CRYSTALS workshops

"Poor Quality Data"

David Watkin
david.watkin@chemistry.oxford.ac.uk

Tetraphenylene – poor quality data

Background

This is a hard, well crystalline organic material. Melting point 232-235°C. The crystals are in the form of prisms terminated with brilliant pyramidal faces. It was selected as a potential test crystal for analysing data collection and processing strategies, since its internal symmetry permits non-crystallographic statistical tests to be applied. A crystal 0.04 x 0.05 x 0.17 mm was Araldited to a fine glass capilliary. A hemisphere of data was collected using the parameter settings selected by the Nonius COLLECT software. Further equivalent data sets were collected with exposure times doubled, halved and quartered. The data was processed with the centred cell parameters, but as primitive monoclinic so that we had the systematic absences available for examination.
The data we have here is for the very fast data collection. The full hemisphere of data took **48** minutes to collect (15519 reflections). The average redundancy is 3, with 4000 reflections having a redundancy of 5 or more.
Chemical formula C24 H16, Space group monoclinic C2/c.

Instructions

Get started as in Exercise 1, with the following differences:
This time:

- Choose the workshop structure "Quick" ("Help"->"Demo")
- the SHELX format input file is called *veryfast.ins*
- choose F^2 refinement instead of F. (Click **Yes**).
- the reflection file is called *veryfast.hkl*
- when the filter dialogue opens, change the minimum I/sigma(I) from 3.0 to –10.0 (ie only reject very negative data).

Foolishly ignore the "*Initial analysis*" selection, then solve the structure with SIR92, using the default settings.
It will not be very successful.

Quit SIR, and when asked **do not** *use* the solution from Sir92. Close the advice dialog that follows.

Now re-select the "*Initial Analysis*" option in the GUIDE. (**OK**)

There may be a few moments delay while the systematic absences are loaded.

Select "*Sigma freq.*" – you will see that there are only about 500 out of 4500 reflections with I>3sigma(I).

Select the "*Absences*" tab – you will see that they are fairly symmetrically distributed about 0,0.

Select "*Wilson Plot*" – the high-angle data don't make any sense.

Right-click a blue cross on the Wilson plot somewhere near rho ($(\sin\theta/\lambda)^2$) = 0.35. Then click the "Reject data" menu item that pops up.

In the filter dialogue, round the $(\sin\theta/\lambda)^2$ upper limit to exactly 0.35.

Close this dialogue, close the Initial Analyses window and then re-run SIR92, **but this time** click the radio-button that says "*Filter Reflections using List28 conditions*".

The structure should now solve.

Note that setting a minimum I/sigma(I) threshold instead of a resolution threshold will not help solve the structure. This is because while high-angle weak reflections are just noise, low angle weak reflections have high information content and are needed for the negative quartets.

On the CRYSTALS command line (bottom left) type:
#SCRIPT XAbsDemo
You will see the systematic absences from the same crystal on two other diffractometers. Note that they are not symmetric about zero. This implies bias in the unobserved reflections, and perhaps also in the weak reflections.

The Guide will now invite you to perform Isotropic refinement.
Try it.
Because of the high R factor, the Guide will not advance past isotropic refinement.
Force anisotropic refinement.
There is little improvement in R, yet if you enable ellipses in the model window they look fair.
But what of the e.s.d's?
Type the following into the CRYSTALS command line:
\DIST
E.S.D. YES
END
Each of the bond lengths will be listed, with its associated e.s.d.

There will be an optimum resolution cutoff, I/sigma(I) cutoff and weighting scheme to give e.s.ds which really reflect the reliability of the model:

Cutoff too high => not enough data.
Cutoff too low => too much noise.

Effect of reflection filters

You can use the *reflection filter* dialog to experiment with different filters, and use the *optimize weights* dialog to experiment with weighting.
Choose "*Filter Reflections*" from the "*Refinement*" menu.

Change the I/sigma(I) threshold from -10.0 to 3.0, and then try some aniso refinement.
The R factor will be quite small, but in the Refinement Summary pane, you will see that there are less than 3 reflections per parameter!

Do a DISTANCE calculation (as above) to see the effect on the e.s.ds. Look at chemically equivalent bonds, and see how their variation compares with their esd.

Command macros

If you find repeatedly typing commands becomes tedious, create a '*Command Macro*':
From the 'Tools' menu, select 'Command Prompt'.
At the prompt, type
EDIT d.txt (you may use 'notepad d.txt' instead).

'd.txt' is an arbitary file name. This will begin the creation of a new file. Type the following lines into the file:

```
#DIST
EXCLUDE H
E.S.D YES
END
```

Save the file ("File"->"Save") and exit ("File"->"Exit")
Close the Command prompt by typing "EXIT".
Now, on the CRYSTALS command line, type:
#USE D.TXT

False minima

When data is short, there is always the worry that the structure is in a false minimum. This can be tested by perturbing the model and seeing if it returns to the same minimum by refinement.

On the "Filter reflections" dialog, set "I/Sigma(I) *is at least* 2", and uncheck any sin(theta) conditions. Perform some anisotropic refinement. If the model looks OK, then from the Tools menu, select 'Probe minima'.

This Script probes the minimum by applying a different random perturbation to all the coordinates of all the atoms and then refining the model. The default rms perturbation is 0.25A in each coordinate, giving an average atomic displacement of 0.4A.

Run the 'minima probe' using the default settings.
At the end, recover the original model by checking the "Use original model option" and click OK.

Increase the rms perturbation to 0.5A (giving 0.9A per atom), and try again.

For amusement, recover a good model, uncheck the "I/sigma(I) is at least" box, and set "I/sigma(I) is at most" to 2.0 (i.e. refine against only the weak data). Now try 'Minima Probing' this. Occasionally, if all the random perturbations happen to be small, the structure may be recovered.

Hydrogen Atoms

If you have lost a good model structure, you may recover one using "Structure" -> "Undo/backup" dialog.

Recover a good starting model, set the "I/sigma(I) is at least" threshold to -10, and uncheck any other thresholds that are set.

Open the Guide, and choose 'Add Hydrogen'. The pink (found) atoms will be seen to be in very poor agreement with the white (computed) sites.
Click "Continue", then remove the hydrogen using "Structure" -> "Remove hydrogen".

Set the filters: I/sigma(I) *is at least* 2.0 and sin(theta) *is at most* 0.35 and try again.
You will see that the atoms found in the difference map are reasonably near to the predicted positions.

Select more anisotropic refinement, but in the dialog box, also enable hydrogen position refinement.
You now have a stable refinement with no restraints yet an observation:parameter ratio of little over two.