Validation of F(obs)/F(Calc) Data with PLATON (Version 20-01-10)

The result of a crystal structure determination is generally available as two files. In the SHELX world they have the extensions .cif and .fcf. Unfortunately, in many cases, only the .cif file with the model parameters is available for structure validation. Many issues cannot be pursued such as the reason for an unexpectedly high R-value that might in fact be due to an incomplete analysis. The availability of the reflection data (.fcf) solves this issue. Recently, it was also shown that fraud is less likely to go undetected with the availability of both files.

The current implementation of CheckCIF/PLATON takes both files (.cif and .fcf) into account when available. The validation report comes as two files with extensions .chk and .ckf. The CheckCIF report (on .chk) is extended with ALERTS generated by the analysis of the reflection data. The .ckf file provides a supporting listing of the various analyses for detailed inspection.

Following are details about the information that is listed in the .ckf listing.

Section 1: General Data
The files on which the analysis is done are listed along with a guestimate on the program that produced the .fcf file. Data names should be identical in both .cif and .fcf. The same applies for the cell dimensions. Also the SHELXL weight parameters are listed as found in the .cif. The latter are used to recalculate R and S values.

Section 2: Deviating I(obs), I(calc) Data
Reflections with a ratio larger than 3.0 of abs[I(obs)-I(calc)] / Sig(I) are listed.
Ratio's are calculated both on the reflection Sigma(I) and sigma's that take the weight parameters into account (SigW(I)). Large values in the last column may indicate reflections that might be effected by systematic error. With a good explanation (e.g. behind the beamstop) those reflections can be taken out of the refinement. The section ends with some statistics.

Section 3: Missing Reflections
The Friedel averaged data set is checked for completeness. All reflections that are missing below a resolution of sin(theta) / lambda < 0.5 are listed. Missing reflections below the minimum resolution as reported in the CIF are starred. The expected intensity (as calculated from the model data in the CIF) of a missing reflection is compared with the largest intensity found in the data set. A reflection with a large ratio may have been left out due to an overflow.

Section 4: Resolution and Completeness Statistics
This section gives the completeness of the Friedel averaged data set as a function of the resolution. The table entries under 'Expected', 'Measured' and 'Missing' are cumulative. The minimum resolution expected for Acta Cryst. papers is indicated.

Section 5: R-value statistics as a function of the resolution
At the end of this section, R-values as reported in the CIF are listed with the values that are calculated from the F(calc) values in the reflection file. Both lines are expected to differ only due to rounding errors. When they differ, the reason might be that the .cif and .fcf were not produced in the same refinement run and/or with different weight parameters.

Section 6: Summary of Reflection Data
A summary is given the reflection data in the CIF. Expected, actual and reported (in the CIF) values are listed.

Section 7: Intensity Distribution
This section visualizes the percentage of observed data (i.e. I > 2 * Sigma(I)) as a function of sin(theta / lambda). Note: percentages are given for three different Sigma levels. Only the 2 * Sigma version is displayed. The slope of this distribution (i.e. the curve consisting of the starred point just before the period) indicates the intensity decay as a function of sin(theta) / Lambda. A fast decay points to disorder. An indication for missed translation symmetry might be a low level of observed data in the first shell.

Section 8: Search for Unaccounted for Twinning
Two searches for possibly missed twinning are done based on differences between F(obs) and F(calc) values. In the first analysis F(calc) is based on the model in the CIF and in the second analysis on the F(calc) values in the CIF. In case of missed twinning, both analyses result in similar proposals for applicable twin laws. When a proper twin law was included in the refinement model, twin laws should be reported only in the first analysis since the twinning contribution is in that case included in the F(calc) values in the CIF.

Section 9: Absolute Structure analysis
This section is present only for non-centrosymmetric structures.

Section 10: Analysis of the Difference Fourier Map
A histogram is displayed of the gridpoint values. The distribution is expected to be symmetrical and centered around zero eA-3.

Unique maximum difference density peaks are listed along with their distance to the nearest atoms in the model.

Alternatively, for each atom in the model, the nearest density peaks are listed.

Similar tabulations are given unique negative density peaks.