

SQUEEZE: An effective cure for the disordered solvent syndrome in crystal structure refinement.

General

See Abstracts:

- M05 (page 66) ACA-94 Atlanta, USA.
- [ACA-98 Arlington, USA](#)

The current version of SQUEEZE has been designed, dimensioned and tested for small moiety structures containing disordered solvent molecules of the type toluene, CH₂Cl₂, tetrahydrofuran, water, methanol etc.. Anions may be treated in the same way. However, take care of the problem of uncertainty of charge balance.

Large voids currently require significant computing in the stage where the size and shape of the solvent accessible void is determined. All calculations are done in the triclinic system (data are expanded automatically when appropriate). Reflection data and FFT-array are stored in memory i.e. large structures (and high symmetry) may require large amounts of memory (change parameter NP21, that defines the available scratch area, globally to a larger value).

Implementation and Use:

SQUEEZE has been implemented as the 'SQUEEZE option' in the program PLATON.

PLATON/SQUEEZE should be compatible with small-moiety structure refinement usage of the popular program SHELXL-97 (or related incarnations).

The program is used as a filter. Input files are:

1. **shelxl.hkl** - (HKLF type 4)
2. **shelxl.res** - (complete set of refined model parameters, including hydrogen atoms but excluding any dummy atoms used to describe the disorder region)

invoke the program with:

platon shelxl.res

give on the prompt >> the instruction:

CALC SQUEEZE

or click on the main PLATON-menu option SQUEEZE.

The result will be in two files

1. shelxl.lis: a listing file giving details of the calculations
2. shelxl.hkp: a modified reflection file against which the ordered structure parameters can be refined (the solvent contribution has been eliminated from the reflection data.)

In order to run SHELXL-97 on the 'solvent-free' F_o^2 data:

(Note: save all files you want to keep)

- cp shelxl.res shelxl.ins

- cp shelxl.hkp shelxl.hkl

- run: **shelxl**

In order to get an .fcf style file ($F_o^2 + F_c^2$ (model + solvent)) you will need shelxl.hkl (= shelxl.hkp) and shelxl.res

run: **platon shelxl.res**

with the interactive option: **CALC FCF**

Final R-values are reported on the basis of the WGHT parameters in the shelxl.res file.

There will be a difference in reflection count as compared to the SHELXL-run due to the differing number of surviving 'observed' reflections.

The procedure (starting from the original reflection data) can be repeated using the newly refined parameters when desired (This may define a 'refined' void area. However, there will be rarely a need to repeat the procedure).

The general procedure (based on a preliminary implementation of the technique) has been described in more detail in: Acta Cryst. (1990), A46, 194 as the 'BYPASS procedure' P. v.d. Sluis & A.L. Spek)

The 'difference-map' improvement potential of this technique has been demonstrated for small molecule structures. The technique should also work for protein data. However, this has not been tested by us as yet with PLATON/SQUEEZE. Current design features may cause problems when tried.

A directory containing test-examples is in

 pub/unix/platon/TEST/SQUEEZE_EXAMPLE

Notes

- The record length of the '.hkp' file has been increased above 80 to accommodate additional data (the exact details are not fixed) including the original intensity and calculated solvent contribution to the structure factors.
- The exact numbers on the SHELXL UNIT instruction are irrelevant for the SQUEEZE procedure.
- The SHELXL-TWIN instruction is not available as yet in PLATON.
- PLATON internally calculates structure factors (using the same scattering factors as used by SHELXL97) for the model given in the shelxl.ins file.
- The solvent contribution to the structure factors is taken as 'error-free'. This means that the 'solvent-free' F_o^2 keep their original su's (esd's).
- The method relies heavily on the quality of the low-order reflections. The dataset should be as complete as possible. Systematic errors may hamper the quality of the results.
- The Contour-Map option in PLATON can be used to inspect the improved difference-map (i.e. calculated with phases including the disordered solvent contribution).
- Current conditions for applicability are:
 1. Reasonable data-resolution (say 25 degrees Mo)
 2. Structure of the known part completed with H-atoms
 3. Disorder density should be well outside the vanderWaals surface of the known structure.
 4. The area to be 'SQUEEZE' should not be too large (say less than in the order of 30% of the unit-cell volume).

Interpretation of the results

1. A successful application of SQUEEZE will show the following results:
 1. A new hkl-file against which a satisfactory refinement of the discrete model can be done (purpose: good geometry, good R-value)
 2. Smooth convergence of the SQUEEZE iteration.
 3. Significant improvement of the R-value in low resolution data. (see table at the end of the listing file).
 4. The number of electrons reported to be found in a void is close to that expected

for the assumed solvent.

5. The difference map peaklist should not contain significant peaks outside the VOID areas. Peaklist on .sqz file.
2. Problems are indicated when
 1. Convergence is unstable
 2. A large number of reflections left out during the iteration process (This may be well indicative for problems with the data, and should be checked for).
 3. Large residual density excursions in the ordered part of the structure.
 3. A report on the use of SQUEEZE should always report for each (significant) independent void:
 1. Where the void is (i.e. x,y,z)
 2. Its volume in Ang**3 and multiplicity.
 3. The number of electrons recovered.
 4. Fo/Fc-listing with Original Fo and Fc including the solvent contribution.
-

Potential Problems and Pitfalls

- Be aware of charge balance problems: SQUEEZED density in the disordered solvent area might contain a charge that can have consequences for the charge, valence and interpretation of the ordered structure part.
- PLATON/SQUEEZE can take care of redundancy of reflection data on the the input '.hkl' file. However, with high symmetry space groups this can lead to some inefficiency. It helps in such cases to supply an averaged, unique dataset (Unfortunately, direction cosines will be unavailable in the latter case for post-absorption correction; of-course a preceding numerical correction for absorption is to be preferred)
- The number of recovered electrons in the solvent area is strongly dependent on the quality of the low-angle reflections. Supply COMPLETE data sets !