



Anna V. Luebben

## PDB2INS – For Setting Up Macromolecular Refinement With SHELXL



## PDB2INS – Why use SHELXL for refinement?

SHELXL facilitates the refinement of macromolecules against high resolution data :

- Enormous flexibility, even refinement of Laue and neutron data.
- Anisotropic refinement (if data:parameter ratio sufficiently high).
- Refinement of (multiple domain) twin data (non – merohedral).
- Refinement against intensities.
- Occupancy refinement.
- Modelling of complicated disorder.
- Calculation of standard uncertainties (esds) for small structures.
- Calculation of  $R_{\text{complete}}$  in addition to  $R_{\text{free}}$ .

Sheldrick G.M. Acta Cryst. **2015**, C71, 3-8.

Gruene T. Hahn, H.W., Luebben A.V. Meilleur F. and Sheldrick, G.M., J. Appl. Cryst **2014**, 47, 462-466.

Luebben, J. and Gruene, T., PNAS **2015**, 112 (29), 8999-9003.

Anna V. Luebben

21.07.2016

1



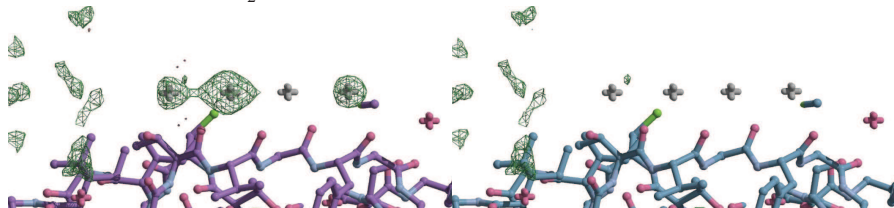
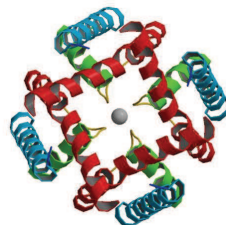
## PDB2INS – Why use SHELXL for refinement?

MthK pore, with kind permission from T. Gruene:

- 3LDC: MthK pore with 100mM K<sup>+</sup>,  $d_{\text{min}} = 1.45\text{\AA}$ .
- SHELXL: calculation of standard uncertainties and correlation coefficients with full matrix least squares.

Fixed occupancies:  
50% K<sup>+</sup> and 50% H<sub>2</sub>O

Refined occupancies:  
(S4)1.00 (S3)1.00 (S2)0.80 (S1)0.92



Köpfer, Song, Gruene, Sheldrick, Zachariae, de Groot, Science **2014**, 346, 352–355.

Anna V. Luebben

21.07.2016

2



## PDB2INS - overview

PDB2INS prepares .ins and .hkl files required for SHELXL:

- .ins file contains crystal data, atomic coordinates and instructions.
- .hkl file contains reflection data.

PDB2INS requires files available from the RCSB protein data base:

- .pdb file containing atom coordinates and general information.
- -sf.cif file containing reflection data.

Anna V. Luebben

21.07.2016

3



- A command line program written in Python.
- Comes as a packaged version.
- Can be used in **interactive modus** where the user is guided through the creation of the file.
- Extensive **command line parser**.
- Fetches files from the **RCSB PDB** server or from the **PDB\_REDO** server.
- **Automatically creates restraints** for natural amino acids, nucleotides and the most common ligands.
- Detection of disulphide bridges, incomplete residues and atoms in alternate locations with appropriate handling by **addition of specific restraints**.
- Making use of the new implementations of **chain IDs in SHELXL** (chainID:Residuenumber):

**A:1234**



### Usage: pdb2ins <filename|@pdbcode> [options]

Available options (with - as a prefix):

- w** followed by the **wavelength**.
- h** followed by 3 or 4 to enter **HKLF** manually.
- c** and six coordinates (a,b,c,alpha,beta,gamma) in Angstroms and Degree (no spaces) to enter the **cell**.
- a** use **anisotropic** displacement data if available.
- o** followed by a filename to specify the **output filename**.
- r** if a PDB code was give with the prefix '@', the **PDB\_REDO** file is fetched
- i** to skip user input.
- b** to **create a .hkl file** (structure factor file).
- d** followed by a filename for a structure factor file (cif format).

```
pdb2ins myPDBfile.pdb -w 1.54178 -h 4 -c 100,20.5,30.5,90,90,90 -i -a -o myFile.ins
```



PDB file

- **Reads PDB file** and extracts all lines with important information.
- For each atom, residue and chain an object is created to allow individual manipulation afterwards without loosing relationship.

User input

- Asks to **verify data** extracted from PDB file or **add missing information**, such as HKLF, wavelength, cell, Z value or space group.
- Asks **additional setting options** such as reset water occupancy or use anisotropic atom data and output file name.

Interpret

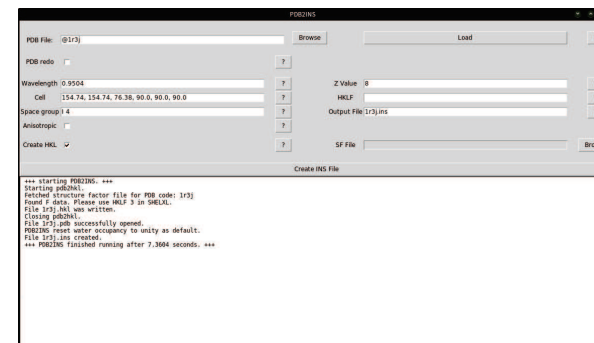
- **Searching for disulphide bridges and terminal residues** by calculating distances between atoms.
- **Calculating the symmetry operations** from the point group and stored symmetry generators.
- **Finding incomplete residues** and the last atom present in them to give appropriate restraints.

INS file

- In different steps the **INS file** is created by first writing a header **with all required crystallographic and residue specific restraints** including a list of residues missing those.
- The main part of the file consists of all **atom data sorted into Residues**. Since SHELXL now recognizes chain IDs all atoms are sorted into residues in the format 'chainID:Residuenumber'.



- Creating restraints from one PDB file as model for e.g. an wild type homologue or neutron data set.
- PDB2INS GUI is available as test version and will be developed into a stable program.





**using interactive mode:**

pdb2ins

**skipping user input interrupts:**

pdb2ins @1r3j -i

**pdb2insGUI**



**Myspecial thanks to:**

George M. Sheldrick

Jens Lübben, Claudia Wandtke, Massimo Sammito

Tim Gruene

All PDB2INS test users

**PDB2INS available at:**

*shelx.uni-goettingen.de*

**aluebbe@chemie.uni-goettingen.de**

GitHub: Av-luebben

**SHELX**

