

## Weights

Weights for reflections may have different roles at different stages in the procedure.

Weights to try to increase the radius of convergence, or to emphasise some feature of the data. (Dunitz Seiler)

Weights to try to reduce the influence of outliers. Sometimes the weight is a function of the residual, so that the refinement is not strictly least-squares. (Robust-Resistant refinement)

Weights to try to reduce bias in the parameter values, and obtain most realistic values for their standard uncertainties. (Quasi-statistical)

1

## Optimal Weights

It is generally accepted that for a correctly weighted valid refinement,  $w\Delta^2 \approx \text{constant}$  as a function of any systematic ranking of the data, e.g. Fo, Fc, index, theta.  
This is rarely the case for purely statistical weights:  $w = 1/\sigma^2 F$

2

## Optimal Weights

Weights are needed which reflect unidentified errors in the data, and unparameterised features of the model

Fudge-factors are applied to statistical weighting, or empirical weights are estimated to make  $\langle w\Delta^2 \rangle$  constant for chosen binnings.

It is usually arranged to make  $S^2$  about unity at the same time.

3

## Goodness of Fit, Weights & Standard Deviations

The SHELX weighting scheme attempts to get a uniform weighted residual by modifying the standard deviations of the observations

The CRYSTALS weighting scheme attempts to model the variation in residual by a low-order polynomial

4

## Dunitz Seiler Weights

Dunitz Seiler, available in SHELXL, up-weight the high angle data.

This discourages the LSQ to fit the low angle data, leading to enhanced Fo-Fc.

Hydrogen atoms contribute most strongly to the low order data, and are thus more visible.

5

## Robust-Resistant Weights

If  $\Delta F$  is large compared with  $\sigma F$ , there is the possibility that the reflection is an outlier, and should be down-weighted.

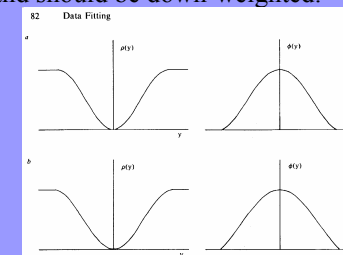


Fig. 6.2. Two robust/resistant fitting functions. (a) Tukey's function:  $\rho(y) = (y^2/2) \cdot [1 - (y/a)^2 + (1/3)(y/a)^3]$ ;  $\phi(y) = [1 - (y/a)^2]^3$ . (b) Andrew's function:  $\rho(y) = 1 - \cos(y/a)$ ;  $\phi(y) = \sin(y/a)/(y/a)$ .

6

## Goodness of Fit, Weights & Standard Deviations

(or, why  $S$  is worthless)

The Goodness of Fit is defined by  $S^2 = \frac{\sum w\Delta^2}{n-m}$

If we let  $w' = w/S^2$  then  $\frac{\sum w'\Delta^2}{n-m} = 1.0$

We can always get a goodness of fit of unity by simply rescaling all the weights

A more useful definition of  $S$  is:  $S^2 = \frac{\sum \Delta^2 / \sigma^2}{n-m}$

7

## Goodness of Fit, Weights & Standard Deviations

The standard uncertainty in a parameter is obtained from the diagonal elements of the LSQ inverse normal matrix

$$\sigma^2 x = (b_{xx})^{-1} \cdot \sum w\Delta^2 / (n-m)$$

For a purely diagonal matrix

$$b_{xx} = \sum w \frac{\partial^2 F}{\partial x^2} \quad \text{and} \quad (b_{xx})^{-1} = 1/b_{xx}$$

$$\sigma^2 x = \frac{\sum w\Delta^2}{n-m} \cdot \frac{\partial^2 F}{\partial x^2}$$

replacing  $w$  by  $w' = w/S^2$  does not affect the variance of  $x$

8

## Weighting Restraints

Weights given to restraints may also have different roles at different times.

At the start of a refinement, the weights may be substantial to try to ensure that the structure falls into the global minimum.

At the end, weights may be set to zero in order to obtain parameter sus based on the X-ray data only.

Some restraints may need to be retained if the data does not define the corresponding parameters well.

9

## Size and time

For Poisson statistics,  $\sigma^2(I) = I$

$I/\sigma(I)$ , signal-to-noise =  $\sqrt{I}$

The average signal:noise improves if the intensities are stronger.

To get good data, get good intensities

Measured intensity  $\propto$  Volume \* time

10

## What is the best crystal size for the collection of X-ray data?

C.H. Gorbitz, Acta Cryst (1999), B55, 1090-1098

Glycine-L-Serine, C<sub>5</sub>H<sub>10</sub>N<sub>2</sub>O<sub>4</sub>

Comparison of crystal as grown, and after grinding to sphere

Original size: 2.20 x 2.0 x 0.80 mm, volume 3.52 mm<sup>3</sup>

Sphere size: 0.30 mm diameter, volume 0.14 mm<sup>3</sup>

$\mu = 0.134 \text{ mm}^{-1}$

11

## Crystal Size

‘The results from (the uncut crystal) are clearly superior to those from the small spherical crystal *with the same exposure time*, and are almost indistinguishable from those on the small crystal even with a ten-fold increase in exposure time’

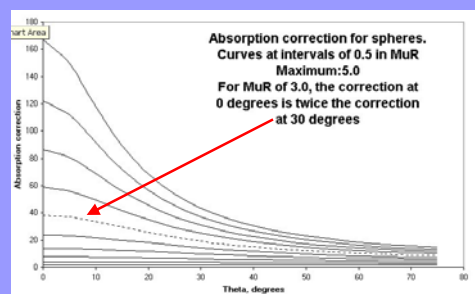
‘It thus appears that *SADABS* does a very good job’

(one can substitute SADABS by most multi-scan methods)

12

## Spherical Absorption Correction

Multiscan methods cannot correct for any underlying theta dependent errors. The correction has been tabulated for spheres and cylinders

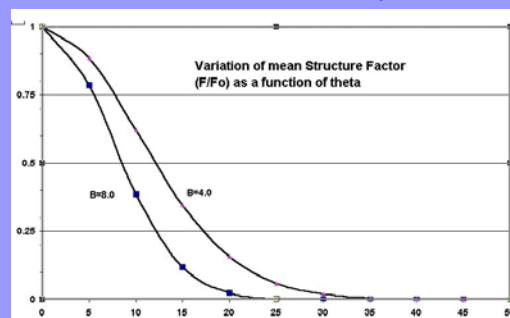


13

## F vs Bragg

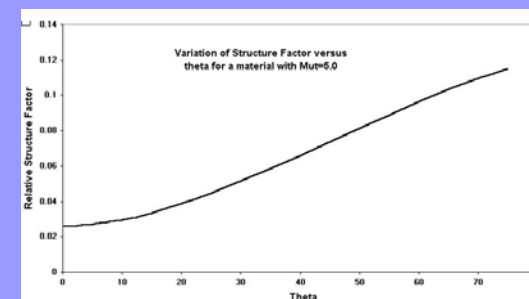
Effect of the isotropic adp on the computed Structure Factor.

$$T = e^{-2B(\sin \theta / \lambda)^2}$$



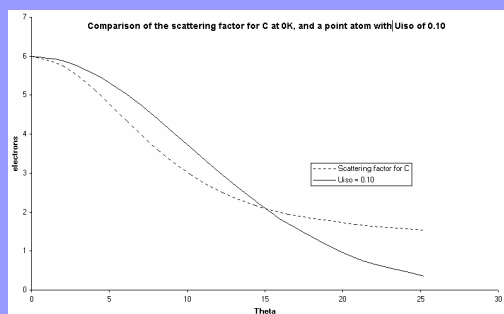
## Effect of absorption on relative Intensity

The effect of uncorrected absorption is to increase the relative intensity of high angle reflections



## Occupancy and adps

These two functions are plotted for a carbon atom at rest and a point atom with  $U_{iso}=0.01$ . Apart from a scale factor (arbitrarily set at 5 for this illustration) it is evident that the two curves have a similar form, particularly at low theta. To determine fractional occupancies, you **must** have high angle data.



## Flack's Parameter

Imagine a non-centrosymmetric crystal that is composed of two discrete parts one of which is related to the other by inversion. Such a crystal is called an inversion twin.

If the relationship is exact, then the diffraction patterns from each half will exactly overlap.

If there are no atoms which have strong anomalous scattering, then both diffraction patterns will be identical.

If there are anomalous scatterers, there will be small differences.

17

## Flack's Parameter

The observed intensity of a diffracted beam will be the sum of the intensities from each part independently.

$$F_{oh}^2 = (1-x) \cdot F_{+h}^2 + x \cdot F_{-h}^2$$

Where the first term is computed from the given model, and the second from its inverse.

18

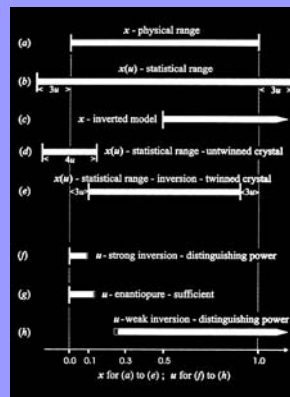
## Flack's Parameter

If  $x$  is zero, the real structure has the same hand as the model.

If  $x$  is unity, the real structure is the inverse of the model.

**However**, remember that  $x$  has a standard uncertainty. Its significance can only be judged with respect to its uncertainty.

19



## Flack's Parameter

This complicated illustration from Flack's paper explains how to interpret the standard uncertainty ( $u$ ).

It shows that before trying to explain the value of  $x$ , you must think about  $u$ .

20

## Flack's Parameter The bottom lines

If the  $su$  is large, we can deduce nothing about the chirality.

For a material suspected of being enantio-pure, a standard uncertainty of no more than 0.04 is needed to before we can interpret  $x$ .

For a material **known** to be enantio-pure, an  $su$  of 0.1 is adequate.

A value of  $x$  sufficiently far from 0 or 1, and **with a small  $su$** , can be interpreted as an inversion twin.

21

## Standard Uncertainties

As was shown earlier, an advantage of full matrix refinement is that it enables 'proper' standard uncertainties (standard deviations) to be computed.

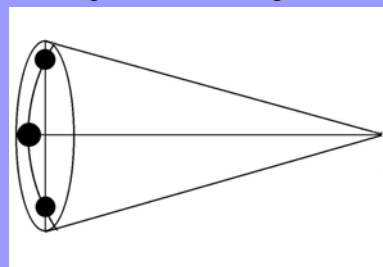
Fiddling with weighting schemes can affect the 'properness', as can being selective about what reflections are used.

However, other physical effects may cause systematic errors in the actual parameter values.

25

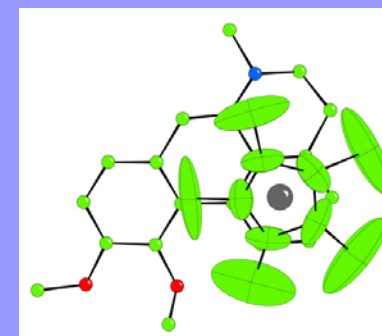
## TLS Calculations

If part of a molecule undergoes rigid-body libration, this leads to an apparent shortening of the bond lengths.



23

## CP\* Compound



24

## Analysis of rigid body motion

Principal axes of the thermal ellipsoids, Å<sup>2</sup>

C 1.	0.0393	0.0521	0.1858	
C 2.	0.0314	0.0616	0.2063	
C 3.	0.0233	0.0700	0.2641	
C 4.	0.0266	0.0801	0.2555	
C 5.	0.0420	0.0842	0.1724	
C 11.	0.0716	0.0949	0.3856	Might be split
C 12.	0.0488	0.0962	0.4511	Might be split
C 13.	0.1059	0.1109	0.6858	Might be split
C 14.	0.0638	0.1430	0.8415	Might be split
C 15.	0.0640	0.1104	0.8912	Might be split

25

## Analysis of rigid body motion

Centre of gravity, centre of libration -0.1809 -0.1139 -0.2056  
Centre for which  $s$  is symmetric : -0.0916 -0.0972 -0.1615

L			T			S		
14.35	0.00	0.00	0.05	0.00	-0.01	0.22	0.00	0.00
0.00	40.14	0.00	0.00	0.09	0.01	0.00	-0.19	0.00
0.00	0.00	249.57	-0.01	0.01	0.03	0.00	0.00	-0.03

The square root of 249° is 15°

This oscillation causes the bond lengths to appear to be shortened.

A correction can be computed

26

## Analysis of rigid body motion

Uncorrected Corrected Difference

Tangential bonds

C5	C4	1.416(5)	1.474	0.058
C5	C1	1.432(5)	1.492	0.060
C4	C3	1.416(5)	1.478	0.062
C3	C2	1.414(5)	1.470	0.056
C2	C1	1.411(5)	1.473	0.062

Radial bonds

C15	C5	1.636(5)	1.705	0.069
C14	C4	1.616(5)	1.680	0.064
C13	C3	1.611(5)	1.682	0.071
C12	C2	1.610(5)	1.676	0.066
C11	C1	1.608(5)	1.672	0.064

27

## Analysis of rigid body motion

In this case, the bond length correction is about ten times the least square estimate of the standard uncertainty.

The standard uncertainties produced by the refinement program need to be interpreted with care.

28

## Standard Uncertainties

For most practical purposes the published *sus* are rather suspect.

If you have a problem that cannot be resolved using Carpenter's Rules, then make sure that you do a very careful experiment – on more than one crystal

29

## Molecular Geometry (Extended) Carpenters Rules

**Acta Cryst A35, (1979), 248-250**

Carpenter gave rules of thumb for estimating the *sus* of derived parameters. If  $\sigma(f)$  is the atomic fractional *su* of an average element, and  $A$  is the corresponding cell edge, then:

$$\begin{aligned}\sigma(x) &= A \cdot \sigma(f) && \text{positional } su \text{ in } A, \text{ often about } 0.002 \\ \sigma(d) &= 1.4 \cdot \sigma(x) && \text{bond } su \text{ in } A, \text{ often about } 0.003 \\ \sigma(a) &= 80 \cdot \sigma(d) && \text{bond angle } su \text{ in degrees, often about } 0.2 \\ \sigma(\tau) &= 1.4 \cdot \sigma(a) && \text{torsion angle } su \text{ in degrees, often about } 0.3\end{aligned}$$

30

## Wrong Structure

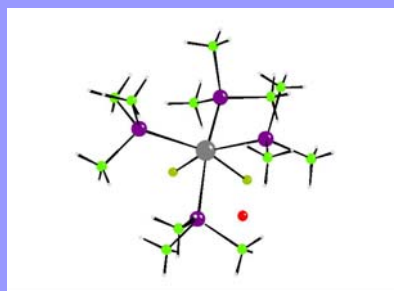
There are very few known instances where refinement of an organic or organo-metallic material reaches a low R value, but is substantially wrong. This security comes from the massive over-determination of the problem.

Post mortem on structures later found to be incorrect are instructive.

31

## Wrong Structure

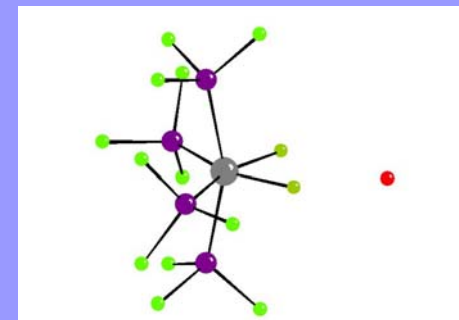
Ged Parkin has studied this series of compounds formed with different halides.



32

## Wrong Structure

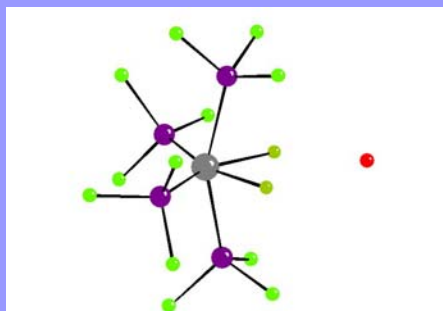
Look at this structure. Is it right or wrong?



33

## Wrong Structure

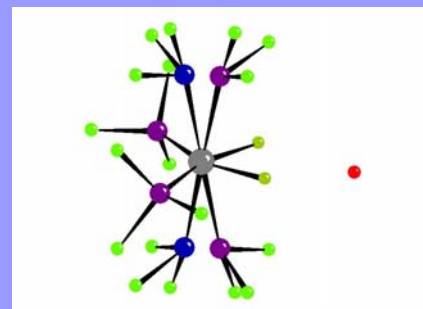
Look at this structure. Is it right or wrong?



34

## Wrong Structure

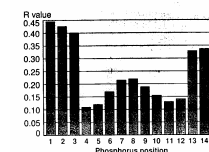
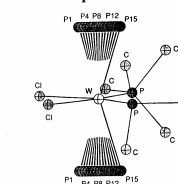
Both solutions



35

## Wrong Structure

Diffraction values for tungsten chloride complex exhibit two minima

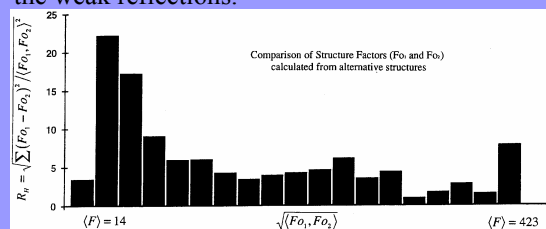


Although the R factors for both solutions are similar, there is no path for least squares between them.

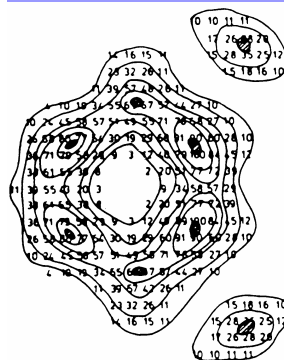
36

## Wrong Structure

It turns out that both solutions yield almost the same R factor. Plotting the  $F_{calc}$  of one solution against those for the other shows the main differences are in the weak reflections.



## Misinterpretation of Maps

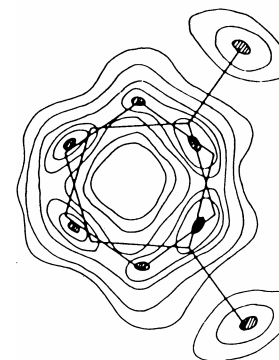


If the R factor fails to drop as low as expected, see if the initial map can be re-interpreted. Look at the map itself.

Initially, the two isolated peaks were regarded as spurious. The benzene ring failed to refine.

38

## Miss-interpretation of Maps



In fact, they, plus confidence in the preparative chemist, provided the key.

The disordered methyl cyclopentadiene refined perfectly.

39