PAUL SCHERRER INSTITUT

PSI

Dr. Tim Grüne :: Paul Scherrer Institut :: tim.gruene@psi.ch

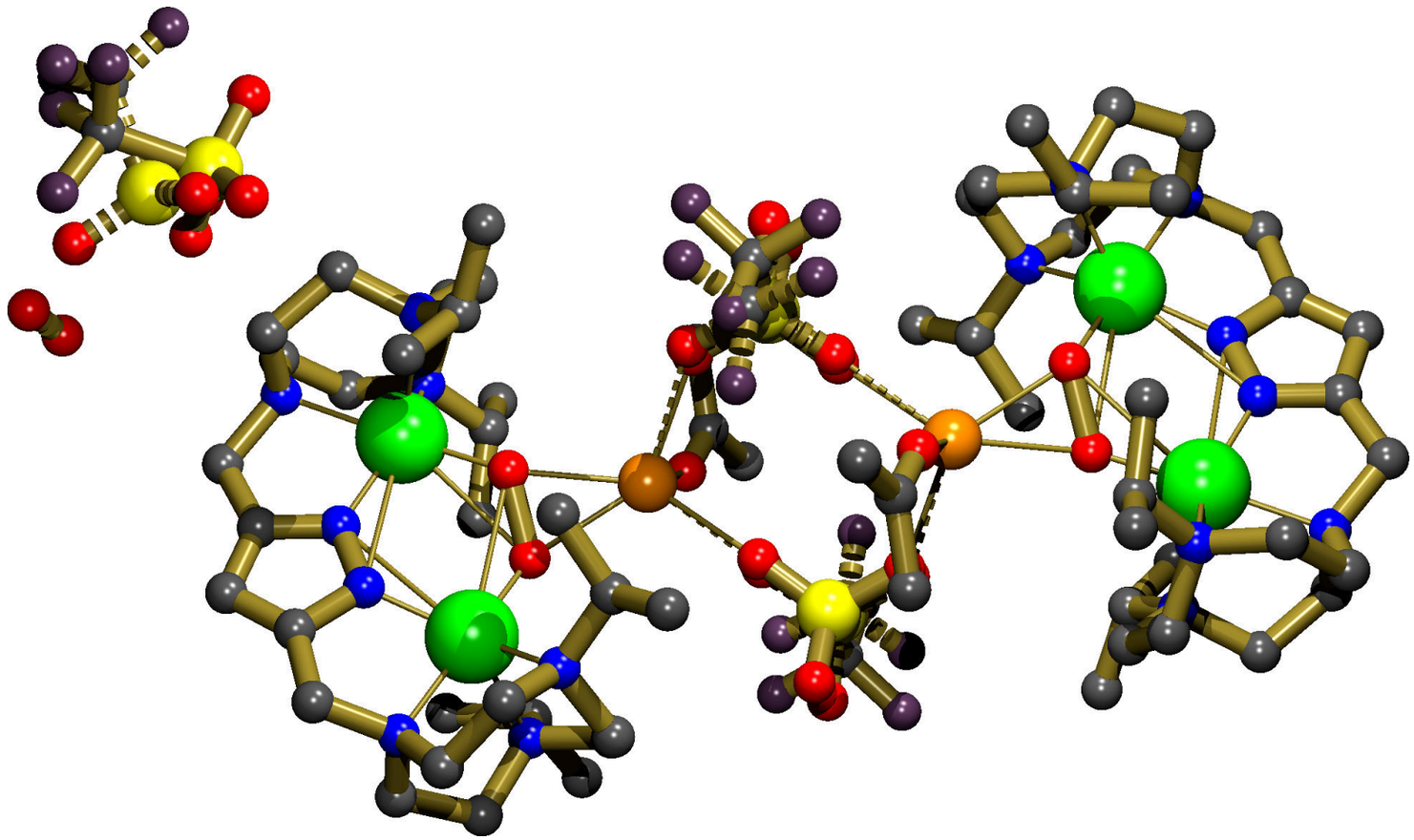
Introduction to Structural Chemistry with (Electron) Crystallography

PSI — Center for Radiopharmaceutical Sciences

13th October 2017

1 - What is a “Structure”

3D Coordinates *aka* 3D Conformation



Bioinorganic Chemistry: $Cu_2^{II}(\mu-\eta^1:\eta^1-O_2)$ *cis*-peroxo
(Dalle *et al.*, J. Am. Chem. Soc. (2014), 136, 7428–7434)

3D Coordinates *aka* 3D Conformation

```
SFAC    C N O Na S Cu H
Na     4     0.876953    0.592795    0.604729  11.00000    0.07618
C      1     0.774060    0.654005    1.000931  11.00000    0.04574
AFIX   43
H      7     0.757349    0.661067    1.053967  11.00000   -1.20000
AFIX    0
CuA    6     0.833839    0.782664    0.759284  11.00000    0.04565
N1A    2     0.801899    0.692307    0.868151  11.00000    0.05391
```

Structure description in SHELX–format (`http://shelx.uni-goettingen.de`)

1. Molecules consist of atoms
2. Atom type (SFAC - scattering factor)
3. Atom coordinate: X, Y, Z
4. Atomic displacement parameter (ADP): thermal vibration

Crystal Structures

- A crystal structures is composed of one or multiple molecules
- The structure provides the coordinates of the atoms (and their “vibration”)
- Precision of bond lengths, bond angles is low compared with e.g. spectroscopic methods
- The three dimensional information is rather unique.

Access to Crystal Structures

Many journals require deposition of model coordinates at a crystallographic data base. Common data bases:

Cambridge Structural Data Bank

Crystallography Open Database

Inorganic Crystal Structure Database

Protein Data Bank

Nucleic Acids Data Bank

CRYSTMET ®

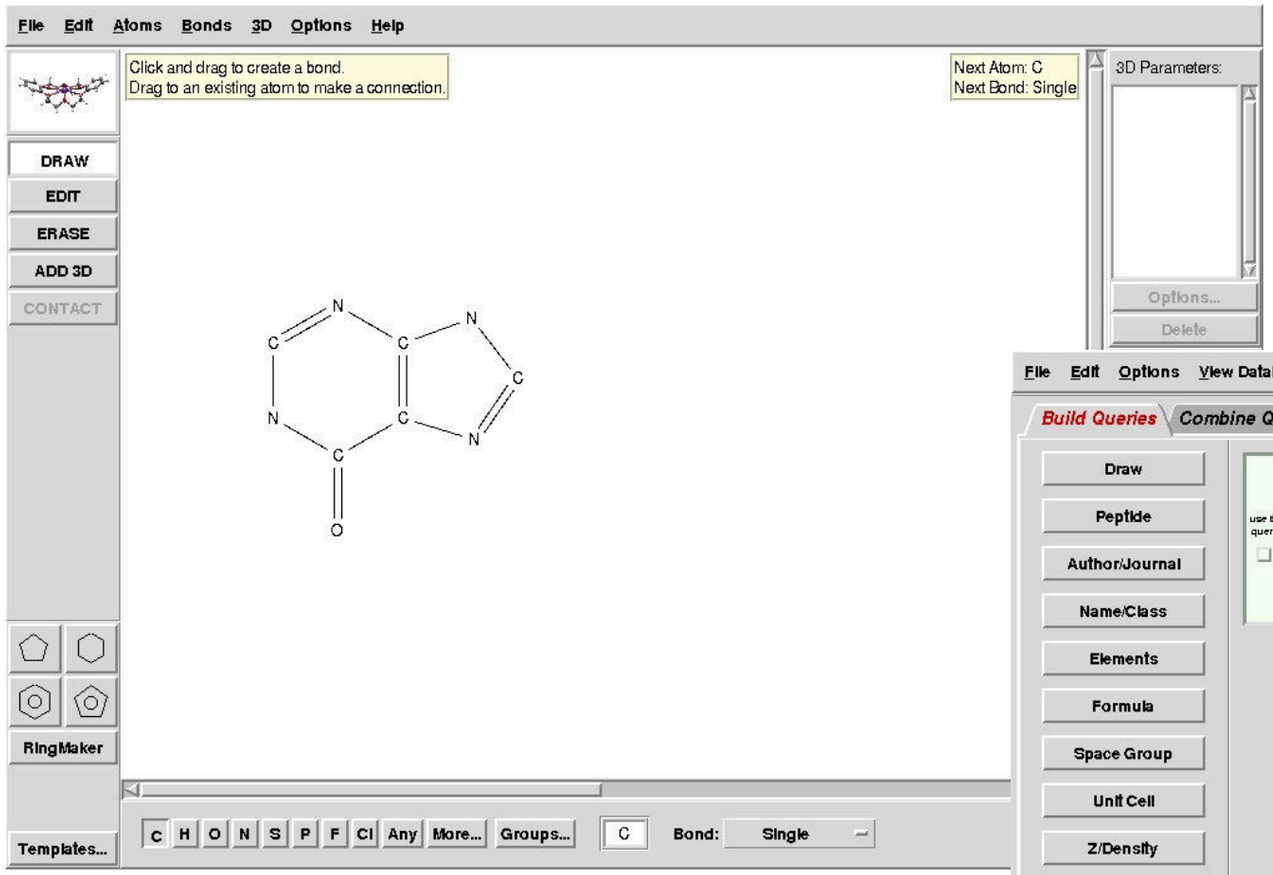
CSD

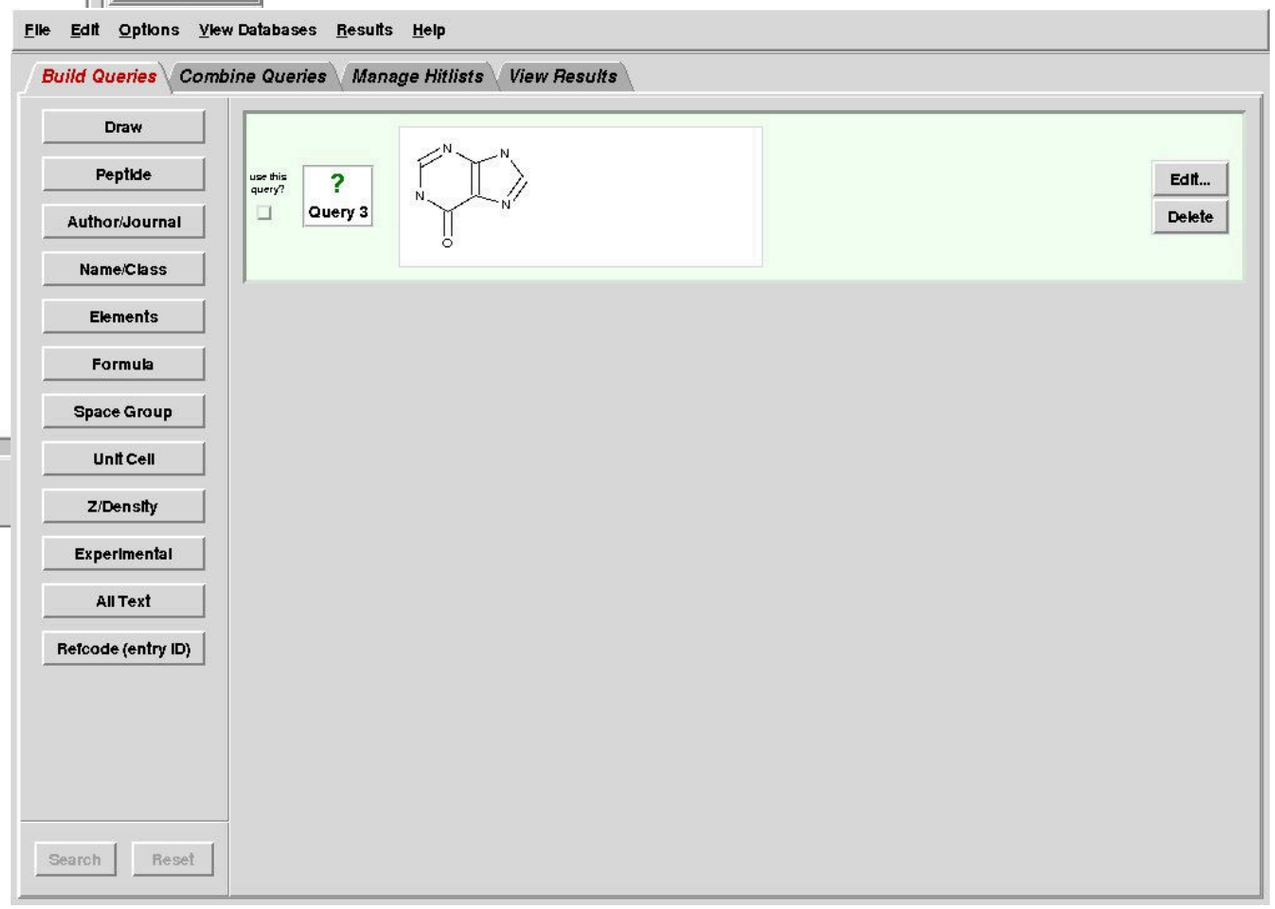
Cambridge Structural Database http://www.ccdc.cam.ac.uk

"The world repository of small molecule crystal structures" for organic und metallo–organic compounds

- Founded 1965
- Crystal structures from X–ray and neutron diffraction (some electron diffraction)
- Single crystal and powder diffraction structures
- Every structure is curated
- More than 800 000 entries, approximately 40 000 per year

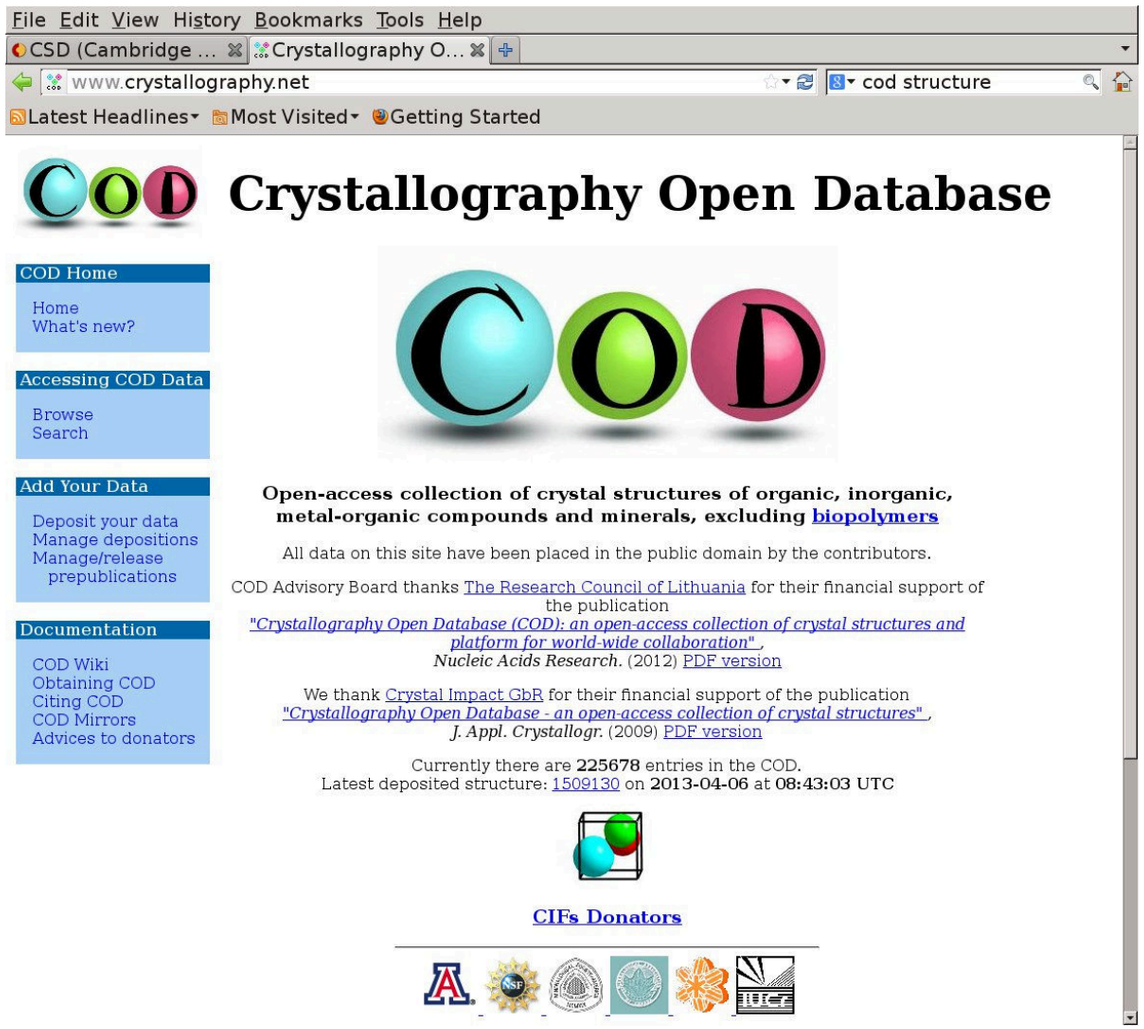
CSD — Comprehensive and Comfortable Search Menu





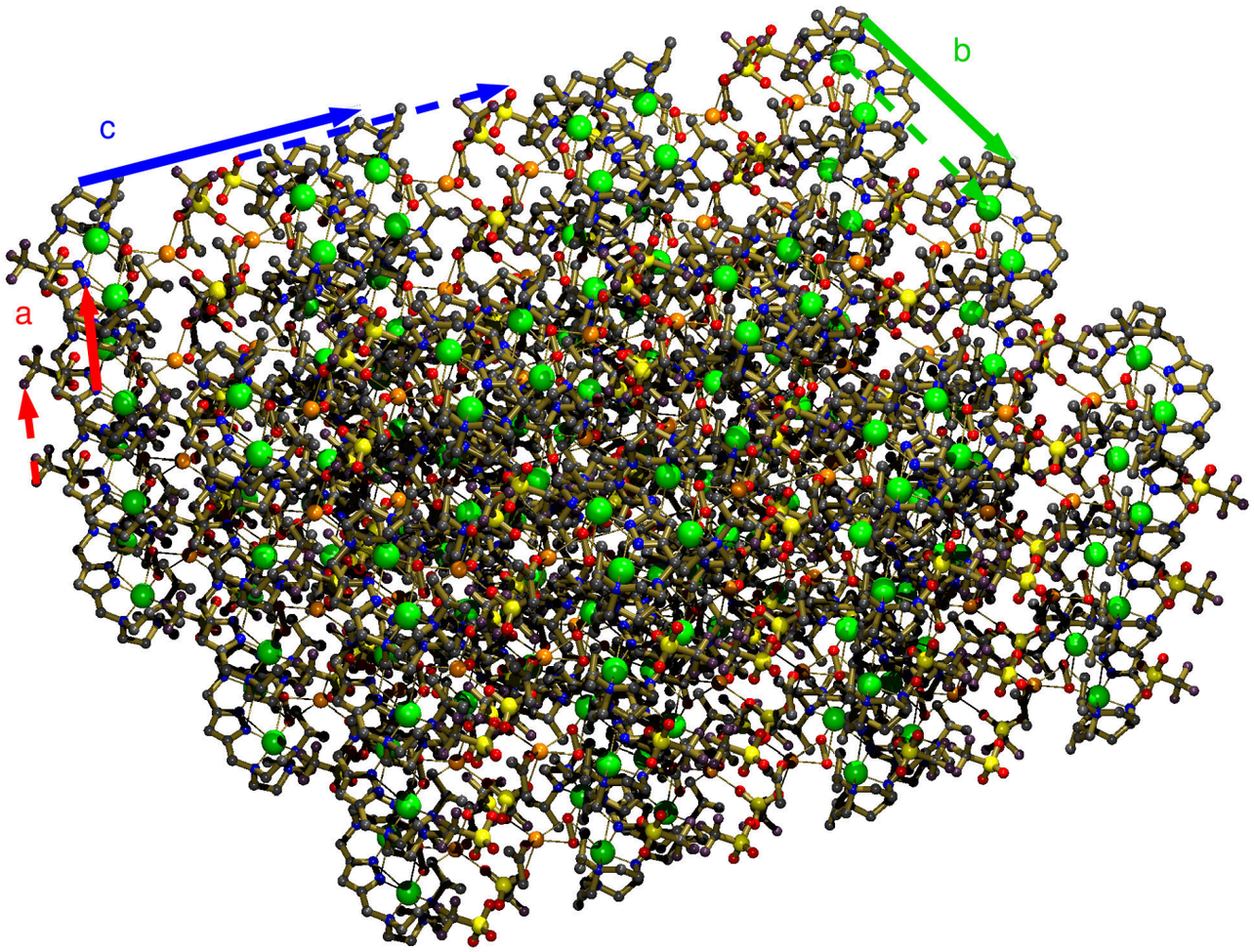
COD — Crystallography Open Database

- http://www.crystallography.net/search.html
- “Open-access collection of crystal structures of organic, inorganic, metal-organic compounds and minerals, excluding biopolymers ”



2 - What is a “Crystal”?

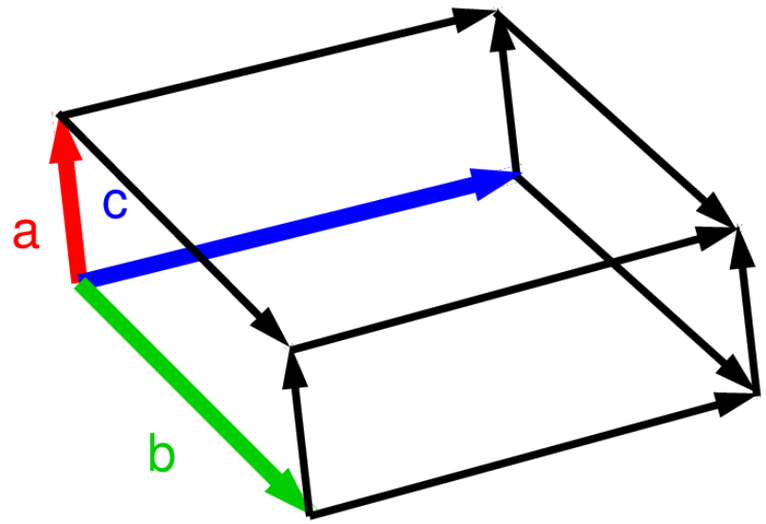
Periodic Packing and Crystal Lattice



- Crystal = Regular packing of one or more molecules
- Regularity expressed by “Unit Cell Vectors” $\vec{a}, \vec{b}, \vec{c}$
- Angles between vectors can $\neq 90°$

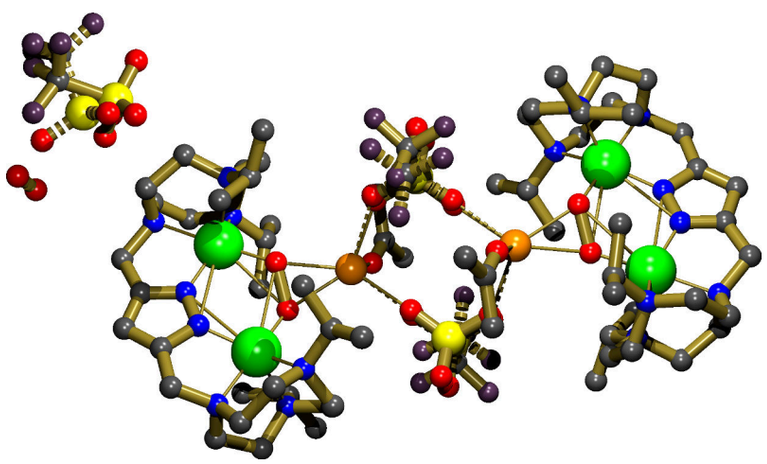
Unit Cell and Spacegroups

- Unit cell is a **concept** to describe the regularity of a crystal



- Unit cell can contain **more than one copy** of molecule
- Several copies lead to **Crystal Spacegroups** (total: 230)
- Both unit cell and spacegroup are experimental results (and thus can be wrong)
- Most frequent spacegroups for **organic structures**: $P2_1/c \approx 34.5\%$, $P\bar{1} \approx 24.7\%$ (Cambridge Structural Database CSD, 2017)
- Most frequent spacegroups for **protein structures**: $P2_12_12_1 \approx 22.5\%$, $P2_1 \approx 16.3\%$ (Protein Data Base PDB, 2017)

Quality of Crystal Structures



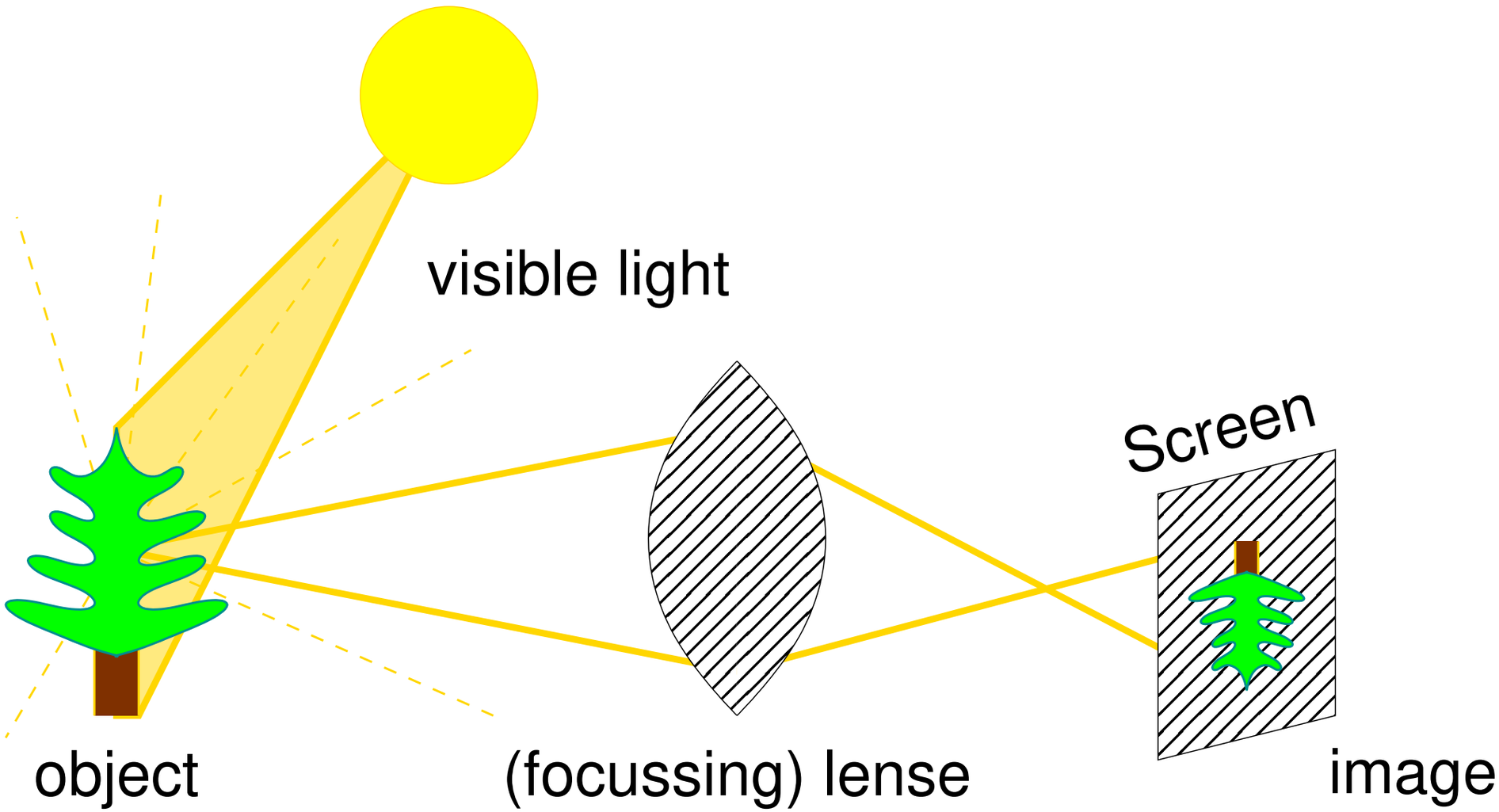
```
      Format: alert-number_ALERT_alert-type_alert-level text

 029_ALERT_3_A _diffrn_measured_fractionˍtheta_full value Low .      0.891 Note
 183_ALERT_1_A Missing _cell_measurement_reflns_used value ....     Please Do !
 184_ALERT_1_A Missing _cell_measurement_theta_min value ......     Please Do !
 185_ALERT_1_A Missing _cell_measurement_theta_max value ......     Please Do !
 224_ALERT_1_A Ueq(Rep) and Ueq(Calc) Differ by -0.058 Ang**2 .       S_3 Check
 224_ALERT_1_A Ueq(Rep) and Ueq(Calc) Differ by -0.052 Ang**2 .       F3_3 Check
 224_ALERT_1_A Ueq(Rep) and Ueq(Calc) Differ by -0.060 Ang**2 .       O1_3 Check
 224_ALERT_1_A Ueq(Rep) and Ueq(Calc) Differ by -0.021 Ang**2 .       O2_3 Check
 224_ALERT_1_A Ueq(Rep) and Ueq(Calc) Differ by -0.028 Ang**2 .       O3_3 Check
 224_ALERT_1_A Ueq(Rep) and Ueq(Calc) Differ by -0.034 Ang**2 .       C_3 Check
 224_ALERT_1_A Ueq(Rep) and Ueq(Calc) Differ by -0.119 Ang**2 .       O2_4 Check
 701_ALERT_1_A Bond    Calc    2.312(5), Rep   2.353(19), Dev..      8.20 Sigma
          NAO    -O2_3          1.555   1.555 ..........      #    6 Check
 701_ALERT_1_A Bond    Calc    2.265(7), Rep   2.45(3), Dev..      26.43 Sigma
          NAO    -O3_3          1.555   2.766 ..........      #    7 Check
 702_ALERT_1_A Angle   Calc    108.7(2), Rep   109.4(9), Dev..      3.50 Sigma
          OB_2   -NAO    -O2_3          1.555   1.555   1.555 #   11 Check
 702_ALERT_1_A Angle   Calc    140.0(2), Rep   143.3(8), Dev..      16.50 Sigma
          OA_2   -NAO    -O2_3          1.555   1.555   1.555 #   12 Check
 702_ALERT_1_A Angle   Calc    98.4(3), Rep    92.1(7), Dev..      21.00 Sigma
          O_5    -NAO    -O2_3          1.555   1.555   1.555 #   13 Check
 702_ALERT_1_A Angle   Calc    137.6(2), Rep   130.5(7), Dev..      35.50 Sigma
          OB_2   -NAO    -O3_3          1.555   1.555   2.766 #   15 Check
 702_ALERT_1_A Angle   Calc    102.3(2), Rep   97.4(8), Dev..      24.50 Sigma
          OA_2   -NAO    -O3_3          1.555   1.555   2.766 #   16 Check
 702_ALERT_1_A Angle   Calc    91.9(3), Rep    101.5(7), Dev..      32.00 Sigma
          O_5    -NAO    -O3_3          1.555   1.555   2.766 #   18 Check
 702_ALERT_1_A Angle   Calc    102.6(3), Rep   104.5(7), Dev..      6.33 Sigma
          O1_3   -S_3    -C_3           1.555   1.555   1.555 #  133 Check
 702_ALERT_1_A Angle   Calc    169.8(6), Rep   159.9(14), Dev..      16.50 Sigma
          S_3    -O2_3   -NAO           1.555   1.555   1.555 #  134 Check
 702_ALERT_1_A Angle   Calc    103.4(3), Rep   102.1(5), Dev..      4.33 Sigma
          O1_4   -S_4    -C_4           1.555   1.555   1.555 #  158 Check
 702_ALERT_1_A Angle   Calc    112.9(4), Rep   110.4(6), Dev..      6.25 Sigma
          F3_4   -C_4    -S_4           1.555   1.555   1.555 #  164 Check
 925_ALERT_1_A The Reported and Calculated Rho(max) Differ by .      5.35 eA-3

INSTRUCTION INPUT via KEYBOARD or LEFT-MOUSE-CLICKS (HELP with RIGHT CLICKS)   Exit

>> Continue (Y/N(Y))                                                  MenuActive
```

- Figures contain no information about data quality
- “Quick” quality indicators: resolution $d_{\text{min}} < 0.84$Å, $R1 \leq 5\%$, $\approx 100\%$ complete
- (In-)Organic Compounds: CheckCIF Report (A.L.Spek, Acta Cryst. 2009, D65, 148-155.)
- Should have neither A– nor B–alerts

Bear in mind: Crystallisation is a **purification** method.

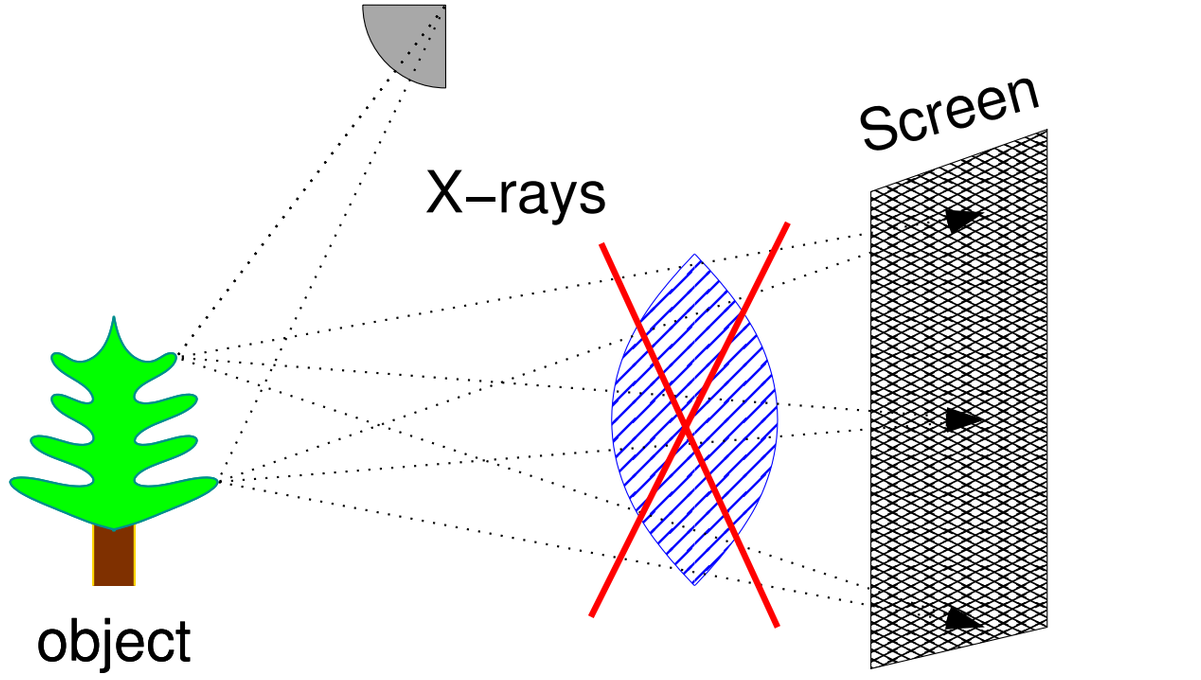
3 - Why Crystals?

Imaging with Visible Light (Light Microscope)



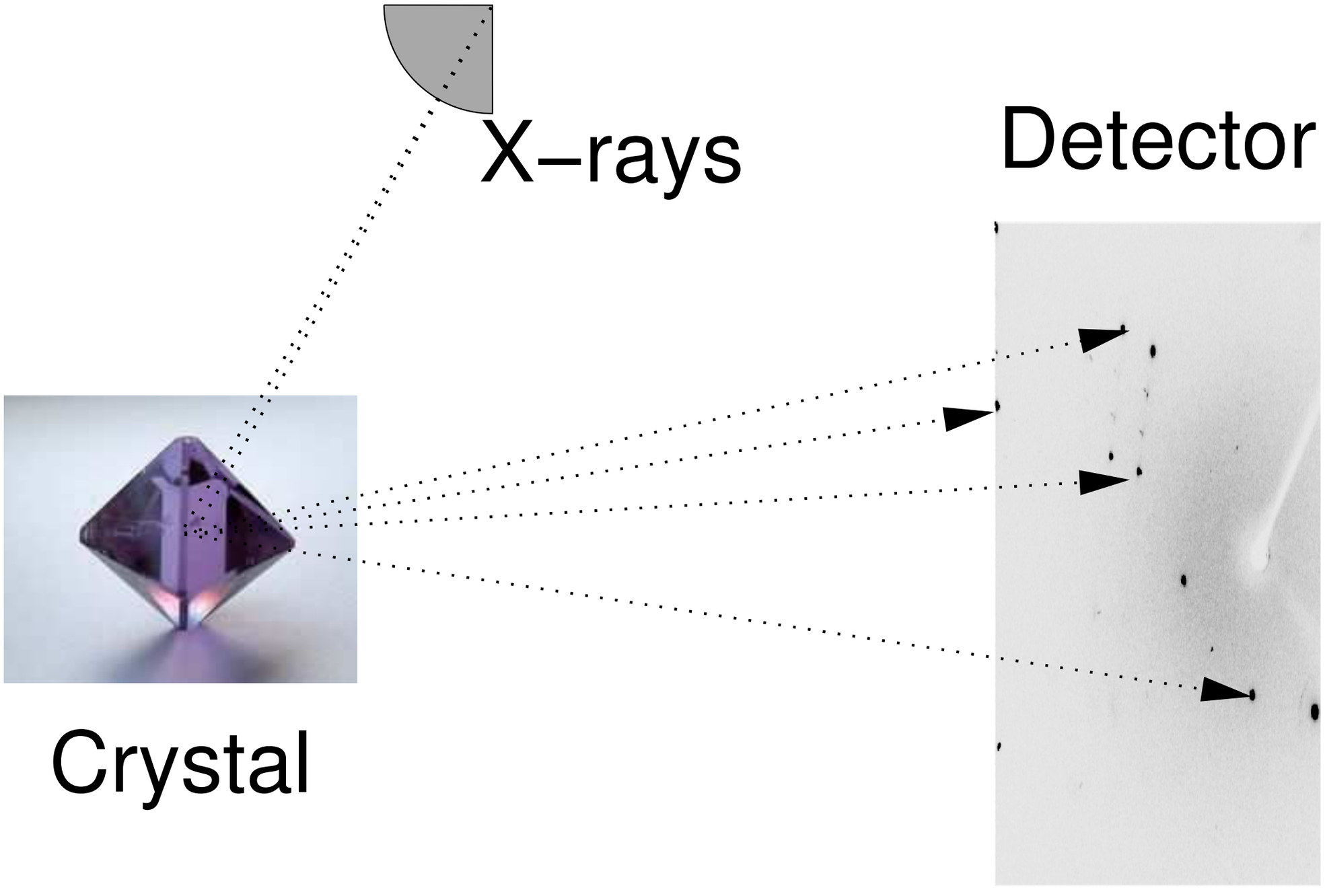
No Imaging with X–rays

- Typically bond lengths are 1–2Å
- Abbé principle to resolve two adjacent points: $\lambda \leq 2d_{\text{min}} \sin \alpha$
- Typical X–ray wavelength: 0.5–2Å



no lense = no image, only "blur"

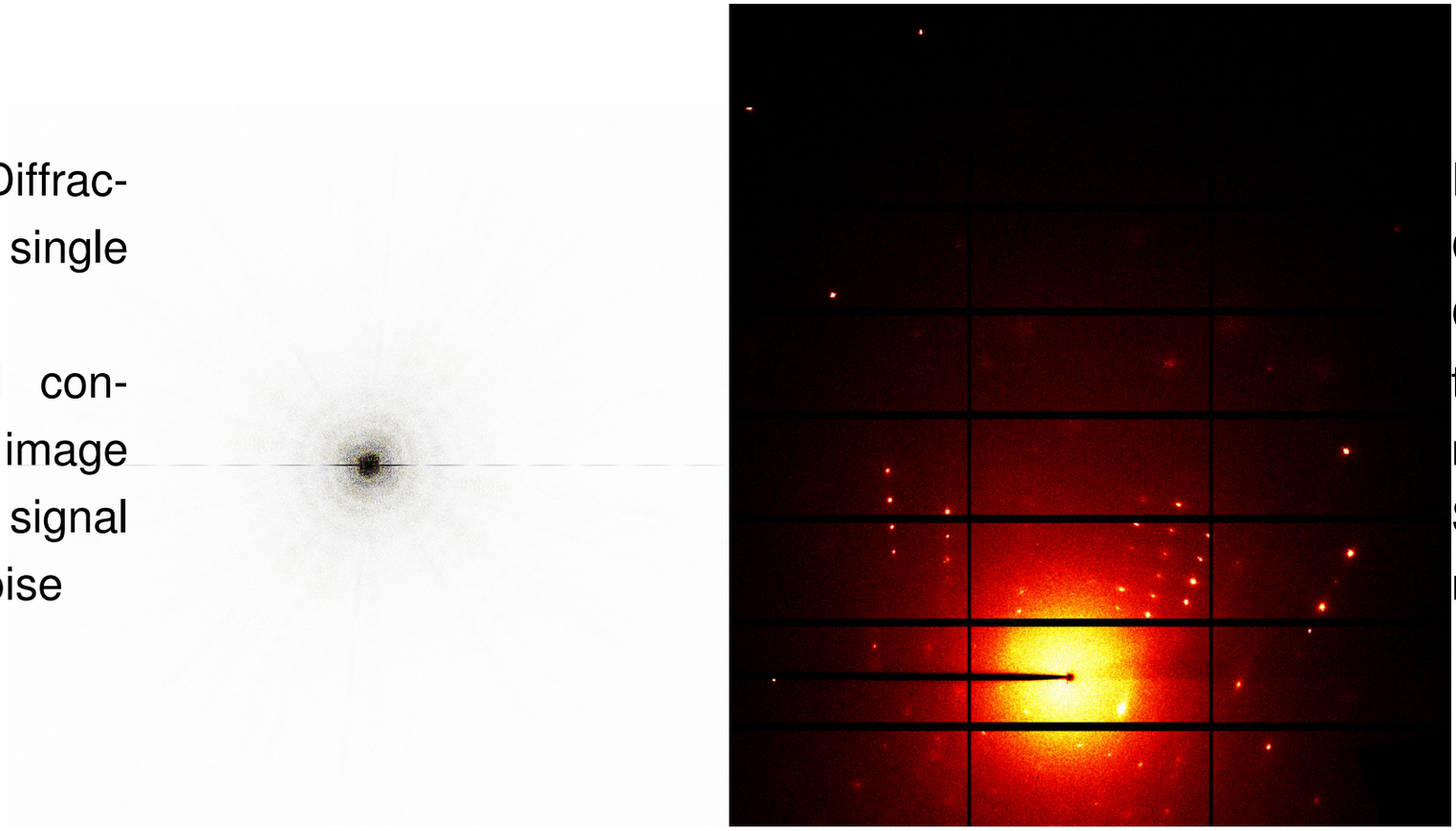
X–ray Diffraction from Crystals



Crystals are Signal Amplifiers

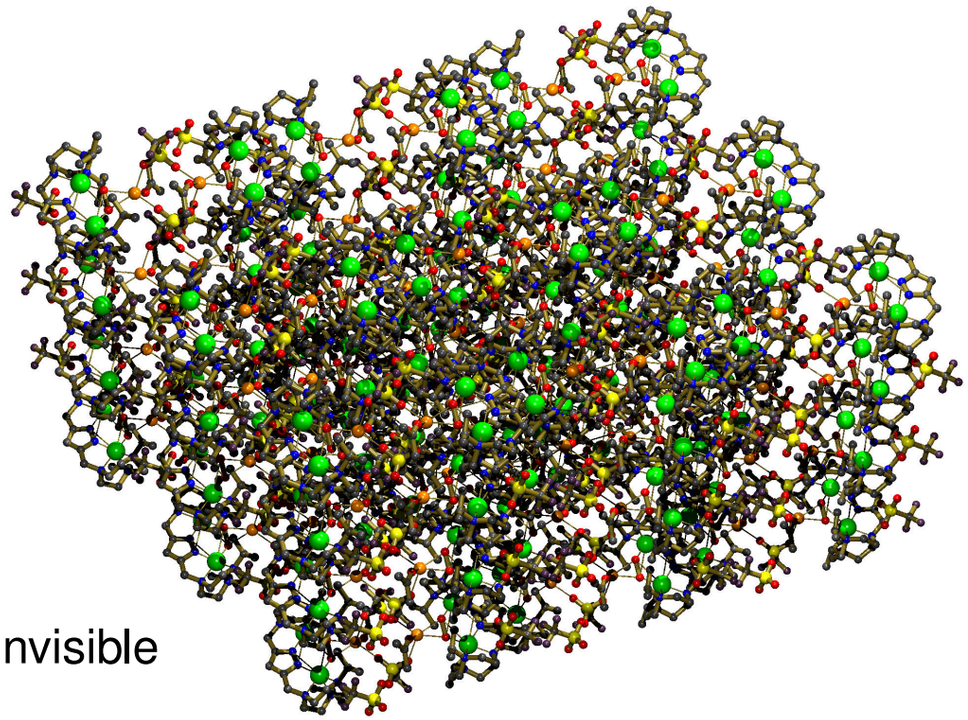
- Single molecules are too small for visualisation:
  - Short wavelength required ≈ 1Å
  - X–rays: no lenses, Electrons: distructive
- Crystals can be used for diffraction instead of direct visualisation

Simulated Diffraction from single molecule:
Each pixel contributes to image formation: signal buried in noise

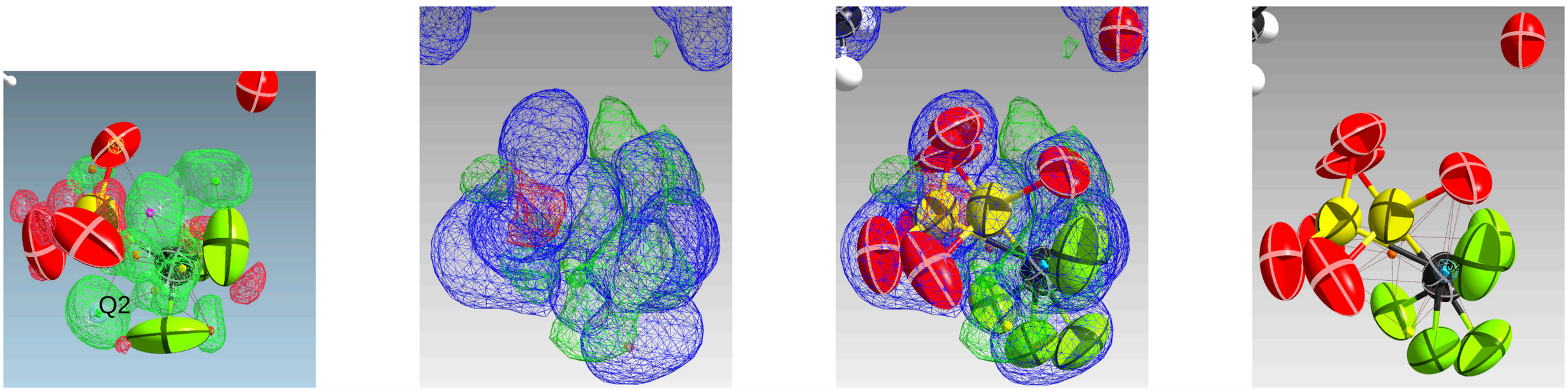


Diffraction from Crystal:
Crystal concentrates signal in reflections (spots)
Signal well above noise

Crystals Structures: Average over all unit cells



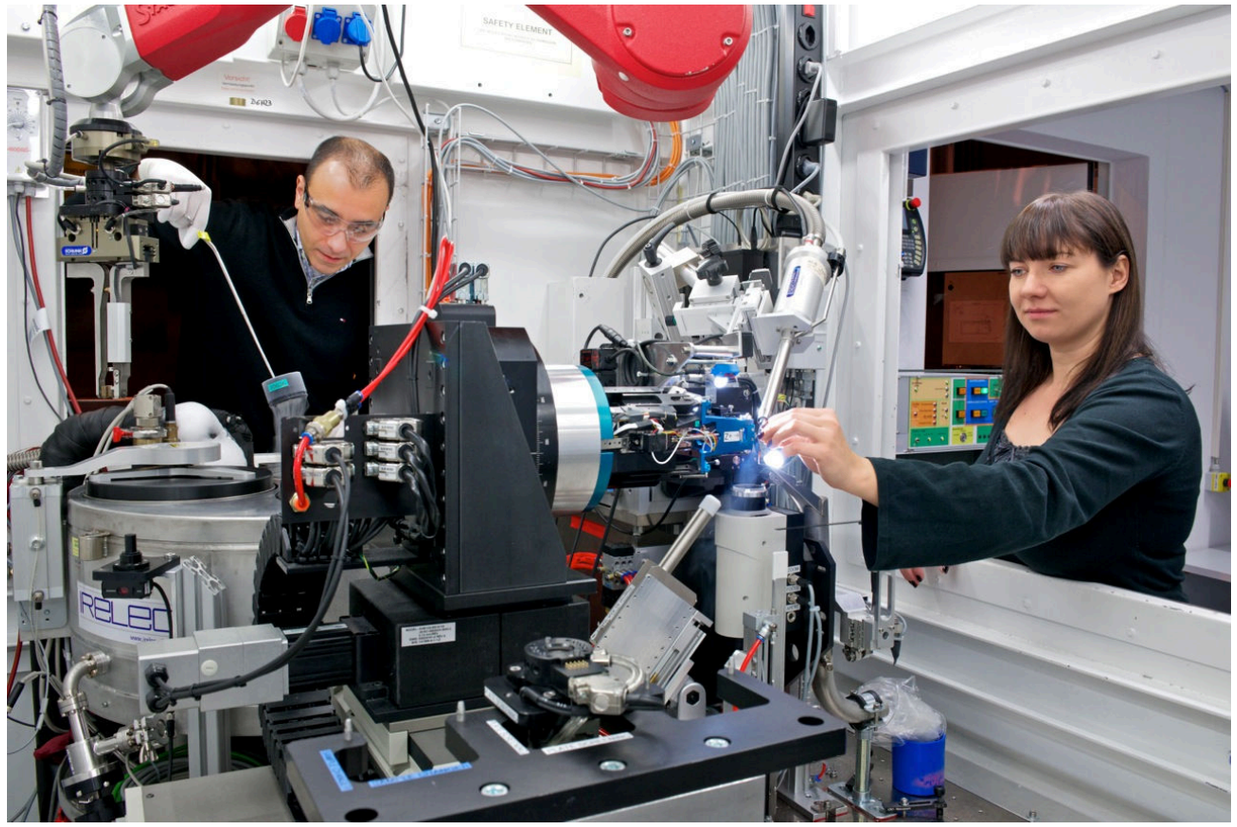
- Crystals structure = average of all unit cells in crystal
- Sometimes, disorder can be modelled
- When disorder becomes too irregular, features become invisible
- (Platons “SQUEEZE” command blinds out disordered regions)



4 - Crystal Structure Determination at PSI

Instruments / X–ray Sources

- Three beamlines, PX-I, PX-II, PX-III
- Optimised for protein crystallography
- Suitable for Organic Compounds



Photograph courtesy Paul Scherrer Institute/Markus Fischer

Applications of Crystal Structures

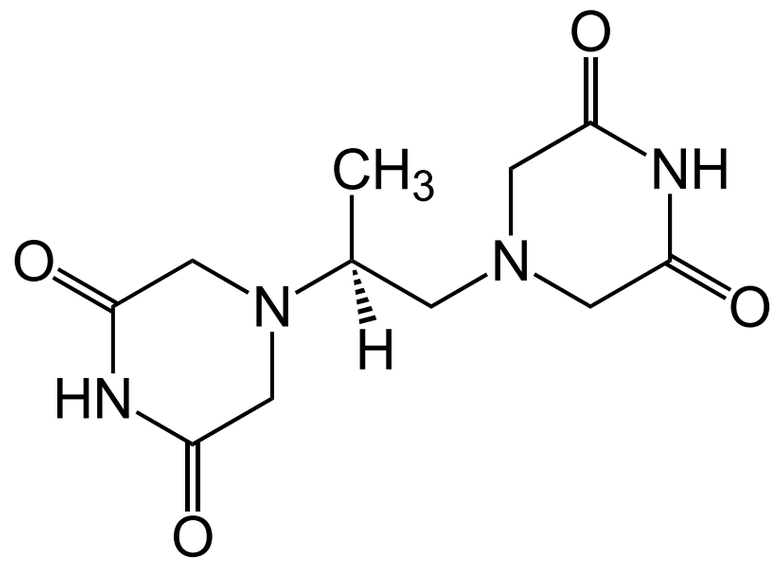
Pretty pictures do not make a reason for Crystallography

1. Confirmation of Synthesis Products
2. Sole method to determine chirality
3. Starting point for Molecular Dynamics (design interaction drug $\leftrightarrow$ target)
4. Basis for new drugs (insulin cocktail)

Applications of Crystal Structures

The therapeutically non-active isomer in a racemate should be regarded as an impurity

E. J. Ariëns, Eur. J. Clin. Pharmacol. (1984), 26, 663–668



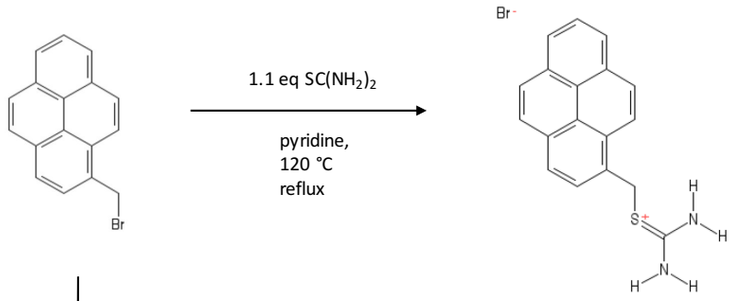
Razemic (dex)razoxane, a cardioprotective agent

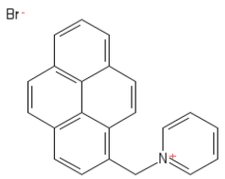
By Jü, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=32957899

X–ray crystallography can determine enantiomeric purity

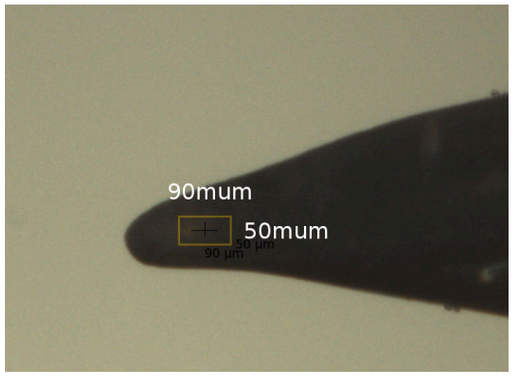
Synthesis Control

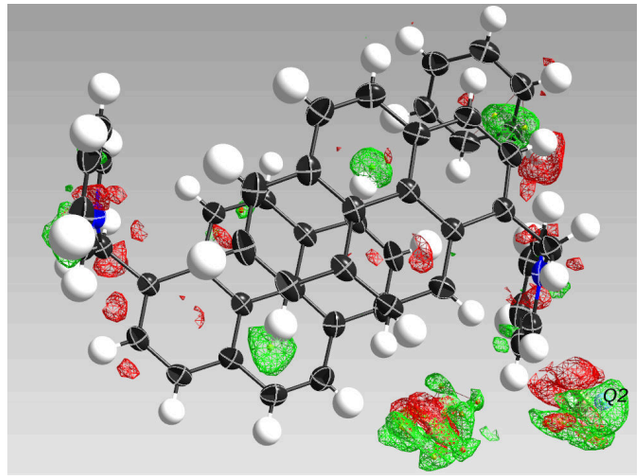
Aim: Methylpyren-diaminomethylen-sulfanium-bromid

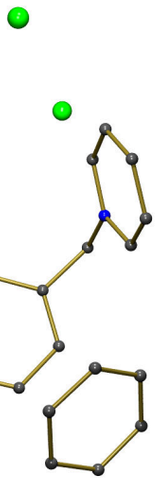




Methylpyren-pyridinium-bromide







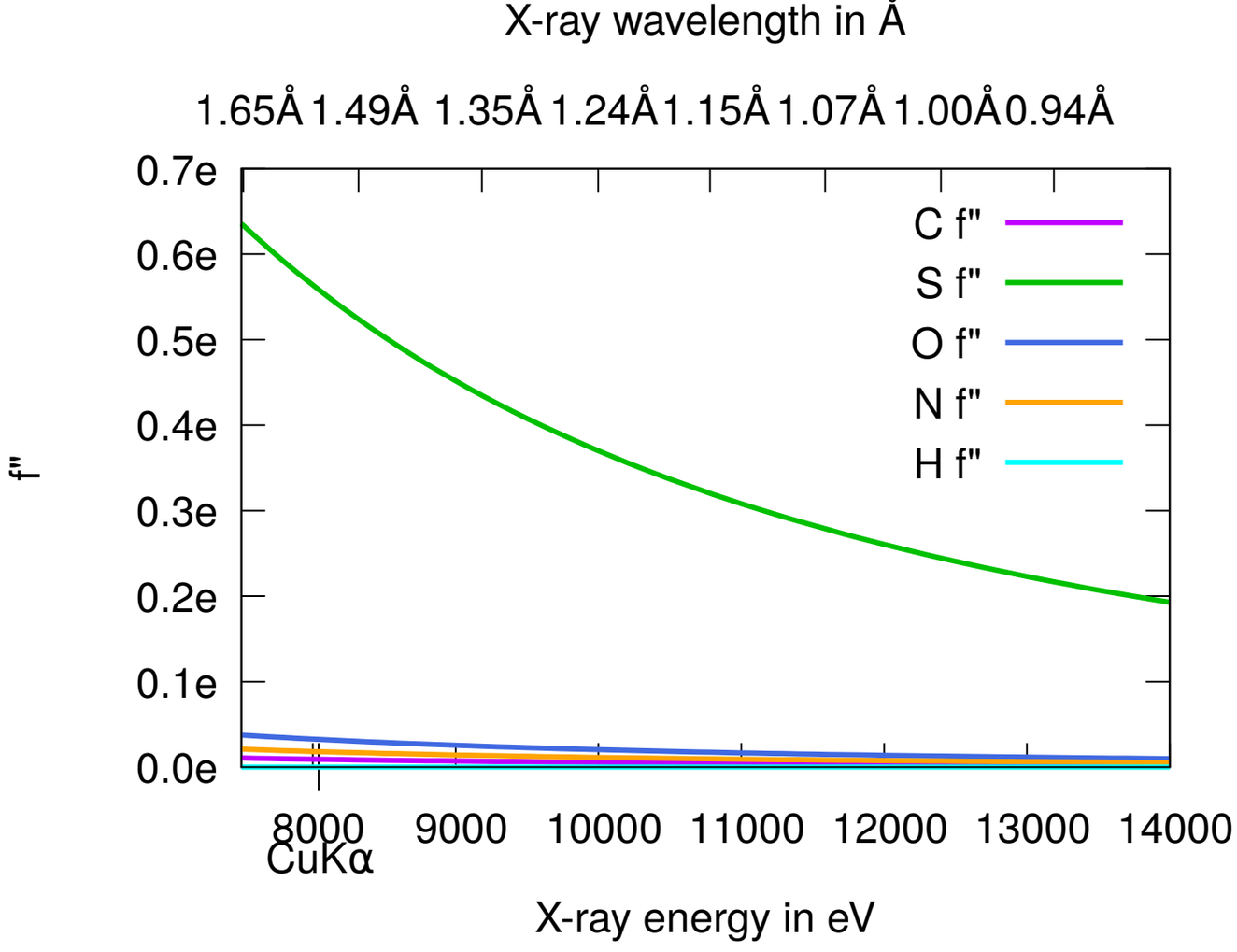
C. Borsa & J. Wennmacher, LBR; da

Very small, poor looking crystals sufficient for reliable structure solution.

R1 = 4.3%, Highest peak: 0.75e^-, Deepest hole: -0.59e^-

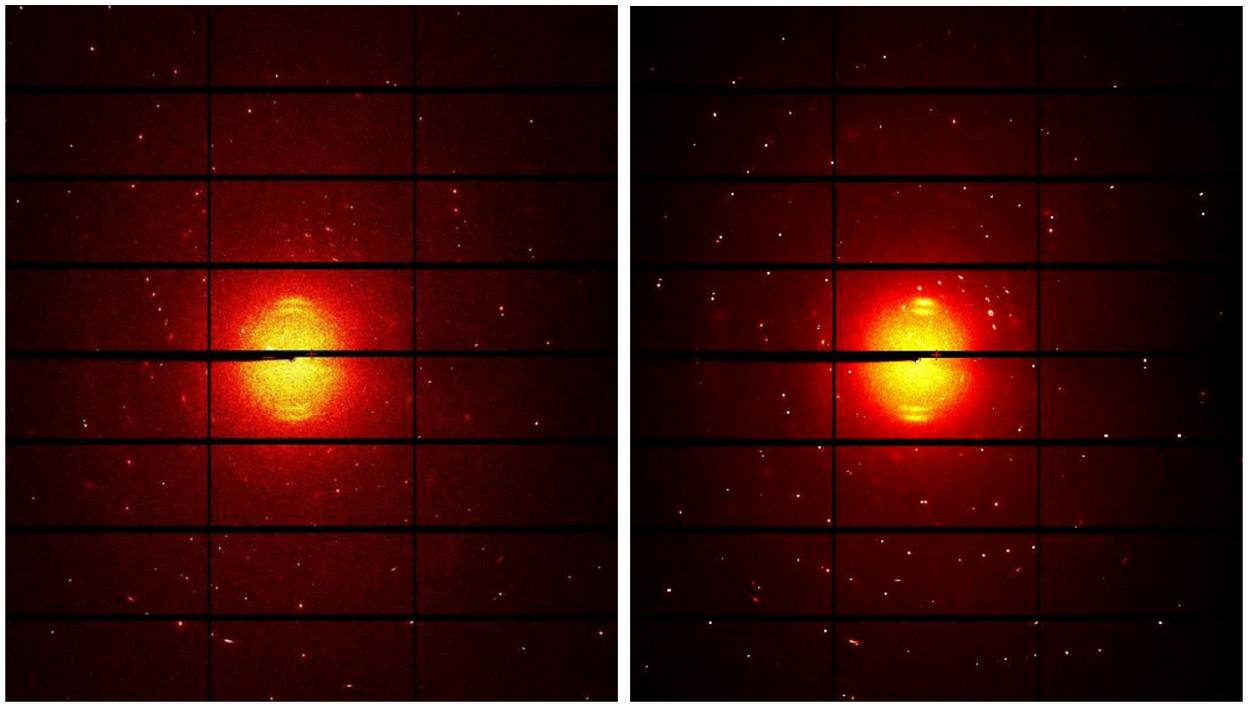
Chirality with only Light Atoms

http://skuld.bmsc.washington.edu/scatter/AS$_f$orm.html



- High resolution data requires short wavelength < 0.8Å
- Anomalous Signal extremely weak for light atoms (CHNO)
- Chirality determined from Anomalous Signal (Absorption)
- Challenge for Synchotron beamlines tunes for protein samples

Data Quality at PX-III



- Organic Compound collected at PX-III, $\lambda = 0.72$Å
- 100% complete data at 0.9Å, 82% data at 0.86Å
- Multiple, > 2, lattices
- Good model statistics, R1=5.1%; chirality determined: Flack x = 0.08(10) Flack Parsons = 0.07(14)

5 - Crystallography with Electrons (instead of X–rays)

Electrons *vs.* X–rays

- X–rays: weak interaction, nearly all do not interact with crystal
- Electrons: strong interaction, short penetration depth
- X–rays minimum crystal size $\approx 5\mu\ m$
- Electrons **maximum** crystal size $\approx 1\mu\ m$

Electrons as Radiation Source

- Crystallography requires an incoming wave (so far: X–rays)
- Electrons are waves, *cf.* de Broglie wavelength: $\lambda = \frac{h}{m_e v_e}$
- Typical energies and wavelengths: 100 keV = 0.05016Å, 200keV = 0.02508Å
- Wavelength much shorter ($\times 1/40$) than X–ray (penetration depth)

An Electron Diffraction Instrument

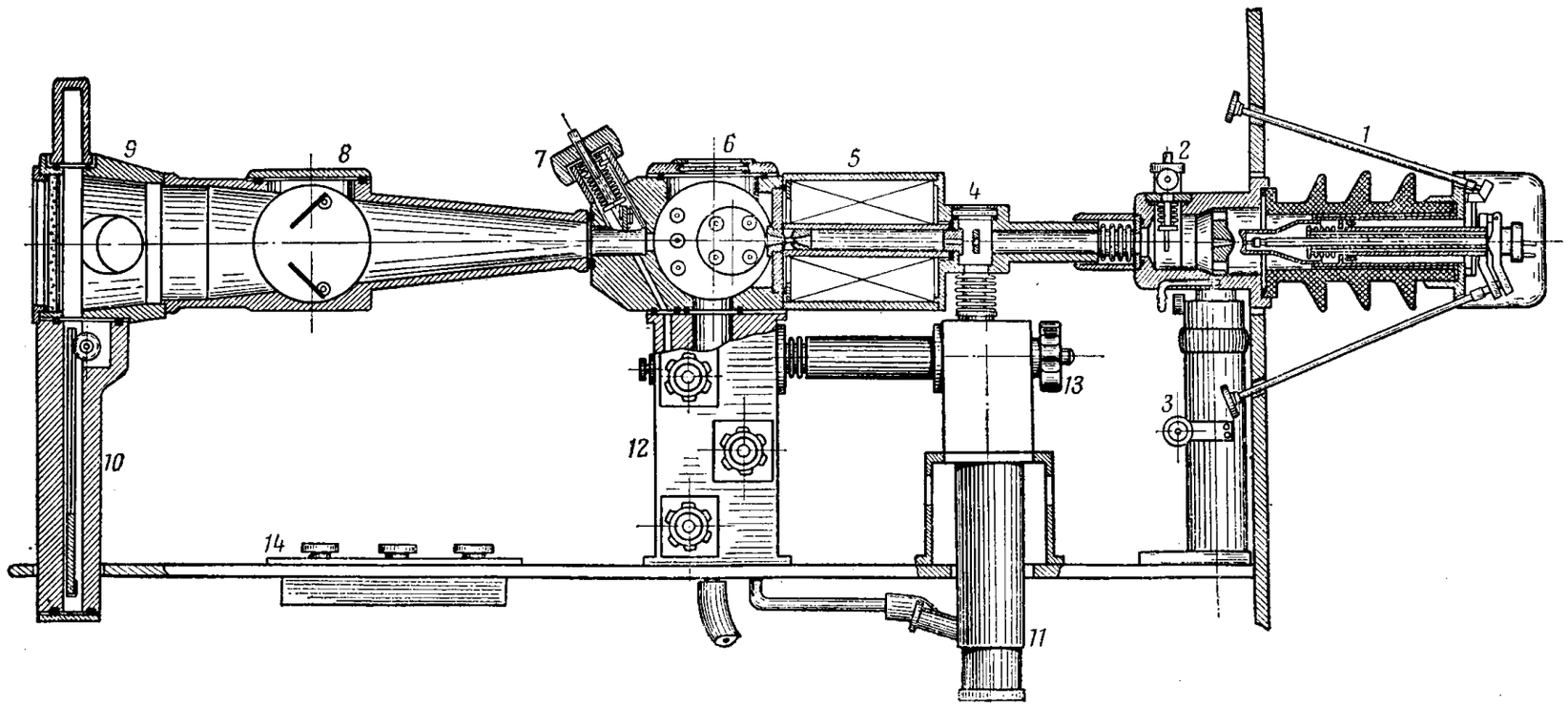
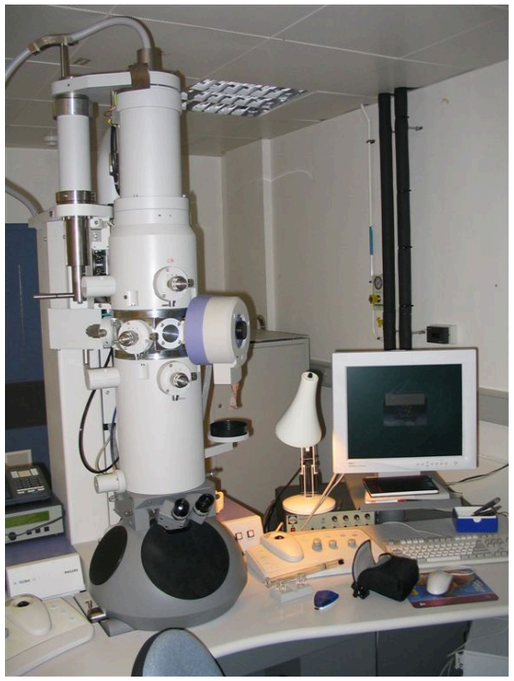
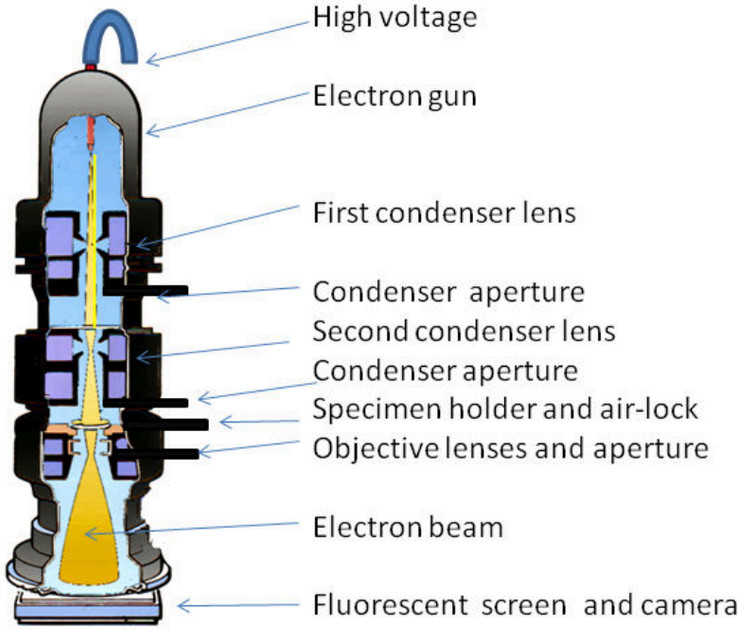


FIG. 123. Electron diffraction camera of the Institute of Crystallography of the Academy of Sciences, U.S.S.R. 1—electron gun, 2—anode, 3—gun support, 4—intermediate chamber, 5—magnetic lens, 6—central chamber, 7—intermediate valve, 8—diffraction section, 9—upper part of photographic chamber, 10—lower part of photographic chamber, 11—high-vacuum pump, 12—fore-vacuum valve block, 13—high-vacuum valve, 14—electrical control panel.

B. K. Vainshtein, “Structure Analysis by Electron Diffraction”, Pergamon Press, 1964

Electron Microscopes





Transmission Electron Microscope

Left: By David J Morgan from Cambridge, UK (Tecnai 12 Electron Microscope), *via* Wikimedia Commons

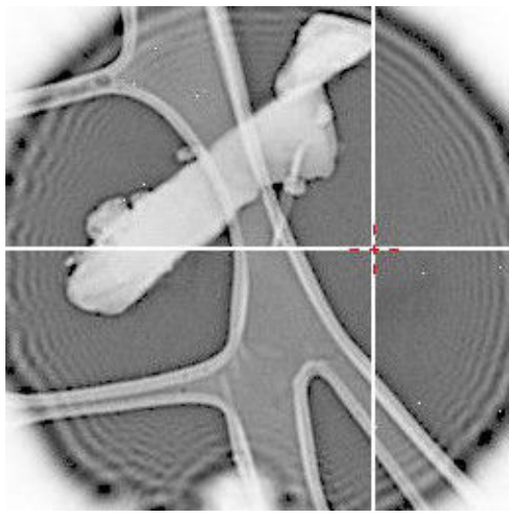
Right: By Dr Graham Beards, *via* Wikimedia Commons

Some Milestones in Electron Crystallography

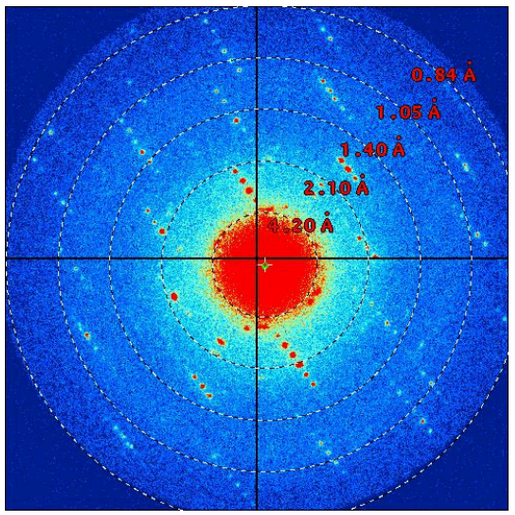
- Pioneers: ZG Pinsker, BK Vainshtein (1940s +; 1990s)
- D. Dorset (1995: Organic Compounds, Structure solution with direct methods)
- U. Kolb (recording of **3D** diffraction patterns, ADT, 1997+)
- X. Zou, S. Hovmöller (recording of 3D diffraction patterns with beam precession, RED, 2008+)

Pharmaceutical I: Curvulone antibiotic / antifungal

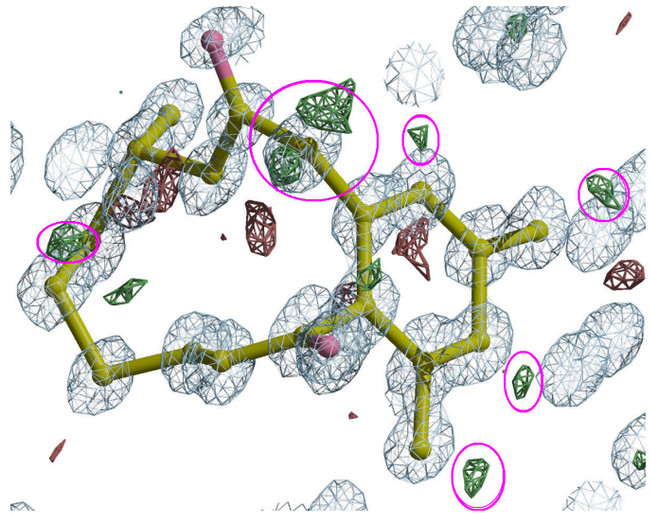
CCDC: IRELOH, Dai et al., Eur. J. Org. Chem (2010), 6928-6937, Sample courtesy Novartis



- Field of view: $3\mu m$
- Crystal: $1.6\mu m \times 400nm$



- $d_{\text{min}} < 0.8$Å
- $P2_12_12_1$: 85% completeness with 3 crystals
- a=8.06Å b=10.00Å c=17.73Å



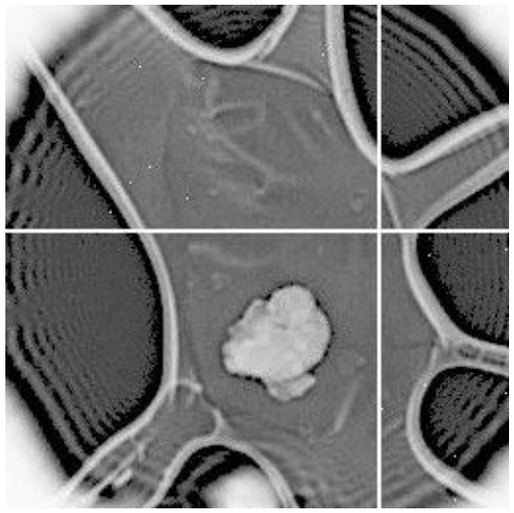
- 1334 reflections, 195 parameters, 156 restraints (RIGU)
- $R1 = 15.5\%, R_{\text{complete}} = 18.5\%$

Despite the poor conventional quality indicators (R-values), the data quality is good enough to show hydrogen positions.

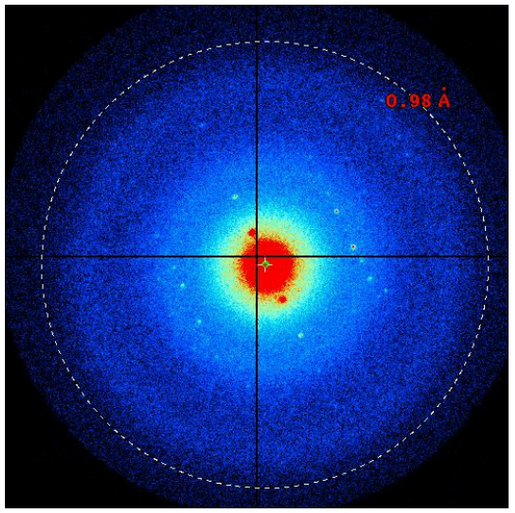
Pharmaceutical II: Epicorazine A

CCDC: EPICZA, Dai et al., Eur. J. Org. Chem (2010), 6928-6937, Sample courtesy Novartis

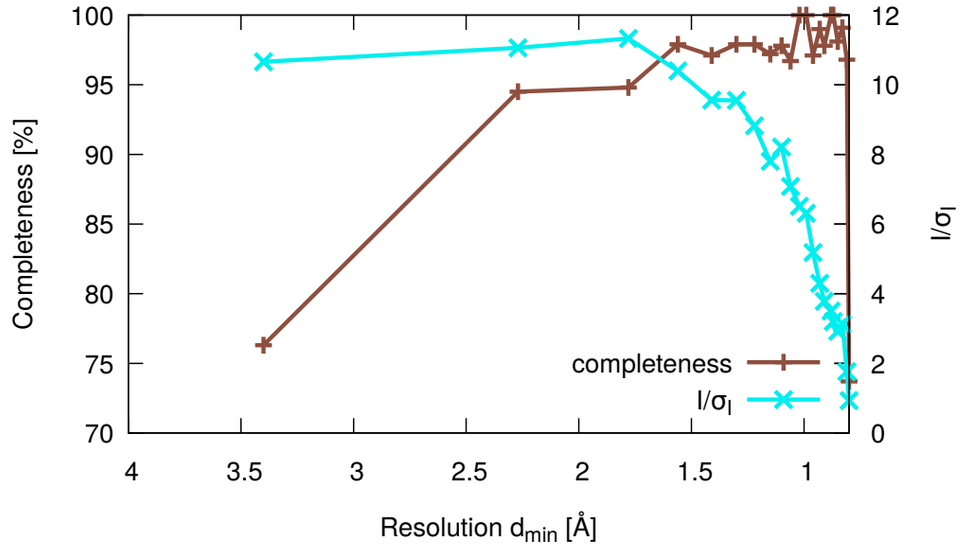
Data quality: recognition of atom types, C *vs.* O *vs.* N *etc.* (CCDC: EPICZA)



- Field of view: $3\mu m$
- Crystal: $400nm$ diameter

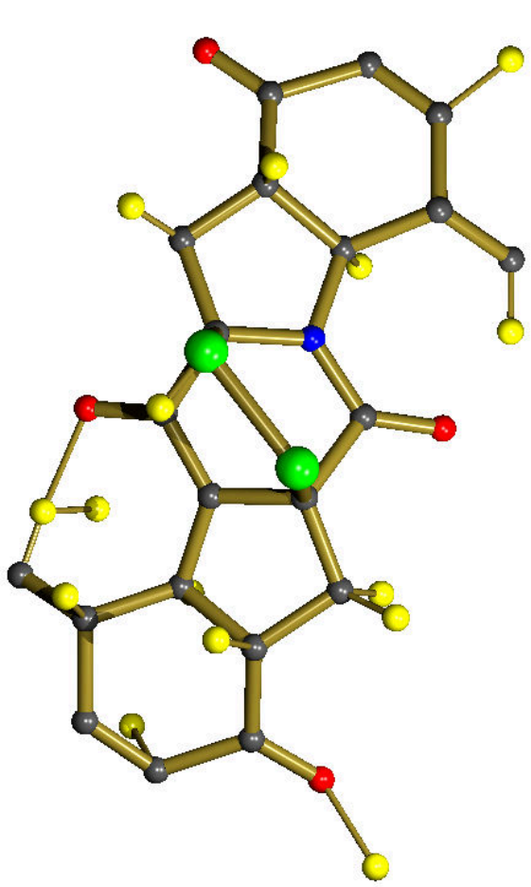


- $d_{\text{min}} = 0.80$Å
- a=10.65Å, b=12.16Å, c=12.83Å
- $P2_12_12_1$: completeness with 4 crystals: 97%

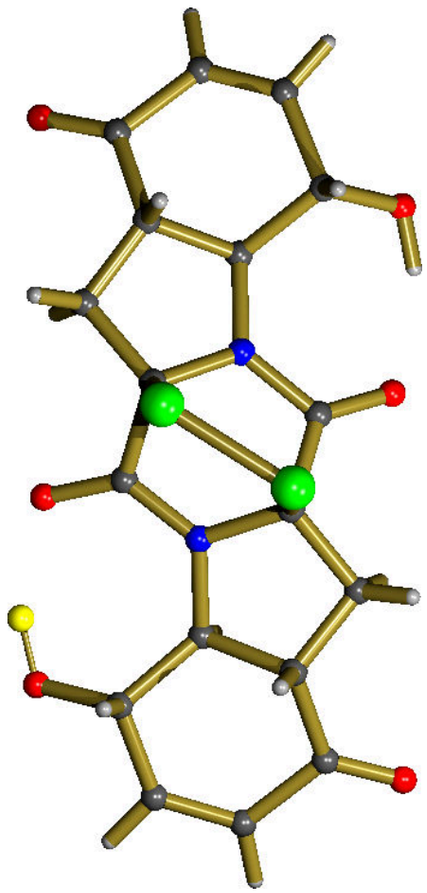


- 3316 refl., 256 param., 267 restraints (RIGU)
- model fit to data: $R1 = 18.8\%, R_{\text{complete}} = 23.2\%$

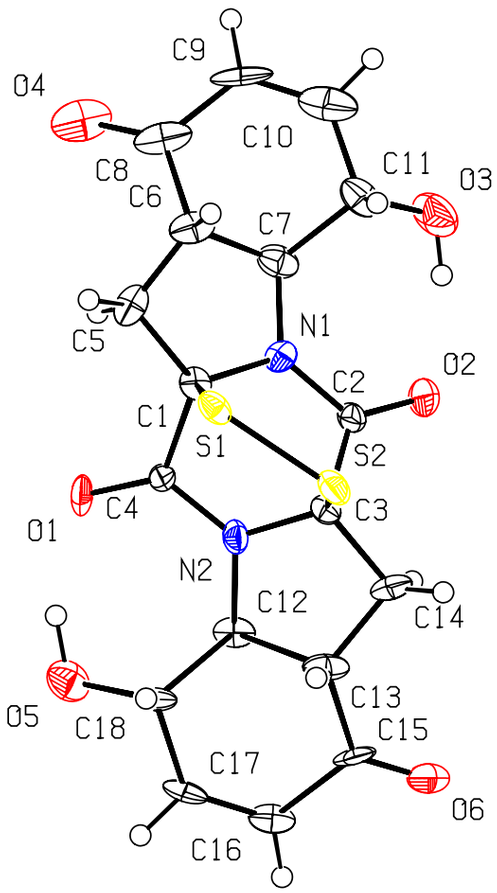
Pharmaceutical II (EPICZA): Structure Solution Process



Direct methods reveal H atoms
=data quality



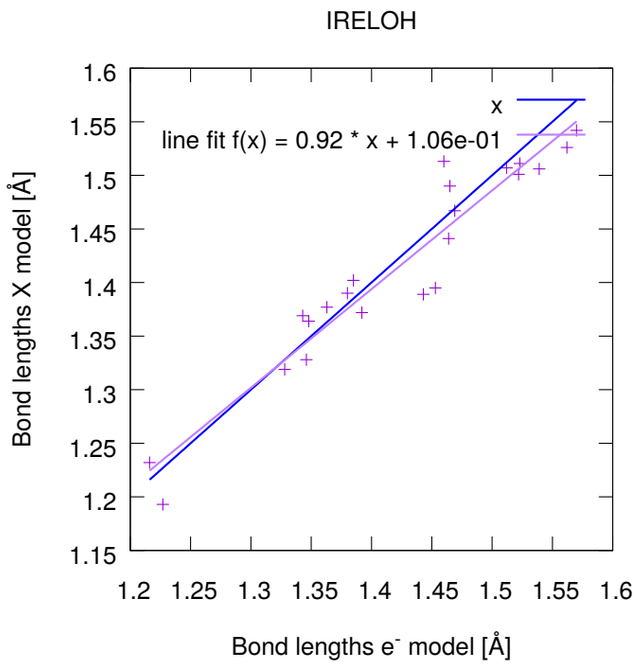
HFIX: all except 1 H
=model quality

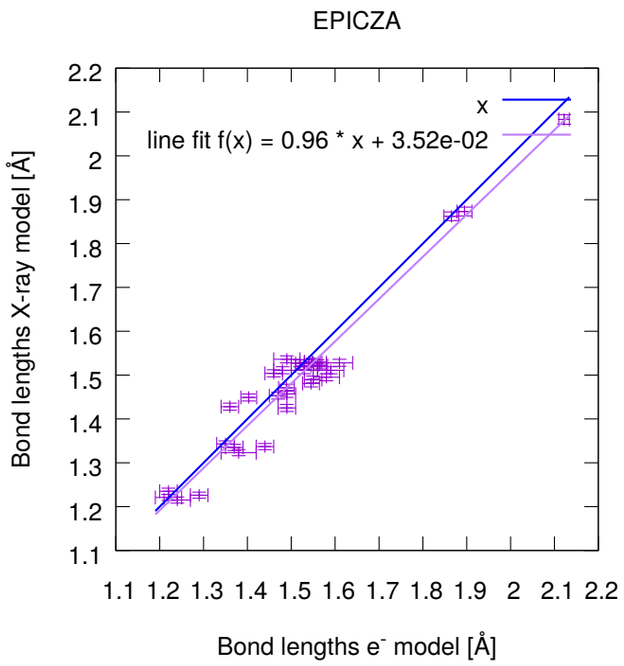


Final Structure

Are Structures from Electron Diffraction Reliable?

- Structures can be solved with X–ray knowledge and methods (D. Dorset, 1995)
- Radiation damage present, but not (always) limiting
- Kinematic approximation sufficient for high quality structures
- Quality indicators **very poor**
- Structure quality acceptable





6 - Acknowledgements

- George Sheldrick (Georg–August–University)
- Trixie Wagner (Novartis)
- Staff at SLS PX–beamlines
- Henning Stahlberg (C–CINA)
- Abrahams group (C–CINA + PSI)