Are X-rays damaging to structural crystallography?

A case study with Xylose Isomerase

Edward Snell

Hauptman-Woodward Medical
Research Institute

Outline

- Where do X-rays go when you illuminate a crystal (Radiation Damage)
- Synchrotron Radiation
- Radiation damage in structural biology
- An example with xylose isomerase
- High resolution data collection
- Where does radiation damage enter the equation
- Implications
- Neutron data collection

Where do the X-rays go when you illuminate a crystal?

Où les rayons X disparaissent quand vous illuminez un cristal ?

Radiation Damage (dommages de rayonnement)

- 1 Å X-ray interaction in a crystal
  - 90% of the X-rays pass straight through (the reason for the beam stop).
  - 8.4% interact by the photoelectric effect. All the X-ray energy is transferred to an electron which is then ejected (main process of radiation damage).
  - 0.8% interact through Compton scattering. The X-ray transfers some of its energy to an atomic electron and a second lower energy photon is released. This forms the incoherent background.
  - 0.8% interact through Thomson (Rayleigh) scattering elastically with no energy loss. This is the X-ray that gives diffraction data.

Processes of radiation damage

Primary, secondary, direct and indirect radiation-damage events in a protein crystal.

The incoming X-ray photons cause primary damage events, represented by darker stars. The paths of secondary radicals are shown by dotted arrows, and the damage events they induce are represented by lighter stars. Direct events occur on the protein molecules, and indirect events occur in the solvent region.

Primary effects are a fact of life, we cannot prevent them. Secondary effects are reduced by cryocooling.

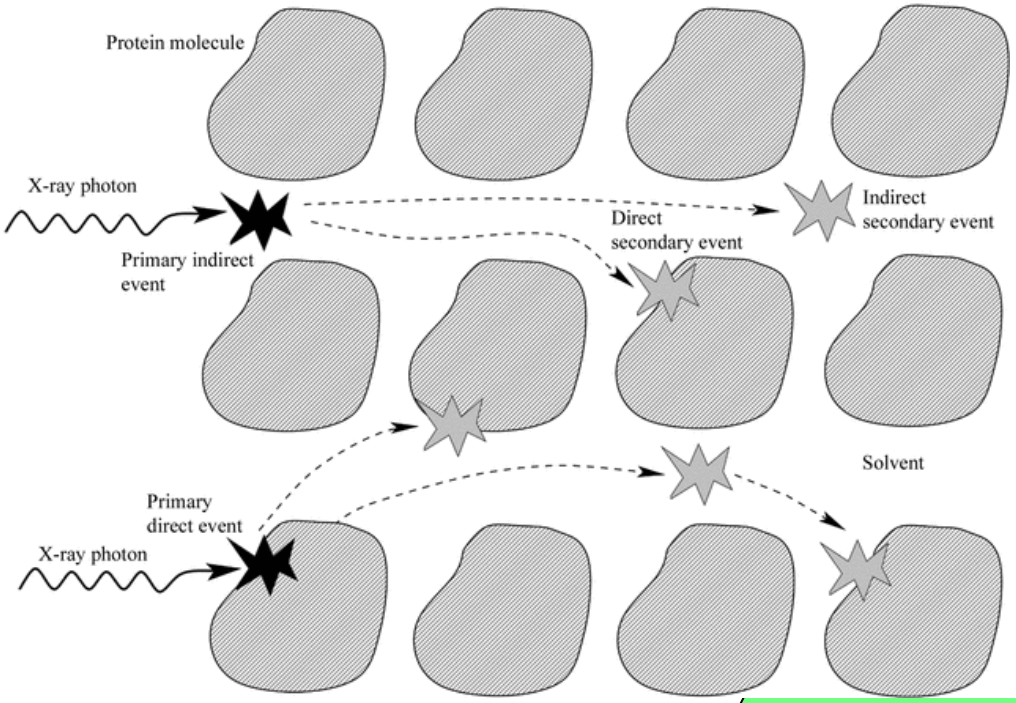
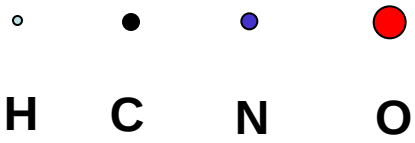
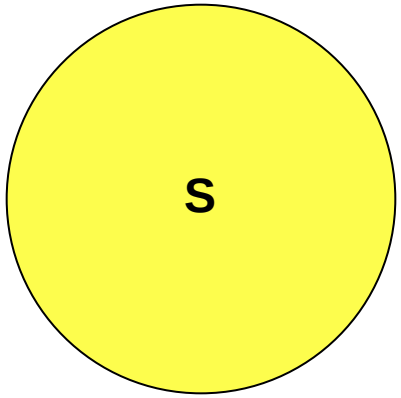
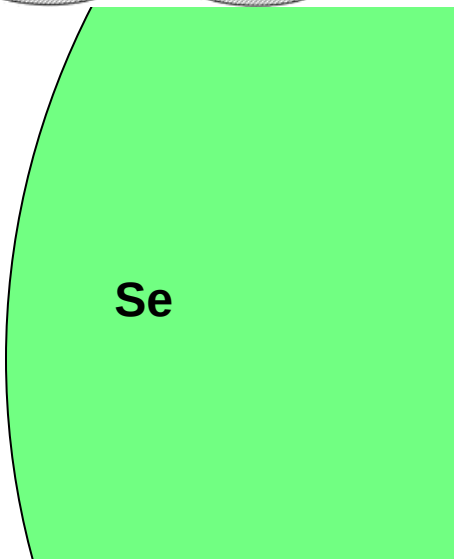


Image from Elspeth Garman

Effect on Other atoms

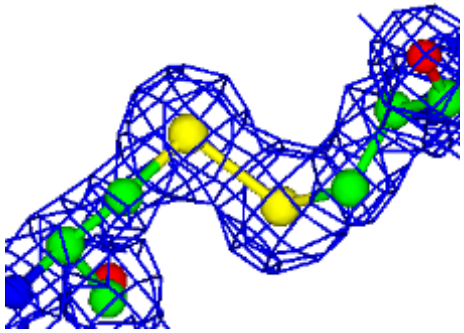






The chemistry: Mobile e- elec affinic sites

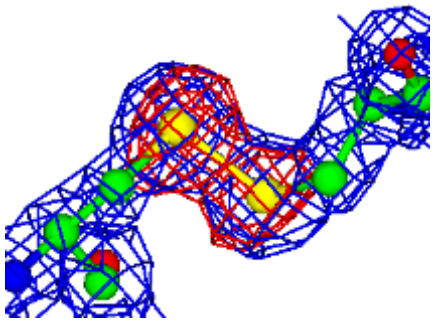
Electron capture	$RSSR + e^-$	$\rightarrow [RSSR]^{\bullet -}$ (400 nm peak)
Disproportionation	$[RSSR]^{\bullet -}$	$\rightarrow RS^- + RS^\bullet$
Protonation	$[RSSR]^{\bullet -} + H^+$	$\rightarrow RSH + RS^\bullet$
Electron loss	$RSSR$	$\rightarrow [RSSR]^{\bullet +} + e^-$
Alkyl loss	$RSSR + e^-$	$\rightarrow RSS^- + R^\bullet$



Specific structural damage

DISULPHIDE BONDS (S-S) MOST SUSCEPTIBLE

Weik *et al* (2000) PNAS 97, 623-628
Burmeister (2000), Acta Cryst D56, 328-341.
Ravelli and McSweeney, (2000) Structure 8, 315-328



X-ray Radiation effect on water

Ionizing radiation can remove an electron from water:

$$H_2O^+ + H_2O \longrightarrow H_3O^+ + OH$$

And the ejected electron

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

The simultaneous formation of H and OH free radicals gives further reactions

$$H + OH \longrightarrow H_2O$$

$$H + H \longrightarrow H_2$$

$$OH + OH \longrightarrow H_2O_2$$

Synchrotron Radiation

Rayonnement de synchrotron

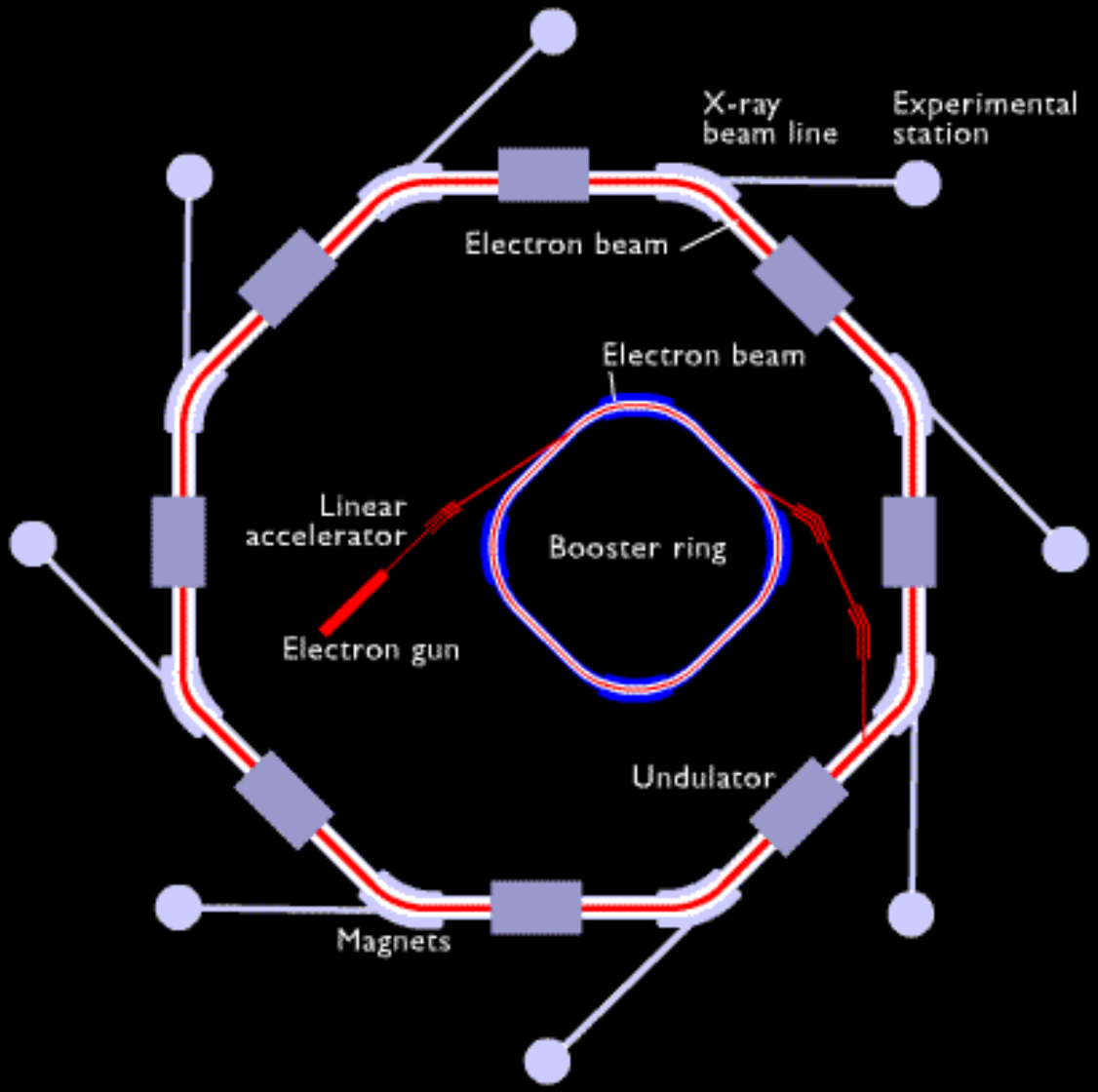
A synchrotron accelerates and stores particles (electrons or protons) moving at speeds close to that of light.

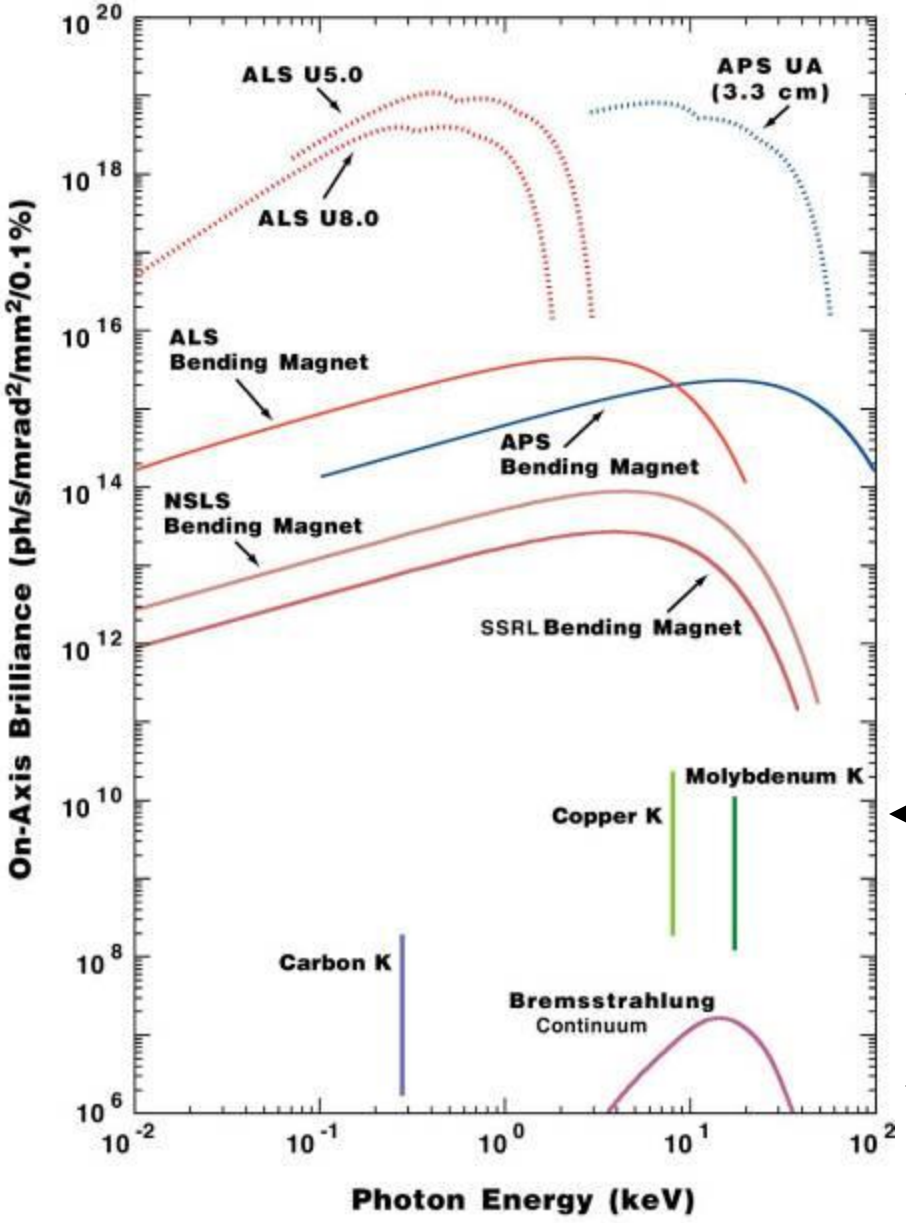
As the particles loose energy they give of electromagnetic radiation.

The particles are steered by magnetic fields.

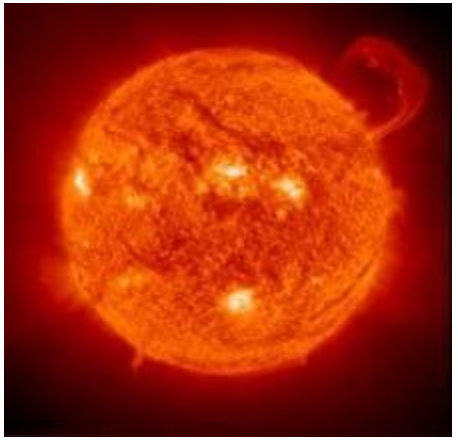
Electromagnetic radiation (photons) is not affected by these fields and is emitted at the tangent to the change in direction.

Insertion devices (undulators and wigglers) ‘amplify’ this radiation





Synchrotron radiation is 10^9 times
More brilliant than the sun
and about 100 million miles closer

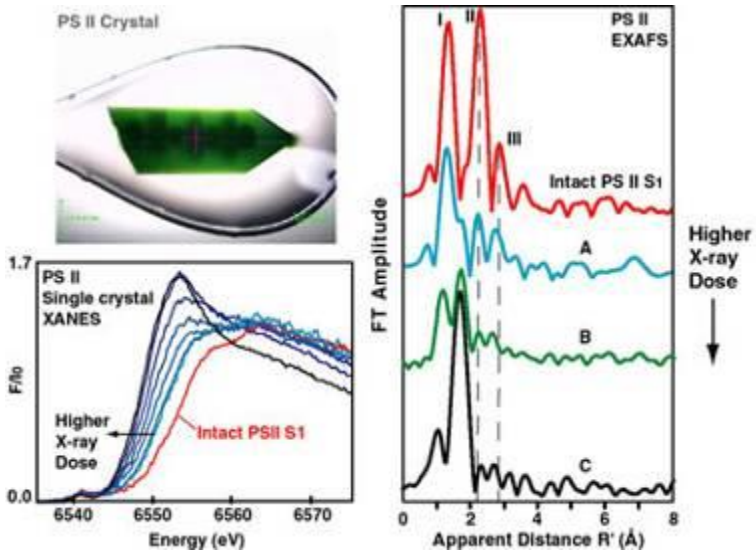


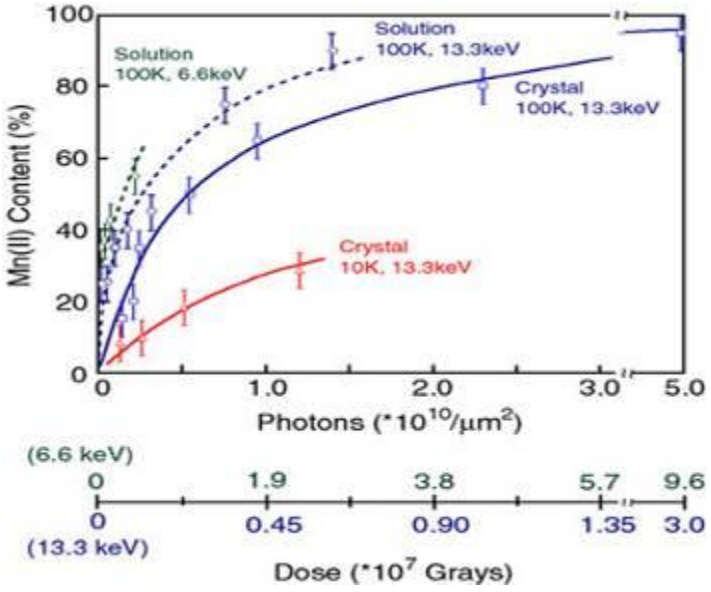
10.97

Radiation Damage in Structural Biology

Exemples de dommages de rayonnement dans la biologie structurale

Case study - Photosystem II





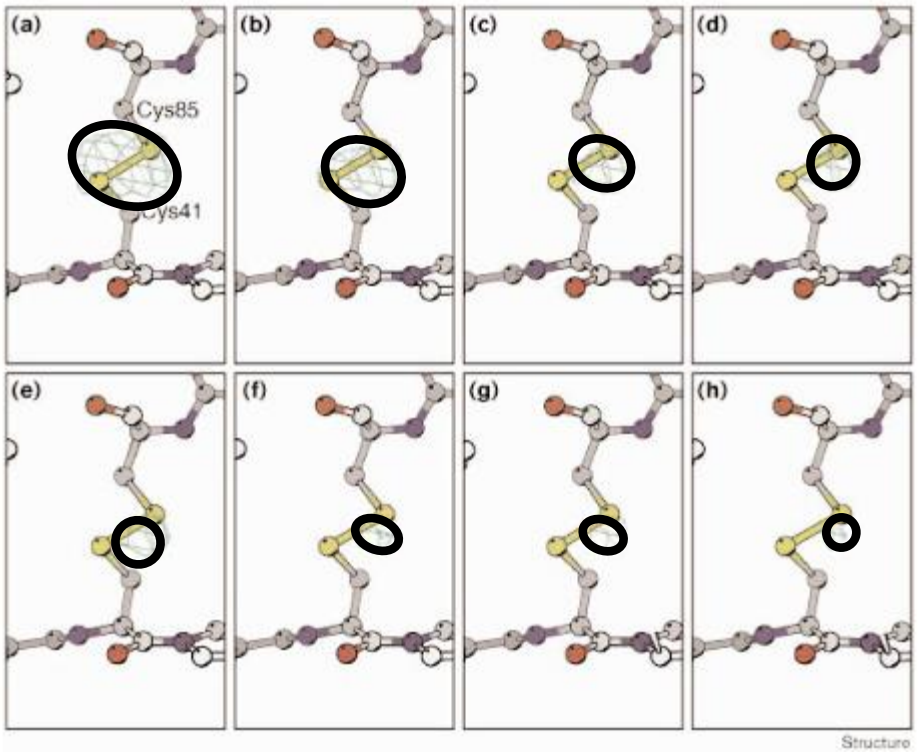
Yano, J et al *Proc. Natl. Acad. Sci. USA* **2005**, *102* 12047-12052

As the X-ray dose increases, the Mn is reduced to Mn(II) as seen by the changes in XANES spectra (left). The changes in the corresponding EXAFS spectra (right) show that the three Fourier peaks characteristic of Mn-bridging-oxo, Mn-terminal, and Mn-Mn/Ca interactions (dashed vertical line) are replaced by one Fourier peak characteristic of a Mn(II) environment.

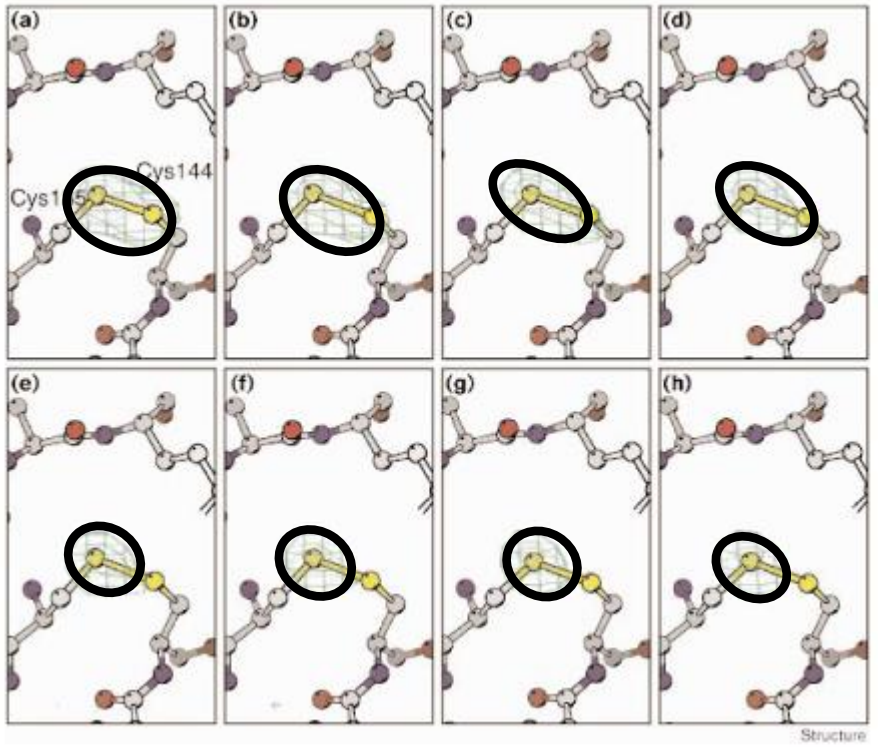
Mn(II) content in the crystals as a function of X-ray irradiation at 13.3 keV (0.933 Å) at 100 K - similar to those during X-ray diffraction data collection. At 66% of the dose (2.3×10^{10} photons/µm^2) compared to the representative average dose of (3.5×10^{10} photons/µm^2) used for crystallography, the crystals contain ~80% Mn(II). (Dashed blue line) The damage profile for solution samples is similar to that seen for crystals. (Dashed green line) The generation of Mn(II) is considerably greater when the x-ray irradiation is at 6.6 keV (1.89 Å) which is the energy at which the anomalous diffraction measurements were conducted. (Solid blue line)

Wing bean chymotrypsin inhibitor disulphides

Cys41-Cys85



Cys144-Cys135

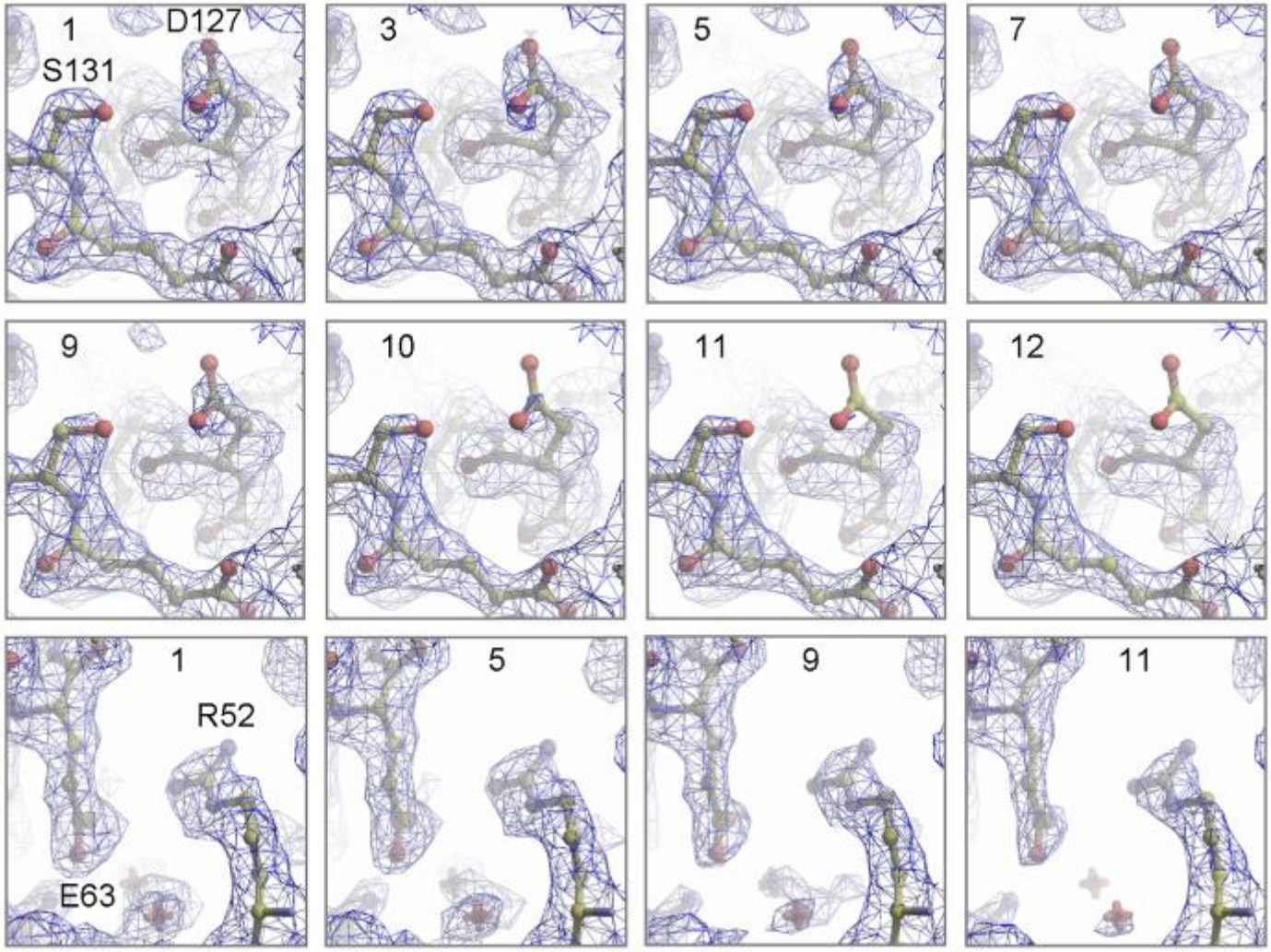


Fo-Fc maps for successive data sets.
Fc with zero occupancy sulphurs.
[Ravelli and McSweeney (2000)]

Thanks to
Elspeth
Garman

Apoferritin electron density

Contoured at 0.2 e/Å^3



Apo1
Asp127
Ser 131

Thanks to
Elspeth
Garman

Apo1
Glu63,
Arg52

Owen, Rudiño-Piñera, Garman. PNAS (2006)

i.e. damage rate is dependent on environment (but not on solvent accessibility –Fioravanti et al JSR 2007.)

Specific structural damage observed:

- Disulphide bridges broken
- Decarboxylation of glutamate and aspartate residues
- Tyrosine residues lose their hydroxyl group
- Methionines: carbon-sulphur bond cleaved

Weik *et al* (2000) PNAS 97, 623-628
Burmeister (2000), Acta Cryst D56, 328-341.
Ravelli and McSweeney, (2000) Structure 8, 315-328.

- Rupture of covalent bonds to heavier atoms:
  C-Br, C-I, S-Hg

Note that if this were due to primary damage alone, damage would be in order of absorption cross sections of atoms, which it is not.

Henderson Limit

- Radiation damage by electrons and X-rays are comparable.
- Electron diffraction patterns fade to ½ their original intensity after 1 electron $Å^{-1}$at room temperature or 5 electron $Å^{-1}$at 77K.
- The amount of energy absorbed per unit weight is expressed in units of gray (Gy). One gray dose is equivalent to one joule radiation energy absorbed per kilogram. One gray is equivalent to 100 rads.
- 5 electrons $Å^{-1}$ is approx $5x10^{7}$ Gy.
- The depth dose curve (maximum dose at ~100 μm) reduces the energy deposition so the effective energy causing the damage is conservatively $2x10^{7}$ Gy.
- X-rays of 1.5 Å give $12x10^{-16}$ Gy per photon m^{-2}.
- The X-ray flux giving rise to $2x10^{7}$ Grays is $1.6x10^{16}$ photons mm^{-2}

(Henderson (1990) Proc. R. Soc. Lond. B. 241, 6-8).

What does it mean practically: Dead Crystals

- Remember,
  - The X-ray flux giving rise to 2×10^7 Grays (dead crystals) is 1.6×10^{16} photons mm^{-2}
- Lab source – crystals at 77K (close enough to 100K)
  - 1×10^8 photons s^{-1} mm^{-2}
    - Dead crystal in ~44,000 hours (5 years – in reality a lot less)
- Synchrotron - crystals at 77K (close enough to 100K)
  - Brookhaven $\sim 0.5 \times 10^{10}$ photons s^{-1} mm^{-2}
    - Dead crystal in ~ 1.5 days
  - Stanford $\sim 1.2 \times 10^{11}$ photons s^{-1} mm^{-2}
    - Dead crystal in ~ 1.5 hours
  - APS $\sim 1.3 \times 10^{13}$ photons s^{-1} mm^{-2}
    - Dead crystal in ~ 4 seconds

A case study with Xylose isomerase

Un exemple avec de l'isomérase de xylose

Understanding crystal growth from an industrial, non-structural, perspective

- Crystallization as a purification mechanism (production)
- Crystallization as a packaging mechanism (dosing)
- Crystallization as an immobilizing mechanism (enzyme action)

Crystallization has been used as an effective means to produce large quantities of industrial enzymes. Industrial enzymes are used in:

- Food industry - production of high fructose corn syrup
- Detergents – removal of protein, starch or fatty oil stains
- Fabric conditioners – cellulases
- Paper, rubber, baking, brewing etc.

Other advantages – available in huge quantities

- Good for model studies with high purity samples available

Xylose Isomerase

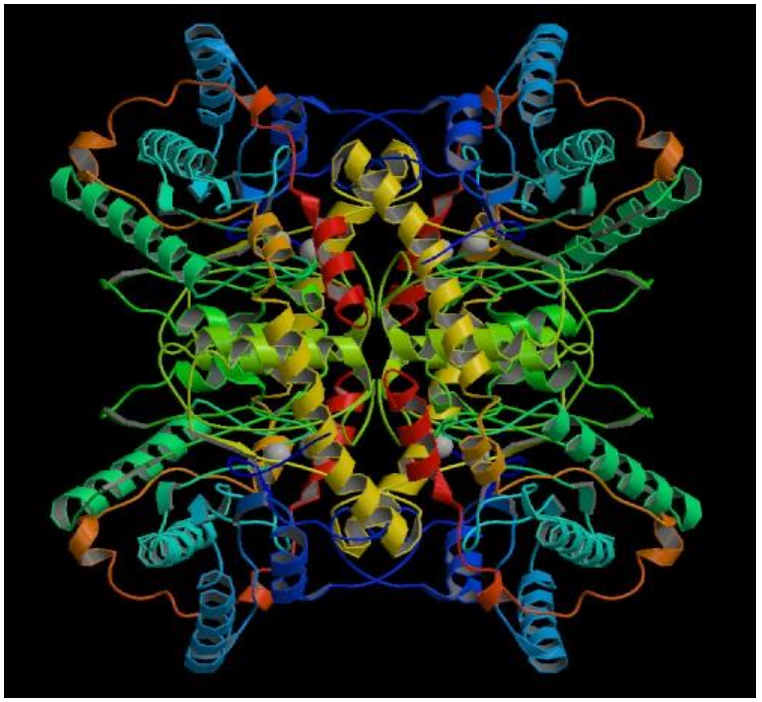
Enzymatic mechanism is a transfer of one H atom from one C atom of the substrate to an adjacent C atom.

Three mechanisms have been proposed – a base-catalyzed proton transfer, a simple hydride shift or a hydride shift mediated by a metal ion.

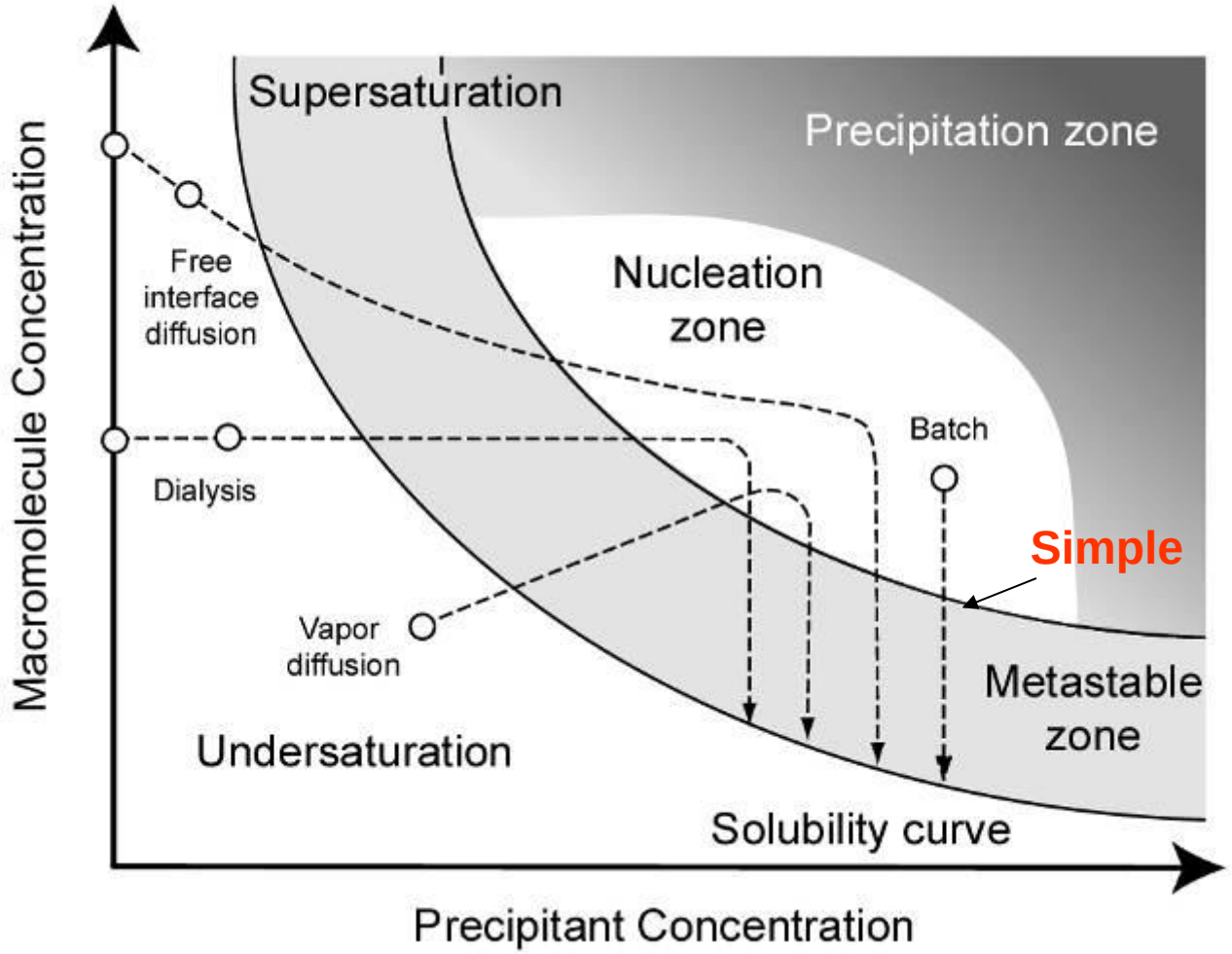
X-ray data, to date, has not revealed the exact mechanism. Neutron has.

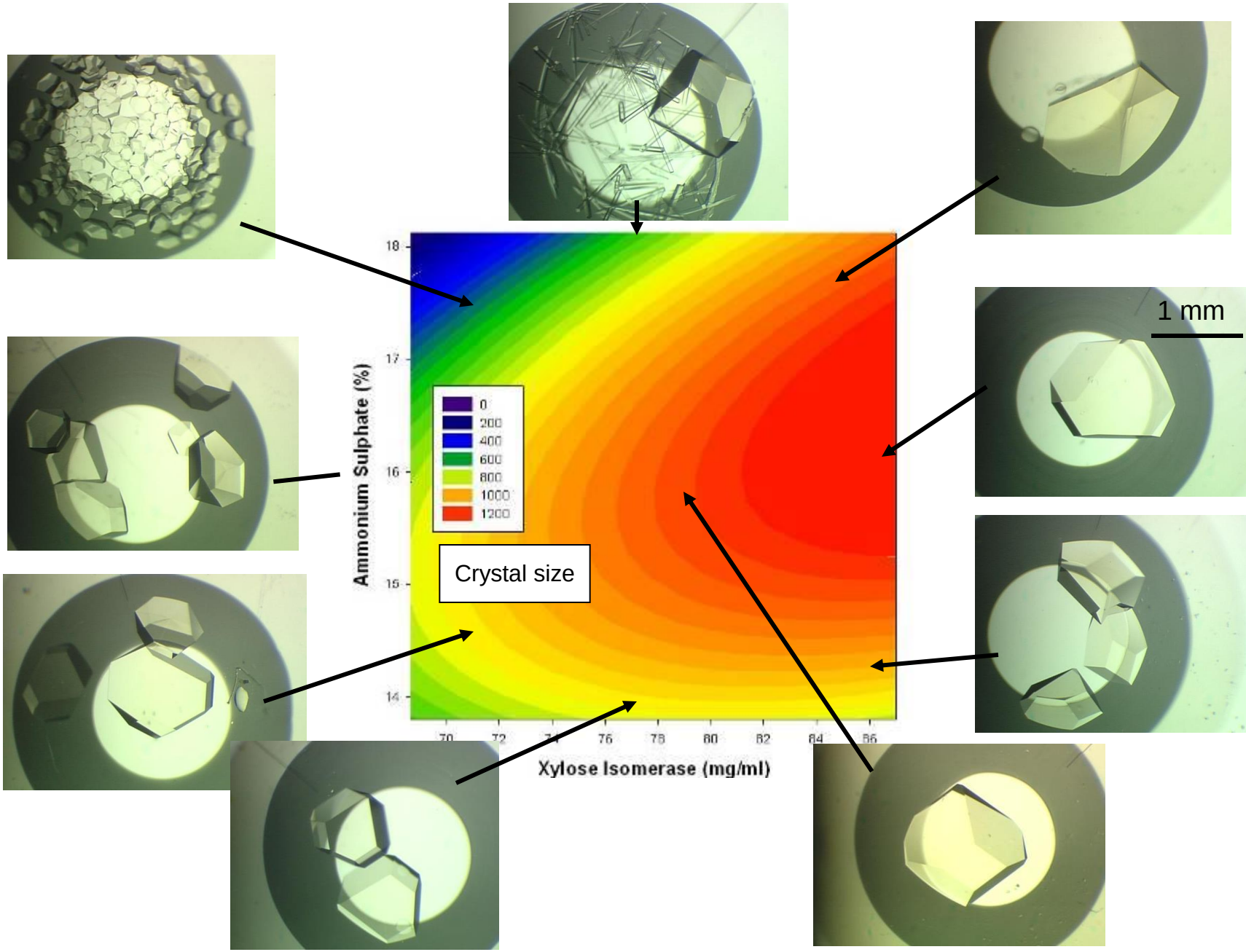
Xylose isomerase is an important industrial catalyst for the production of fructose.

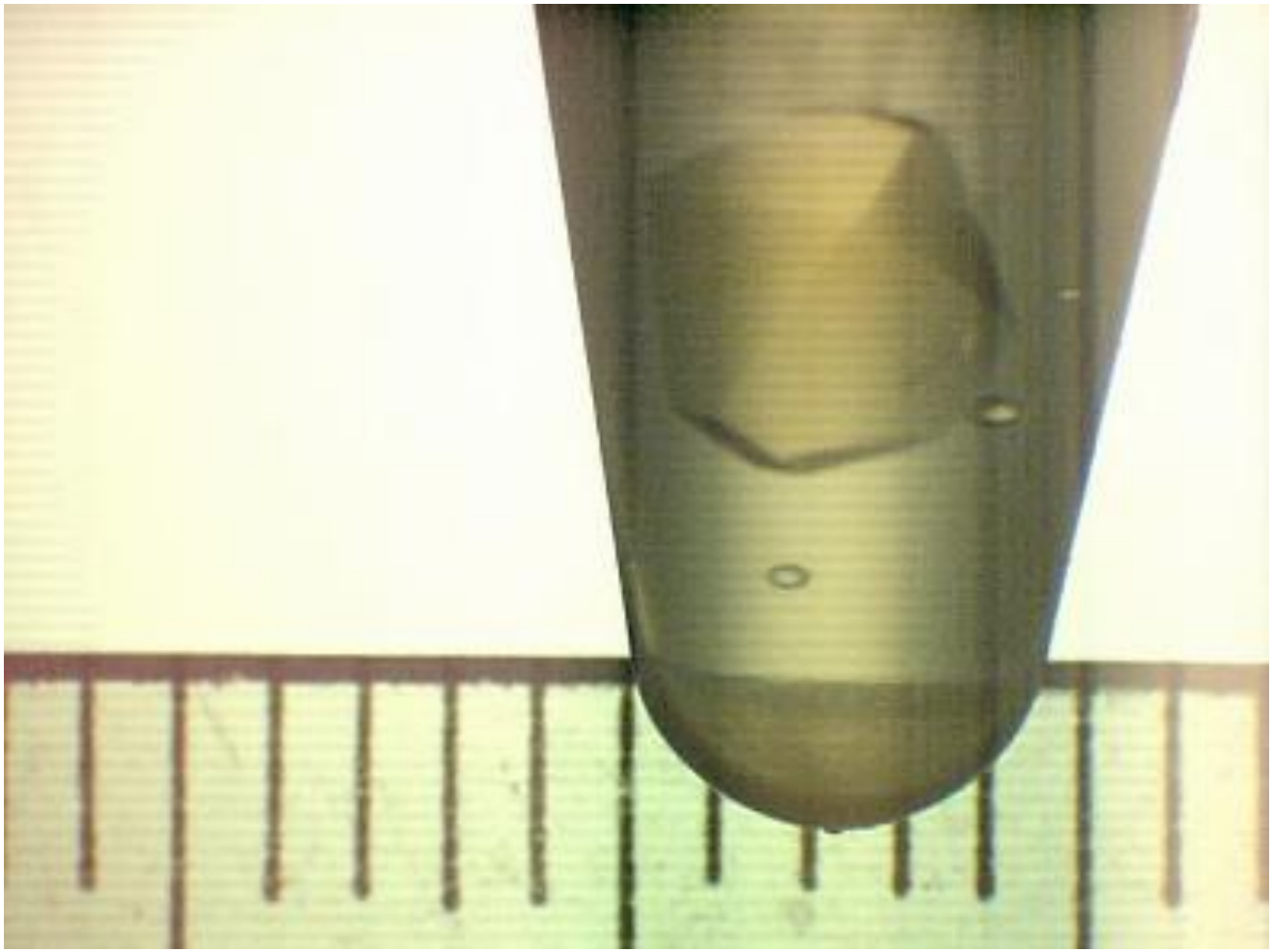
It is a homotetramer ~ 172 KDa



Method of Growth




Ammonium Sulphate (%)
Xylose Isomerase (mg/ml)
0
200
400
600
800
1000
1200
Crystal size
1 mm



Crystallization for cryocooling

- Mutant *Streptomyces rubiginosus* Xylose Isomerase
- Completely new conditions were employed each incorporating different cryoprotectants.
- Crystals grew rapidly, over a few days so the process took less than a month.
- Several conditions did well but one was outstanding:

- Well solution
  - Preciptiant: 2-propanol (4-13%)
  - Buffer: 50 mM HEPES pH 7.0
  - Salt 50 mM MgCl2
  - Mix into 0.5 ml well with 100 μl Cryoprotectant: Ethylene Glycol (20-30%) to bring volume to 0.6 ml

Protein solution

100 mg/ml in H2O with 50 mg/L MgCl2 (no typos) mixed 1:1 with well solution.

Crystals looked like diamonds and diffracted off the edge of our laboratory detector.

High resolution data collection

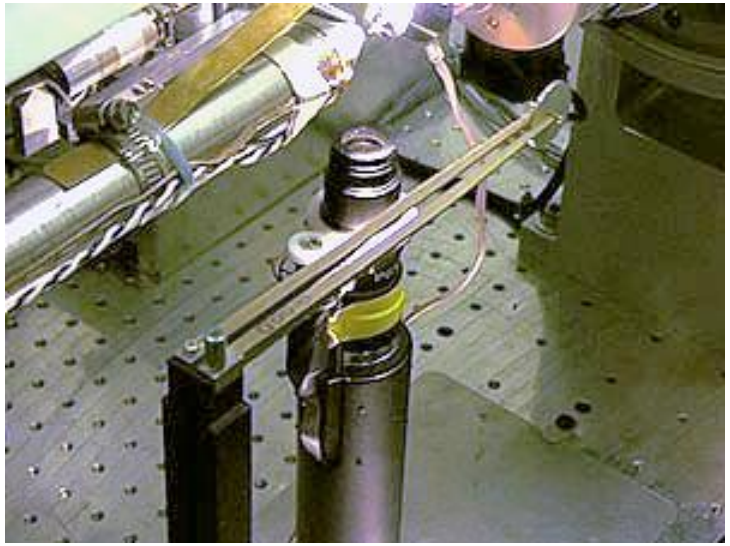
Collecte de données de haute résolution

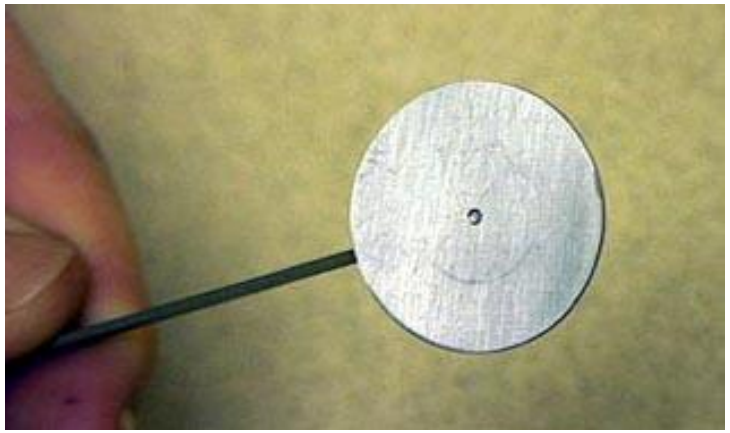
Data collection

For the high resolution data collection a somewhat larger beamstop was used.

A CCD can cope with overloaded reflections that can damage an image plate. However the overload can spill over into neighbouring pixels and eventually produce streaked lines. The large beamstop prevents this but is not typically necessary.

The cell dimensions for the crystals were, 92.7, 97.9 and 102.2 A which would have made spot separation difficult at such high resolution. Fortunately nature smiled on us and gave a space group of I222 – every second reflection was missing.

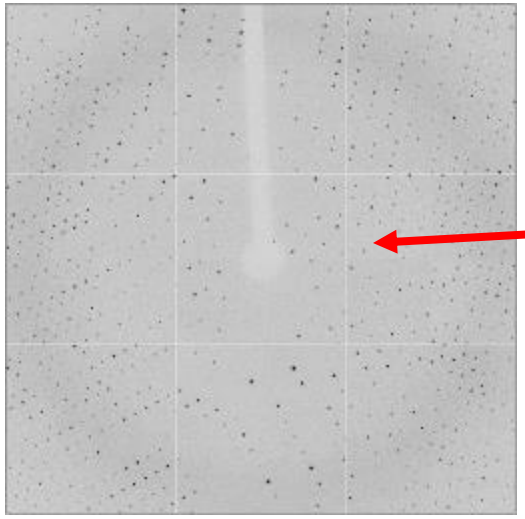




X-ray data

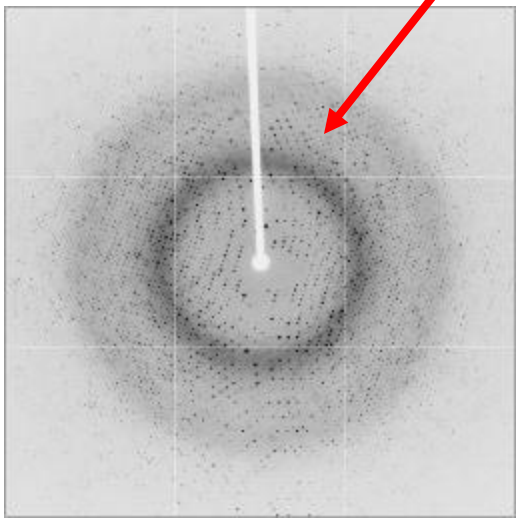
- SSRL beamline 9-1
- Beautiful diffraction to about 0.9 A
- Collected a high resolution pass followed by a low resolution pass to cover the complete dynamic range.
  - The reasoning was that the radiation damage caused the high resolution data to go first and we could collect the ‘undamaged’ low resolution data last.

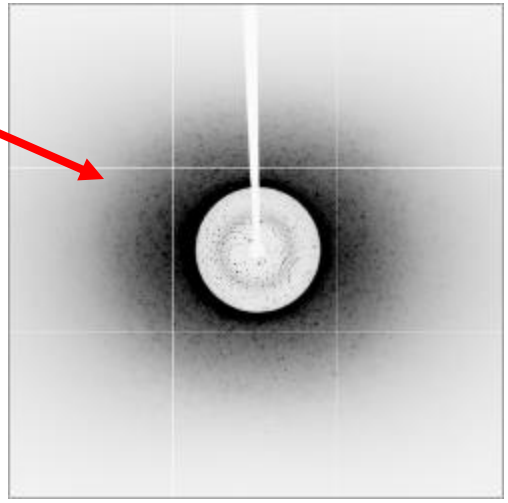
  - Problem – The low resolution data would not scale to the high resolution data (which presented it’s own processing problems). In the time available each had been processed as collected but not scaled together. The data had a low resolution hole and had to be discarded.
  - Lesson learned – Radiation damage is a global process. It is first observed in the high resolution data but is affecting all the data.
  - New protocol – Low resolution data collection is very fast, collect it first with minimal exposure then extend the resolution with high resolution data collection.

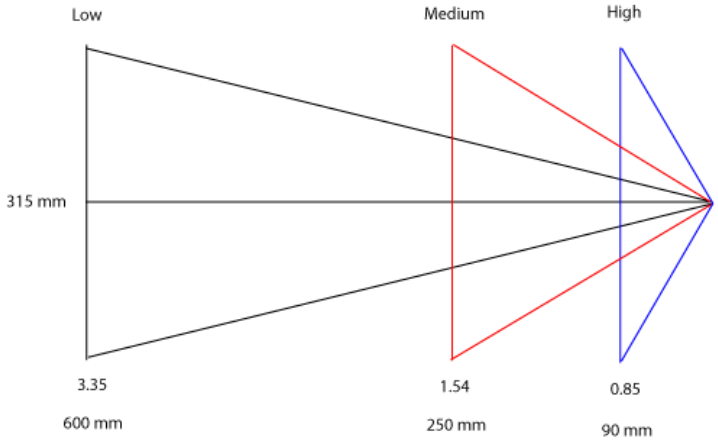


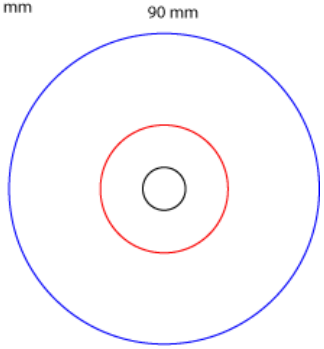
SSRL Beamline 11-1, ADSC Quantum-315 CCD detector
Low, medium and high-resolution data collection, 0.8550A

- Low: Crystal-to-detector 600 mm, 2 degree rotation, 1s exposure, 80 images, 160 degrees of data.
- Medium: Crystal to detector 250 mm, 2 degree rotation, 2 s exposure, 75 images, 150 degrees of data.
- High: Crystal to detector 90 mm, 0.5 degree rotation, 4 s exposure, 720 images, 360 degrees of data









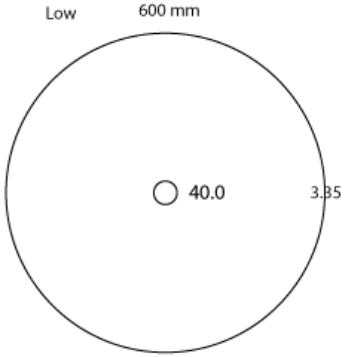
Real Space

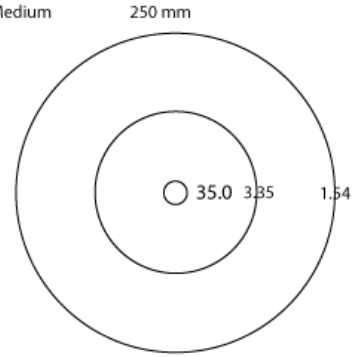
Medium to high overlap area = 13%

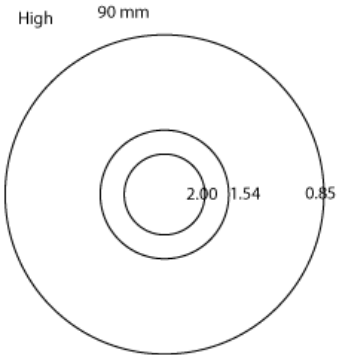
Low to medium overlap area = 17%

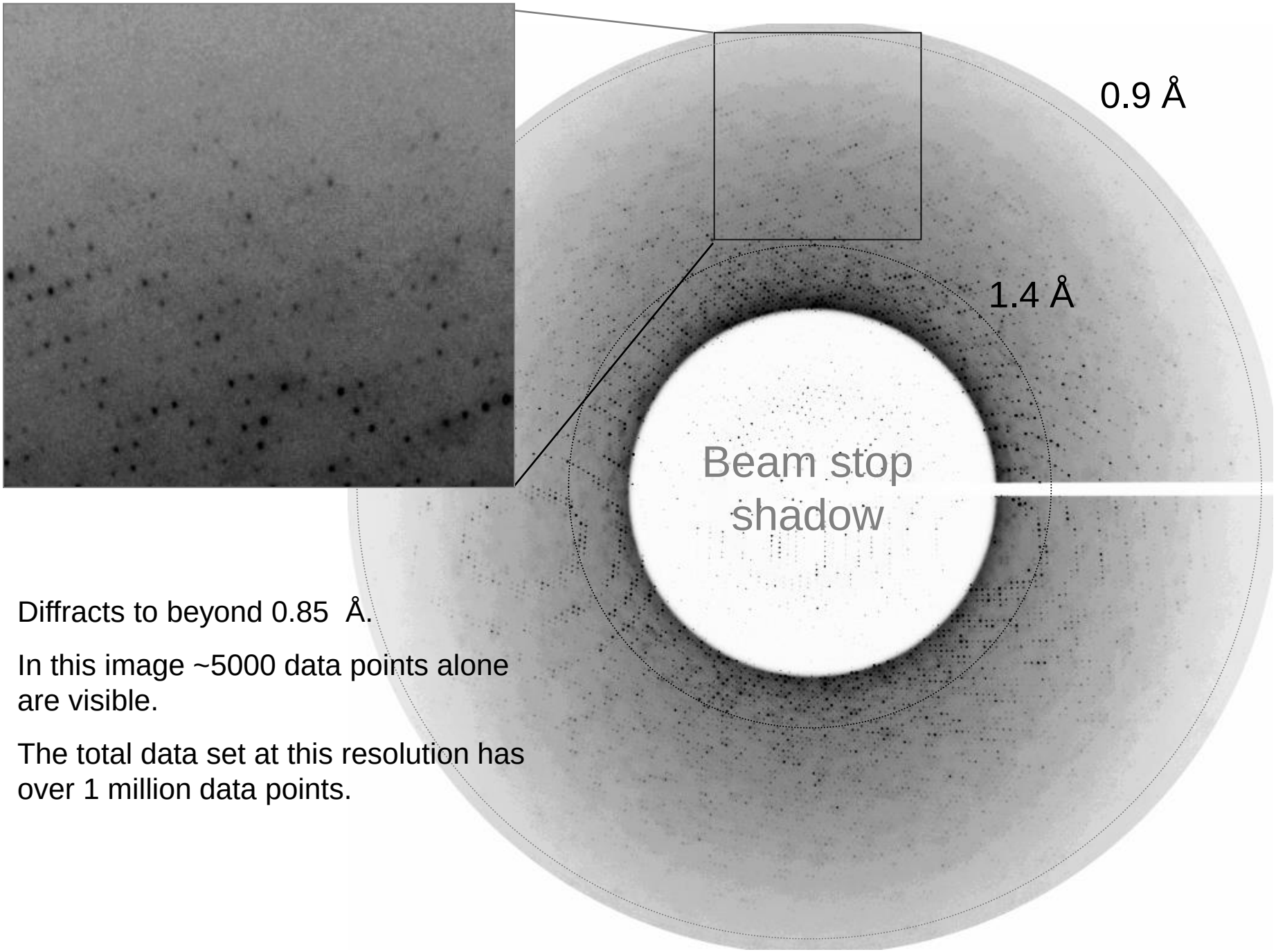
Physical space sampled

Resolution space sampled









Diffracts to beyond 0.85 Å.

In this image ~5000 data points alone are visible.

The total data set at this resolution has over 1 million data points.

Lower limit	Upper limit	% of of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
40.00	2.97	0.1	0.2	0.2	0.3	0.5	1.1	5.3	94.7	100.0
2.97	2.36	0.2	0.4	0.6	1.1	1.7	3.3	18.6	81.4	100.0
2.36	2.06	0.3	0.7	1.2	1.6	2.5	5.2	26.1	73.8	99.9
2.06	1.87	1.0	1.5	1.8	2.1	2.9	5.0	15.3	84.7	100.0
1.87	1.74	1.9	2.2	2.4	2.6	3.2	4.5	8.0	92.0	100.0
1.74	1.64	2.4	2.7	3.0	3.1	3.7	5.3	10.1	89.9	100.0
1.64	1.56	2.1	2.5	3.0	3.4	4.2	6.6	13.1	86.9	100.0
1.56	1.49	2.7	3.4	3.9	4.3	5.7	8.7	16.1	83.9	100.0
1.49	1.43	2.7	3.5	4.3	5.0	6.2	10.0	19.9	80.1	100.0
1.43	1.38	3.0	3.9	4.7	5.5	7.3	11.7	23.0	77.0	100.0
1.38	1.34	3.2	4.1	5.2	6.2	8.0	12.9	25.3	74.7	100.0
1.34	1.30	3.4	4.5	5.6	6.7	9.0	14.6	28.3	71.7	100.0
1.30	1.27	3.4	4.9	6.3	8.0	10.6	17.1	32.1	67.9	100.0
1.27	1.23	3.7	5.1	6.5	7.9	10.5	17.7	34.2	65.8	100.0
1.23	1.21	3.9	5.6	7.4	8.9	12.0	19.8	37.9	62.1	100.0
1.21	1.18	3.8	5.5	7.2	9.0	12.2	20.7	40.2	59.8	100.0
1.18	1.16	4.2	5.9	7.7	9.5	12.9	22.1	41.2	58.8	100.0
1.16	1.14	4.1	6.3	8.2	9.9	13.8	23.5	44.4	55.6	100.0
1.14	1.12	3.8	6.3	8.7	11.0	15.6	25.9	47.9	52.1	100.0
1.12	1.10	3.7	6.1	8.7	11.3	16.2	28.3	52.5	47.5	100.0
1.10	1.08	4.3	7.3	10.7	13.6	19.1	32.6	57.8	42.2	100.0
1.08	1.06	4.1	7.2	10.8	14.1	20.7	35.5	62.0	38.0	100.0
1.06	1.05	4.4	8.6	12.2	16.3	23.4	39.5	65.1	34.9	100.0
1.05	1.03	4.5	9.0	13.7	18.3	26.7	43.7	70.9	29.1	100.0
1.03	1.02	4.6	9.4	14.7	19.8	29.5	49.4	76.6	23.4	100.0
1.02	1.00	5.1	10.6	16.8	22.5	32.7	53.4	79.9	20.1	100.0
1.00	0.99	5.5	11.7	18.5	24.7	35.9	57.2	83.7	16.3	100.0
0.99	0.98	5.9	12.9	20.9	28.3	40.6	62.8	87.0	13.0	100.0
0.98	0.97	6.2	13.9	22.1	30.7	44.3	66.9	89.1	10.9	100.0
0.97	0.96	6.5	15.3	24.3	32.8	46.8	70.4	91.6	8.4	100.0
0.96	0.95	6.5	16.1	26.1	35.6	50.8	74.0	93.5	6.4	99.9
0.95	0.94	7.3	17.2	28.5	38.3	54.2	77.8	95.3	4.7	100.0
0.94	0.93	8.8	20.0	32.6	43.6	59.7	81.5	96.4	3.6	99.9
0.93	0.92	9.0	21.0	33.5	45.0	62.5	84.7	97.7	2.2	99.9
0.92	0.91	9.8	23.5	37.6	49.8	66.1	87.2	97.9	1.9	99.9
0.91	0.90	10.4	25.7	41.7	53.5	70.2	89.7	98.7	1.1	99.9
0.90	0.89	11.7	27.8	44.5	57.3	74.6	92.1	99.3	0.6	99.9
0.89	0.89	13.0	30.1	47.3	60.7	77.2	93.8	99.4	0.4	99.8
0.89	0.88	13.5	31.5	49.6	63.3	79.5	94.4	99.3	0.4	99.7
0.88	0.87	12.2	28.3	44.2	55.5	68.5	79.5	82.7	0.2	82.9
All hkl		5.1	10.5	16.1	20.9	28.4	40.6	56.4	43.2	99.5

Shell limit	Lower Angstrom	Upper Angstrom	Average I	Average error	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac
40.00	2.97		29658.7	971.4	132.0	0.991	0.073	0.094
2.97	2.36		12008.8	485.4	94.7	0.967	0.071	0.090
2.36	2.06		8302.1	360.0	84.7	0.878	0.067	0.084
2.06	1.87		4912.8	165.1	42.0	0.947	0.068	0.085
1.87	1.74		2938.6	65.3	16.0	1.000	0.067	0.081
1.74	1.64		2071.2	47.7	13.1	0.948	0.069	0.081
1.64	1.56		1600.1	39.3	12.5	0.944	0.074	0.085
1.56	1.49		1232.0	31.8	11.5	0.932	0.078	0.090
1.49	1.43		956.9	26.4	10.6	0.874	0.079	0.088
1.43	1.38		778.3	22.8	10.1	0.871	0.084	0.091
1.38	1.34		668.3	19.0	9.9	0.997	0.090	0.097
1.34	1.30		602.1	17.9	9.8	1.042	0.096	0.103
1.30	1.27		515.0	16.4	9.5	1.065	0.096	0.103
1.27	1.23		487.6	16.0	9.6	1.028	0.098	0.104
1.23	1.21		442.7	15.1	9.4	0.982	0.101	0.105
1.21	1.18		416.9	14.6	9.4	0.935	0.103	0.105
1.18	1.16		399.4	14.4	9.4	0.904	0.103	0.106
1.16	1.14		362.8	13.7	9.3	0.888	0.109	0.111
1.14	1.12		318.1	12.9	9.2	0.901	0.121	0.124
1.12	1.10		283.6	12.3	9.1	0.917	0.134	0.138
1.10	1.08		244.5	11.6	9.0	0.927	0.146	0.147
1.08	1.06		212.0	11.0	8.9	0.921	0.155	0.158
1.06	1.05		183.2	10.2	8.8	0.968	0.169	0.171
1.05	1.03		158.7	9.9	8.7	0.969	0.150	0.142
1.03	1.02		131.5	9.4	8.6	0.998	0.158	0.138
1.02	1.00		115.2	9.2	8.5	0.997	0.175	0.152
1.00	0.99		100.2	9.0	8.4	1.019	0.200	0.177
0.99	0.98		85.7	8.8	8.4	1.020	0.228	0.199
0.98	0.97		76.0	8.6	8.4	1.047	0.256	0.220
0.97	0.96		67.7	8.6	8.4	1.041	0.284	0.251
0.96	0.95		60.1	8.6	8.4	1.055	0.319	0.281
0.95	0.94		54.2	8.6	8.4	1.051	0.353	0.309
0.94	0.93		46.9	8.6	8.5	1.063	0.407	0.350
0.93	0.92		42.1	8.6	8.5	1.052	0.451	0.397
0.92	0.91		37.5	8.6	8.5	1.050	0.503	0.438
0.91	0.90		33.0	8.6	8.5	1.056	0.568	0.492
0.90	0.89		29.0	8.6	8.5	1.054	0.647	0.561
0.89	0.89		26.0	8.6	8.5	1.041	0.708	0.610
0.89	0.88		23.9	8.5	8.5	1.045	0.762	0.651
0.88	0.87		21.9	8.9	8.9	1.043	0.820	0.699
All reflections			1823.5	65.6	17.6	0.978	0.078	0.093

Journal of
Applied
Crystallography

ISSN 0021-8898

Received 6 September 2000
Accepted 22 November 2000

Global indicators of X-ray data quality

Manfred S. Weiss

Institute of Molecular Biotechnology, Department of Structural Biology and Crystallography, PO Box 100813, D-07708 Jena, Germany. Correspondence e-mail: msweiss@imb-jena.de

Global indicators of the quality of diffraction data are presented and discussed, and are evaluated in terms of their performance with respect to various tasks. Based on the results obtained, it is suggested that some of the conventional indicators still in use in the crystallographic community should be abandoned, such as the nominal resolution $d_{\min}$ or the merging R factor $R_{\rm merge}$, and replaced by more objective and more meaningful numbers, such as the effective optical resolution $d_{\rm eff,opt}$ and the redundancy-independent merging R factor $R_{\rm r.i.m.}$. Furthermore, it is recommended that the precision-indicating merging R factor $R_{\rm p.i.m.}$ should be reported with every diffraction data set published, because it describes the precision of the averaged measurements, which are the quantities normally used in crystallography as observables.

Highly redundant data can have artificially high R_{merge} values

Rmerge software, Bob's scaling software and now Scala incorporate this.

Table 2
Global quality indicators that can be derived from a diffraction data set in which equivalent reflections have not been merged.

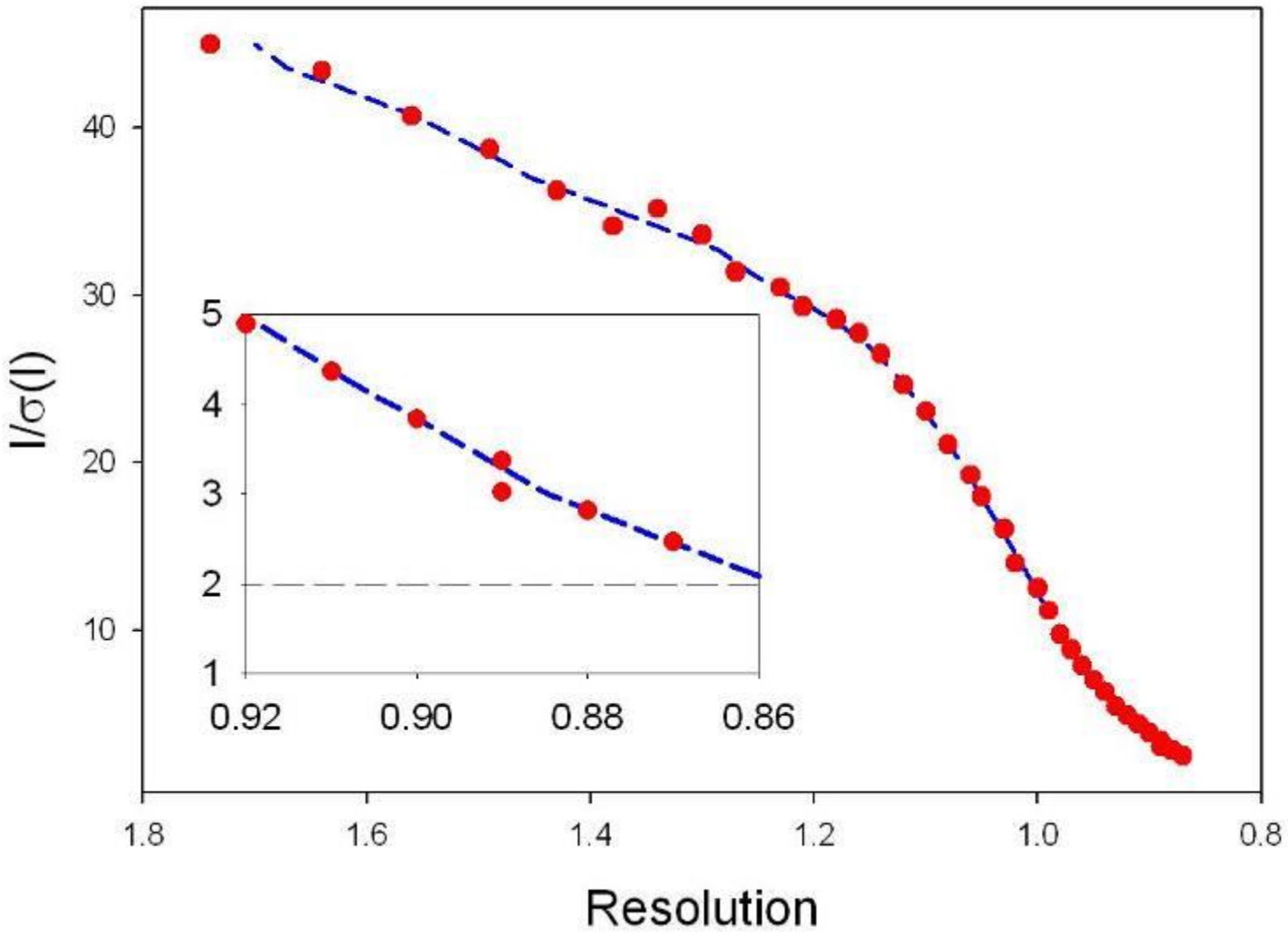
(1)	Redundancy	N
(2)	Merging R factor	$R_{\rm merge} = \sum_{hkl}\sum_{i} \lvert I_i(hkl) - \overline{I(hkl)}\rvert \Big/ \sum_{hkl}\sum_{i} I_i(hkl)$
(3)	Redundancy-independent merging R factor	$R_{\rm r.i.m.} = \sum_{hkl}[N/(N-1)]^{1/2} \sum_{i} \lvert I_i(hkl) - \overline{I(hkl)}\rvert \Big/ \sum_{hkl}\sum_{i} I_i(hkl)$
(4)	Pooled coefficient of variation	$\mathrm{PCV} = \sum_{hkl}\{[1/(N-1)]\sum_{i} \lvert I_i(hkl) - \overline{I(hkl)}\rvert^2\}^{1/2} \Big/ \sum_{hkl}\sum_{i} I_i(hkl)$
(5)	Precision-indicating merging R factor	$R_{\rm p.i.m.} = \sum_{hkl}[1/(N-1)]^{1/2} \sum_{i} \lvert I_i(hkl) - \overline{I(hkl)}\rvert \Big/ \sum_{hkl}\sum_{i} I_i(hkl)$
(6)	R factor between data subsets I_1 and I_2	$R_{\rm mrgd\text{-}\mathit{I}} = \sum_{hkl} \lvert \overline{I_1(hkl)} - \overline{I_2(hkl)}\rvert \Big/ 0.5\sum_{hkl}[\overline{I_1(hkl)} + \overline{I_2(hkl)}]$

Summary of various R-factors by shells

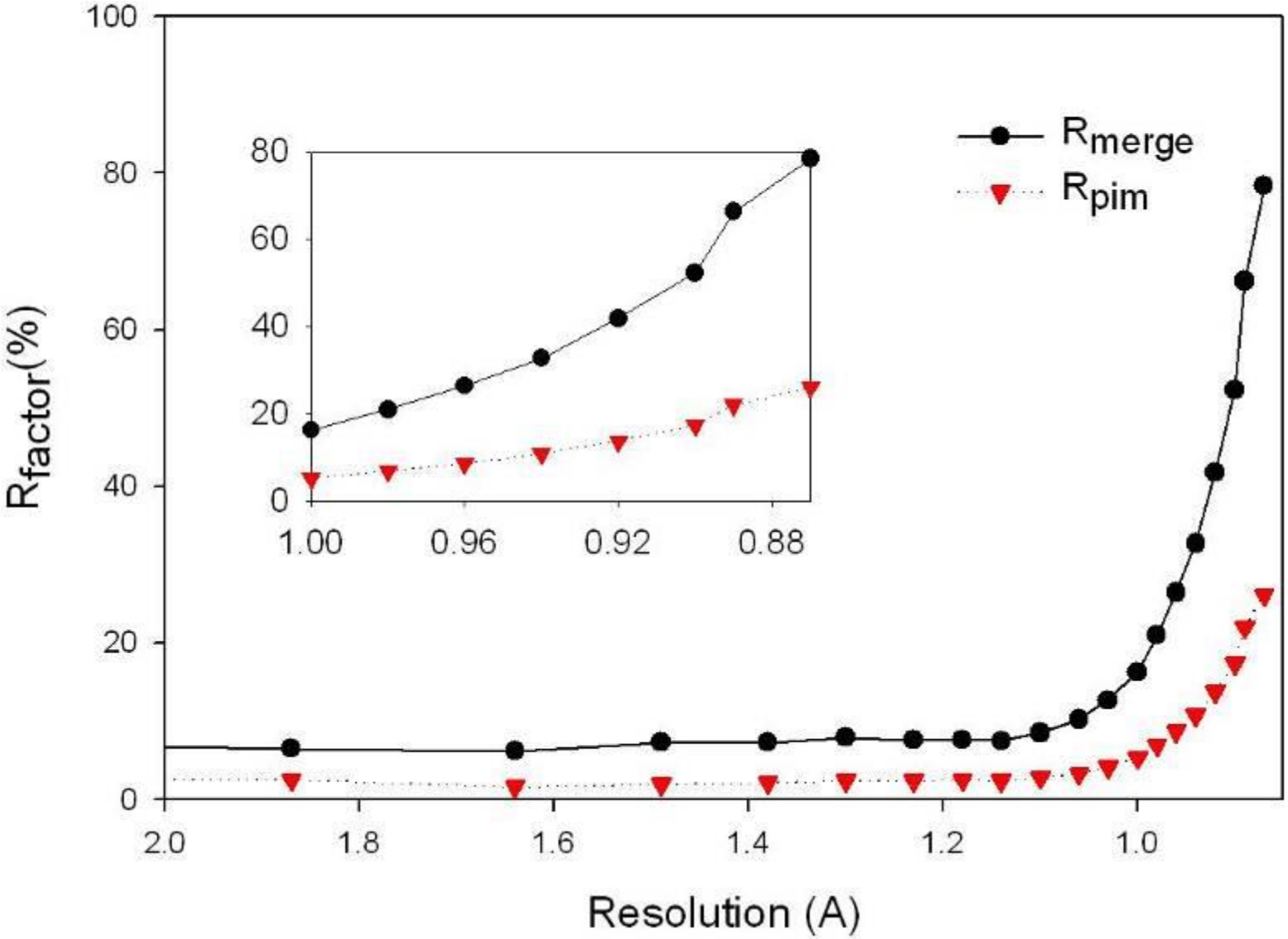
****

(in all sums except Imean and Smean single measurements are excluded.)

Resolution bins		# of	Average	Average	Average	Rm	Rsqm	Rrim	Rpim
Lower	Upper	refs	Red.	I	err				
30.00	2.36	19431	9.12	21022.9	2135.2	0.073	0.089	0.077	0.024
2.36	1.87	19039	8.73	6669.7	698.1	0.065	0.080	0.070	0.025
1.87	1.64	18940	16.38	2529.5	250.2	0.062	0.081	0.064	0.016
1.64	1.49	18867	14.87	1429.4	155.2	0.073	0.095	0.075	0.019
1.49	1.38	18858	12.88	883.3	102.7	0.073	0.094	0.076	0.021
1.38	1.30	18798	11.61	647.7	73.7	0.079	0.099	0.083	0.024
1.30	1.23	18796	10.18	514.0	54.9	0.076	0.091	0.079	0.024
1.23	1.18	18782	10.07	445.3	50.1	0.076	0.092	0.080	0.025
1.18	1.14	18775	10.07	401.0	47.6	0.075	0.092	0.079	0.024
1.14	1.10	18712	10.03	320.8	42.5	0.085	0.103	0.090	0.028
1.10	1.06	18753	9.99	246.2	37.8	0.102	0.123	0.108	0.033
1.06	1.03	18725	9.95	184.7	33.1	0.127	0.151	0.134	0.042
1.03	1.00	18765	9.92	134.6	30.6	0.163	0.192	0.172	0.053
1.00	0.98	18697	9.88	101.5	29.3	0.210	0.241	0.221	0.069
0.98	0.96	18693	9.88	79.5	28.5	0.264	0.295	0.279	0.087
0.96	0.94	18710	9.81	63.8	28.3	0.327	0.352	0.345	0.108
0.94	0.92	18706	9.77	50.1	28.4	0.418	0.424	0.441	0.138
0.92	0.90	18610	9.75	39.8	28.5	0.523	0.494	0.552	0.174
0.90	0.89	18744	9.71	31.4	28.6	0.662	0.570	0.699	0.220
0.89	0.87	18624	9.62	26.0	28.4	0.784	0.624	0.828	0.261
All reflections		376025	10.61	1830.8	199.7	0.075	0.093	0.079	0.024



3,991,720 reflections, average redundancy 10.6, 376,419 unique


100
80
60
40
20
0
2.0
1.8
1.6
1.4
1.2
1.0
Resolution (A)
Rfactor(%)
Rmerge
Rpim
80
60
40
20
0
1.00
0.96
0.92
0.88

Molecular replacement and removal of model bias

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

ARP/wARP and molecular replacement

Anastassis Perrakis,[a]* Maria Harkiolaki,[b] Keith S. Wilson[b] and Victor S. Lamzin[c]

[a]Department of Molecular Carcinogenesis, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands, [b]York Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5DD, England, and [c]EMBL, c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany

Correspondence e-mail: perrakis@nki.nl

The aim of *ARP/wARP* is improved automation of model building and refinement in macromolecular crystallography. Once a molecular-replacement solution has been obtained, it is often tedious to refine and rebuild the initial (search) model. *ARP/wARP* offers three options to automate that task to varying extents: (i) autobuilding of a completely new model based on phases calculated from the molecular-replacement solution, (ii) updating of the initial model by atom addition and deletion to obtain an improved map and (iii) docking of a structure onto a new (or mutated) sequence, followed by rebuilding and refining the side chains in real space. A few examples are presented where *ARP/wARP* made a considerable difference in the speed of structure solution and/or made possible refinement of otherwise difficult or uninterpretable maps. The resolution range allowing complete autobuilding of protein structures is currently 2.0 Å, but for map improvement considerable advances over more conventional refinement

Acta Cryst. (2001). D**57**, 1445–1450

2.1. Automatically building a new model – *warpNtrace*

This mode requires the native diffraction data to extend to a spacing better than 2.0 Å. If the resolution is between 2.0 and 2.3 Å and the starting model is good and/or the solvent content is high, then this approach is worth a try. If the model is particularly bad (*i.e.* it was very difficult or even unexpected to find a molecular-replacement solution) or incomplete (less than 2/3 of the final model), data to a resolution higher than 2.0 Å might be necessary.

2.1.1. Direct use of the molecular-replacement model in *warpNtrace* cycles. The available model is here fed directly into the *warpNtrace* procedure. A new map is calculated after a single refinement cycle that is performed mainly to obtain reliable σ_A weights (Read, 1986) and then a new model is built automatically. An important element is that the stereochemistry of the molecular-replacement model is completely ignored. The atoms of the model are used solely as guides for the autotracing, but the autotracing algorithm does *not* compare ambiguous areas against the existing model. This might appear to be a limitation of the algorithm, since prior knowledge is ignored, but in reality presents a vital means of minimizing model bias. This simplistic procedure (Fig. 2*b*) can yield impressive results, as depicted in §3.1.

Refinement protocol

- Refmac starting at low resolution isotropic refinement.
- Refmac gradually increasing to high resolution.
- Refmac at high resolution go anisotropic.
- Make best guesses at occupancies.
- Final Refmac, multiple occupancies with guesses on actual occupancy, anisotropic refinement, all data.
- Use shelxpro to convert pdb to shelxl.ins file
- Refine isotropically at medium resolution with multiple occupancies assigned to free variables.
- Gradually increase resolution and refine occupancies.
- Go anisotropic, wait for several crashes and sort out the problems.
- Add hydrogens, wait for several more crashes
- Celebrate, publish – soon. (célébrez, éditez – bientôt)

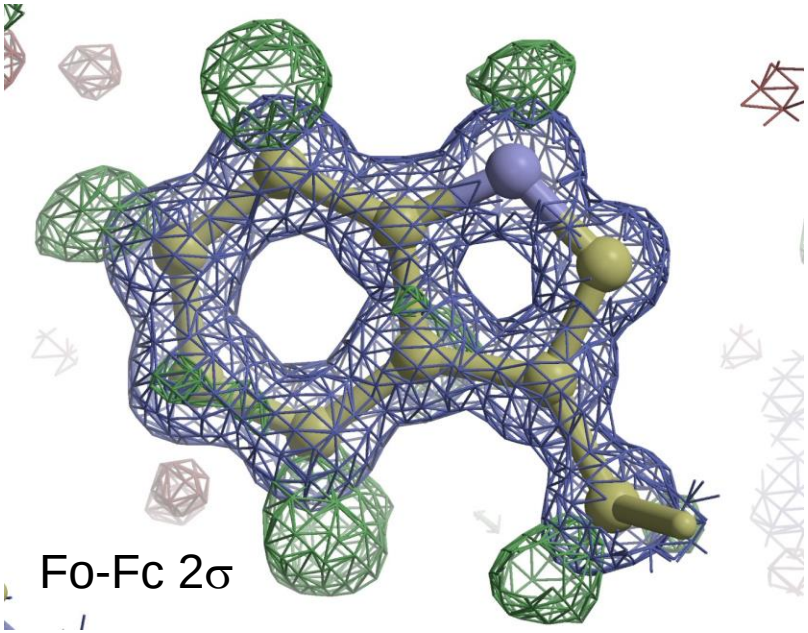
Shelx – Current status

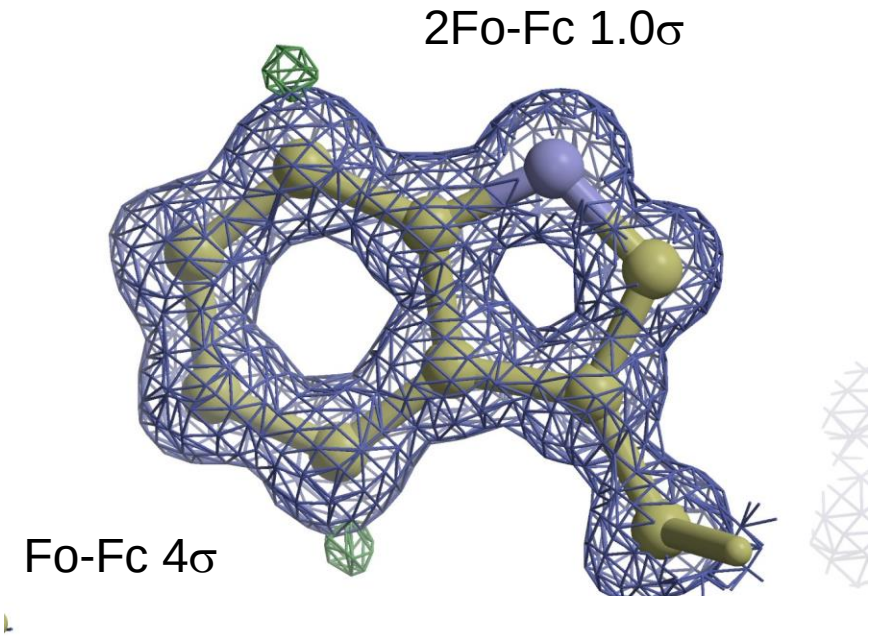
- R =9.9% for all data
- R_{free} =11.1% for all data

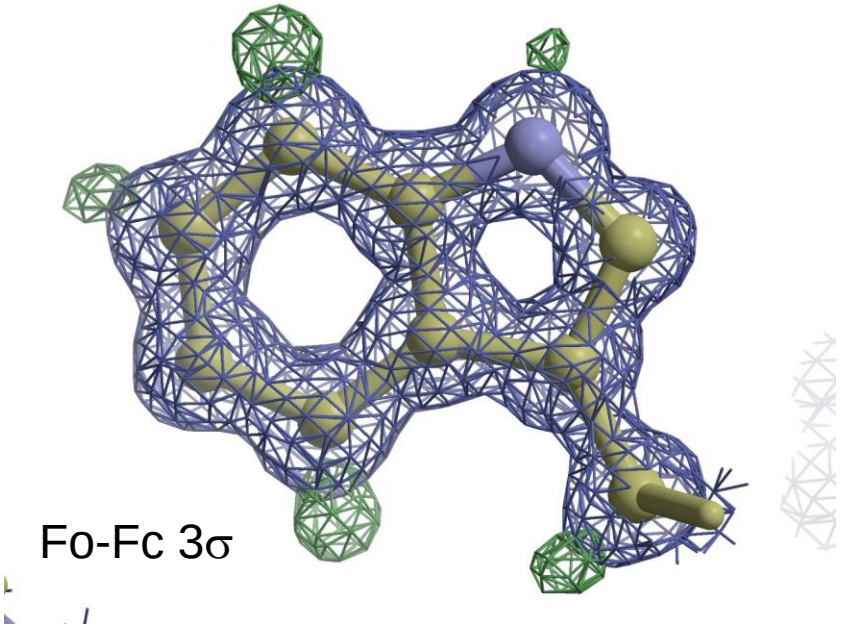
- Hydrogen's on potentially protonated sites are not be generated.

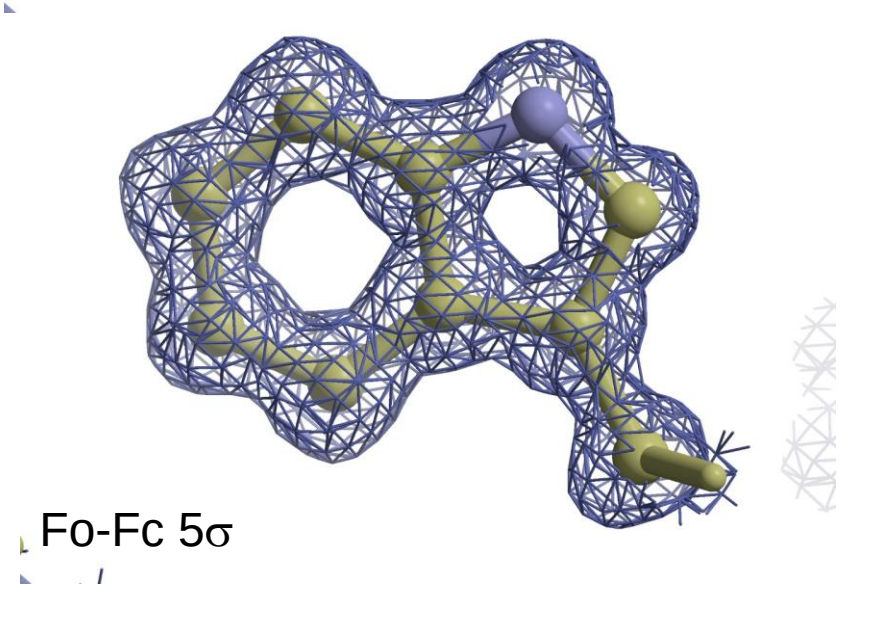
- Have to use 2Fo-Fc = 1σ otherwise atoms appear as unconnected spheres

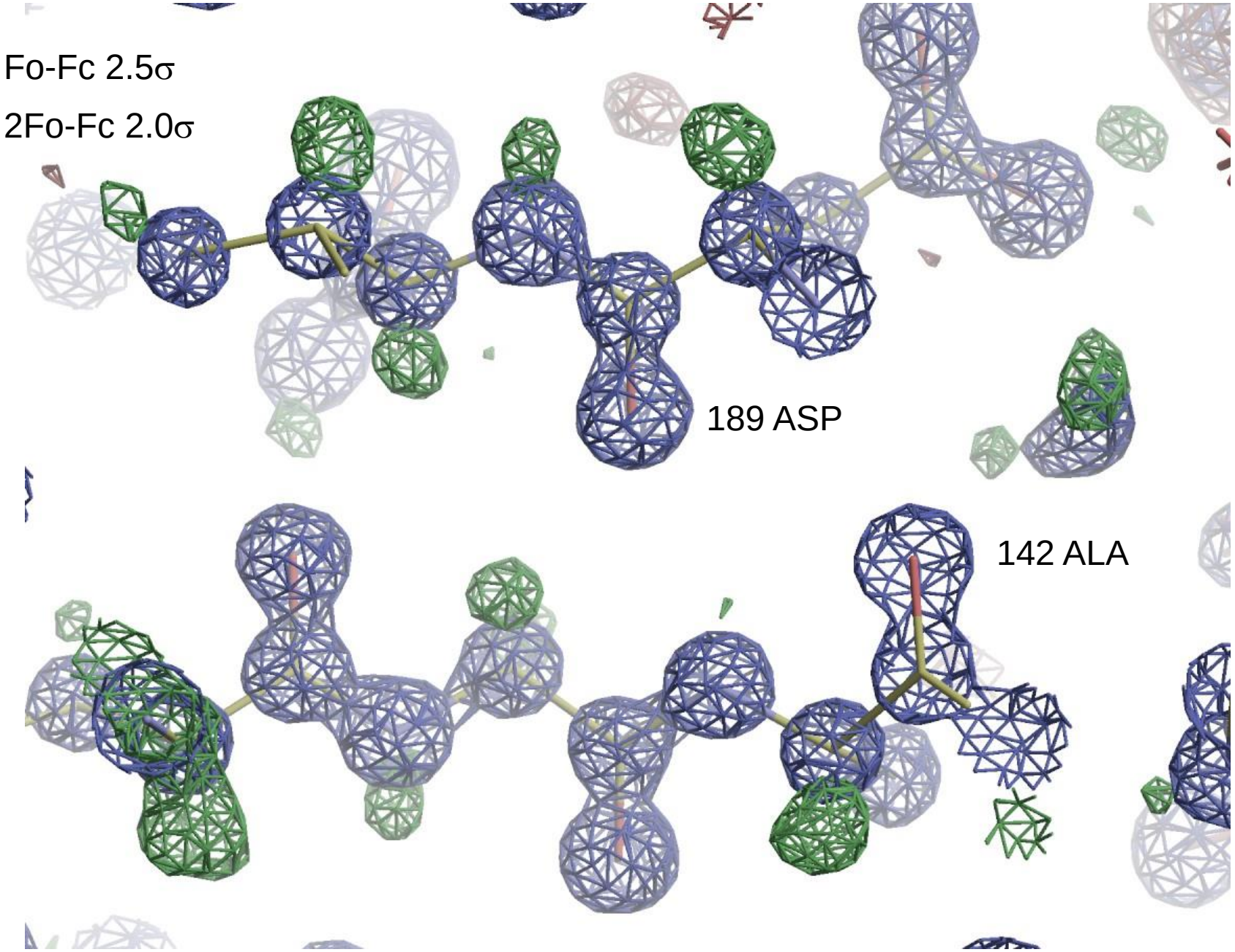
- The active sight – multiple metal positions.


Fo-Fc 2σ


2Fo-Fc 1.0σ
Fo-Fc 4σ


Fo-Fc 3σ

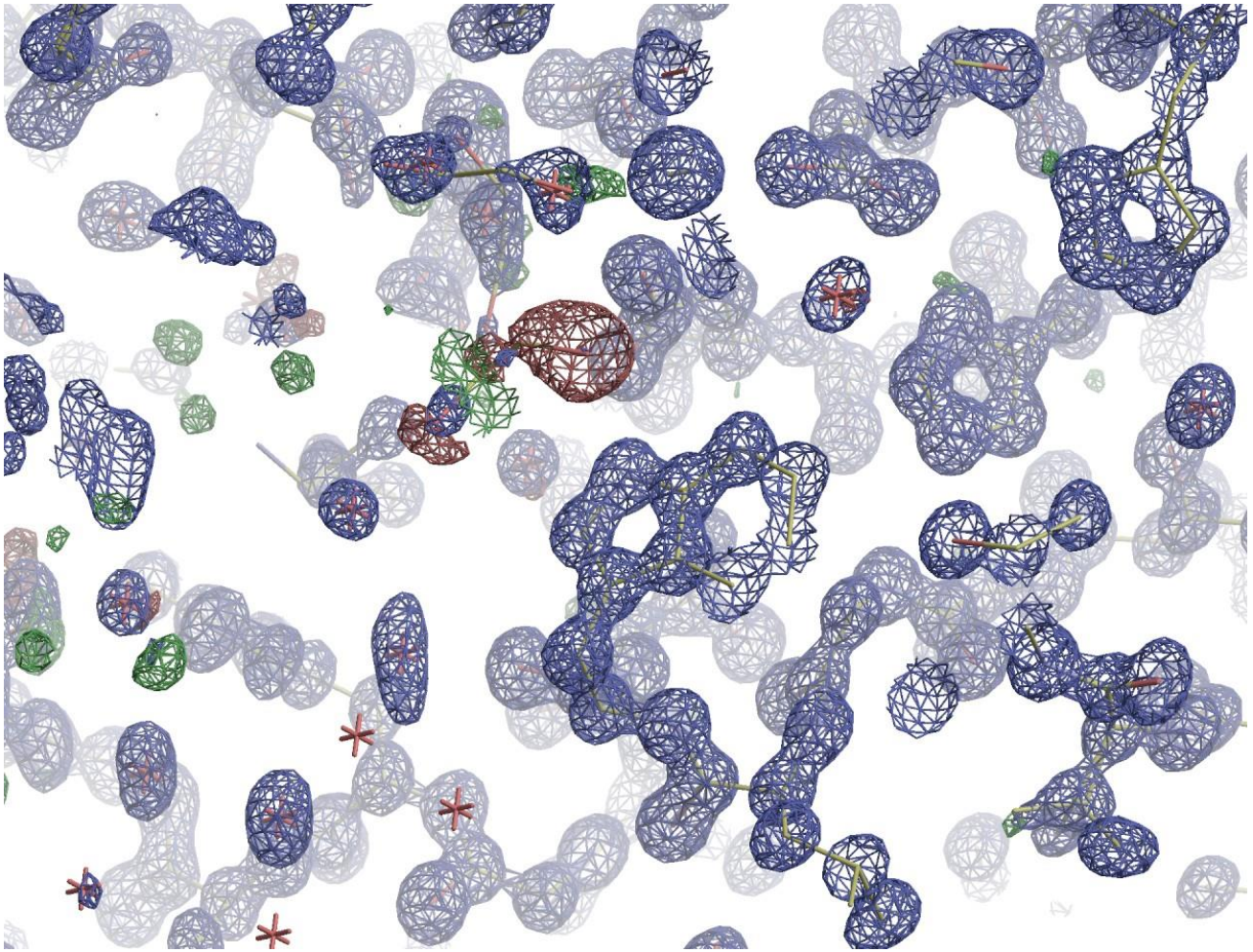

Fo-Fc 5σ

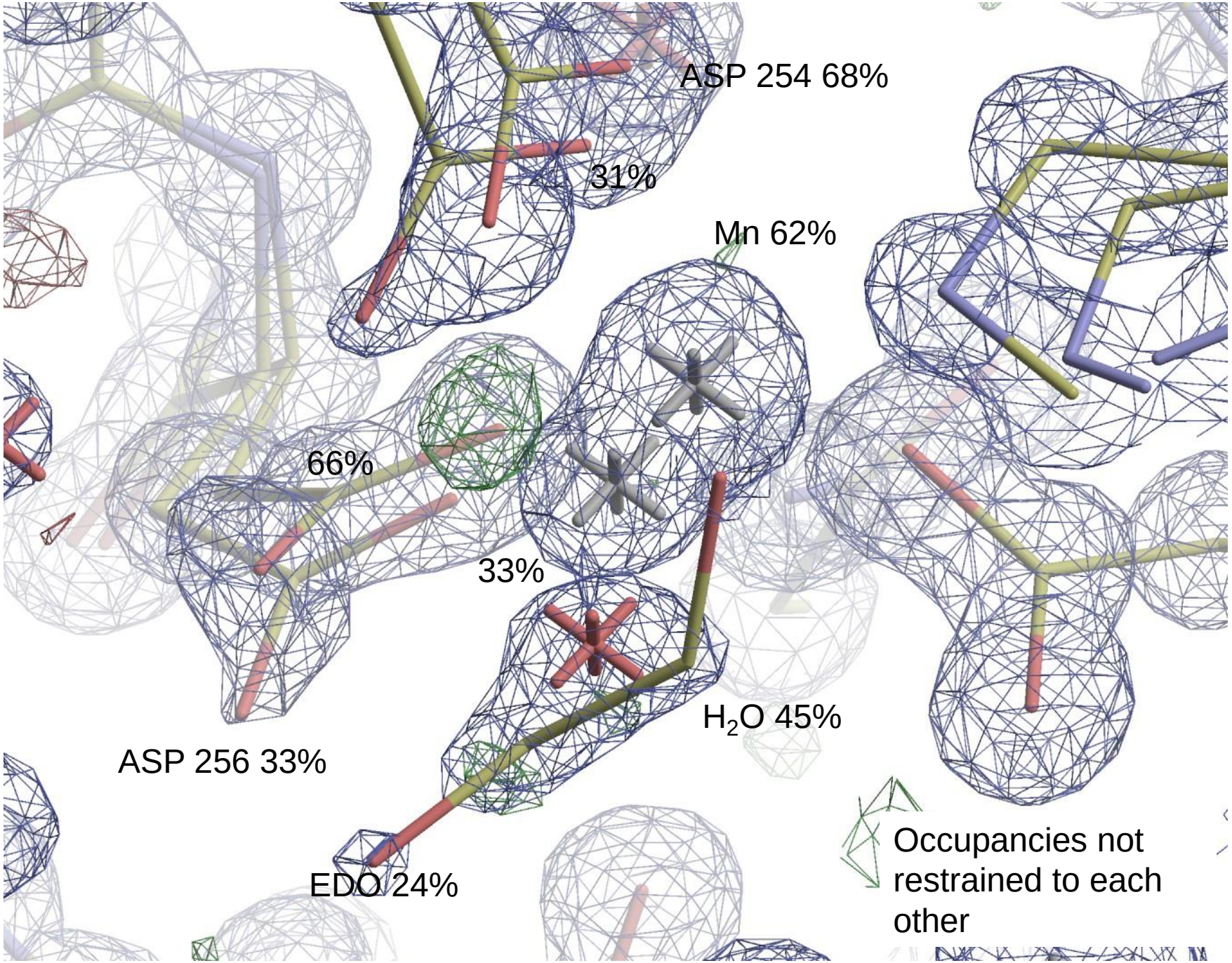

Fo-Fc 2.5σ
2Fo-Fc 2.0σ
189 ASP
142 ALA

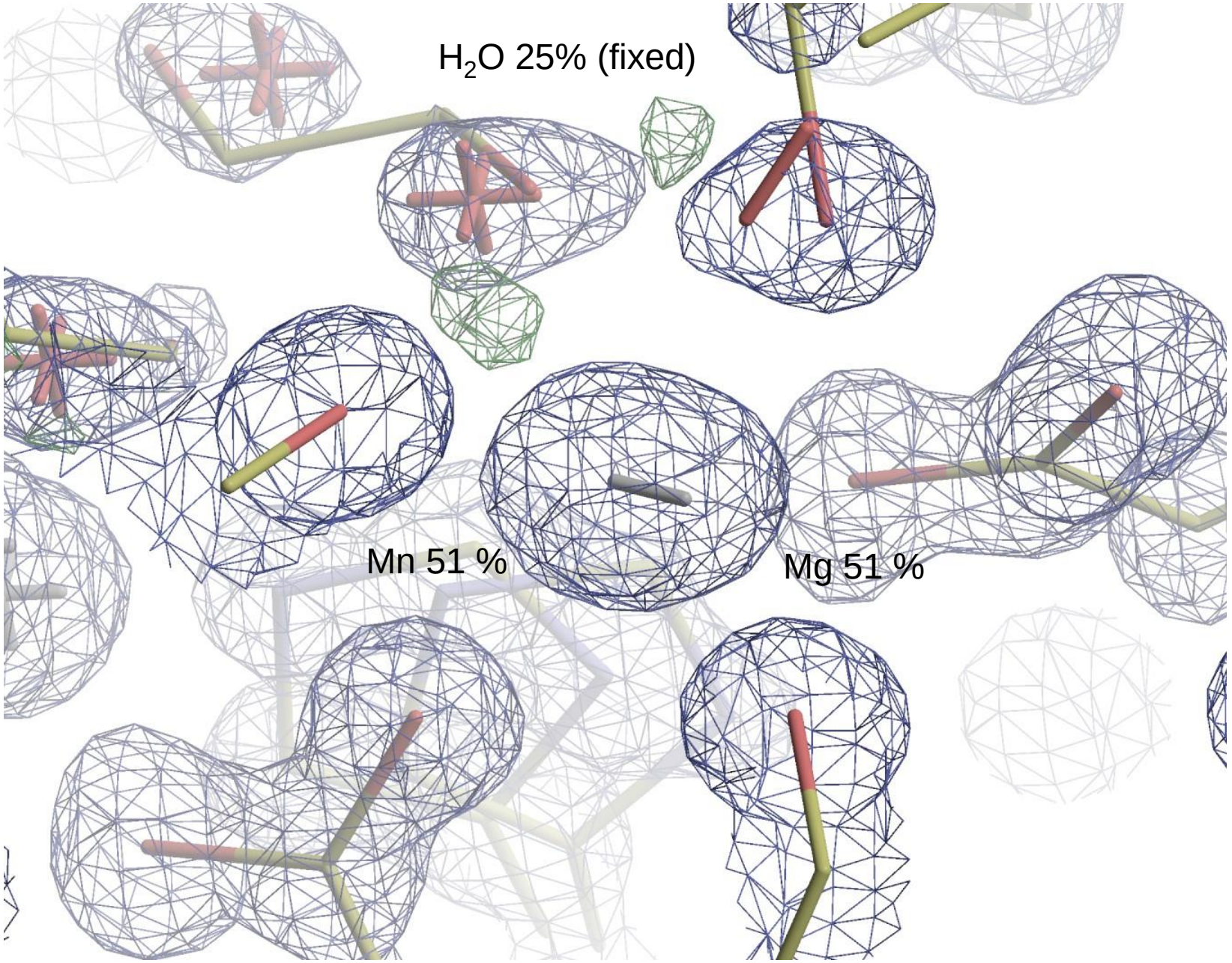
2Fo-Fc 1.0σ

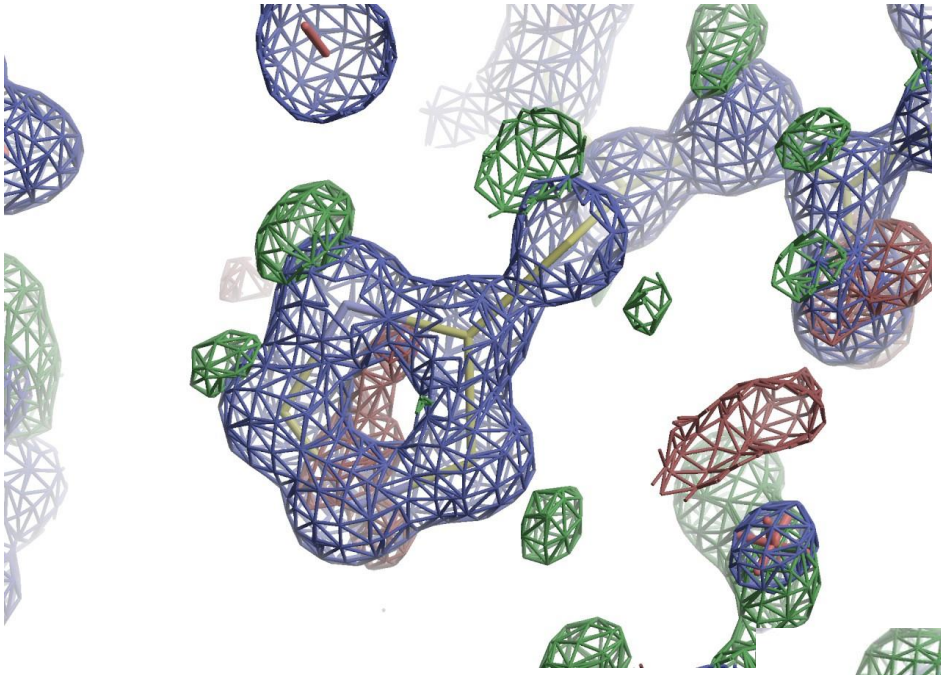
Fo-Fc 5σ

Accidentally changed the occupancy of a water and ethylene glycol



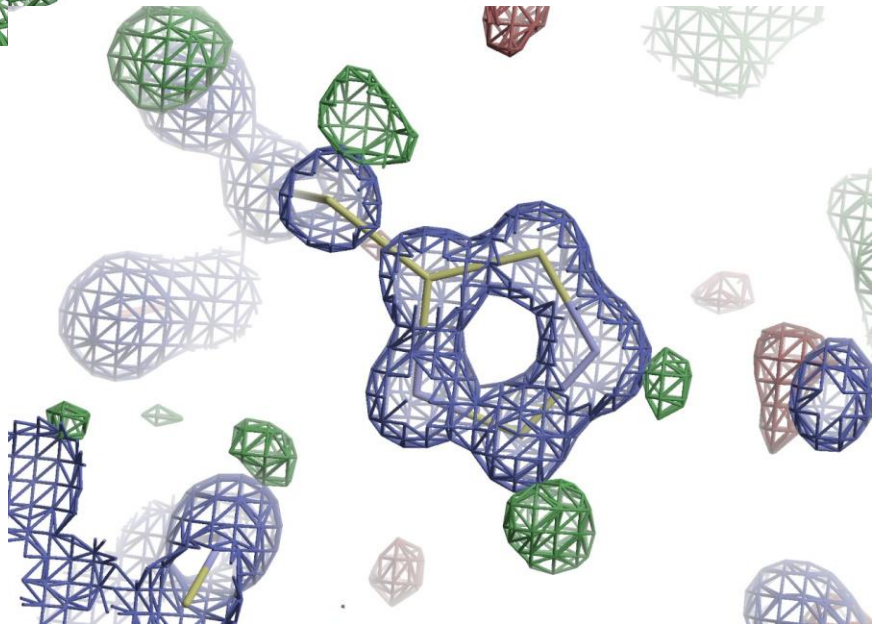

ASP 254 68%
31%
Mn 62%
66%
33%
H_2O 45%
ASP 256 33%
EDO 24%
Occupancies not restrained to each other

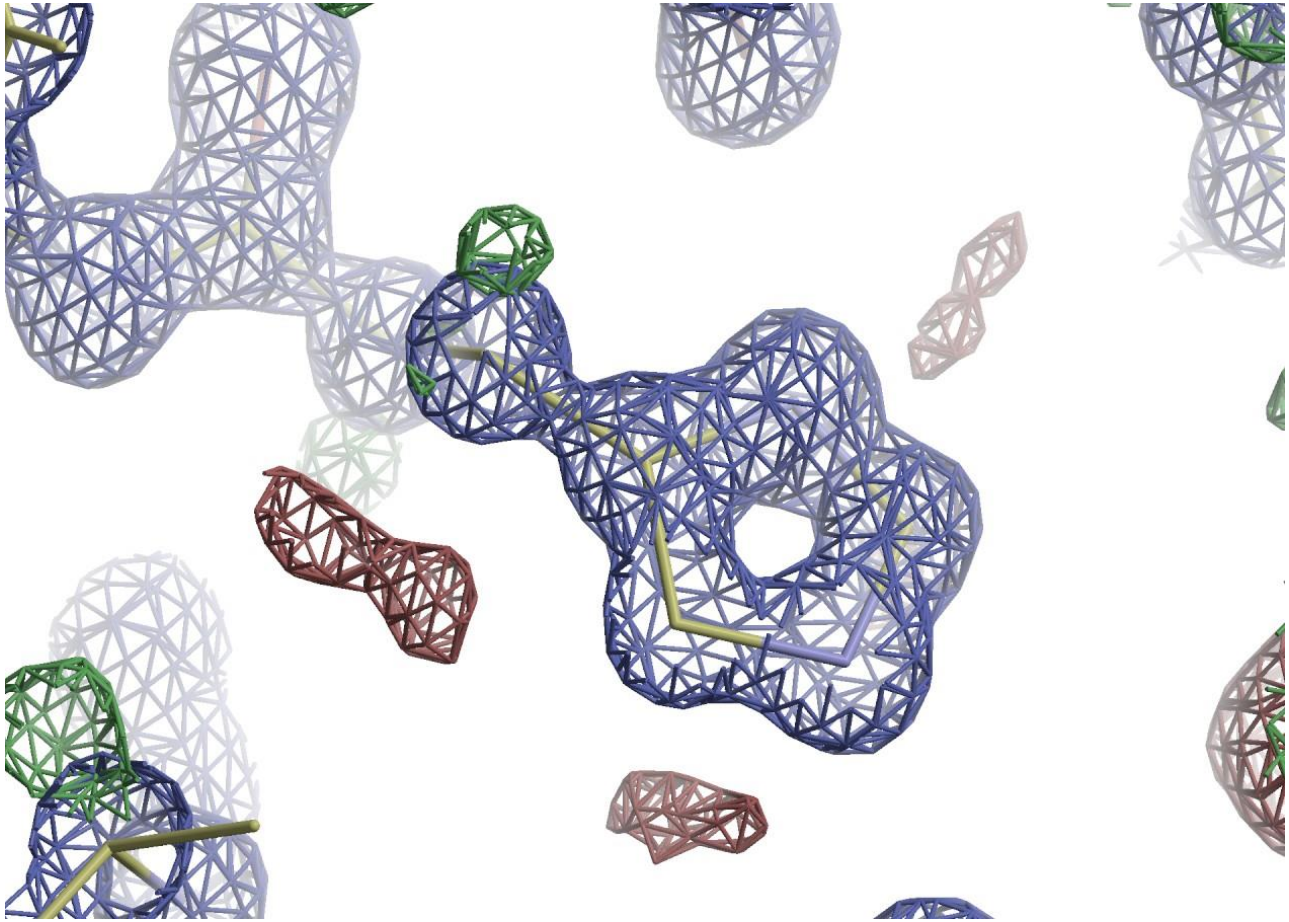

H_2O 25% (fixed)
Mn 51 %
Mg 51 %



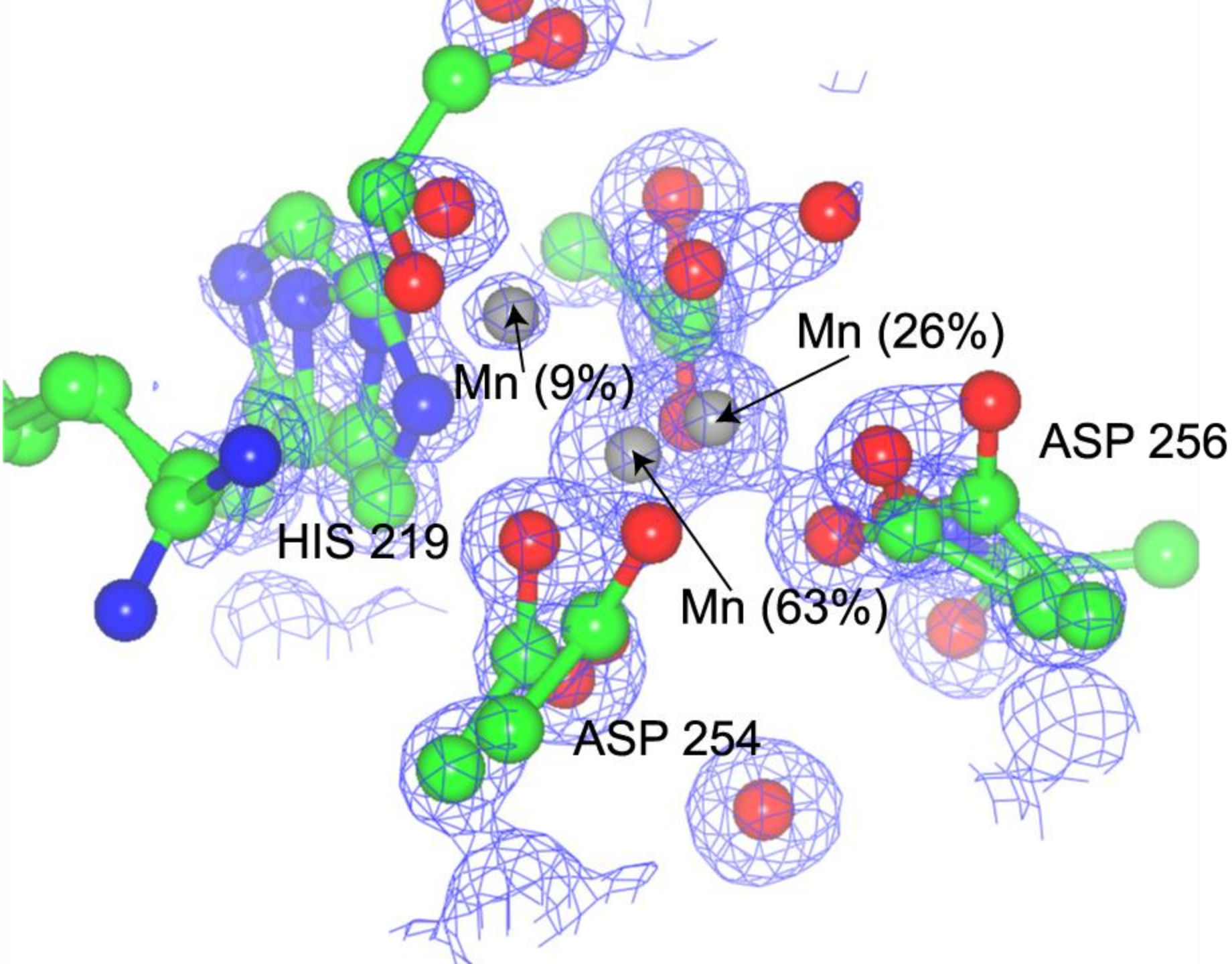
Protonation state of His 53 key to mechanism. Can we determine this?

Possibly, can see electrons on other His residues.



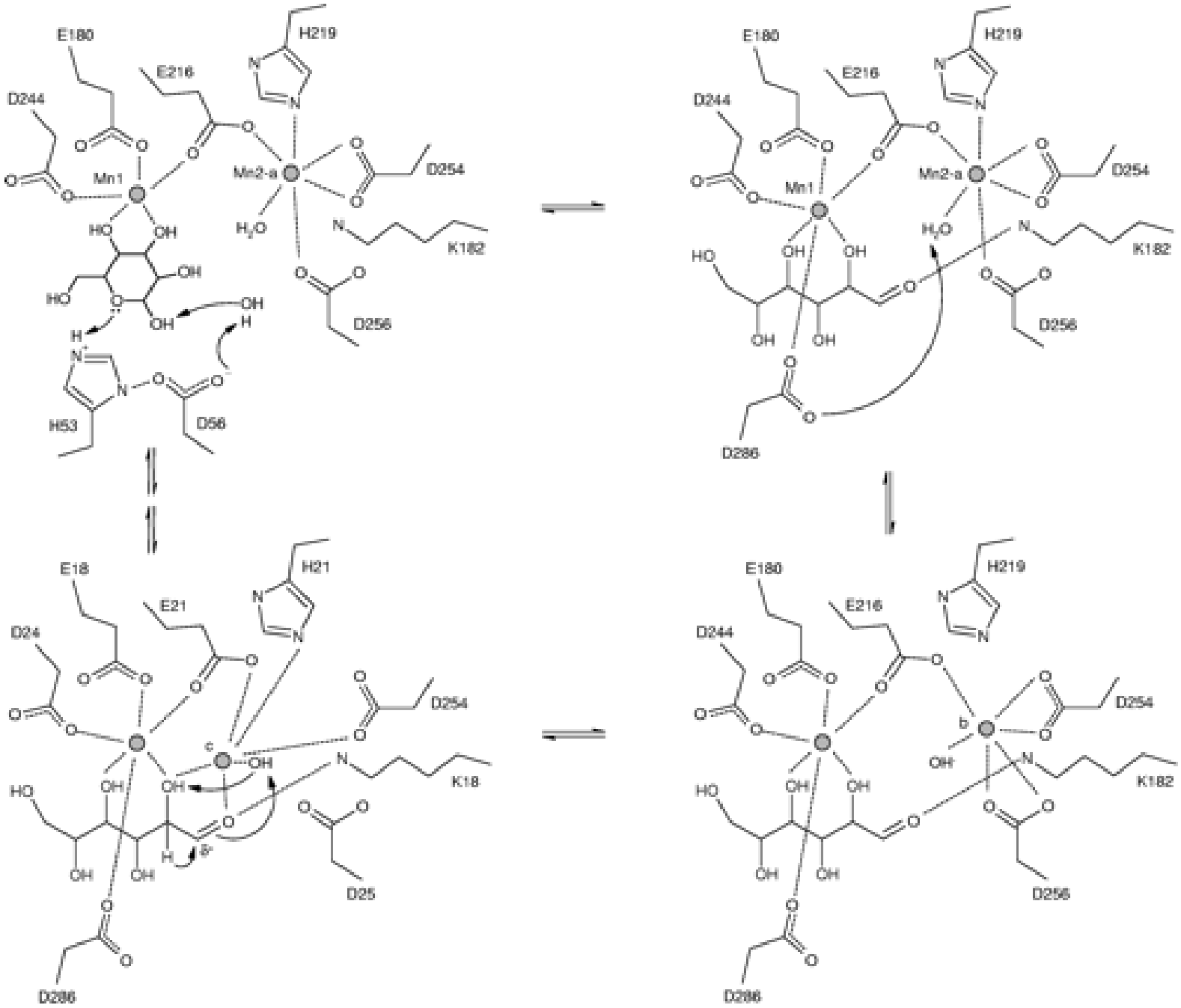


His 53 – no evidence for electrons. Case not conclusive but in agreement with neutron data

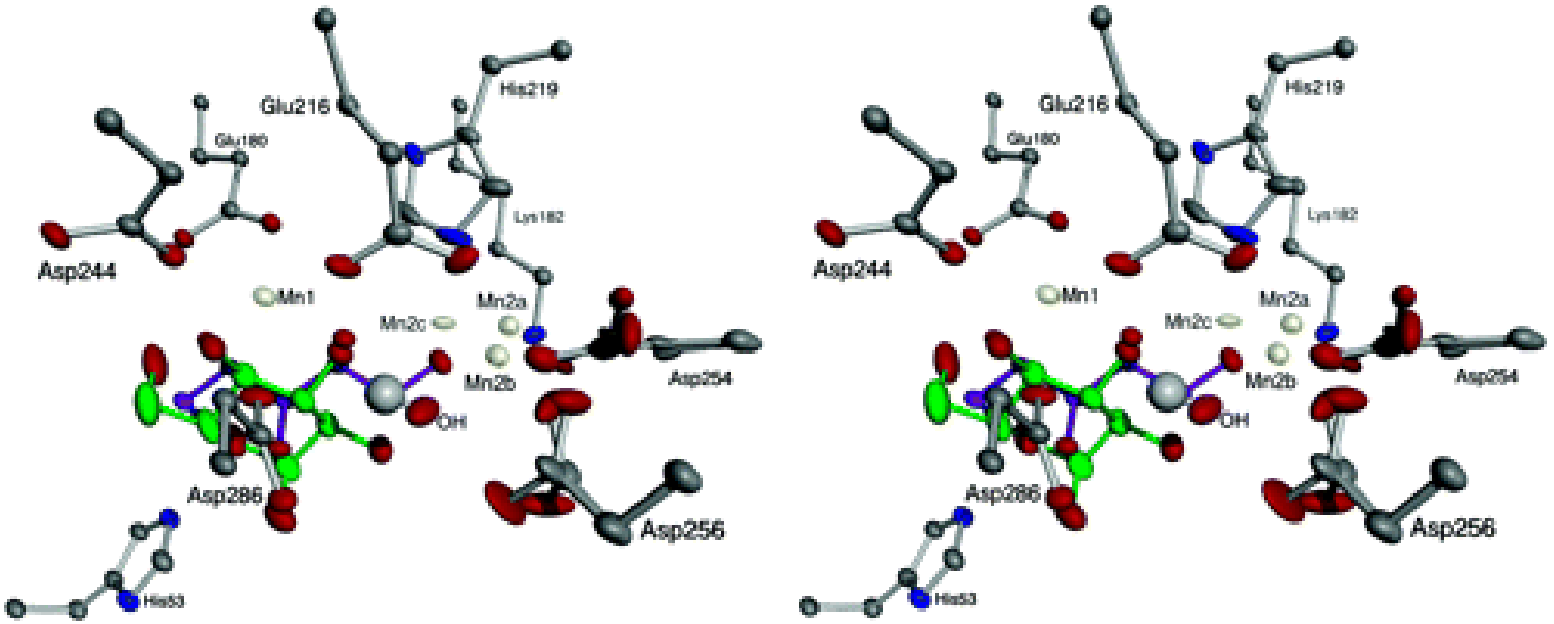

Mn (9%)
Mn (26%)
ASP 256
HIS 219
Mn (63%)
ASP 254

Compare to published results

Comparer aux résultats édités


E180
H219
E216
D244
Mn1
Mn2-a
D254
H2O
K182
D256
H53
D56
D286
E18
H21
E21
D24
K18
D25
b
c
β'

Three different Mn sites

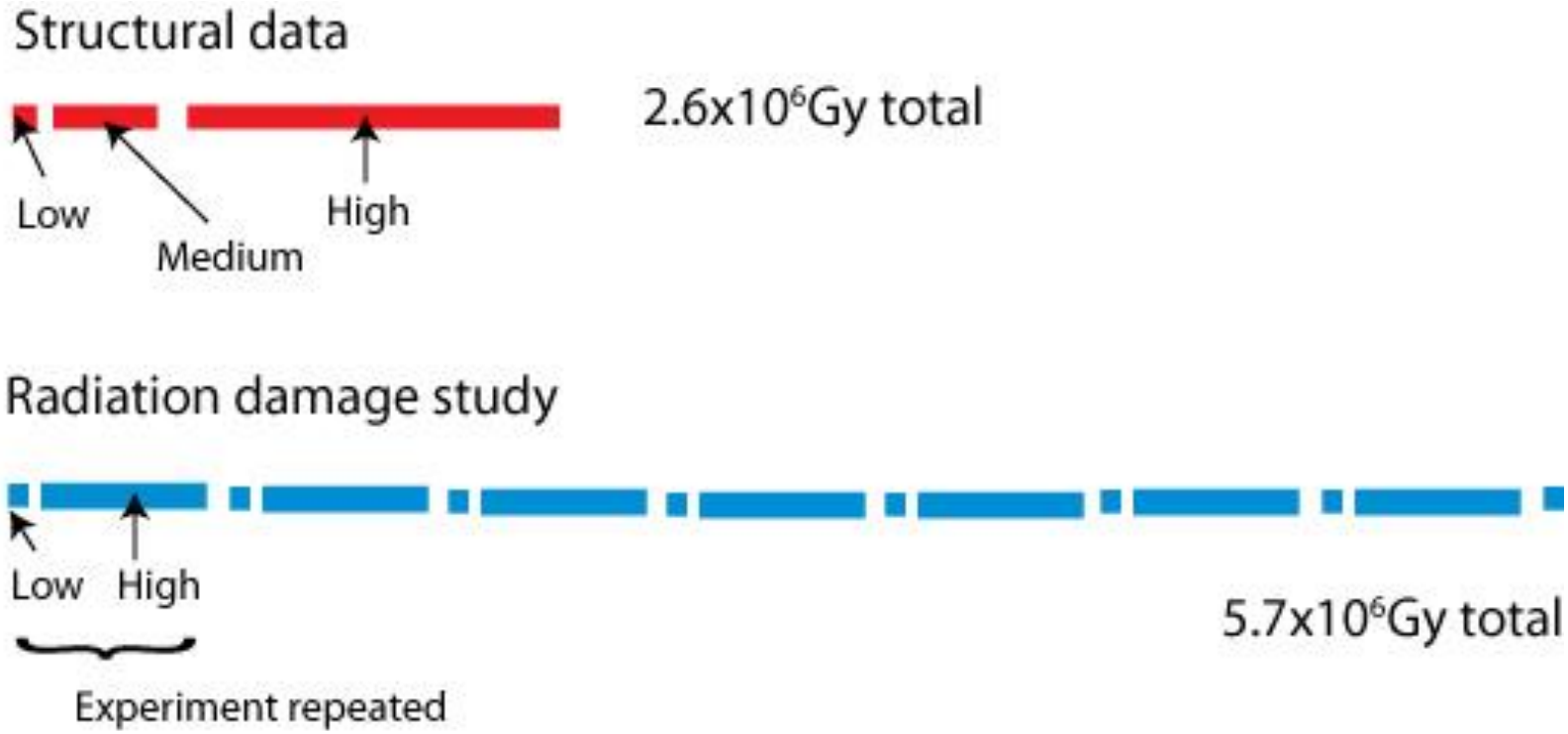


However, the enzyme has a low rate turnover. The three partial occupancy metal sites for one of the metal positions are postulated to explain this.

Xylose isomerase in substrate and inhibitor michaelis states: atomic resolution studies of a metal-mediated hydride shift. Fenn, RInge and Petsko, Biochemistry 2004, 6464-6474.

Where is the radiation damage?

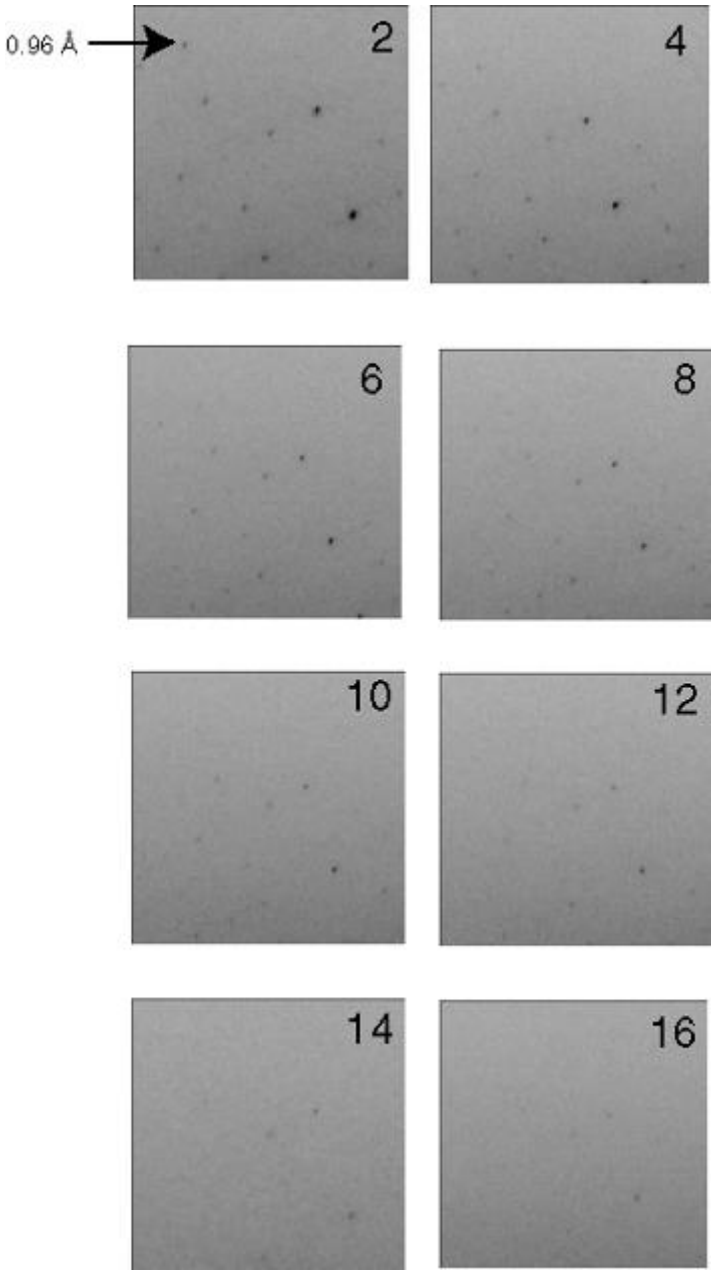
Là où sont les dommages de rayonnement


Structural data
2.6x10⁶Gy total
Low
Medium
High
Radiation damage study
Low
High
5.7x10⁶Gy total
Experiment repeated

The Numbers – Radiation Damage Datasets

High resolution partial data set (0.9 Å)								
Data set	2	4	6	8	10	12	14	16
R_{factor}	6.7(45.8)	6.7(54.7)	6.9(57.5)	7.2(59.4)	7.5(85.2)	8.0(68.7)	8.0(73.5)	8.0(-)
I/σ(I)	8.9(1.6)	8.5(1.2)	8.6(1.0)	8.3(0.8)	8.3(0.7)	8.0(0.6)	7.8(0.6)	7.7(0.5)
Completeness (%)	24.8(24.8)	24.8(23.2)	24.5(19.6)	24.1(15.3)	23.6(10.9)	23.0(7.0)	22.2(3.1)	21.7(1.4)
Redundancy	1.4(1.4)	1.4(1.3)	1.4(1.2)	1.3(1.2)	1.3(1.1)	1.3(1.1)	1.3(1.0)	1.3(1.0)
Mosaicity (º)	0.17	0.17	0.17	0.16	0.16	0.16	0.16	0.16
B_{factor}	6.04	6.35	6.70	6.85	7.25	7.54	7.85	8.13
Medium resolution complete data set (1.2 Å)								
Data set	3	5	7	9	11	13	15	
R_{factor}	7.5(22.5)	7.5(24.7)	7.7(27.3)	7.6(30.1)	7.9(33.4)	7.9(37.3)	7.8(41.7)	
I/σ(I)	16.8(5.0)	16.6(4.7)	16.4(4.3)	16.6(3.9)	16.1(3.3)	15.4(2.8)	15.3(2.4)	
Completeness (%)	99.7(99.3)	99.7(99.4)	99.7(98.9)	99.7(99.1)	99.7(98.4)	99.6(96.8)	99.4(93.7)	
Redundancy	3.6(3.2)	3.6(3.3)	3.5(3.2)	3.5(3.1)	3.5(2.8)	3.5(3.0)	3.5(2.8)	
Mosaicity (º)	0.14	0.14	0.14	0.14	0.14	0.14	0.14	
B_{factor}	8.77	8.83	9.07	9.61	9.83	10.31	10.86	

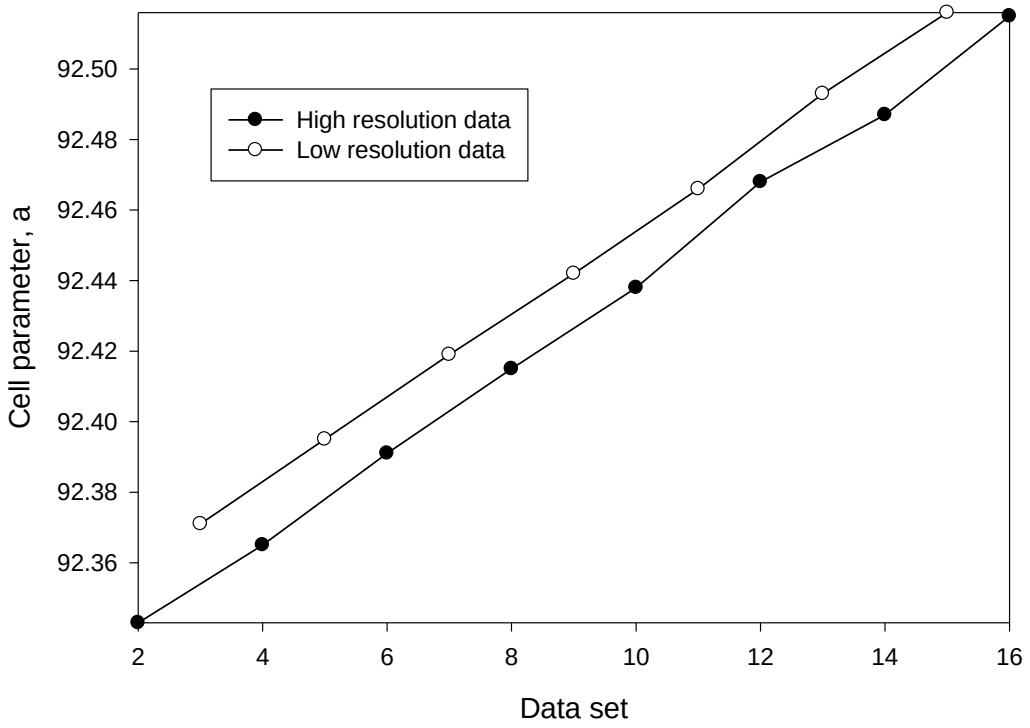
With each data set R_{factor} increases, signal-to-noise, completeness, and redundancy decreases. The mosaicity is unchanged, we are just seeing the beam contributions. The B_{factor} increases.

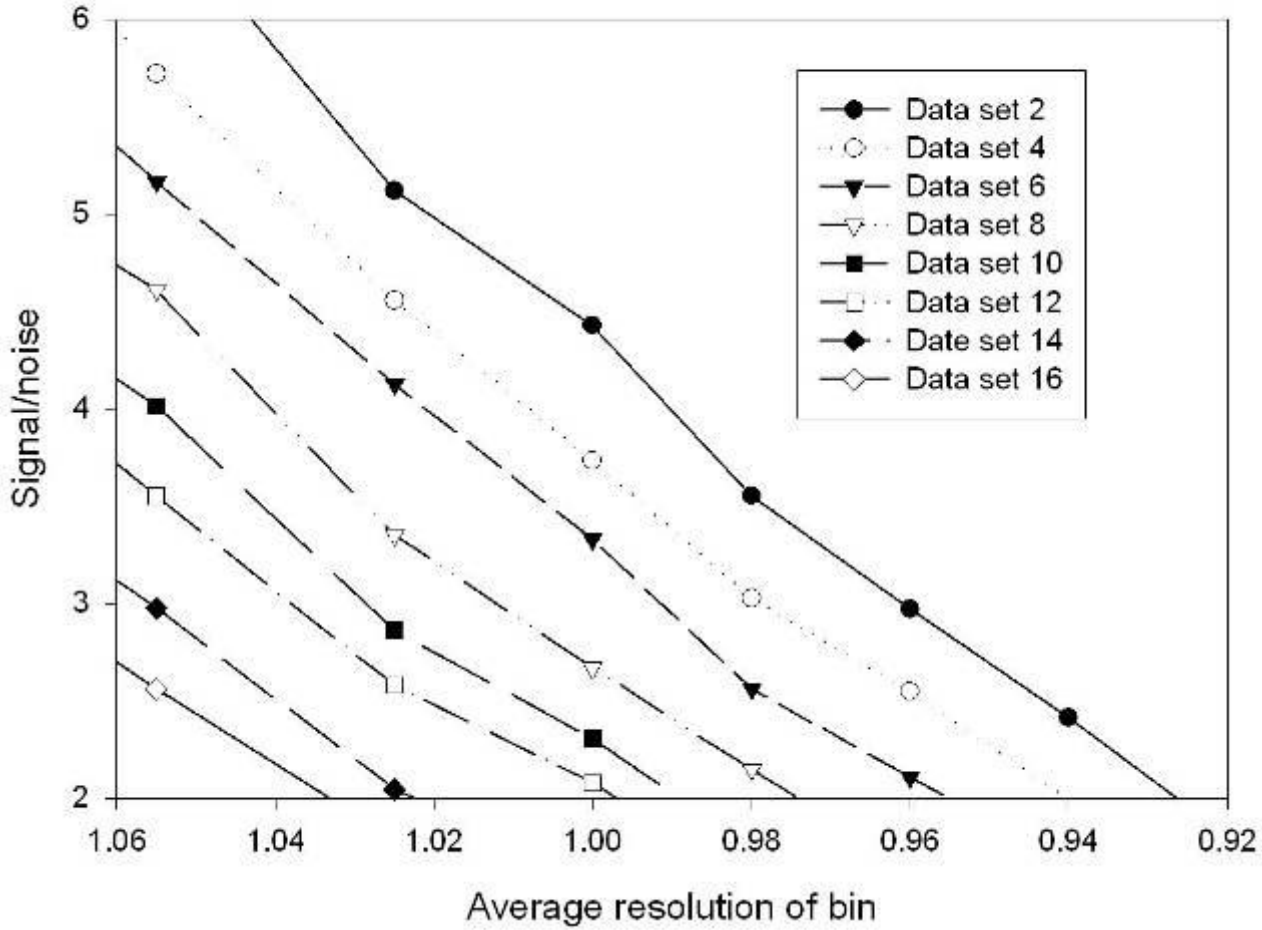


The Images

Same portion of high resolution data showing gradual decay of reflections.

Note that the background radiation remains constant

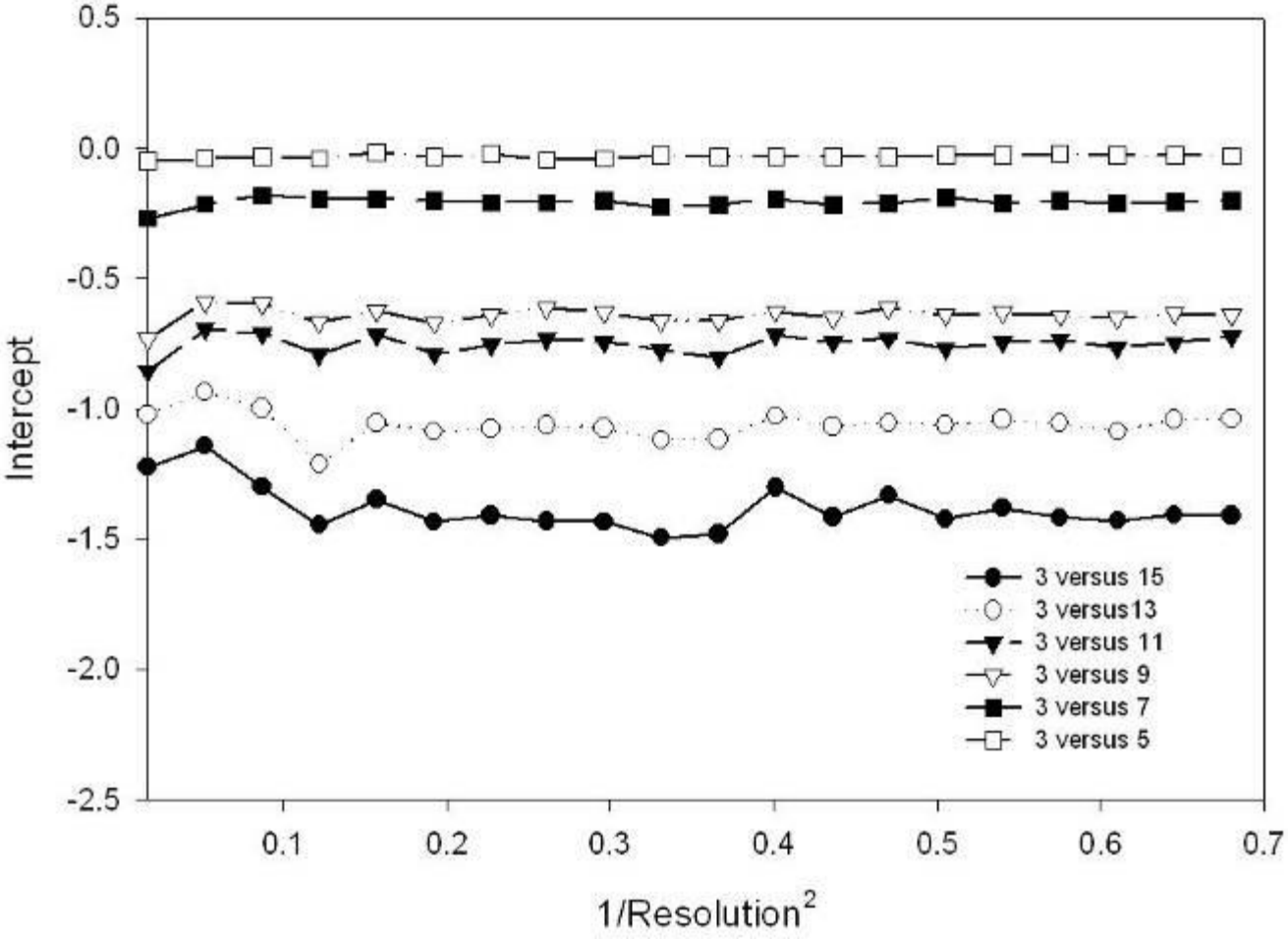


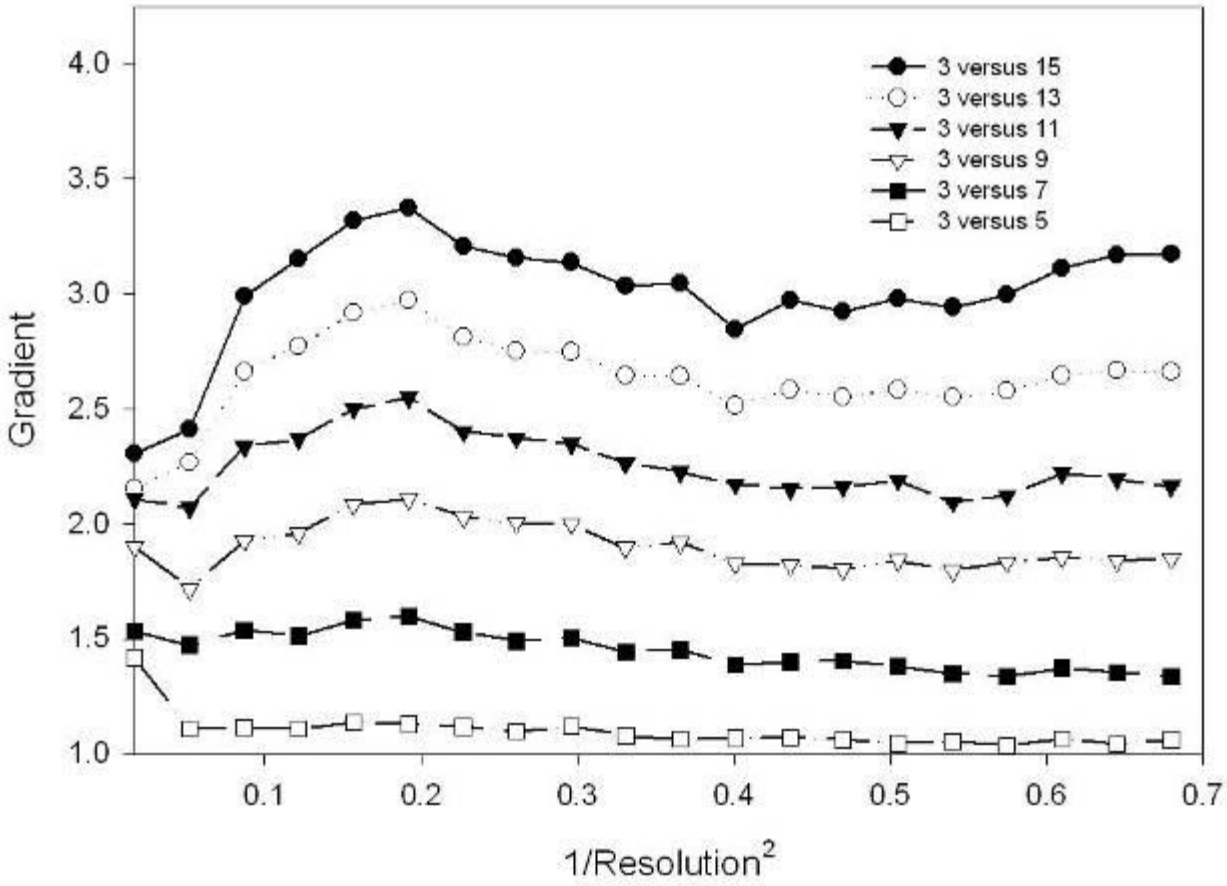

Signal/noise
6
5
4
3
2
1.06
1.04
1.02
1.00
0.98
0.96
0.94
0.92
Average resolution of bin
Data set 2
Data set 4
Data set 6
Data set 8
Data set 10
Data set 12
Date set 14
Data set 16

The normal probability plot

- The normal probability plot (Chambers 1983) is a graphical technique for assessing whether or not a data set is approximately normally distributed. The data are plotted against a theoretical normal distribution in such a way that the points should form an approximate straight line. Departures from this straight line indicate departures from normality.
- Normal probability plots in crystallography indicate structural changes between data sets if the intercept and gradient of the plot diverge from zero and one respectively. In the xylose isomerase case there is a decrease in the intercept and increase in the gradient as a function of dose. Structurally significant changes are occurring.

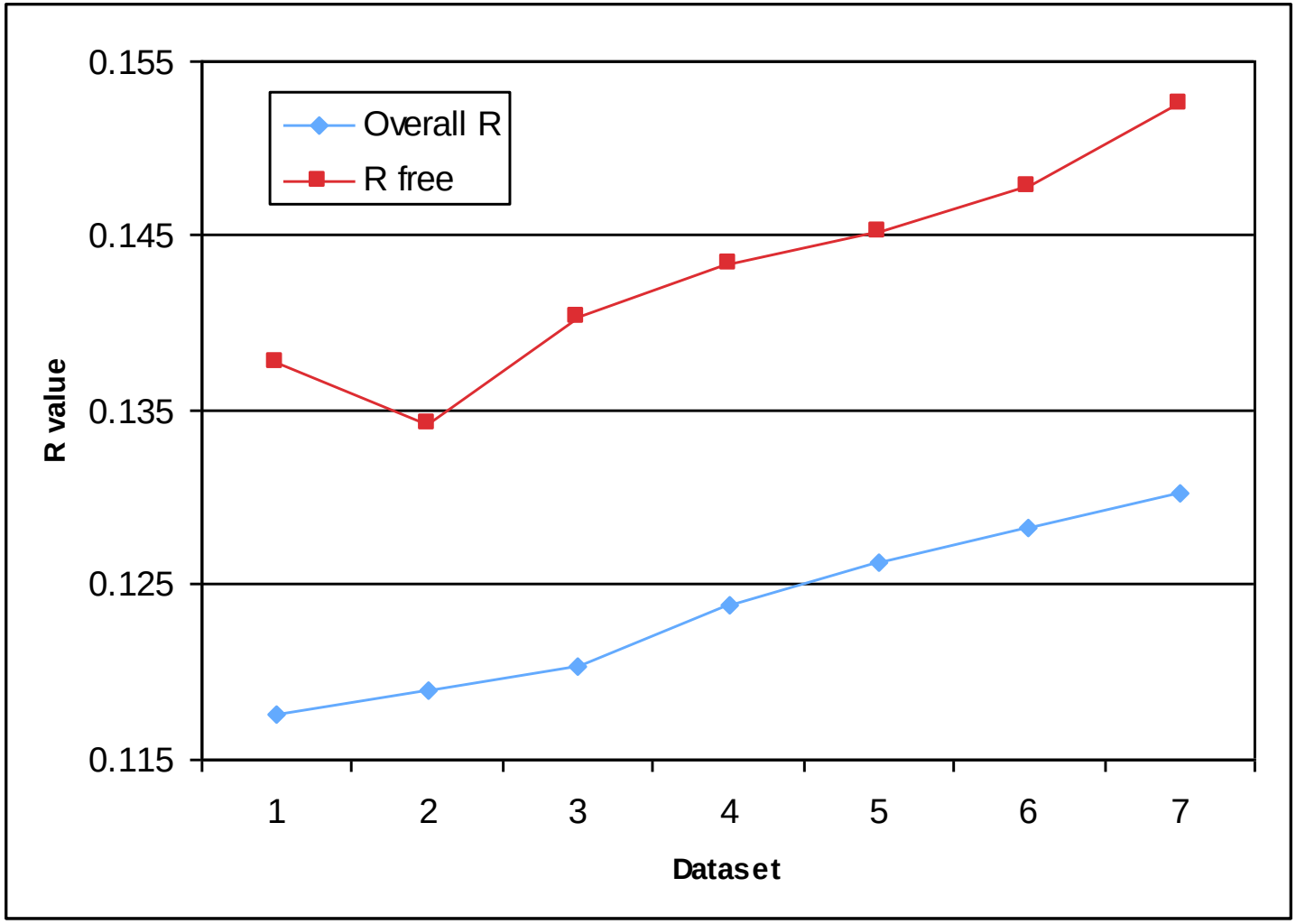
- Howell, P. L.; Smith, G. D., Identification of heavy-atom derivatives by normal probability methods. Acta Cryst A 1992, 25, 81-86.

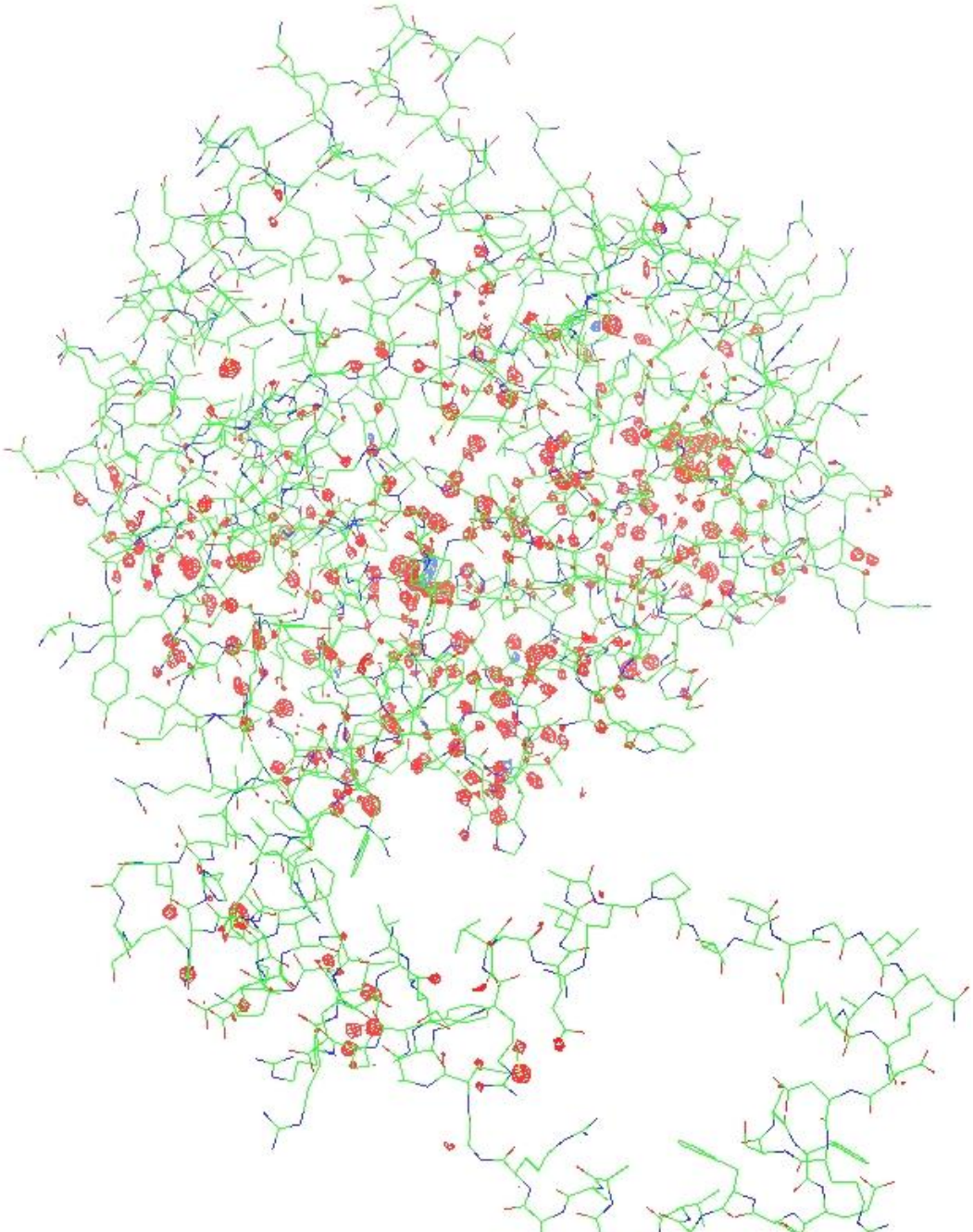

0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
Intercept
0.1
0.2
0.3
0.4
0.5
0.6
0.7
1/Resolution^2
3 versus 15
3 versus13
3 versus 11
3 versus 9
3 versus 7
3 versus 5

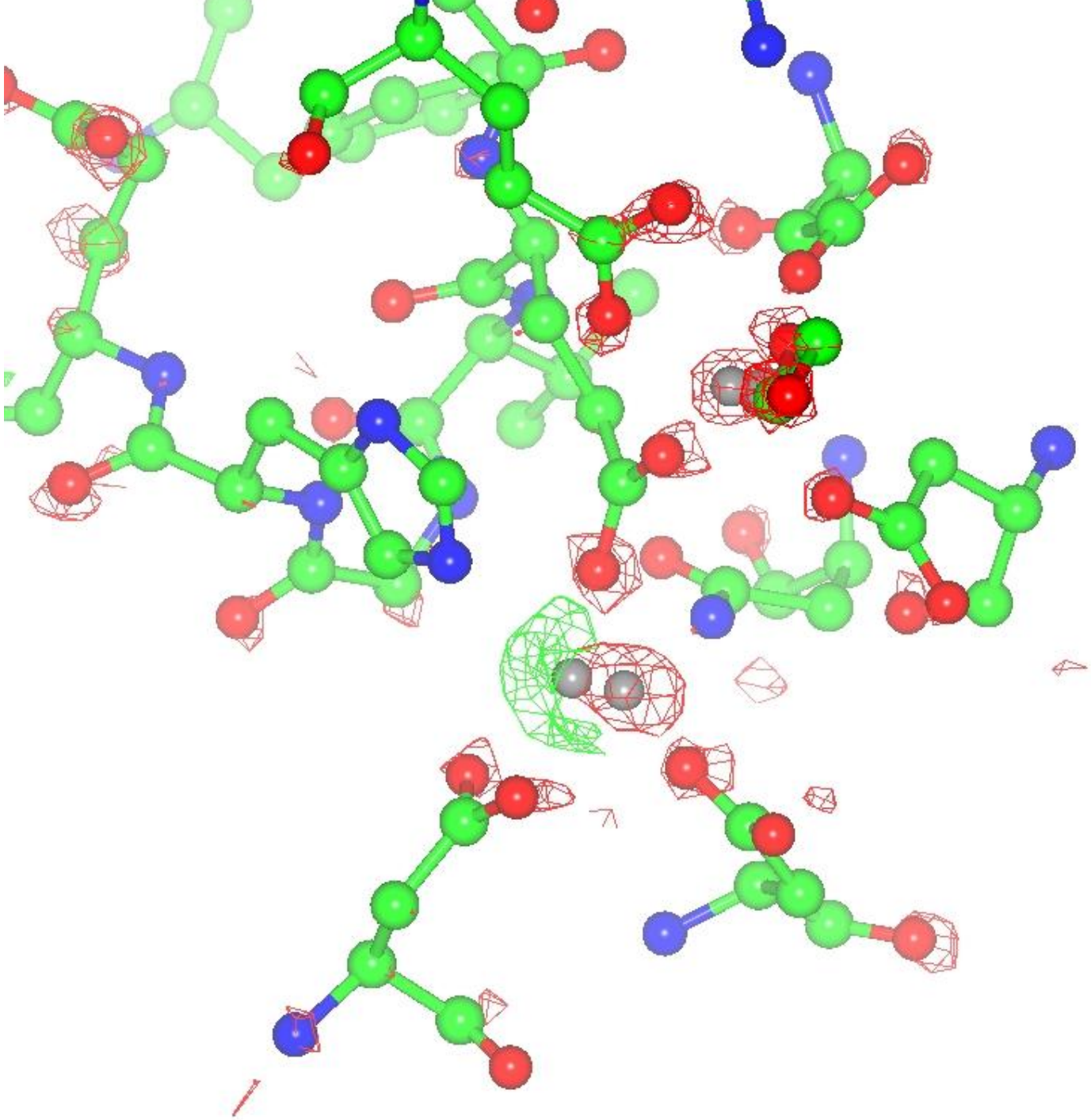

3 versus 15
3 versus 13
3 versus 11
3 versus 9
3 versus 7
3 versus 5
Gradient
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
1/Resolution^2

Are there structural consequences?

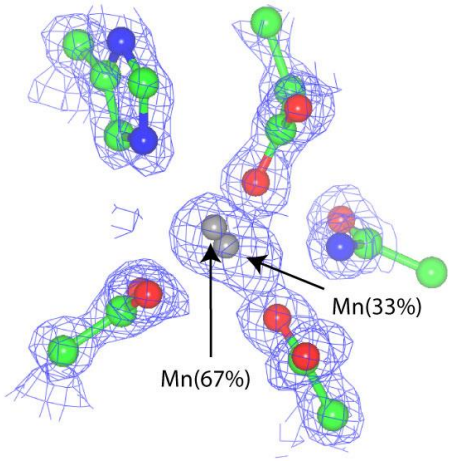
Yes, but need to determine the structure to see what those consequences are.


0.155
0.145
0.135
0.125
0.115
R value
Overall R
R free
1
2
3
4
5
6
7
Dataset

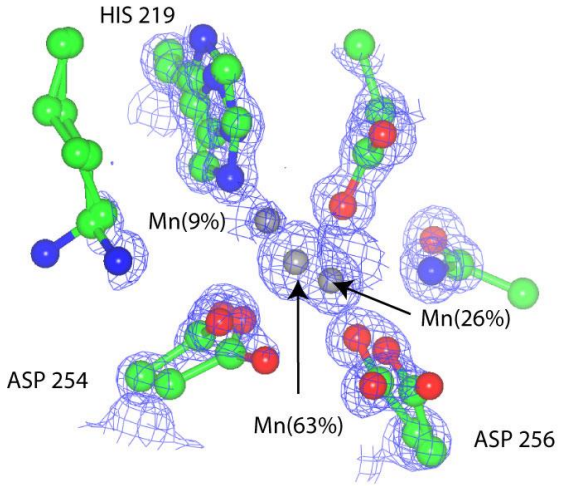




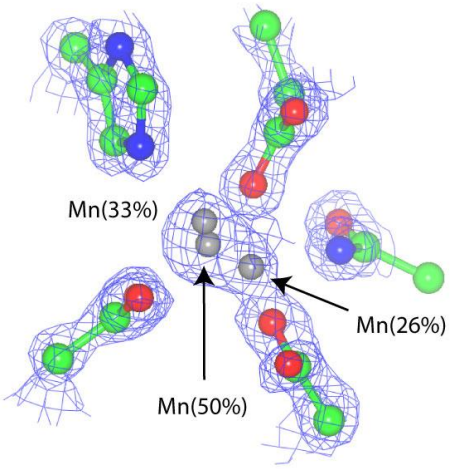
High-resolution data set



(a) 0.8 x 10^6 Gy for dataset



(b) 2.6 x 10^6 Gy dose for dataset



(c) 5.7x 10^6 Gy total exposure, 0.8 x 10^6 Gy for dataset

Radiation damage set 3.

Radiation damage set 15.

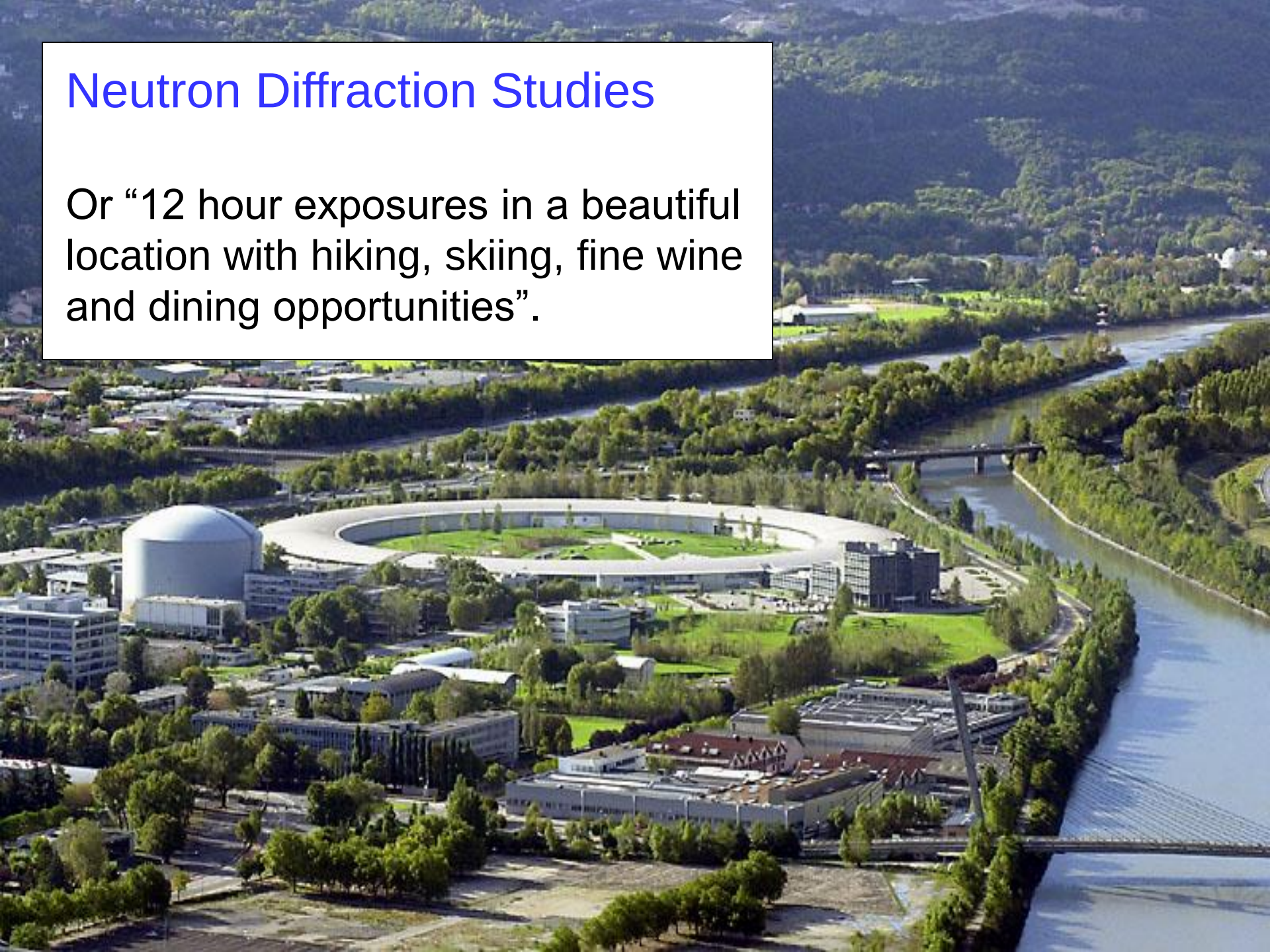
Mechanistic implications

- The proposed reason for low enzymatic turnover, i.e. the three alternate metal sites is not a biological ‘truth’.
- The alternative sites and change in occupancy is driven by the observation.
- As a function of X-ray dose, multiple sites appear and occupancy changes.
- The measurement drives the observation!

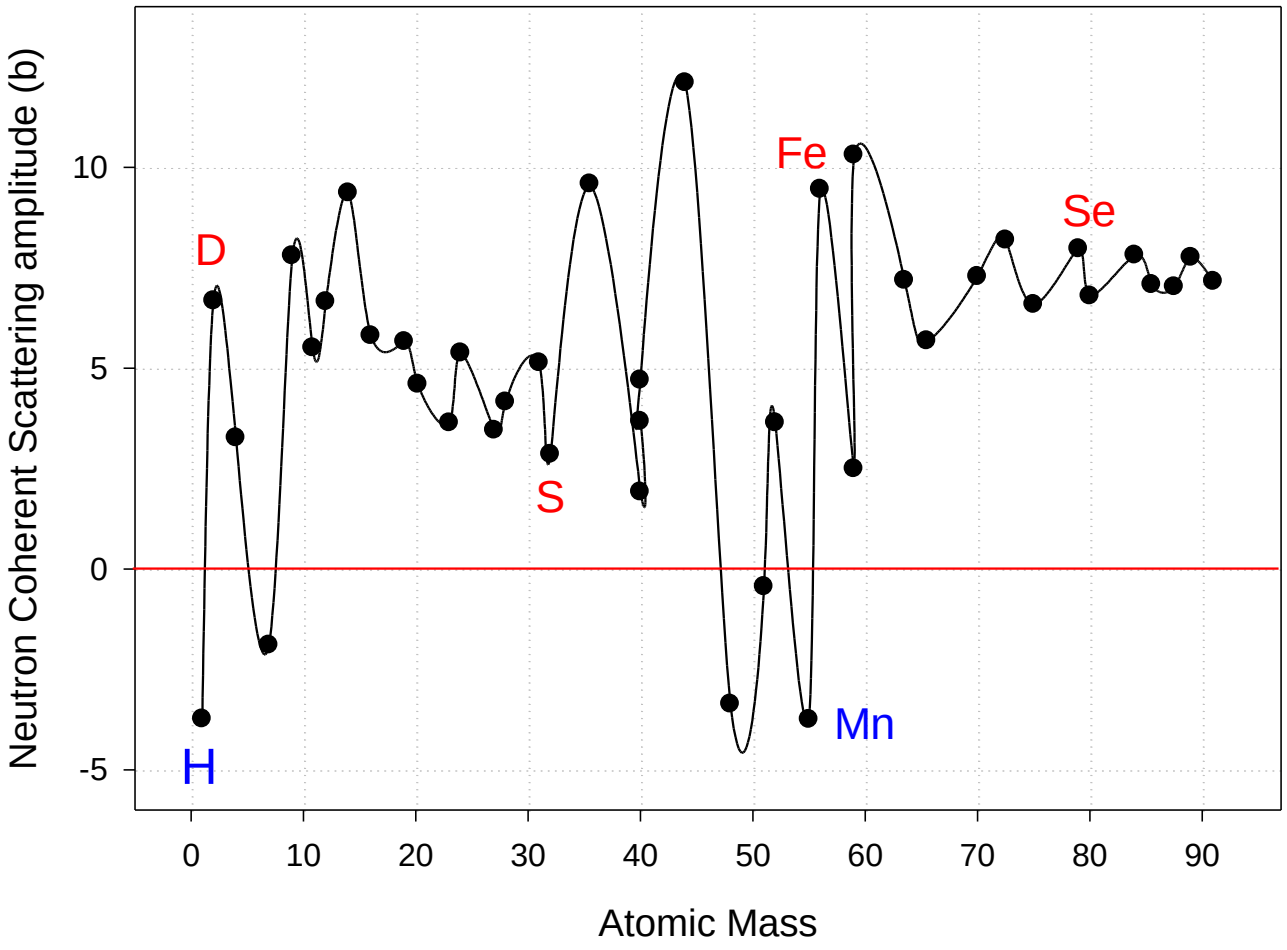
How to get round this?

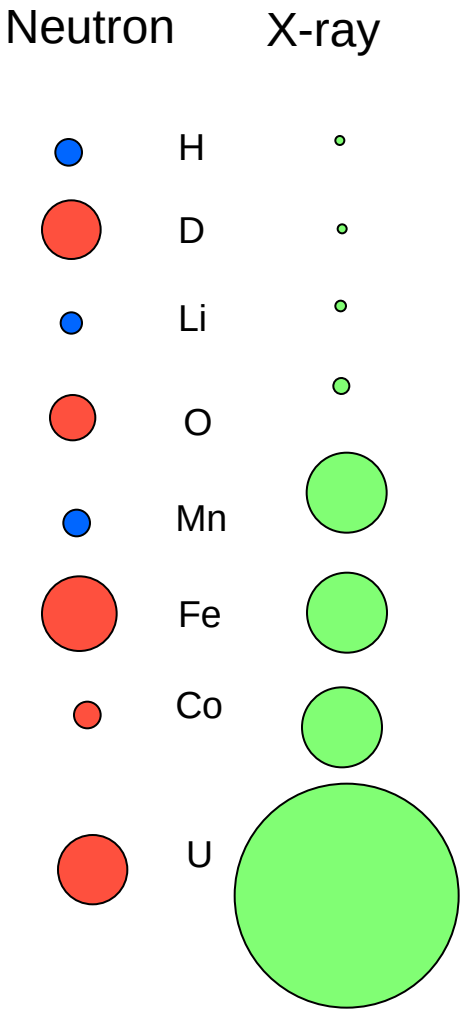
Neutron Diffraction Studies

Or “12 hour exposures in a beautiful location with hiking, skiing, fine wine and dining opportunities”.



Or, to answer it more graphically, what a neutron sees:





Why make use of neutrons?

- For neutron diffraction the scattering amplitudes vary from element to element in a non systematic way - atoms of similar atomic mass can be easily distinguished.
- The scattering amplitude of hydrogen is of the same order of magnitude as the amplitudes of other atoms typically found in biological molecules - hydrogen atoms can be seen thereby;
  - revealing whether a particular acidic group is dissociated or has a hydrogen atom bound to it,
  - discriminating between water and hydroxyl anion in the active site of an enzyme,
  - determining the orientation of a water molecule etc.
- Deuterium and hydrogen have opposite sign scattering amplitudes enabling contrast matching techniques.
- Radiation damage is not a concern.

Problems with neutrons:

Neutron sources have low fluxes:

- For example, the LADI (Laue Diffractometer) experimental station at Insitiute Laue Langevin has a flux of $3x10^7$ neutrons cm^{-2} s^{-1} for a partially monochromatised beam (l=3.5 A, $\delta\lambda/\lambda$=20%). A monochromatic beam from a wiggler source on a synchrotron has 10 orders of magnitude greater flux.

Neutrons are weakly scattered

- Neutrons are electrically neutral and interact weakly with matter, they are scattered by the nucleus and unpaired electrons.

Combine neutrons and X-rays

Spallation Neutron Source, Oak Ridge Tennessee

Summary

- Radiation damage alters the structure.
- It can provide misleading results.
- It occurs as a function of dose (or resolution).
- Neutron offer a non-ionizing way of benchmarking the atomic positions to determine the degree of damage present in the X-ray structure.

Acknowledgements

- Kristin Wunsch
- Mark van der Woerd
- Russell Judge
- Ann Wojtaszczyk
- Dean Myles
- Flora Meilleur
- The staff at Hauptman-Woodward Medical Research Institute