Dealing with Resolution Anisotropy and Considerations for High Resolution Structure Determination

Dmitry Lyumkis
Salk Institute for Biological Studies
American Crystallographic Association, cryo-EM Workshop, August 11, 2020
Outline

• Preferred orientation — the origin of resolution anisotropy
• Resolution anisotropy in cryo-EM reconstructions (caused by preferred orientation) and its evaluation with 3D FSCs
• Combatting anisotropy using tilted data collection strategies
• The (unexpected) effect of sampling (in)homogeneity on global resolution
• Evaluating sampling (in)homogeneity using the SCF criterion
• Considerations for high-resolution
Origin of resolution anisotropy in cryo-EM: The idealized behavior of particles on a grid

Origin of resolution anisotropy in cryo-EM: The reality ... Particles adhere at the air-water interface.

<table>
<thead>
<tr>
<th>Sample # Name</th>
<th>Example cross-sectional schematic diagram</th>
<th>Sample # Name</th>
<th>Example cross-sectional schematic diagram</th>
<th>Sample # Name</th>
<th>Example cross-sectional schematic diagram</th>
<th>Sample # Name</th>
<th>Example cross-sectional schematic diagram</th>
</tr>
</thead>
<tbody>
<tr>
<td>1* 32 kDa Kinase</td>
<td><img src="image1.png" alt="Image" /></td>
<td>14* Neural Receptor</td>
<td><img src="image2.png" alt="Image" /></td>
<td>27* IDE</td>
<td><img src="image3.png" alt="Image" /></td>
<td>38* Apoferritin (0.5 mg/mL)</td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td>4** Hemagglutinin</td>
<td><img src="image5.png" alt="Image" /></td>
<td>17* Protein with Bound Lipids (deglycosylated)</td>
<td><img src="image6.png" alt="Image" /></td>
<td>30* GDH</td>
<td><img src="image7.png" alt="Image" /></td>
<td>39* Apoferritin with 0.5 mM TCEP</td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
<tr>
<td>5* HIV-1 Trimer Complex 1</td>
<td><img src="image9.png" alt="Image" /></td>
<td>18 Protein with Bound Lipids (glycosylated)</td>
<td><img src="image10.png" alt="Image" /></td>
<td>31* GDH</td>
<td><img src="image11.png" alt="Image" /></td>
<td>40 Protein with Carbon Over Holes</td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
<tr>
<td>6* HIV-1 Trimer Complex 1</td>
<td><img src="image13.png" alt="Image" /></td>
<td>19* Lipo-protein</td>
<td><img src="image14.png" alt="Image" /></td>
<td>32* GDH + 0.001% DDM (2.5 mg/mL)</td>
<td><img src="image15.png" alt="Image" /></td>
<td>41 Protein and DNA Strands with Carbon Over Holes</td>
<td><img src="image16.png" alt="Image" /></td>
</tr>
<tr>
<td>7* HIV-1 Trimer Complex 2</td>
<td><img src="image17.png" alt="Image" /></td>
<td>20 GPCR</td>
<td><img src="image18.png" alt="Image" /></td>
<td>33* DNAB Helicase-helicase Loader</td>
<td><img src="image19.png" alt="Image" /></td>
<td>42* T20S Proteasome</td>
<td><img src="image20.png" alt="Image" /></td>
</tr>
<tr>
<td>10* Stick-like Protein 1</td>
<td><img src="image21.png" alt="Image" /></td>
<td>21* Rabbit Muscle Aldolase (1mg/mL)</td>
<td><img src="image22.png" alt="Image" /></td>
<td>34* Apoferritin</td>
<td><img src="image23.png" alt="Image" /></td>
<td>43* T20S Proteasome</td>
<td><img src="image24.png" alt="Image" /></td>
</tr>
<tr>
<td>12* Stick-like Protein 2</td>
<td><img src="image25.png" alt="Image" /></td>
<td>22* Rabbit Muscle Aldolase (6mg/mL)</td>
<td><img src="image26.png" alt="Image" /></td>
<td>35* Apoferritin</td>
<td><img src="image27.png" alt="Image" /></td>
<td>44* T20S Proteasome</td>
<td><img src="image28.png" alt="Image" /></td>
</tr>
<tr>
<td>13* Neural Receptor</td>
<td><img src="image29.png" alt="Image" /></td>
<td>25* Protein in Nanodisc (0.58 mg/mL)</td>
<td><img src="image30.png" alt="Image" /></td>
<td>36* Apoferritin</td>
<td><img src="image31.png" alt="Image" /></td>
<td>45* Mtb Proteasome</td>
<td><img src="image32.png" alt="Image" /></td>
</tr>
<tr>
<td>27* IDE</td>
<td><img src="image33.png" alt="Image" /></td>
<td>38* Apoferritin (0.5 mg/mL)</td>
<td><img src="image34.png" alt="Image" /></td>
<td>39* Apoferritin with 0.5 mM TCEP</td>
<td><img src="image35.png" alt="Image" /></td>
<td>40 Protein with Carbon Over Holes</td>
<td><img src="image36.png" alt="Image" /></td>
</tr>
<tr>
<td>41 Protein and DNA Strands with Carbon Over Holes</td>
<td><img src="image37.png" alt="Image" /></td>
<td>42* T20S Proteasome</td>
<td><img src="image38.png" alt="Image" /></td>
<td>43* T20S Proteasome</td>
<td><img src="image39.png" alt="Image" /></td>
<td>44* T20S Proteasome</td>
<td><img src="image40.png" alt="Image" /></td>
</tr>
<tr>
<td>45* Mtb Proteasome</td>
<td><img src="image41.png" alt="Image" /></td>
<td>46 Protein on Streptavidin</td>
<td><img src="image42.png" alt="Image" /></td>
<td>47* Mtb Proteasome</td>
<td><img src="image43.png" alt="Image" /></td>
<td>48 Protein on Streptavidin</td>
<td><img src="image44.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Outline

- Preferred orientation — the origin of resolution anisotropy
- Resolution anisotropy in cryo-EM reconstructions (caused by preferred orientation) and its evaluation with 3D FSCs
- Combatting anisotropy using tilted data collection strategies
- The (unexpected) effect of sampling (in)homogeneity on global resolution
- Evaluating sampling (in)homogeneity using the SCF criterion
- Considerations for high-resolution
Interaction at the air-water interface results in: “preferred particle orientation” — Hemagglutinin (HA) trimer

Interaction at the air-water interface results in: smearing and elongation along Z-axis (loss of Z resolution)

Fourier Shell Correlation (FSC): a quantitative way to evaluate resolution in cryo-EM

- Calculate the FSC between two half maps as a function of spatial frequency over the entire Fourier shell

Fourier Shell Correlation (FSC): a quantitative way to evaluate resolution in cryo-EM

- Calculate the FSC between two half maps as a function of spatial frequency over the entire Fourier shell

Fourier Shell Correlation (FSC): a quantitative way to evaluate resolution in cryo-EM

- Calculate the FSC between two half maps as a function of spatial frequency over the entire Fourier shell

Fourier Shell Correlation (FSC): a quantitative way to evaluate resolution in cryo-EM

- Calculate the FSC between two half maps as a function of spatial frequency over the entire Fourier shell

Fourier Shell Correlation (FSC): a quantitative way to evaluate resolution in cryo-EM

- Calculate the FSC between two half maps as a function of spatial frequency over the entire Fourier shell

Fourier Shell Correlation (FSC): a quantitative way to evaluate resolution in cryo-EM

- Calculate the FSC between two half maps as a function of spatial frequency over the entire Fourier shell.

Fourier Shell Correlation (FSC): a quantitative way to evaluate resolution in cryo-EM

- Calculate the FSC between two half maps as a function of spatial frequency over the entire Fourier shell

3D FSCs: a quantitative way to evaluate directional resolution and anisotropy in single-particle cryo-EM

- Calculate the FSC between two half maps as a function of the direction (and looking at cross-correlations over a local cone of data and not the entire shell)

3D FSCs: a quantitative way to evaluate directional resolution and anisotropy in single-particle cryo-EM

- Calculate the FSC between two half maps as a function of the direction (and looking at cross-correlations over a local cone of data and not the entire shell)

Calculate the FSC between two half maps as a function of the direction (and looking at cross-correlations over a local cone of data and not the entire shell)

3D FSCs: a quantitative way to evaluate directional resolution and anisotropy in single-particle cryo-EM

- Calculate the FSC between two half maps as a function of the direction (and looking at cross-correlations over a local cone of data and not the entire shell)

3D FSCs: a quantitative way to evaluate directional resolution and anisotropy in single-particle cryo-EM

- Calculate the FSC between two half maps as a function of the direction (and looking at cross-correlations over a local cone of data and not the entire shell)

3D FSCs: a quantitative way to evaluate directional resolution and anisotropy in single-particle cryo-EM

- Calculate the FSC between two half maps as a function of the direction (and looking at cross-correlations over a local cone of data and not the entire shell)

3D FSCs: a quantitative way to evaluate directional resolution and anisotropy in single-particle cryo-EM

- Calculate the FSC between two half maps as a function of the direction (and looking at cross-correlations over a local cone of data and not the entire shell)

Evaluating resolution anisotropy: the 3D FSC

Anisotropic case

Aiyer et al. (2017). Nat Methods, 14(8), 793-796.

Isotropic case

Tan et al. (2017). Nat Methods, 14(8), 793-796.
Implementation of 3D FSCs in Chimera

Aiyer et al. *in press*. MiMB.
Web-server for remote 3D FSC processing: https://3dfsc.salk.edu

Remote 3DFSC Processing Server

Required Fields

- **Jobname**
  - Jobname

- **Apix**
  - Apix

Optional Fields

- **Mask file**
  - Choose File
  - No file chosen

- **Cone angle**
  - 20.0

- **FSC Cutoff**
  - 0.143

- **Sphericity Threshold**
  - 0.5

- **High-pass filter (Angstrom)**
  - 150.0

Submit Job Reset Form

www.github.com/LyumkisLab/3DFSC
Outline

- Preferred orientation — the origin of resolution anisotropy
- Resolution anisotropy in cryo-EM reconstructions (caused by preferred orientation) and its evaluation with 3D FSCs
- Combatting anisotropy using tilted data collection strategies
- The (unexpected) effect of sampling (in)homogeneity on global resolution
- Evaluating sampling (in)homogeneity using the SCF criterion
- Considerations for high-resolution
Tilting the grid in the microscope helps to alleviate resolution anisotropy caused by preferred specimen orientation.

Tan et al. (2017). Nat Methods, 14(8), 793-796.
Why do tilts work?
Filling in more Fourier space density (better Fourier sampling)
Tilting ameliorates resolution anisotropy and improves maps.
Tilting ameliorates resolution anisotropy and improves maps of HIV-1 intasome complexes (from Thursday ACA session).

Anisotropic reconstruction of untilted data.

(more) Isotropic reconstruction of tilted data.

Passos, Li et al. (2020). Science
Examples of tilted data collection strategies in the literature

~100 MDa Nuclear Pore Complex


~43 kDa Protein Kinase A

Possible limitations associated with tilted data collection strategies

1. Defocus gradient across the micrograph

2. More beam-induced movement

3. Thicker ice
Compensating for limitations of tilts using computational image processing: AAV

Working hypothesis: most of the problems associated with tilts could be ameliorated using computational strategies
Outline

- Preferred orientation — the origin of resolution anisotropy
- Resolution anisotropy in cryo-EM reconstructions (caused by preferred orientation) and its evaluation with 3D FSCs
- Combatting anisotropy using tilted data collection strategies
- The (unexpected) effect of sampling (in)homogeneity on global resolution
- Evaluating sampling (in)homogeneity using the SCF criterion
- Considerations for high-resolution
High-resolution dataset, with multiple “preferred orientations” for characterizing the effects of sampling

75K particles

Dataset could be subdivided into four groups by orientation

- 75K particles
- 10K particles
- 10K particles
- 10K particles

10K particles orientation 1
10K particles orientation 2
10K particles orientation 3
Reconstructions differing only in orientation distributions have distinct SSNR profiles.
Outline

- Preferred orientation — the origin of resolution anisotropy
- Resolution anisotropy in cryo-EM reconstructions (caused by preferred orientation) and its evaluation with 3D FSCs
- Combatting anisotropy using tilted data collection strategies
- The (unexpected) effect of sampling (in)homogeneity on global resolution
  - Evaluating sampling (in)homogeneity using the SCF criterion
- Considerations for high-resolution
Decoupling sampling with noise variance

Previously ...

\[
SSNR(\vec{k}) = N \frac{1}{2k N(\vec{k})}
\]


Derivation decouples a sampling factor from the per-particle SSNR

\[
SSNR(\vec{k}) = N \frac{SCF}{2k N(\vec{k})} = \frac{N}{2k} \times SCF \times \frac{1}{N(\vec{k})}
\]

Dataset signal power / noise power at different spatial frequencies

Average number of times lattice point \( k \) is sampled

Per-particle SSNR

\[
SCF \equiv \frac{1}{< 1/(2k \text{ sp})>}
\]

Geometrical factor
Evaluating SCF behavior using different Fourier sampling schemes

Decrement of the SCF for side-like views: (fully sampled)

$\frac{8}{\pi^2} = 0.81$
Decrement of the SCF for top-like views: 
(Incompletely sampled)
Knowing the SCF can predict the attenuation of the SSNR: side-like distributions, varying phi-angle modulation.
Knowing the SCF can predict the attenuation of the SSNR: side-like distributions, varying phi-angle modulation

<table>
<thead>
<tr>
<th>sampling map</th>
<th>multiplicative shift from SCF</th>
<th>multiplicative shift from SSNR (HA)</th>
<th>multiplicative shift from SSNR (Apo ferritin)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Side</td>
<td>0.81</td>
<td>0.85</td>
<td>0.80</td>
</tr>
<tr>
<td>Modulation λ=0.4</td>
<td>0.75</td>
<td>0.76</td>
<td>0.73</td>
</tr>
<tr>
<td>Modulation λ=0.6</td>
<td>0.65</td>
<td>0.64</td>
<td>0.62</td>
</tr>
<tr>
<td>Modulation λ=0.8</td>
<td>0.48</td>
<td>0.49</td>
<td>0.47</td>
</tr>
<tr>
<td>Modulation λ=1.0</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
</tr>
</tbody>
</table>
Knowing the SCF can predict the attenuation of the SSNR: top-like distributions, varying cone size with 3% “sprinkles”
Knowing the SCF can predict the attenuation of the SSNR: top-like distributions, varying cone size with 3% “sprinkles”

<table>
<thead>
<tr>
<th>sampling map</th>
<th>multiplicative shift from SCF (formula)</th>
<th>multiplicative shift from SCF (numerical)</th>
<th>multiplicative shift from SSNR (HA)</th>
<th>multiplicative shift from SSNR (Apoloferitin)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5° cone, 3% random</td>
<td>0.03</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>30° cone, 3% random</td>
<td>0.06</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>45° cone, 3% random</td>
<td>0.10</td>
<td>0.12</td>
<td>0.11</td>
<td>0.07</td>
</tr>
</tbody>
</table>
GUI for calculating SCF values:
Example with top-like distribution, untilted and tilted

SCF* = 0.02

SCF* = 0.49

Baldwin, P. R., & Lyumkis, D. (2020). bioRxiv, 3(1), 2020.06.08.140863.


www.github.com/LyumkisLab/SamplingGui
Outline

- Preferred orientation — the origin of resolution anisotropy
- Resolution anisotropy in cryo-EM reconstructions (caused by preferred orientation) and its evaluation with 3D FSCs
- Combatting anisotropy using tilted data collection strategies
- The (unexpected) effect of sampling (in)homogeneity on global resolution
- Evaluating sampling (in)homogeneity using the SCF criterion
- Considerations for high-resolution
Considerations for optimizing cryo-EM resolution

- **Microscope** (e.g. JEOL, TFS, Krios G3/G4, Arctica, etc) **microscope settings** (e.g. KV, aperture size), **microscope alignment** (coma-free, parallel illumination)
- **Ancillary hardware**: cold or warm FEG, energy filter, aberration correctors, phase plates
- **Detector and detector DQE**: Falcon4, Falcon3, K3, K2, CCD, etc
- **Detector settings**: counting vs. super-resolution, dose rate, frame rate
- **Imaging and data collection**: magnification (and pixel size), defocus range, image shift or stage position, speed vs accuracy (without compromising quality), hole selection, with or w/o stage tilt
- **Grid / data collection quality**: thick/thin ice, how much drift, sample dispersity, substrate support, grid type
- **Image processing**:
  - Software (Relion, CryoSparc, CisTEM, Warp/M, EMAN, etc)
  - Initial motion correction and CTF estimation (global, patch-based, per-particle)
  - Classification strategy
  - Post-processing (CTF refinement, particle polishing [movement, B-factor weighting], Ewald correction, density modification)
- **Fourier space sampling**
- **Sample**
Considerations for optimizing cryo-EM resolution: 1.2 Å apoferritin (cold FEG)

Cryo-EM data collection parameters

<table>
<thead>
<tr>
<th>Data collection and processing</th>
<th>Apoferritin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electron Gun</td>
<td>CFEG</td>
</tr>
<tr>
<td>Detector</td>
<td>Falcon4</td>
</tr>
<tr>
<td>Magnification</td>
<td>270k</td>
</tr>
<tr>
<td>Energy filter slit width (eV)</td>
<td>10</td>
</tr>
<tr>
<td>Voltage (kV)</td>
<td>300</td>
</tr>
<tr>
<td>Flux on detector (e/pix/sec)</td>
<td>4.5</td>
</tr>
<tr>
<td>Electron exposure on sample (e-/Å²)</td>
<td>40</td>
</tr>
<tr>
<td>Target defocus range (μm)</td>
<td>0.3-0.9</td>
</tr>
<tr>
<td>Calibrated pixel size (Å)</td>
<td>0.457</td>
</tr>
<tr>
<td>Symmetry imposed</td>
<td>0</td>
</tr>
<tr>
<td>Number of collected movies</td>
<td>3370</td>
</tr>
<tr>
<td>Initial particle images (no.)</td>
<td>428590</td>
</tr>
<tr>
<td>Final particle images (no.)</td>
<td>363126</td>
</tr>
<tr>
<td>Map resolution at FSC=0.143 (Å)</td>
<td>1.22</td>
</tr>
</tbody>
</table>

Considerations for optimizing cryo-EM resolution:
1.2 Å apoferritin (warm FEG)

- Titan Krios (Schottky FEG)
- Voltage: 300
- Monochromator + Cs corrector
- Falcon 3 detector
- 0.492 Å/pix
- Electron dose (e-/Å^2): 50
- Software: Relion
- Number of particles: 1,090,676
- Resolution: 1.25 Å

Considerations for optimizing cryo-EM resolution:
1.3 Å apoferritin (cold FEG)

- cryoARM300 (JEOL)
- Voltage: 300
- Cold FEG
- K2 detector
- In-column omega filter
- 0.49 Å/pix
- Electron dose (e-/Å²): 88
- Software: M
- Number of particles: 120,295
- Resolution: 1.34 Å

Considerations for optimizing cryo-EM resolution: 1.3 Å apoferritin (warm FEG)

Dataset details:
EFTEM NanoProbe
- 15 eV slit
- mag. x215k
- spot 4
- C2 50 um
- beam diameter 1.55 um

K3 SperRes
- phys pixel 0.40 Å
- superres pixel 0.20 Å
- exp. rate 3.42 e/physpix/s
- exp. time 3.036 s
- total exp. 64.5 e/Å²
- frames 76
- defocus range 0.2 - 0.8 um
- movies 2,391
Try to understand your limitations

First 2 e-/Å²

Exposure Dependency of the Mean Intensity of the Unmerged Integrated Reflections

Summary

• Resolution anisotropy originates from preferred orientation and inhomogeneity in Fourier sampling
• 3D FSCs provide quantitative measurements of resolution anisotropy
• Tilts help to overcome resolution anisotropy
• Inhomogeneity in Fourier sampling also causes attenuation of global resolution: incomplete sampling worse than inhomogeneous (but complete) sampling
• The Sampling Compensation Factor (SCF) estimates the amount of global resolution attenuation
• Many factors to optimize for high-resolution … toward sub-ÅÅ reconstructions!
Acknowledgements and Support

Yong Zi Tan  
Phil Baldwin  
Sriram Aiyer

Collaborators
Niko Grigorieff
Ben Himes
Mavis Agbanjie-McKenna
Robert McKenna

Scripps/Salk Consortium
Jamie Williamson
Travis Berggren
Martin Hetzer

Computational support
Phil Baldwin
JC Ducom

Cryo-EM support
Bill Anderson

Funding
Leona M. and Harry B. Helmsley Charitable Trust Grant

Salk Institute for Biological Studies
National Institute of Allergy and Infectious Diseases

NIH Directors Early Independence Award