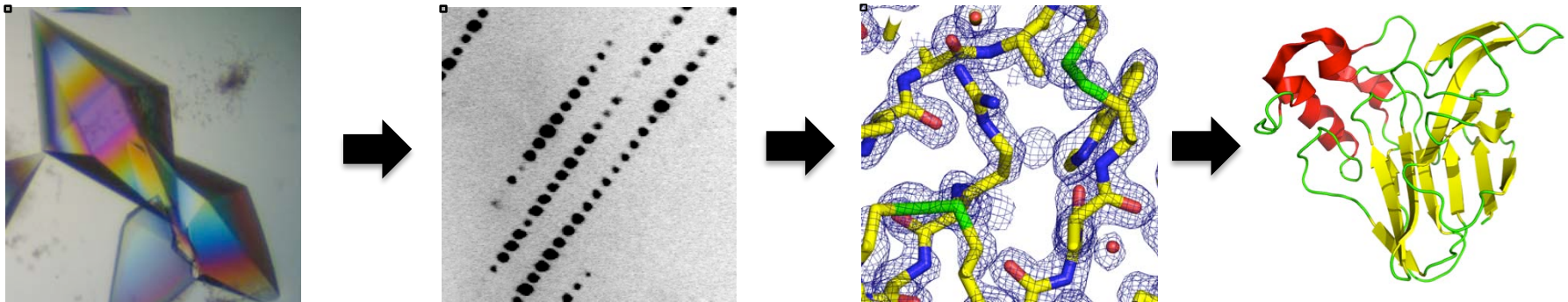


# Soft X-rays in the Home Lab: Native SAD Phasing with Chromium and Copper Radiation

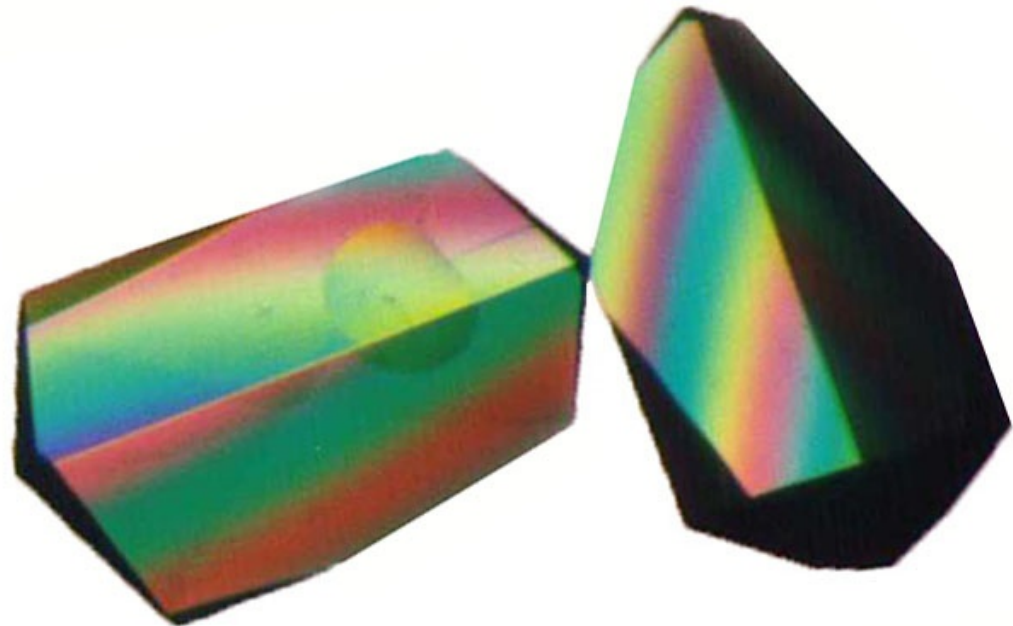
J.W. Pflugrath and Mark Del Campo  
Rigaku Americas

[Jim.Pflugrath@Rigaku.com](mailto:Jim.Pflugrath@Rigaku.com)



# Outline of the seminar

- Hardware
- Software
- Examples
- Speculation



“... on the shoulders of giants.”



- $\text{CuK}_\alpha$  or  $\text{CrK}_\alpha$  radiation from a home source
  - Blow, D.M. Proc. of the Royal Society 247, 302-336 (1958).
  - Hendrickson, W.A. and Teeter, M.M. Nature 290, 107-113 (1981).
  - Wang, B.C. Methods in Enzymology, 115, 90-112 (1985).
  
- Synchrotron radiation near the sulfur absorption edge
  - Stuhrmann, S., Bartels, K.S., Braunwarth, W., Doose, R., Dauvergne, F., Gabriel, A., Knöchel, A., Marmotti, M., Stuhrmann, H.B., Trame, C. and Lehmann, M.S. J. Synchrotron Rad. 4, 298-310 (1997).
  - Behrens, W., Otto, H., Stuhrmann, B. and Heyn, M.P. Biophys. J. 75, 255-266, (1998).
  
- Synchrotron radiation at 1.54 or 2.0 Å wavelength
  - Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. and Sheldrick, G.M. J. Mol. Biol. 289, 83-92 (1999).
  - Cianci, M., Rizkallah, P.J., Olczak, A., Raftery, J., Chayen, N.E., Zagalsky, P.F. and Helliwell, J.R. Acta Cryst. D57, 1219-1229 (2001).

# Background

## History of Sulfur Phasing:

1981: Hendrickson & Teeter (Nature), crambin (4.72kDa) with S-RAS (resolved anomalous scattering) method.

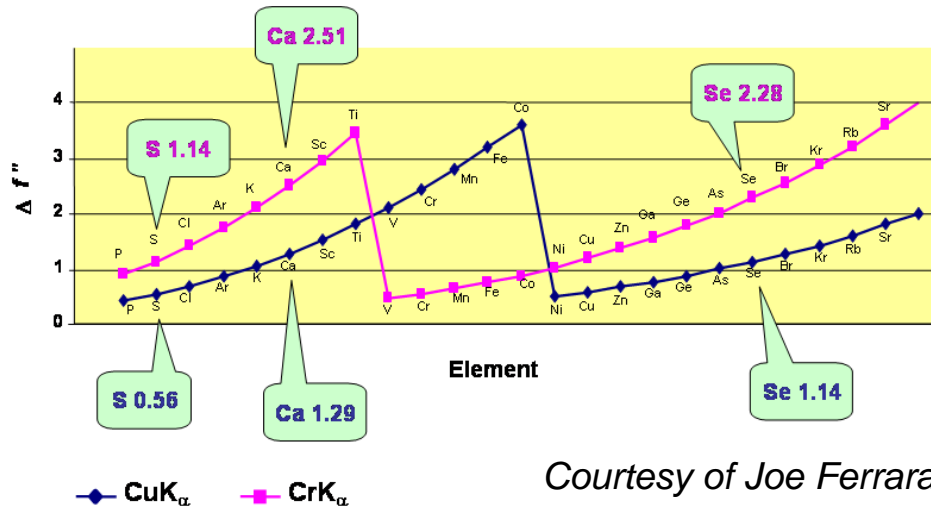
1985: BC Wang (Methods in Enzymology), simulated a S-SAD phasing of protein RHE using iterative single-wavelength anomalous scattering (ISAS) method.

1999: Dauter (Acta Cryst D), feasibility study on lysozyme.

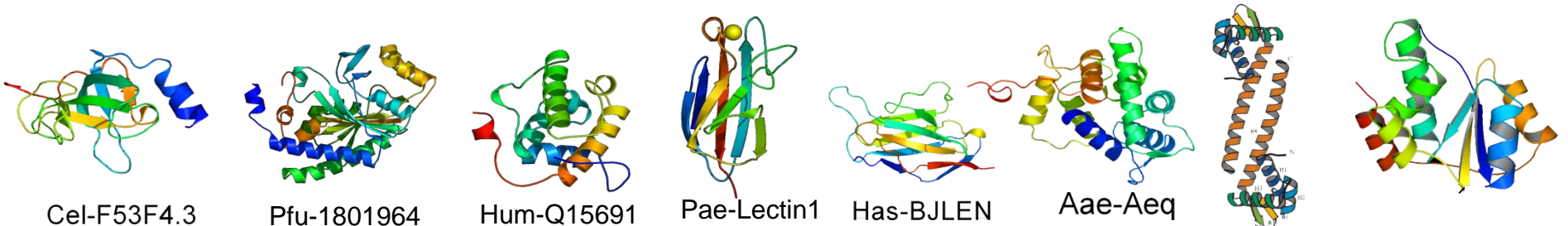
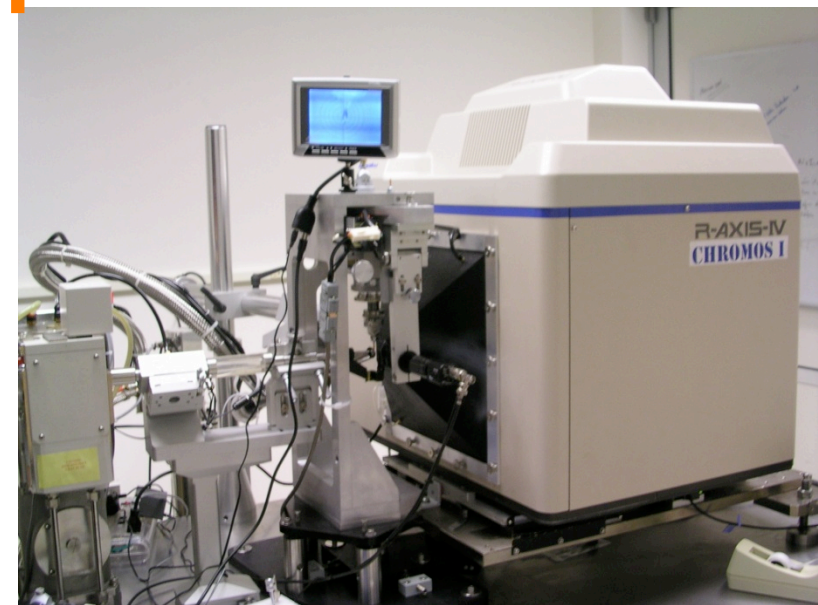
2000: Liu, etc (Protein Sci), new structure of obelin,  $\lambda=1.74\text{\AA}$ , ISAS method.

# Background

## Structures determined by Sulfur-SAD phasing using Cr X-rays



Courtesy of Joe Ferrara



SSO10A  
 PFU 542154  
 50 residues / sulfur



Courtesy of Zhi-Jie Liu, IBP

# Our Home Lab Setup

- FR-E+ Dual Wavelength
- VariMax DW Multilayer Optic
- HTC Imaging Plate
  - Helium Beam Path



# X-ray Generator: FR-E+ SuperBright



- Highest brilliance microfocus rotating anode generator
  - 2.475 kW power (45 kV 55 mA)
  - 70  $\mu\text{m}$  focal spot size
- Delivers  $16 \times 10^{10}$  Cu  $K\alpha$  photons/second/ $\text{mm}^2$ 
  - for a 100  $\mu\text{m}$  sample at  $4.8 \text{ mR}$

# μXG Performance Comparison

	Rigaku MM-007 HF VariMax-HF	Rigaku FR-E+ VariMax-HF	Rigaku MM-003
Focal Spot Size (μm)	Ø70	Ø70	Ø30
Power (kW)	1.2	2.475	0.03
Beam Size at Sample (μm)	208	208	100
Divergence (mR)	4.8	4.8	7.3
Useable flux w/ 0.1 mm aperture (X-ray/sec/mm <sup>2</sup> )	7.8 x 10 <sup>10</sup>	16 x 10 <sup>10</sup>	1.4 x 10 <sup>10</sup>
<b>Normalized Useable Flux (RU + Blue)</b>	<b>19.5</b>	<b>40</b>	<b>3.4</b>

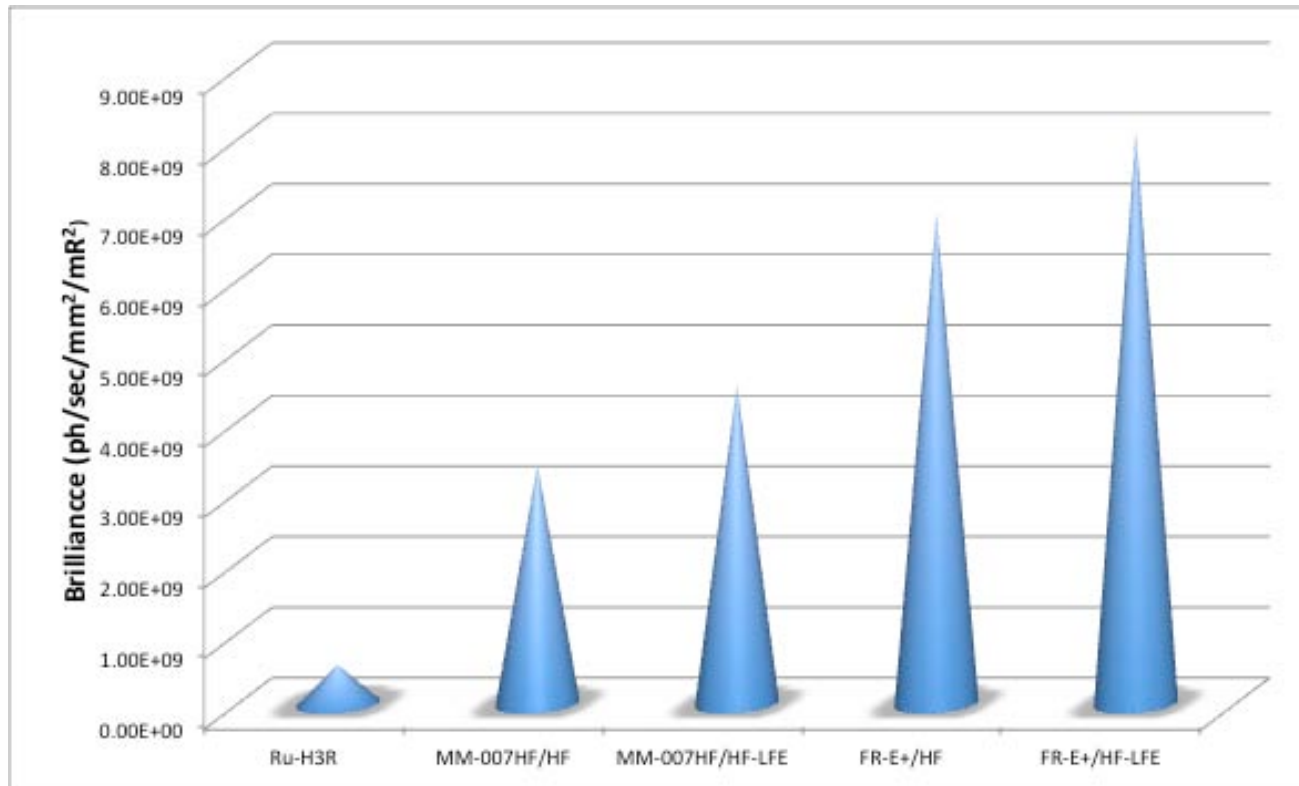


# μXG Performance Comparison

micro-focus rotating anode

	Rigaku MM-007 HF VariMax-HF	Rigaku FR-E+ VariMax-HF	Rigaku MM-003
Focal Spot Size (μm)	Ø70	Ø70	Ø30
Power (kW)	1.2	2.475	0.03
Beam Size at Sample (μm)	208	208	100
Useable flux w/ 0.1 mm aperture (X-ray/sec/mm <sup>2</sup> )	7.8 x 10 <sup>10</sup>	16 x 10 <sup>10</sup>	1.4 x 10 <sup>10</sup>
Divergence (mR)	4.8	4.8	7.3
Brilliance (X-ray/sec/mm <sup>2</sup> /mR <sup>2</sup> x 10 <sup>9</sup> )	3.4	7.0	0.26
<b>Normalized Brilliance (RU + Blue)</b>	<b>6.07</b>	<b>12.5</b>	<b>0.47</b>

# Home Lab: Like a Bending Magnet Source

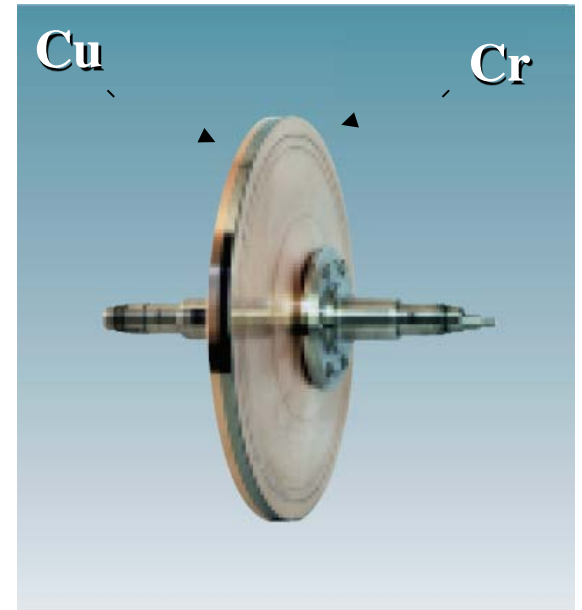


# Generator: FR-E+ DW SuperBright

- Two wavelengths in one generator
  - Cr and Cu rotating anode
  - Switchable electron gun



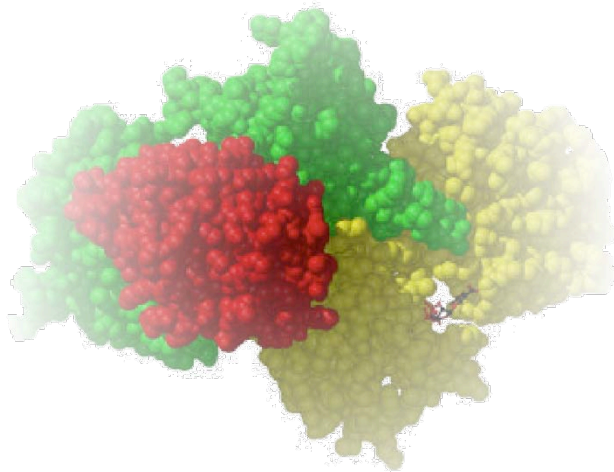
Ultimate HomeLab w/ FR-E+ DW SuperBright



Chrome Phasing  
without  
Derivatization

# Why use Cr radiation?

15 P 30.974	16 S 32.06	17 Cl 35.453
33 As 74.922	34 Se 78.96	35 Br 79.904
51 Sb 121.757	52 Te 127.60	53 I 126.905

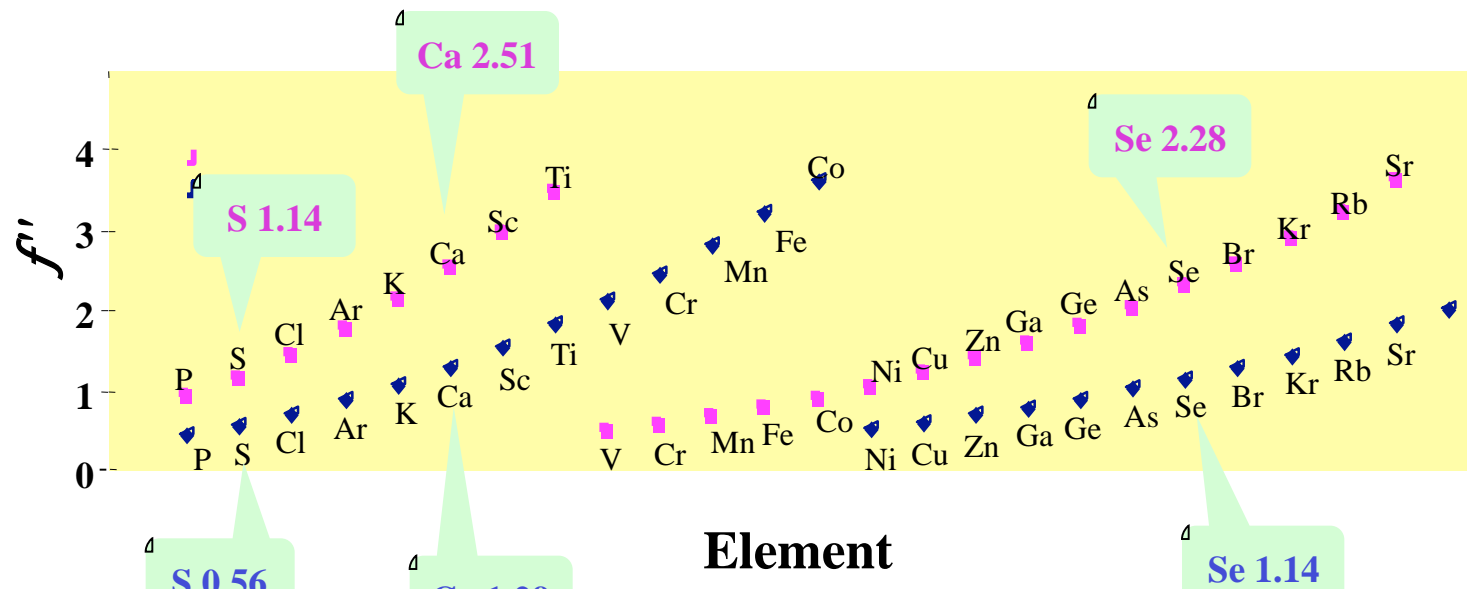


**Rigaku**

- Problem
  - Scientists would like to speed up the process of solving structures and overcome the difficulty of expressing and crystallizing selenomet proteins
- Solution
  - Enhance the anomalous signal from sulfur, which is present in nearly all proteins (certainly present in the proteins for which we expect to use selenomet)
- In addition
  - Enhance the anomalous signal from some native metals, present in nearly 1/3 of all proteins
  - Enhance the anomalous signal from heavy atom derivatives, for example  $\Delta f' \approx 12.6 e^-$  for iodine

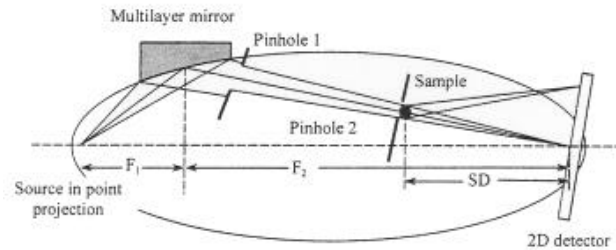
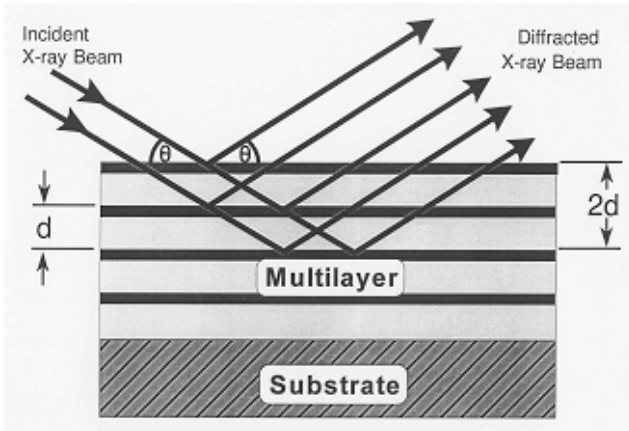
# Native SAD Phasing with Cu or Cr

- Cr SAD phasing takes advantage of increasing of anomalous scattering of many intrinsic elements like S
- Obtain the phase of protein data without any derivatization.
  - Make the density map available before sending your crystal to synchrotron for the final high resolution data.

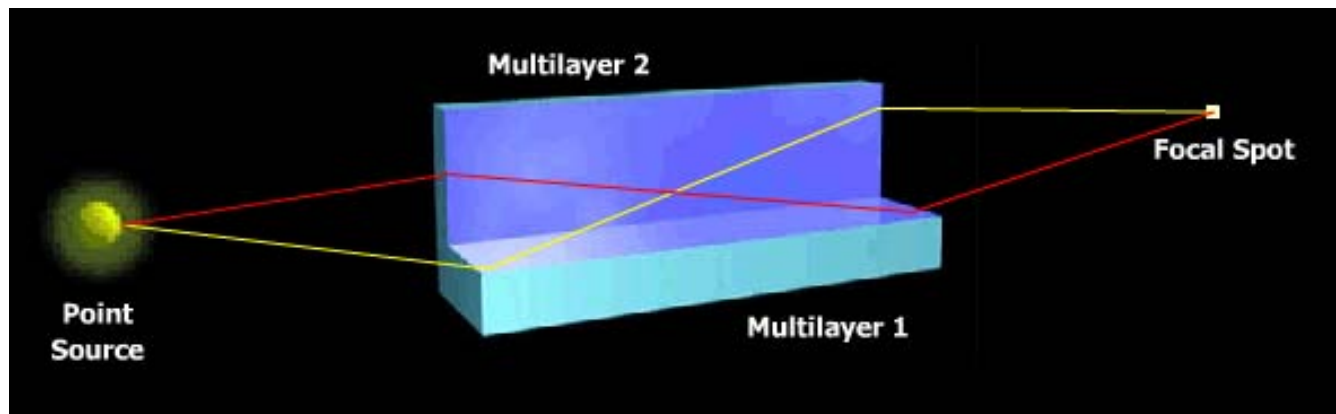
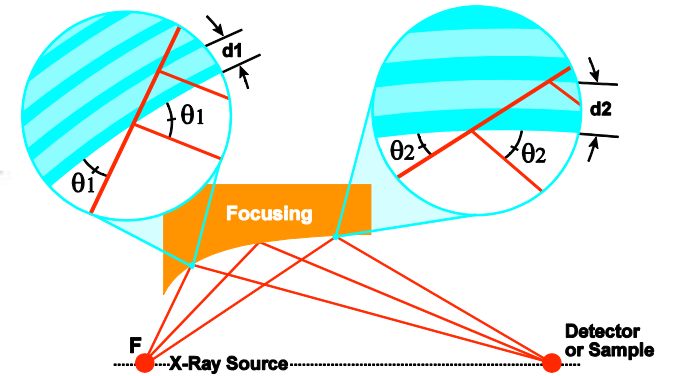


◆ Cu K $\alpha$     ■ Cr K $\alpha$

# Osmic Multilayer Optics

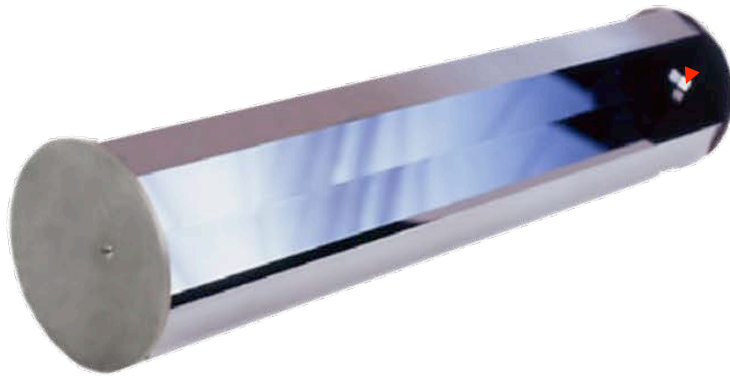


Focusing Multilayer Optic



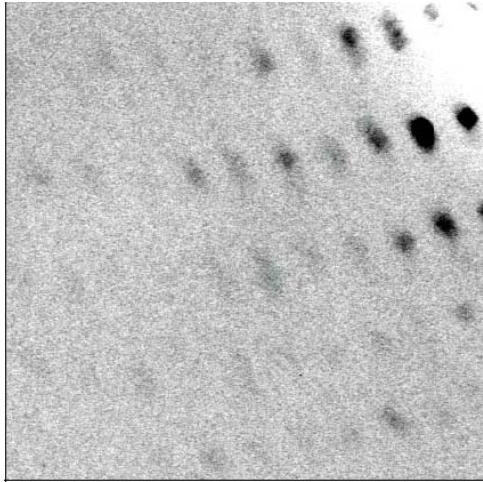
# Osmic VariMax™ Optics

Variable  
Divergence Slit



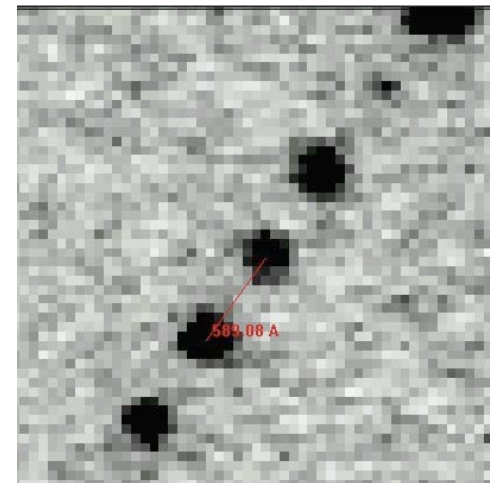
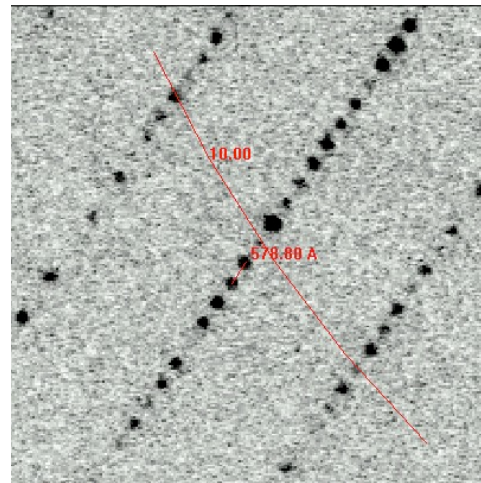
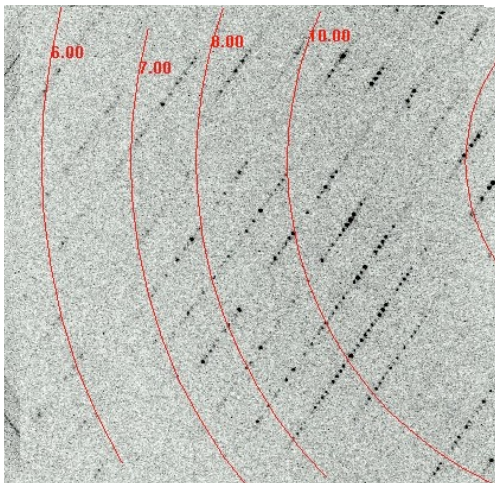
- VariMax HR
- VariMax HF
- VarMax VHF

# VariMax™ for Large Unit Cells



- 50S Ribosomal Subunit

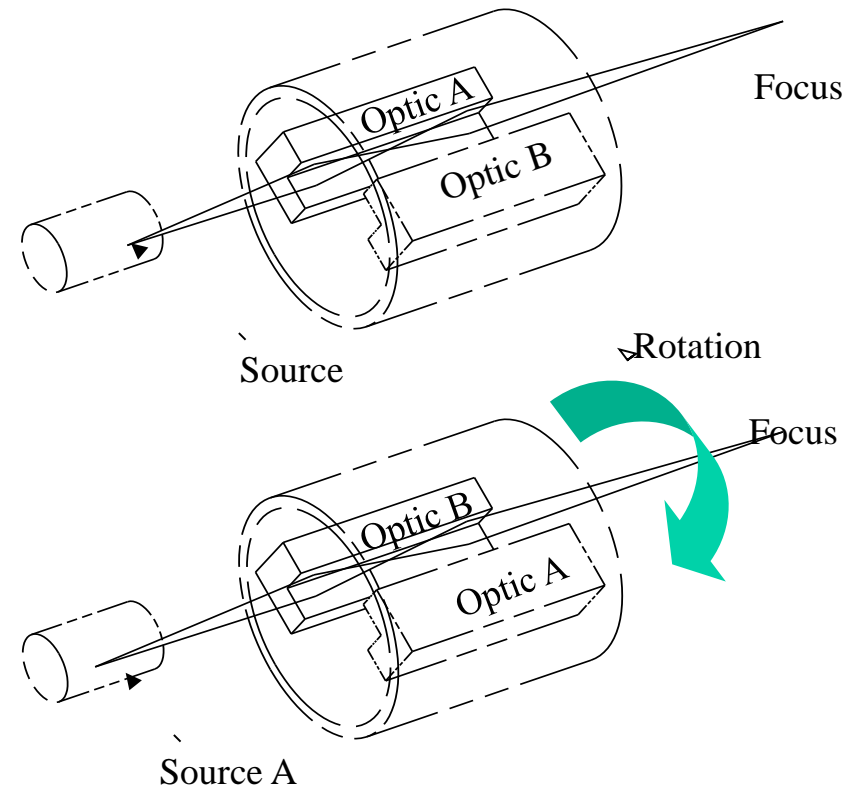
- Peter Moore, Tom Steitz and Martin Schmeing, Yale University
- Space group:  $C222_1$
- Unit cell dimensions
  - 216.98 302.30 578.28
  - 90.00 90.00 90.00





# VariMax-DW

- Multiple wavelengths
- With preferred take-off angle
- Pre-aligned to each source spot
- Cross at sample position



# DualWave Mirror VariMax DW

*Mirro  
VariMax DW*



Mo setting

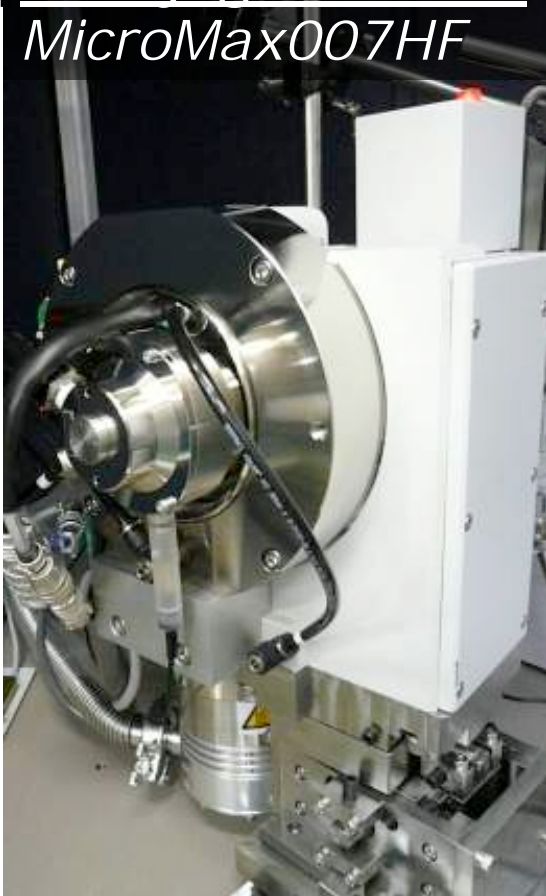


Cu setting

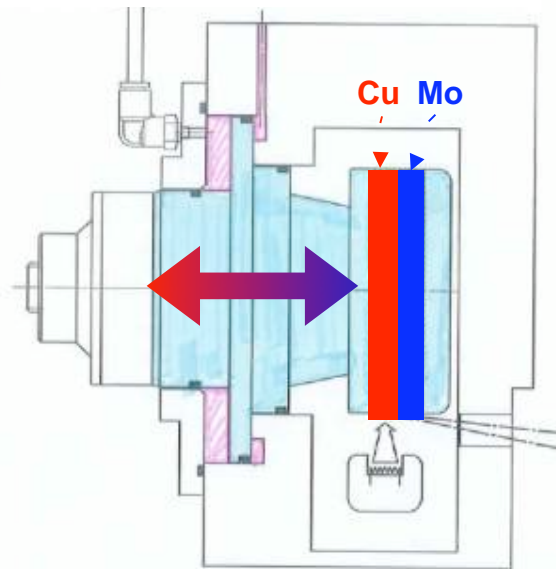
Just rotate the handle

# DualWave X-ray generator MicroMax007HF

X-ray generator  
MicroMax007HF



Target  
DW Target



Controller

<How to change>

1. X-ray OFF
2. Switch to position
3. X-ray ON

Unnecessary to remove  
target

# Imaging Plate Area Detectors

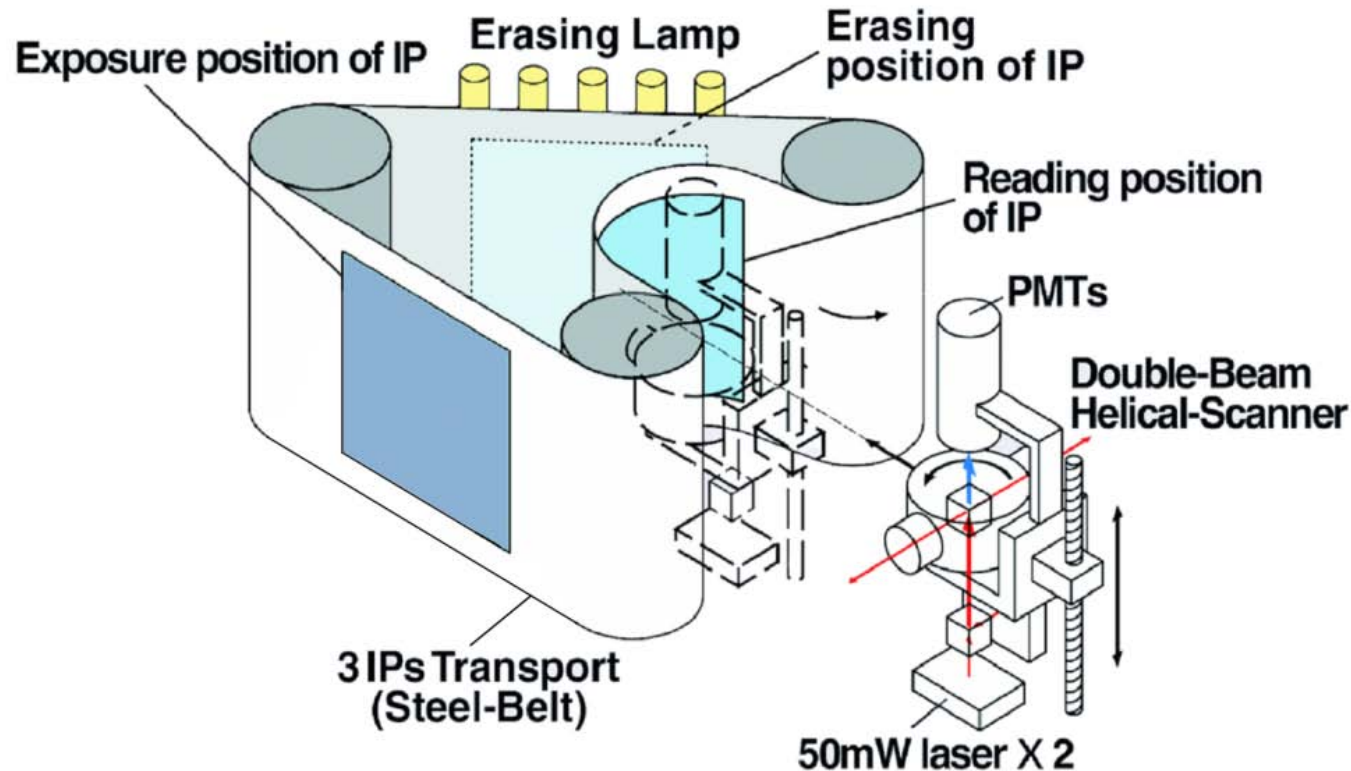


# Imaging Plate Benefits

- Very low background noise
- Large solid angle
- No image distortion (no FOT)
- No calibration except for PMTs
- Wide dynamic range
- Available helium beam path
  - For chromium phasing applications
  - No large internal air gap behind the faceplate
- Supported by CrystalClear™, HKL2000™, MOSFLM and XDS

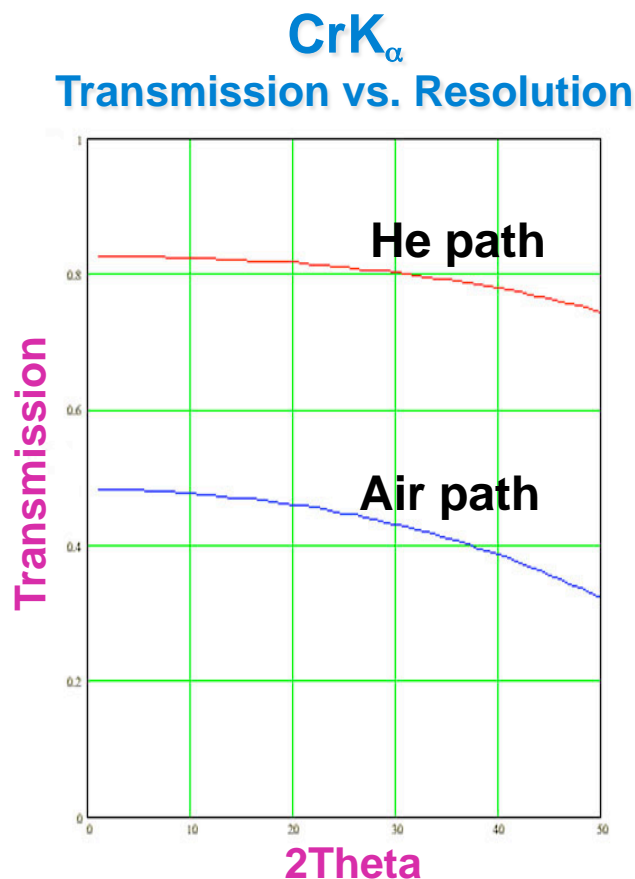


# R-AXIS HTC



- 3 IPs: Expose, Read, Erase simultaneously
- As fast as 1<sup>st</sup> generation CCD detectors
- 30 s read time

# The Absorption Problem



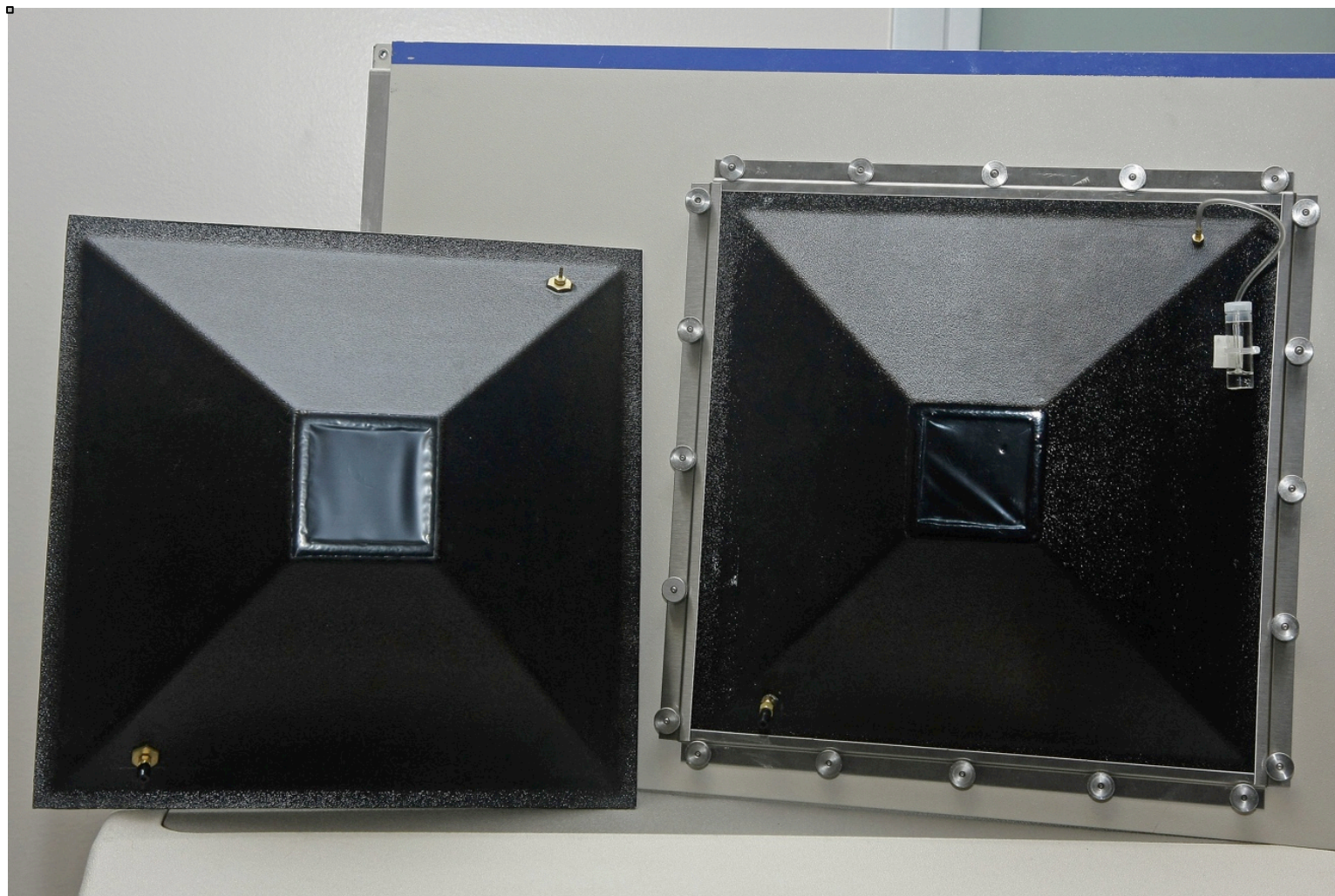
- More than 50% of the flux is lost
  - Without He<sub>(g)</sub> path (d= 135 mm)
- Modified hardware setup
  - Essential to preserve small anomalous signals
  - He<sub>(g)</sub> path is critical to success
- Minimization of the free air path is important

# Helium Beam Path





# Helium Beam Path



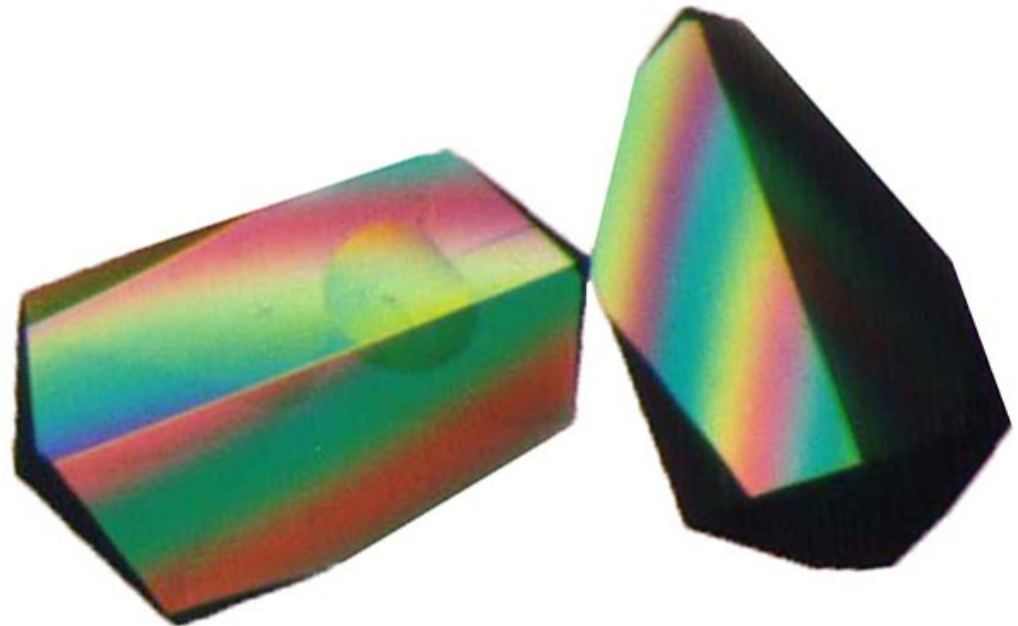
# Our Home Lab Setup

- FR-E+ Dual Wavelength
- VariMax DW Multilayer Optic
- HTC Imaging Plate
  - Helium Beam Path



# Outline of the seminar

- Hardware
- **Software**
- Examples
- Speculation



Acta Crystallographica Section D

**Biological  
Crystallography**

ISSN 0907-4449

**Wladek Minor,<sup>a\*</sup> Marcin  
Cymborowski,<sup>a</sup> Zbyszek  
Otwinowski<sup>b</sup> and Maksymilian  
Chruszcz<sup>a</sup>**

<sup>a</sup>Department of Molecular Physiology and  
Biological Physics, University of Virginia,  
Charlottesville, VA 22903, USA, and

<sup>b</sup>Department of Biochemistry, UT Southwestern  
Medical Center at Dallas, Dallas, TX 75235,  
USA

Correspondence e-mail:  
wladek@iwonka.med.virginia.edu

## ***HKL-3000*: the integration of data reduction and structure solution – from diffraction images to an initial model in minutes**

A new approach that integrates data collection, data reduction, phasing and model building significantly accelerates the process of structure determination and on average minimizes the number of data sets and synchrotron time required for structure solution. Initial testing of the *HKL-3000* system (the beta version was named *HKL-2000\_ph*) with more than 140 novel structure determinations has proven its high value for MAD/SAD experiments. The heuristics for choosing the best computational strategy at different data resolution limits of phasing signal and crystal diffraction are being optimized. The typical end result is an interpretable electron-density map with a partially built structure and, in some cases, an almost complete refined model. The current development is oriented towards very fast structure solution in order to provide feedback during the diffraction experiment. Work is also proceeding towards improving the quality of phasing calculation and model building.

Received 30 March 2006

Accepted 26 May 2006

## HKL3000 – Structure tab uses:

CCP4 suite --- Many authors

Collaborative Computational Project, Number 4. 1994.

"The CCP4 Suite: Programs for Protein Crystallography". Acta Cryst. D50, 760-763

Data Analysis (signal high enough?)

Find sites: SHELXD --- Sheldrick

Phase: MLPHARE + DM --- Otwinowski; Cowtan

Auto-build: ARP/wARP --- Perrakis, Cohen, Lamzin, et al.

REFMAC: --- Murshudov, Vagin, Steiner, et al.

Project  
Data  
Summary  
Index  
Strategy  
Integrate  
Scale

**Structure**

Macros



HKL-3000 v703b.db030.ph106 Package Licensed to Joseph Ferrara at Rigaku/MSC

File Options Site Configuration Crystal Information Report Help

Project Collect Data Summary Index Strategy Integrate Scale Structure Publication Macros Credits Copyrights

Scaled Sets  
1. output.sca

Space Group P43212  
# of Residues 129 / 129  
Molecular Weight 14313.65  
Anomalous Atom S  
# of Anomalous Atoms 10 / 10  
Wavelength 1.54178  
Resolution 25.00 1.50  
Solvent Content 0.38 Estimate

Auto  
Abort  
History  
Advanced

Finding Sites - Shelx  
Data Analysis Find Sites Sites View Phase NCS Build Refine

Anomalous Signal to Noise vs. Resolution

Resolution (A)	Signal to Noise
6.67	3.8
5.00	4.5
4.00	3.5
3.33	2.8
2.86	2.0
2.50	1.8
2.22	1.9
2.00	1.8
1.82	1.5
1.67	1.3
1.54	1.1

Completeness vs. Resolution

Resolution (A)	Completeness (%)
6.67	90
5.00	98
4.00	99
3.33	99
2.86	99
2.50	99
2.22	99
2.00	99
1.82	98
1.67	95
1.54	85

Data Analysis  
Check Twinning

## feature articles

---

Acta Crystallographica Section A

### Foundations of Crystallography

ISSN 0108-7673

Received 5 July 2007

Accepted 7 September 2007

## A short history of *SHELX*

George M. Sheldrick

Department of Structural Chemistry, University of Goettingen, Tammannstrasse 4, D-37077  
Goettingen, Germany. Correspondence e-mail: gsheldr@shelx.uni-ac.gwdg.de

An account is given of the development of the *SHELX* system of computer programs from *SHELX-76* to the present day. In addition to identifying useful innovations that have come into general use through their implementation in *SHELX*, a critical analysis is presented of the less-successful features, missed opportunities and desirable improvements for future releases of the software. An attempt is made to understand how a program originally designed for photographic intensity data, punched cards and computers over 10000 times slower than an average modern personal computer has managed to survive for so long. *SHELXL* is the most widely used program for small-molecule refinement and *SHELXS* and *SHELXD* are often employed for structure solution despite the availability of objectively superior programs. *SHELXL* also finds a niche for the refinement of macromolecules against high-resolution or twinned data; *SHELXPRO* acts as an interface for macromolecular applications. *SHELXC*, *SHELXD* and *SHELXE* are proving useful for the experimental phasing of macromolecules, especially because they are fast and robust and so are often employed in pipelines for *high-throughput* phasing. This paper could serve as a general literature citation when one or more of the open-source *SHELX* programs (and the Bruker AXS version *SHELXTL*) are employed in the course of a crystal-structure determination.

© 2008 International Union of Crystallography  
Printed in Singapore – all rights reserved

Structure

Data Analysis

Find Sites

Phase

Build

Refine



HKL-3000 v703b.db030.ph106 Package Licensed to Joseph Ferrara at Rigaku/MS

File Options Site Configuration Crystal Information Report Help

Project Collect Data Summary Index Strategy Integrate Scale Structure Publication Macros Credits Copyrights

**Finding Sites - Shelx**

Data Analysis Find Sites Sites View Phase NCS Build Refine

Scaled Sets  
1. output.sca

Space Group P43212  
# of Residues 129 / 129  
Molecular Weight 14313.65  
Anomalous Atom S  
# of Anomalous Atoms 10 / 10  
Wavelength 1.54178  
Resolution 25.00 1.50  
Solvent Content 0.38 Estimate

Auto  
Abort  
History  
Advanced

Status: Try 78 CC All/Weak 15.16 / 0.59, best 44.36 / 26.31  
Time Elapsed: 00:01:56

**CC vs. PATFOM**

**CC All vs. Count**

High Resolution Limit 2.0

Find Sites  
Finish

Infinite Number of Cycles



Structure

Data Analysis

Find Sites

Phase

Build

Refine



HKL-3000 v703b.db030.ph106 Package Licensed to Joseph Ferrara at Rigaku/MS

File Options Site Configuration Crystal Information Report Help

Project Collect Data Summary Index Strategy Integrate Scale Structure Publication Macros Credits Copyrights

Scaled Sets  
1. outputsca

Space Group P43212  
# of Residues 129 / 129  
Molecular Weight 14313.65  
Anomalous Atom S  
# of Anomalous Atoms 10 / 10  
Wavelength 1.54178  
Resolution 25.00 1.50  
Solvent Content 0.38 Estimate

Auto  
Abort  
History  
Advanced

Finding Sites - Shelx: Done

Data Analysis Find Sites Sites View Phase NCS Build Refine

Last  
Best  
Use Sites

run 1 cycle 101:	44.25	/	26.55	pf	3.16
run 1 cycle 63:	44.36	/	26.31	pf	3.36
run 1 cycle 85:	14.32	/	7.13	pf	0.78
run 1 cycle 54:	17.36	/	6.81	pf	0.94
run 1 cycle 12:	19.14	/	6.44	pf	0.91

Symmetry  NCS  Center  Measure Distance:

Table

Delete

Below 0.10

MLPHARE

**ISOMORPHOUS REPLACEMENT AND  
ANOMALOUS SCATTERING**

Proceedings of the CCP4 Study Weekend  
25-26 January 1991

Compiled by W. Wolf, P.R. Evans and A.G.W. Leslie

MAXIMUM LIKELIHOOD REFINEMENT OF HEAVY ATOM PARAMETERS

Zbyszek Otwinowski

Howard Hughes Medical Institute  
and Department of Molecular Biophysics and Biochemistry,  
Yale University, New Haven, CT 06514, USA

**Introduction**

Analysis of isomorphous replacement (MIR) data has two major steps. The first is refinement of the heavy atom substitution parameters. The second is application of the results of the first step to calculate probabilities of all possible phases and from that distribution the centroid phase and figure of merit.

**Rigaku**

Otwinowski (1991) Proc CCP4 SW 1991, 80-86.

DM

Daresbury Laboratory

**JOINT CCP4 AND ESF-EACBM**

**NEWSLETTER ON**

**PROTEIN CRYSTALLOGRAPHY**



An informal Newsletter associated with the BBSRC Collaborative Computational Project No.4 on Protein Crystallography and the ESF Network of the European Association of the Crystallography of Biological Macromolecules.

Number 31

November 1994

### Contents

'dm': An Automated Procedure for Phase Improvement  
by Density Modification.

Dr. K. D. Cowtan, University of York, Heslington, York, YO1 5DD.

#### Density Modification: The Problem.

The techniques of protein structure solution by X-ray diffraction methods are now well established and are the method of choice for obtaining structural information about many types of biological macromolecule. In many cases however a structure solution by X-ray methods is still far from routine, and in particular the step of obtaining phases for the measured x-ray intensities can present problems. Experimental techniques, such as the measurement of intensities from single or multiple isomorphous derivatives or the differences due to anomalous scattering give estimates to some of the phase information, which may or may not be enough to deduce the structure of the molecule.

**Rigaku**

- Cowtan (1994) Joint CCP4 ... 31, 34-38.

Structure

Data Analysis

Find Sites

Phase

Build

Refine



HKL-3000 v703b.db030.ph106 Package Licensed to Joseph Ferrara at Rigaku/MSC

File Options Site Configuration Crystal Information Report Help

Project Collect Data Summary Index Strategy Integrate Scale Structure Publication Macros Credits Copyrights

Scaled Sets  
1. output.sca

Space Group P43212  
# of Residues 129 / 129  
Molecular Weight 14313.65  
Anomalous Atom S  
# of Anomalous Atoms 10 / 10  
Wavelength 1.54178  
Resolution 25.00 1.50  
Solvent Content 0.38 Estimate

Auto  
Abort  
History  
Advanced

Phasing - DM: Done

Data Analysis Find Sites Sites View Phase NCS Build Refine

Figure of Merit vs. Resolution

Resolution (Å)	Figure of Merit (Blue)	Figure of Merit (Red)
10.0	0.78	0.15
6.67	0.82	0.32
5.00	0.88	0.30
4.00	0.90	0.28
3.33	0.88	0.27
2.86	0.88	0.27
2.50	0.88	0.33
2.22	0.88	0.33
2.00	0.88	0.32
1.82	0.85	0.28
1.67	0.82	0.27
1.54	0.78	0.15

Phasing Power vs. Resolution

Resolution (Å)	Phasing Power
10.0	1.0
6.67	1.55
5.00	1.1
4.00	1.0
3.33	1.1
2.86	1.0
2.50	1.0
2.22	1.0
2.00	1.0
1.82	0.8
1.67	0.8
1.54	0.5

Phasing High Resolution Limit 1.5  
 Fast Mode

Check Hand  
Phase  
Display Map

Use Inverse Sites

# Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7

Gerrit Langer<sup>1,3</sup>, Serge X Cohen<sup>2,3</sup>, Victor S Lamzin<sup>1</sup> & Anastassis Perrakis<sup>2</sup>

<sup>1</sup>European Molecular Biology Laboratory, c/o DESY, Notkestrasse 85, Hamburg 22607, Germany. <sup>2</sup>Department of Molecular Carcinogenesis, Netherlands Cancer Institute, Plesmanlaan 121, Amsterdam 1066CX, The Netherlands. <sup>3</sup>These authors contributed equally to this work. Correspondence should be addressed to A.P. (a.perrakis@nki.nl) and V.S.L. (victor@embl-hamburg.de).

Published online 19 June 2008; doi:10.1038/nprot.2008.91

**ARP/wARP is a software suite to build macromolecular models in X-ray crystallography electron density maps. Structural genomics initiatives and the study of complex macromolecular assemblies and membrane proteins all rely on advanced methods for 3D structure determination. ARP/wARP meets these needs by providing the tools to obtain a macromolecular model automatically, with a reproducible computational procedure. ARP/wARP 7.0 tackles several tasks: iterative protein model building including a high-level decision-making control module; fast construction of the secondary structure of a protein; building flexible loops in alternate conformations; fully automated placement of ligands, including a choice of the best-fitting ligand from a 'cocktail'; and finding ordered water molecules. All protocols are easy to handle by a nonexpert user through a graphical user interface or a command line. The time required is typically a few minutes although iterative model building may take a few hours.**

Langer et al. (2008) Nature Protocols 3, 1171-1179.

## research papers

Acta Crystallographica Section D  
**Biological  
Crystallography**  
ISSN 0907-4449

**Alexei A. Vagin, Roberto A.  
Steiner,† Andrey A. Lebedev, Liz  
Potterton, Stuart McNicholas,  
Fei Long and Garib N.  
Murshudov\***

Structural Biology Laboratory, Department of  
Chemistry, University of York, York YO10 5YW,  
England

† Current address: IFOM – The FIRC Institute of  
Molecular Oncology, Via Adamello 16, 20139  
Milano, Italy

Correspondence e-mail: garib@ysbl.york.ac.uk

## ***REFMAC5* dictionary: organization of prior chemical knowledge and guidelines for its use**

One of the most important aspects of macromolecular structure refinement is the use of prior chemical knowledge. Bond lengths, bond angles and other chemical properties are used in restrained refinement as subsidiary conditions. This contribution describes the organization and some aspects of the use of the flexible and human/machine-readable dictionary of prior chemical knowledge used by the maximum-likelihood macromolecular-refinement program *REFMAC5*. The dictionary stores information about monomers which represent the constitutive building blocks of biological macromolecules (amino acids, nucleic acids and saccharides) and about numerous organic/inorganic compounds commonly found in macromolecular crystallography. It also describes the modifications the building blocks undergo as a result of chemical reactions and the links required for polymer formation. More than 2000 monomer entries, 100 modification entries and 200 link entries are currently available. Algorithms and tools for updating and adding new entries to the dictionary have also been developed and are presented here. In many cases, the *REFMAC5* dictionary allows entirely automatic generation of restraints within *REFMAC5* refinement runs.

Received 19 April 2004  
Accepted 22 September 2004

Murshudov et al. (1997) *Acta Cryst D* 53, 240-255.  
Vagin et al. (2004) *Acta Cryst D* 60, 2184-2195.

Structure

Data Analysis

Find Sites

Phase

Build

Refine



HKL-3000 v703b.db030.ph106 Package Licensed to Joseph Ferrara at Rigaku/MSC

File Options Site Configuration Crystal Information Report Help

Project Collect Data Summary Index Strategy Integrate Scale Structure Publication Macros Credits Copyrights

**Building Model - ARP/wARP**

Data Analysis Find Sites Sites View Phase NCS Build Refine

Scaled Sets  
1. output.sca

Space Group: I222  
# of Residues: 388 / 388  
# of Molecules: 1 Estimate  
Resolution: 50.00 1.29

HKL file: hkl\_phase.mtz  
PDB file: hkl\_build\_3.pdb  
Use: ARP/wARP ?  
# of Build Cycles: 3 One Cycle  
 Use Existing Model  
 Use Sequence

Auto  
Abort  
History  
Advanced

Status:  
Time Elapsed: 00:09:35

```
> Build model start 14:44:02 Nov 16, 2011 ...
Directory: build_model_7
Map in hkl_phase.mtz

Refine cycle 0: dummy atoms included R = 0.334 Rfree = 0.371

Build main cycle: 1

Build cycle 1: 306 aa (79%) in 17 chains, 78 aa (20%/m) in the longest chain
Build cycle 2: 329 aa (85%) in 9 chains, 88 aa (23%/m) in the longest chain
Build cycle 3: 338 aa (87%) in 8 chains, 80 aa (21%/m) in the longest chain
Build cycle 4: 339 aa (87%) in 7 chains, 90 aa (23%/m) in the longest chain
Build cycle 5: 339 aa (87%) in 8 chains, 81 aa (21%/m) in the longest chain
Build cycle end: chains 7, 332 aa (86%)
        6 chains have been docked - 323 aa (86%)
Refine cycle 1: dummy atoms included R = 0.286 Rfree = 0.307
```

> Secondary structure statistics

Chain A

MNYGPTPEDRF TFG LWT V G W Q G R D P F G D A T R R A L D P V E S V R R L A E L G A H G V T F H D D D L  
F G S S D S E R E E H V K R F R Q A L D D T G M K V P M A T T N L F T H P V F K D G G F T A N D R D V R R Y A L R K  
R N I D L A V E L G A E T Y V A W G G R E G A E S G G A K D V R D A L D R M K E A F D L L G E Y V T S G G Y D I R F  
E P K P N E P R G D I L L P T V G H A L A F I E R L E R P E L Y G V N P E V G H E G M A G L N F P H G I A Q A L W A K  
L E L M L N C S N P W U P C D L E F S A C R I C A A M L R L L E S A L S U P C C D L E F F S P D S T E P S

Build Model  
Display Model

## research papers

Acta Crystallographica Section D

**Biological  
Crystallography**

ISSN 0907-4449

**P. Emsley,<sup>a\*</sup> B. Lohkamp,<sup>b</sup>  
W. G. Scott<sup>c</sup> and K. Cowtan<sup>d</sup>**

<sup>a</sup>Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, England, <sup>b</sup>Department of Medical Biochemistry and Biophysics, Karolinska Institute, SE-171 77 Stockholm, Sweden, <sup>c</sup>Department of Chemistry, University of California, 1156 High Street, Santa Cruz, CA 95064, USA, and <sup>d</sup>Department of Chemistry, University of York, Heslington, York YO10 5DD, England

Correspondence e-mail:  
paul.emsley@bioch.ox.ac.uk

**Features and development of *Coot***

*Coot* is a molecular-graphics application for model building and validation of biological macromolecules. The program displays electron-density maps and atomic models and allows model manipulations such as idealization, real-space refinement, manual rotation/translation, rigid-body fitting, ligand search, solvation, mutations, rotamers and Ramachandran idealization. Furthermore, tools are provided for model validation as well as interfaces to external programs for refinement, validation and graphics. The software is designed to be easy to learn for novice users, which is achieved by ensuring that tools for common tasks are 'discoverable' through familiar user-interface elements (menus and toolbars) or by intuitive behaviour (mouse controls). Recent developments have focused on providing tools for expert users, with customisable key bindings, extensions and an extensive scripting interface. The software is under rapid development, but has already achieved very widespread use within the crystallographic community. The current state of the software is presented, with a description of the facilities available and of some of the underlying methods employed.

Received 9 June 2009  
Accepted 26 February 2010

Emsley et al. (2010) *Acta Cryst D* 66, 486-501.



# Thaumatococcus, Rigaku R-Axis IV++, MM007HF

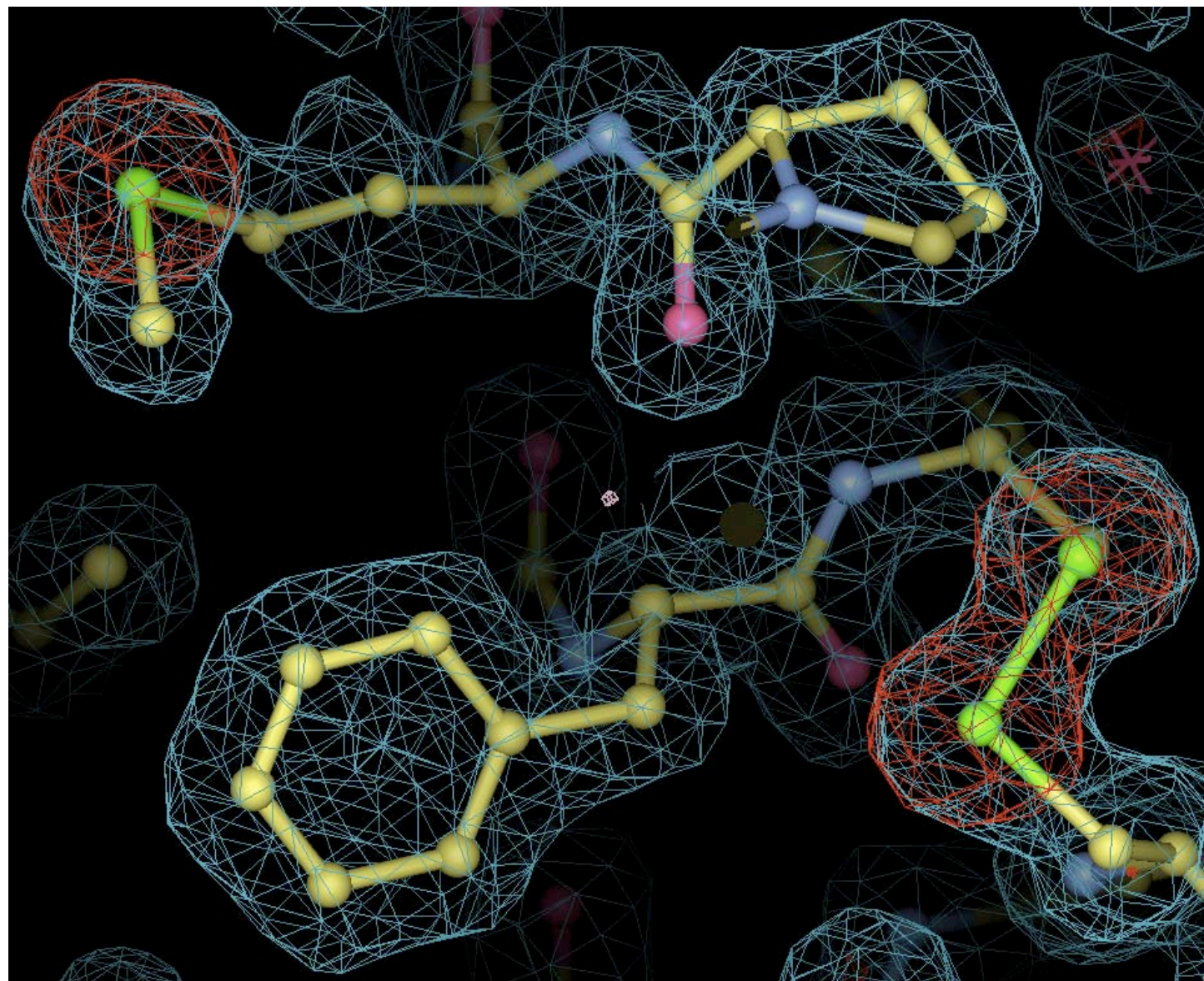
207 aa, 17 sulfurs, 300 images, 90 degrees, 10 hours

## Summary of data collection statistics

---

Spacegroup	P4 <sub>1</sub> 2 <sub>1</sub> 2		
Unit cell dimensions	57.81	57.81	150.10
	90.00	90.00	90.00
Resolution range	19.72 - 1.50		(1.55 - 1.50)
Total number of reflections	243628		
Number of unique reflections	40409		
Average redundancy	6.03		(2.48)
% completeness	96.8		(72.9)
Rmerge	0.029		(0.168)
Rmeas	0.031		(0.208)
RmeasA (I+,I- reflns kept apart)	0.030		(0.201)
Reduced ChiSquared	1.10		(0.80)
Output <I/sigI>	35.7		(5.4)

**Electron density of thaumatin map from ARP/wARP near phenylalanine residue Phe203. Nearby are Pro111, Met112, and the disulfide formed by Cys9 and Cys204.**



# Examples

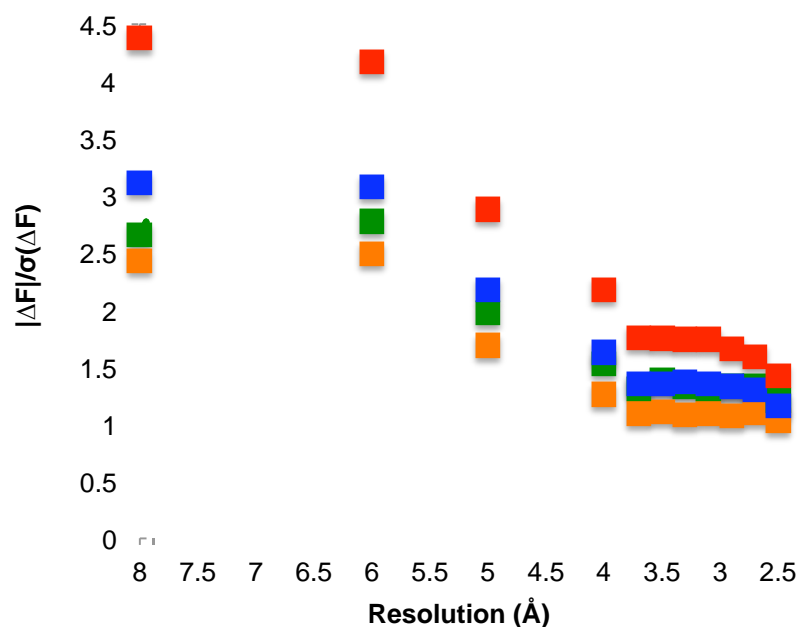
- **Thaumatin**
  - P4<sub>1</sub>2<sub>1</sub>2      58 x 58 x 150
  - 207 residues, 17 sulfurs
- **Lysozyme (30 minute data collection)**
  - P4<sub>3</sub>2<sub>1</sub>2      77 x 77 x 38
  - 129 residues, 10 sulfurs
- **Glucose Isomerase and a 2<sup>nd</sup> one**
  - I222      93 x 98 x 102
  - 388 residues, 1 manganese & 1 calcium

# Thaumatin - Chromium radiation S-SAD

Spacegroup	P4 <sub>1</sub> 2 <sub>1</sub> 2			
Unit cell dimensions	57.95	57.95	150.10	
	90.00	90.00	90.00	
Mosaicity	0.53			
Resolution range	31.70 - 2.50			
Number of 0.3 deg images	600	266	200	150
Exposure time (hours)	5.00	2.21	1.67	1.25
Total number of reflections	119594			30759
Number of unique reflections	9436			9295
Average redundancy	12.67			3.31
% completeness	99.9			98.6
Rmerge	0.032			0.020
Rmeas	0.034			0.024
RmeasA (I+,I- reflns kept apart)	0.030			0.017
Output <I/sigI>	75.0			43.9

# Thaumatococcus - Chromium radiation S-SAD

## Anom S/N vs Reso



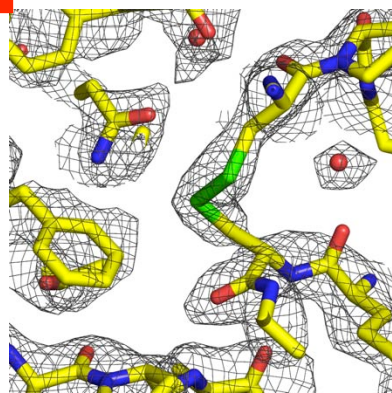
## Chromium Phasing parameters

# of Img	Data Used	Avg. Redundancy	Find Sites Res. (Å)	FOM after DM
<b>600</b>	<b>180°</b>	<b>12.7</b>	<b>2.5</b>	<b>0.85</b>
<b>266</b>	<b>79.8°</b>	<b>5.8</b>	<b>2.5</b>	<b>0.83</b>
<b>200</b>	<b>60°</b>	<b>4.4</b>	<b>2.5</b>	<b>0.83</b>
<b>150</b>	<b>45°</b>	<b>3.3</b>	<b>2.5</b>	<b>0.83</b>

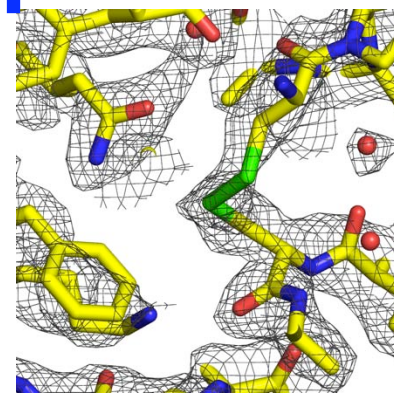
# Thaumatococcus - Chromium radiation S-SAD

Maps after solvent-flattening  
Models from ARP/wARP auto-build

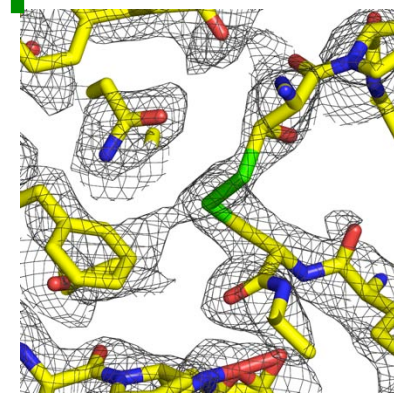
600 (5 hr)



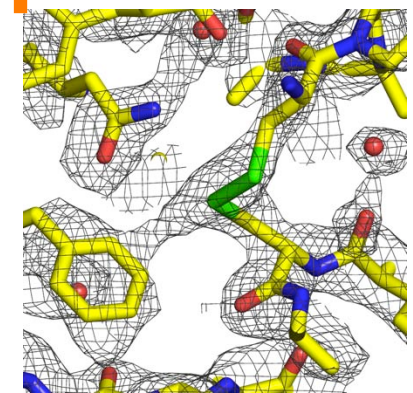
266



200



150 (1.25 hr)

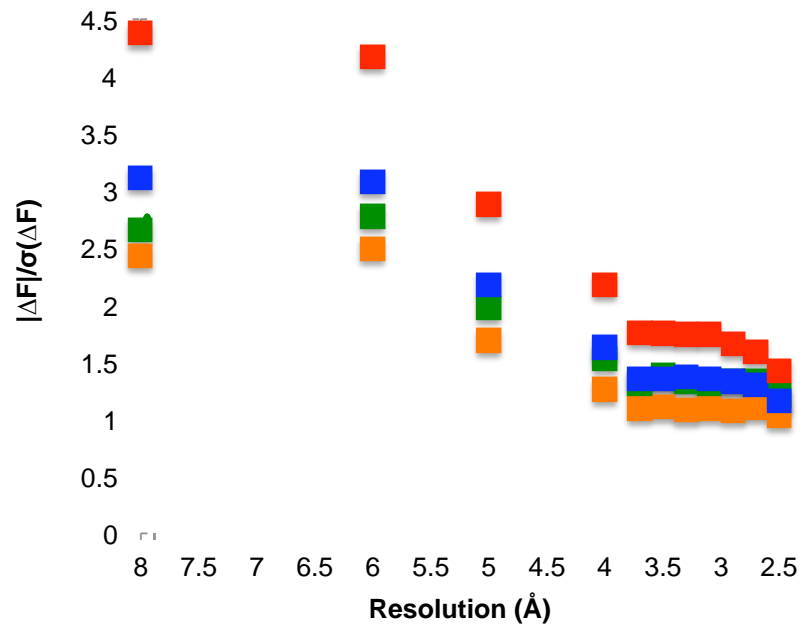


# Thaumatococcus - Copper radiation S-SAD

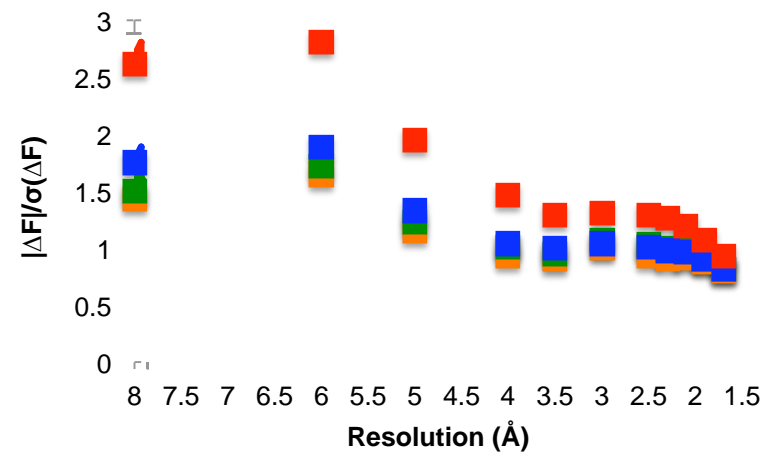
Spacegroup	P4 <sub>1</sub> 2 <sub>1</sub> 2		
Unit cell dimensions	57.88	57.88	150.0
	90.00	90.00	90.00
Mosaicity	0.45		
Resolution range	24.47 - 1.70		
Number of 0.3 deg images	600		150
Exposure time (hours)	5.00	2.21 1.67	1.25
Total number of reflections	384288		89184
Number of unique reflections	28883		28390
Average redundancy	13.30		3.14
% completeness	99.9		98.4
Rmerge	0.035		0.023
Rmeas	0.037		0.028
RmeasA (I+,I- reflns kept apart)	0.036		0.025
Output <I/sigI>	56.7		30.0

# Thaumatin - Cr & Cu radiation S-SAD

## Cr: Anom S/N vs Reso



## Cu: Anom S/N vs Reso





# Lysozyme, Rigaku R-Axis HTC, FR-E+

129 aa, 10 sulfurs, 60 images, 90 degrees, **30 minutes**

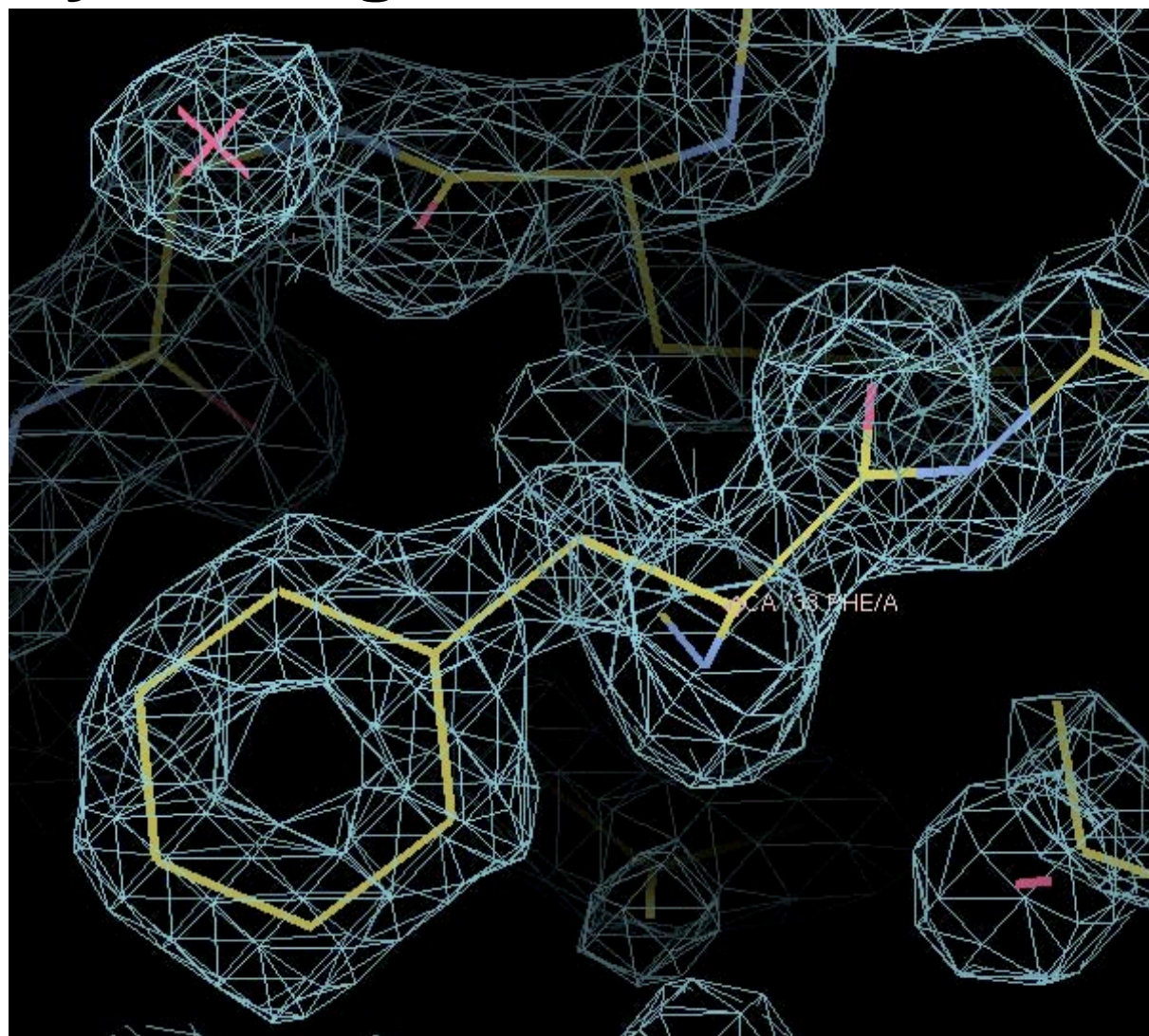
Summary of data collection statistics

```

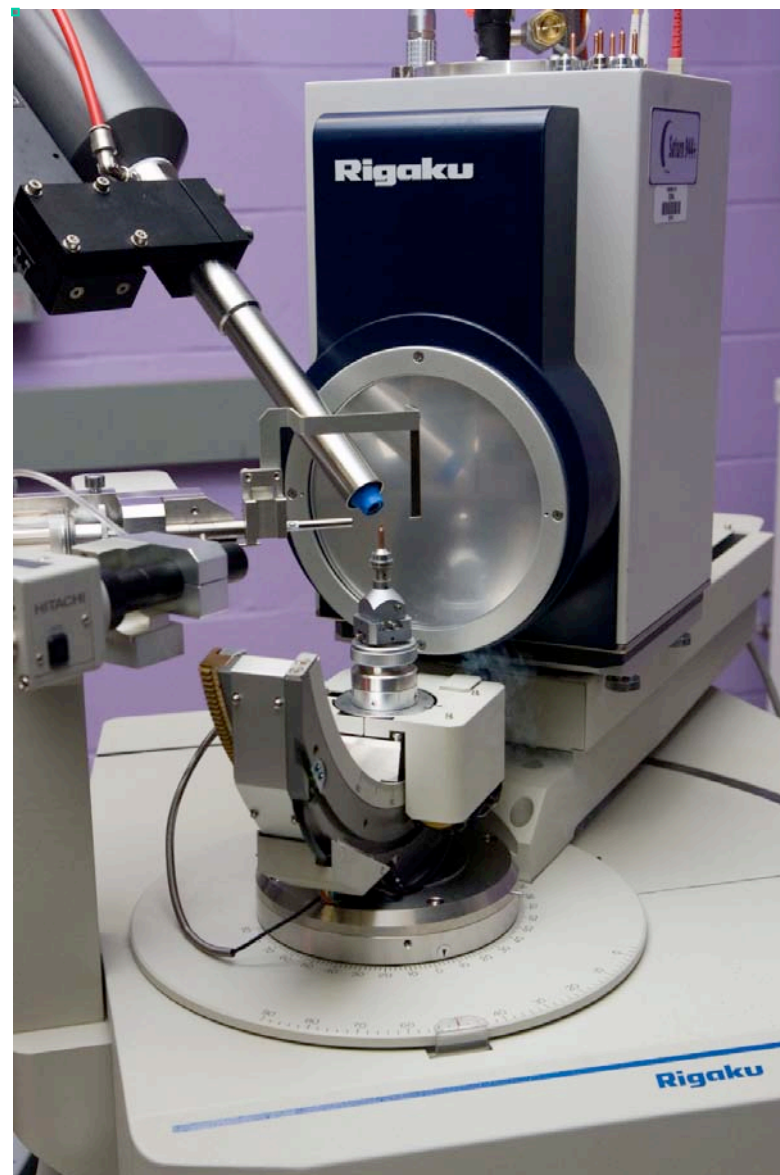
-----
Spacegroup                P43212
Unit cell dimensions      77.56    77.56    38.04
                          90.00    90.00    90.00
Resolution range         18.81 - 1.70    (1.76 - 1.70)
Total number of reflections 88041
Number of unique reflections 13250
Average redundancy        6.64                (6.39)
% completeness           100.0              (99.8)
Rmerge                    0.035              (0.276)
Rmeas                     0.038              (0.300)
RmeasA (I+,I- reflns kept apart) 0.036              (0.299)
Reduced ChiSquared        1.21                (1.04)
Output <I/sigI>           34.2                (6.0)

```

# Lysozyme, Rigaku R-Axis HTC, FR-E+



# Example 3A Glucose isomerase, Rigaku MM007HF, Saturn 944+



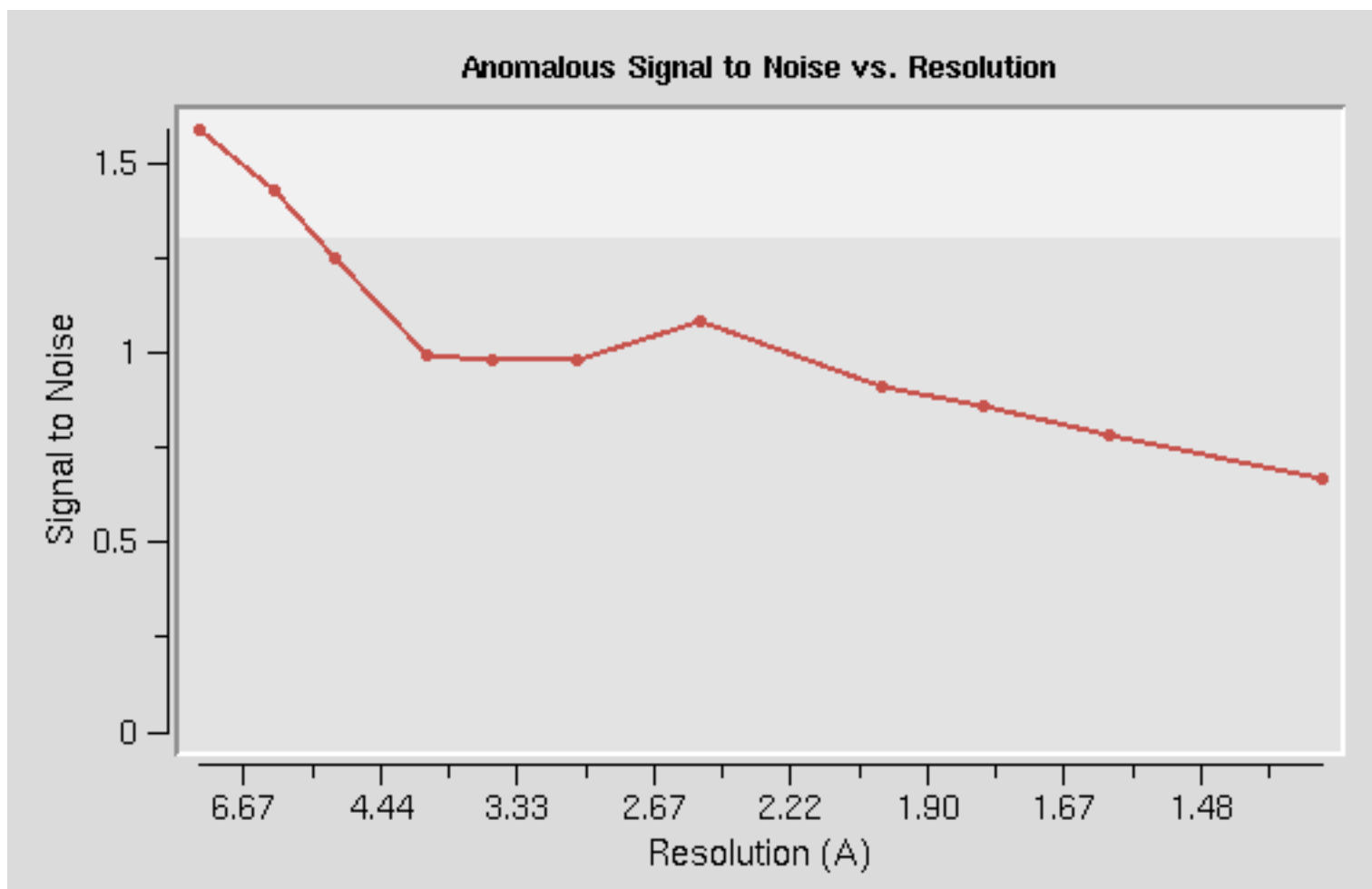
## Example 3 Glucose isomerase, Rigaku MM007HF, Saturn 944+

388 aa, 2 Mn?Ca, 1080 images, 540 degrees, 10 hours

### Summary of data collection statistics

```
-----
Spacegroup                I222
Unit cell dimensions      92.64    98.36    102.46
                          90.00    90.00    90.00
Resolution range         22.90 - 1.40    (1.45 - 1.40)
Total number of reflections 844698
Number of unique reflections 92277
Average redundancy        9.15                (6.64)
% completeness           99.8                (98.4)
Rmerge                   0.047                (0.196)
Rmeas                    0.050                (0.213)
RmeasA (I+,I- reflns kept apart) 0.050                (0.214)
Reduced ChiSquared       1.17                (0.65)
Output <I/sigI>          31.8                (6.8)
```

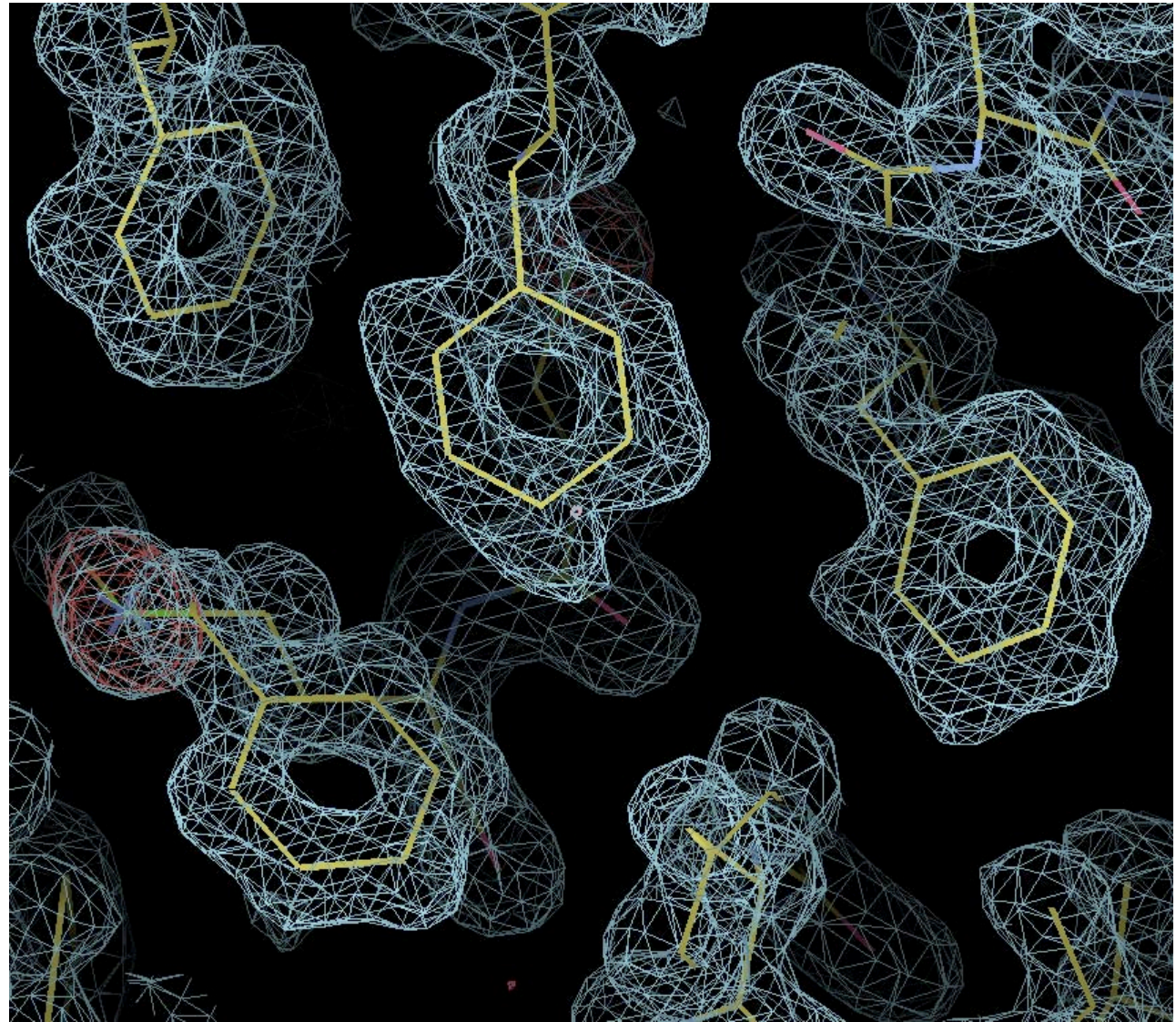
## Example 3 Glucose isomerase, Rigaku MM007HF, Saturn 944+



# Example 3 Glucose isomerase, Rigaku MM007HF, Saturn 944+

Map after  
MLPHARE +  
DM

No refinement

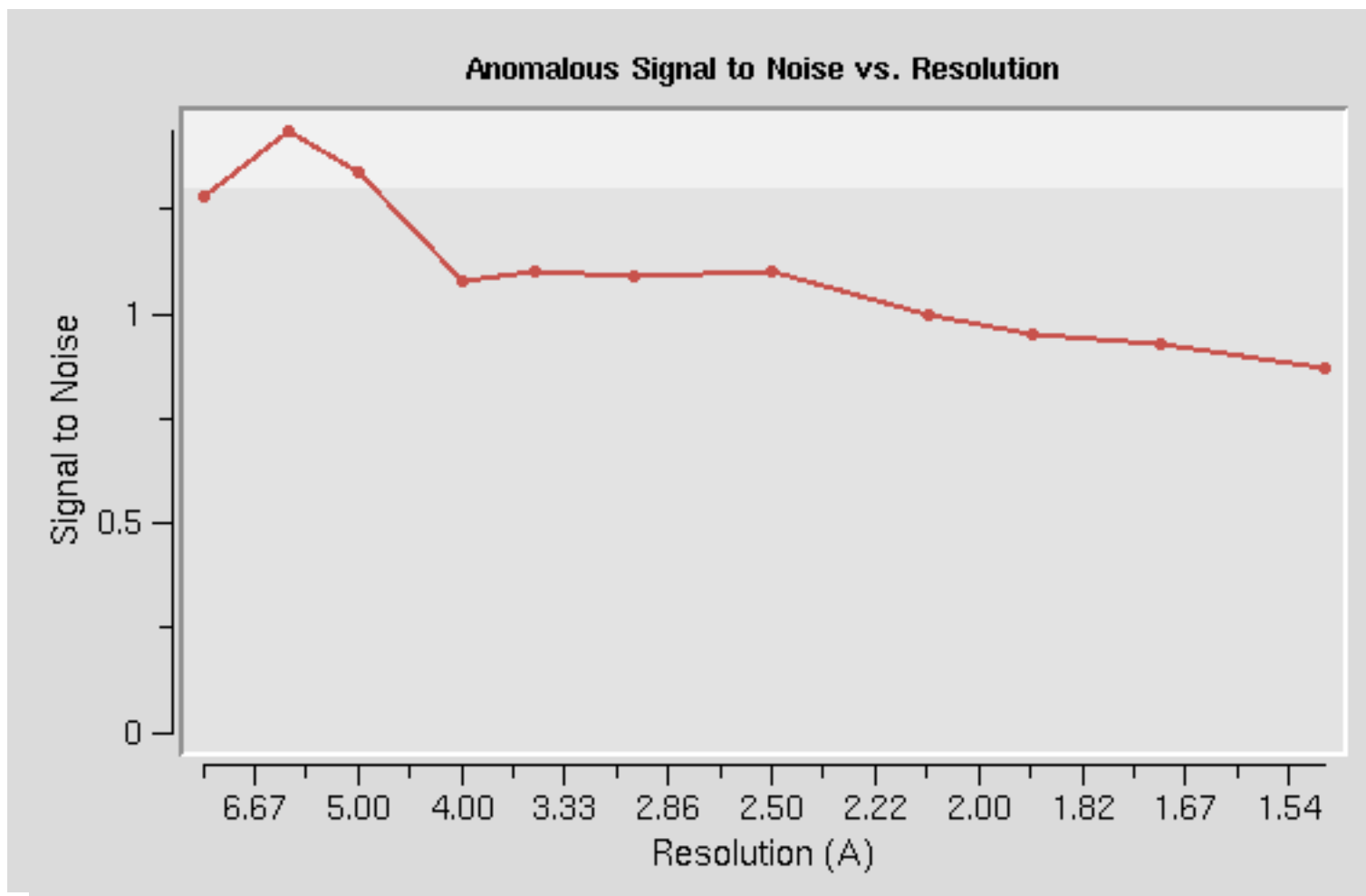


Example 3B Glucose isomerase, Rigaku FR-E+, R-AXIS HTC  
 388 aa, 2 Mn?Ca, 468 images, 344 degrees, 3.9 hours

Summary of data collection statistics

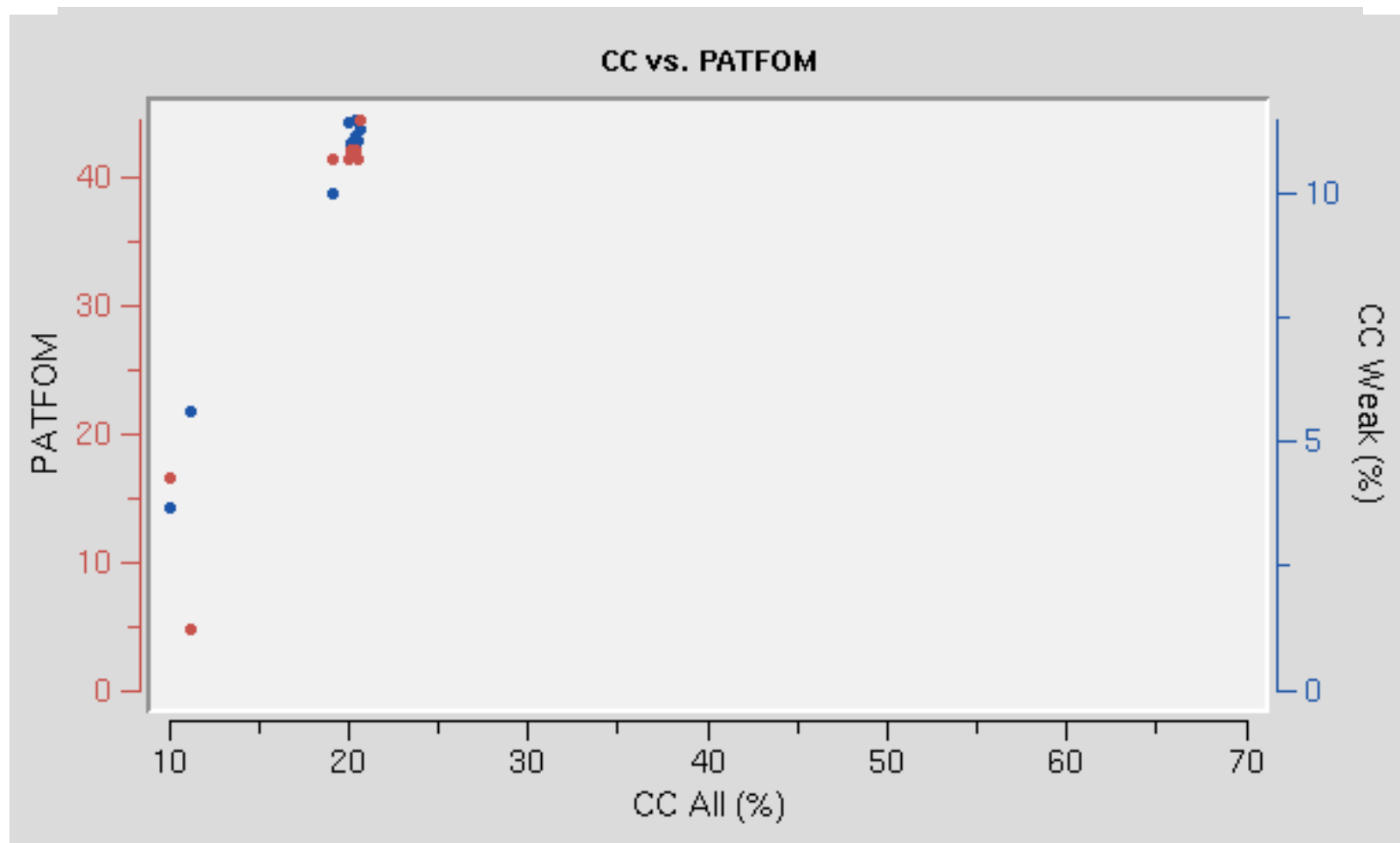
```
-----
Spacegroup                I222
Unit cell dimensions      92.79    97.67    102.56
                          90.00    90.00    90.00
Resolution range         21.61 - 1.50    (1.55 - 1.50)
Total number of reflections 968113
Number of unique reflections 71594
Average redundancy       13.52          (13.32)
% completeness          95.8          (91.3)
Rmerge                   0.053         (0.345)
Rmeas                    0.055         (0.359)
RmeasA (I+,I- reflns kept apart) 0.055         (0.362)
Reduced ChiSquared       1.24          (0.91)
Output <I/sigI>         34.5          (6.0).
```

## Example 4 Glucose isomerase, Rigaku FR-E+, R-AXIS HTC





## Example 4 Glucose isomerase, Rigaku FR-E+, R-AXIS HTC



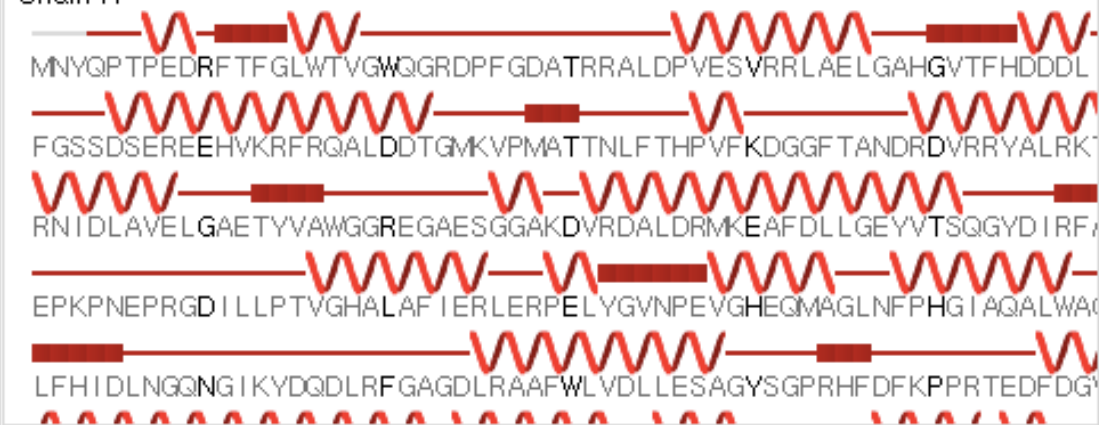
## Example 4 Glucose isomerase, Rigaku FR-E+, R-AXIS HTC

Build main cycle: 3

Build cycle 1: 373 aa (96%) in 6 chains, 126 aa (32%/m) in the longest chain  
 Build cycle 2: 373 aa (96%) in 7 chains, 134 aa (35%/m) in the longest chain  
 Build cycle 3: 373 aa (96%) in 6 chains, 126 aa (32%/m) in the longest chain  
 Build cycle 4: 376 aa (97%) in 5 chains, 134 aa (35%/m) in the longest chain  
 Build cycle 5: 375 aa (97%) in 6 chains, 126 aa (32%/m) in the longest chain  
 Build cycle end: chains 5, 371 aa (96%)  
 5 chains have been docked - 371 aa (96%)  
 Refine cycle 15: dummy atoms included R = 0.273 Rfree = 0.296

> Secondary structure statistics

Chain A



MNYQPTPEDRF TFG LWTVGWQGRDPFGDATRRALDPVESVRR LAELGAHGVTFHDDDL  
 FGSSDSEREEHVKRFRQALDDTGKVPMTTNLF THPVFKDGGFTANDRDVRRYALRK  
 RNIDLAVELGAETYVAVWGGREGAESGGAKDVRDALDRMKEAFDLLGEYVTSQGYDIRF  
 EPKPNEPRGDI LLPTVGHALAFIERLERPELYGVNPEVGHEGMAGLNFPHGIAQALWAG  
 LFHIDLNGQNGIKYDQDLRF GAGDLRAAFWLVDLLESAGYSGPRHFDFKPPRTEDFDG

Refine cycle 16: dummy atoms included R = 0.236 Rfree = 0.266  
 Refine cycle 17: dummy atoms included R = 0.204 Rfree = 0.240

# Conclusions

Do a careful experiment:

Helium beam path for Cr, not necessary for Cu

Redundancy as low as 3 to 4 on I<sup>+</sup>, I<sup>-</sup> with Cu;  
lower with Cr

Even low  $\Delta F/\sigma_{\Delta F}$  can work

SHELXD is amazing!

Any phase improvement is helpful in auto-tracing

# Speculation

Why does low-redundancy with a Home Lab and Imaging Plate work so well?

Stability?

Less demand on synchronization?

No funky detector calibration?

Dose: Does it matter if one can solve the structure with less than 1 minute of exposure?

Photons are photons

# Acknowledgements

- Rigaku Americas
  - Joe Ferrara
  - Vijaya Madakasira
  - Mark Pressprich
  - Angela Criswell
- HKL Research
  - Wladek Minor
  - Marcin Cymborowski
- The Crystallographic Community and all those smart people who created great software!

Thank you!