Presenter: Emre Brookes

Topic:

Small angle X-ray and Neutron Scattering

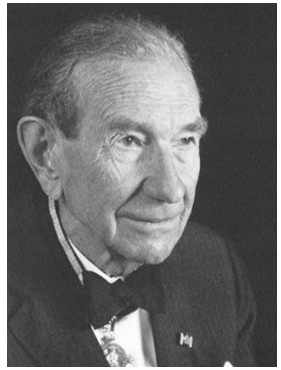
Outline - SAS

- Experimental setup
- Theory
- Modeling
- Software
- Practical considerations, Sample preparation etc.
- Possibly other techniques introduced as time permits

Kratky plot

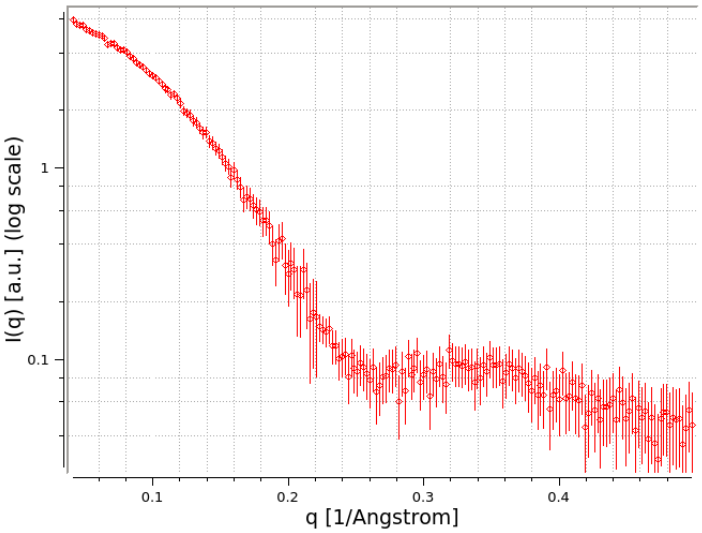
SAS provides a sensitive means to evaluate *the degree of compactness* of a protein:

- To determine whether a protein is globular, extended or unfolded
- To monitor the folding or unfolding transition of a protein

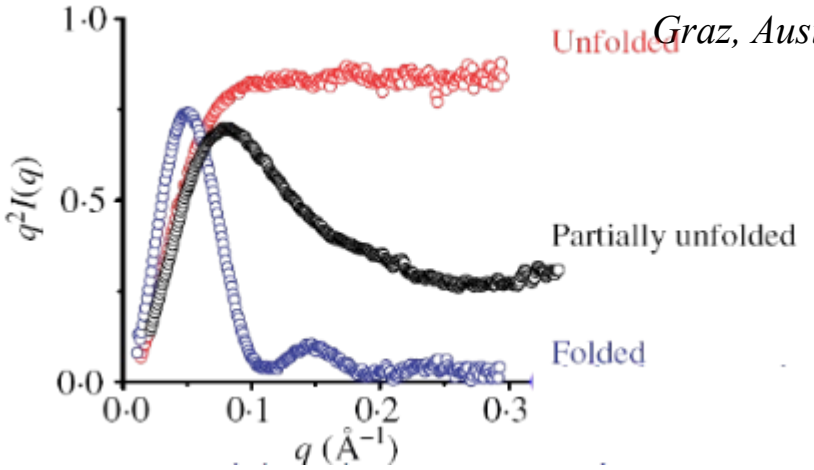


Prof. Otto Kratky 1902-1995 Graz, Austria

This is most conveniently represented using the Kratky plot:



$q^2 I(q)$ versus q



Putnam, D., et al. (2007) Quart. Rev. Biophys. 40, 191-285.

Folded particle : *bell-shaped curve* (asymptotic behavior $I(q) \sim q^{-4}$)

Random polymer chain : *plateau* at large q-values (asymptotic behavior $I(q) \sim q^{-2}$)

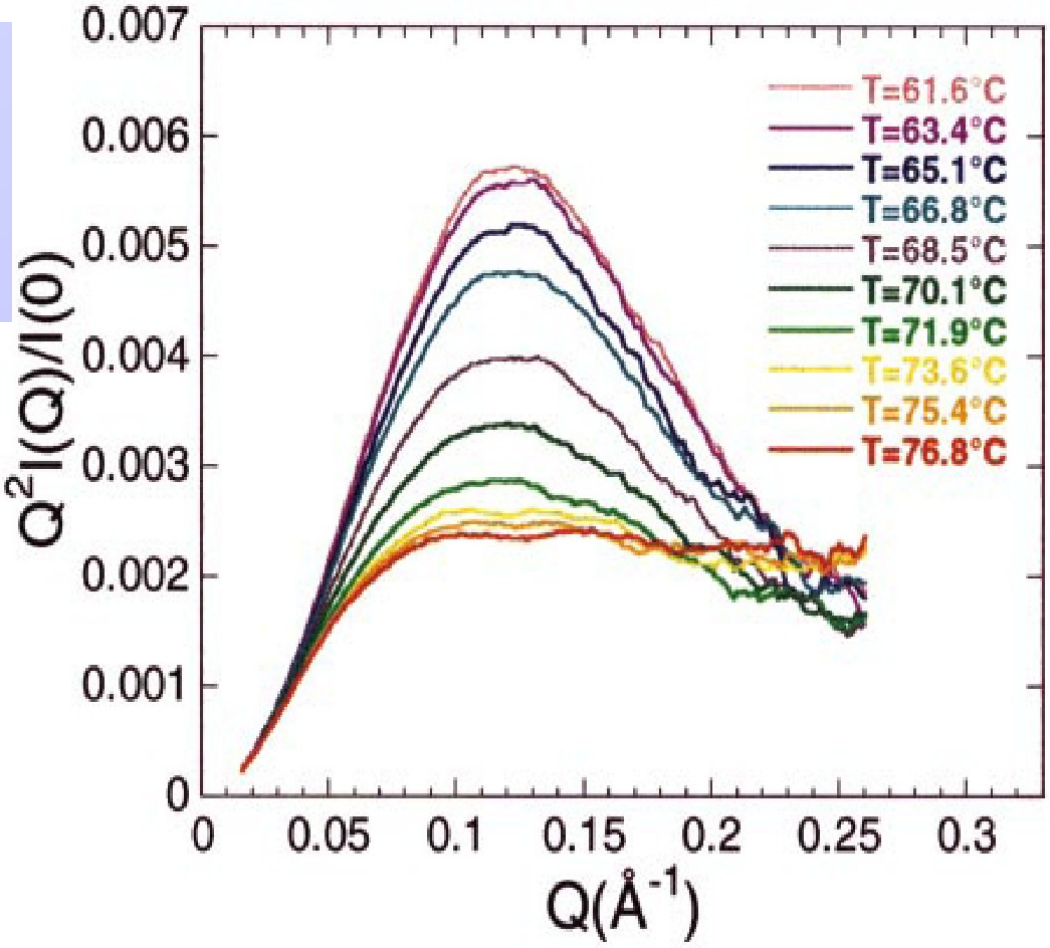
Extended polymer chain : *increase* at large q-values (asymptotic behavior $I(q) \sim q^{-1.x}$)

Kratky plots of (partially) folded proteins

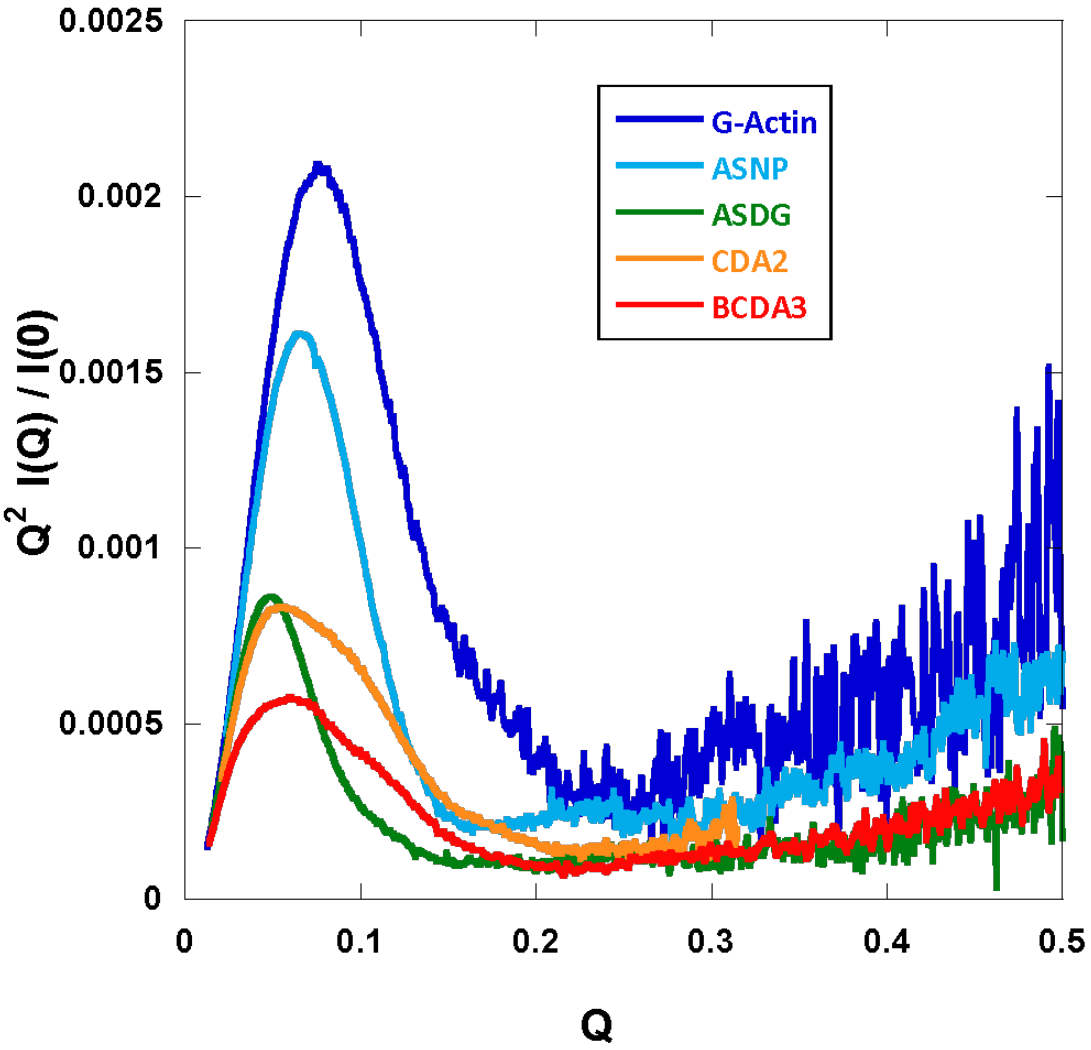
Pérez et al., J. Mol. Biol.(2001), 308, 721-743

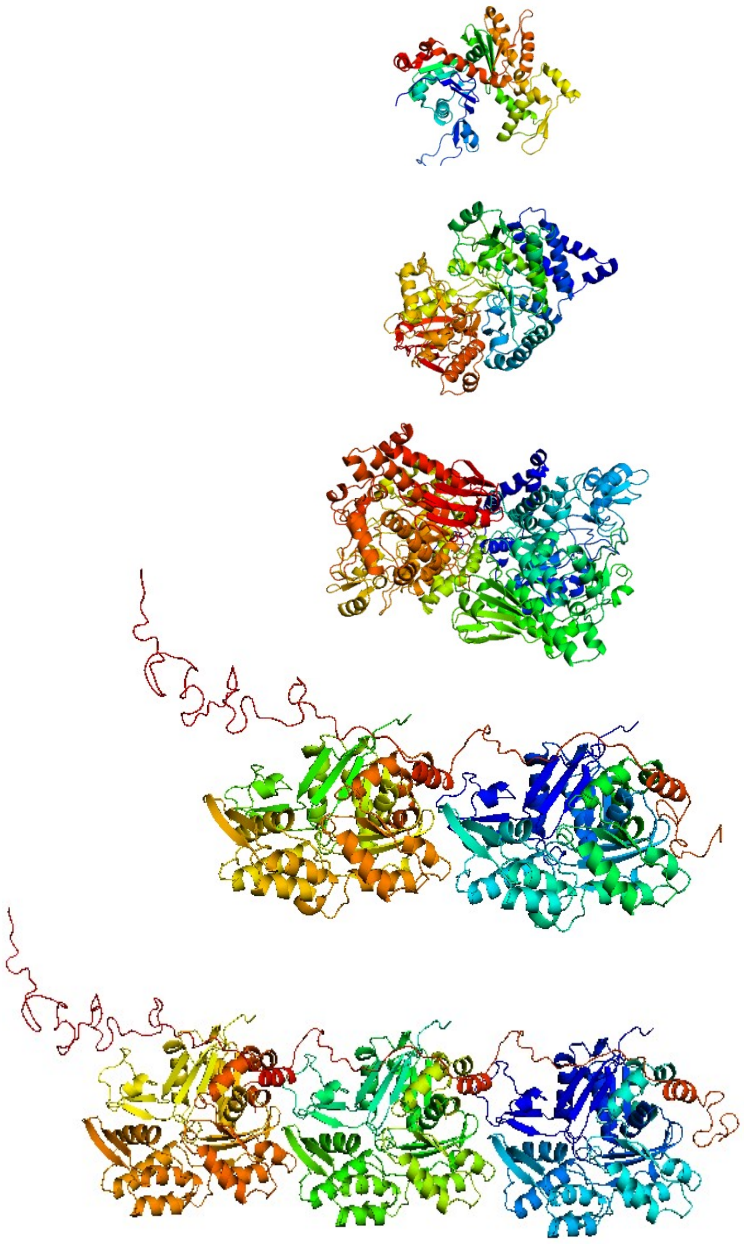
In practice, thin Gaussian chains do not exist.

In spite of the plateau at T=76°C, NCS is not a Gaussian chain when unfolded, but a thick chain with persistence length.



Kratky plots of folded proteins

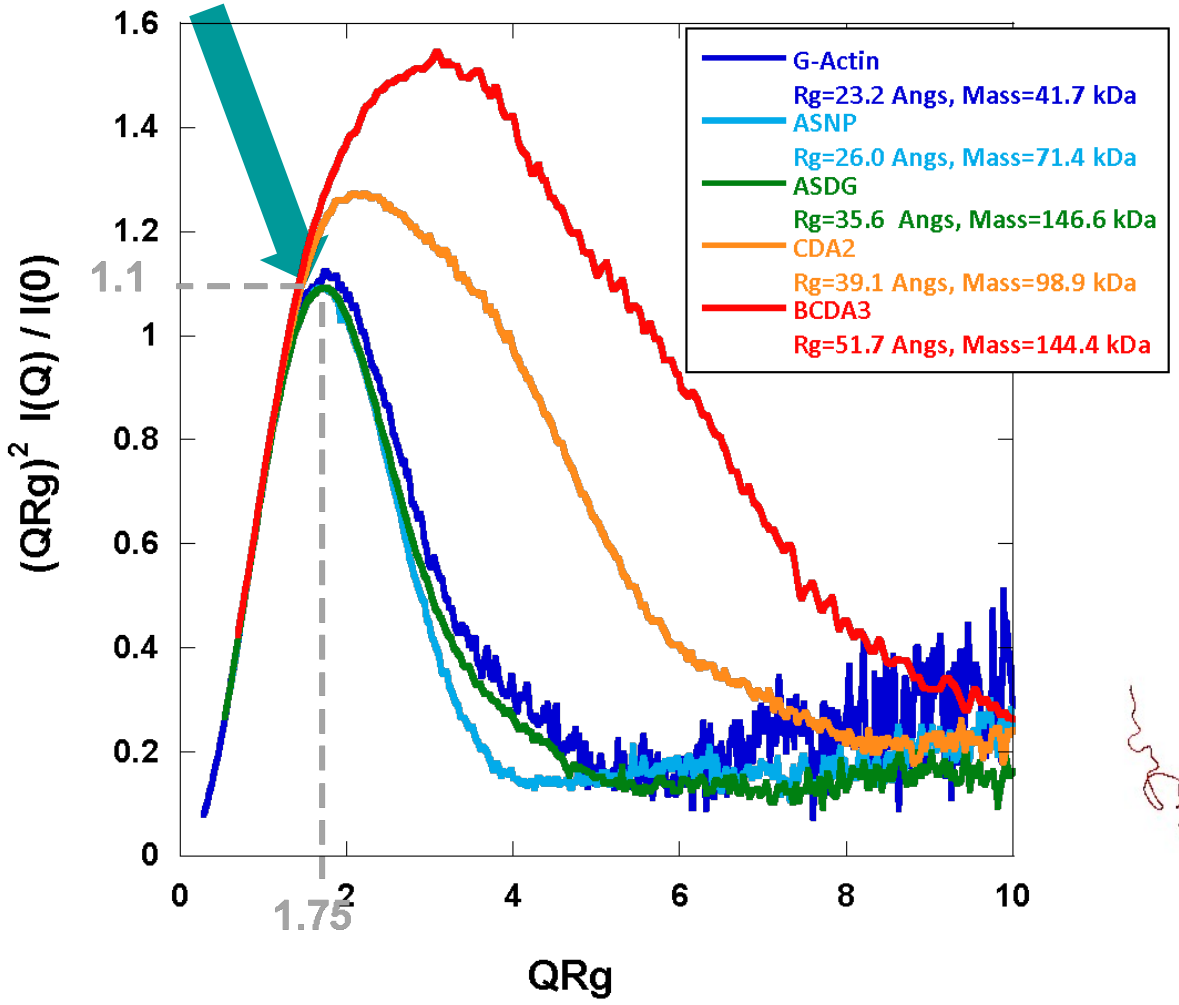


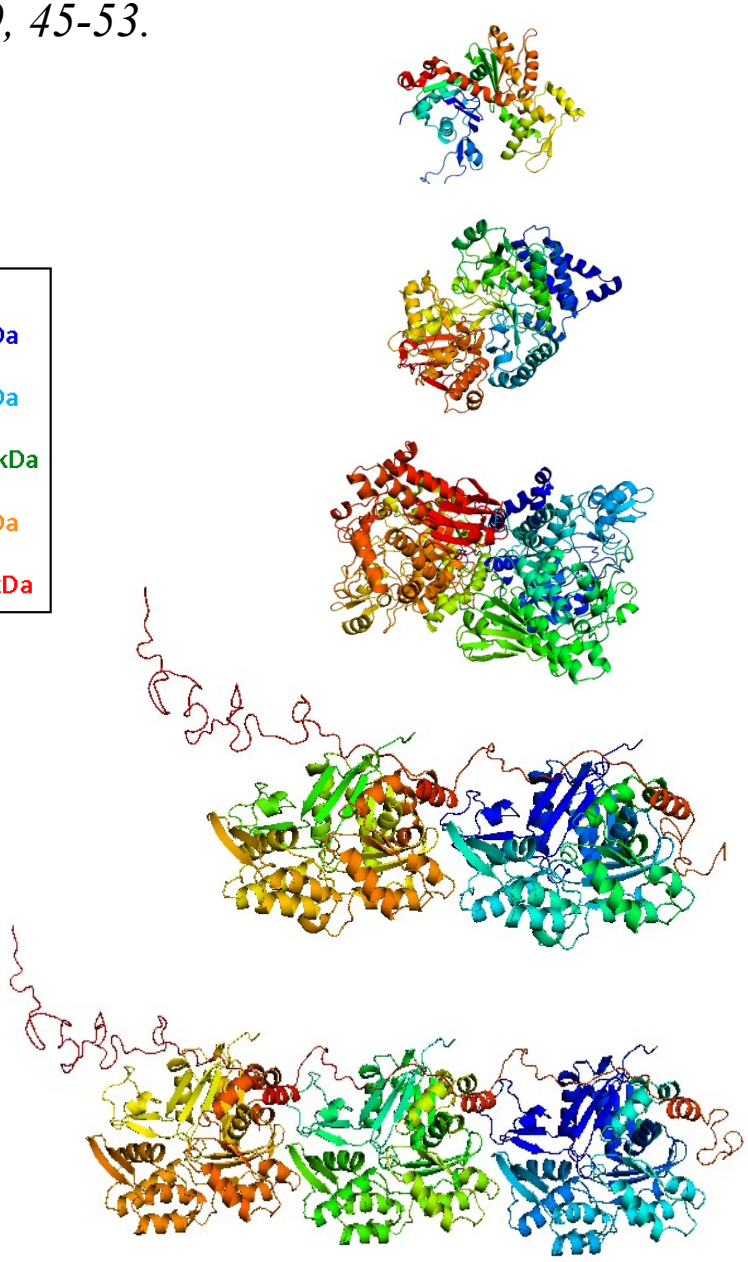


Dimensionless Kratky plots of folded proteins

Durand et al. (2010), J. Struct. Biol. 169, 45-53.

For globular structures, DLKPs fold into the same maximum

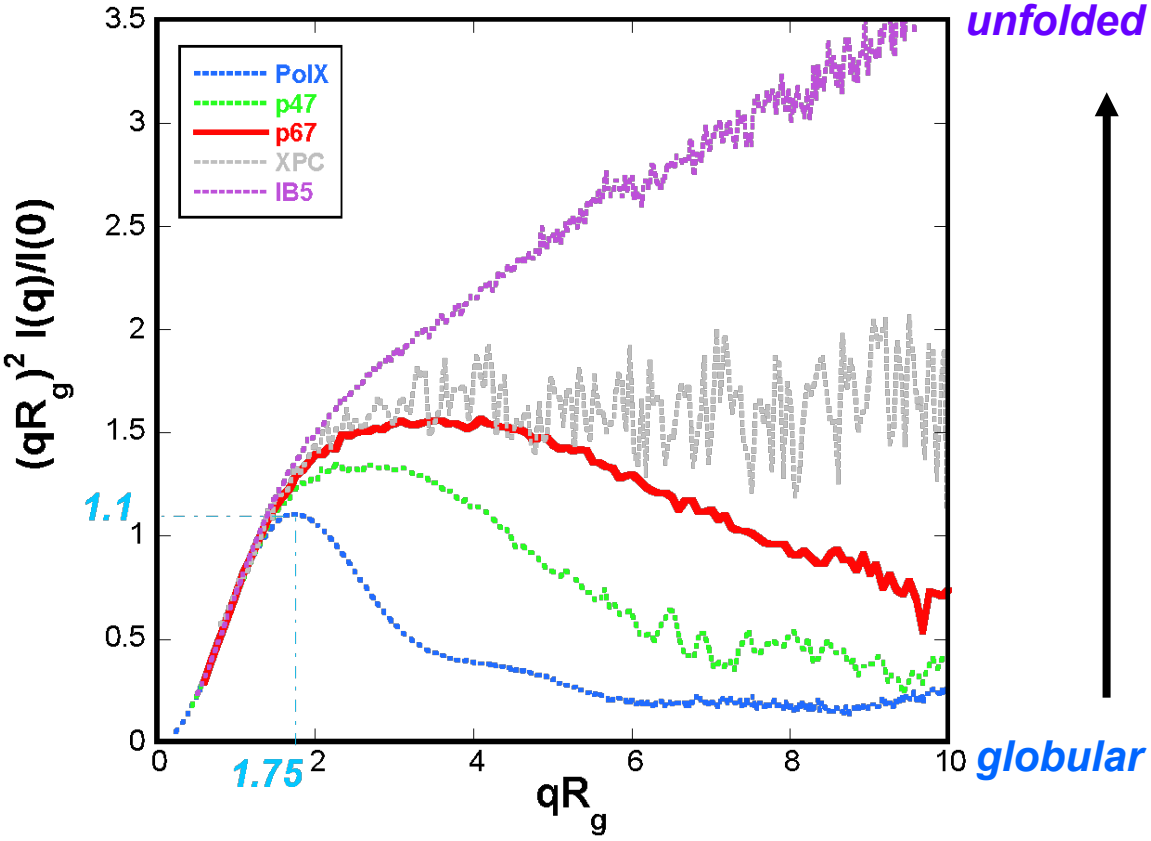


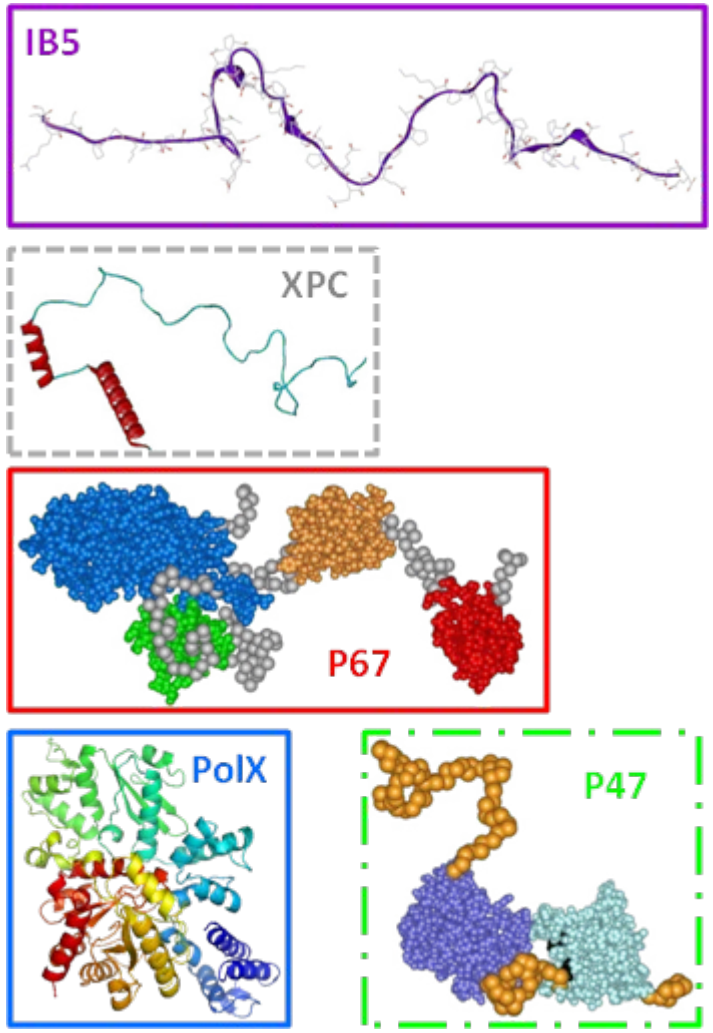


The maximum value on the dimensionless bell shape tells if the protein is globular.

Dimensionless Kratky plots of (partially) folded proteins

Receveur-Bréchot V. and Durand D (2012), Curr. Protein Pept. Sci., 13:55-75.





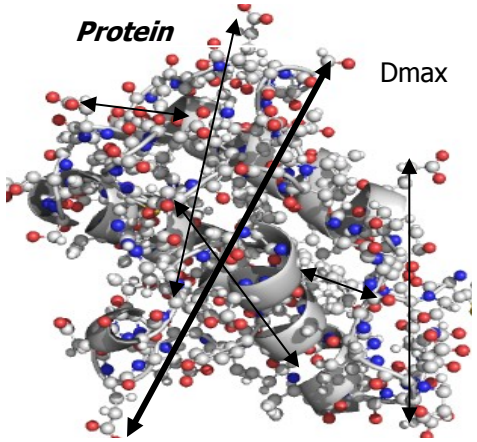
The bell shape vanishes as folded domains disappear and flexibility increases.

The curve increases at large q as the structure extends.

Distance distribution function $p(r)$

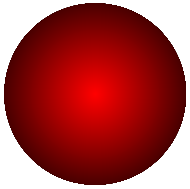
$p(r)$ is obtained by "histogramming" the distances between any pair of scattering elements within the particle.

(weighted by scattering density)



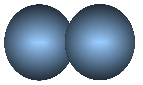
$p(r)$ vanishes at $r = D_{max}$

Cylinder

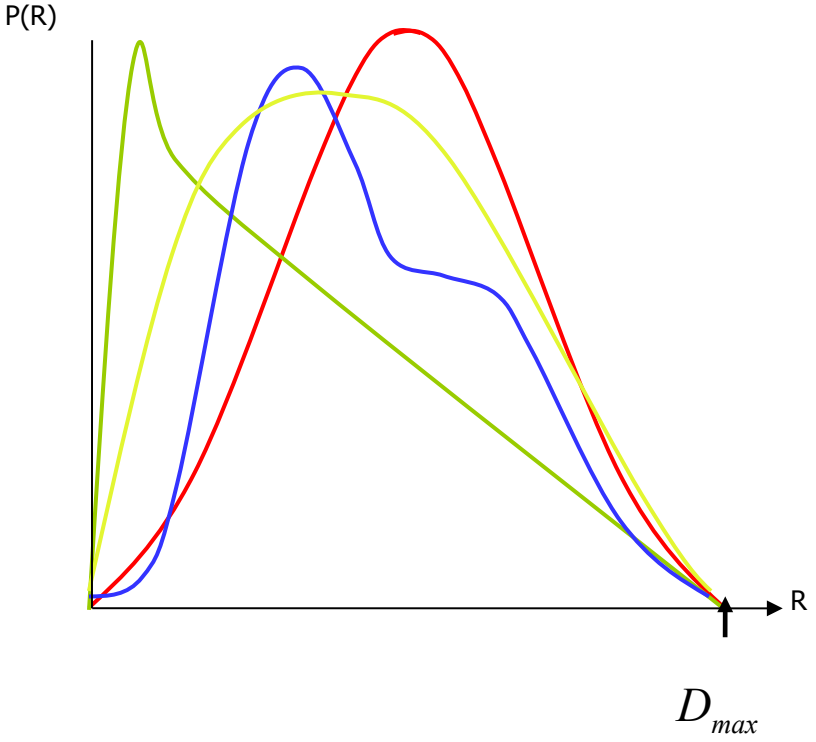


Solid sphere

Disc



Domains



The distance distribution function characterizes the shape of the particle **in real space**

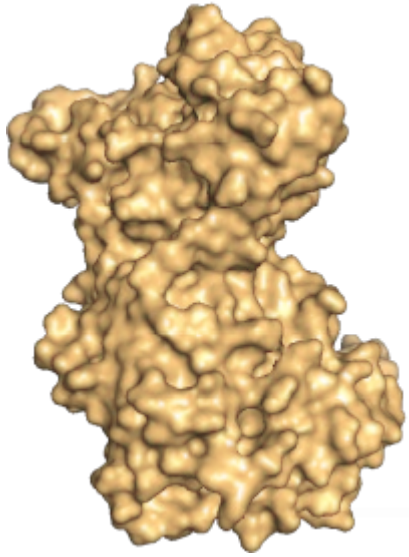
Autocorrelation function

$$\gamma(\mathbf{r}) = \rho(\mathbf{r}) * \rho(-\mathbf{r}) = \int_{V_\mathbf{u}} \rho(\mathbf{r}+\mathbf{u})\rho(\mathbf{u})\, dV_\mathbf{u}$$

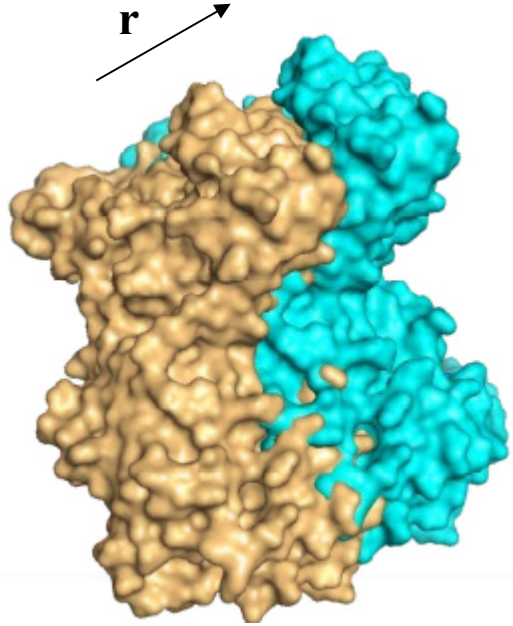
$\rho(\mathbf{r}) = \rho$ Assuming uniform scattering density inside of “Molecule”, 0 o.w.

$$\Rightarrow \gamma(\mathbf{r}) = \rho^2 V_{\text{overlap}}(\mathbf{r}) \qquad \gamma(0) = \rho^2 V$$

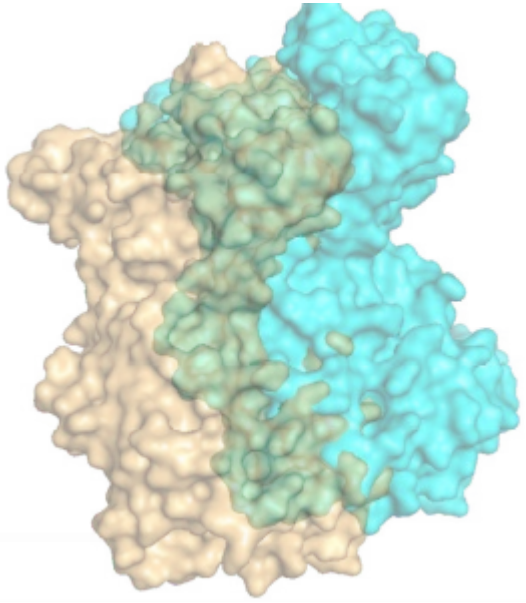
Molecule



Ghost



Molecule ∩ Ghost = overlap



Images: Patrice Vachette

Autocorrelation function

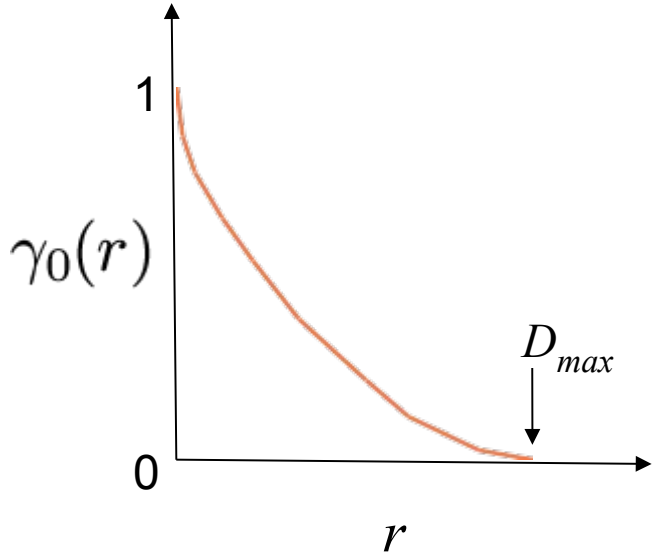
$$\gamma(\mathbf{r}) = \rho(\mathbf{r}) * \rho(-\mathbf{r}) = \int_{V_\mathbf{u}} \rho(\mathbf{r}+\mathbf{u})\rho(\mathbf{u})\, dV_\mathbf{u}$$

$\rho(\mathbf{r}) = \rho$ Assuming uniform scattering density inside of “Molecule”, 0 o.w.

$$\Rightarrow \gamma(\mathbf{r}) = \rho^2 V_{\text{overlap}}(\mathbf{r}) \qquad \gamma(0) = \rho^2 V$$

$\gamma(r) = \langle \gamma(\mathbf{r}) \rangle$ Spherical average

$$\gamma_0(r) = \frac{\gamma(r)}{\gamma(0)}$$ Characteristic function

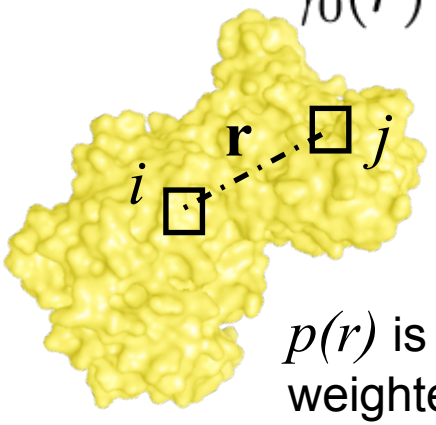


$\gamma_0(r)$ Probability of finding a point within the molecule at a distance r from a given point

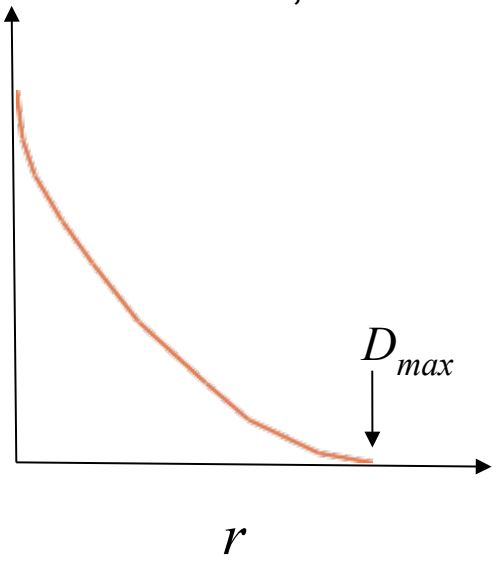
Distance (pairwise) distribution function

$\rho(\mathbf{r}) = \rho$ Assuming uniform scattering density inside of “Molecule”, 0 o.w.

$\Rightarrow \gamma(\mathbf{r}) = \rho^2 V_{\text{overlap}}(\mathbf{r})$ $\gamma(0) = \rho^2 V$



$\gamma_0(r)$ Probability of finding within the molecule a $\gamma_0(r)$ point j at a distance r from a given point i



$p(r)$ is the distribution of distances between all pairs of points within the molecule weighted by the respective scattering densities.

→ number of elementary volumes $i \propto V$

→ number of elementary volumes $j \propto r^2$

→ number of pairs (i,j) separated by the distance $r \propto V\, r^2\, \gamma_0(r)$

$$p(r) = \rho^2 V r^2 \gamma_0(r) = r^2 \gamma(r)$$

Distance distribution function *p(r)*

US-SOMO: SAS Functions

PDB Filename: 1HEL

Definition files:

SAS I(q) Functions:

P(r) vs. r Functions:

Load P(r) Distribution | Load Plotted P(r) | Clear P(r) Distribution | Legend | Width

Bin size (Angstrom): 2

Smoothing: 0

Raw | SAXS | SANS | Normalize

Residue contrib. range (Angstrom): 40 50 Display

Compute P(r) Distribution 100%

File

Preparing file 1HEL model 1 for p(r) vs r plot in SAXS mode, Normalized.

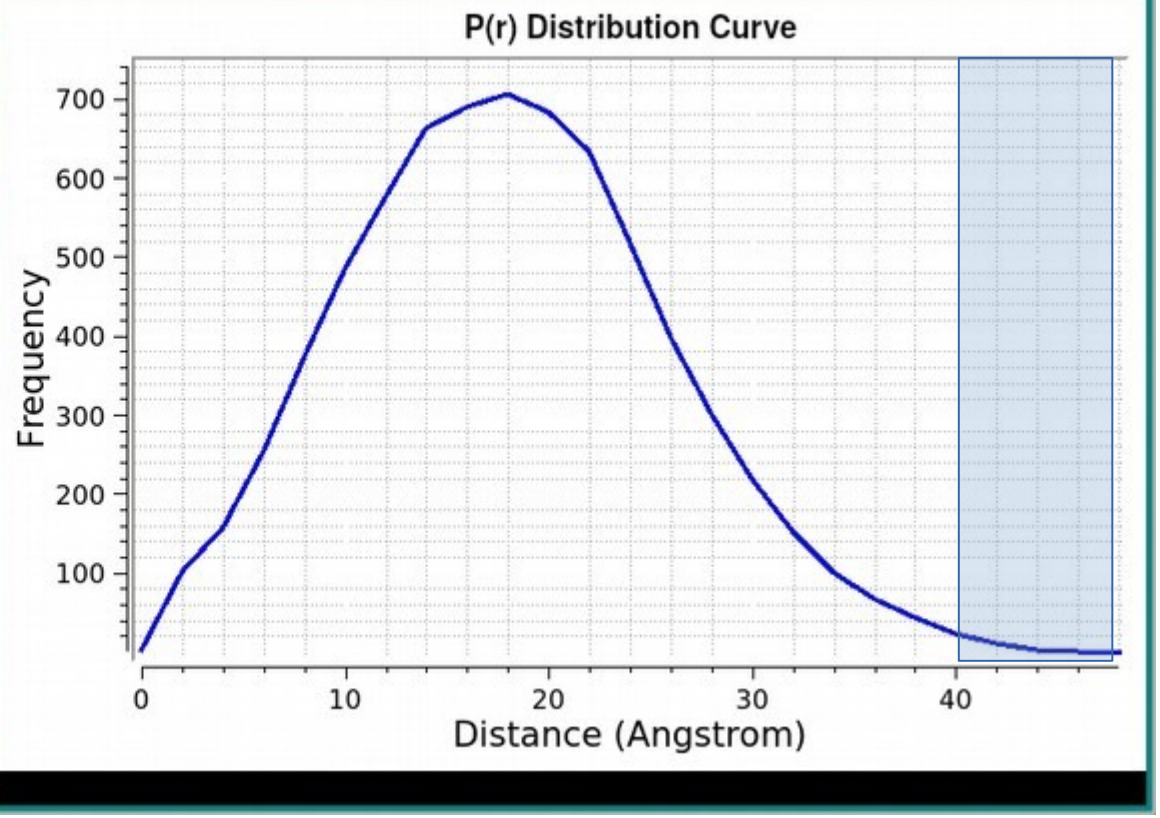
Number of atoms 1001. Bin size 2.

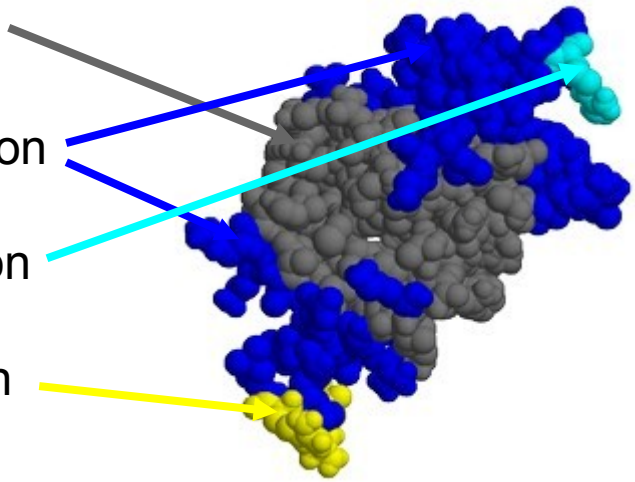
P(r) curve file: /root/ultrascan//somo/saxs/1HEL_1b2.sprr_x created.

1HEL Molecular weight 14306.8 (computed from pdb)

Stop | Open Options Panel

Help | Close





UltraScan-SOMO https://somo.aucsolutions.com

Relation between $p(r)$ and $I(q)$

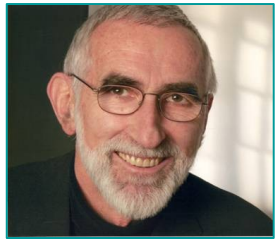
$$I(q) = 2\pi^2 \int_0^\infty p(r) \frac{\sin(qr)}{qr} dr$$

$$p(r) = \frac{1}{2\pi^2} \int_0^\infty I(q)\, qr \sin(qr)\, dq$$

However, direct calculation of $p(r)$ from $I(q)$ is made difficult and risky by $[q_{min}, q_{max}]$ truncation and data noise effects.

Back-calculation of the Distance Distribution Function

Glatter, O. *J. Appl. Cryst.* (1977) **10**, 415-421.



Prof. Otto Glatter
Guinier Prize 2012
Graz, Austria

Main hypothesis : the particle has a finite size, characterised by D_{max}.

D_{max} is proposed by the user
$p(r)$ is described over $[0, D_{Max}]$ by a linear combination of M orthogonal functions

$$p_{\text{theoret}}(r) = \sum_{n=1}^{M} c_n \varphi_n(r)$$

$I(q)$ is calculated by Fourier Transform of $p_{theoret}(r)$

$$I(q) = 2\pi^2 \int_0^{D_{\max}} p_{\text{theoret}}(r) \frac{\sin(qr)}{qr} dr$$

Dr. Dmitri Svergun
Hamburg, Germany

Svergun (1988) : program "GNOM"
M ~ 30 – 100; ill-posed least squares regularization method
“Perceptual criteria” : smoothness, stability, absence of systematic deviations

- Each criterium has a predefined weight
- The solution is given a score calculated by comparison with “ideal values”

Back-calculation of the Distance Distribution Function

$$p(r) = \frac{r^2}{2\pi^2 r_e^2 \varphi} \int_0^\infty q^2 I(q) \frac{\sin(qr)}{qr} dq$$

$$p_{\text{theoret}}(r) = \sum_{n=1}^{M} c_n \varphi_n(r)$$

The method of Glatter requires prior estimates of D_{max} and a regularization parameter.

Steen introduced methods to determine these parameters automatically using Bayesian analysis methods.

Thomas Bayes
1701-1761

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}$$

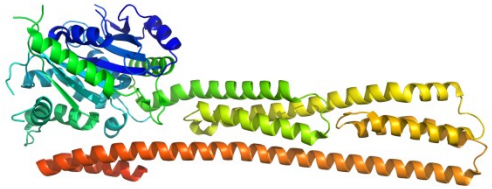
Data – $I(q)$
Generative model (Glatter)
Priors – D_{max} , α

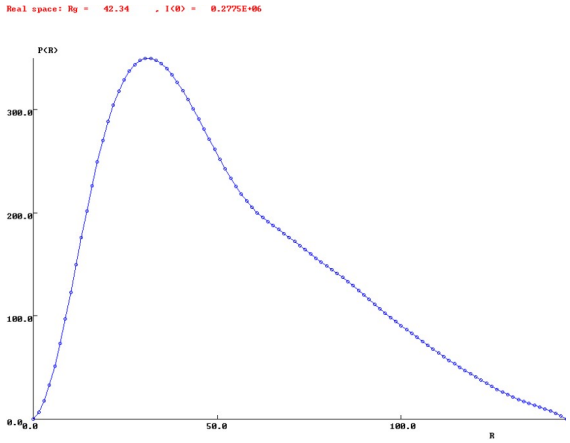
Bayesian inference:
Prior dists → sample → generate → filter against data → Posterior dists

Steen Hansen, J. Appl. Cryst. (2000). 33, 1415-1421

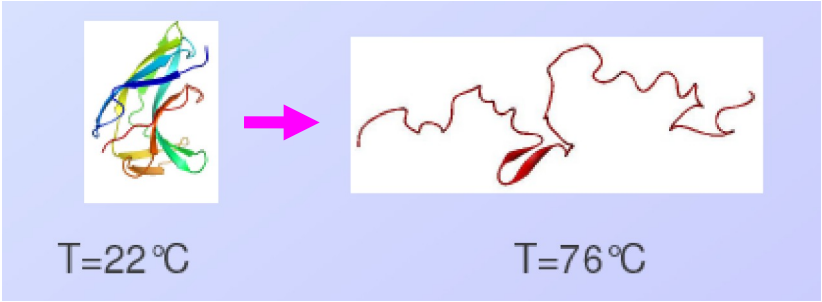
Distance Distribution Function

GBP1

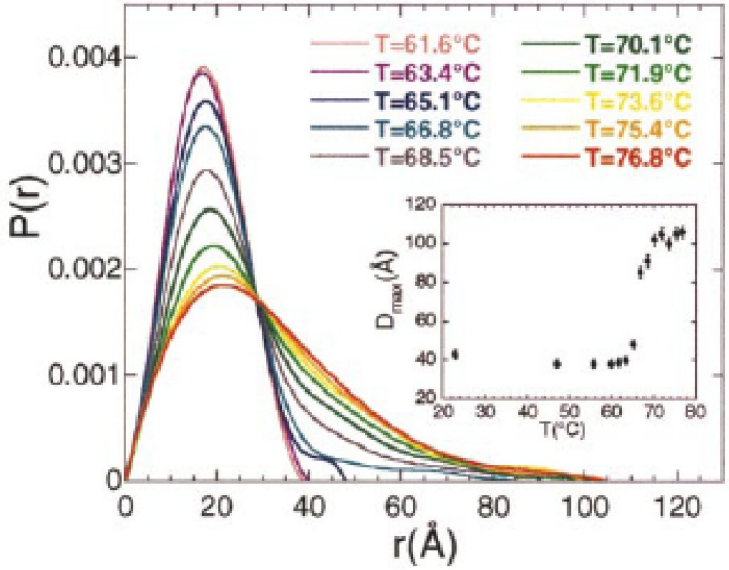




Heat denaturation of Neocarzinostatin



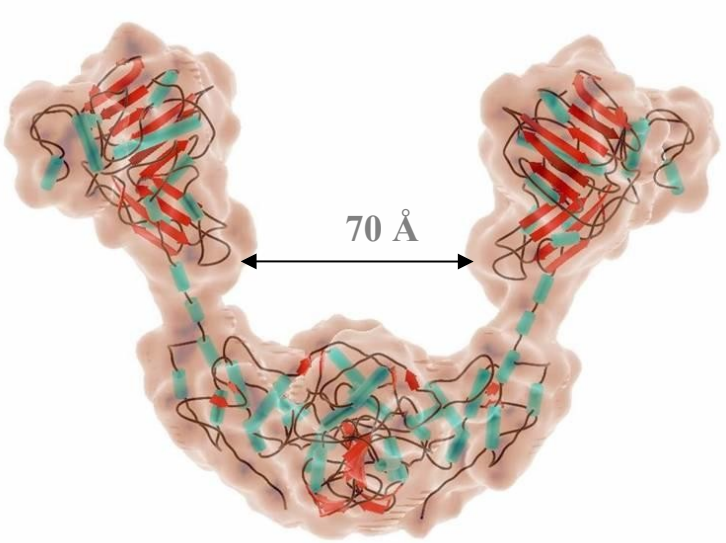
Pérez et al., J. Mol. Biol. (2001) 308, 721-743



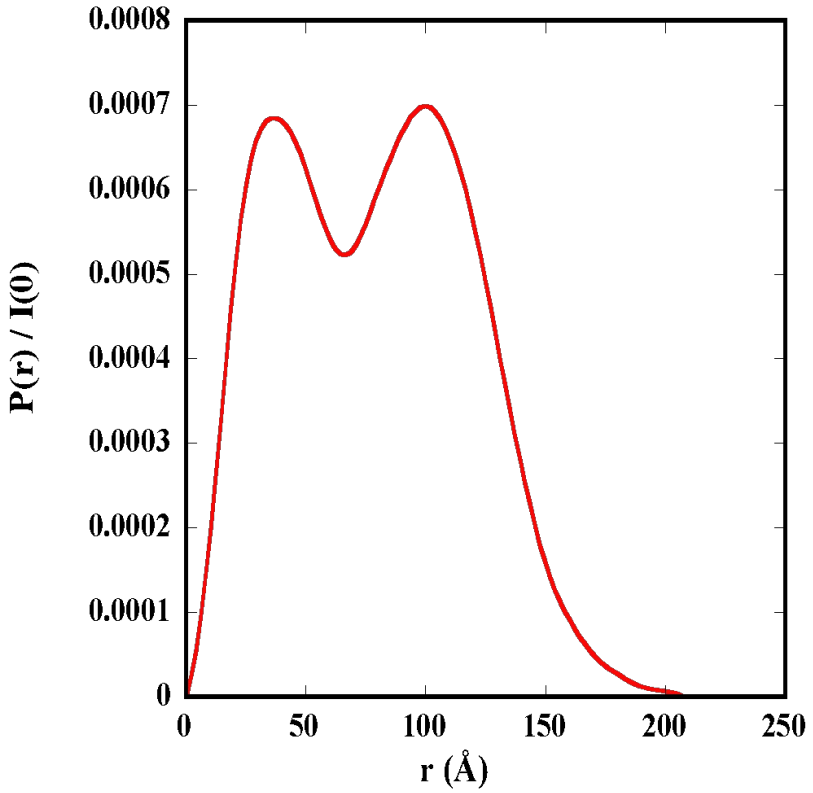
Distance Distribution Function

M. Graille et al., Structure (2008), 16, 360-370.

Topoisomerase VI



Bimodal distribution



Distance Distribution Function

The radius of gyration and the intensity at the origin can be derived from p(r) using the following expressions :

$$R_g^2 = \frac{\int_0^{D_{\max}} r^2 p(r)\, dr}{2\int_0^{D_{\max}} p(r)\, dr} \qquad I(0) = 4\pi r_e^2 \varphi \int_0^{D_{\max}} p(r)\, dr$$

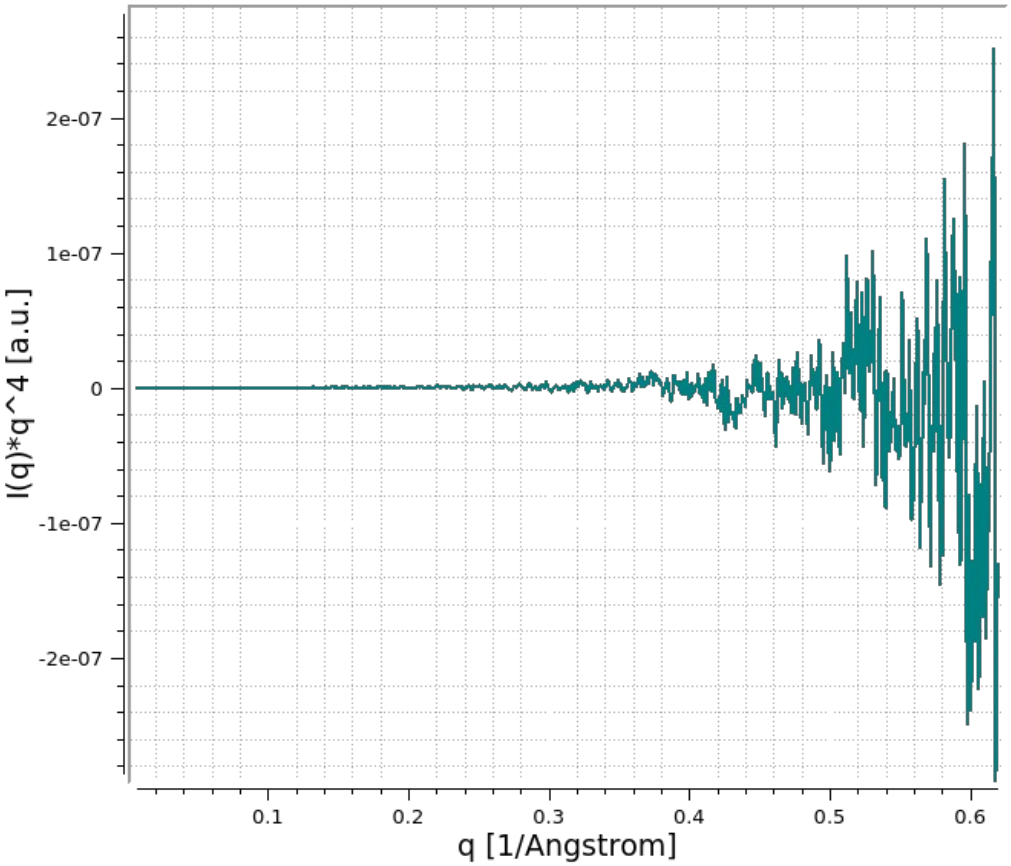
This alternative estimate of R_g makes use of the whole scattering curve, and is less sensitive to interactions or to the presence of a small fraction of oligomers.

Comparison of estimates from Guinier analysis and from $P(r)$ is a useful cross-check.

Porod's law

$$I \sim q^{-4}$$

Intensity decay is proportional to q^{-4} at higher angles for globular particles of uniform density.



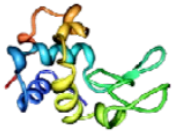
Günther Porod 1919-1984 Graz, Austria

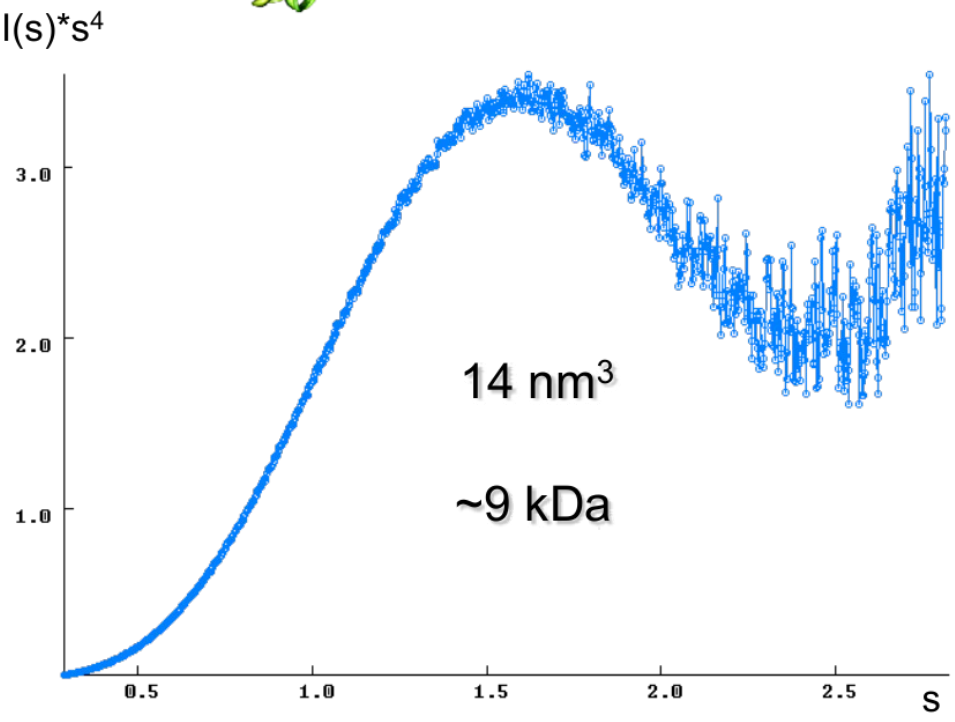
$$V_P = \frac{2\pi^2 I(0)}{\int_0^\infty [I(q) - K_4]\, q^2\, dq}$$

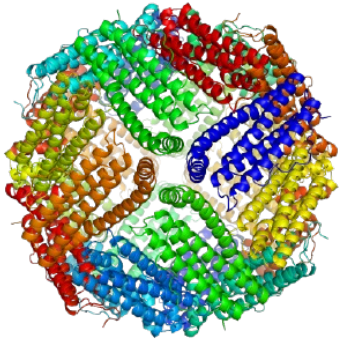
Porod's law

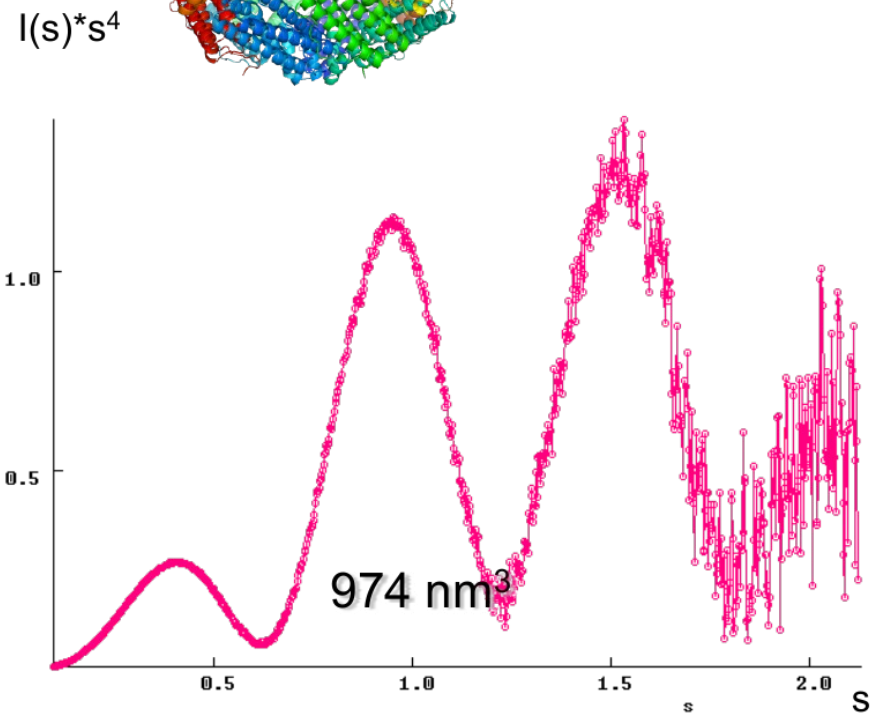
$$V_P = \frac{2\pi^2 I(0)}{\int_0^\infty [I(q) - K_4]\, q^2\, dq}$$

K_4 is a constant determined to ensure the asymptotic decay of $I(q)$ is proportional to q^{-4}







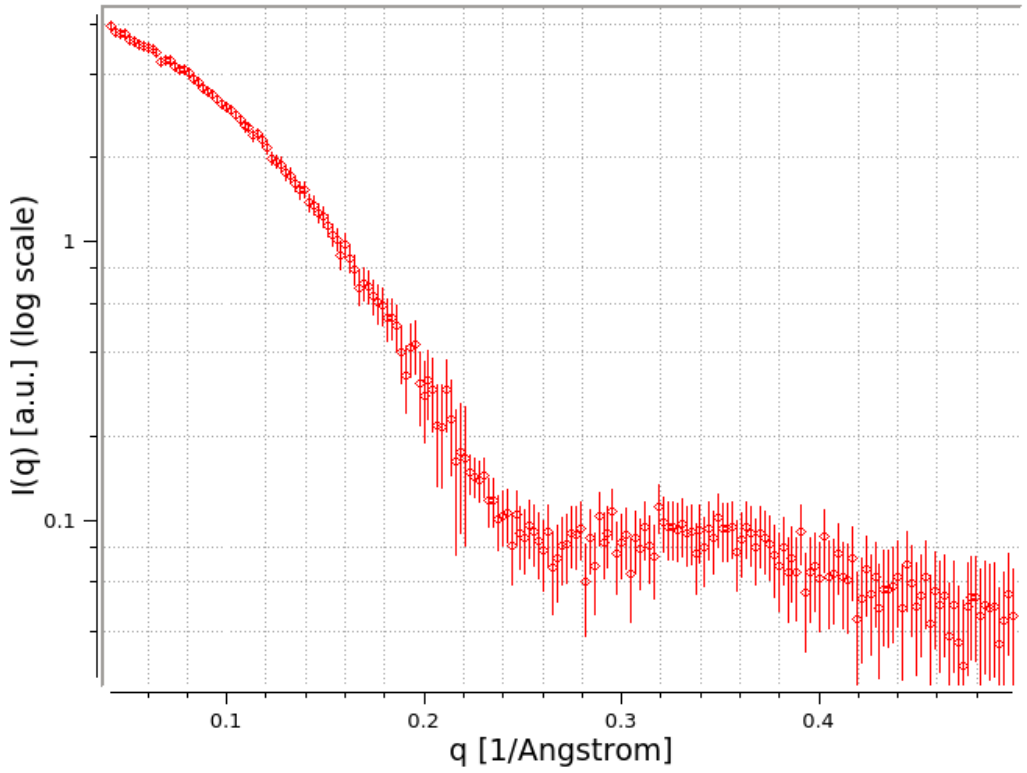


Images courtesy Al Kikhney, EMBL

Information in a SAS curve

Svergun, D.I. & Koch, M.H.J. (2003) Small-angle scattering studies of biological macromolecules in solution. Rep. Prog. Phys. 66 1735-82

- Shannon channels = $D_{max} \cdot q\text{-}range / \pi$
- "the number of [obtainable parameters] typically does not exceed **10–15**"



Lyzosyme $D_{max} \sim 48$ Angstroms

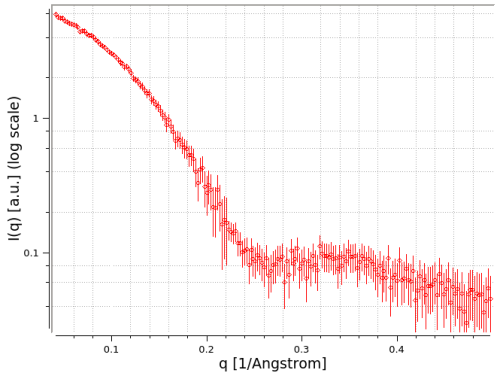
Shannon channels = $48 * 0.5 / \pi \sim 8$

Information in a SAS curve

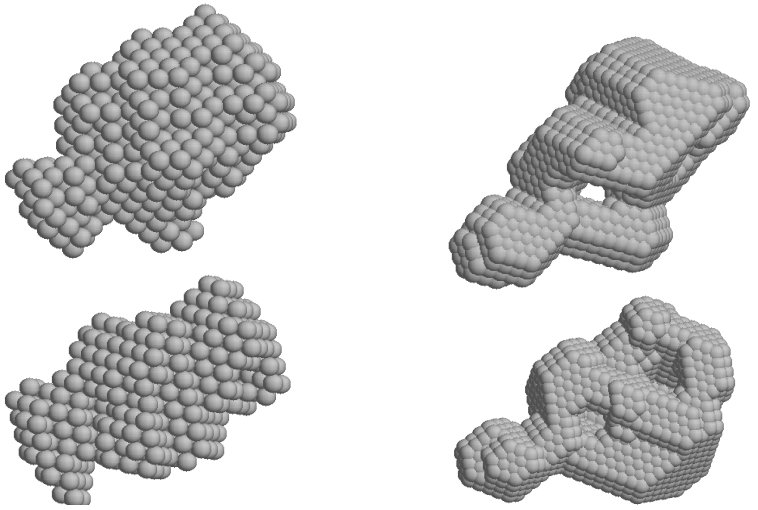
V. V. Volkov and D. I. Svergun (2003). Uniqueness of ab-initio shape determination in small-angle scattering. J. Appl. Cryst. 36, 860-864.

DAMAVER is a set of programs to align ab initio low resolution models (e.g. provided by DAMMIN, DAMMIF and/or GASBOR), select the most typical ("probable") one and build an averaged model.

- DAMSEL: compare all models, find most probable one and outliers
- DAMSUP: align all models with the most probable one
- DAMAVER: average aligned models and compute probability map
- DAMFILT: filter the averaged model at a given cut-off volume
- DAMSTART: generates from the averaged model an input file with fixed core for DAMMIN (for those who want to refine the averaged model)

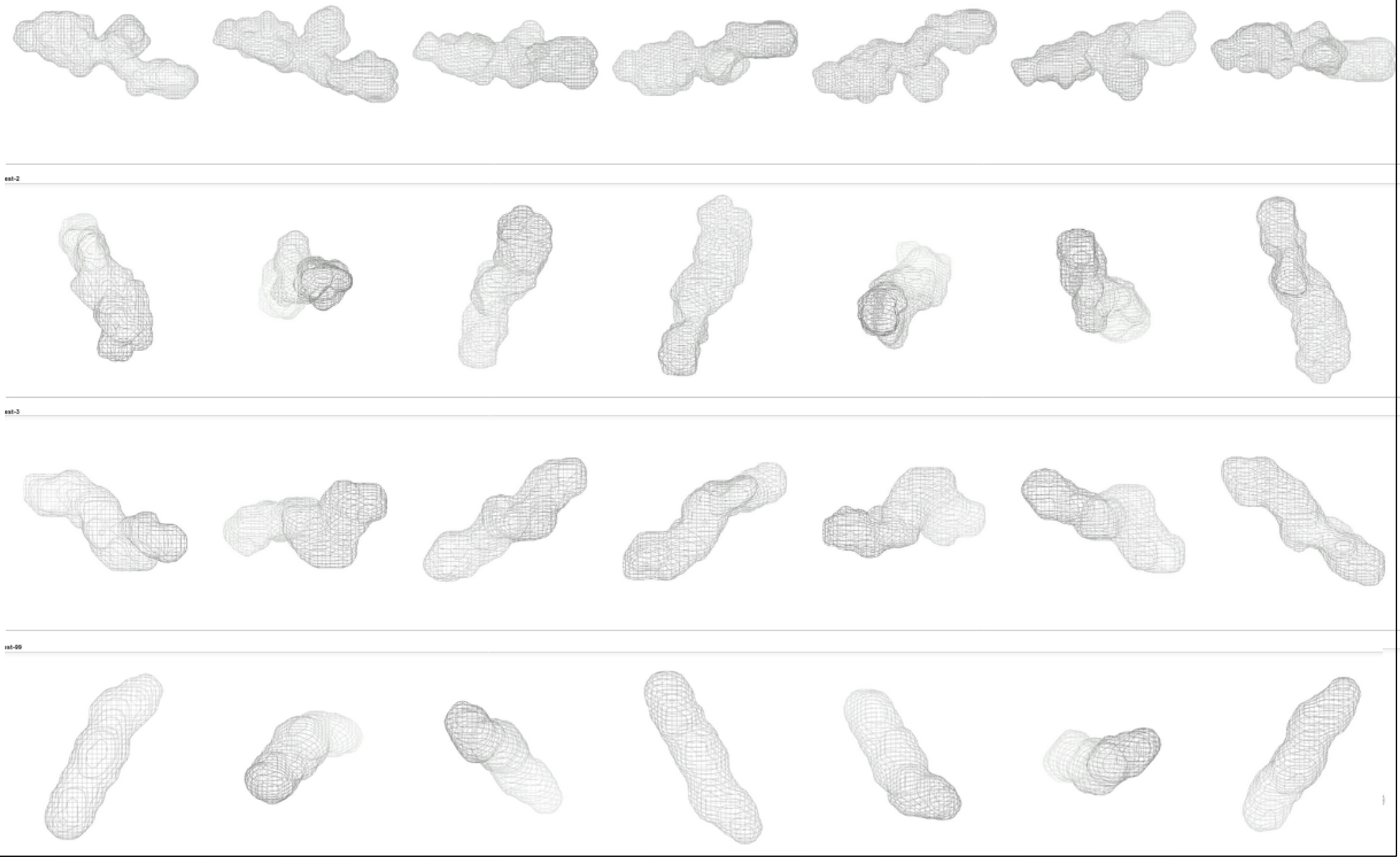


→

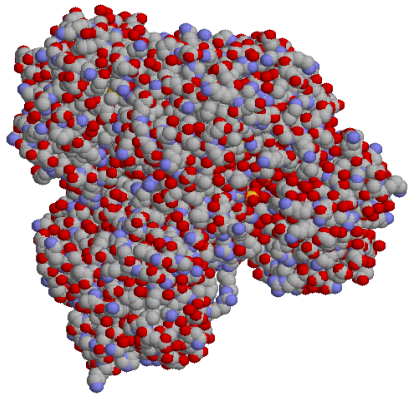


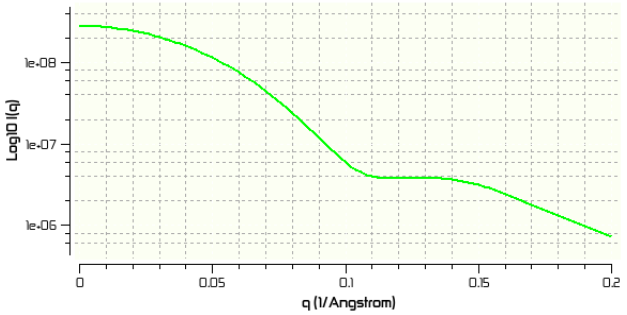
Ab initio models

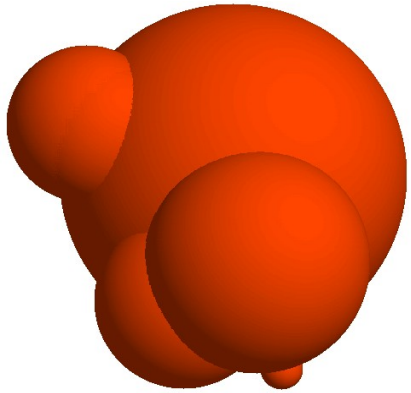
TGFβ / APS ID-12 / DAMMIN



Parsimonious modeling







Brookes, E., Parsimonious Spatial Models from Small Angle Scattering of Biological Macromolecules, SAS 2012, Sydney

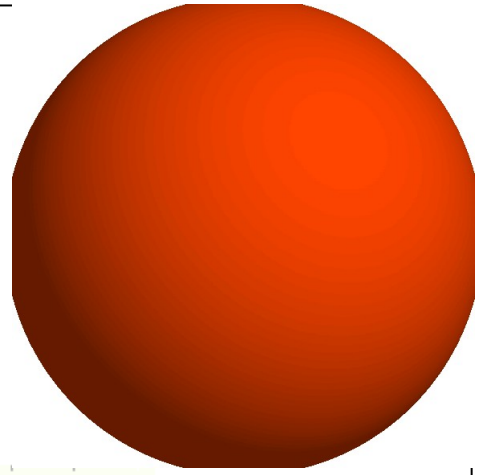
Parsimonious modeling

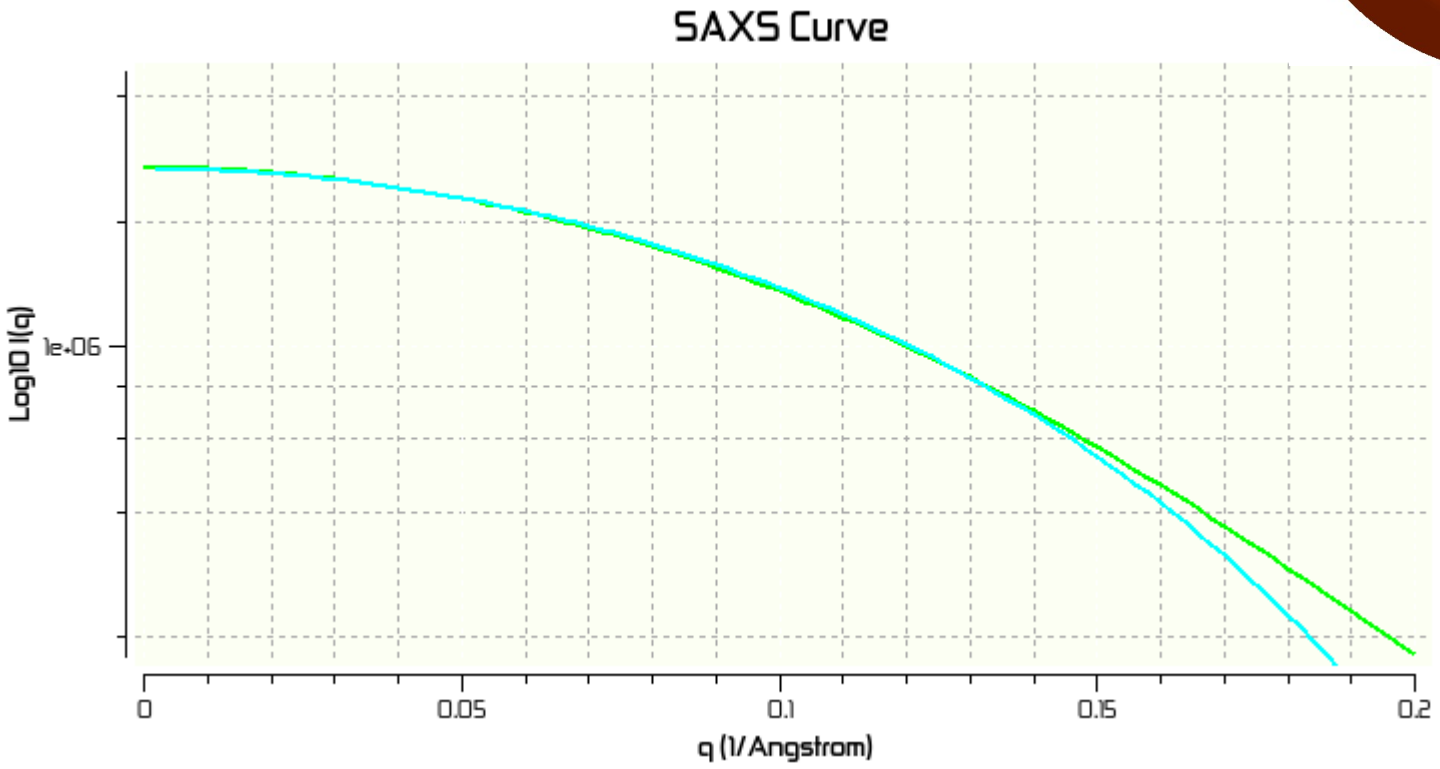
PDB	MW in Daltons	Description
8RAT.PDB	13,683.87	CRYSTALLOGRAPHIC STUDIES OF THE PROTEIN RIBONUCLEASE-A
1A4V.PDB	14,152.00	ALPHA-LACTALBUMIN
1DWR.PDB	17,682.20	MYOGLOBIN (HORSE HEART) WILD-TYPE COMPLEXED WITH CO
1HCO.PDB	32,279.78	HUMAN CARBONMONOXY HAEMOGLOBIN
1BEB.PDB	35,305.26	BOVINE BETA-LACTOGLOBULIN
1CTS.PDB	49,129.58	CITRATE SYNTHASE
2CGA.PDB	51,318.72	BOVINE CHYMOTRYPSINOGEN
1GZX.PDB	64,575.52	OXY T STATE HAEMOGLOBIN: OXYGEN BOUND AT ALL FOUR HAEMS
5LDH.PDB	74,917.32	ACTIVE TERNARY COMPLEX OF PIG HEART LACTATE DEHYDROGENASE WITH S-LAC-NAD
2GD1.PDB	144,427.77	OXIDOREDUCTASE(ALDEHYDE(D)-NAD(A))
1GD1.PDB	147,077.69	HOLO-GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE FROM BACILLUS STEAROTHERMOPHILUS
1ADO.PDB	157,287.20	FRUCTOSE 1,6-BISPHOSPHATE ALDOLASE FROM RABBIT MUSCLE
1OVA.PDB	169,965.56	UNCLEAVED OVALBUMIN

Brookes, E., Parsimonious Spatial Models from Small Angle Scattering of Biological Macromolecules, SAS 2012, Sydney

Parsimonious modeling

1A4V - 1 sphere

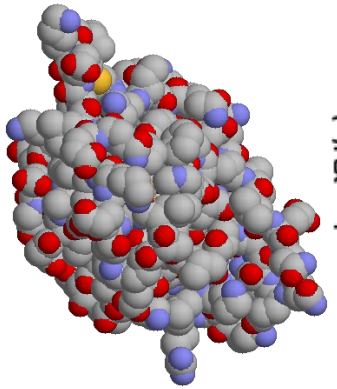


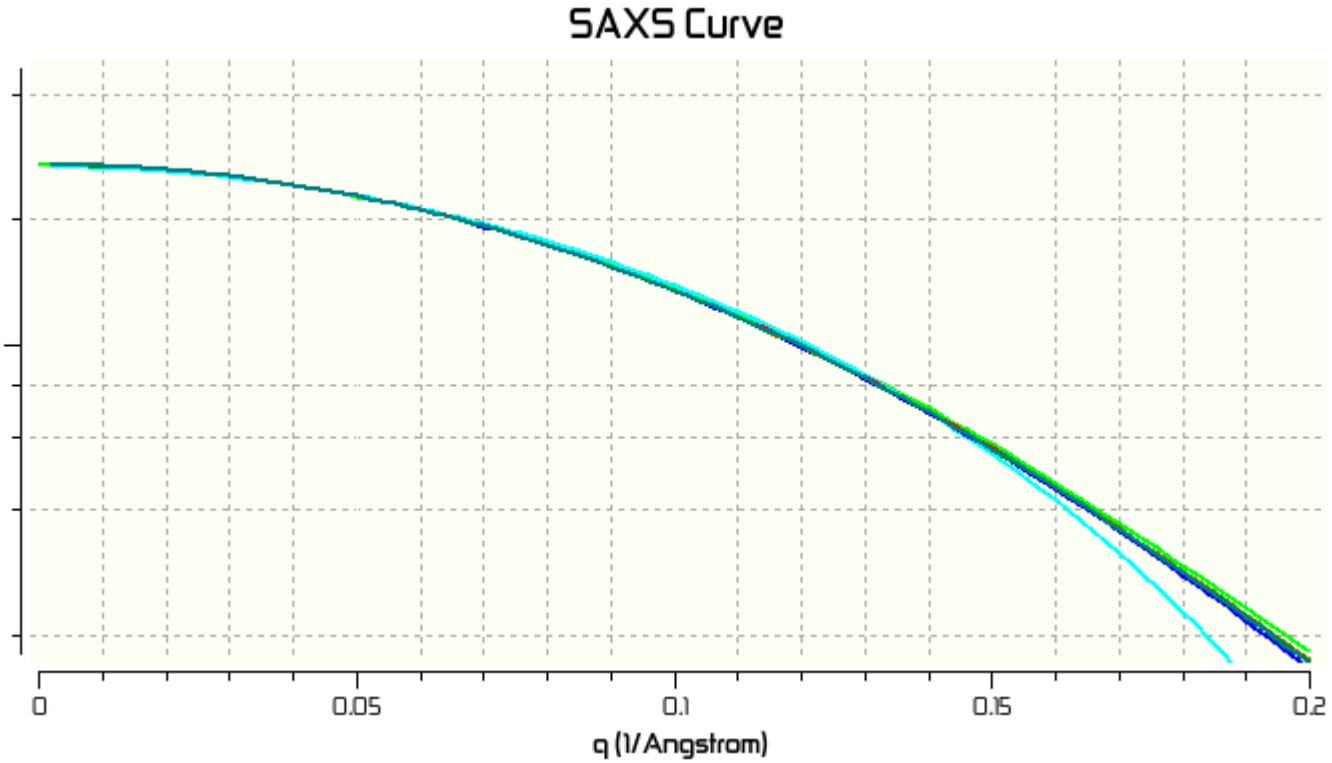


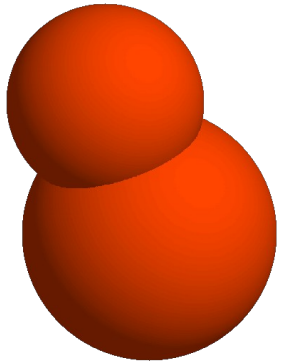
Brookes, E., Parsimonious Spatial Models from Small Angle Scattering of Biological Macromolecules, SAS 2012, Sydney

Parsimonious modeling

1A4V - 7 spheres





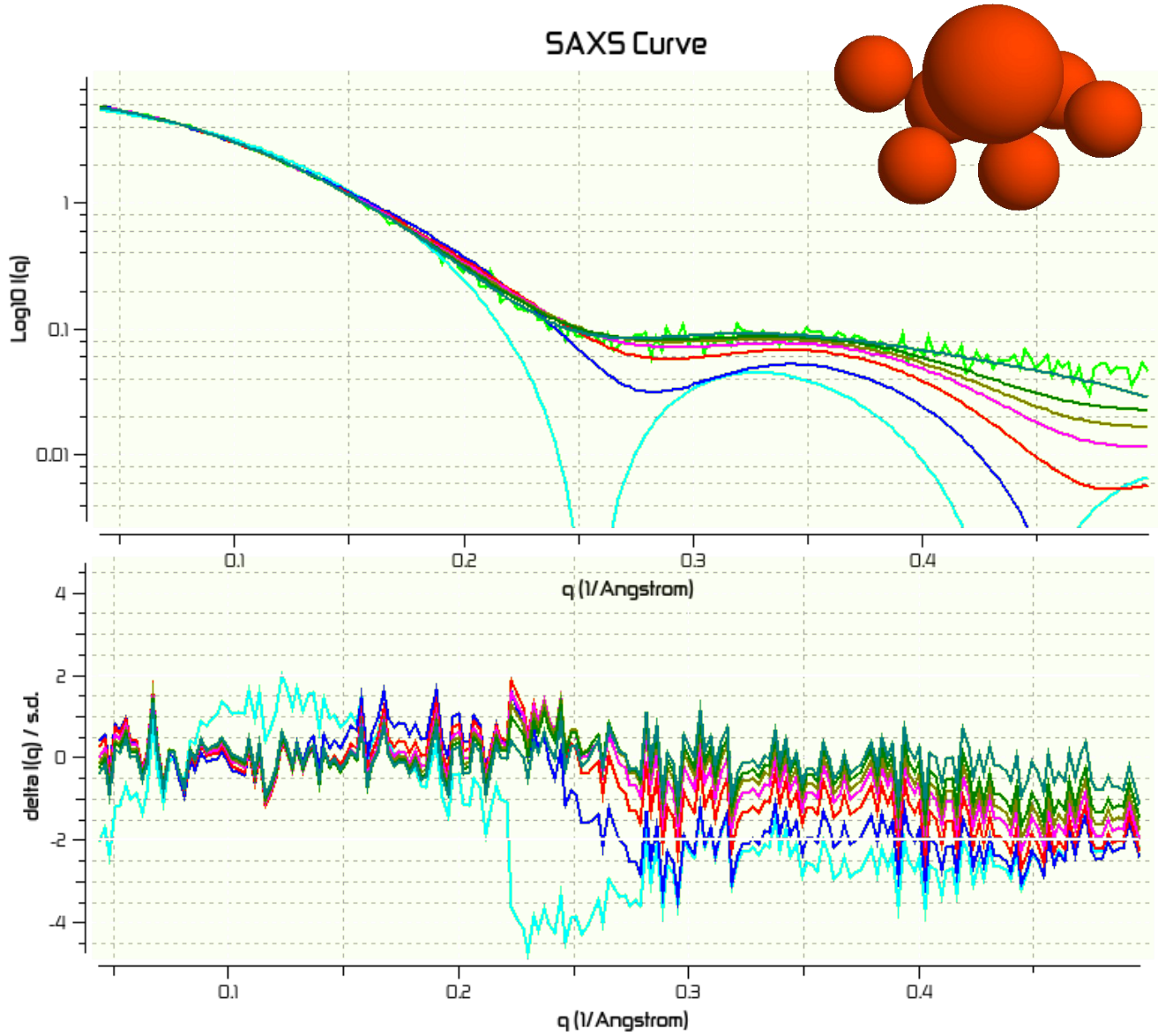


Model name	D(tr) [cm/sec^2]	Rg [nm]	Max extensions X [nm]	Y [nm]	Z [nm]	Axial ratios X:Z	X:Y	Y:Z
1A4V_1-db_1sa-10_1	1.210e-06	1.37	3.54	3.54	3.54	1.00	1.00	1.00
1A4V_1-db_2sa-10_1	1.200e-06	1.45	4.74	3.15	3.15	1.50	1.50	1.00
1A4V_1-db_3sa-10_1	1.140e-06	1.52	4.69	3.50	2.94	1.60	1.34	1.19
1A4V_1-db_4sa-10_1	1.170e-06	1.48	4.67	3.15	3.08	1.51	1.48	1.02
1A4V_1-db_5sa-10_1	1.140e-06	1.51	4.70	4.01	3.23	1.46	1.17	1.24
1A4V_1-db_6sa-10_1	1.140e-06	1.50	4.87	3.69	3.28	1.49	1.32	1.12
1A4V_1-db_7sa-10_1	1.160e-06	1.49	4.74	3.21	3.21	1.48	1.48	1.00
1A4V_1-so	1.137e-06	1.48	5.67	3.54	3.36	1.69	1.60	1.05

Brookes, E., Parsimonious Spatial Models from Small Angle Scattering of Biological Macromolecules, SAS 2012, Sydney

Parsimonious modeling

- n-vector of spheres
- arbitrarily placed
- *I(q)* reproduced with by ~7 spheres for test cases
- Physical extents reasonably matched
- overlap intensity issue
- connectivity



Brookes, E., Parsimonious Spatial Models from Small Angle Scattering of Biological Macromolecules, SAS 2012, Sydney

Parsimonious modeling

- Extensible pool of parametrized shapes
  - Spheres, spheroids, ellipsoids, cylinders, tori
- Input I(q), chosen shapes from pool
- → Parsimonious model

Brookes, E., Progress in Parsimonious Spatial Modeling of Biological SAS Experimental Data, ACA 2014, Albuquerque

Parsimonious modeling

Parsimonious Spatial Modeling

Data files to process — Browse... lyzexp.dat

Main model type controls

Model type list (0=sphere,1=cylinder,2=spheroid,3=ellipsoid,4=torus)	0,1,1,1,1
Compute models for all lengths upto above	☑
Compute models for all unique combinations	☑
Sample electron density [e/Angstrom^3]	0.425
Buffer electron density [e/Angstrom^3]	0.334
Grid size [Angstrom]	4
Automatically compute max	

q range editing control

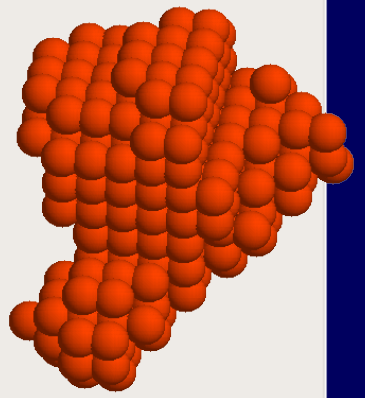
Supplementary control

Genetic algorithm cont

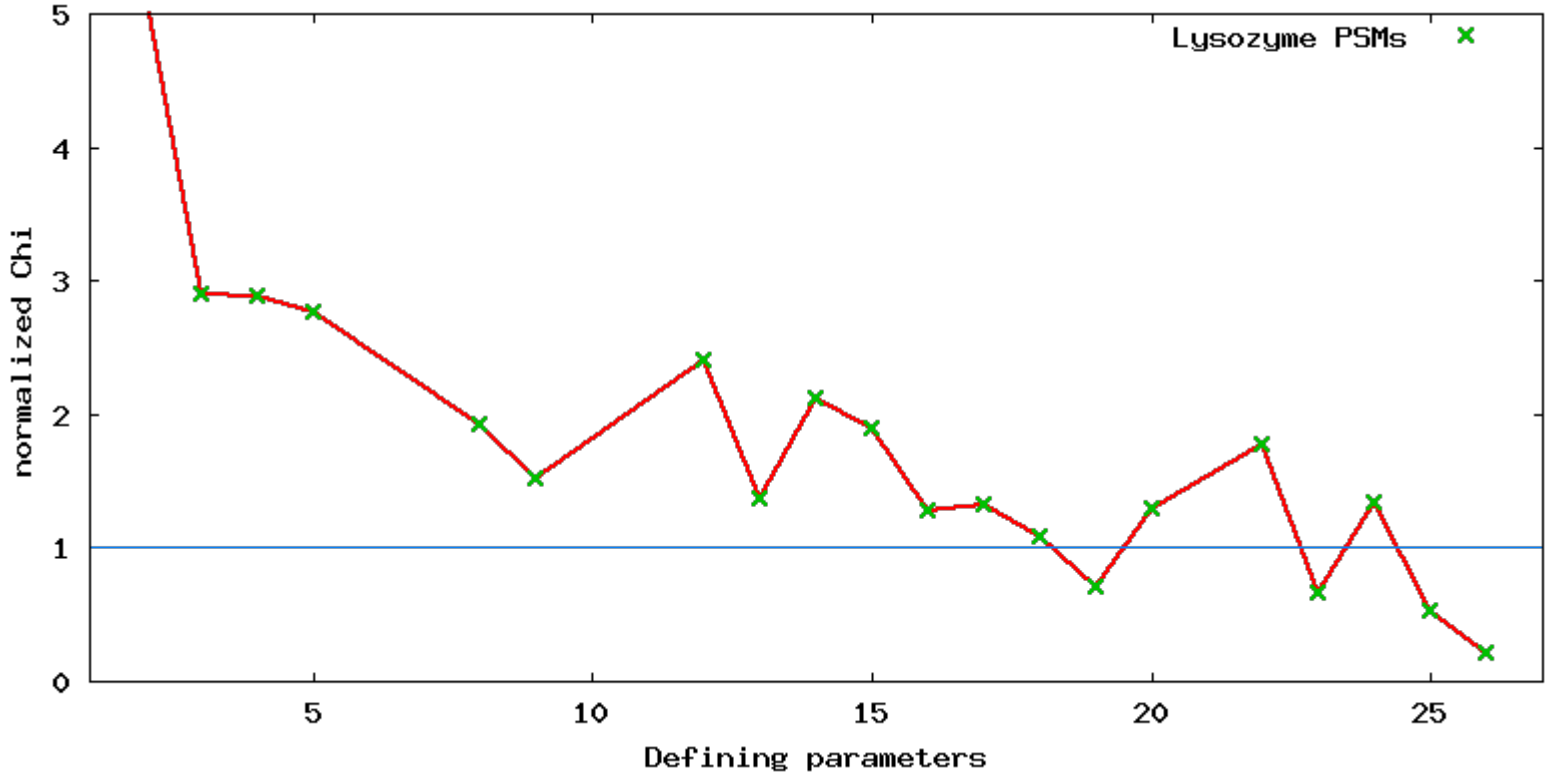
Miscellenaous controls

Submit | Reset to default values

```
started
processing lyzexp.dat
computing approximate maximum dimension of lyzexp
approximate maximum dimension of lyzexp is 26
start GA for lyzexp_sphere
end GA for lyzexp_sphere model nchi 5.197
start GA for lyzexp_cylinder
end GA for lyzexp_cylinder model nchi 3.137
start GA for lyzexp_sphere_sphere
end GA for lyzexp_sphere_sphere model nchi 2.769
start GA for lyzexp_sphere_cylinder
end GA for lyzexp_sphere_cylinder model nchi 2.794
start GA for lyzexp_cylinder_cylinder
end GA for lyzexp_cylinder_cylinder model nchi 2.458
start GA for lyzexp_sphere_sphere_sphere
end GA for lyzexp_sphere_sphere_sphere model nchi 1.915
start GA for lyzexp_sphere_sphere_cylinder
end GA for lyzexp_sphere_sphere_cylinder model nchi 2.341
start GA for lyzexp_sphere_cylinder_cylinder
end GA for lyzexp_sphere_cylinder_cylinder model nchi 1.097
start GA for lyzexp_cylinder_cylinder_cylinder
end GA for lyzexp_cylinder_cylinder_cylinder model nchi 0.3411
start GA for lyzexp_sphere_sphere_sphere_sphere
```

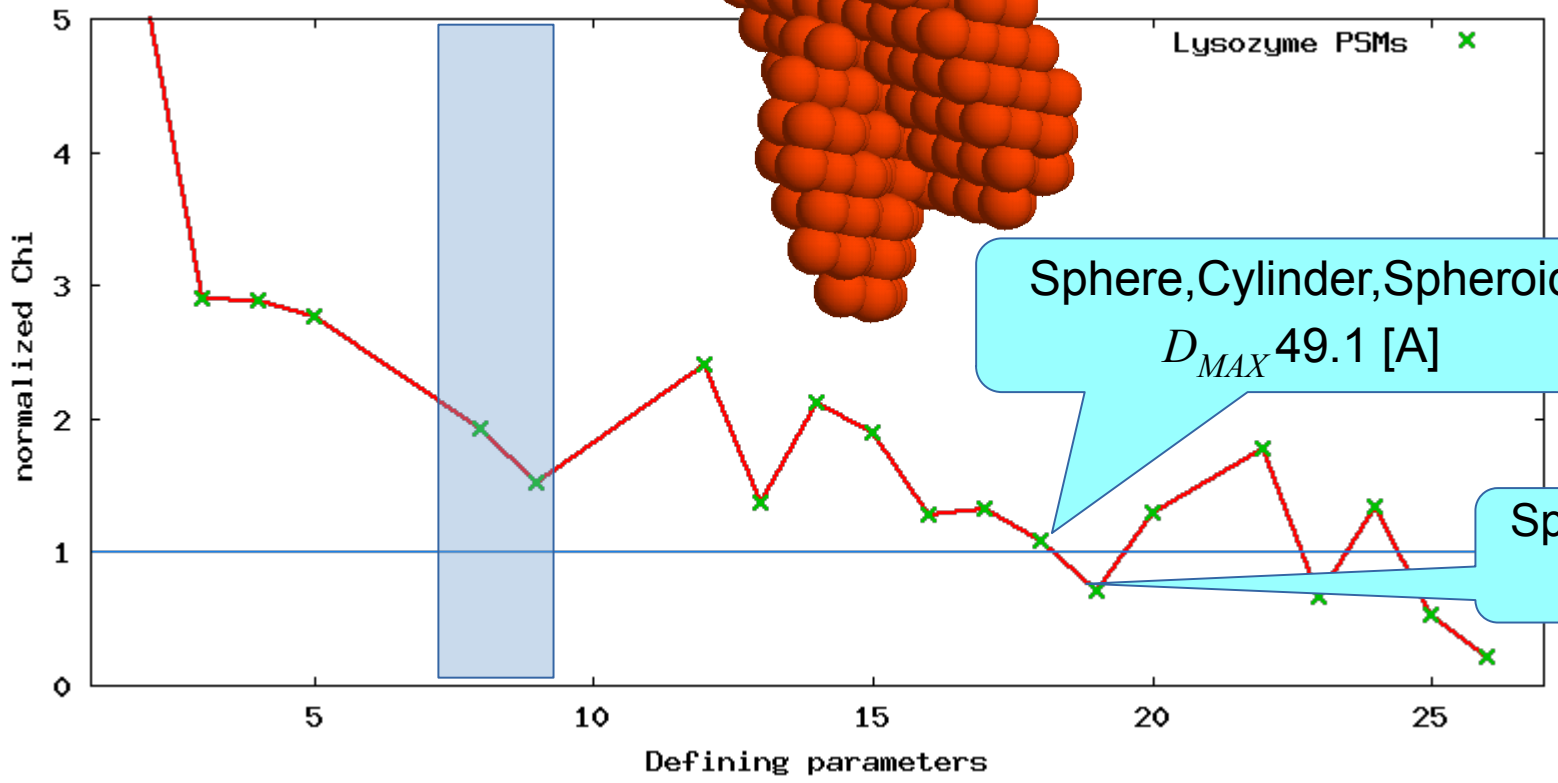


Parsimonious modeling

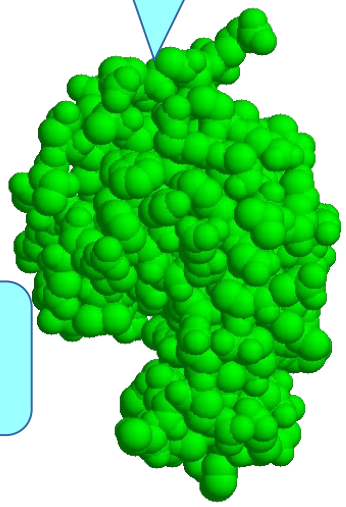


Brookes, E., Progress in Parsimonious Spatial Modeling of Biological SAS Experimental Data, ACA 2014, Albuquerque

Parsimonious modeling

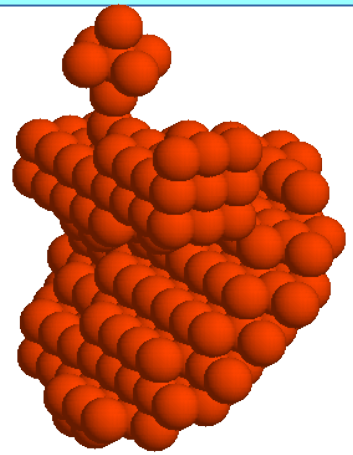


1HEL.pdb
D_{MAX} 48 A



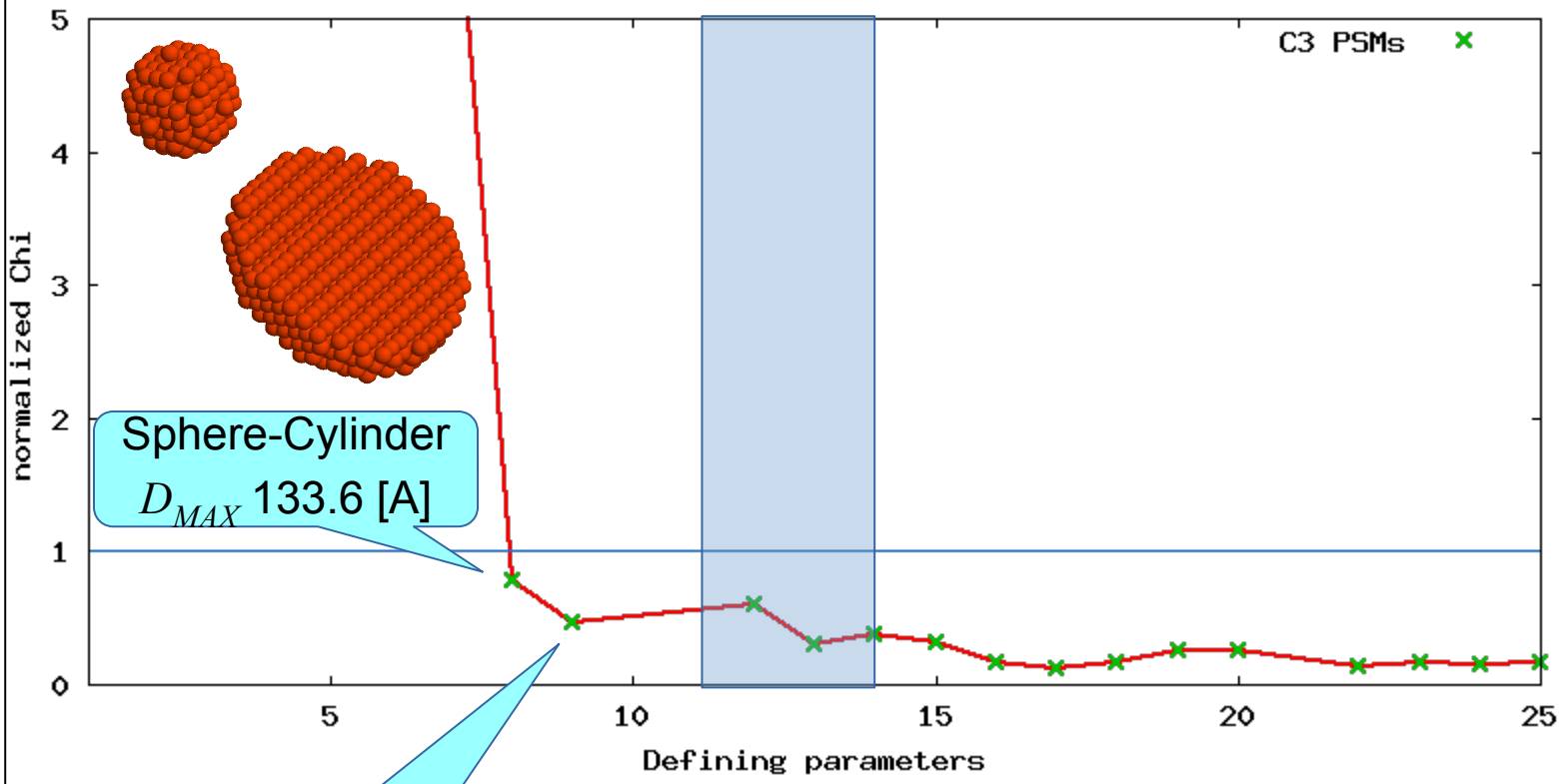
Sphere,Cylinder,Spheroid
D_{MAX} 49.1 [A]

Sphere,Cylinder,Ellipsoid
D_{MAX} 49.2 [A]



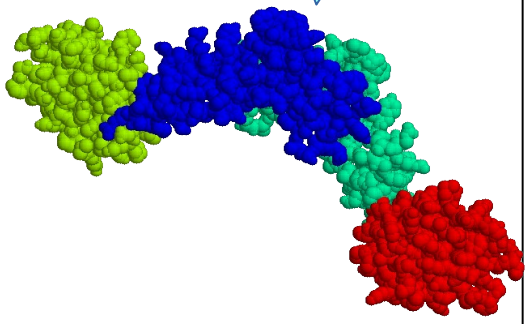
Brookes, E., Progress in Parsimonious Spatial Modeling of Biological SAS Experimental Data, ACA 2014, Albuquerque

Parsimonious modeling

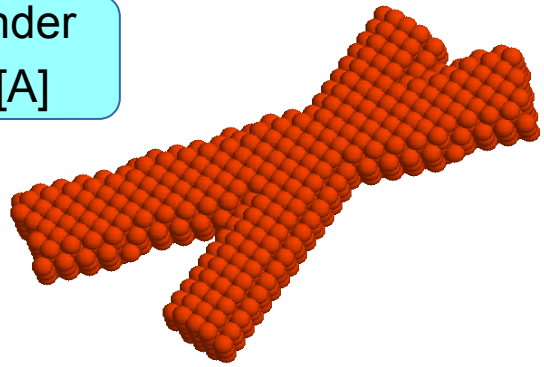


Sphere-Cylinder
D_{MAX} 133.6 [A]

Homology Model
D_{MAX} 142 [A]



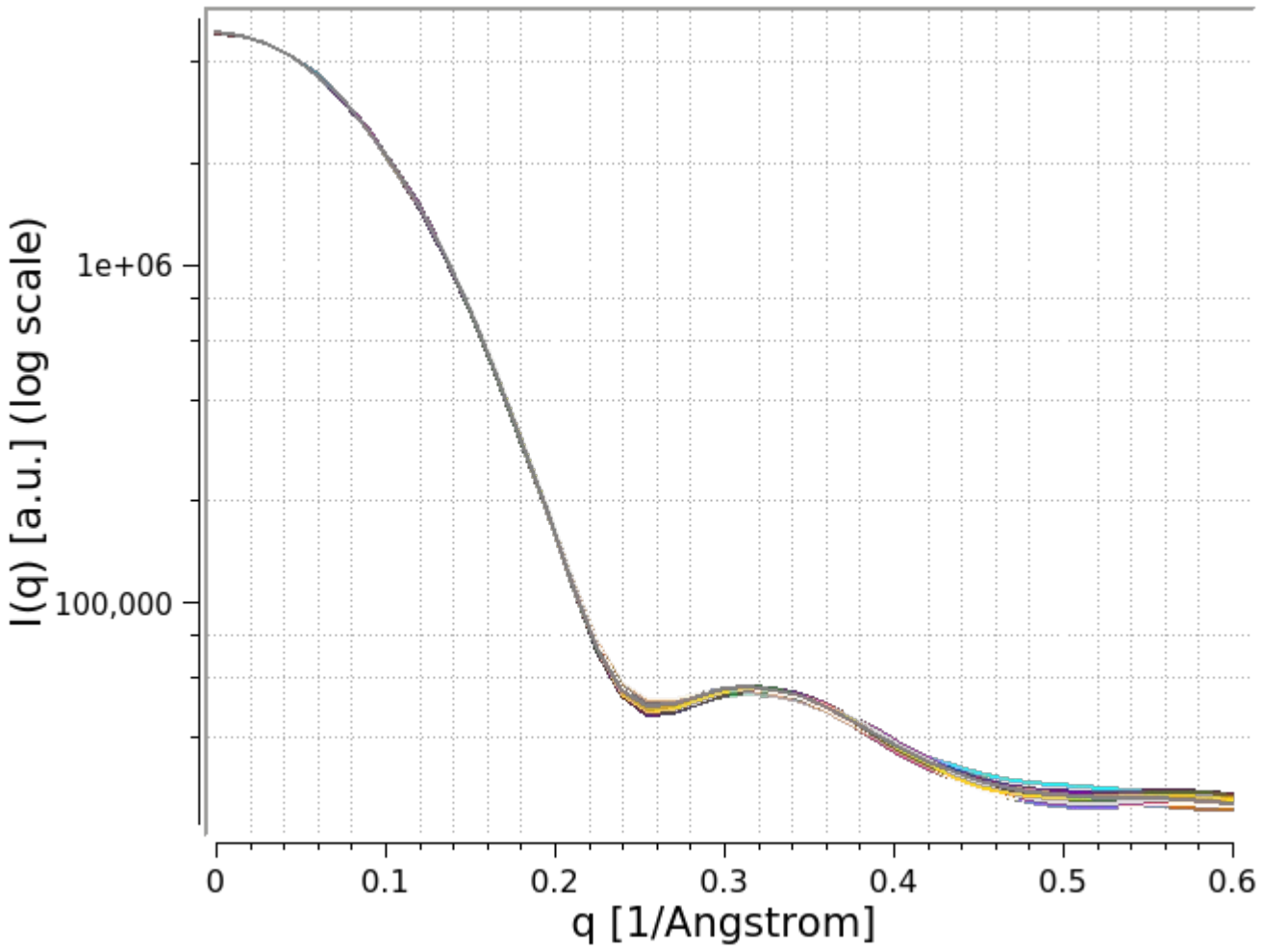
Cylinder-Cylinder
D_{MAX} 138.6 [A]



Brookes, E., Progress in Parsimonious Spatial Modeling of Biological SAS Experimental Data, ACA 2014, Albuquerque

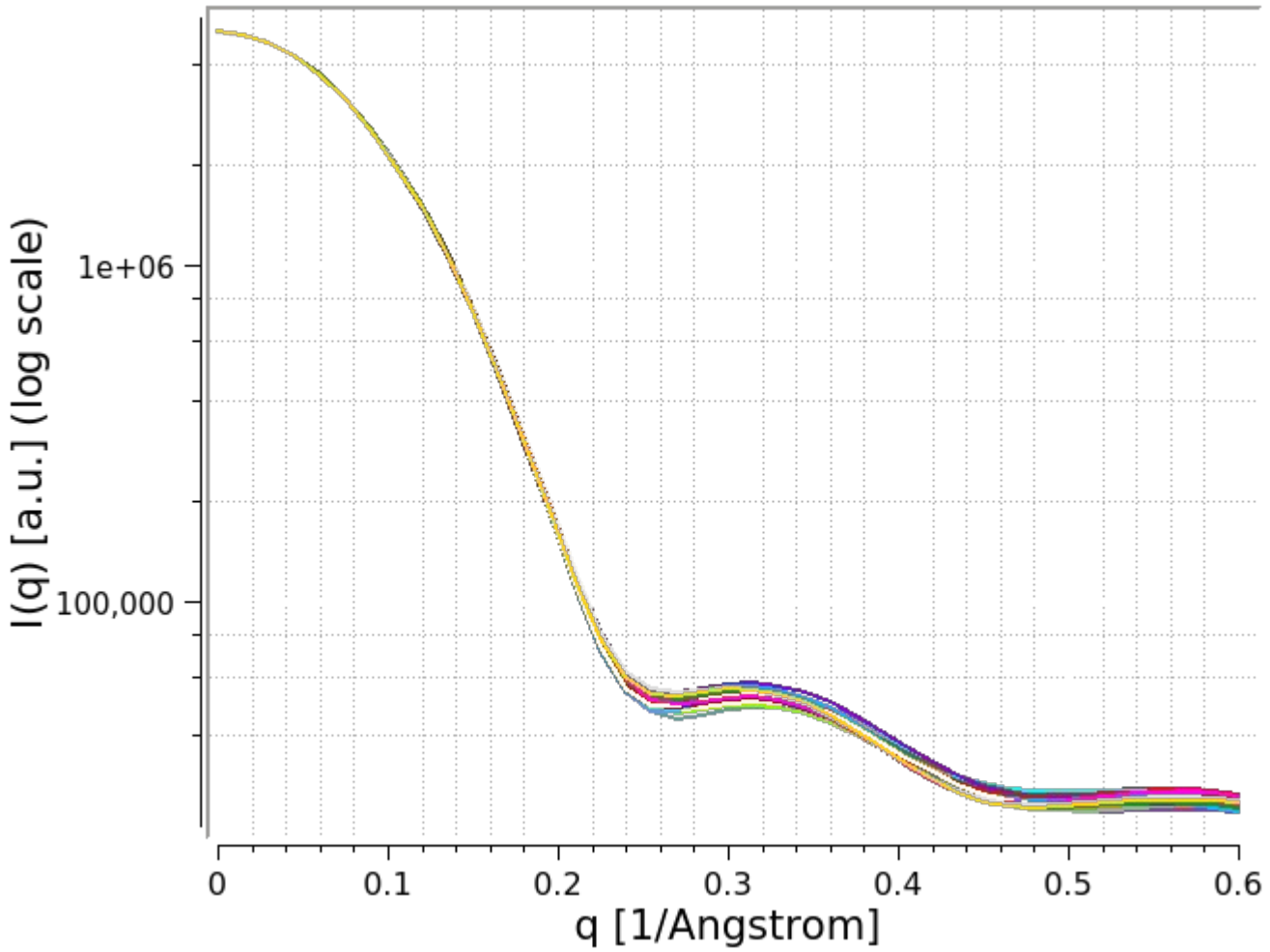
SAS is very sensitive

1HEL
50 frames DMD
10 ps/frame



SAS is very sensitive

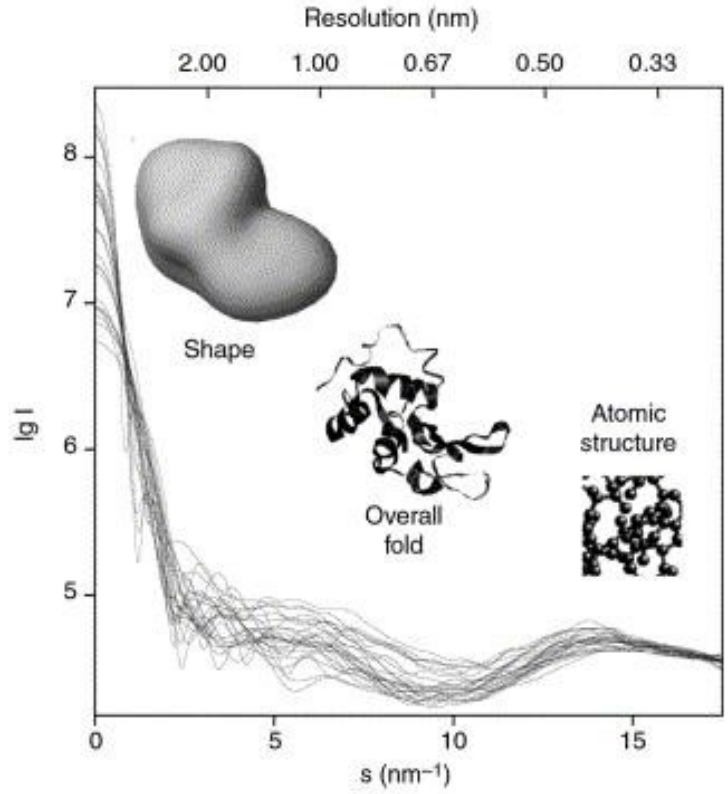
1AKI
48 frames NAMD
100mM NaCl
10 ps/frame



Information in a SAS curve

Low IC: $D_{max} \cdot q\text{-}range / \pi$

High sensitivity

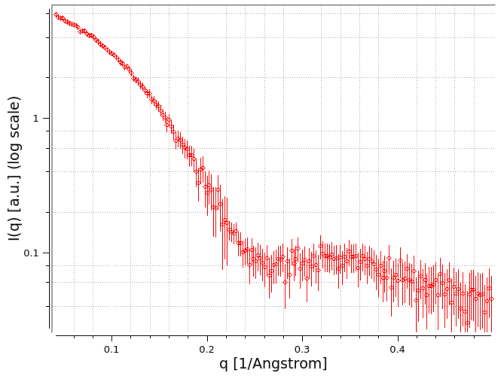


GFDL, https://en.wikipedia.org/w/index.php?curid=13264733

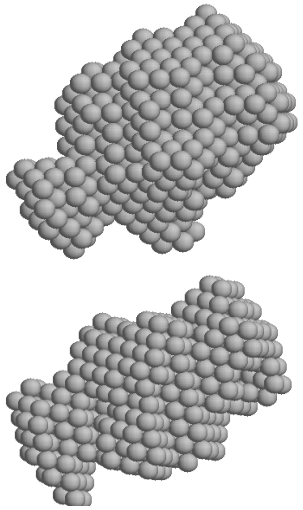
Ab-inito reconstruction

DAMMIN

D. I. Svergun (1999) Biophys J. 2879-2886.

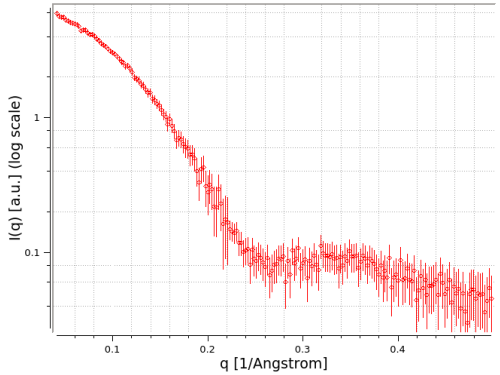


→

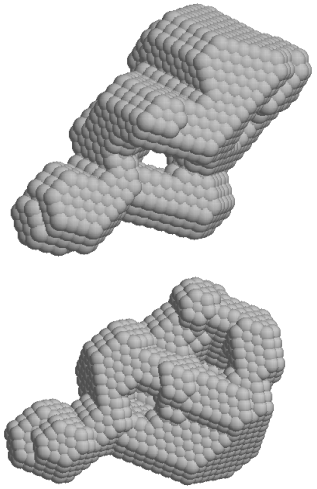


DAMMIF

Franke, D. and Svergun, D.I. (2009) J. Appl. Cryst., 42, 342-346.



→

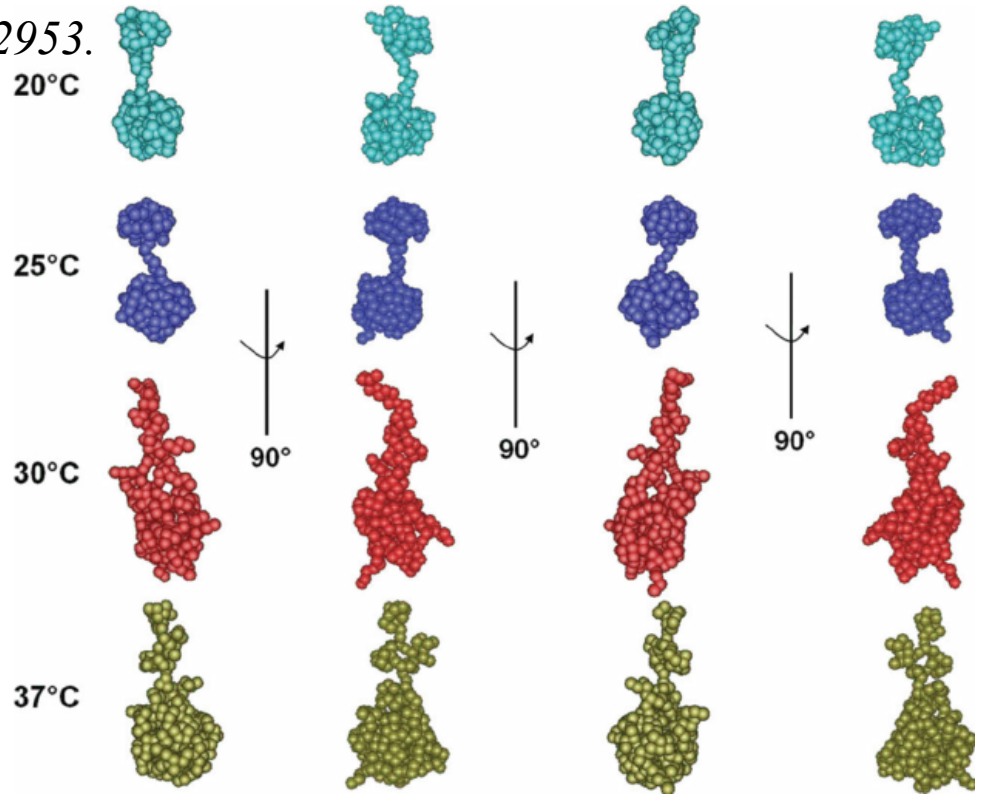


Ab-inito reconstruction

GASBOR

Svergun, D.I. et al. (2001) Biophys. J., 80, 2946-2953.

ab initio reconstruction of protein structure
by a chain-like ensemble of dummy residues



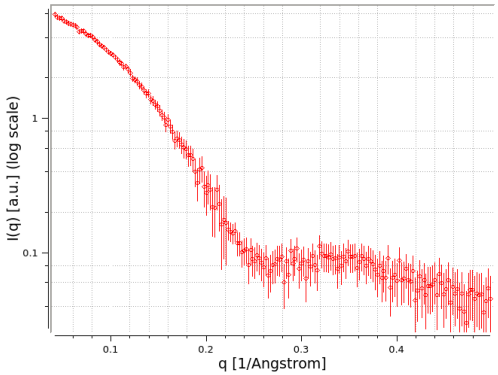
Various views of the ab initio 3D models obtained using GASBOR and by averaging ten single models for each sample by using DAMAVER for rSdrFB1-4 at different temperatures.

Dipoto, Antonella et al., (2015). Appl. Microbiology and Biotech.

Ab-inito reconstruction

DENSS

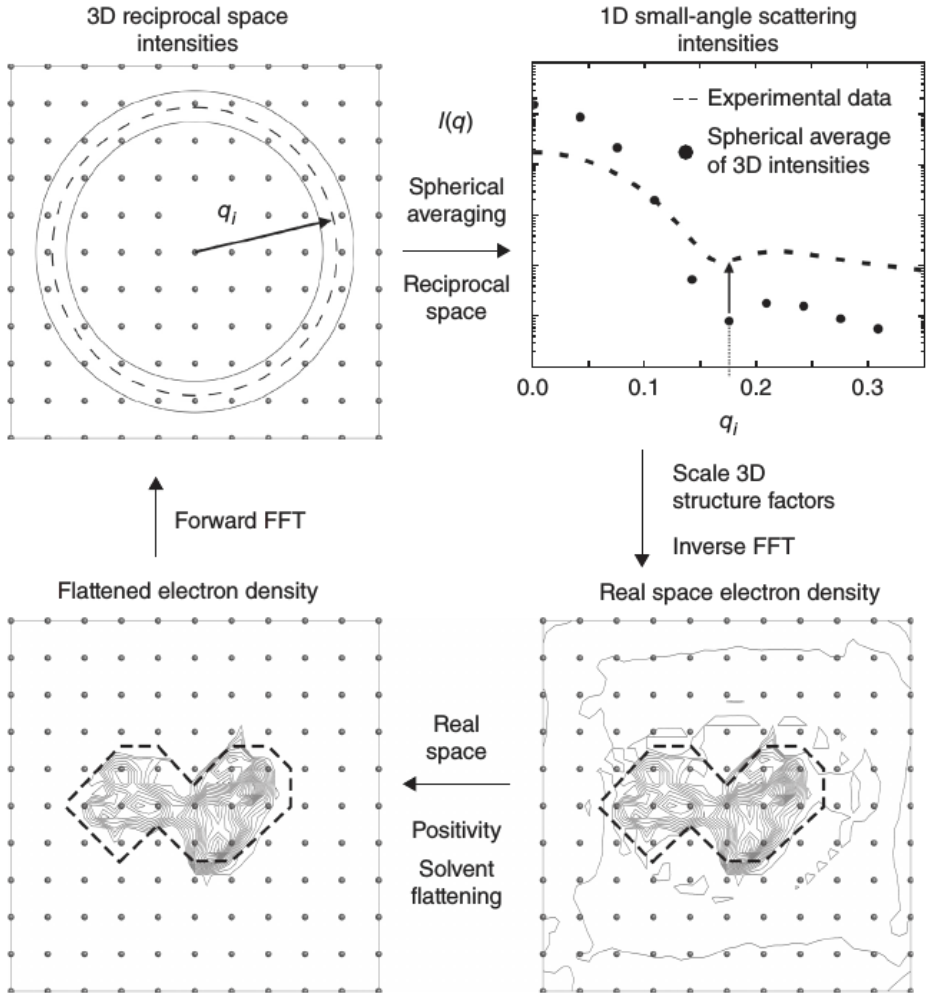
T. D. Grant (2018) Nat. Meth. 15:3 191-193



$$f(q) = \int_{V_\mathbf{r}} \rho(\mathbf{r})\, dV_\mathbf{r}$$

$$\mathbf{F(q)} = \sum_{i=1}^{N} f_i\, \mathbf{e}^{\mathbf{i} r_i q}$$

$$\mathbf{I(q)} = \mathbf{F(q)} \cdot \mathbf{F^*(q)}$$



As with DAMMIN, DAMMIF, GASBOR →stochastic, averaging of multiple models

Information content revisited

Jochen S Hub. Curr. Op. in Struct. Bio. 2018, 49:18-26

“the interpretation of solution scattering data by computational methods is complicated by the low information content of the data, by scattering contributions from the hydration layer, and by unknown systematic errors.”

“The physical information in atomistic force fields complements the low-information SWAXS data; explicit-solvent MD may be used to predict solvent scattering, and the MD-related sampling methods may guide the structure refinement against SWAXS data.”

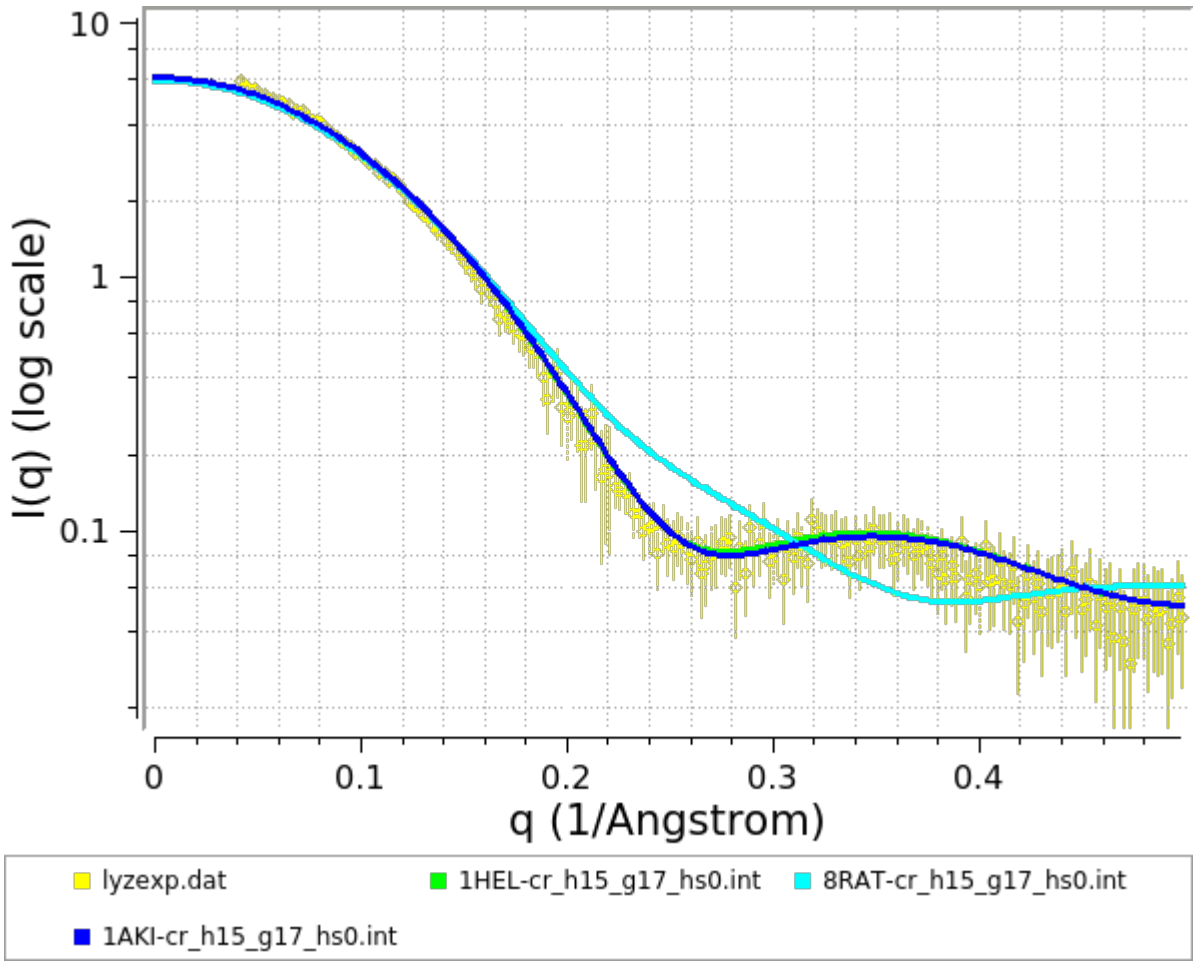
“Because SWAXS curves are smooth and one-dimensional (1D), they contain quite a limited amount of information. How the information is distributed over the q-range is a matter of ongoing research, but it is generally accepted that experimental SWAXS curves do not contain more than 10–30 independent data points. Hence, the number of backbone angles of biomolecules exceeds the number of independent data points of SWAXS curves by roughly two orders of magnitude. This precludes any straightforward fitting of protein structures against SWAXS data, but instead it leads to a high risk of overfitting.”

Comparing models against data

1HEL

8RAT

1AKI

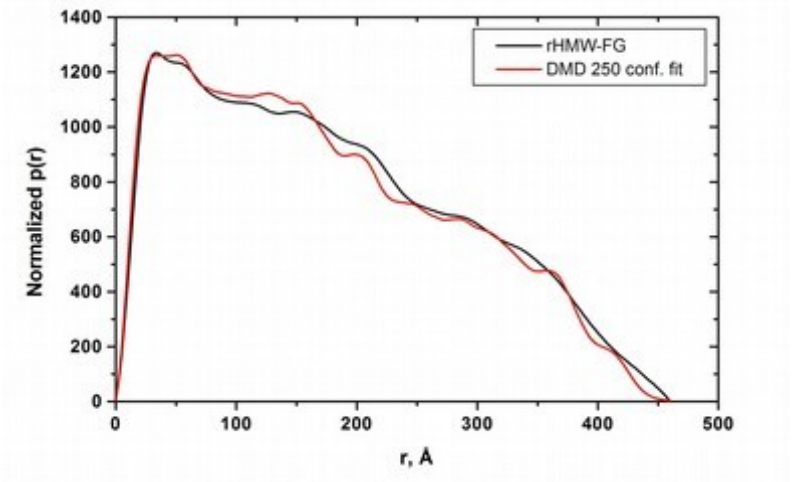


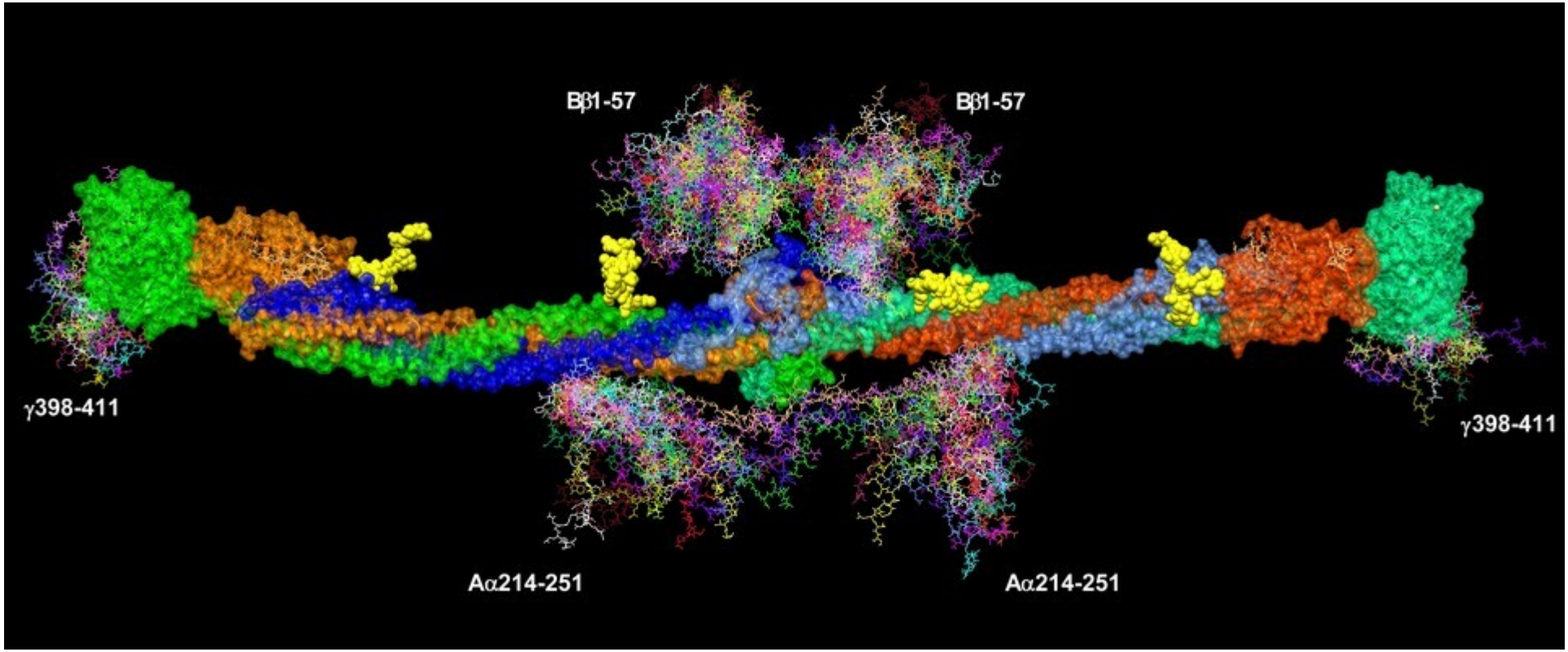
Sorted best fit results
0.914138 : "1HEL-cr_h15_g17_hs0_03.int"
0.922224 : "1AKI-cr_h15_g17_hs0_03.int"
1.2017 : "8RAT-cr_h15_g17_hs0_03.int"
Best fit model: "1HEL-cr_h15_g17_hs0_03.int" nchi 0.914138

NNLS results:
1AKI-cr_h15_g17_hs0_03.int 0
1HEL-cr_h15_g17_hs0_03.int 0.726835
8RAT-cr_h15_g17_hs0_03.int 0.273165
Scaling factor: 1 chi^2=143.506 df=195
nchi=0.857863

Comparing models against data

Mattia Rocco et al.
Fibrinogen is an important component of the coagulation cascade, as well as a major determinant of blood viscosity and blood flow
A centrosymmetric dimer made by 3 pairs of chains
US-SOMO/DMD simulations of the conformational variability for comparison to experimental data

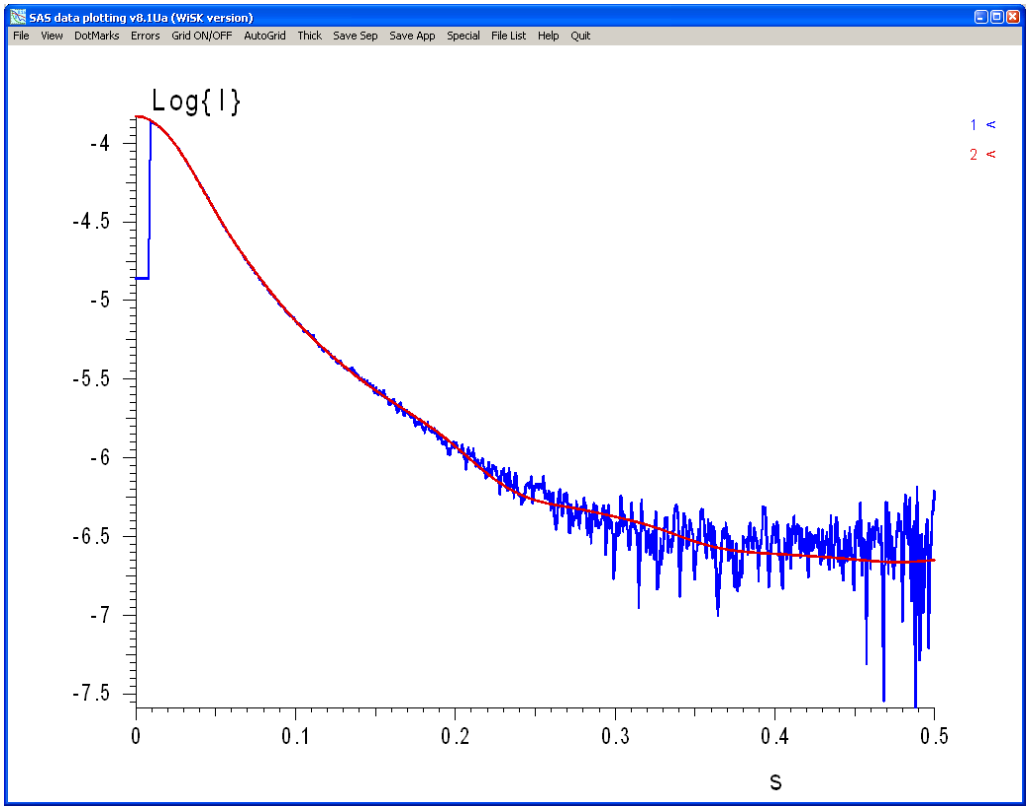


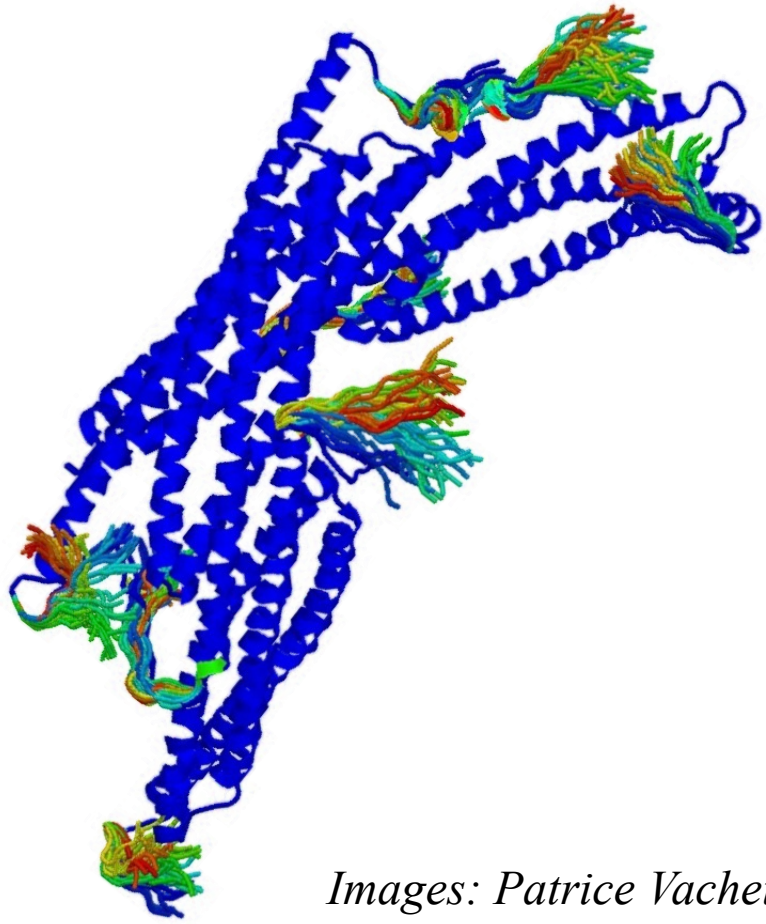


Images: Mattia Rocco

Comparing models against data

Group of M. Milani, University of Milano
Smac-DIABLO a dimeric protein involved in apoptosis (programmed cell death)
chain: 192 residues, MM= 21.8 kDa
Final refinement of model by addition of N- and C-termini using US-SOMO/DMD



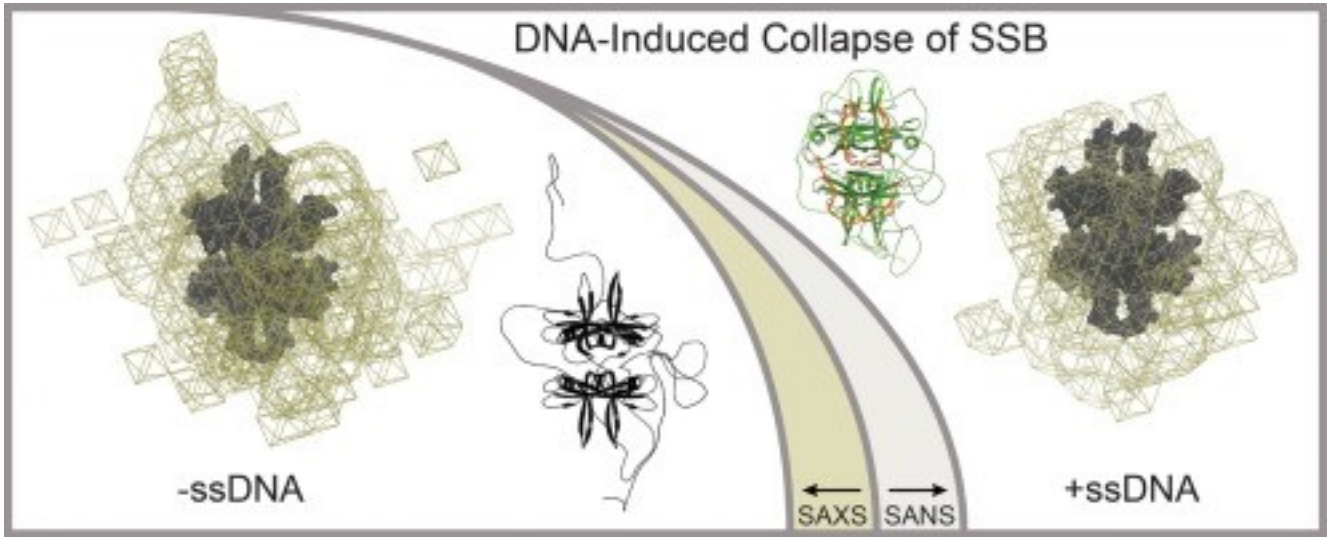


Images: Patrice Vachette

Mastrangelo, Eloise et al. Biophys. J. vol. 108,3 (2015): 714-23. doi:10.1016/j.bpj.2014.11.3471

Comparing models against data

Mathew Green et al, University of Nottingham
SSB binds individual strands of DNA
Critical role in DNA metabolism: Replication, recombination & repair
Intrinsically disordered
US-SOMO/DMD used to create conformations to screen against SAXS data



Matthew Green et al. J. of Mol. Bio., 428:2 357-364 2016

Expanding conformational space

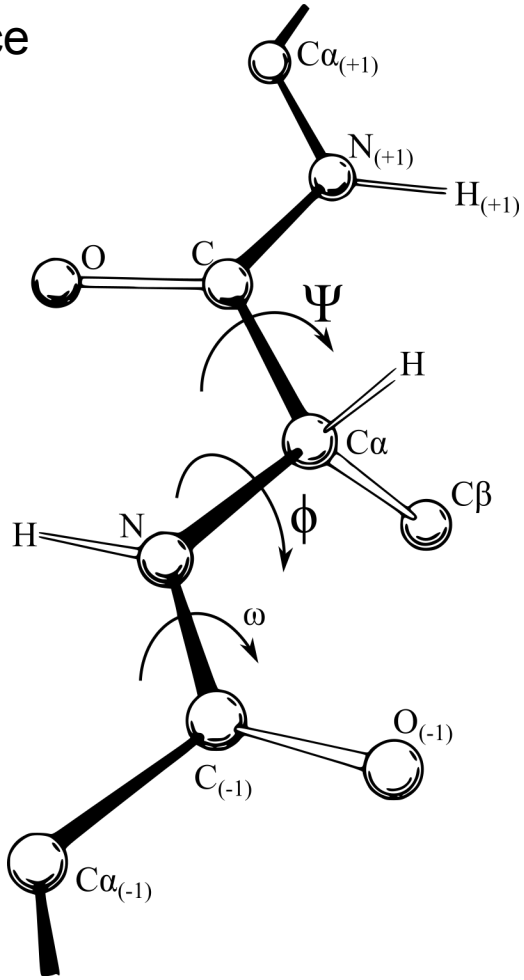
MD – frequently very high computational demand to cover conformational possibilities without restraints

MC – can cover conformational space faster

Often not accessible to a biologist without a steep learning curve…
convenience tools are available to help.

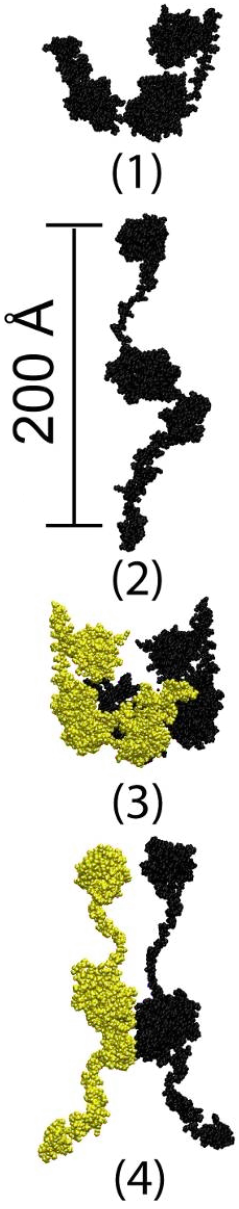
Expanding conformational space – Dihedral angles

Expanding conformational space
Dihedral angle MD & MC



By Dcrjsr, vectorised Adam Rędzikowski - Own work, CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=24585750

HIV-1 Gag

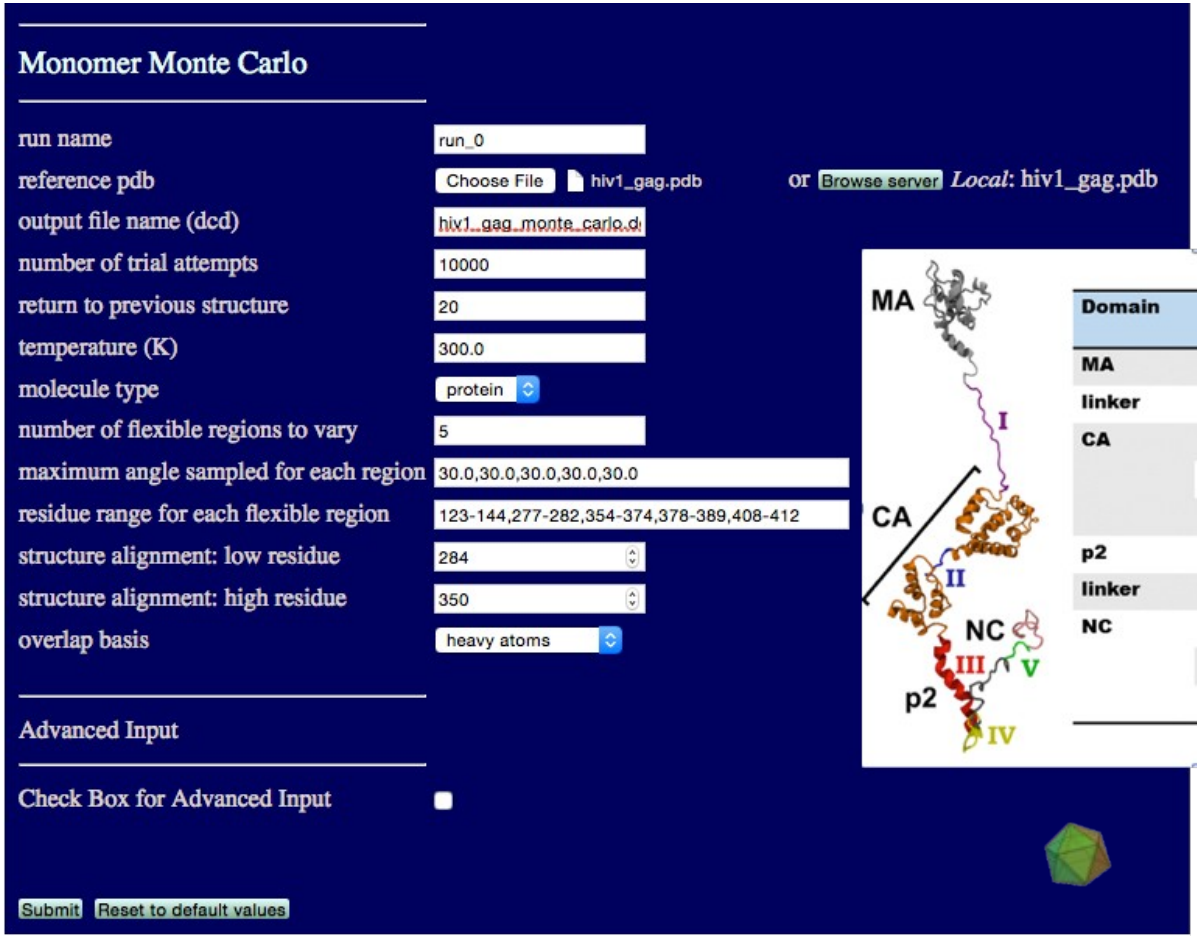


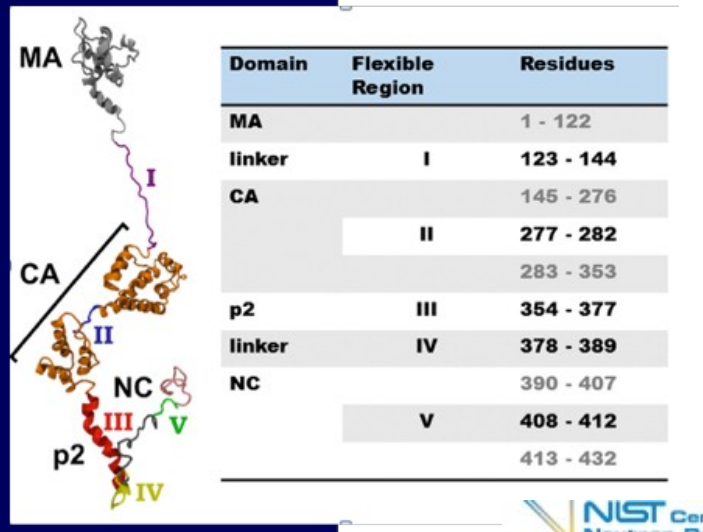
Siddhartha A.K. Datta et al., J Mol Biol. 2011; 406(2): 205–214.

Expanding conformational space - SASSIE

Developed to enable NCNR user community to efficiently develop molecular models for the neutron/X-ray scattering/reflectivity experiments.

www.smallangles.net/sassie





Domain	Flexible Region	Residues
MA		1 - 122
linker	I	123 - 144
CA		145 - 276
	II	277 - 282
		283 - 353
p2	III	354 - 377
linker	IV	378 - 389
NC		390 - 407
	V	408 - 412
		413 - 432

NIST Center for Neutron Research

NIST National Institute of Standards and Technology
Technology Administration, U.S. Department of Commerce

Generate Ensemble

Energy Minimize & Calculate I(q)

Compare to Exp.: Chi-square Filter Density Plot

Joseph E. Curtis et al., Comp. Phys. Comm., 2012, 183(2), 382–389

Rigid body modeling

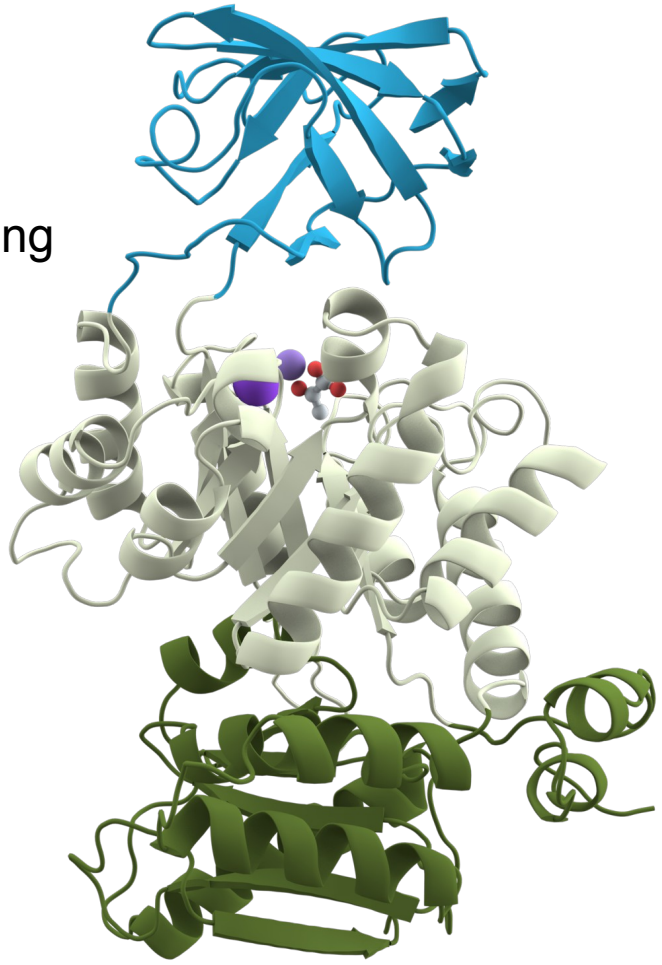
Structure of subunits are known

Arbitrary complex can be constructed by moving and rotating

Verify no steric clashes

→ scattering data subunits

+ contacts (chemical shifts by NMR or mutagenesis)
+ distances between residues (FRET or mutagenesis)
+ relative orientation (RDC by NMR)



Pyruvate kinase 1PKN

Petoukhov et al. 2006 Eur. Biophys J 35:567

Software landscape

Software for “data reduction”, “visualization”, “model fitting”, various “analysis” …
Grouped packages and stand alone components

ATSAS – Dmitri Svergun group
Scatter – Rob Rambo
BioXTAS Raw – Jesse Hopkins
SASView – multiple contributors
CCP-SAS – SCT/SCTPL / US-SOMO / SASSIE & others – multiple contributors
more at *http://smallangle.org/content/software*

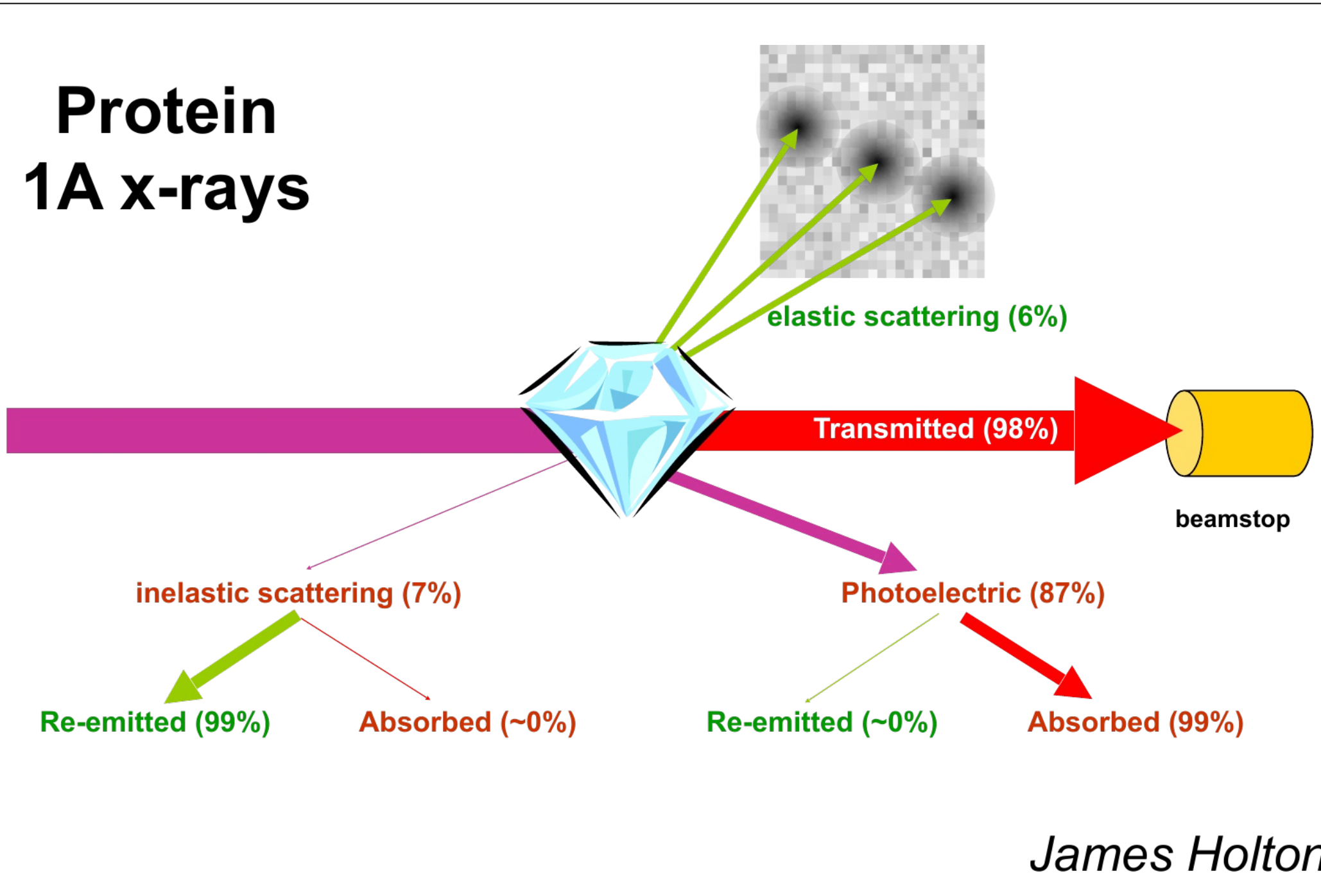
Table 1

Incomplete list of methods for predicting SWAXS curves from structural models: Fitting of hydration layer required ($\delta\rho_{fit}$, including method that ignore the hydration layer), using tabulated reduced form factors (f_{red}), resolution [atomistic or coarse grained (CG)], fluctuations included, free availability [Download (D), web server (W)]. Additional software is listed in Refs. [63,64]

ID	Name/authors	Year	$\delta\rho_{fit}/f_{red}$	Resol.	Fluct.	Avail.	Refs.
Implicit solvent methods							
1	CRYSOL	1995	Yes/yes	atom.	–	D/W	[25]
2	ORNL-SAS	2007	Yes/yes	atom.	–	D	[65]
3	SoftWAXS	2009	Yes/–	atom.	–	D	[66]
4	Fast-SAXS-pro	2009	Yes/yes	CG	Yes	D/W	[30,36]
5	FoXS	2010	Yes/yes	atom.	–	D/W	[67,29]
6	PHAISTOS	2010	Yes/yes	CG	–	D	[68]
7	AquaSAXS/AquaSol	2011	Yes/yes	atom.	–	W	[27]
8	SASbtx/Zernike	2012	Yes/–	atom.	–	W	[69]
9	RISM-SAXS	2014	–/yes	atom.	–	D	[70]
10	BCL::SAXS	2015	Yes/yes	atom.	–	D	[71]
11	Pepsi-SAXS	2017	yes/yes	atom.	–	D	[72•]
Explicit solvent methods							
12	SASSIM/Sassena	2002	–/yes	atom.	Yes	D	[73]
13	MD-SAXS	2009	–/–	atom.	Yes	–	[74,75]
14	AXES	2010	Yes/–	atom.	–	W	[26]
15	HyPred	2011	–/–	atom.	–	W	[76]
16	Park et al.	2009	–/–	atom.	–	–	[77]
17	Köfinger &Hummer	2013	–/–	atom.	Yes	D	[78]
18	WAXSiS	2014	–/–	atom.	Yes	D/W	[38,79]

Table from Jochen S Hub. Curr. Op. in Struct. Bio. 2018, 49:18-26

Where do Photons go?



Practical considerations

Sample requirements for (SAXS) solution scattering

- size: >5kD
- purity: highly monodisperse !
- concentration: 0.25 – 10mg/ml (higher for small proteins and intermediate angle data
- sample volume 15-50 ul ;(so only a fraction of 1mg protein needed for a starting experiment!)
- enough material for at least 3 concentrations
- matching buffer solution is very important (lower salt better)
- most buffer components tolerated (e.g. glycerol (<30%) and salt (<0.5M) are OK)
- S-reducing agent can help protein to stay intact under irradiation

Additional requirements for time-resolved measurements

- lots of sample (at least 10mg, better more)
- sufficiently large change between initial and final state
- pre-characterization of kinetics by other techniques

Courtesy: Thomas Weiss, SLAC

Practical considerations

A good SAXS experiment starts in your home lab

- every protein has its own “personality”
  - the more you know about your protein the better you can select the data acquisition parameters (buffer composition, pH, additives ….)

- Characterize your protein as much as possible with biochemical means
  - check for possible oligomerization with concentration
    - in case of complexes: for dissociation under dilution
  - determine highest concentration the protein is stable (and how long?)
  - simulate shipping conditions (e.g. freezing & thawing) and check sample quality afterwards

- know your numbers
  - sequence and MW
  - extinction coefficient and concentration of your stock solution

Courtesy: Thomas Weiss, SLAC

Practical considerations

Monodispersity

- check your samples:
  - Good solubility (clear solution), no obvious precipitates
  - Single species on native gels
  - SDS-PAGE should show no contamination
  - Single symmetric peak on an SEC column
- Other analytical techniques:
  - Dynamic light scattering (DLS)
  - Analytical ultracentrifugation
  - Mass spectrometry SEC-MALLS

Buffer conditions

- use a low salt concentration if possible
- for proteins PBS buffer is usually a good choice
- consider additives to prevent radiation damage (DDT, TCEP, Glycerol …)
- bring plenty of matched buffer

Courtesy: Thomas Weiss, SLAC

Practical considerations

Before coming to SSRL

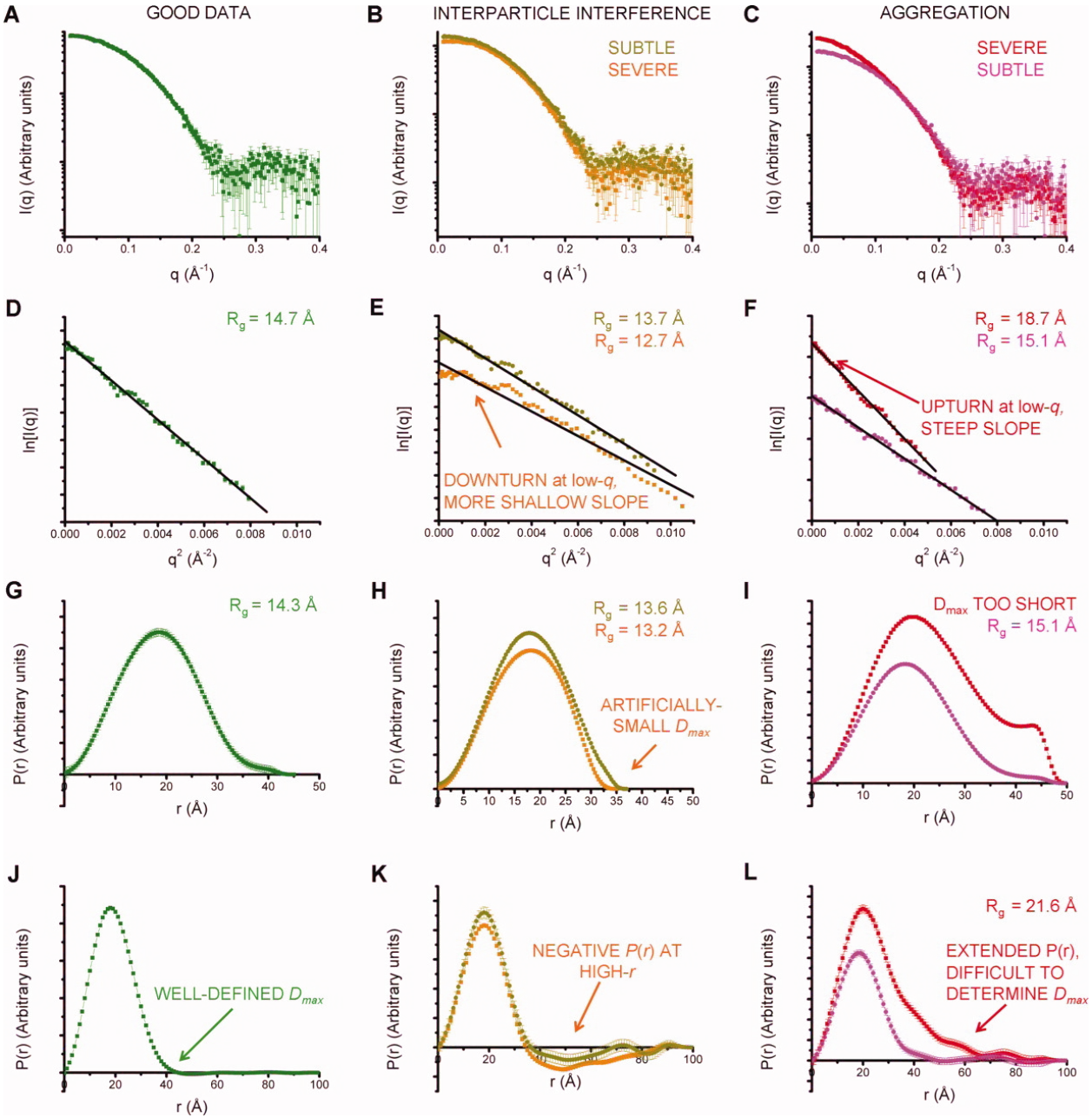
- provide accurate information in the beamtime request form
- ask beamline staff if you are unsure or have questions
- contact your beamline staff before experiment just in case something changed

At the beamline

- understand how the data collection works and how to load your samples
- take plenty of buffer images
- take advantage of the online data reduction: **monitor what's happening!**
- consider sample recovery for post exposure analysis
- bring additional radical scavengers in case of unexpected radiation damage

Courtesy: Thomas Weiss, SLAC

Practical considerations – Data Quality



Jacques D.A. and Trewhella. Prot. Sci. 2010 19(4):642-657

Immediate data quality checks

- aggregation:
  - upturn at low q
  - residuals in guinier plot will show upward curvature
- interparticle repulsion:
  - downturn at low q
  - residuals in guinier plot will show downward curvature
  - will increase with concentration

Checks with the p(r) function

- determine Dmax
  - no “nose-diving” !
  - no excessive oscillation around 0
  - rule of thumb: Dmax ≈ 3* Rg
  - Switch off P(dmax)=0 and use large Dmax to estimate
- determine Rg
  - should compare well with Rg from Guinier

Courtesy: Thomas Weiss, SLAC

Practical considerations – Aggregated data

What if your Sample is Aggregated?

- centrifuge your sample (ideally keep it cold)
- dilute and centrifuge
- filter
- add more DTT if radiation damage is the problem
- run sample through SEC column if time permits
- change buffer condition (if you have enough material)

Courtesy: Thomas Weiss, SLAC

Practical considerations – Finding a SAXS beamline

lightsources.org

Light sources of the world

There are more than 50 light sources in the world (operational, or under construction). This page lists all the members of the lightsources.org collaboration.

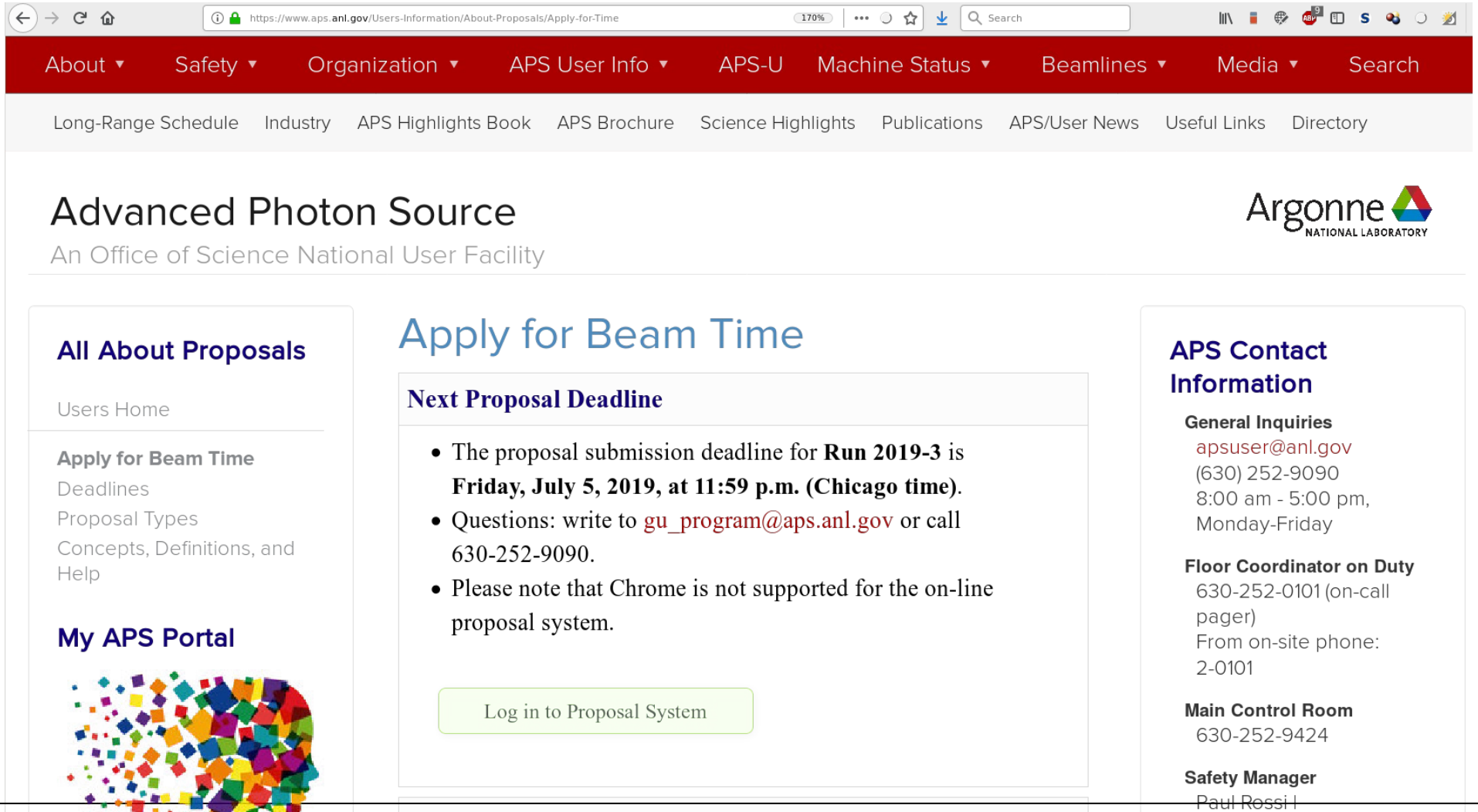
Orange pins on the map represent members of the lightsources.org collaboration.

Practical considerations – Finding a SAXS beamline

Practical considerations – Finding a SAXS beamline

Steps are generally the same:

- Find a beamline
- Talk with a beamline scientist
- Register and submit a proposal
- If you are going to do the experiment yourself
  - Safety training etc.



Practical Considerations

If you are awarded time

- Bring a **TEAM!**
- Bring extra samples (ask colleagues).
- Expect to work every hour of your allocation!
  - e.g. if you have 2 days beamtime scheduled, expect to have someone working at the beamline 48 hours

Argonne NATIONAL LABORATORY

APS at Argonne National Laboratory

diamond

Practical considerations

SANS beamlines

NCNR, NIST, Maryland
HFIR, ORNL, Tennessee
ISIS, RAL, UK
ILL, Grenoble, France
ANSTO, Sydney, Australia
ESS, Lund, Sweden (2025)
others...

NIST

Search NIST | NIST MENU

NIST CENTER FOR NEUTRON RESEARCH

Logon to your NCNR-IMS account

Obtaining Beam Time

Arrange a visit to NCNR +

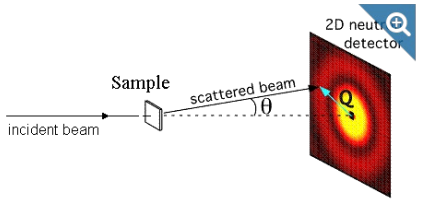
Planning Your Experiment +

Live Data

About NCNR +

Neutron Instruments −

Small Angle Neutron Scattering (SANS)



Small-Angle Neutron Scattering (SANS) probes material structure on the nanometer (10^{-9} m) to micrometer (10^{-6} m) scale. Structures on this length scale are critical to the performance of advanced engineering materials.

SANS Instruments

- CHRNS VSANS
- CHRNS 30m SANS
- CHRNS USANS
- nSoft 10m SANS
- NG7 30m SANS

OAK RIDGE National Laboratory

About Us · User Facilities · Science & Discovery · News · Events · ORNL Careers

Our People

Search

Science and Discovery

REQUEST BEAM TIME

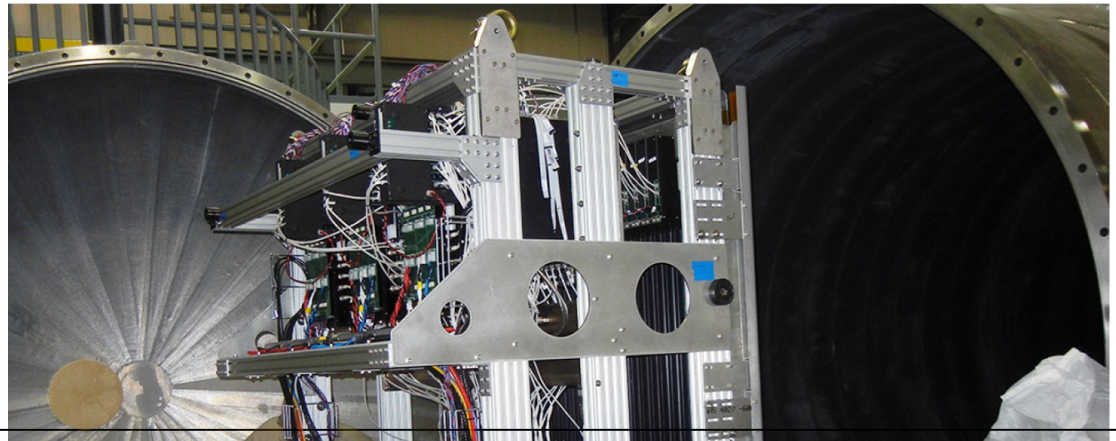
Home | About | Future | Science | For Users | For Industry | Publications | Instruments

News | Events

Home » Biological Small-Angle Neutron Scattering Instrument

Biological Small-Angle Neutron Scattering Instrument BIO-SANS | CG-3 | HFIR

Overview | Team | User Guidance | Live Instrument Data | Publications | Gallery | Spec Sheet | Contact Us



SAXS vs SANS

	SAXS	SANS
Features	msec resolution for time-resolved measurements	D labeling and H/D contrast variation
	Superior q-resolution	Magnetic scattering
	Anomalous scattering (ASAXS)	Conducive to extreme environments
	Small sample size	Nondestructive
Complications	Radiation damage to some samples	Incoherent scattering
	Parasitic scattering	H/D isotope effects
	Fluorescence	
	Beam stability	

Charles Glinka, NIST

ASAXS

Anomalous SAXS:
Allows limited contrast variation when the adsorption edge of one of the constituent elements is at an accessible energy range.

Theory pioneered by Heinrich B. Stuhrmann:

Q. Rev. Biophys. 14, 433 (1981)
Adv. Polym. Sci. 67, 123 (1985)

Heinrich B Stuhrmann
Guinier prize 2006

Stuhrmann analyzed metal containing proteins such as hemoglobin, ferritin, and the anomalous effect on the radius of gyration of DNA near the absorption edge of counterions.

In the case of the large subunit of ribosome (1500 kD), measurements near phosphorous K-edge allowed separation of all three partial intensities. *Stuhrmann. J. Appl. Cryst. 2007. 40:s23*

M. Sztucki et al., 2012 Eur. Phys J. Special Topics. 208:319-331

ASAXS

$$f(\lambda) = f_0 + f'(\lambda) + \mathbf{i} f''(\lambda)$$

$$|f| = [(f_0 + f')^2 + f''^2]^{\frac{1}{2}}$$

KK:
Re/Im of Fourier
1-1 Even odd

Stuhrmann 1981: *f''* via absorption vs wavelength for bound iron.
f' via *f''* using the Kramers-Kroning relation…
Tabulated values are available for most elements.
Corrected I(q) curves were produced, compared.
Multipole expansion for scattering density
→ distance distributions for iron were estimated.

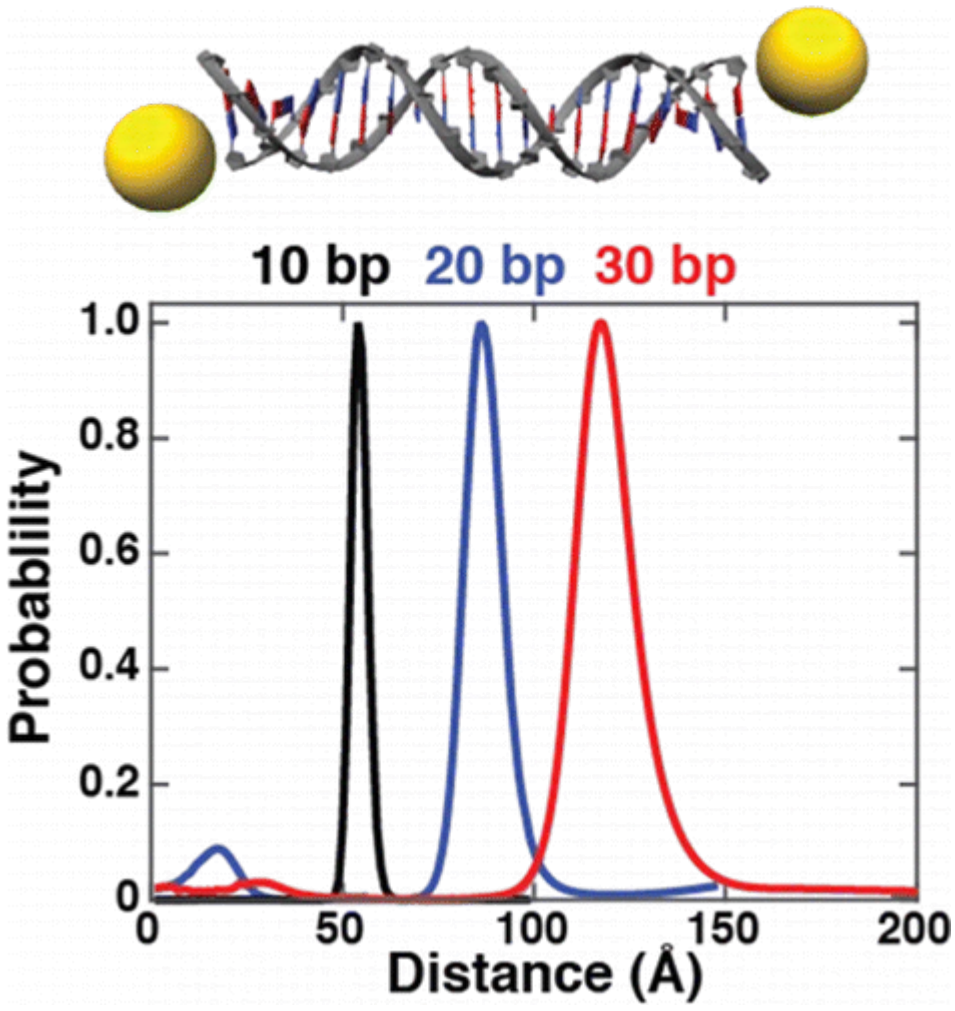
Generalized in *V. J. Pinfield and D.J. Scott. PLoS ONE. 2014 9(4): e95664*

Table 2. Distances between label atoms or nanocrystals.

Molecule	Actual distance between labels/Å	Calculated distance between labels/Å
10 bp DNA, atom labels	37.3	----
10 bp DNA, nanocrystal	50.5	51
20 bp DNA, nanocrystal	60.7	61
50 bp DNA, nanocrystal	142.0	143
100 bp DNA, nanocrystal	269.6	270
200 bp DNA, nanocrystal	672.0	673

The distance between the label atoms or nanocrystals, as defined in the coordinate files, and determined by the anomalous SAXS simulation.
doi:10.1371/journal.pone.0095664.t002

ASAXS

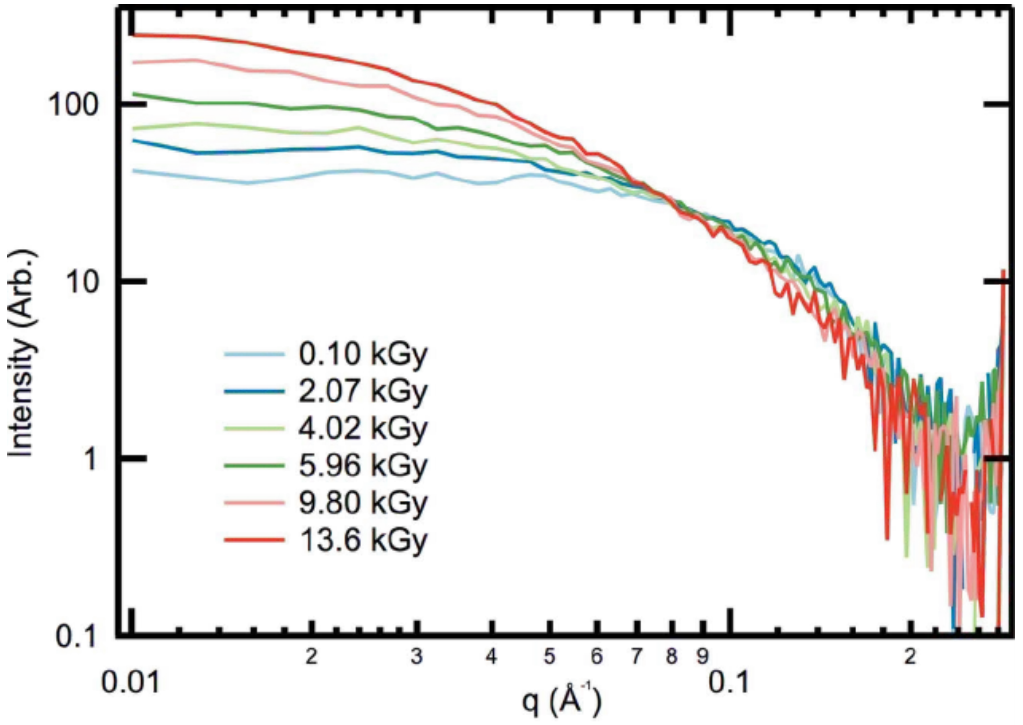


APS
ID-12

Thomas Zettl et al. Nano Letters 2016 16 (9), 5353-5357

SAXS – Radiation damage

X-ray-induced radiation damage can cause macro- molecule aggregation, fragmentation, conformation changes and unfolding, all of which can be detected by SAXS. Radiation damage is therefore a major obstacle for SAXS, and descriptions of dedicated biological SAXS beamlines acknowledge the need to check for and avoid radiation damage.



Radiation damage in most contexts is a function of Dose (Gy = J kg^{-1}).

$$\text{Dose} = \frac{ftAE_\gamma}{\rho l}$$

f – flux density
t – exposure time
A – fraction of incident energy absorbed
E_γ – energy of photon
ρ – sample density
l – path length

Jesse Hopkins and Robert Thorne, J. Appl. Cryst. (2016) 49:880-890

SAXS – Radiation damage

Minimize:

- Reduce exposure time

$$\text{Dose} = \frac{ftAE_{\gamma}}{\rho l}$$

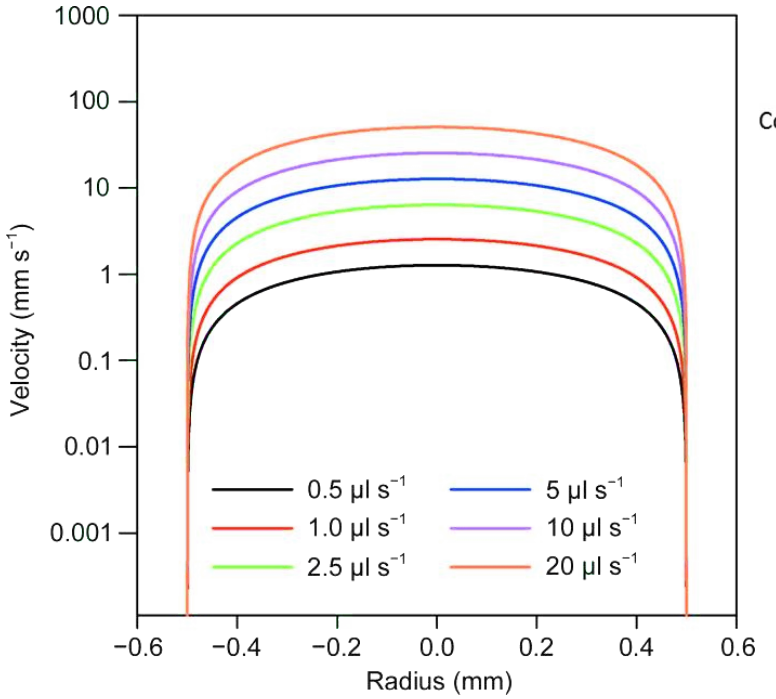
- Decrease volume irradiated
  - Oscillating or continuous flow
  - Defocusing the beam

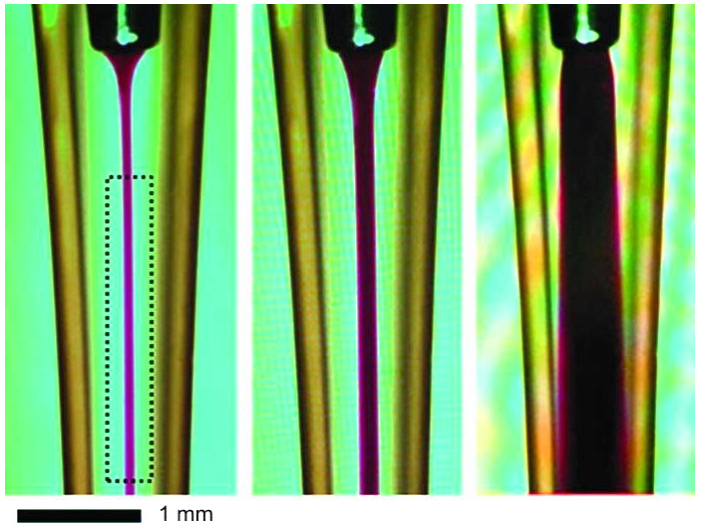
- Buffer additives to competitively bind with free radicals or by inhibit aggregation
  - Glycerol

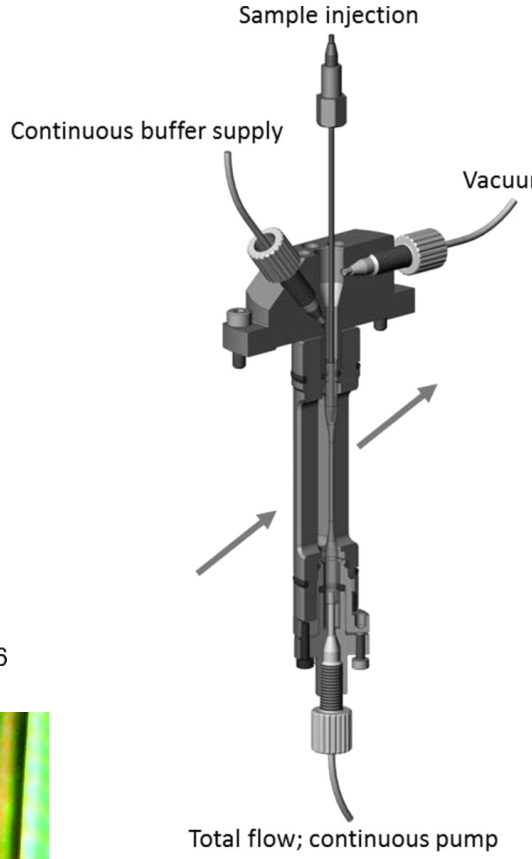
- Cryo-SAXS

Jesse Hopkins and Robert Thorne, J. Appl. Cryst. (2016) 49:880-890

SAXS - Coflow

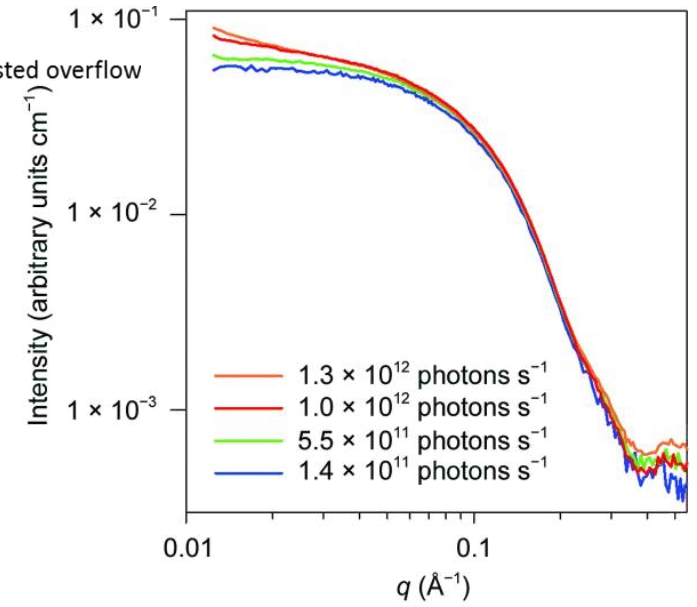


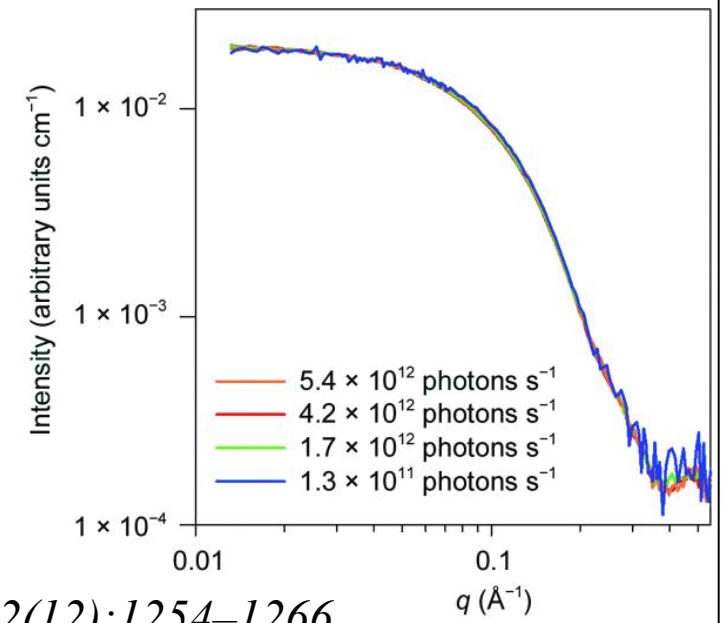




ANSTO / SAXS/WAXS
APS / BioCAT
SSRL / BL4.2
SOLEIL / SWING
MAX IV / CoSAXS

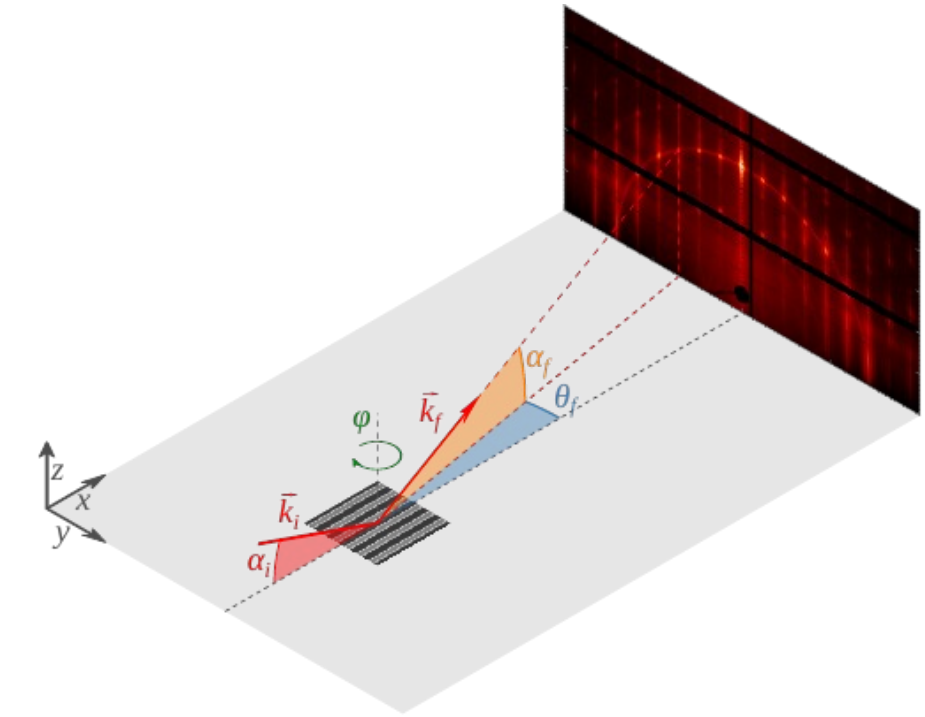
RNase A



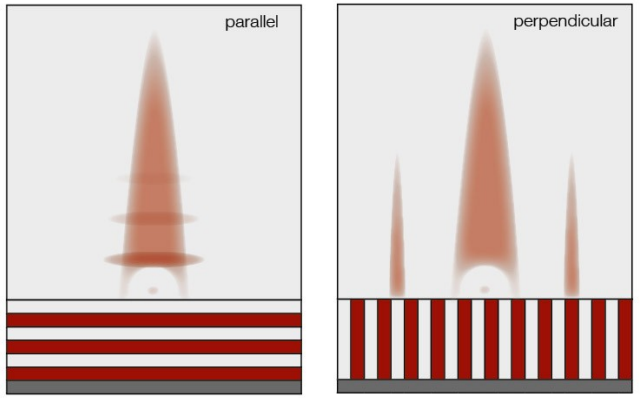


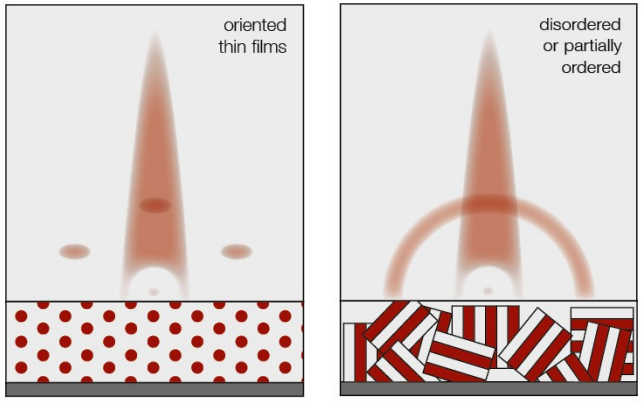
Nigel Kirby et al., Acta Cryst. D Struct. Biol. 2016 72(12):1254–1266.

Surface studies - GISAS

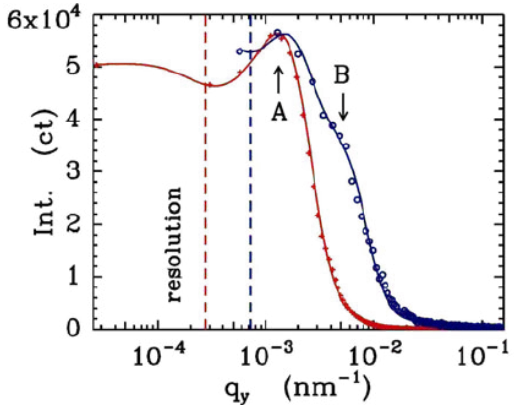


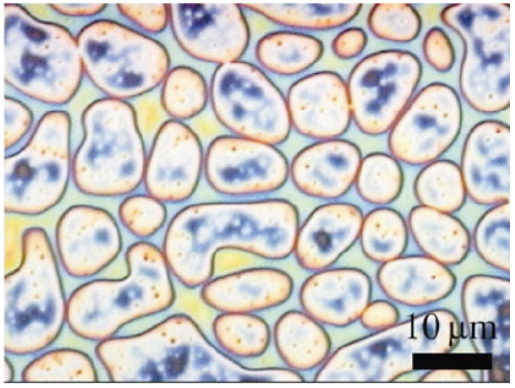
Müller-Buschbaum P. (2009) A Basic Introduction to Grazing Incidence Small-Angle X-Ray Scattering. Lecture Notes in Physics, vol 776, Springer





https://wiki.anton-paar.com/en/grazing-incidence-small-angle-x-ray-scattering-gisaxs



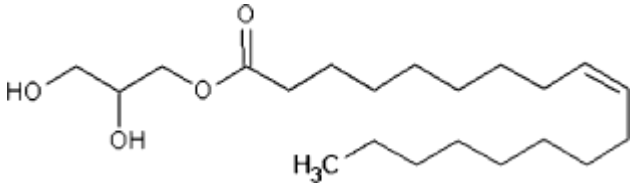


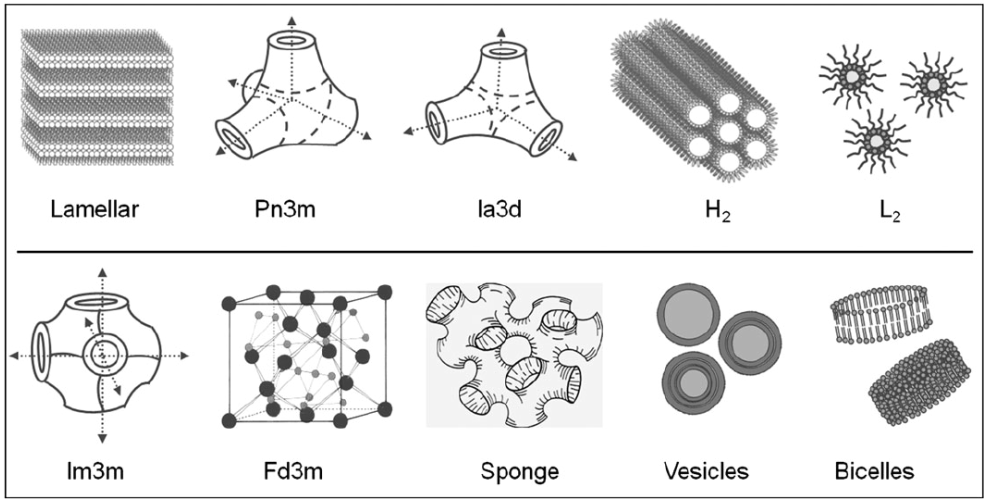
Blend films of PS and PnBA
"A" most prominent in-plane length

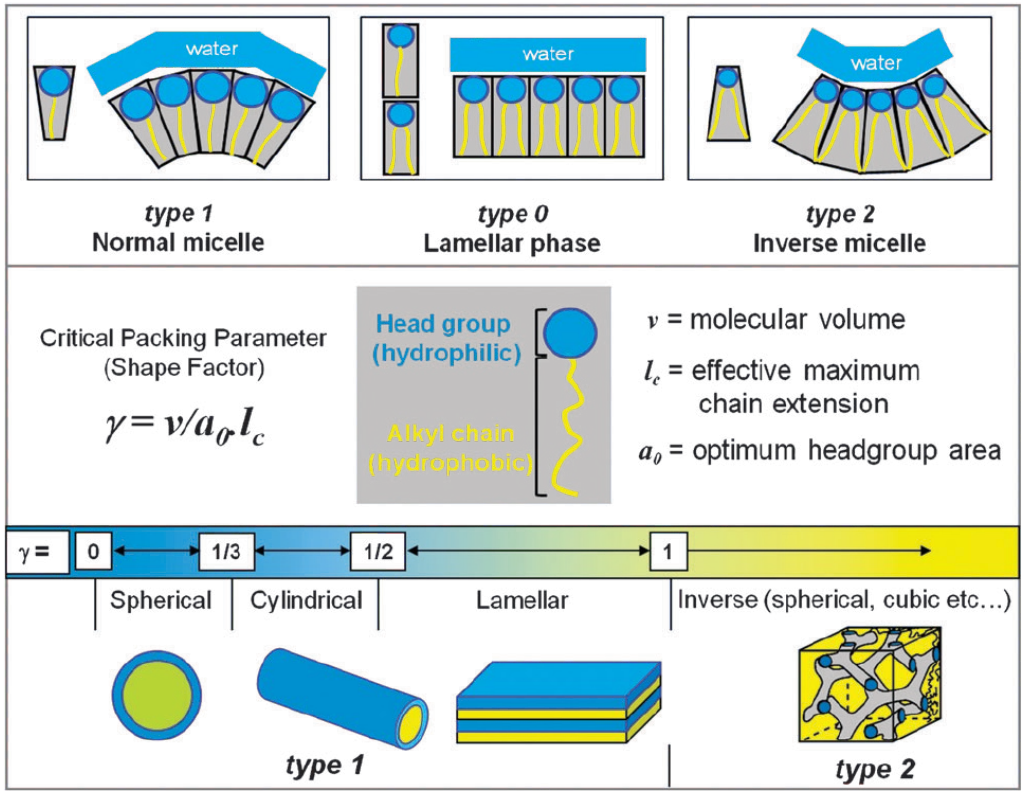
Müller-Buschbaum P., Prog. in Colloid & Polymer Sci. 2006 doi:10.1007/2882_031

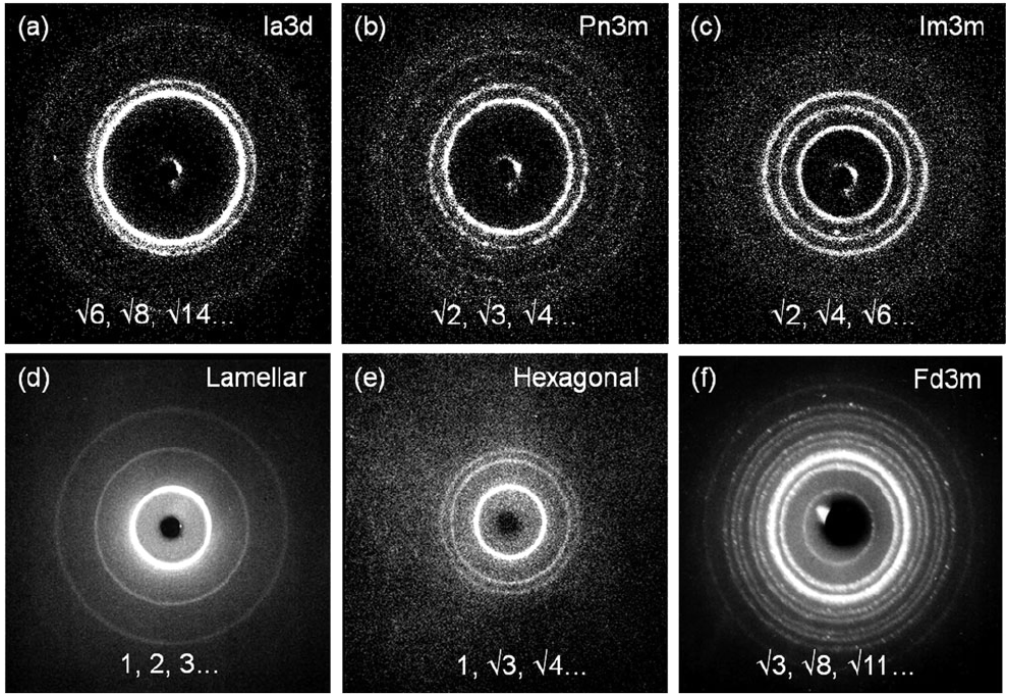
Liquid crystal SAXS

Liquid crystal structures
e.g. Monoolein thermotropic and lyotropic









C. V. Kulkarni et al. Phys. Chem. Chem. Phys. 2011, 13:3004-3021

Monodispersity revisited

Svergun – 2003

Shannon channels = $D_{max} \cdot q\text{-}range / \pi$

"the number of [obtainable parameters] typically does not exceed **10–15**"

Hub – 2018

"… generally accepted that experimental SWAXS curves do not contain more than 10–30 independent data points."

Monodispersity → maximize information content / species

Even if you purify immediately before SAXS measurements and inject each fraction or a pool of fractions, you still have a chance that the sample will either aggregate or degrade during operations

SEC-SAXS

- High pressure liquid chromatography or FPLC (Fast protein liquid chromatography) on line with the SAXS cell
- Individual peaks are more likely to be monodisperse
- First use paper, available to users who could self-manage FPLC
  - *Mathew, E., Mirza, A., & Menhart, N. (2004).* Liquid-chromatography-coupled SAXS for accurate sizing of aggregating proteins. *J. Synchrotron Rad. 11, 314-318.*
- First setup with user HPLC support
  - *David, G. & Pérez, J. (2009).* Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J. Appl. Cryst. 42, 892-900*
- Implementations (not guaranteed exhaustive)
  - ID-18 BioSAXS/APS
  - BL4.2/SSRL
  - CHESS/MacCHESS
  - SWING/SOLEIL
  - BM-29/ESRF
  - I22/Diamond
  - P12/Petra
  - SR13 ID01/Australian Synchrotron

SEC-SAXS

- **Inline HPLC/MALS system**

SWING/SOLEIL

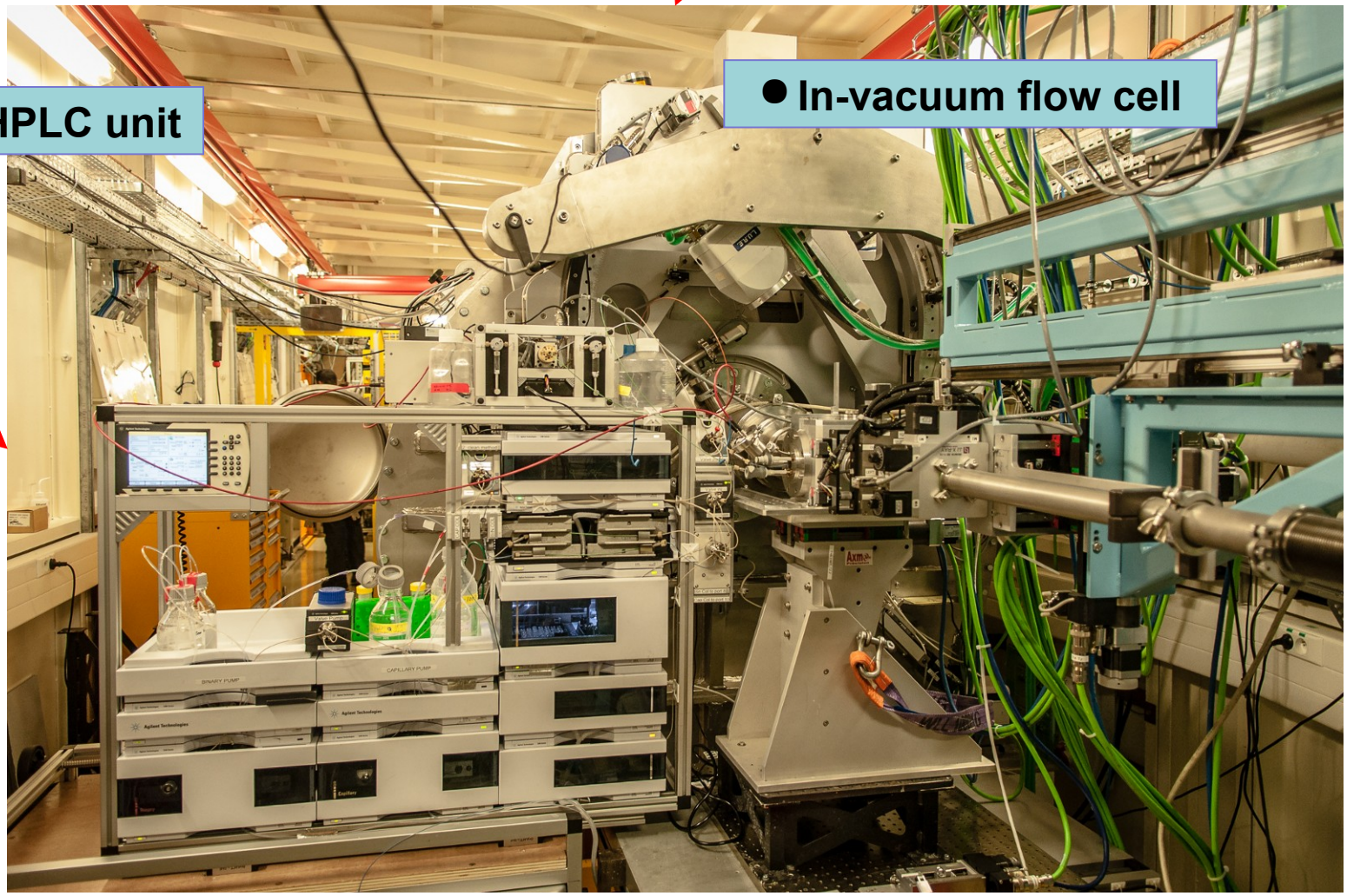
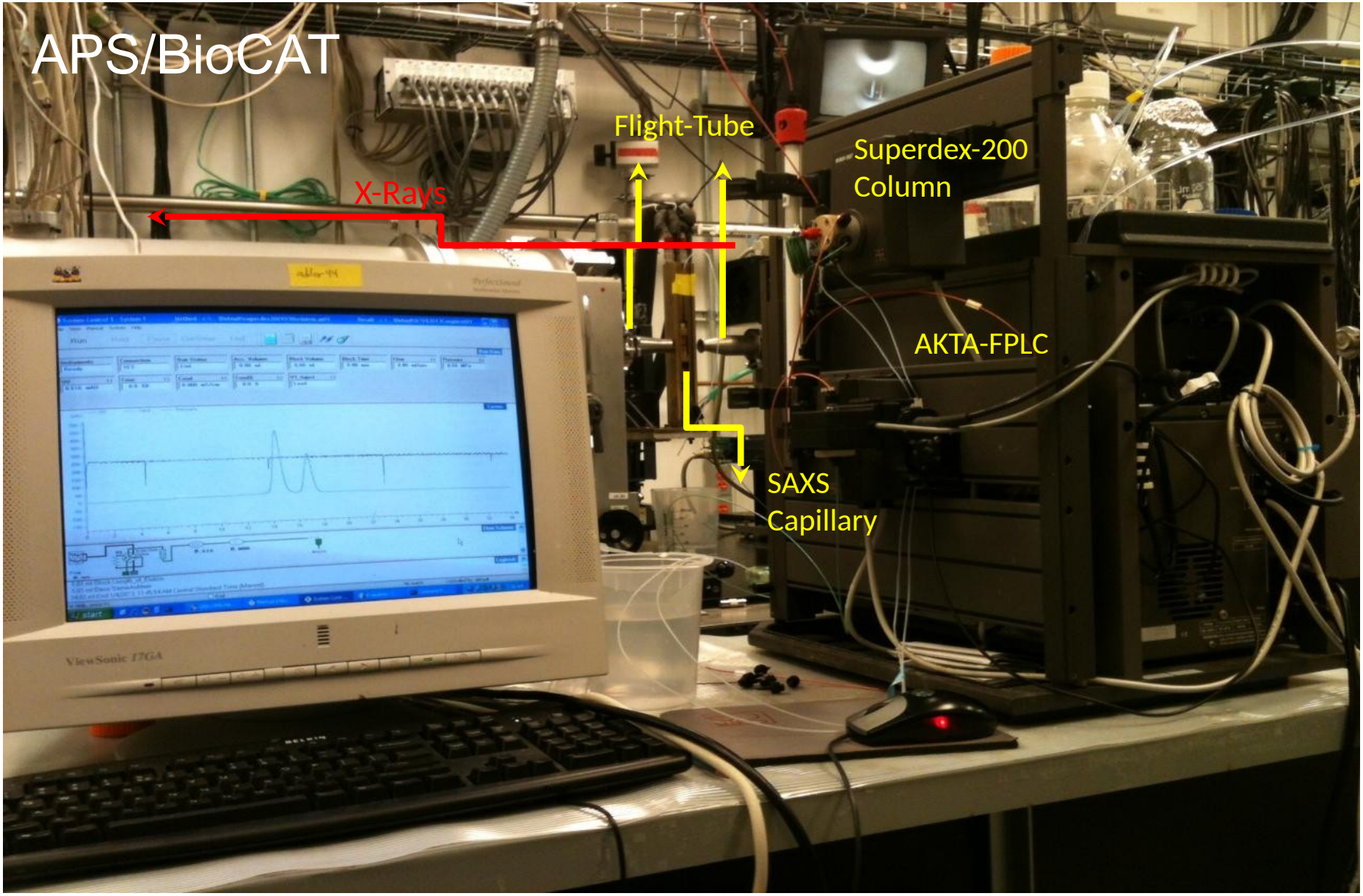


Photo credit: Javier Perez

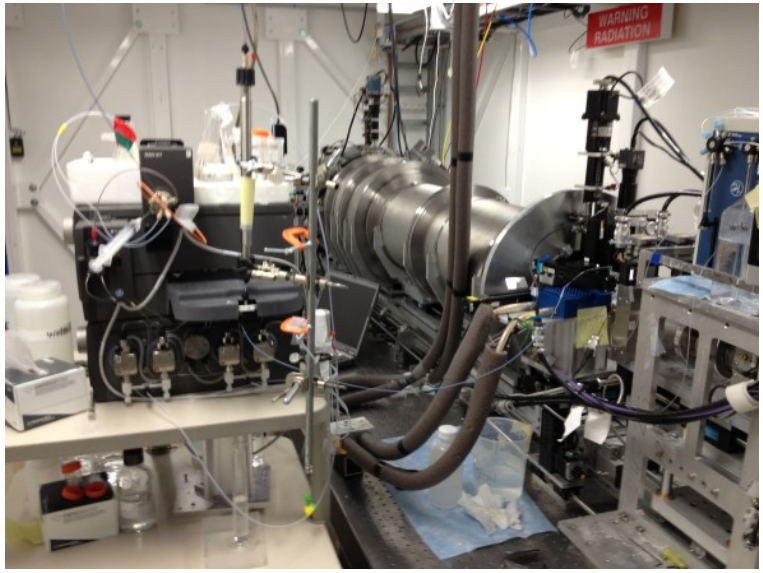
SEC-SAXS



Slide Credit: Srinivas Chakravarthy

SEC-SAXS

SSRL/BL4.2



- The online FPLC-SAXS system at the BioSAXS beamline BL4-2
- consists of an Akta Ettan with low volume (2.5ml) SEC columns:
  - Superdex 200
  - Superose 6
  - Or bring your own

- The system uses the same flow path as the
- regular “autosampler” setup at the beamline:
  - rapid switch-in of the FPLC system during normal data collection
  - FPLC-SAXS and “autosampler” results can be compared quickly

- sample requirement:
  - typically 50ul of 5mg/ml sample
  - each run requires 3 ml of buffer and takes roughly an hour

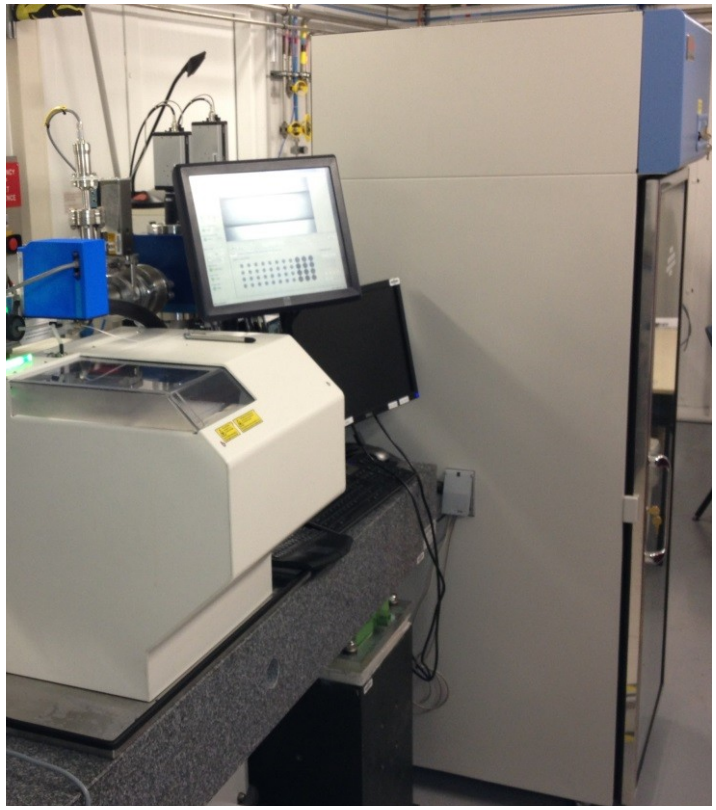
- Automated data analysis scripts allow easy tracking of experimental results during experiment
- More information on our website:
  - **http://www-ssrl.slac.stanford.edu/~saxs/**

Slide Credit: Thomas Weiss

SSRL

SEC-SAXS

BM29/ESRF



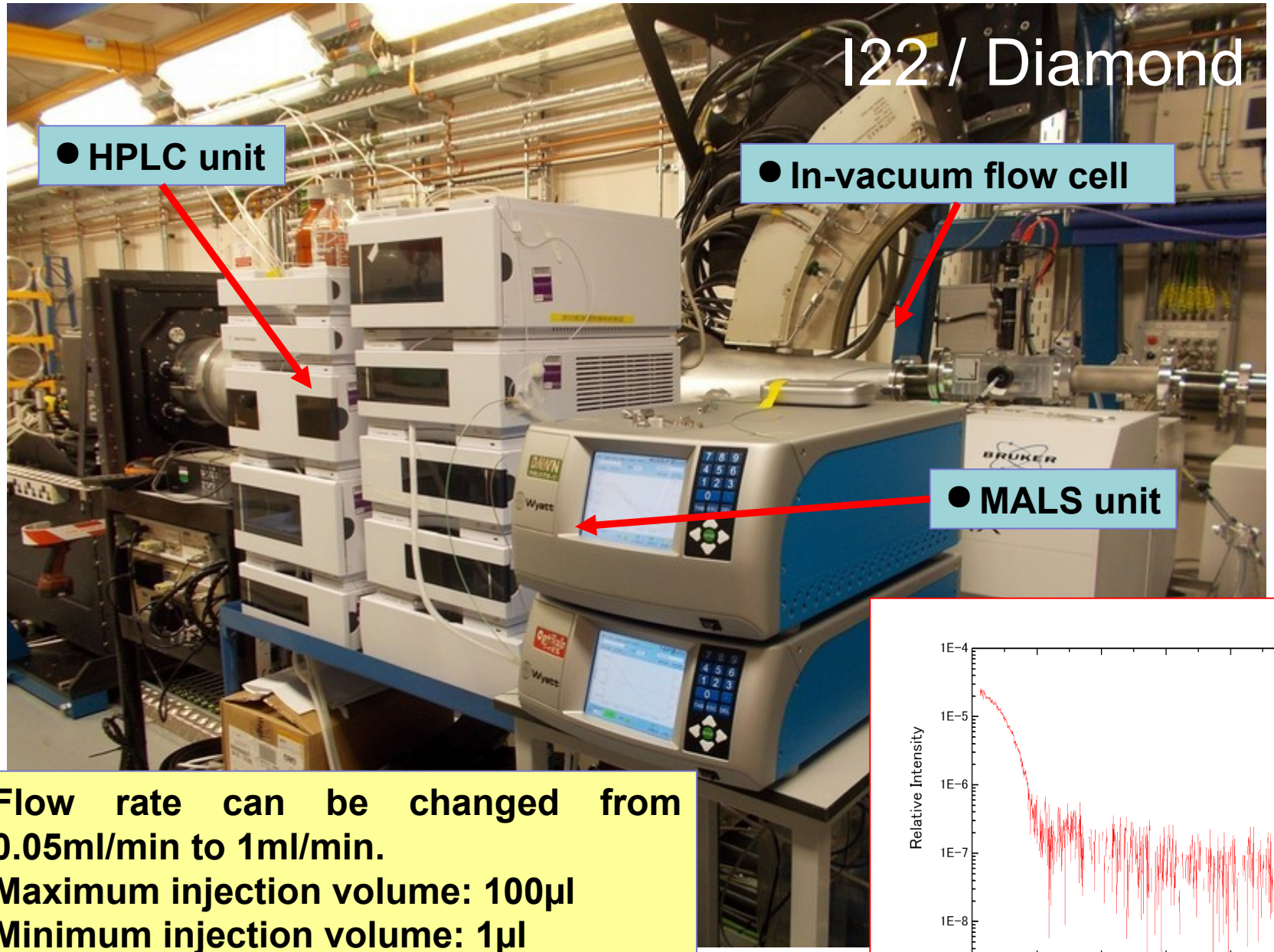
- Automated switching between SEC and SC for efficient use
-

- Integrated Sample changer and Online-SEC
  - SEC units housed in temperature controlled cabinet (4 -25 °C)

Slide Credit: Adam Round

ESRF EMBL

SEC-SAXS



- Flow rate can be changed from 0.05ml/min to 1ml/min.
- Maximum injection volume: 100µl
- Minimum injection volume: 1µl

Slide Credit: Katsuaki Inoue

SEC-SAXS

- Separate immediately before measuring
- Individual peaks are more likely to be monodisperse
- Now available as primary method of analysis at multiple beamlines

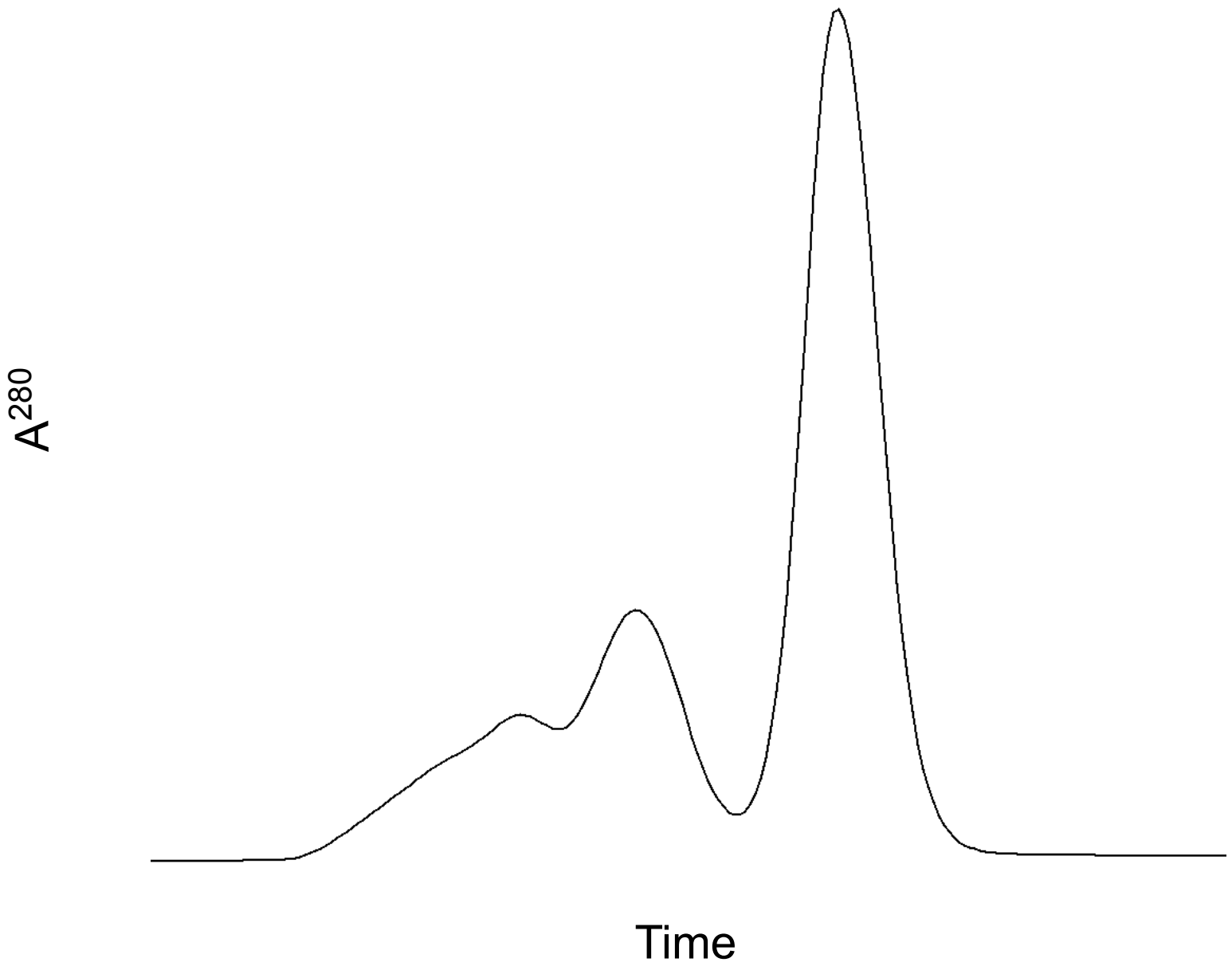
- Conclusion of ACA 2014 session 4.2.4 [ACA Reflexions Fall 2014]:

"The consensus that emerged was that SEC-SAXS may become the standard data collection strategy for biological samples, as a large number of samples that were heretofore believed to be monodisperse have been shown to be polydisperse when analyzed with online SEC-SAXS setups."

ACA Structure Matters — Albuquerque ACA Meeting — Fall 2014

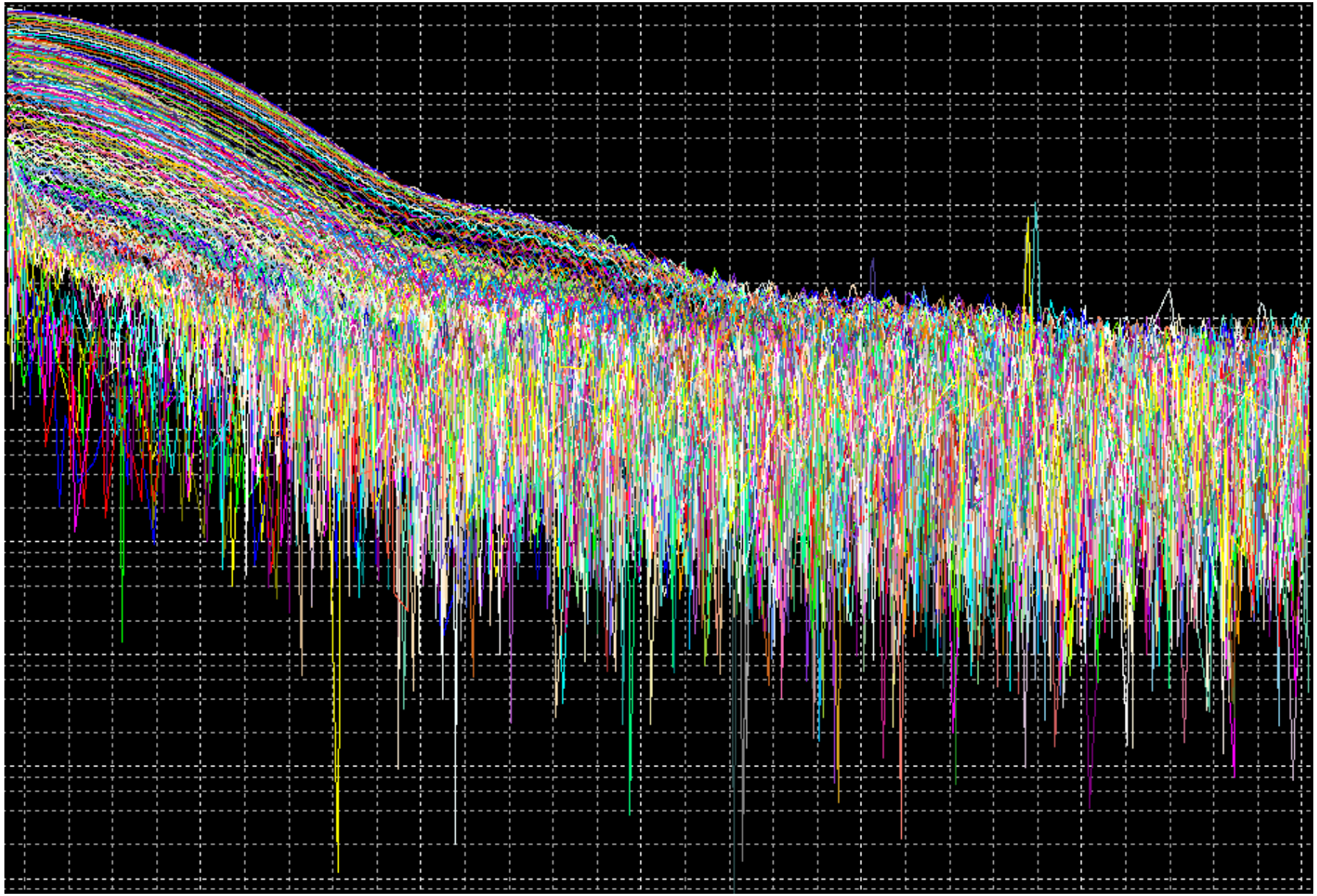
Nicolas Wolff, Sophie Zinn-Justin, Nigel Kirby, Alvin Acerbo, Srinivas Chakravarthy, Javier Pérez, Alexey Kikhney, Emre Brookes, Adam Round, David Lambright.

UV Trace Aldolase



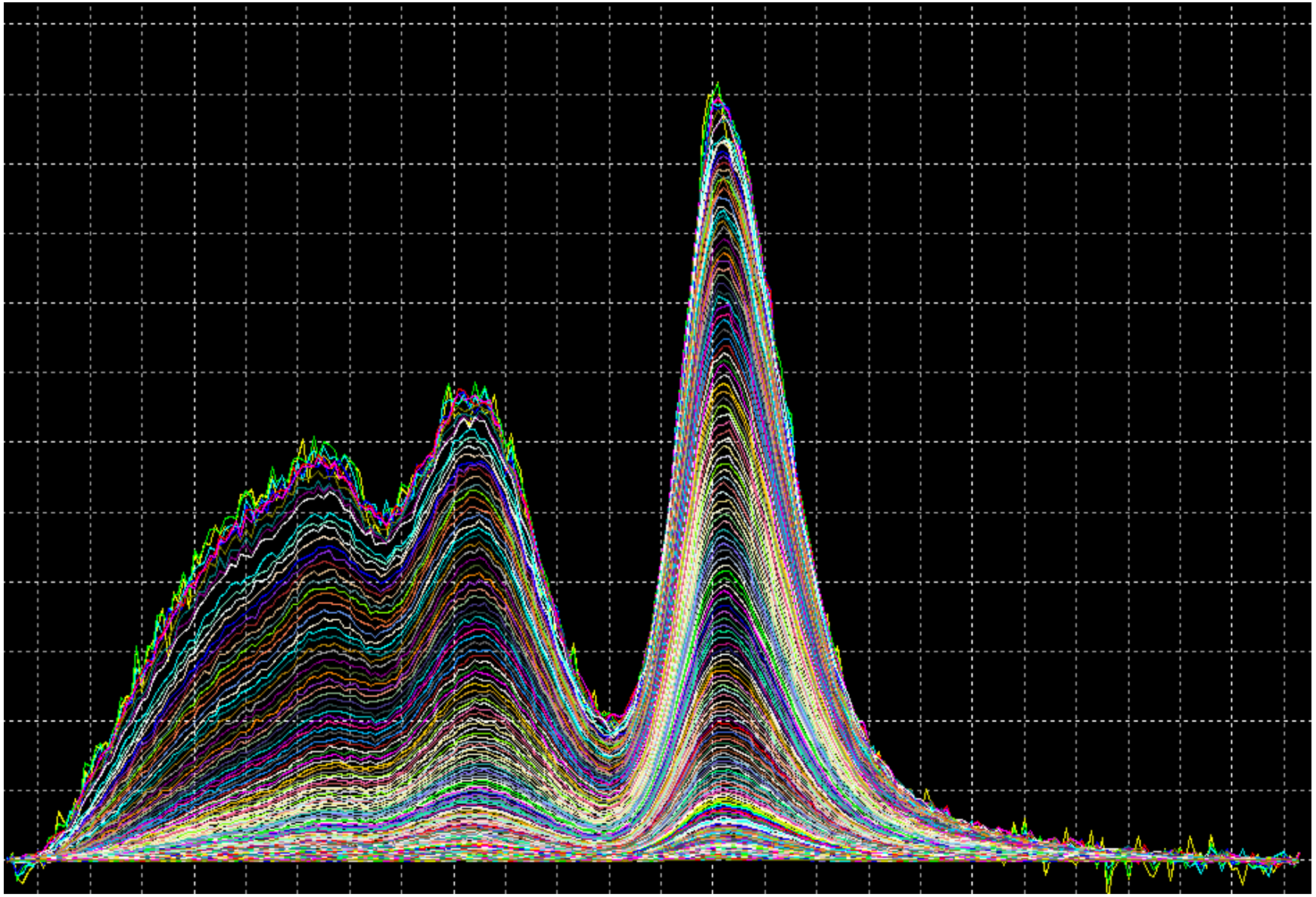
Brookes et al, J. Appl. Cryst. 46 (2013) 1823-33

SEC-SAXS I(q) profiles Aldolase



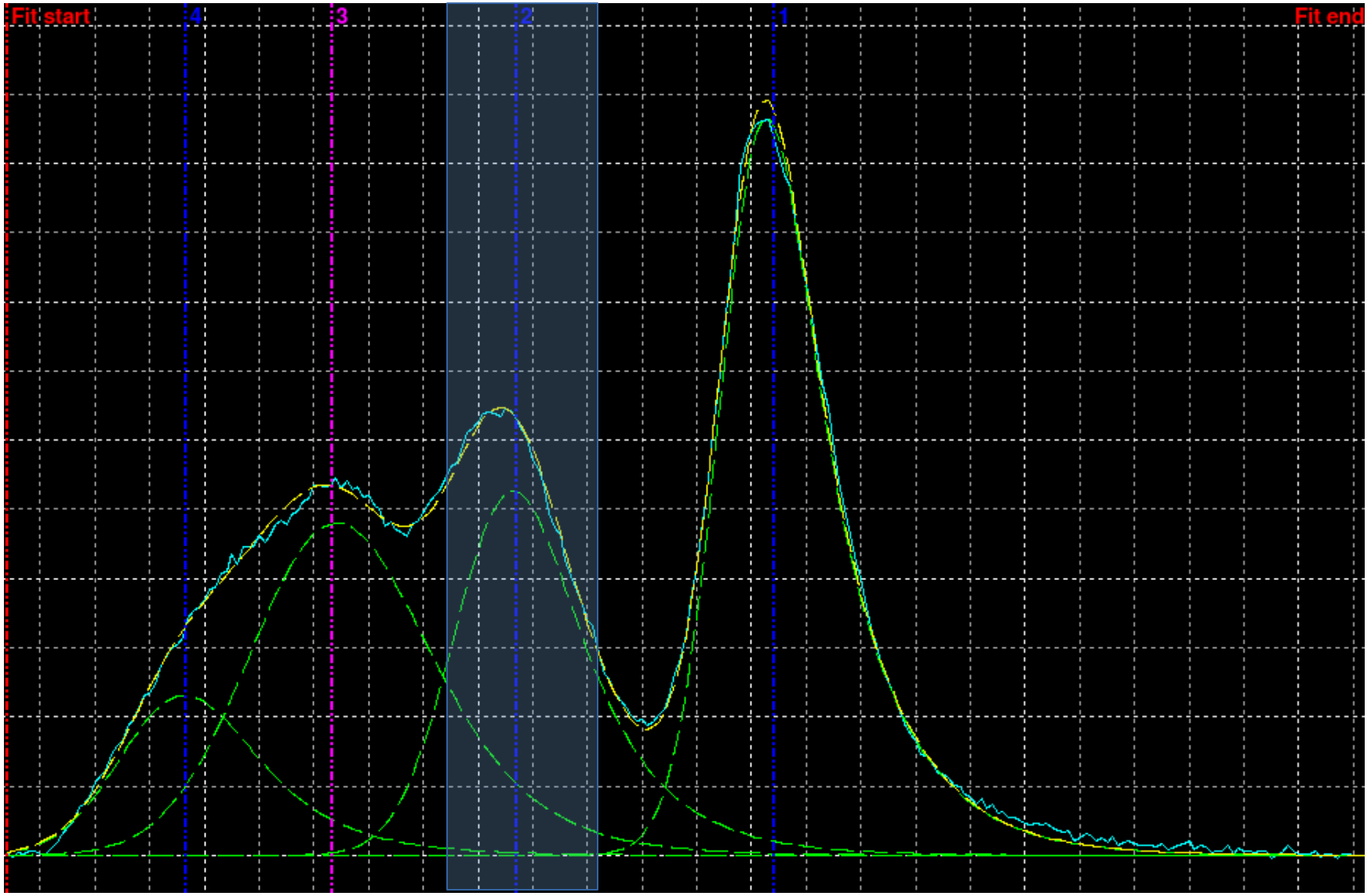
Brookes et al, J. Appl. Cryst. 49 (2016) 1827-41

SEC-SAXS I(t) profiles Aldolase



Brookes et al, J. Appl. Cryst. 49 (2016) 1827-41

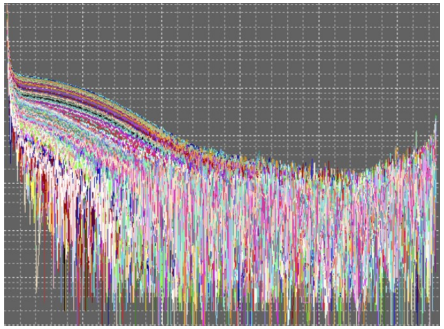
Aldolase Gaussian fit of one I(t)



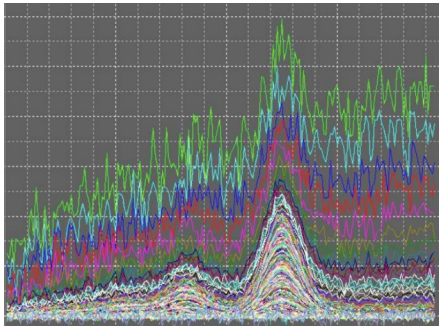
Brookes et al, J. Appl. Cryst. 49 (2016) 1827-41

SEC-SAXS deconvolution of peaks

Collect SEC-SAXS data

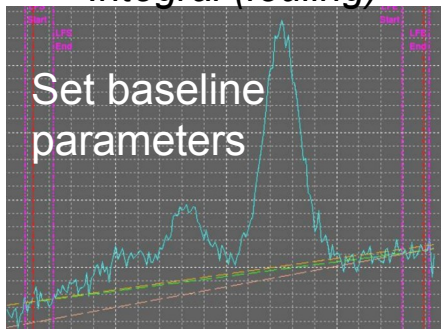


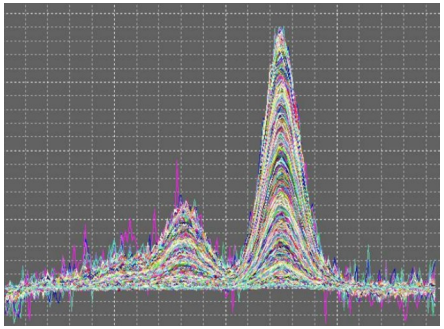
Make I(t)



Select typical curve

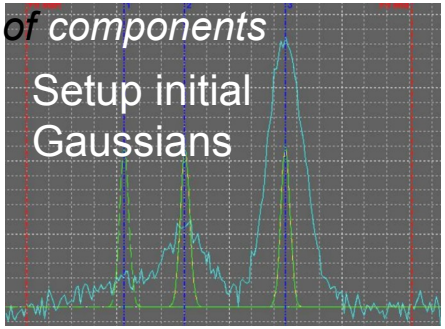
Linear (drift)
Integral (fouling)





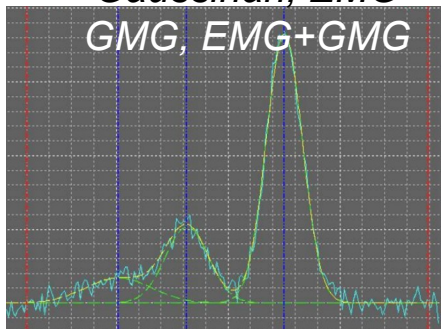
optional SVD to inform number of components

Select typical curve

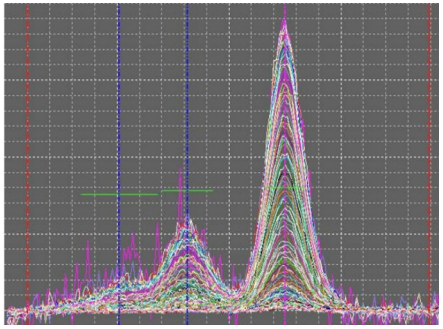


Gauss Fit

Gaussinan, EMG GMG, EMG+GMG

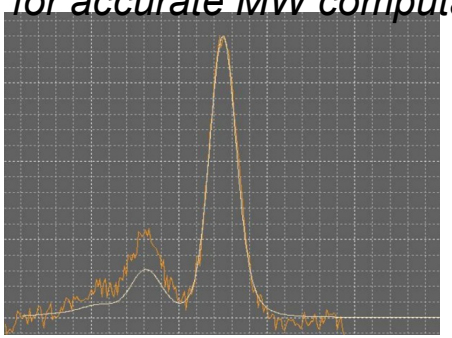


Global Gauss Fit



Opt. apply to conc. curve

UV or RI, simultaneously fit for accurate MW computations

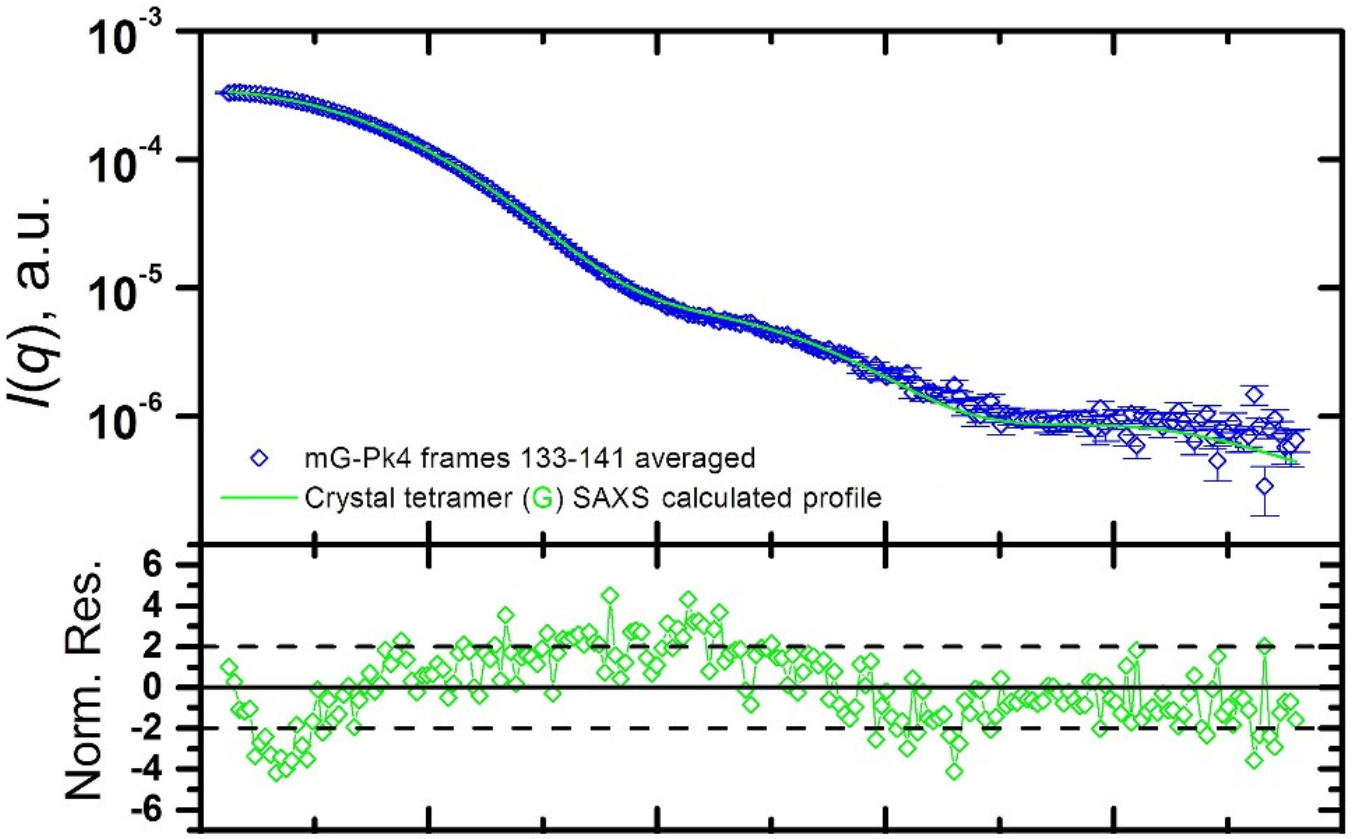


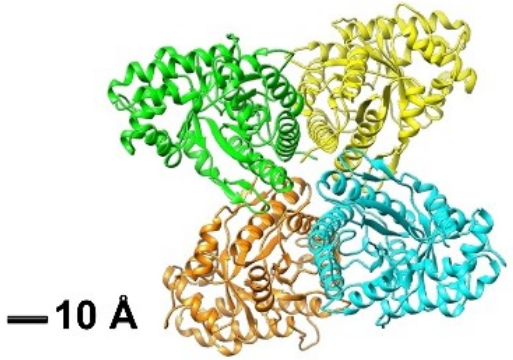
Make I(q)

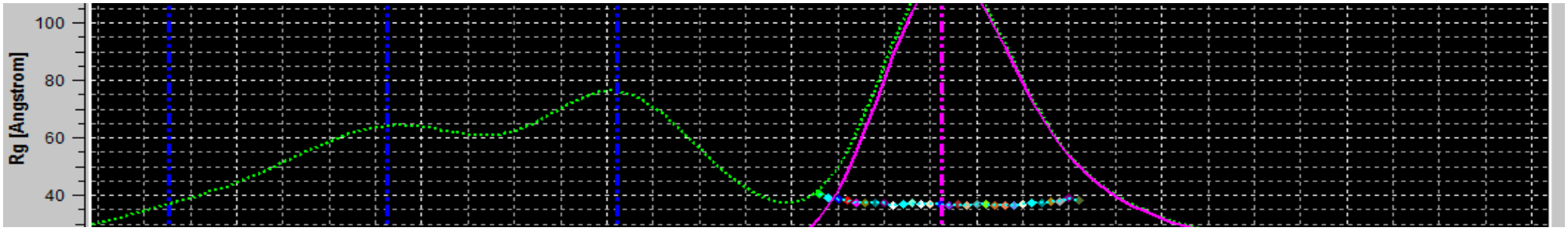
Set of I(q) curves for each Gaussian peak

Brookes et al, J. Appl. Cryst. 46 (2013) 1823-33
Brookes et al, J. Appl. Cryst. 49 (2016) 1827-41

Deconvolution of Aldolase & Model

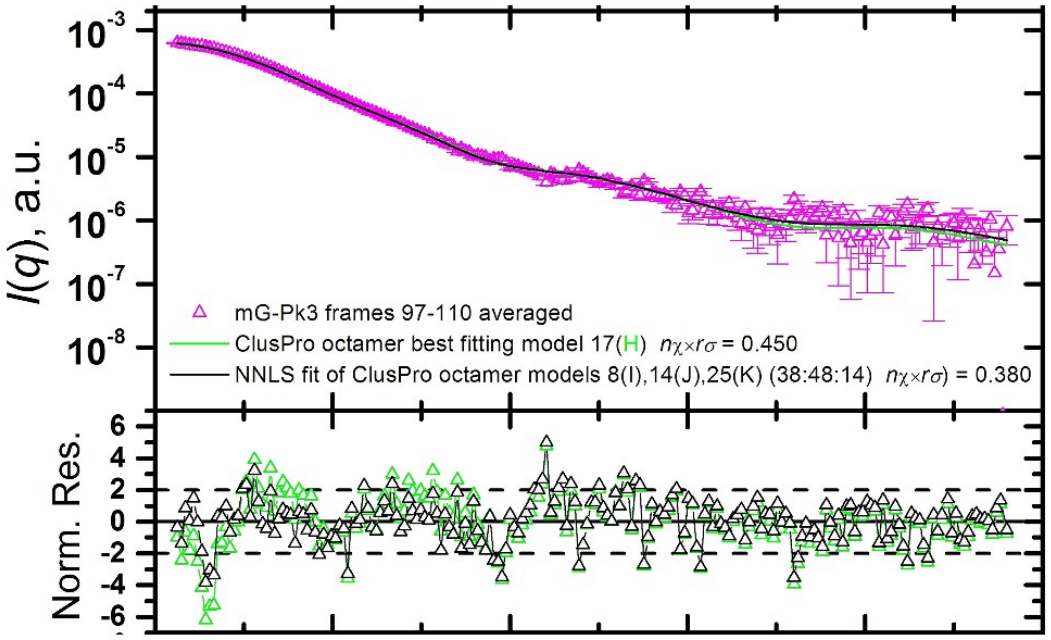


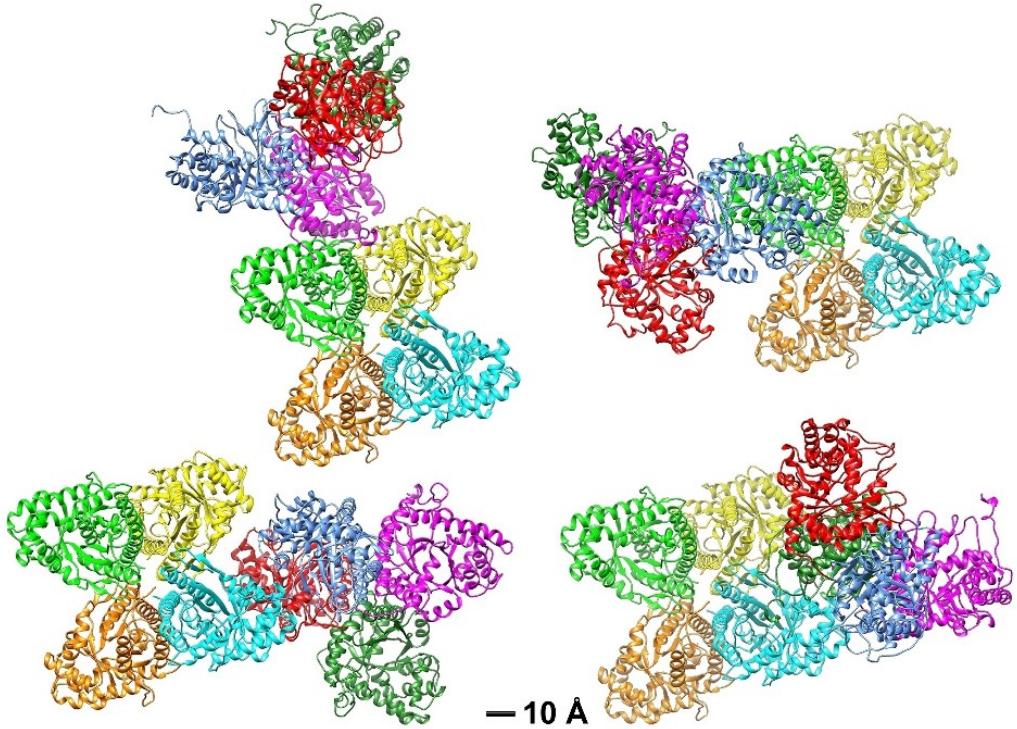


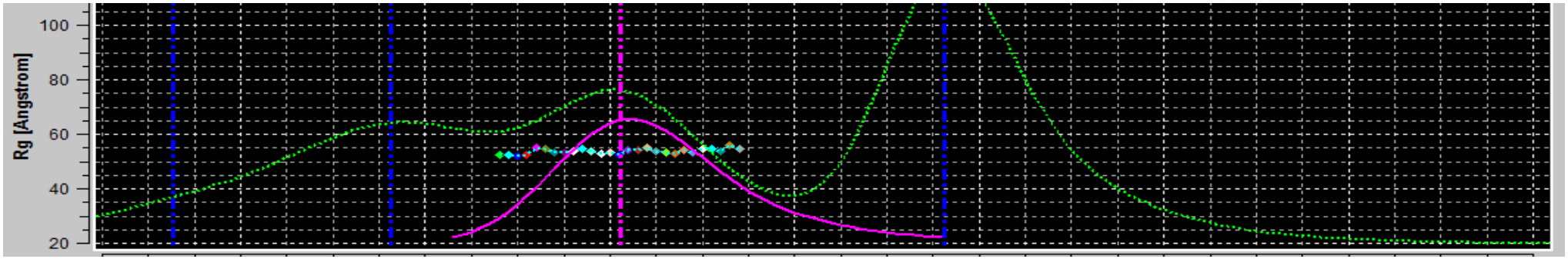


Brookes et al, J. Appl. Cryst. 49 (2016) 1827-41

Deconvolution of Aldolase & Model

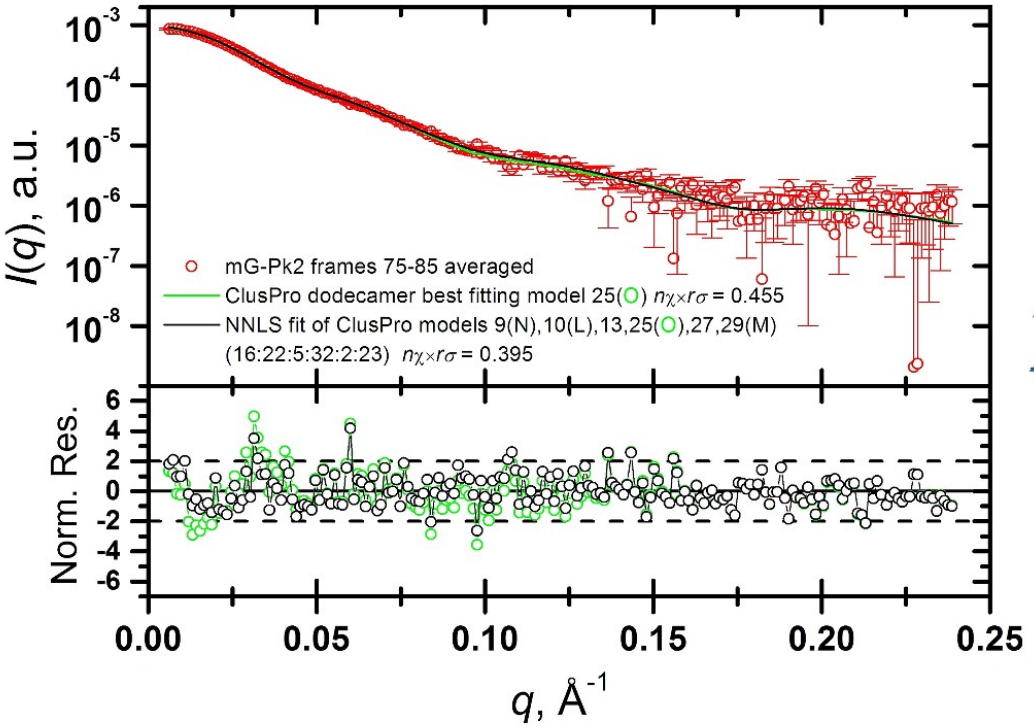


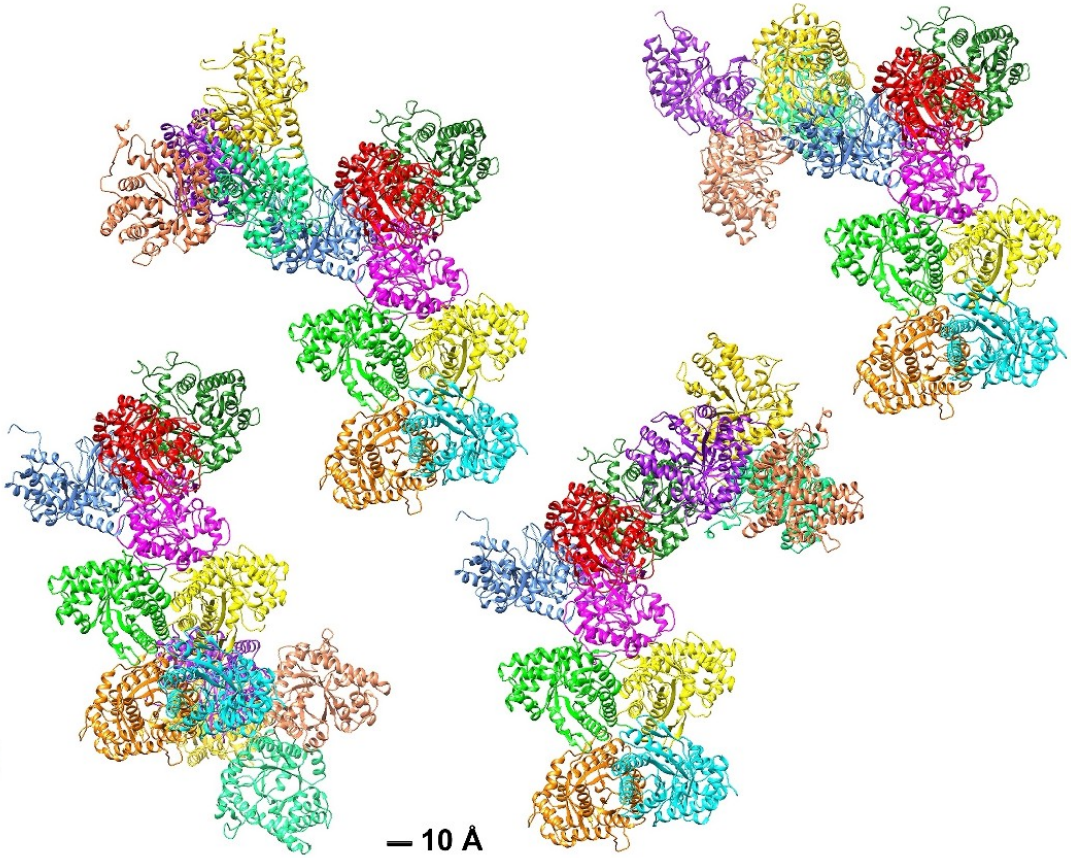


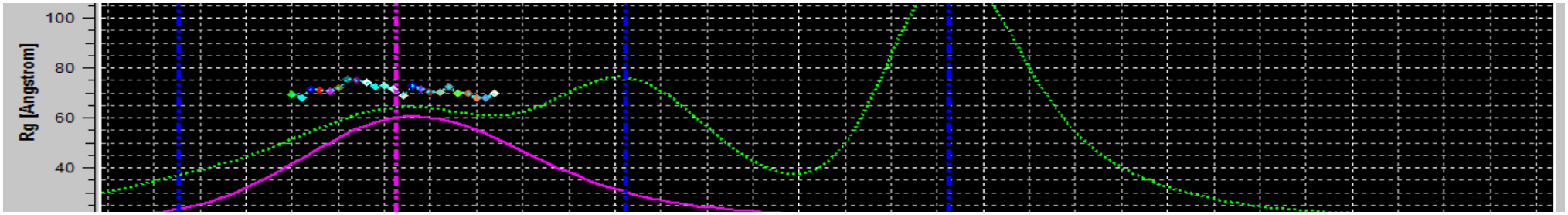


Brookes et al, J. Appl. Cryst. 49 (2016) 1827-41

Deconvolution of Aldolase & Model

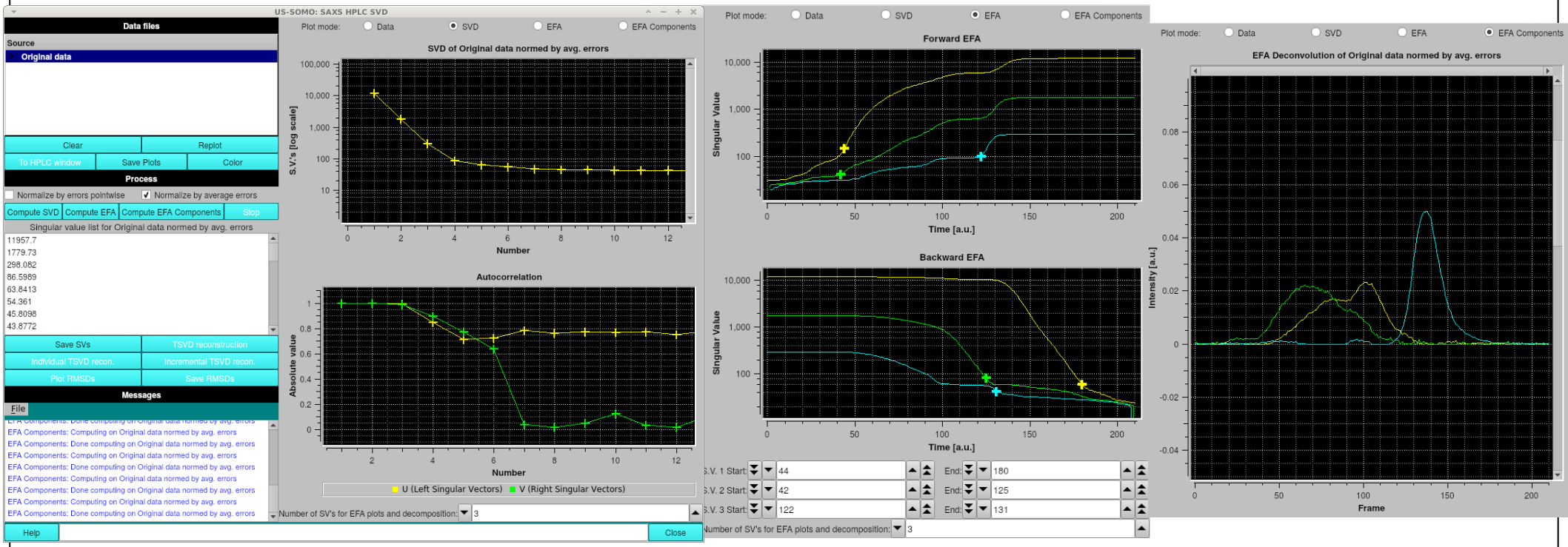






Brookes et al, J. Appl. Cryst. 49 (2016) 1827-41

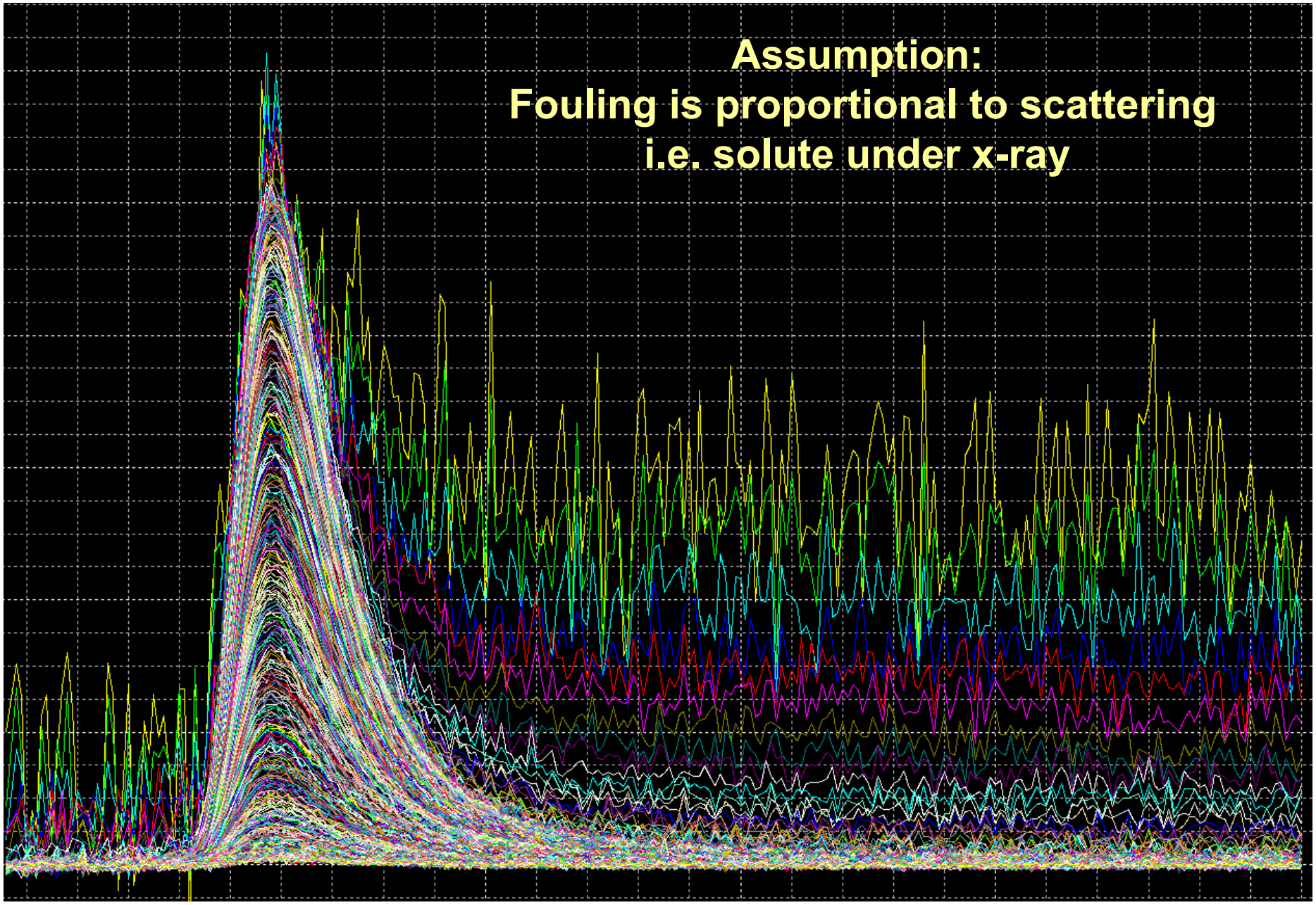
SVD/EFA



Hopkins et al. (2017) J. Appl. Cryst. 50(5) 1545-53
Meisburger et al. (2016) J. Am. Chem. Soc. 138 6506-16

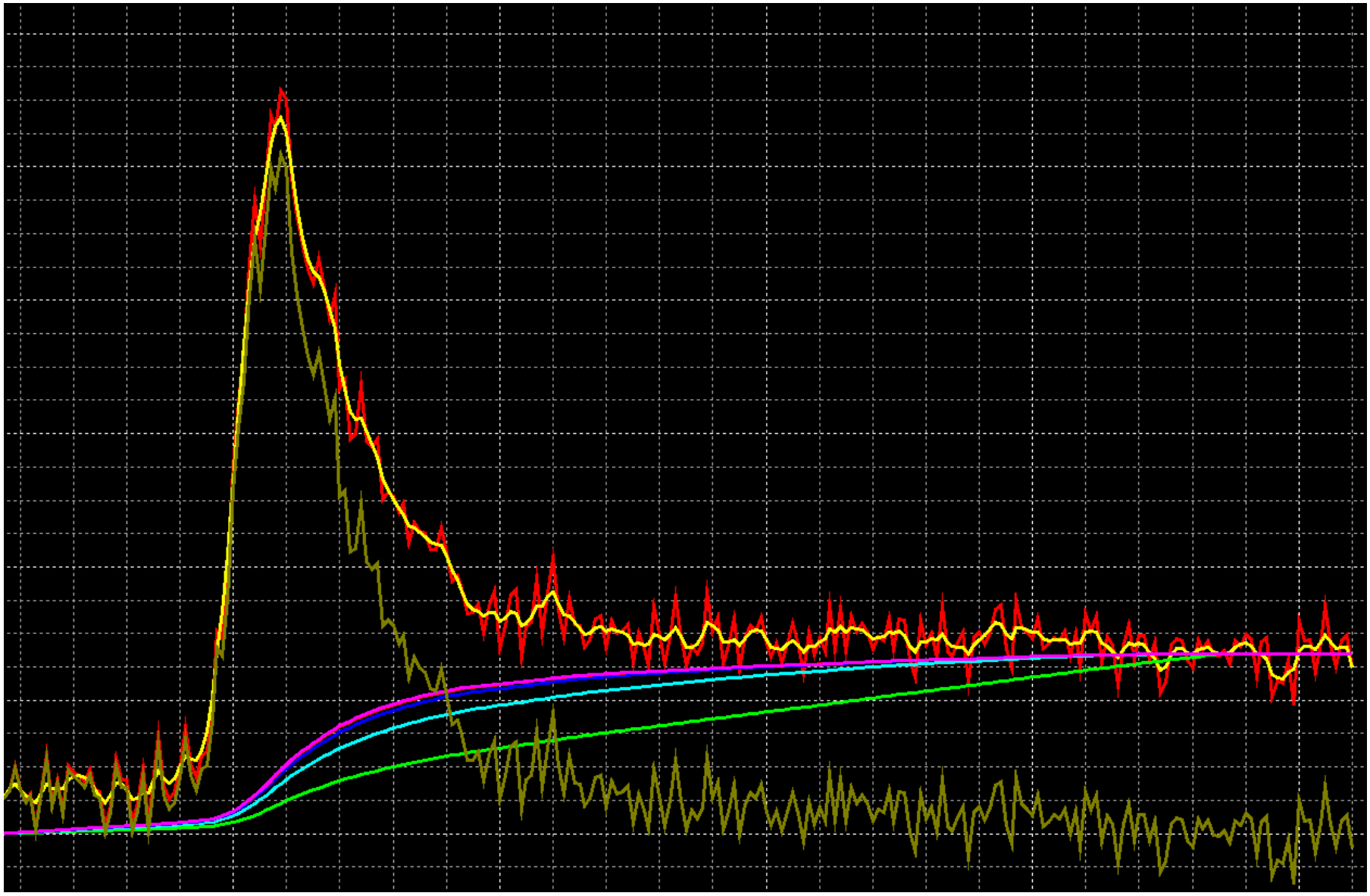
Brookes et al, J. Appl. Cryst. 49 (2016) 1827-41

SEC-SAXS I(t) profiles Lyzosyme with fouling



Brookes et al, J. Appl. Cryst. 49 (2016) 1827-41

I(t) fouling correction

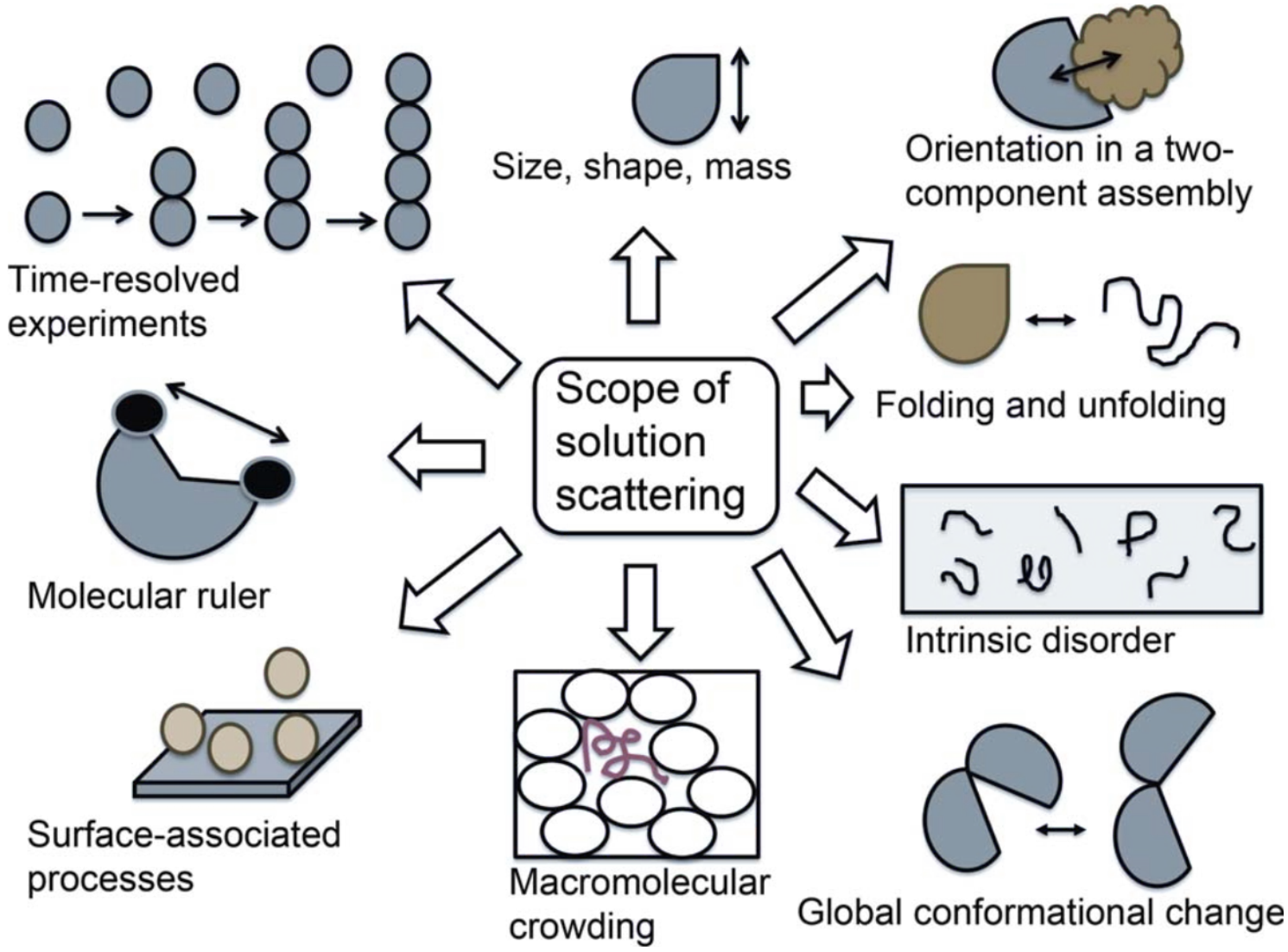


Brookes et al, J. Appl. Cryst. 49 (2016) 1827-41

SEC-SAXS

- If you have sufficient sample – use SEC-SAXS for biological macromolecules
- If you have true baseline separation, excellent, you should probably be ok simply taking the peak data, but global Gaussian decomposition will use all of the data
- If you do not have true baseline separation, be very careful and you should use these techniques

Scope of solution scattering



B. N. Chaudhuri. Prot. Sci. 2015 24:267-276

XFELs – SLAC/LCLS

XFELs

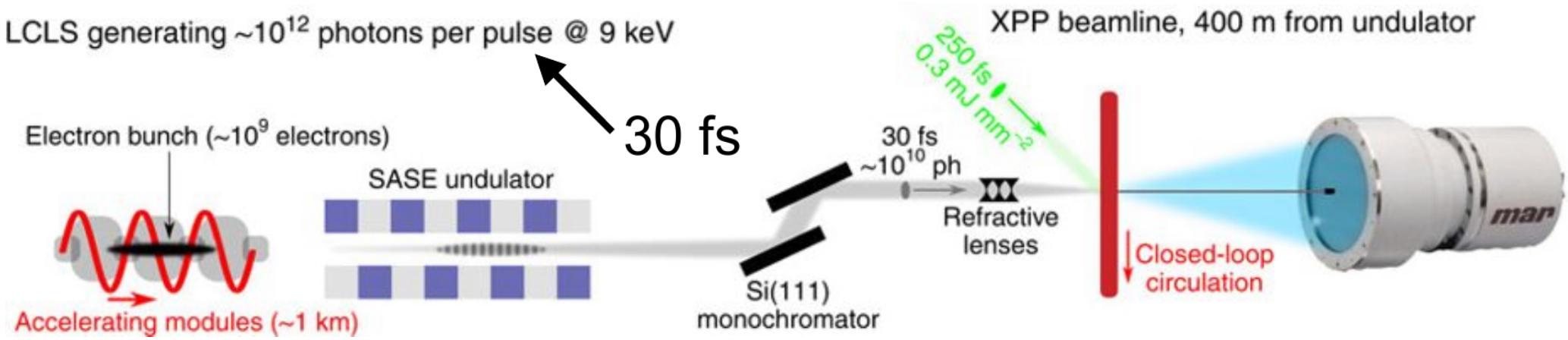
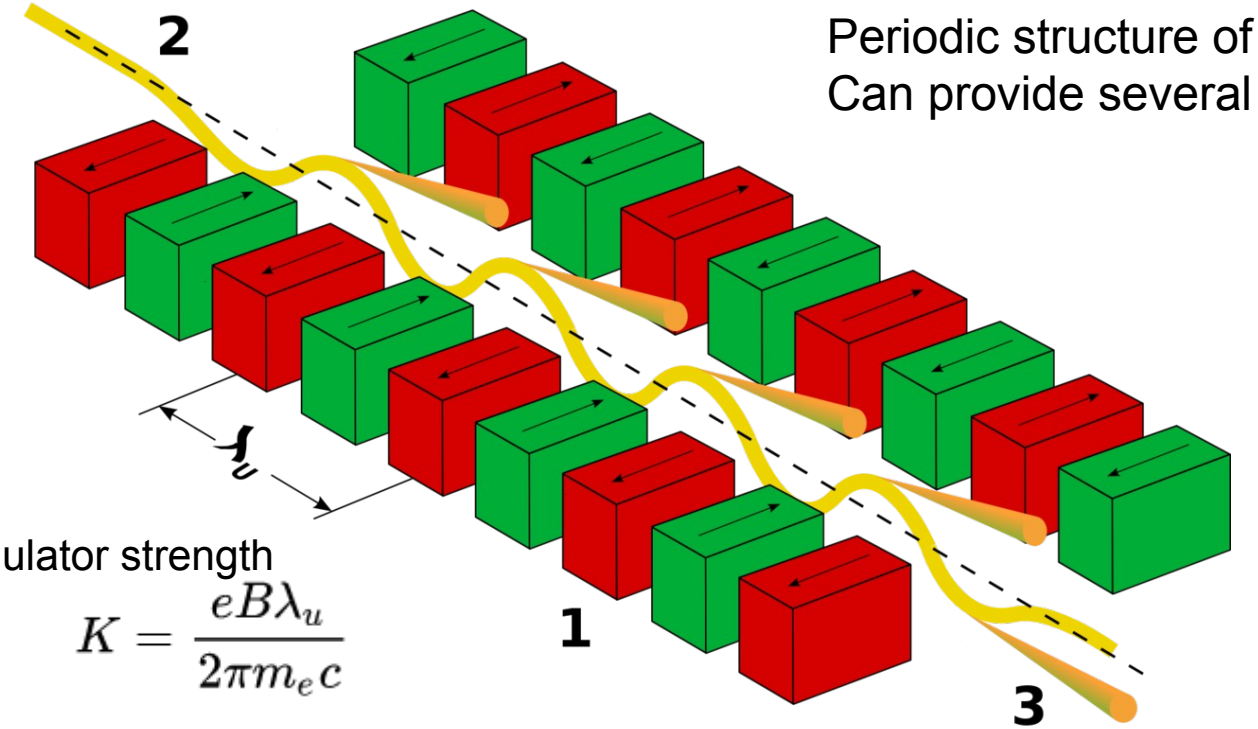


Image: Pearson

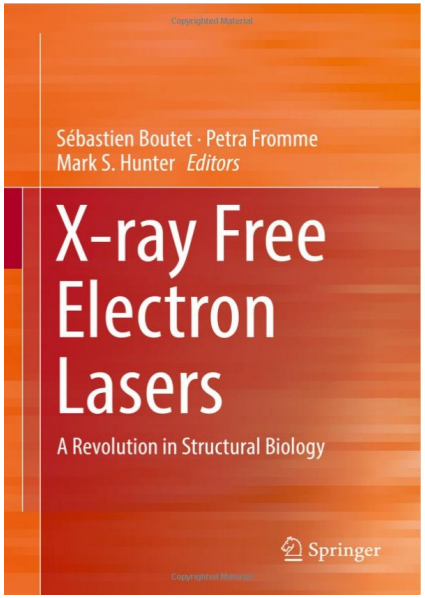


Periodic structure of dipole magnets
Can provide several orders of magnitude higher flux

Undulator strength

$$K = \frac{eB\lambda_u}{2\pi m_e c}$$

CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=537945



XFEL SAXS/WAXS

2 point correlations : $p(r)$ → 4 point correlations promises possibility of better 3d models

Still early days.

Few XFELs instruments

Serial access (vs synchrotron)

lightsources.org

European XFEL

Elettra Sincrotrone Trieste

FERMI at ELETTRA

DESY.

Flash at DESY

SLAC NATIONAL ACCELERATOR LABORATORY

LCLS at SLAC

PAL

POHANG ACCELERATOR LABORATORY

PAL-XFEL

PAUL SCHERRER INSTITUT PSI

Swiss FEL at PSI

Schools

Argonne NATIONAL LABORATORY

Beyond Rg Bio

https://small-angle.aps.anl.gov/future-courses#BeyondRgBio

OAK RIDGE National Laboratory

National School on Neutron and X-Ray Scattering

https://neutrons.ornl.gov/nxs

HERCULES European School

Organized by Université Grenoble Alpes & Grenoble Institute of Technology

UNIVERSITÉ Grenoble Alpes Grenoble INP

http://hercules-school.eu/

EMBO

Practical courses

http://embo.org/funding-awards/courses-workshops/practical-courses

Books

"La diffraction des rayons X aux très petits angles: Application a l'etude de phénomènes ultramicroscopiques":

A. Guinier (1939), Ann. de Phys., 11:12

pdf in course papers

"Small Angle Scattering":

A. Guinier and A. Fournet, (1955), in English, ed. Wiley, NY

"Small Angle X-Ray Scattering":

O. Glatter and O. Kratky (1982), Academic Press.

pdf available

http://physchem.kfunigraz.ac.at/sm/Software.htm

"Structure Analysis by Small Angle X-ray and Neutron Scattering":

L.A. Feigin and D.I. Svergun (1987), Plenum Press.

pdf available

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/reprints/feigin_svergun_1987.pdf

"Neutrons, X-Rays and Light, Scattering methods applied to soft condensed matter":

P. Lindner and T. Zemb Eds, (2002) Elsevier, North-Holland.

The Proceedings of the SAS Conferences held every three years are usually published in the Journal of Applied Crystallography