

Structure Refinement

The structure solution from direct or Patterson methods is frequently already very good. However, the coordinates are not quite accurate, the atom types of some or all atoms have been assigned incorrectly (if at all), and details of the structure are missing (H-atoms, disorders, solvent molecules, *etc.*).

The atomic positions in the first solution are not the direct result of the diffraction experiment, but an interpretation of the electron density function calculated from the measured intensities and the “somehow determined” phase angles.

Better phases can be calculated from the atomic positions, which allow re-determining of the electron density function with a higher precision. From the new electron density map, more accurate atomic positions can be derived, which lead to even better phase angles, and so forth. In every such cycle, adjustments to the atomic model are made: atom types are changed, missing atoms are introduced, etc.

This whole process is called structure refinement.

Structure Refinement: Three Main Steps

Close examination of the F_o-F_c **map** helps to introduce new atoms and remove “bad” old ones.

Once all non-hydrogen atoms are found, the atoms can be refined anisotropically.

Once the model is anisotropic, the hydrogen atom positions can be determined or calculated.

Structure Refinement: Electron Density Maps

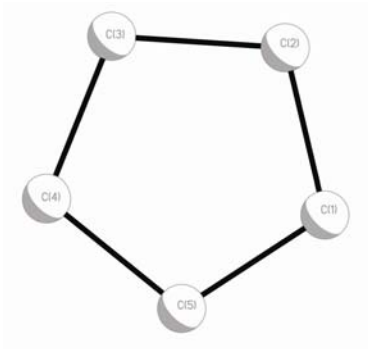
Different electron density maps:

F_o map: Electron density calculated using observed structure factors combined with phases calculated from the atomic model. This map shows the observed electron density; its accuracy depends largely on the accuracy of the phases.

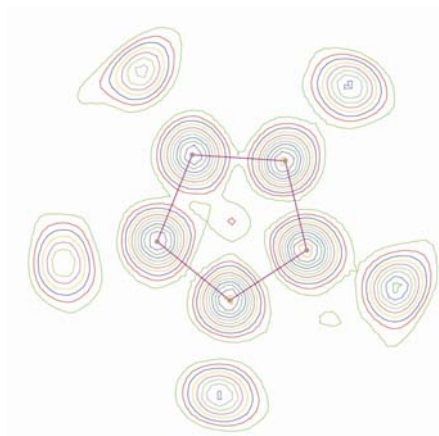
F_c map: Electron density calculated from structure factors and phases calculated from the atomic model. This map shows the electron density according to the model only.

$F_o - F_c$ map: The difference between the two maps. Has close-to-zero values for parts of the model that are consistent with the experimental density, large positive values at places where the model should have an atom, but does not, and large negative values at places where the model has an atom, which should not be there. Weaker positive or negative values for the $F_o - F_c$ map could point to wrongly assigned atom types.

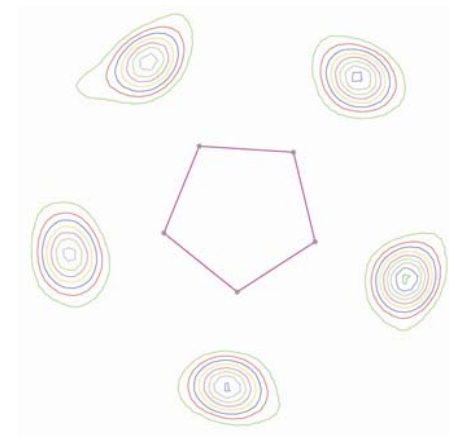
Structure Refinement: Electron Density Maps



Model: A Cp ring



F_o map

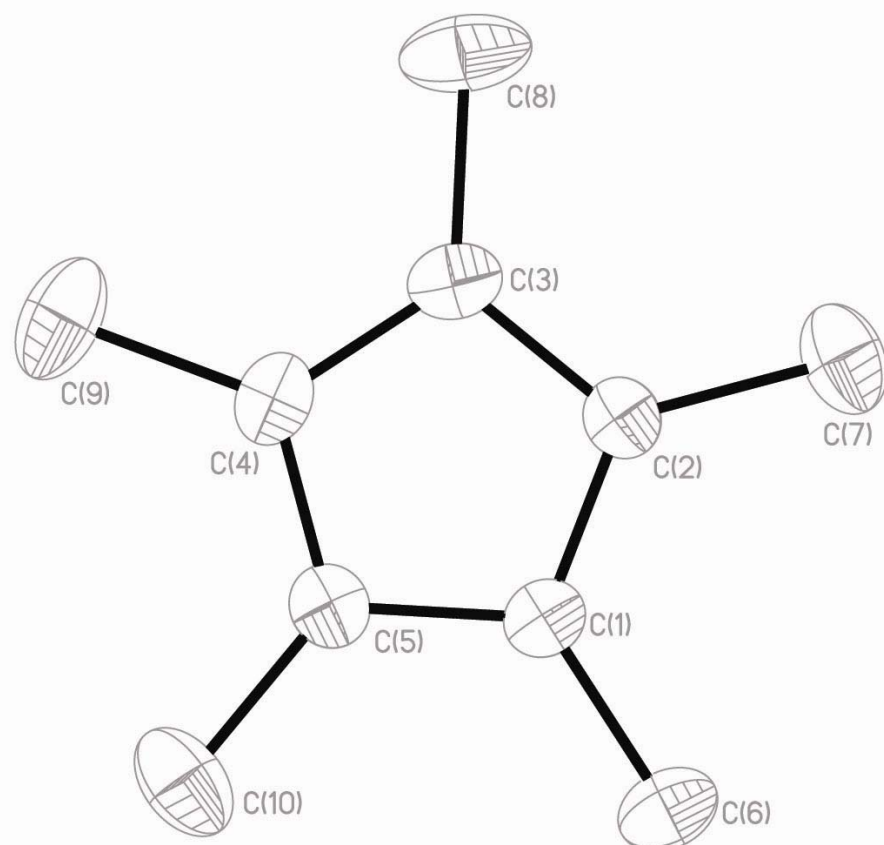
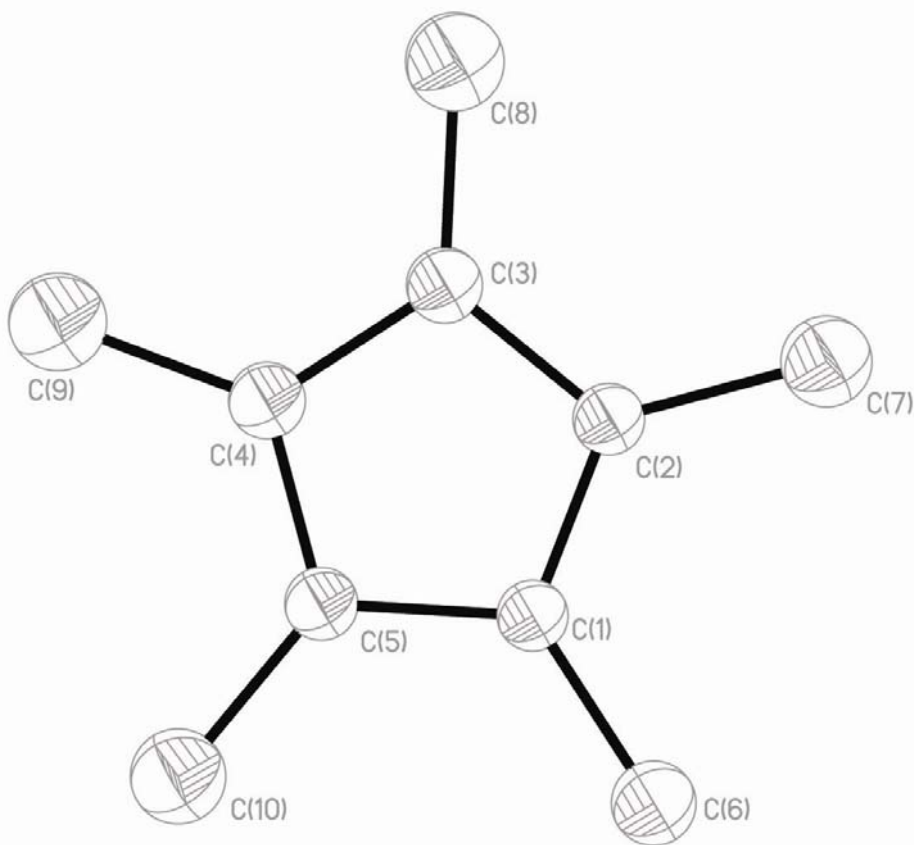


$F_o - F_c$ map

Examination of the electron density maps shows that ligand in the current model is not a Cp ring, but in fact a Cp* ring.

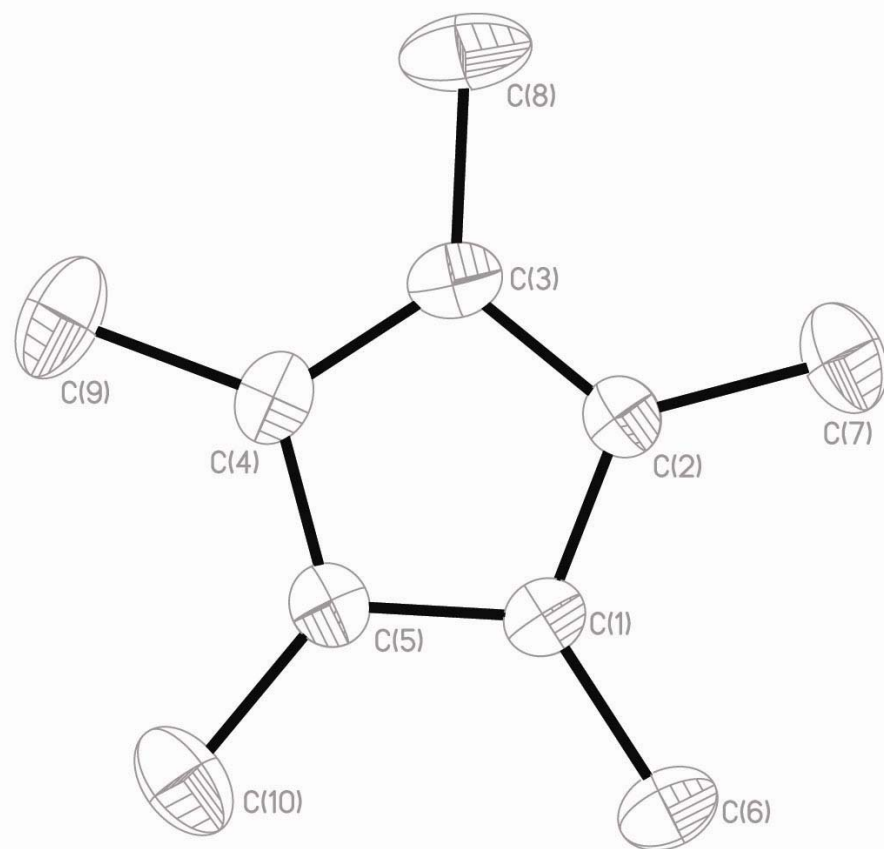
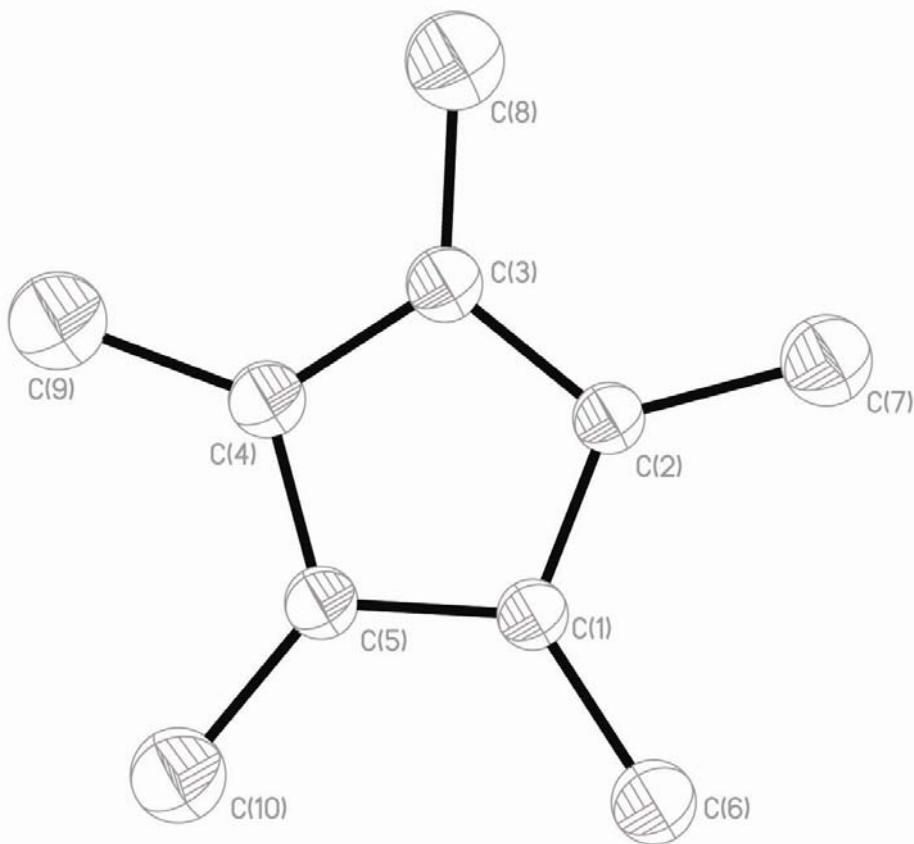
Structure Refinement: Anisotropic Displacement Parameters

It is reasonable to assume that atoms move anisotropically, i.e. with different amplitudes in different directions. Instead of describing an atom as a sphere, it is described as an ellipsoid; the elements of the matrix describing the ellipsoid are called anisotropic displacement parameters (ADP):



Structure Refinement: Anisotropic Displacement Parameters

The size of the ellipsoids (or spheres in the isotropic case) is chosen so that the ellipsoid includes 50% electron density of the atom type in the model. In the example below each ellipsoid or sphere includes 3 electrons. Thus, abnormally large or small ellipsoids can point out incorrectly assigned atom types or other problems with the model.



Evaluation of the Model

The model should only be altered if a change improves its quality.
How to judge quality of the model?

Least-squares approach:

By means of Fourier transformation, a complete set of structure factors is calculated from the atomic model. The calculated intensities are then compared with the measured intensities, and the best model is that, which gives the smallest value for the minimization function M .

$$M = \sum w(F_o^2 - F_c^2)^2$$

F : structure factor; o : observed; c : calculated; w weighting factor (derived from σ).

Residual Values: the R factors

$wR2$: Most closely related to refinement against F^2 .

$$wR = \left[\frac{\sum w(F_o^2 - F_c^2)^2}{\sum wF_o^2} \right]^{1/2}$$

$R1$: Most popular one, based on F .

$$R = \frac{\sum \left| |F_o| - |F_c| \right|}{\sum |F_o|}$$

GooF: S is supposed to be > 1.0

$$S = \left[\frac{\sum w(F_o^2 - F_c^2)^2}{(N_R - N_P)} \right]^{1/2}$$

F : structure factor; o : observed; c : calculated; w weighting factor (derived from σ).
 N_R : number of independent reflections; N_P : number of refined parameters.

Parameters

For every atom: x , y , z coordinates and one (isotropic) or six (anisotropic) displacement parameters.

For every structure: overall scale factor osf .

Possibly additional scale factors, maybe a *Flack-x*-parameter.

Atom types are also parameters, even though they are not refined. Incorrectly assigned atom types can cause quite some trouble.

Altogether: The number of parameters is roughly ten times the number of independent atoms in a structure.

The IUCr says: For a stable refinement: the data-to-parameter-ratio should be > 8 for non-centrosymmetric structures and > 10 for centrosymmetric structures. \rightarrow ca. 0.81 \AA or $2\theta = 50^\circ$ (Mo).

Constraints and restraints improve the data-to-parameter-ratio: constraints remove parameters, restraints add to data.

Constraints

Constraints are mathematical equations, rigidly relating two or more parameters or assigning fixed numerical values to certain parameters, hence reducing the number of independent parameters to be refined.

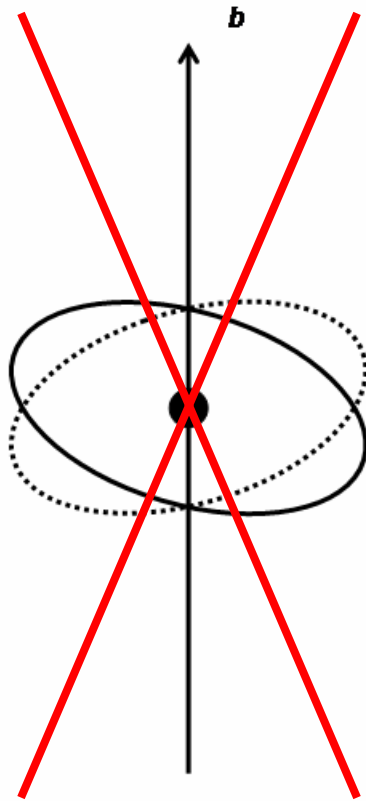
Site occupation factors are constraints present in every structure. Even for disordered atoms the sum of the occupancies is constrained to add up to 1.0.

Hydrogen atoms can be “placed on mathematically calculated positions and refined using a riding model”.

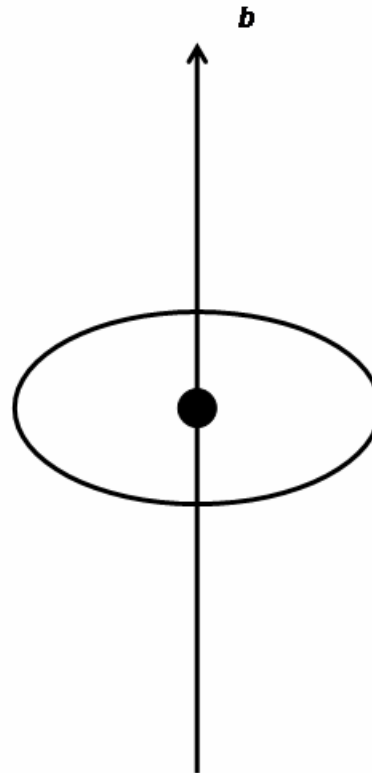
That means *X-H distances* and *H-X-H* or *H-X-Y angles* are constraint to certain values. Not the hydrogen *positions*!

Atoms on special position require constraints for their coordinates, occupancies and sometimes also their ADPs:

Special Position Constraints



No good



Good

An atom on a twofold axis along b .
A 180° rotation must not change the position of the atom or the shape of the thermal ellipsoid.

From the first condition follows:
 $(x, y, z) = (-x, y, -z)$, which is only true for $x = z = 0$.

The second condition dictates:
 $(U^{11}, U^{22}, U^{33}, U^{23}, U^{13}, U^{12}) =$
 $(U^{11}, U^{22}, U^{33}, -U^{23}, U^{13}, -U^{12})$, which is only true for $U^{23} = U^{12} = 0$.

The most popular refinement program generates special position constraints automatically. Big relief!

Restraints

Restraints are assumptions used to introduce chemical or physical information into a refinement as additional experimental observations. Restraints are treated as data (with a standard uncertainty).

Minimization Function including restraints:

$$M = \sum w(F_o^2 - F_c^2)^2 + \sum 1/\sigma^2 (R_t - R_o)^2$$

F : structure factor; o : observed; c : calculated; w weighting factor; σ : standard uncertainty assigned to the restraint; R_t : target value for restraint quantity; R_o : actual value.

Restraints should be used with great care and only if justified. When appropriate, however, they should be used without hesitation, and having more restraints than parameters in a refinement is nothing to be ashamed of.

Geometrical Restraints

One can restrain a group of atoms to lie on a common plane (e.g. a phenyl ring) or an interatomic distance or a bond angle to a certain target value (direct distance restraints).

Alternatively, one can restrain equivalent distances to be equal (relative distance restraints, aka “similarity restraints”).

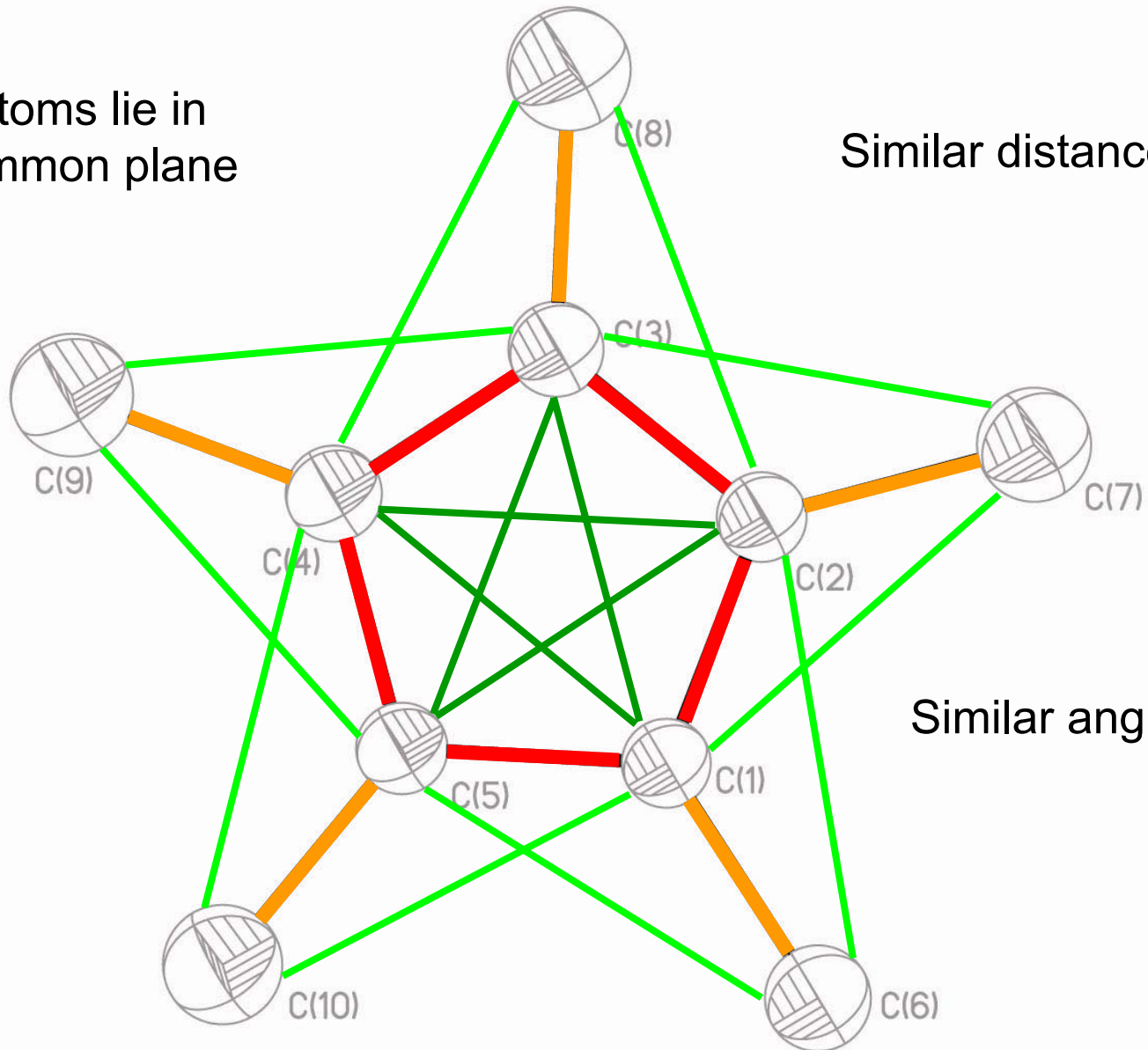
Advantage of relative distance restraints: No need for “outside” information, refinement converges well (*esp.* for $Z' > 1$).

Disadvantage: underestimated standard uncertainties of bond lengths and angles. And it is too easy to refine a structure in a space group with too low symmetry.

Geometrical Restraints

All 10 atoms lie in
one common plane

Similar distances



Similar angles

ADP Restraints

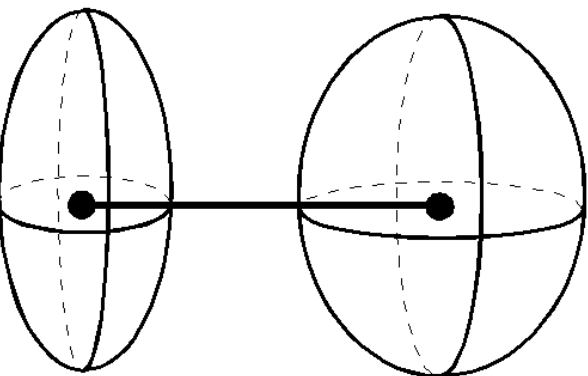
It is fair to assume that atoms bound to one another move similarly, both in direction and amount.

One can restrain the anisotropic displacement parameters of two atoms *in the direction of the bond* between them to be equal within a given standard uncertainty (e.g. 0.01). This is called a “rigid bond restraint”.

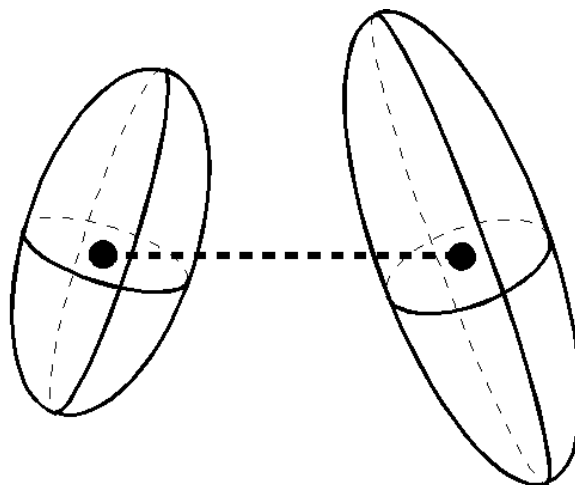
It can be assumed that atoms that are bound to one another move in similar directions with approximately similar amplitudes. Therefore, one can restrain atoms close to one another to have the same U^{ij} components within a given standard uncertainty (e.g. 0.04). This is called a “similar-ADP restraint”.

The second assumption is much bolder than the first one, hence the much larger standard uncertainty.

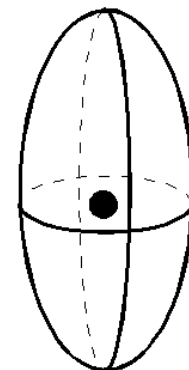
ADP Restraints



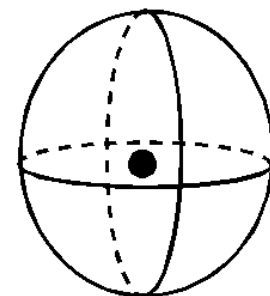
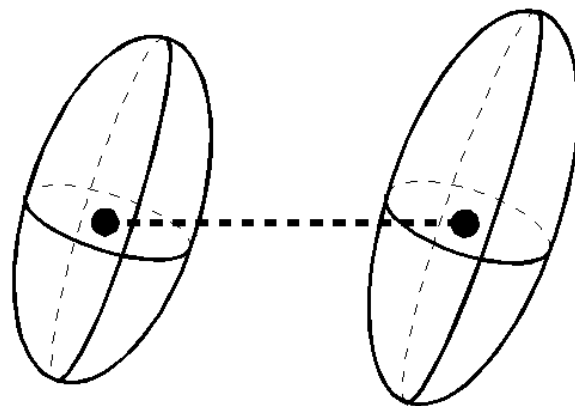
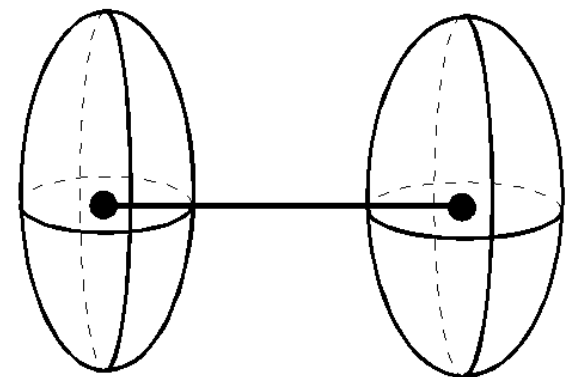
rigid bond



similar-ADP

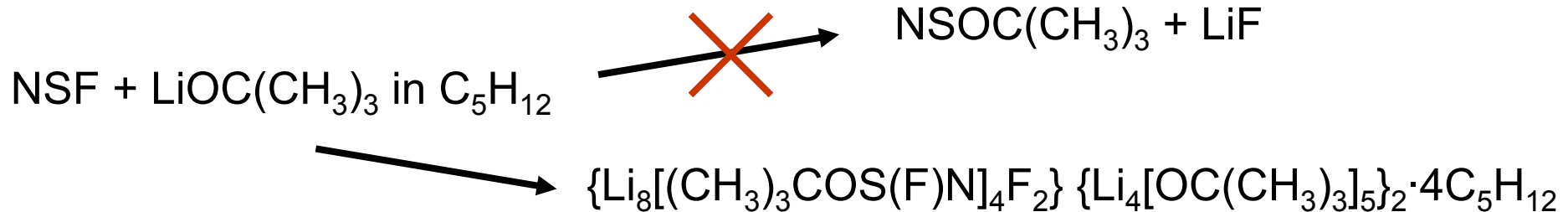


isotropic behavior



Problems and Pitfalls

The big advantages of X-ray structure analysis are that it is very objective and can determine the structure of even very complex molecules. E.g.



Disadvantages are that you do need crystals and that the accuracy of bond lengths and angles is relatively low (ca. 0.02 to 0.001 Å). With some spectroscopic methods (e.g. rotational spectroscopy), interatomic distances can be determined 100 times more accurately.

Another disadvantage is that the structure determined is the **spatial average** over the entire crystal and the **temporal average** over the data collection time (hours). This makes it impossible to study dynamic processes. However, those processes can leave their traces behind...

Problems and Pitfalls: Wrong Space Group

At least 1% of the crystal structures are published in the wrong space group. Frequently, symmetry elements (esp. inversion centers) have been overlooked.

When an inversion center is overlooked, the consequences can be very serious. The refinement is mathematically unstable (saddle point instead of minimum). Bond lengths that should be equal fly apart (e.g. $1.52 \text{ \AA} \rightarrow 1.38 + 1.66 \text{ \AA}$) and the same happens to the displacement parameters of atoms related by the center.

It is not so bad when other symmetry elements are overlooked. The refinement is stable but too many parameters are refined against too many data. The structure is not distorted, but is less precise than if the data had been averaged in the correct Laue group and fewer parameters used to describe it.

Common cases of wrong space groups

$P1 \rightarrow P\bar{1} \rightarrow C2/m \rightarrow R\bar{3}m$

$C2 \rightarrow C2/c \rightarrow R\bar{3}c$

$C2 \rightarrow Fdd2$

$C2 \rightarrow R32$

$C2 \rightarrow I222$

$Cc \rightarrow Fdd2$

Figure courtesy of George M. Sheldrick.

Often an inversion center or a centered lattice with higher symmetry are overlooked. A long series of papers by Marsh and others document wrong space groups reported in the literature.

A particularly revealing example was a $P1$ space group corrected by Marsh (1999) to $C2$; this correction was then corrected to $Fdd2$ by Spek in 2000 (*Acta Cryst.* B56 744).

CheckCIF, used to check all small molecule structures submitted to *Acta Cryst.*, would have caught this example.

Some innocuous-looking cells

$$\begin{array}{l} \bar{P}1 ? \\ Z = 2 \end{array} \quad \begin{array}{l} a = 4.982, b = 12.133, c = 12.871 \text{ \AA}, \\ \alpha = 67.67, \beta = 78.84, \gamma = 78.14^\circ \end{array}$$

$$\rightarrow \begin{array}{l} C2/c \\ Z = 4 \end{array} \quad \begin{array}{l} a = 23.748, b = 4.982, c = 13.936 \text{ \AA}, \\ \alpha = 90, \beta = 122.08, \gamma = 90^\circ \end{array}$$

$$\begin{array}{l} C2/c ? \\ Z = 4 \end{array} \quad \begin{array}{l} a = 11.997, b = 6.928, c = 13.574 \text{ \AA}, \\ \alpha = 90, \beta = 90, \gamma = 90^\circ \end{array}$$

$$\rightarrow \begin{array}{l} \bar{P}3c1 \\ Z = 4 \end{array} \quad \begin{array}{l} a = 6.928, b = 6.928, c = 13.574 \text{ \AA}, \\ \alpha = 90, \beta = 90, \gamma = 120^\circ \end{array}$$

$$\begin{array}{l} C?? \\ Z = 12 \end{array} \quad \begin{array}{l} a = 16.207, b = 26.937, c = 6.823 \text{ \AA}, \\ \alpha = 90, \beta = 106.32, \gamma = 90^\circ \end{array}$$

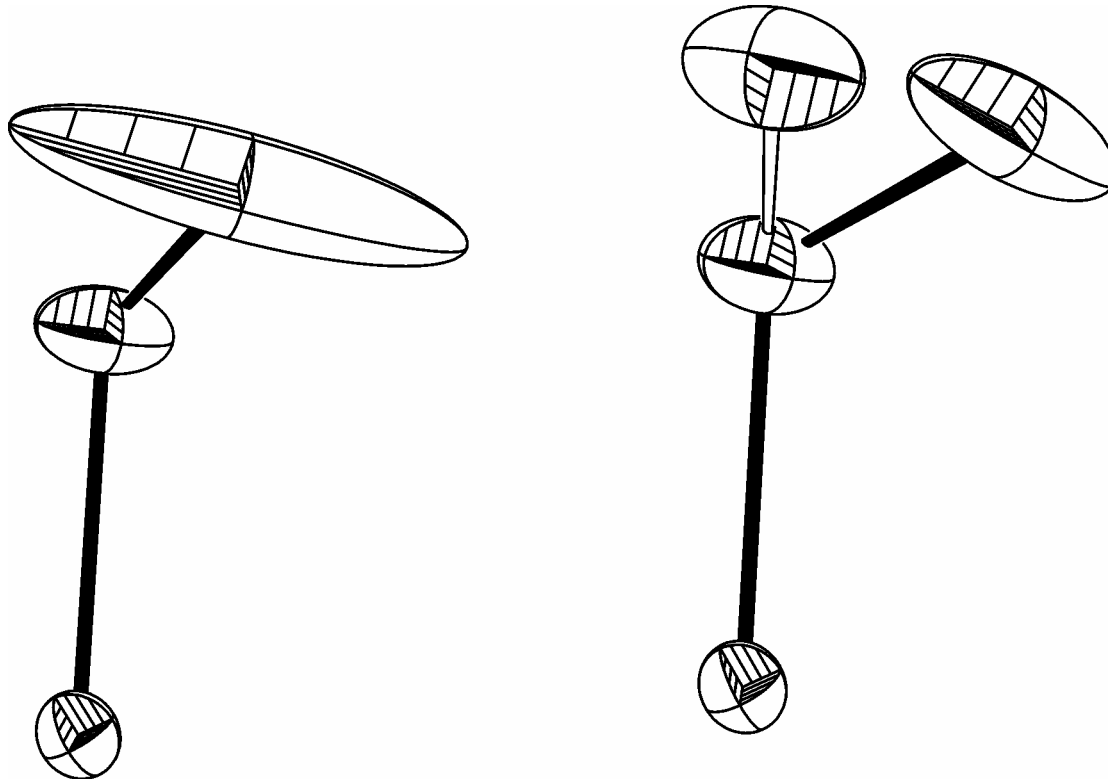
$$\rightarrow \begin{array}{l} \bar{R}3 \\ Z = 6 \end{array} \quad \begin{array}{l} a = 15.718, b = 15.718, c = 15.718 \text{ \AA}, \\ \alpha = 117.94, \beta = 117.94, \gamma = 117.94^\circ \end{array}$$

Figure courtesy of George M. Sheldrick.

Displacement Parameters

As mentioned earlier, after anisotropic refinement, atoms are represented by 50% probability ellipsoids. Abnormally large or small ellipsoids, as well as ellipsoids with strange shapes, tell us something about problems with the refinement.

Two or more positions of a disordered atom can be represented by a large displacement ellipsoid. *E.g.* A disordered ethyl group:



Displacement Parameters

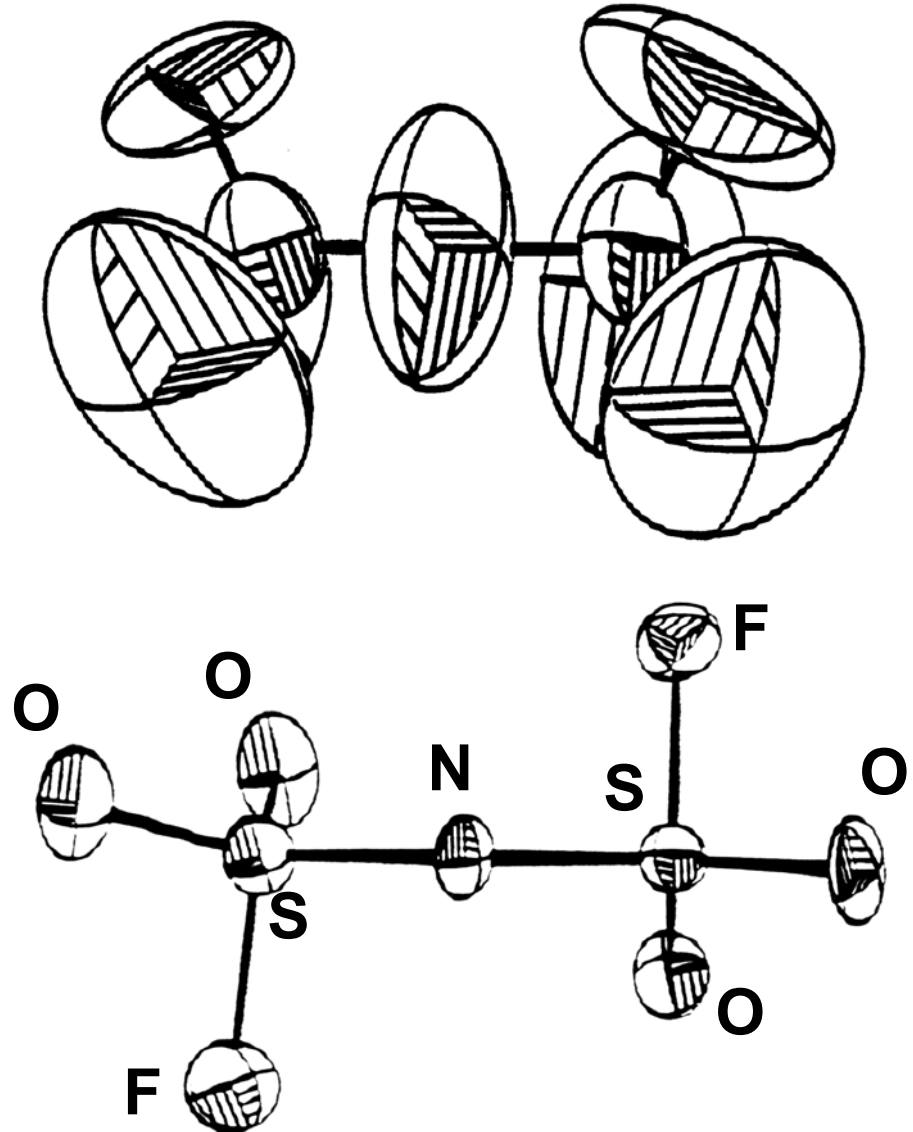
Thermal motion of atoms in the crystal can also lead to large ADPs.

The very large ellipsoids in the anion $\text{N}(\text{SO}_2\text{F})_2^-$ at room temperature can be explained either with strong atomic movement or disorder.

In this case, the normal staggered conformation at 112K and the ecliptic conformation at 290K make this look more like a disorder. This is corroborated by the bond lengths:

290K: S-O and S-F all similar between 1.24 and 1.36 Å

112K: S-F 1.57, S=O 1.42 Å



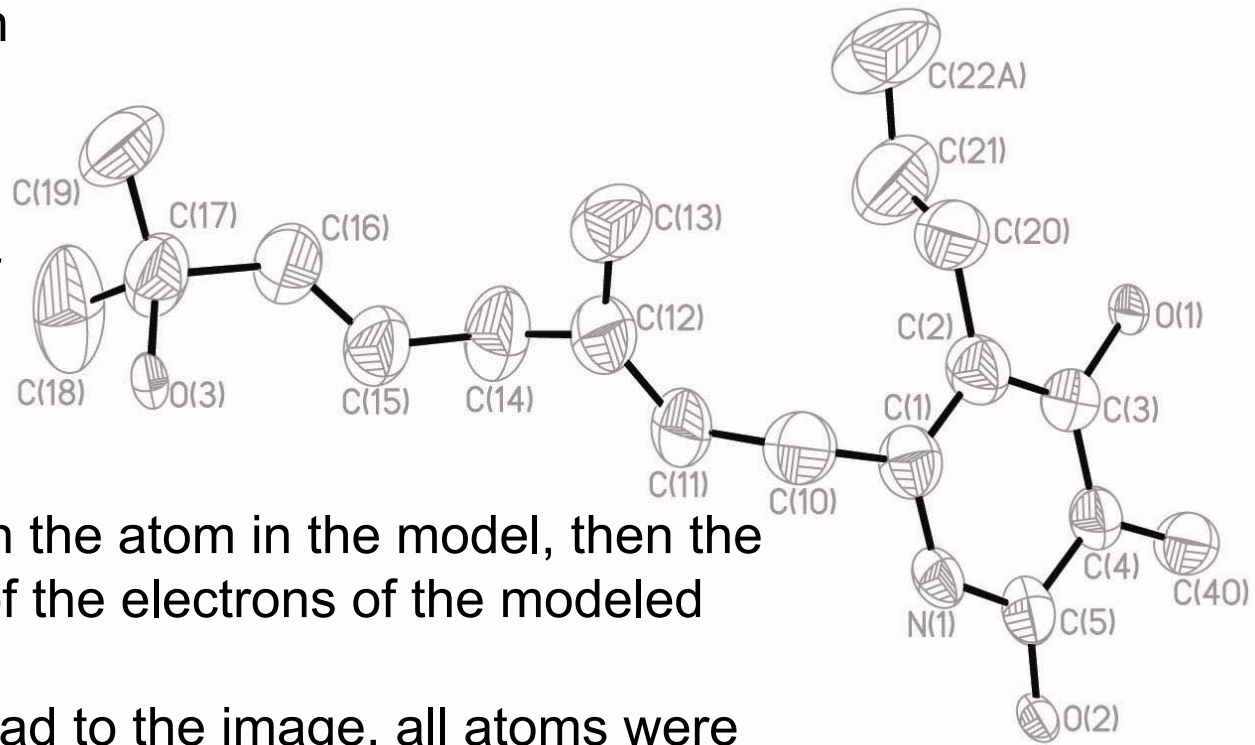
More on Displacement Parameters

The ADPs can help with the assignment of atom types:

Each thermal ellipsoid is drawn around 50% of the electrons of the element in the model.

If the atom in the real molecule is heavier than the atom in the model, then the volume including 50% of the electrons of the modeled element is smaller.

In the refinement that lead to the image, all atoms were refined as carbon, even though there are three oxygen atoms and one nitrogen atom in the molecule. As carbon has 6 electrons, all ellipsoids are calculated to contain 3 electrons from the calculated electron density. Nitrogen and oxygen have more electrons than carbon, making the electron density higher at oxygen and nitrogen sites than at places where we have a carbon atom. That leads to a smaller volume of the corresponding ellipsoid, as 3 electrons are only 43% of a nitrogen and only 38% of an oxygen.



Libration

The higher the temperature, the shorter bond lengths appear to be, even though the unit cell gets slightly larger at the same time. The reason for this is libration. Especially terminal atoms show approximately circular motion. In anisotropic refinement, this motion is fitted as an ellipsoid, the center of which lies inside the circle of motion, while the atom is located on its periphery. This effect makes the bond distance appear shorter.

Depending on the movement of the atoms, the correction that needs to be applied is somewhere between 0.001 – 0.1 Å. Libration is much stronger at higher temperatures and particularly pronounced for hydrogen atoms, small ions, like NO_3^- , BF_4^- , ClO_4^- , PF_6^- and for $-\text{CF}_3$ groups.

$$\Delta r \approx \frac{\Delta U}{2r} = \frac{[U_B - U_A]}{2r}$$

For most low-temperature structures, the effects of libration are smaller than the standard uncertainties of the bond lengths and it is not necessary to correct for it.

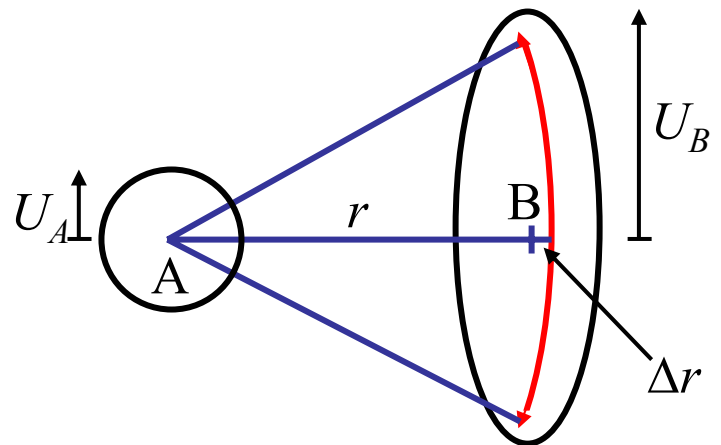


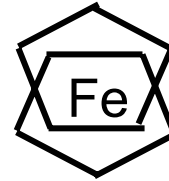
Figure courtesy of George M. Sheldrick.

Ferrocene: Ecliptic or Staggered?



D_{5h}

or



D_{5d}

Figure courtesy of George M. Sheldrick.

The first X-ray structure analysis (in 1956 at room temperature) showed $Z = 2$ in $P2_1/c$. Thus, the molecule must sit on an inversion center. Therefore it also “must” be staggered (D_{5d}). Electron scattering in the gas phase (1968) showed approximate D_{5h} geometry.

The high-temperature X-ray data can be understood as disorder of a D_{5h} structure, as the ADPs of the carbon atoms remain large even at somewhat lower temperature (173 K, 1979).

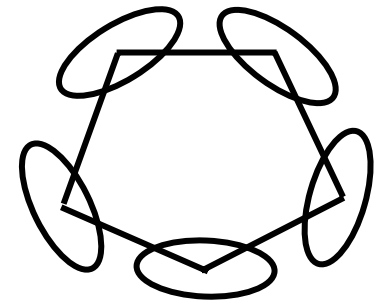


Figure courtesy of George M. Sheldrick.

Two more low-temperature structures have been determined later: A triclinic phase at 164K was found to be almost D_{5h} (deviation $\Delta\phi$ ca. 9°) and an orthorhombic phase at 98 K was exactly D_{5h} within the experimental error.

Hydrogen Atoms

It is electron density that we measure by X-ray diffraction. The heavier an atom is and the more electrons it has, the stronger is its effect on the diffraction pattern. This also means that, especially in the presence of heavy atoms, light atoms are somewhat more difficult to localize. The lightest atom of all is hydrogen: it has only one electron, localized away from the nucleus. Therefore, hydrogen atoms are notoriously difficult to detect with X-ray diffraction methods.

Relatively high electron density between the atoms and libration effects make X—H bonds appear too short.

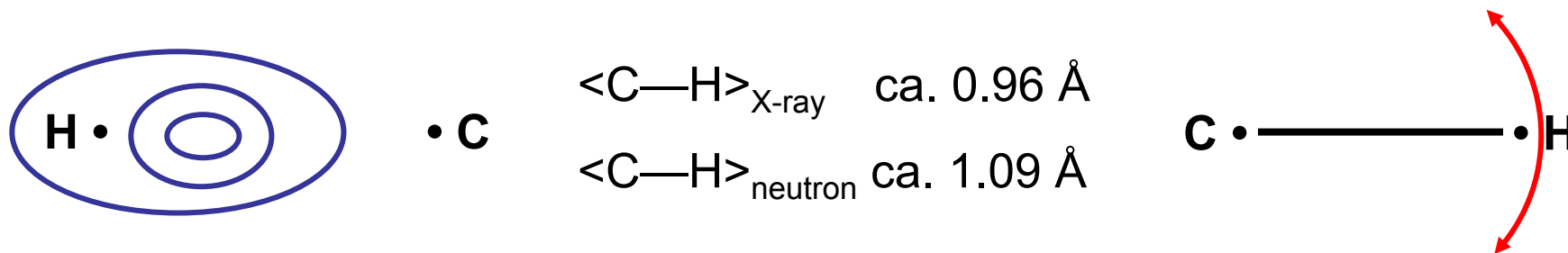


Figure courtesy of George M. Sheldrick.

Hydrogen Atoms

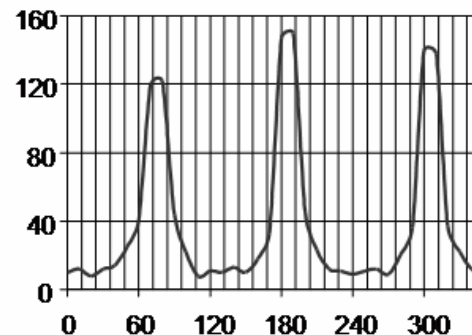
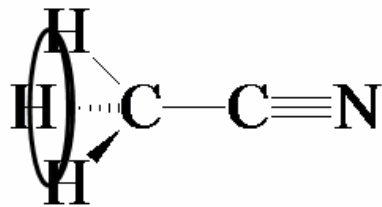
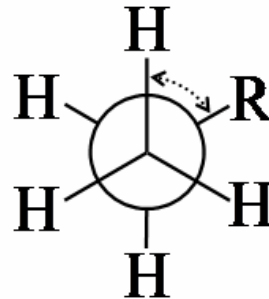
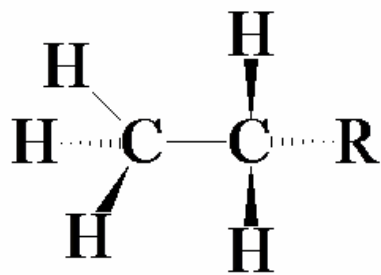
Especially for hydrogen atoms bound to carbon it is frequently possible to calculate the hydrogen positions from the coordinates of the atoms the hydrogen atoms are attached to, as the standard bond lengths and angles are well known.

Hydrogen atoms of water molecules must be detected in the experimental electron density or else they cannot be included into the model.

Even more difficult to detect can be hydrogen atoms in heavy metal hydrides. The sometimes relatively strong Fourier truncation ripples close to heavy atom positions can overpower the rather weak electron density maxima representing the hydrogen atoms. Very accurate and especially complete high quality data and proper scaling are required to distinguish those hydrogen atoms from the background noise.

Hydrogen Atoms

In most cases, the positioning of hydrogen atoms bound to carbon in an atomic model during the refinement of an X-ray crystal structure is done entirely without any or only very little direct information from the diffraction experiment. Hydrogen atoms on aromatic carbons, CH, CH₂ groups, most CH₃ groups are easy. CH₃ in acetonitrile, Cp* or toluene: torsion angle needs to be determined.



MIT OpenCourseWare
<http://ocw.mit.edu>

5.069 Crystal Structure Analysis
Spring 2008

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.