Electron Diffraction of Biological Macromolecules

Bioinformatics and X–Ray Structural Analysis

Tim Gruene
Paul Scherrer Institute
tim.gruene@psi.ch
1 Crystals and Diffraction
1.1 Structure Determination by Single Crystal Diffraction

- Diffraction spots: interaction between wave and **crystal**

- Experimental result: **Position** and **Intensity** for each spot
1.2 Spot Position and Spot Intensity

- Spots positions according to Laue Conditions and orientation of Unit Cell:
  
  \[(\vec{S}_o - \vec{S}_i) . \vec{a} = h\]
  \[\text{and } (\vec{S}_o - \vec{S}_i) . \vec{b} = k\]
  \[\text{and } (\vec{S}_o - \vec{S}_i) . \vec{c} = l\]

- Monochromatic wave: \(\vec{S} = (S_o - S_i)\) can be calculated from experimental geometry

- Spot position \(\Leftrightarrow\) Crystal lattice

- Spot intensity \(\Leftrightarrow\) Unit cell content
1.3 Data Collection ... → ... Structure Refinement

- Structure determination: atom coordinates refined against idealized amplitudes $|F_{\text{ideal}}(hkl)|$

- Relationship amplitudes and intensities: $|F_{\text{ideal}}(hkl)|^2 \propto I_{\text{ideal}}(hkl)$

- Detector signal = experimental intensity $I_{\text{exp}}(hkl)$

- Data processing: from detector signal to amplitudes
1.4 Data Processing and Scaling

For X–rays*:

\[ I_{\text{exp}}(hkl) = \frac{e^4}{m_e^2 c^4} \frac{\lambda^3 V_{\text{crystal}}}{V_{\text{unit cell}}} I_0 L P T E \cdot I_{\text{ideal}}(hkl) \]

**Data Integration**  Extraction of \( I_{\text{exp}} \) from detector: intensity counts after background subtraction — largely independent from radiation source

**Data Scaling**  Conversion from \( I_{\text{exp}} \) to \( I_{\text{ideal}} \): reduction of experimental errors, crystal shape, detector properties, . . . — depends on type of radiation

2 Types of Radiation

For atomic structure solution by crystallography:

1. X–rays

2. neutrons

3. electrons
3 Differences between Types of Radiation

1. Calculation of $|F_{\text{calc}}(hkl)|$ from atom coordinates

2. Conversion from $I_{\text{exp}}(hkl)$ to $|F_{\text{ideal}}(hkl)|$

3. X–rays and neutrons: $|F_{\text{ideal}}(hkl)| \propto \sqrt{I_{\text{exp}}(hkl)}$
3.1 Types of Radiation — X-rays

1. most advanced (pipelines from data collection to structure refinement)

2. typical wavelength: \( \lambda = 0.8\text{–}1.9\text{Å} \)

3. standard structure determination

4. PDB (Protein Data Base):
   - 80,000 X-ray structures
   - 80 neutron structures
   - 60 electron structures

Annual Growth of the PDB (X-ray)
3.2 Types of Radiation — neutrons

1. (virtually) no radiation damage

2. requires large crystals ($\geq 1\text{mm}^3$)

3. visualisation of hydrogen atoms

4. adjacent elements (e.g. $K^+$ vs. $Cl^−$, $Zn^{2+}$ vs. $Cu^+$)

5. structure determination from radiation sensitive samples (Photosystem II)

PDB ID 2ZOI: D/H exchange in $\beta$–strand

3.3 Types of Radiation — electrons

1. strong interaction compared with X–rays: good for very small crystals ($\ll 1\,\mu m$ thickness)

2. typical wavelength: $200\,keV = 0.0251\,\text{Å}$: flat Ewald sphere

3. charge enables electron optics: imaging and diffraction

4. new phasing possibilities

Diffraction of nanocrystals
(van Genderen et al., Acta Cryst A72 (2016))

Inset: HIV to scale, courtesy Thomas Splettstoesser, en.wikipedia.org
3.4 Goal of Diffraction Experiment

- Fit molecule into density $\rho(x, y, z)$ to determine atomic structure

- $\rho(x, y, z) = \sum_{h,k,l} |F_{\text{ideal}}(hkl)| e^{i\phi(hkl)} e^{-2\pi i (hx + ky + lz)}$
3.5 Crystallographic Maps

- After phasing, diffraction data provide density maps $\rho(x, y, z)$

- The type of map depends on the interaction

<table>
<thead>
<tr>
<th>Radiation</th>
<th>Interaction</th>
<th>Map type</th>
</tr>
</thead>
<tbody>
<tr>
<td>X–ray</td>
<td>$e^-$</td>
<td>electron density map</td>
</tr>
<tr>
<td>n</td>
<td>nucleus</td>
<td>nucleic “density” map</td>
</tr>
<tr>
<td>$e^-$</td>
<td>$p + e^-$</td>
<td>Coulomb potential $\approx$ electron density map</td>
</tr>
</tbody>
</table>

- Macromolecules can be built into the maps “as usual”
4 Applications for Electron Diffraction

1. Diffraction & Radiation Damage

2. Nanocrystals have less Defects

3. Powder contains Single Nanocrystals

4. Seemingly failed Crystallisation Attempts contain Nanocrystals
4.1 X–rays Scattering and Electron Scattering

- Probability of inelastic scattering: $10^{-4}$
- Deposited energy: 10keV
- Probability of elastic scattering: $10^{-5} = 10^{-4}/10$
- Damage per diffracted photon: 100keV
4.2 X–rays Scattering and Electron Scattering

- X–rays (10keV)
  - Probability of inelastic scattering: $10^{-4}$
  - Deposited energy: 10keV
  - Probability of elastic scattering: $10^{-5} = 10^{-4}/10$
  - Damage per diffracted photon: 100keV

- $e^-$ (200keV)
  - Probability of inelastic scattering: 30%
  - Deposited energy: 20eV
  - Probability of elastic scattering: 10%
  - Damage per diffracted electron: 60eV = 0.06keV

2,000 times more damage with X–rays
4.3 X–rays Scattering and Electron Scattering

- Small Crystals very radiation sensitive
- X–rays mostly pass through (99.99%): beamstop
- X–rays mostly damage (10:1).
4.4 Small Crystals

**X–rays**

You *can* measure nanocrystals. You need

- Free Electron Laser (XFEL, SwissFEL, . . .)
- 10,000 – 100,000 crystals, $V \approx 5 \text{ ml}$
- Special Software, Computational Demands

**Electrons**

You *must* measure nanocrystals. You need

- Electron Microscope
- 1–2 nanocrystals
- standard software (XDS, SHELX, Refmac5, . . .)
- Direct Pixel Detector helps (Timepix, Dectris Eiger, . . .)
4.5 Joint Venture: Free Electron Lasers and Electron Diffraction

- Beamtime for FELs will be very competitive

- Only few end stations available

- Electron Microscopes are “more abundant”

- Sample quality can be pre-assessed with Electron Diffraction

  ⇒ FEL have more time to time–resolved studies

  ⇒ Electron diffraction enhances the through-put of FELs
4.6 Applications: Better Ordered Crystals

- “long” range disorder
- worse with larger crystals
- worse with freezing
- nanocrystals: often better defined spots
4.7 Applications: Single Crystals

- Powder samples often contain single nanocrystals suitable for electron diffraction

- Usually too small for conventional crystallography

- Highly interesting for the pharmaceutical industry
4.8 Applications: Crystals at all!

- macromolecules are difficult to crystallise

- in particular: membrane proteins

- Large fraction of clear drops actually contains nanocrystals —Stevenson et al., PNAS (2014) 111, 8470–8475
5 Instruments for Electron Diffraction
5.1 Electron Microscopes
5.2 Direct Pixel Detectors

Monolithic direct electron detector:
- damage prone
- Small point spread
- Low dynamic range

Hybrid pixel detector:
- radiation hard
- Larger point spread
- High dynamic range

Ideal for imaging

Ideal for diffraction

(Courtesy Prof. Abrahams)
5.3 Direct Pixel Detectors

Direct Pixel Detectors have no electronic noise, only background scattering.

Cross-section with spots

Cross-section without spots
5.4 The Timepix Detector

- Timepix assembly:
- ASI read–out
- Electronics outside vacuum
- Peltier cooling of detector $\pm 0.1K$
- $512 \times 512$ and $1024 \times 1024$ pixel versions
- linear: $1–10,000 \ e^- / frame$
- read–out: up to 120 frames /s
- radiation hard
5.5 Electron Microscope: Imaging Mode

Plane Wave  Object  Lense  Image Plane (Detector)

Rays of \textit{equal origin} focus on detector
5.6 Electron Microscope: Imaging Mode

Detector noise and radiation sensitivity require low contrast images

5.7 Electron Microscope: Diffraction Mode
5.8 Electron Microscope: Diffraction Mode

Image at Backfocal Plane = \| \text{Fouriertransform of object} \| \\

If object = crystal: 

diffraction spots according to Laue conditions:

Rays of equal direction focus on detector
6 Example Data and Example Structures


2. Lysozyme (Manuscript in preparation)
6.1 Carbamazepine

- Drug for epilepsy
- Small organic compound $C_{15}H_{12}N_{2}O$
- Well known structure used as test case (El Hassan et al., Crystal Growth and Design (2013), 13, 2887–2896)
6.2 Carbamazepine Data

Crystal Size: $1.2 \times 0.8 \times 0.2$
Rotation range: $51^\circ$
Dose: $4.0e^-/\text{Å}^2$
Space group: $P2_1/n$
Resolution: $8.7$–$0.8$ (0.85–0.80)Å
Completeness: $45\%$ (46%) $R_{\text{merge}}$: $8.4\%$ (35.8%) $I/\sigma_I$: 5.6 (1.8) $R_{\text{complete}}$: 28.0 %
6.3 Carbamazepine Structure Solution

- Solved with direct methods, *i.e.* no chemical information
- No atoms missed, no atoms too many
- Only 4 wrongly assigned atom types
6.4 Carbamazepine Structure Solution

Low data completeness affects map quality despite atomic resolution
6.5 Solving Macromolecular Data: Lysozyme

- Currently PDB holds three entries from 3D electron diffraction


2. 5A3E: Lysozyme, $P2_1$, 2.5Å [Nannenga et al. Nat. Meth. (2014) 11, 927]

3. 3J7U: Catalase, $P2_12_12_1$, 3.2Å [Yonekura et al. PNAS (2015) 112, 3368–3373]

Structures 1+2 are collected from μ crystals; Structure 3 was solved by merging data from 99 crystals.
6.6 Data from a single Lysozyme nanocrystal

- Collected 40° before radiation damage destroyed crystal
- Crystal thickness ≈ 100nm
- Data processed with RED and with XDS in $P1$

<table>
<thead>
<tr>
<th></th>
<th>RED</th>
<th>XDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell</td>
<td>32.3Å 69.7Å 105.6Å</td>
<td>32.1Å 70.9Å 104.0Å</td>
</tr>
<tr>
<td></td>
<td>93.6° 92.0° 90.1°</td>
<td>93.4° 91.9° 91.1°</td>
</tr>
<tr>
<td>Resolution</td>
<td>32.3–2.4 (2.5–2.4)</td>
<td>32.1 –2.2 (2.3–2.2)</td>
</tr>
<tr>
<td>$I/\sigma_I$</td>
<td>21.2 (8.8)</td>
<td>6.7 (1.4)</td>
</tr>
<tr>
<td>$R_{\text{merge}}$</td>
<td>7.9% (13.7%)</td>
<td>28.9% (49.8%)</td>
</tr>
<tr>
<td>Completeness</td>
<td>4.1% (0.1%)</td>
<td>20.7% (20.8%)</td>
</tr>
<tr>
<td># refl.</td>
<td>1897</td>
<td>14148 (2571)</td>
</tr>
<tr>
<td># unique refl.</td>
<td>1568</td>
<td>9542 (1539)</td>
</tr>
</tbody>
</table>
6.7 Spacegroup of Lysozyme nanocrystal

- Cell: $32.1 \times 70.9 \times 104.0, 93.4^\circ 91.9^\circ 91.1^\circ$

- XDS suggests: $P 2 \overline{1} 1$

- PDB ID 4R0F: $P2_12_12$ with $104.63 \times 66.49 \times 31.65$

Possible explanations:

1. $\alpha$ angle distorted because of erroneous parameters (distance, frame width, rotation range, image distortions)

2. Macrocrystal induces more rigid packing $\Rightarrow$ enforces higher symmetry

$\Rightarrow$ Currently an open question
6.8 Lysozyme: Model Bias

Refined map from Refmac5

Refined map from Shelxl
6.9 Lysozyme: is it Real? (I)

Refined map from Shelxl (zoomed)  Same map 4x NCS averaged
6.10 Lysozyme: is it Real? (II)

- Purple: Molecular replacement including side chains
- Green: Molecular replacement with poly-Ala model; side chains autobuilt with Buccaneer
- Autobuilding uses sequence information and data.
- Many side chains consistent
7 Phasing with Images
7.1 The Crystallographic Phase Problem

\[
\rho(x, y, z) = \sum_{h,k,l} |F_{\text{ideal}}(hkl)| e^{i\phi(hkl)} e^{-2\pi i (hx + ky + lz)}
\]

- Diffraction experiment measured amplitudes \( |F_{\text{ideal}}(hkl)| \)
- Phases \( \phi(hkl) \) “get lost”
- Phasing methods:
  1. Molecular Replacement
  2. SAD/MAD — Single-/Multi–wavelength anomalous dispersion
  3. SIRAS — Isomorphous replacement with anomalous dispersion
7.2 Electron Microscope Imaging

1. Record *many* images

2. Classify, group, and reduce noise

3. Find orientations

4. Reconstruct 3D electron density

(EMDB 3281, *A chimeric sapovirus capsid*)
7.3 Indexing Diffraction Data

- Diffraction Data can be indexed

⇒ Unit Cell Dimensions and (often) Space group are known without solving the structure

⇒ Place single atom at unit cell corners and create projections from “single atom map” from all orientations
7.4 EM Imaging from Crystals

Projected Lysozyme Density
210Å × 210Å

Projected Single Atom Density
210Å × 210Å
Same Orientation
Match: 3.6%

Projected Single Atom Density
210Å × 210Å
Different Orientation
Match: 0.8%

In Image Mode, Contrast between 3.6% and 0.8% too low
(Simulated Data)
7.5 EM Imaging from Crystals in Fourier Space

Image of Lysozyme Crystal after Fourier Transform
210Å × 210Å

Projected Single Atom Density after Fourier Transform
210Å × 210Å
Same Orientation
Match: 65%

Projected Single Atom Density after Fourier Transform
210Å × 210Å
Different Orientation
Match: 14%

Contrast after Fourier Transformation enables selection of correct orientation
(Simulated Data)

February 2016
Uni Konstanz
8 Phasing with EM Images I

Images

FFT

Match Orientation

Orientations for e.g. 100 images:
Reconstruction

Low resolution 3D map

Diffraction

Indexing

32Å 70Å 104Å

Calculate Single Atom Projections

FFT
9 Phasing with EM Images II

- Phases from Electron imaging **low resolution** but **accurate**
- Phases can be extended to high resolution

3D map

phases

phase extension

high resolution map

low resolution
accurate phases

high resolution intensities
10 Experimental Considerations

- Ewald “plane”
- Dynamic scattering
- Instrumental limitations
10.1 X–rays: The Ewald Sphere

\[ \lambda = 1 \text{Å}, \text{“normal” resolution: } 2\theta_{\text{max}} = 40^\circ \]

- Assume: wrong detector distance
- Diffraction spot calculated wrongly (red circle)
- Reciprocal lattice becomes distorted

Curvature of the Ewald sphere gauges the diffraction geometry
10.2 Electrons: The Ewald “Plane”

- Typical X–ray wavelength $\lambda_X = 1\text{Å}$
- Typical $e^-$ wavelength $\lambda_e = 0.025\text{Å}$
- Radius of Ewald sphere 40x greater
- Ewald sphere nearly flat
10.3 Electrons: The Ewald “Plane”

\[ 1/\lambda_e = 1/0.025\text{Å} \]

“normal” resolution:

\[ 2\theta_{\text{max}} = 1^\circ \]

- opening angle of highest resolution reflections \( \approx 1^\circ \)
- Ewald sphere virtually flat
- Without curvature: impossible to refine both detector distance and cell
10.4 Electrons: The Ewald “Plane”

- Detector distance and unit cell parameters are strongly related
- Wrongly set distance can lead to incorrect bond lengths
- Distance refinement with X–ray data routine
- Distance refinement with electron data = unstable
- Distance calibration from powder sample
10.5 Distance Calibration

- Bragg's law: \( \lambda = 2d \sin \theta \); \( d, \lambda \) are known
10.6 Dynamic Scattering

- Kinematic Theory of Diffraction: Every photon / electron / neutron scatters once in the crystal

\[ |F_{\text{ideal}}(hkl)| \propto \sqrt{I_{\exp}(hkl)} \]

- Dynamic Scattering: Multiple Scattering events occur

- Electron Diffraction: Multiple Scattering occurs even with nanocrystals
10.7 Dynamic Scattering

- Outgoing ray $\vec{S}_1^1$ acts as incoming ray for reflection $\vec{S}_o^2$.
- Re-reflection with 10% probability at 50 nm path length.
10.8 Dynamic Scattering

Laue Conditions (accordingly \( \vec{b} \) and \( \vec{c} \)):

\[
\begin{align*}
(\vec{S}_1^o - \vec{S}_i) \cdot \vec{a} &= h_1 \\
(\vec{S}_2^o - \vec{S}_1^o) \cdot \vec{a} &= h_2 \\
(\vec{S}_2^o - \vec{S}_i) \cdot \vec{a} &= h_1 + h_2
\end{align*}
\]

Experimental Intensities by superposition of two reflections:

\[
I_{\text{exp}}(h_2k_2l_2) = \left| F_{\text{ideal}}(h_2k_2l_2) + \alpha F_{\text{ideal}}(h_1k_1l_1) \right|
\]

- \( \alpha < 1 \): 0.1 for 50 nm path length
- \( (h_1k_1l_1) \) strong and \( (h_2k_2l_2) \) weak \( \Rightarrow \) wrong estimate for \( |F_{\text{ideal}}(h_2k_2l_2)| \)
- affects high resolution data
10.9 Dynamic Scattering for Organic Crystals

- Presence in Macromolecular Diffraction data currently discussed in literature
- Some claim it is negligible
- Experimental evidence equivocal
- Treatment (scaling / refinement) should be improved
10.10 Other Instrumental limitations

- Electron Microscopes not designed for accurate sample rotation

- Rotation axis not linked to Camera read–out

- Lense system rotates (diffraction) image: rotation axis unknown

- Sample holder not desiged for 180° rotation
10.11 SwissFEDI

Swiss Free Electron Diffraction Instrument

- Horizontal beam

- 15–18m instrument length
  1. Reduced Cross-talk between magnetic lenses
  2. No optical enlargement of detector distance: $1^\circ - 2^\circ$ opening angle covers $20 \times 20cm^2$ detector area

- Sample holder designed for sample rotation
11 Acknowledgements

- Wei Wan (Stockholm)
- Kay Diederichs (Konstanz)
- Jan Pieter Abrahams (Basel / PSI)
- George Sheldrick (Göttingen)