Methods in Chemistry III – Part 1
Modul M.Che.1101 WS 2010/11 – 9

Modern Methods of Inorganic Chemistry

Mi 10:15-12:00, Hörsaal II

George Sheldrick

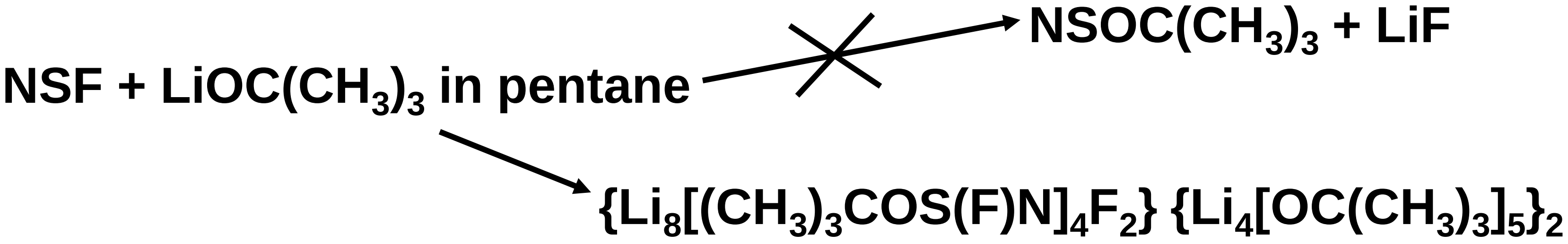
gsheldr@shelx.uni-ac.gwdg.de

Steps in an X-ray structure determination

1. Grow crystals
2. Collect X-ray data (reflection positions and intensities)
3. Assign reflection indices *h*, *k* und *ℓ* and determine unit-cell
4. Determine space group
5. Find the phases of the reflections by direct or Patterson methods
6. Calculate electron density and assign atoms
7. Complete and refine structure
8. Calculate bond lengths and angles and make pictures of structure
9. Verify structure (CHECKCIF) and deposit it
10. Write paper

The result of an X-ray structure determination

The great advantages of X-ray structure determination in contrast to spectroscopic methods is that it is very objective and that highly complex and completely unexpected structures can be determined, e.g.:



$NSOC(CH_3)_3 + LiF$

$NSF + LiOC(CH_3)_3$ in pentane

$\{Li_8[(CH_3)_3COS(F)N]_4F_2\}\ \{Li_4[OC(CH_3)_3]_5\}_2$

Chemical fantasy can be useful in such cases. The experimentally determined electron density is always blue! In spite of this, the final result is usually unambiguous. Many structures are unproblematic, but some (about 10%) require considerable patience and experience.

Incorrect space groups

At least 1% of the structures in the crystallographic databases appear to have been determined in the wrong space group; most often symmetry elements, especially inversion centers, have been overlooked.

When an *inversion center* is overlooked, the refinement of the structure is mathematically unstable. Bond lengths and atomic displacement (thermal) parameters, which would have been made equivalent by the inversion center, fly apart. The structure is simply wrong.

When *other symmetry elements* are missing, the result is that (for the example of a twofold axis) twice as many parameters as necessary are refined, but against double the number of data. The structure is not wrong, but a refinement in the correct space group against symmetry averaged data would have been more precise.

Other possible problems

1. The *precision* is relatively modest. The estimated standard deviations (esds) of interatomic distances lie in the range 0.02 – 0.001 Å (1.0 – 0.05%). Under favorable conditions pure rotational and other forms of spectroscopy can determine bond lengths two orders of magnitude more precisely.

2. A crystal structure represents an average over millions of unit-cells and a relatively long time (the time taken for the data collection, usually several hours); dynamic processes cannot be determined directly by means of X-ray diffraction, but they may still leave their mark on the structure.

3. When the contents of each unit-cell are not identical, but the discrepancies are of a statistical nature, parts of the structure are considered to be *disordered*.

4. The crystal might not be a single crystal but *twinned.*

The heavy atom method

By interpretation of the Patterson function, it is possible to find the positions of a small number of heavy atoms in the cell. These enable the calculated structure factor $\mathbf{F}_c$ (a complex number) or the corresponding amplitude F_c and phase ϕ_c to be calculated for each reflection. A difference Fourier synthesis with coefficients $|F_o - F_c|$ and phases ϕ_c enables further atoms to be found as peaks; the approximation $\phi = \phi_c$ becomes better as the structure becomes more complete. The full structure may be found by iterative application of such difference maps.

Although it is no longer considered necessary to prepare heavy atom derivatives such as bromobenzoates to solve small molecule structures (*direct methods* are so effective), heavy atom derivatives are still useful in macromolecular crystallography.

The Patterson function

$$F^2 \underset{\text{FT}}{\overset{\text{FT}}{\rightleftarrows}} \text{Patterson}$$

The Patterson function is a Fourier synthesis performed with all phases set to zero and the measured F^2-values as coefficients. The dimensions of the unit-cell and the lattice type of the Patterson are identical with those of the crystal structure. The Patterson is a vector map: each atom-atom vector in the structure corresponds to a peak in the Patterson. The peak height is proportional to the product of the atomic numbers of the two atoms and to the number m of vectors that are equal in magnitude and direction and so overlap:

$$H \propto m Z_i Z_j$$

Usually the origin peak (height proportional to $\sum Z_i^2$) is scaled to 999, so that the peak heights of the remaining peaks are given by:

$$H = 999\, m Z_i Z_j / \sum Z_i^2$$

This summation is over all atoms in the unit-cell.

One heavy atom in P$\bar{1}$

In this space group, for each atom at x, y, z there is a symmetry equivalent atom at $-x$, $-y$, $-z$ and thus a Patterson peak at $2x$, $2y$, $2z$.

Crystals of $C_{32}H_{24}AuF_5P_2$ have two formula units in the unit-cell. The two gold atoms are related to each other by the inversion center. In addition to the origin peak (999) there is a peak of height 374 with X=0.318, Y=0.471, Z=0.532, and of course a peak at $-X$, $-Y$, $-Z$, that are much higher than the rest (≤145). This peak height is in good agreement with the calculated value of 377 for a Au—Au vector.

In order to calculate x from $2x$ it would be far too simple just to divide by 2! We have to take into account that there must also be peaks in the next unit-cell of the Patterson with $X = 2x+1$ etc. Thus we calculate the possible gold coordinates as follows:

x = 0.318/2 or 1.318/2 = 0.159 or 0.659
y = 0.471/2 or 1.471/2 = 0.236 or 0.736
z = 0.532/2 or 1.532/2 = 0.266 or 0.766

These *eight* possible positions correspond to the *eight* possible ways of putting the origin on an inversion center in space group P$\bar{1}$.

Two independent heavy atoms in $P\bar{1}$

The compound $[C_{24}H_{20}S_4Ag]^+$ $[AsF_6]^-$ crystallizes in $P\bar{1}$ with a cell volume of 1407 Å^3. Two formula units per cell (72 atoms) correspond to 19.5 Å^3 per atom. It is useful to calculate a difference vector table:

	x_1	$-x_1$	x_2	$-x_2$
x_1	0	$-2x_1$	x_2-x_1	$-x_1-x_2$
$-x_1$	$2x_1$	0	x_1+x_2	x_1-x_2
x_2	x_1-x_2	$-x_1-x_2$	0	$-2x_2$
$-x_2$	x_1+x_2	x_2-x_1	$2x_2$	0

$\pm(x_1+x_2)$ and $\pm(x_2-x_1)$ have $m = 2$, the null-vectors (0) have $m = 4$;

With $Z(\mathrm{Ag}) = 47$, $Z(\mathrm{As}) = 33$, $\sum Z^2 = 11384$ and $H = 999\, m\, Z_i Z_j / \sum Z^2$ the following Patterson peaks can be predicted:

Ag—As $m = 2$ height = 272;
Ag—Ag $m = 1$ height = 194;
As—As $m = 1$ height = 96.

Interpretation of the Patterson

	X	*Y*	*Z*	height	interpretation
1	0	0	0	999	origin
2	0.765	0.187	0.974	310	x(Ag)+x(As)
3	0.392	0.099	0.325	301	x(Ag)–x(As)
4	0.159	0.285	0.298	250	2x(Ag)
. .	. .	. .	. .	. .	. .
. .	. .	. .	. .	. .	. .
14	0.364	0.077	0.639	102	2x(As)

A consistent solution would be:

Ag: $x = 0.080$, $y = 0.143$, $z = 0.149$
As: $x = 0.682$, $y = 0.039$, $z = 0.820$

For the first atom (Ag) there are 8 equivalent solutions, so one can simply divide 2x by 2. The second atom (As) must however be consistent with the Ag, so x(Ag) is subtracted from x of peak 2 and then checked whether this As position agrees with the other vectors.

Two independent heavy atoms in $P2_1$

	$x_1\ y_1\ z_1$	$x_2\ y_2\ z_2$	$-x_1\ \tfrac{1}{2}+y_1\ -z_1$	$-x_2\ \tfrac{1}{2}+y_2\ -z_2$
$x_1\ y_1\ z_1$	0	$-x_1+x_2\ -y_1+y_2\ -z_1+z_2$	$-2x_1\ \tfrac{1}{2}\ -2z_1$	$-x_1-x_2\ \tfrac{1}{2}-y_1+y_2\ -z_1-z_2$
$x_2\ y_2\ z_2$	$x_1-x_2\ y_1-y_2\ z_1-z_2$	0	$-x_1-x_2\ \tfrac{1}{2}+y_1-y_2\ -z_1-z_2$	$-2x_2\ \tfrac{1}{2}\ -2z_2$
$-x_1\ \tfrac{1}{2}+y_1\ -z_1$	$2x_1\ \tfrac{1}{2}\ 2z_1$	$x_1+x_2\ \tfrac{1}{2}-y_1+y_2\ z_1+z_2$	0	$x_1-x_2\ -y_1+y_2\ z_1-z_2$
$-x_2\ \tfrac{1}{2}+y_2\ -z_2$	$x_1+x_2\ \tfrac{1}{2}+y_1-y_2\ z_1+z_2$	$2x_2\ \tfrac{1}{2}\ 2z_2$	$-x_1+x_2\ y_1-y_2\ -z_1+z_2$	0

In the table –½ is replaced by +½, because they are equivalent.

The Patterson peaks can be summarized as follows:

0 0 0	$m = 4$	Origin
$\pm\{\,2x_1\ \ \tfrac{1}{2}\ \ 2z_1\,\}$ $\pm\{\,2x_2\ \ \tfrac{1}{2}\ \ 2z_2\,\}$	$m = 1$ $m = 1$	Harker-section at y = ½
$\pm\{\,x_1-x_2\ \ y_1-y_2\ \ z_1-z_2\,\}$ $\pm\{\,x_1-x_2\ \ -y_1+y_2\ \ z_1-z_2\,\}$ $\pm\{\,x_1+x_2\ \ \tfrac{1}{2}+y_1-y_2\ \ z_1+z_2\,\}$ $\pm\{\,x_1+x_2\ \ \tfrac{1}{2}-y_1+y_2\ \ z_1+z_2\,\}$	$m = 1$ $m = 1$ $m = 1$ $m = 1$	Cross-vectors

The symmetry of the Patterson function

Because for each vector i → j there is always a vector j → i, the Patterson must possess an *inversion center*.

In general the symmetry of the Patterson is determined by the symmetry of the diffraction pattern, which also possesses an inversion center. Glide planes of the space group become, as in reciprocal space, normal mirror planes, and screw axes become rotation axes without translation. In other words, the Patterson exhibits the same symmetry as the Laue group.

Harker sections

The symmetry of the space group leads to an accumulation of Patterson peaks in particular planes (or lines). E.g.

P2_1: Atoms at $x\ y\ z$; $-x\ \tfrac{1}{2}+y\ -z$. Self vectors at $2x\ \ \tfrac{1}{2}\ \ 2z$,
Harker section at $Y = \tfrac{1}{2}$.

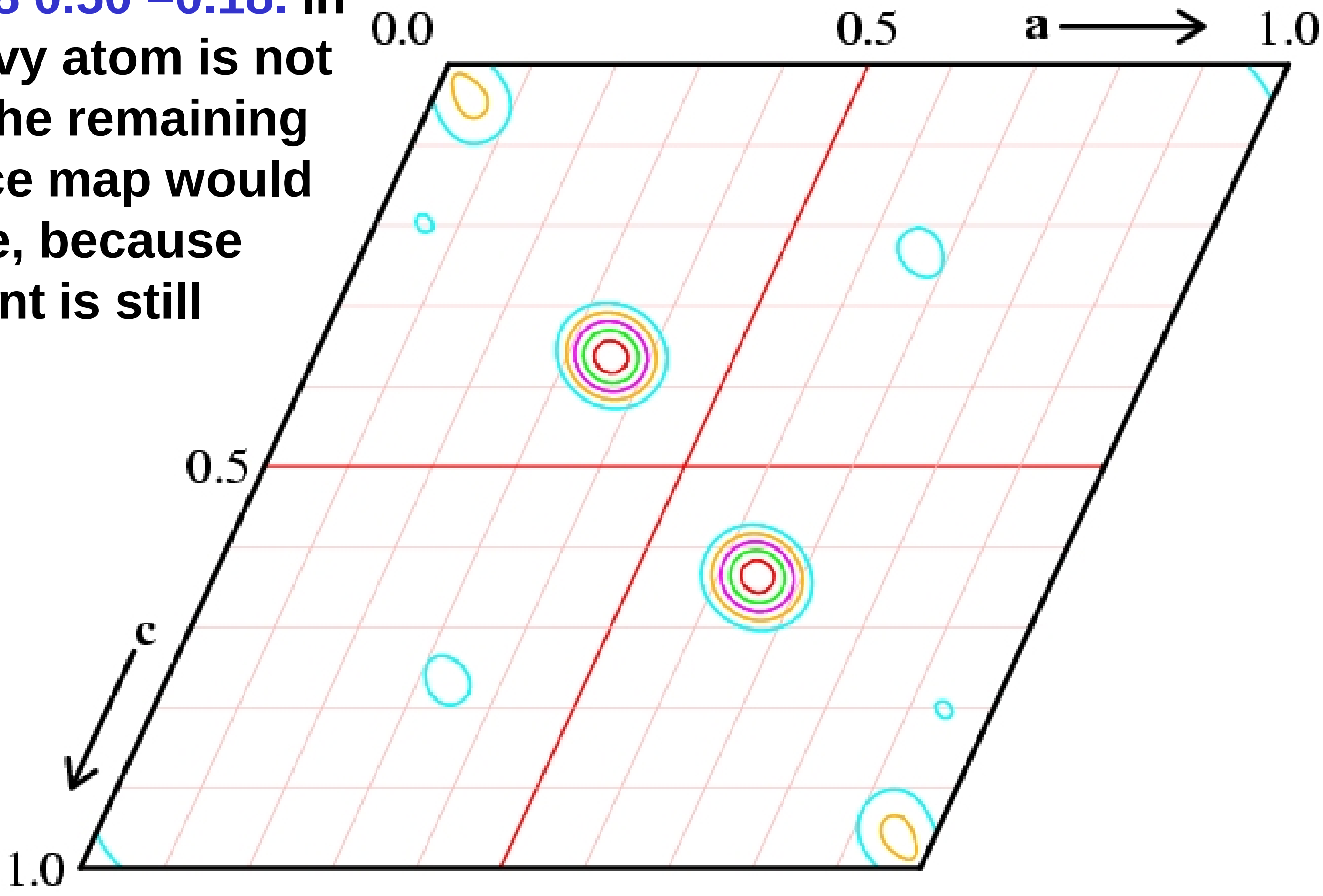
P2: Atoms at $x\ y\ z$; $-x\ y\ -z$. Self vectors at $2x\ \ 0\ \ 2z$,
Harker section at $Y = 0$.

P*m*: Atoms at $x\ y\ z$; $x\ -y\ \ z$. Self vectors at $0\ \ 2y\ \ 0$,
Harker line at $X = 0$, $Z = 0$.

P2 and P*m* have identical systematic absences but can be distinguished from each other with the help of Harker sections. The 2_1-axis in P2_1 gives rise to a Harker section at $Y = \tfrac{1}{2}$; the systematic absences $k \neq 2n+1$ would however be a more reliable method to determine this space group

One heavy atom in $P2_1$

An organoselenium compound with one independent selenium atom (two in the unit-cell) in $P2_1$ gave rise to the following $Y = ½$ Harker section. The peaks at $\pm(2x\ ½\ 2z) = \pm(0.36\ 0.50\ 0.36)$ correspond to a Se-atom at 0.18 0 0.18 and its symmetry equivalent at –0.18 0.50 –0.18. In this case one heavy atom is not sufficient to find the remaining atoms; a difference map would be a double image, because the Se-arrangement is still centrosymmetric!



Problems of the heavy atom method

Often one heavy atom suffices, with the help of a Fourier synthesis with amplitudes $|F_o - F_c|$ and phases ϕ_c, to find the remaining atoms (iteratively). Exceptions are non-centrosymmetric space groups in which the heavy atom arrangement is centrosymmetric; then all phases ϕ_c are 0º or 180º and the resulting electron density possesses an inversion center and so is a double image.

Another problem – also for structures solved by direct methods – is the assignment of the atoms, because the electron density is always blue. The density is proportional to the atomic number, but isoelectronic atoms are particularly difficult to distinguish.

Exercises

1. An atom at x, y, z in the space group $P2_1/c$ generates three symmetry equivalents: $-x, \tfrac{1}{2}+y, \tfrac{1}{2}-z$; $-x, -y, -z$; $x, \tfrac{1}{2}-y, \tfrac{1}{2}+z$. Express the coordinates of the Patterson maxima generated from a single atoms by means of a 4x4 difference table.

2. Combine these coordinates into a table with columns X, Y, Z and m. Symmetry equivalent vectors should be indicated with brackets and '±' symbols as in the following table.

3. The following strong Patterson peaks were observed for an organogold compound in the space group $P2_1/c$. Where is the gold atom (there are several equally good correct answers)?

	X	*Y*	*Z*	Höhe
	0.00	0.00	0.00	1.0
± {	0.16	0.50	0.14 }	0.4
± {	0.00	0.04	0.50 }	0.4
± {	0.16	±0.46	0.64 }	0.2