Introduction to biological small angle scattering

Frank Gabel (IBS/ILL)

HERCULES Specialized Course 16
(September 15th 2014)

Length-scales and tools in structural biology

SAS bridges the gap between atomic resolution (NMR and crystallography) and the light microscope
Objects that can be studied by SAS

Atomic resolution:

NMR and/or Crystallography

SAS cannot determine *de novo* the positions of individual atoms/residues in biomacromolecules on the Angstrom scale.

Biological small angle scattering is growing exponentially…
Scattering basics: Huygens-Fresnel principle

Incoming X-ray/neutron wave

Many scattering centers, FOURIER transform:

\[ I(Q) = \left\langle \left| \sum_j b_j e^{-i Q \cdot r_j} \right|^2 \right\rangle \]

\[ Q = \left| \vec{k} - \vec{k}' \right| = \frac{4\pi}{\lambda} \sin \theta \]

Reciprocal relationship between real space and the diffraction pattern

Many scattering centers, Fourier transform:

\[ I(Q) = \left\langle \left| \sum_{i,j} b_i b_j e^{-i Q (r_i - r_j)} \right|^2 \right\rangle \]

Orientational averaging

\[ I(Q) = \sum_{i,j} b_i b_j \frac{\sin Q(r_i - r_j)}{Q(r_i - r_j)} \]

Debye equation (1915)
Why “small” angle scattering?


SAS sample conditions and information obtained

Neutrons, X-rays

solutions ~ mg/ml volume ~ 10-200 μL mass > 0.1 mg

Global properties:
Radius of gyration, molecular weight

structural details

SAS sensitivity:
- Macromolecules 1 kDa ... ~ MDa
- Linear dimension 10 Å ... ~ 1000 Å

Information obtained:
1) Oligomeric state of macromolecules
2) Shape or conformation (globular, stick etc...)
3) Interaction of different macromolecules
4) Variation of points (1)-(3) as a function of pH, salt, ligands, T, p, ...
5) **Contrast variation**: visualisation of individual sub-units in situ
Concept of scattering density and contrast

In vacuo:

\[ I(Q) = \sum_{j} b_{j} e^{-\Psi_{j}} \]

\[ I(Q) = \left[ \int \frac{b_{j}}{V} e^{-\Phi_{j}} dV \right]^{2} \approx \int \rho_{\text{protein}} e^{-\Phi_{j}} dV \]

\[ I(Q) = N \left( \int \rho_{\text{protein}}^{*} e^{-\Phi_{j}} dV \right) \]

In solution:

\[ I(Q) = \left( \int \rho_{\text{protein}} \rho_{\text{solvent}} e^{-\Phi_{j}} dV \right) \]

Continuum approximation:

\[ \frac{b_{j}}{V_{j}} \approx \rho_{\text{protein}} = \text{const} \]

-How are scattering densities calculated?
-Under which conditions is the approximation valid?

Ideal solutions: no inter-particle effects, only form-factors

Model-free parameters
Guinier approximation and radius of gyration

\[ I(Q) \approx I(0) \exp \left(-\frac{1}{3} R_g^2 Q^2 \right) \quad R_g Q \leq 1...1.3 \]
(from expansion of Debye equation)

\[ \ln[I(Q)] \approx \ln[I(0)] - \frac{1}{3} R_g^2 Q^2 \]

Radius of gyration:

\[ R_g^2 = \frac{1}{M} \sum_i m_i r_i^2 \]

For a given molecular weight, a sphere has the smallest \( R_g \), i.e. it is the most compact object

Calibrating molecular weight (\( M_r \)) with water by SANS

Often important for the study of oligomeric/association states and flexible systems

\[ \frac{I(0)}{I_{inc}(0)} = 10^{-3} f \frac{4 \pi T_s}{1-T} CM_{\alpha} N_d \left[ \left( \sum b_i - \rho_s V \right) / M_r \right] \]

\[ M_r = \frac{1-T}{f \frac{4 \pi T_s}{1-T}} \frac{I(0)}{I_{inc}(0)} 10^{\frac{3}{2}} \left[ \left( \sum b_i - \rho_s V \right) / M_r \right]^{-2} \]

C: concentration in mg/ml
f: correction factor for anisotropy of incoherent scattering
T: water transmission
\( T_s \): sample transmission
\( t \): pathlength in cm
I(0): coherent scattering in forward direction
I_{inc}(0): incoherent scattering from water

Relative measurement of molecular weight ($M_r$) by SAS

\[ I(Q) \propto \left| \sum_j b_j e^{-iQ \cdot r_j} \right|^2 \]

\[ I(0) \propto \left| \sum_j b_j \right|^2 \]

\[ I(0) \propto N \left| \sum_j b_j \right|^2 \]

\[ N = C_n V / M_r \]

\[ \left| \sum_j b_j \right|^2 \propto M_r^2 \rightarrow I(0) \propto CM_r \]

Relative calibration to a known standard (BSA, lysozyme) in same buffer conditions:

\[ I_{prot}(0) = I_{standard}(0) \frac{C_{prot} M_{prot}}{C_{standard} M_{standard}} \]

Characteristic and distance distribution function
Characteristic and distance distribution function $\rho(r)$

Debye equation in integral form

$$I(Q) = \int \int \rho(r_1)\rho(r_2) \frac{\sin Q(r_1 - r_2)}{Q(r_1 - r_2)} dV_1 dV_2$$

$$\gamma_0(r) = \frac{\gamma(r)}{\gamma(0)} = \frac{\langle V_c(r) \rangle}{V}$$

$V_c$ is the shifted volume.

$$(r^2)^2 = 2 \mathcal{V}$$

$$(r^2)^2 = 2 \mathcal{V}$$

Figure 2.2. Characteristic function $\gamma_0(r)$ and distance distribution function $\rho(r)$ for a solid sphere of radius $R$.


SAS provides information on distances on different length-scales

Examples of $I(Q)$ and $p(r)$

\[ I(Q) = \left[ \frac{3 \sin(QR) - QR \cos(QR)}{(QR)^3} \right]^2 \]


Some words about resolution…

Nominal definition: \[ R = \frac{2\pi}{Q_{\text{max}}} \]

Inter-subunit distance with a precision of about 1-2 Å!

Inter-subunit position/orientations less well-defined!
Modelling

Sophisticated data analysis and modelling

- *Ab initio* structure analysis

- Rigid body modelling

- Validation of structural models
Different possibilities of modelling

Monodispersity is paramount (AUC, gels, SEC-MALLS)!

Flexible systems

Important check: molecular weight!
Oligomeric equilibria

Again: important check is molecular weight!

Fig. 21. The experimental $P(r)$ of the rabbit 15-lipoxygenase-1 (black). Theoretical $P(r)$ calculated for a mixture of two protein conformations adopting different extensions in the N-terminal regions with the indicated occupancies (red) and for a mixture of 80% monomeric and 20% oligomeric assemblies (blue) are shown. The crystal structures used in the $P(r)$ calculation (PDB id 1lox; Gillmor et al. 1997) using in the $P(r)$ calculation are shown as surfaces using the same coloring (after Hammel et al. 2004b).

Neutrons vs. X-rays
SAXS vs. SANS: scattering processes

- X-ray scattering length is proportional to number of electrons
- Neutron scattering length depends irregularly on atom and isotope

Atoms have a form-factor for X-rays but nuclei don’t for neutrons...

Practical calculation of scattering densities

Example glycine in H$_2$O:

\[ \rho = \left[ 2 \times 0.67 + 0.94 + 0.58 + 3 \times (-0.37) \right] 10^{-14} \text{cm}^{-2} \]

\[ = 2.68 \times 10^{10} \text{cm}^{-2} \]

\[ = 2.68 \times 10^{10} \text{cm}^{-2} \]

\[ \sum_j b \rho_{\text{protein}} = \sum_j b_j \]
SAXS vs. SANS: some practical aspects

<table>
<thead>
<tr>
<th></th>
<th>sample amount</th>
<th>flux</th>
<th>contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAXS</td>
<td>~15 µl @ 1-10 mg/ml</td>
<td>high</td>
<td>weak</td>
</tr>
<tr>
<td>SANS</td>
<td>~150 µl @ 1-10 mg/ml</td>
<td>low</td>
<td>high (D$_2$O)</td>
</tr>
</tbody>
</table>

parasitic scattering at low Q

flat background

SAXS, ID02, 1s exposure time

SANS, D22, 20min exposure time

(H$_2$O)  (D$_2$O)

SANS vs. SAXS instruments

D22 (ILL): SANS

BM29 (ESRF): SAXS

Quartz cuvette (SANS)

Capillary (SAXS)

Exposure times:
~ 10-20 minutes (D22)
~ 1-10 seconds (ID02)
~ 10-100 seconds (BM29)

Including sample change:
~ 10-20 minutes (D22)
~ 5 minutes (ID02)
~ 2-3 minutes (BM29)
Understanding contrast: destructive interference in SANS

Natural Contrast in SANS

Homogeneous macromolecules can be matched, i.e. made invisible!!!
Not so easy with SAXS…
Contrast in SAXS


**Figure 4.5.** Possibilities of contrast variation in X-ray (a) and neutron (b) scattering. Lines (1)–(4) correspond to low-density lipoproteins, proteins, RNA, and DNA. The scattering density of solutes in H₂O/D₂O mixtures increases slightly due to H/D exchange. Matching points are denoted by crosses.

An analogon in optics: refractive index
Artificial contrast using deuteration

Protein deuteration not complete but only ~75%!

Talk by Michael Härtlein...

Careful at high D$_2$O levels in the solvent: favours oligomerisation/aggregation!

Protein-protein complexes

Jacques and Trewhella (2010)
Prot. Sci. 19, 642-657

Petoukhov and Svergun (2006)
Practical guidelines

Putnam et al. (2007)
Putnam et al. (2007)
Combination of SAXS/SANS and with other techniques: recent highlights
Example 1: small protein-RNA complex

Translational repression in *D. melanogaster* dosage compensation

- *msl-2* translation is repressed by binding of UNR and SXL to the *msl-2* 3' UTR in female flies
- Minimal complex: 18-mer RNA motif (F-site) in *msl-2* mRNA is recognized by UNR cold shock domain 1 (CSD1) and SXL RR11-RRM2

Bennett et al Cell (2005); Duncan et al Genes Dev (2006); Abaza et al Genes Dev (2006); Patalano et al Development (2009)

**SAXS/SANS-specific information**

**SAXS:** overall envelope

**SANS:** Internal information

Program **MONSA:** Svergun (1999) *Biophys. J.* 76, 2879-2886.
EXAMPLE 1:
A large protein/RNA complex

The structure of the box C/D enzyme reveals regulation of RNA methylation. Nature 502(7472), 519-523

RNA modifications:
snoRNPs, snoRNAs and box C/D

snoRNP = “Small nucleolar Ribonucleo-Protein”

Only in archaea and eukaryotes, not in bacteria

snoRNP = sRNP in archaea

“Guide RNA” (> 100 in humans!)

≈ 390 kDa

Two different architectures have been proposed:

• Where is the sRNA situated?
• What is mechanism of methylation?
• Why two assymmetric methylation sites?
Information contained in SANS data: positions of FIB proteins within the complex

Important restraints for the atomic models!

Family of refined structures

Apo (no rRNA substrate)

Holo (with rRNA substrate)

Large conformational change upon substrate binding!
Concluding remarks

A few practical comments…

- use SAXS for homogeneous systems composed of a single body
- SAXS is better suited for high-throughput
- SANS good for complex systems (protein-DNA/RNA, membrane proteins…)
- SANS has no radiation damage
- neutrons only possible at large facilities (no “home sources” for the moment!)
- request for beam-time is generally via an electronic proposal system
- deadlines are usually twice a year, beamtime is attributed some months later
- BAG (“Block allocation group”) systems allow more flexible access
- for continuation proposals, reports need to be submitted regularly

- experiments need to be prepared with great care (i.e. isotopic effect of D₂O)!!
- “local contacts”, often beamline resonsibles, assist during experiments
- access (for non-industrial use) is in general free
- no maintenance, user friendly (software etc…)
Literature

**Basics (scattering, quantum mechanics):**
- The Feynman lectures on Physics, Volume 3: Quantum mechanics (Addison Wesley, 2006)

**Books on small angle (neutron) scattering:**
- Guinier/Fournet: Small angle scattering of X-rays (John Wiley & Sons, 1955)
- Serdyuk, Zaccai, Zaccai: Methods in molecular biophysics (Cambridge University Press, 2007)

**Reviews on SAXS/SANS:**