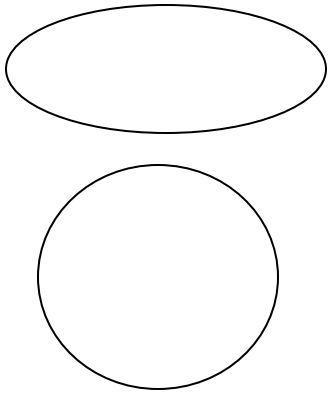
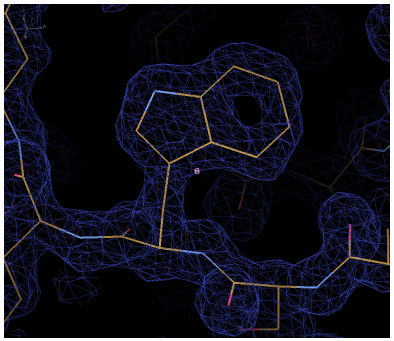
Biological Small Angle X-ray Scattering (SAXS)

March 10, 2014



Structural Biology



Shape

Atomic Detail

Dynamic Light Scattering

Electron Microscopy

Cryo-Electron Microscopy

X-ray Crystallography

Small Angle X-ray Scattering

Wide Angle X-ray Scattering

X-FEL?

Why bother with low resolution?

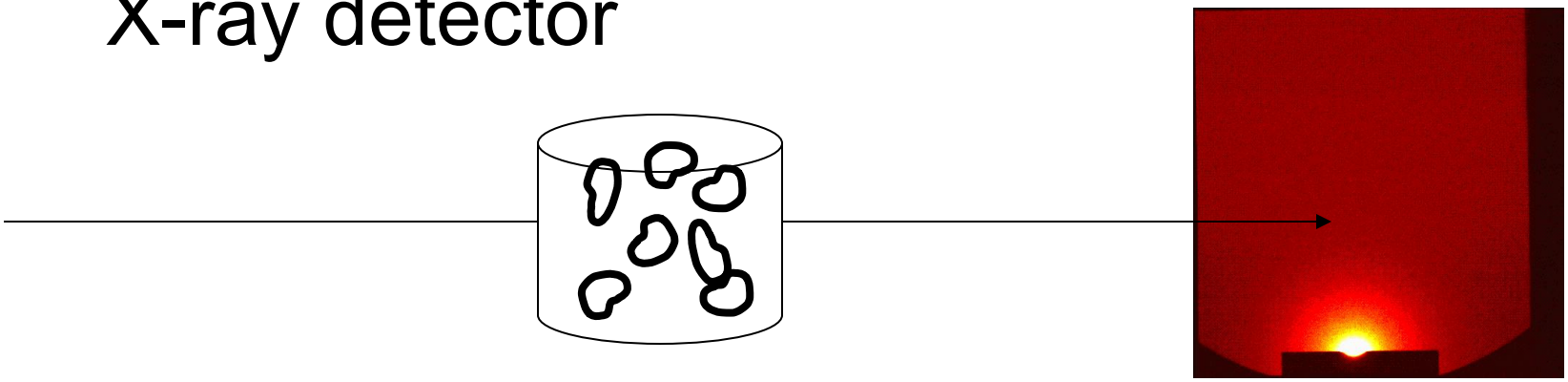
- Crystallization not possible
- Too large for NMR
- Dynamic system
- Time resolved studies

Questions that SAXS can address

- What is the effective association state (is it a monomer, a dimer . . .) of a protein or other macromolecule in solution or within a membrane environment?
- What is its shape or conformation (is it compact and globular or ellipsoidal, long and narrow, or flat and broad, star-shaped or branched . . . is its structure in solution similar to its crystal structure . . .)?
- Do different macromolecules in solution interact to form a complex or not?
- How the complex vary as a function of solvent conditions (pH, salt, ligand, temperature . . .)? Are there modifications in association state or conformational changes?

What is it?

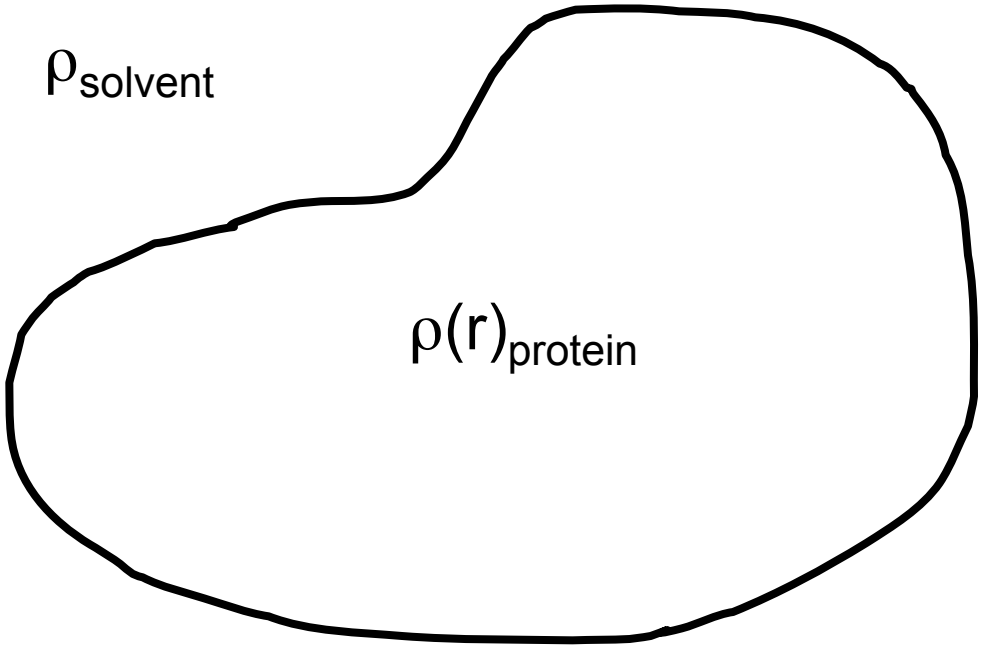
- X-rays are shot through a solution of protein
- The protein scatters the incoming X-rays
- The scattered X-rays are collected on a X-ray detector



Information content in scattering

- Solution scattering deals with proteins that are randomly orientated in solvent
- Three dimensional information compressed to one dimension
- A one dimensional self-correlation function of the protein is obtained from the Fourier transform of the scattering function

Electron density contrast

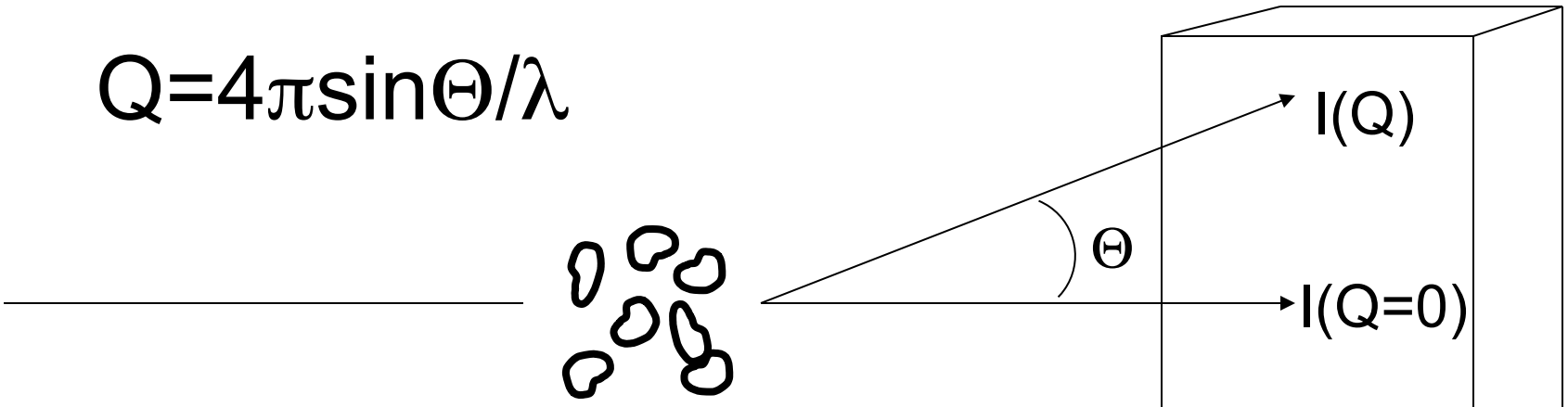


$$\Delta\ \rho(r) = \rho(r)_{protein} - \rho_{solvent}$$

Q?

- Describes scattering angle independent of wavelength so multiple experiments can be compared.

$$Q=4\pi\sin\Theta/\lambda$$



λ = wavelength of X-ray

Θ = angle from incident beam

X-ray Detector

Q=0?

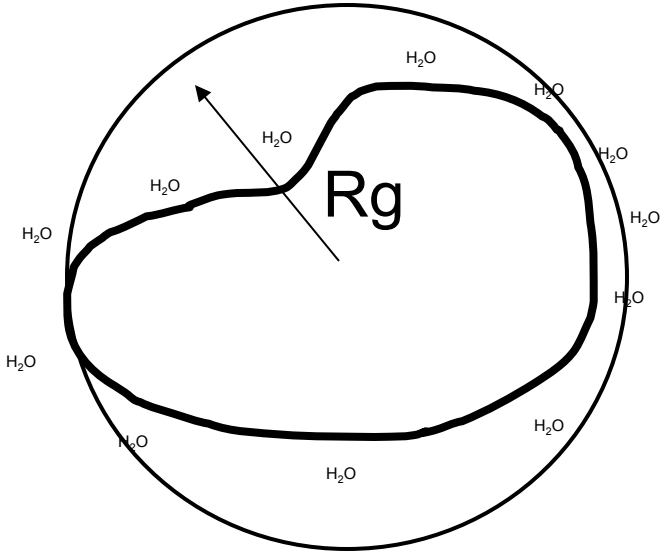
- Proportional to the square of the molecular mass of the macromolecule at a defined concentration

Physical parameters

- Since scattering can be mathematically modeled certain physical properties can be determined by scattering intensity
- Directly measurable using SAXS
  - Radius of gyration
  - Particle volume
  - Molecular weight
  - Distance distribution function $\rho(r)$

Calculation of radius of gyration

Radius of gyration is the root-mean square of the distance of all electrons from the center of gravity

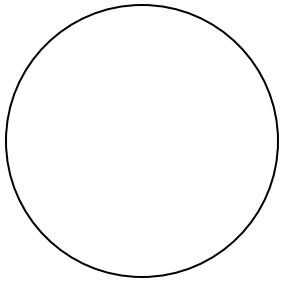


$$I(Q)=I(Q=0)\exp(-Q^2R^2_g/3)$$

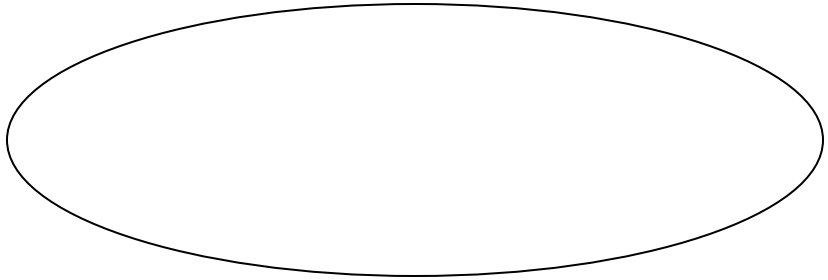
Where $Q=4\pi\sin\Theta/\lambda$

Molecule shape

- $R^2_g = (3/5)R^2$ for spheres

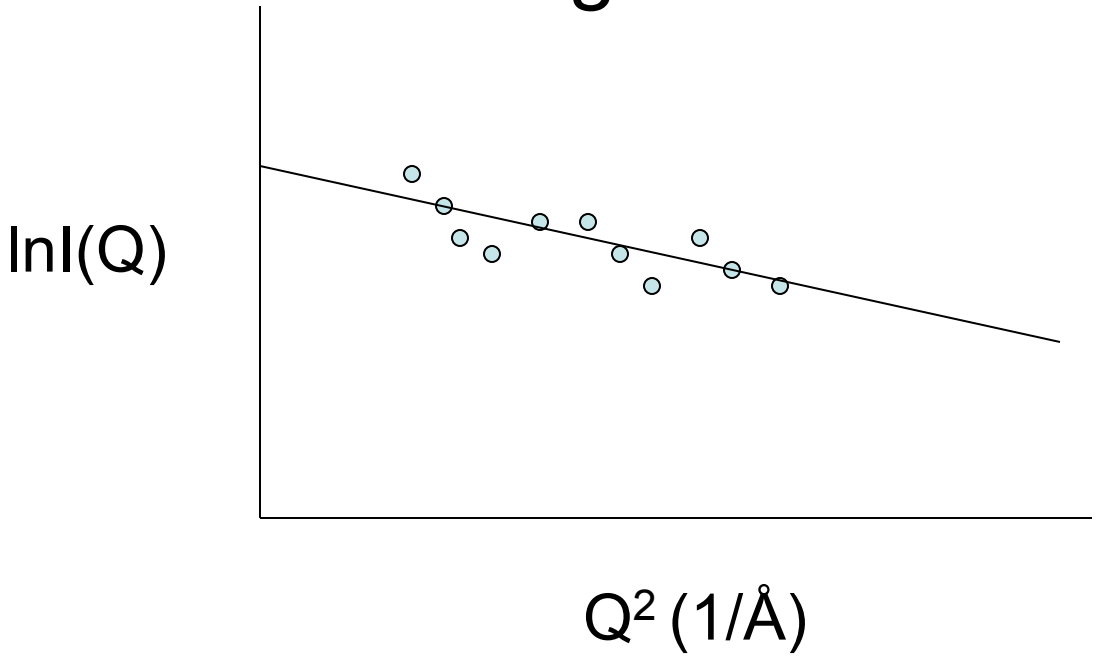


- $R^2_g = (a^2 + b^2 + c^2)/5$ for ellipsoid



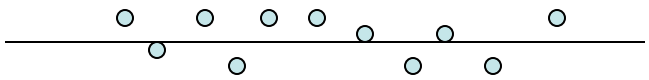
Guinier Analysis

- Radius of gyration
- Titration curves
- Molecular weight estimate

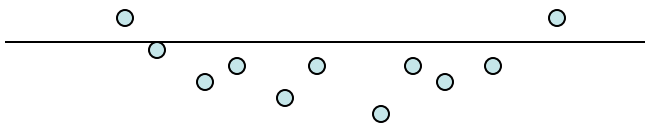


Guinier Analysis (cont.)

- Guinier plot is useful for detecting sample problems
  - Poly-disperse sample
  - Extreme deviation from globular shape
- These defects are more obvious by examination of residuals
  - Plot of deviation from linearity



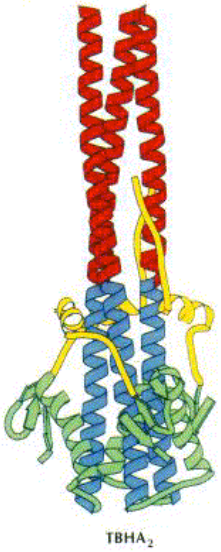
Normal Distribution



Suspicious Distribution

Why do I care about radius of gyration?

- Suppose you wanted to know if a protein dimerized in solution
- Does the radius of gyration differ from what would be predicted?
  - Protein may be elongated
- Test conditions to alter the Rg
  - Addition of ligands
  - Physical properties of solution
    - pH, ionic strength

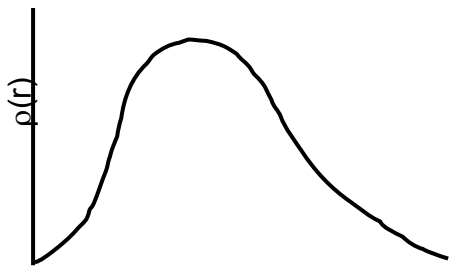


Distance distribution function $\rho(r)$

- Histogram of distances of all possible electron pairs
- Fourier inversion of scattering curve

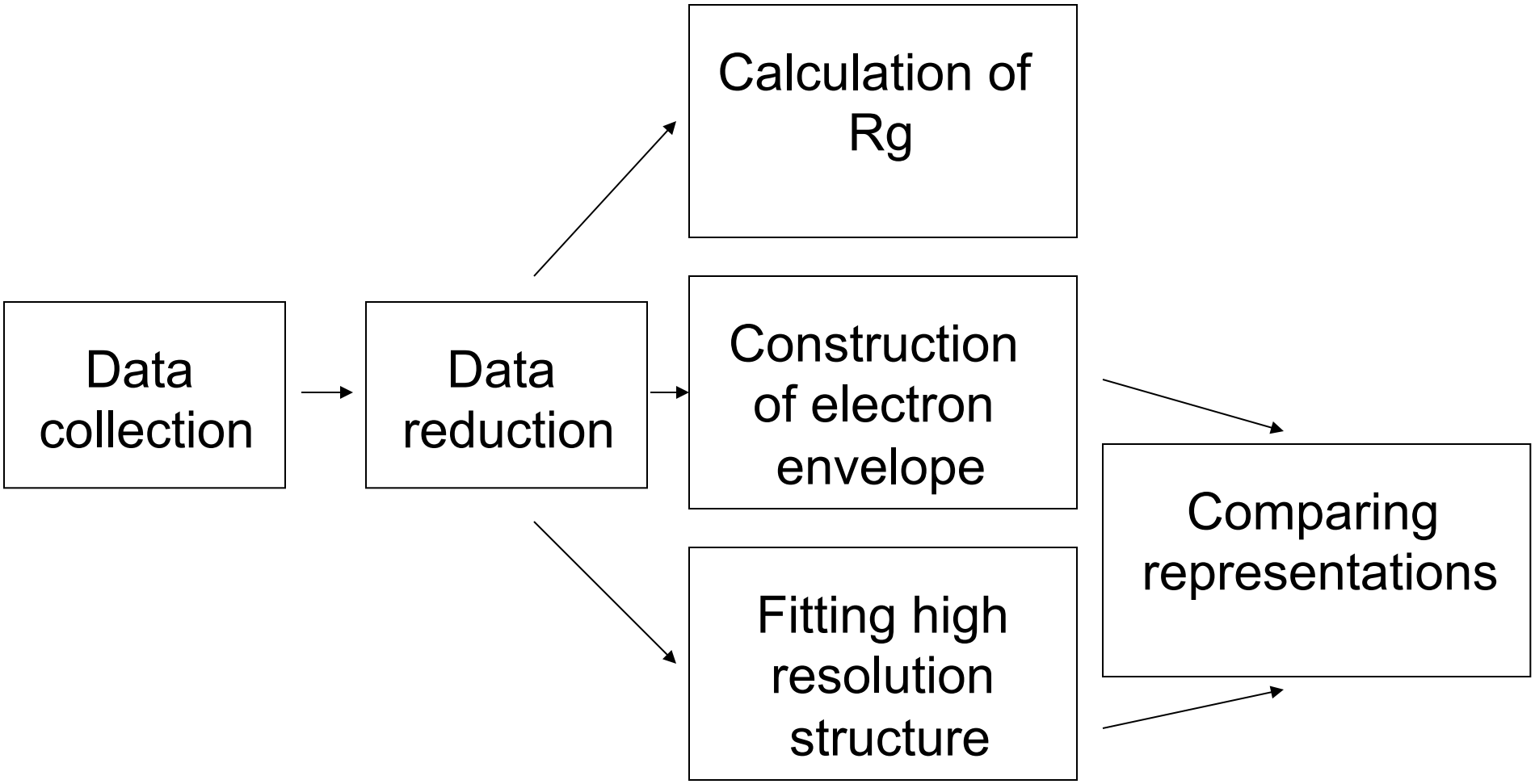
$$\rho(r) = 1/2\pi^2 * (\int_0^\infty I(h) * hr * \sin hr * dh) \qquad h = 4\pi \sin\Theta/\lambda$$

- Allows calculation of largest particle dimension



Electron pair distance distribution (Å)

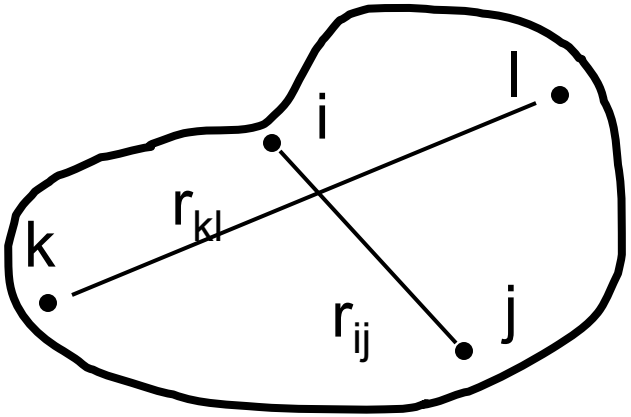
Construction of 3D representation

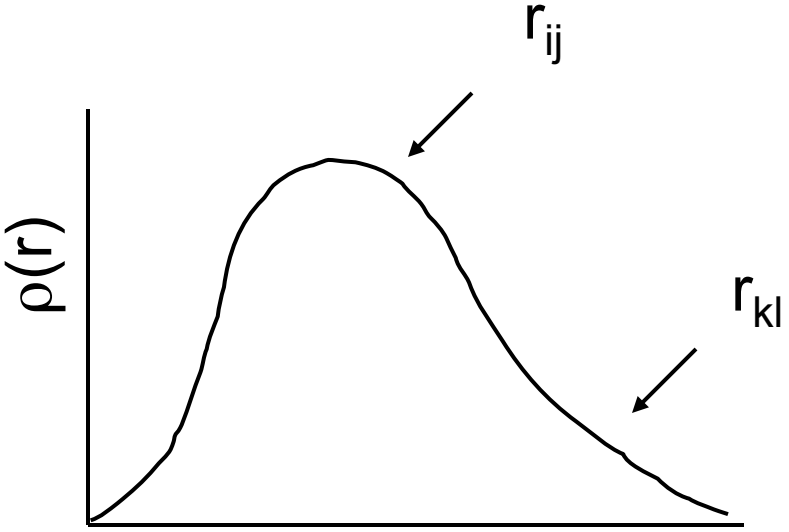


Structure using SAXS

- Calculation of electron density envelope
  - Requires accurate determination of maximum radius
- Orientation of known atomic structures
  - Individual domains of large proteins are easier to solve the structures of then the full molecule
  - Using SAXS and biochemical knowledge the domains can be orientated

Determining the shape of a molecule





Electron pair distance distribution (Å)

$$I(Q)=\sum_i\sum_j \Delta\rho(ri)\ \Delta\rho(rj)\sin[Q(ri-rj)]/Q(ri-rj)$$

Molecule shape and distribution

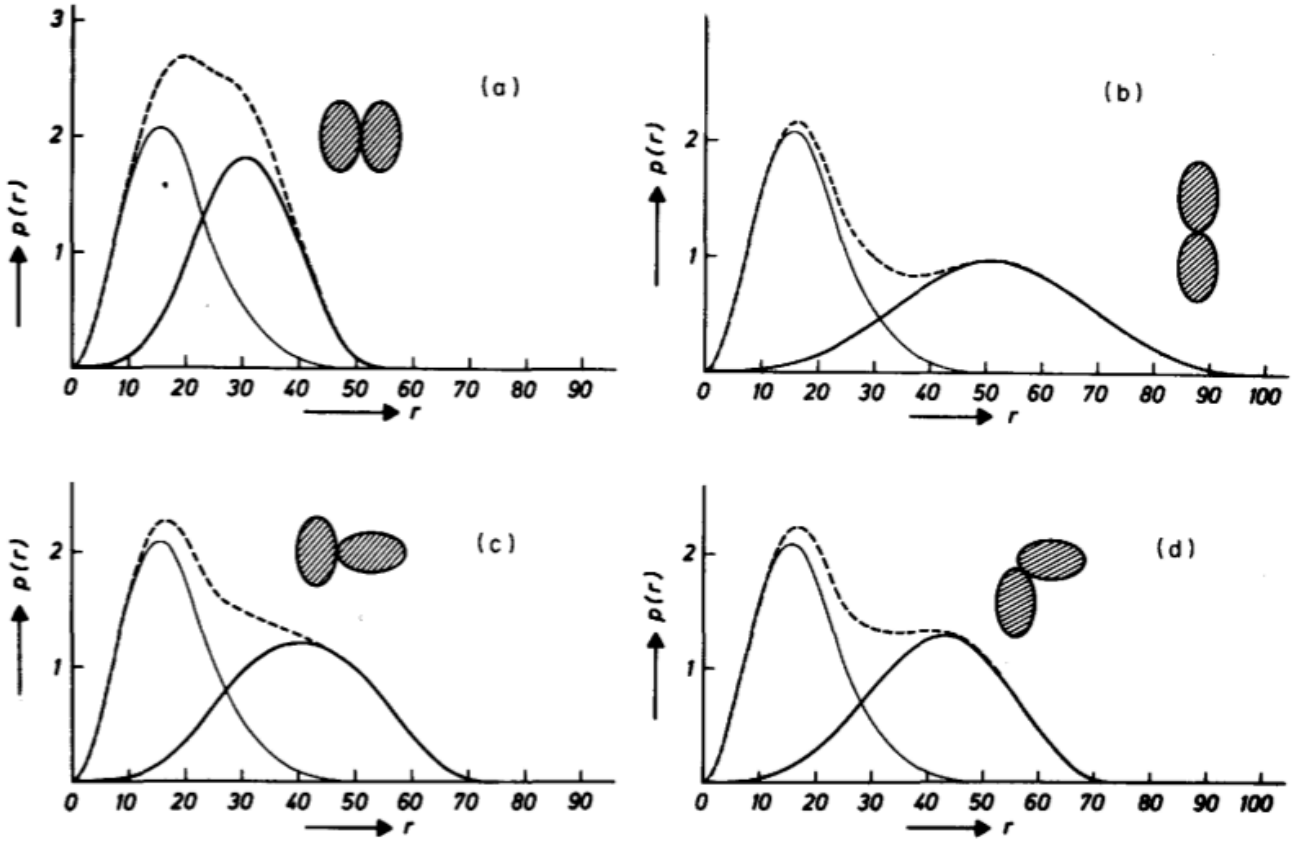


FIG. 9. Distance distribution function $p(r)$ from dimer models built from prolate ellipsoids. —— monomers, – – – – dimers, —— difference between dimers and monomers. (a) parallel formation, (b) linear formation, (c) T-type, (d) L-type.

Molecule shape and distribution

- Trial and error process to determine model that agrees with the data
- This doesn't prove that the model is correct, only that it is equivalent

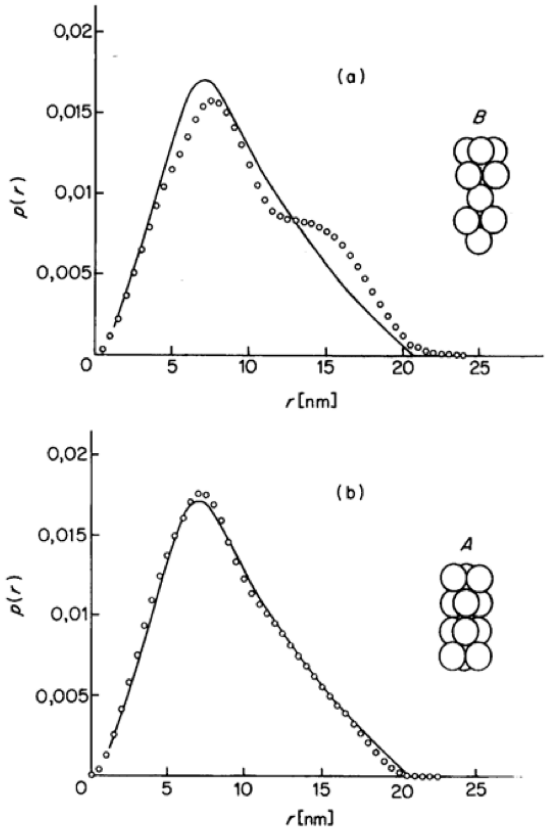


FIG. 9. Distance distribution function $p(r)$ of the haemocyanin of *Astacus leptodactylus* (○○○) compared with the calculated $p(r)$-function of (a) model *B* and (b) of model *A*.

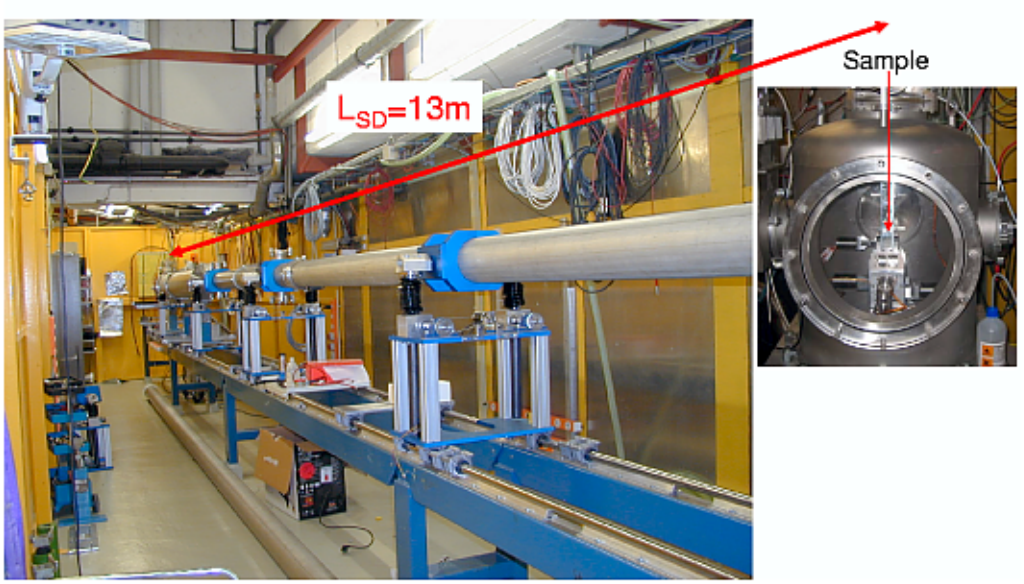
Data Collection

- Mono-disperse protein
  - Filter
  - Centrifuge
  - >1 mg/ml
- Matching buffer for blank
  - Preferably from the same sample
- Addition of agents to decrease radiation damage

Where can you collect this data?

- Synchrotrons
  - Brookhaven National Laboratory (NSLS), Long Island
  - Cornell High Energy Synchrotron Source, Ithaca
  - Advanced Photon Source, Chicago
  - ESRF, Grenoble
- Small X-ray generators
  - Rigaku
  - Bruker

Data Collection

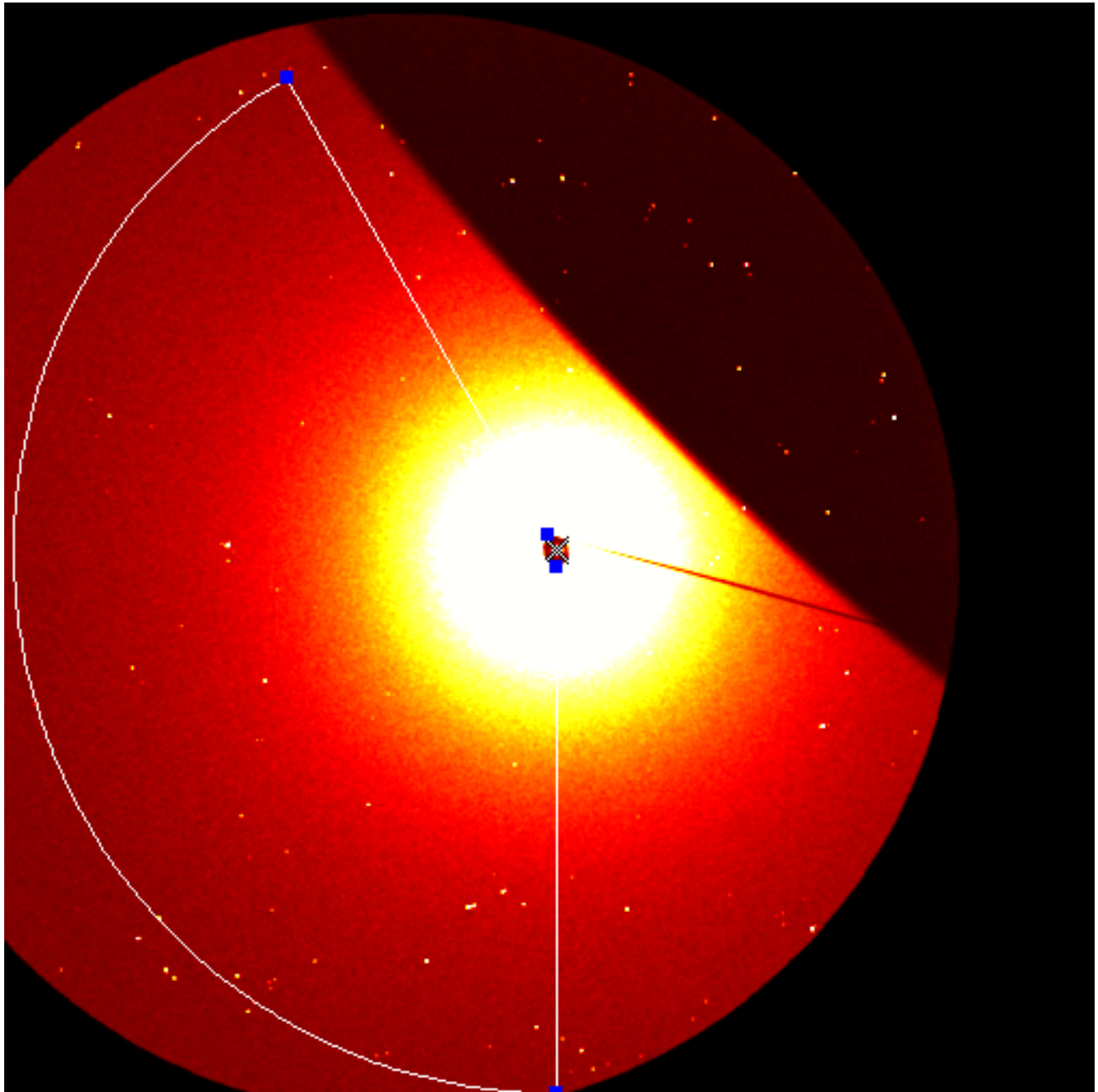


- Sample positioned at end of long tube which is under vacuum
- Images are collected with exposures of 10s-180s

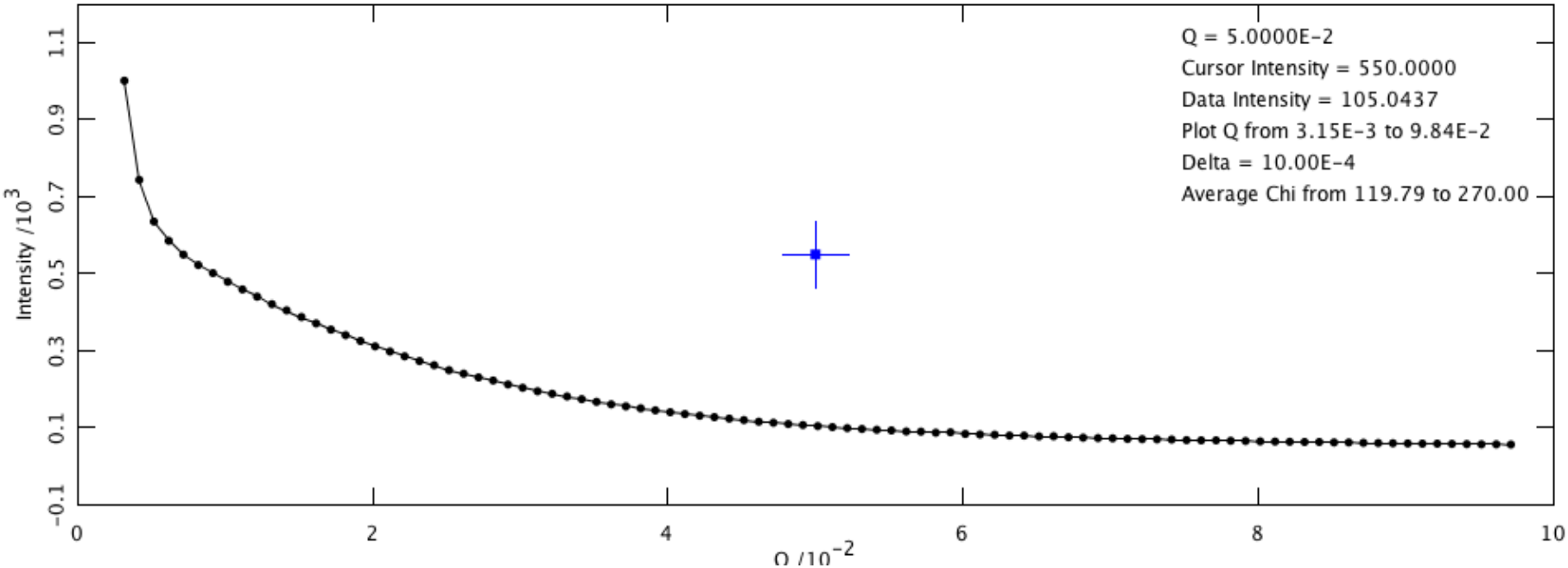
Processing of Images

- Several software programs are available
  - Datasqueeze
  - View.gtk
- Average data in arcs and store in bins by Q
- Subtract buffer data

Image with false color

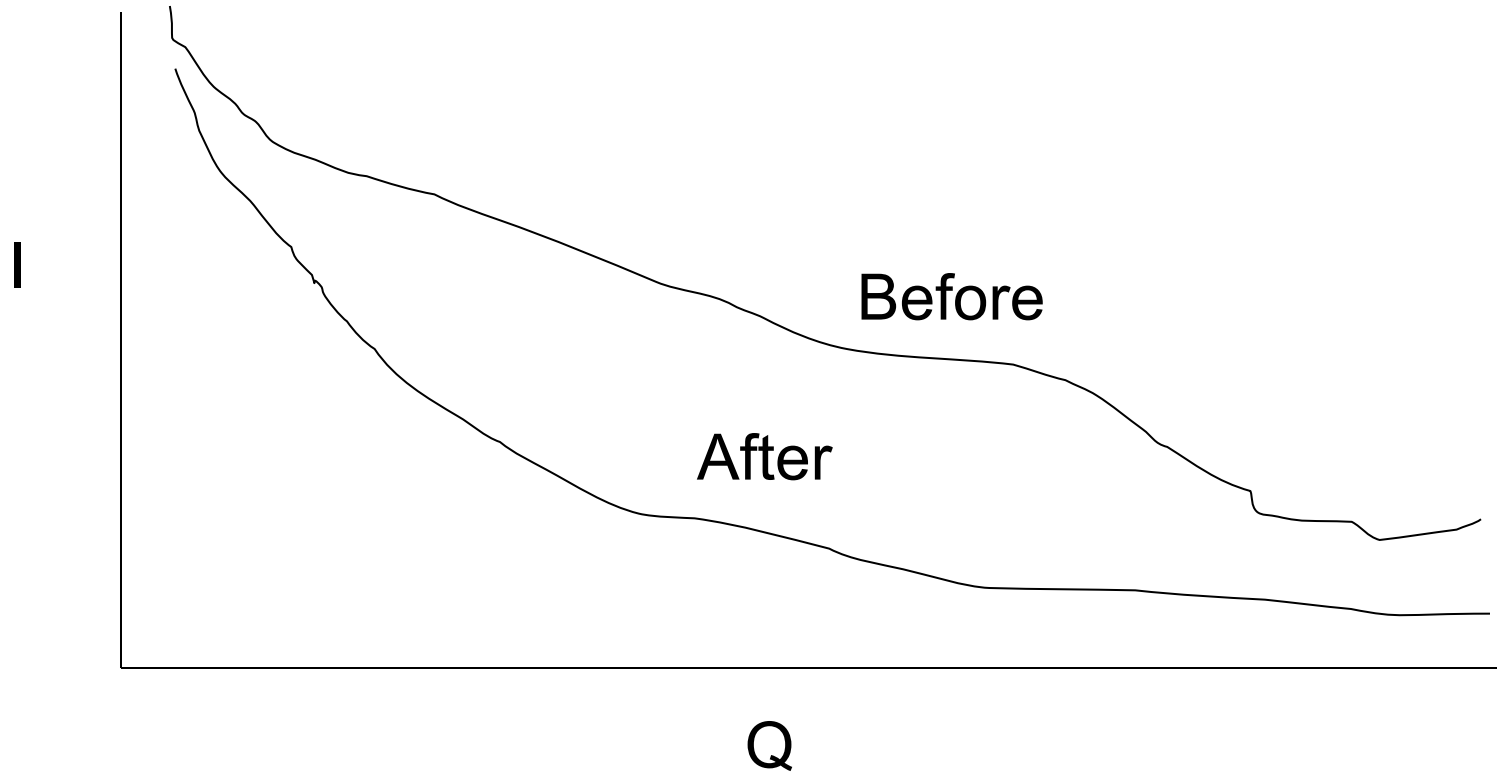


Average intensity plot verse Q



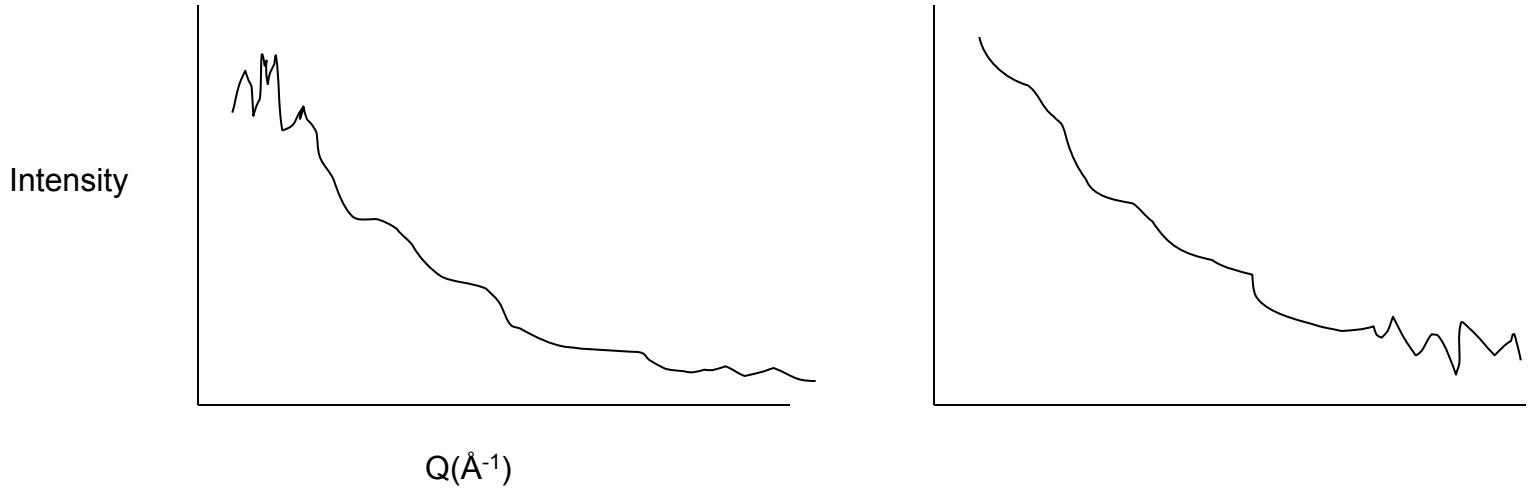
Subtract Buffer

Scattering only from protein



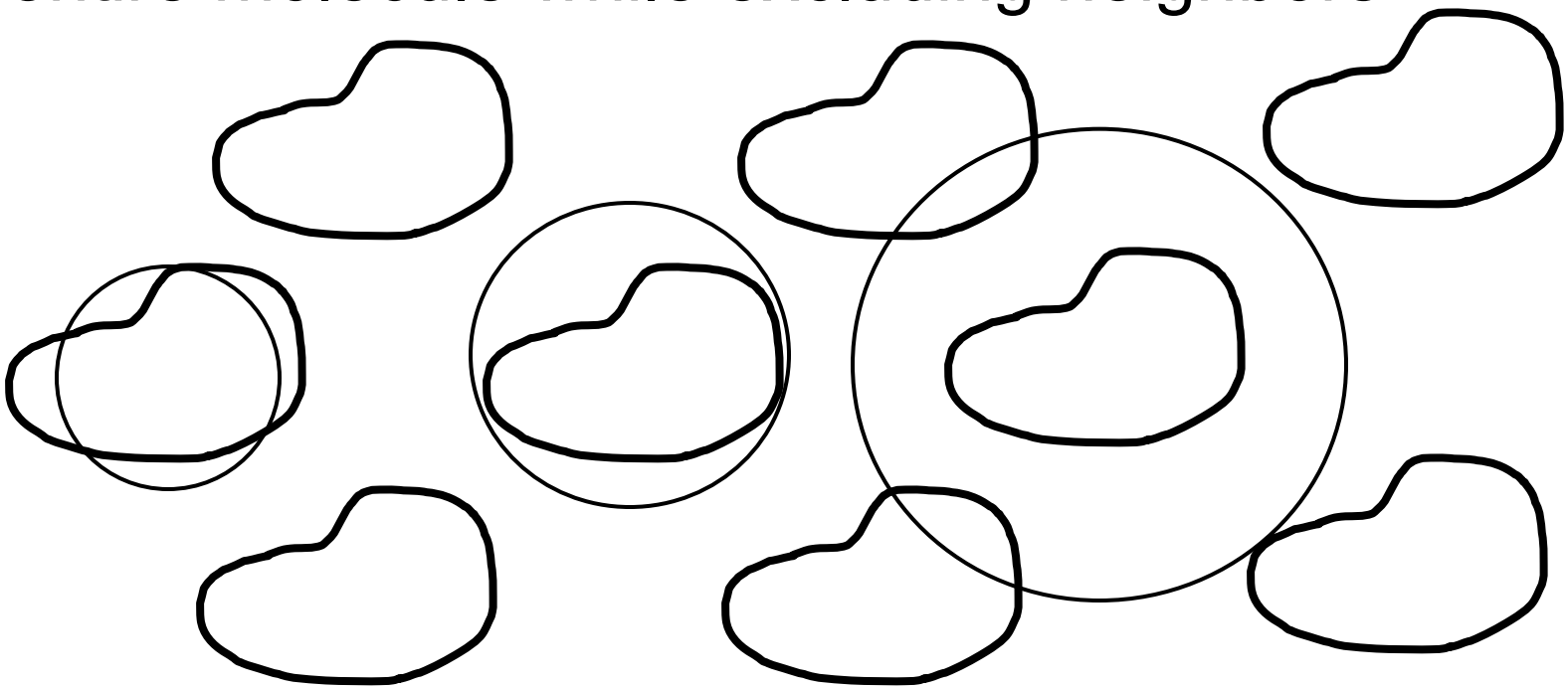
Merging of exposures and concentration

- Longer exposures give more signal at higher Q (more detail)
- X-ray detectors can become saturated at longer exposures

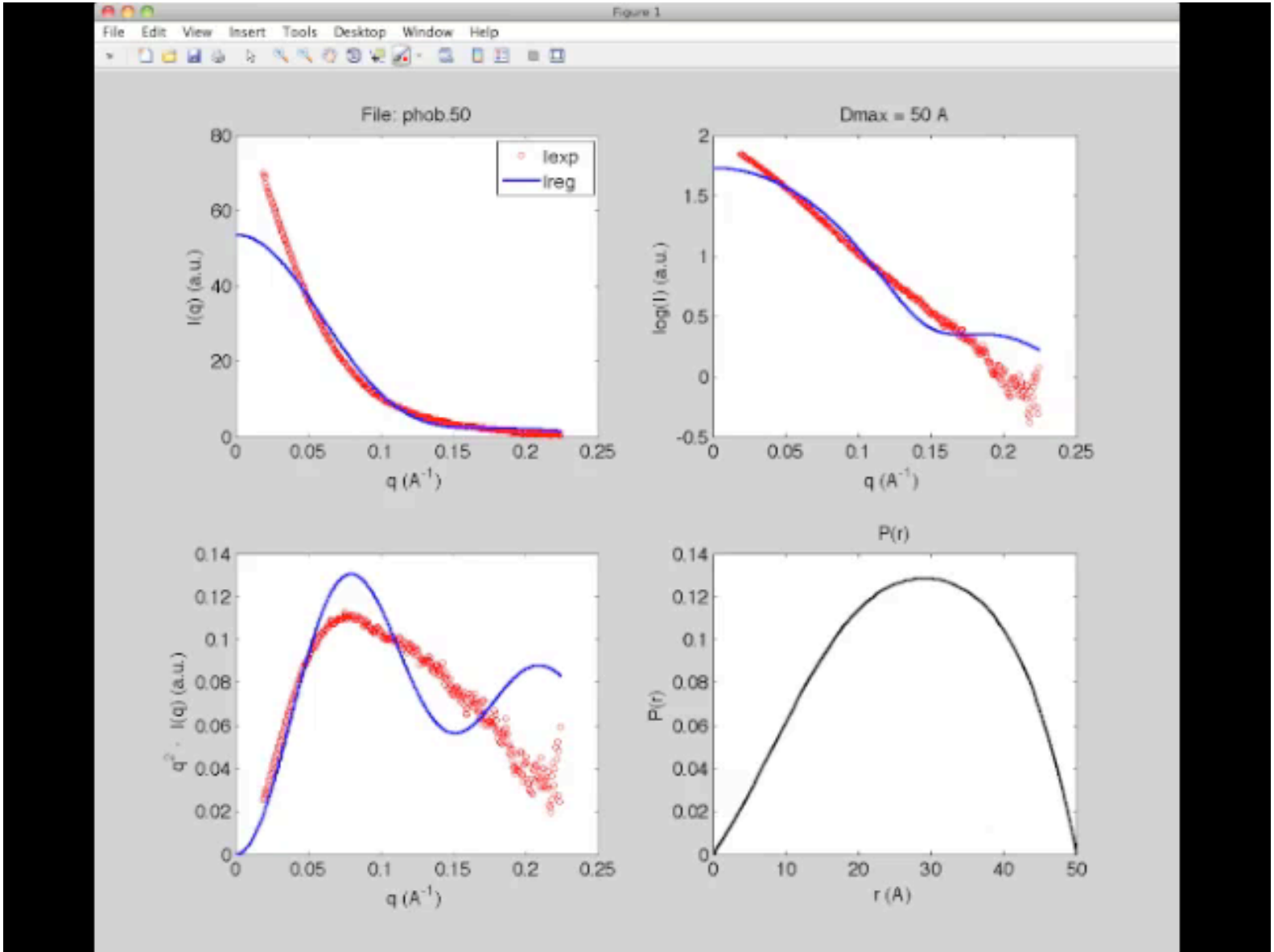


Calculation of electron envelope

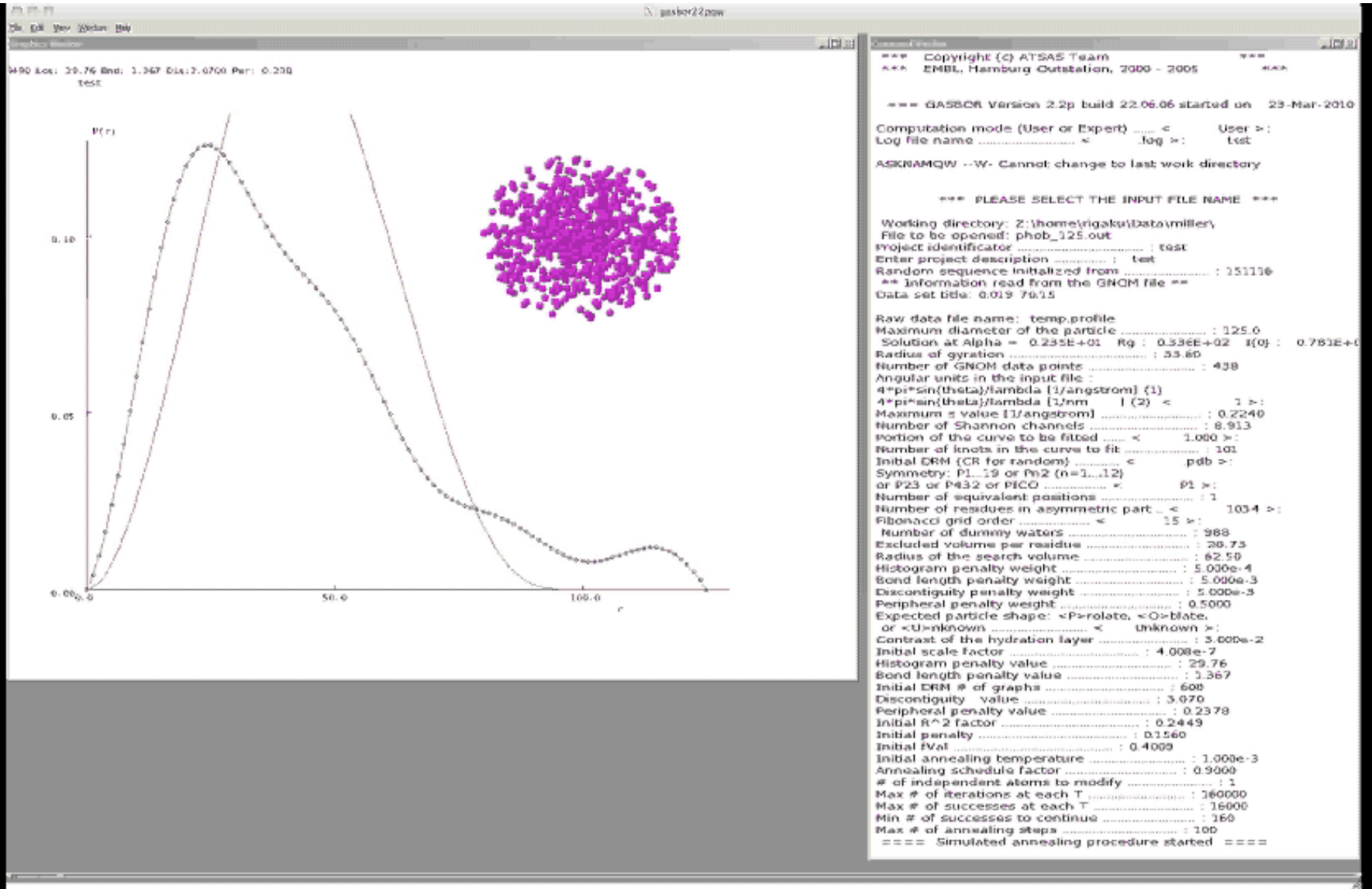
- Determination of protein radius (Dmax)
  - Dmax must be large enough to contain entire molecule while excluding neighbors



Fitting Dmax to data



Electron density



Fitting atomic resolution structure to data

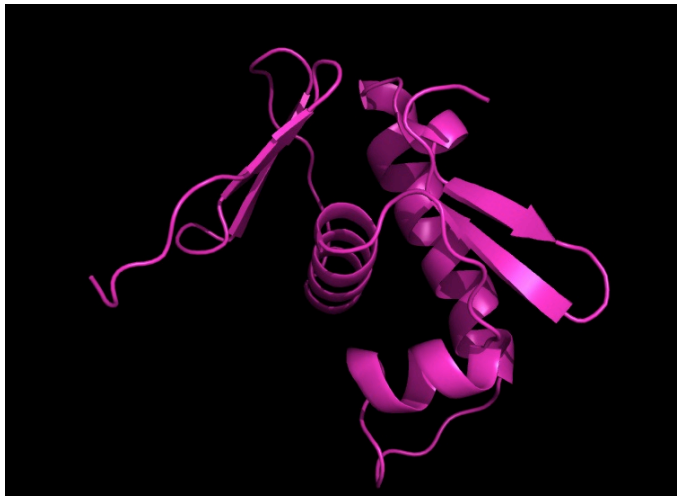
- Large proteins are composed of multiple domains whose structure can be determined using NMR or X-ray crystallography
- Structure of homologous proteins can be used as framework

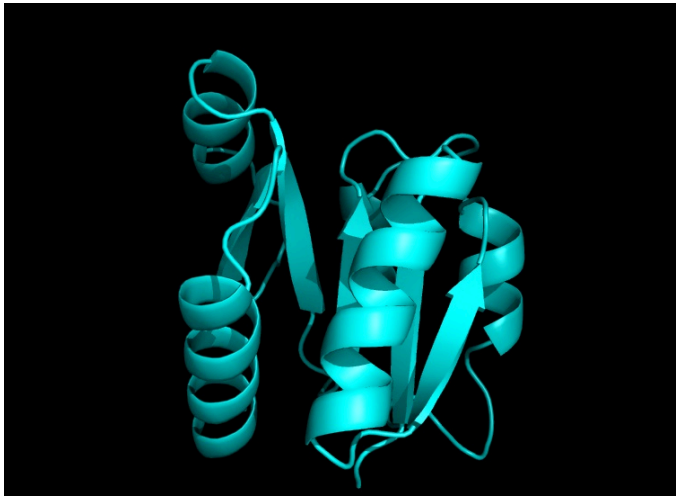
Fitting

- Take existing structures and orient them so the calculated scatter agrees with the data
- Normally run multiple times
- Complexes are averaged and outliers are rejected
- Success is measured by RMSD between structures

Fitting domains

- Use individual domains to recreate full structure
  - Create scattering curve from each domain
  - Determine orientation in which scatter matches measured profile



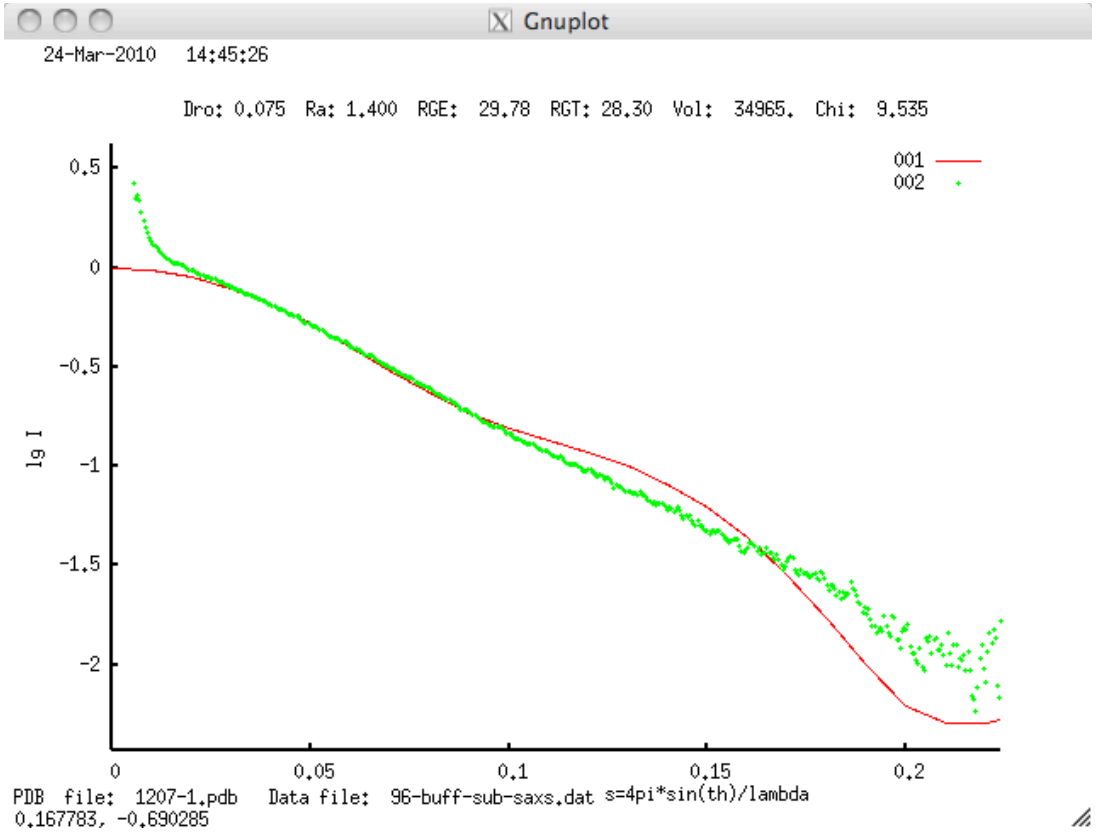


Construction of envelope

- Molecules are arranged to best match experimental data

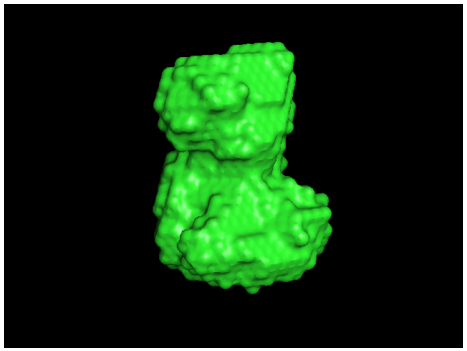
Red = calculated

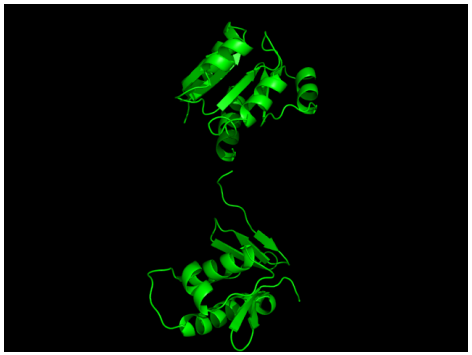
Green = experimental

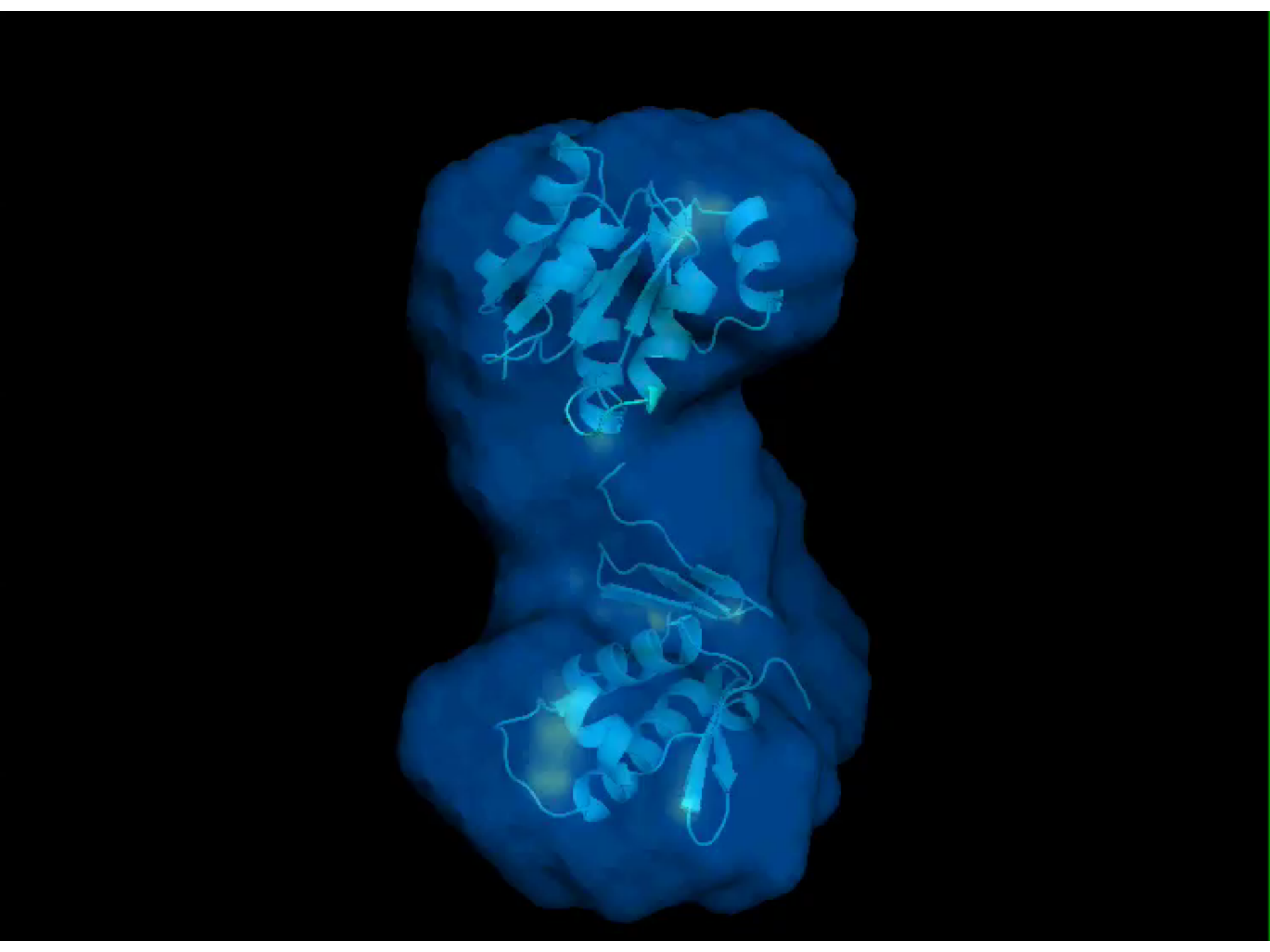


Putting together envelope and reconstruction

- Multiple electron density envelopes are averaged
- Multiple reconstructions are compared and a consensus molecule is selected







Is it correct?

- Agreement between envelope and reconstruction is a good sign
- SAXS data typically serves as the basis to conduct biochemical analysis to verify
- It’s a good start

Practical applications

Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS)

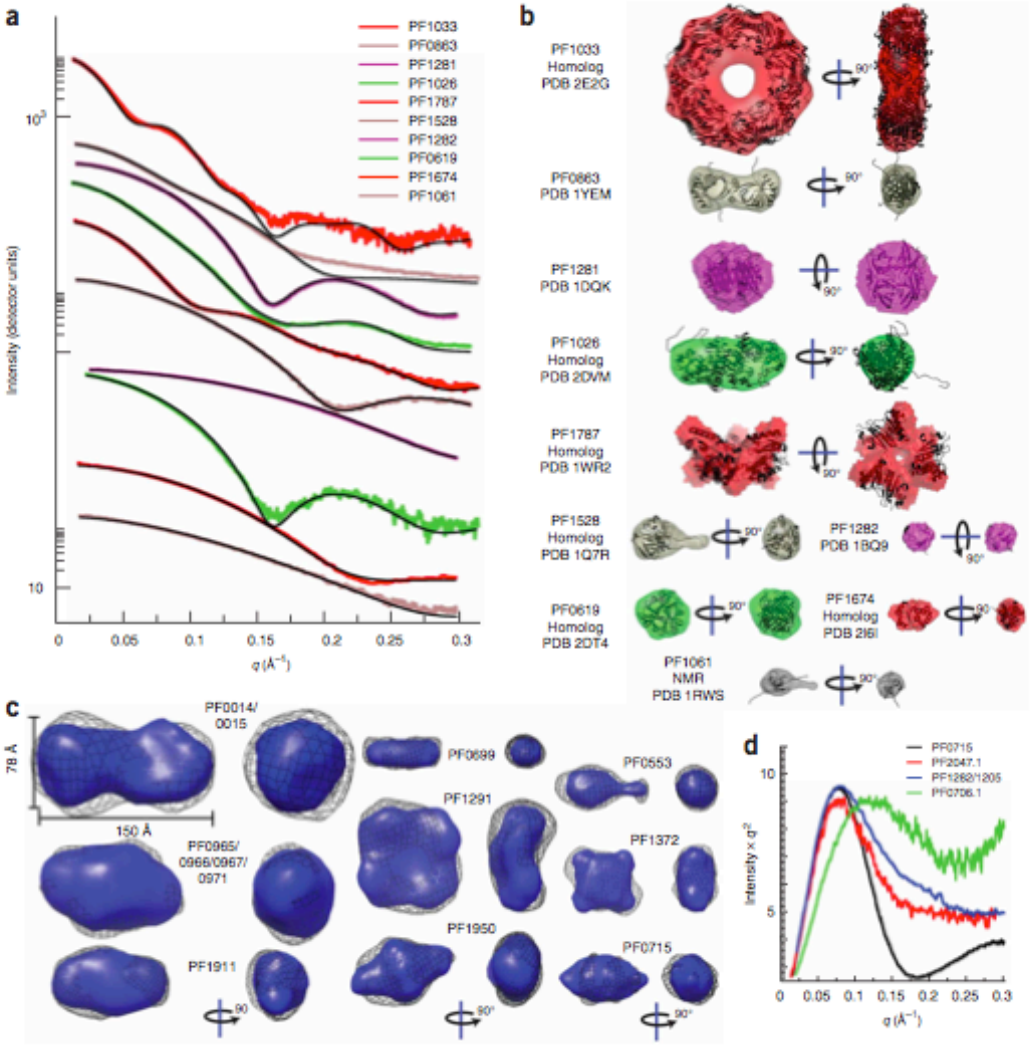
Greg L Hura[1,6], Angeli L Menon[2,6], Michal Hammel[1,6], Robert P Rambo[1], Farris L Poole II[2], Susan E Tsutakawa[3], Francis E Jenney Jr[2,4], Scott Classen[1], Kenneth A Frankel[1], Robert C Hopkins[2], Sung-jae Yang[2], Joseph W Scott[2], Bret D Dillard[2], Michael W W Adams[2] & John A Tainer[3,5]

We present an efficient pipeline enabling high-throughput analysis of protein structure in solution with small angle X-ray scattering (SAXS). Our SAXS pipeline combines automated sample handling of microliter volumes, temperature and anaerobic control, rapid data collection and data analysis, and couples structural analysis with automated archiving. We subjected 50 representative proteins, mostly from *Pyrococcus furiosus*, to this pipeline and found that 30 were multimeric structures in solution. SAXS analysis allowed us to distinguish aggregated and unfolded proteins, define global structural

the amount of sample required. New algorithms have been developed that can identify accurate shapes and assemblies based on the scattering data[4,10,11]. Notably, SAXS analyses can build on and be combined with other results to test experimental hypotheses and computational models[4].

Though lower in spatial resolution than crystallography or NMR spectroscopy, SAXS offers fundamental advantages for high-throughput structural analyses: structural measurements are carried out in solution, sample preparation is simple, quality global parameters can be obtained for most samples, and SAXS

Practical applications



Practical applications

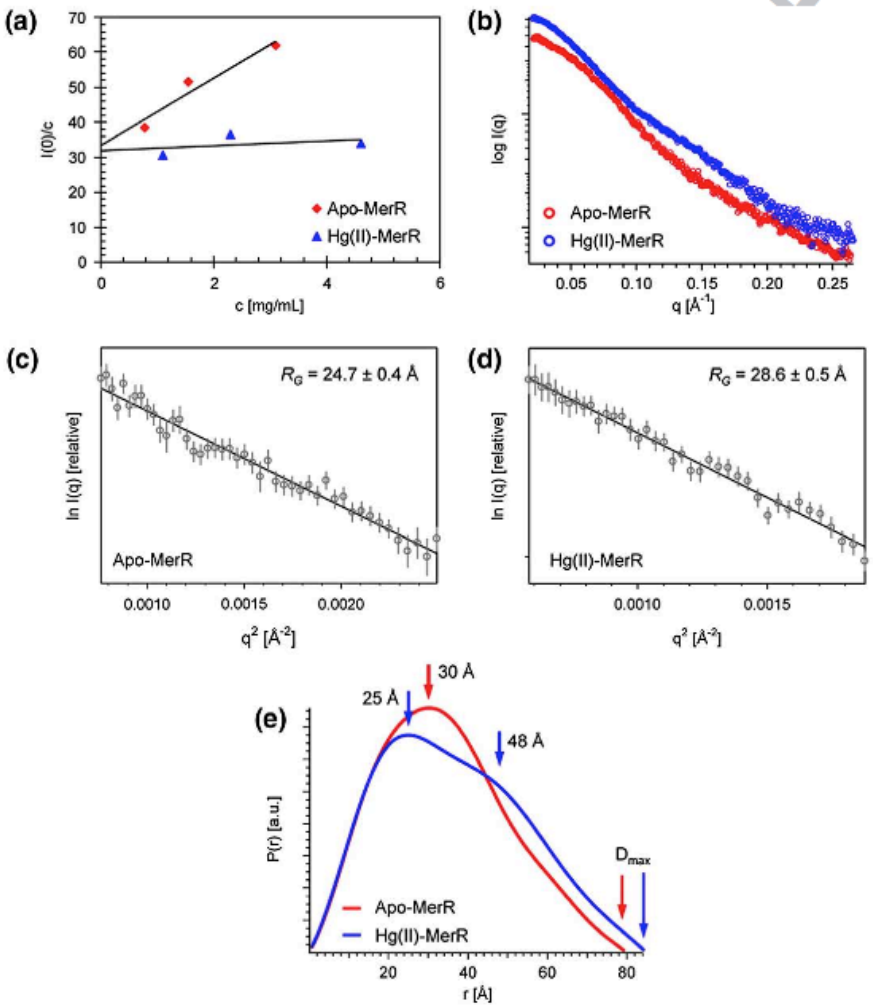
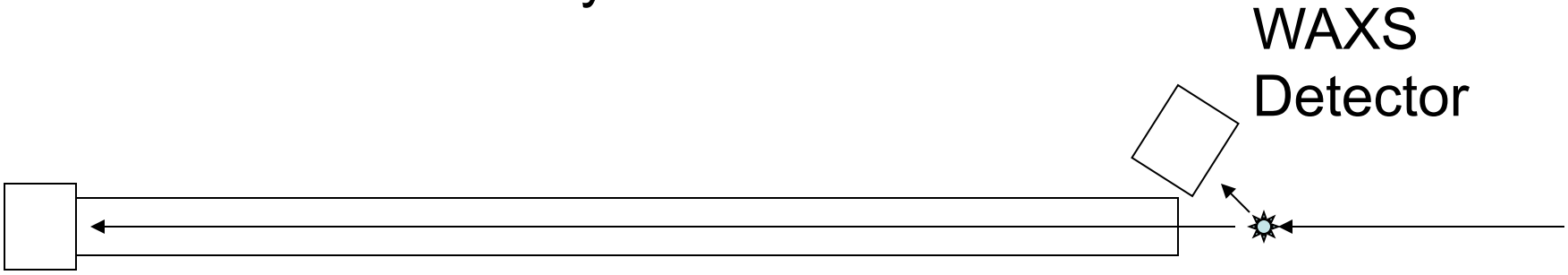


Fig. 1. SAXS data for apo-MerR and Hg(II)-MerR. (a) Scattering intensity at zero angle $I(0)$ normalized to concentration $I(0)/c$ *versus* protein concentration c. (b) Experimental scattering profiles. (c) Linearization plots of the low-q region to obtain R_g by the Guinier approximation for apo-MerR and (d) Hg(II)-MerR. Extrapolation of $I(q)$ to $q \rightarrow 0$ yields $I(0)$. (e) Distance distribution functions $P(r)$ for apo-MerR (red) and Hg(II)-MerR (blue) normalized to unity from indirect Fourier transformation of scattered intensities using the program GNOM.[35] The distances with the highest probability, as indicated by broken lines, were 30 Å for apo-MerR and 25 Å for Hg(II)-MerR, with an additional shoulder at ~48 Å, consistent with the expected separation of DBDs. D_{max} (arrows) for apo-MerR was 80 Å and shifted to 85 Å for Hg(II)-MerR.

SAXS for atomic resolution?

- Wide Angle X-ray Scattering (WAXS)
  - Higher resolution (~3Å)
  - Collected similarly to SAXS



- Used in addition to SAXS data