In Meso In Situ Serial Crystallography Workshop
Nov. 17-19, 2015

Wir schaffen Wissen – heute für morgen

Paul Scherrer Institut, Swiss Light Source
Meitian Wang

Data collection strategy: from single-crystal to multi-crystal and serial crystallography

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,)	2×10^{11} <–> 2×10^{12}	2×10^{11} <–> 2×10^{12}	5×10^{11}
Beamsize, μm^2 (with focusing, slits)	**2 × 1 <–> 100 × 100 (fast beam size change)**	**10 × 10 <–> 100 × 100**	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

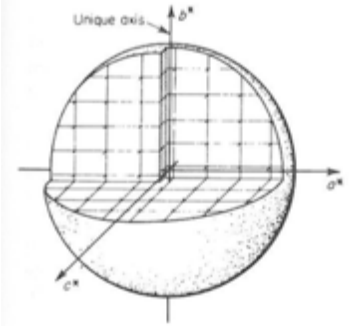


Figure 6.6. Unique volume in reciprocal space for a monoclinic crystal.

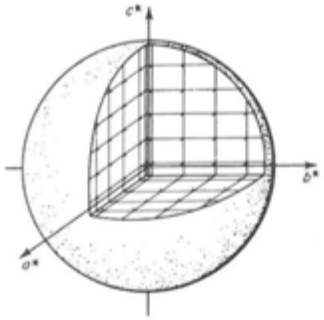
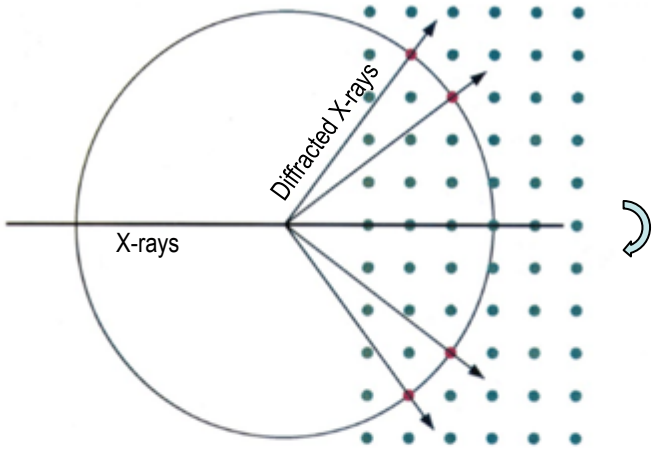


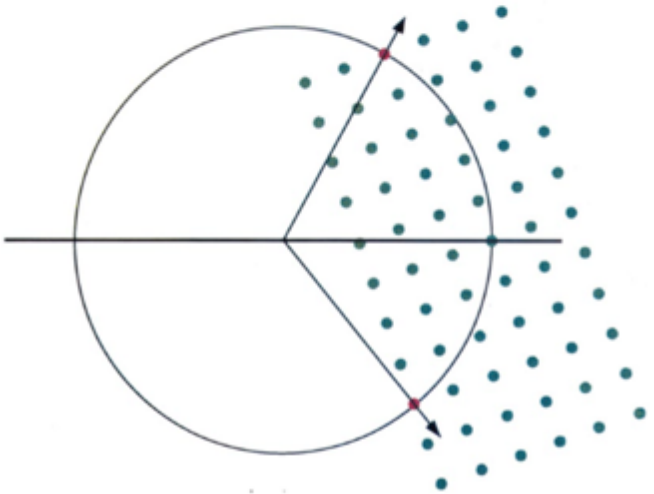
Figure 6.7. Unique volume in reciprocal space for an orthorhombic crystal.

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*); $90 + \theta_{max}$ (*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

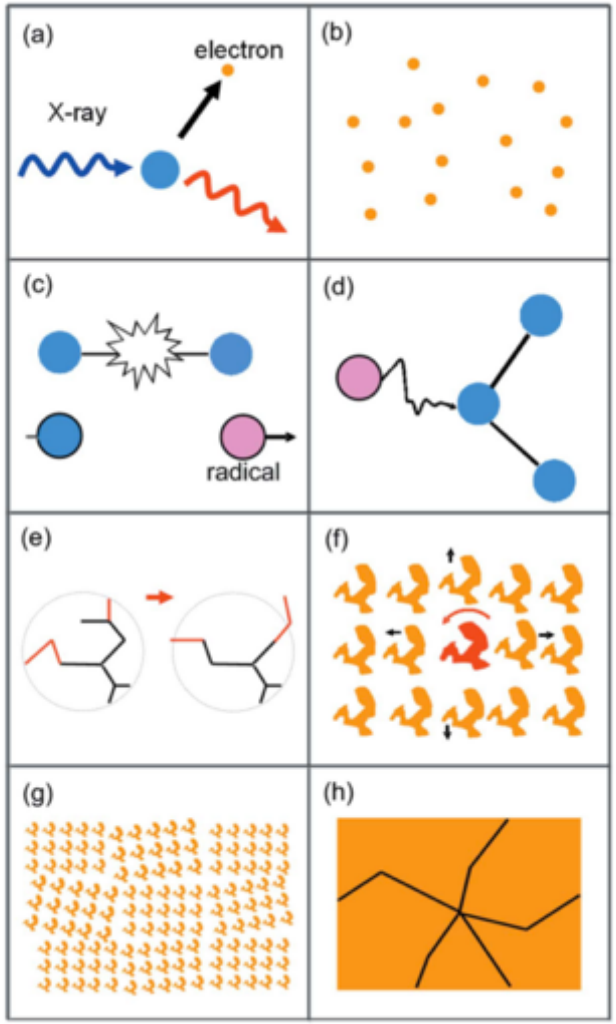




Rotation method and rotation range

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

Radiation Damage



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

Room temperature 298 K

(0.1 – 0.5 MGy)

- Owen, *et al. Acta Cryst.* **D68**, 810 (2012)
- Warkentin, *et al. J. Synchrotron Rad*, **20**, 7 (2013)

Cryo-temperature 100 K

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)
- Does estimation, Holton, *J. Synchrotron Rad.* **16**, 133 (2009)

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

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

Radiation Damage and Single-, Micro-, Serial- Crystallography

Radiation damage estimation at 100 K

$Dose = (t_{expo} \times flux) / (k_{dose} \times I_{H\text{-}beam} \times I_{B\text{-}beam})$; $k_{dose} = 2000\lambda^{-2}$ (Gy, sec, photon, μm)

e.g. 12.4 keV (1.0Å), 4×10^{11} photon/sec,
dose-rate =

4×10^{11} p/s / (2000 × **100 μm × 100μm**) = 0.02 MGy/s	→ one xtal	Single crystal crystallography
4×10^{11}p/s / (2000 × **10 μm × 10 μm**) = 2 MGy/s	→ a few xtals	Micro-crystallography
4×10^{11}p/s / (2000 × **5 μm × 5 μm**) = 8 MGy/s	→ tens xtals	Serial crystallography
4×10^{11}p/s / (2000 × **1 μm × 5 μm**) = 40 MGy/s	→ hundreds xtals	Serial crystallography
4×10^{11}p/s / (2000 × **1 μm × 1 μm**) = 200 MGy/s	→ thousands xtals	Serial femtosecond crystallography (xFEL)

Does estimation, Holton, *J. Synchrotron Rad.* **16**, 133 (2009)

Intensity Data Collection: Counting Statistics

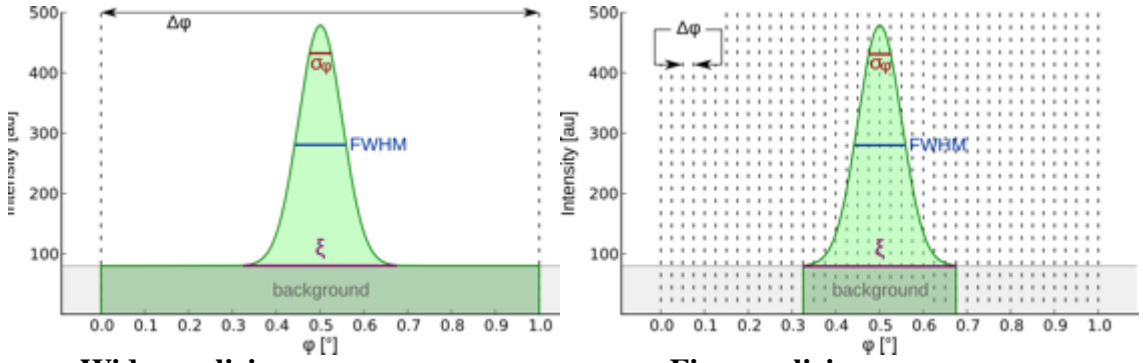
Random errors, counting statistics

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

$$I = N_p - N_b \qquad \textbf{Signal is the difference}$$

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

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

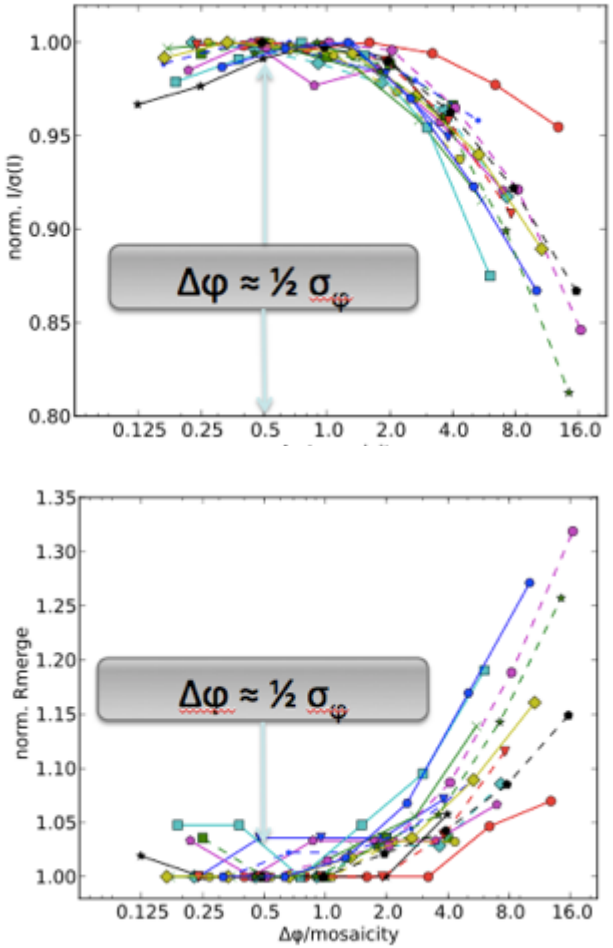


Wide φ-slicing

- 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

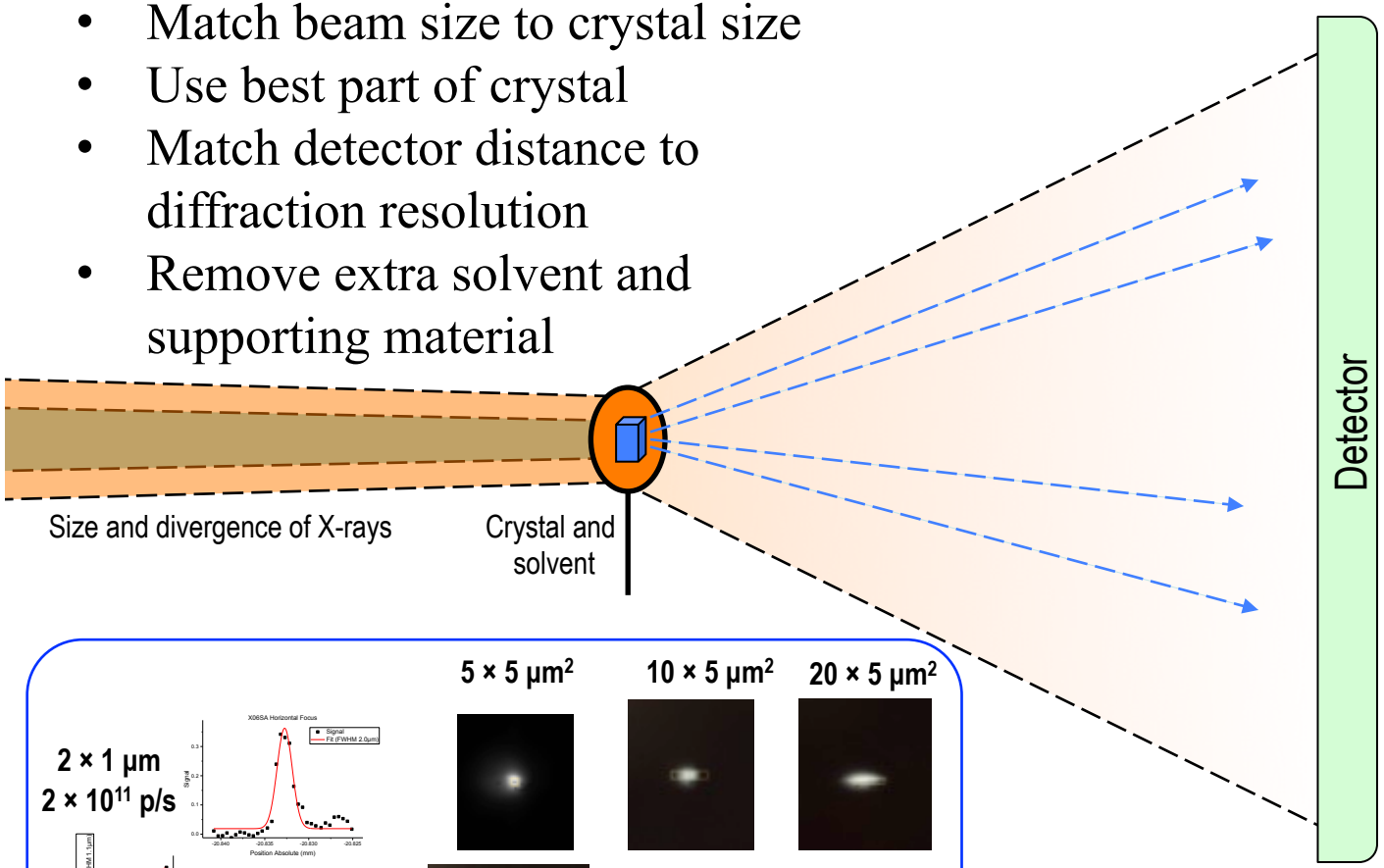


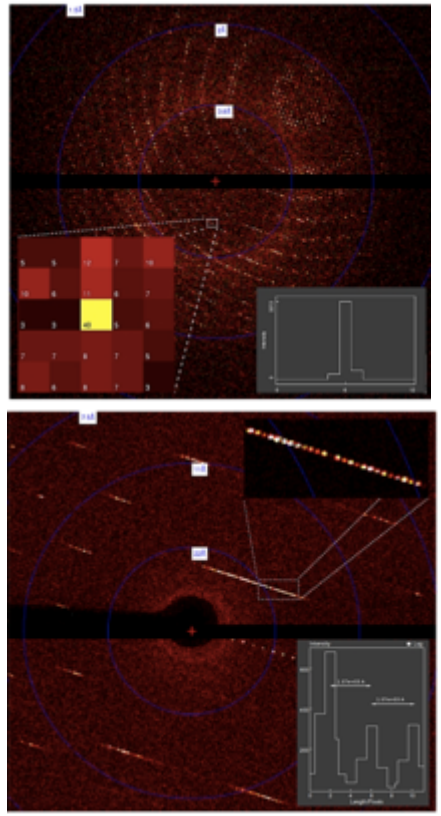
Fine-phi slicing data collection is enabled by the pixel array detector (PILATUS, EIGER), which has single-photon sensitivity and no readlout noise

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

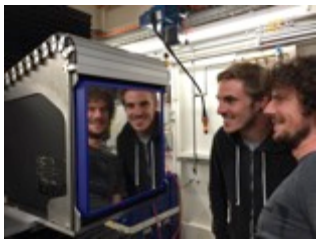
Intensity Data Collection: Reduce Background

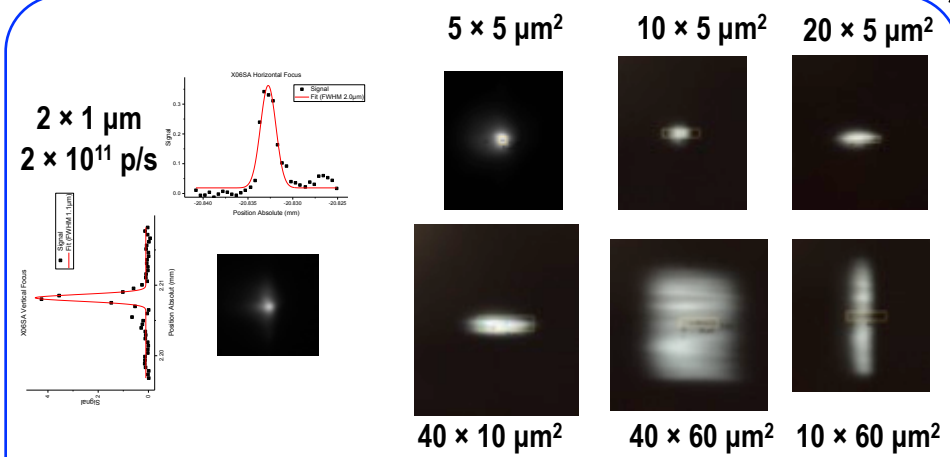
- Match beam size to crystal size
- Use best part of crystal
- Match detector distance to diffraction resolution
- Remove extra solvent and supporting material





Spot resolution with EIGER 16M detector (75 µm pixel, zero point-spread function)) at beamline PXI, SLS





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

In Meso Methods in Membrane Protein Crystallography

- The lipid cubic phase (LCP) or in meso method for crystallizing membrane proteins has delivered over 260 structures of integral membrane proteins (12% of the membrane protein structures in the PDB).
- The method is experiencing explosive growth as half of the 260 structures was deposited in the last two year.
- Because of the sticky and viscous nature of the mesophase, the *harvesting process* is slow and inefficient with the loss of valuable crystals.

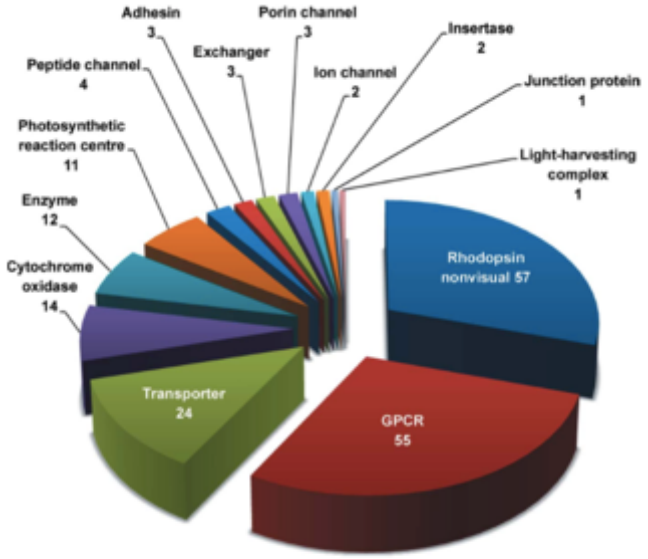


Figure 1
Distribution by biological function or activity of integral membrane proteins and peptides crystallized by the *in meso* method that have yielded crystal structures and records in the Protein Data Bank. The data correspond to the entries in Table 1 and were sourced from the Protein Data Bank in September 2014.

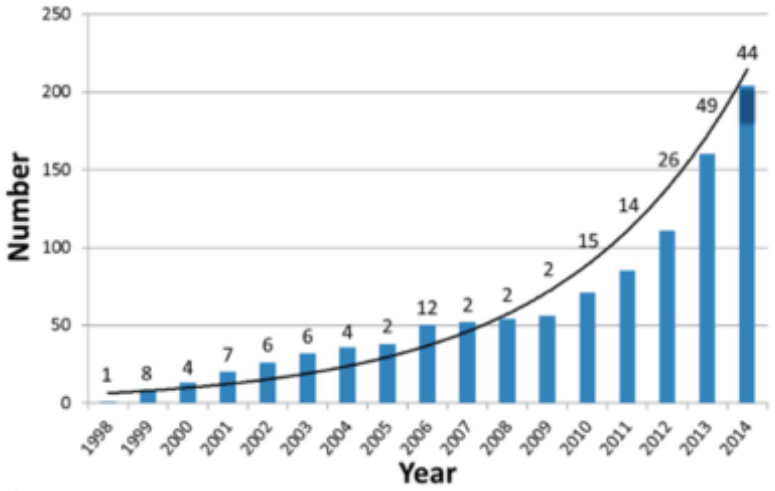
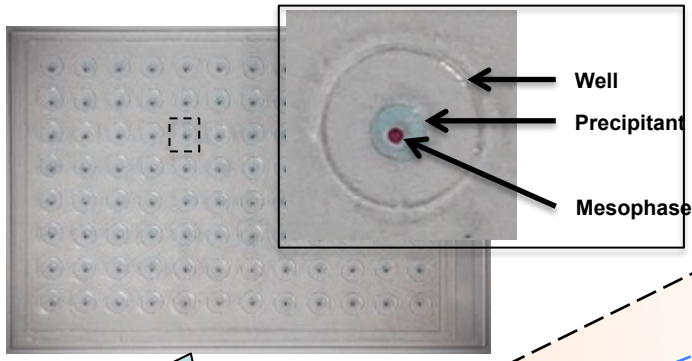


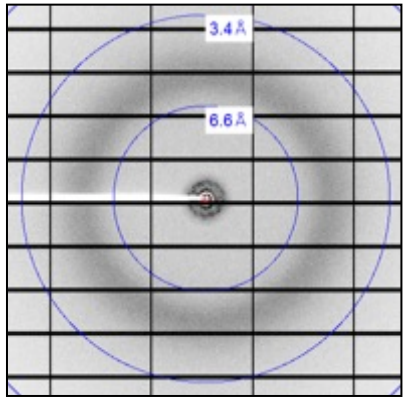
Figure 2
Annual cumulative number of released PDB records for integral membrane-protein and peptide structures solved with crystals grown by the *in meso* method. The number of records released each year is indicated. The figure for 2014 is estimated based on a count of 32 recorded up until September 2014. The line is drawn to guide the eye and takes the form $y = 5.13\exp(0.22x)$.

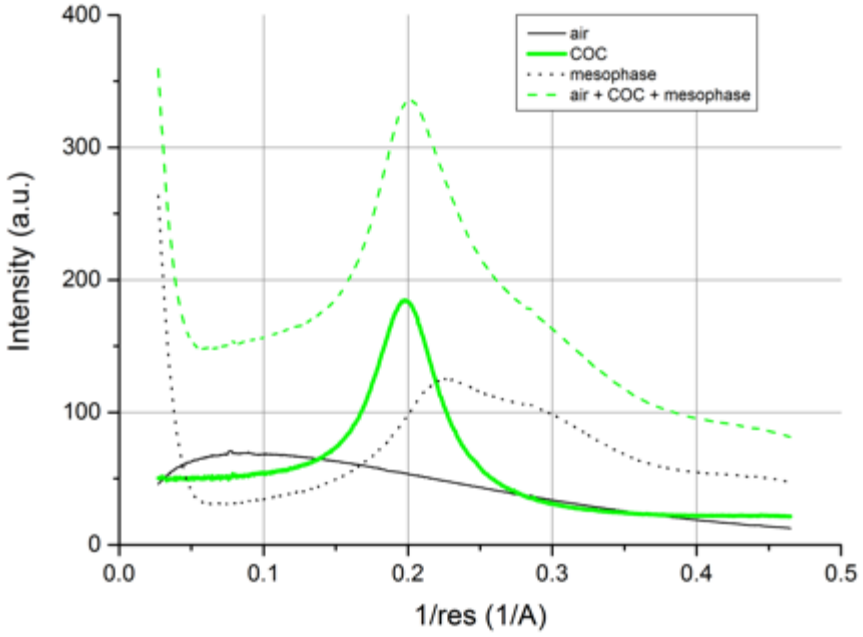
M. Caffrey, *Acta Cryst*. **F71**, 3 (2015)

From Well and Loop to *in situ* (IMISX)



Sandwich plate with Cyclic Olefin Copolymer film (COC)

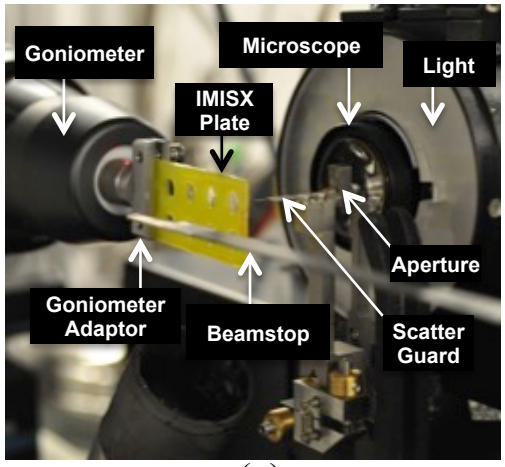




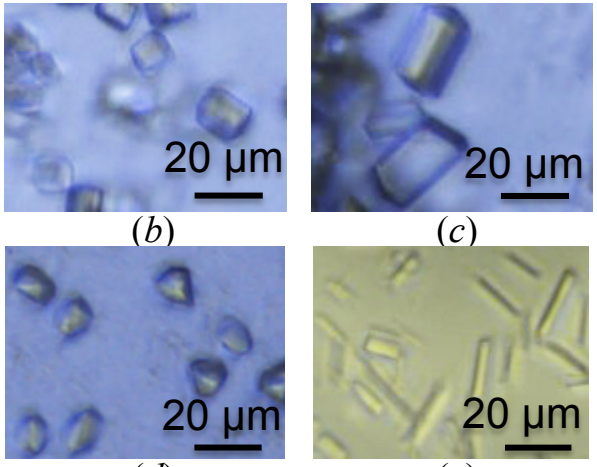
X-ray diffraction background

IMISX Data Collection at Room Temperature

Experimental setup

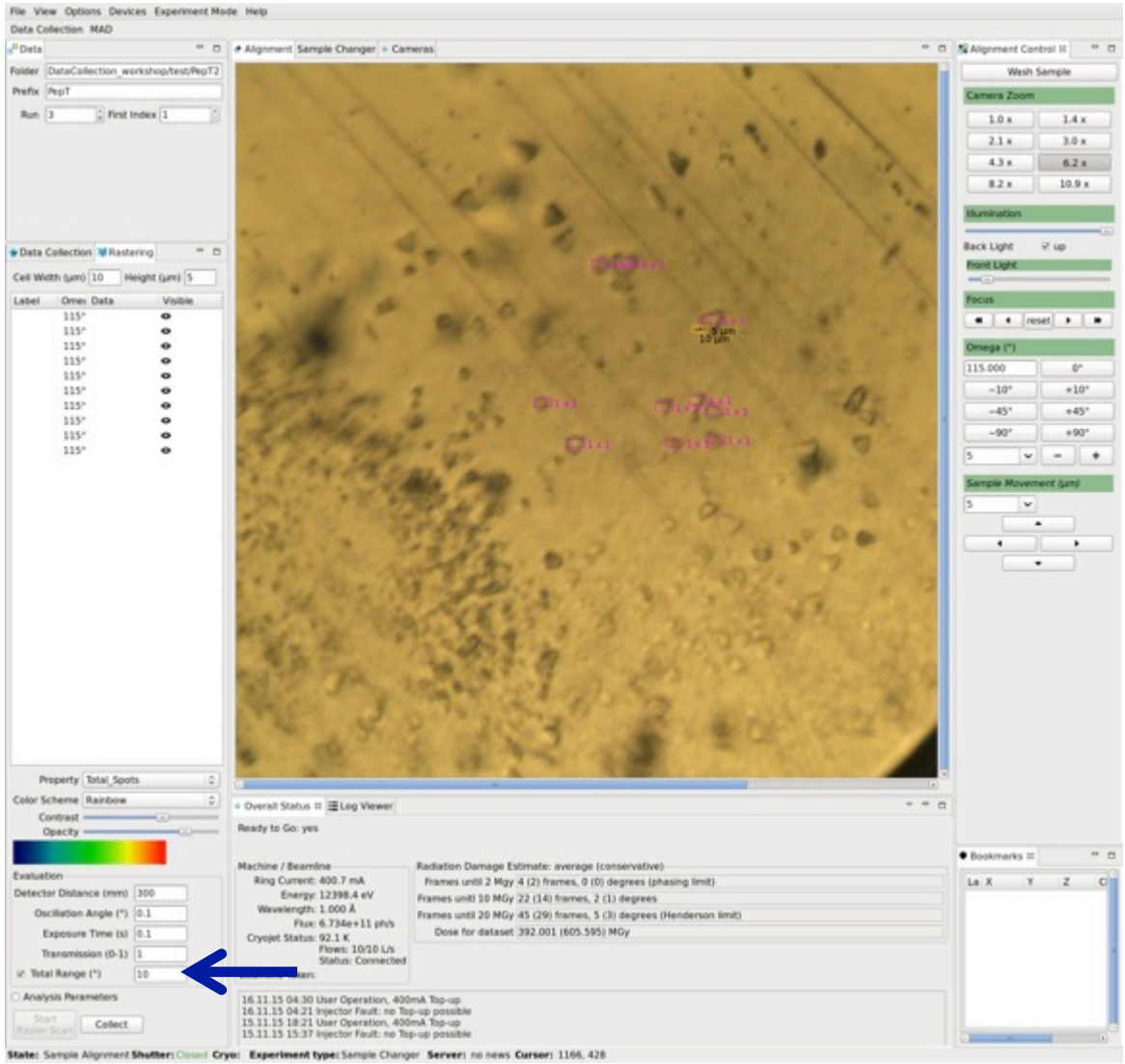


(*a*)



(*b*) (*c*) (*d*) (*e*)

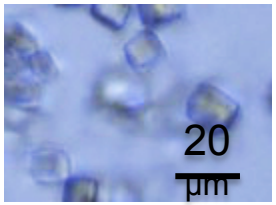
Crystals viewed through an on-axis microscope



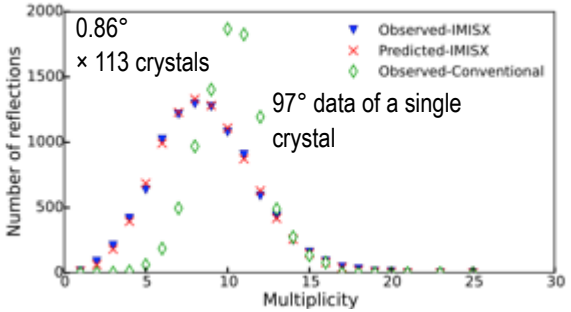
Semi-automated “select and collect” protocol in DA+ GUI

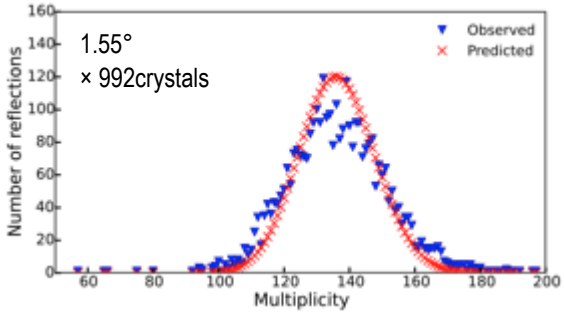
IMISX RT Example: S-SAD Phasing with Lysozyme

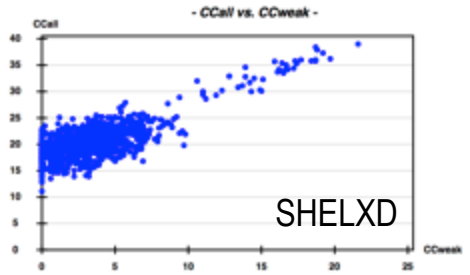
Lysozyme in LCP (Chia-Ying Huang)
Crystal size: ~ **10 × 10 × 20** μm^3
Wavelength 1.7 Å
Beam size 30 × 10 μm^2
In situ room temperature
20 × 0.1° frames per crystal
1290 crystals collected (Vincent Olieric)
1171 datasets processed
992 datasets merged (Kay Diederichs)

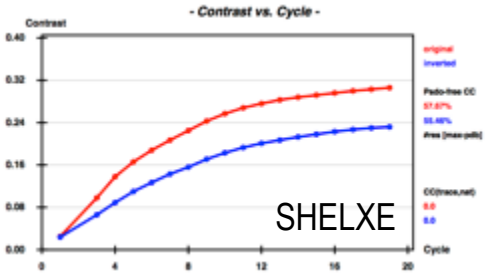


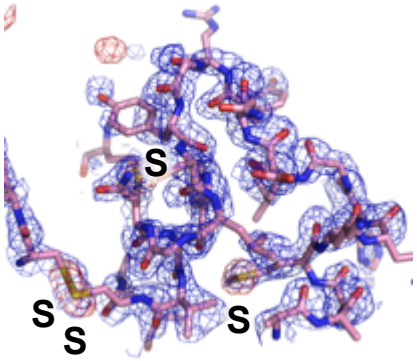
Multiplicity of reflections











Experimental phased map

In merged data sets obtained with ***randomly oriented*** crystals, the multiplicities follow a binomial distribution.

$$B(n,p,k) = \binom{n}{k} p^k (1-p)^{n-k}$$

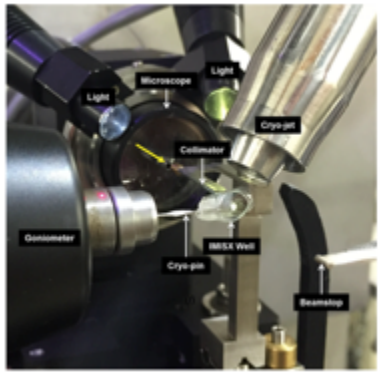
n is the number of asymmetric units in recip. space
p is the effective oscillation range of dataset / 180
k is the multiplicity

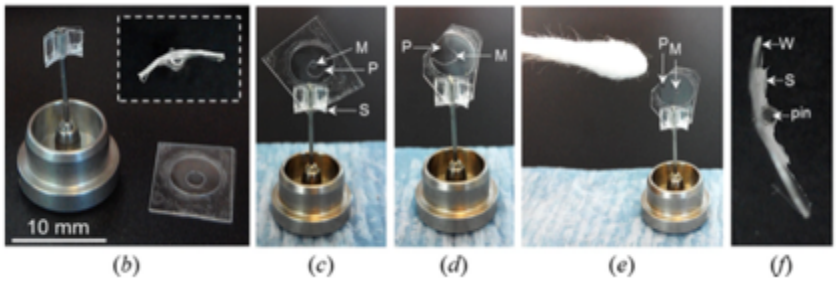
IMISX method, Huang, *et al. Acta. Cryst.* **D71**, 1238 (2015)

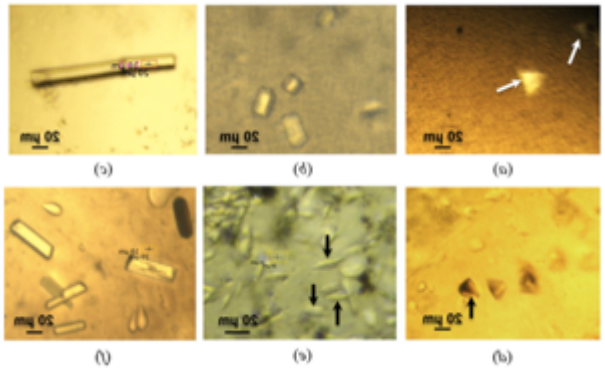
Room Temperature is Cool, But Cryo is Even Cooler

The cryogenic advantages (IMISXcryo)

- 50 times longer crystal lifetimes with X-rays
- Keep crystals at their best state
- Prepare crystals in advance of beamtime
- Simplify crystal storage and transportation
- Compatible with sample changers (vial)
- IMISXcryo specific, crystals are visible under crvo



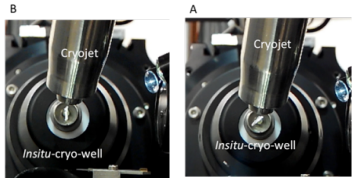




Precipitant removal technique → Visible crystals *in meso* in cryo

IMISXcryo *vs.* loop-harvesting

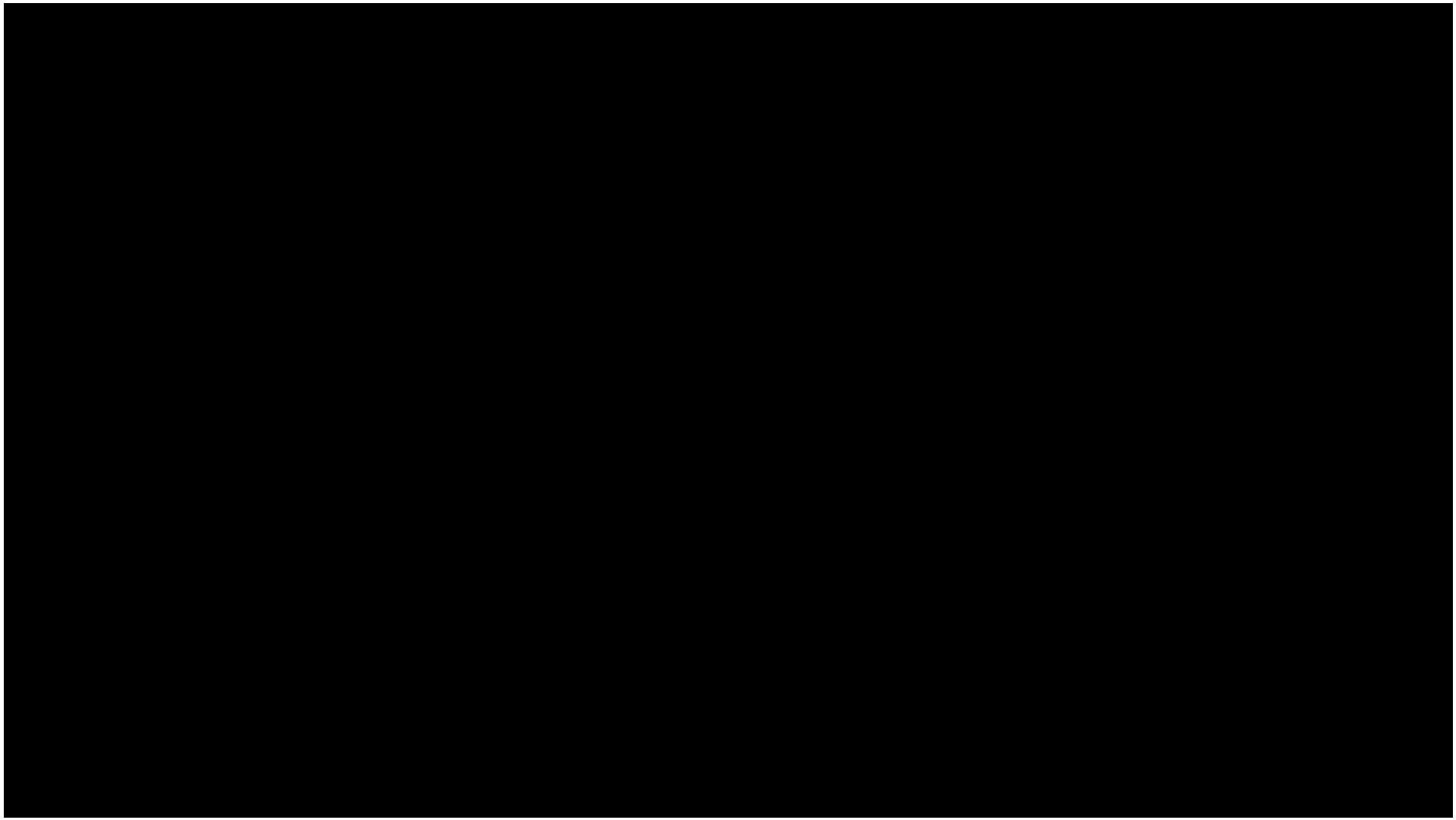
- Limited rotation range ~ +/– 45°
- Preferential crystal orientation



X-ray

Curved-well technique → Extend rotation range

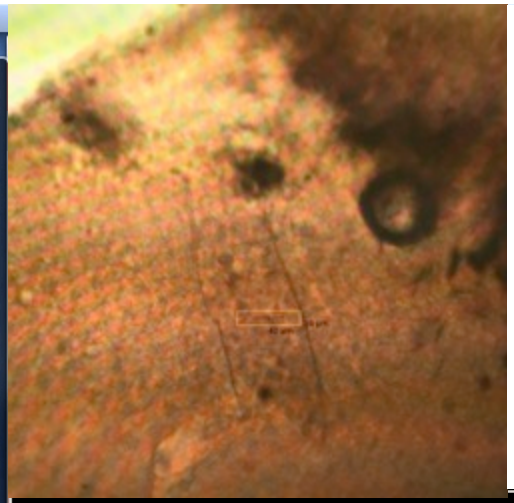
Automated Sample Changing with CATS System



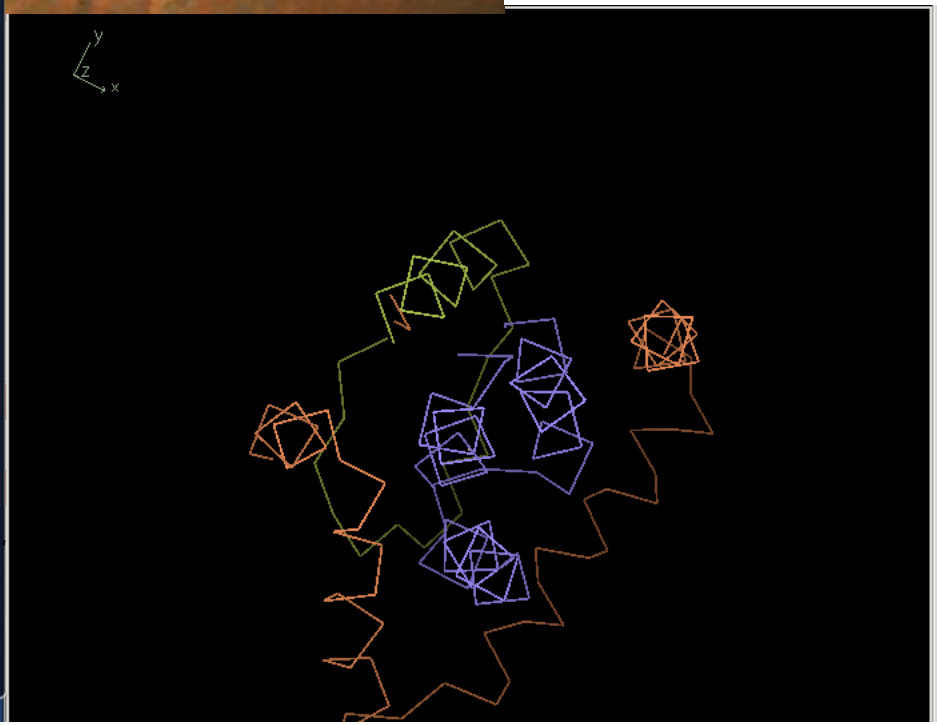
Sample exchange as standard loop-harvested crystals in pin/vial

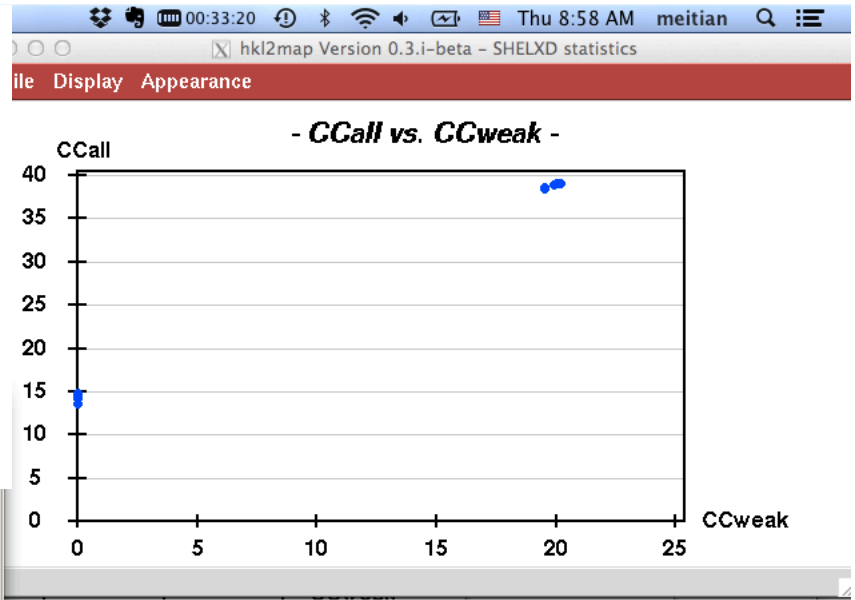
20 – 100 crystals per well, 10 or 16 wells per puck

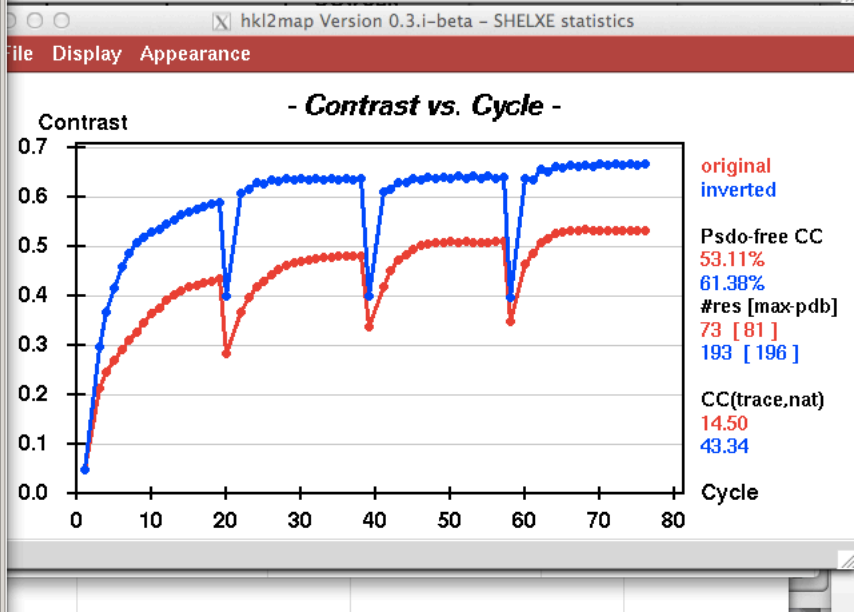
IMISXCryo Example One: One Crystal W-SAD Phasing



- 150 × 50 ×10 μm^3
- +/− 70° data
- SG. *I*222
- 1.9 Å, I/σ = 1.8, CC_{ano}= 0.26
- Completeness 91.7%
- SHELXC/D/E

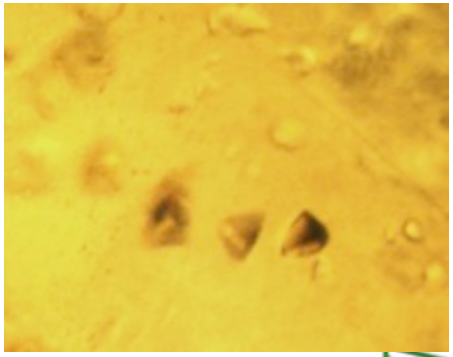


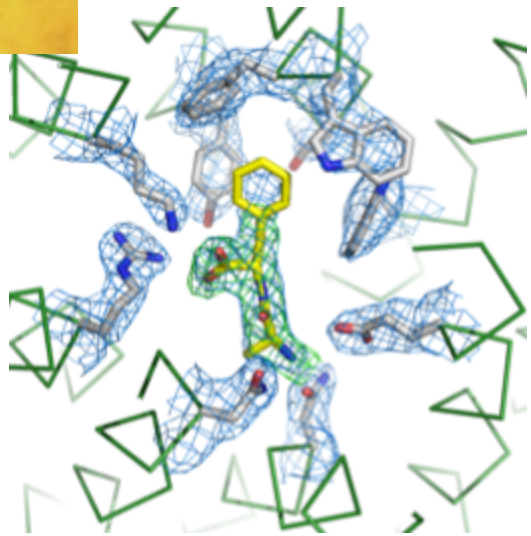




IMISXCryo Example Two: Two Crystals Native

Protein	$PepT_{st}$ (peptide transporter)
Number of crystals	2
Data per crystal (°)	60
Crystal size (μm^3)	20 × 20 × 30
Beam size (μm^2)	10 × 18
Space group	$C222_1$
Unit cell (Å)	100.2, 109.5, 111.5
Resolution (Å)	50 – 2.4 (2.46 – 2.4)
R_{meas}	0.122 (0.999)
$R_{p.i.m.}$	0.059 (0.487)
I/σ	9.3 (2.2)
$CC_{1/2}$	0.99 (0.62)
Completeness (%)	99.4 (98.8)
Multiplicity	4.3 (4.2)

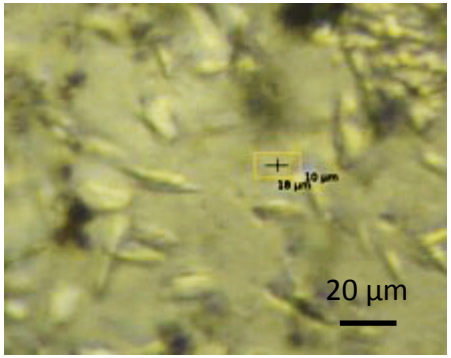


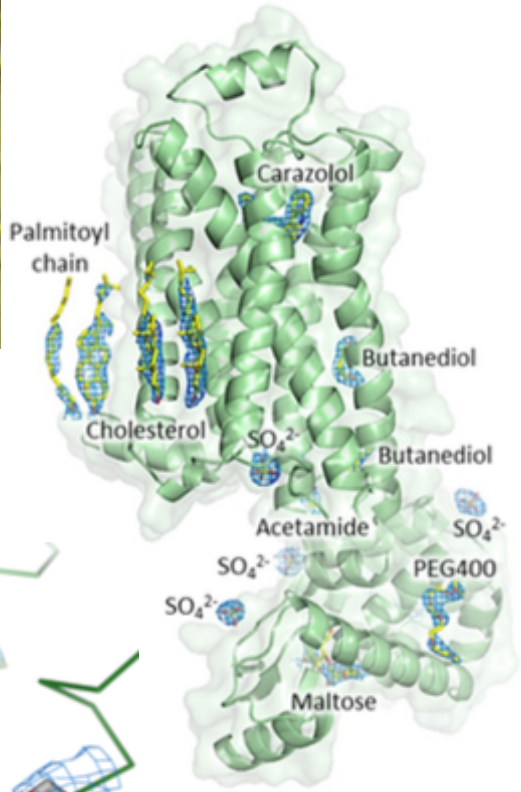


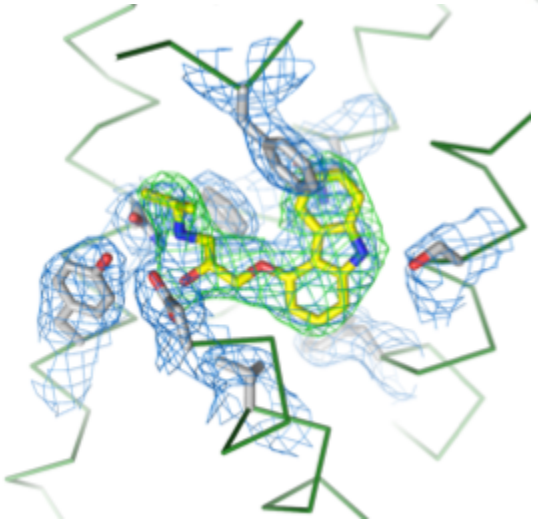
Diffraction resolution is the same as the best resolution obtained in previous published structure with loop-harvested crystals

IMISXCryo Example Three: 104 Crystals GPCR

Protein	Rock2 (Human β2-adrenoreceptor)
Number of crystals	104 of 149 in one well
Data per crystal (°)	3
Crystal size (μm^3)	**5 × 10 × 30**
Beam size (μm^2)	10 × 18
Space group	*C*2
Unit cell (Å)	108.0, 170.6, 40.4
Resolution (Å)	50 – 2.5 (2.57 – 2.5)
R_{meas}	0.203 (2.084)
$R_{p.i.m.}$	0.085 (0.879)
I/σ	7.3 (1.1)
$CC_{1/2}$	0.99 (0.21)
Completeness (%)	95.1 (91.0)
Multiplicity	5.7 (5.6)

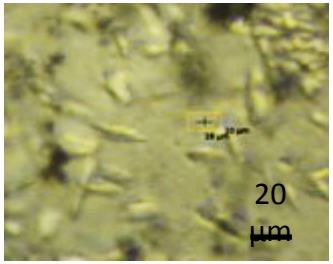


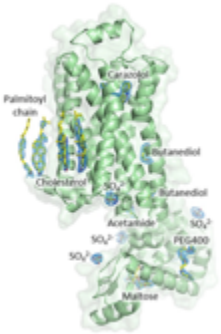


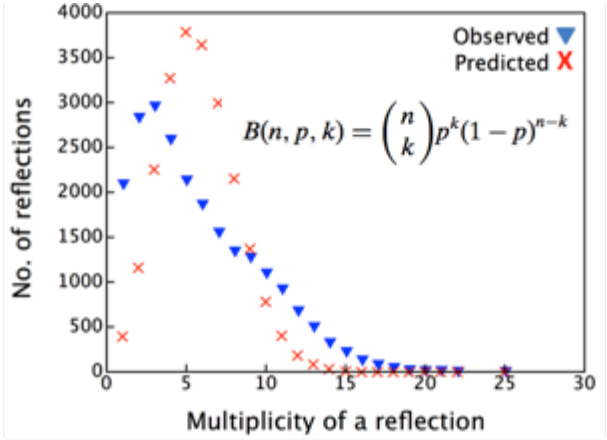


IMISXCryo Example Three: 104 Crystals GPCR

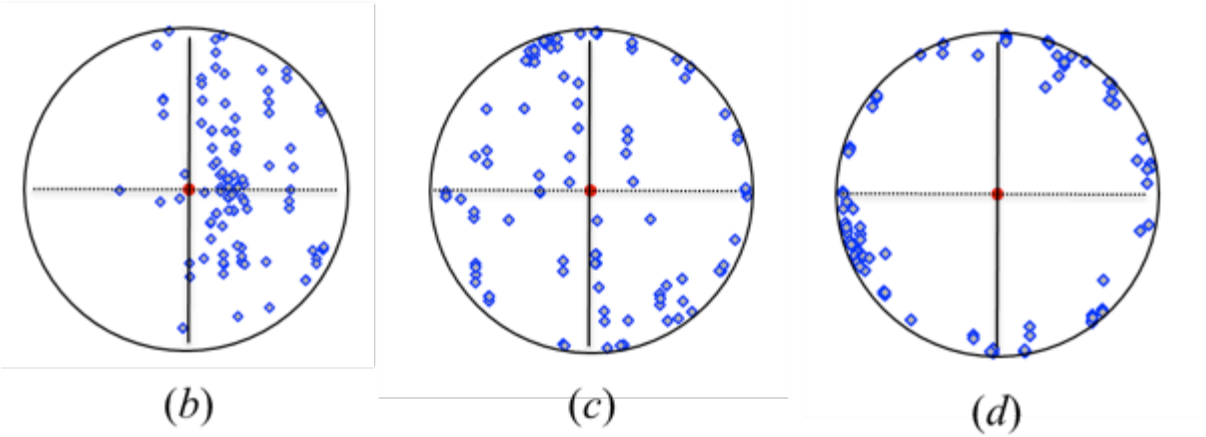
Preferential crystal orientations as shown in the screwed multiplicity distribution from the Binomial





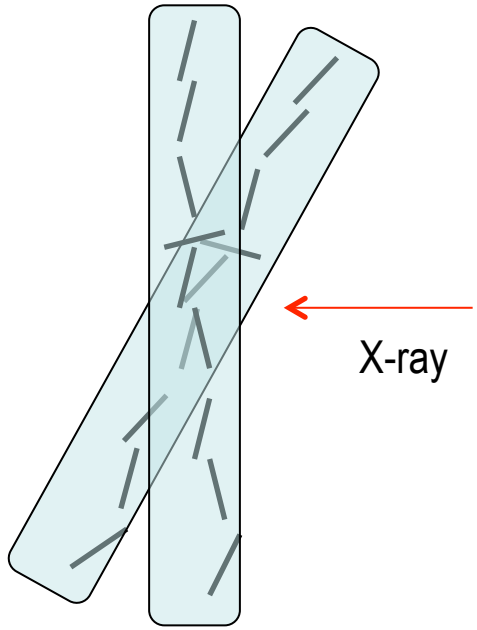


(a)



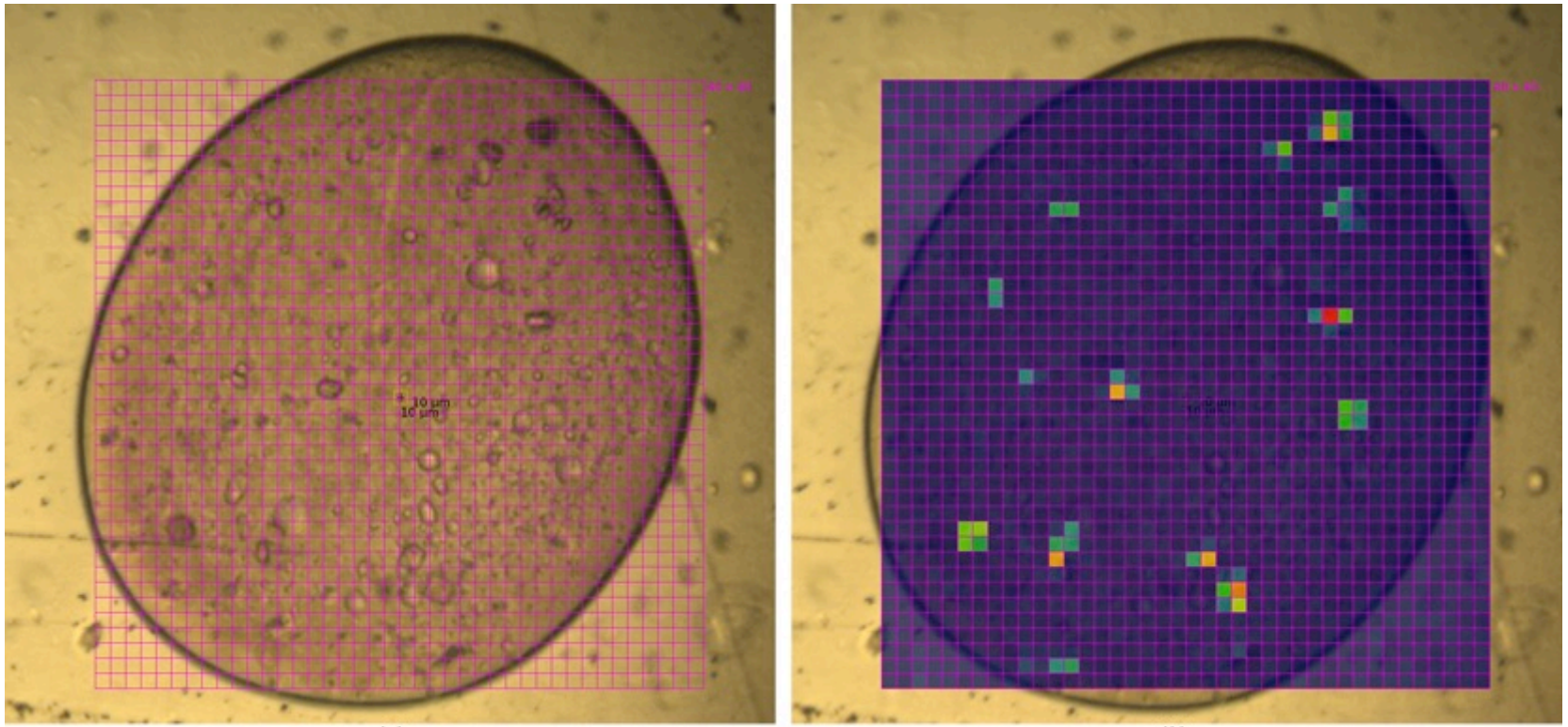
(b) (c) (d)

Projection of the unit-cell a, b, and c axis along the X-ray beam direction (red dot)



IMISX plate

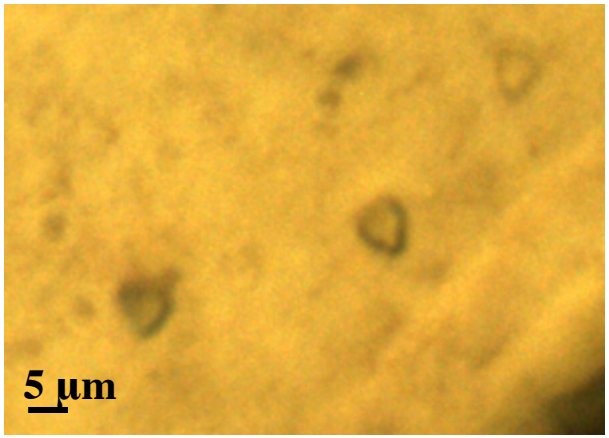
Fast Rastering of LCP Bolus at 100K

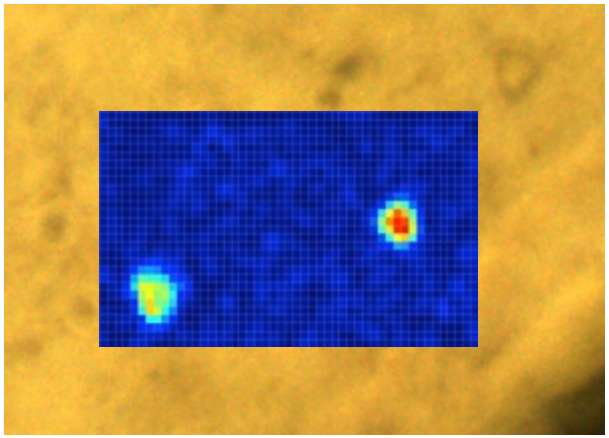


(a) (b)

Cryogenic data collection allows diffraction based crystal searching and centering
Rastering of LCP bolus in 64 second
40 × 40 grid-scan with beam size of 10 × 10 μm^2 and a PILATUS 6M operated at 25 Hz

Serial Micron-Crystallography





50 × 30 grid-scan in 1 µm steps (beam size of 3 × 2 µm^2) on membrane protein microcrystals grown in LCP (C.-Y. Huang from M. Caffrey lab)

Number of crystals	25
Data per crystal (°)	10
Crystal size (µm^3)	**5 × 5 × 5**
Beam size (µm^2)	8 × 8
Space group	$C222_1$
Unit cell (Å)	102.4, 109.8, 111.4
Resolution (Å)	50 – 2.8 (2.87 – 2.8)
R_{meas}	0.440 (1.662)
$R_{p.i.m.}$	0.145 (0.549)
I/σ	6.2 (1.6)
$CC_{1/2}$	0.983 (0.447)
Completeness (%)	99.8 (100.0)
Multiplicity	9.2

Complete 2.8 Å data set merged from 25 crystals measured with the EIGER 16M at beamline X06SA, SLS

New Challenges in Data Processing and Merging

Process each data set with XDS

Merge data sets together with XSCALE

How to select data sets?

How to evaluate quality of the merged data?

New Challenges in Data Processing and Merging

Se-SAD phasing

210 crystals collected with 15° data each

86 crystals

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE >= -3.0 AS FUNCTION OF RESOLUTION

RESOLUTION LIMIT	NUMBER OF REFLECTIONS OBSERVED	UNIQUE	POSSIBLE	COMPLETENESS OF DATA	R-FACTOR observed	R-FACTOR expected	COMPARED	I/SIGMA	R-meas	CC(1/2)	Anomal Corr	SigAno	Nano
12.08	8708	375	381	98.4%	16.8%	17.0%	8705	27.26	17.2%	99.5*	93*	3.406	144
8.54	17049	690	690	100.0%	14.3%	16.7%	17049	23.57	14.7%	98.9*	88*	2.884	300
6.97	23497	897	897	100.0%	18.3%	19.5%	23495	20.06	18.7%	99.3*	80*	2.506	402
6.04	23909	1060	1060	100.0%	27.7%	27.5%	23905	14.03	28.3%	98.3*	75*	2.068	477
5.40	30161	1193	1193	100.0%	33.0%	32.6%	30158	13.34	33.6%	98.1*	64*	1.686	547
4.93	34059	1336	1336	100.0%	28.3%	28.7%	34059	14.87	28.9%	98.9*	60*	1.570	620
4.56	36934	1456	1456	100.0%	27.1%	27.0%	36934	15.98	27.6%	99.0*	50*	1.455	680
4.27	34601	1536	1536	100.0%	27.8%	27.4%	34594	15.44	28.4%	98.9*	45*	1.379	719
4.03	40747	1646	1647	99.9%	36.1%	35.7%	40744	13.58	36.8%	97.7*	36*	1.186	771
3.82	43056	1725	1725	100.0%	48.8%	45.7%	43055	11.00	49.8%	97.1*	35*	1.175	814
3.64	47505	1873	1873	100.0%	59.7%	55.3%	47505	9.70	60.8%	95.1*	30*	1.131	888
3.49	48589	1900	1901	99.9%	79.8%	76.9%	48589	7.59	81.3%	92.9*	19*	1.006	905
3.35	43934	1993	1993	100.0%	93.2%	88.7%	43926	6.06	95.3%	87.3*	18*	0.955	942
3.23	50511	2090	2091	100.0%	104.8%	103.5%	50508	5.40	107.0%	88.0*	11	0.869	993
3.12	53528	2158	2158	100.0%	130.3%	129.7%	53525	4.46	132.9%	78.9*	12*	0.885	1028
3.02	56010	2217	2217	100.0%	158.3%	160.1%	56007	3.69	161.4%	76.7*	10	0.872	1059
2.93	58278	2311	2311	100.0%	185.2%	191.0%	58278	3.10	188.8%	67.9*	3	0.779	1110
2.85	59034	2339	2339	100.0%	225.2%	231.4%	59034	2.70	229.7%	61.9*	7	0.799	1124
2.77	56347	2500	2500	100.0%	269.5%	278.7%	56344	2.18	275.5%	49.0*	5	0.771	1201
2.70	56162	2458	2458	100.0%	372.6%	387.3%	56157	1.66	380.7%	36.3*	1	0.725	1176
total	822619	33753	33762	100.0%	47.9%	47.8%	822571	8.30	48.8%	99.1*	39*	1.141	15900

183 crystals

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE >= -3.0 AS FUNCTION OF RESOLUTION

RESOLUTION LIMIT	NUMBER OF REFLECTIONS OBSERVED	UNIQUE	POSSIBLE	COMPLETENESS OF DATA	R-FACTOR observed	R-FACTOR expected	COMPARED	I/SIGMA	R-meas	CC(1/2)	Anomal Corr	SigAno	Nano
12.08	17470	375	381	98.4%	40.4%	33.7%	17470	30.41	40.8%	99.8*	93*	4.126	144
8.54	31967	690	690	100.0%	35.4%	31.9%	31967	26.44	35.9%	99.2*	87*	3.162	300
6.97	43436	897	897	100.0%	41.5%	38.8%	43436	22.55	42.0%	99.2*	84*	2.778	403
6.04	45097	1060	1060	100.0%	63.3%	63.0%	45097	15.77	64.1%	98.5*	77*	2.298	481
5.40	56459	1193	1193	100.0%	71.9%	72.8%	56459	14.87	72.7%	98.6*	69*	1.904	550
4.93	65125	1336	1336	100.0%	62.8%	62.3%	65125	16.47	63.4%	99.1*	61*	1.701	620
4.56	70460	1456	1456	100.0%	59.4%	57.0%	70460	17.58	60.1%	99.2*	53*	1.556	680
4.27	66429	1536	1536	100.0%	61.9%	57.3%	66429	16.86	62.7%	98.9*	48*	1.410	726
4.03	77570	1647	1647	100.0%	77.4%	73.0%	77570	14.84	78.3%	98.1*	38*	1.263	775
3.82	82353	1725	1725	100.0%	100.7%	92.9%	82353	11.96	101.8%	97.5*	35*	1.200	815
3.64	90566	1873	1873	100.0%	114.0%	105.2%	90566	10.55	115.2%	97.3*	31*	1.169	888
3.49	91784	1901	1901	100.0%	153.1%	146.9%	91784	8.19	154.7%	96.0*	23*	1.022	906
3.35	83511	1993	1993	100.0%	170.9%	165.8%	83511	6.59	173.0%	93.6*	22*	0.972	949
3.23	94599	2091	2091	100.0%	187.0%	186.9%	94599	5.87	189.1%	92.6*	9	0.897	997
3.12	99286	2158	2158	100.0%	217.8%	221.9%	99286	4.81	220.2%	88.6*	14*	0.931	1031
3.02	102888	2217	2217	100.0%	260.0%	268.8%	102888	3.99	262.9%	85.5*	10	0.888	1061
2.93	107529	2311	2311	100.0%	293.6%	307.4%	107529	3.33	296.8%	81.1*	7	0.815	1110
2.85	107838	2339	2339	100.0%	336.7%	354.5%	107838	2.87	340.4%	77.4*	5	0.800	1124
2.77	103440	2500	2500	100.0%	451.9%	478.5%	103440	2.28	457.4%	69.5*	5	0.759	1203
2.70	100279	2458	2458	100.0%	608.0%	650.5%	100279	1.72	615.6%	56.2*	5	0.752	1181
total	1538086	33756	33762	100.0%	87.9%	85.3%	1538086	9.10	88.9%	99.5*	43*	1.204	15944

Conclusion

IMISX method allows high-throughput data collection at both room and cryogenic temperatures

IMISXcryo method delivers high quality diffraction data without slow and inefficient crystal harvesting process

IMISXcryo method enables serial crystallography with micron-sized crystals at synchrotron sources

Acknowledgement

IMISX method development
Chia-Ying Huang, Pikyee Ma, Nicole Howe, Lutz Vogeley, Martin Caffrey (Trinity College Dublin)
Vincent Olieric, Ezequiel Panepucci (SLS)

Data processing and merging
Kay Diederichs, Greta Assmann, Wolfgang Brehm (University of Konstanz)
Rangana Warshamanage (SLS)

PX beamline development at SLS
The MX group

PAUL SCHERRER INSTITUT

PSI

Thank you for your attention!