Albert-Einstein-Str. 15 D-12489 Berlin, Germany

Helmholtz-Zentrum Berlin für Materialien und Energie

Macromolecular Crystallography (HZB-MX)

D-12489 Berlin, Germany msweiss@helmholtz-berlin.de

Data Collection Strategies

Manfred S. Weiss











Data Collection Strategies





or

How to Avoid Collecting Suboptimal X-ray Diffraction Data ?









Manfred S. Weiss

Helmholtz-Zentrum Berlin für Materialien und Energie Macromolecular Crystallography (HZB-MX) Albert-Einstein-Str. 15 D-12489 Berlin, Germany msweiss @helmholtz-berlin.de

Structure Determination











The collection of diffraction data is the last real **experiment** that is conducted before the determination and refinement of the structure.

The factors involved in diffraction data collection are complex and require some thought in order to produce the highest quality data set possible.



The quality of the diffraction data ultimately determines the quality of the resulting structure.





Two objectives:



to collect a complete data set







to collect the best possible data set

















h, k, I Miller indices I(h,k,l) intensity σI(hkl) error in I



















HZB-MX BL14.2 MARCCD-225 $\lambda = 0.918 \text{ Å}$ F = 150 mm 90 images 0.5°/ image 2 sec / image

Thaumatin, space group P4₁2₁2, a=50 Å, c= 150 Å





Crystal





Diffraction as Reflection (Bragg)



















$$a^* = 1/a, a^* \perp b$$

 $b^* = 1/b, b^* \perp a$

d* = 1 a* + 1 b*

(11)

 $a^* = 1/a, a^* \perp b$

 $b^* = 1/b, b^* \perp a$

















d* = 1 a* + 2 b*

(12)

 $a^* = 1/a, a^* \perp b$

 $b^* = 1/b, b^* \perp a$

















d* = 1 a* + 3 b*

(13)



















$$a^* = 1/a, a^* \perp b$$

 $b^* = 1/b, b^* \perp a$





 $a^* = b \times c / V$, $a^* \perp b$, c, $a^* \bullet a = 1$

 $b^* = a \times c / V$, $b^* \perp a, c, b^* \bullet b = 1$

 $c^* = a \times b / V$, $c^* \perp a, b$, $c^* \bullet c = 1$



 $V^* = 1 / V$





The Purpose of the Reciprocal Lattice



BioStruct-X Course, 02.09.2013, Budapest, Hungary

HZ

elmholtz

Zentrum Berlin













every set of planes (characterized by the normal vector d) corresponds to a reciprocal lattice point d* (d* = 1 / d)





every set of planes (characterized by the normal vector d) corresponds to a reciprocal lattice point $d^* (d^* = 1 / d)$









 $\mathbf{A}^* = \mathbf{h} \cdot \mathbf{a}^* + \mathbf{k} \cdot \mathbf{b}^* + \mathbf{I} \cdot \mathbf{c}^*$ (h, k, I: Miller indices)





every set of planes (characterized by the normal vector d) corresponds to a reciprocal lattice point d* (d* = 1 / d)



 $\mathbf{A}^* = \mathbf{h} \cdot \mathbf{a}^* + \mathbf{k} \cdot \mathbf{b}^* + \mathbf{I} \cdot \mathbf{c}^*$ (h, k, I: Miller indices)







every set of planes (d) corresponds to an X-ray reflection





every set of planes (characterized by the normal vector d) corresponds to a reciprocal lattice point d* (d* = 1 / d)









every set of planes (d) corresponds to an X-ray reflection

 $\mathbf{A}^* = \mathbf{h} \cdot \mathbf{a}^* + \mathbf{k} \cdot \mathbf{b}^* + \mathbf{I} \cdot \mathbf{c}^*$ (h, k, I: Miller indices)

every reciprocal lattice point (d*) corresponds to an X-ray reflection

The Ewald Construction







- only when OP is a reciprocal lattice point, a reflection can be observed
- the Ewald-construction is nothing but a graphical representation of Bragg's Law.

















On the Detector?





The Rotation Method





The Rotation Method





Diffraction data are nowadays typically collected using the rotation method and an area detector.









The Rotation Method





Diffraction data are nowadays typically collected using the rotation method and an area detector.









- the rotation method and an area detector.
- A set of diffraction images with a small rotation increment (e.g. $\Delta \phi = 1.0^\circ$) is collected and constitutes the raw data.

A Small Rotation Range





On the Detector?





Compare this to a Real Diffraction Image















The Problem of Overlaps





The Effect of Mosaicity





The Blind Region





The Blind Region





 reflections in the blind region cannot be collected by rotation around a single axis







The Blind Region




The Blind Region





reflections in the blind region cannot be collected by rotation around a single axis



the size of the blind region is determined by the maximum diffraction angle θ_{max}



The Blind Region





- P
- 0





- reflections in the blind region cannot be collected by rotation around a single axis
- * the size of the blind region is determined by the maximum diffraction angle θ_{max}
- in order to collect a complete data set, one has to collect a pass around a second rotation axis of at least 2θ_{max}

The Blind Region





- P
- 0





- reflections in the blind region cannot be collected by rotation around a single axis
- * the size of the blind region is determined by the maximum diffraction angle θ_{max}
- ✤ in order to collect a complete data set, one has to collect a pass around a second rotation axis of at least $2\theta_{max}$
- the presence of symmetry helps unless the symmetry axis lies in the blind region













Any symmetry element present in real space will also be present in reciprocal space (a screw axis in real space will become a rotation axis in reciprocal space).





Any symmetry element present in real space will also be present in reciprocal space (a screw axis in real space will become a rotation axis in reciprocal space).





In addition, the reciprocal space contains an inversion center in the origin (Friedel's law).

The Reflection Sphere



















No symmetry



Measure the whole reflection sphere



$$-h \rightarrow +h$$
$$-k \rightarrow +k$$
$$-l \rightarrow +l$$







Inversion Center (Friedel)

Measure half the reflection sphere



$$-h \rightarrow +h$$
$$-k \rightarrow +k$$
$$0 \rightarrow +l$$









Inversion Center (Friedel) + 2-fold rotation axis along c



Measure a quarter of the reflection sphere



$$-h \rightarrow +h$$

$$0 \rightarrow +k$$

$$0 \rightarrow +l$$









Inversion Center (Friedel) + 4-fold rotation axis along c



Measure an octant of the reflection sphere



$$\begin{array}{c} 0 \rightarrow +h \\ 0 \rightarrow +k \\ 0 \rightarrow +l \end{array}$$



How much do we need to collect?













Crystal Class	Point Group	Rotation Axis	Standard Data	Anomalous Data
Triclinic	1	any	180°	180° + 2 θ _{max}
Monoclinic	2	b*	90°	180°
		a*, c*	180°	180° + 2θ _{max}
Orthorhombic	222	a*, b*, c*	90°	90°
Tetragonal	4	C*	90°	90°
		a*, b*	90°	90° + θ _{max}
	422	С*	45°	45°
		a*, b*	90°	90°
Trigonal	3	C*	60°	60° + 2θ _{max}
		a*, b*	90°	90° + θ _{max}
	32	C*	30°	30° + θ _{max}
		a*, b*	90°	90°
Hexagonal	6	C*	60°	60°
		a*, b*	90°	90° + θ _{max}
	622	C *	30°	30°
		a*, b*	90°	90°
Cubic	23	any	≈ 60°	≈ 70°
	432	any	≈ 35°	≈ 45°

Z. Dauter (1997). *Meth. Enzymol.* **276**, 326-344.

Strategy













A decision has to be made where to start the data collection (φ₀), how many images to collect and which rotation increment to use (Δφ) in order to avoid overlaps.

Strategy





P





- A decision has to be made where to start the data collection (φ₀), how many images to collect and which rotation increment to use (Δφ) in order to avoid overlaps.
- Other decisions: maximum resolution, wavelength, detector distance, beamstop distance, exposure time, etc.



Ravelli et al. (1997). J. Appl. Cryst. 30, 551-554.

Alternative Indexing





In some space groups the reflections can be indexed in different ways.









Alternative Indexing





In some space groups the reflections can be indexed in different ways.



It is important to take this into account, when trying to collect a complete data set from more than one crystal.



Alternative Indexing













Space Group	Reindexing Transformation	Reindexing Matrix	
P4, P4 ₁ , P4 ₂ , P4 ₃	hkl → kh-l	01010000-1	
14, 14 ₁	hkl → hk-l	10001000-1	
P3, P3 ₁ , P3 ₂	hkl → -h-kl	-1000-10001	
	or hkl → kh-l	01010000-1	
	or hkl → -k-h-l	0-10-10000-1	
R3	hkl → kh-l	01010000-1	
P321, P3 ₁ 21, P3 ₂ 21	hkl → -h-kl	-1000-10001	
P312, P3 ₁ 12, P3 ₂ 12	hkl → -h-kl	-1000-10001	
P6, P6 ₁ , P6 ₂ , P6 ₃ , P6 ₄ , P6 ₅	hkl → kh-l	01010000-1	
P23, P2 ₁ 3	hkl → k-hl	010-100001	
123, 12 ₁ 3	hkl → k-hl	010-100001	
F23	hkl → k-hl	0 1 0 -1 0 0 0 0 1	

Z. Dauter (1999). Acta Cryst. **D55**, 1703-1717.

Summary – Part 1





Data Collection – Sources of Error



















Slits and Collimation











Rotation Axis (Spindle)





Rotation Axis (Spindle)





Rotation Axis (Spindle)

























Beamstop

















What Should You Do?













Before the Experiment





What kind of experiment do you want to do?

- ⇒ Native data collection for molecular replacement
- ⇒ Heavy atom derivative data collection for phasing
- ⇒ MAD or SAD data collection for phasing
- ⇒ High resolution native data collection
- ⇒ Screening for ligand binding
- ⇒ Other









Different Requirements













	Molecular replacement	Anomalous Phasing	High-resolution refinement	Ligand search
Accuracy	+	++++	++	++
Low-resolution completeness	+++	+++	++	++
Resolution	+	+	+++	++
Overall completeness	++	++	++	++
Automation	++	+	++	+++

Consider the Expected Signal Strength





SIR:

$$R = 100 \cdot \Sigma_{hkl} ||F_{PH}| - |F_{P}|| / \Sigma_{hkl} |F_{P}|$$
 15-30%

 MAD:
 $R_{anom} = 200 \cdot \Sigma_{hkl} ||I^+ - |I^-| / \Sigma_{hkl} ||I^+ + |I^-|$
 ~5%

 S-SAD:
 ~1%







Before the Experiment





What kind of experiment do you want to do?

- ⇒ Native data collection for molecular replacement
- ⇒ MAD or SAD data collection for phasing
- ⇒ Heavy atom derivative data collection for phasing
- ⇒ High resolution native data collection
- ⇒ Screening for ligand binding
- ⇒ Other







- ⇒ Trust your beamline scientist
- ⇒ Carry out a test data collection

Starting the Experiment





Check cryo-system alignment and distance














- Check cryo-system alignment and distance
- Mount crystal and center it













- Check cryo-system alignment and distance
- Mount crystal and center it
 - ⇒ Use the right cryo-protectant for the crystal
 - ⇒ Match loop size to crystal size









- Check cryo-system alignment and distance
- Mount crystal and center it



⇒ Match loop size to crystal size







⇒ consider different mounting methods







- Check cryo-system alignment and distance
- Mount crystal and center it
- Match beam size to crystal size











- Check cryo-system alignment and distance
- Mount crystal and center it
- Match beam size to crystal size
- Collect two test diffraction images and index them
 - ⇒ Inspect the diffraction pattern for split reflections or multiple lattices
 - ⇒ Check anisotropy of diffraction









- Check cryo-system alignment and distance
- Mount crystal and center it
- Match beam size to crystal size
- Collect two test diffraction images and index them
- Choose data collection wavelength
 - ⇒ Measure an X-ray fluorescence spectrum for MAD/SAD













- Check cryo-system alignment and distance
- Mount crystal and center it
- Match beam size to crystal size
- Collect two test diffraction images and index them
- Choose data collection wavelength
- Choose the detector distance
 - ⇒ Use the whole detector surface
 - \Rightarrow Rule of thumb: d_{min} = visible diffraction + 0.2 Å







- Check cryo-system alignment and distance
- Mount crystal and center it
- Match beam size to crystal size
- Collect two test diffraction images and index them
- Choose data collection wavelength
- Choose the detector distance
- Choose the total rotation range
 - Run a strategy program in order to get a complete data set as quickly as possible
 - ⇒ If you have time and the crystal does not decay, collect redundant data







- Check cryo-system alignment and distance
- Mount crystal and center it
- Match beam size to crystal size
- Collect two test diffraction images and index them
- Choose data collection wavelength
- Choose the detector distance
- Choose the total rotation range
- Choose the rotation increment
 - ⇒ Run a strategy program to avoid overlapping reflections
 - ⇒ Fine slicing *vs.* wide slicing







- Check cryo-system alignment and distance
- Mount crystal and center it
- Match beam size to crystal size
- Collect two test diffraction images and index them
- Choose data collection wavelength
- Choose the detector distance
- Choose the total rotation range
- Choose the rotation increment
- Choose exposure time and/or attenuation
 - ⇒ Avoid overloaded reflections









- Check cryo-system alignment and distance
- Mount crystal and center it
- Match beam size to crystal size
- Collect two test diffraction images and index them
- Choose data collection wavelength
- Choose the detector distance
- Choose the total rotation range
- Choose the rotation increment
- Choose exposure time and/or attenuation
- Start data collection





During the Experiment





Process the data as you collect them









During the Experiment





- Process the data as you collect them
- Try to solve your structure as soon as possible







Summary – Part 2







The collection of diffraction data is the last real **experiment** that is conducted before the determination and refinement of the structure.

The factors involved in diffraction data collection are complex and require some thought in order to produce the highest quality data set possible.



The quality of the diffraction data ultimately determines the quality of the resulting structure.













Thank you for your attention