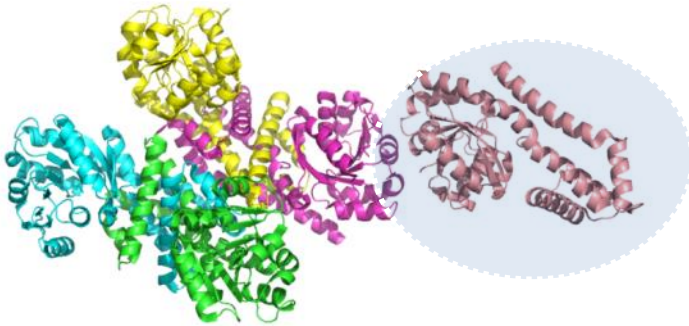


Solution envelope from BcR38B-21.20-SeMa-Gf (3gt0)



~165A

Correct position for
5th chain

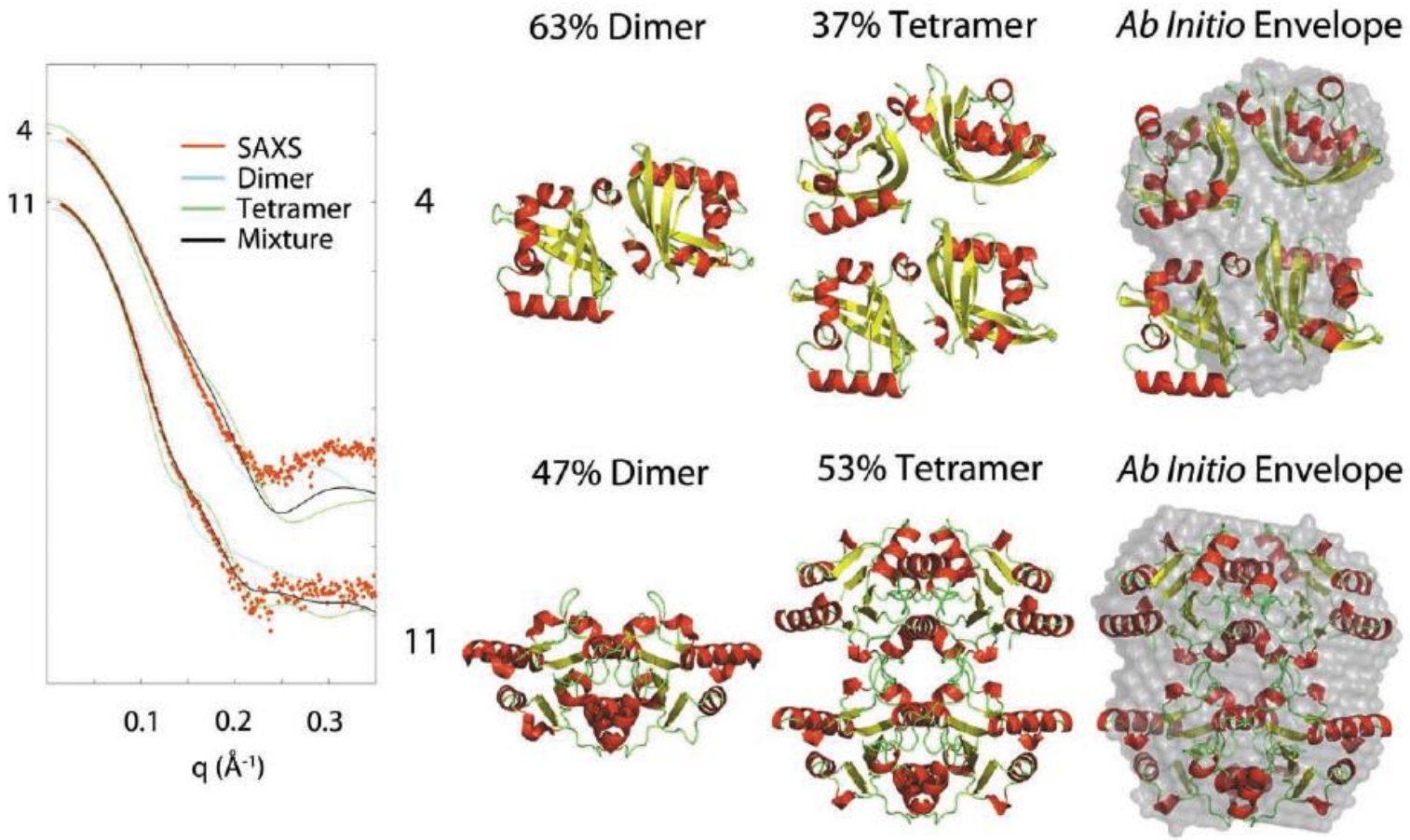


~165A

Biological unit based
on 2izz and SAXS

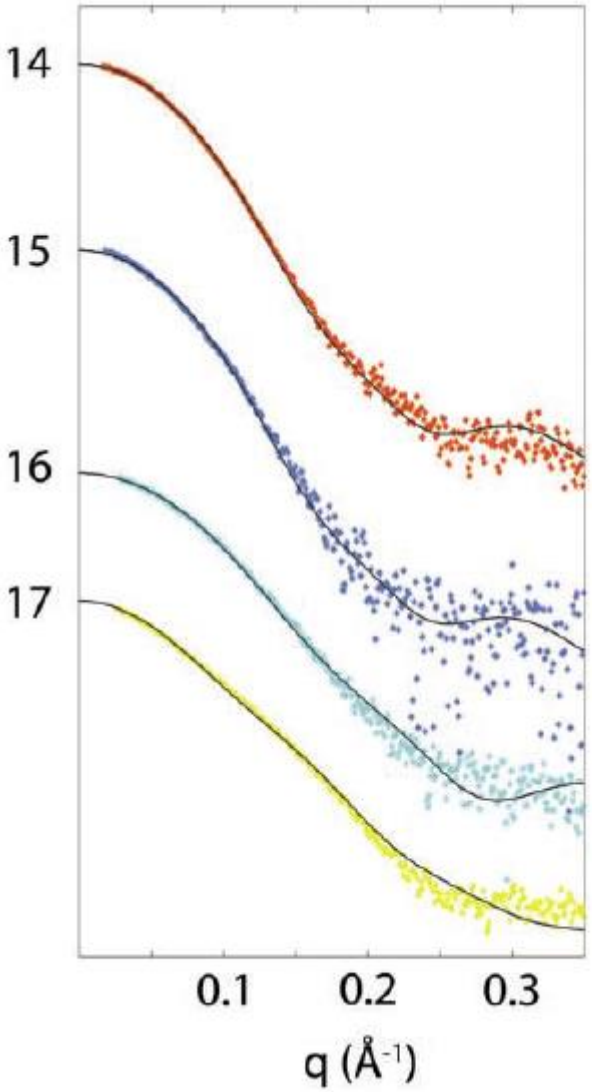
SAXS can identify a solution oligomer that may be different from the crystallographic one.

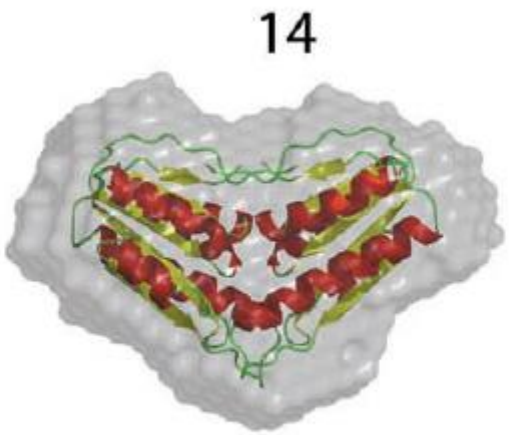
Identification of mixtures : If you know the structure you can identify an oligomer mixture

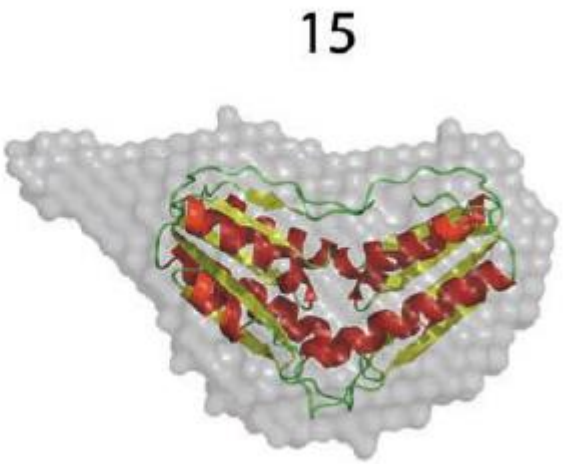


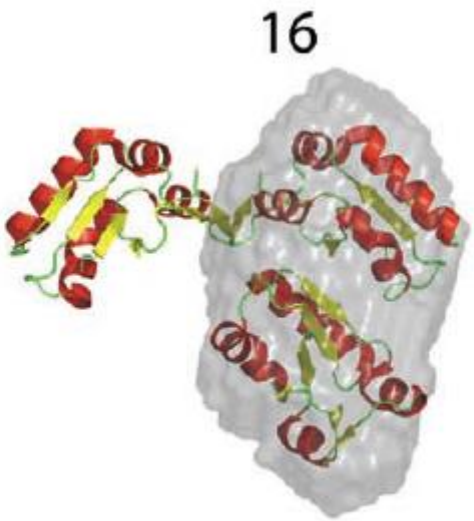
Analyzed using the program Oligomer: http://www.embl-hamburg.de/biosaxs/manual_oligomer.html

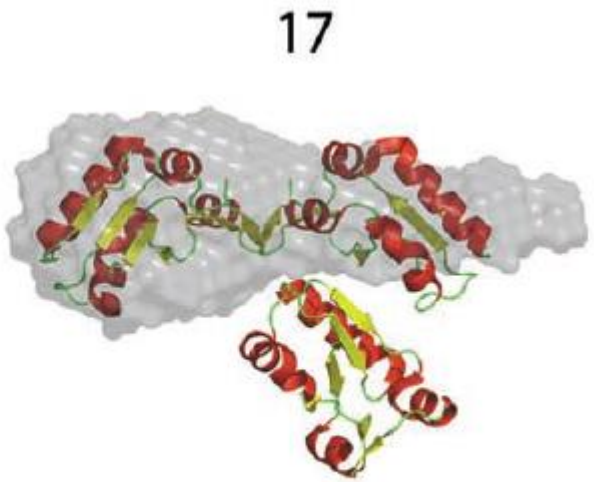
Alternative constructs





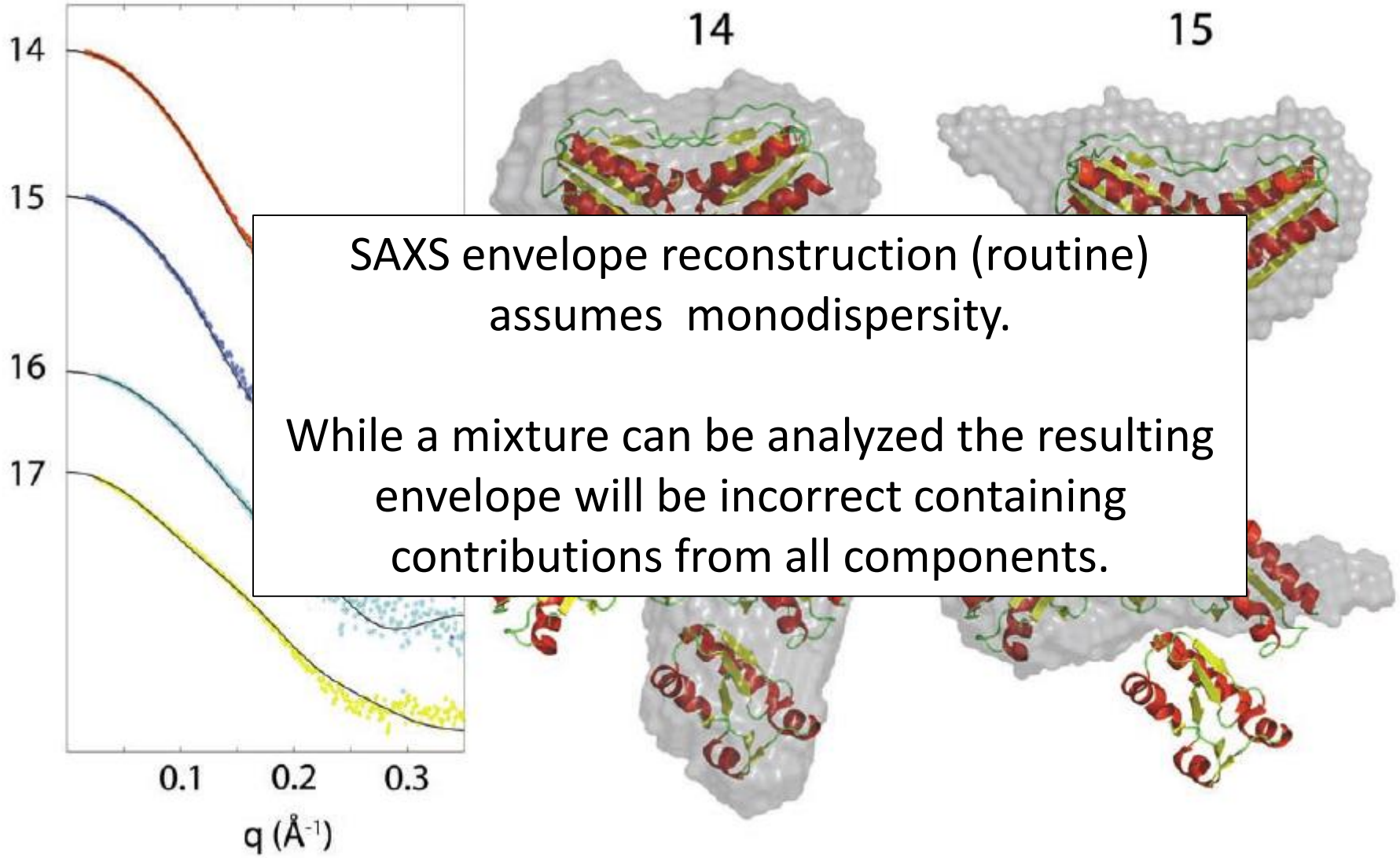






Absolute chi's depend on error model. Relative chi's can distinguish right from wrong.

Alternative constructs



SAXS envelope reconstruction (routine) assumes monodispersity.

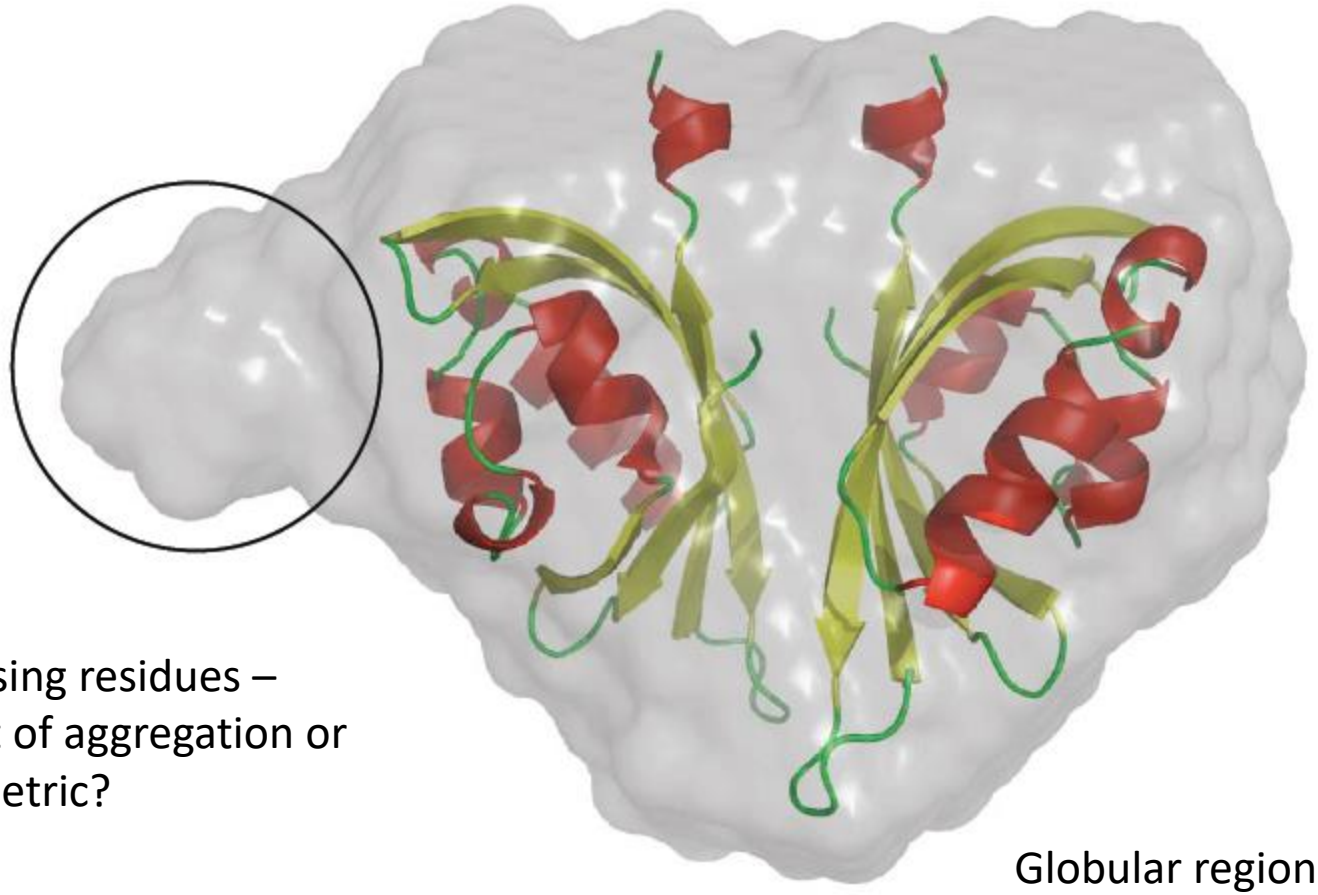
While a mixture can be analyzed the resulting envelope will be incorrect containing contributions from all components.

Absolute chi's depend on error model. Relative chi's can distinguish right from wrong.

SAXS is a simple experiment
but a powerful one

It is easily interpreted but it has
limitations

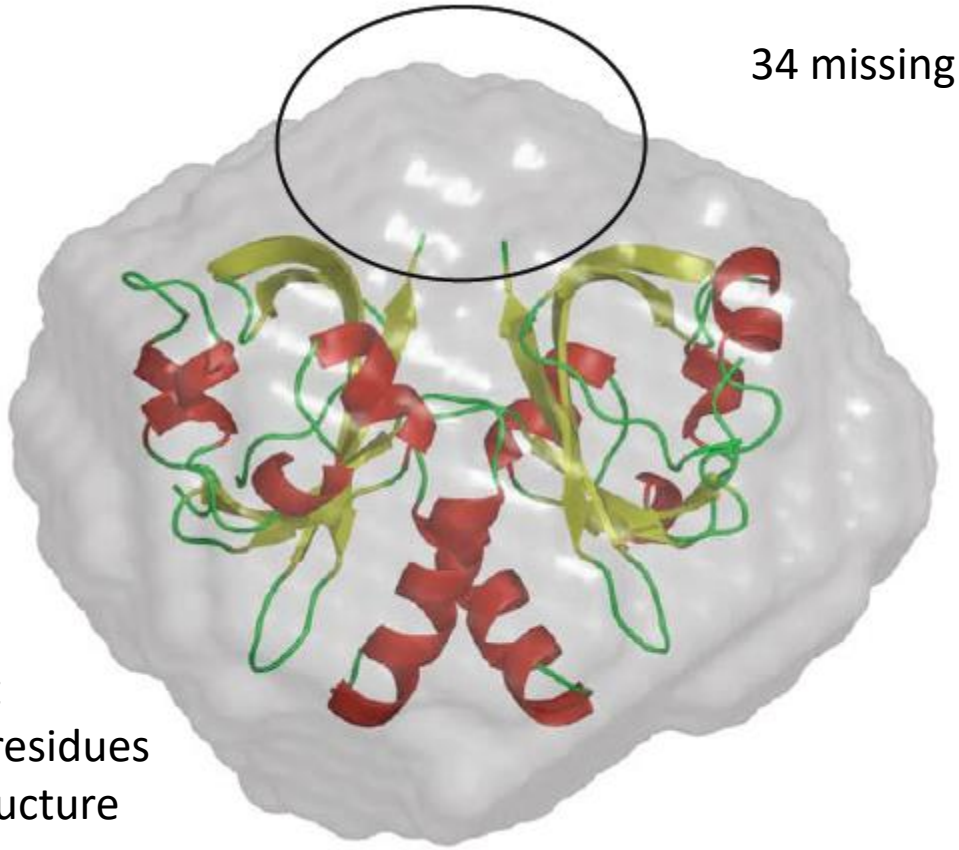
Diguanylate cyclase



12 missing residues – artifact of aggregation or asymmetric?

Globular region fits well

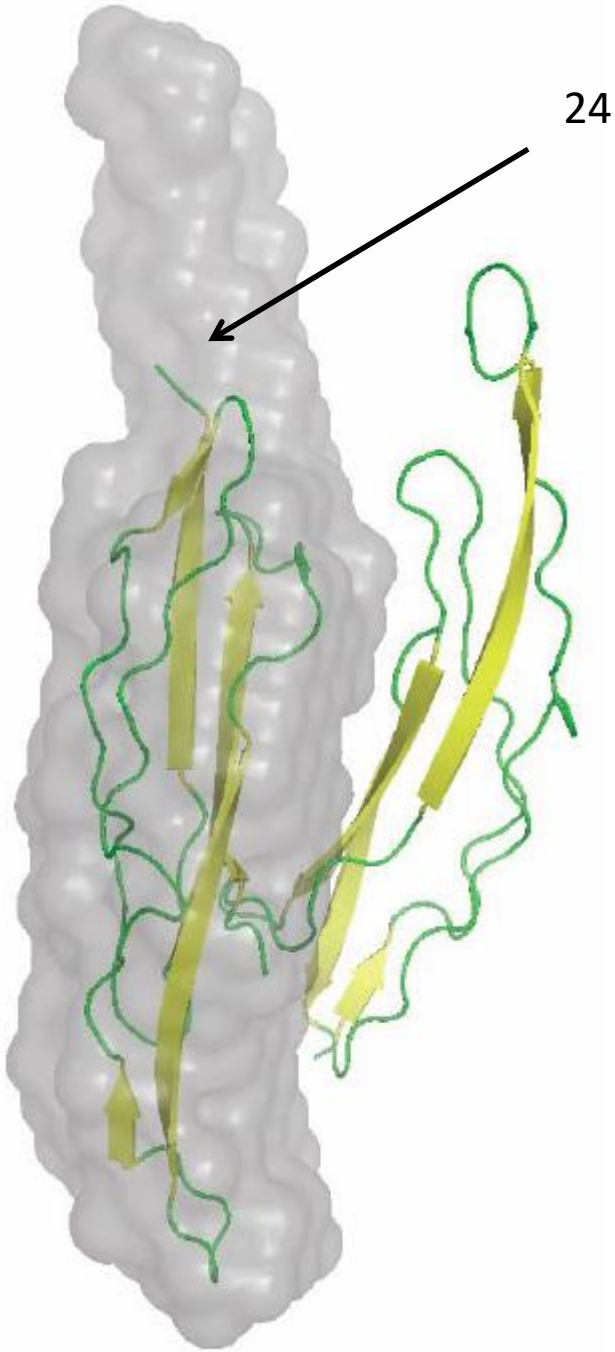
Sensory Box/GGDEF Protein Family



34 missing residues

When a significant percentage of the residues are missing in a structure positioning within an envelope may be ambiguous – *a potato is a potato*.

MucBP Domain of PEPE_0118



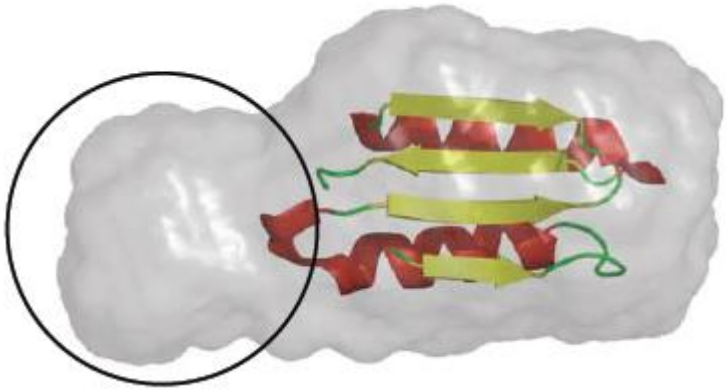
Biological unit was thought to be a dimer from crystallography.

Solution state is not.

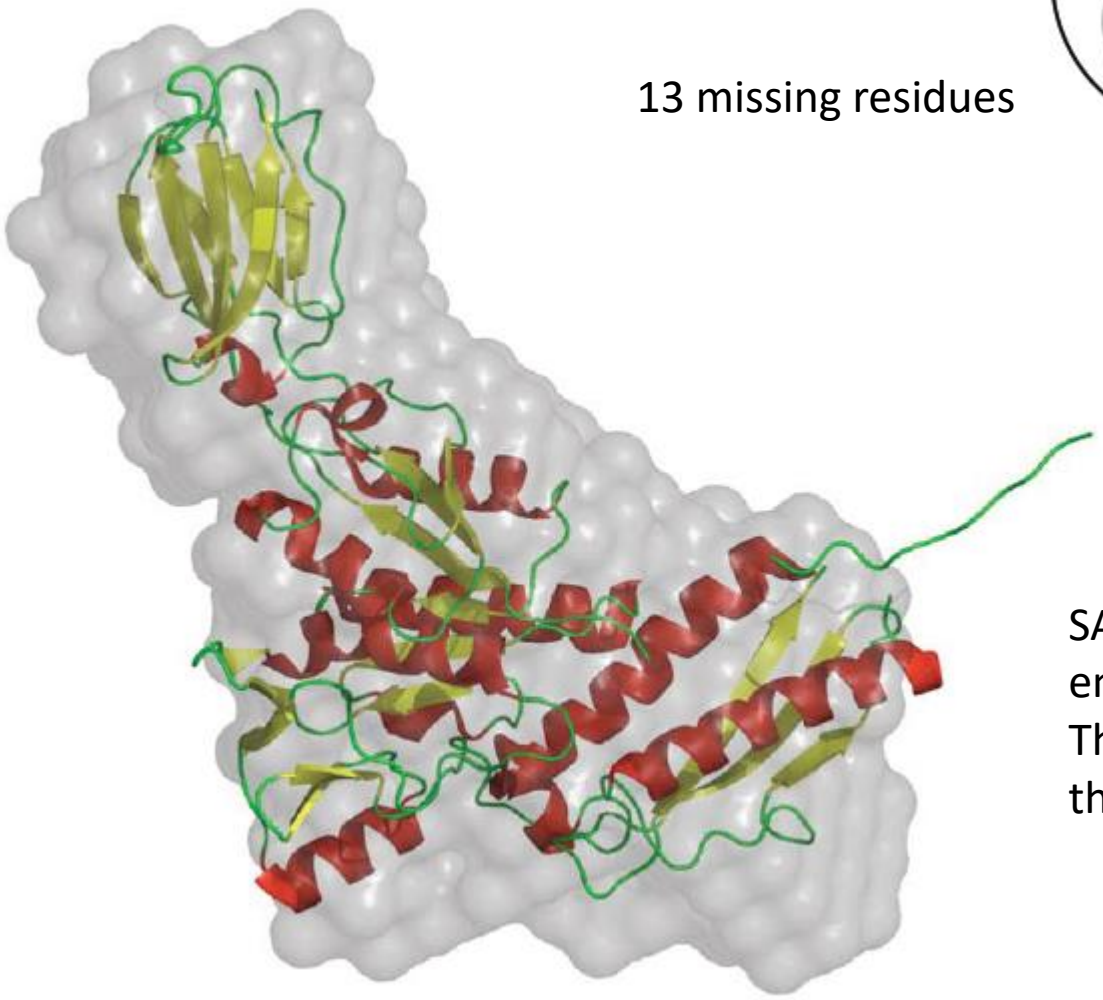
The biological state is not necessarily the solution or crystallographic state.

In this case the asymmetry allowed fitting

Size matters



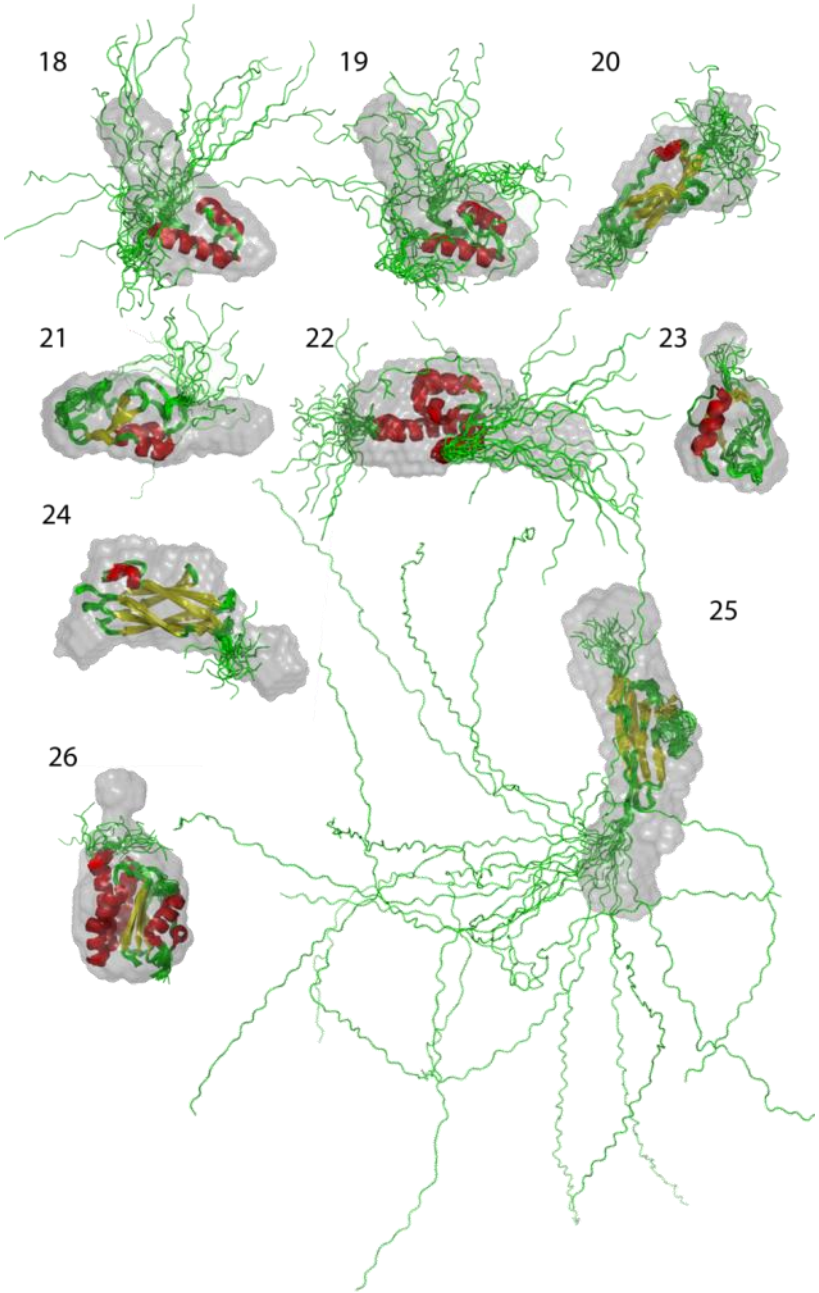
13 missing residues

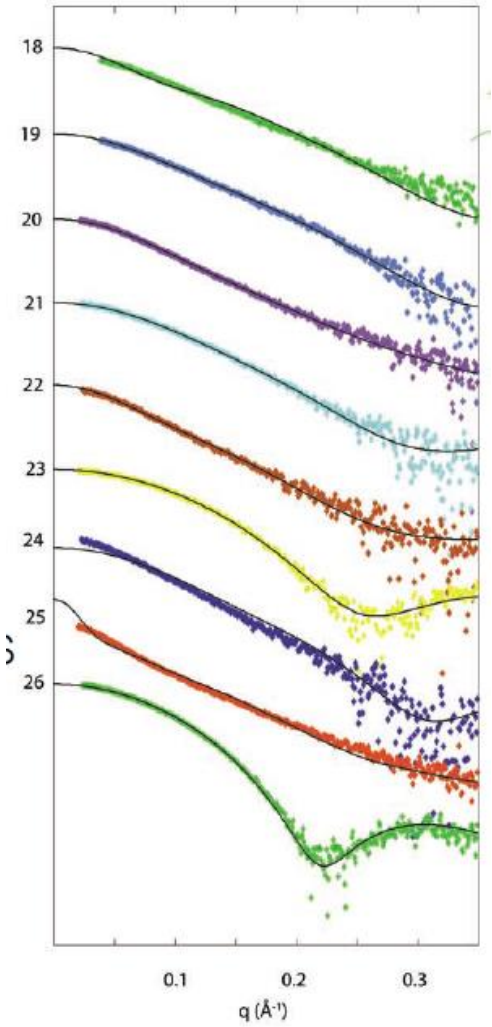


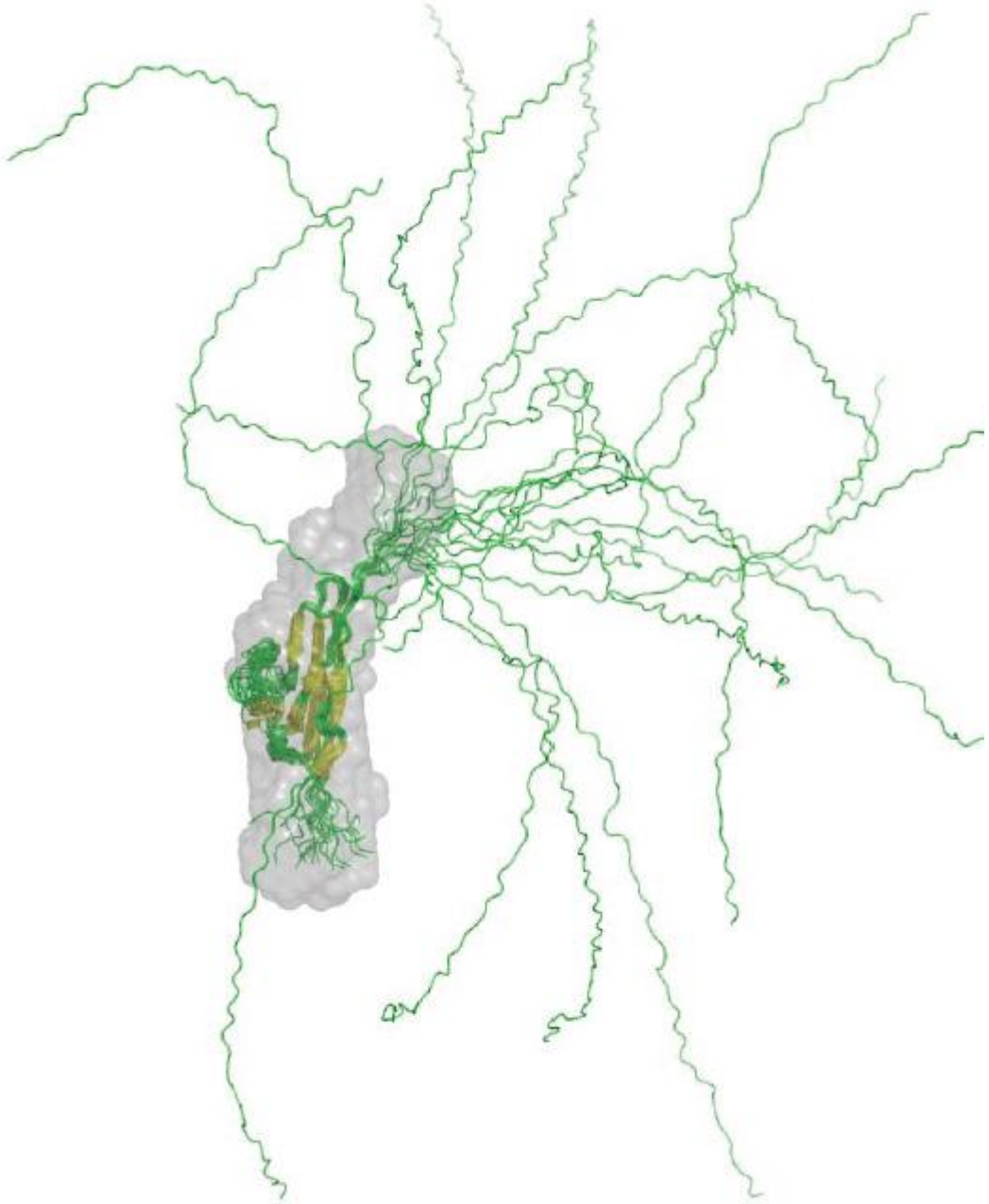
SAXS is not just about shape of the envelope but also it's overall size. The envelope produced reflects the size of the sample.

SAXS is complementary to NMR

Comparing NMR structures







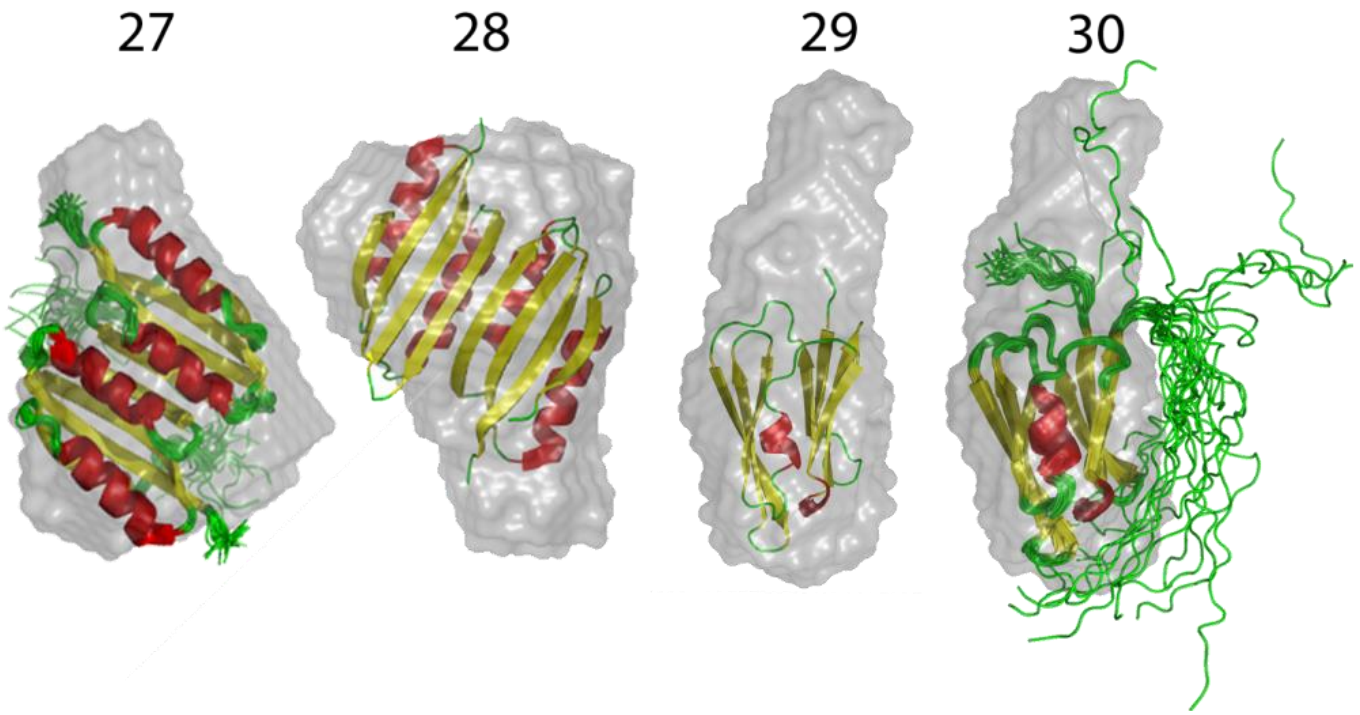
Protein of Unknown Function

"Core" domain seems to be in agreement, but disorded region highly incompatible.

Which is correct, NMR or SAXS?

SAXS is complementary to Crystallography and NMR

Comparing NMR and X-ray structures



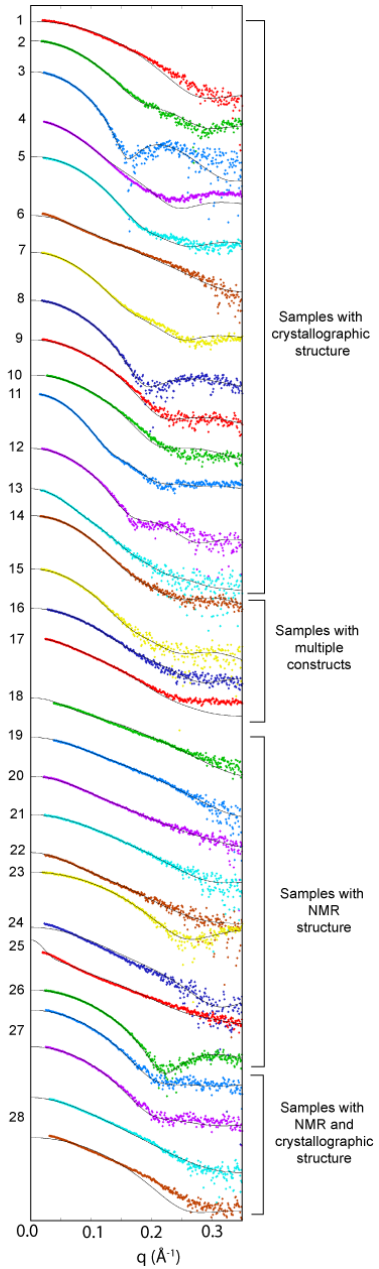
How robust is it?

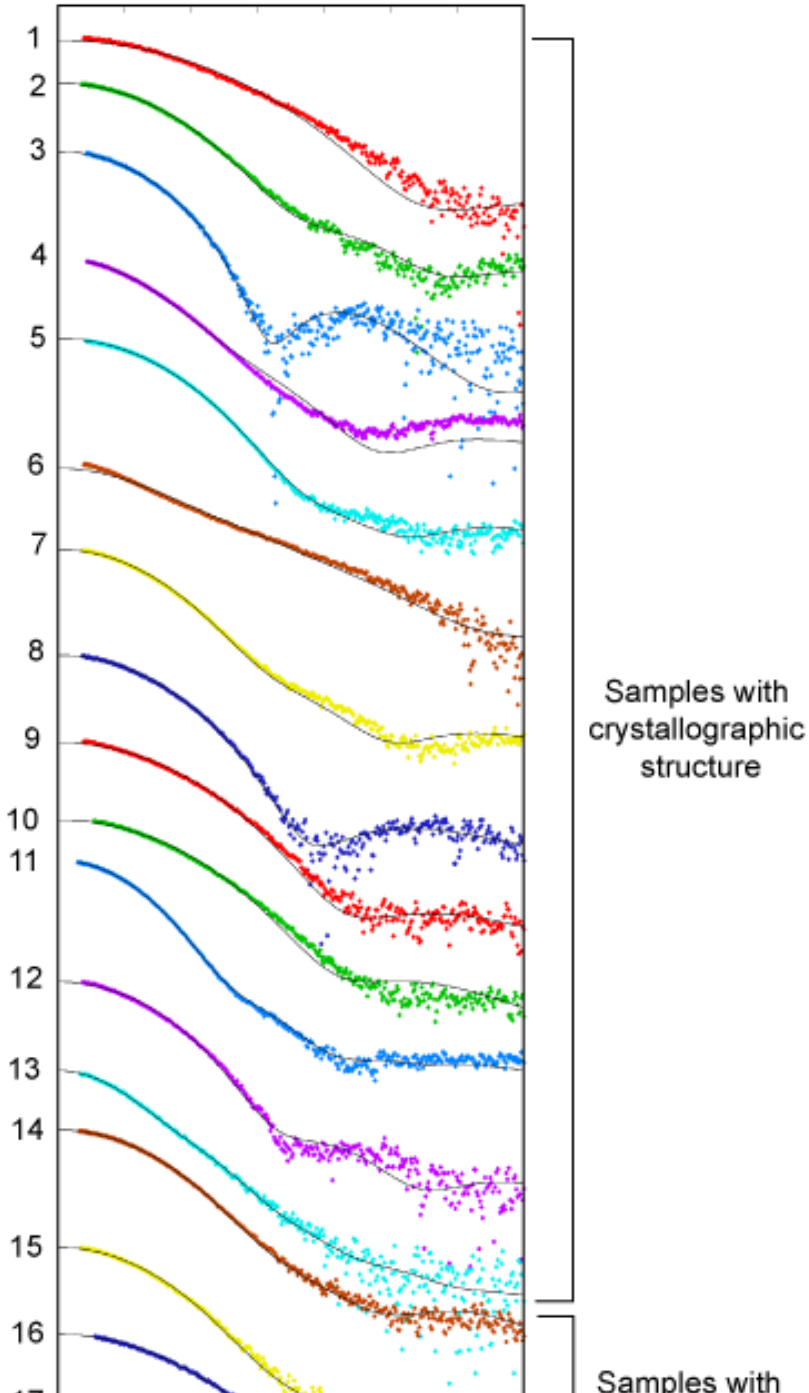
Your answer will depend on your age and experience.

Key developments:

Area	Detail	Year
Algorithm development.	Spherical harmonics/Monte Carlo	1970's/1990's
Synchrotron sources.	Second and third generation	1980's
Low noise rapid detectors.	CCD's, Pixel Arrays	1990's
Computational power	Machines and software	2000's

SAXS today benefits from each of these developments.


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
0.0
0.1
0.2
0.3
q (Å-1)
Samples with crystallographic structure
Samples with multiple constructs
Samples with NMR structure
Samples with NMR and crystallographic structure

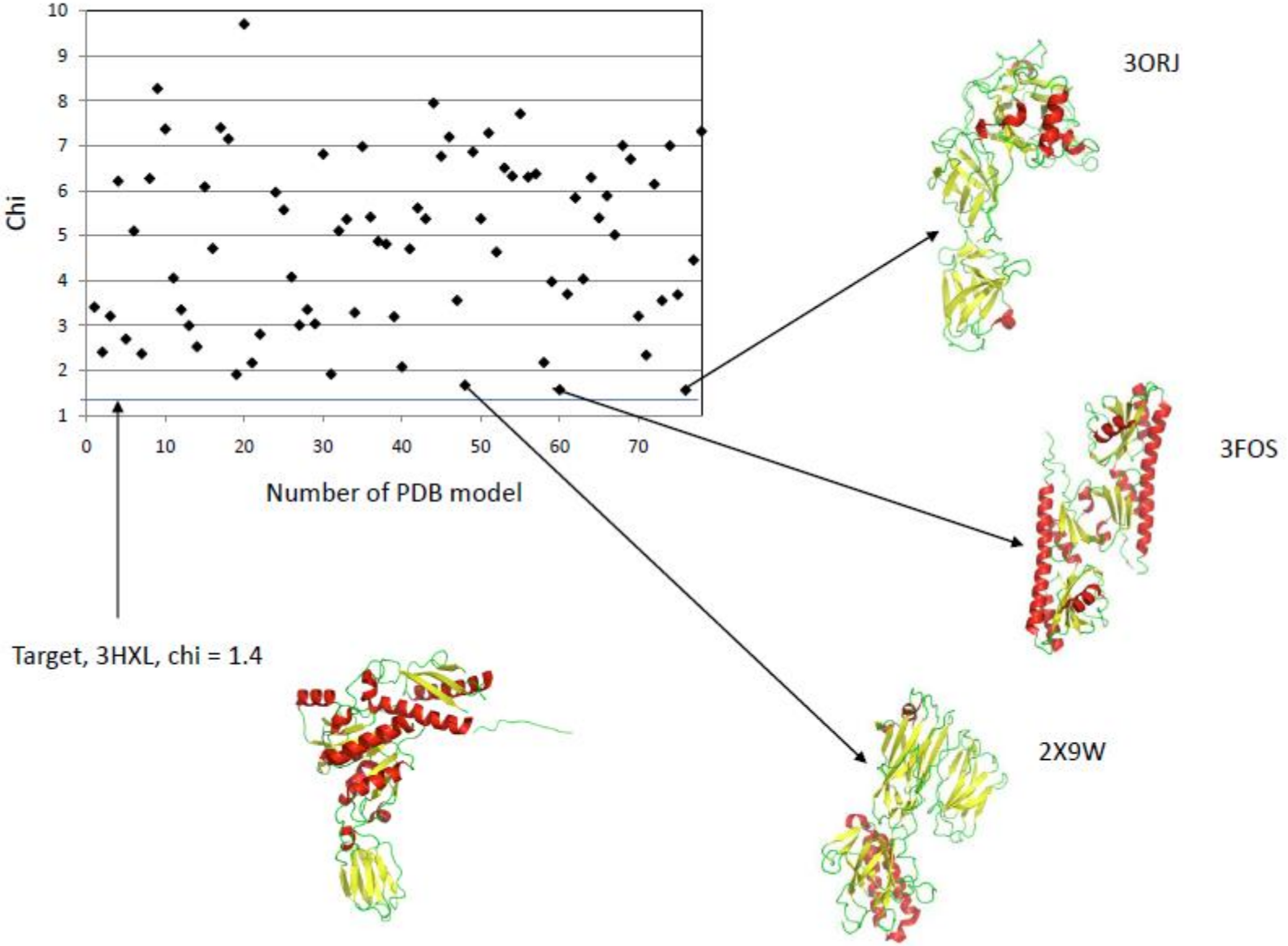

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Samples with crystallographic structure
Samples with

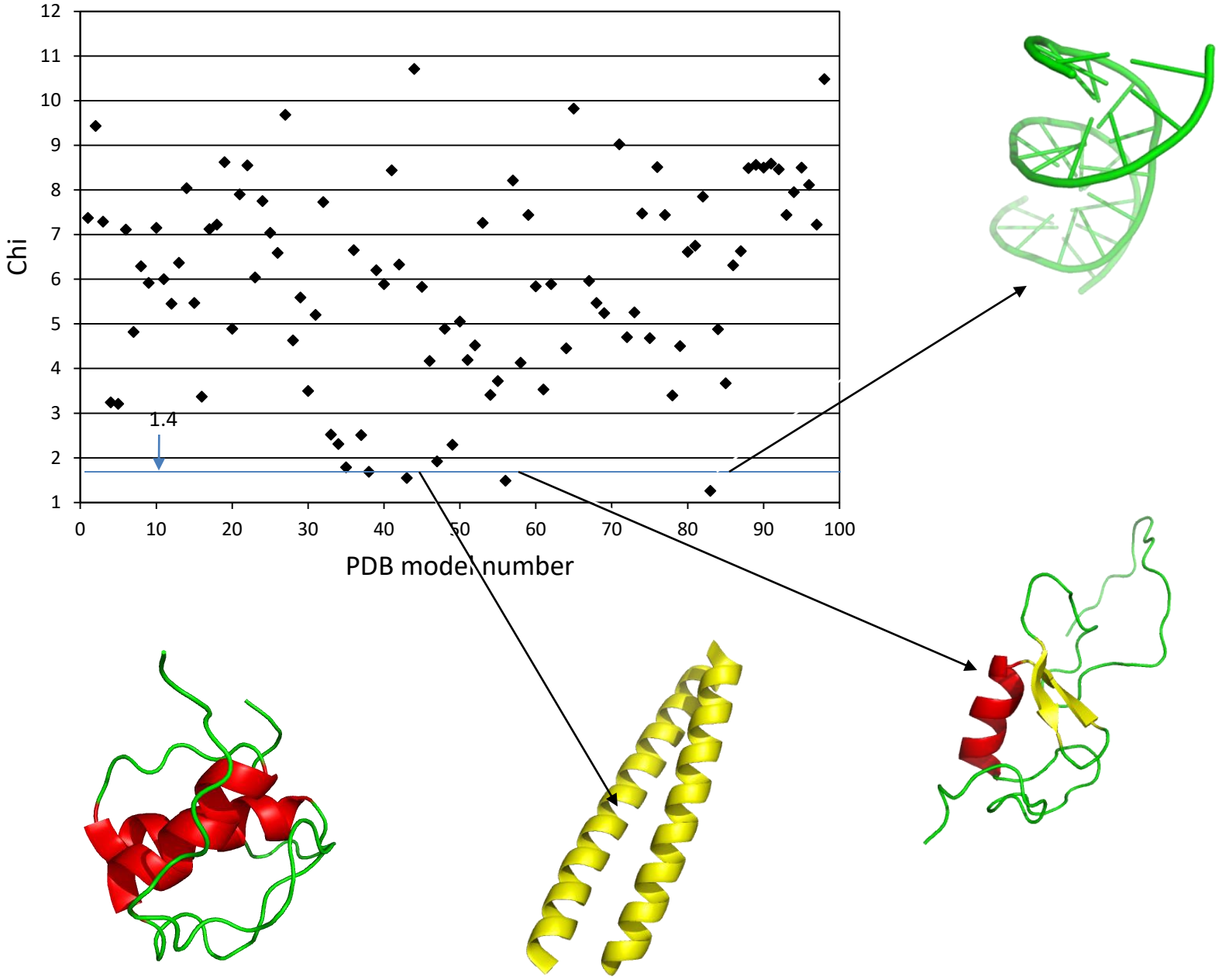
In all cases where we have:

(1) structural information and
(2) good SAXS data

the reconstruction has always accurately represented the envelope of the structure

	Samples where crystallographic structures were available																												
1	**4.2**	15.4	24.5		14.4	5.2	16.3	19.3	7.8	10.0		17.5	23.6	16.6	16.6	9.5		7.1	3.7	3.0	1.4	6.7	3.9	1.7	10.4	9.5	12.2	6.7	4.2	3.5
2	32.1	**2.6**	13.7		3.1	20.6	1.5	5.2	6.8	5.6		3.0	17.4	1.6	1.6	8.9		32.2	21.6	28.3	33.1	19.5	29.2	31.7	31.9	16.7	8.6	5.8	18.2	26.8
3	33.8	7.3	**1.6**		9.5	26.6	5.9	2.5	17.2	15.6		4.9	9.0	5.5	5.5	17.4		31.2	25.9	30.0	33.6	24.7	32.3	32.7	29.9	25.0	20.0	17.6	25.5	31.0
4	76.5	10.6	16.5	**2.6**	12.2	56.1	6.1	12.0	30.5	26.5		10.3	16.0	6.9	6.8	27.5		72.8	55.2	66.9	76.7	50.7	71.7	74.4	67.0	50.5	36.9	31.0	52.0	68.0
5	74.4	3.7	20.1		**2.2**	56.5	7.5	3.8	14.9	9.6		2.7	31.9	5.9	6.3	27.5		82.0	59.0	70.2	78.7	55.0	69.1	75.8	73.9	41.2	23.7	15.6	51.4	63.8
6	18.3	20.5	32.3		20.5	**6.1**	20.5	25.9	17.2	17.8		23.1	25.7	20.9	21.0	7.1		5.8	3.7	6.4	11.8	4.2	16.9	11.2	6.7	19.1	18.9	16.4	8.5	17.2
7	57.6	6.5	13.1		8.9	39.8	**3.8**	10.5	22.0	18.8		7.0	14.4	4.9	4.5	18.3		51.0	38.7	48.6	56.7	34.2	53.3	55.2	47.7	37.4	26.7	22.1	36.2	50.8
8	34.4	3.8	5.1		2.7	24.5	4.2	**2.0**	10.1	8.2		3.6	12.9	3.6	3.7	12.3		32.1	24.4	29.8	34.6	21.9	31.7	33.7	30.9	20.3	13.9	10.9	22.2	29.9
9	18.9	4.1	18.1		3.1	10.7	4.8	7.9	**3.7**	3.6		5.9	18.6	4.5	4.8	7.2		21.2	12.7	16.0	19.5	11.7	16.0	18.7	20.0	7.1	3.4	3.4	8.1	14.5
10	20.4	4.9	22.4		3.0	12.2	5.8	10.4	4.1	**4.2**		7.5	20.8	5.7	6.1	7.9		25.5	15.1	19.3	22.2	13.7	17.3	21.2	21.9	8.0	3.5	3.5	9.5	15.4
11	94.2	37.1	19.8		41.6	77.8	31.1	26.3	59.4	56.0	**3.0**	31.0	15.7	30.8	30.7	54.9		84.3	75.8	86.4	93.8	71.1	91.1	92.2	78.1	75.8	65.2	60.0	75.6	88.4
12	33.2	3.2	4.2		4.6	23.8	2.7	3.1	12.8	10.9		**2.9**	9.0	2.5	2.5	11.9		29.2	23.3	28.9	33.3	21.0	31.3	32.2	28.1	21.9	15.9	13.1	22.4	29.5
13	26.4	9.3	7.9		10.4	19.2	7.5	8.0	15.1	14.1		8.3	**1.4**	7.7	7.6	11.1		20.0	18.0	22.4	25.7	15.8	25.2	25.0	18.5	20.1	16.6	15.2	18.6	24.3
	Samples where multiple constructs and crystallographic structures were available																												
14	41.6	3.5	9.1		4.6	28.0	1.7	6.0	13.0	10.6		3.9	13.9	**1.9**	1.8	12.0		37.7	27.7	35.4	41.6	24.6	38.3	40.2	37.5	25.1	16.8	13.2	25.4	36.1
15	19.3	2.5	4.1		2.7	12.8	1.7	3.0	6.8	5.8		2.5	5.5	1.9	**1.8**	5.2		16.3	12.4	16.2	19.1	10.9	18.0	18.5	16.4	12.2	8.4	6.9	11.9	16.9
16	8.9	3.8	12.5		3.2	4.3	4.5	7.1	3.0	2.8		5.6	12.0	4.7	4.8	**1.7**		10.5	5.8	6.5	8.5	5.3	7.3	8.1	8.6	4.2	2.5	2.6	2.9	6.5
17	11.8	9.7	21.2		9.2	3.4	10.1	14.4	7.6	7.5		12.2	18.1	10.3	10.4	2.5	**2.1**	10.8	5.3	5.6	9.8	3.7	10.0	9.7	8.6	8.8	7.8	7.1	2.3	9.3
	Samples where NMR structures were available																												
18	7.0	16.5	26.3		15.6	2.0	17.1	20.6	9.8	11.5		18.6	22.1	17.4	17.5	7.9		**4.7**	**1.4**	1.3	2.3	3.8	6.6	2.0	2.7	11.4	13.6	8.7	4.8	6.3
20	10.2	13.0	22.8		12.6	1.9	13.5	17.2	9.2	9.8		15.2	19.4	13.7	13.8	4.2		6.3	2.0	**1.7**	6.0	2.2	9.1	5.5	5.5	10.7	11.1	8.5	3.2	9.3
21	5.2	14.8	24.3		13.7	3.4	15.6	18.6	7.6	9.5		16.9	21.8	15.8	16.0	7.9		5.4	2.4	1.8	**1.5**	4.6	4.7	1.7	5.7	9.0	11.7	6.6	3.5	4.5
22	6.3	6.6	12.0		6.4	1.6	6.8	8.8	5.3	5.3		7.7	10.1	6.9	6.9	1.8		3.9	1.4	1.8	4.3	**1.6**	5.6	4.0	3.2	5.9	5.7	5.0	1.8	5.6
23	1.6	10.1	16.9		8.9	6.6	10.9	12.5	3.7	5.6		11.5	17.1	11.0	11.1	8.1		8.0	5.7	5.3	3.6	7.1	**1.4**	4.3	10.8	4.1	7.3	3.5	4.5	1.8
24	7.8	6.3	12.7		6.1	1.8	6.4	8.9	5.7	5.6		7.6	10.2	6.6	6.7	1.6		6.0	2.0	4.2	6.5	2.0	6.9	**5.9**	5.2	6.4	5.6	5.4	2.0	6.6
25	18.3	15.8	24.8		16.2	7.9	15.2	19.7	15.9	15.7		17.6	17.5	15.5	15.5	6.1		6.3	5.1	9.0	14.2	3.5	17.0	13.6	**7.4**	17.0	15.7	15.3	8.3	16.8
26	16.4	8.7	26.0		4.9	13.0	10.8	12.7	1.9	2.3		11.2	24.8	10.3	10.8	11.8		26.4	16.8	18.1	19.8	15.9	13.5	18.9	21.9	**2.1**	3.1	2.9	10.9	11.3
	Samples where both crystallographic and NMR structures were available																												
27	13.9	2.4	10.2		2.5	8.3	2.5	4.0	2.7	2.5		3.0	10.7	2.2	2.3	4.8		14.2	9.0	10.5	13.6	8.1	12.0	13.3	12.2	5.8	**2.5**	**2.4**	6.4	11.1
28	8.3	13.1	23.9		12.2	1.7	13.6	17.5	7.6	8.7		15.4	19.5	13.9	14.0	3.8		6.1	2.0	1.8	4.5	1.8	7.3	4.0	3.1	9.1	10.3	6.9	**2.3**	**7.4**


Chi
Number of PDB model
Target, 3HXL, chi = 1.4
3ORJ
3FOS
2X9W


Chi
PDB model number
1.4
12
11
10
9
8
7
6
5
4
3
2
1
0
10
20
30
40
50
60
70
80
90
100

How robust is it?

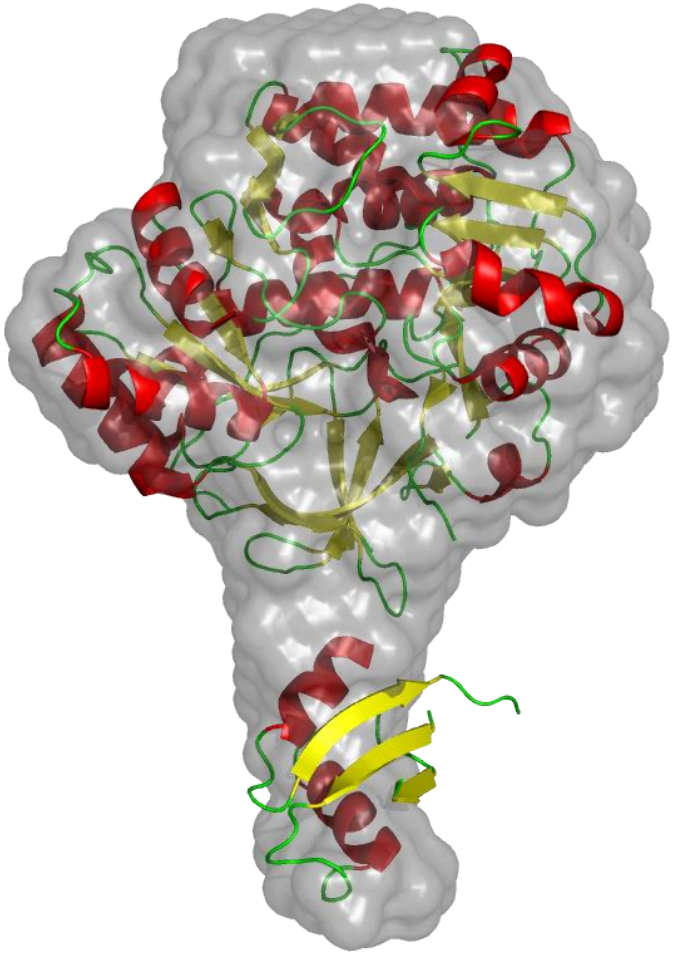
There are several metrics that can be used to determine the quality of data and correctness of the envelope.

How robust is it?

If you already have some structural knowledge it us very robust.

But what if you don’t?

One example, comparing Structural Blast Results



The envelope of the unknown structure confirms structural homology to sequence homology

Other examples with SAXS

- One can think of many experiments where an envelope would be useful information.
- For example, by using multiple constructs, components of a structure could be put in their relative 3D environment.
- Mutational studies on the predicted surfaces of complex contacts could be structurally tested.
- Many, many applications.

Other examples with SAXS

- SAXS can be used to analyze natively unfolded proteins.
- It can identify aggregation as a function of biochemical conditions.
- It can measure the B_{22}, the attraction/repulsion of protein samples.
- It can characterize quantitatively how well folded a sample is (and as a function of biochemical conditions)
- It can be used in a time resolved manner (at least with a synchrotron source).

The End (almost)

Take home message

- With good data, SAXS is complementary to X-ray crystallography, NMR, or other structural methods.
- It builds on the information other techniques provide.
- Without complementary structural information SAXS provides basic data but envelope reconstruction cannot be completely validated (although our success is pretty good).
- It can be used on it's own as a hypothesis generator and, with careful experimental design, to test a hypothesis.

- **Remember it's a low resolution technique.**

Acknowledgements

Hauptman-Woodward Medical Research Institute

Thomas Grant, Joseph Luft, Raymond Nagel, Elizabeth Snell, Jennifer Wolfley and George DeTitta

University of Rochester

Elizabeth Grayhack, Eric Phizicky, and staff

Stanford Synchrotron Radiation Lightsource

Hiro Tsuruta, Aina Cohen, Mike Soltis, and the beamline staff

University of Oxford

Yoonjoo Choi and Charlotte Deane

Support and Funding

NIH, NSF and DoD support to Edward Snell

Thank you and questions?

esnell@hwi.buffalo.edu