Selected Topics in Protein Structure and Function

- Structural Determination of Membrane Proteins

Presented by: Jyh-Yeu (Eric) Lee, Assistant Professor, BMI
Preparatory Readings

• Lipid cubic phase / bicelle crystallization
  • doi:10.1107/S2053230X14026843

• X-ray microdiffraction
  • doi:10.1039/B618173B

• Cryo-EM: 2D microcrystals and single particles
  • doi:10.1016/j.cocis.2018.01.010
  • doi:10.1126/science.aat4346
  • doi:10.1038/nmeth.3700

• Fusion protein
  • doi:10.1107/S2053230X15011061
Lecture Outline

1. Why studying membrane proteins?

2. Challenges in membrane protein structural biology

3. Ways to study membrane protein structures
   • X-ray crystallography
   • Electron microscopy
   • Computer simulation, NMR, mass spec, cross-linking, ...

4. Strategies in structural determination of membrane proteins
1. Why studying membrane proteins?

- Encoded by some 20-30% genes in typical genome.
- Major components of the mosaic lipid bilayers in cellular membranes
- Mediate cell-to-cell communication and signaling events.
- Disruptions or mutations in humans have been implicated in diseases, such as cardiovascular and metabolic diseases, cancer, rare genetic diseases, ...
Membrane proteome (human)

(Almén et al, BMC Biol, 2009)
Half drug targets are membrane proteins.

(Yildirim et al, Nat Biotech, 2007)
Available atomic/near-atomic models of membrane proteins (2019-3-8)

PDB Statistics: Overall Growth of Released Structures Per Year
Available atomic/near-atomic models of membrane proteins (2019-3-8)
2. Challenges in membrane protein structural biology

• Naturally occurred proteins exist in low abundance, with only a few exceptions (e.g., bacteriorhodopsin or aquaporin), and form complexes.

• *E. coli* is often not suitable for producing recombinant membrane proteins of eukaryotic origins.

• No so-called standard protocol of protein extraction, largely due to the complexity of protein-lipid interaction.

• Protocols of purification, crystallization, and *in vitro* reconstitution remain empirical for individual cases.
Choosing the appropriate expression hosts for recombinant proteins

(Zorman et al, Curr Opin Struct Biol, 2015)
Things to consider for membrane protein extraction and purification

- Cell disruption
- Solubilization agent
  - Detergents
  - Polymers
- Protein engineering
- Column chromatography
- *In vitro* reconstitution
3. Ways to study membrane protein structures: **optics & spectroscopy**

**Resolution limits:**

- **FRET**
- **FTIR**
- **CD**

![Diagram showing resolution limits and microscopy techniques](image)
Electron beam

If \( F(X) = \text{FT}[f(x)] \), then \( f(x) = \text{IFT}[F(X)] \)

where \( \text{FT} = \text{Fourier transform} \) & \( \text{IFT} = \text{Inverse Fourier transform} \).

Note: important in X-ray crystallography and 3D reconstruction algorithms.
X-ray crystallography: crystallization

Coarse vapor diffusion

X-ray crystallography: crystallization

Bicelle (stacked 2-D crystals)

(Ujwal & Bowie, Methods, 2011)
Purified detergent-solubilized membrane protein (On Ice)

Pipette to mix and incubate on ice for 30'

Protein-Bicelle Mixture

Crystallization trials using standard set up including robotics

(Ujwal & Bowie, Methods, 2011)
\[ [\text{Protein}] = 4-10 \text{ mg/mL} \]

\[ \Delta T = \sim 3 \degree C \quad (20-23 \degree C) \]

\[ \Delta [\text{AmSO}_4] < 0.5 \text{ M (Precipitant)} \]
X-ray crystallography: crystallization

*In meso* lipid cubic phase

(Caffrey, Acta Cryst F, 2015)
(Caffrey, Acta Cryst F, 2015)
(Caffrey, Acta Cryst F, 2015)
(Caffrey, Acta Cryst F, 2015)
X-ray crystallography: micro-diffraction

**LCP**

- Bacteriorhodopsin
- Bacteriorhodopsin
- Cholesterol
- Bacteriorhodopsin (X-ray damaged)
- Lysozyme

(Cherezov & Caffrey, Faraday Discuss, 2007)
(Warren et al, in "The Next Generation in Membrane Protein Structure Determination", 2016)
Accordingly, we sought to exploit other characteristics of the assorted phases to assist in their identification. Under the prevailing experimental conditions the bulk cubic phase always gave rise to discrete Bragg reflections or spots (Fig. 3(a)). These arise because the domain size of the cubic phase exceeds the diameter of the beam \( r \approx 5 \text{ mm} \) and the thickness of the sample-holding microcell \( 25 \text{ mm} \). Accordingly, a spotty pattern plus the corresponding d-spacings of the discrete reflections were used to identify the cubic phase.

In contrast, the lamellar (L\(_{\alpha}\)) phase tended to produce powder diffraction rings or arcs (Fig. 3(b)). Such patterns were never observed with the cubic phase in the microcells. Accordingly, powder-like diffraction along with a d-spacing range that is characteristic of the phase were used as hallmarks or signatures of the lamellar phase.

Our working hypothesis posits that a lamellar phase acts as a conduit between the bulk cubic phase and the crystal. Thus, diffraction characteristic of the lamellar phase was looked for in the vicinity of crystals growing in the cubic phase housed in Fig. 3.

---

1s exposure 10s exposure

*(Cherezov & Caffrey, Faraday Discuss, 2007)*
X-ray crystallography: micro-diffraction

*Bicelle*

A  
50-500 µm  
25-30 Å

B  
100-300 µm  
7-10 Å

C  
50-150 µm  
3.5-4 Å
- Long exposure
  2-5 sec @ APS
  30 sec @ ALS

- Radiation damage
  3-5 frames (< 5°)

- Signal (I/σ = 1-1.5 at 3.9-4Å)
X-ray crystallography: data process

reading from a file: C3_1_0020.x
reading from a file: C3_2_0001.x
reading from a file: C3_2_0002.x
reading from a file: C3_2_0003.x
reading from a file: C3_2_0004.x
reading from a file: C3_2_0005.x
reading from a file: C3_2_0006.x
reading from a file: C3_2_0007.x
reading from a file: C3_2_0008.x
reading from a file: C3_2_0009.x
reading from a file: C3_2_0010.x
reading from a file: C3_2_0011.x
reading from a file: C3_2_0012.x
reading from a file: C3_2_0013.x
reading from a file: C3_2_0014.x
reading from a file: C3_2_0015.x
reading from a file: C3_2_0016.x
reading from a file: C3_2_0017.x
reading from a file: C3_2_0018.x
reading from a file: C3_2_0019.x
reading from a file: C3_2_0020.x

reading from a file: B3_1_0020.x
reading from a file: B3_2_0001.x
reading from a file: B3_2_0002.x
reading from a file: B3_2_0003.x
reading from a file: B3_2_0004.x
reading from a file: B3_2_0005.x
reading from a file: B3_2_0006.x
reading from a file: B3_2_0007.x
reading from a file: B3_2_0008.x
reading from a file: B3_2_0009.x
reading from a file: B3_2_0010.x
reading from a file: B3_2_0011.x
reading from a file: B3_2_0012.x
reading from a file: B3_2_0013.x
reading from a file: B3_2_0014.x
reading from a file: B3_2_0015.x
reading from a file: B3_2_0016.x
reading from a file: B3_2_0017.x
reading from a file: B3_2_0018.x
reading from a file: B3_2_0019.x
reading from a file: B3_2_0020.x

reading from a file: B13_1_0020.x
reading from a file: B13_1_0021.x
reading from a file: B13_1_0022.x
reading from a file: B13_1_0023.x
reading from a file: B13_1_0024.x
reading from a file: B13_1_0025.x
reading from a file: B13_1_0026.x
reading from a file: B13_1_0027.x
reading from a file: B13_1_0028.x
reading from a file: B13_1_0029.x
reading from a file: B13_1_0030.x
reading from a file: B13_1_0031.x
reading from a file: B13_1_0032.x
reading from a file: B13_1_0033.x
reading from a file: B13_1_0034.x
reading from a file: B13_1_0035.x
reading from a file: B13_1_0036.x
reading from a file: B13_1_0037.x
reading from a file: B13_1_0038.x
reading from a file: B13_1_0039.x
reading from a file: B13_1_0040.x
Cryo-electron microscopy (cryo-EM)

- Electron microscopy of specimens (usually protein and/or DNA/RNA samples) in their “natural”, “hydrated” environment with out artificial contrast.

- Specimens are “vitreously” frozen and imaged at $\text{N}_2(\text{l})$ or $\text{He}(\text{l})$ temperature.

- Most importantly, NOT cryo-preparation of TEM!
Cryo-electron microscopy (cryo-EM)

- Electron crystallography
  - 2-D crystals
  - Helical crystals

- Single-particle cryo-EM

- Electron cryotomography (cryo-ET)
Image

a) Bright field imaging

b) Selected area diffraction

Diffraction
Only 2-D projections are recorded \((x, y)!!\)
(Amos et al, Prog Biophys Mol Biol, 1983)
Cryo-EM: 2-D crystals and MicroED

2-D crystallization by lipid-monolayer technique

**Lipid Monolayer**

**Step 1**

4 °C

**Step 2**

RT

**Step 3**

RT

**Step 4**

RT

- His-tagged Protein
- Ni-chelated Lipid
- Phospholipid
- Detergent
- Bio-Beads
- Nucleotide
2D crystallization by reconstitution

Protein in detergent solution

Add lipid

Remove detergent

2D crystals!!
Calculation of Amplitudes & Phases

<table>
<thead>
<tr>
<th>H</th>
<th>K</th>
<th>A</th>
<th>φ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-3</td>
<td>771.2</td>
<td>-122.9</td>
</tr>
<tr>
<td>1</td>
<td>-2</td>
<td>3836.0</td>
<td>-148.3</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>5184.4</td>
<td>9.6</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>599.8</td>
<td>-17.9</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>5953.2</td>
<td>15.9</td>
</tr>
<tr>
<td>2</td>
<td>-7</td>
<td>17.2</td>
<td>83.4</td>
</tr>
<tr>
<td>2</td>
<td>-4</td>
<td>41.4</td>
<td>-125.0</td>
</tr>
<tr>
<td>2</td>
<td>-3</td>
<td>1485.4</td>
<td>52.6</td>
</tr>
<tr>
<td>2</td>
<td>-2</td>
<td>2491.6</td>
<td>-152.5</td>
</tr>
</tbody>
</table>

FT
IFT

Cross-Correlation Map
S/N-enhanced Image

Correlation Averaging
real space

Fourier space

25Å
20Å
15Å

15Å
20Å
25Å
### 0° Non-tilting Images

<table>
<thead>
<tr>
<th>H</th>
<th>K</th>
<th>L</th>
<th>A</th>
<th>φ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-4</td>
<td>0</td>
<td>402.2</td>
<td>-118.7</td>
</tr>
<tr>
<td>1</td>
<td>-3</td>
<td>0</td>
<td>771.2</td>
<td>-122.9</td>
</tr>
<tr>
<td>1</td>
<td>-2</td>
<td>0</td>
<td>3836.0</td>
<td>-148.3</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>5184.4</td>
<td>9.6</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>599.8</td>
<td>-17.9</td>
</tr>
<tr>
<td>2</td>
<td>-7</td>
<td>0</td>
<td>17.2</td>
<td>83.4</td>
</tr>
<tr>
<td>2</td>
<td>-4</td>
<td>0</td>
<td>41.4</td>
<td>-125.0</td>
</tr>
<tr>
<td>2</td>
<td>-3</td>
<td>0</td>
<td>1485.4</td>
<td>52.6</td>
</tr>
<tr>
<td>2</td>
<td>-2</td>
<td>0</td>
<td>2491.6</td>
<td>-152.5</td>
</tr>
</tbody>
</table>

### ±70° Tilted Images

<table>
<thead>
<tr>
<th>H</th>
<th>K</th>
<th>L</th>
<th>A</th>
<th>φ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2529.0</td>
<td>182.4</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0</td>
<td>445.1</td>
<td>14.3</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>0</td>
<td>33.5</td>
<td>268.0</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>0</td>
<td>16.0</td>
<td>127.6</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>599.8</td>
<td>-17.9</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0</td>
<td>21.8</td>
<td>208.9</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>0</td>
<td>7.6</td>
<td>336.0</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>0</td>
<td>5.0</td>
<td>353.0</td>
</tr>
<tr>
<td>-6</td>
<td>-1</td>
<td>0</td>
<td>13.9</td>
<td>355.9</td>
</tr>
</tbody>
</table>

### 0° 2D Data (partial)

![Diagram of 2D Data (partial)](image)

### +20° 2D Data (partial)

![Diagram of +20° 2D Data (partial)](image)

### -20° ~ +20° 3D Data (partial)

![Diagram of 3D Data (partial)](image)
(Martynowycz & Gonen, Curr Opin Colloid Interf Sci, 2018)
(Bai et al, eLife, 2013)
(Bai et al, eLife, 2013)
(Cheng et al, Cell, 2015)
(Cheng et al, Cell, 2015)
(Cheng et al, Cell, 2015)
(Cheng et al, Cell, 2015)
4. Strategies in structural determination of membrane proteins

- Fusion proteins
- Antibody
- Ligands
- Library of small molecules
- Protein re-engineering
- ...

uOttawa.ca
Analogue approach has subsequently been used for a number of other MHC class II–peptide complexes, as well as T-cell receptor (TCR)–peptide and TCR–peptide–MHC class II ternary complexes (Reddy Chichili et al., 2013b). Another successful example involved the nuclear LIM (Lin-11/Islet-1/Mec-3) domain-containing zinc-binding transcription factors. An 11-residue Gly/Ser-rich linker was used to link the LIM domain of LMO4 to the C-terminus of LDB1 (LIM-domain binding protein 1), facilitating structure determination by both NMR and crystallography (Deane et al., 2003, 2004). Again, (Kobe et al, Acta Cryst F, 2015).
A glycine-rich linker (Fremont et al., 1996) has been used for a number of other MHC class II–peptide complexes, as well as T-cell receptor (TCR)–peptide and TCR–peptide–MHC class II ternary complexes (Reddy Chichili et al., 2013). Another successful example involved the nuclear LIM (Lin-11/Islet-1/Mec-3) domain-containing zinc-binding transcription factors. An 11-residue Gly/Ser-rich linker was used to link the LIM domain of LMO4 to the C-terminus of LDB1 (LIM-domain binding protein 1), facilitating structure determination by both NMR and crystallography (Deane et al., 2003, 2004). Again, an analogous approach has been used for a number of other MHC class II–peptide complexes, as well as T-cell receptor (TCR)–peptide and TCR–peptide–MHC class II ternary complexes (Reddy Chichili et al., 2013).

Figure 2

Examples of successful application of the heterologous fusion-protein approach. The structures are not shown on the same scale. (a) Cartoon diagram of the structure of HTLV-1 gp21 (subunits are shown in different colours) fused at the N-terminus to MBP (in different shades of grey) with a three-Ala linker (red; PDB entry 1mg1; Kobe et al., 1999). (b) Cartoon diagram of the structure of 2-adrenergic receptor (2AR; blue) with T4 lysozyme (T4L; grey) inserted into a loop in 2AR (Rosenbaum et al., 2007; PDB entry 2rh1). (c) Avi evo of crystal packing interactions for the 2AR-T4L fusion protein [shown and coloured as in (b)]. Note the crystal contacts between the fusion partner T4L and the soluble portion of 2AR. (d) Cartoon diagram of the structure of the complex of the extracellular domains of TLR1 (green) and TLR2 (blue) (Jin et al., 2007; PDB entry 2z7x). Both proteins are fused at the C-terminus to VLR as the fusion partner (grey). All structure figures were produced with PyMOL (Schrödinger).
Discussion Topics and Brainstorming

1. How does the use of direct detector benefit structural biology?

1. Why do we need to do motion correction for cryo-EM data?

1. How can we use fluorescent probes to facilitate structural determination?

1. How can we use nanoparticle technology (e.g., nanodiscs, SMA, ...) in structural biology?

2. What can we do to examine the dynamic mode of structural biology?