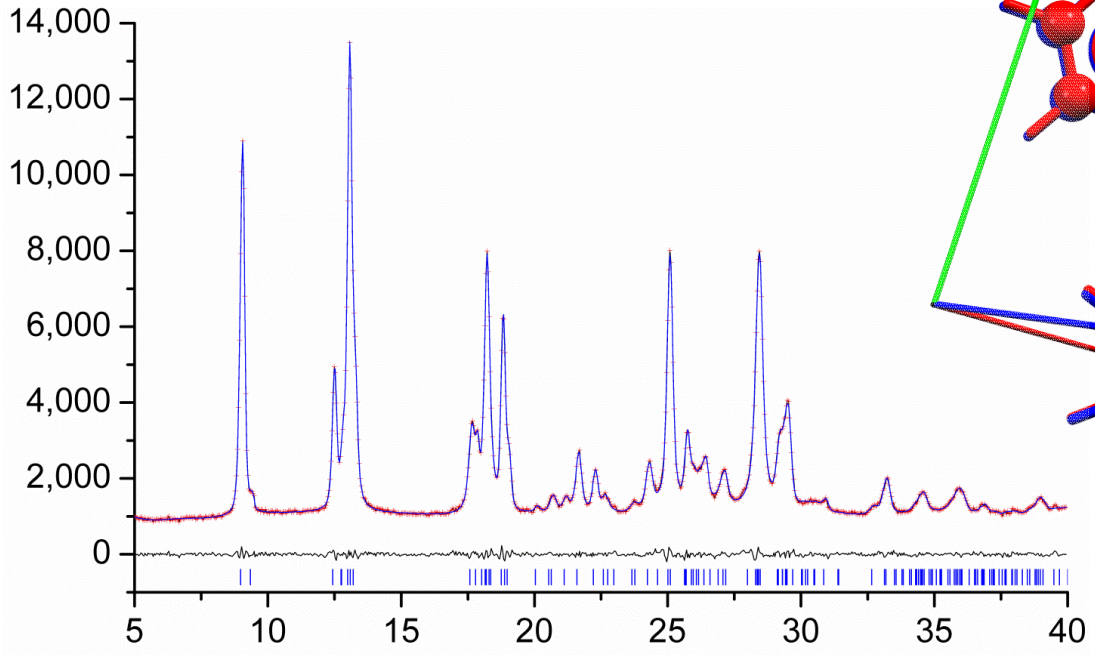
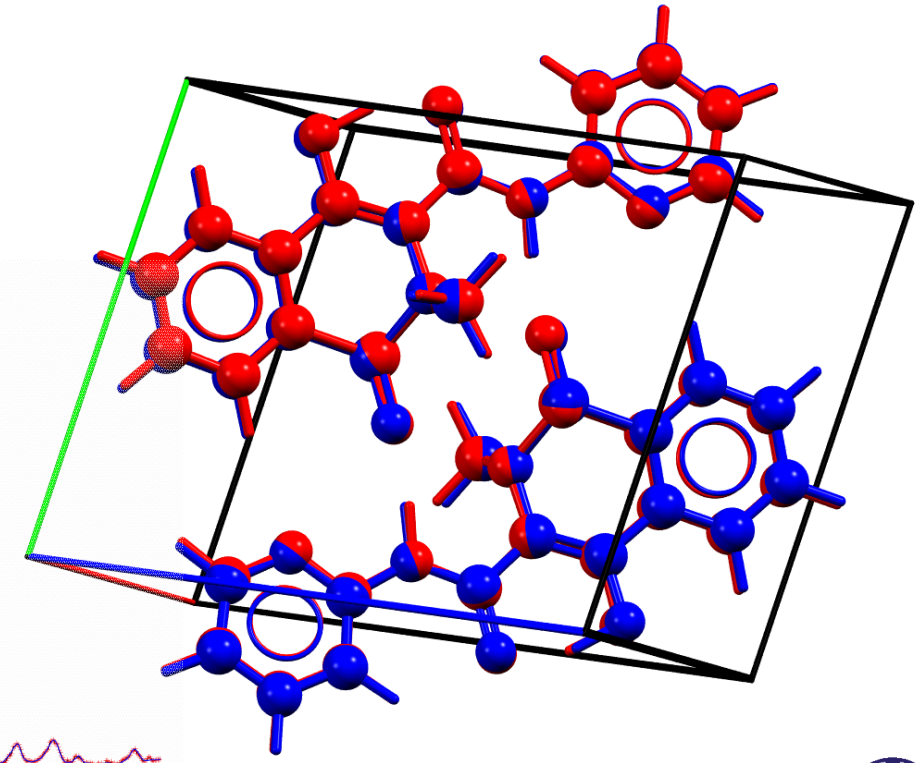
Verification of Crystal Structures Determined from XRPD—the Role of DFT-D

Jacco van de Streek
jacco.vandestreek@sund.ku.dk





This document was presented at PPXRD - Pharmaceutical Powder X-ray Diffraction Symposium

Sponsored by The International Centre for Diffraction Data

PPXRD Website – www.icdd.com/ppxrd

ICDD Website - www.icdd.com

Outline

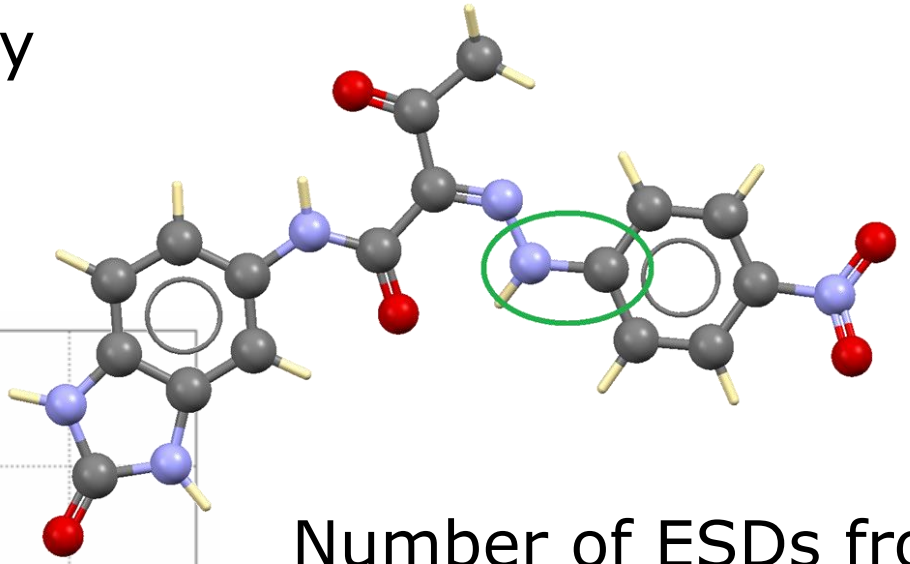
1. Reliability: is a crystal structure from XRPD correct or not?

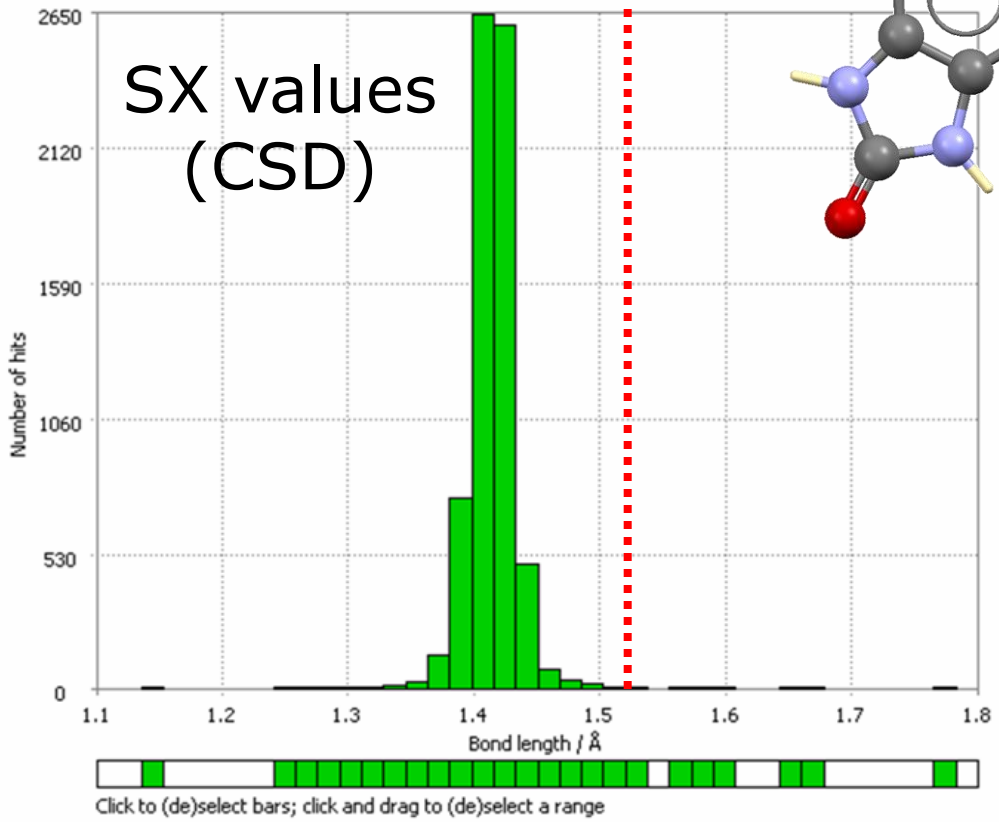
2. Precision: for a correct crystal structure from XRPD, how precise are *e.g*. the bond lengths?

3. Future directions
- Temperature
- Hydrogen atoms
- ss-NMR
- Space group validation

Mogul z-scores

= value in query



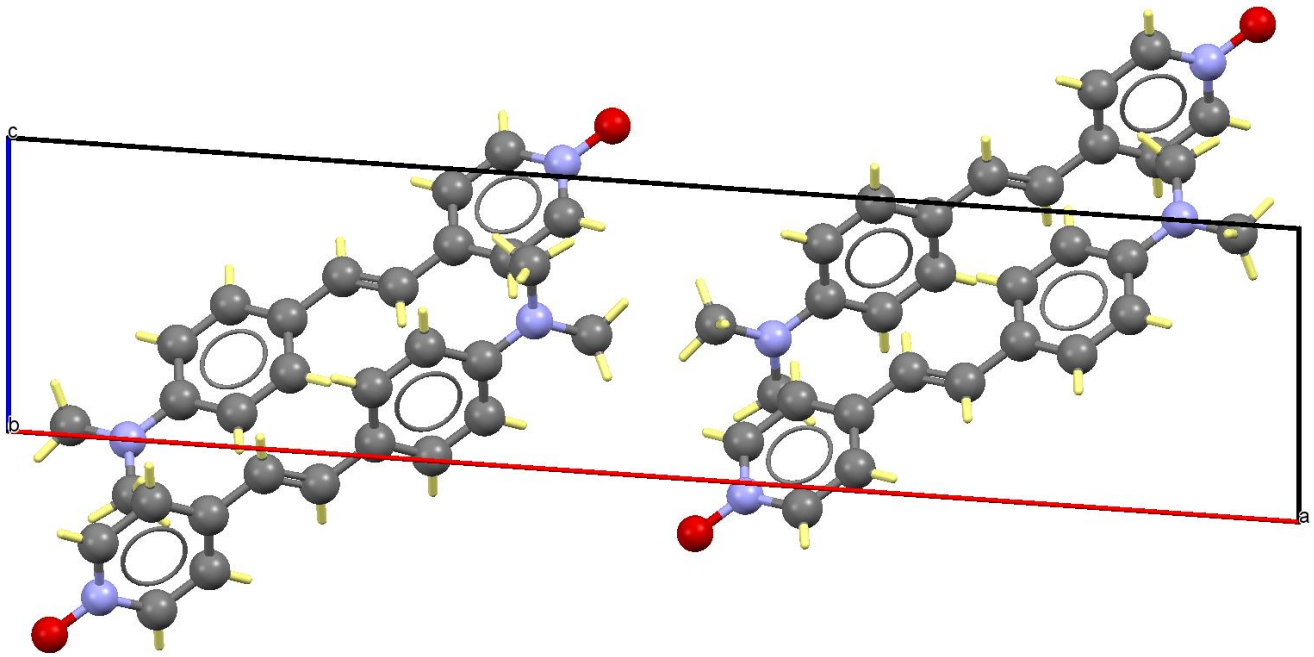


Number of ESDs from mean = *z*-score

Each bond has a *z*-score
Each angle has a *z*-score

“Precise” *vs* “Reliable”

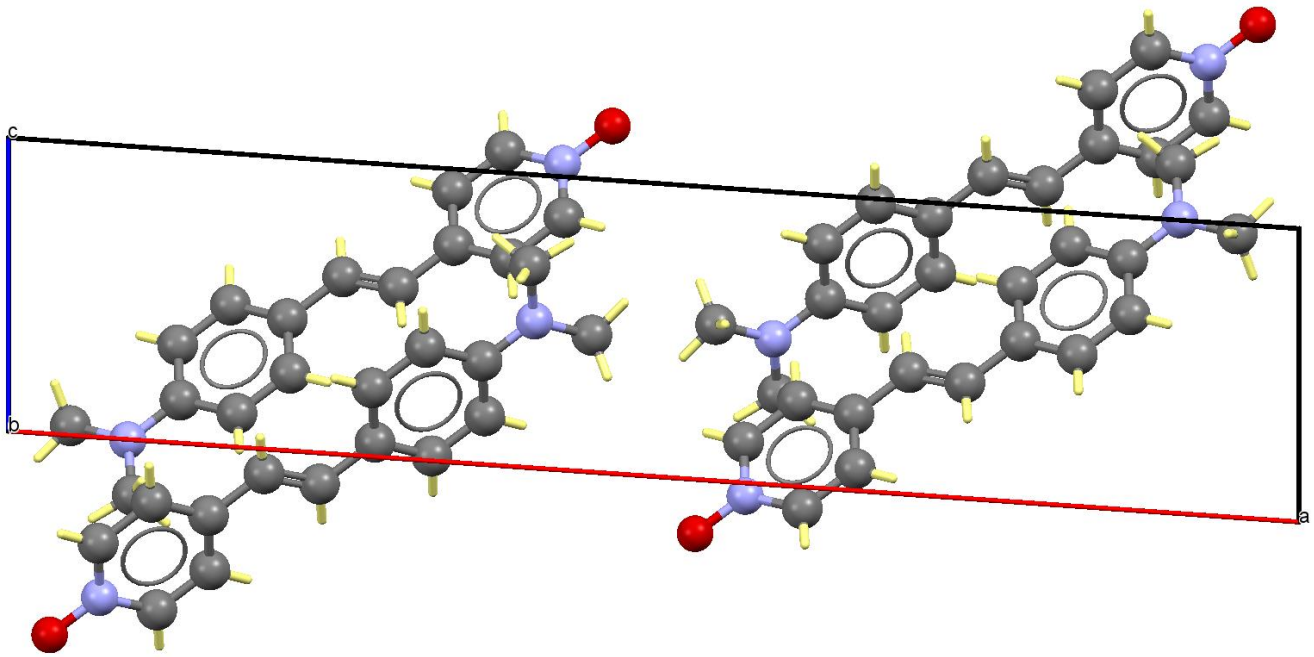
"Precise" *vs* "Reliable"



IJEKAJ, from XRPD

Maximum *Mogul z*-score < 3

"Precise" *vs* "Reliable"

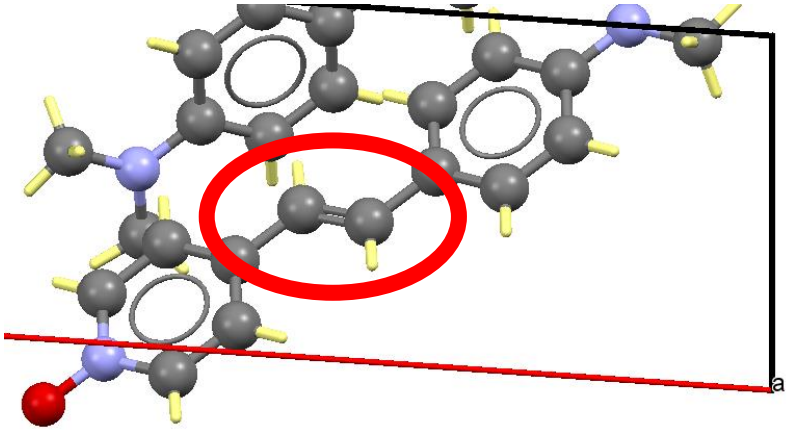


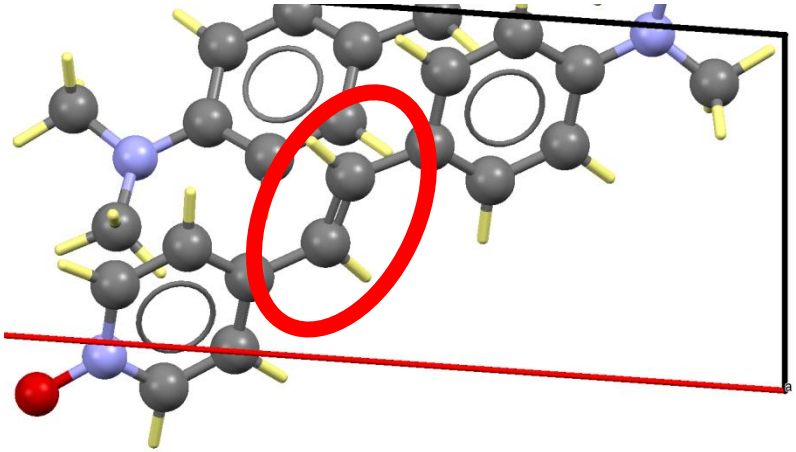
IJEKAJ, from XRPD

Maximum *Mogul z*-score < 3

Is this structure precise? ✓
Is this structure reliable?

"Precise" *vs* "Reliable"





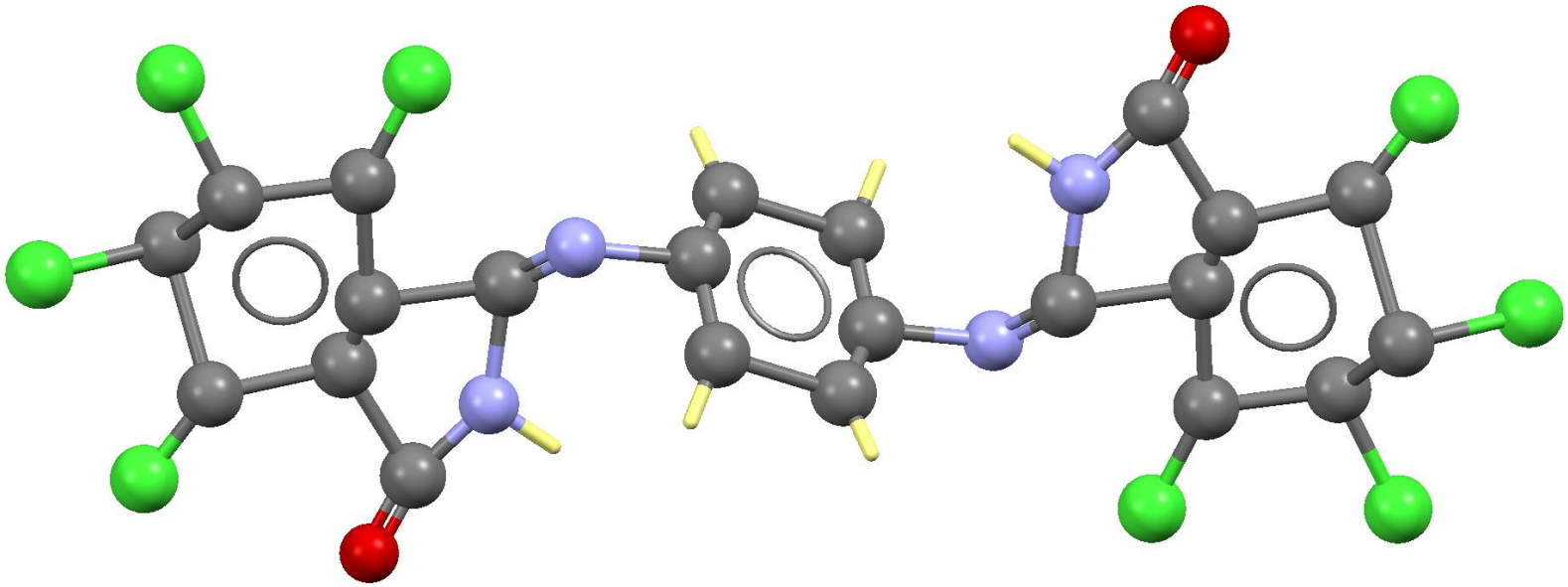
IJEKAJ, from XRPD

Maximum *Mogul* *z*-score < 3

Is this structure precise? ✓

Is this structure reliable? ✗

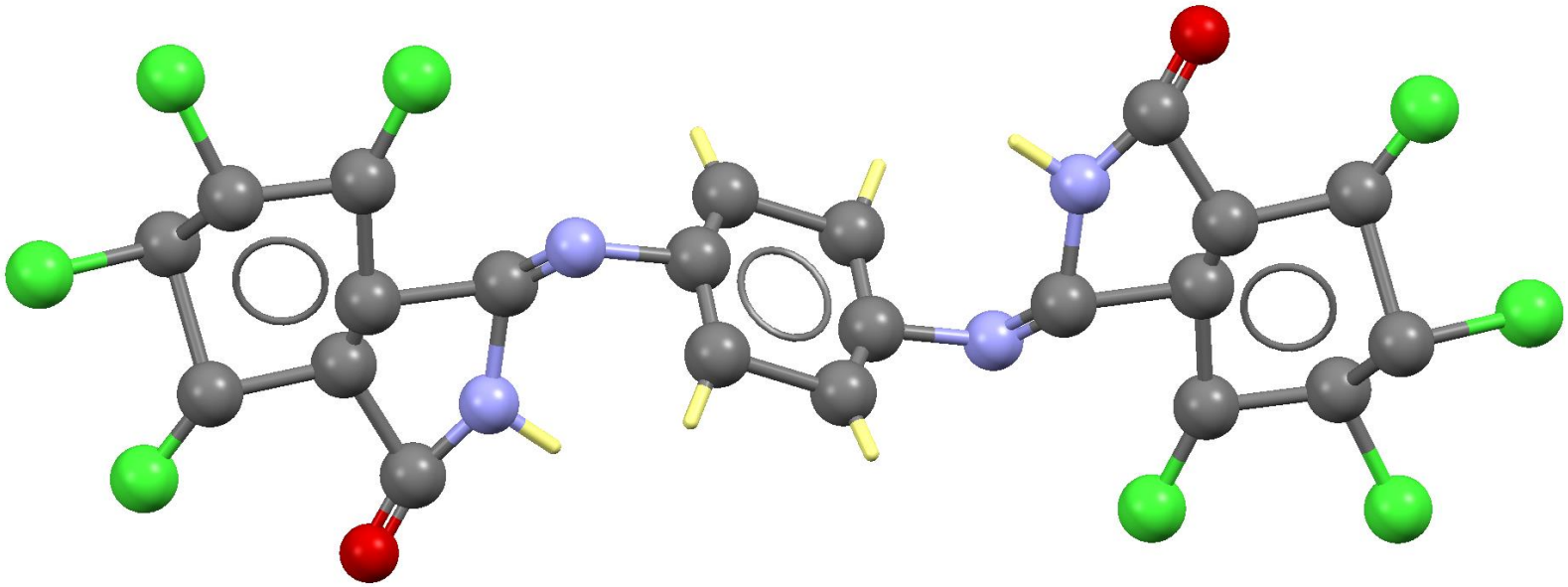
"Precise" *vs* "Reliable"



FEGDOL01, from XRPD

Maximum *Mogul* *z*-score > 25

"Precise" *vs* "Reliable"



FEGDOL01, from XRPD

Maximum *Mogul* *z*-score > 25

Is this structure precise? ✘
Is this structure reliable?

"Precise" *vs* "Reliable"

Correct position of the molecule
Correct orientation of the molecule
Correct connectivity
Correct unit cell, space group

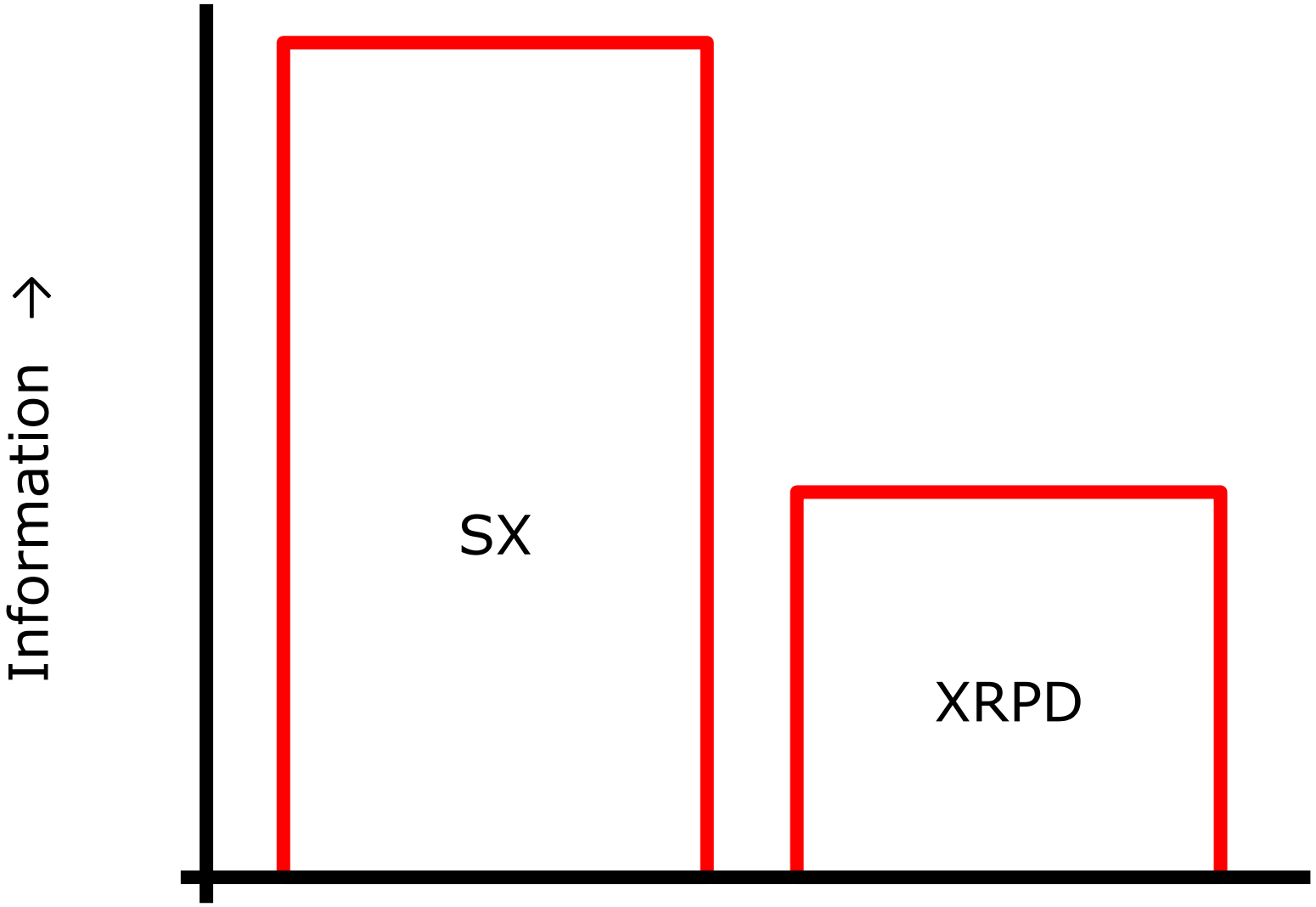
FEGDOL01, from XRPD

Maximum *Mogul* *z*-score > 25

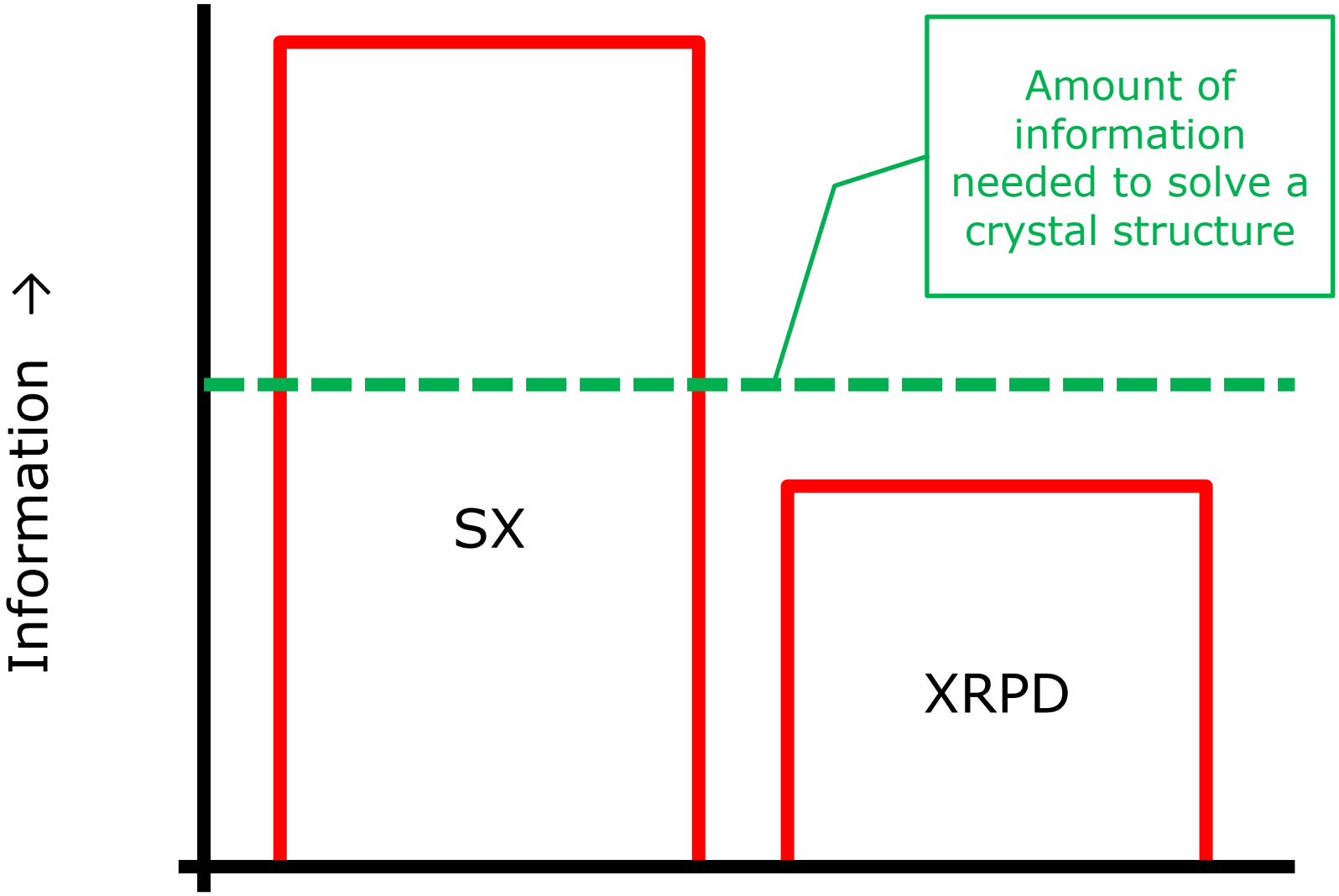
Is this structure precise? ✘
Is this structure reliable? ✓

How to determine *precise and reliable* molecular structures from XRPD?

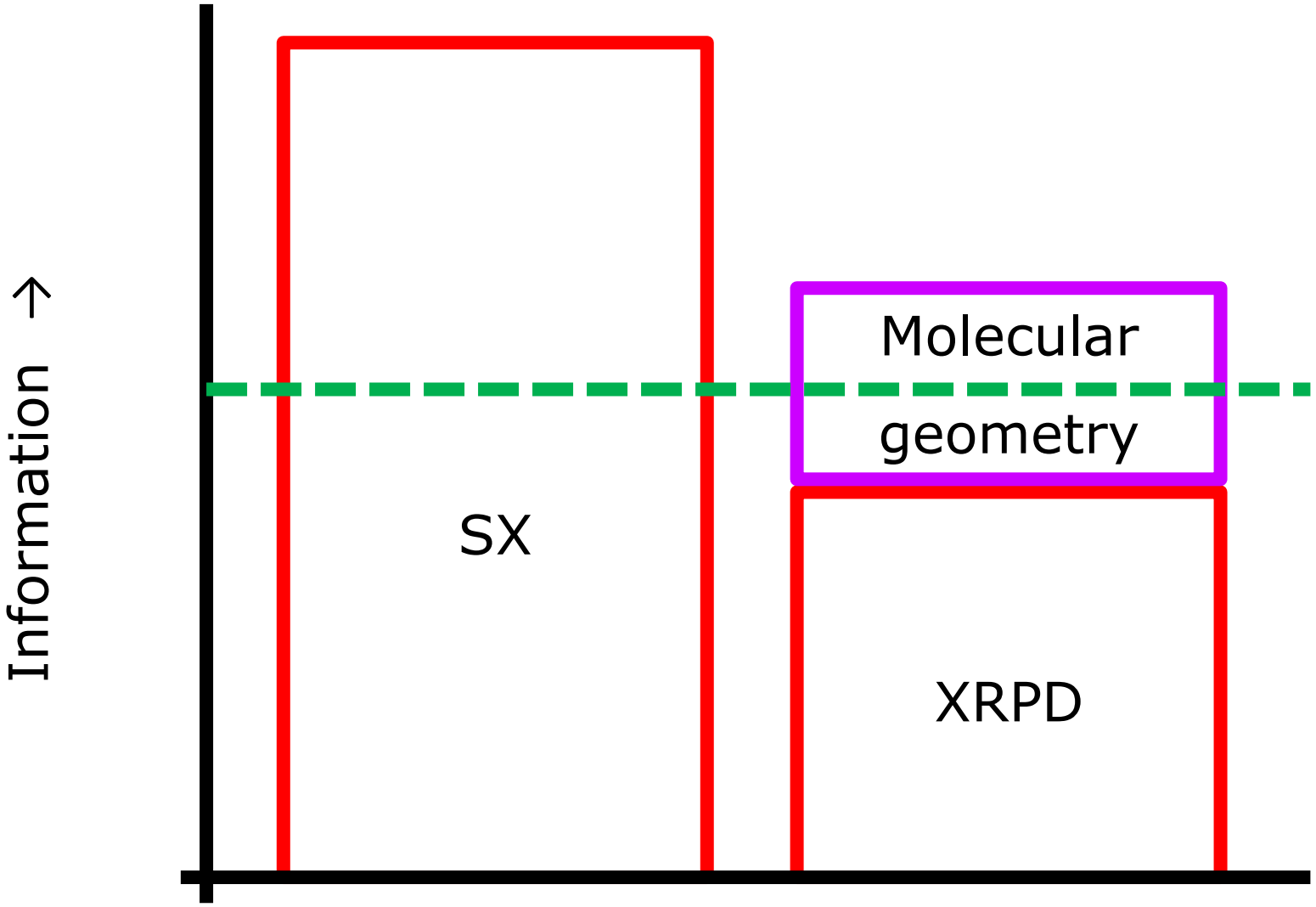
How to determine molecular structures from XRPD?



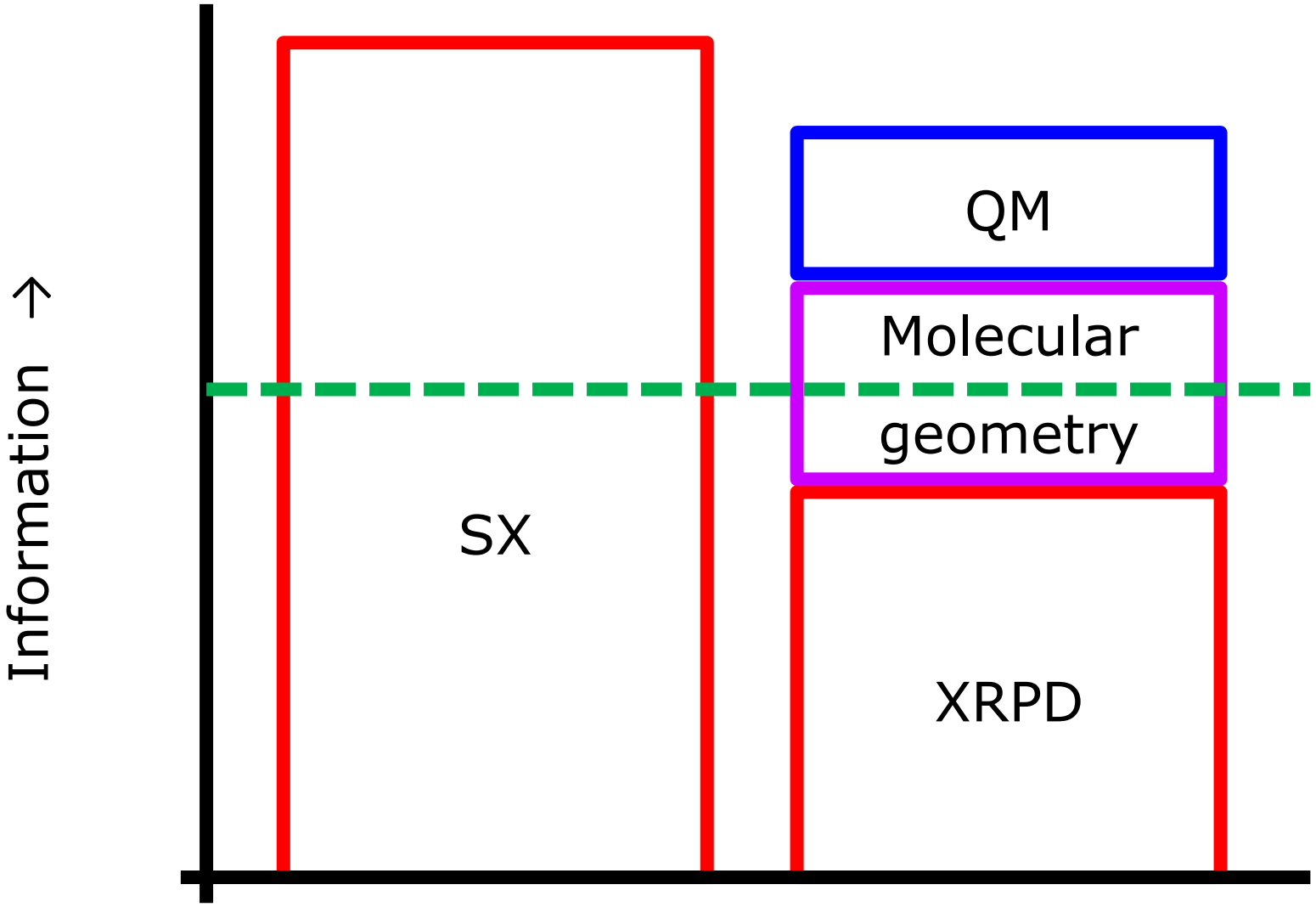
How to determine molecular structures from XRPD?



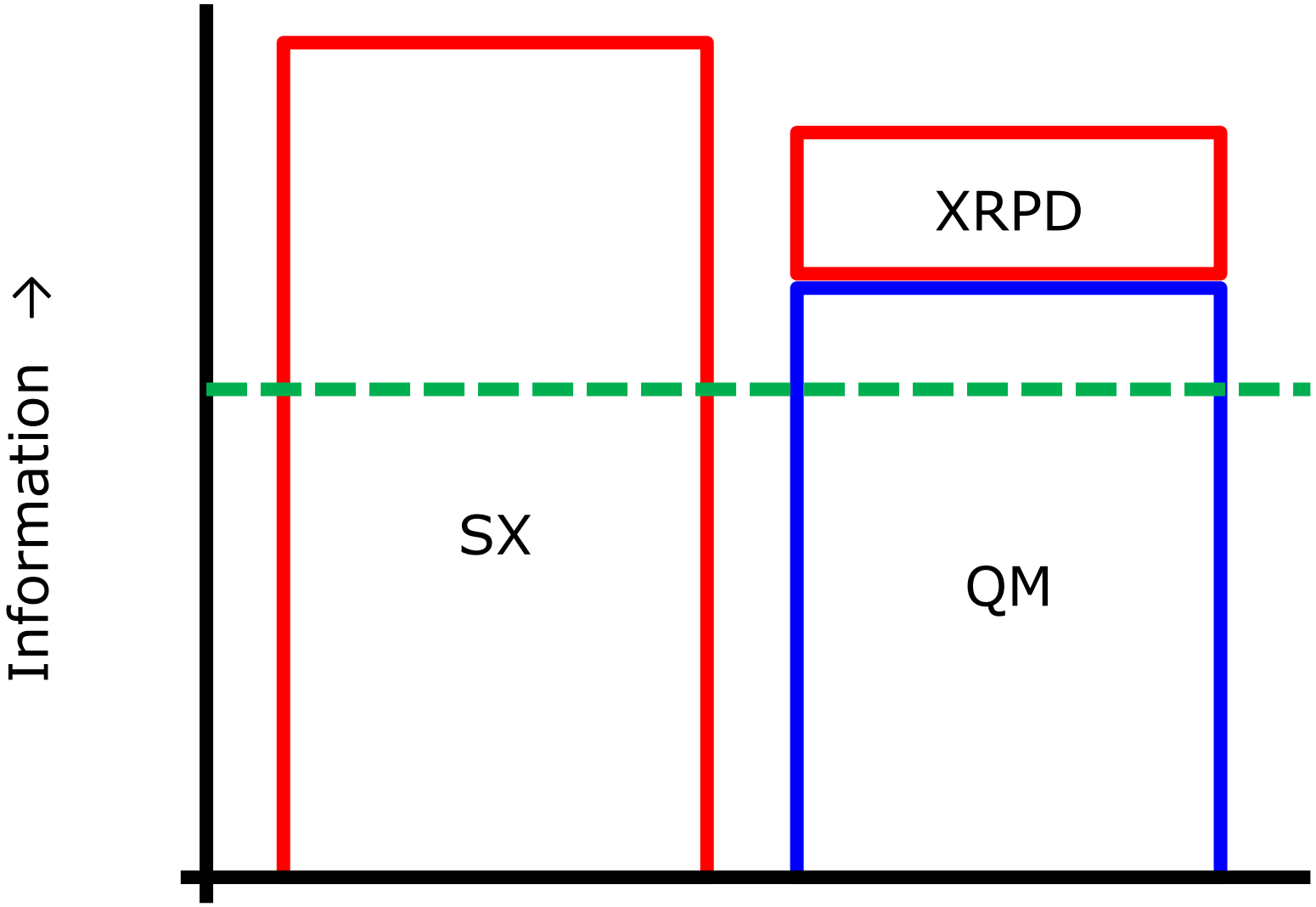
How to determine molecular structures from XRPD?



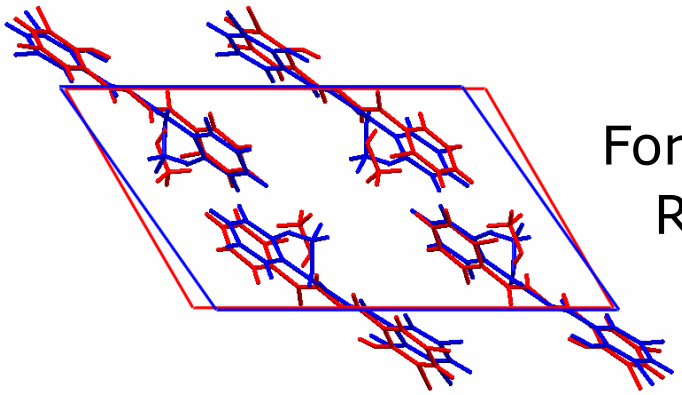
How to determine *precise and reliable* molecular structures from XRPD?



Crystal Structure Prediction



Dispersion-corrected DFT (DFT-D)

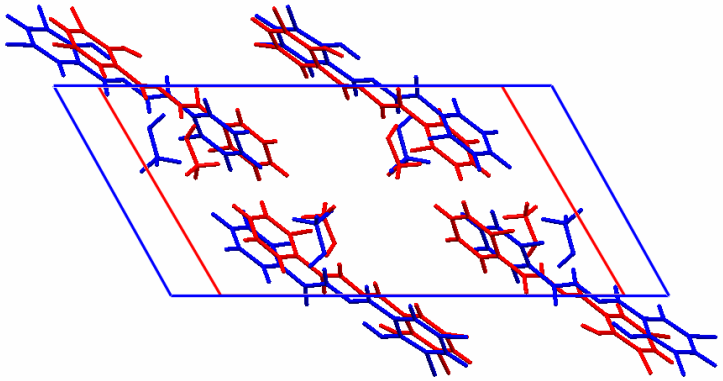


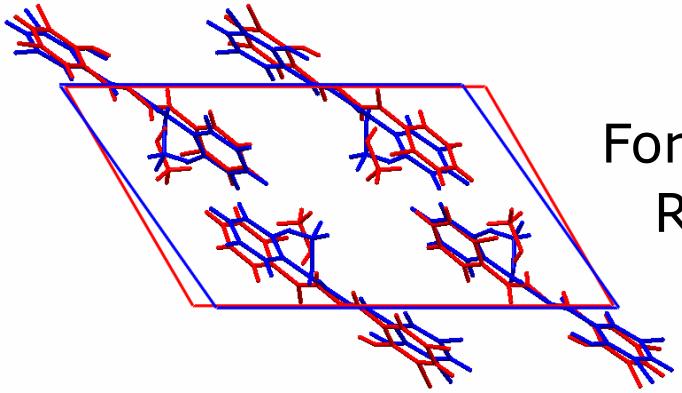
Force fields...
RMS 0.5 Å
10 sec

Minimised
Experiment

Dispersion-corrected DFT (DFT-D)

Pure DFT...
RMS 0.8 Å
100 hrs



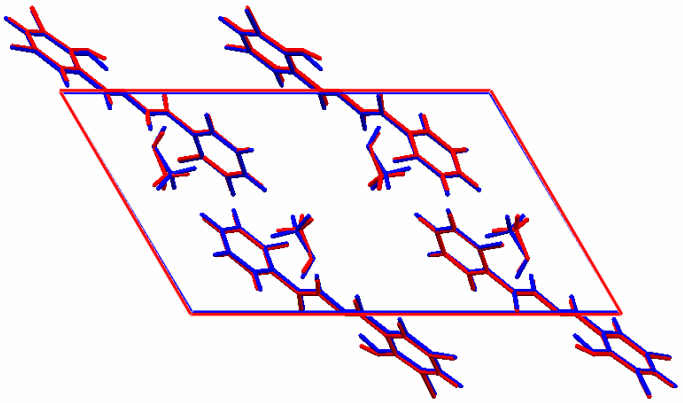


Force fields...
RMS 0.5 Å
10 sec

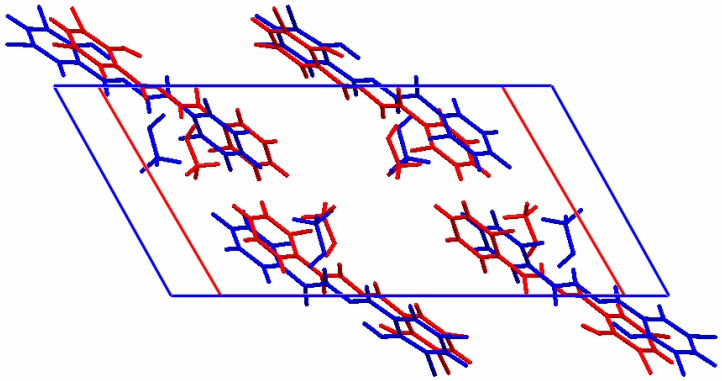
Minimised
Experiment

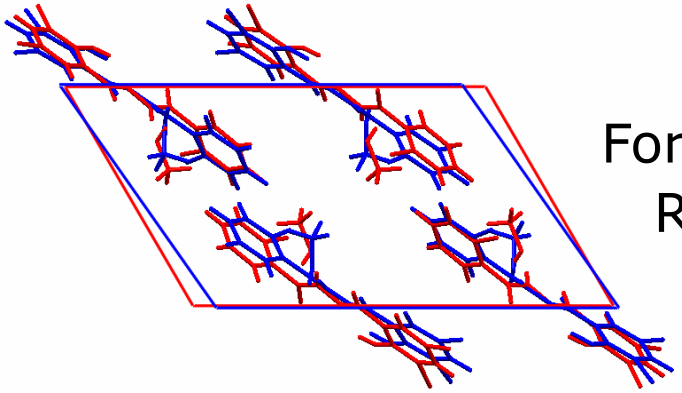
Dispersion-corrected DFT (DFT-D)

Dispersion-corrected DFT…
RMS 0.1 Å
100 hrs



Pure DFT...
RMS 0.8 Å
100 hrs





Force fields...
RMS 0.5 Å
10 sec

Minimised
Experiment

Dispersion-corrected DFT (DFT-D)

$$E_{\text{lattice}} = E_{\text{DFT}} + E_{\text{VdW}}$$

VASP or *CASTEP*

- PAW potentials
- Plane-wave basis set
- GGA – PBE
- 520 eV energy cut-off
- 0.07 Å^{-1} *k*-point spacing
- Grimme dispersion correction
- Static, T = 0 K

S. Grimme, J. Antony, S. Ehrlich & H. Krieg (2010) *J. Chem. Phys.* **132**, 154104
G. Kresse & J. Hafner (1993) *J. Phys. Rev. B* **47**, 558-561
Clark, Segall, Pickard, Hasnip, Probert, Refson & Payne (2005) *Z. Kristallogr*. **220**, 567

XRPD + DFT-D

The combination of XRPD + DFT(-D) has been around for some time and is fairly common these days.

Neumann, Tedesco, Destri, Ferro & Porzio (2002) *J. Appl. Cryst*. **35**, 296-303.

Avila, Mora, Delgado, Contreras, Fitch & Brunelli (2008) *Acta. Cryst*. **64**, 217-222.

Florence, Bardin, Johnston, Shankland, Griffin & Shankland (2009) *Z. Krist*. **30**, 215-220.

Bekö, Thoms, Brüning, Alig, Van de Streek, Lakatos, Glaubitz & Schmidt (2010) *Z. Krist*. **225**, 382-387.

Blanton, Rajeswaran, Stephens, Whitcomb, Misture & Kaduk (2011) *Powder Diffr*. **26**, 313-320.

Book "Uniting Electron Crystallography and Powder Diffraction" (2012) Chapter "Powder Diffraction+Computational Methods" by L'ubomír Smrčok.

Reliable Structures with DFT-D

Reproduction of Molecular Crystal Structures

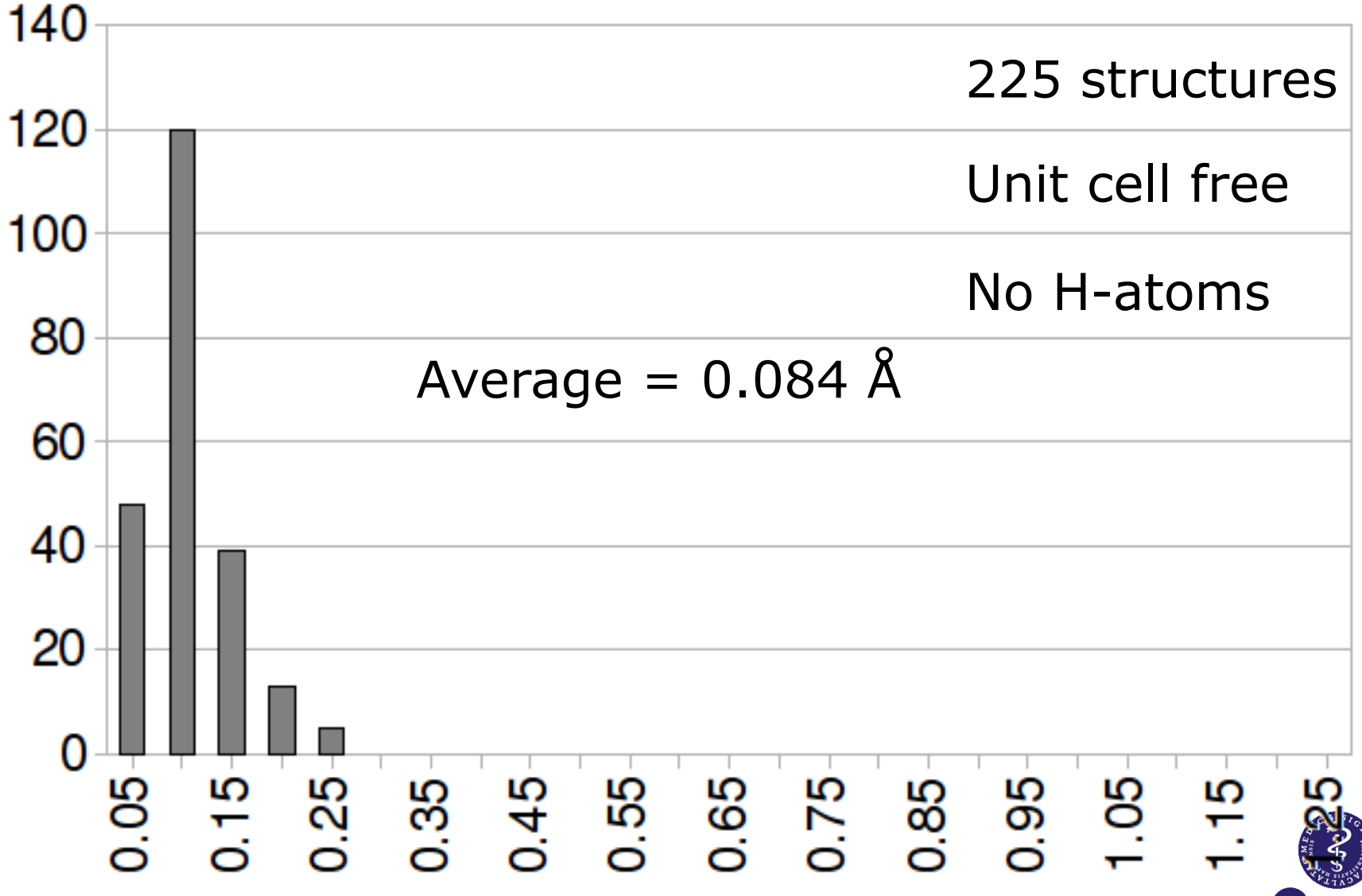
225 high-quality single-crystal structures from the August 2008 issue of *Acta Crystallographica* E were downloaded (Open Access!) and energy-minimised

225 experimental single-crystal structures...
225 energy-minimised structures...

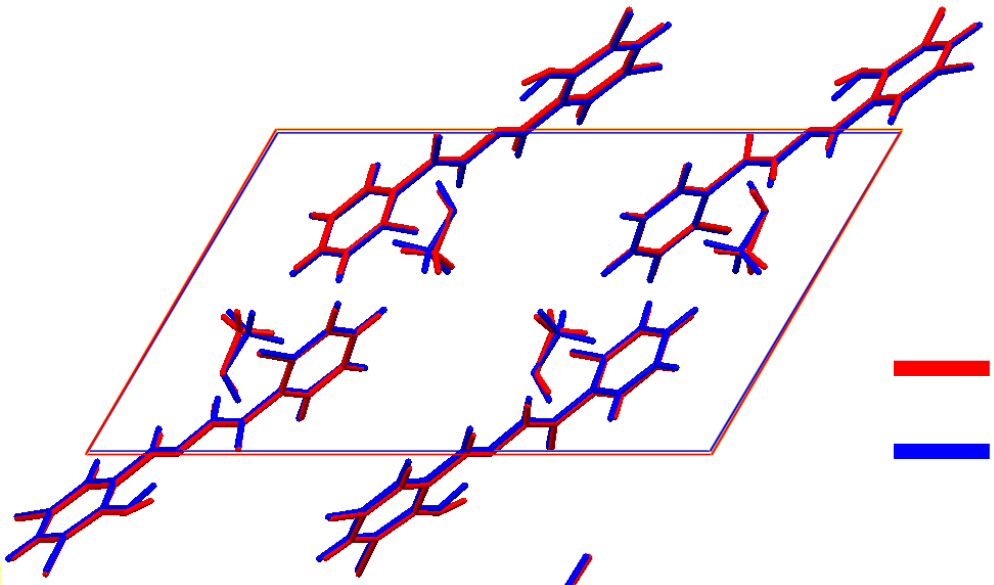
How well are the experimental structures reproduced?

J. van de Streek & M. A. Neumann (2010). *Acta Cryst.* B**66**, 544-558

RMS Cartesian Displacement

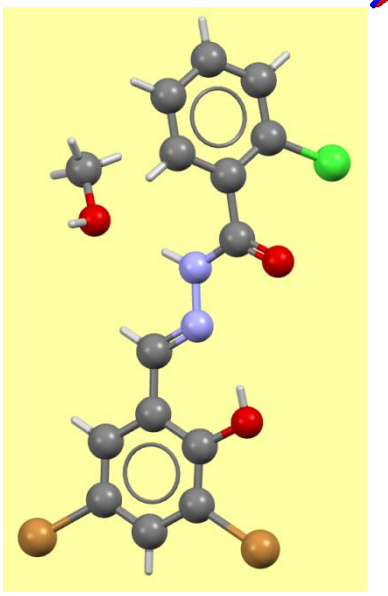


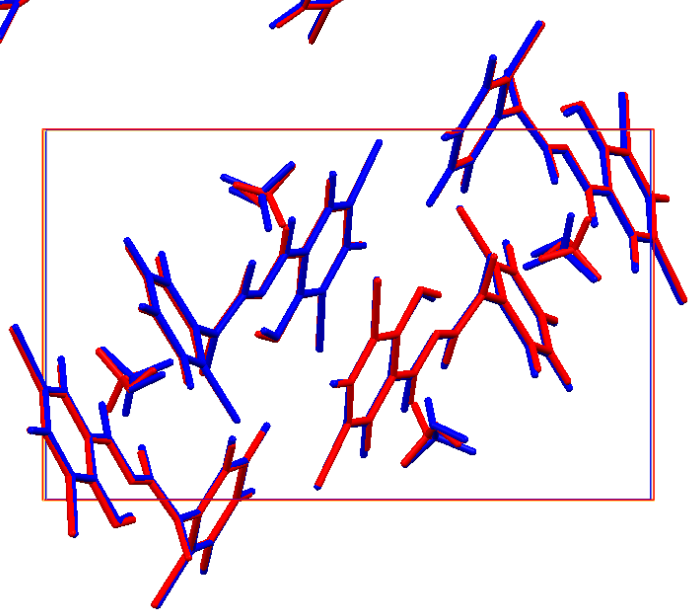
0.084 Å

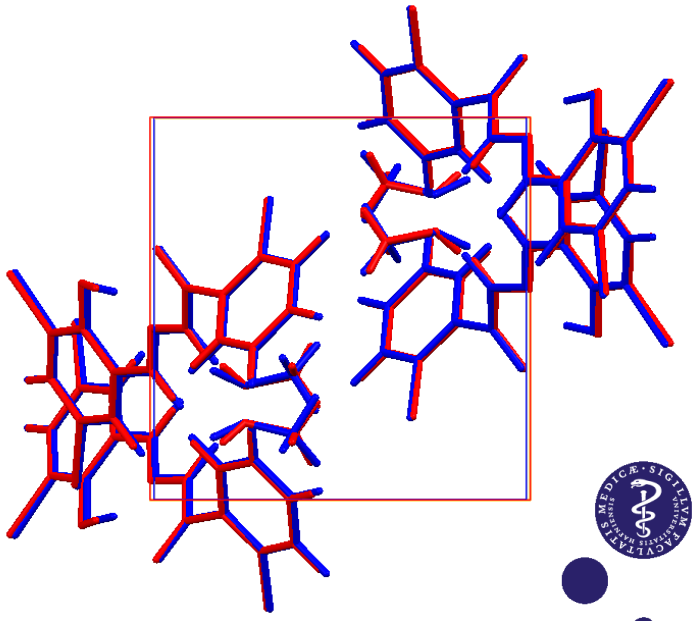


Unit cell free
ΔV = -3%

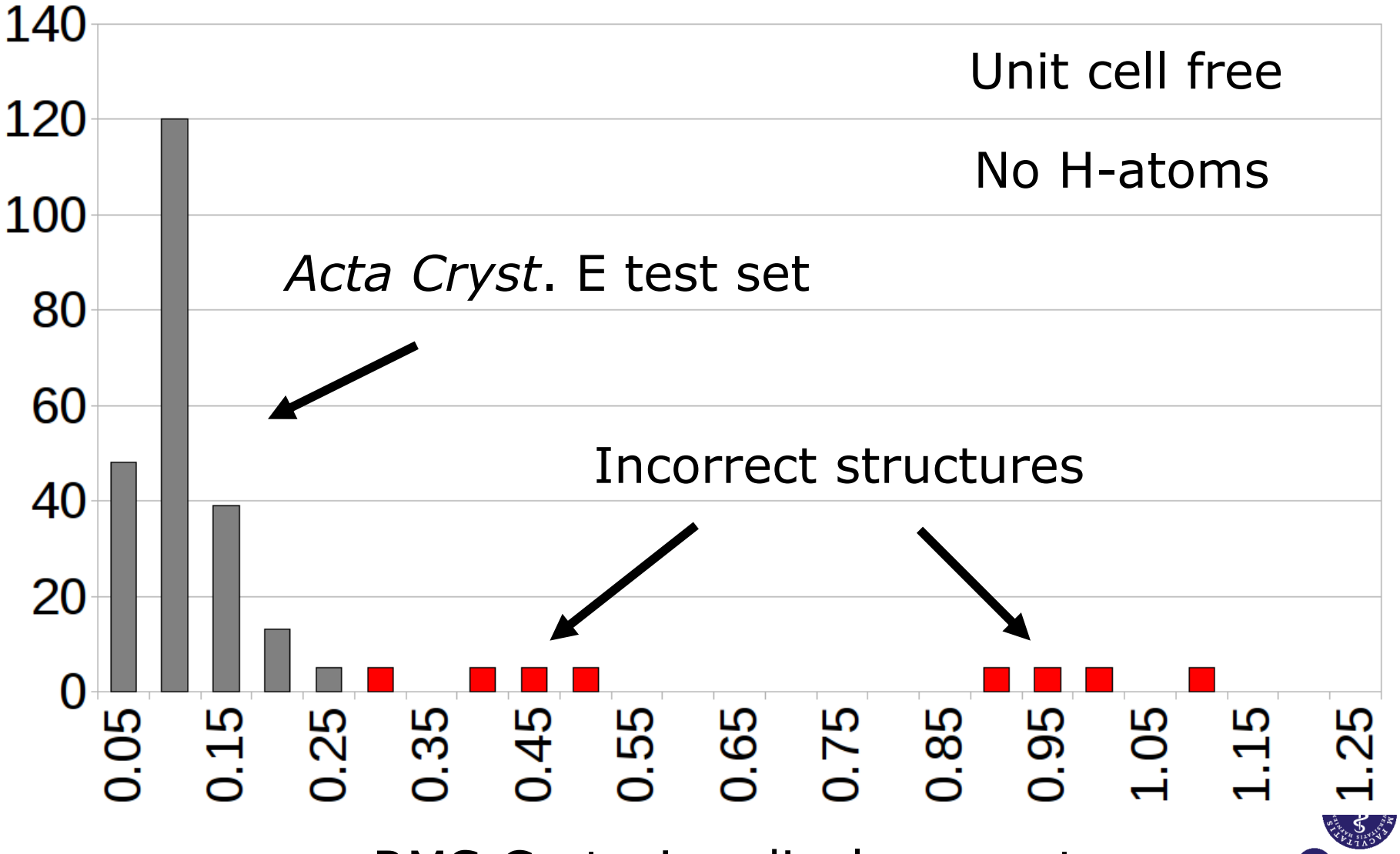
Experiment
Minimised







What about Wrong Structures?



Reliability

Because of this validation study, we can determine the correctness and precision of crystal structures from XRPD semi-quantitatively

$$SX \approx DFT\text{-}D = Virtual\ SX$$

$$XRPD \rightarrow DFT\text{-}D$$

$$XRPD \rightarrow Virtual\ SX$$

Let's do exactly that…

In 2010
225 SX structures minimised with DFT-D

In 2014
215 XRPD structures minimised with DFT-D

Let's do exactly that…

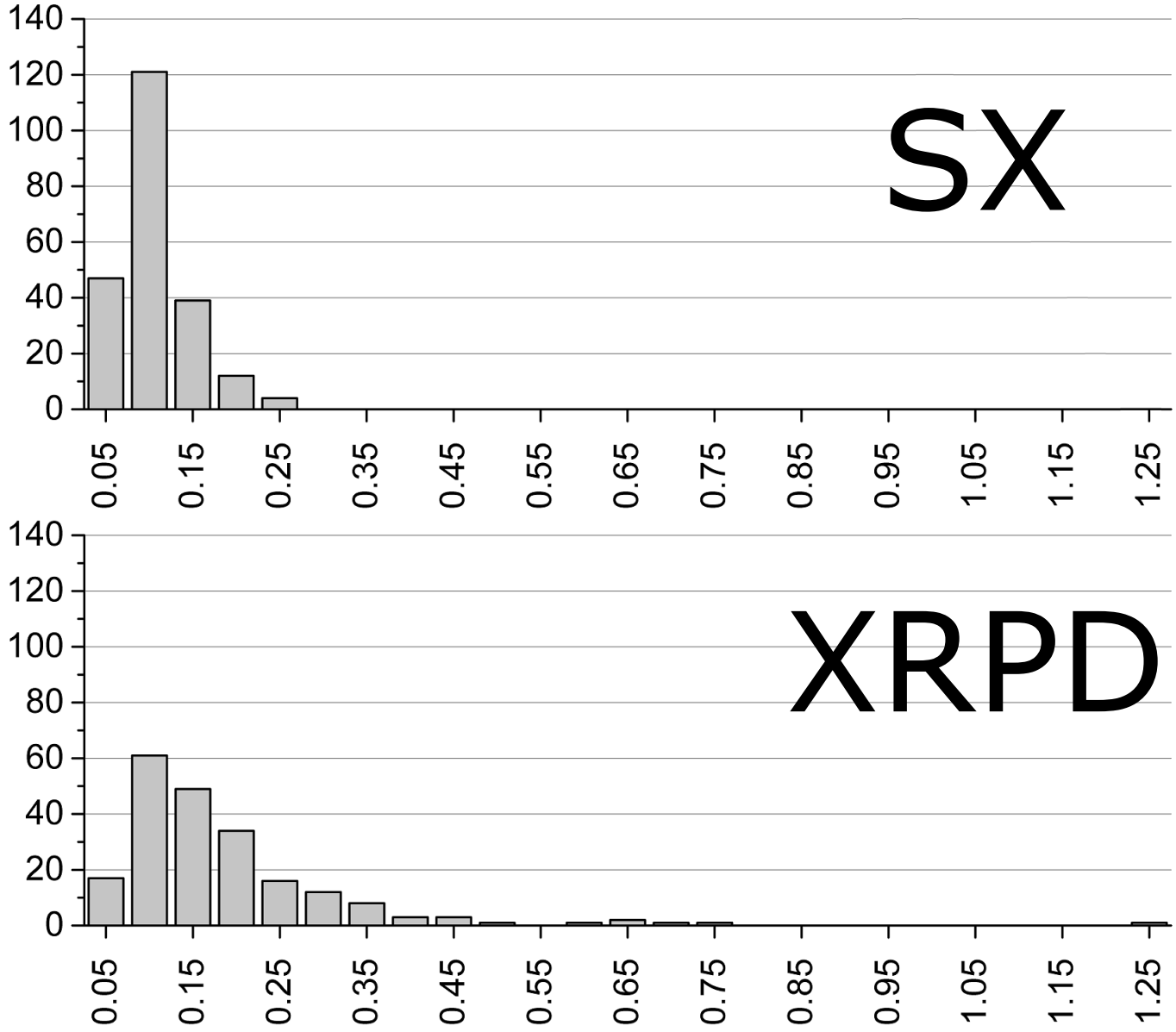
215 XRPD structures from all IUCr journals were retrieved from the CSD and energy-minimised

215 experimental XRPD structures...
215 energy-minimised structures...

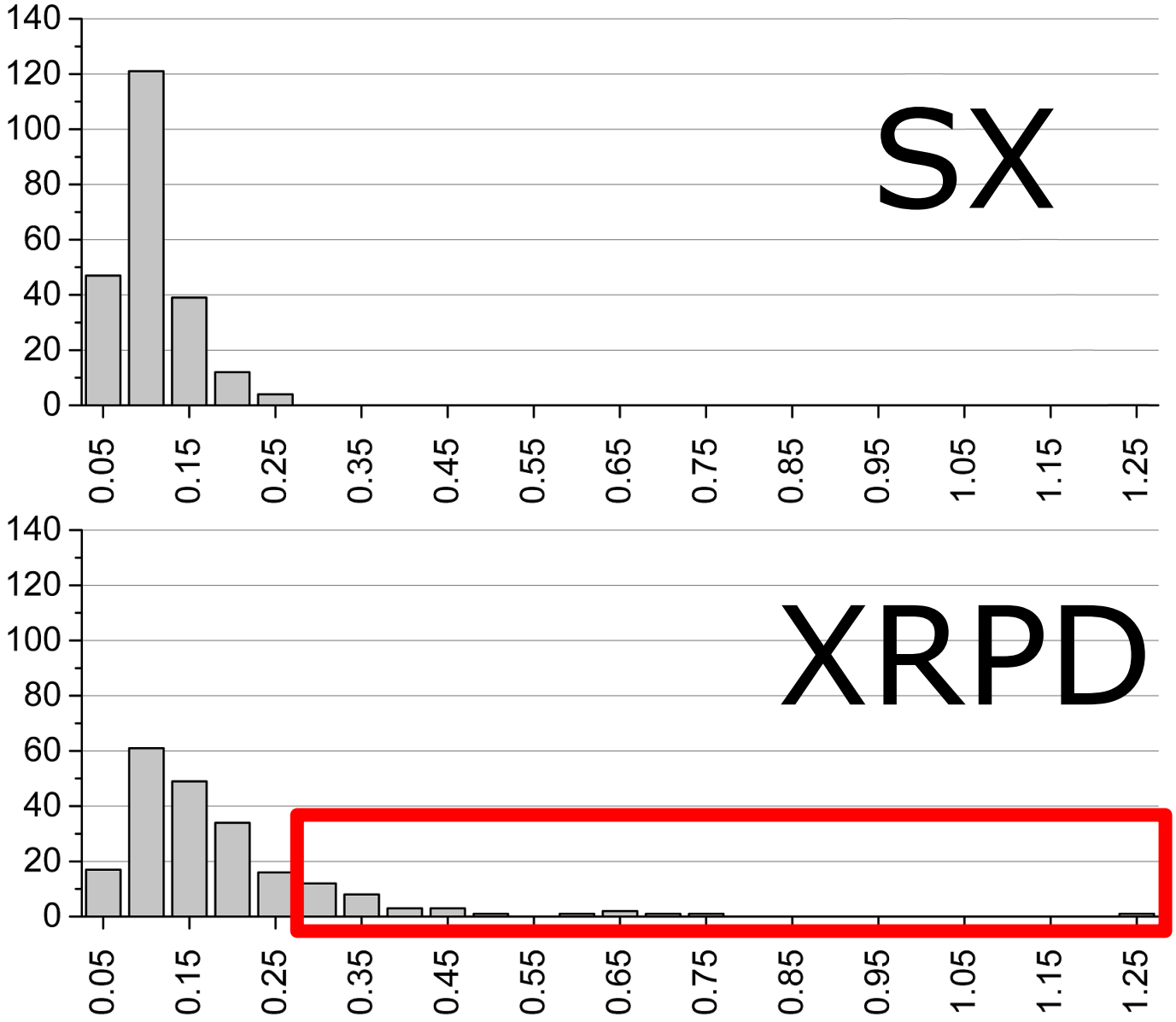
How well are the experimental structures reproduced?

J. van de Streek & M. A. Neumann (2014). *Acta Cryst.* B**70**, 1020-1032

RMS Cartesian Displacement



RMS Cartesian Displacement



High RMSCD values

After thorough analysis of all “suspicious” structures individually, roughly three categories:

1. Correct, but less precise

2. Error in the H atom positions

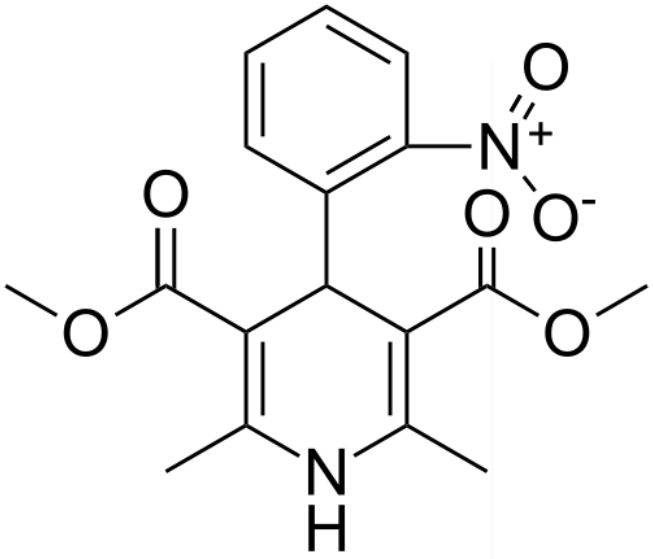
3. Minor error in one or two non-H atom positions

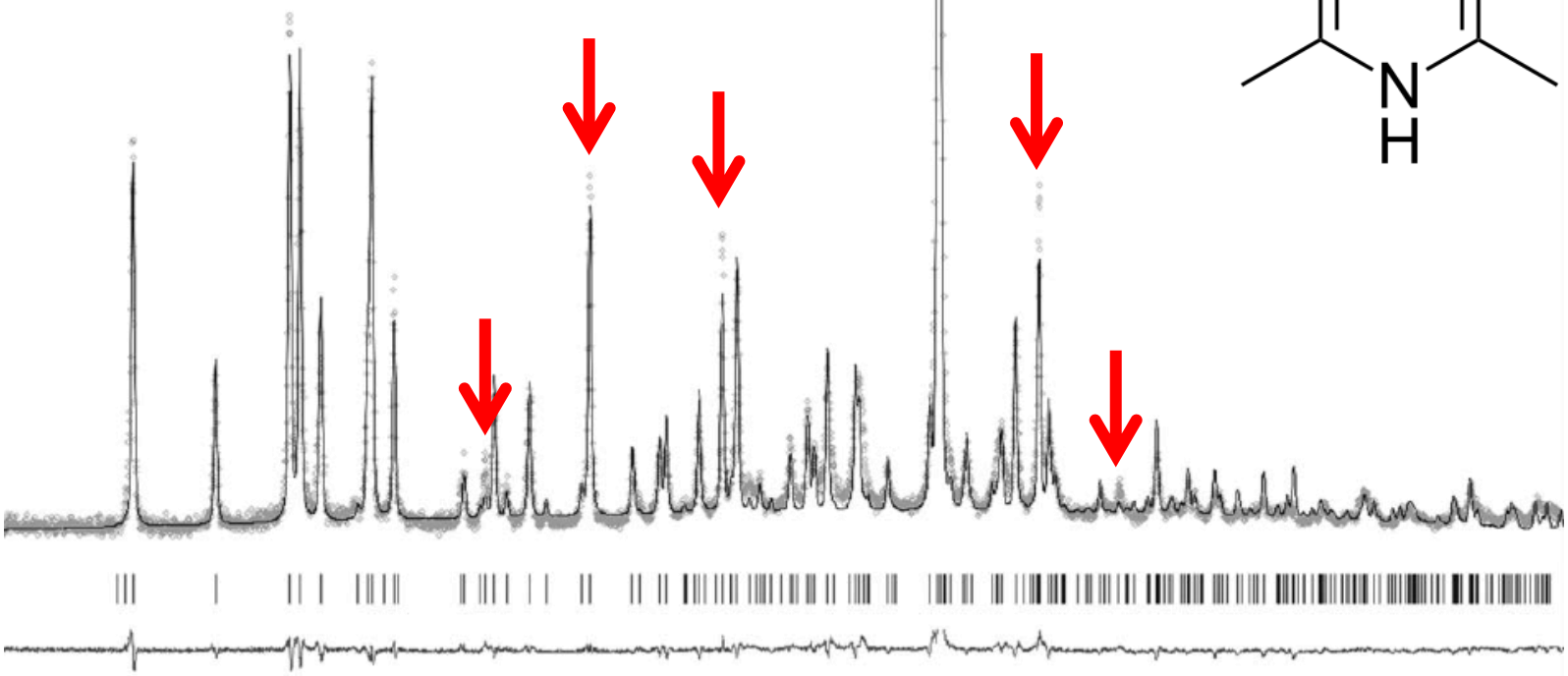
An examples of each category...

Example for case 1: Correct, but less precise

1. Correct, but less precise

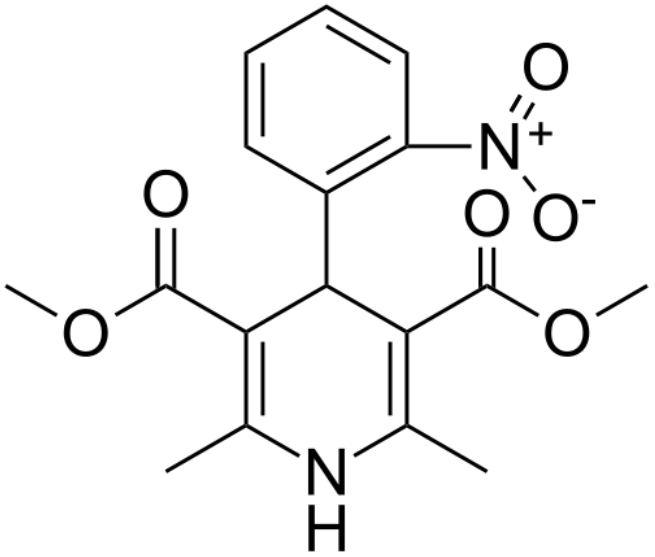
BICCIZ01 (2011)
RMSCD = 0.34 Å
χ^2 = 20

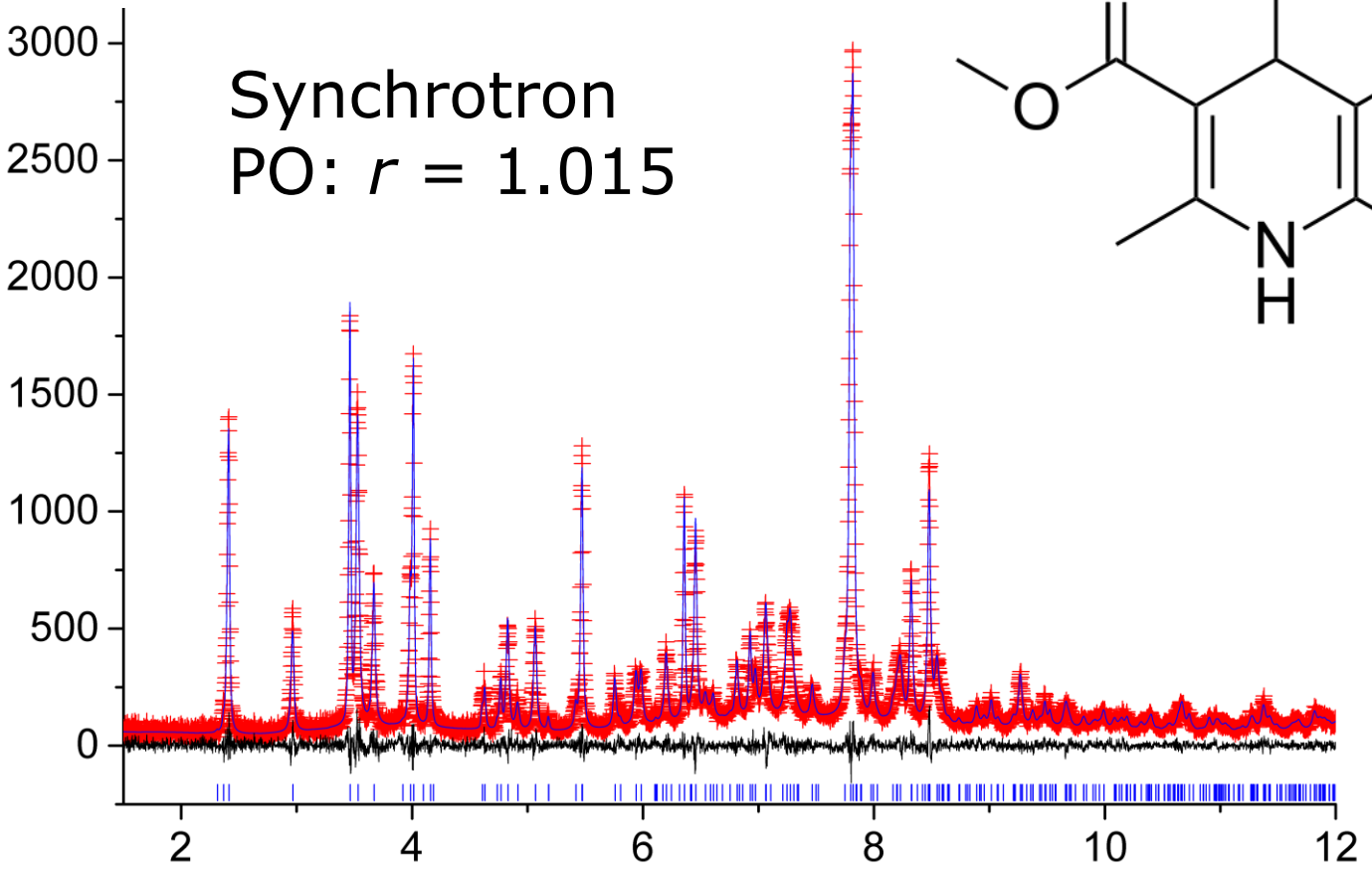




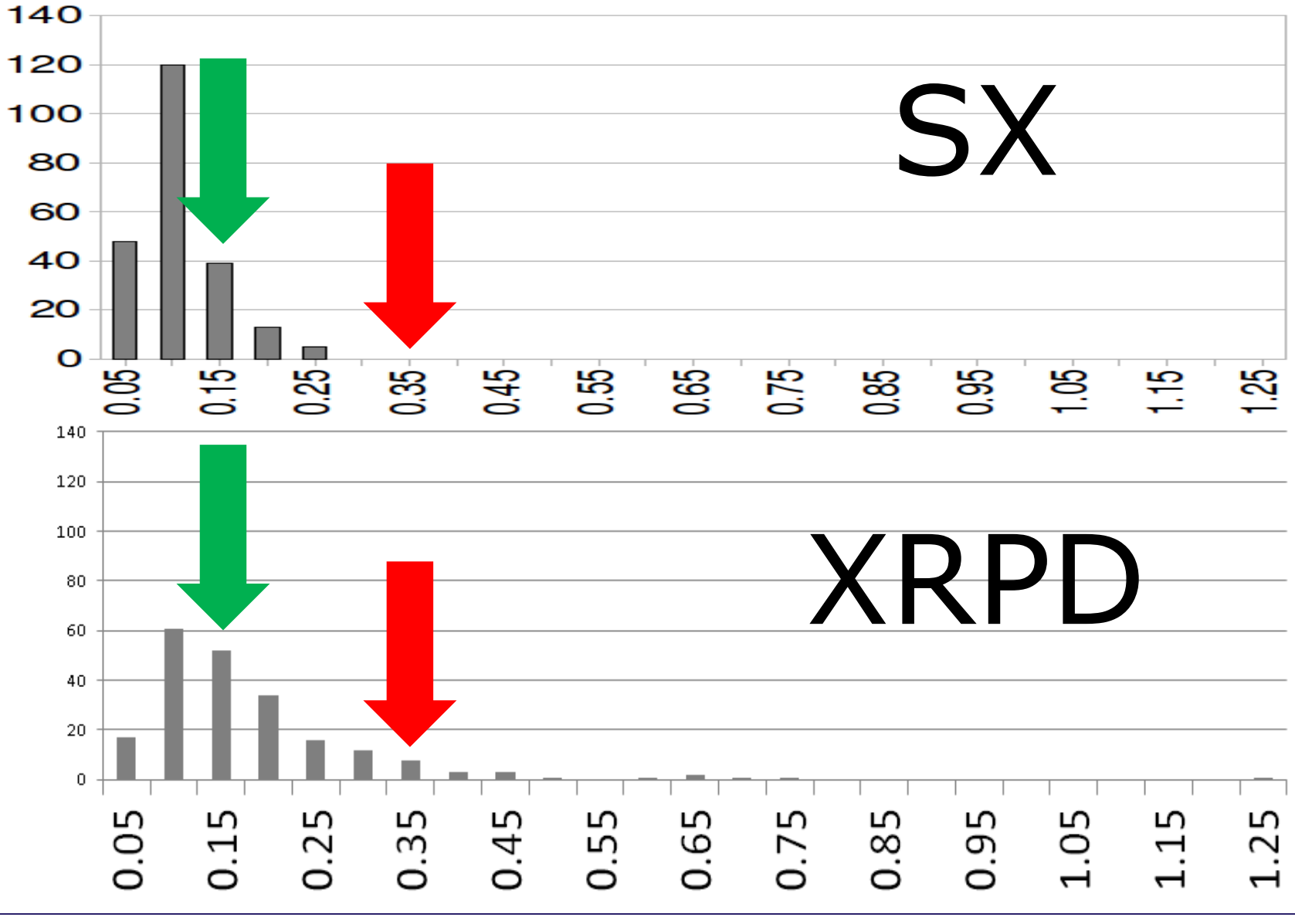
1. Correct, but less precise

BICCIZ01 re-refined with *TOPAS*
RMSCD = 0.13 Å
χ^2 = 2.2

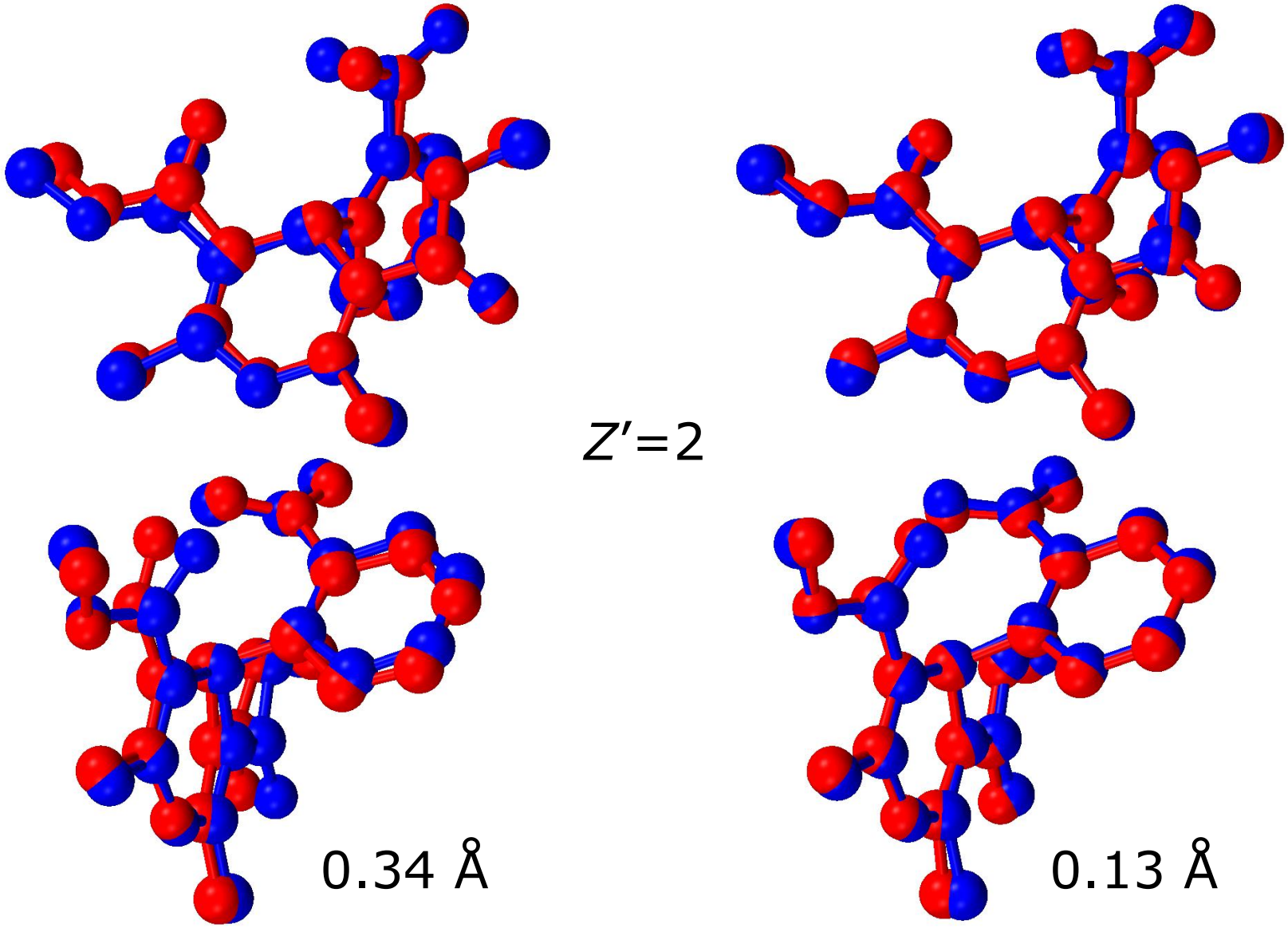




RMS Cartesian Displacement



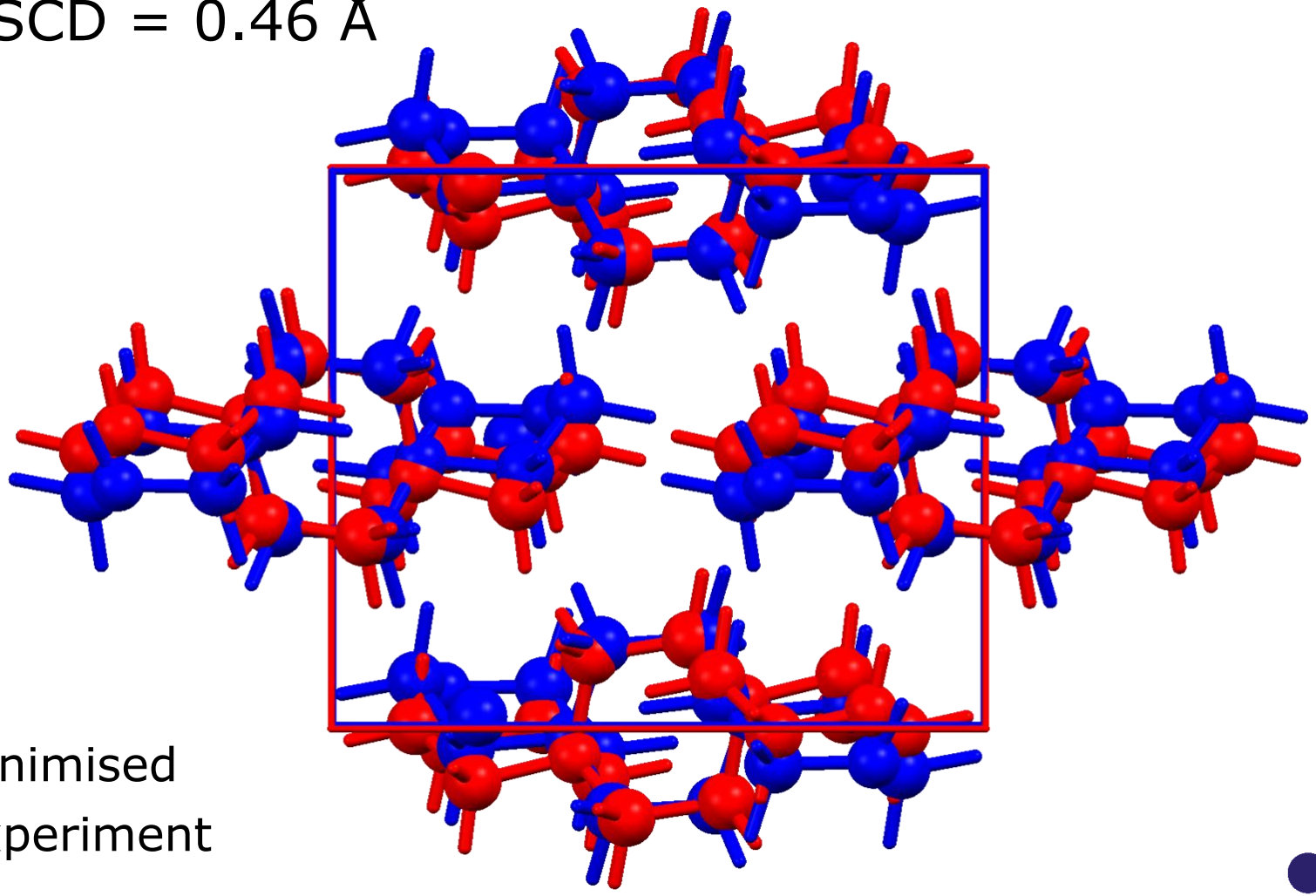
RMS Cartesian Displacement



Example for case 2: Error in H-atom positions

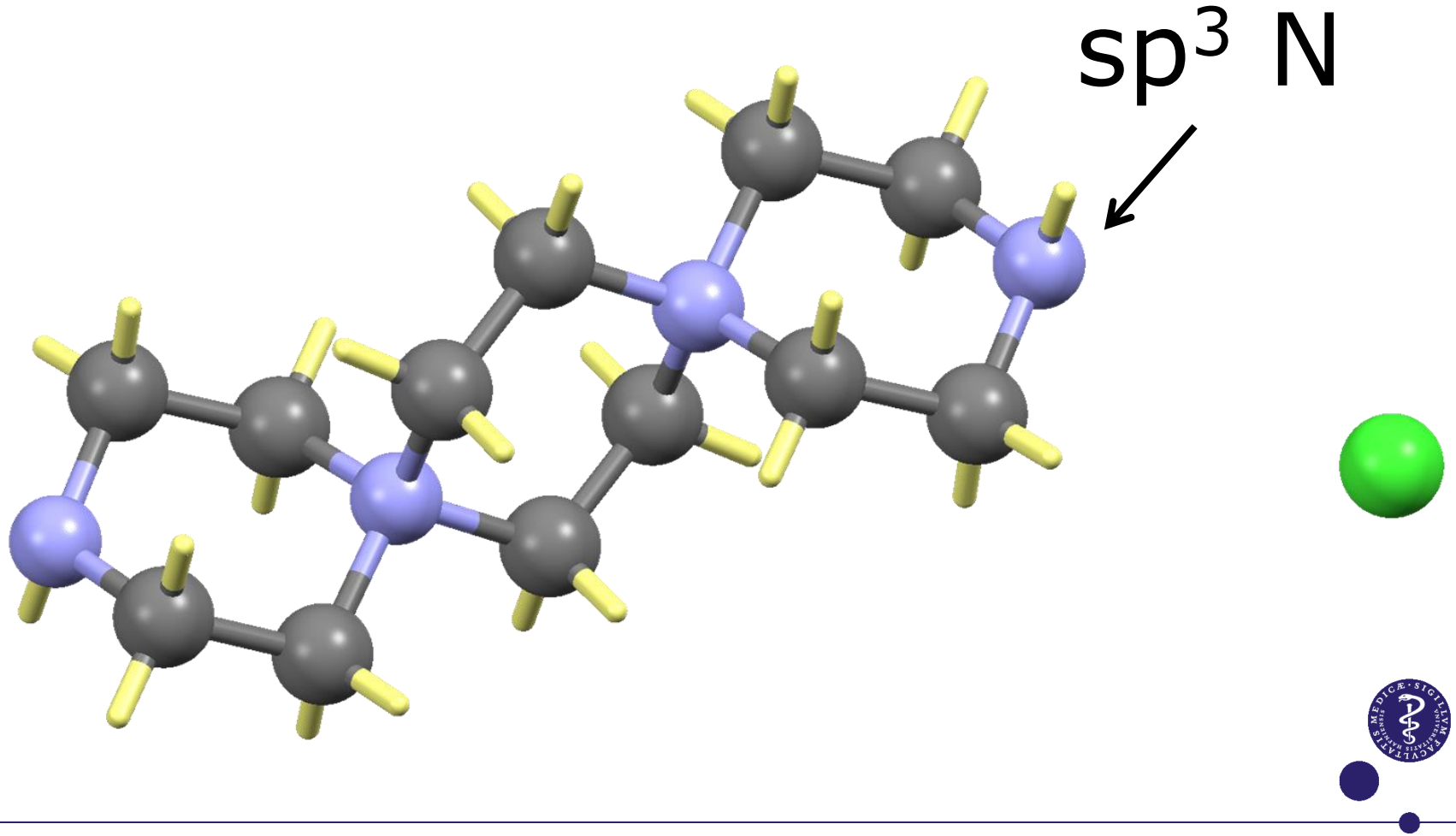
2. Error in H-atom positions

GOLTUW (1999), Rietveld refinement fine
RMSCD = 0.46 Å



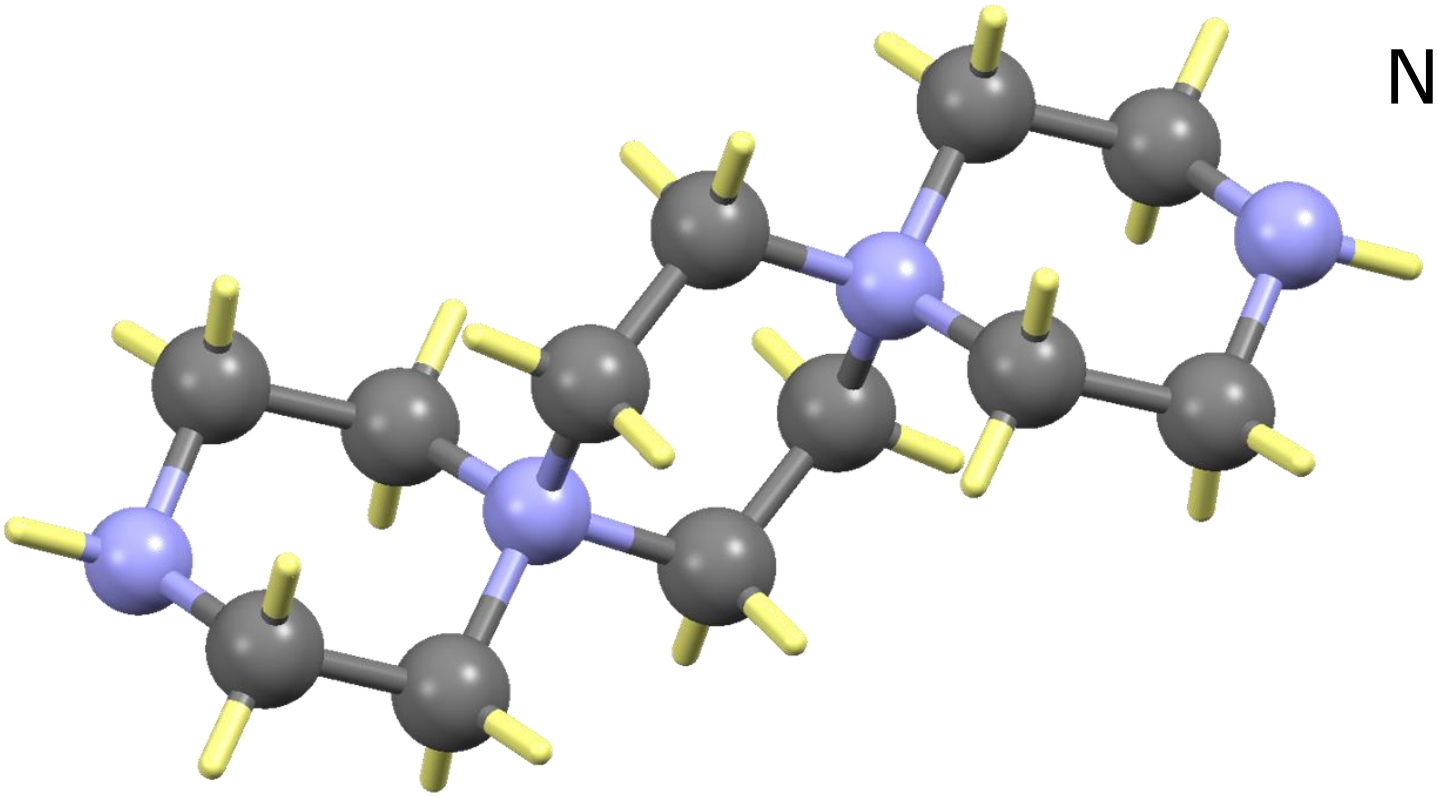
2. Error in H-atom positions

GOLTUW (1999)
RMSCD = 0.46 Å



2. Error in H-atom positions

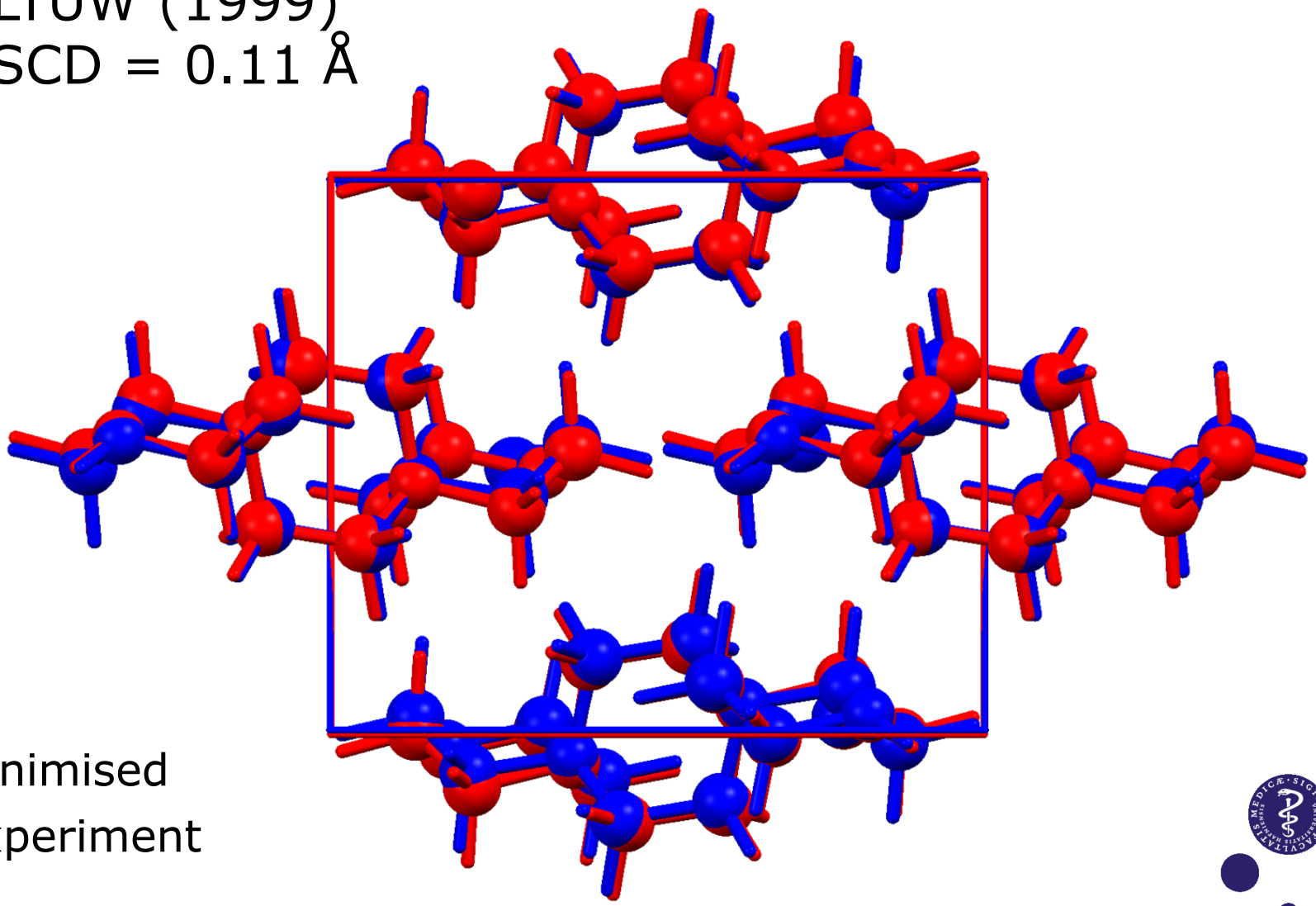
GOLTUW (1999)
RMSCD = 0.11 Å, 3.6 kcal/mol more favourable



No H bond

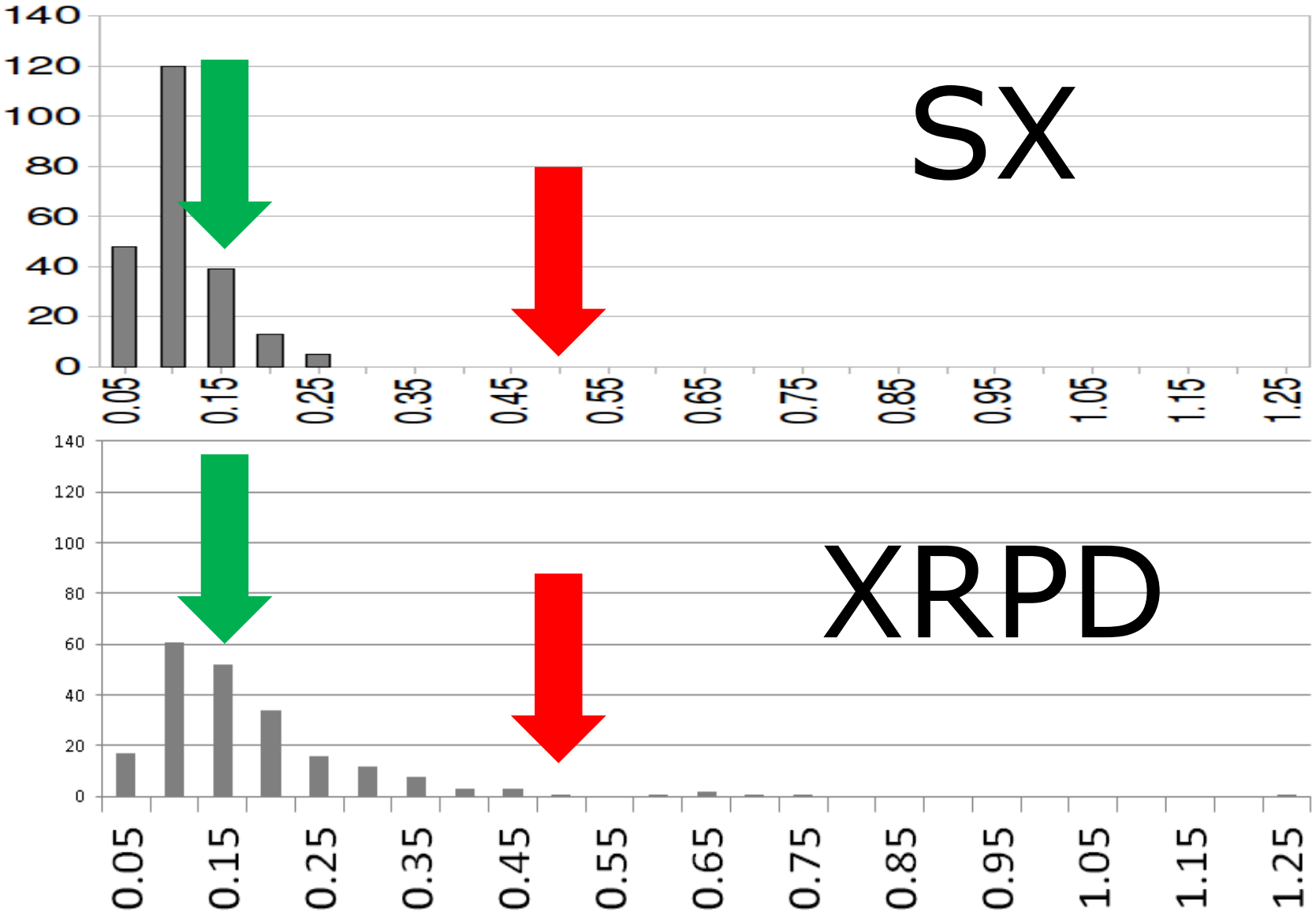
2. Error in H-atom positions

GOLTUW (1999)
RMSCD = 0.11 Å



Minimised
Experiment

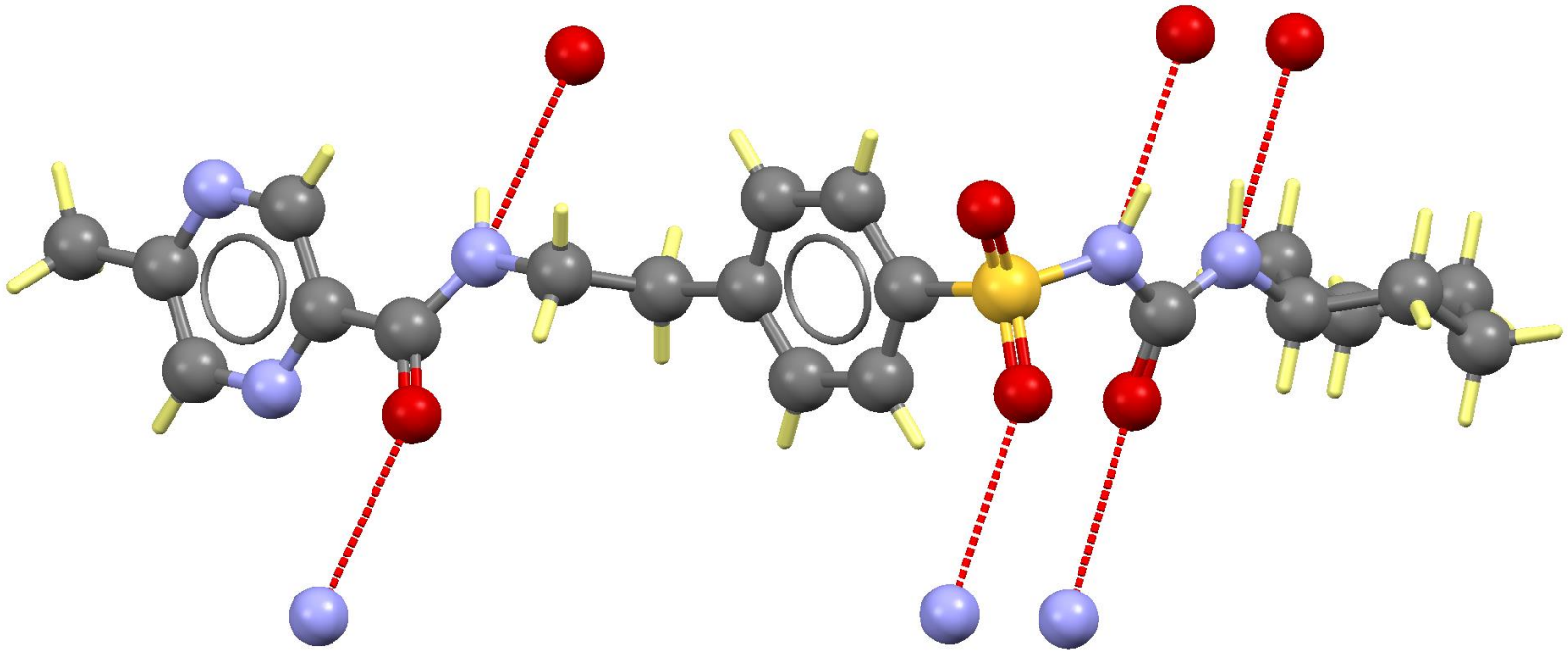
RMS Cartesian Displacement



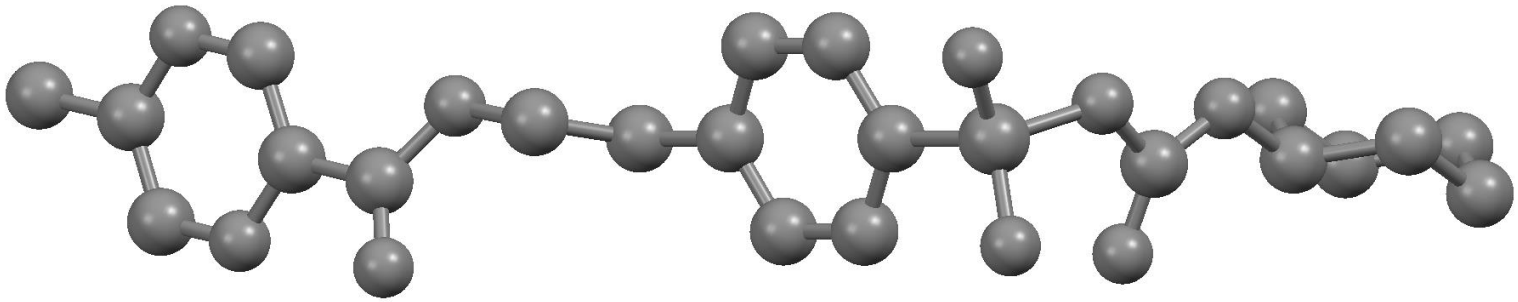
Example for case 3: Error in non H-atom position

3. Minor Error in Non-H atom Positions

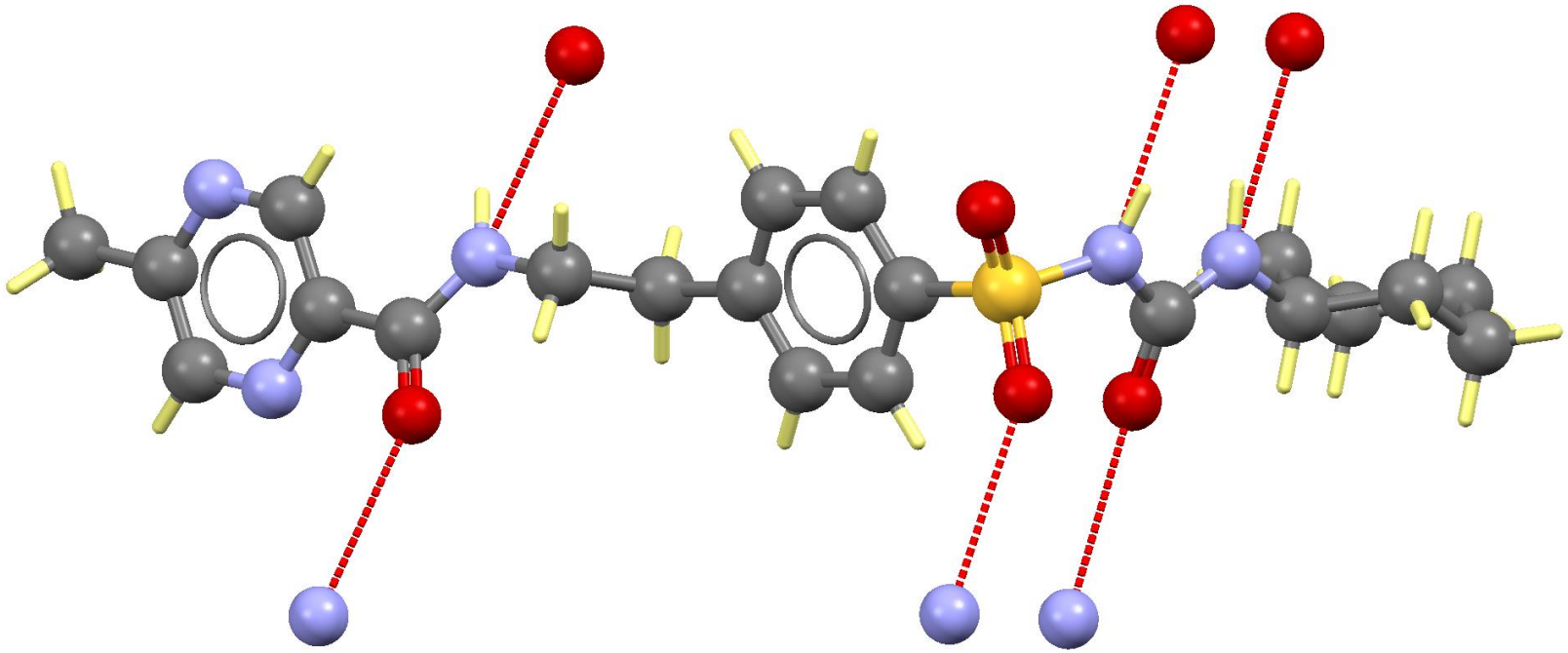
SAXFED (Glipizide, 2005)
RMSCD = 0.72 Å



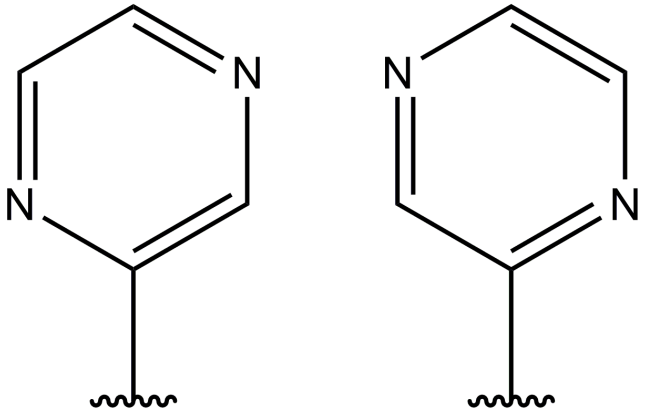
What XRPD Sees

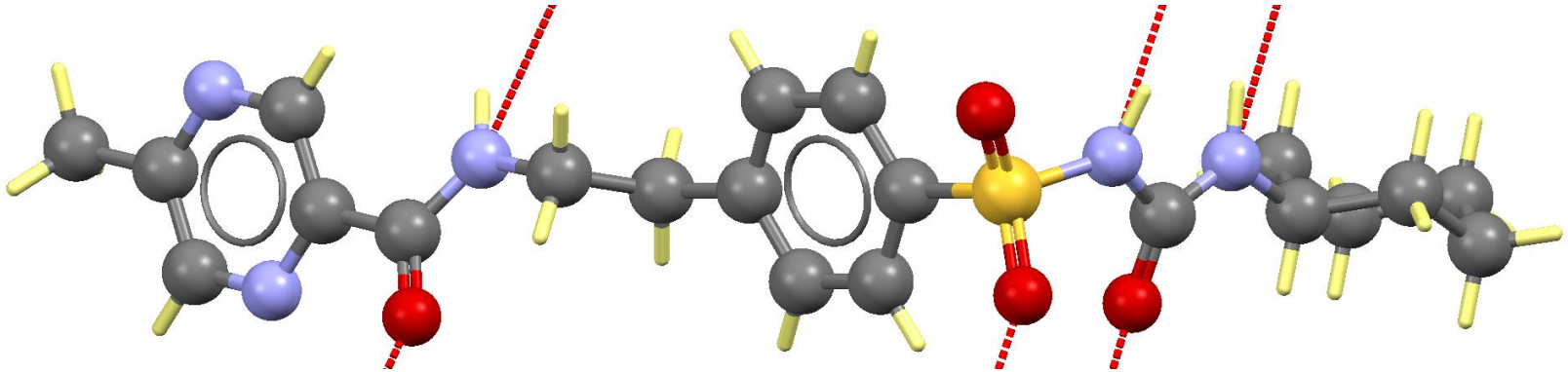


Chemist's impression



Ambiguity



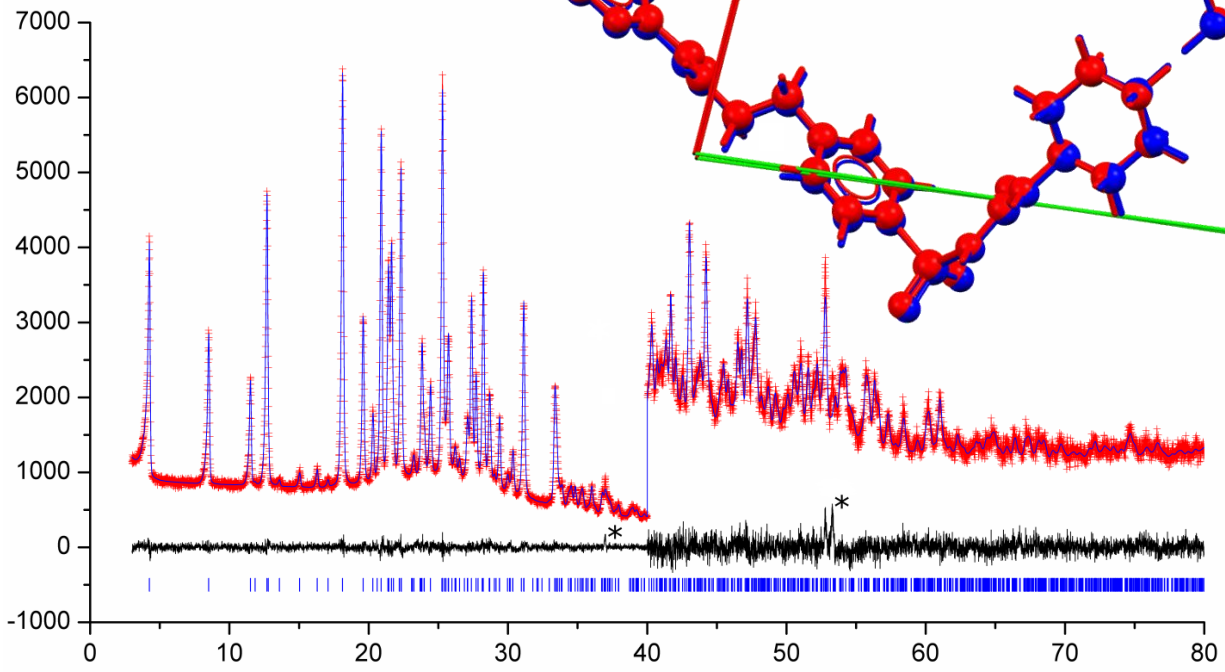


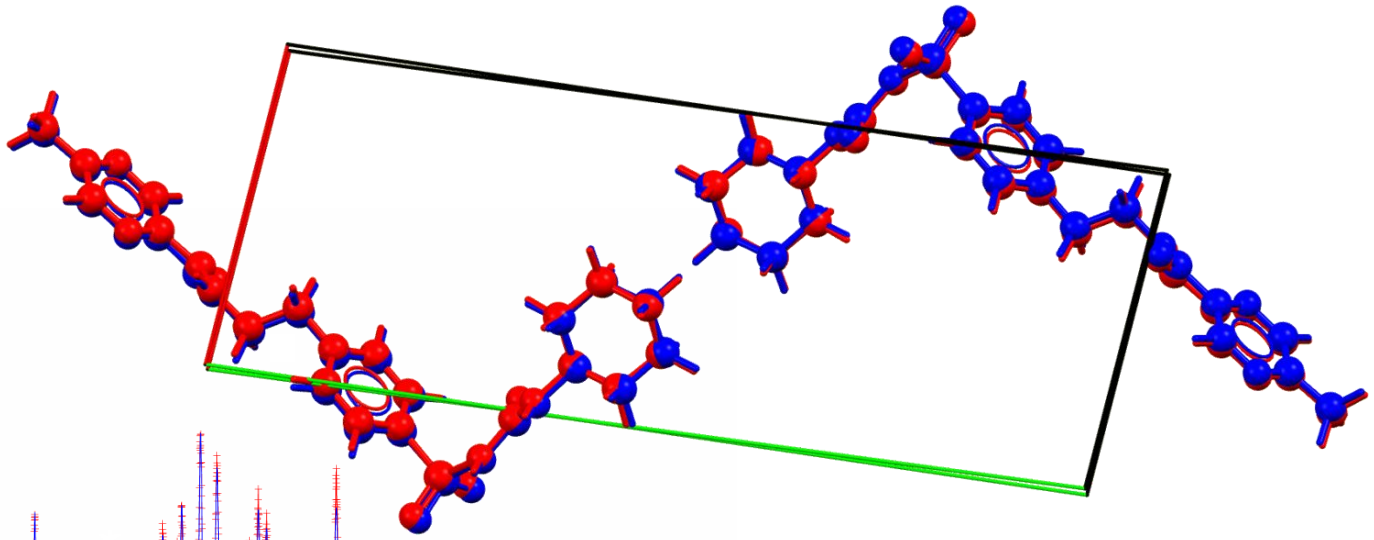
The pyrazyl ring can be turned over 180°:
N and CH: same number of electrons

SAXFED

SAXFED (2005)
RMSCD = 0.13 Å , 3.6 kcal/mol more favourable

Lab, no PO

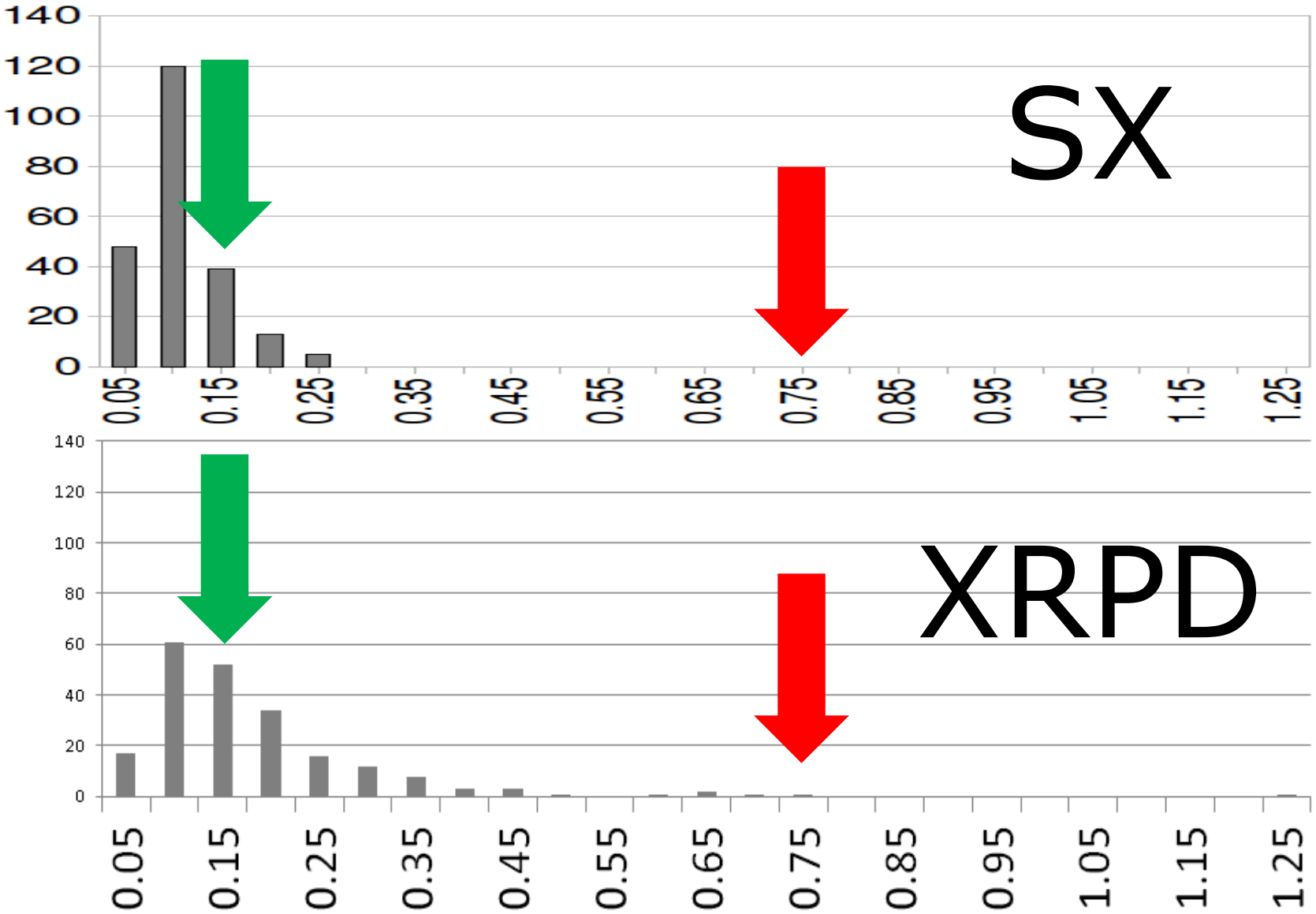




Experiment
Calculated

* = NaCl

RMS Cartesian Displacement



Errors in Molecular Crystal Structures

	SX	PXRD
1. Inaccurate	0	5
2. H-atom	3	5 (+6 missing)
3. Non H-atom	0	9
4. Space group	0	5
	1.3%	8.8%

All errors are minor (C *vs* N, H atom)

Preferred Orientation

Upon closer inspection, preferred orientation (PO) is the greatest source of problems and uncertainties.

PO modifies your experimental data to fit your model.

PO does not just redistribute information, it reduces the information.

Be suspicious when a structure has PO!

If you are prepared to ignore the experimental data in favour of your model, why did you measure the experimental data in the first place?

Precise Structures with DFT-D

Precision with DFT-D

Validating the crystal structure is done *after* the Rietveld refinement: it does not influence the Rietveld process.

This is a pity: the DFT-D contains a lot of independent information, can this information be used *as part of* the Rietveld refinement?

I.e. can the independent information from the DFT-D be merged into the Rietveld refinement to *complement* the experimental data to make the final result more accurate?

Accuracy with DFT-D

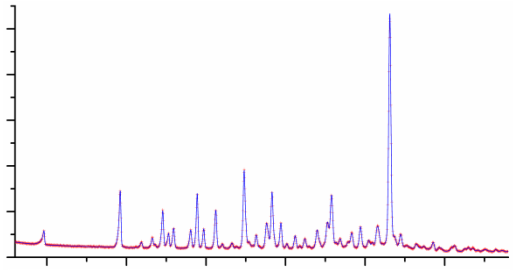
Use the bond lengths and bond angles from the DFT-D minimised crystal structure as restraints:

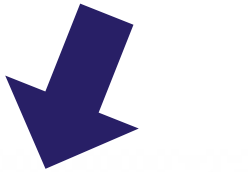
"Polymorph-dependent restraints"

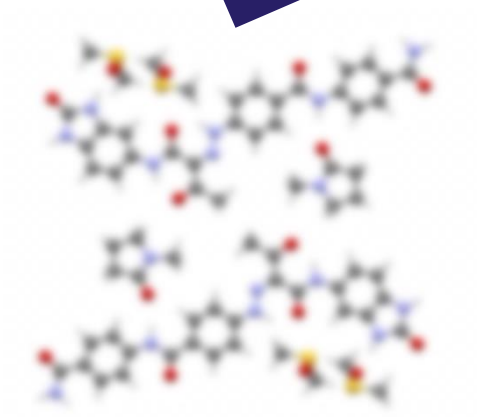
Only *after* the structure has been validated as being correct, otherwise you are biasing your refinement.

XRPD + DFT-D

XRPD

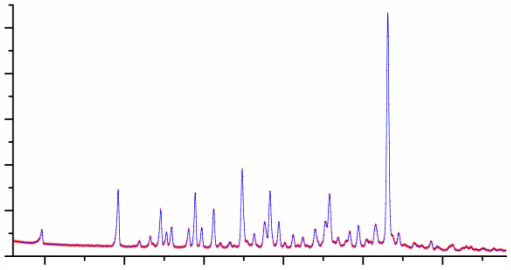




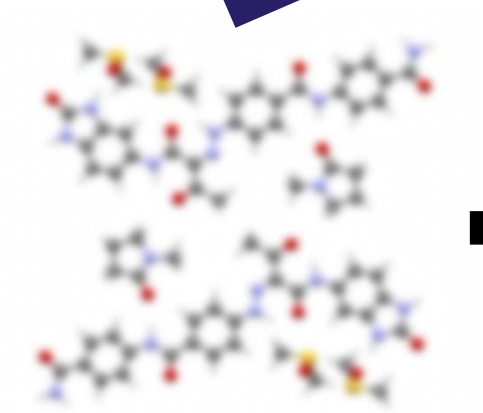


XRPD + DFT-D

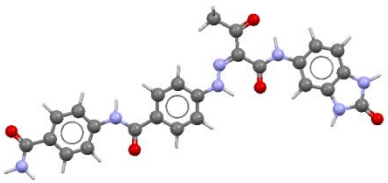
XRPD



DFT-D

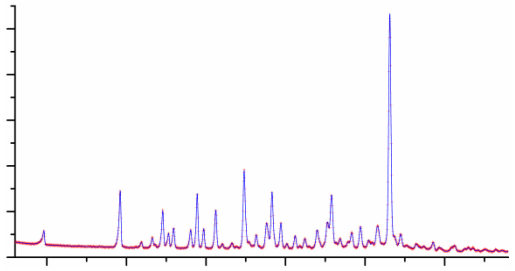


+

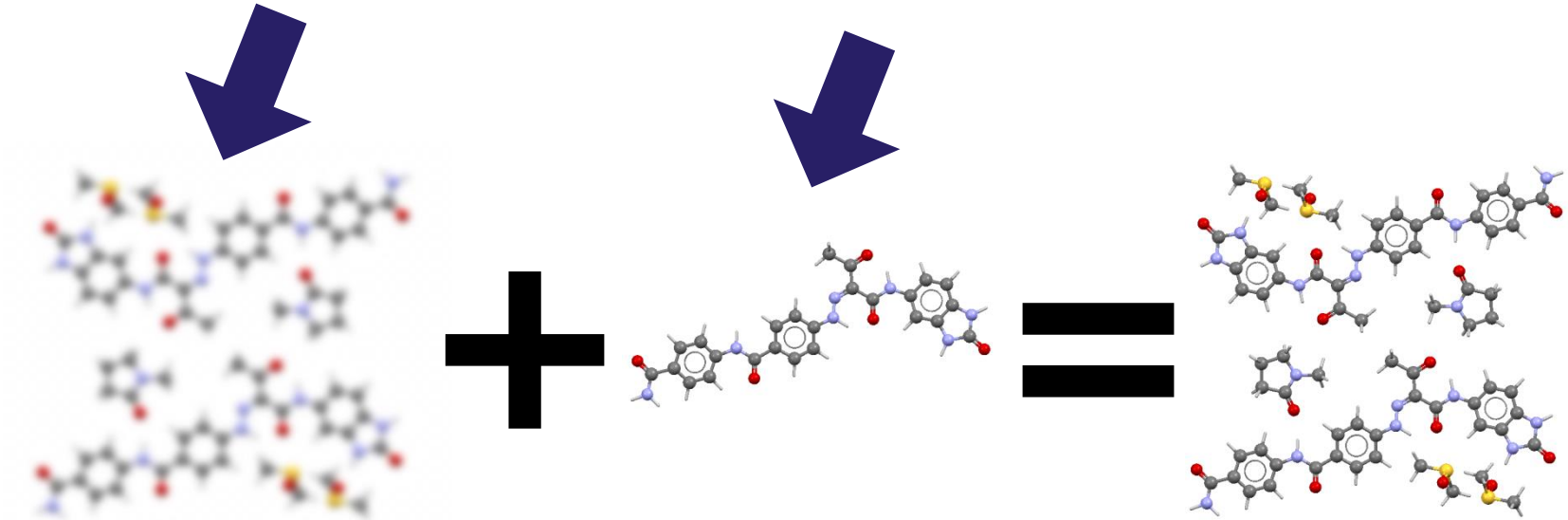


XRPD + DFT-D

XRPD



DFT-D



Accuracy with DFT-D

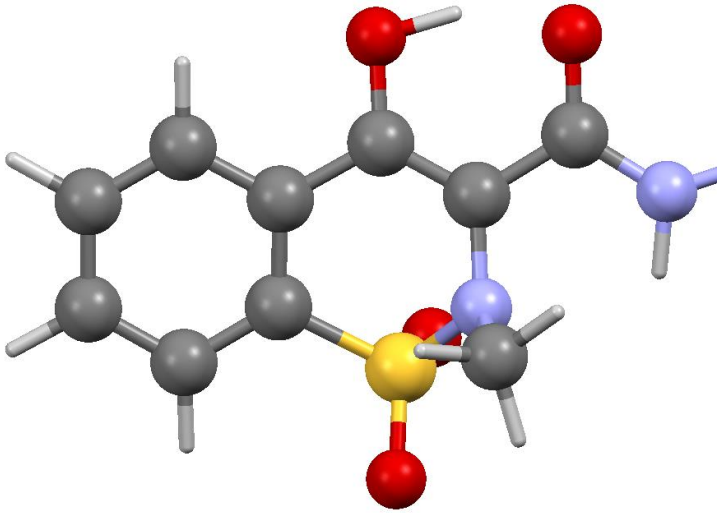
Polymorph-dependent restraints in *TOPAS*:

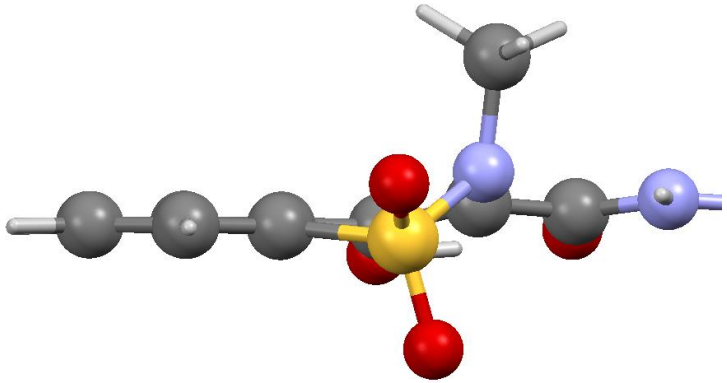
```
Distance_Restrain( N1   C2,  1.47872, 1.47998`_0.00610, 0, 10000 )
Distance_Restrain( N1   C3,  1.47894, 1.48956`_0.00690, 0, 10000 )
Distance_Restrain( N1   C4,  1.48941, 1.48492`_0.00524, 0, 10000 )
Distance_Restrain( C2   C5,  1.50425, 1.47233`_0.00715, 0, 10000 )
Distance_Restrain( C2   H6,  0.95,      0.96054`_0.01471, 0, 10000 )
Distance_Restrain( C2   H7,  0.95,      0.94072`_0.01347, 0, 10000 )
Distance_Restrain( C3   C8,  1.50403, 1.49550`_0.00524, 0, 10000 )
Distance_Restrain( C3   H9,  0.95,      0.95970`_0.01483, 0, 10000 )
…
Angle_Restrain( C2   N1   C3, 111.15614, 115.12083`_0.35599, 1, 1 )
Angle_Restrain( C2   N1   C4, 112.79224, 112.04806`_0.36718, 1, 1 )
Angle_Restrain( C3   N1   C4, 114.20513, 113.81510`_0.39248, 1, 1 )
Angle_Restrain( N1   C2   C5, 112.35920, 113.55737`_0.35977, 1, 1 )
Angle_Restrain( N1   C2   H6, 111.80674, 113.25174`_1.01145, 1, 1 )
…
```

Planarity Restraints

The DFT-D tells you which atoms are in the same plane, so the planarity restraints are also based directly on the DFT-D calculations

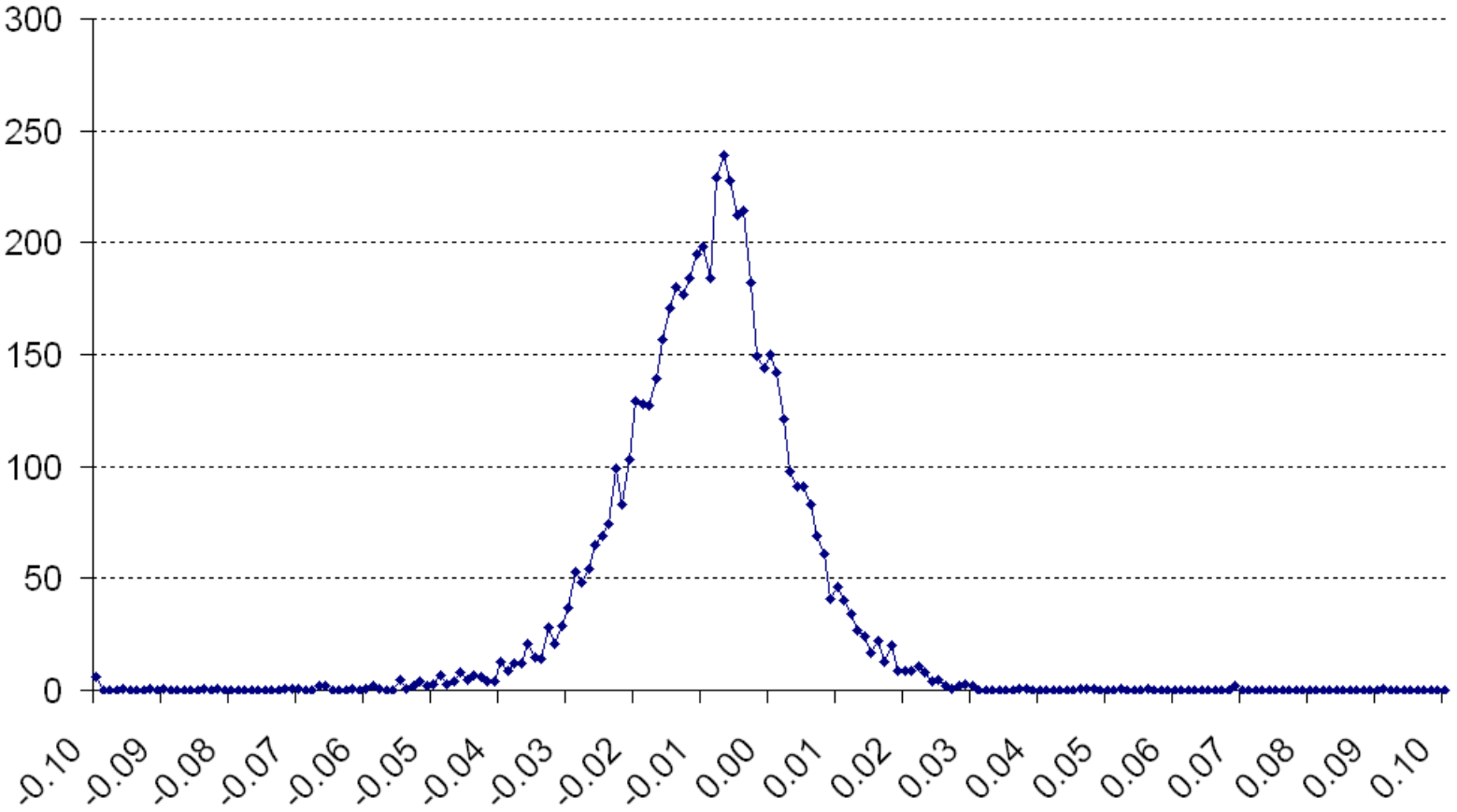
Flatten(C5 C15 H27 C26 O40 C38 H47, , 4.17658429`_5.92244831, 0, 100000)





Accuracy with DFT-D

Average absolute difference over 5,778 bonds from *Acta* E test set: 0.013 Å (non-H atoms only)



Bond length deviations: SX – DFT-D [Å]

Hydrogen Atoms

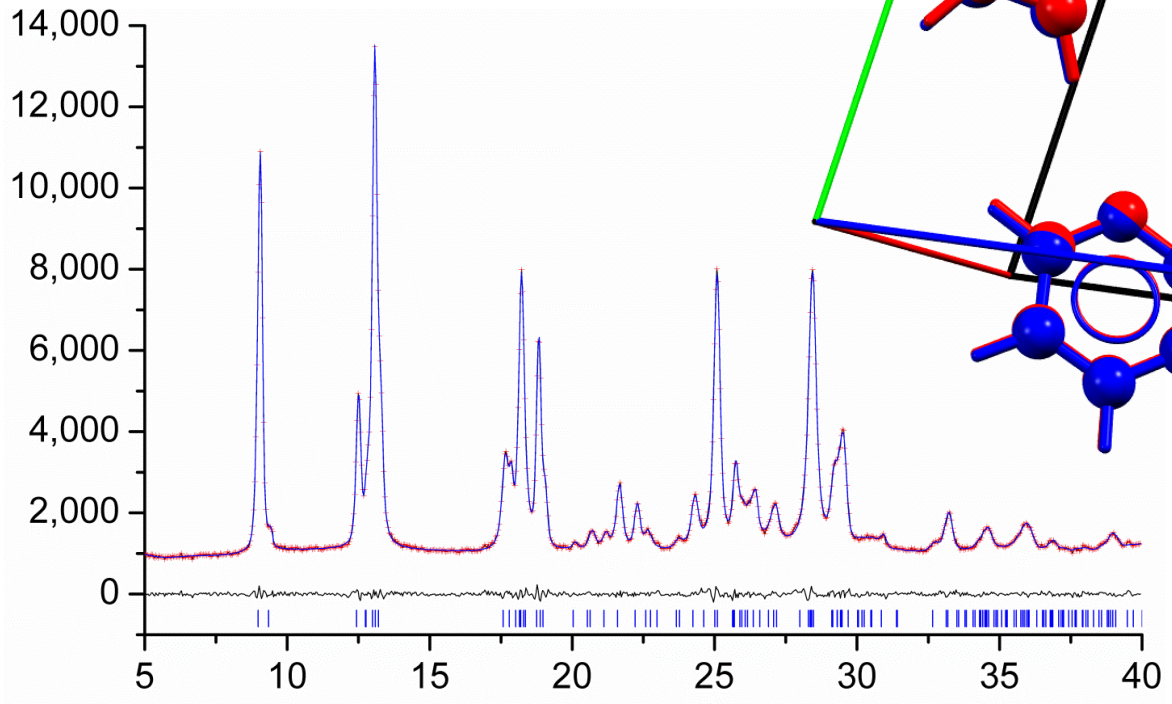
For the hydrogen atoms, restraints are not always sufficient.

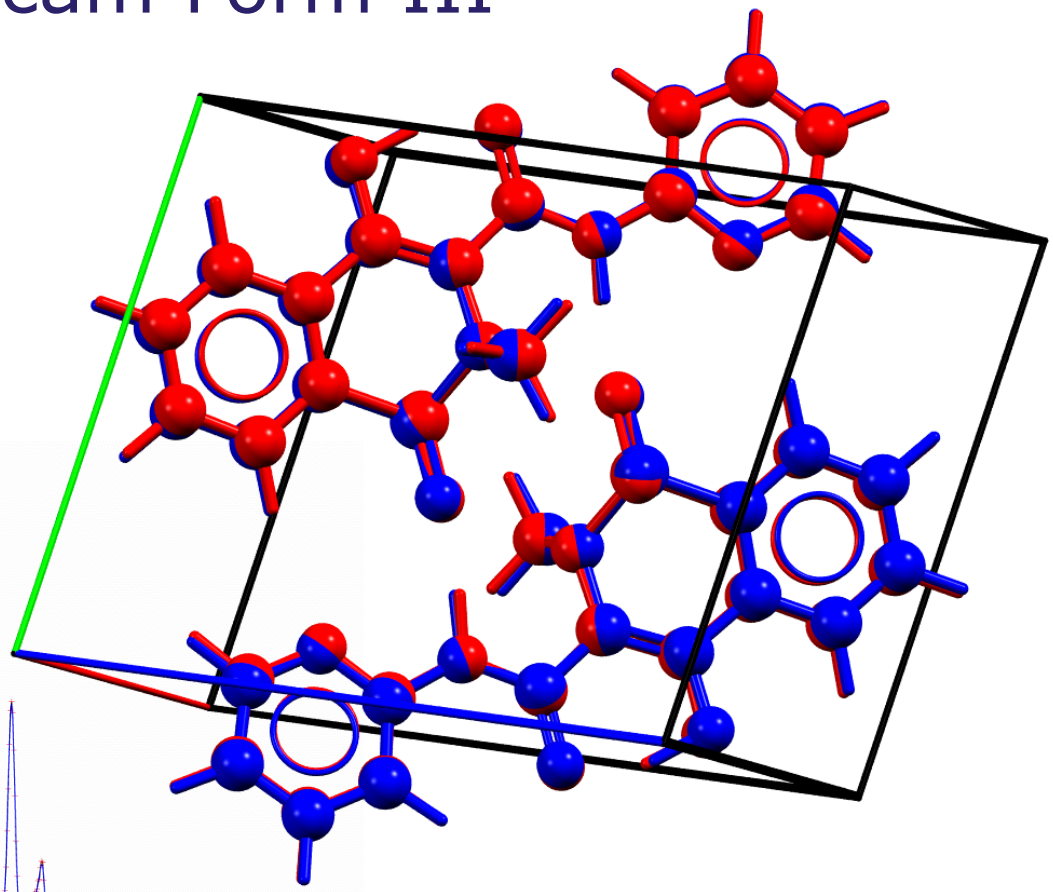
Better solution: energy-minimise hydrogen-atom positions with DFT-D while non-hydrogen atoms and unit cell kept fixed at their experimental values.

Deringer, Hoepfner & Dronskowski (2012) *Cryst. Growth Des*. **12**, 1014–1021

Piroxicam Form III

Lab, no PO



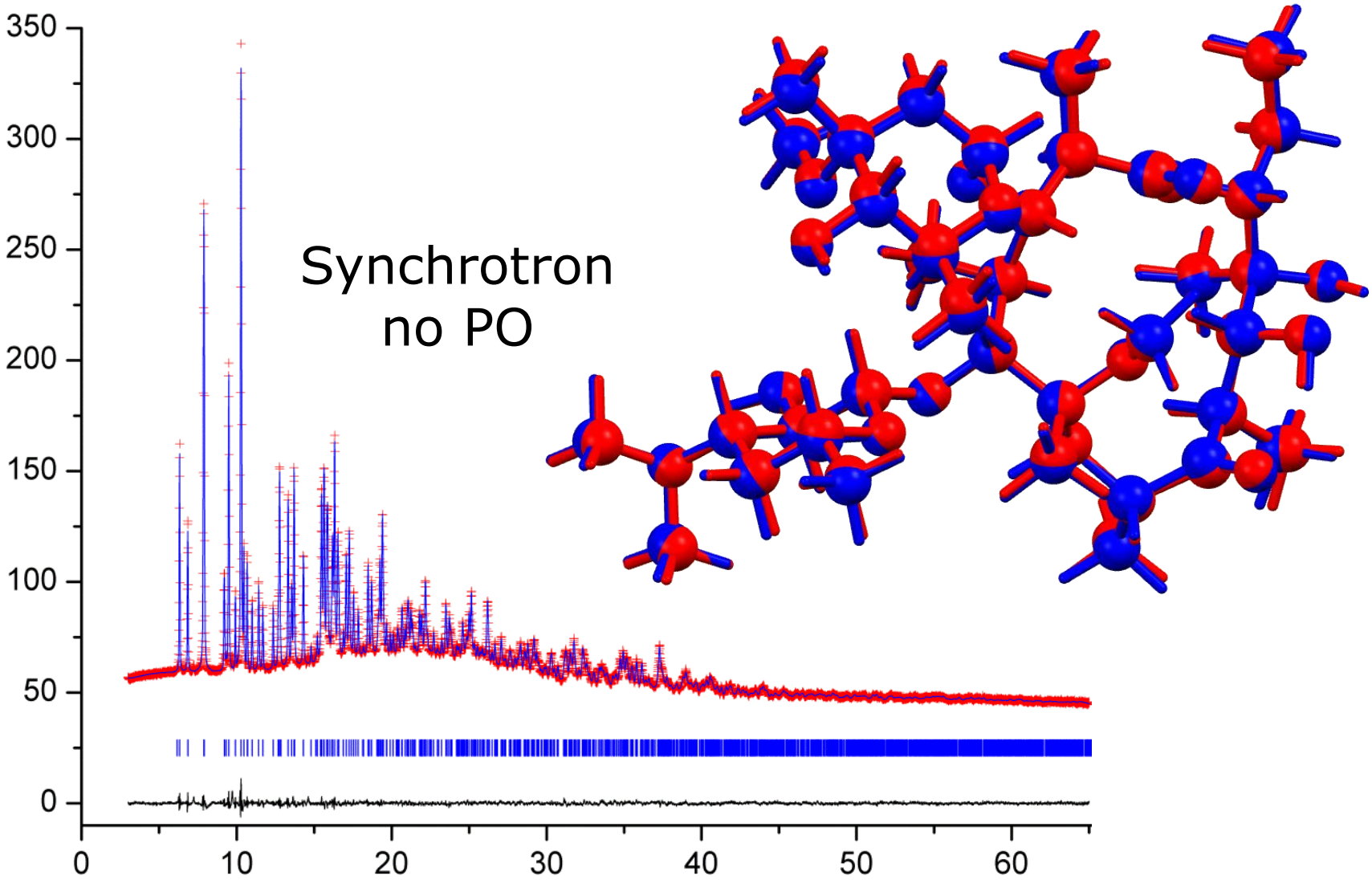


Experiment

Calculated

Naelapää, Van de Streek, Rantanen & Bond (2012). *J. Pharm. Sci.* **101**, 4214-4219

Clarithromycin Trihydrate



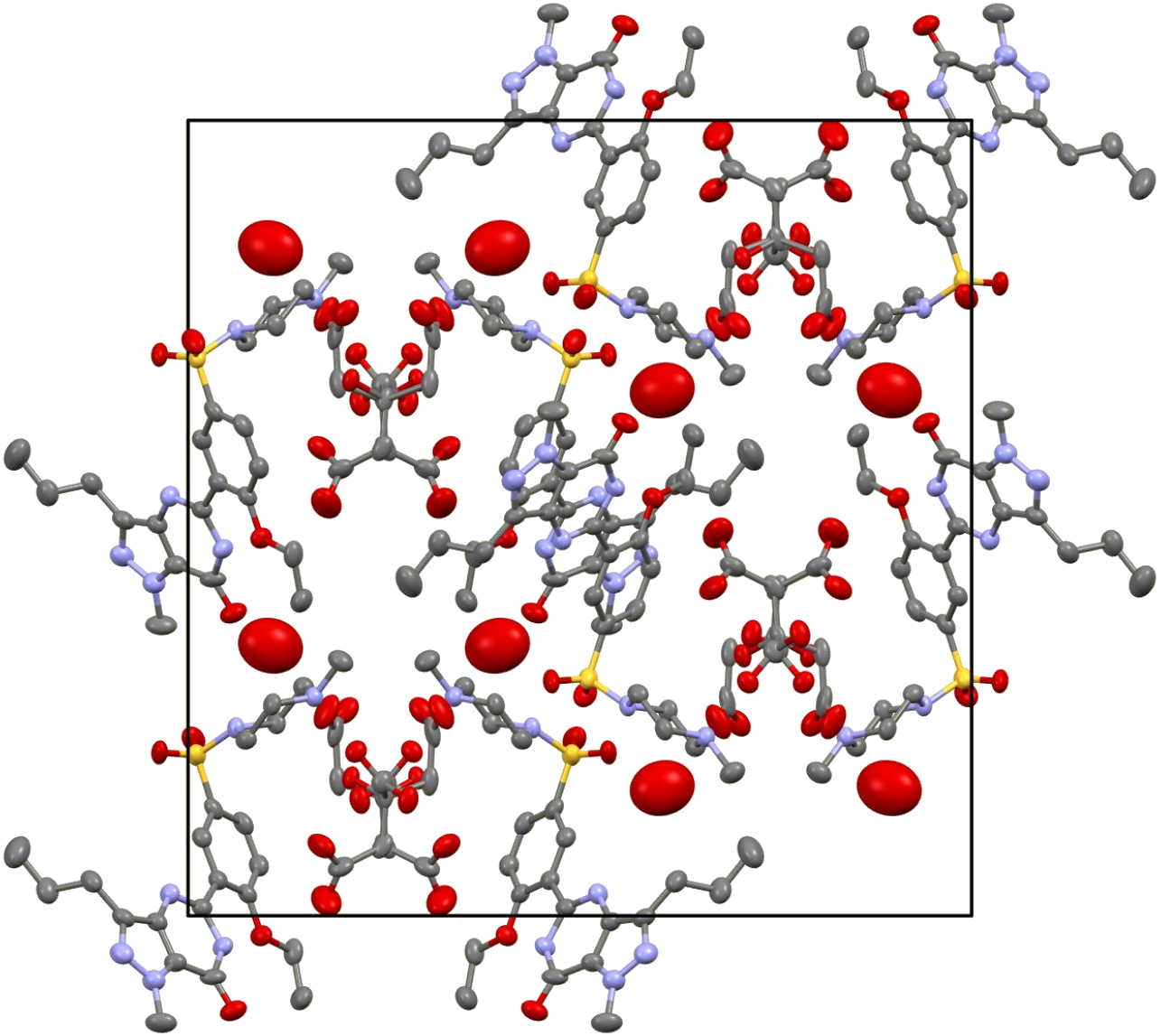
J. van de Streek (2012). *Acta Cryst.* C**68**, o369-o372

Future Directions

Future Directions

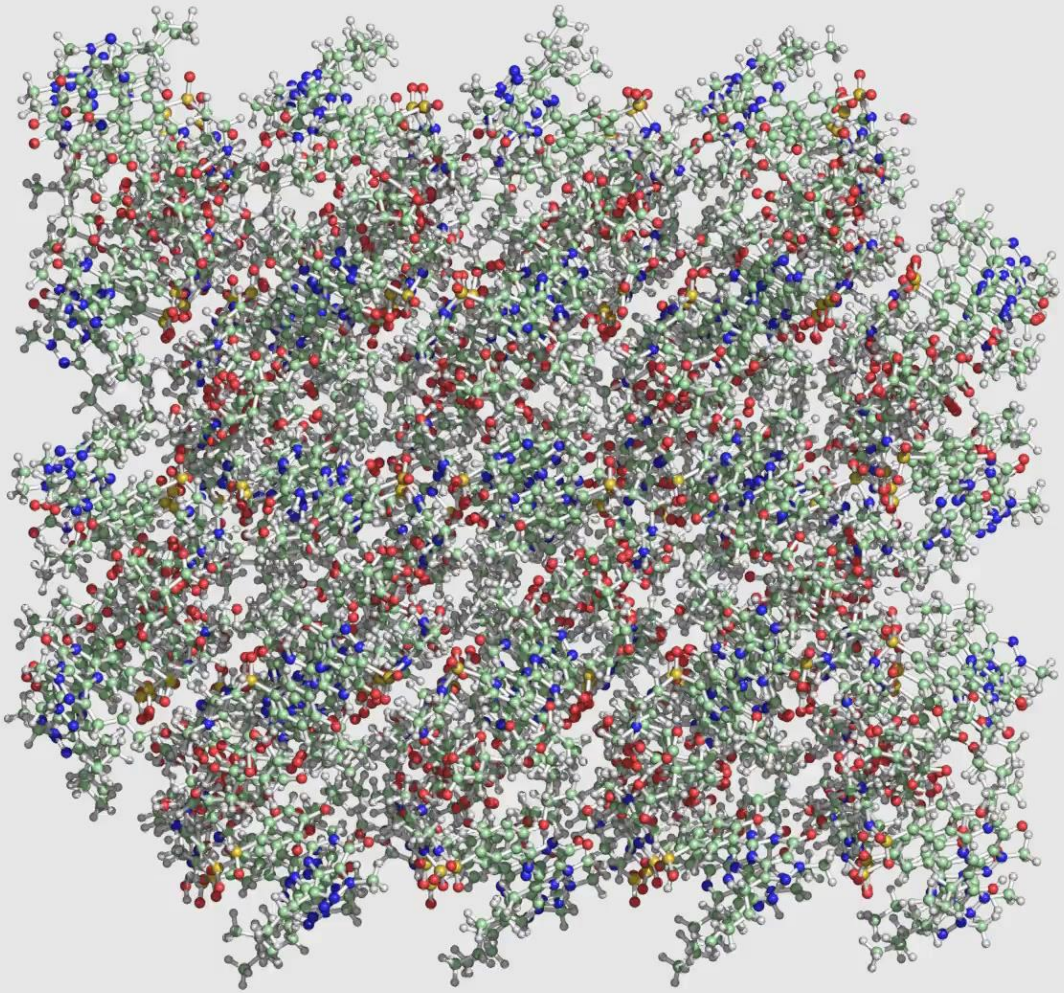
Temperature
Hydrogen atoms
ss-NMR
Space group validation

Temperature

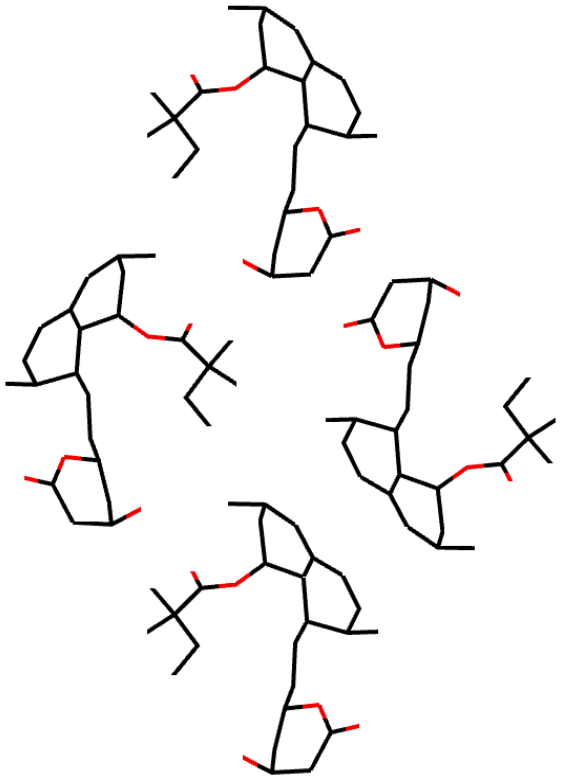


Viagra
T = 173 K
Mobile H_2O

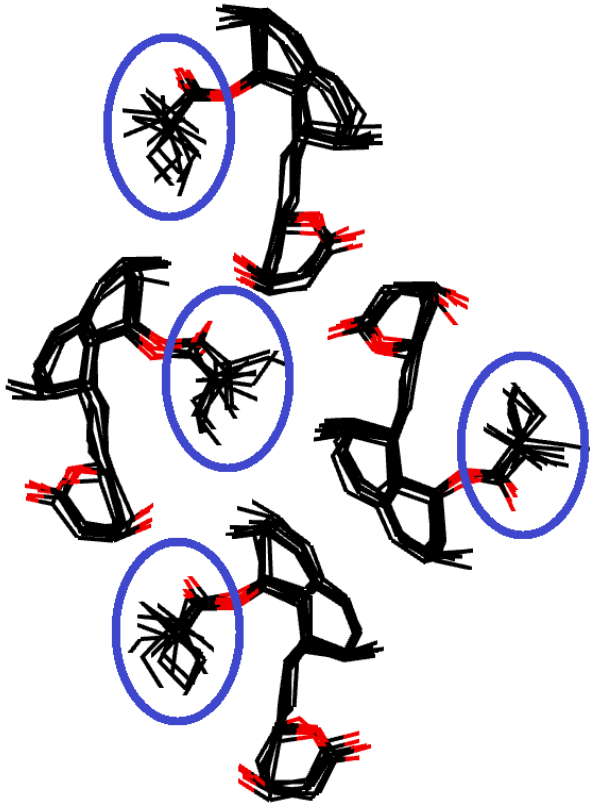
Temperature: Molecular Dynamics



Disorder in Simvastatin



T = 0 K

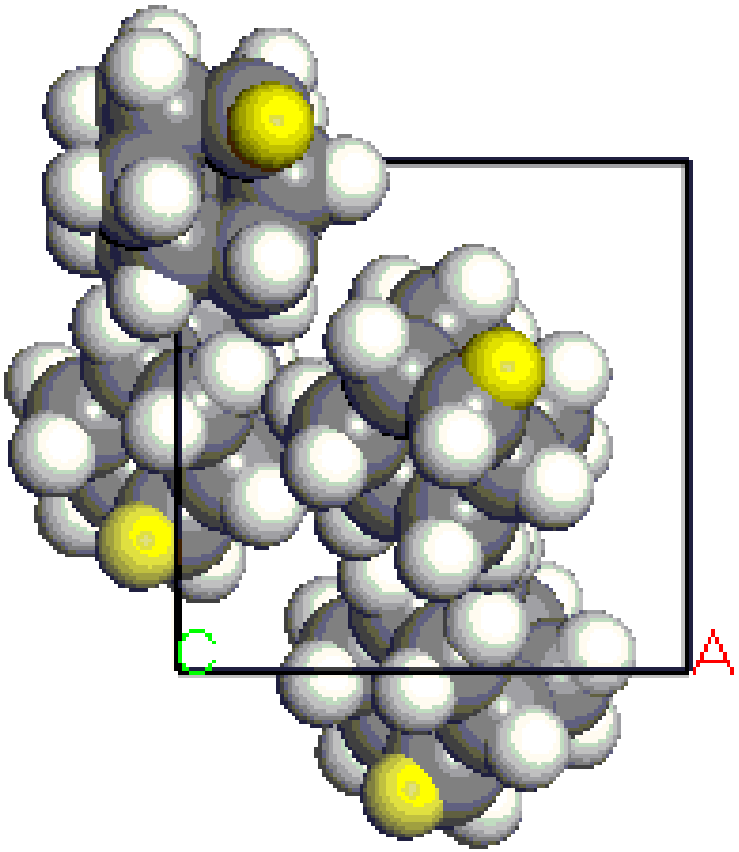


T = 298 K

(from MD)

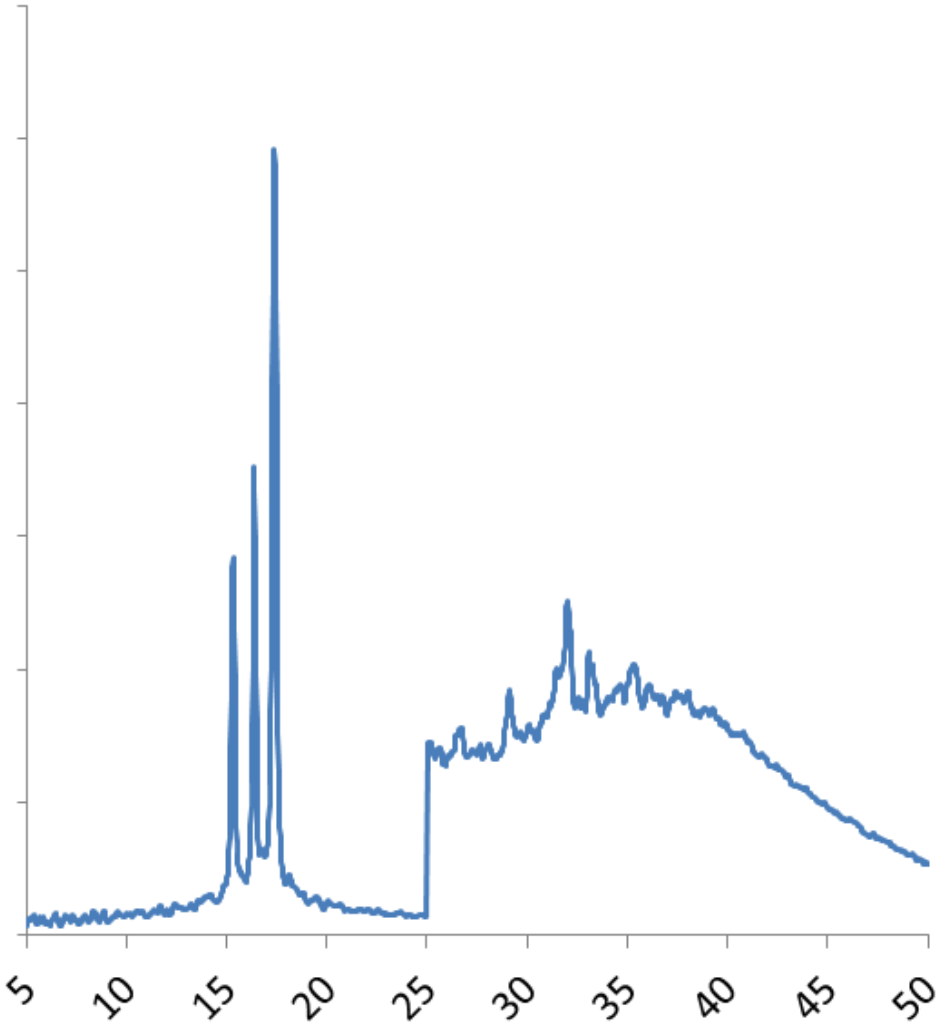
Rotator Phase

Adamantane

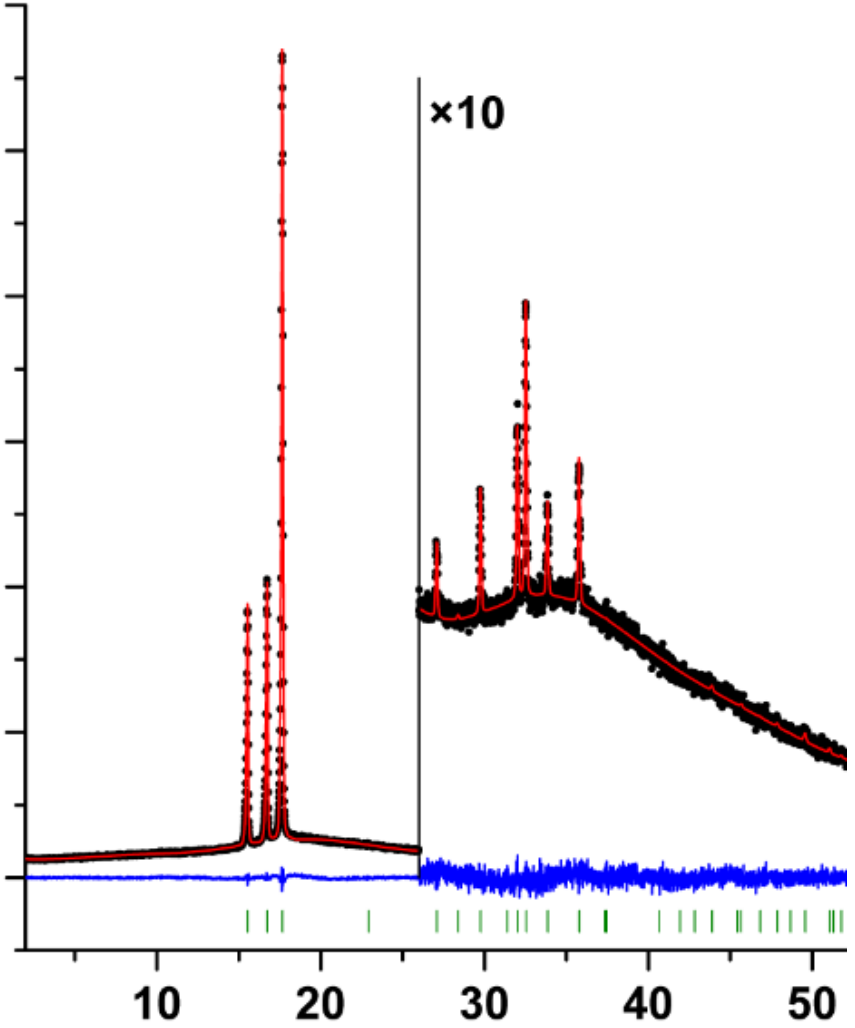


XRPD Rotator Phase Cis-Inositol

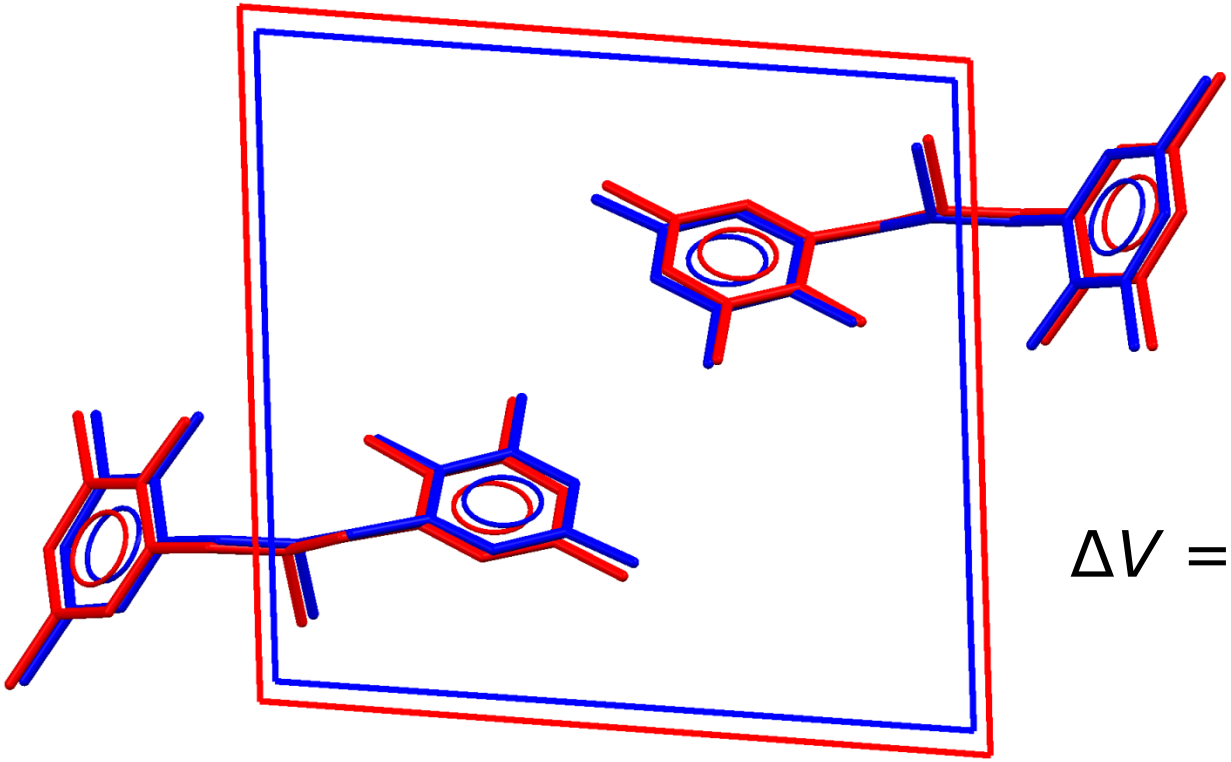
Simulated



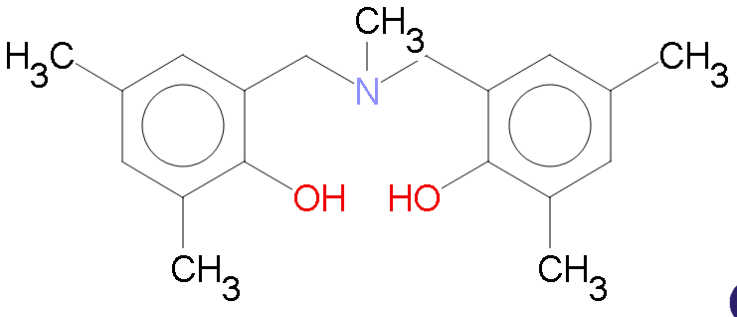
Experiment



Unit-cell Expansion



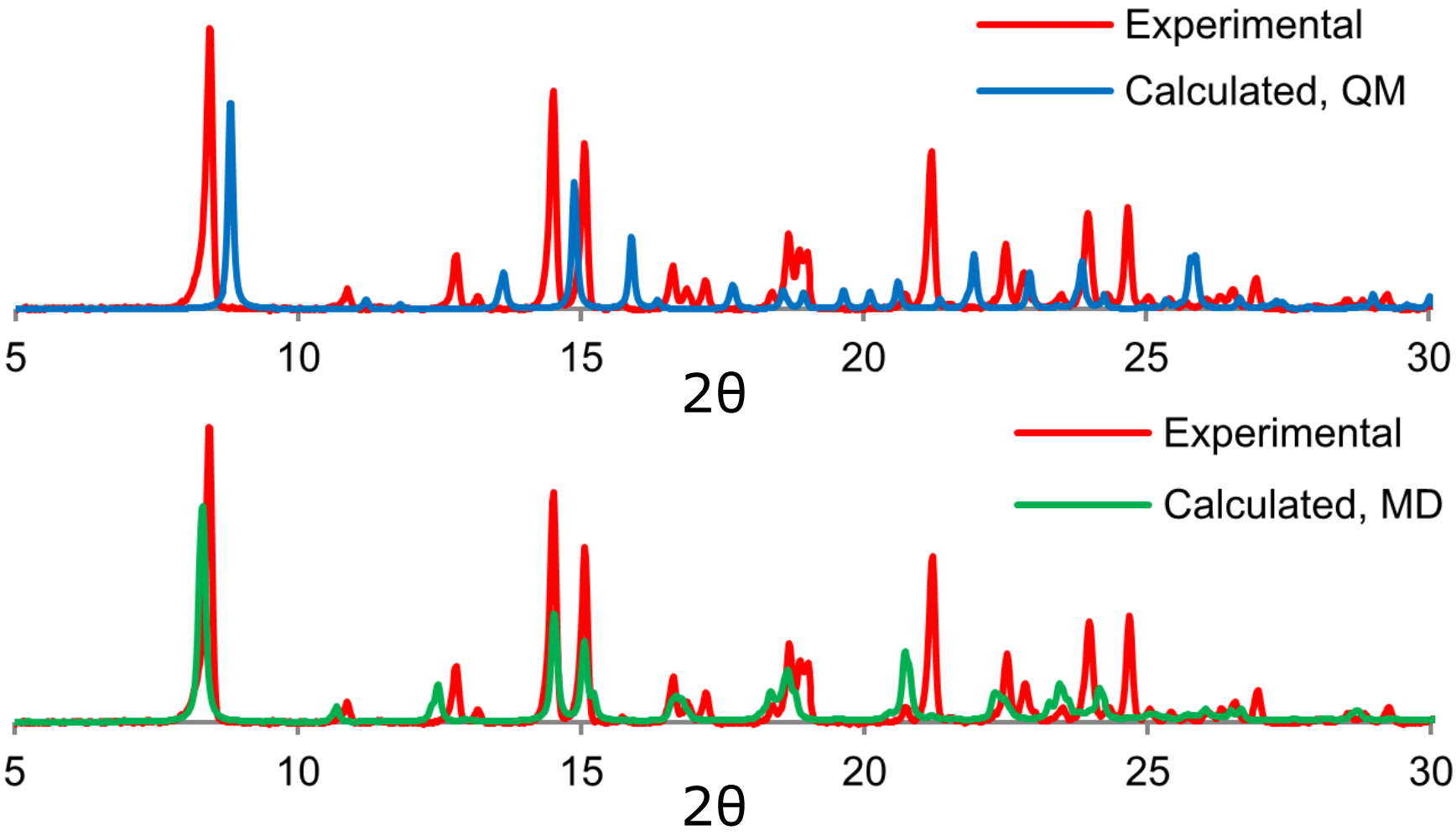
ΔV = -12%



Minimised, T = 0 K

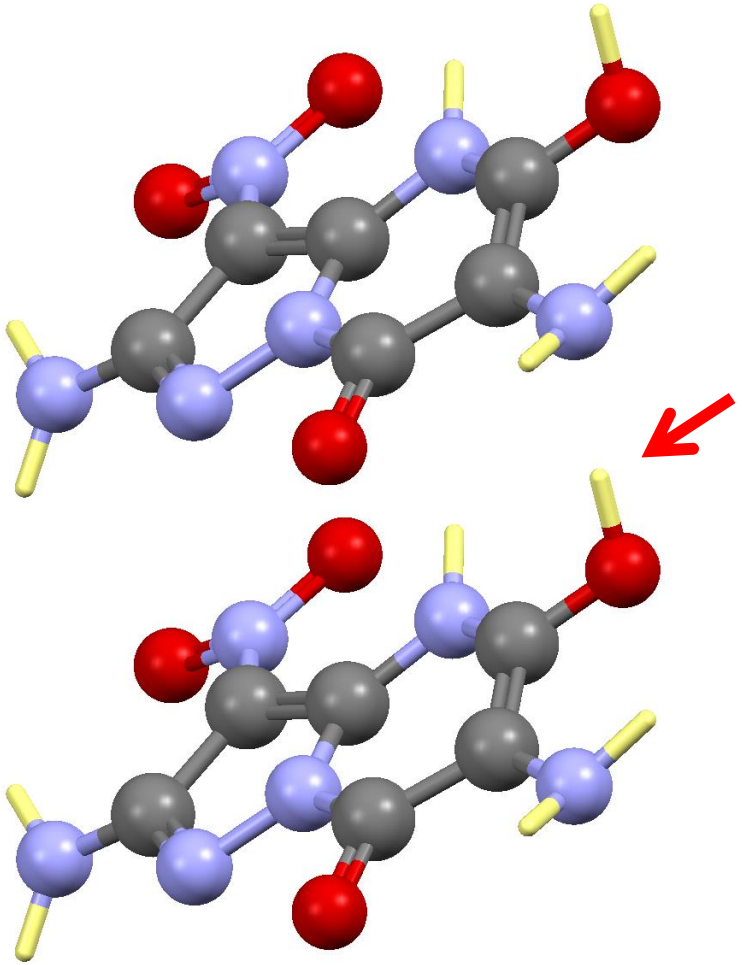
Experiment, T = 298 K

Unit-cell Expansion

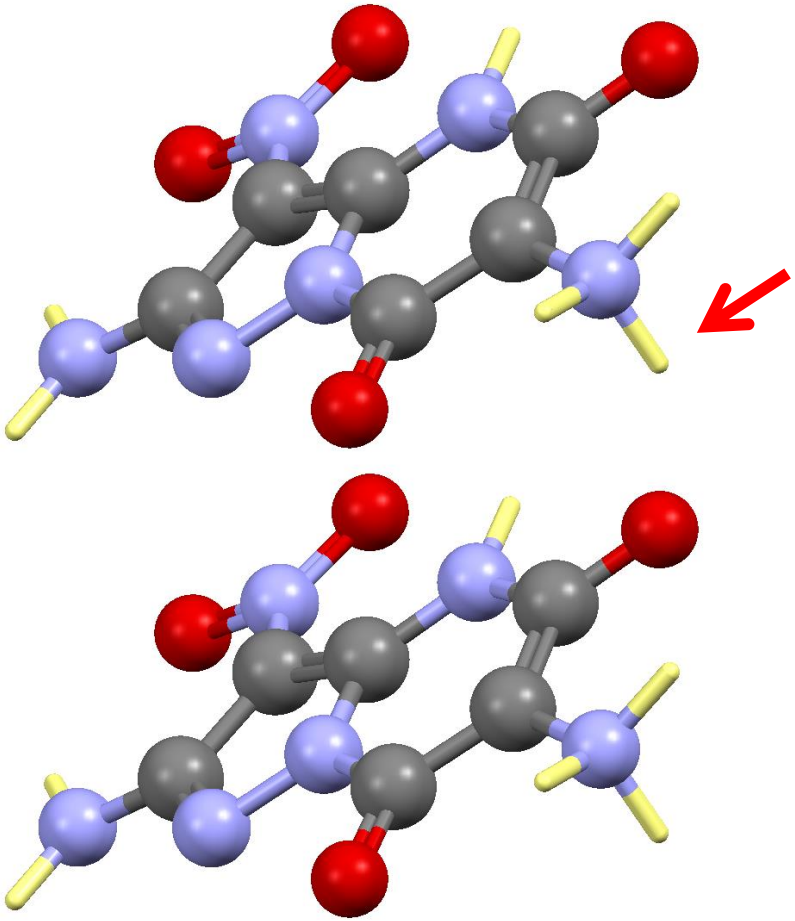


Important to match peak positions of predicted crystal structures to experimental powder patterns

Future Directions: Protons

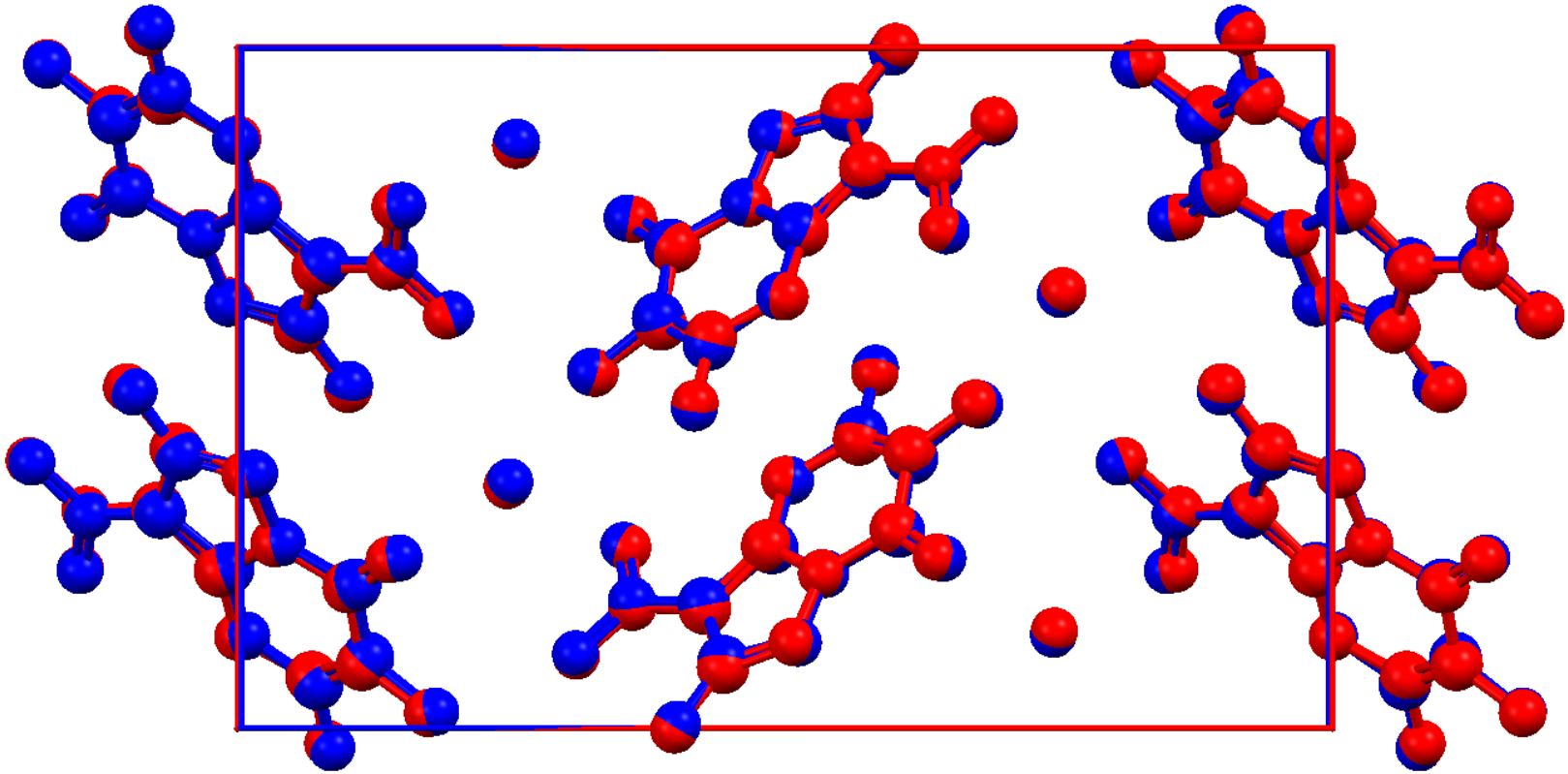


LIPWEM
Neutron data



LIPWEM
DFT-D

Future Directions: Protons



The non-H atoms are not affected:
RMSCD = 0.06 Å

Future Directions: ss-NMR

XRPD weaknesses:
- H atoms
- *Z'*
- Disorder
- PO

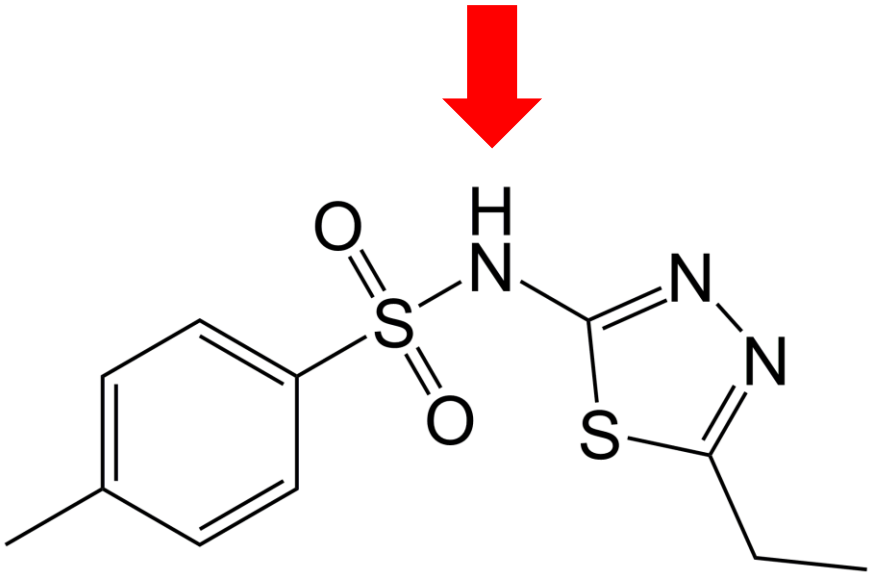
ss-NMR strengths:
- H atoms
- *Z'*
- Disorder
- PO

Future Directions: ss-NMR

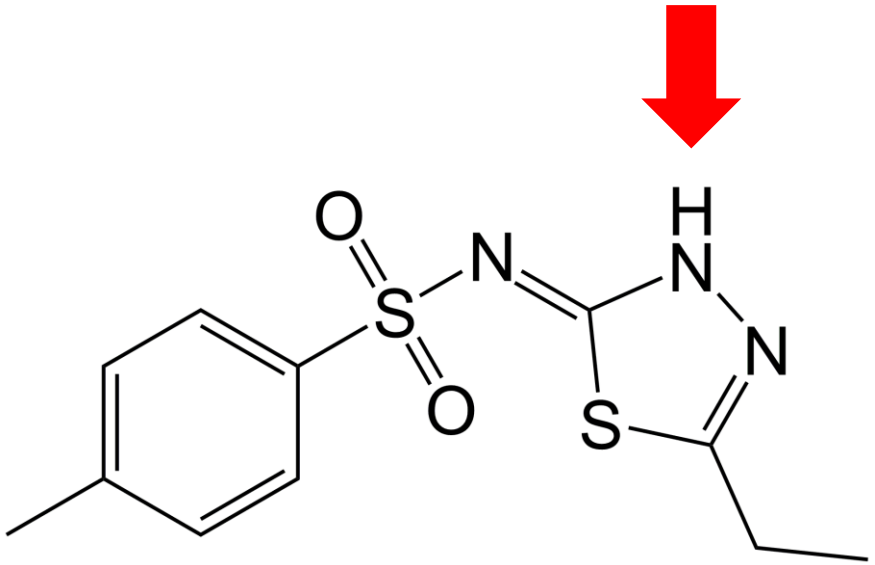
For XRPD, molecular crystal structures, DFT and ss-NMR, see papers by *e.g.*:

Robin K. Harris
Kenneth D. M. Harris
Graeme M. Day
Lyndon Emsley
Chris J. Pickard

UKIRAI (2010)

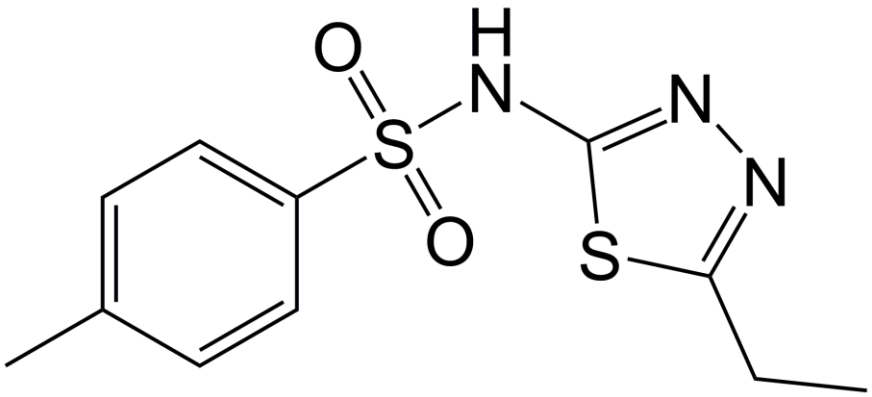


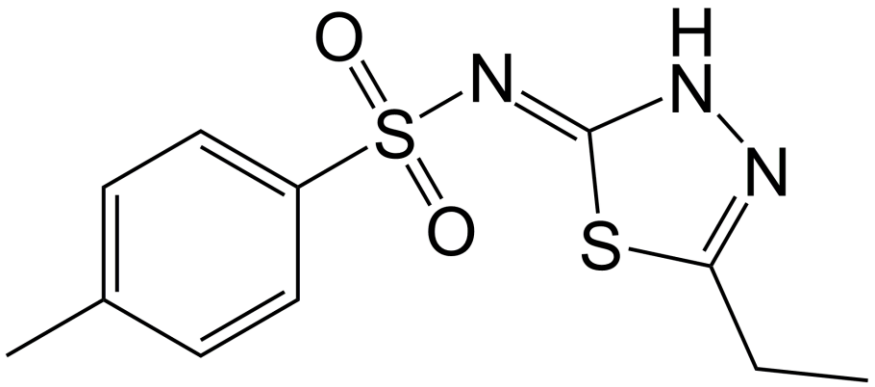
As published
No H-bond
Short N···O contact

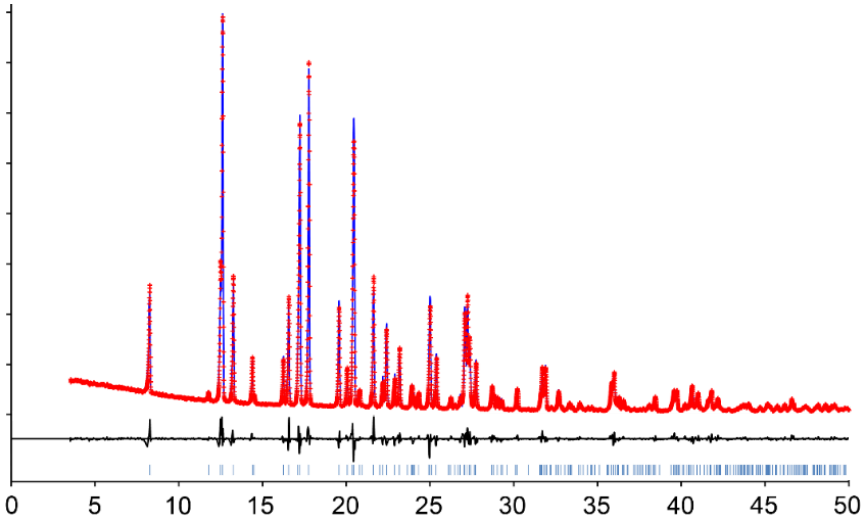


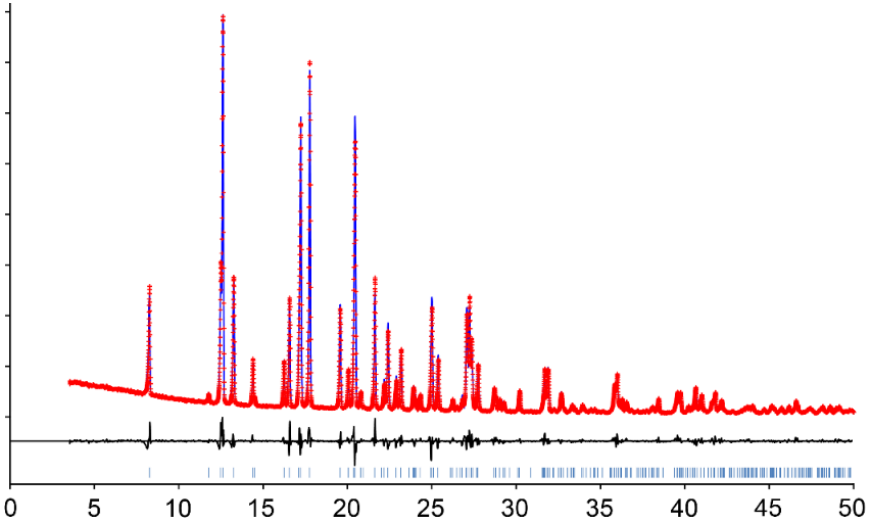
Alternative tautomer:
N–H···O H-bond

UKIRAI: XRPD







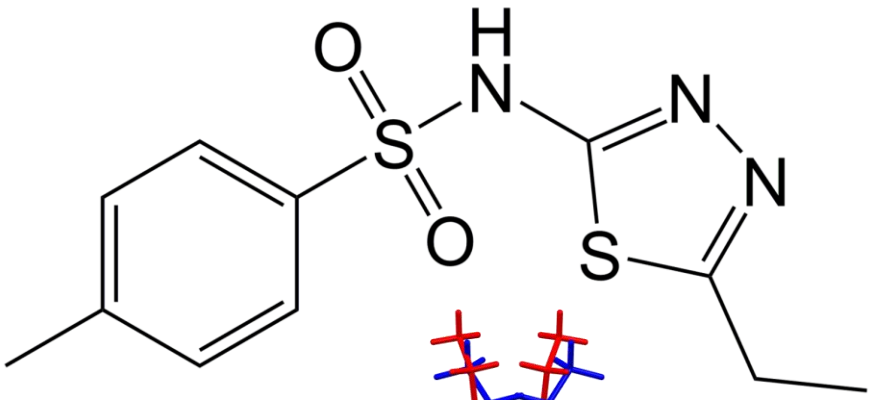


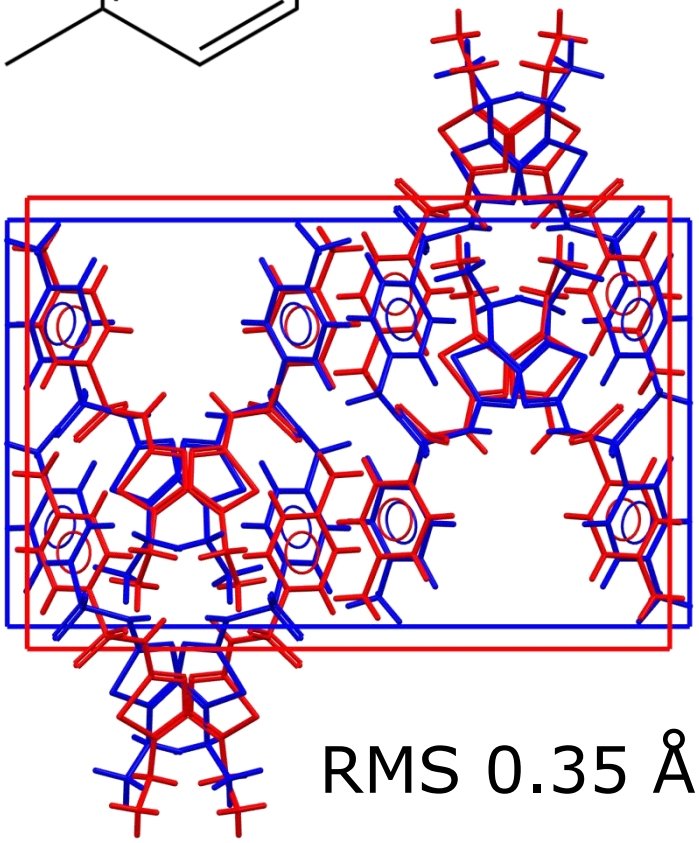
Equal Figures of Merit...

Experiment

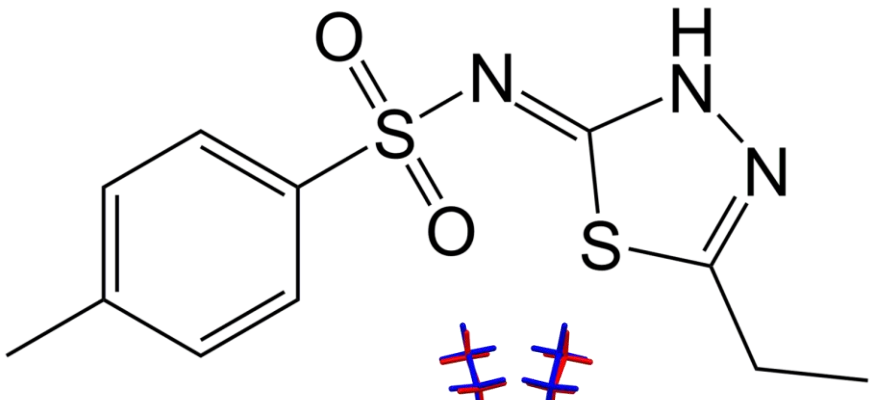
Calculated

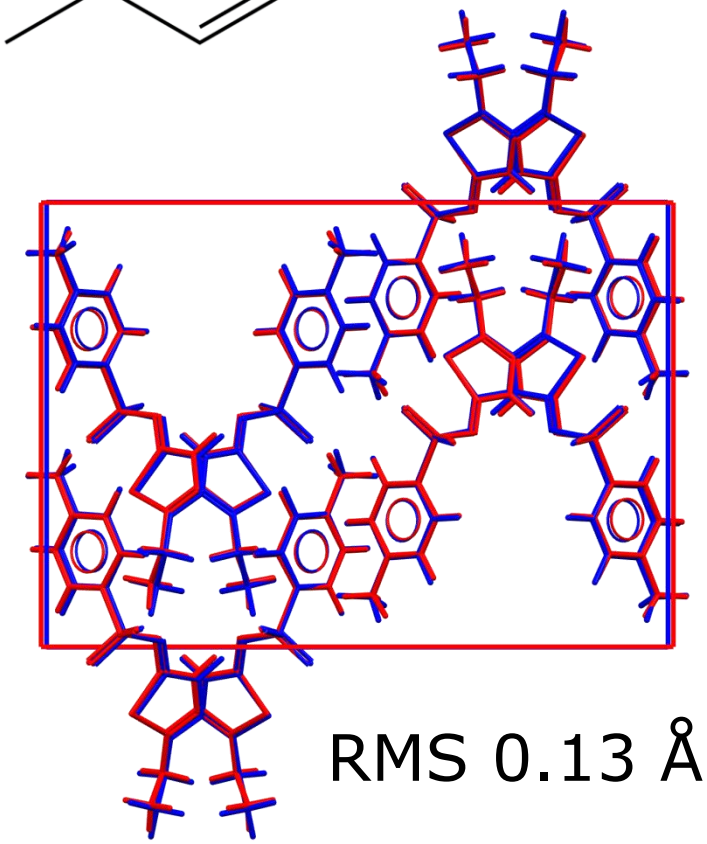
UKIRAI: DFT-D





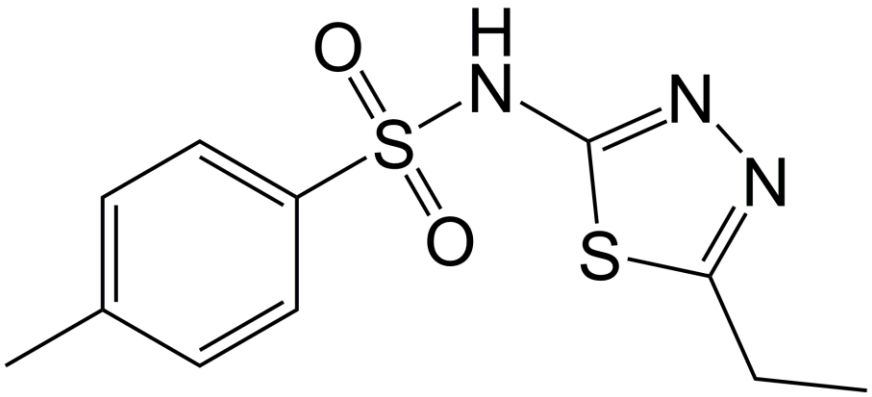
RMS 0.35 Å

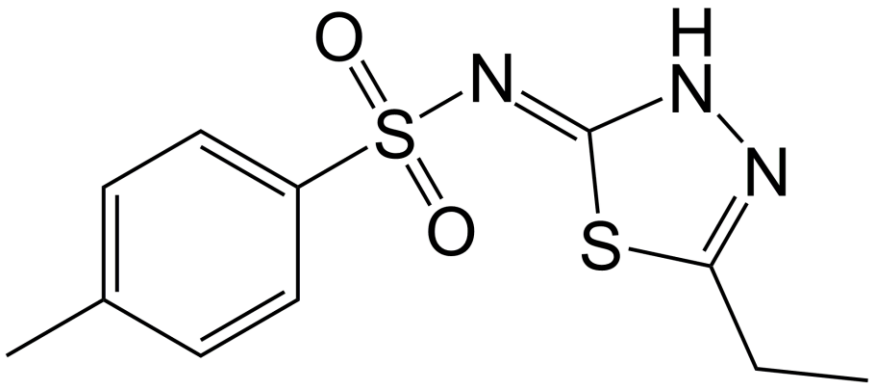




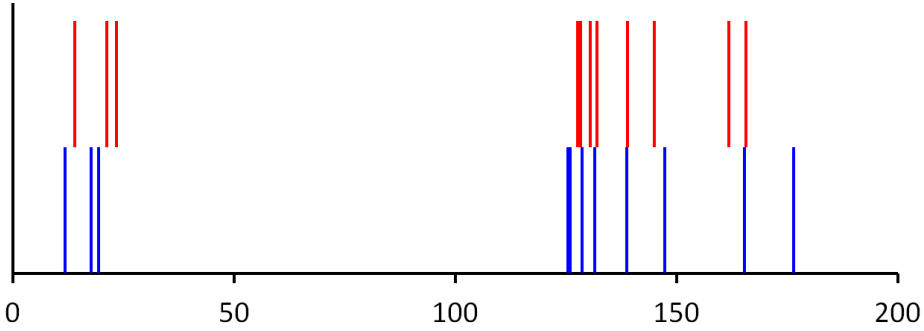
RMS 0.13 Å

UKIRAI: ss-NMR

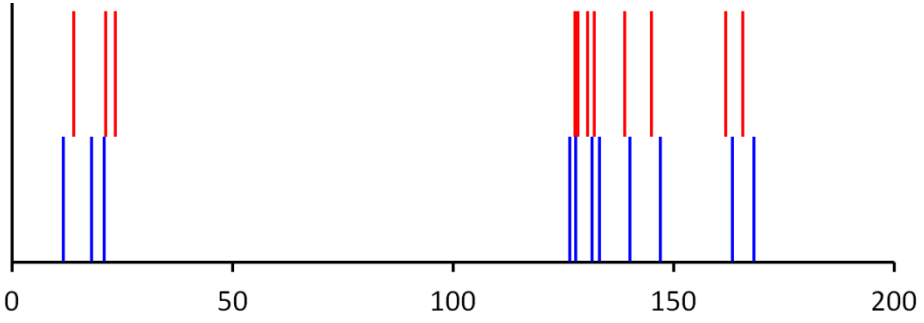




RMSD = 4.1 ppm



RMSD = 1.9 ppm



^{13}C isotropic chemical shift / ppm

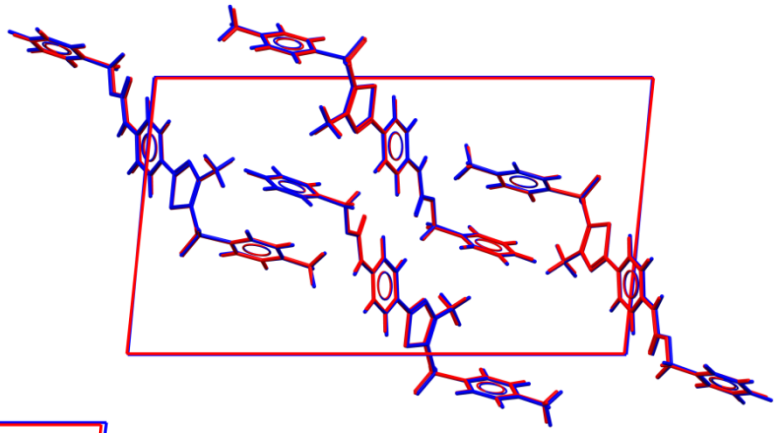
Experiment

Calculated

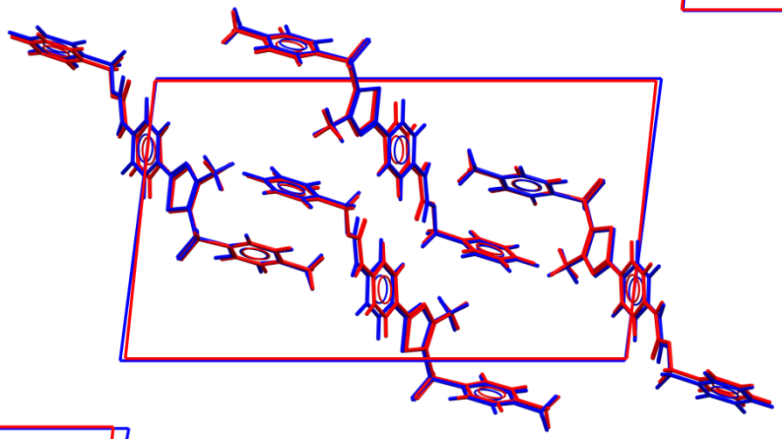
Tailor-Made Force Fields

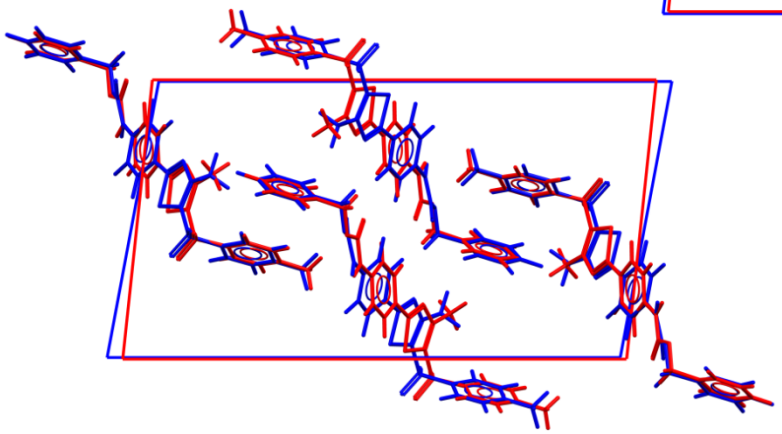
Experiment
Calculated

DFT-D
0.06 Å
100 hrs



TMFF
0.19 Å
10 sec



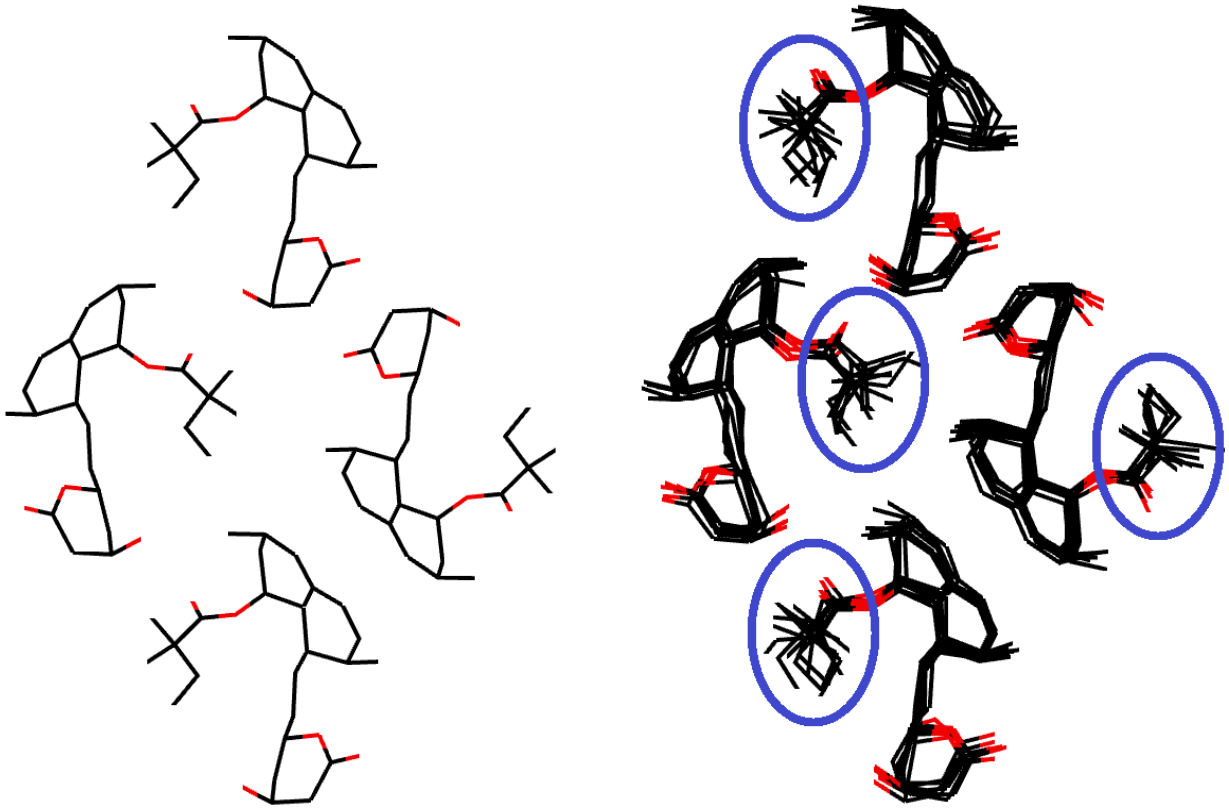


COMPASS
0.33 Å
10 sec

M. A. Neumann (2008) *J. Phys. Chem.* **112**, 9810-9829

MD with Tailor-Made Force Fields

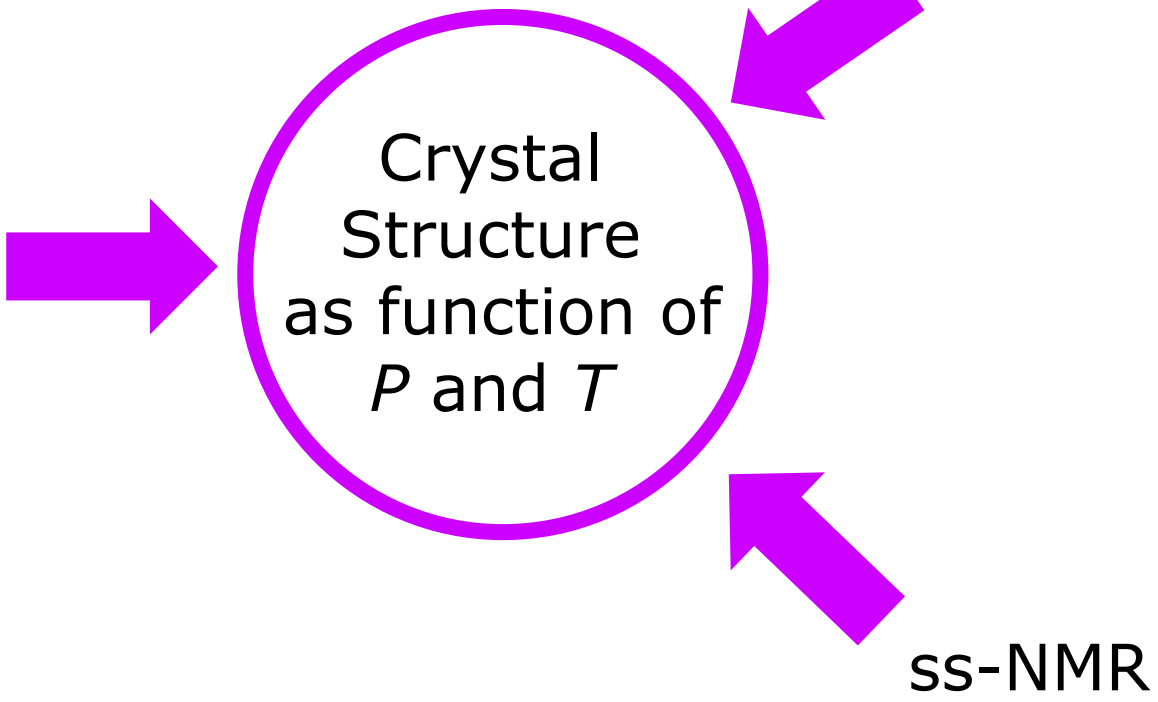
Will allows us to tackle
- disorder (*i.e.* longer time scales)
- amorphous phases (*i.e.* longer length scales)
- ...



Final Aim

DFT-D

1. Initial model
2. Validation
3. Polymorph-dependent restraints
4. ss-NMR
5. TMFF



In the absence of SX data

Conclusions

The combination of XRPD and DFT-D allows for precise and reliable molecular crystal structures.

XRPD structures are less reliable and less precise than SX structures. Missed space-group symmetry.

Preferred orientation is a major problem.

ss-NMR + DFT-D complements XRPD very well.

MD with tailor-made force fields will allow us to tackle dynamic and non-periodic structures.

Acknowledgements

Bill David
Ken Shankland
Martin U. Schmidt
X-Ray Laboratory Dr. Ermrich
And many others...

Marcus A. Neumann
Kristoffer E. Johansson
Xiaozhou Li
Anders S. Larsen

VILLUM FONDEN

THE LUNDBECK FOUNDATION

UNIVERSITY OF COPENHAGEN