Small-Angle X-ray Scattering

Bio5325, Fall 2007

Tom Ellenberger

SAXS vs. X-ray Diffraction

- SAXS and x-ray diffraction are fundamentally similar.
- Both methods make use of a collimated, intense beam of x-rays to obtain structural information about the sample.
- Differences arise from making measurements of target molecules in solution (SAXS) or embedded in a crystal (diffraction).
- Solution scattering arises from tumbling molecules and it is radially symmetric (isotropic).
- X-ray diffraction from a crystal yields much higher resolution and a better signal-to-noise ratio (crystal acts as amplifier of scattering intensity sampled at discrete points).
- SAXS analysis can be applied to flexible proteins that don't easily crystallize.

SAXS vs. X-ray Diffraction

- SAXS does not require crystals and is a natural for understanding systems having substantial flexibility.
- SAXS data collection is rapid (seconds).
- SAXS requires microliters of a ~1-20 mg ml^{-1} solution of protein. Very economical.
- The precision/accuracy of SAXS structural analysis is inherently limited by a small number of observables. Even less information is available from SAXS than single particle EM.
- SAXS in combination with x-ray crystallographic data can be very powerful for the analysis of large multi-component systems.

Small angle scattering of x-rays

- Basis for scattering of x-rays is same in SAXS and single crystal diffraction.
- Thompson scattering is main component of x-ray scattering by macromolecules. It results from the elastic scattering of x-rays by electrons.
- In a typical SAXS or diffraction experiment, Thompson scattering approximates scattering by free (nonbonded) electrons.

Flexibility and Disorder in Crystals

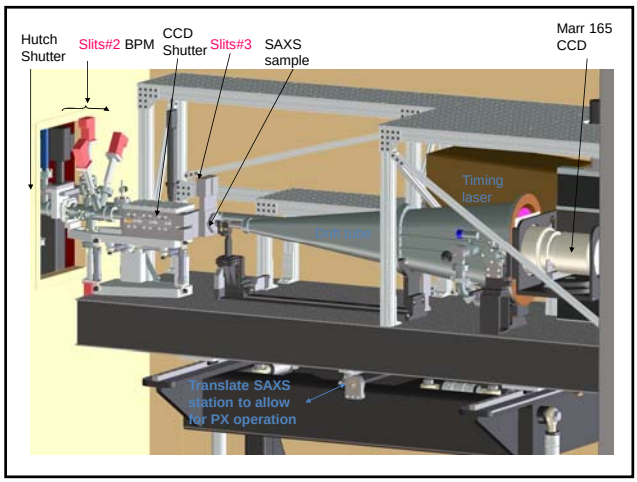
- Protein crystals are not perfectly ordered. Static and dynamic disorder are present.
- Disorder results in diffuse scatter around the positions of Bragg reflections.
- Disorder can be modeled by an atomic B-factor (temperature factor).
- Major conformations represented in a crystal can be independently modeled and assigned relative weights during crystallographic model refinement. High resolution data are required to justify this approach.

SAXS Data Collection

- SAXS is a contrast method.
- Scattering signal is derived from the difference between average electron density of the solvent (~0.33 $e^-/Å^3$ for water) and solute (e.g., ~0.44 $e^-/Å^3$ for protein).
- Thus, electron density contrast ($\Delta\rho$) is affected by solvent composition and by the sample concentration.

Small Angle X-ray Scattering at the SIBYLS beamline (12.3.1)

Advanced Light Source, LBNL



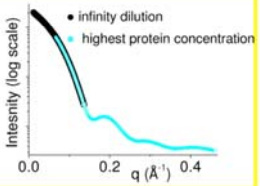
SAXS Intensity Curve

• Scattering intensity $I(q)$ is radially symmetric and resolution -dependent.

$q = \frac{(4\pi \sin\theta)}{\lambda}$ $\quad q = \frac{2\pi}{d}$ where $1/d$ is reciprocal resolution

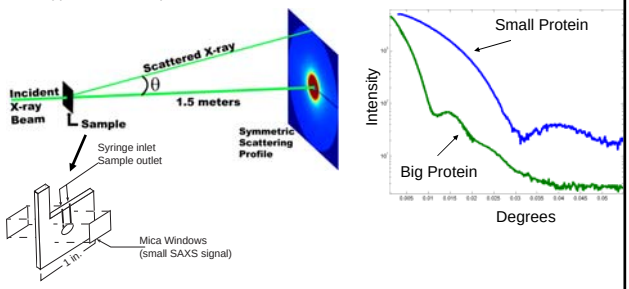
IDEAL SCATTERING

To obtain an ideal scattering curve for the entire q range, the scattering profile must be extrapolated to infinite dilutions at low resolution (q<0.1 Å^{-1}) and merged with the scattering profile for larger angles (q>0.1 Å^{-1}). Accurate data for the large angles can be obtained by measuring higher concentrations, using longer exposure times, and/or decreasing the sample-to-detector distance.



If the SAXS data is collected from monodisperse samples that are free of aggregates and interparticle interference, the reconstruction of the solution structure can proceed.

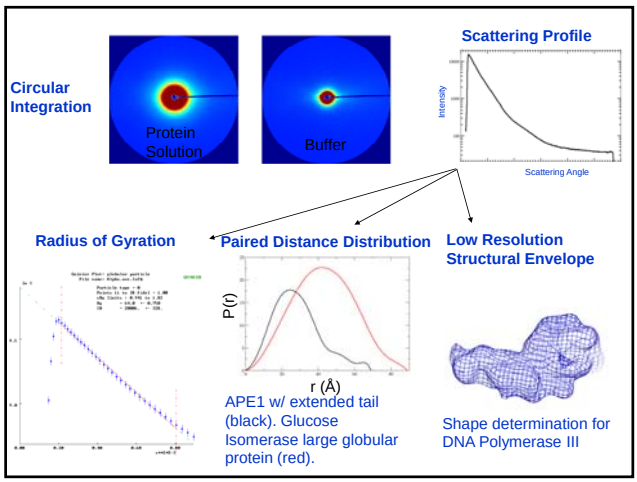
The Typical SAXS experiment



• Sample Requirements: 15 μL volume of 1-15 mg ml^{-1} protein.
• Protein solution must be homogeneous (monodisperse).
• Buffer solution—electron density should not exceed protein electron density.

SAXS Data Collection

- Data from a matched buffer blank is subtracted from scattering by experimental sample. A protein sample is typically dialyzed against a buffer that serves as the blank.
- Must measure blank precisely using same sample cuvette. Small differences in buffer composition will significantly affect SAXS data.
- Low angle scattering data may be contaminated by primary (unscattered) x-rays that miss the beamstop.
- Careful setup of the experimental station is crucial to success of a SAXS experiment.

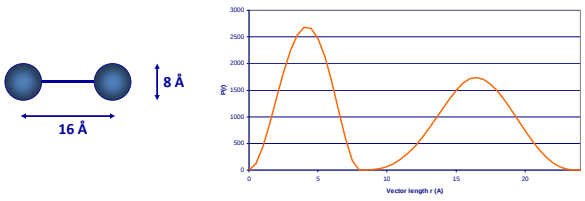


SAXS Intensity Curve

- Scattering $I(q)$ is radially symmetric.
- Difficult to confirm quality of $I(q)$ data from inspection. No SAXS equivalent of an R-factor.
- Concentrated solutions exhibit "interference" between molecules. Must check sample dilutions to obtain a consistent scattering profile (after normalization for protein conc.).
- At low resolution, SAXS data are dominated by the single size parameter for the molecule.

3 Model-Independent Parameters

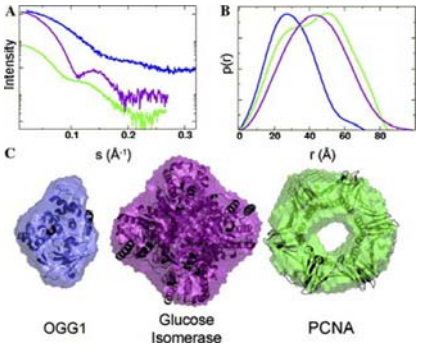
- D_{max} – Maximum inter-atomic distance
- R_G – Radius of gyration
  - Mass-weighted average radius
- P(r) – Paired vector distribution function
  - Histogram of all inter-atomic distance vectors.



Molecular Size and Shape

- Radius of gyration (R_G) is the x-ray analog of the hydrodynamic (Stokes) radius.
- R_G is highly shape-dependent and a poor measure of molecular weight (volume).
- R_G^2 corresponds to the average square distance of each scatterer (electron) from the center of the molecule.
- Sphere of radius r has $R_G = r(3/5)^{1/2}$

SAXS *vs*. Molecular Shape



Guinier Approximation

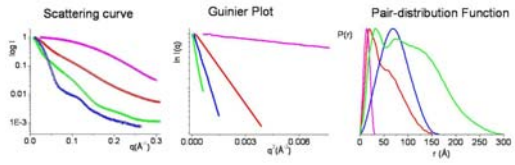
- At low resolution ($qR_G < 1.3$ for globular proteins), scattering can be related to the Guinier approximation:

$$I(q) = I(0)exp[-(q^2 R_G{}^2)/3]$$

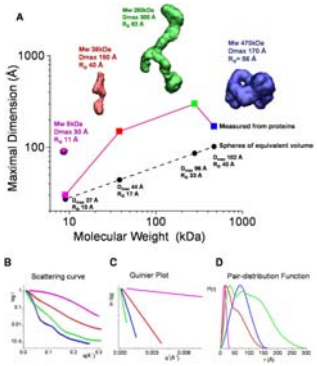
- Particles with large R_G will give rise to scattering with a small central peak, samples with small R_G will give rise to a large central peak.

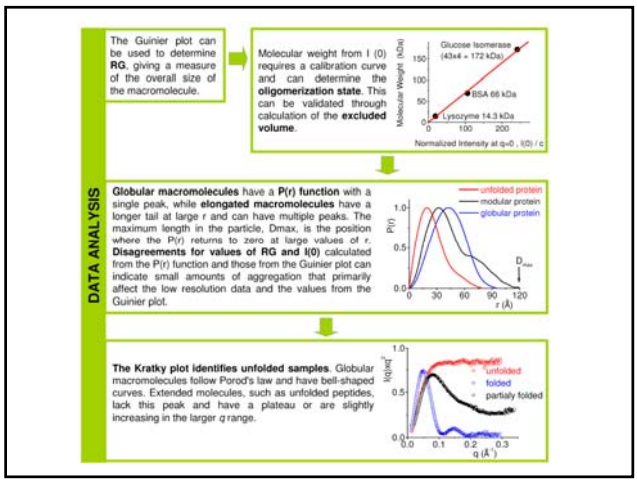
Guinier Plot

- The Guinier plot of $\log(I_{(q)})$ *vs*. q^2 will give a straight line from which R_G and $I(0)$ can be extracted.
- A linear Guinier plot indicates a well behaved globular sample. Aggregation or oblong shape will cause nonlinearity.



Molecular Size and Shape





Molecular Size and Shape

- The lowest resolution scattering $I(0)$, measured at zero angle ($q = 0$) and placed on an absolute scale is equal to the square of the number of electrons ($\propto$ mass).
- $I(0)$ is coincident with the direct beam and cannot be measured directly. It is determined by extrapolation of the scattering curve.
- Higher resolution scattering data contains information about the molecular shape.

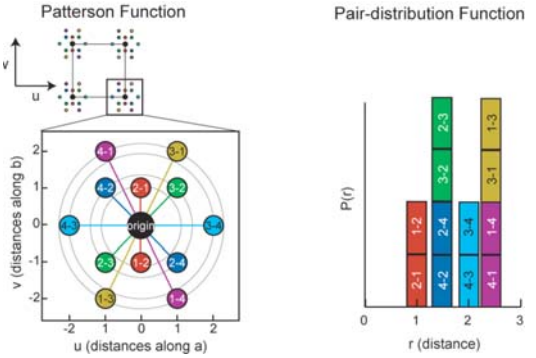
Pair-Distribution Function

- The pair-distribution function $P_{(r)}$ is the SAXS analog of the Patterson function.
- This autocorrelation function is directly calculated by a Fourier transform of the scattering curve.
- $P_{(r)}$ is radially averaged and lacks vectors corresponding to intermolecular distances (thus, no origin peak).

Pair-Distribution Function

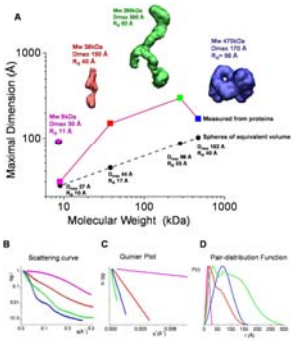
- $P_{(r)}$ is typically constrained to be zero at r = 0 Å and at $r \geq D_{MAX}$.
- Because $P_{(r)}$ is small in the vicinity of D_{MAX}, errors in estimates of D_{MAX} are difficult to detect.
- $P_{(r)}$ makes use of the full range of q, so it can be used to estimate a "real-space" value of R_G that does not suffer from aggregation artifacts prevalent at low resolution.

Pair-Distribution Function



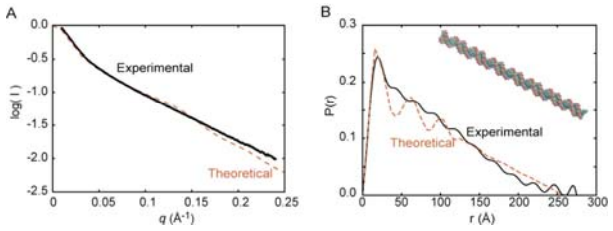
Pair-Distribution Function

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



Pair-Distribution Function

76mer Duplex DNA



Solvent Contrast Variation

- Scattering is mainly determined by the boundary between the solute and surrounding solvent.
- Most internal features of the molecule can be ignored for $q \leq 0.2$
- For multi-component systems, a choice of high density solvent (glycerol, salts, sugars) can be used to mask out one component.
- This method is particularly promising for nucleic acids in complex with proteins.

Structural Modeling Using SAXS

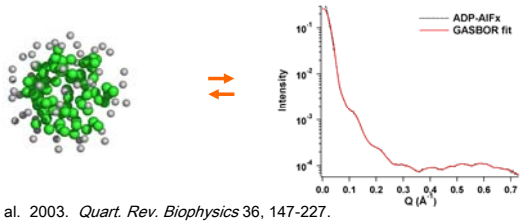
- SAXS analysis using known structures as constraints:
  - Identify biological multimers (size, shape) by docking of known structure(s).
  - Document conformational states that have/have not been observed in a crystal.
  - Analyze "unstructured" protein regions that are not represented in structures of truncated proteins.

Calculating Molecular Envelopes

- De novo calculation of molecular shape from SAXS data is an underdetermined problem.
- Few parameters are experimentally measured: R_G, D_{MAX}, molecular size.
- Shape of experimentally measured scattering curve $I(q)$ does not uniquely constrain molecular shape.
- Errors/noise in experimental data limit model precision and accuracy.
- D_{MAX} is estimated from noisy/weak scattering at high (q) values.

GASBOR (D. Svergun, EMBL Hamburg)

- Beads on a string - One scatterer / amino acid
- Simulated annealing to match to experiment
- Constraints on modeling:
  - Packing & connectivity (3.8 Å between spheres, C_α-C_α
  - Symmetry (if present). No symmetry info in SAXS data.



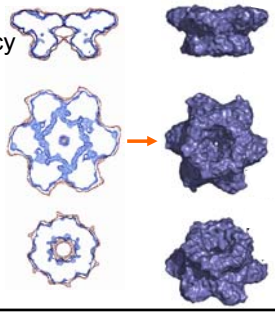
Koch et al. 2003. *Quart. Rev. Biophysics* 36, 147-227.
Konarev et al. 2003. *J. Appl. Cryst.* 36, 1227-1282.

DamAver – analysis of *ab initio* models

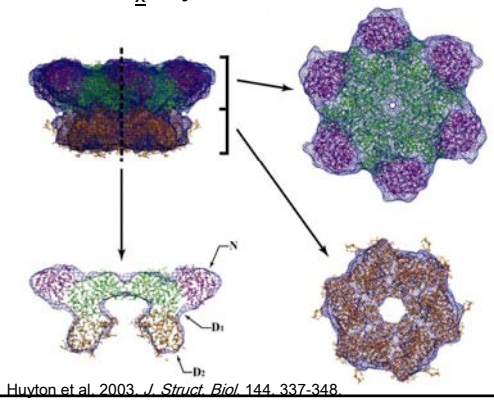
- Superposition of multiple runs from different random number seeds
- Comparison by Normalized Spatial Discrepancy
- Filter based on occupancy

Red = aligned, superimposed, summed runs

Blue = filtered, "most probable" model with highest voxel occupancies



Comparison of p97 SAXS structure to ADP•AlF_x crystal structure



Huyton et al. 2003. *J. Struct. Biol.* 144, 337-348.

Tomorrow's Journal Club

- Davies et al. 2005. Conformational changes of p97 during nucleotide hydrolysis determined by small-angle x-ray scattering. *Structure* 13, 183-195.