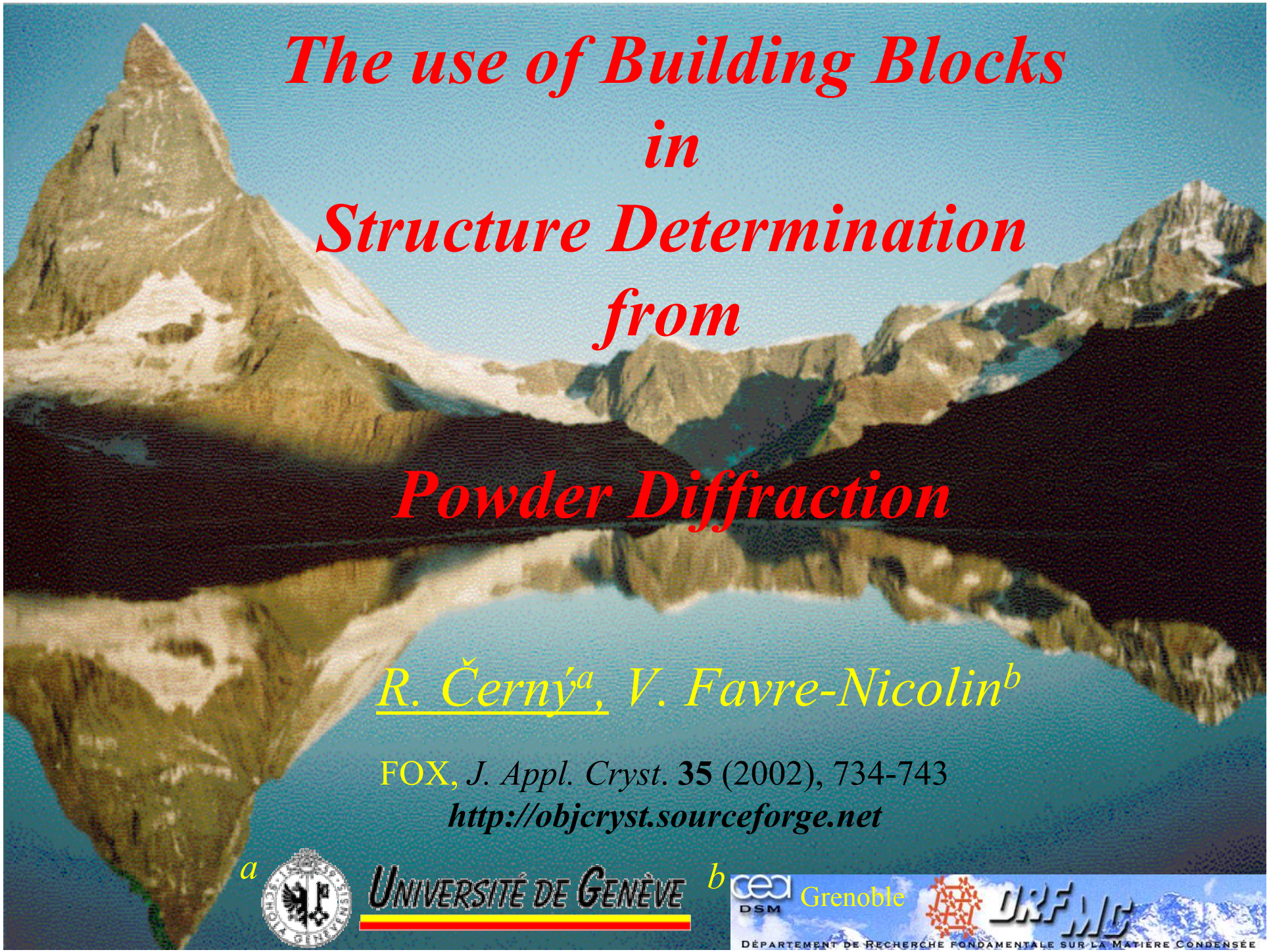


The use of Building Blocks in Structure Determination from Powder Diffraction

R. Černý[a], V. Favre-Nicolin[b]

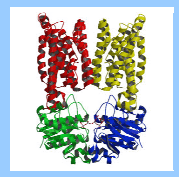
FOX, *J. Appl. Cryst.* **35** (2002), 734-743
http://objcryst.sourceforge.net

a UNIVERSITÉ DE GENÈVE *b* CEA DSM Grenoble DRFMC
DÉPARTEMENT DE RECHERCHE FONDAMENTALE SUR LA MATIÈRE CONDENSÉE

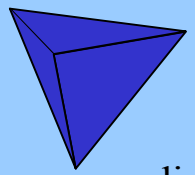
Building Blocks

Well defined sub-units of the crystal

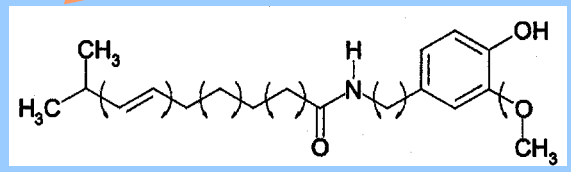
isolated atoms



chains of amino-acids



coordination polyhedra



molecules, fragments

Advantages:

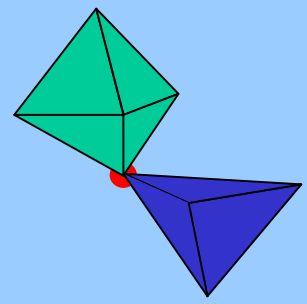
- Lower number of free parameters :
  ($3 \times 5 = 15$ against $3 + 3 = 6$) for a tetrahedron
- Direct crystal-chemistry interpretation allowing *a priori* definition of blocks

Benefits:

- Meaningless structural configurations cannot occur
- Range of convergence of the correct model is larger
- Hydrogen atoms can be included
- Active use of the chemical information and intuition

Problems:

- Description of blocks
- Connectivity between the blocks



- Structural solution is biased

Refinement:

(Semi -) Rigid Bodies

- Common tool in single crystal diffraction

  C. Scheringer, *Acta Cryst*. **16** (1963) 546-550

- Distance and orientation constraints and restraints

  D. Watkin, *Acta Cryst*. **A50** (1994) 411-437

- **TLS** thermal vibration model (**T**ranslational and **L**ibrational harmonic motion + **S**crew tensor)

- Becoming un-avoidable in powder diffraction

  R. Dinnebier, *Powder Diffraction* **14** (1999) 84-92

- Distance least-squares refinement

  Ch. Baerlocher, XRS-82. *The X-ray Rietveld system*, 1982, ETHZ Zürich

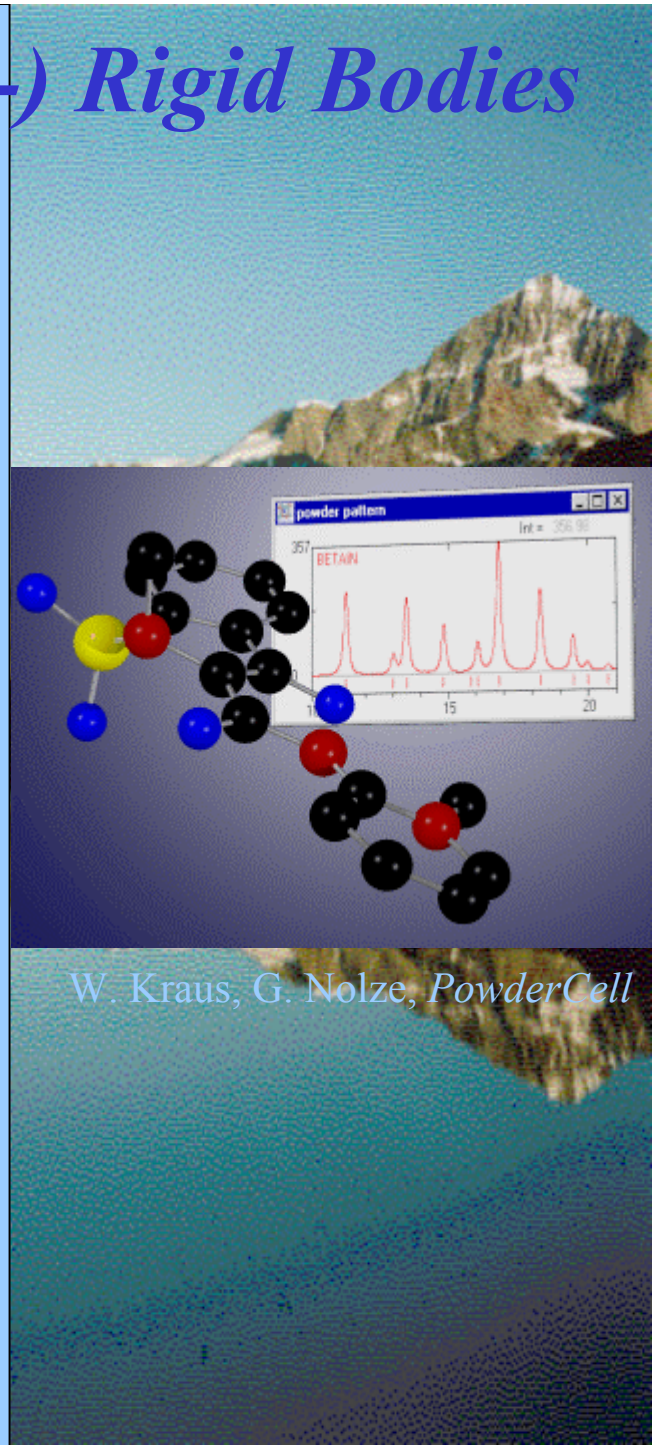
- *Ab-initio* least-squares refinement

  R. Bianchi *et al.*, Acta Cryst. A37 (1981) 65-71 … RISCON

  N. Masciocchi *et al.*, *J. Appl. Cryst*. 27 (1994) 426-429 … P-RISCON

- Modeling against small-angle scattering data

  D.I. Svergun, *Acta Cryst*. **A50** (1994) 391-402



(Semi -) Rigid Bodies

Analysis of Patterson maps:

- Locating vector sets of known fragments.

  C.E. Nordman, K. Nakatsu, *J. Am. Chem*. Soc. **85** (1963) 353-354 … DIRDIF (ORIENT)

- Convolution molecule method.

  W. Hoppe, *Z. Elektrochem*. **61** (1957) 1076-1083
  R. Huber, *Acta Cryst*. **19** (1965) 353-356

- Rotational and translation search … PATSEE.

  E. Egert, *Acta Cryst*. **A39** (1983) 936-940

Analysis of E-maps:

- Interpretation of E-maps by locating known fragments

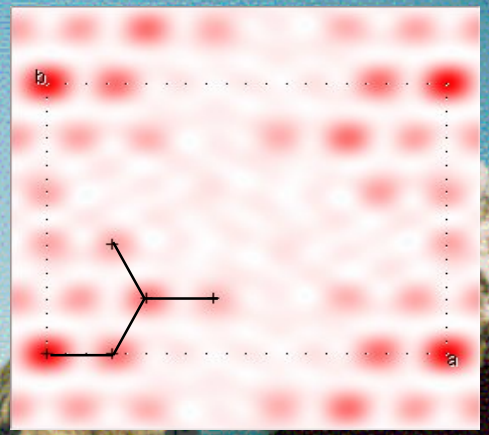
  F. Pavelčík *et al*., *Acta Cryst.* **D58** (2002) 275-283

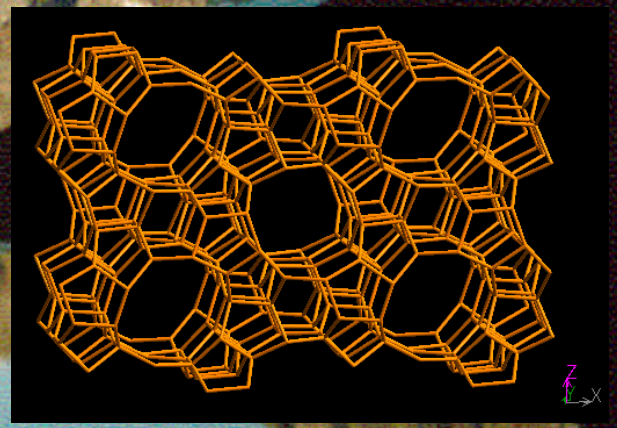
- Search for fragments of known framework topology, zeolites … FOCUS.

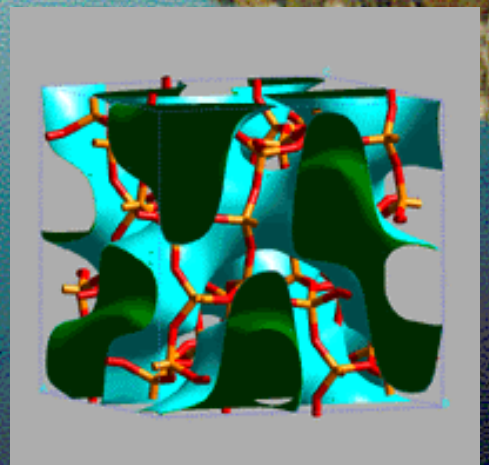
  R.W. Grosse-Kunstleve *et al*., *J. Appl. Cryst.* **30** (1997) 985-995

- Structure envelopes

  S. Brenner *et al*., *J. Appl. Cryst*. **30** (1997) 1167-1172







(Semi -) Rigid Bodies

Powder patterns decomposition: A. Altomare *et al.*, *J. Appl. Cryst.* **32** (1999) 339-340 … EXPO

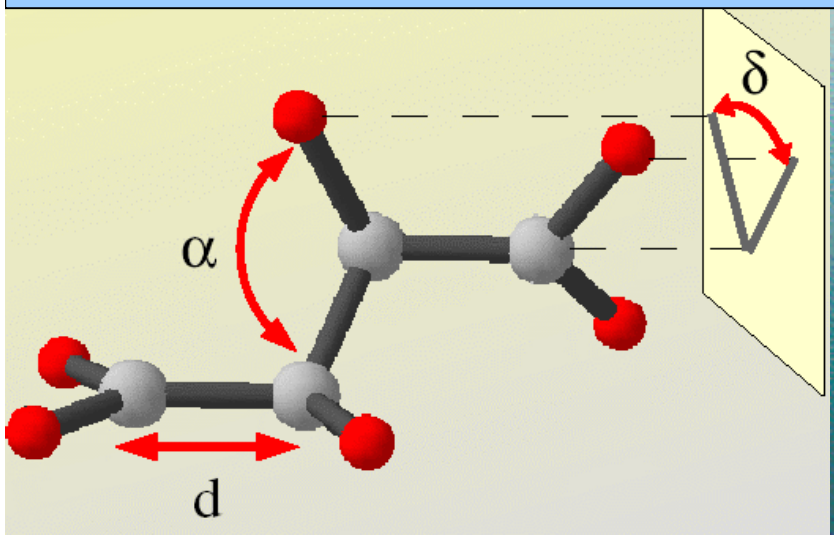
$$\langle |F_{\mathbf{h}}| \mid |F_{p\mathbf{h}}| \rangle = |F_{p\mathbf{h}}|^2 + \sum_j f^2_j(\mathbf{h})$$

Global optimization of the model in direct space:

- Locating known blocks when minimizing cost function (diffraction and/or other observations). Rapidly developing in powder diffraction for last 15 years.

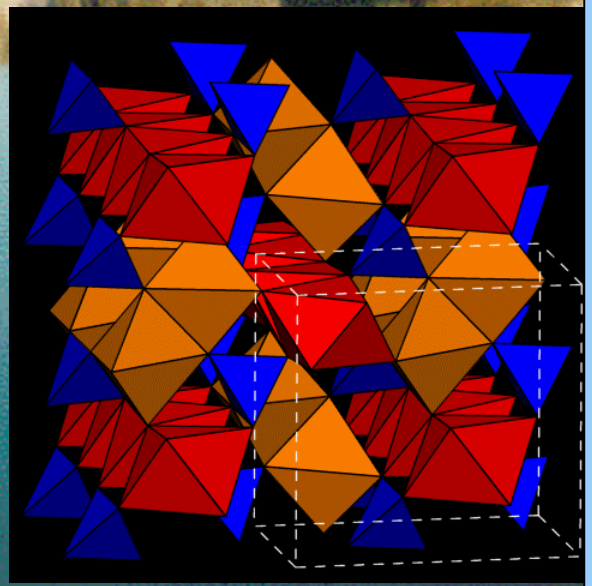
Easy for molecular structures:

- Connectivity is known (plane molecular formula)
- Description in terms of internal coordinates (bond distances, bond angles and torsion angles)



Difficult for non-molecular structures:

- Coordination of atoms is *a-priori* unknown
- Connectivity of blocks is *a-priori* un-known
- Description in terms of int. coo. still possible



Using information on connectivity:

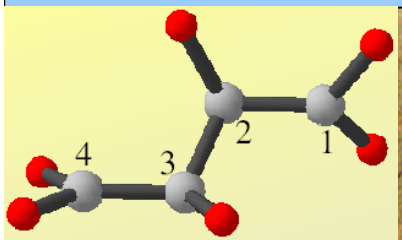
- T-T distances and angles in zeolites
  M.W. Deam, J.M. Newsam, *Nature* **342** (1989) 260-262 …ZEFSA
- Potential energy
  H. Putz et al., *J. Appl. Cryst.* **32** (1999) 864-870 … ENDEAVOUR

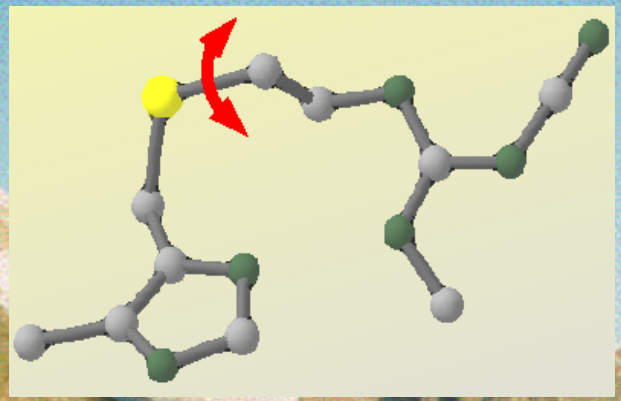
Building Blocks : Description

Atomic coordinates + constraints

Internal coordinates (Z – matrix)

Pitfalls of internal coordinates (Z – matrix)





Atoms are deduced from previous atoms:

- The first atom in the Z-matrix must also be the first to be found
- The convergence can depend on the order of the atoms in the Z-matrix !

```
Z-Matrix
10
C    1
C    1  1.5
C    2  1.5   1  110
C    3  1.5   2  110   1  0
O    1  1.2   2  120   3  0
O    1  1.2   2  120   5  180
O    2  1.4   1  110   3  180
O    3  1.4   2  110   4  180
O    4  1.2   3  120   2  0
O    4  1.2   3  120   9  180
```

number of atoms

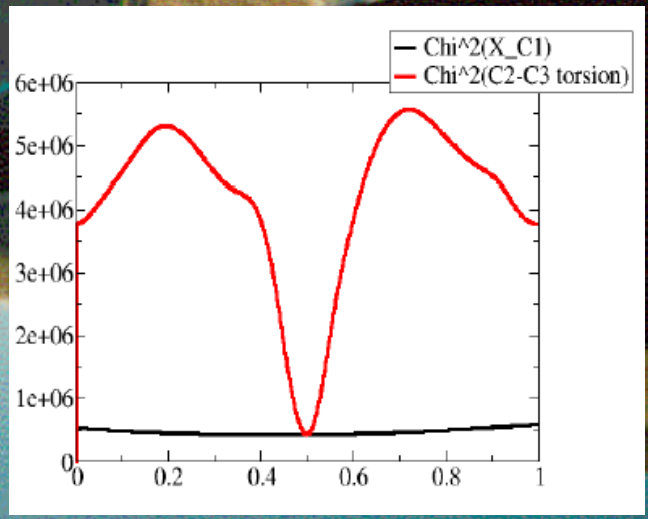
free torsion angles

atom type

bond ***length*** with atom #

bond ***angle*** with atom #

dihedral angle with atom #



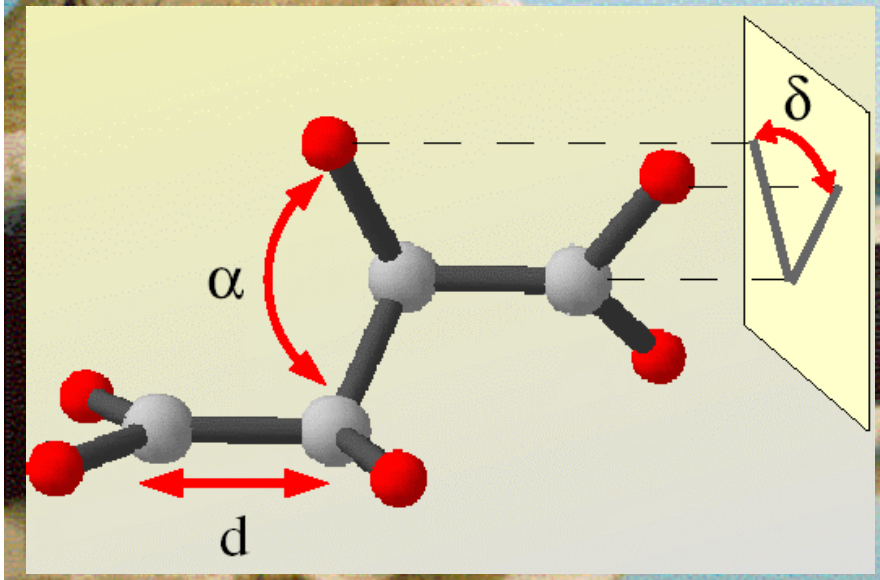
A torsion angle has a much narrower minimum than a translational parameter of a free atom

The internal coordinates approach (Z-matrix) reduces the parameter space to explore, but makes it (much) more difficult to find the solution.

Building Blocks : Description

Flexible approach = isolated atoms + restraints (presented by VFN in Durban)

Idea: keep all the coordination information, but **without its pitfalls**



- *this modelization is **independent from the order of the atoms***
- ***any type of restraint** can be introduced*
- ***any type of movement** can be directly done (no need to compute complex torsions)*

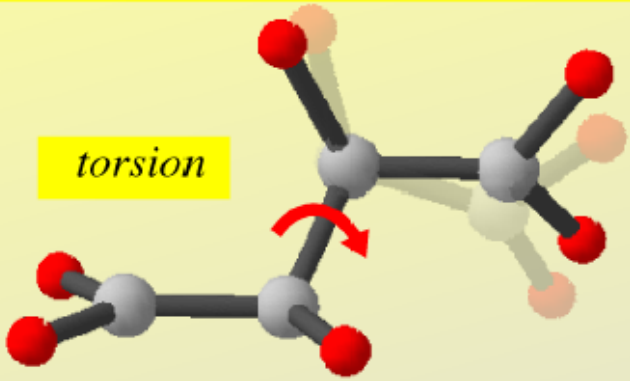
All atom positions are directly defined by their xyz coordinates ***and*** the coordination information is introduced by **restraints** on:

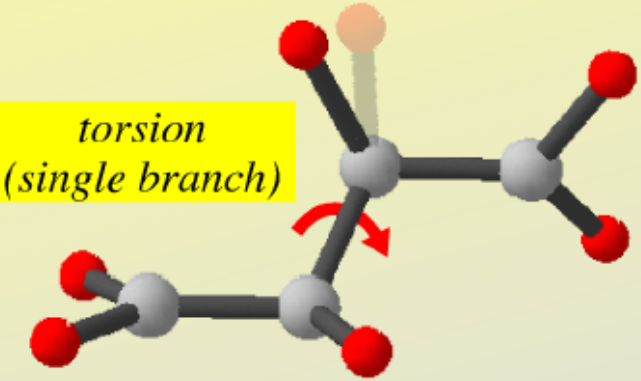
- ***bond lengths*** $\chi^2 = \dfrac{(d - d_0)^2}{\sigma_d^2}$
- ***bond angles*** $\chi^2 = \dfrac{(\alpha - \alpha_0)^2}{\sigma_\alpha^2}$
- ***dihedral angles*** $\chi^2 = \dfrac{(\delta - \delta_0)^2}{\sigma_\delta^2}$

Building Blocks : Description

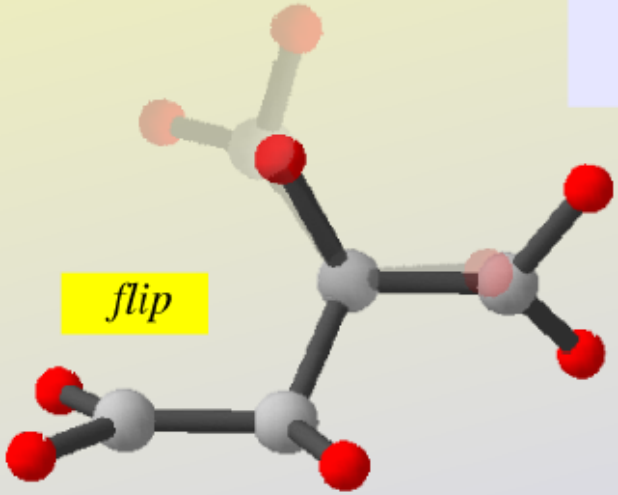
Flexible approach - random moves

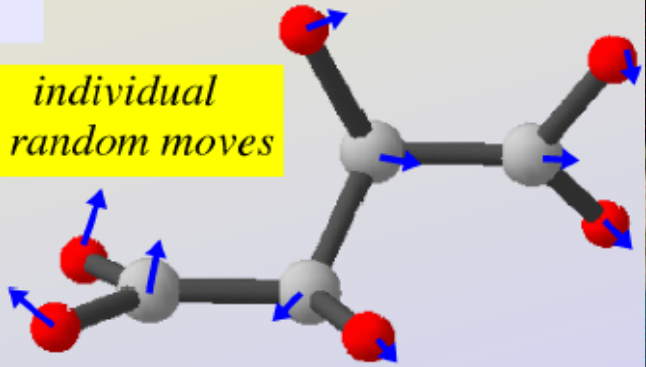
*With atoms defined independently, it is vital to have **intelligent moves** that do not break the restraints*





*All torsion & flip moves that do not break restraints are **automatically identified***

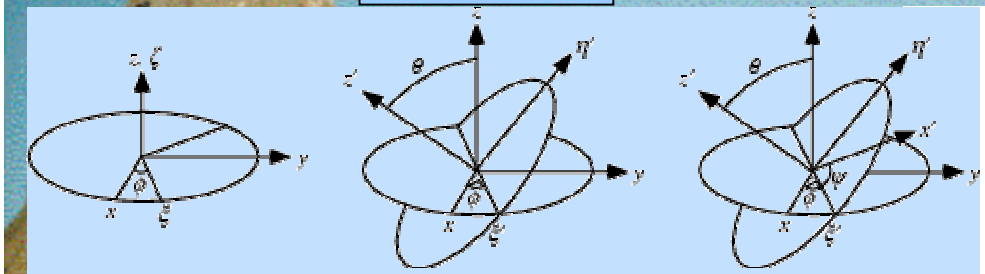




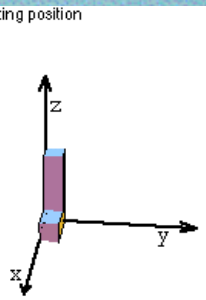
After each random move, a Monte-Carlo test is made on the total internal restraint cost to see if the configuration is kept

Building Blocks : Rotations

Euler angles



$$\begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix},$$



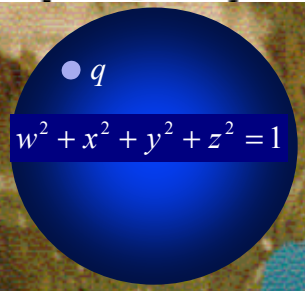
Gimble lock :
Coincidence of inner most and outmost gimbles' rotation axes; loss of degree of freedom.
Rotations around one axis :
Ambiguity of rotations (π, -π).

Euler parameters - quaternions

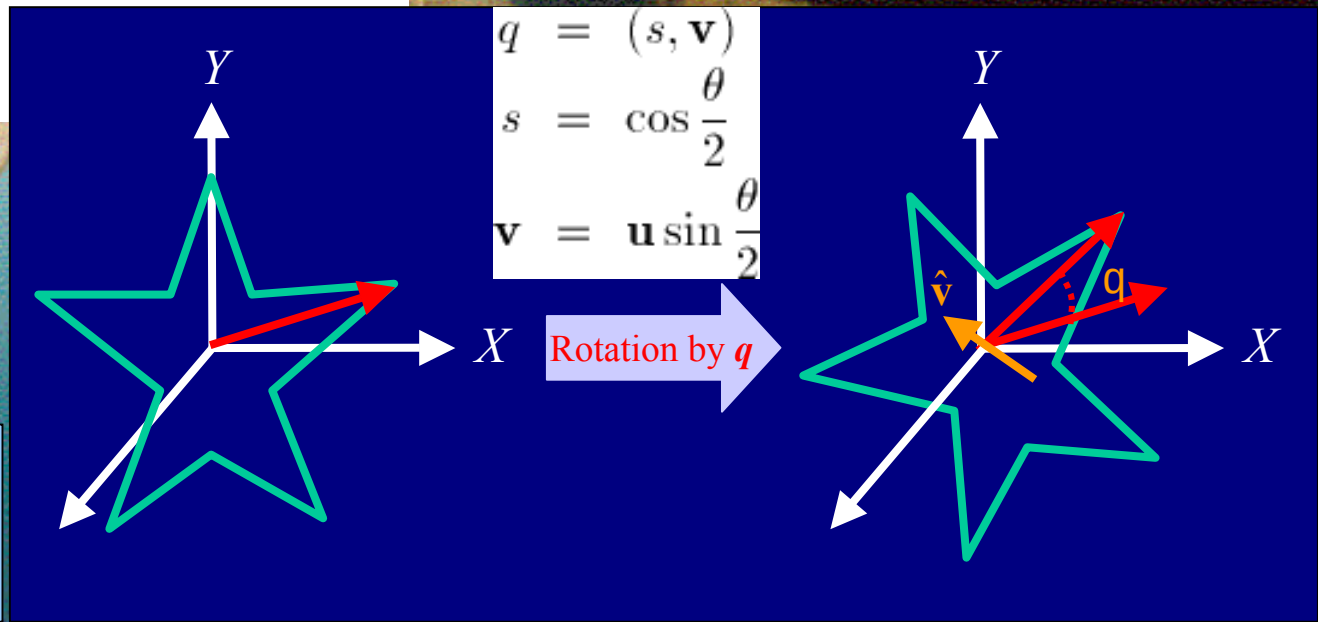
$$\mathbf{q} = w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = [\, x \quad y \quad z \quad w \,] = (s, \mathbf{v})$$

$$s = w$$

$$\mathbf{v} = [\, x \quad y \quad z \,]$$



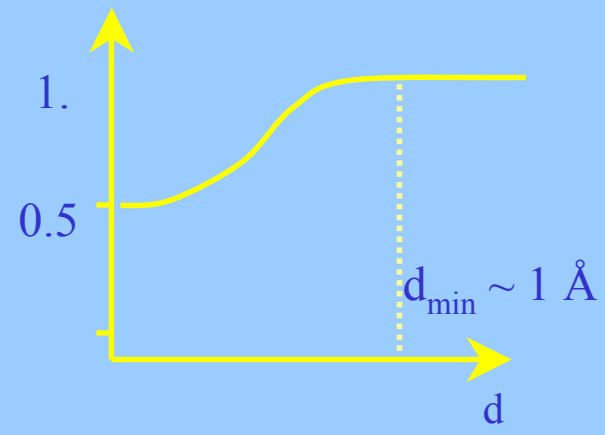
- One-to-one mapping between quaternions and rotations.
- Homogeneous distribution of quaternions on the 4D sphere.



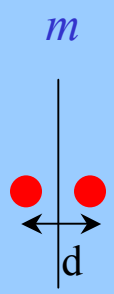
Connectivity of blocks and Special Positions: Dynamical Occupancy Correction

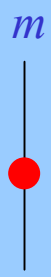
$$\text{Occupancy}^{-1} = 1 + \Sigma_{\text{neighbours closer then } d_{min}} (d_{min} - d_i)/ d_{min}$$

Dynamical occupancy

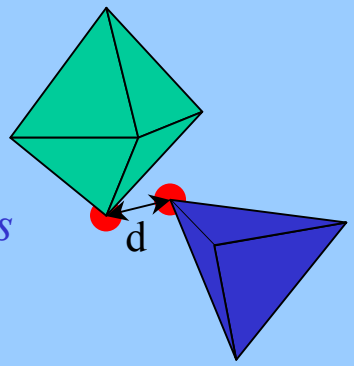


Symmetry

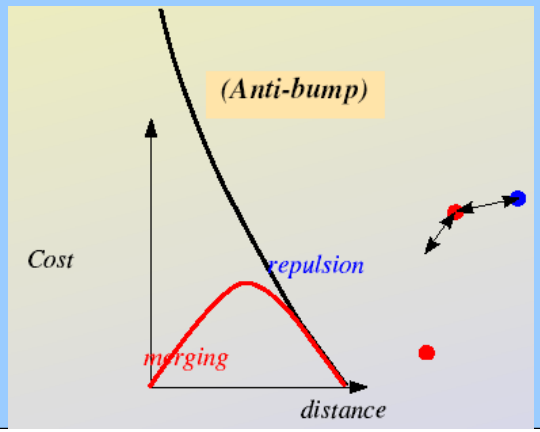


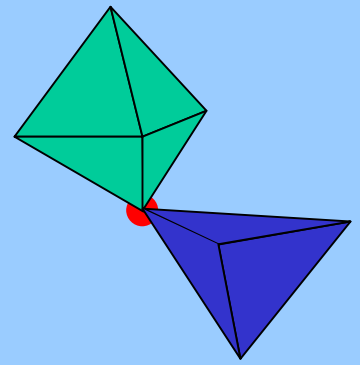


Shared atoms

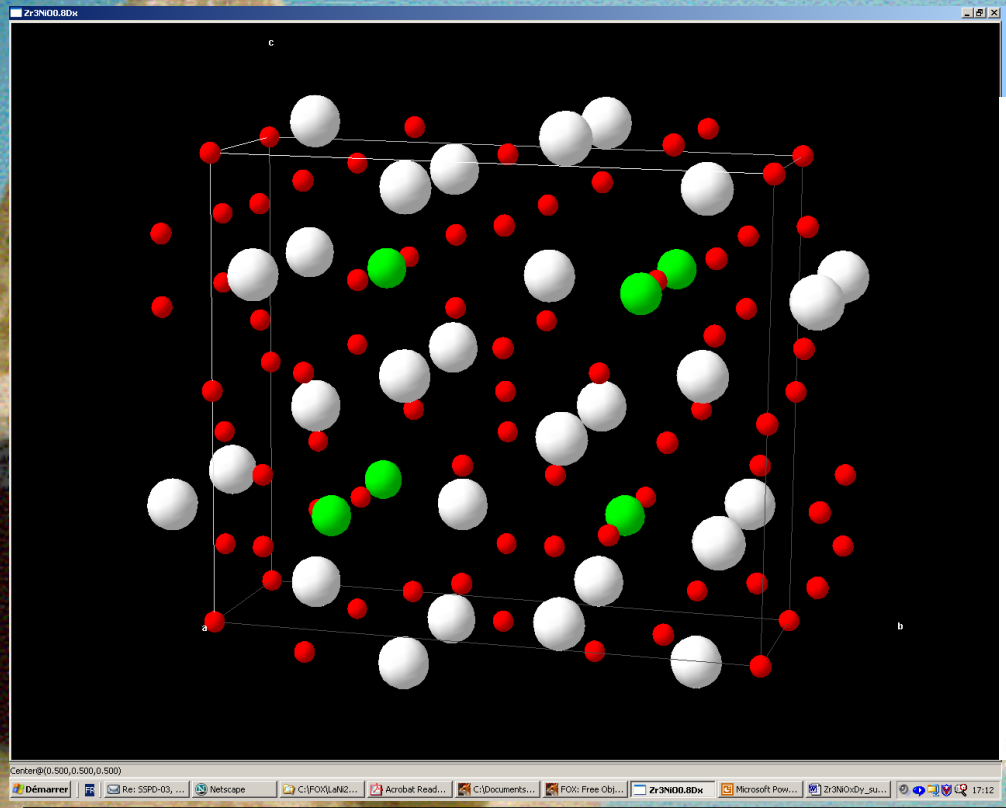


- ***Smooth*** *correction*
- ***Merging*** *of identical atoms*

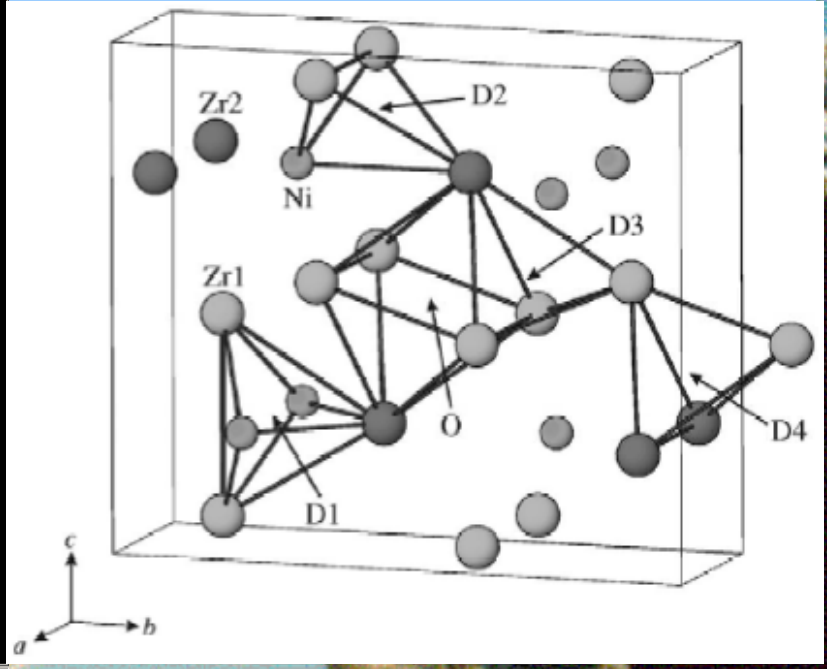




Merging excess atoms : $Zr_3NiO_xD_y$



Zavaliy I.Yu., Cerny R., I.V. Koval'chuck, I.V. Saldan
J. Alloys and Compounds, in press



Merging excess hydrogen (deuterium) atoms by Dynamical Occupancy Correction:

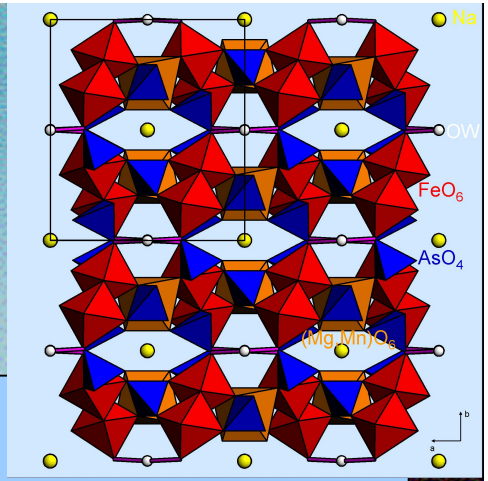
- The amount of absorbed hydrogen is not exactly known.
- The hydrogen atoms are distributed on interstitial sites, which can be predicted, but
- Not all sites are occupied (2 Å exclusion rule) and
- The sites are often partly occupied

so

- Introduce more isolated hydrogen atoms and - Merge the excess by DOC

Merging excess polyhedra : $NaFe_2(Mg_x,Mn_{1-x})[AsO_4]_3 \cdot H_2O$

Structure without knowing exact composition



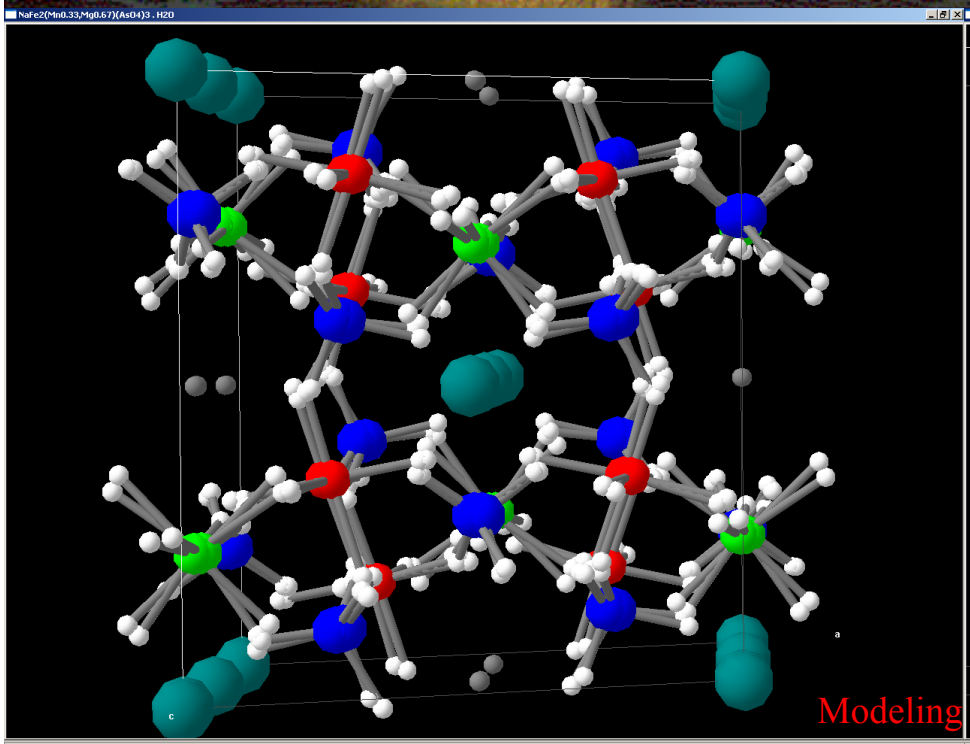
Single crystal data, composition known only qualitatively.

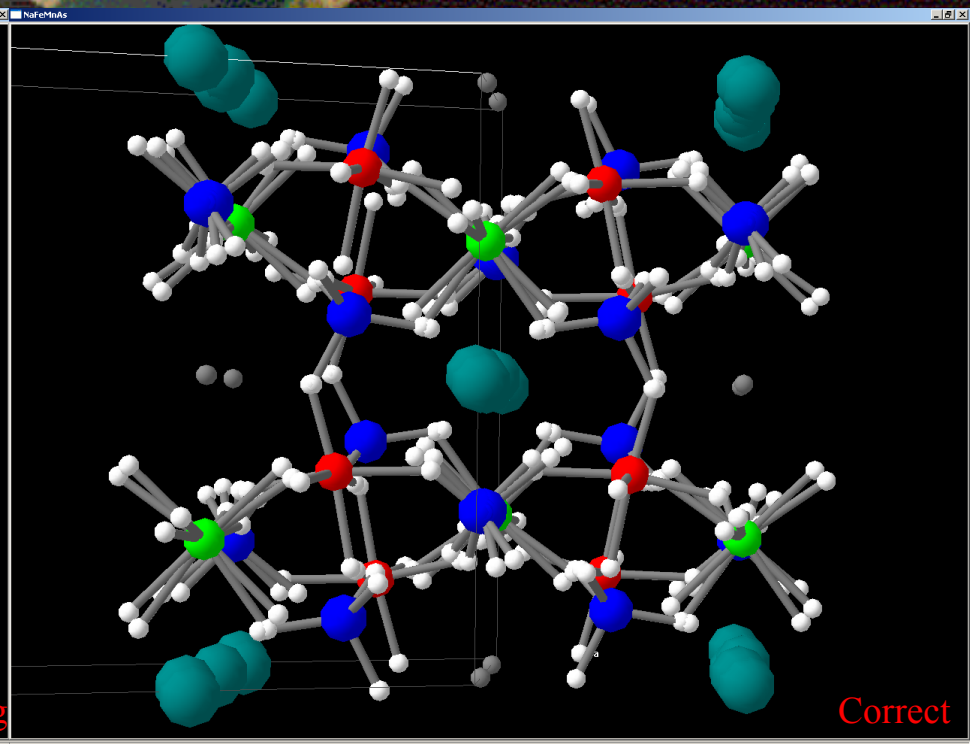
a = 12.181(1), b = 12.807(1), c = 6.6391(5) Å, β = 112.441(9)°, $C2/c$

Modeling: 3 tetrahedra AsO_4, 2 octahedra FeO_6, 1 octahedron MnO_6, 1 atom Na, 1 H_2O

Correct : 2 tetrahedra AsO_4, 1 octahedra FeO_6, 1 octahedron $(Mg,Mn)O_6$, 1 atom Na, 1 H_2O

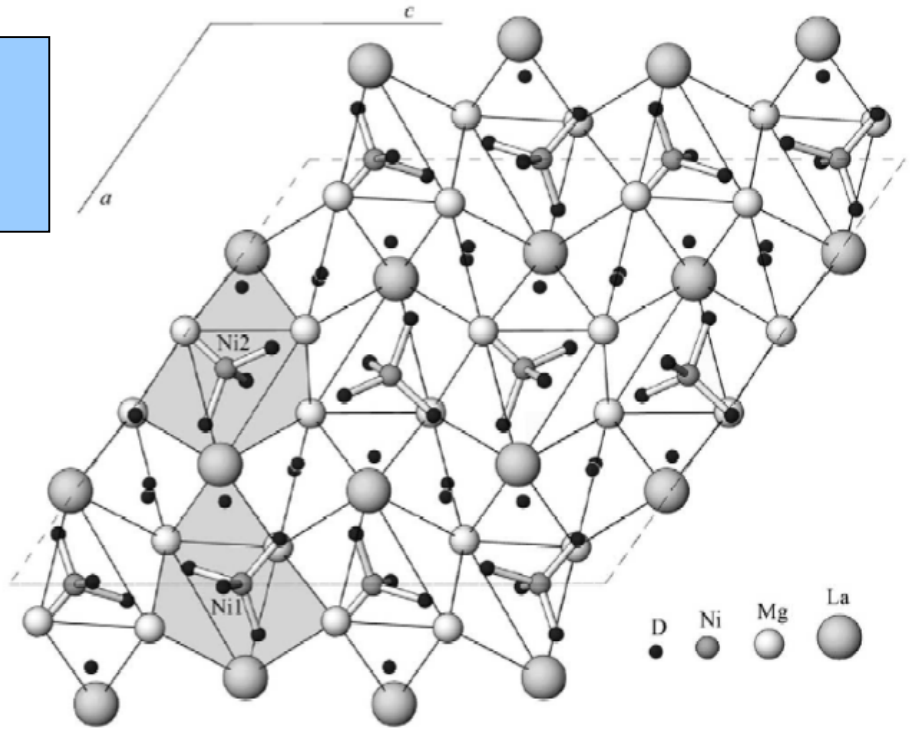
Excess polyhedra merged by DOC. No solution by Direct Methods when composition wrong!





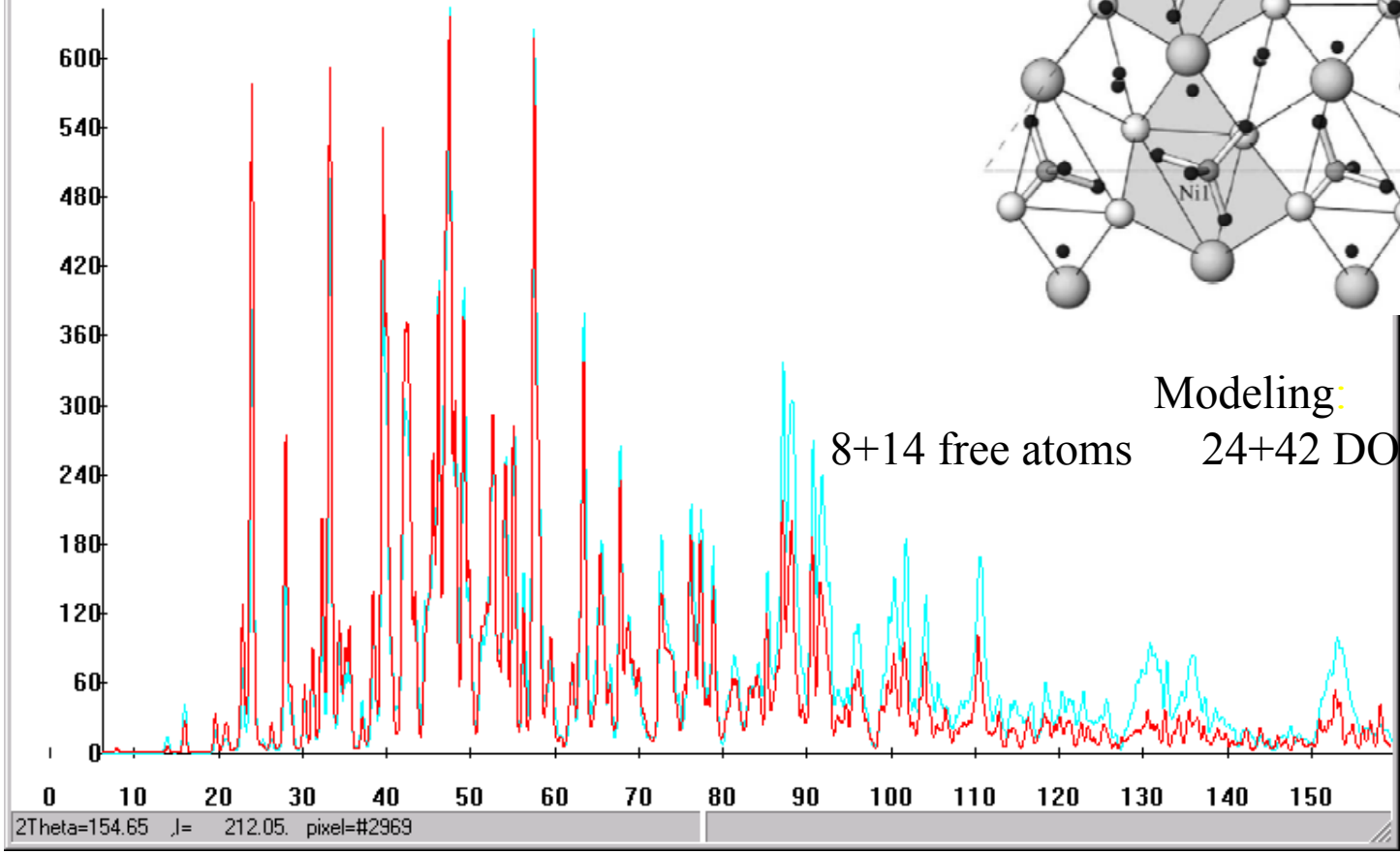
Isolated atoms : $LaMg_2NiD_7$

a = 13.9785 ***b*** = 4.7025 ***c*** = 16.0244 **β** = 125.238,
$P2_1/c$, metal atoms pseudosymmetry *Cmcm*
8 metal and 14 deuterium atoms

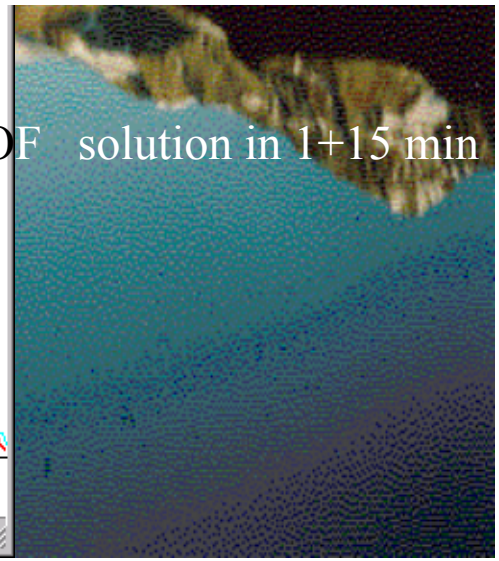


P21/c

Neutron pattern (ILL-D2B), λ = 1.594 Å



Modeling:
8+14 free atoms 24+42 DOF solution in 1+15 min

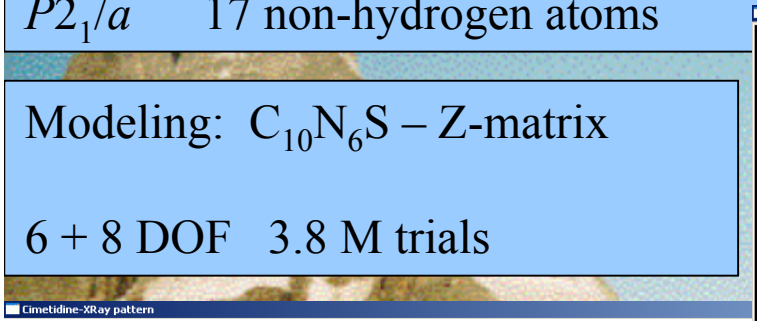


Molecule : Cimetidine

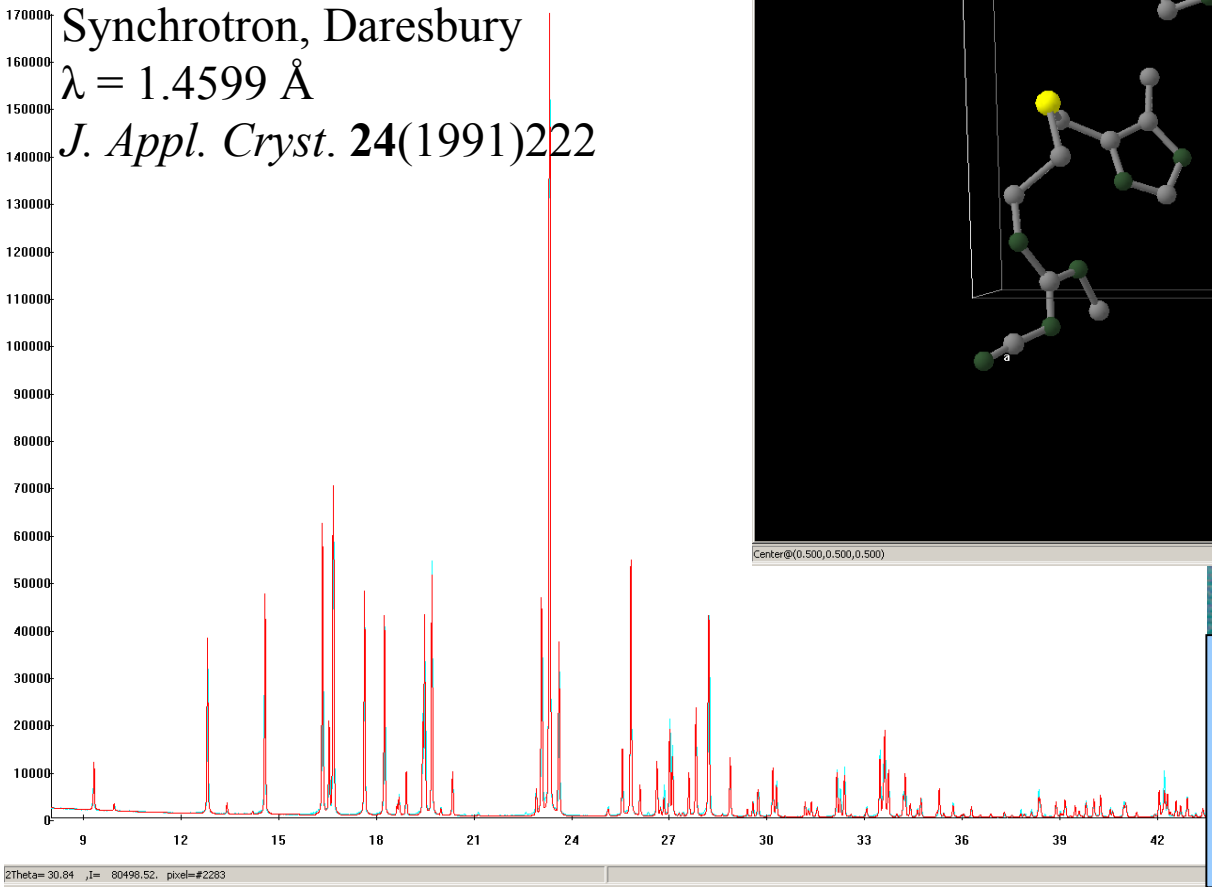
a = 10.3941 ***b*** = 18.8188 ***c*** = 6.8249 **β** = 106.44
$P2_1/a$ 17 non-hydrogen atoms

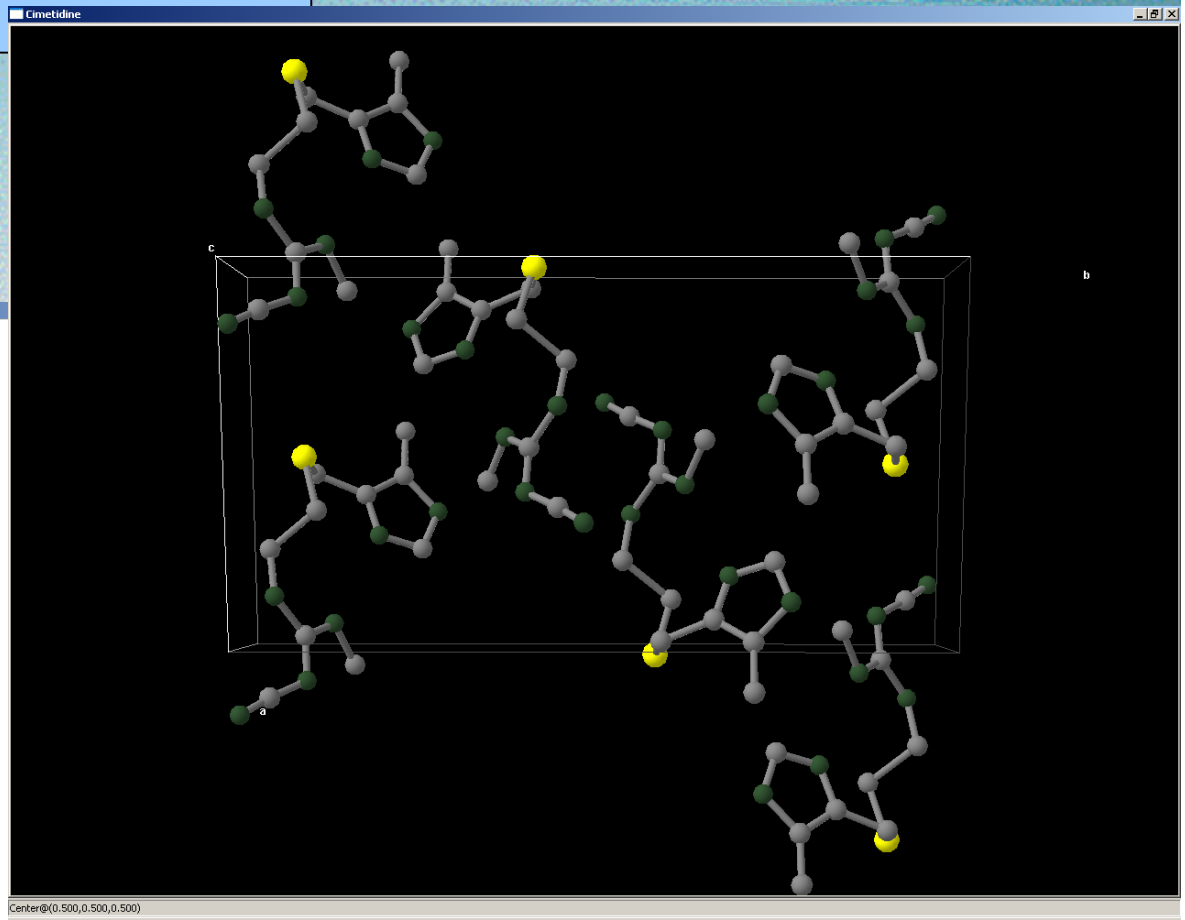
Modeling: $C_{10}N_6S$ – Z-matrix

6 + 8 DOF 3.8 M trials



Synchrotron, Daresbury
λ = 1.4599 Å
J. Appl. Cryst. **24**(1991)222

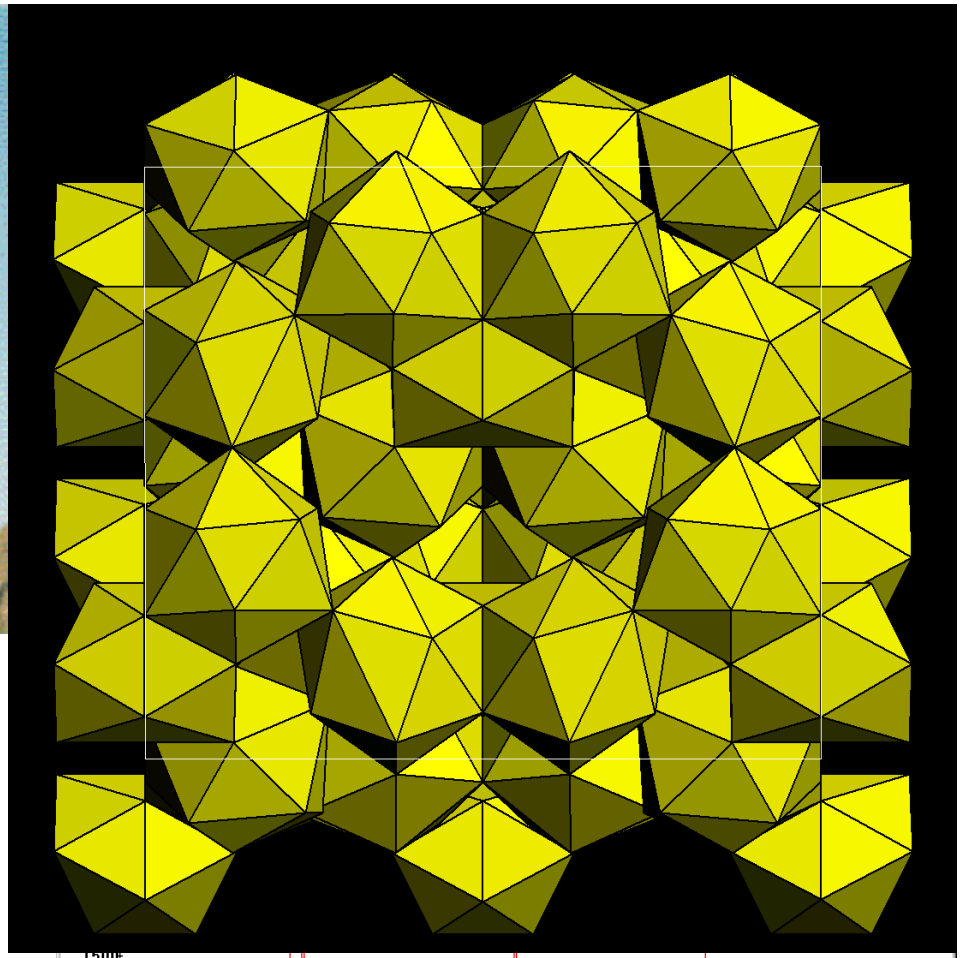




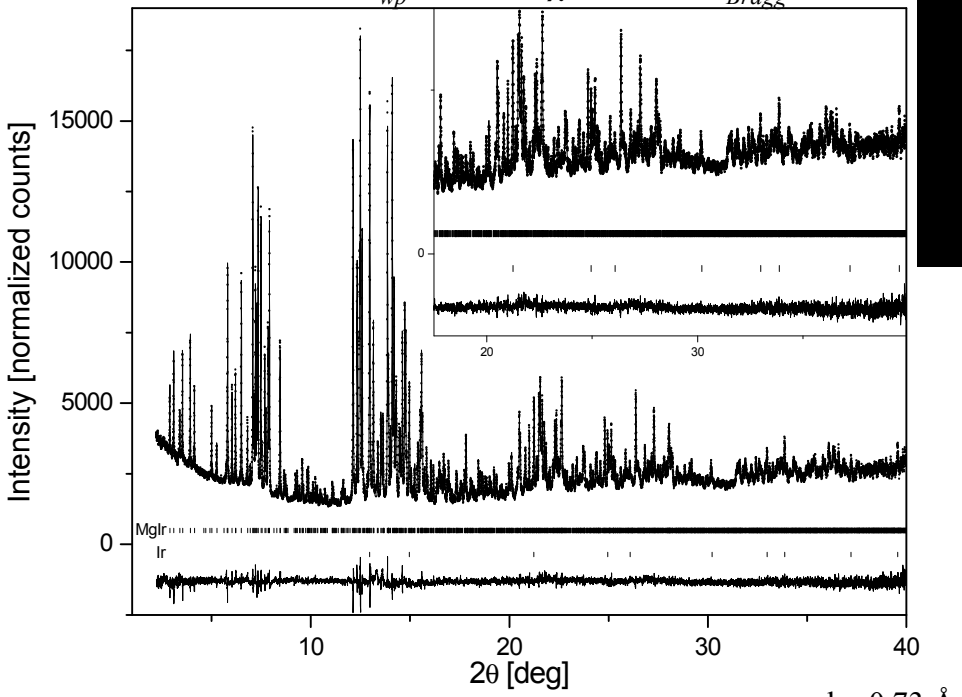
Modeling: $C_{10}N_6S$ – isolated atoms + bonding restraints
6 + 51 DOF + 40 restraints
1.85 M trials

Polyhedra : MgIr

a = 18.4648 ***b*** = 16.1701 ***c*** = 16.8175, *Cmca*
13 Ir and 12 Mg atoms
Frank-Kasper intermetallic compound

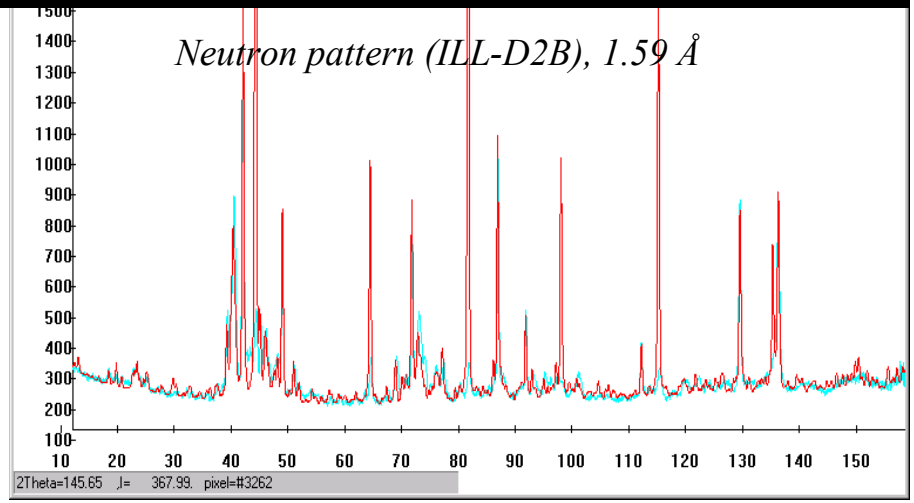


SNBL, 0.50012 Å, R_{wp} = 0.094, χ^2 = 3.02, R_{Bragg} = 0.05



d = 0.73 Å

Neutron pattern (ILL-D2B), 1.59 Å



Modeling MgIr

Modeling:

Free atoms

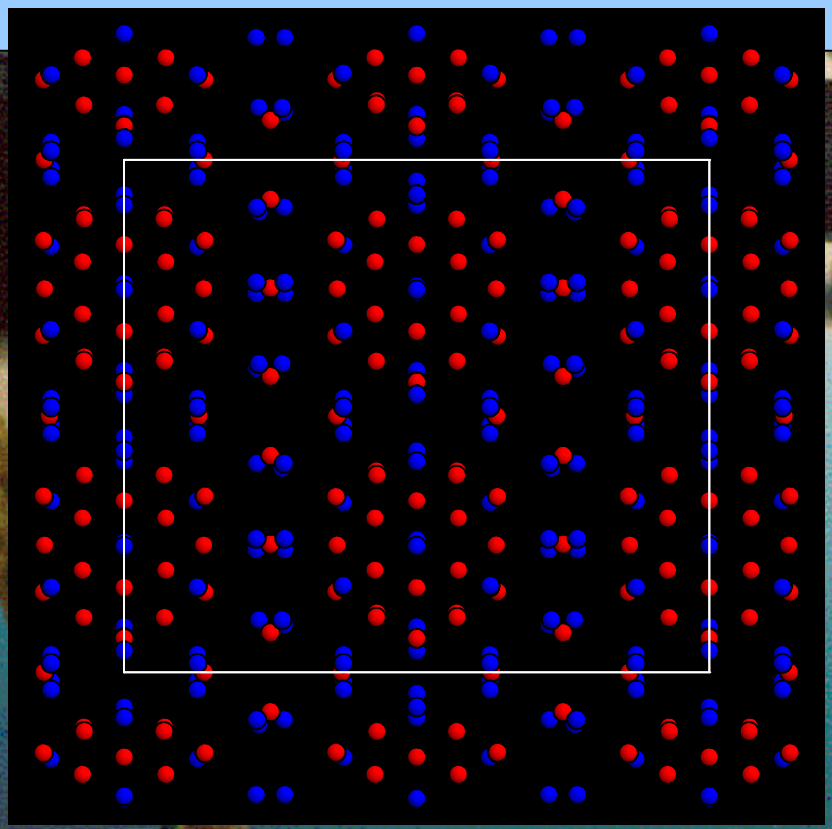
13 Ir atoms introduced on general positions and quickly localized.
All Ir atoms fixed and 12 Mg atoms introduced and localized. Any additional atom merged.

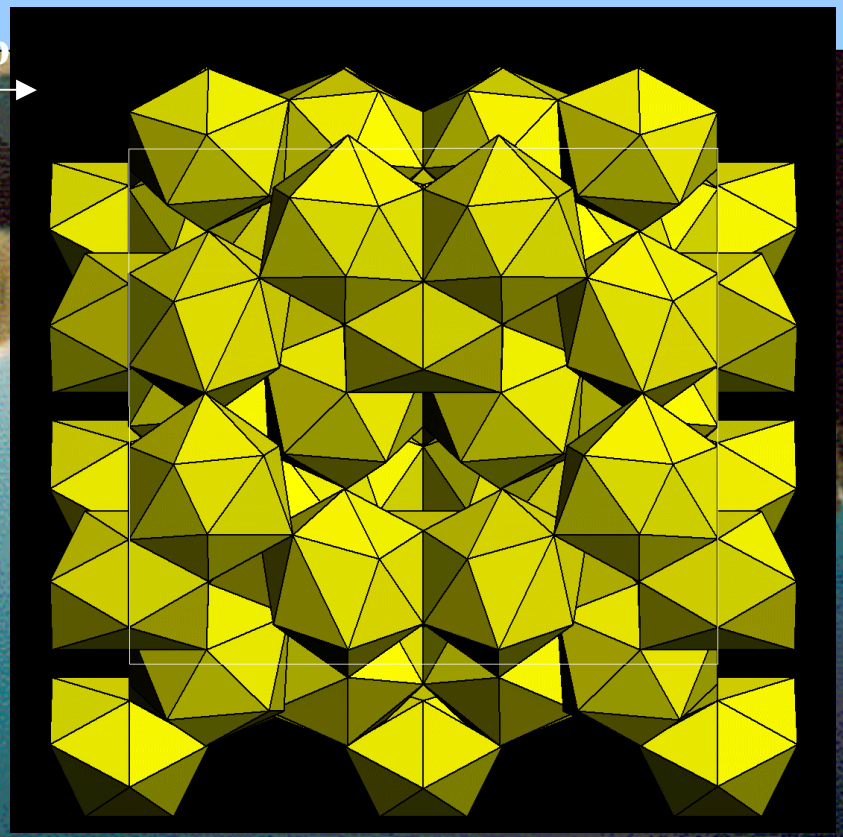
13+12 free atoms = 39+36 free p. solution in 15+160 min

Polyhedra

7 icosahedra $(Ir, Mg)_{12}$ introduced on general position and localized.
Ligands identification difficult – needs more programming.

7 icosahedra = 42 free p. solution in 120 min



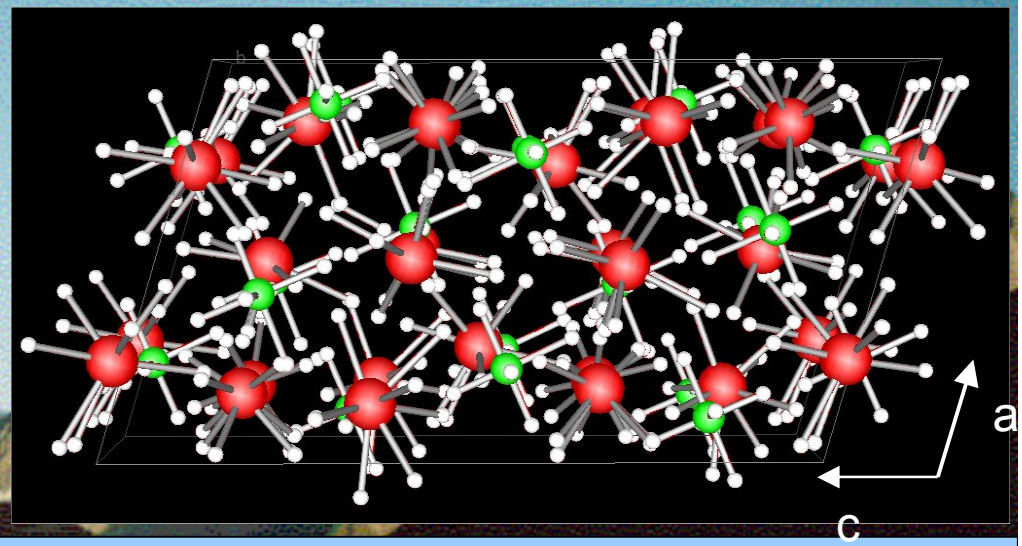


Polyhedra : $Sr_5V_3(F/O/OH/H_2O)_{22}$

a = 11.2428, ***b*** = 8.19535, ***c*** = 19.94831,
b = 106.72, $P2_1/c$, 30 atoms

8 polyhedra + 22 free atoms =
48 + 66 DOF solution in 5 + 10 min

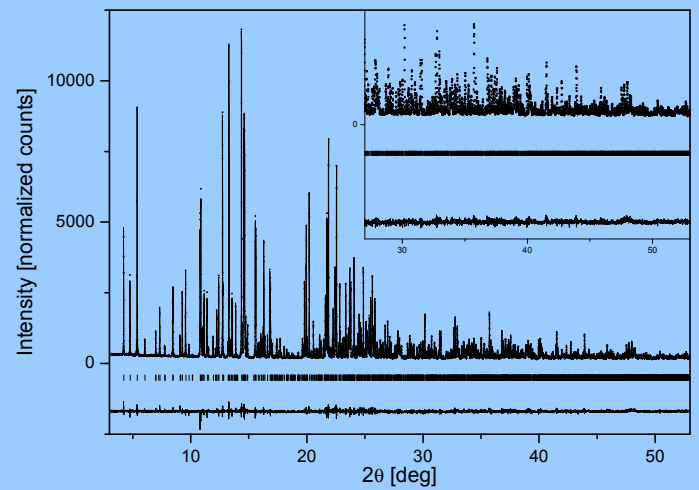
simulated data, synchrotron, λ = 0.79764 Å

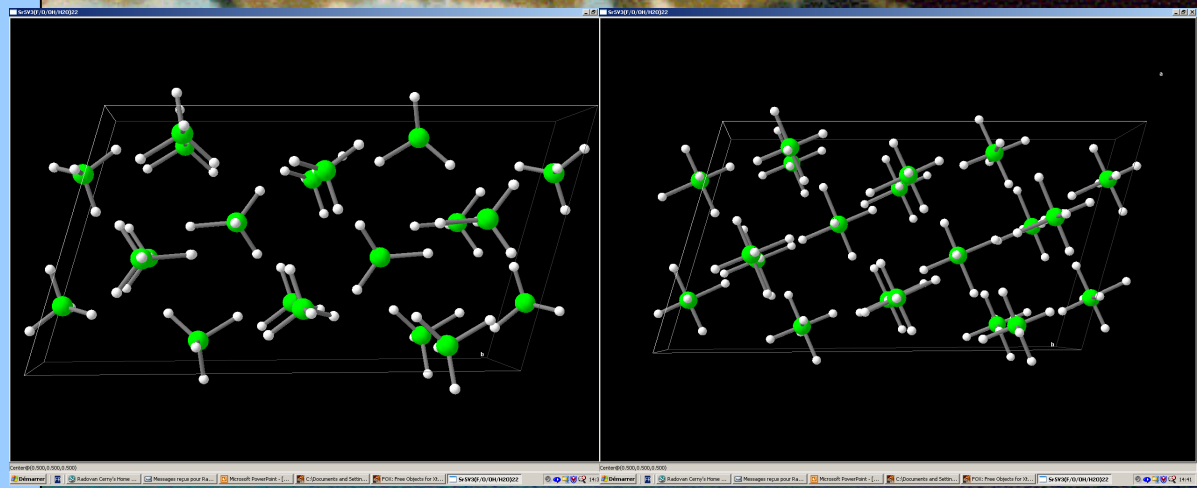


Modeling:

3 polyhedra SrO_9 and 2 polyhedra SrO_{11} and

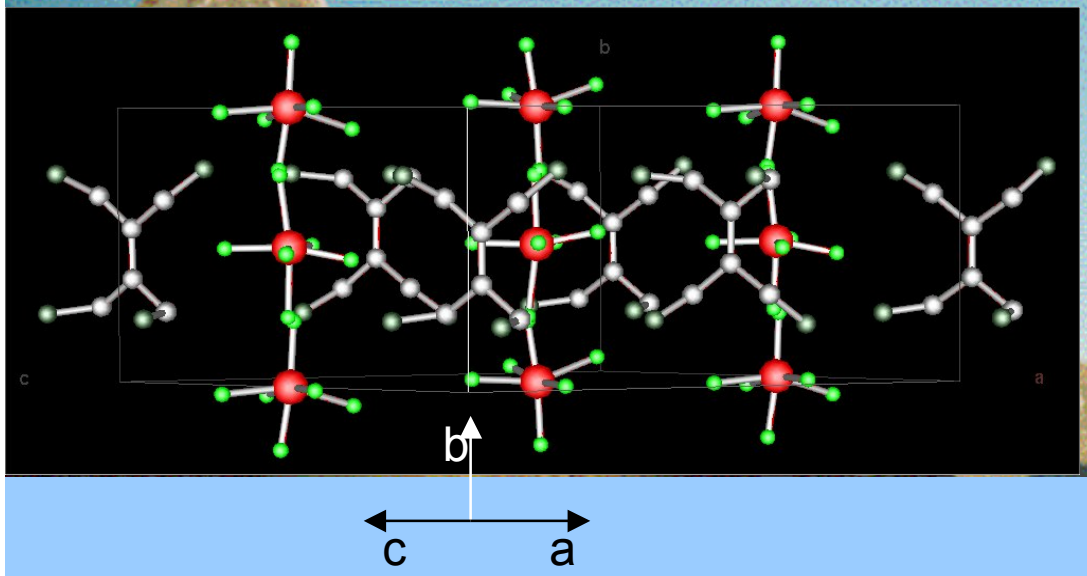
3 tetrahedra VO_6 or 3 octahedra VO_6

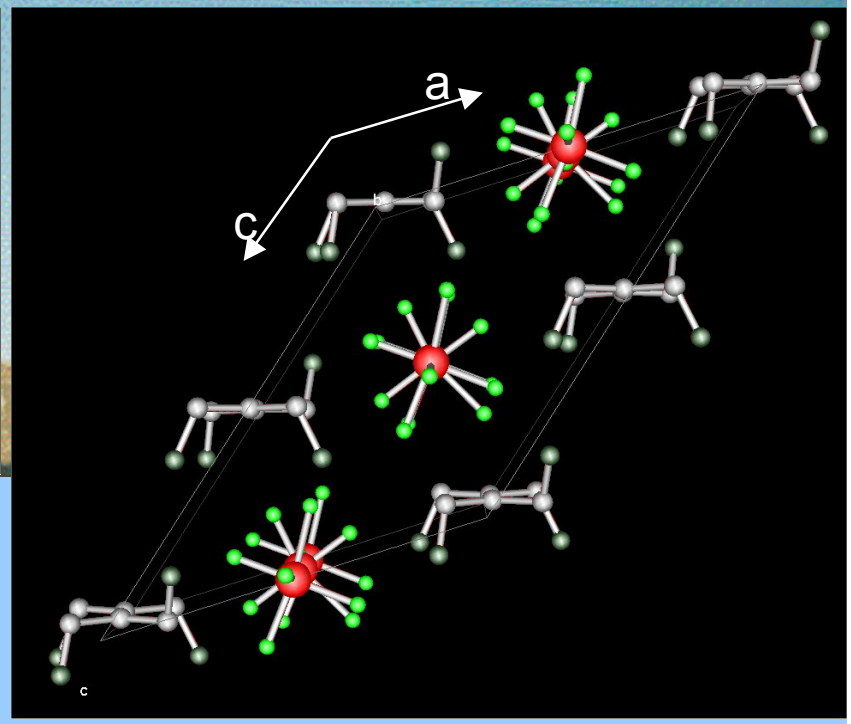




R_{wp} = 0.195, χ^2 = 6.12, R_{Bragg} = 0.103

Mixed : $Al_2F_{10}[C_6N_4H_20]$





Powder data, laboratory, CuK_α, ***a*** = 10.3258(9), ***b*** = 7.3960(6), ***c*** = 13.243(1) Å, b = 139.901(2)° , *Pc*

Modeling in *P2/c*: 2 octahedra AlF_6, 1 molecule $C_6N_4H_{20}$

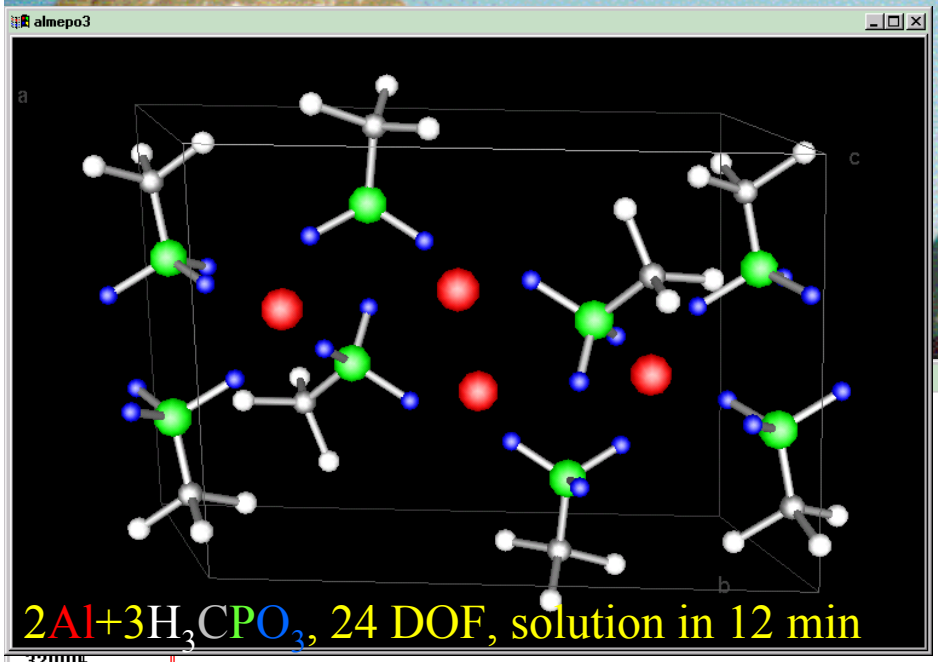
Both octahedra quickly and correctly localized, non-centrosymmetric molecule localized on the inversion centre. The structure transformed to the non-centrosymmetric space group and refined.

2 polyhedra + 1 semi-rigid molecule = 12 + 6 free params. solution in 5 min

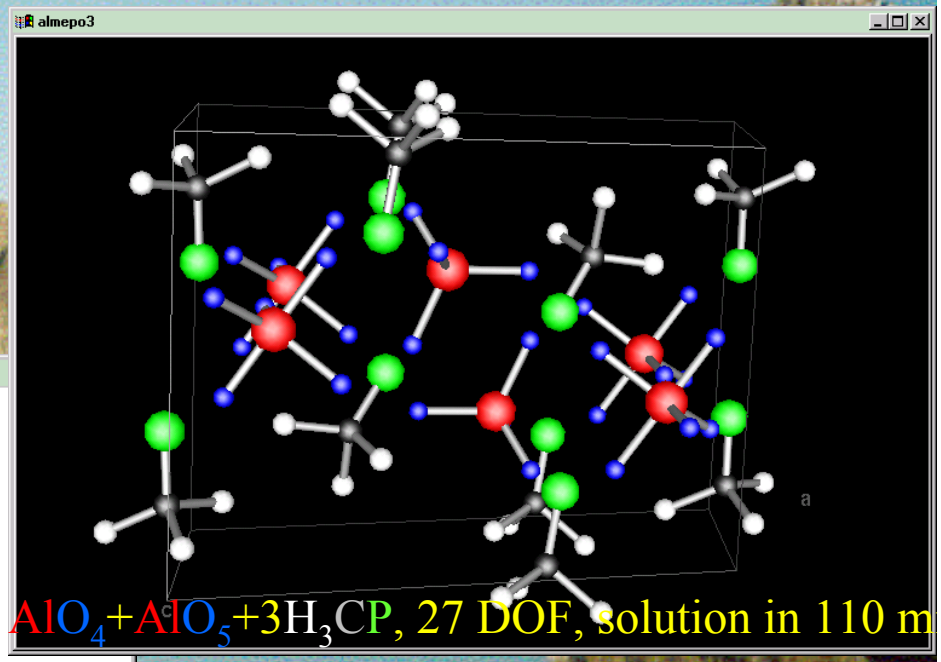
Mixed : $Al_2(CH_3PO_3)_3$

Edgar *et al.* (2002) *Chem. Commun.* **8** (2002) 808-809

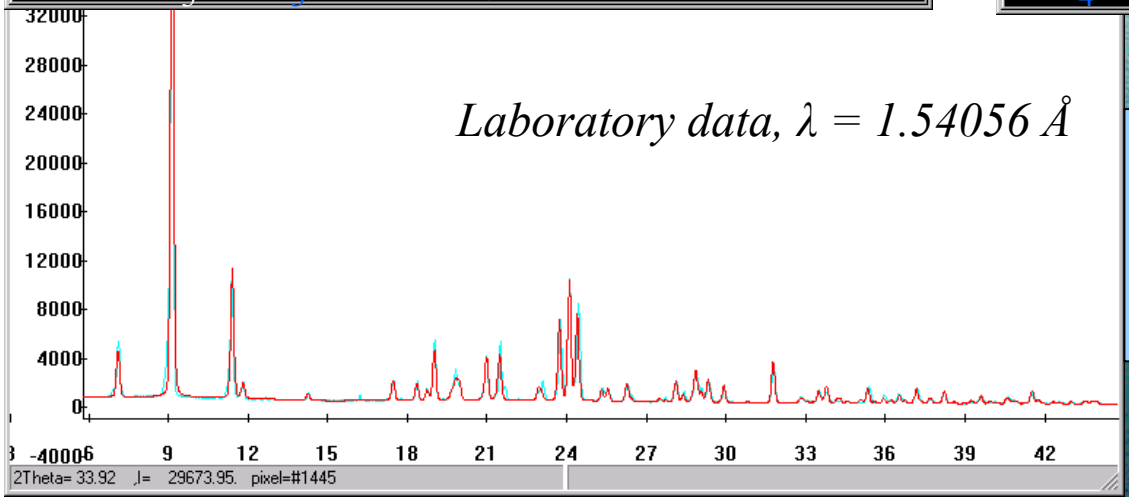
a = 13.2979 ***b*** = 9.6570 ***c*** = 5.0725 **α** = 89.601 **β** = 111.242 **γ** = 92.088, $P\ \bar{1}$, 26 atoms



2Al+3H_3CPO_3, 24 DOF, solution in 12 min



AlO_4+AlO_5+3H_3CP, 27 DOF, solution in 110 m



Laboratory data, λ = 1.54056 Å

Buildings blocks deduced from $_{27}Al$, $_{13}C$ and $_{31}P$ NMR spectra :

3 C-PO3 units and 2 Al atoms

Errors in Building Blocks ?

Maximum likelihood aproach (presented by VFN in Durban)

In a "classical approach" :

$$\sigma^2 = y_{obs}$$

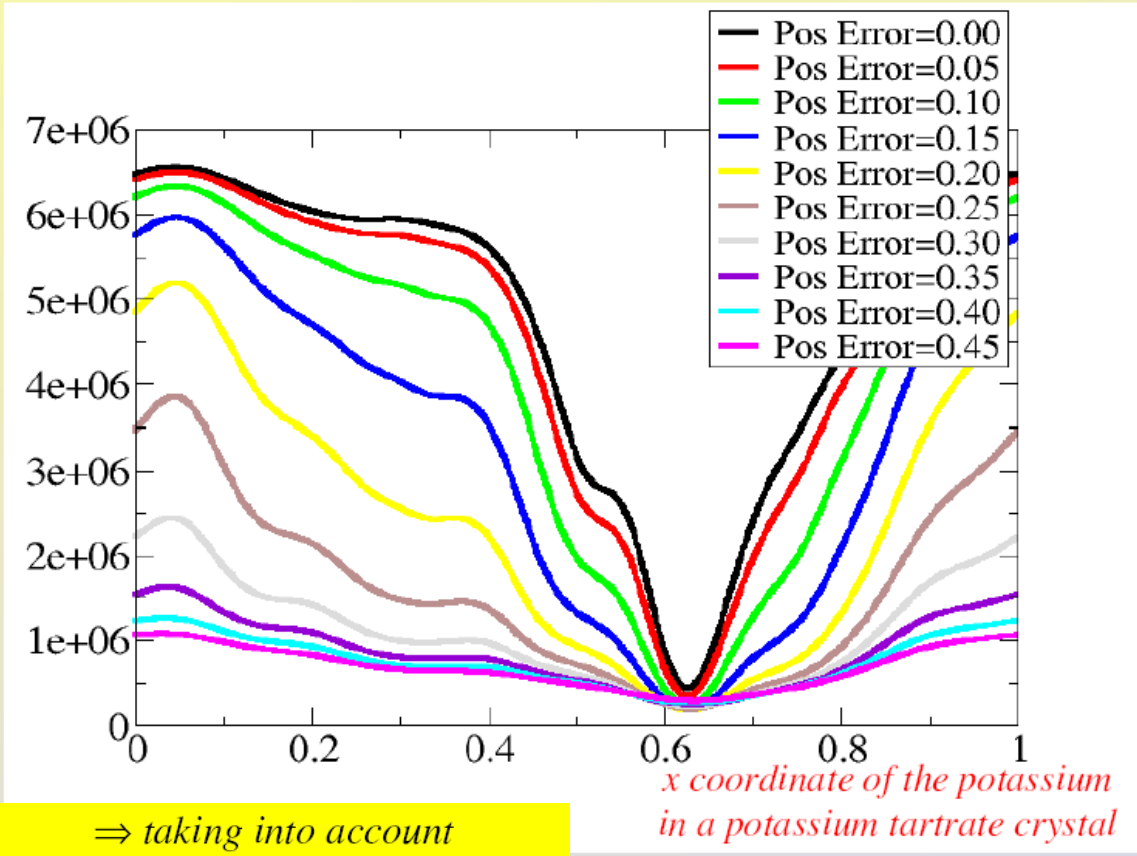
$$\chi^2 = \sum \frac{(y^i_{obs} - y^i_{calc})^2}{\sigma^2_i}$$

Assumes that the model can fit perfectly the observed data during the global optimization.
But there are positional errors in the model, because the model is not yet correct!
And some building blocks can be missing!

$$\sigma^2_i = \sigma^2_{calc} + \sigma^2_{obs}$$

$$\chi^2 = \sum \frac{(y^i_{obs} - \langle y^i_{calc} \rangle)^2}{\sigma^2_i}$$

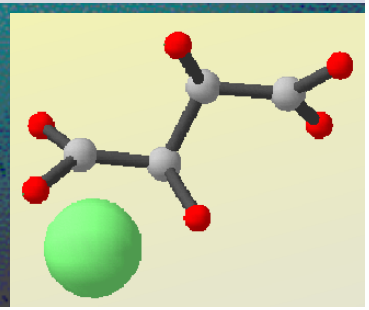
Hypersurface : $\chi^2 = f(\text{parameters})$



x coordinate of the potassium in a potassium tartrate crystal

⇒ *taking into account* ***random positionnal errors*** :

- ***increases the width of the global minimum*** *(for small errors)*
- ***flattens the hypersurface*** *for large errors*

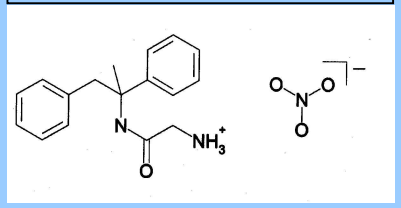


Solving structures per partes ?

Missing atoms (*blur*) do not contribute to calculated structure factor, but increase the σ^2_{calc}.

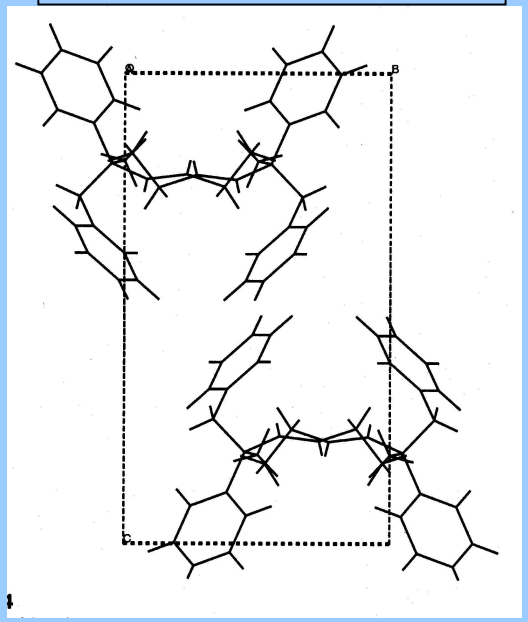
A.J. Markvardsen, W.I.F. David and K. Shankland, *Acta Cryst.* **A58** (2002) 316 - 326

remacemide nitrate

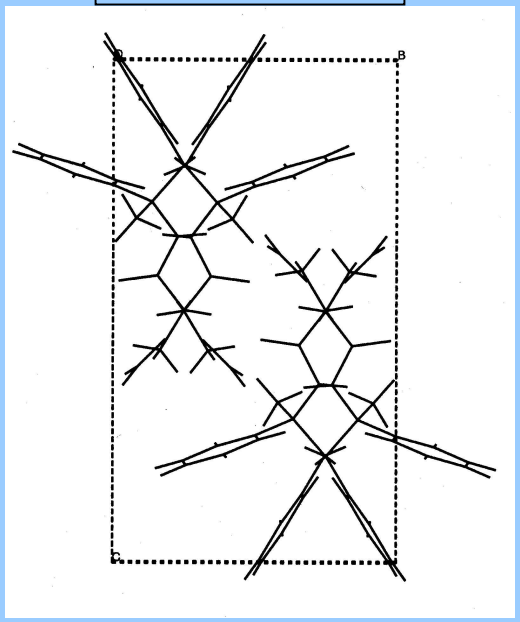


Nitrate ion contains ~ 18 % of the total scattering power

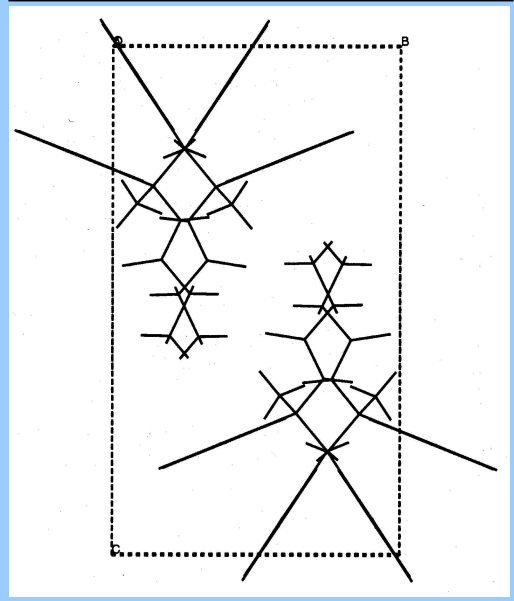
LS solution without NO_3



correct solution



ML solution without NO_3, but taking into account its contribution to σ^2_{calc}.



Summary

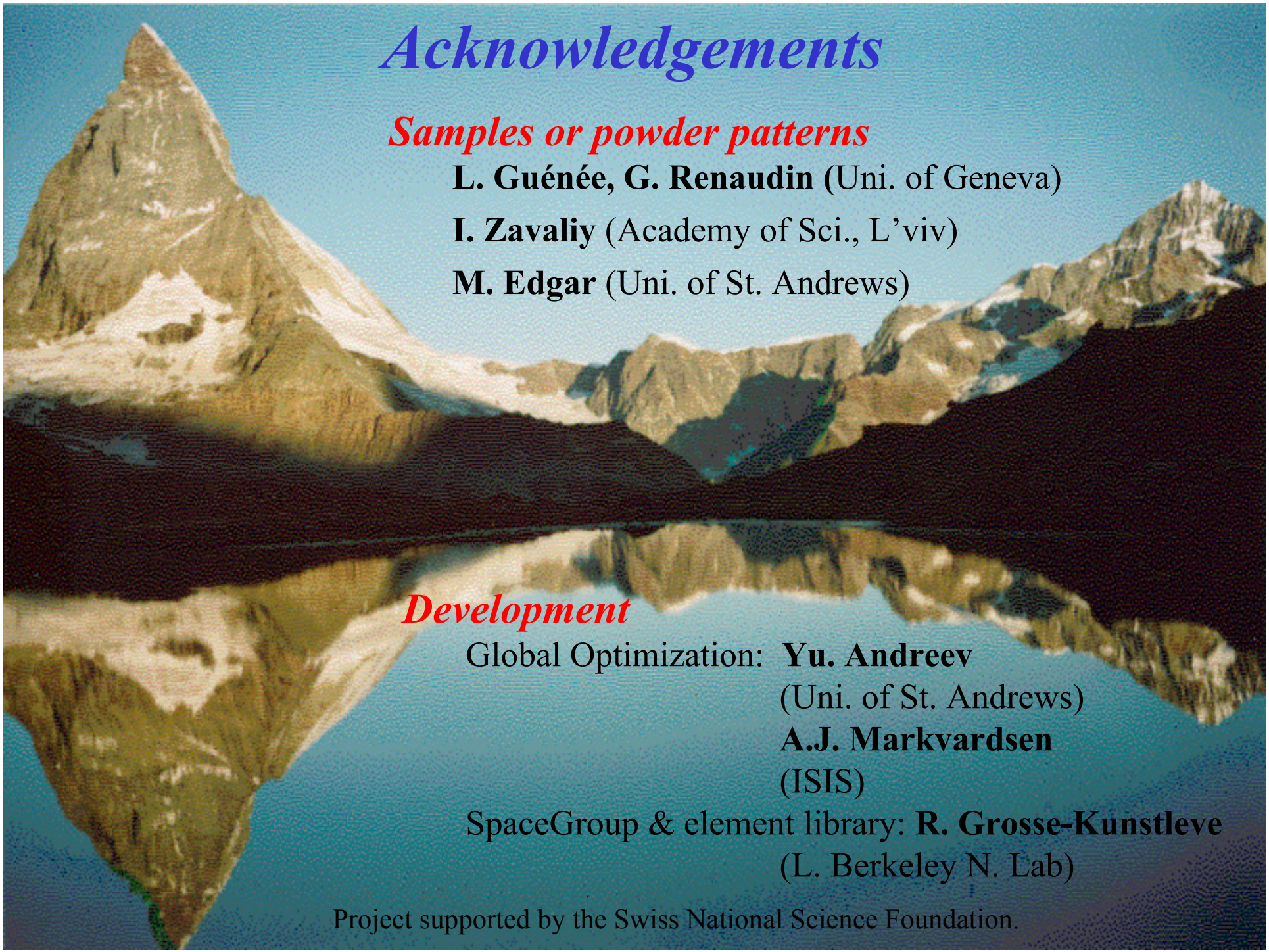
FOX - program for the global optimization of crystal structures (powder diffraction)

- ***any*** combination of ***building blocks*** :
  isolated atoms, molecules, polyhedrons
- ***no*** assumption on the actual ***connectivity***
- ***no*** *a priori* knowledge about ***special positions***

Future developments :

- Optimizing individual building blocks separately in different worlds of parallel tempering?
- "Genetic" approach, to copy "best" partial structures between the different parallel optimizations?
- Use additional information (potential energy, bond valences etc.) for searching the correct connectivity between blocks.

Favre-Nicolin V., Černý R., *J. Appl. Cryst.* **35** (2002), 734-743
Availability (Windows, Linux): ***http://objcryst.sourceforge.net***



Acknowledgements

Samples or powder patterns

L. Guénée, G. Renaudin (Uni. of Geneva)

I. Zavaliy (Academy of Sci., L'viv)

M. Edgar (Uni. of St. Andrews)

Development

Global Optimization: **Yu. Andreev**
(Uni. of St. Andrews)
A.J. Markvardsen
(ISIS)

SpaceGroup & element library: **R. Grosse-Kunstleve**
(L. Berkeley N. Lab)

Project supported by the Swiss National Science Foundation.