Uses of X-ray Crystallography

September, 2004

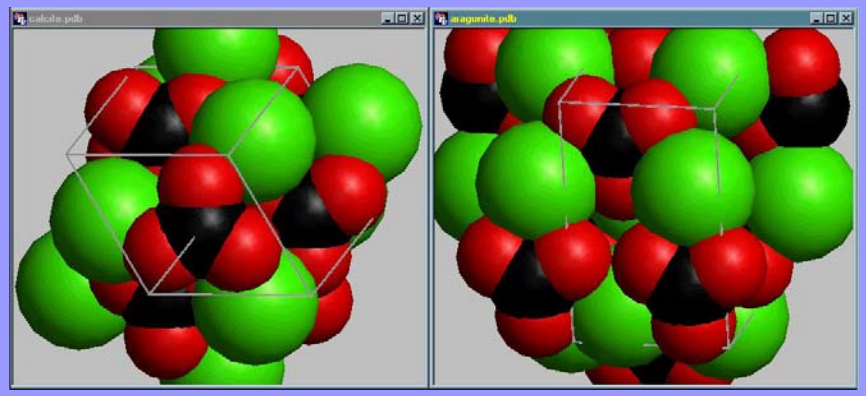
This lecture will cover:

- Health and Safety
- What we can use the method for
- Outline of theory
- Outline of practice

Health and Safety

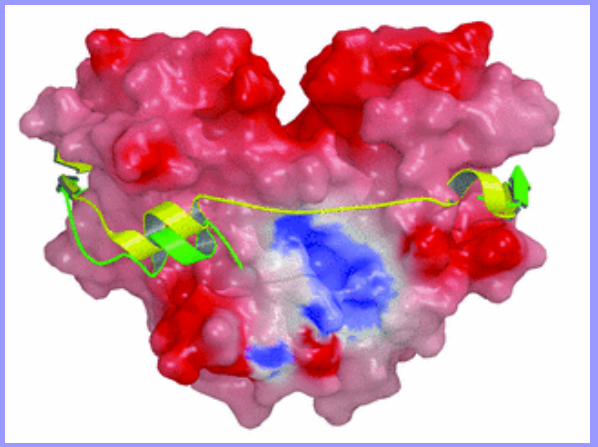
- X-rays have the potential to cause serious radiation burns.
- The equipment is safety-interlocked to current standards.
- The interlocks are checked weekly.
- People wishing to use the equipment must be trained and registered.
- Members of chemistry may observe the equipment in use without registering.

Historical View of Crystallography



The origins are in mineralogy. The spatial arrangements of atoms explain the properties

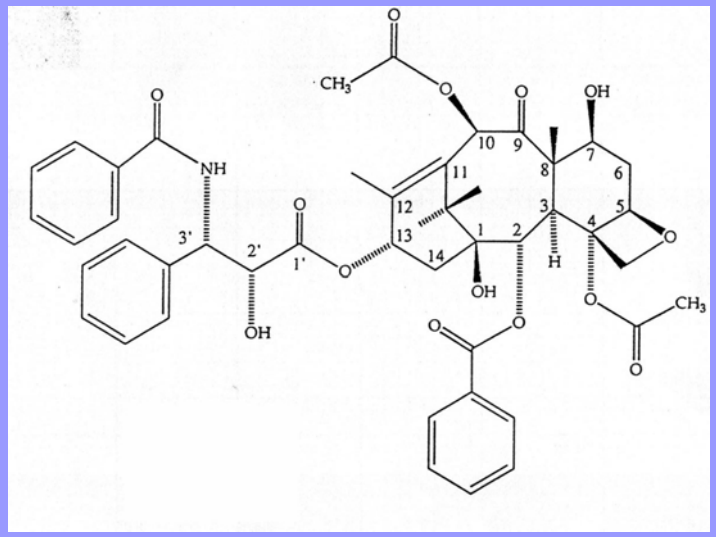
And the other end of the scale -



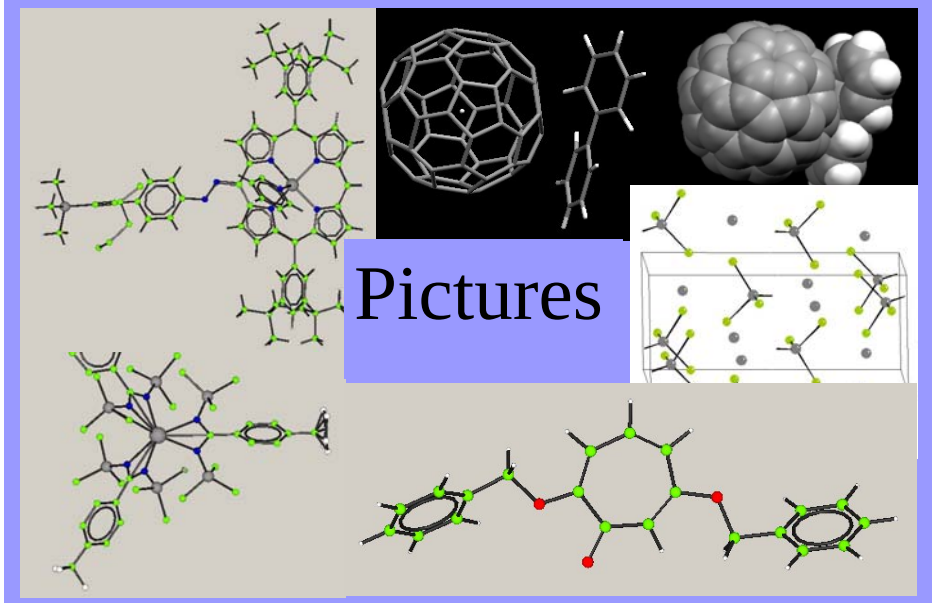
Interaction of a Plague pathogen with a protein molecule.

A role in Chemistry:

How would you characterise this molecule?

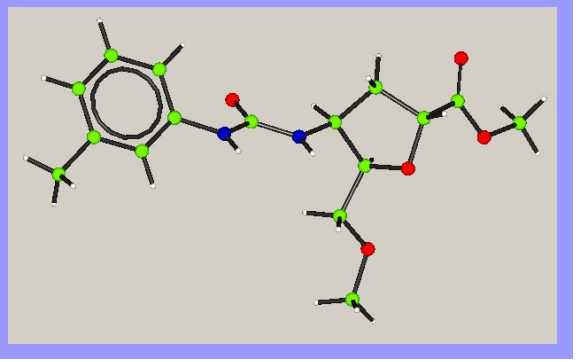


What do we want?

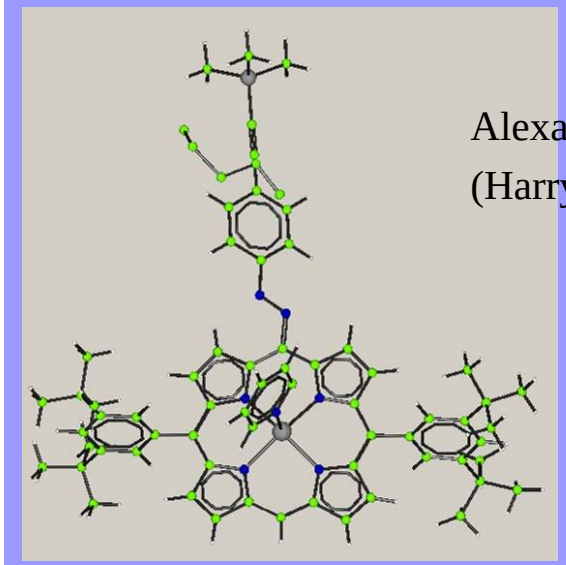


Why do we want crystal structures?

- Gross structure – atom types and connectivity
- Stereo chemistry
- Fine detail – bond lengths, angles and torsion angles
- Electron density distributions
- Crystal Structure



Have we made what we thought we had made?



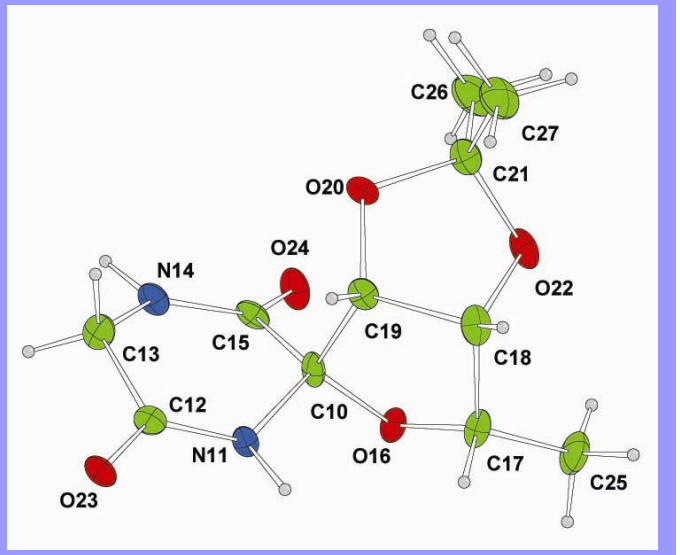
Alexander Kripovacik

(Harry Anderson Group)

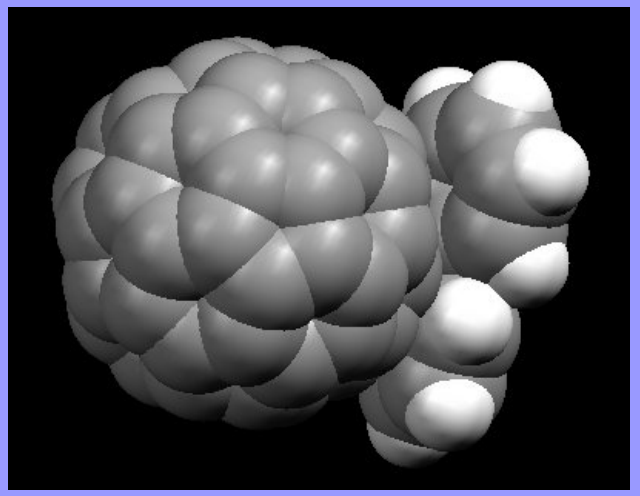
Is the stereo-chemistry correct?

By Mattias Muller

(for George Fleet Group)



What does it look like?



Jane Haggit (Malcolm Green Group)

Single Crystal X-ray Structure Analysis

Pros:

Usually gives definitive 3-dimensional information

Usually gives all of the structure

Usually very fast – 1hr-24 hrs.

Inexpensive on consumables – £10-£50 per structure

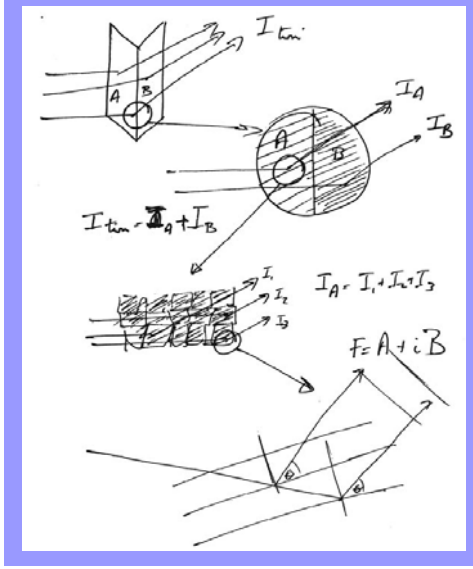
Cons:

Material is in solid state

Needs single crystals

May fail altogether

The Physics

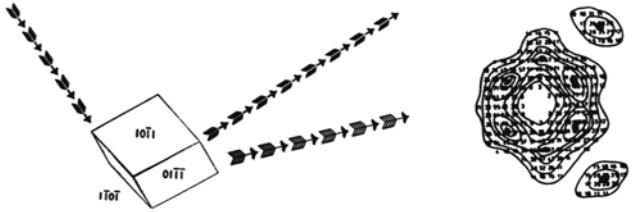


The crystal acts as a 3D diffraction grating. Monochromatic X-rays falling on the crystal interfere constructively when the Bragg angle is satisfied.

The emergent beams have an intensity and a phase

The Maths

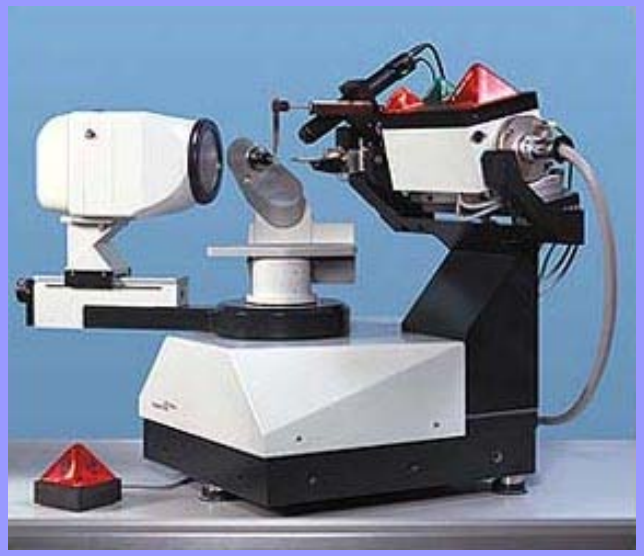
$$\rho(xyz) = \frac{1}{V}\sum_h\sum_k\sum_l |F(hkl)|e^{-2\pi i(hx+ky+lz-\alpha(hkl)}$$



$$F(hkl) = \frac{V}{abc} = \int_o^a\int_o^b\int_o^c \rho(xyz)e^{2\pi i(hx/a+ky/b+lz/c)}dxdydz$$

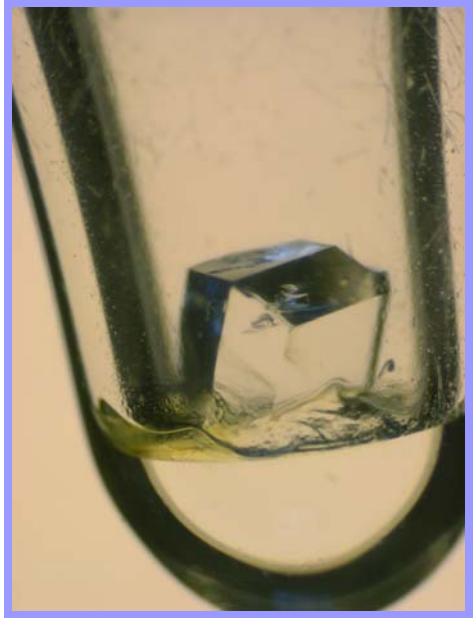
$$F(hkl) = \sum_o^n f_j e^{2\pi i(hx_j/a+ky_j/b+lz_j/c)}$$

The Machine



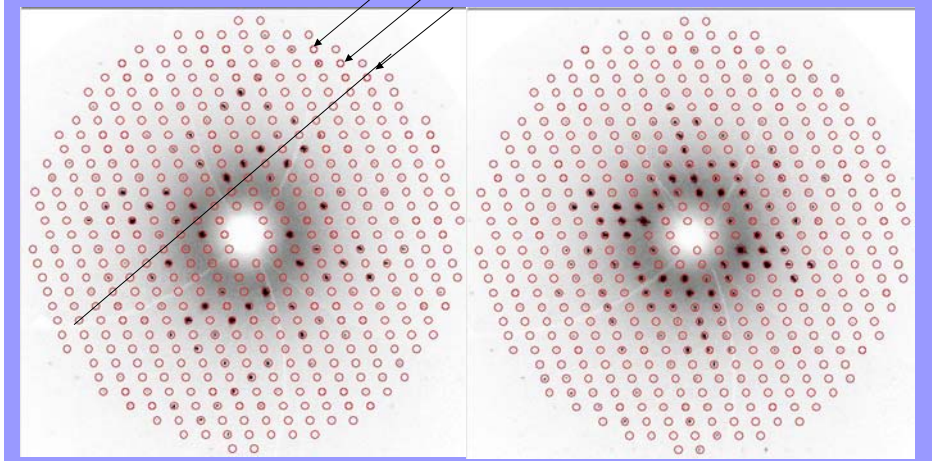
Nonius KCCD Diffractometer

The Sample



The Diffracted Image

The diffraction pattern is a regular lattice

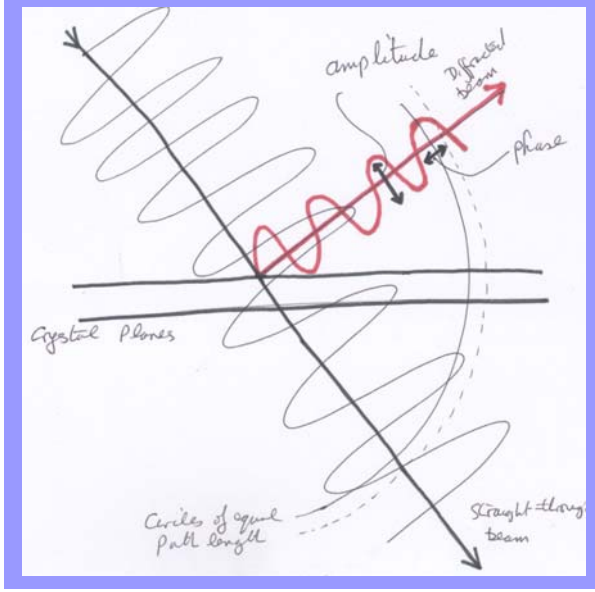


The Data

```
11.471     8.223    14.272    90.000    93.177    90.000 p2

 0   0   1  8167.4    225.5      0   1   5  1570.9     39.8
 0   0   2  1832.6     95.3      0   1   6   826.1     18.2
 0   0   3  1282.5     62.6      0   1   7   751.8     15.3
 0   0   4   466.4     21.3      0   1   8   538.8     11.7
 0   0   5   472.4     19.8      0   1   9   741.9     26.2
 0   0   6  1676.1     67.2      0   1  10    78.7      8.0
 0   0   7   198.8      8.3      0   1  11   483.1     15.7
 0   0   8   564.3     15.9      0   1  12   105.1     10.4
 0   0   9    69.7     27.7      0   1  13   485.8     24.2
 0   0  10     9.3      8.4      0   1  14   308.0     16.6
 0   0  11   146.8     14.7      0   1  15    26.2      9.1
 0   0  12    23.2     11.5      0   1  16    39.4     10.2
 0   0  13     8.4     10.9      0   1  17    48.2     10.9
 0   0  14   213.1     19.8      0   2   1  2997.1     83.6
 0   0  15   195.4     23.8      0   2   2  4156.8    102.9
 0   0  16    42.5     15.2      0   2   3   803.0     20.5
 0   0  17    44.0     14.5      0   2   4   696.9     17.3
 0   1   1  9068.2    212.0      0   2   5   142.5      4.7
 0   1   2  1097.4     28.8      0   2   6   533.1      9.7
 0   1   3  2421.4     60.5      0   2   7  2357.4     44.7
 0   1   4  2995.3     69.9
```

The experiment yields the intensity of many thousands of unique ‘reflections’

The Problem



The phase of a diffracted beam is actually the phase lag with respect to the incident beam.

It is not easy to measure.

The Solution

We cannot easily measure the phases, but in the 1950’s it was realised that there must be relationships between the phases, so that if a few are known, the rest can be estimated.

Modern DIRECT METHODS programs ‘guess’ some starting phases, and then try to select the best guess. This is measured by ‘figures of merit’

What do we Get?

```
C1 C 0.7256(3) 0.0278(5) 0.9619(2) 0.0350 1.0000 Uani  . .
C2 C 0.6582(3) -0.1404(5) 0.9635(3) 0.0381 1.0000 Uani  . .
C3 C 0.6803(4) -0.2686(6) 1.0480(3) 0.0477 1.0000 Uani  D .
C4 C 0.7626(4) -0.2240(6) 1.1292(3) 0.0486 1.0000 Uani  D .
C5 C 0.8268(3) -0.0545(6) 1.1241(3) 0.0416 1.0000 Uani  . .
C6
C7
N8
O9
O10 O 1.0102(4) 0.0444(6) 1.1978(3) 0.1055 1.0000 Uani  . .
N11 N 0.6986(3) 0.1659(5) 0.8739(2) 0.0434 1.0000 Uani  . .
O12 O 0.6341(3) 0.3010(4) 0.8924(3) 0.0704 1.0000 Uani  . .
O13 O 0.7396(3) 0.1356(4) 0.7858(2) 0.0702 1.0000 Uani  . .
I14 I 0.52258(2) -0.19874(4) 0.84666(18) 0.0502 1.0000 Uani  . .
H31 H 0.636(3) -0.371(5) 1.048(3) 0.0500 1.0000 Uiso  D .
H41 H 0.773(3) -0.297(5) 1.182(3) 0.0500 1.0000 Uiso  D .
H71 H 0.850(3) 0.344(4) 0.988(3) 0.0500 1.0000 Uiso  D .
H72 H 0.878(3) 0.324(4) 1.097(2) 0.0500 1.0000 Uiso  D .
H73 H 0.950(3) 0.247(5) 1.015(3) 0.0500 1.0000 Uiso  D .
loop_
_atom_site_aniso_label
_atom_site_aniso_U_11
_atom_site_aniso_U_22
_atom_site_aniso_U_33
_atom_site_aniso_U_23
_atom_site_aniso_U_13
_atom_site_aniso_U_12
C1 0.0412(19) 0.033(2) 0.0303(16) 0.0002(14) -0.0046(14) 0.0084(16)
C2 0.0430(19) 0.039(2) 0.0321(17) -0.0055(15) -0.0023(14) 0.0039(16)
C3 0.055(2) 0.037(2) 0.050(2) 0.0050(19) -0.0030(19) -0.0004(18)
C4 0.061(3) 0.045(3) 0.039(2) 0.0072(18) -0.0067(19) 0.010(2)
C5 0.044(2) 0.048(2) 0.0325(17) -0.0023(17) -0.0068(15) 0.0105(17)
C6 0.0399(19) 0.037(2) 0.0404(18) -0.0008(17) -0.0030(15) 0.0049(17)
C7 0.059(3) 0.059(3) 0.053(3) -0.001(2) -0.011(2) -0.010(2)
N8 0.068(3) 0.058(2) 0.057(2) -0.0046(19) -0.0292(19) 0.018(2)
O9 0.115(3) 0.114(3) 0.0426(17) -0.0150(19) -0.0262(18) 0.033(2)
O10 0.074(2) 0.132(4) 0.106(3) 0.020(3) -0.052(2) -0.018(2)
N11 0.0502(18) 0.042(2) 0.0371(17) 0.0016(14) -0.0105(14) -0.0017(15)
O12 0.084(2) 0.057(2) 0.069(2) 0.0139(16) -0.0030(17) 0.0286(18)
O13 0.109(2) 0.066(2) 0.0362(15) 0.0074(14) 0.0056(15) 0.0065(18)
I14 0.05144(18) 0.05178(19) 0.04654(17) -0.00924(12) -0.01050(11) -0.00438(13)
```

Hundreds of numbers!

Exchange of Results

This is usually done via text (ASCII) files. A program is needed to convert the data into pictures and tables.

- Cif files: Crystallographic Information File format. This has become an international standard.
- SHELX files: After the name of a widely used crystallographic program.
- PDB files: Protein Data Bank format. Often used by modelling and graphics programs.

Graphics Programs

There are dozens available from free viewers to expensive modelling packages

Mercury: Free from CCDC. Has some knowledge of symmetry operations. Reads cif files. Will output bitmap files.

Encifer: Free from CCDC. As above, but also enables you to read, edit and validate the syntax of cif files. Only outputs cif files.

ViewerLite: Free from Accelrys. Reads PDB files, will output JPG,TIF etc.

ORTEP: Free with WINGX. A fully functional crystallographic plotting program. Reads cif and SHELX files.

DEMO

Structure solution from data input to final picture and tables

Converting cif files to Tables

Microsoft Excel. You will need to use a text editor to isolate the bit of the cif you want to process.

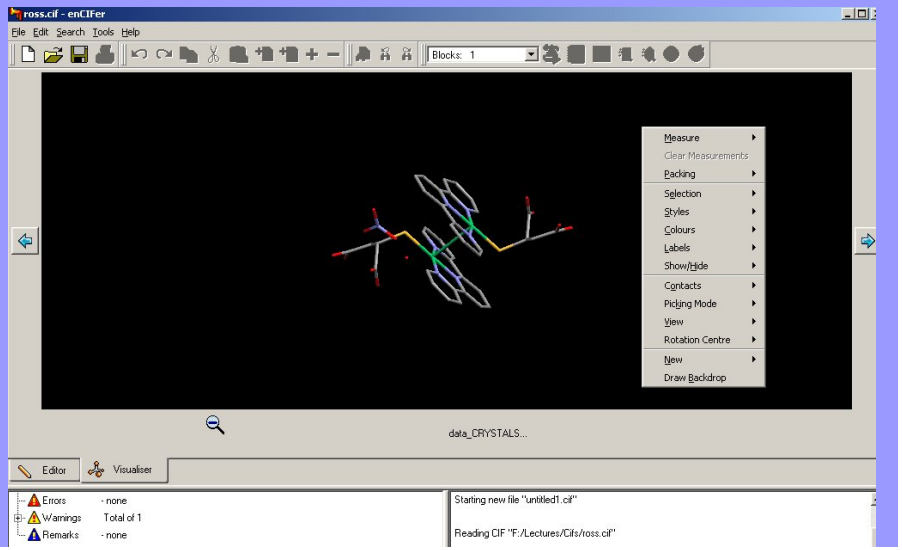
Microsoft Word. You will have to fiddle with left, right and decimal tabs to make a nice table

Use the IUCr web service ***printcif***.

http://journals.iucr.org/services/cif/printcif.html

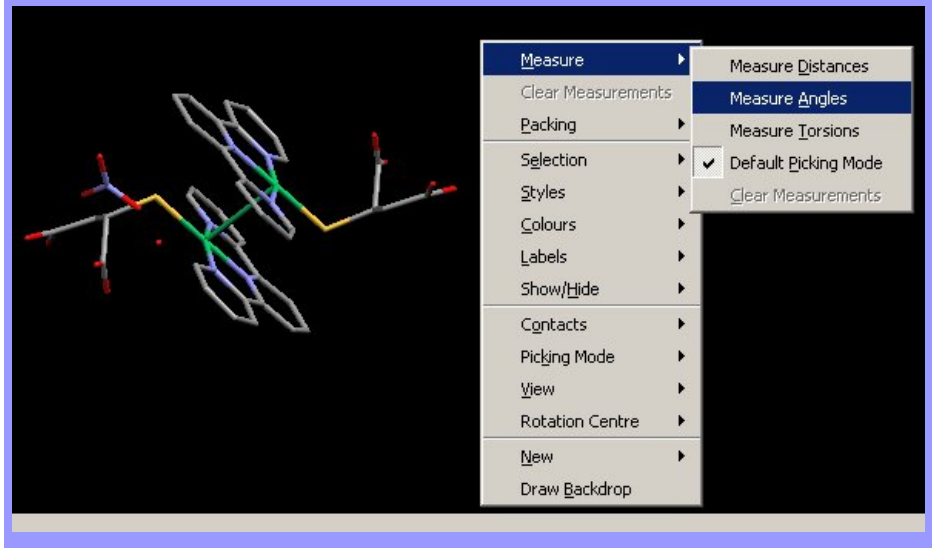
Encifer

Visualisation module



Encifer

Measuring geometric parameters



Validating the Contents of a cif file

The IUCr has both defined the syntax of a cif file, and set criteria for the quality of the data contained in the file. This can be validated with:

Checkcif as a web service.

http://journals.iucr.org/services/cif/checkcif.html

PLATON in WINGX. The program can be downloaded from:

http://www.chem.gla.ac.uk/~louis/software/wingx/

PLATON Validation

```
#===============================================================================
# PLATON/CHECK-(181101) versus check.def version of 16/11/01 for entry: CRYSTALS
# Data From: publish.cif - Data Type: CIF         Bond Precision   C-C = 0.0047 A
#
# CELL  0.71073   9.724   5.631  16.284   90.00   92.01   90.00    891.08
# SpaceGroup from Symmetry P21             Hall: P 2yb
#                 Reported P 1 21 1              ?
# MoietyFormula C15 H24 N4 O7
#      Reported C15 H24 N4 O7
# SumFormula    C15 H24 N4 O7
#   Reported    C15 H24 N4 O7
# Mr         =    372.38[Calc],    372.38[Rep]
# Dx,gcm-3   =     1.388[Calc],     1.388[Rep]
# Z          =          2[Calc],         2[Rep]
# Mu (mm-1)  =     0.111[Calc],     0.111[Rep]
# F000       =     396.0[Calc],     396.0[Rep]
# Reported   T limits: Tmin=0.990              Tmax=1.000 'MULTI-SCAN'
# Calculated T limits: Tmin=0.997 Tmin'=0.255  Tmax=0.999
# Reported   Hmax= 12, Kmax=  7, Lmax= 21, Nref= 2204       , Th(max)= 27.47
# Calculated Hmax= 12, Kmax=  7, Lmax= 21, Nref= 2262( 4105), Ratio= 0.97( 0.54)
# R= 0.0542( 2204), wR2= 0.1059( 2204), S = 0.947, Npar= 244
#===============================================================================
>>> The Following ALERTS were generated <<<

063_ALERT A Crystal Probably too Large for Beam Size .....      12.30 mm
220_ALERT A Large Non-Solvent    C     Ueq(max)/Ueq(min) .       4.83 Ratio
#===============================================================================
028_ALERT B -diffrn-measured-fraction-theta-max  low .....        0.97
222_ALERT B Large Non-Solvent    H     Ueq(max)/Ueq(min) .       4.36 Ratio
242_ALERT B Check Low  U(eq) as Compared to Neighbors ....        C(16)
411_ALERT B Short Inter H...H Contact: H(101) ..  H(211) =       2.06 Ang.
432_ALERT B Short Inter X...Y Contact: O(19)  ..  C(25)  =       2.89 Ang.
#===============================================================================
089_ALERT C Poor Data / Parameter ratio ..................       9.03
125_ALERT C No _symmetry_space_group_name_Hall given .....          ?
142_ALERT C su on b - Axis Small or Missing (x 100000) ...         20 Ang.
143_ALERT C su on c - Axis Small or Missing (x 100000) ...         50 Ang.
```

PLATON Validation

Alerts are ranked in severity.

A level: Must either be corrected or explained in the text of a publication.

B level: Must be corrected or explained in a note to the editor.

C level: May indicate a problem.

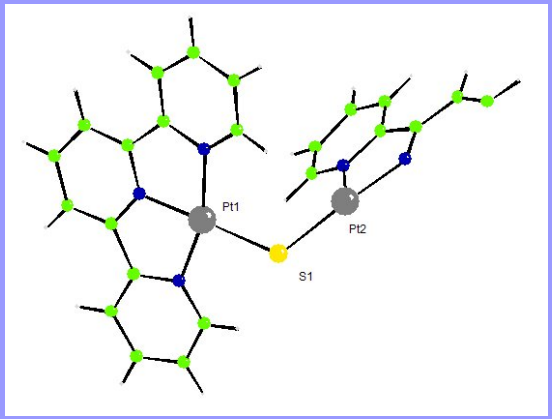
```
>>> The Following ALERTS were generated <<<

063_ALERT A Crystal Probably too Large for Beam Size .....      12.30 mm
220_ALERT A Large Non-Solvent    C     Ueq(max)/Ueq(min) .      4.83 Ratio
#=============================================================================
028_ALERT B -diffrn-measured-fraction-theta-max  low .....       0.97
222_ALERT B Large Non-Solvent    H     Ueq(max)/Ueq(min) .      4.36 Ratio
242_ALERT B Check Low  U(eq) as Compared to Neighbors ....       C(16)
411_ALERT B Short Inter H...H Contact: H(101) ..  H(211) =      2.06 Ang.
432_ALERT B Short Inter X...Y Contact: O(19)  ..  C(25)  =      2.89 Ang.
```


CCDC – The Cambridge Crystallographic Data Centre

CCDC maintains and distributes the CSD (Cambridge Structural Database), plus software for searching, viewing and analysing crystal structures.

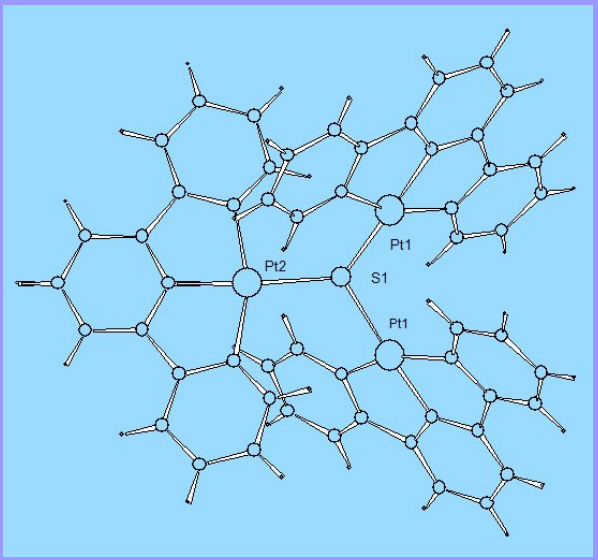
Example: ½ molecule in asymmetric unit

The crystallographic entity is ½ of the molecule. The second half (which shares the S) is generated by crystallographic symmetry.



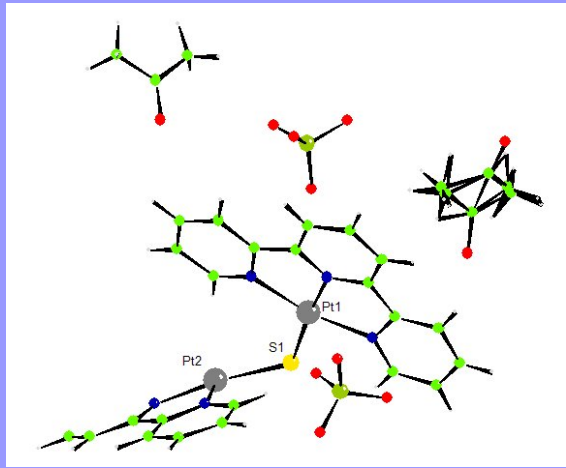
Example: The full molecule

The completed molecule has a propeller shape



Example: Solvent of crystallisation

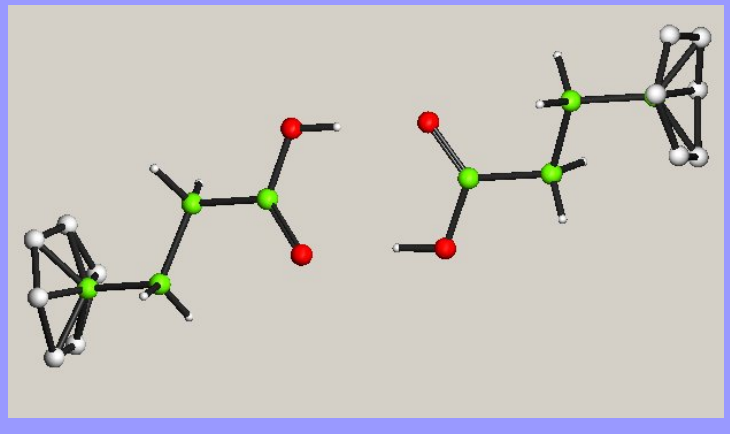
As well as the ClO_4 counter ions, there is acetone of solvation.



Example: Disorder

There are 2 molecules in the asymmetric unit, forming a hydrogen bonded dimer.

Both terminal CF_3 groups are disordered.



Advanced X-ray Course

Anyone intending to do their own X-ray analyses, or anyone simply wishing to learn more about crystallography, should come to the 2-day (6hr) intensive course.

Possible dates:

Monday & Tuesday 18 & 19 October