PAUL SCHERRER INSTITUT

PSI

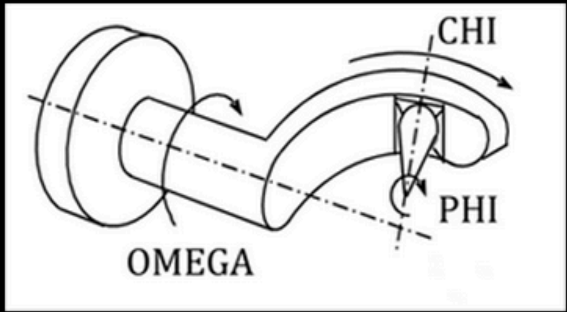
Workshop on Advanced Data Collection with Multi-Axis Goniometry

Wir schaffen Wissen – heute für morgen

Paul Scherrer Institut, Swiss Light Source

Meitian Wang

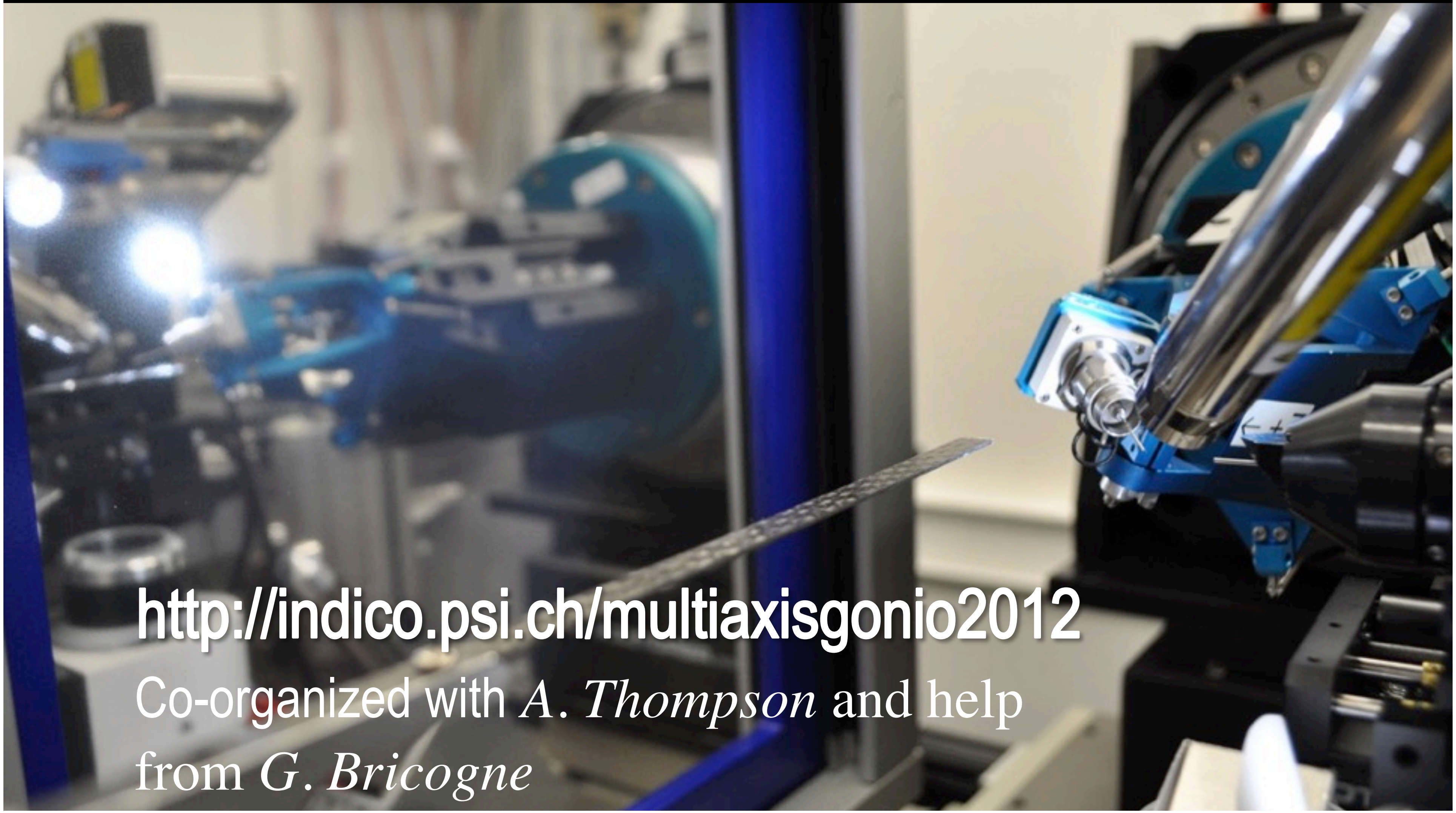
A Better Data Collection Strategy: You Can Always Use a Lower Dose



Workshop on Advanced Data Collection with Multi-Axis Goniometer and Single-Photon Counting Detector

6-8 November 2012 *Paul Scherrer Institut*

Europe/Zurich timezone



http://indico.psi.ch/multiaxisgonio2012

Co-organized with *A. Thompson* and help from *G. Bricogne*

Three PX Beamlines at the Swiss Light Source

Beamline	PXI (X06SA)	PXII (X10SA)	PXIII (X06DA)
Source	U19	U19	2.9T Superbend
Energy range	6.0 – 17.5 keV	6.5 – 20.0 keV	5.5 – 17.5 keV
Flux, phs/s (12.4 keV, focused beam)	2×10^{11} <–> 2×10^{12}	2×10^{12}	5×10^{11}
Beamsize, μm^2 (with focusing, slits)	**2 × 1 <–> 100 × 100 (fast beam size change)**	50 × 10 30 × 10, 20 × 10, 10 × 10	80 × 45 μm^2
Goniometer	Micro-diffractometer (SmarGon)		**Multi-axis, PRIGo (SmarGon)**
Detector	**EIGER 16M**	**PILATUS 6M**	**PILATUS 2M**
Data collection time	2 – 3 minutes		
Sample changer	IRELEC CATS		
Industrial usage	15%	50%	40%

Geometric Data Collection

Unique volume and unique reflections

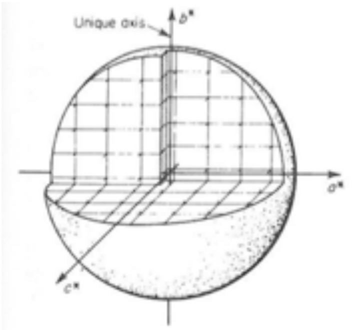


Figure 6.6. Unique volume in reciprocal space for a monoclinic crystal.

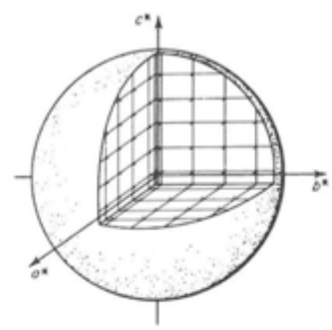
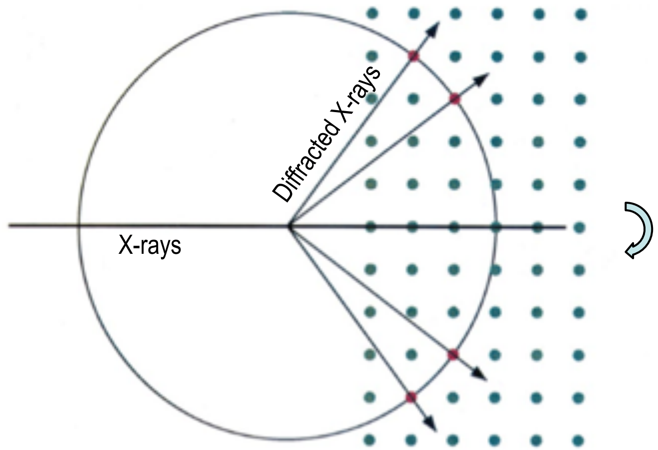


Figure 6.7. Unique volume in reciprocal space for an orthorhombic crystal.



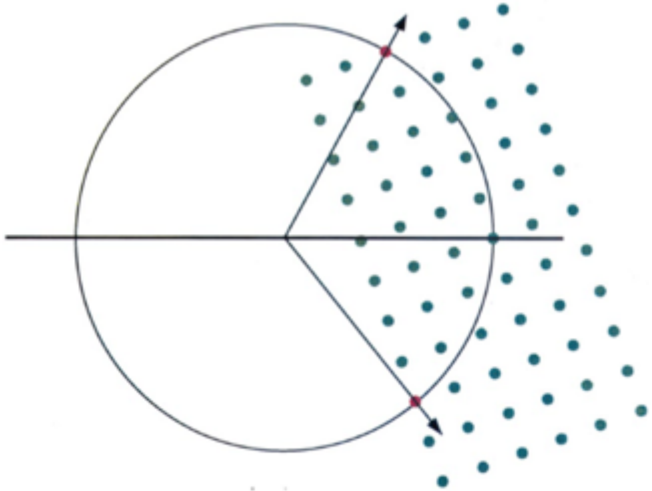


Table 1
Rotation range (°) required to collect a complete data set in different crystal classes.

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

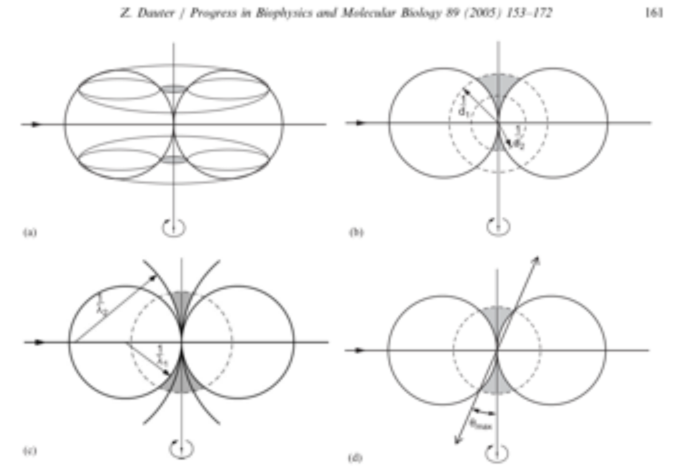
Point group	Native data	Anomalous data
1	180 (any)	180 + $2\theta_{max}$ (any)
2	180 (*b*); 90 (*ac*)	180 (*b*); 180 + $2\theta_{max}$ (*ac*)
222	90 (*ab* or *ac* or *bc*)	90 (*ab* or *ac* or *bc*)
4	90 (*c* or *ab*)	90 (*c*); 90 + θ_{max} (*ab*)
422	45 (*c*); 90 (*ab*)	45 (*c*); 90 (*ab*)
3	60 (*c*);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

Rotation method and rotation range

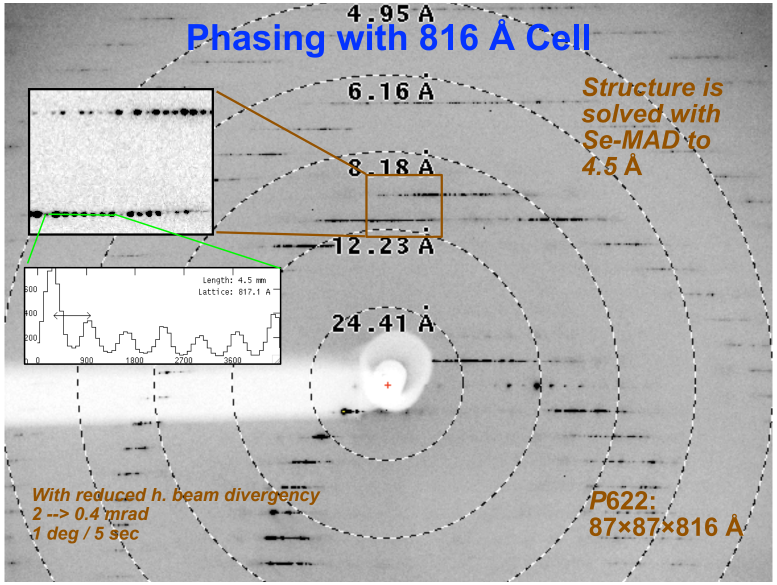
Stout and Jensen (1989), Dauter, *Acta Cryst. D***55,** 1703 (1999)

Geometric Data Collection with Multi-Axis Goniometry

- **Avoid blind region**

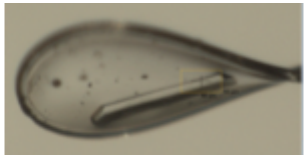


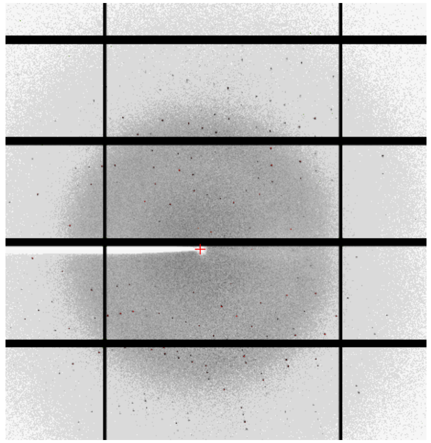
- **Align long unit-cell axis to avoid spots overlaps**



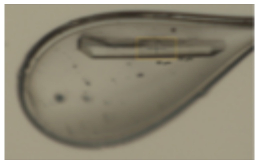
- **Align even-fold rotation axis to record Bijvoet pairs on the same image**

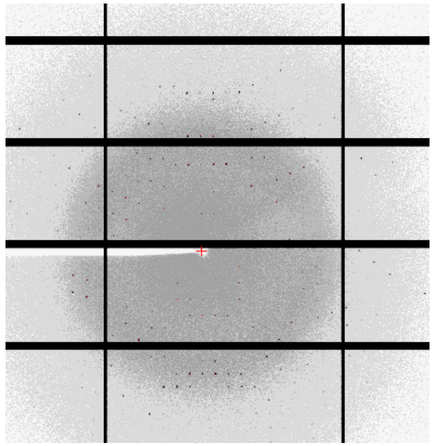
Misaligned:





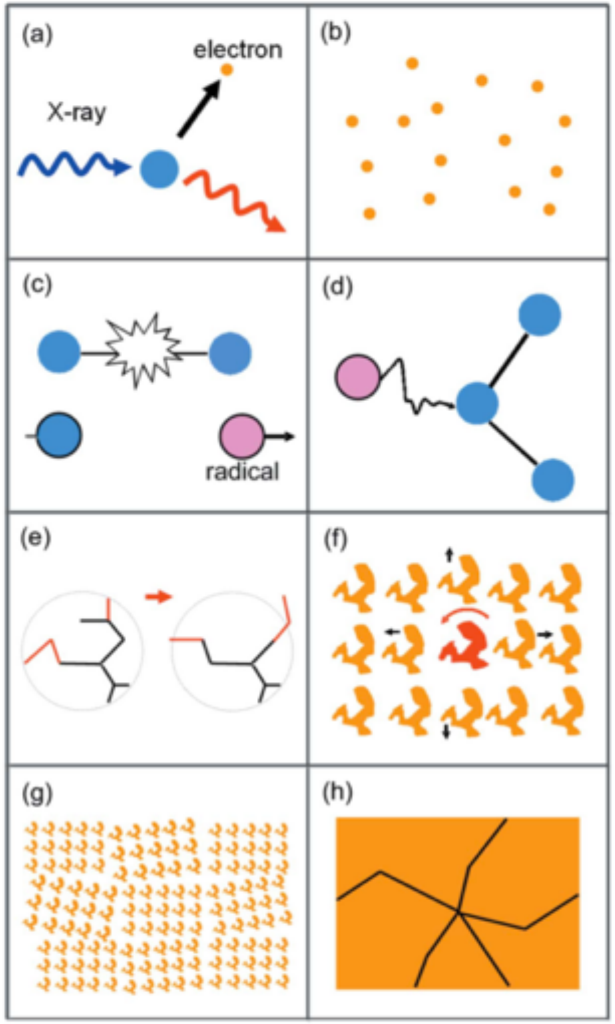
Aligned to twofold:





(XOAlign, Stratcalc)

Radiation Damage



M. Warkentin *et al. J. Synchrotron Rad.* **20,** 7 (2013)

Native data collection (20 MGy)

- Henderson, *Proc. R. Soc. B.* **241**, 6 (1990)
- Owen, *et al. Proc. Natl. Acad. Sci. USA*, **103**, 4912 (2006)

Experimental phasing (< 5 MGy)

- Holton, J. M. *J. Synchrotron Rad.* **14**, 51 (2007)
- Olieric, *et al.* *Acta Cryst.* **D63**, 759 (2007)

Rule of thumb

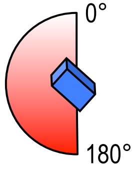
- Resolution dependency of 10 MGy / Å, Howells *et al. J. El. Spect. & Rel. Phen.* **170**, 4 (2009)
- Scaling *B* factor *vs.* dose of ~ 1 *B*-factor / MGy, Kmetko *et al. Acta Cryst.* **D62**, 1030 (2006)
- Does estimation, Holton, *J. Synchrotron Rad.* **16**, 133 (2009)

  $Dose = (t_{expo} \times flux) / (k_{dose} \times I_{H\text{-}beam} \times I_{B\text{-}beam})$

  $k_{dose} = 2000\lambda^{-2}$

"Traditional" Data Collection Strategy

High-dose low multiplicity data collection strategy
(180° data within dose limit)



Intensity Data Collection: Counting Statistics

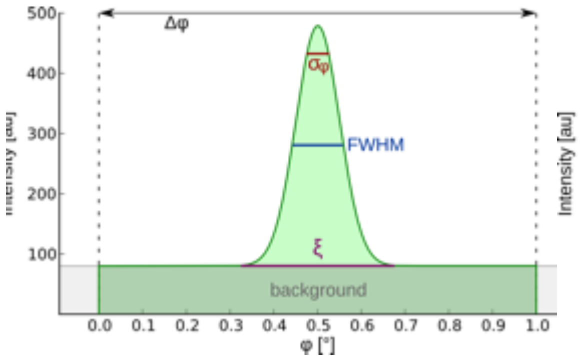
Random errors, counting statistics

$\sigma_{count} = N^{1/2}$

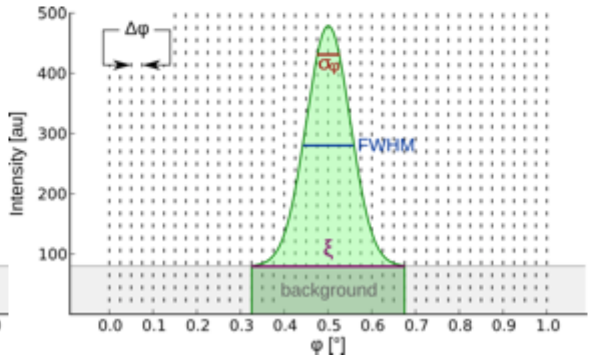
$I = N_p - N_b$ Signal is the difference

$\sigma_I = (\sigma_p{}^2 + \sigma_b{}^2)^{1/2}$

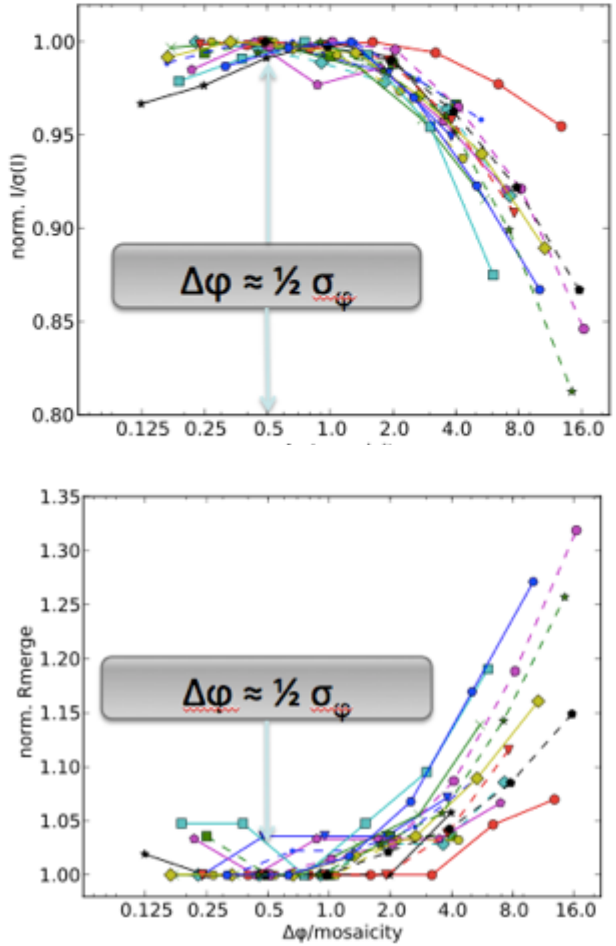
$\sigma_I = (N_p + N_b)^{1/2}$ Uncertainty is the sum



Wide φ-slicing
- Large $\Delta\varphi$ ($\Delta\varphi > \xi$)
- Large overlap of reflections and background along φ
- Few images



Fine φ-slicing
- Small $\Delta\varphi$ ($\Delta\varphi << \xi$)
- Minimal overlap of reflections and background along φ
- Many images

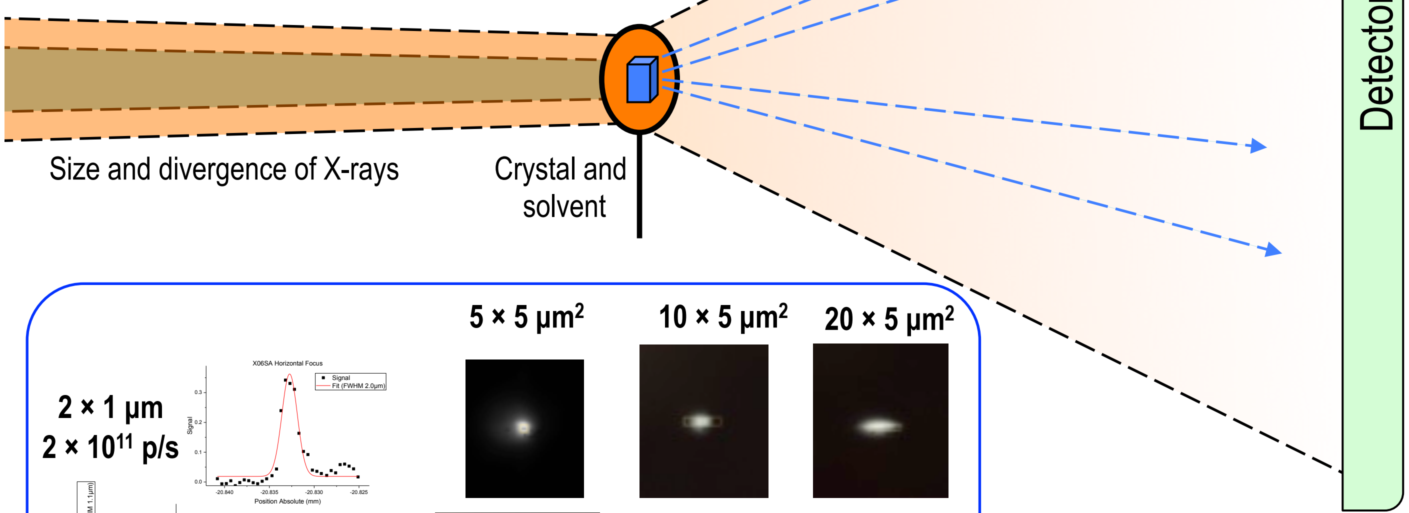


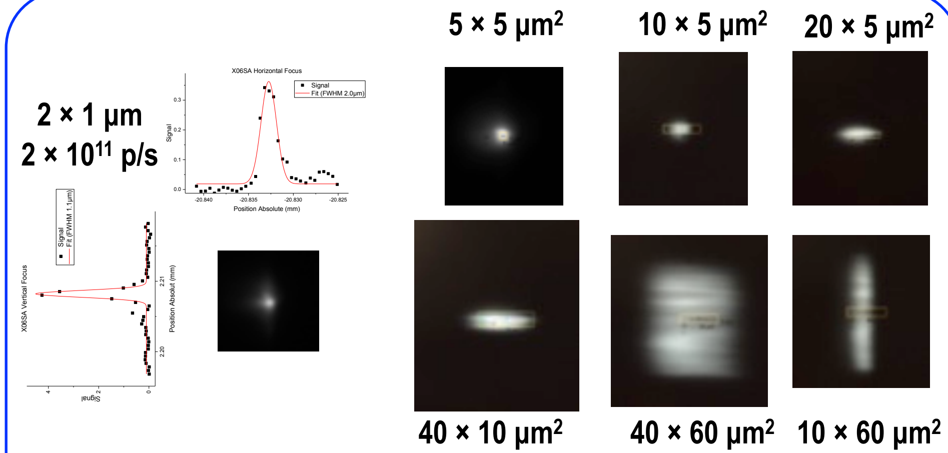
Fine-phi slicing data collection is enabled by the pixel array detector (PILATUS, EIGER), which has single-photon sensitivity and no readlout noise

Mueller, *et al*, *Acta Cryst*. (2012) **D68**, 42

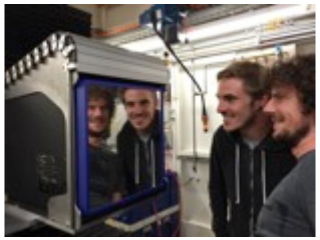
Intensity Data Collection: Reduce Background

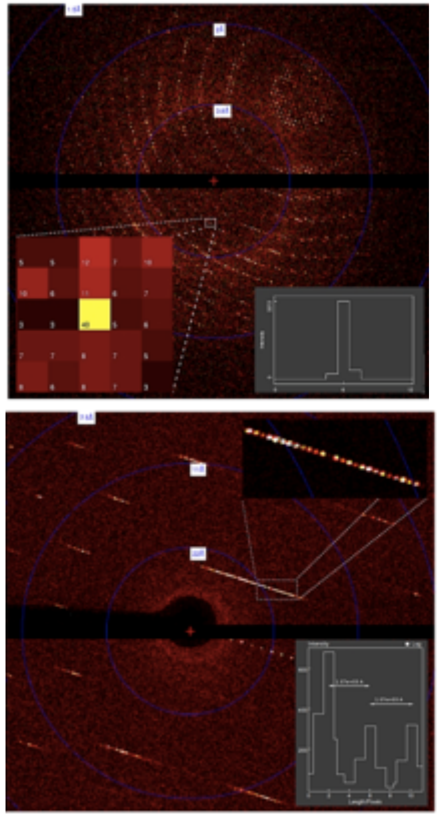
- Match beam size to crystal size
- Use best part of crystal
- Match detector distance to diffraction resolution
- Remove extra solvent
- Focus beam on detector





Variable beam size from 1 to 200 μm at beamline PXI, SLS





Spot resolution with EIGER 16M detector (75 μm pixel) at beamline PXI, SLS

Precision and Accuracy of the Intensity Measurement

- *Precision* of measurements tells us how much they differ *from each other*
- To increase *precision* → repetitive measurements (high multiplicity) within radiation damage limit
- *Accuracy* of measurements tells us how much they differ *from the truth*, but the truth is not known
- To improve *accuracy* → reduce systematic errors: beam stability, diffraction geometry, crystal orientation, absorption and variation, detector non-ideality, …
- <u>All data quality indicators measure precision, not accuracy!</u>

Diederichs, K. *Methods in Molecular Biology*, **1320**, 147 (2015)

Intensity Data Collection: Systematic Errors

Random errors, counting statistics

$\sigma_{count} = N^{1/2}$

$I = N_p - N_b$ Signal is the difference

$\sigma_I = (\sigma_p{}^2 + \sigma_b{}^2)^{1/2}$

$\sigma_I = (N_p + N_b)^{1/2}$ Uncertainty is the sum

Systematic errors, instrumental instability

$\sigma_{sys} = k_1 I$ (typical range of k_1 is 0.001 to 0.01)

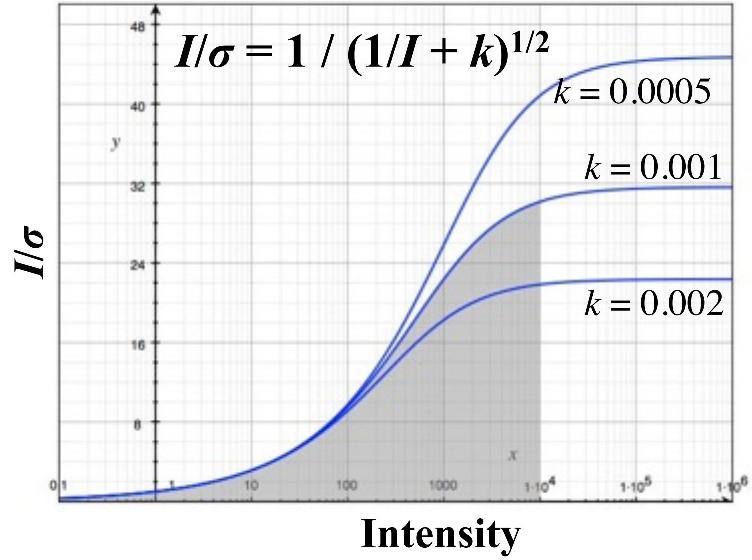
$\sigma = (\sigma_{count}{}^2 + \sigma_{sys}{}^2)^{1/2}$

$\sigma_{count} \sim I^{1/2}$; $\sigma_{\mathrm{sys}} \sim I$

$\sigma = (I + kI^2)^{1/2}$

$I/\sigma = I / (I + kI^2) = 1 / (1/I + k)^{1/2}$

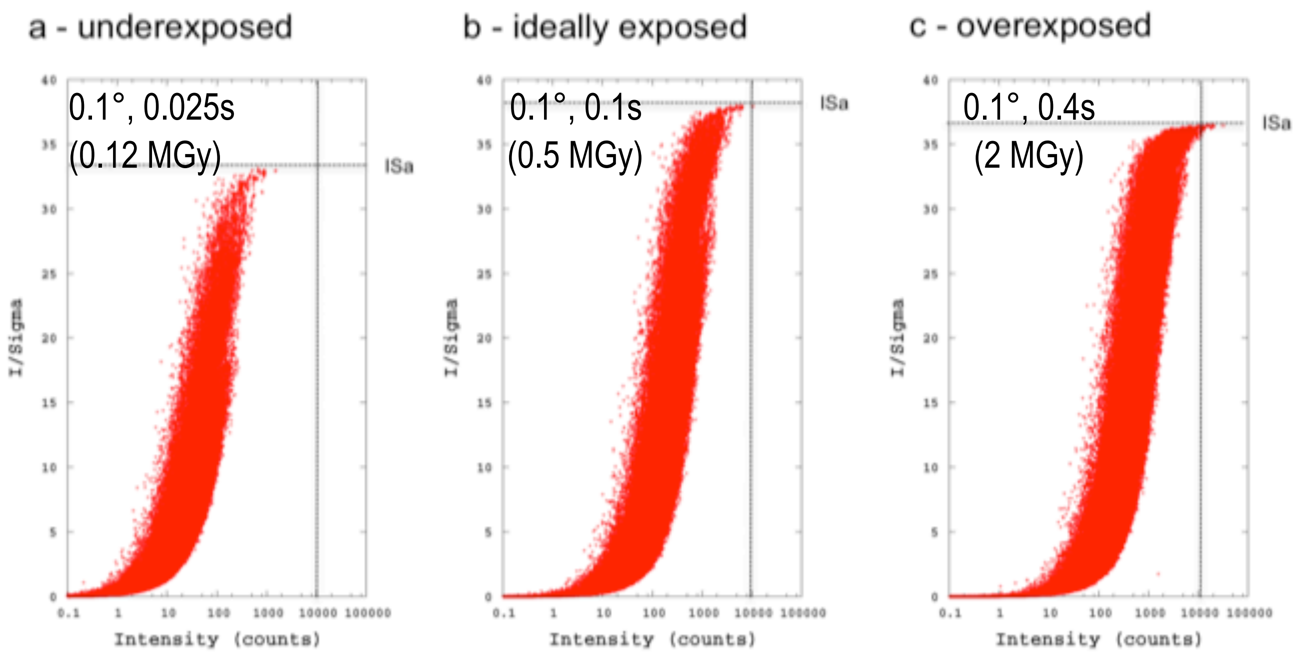
Asymptotic behavior of *I/σ* (*ISa*)



Achievable *I/σ* is limited by systematic errors, *ISa*

Diederichs, *Acta Cryst*. **D66,** 733 (2010)

Best exposure? You Can Always Use a Lower Dose



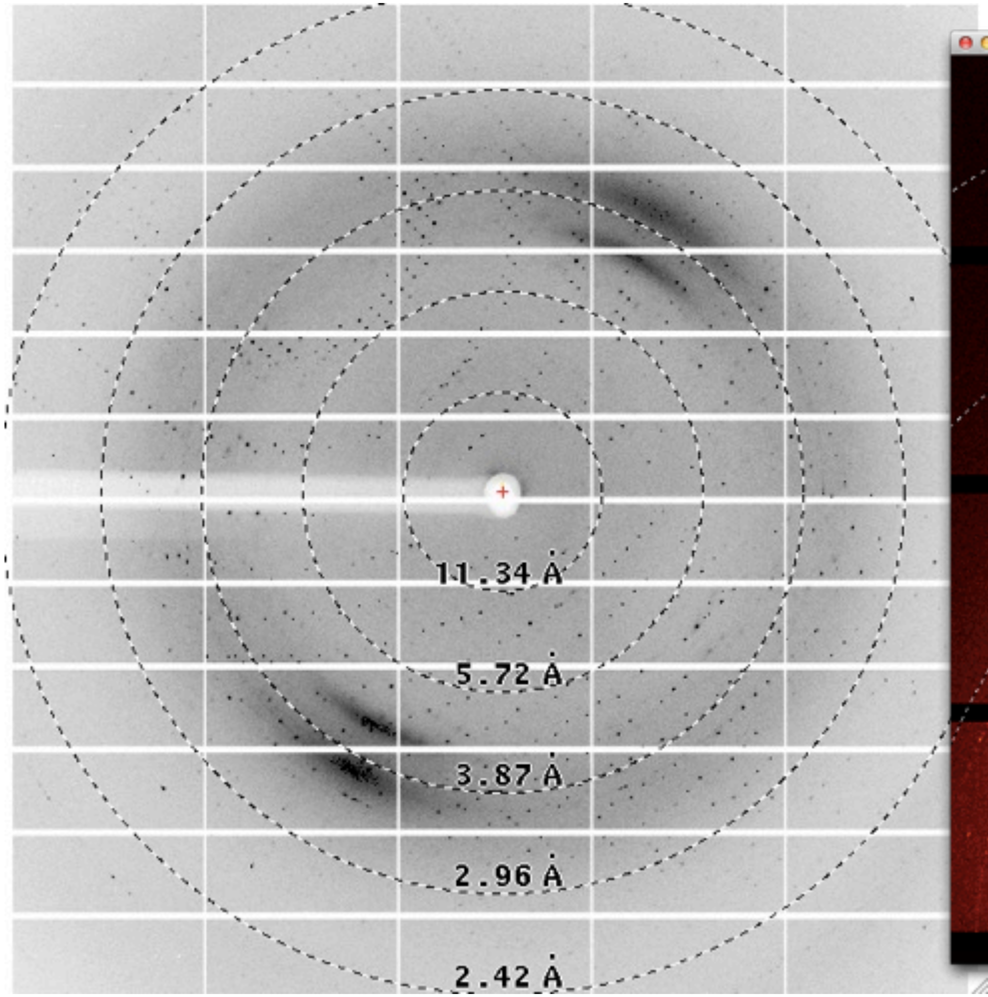
$(I/\sigma)^{asymptotic}$ plots of data sets of 180° collected at 6 keV with 1.5×10^{10} photons/s

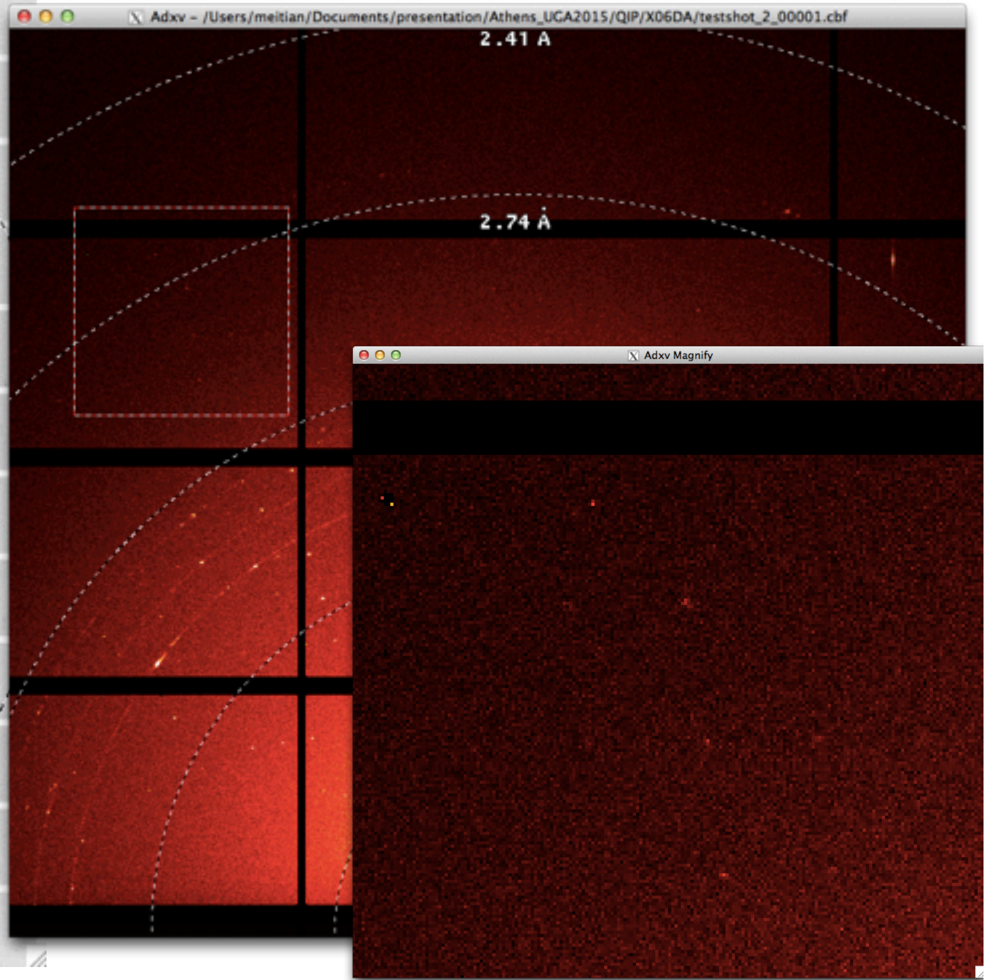
- Achievable *I/σ* at low-med resolution (phasing) is limited by systematic error
- Achievable *I/σ* at high resolution (refinement) is limited by counting statistics
- Low *ISa* means high systematic error
- Good datasets have *ISa* > 30, bad ones < 10

Diederichs, *Acta Cryst*. **D66,** 733 (2010)

PAUL SCHERRER INSTITUT PSI

Lower Dose Example



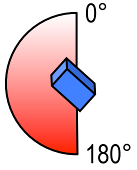


1 sec exposure with 1.5 × 10^{11} ph/sec
1° oscillation at beamline X10SA, SLS

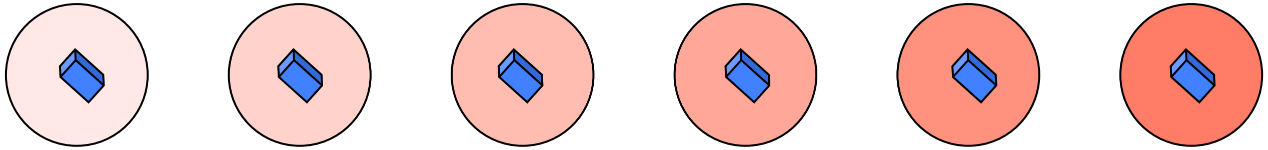
1 sec exposure with 1.5 × 10^{10} ph/sec
0.5° oscillation at beamline X06DA, SLS

Dose-slicing Data Collection Strategy

High-dose low multiplicity data collection strategy
(180° data within dose limit)

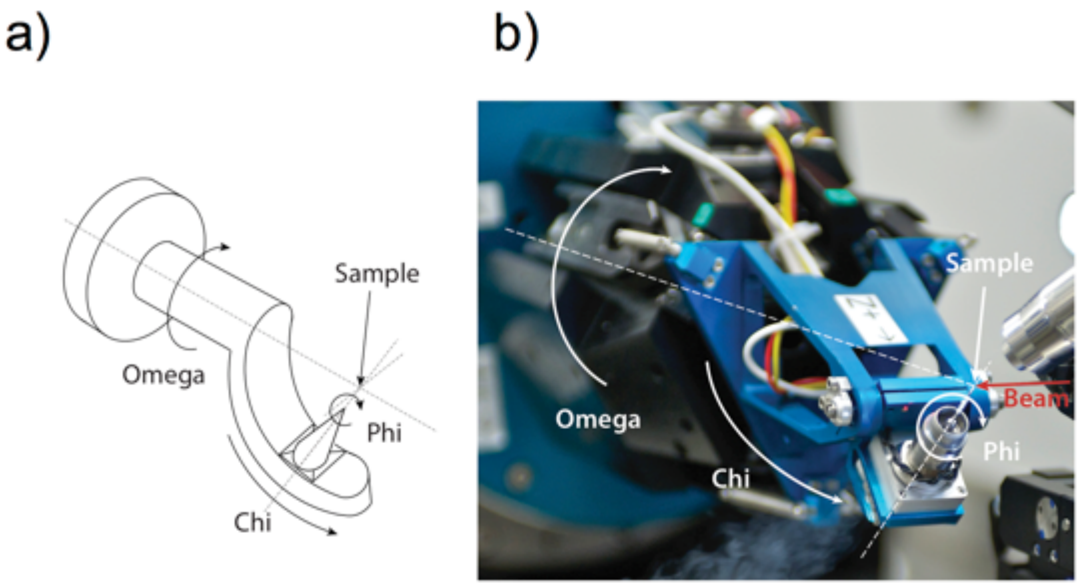


Dose-slicing high multiplicity data collection strategy, Liu *et al. Acta Cryst.* **A67**, 544, (2011)
(*n* × 360° data within dose limit)



Intensity Data Collection: Reduce Systematic Errors

Multi-orientation data collection with multi-axis goniometer



Manipulating crystal orientation changes

1) **Diffraction geometry**
2) **X-ray absorption path**
3) **Position of reflections on detector**

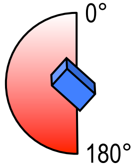
Sampling diffraction geometry, absorption path, and detector surface to average out systematic measurement errors

Waltersperger *et al. J. Synchrotron Rad.* **22**, 895 (2015).

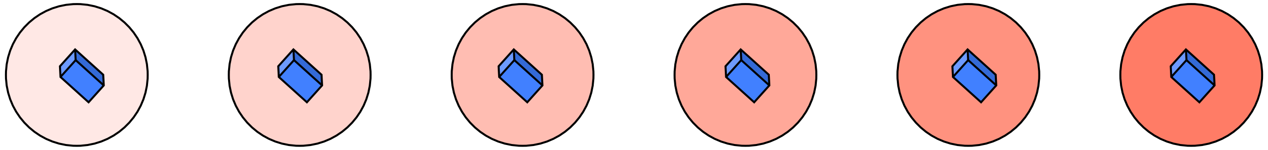
Weinert *et al. Nature Methods* **12,** 131 (2015)

Multi-Orientation Data Collection Strategy

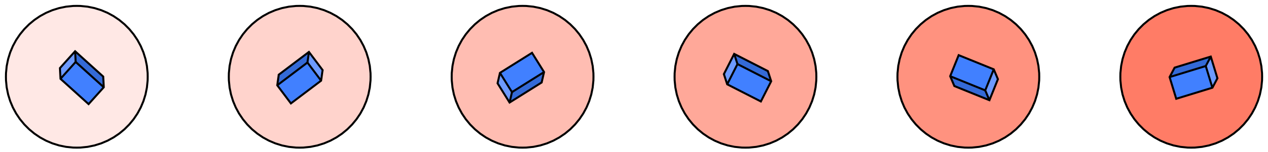
High-dose low multiplicity data collection strategy
(180° data within dose limit)



Dose-slicing high multiplicity data collection strategy, Liu *et al. Acta Cryst.* **A67**, 544, (2011)
(*n* × 360° data within dose limit)

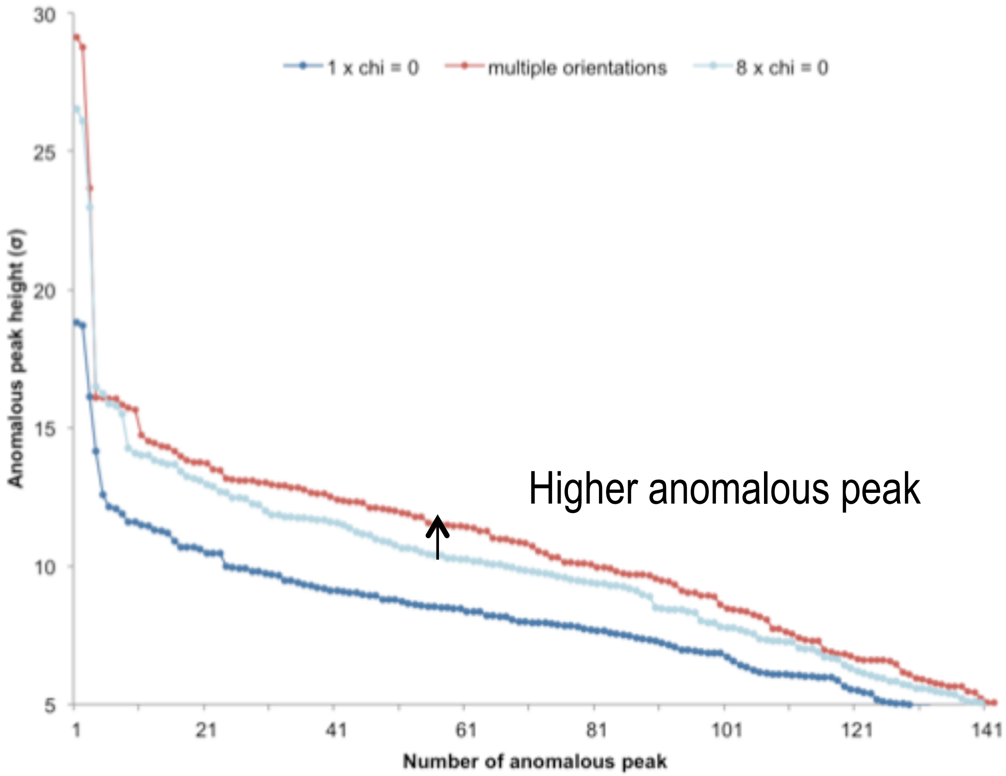


Multi-orientation high multiplicity data collection strategy, Weinert *et al. Nature Methods* **12**, 131, (2015)
(*n* × 360° data in multiple crystal orientation within dose limit)



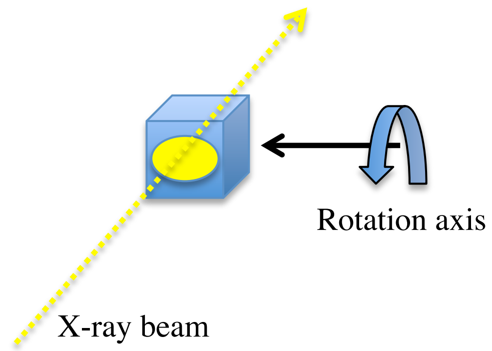
Advantage of Multi-Orientation Strategy

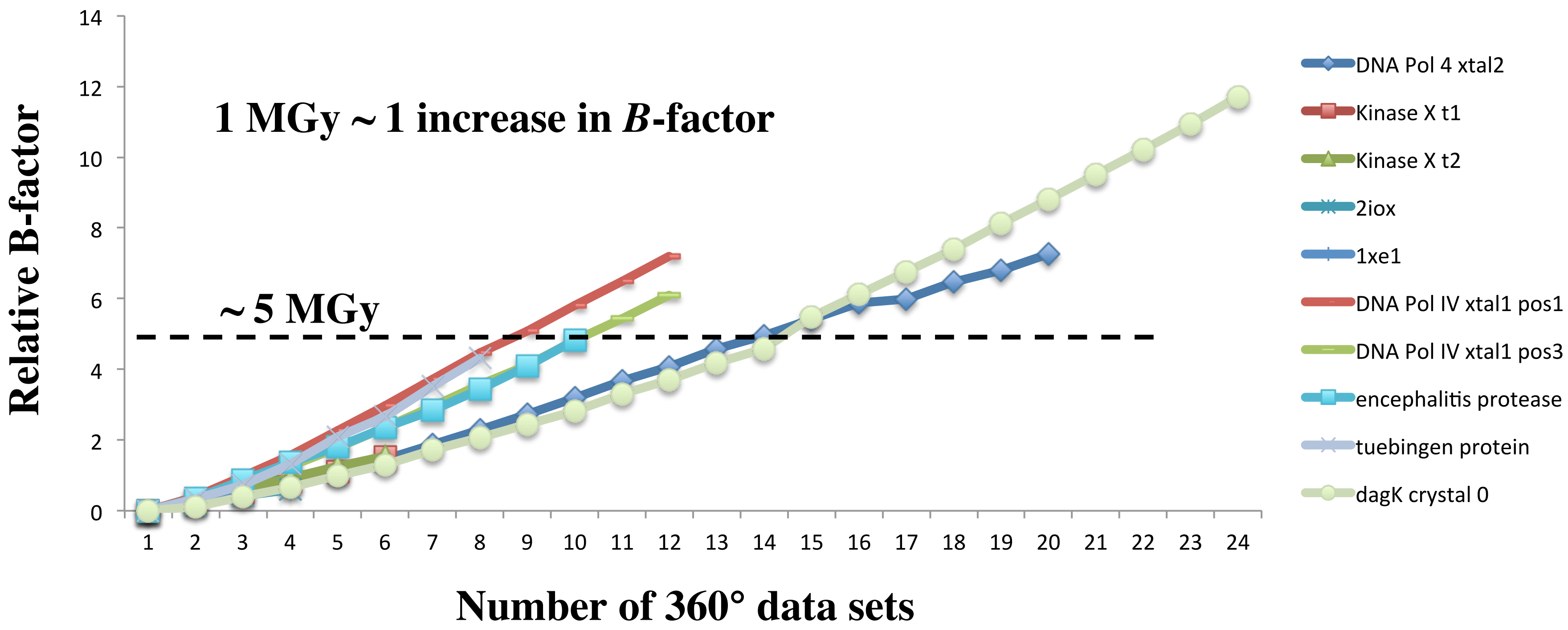
Sample	Dataset	Orientation, *chi*	Time	Native-SAD
Position 1	0.1° / 0.1 s, 720°	0°, 10°, 20°, 30°	48 min	✔
Position 2	0.1° / 0.1 s, 720°	0° (*i.e.* one orientation)	48 min	✖



Low-Dose Strategy Allows Radiation Damage Control

Crystal size 100 × 100 × 100 μm³
X-ray beam size 90 (h) × 45 (v) μm³
6 keV, 1.5 × 10^{10} ph/sec
➔ about 0.5 MGy/360 sec, 360° data

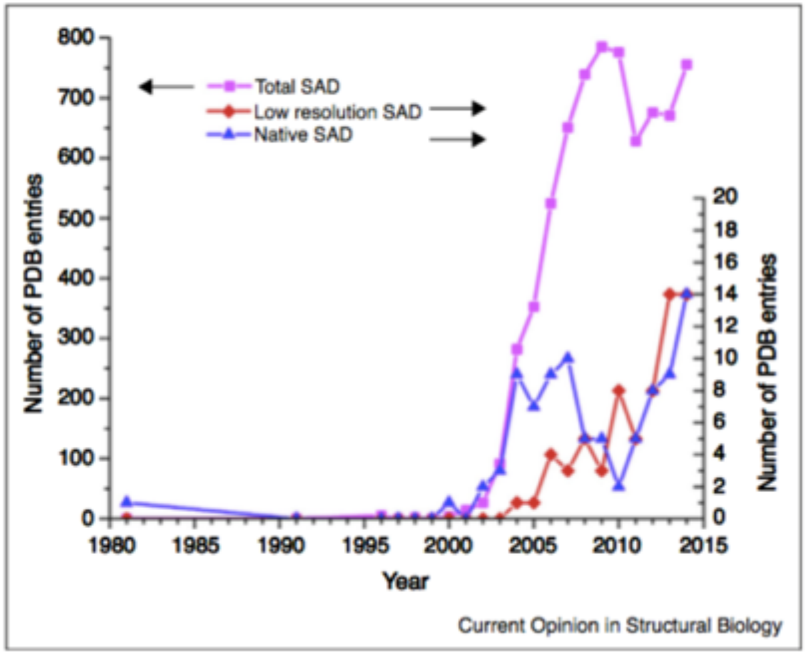




Application in Experimental Phasing

- Phasing by anomalous dispersion has been and still is by far the most popular *de novo* structure determination method

Figure 1



SAD structures in the PDB. PDB entries are as of 31 December 2014. *De novo* low-resolution SAD structures as defined here have $d_{min} \geq 3.5$ Å. *De novo* native-SAD structures as defined here have no preceding PDB deposits and have not contained atoms heavier than atomic number 20.

Liu & Hendrickson Curr. Opin. Struc. Bio. **34,** 99 (2015)

Current challenges

- Low-res SAD with $d > 3.5$ Å
- Native-SAD with $Z < 20$ (Vincent Olieric's talk in Session II)

Conclusion

Distributing tolerable X-ray does into multiple data sets measured at different crystal orientation can improve accuracy of intensity measurement

Low-dose, high-multiplicity, multi-orientation data collection strategy helps in abstracting weak anomalous signals for experimental phasing

Development in synchrotron beamline X-ray optics, multi-axis goniometer and pixel array detector has changed MX data collection strategies

Acknowledgement

X06SA micro-focusing upgrade

Clemens Schulze-Briese, Claude Pradervand, Roman Schneider, James Leuenberger, and MX group

EIGER 16M

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Kommission für Technologie und Innovation KTI

Stefan Brandstetter, Clemens Schulze-Briese, Andreas Förster, Oliver Bunk, Arnau Casanas, and MX group

PRIGo/SmarGon/Native-SAD

Tobias Weinert, Vincent Olieric, Sandro Waltersperger, Claude Pradervand,, Ezequiel Panepucci

Data processing

Kay Diederichs

Global Phasing Ltd.

Many many collaborators providing test crystals and ideas!

Thank you for your attention!