Structure Refinement

Martin Lutz

Bijvoet Center for Biomolecular Research, Crystal and Structural Chemistry, Utrecht University, Utrecht, The Netherlands.
m.lutz@uu.nl

Introduction

Generally speaking, least-squares refinement is the optimization of a model with respect to the observed data. In crystallography the most used minimization function is

$$M = \sum (Y_{obs} - Y_{calc})^2$$

One has to define what is considered as observation: I, F^2 or F which are all obtained by data reduction. Another option is to use the raw pixel intensities from area detector measurements.
For the refinement the following matrix procedure is usually employed:

$$\mathbf{A}^T\mathbf{W}\mathbf{y} = \mathbf{A}^T\mathbf{W}\mathbf{A}\hat{\mathbf{x}}$$

A is the design matrix. **W** the weight matrix which is derived from the variance of the observations. With constraints, certain parameters are removed from the refinement. With restraints additional knowledge is added as extra information.
All observations should be independent. Also the errors should be independent and should approximately follow a normal distribution.

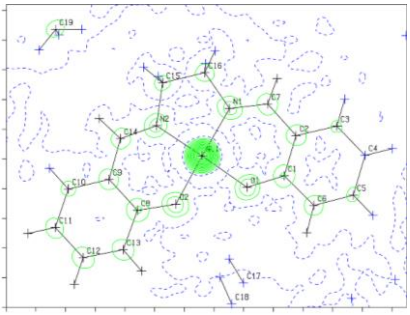

Figure 1. Fourier map in the molecular plane

The refined model consists in most cases of spherical atomic scattering factors. This can sometimes be improved by a nonspherical model. A special situation occurs with diffuse electron density which is caused by disorder. An atomic model is here inappropriate in some instances. Nevertheless has the proper treatment of disorder significant influence on the ordered part of the structure.
Before a refinement is started, the model should be as complete as possible.

Weights

At the end of the refinement, when the model is complete, the weights can be adapted. Several weighting schemes have been suggested to highlight certain structural features. A proper weighting scheme also contributes to the stability of the refinement procedure.

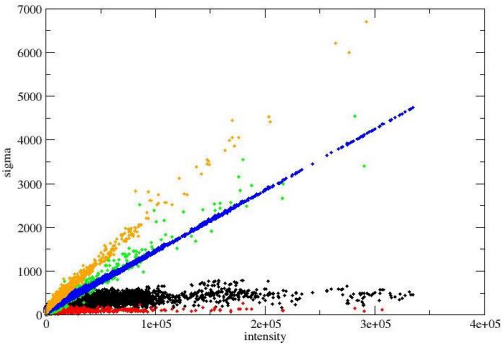

Figure 2. Influence of the data reduction procedure on the sigma's of reflections

Validation

Least-squares refinement is a very stable process and leads to results under nearly all circumstances. Still, not all results are correct. It is up to the researcher to validate the structural model. Really essential in the structure validation is the chemical knowledge. Additionally, there are mathematical and graphical tools available for the validation of the structure. It is important to carefully inspect the listing file of the refinement program, especially the correlation matrix and the standard uncertainties of the parameters. Graphical tools are useful for preparing scatter plots, normal probability plots and histograms. These graphical tools can also be used for the detection of outliers. Outliers can indicate shortcomings in the model or measurement errors. It is essential that outliers are properly handled.

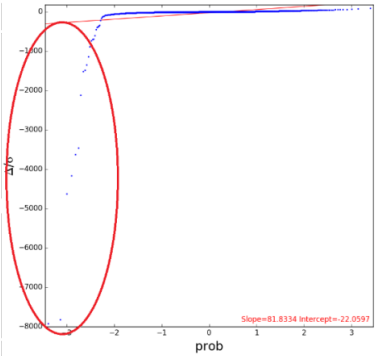

Figure 3. Normal probability plot indicating a large number of "too weak" observations.

After the last cycles of the refinement, the residual density map should be flat and featureless. Peaks and holes in the map are an indication that the model is not yet optimal. In the end, the structural model should be "fit for purpose". Further improvement is sometimes limited by data quality and resolution. A careful description of the model shortcomings is recommended in the publication process.

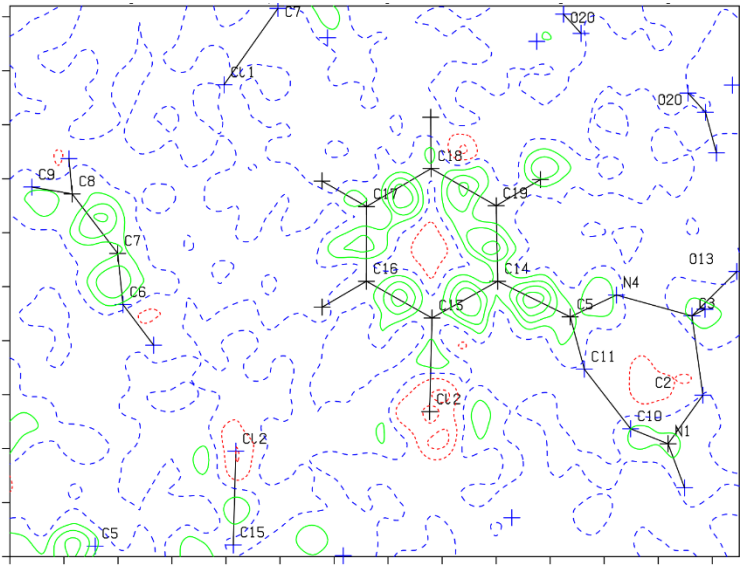

Figure 4. Residual density map in the plane of a phenyl ring.

References

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

D. Watkin (2008). *J. Appl. Cryst.* **41**, 491-522.

E. Prince (1994). Mathematical Techniques in Crystallography and Materials Science. 2nd edition. Springer Verlag, Berlin.

W.C. Hamilton (1964). Statistics in Physical Science. The Ronald Press Company, New York.