

CRYSTALS Workshop

Charge Flipping and Space Group Determination

Arie van der Lee, Institut Européen des Membranes, Montpellier, France
Lukáš Palatinus, Laboratoire de Cristallographie, EPFL, Switzerland

Contents

1	Introduction	1
2	A simple and fast example - bt6337	1
3	Change of symmetry - CCDC 677069	2
4	Llewellyn's space group problem revisited	3
5	Unclear symmetry - Florina's compound	4
6	Choosing between symmorphic space groups - CSD-FOYTAO01	5
7	What to do if things go wrong	5

1 Introduction

This workshop introduces structure solution by charge flipping such as it is interfaced by *CRYSTALS*. It shows in addition how the space group is determined *after* the structure solution step. The *CRYSTALS* interface prepares the input file for the charge flipping program *SUPERFLIP*, which comes with the *CRYSTALS* distribution kit, but which can also be downloaded from the EPFL server at superspace.epfl.ch/superflip/.

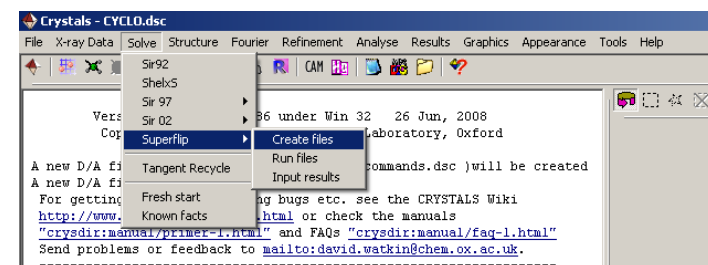
The structure solution methodology of charge flipping is fundamentally different from that of direct methods. Contrary to direct methods, space group information is not important for charge flipping, since structures are solved in *P1*. The chemical formula is not needed as well for the structure solution as such,

although it is advantageous to know at least the chemical elements that are present in the structure for the subsequent automatic map interpretation.

2 A simple and fast example - bt6337

The data in this section were taken from the IUCr server (code bt6337). Start the *CRYSTALS* program, change the working directory (File → New Working Dir) to where the data are installed, normally under `\Wincrys\demo\chargeflip\IUCRbt6337`. Import cell parameters and chemical element info via File → Open Shelx Structure. The structure is of course not there, but most data reduction programs export a short *SHELX*-type structure containing this information to be used for structure solution. Choose the bt6337.ins file, answer the different questions as you like and import at the same time the reflection file bt6337.hkl. Note that *CRYSTALS* still asks for a space group - *P1* here - although this is strictly speaking not necessary. The only utility for charge flipping is that the reflection data are averaged according to their corresponding Laue symmetry (advantageous only if the Laue symmetry is higher than triclinic), which enhances slightly the quality of the resulting electron density map. Note also that a dummy elemental composition is used, here $C_1Cl_1N_0O_0Zn_1$ instead of the final $C_{17}H_{23}Cl_2N_3O_2Zn$.

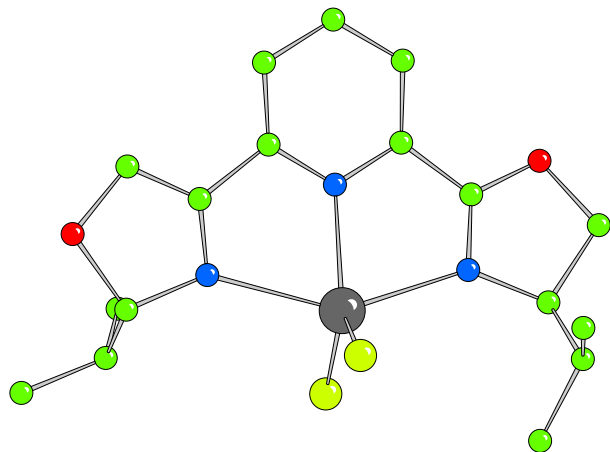
Now try to solve the structure through charge flipping. Choose Solve → Superflip → Create files and answer 'yes' to the question whether a back-up of the *CRYSTALS* database has to be made. The interface is still under construction; in a few cases things may go wrong and then you can rely on your back-up. The 'Superflip preparation' pop-up window shows default values which should work in most cases. Obviously, use 'Superflip parameters for difficult parameters' if it does not work using default parameters. For advanced users the possibility exists to edit the input file, which is not more difficult than a *SHELX* or a *SIR* input file, but you need at least the *SUPERFLIP* documentation.



For the moment, press ok and see that *CRYSTALS* launches Superflip in a separate DOS-box which should run for a few seconds, before returning to *CRYSTALS*. *CRYSTALS* analyses the log-file from *SUPERFLIP* and concludes that the structure is solved and that there is no good symmetry agreement between the resulting electron density map and possible symmetry operators in a triclinic lattice. The only possible symmetry operator in a triclinic lattice is - besides the trivial identity operator - an inversion center, so we conclude that the space group is most possibly *P1*. Continue now analysing the symmetry and *CRYSTALS* concludes that the original space group is confirmed. Press ok and *CRYSTALS* launches *EDMA*, the program that analyses the electron density map and attributes element types to the peak maxima. A structure is read back into *CRYSTALS* and you can verify that the original composition is changed to a one which is close to the final one. It may happen that the oxygens or nitrogens are not correctly attributed, but this will be fixed during refinement.

Answer now the following questions; assembling the molecule is a good idea, since *EDMA* does not do any assembling. You can change as well the element types, if *EDMA* was in error. There is a possibility to do tangent recycling using *SIR92*; this may improve the structure, but it appears that it is in some cases dangerous, since *SIR*'s space group interpretation routine does not always give the same symmetry operations as does *CRYSTALS* or *SUPERFLIP*, so it may destroy your structure.

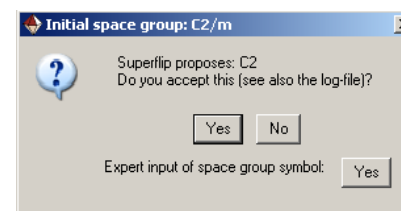
At the end of the chain of questions you should have a structure with 1-4 atoms missing and upon refinement of the scale factor an *R*-factor of about 0.18, sufficient for completing and refining the structure.



3 Change of symmetry - CCDC 677069

Although the aim of the synthesis was indeed to obtain a chiral compound (J.-C. Rossi, M. Marull, N. Larcher, J. Taillades, R. Pascal, A. van der Lee, and P. Gerbier, *Tetrahedron Asym.* **19**, 876-883 (2008)), crystallization may finally result in a racemic mixture if the solution is not enantiomeric pure. The systematic extinctions for the data collected with Cu-radiation point to either *C2* or *C2/m* as possible space groups. See what happens if *C2/m* is selected as initial space group, neglecting knowledge of the synthesis conditions, chemical formula, and *E*-statistics.

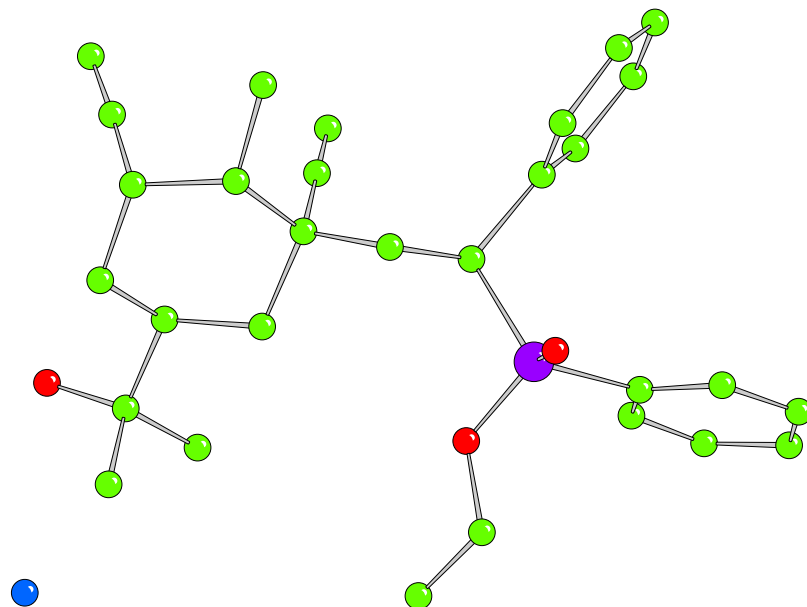
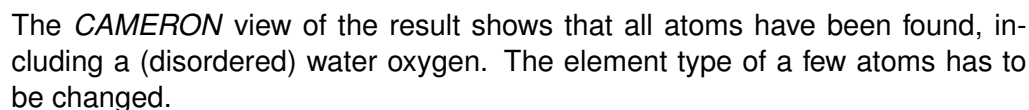
Change the working directory to `\Wincrys\demo\chargeflip\CCDC677069` and import the Shex and hkl file. Keep *C2/m* as space group and merge the Friedel reflections when importing the reflections. Set up the *SUPERFLIP* input file as explained in the first example and run *SUPERFLIP*. The calculation converges correctly and *SUPERFLIP* proposes a space group change from *C2/m* to *C2*:



In rare occasions the space group symbol is not correctly type set; the expert input of space group symbol option allows then to correct the symbol. Since the space group symbol has changed, the reflection file has to be re-imported in order to take account of different extinction condition or to keep the Friedel reflections. The latter is the case here:



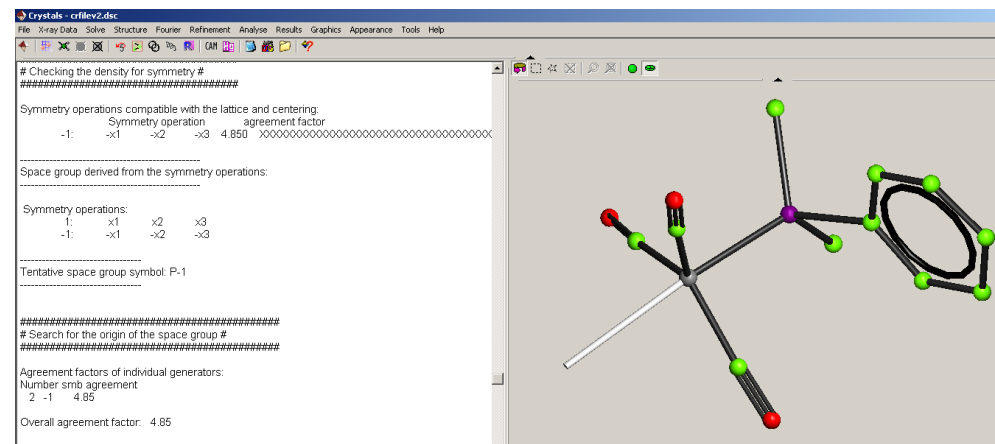
Re-import the reflection data and do not merge the Friedel reflections. The data have been measured on purpose with Cu-radiation and since the heaviest atom in the structure is phosphorous, an accurate estimate of the Flack parameter can be obtained. Proceed as before and let *EDMA* attribute atom types to the maxima found in the electron density map. Letting *CRYSTALS* assemble the molecule is again a good idea:



Note that the resulting structure may differ slightly from one *SUPERFLIP* run to another, since the starting point of the structure solution process is a random map with a zero mean density.

In 2005 there was during a *CRYSTALS* workshop a session dedicated to Llewellyn's compound under the name 'Which space group'? It was expected that the compound should crystallize in $P\bar{1}$. However, structure solution using *SIR92* with default parameters in this space group did not give the structure. As an alternative other parameters for *SIR92* were proposed (which eventually solved the structure), structure development using Fourier refinement with the very mediocre *SIR92* solution, or a structure solution in $P1$ using default parameters which was considered to be a '*useful trick*', '*giving more flexibility to Direct Methods*'. The latter method effectively works, but it is rather cumbersome to go back to the actual space group $P\bar{1}$.

It is interesting to see how *SUPERFLIP* performs with this rather weak reflection data. Import the .ins and .hkl files in the directory \Wincrys\demo\chargeflip\Llewellyn and try to solve the structure with charge flipping. Answer all questions by their default values and see that *SUPERFLIP* solves the structure in a couple of seconds. The element type of all atoms is correctly assigned and there are no atoms missing:



The left pane of the *CRYSTALS* screen shows a part of the log-file. It can be seen that the *symmetry agreement factor* for the inversion center is 4.85 %. The symmetry agreement factor ϕ_{sym} expresses the 'degree of presence' of a symmetry operator S in an electron density map $\rho(\mathbf{r})$. If $\phi_{\text{sym}}(S) = 0\%$ then $\rho(\mathbf{r})$ is perfectly symmetrical with respect to S , and if $\phi_{\text{sym}}(S) = 100\%$ then $\rho(\mathbf{r})$ is completely random with respect to S .

5 Unclear symmetry - Florina's compound

There are cases where spacegroup determination using the established methods, *i.e.* the analysis of systematic absences, completely fails. This can be due to several reasons: the data may be weak which makes that the distinction between an 'observed' and a 'non-observed' reflection is difficult to make. It is also possible that due to multiple reflections inside the crystal intensity is found for reflections that are normally 'forbidden' according to the space group symmetry. Twinning may hinder the established analysis, as well as the presence of local non-crystallographic symmetry.

A nice illustration is given by the space group determination of the crystal structure of flo19 (Dimutru, Legrand, Van der Lee & Barboiu, M., 2008, in preparation). The data are weak ($\langle I/\sigma(I) \rangle = 6.81$ for all data) and a large part of the structure gives in projection down the *c*-axis an apparent C-centering. The systematic absences suggest $P-2_12_1/n$ as diffraction symbol, but this is a non-existing symbol in Laue class *mmm* ($R_{\text{int}}(\text{mmm})=0.06$).

class	condition	$\langle I/\sigma(I) \rangle$		# refl.		t/f
		.true.	.false.	.true.	.false.	
<i>h</i> 00	<i>h</i> = 2 <i>n</i>	1.84	0.35	3	5	5.23
0 <i>k</i> 0	<i>k</i> = 2 <i>n</i>	52.20	5.69	2	2	9.18
00 <i>l</i>	<i>l</i> = 2 <i>n</i>	8.93	1.04	41	43	8.61
0 <i>k</i> <i>l</i>	<i>l</i> = 2 <i>n</i>	8.04	5.73	428	428	1.40
0 <i>k</i> <i>l</i>	<i>k</i> = 2 <i>n</i>	6.19	7.68	458	398	0.81
0 <i>k</i> <i>l</i>	<i>k</i> + <i>l</i> = 2 <i>n</i>	7.20	6.57	426	430	1.10
<i>h</i> 0 <i>l</i>	<i>l</i> = 2 <i>n</i>	8.92	8.15	230	231	1.09
<i>h</i> 0 <i>l</i>	<i>h</i> = 2 <i>n</i>	6.05	10.50	204	257	0.58
<i>h</i> 0 <i>l</i>	<i>h</i> + <i>l</i> = 2 <i>n</i>	9.27	7.81	229	232	1.19
<i>h</i> <i>k</i> 0	<i>k</i> = 2 <i>n</i>	4.72	16.24	41	38	0.29
<i>h</i> <i>k</i> 0	<i>h</i> = 2 <i>n</i>	5.43	13.91	34	45	0.39
<i>h</i> <i>k</i> 0	<i>h</i> + <i>k</i> = 2 <i>n</i>	19.49	1.71	38	41	11.43

For standard space group determination programs such a data set is very problematic. Space groups that are more the less compatible with the diffraction symbol include $Pmmn$, $Pm2_1n$, $P2_1mn$, and $P22_12_1$. Before going on (and don't read what follows), amuse yourself by trying to solve the structure with the traditional programs *SHELX* and *SIR*. The final chemical formula is (not needed for *SUPERFLIP*) $\text{C}_{62}\text{H}_{46}\text{N}_{14}$, $Z=1$.

Pick up the data in `\Wincrys\demo\chargeflip\florina (flo19.hkl and flo19.ins)`

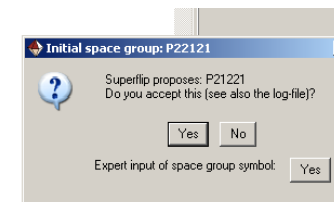
and solve the structure using charge flipping starting with *e.g.* space group $P22_12_1$ (Note that *CRYSTALS* can not handle space group symbol $Pm2_1n$ although it is a legal *IT* setting of standard setting $P2_1mn$). Surprisingly, *SUPERFLIP*, proposes space group $P2_122_1$ as most likely candidate, based on the symmetry analysis of the electron density map, resulting from charge flipping.

Tentative space group symbol: P21221

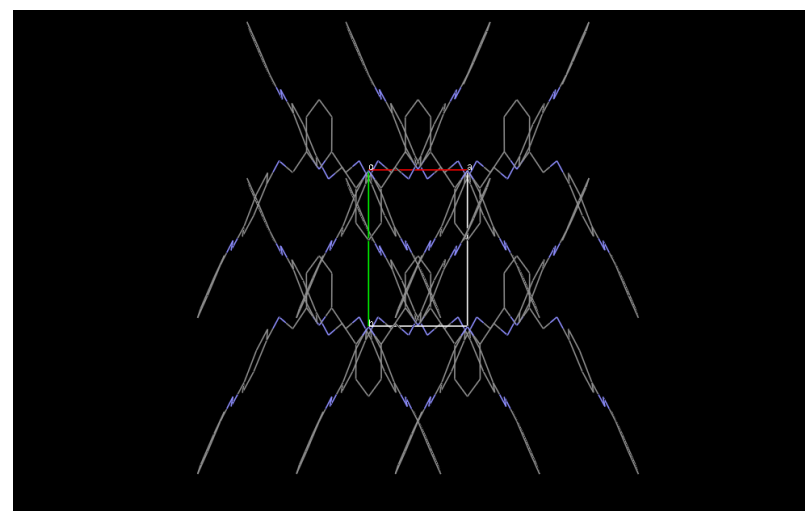
Search for the origin of the space group #
#####

Agreement factors of individual generators:

Number	sub	agreement
2	21	1.22
3	2	2.46



The reflection conditions corresponding to the 2_1^a screw axis are difficult to detect automatically because of the weak data set. The apparent n^c glide plane, however, is a result of the packing of the structure. The wireframe representation below shows that the phenyl moieties projected down the *c*-axis display an apparent C-centering.

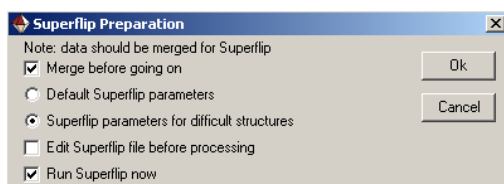


6 Choosing between symmorphic space groups - CSD-FOYTAO01

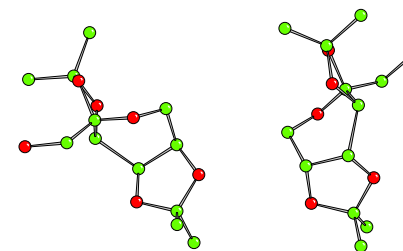
A space group is symmorphic if all symmetry operations can be described in terms of rotation axes and reflection planes all through the same point (including roto reflections), thus without screw axes and glide planes. This means that such a space group cannot be distinguished by means of its systematic absences, only by their absence. In each crystal system there are a number of space groups without any extinction conditions, *e.g.* if the Laue symmetry is $4/m$ then the space group can be $P4$, $P\bar{4}$, or $P4/m$. The established space group determination algorithms make in these cases the distinction usually on the basis of E -statistics and/or their relative frequency in data bases such as the CSD or the ICSD. Using the present space group determination algorithm, the question is not important, since it is not based on the analysis of systematic extinction conditions.

Slightly related is the case where there is an ambiguity between a non-centrosymmetric and a centrosymmetric non-symmorphic space group with the same extinction conditions, *e.g.* $Pn2_1a$ and $Pnma$, since the inversion center does not generate a new systematic extinction condition - being a non-symmorphic symmetry operator. Using *SUPERFLIP*'s structure solution method, only one run is sufficient in order to find the proper space group.

We present here the case of the structure of CSD code FOYTAO01 using data kindly put at our disposal by Dr. M. Bolte (Frankfurt). Import the cell parameters, element types, and reflection data from the files that can be found in `\Wincrys\demo\chargeflip\FOYTAO01`. In order to test the algorithm, change the space group in the space group tab from $P4$ to $P\bar{4}$. Launch *SUPERFLIP* using default options and note that it now takes much longer to solve the structure. The iterations continue up to 10000 cycles and will then stop. If you are fed up after say 1000-2000 cycles, you can stop the run by pressing CTRL-C (when the DOS-box is active). The *CRYSTALS* script will terminate by saying that the structure has not been solved. Restart *SUPERFLIP*, but now using the option for difficult structures:



The structure should solve now much faster and you can check that *SUPERFLIP* proposes the correct space group $P4$. The two independent molecules in the asymmetric unit are found and all element types correctly assigned.



7 What to do if things go wrong

SUPERFLIP is a program in continuous development and things may go wrong at unexpected times. Equally well, the *CRYSTALS* interface for *SUPERFLIP* may sometimes hang or do other things wrong. There can be large number of reasons why *SUPERFLIP* does not do what you want it to do. How do you recognise when it goes wrong? Normally, when launching the structure solution calculation - after the preparation step - a DOS-window pops up and you can see how the iterations evolve. If there is something wrong, the DOS-window flashes shortly before returning to *CRYSTALS* and you get a message in the log-window that the structure is not solved. You can get more information on what exactly went wrong by launching *SUPERFLIP* from the command line. Open for that a dos-window and go to the directory where the structure calculation are done. Launch *SUPERFLIP* by typing on the command line: `[path]superflip sflip.inflip` where `[path]` is the `[path]` where *SUPERFLIP* is installed, normally in the `wincrys` directory.

If you don't know how to do this, you can try one trick from inside *CRYSTALS*. Select again *SUPERFLIP* from the solve menu, but in the preparation step select 'Edit Superflip file before processing'.

