Interpretation of SAXS Data

Dr. Tetsuro Fujisawa
Biometal Science Laboratory
RIKEN SPring-8 Center
• Ideal optics at the expense of flux
  • High-resolution data collection with minimum smearing
    – Low parasitic scattering

• Relatively small beam size at the fixed energy (13.7 keV) (0.4×0.8 mm)
  – Time-resolved experiments with continuous flow apparatus
  – High-pressure small-angle X-ray scattering

• Moderate flux (ca. mid 10^{11} photons)
  – Diluted protein concentration
  – Time-resolved experiments with stopped flow experiment
Principle of small-angle scattering

Contrast of electron density is the source of scattering from protein complex.

\[ S = 2 \sin \theta / \lambda \]

\[ i_1(S) = \int \int \Delta \rho(r_1) \Delta \rho(r_2) \frac{\sin(2\pi r_{12}S)}{2\pi r_{12}S} dV_{r_1} dV_{r_2} \]

\[ i_1(S) = \sum_{i=1}^{N} \sum_{j=1}^{N} f_i(S) f_j(S) \frac{\sin(2\pi r_{ij}S)}{2\pi r_{ij}S} \]

In solution

In vacuum

\[ \Delta \rho(r) = \rho(r) - \rho_0 \]

\[ \bar{\rho} = 0.43 \]

\[ \rho_0 = 0.335 \]
Solution scattering deals with a system;

- Isotropic orientation against incident x-ray
- Equilibrium between solute and solvent
- *Time average* = *Spatial average*
A reliable instrument on a **synchrotron source** with a **low noise detector** is **INDISPENSIBLE**

SPring-8/BL-45XU SAXS

**Sample requirements**
- Sample volume: 50 μL
- Concentration: 0.1-10 mg/ml
- Molecular weight: few kDa to hundreds Mda

**Mono-disperse sample**

**Identical condition for sample and buffer scattering data collection**

**HP and time-resolved measurements**

**SPring-8/BL-45XU SAXS**

**X-ray** $\lambda = 0.9 \text{ Å, 13.8keV}$

**Synchrotron source**

**II-CCD detector**

**3600mm**

**Sample**

**Identical condition for sample and buffer scattering data collection**

$I(S)$

$S = 2\sin \theta / \lambda$
Development of fast mixing flow cell in sub-milli time region

In order to satisfy the identical condition for buffer subtraction, we use remote bulb switch with continuous liquid flow.
Why PX users started to use SAXS?

Protein structure and SAXS

Gauss function
Hydrodynamic values
Rg & I(0)

Shapes
Protein solution dynamics seen by SAXS

Beer’s law: In case of dilute solutions, scattering from the mixture of protein conformers are sum of scattering from each.

\[ I_{obs}(0) = \sum_{n} N_n I_n(0) = \sum_{n} N_n F_n^2(0) \]

\[ Rg_{obs}^2 = \frac{\sum_{n} N_n F_n^2(0) Rg_n^2}{\sum_{n} N_n F_n^2(0)} \]

\[ I(S) \approx I(0) \exp \left( -\frac{4\pi^2}{3} Rg^2 S^2 \right) \]

At small S all scattering curves are Gauss function.

Guinier’s law
Singular Value Decomposition (SVD) analyses

\[
[F] = [S] \cdot [C] = [U][W][V^T]
\]

Singular value decomposition shows the number of components incorporated in data matrix.
Determination of each scattering curves by global fitting

\[
\begin{align*}
[Y] &= [C][S] \\
[F] - [C][S] &\Rightarrow \min
\end{align*}
\]

When matrix \([C]\) is known, the \([S]\) can be determined.

\[
\begin{align*}
[A] &= [A]_0 \exp(-k_1 t) \\
[B] &= \frac{k_1}{k_2 - k_1} \{\exp(-k_1 t) - \exp(-k_2 t)\} \\
[C] &= [A]_0 \left(1 - \frac{1}{k_2 - k_1} (k_2 \exp(-k_1 t) - k_2 \exp(-k_1 t))\right)
\end{align*}
\]

\[
\begin{align*}
[\hat{S}] &= \left[C^T C\right]^{-1} \left[C^T\right][F] \\
[\hat{C}] &= [F][S^T \left[SS^T\right]^{-1}
\end{align*}
\]

Time resolved experiments can resolve each scattering curves because the change of fraction can be easily calculated.

The number of parameter to be determined is 3 from \(m\text{frame} \times 3\) parameters.
Cytochrome C refolding

Prof. Satoshi Takahashi’s group (Osaka Univ.)

150 µ sec
High resolution protein structures from NMR/PX and SAXS
Reason 1: Simulation from PDB has been greatly improved.

Freely available software packages (ATSAS, HASY Lab, EMBL)

Rigid body modeling of symmetric oligomers and subunits

- **GLOBSYM,MASHA** (ATSAS package, Svergun, Petoukhov, Konarev)

Add Missing Loops and Domains to High and Low Resolution Models of Proteins

- **CREDO** (Petoukhov, Svergun)

*Ab initio* Reconstruction of a Protein Structure by a Chain-Like Ensemble of Dummy Residues

- **GASBOR** (Petoukhov, Svergun,Koch)

http://www.embl-hamburg.de/ExternalInfo/Research/Sax/software.html
Small-angle scattering comprises bound water

\[ I(S) = \left\langle \left| A(\vec{S}) \right|^2 \right\rangle_{\Omega} = \left\langle \left| A(\vec{S}) - \rho_S E(\vec{S}) + \delta \rho_b B(\vec{S}) \right|^2 \right\rangle_{\Omega} \]

Scattering should be taken into account hydration shell especially for small globular proteins.

Fitting parameters from crysol

<table>
<thead>
<tr>
<th>$\delta \rho_b$</th>
<th>Shell</th>
<th>Vol</th>
<th>$\chi$</th>
<th>Hydra</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>3.16</td>
<td>16067</td>
<td>1.702</td>
<td>0.515</td>
</tr>
</tbody>
</table>

A(S): atomic scattering in vacuum

E(S): scattering from the excluded volume

B(S): scattering from the hydration shell

$\delta \rho_S$: contrast from bulk water (0-0.09 e/Å³)

Thickness of hydration shell: 3.16-3.3 Å
Reason 2: Low resolution shape determination is robust

Ab initio Shape Determination using a Single Phase Dummy Atom Model

- by Simulated Annealing (DAMMIN, Svergun)
- by Genetic Algorithm (DALAI_GA, Chacon)
- by Monte Carlo style give 'n' take algorithm (SAXS3D, Walther)

Independent of algorithms bead model converged to the average structure.

Ab initio determination from 2 dimensional data to 3 dimensional shape with constraints of non-negativity

\[ \sigma = \frac{\text{Maximum dimension of } \rho + \text{blank}}{\text{Maximum dimension of } \rho} \]

\( \sigma > 5 \) phase can be retrieved

Over sampling can retrieve phase information

J. Miao et al. (2001) PNAS, 98, 6641

Reason 3: The fluctuating part also contributes the SAXS models.

SAXS successfully determined protein shape prior to protein crystallography.
High resolution protein structures from NMR/PX and SAXS

Example of rigid body refinement

**CooA, A Transcriptional Activator Belonging to CAP Family**


*R. rubrum* synthesizes an enzyme system to generate energy from CO oxidation.

He et al. (1999, 2000)

\[ \text{CO + H}_2\text{O} \rightarrow \text{CooS, CooH, CooF} \rightarrow \text{CO}_2 + \text{H}_2 \]

---

**Diagram:**

- **CooA**
  - CO
  - DNA
  - RNAP
  - Transcription of CO-Oxidizing Proteins

- **CRP**
  - CAMP
  - DNA
  - RNAP
  - Transcription

**Legend:**

- CO
- DNA
- RNAP
- CAMP

**Text:**

Effector-Induced Conformational Changes

Specific Binding to Target DNA
## Known Crystal Structures of CAP Family

<table>
<thead>
<tr>
<th></th>
<th>- cAMP</th>
<th>+ cAMP</th>
<th>+ cAMP &amp; DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CRP</strong></td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td><strong>CooA</strong></td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
</tbody>
</table>

A: McKay et al. (1981)  
B: Schultz et al. (1981)  
C: Lanzilotta et al. (2000)

Exact activation mechanisms remained unknown …..
Proposed Regulation Mechanisms of CooA

Linear-Bent Hypothesis

Effector-Free (Linear) → Effector-Bound (Bent) → + Effectors & DNA

Lanzilotta et al. (2000), Chan et al. (2000)

CO-Dependent Hinge-Bending

Dramatic Switching from the Linear- to Bent-Conformations

Global Repositioning of the DNA-Binding Domains
Concentration Dependence of $R_g$ and $I(0)$

Interference-Free SAXS Parameters (at Infinite Dilution)

<table>
<thead>
<tr>
<th></th>
<th>$R_g$ (0) (Å)</th>
<th>$I(0,0)$ (a.u.)</th>
<th>$M_w$ (kDa)</th>
<th>$A_2$ (10^{-4} ml mol g^{-2})</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO-Free CooA</td>
<td>25.3 ± 0.5</td>
<td>3009 ± 42</td>
<td>56.0 ± 0.5</td>
<td>-0.39 ± 0.25</td>
</tr>
<tr>
<td>CO-Bound CooA</td>
<td>25.8 ± 0.6</td>
<td>3051 ± 45</td>
<td>57.6 ± 0.2</td>
<td>-5.48 ± 0.13</td>
</tr>
</tbody>
</table>

Negative Values in $A_2$ → Attractive Interparticle Interactions

Drastic Changes in $A_2$ → CO-Dependent Conformational Change
The CO-dependent conformational change of CooA is not so drastic as suggested in ‘Linear-Bent Hypothesis’.

Similarity of Molecular Shapes between CO-Free and CO-Bound CooA

Experimental $P(r)$s

Simulated $P(r)$s

$I(S) = 4\pi \int_0^\infty P(r) \frac{\sin(2\pi r S)}{2\pi r S} dr$  

$R_g = 27.0 \pm 0.2 \text{ Å} (1.00)$

$R_g = 25.8 \pm 0.2 \text{ Å} (0.95)$
Neither the crystal structure nor homodimer model was consistent with the experimental SAXS profile of CO-free CooA.

CO-free CooA does not adopt the linear-conformation in solutions.
The heterogeneous positioning of the DNA-binding domains suggests a structural flexibility of the hinge-region.

Amino Acid Sequence of the Hinge-Region

Heme Domain << LMFHDIKQRIA >> DNA-Binding Domain

The problem of backbone connectivity after refinement
Lacking some sequence in high-resolution structure

Exploring Conformational Space by Sampling the Dihedral Angles of the Hinge-Region (Phe132, His133 and Asp134)
74^3 of conformations were evaluated under the assumption of two-fold symmetry.
**Best-Fit Structure**

\[
\chi = \left[ \sum_i \left( I_{\text{exp}}(S_i) - I_{\text{model}}(S_i) \right)^2 / \sigma(S_i)^2 \right]^{-1/2}
\]

- CO-Free CooA
  - \( \chi = 2.3 \)
- CO-Bound CooA
  - \( \chi = 3.0 \)
- Bent-Conformation!!
Our Proposal for the Regulation Mechanism of CooA


Positional Changes in DNA-binding Sites (Arg177, Gln178 and Ser181)

- Recognition of Target DNA?

The C-terminus of the DNA-binding domains is rich in Asp content and retains an extensive acidic nature.

- Significant Decrease in $A_2$ value

Recognition of Target DNA or RNAP?
High hydrostatic pressure on protein solutions and SAXS

Search for the relation ship between hydrodynamic parameters and thermodynamic parameters
High pressure axis as a tool for studying enzymes

- Shift of $Keq$
- Shift of $k$
- Fluctuation

Developing SAXS system in order to correlate shape change with macroscopic physico-chemical values.

Applying pressure induces:
- Ion-pair formation
  \[ \text{H}_2\text{O} \Leftrightarrow \text{H}^+ + \text{OH}^- \] 21.3 ml/mol
- Hydrophobic hydration
  \[ \text{(CH}_4\text{)}_{\text{hexane}} \Leftrightarrow \text{(CH}_4\text{)}_{\text{water}} \] 22.7 ml/mol
- Protein dissociation
  LDH (M$^+ \rightarrow$ 4M) 500 ml/mol
- Apo to holo transition (LDH)
  LDH(apo$\rightarrow$holo) 390 ml/mol

Little effect on hydrogen bonding..

Force generation
=volume change

Reaction could be controlled by hydrostatic pressure
Recent HP NMR studies reveal the fraction of high energy conformers

Example of preliminary high-pressure SAXS measurement
Lactate dehydrogenase desociation  (Fujisawa, Kato, Inoko, Biochemistry, 1999)

- Fluorescence study suggested monomer at 200 MPa while SAXS determined dimer configuration.
- Dissociation by high pressure did not follow volume minimum pathway.

The use of SVD method for SAXS HP data revealed intermediate of dissociation.

- Fluorescence study suggested monomer at 200 MPa while SAXS determined dimer configuration.
- Dissociation by high pressure did not follow volume minimum pathway.
Concluding remarks

- SAXS sees time and spatial averaged shape of protein complex, which inevitably comprises bound waters.
- The shift of equilibrium by rapid mixing experiments or pressure perturbations is effective for detecting minor conformers.
- Rigid body refinement of multi-domain protein by SAXS can play more and more important roles: *SAXS selected the structure model proposed by PX.*

In the SAXS workshop ’SAXS in the 21st century’ Dr. MHJ Koch pointed that

**Small-angle scattering**

offers a bridge between low resolution methods including hydrodynamic and thermodynamic methods and high resolution or theoretical models.

If you want to collaborate with us, please mail me first.

E-mail: fujisawa@spring8.or.jp