Rigid Body Refinement: Applications to Proteins

Dr. Kate Brown
Imperial College London, UK
The University of Texas at Austin, US
Rigid Body Modeling is Great! (in general)

- Gives you context to your shape
- Can be done using coordinates from crystal/NMR structures and/or homology models
- Does not need to be “complete” (e.g., you can align partial models into shapes)
- Enhances your ability to make *ab initio* predictions of portions of your structure for which you have no other information
There are many examples of the use of Rigid Body Modeling Using SAXS

- Illustrate with some published examples
  - Brown and Fairweather labs – to predict protein complexes and multi-domain assemblies
  - Tainer Lab – to understand complex behaviour in complex system with conformational changes
  - Jackson lab – to study mechanisms of protein folding
  - See also recent review by Hayden and Dmitri Mertens HD, Svergun Dl., *J. Struct. Biol.* 2010, **172**, 128-41.
Case Study 1:
Oligomerization of Tetanus Toxin Fragment C ($H_C$)

Structure of Fragment C ($H_C$)

Tetanus toxin

![Diagram of Tetanus toxin with fragments HN, HC, HCN, and HCC]

**HCN domain**

- Cys869
- Cys1077
- Cys1093
- Cys1301

**HCC domain** (ganglioside binding)

- 6 x His tag
- NH$_2$
- COOH

- S-S bonds

**Diagram of HCN domain**

- Tetanus toxin

**Diagram of HCC domain**

- Ganglioside binding
Molecular Mass Comparison of Wild-type H_C and Site-Directed Mutants of Cysteine Residues
Analytical gel filtration

(a) wt $H_C$
- Cys1301 Ala
- Cys1077 Ala
- Trimer / Tetramer
- Monomer

(b) Cys869 Ala
- Cys1093 Ala
Mass spectrometric data of **Polydispersed** $H_C$

Mass spectrometric data of **Monomer** $H_C$ fraction from SEC
Mass Spectrometric Data of Dimer
$H_C$ fraction from SEC
SAXS data of Hc

Monomer (1)

Dimer (2)

Polydispersed samples of increasing concentrations (3-6)
## SAXS results

<table>
<thead>
<tr>
<th>$H_c$ sample</th>
<th>$R_G$ (nm)</th>
<th>$D_{max}$ (nm)</th>
<th>$V_{Porod}$ (nm$^3$)</th>
<th>$MW_{exp}$ (kDa)</th>
<th>% mon.</th>
<th>% dimer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monomer</td>
<td>2.90 ± 0.05</td>
<td>9.5 ± 1</td>
<td>76 ± 7</td>
<td>55 ± 5</td>
<td>90.1</td>
<td>9.9</td>
</tr>
<tr>
<td>Dimer</td>
<td>4.10 ± 0.05</td>
<td>13 ± 1</td>
<td>175 ± 10</td>
<td>80 ± 10</td>
<td>0.0</td>
<td>100</td>
</tr>
<tr>
<td>Polydispersed</td>
<td>3.40 ± 0.10</td>
<td>13.0 ± 1</td>
<td>96 ± 10</td>
<td>62 ± 10</td>
<td>64.2</td>
<td>35.8</td>
</tr>
<tr>
<td>Polydispersed</td>
<td>3.80 ± 0.10</td>
<td>13.0 ± 1</td>
<td>116 ± 10</td>
<td>68 ± 10</td>
<td>42.6</td>
<td>57.4</td>
</tr>
<tr>
<td>Polydispersed</td>
<td>3.90 ± 0.05</td>
<td>13.0 ± 1</td>
<td>140 ± 10</td>
<td>75 ± 10</td>
<td>20.5</td>
<td>79.5</td>
</tr>
<tr>
<td>Polydispersed</td>
<td>4.00 ± 0.05</td>
<td>13.0 ± 1</td>
<td>155 ± 10</td>
<td>77 ± 10</td>
<td>14.1</td>
<td>85.9</td>
</tr>
</tbody>
</table>
SAXS Derived Model of the $H_C$ Dimer

- Dimer shape calculated by DAMMIN, imposing 2-fold symmetry
- Known structure (monomer) docked into dimer shape by SASREF
- Two orientations fitted the scattering data
- Only one fitted with known biochemical data
Case Study 2: Rigid Body Modeling in SAXS-derived Shapes of Surface Layer Proteins and Toxin B from *C. difficile*


Four distinct structural domains in *Clostridium difficile* toxin B visualized using SAXS. Albesa-Jové et al., *J. Mol. Biol.* 2010 Mar 12;396(5):1260-70
Structural Goals of These Projects

• To determine the SAXS shapes of *C. difficile* proteins important for the virulence of the organism and align known structures/models into these shapes.

• The first project involved the characterization of a complex between two surface layer proteins (SLPS, ca. 80 kDa = 35 kDa LMWSLP + 45 kDa HMWSLP). We had a crystal structure and SAXS data for a truncated form of the LMWSLP, and SAXS data for full-length LMWSLP and the complex of LMWSLP + HMWSLP.
SLP Strategy

• The truncated crystal structure was fit into its SAXS shape with CRYSOL.

• DAMMIN was used to calculate shapes of the truncated and full-length LMWSLP and MONSA was used for the complex. 20 models were generated for each sample and averaged with DAMAVER.
Characterisation: Solution structures the S-Layer proteins
Toxin B Strategy

• Full-length toxin B is 2366 residues. It is made up of four domains. We had either crystal structures or homology models of three of portions of three of the four domains.

• DAMMIN was used to calculate shape of full-length Toxin B. 20 models were generated for each sample and averaged with DAMAVER.

• We used SITUS (http://situs.biomachina.org/) to align the three domains into the shape of TcdB.
Scattering curve and $P(r)$ plot for Toxin B

Figure 4

(a) Intensity vs. $q(\text{Å}^{-1})$

(b) $P(r)$ (Normalized) vs. $r$ (Å)
SAXS Models of Toxin B and Cryo-EM Model of Toxin A

A different conformation or partially Unfolded protein?

BUT NOT: “because ab initio envelope calculations from scattering data can be problematic in flexible systems in which domain orientations differ between conformers”

Pruitt et al., *Proc Natl Acad Sci U S A.* 2010 107, 13467-7
Hexameric structures of the archaeal secretion ATPase GspE and implications for a universal secretion mechanism

Atsushi Yamagata and John A Tainer*

Department of Molecular Biology, The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA, USA
Dynamic Hexameric Assembly in Solution

- *Archaeoglobus fulgidus* secretion superfamily ATPase (afGspE) was studied using crystallographic methods and SAXS. This protein hydrolyses ATP + Mg$^{2+}$ -> ADP + Pi to drive secretion processes.

- Crystal structures of afGspE demonstrate that it is a hexamer and that complexes with ATP-like compounds and Mg$^{2+}$ have mixed conformational states of the two domains which form each monomer – closed and open states
Orange = open
Blue = closed
SAXS Studies

• Use SAXS to study conformational changes in solution upon addition of ATP and post-hydrolysis.

• SAXS data were collected of:
  – Protein + 2 mM ADP-PNP + 10 mM Mg$^{2+}$
  – Protein + 2 mM ADP + 10 mM Mg$^{2+}$
Initial Data Analysis (II)

ADP-PNP ADP

• $R_g$ 51.0 Å 47.3 Å

• $D_{max}$ 148 Å 146 Å

• These data suggest that the hexameric rings are similar in size but the decrease in $R_g$ suggests that a conformational change has taken place.

SO MODEL POSSIBILITIES!!
Strategy for Interpreting SAXS data

• Use the crystal structure of the mixed form of afGspE and a crystal structure of a related protein to create a model of an all closed form and an all open form.

• Calculate scattering curves for all three forms: mixed, open and closed and see how well each curve (or combination of curves) fits the scattering data.
Closed form

• AMP-PNP-afGspE shows differences from the crystal structure (mixed form)
• Best fit is to a completely closed form – no other models fit better
• The same closed forms were generated with other non-hydrolysable ATP homologues, ATPγS and ADP-Vi, all with $\chi^2 \sim 5$. 
ADP-afGspE SAXS data

• Data somewhat resembles the crystal structure (mixed form)
• But data can also be modeled using a combination of mixed, open and closed states
• Fits of data to curves are not as good as AMP-PNP-afGspE in comparison. Why?
  1) solution structure is even more flexible
  2) models may need more subtle adjustments, they do not fully model ADP-afGspE SAXS curve
  3) However, no strong evidence that this complex looks like the closed form
Case Study 4:
Using SAXS to Study “Knots” in Proteins

Folding studies of two knotted methyltransferases: YibK & YbeA

YibK: 156 aa, CO=16/17, RCO=0.11

YbeA: 152 aa, CO=17, RCO=0.11
Strategy

• How do these knots form? Test knot formation by fusing another protein to the N- or C-terminus or both.

• SAXS models were made with GASBOR and rigid body modeling done with BUNCH
Creating the most deeply knotted proteins known

Mallam, Onuoha, Grossmann and Jackson (2008) Molecular Cell 30, 642-648
N- and C-terminal fusions of YibK & YbeA: Native-like dimers? SAXS & ITC

Surprisingly, it is possible to add a large structured domain to both N & C termini of YibK and YbeA to generate stable folded fusion proteins which bind cofactor and have structures similar to those expected.

These fusions have the deepest knots known.

ITC of the fusion proteins show they all bind to the cofactor with wild-type like affinities.