X-ray studies on multiple single crystals

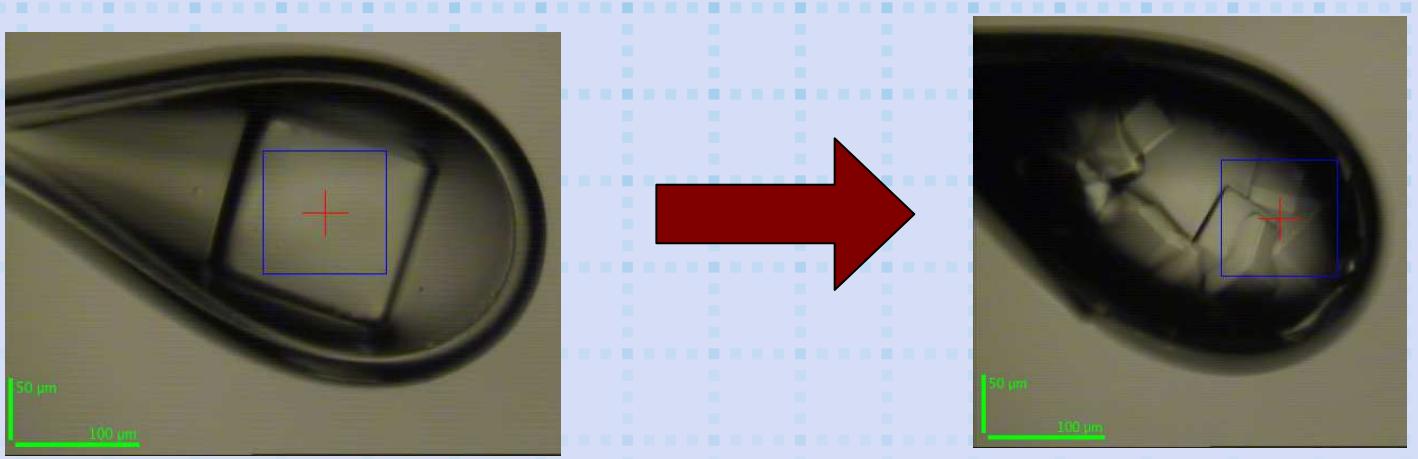
Dr. Karthik S. Paithankar

Laboratory of Molecular Biophysics
Department of Biochemistry
University of Oxford, U.K.

karthik.paithankar@bioch.ox.ac.uk

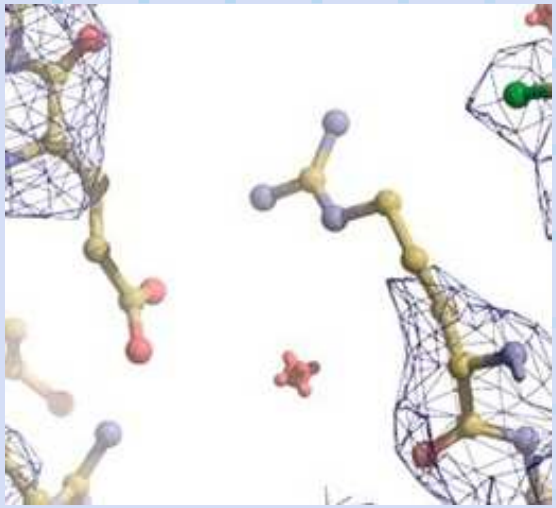
Objectives

- Adapt methods from material sciences to macromolecular crystallography
- Test the feasibility of the new methods for MX to combat radiation damage
- HOWTO: Take multiple crystals (2, 3, ... n) in a single loop and collect data – oriented randomly
- Index the diffraction pattern and utilize the information from all the crystals



Why ? Specific structural damage

Radiation damage in crystallography even at 100 K



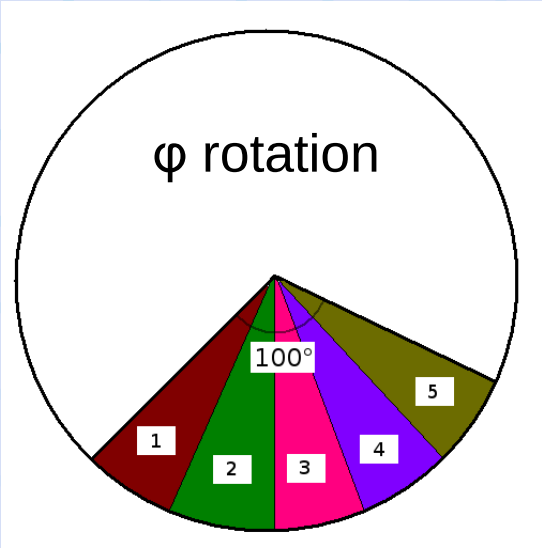
Usual: Collect single crystal dataset for 100°

Proposed: collect 20° of data from 5 different crystals simultaneously and combine them

20° × 5 crystals = 100°

Advantages:

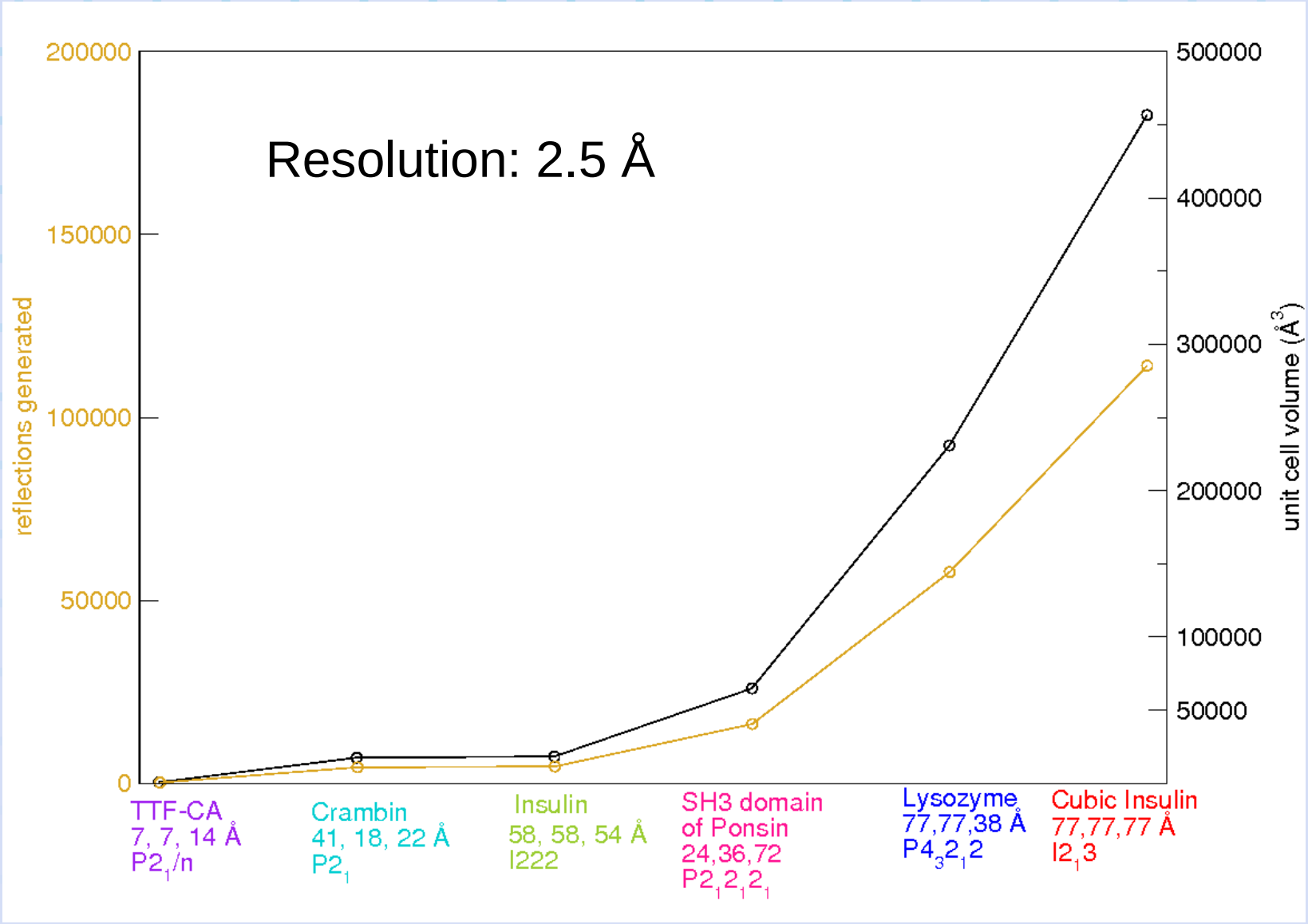
- ✓ All the 5 crystals are at the beginning of decay
- ✓ Lower absorbed dose per crystal
- ✓ Higher quality data
- ✓ Metal centers, active site preserved
- ✓ Extract correct biological information



Potential applications of multiple crystal diffraction

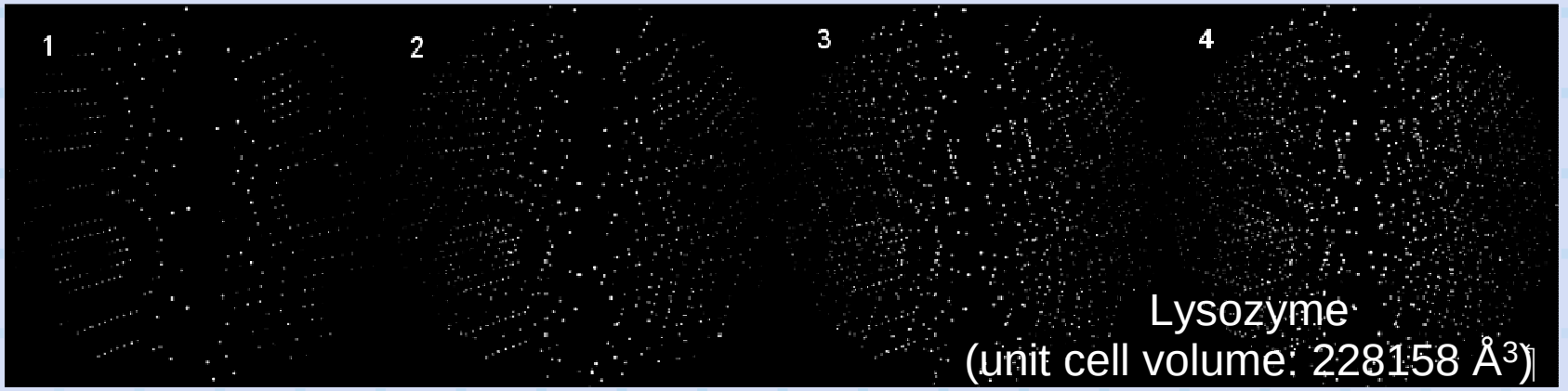
- Micro crystals used in structure solution.
- Minimize crystal handling (less mechanical damage)
- Alleviate radiation damage problems

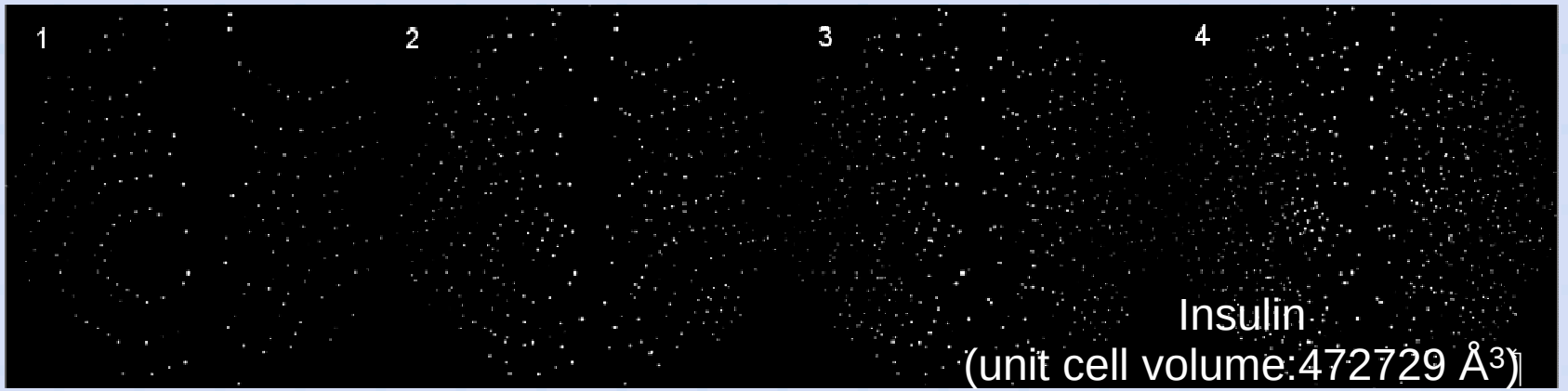
Spot overlap: From material science to MX



Simulations of diffraction patterns (with PolyXSim)

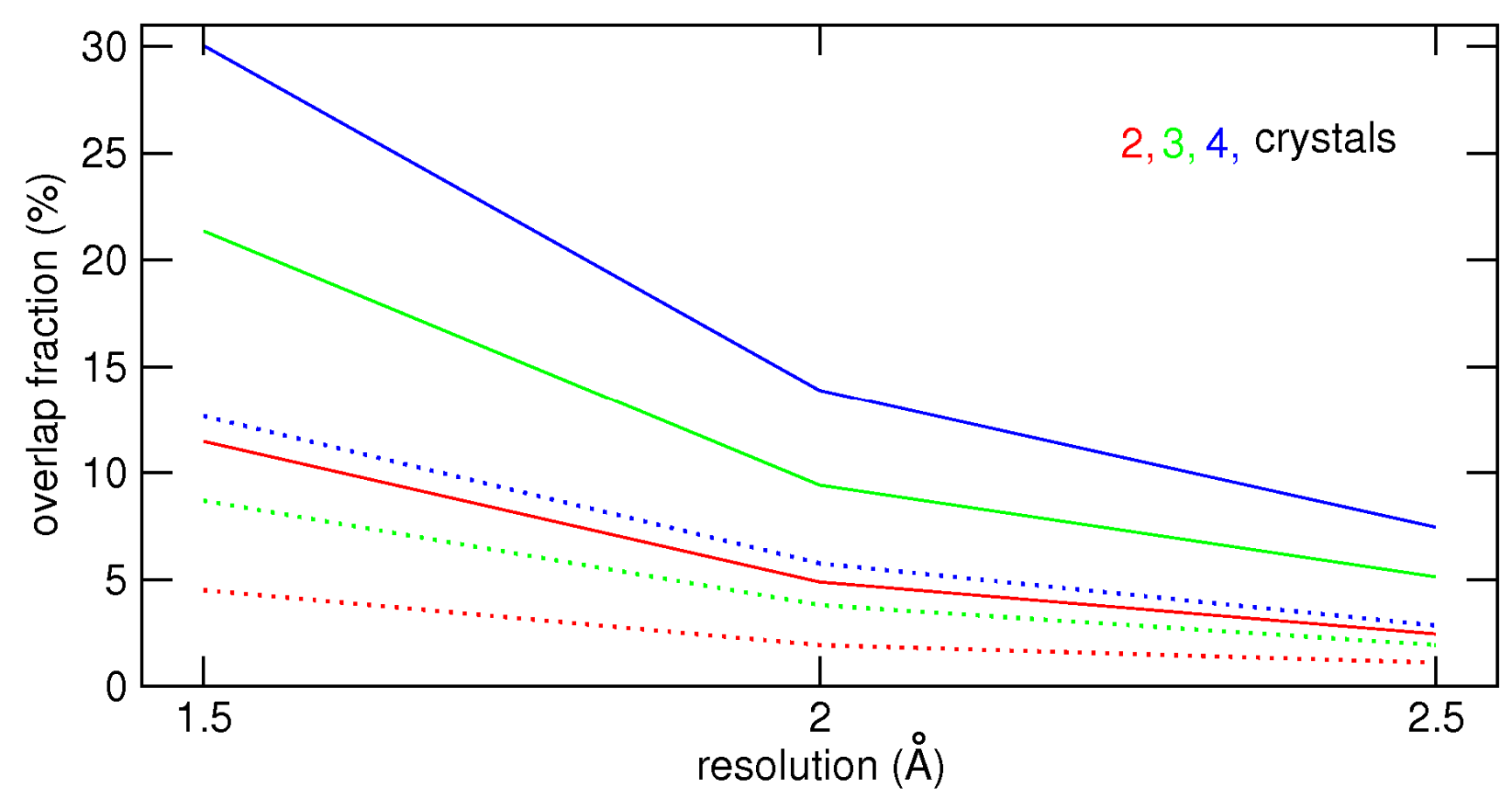
Diffraction patterns with 1, 2, 3 and 4 crystals of lysozyme (top) and insulin (bottom), $\Delta\varphi = 0.5^{o}$



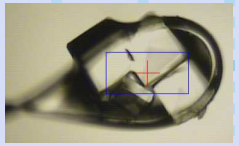


Overlap fraction with multiple crystals

Dotted lines → Lysozyme
Solid lines → Insulin



Overview of the methodology



RADDOSE

Data acquisition

Harvest reflections

Integrated intensities

hkl, I, σI

Indexing

3D – structure

MX tools

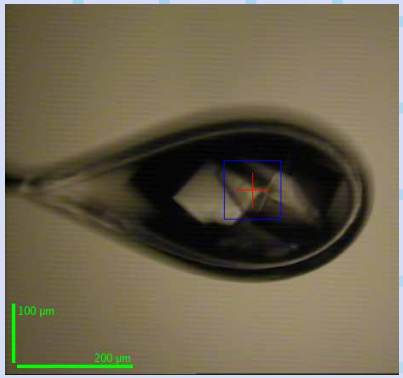
Crystal Orientation

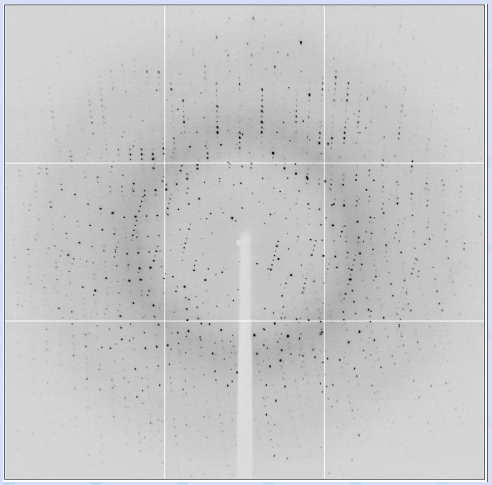
Spot prediction

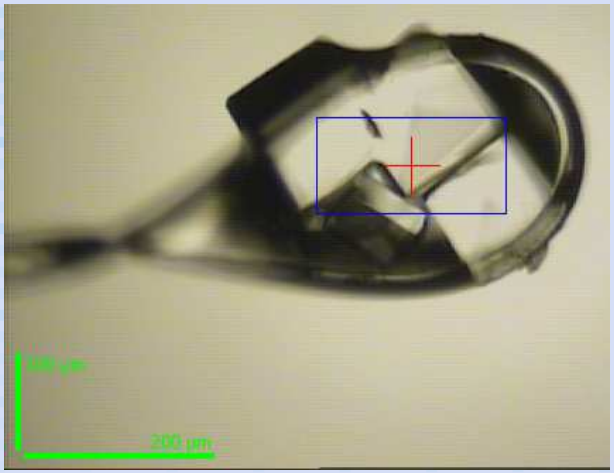
Material scientists

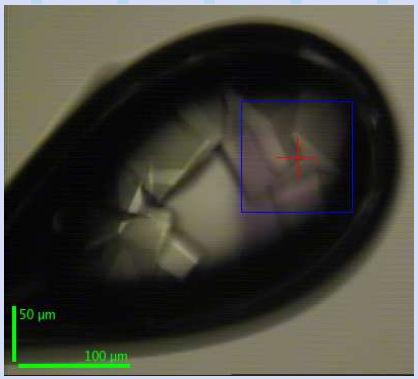
Experiments: Multiple crystals in X-ray beam

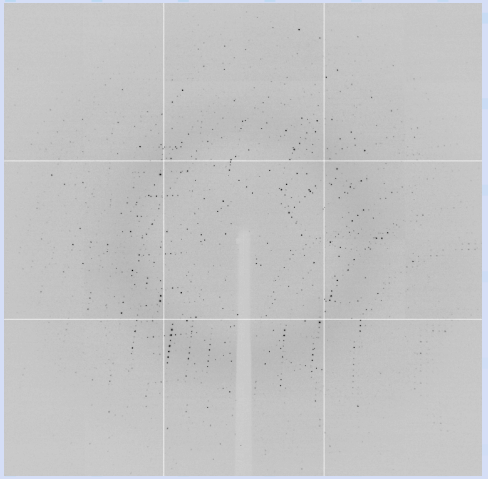
Data of crystals from lysozyme and insulin collected at ESRF (ID 14-4)











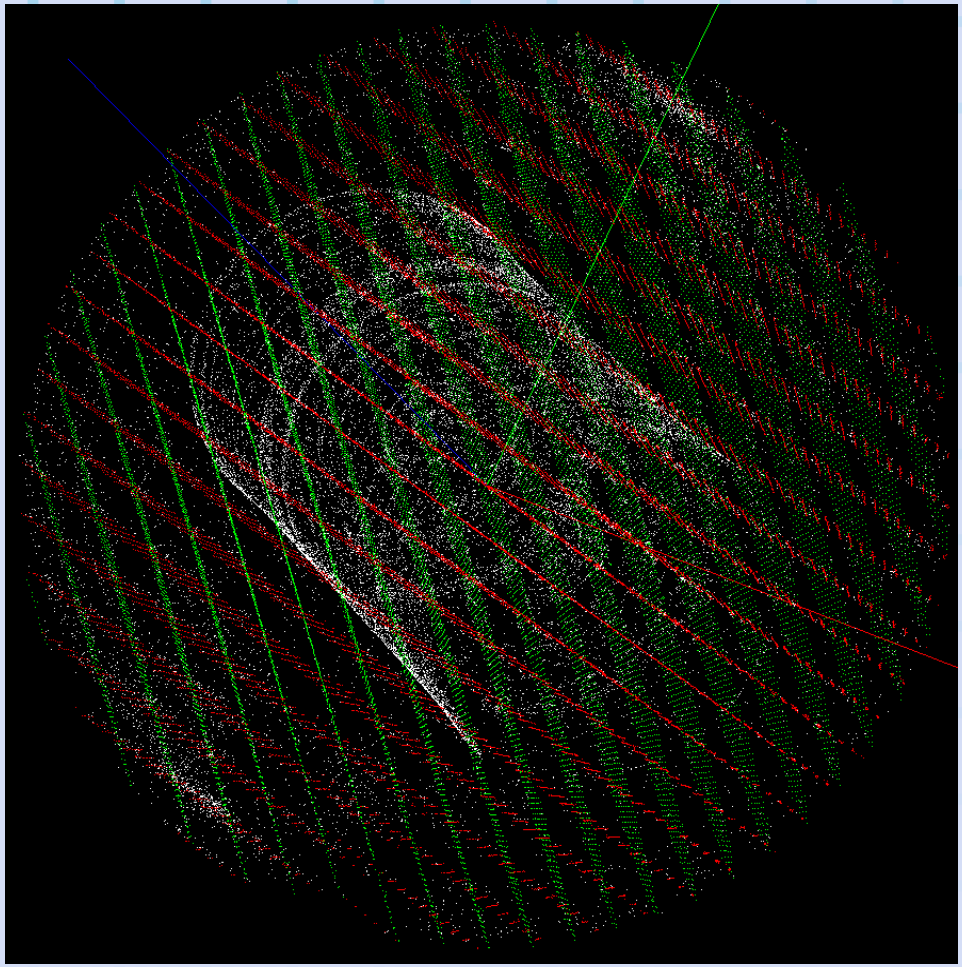
Δφ: 0.5°

Data collection statistics

Protein crystal	Lysozyme	Insulin
Total number of crystals in the exposed	1, 2, 3, 4, 7	1,2,3,4
Wavelength (Å)	0.939	0.939
Space group	$P2_12_12_1$	$I2_13$
Unit-cell dimensions (Å)	78.5 78.5 36.9	78.1 78.1 78.1
Resolution at the detector edge (Å)	1.8	1.8
Oscillation angle (°)	0.5	0.25 – 0.5
Exposure time (s)	1	1

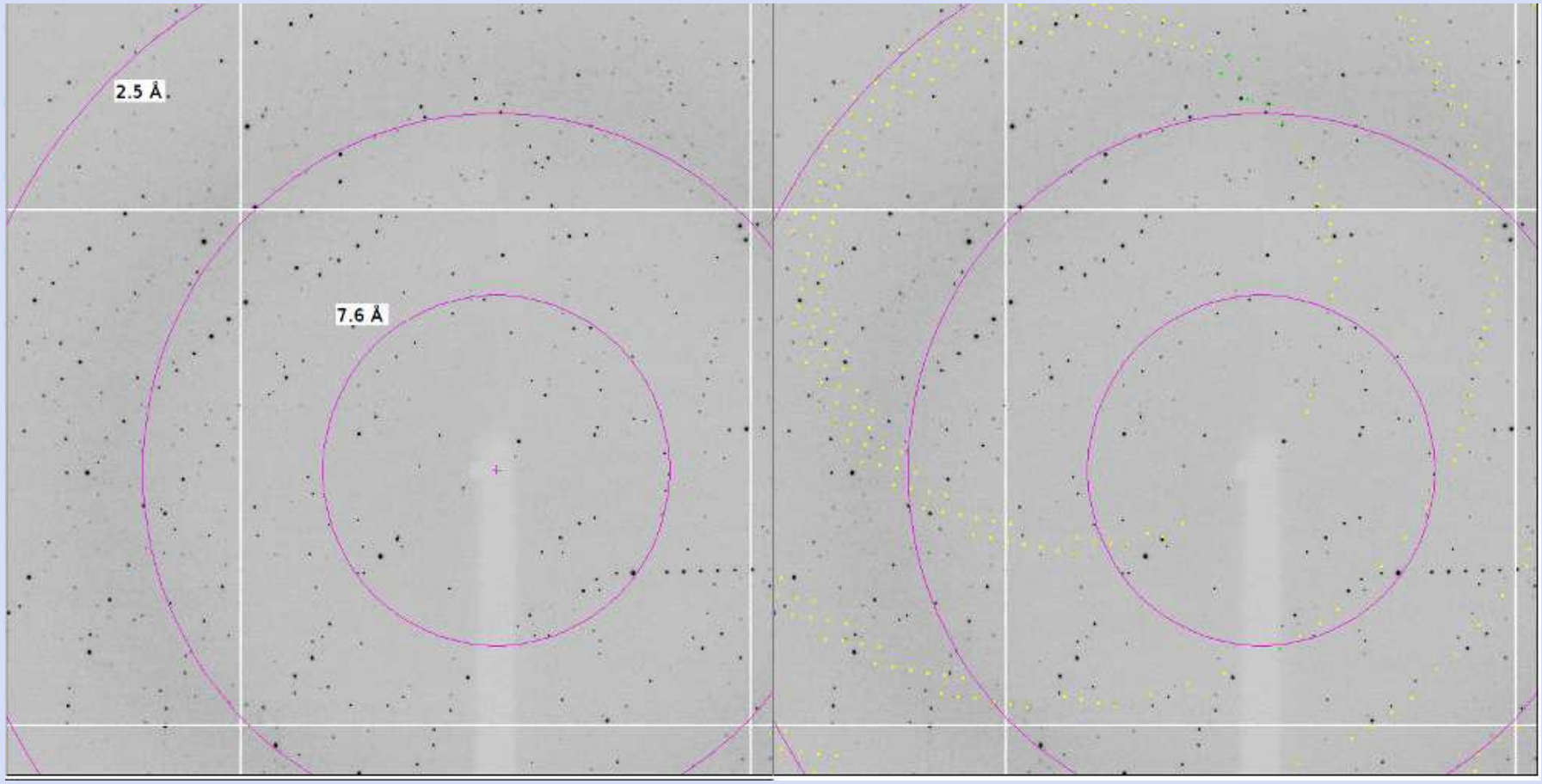
Data reduction in practice

- Search for peaks above a certain threshold and construct an array
- Indexing with Grainspotter
- OUTPUT: orientation of each crystal in the ensemble
- The orientation of each crystal lattice is provided to MOSFLM by means of the U (orientation) matrix
- Two crystal lysozyme
- Red and green indicate the two lattices

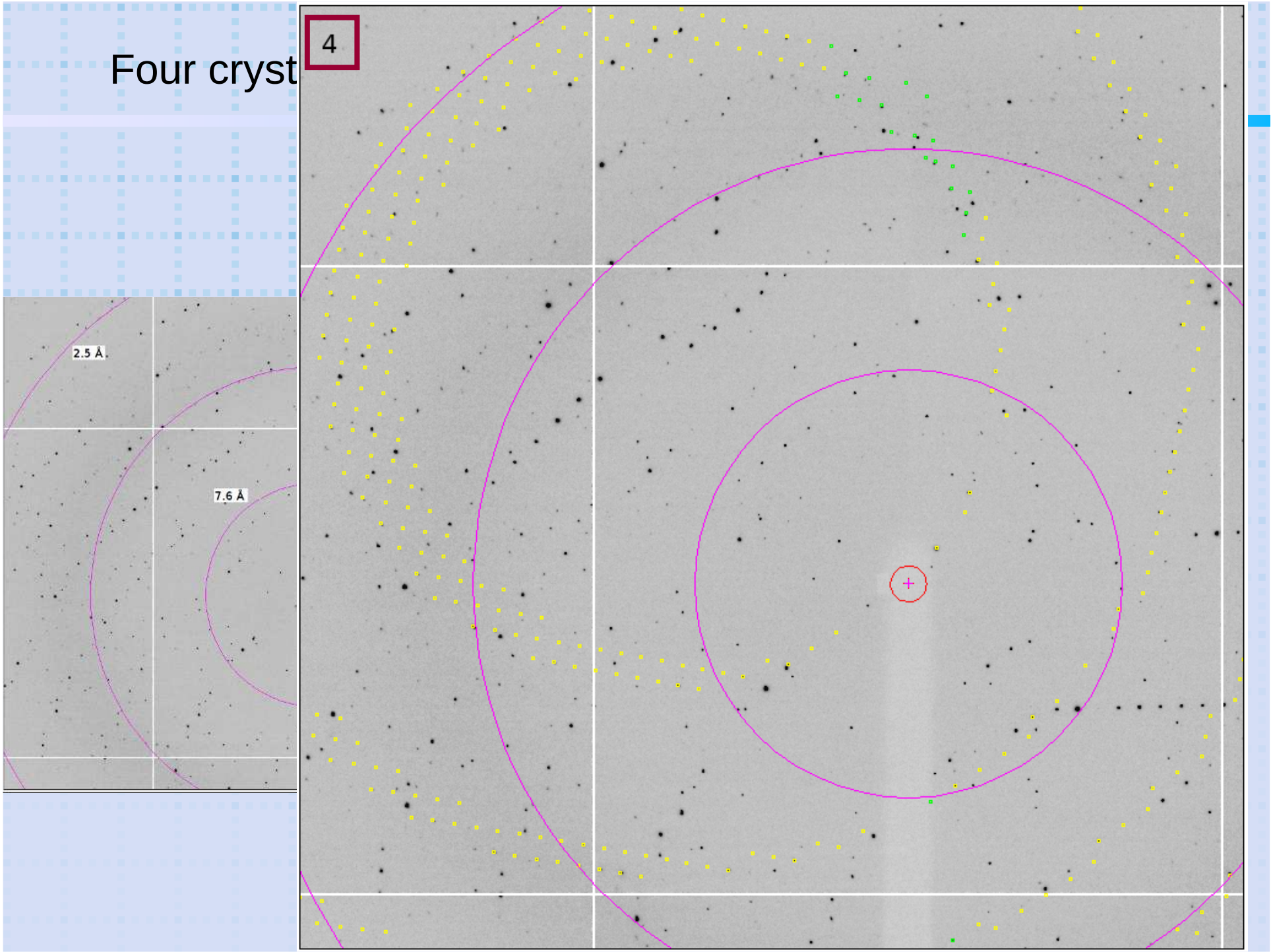


www.totalcryst.dk

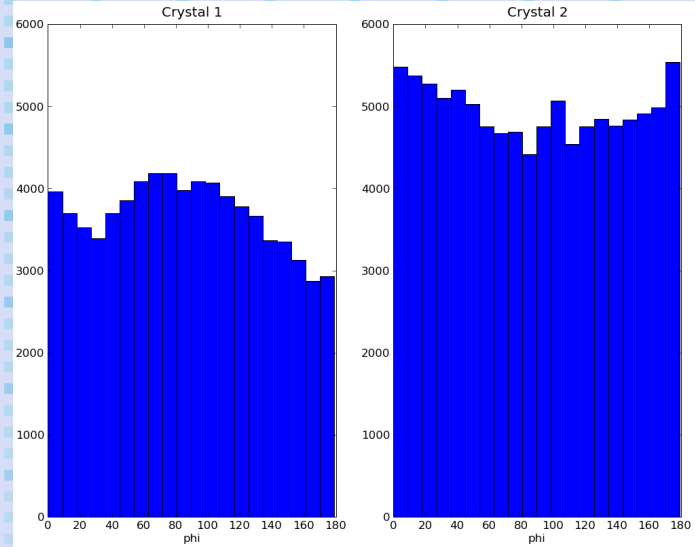
Four crystal lysozyme dataset



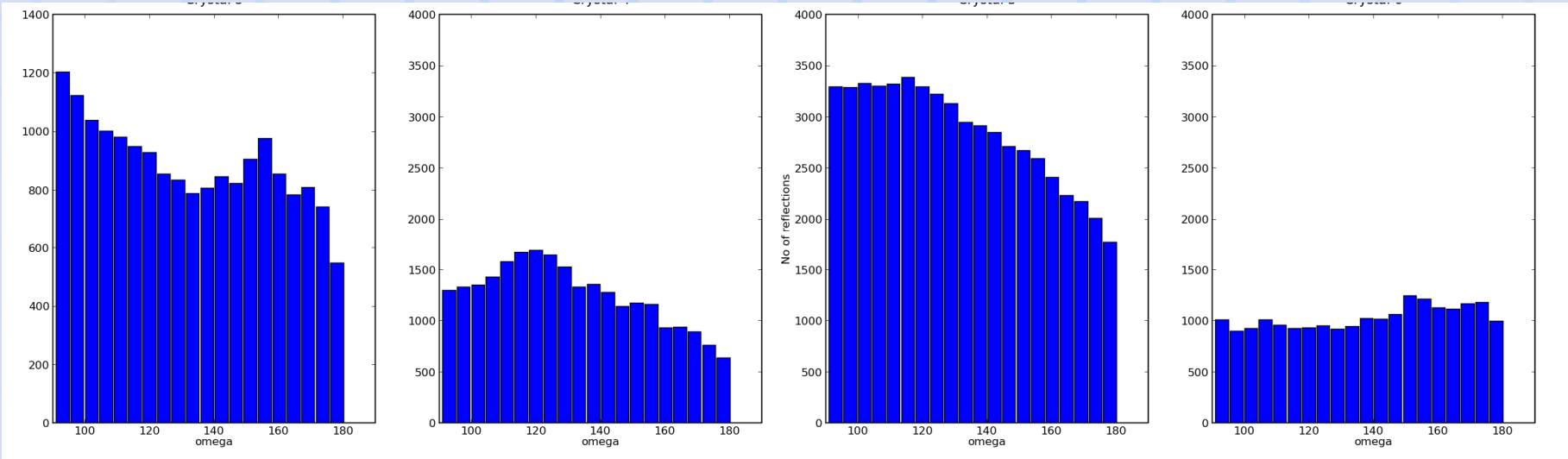
Four cryst



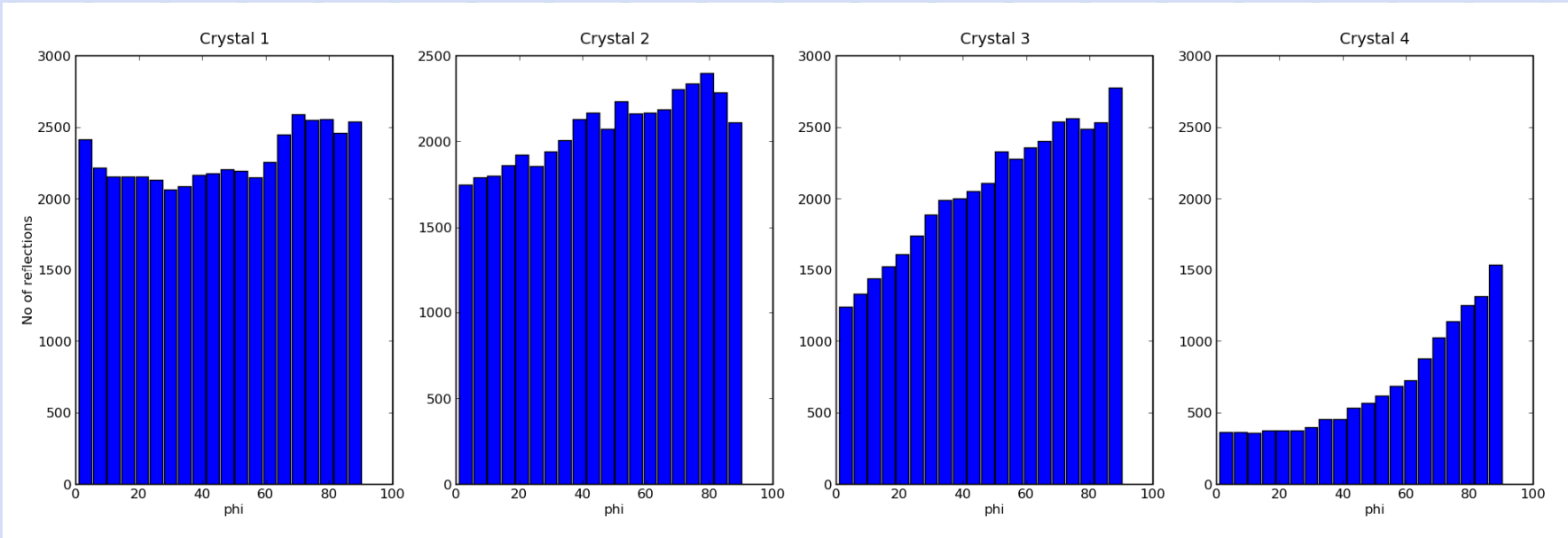
Number of reflections per crystal for insulin 2 crystal ensemble



Number of reflections per crystal for lysozyme 7 crystal ensemble



Number of reflections per crystal for insulin 4 crystal ensemble



Owing to poor diffraction data could not be integrated for this crystal

Compare individual and merged datasets: Insulin 4 crystals

- To determine whether improvement in data could be obtained by merging data from different crystals
- The individual datasets contain more information for ranges studied in spite of random orientation of the different crystals

Crystal number	Image numbers ($\Delta\varphi$= 0.5°)	Completeness	Multiplicity	R_{meas0}	I/σI	Number of reflections: Total (Unique)
1	1-15	60.8(59.4)	1.4(1.4)	0.13(0.3)	8.2(3.3)	4546(3307
2	1-15	60(60)	1.4(1.4)	0.18(0.66)	6.2(1.5)	5369(3748)
3	1-15	62.2(60.5)	1.4(1.4)	0.08(0.2)	11.0(4.4)	5401(3911)
Merged	1-5 (× 3)	53.2(51.9)	1.4(1.4)	0.3(0.7)	4.8(1.5)	4545(3354)
		~14 %				
1	1-60	99.3(99)	3.5(3.6)	0.09(0.2)	15.6(6.1)	19145(5410)
2	1-60	99.4(99)	3.5(3.6)	0.1(0.3)	13.1(5.1)	19146(5420)
3	1-60	98.5(99.4)	3.6(3.7)	0.07(0.2)	16.6(6.7)	22304(6201)
Merged	1-20(× 3)	99(98.9)	3.5(3.5)	0.4(0.6)	8.5(2.5)	21867(6310)
		~1 %				

A fourth crystal was identified but diffracted poorly

Integration of insulin 3 crystals (separately 1º to 180º)

Crystal identifier	I	II	III (1-360)	III (1-180)
Resolution range (Å)	55.1-1.9(2-1.9)	55.1-1.9(2-1.9)	55.1-1.9(2.0-1.9)	
Number of reflections:				
Total	67939	68291	68368	34010
Unique	6399	6396	6393	6344
Completeness	100(100)	100(100)	100(100)	99.8(100)
Multiplicity	10.6(10.9)	10.7(11)	10.7(11)	5.4(5.5)
R_{meas}	0.1(0.3)	0.07(0.12)	0.25(1.1)	0.1(0.16)
$I/\sigma I$	17.8(7.1)	29.7(19.4)	17.2(11.9)	14.6(8.9)
Wilson B factor ($Å^2$)	16.6	18.4	17.7	17.0

Statistics for 3 crystals from the lysozyme 4 crystals ensemble

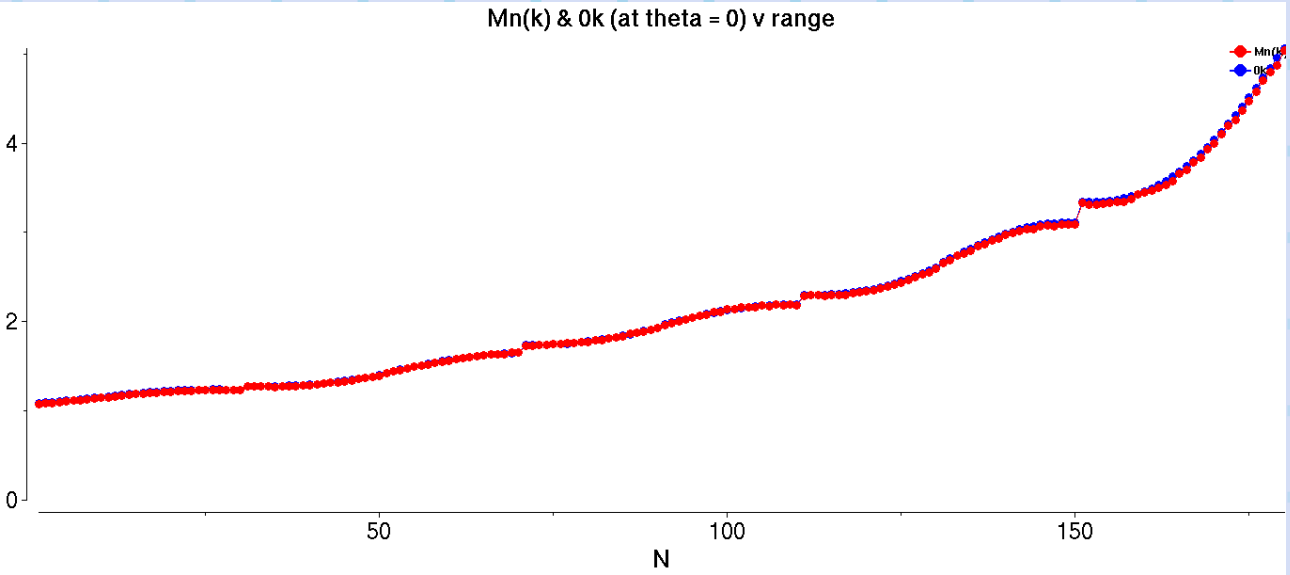
Crystal	Image range ($\Delta\varphi$= 0.5°)	Completeness	Multiplicity	R_{meas0}	I/σI	Number of reflections: Total (Unique)
1	1-36	76.3(78.1)	1.8(1.8)	0.14(0.4)	5.3(2.1)	12846(7217)
1	37-72	66.8(70.1)	2.0(2.0)	0.16(0.5)	4.3(1.8)	12846(6329)
1	73-108	69.7(71.4)	1.9(2.0)	0.17(0.5)	4.2(1.9)	12746(6556)
1	109-144	72.0(73.7)	1.9(1.9)	0.16(0.5)	7.1(2.1)	12887(6680)
1	145-180	81.2(81.6)	1.7(1.7)	0.12(0.35)	5.0(2.7)	12835(7552)
2	1-36	58.7(60.4)	2.3(2.3)	0.04(0.1)	15.4(7.2)	12705(5581)
2	37-72	72.1(74.6)	1.9(1.9)	0.07(0.12)	11.8(6.0)	12812(6699)
2	73-108	73.9(77.7)	1.9(1.8)	0.08(0.15)	9.7(5.0)	12887(6842)
2	109-144	70.6(71.3)	1.9(2.0)	0.08(0.17)	9.9(5.3)	12889(6627)
2	145-180	51.6(52.7)	2.7(2.7)	0.07(0.17)	12.1(6.1)	12890(4822)
3	1-36	37.9(40.7)	3.3(3.4)	0.27(1.9)	4.1(0.8)	12973(3887)
3	37-72	59.5(60.6)	2.3(2.3)	0.12(0.36)	6.6(2.6)	13122(5822)
3	73-108	72.1(76.8)	1.9(1.8)	0.07(0.16)	11.4(6.0)	13023(6691)
3	109-144	71.0(72.3)	2.0(2.0)	0.05(0.09)	14.9(8.8)	12949(6587)
3	145-180	65.8(67.2)	2.1(2.1)	0.06(0.12)	16.8(8.8)	13079(6155)

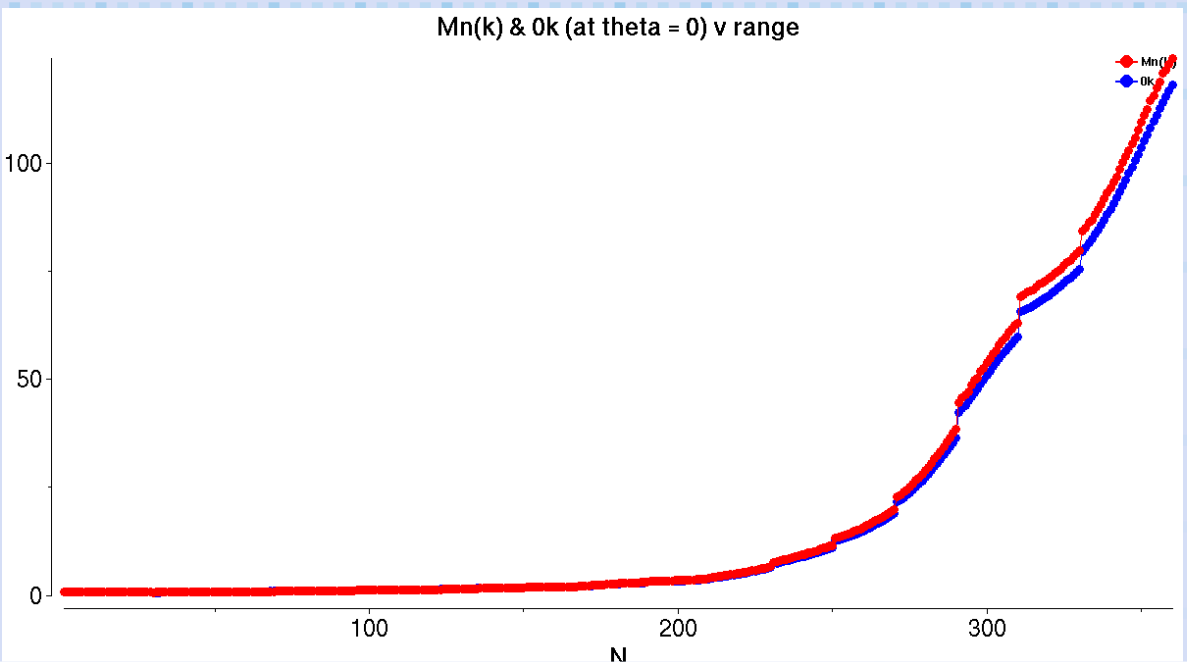
Lysozyme: 78.3 78.3 36.9 90° 90° 90°

Statistics for 3 crystals from the Insulin 3 crystals ensemble

Crystal	Image range ($\Delta\varphi$= 0.5°)	Completeness	Multiplicity	R_{meas0}	I/σI	Number of reflections: Total (Unique)
1	1-72	91(93)	2.4(2.3)	0.1(0.3)	7.9(3.1)	13490(5730)
1	73-144	92.3(93.4)	2.3(2.3)	0.1(0.3)	7.4(3.4)	13506(5836)
1	145-216	93.6(93)	2.3(2.3)	0.1(0.36)	7.9(2.7)	13549(5964)
1	217-288	91.5(93.1)	2.3(2.3)	0.1(0.36)	7.3(2.8)	13431(5796)
1	289-360	96(95.6)	2.2(2.3)	0.1(0.4)	7.1(2.9)	13530(6098)
2	1-72	90.1(91.7)	2.4(2.4)	0.08(0.15)	11.7(6.1)	13552(5698)
2	73-144	96.9(97.1)	2.2(2.2)	0.06(0.1)	13.1(7.4)	13595(6134)
2	145-216	90(94.1)	2.4(2.3)	0.07(0.1)	14.7(8.8)	13562(5650)
2	217-288	92.5(91.7)	2.3(2.4)	0.06(0.1)	16(10.8)	13565(5845)
2	289-360	92.5(89.6)	2.3(2.4)	0.06(0.08)	18.5(13)	13564(5917)
3	1-72	84(88.9)	2.6(2.5)	0.07(0.1)	14.5(9.1)	13627(5274)
3	73-144	93.4(94.8)	2.3(2.3)	0.07(0.1)	13.6(7.5)	13619(2016)
3	145-216	97.5(98.5)	2.2(2.2)	0.1(0.2)	8.9(4.8)	13588(6220)
3	217-288	97.2(99)	2.2(2.2)	0.3(0.99)	3.5(1.4)	13452(6213)
3	289-360	82.9(82.5)	1.7(1.7)	1.7(5.3)	0.4(0.1)	8700(5200)

Crystal 3 probably damaged at exposure 200





Compare total and merged datasets: Insulin 3 crystals

Crystal identifier	I	II	III (1-360)	III (1-180)
Resolution range (Å)	55.1-1.9(2-1.9)	55.1-1.9(2-1.9)	55.1-1.9(2.0-1.9)	
Number of reflections:				
Total	67939	68291	68368	34010
Unique	6399	6396	6393	6344
Completeness	100(100)	100(100)	100(100)	99.8(100)
Multiplicity	10.6(10.9)	10.7(11)	10.7(11)	5.4(5.5)
R_{meas}	0.1(0.3)	0.07(0.12)	0.25(1.1)	0.1(0.16)
I/σI	17.8(7.1)	29.7(19.4)	17.2(11.9)	14.6(8.9)
Wilson B factor (Å^2)	16.6	18.4	17.7	17.0

Image range X 3 crystals	Completeness	Multiplicity	R_{meas0}	I/σI	Number of reflections: Total (Unique)
1-72	99(99)	6.4(6.5)	0.37(0.45)	17.4(8.7)	40665(6390)
73-144	99.9(100)	6.4(6.5)	0.36(0.43)	18.2(8.5)	40710(6385)
145-216	100(100)	6.3(6.5)	0.38(0.45)	16.5(8.9)	40685(6410)
217-288	100(100)	6.3(6.4)	0.4(0.7)	16.3(9.1)	40387(6411)

HEWL 7 crystals analysis

All seven crystals were uniquely identified but only four of them could be integrated

Crystal	Image range ($\Delta\varphi$= 0.5°)	Completeness	Multiplicity	R_{meas0}	I/σI	Number of reflections: Total (Unique)
4	1-36	68.9(70.6)	2.1(2.0)	0.04(0.06)	20.0(11.8)	11355(5511)
4	37-72	57.4(62.4)	2.4(2.3)	0.04(0.06)	23.8(14.6)	11362(4647)
4	73-108	64(67.2)	2.2(2.1)	0.04(0.07)	20(12.1)	11366(5166)
4	109-144	71.4(73.2)	2.0(2.0)	0.06(0.12)	12.9(7.5)	11309(5668)
4	145-180	69(72)	2.0(2.0)	0.07(0.17)	11.3(5.3)	11049(5472)
5	1-36	62.7(65.6)	2.0(1.6)	0.3(1.3)	2.2(0.7)	11988(5875)
5	37-72	63.7(67.4)	2.0(1.6)	0.12(0.5)	5.9(1.4)	12267(5986)
5	73-108	72.7(69.7)	1.8(1.5)	0.2(0.8)	2.9(1.0)	12312(6803)
5	109-144	70(66)	1.7(1.4)	0.6(3.6)	1(0.2)	11519(6617)
5	145-180	68.7(55.80	1.8(1.7)	0.8(5.7)	0.8(0.1)	11811(6544)
6	1-36	82.6(68.3)	1.6(1.5)	0.12(0.3)	6.1(2.4)	12575(7789)
6	37-72	71.6(69.8)	1.8(1.5)	0.2(0.6)	3.3(1.4)	12516(68170
6	73-108	80.9(70)	1.6(1.5)	0.3(1.1)	2.5(0.8)	12255(7562)
6	109-144	78.8(73)	1.6(1.4)	0.4(1.7)	1.6(0.5)	11999(7350)
6	145-180	52.3(47.1)	1.5(1.3)	2.2(8.1)	1.5(1.3)	7436(4837)
7	1-36	67.6(61.7)	1.8(1.6)	0.49(1.9)	1.2(0.3)	11439(6406)
7	37-72	74.3(67.6)	1.6(1.4)	0.5(2.2)	1.0(0.3)	11056(6789)
7	73-108	71.2(62.4)	1.6(1.5)	0.5(2.6)	0.8(0.2)	10236(6451)
7	109-144	64.3(53)	1.7(1.7)	0.6(5)	0.6(0.1)	10453(5991)
7	145-180	59.7(55.1)	1.9(1.6)	0.5(22.8)	0.8(0.1)	10583(5643)

Refinement statistics

Low resolution limit	55.73	55.73	2.00
High resolution limit	1.90	6.00	1.90
R-merge	0.100	0.043	0.291
R-meas	0.107	0.046	0.312
Observations			
Total	65257	2149	9473
unique	8598	303	1261
Mean((I)/sd(I)	19.9	40.6	9.3
Completeness	89.5	83.5	91.4
Multiplicity	7.6	7.1	7.5

20 cycles minimization using REFMAC5 (no waters):

cell	78.8 78.8 36.9
space group	$P4_32_12$
solvent content	38 %
No. of protein atoms	1001 (129 residues)
B-factor	9.4 Å2
R-factor/R-free	24.4/29.8
rmsd bond lengths	0.02 Å
angles	1.7 °

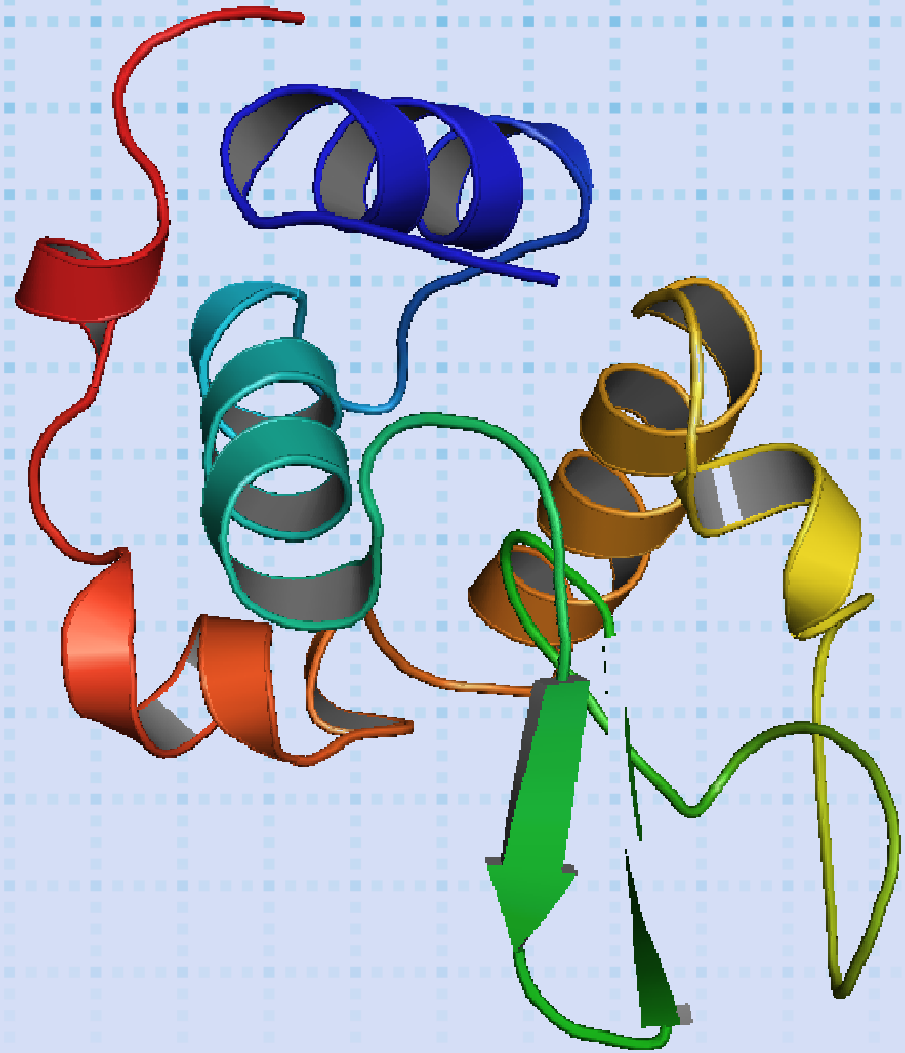
Structure determination of Lysozyme using 2/4 crystals (2 x45°)

Integration

Molecular replacement

Rigid-body refinement

Restrained refinement



3D- structure of lysozyme

R-factor/ R-free

24.4/29.8

MX tools

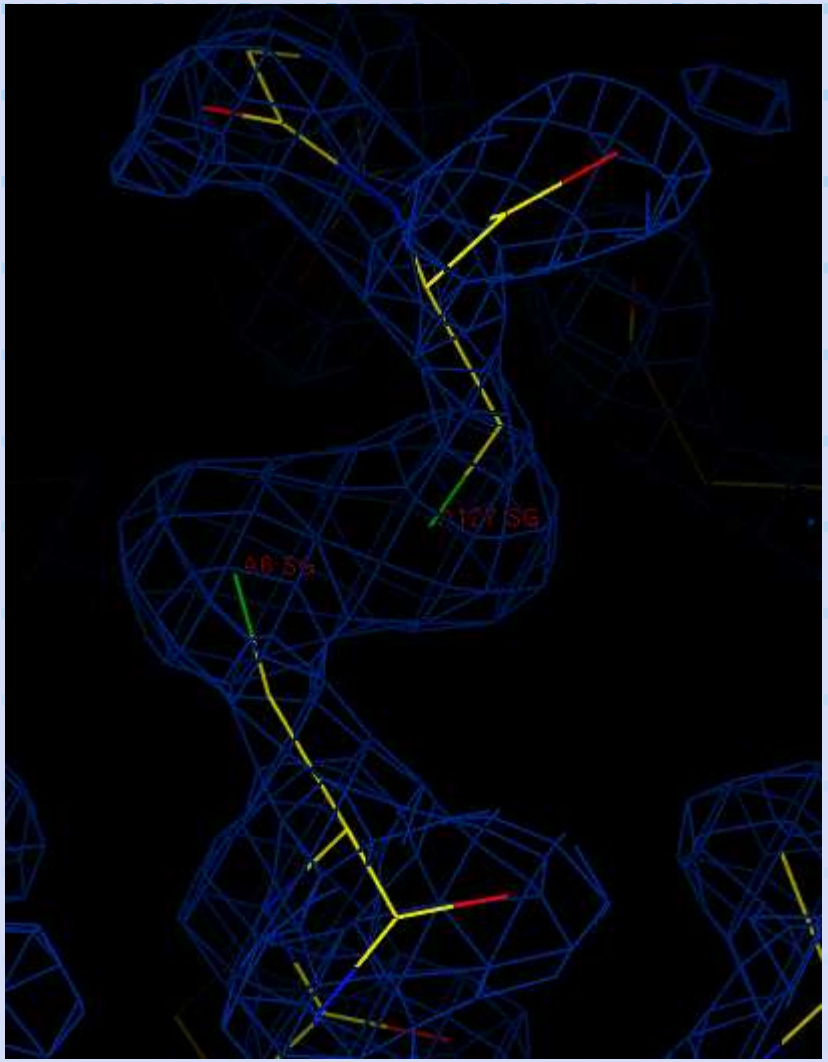
Structure determination of Lysozyme using 2/4 crystals (2 x45°)

Integrated intensities

Molecular replacement

Rigid-body refinement

Restrained refinement



Disulphide bridge

A6 – A127

2fo -fc map at

1.3 σ

Image created

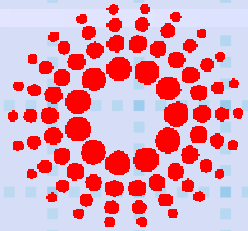
with ONO

MX tools

Summary – X-ray studies on multiple crystals

- Simulations show spot overlap is not huge problem; due to random overlaps and not systematic overlap
- Data were collected with multiple single crystals in a single loop
- Possible to index unknown lattice but collecting few exposures with only one crystal or select an exposure with one strong lattice
- Extraction of data from all the different crystals achieved
- Since different orientations lose different spots - the completeness falls up to about 14 % (in the range studied)
- The combination of data from multiple crystals compensates for the loss of redundancy owing to rejected spots
- Structure solution from multiple crystals of Lysozyme (and Insulin)
- FUTURE PROSPECTS: Practical application of this methodology to data from micro-crystals obtained from a protein with unknown structure

Acknowledgements



European Synchrotron Radiation Facility

Henning Osholm Sørensen & Jon Wright

All of Team – Total Cryst

BAG beam time

SIXTH FRAMEWORK PROGRAMME

Elspeth F. Garman