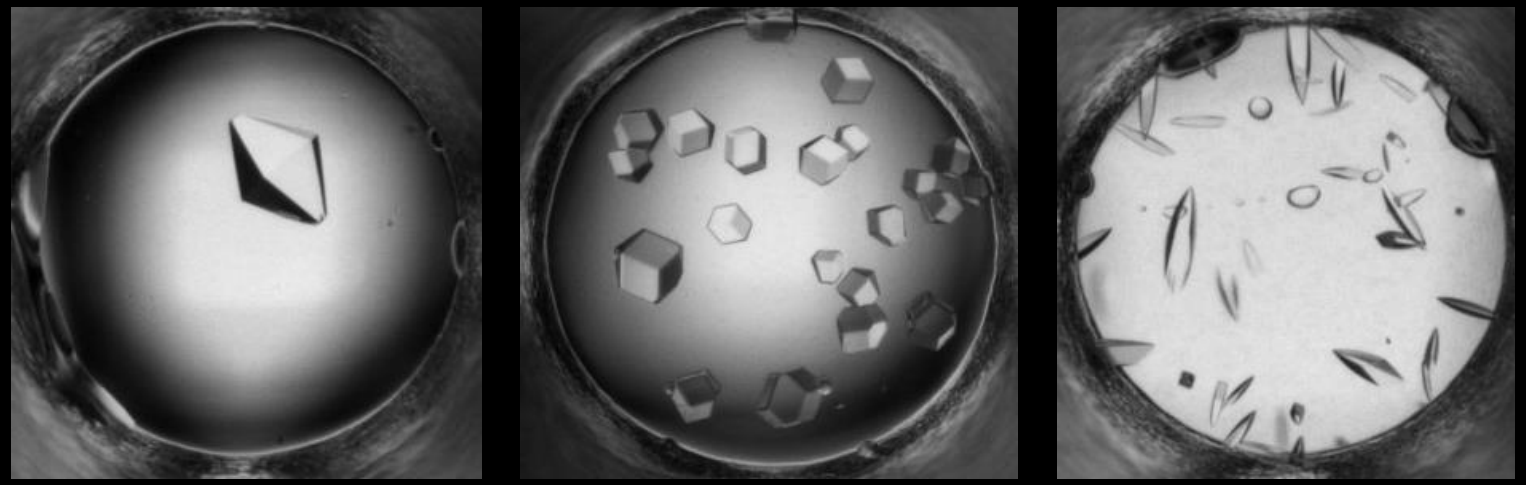
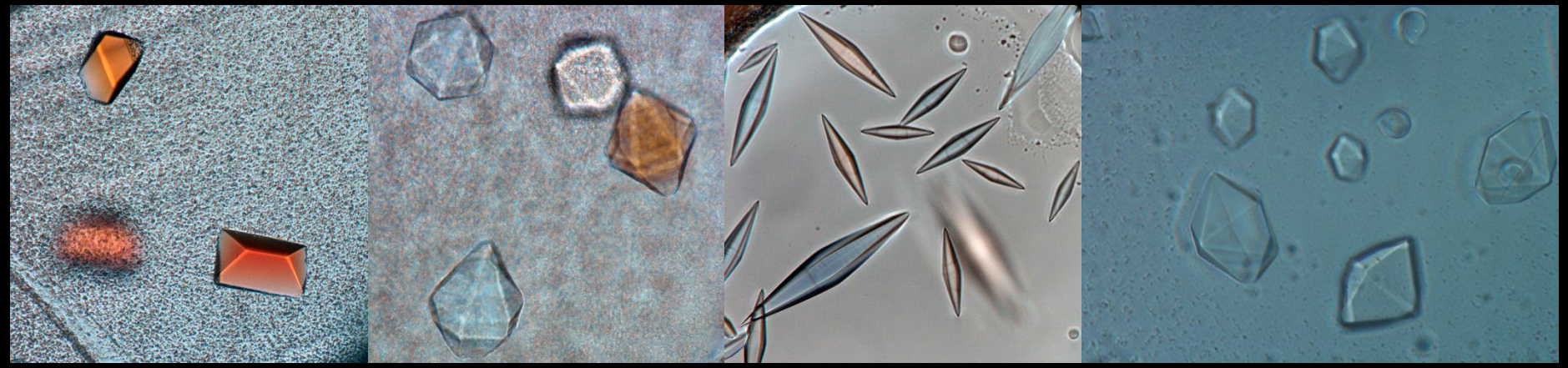
Small-Angle Scattering as a complementary technique in structural biology

Edward H. Snell
CEO Hauptman-Woodward Medical Research Institute



Crystallography Requires Crystals



No crystal …

No crystallography ….

No crystallographer ….

Jean Baptiste Fourier (1763-1830)

In 1807 came up with an idea ….

Any periodic function could be rewritten as a weighted sum of sines and cosines of different frequencies.

This was not regarded as possible by other mathematicians of the time and it was not until 1878 that the idea was published in English.

The name of the idea is the Fourier series

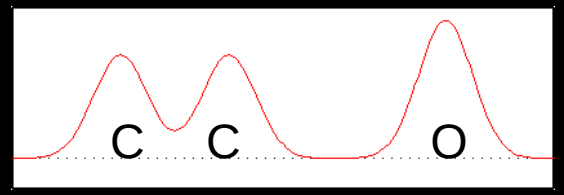
To change a signal to a Fourier series we perform a Fourier transform. To change the Fourier transform to a signal we perform an inverse Fourier transform.

Fourier's theorem is not only one of the most beautiful results of modern analysis, but it may be said to furnish an indispensable instrument in the treatment of nearly every recondite question in modern physics.

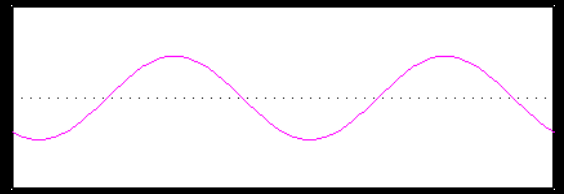
Lord Kelvin (1824-1907)

Other Lord Kelvin quotes

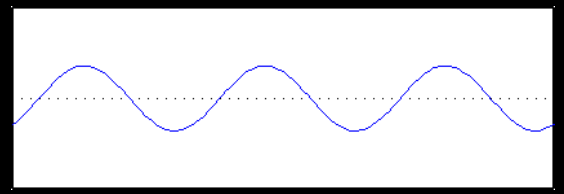
- Heavier-than-air flying machines are impossible.
- Radio has no future.
- In science there is only physics; all the rest is stamp collecting.



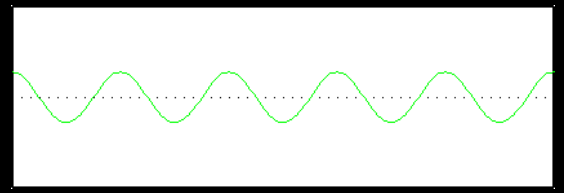
A unit cell of a crystal with 2 carbons and an oxygen



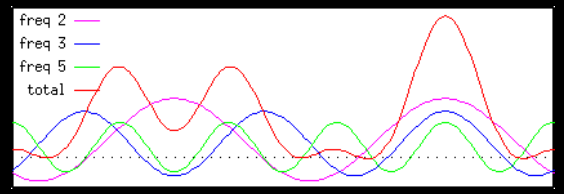
A cosine wave with frequency of 2, one peak represents the oxygen and the other the two carbons



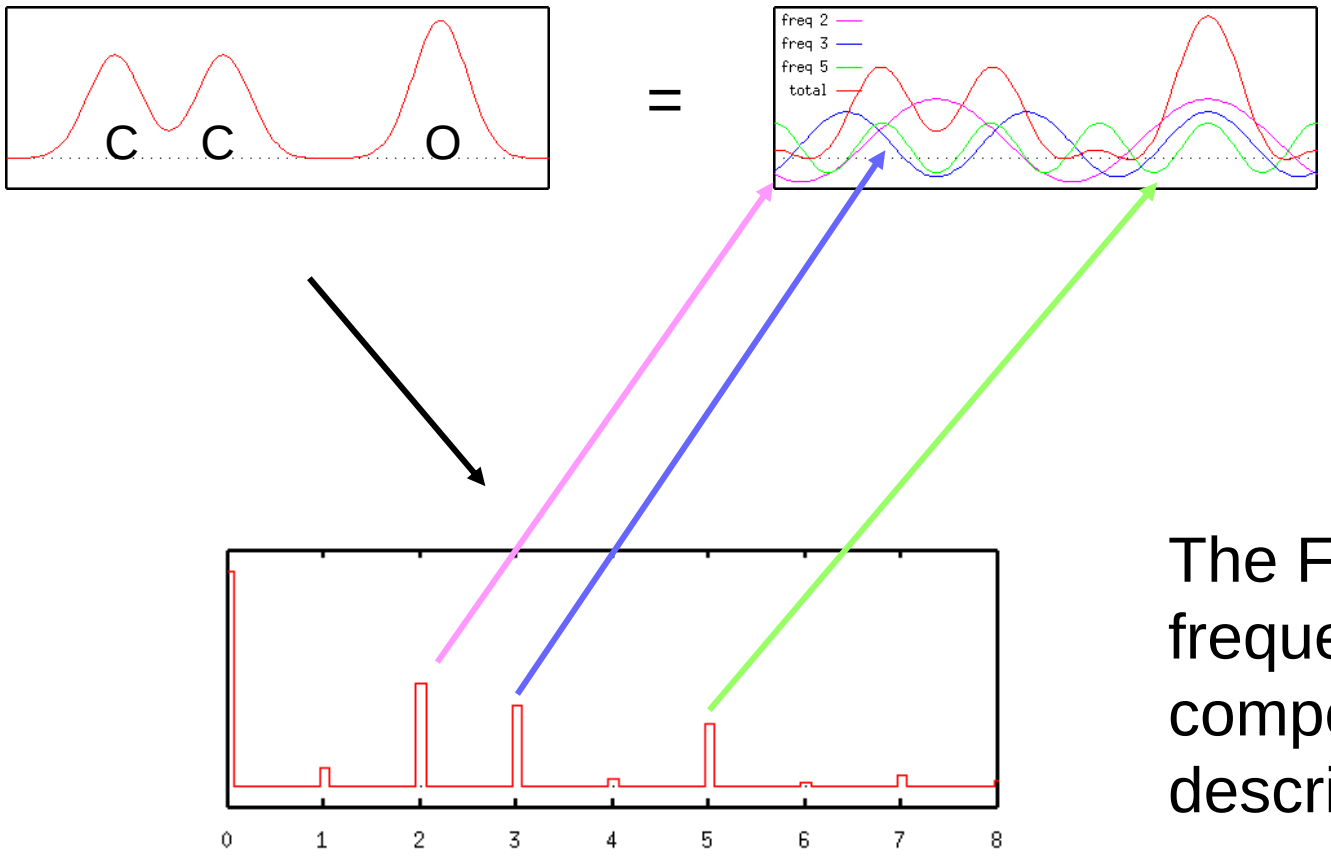
Add a cosine wave with frequency of 3, three repeats across the crystal. Note the phase is different, it starts in a different place.



Add a third cosine wave with frequency of 5, with the peaks lined up on the carbons



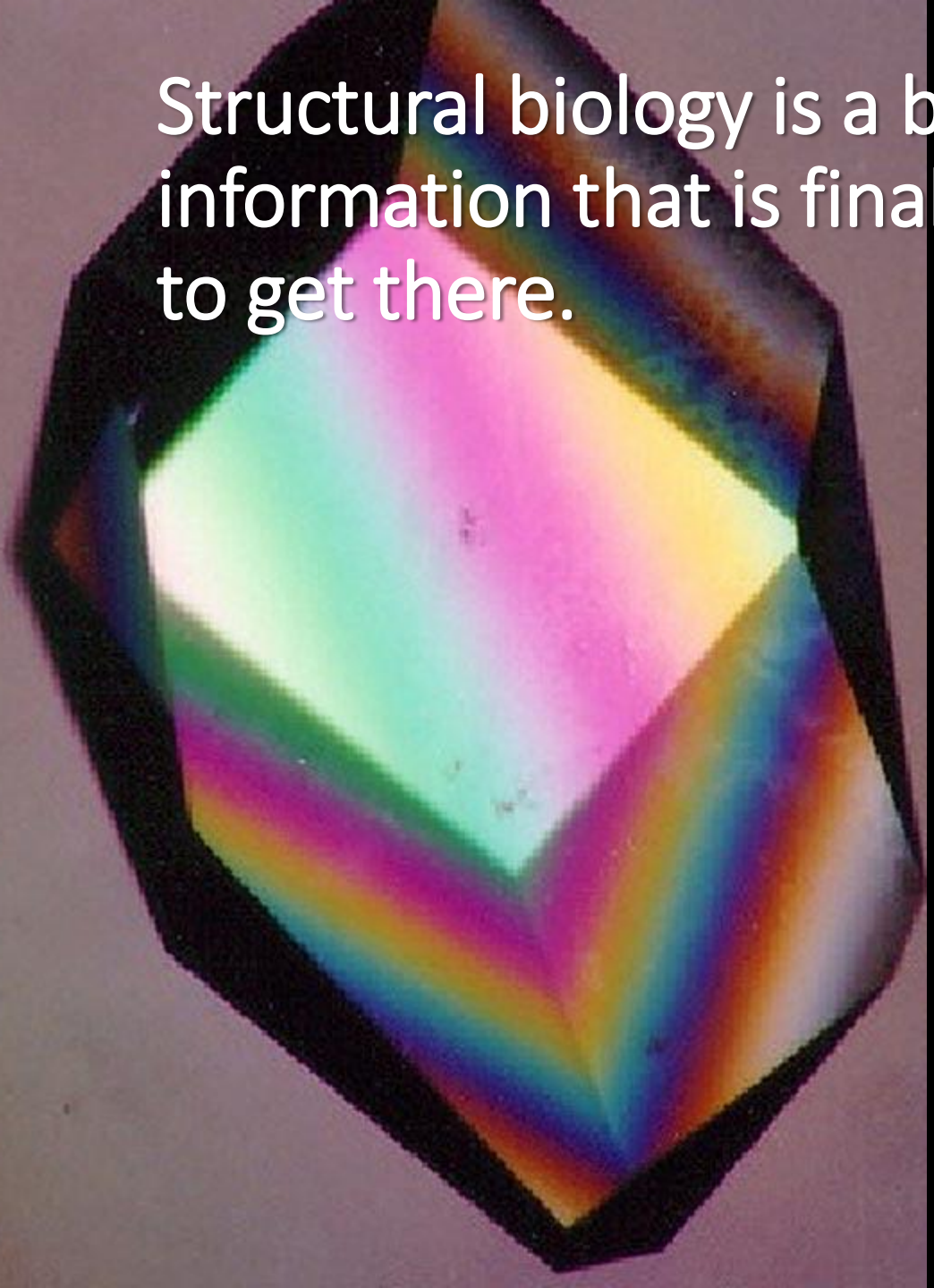
Add all the waves and the result is the original unit cell of the crystal

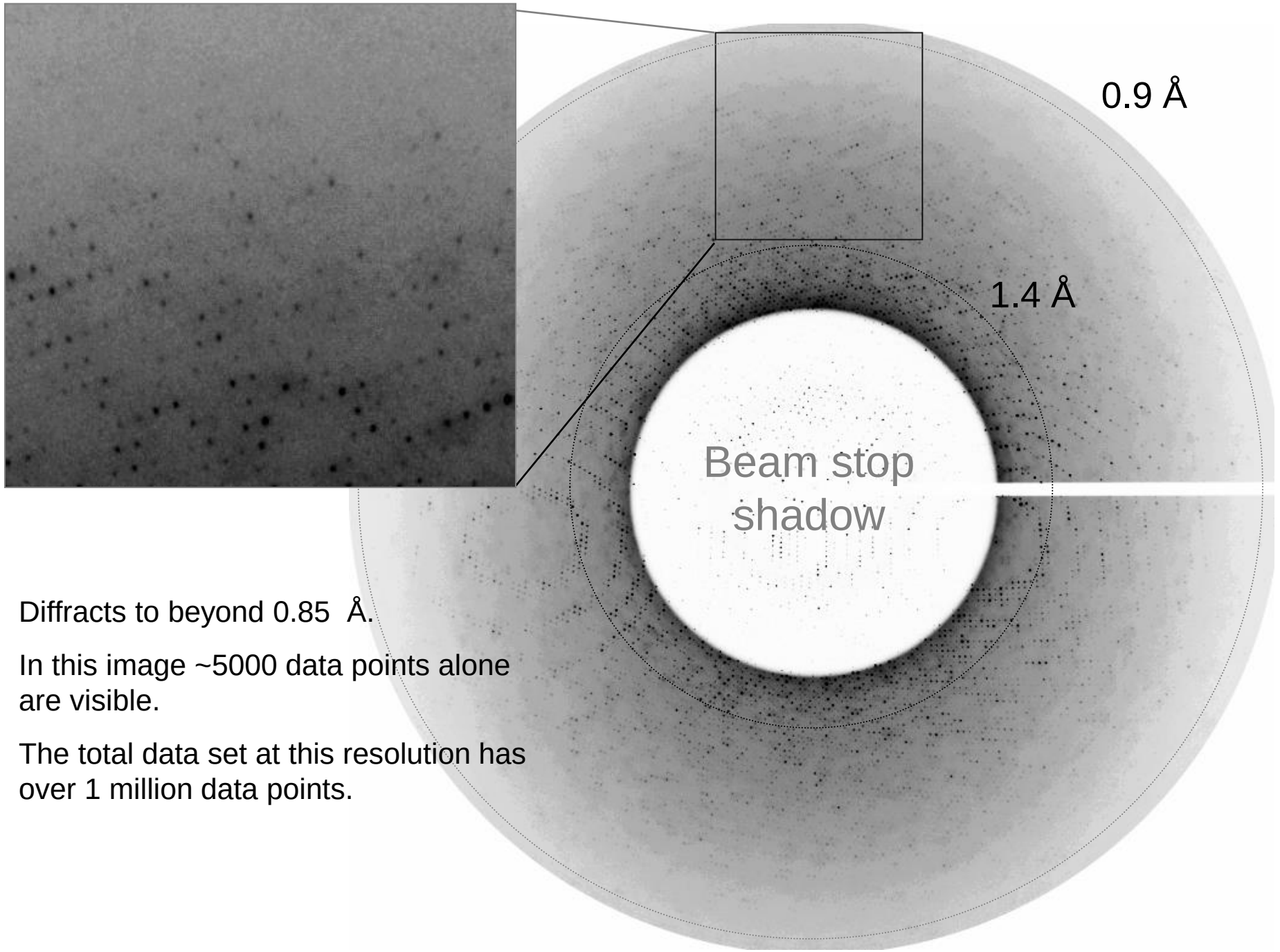


The Fourier transform of the unit cell

The Fourier transform gives frequency information on the components of the waves that describe the real space.

Structural biology is a beautiful pursuit, not only in the information that is finally produced, but also the journey to get there.





Diffracts to beyond 0.85 Å.

In this image ~5000 data points alone are visible.

The total data set at this resolution has over 1 million data points.


Crystallize
Now

No crystal …

No crystallography ….

No crystallographer ….

However …

- It is possible to get low resolution structural information from a protein or complex in solution.
- This can tell you about the foldeness and dynamics of the system.
- It can position known structural information in a complex.
- It can determine the area sampled by flexible regions not resolvable crystallographically.
- It is not limited to the chemistry where crystallization occurs.
- It can determine if gross structural changes occur.
- It can be used to provide information to guide crystallization
- New algorithms may enable direct electron density determination.

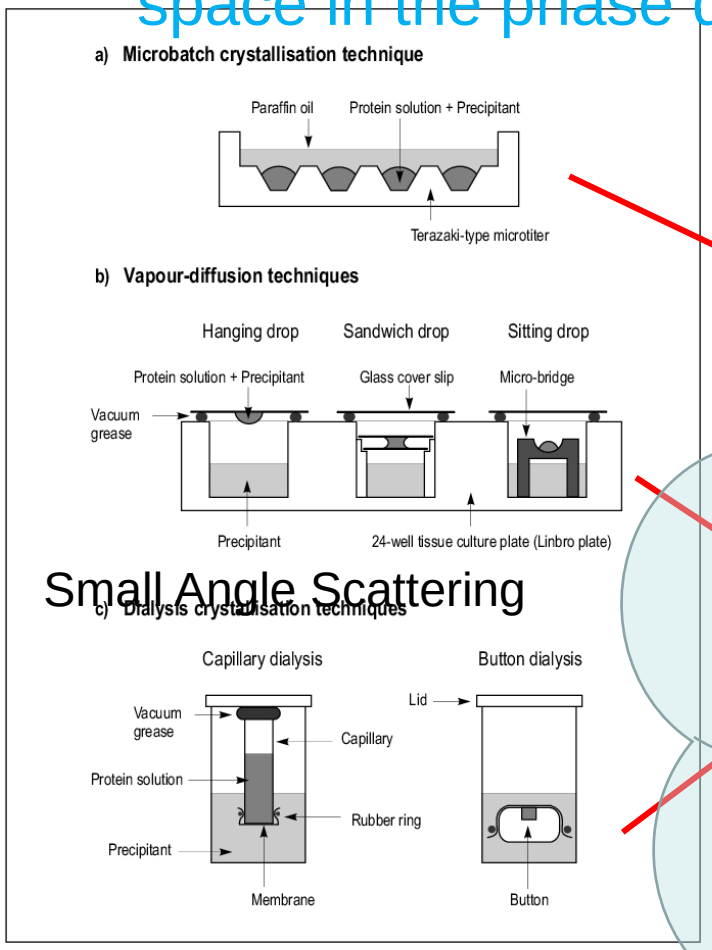
Introduction to Small Angle Solution Scattering (X-ray or Neutron)

SAXS and SANS

(one of several complementary techniques)

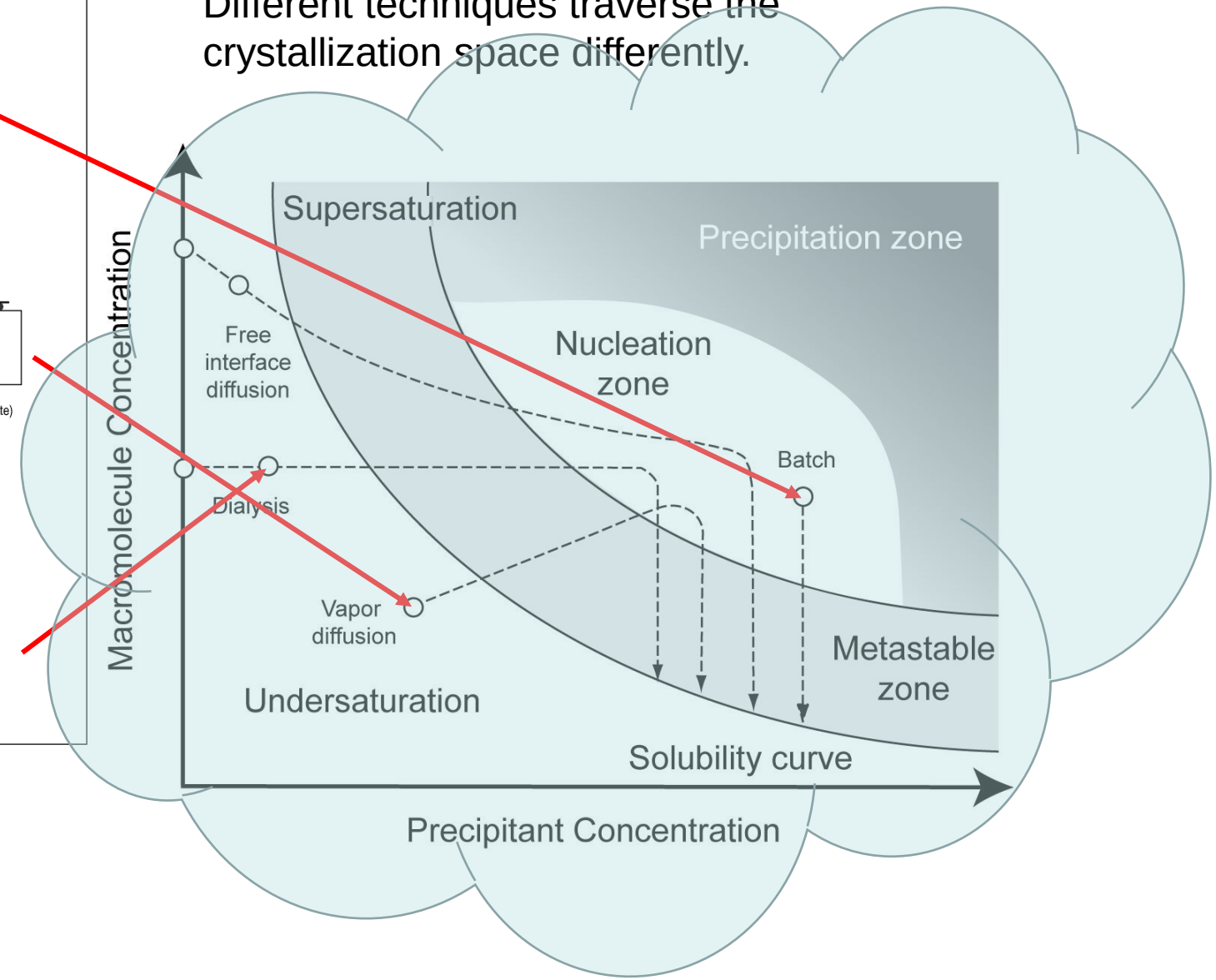
Mandatory phase diagram slide

Small Angle Scattering can provide information from all chemical space in the phase diagram but we will focus on undersaturation

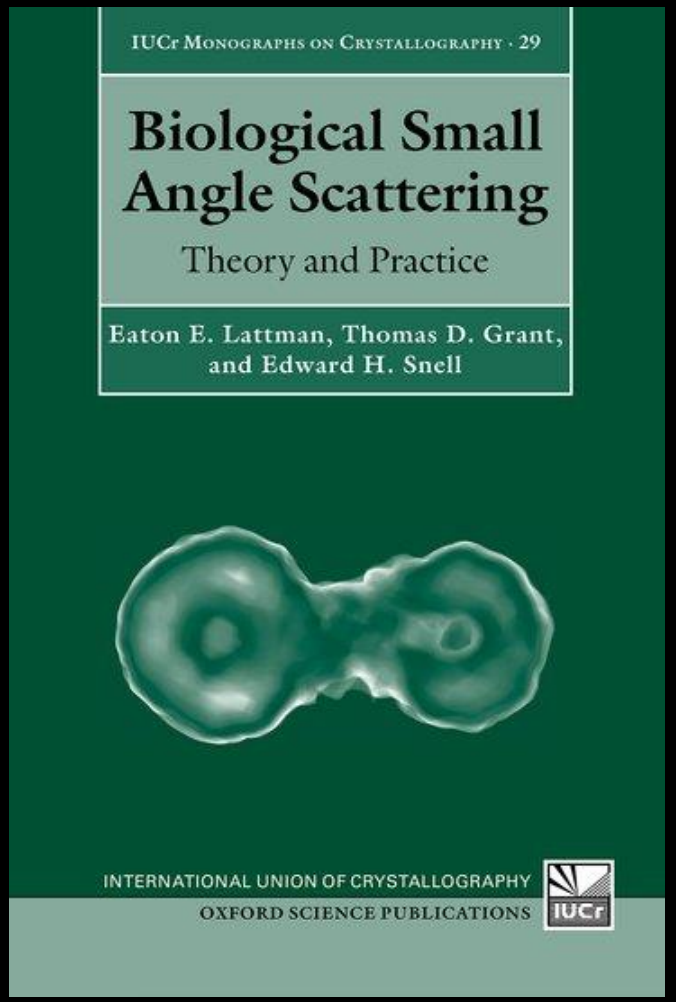


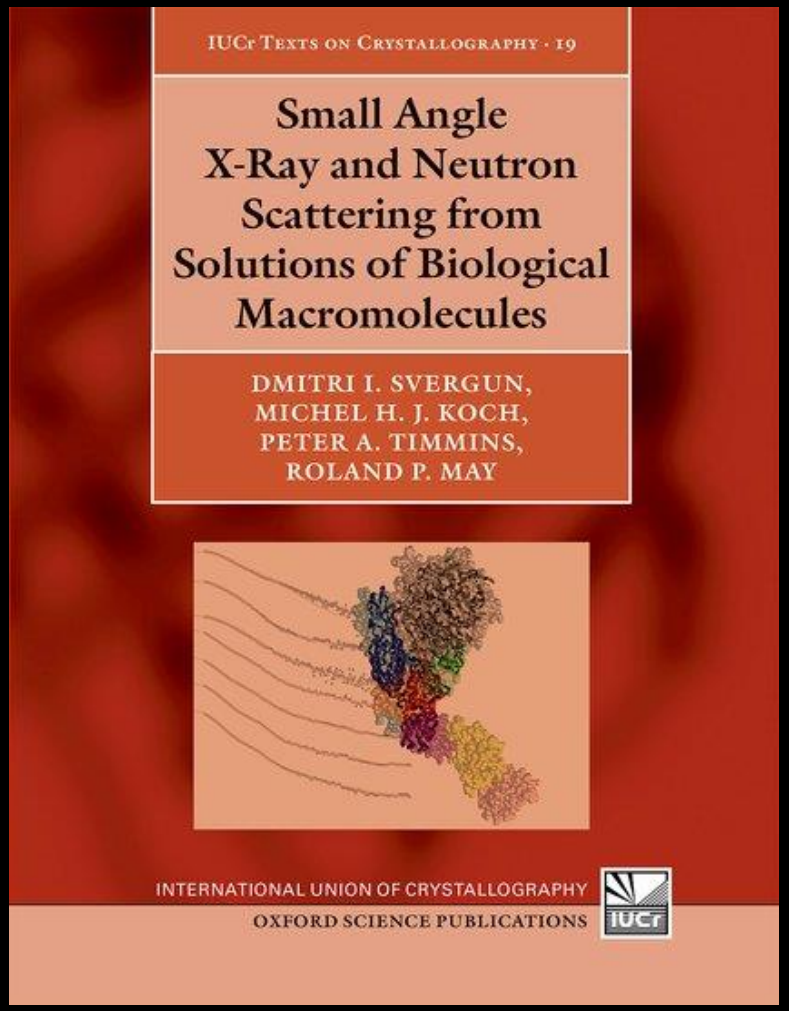
Small Angle Scattering

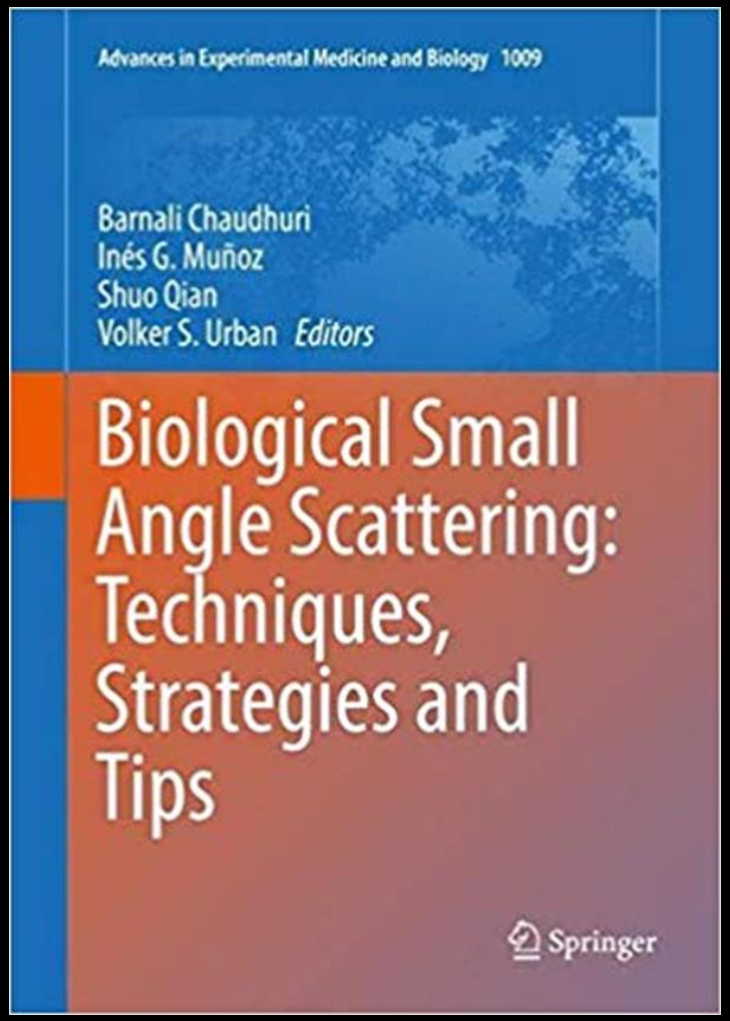
Different techniques traverse the crystallization space differently.



Good books are available





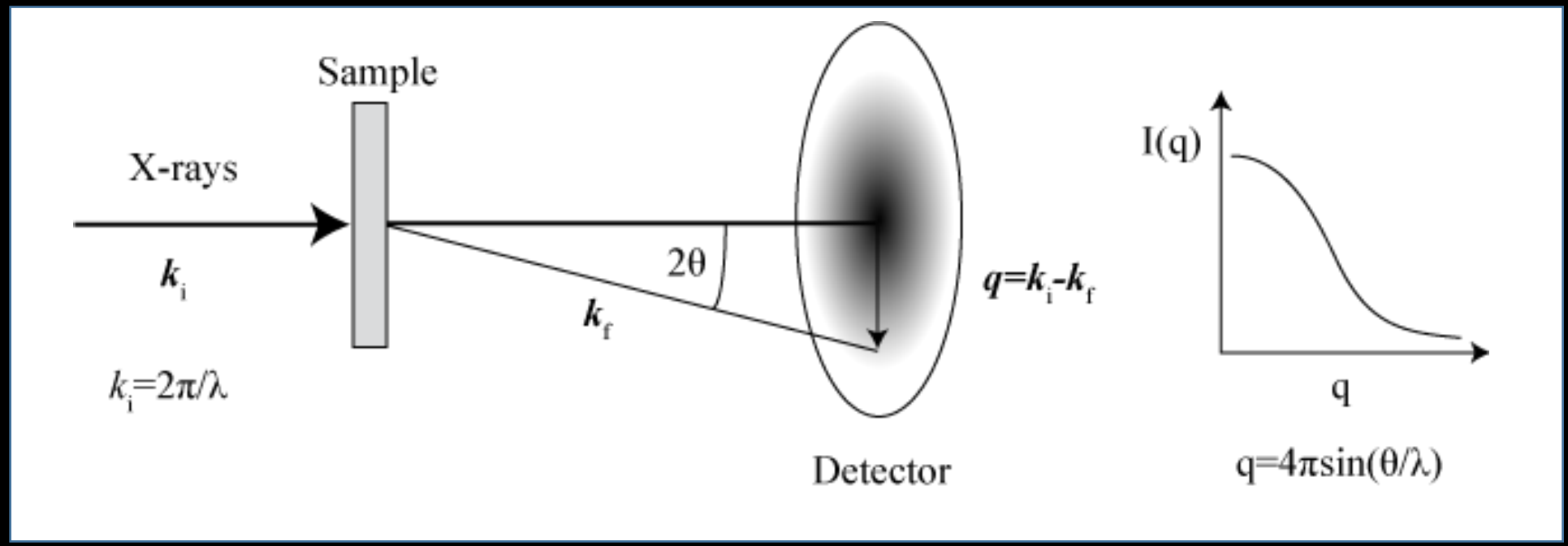


- Excellent practical guide - X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution Putnam et al Quarterly Reviews in Biophysics, 2007

Small Angle Scattering

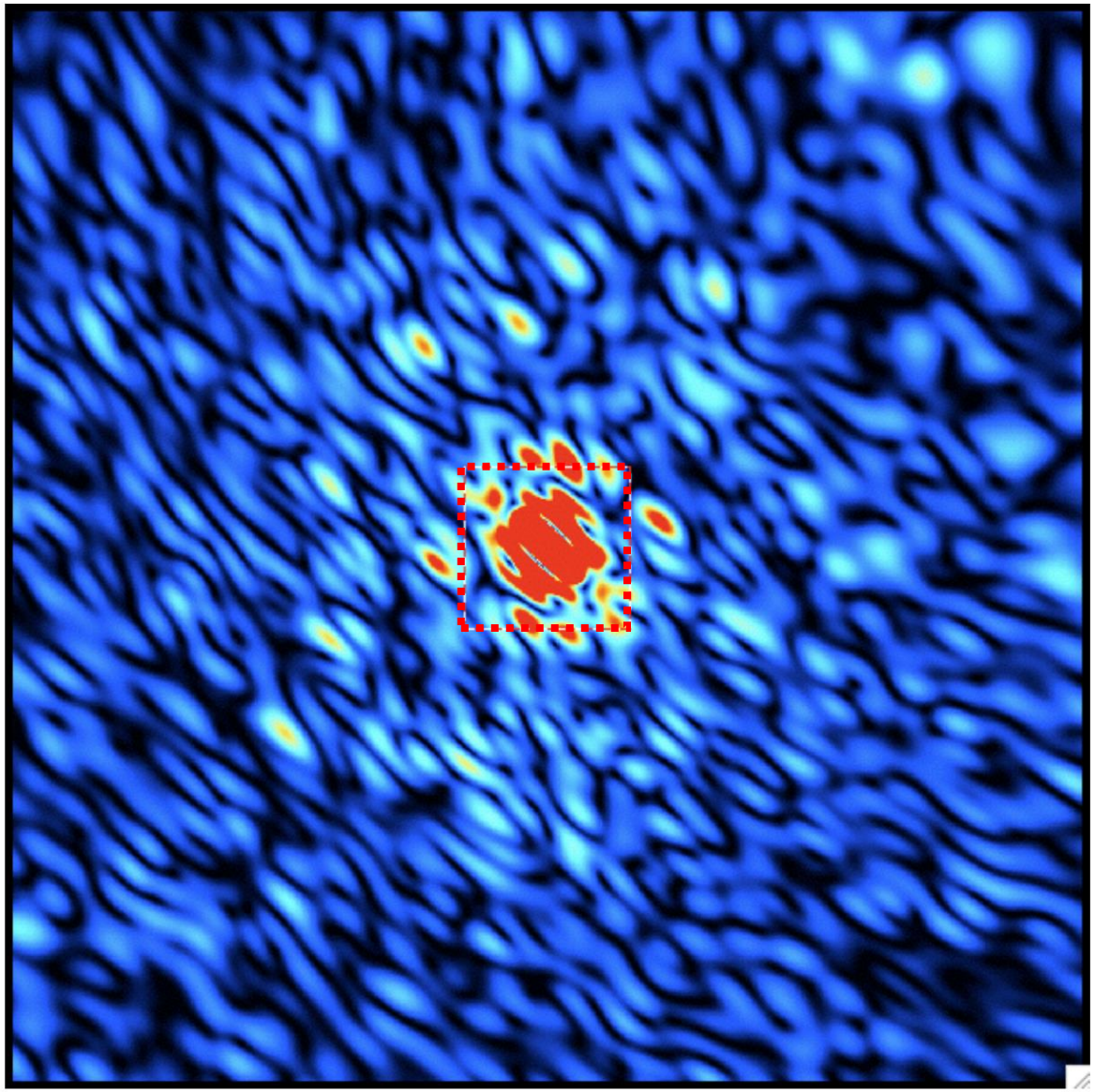
- A solution is illuminated with a parallel, monochromatic X-ray or neutron beam, and the scattered radiation is collected on a detector placed far back from the specimen. Because the solution is homogeneous and isotropic, the observed scattering pattern is circularly symmetric. The full pattern, a three-dimensional function in diffraction space, is spherically symmetric.
- The term solution scattering is applied to the general phenomenon, with the term small angle scattering reserved for the most common application in which observations are confined to radiation scattered within a small angular cone around the main beam.
- Both X-rays and neutrons are used with the terminology Small-Angle X-ray scattering (SAXS) and Small-Angle Neutron Scattering (SANS).
- For the most part, discussions on SAXS and SANS are interchanagable but each has specific advantages and disadvantages

SAXS is everything behind the beamstop

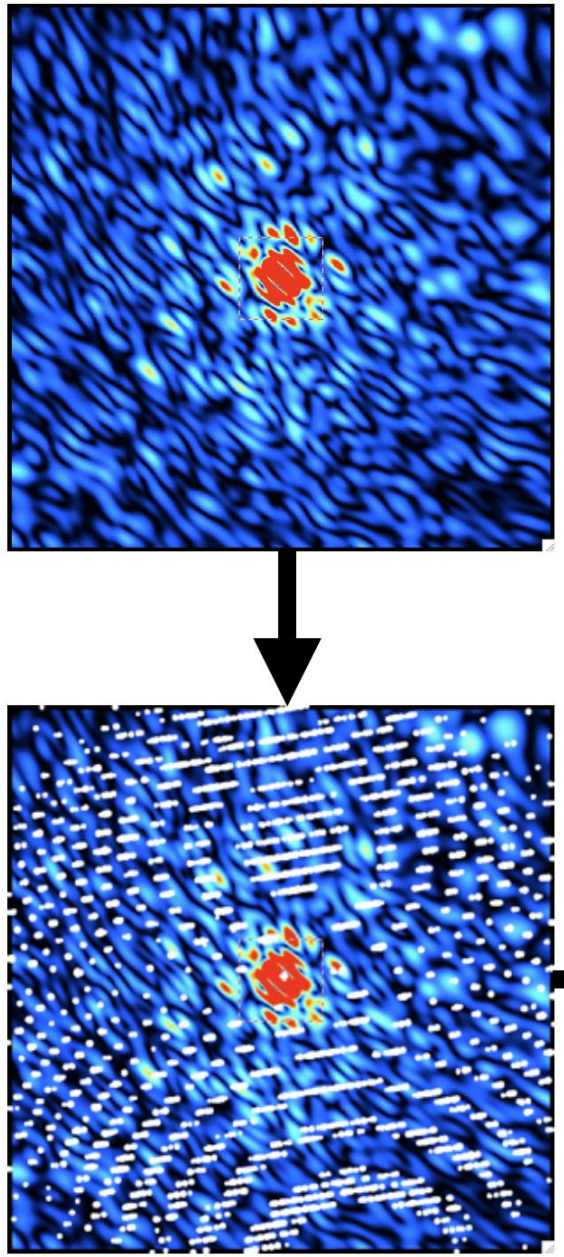


- Particles in solution tumble – spherically averaged intensity is recorded.
- Radial integration results in one dimensional SAXS profile.
- Larger particles scatter at smaller angles.
- Analysis of the 1D profile yields information about size and shape.

Molecular Transform

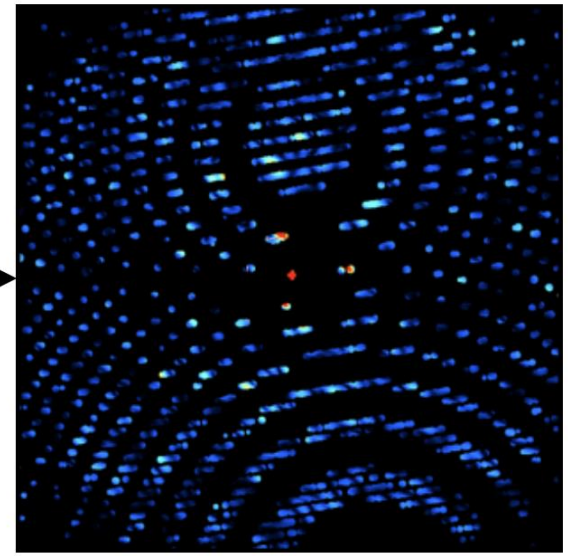


See Intro_to_SAXS.pdf at www.BioXFEL.org
Seminars by Thomas Grant

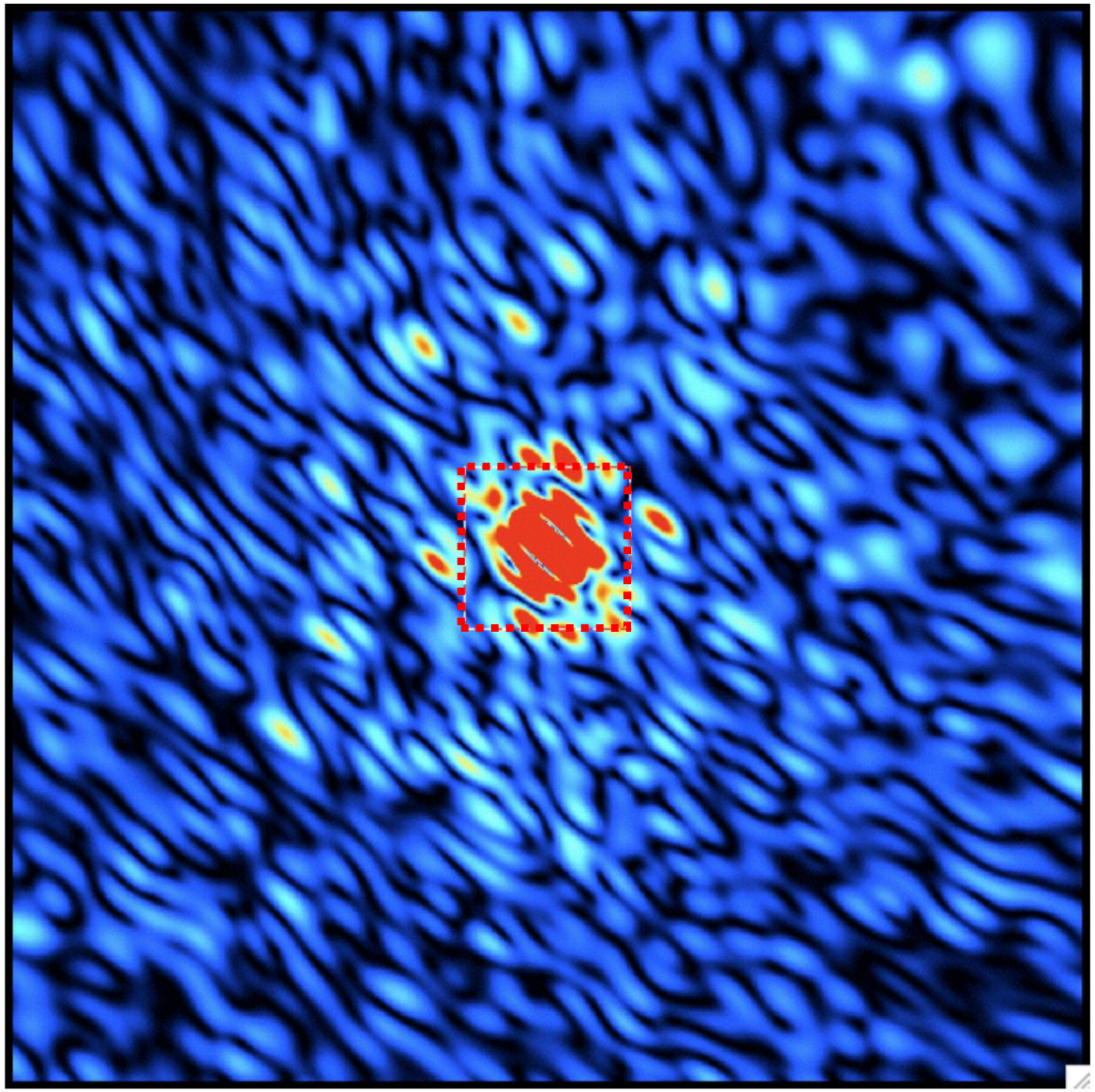


Molecular Transform

Bragg Sampling from
X-ray Crystallography

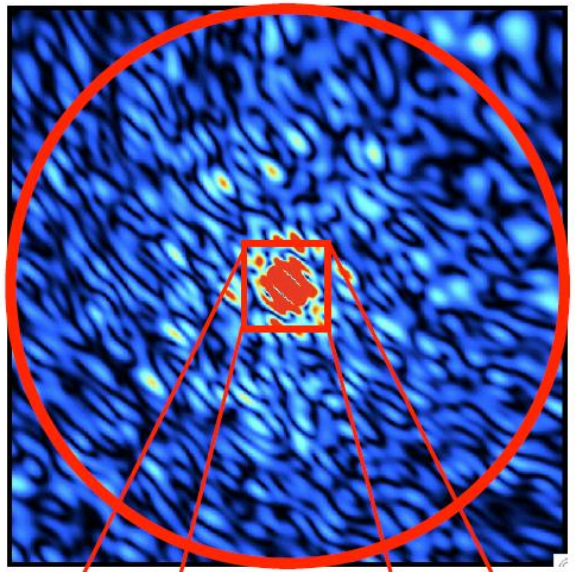


Molecular Transform



See Intro_to_SAXS.pdf at www.BioXFEL.org
Seminars by Thomas Grant

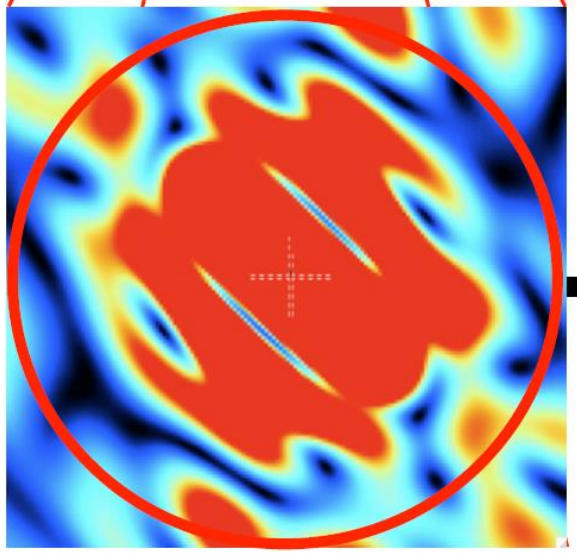
2Å

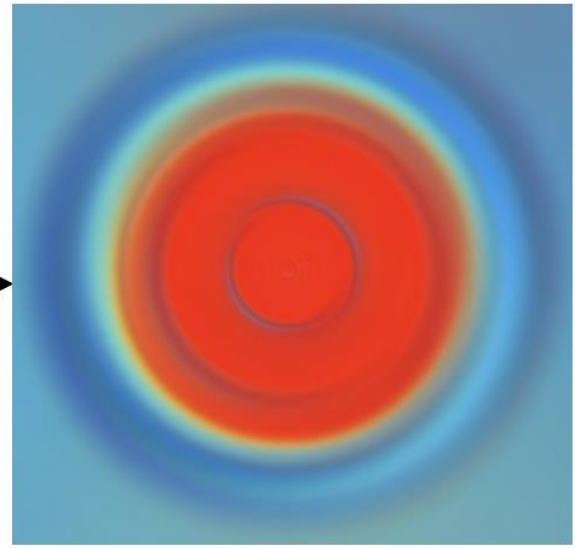


Molecular Transform

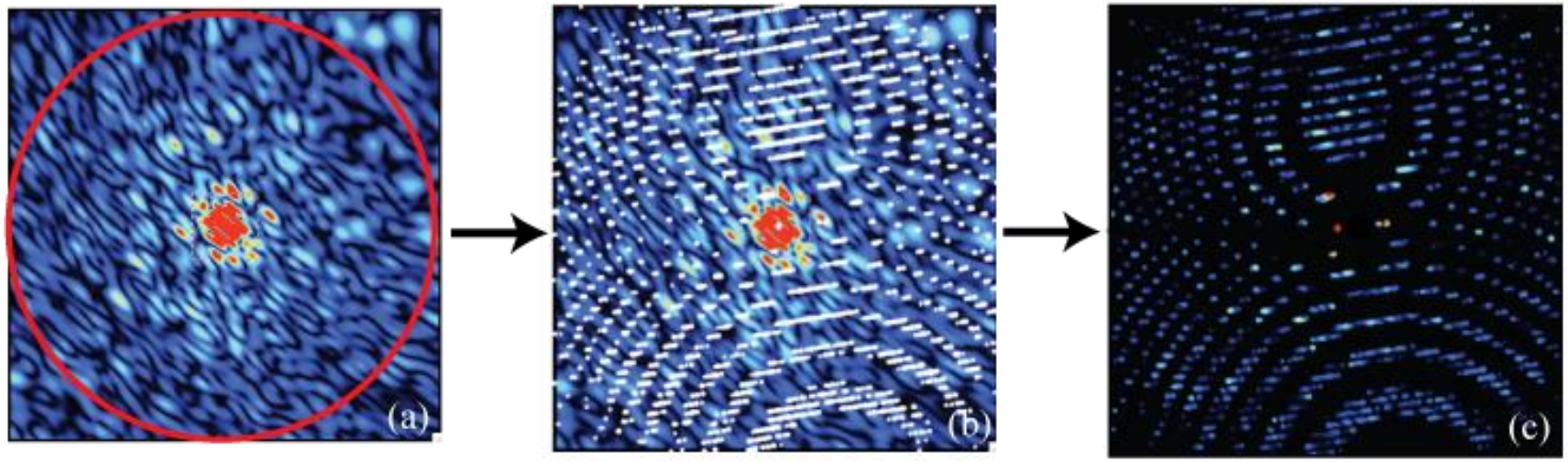
Spherical averaging from solution of tumbling molecules

20Å

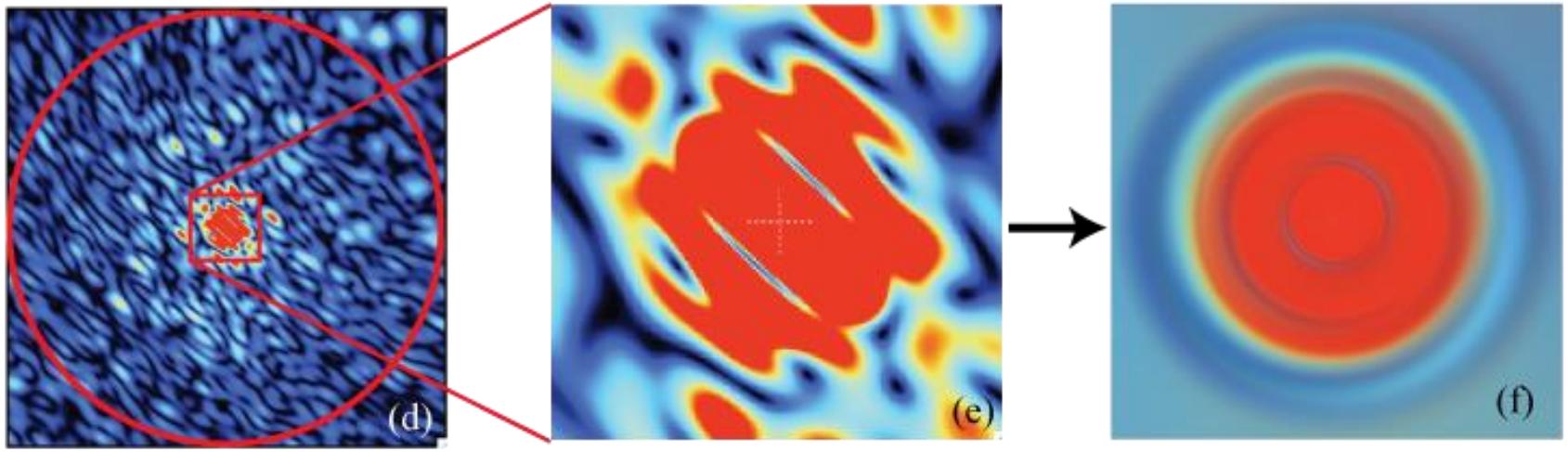




Single Crystal Diffraction



Small Angle Scattering

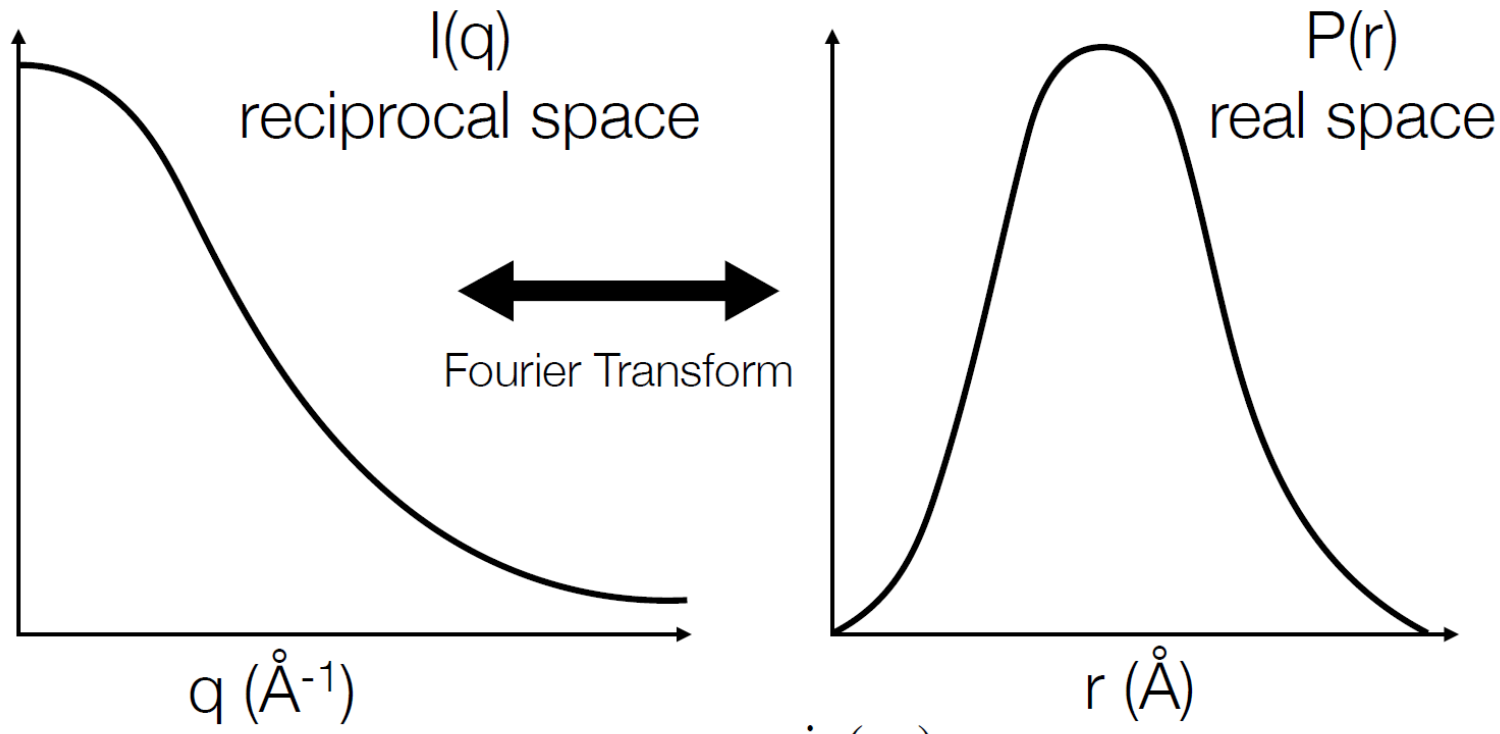


2Å

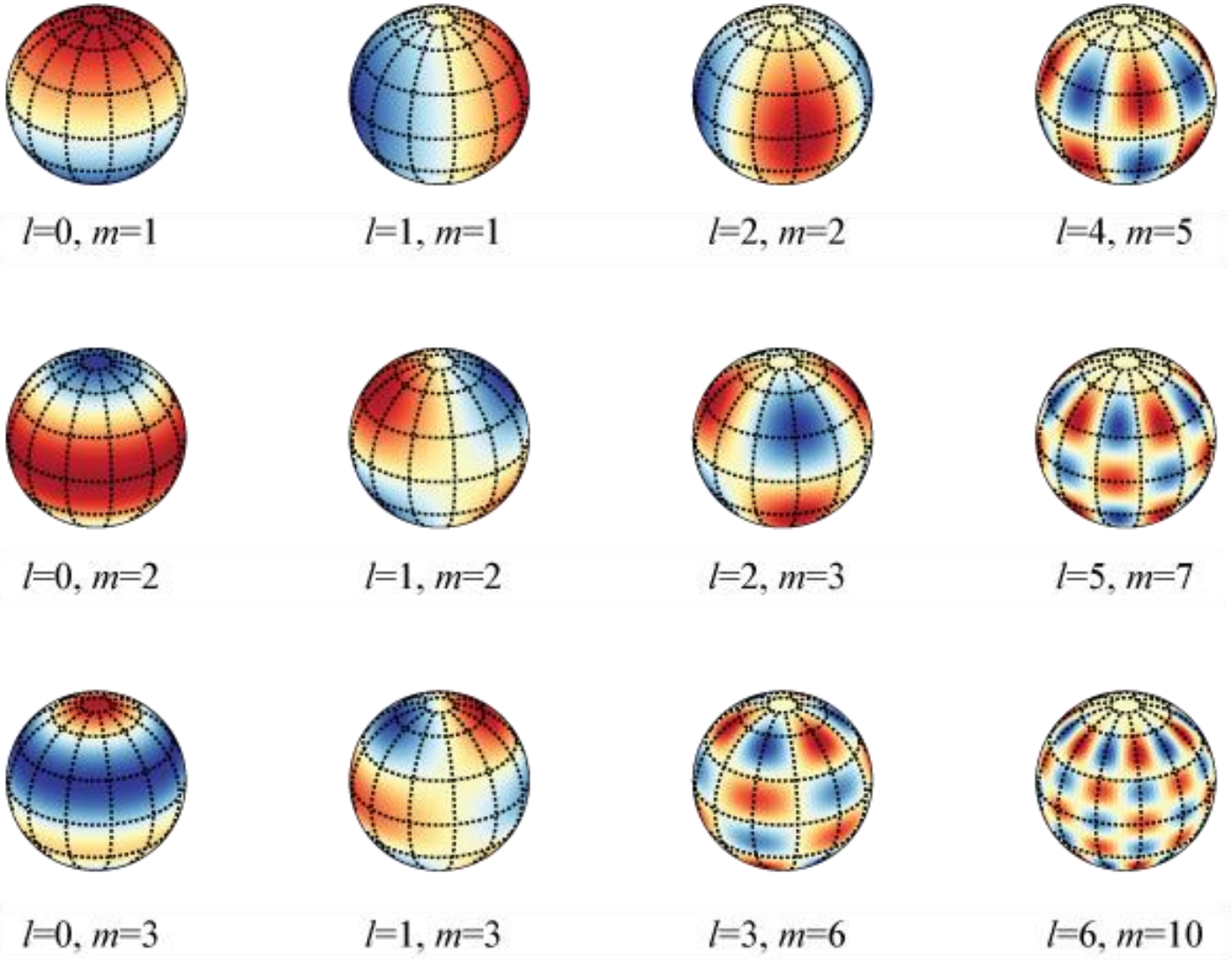
20Å

Scattering is in Fourier space, transform to real space

(The theoretical basis of SAXS is less complex that of crystallography)



$$I(q) = \int p(r) \frac{\sin(qr)}{qr} dr$$

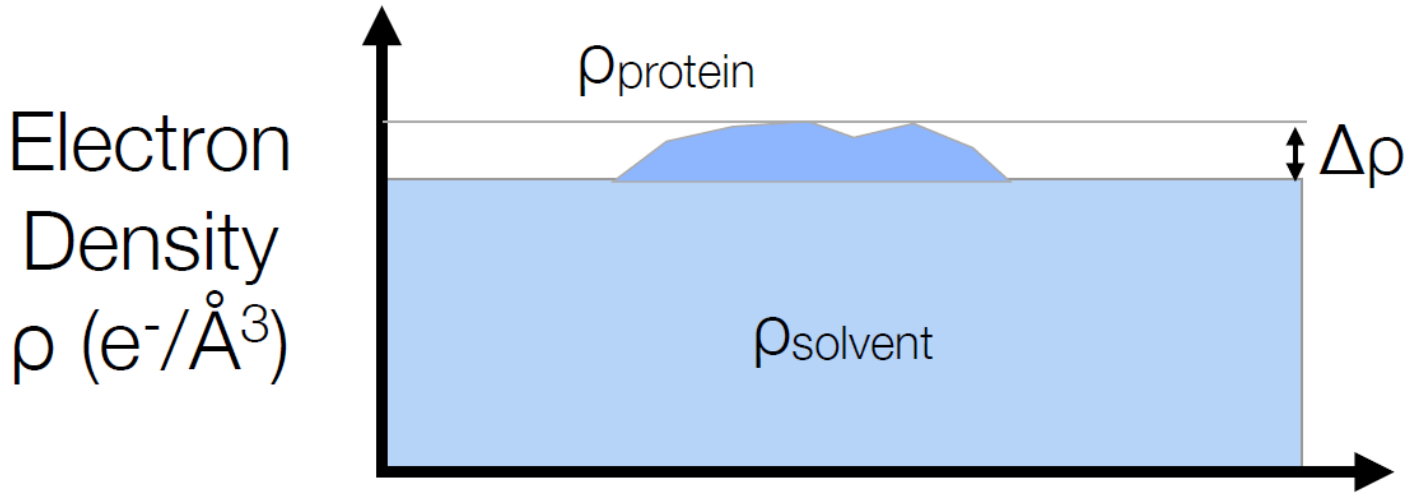


Spherical harmonics – the Fourier Transform of SAXS

An efficient means of three dimensional Fourier transform

SAXS is a Contrast Technique

- SAXS is a contrast method, i.e. it depends on the square of the difference in the electron density between the molecule and the solvent



$(\Delta\rho)^2 = (\rho_{protein} - \rho_{water})^2 = (0.44\text{-}0.33)^2 \approx$ 10% above background

research papers

STRUCTURAL BIOLOGY

ISSN 2059-7983

2017 publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution: an update

Jill Trewhella,[a]* Anthony P. Duff,[b] Dominique Durand,[c] Frank Gabel,[d] J. Mitchell Guss,[a] Wayne A. Hendrickson,[e] Greg L. Hura,[f] David A. Jacques,[g] Nigel M. Kirby,[h] Ann H. Kwan,[a] Javier Pérez,[i] Lois Pollack,[j] Timothy M. Ryan,[h] Andrej Sali,[k] Dina Schneidman-Duhovny,[l] Torsten Schwede,[m] Dmitri I. Svergun,[n] Masaaki Sugiyama,[o] John A. Tainer,[p] Patrice Vachette,[c] John Westbrook[q] and Andrew E. Whitten[b]

Received 16 June 2017
Accepted 7 August 2017

Edited by M. Czjzek, Station Biologique de Roscoff, France

Keywords: small-angle scattering; SAXS; SANS; biomolecular structure; proteins; DNA; RNA; structural modelling; hybrid structural modelling; publication guidelines; integrative structural biology.

Supporting information: this article has supporting information at journals.iucr.org/d

[a]School of Life and Environmental Sciences, The University of Sydney, NSW 2006, Australia, [b]ANSTO, New Illawarra Road, Lucas Heights, NSW 2234, Australia, [c]Institut de Biologie Intégrative de la Cellule, UMR 9198, Bâtiment 430, Université Paris-Sud, 91405 Orsay CEDEX, France, [d]Université Grenoble Alpes, Commissariat à l'Energie Atomique (CEA), Centre National de la Recherche Scientifique (CNRS), Institut de Biologie Structurale (IBS), and Institut Laue-Langevin, 71 Avenue des Martyrs, 38000 Grenoble, France, [e]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA, [f]Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA, [g]University of Technology Sydney, iThree Institute, 15 Broadway, Ultimo, NSW 2007, Australia, [h]Australian Synchrotron, 800 Blackburn Road, Clayton, VIC 3168, Australia, [i]Synchrotron SOLEIL, L'Orme des Merisiers, Saint-Aubin BP48, 91192 Gif-sur-Yvette CEDEX, France, [j]School of Applied and Engineering Physics, Cornell University, Ithaca, NY 14853-2501, USA, [k]Department of Bioengineering and Therapeutic Sciences, Department of Pharmaceutical Chemistry, and California Institute for Quantitative Biosciences (QB3), University of California San Francisco, San Francisco, California, USA, [l]School of Computer Science and Engineering, Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel, [m]Biozentrum, University of Basel and SIB Swiss Institute of Bioinformatics, Basel, Switzerland, [n]European Molecular Biology Laboratory (EMBL) Hamburg, c/o DESY, Notkestrasse 85, 22607 Hamburg, Germany, [o]Research Reactor Institute, Kyoto University, Kumatori, Sennan-gun, Osaka 590-0494, Japan, [p]Basic Science Research Division, Molecular and Cellular Oncology, MD Anderson Cancer Center, University of Texas, Houston, Texas, USA, and [q]Department of Chemistry and Chemical Biology, Rutgers University, New Brunswick, NJ 07102, USA. *Correspondence e-mail: jill.trewhella@sydney.edu.au

In 2012, preliminary guidelines were published addressing sample quality, data acquisition and reduction, presentation of scattering data and validation, and modelling for biomolecular small-angle scattering (SAS) experiments. Biomolecular SAS has since continued to grow and authors have increasingly adopted the preliminary guidelines. In parallel, integrative/hybrid determination of biomolecular structures is a rapidly growing field that is expanding the scope of structural biology. For SAS to contribute maximally to this field, it is essential to ensure open access to the information required for evaluation of the quality of SAS samples and data, as well as the validity of SAS-based structural models. To this end, the preliminary guidelines for data presentation in a publication are reviewed and updated, and the deposition of data and associated models in a public archive is recommended. These guidelines and recommendations have been prepared in consultation with the members of the International Union of Crystallography (IUCr) Small-Angle Scattering and Journals Commissions, the Worldwide Protein Data Bank (wwPDB) Small-Angle Scattering Validation Task Force and additional experts in the field.

1. Introduction

The objective of publishing the preliminary guidelines for biomolecular small-angle scattering (SAS) experiments (Jacques, Guss, Svergun *et al.*, 2012; Jacques, Guss & Trewhella, 2012) was to provide a reporting framework so that 'readers can independently assess the quality of the data and the basis for any interpretations presented'. The focus was on

OPEN ACCESS

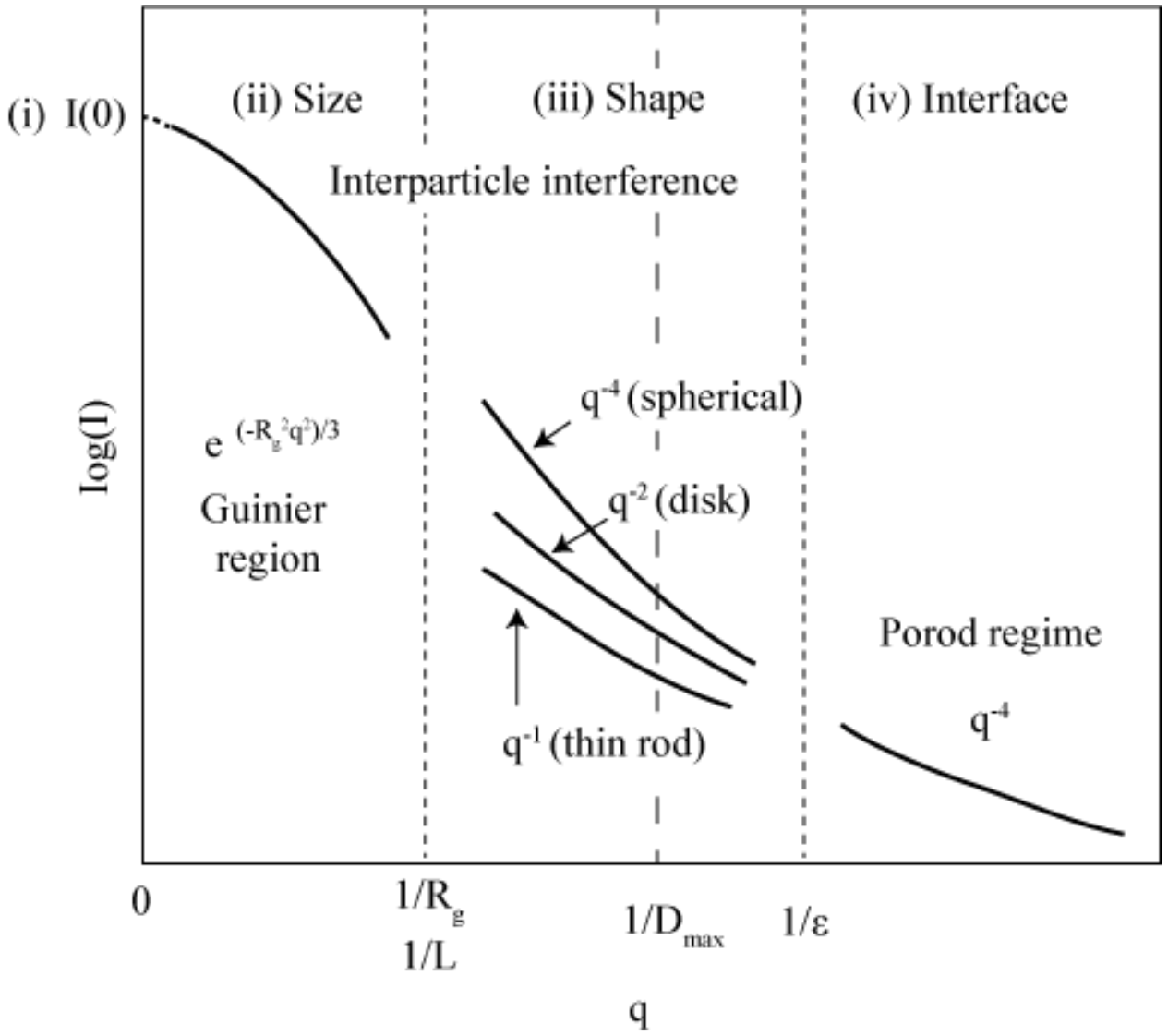
Good practical guidelines are already available.

- Publication guidelines are available and provide good advice over best practices – e.g. Trewhella et. al Acta Cryst. D72, 710-728, 2017).

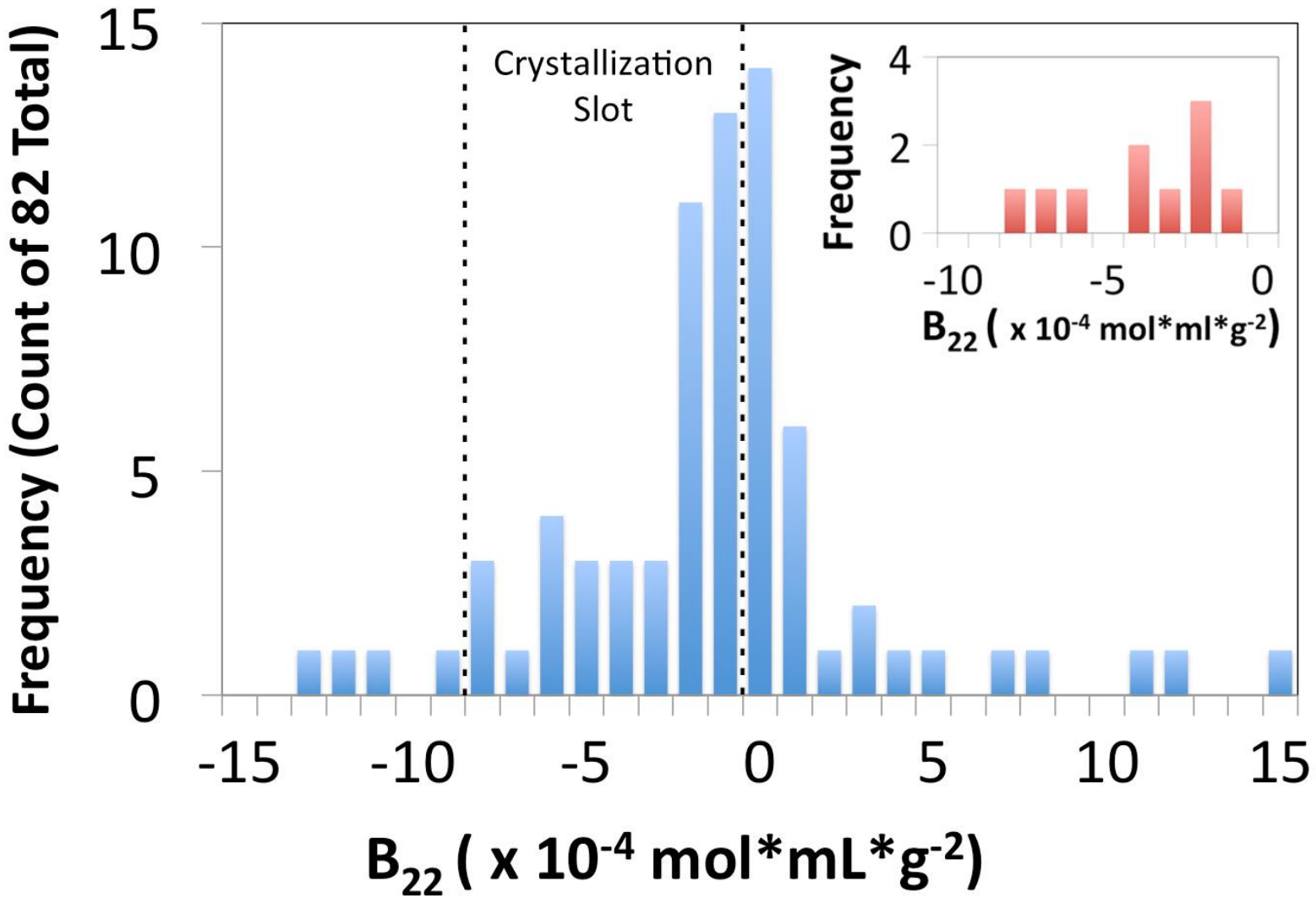
- It is highly recommended to follow these guidelines and while not all will be applicable, following those that do apply ensures the highest quality data and that conclusions are from the sample and not errors in the process.

The Shape of the Scattering Curve is important

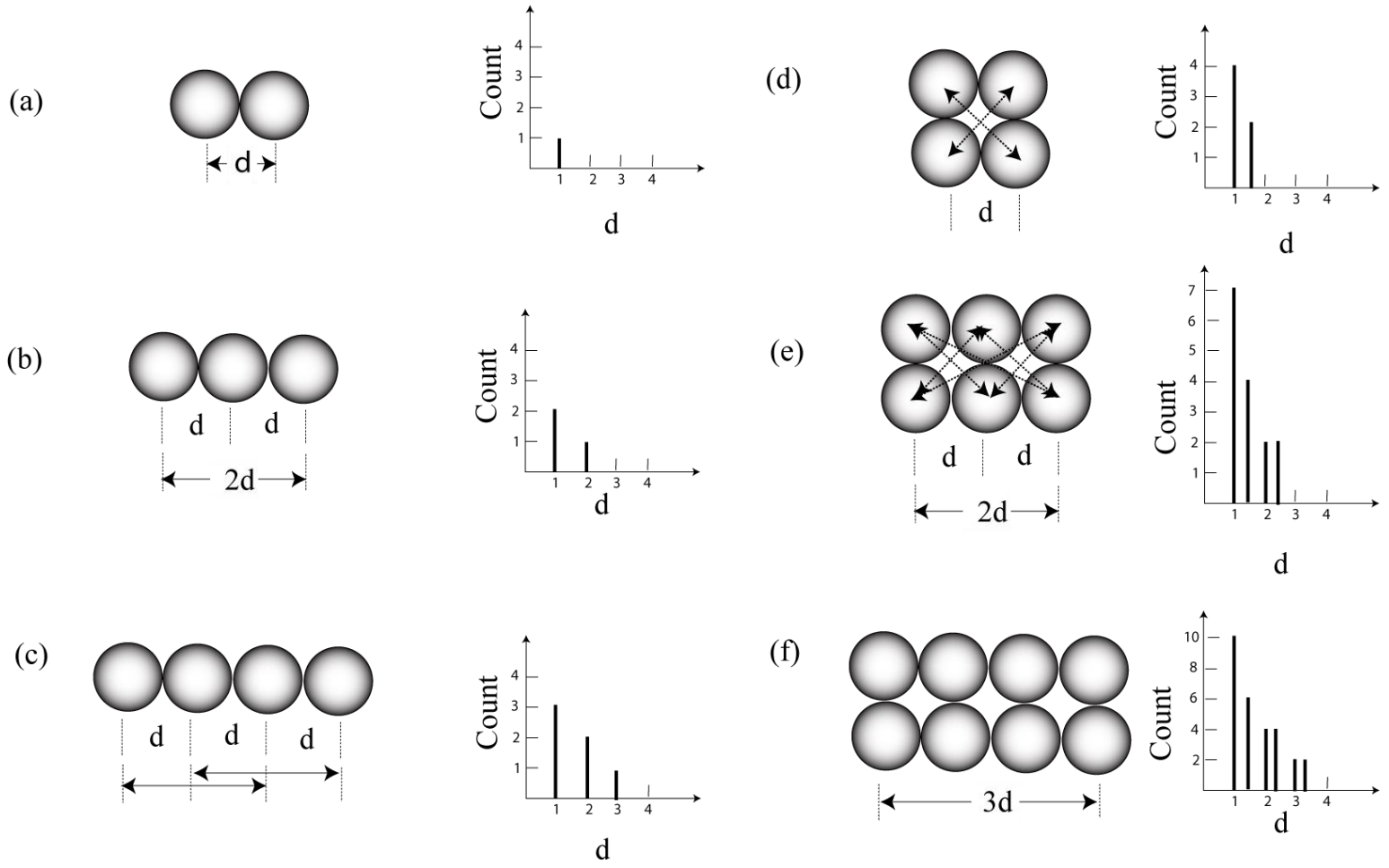
but not the absolute intensity



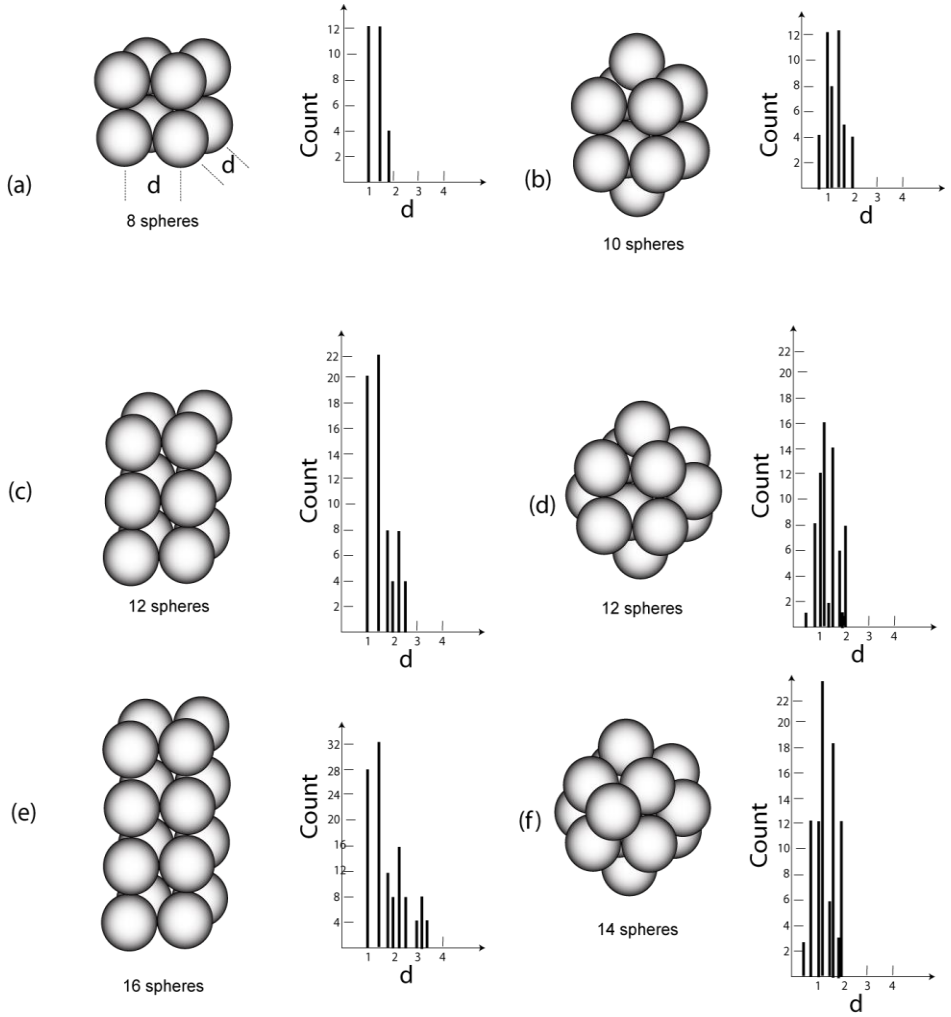
The shape provides the information that comes from SAXS


Frequency (Count of 82 Total)
15
10
5
0
Crystallization Slot
-15
-10
-5
0
5
10
15
B22 (x 10-4 mol*mL*g-2)
Frequency
4
2
0
-10
-5
0
B22 (x 10-4 mol*ml*g-2)

Two-dimensional examples



Extension to Three dimensions

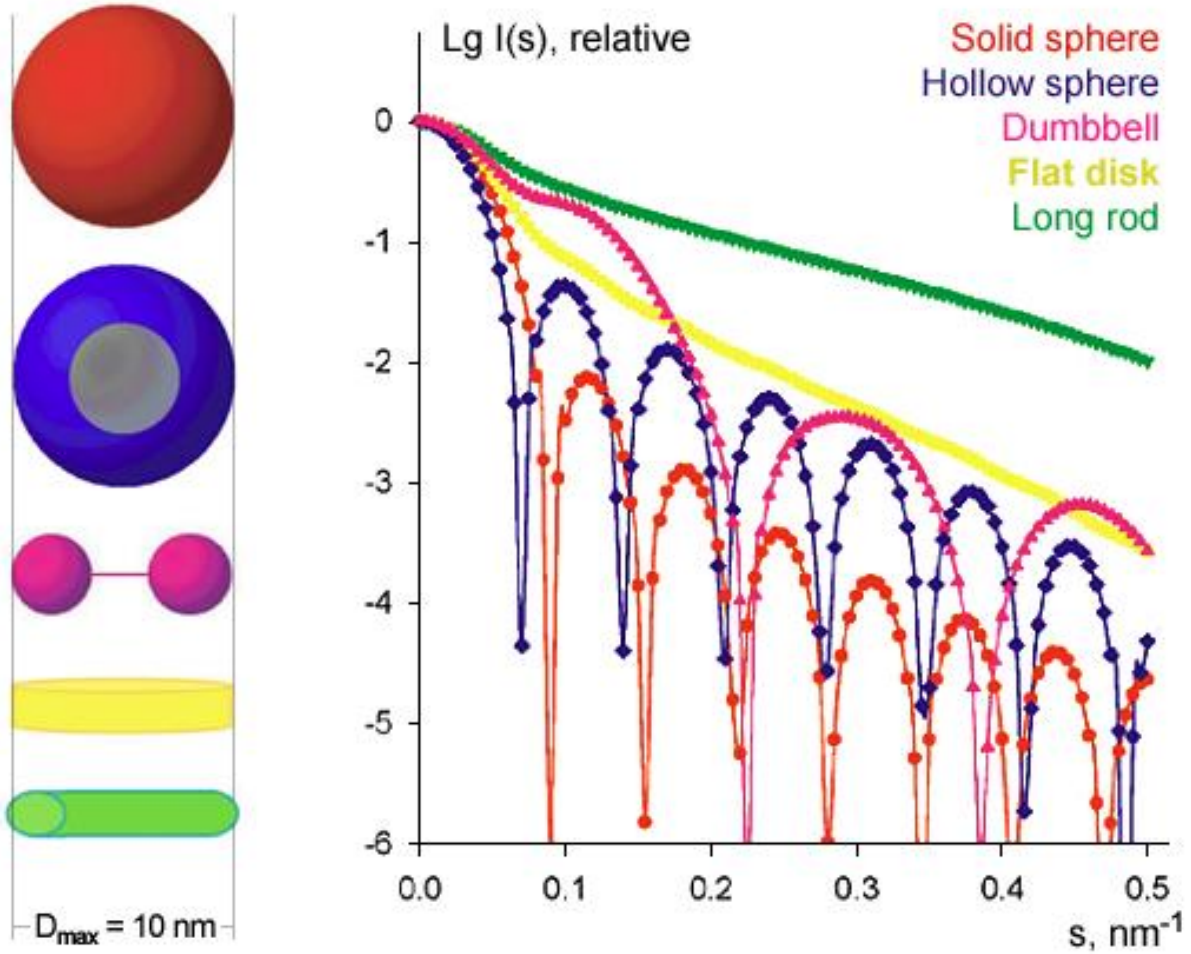


Fourier space yields frequency of interatomic scattering vectors as a function of the length of the vector

P(r) (**Pair distribution function**) plot is simply the histogram of interatomic scattering

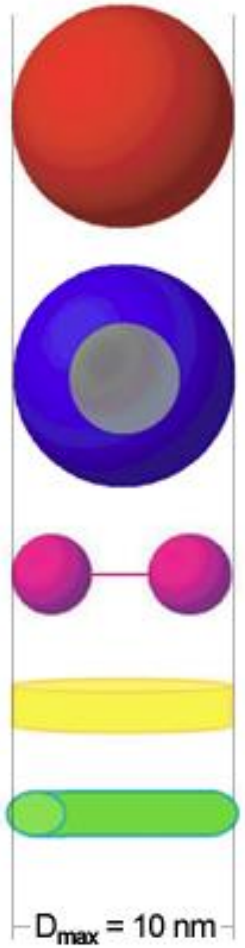
Larger compact molecules have a high distribution at lower angle (consider detector distance etc.)

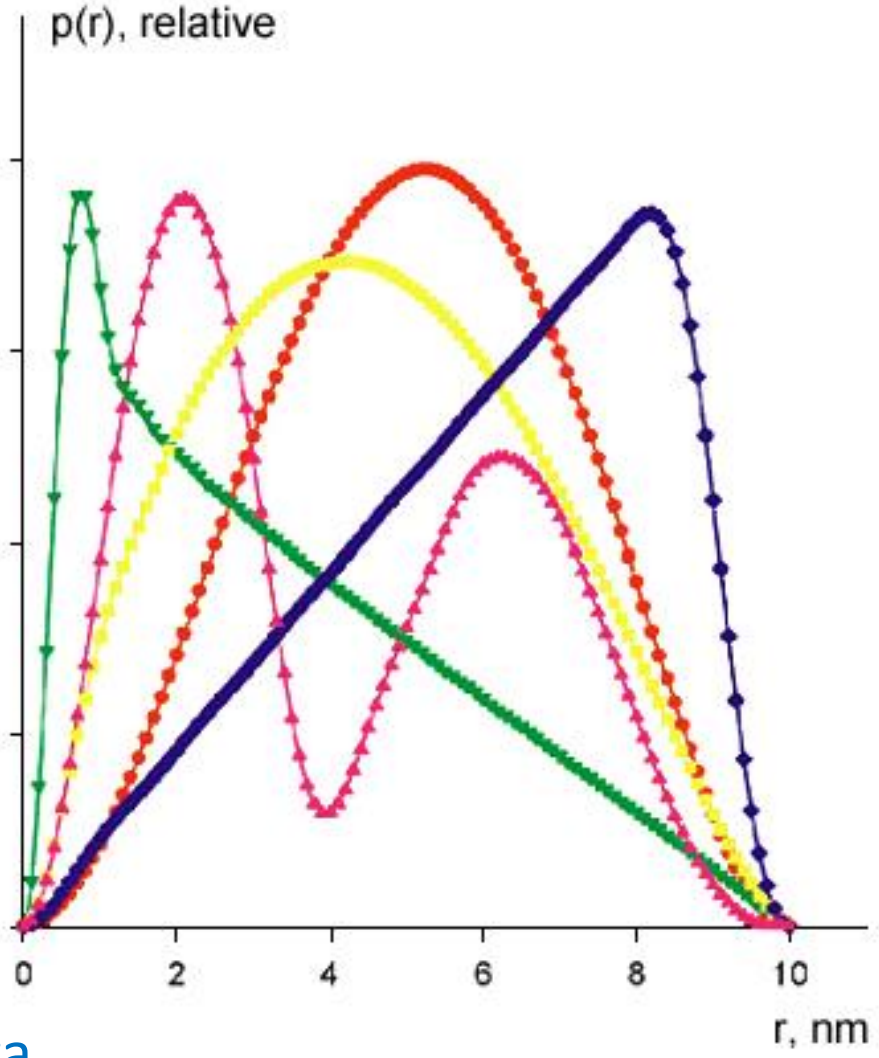
Data



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)

Basics of SAXS/SANS

- The pair distribution function describes the interatomic scattering between all atoms making up the sample.
- The pair distribution function is the Fourier transform of the scattering.
- The raw data comes from a difference technique.
- Shape but not intensity provides the data.

SAXS consists of intensity due to the from factor and interparticle contributions

- Equation for scattering intensity:

$$I(q) = F(q) * S(q)$$

Experimental Intensity

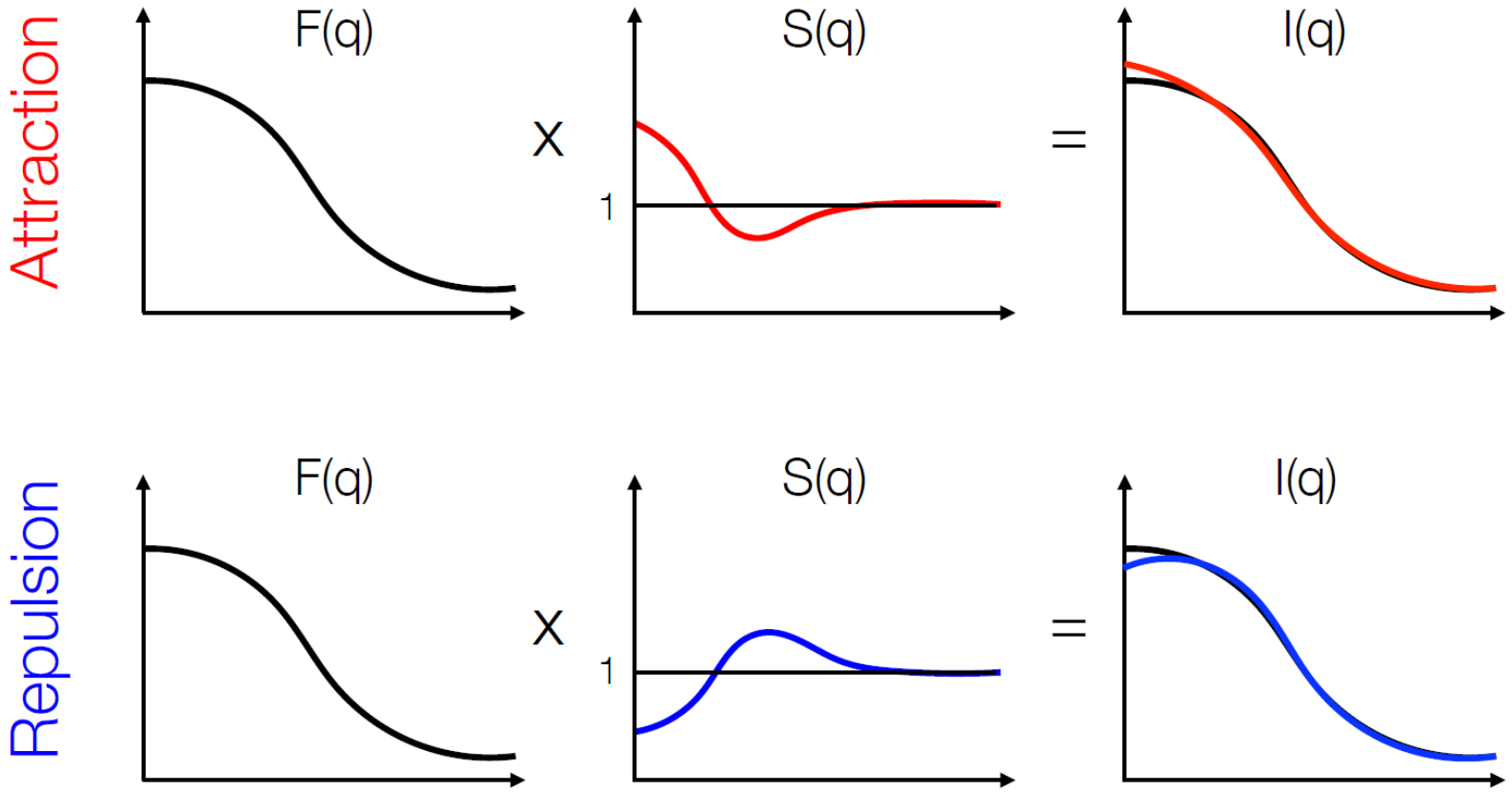
Form factor of particle

Structure factor of solution

- Form factor describes *intra*particle interactions, i.e. size and shape
- Structure factor describes *inter*particle interactions, i.e. repulsion/attraction
- Ideally a monodisperse solution for SAXS should have no interparticle interactions, i.e. S(q) = 1

Interparticle Interactions

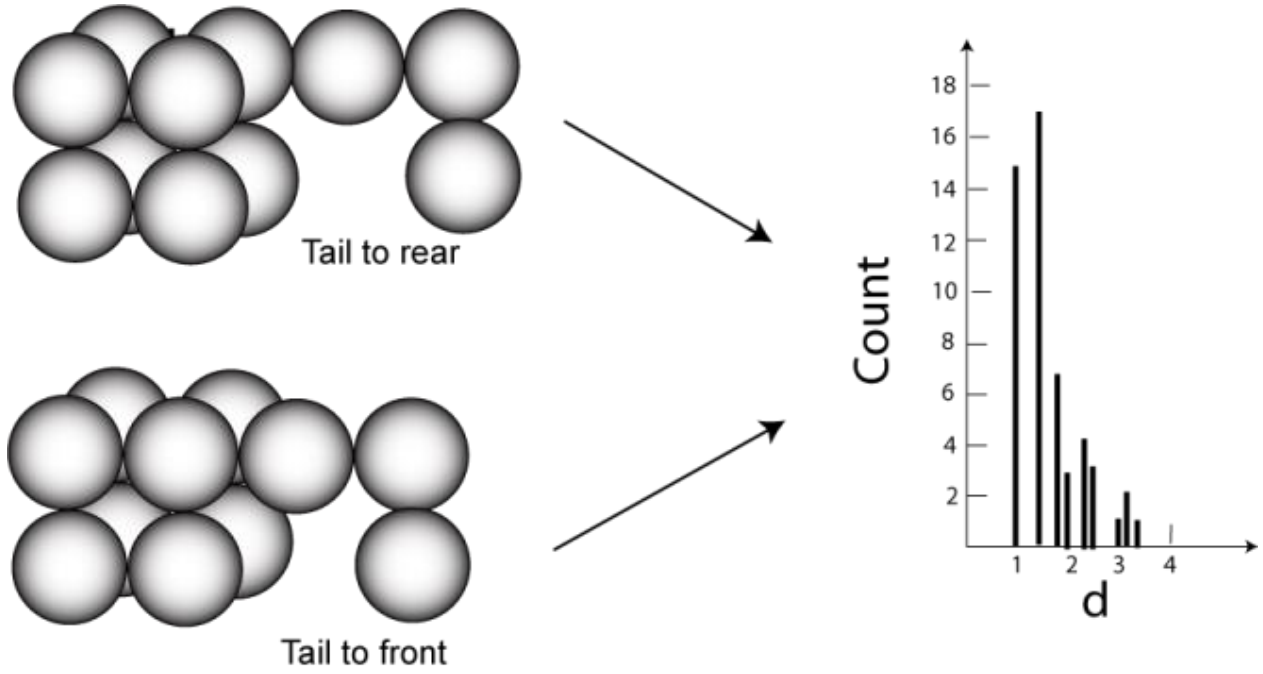
S(q) ≠ 1 affects low q data most



Similar to data from light scattering and can be used in the same manner

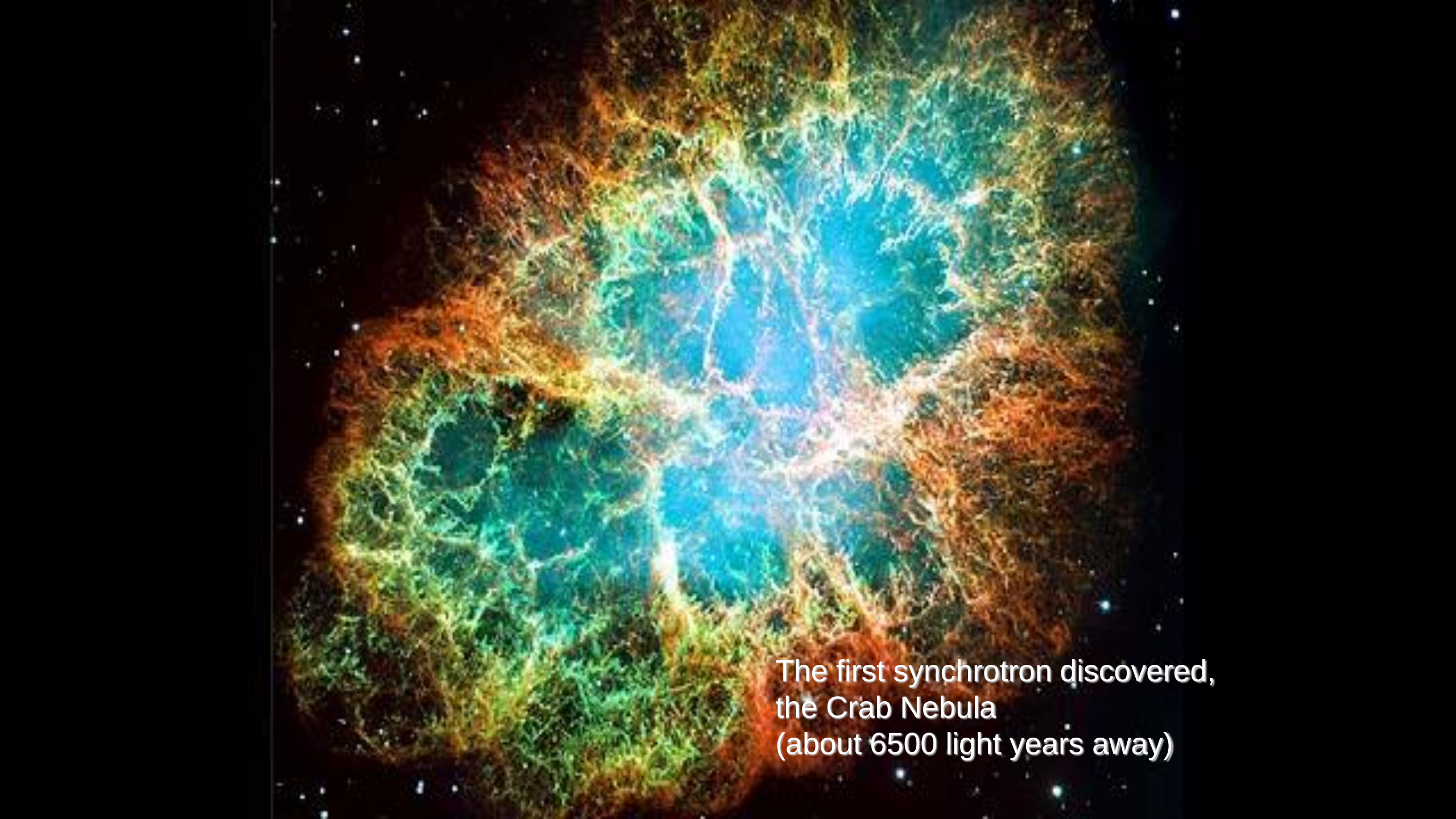
What can possibly go wrong?

SAXS data has to be interpreted

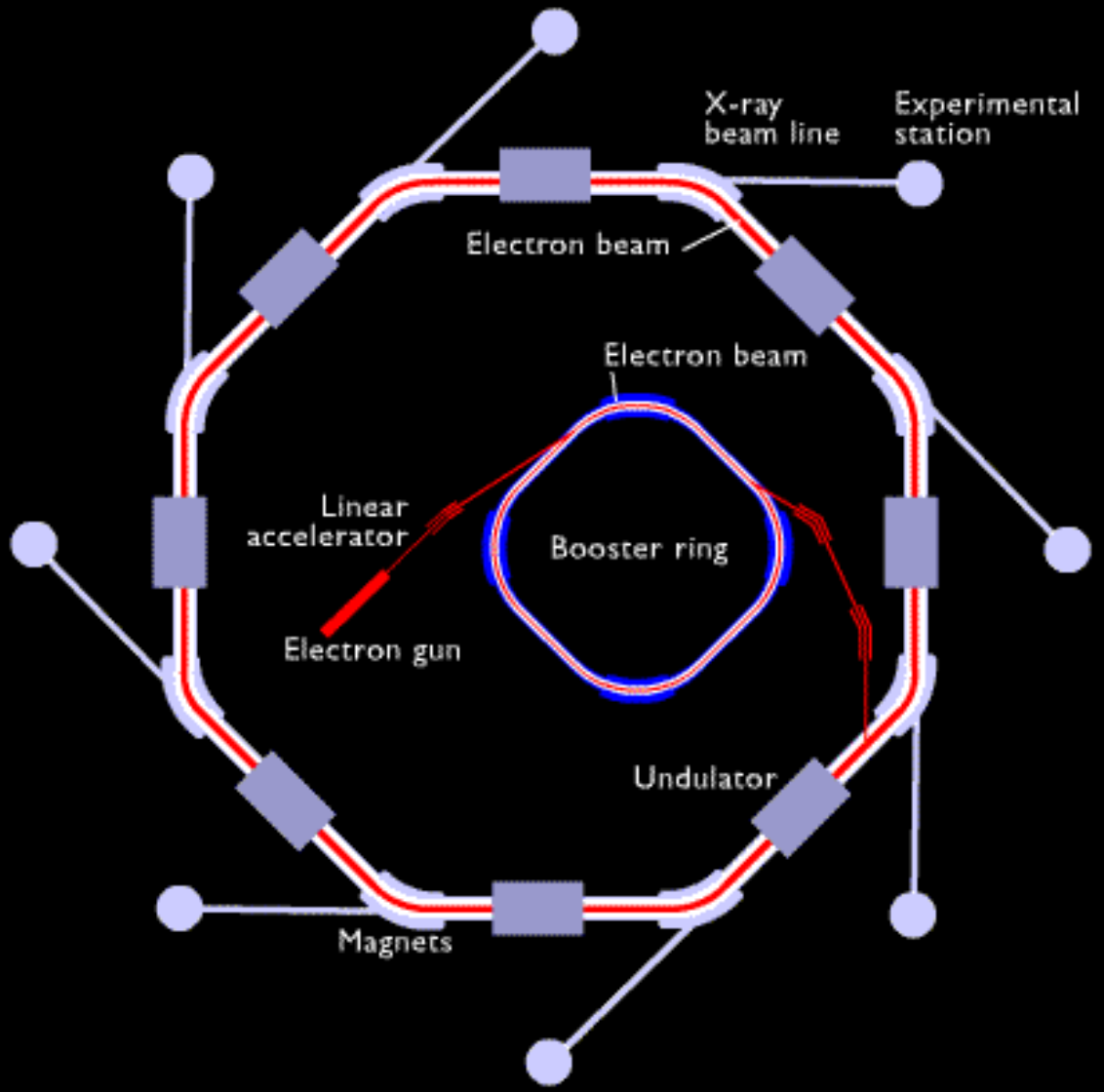


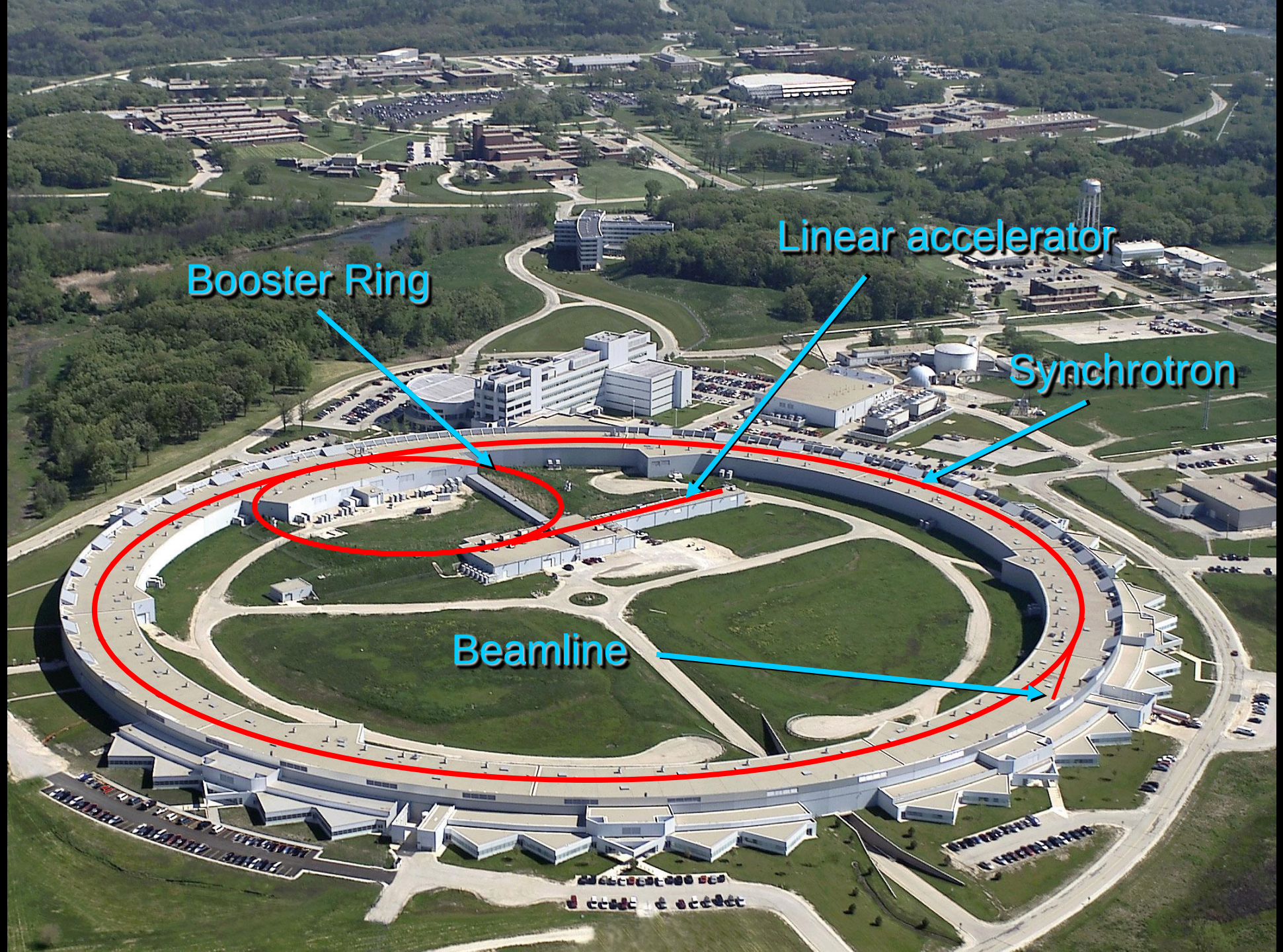
Sometimes a unique reconstruction is not available.

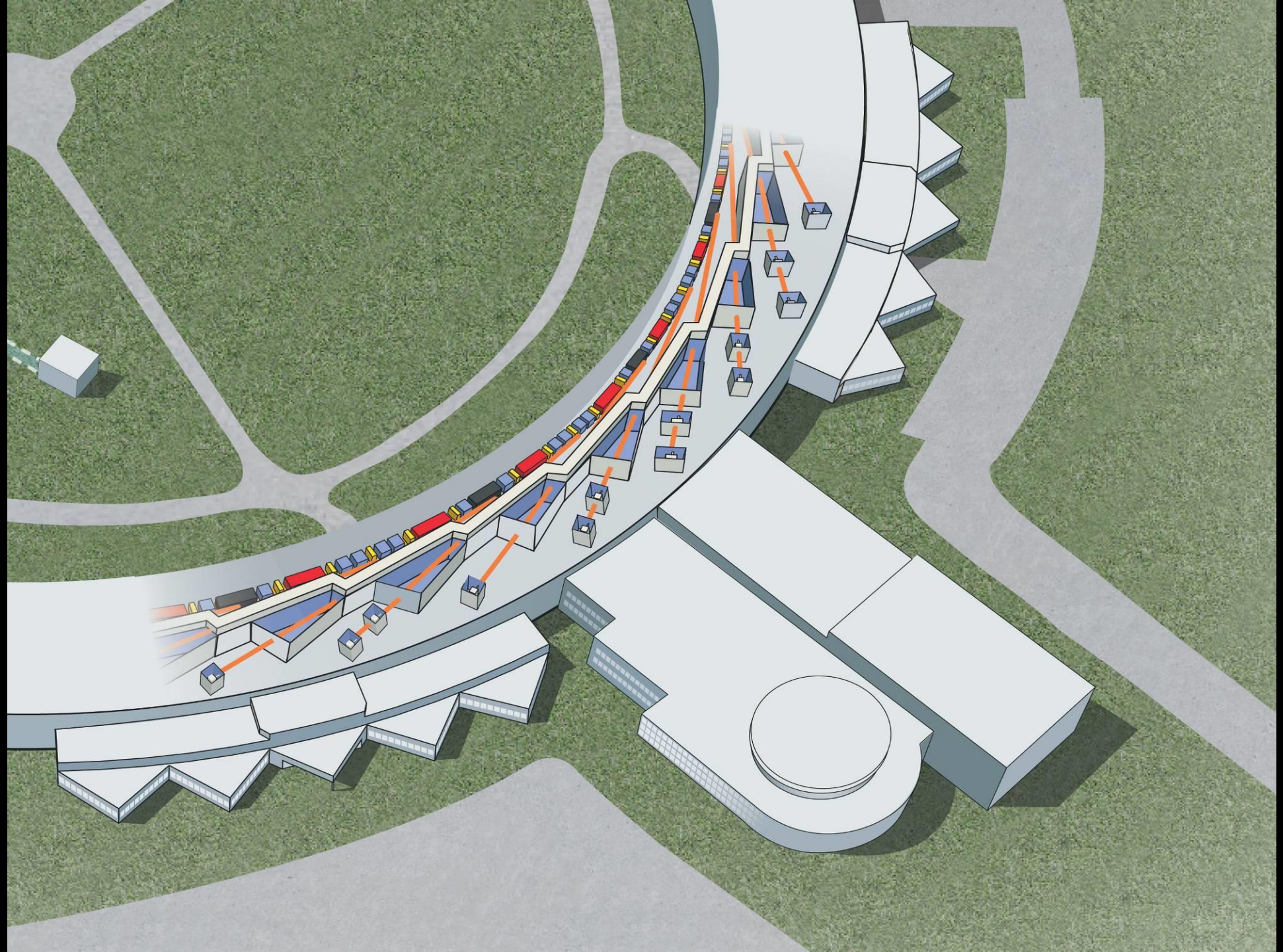
The experiment

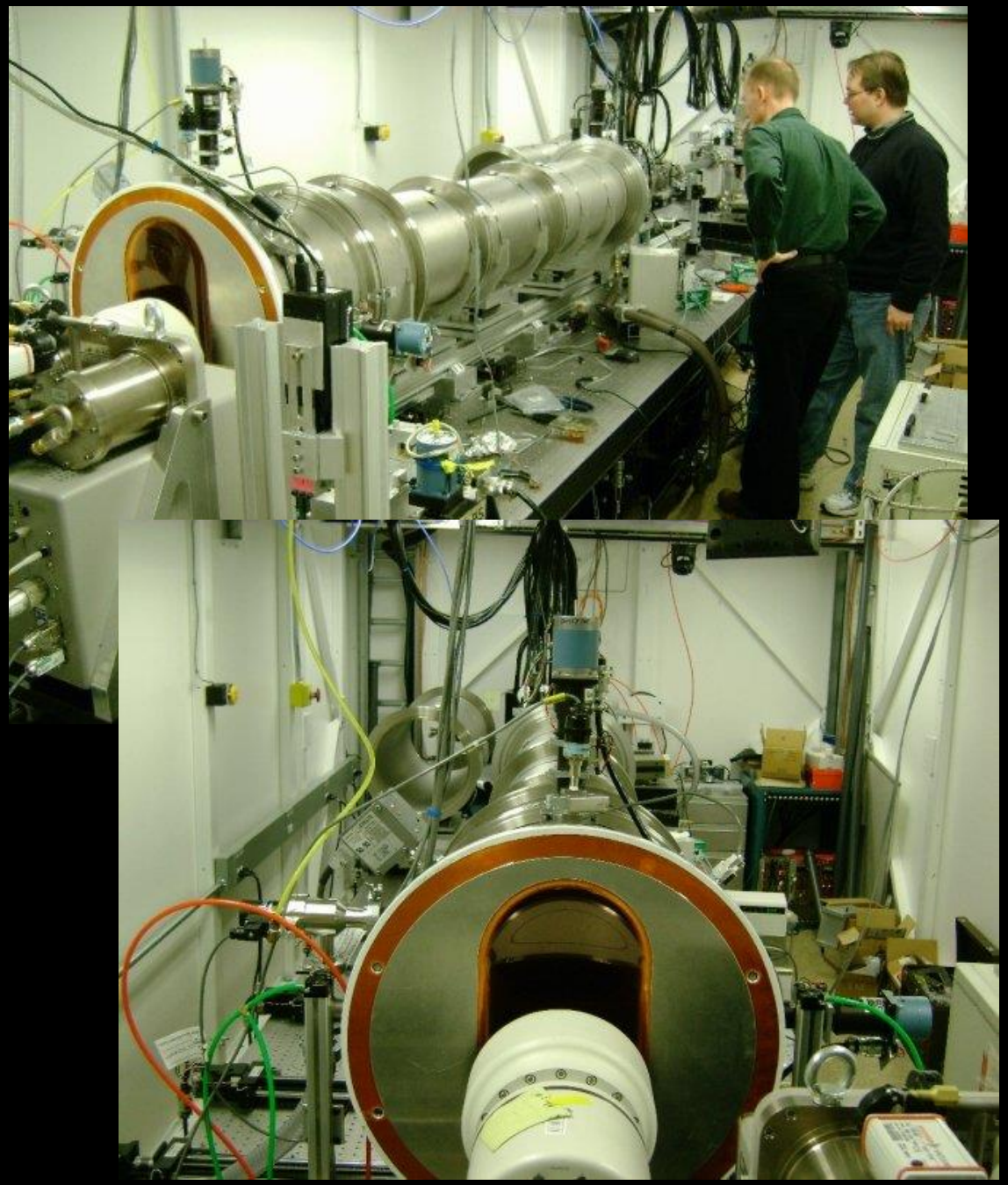


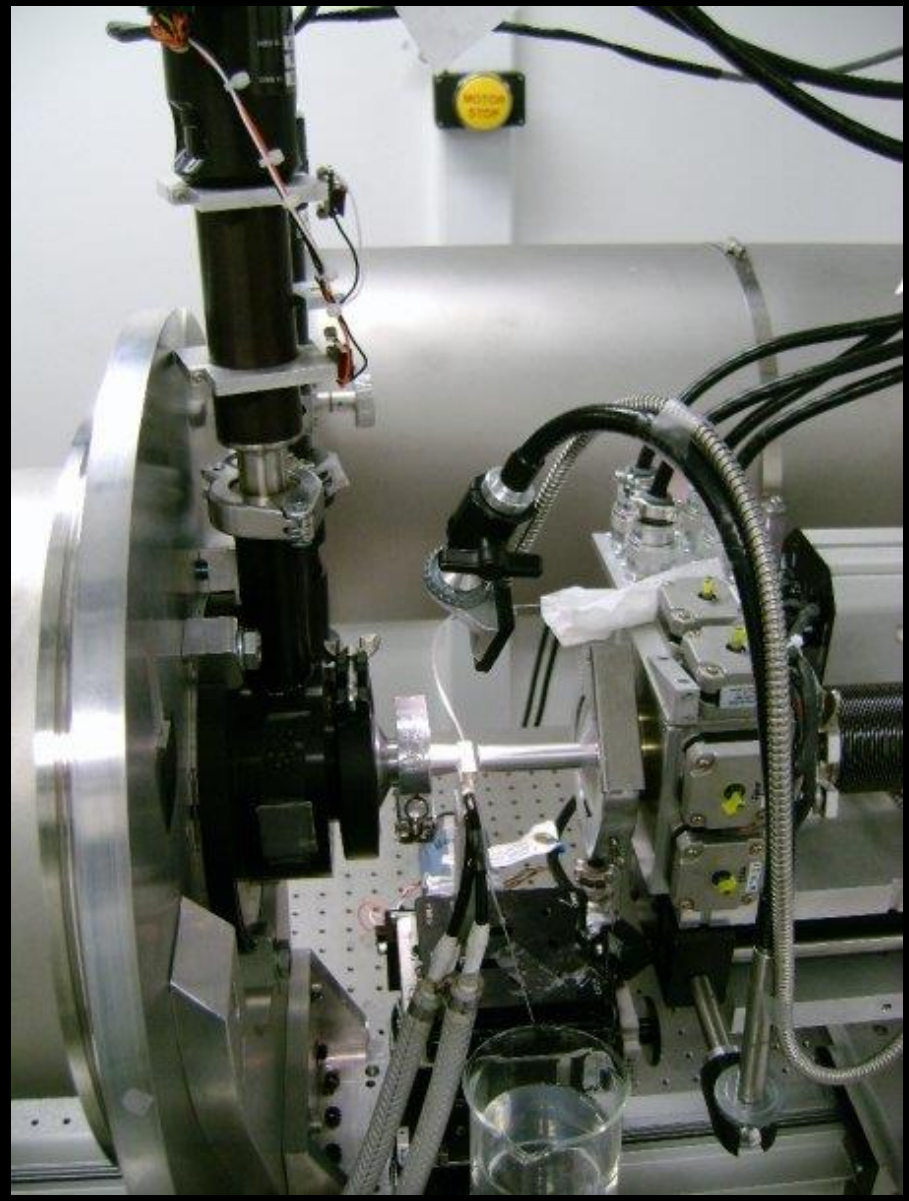
The first synchrotron discovered,
the Crab Nebula
(about 6500 light years away)


X-ray
beam line
Experimental
station
Electron beam
Electron beam
Linear
accelerator
Booster ring
Electron gun
Undulator
Magnets


Booster Ring
Linear accelerator
Synchrotron
Beamline







Stanford Synchrotron Lightsource
Beamline 4-2

High throughput protocol

Up to 12 different PCR strips.

3-7 different concentrations per sample.

For high-throughput studies, 2 samples per strip, 24 samples in total

Start with buffer then lowest concentration first. End with buffer

8 exposures, 1-2s each dependent on sample molecular weight, buffer and concentration.

Oscillate sample to minimize radiation damage

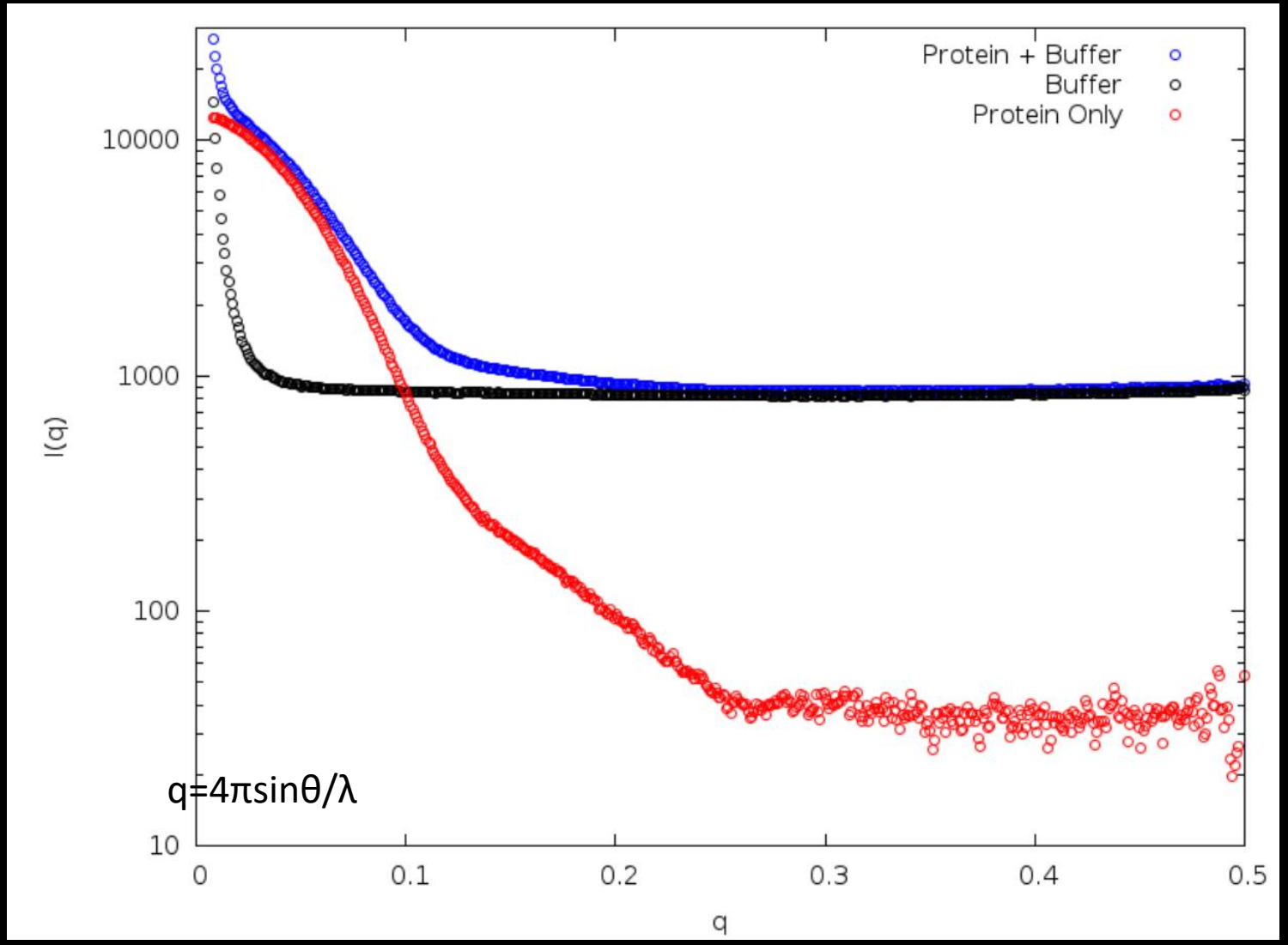
Repeat the buffer.

Load next sample

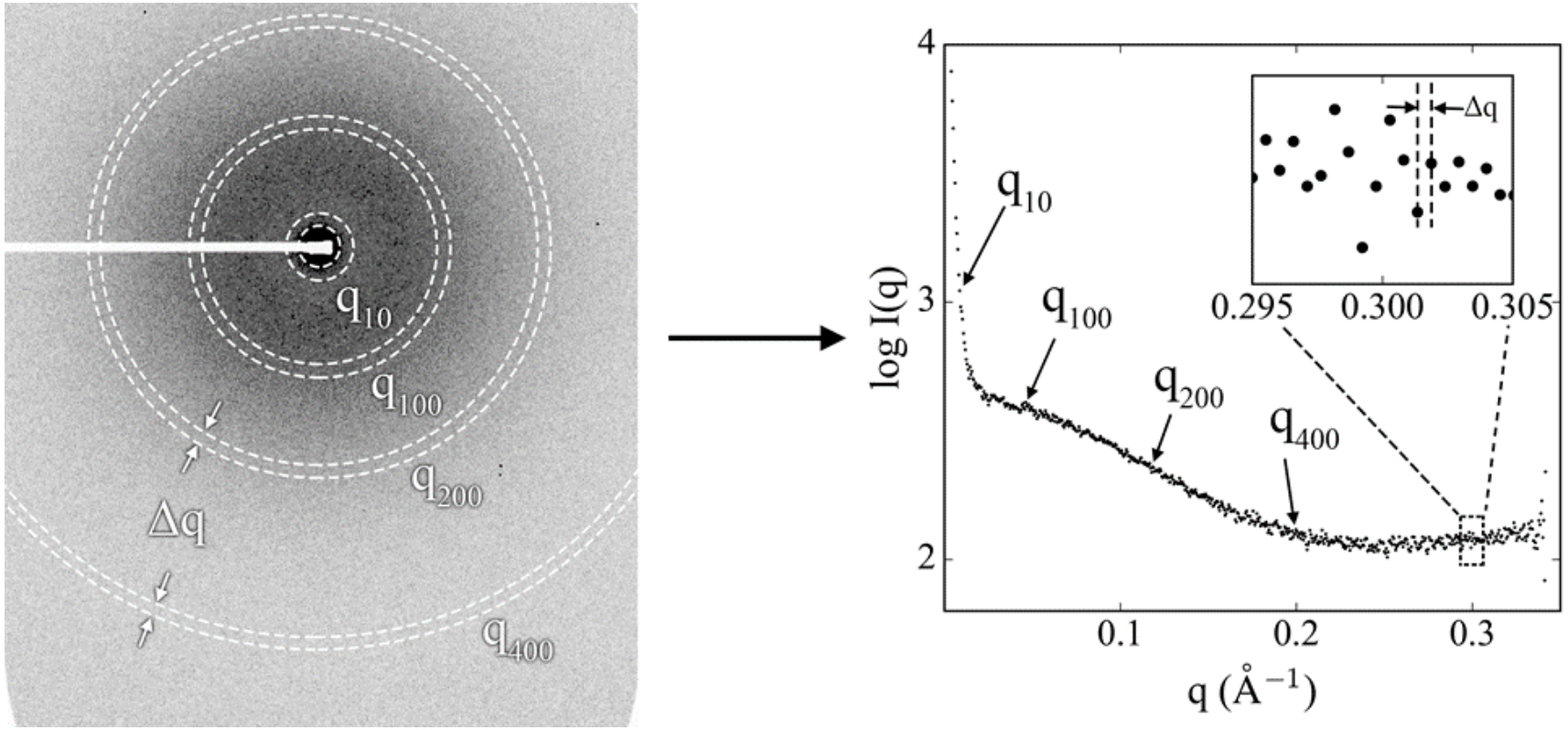
Time per concentration series – approximately 10 to 15 minutes. In high-throughput mode 24 samples in 3 to 4 hours.

Enables two important things – eat and sleep!

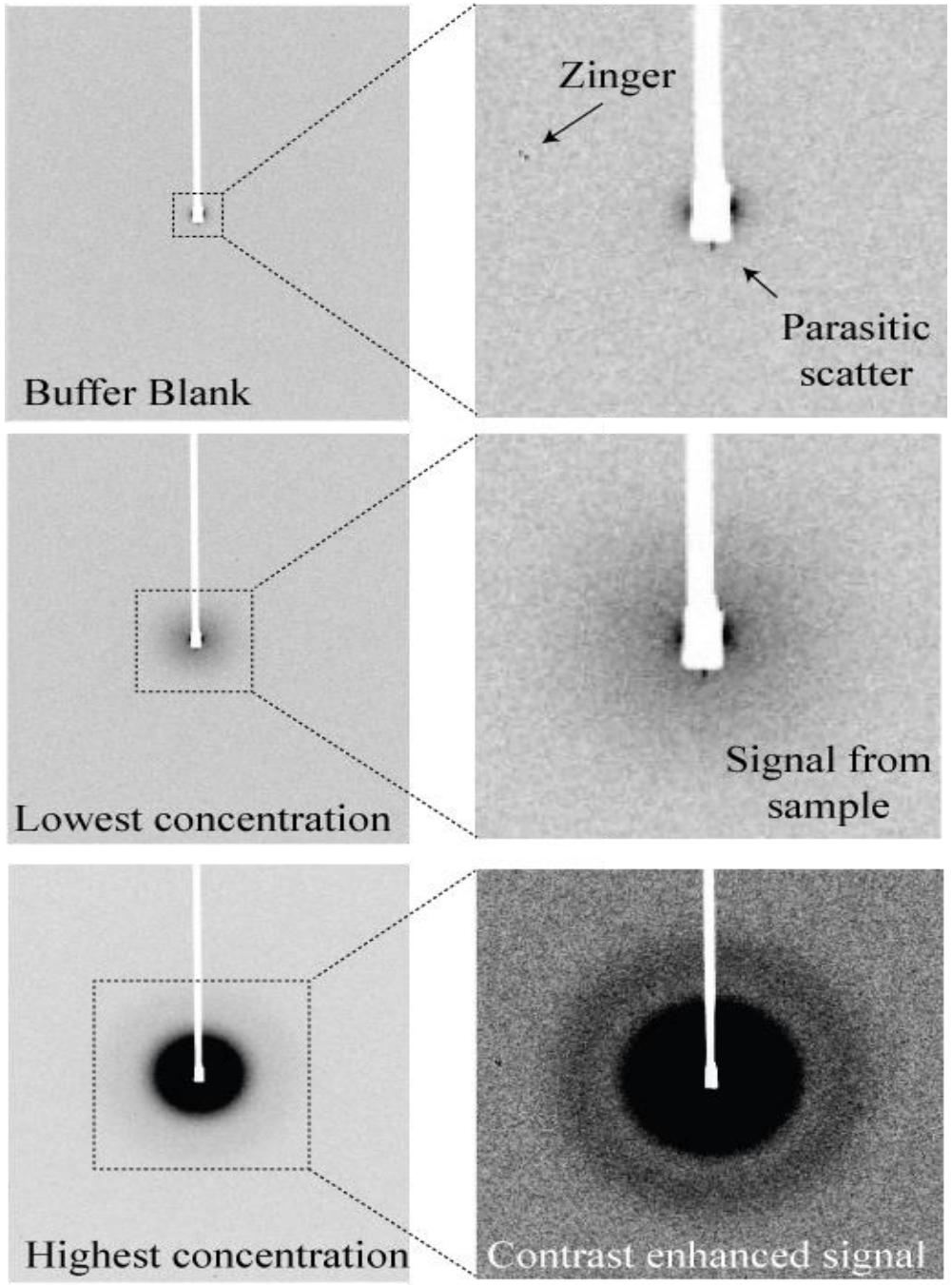
SAXS data is the sample data with the buffer signal subtracted



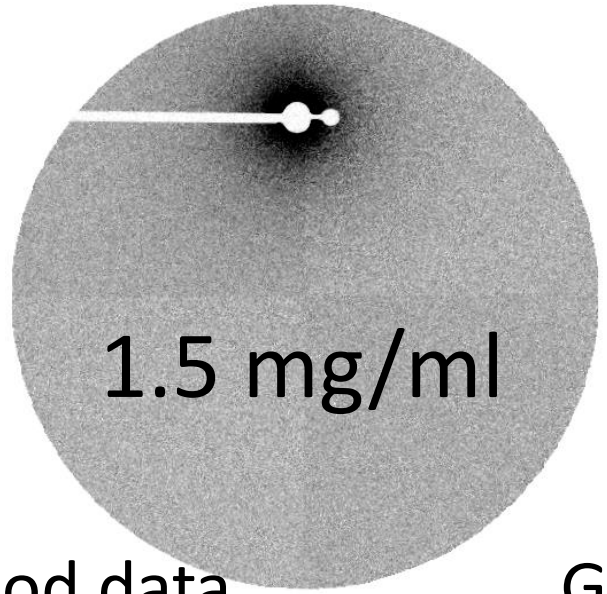
Radial integration with significant oversampling



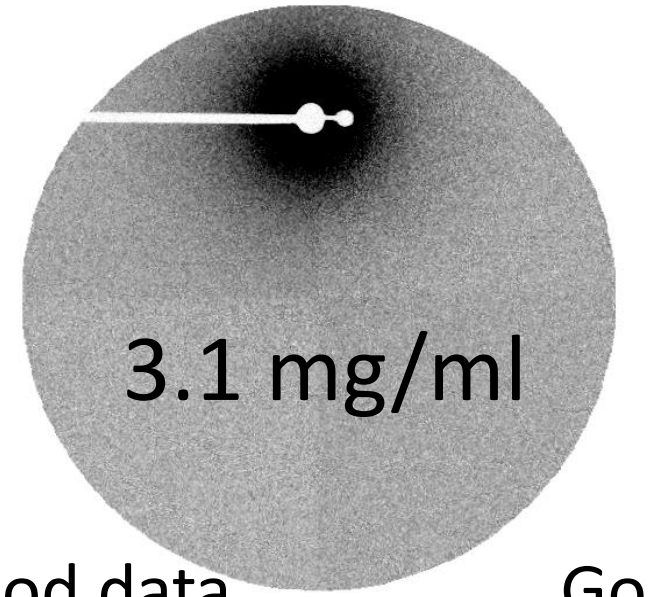
- SAXS/SANS is a continuous sampling that can be considered analogous to crystallography with one dimensional indicies (hkl's) that are sampled discretely, but also contribute to the overall scattering.
- The scattering data is extremely precise because each point (Shannon channel) is greatly over sampled.



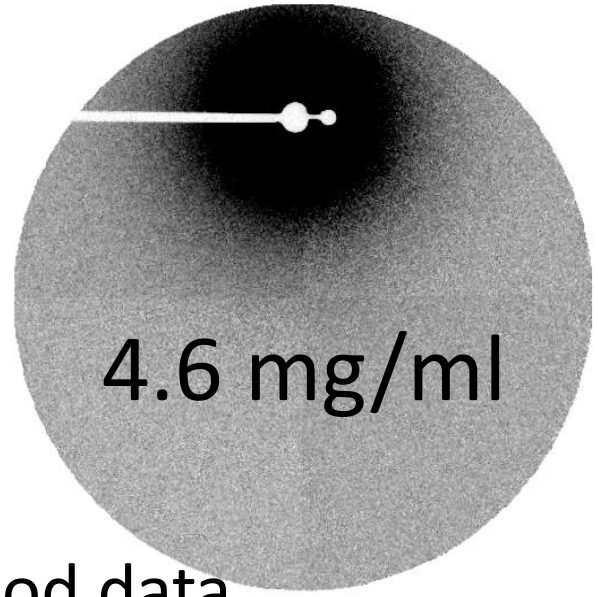
Small angle scattering data near the beamstop



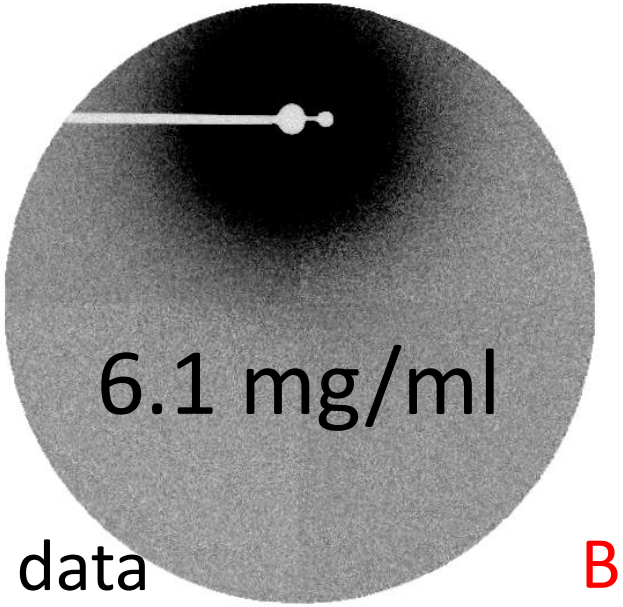
Good data



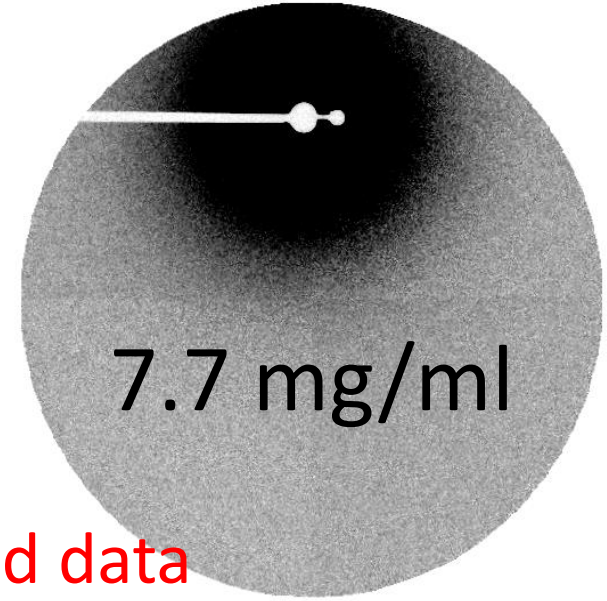
Good data



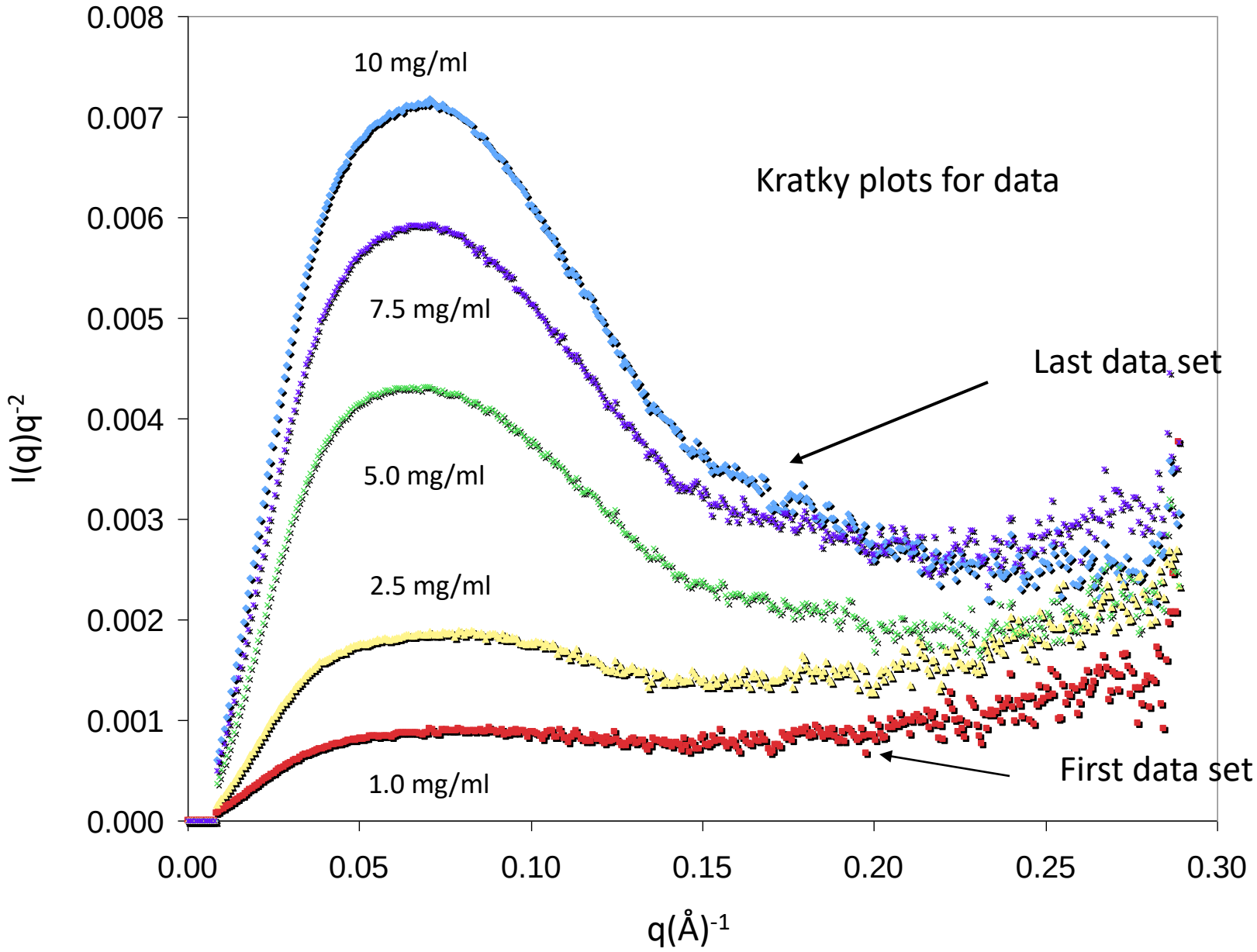
Good data



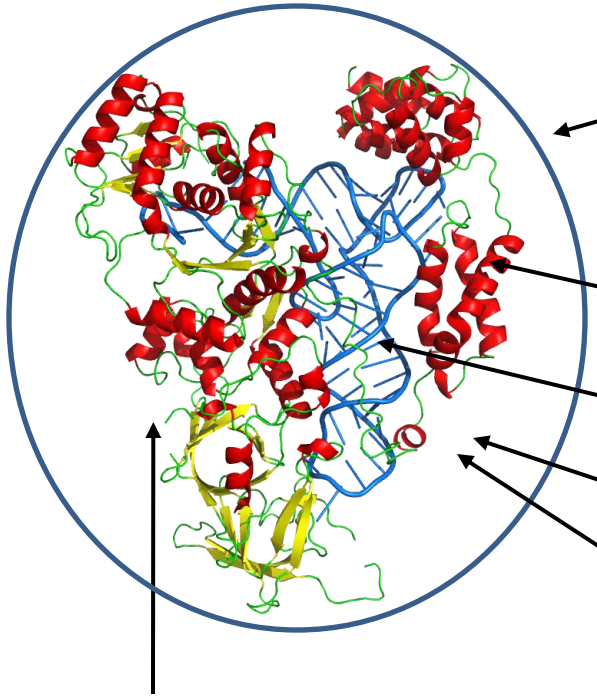
Good data



Bad data


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

Eukaryotic Gln tRNA synthetase



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

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

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

Practicing safe SAXS

Safe SAXS

(Be prepared)

Sample preparation

- For SAXS, a rule of thumb is that the concentration in mg/ml multiplied by the molecular weight in kDa should be greater than 100.
- At least three concentrations are used to detect potential concentration-dependent aggregation or oligomerization processes - more concentrations are preferable.
- SANS requires almost an order of magnitude more sample, and if X-ray scattering has been used to detect concentration-dependent effects beforehand, it does not usually require a concentration series. Several different contrast points are usually recorded by appropriately varying the ratio of D_2O to H_2O.

Sample preparation

- The sample should be well characterized and free from impurities.
  - For example Jefferies et al. (2016) note that a 15 kDa protein sample purified to 98% that contains a 2% 100 kDa impurity will generate a forward scattering intensity that is almost twice that expected from the pure 15 kDa sample.
- The sample should be verified that it is the sample wanted!
- It is good practice to be sure of the sample produced, e.g. by a suitable assay, mass spectrometry, or other technique.

Sample preparation

- SDS-PAGE acts as a check on the approximate molecular, Native PAGE can be used to determine if a sample is homogeneous or may consist of a range of structural species.
- UV spectrophotometry of the elution volume can reveal the presence of oligomeric species or aggregation.
- Oligomerization or aggregation can be a time dependent process .
- Concentration dependent aggregation does not preclude successful SAXS data collection; it just limits the concentration range that should be studied.
- Consider changes over time.
- Static Light Scattering (SLS) can be used as a check on the molecular weight of the sample in solution, and how this may change with concentration or buffer choice.

SAXS is a difference technique - Buffer considerations

- The SAXS signal is proportional to the sample concentration and the difference in density between the sample and the buffer blank.
- The buffer should be designed to minimize its own scattering component, for example, by avoiding high atomic number elements.
- The buffer should also stabilize the sample.
- The first rule of a buffer for SAXS is that the sample must be stable in the buffer.
- Salts can be used to decrease the long-range electrostatic repulsion between the macromolecules at the measurement conditions.
- Adjusting the pH of the buffer toward the protein’s isoelectric point can also decrease long-range electrostatic repulsion but can also lead to increased aggregation.
- Buffer choice is a critical component of the experiment.

Sample transport

- It is not recommended to ship samples at non-controlled temperature conditions or as a pre-frozen concentration series. In the former sample denaturation can occur and for the latter, aggregation can result.
- Many beamlines offer at least basic wet laboratory facilities - the sample can be freshly prepared either by thawing and diluting the flash frozen sample or by concentrating a dilute solution of the sample.
- Spin down samples down before data collection.
- If size exclusion chromatography is to be employed it is better to use a column that has already proved successful in sample preparation.
- Feedback during the data collection can indicate that the sample or buffer may need to be changed.
- Centrifugal filters for small volumes at appropriate molecular weights can be very useful.
- Sample lifetime is also a consideration. Some samples will degrade or aggregate over time so it is always preferable to use as fresh a sample preparation as possible.

Other tips

- Many experiments work well the second or third time simply because a problem seen the first time was not noticed or could not be addressed at the beamline.
- Standard samples can check on the performance of the instrument, to ensure that the first use provides good data, and to act as a calibration standard for the sample of interest.
- There are many standards recommended including chicken egg white lysozyme or bovine serum albumin.
- Any well-characterized, stable, globular protein that is not prone to oligimerization can be used.
- Access to buffer components and potentially additives that can reduce radiation damage or minimize aggregation is useful.

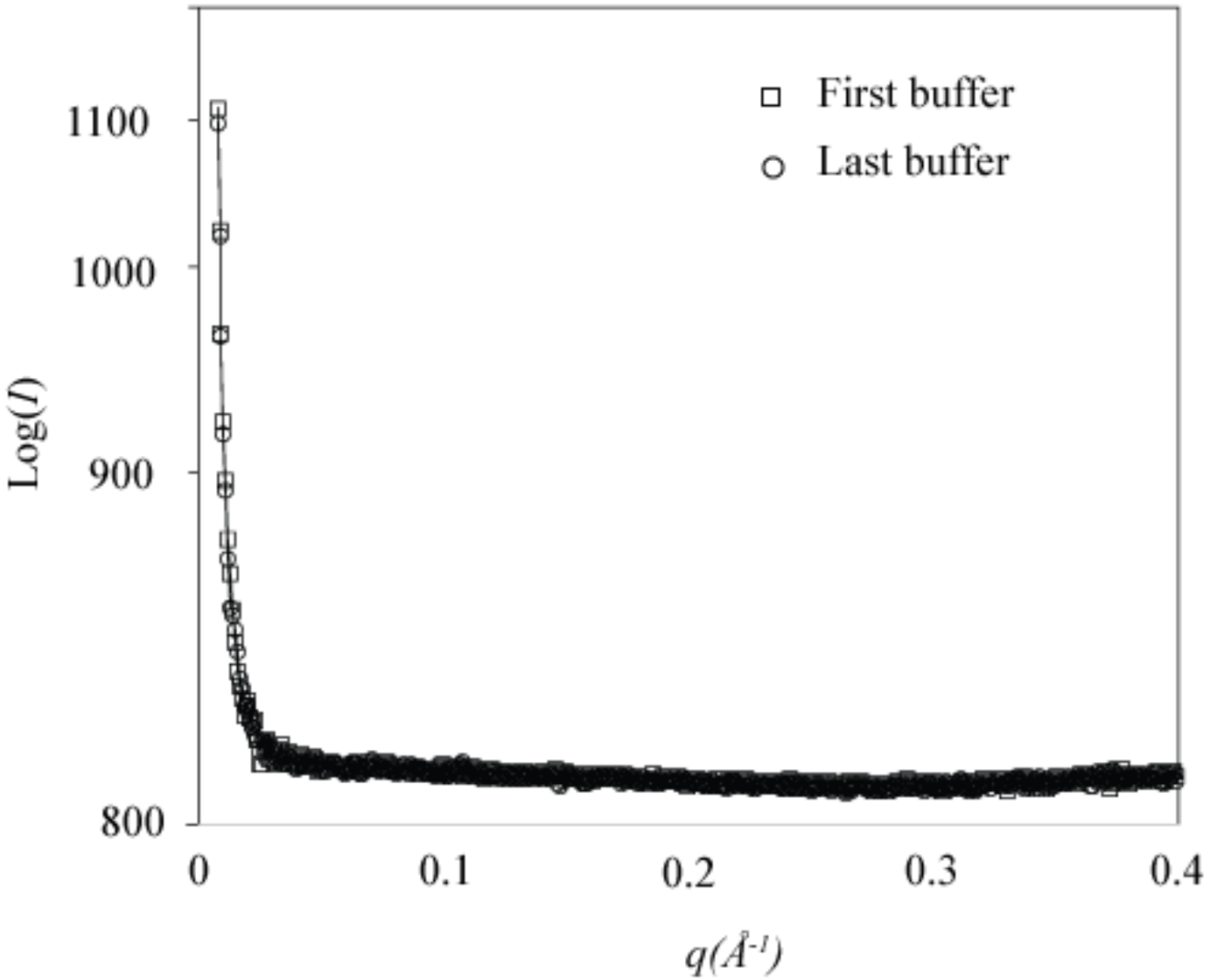
But the most importantly

Time taken to clean the sample cell is never wasted time

Safe SAXS

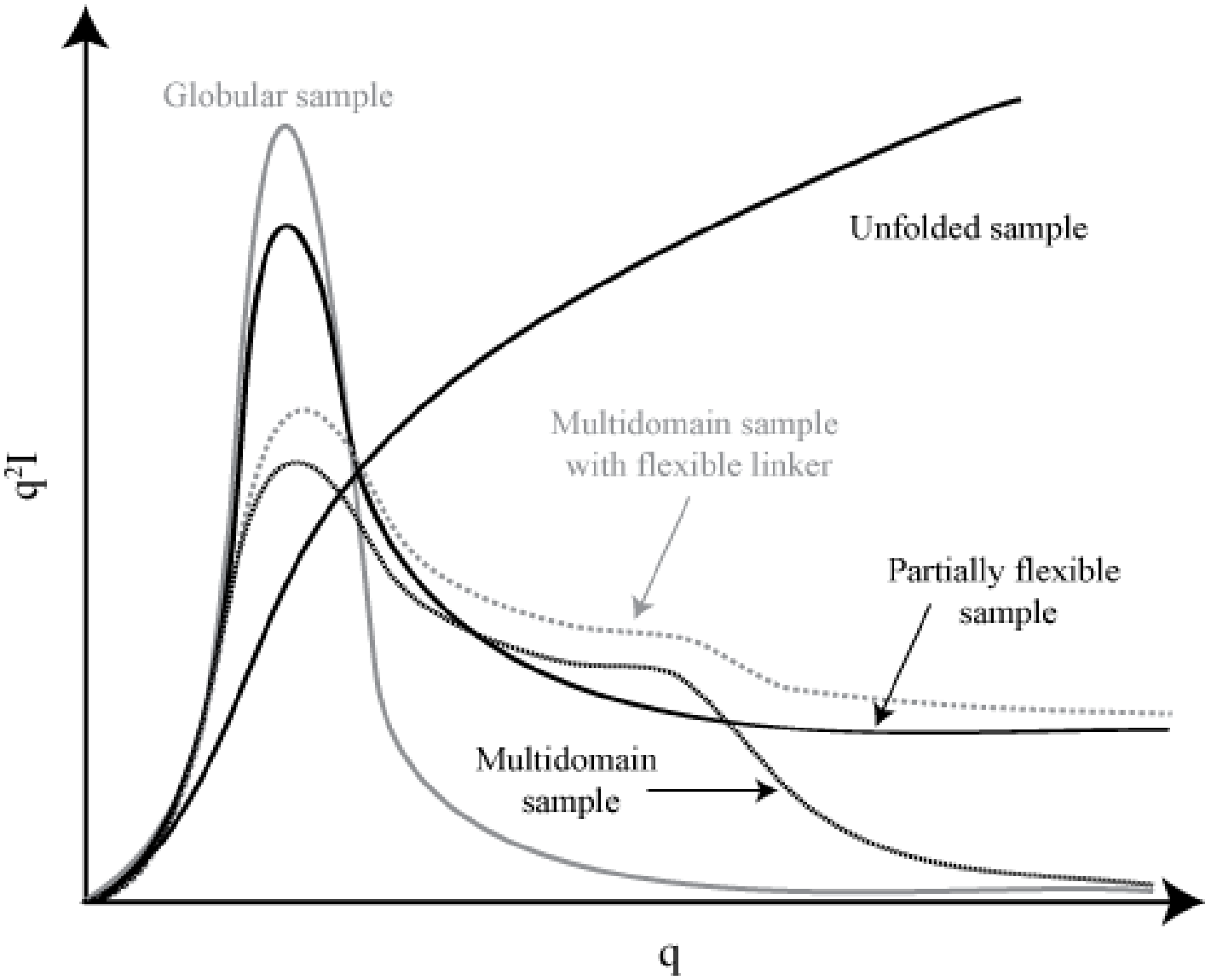
(Check your data)

Buffer analysis before and after sample collection can identify issues



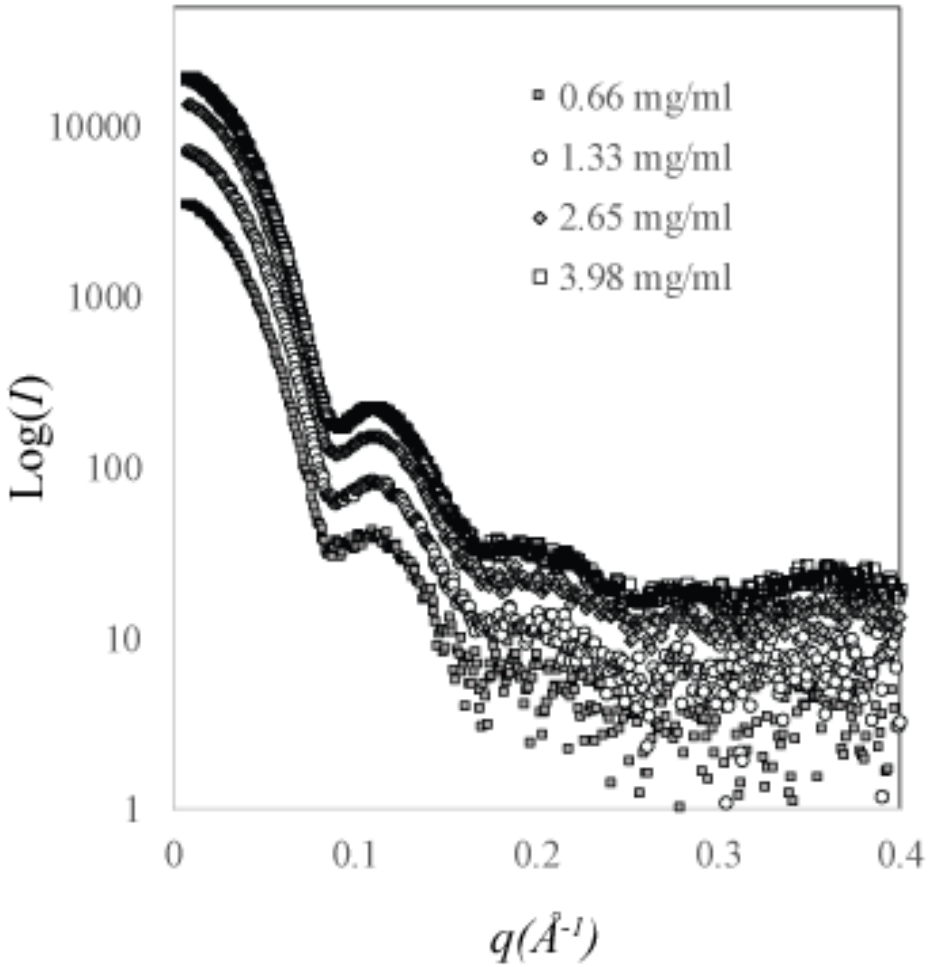
- Good practice is to record a buffer blank before and after sample collection.
- This ensures that any buildup of material on the sample cell wall (e.g. due to radiation damage) can be identified.
- It also ensures that a good buffer blank is available should problems (e.g. air bubbles) occur in the first sample.
- Buffers before and after should overlay perfectly.

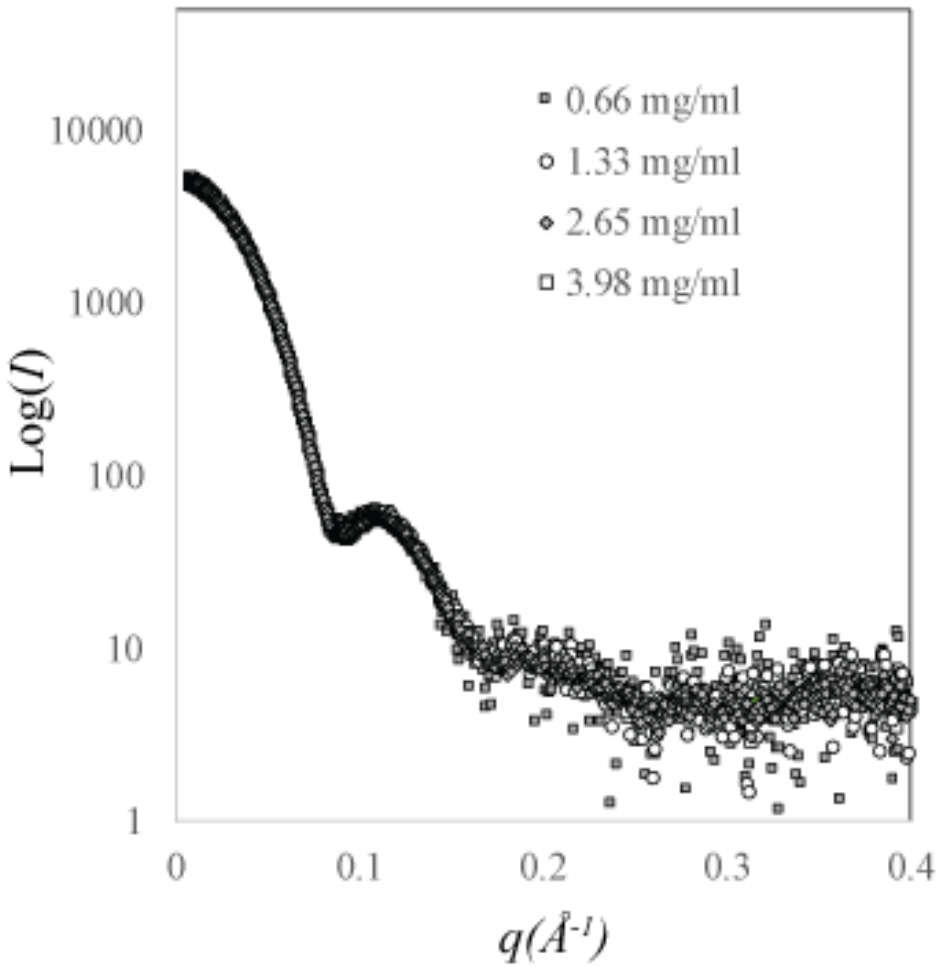
SAXS immediately provides sample characterization



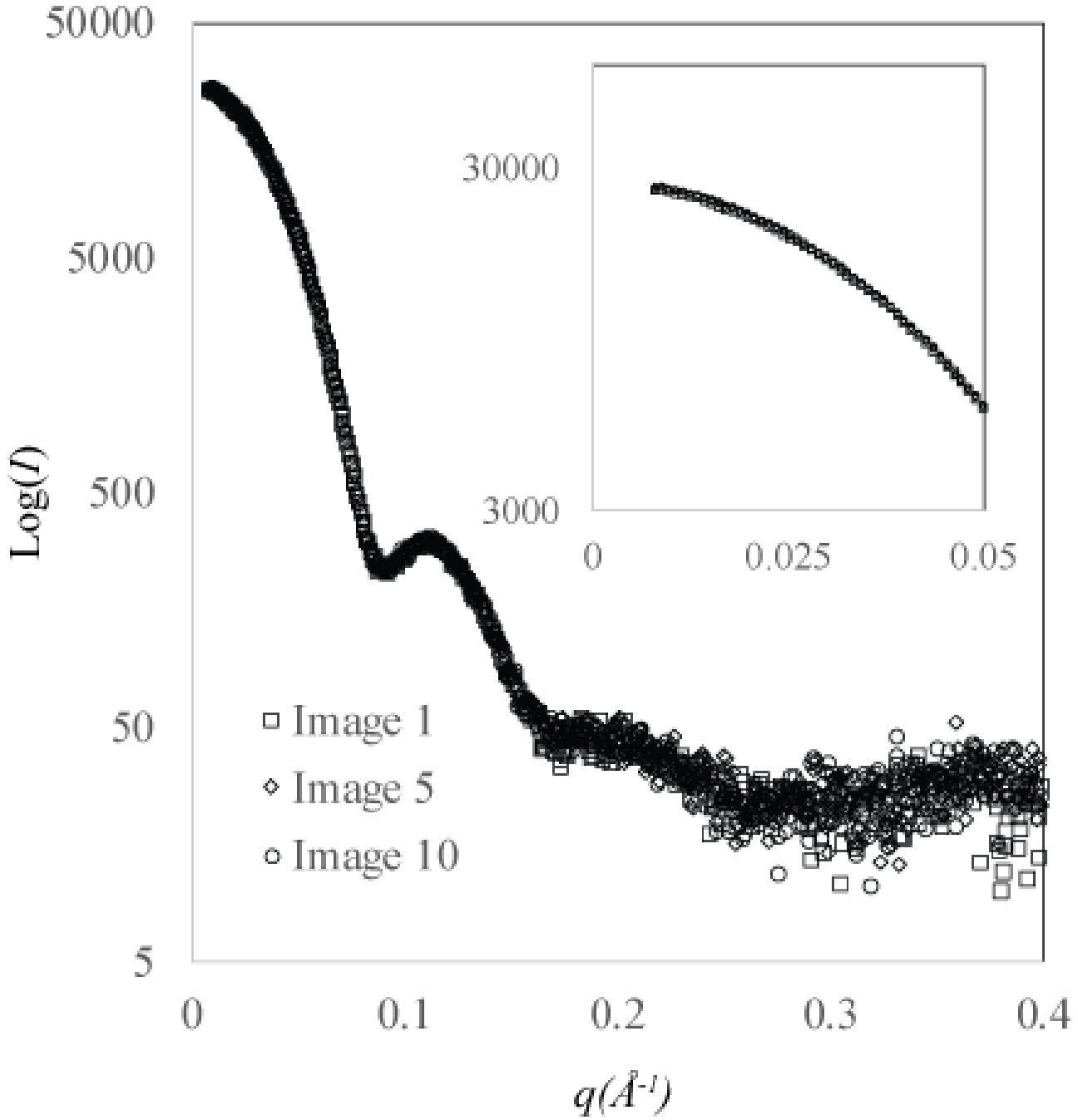
- The scattering profile can characterize the sample in terms of its globularity, domain structure, and flexibility.
- If the sample is unfolded or partially flexible it is difficult to continue with any analysis.
- This is called the Kratky plot.

When scaled by concentration overlap should be perfect





No concentration dependent changes



- Perfect overlay as a function of image – no radiation damage.
- Radiation chemistry causes charges which cause aggregation.
- It also causes bond breakage which can result in unfolding.

Sample characterization: Guinier approximation

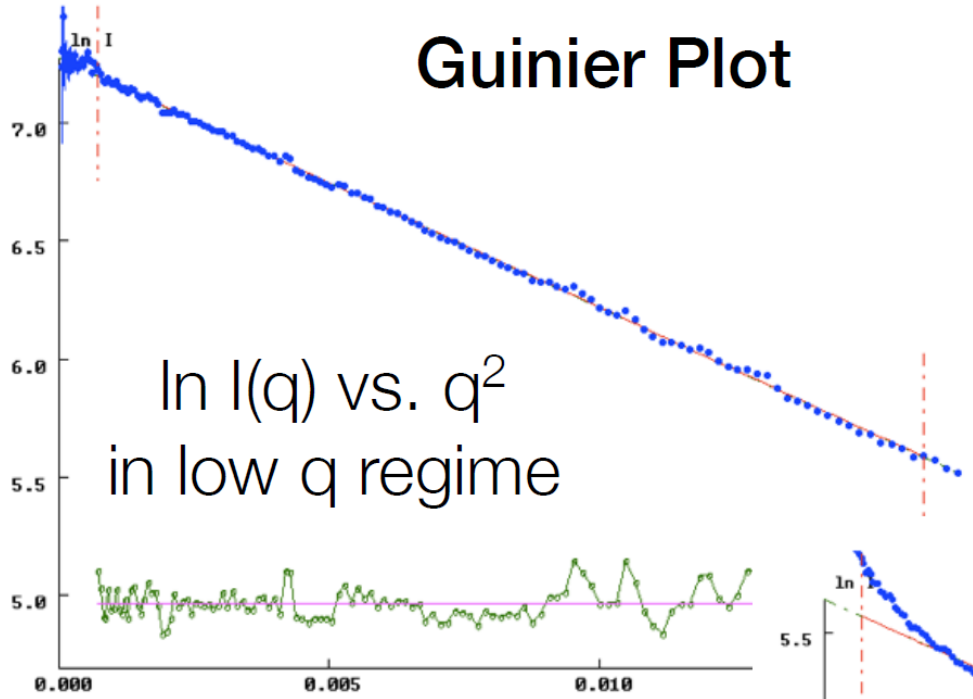
- Developed by André Guinier in 1939.
- As q → 0, intensity can be approximated by:

$$I(q) = I_0 e^{-q^2 R_g^2/3}$$

$$\ln I(q) = \ln I_0 - \frac{R_g^2}{3} q^2$$

y = b + m * x

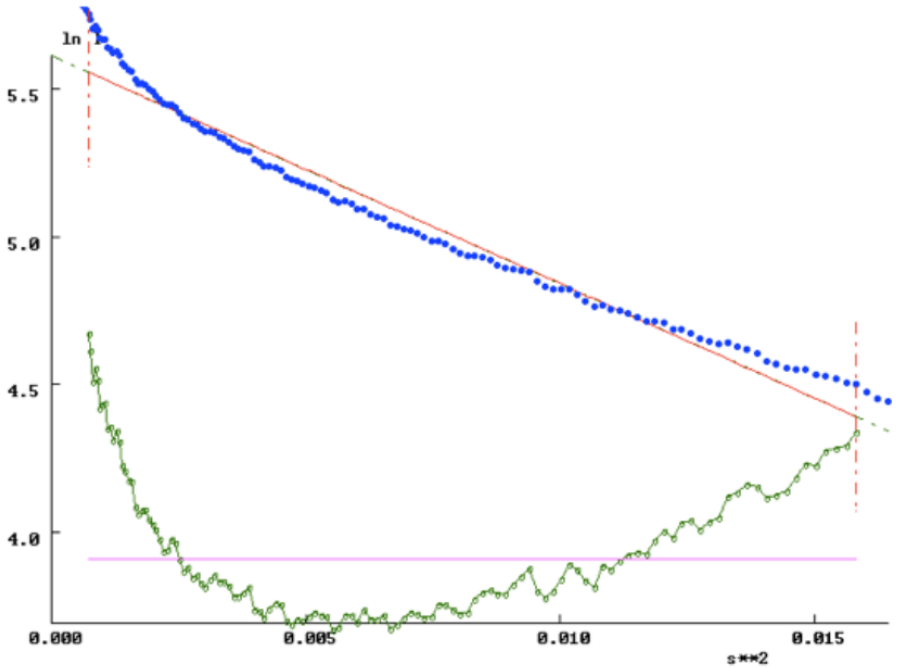
Guinier Plot



ln I(q) vs. q^2
in low q regime

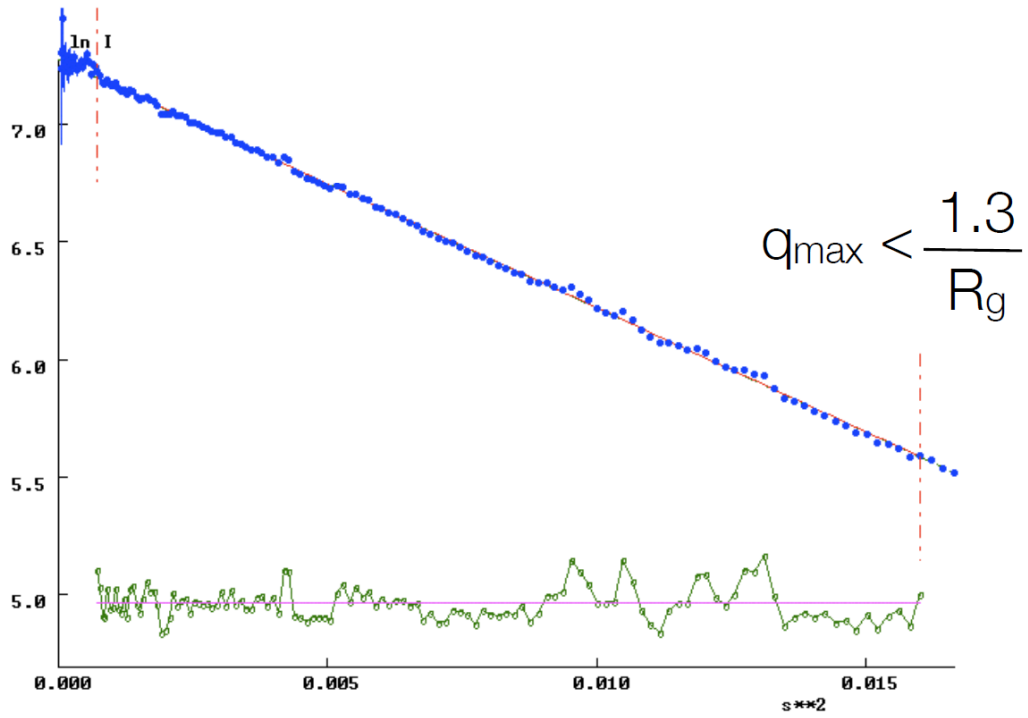
Straight line shows
no aggregation:
Can determine R_g
from slope of line,
I(0) from intercept

Curved line shows
attraction or aggregates
present:
No SAXS processing
should be done



Approximation only valid over a certain region of scattering space

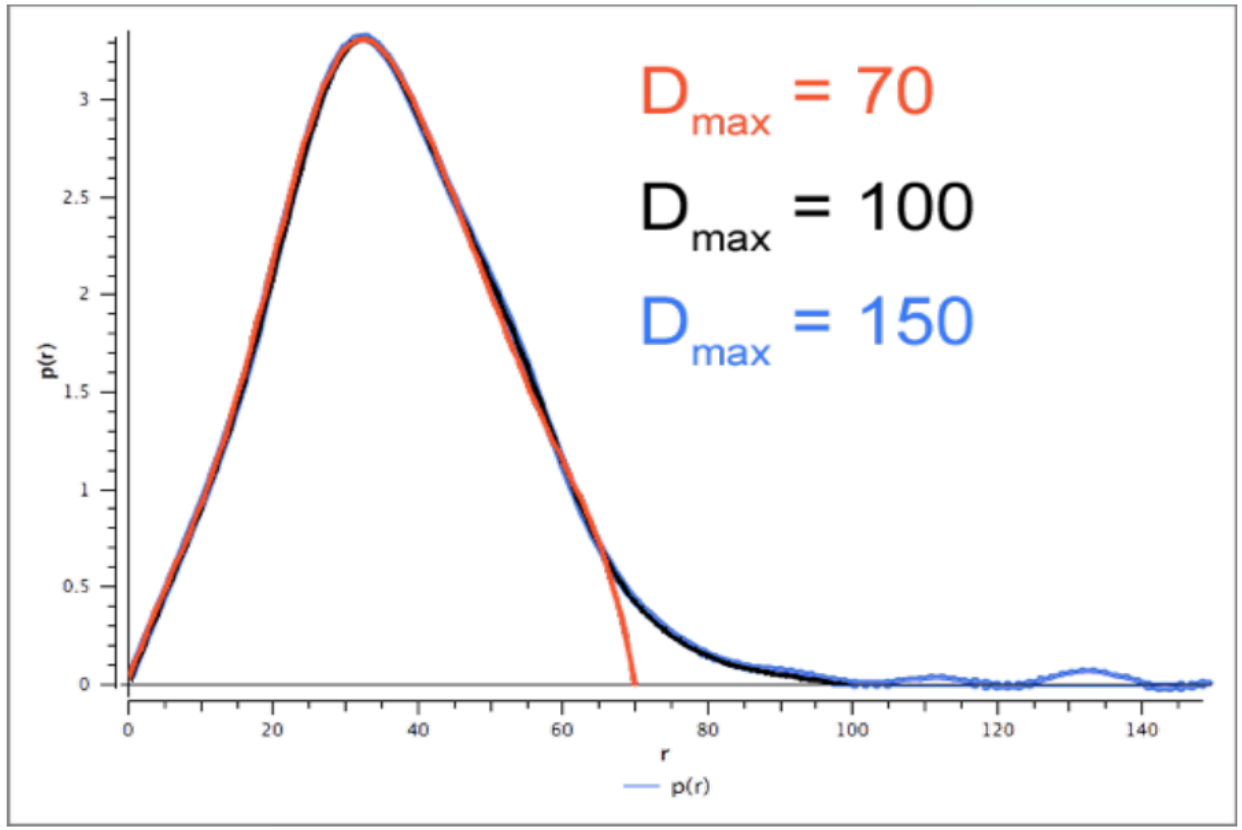
$$q_{min} < \frac{\pi}{D_{max}}$$



$$q_{max} < \frac{1.3}{R_g}$$

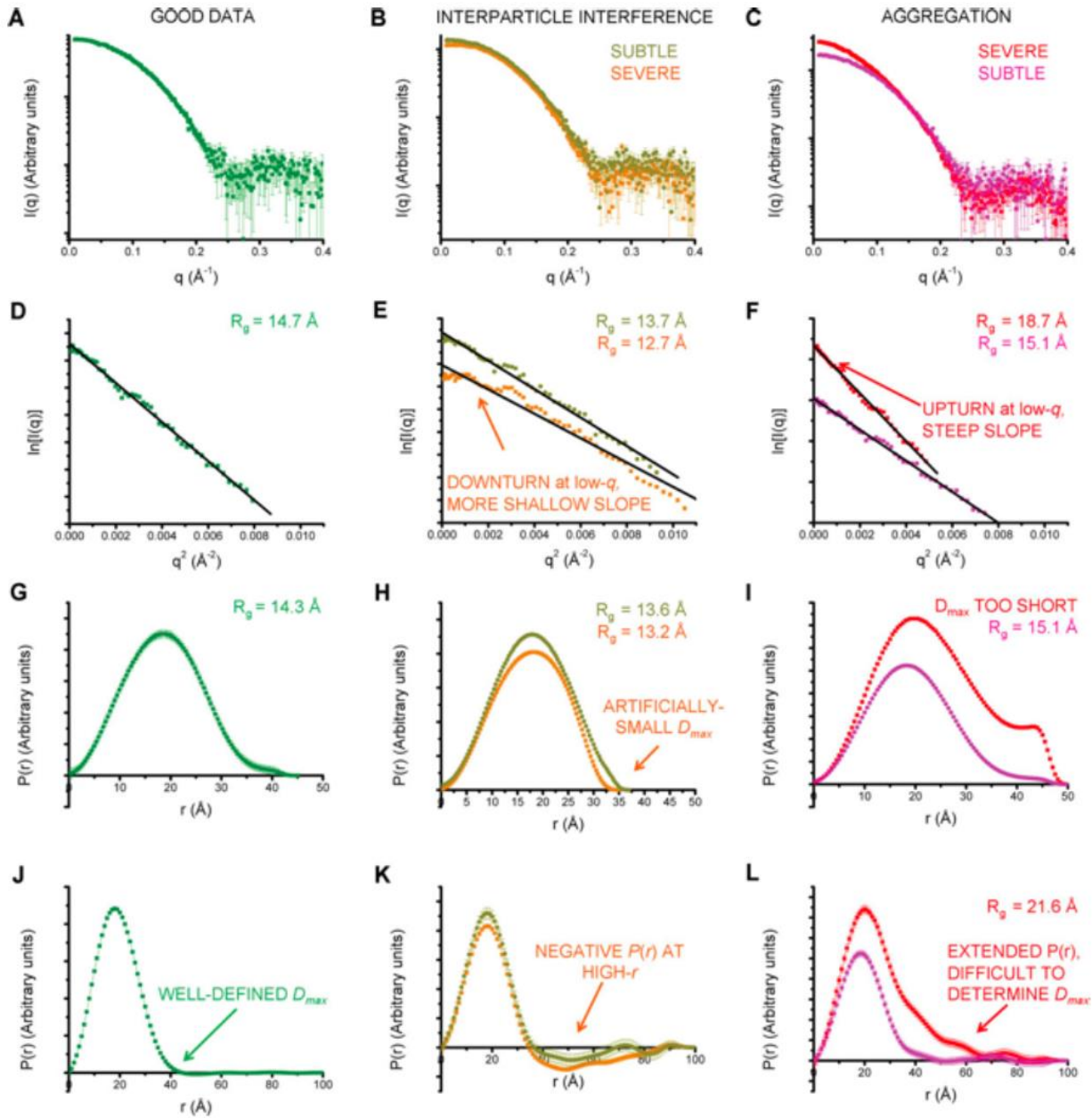
Pair distribution function is used to determine the maximum particle dimension

- Can be used to determine D_{max}
- P(r) should gradually fall to zero at D_{max}
- Underestimated D_{max} appears as abrupt, forced descent to zero
- Starting with large values should identify a decent estimate of D_{max}, given good quality data
- Errors in D_{max} can be large, (~10 - 20%) for good data



The maximum particle dimension is given by the distance between the furthest interatomic scattering

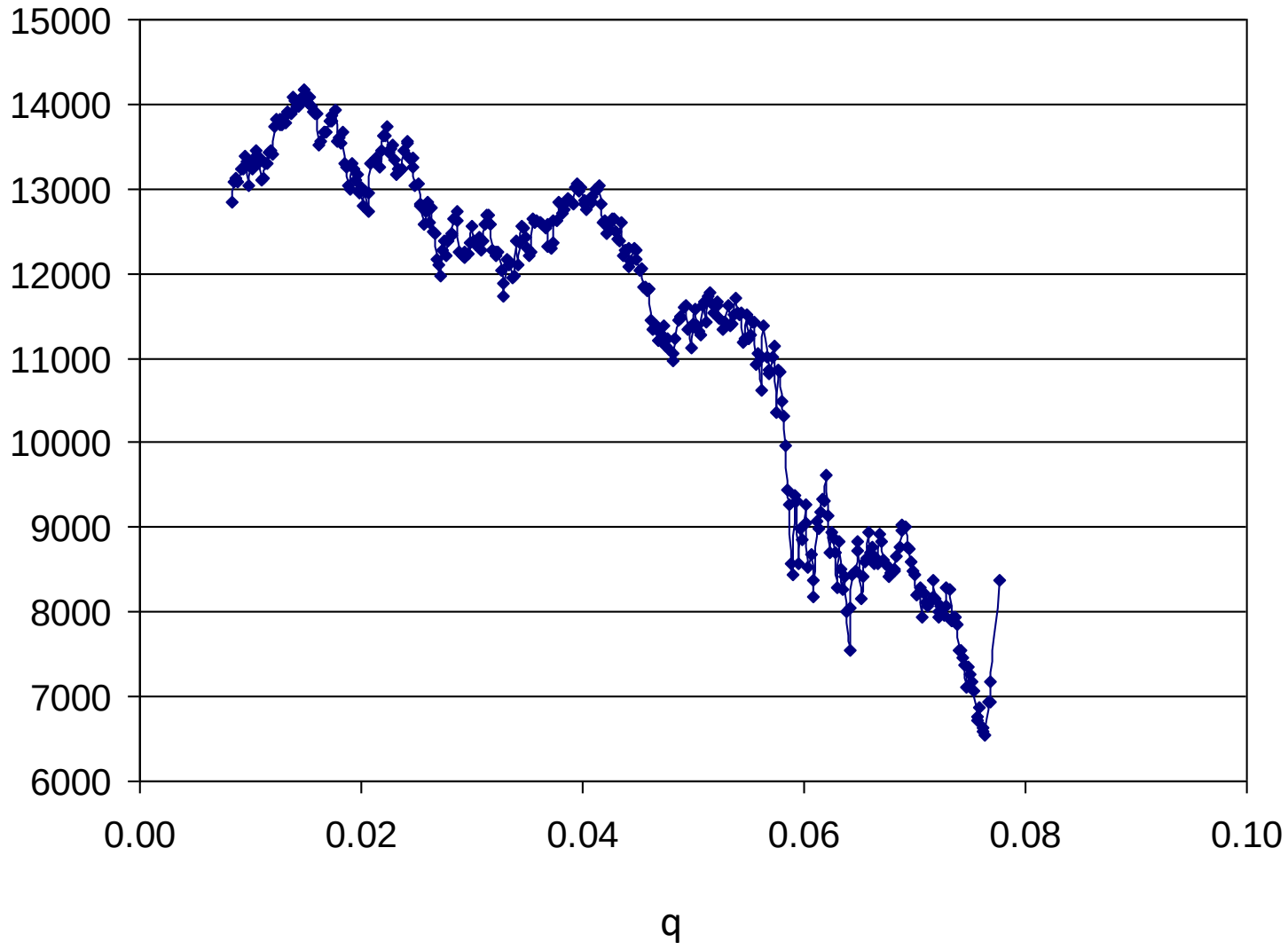
Sample quality greatly affects data analysis



Jacques and Trewhella, 2010 *Protein Science* Review

A limitation of the technique is that good or ‘bad’ data can produce a result

Lets take some '*scattering*' data

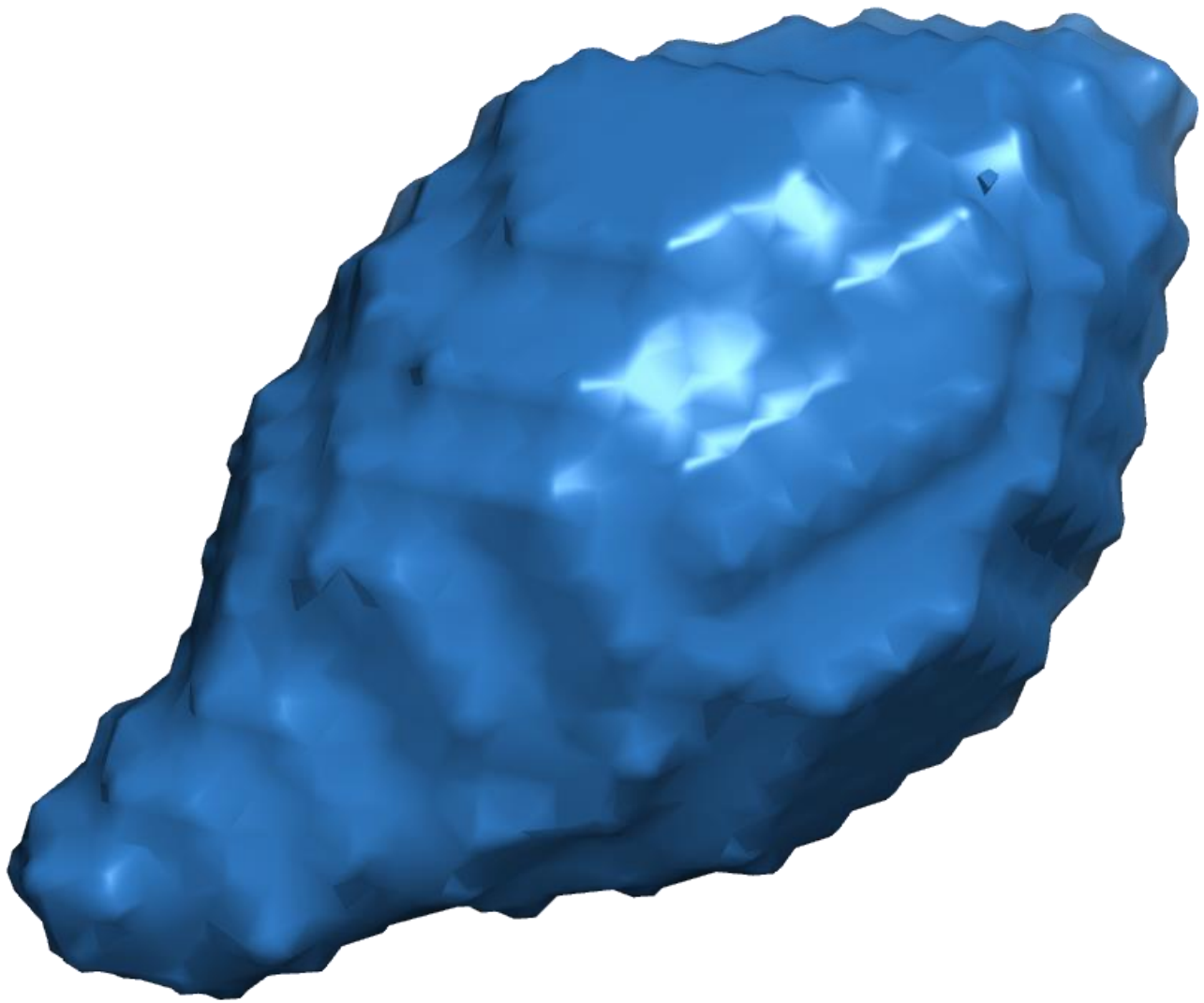


Note for SAXS experts, don't scream about the data quality – bare with me on this.

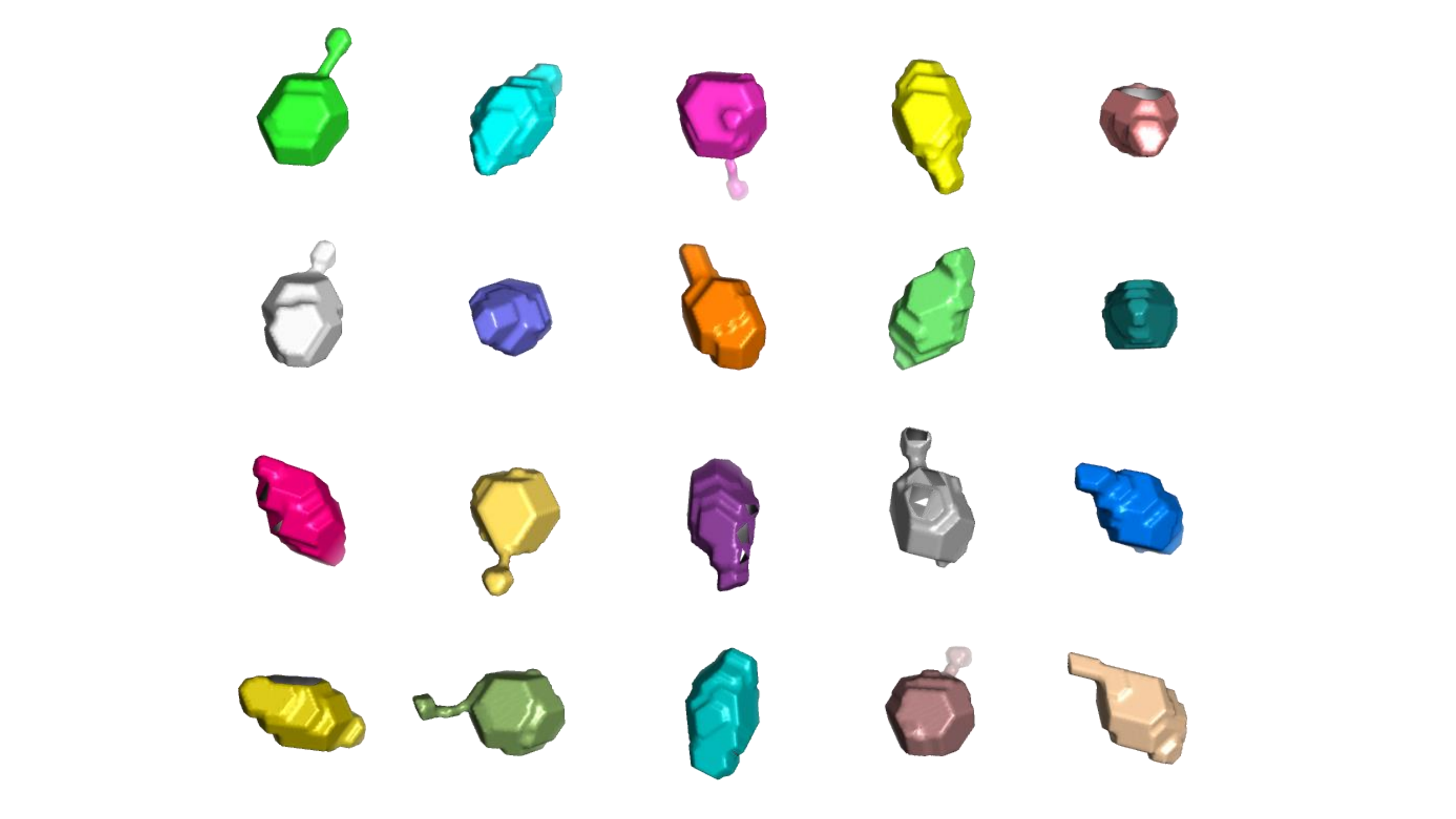
Envelope Reconstruction (typical procedure)

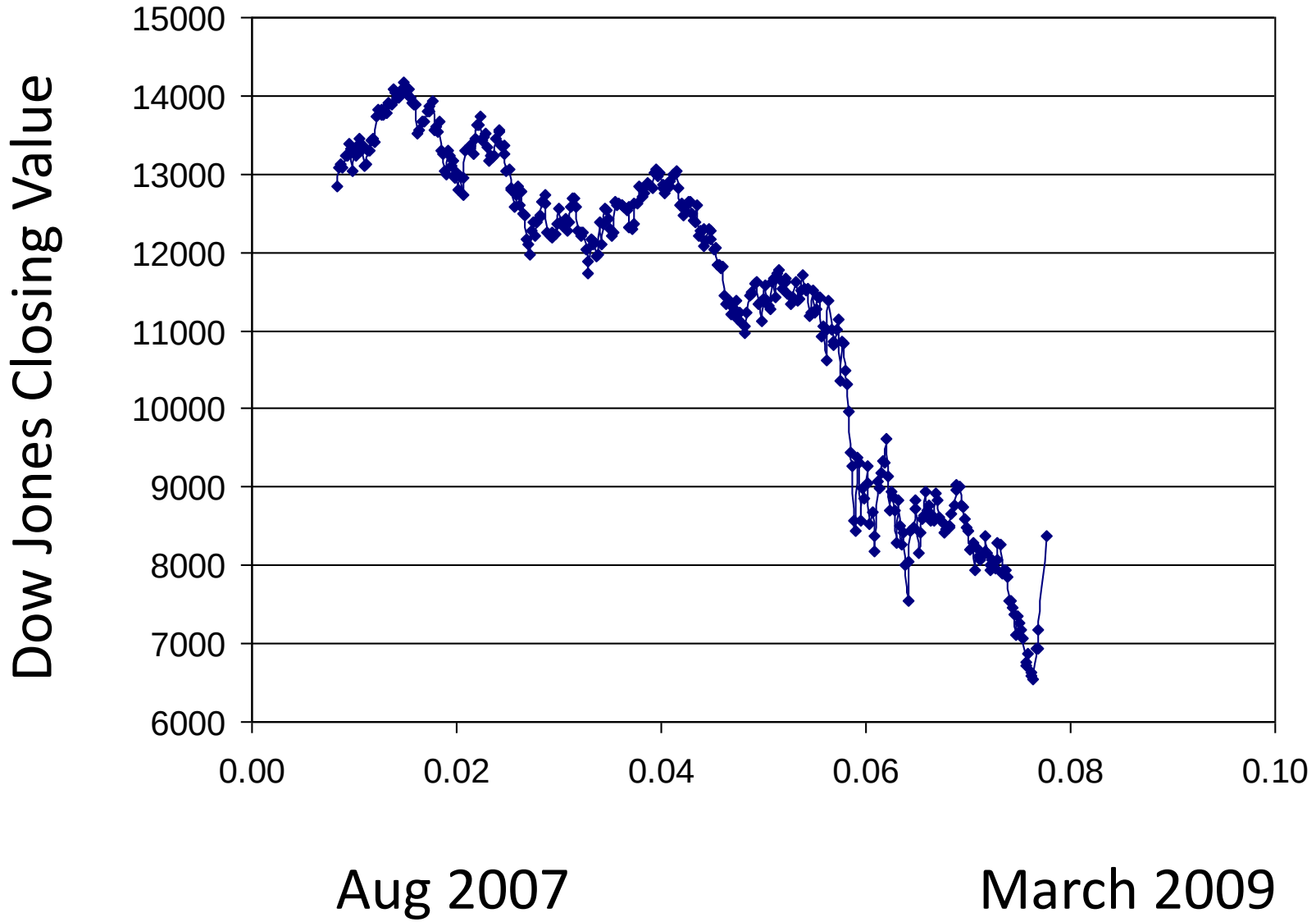
- Produce 10-20 *ab initio* reconstructions
- Determine the most probable model, i.e. the least different from the rest and align all to this.
- Estimate the similarity of the models using the Normalized Spatial Discrepancy (NSD)
  - Average NSD ~ 0.5 implies good stability of solution
  - Average NSD ~ 0.7-0.9 implies fair stability
  - Average NSD > 1.0 implies poor stability.
- NSD can yield some idea of flexibility or possible oligomeric mixtures.
- DAMAVER can be used to select the most populated volume from all reconstructions.

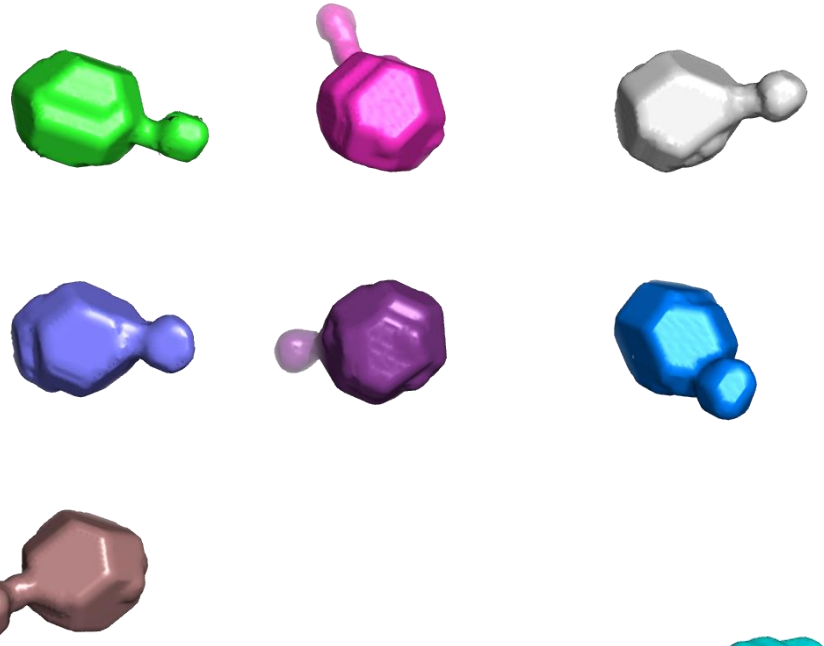
Terminology note – Envelope is the boundary of the protein and the buffer, reconstruction is the term used to describe determining that envelope



NSD = 0.613, 20 reconstructions




Dow Jones Closing Value
15000
14000
13000
12000
11000
10000
9000
8000
7000
6000
0.00
0.02
0.04
0.06
0.08
0.10
Aug 2007
March 2009



Actually two populations

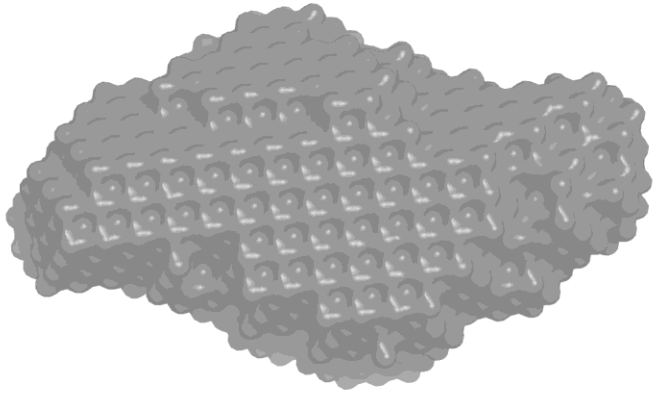
A Bull or a Bear market!

Both are correct, i.e. they explain the scattering data

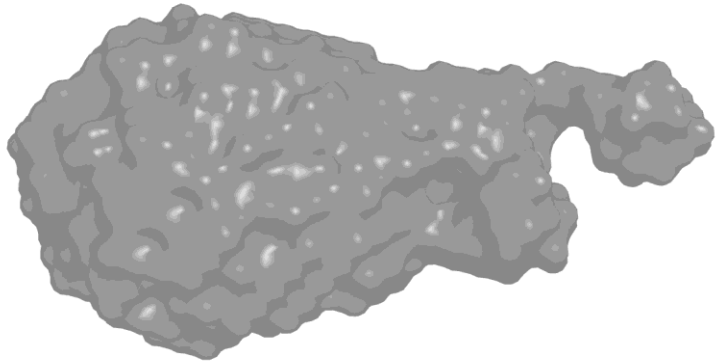
What can SAXS/SANS provide when done well?

SAXS can determine *ab initio* molecular envelopes

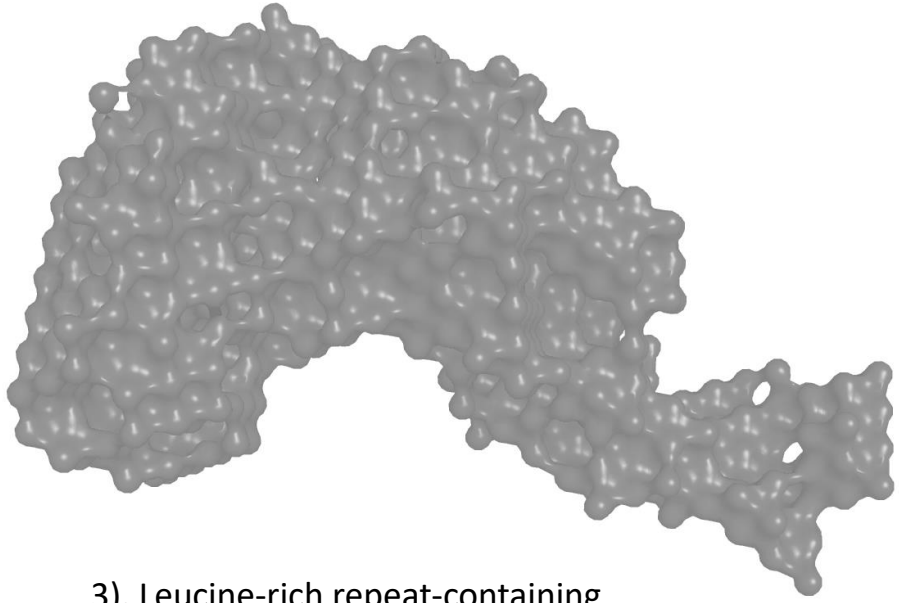
Ab intio envelopes



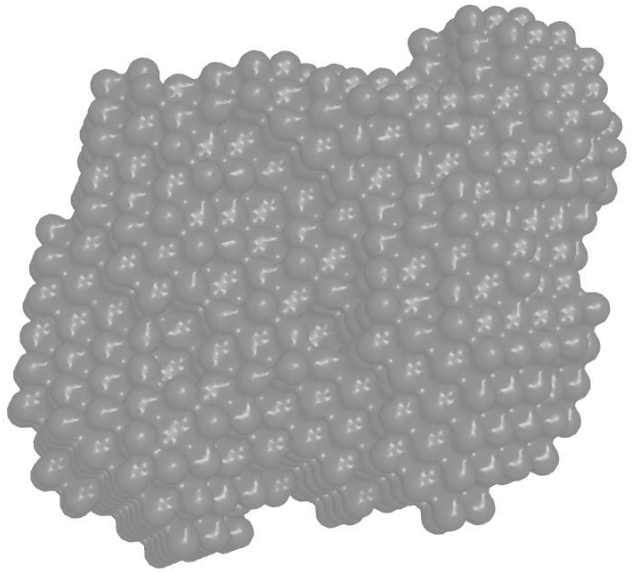
1). alr0221 protein from Nostoc (18.6 kDa)



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



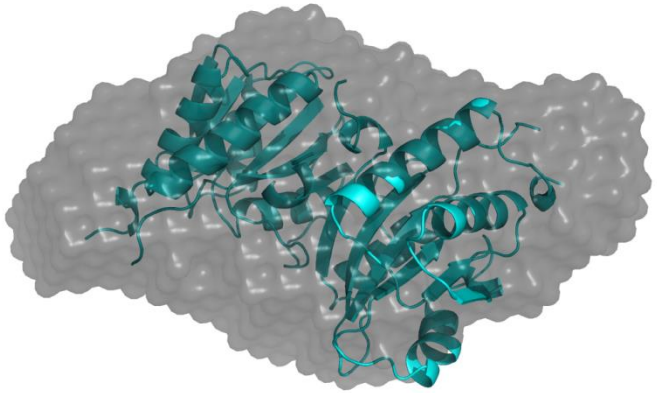
3). Leucine-rich repeat-containing protein LegL7 (39 kDa)



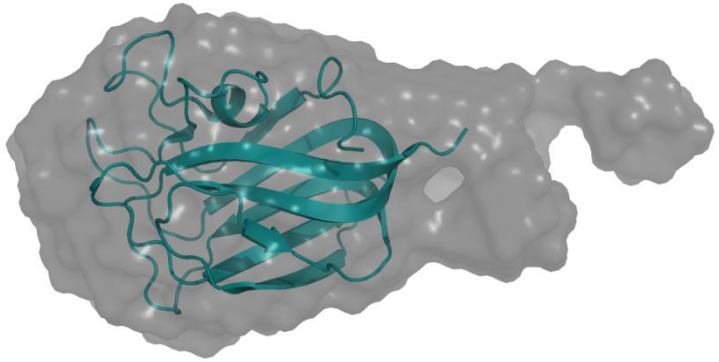
4). *E. Coli.* Cystine desulfurase activator complex (170 kDa)

Ab initio envelopes are compatible with structural models

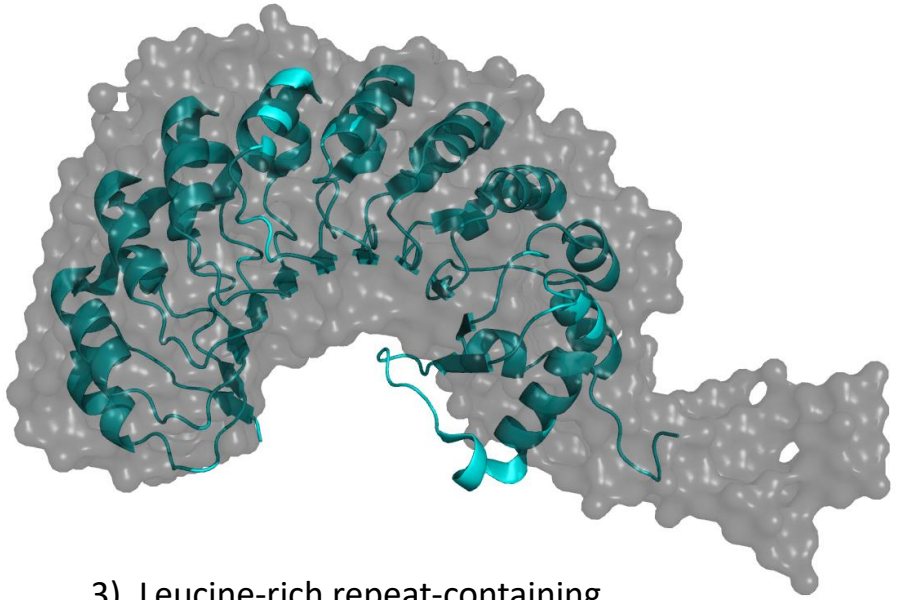
Overlaid with subsequent X-ray structures



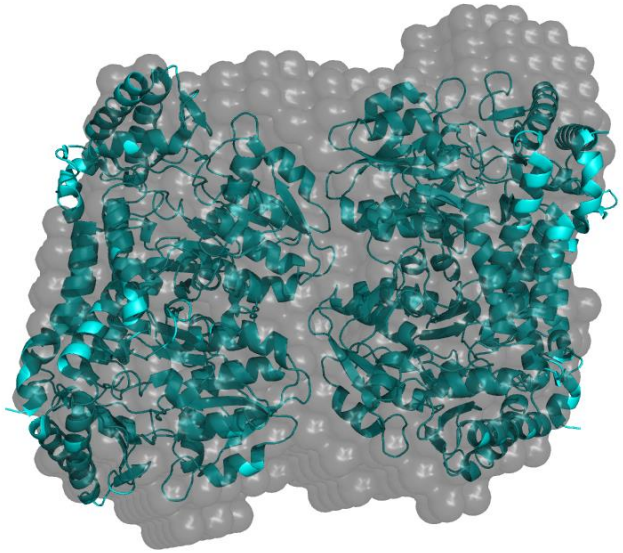
1). alr0221 protein from Nostoc (18.6 kDa)



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



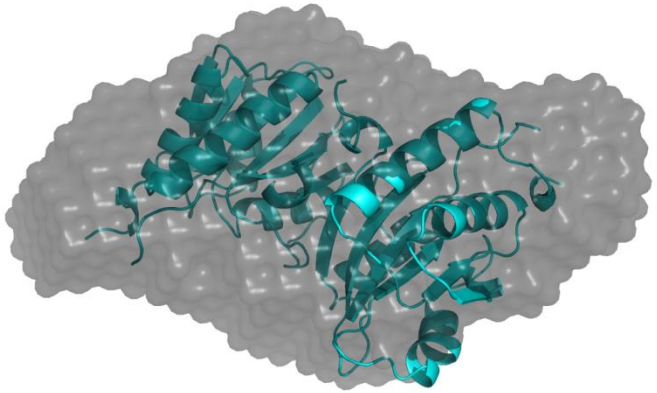
3). Leucine-rich repeat-containing protein LegL7 (39 kDa)



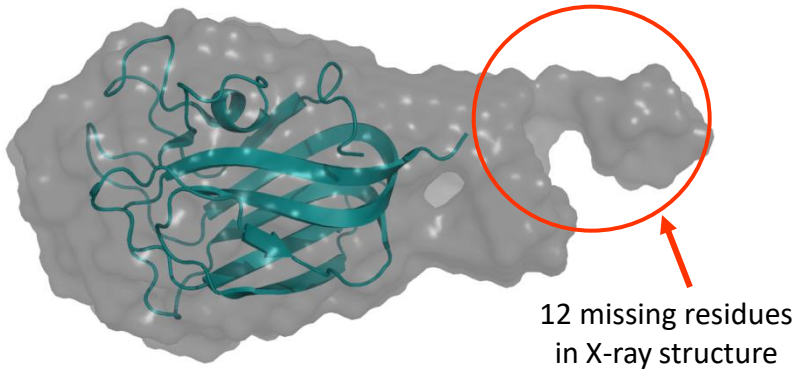
4). *E. Coli.* Cystine desulfurase activator complex (170 kDa)

And they provide extra information on residues present in the construct but structurally undefined

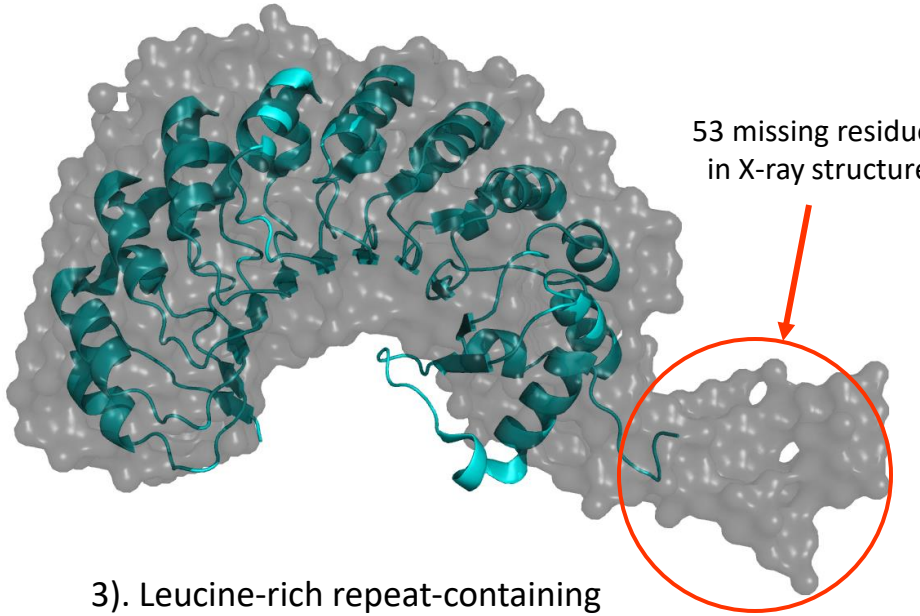
And data on what was missing …



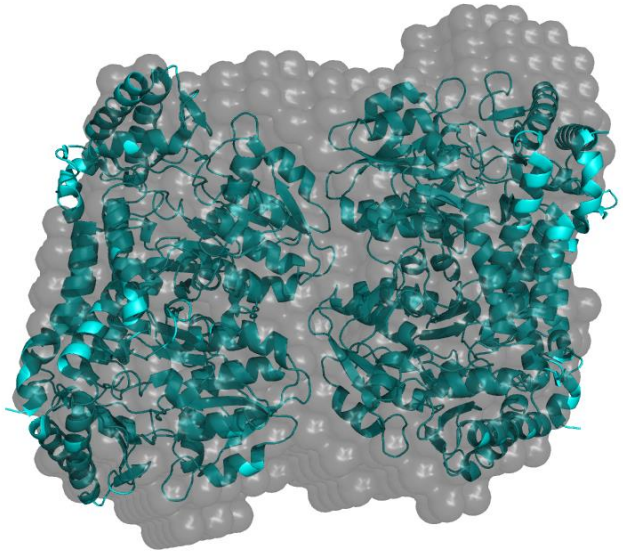
1). alr0221 protein from Nostoc (18.6 kDa)



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

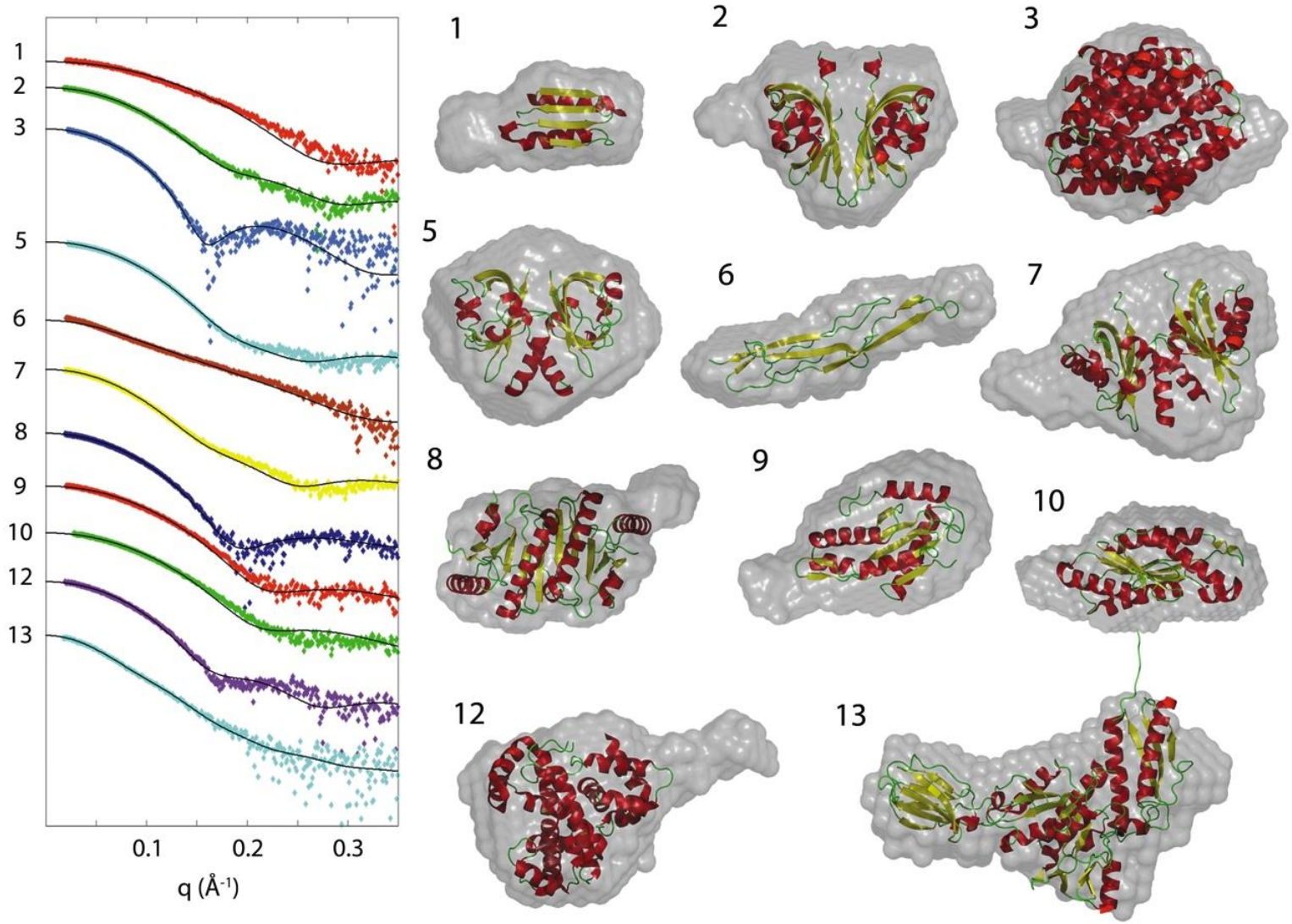


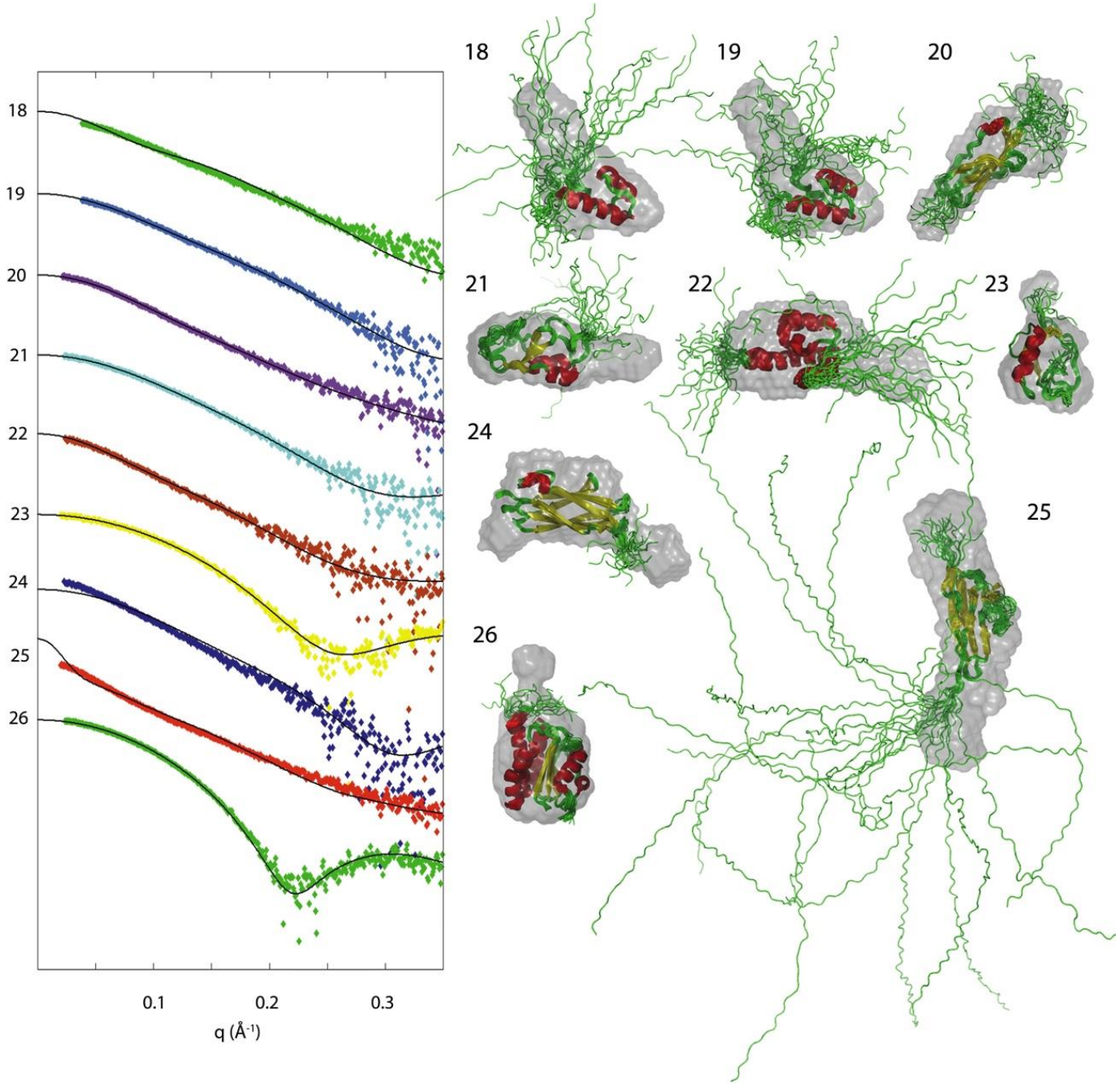
3). Leucine-rich repeat-containing protein LegL7 (39 kDa)



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

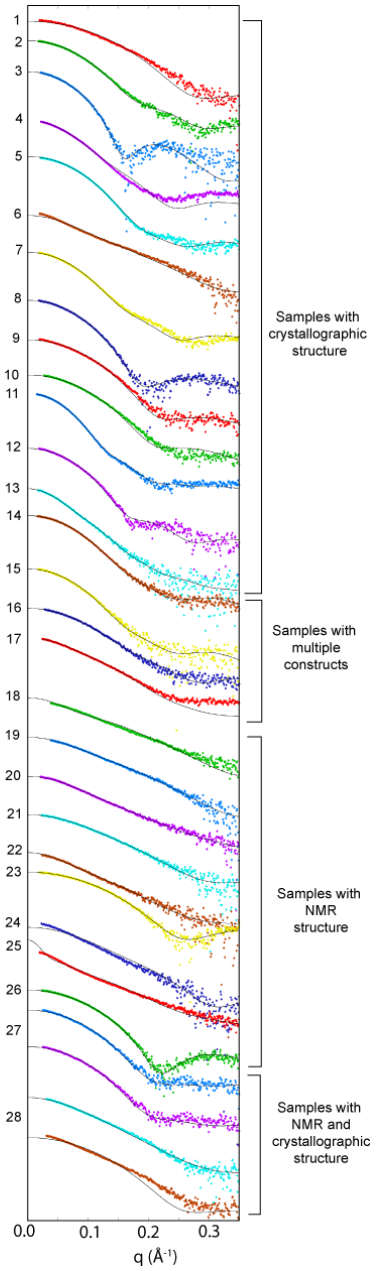
Comparing X-ray structures

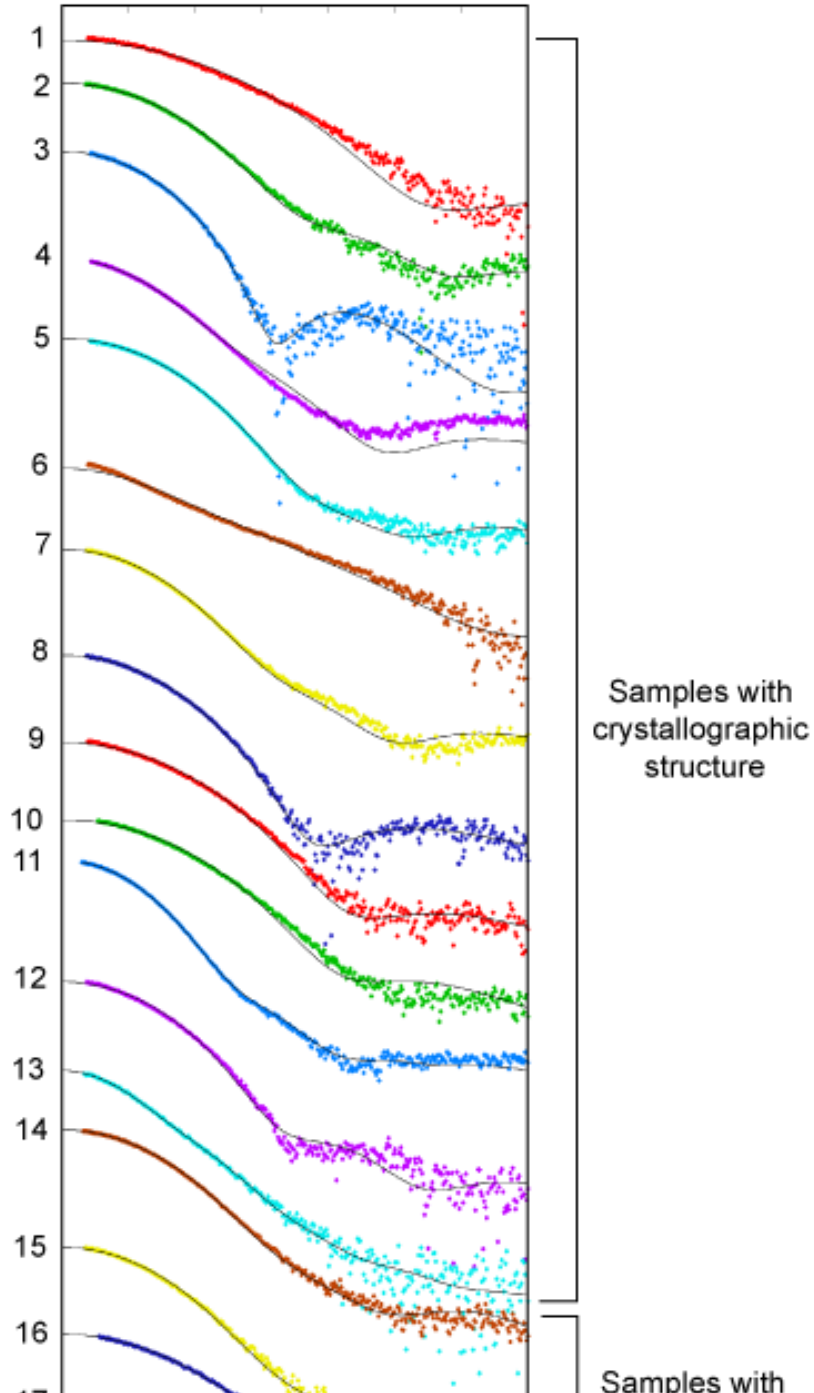




Comparing NMR structures

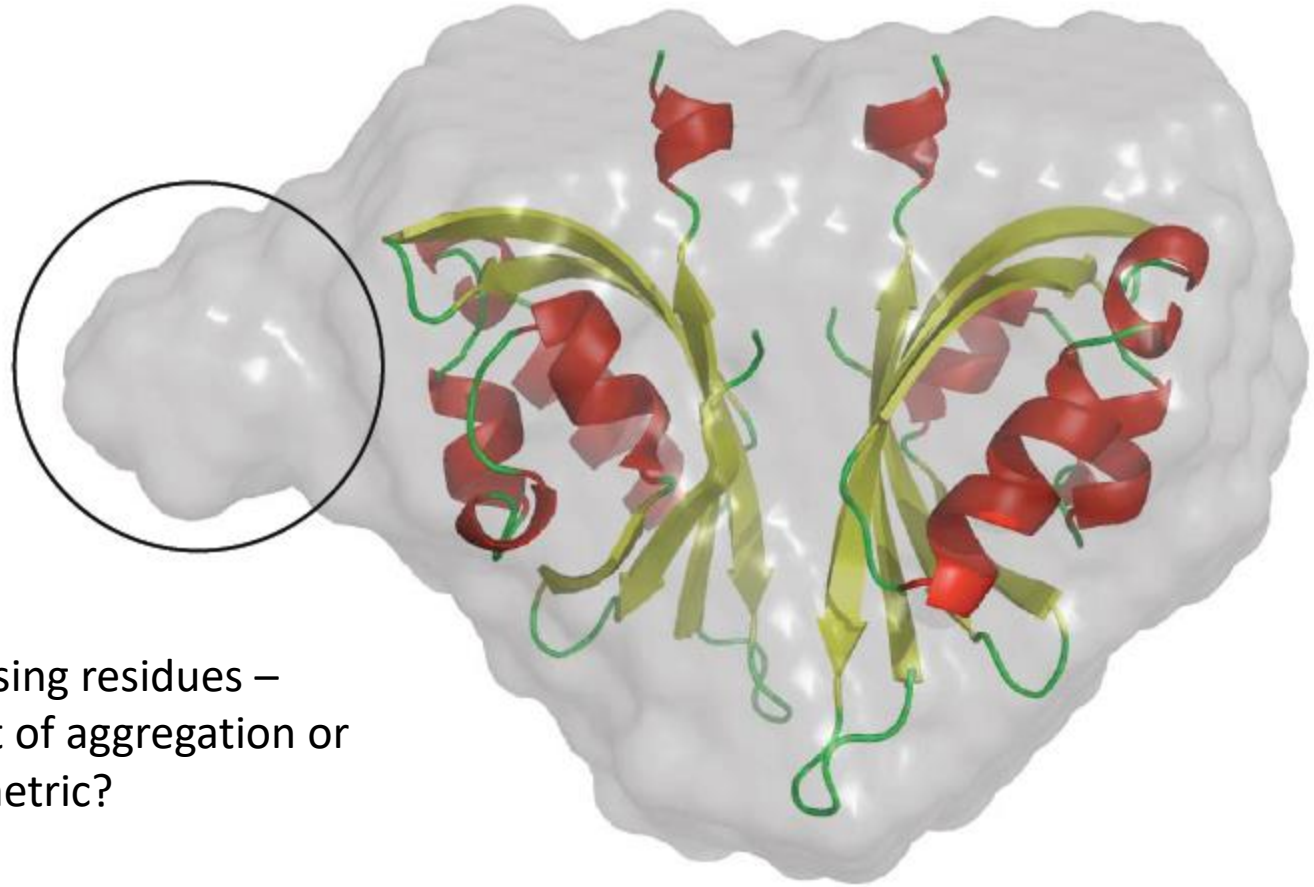
20 lowest energy Conformations shown


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

SAXS may provide more questions

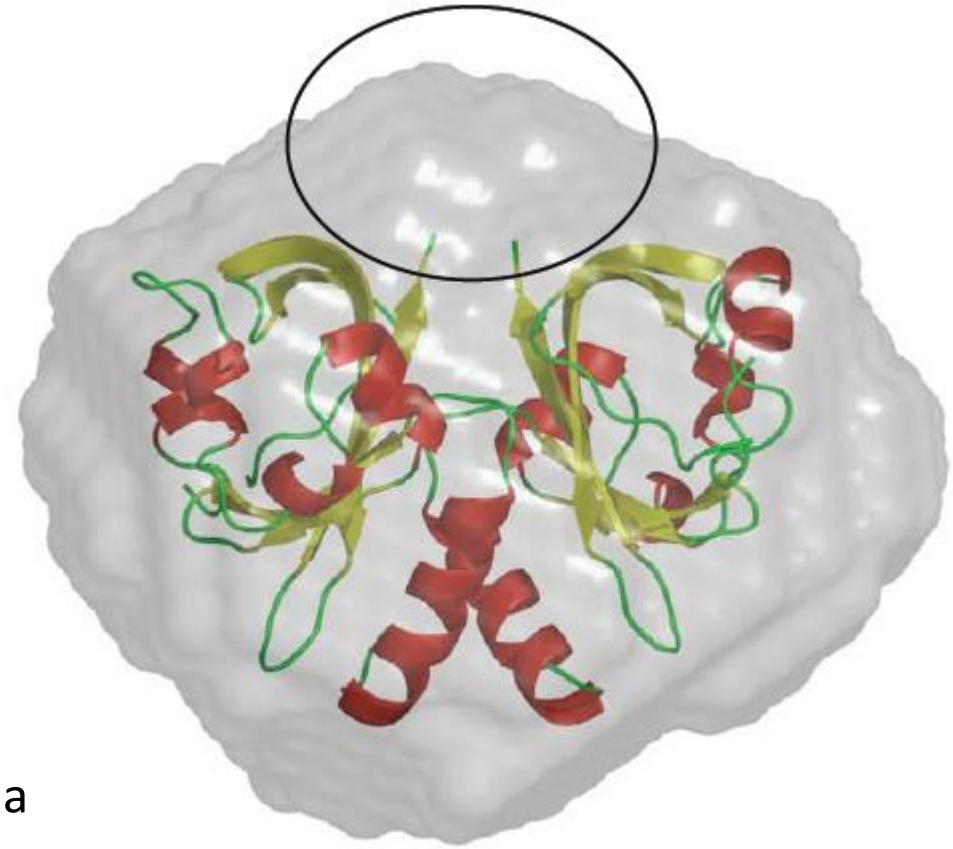
Diguanylate cyclase



12 missing residues – artifact of aggregation or asymmetric?

12 missing residues – artifact of aggregation or asymmetric?

Sensory Box/GGDEF Protein Family



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

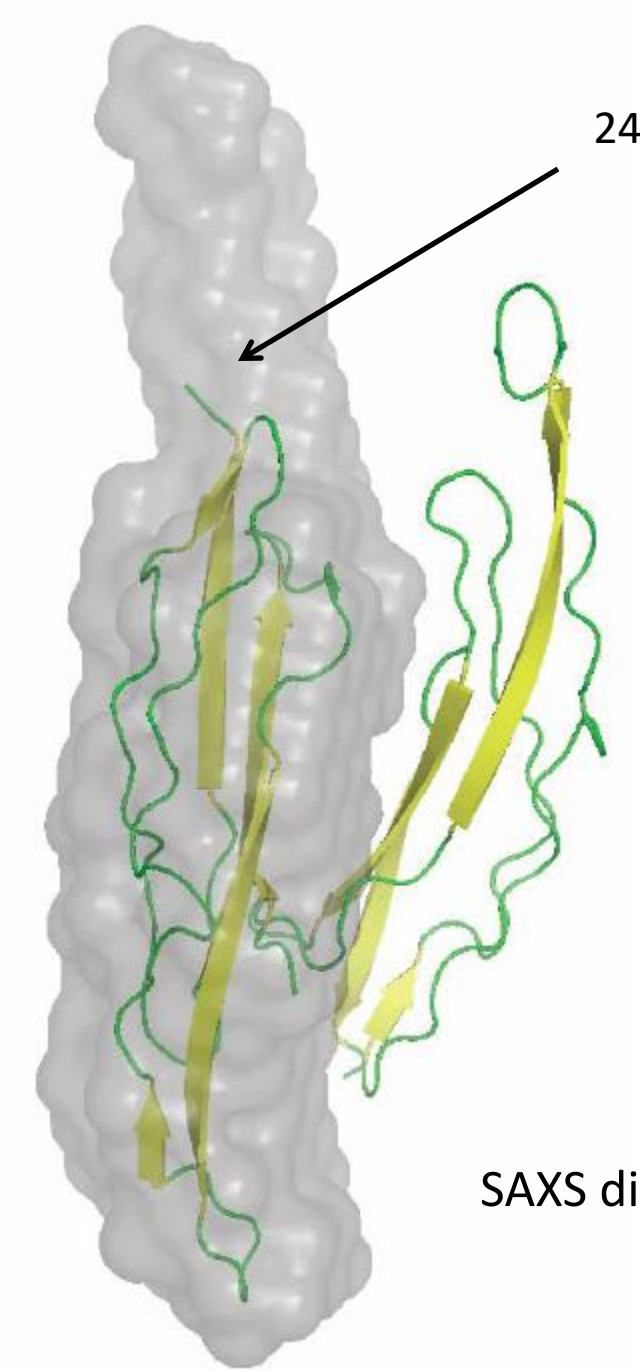
SAXS may be ambiguous

MucBP Domain of PEPE_0118

Biological unit was thought to be a dimer from crystallography.

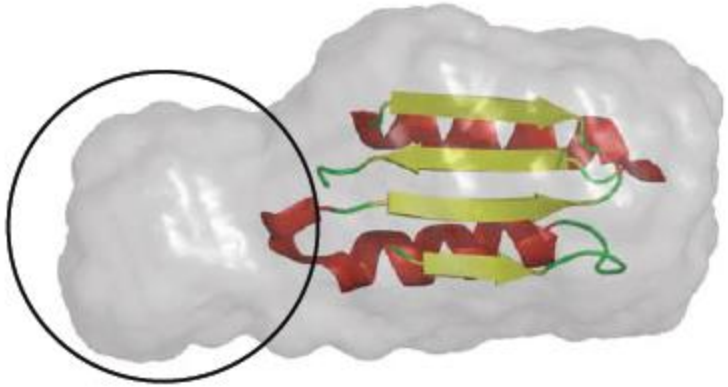
Solution state is not.

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

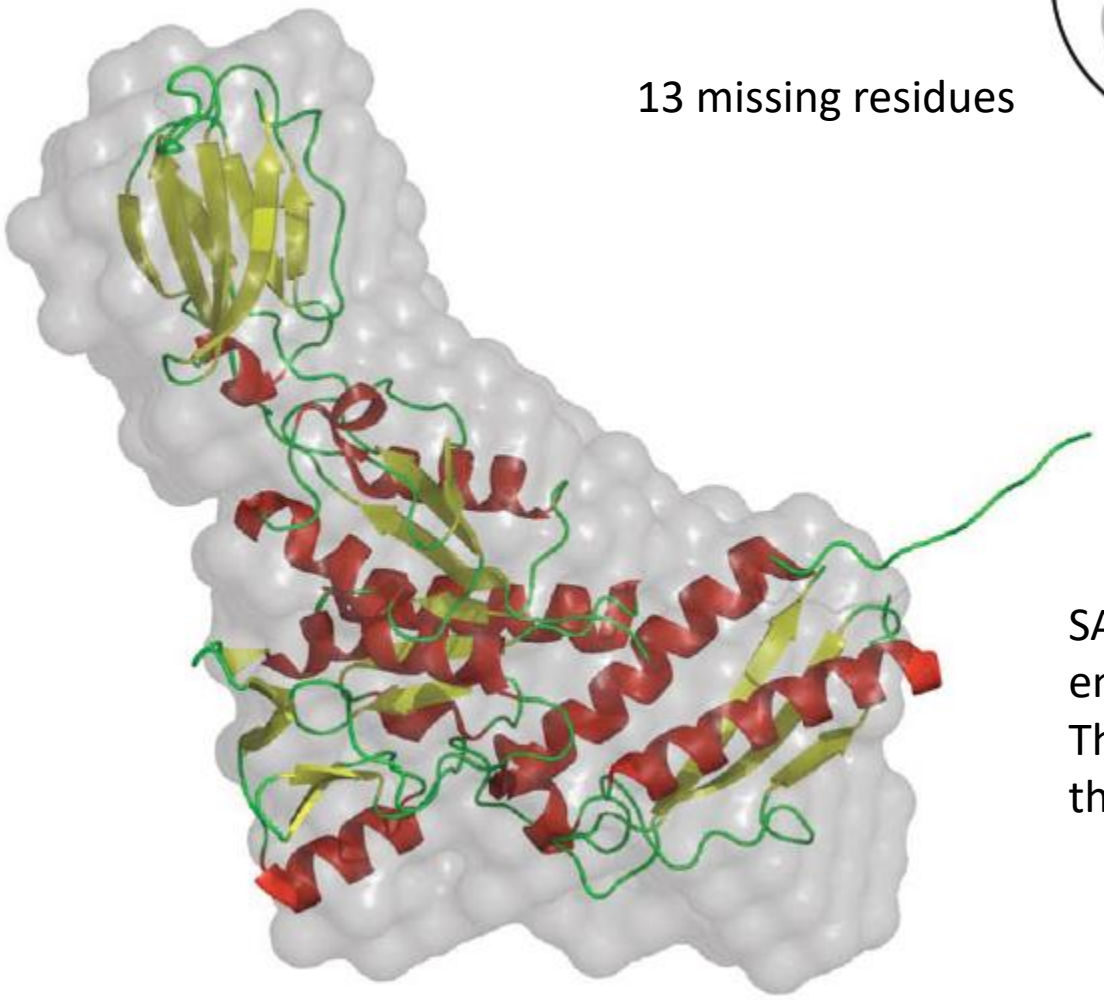


SAXS distinguishes solution states

Size matters



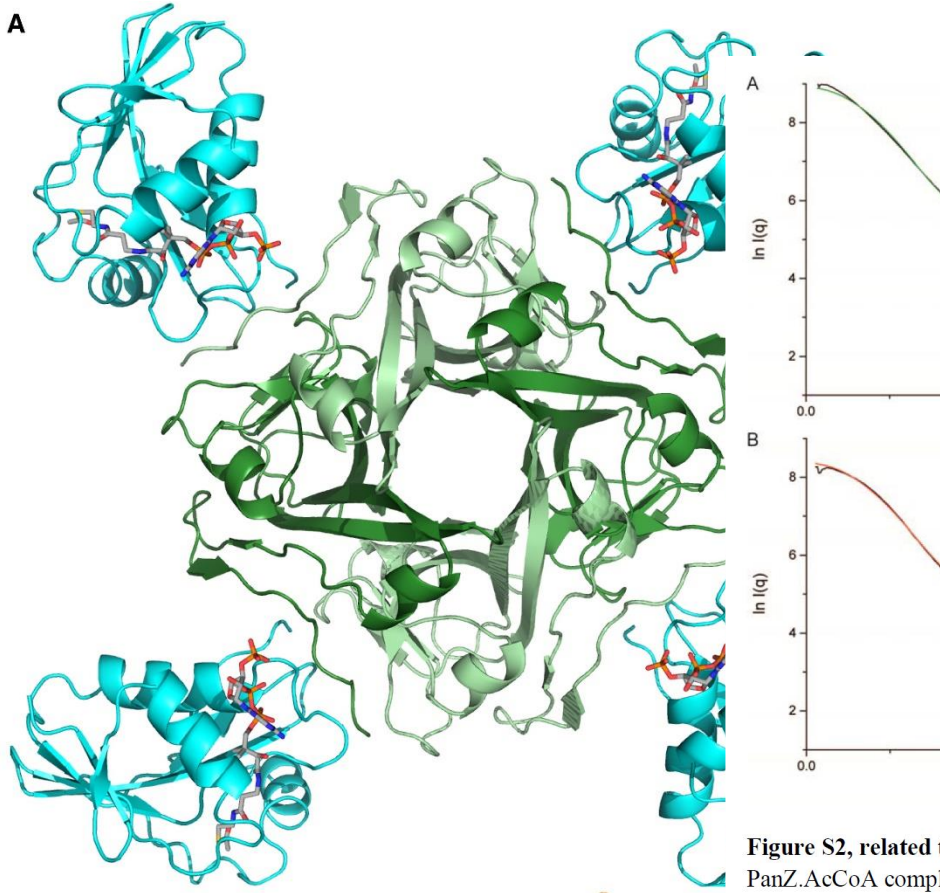
13 missing residues

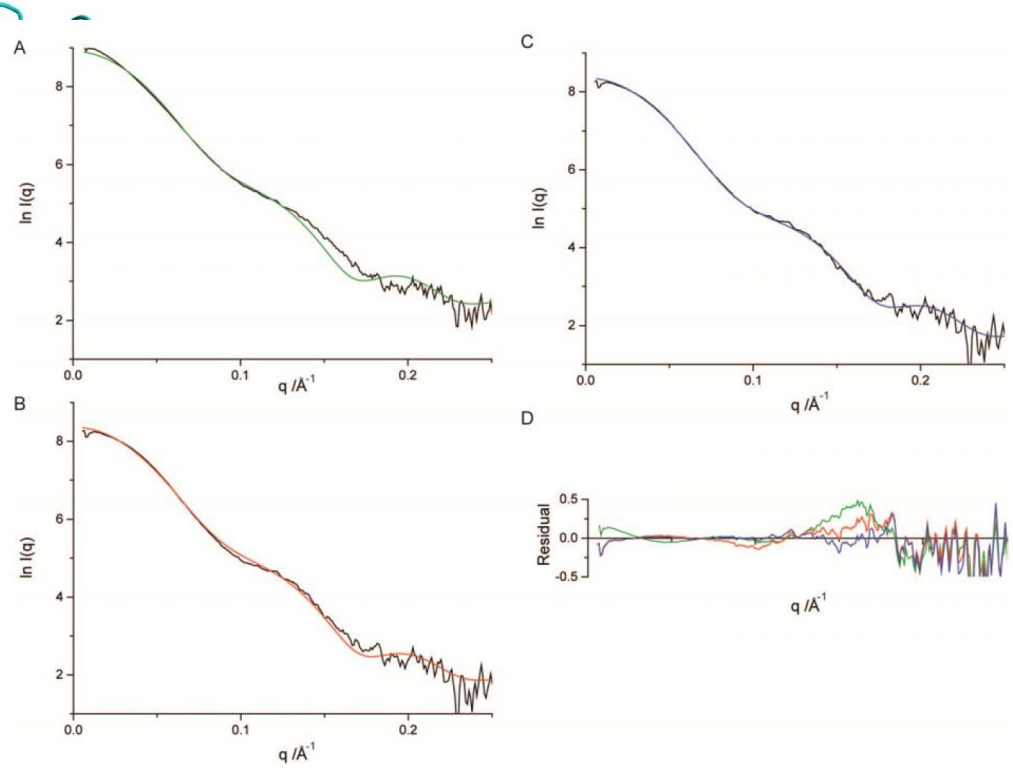


SAXS is not just about shape of the envelope but also it's overall size. The envelope produced reflects the size of the sample.

But ab initio shape reconstruction is the least useful capability for SAXS

- It is possible to accurately model a SAXS or SANS profile
- SAXS and SANS data provides
  - Molecular mass M
  - Radius of gyration R_g
  - Porod invariant Q
  - Particle volume V
  - Maximum particle dimension D_{max}
  - Particle surface area S
  - Correlation length l_c
  - Volume of correlation V_c
- SAXS can be used to test hypothesis but not validate them.





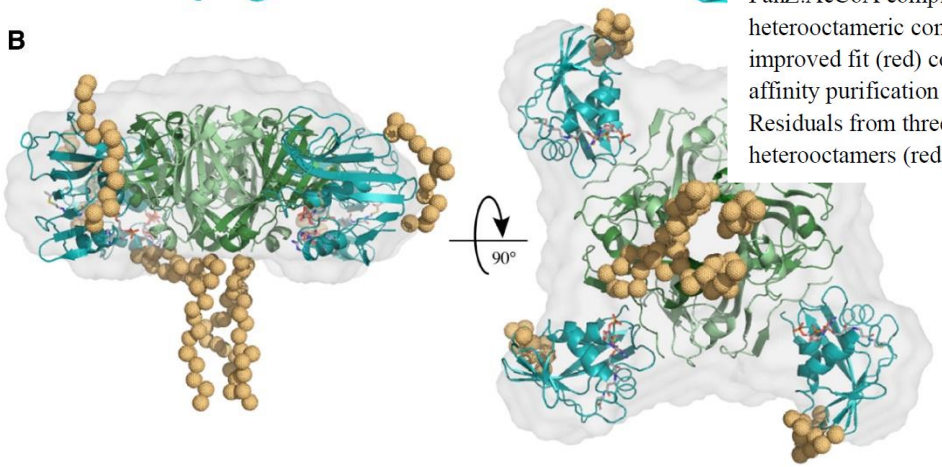
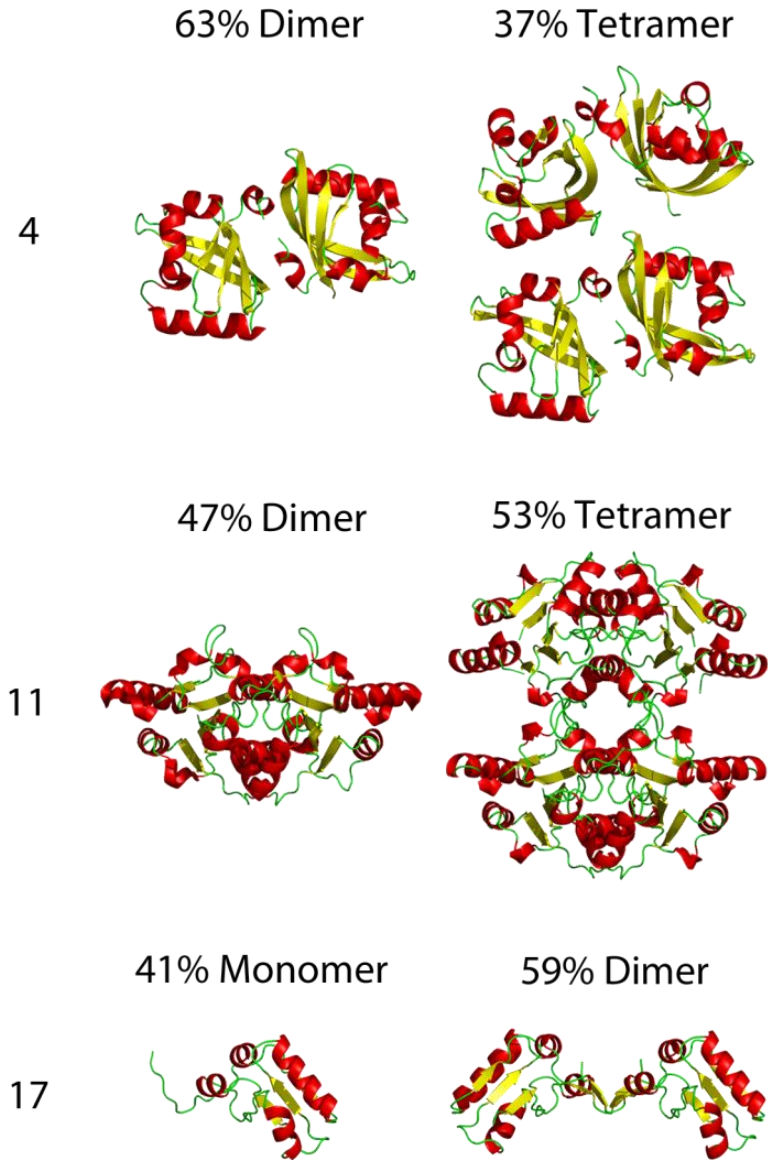
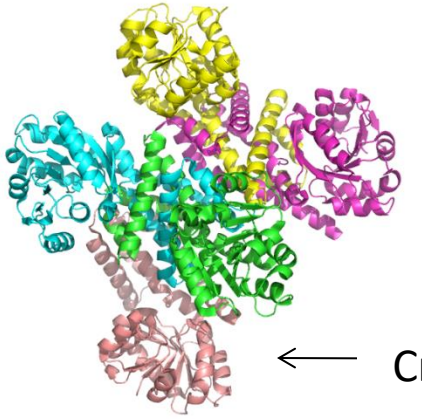


Figure S2, related to figure 2 SAXS data and fitting. **A** Raw SAXS data for the PanD(T57V)-PanZ.AcCoA complex (black) compared with predicted data for the crystallographically resolved heterooctameric complex (green). **B** Inclusion of a population of dimers of heterooctamers leads to an improved fit (red) compared to the monomer. **C** Subsequent inclusion of the eight *C*- and *N*-terminal affinity purification tags using a coarse-grained model leads to a further improved fit (blue). **D** Residuals from three sequential rounds of data fitting: heterooctamer (green), inclusion of dimer of heterooctamers (red), inclusion of affinity tags (blue).

The Structure of the PanD/PanZ Protein Complex Reveals Negative Feedback Regulation of Pantothenate Biosynthesis by Coenzyme A Monteiro, et al., Chemistry & Biology Volume 22, Issue 4, Pages 492-503 (April 2015)



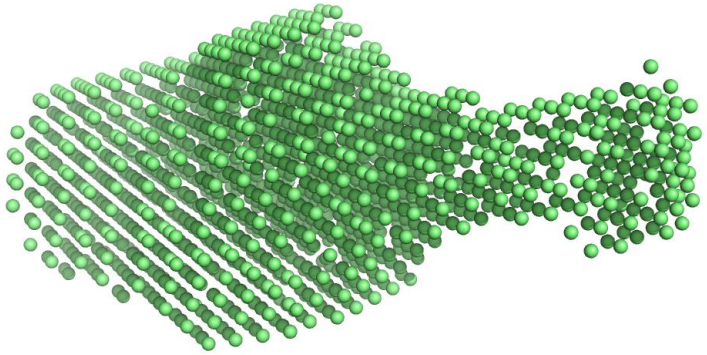
Identification of mixtures if you know the initial structure (another story)



2izz from the PDB
(5 chains in PDB)

← Crystal packing artifact

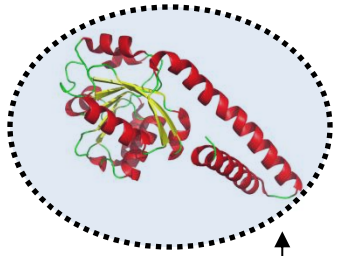
Solution envelope from BcR38B-21.20-SeMa-Gf (3gt0)



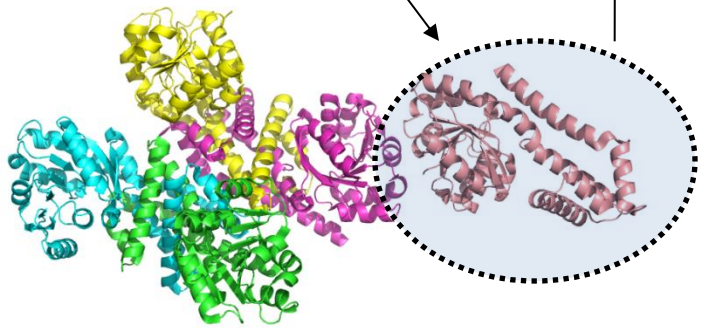
~165A

Another story

3gt0 from the PDB



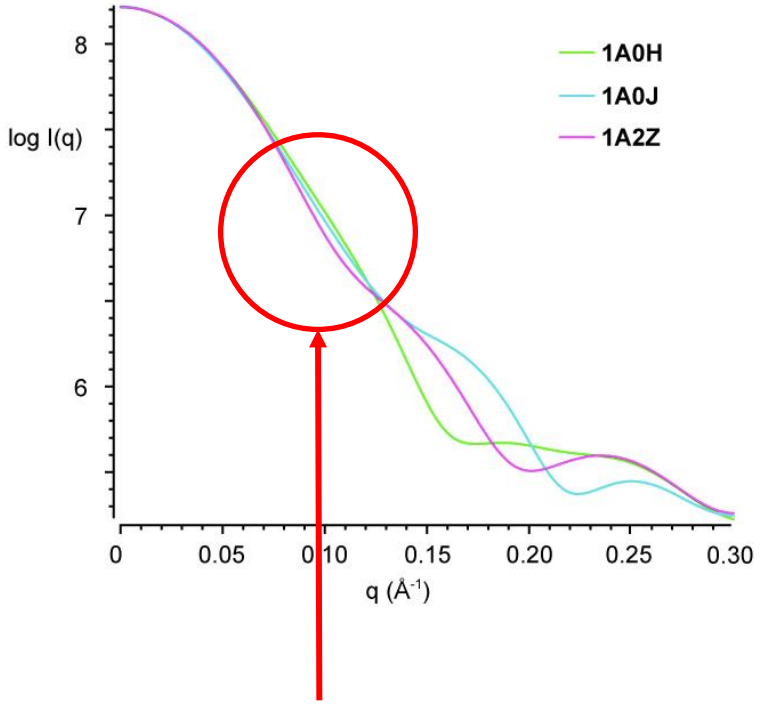
Correct position for 5th chain

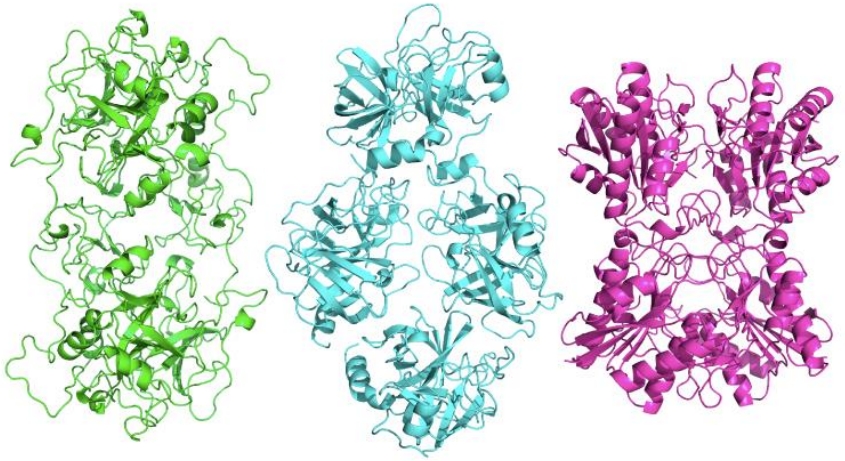


~165A

Biological unit based
on 2izz and SAXS

Structures with very similar radius of gyrations can have very different scattering curves





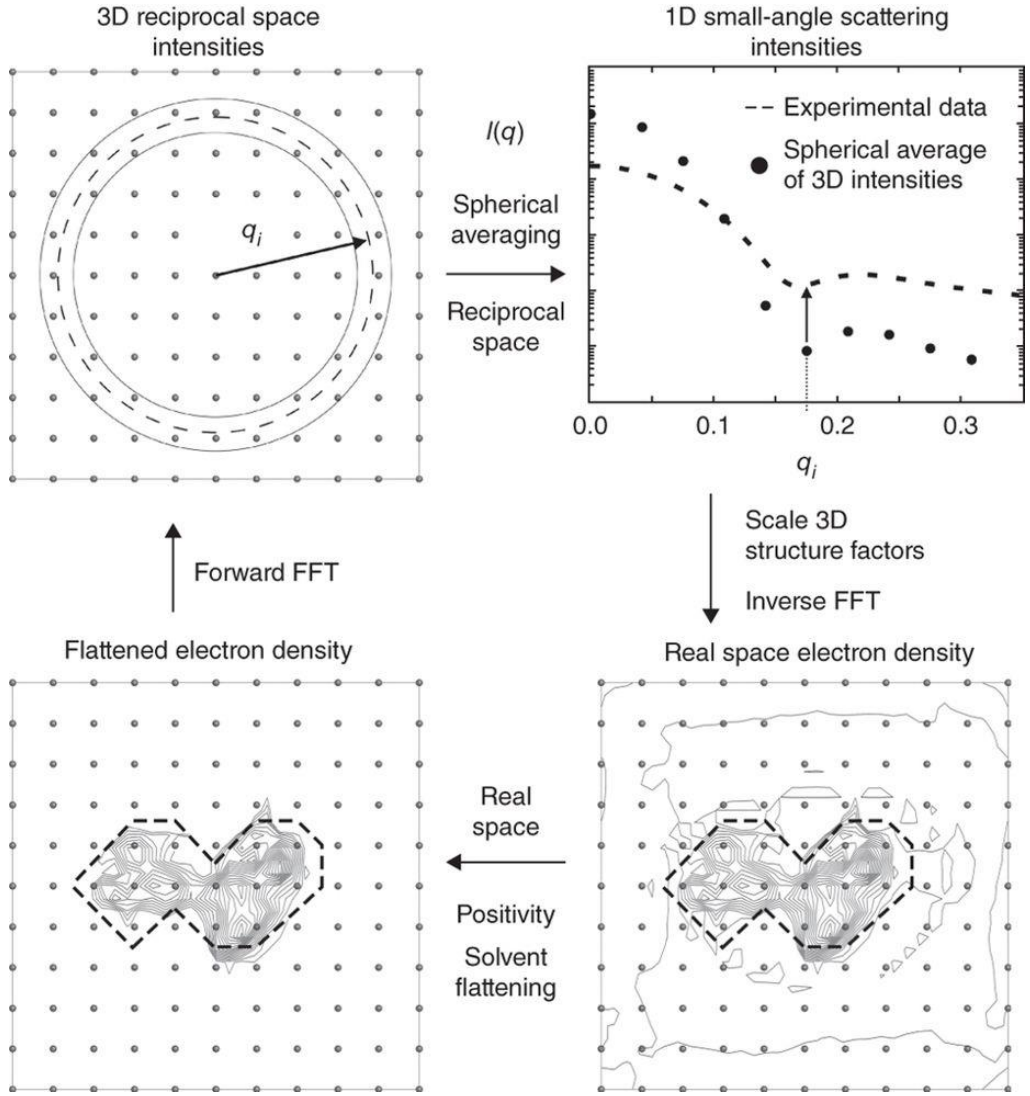
1A0H
MW: 96.8 kDa
Rg: 31.5 Å

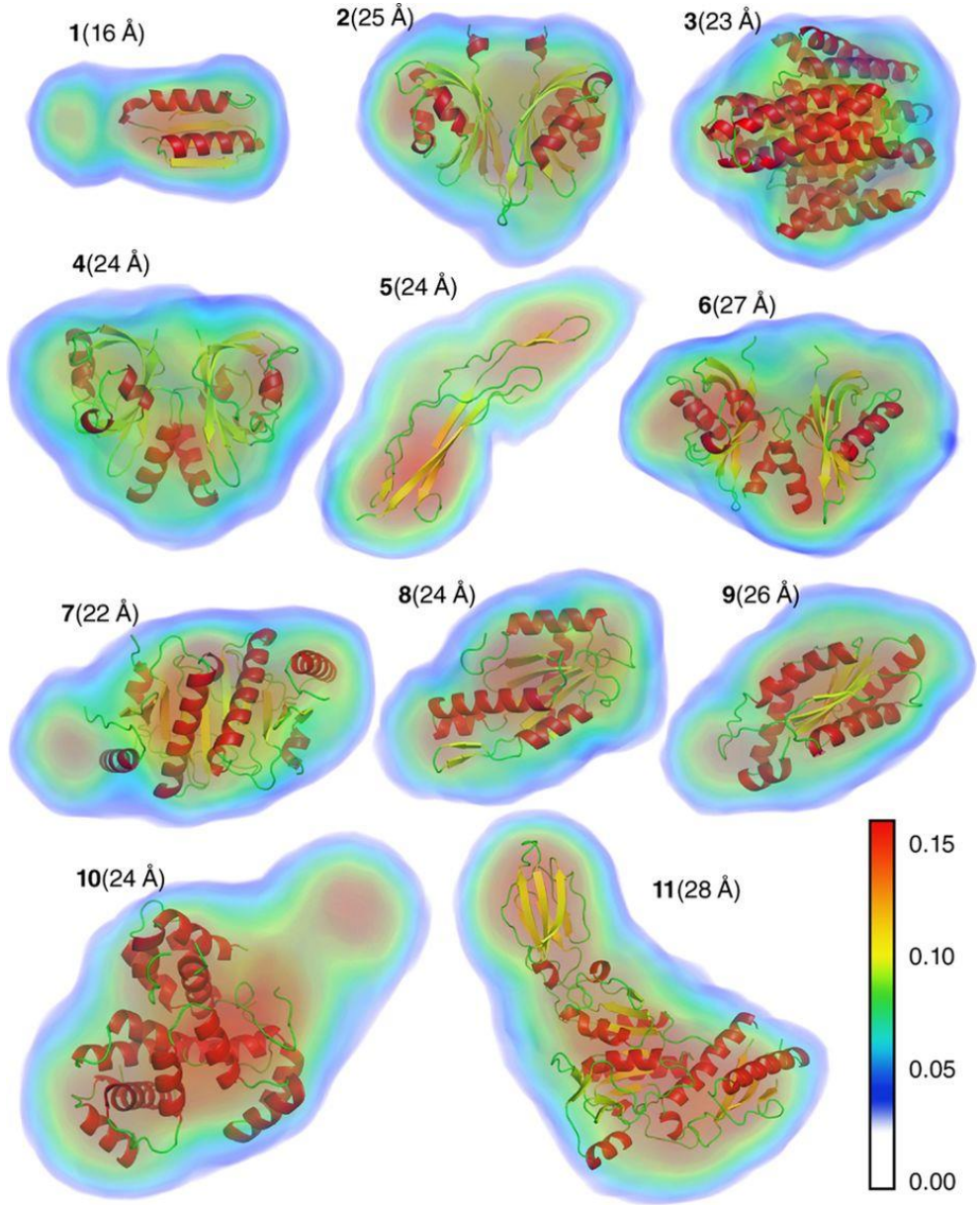
1A0J
MW: 96.6 kDa
Rg: 32.4 Å

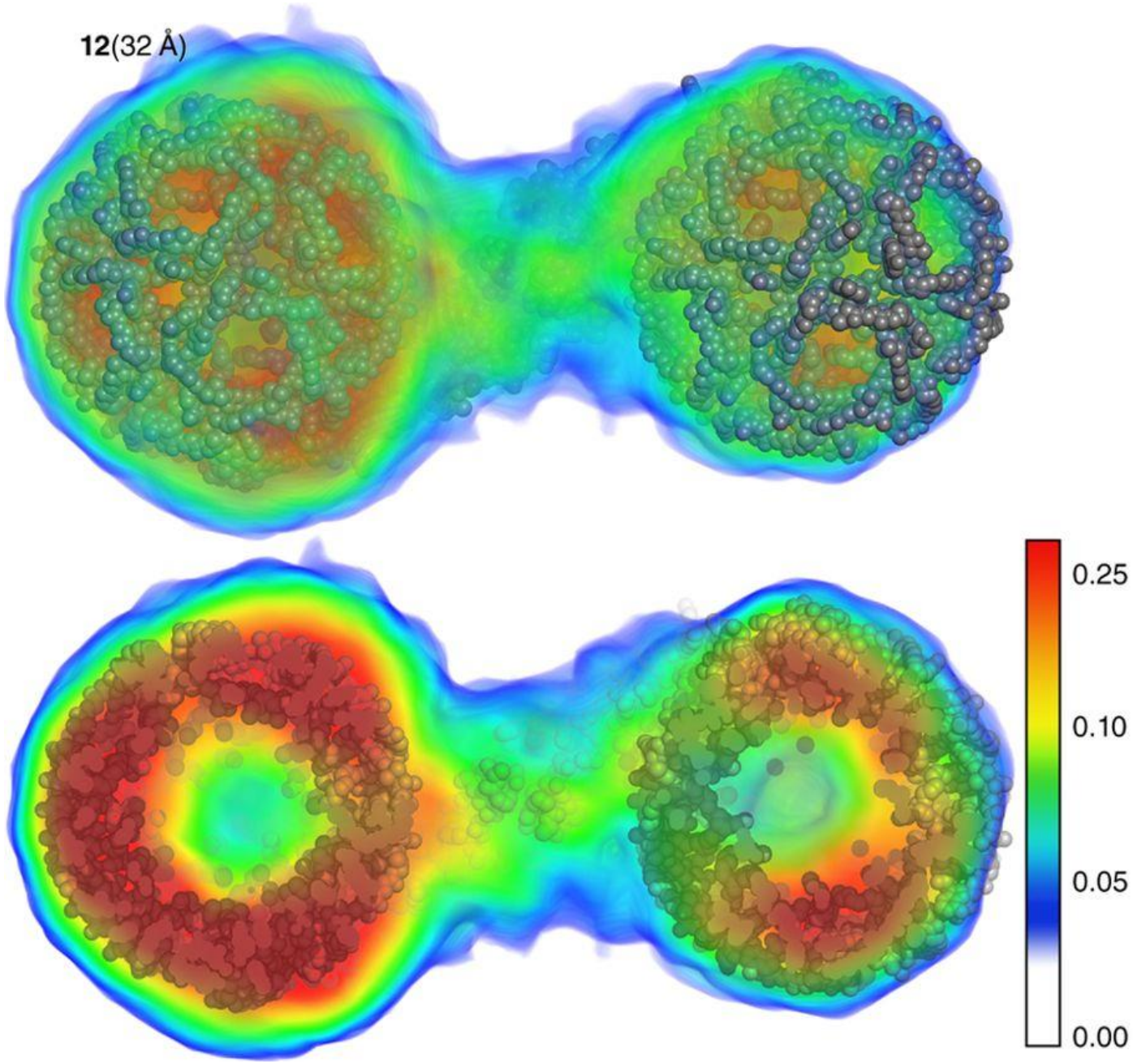
1A2Z
MW: 99.4 kDa
Rg: 30.9 Å

Note that even this is a significant difference in SAXS data

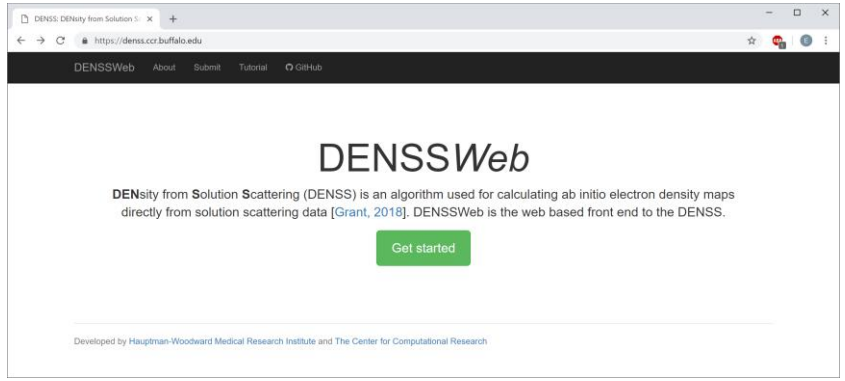
Direct electron density determination

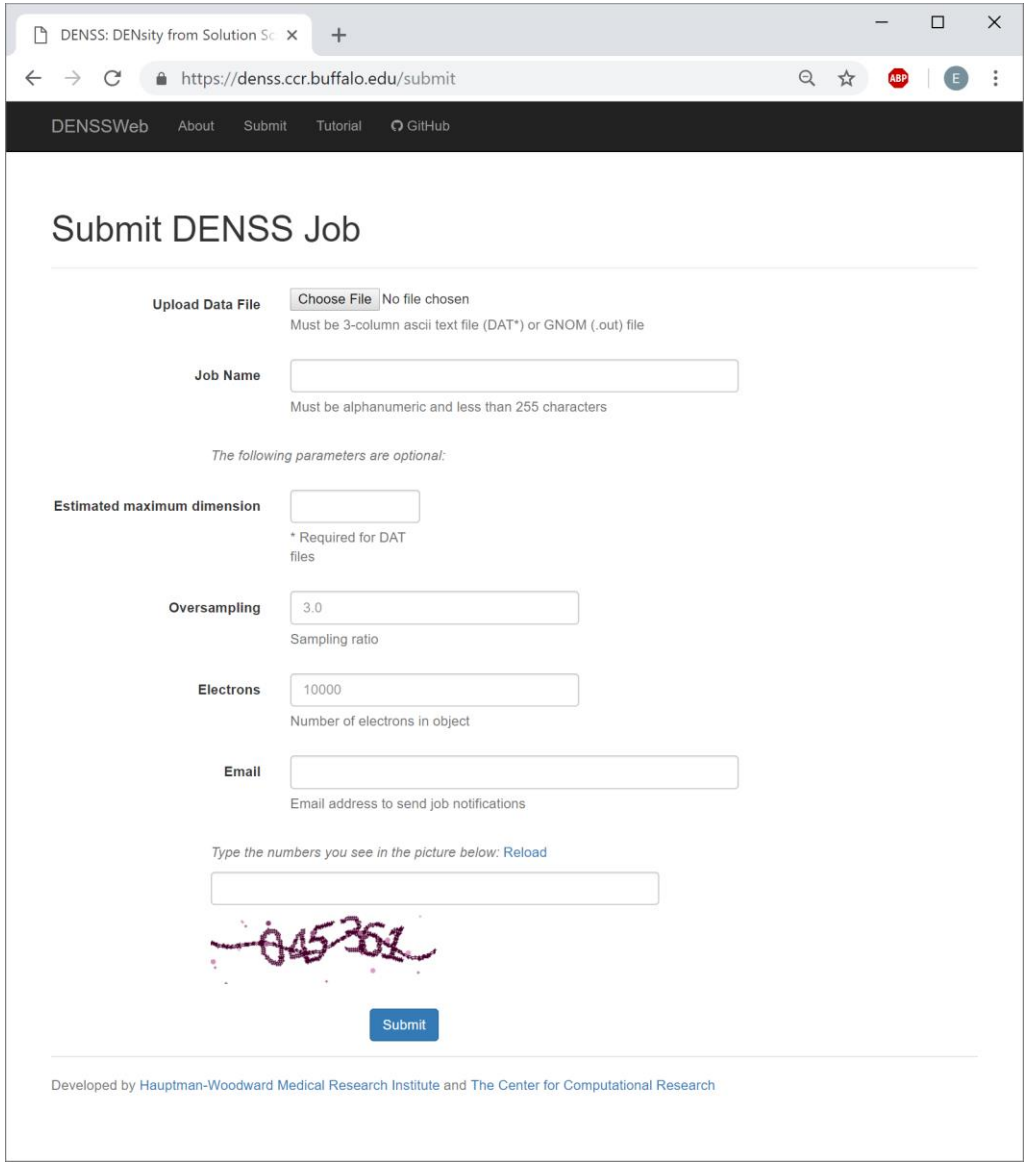

3D reciprocal space intensities
q_i
$I(q)$
Spherical averaging
Reciprocal space
1D small-angle scattering intensities
Experimental data
Spherical average of 3D intensities
0.0
0.1
0.2
0.3
q_i
Scale 3D structure factors
Inverse FFT
Real space electron density
Real space
Positivity
Solvent flattening
Flattened electron density
Forward FFT


1(16 Å)
2(25 Å)
3(23 Å)
4(24 Å)
5(24 Å)
6(27 Å)
7(22 Å)
8(24 Å)
9(26 Å)
10(24 Å)
11(28 Å)
0.15
0.10
0.05
0.00



Electron density map reconstruction of sample 12 endophilin–A1 BAR domain interacting with arachidonyl–CoA micelles from experimental solution scattering data.





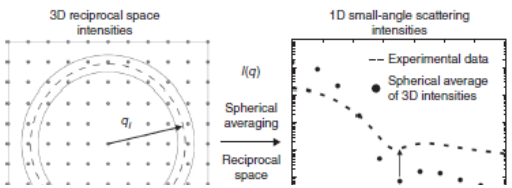
Ab initio electron density determination directly from solution scattering data

Thomas D Grant[1,2]

Using a novel iterative structure factor retrieval algorithm, here I show that electron density can be directly calculated from solution scattering data without modeling. The algorithm was validated with experimental data from 12 different biological macromolecules. This approach avoids many of the assumptions limiting the resolution and accuracy of modeling algorithms by explicitly calculating electron density. This algorithm can be applied to a wide variety of molecular systems.

data rely exclusively on modeling. Many modeling algorithms implicitly assume that electron density is uniform inside the particle envelope[8–12]. This assumption breaks down at resolutions better than 10 Å, where fluctuations in electron density contribute substantially to scattering and for particles containing molecules of different densities or large scale conformational dynamics[13,14]. The ability to resolve multiple densities, high-resolution features and solution dynamics would be facilitated by the calculation of electron density maps directly from solution scattering data.

The electron density of an object can be obtained from its 3D complex-valued structure factor through an inverse Fourier transform requiring both the structure factor amplitude and phase. In a



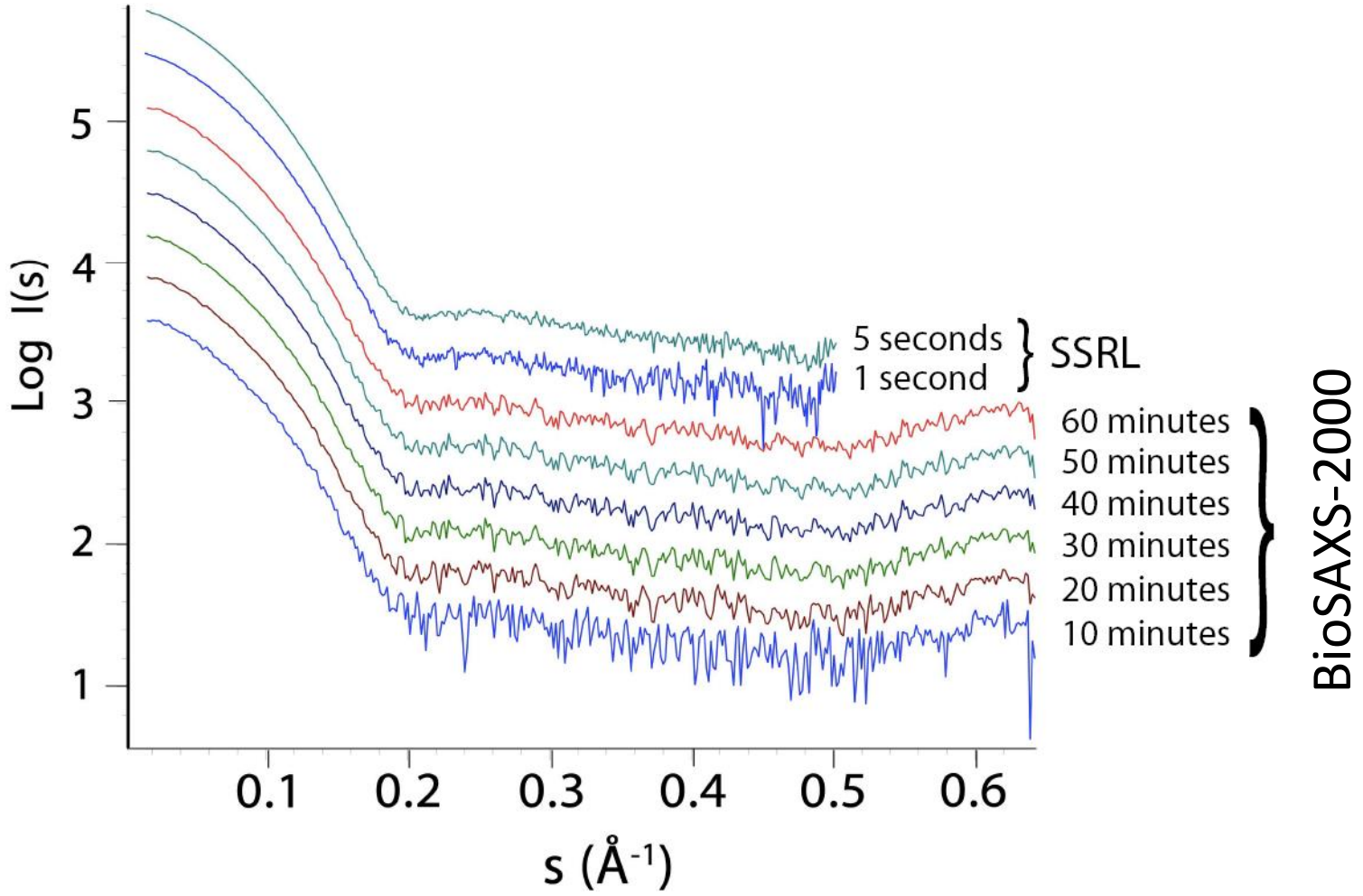
Grant, Nature Methods, 15(3), 191-193 (2018)

http://Denss.ccr.buffalo.edu

SAXS in the laboratory


Rigaku
BioSAXS-2000
Automatic Sample Changer
Rigaku
FR-X
Rigaku
Rigaku

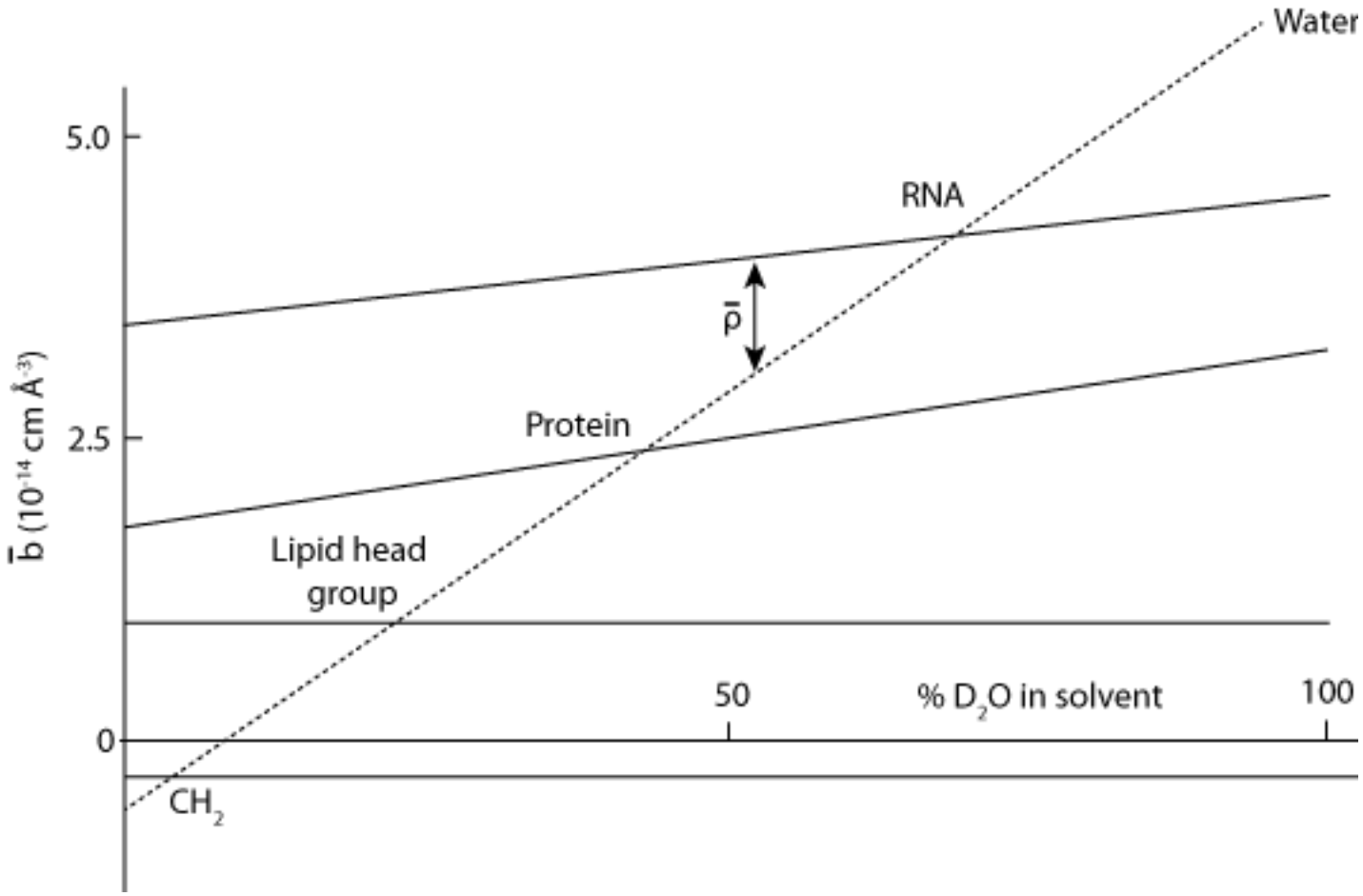
Comparison of synchrotron and laboratory SAXS data

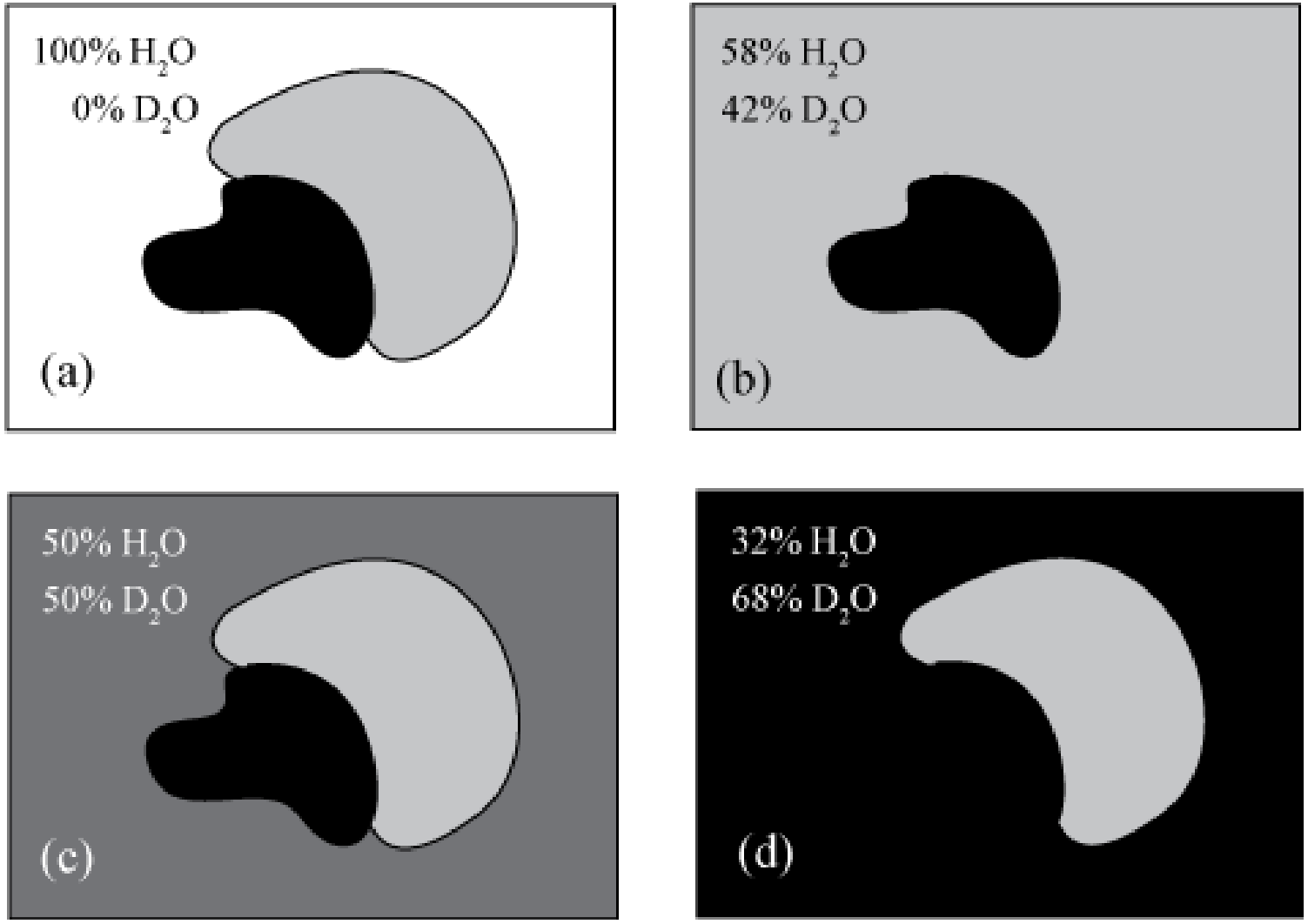


The shape of the scattering curve rather than the absolute value is the data (assuming the signal is above the noise)

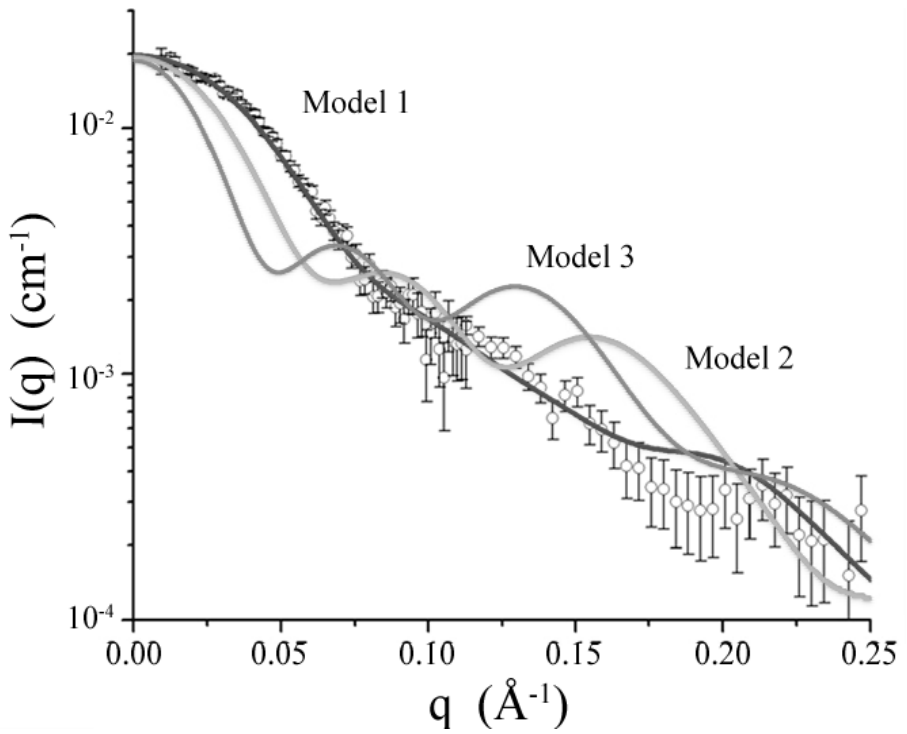
Small Angle Scattering with Neutrons

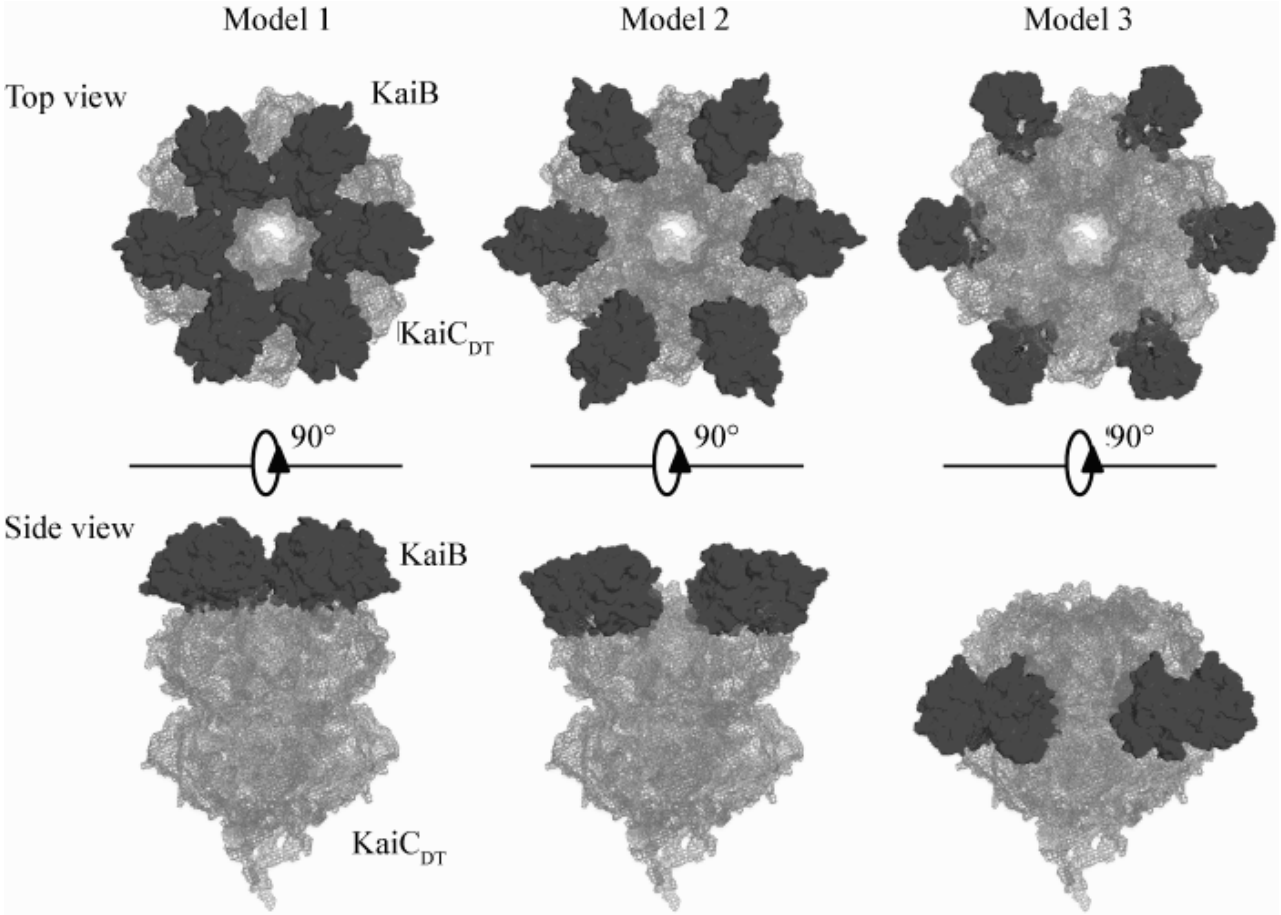
Contrast matching (more difficult in the X-ray case)




100% H_2O
0% D_2O
(a)
58% H_2O
42% D_2O
(b)
50% H_2O
50% D_2O
(c)
32% H_2O
68% D_2O
(d)

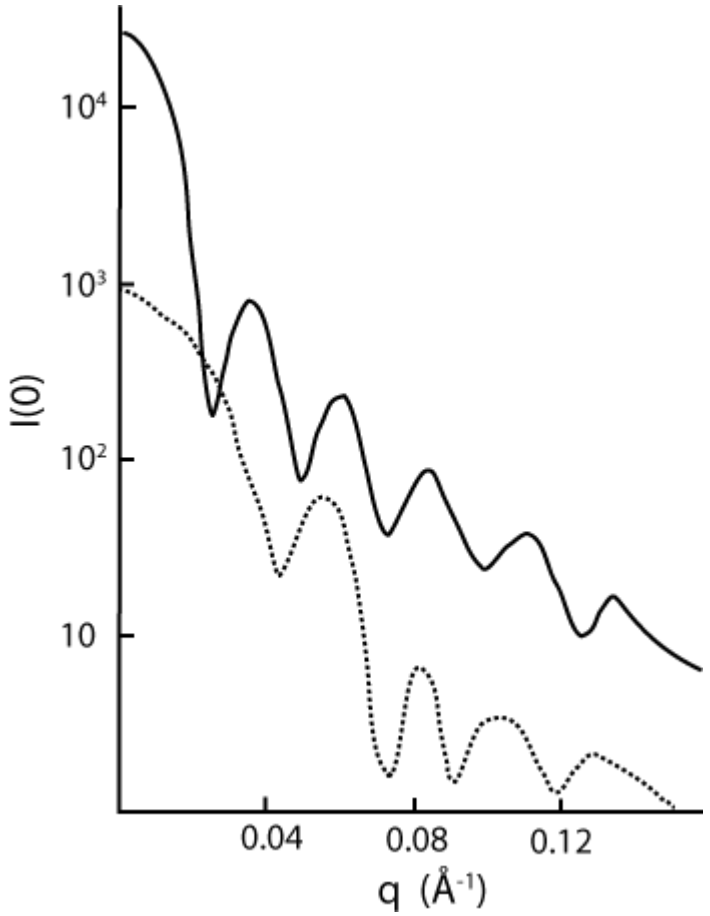
A deuterated sample can provide contrast from a non-deuterated sample



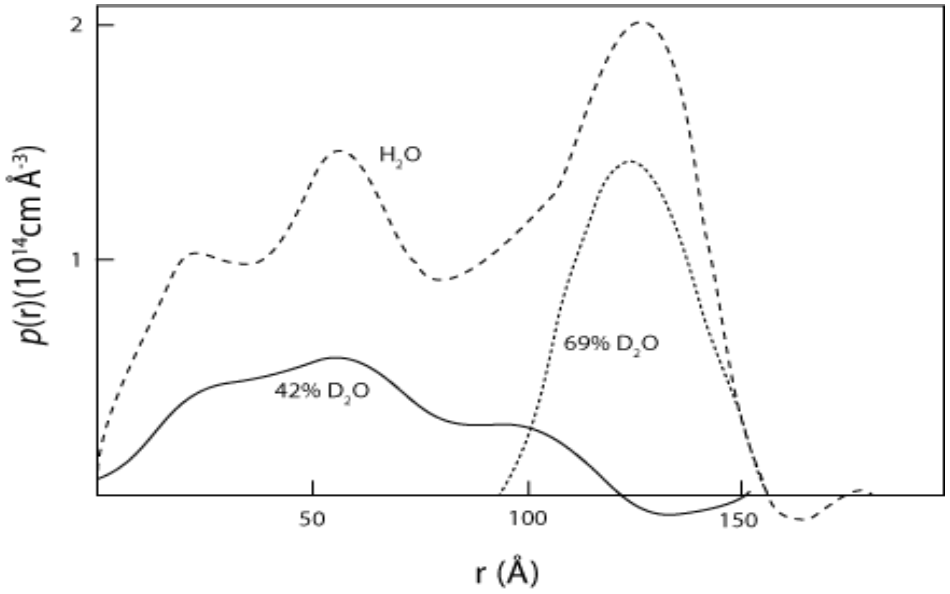


The central oscillator that generates the circadian rhythm in the cyanobacterium Synechococcus elongatus comprises three proteins KaiA, KaiB, and KaiC4. The KaiB–KaiC interaction is a key event during oscillation of the cyanobacterial circadian protein system (Sugiyama et al., 2016)

Classic application – understanding the structure of viruses



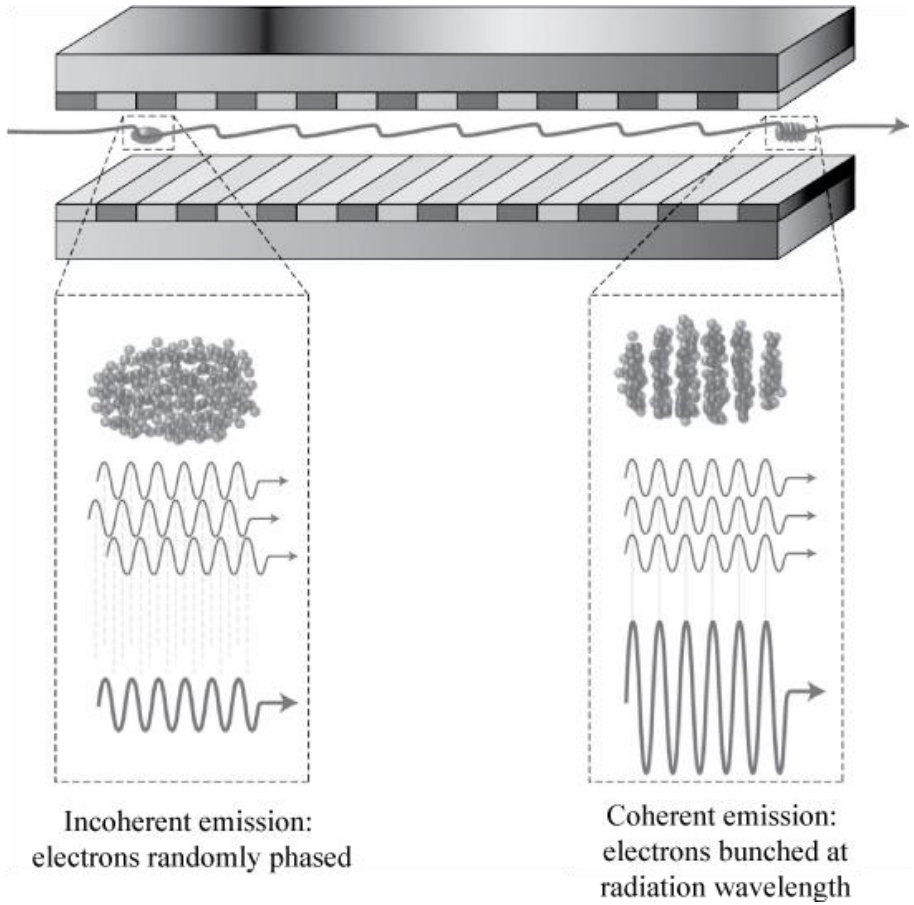
Scattering curve from Southern Bean Mottle virus in solutions of different D_2O content. The continuous line with 69.5% D2O and scattering mostly due to the protein shell and the dashed line with solvent content 42% D_2O and the scattering mostly by the nuclein acid (RNA). The subsidiary maxima are shifted to a larger q which indicate that the sphere that approximates the volume occupied by the RNA has a smaller diameter than the virus (Chauvin et al., 1976).



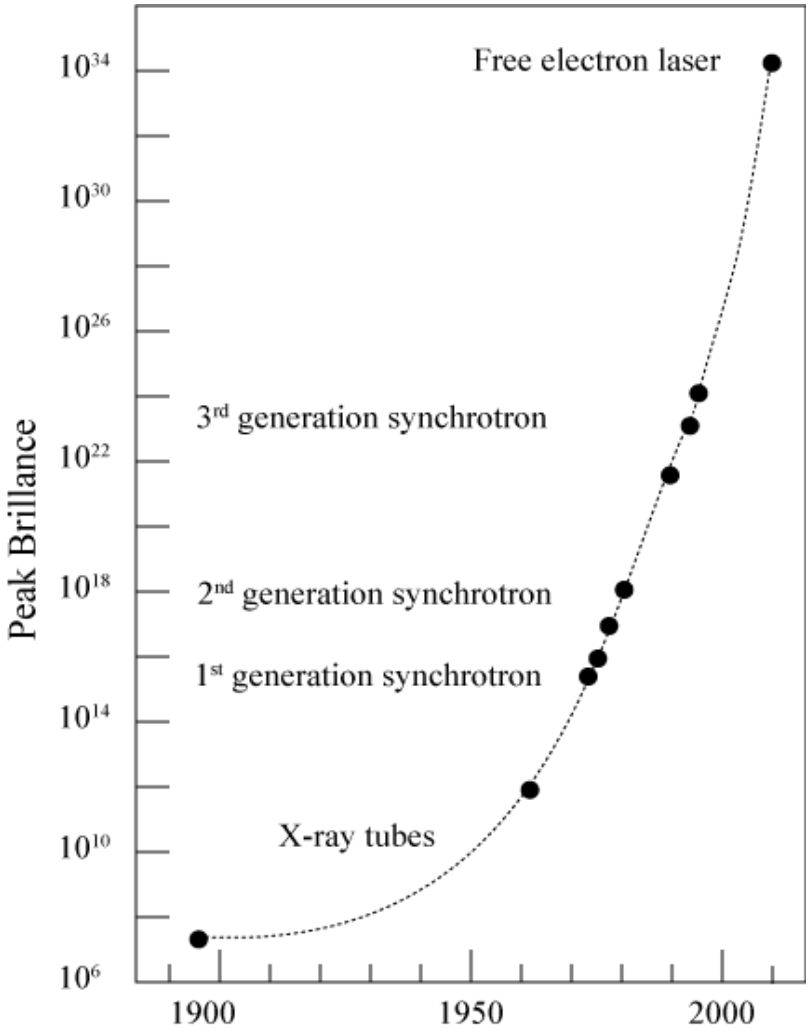
The near future

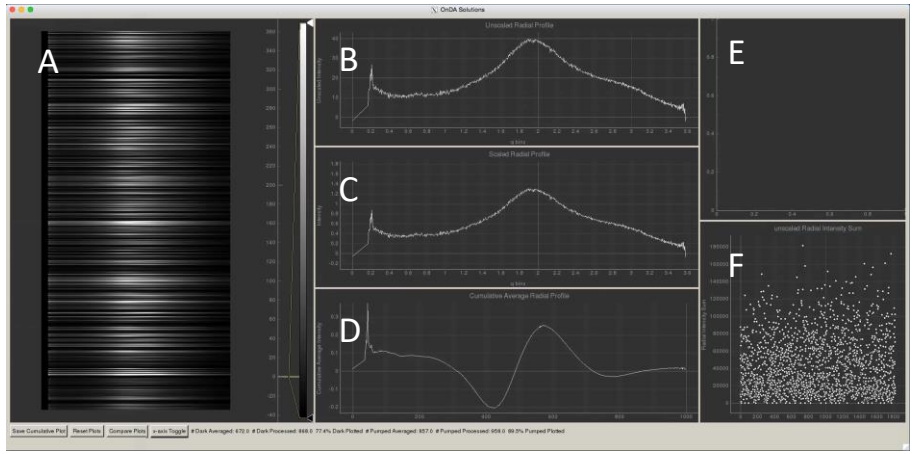
The Linear Coherent Light Source XFEL

X-ray free electron lasers



SAXS experiments at XFELS capture short time points at low concentrations



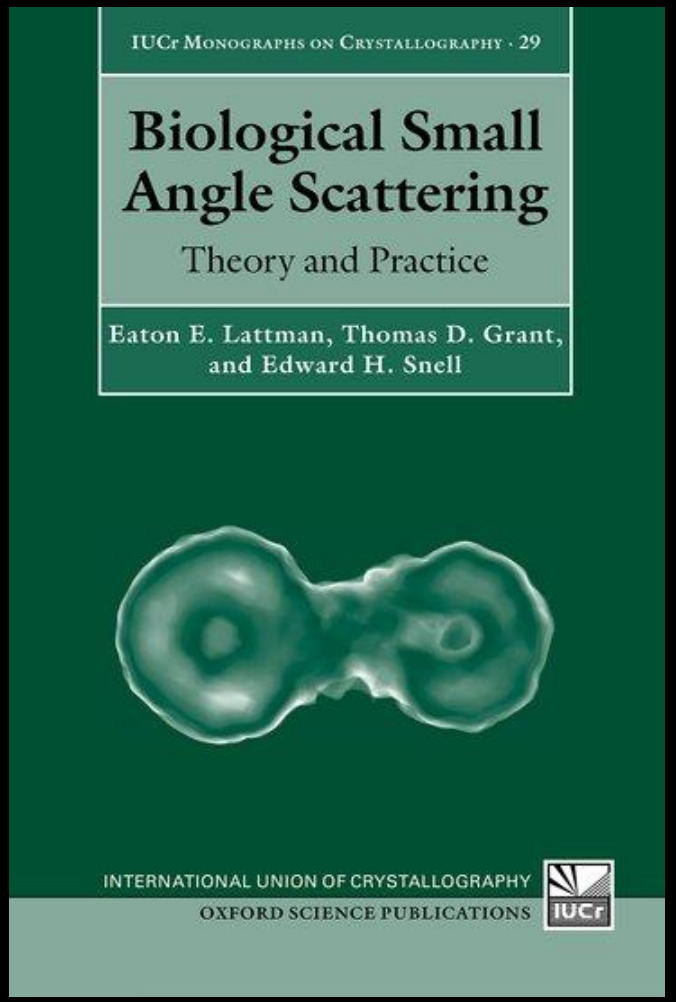


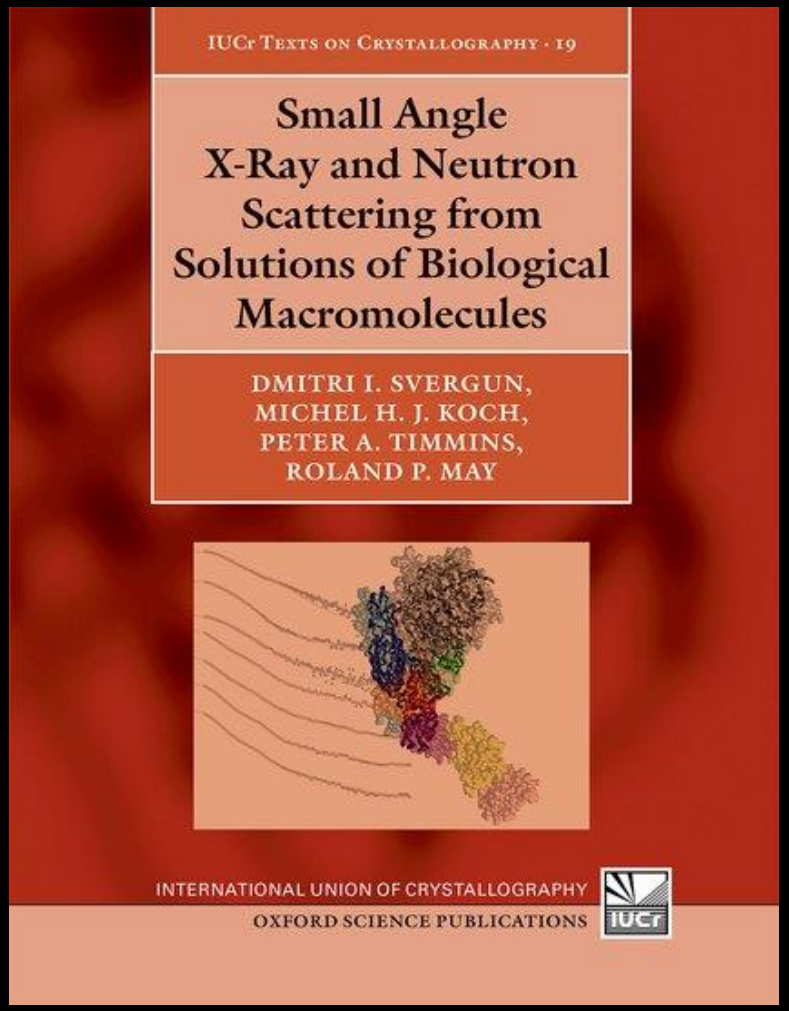
Time resolved studies and potentially extension of resolution

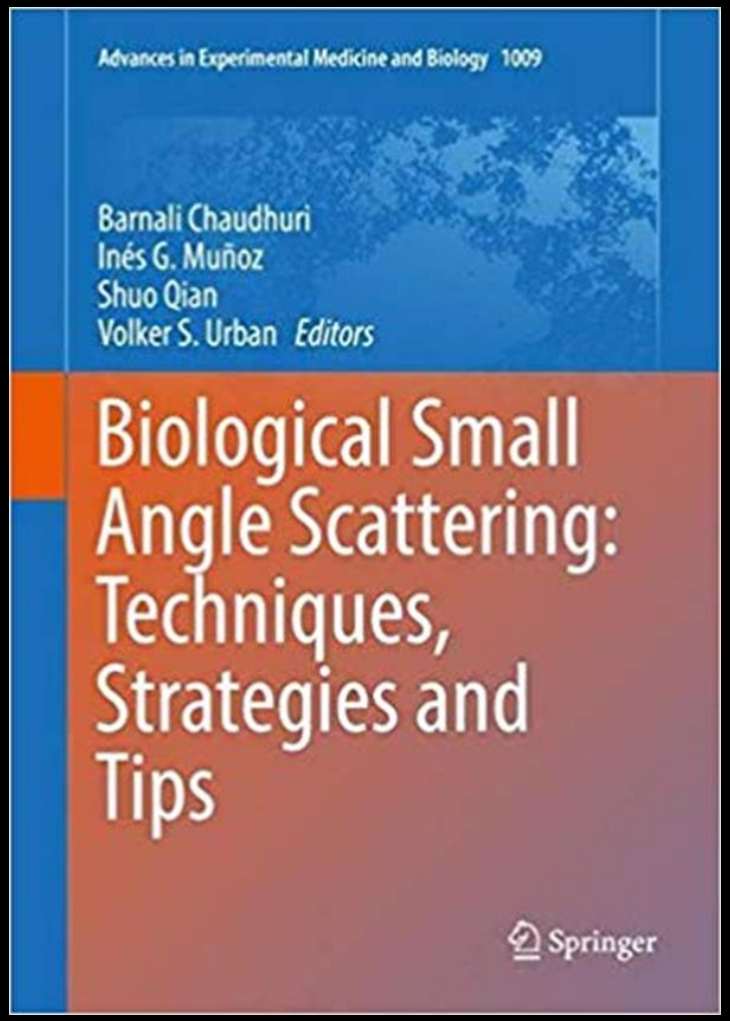
Summary

- SAXS is a solution technique.
- It can characterize a sample to determine if crystallization should be attempted and the potential level of difficulty.
- When other structural information is known it is a powerful complementary technique.
- It can reveal the solution oligomer and the spatial sampling of flexible regions.
- It’s easy to make mistakes with it and preparation is critical.
- It can be done in parallel with crystallization.
- **It can be done under physiological conditions.**

Good books are available







- Excellent practical guide - X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution Putnam et al Quarterly Reviews in Biophysics, 2007

Thank you and questions?

esnell@hwi.buffalo.edu