Cheiron School 2010

Protein Crystallography

Macromolecular crystallography (MX)

SPring-8 / JASRI

Structural Biology Group

Takashi Kumasaka

Contents

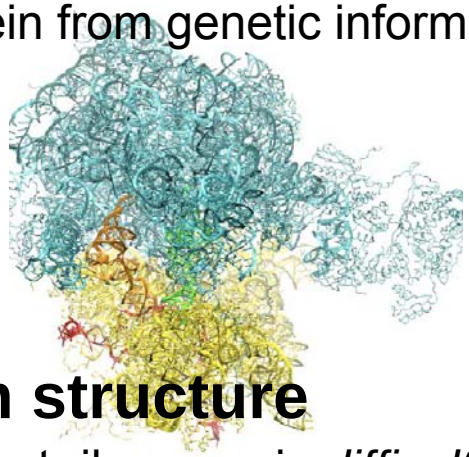
1: Introduction to MX

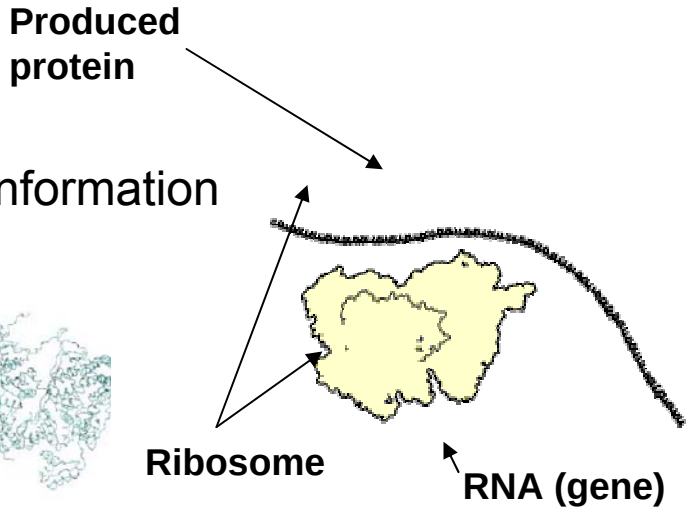
Structural study of Ribosome

Ribosome …
plays a factory to produce protein from genetic information

Central Dogma
DNA -> RNA -> Protein

A major target of antibiotic drug





To reveal function from structure

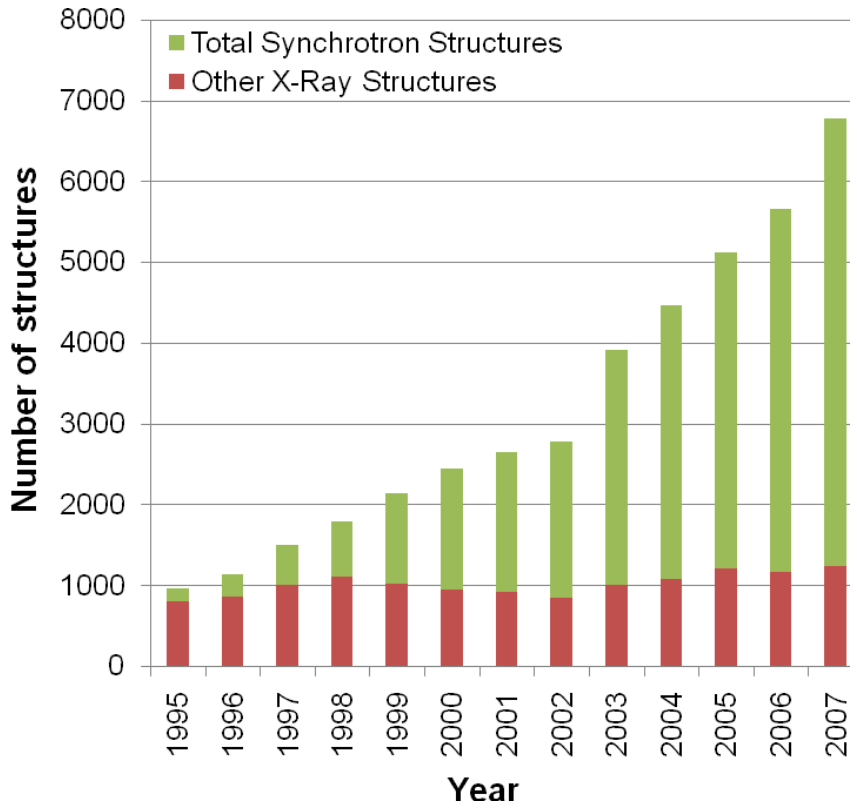
Prof. Ada Yonath
2009 Nobel Prize

It is considered that ribosome is *difficult to be crystallized* because of its huge size. Prof. Yonath started its trial from 1980's when the other prize winners (Profs Steitz and Ramakrishnan) did not undertake it.
> Structures determined at 2000.

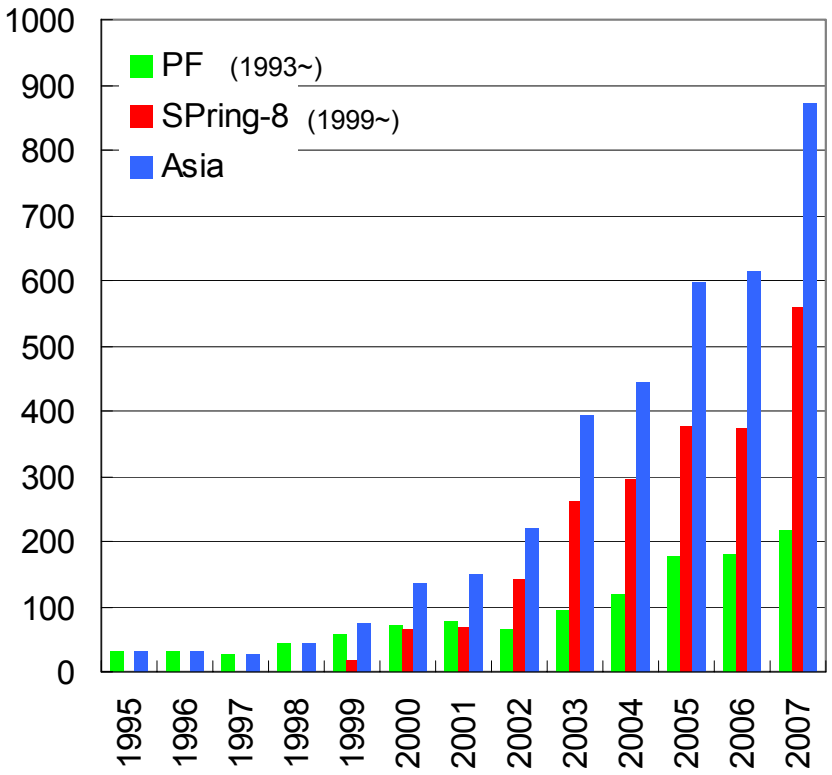
Used synchrotron
ESRF ID2, ID14-2, -4, ID29
APS 19-ID
CHESS F1
DESY BW6 & BW7B
PF BL6A… and many others

Crystal structure and synchrotron

Number of determined structures

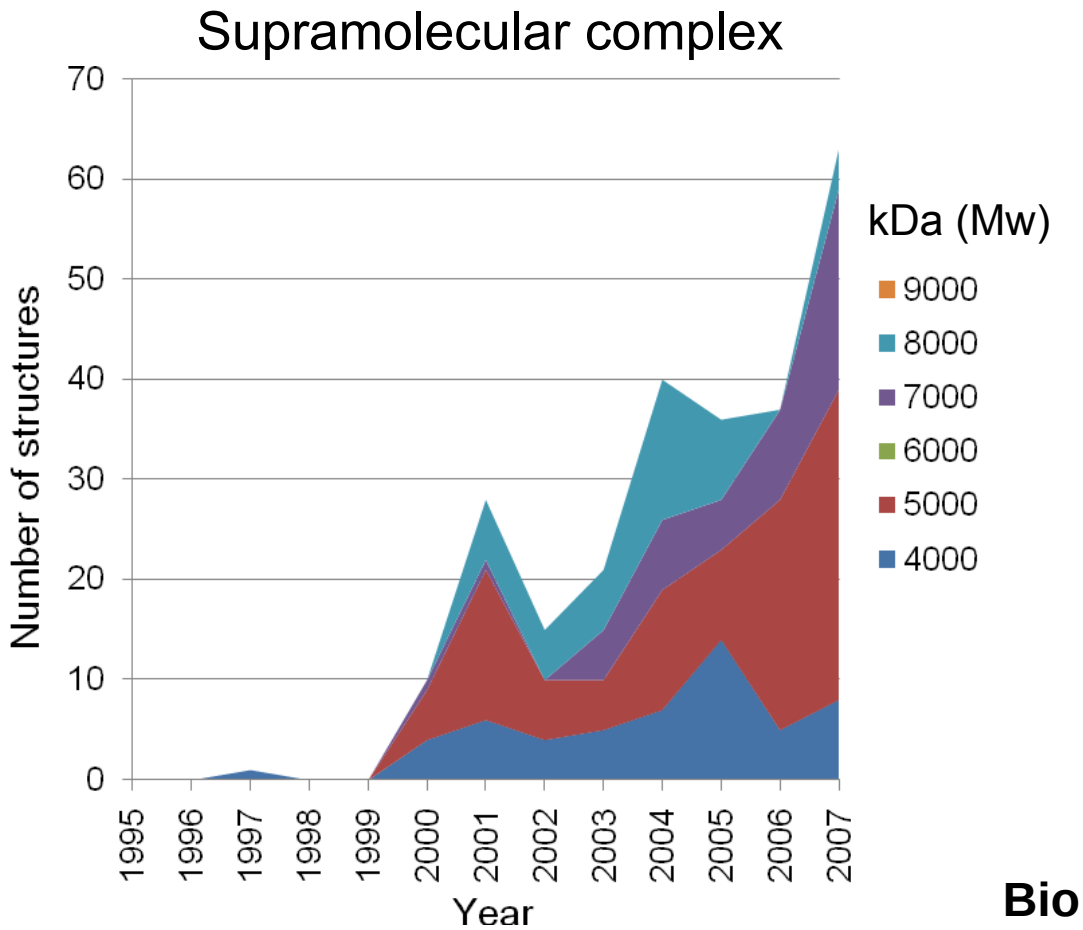


Asian contribution

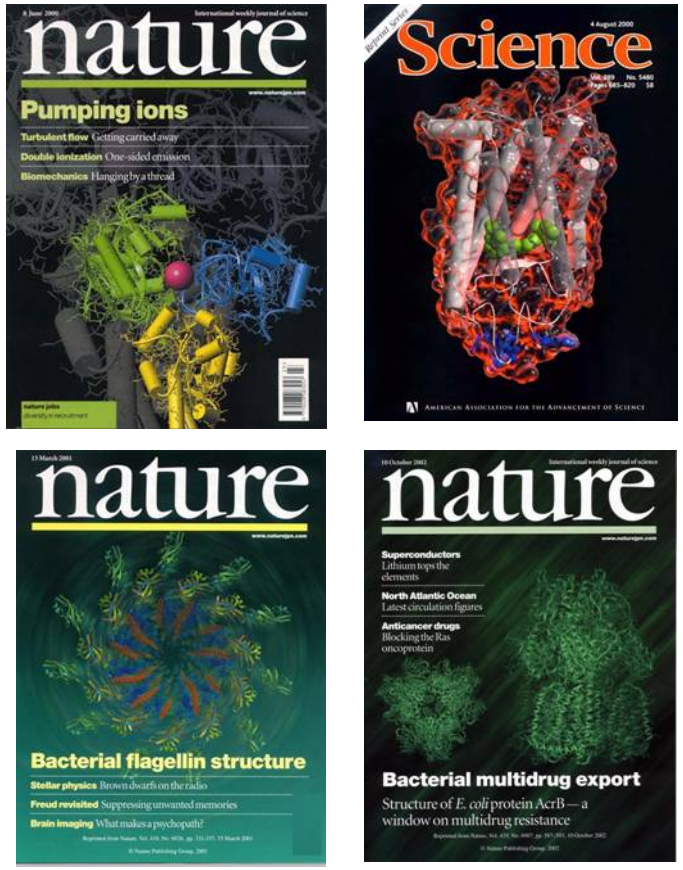


Nowadays, most structures are determined using SOR.

Determination of important and complex structures



3G SOR went into this field from 2000, and accelerates large molecule analysis.

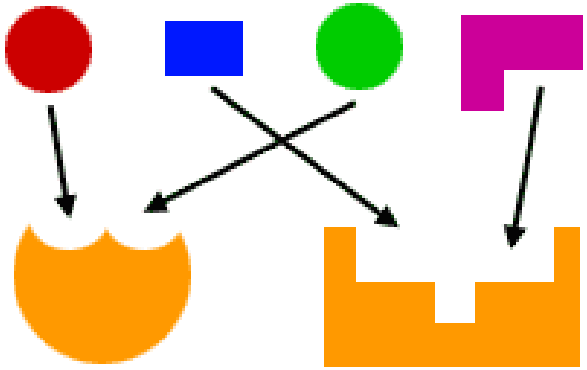


Biologically important proteins including membrane proteins:
Calcium pump, Rhodopsin, Bacterial flagella, Drug effilux protein and so forth.

Application to drug discovery

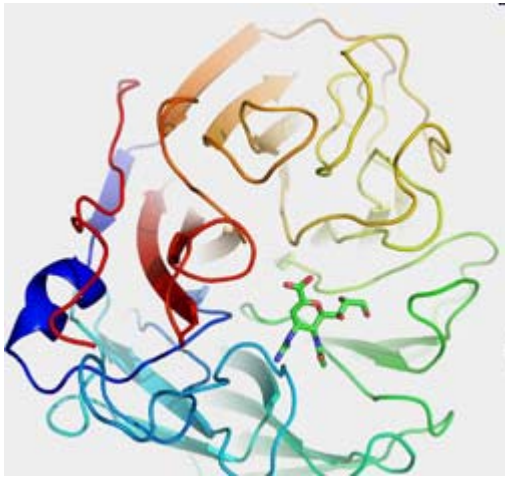
Enzyme

Reaction selectivity < Key and keyhole

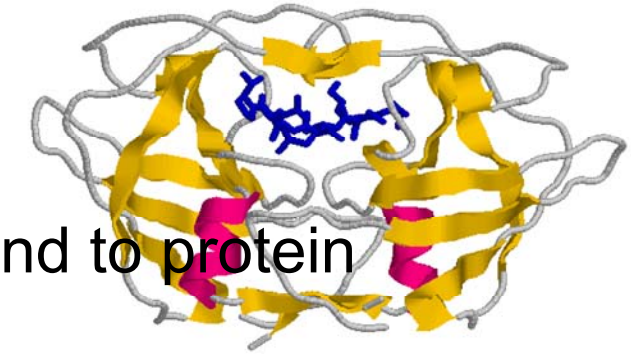


Only works a key (chemicals) can entry into the keyhole (enzyme)

Marked results



Relenza (Influenza)

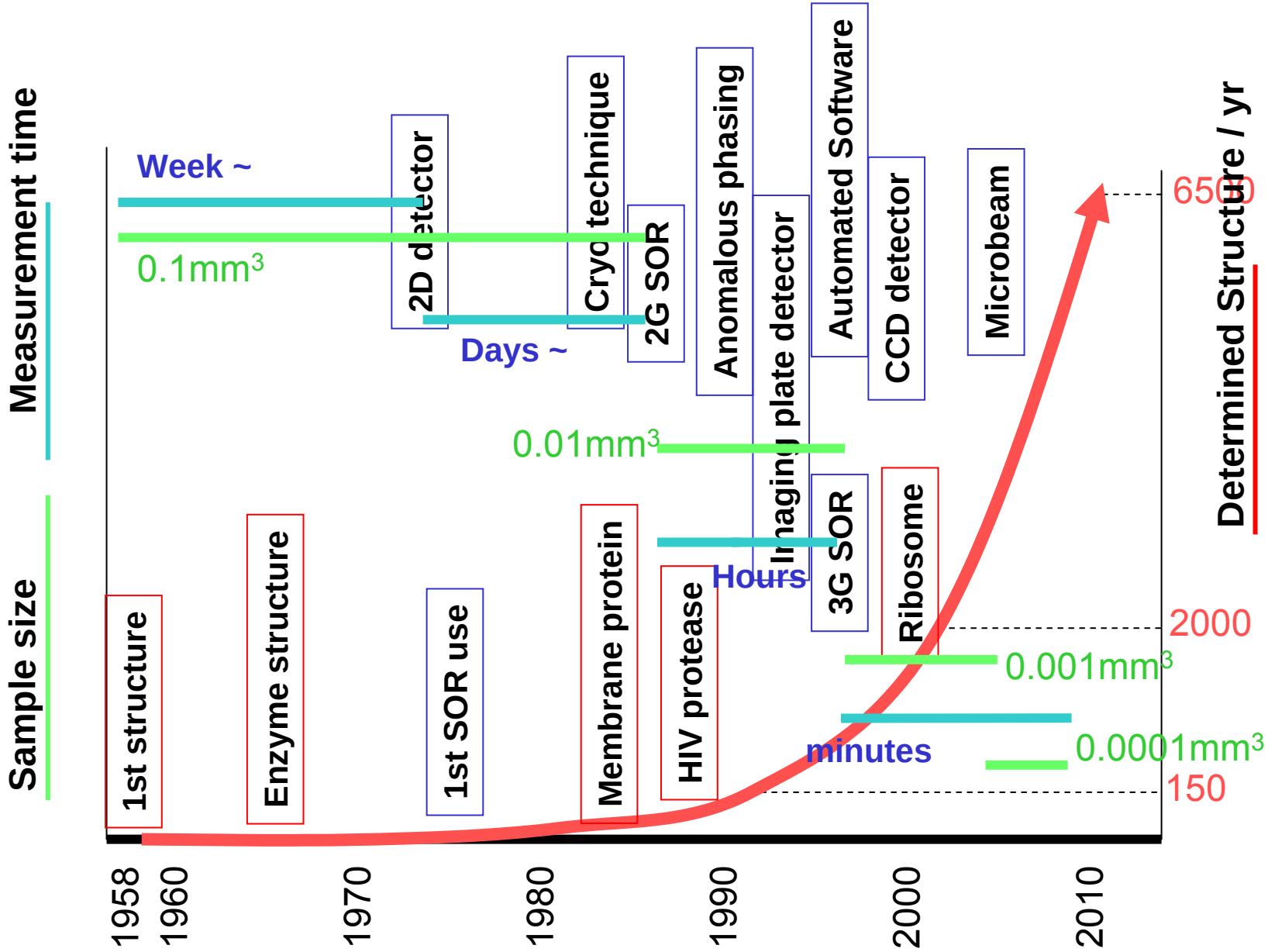


Indinavir (AIDS)

Drug Design: Chemicals recognize and bind to protein structure to regulate protein function.

- **National interests and drug discovery.**
- **Keen competition in drug development.**
- **Importance structural analyses of drug-target proteins**

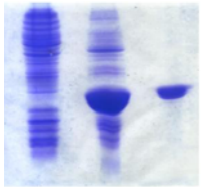
History of development in MX



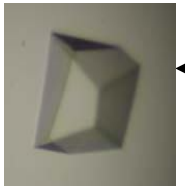
Advances in Protein Crystallography by Synchrotron Radiation

Steps in crystallographic analysis

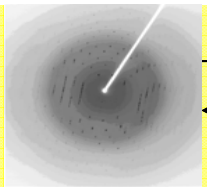
Sample preparation



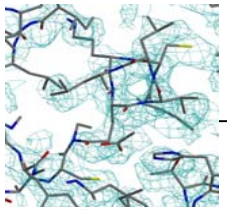
Crystallization



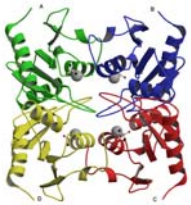
Diffraction Measurement



Phasing



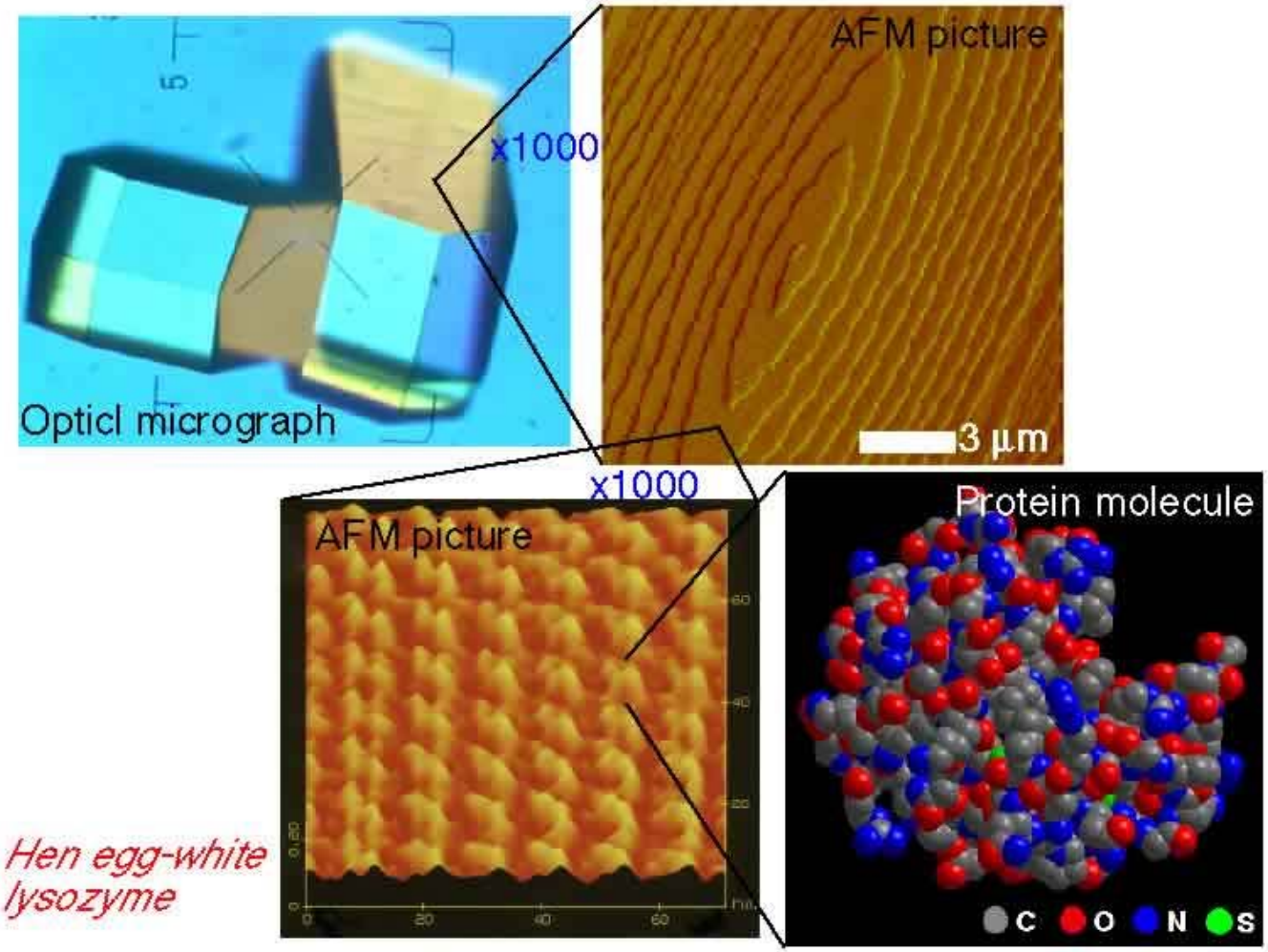
Structure Refinement



	Before SOR	After SOR	3G SOR
Small amount of samples for	10 mg~	0.1 mg~	
Small crystalline size	0.1 mm^3~	0.01 mm^3~	0.001 mm^3~
High speed data collection	**day ~ week**	**20 min ~**	**5 min ~**
New phasing method	month ~ year	Day ~ a few months	
Automated refinement by high resolution data	Month ~ year	Day ~ weeks	

Synchrotron data collection
> effective to not only X-ray measurement but also all other exp. steps in scale down / time reduction / high resolution.

2: Methodology of MX
2.1: Protein crystallization



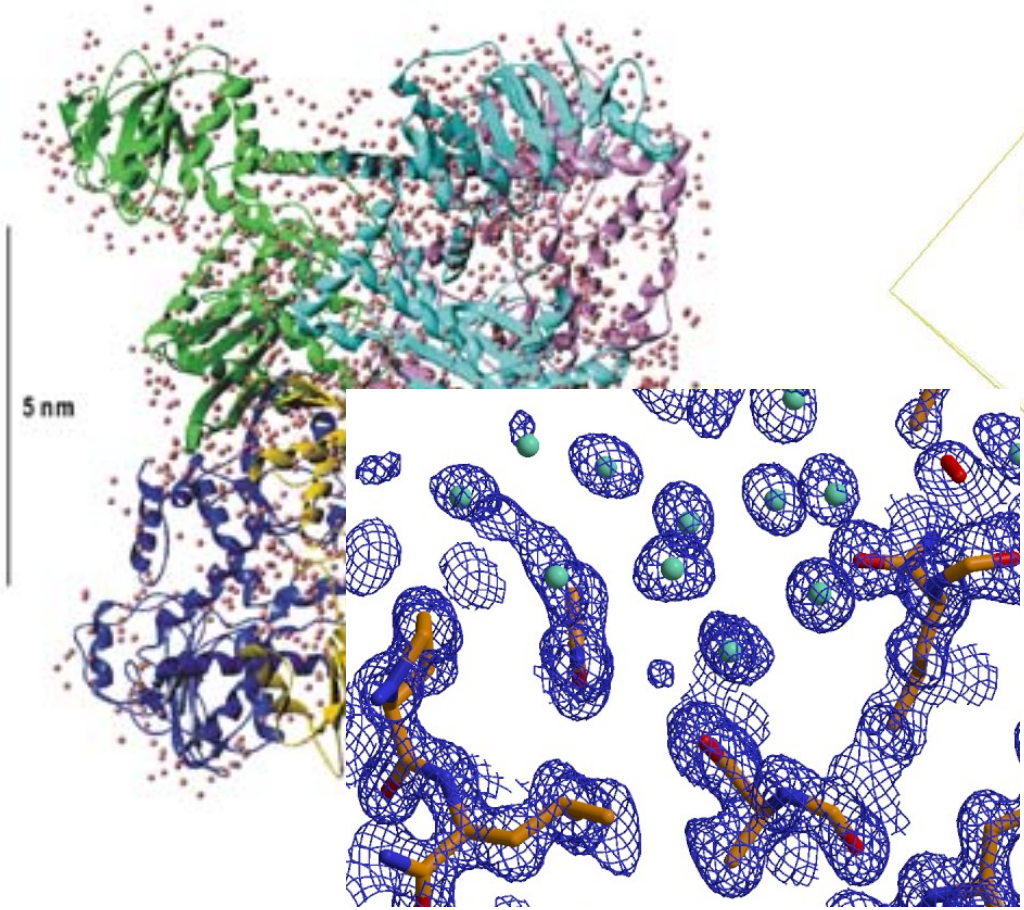
http://www.cir.tohoku.ac.jp/sazaki-p/%20Web_pages/Protein_crystallization.html

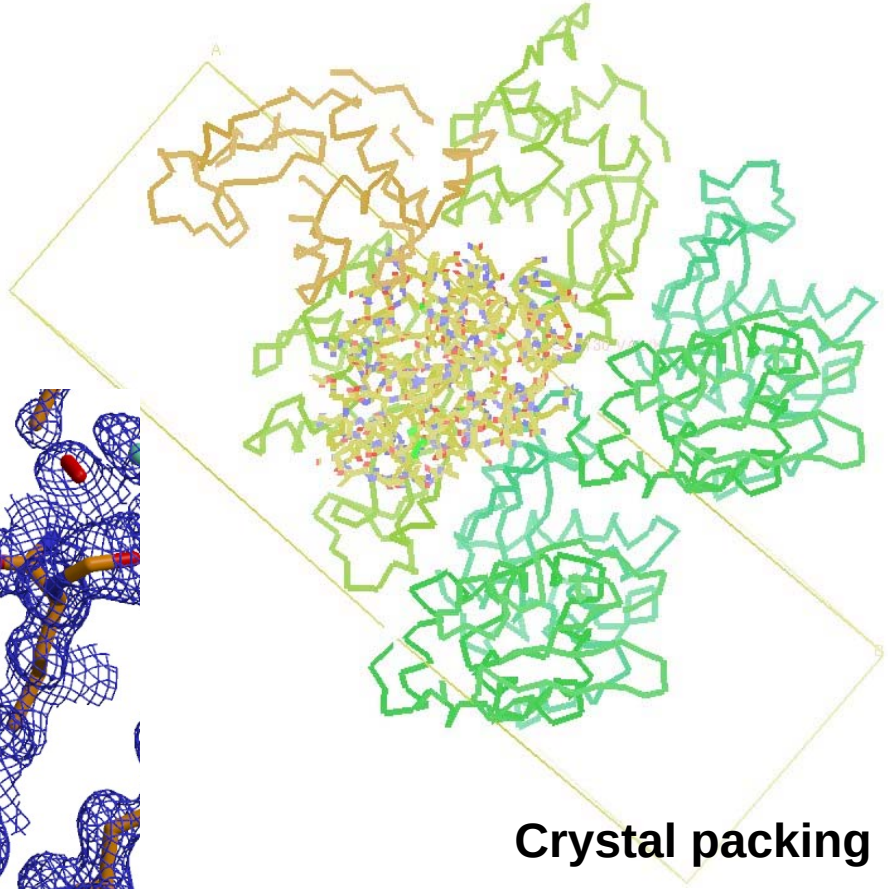
Hydration and packing

<u>Dissolution</u>:
Protein-water interaction
Water-water interaction

→ **Dehydration**

<u>Condensation</u>:
Protein-protein interaction





Crystal packing

Nakasako, M. *Philos. Trans. R. Soc. Lond. B* 359, 1191-204 (2004)

Factors determining protein solubility

Protein aggregation

Electrostatic interaction ~ oriented/ordered interaction

Hydrophobic interaction ~ random aggregation

・Kosmotropic ions: multivalent anion etc.

Hofmeister series: Salting-out effect

Stabilizing and enhancing water-water network

> Aggregate hydrophobic patches in proteins

・Organic solvent / uncharged polymer: acetone, DMSO, PEG etc.

Lower permittivity > Coulomb repulsion > aggregation

excluded volume effect in polymer > aggregation by osmotic pressure

・Detergents

Hydrophilization of protein hydrophobic regions > solubility ↑

・Chaotropic ions: larger ionic radii monovalent anion etc.

Guanidium ion, Urea, Iodine etc

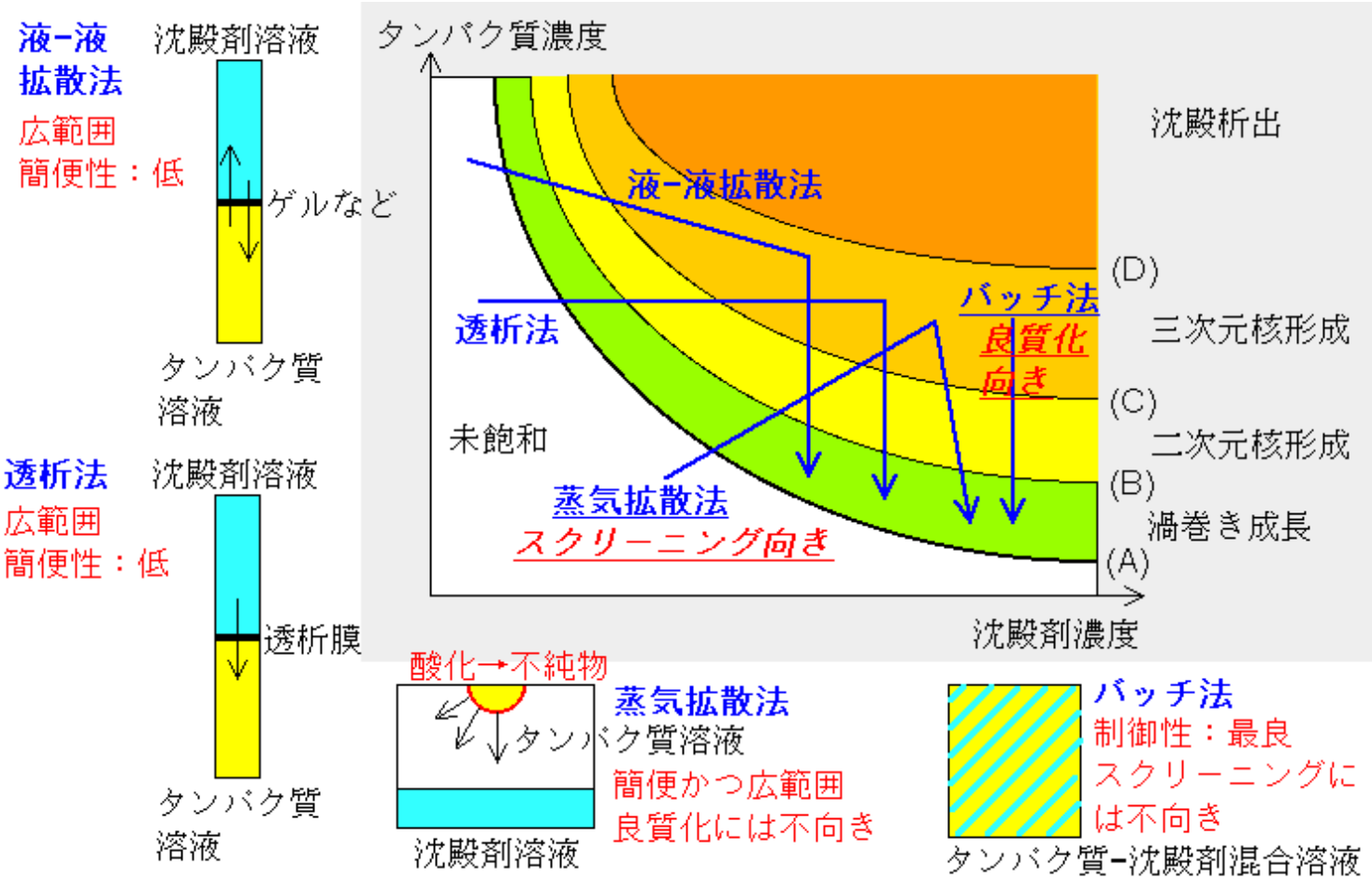
Destroy water cluster > enhance solvent entropy

・Amino acids / polyamine

Direct interaction with protein > Reduce protein-protein interaction

・pH Shift the surface charge in protein > modulate electrostatic interaction

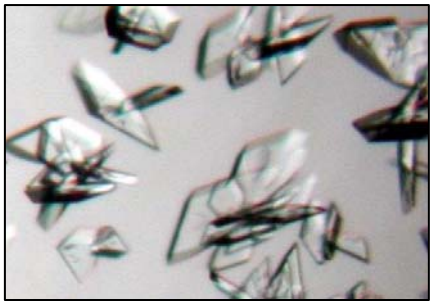
Phase Diagram



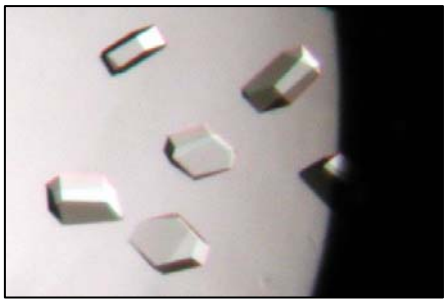
http://www.cir.tohoku.ac.jp/sazaki-p/%20Web_pages/Protein_crystallization.html

Seeding technique

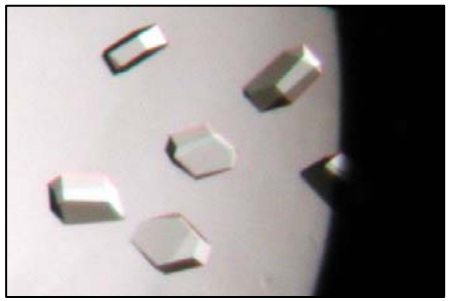
Optimizing crystal seed formation



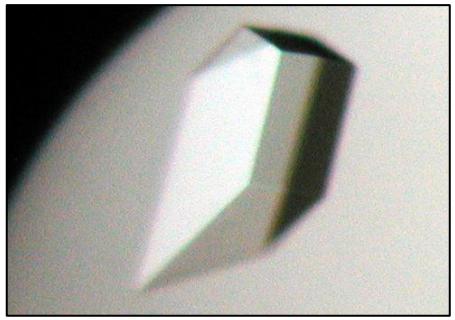
Clusters



Single crystals



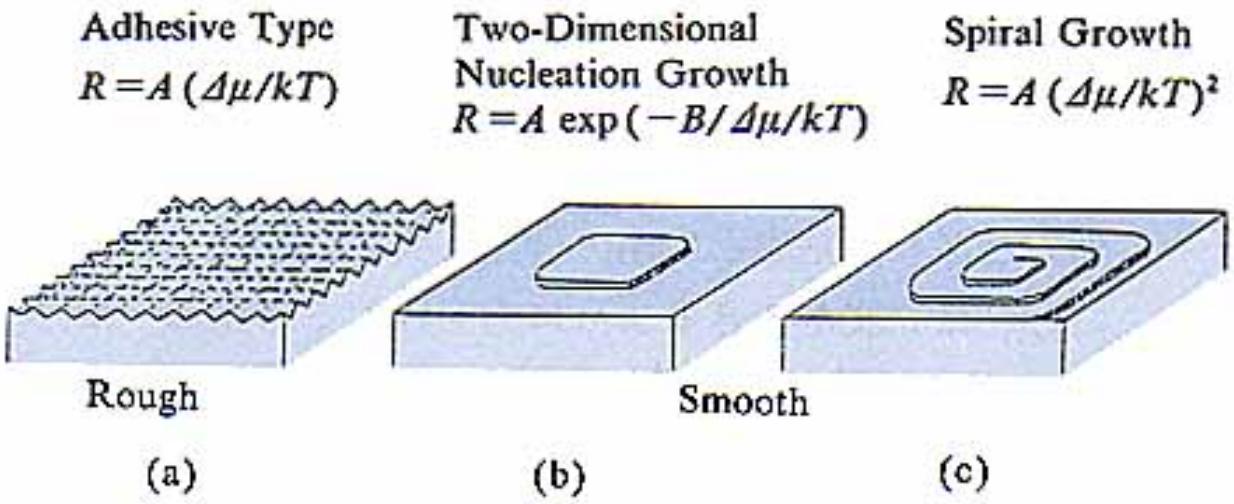
Small crystals

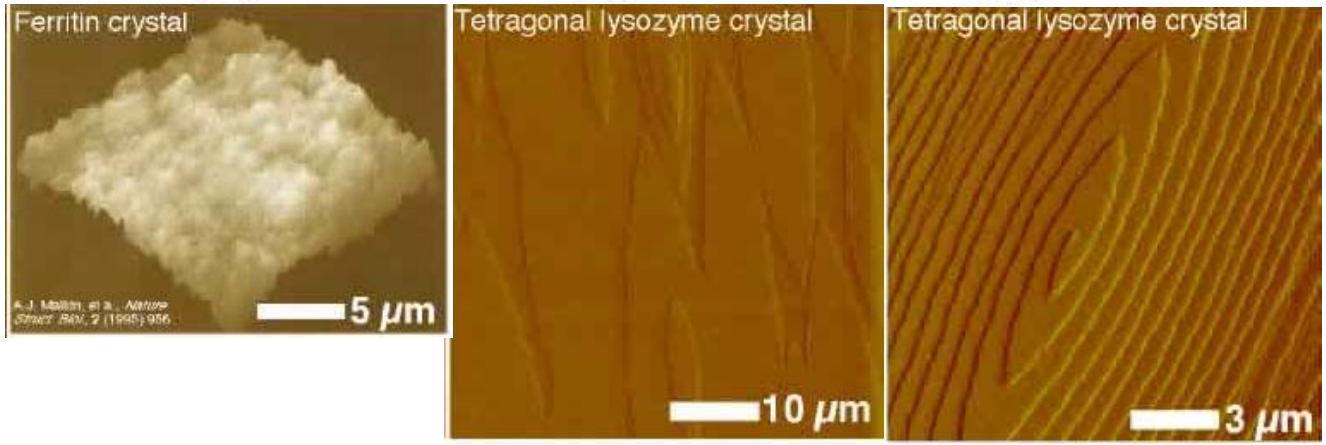


Size improvement

Macro Seeding / Micro Seeding / Streak Seeding / Heteroseeding

Type of crystal growth





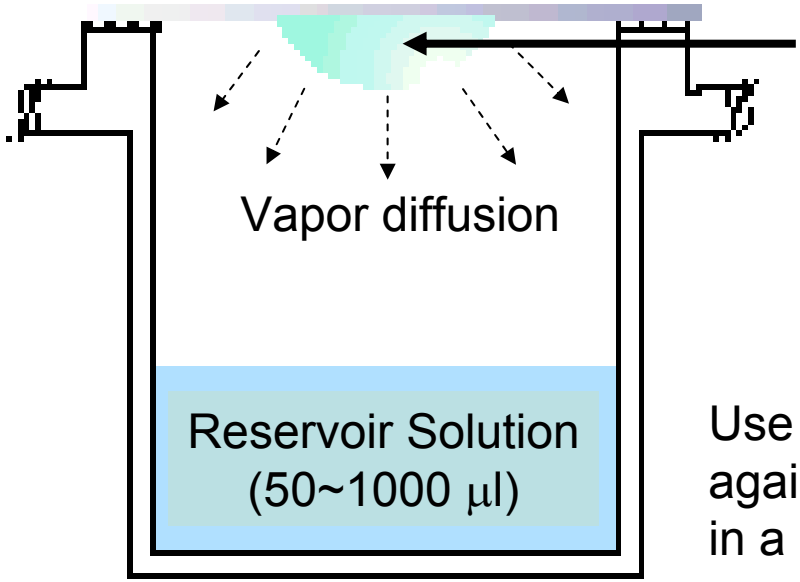
High ← Driving force of crystallization → Low

http://www.cir.tohoku.ac.jp/sazaki-p/%20Web_pages/Protein_crystallization.html

Crystallization method

Use the equilibrium against concentration gradient.

- Hanging drop vapor diffusion



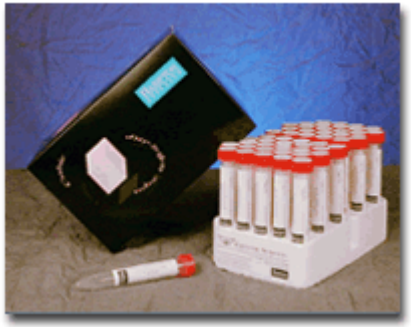
Protein Solution
(5~50 mg/ml,
1~10 μl)

Hanging drop vapor diffusion

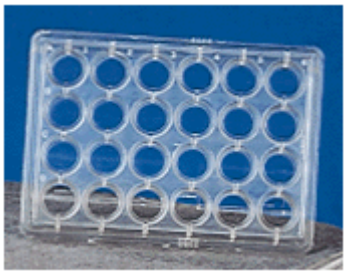
Use the equilibrium against concentration gradient in a sealed container.

- Sitting drop vapor diffusion
- Dialysis method: liquid-liquid contact via semipermeable membranes.
- Batch method: Simply mixing protein solution and precipitant.

Screening of crystallization condition



1. Initial Screening
   - Sparse Matrix Method
   - Screening kits available commercially
     - incl. 48~50 conditions
2. Optimization
   - Fine tuning solution parameters: Conc., pH, additives…



	1	2	3	4	5	6	
A	A.S. 1.0M	A.S. 1.1M	A.S. 1.2M	A.S. 1.3M	A.S. 1.4M	A.S. 1.5M	pH 7.0
B	A.S. 1.0M	A.S. 1.1M	A.S. 1.2M	A.S. 1.3M	A.S. 1.4M	A.S. 1.5M	pH 8.0
C	PEG 10 %	PEG 11 %	PEG 12 %	PEG 13 %	PEG 14 %	PEG 15 %	pH 7.0
D	PEG 10 %	PEG 11 %	PEG 12 %	PEG 13 %	PEG 14 %	PEG 15 %	pH 8.0

3. Heterogeneity check of sample solution
   - Impurity: Electrophoresis, Analytical HPLC
   - Instability: Enzymatic assay, Mass spectrometry
   - Oligomeric assembly: DLS (dynamic light scattering)

From precipitants to crystals

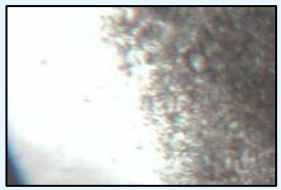
0

Clear drop

1

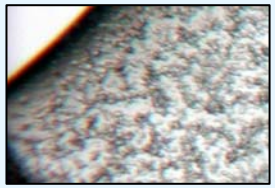
Brown aggregates

2



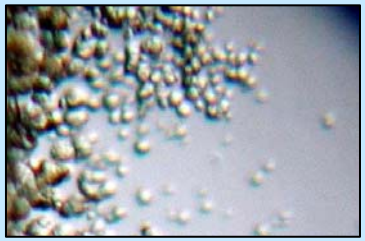
White aggregates

3



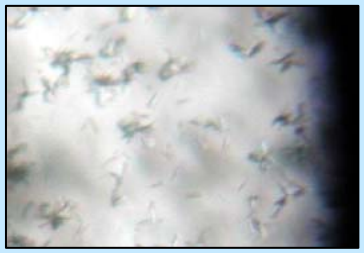
Transparent aggregates

4



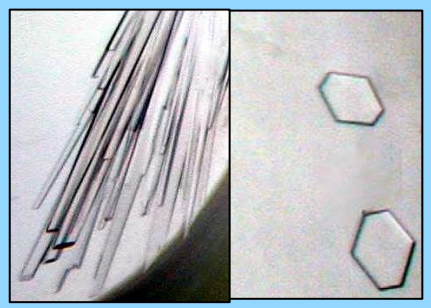
Non-crystalline particles

5



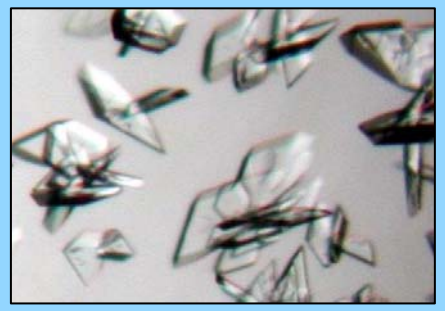
Microcrystals

6



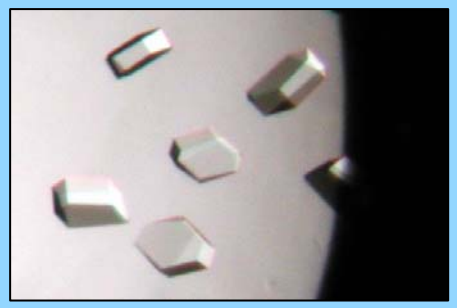
Needles Plates

7



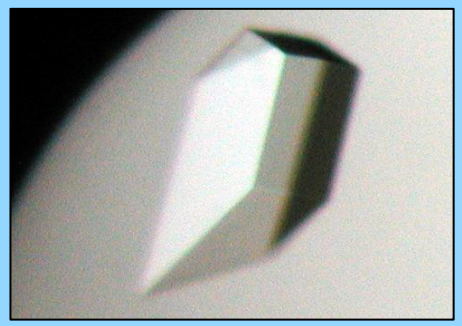
Clusters

8



Small crystals (0.05-0.2mm)

9



Large crystals (0.2mm ~)

Crystallization robot

自動結晶化観察ロボットシステム「TERA」 **RIKEN SPring-8**

ディスペンサーロボット **Dispenser robot**

結晶の顕微鏡写真

Microscopic image

結晶化ウェル自動観察装置

Automated monitoring system

Tabletop dispenser robot
(Thermo Scientific Hydra)

2.2: Data Collection

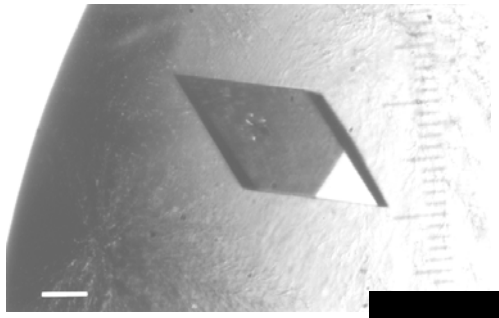
The Principle of X-ray Crystallography

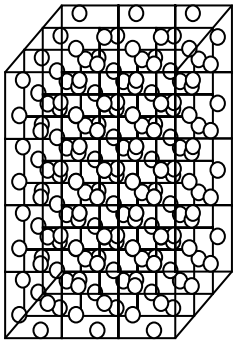
Real Space

$f\,(x\,y\,z)$

Atomic Scattering Factor
(Real Number)

Crystal





Atomic Coordinate

X-ray Diffraction

Fourier Transform

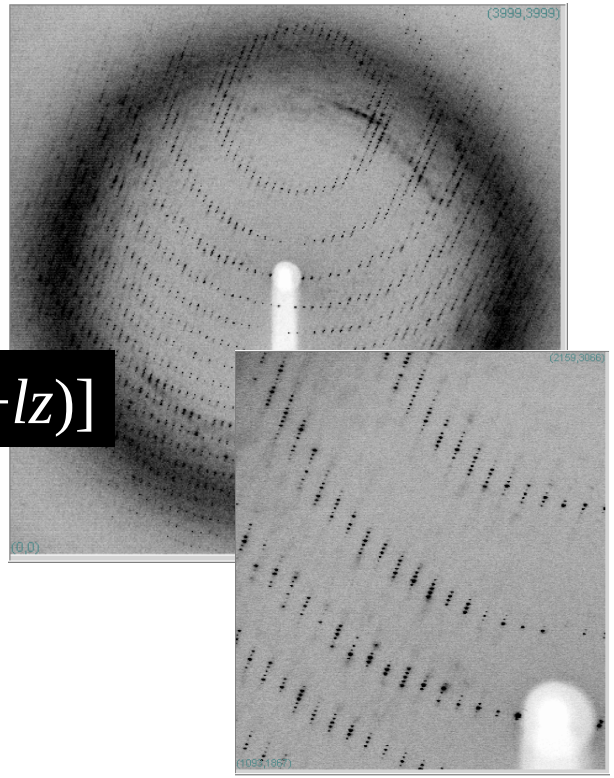
$$\boldsymbol{F}(hkl) = \Sigma\, f_j \exp[2\pi i(hx+ky+lz)]$$

Computation

Reciprocal Space

$\boldsymbol{F}\,(h\,k\,l)$

Structure Factor
(Complex Number)



Diffraction Spot

Crystal systems and space groups

7 systems & 230 space group

Triclinic: $P1$. $a \neq b \neq c$, $\alpha \neq \beta \neq \gamma \neq 90^\circ$

Monoclinic: $P2$, $P2_1$, $C2$. $a \neq b \neq c$, $\alpha = \beta = 90$, $\gamma \neq 90^\circ$

Orthorhombic: $P222$, $P2_12_12_1$, $F222$. $a \neq b \neq c$, $\alpha = \beta = \gamma = 90^\circ$

Trigonal: $P3$, $P3_1$, $P3_121$. $a = b \neq c$, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$

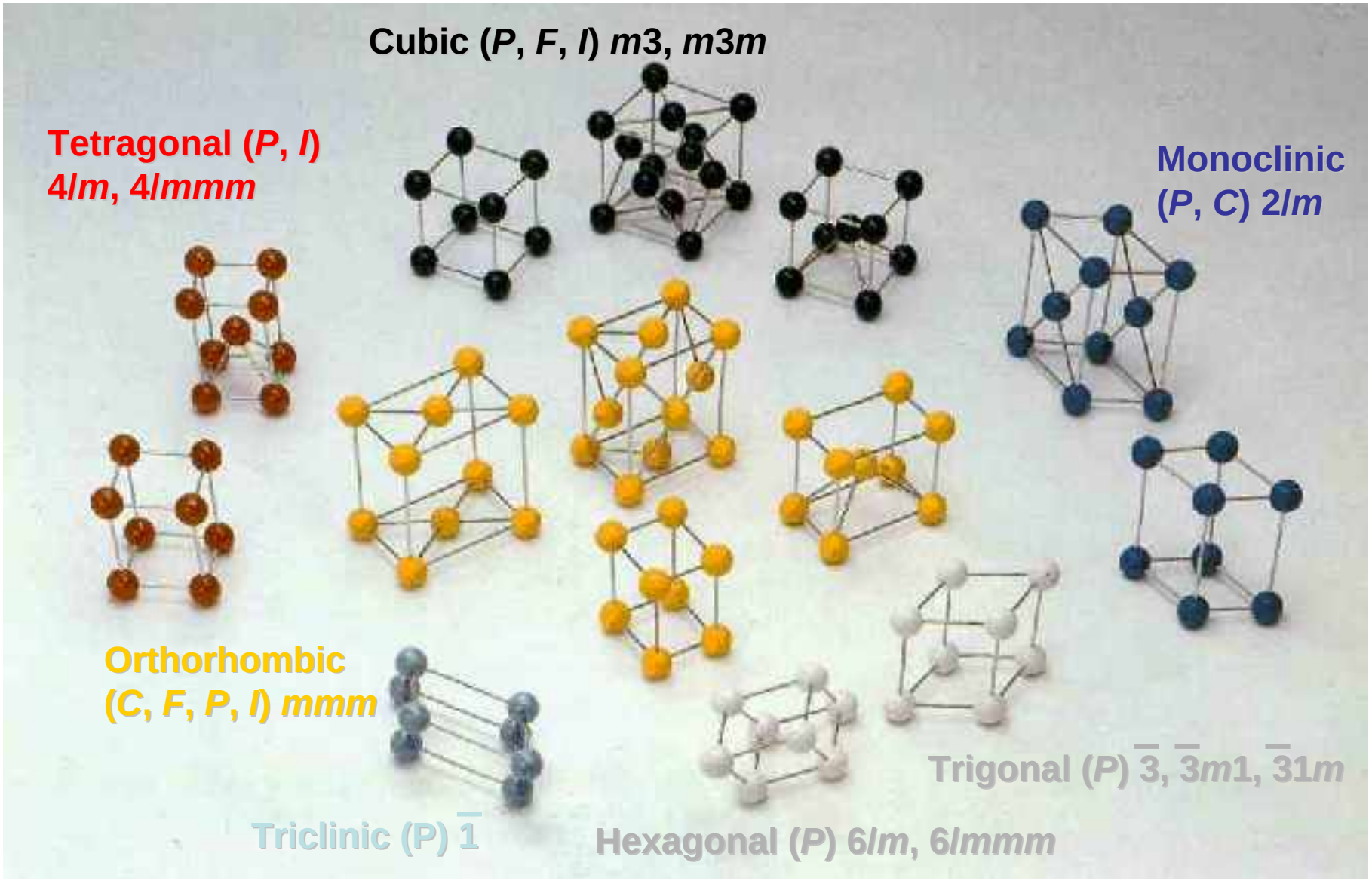
Hexagonal: $P6$, $P6_1$, $P6_122$. $a = b \neq c$, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$

Tetragonal: $P4$, $P4_1$, $P4_12_12$, $I4_122$. $a = b \neq c$, $\alpha = \beta = \gamma = 90^\circ$

Cubic: $P43$, $F432$, $P4_13_12$, $I432$. $a = b = c$, $\alpha = \beta = \gamma = 90^\circ$

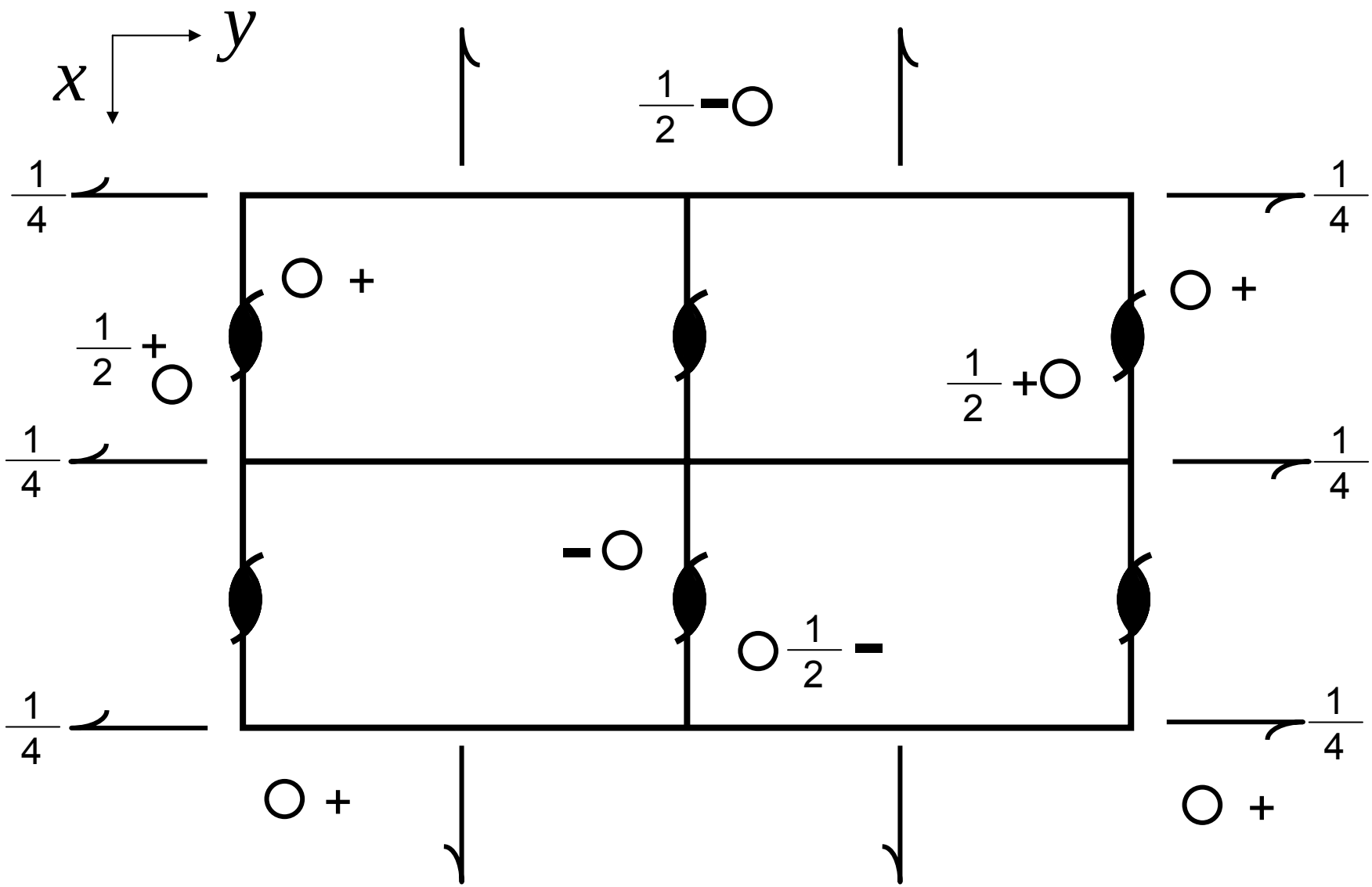
65 Enantiomorphic space groups (which do not have a center of symmetry due to the absence of inversions, mirrors, and glide plane elements) out of 230 all space groups, because protein molecules do not have such symmetry.

Bravais lattice and Laue group



Space group $P2_12_12_1$

1: x, y, z, 2: $\frac{1}{2}-x, y, \frac{1}{2}+z$,
3: $\frac{1}{2}+x, \frac{1}{2}-y, z$, 4: $x, \frac{1}{2}+y, \frac{1}{2}-z$



Relationship between real and reciprocal spaces

*P*2 (2-fold axis along y-axis through origin) > $\rho(x,y,z)=\rho(-x,y,-z)$

$$
\begin{aligned}
\mathbf{F}(hkl) &= V\int_{\mathrm{Cell}} \rho(xyz)\exp[2\pi i(hx+ky+lz)] \\
&= V\int_{\mathrm{Half-Cell}} \rho(xyz)\{\exp[2\pi i(hx+ky+lz)]+\exp[2\pi i(-hx+ky-lz)]\} \\
&= \mathrm{F}(\bar{h}k\bar{l})
\end{aligned}
$$

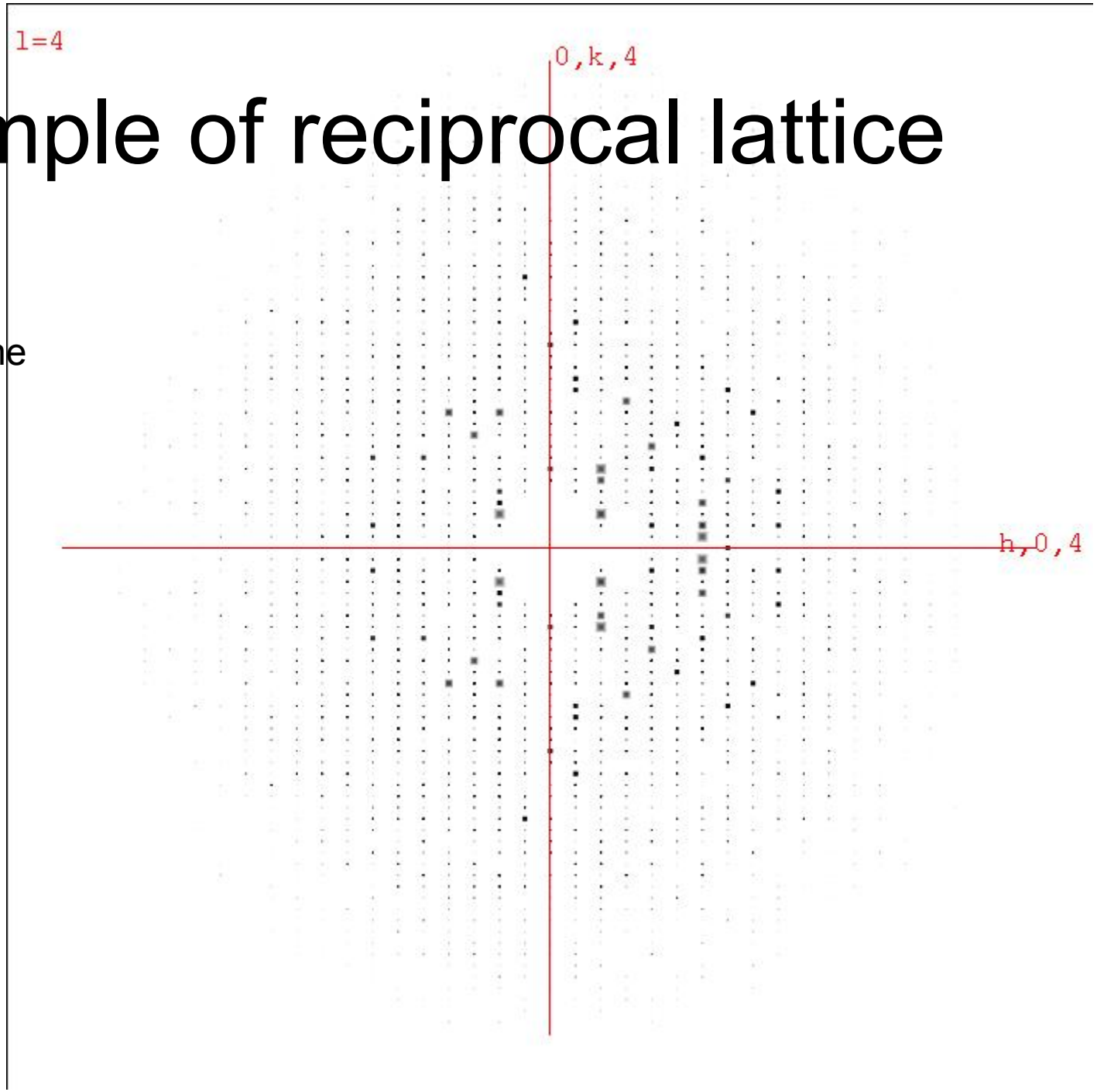
Other rules:

1) Screw axis shows same symmetry of its rotation with systematic absences
   ex) $P2_1$, F(hkl) = F(-h,k,-l), F(0,k,0) = 0 (k=2n)

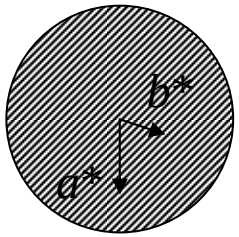
2) Complex lattice…
   ex) C-base centered, F = 0 (h+k = Odd)

An example of reciprocal lattice

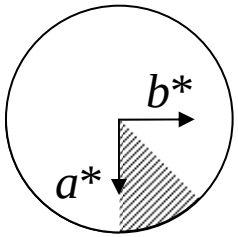
PDB ID: 1HF4
Egg white lysozyme
Space group: $P2_1$
Lattice constant:
$a = 27.94$
$b = 62.73$
$c = 60.25$
$\alpha = 90.0$
$\beta = 90.76$
$\gamma = 90.0$



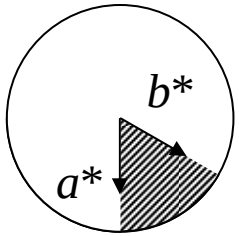
Asymmetric unit in reciprocal space



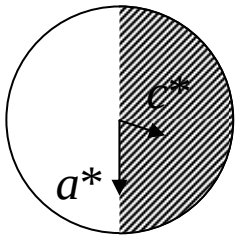
Triclinic
1
h k l
± ± +



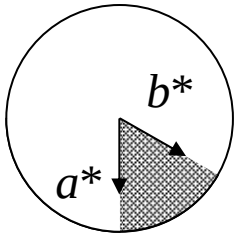
Tetragonal
4/mmm
+ + + (h≥k)



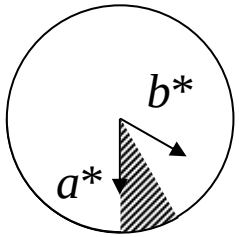
Hexagonal
6/m
+ + +



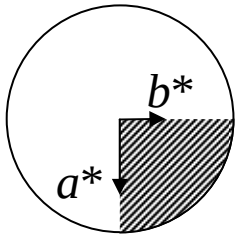
Monoclinic
2/m
± + +



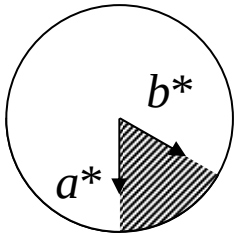
Trigonal
3
+ + ±



Hexagonal
6/mmm
+ + + (h≥k)

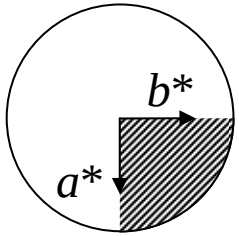


Orthorhombic
mmm
+ + +

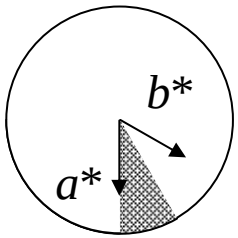


Trigonal
3m1
+ + +

Cubic
m3
+ + + (h≥l & k≥l)



Tetragonal
4/m
+ + +



Trigonal
31m
+ + ± (h≥k)

Cubic
m3m
+ + + (h≥k≥l)

X-ray diffraction data collection

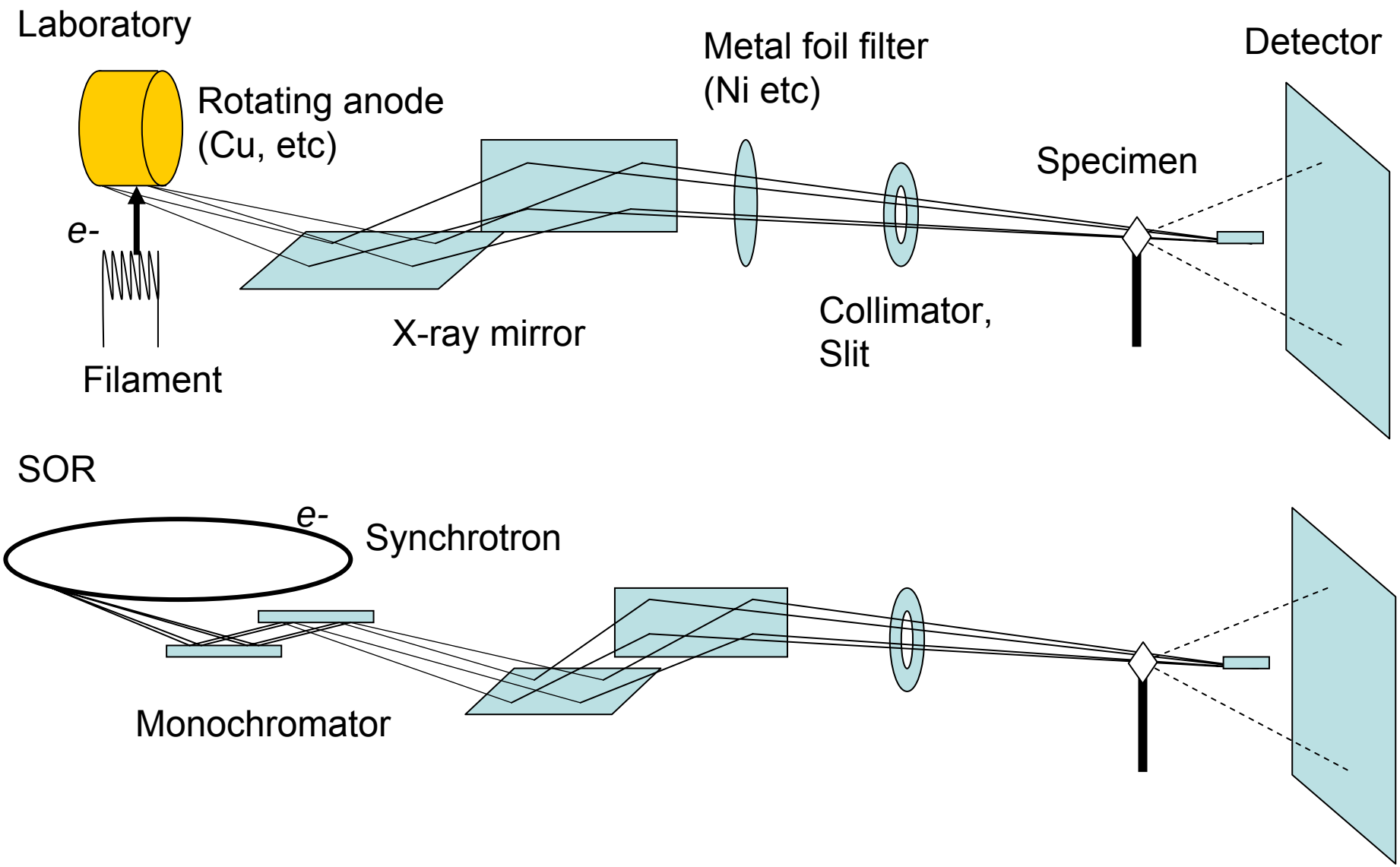
Essentials in high quality data collection:

Incident X-ray: Intensity, Divergence, Wavelength

Detector: Detection accuracy, Speed, Image resolution

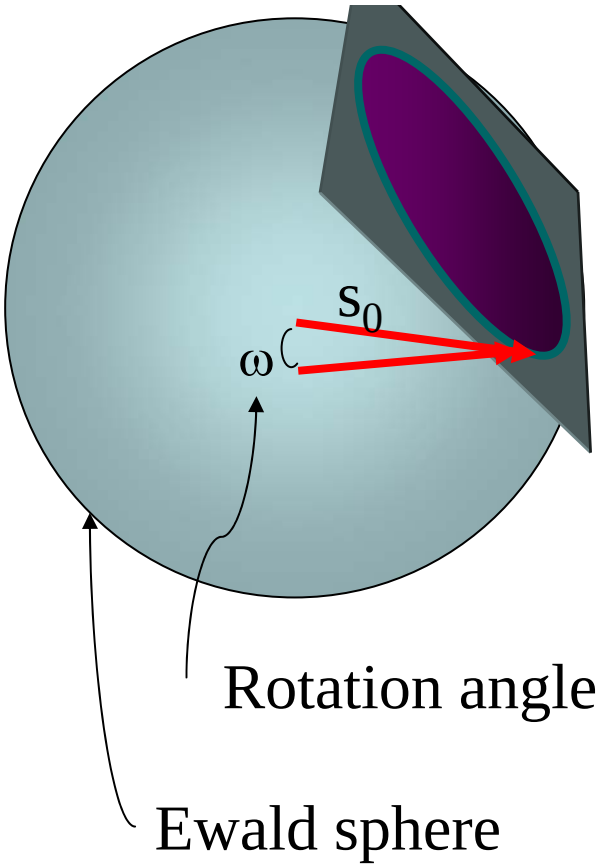
Crystal: Crystalline order, Size, Radiation resistance

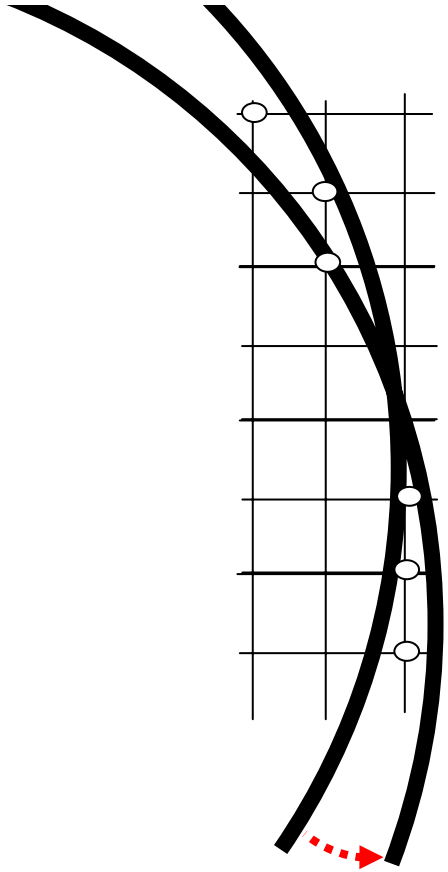
Experimental setup



Oscillation method

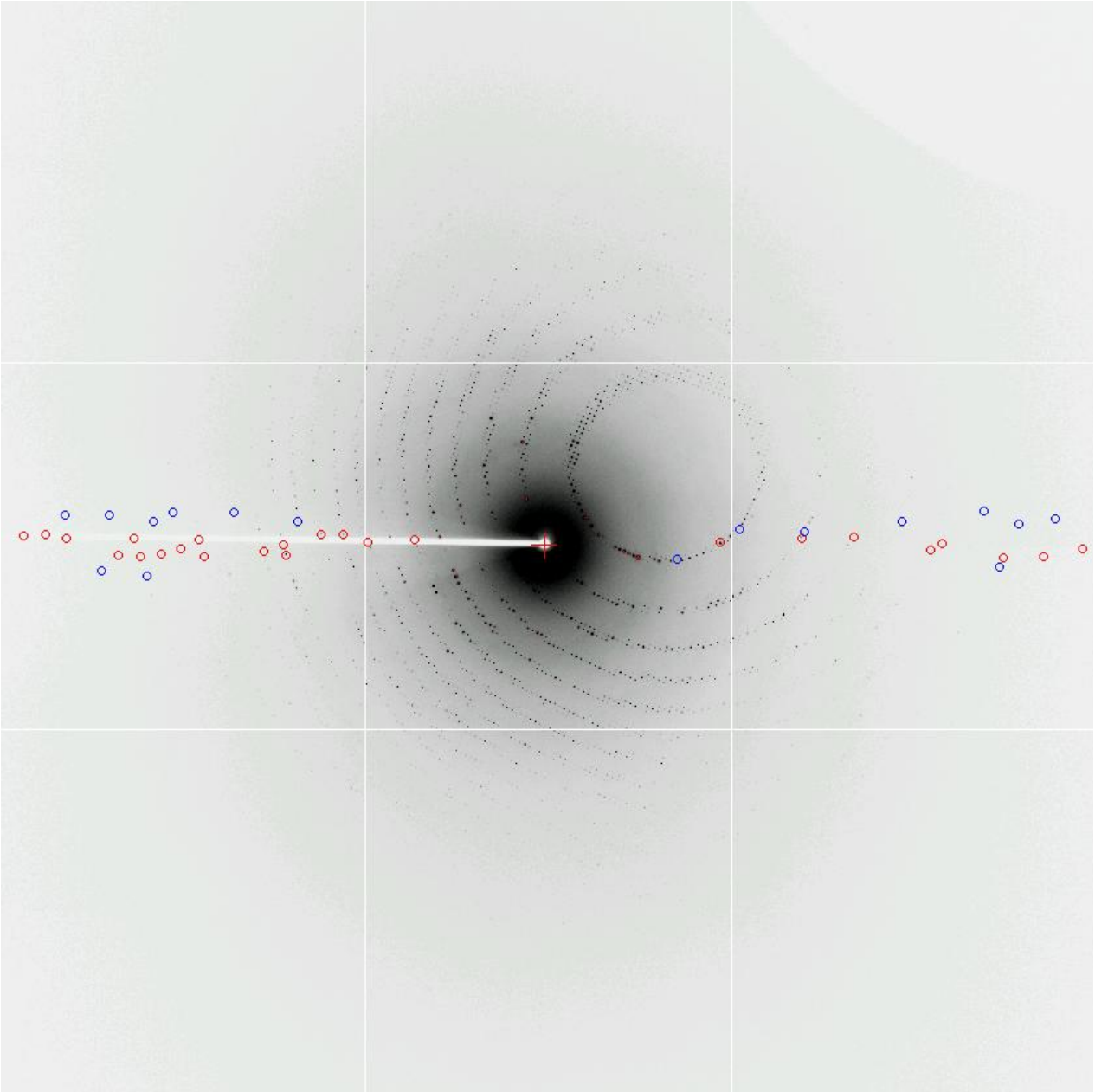
To record all individual reciprocal spots, the crystal is rotated with a step-width around one axis. The step-width images are processed to obtain a data set.



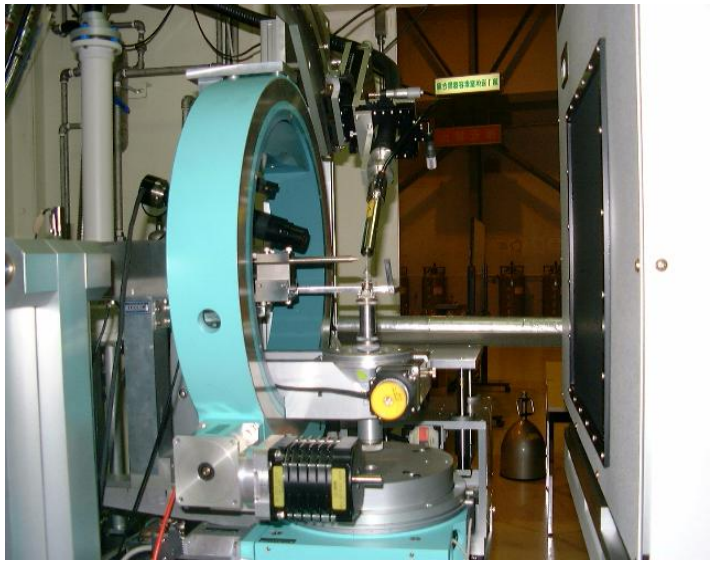


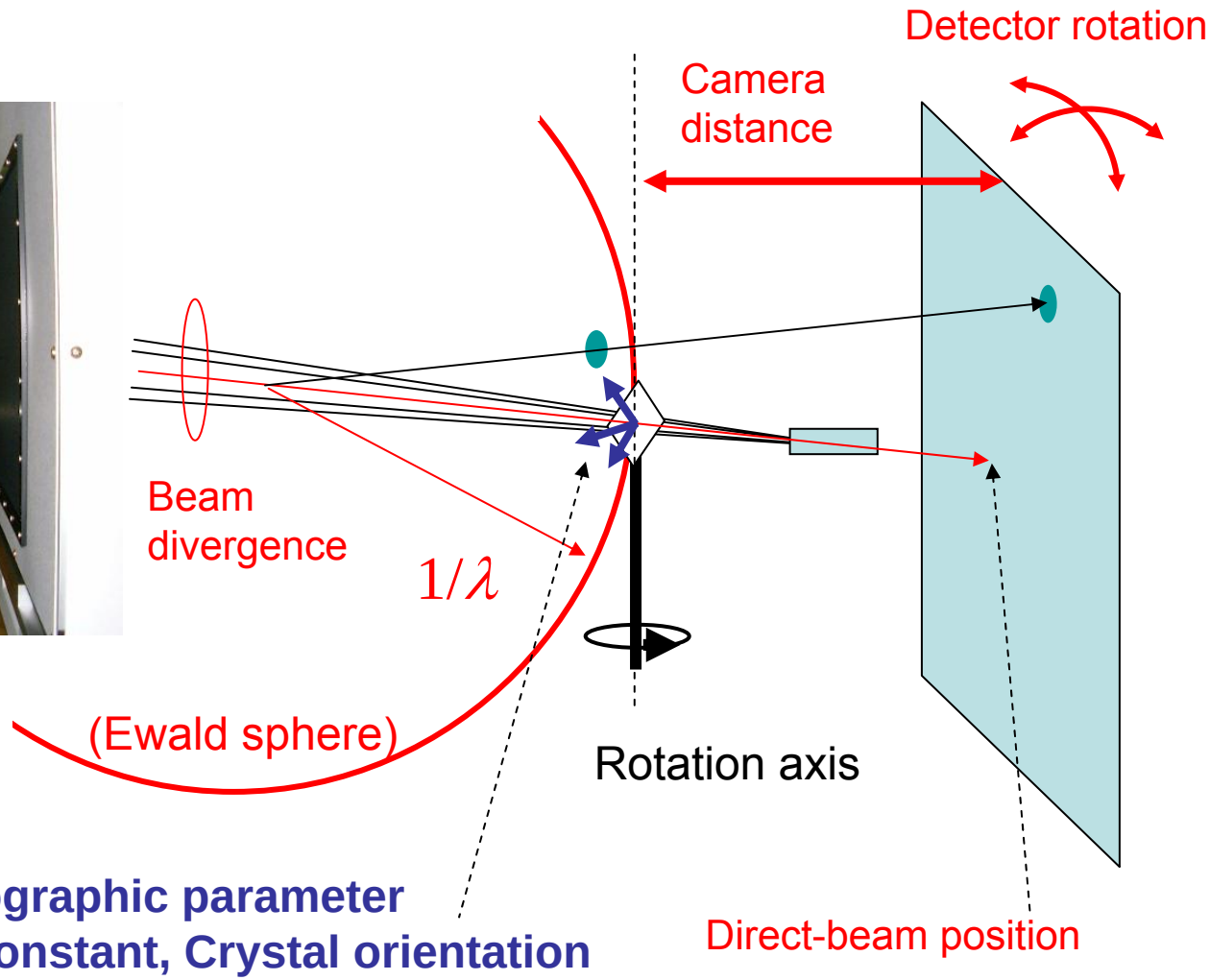
When a reciprocal points across the Ewald sphere, its intensity profile is recorded on detector.

A series of images



Parameters in oscillation method





Diffraction image processing

Obtain index (hkl) and intensity (*I*) of each diffraction spot
In collected from single wavelength & single crystal

Software

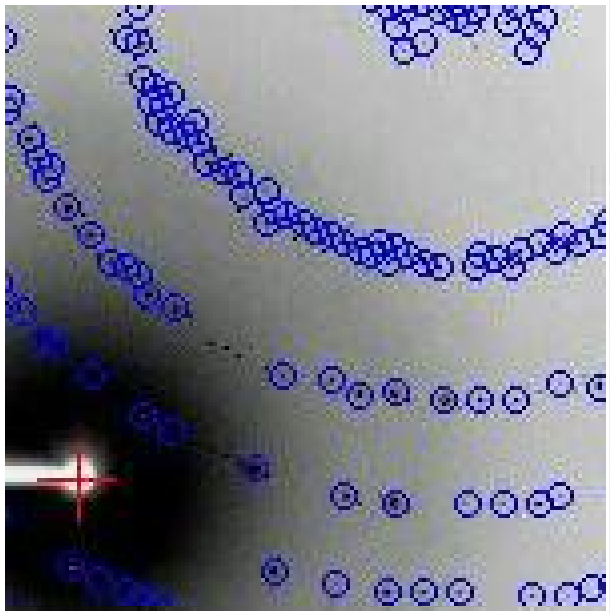
MOSFLM, XDS (free software)
HKL2000, CrystalClear

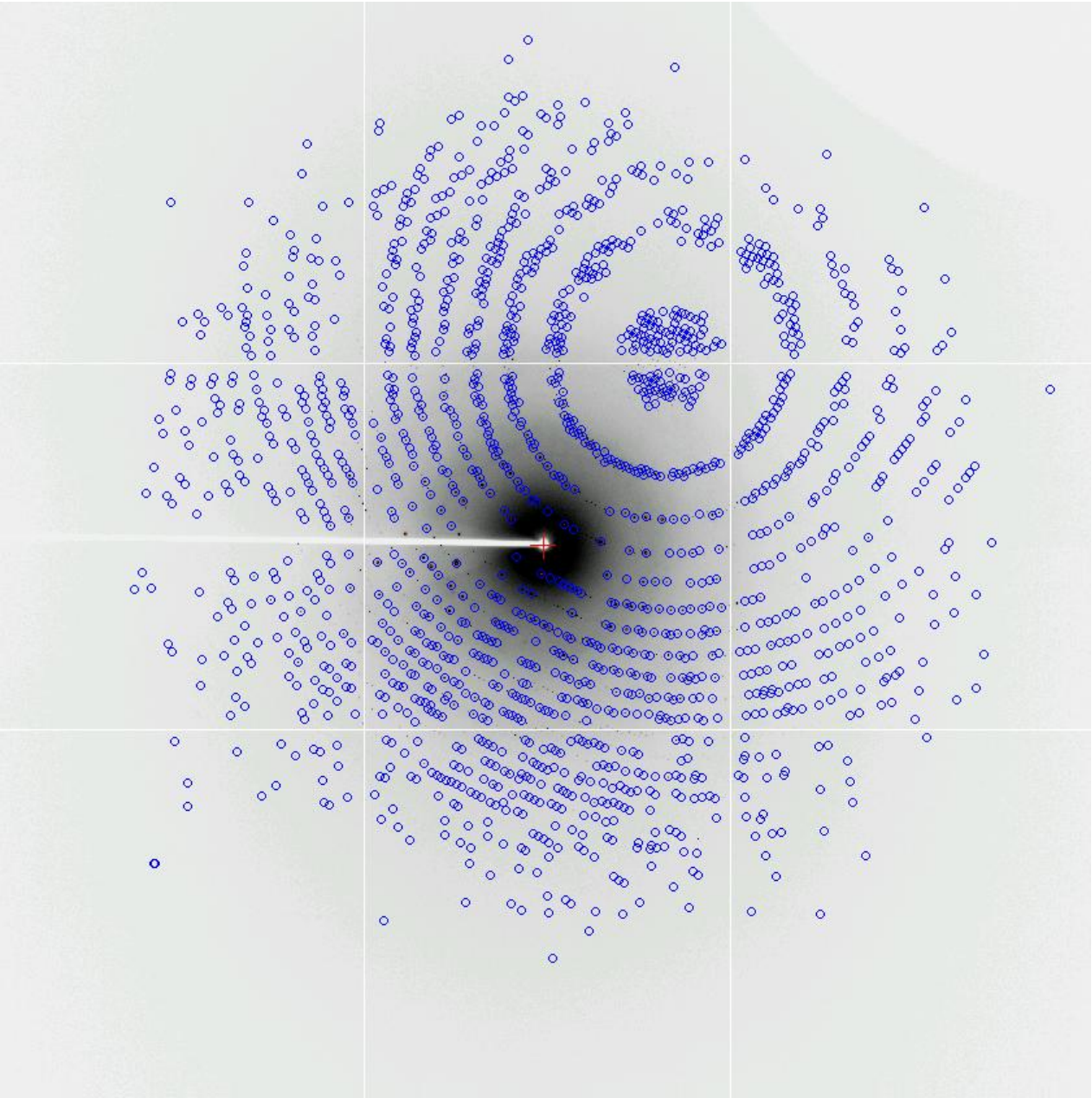
Steps of image processing

Indexing:	Determine parameters incl. lattice const.
Integration:	Calculating peak intensity
Scaling:	Merging & averaging equivalent reflections

Spot Finding

Find spots and calculate and record its coordinates on detector.





Autoindexing

Using spot positions, deduce possible crystal system and lattice parameters.

Choose a solution:

Soln	Least Sq	Spacegrp	Bravais	Lattice	a	b	c	Volume	α	β	γ
7	0.23	75	tetrago	P	77.02	77.02	37.44	222091	90.00	90.00	90.00
9	0.20	21	orthorh	C	108.87	108.97	37.44	444181	90.00	90.00	90.00
11	0.23	16	orthorh	P	37.44	77.01	77.03	222090	90.00	90.00	90.00
12	0.04	5	monocli	C	108.87	108.97	37.44	444181	90.00	90.00	90.00
13	0.12	3	monocli	P	37.44	77.01	77.03	222090	90.00	90.13	90.00
13b	0.17	3	monocli	P	37.44	77.01	77.03	222090	90.00	90.13	90.00
14	0.00	1	triclin	P	37.44	77.01	77.03	222090	89.95	89.87	89.90

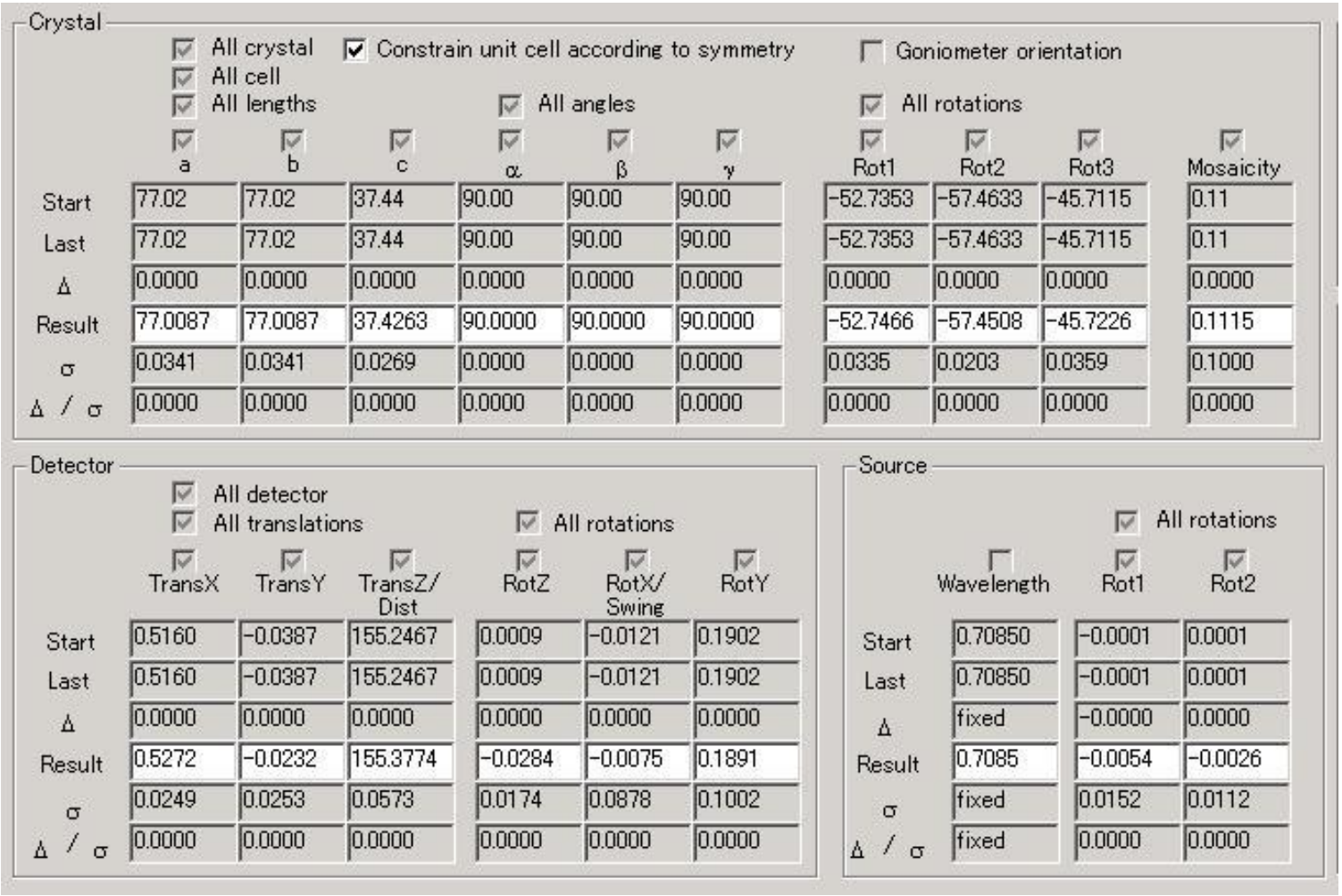
Lattice type

Lattice constant

Agreement between observed and calculated spot position

Refinement

Various parameters are optimized using spot positions



Crystal

All crystal ☑ Constrain unit cell according to symmetry ☐ Goniometer orientation
All cell
All lengths ☑ All angles ☑ All rotations

	a	b	c	α	β	γ	Rot1	Rot2	Rot3	Mosaicity
Start	77.02	77.02	37.44	90.00	90.00	90.00	-52.7353	-57.4633	-45.7115	0.11
Last	77.02	77.02	37.44	90.00	90.00	90.00	-52.7353	-57.4633	-45.7115	0.11
Δ	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Result	77.0087	77.0087	37.4263	90.0000	90.0000	90.0000	-52.7466	-57.4508	-45.7226	0.1115
σ	0.0341	0.0341	0.0269	0.0000	0.0000	0.0000	0.0335	0.0203	0.0359	0.1000
Δ / σ	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Detector

All detector
All translations ☑ All rotations

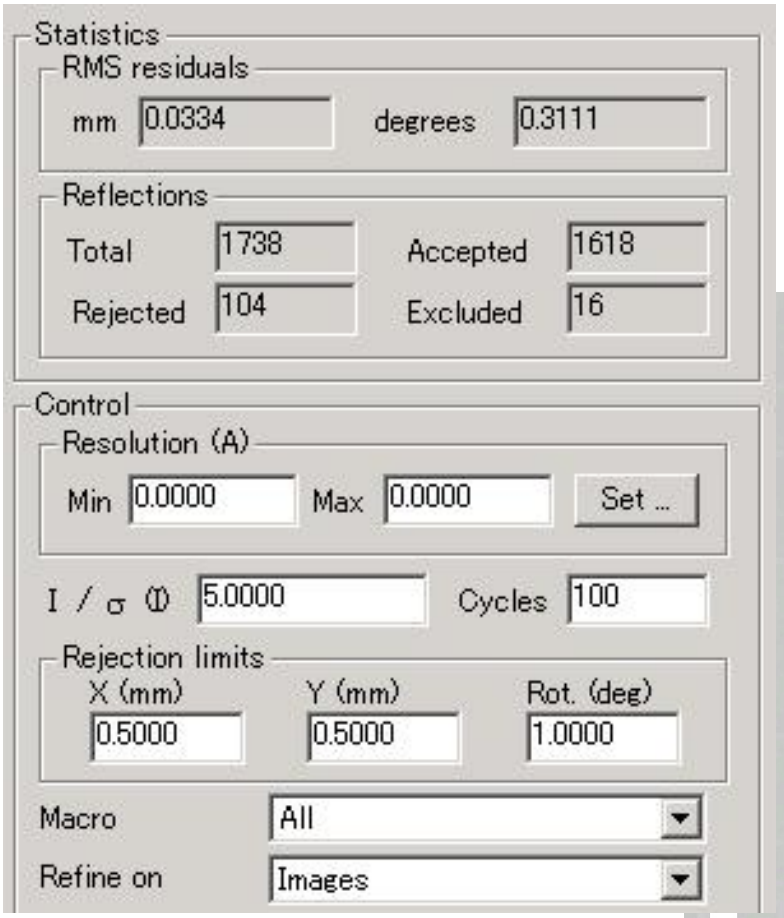
	TransX	TransY	TransZ/ Dist	RotZ	RotX/ Swing	RotY
Start	0.5160	-0.0387	155.2467	0.0009	-0.0121	0.1902
Last	0.5160	-0.0387	155.2467	0.0009	-0.0121	0.1902
Δ	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Result	0.5272	-0.0232	155.3774	-0.0284	-0.0075	0.1891
σ	0.0249	0.0253	0.0573	0.0174	0.0878	0.1002
Δ / σ	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

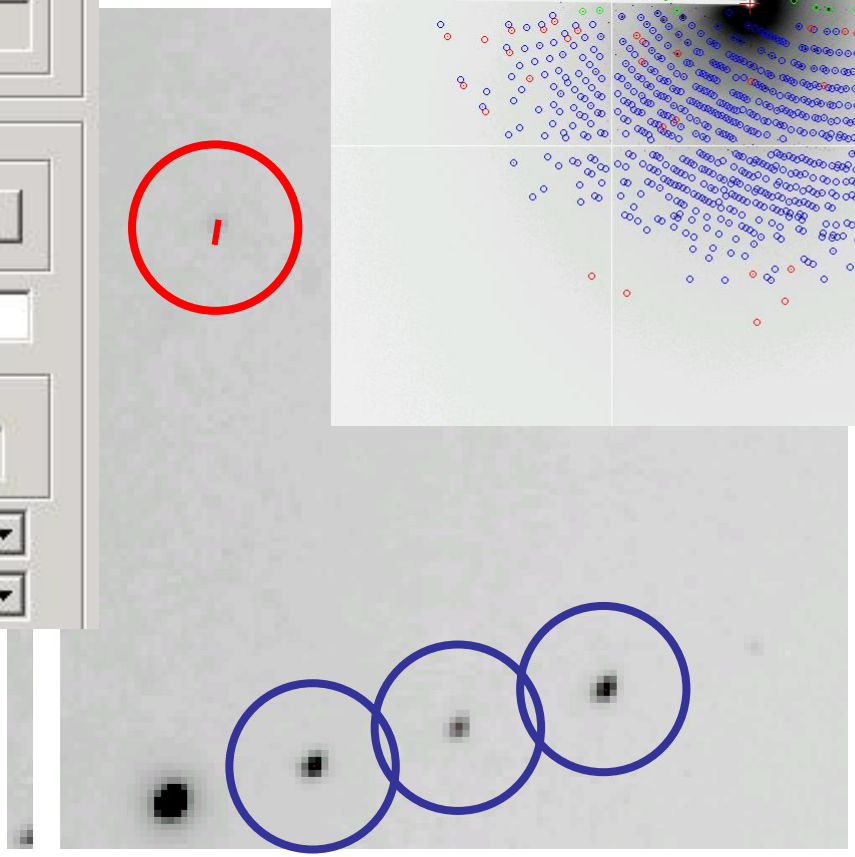
Source

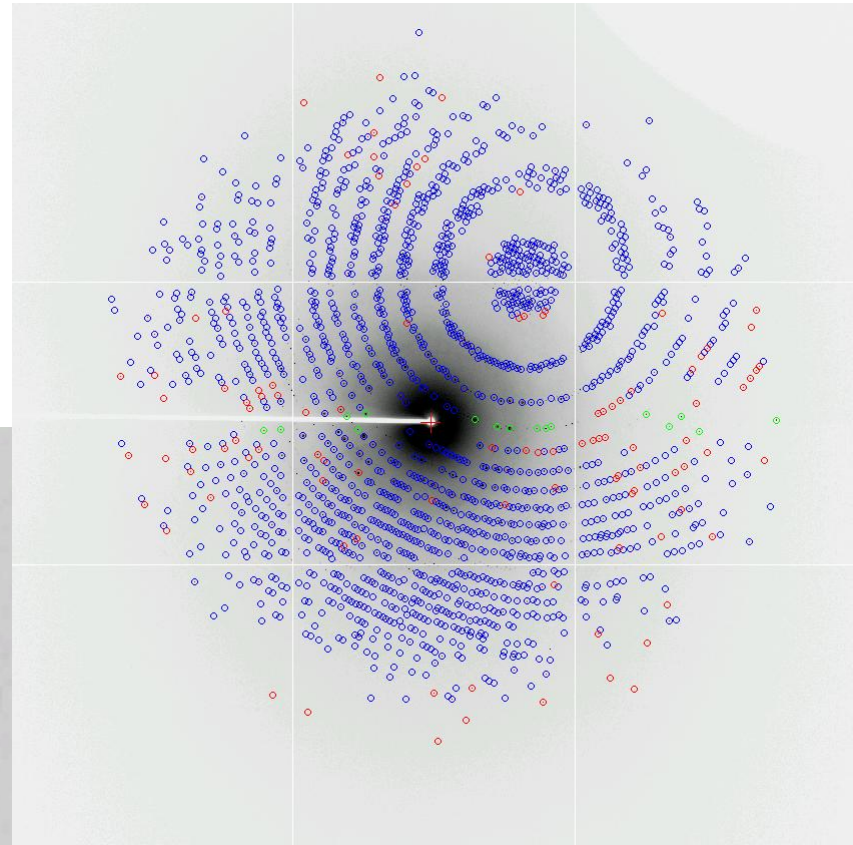
All rotations

	Wavelength	Rot1	Rot2
Start	0.70850	-0.0001	0.0001
Last	0.70850	-0.0001	0.0001
Δ	fixed	-0.0000	0.0000
Result	0.7085	-0.0054	-0.0026
σ	fixed	0.0152	0.0112
Δ / σ	fixed	0.0000	0.0000

Refinement

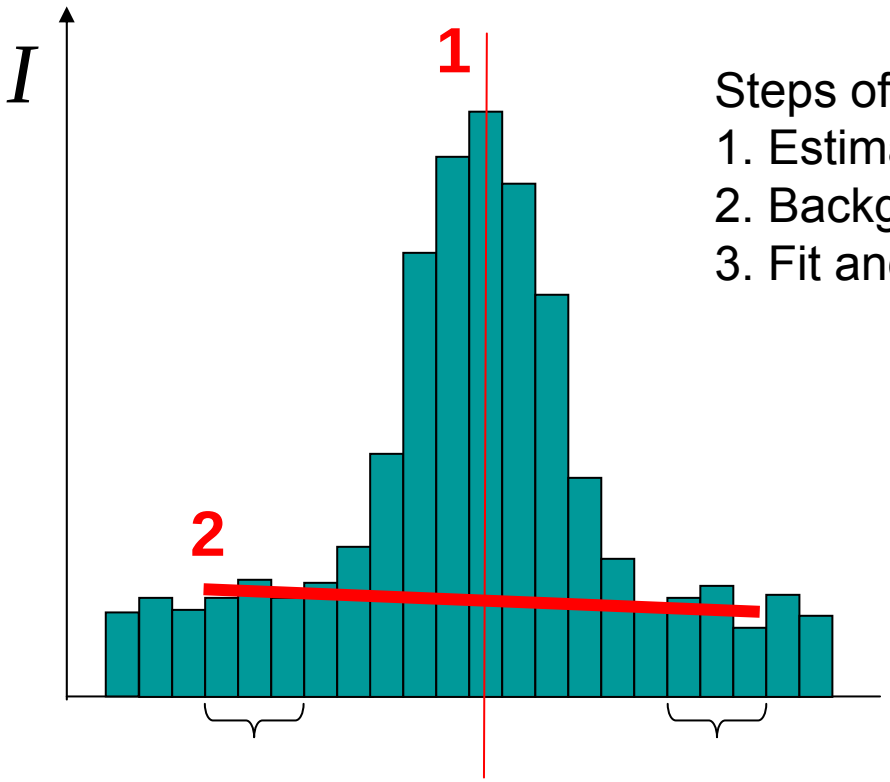

Statistics
RMS residuals
mm 0.0334
degrees 0.3111
Reflections
Total 1738
Accepted 1618
Rejected 104
Excluded 16
Control
Resolution (A)
Min 0.0000
Max 0.0000
Set ...
I / σ (I) 5.0000
Cycles 100
Rejection limits
X (mm) 0.5000
Y (mm) 0.5000
Rot. (deg) 1.0000
Macro All
Refine on Images





Integration

Integrate diffraction spot profile.



Steps of integration:
1. Estimate correct spot positions
2. Background estimation
3. Fit and integrate by averaged reflection profile

Scaling

Equivalent intensity among symmetrically equivalent reflections

ex. $P2_1$; (x, y, z), $(\bar{x}, y+1/2, \bar{z})$

$$I(h\ k\ l) = I(\bar{h}\ k\ \bar{l})$$

$$I(\bar{h}\ \bar{k}\ \bar{l}) = I(h\ \bar{k}\ l)$$

Estimate scale and falloff factor in each plate

Variation of incident intensity, absorption by crystal, etc.

during one data set

Rmerge overall:

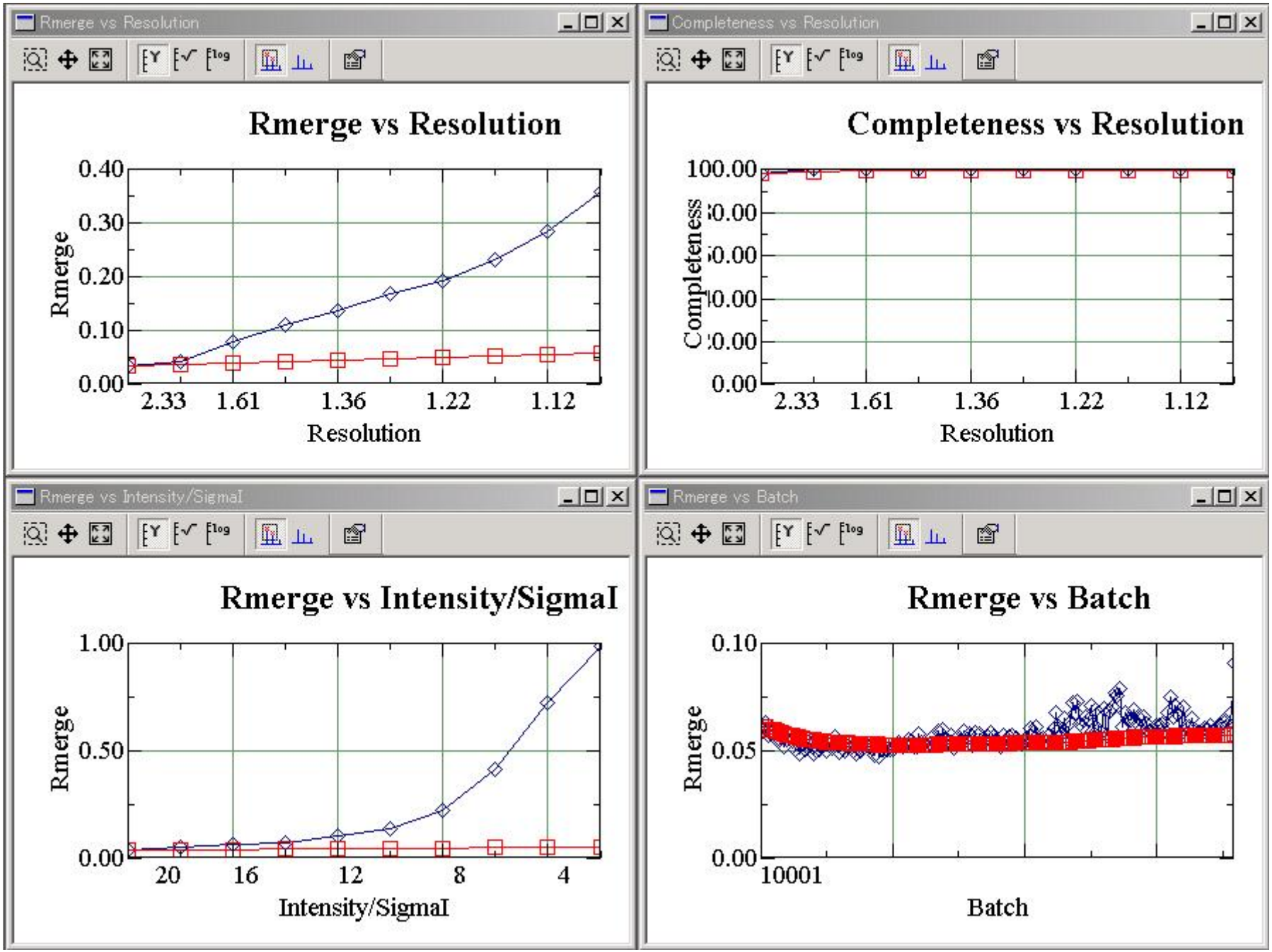
Measures the agreement of symmetry related observations of a reflection.

Rmerge in the last shell:

Rmerge in the highest resolution shell.

I/sigma:

A measure of the signal to noise ratio.


Rmerge vs Resolution
Rmerge
0.40
0.30
0.20
0.10
0.00
2.33
1.61
1.36
1.22
1.12
Resolution
Completeness vs Resolution
Completeness
100.00
0.00
2.33
1.61
1.36
1.22
1.12
Resolution
Rmerge vs Intensity/SigmaI
Rmerge
1.00
0.50
0.00
20
16
12
8
4
Intensity/SigmaI
Rmerge vs Batch
Rmerge
0.10
0.05
0.00
10001
Batch

Theory of error ~ Signal-Noise Ratio; S/N

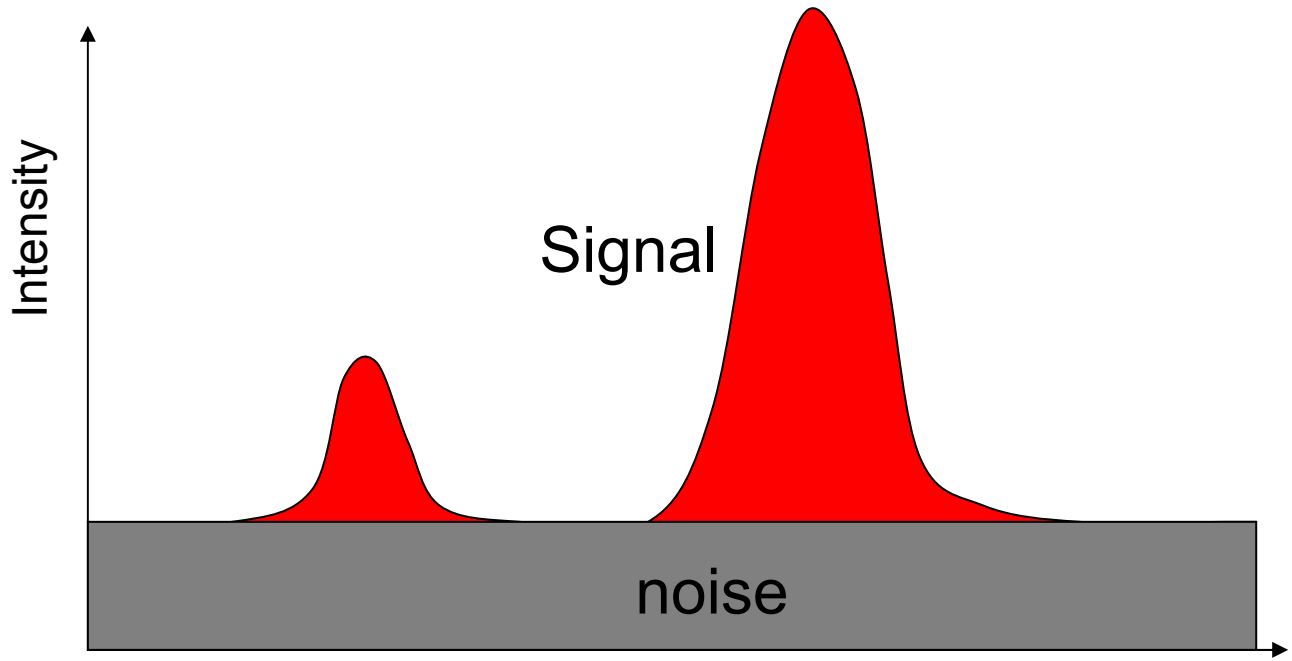
Signal: Diffraction intensity ~ Dose dependent

Noise: Radiation damage ~ Dose dependent

Scattering noise ~ Dose dependent

Detector dark noise ~ Time dependent

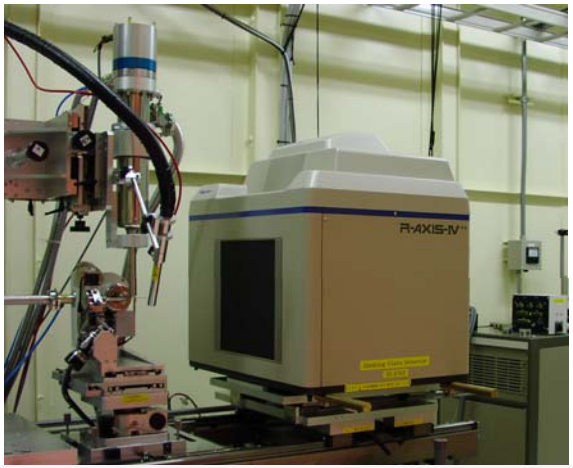
Detector readout noise ~ Image number dependent



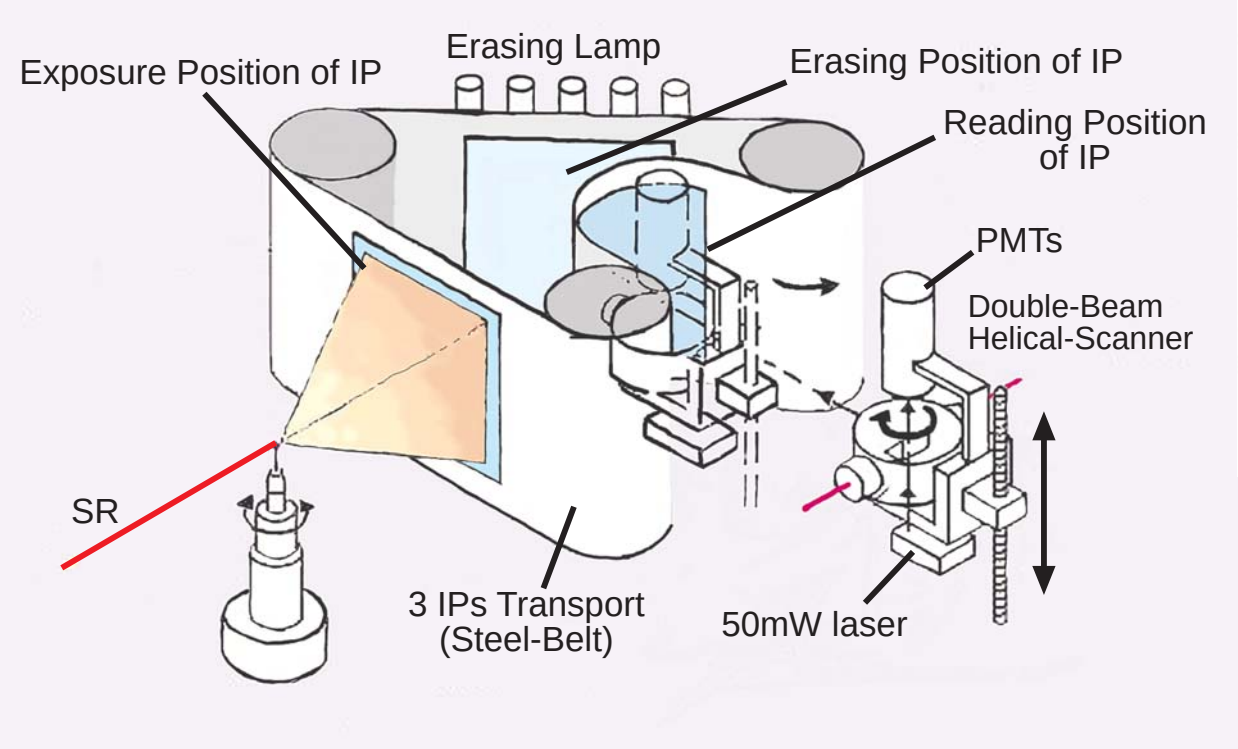
2D detectors for MX

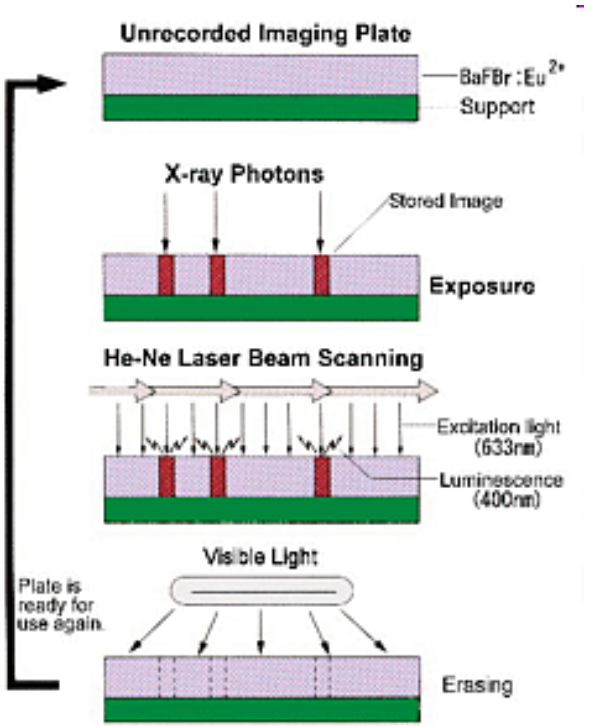
	CMOS	CCD		Amorphous Selenium	Silicon Pixel	IP
		Indirect	Direct			
Area size (100-400mm)	○ Multi-element	○ ME+FOT	△ cm sq. order	◎ by processing tech.	○ ME	◎
Resolution (50-100μm)	◎ Few-200 μm Phosphor	○ 10 - 100 μm FOT&phosphor	◎ Few μm	○ 100~200 μm	△ ~200 μm	○ 50 μm~
Readout Speed	◎ Sub mSec Continuous readout	○ Sec		○ Sec	◎ Real time Counting	△ Min
Sensitivity	△～◎ Phosphor & Window	△～◎ Phosphor		△	△～◎ Low for high E photon	◎
Noise	△～○ Relatively high readout & dark noise	○ Successful Cooling Phosphor/FOT/Window		△ Higher noise by polycrystalline	◎ Counting (counting loss at high dose)	△ Stray light of laser / Loss of fluorescence Capture
Skew	◎	△ FOT	◎ Direct	◎ Direct	◎ Direct	○～◎ Geometry at readout
Dynamic range	△ ~12bit	○ ~16 bit		○ ~16 bit	◎ ∞ (Counting)	◎ ~20 bit
Cost	◎ Versatile Processing technology	△～○ Complex system	◎ Cheap but small	? Expecting Future development	△ Original tech. and monopolistic	◎ Simple and matured technology

Imaging plate



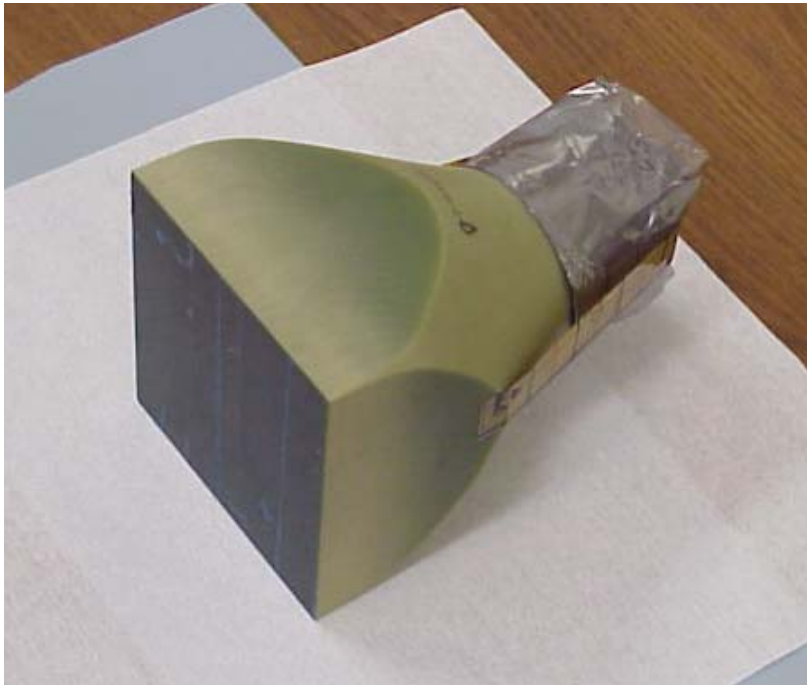
Plastic X-ray sensitive film
Photostimulated luminescence by $BaFBr:Eu^{2+}$

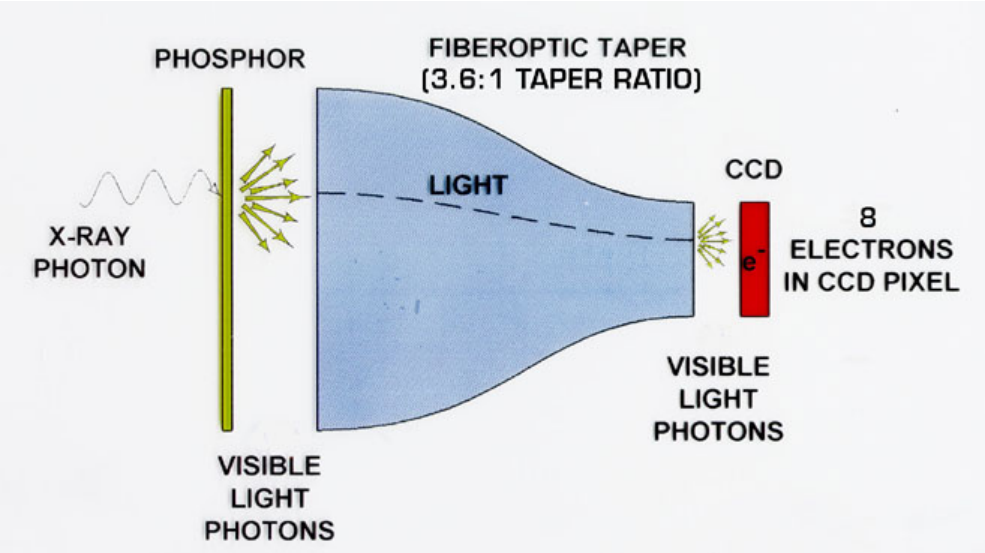


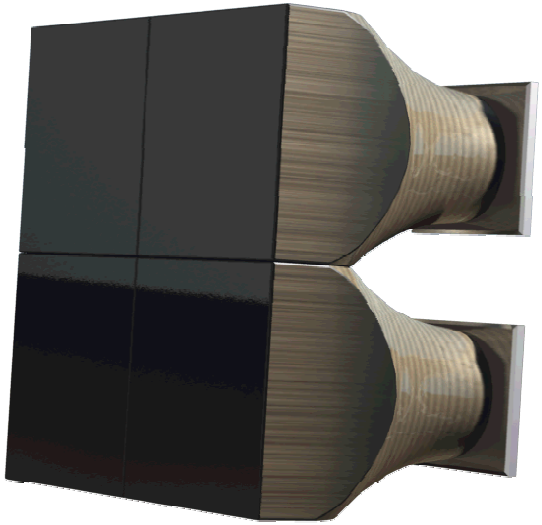


(Rigaku, Japan)

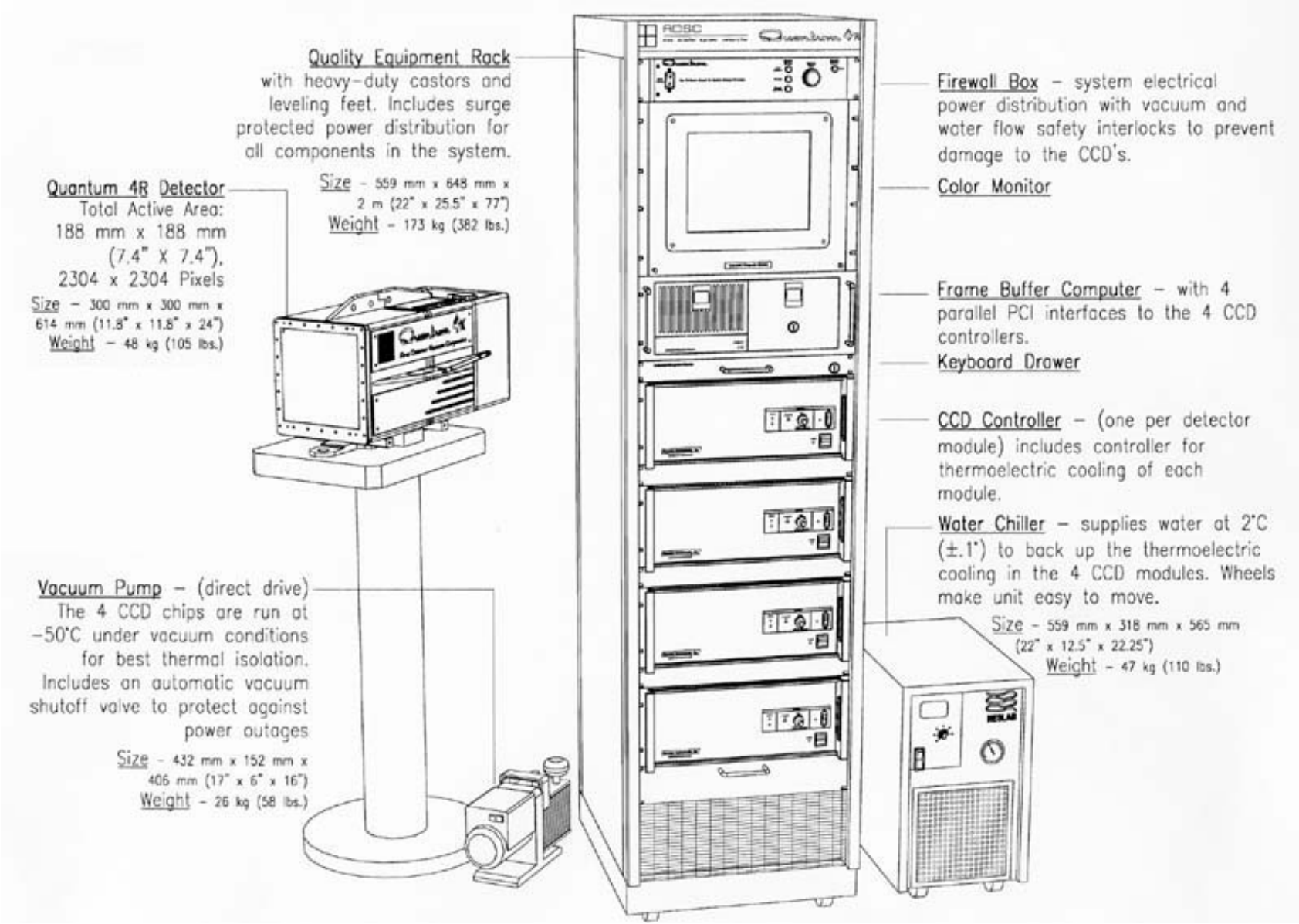
CCD Detector







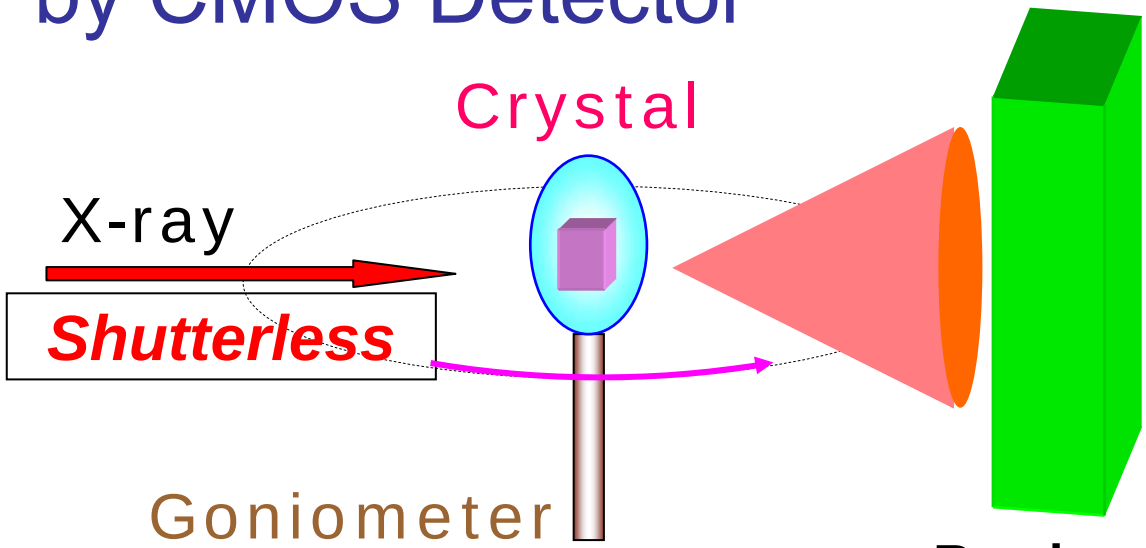
CCD detector system



Area Detector Systems Co. http://www.adsc-xray.com

Continuous rotation data collection by CMOS Detector

Hasegawa (JASRI) &
Yamamoto (RIKEN)
Hamamatsu Photonics



Rotate with a constant speed

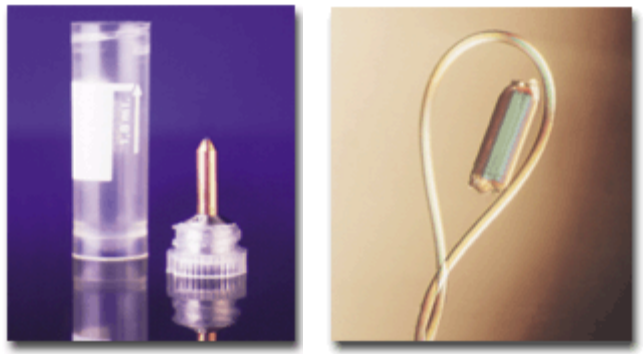
Read out images with a constant frame rate

High throughput and/or Fine slice data collection

Specification	Hamamatsu C10158DK	ADSC Q210
Scintillator	CsI:Tl	Gd_2O_2S:Tb
Pixel size [mm^2]	50 x 50	51 x 51
Detector area [mm^2]	118.8 x 118.8	210 x 210
Output data [bits]	14	16
Dynamic range	6,000	14,100
Dead time due to readout	14 msec / pixel	1.1 sec / frame

Cryocrystallographic technique

Prevent thermal degradation of sample
diffusion and reaction of free radicals
at cryogenic temperature (30 – 100 K)
using cold N_2/He gas stream



Sample Mount Pin & Cryoloop

Diffraction power of crystal

Darwin's Formula

$$E(\mathbf{h}) = \frac{I_0}{\omega}\lambda^3 \frac{e^4}{m^2c^4}\frac{P \cdot L \cdot A \cdot V_x}{V^2} \cdot \left|F(\mathbf{h})\right|^2 \cdots$$

I_0: Incident intensity, ω: Angular velocity of crystal rotation, λ: Wavelength,
e: Charge of electron, P: Polarization factor (= $(1+\cos^2 2\theta)/2$),
L: Lotentz factor (= $1/\sin\theta$ when spindle x-ray),
A: Absorption coefficient, V_x: Crystal volume, V: Lattice volume

In case of protein crystal…

- High solvent contents (25 ~ 75%)
- Large unit cell

> Weak diffraction power ~ Low resolution

Crystal packing ~ molecular vibration ~ resolution

Relationship with B-factor (DWF)

Vibration in solution > Movie

Packing density V_M:

$V_M = V_{cell} / Mw_{cell}$

High density (small V_M) > High Rigidity

(Kantardjieff & Rupp, 2003)

Packing control by humidity control

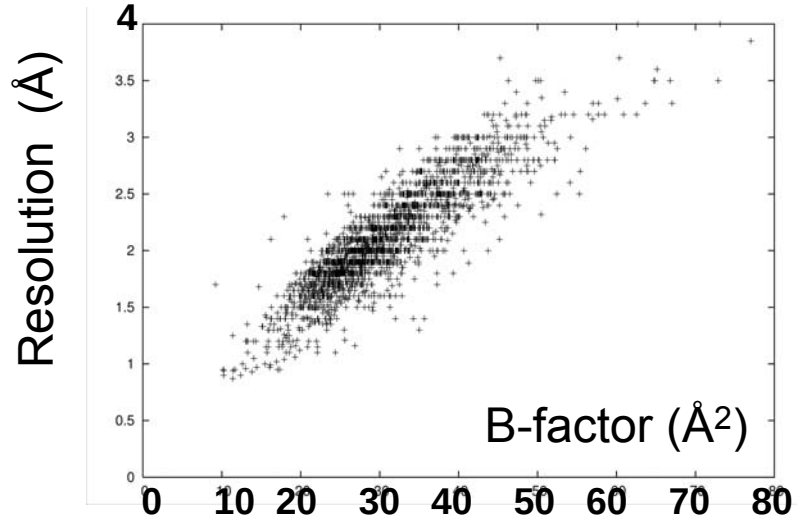
FMS (Free Mounting System)

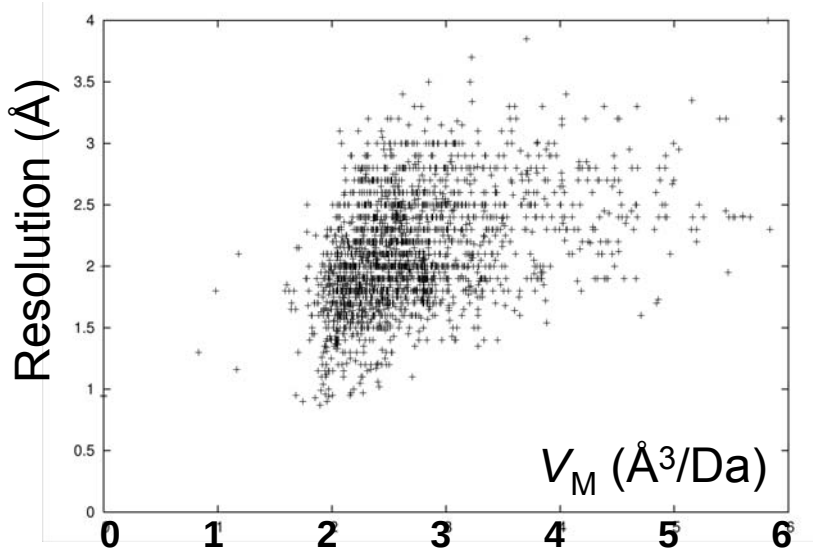
> lower humidity around crystal

> dehydration

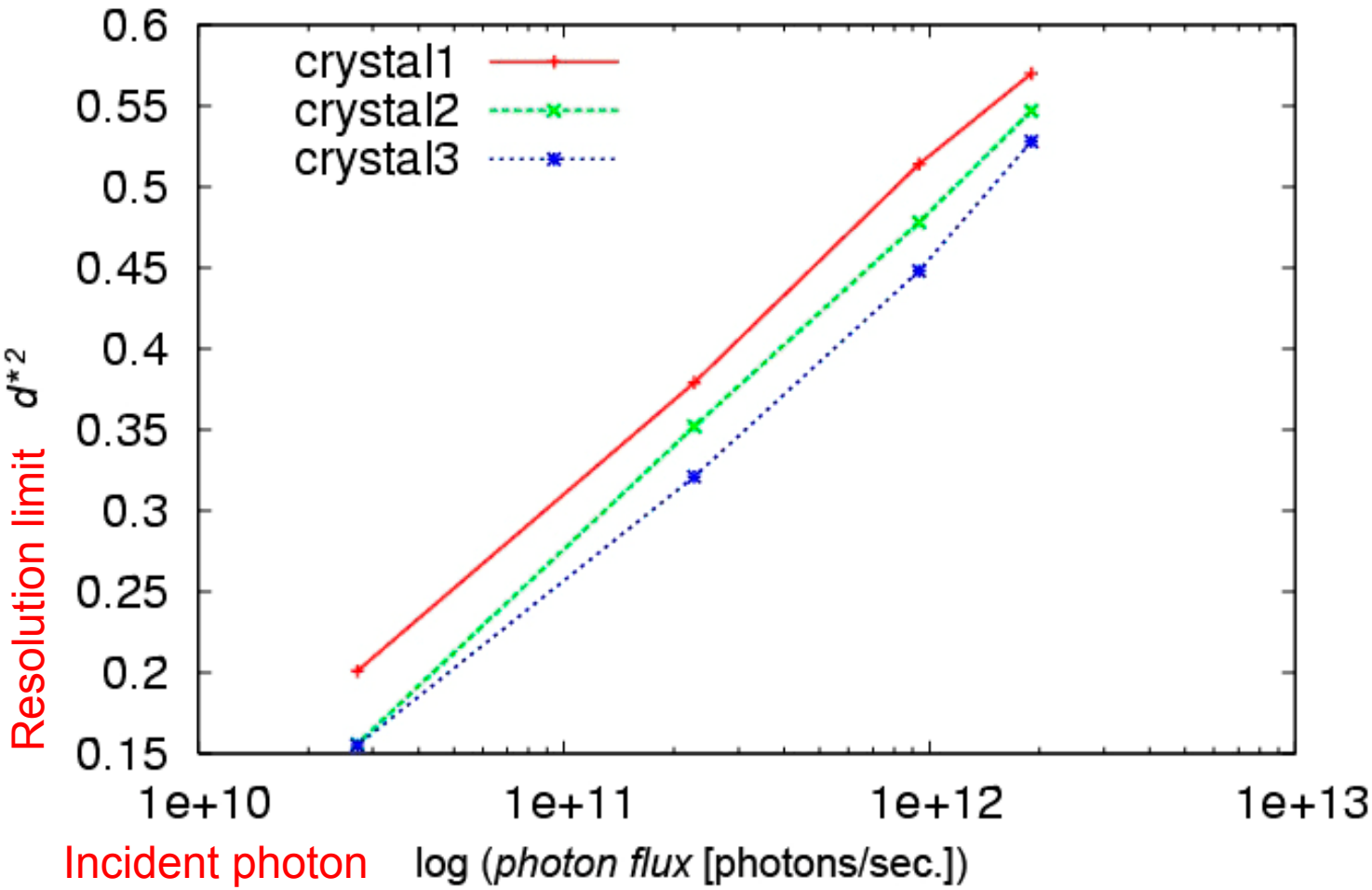
> induce phase transition

(Kiefersauer et al., 2000)

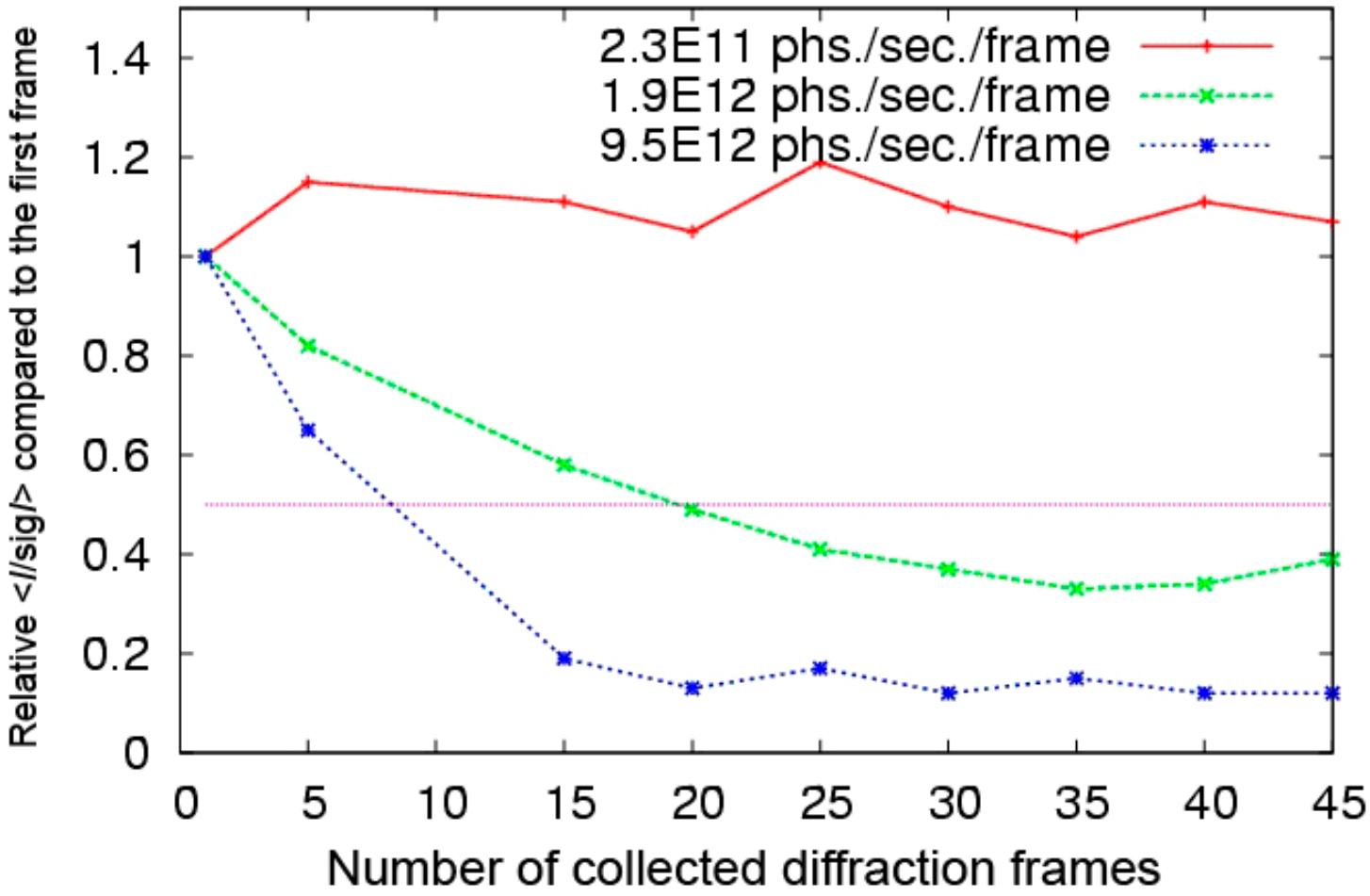




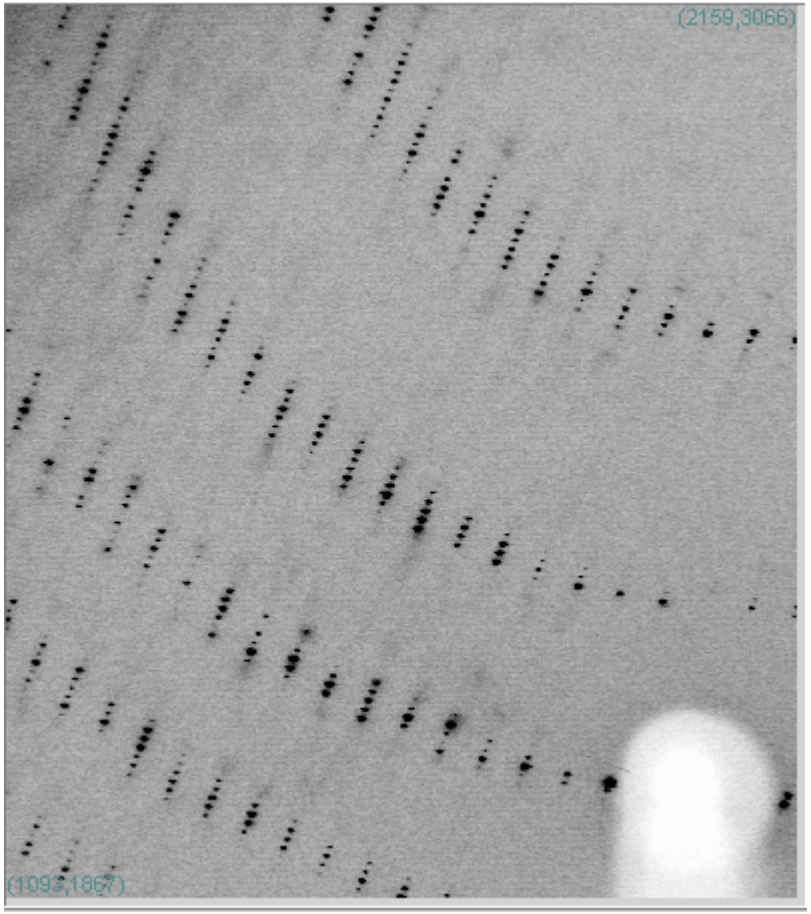
Resolution and incident intensity



Reduction by radiation damage

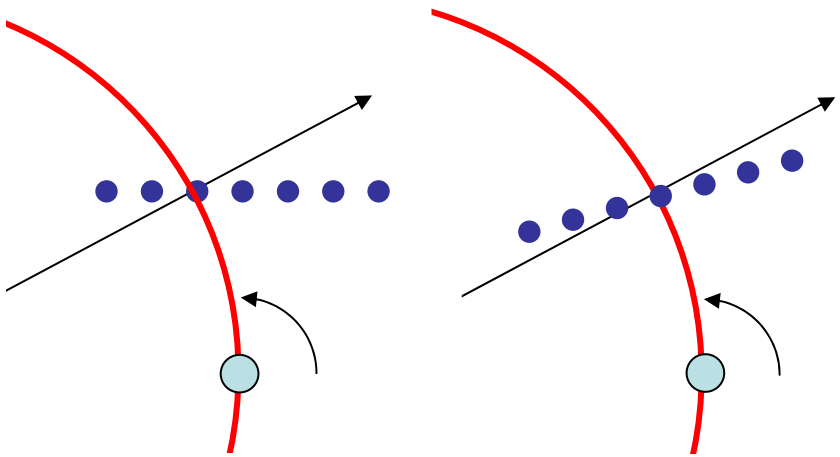


Reflection overlaps

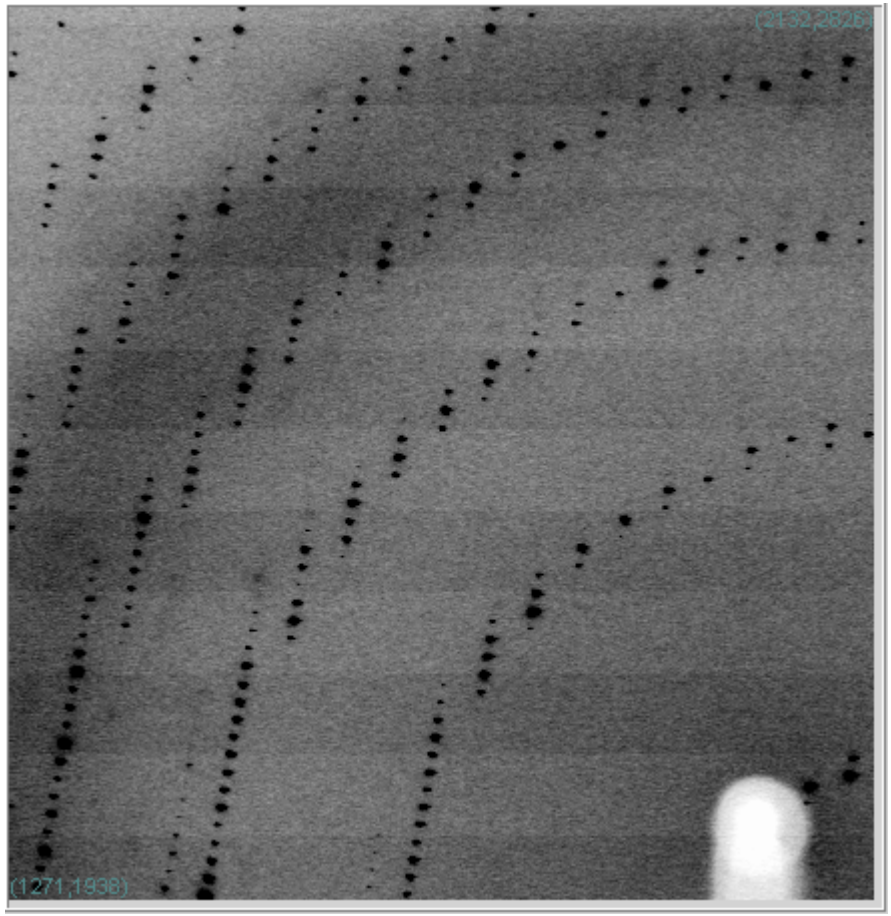


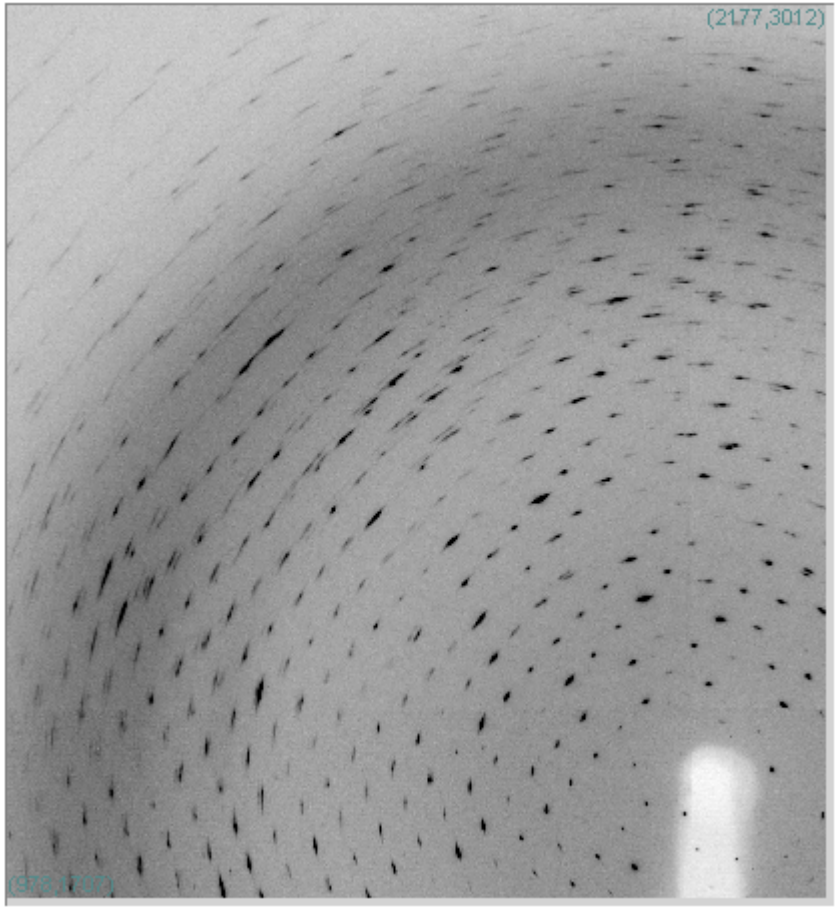
Longer lattice constant gives narrower spacing of adjacent reflections.

Long axis should be placed along rotation axis.



Mosaic spread

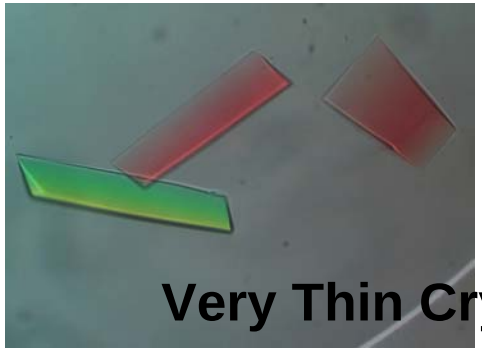




Spot sharpness depends on crystalline order.

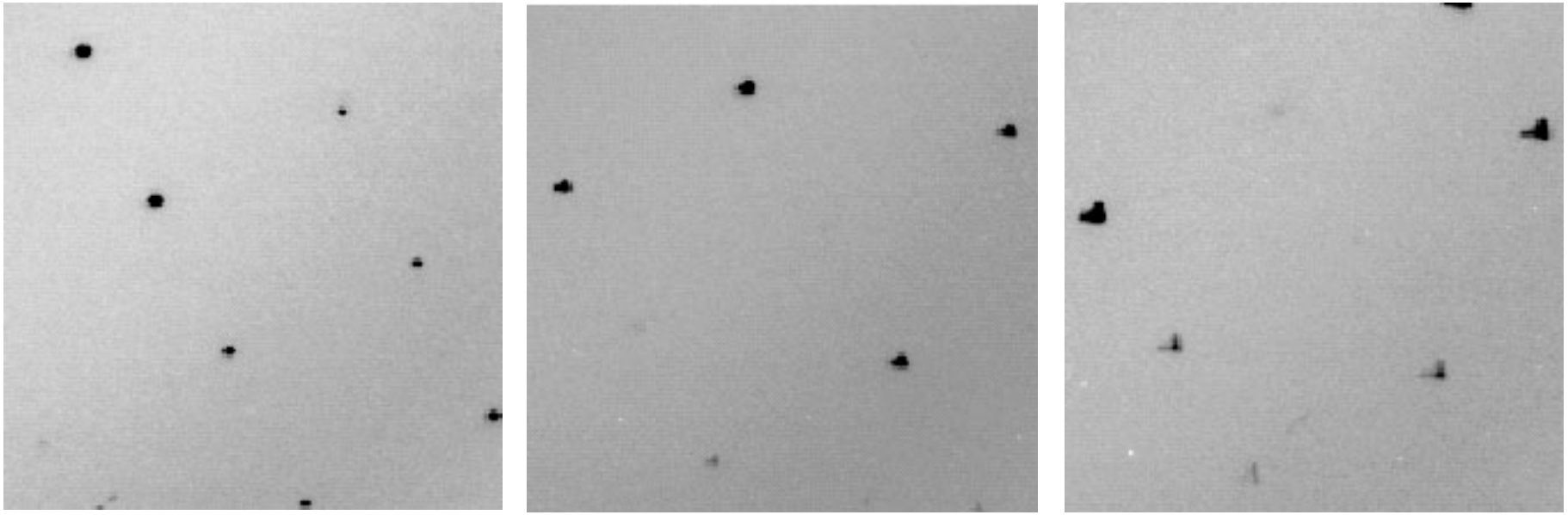
Radiation damage

Bacterial flagelin F41 Crystal
@ SPring-8 BL41XU



Very Thin Crystal (≒10μm)

1st frame ***15min.*** ***25min.***



Total Flux at Sample **≒ 10^{13} photons/sec/mm^2**

F.A. Samatey, K.Imada, S.Nagashima, K.Namba (ERATO)

Interaction between photon and protein

Primary Effect

・Absorbed photon energy > **Temperature increment**

・Photoelectron formation

> **Chemical reduction / Reactive radical formation**

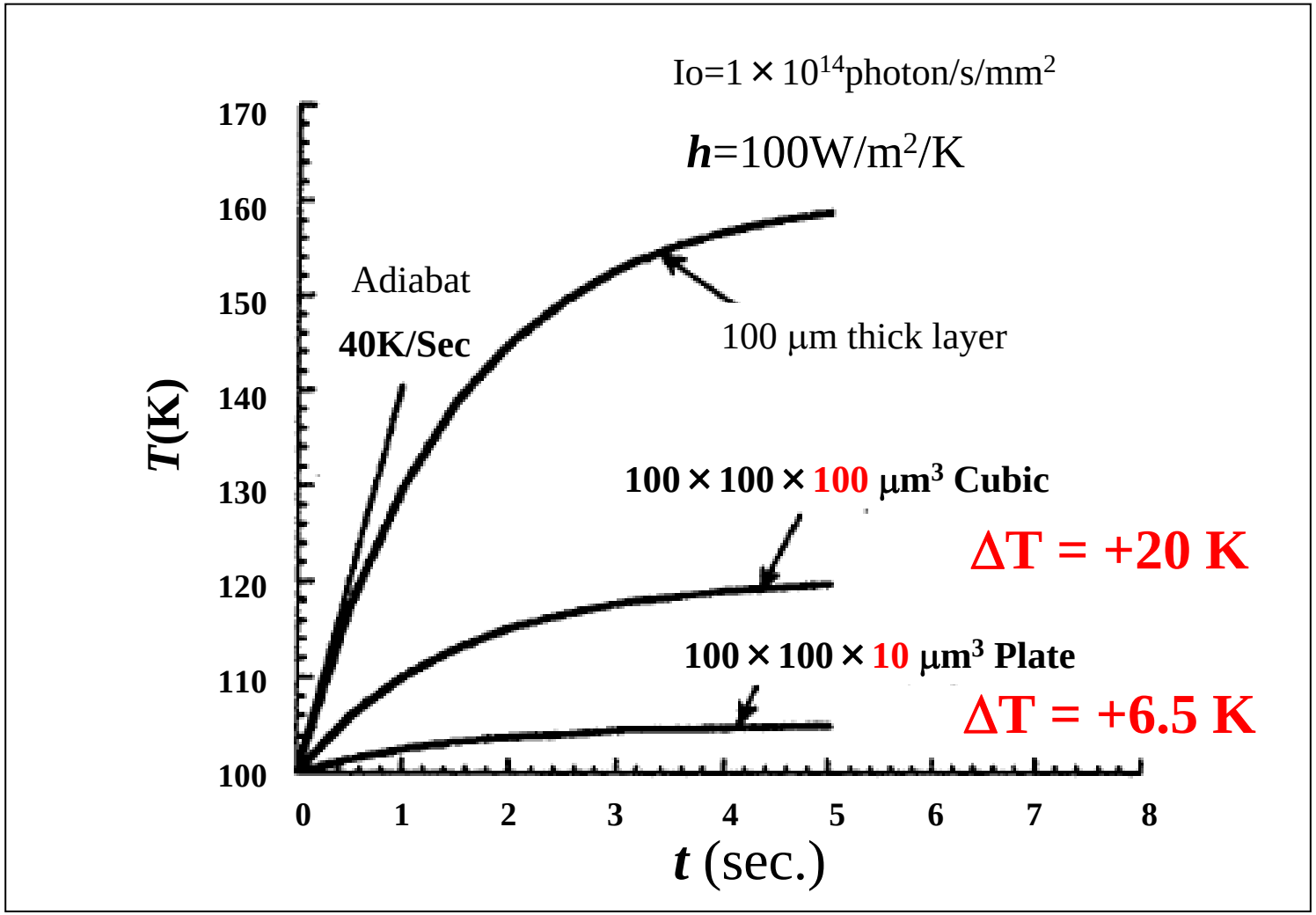
X-ray dose	**dependent**
Temperature / Time	independent

Secondary Effect

・Chemical reaction by free radicals

X-ray dose, temperature / time dependent

Radiation induced temperature increment under cryogenic condition



Kuzay *et al.* Acta Cryst. (2001). D57, 69-81

Radiation induced formation of reactive radicals (1)

Water

$$H_2O \xrightarrow{h\nu} H_2O^+ + e^-$$

$$H_2O^+ + H_2O \longrightarrow H^+ + OH^\bullet$$

$$e^- + H_2O \longrightarrow H_2O^-$$

$$H_2O^- + H_2O \longrightarrow H^\bullet + OH^-$$

$$H_2O \xrightarrow{h\nu} \underline{H}^\bullet + \underline{OH}^\bullet$$

Radiation induced formation of reactive radicals (2)

Disulfide bridge

$$(p)\text{-}CH_2SSCH_2\text{-}(p) \xrightarrow{h\nu} (p)\text{-}CH_2SS^{\bullet +}\ CH_2\text{-}(p) + e^-$$

$$(p)\text{-}CH_2SSCH_2\text{-}(p) + \underline{OH}^{\bullet} \longrightarrow (p)\text{-}CH_2SS^{\bullet +}\ CH_2\text{-}(p) + OH^-$$

$$(p)\text{-}CH_2SS^{\bullet +}\ CH_2\text{-}(p) + OH^- \longrightarrow (p)\text{-}CH_2SOH + (p)\text{-}CH_2S^{\bullet}$$

$$(p)\text{-}CH_2SSCH_2\text{-}(p) + \underline{OH}^{\bullet} \longrightarrow (p)\text{-}CH_2SOH + (p)\text{-}CH_2S^{\bullet}$$

Cysteine

$$(p)\text{-}CH_2SH + e^- \longrightarrow (p)\text{-}CH_2^{\bullet} + SH^-$$

$$(p)\text{-}CH_2SH + \underline{H}^{\bullet} \longrightarrow (p)\text{-}CH_2^{\bullet} + H_2S$$

Burmeister, Acta Cryst. (2000). D56, 328-341

Radiation induced formation of reactive radicals (3)

Aspartate & Glutamate

$$(p)\text{-}CH_2CH_2COO^- \xrightarrow{h\nu} (p)\text{-}CH_2C^{\bullet +}\ H_2COO^- + e^-$$

$$(p)\text{-}CH_2C^{\bullet +}\ H_2COO^- \longrightarrow (p)\text{-}CH_2CH_2^{\bullet} + CO_2$$

Tyrosine

$$(p)\text{-}CH_2C_6H_4OH \xrightarrow{h\nu} (p)\text{-}CH_2C_6H_4OH^{\bullet +} + e^-$$

$$(p)\text{-}CH_2C_6H_4OH^{\bullet +} \longrightarrow (p)\text{-}CH_2C_6H_4O^{\bullet} + H^+$$

Methionine

$$(p)\text{-}CH_2CH_2SCH_3 + 2\underline{H}^{\bullet} \longrightarrow (p)\text{-}CH_2\ CH_3 + CH_3SH$$

Burmeister, Acta Cryst. (2000). D56, 328-341

Dose limit

Estimated dose limit for ionizing radiation

1.3×10^{17} keV/mm^{-3}

1×10^{16} photon/mm^{2}@12.4keV

（Henderson, R. (1990). *Proc. R. Soc. London Ser. B*, **241**, 6-8.）

4×10^{17} keV/mm^{-3}

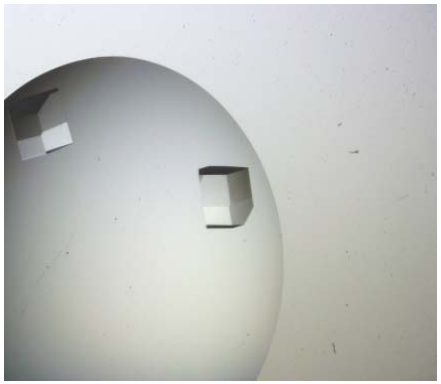
（Gonzalez, A., Nave, C. (1994). *Acta Cryst. D* **50**, 874-877.）

5×10^{16} photon/mm^{2}@12.4keV

（Sliz, P.,Harrison,S.C.Rosenbaum,G., (2002). *Structure*, **11**, 13-19.）

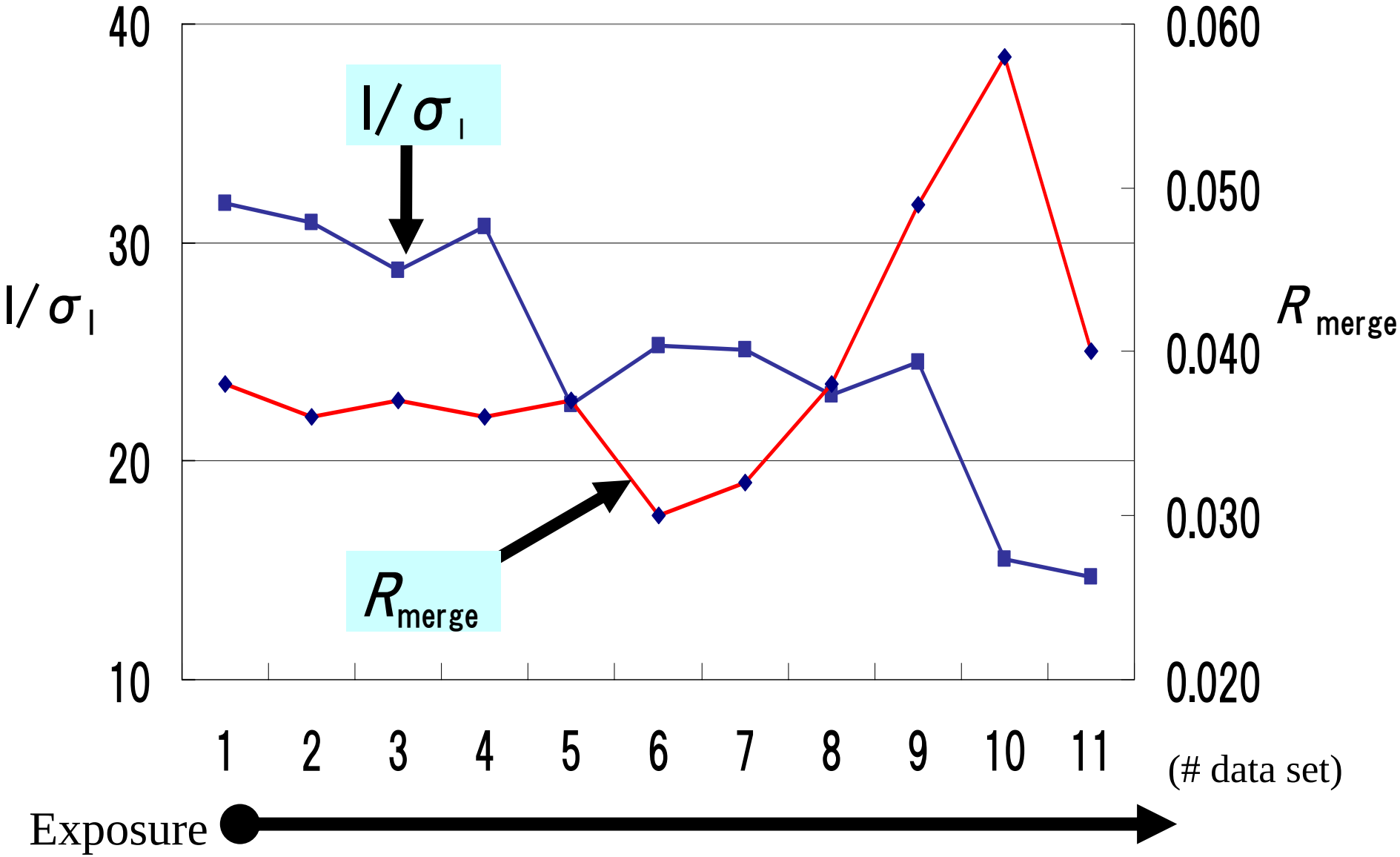
Estimation of radiation damage

Sample	Hen egg white lysozyme
Space group	$P4_32_12$
Lattice	$a=b=78.54$,$c=37.77$Å
	α＝β＝γ＝90°

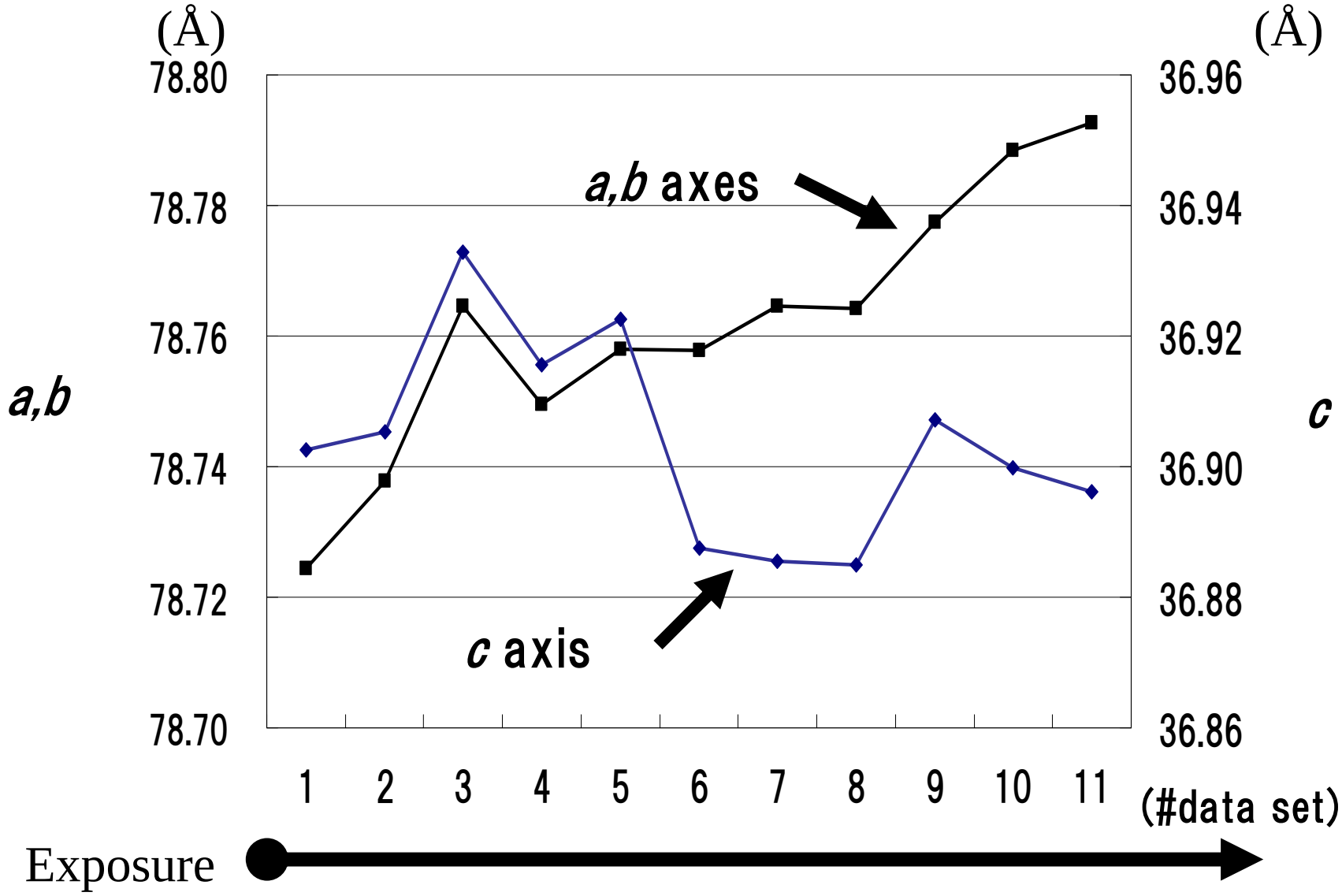


BL	SR (SPring-8-BL45XU)	
Data set	11 (1～11)	1(12)
image/set	95	245
Wavelength	1.02Å	1.02Å
Oscillation angle	1.0°	1.0°
Camera distance	150mm	220mm
Exposure	5秒	60秒
Detector	Jupiter210 (CCD)	RAXIS-V (IP)

Variation of I/σ_I and R_{merge}



Variation of lattice constants



Radiation damage in real space

electron density at 1.6 Å resolution

Data set

1

R-factor 18.51 %

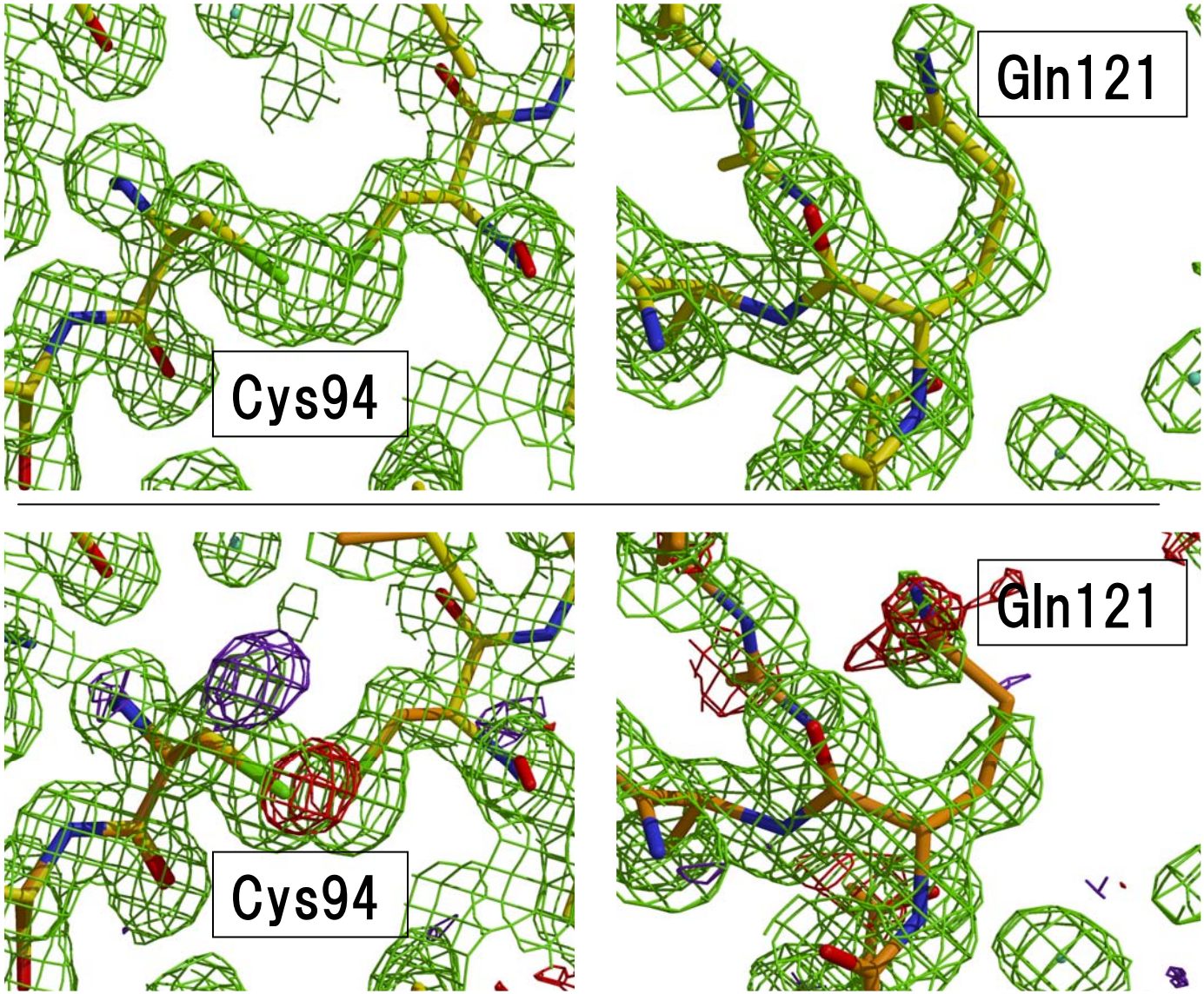
R_{free} 20.64 %

fo-fc	2.5 σ
fo-fc	−2.5 σ
2fo-fc	1.0 σ

12

R-factor 18.50 %

R_{free} 20.68 %



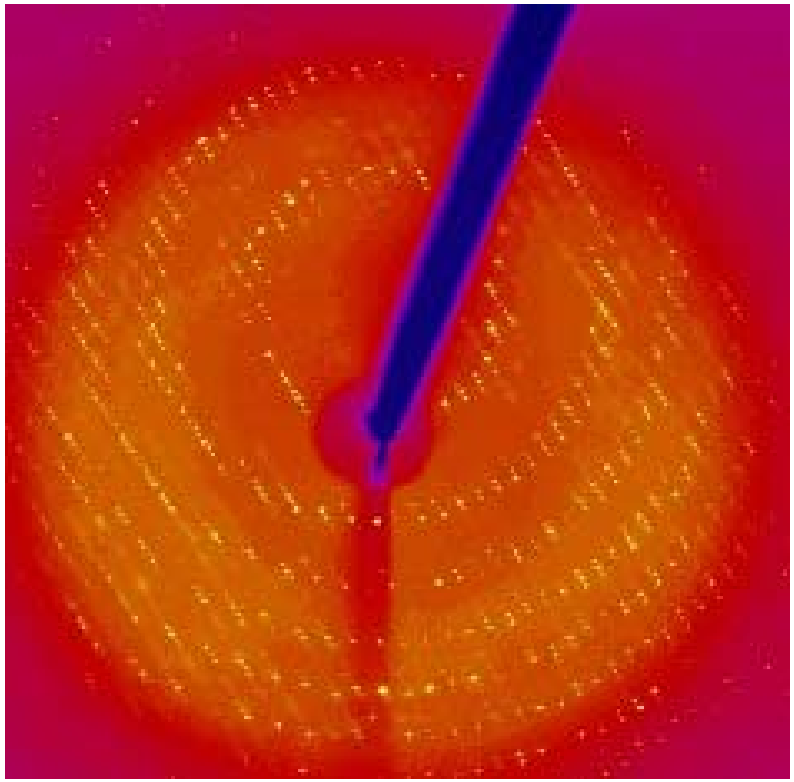
2.2: Phasing

Crystallographic phase problem

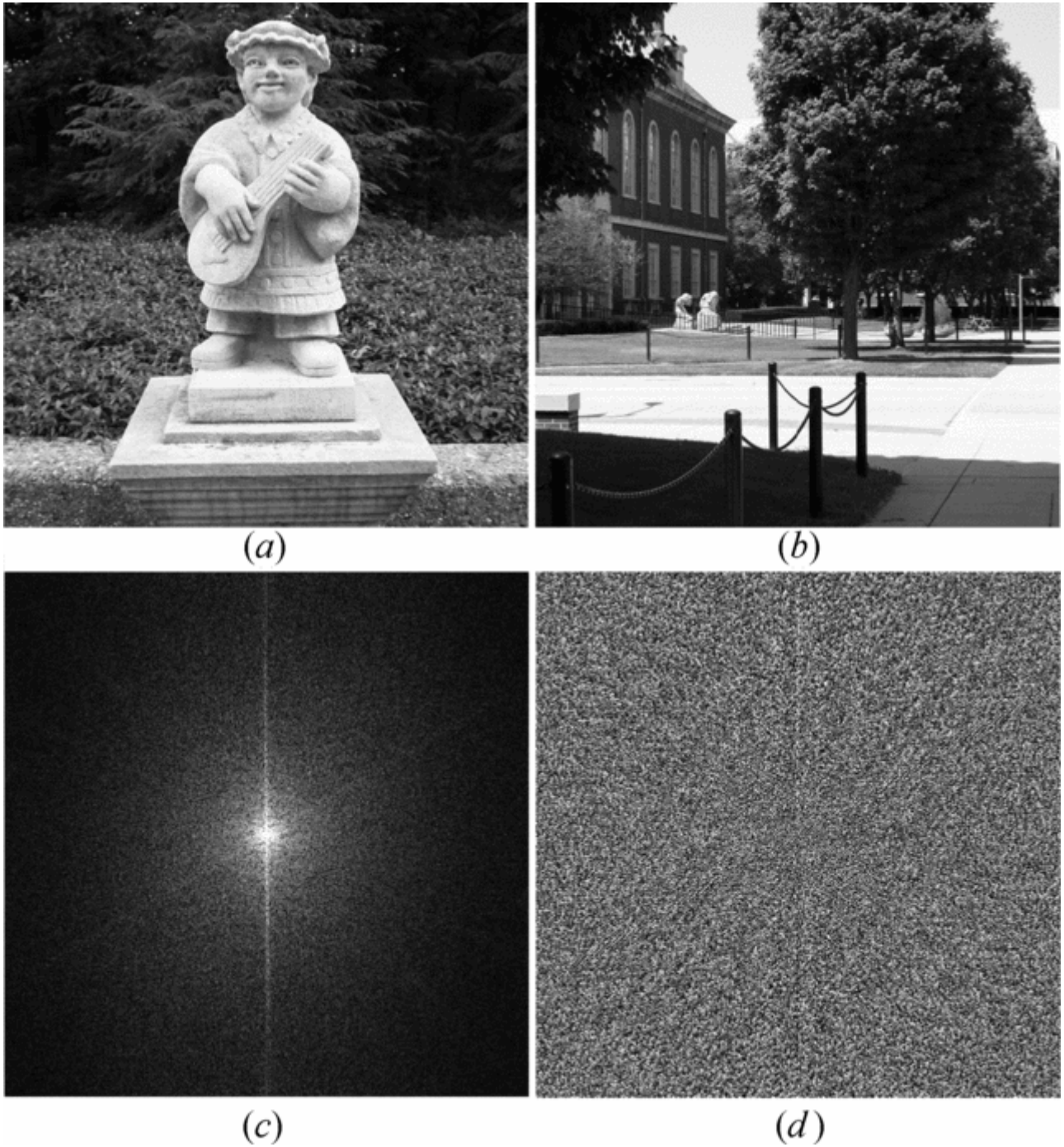
Diffraction intensity is only measureable,
but its phase information is completely lost.

$$I(hkl) = \mathbf{F}(hkl)\,\mathbf{F}^*(hkl)$$

$$\mathbf{F}(hkl) = |\,\mathbf{F}(hkl)\,|\,\exp i\alpha$$



Solving methods
1. Direct method
2. Isomorphous replacement (IR)
3. Molecular replacement (MR)



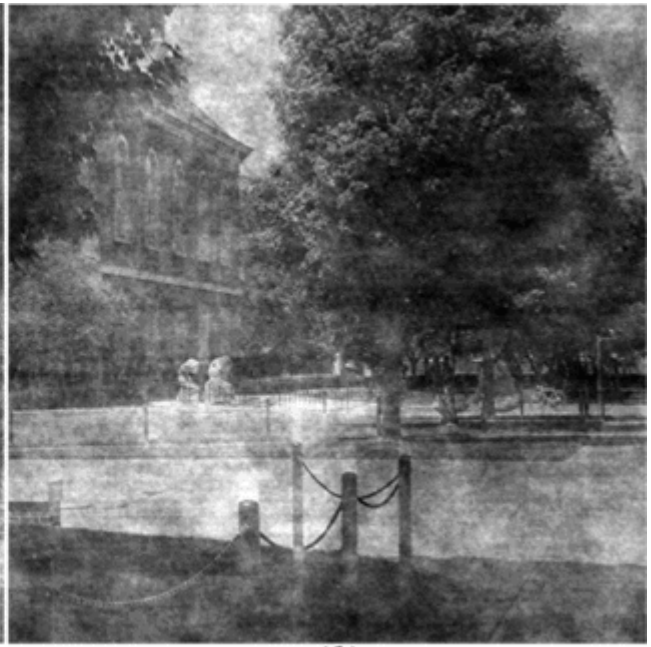
|F| of (a)

α of (a)

Albert & Lecomte, J. Appl. Cryst. 40, 1153 (2007)

|F| of (b)
α of (a)

(*a*)

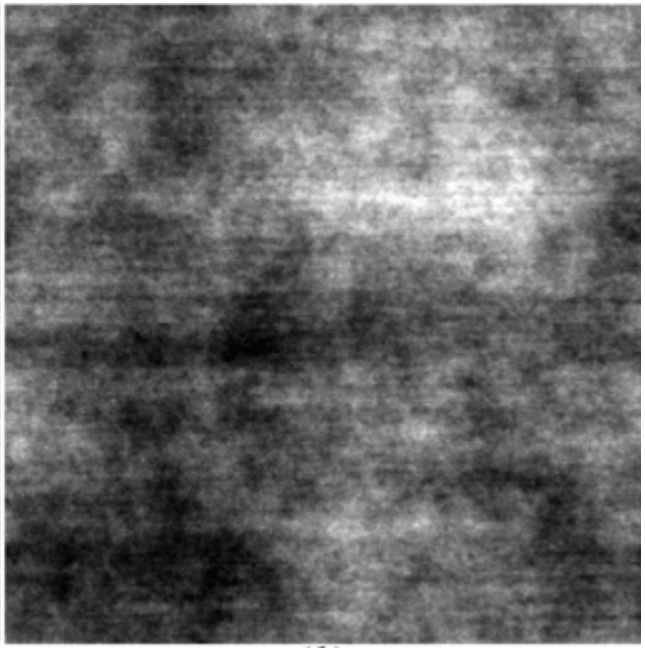


(*b*)

|F| of (a)
α of (b)

random |F|
α of (a)

(*a*)



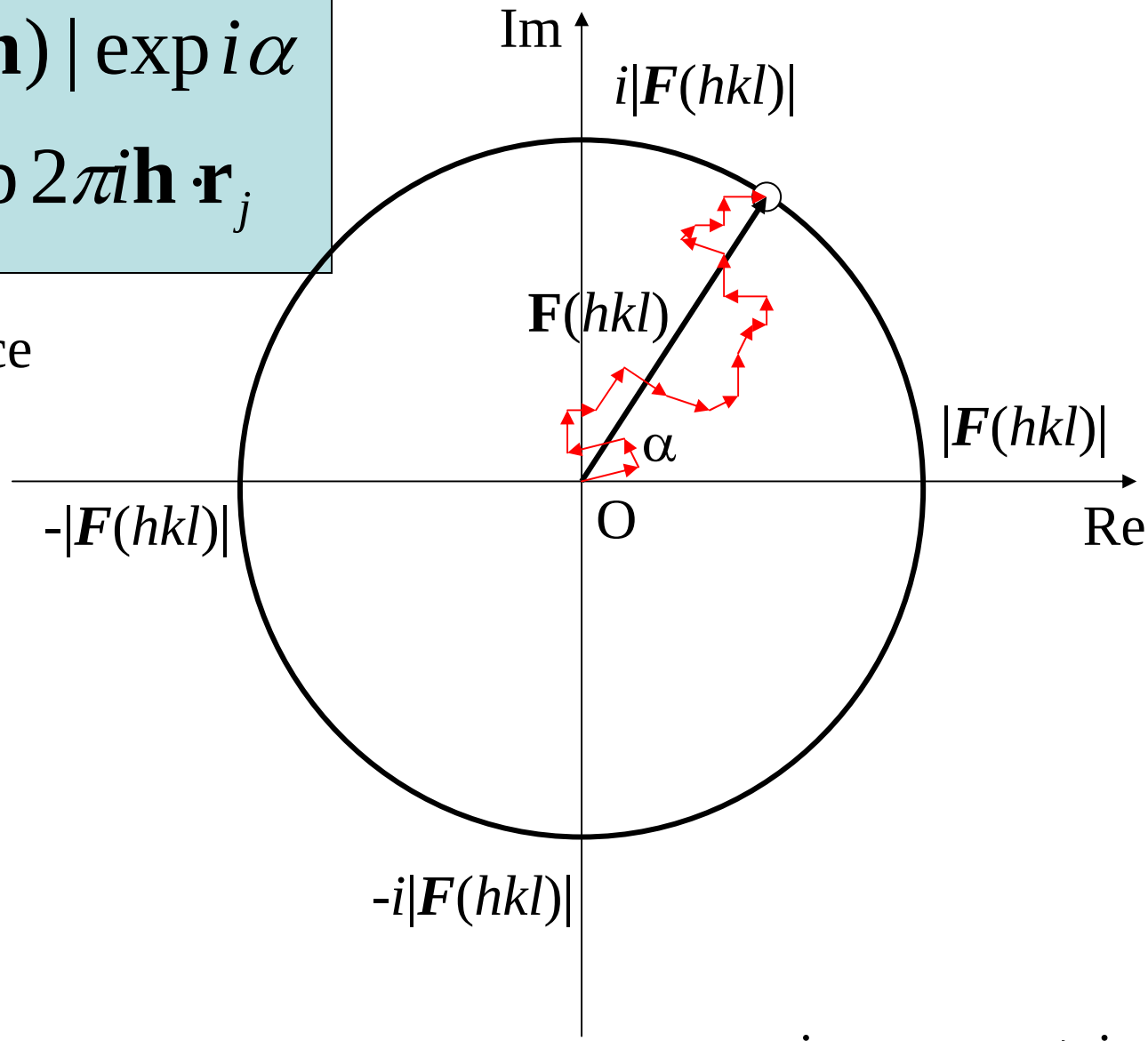
(*b*)

random |F
α of (b)

Harker Diagram ~ Structure factor and phase

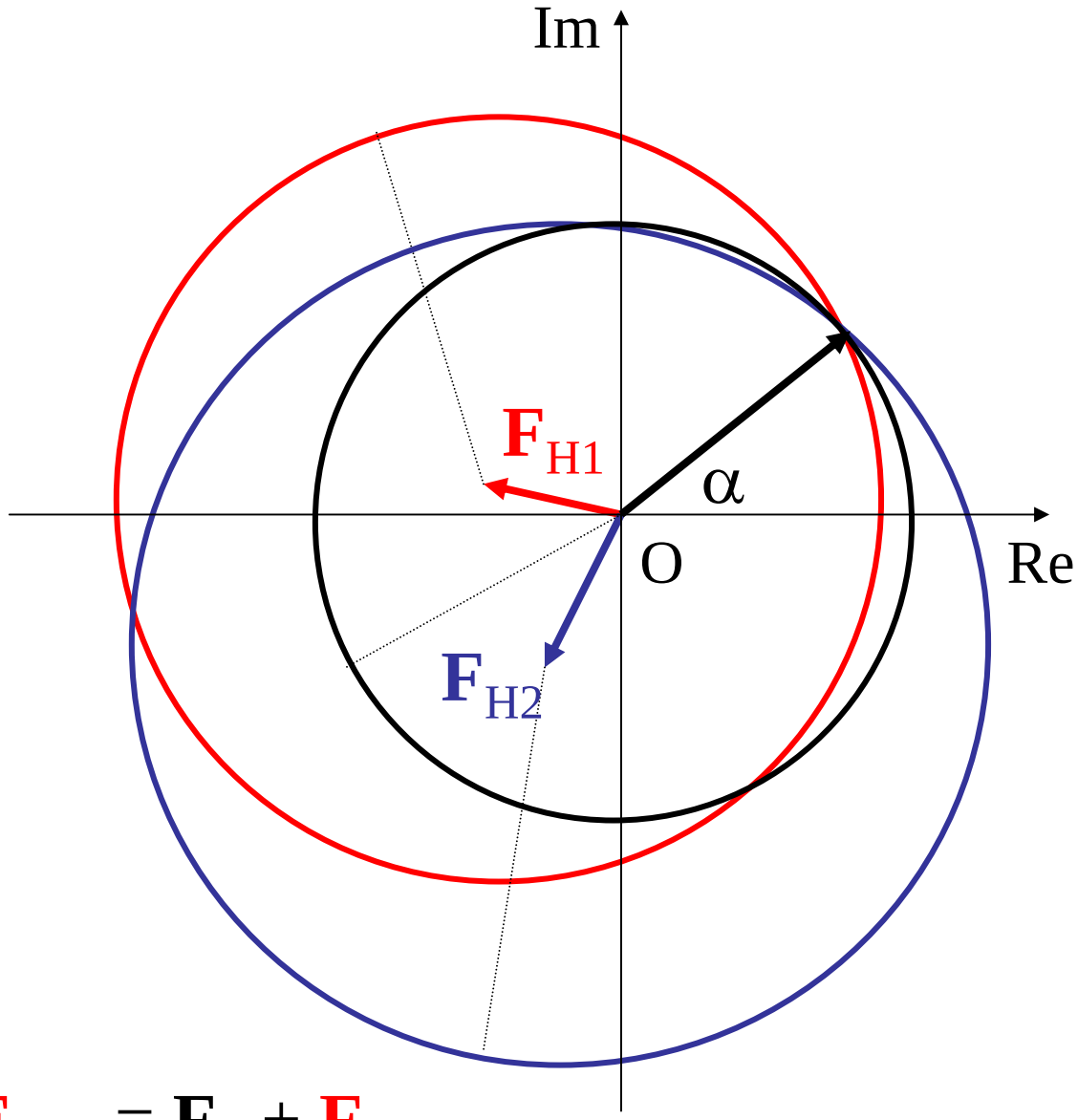
$$\mathbf{F}(\mathbf{h}) = |\mathbf{F}(\mathbf{h})| \exp i\alpha$$
$$= \sum f_j \exp 2\pi i \mathbf{h} \cdot \mathbf{r}_j$$

Complex space



$\exp i\alpha = \cos\alpha + i \sin\alpha$

Isomorphous replacement

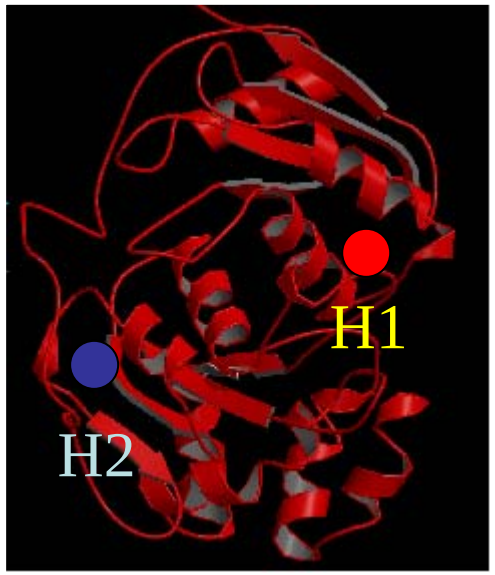


$|\mathbf{F}_P|$: Native

$|\mathbf{F}_{PH1}|$: Derivative 1

$|\mathbf{F}_{PH2}|$: Derivative 2

$\mathbf{F}_P = |F_P| \exp i\alpha$



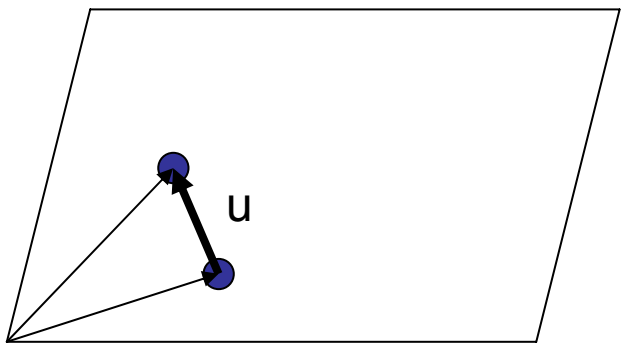
$\mathbf{F}_{PH1} = \mathbf{F}_P + \mathbf{F}_{H1}$

Patterson function

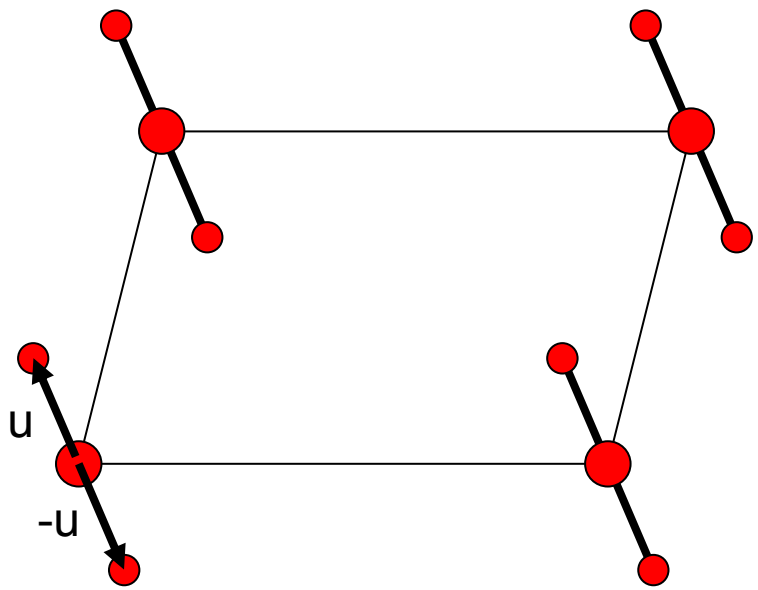
Directly calculated from intensity without phase.
The function shows self correlation of electron density.

$$P(\mathbf{u}) = \int_{cell} \rho(\mathbf{r})\rho(\mathbf{r}+\mathbf{u})\mathrm{d}^3\mathbf{r} = \frac{1}{V}\sum_{\mathbf{h}} |F(\mathbf{h})|^2 \exp(-2\pi i \mathbf{h}\cdot\mathbf{u}).$$

Interatomic vector $\mathbf{u} = (u\ v\ w)$



Real space



Patterson space

In case of few atoms in cell, their coordinates are determined from Patterson function.

Characteristics of Patterson function

1. Even function: $P(\mathbf{u}) = P(-\mathbf{u})$

2. Screw axis in real space > Rotation axis

3. Harker line / Harker section

$P2_1$: (x, y, z), $(-x, y+1/2, -z)$

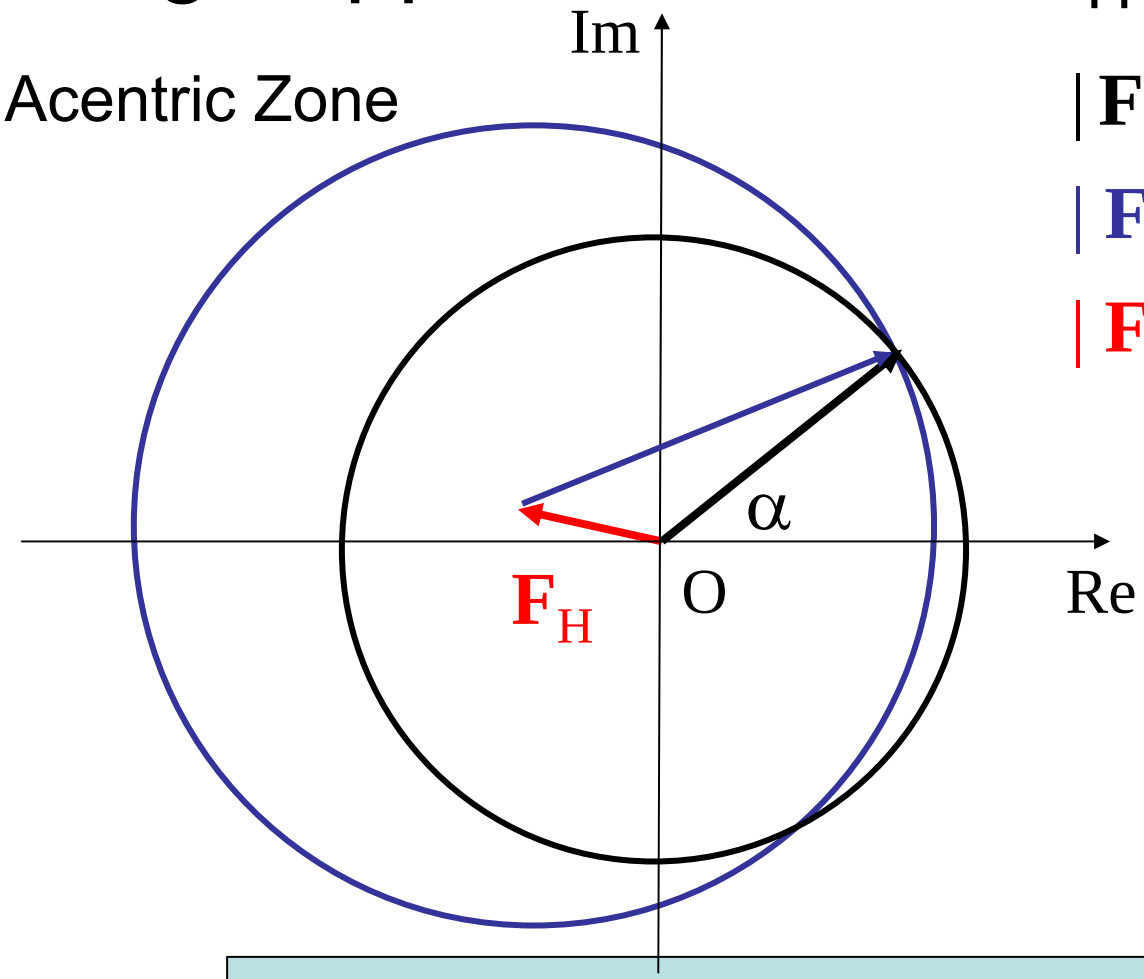
$(u, v, w) = (-2x, 1/2, -2z)$

4. Correspond to mathematical convolution

$$f(t) * g(t) = \int_0^t f(t-\tau) g(\tau) d\tau$$

$f(t)*f(t)$: Self correlation
$f(t)*g(t)$: Cross corr.

Rough approximation of F_H from data



$|\mathbf{F}_P|$, α_P: Native

$|\mathbf{F}_{PH}|$, α_{PH} : Derivative

$|\mathbf{F}_H|$, α_H : Heavy atom

$$\mathbf{F}_P = |F_P| \exp i\alpha$$

$$\Delta F = \left|F_{PH} - F_P\right| = \left|F_H \cos(\alpha_{PH} - \alpha_P) - 2F_P \sin^2\left\{\frac{\alpha_P - \alpha_{PH}}{2}\right\}\right|$$

$$\sim \left|F_H \cos(\alpha_{PH} - \alpha_P)\right|$$

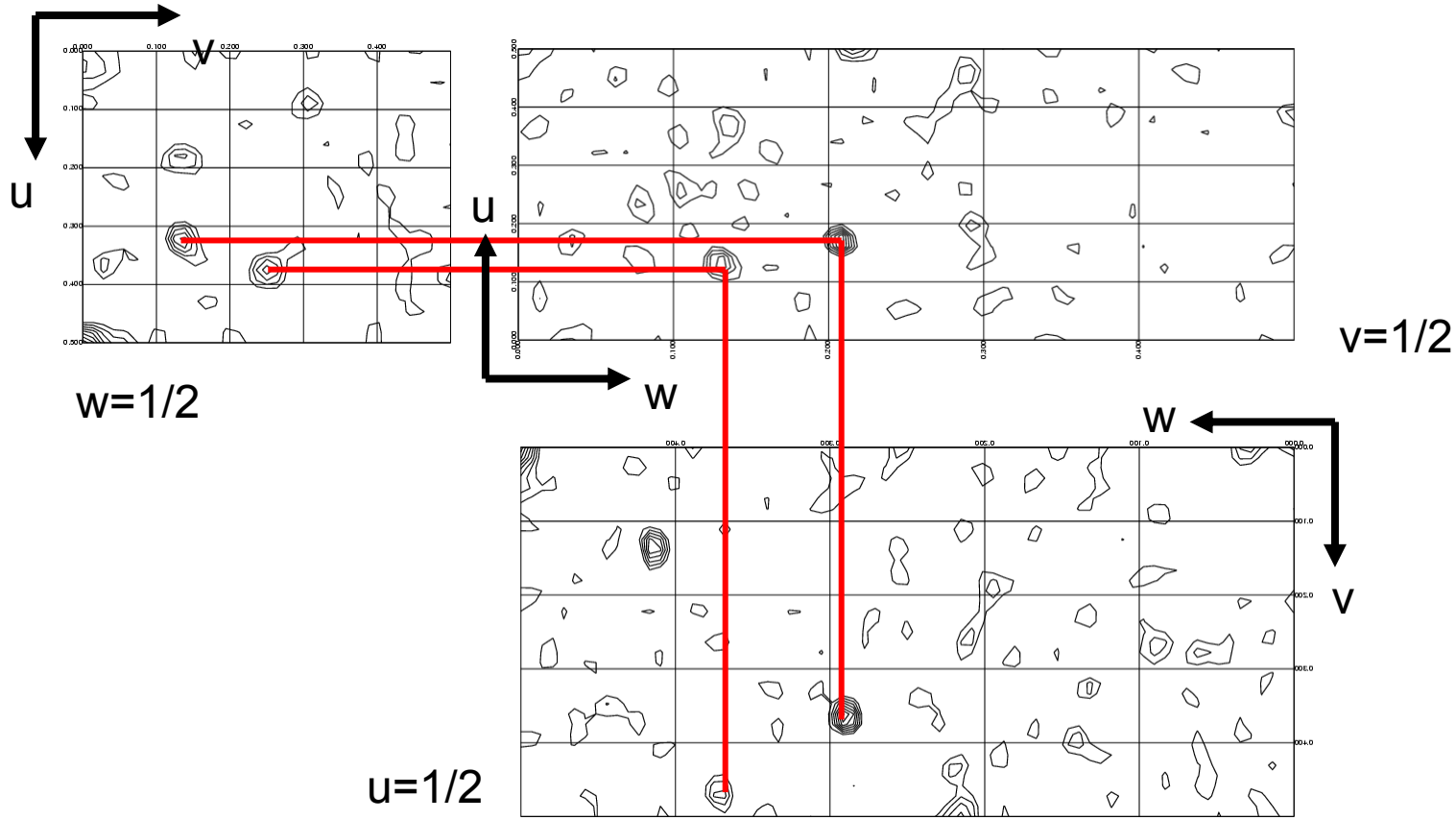
Harker Section

$P2_12_12_1$ 1: x, y, z, 2: $\frac{1}{2}-x, -y, \frac{1}{2}+z$,
3: $\frac{1}{2}+x, \frac{1}{2}-y, -z$, 4: $-x, \frac{1}{2}+y, \frac{1}{2}-z$

Patterson Peaks $p(u,v,w)$

2-1: $\frac{1}{2}-2x, -2y, \frac{1}{2}$	2-1: $\frac{1}{2}-2x, 2y, \frac{1}{2}$
3-1: $\frac{1}{2}, \frac{1}{2}-2y, -2z$	3-1: $\frac{1}{2}, \frac{1}{2}-2y, 2z$
4-1: $-2x, \frac{1}{2}, \frac{1}{2}-2z$	4-1: $2x, \frac{1}{2}, \frac{1}{2}-2z$
3-2: $2x, \frac{1}{2}, -2z-\frac{1}{2}$	3-2: $2x, \frac{1}{2}, \frac{1}{2}+2z$
2-4: $\frac{1}{2}, -\frac{1}{2}-2y, 2z$	2-4: $\frac{1}{2}, \frac{1}{2}+2y, 2z$
4-3: $-\frac{1}{2}-2x, -2y, \frac{1}{2}$	4-3: $\frac{1}{2}+2x, 2y, \frac{1}{2}$

Relationship among Harker peaks

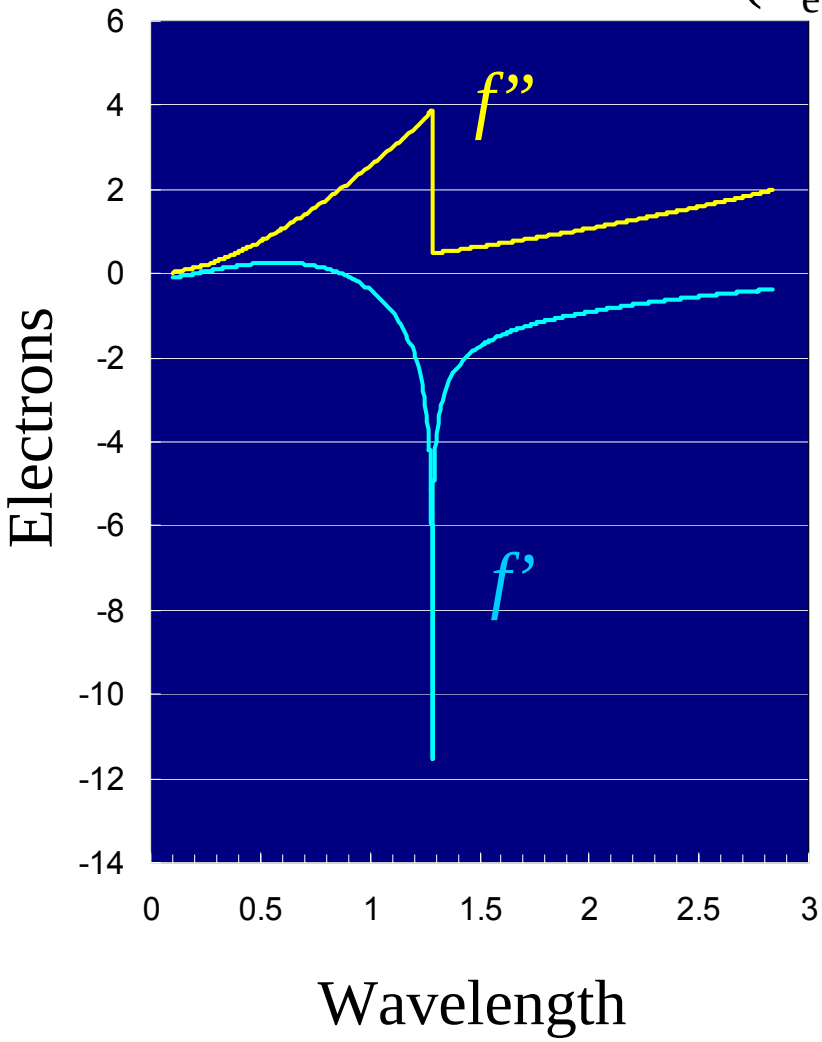


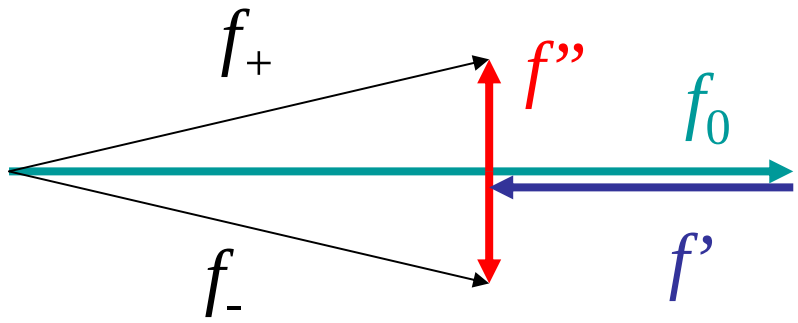
1: (x, y, z)	Patterson space (u, v, w)
2: $(\frac{1}{2}-x, -y, \frac{1}{2}+z)$	3-1 $(\frac{1}{2}, \frac{1}{2}-2y, -2z)$, 2-4 $(\frac{1}{2}, -\frac{1}{2}-2y, 2z)$
3: $(\frac{1}{2}+x, \frac{1}{2}-y, -z)$	4-1 $(-2x, \frac{1}{2}, \frac{1}{2}-2z)$, 3-2 $(2x, \frac{1}{2}, -\frac{1}{2}-2z)$
4: $(-x, \frac{1}{2}+y, \frac{1}{2}-z)$	2-1 $(\frac{1}{2}-2x, -2y, \frac{1}{2})$, 4-3 $(-\frac{1}{2}-2x, 2y, \frac{1}{2})$

Anomalous Phasing

Anomalous Effect

Anomalous term of zinc (Z_e = 32)

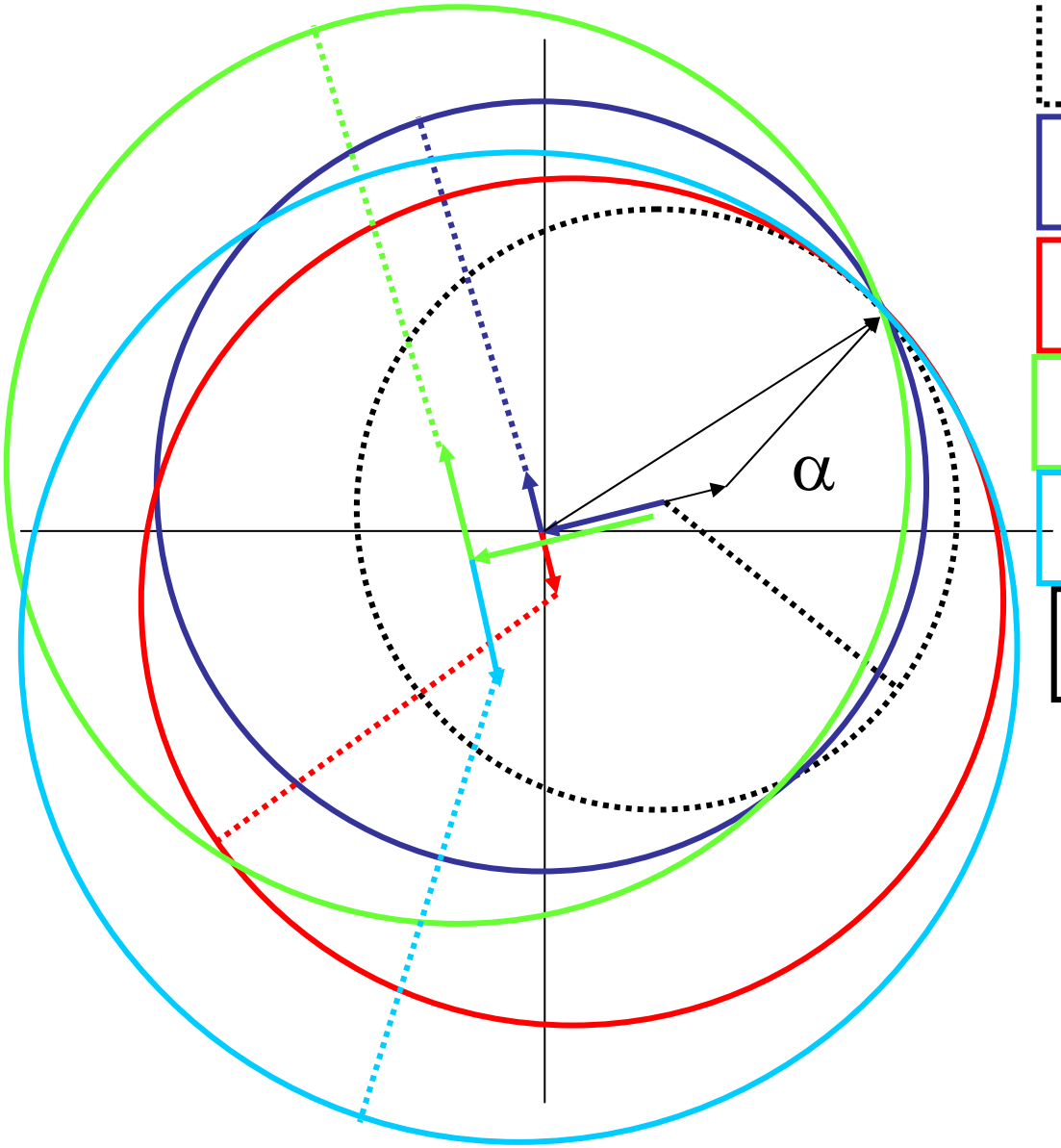




$$f_{\pm} = f_0 + f' \pm i f''$$

Smaller than usual
heavy atom effects
↓
Need high quality data

2 Wavelength MAD



$|F|$: Native

$|F_{\lambda 1}(\mathbf{h+})|$: Anomalous

$|F_{\lambda 1}(\mathbf{h-})|$: Anomalous

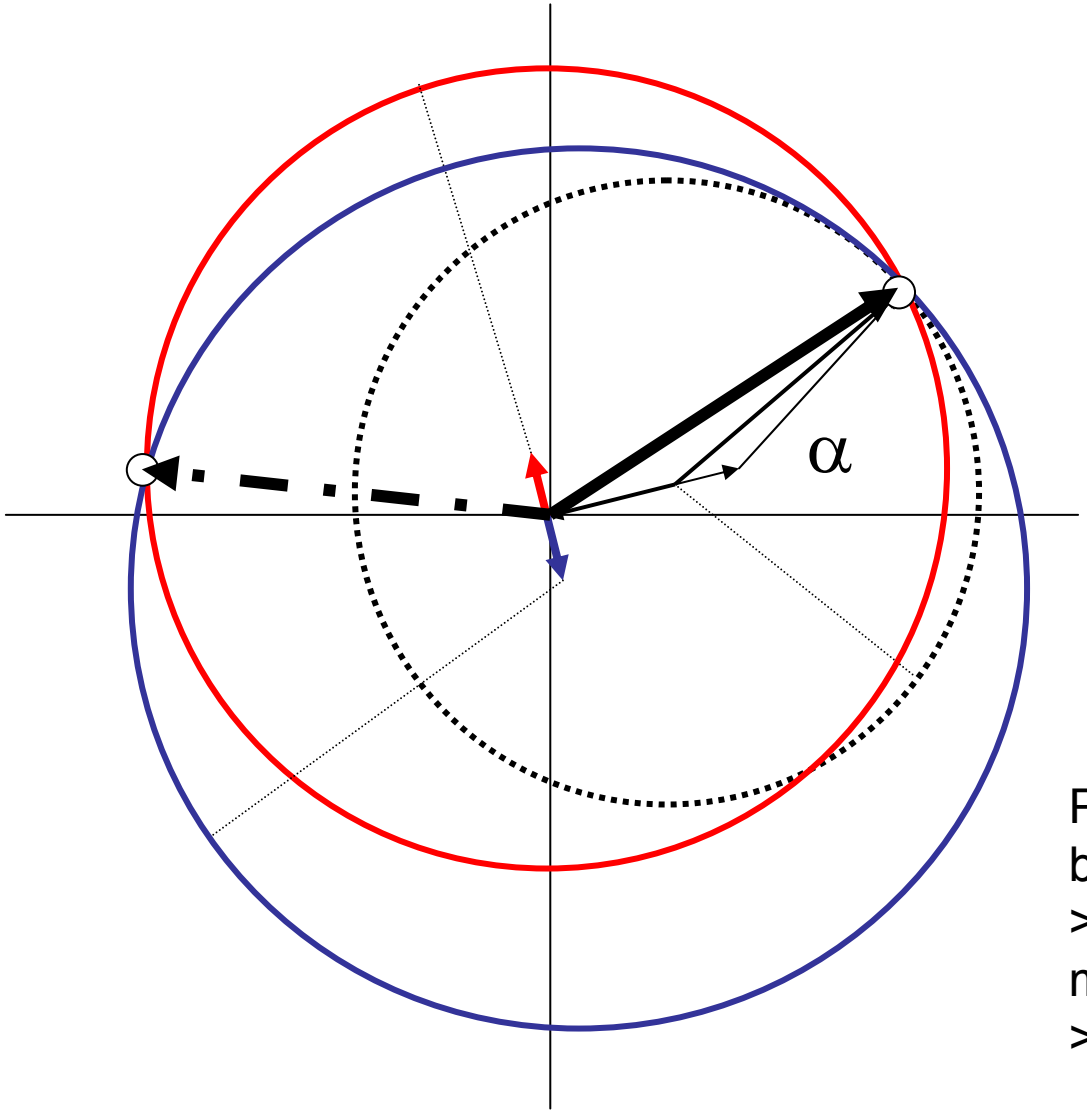
$|F_{\lambda 2}(\mathbf{h+})|$: Anomalous

$|F_{\lambda 2}(\mathbf{h-})|$: Anomalous

$\mathbf{F}_P = |F_P| \exp i\alpha$

Single phase solution can be determined.

SAD



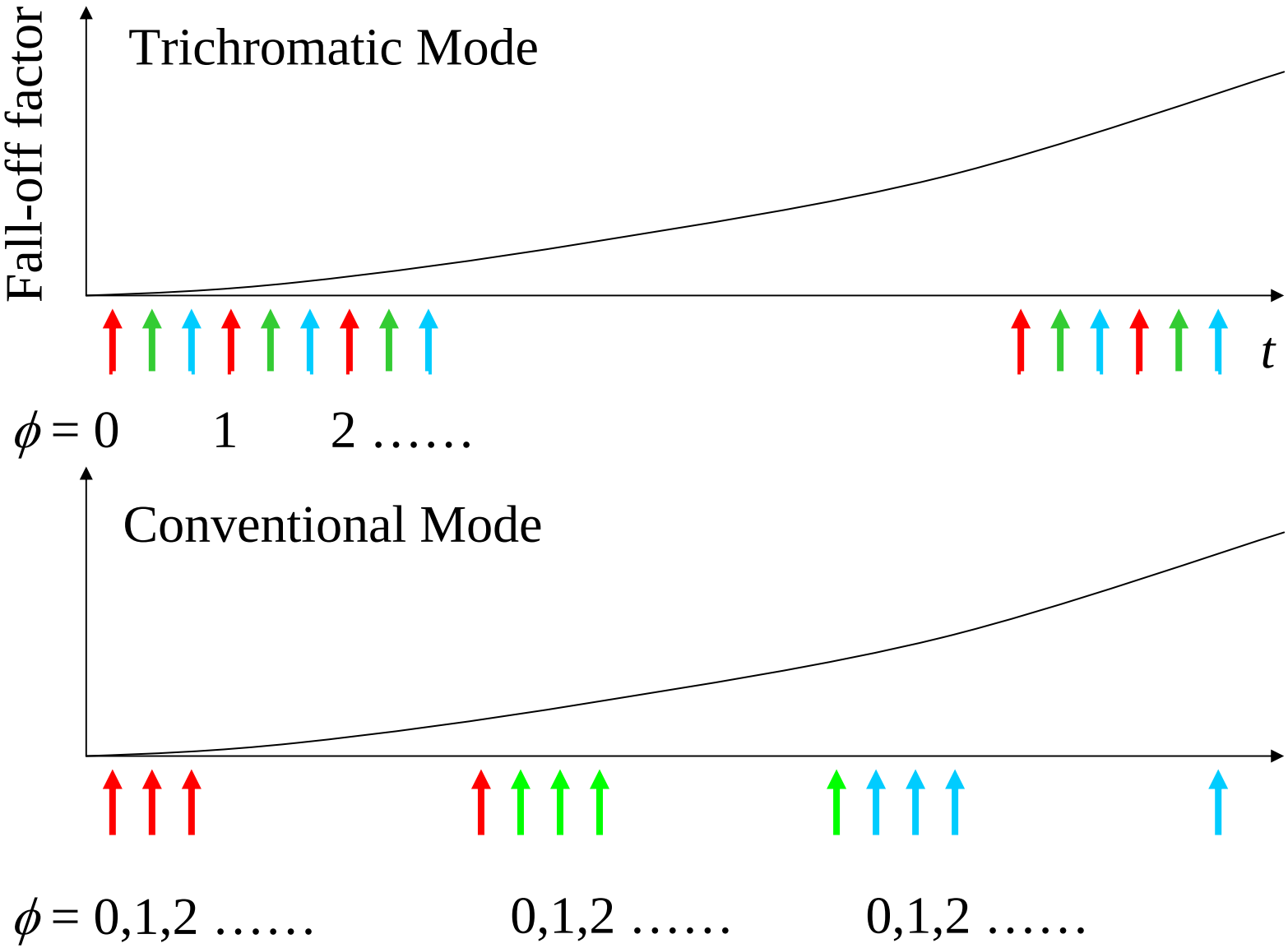
| F |: Native

| F_λ (**h+**) | : Anomalous

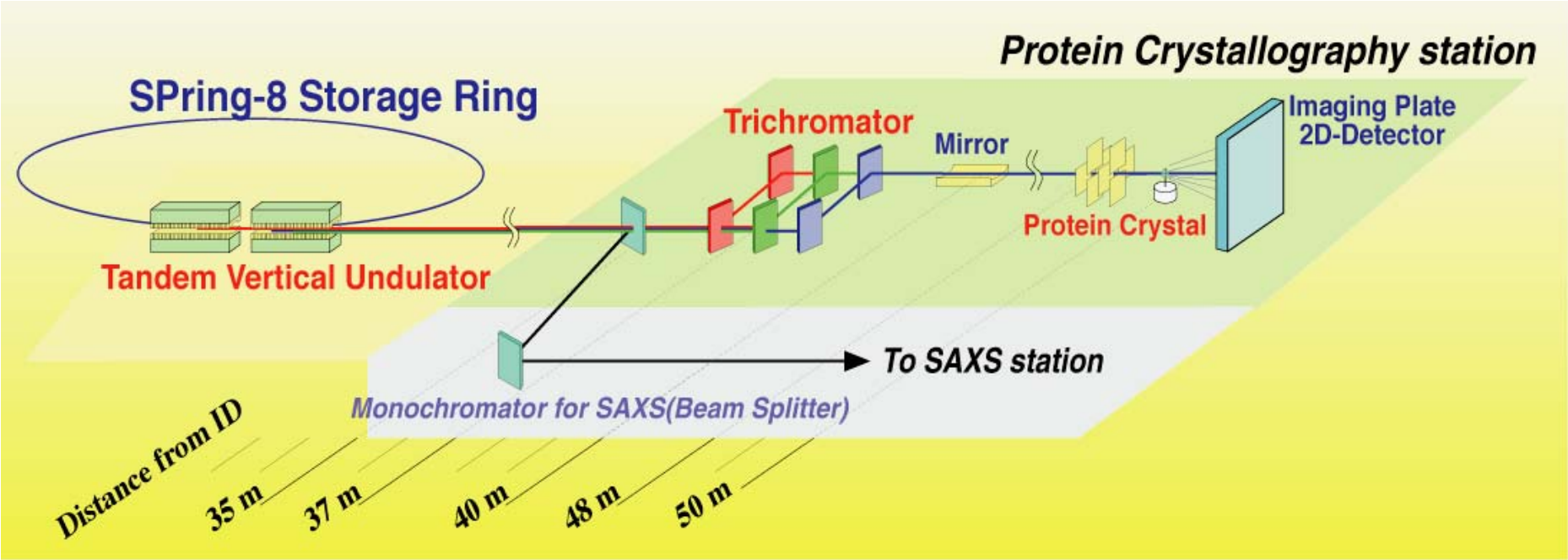
| F_λ (**h−**) | : Anomalous

Phase probability function shows bimodal.
>> Phase improvement by density modification
>> High precision data collection

Effect of Trichromatic Data Collection



Overview of Beamline BL45XU



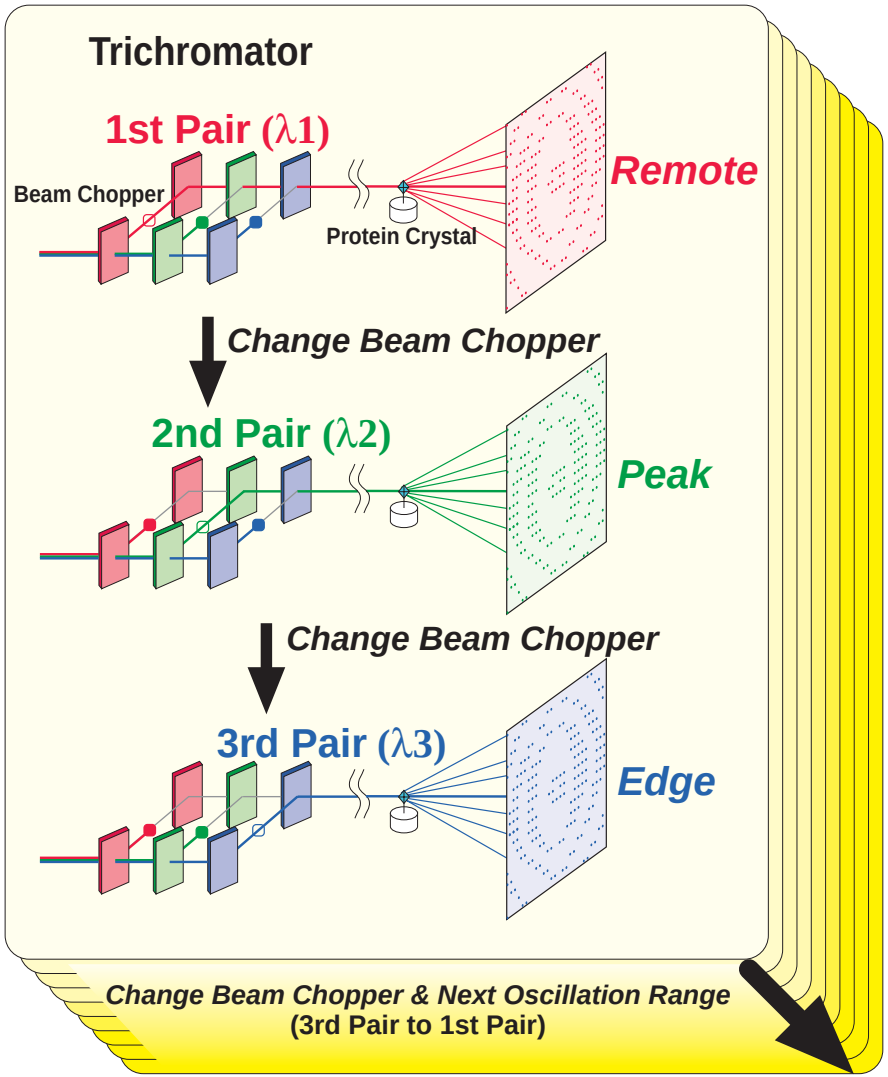
Trichromatic concept for optimizing MAD-experiment

- **Tandem vertical undulator (for High Quality Beam)**
- **Trichromator (for Rapid Data Collection)**

(Yamamoto, Kumasaka, Fujisawa, Ueki, 1996)

Trichromatic Concept for MAD-experiment

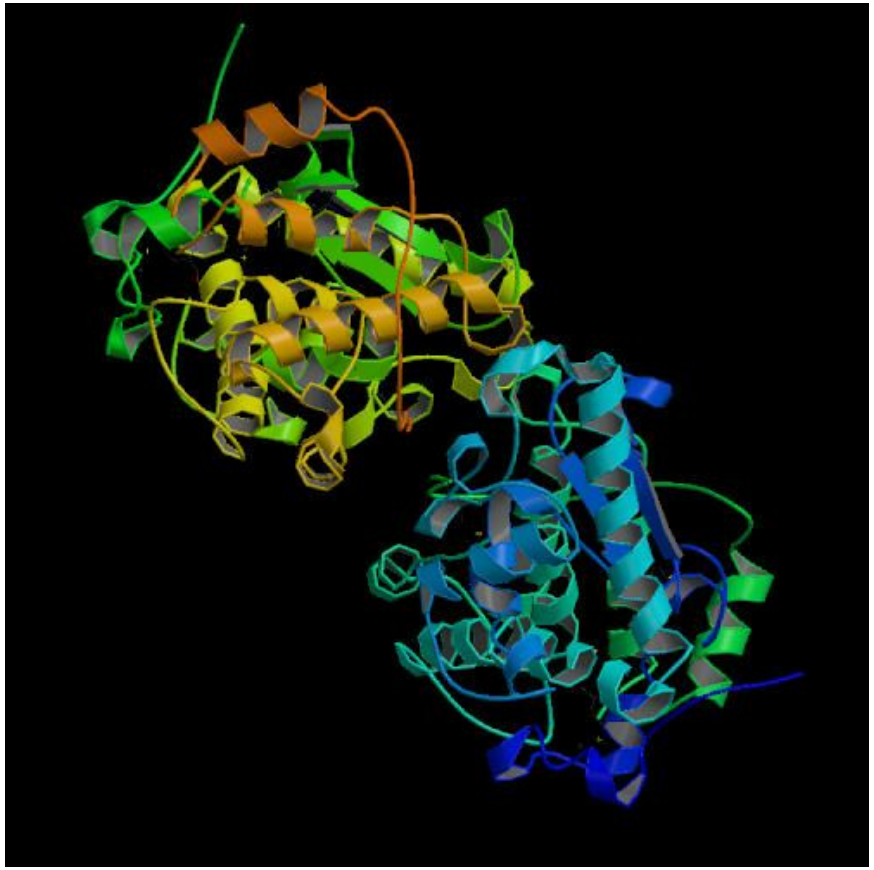
Three-wavelengths data will be taken simultaneously for **the identical protein crystal** without changing the setting by **"Trichromator"**.



Bacterial Chitosanase (Mw 34k, 7 SeMet)

Source: Gram-Negative Bacterium (*Matsuebacter chitosanotabidus* 3001)
Function: Hydrolysis of glycosidic bonds of chitosan
(○GlcN-GlcN, GlcN-GlcNAc, GlcNAc-GlcN, ×GlcNAc-GlcNAc)

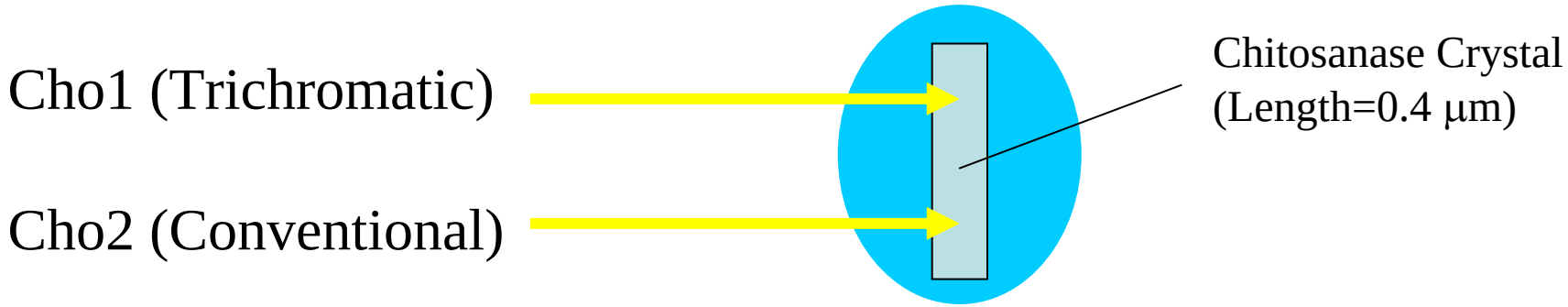
$P2_12_12_1$
a = 51.5, b = 56.2, c = 206.8 Å
1.7 Å Resolution
Two monomers / Asym. unit



Effect of Trichromatic Data Collection

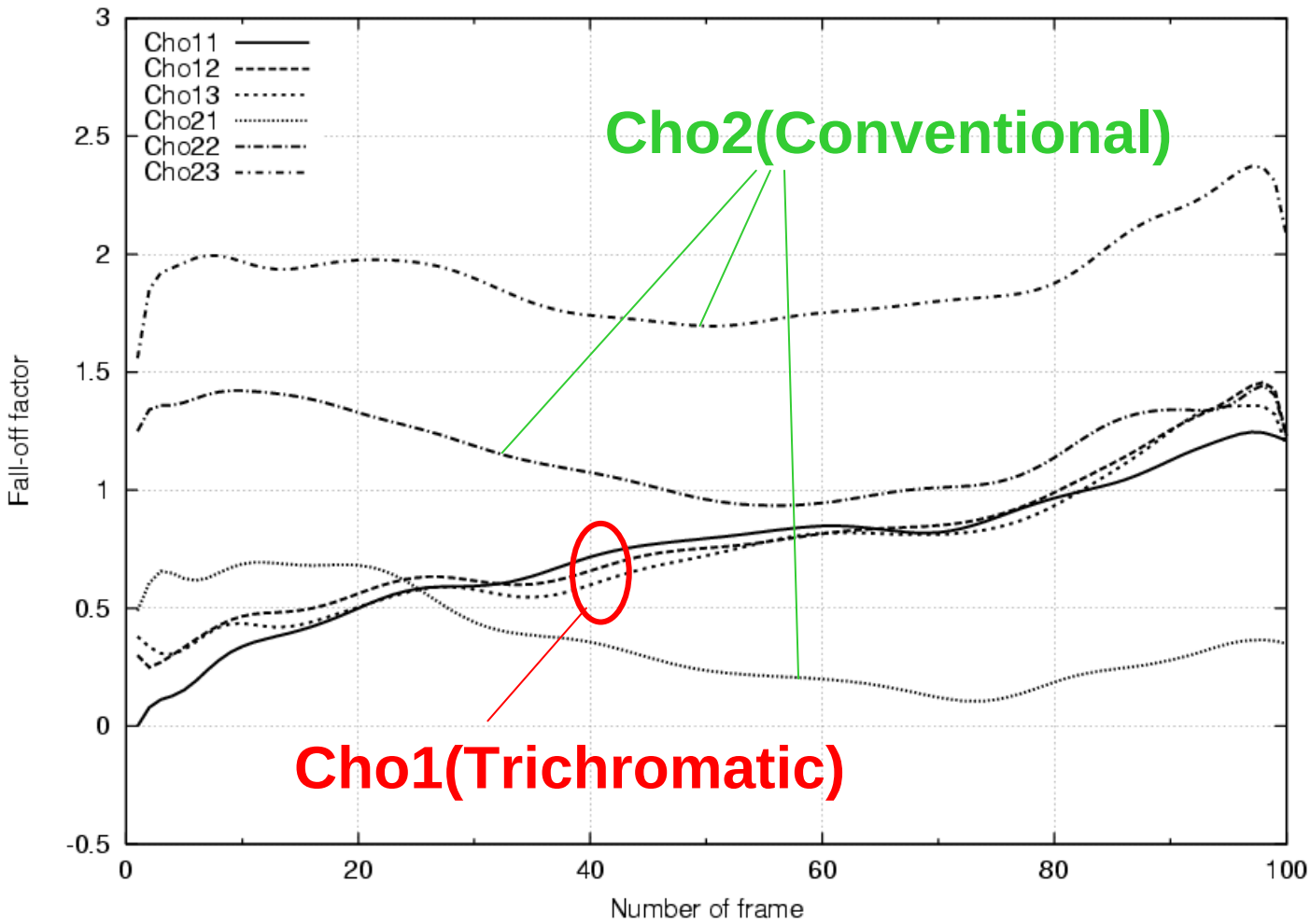
Data Collection Statistics

Data	Observations	Individuals	I / σ	R_{merge}	R_{iso}	B
Cho1: Remote	260,402	65,579	18.5	0.049	—	—
Peak	269,821	66,482	17.4	0.053	0.057	0.08
Edge	269,362	66,428	17.5	0.084	0.048	0.11
Cho2: Remote	261,577	65,695	19.8	0.045	—	—
Peak	263,567	66,480	18.7	0.048	0.064	0.09
Edge	263,387	66,449	17.1	0.076	0.078	-0.39



Effect of Trichromatic Data Collection

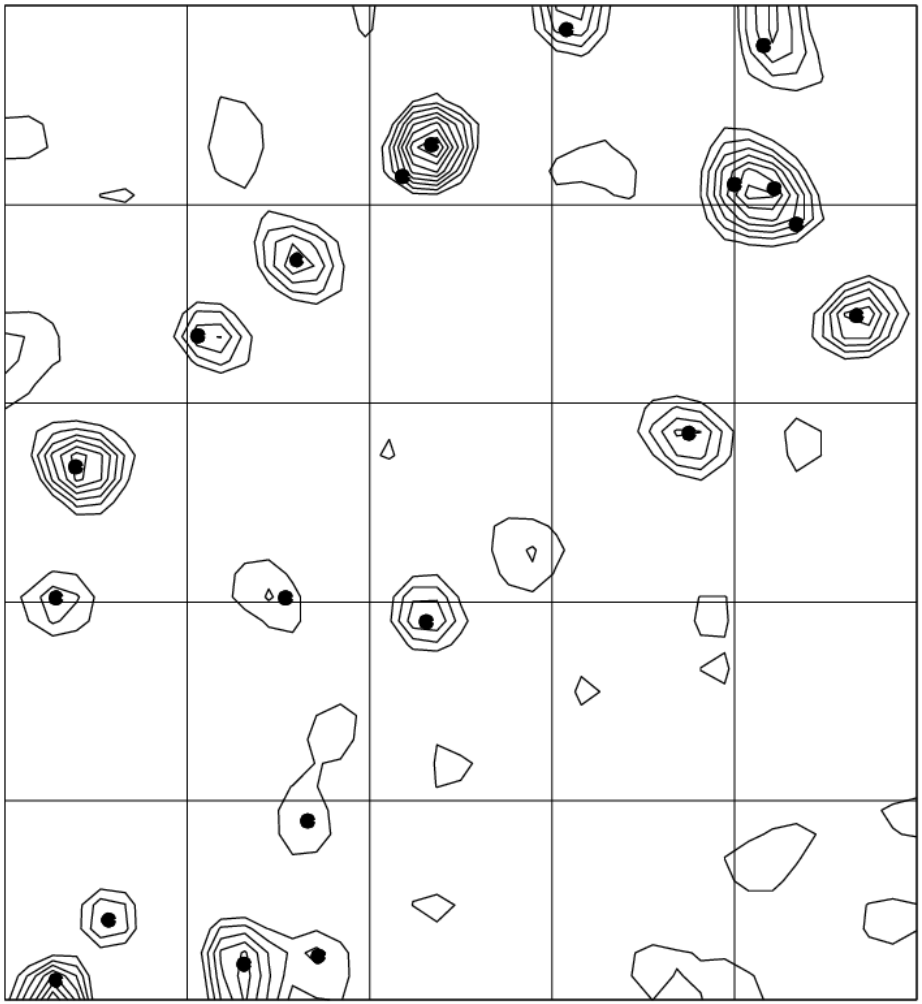
Variation of Fall-off factor



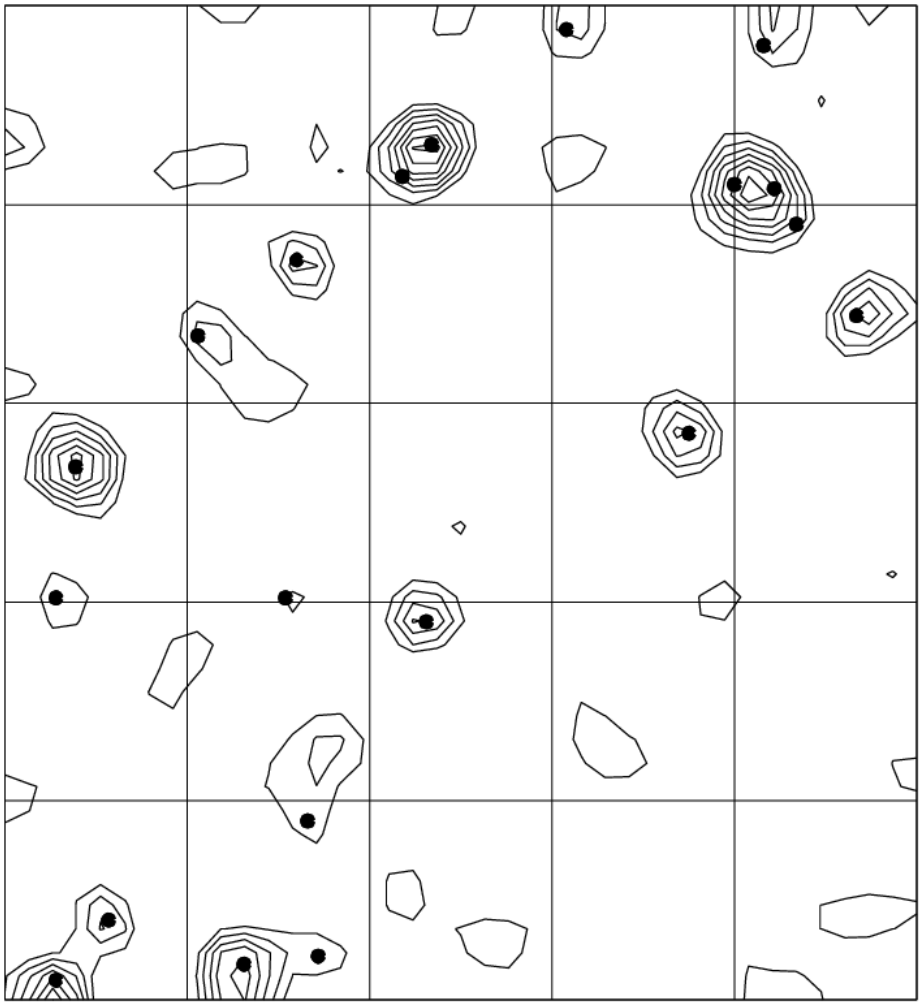
Effect of Trichromatic Data Collection

Comparison with dispersive Patterson maps

Cho1 (Trichromatic)



Cho2 (Conventional)



Harker section (w = ½)

Effect of Trichromatic Data Collection

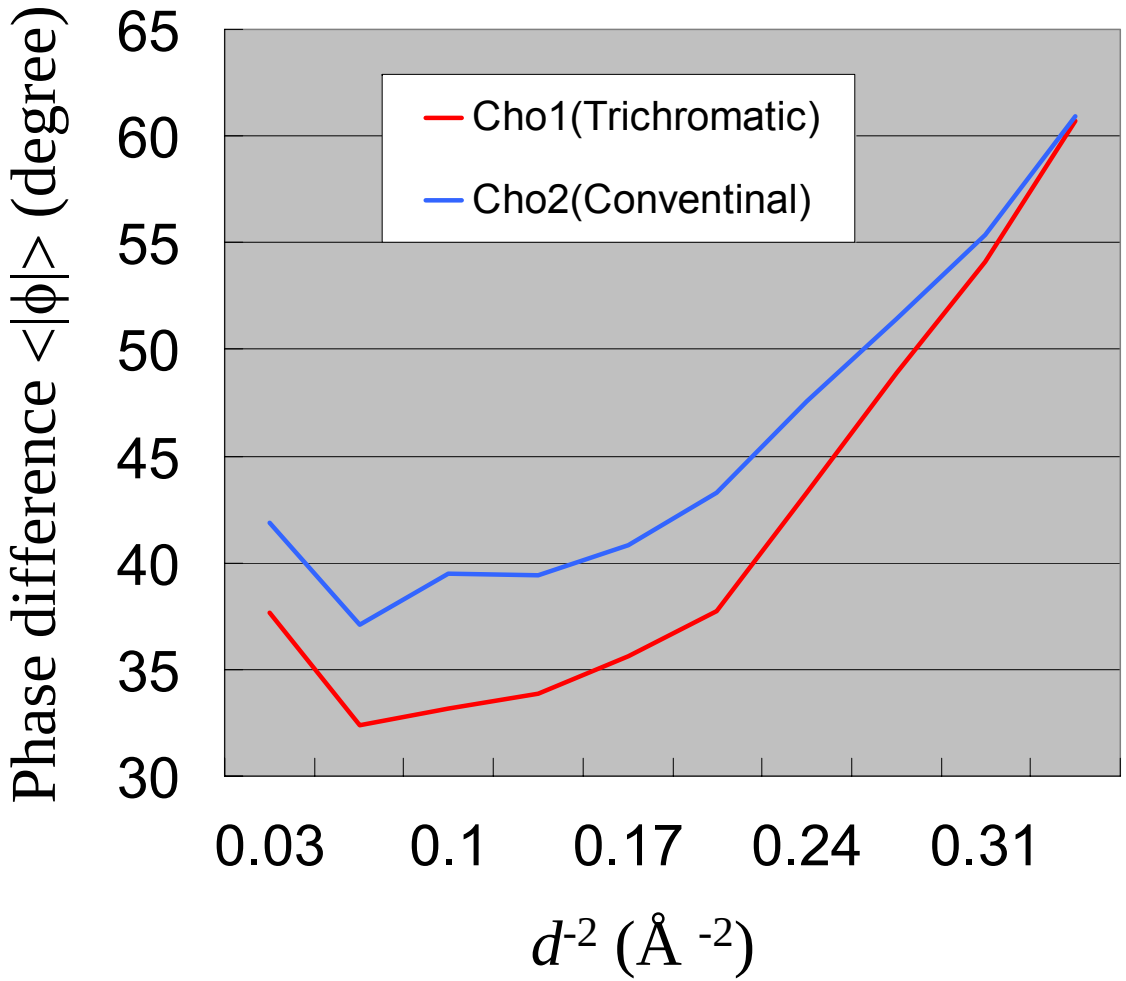
Phasing Statistics (20 – 1.7 Å)

Data		Cho1			Cho2	
	Remote	Peak	Edge	Remote	Peak	Edge
R_{Cullis} (iso)#		0.82 / 0.84	0.83 / 0.88		0.78 / 0.83	0.76 / 0.86
R_{Cullis} (ano)	0.94	0.91	0.99	0.94	0.91	0.99
Lack of closure (iso)#		8.9 / 14.0	8.1 / 12.5		11.4 / 14.7	10.3 / 16.8
Lack of closure (ano)	8.98	16.56	7.32	8.11	15.91	6.37
Figure of merit	0.6057			0.6167		
Phasing power#		1.22 / 0.81	1.19 / 0.82		1.40 / 0.90	1.38 / 0.89
$<\|\Delta\phi\|>$*	44.2	(33.9)		47.8	(39.4)	

#: Acentric and centric values before and after slash.
*: Phase difference against phases calculated from refined model
Parenthesis show the values within the range of 10-2.5 Å.

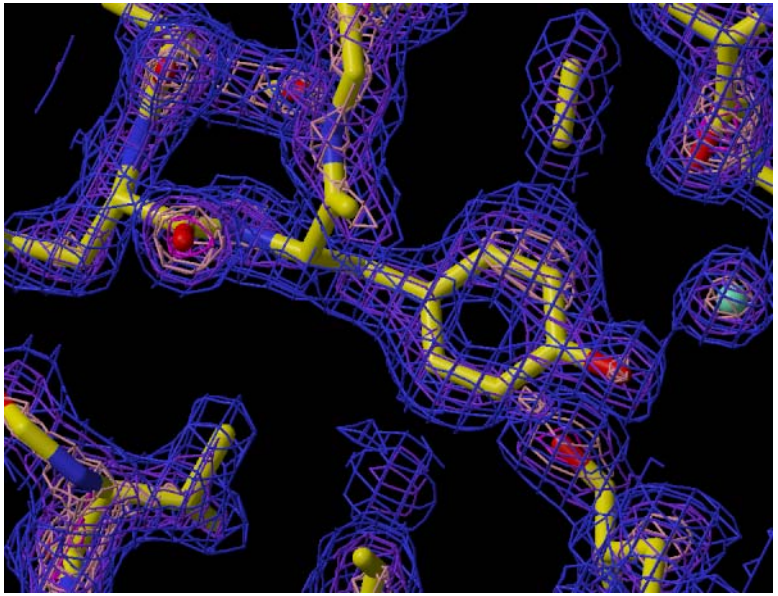
Effect of Trichromatic Data Collection

Phase difference against true phase

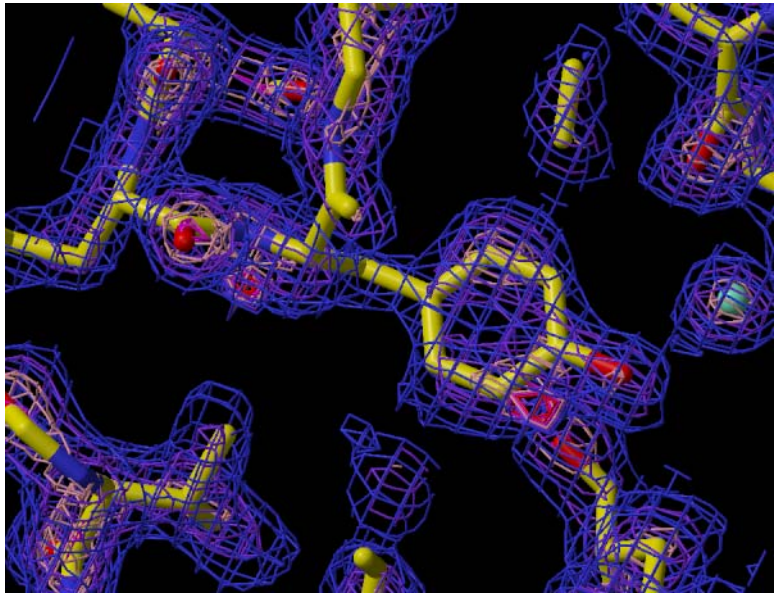


Effect of Trichromatic Data Collection

Quality of electron density map



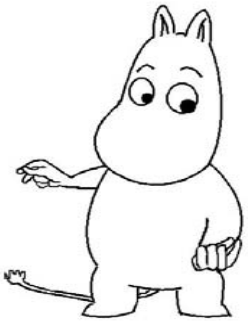
Cho1 (Trichromatic)



Cho2 (Conventional)

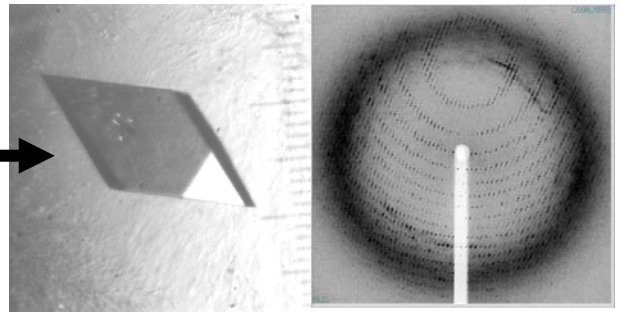
Tyr 165 (Chitosanase A-chain)
1.7 Å MAD phase (without any phase modification)

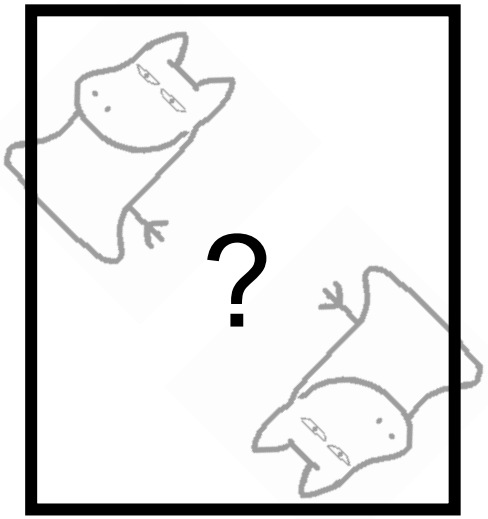
Molecular replacement



Known determined structure

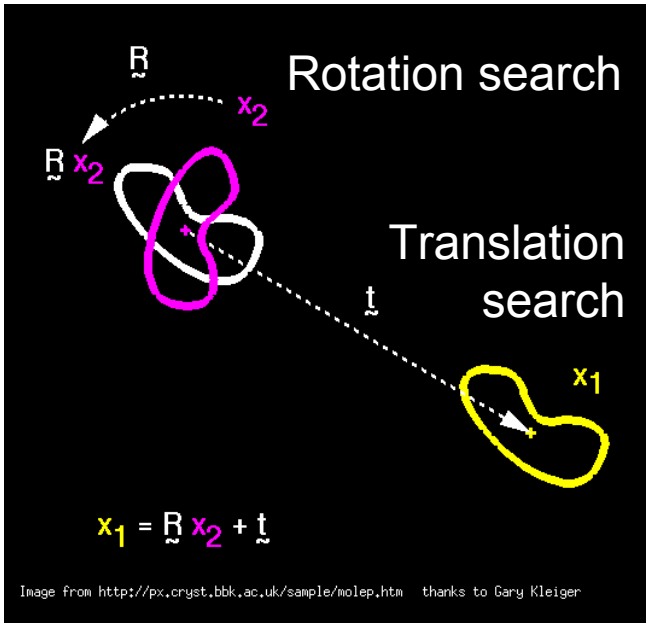
Unknown
but probably similar structure





To solve unknown structure, a known structure is used as a approximation.

The known structure will be selected by sequence similarity. Highest sequence similarity might gives highest structural similarity.



How to pack the molecules into the cell ?
> 6-D search

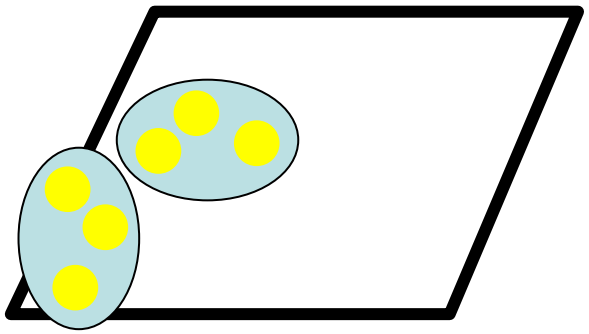
Patterson function

Intramolecular・Intermolecular vectors

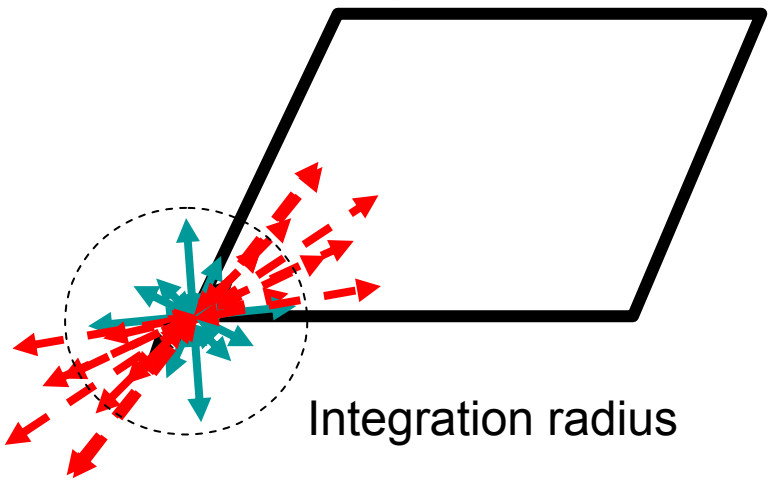
Real space

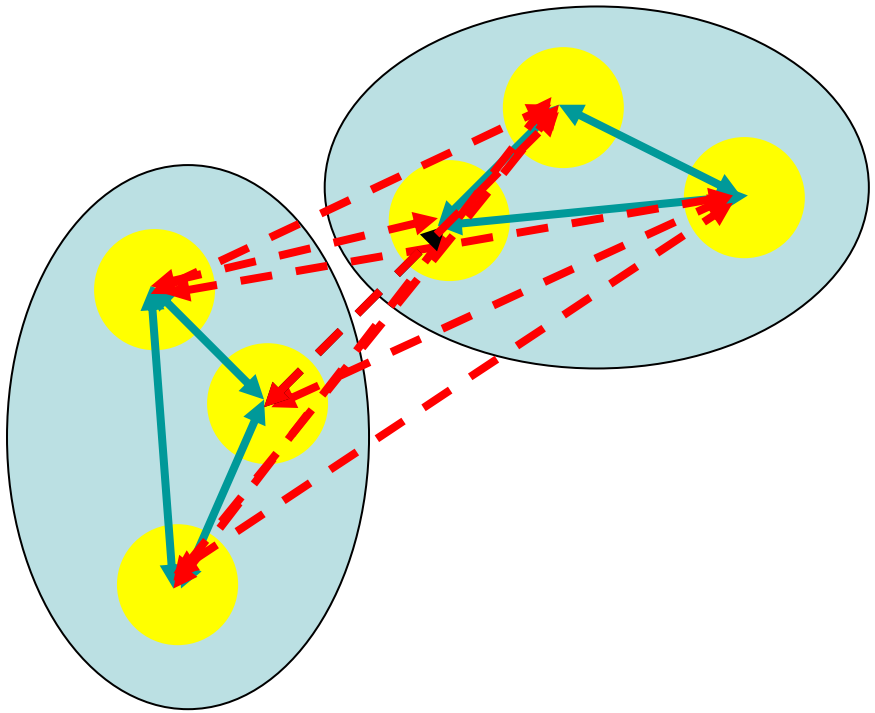
$$\rho(\mathbf{r}) = \int_{\mathbf{S}} \boldsymbol{F}(\mathbf{S}) \exp[-2\pi i\, \mathbf{S} \cdot \mathbf{r}]\, d\mathbf{S}$$

$$P(\mathbf{u}) = \int_{\mathbf{S}} |\boldsymbol{F}(\mathbf{S})|^2 \exp[-2\pi i\, \mathbf{S} \cdot \mathbf{u}]\, d\mathbf{S}$$

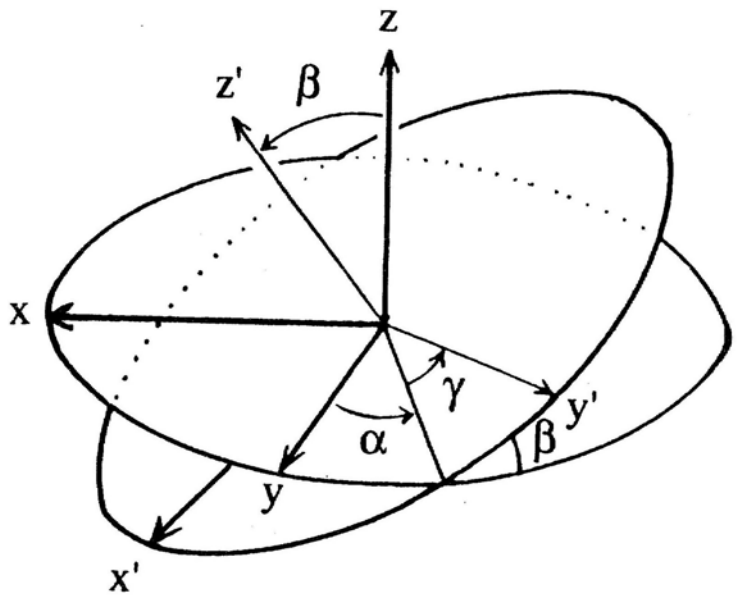


Patterson space





Euler angles, $\alpha\ \beta\ \gamma$

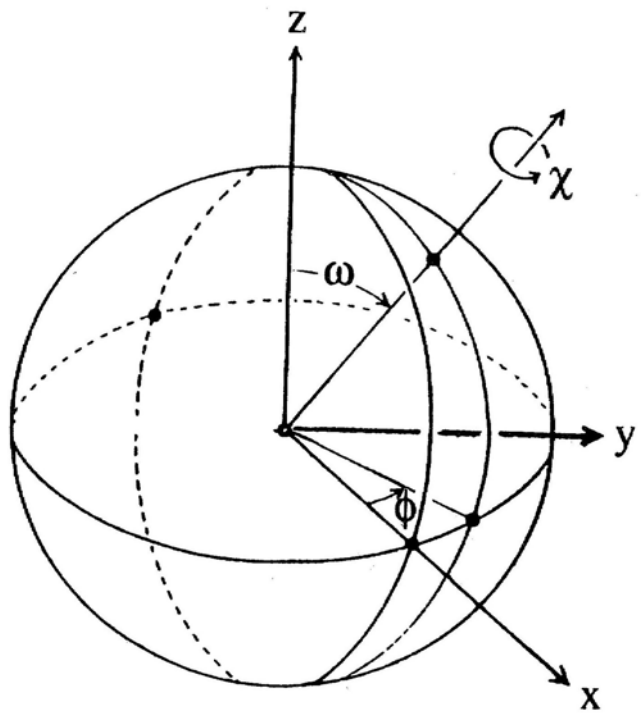


Rotation axis: z'' x' z

$$\begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Order in 3 step rotation: 3 2 1

Polar angles, $\phi\psi\kappa$



z''''	y'''	z''	-y'	-z
$\begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} \cos\varphi & 0 & -\sin\varphi \\ 0 & 1 & 0 \\ \sin\varphi & 0 & \cos\varphi \end{pmatrix}$	$\begin{pmatrix} \cos\kappa & -\sin\kappa & 0 \\ \sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{pmatrix}$	$\begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$
5	4	3	2	1

Real rotation ~ κ

An example of rotation function

α	β	γ	x	y	z	Correlation Coefficient	R-factor
30.37	54.61	351.97	0.000	0.000	0.000	16.0	48.9
59.63	125.39	171.97	0.000	0.000	0.000	16.0	48.9
27.57	41.41	20.51	0.000	0.000	0.000	9.2	51.1
62.43	138.59	200.51	0.000	0.000	0.000	9.2	51.1
17.43	98.67	334.32	0.000	0.000	0.000	7.2	51.7
72.57	81.33	154.32	0.000	0.000	0.000	7.2	51.7
41.73	139.11	197.95	0.000	0.000	0.000	7.7	52.1
48.27	40.89	17.95	0.000	0.000	0.000	7.7	52.1
81.84	98.18	226.67	0.000	0.000	0.000	8.2	51.6
8.16	81.82	46.67	0.000	0.000	0.000	8.2	51.6

Modeling & refinement of structure

Modeling: Construct molecular model to fit obtained electron density using interactive molecular graphics software or automated modeling software.

Refinement: Optimization of observed and calculated F data by shifting atomic coordinates.

R-factor: Crystallographic Reliability-factor

$$R1=\Sigma\ ||Fo|-|Fc(\mathbf{r})||/\Sigma\ |Fo|$$

$$wR2=(\Sigma\ w(|Fo|^2-|Fc(\mathbf{r})|^2)^2/\Sigma\ w(|Fo|^2)^2)^{1/2}$$

Cross validation of R-factor (R_{free})

Refinement of structural model

1) Unrestraint refinement

Only using R-factor refinement

in case of ultra-high resolutions (0.8 A or higher)

2) Restraint refinement

Coupled with molecular mechanics

Model validity is also guaranteed by low energy

~ structural stability

Target function

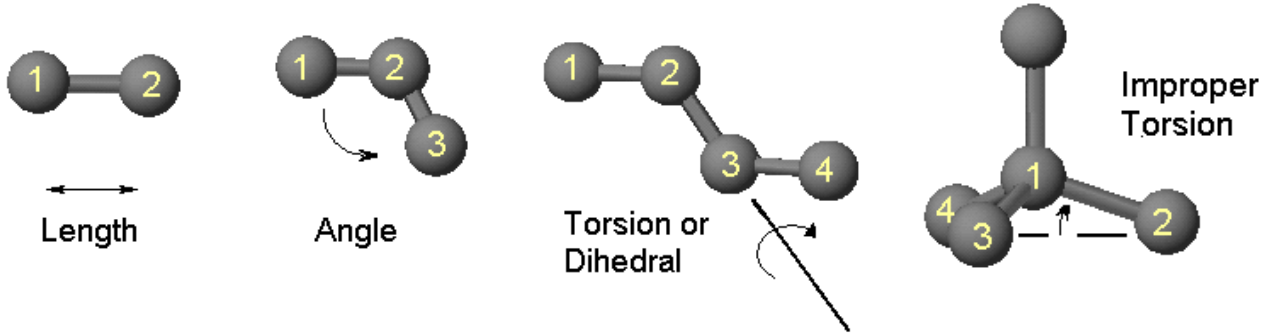
$$E = E_{\mathrm{chem}} + w_{\mathrm{xray}} E_{\mathrm{xray}}$$

$$E_{\mathrm{xray}} = \sum_{\mathbf{h}} \left| F_{\mathrm{O}}(\mathbf{h}) - k F_{\mathrm{C}}(\mathbf{h}) \right|^2$$

Basics of molecular mechanics (MM)

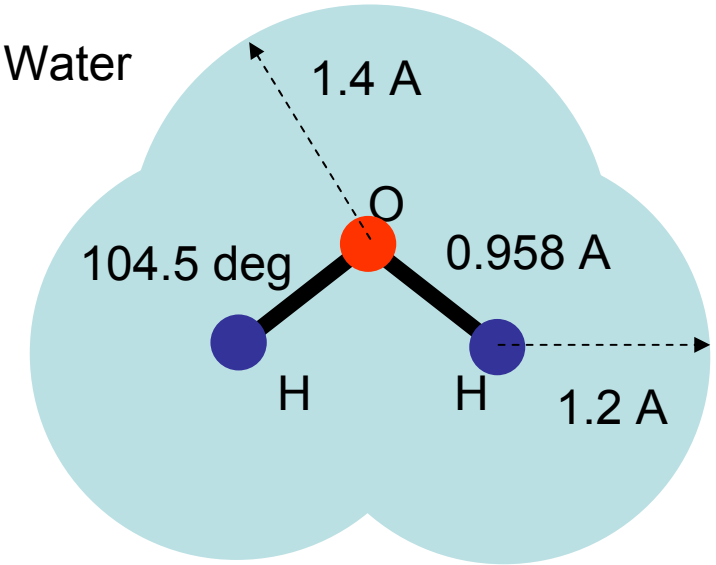
Energy calculation of atomic bonds and interactions by classical mechanics.

bond length
bond angle
dihedral/torsion
improper dihedral
van der Waals
electrostatic



$$E_{chem} = E_{bond} + E_{angle} + E_{dihe} + E_{impr} + E_{vdW} + E_{elec}$$

e.x. Water



$$E_{bond} = k_{bond} (r - r_{ide})^2$$
$$E_{angle} = k_{angle} (\theta - \theta_{ide})^2$$

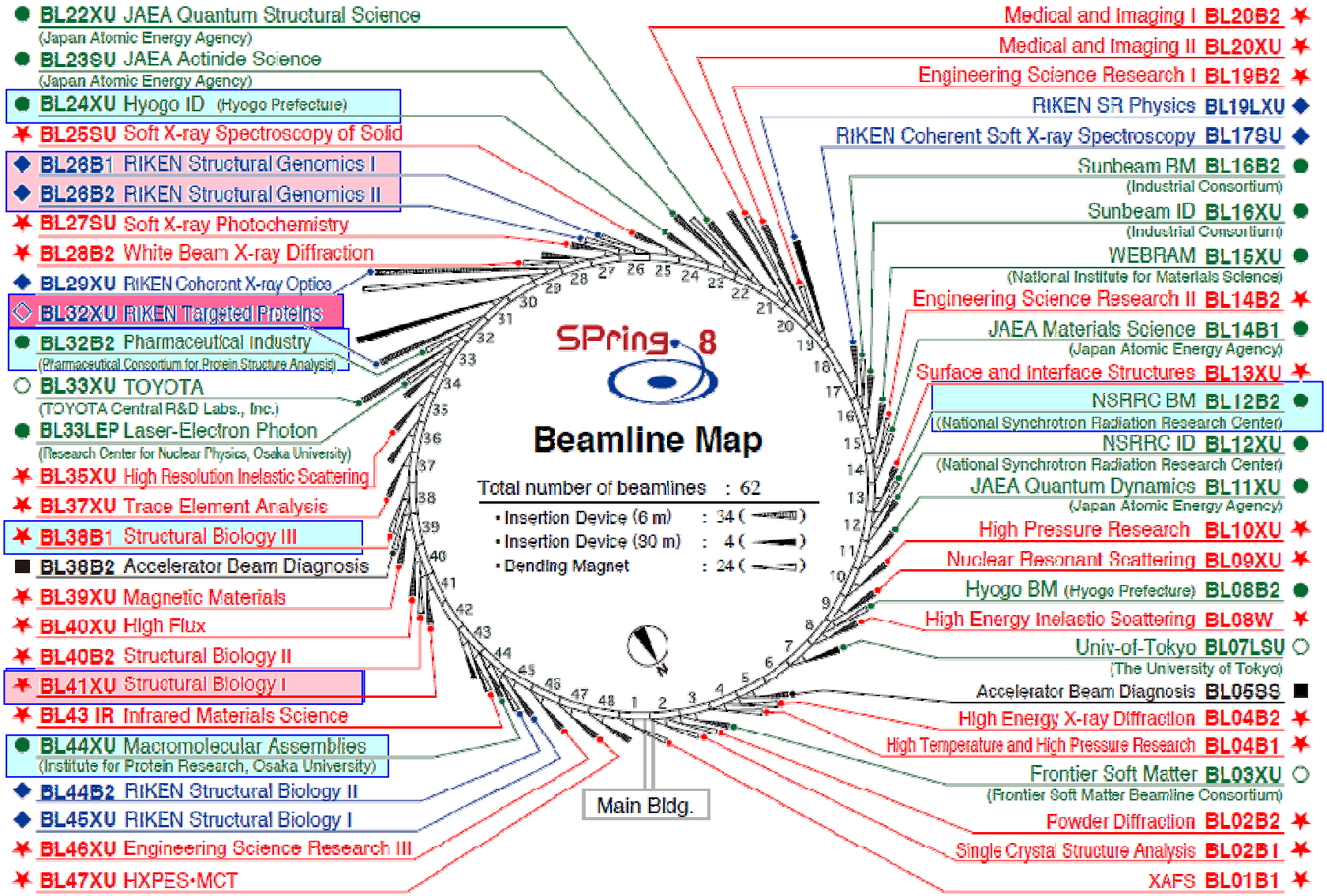
Force Field

TIPS3P Parameter

Bond O-H	450.0	0.9572
Bond H-H	0.0	1.5139
Angle H-O-H	55.0	104.52

3: Recent advances in MX beamlines

MX Beamlines at SPring-8



■ 9 of 53 beamlines are dedicated to MX

Beamlines and User Accessibility

1. Public Beamlines (BL41XU, BL38B1; JASRI)
   Academic use + Proprietary use (incl. Mail-in service)

2. Contract Beamline (BL44XU; Osaka Univ.)
   Academic use

   Contract Beamline (BL24XU; Hyogo Pref.)
   Academic use + Partially opened to proprietary use

3. RIKEN Beamlines (BL26B1&B2, BL32XU; RIKEN)
   RIKEN's academic research + Partially opened to public use (20%)

4. Pharmaceutical Industrial Beamline (BL32B2; PcProt)
   Fully operated for propietary use
   by the members of
   Japan Pharmaceutical Manufacturers Association (JPMA)

Synchrotron MX

Brilliant synchrotron radiation facilitates MX research

1. For cutting edge research

   High precision data collection

   for Micro-crystal & Large unit-cell samples

2. For structural genomics approach

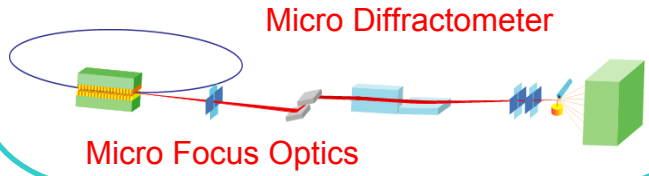
   Automated and rapid data collection

   for High throughput screening

SPring-8 MX beamline complex

Micro Focus (BL32XU)

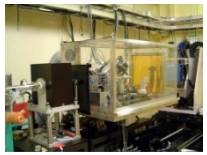
- Micro Focus Optics for Micro-Crystal
  < 10 μm
- Sample Handling for Micro-Crystal



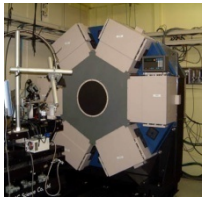
SPring-8

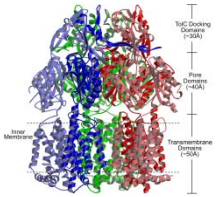
Large Molecular Complex (BL44XU)

- Parallel Beam for Large Unit Cell
  > 500 Å



P2 station for Virus



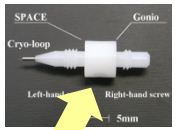


High-speed Network

Remote Access

High Throughput (BL26B1/2, BL32B2, BL38B1)

- Stable bending magnet beam
- Beamline automation
- Mail-in Data Collection

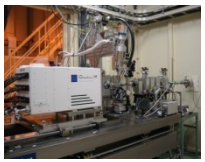


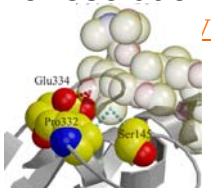
Data Server

- Large Data Storage
- On-line Analysis
- Data base

High-resolution Analysis (BL41XU)

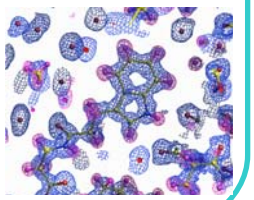
- High Precision Data collection
- Sub-atomic resolution





Substrate Complex

/isualization of H-atom

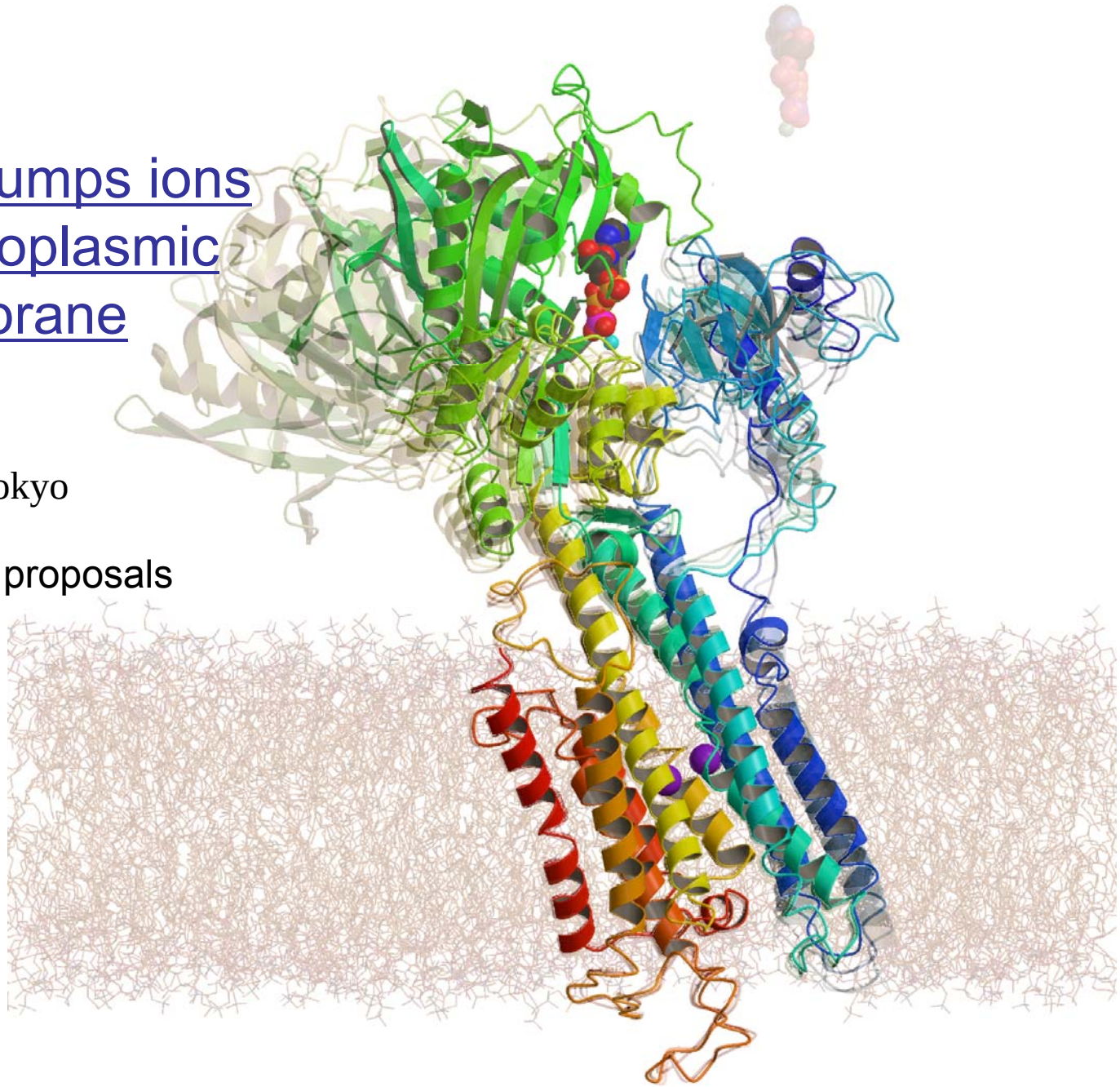


Recent topics

Ca^{2+}-ATPase pumps ions across the sarcoplasmic reticulum membrane

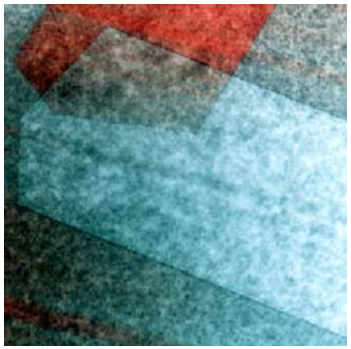
Chikashi Toyoshima
IMCB, The Univ. of Tokyo

A result of long term proposals



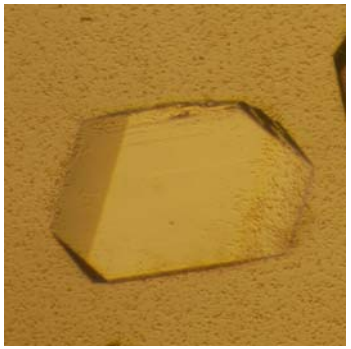
Crystals of Ca^{2+}-ATPase in various states

E1·2Ca^{2+}



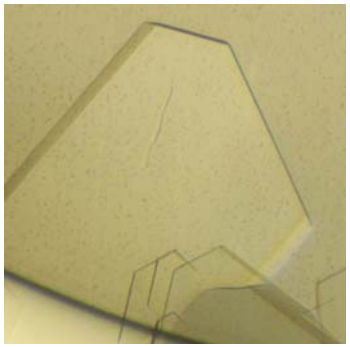
C2
1.95 Å

E1·ATP (AMPPCP)



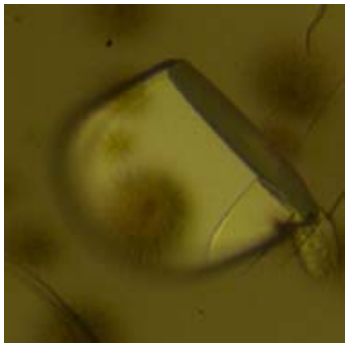
C2
2.5 Å

E1P·ADP (AlF_4^-·ADP)



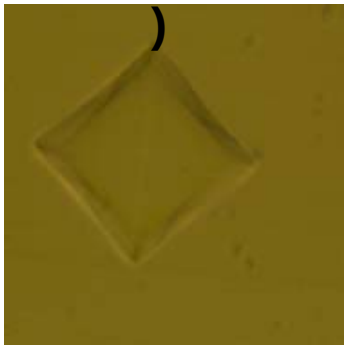
C2
2.4 Å

E2·Pi (MgF_4^{2-})

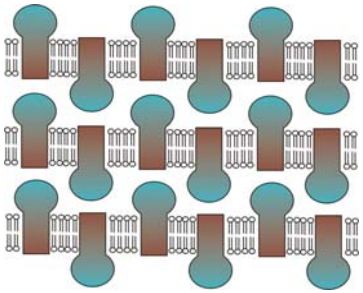


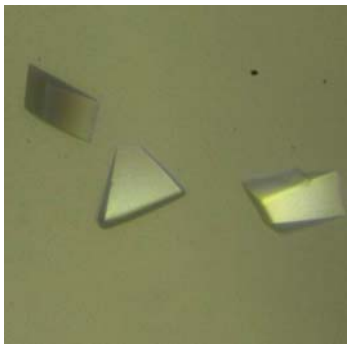
$P2_1$
2.3 Å

E2 (TG,BHQ,CPA)

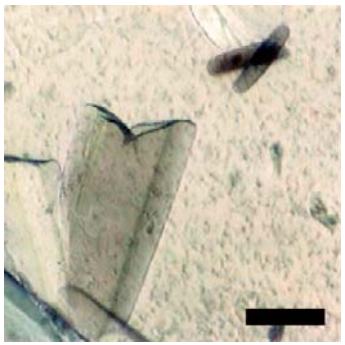


$P4_12_12$
2.1 Å

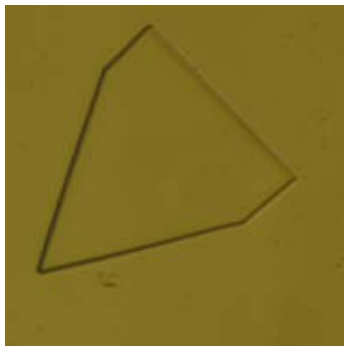




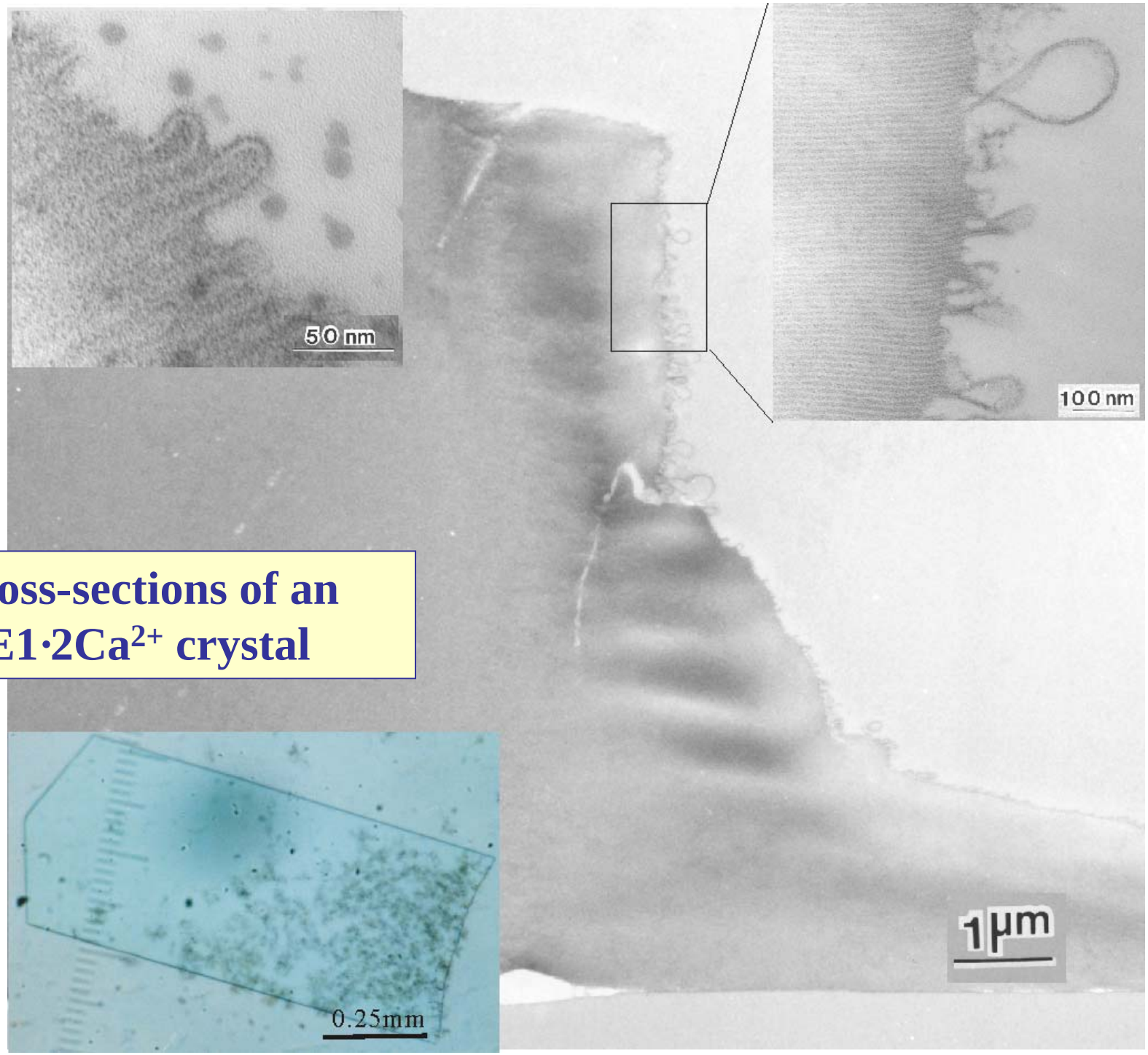
$P2_1$
2.9 Å



C2
2.9 Å



$P2_1$
3.3 Å



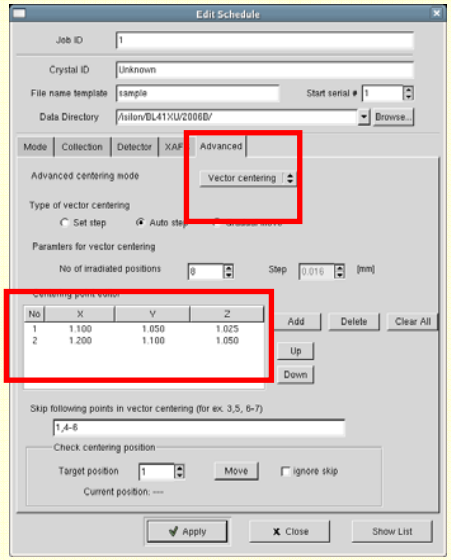
cross-sections of an E1·2Ca^{2+} crystal

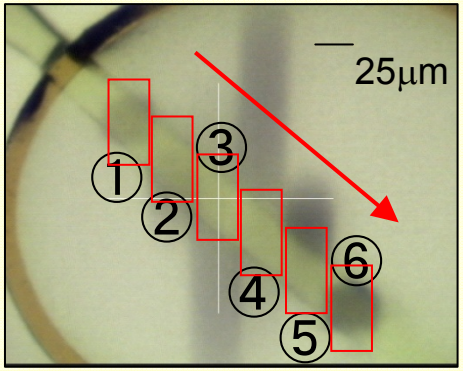
Data acquisition tools

Multiple X-ray exposures on one/several crystals
to control serious radiation damage during data collection
to find better diffracting crystals or crystal segments

・Vector Centering：

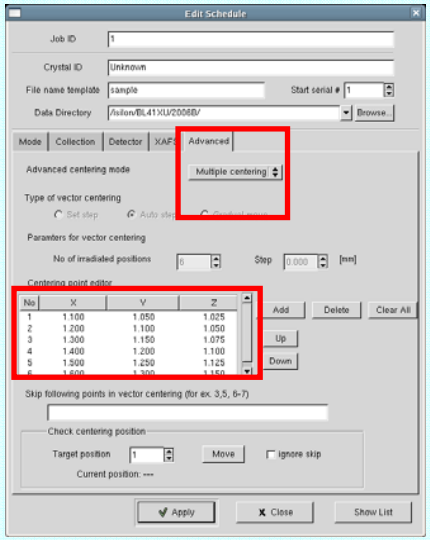
Changing exposure position linearly in every constant frames on *One crystal*

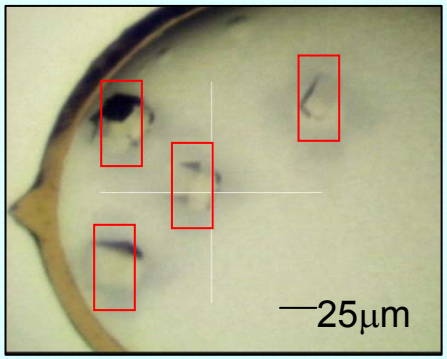


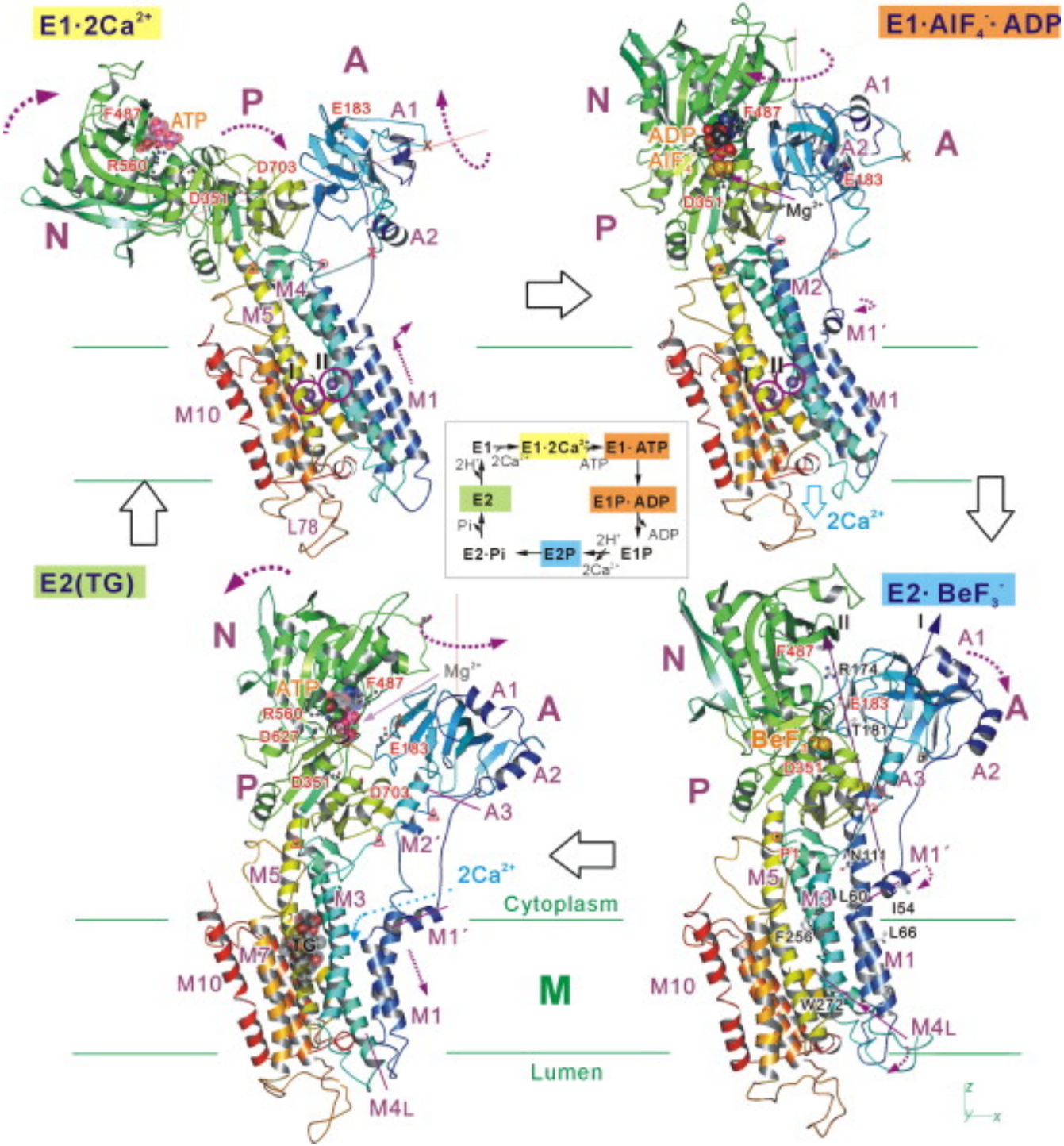


・Multiple Centering：

Multiple exposure positions using several crystals in *One cryo-loop*

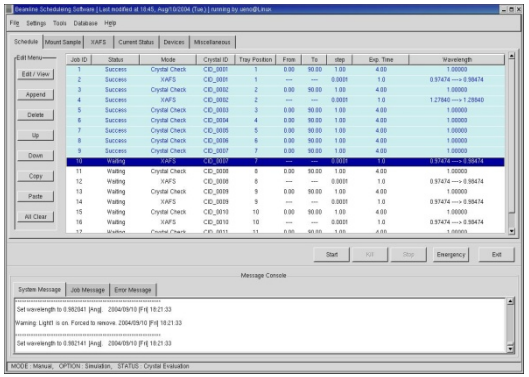





E1·2Ca2+
E1·AlF4-·ADP
E2(TG)
E2·BeF3-
A
P
N
ATP
F487
R560
D703
D351
E183
A1
A2
A3
M1
M1´
M2
M2´
M3
M4
M4L
M5
M7
M10
L78
TG
Mg2+
ADP
AlF4-
D627
2Ca2+
Cytoplasm
Lumen
M
BeF3-
R174
T181
N111
L60
I54
L66
F256
W272
E1
E1·2Ca2+
E1·ATP
E2
E1P·ADP
E2·Pi
E2P
E1P
2H+
ATP
ADP
Pi

Remote Access
Convenient use without visitation

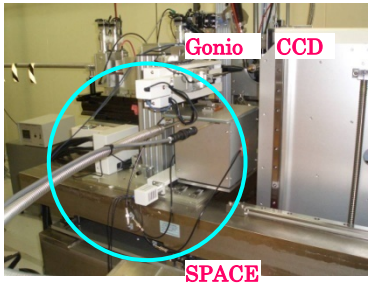
Automation of beamline control



Beamline control software BSS

- Integrated control of beamline optics and diffractometer.
- Automatic set-up of measurement condition.
- Multiple measurements proceed in the scheduled order.

Automation of sample mounting



Sample changer SPACE

- Mount specimens on goniometer instead of users.
- Combination of BSS & SPACE enables automatic data collection with exchanging crystals.

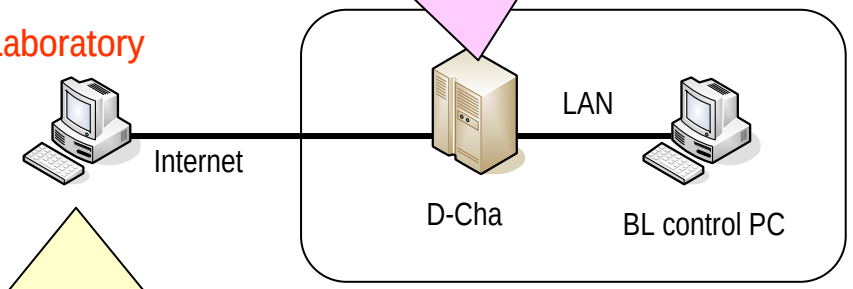
Mail-in data collection

- Send crystal to SPring-8
- Edit measurement condition with D-Cha
- Collect data with a help of BL operator

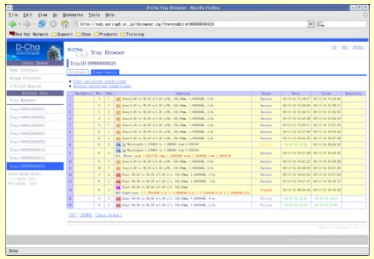
Beamline database D-Cha

Sharing information between users' lab & SPring-8

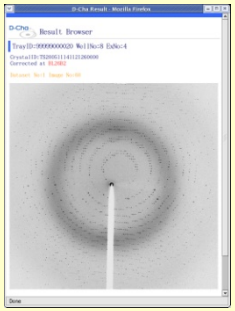
- Crystal related information
- Measurement condition & schedule
- Recorded images



Web interface of D-Cha



Measurement schedule

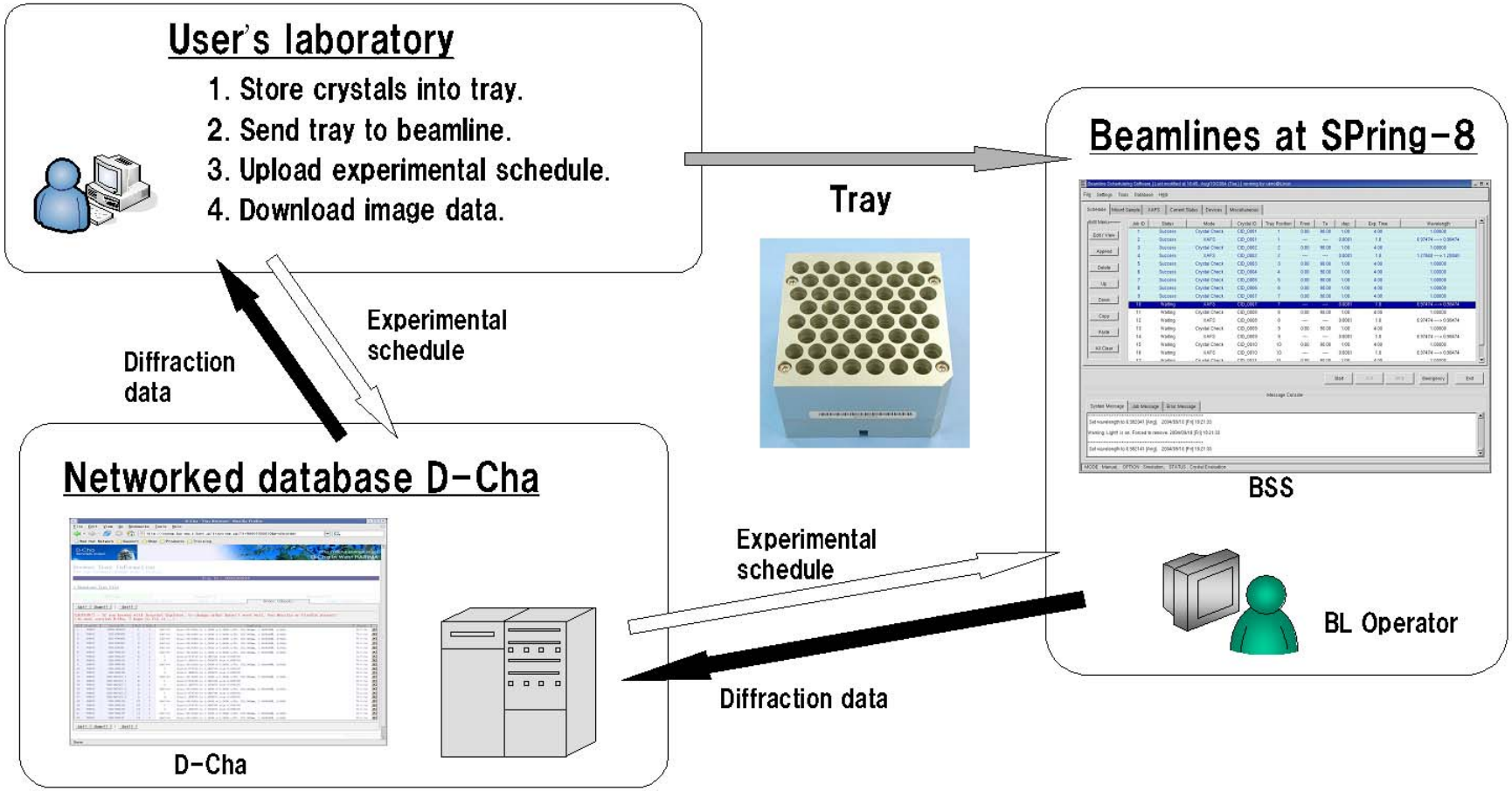


Diffraction image

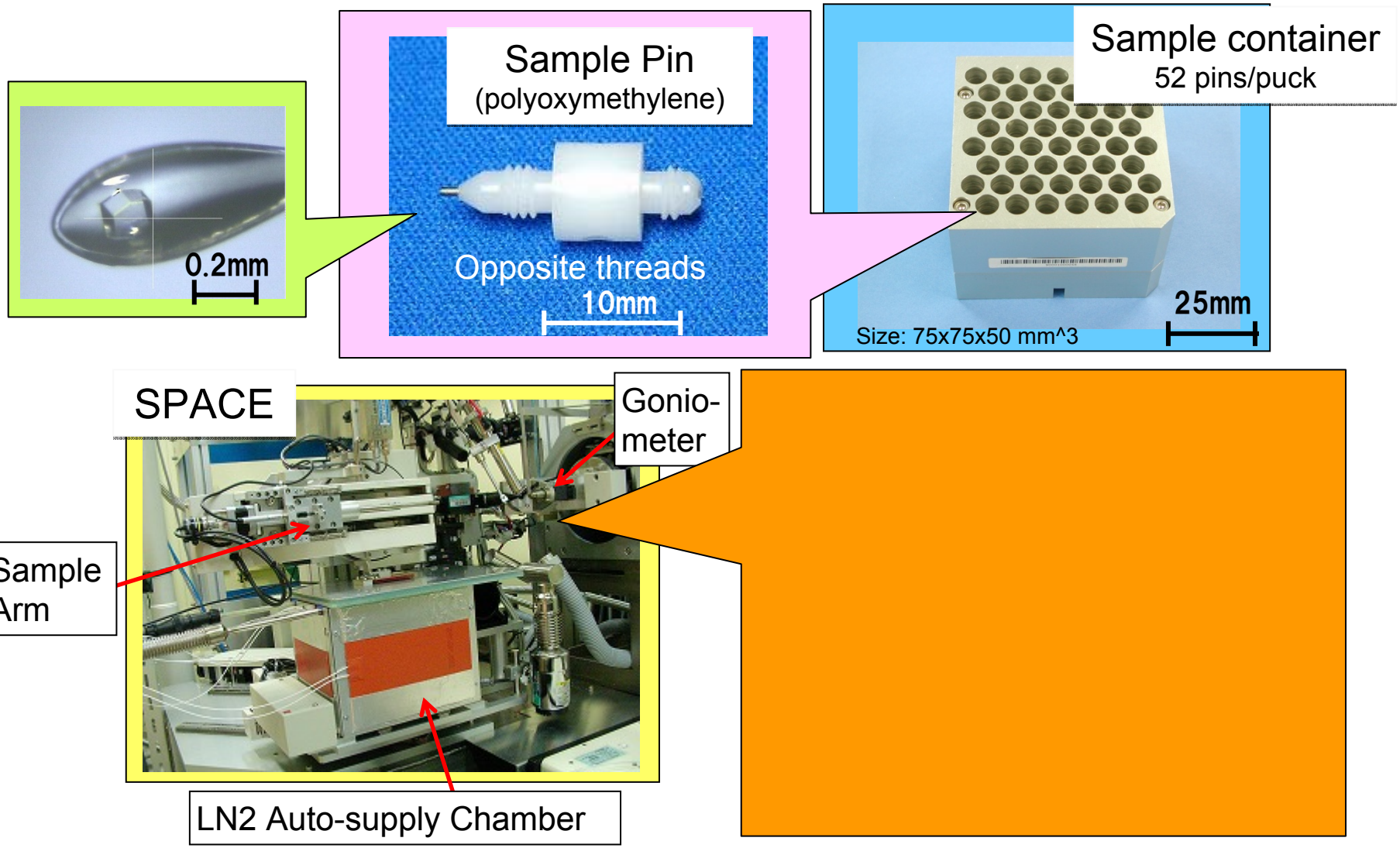
BL Automation

Samples & data management

D-Cha (Database for Crystallography with Home-lab. Arrangement)
Crystal information database, Schedule editor, Image browser, etc. on the Internet

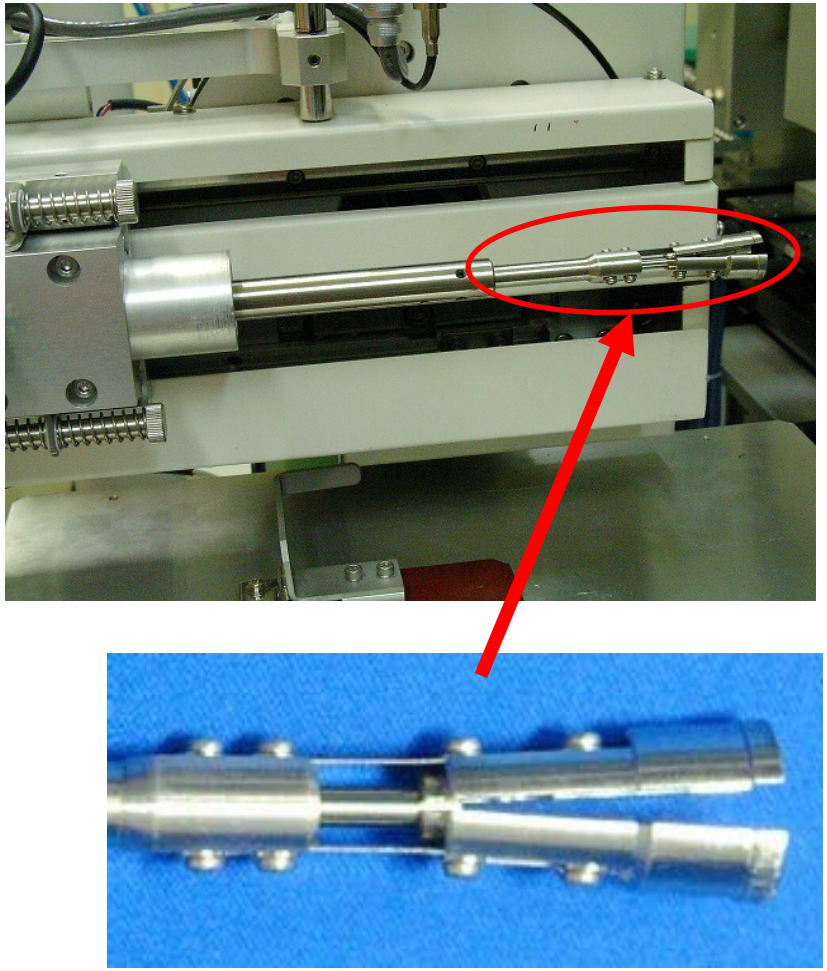


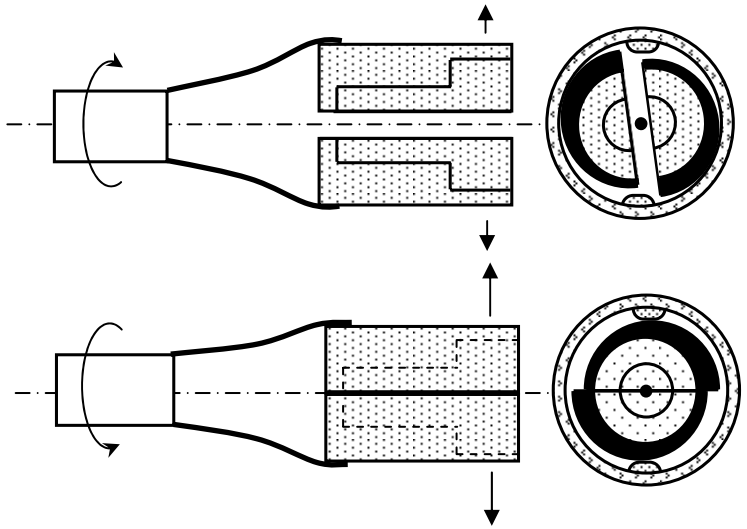
SPACE – SPring-8 Precise Automatic Cryo-sample Exchanger



(Ueno, Kanda, Ida, Kumasaka, Yamamoto et al.)

New attachment for Hampton/SPINE pins





(Murakami, Ueno, Yamamoto et al., Patent #2009-115652)

Remote access at SPring-8 MX beamlines

Basic concept

Make use of sample changer SPACE, integrated beamline control software *BSS* and beamline database *D-Cha* to make all operations required in MX data collection from remote env.

Characteristic features of remote access at SPring-8

- Spec fulfilling the strict radiation safety regulation in Japan
  - Not use the NX client for secure access to BL inside gateway.
  - Remote client & local server architecture with the original protocol.
  - Authentication gateway and operation restricts units are installed between the client-server to ensure safety of remote access.
- New original protocol developed as a SPring-8 standard

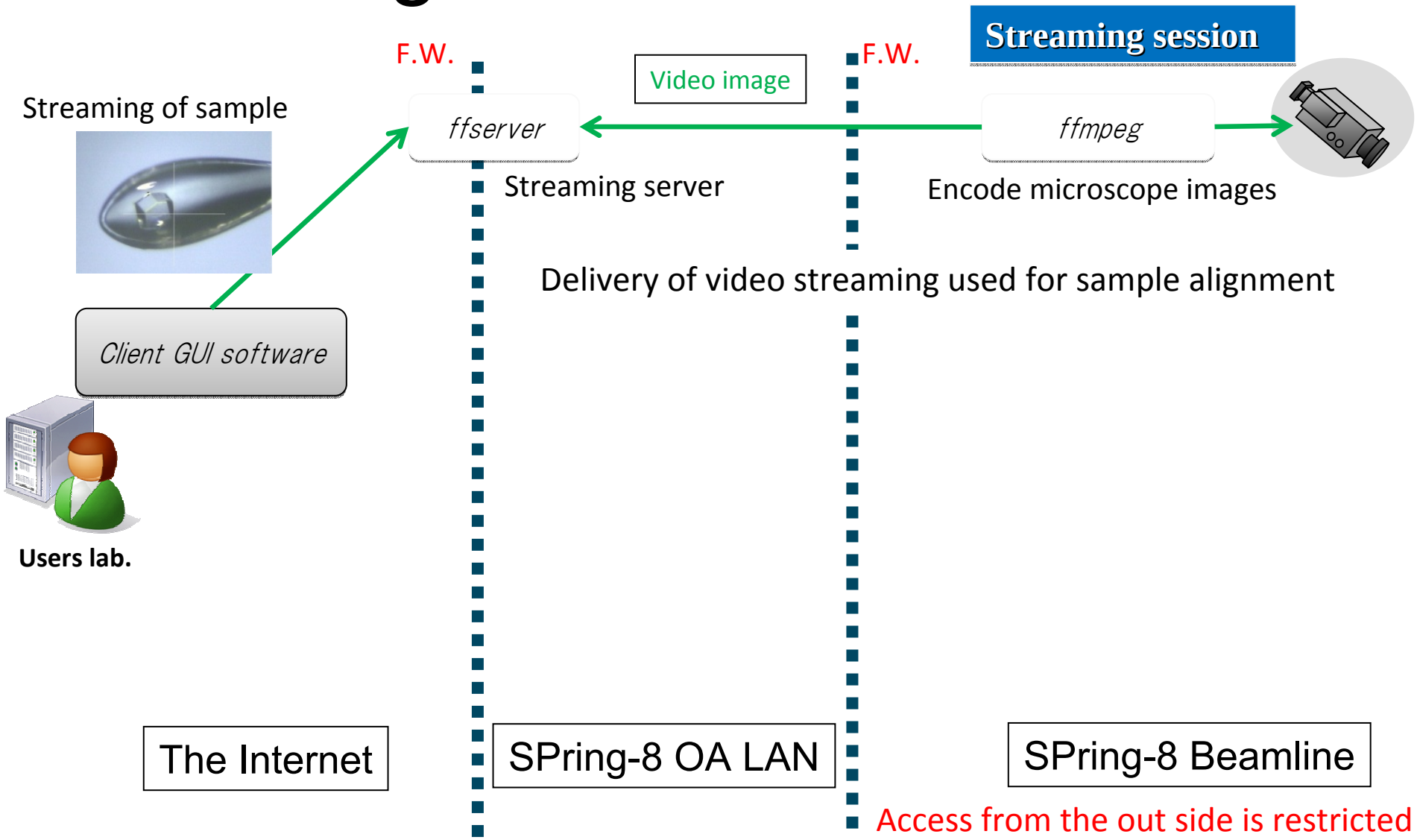
  For the future use other than MX BL…

  Remote access composed of three session (network connection)

  1. Streaming session
  2. Device control session <<<
  3. Result view session

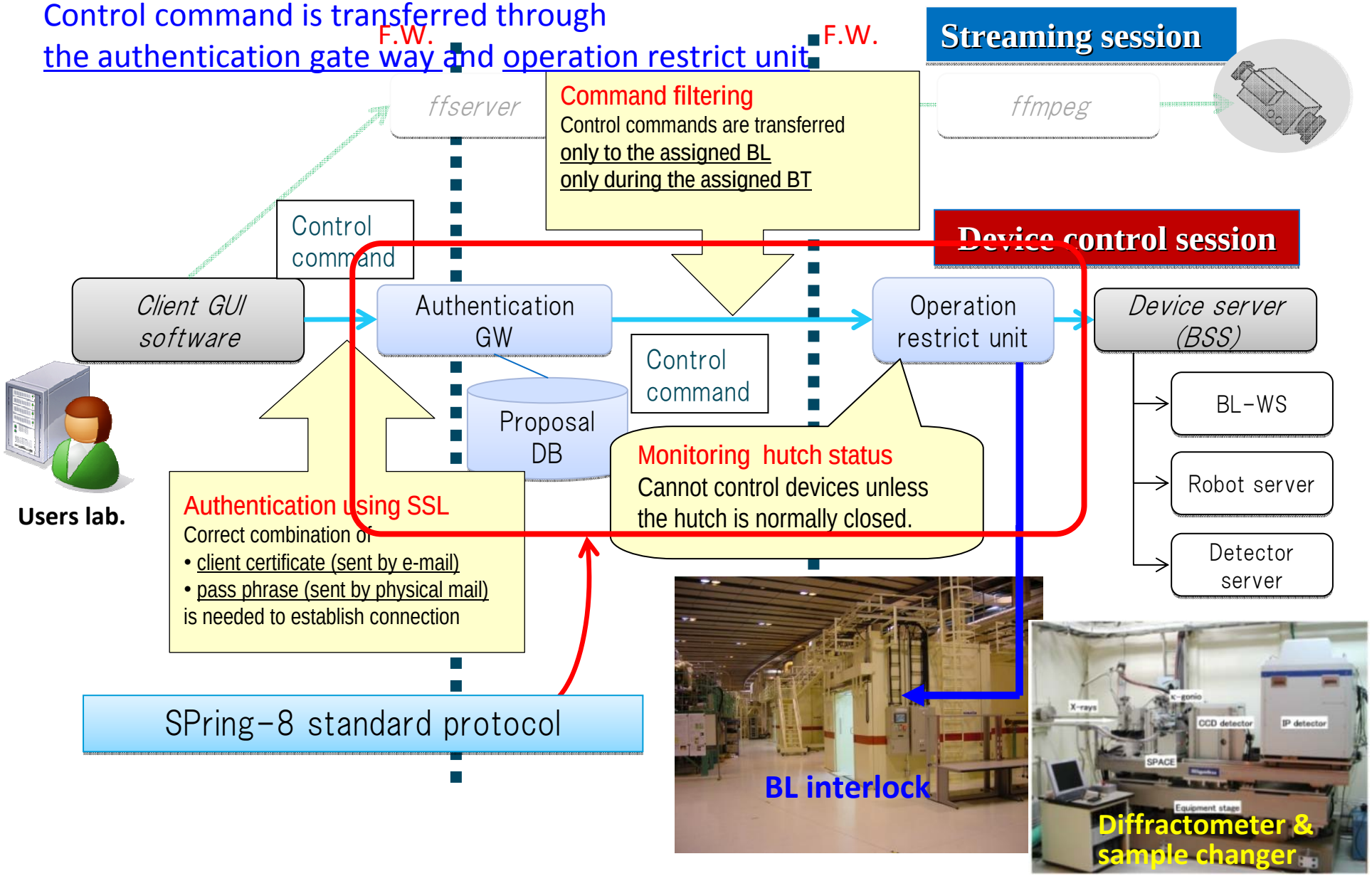
Hasegawa, Ueno, Furukawa et al. (2010)

Streaming session



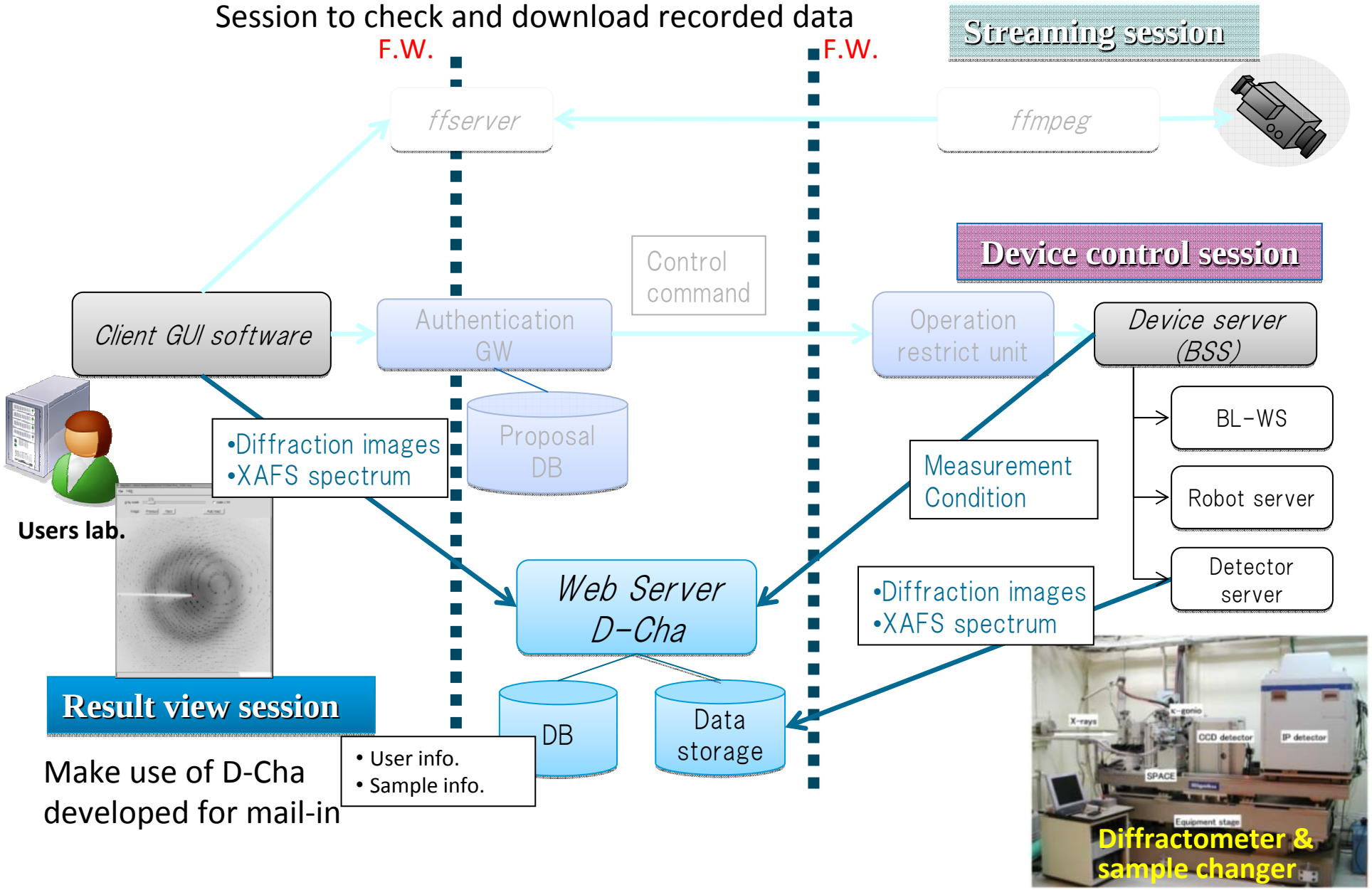
Device control session

Control command is transferred through
the authentication gate way and operation restrict unit



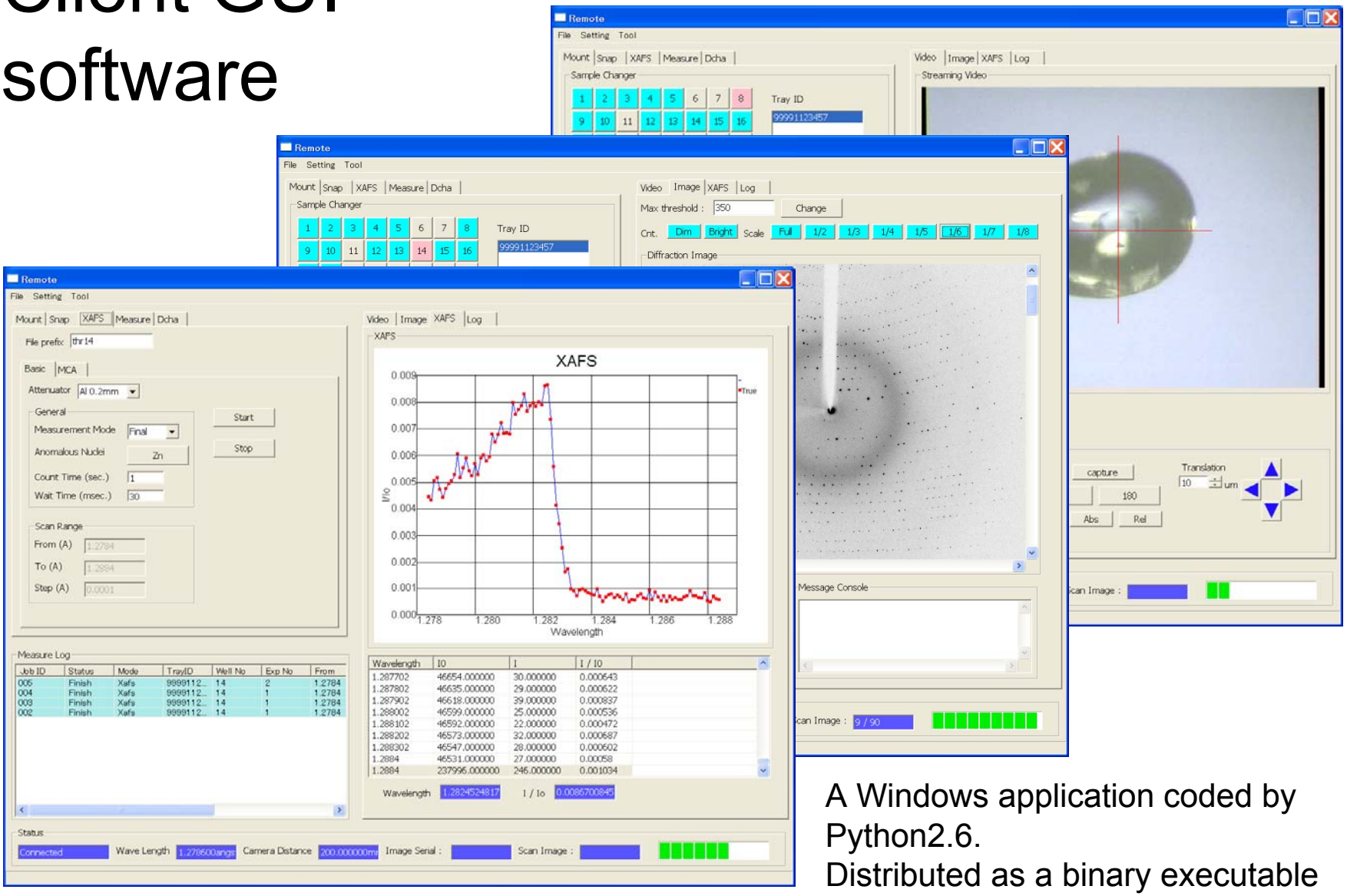
Result view session

Session to check and download recorded data



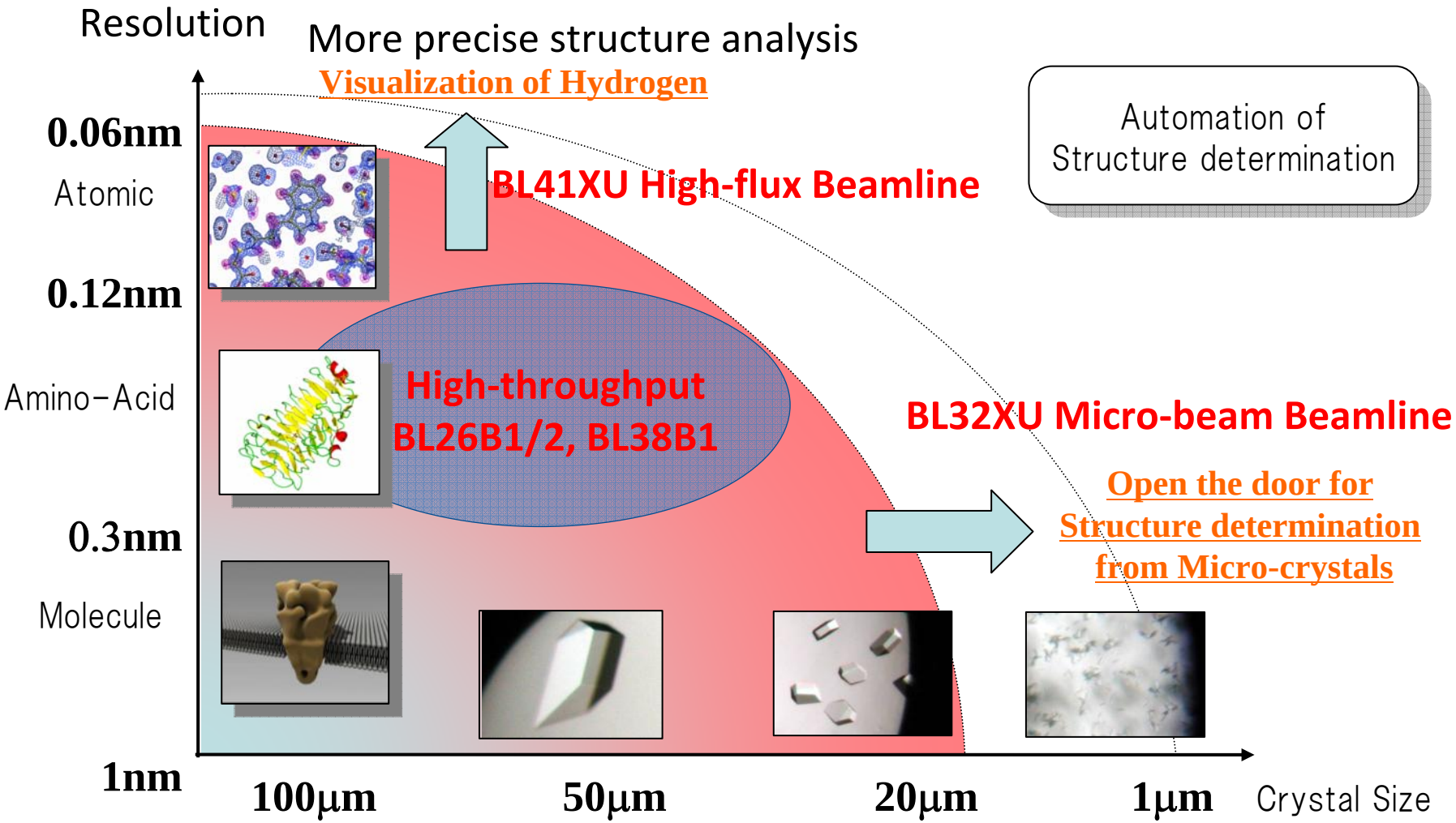
Make use of D-Cha developed for mail-in

Client GUI software



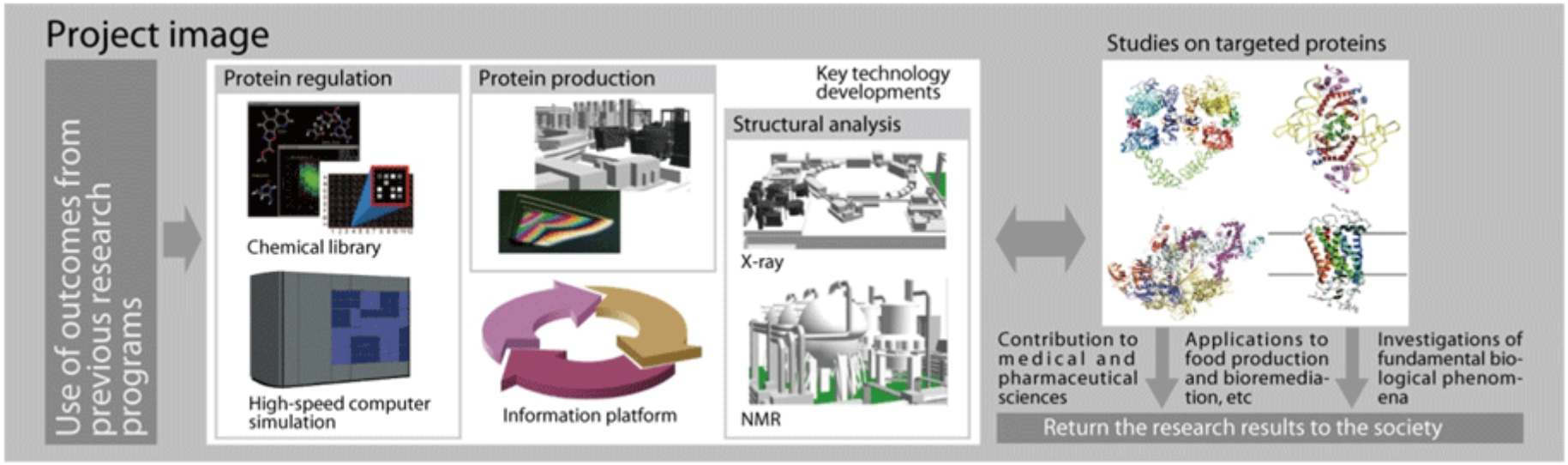
A Windows application coded by Python2.6.
Distributed as a binary executable

Get more structures and details



RIKEN Targeted Proteins beamline BL32XU
for Targeted Proteins Research Program (TPRP)

- What is TPRP ?
  - Grant: A national project promoted by MEXT, Japan
  - Aims: To reveal the structure and function of proteins that have great importance in both academic research and industrial application.
  - Research Themes:
    Targeted Proteins Research:
    Fundamental Biology / Medicine & Pharmacology / Food & Environment
    Technology Development:
    Protein Production / Structural Analysis / Chemical Regulation / Information Platform
- Beamline Construction
  - Kunio Hirata, Masaki Yamamoto et al. (RIKEN)

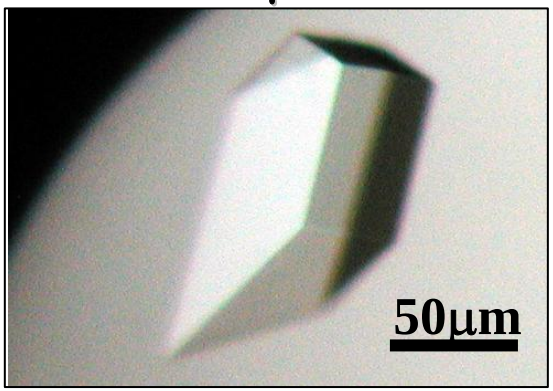


Development of micro-beam beamline

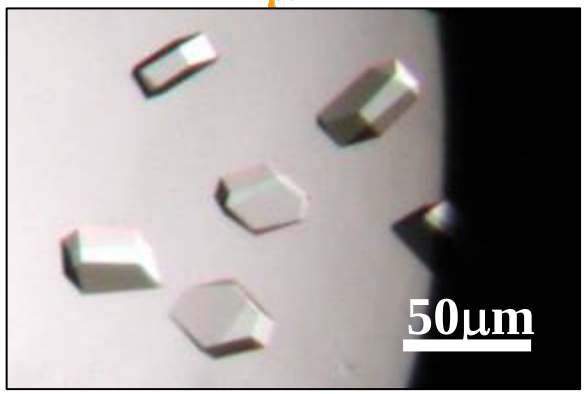
X-ray crystallography of proteins related to human disease and aging.

Micro-beam optimized for Micro-crystal

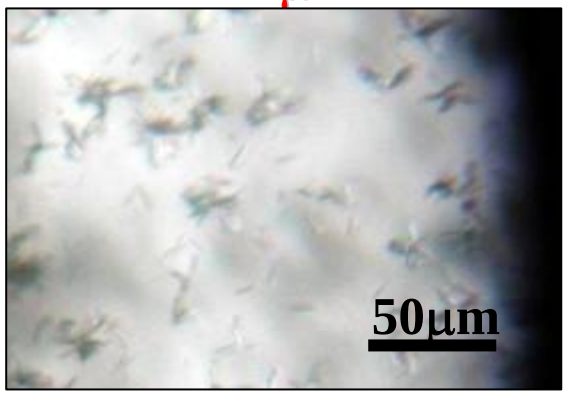
Standard
>50μm

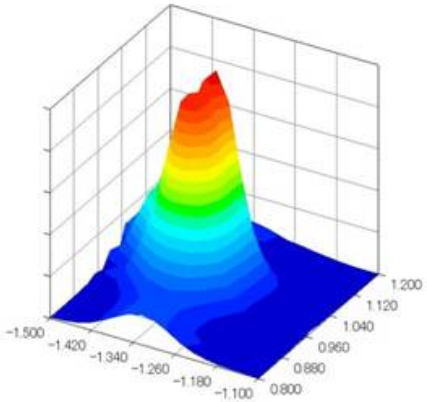


Current Limit
20~30μm



Micro-crystal
<10μm





Beam profile of SPring-8 BL41XU

Target Crystals

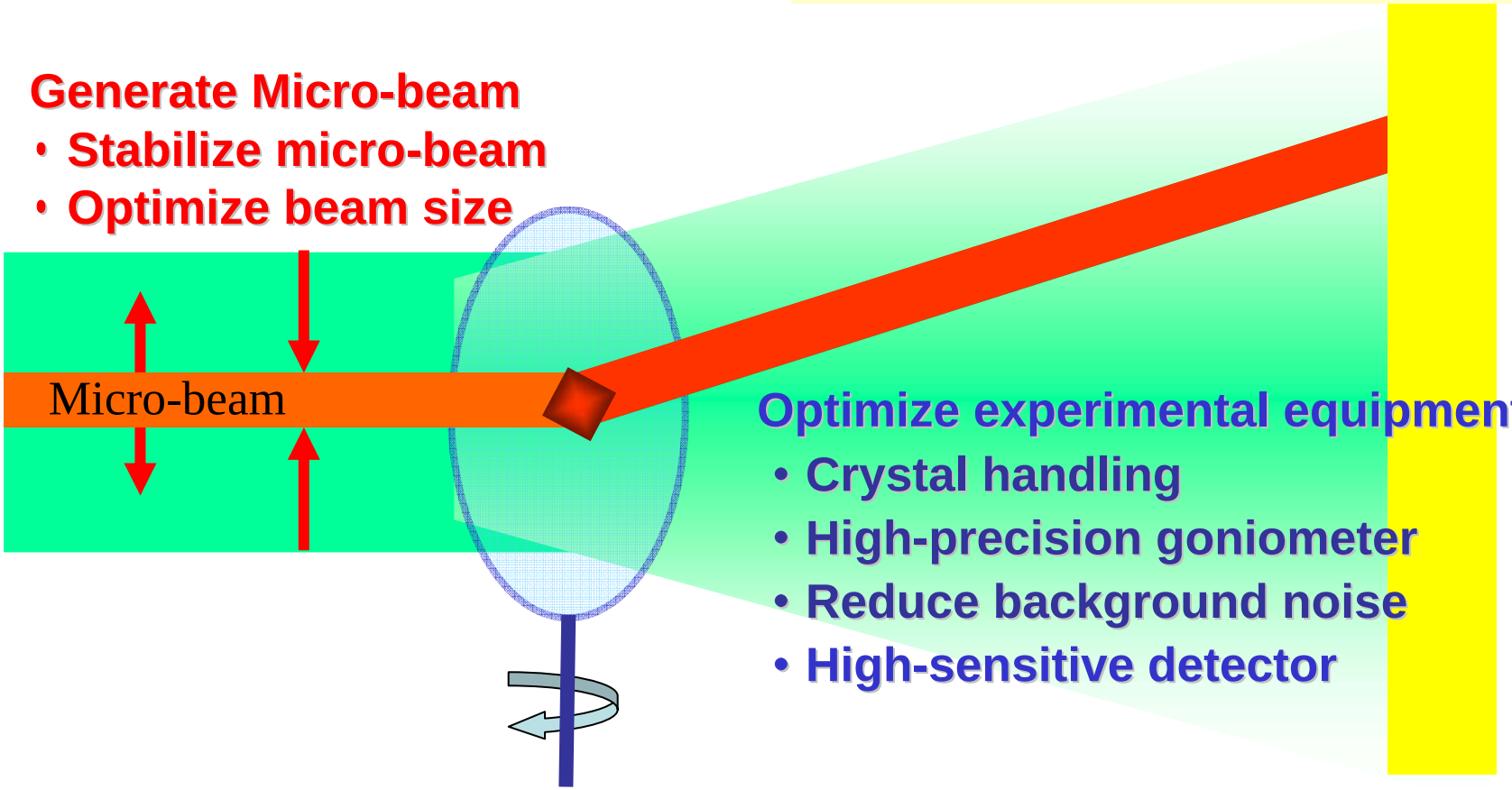
	Current	Target Beam Size	
• Beam Size	30×30	1×1	μm^2
• Flux density	10^9	$>10^{10}$	photons/sec./μm^2

R&D target for Micro-crystallography

Micro-crystal
- **Small size crystal（<10mm）**
- **Weak signal（10^6copies）**

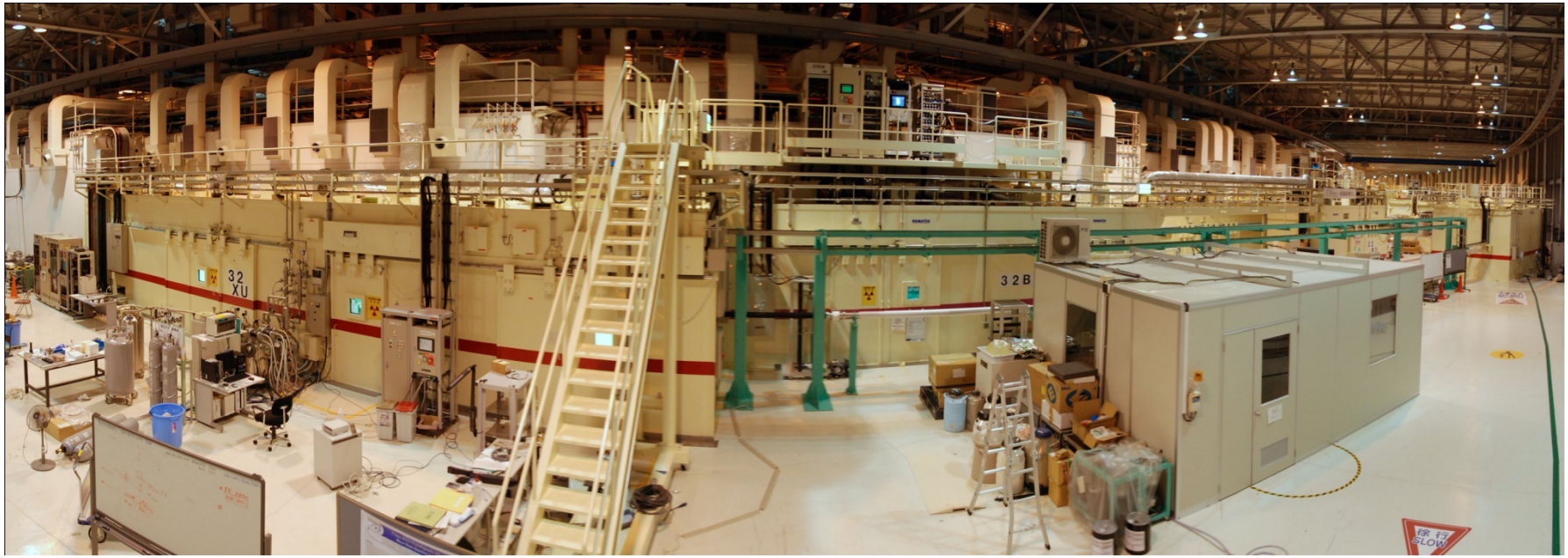
Maximize signal-to-noise ratio
- **Generate micro-beam**
- **Optimize experimental equipments**

Generate Micro-beam
- **Stabilize micro-beam**
- **Optimize beam size**



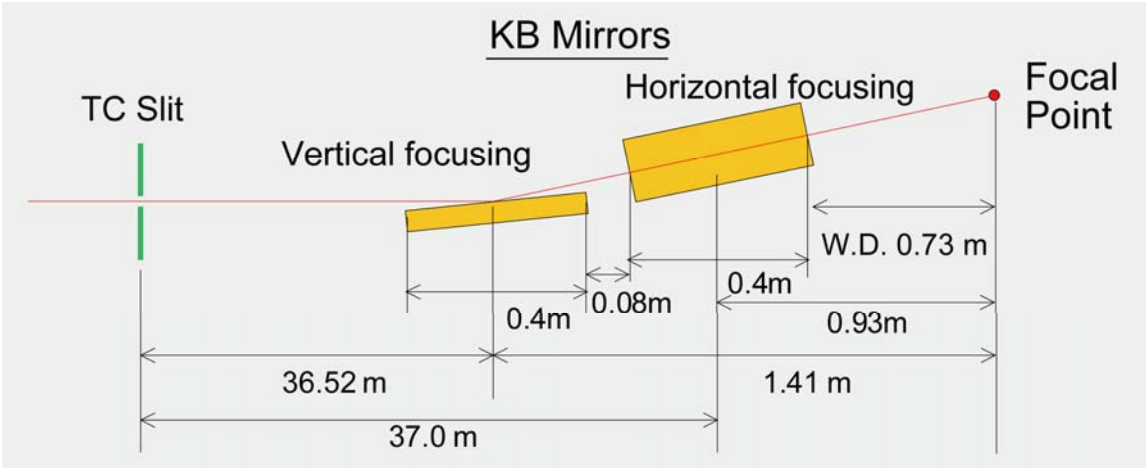
Optimize experimental equipments
- **Crystal handling**
- **High-precision goniometer**
- **Reduce background noise**
- **High-sensitive detector**

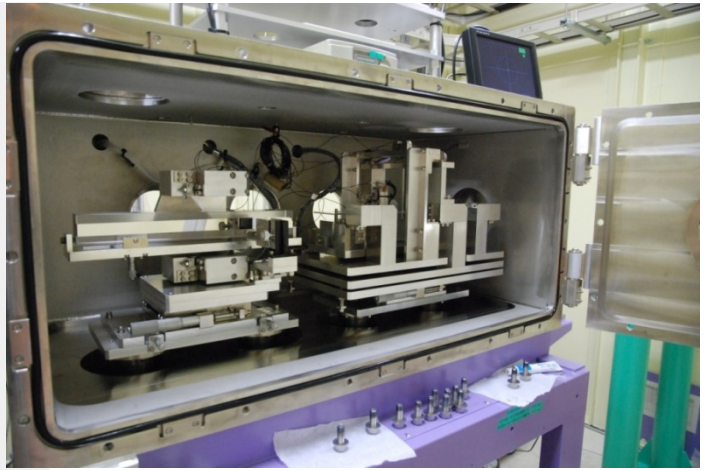
Design concept of BL32XU



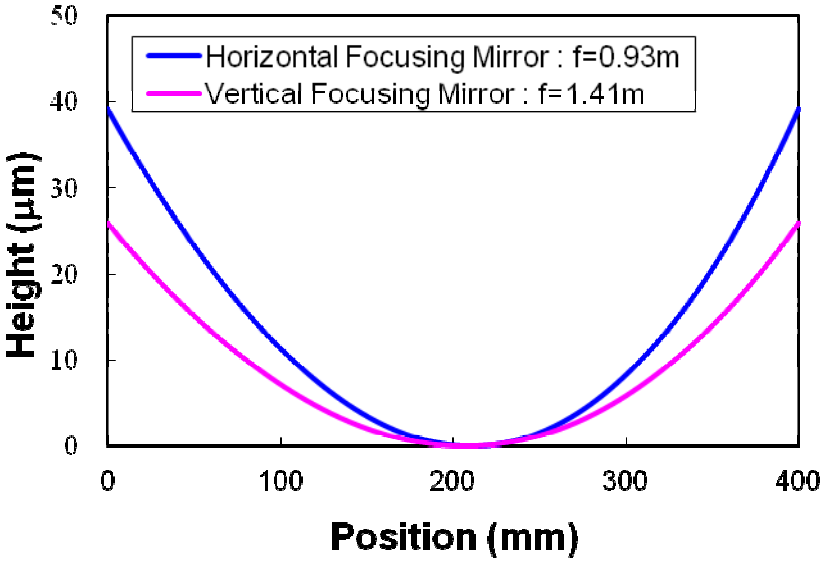
1. Brilliant source
2. Simple components
3. Focusing X-rays with large magnification factor
4. Changeable beam size at sample position

EEM-mirrors for 1 um focusing





Designed mirror surface shape

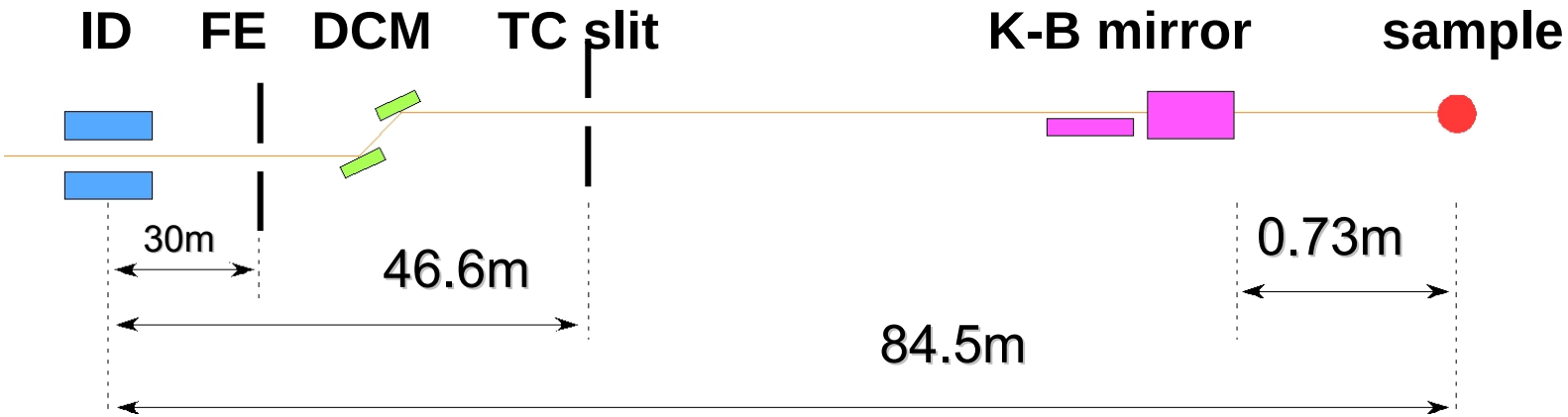


Kirkpatrick-Baez Mirror

Mirror shape :	Elliptical
Mirror length :	400 mm
Energy range :	8-20 keV
Mirror material :	SiO_2
Mirror surface :	Pt-coated
Glancing angle :	3.5mrad

Design of focusing optics

- Virtual light source is TC-Slit (located at 36m upstream of 1st mirror)
- Pt-coated elliptical mirrors with K-B (Kirkpatrick-Baez) configuration
- Magnification factors: 26 in vertical, 40 in horizontal
- Beam divergence at sample position < 2 mrad
- Available X-ray energy range: 8 - 20 keV, especially high-flux at 12.4 -13.8keV

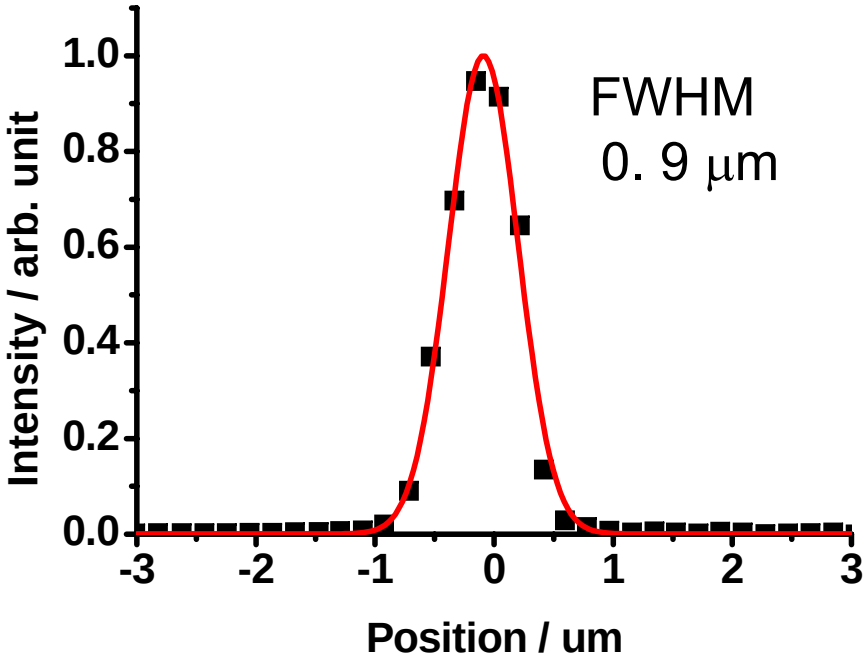


Beam size @ sample	1(H) x 1(V) μm^2	20(H) x 19(V) μm^2
TC slit size	40(H) x 26(V) μm^2	800(H) x 500(V) μm^2
Photon flux@12.4 keV	6×10^{10} photons/s	2×10^{13} photons/s

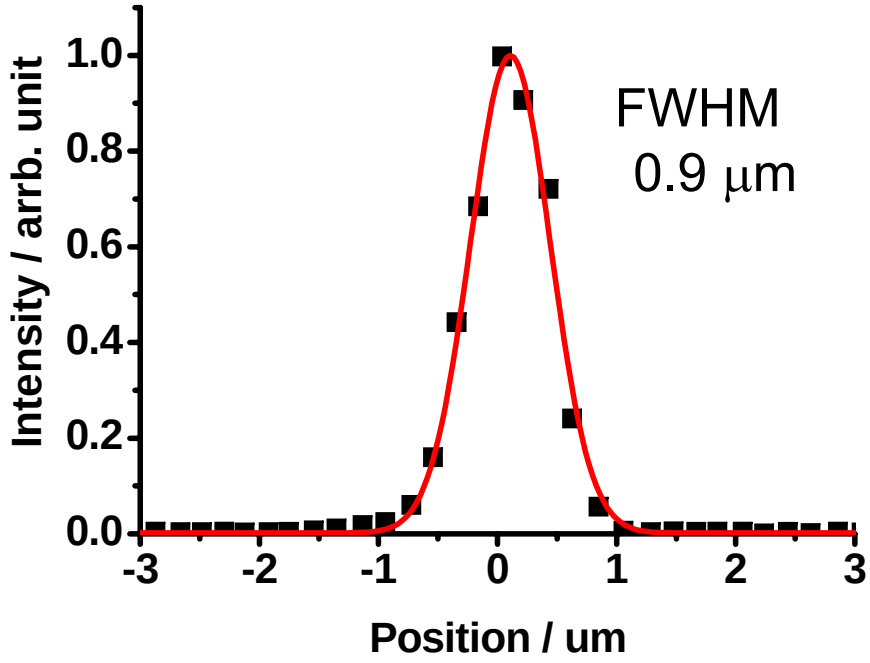
Glancing angle is designed at 3.5mrad

Achieved beam size（2009/11/27）

Horizontal beam profile



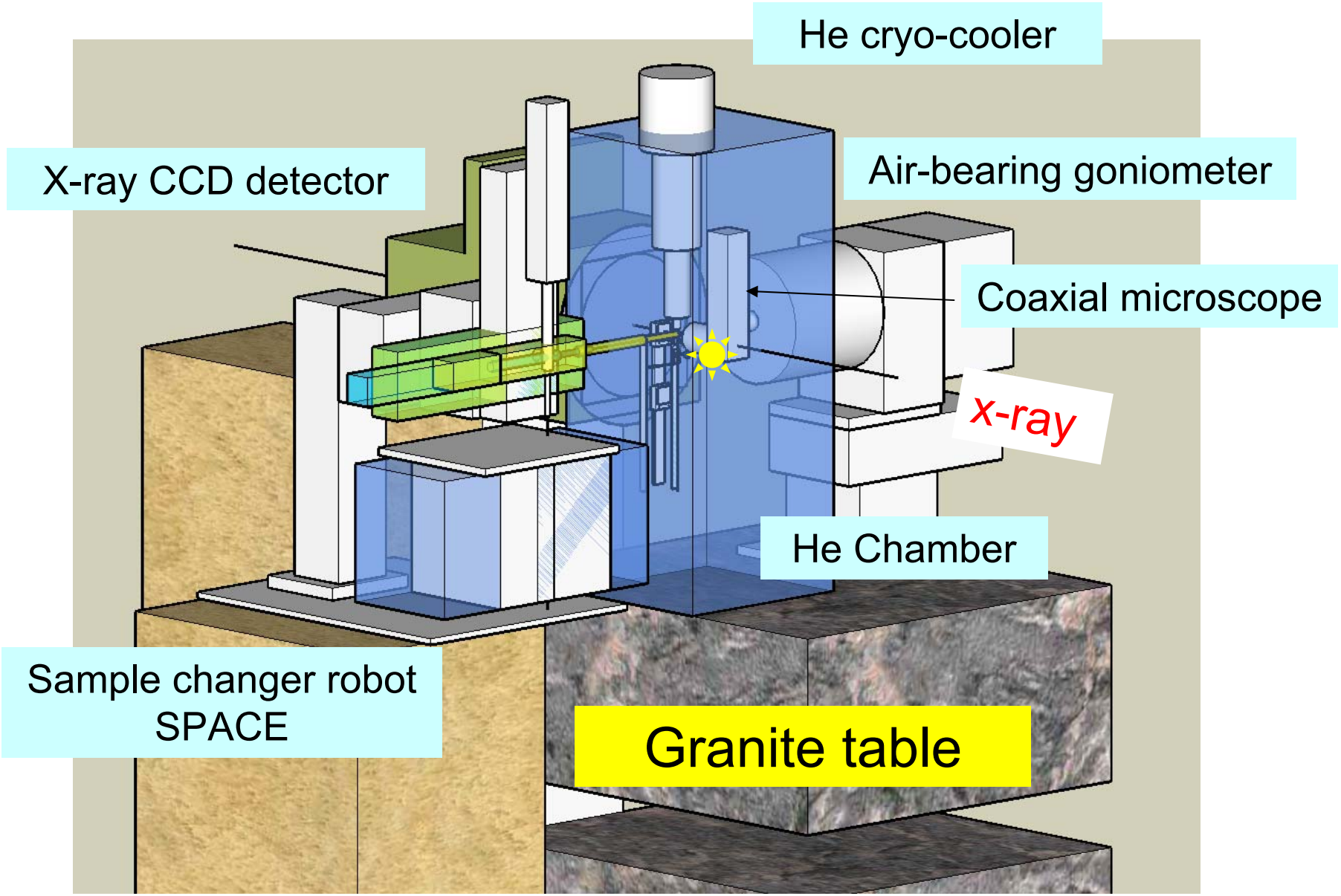
Vertical beam profile



Focused photon flux : 6.2×10^{10} photons/sec

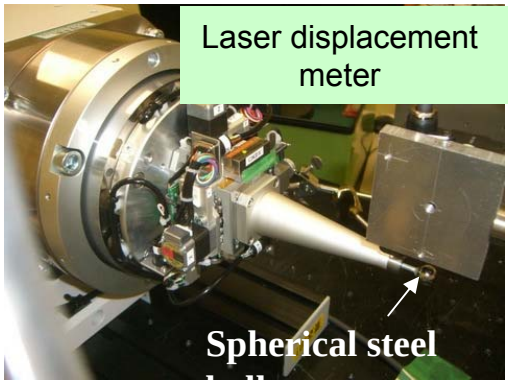
The smallest & highest flux density in the world

Micro-crystal diffractometer

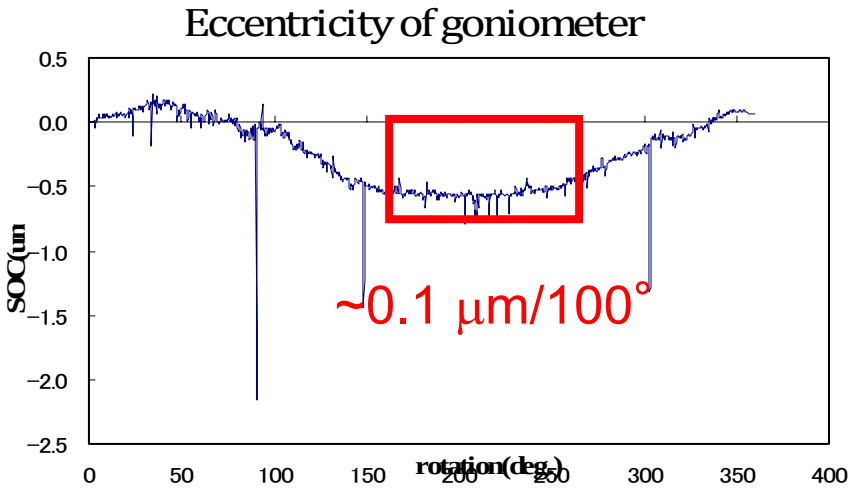


Air-bearing goniometer

- High-precision spindle axis with air-bearing unit
- Hi-speed rotation useful for fast centering, inverse beam geometry etc.

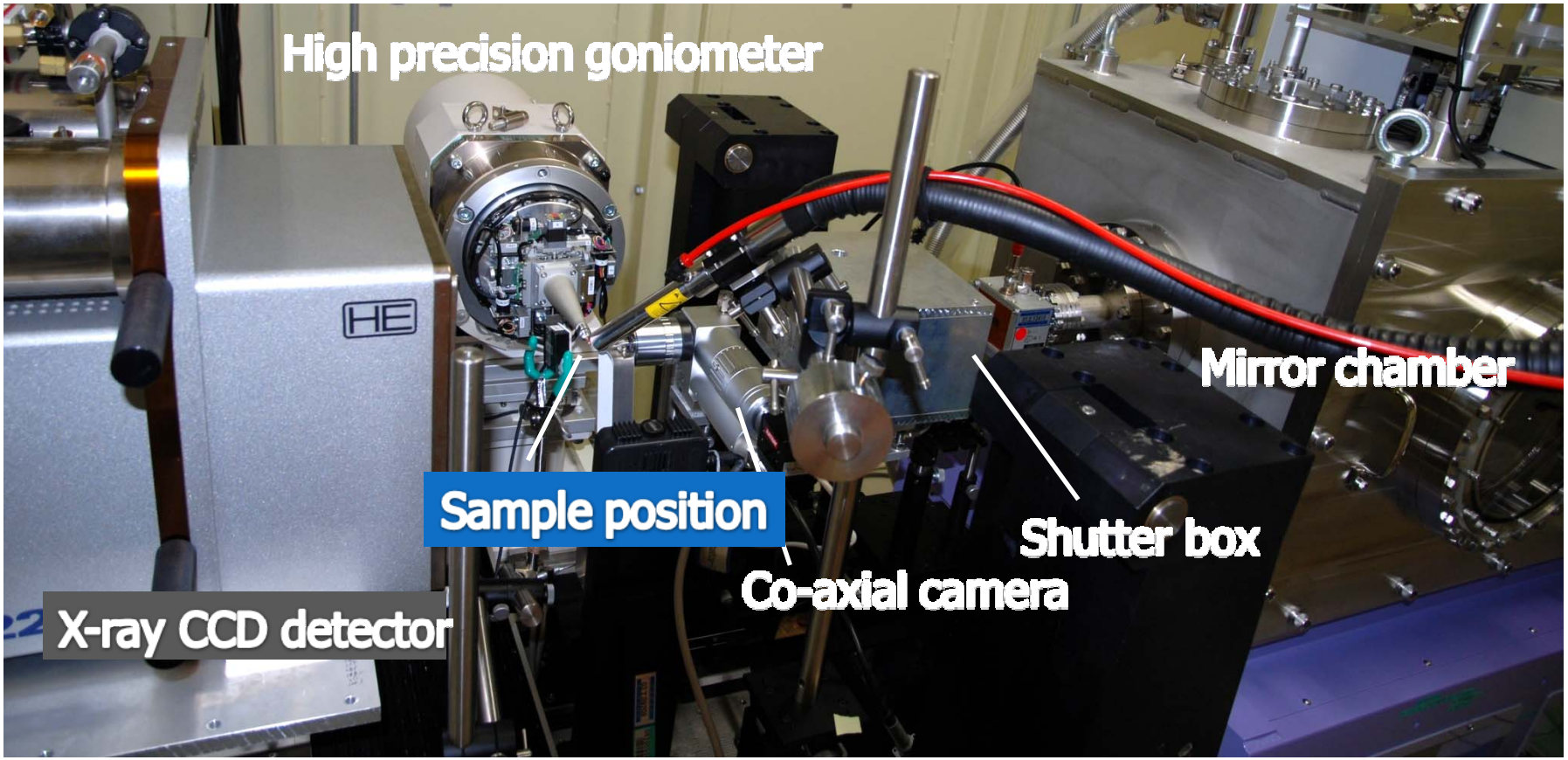


Eccentricity < 0.7μm/360°



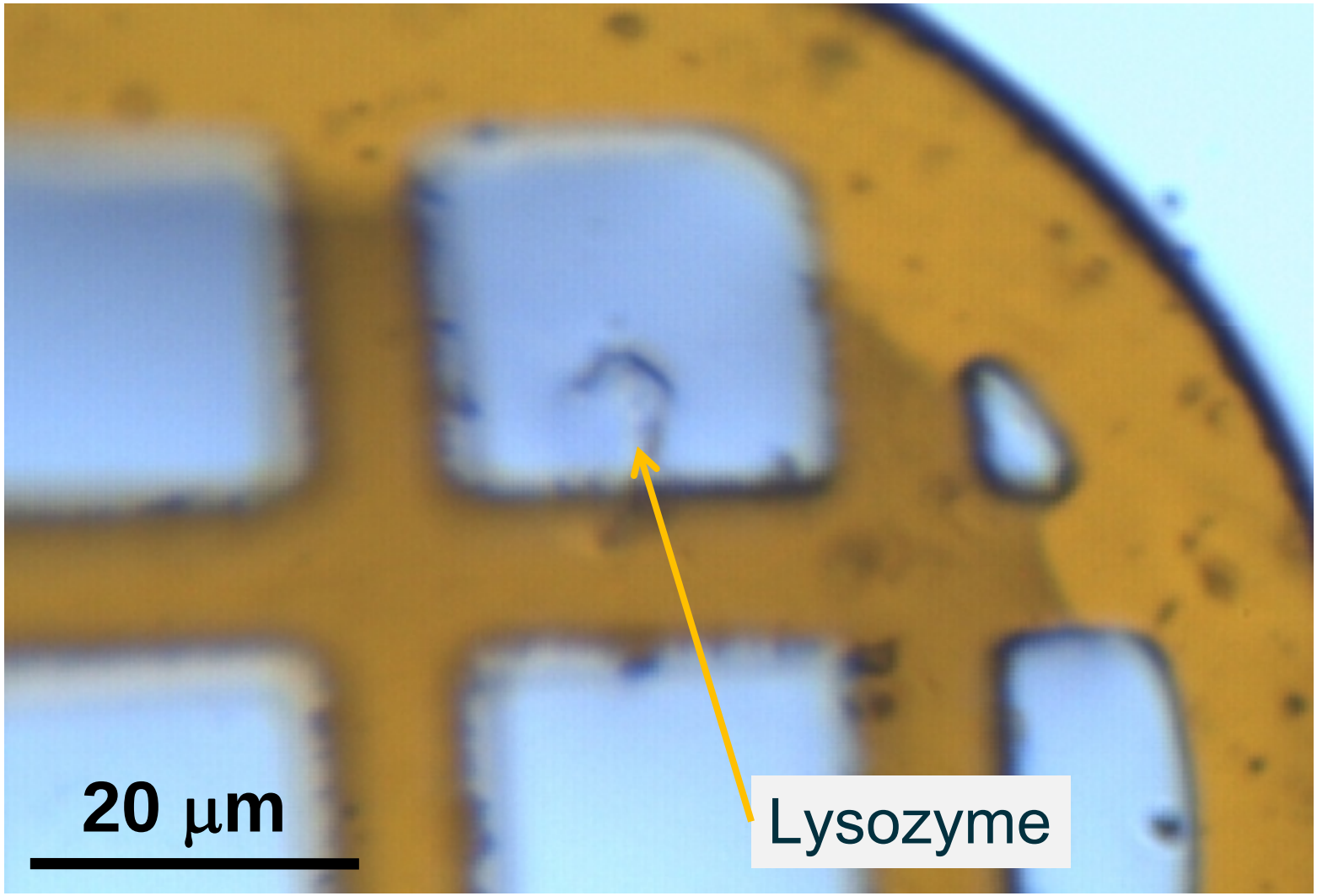
（KOHZU PRECISION Co., LTD.）

Tentative diffractometer setting



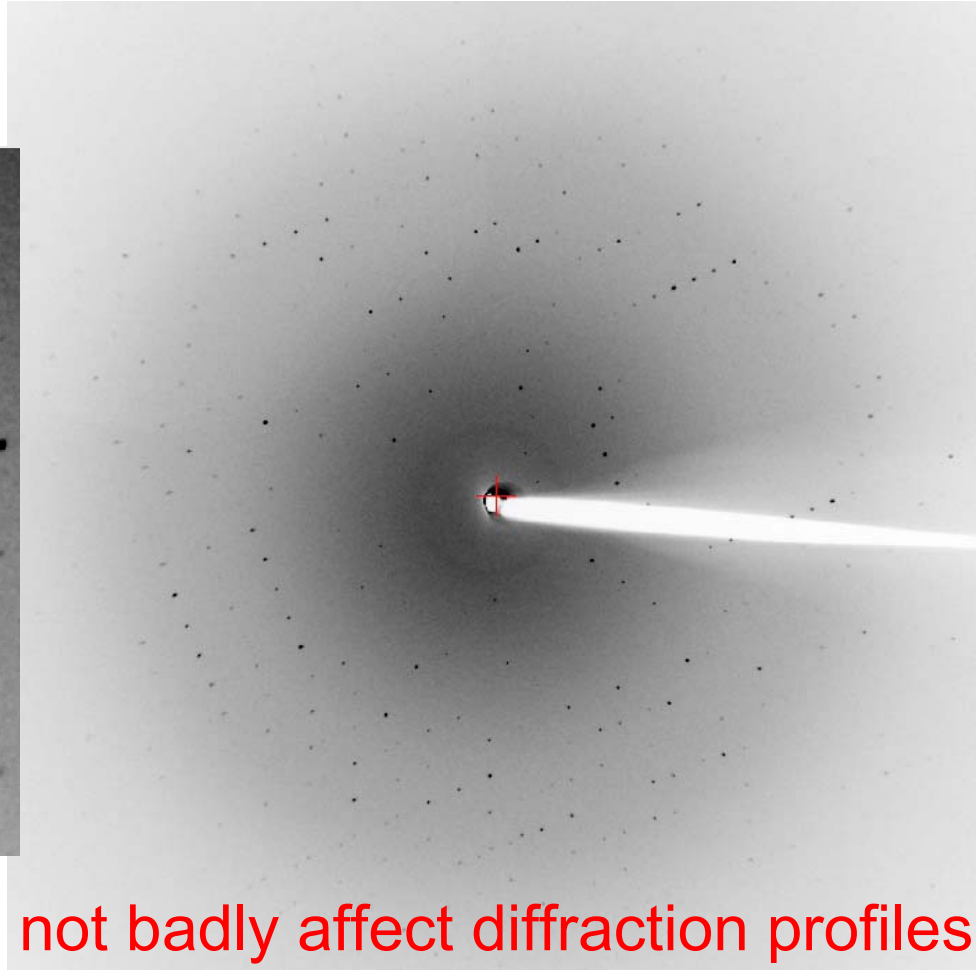
Focusing mirror -> Ion chamber -> Shutter -> Co-axial sample camera -> Collimator -> Back light -> Beam stopper

The first crystal onto the 1μm beam



The first diffraction image (09/12/04)

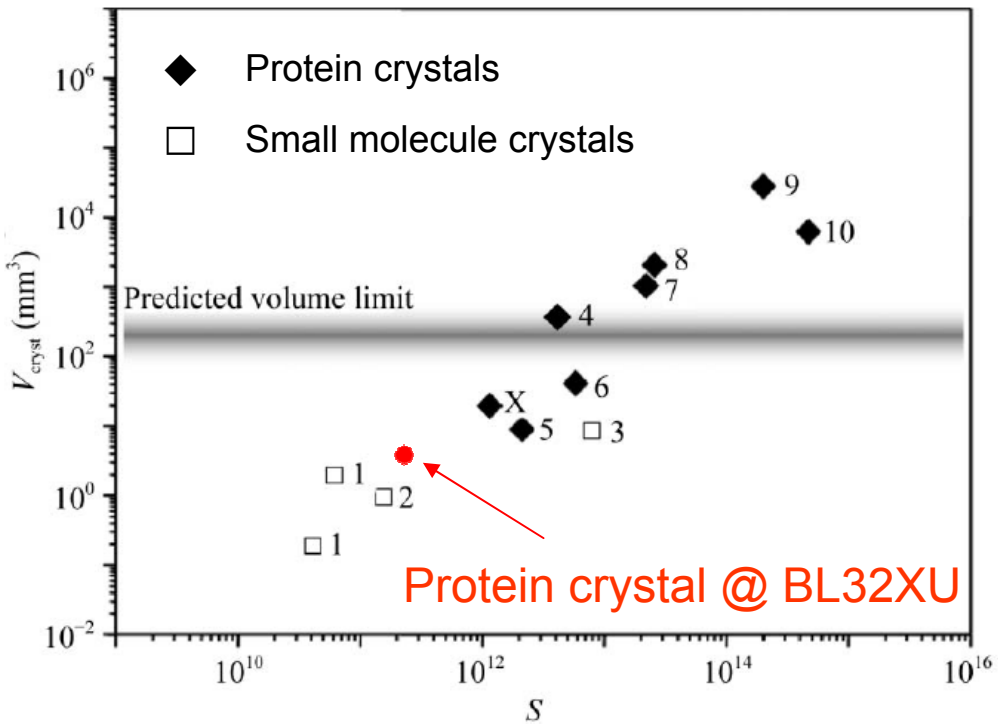
Crystal	Lysozyme 5um crystal
Beam property	1mm square, 2.6 x 10^{10} photons/sec.
Exposure time	1 sec.
Resolution limit	2.0 Å



Larger beam divergence did not badly affect diffraction profiles

Data collection limit by crystal size

Acta Cryst. (2008), D64, 158-166



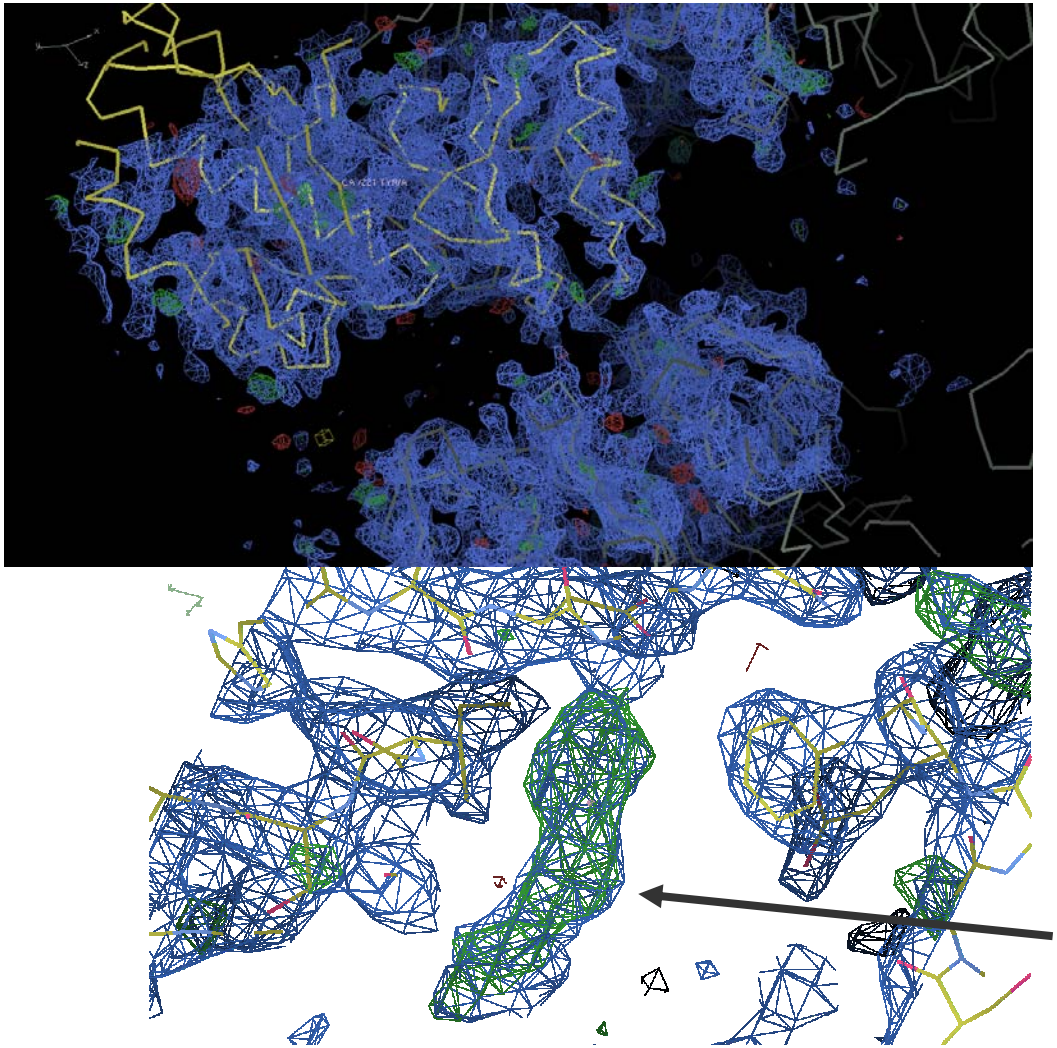
Formula of diffraction power

$$S = (F_{000} / V_{cell})^2 \times \lambda^3 \times V_{cryst}$$

We collect a 2 Å resolution data from 2 um lysozyme crystal.

BL32XU open the new field of Protein micro-crystallography

A recent result of a structure determination



2.6 Å $C222_1$,
1 degree x 90 images,
from a single crystal
(0.015 x 0.012 x 0.012 mm).
1 μ m beam, 1 sec exposures

Completeness=96.5%
(93.9%, 2.69-2.6A)

Apr. 2010

Drs. Sengoku and Bessho
(RIKEN)

Extra electron density
of a drug candidate.