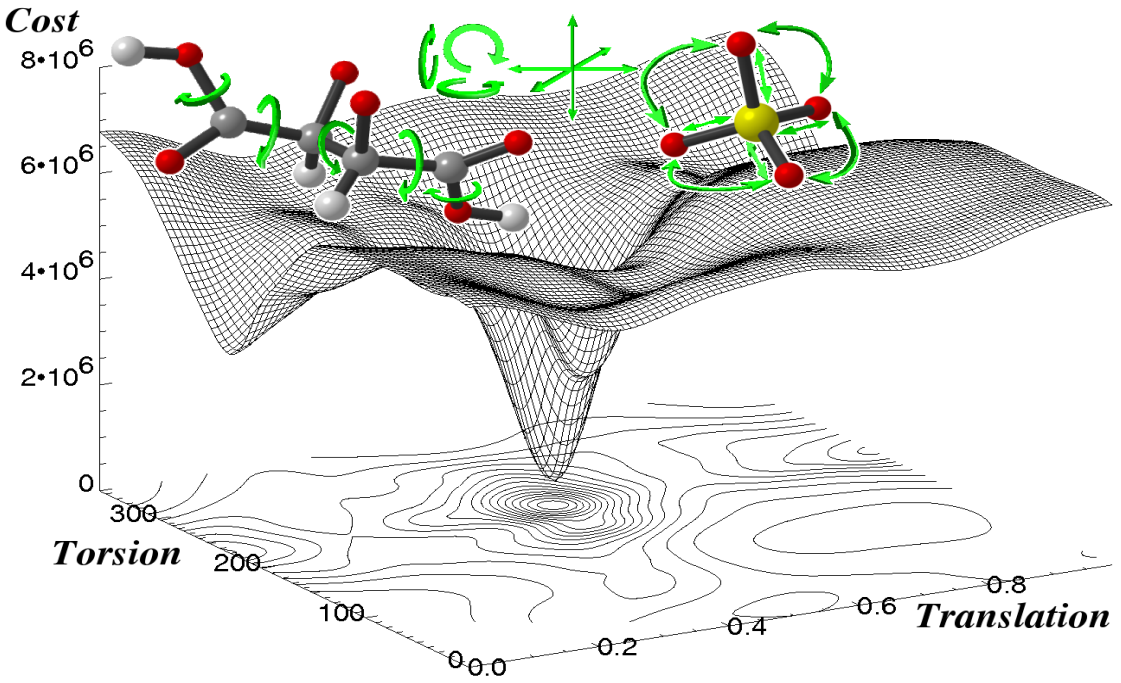
Structure solution in direct space using Fox and smart restraints

Vincent Favre-Nicolin

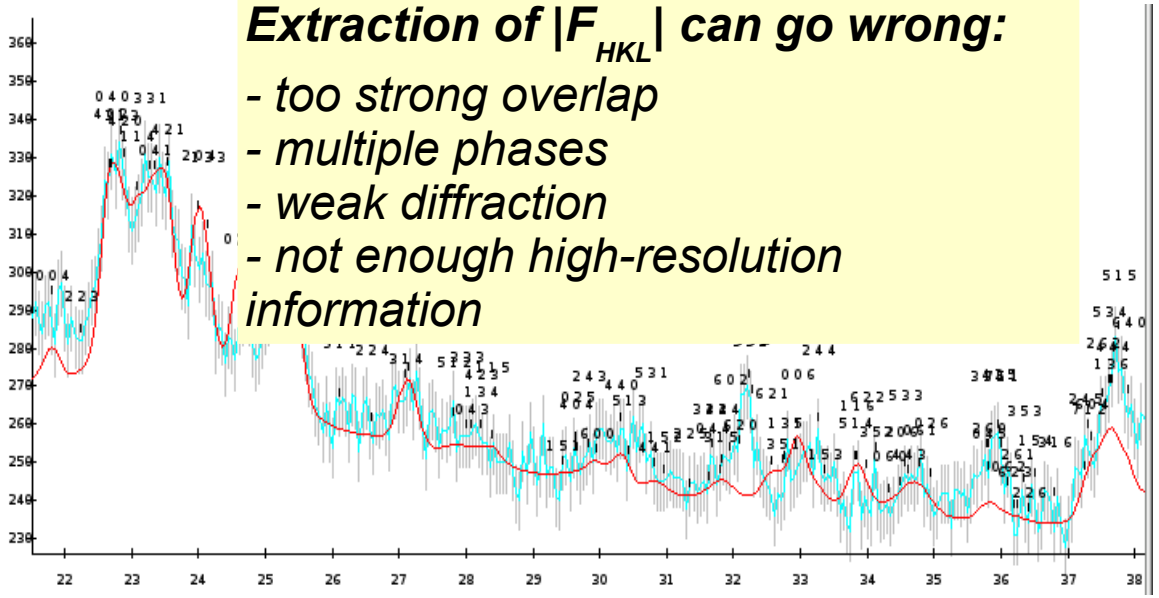
CEA, INAC, SP2M & Université Joseph Fourier – Grenoble, France

http://vincefn.net/reciprocs/



Introduction: Structure Determination in Direct Space

Real (Direct)-Space Methods vs. Reciprocal-Space (Direct) Methods



Powder diffraction :
Extraction of $|F_{HKL}|$ can go wrong:

- too strong overlap
- multiple phases
- weak diffraction
- not enough high-resolution information

Single Crystal:

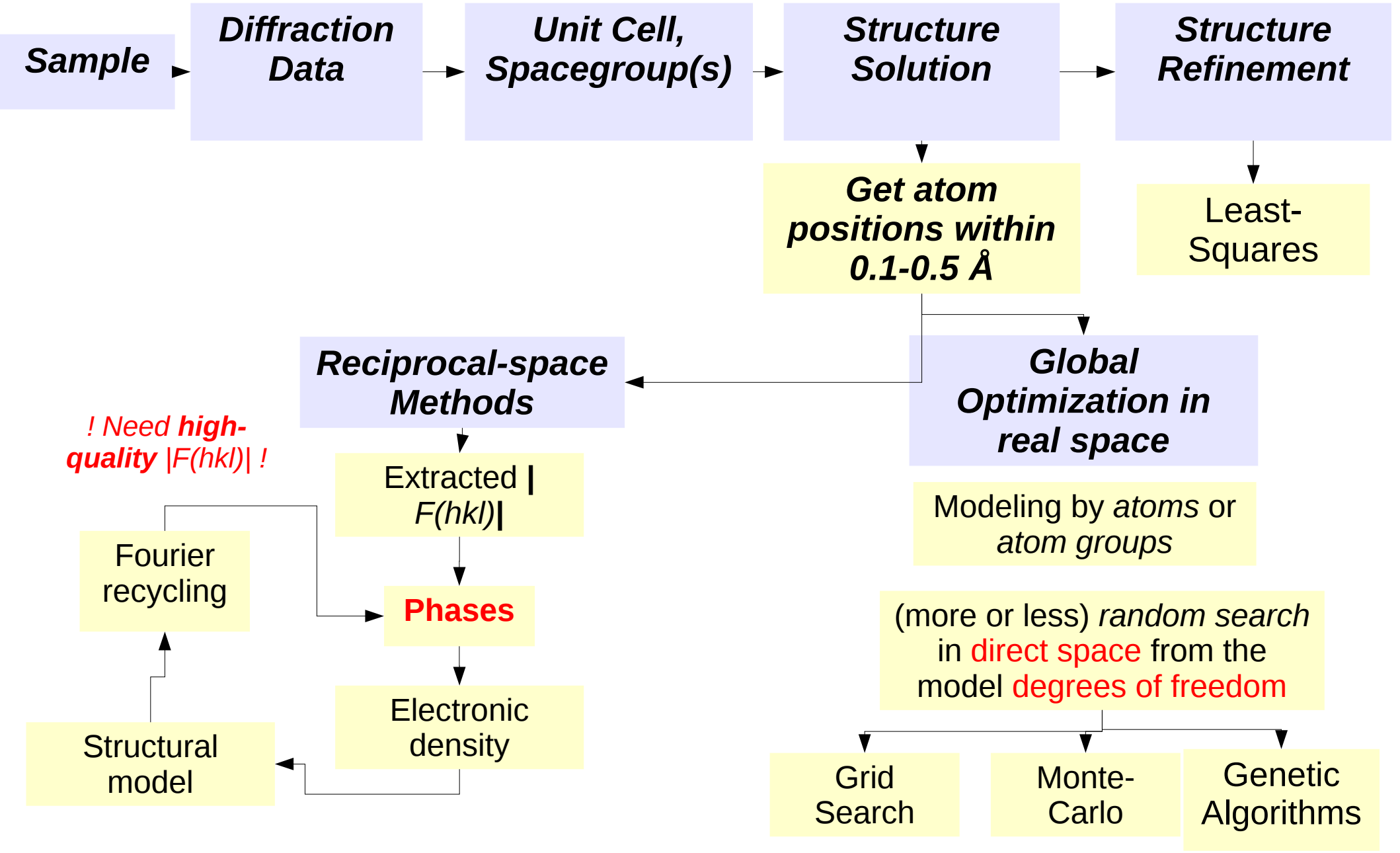
- twinning/reflexion overlap
- no high resolution data available

*Direct methods are powerful (using full crystallographic formalism to derive the electronic density in seconds) but may not **recover from bad structure factors***

*Real-Space structure solution: try **many** configurations until a satisfactory one is found*
*=> **brute-force approach** enabled by the increase in computing power*
*A **basic** but **robust** approach to structure solution*

Limited requirements on data resolution :
need more observed |Fhkl| than parameters (preferably many more)
usually, a resolution of 2.5 Å is enough for most small molecules/inorganic structures, less if rigid bodies are used

Solving Structures



Specifics of Fox

Parametrization

- inorganic or organic materials
- description using atoms, polyhedra, molecules
- automatic, smooth correction of special positions

Data

- powder pattern (X-Ray, neutron, multi-phase, TOF, electron)
- single crystal
- joint optimization with several data sets
- use integrated prof les (no need to extract F(*hkl*))

Algorithms

- Parallel Tempering (Simulated Annealing)
- yields multiple solutions
- expandable to new algorithms

Other uses

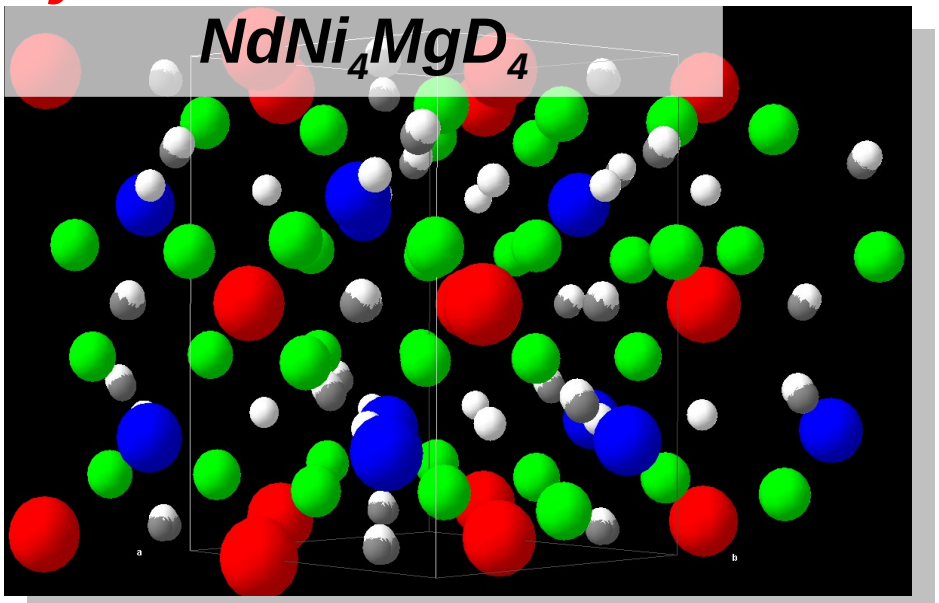
- display Fourier maps (from gsas/expgui or internally)
- simulation of powder & single crystal diffraction
- display data (crystal, powder) from CIF f le

Availability

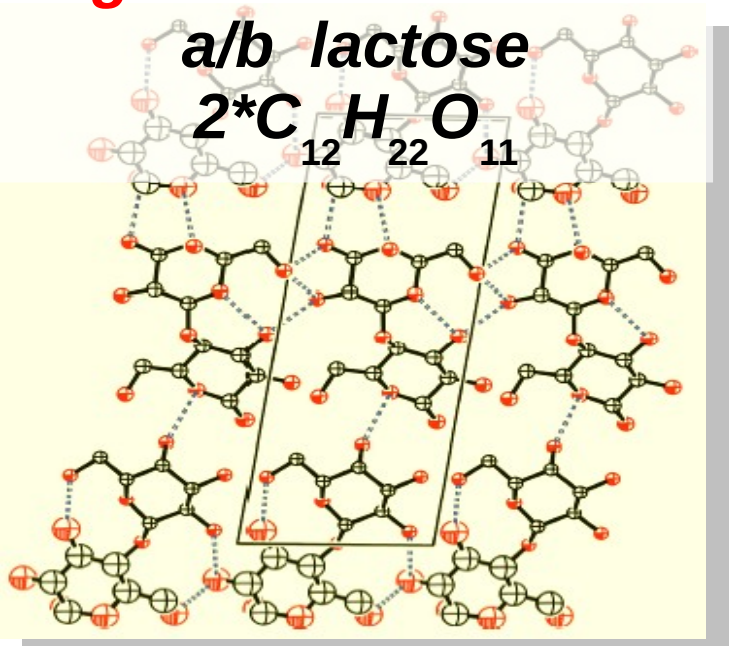
- free (http://objcryst.sourceforge.net)
- open source (GPL)
- available for *Linux*, *MacOS X* and *windows*

Examples of structures solved

Hydrides:

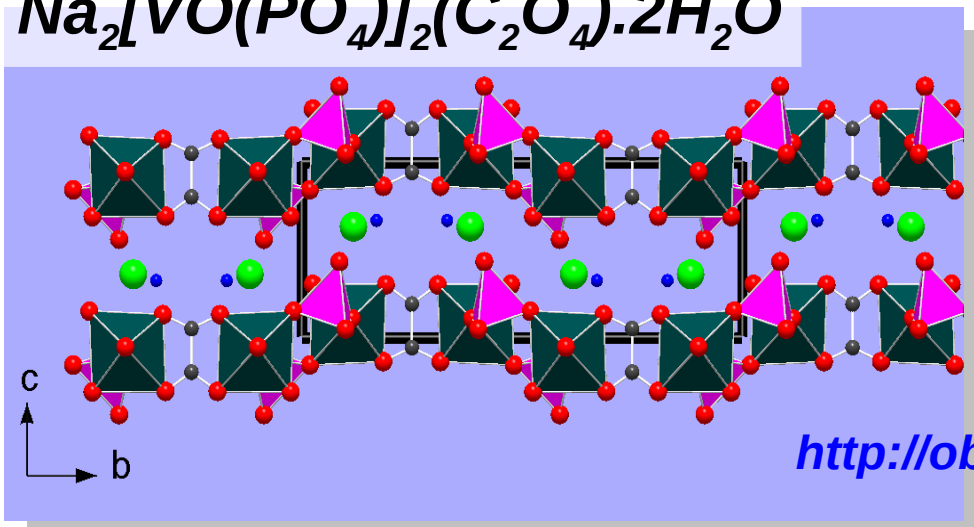


Organic



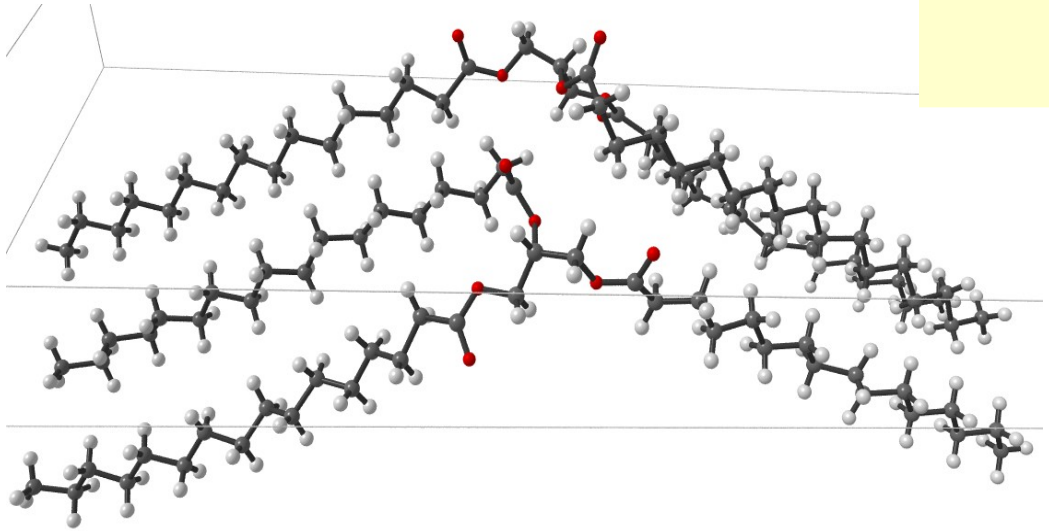
Inorganic:

$Na_2[VO(PO_4)]_2(C_2O_4).2H_2O$



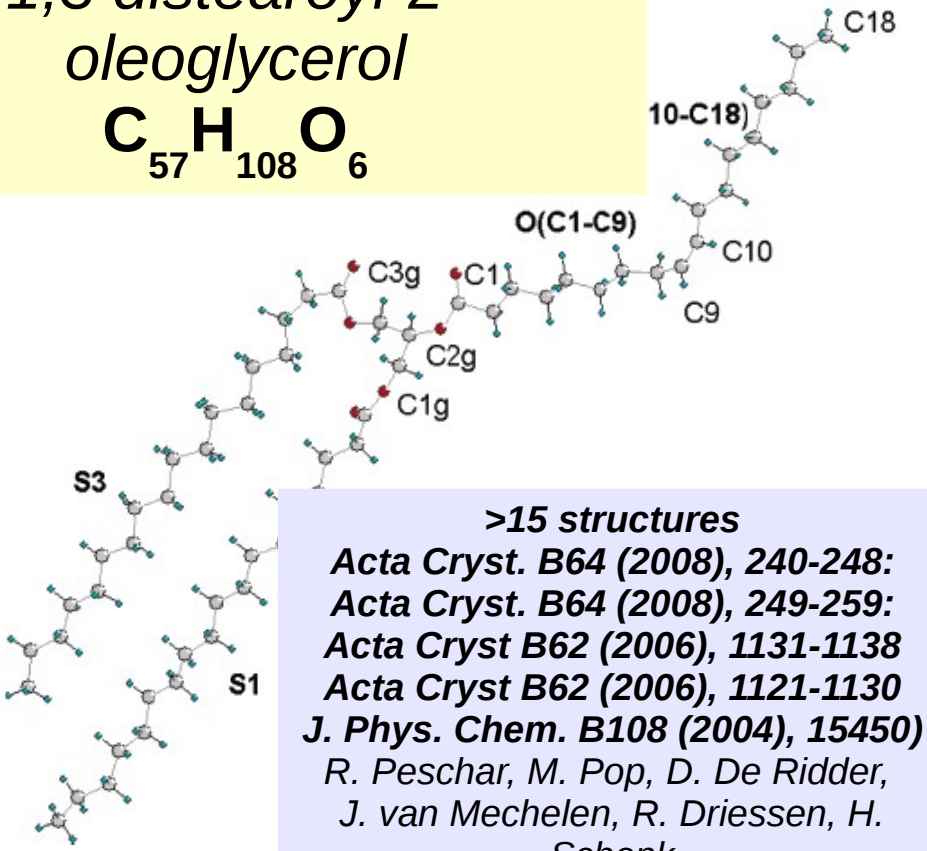
http://objcryst.sourceforge.net/Fox/FoxBiblioStructures

Triglycerides



1,3 distearoyl-2-oleoglycerol

$C_{57}H_{108}O_6$



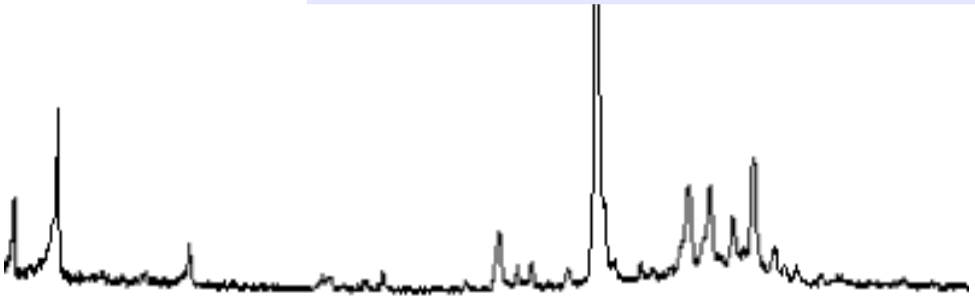
>15 structures
Acta Cryst. B64 (2008), 240-248:
Acta Cryst. B64 (2008), 249-259:
Acta Cryst B62 (2006), 1131-1138
Acta Cryst B62 (2006), 1121-1130
J. Phys. Chem. B108 (2004), 15450)
R. Peschar, M. Pop, D. De Ridder, J. van Mechelen, R. Driessen, H. Schenk
FOX + ORGANA

β' PSP (1,3-di-n-hexadecanoyl-2-n-octadecanoyl glycerol)

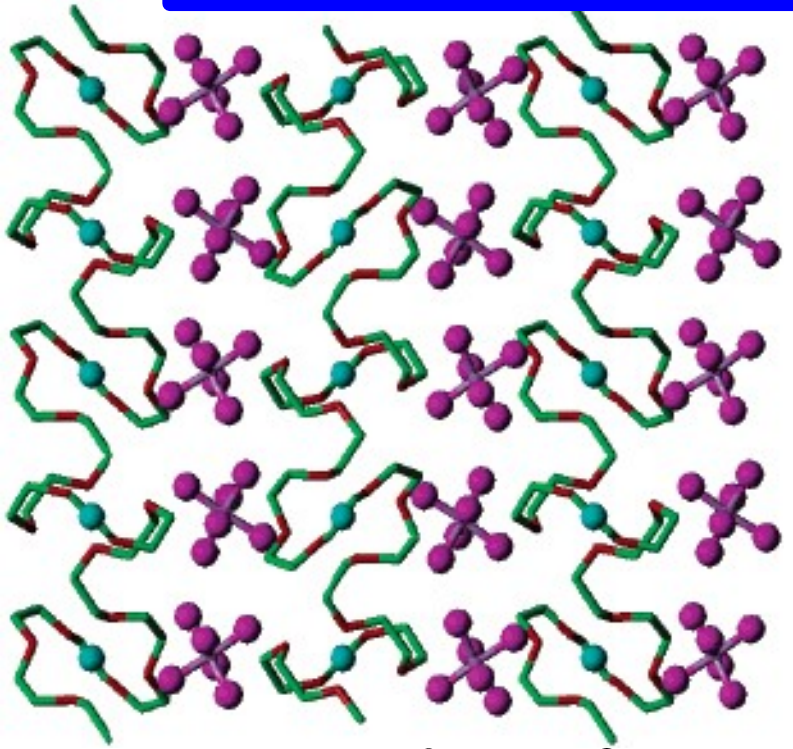
$C_{53}H_{102}O_6$

up to 56 non-H free torsion angles !

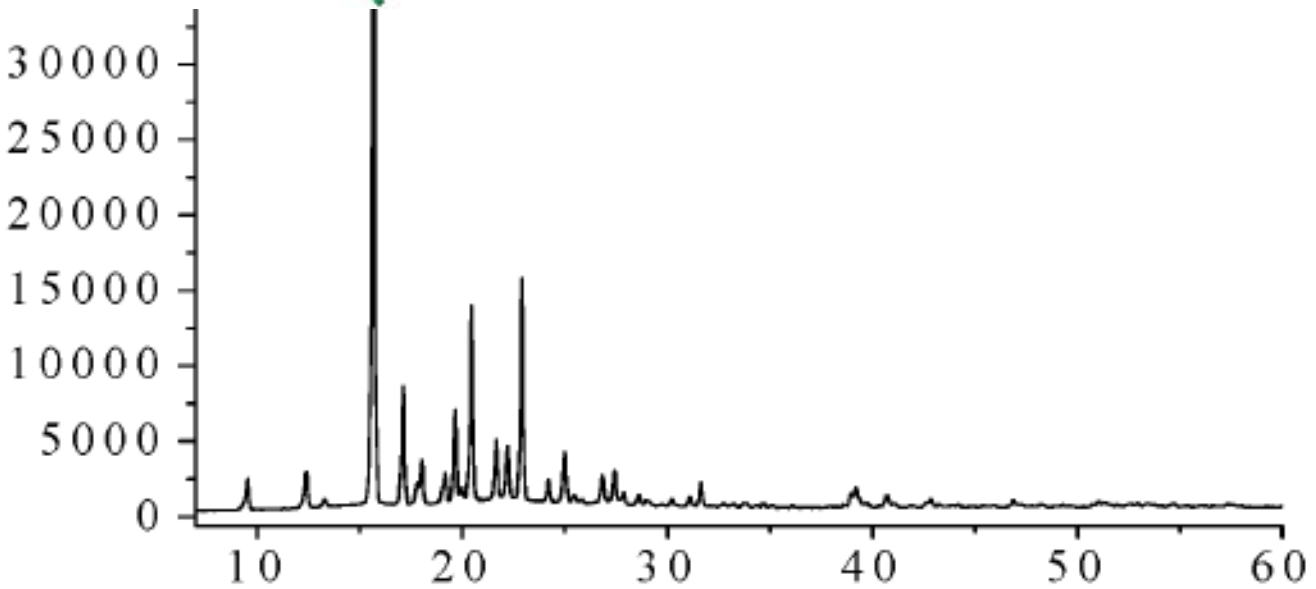
FOX > 2 months



Polymer Electrolyte β-PEO_6:$LiAsF_6$



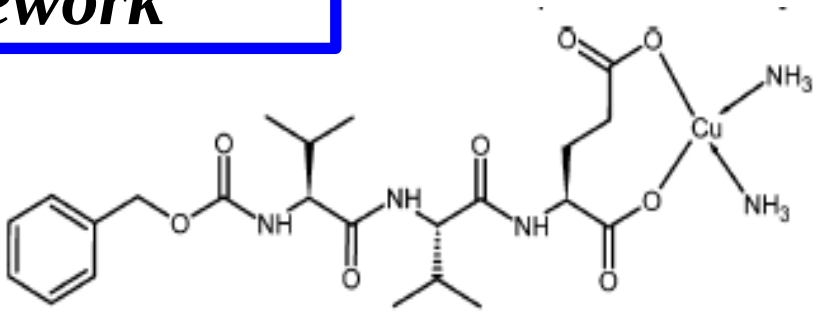
27 non-H atoms
JACS 127 (2005), 12176-12177
E. Staunton, Yu. G. Andreev, and P. G. Bruce



Metal-Peptide Framework

JACS (02/2008) DOI:10.1021/ja0762588

A. Mantion, L. Massüger, P. Rabu, C. Palivan, L. B. McCusker & A. Taubert



37 independent atoms
~ 20 internal DOF, 1 flexible 9-atom ring

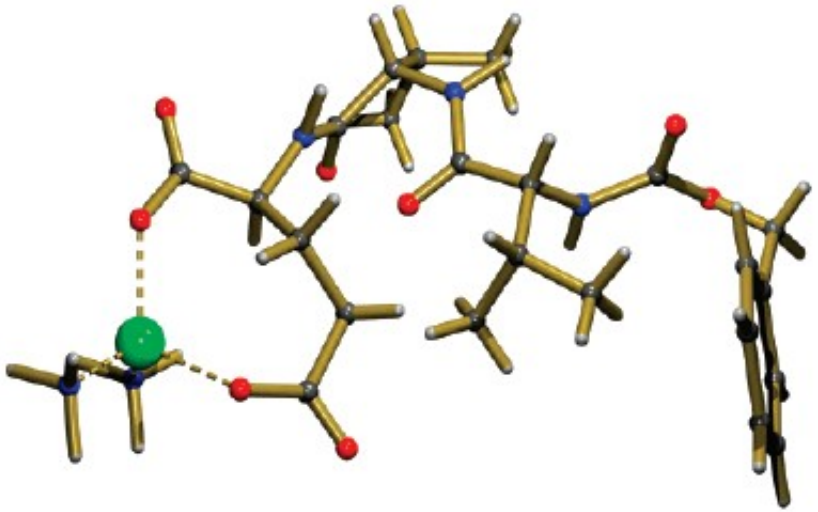


Figure 9. Crystal structure of MPF-9 generated by FOX and before Rietveld refinement, including hydrogen atoms.

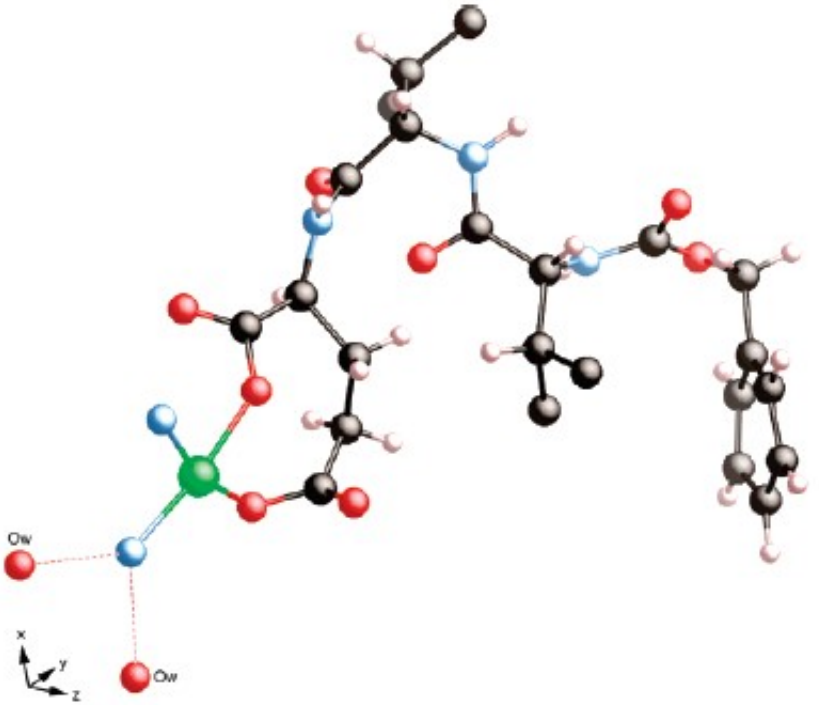
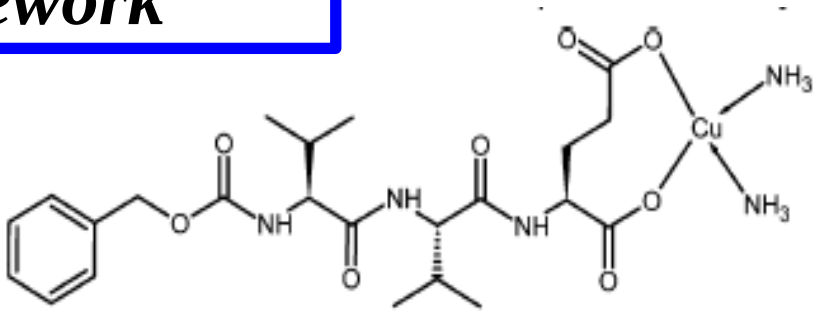


Figure 10. Crystal structure of MPF-9 after refinement of the model generated by FOX, including the water molecule.

Metal-Peptide Framework

JACS (02/2008) DOI:10.1021/ja0762588

A. Mantion, L. Massüger, P. Rabu, C. Palivan,
L. B. McCusker & A. Taubert



37 indep atoms
~ 20 internal DOF, 1 flexible 9-atom ring

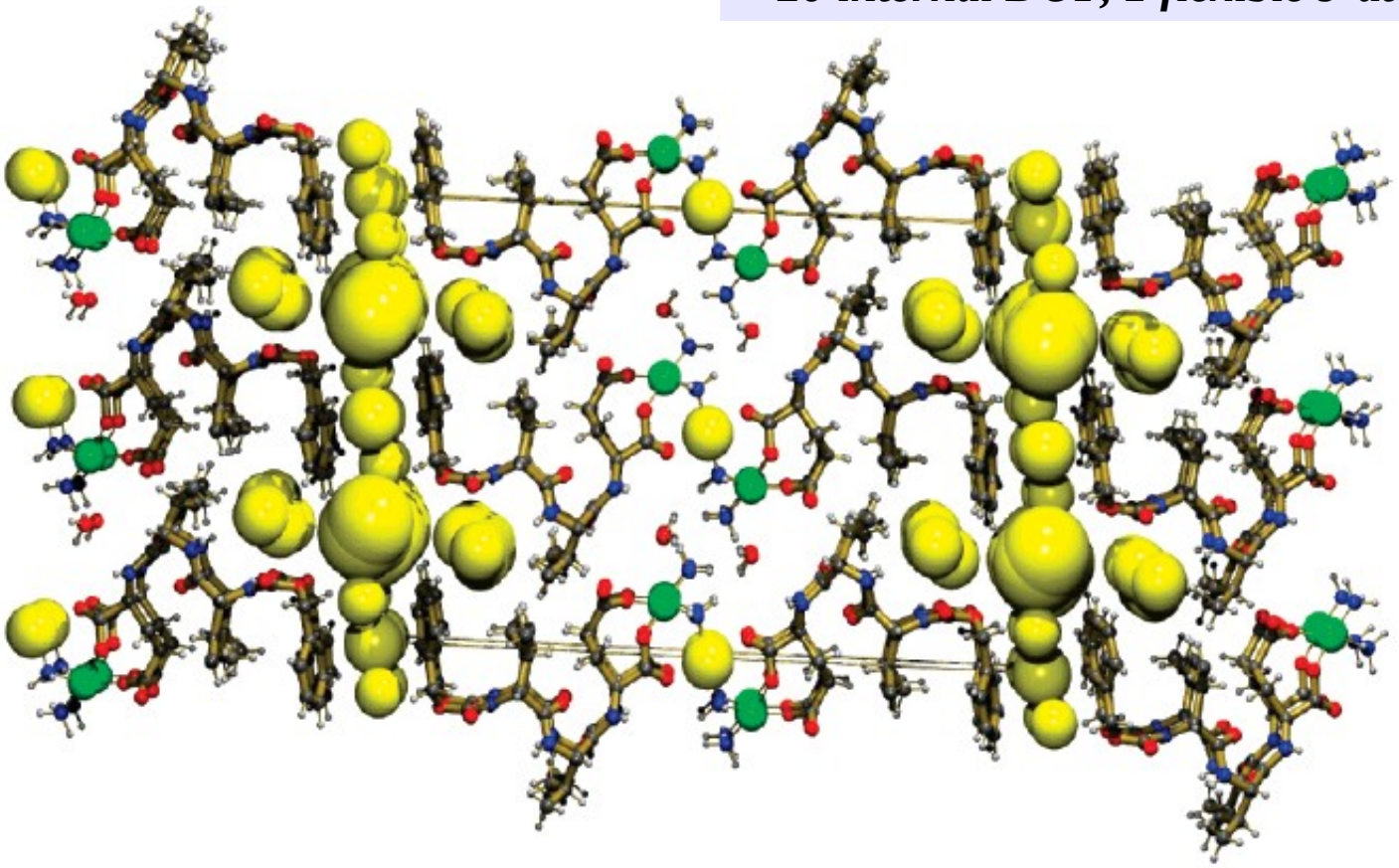


Figure 15. Packing diagram of MPF-9. Yellow balls indicate the voids.

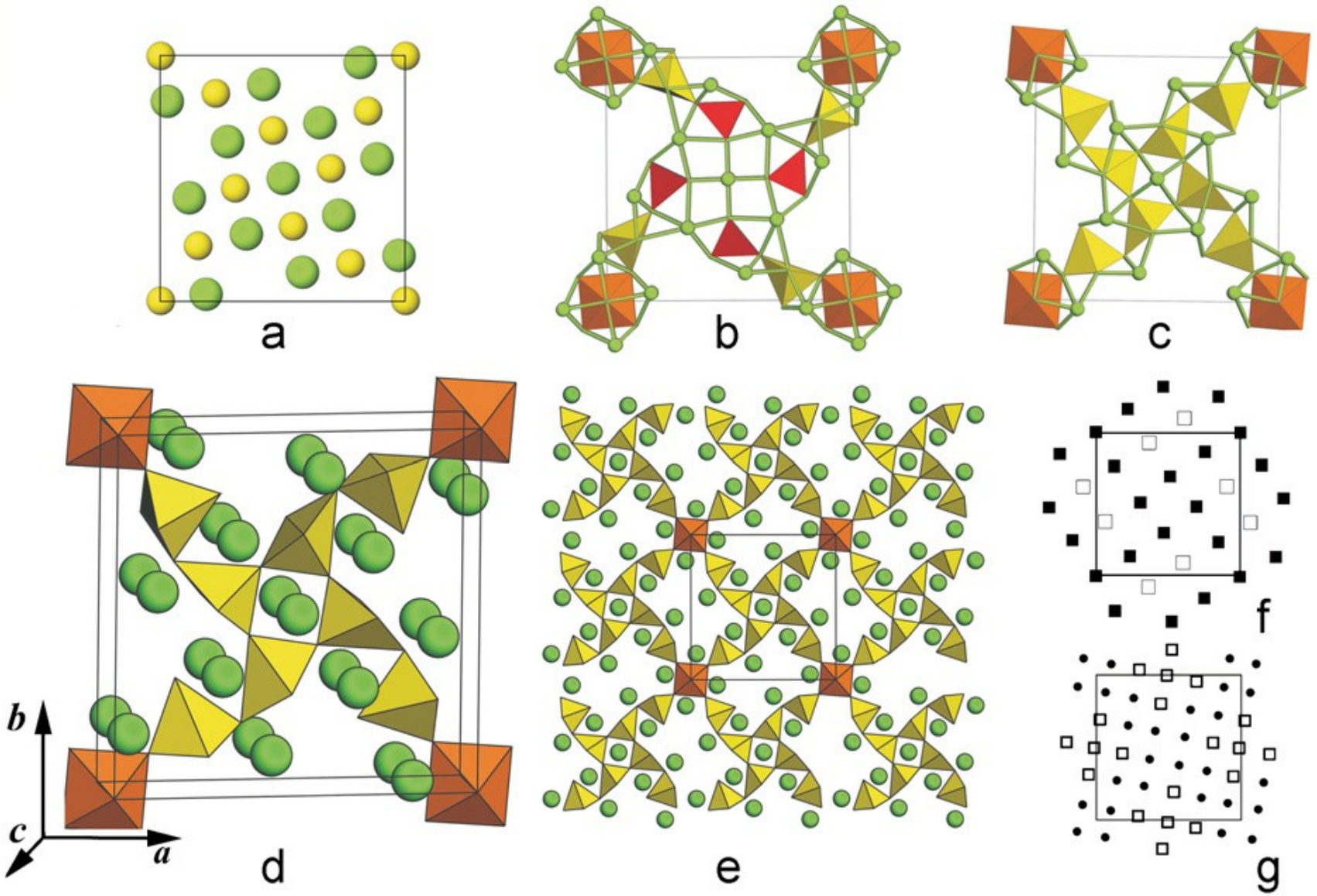
Electron diffraction

$Pb_{13}Mn_9O_{25}$ precession electron diffraction data
*P*4/*m*, *Z* = 1

J. Hadermann et al., *Ultramicroscopy* **110** (2010) 881–890

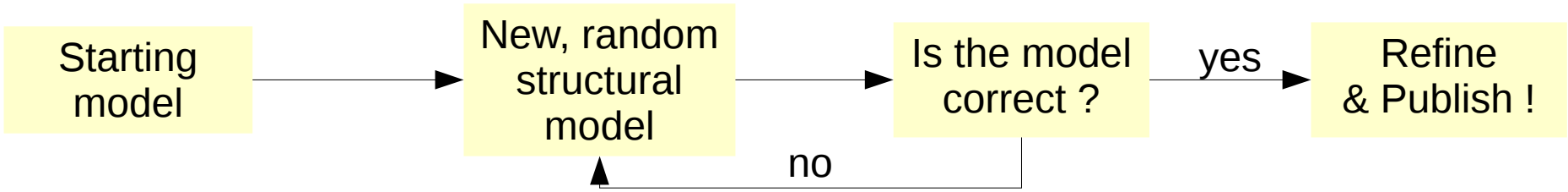
a) Pb and Mn from direct methods (SIR2008)
b,c) O localized by FOX using:
- antibumps
- BVS cost function

c,d) Rietveld and DFT confirmed the correct model



Real-Space Exploration ?

Basic Direct-space Algorithm (trial & error):



A good program requires :

... this works but is only adequate for the very patient crystallographer !

Criteria *to test the validity of the the model*

An ergodic algorithm *which:*

- *can explore* ***every possible*** *structural model*
- *will spend more time "close" to the real solution (efficient* ***biasing****)*

A flexible modeling of the structure*:*

- *allowing to reduce the number of parameters (****Degrees of Freedom****)*
- *able to describe any structure*
- *allow easy configuration changes*

Criteria for Minimization

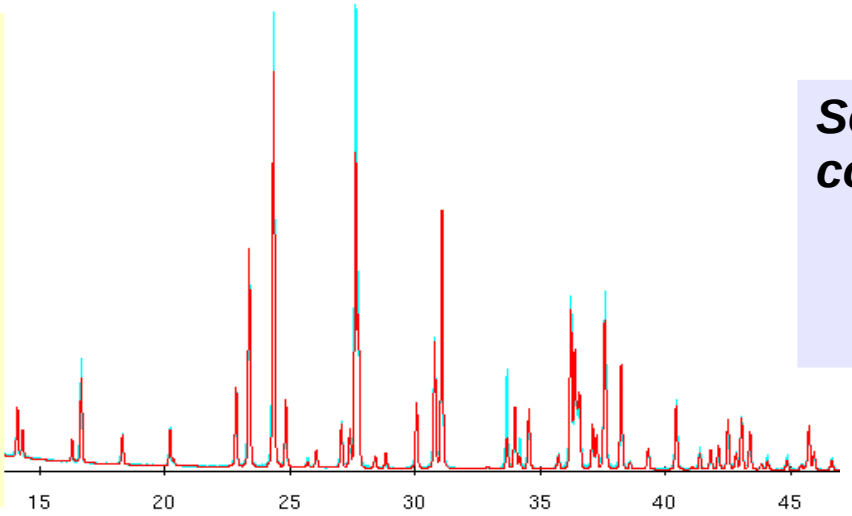
Criteria to Evaluate Trial Structural Models Diffraction Data

R-factor

$$R_{wp} = \sqrt{\frac{\sum w_i \left(I_i^{obs} - I_i^{calc}\right)^2}{\sum w_i \left(I_i^{obs}\right)^2}}$$

or c^2

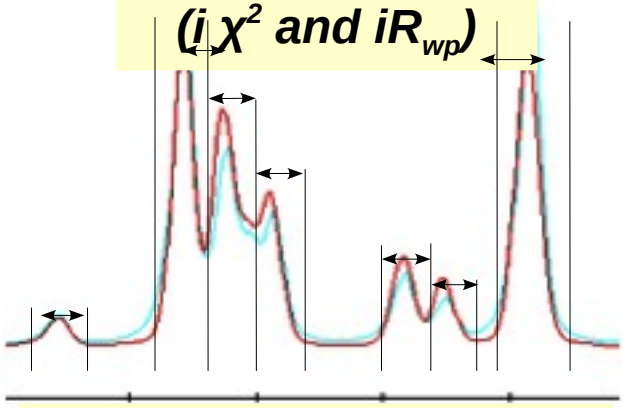
$$\chi^2 = \sum \frac{1}{\sigma_i^2} \left(I_i^{obs} - I_i^{calc}\right)^2$$



Several datasets can be combined:

X-rays & neutrons
several wavelengths,
Several temperatures ...

integrated profiles (iχ^2 and iR_{wp})



Integrated profiles allow to avoid the requirement of a perfect description of profiles

Why not use extracted structure factors (faster & equivalent) ?

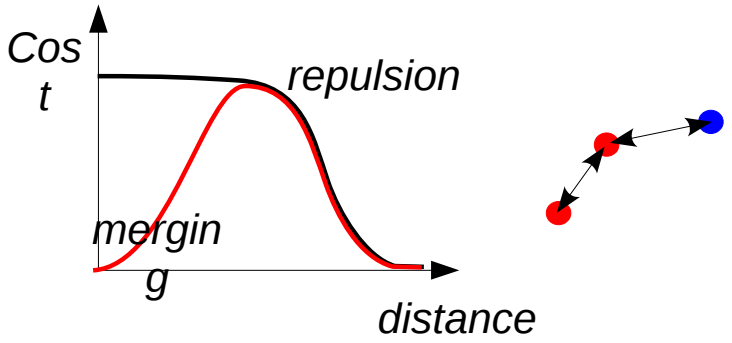
*- This would require a **perfect description** of the profiles and background, which can be difficult ("real" samples, with ill profiles and multiple phases, background difficult to "guess" for close-packed reflections).*

- Direct-space algorithms are necessary for samples where the extraction of structure factors is difficult

- with "integrated profiles", the full pattern is not calculated and the speed is equivalent to extracted structure factors

Criteria to Evaluate Trial Structural Models
AntiBump Restraint

Anti-bump



*An **AntiBump** function allows the repulsion of atoms while permitting the "merging" of identical atoms on special positions or connecting several polyhedra*

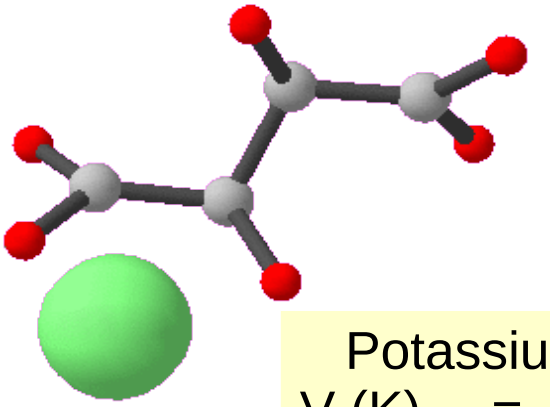
Energy calculations ?
Either internal energy for molecules or global for the entire unit cell

… But energy calculations are extremely costly from a computation point of view

Criteria to Evaluate Trial Structural Models
Bond Valence

$$V_i = \sum_{neighbours\ j} e^{\frac{(R_1 - r_{ij})}{0.37}}$$

The valence of each atom depends on its neighbours and their distance :



Potassium tartrate :
$V(K)_{calc}$ = + 0.991 OK!

***Warning**: bond valence / AntiBump calculations use a lot of computing power (as much or more than structure factor computation)*

=> only use them if the diffraction data is not of sufficient quality to solve the structure

Criteria to Evaluate Trial Structural Models
Combining Several Criteria

Problem: different criteria will have different scales !!

When combining experimental data, χ^2 can be summed:

$$\chi^2 = \sum_{data\,1} \frac{1}{\sigma_i^2}\left(I_i^{obs} - I_i^{calc}\right)^2 + \sum_{data\,2} \frac{1}{\sigma_i^2}\left(I_i^{obs} - I_i^{calc}\right)^2 + \ldots$$

=> avoid using R-factors which cannot be summed

Fringe benefit: using χ^2 makes you ready for maximum likelihood (ML)

Sometimes combining « incompatible » criteria (χ^2, energy, antibump) is necessary
=> finding the correct scale can be difficult.
=> correct scale factors can be guess if you know the 'target' values :
e.g. χ^2 should converge towards ***Nobs*** (Goodness-Of-Fit=1), antibump towards 0, etc..

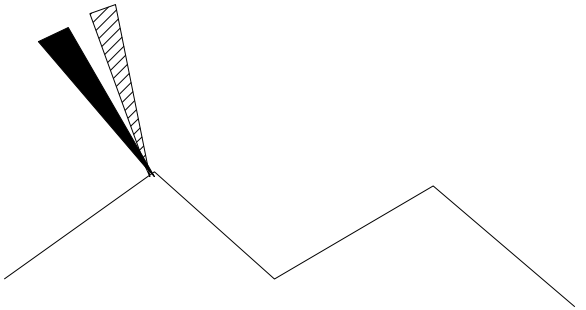
Sometimes scaling different data sets is necessary (e.g. combine powder diffraction data from synchrotron and neutron) : statistically, no scale should be applied, but for « global optimisation » algorithms rules may be bent (see later ML slides)

Model Building:
Real Space Parametrization

Structural Description from Building Blocks: the Z-matrix approach

The number of trials required varies exponentially with the number of parameters
=> ***need to use all the*** **a priori** ***information about the atomic coordination***

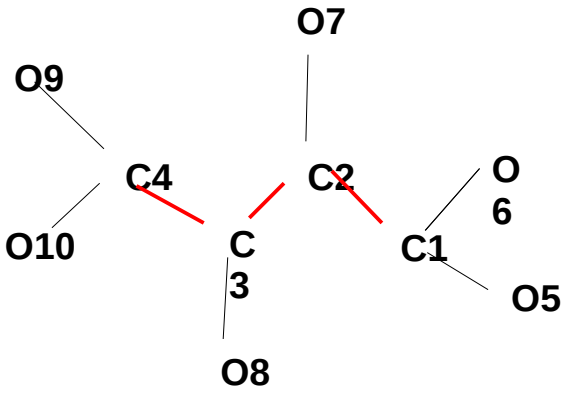
Building blocks ***for the crystal structure***



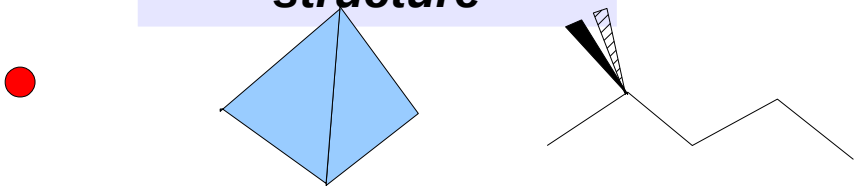
atom

Structural Description from Building Blocks: the Z-matrix approach

The number of trials required varies exponentially with the number of parameters
=> *need to use all the* a priori *information about the atomic coordination*



***Building blocks* for the crystal structure**



atom ***Polyhedron*** ***molecule***

Description from bond lengths, bond angles and dihedral angles

Z-Matrix

number of atoms

```
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
```

atom type

free torsion angles

bond ***length*** with atom #

bond ***angle*** with atom #

dihedral angle with atom #

30 -> 6+5=11 DOF

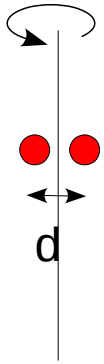
Dynamical Occupancy Correction

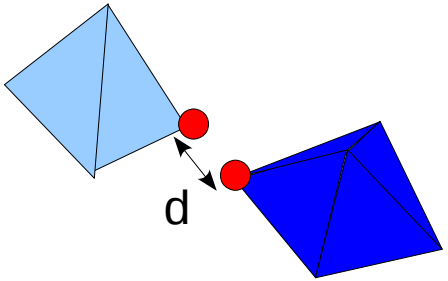
*Inorganic structures often have atoms in **special positions**, and have **atoms common to several polyhedra**.=> New algorithm which must:*

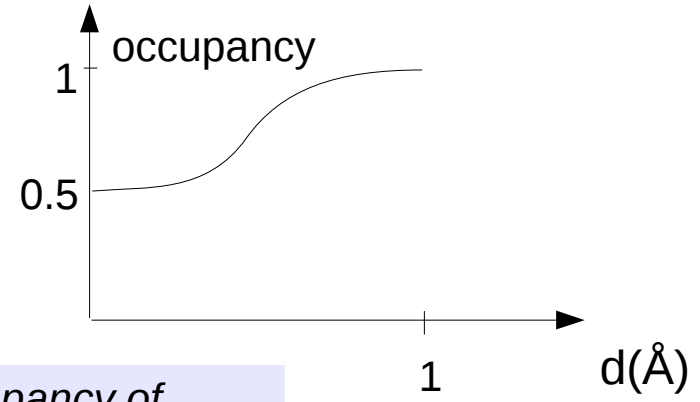
*- correctly describe atoms in special positions **without a priori knowledge***

- allow the atoms to move continuously to and from the special positions

- not depend on the type of compound or the modelling chosen (atoms, polyhedra, molecules)



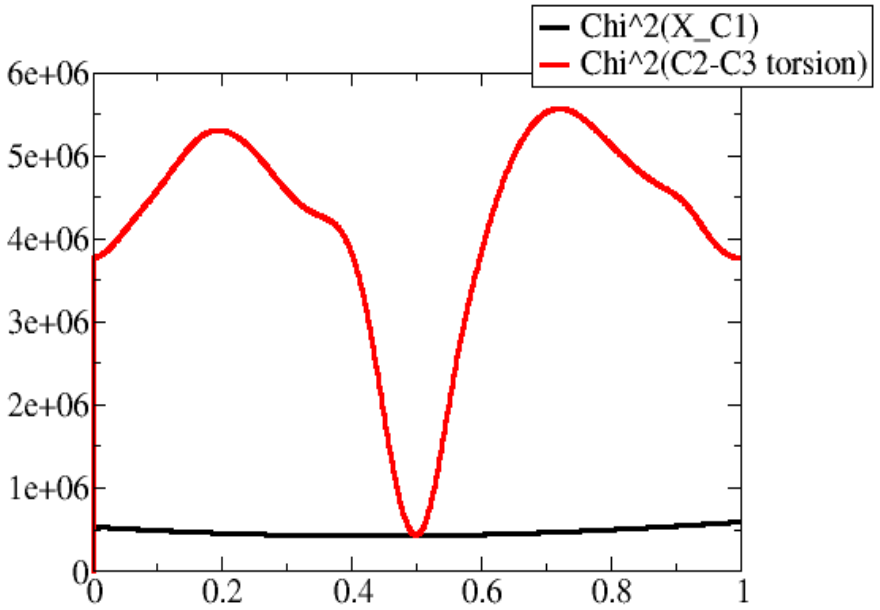


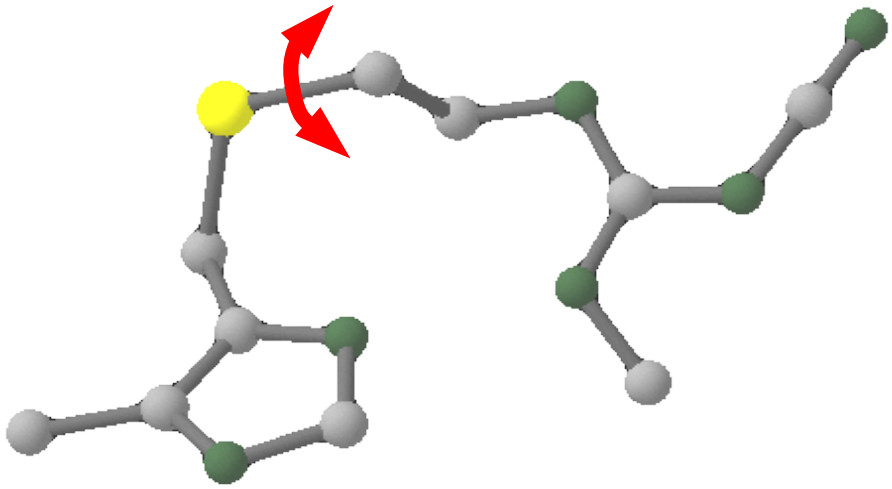


Diminish the occupancy of overlapping atoms as they are getting closer

! Warning ! Do not use the DOC if no special positions or shared atoms are expected, as it slows down the computation.

Pitfalls of internal coordinates (z-matrix)





a **torsion angle** *(moving many atoms) has a* ***much narrower minimum*** *than a translation parameter of an individual atom*

=> even if the number of degrees of freedom diminishes, the global minimum is much narrower

Atoms are deduced from previous atoms => the first atoms 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

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

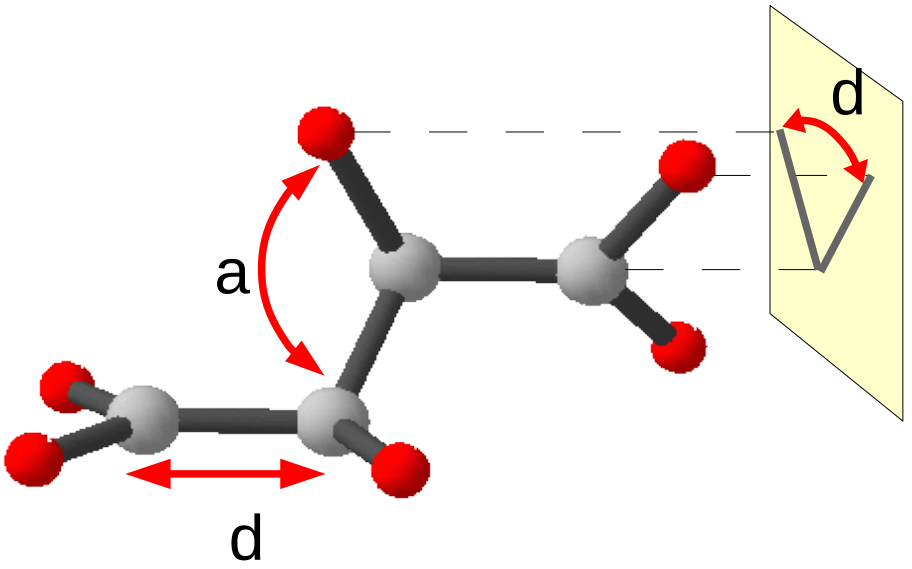
Flexible Approach using Restraints

idea**: keep all the coordination information, but with a **flexible approach

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

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

The orientation of the molecule is defined by a **quaternion** (to avoid "gimbal lock" angles)

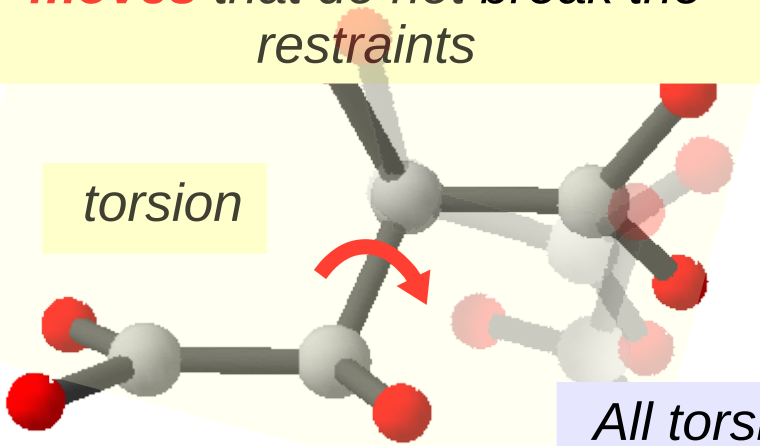


- *this modelization is **independent from the order of the atoms***
- ***any type of restraint** can be introduced*
- ***any type of movement** can be directly done (no need to compute complex torsions)*
- *any **cycle** can be defined*

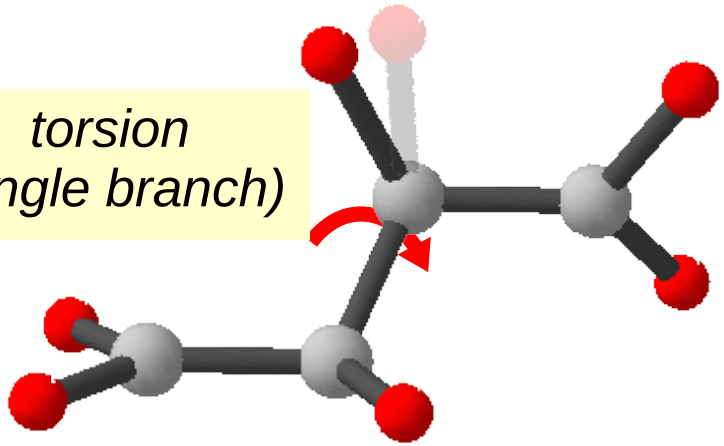
Making the Smart Moves

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

torsion

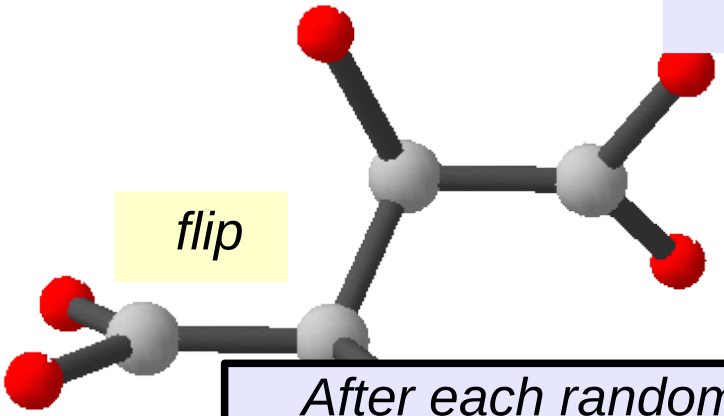


torsion (single branch)

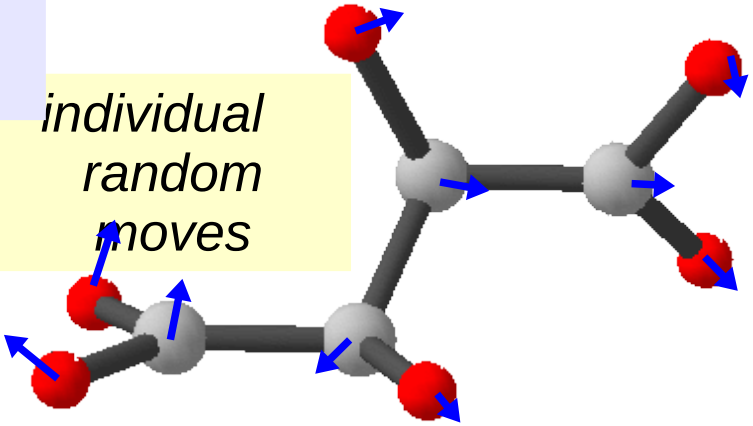


*All torsion & flip moves that do not break restraints are **automatically identified***

flip



individual random moves



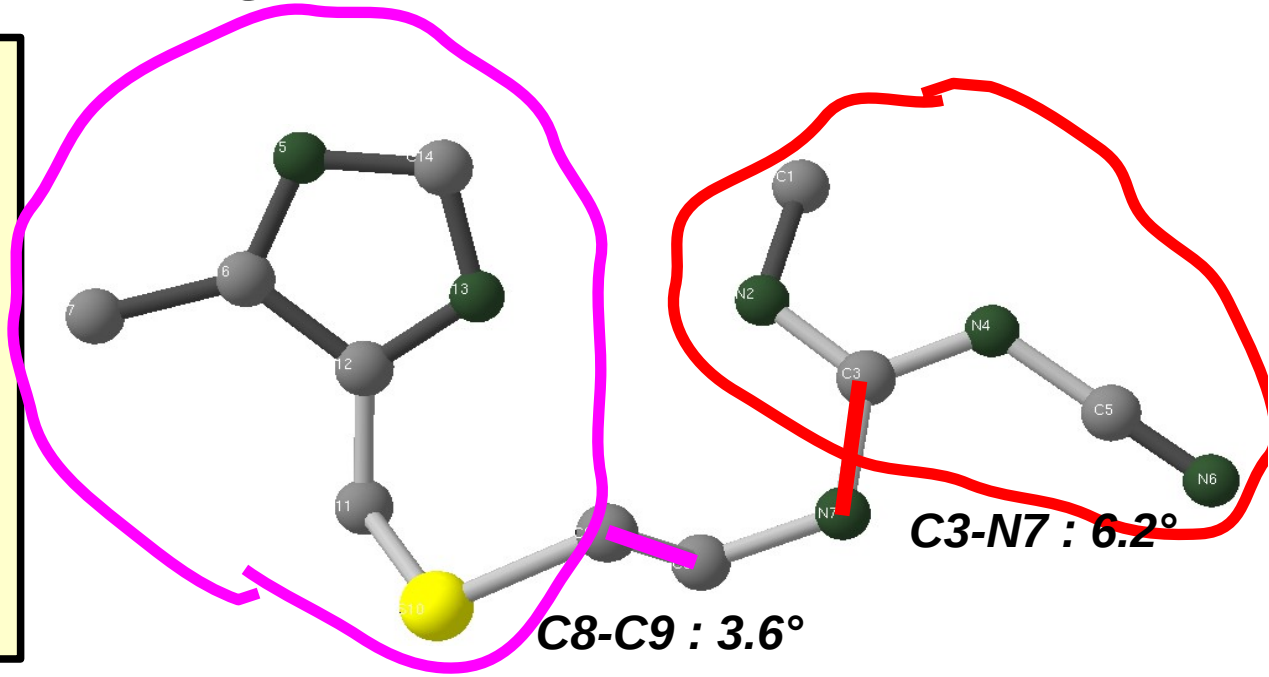
After each random move, a test is made on the total internal restraint cost to see if the configuration is kept

Adaptative Conformation Changes

Random torsion angle changes :
- rotate the ***smallest fragment***
- ***tune the max. rotation*** *so that the* ***average displacement is 0.1Å****.*

- same for bond angle changes

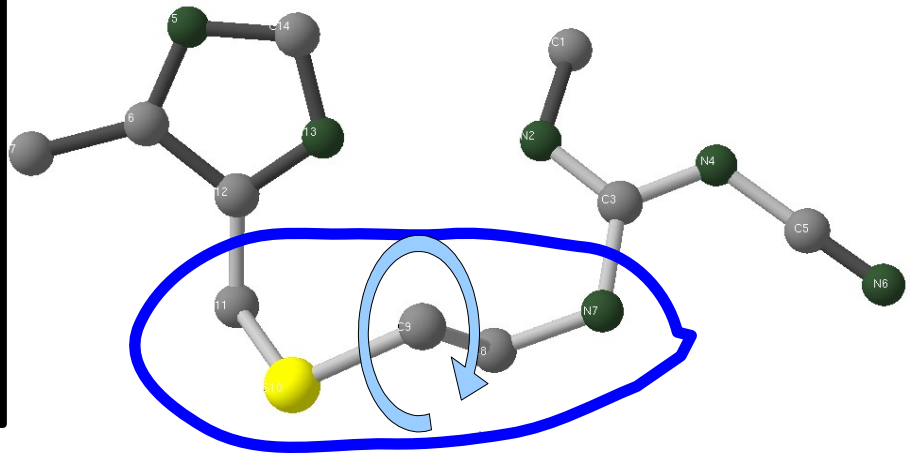
- tune global rotation of molecule



" Twist " mode :
alter an ***internal*** *part of a chain/cycle*

=> long chains, flexible cycles

TODO*: determine "* ***soft modes*** *" of the molecule and use them to distort the molecule (computationnally costly)*



Cimetidine

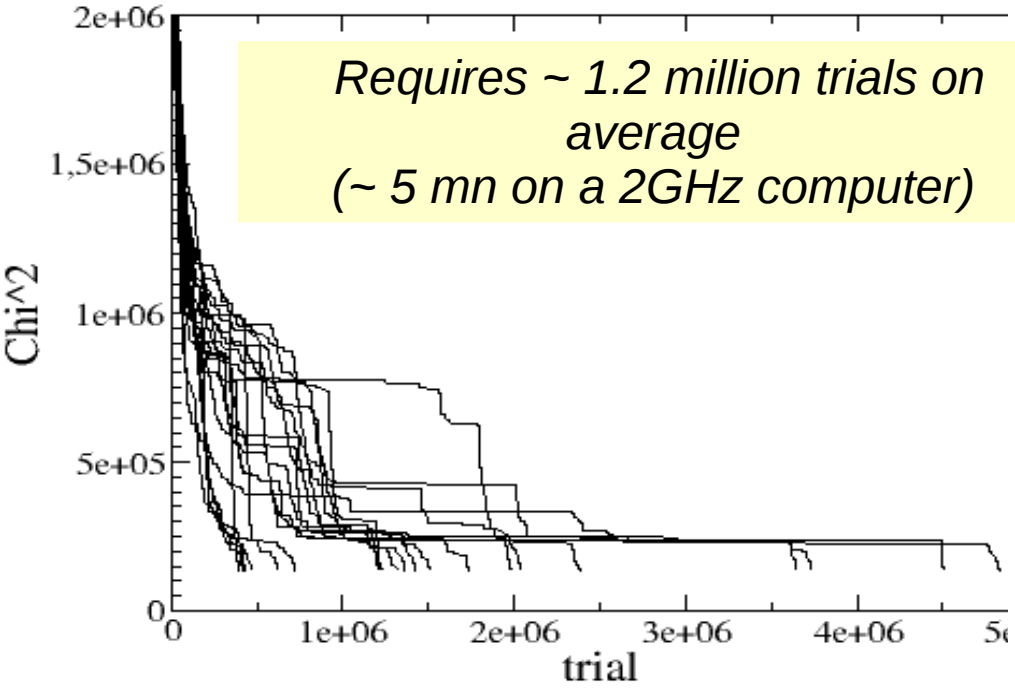
A well-known testcase for ab initio structure determination from powder diffraction

17 non-H atoms (8+6=14 DOF)

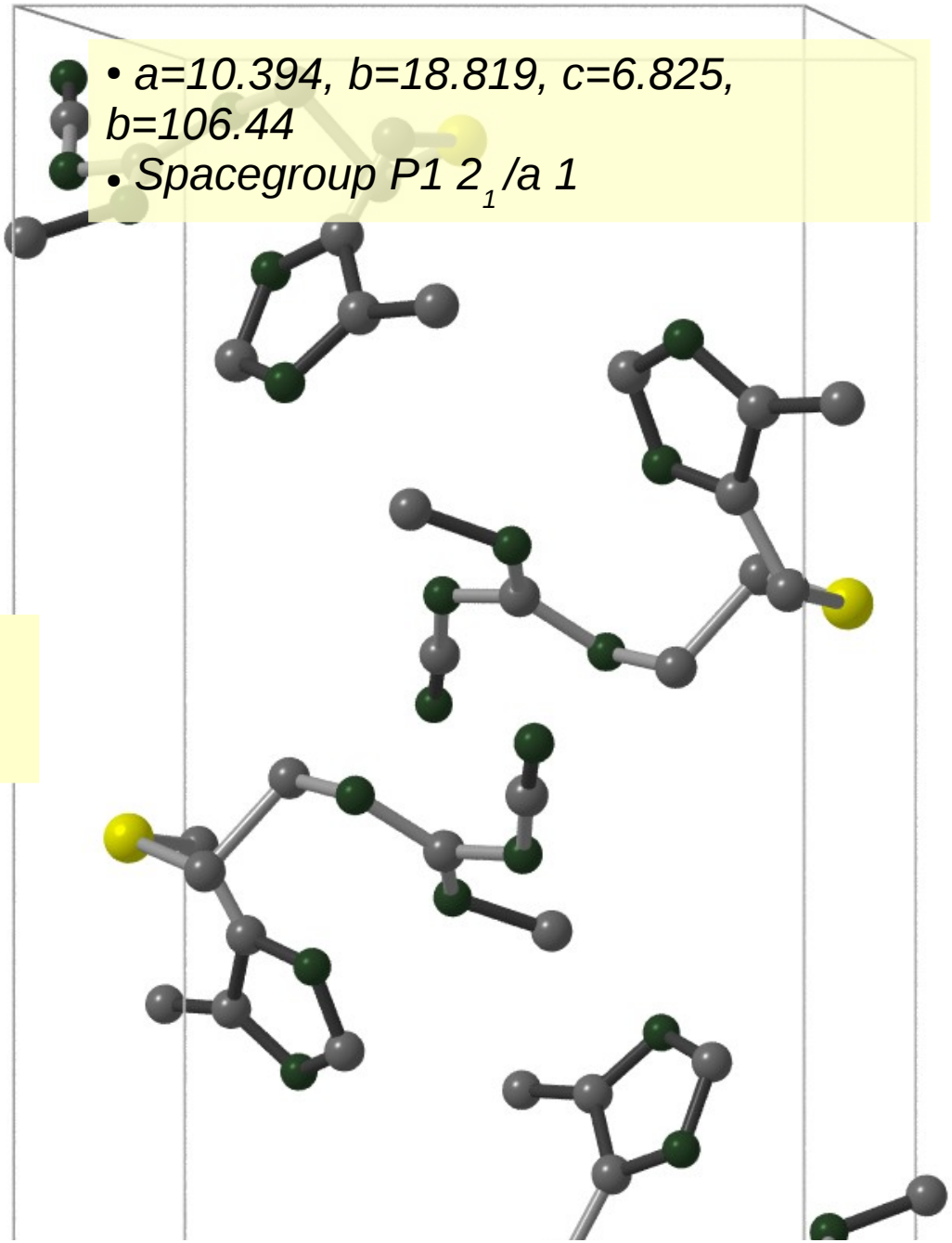
Cernik et al. *J. Appl. Cryst* ***24*** (1991), 222

- *a=10.394, b=18.819, c=6.825, b=106.44*
- *Spacegroup P1 2_1/a 1*

Results : ***up to 3 times faster than with the z-matrix model***



Requires ~ 1.2 million trials on average
(~ 5 mn on a 2GHz computer)



Reverse Monte-Carlo

Parametrization -> Degrees Of Freedom

starting configuration

random change of parameters

evaluation of the new configuration: ***Cost (C)***

is configuration better ? $C_n < C_{n-1}$

yes

keep configuration

no

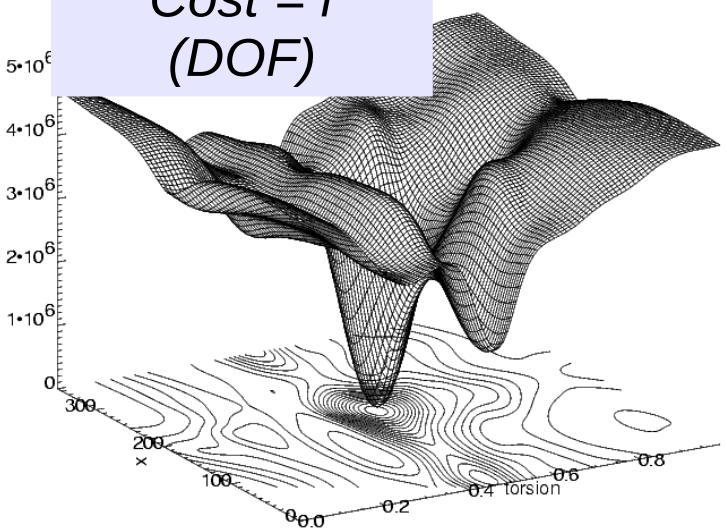
keep configuration with probability:

$$P = e^{\frac{-\Delta C}{T}}$$

Temperature of the algorithm

Generate a distribution of configurations following ***Boltzmann****'s law*

Hypersurface Cost = f (DOF)



Reverse Monte-Carlo

Parametrization -> Degrees Of Freedom

starting configuration

random change of parameters

evaluation of the new configuration: ***Cost (C)***

is configuration better ? $C_n < C_{n-1}$

yes

keep configuration

no

keep configuration with probability:

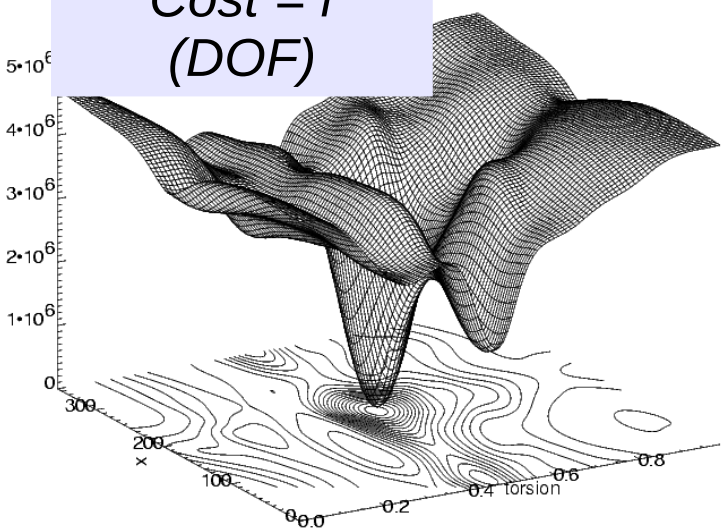
$$P = e^{\frac{-\Delta C}{T}}$$

Temperature of the algorithm

Generate a distribution configurations following ***Boltzmann****'s law*

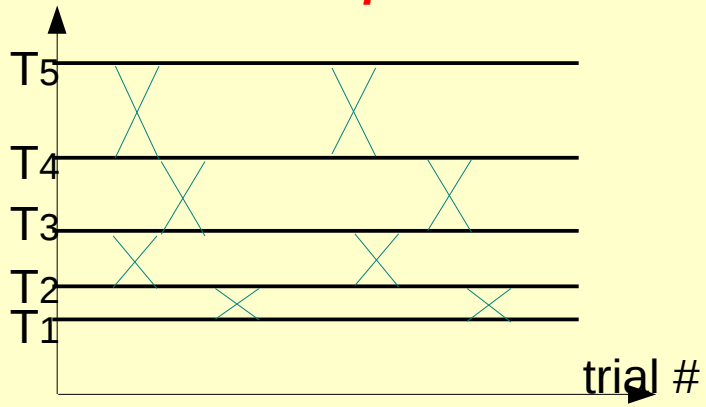
Simulated Annealing *: decrease T as a function of the trial number to converge towards the global minimum*

Hypersurface Cost = f (DOF)



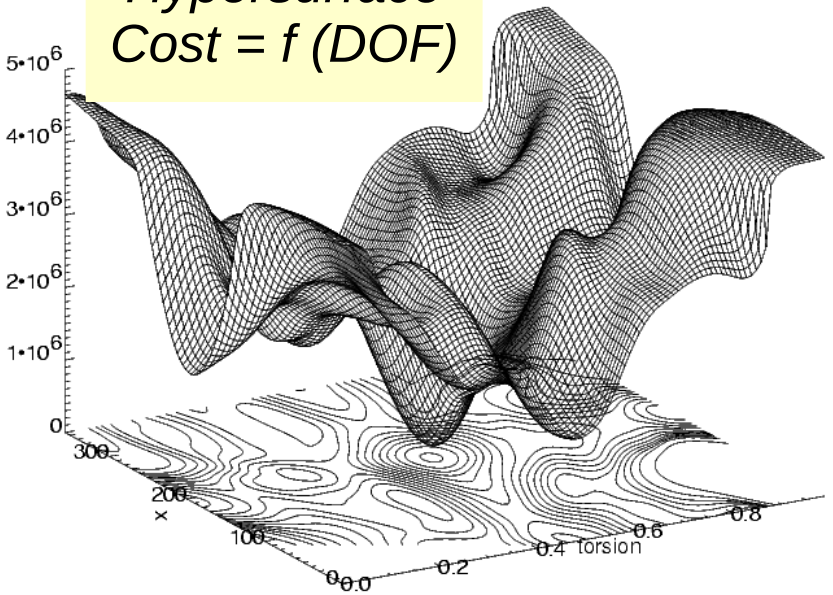
Parallel Tempering & Annealing Temperatures

simultaneous optimization at different temperatures



Using several parallel optimizations at different temperatures ensures that ***the algorithm can get out of any local minimum****.*
Furthermore, it does not require to predict an adequate decrease rate for the temperature.

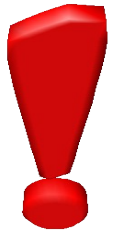
Hypersurface
Cost = f (DOF)



To automatically choose the temperatures, in each parallel optimization ***it is the average atomic displacement per random move which is imposed, from 0.01 to 1 Å****.*
The Temperature is then tuned so that in each "world" ***the acceptance rate of new configurations is from 10 to 30%****.*

Maximum Likelihood & Global Optimization

Maximum Likelihood



WARNING :
Approximations !
(Theorists hold your fire !)

In a "classical approach" :

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

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

*assumes that the model can fit **perfectly** the observed data.*
*But there can be **errors** in the model !*

*typically **positionnal errors** during the search for a structure solution*

with a positionnal error measured by: $D(\vec{k}) = \langle \cos(2\pi \vec{k}.\Delta\vec{r}) \rangle$

*introduce a **variance** on the calculated structure factor*

$$\sigma^2_{calc} = \left(1 - D^2\right) \sum_{atoms} f^2_j$$

*Use the "**most likely**" calculated structure factor*

$$\langle F_{calc} \rangle = D F_{calc}$$

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

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

Application to Global Optimization

1st application:

incomplete model

missing atoms (H's, solvant) do not contribute to the ***Structure Factor*** *but increase the* ***variance***

$$D(\vec{k}) = \langle \cos(2\pi \vec{k}.\Delta\vec{r}) \rangle = 0$$

$$\langle F_{calc} \rangle = DF_{calc} = 0$$

$$\sigma^2_{calc} = (1 - D^2) \sum_{atoms} f_j^2$$

Markvardsen,Acta Cryst A58(2002)

2nd application:

model errors

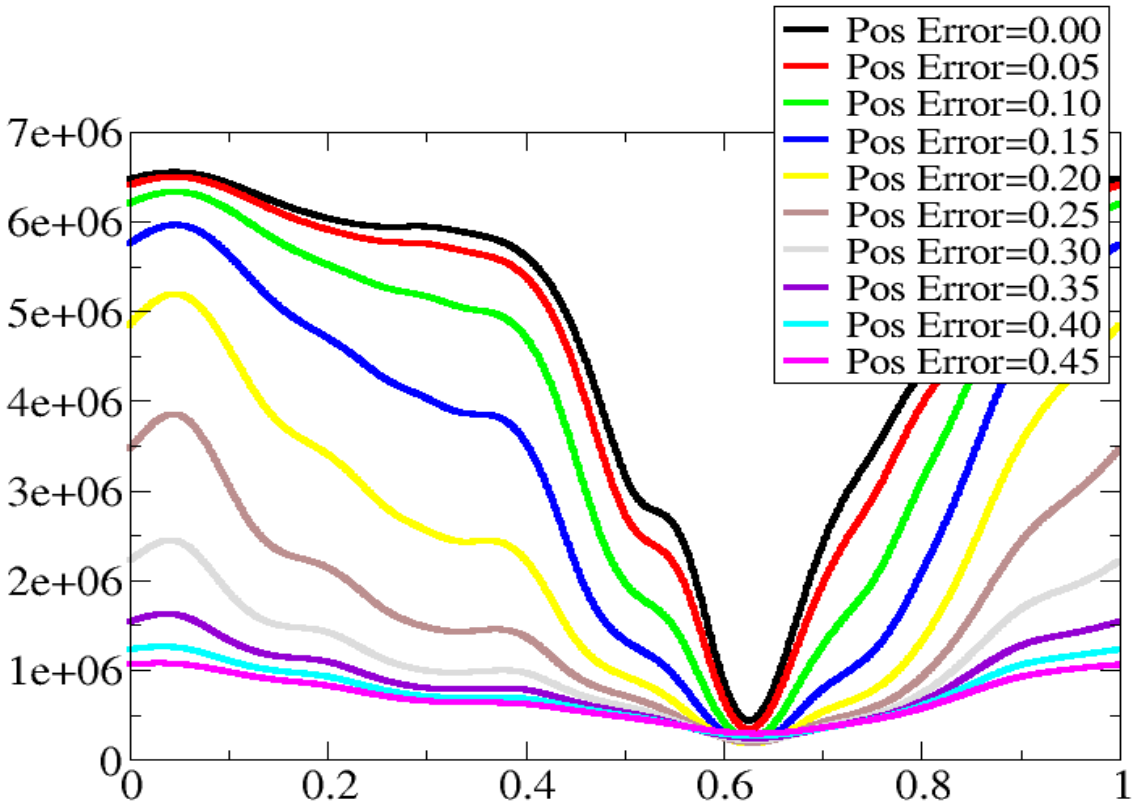
Atoms are always misplaced during a global optimisation

taking into account ***random positionnal errors*** *should yield a* ***better agreement between the incorrect model and the observed diffraction data***

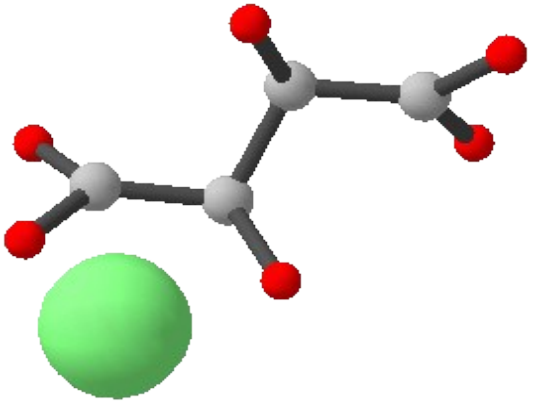
can it help its ***convergence*** *?*

Hypersurface as a function of positionnal error

$$Hypersurface : \chi^2 = f\,(parameters)$$



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*

Rules to find a structure solution: check multiple solutions

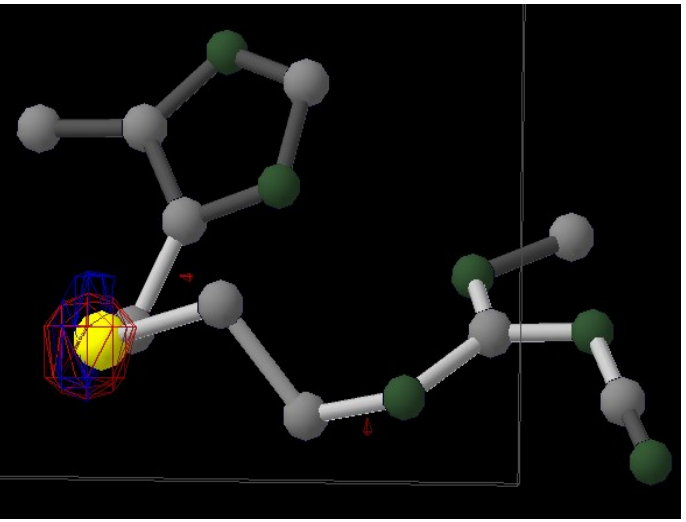
Look at multiple solutions => estimate confidence in " solution "

1) Compare the χ^2 and Rwp

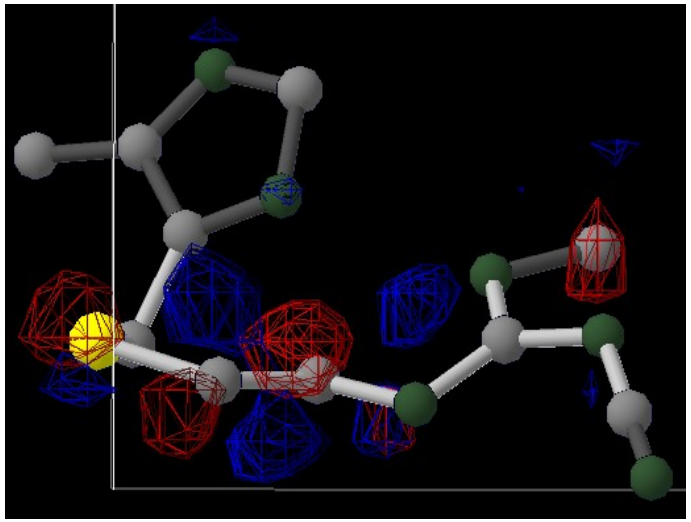
*2) Use **Fourier Difference** Maps to check differences (requires at least 1.5Å resolution data)*

Use the same contours for all solutions

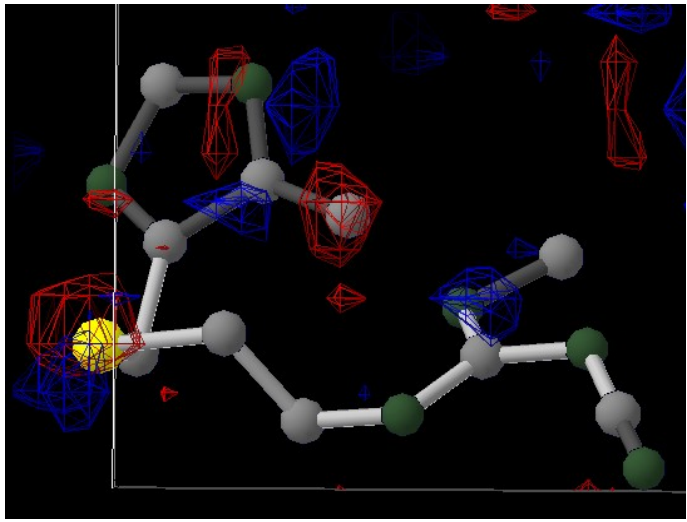
Fo-Fc, +1 and -1 contours



Correct



Wrong conformation of internal chain



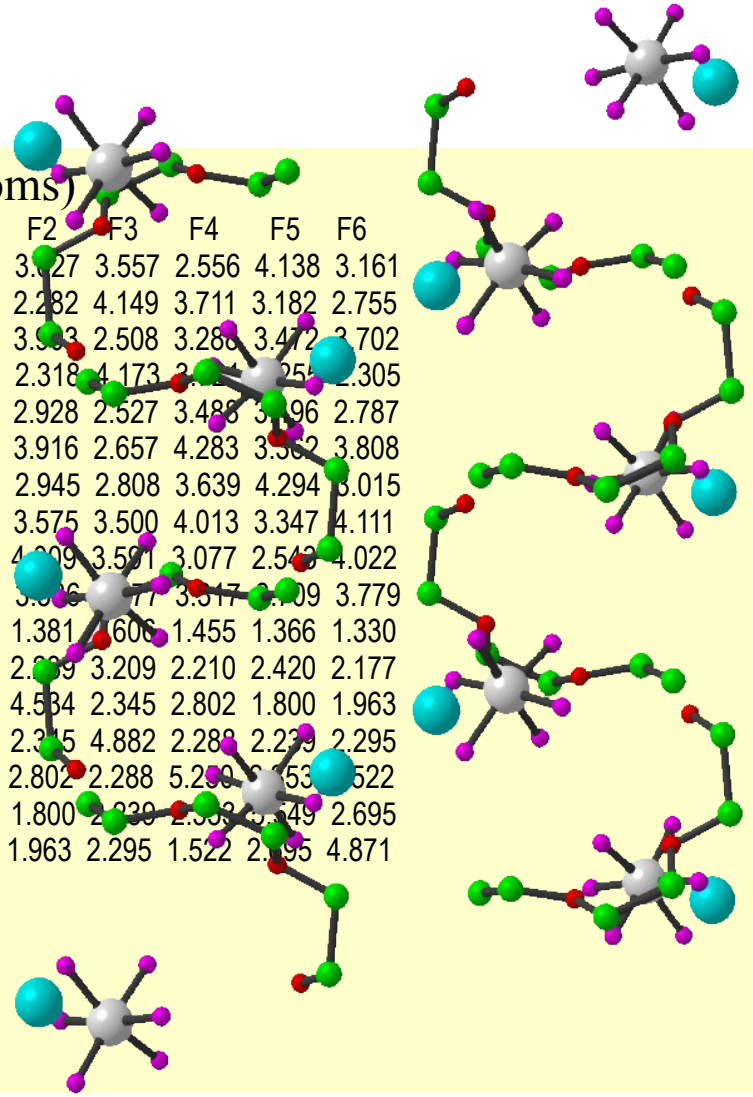
Wrong position for side CH_3 group

Interatomic distances

Check distances, overlap between atoms

Table of minimal distances between all components (atoms)

	Li	C5	C6	O3	C4	C3	O2	C2	C1	O1	As	F1	F2	F3	F4	F5	F6
Li	5.380	3.455	3.689	2.300	2.659	3.079	2.661	1.219	1.727	2.865	3.693	2.889	3.627	3.557	2.556	4.138	3.161
C5	3.455	5.160	1.571	1.436	2.322	3.469	3.530	3.789	3.074	1.829	3.175	2.389	2.282	4.149	3.711	3.182	2.755
C6	3.689	1.571	5.059	2.475	3.670	4.275	2.969	3.044	2.123	0.874	4.048	3.447	3.903	2.508	3.288	3.472	2.702
O3	2.300	1.436	2.475	4.893	1.471	2.472	2.959	2.781	3.687	3.165	2.873	3.612	2.318	4.173	3.721	3.255	2.305
C4	2.659	2.322	3.670	1.471	4.174	1.522	2.426	3.003	3.962	4.098	2.890	3.827	2.928	2.527	3.488	3.496	2.787
C3	3.079	3.469	4.275	2.472	1.522	2.890	1.349	2.428	3.481	3.958	4.383	3.754	3.916	2.657	4.283	3.362	3.808
O2	2.661	3.530	2.969	2.959	2.426	1.349	3.693	1.457	2.435	2.707	4.101	3.683	2.945	2.808	3.639	4.294	3.015
C2	1.219	3.789	3.044	2.781	3.003	2.428	1.457	4.955	1.067	2.379	4.260	3.606	3.575	3.500	4.013	3.347	4.111
C1	1.727	3.074	2.123	3.687	3.962	3.481	2.435	1.067	4.115	1.391	3.271	3.844	4.009	3.591	3.077	2.543	4.022
O1	2.865	1.829	0.874	3.165	4.098	3.958	2.707	2.379	1.391	4.398	4.048	3.344	3.336	3.077	3.317	3.709	3.779
As	3.693	3.175	4.048	2.873	2.890	4.383	4.101	4.260	3.271	4.048	5.570	1.607	1.381	3.606	1.455	1.366	1.330
F1	2.889	2.389	3.447	3.612	3.827	3.754	3.683	3.606	3.844	3.344	1.607	3.077	2.239	3.209	2.210	2.420	2.177
F2	3.627	2.282	3.903	2.318	2.928	3.916	2.945	3.575	4.009	3.336	1.381	2.239	4.534	2.345	2.802	1.800	1.963
F3	3.557	4.149	2.508	4.173	2.527	2.657	2.808	3.500	3.591	3.077	1.606	3.209	2.345	4.882	2.288	2.239	2.295
F4	2.556	3.711	3.288	3.721	3.488	4.283	3.639	4.013	3.077	3.317	1.455	2.210	2.802	2.288	5.290	3.753	3.522
F5	4.138	3.182	3.472	3.255	3.496	3.362	4.294	3.347	2.543	3.709	1.366	2.420	1.800	2.239	3.538	3.549	2.695
F6	3.161	2.755	3.702	2.305	2.787	3.808	3.015	4.111	4.022	3.779	1.330	2.177	1.963	2.295	1.522	2.695	4.871



Polymer Electrolyte β-PEO$_6$:LiAsF$_6$

E.Staunton, Yu. Andreev, P. Bruce *JACS* ***127*** (2005), 12176

Top problems for ab initio structure determination:

(assuming data and unit cell correctly

#1 Model is wrong:

- wrong formula
- incorrect restraints
- wrong spacegroup
- missing solvant
...

Solution:
*1- check **real** composition (mass spectroscopy, EDX microscope)*
2- Add new atoms/polyhedra if they are missing

(The first structure solved with Fox, $CsOH.H_2O$
...was supposed to be a Cs hydride !)

#2 Preferred Orientation

Solution:
1- Collect new data
...
9- Collect new data
*10- Search for preferred orientation parameters **ab initio***

Phys. Rev. B 76 (2007), 092104

#3 Not enough data

Solution:
Add more restraints: antibump, bond length
Use other methods to gather restraints (NMR)

#4 Need a faster computer ?

Check with litterature if more trials are <u>really</u> required or email the author for advice !

Rules to find a structure solution: be a flexible User

***Restraints** must be used to reduce parameter space*

*... but **too many restraints** can **slow** or **prevent** a structure solution*

e.g.: a combination of strong antibump and angular restraints can make very difficult to go from one local minimum to the global one.

Imagine the “molecule” to be solved is: a man & a chair.
The “solution” is:
the man, sitting on the chair, in the Prado Museum (Madrid)

The random starting location is: Grenoble railway station.

To “ speed up ” the solution, you impose as much restraints as you can, i.e. “ the man must be sitted on the chair at all times ”

=> Rigid groups should be used scarcely... and do not generally speed up the convergence
*=> **if the algorithm “ distorts ” your molecule during the optimization, it's for your own good (honest !)***
=> the correct conformation comes from the data, not the number of restraints
=> NB: different rules apply if data is of bad quality

Rules to find a structure solution: no high-resolution data

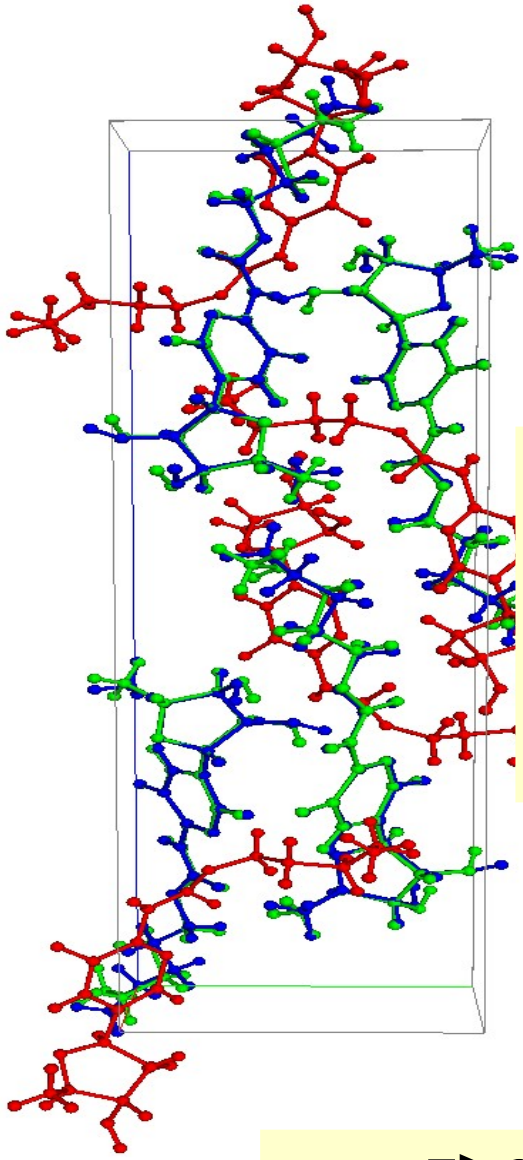
“ Solving the structure ” means finding all the atomic positions with an error of ~ 0.1Å.
=> high resolution data (1Å and higher) is not needed...
increasing the resolution by 25% doubles the computing time !

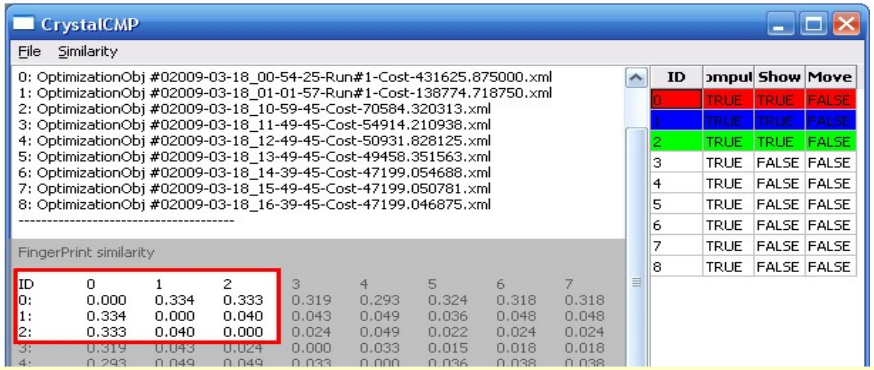
Giving high-resolution data is like giving your address to a friend
*....by insisting on the **exact pattern of colours from your garden's flowers***
*... instead of **just giving the address & colour** of the house*

Most of the time, a 2.5 Å resolution is enough.
...sometimes 1.5Å.

Of course you still need the high-resolution data for the least squares refinement !

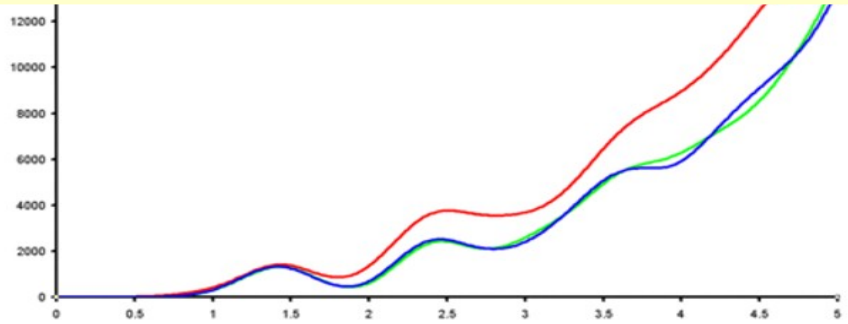
Comparing crystal structures





Sometimes the score (Chi^2, Rwp) is not enough to decide which solution is correct

=> to browse among solutions a tool has been developed (Jan Rohlíček, Michal Hušák, ICT Prague):
http://crystalcmp.sf.net

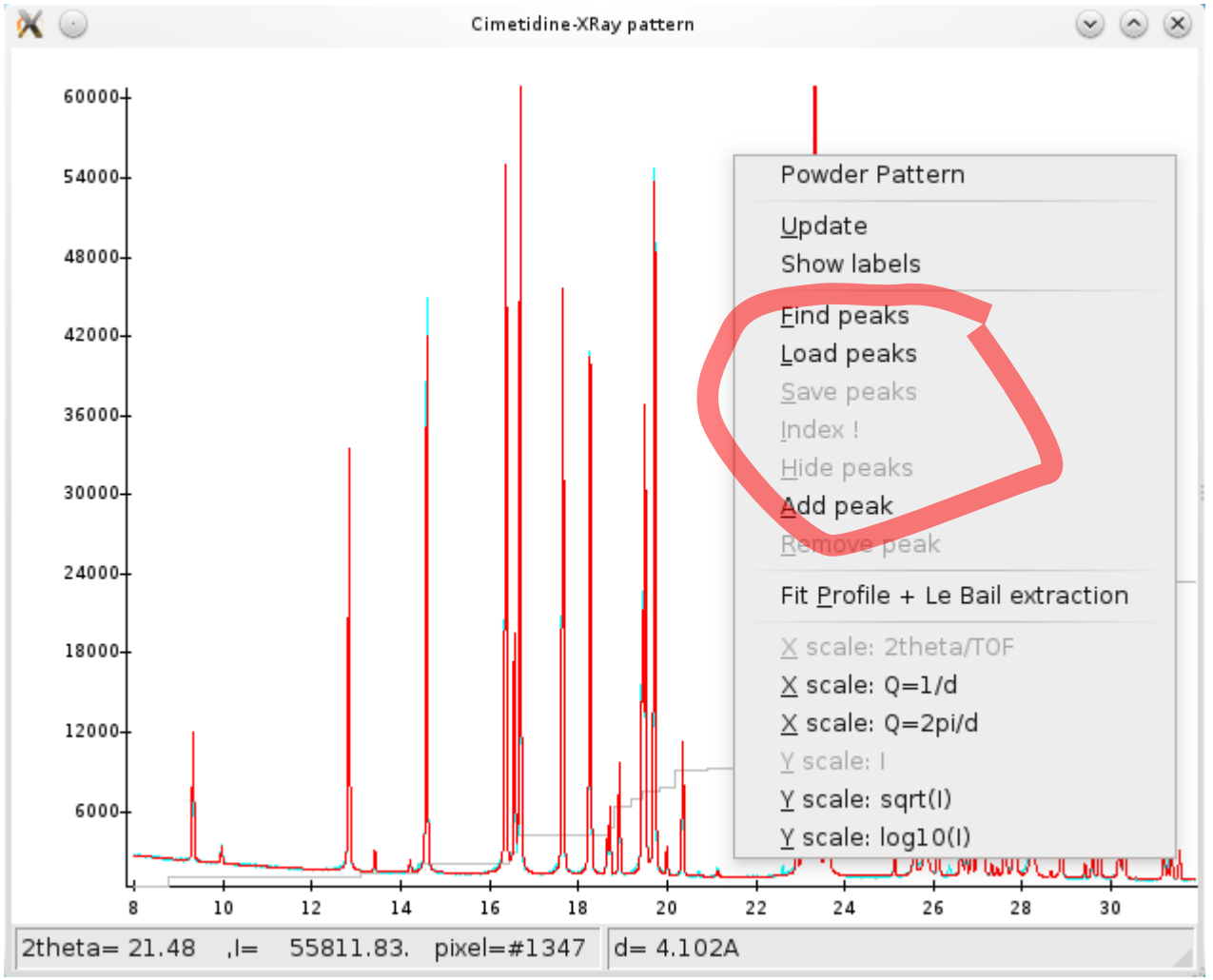


=> allows to superpose several structures
=> also compares a 'fingerprint' of all molecules

Fox: recent features

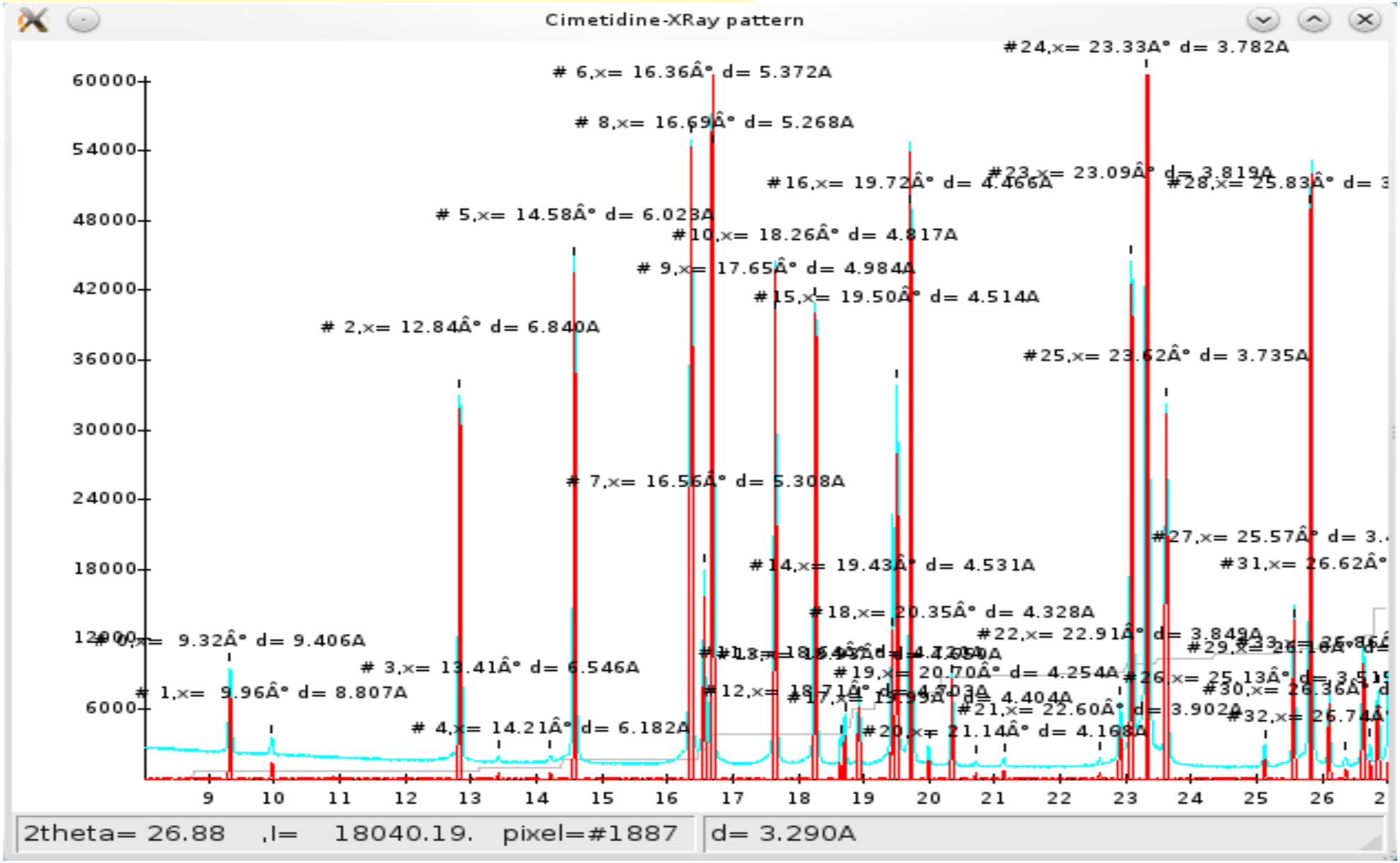
Auto-indexing

- *Search for peaks*

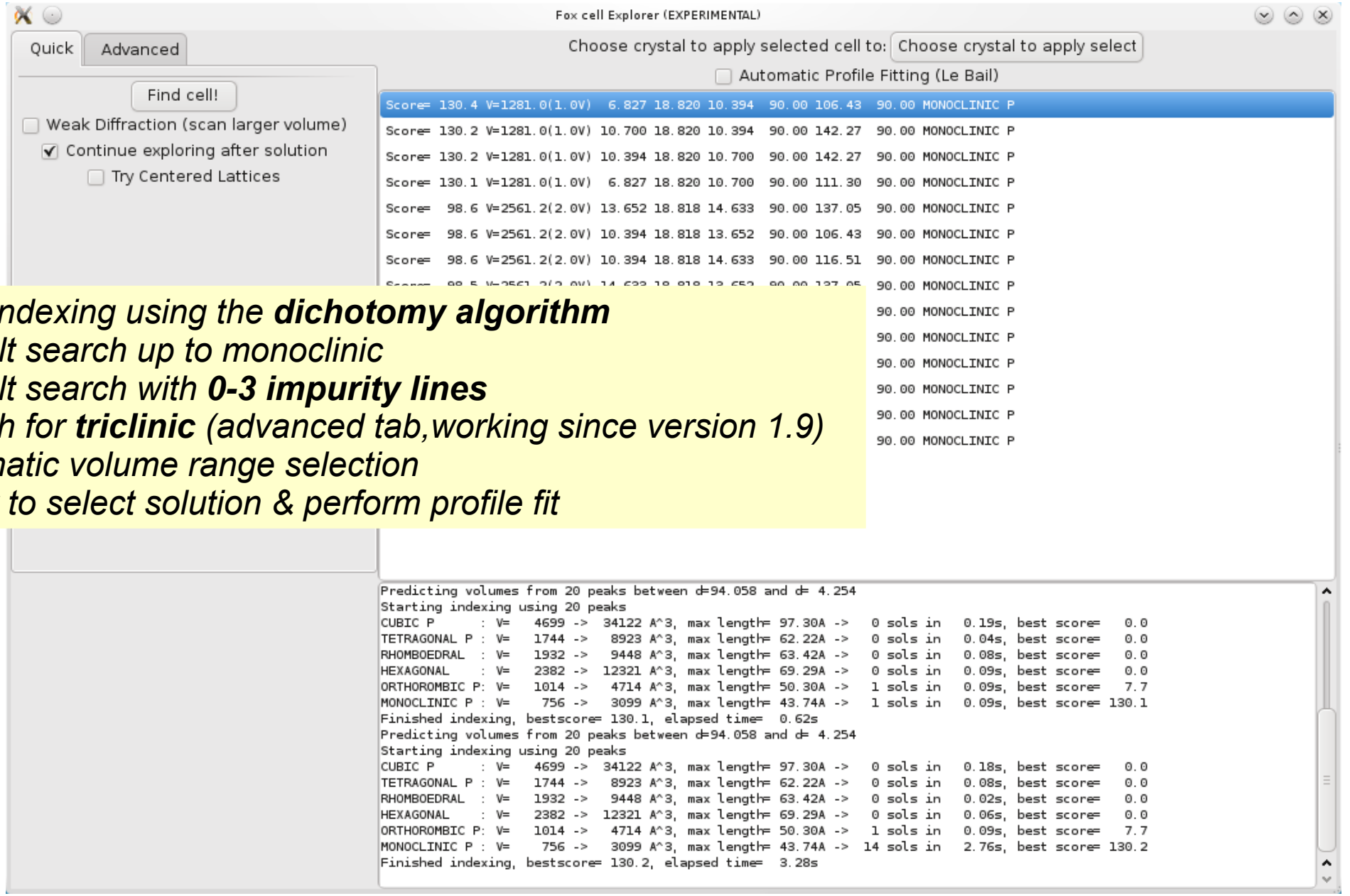


Auto-indexing

- *Search for peaks*



Auto-indexing



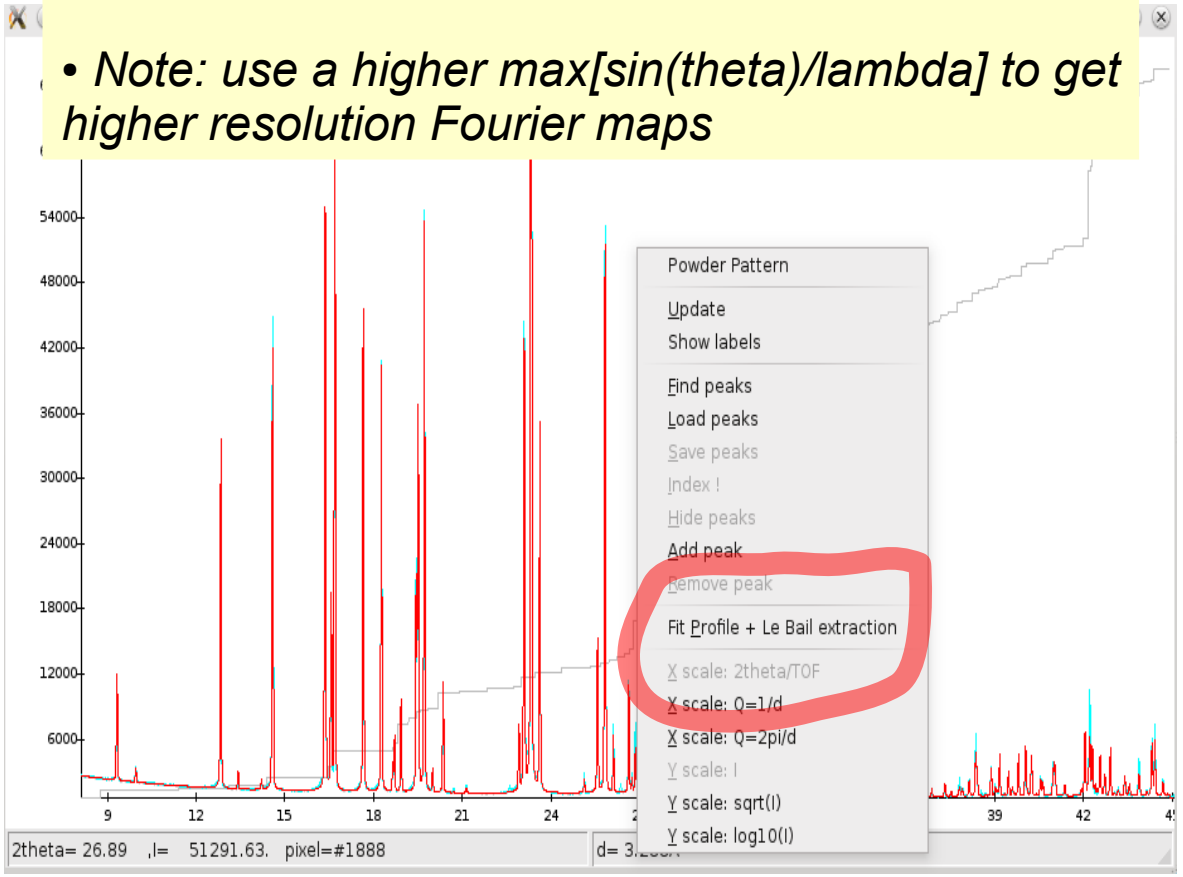
- *Auto-indexing using the **dichotomy algorithm***
- *Default search up to monoclinic*
- *Default search with **0-3 impurity lines***
- *Search for **triclinic** (advanced tab,working since version 1.9)*
- *Automatic volume range selection*
- *Ability to select solution & perform profile fit*

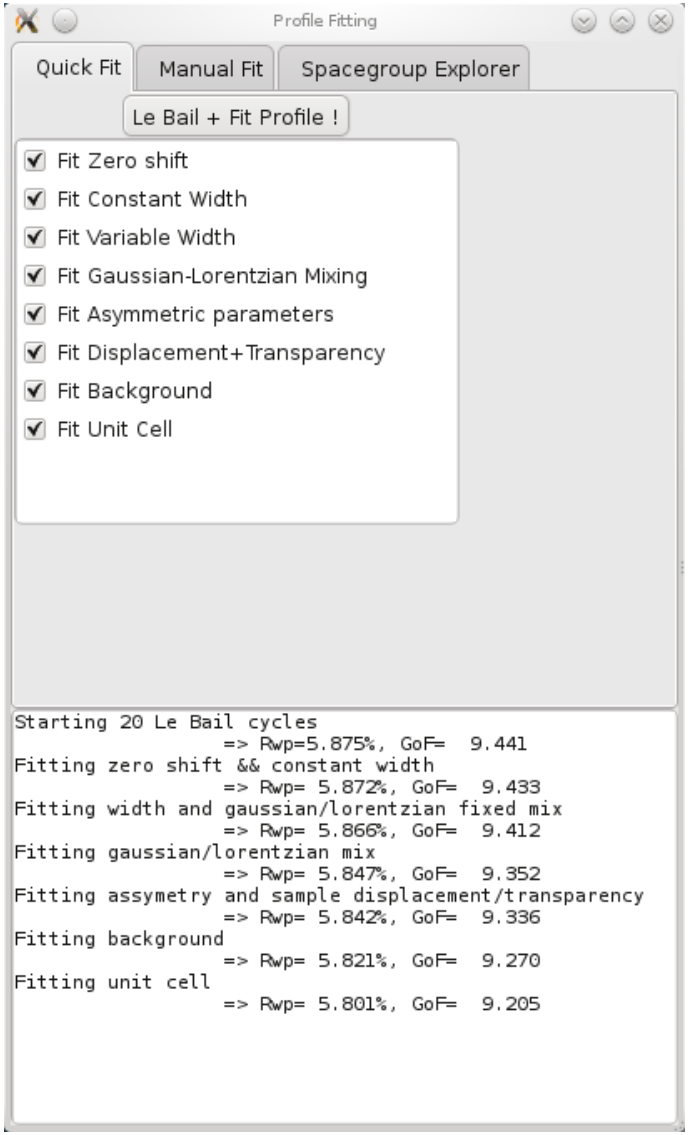
Profile fitting & Le Bail extraction

- *Fox is not too sensitive to exact profiles but it is still better to use fitted ones (for large overlaps)*

- *By default Fox refines parameters with increasing complexity (width -> symmetric profile -> asymmetric -> background -> cell...)*

- *Note: use a higher max[sin(theta)/lambda] to get higher resolution Fourier maps*

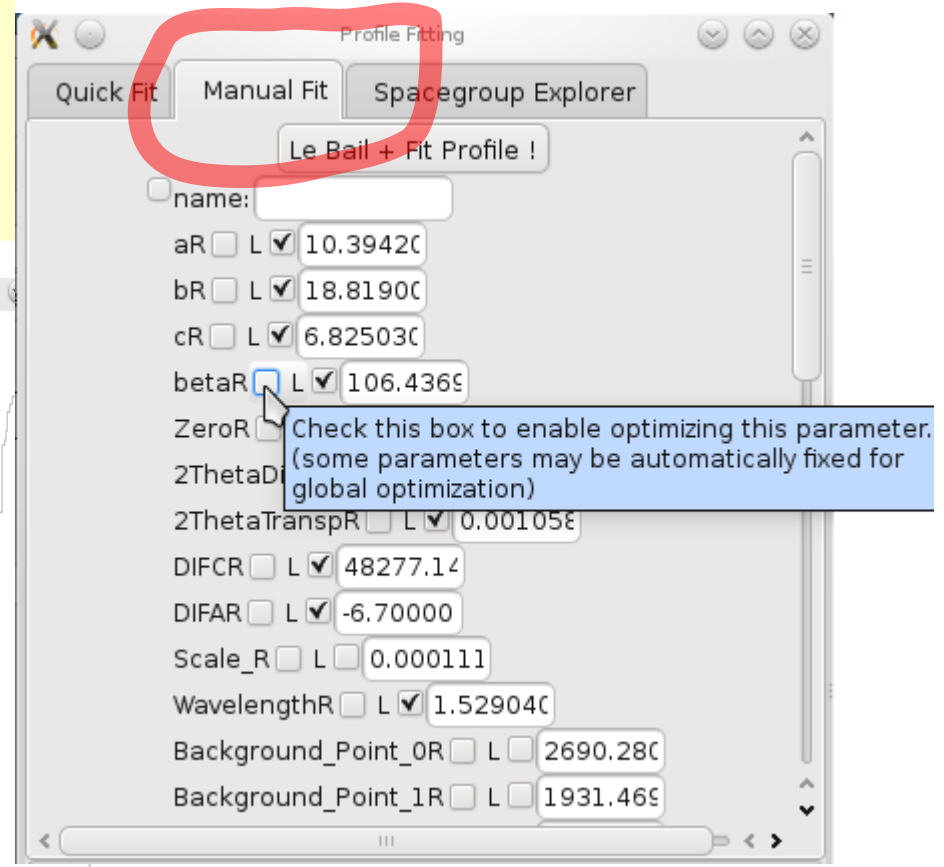


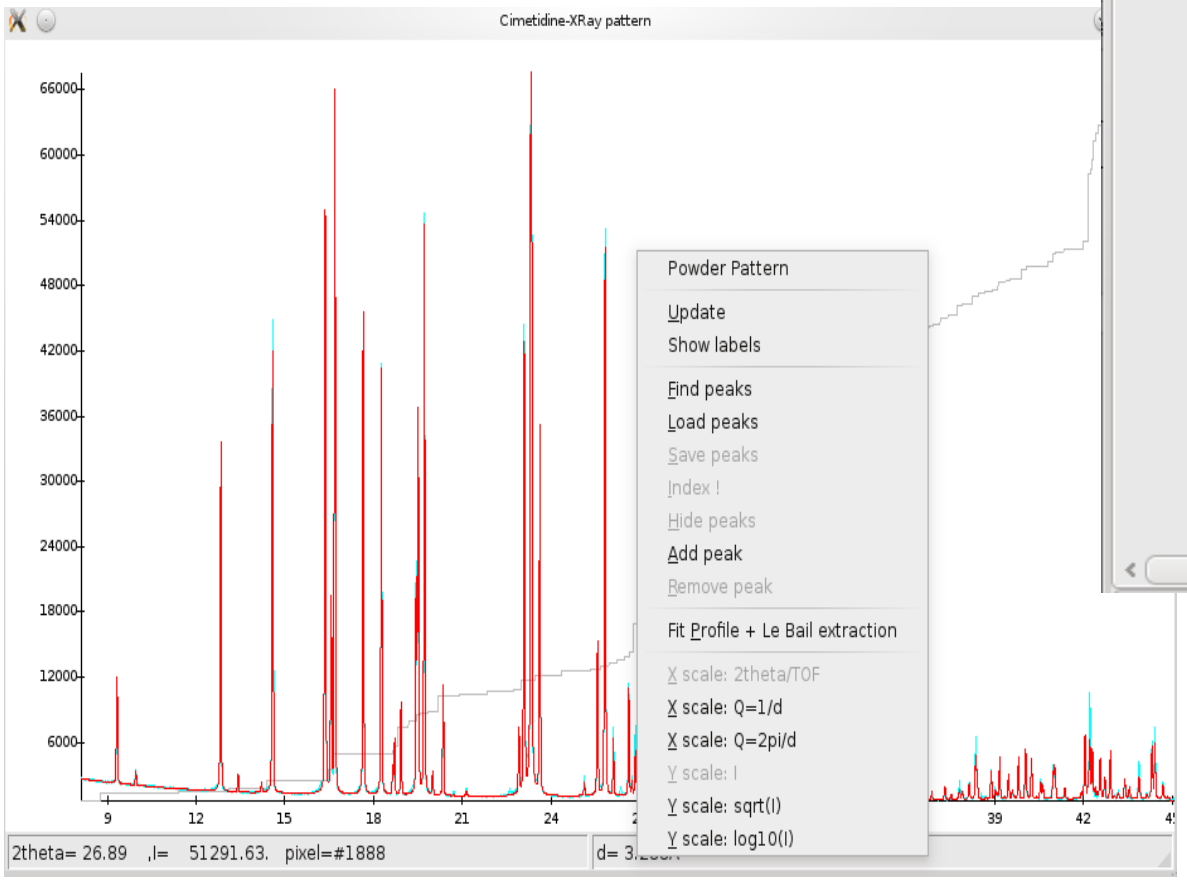


Profile fitting & Le Bail extraction

- *It is also possible to select individual parameters for a manual fit*

- *NB: enable tooltips and read them !*

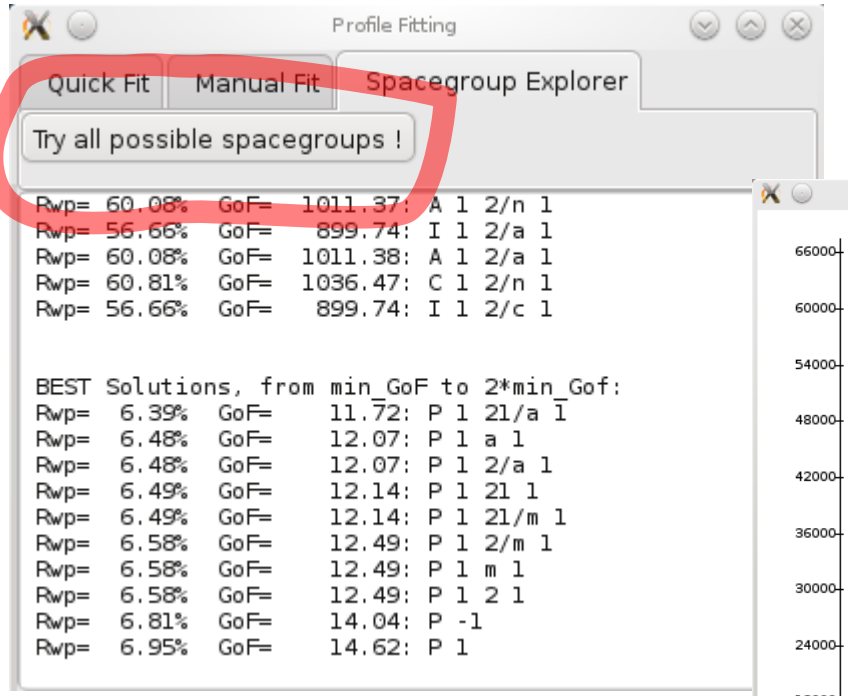


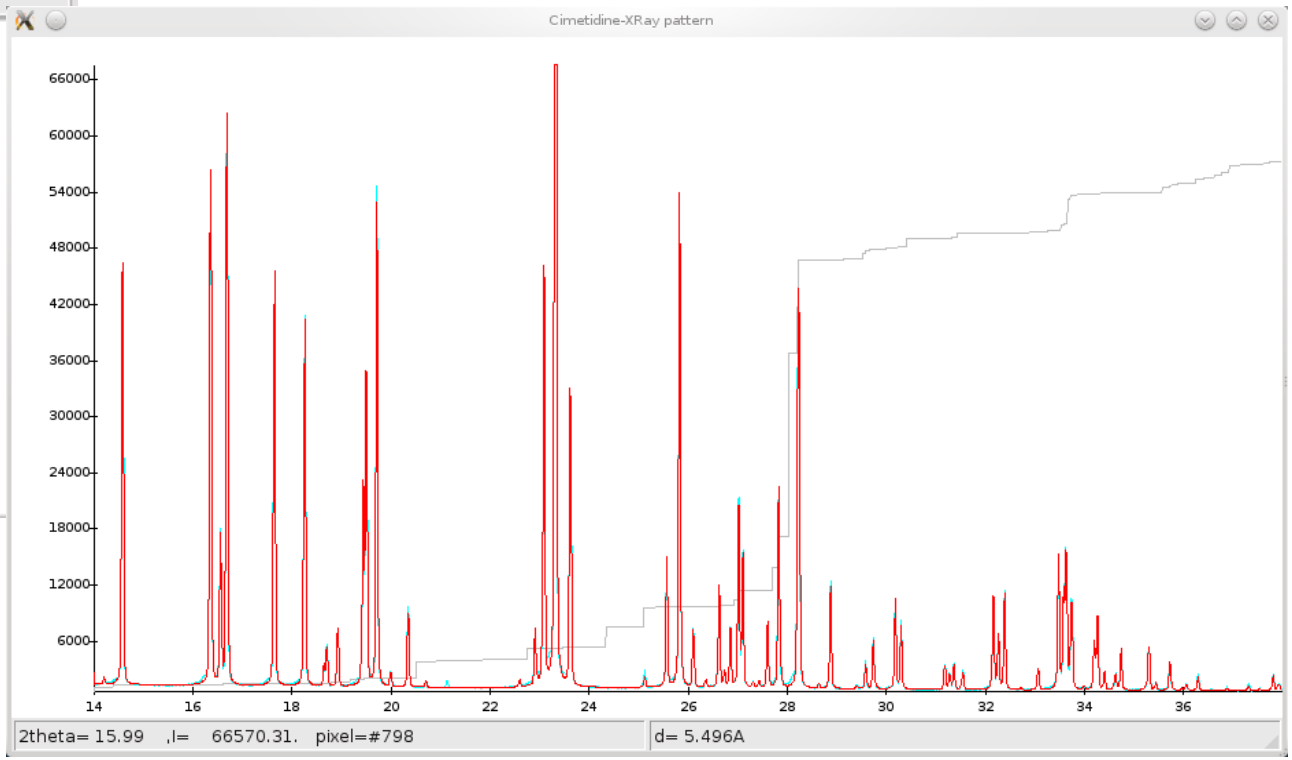


Profile fitting: spacegroup explorer

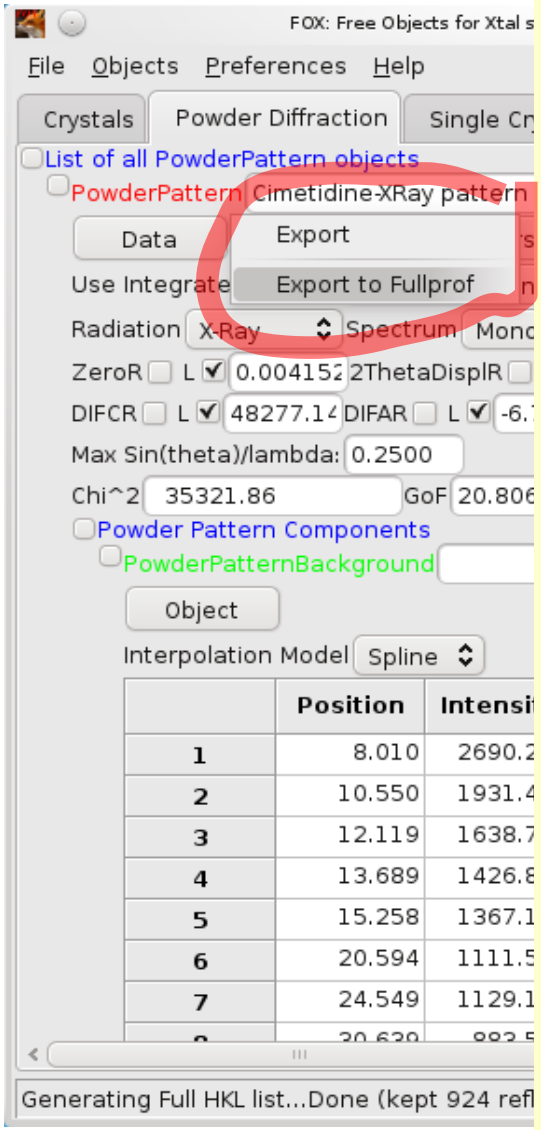
- *Performs a profile fitting for all spacegroup settings allowed by the unit cell (37 for monoclinic cells, 478 for cubic cells… it can take a while).*

- *Spacegroups are listed by increasing GoF up to to 2*min(GoF)*



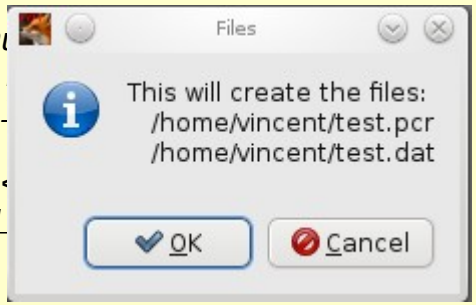


Fullprof export

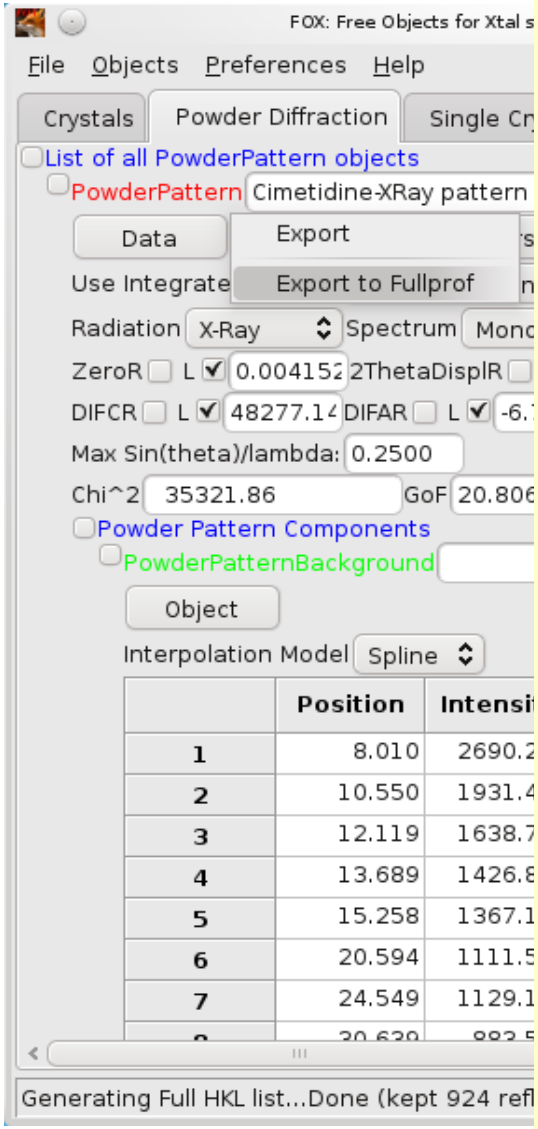



```
• 1 !Number of refined parameters
•! Zero Code Sycos Code Sysin Code Lambda Code More -> Patt #1
• 0.00415153 0.0 0.00119514 0.0 0.00105758 0.0 0.000  0.0  0
•!-------------------------------------------------------------------------------
--
•!  Data for PHASE number:   0  ==> Current R_Bragg for Pattern#  1:     0.00

•!-------------------------------------------------------------------------------
--
•Cimetidine
•!Nat Dis Ang Jbt Isy Str Furth  ATZ Nvk More
•17  17  20   0   0   0    0    1.0  0   1
•!Jvi Jdi Hel Sol Mom Ter  N_Domains
•  0   3   0   0   0   0      0
•!Contributions (0/1) of this phase to the  patterns
• 1
•!Irf Npr Jtyp  Nsp_Ref Ph_Shift for Pattern#0
•  0   0   0      0      0
•! Pr1    Pr2    Pr3   Brind.   Rmu                                  Pattern#0
•  1.0    1.0    1.0    1.0      0
•! Max_dst(dist) (angles)  Bond-Va
•    2.7000      1.5000        0
•P 1 21/a 1                    '
•!Atom Typ X Y Z Biso Occ In Fin N_
•C1 C 0.566876 0.342198 0.343369 3
•       0 0 0  0    0
•N2 N 0.448664 0.371933 0.388352 3 1 0  0   0   0
•       0 0 0  0    0
•C3 C 0.454253 0.405903 0.550729 3 1 0  0   0   0
•       0 0 0  0    0
•N4 N 0.586007 0.407123 0.696493 3 1 0  0   0   0
•       0 0 0  0    0
•C5 C 0.624978 0.44932 0.859614 3 1 0  0   0   0
•       0 0 0  0    0
•N6 N 0.663486 0.486503 1.01791 3 1 0  0   0   0
•       0 0 0  0    0
•N7 N 0.367983 0.448348 0.627608 3 1 0  0   0   0
•       0 0 0  0    0
```



Fullprof export

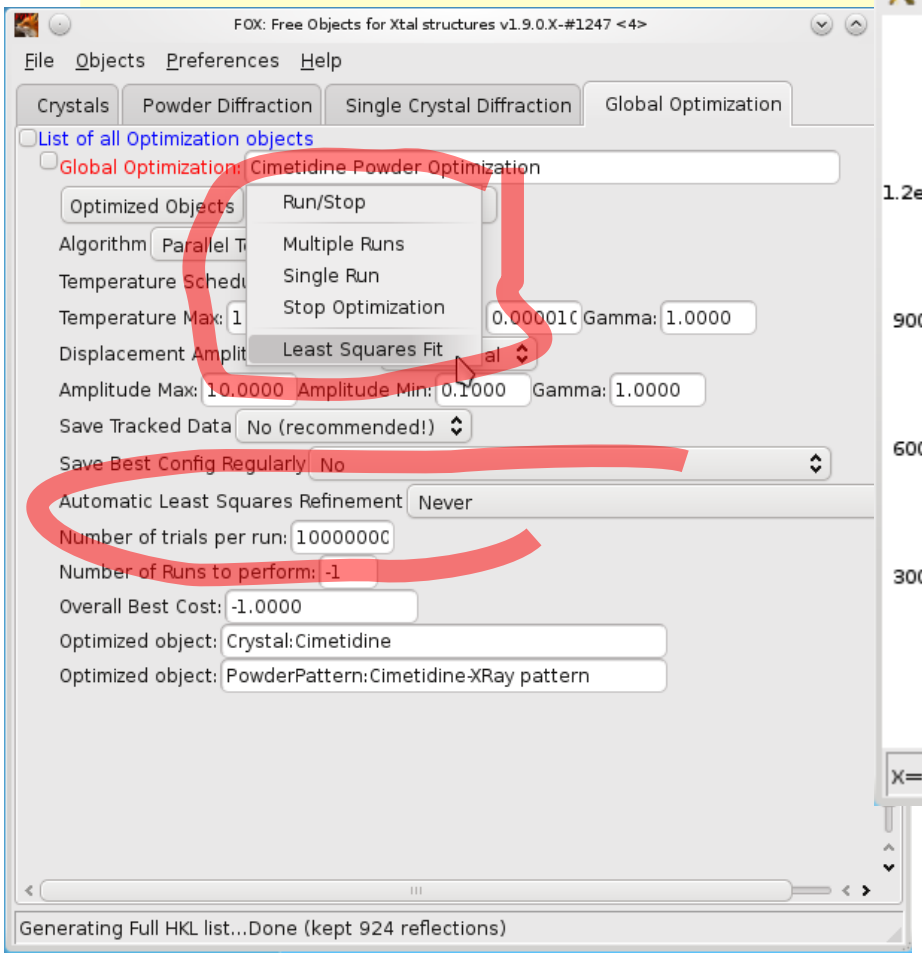


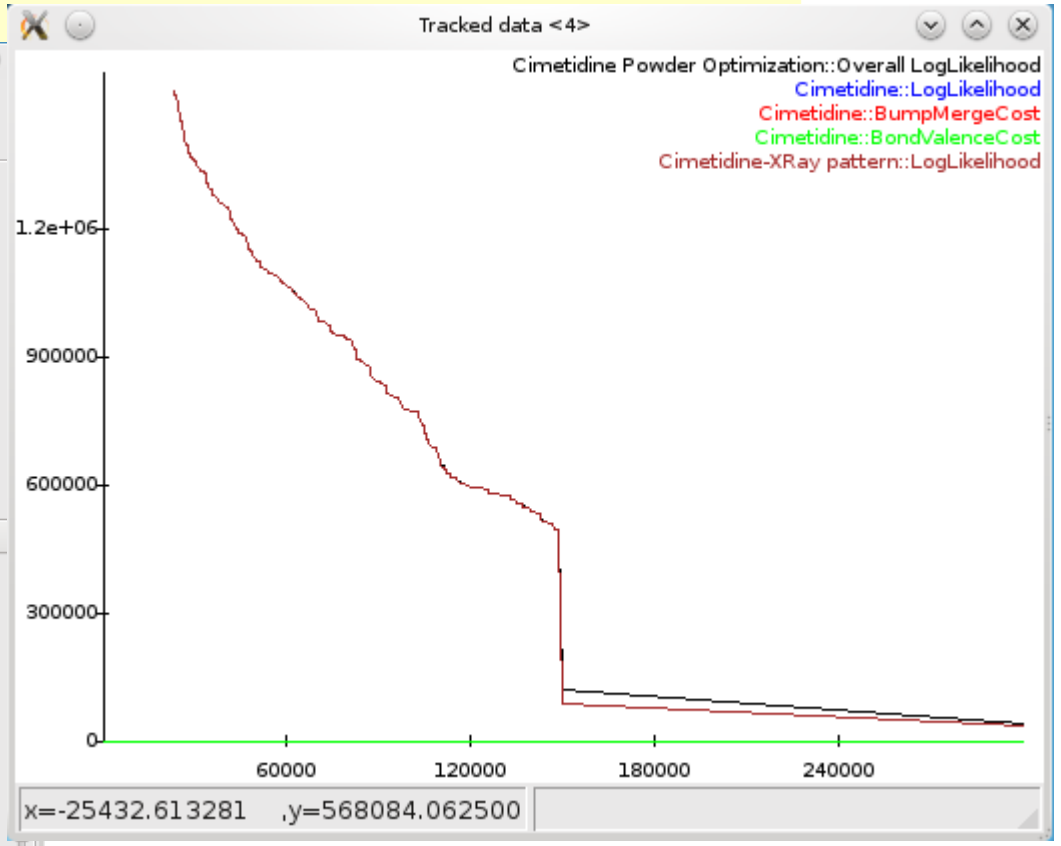

```
!     U      V      W      X      Y     GauSiz     LorSiz Size-Model
0.00341898 8.85951e-05 0.00156559   0.0   0.0      0.0        0.0
    0.0   0.0   0.0   0.0   0.0      0.0        0.0
!     a           b           c         alpha       beta        gamma       #Cel
Info
10.3942 18.819 6.82503 90 106.437 90
    0.0         0.0       0.0         0.0         0.0         0.0
! Pref1 Pref2 alpha0 beta0 alpha1 beta1 ?
   0.0   0.0    0.0   0.0    0.0   0.0
   0.0   0.0    0.0   0.0    0.0   0.0
!Soft distance constraints
C1 N2 1 0 0 0 1.47 0.01
N2 C3 1 0 0 0 1.4 0.01
C3 N4 1 0 0 0 1.4 0.01
N4 C5 1 0 0 0 1.37 0.01
C5 N6 1 0 0 0 1.13 0.01
C3 N7 1 0 0 0 1.4 0.01
N7 C8 1 0 0 0 1.47 0.01
C8 C9 1 0 0 0 1.54 0.01
C9 S10 1 0 0 0 1.82 0.01
S10 C11 1 0 0 0 1.82 0.01
C11 C12 1 0 0 0 1.5 0.01
C12 N13 1 0 0 0 1.37 0.01
N13 C14 1 0 0 0 1.37 0.01
C14 N15 1 0 0 0 1.37 0.01
N15 C16 1 0 0 0 1.37 0.01
C16 C17 1 0 0 0 1.5 0.01
C16 C12 1 0 0 0 1.45 0.01
!Soft angle constraints
C1 N2 C3 1 1  0 0 0  0 0 0 108.862 0.572958
N2 C3 N4 1 1  0 0 0  0 0 0 119.748 0.572958
C3 N4 C5 1 1  0 0 0  0 0 0 119.748 0.572958
N4 C5 N6 1 1  0 0 0  0 0 0 180 0.572958
N2 C3 N7 1 1  0 0 0  0 0 0 119.748 0.572958
N4 C3 N7 1 1  0 0 0  0 0 0 119.748 0.572958
C3 N7 C8 1 1  0 0 0  0 0 0 108.862 0.572958
N7 C8 C9 1 1  0 0 0  0 0 0 108.862 0.572958
C8 C9 S10 1 1  0 0 0  0 0 0 108.862 0.572958
```

Least squares

Least squares refinement can be performed:

- *For profile fitting*
- *After optimization (only the structure is refined, no parameter choice)*
- ***Automatically during global optimization***

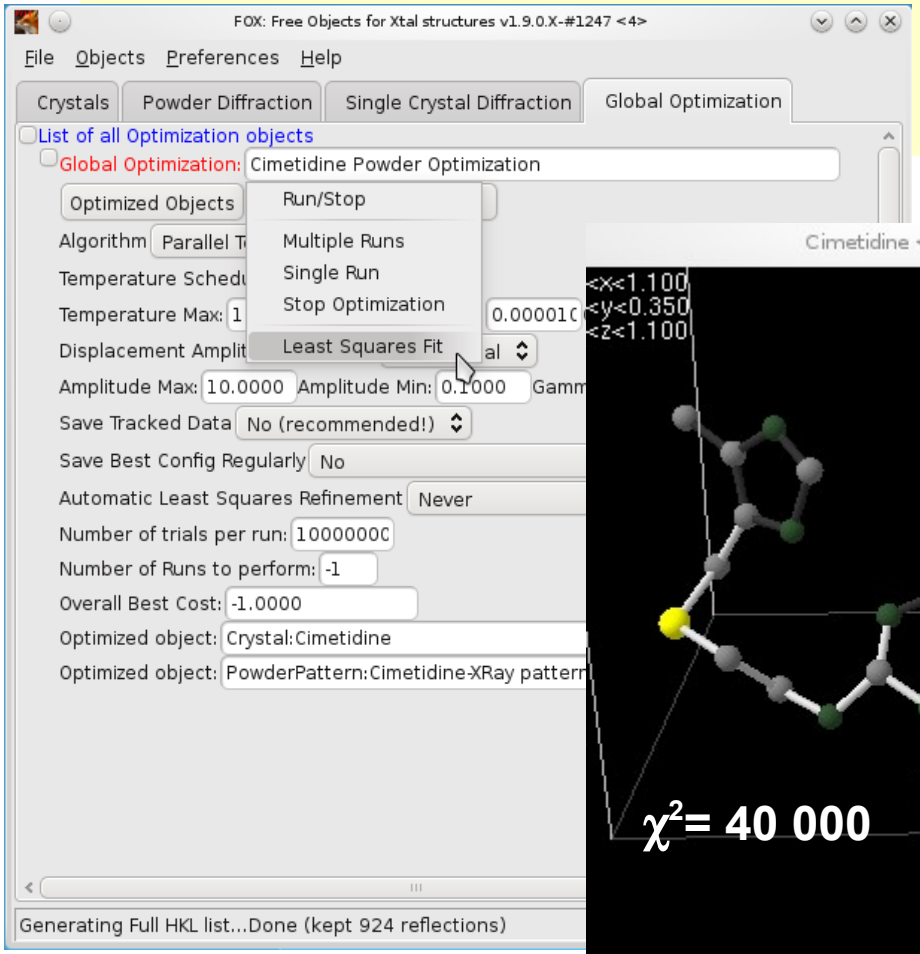


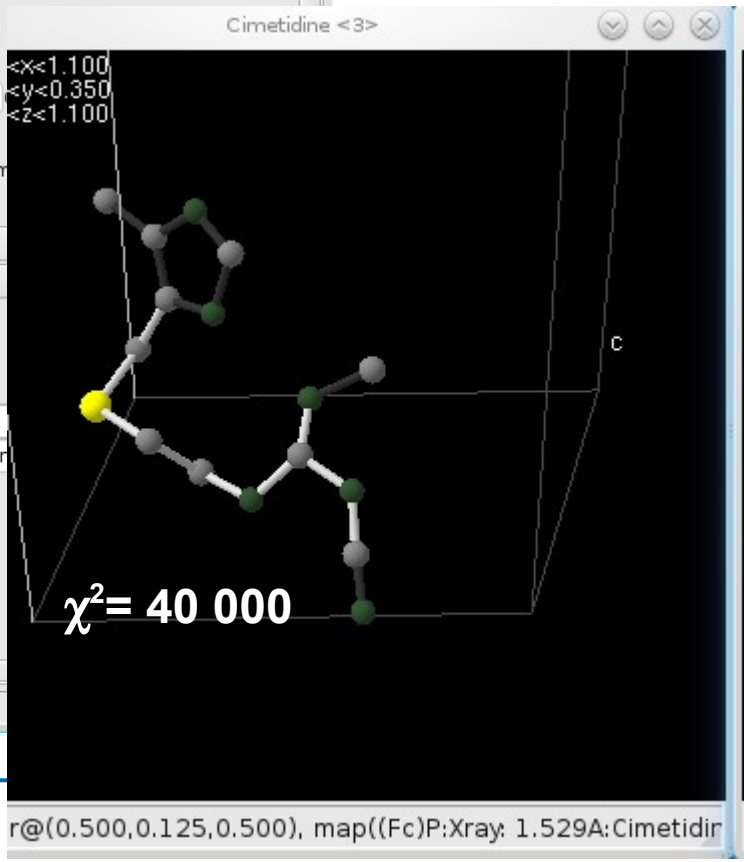


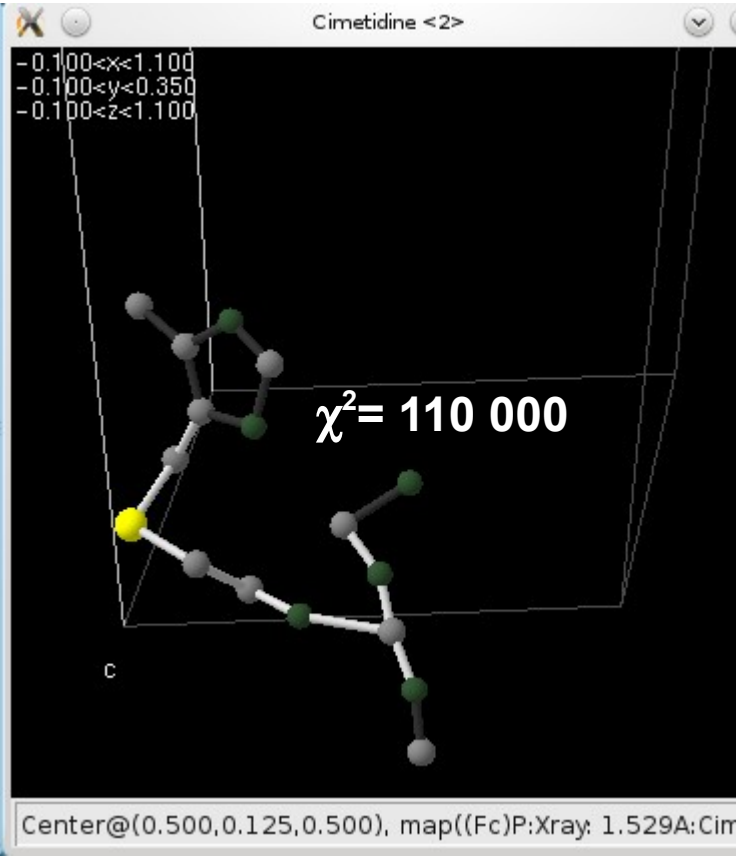
Least squares

Least squares refinement can be performed:

- *For profile fitting*
- *After optimization (only the structure is refined, no parameter choice)*
- ***Automatically during optimization***



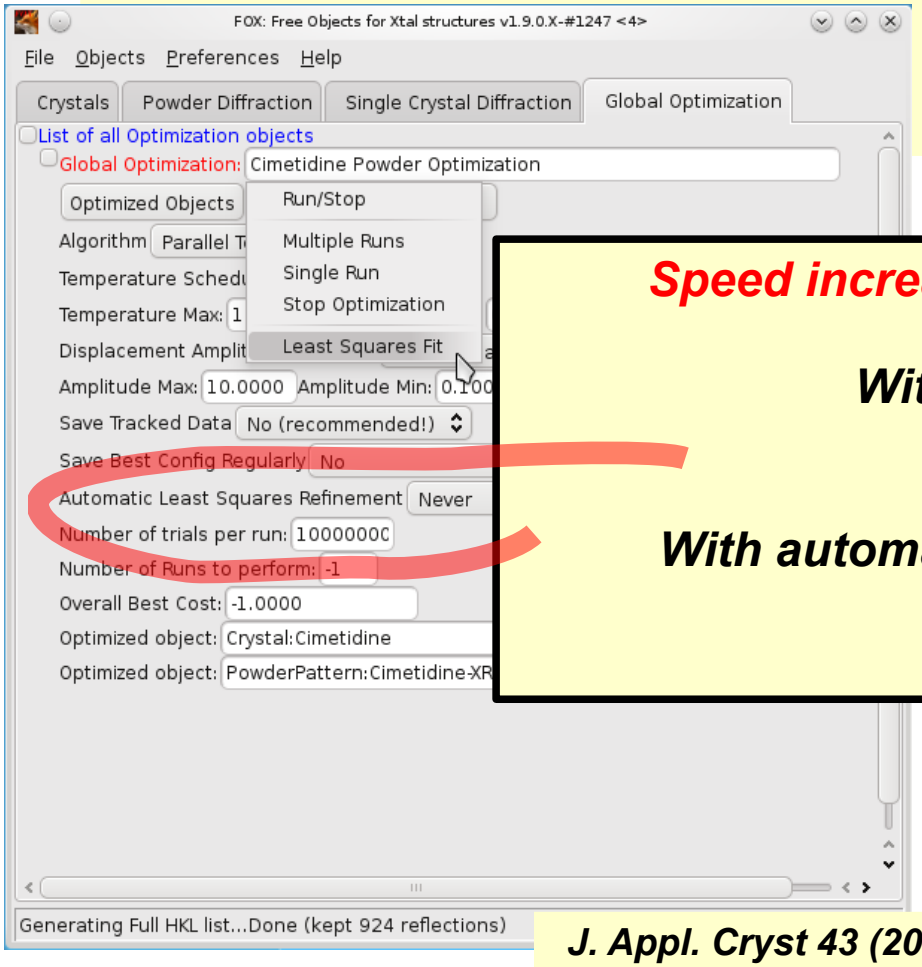




Least squares

Least squares refinement can be performed:

- *For profile fitting*
- *After optimization (only the structure is refined, no parameter choice)*
- ***Automatically during optimization***



Speed increase for structure solution (Cimetidine) :

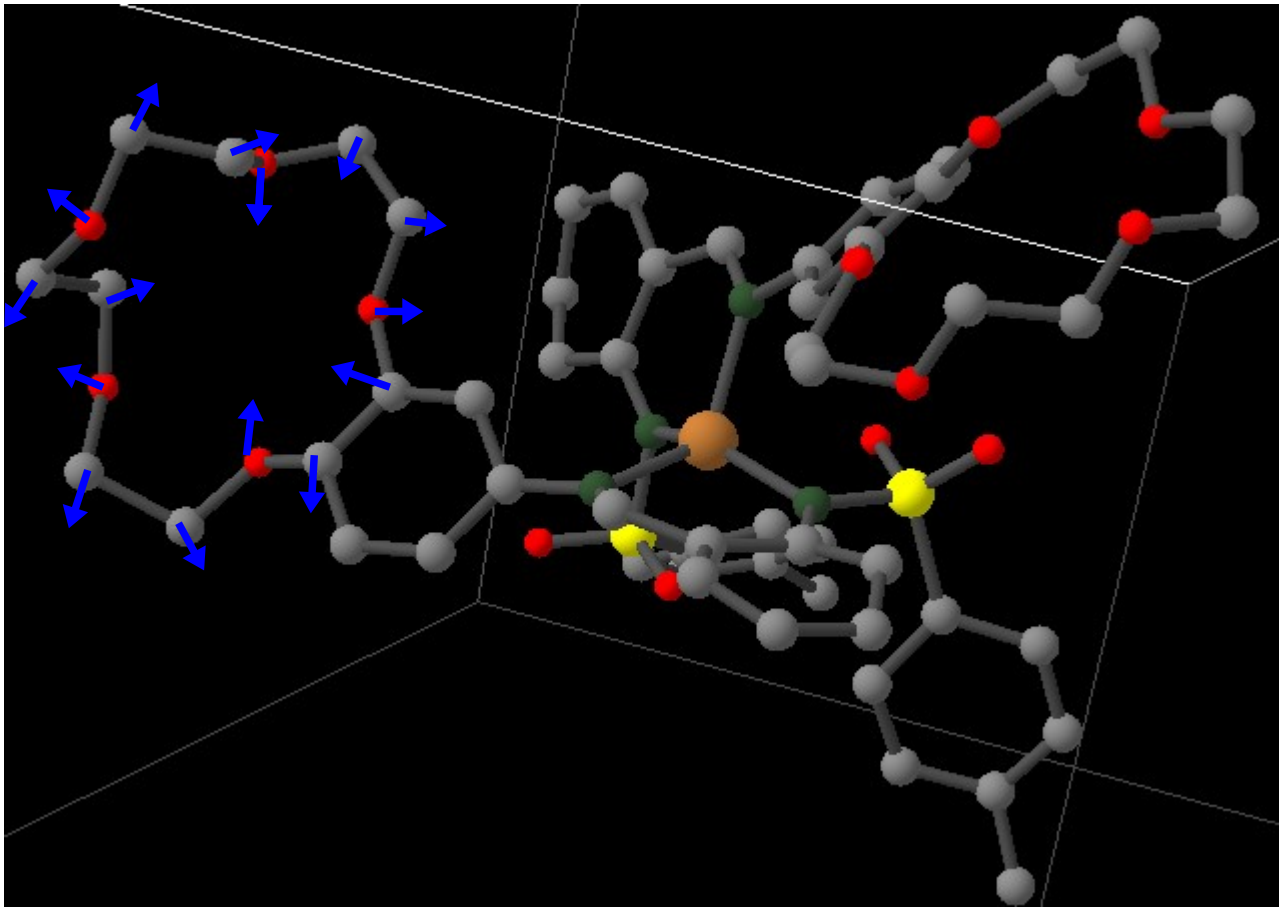
Without automatic least-squares:
1.6×10^6 trials

With automatic least-squares every 150 000 trials:
6×10^5 trials (~2 minutes)

J. Appl. Cryst 43 (2010),401 : suggests global optimization can be performed by doing only downhill minimization from random starting points...

Model Building: Molecular Dynamics for Flexible Cycles

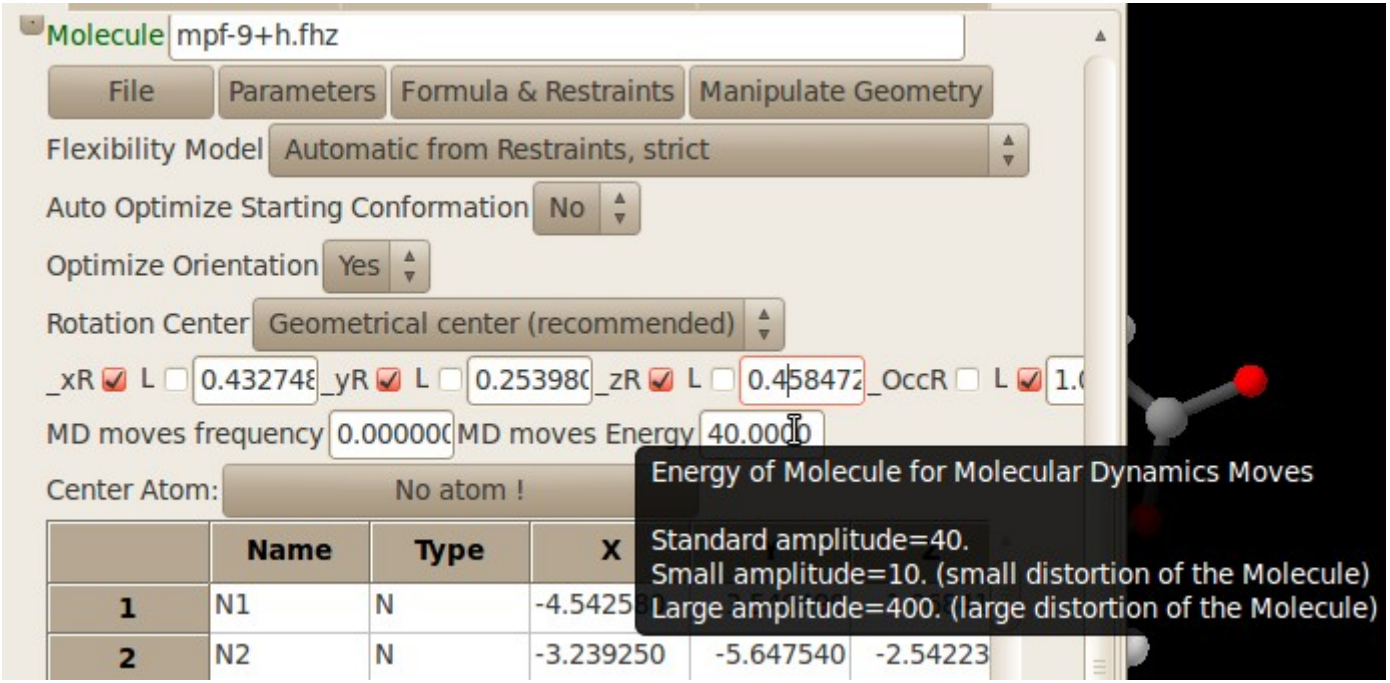
Using Molecular Dynamics



Atoms in "restrained" groups are moved using molecular dynamics principles :

- *Each atom is given a random vector speed*
- *The overall Energy is $E_{kinetic} + E_{restraints}$*
- *Atoms are moved according to standard mechanics (force=gradient of $E_{restraints}$)*

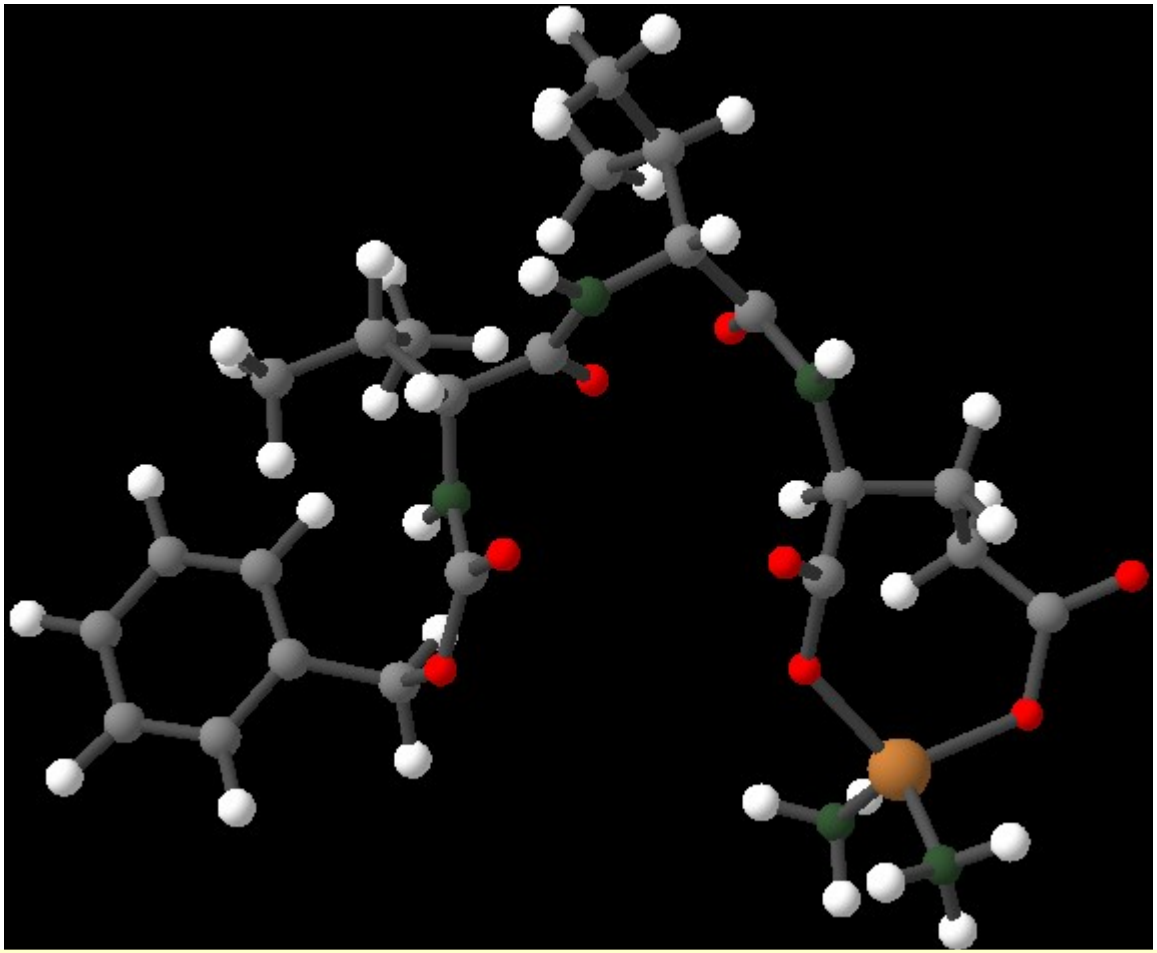
Using Molecular Dynamics



*MD moves are **computationally expensive***
=> they are only tried once in a while
*=> the **frequency** can be chosen (by default: 0=never)*
*=> the **relative energy** of the molecule can be chosen to avoid too much distortion*

… But remember that SOME DISTORTION IS NECESSARY to reach the 'true' conformation of the Molecule, starting from an incorrect one...

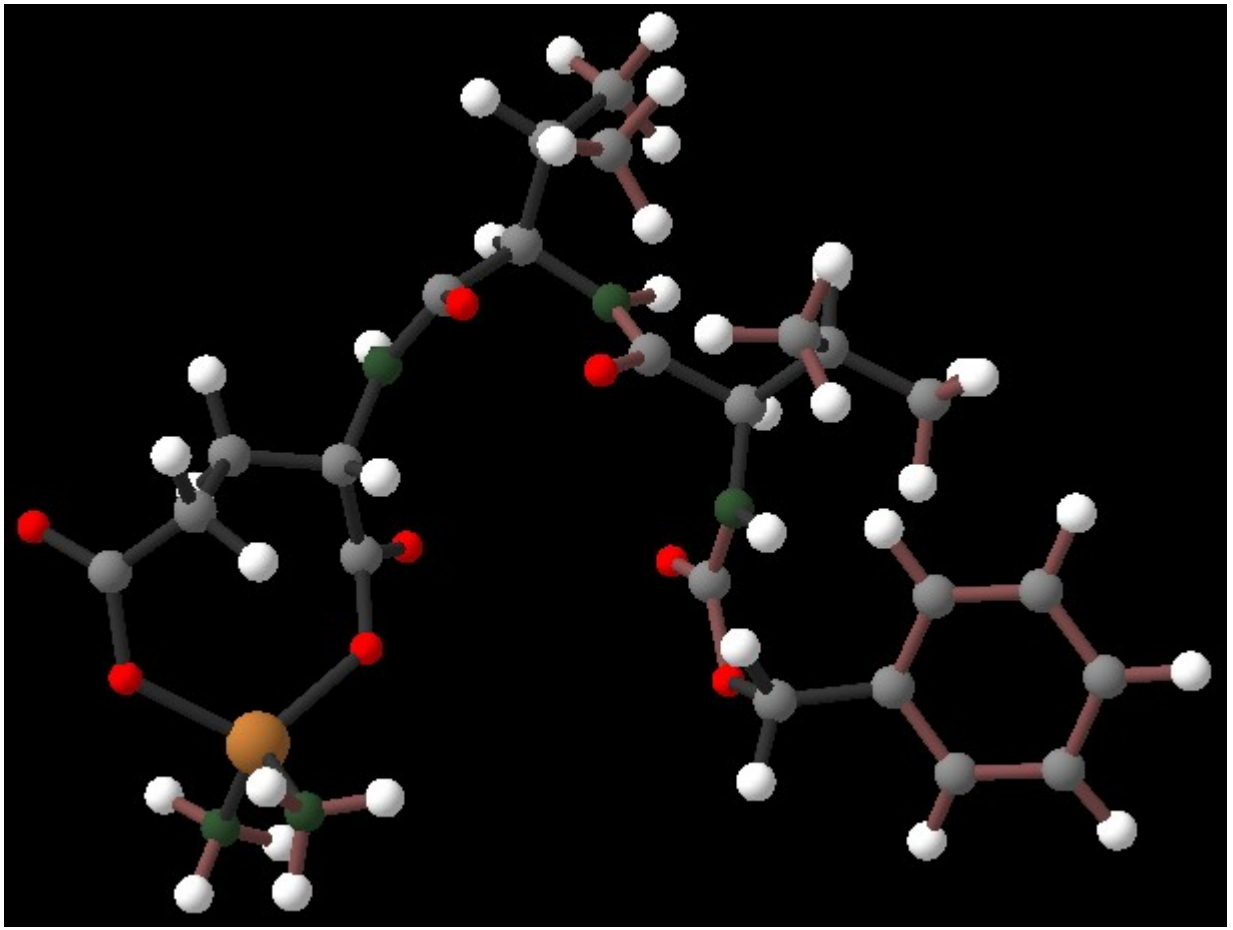
Using Molecular Dynamics + least squares



MD moves allow to solve complex, flexible structures with large cycles...
...but it can take a long time !

Using periodic least squares greatly helps the convergence, as the least squares algorithm moves all the atoms individually (taking into account restraints) and is not limited by simple moves, or a z-matrix description

Molecular Dynamics + least squares + rigid bodies



SOME DISTORTION IS NECESSARY... but sometimes you really want to avoid it
=> You can create 'rigid groups' of atoms that will only be translated/rotated as a rigid body, even during least squares.

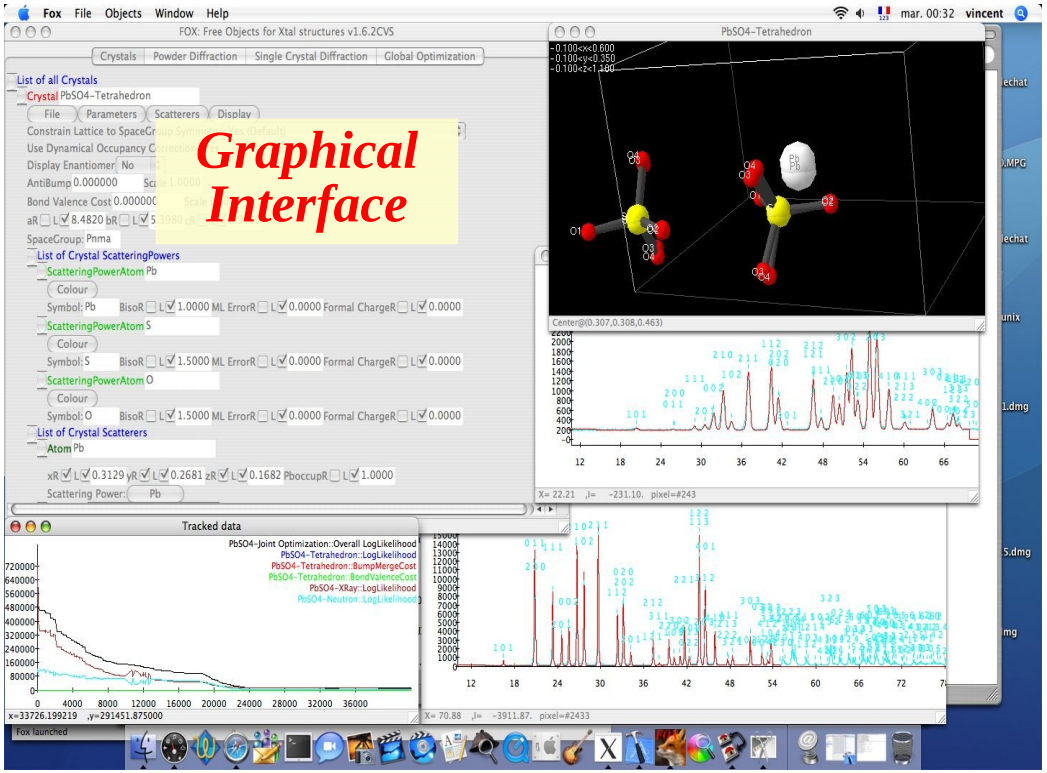
Using Fox

From the wiki: http://objcryst.sf.net

- **FOX**
  - Fox Home Page (wiki)
  - SourceForge Project
- **About FOX**
  - Download
  - Install
  - Screenshots
  - Biblio: Fox References
  - Biblio: Structures solved
  - Mailing List
  - FAQ
- **Using FOX**
  - Tutorials
  - FOX Manual (intro.)
  - - Crystal Structures
  - - Powder Diffr. data
  - - Single Crystal data
  - - Optimization Algo.
- **FOX Development**
  - Current Development
  - Features Requests
  - ObjCryst++ API
  - Getting FOX from SVN
  - Browse Code Repository

Download & Install (Linux, MacOS X, Windows)

Manual, Tutorials



command-line usage:

```
Fox example/pbso4-joint.xml --nogui --randomize -n 100000 --nbrun 10
     --finalcost 1000 -o test.xml
```

Outlook & Acknowledgements

After 10 years... Fox can now index, fit profiles, … and still solve structures !

Projects:

- ***Update tutorials***
- ***(much faster) least squares refinements***
- ***Efficient protein flexibility***
- ***More tests for Fox.Grid***
- ***Inter-atomic restraints (complicated) ?***
- ***FoxFlip (charge flipping) ??***
- ***Contribute to pyobjcryst***

Thanks to:

- ***Radovan Cerny (U. Geneva)***
- ***Jan Rohlíček, Michal Hušák (ICT Prague)***
- ***Mark Pitt (TOF), Anders Markvardsen (help with Max Likelihood)***
- ***Brian Toby + Michael Polyakov (Fourier maps display)***

- ***Lachlan Cranswick... for too many reasons to list...***

Improvements depend on user feedback !!! Send feedback, feature requests !

open-source project => add your contribution

*=> help by testing "development" versions (subscribe to the **mailing list**!)*

*=> send **bug reports** (also for cif import in openbabel)*

Get FOX from http://objcryst.sourceforge.net