Automatic data collection at SPring-8/Japan

Kunio Hirata

RIKEN/SPring-8

Macromolecular Crystallography (MX) Beamlines at SPring-8

Bending magnet beamlines BL26B1/B2

Developed for structural genomic research in 2000s

Automation at SP8 MX started from here

- Flux ~10^{11} (ph/s)
- Data collection from well diffracting crystal

Undulator BL, BL45XU

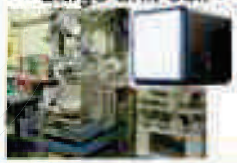
Automated data collection, High flux beam

- High flux 2 x 10^{13} ph/s @ 20 x 20 μm
- Fully automated data collection, ZOO system

Undulator BL, BL44XU

Low divergence beam

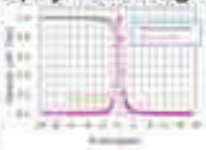
- Macromolecular complex
- Large unit cell > 500 Å

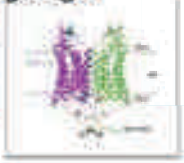


Undulator BL, BL32XU

Micro-beam minimum 1μm

- Micro-focus beam (1 μm; 6 x 10^{10} ph/s)
- Micro-crystallography / challenging target





Undulator BL, BL41XU

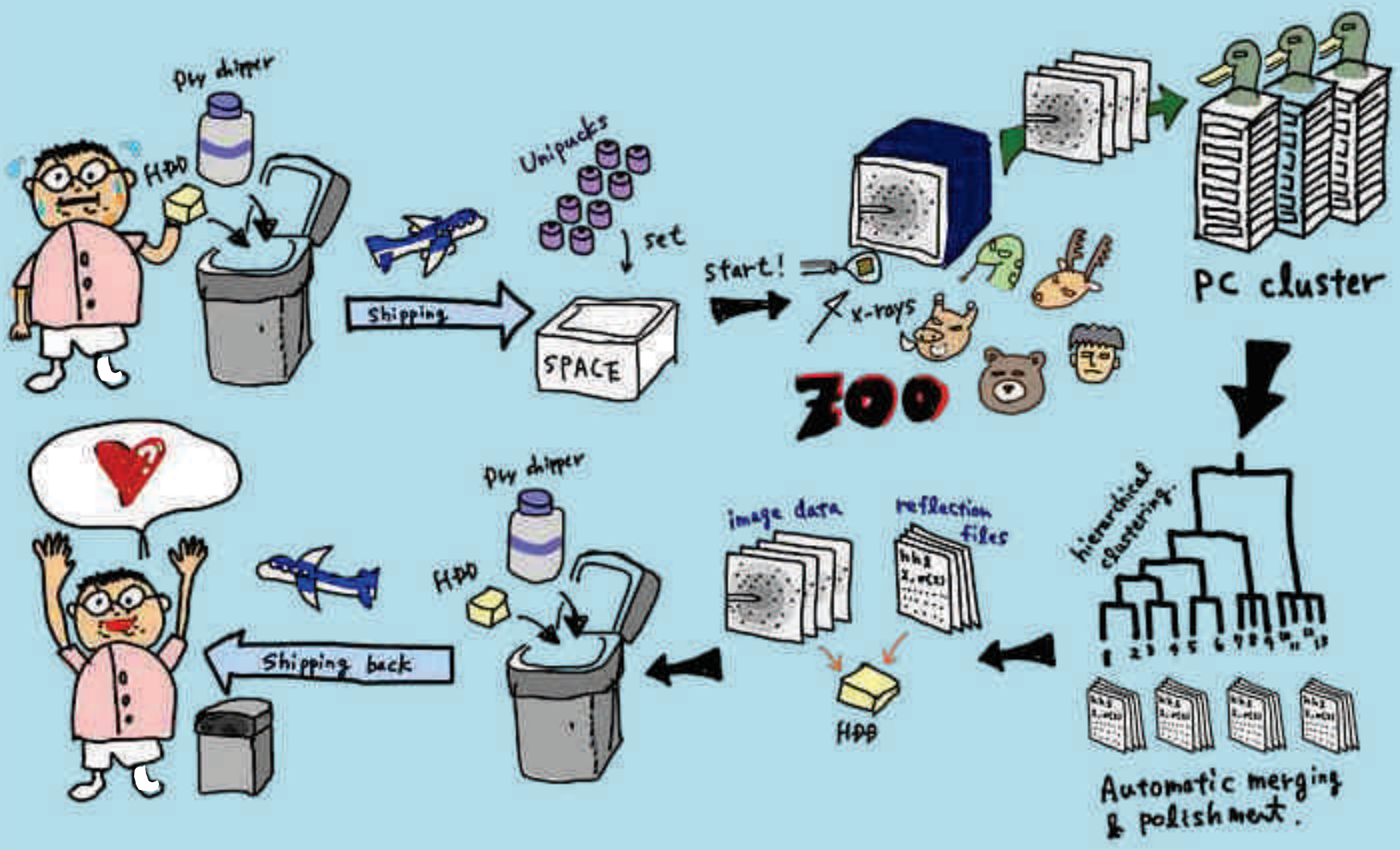
High flux beam

- High flux (4 x 10^{13} ph/s @ 2 x 20 μm)
- Ultra high resolution (λ ~ 0.5 Å, ID 3rd H)

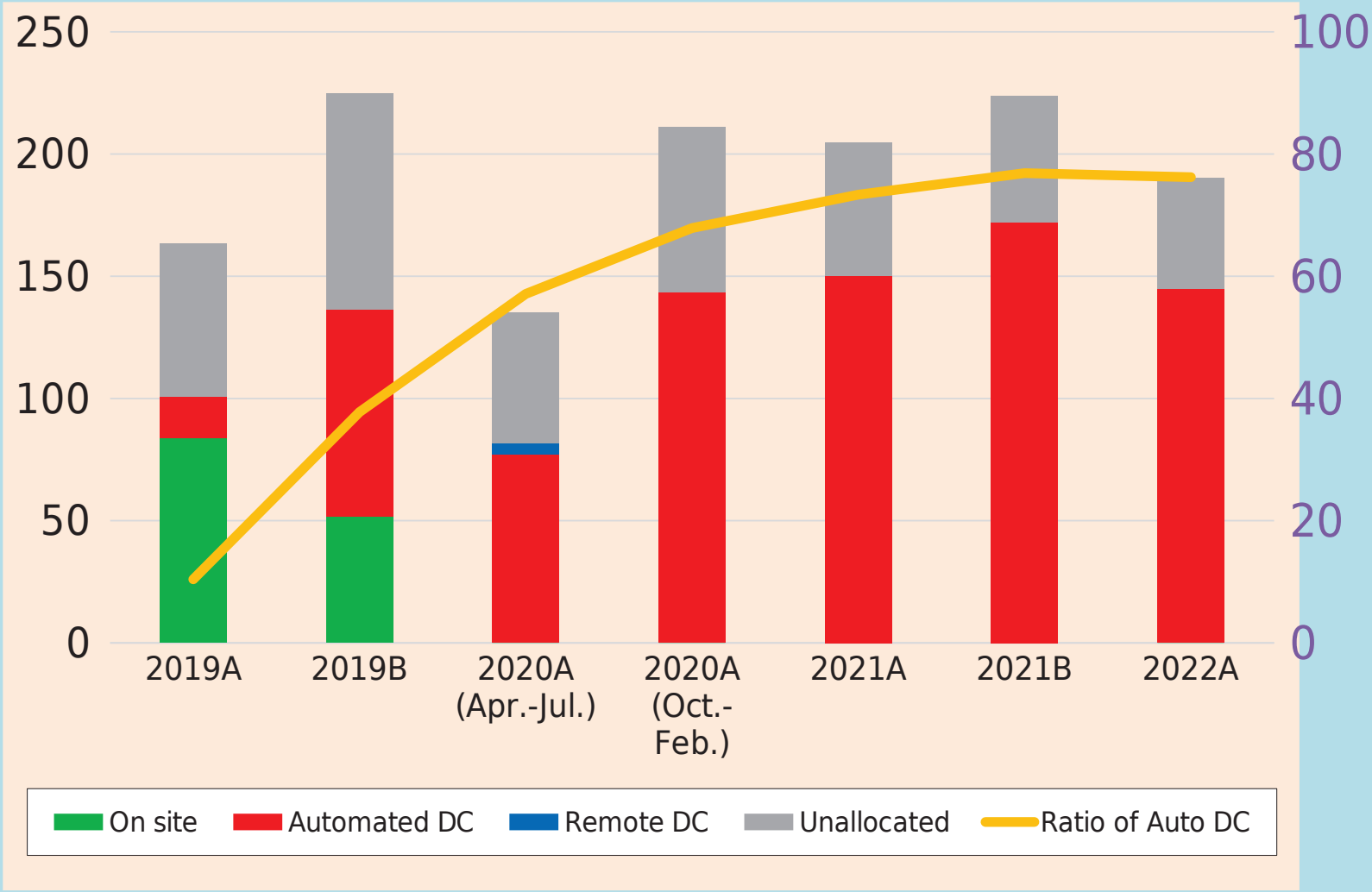
Operated by three organizations, but with common data collection environment, data collection software, sample changer etc.

ZOO system

Hirata, K., Yamashita, K. *et al*. (2019). *Acta Cryst (2019). D75*, 1–13.



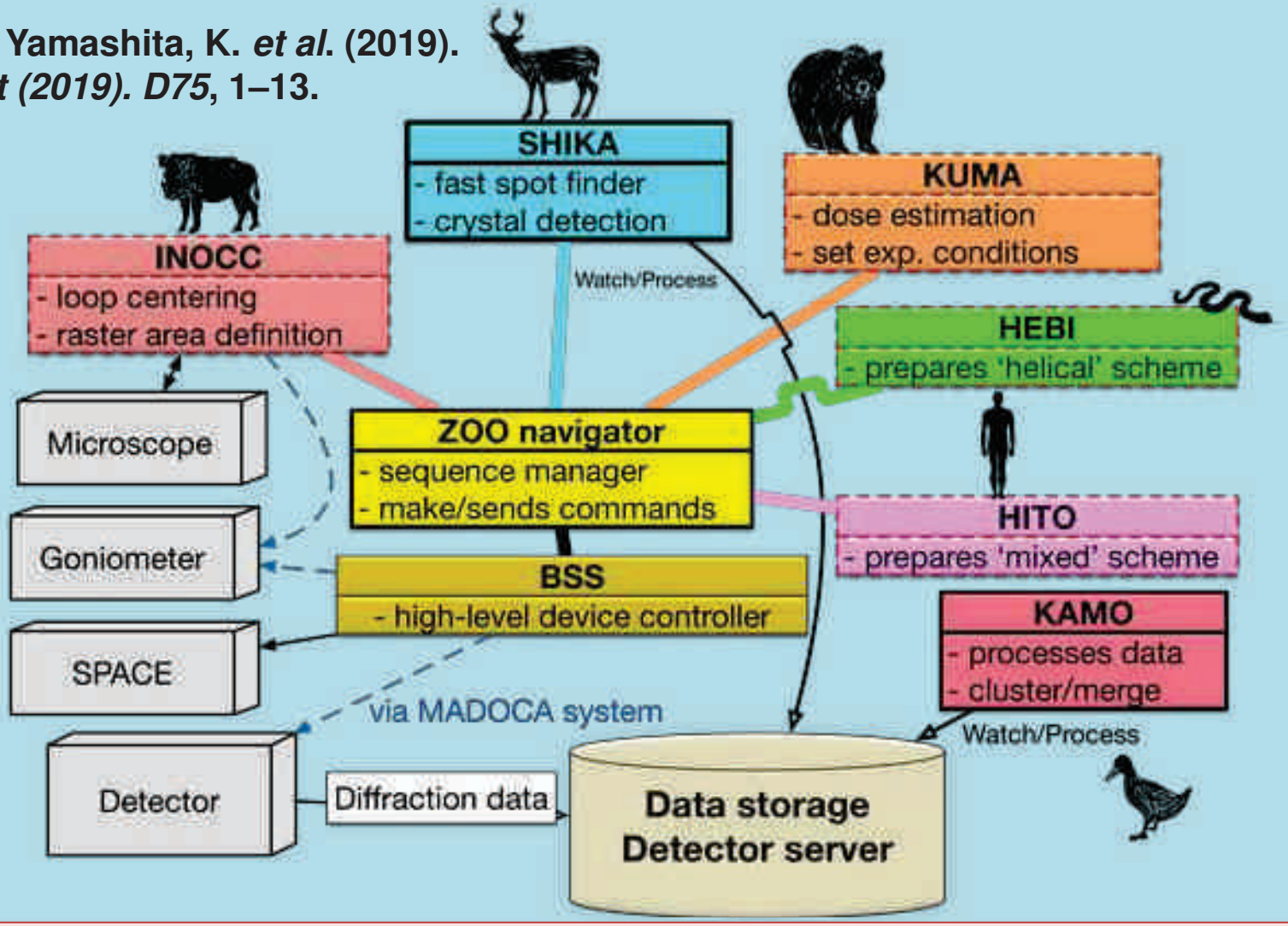
Transition to ‘unattended experiments’



80% public users conduct automatic data collection

ZOO system

Hirata, K., Yamashita, K. *et al*. (2019).
Acta Cryst (2019). D75, 1–13.



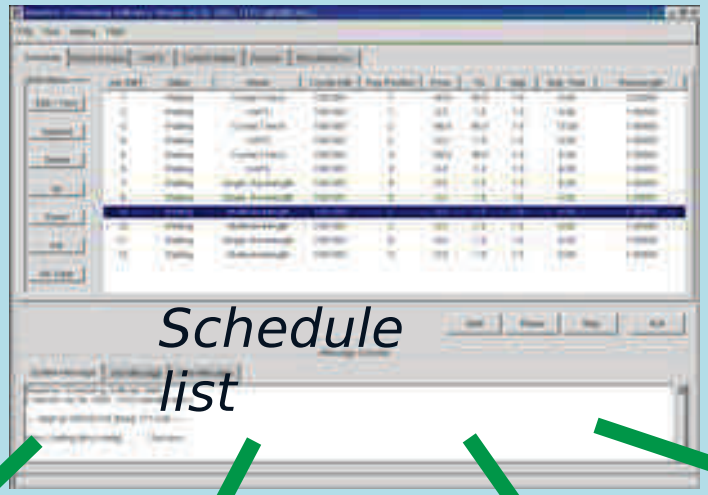
Program package with 'animal names' in

BSS (Beamline Scheduling Software)

Standard GUI for all MX beamlines at SPring-8

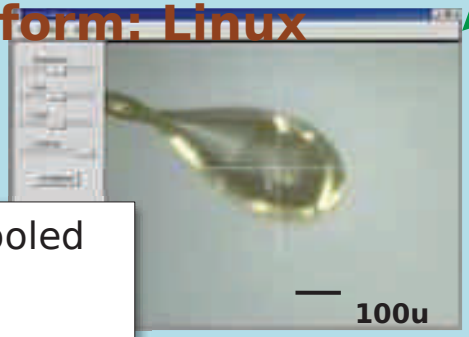
- **All-in-one control**
- **Job list for multiple conditions**
- **Load text BL configuration file**
- **Language & library: C, GTK+2, OpenGL, V4L2**
- **Platform: Linux**

Like MXCuBe?



Schedule list

Diffraction measurement



Cryo-cooled protein crystal

Sample Exchange, Centering

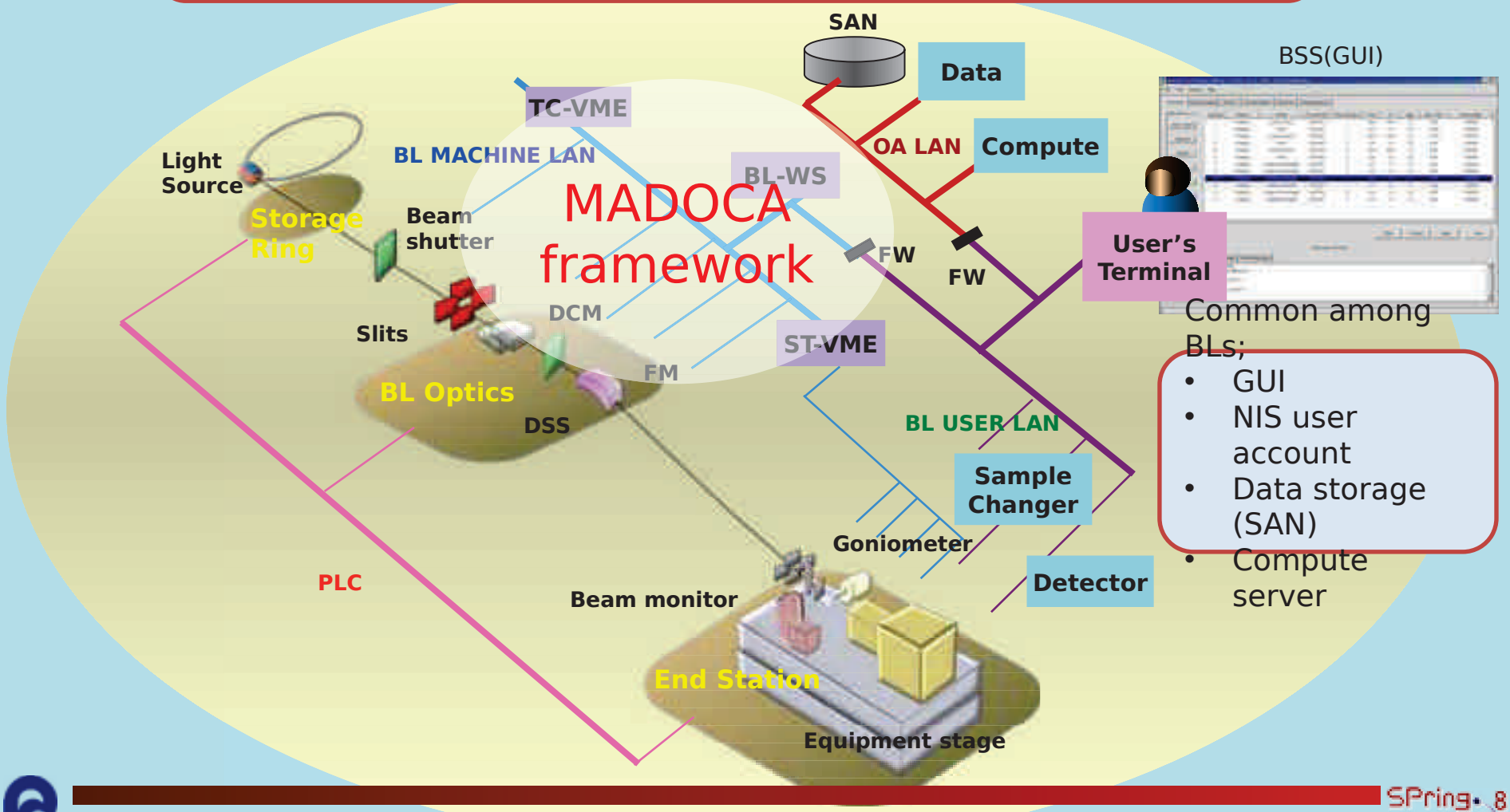
Exp. Conditions

XAFS measurement

G. Ueno et al., J. Synchrotron Rad. (2005). 12, 380-384

Control system for SPring-8 MX beamlines

- Unified beamline control with a GUI under C/S architecture
- Common storage & compute server on the network
- SR, BL, end station are seamlessly controlled via MADOCA

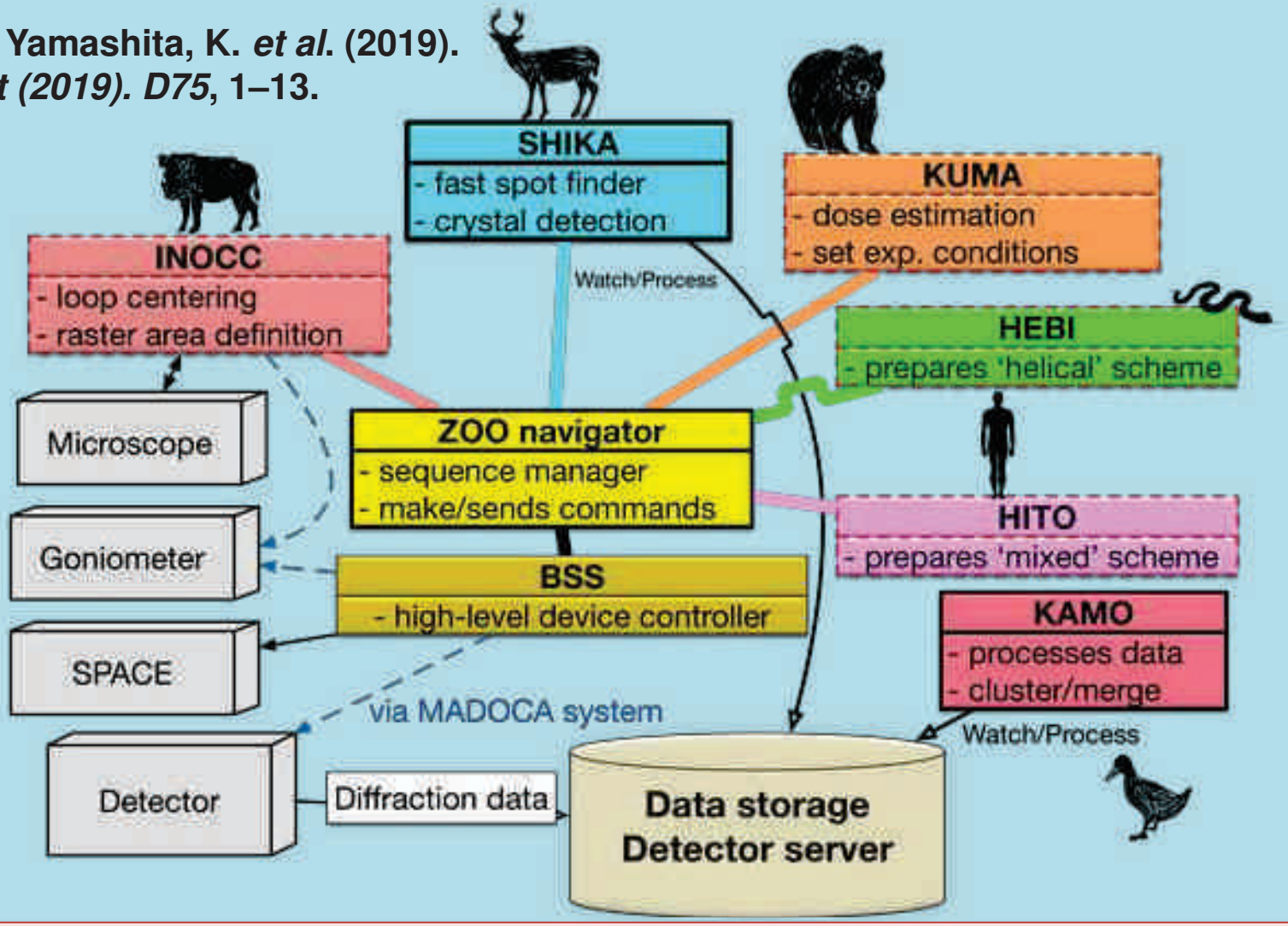


Common among BLs;

- GUI
- NIS user account
- Data storage (SAN)
- Compute server

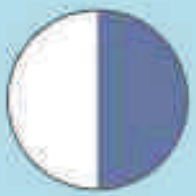
ZOO system

Hirata, K., Yamashita, K. *et al*. (2019).
Acta Cryst (2019). D75, 1–13.



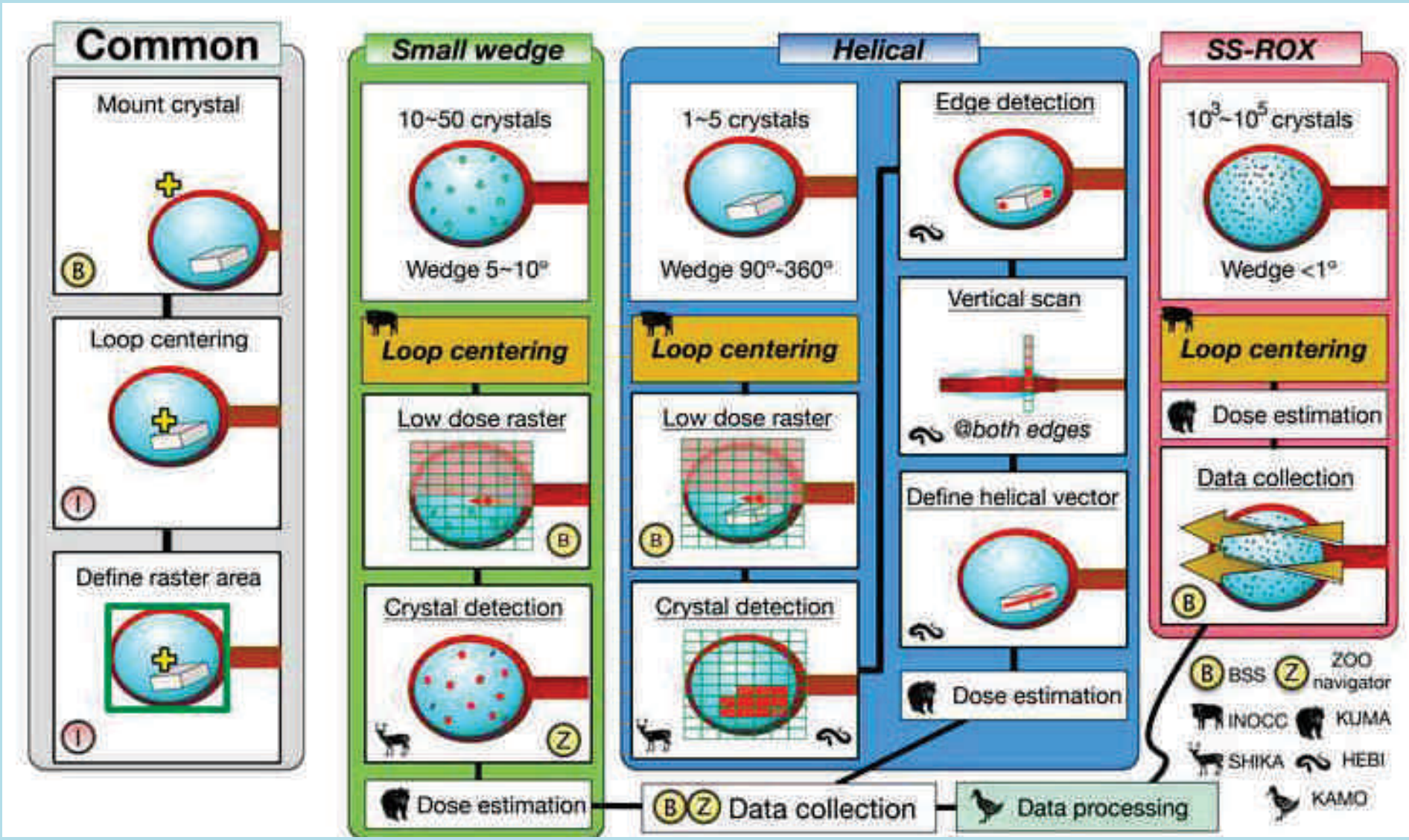
Program package with 'animal names' in

Available schemes in ZOO

Scheme		Rot. (/xtal)
”Small wedge” (SWSX)	 10–300 crystals	 5-10 deg.
“helical” or “single”	 single or a few crystals	 90-360 deg.
“mixed”	 small & large crystals	 5-10 deg. + 30-360 deg.
“SSROX”	 300-10,000 crystals	 < 1 deg.

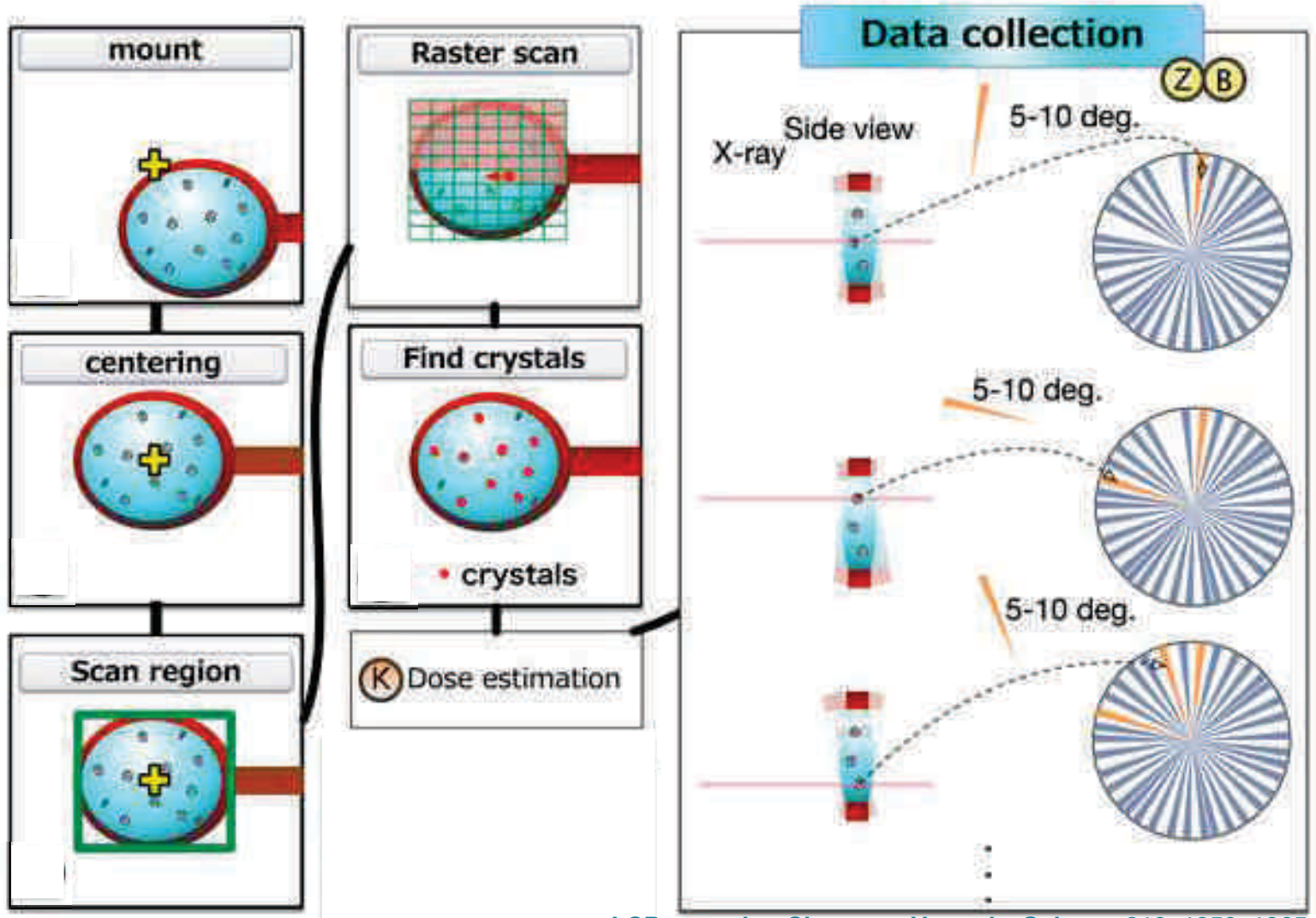
Users should choose ‘scheme’ for each cryo-pin.

Available schemes in ZOO



“Small Wedge Synchrotron Crystallography”

Many randomly oriented crystals on each loop



LCP crystals : Cherezov, V. *et al.* : Science 318, 1258–1265 (2007).

Many sub-datasets are merged into the final one in SWSX

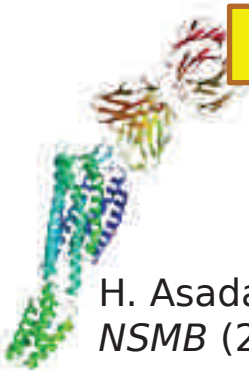
Membrane protein structure analysis with ZOO

Orexin 2 receptor
2.0 Å

251

R. Suno et al.
Structure(2018).

Angiotensin II receptor
3.2 Å



241

H. Asada et al.
NSMB (2018).

M2 muscarinic receptor
2.3Å

598
Hg-SAD

R. Suno *et al*.
Nat. Chem. Biol.(2018)

Human ETB receptor
2.0 Å

483

W. Shihoya et al.
Nat. Commun.(2018)

5-HT2A Receptor
2.7 Å

265

K. Kimura *et al*.
Nat. Chem. Biol.
(2019)

# of datasets

Prostaglandin E receptor(EP4)
3.2 Å

203

Y. Toyoda *et al*.
Nat. Chem. Biol. (2018)

Prostaglandin E subtype EP3
3.2 Å

683

K. Morimoto *et al*.
Nat. Chem. Biol. (2018)

β2AR regulation
3.2 Å

125

X. Liu et al.
Science(2019)

Melanocortin-4 receptor
2.8 Å

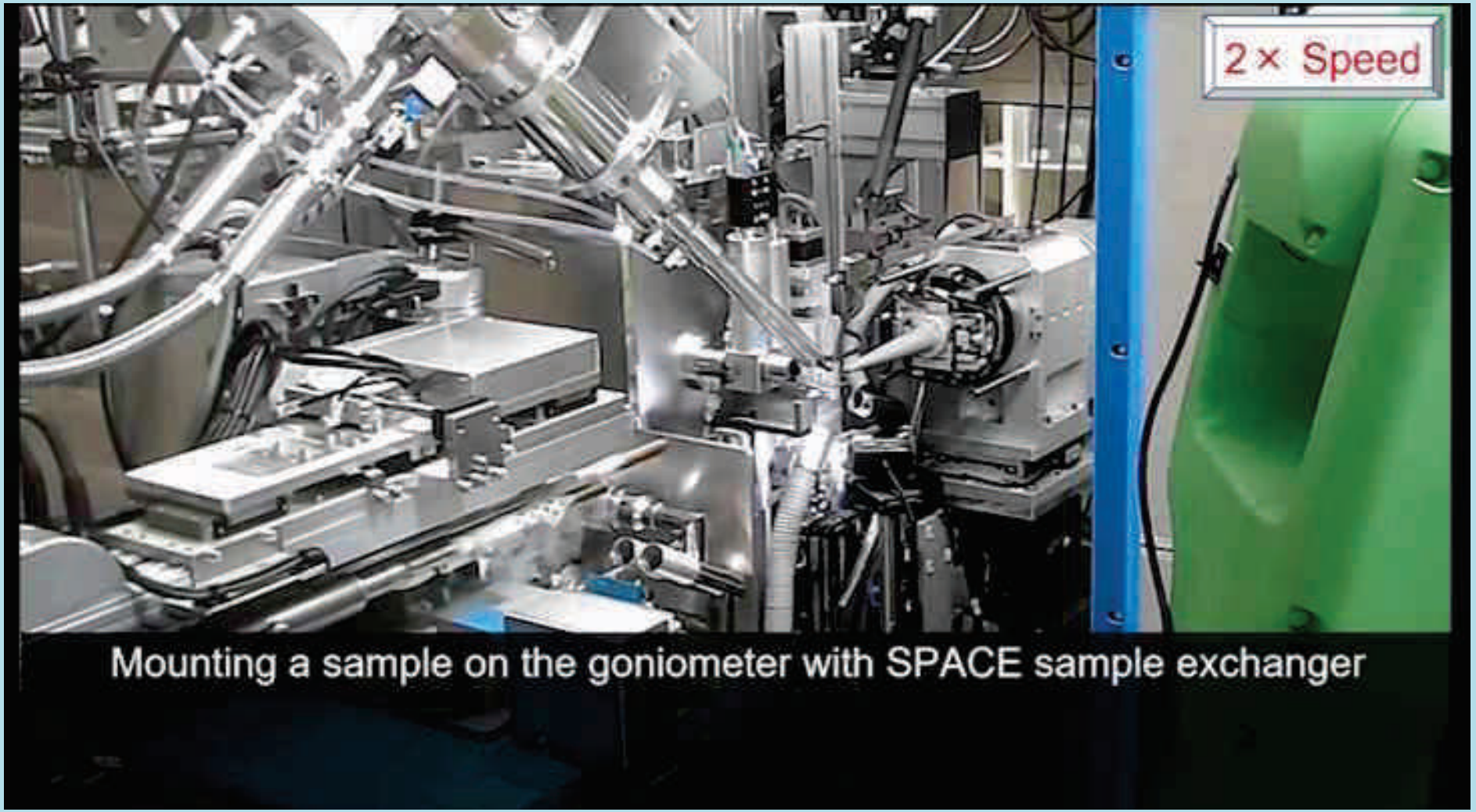
83

J. Yu et al.
Science(2020)

“Quantitative changes” make “Qualitative

Helical data collection

Movie at BL45XU



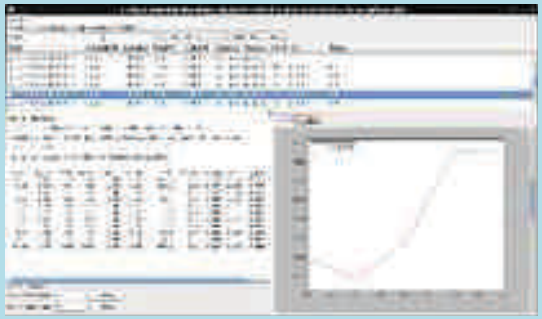
Keitaro Yamashita

KAMO

(Automatic data processing)

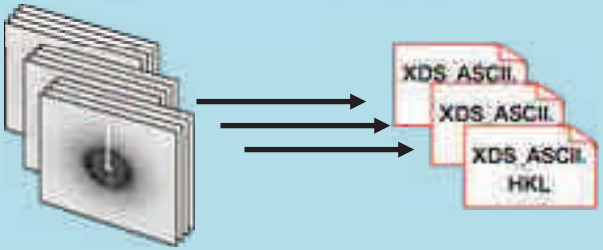
鴨

- Index/Integrate each dataset
- Grouping datasets (cells)
- Determination of space group
- Clustering (cell/intensity CC)
- Selecting good cluster (Completenes, Redundancy)
- Merging datasets
- Outlier detection/rejection



Merge & polish

selected clusters

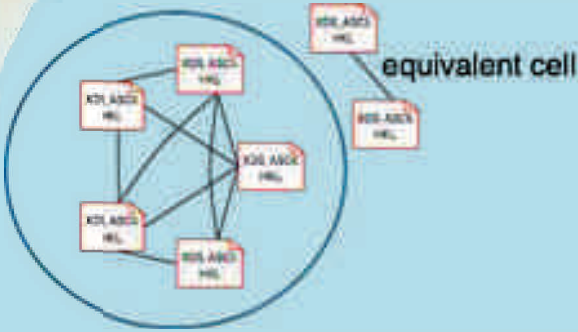


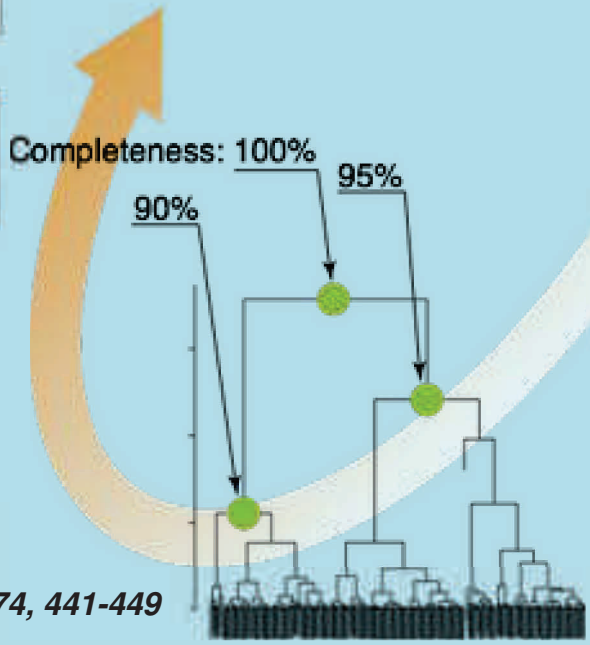
Data processing

XDS (or DIALS)

Grouping

crystals with equivalent cells





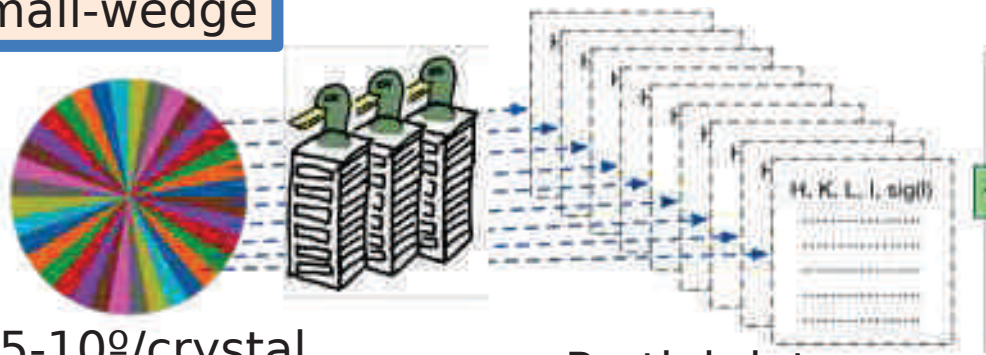
Clustering

- BLEND (cell parameters)
- CC of intensities

Yamashita, K., (2018). *Acta Cryst (2018). D74, 441-449*

Data back to users

Small-wedge



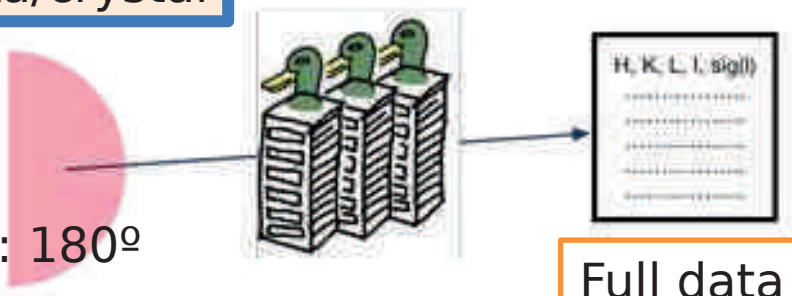
5-10º/crystal

Partial data

Clustering

Full datasets

Full data/crystal

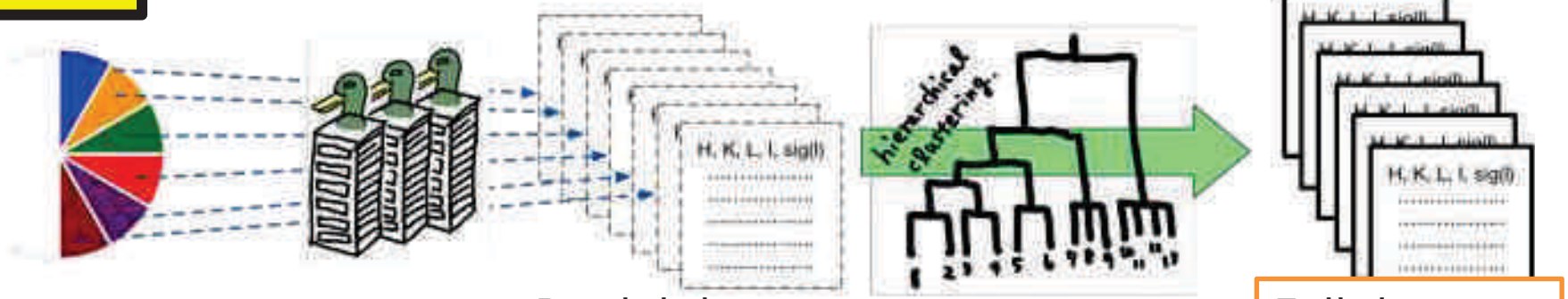


Ex: 180º

Full data

BUTTAGIRI

Clustering

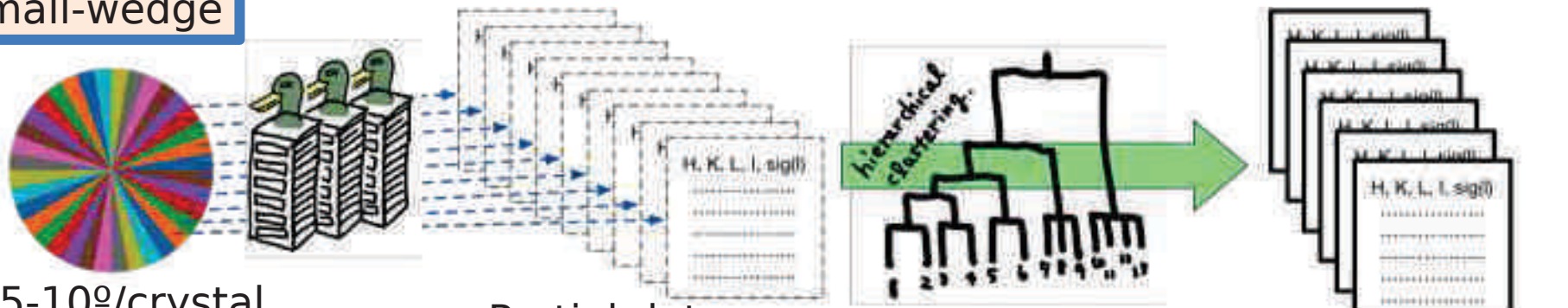


Partial data

Full datasets

Data back to users

Small-wedge



5-10º/crystal

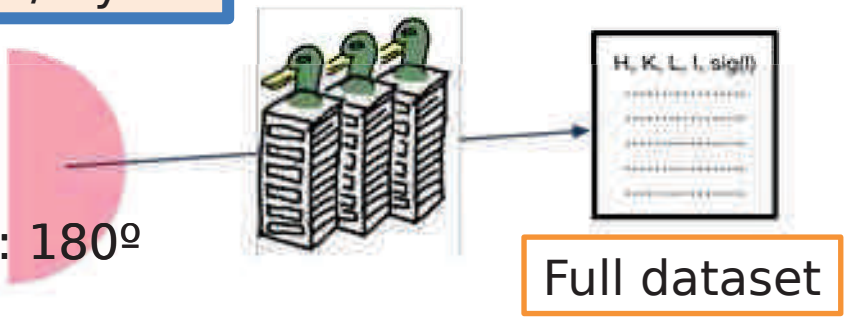
Partial data

Clustering

Full datasets

Data back to users

Full data/crystal



Ex: 180º

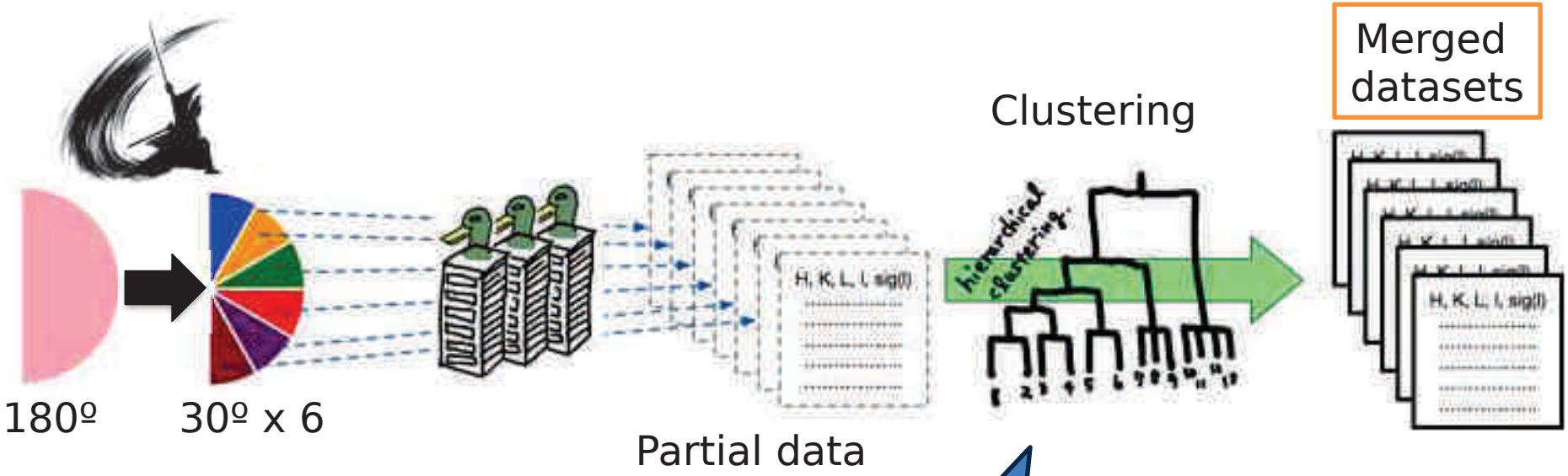
Full dataset

BUTTAGIRI process

(Chunk data processing?)

divides ‘complete data’ to ‘small wedge data’

(Ex. ： Full data 180º = 30º x 6 wedges)



Hierarchical clustering

(Select better data in ‘merging process’)

All types of processed/merged datasets is shipped back

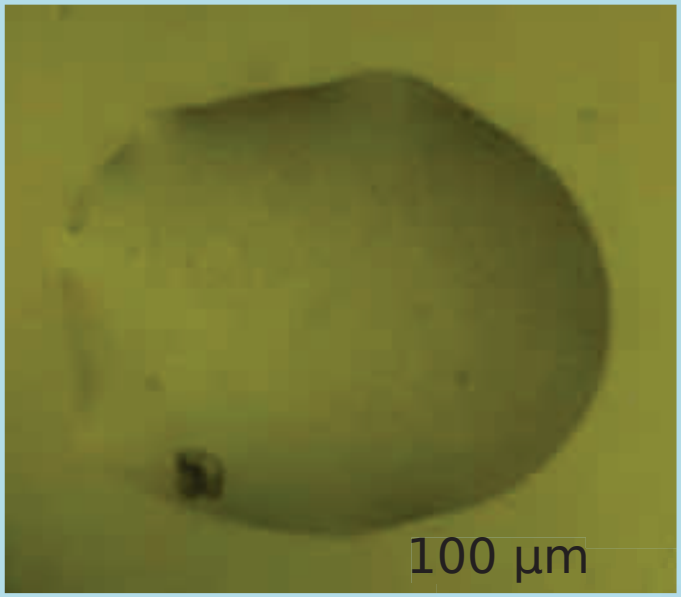
problem solved by ZOO

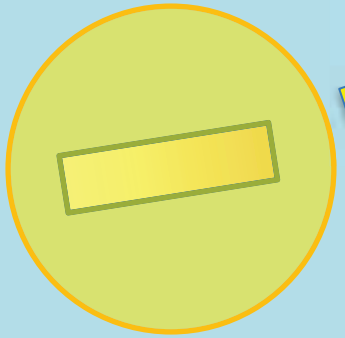
Courtesy of Prof. Sengoku

Data collection from ‘inhomogeneous crystal’

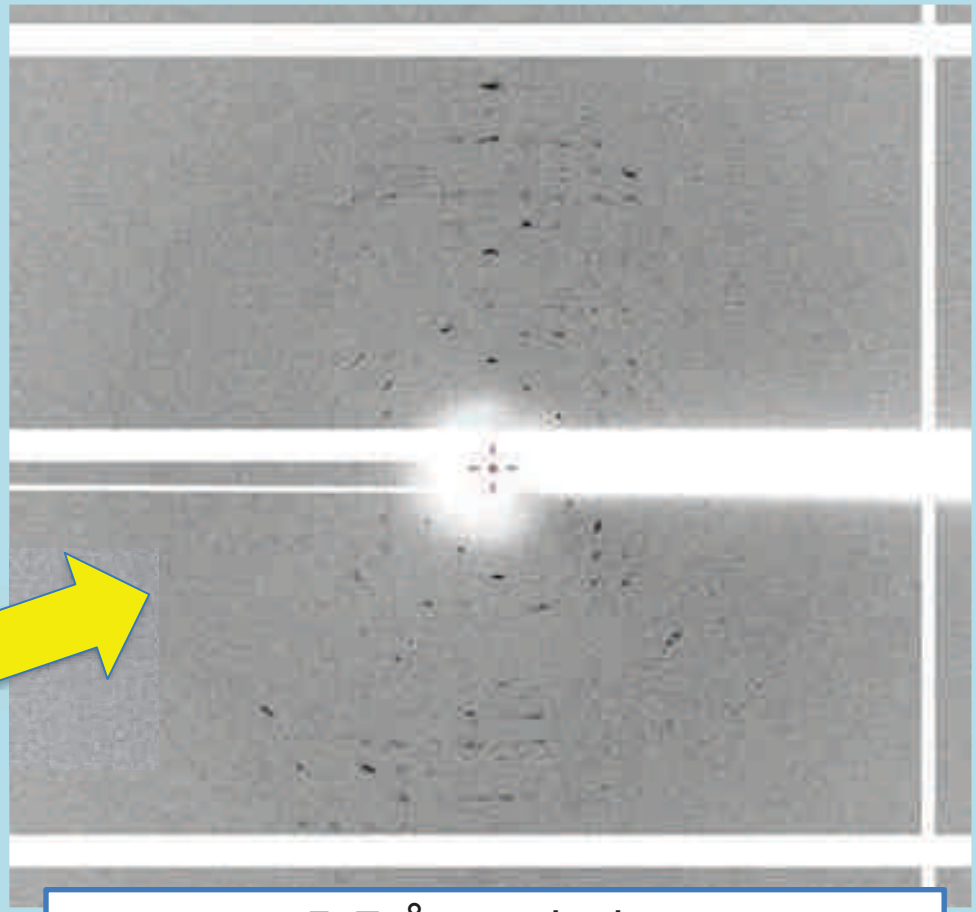
Is a human superior to ‘automatic data collection’, especially for ‘inhomogenous crystal’?

This case is one of the answers to this question.





Expose 200 µm beam



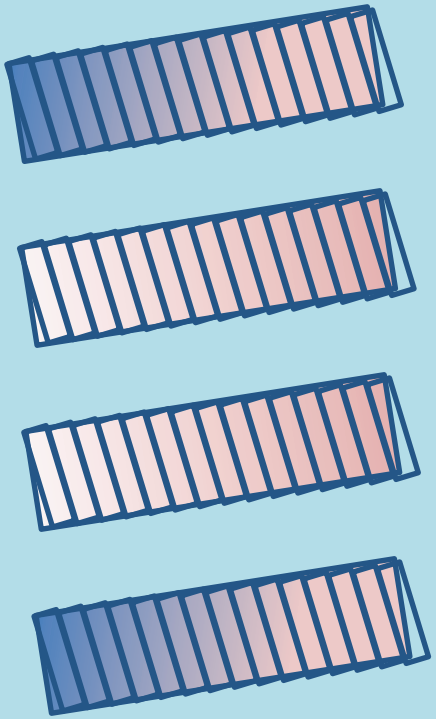
5-7 Å resolution
Diffraction shows ‘elongated’ peaks

problem solved by ZOO

Courtesy of Prof. Sengoku

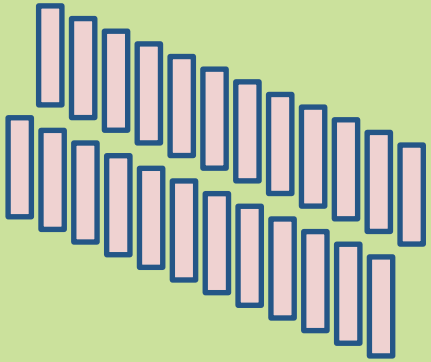
ta collection from ‘inhomogeneous cryst

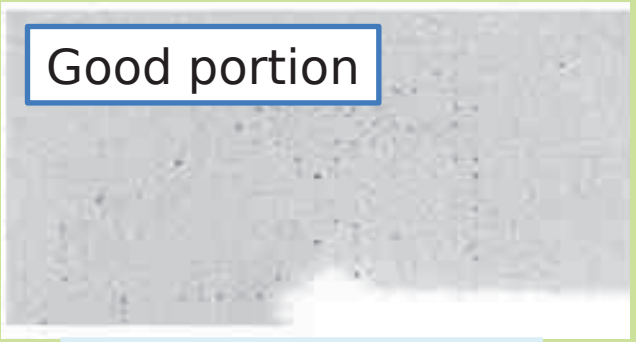
Helical 720º
→ BUTTAGIRI



Clustering

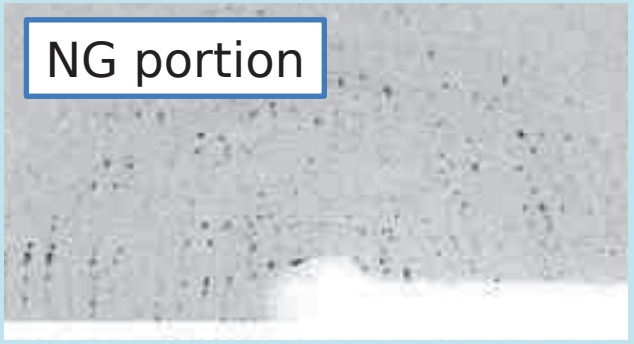
Better datasets





2.8 Å resolution

Worse datasets



11 crystals

Automation can solve structure using ‘inhomogeneous’ crystals

T. Kato *et al*. Nat. Chem. 12, 1081-1088(2020)

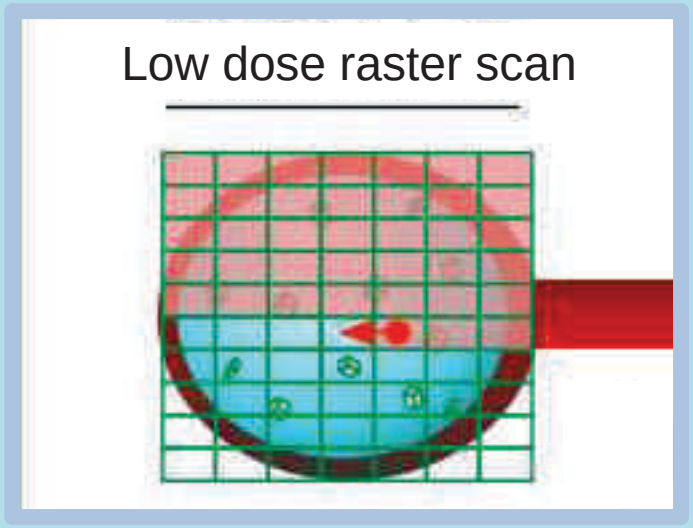
Mixed scheme(HITO)

objective :

To apply optimal scheme for each crystal with different size and configuration

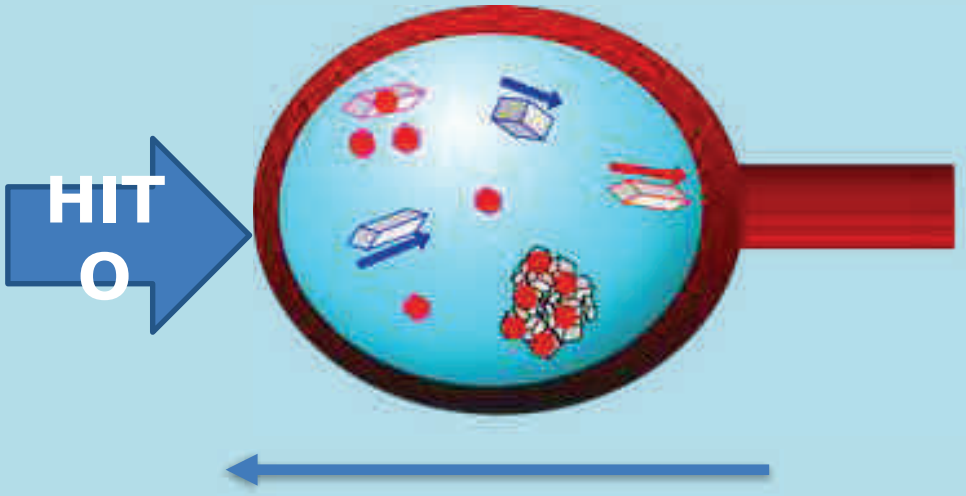
Small crystal → “SWSX”
Large crystal → “helical”

ZOO always does



SHIKA makes a crystal map

HITO recognizes crystal size/configuration



Rotation axis

HITO can choose ‘suitable scheme’ for each crystal

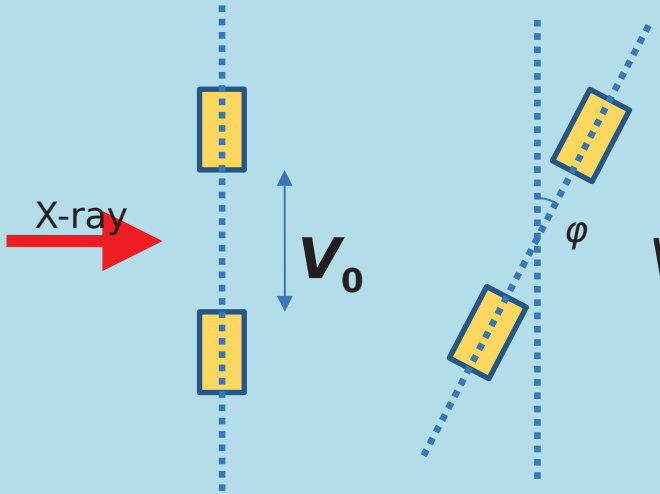
What HITO does?

Rotation axis

Crystal map from raster scan (from SHIKA)

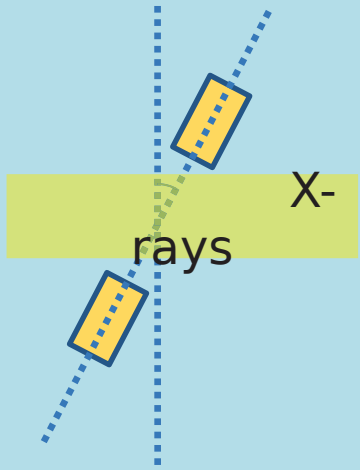
Find ‘overlapped crystals’ along with a rotation axis and calculate gaps between them.

How to calculate ‘oscillation range’



inclination φ *[deg.]*

$$V_x = V_0 \cdot \cos\varphi$$



(Current)

$$V_0 \cdot \cos\varphi \quad > 2V_{beam}$$

(beam size)

Ex) BL32XU

V beam size **15 µm**

$V_0 \cdot \cos\varphi = 2 \times 15$

V_0 is from a crystal map from raster

Ex) $V_0 = 50$ µm

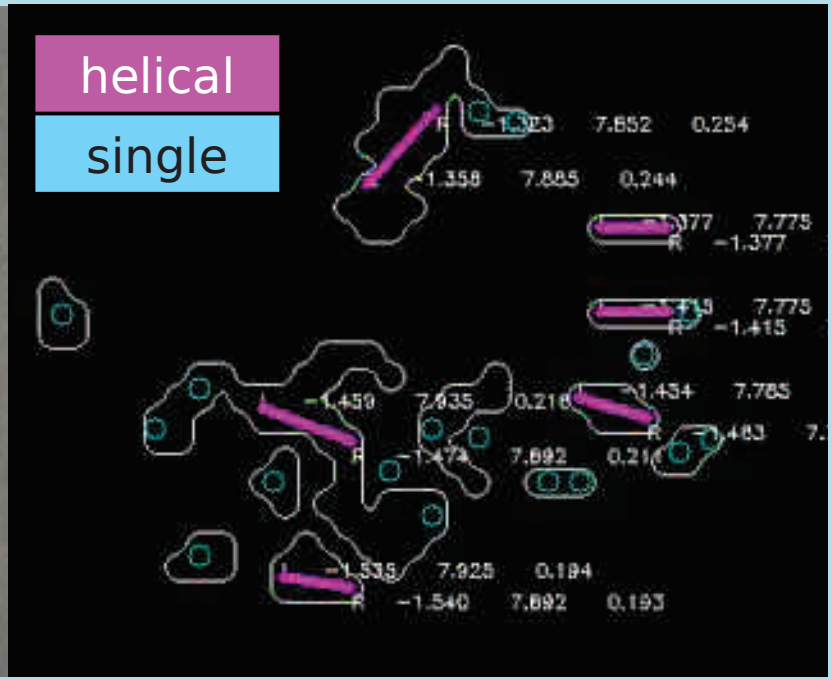
Φ ~ ±53º

※Crystal thickness -> X

HITO .vs. LCP crystals

Company users @ BL45XU (Oct 2021）

LCP crystals in a plate



HITO scheme map

Rotation data from mixed mode

PIN	Osc. range
Helical(no-align)	40° x 5
Helical part(align)	50° x 1
Multi	10° x 6
Single(full)	30° x 4
Single(no-align)	20° x 3
Single(align)	70° x 1, 50° x 1

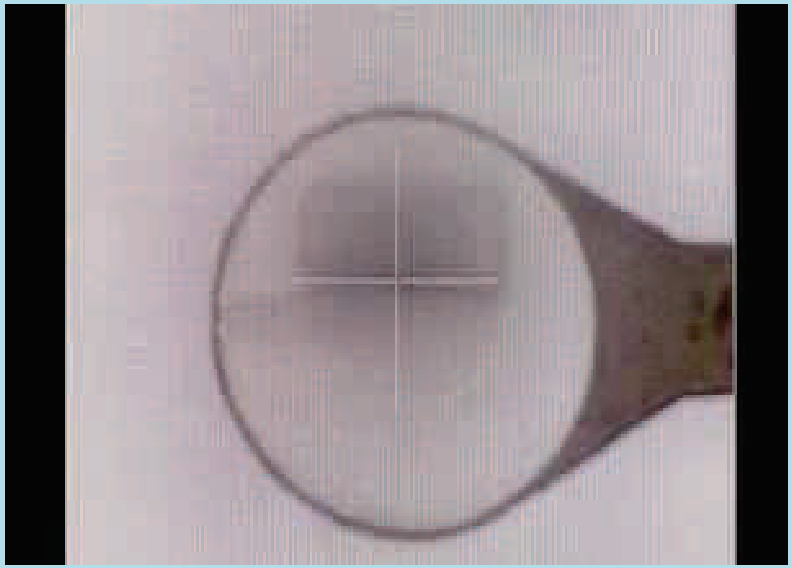
Rotation data of 610º was collected from 1 loop.

User's feedback

8 times more redundancy (cf. SWSX)

Serial Synchrotron Rotation crystallography (SSROX)

2D raster scan with goniometer rotation



- Micro focused beam (1~5 μm)
- Translation+Rotation
  - ±45 deg./line
- High dose expose
  - 10 MGy/pts

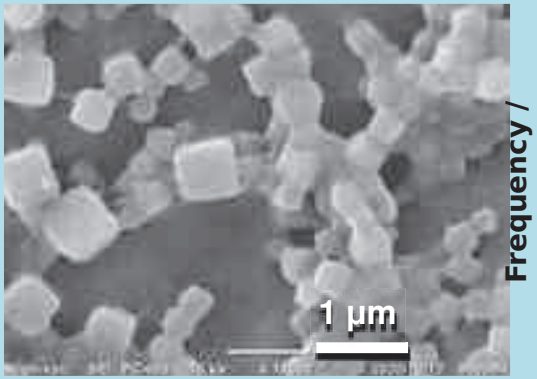
J. Synchrotron Rad. (2017). 24, 29–41

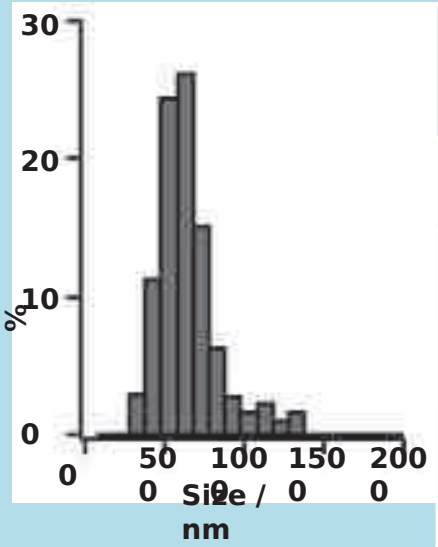
- **Still measurement requires exceedingly more crystals with the ‘monochromatic’ beam.**

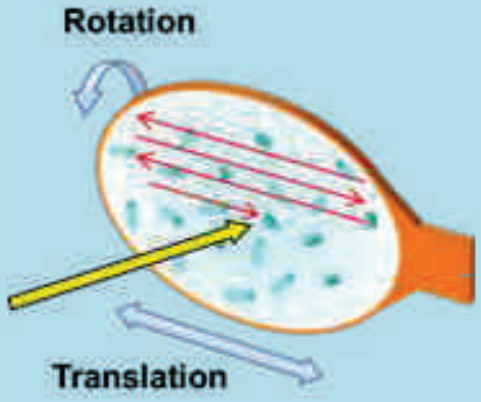
A single snapshot from a single crystal

Nano-sized PhC with SSROX

580 ± 230 nm





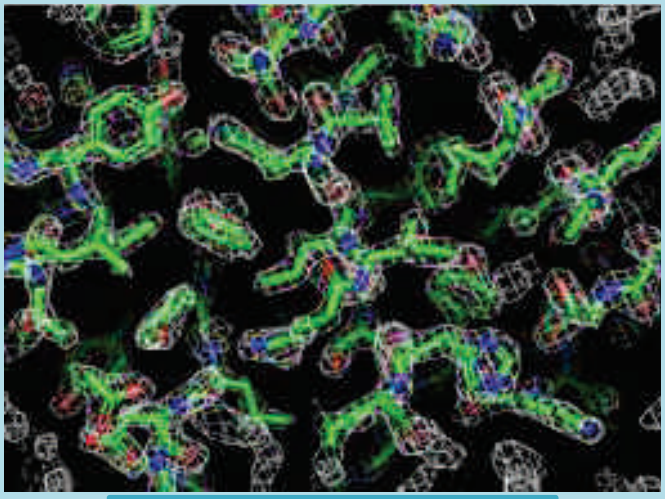


BL32XU	SSROX params.
Beam size	1 μm
Dose/point	10 MGy
Translation	360 μm
Rotation	-45 deg. ~ +45 deg.
# of frames	360
Translation / frame	1.0 μm
Oscillation width	0.25 deg. / frame
Exposure time	0.01 sec / frame
Attenuation	None.

Automated data collection with ZOO under cryo-condition

SSROX data using 600 nm crystals

# of loops	4
# of collected frames	330,480
# of hits	35,204
# of indexed	8,494
Resolution	1.95 Å
<I/σI> (1.96-1.95Å)	7.3 (1.23)
CC(1/2) (1.96-1.95Å)	0.9908 (0.6707)



1.95 Å resolution

Cell dimensions a~105Å (*I*23)

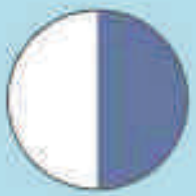
Crystal size 600 nm = 6000 Å ~ 57 copies

57 copies ^ 3 = **1.9×10^5 copies**

8,494 crystals : **1.6×10^9 copies**

Scientific Reports **volume 12**, Article number: 16031 (20

Available schemes in ZOO

Scheme		Rot. (/xtal)
”Small wedge” (SWSX)	 10–300 crystals	 5-10 deg.
“helical” or “single”	 single or a few crystals	 90-360 deg.
“mixed”	 small & large crystals	 5-10 deg. + 30-360 deg.
“SSROX”	 300-10,000 crystals	 < 1 deg.

Possible ‘goniometer based’ experiments are available

Summary

- After COVID
  - Ratio of automatic data collection rapidly increases
- After CryoEM appears
  - The number of users is getting smaller
  - Especially for membrane proteins
- Nano-crystallography
  - 600 nm crystals at atomic resolution
  - Each crystal: unit cells 2 x 10^5 copies
  - Total 1.6 x 10^9 copies
- Near future
  - **Towards** “No spread sheet” experiments

Acknowledgements

JASRI
Kazuya Hasegawa
Takashi Kumasaka
Seiki Baba
Nobuhiro Mizuno
Hironori Murakami
Yuki Nakamura

RIKEN
Keitaro Yamashita
Go Ueno
Yoshiaki Kawano
Naoki Sakai
Masaki Yamamoto

Tokyo University
Takafumi Kato
Wataru Shihoya
Kaoru Kumazaki
(Astellas)
Hideaki Kato
Yoshiki Tanaka
(NAIST)
Tomohiro Nishizawa
Osamu Nureki
Keitaro Yamashita

RIKEN Yokohama
Hiroaki Tanabe

Kyoto University
Ryoji Suno
Hidetsugu Asada
Im Dohyun
Yosuke Toyoda
Takuya Kobayashi
So Iwata

DECTRIS
Supporting staffs