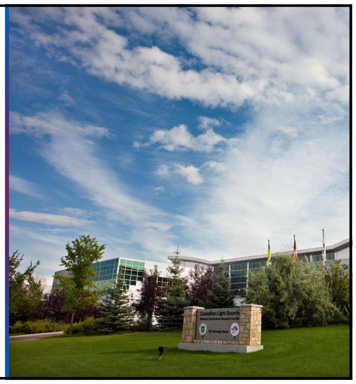
10th Annual
CLS MX Data Collection School
VIRTUAL EDITION

# Data Collection Strategies A practical Guide

Michel Fodje

Canadian Light Source · Centre canadien de rayonnement synchrotron



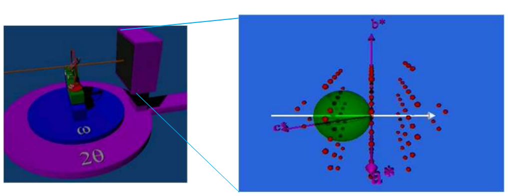
## Outline

- The goal of data collection
- Steps involved
- Data collection parameters
- Practical tips on collecting good quality data

Canadian Light Source · Centre canadien de rayonnement synchrotron

2

## Data Collection – Rotation Method




Real Space

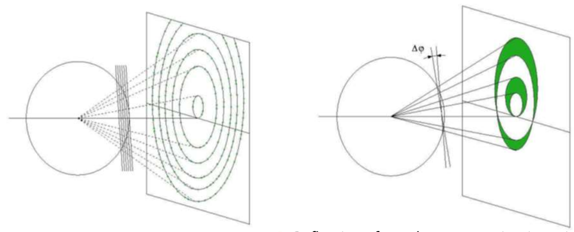
Reciprocal Space

Canadian Light Source · Centre canadien de rayonnement synchrotron

11/12/2020

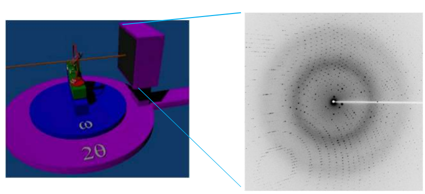
3

## Data Collection – Ewald Sphere




- Reflections from the same region in reciprocal lattice form a lune as crystal rotates

## Data Collection - The Rotation Method




11/12/2020

## Preamble

- Data collection is the last experimental step in structure determination
- All down-stream steps will *benefit* or *suffer* from choices made during data collection
- It is not always possible to repeat it
- Therefore
  - ***It is very important to carefully select parameters in order to measure the best data possible on the crystal.***
  - ***Goal is to obtain best possible statistics whilst using as few photons as possible.***

11/12/2020

## What is the best data possible?

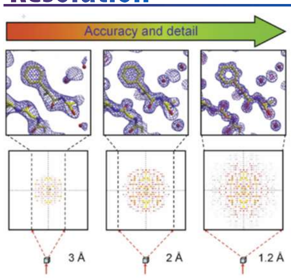
- Achieve highest resolution the crystal will diffract
- Measure 100% of all possible reflections to that resolution
  - An incomplete high-resolution dataset can be worse than a complete lower resolution dataset.
- Achieve highest possible data accuracy and precision
  - within the constraints of factors beyond your control.
- Collect data as efficiently as possible

- Resolution
- Completeness
- Mosaicity
- Overlaps
- Overloads
- Radiation damage
- Background

11/12/2020 Canadian Light Source Centre canadien de rayonnement synchrotron

THE BRIGHTEST LIGHT IN CANADA | lightsource.ca 7

## Resolution



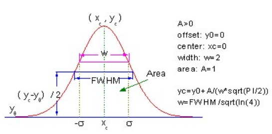

- Higher resolution reflections are weaker, more difficult to measure, and appear at wider angles
- The higher the resolution, the more detail you'll be able to see in the electron density map.

Canadian Light Source Centre canadien de rayonnement synchrotron

THE BRIGHTEST LIGHT IN CANADA | lightsource.ca 8

## Mosaicity

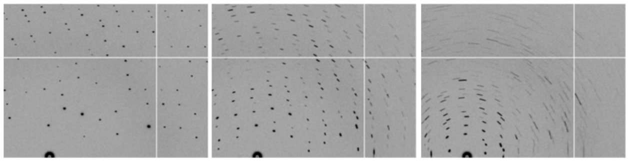
- Standard deviation/FWHM of reflection profile. *Angular range over which Bragg planes diffract.*
- The diffraction pattern of imperfect crystals is the sum of the diffraction patterns originating from mosaic blocks with slightly different orientations.
- The mosaicity of good quality protein crystals are between 0.1 – 0.5 deg
- XDS uses **σ**, HKL2000/iMosflm uses FWHM = 2.34 × σ




Canadian Light Source Centre canadien de rayonnement synchrotron

THE BRIGHTEST LIGHT IN CANADA | lightsource.ca 9

## Effect of Mosaicity on diffraction



**Mosaicity = 0.2˚** **Mosaicity = 2˚** **Mosaicity = 6˚**

Canadian Light Source Centre canadien de rayonnement synchrotron

THE BRIGHTEST LIGHT IN CANADA | lightsource.ca 10

## Data Collection Steps

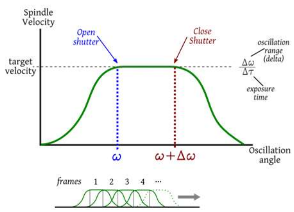
- Setup the beamline parameters
- Select and mount a crystal
- Align the crystal into the beam
- Characterize sample and determine optimal parameters
- Collect diffraction data frames according to a set of parameters
- Process the data immediately.

Canadian Light Source Centre canadien de rayonnement synchrotron

THE BRIGHTEST LIGHT IN CANADA | lightsource.ca 11

## Data Collection Parameters

- Starting ω angle
- ω - rotation range
- Δω, oscillation range per image
- Exposure time t, per image



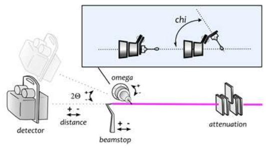

Canadian Light Source Centre canadien de rayonnement synchrotron

11/12/2020

THE BRIGHTEST LIGHT IN CANADA | lightsource.ca 12

## Data Collection Parameters

- Photon energy/wavelength
- Maximum diffraction resolution/detector Distance
- Detector swing-out (2θ)
- Beam-stop distance
- Beam Attenuation
- Spindle Offset (chi)/Kappa angle




Canadian Light Source Centre canadien de rayonnement synchrotron 11/12/2020 13 THE BRIGHTEST LIGHT IN CANADA | lightsource.ca

## Practical Tips – Selecting the Crystal

- Should be single and intact (not split).
- Do not confuse split-crystals with twins
- Collecting and processing a few frames 90° apart (screening) will reveal problems with the crystal
- If you have many crystals, try several, as crystals within the same drop may have differences in mosaicity/resolution.

Canadian Light Source Centre canadien de rayonnement synchrotron 11/12/2020 14 THE BRIGHTEST LIGHT IN CANADA | lightsource.ca

## Practical Tips – Rotation Range

Rotation range (°) required to collect a complete data set in different crystal classes

The direction of the spindle axis is given in parentheses; ac means any vector in the ac plane.

| Point group | Native data | Anomalous data |
|---|---|---|
| 1 | 180 (any) | 180 + $2\theta_{max}$ (any) |
| 2 | 180 (b); 90 (ac) | 180 (b); 180 + $2\theta_{max}$ (ac) |
| 222 | 90 (ab or ac or bc) | 90 (ab or ac or bc) |
| 4 | 90 (c or ab) | 90 (c); 90 + $\theta_{max}$ (ab) |
| 422 | 45 (c); 90 (ab) | 45 (c); 90 (ab) |
| 3 | 60 (c); 90 (ab) | 60 + $2\theta_{max}$ (c); 90 + $\theta_{max}$ (ab) |
| 32 | 30 (c); 90 (ab) | 30 + $\theta_{max}$ (c); 90 (ab) |
| 6 | 60 (c); 90 (ab) | 60 (c); 90 + $\theta_{max}$ (ab) |
| 622 | 30 (c); 90 (ab) | 30 (c); 90 (ab) |
| 23 | ~60 | ~70 |
| 432 | ~35 | ~45 |

*Dauter, Acta Cryst. (1999). D55, 1703-1717 [ doi:10.1107/S0907444999008367 ]*

- Required for complete data, depends on crystal orientation, symmetry and detector 2θ.

Canadian Light Source Centre canadien de rayonnement synchrotron 11/12/2020 15 THE BRIGHTEST LIGHT IN CANADA | lightsource.ca

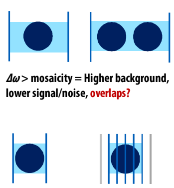
## Practical Tips – Multiplicity/Redundancy

- The average number of times a given reflection is independently measured.
- Higher multiplicity → higher signal-to-noise ratios → better data quality → GOOD
  - Improves anomalous signal (MAD/SAD)
- Increase Rotation range beyond minimum required to increase multiplicity
  - Try to get at least 4-fold multiplicity in all shells
  - ***But first make sure you will get 100% data before the crystal decays.***

Canadian Light Source Centre canadien de rayonnement synchrotron 11/12/2020 16 THE BRIGHTEST LIGHT IN CANADA | lightsource.ca

## Practical Tips - Oscillation Range (*Δω*)

- Choice of delta affects signal/noise ratio.
- "thin"-slicing (*fine-slicing*) improves signal/noise, allows more accurate refinement of spot positions (with 3D integration), more accurate anomalous signal.
- Beware of readout-noise and shutter jitter noise.
- 2D integration programs prefer "thick"-slicing.



*Δω* > mosaicity = Higher background, lower signal/noise, overlaps?

*Δω* ~ mosaicity or *Δω* < mosaicity = good signal/noise

See *M. Mueller, M. Wang, and C. Schulze-Briese, (2012) Acta Cryst.* ***D68****, 42-56.*

Canadian Light Source Centre canadien de rayonnement synchrotron 11/12/2020 17 THE BRIGHTEST LIGHT IN CANADA | lightsource.ca

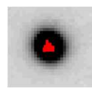
## Practical Tips - Oscillation Range (*Δω*)

**Δω=0.2˚** **Δω=1˚** **Δω=2˚ Overlaps**

Canadian Light Source Centre canadien de rayonnement synchrotron 11/12/2020 18 THE BRIGHTEST LIGHT IN CANADA | lightsource.ca

## Practical Tips – Exposure Time & Attenuation

- The higher the exposure time, the higher the reflection intensities
- Higher exposure times → higher signal-to-noise ratios → better data quality → GOOD
  - Doubling exposure, improves signal to noise by √2
  - Longer exposure reduces impact of high frequency noise in beam intensity and position.
- Higher exposure → faster radiation damage → BAD
- Higher exposure → overloaded detector pixels → BAD
- Attenuate the beam if the spots are overloaded and the exposure time is already low.
- ***It is better to collect a 100% complete data of lower intensity than to collect incomplete data with high intensities and high signal-to-noise ratios.***





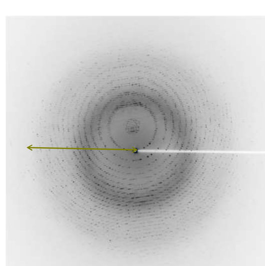
## Practical Tips – Detector Distance/ Offset

- Determines the maximum resolution spots recordable
- Longer distance → better signal-to-noise ratio:
  - The background area increases as square of distance
  - Reflection profiles increase less.
- Spots are better resolved at longer distances. May need to sacrifice resolution to get more complete data.
- 2θ Offset may be used to increase resolution limit without reducing distance.
- Avoid 2θ for MAD/SAD anomalous data
  - Anomalous completeness suffers.
- Minimum rotation range required for complete data is higher if 2θ is used.

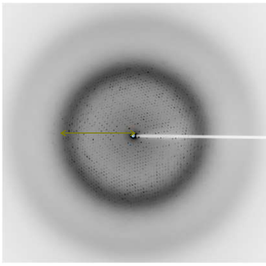


## Practical Tips – Detector Distance

- visible diffraction spots should cover at least ~80% of detector radius but not extend to the edges
- Use , zoom, brightness and contrast adjustments to inspect images.



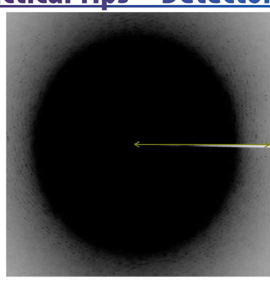
## Practical Tips – Detector Distance



- Detector too close
  - ***move back.***



## Practical Tips – Detector Distance

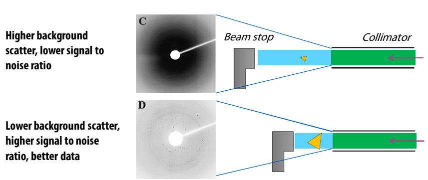


- Detector too far
  - ***move in.***



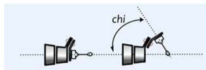
## Practical Tips – Beam-stop distance

- Placing beam stop closer to crystal reduces background scatter and improves signal.
- But may affect low resolution data



## Practical Tips – Spindle Offset

- Kappa & Chi Angles
- Used to offset the crystal in order to collect complete data in the "blind-region"
  - Blind region depends only on diffracting angle (resolution and wavelength).
- Used to reorient crystal (cell-axis along the spindle) to avoid overlaps
- Can be used to align symmetry axis along spindle to collect Bijvoet pairs on same image.
  - For optimizing anomalous signal



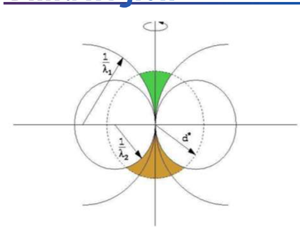


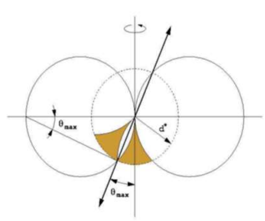

## Practical Tips – Energy / Wavelength

- Crucial for Anomalous data (SAD/MAD)
- For MAD/SAD, always use Fluorescence Scan rather than theoretical peak and inflection point values.
- Also determines maximum resolution measurable,
  - High energy (short wavelength) = Higher resolution,
  - Low energy (long wavelength) = lower resolution for same detector distance.
- Air absorption increases exponentially with decrease in energy (long wavelength).
- Most detectors are less efficient at lower energies.



## Blind Region





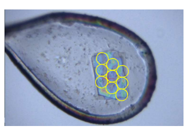
- Reorienting Omega axis can help
- Blind region is smaller with shorter wavelength (higher energy)

## Practical Tips – Other factors

- Beam size, beam divergence, photon flux*.
  - Not always within your control
- Sample temperature:
  - Room temperature (if your crystals degrade on freezing)
  - 100 K (Nitrogen vapor)
- Helium cone can be used to reduce air-absorption at low energies.

## Selecting the best parameters for your Crystal

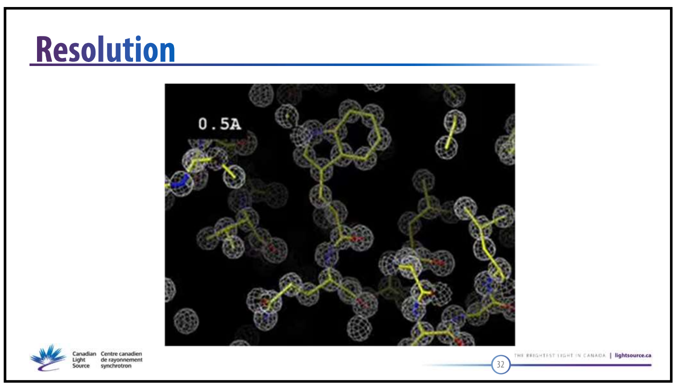
- Select and Characterize the best crystal among many
- Determine Parameters (Strategy) for data collection.
- Collect a few frames 90-degrees apart
- Inspect the images for multiple/split crystals , overloads, detector coverage
- Process them using a strategy program (AutoProcess, Mosflm, HKL2000) etc.
- Raster Screening:
  - Select the best part of a large crystal to collect on with a small beam.

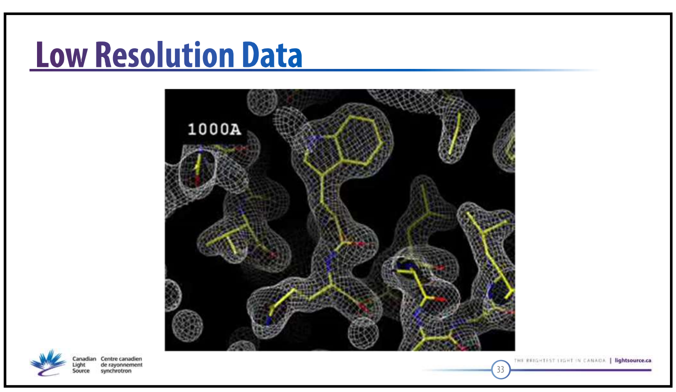


## Finally

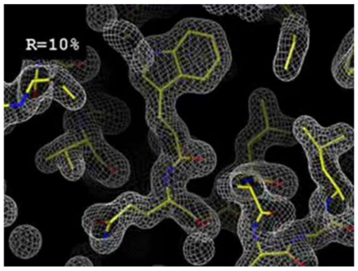
- Start processing your data as soon as you have enough images
  - Helps identify problems/errors in parameters before you have used all your time or radiation-damaged your crystal.
- Fully process your data before you leave the synchrotron
- Back-up your data (both raw and processed)
- ***Acknowledge the beamline/Synchrotron if you use the data in anyway to help solve the structure, even if it is not the final data you deposit.***


James Holton's Videos
https://bl831.als.lbl.gov/~jamesh/movies/


Resolution
0.5A

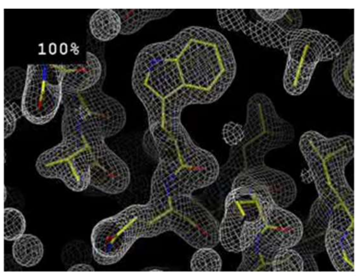

Low Resolution Data
1000A

## Data Quality – R-factor



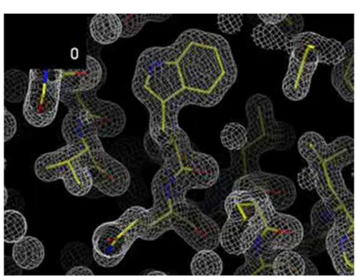

34

## Completeness

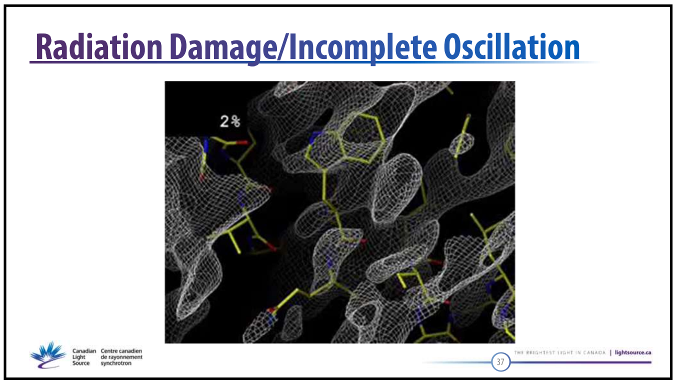



35

## Overloads




36


Radiation Damage/Incomplete Oscillation
2%
37
lightsource.ca