# CRYSTALS: Refinement and Validation Tools

**David J. Watkin**, Richard I. Cooper and Amber L. Thompson

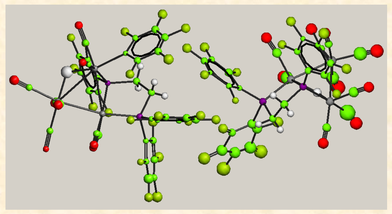
**Chemical Crystallography, Department of Inorganic Chemistry, Oxford.**

**Contact – David.Watkin@chem.ox.ac.uk**

*A new version of* ***CRYSTALS*** *will be released late Autumn 2009†*

## Understanding Poor Data

A non-Uk, non-Kccd user, non-CRYSTALS user e-mailed to say that he had was having problems refining a structure, and could CRYSTALS[1] do a better job?



Crystal Data[2]

$C_{63}\ H_{10}\ F_{35}\ O_{17}\ P_4\ Ru_6$

$P2_1/a$, 21.54 12.54 28.76 β=93.52

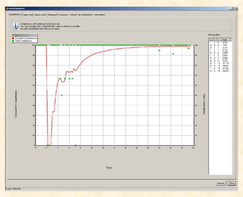
17845 reflections, 155 non-H atoms

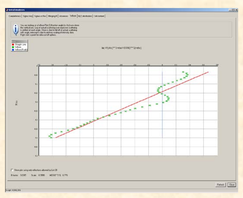
Two different molecules per asymmetric unit

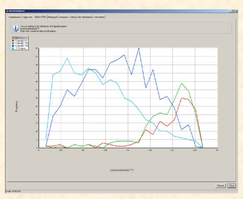
R1=50.4% (17845 data)

R1=12.4 (3258 with F>4sigma(F))

**Normal Data**







Simple plots may indicate the over-all quality of the data. Characteristic shapes are easily recognised.

**Completeness vs resolution**

Low completeness in high resolution shells often points to failures in the reflection searching process, since zero or negative intensities should count towards the completeness.

**Wilson Plot**

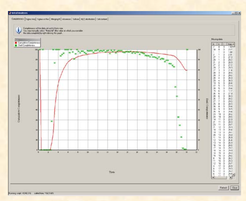
The gradient of the Wilson plot is a measure of the overall temperature factor, and should normally correspond to a value of about 0.02 to 0.05.

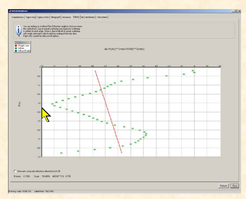
A very low or negative value may be symptomatic of neglect of a theta dependent absorption correction, or a serious failure in the data processing.

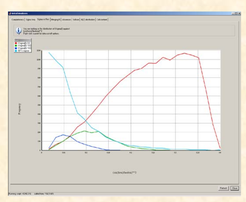
**Intensity distribution vs resolution**

Most of the reflection in the "Poor Data" set have I<3Sigma(I). Though weak data can be important,[4] in this case it overwhelms the good data.

**Poor Data**

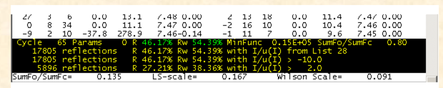


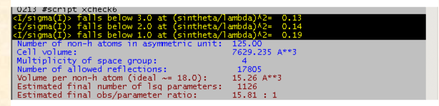


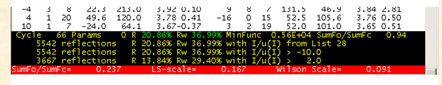


## Refinement

- Low completeness alone does not indicate a bad data set. In the absence of space group ambiguities, disorder, twinning etc, completeness as low as 70% can still give a reliable structure. When it is due to weak data, there will be problems with the refinement.
- The Wilson Plot and the text diagnostic (right), suggest that the high angle data are pretty worthless. These suggest trying a refinement using only the data with Theta < 18.5º.







Refinement with all data. Either the model is seriously wrong, or there is a fundamental problem with the data.

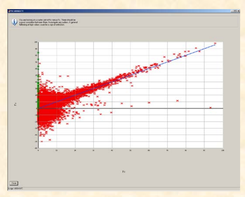
Both the Fo-Fc plot and the NPP seem to suggest that there may be two different kinds of data present in the reflection file.

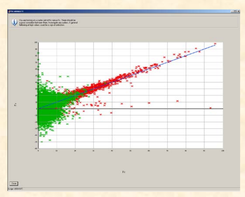
Refinement using all data with theta<18.5º. Many of the outliers are excluded.

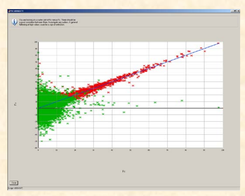
Refinement using all data with I>3Sigma(I).

All the outliers are now excluded. This will be a good thing if they are invalid data, but a seriously bad thing if they are the result of failures in the model.

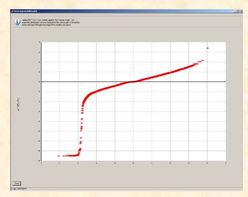
**Fo vs Fc**

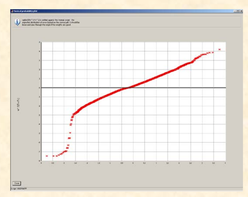


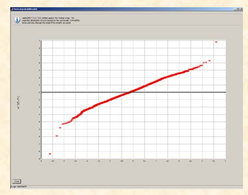




**Normal Probability**





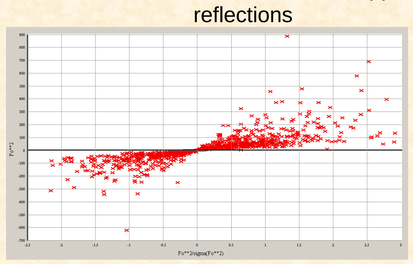


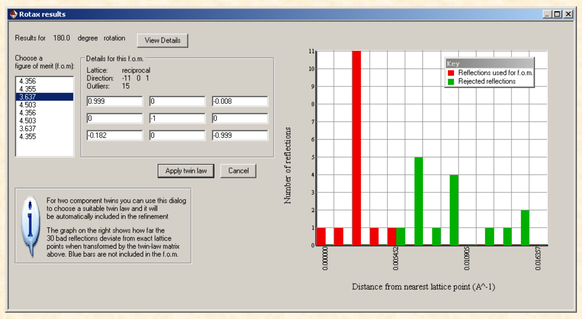
## Other Diagnostics

The Fo vs Fc plot shows several features

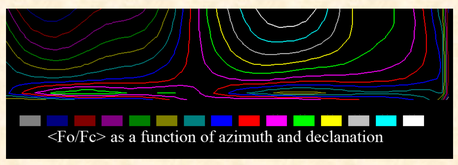
- The middle and high intensity points tend to lie above the unit gradient line. This is sometimes symptomatic of twinning.
- There is a great ill-defined mass of weak data with little correlation between Fo and Fc.
- There is a curious band of reflections where Fo is much less than Fc. These may be causing the other reflections to lie above the unit gradient.

ROTAX[3] fails to reveal any convincing evidence for twinning (right). Generally, twinning causes Fo to be larger than expected due to contributions from overlapping reflections






The scatter plot of the systematic absences (above) is reasonably symmetric about zero. There are no strong reflections, whose presence might have indicated either an incorrect space group or twinning.




A plot of <Fo/FC> against azimuth and declination (below, left) shows a systematic variation over reciprocal space. This is consistent with a failure in the multi-scan inter-frame scaling, perhaps indicating inadequate redundancy, or a slow movement of the crystal with time.

12402 Reflections processed

Minimum ratio = 0.753 at Phi(p)= 24.6 Mu(p)= 25.6 Phi(s)=178.6 Mu(s)= 25.6

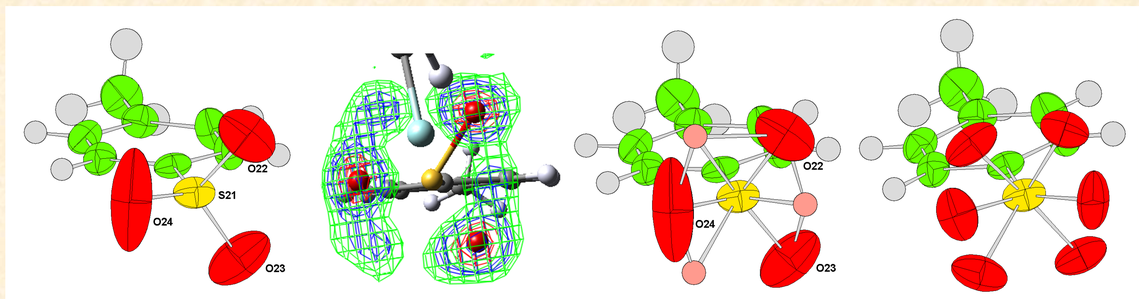
Maximum ratio = 1.153 at Phi(p)= 26.8 Mu(p)= 0.0 Phi(s)=149.2 Mu(s)= 0.0

Average ratio = 0.966

## Disorder

Ever since the invention of mail- and news-groups, the treatment of disorder has been a recurrent topic.

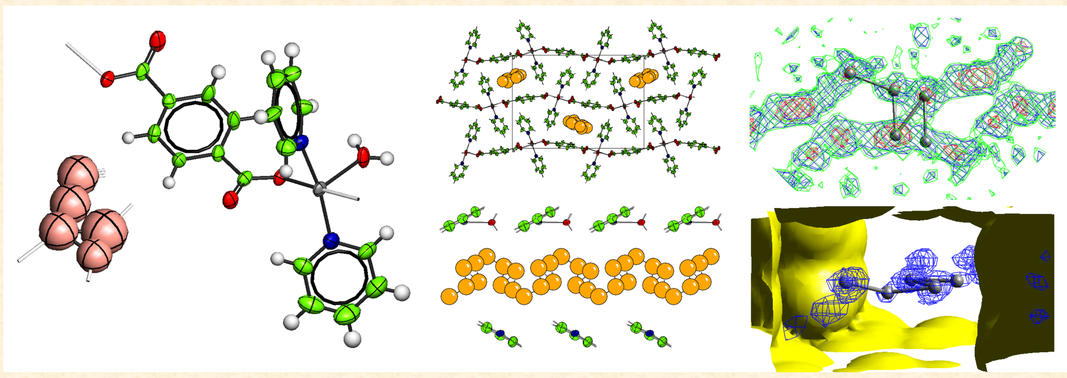
The nature of some of the questions and some of the answers suggest many new crystallographers do not understand what a mine field is hidden by the term "disorder".




If disorder seems likely, the electron density map may indicate whether a discrete-atom model is appropriate (above).[5]

However, there is no *a priori* reason why a resolved atomic model should be suitable.

In the metal organic framework material (below),[6] a packing diagram shows that the 5 Q-peaks form part of a chain running right through the crystal. The difference density in this region is a "soup", best represented with the discrete Fourier transform of the electron density in the cavity between the well-resolved atoms.



1. CRYSTALS. Betteridge, P.W., Carruthers, J.R., Cooper, R.I., Prout, K. & Watkin, D.J. (2003). *J. Appl. Cryst.*, **36**, 1487.
2. Sánchez-Cabrera, G., Leyva, M. A., Zuno-Cruz, F. J, Hernández-Cruz, M. G., Rosales-Hoz, M. J (2009). *J. Organomet. Chem.*, **694**, 1949–1958.
3. ROTAX. Parsons, S., Gould, R.O. and Cooper, R.I. & Watkin, D. J. (2002). *J. Appl. Cryst.*, **35**, 168–174; Re-implemented in CRYSTALS.
4. Walker, M., Pohl, E., Herbst-Irmer, R., Gerlitz, M., Rohr, J., & Sheldrick, G. M. (1999). *Acta Cryst.*, **B55**, 607–616.
5. Rauf, W., Thompson, A. L. & Brown, J. M. (2009). *Chem. Commun.*, 3874–3876.
6. Du, Y., Thompson, A. L., O'Hare, D. (2008). *Chem. Commun.*, 5987–5989.

† Richard Cooper has recently re-worked the graphical interface to give dramatically improved performance with most graphics cards. This will be available together with lots of other goodies shortly.

If you would like to try the beta version, contact us by e-mail.