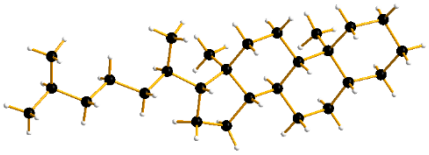
Determination of absolute configuration using X-ray diffraction

Determination of absolute configuration using X-ray diffraction

Simon Parsons

School of Chemistry and Centre for Science at Extreme Conditions, The University of Edinburgh, King's Buildings, West Mains Road, Edinburgh, EH9 3FJ, UK. S.Parsons@ed.ac.uk.

Methods for determination of absolute structure using X-ray crystallography are described, with an emphasis on applications for absolute configuration assignment of enantiopure light-atom organic compounds.



Key words: absolute structure, absolute configuration, chirality, X-ray crystallography

Abstract

Methods for determination of absolute structure using X-ray crystallography are described, with an emphasis on applications for absolute configuration assignment of enantiopure light-atom organic compounds. The ability to distinguish between alternative absolute structures by X-ray crystallography is the result of a physical phenomenon called resonant scattering, which introduces small deviations from the inherent inversion symmetry of single-crystal X-ray diffraction patterns. The magnitude of the effect depends on the elements present in the crystal and the wavelength of the X-rays used to collect the diffraction data, but it is always very weak for crystals of compounds containing no element heavier than oxygen. The precision of absolute structure determination by conventional least squares refinement appears to be unduly pessimistic for light-atom materials. Recent developments based on Bijvoet differences, quotients and Bayesian statistics enable better and more realistic precision to be obtained. The new methods are sensitive to statistical outliers, and techniques for identifying these are summarised.

1. Introduction

The space group, which defines the symmetry relationships that exist between the atoms composing a crystal, is one of the most fundamental features of a crystal structure. The translational symmetry that relates one unit cell to another is part of the space group symmetry. If two molecules within the unit cell are related by a symmetry operation such as a screw axis or reflection, those operations are also part of the space group.

The inversion centre is a very common symmetry operation in crystal structures. Those structures with space groups which possess an inversion centre are called *centrosymmetric*, while those which lack an inversion centre are *non-centrosymmetric*. The same terms can be applied to the space groups themselves. Inversion centres are common because they often form a low-energy way for molecules to interact. However, an inversion centre would relate a chiral molecule to its enantiomer, and so an enantiopure compound must always crystallise in a non-centrosymmetric space group.

A non-centrosymmetric crystal structure cannot be superimposed on its inverted image, and determination of *absolute structure* amounts to assigning a particular non-centrosymmetric crystal structure to one of two possible structures which are related by inversion. The issue of absolute structure is relevant only to non-centrosymmetric crystal structures.

The inverted form of a crystal structure containing one enantiomer is a structure containing the opposite enantiomer. Therefore determination of the absolute structure of an enantiopure molecular crystal can be used to establish the absolute configuration of the molecules that comprise it. Notice that the word 'absolute' is being used here in two different contexts. To quote Howard Flack:[1] '*Absolute structure* is a crystallographer's term and applies to non-centrosymmetric crystal structures. *Absolute configuration* is a chemist's term and refers to chiral molecules.'

Inversion of a non-centrosymmetric crystal structure also leads to inversion of its diffraction pattern. In principle, therefore, the distinction between two possible absolute structures can be made by comparing the original and inverted diffraction patterns calculated from the original and inverted structural models with the one that is measured experimentally. The calculated pattern that agrees better with the experimental data defines which absolute structure is the correct one.

This would be simple were it not for the fact that X-ray diffraction patterns themselves are, at least approximately, centrosymmetric, *i.e.* the intensities of the reflections with Miller indices *hkl* and $\bar{h}\bar{k}\bar{l}$ are the same ($\bar{h}$ means $-h$). This is called Friedel's Law, and it arises because the sets of Miller planes *hkl* and $\bar{h}\bar{k}\bar{l}$ are identical. Inverted images of a

centrosymmetric diffraction pattern are, of course, the same, and if Friedel's law held exactly it would be impossible to draw the distinction described in the previous paragraph.

Fortunately, an effect called resonant scattering (or anomalous scattering or dispersion) introduces deviations from Friedel's law. The source of the effect is absorption of X-ray photons by excitation of the core electrons of the atoms of the crystal. An excellent web-site giving further details of the physical origin of the effect and its applications in macromolecular crystallography is available.[2]

Two features of resonant scattering are relevant to absolute structure determination. First, the effect is small compared to the contribution of non-resonant scattering. Some illustrative figures are given in Table 1. The contribution to the atomic scattering factor which introduces deviations from Friedel's law is given the symbol f'' and this should be compared to the value of the non-resonant contribution (f). Secondly, its magnitude depends on the wavelength of the X-rays used to measure the diffraction pattern and on the elements present in the crystal. Data for the two most common radiations used for X-ray diffraction (Mo-Kα and Cu-Kα) are given in Table 1.

Table 1: Resonant (f) and non-resonant (f' and f'') scattering factors for C and Cl. Note that the non-resonant scattering factors (f) are not dependent on wavelength, but do depend on resolution, while the opposite is the case for f' and f''. Data from ref. [3].

Atom	Radiation	$f(d = 5$ Å)/e	$f(d = 1$ Å)/e	f'/e	f''/e
Carbon	Mo *K*α	5.107	1.114	0.0033	0.0016
	Cu *K*α			0.0181	0.0091
Chlorine	Mo *K*α	15.234	4.023	0.1484	0.1585
	Cu *K*α			0.3639	0.7018

Resonant scattering effects are smallest for the 'light atoms' of the first two periods of the periodic table, and absolute structure determination therefore presents a particular challenge in exactly the area where it is most important, in organic chemistry. The problem is especially critical for compounds used in pharmaceutical applications, where enantiomers may show very different biological activities (*e.g.* see ref. [4] for a compilation of odours of selected enantiomers). In the past, absolute structure determination for purely light-atom compounds has usually only been possible by preparing a derivative containing either a heavy atom such as chlorine, or a group of known chirality. This is no longer the case, and one of the aims of this paper is to summarise recent progress that enables confident assignment of absolute structure even for hydrocarbons.

2. Further space group considerations for enantiopure crystal structures

Symmetry operations may be classified as *proper* or *improper*. Proper operations include pure rotations, pure translations and screw axes, which are a combinations of translation and rotation. The other operations which occur in crystal structures are inversion centres, mirror planes and other rotoinversions ($\bar{3}$, $\bar{4}$ and $\bar{6}$, equivalent to S_6, S_4 and S_3 operations in Schönflies notation) and glide planes, which are a combination of translation and reflection. These are all improper operations.

A chiral molecule is transformed into its enantiomer by an improper operation. Improper operations can therefore not occur in the space group of an enantiopure crystal. The restriction on inversion symmetry was referred to in Section 1, but the same applies to mirror planes, rotoinversions and glide planes. The most common space groups for molecular compounds, $P2_1/c$, $P\bar{1}$ and $C2/c$ are not possible for enantiopure crystal structures because they contain inversion centres.

Many non-centrosymmetric space groups, such as $Pna2_1$ and Cc also contain improper operations, and these too are impossible for enantiopure crystals. These space groups are, nevertheless, non-centrosymmetric and the absolute structure of any crystal structure forming in them still needs to be established. In space groups $Pna2_1$ and Cc the absolute structure would define the polarity of the crystal structure with respect to the crystal morphology rather the chirality or absolute configuration of the component molecules.

Chirally pure compounds may crystallise in the space groups containing proper operations only, the so-called Söhncke space groups. 65 of the 230 space groups fall into this category, and the most common are $P2_12_12_1$, $P2_1$ and $P1$. It is straightforward to recognise a Söhncke group: if after removal the first letter of the space group symbol all remaining characters are positive numbers then the space group is a Söhncke group and able to accommodate an enantiopure crystal structure. For example:

$P2_12_12_1$: Removing the *P* gives $2_12_12_1$. These characters are only positive numbers so this is a Söhncke group.

$P2_1/c$: Removing the *P* gives 2_1/c. The '/' and '*c*' are not numbers, so this is not a Söhncke group, and a compound with asymmetric centres crystallising in this space group would be a racemate.

$I\bar{4}$: Removing the I gives $\bar{4}$. The '$\bar{4}$', which is also sometimes written '−4', is not a positive number, so this is not a Söhncke group either. $I4$ and $I4_1$, on the other hand, are.

The term 'chiral space group' is sometimes used in the literature to mean a space group capable of accommodating a chirally pure crystal structure, *i.e.* a Söhncke group. This

usage has been criticised by Howard Flack because the term literally implies that the space group itself is chiral.[5] That is, if the symmetry operations are considered as objects then the space group is not superimposable on its mirror image. While this is true of a space group like $P3_1$, because the mirror image of a 3_1 operation is a 3_2 operation, it is not true of $P2_1$ (the mirror image of 2_1 is 2_1) even though both are Söhncke groups. There are 22 space groups, forming 11 'enantiomorphic pairs' for which the term 'chiral space group' is appropriate. $P3_1$ and $P3_2$ is one example of such a pair; $P4_12_12$ and $P4_32_12$ is another. A full list is available in most crystallographic text books and ref. [1].

A non-centrosymmetric crystal structure can usually be inverted by simply multiplying all the fractional coordinates by −1. There are two cases where additional steps are required:

i. If the space group belongs to one of the 11 enantiomorphic pairs. In this case the space group also needs to be changed to its enantiomorphic partner, so that if a structure in $P3_1$ is to be inverted the coordinates need to be multiplied by −1 *and* the space group should be changed to $P3_2$.

ii. If the space group is one of *Fdd2*, $I4_1$, $I4_122$, $I4_1md$, $I4_1cd$, $I\bar{4}2d$ or $F4_132$. In these cases, in addition to multiplication of the coordinates by −1, an additional origin shift is required as listed in ref. [6]. Note that of these only $I4_1$, $I4_122$ and $F4_132$ are Söhncke groups; all three are relatively rare for non-macromolecular materials.

3. Friedel pairs and Bijvoet pairs

If resonant scattering effects are large enough to be observable, the symmetry of the X-ray diffraction pattern of a crystal is directly related to the space group of the crystal structure. For a crystal structure in the centrosymmetric space group $P2_1/c$, reflections hkl, $\bar{h}k\bar{l}$, $h\bar{k}l$ and $\bar{h}\bar{k}\bar{l}$ are all related by symmetry and have the same intensities even if resonant scattering effects are substantial. Notice that for this centrosymmetric structure the pair of reflections hkl and $\bar{h}\bar{k}\bar{l}$ are related by symmetry and Friedel's law would hold rigorously.

If the space group is $P2_1$ the set of equivalences seen in $P2_1/c$ splits into two. The reflections hkl and $\bar{h}k\bar{l}$ form one symmetry-equivalent set; reflections $h\bar{k}l$ and $\bar{h}\bar{k}\bar{l}$ form another. In terms of symmetry, the point group of the diffraction pattern is has dropped from $2/m$ to 2 (or from C_{2h} to C_2 in Schönflies notation), and the equivalent reflections are related by the two-fold axis. Loss of the inversion and mirror symmetry that was present in $2/m$ means that the pairs of reflections (hkl and $\bar{h}\bar{k}\bar{l}$) and (hkl and $h\bar{k}l$) are no longer related by symmetry, so that Friedel's law does not strictly apply. As described in Section 1, the

difference in the intensities of these two pairs of reflections depends on the magnitude of the resonant scattering.

The reflections *hkl* and $\bar{h}\bar{k}\bar{l}$ are referred to as a *Friedel pair*.[7] *Bijvoet pairs* are pairs of reflections which are related to the Friedel pair by symmetry.[8] In the example in $P2_1$, the Friedel pair is equivalent to *hkl* and $h\bar{k}l$, and this would be a Bijvoet pair. In practice, however, the two terms *Friedel pair* and *Bijvoet pair* are often used interchangeably.[9] The differences in intensities between reflections *hkl* and $\bar{h}\bar{k}\bar{l}$ and their symmetry equivalents are called a *Bijvoet differences.*

Bijvoet pairs and differences are named after the Dutch crystallographer Johannes Bijvoet who was the first to use resonant scattering to establish the absolute configuration of a material in his structure determination of sodium rubidium tartrate in 1951.[10] The pronunciation of *Bijvoet* is (approximately) 'Bay-voot', to rhyme with 'boot', but is given more exactly online in ref. [11].

4. Flack's method of absolute structure refinement

The method for absolute structure determination most commonly used today is based on a formulation first described by Howard Flack.[12] It is usually applied towards to end of structure refinement, when the analyst has an essentially complete model, with non-H-atoms modelled with anisotropic displacement parameters, all the H-atoms located and any disorder modelled.

In Flack's method, the sample is considered to be a twin or composite composed of a *reference domain*, which has the absolute structure of the current refinement model, and a second domain in which the absolute structure is inverted. The model thus contains both possible absolute structures, and the absolute structure of the sample is found by determining the relative proportion of the inverted domain present. This proportion is given the symbol *x*, and it is called the *Flack parameter*.

The value of *x* has a physically meaningful value in the range of 0 to 1, and represents the fraction of the inverted structure present in the crystal. A value of $x = 0$ implies that none of the crystal is in the inverted form and therefore the model has the correct absolute structure; if $x = 1$ then all of the crystal is in the inverted form and the model should be inverted. Intermediate values of *x* point to inversion or 'racemic' twinning where some domains of the physical crystal contain one enantiomer, and other domains contain its inverse.

The value of *x* is determined as part of structure refinement. The intensity of a reflection *hkl* from a single crystal composed entirely of the reference domain would be

$|F_{\text{single}}(hkl)|^2$. The symbol F refers to the structure factor [the intensity of a reflection hkl is proportional to $|F^2(hkl)|$], and the subscript 'single' indicates that the quantity is computed from an un-twinned model consisting only of one enantiopure domain. If the crystal were composed entirely of the inverted domain then, as described in Section 1, the diffraction pattern would be inverted, and the intensity of that reflection would be $|F_{\text{single}}(\bar{h}\bar{k}\bar{l})|^2$. For the twinned crystal of Flack's model each reflection hkl is considered to have a contribution from both domains, so that the intensity of each reflection modelled as

$$|F_{\text{twin}}(hkl)|^2 = (1-x)\,|F_{\text{single}}(hkl)|^2 + x\,|F_{\text{single}}(\bar{h}\bar{k}\bar{l})|^2 \qquad [1]$$

5. The Precision of the Flack Parameter

In a crystallographic structure refinement a model consisting of the atoms and their coordinates and displacement parameters and an overall scale factor (which places the measured intensities on an absolute scale) is optimised so as to minimise the difference between the measured and calculated structure factor magnitudes or their squares. Other parameters may be added to the model, including occupancies if the structure is disordered, and the Flack parameter if the structure is non-centrosymmetric.

For most conventional non-macromolecular crystal structure determinations the number of observations typically exceeds the number of model parameters by a factor of about 10. This *over-determination* means that the value of a parameter is determined with an estimate of its precision, its *standard uncertainty.* Standard uncertainties are quoted in brackets following the value of a parameter, for example 1.520(4) Å for a bond length. The figure in brackets refers to the last quoted decimal place, and in this example the standard uncertainty on the measurement of 1.540 Å is 0.004 Å.

It is important to interpret the value of the Flack parameter in the context of its standard uncertainty. For example, a value of x = 0.2(8) has such a large standard uncertainty (0.8) that one neither knows whether the crystal is twinned by inversion or not, nor whether it should be inverted.

Flack and Bernardinelli considered how small the standard uncertainty, u, of x should be before any conclusion regarding absolute structure can be made.[13] They concluded that even if a compound is known to be enantiopure, the value of u should be less than 0.1 before any conclusions regarding absolute structure can be made. If the enantiopurity of the sample is unknown then the value of u should be less than 0.04. The value of x should also be within $2u$ of zero.

Note that experimental measurements of a quantity which is physically zero are expected to lie in a statistical distribution centred about zero because of the random errors present in the measurements. It is thus quite common to obtain refined Flack parameters that lie slightly outside the physical range of 0-1, *e.g.* −0.03(2). The *statistical range* of *x* is −3*u* to 1+3*u*. Recalling that 99.98% of the area under a Gaussian probability distribution function (pdf) lies within ±3*u* of the mean, this range captures values of *x* contained under Gaussian pdfs centred at the physical limits *x* = 0 and 1.

6. *FRIEDIF*

The ability to achieve a low standard uncertainty for the Flack parameter depends on the resonant scattering effects having sufficient magnitude to lead to measurable Bijvoet differences. As described in Section 1, this depends on the chemical elements present in the crystal and the wavelength of the X-rays used to collect the diffraction data. The magnitude of resonant scattering effects in a given experiment can be conveniently quantified by the *FRIEDIF* parameter. This is defined as the expected root mean square Bijvoet difference (*D*) divided by the expected average Bijvoet pair intensity (*A*), multiplied by 10^4 to place it on a convenient scale.

$$FRIEDIF = 10^4 \times \frac{< D^2 >^{\frac{1}{2}}}{< A >} \qquad [2]$$

Flack and Shmueli showed that the expectation values involved can be calculated statistically using only the formula of a compound and the wavelength used for data collection.[14]

If *FRIEDIF* has a value of about 80 or more, absolute structure determination presents little problem (Fig. 1).[15] However, resonant scattering effects for elements such as C, N and O are so small in comparison to the non-resonant scattering factors (*f*) even for Cu *K*α radiation, that values of *FRIEDIF* for compounds containing only these elements are well below 80, making it is difficult to determine the Flack parameter with sufficient precision to establish absolute structure for many organic compounds.

For example, the value of *FRIEDIF* for the amino acid *L*-alanine ($C_3H_7NO_2$) with Cu *K*α radiation is only 34. Accordingly, the value of the Flack parameter obtained from a conventional least squares refinement of *L*-alanine was −0.04(27).[16] The data set was of excellent quality, yet the precision of the Flack parameter is too low to enable a definitive statement to be made regarding the absolute structure.[13] The level of precision obtained here is consistent with a broader survey by Bernardinelli and Flack (Fig. 1).

One option might appear to be to carry out data collections with still longer X-ray wavelengths, such as Cr *K*α radiation (λ = 2.2909 Å). The values of *f″* for C, N and O at this wavelength are about double those for Cu *K*α. However, this approach compromises the maximum practically attainable resolution, which is about 1.2 Å for Cr *K*α, compared to about 0.8 Å for Cu *K*α, while systematic errors due to absorption also become substantial.

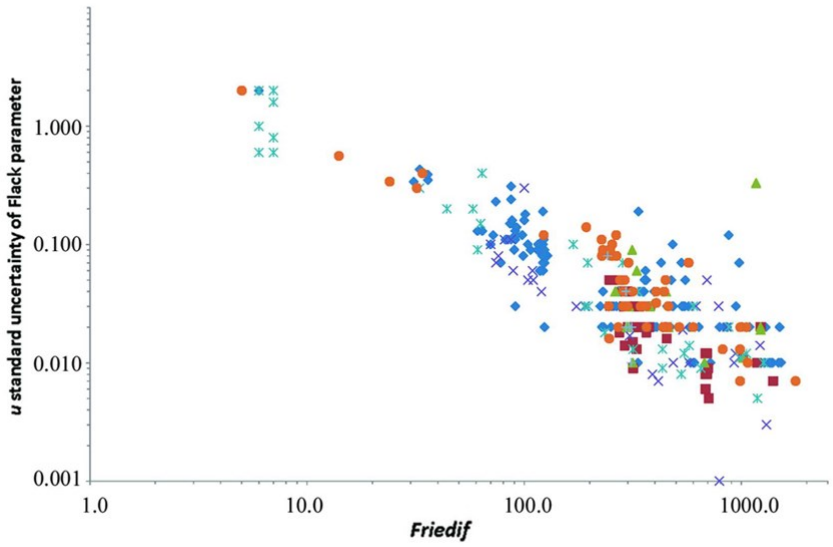


Figure 1: Plot of *u* versus *FRIEDIF* on logarithmic axes for non-centrosymmetric structures. Figure taken from ref. [15] with permission; the different symbols in the figure refer to the discussion in that paper.

7. A statistical anomaly in the precision of x for light atom structures

Thompson and Watkin plotted a histogram of the Flack parameters obtained by conventional least squares refinement for 150 structures for which *FRIEDIF* ranged between 3.4 and 10.8.[17] The values of *u* expected on the basis of Fig. 1 would have been 0.5 or higher, and since the resonant scattering was so low, it would be expected that the histogram would reflect a complete lack of knowledge of the absolute structures of the materials in question and be centred about *x* = 0.5. The histogram actually obtained is shown in Fig. 2a, and though it is broad, it can be seen that it is centred about a mean value of 0.027. As Thompson and Watkin conclude, 'despite the … weakness of the anomalous signal, on average it is observable.'

A consistent observation was made by Parsons, Flack and Wagner, who calculated values of *x*/*u* for around 20 high-quality data sets collected on materials of known enantiopurity.[16] The values of *x*/*u* for these materials should be scattered randomly about zero following a unit Gaussian probability distribution. The actual distribution obtained is

shown in Fig. 2b, and it can be seen that it does not fit the unit Gaussian distribution very well at all. The values of *x*/*u* are much more tightly distributed about zero than the standard uncertainties of the individual measurements would have suggested. This disagreement can be quantified by calculating the reduced χ^2, which has an ideal value of unity. The data in the histogram yield a value of only 0.03, suggesting that the values of *u* are overestimated by a factor of about 5.5.

Thus, although Flack parameters for light-atom compounds have high standard uncertainties, they tend to give the correct indication of absolute configuration. The standard uncertainty obtained by conventional least squares also appears to be overestimated. Methods which yield more realistic estimates of precision are described in the following sections.

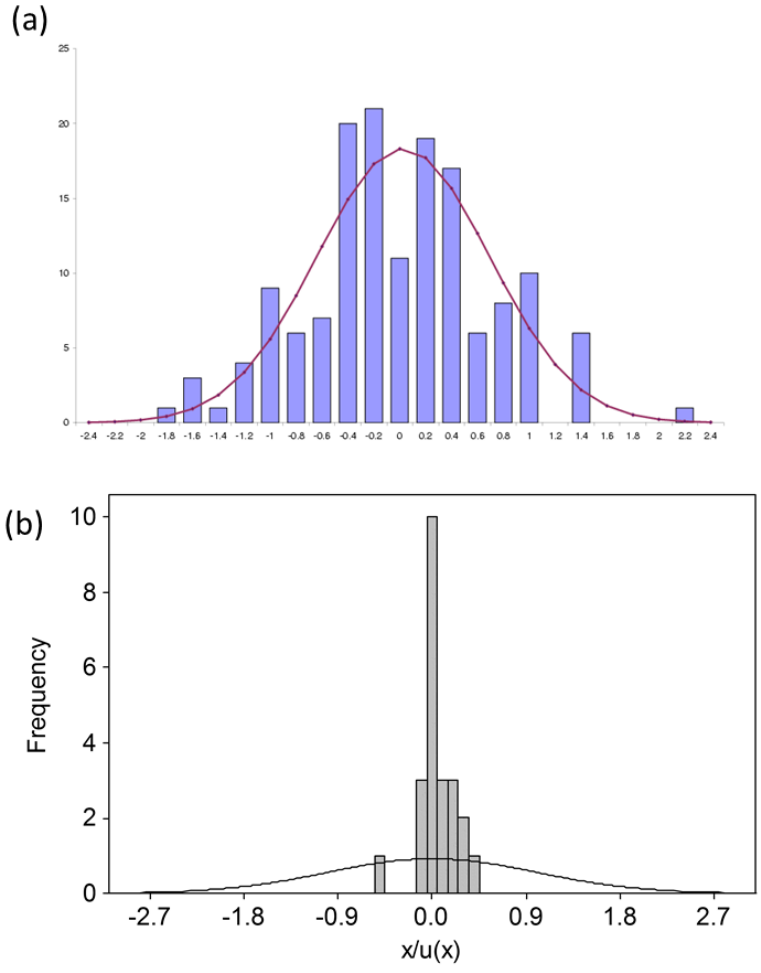


Figure 2: (a) Histogram showing the distribution of the Flack parameter for 150 light-atom structures. A theoretical Gaussian distribution (based on the sample mean and standard deviation) is overlayed to guide the eye. Figure taken from ref. [17] with permission. (b) Histogram of *x*/*u* determined for 23 structures in ref. [16] compared to a unit Gaussian probability density function.

8. Use of Bijvoet Differences and Quotients[16]

There are normally hundreds or thousands of individual Bijvoet differences in a data set (the .hkl file for SHELX[18] users). The observed Bijvoet difference for any reflection *hkl* is given by

$$D_{\text{obs}}(hkl) = |F_{\text{obs}}(hkl)|^2 - |F_{\text{obs}}(\bar{h}\bar{k}\bar{l})|^2 \quad [3]$$

Values of $D_{\text{obs}}(hkl)$ and estimates of their uncertainties can be readily extracted from a data set.

Calculated values of the Bijvoet differences, $D_{\text{model}}(hkl)$, can be obtained from the refinement model. Recalling that in Flack's method of absolute structure refinement a crystal is envisaged to be an inversion twin (Equ. 1), the analogue of Equ. 3 involving quantities calculated from a refinement model becomes

$$D_{\text{model}}(hkl) = |F_{\text{twin}}(hkl)|^2 - |F_{\text{twin}}(\bar{h}\bar{k}\bar{l})|^2 \quad [4]$$

Substitution of Equ. 1 written out for both $|F_{\text{twin}}(hkl)|^2$ and $|F_{\text{twin}}(\bar{h}\bar{k}\bar{l})|^2$ into this expression gives

$$D_{\text{model}}(hkl) = (1-2x)[|F_{\text{single}}(hkl)|^2 - |F_{\text{single}}(\bar{h}\bar{k}\bar{l})|^2] = (1-2x)D_{\text{single}}(hkl) \quad [5]$$

$F_{\text{single}}(hkl)$ refers to the structure factor of a reflection with indices *hkl* calculated from the refinement model with *x* set to zero. Ideally, the sets of D_{obs} and D_{model} are the same, but the former is subject to random and systematic measurement errors, so the agreement is never perfect.

The value of *x* can be obtained following a structure refinement in which it is not part of the model by recognising that Equ. 5 is a straight-line equation in the form *y* = *mx*. A plot of $D_{\text{obs}}(hkl)$ (which comes from the data set) against $D_{\text{single}}(hkl)$ (which comes from the model) is a straight line with gradient (1-2*x*). The gradient of the plot, and its uncertainty, can be extracted by simple linear regression using weights based on the uncertainties of D_{obs}, and the values of *x* and *u* then calculated.

Quotients, *Q*, where

$$Q_{\text{obs}}(hkl) = \frac{|F_{\text{obs}}(hkl)|^2 - |F_{\text{obs}}(\bar{h}\bar{k}\bar{l})|^2}{|F_{\text{obs}}(hkl)|^2 + |F_{\text{obs}}(\bar{h}\bar{k}\bar{l})|^2} = \frac{D_{\text{obs}}(hkl)}{2A_{\text{obs}}(hkl)} \quad [6]$$

$$Q_{\text{model}}(hkl) = (1-2x)\frac{|F_{\text{single}}(hkl)|^2 - |F_{\text{single}}(\bar{h}\bar{k}\bar{l})|^2}{|F_{\text{single}}(hkl)|^2 + |F_{\text{single}}(\bar{h}\bar{k}\bar{l})|^2} = (1-2x)Q_{\text{single}}(hkl) \quad [7]$$

where A_{obs} is the average Bijvoet pair intensity, can be used in a similar way. Use of quotients has the advantage that *Q* is independent of the overall scale factor and, in principle at least, suffers from reduced absorption and extinction effects. In practice the results of using *D* or *Q* are usually similar; SHELXL-2012 and subsequent versions employ the quotient method.[18]

Fig. 3a shows an example of a quotient plot for the data set of *L*-alanine referred to in Section 6, for which a value of $x = -0.04(27)$ had been obtained. The gradient of the fitted line (= $1 - 2x$) is 0.984(68), giving $x = 0.01(3)$, which is substantial improvement in precision. Parsons, Flack and Wagner showed that the improved performance was sustained over 23 test data sets, with a reduced χ^2 much closer to 1.[16] The value of x for cholestane ($C_{27}H_{48}$, *FRIEDIF* = 9) was found to be −0.01(13) from quotients (Fig. 3b) and −0.02(8) from differences, the latter defining the absolute structure even for as a challenging a system as a hydrocarbon.

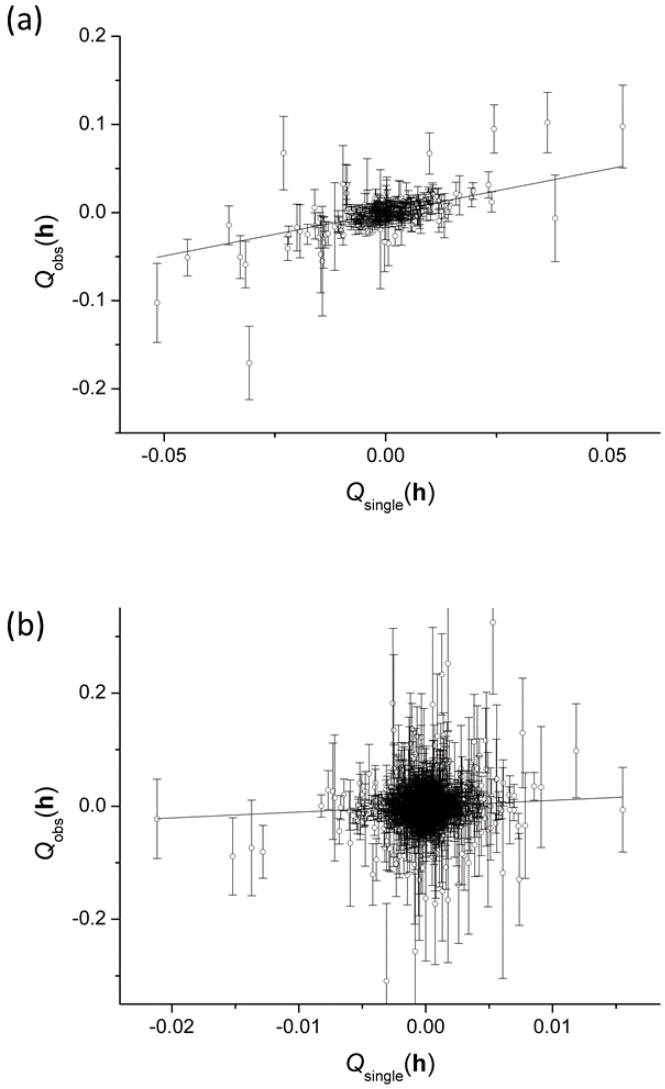


Figure 3: Plots of Q_{obs} against Q_{single} for (a) *L*-alanine and (b) cholestane. The gradient (= $1-2x$) can be used to evaluate x. Note that $u(x) = u(\text{gradient})/2$. The bold **h** in the axis labels stands for the Miller indices *hkl*.

The methods described above are applied after a structure has been refined. This approach carries the risk that it fails to allow x to refine along with the scale factor, atomic

positions and displacement parameters, and so any inter-dependence (correlation) of these parameters and *x* will be unaccounted for. An advantage of the approach described here, which resolves this difficulty, is that the observed values of the differences or quotients can be included as restraints in the main structure refinement.[19] The results obtained are very similar those from the post-refinement procedure, indicating that correlations involving *x* and the other parameters can be neglected. Results obtained in a recent survey by Watkin & Cooper support this conclusion.[20]

9. Bayesian Methods[21] and the probability that a proposed absolute structure is correct

Thomas Bayes was an 18th Century English clergyman who invented a method for combining probabilities, called Bayes' Theorem, which now forms the basis of a school of statistical thought which takes a probabilistic approach to data analysis. For example, the equations of least squares can be derived using a Bayesian approach by recasting the data-fitting problem into one which finds the most probable set of parameters given a set of measured data.

In the context of absolute structure determination Bayes' Theorem enables the probability that a proposed absolute structure is correct to be evaluated using the equation:

$$p(x = 0 \mid D_{\text{obs}}) = \frac{p(D_{\text{obs}} \mid x = 0)p(x = 0)}{p(D_{\text{obs}})} \qquad [8]$$

The bar, '|', means 'given', so $p(D_{\text{obs}}|x = 0)$ means the probability of the set of D_{obs} given $x = 0$; the way this is calculated is described below. In the crystallographic literature, the symbol *x* in Equ. 8 is usually replaced by *y* and referred to as the Hooft parameter, which can be thought of as the Flack parameter determined using Bayesian methods. The physical interpretation of *x* and *y* are, however, the same and so in the interests of simplicity the symbol *x* will be used below.

The term $p(D_{\text{obs}}|x = 0)$ in Equ. 8 can be obtained by calculating the values of $|F_{\text{single}}(hkl)|^2$ from a refinement model where the Flack parameter is fixed to 0. For each Bijvoet difference in the data set the quantity

$$z_{hkl} = \frac{D_{\text{single}}(hkl) - D_{\text{obs}}(hkl)}{u(D_{\text{obs}}(hkl))} \qquad [9]$$

can be defined, which is the number of standard deviations difference between the value of $D_{\text{single}}(hkl)$ calculated from the refinement model and the observed value $D_{\text{obs}}(hkl)$. The uncertainty of $D_{\text{obs}}(hkl)$ is obtained by propagation error from the component uncertainties in $|F_{\text{obs}}(hkl)|^2$ and $|F_{\text{obs}}(\bar{h}\bar{k}\bar{l})|^2$:

$$u(D_{\text{obs}}(hkl)) = \sqrt{u^2(|F_{\text{obs}}(hkl)|^2) + u^2(|F_{\text{obs}}(\bar{h}\bar{k}\bar{l})|^2)} \qquad [10]$$

For example if in a data set the values of $|F_{\text{obs}}|^2$ and $|F_{\text{single}}|^2$ for a Bijvoet pair were as shown below

h	k	l	$\lvert F_{\text{single}}(hkl)\rvert^2$	$\lvert F_{\text{obs}}(hkl)\rvert^2$	$u(\lvert F_{\text{obs}}(hkl)\rvert^2)$
$\bar{6}$	2	1	68.55	65.70	0.71
6	2	1	67.61	64.50	0.71

then z_{hkl} for this Bijvoet pair would be

$$z_{621} = \frac{(67.61 - 68.55) - (64.50 - 65.70)}{\sqrt{0.71^2 + 0.71^2}} = 0.259$$

The necessary data for this calculation are available in, for example, an .fcf file generated by a LIST 4 instruction in SHELXL.

If the measurement errors of $|F_{\text{obs}}|^2$ are assumed to follow a Gaussian distribution, which is expected to be the case for experimental data, then the probability of obtaining a particular value of z_{hkl} is

$$p(z_{\text{hkl}}) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-z_{\text{hkl}}^2}{2}\right) \qquad [11]$$

or 0.386 in the example. If these probabilities are calculated for all the Bijvoet pairs in the data set the overall value of $p(D_{\text{obs}}|x = 0)$ is obtained by just multiplying all of the individual values of $p(z_{\text{hkl}})$ together:

$$p(D_{\text{obs}} \mid x = 0) = \prod_{\substack{\text{all Bijvoet} \\ \text{pairs, hkl}}} p(z_{\text{hkl}}) \qquad [12]$$

The capital 'pi' symbol, $\prod$, in Equ. 12 stands for a product in the same way that $\sum$ would indicate a summation.

The term $p(x = 0)$ in Equ. 8 is the probability that $x = 0$. If the compound is known to be enantiopure then x could be 0 or 1 with equal probability, so $p(x = 0) = 0.5$.

The term $p(D_{\text{obs}})$, is the probability of obtaining the observed data. Although the physical significance of this quantity is somewhat difficult to grasp, it can be evaluated in a number of ways. The most general is to recognise that the probability that x has some value is 1, so that the total probability function $p(\text{x}|D_{\text{obs}})$ calculated for all values of x between 0 and 1 must be normalised, *i.e.*

$$\int_{x=0}^{x=1} p(x \mid D_{\text{obs}}) dx = 1 \qquad [13]$$

However, if the crystal is known to be enantiopure, so that $x = 0$ or $x = 1$, it becomes

$$p(D_{\text{obs}}) = p(D_{\text{obs}} \mid x = 0)p(x = 0) + p(D_{\text{obs}} \mid x = 1)p(x = 1) \qquad [14]$$

which can be evaluated as described above.

Hooft and co-workers have shown that the 'true-false' probabilities obtained using the Bayesian approach are decisive even for cases where *FRIEDIF* is well below 80. In the example of *L*-alanine the probability that the refinement model was correct [*P*2(true)] was 1.000 (i.e. 100%, corresponding to complete certainty) based on 285 Bijvoet pairs. The two-way true-false test can be extended to a three-way test to include the possibility of racemic twinning. For *L*-alanine the probabilities that the sample was a racemic twin or incorrect were *P*3(rac-twin) = 10^{-30} and *P*3(false) = 10^{-121}, respectively. Even in the case of cholestane, for which 3895 Bijvoet pairs were measured, *P*2(true) = 1.000, *P*3(rac-twin) = 3 x 10^{-6} and *P*3(false) = 4 x 10^{-22}.

The probability of any value of x can be calculated by applying the methods described above, which can thus be taken further to build a complete probability distribution function. It is mathematically convenient to describe the distribution in terms of $1-2x$ rather than x itself, and the resulting probability density function in similar in appearance to a Gaussian distribution. The standard formulae of statistics can be applied to resulting function to extract the expectation value $\langle 1-2x \rangle$ and its uncertainty, and thereby the value of the Flack (or Hooft in this context) parameter and its uncertainty. For *L*-alanine and cholestane the results are 0.01(4) and −0.02(10), respectively.

The values of x and its uncertainty obtained by the methods described in this and the previous sections are usually found to be the same or very similar. The reason for this is that the use of least squares to fit the values of D_{obs} and D_{single} or Q_{obs} and Q_{single} carries the implicit assumption that the measurement errors are Gaussian, which is the same assumption applied in Equ. 11.

The two and three way tests and evaluation of the Flack parameter using Bayesian methods has been implemented in the programs PLATON[22] and CRYSTALS.[23]

10. Validation and problem cases

Absolute structure determination using either of the methods described in Sections 8 and 9 is robust against small errors in the structural model, but is sensitive to the presence of errors in the measured Bijvoet differences. These may be undetected systematic effects, such as a small amount of twinning,[20] or measurements which are simply wrong for some

reason, so-called outliers. It is usual to filter-out suspicious data such as those Bijvoet differences for which $D_{obs}(hkl)$ is more than twice the maximum theoretical Bijvoet difference calculated for the structure under investigation.[22] Bijvoet pairs for weak reflections give high values of $Q_{obs}(hkl)$ because the denominator of Equ 6 is small rather than because they contain significant resonant signal, and these too should also be omitted from calculations based on quotients.[16]

Even when filters are applied it is not unusual to find a few remaining outliers, which can be detected by plotting D or Q_{obs} against D or Q_{single}. Fig. 4 shows an example where omission of the three outliers indicated with the circles changed x from 0.25(5) to 0.06(8). Looking at the plots used to evaluate x in this way is an informative means for validation, giving an often salutary view of the data from which conclusions about absolute structure are being drawn. A number of further examples of problematic structures have been given with extensive analysis and discussion by Watkin and Cooper.[20]

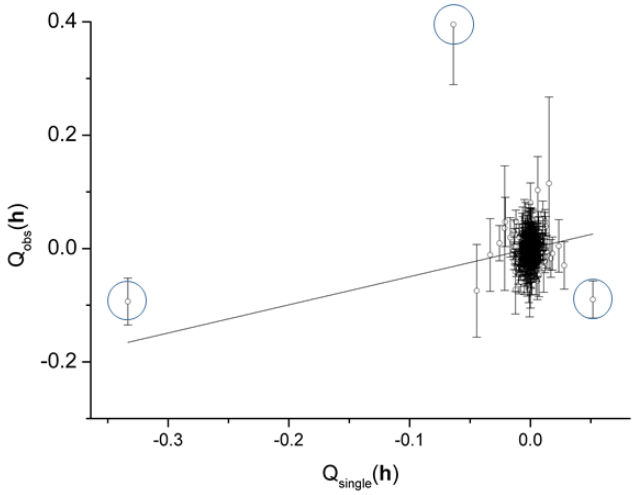


Figure 4: The value of x is sensitive to outliers. Removal of the three data points indicated changes x from 0.25(5) to 0.06(8).

It is also important to assess the validity of the assumption of Gaussian measurement errors. One way to do this would be to compare a histogram of $D_{obs} - D_{model}/u(D_{obs})$ or $Q_{obs} - Q_{model}/u(Q_{obs})$ with a unit Gaussian distribution (Fig. 5a), but a more quantitative approach is to calculate a Gaussian probability plot which converts the comparison into a straight-line which ideally has unit gradient and passes through the origin (Fig. 5b).[24] Deviations from linearity indicate a departure from Gaussian behaviour, possibly because of systematic errors which have not been accounted for in data reduction or the refinement model. A

similar conclusion might be drawn if the various different methods described here for evaluation of *x* give substantially different results.[20]

Recent work has shown that iterative reweighting of the straight-line fitting of the method described in Section 8 can help to reduce the effect of outliers or non-ideal error distributions characteristic of lower-quality data-sets.[20] However, an advantage of the Bayesian method outlined in Section 9 is that non-Gaussian behaviour can be accounted for by simply applying an alternative probability distribution function in place of Equ 11. A practically suitable choice has been found to be the Student-*t* distribution, which, having broader tails than a Gaussian distribution, can better accommodate outlying data points (Fig. 6).[25] The shape and tails of a Student-*t* distribution are controlled through a parameter ν, which is called the number of degrees of freedom. A distribution with ν above about 15 is very similar to a Gaussian, but values below 10 or even 5 indicate substantial deviations from ideality.

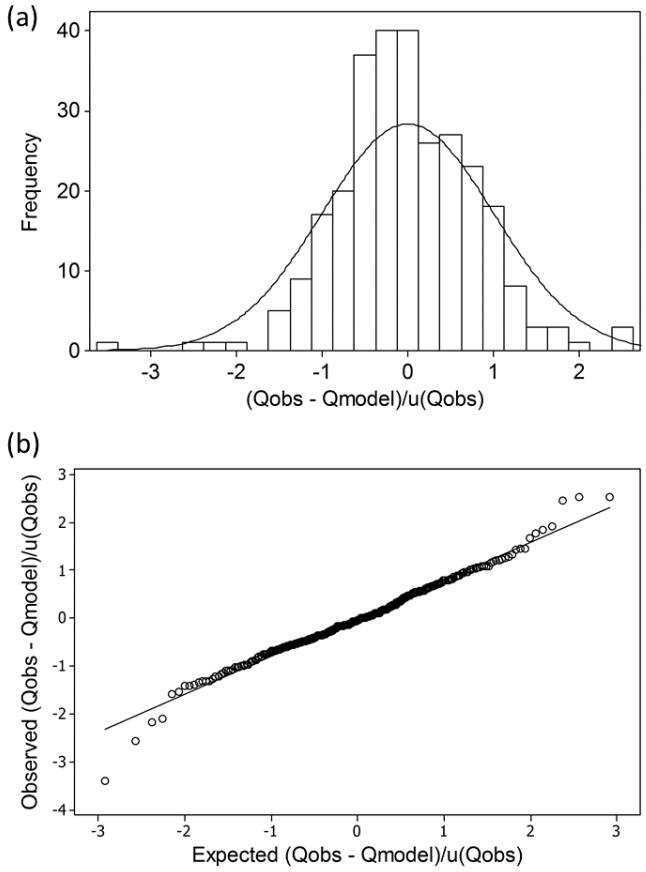


Figure 5: Two methods for comparing normalised deviates with a Gaussian distribution. (a) Simply compare a histogram with a Gaussian distribution function, or (b) convert the comparison into a straight line with a Gaussian probability plot. The line in (b) is has a gradient of 0.80, an intercept of −0.014 and a correlation coefficient of 0.992. Ideal values are 1, 0 and 1, respectively. It would be sensible to omit or down-weight the 7 points which deviate substantially from the line, and to rescale the weights to achieve a gradient is closer to 1.

Use of a Student-*t* distribution enables a robust estimates of *x* and its uncertainty to be obtained for problem data sets.[25] Procedures have been devised for automatic determination of the most appropriate value of ν; the validity of the chosen value of ν can be assessed using a Student-*t* probability plot,[26] which, like the Gaussian probability plot referred to above, should be a straight line of unit gradient passing through the origin. In fact the value of ν is a useful validation criterion in its own right. Values below about 10 ideally need some explanation of what was wrong with the data.

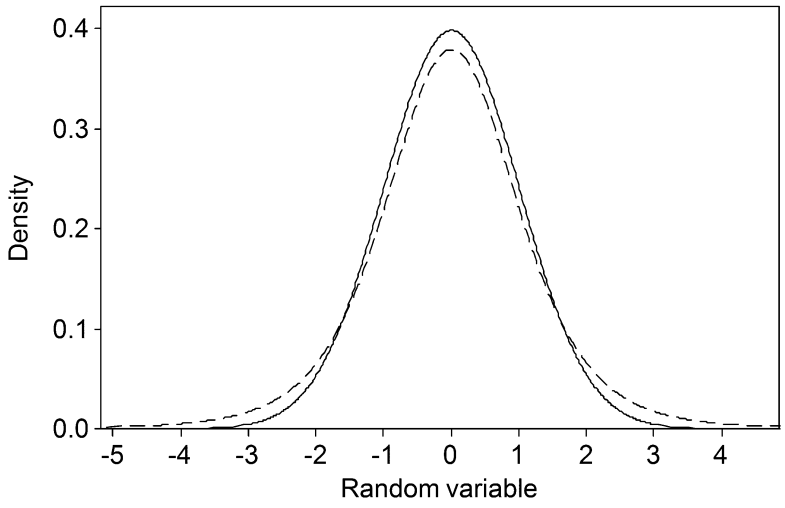


Figure 6: Comparison of a unit Gaussian distribution (full line) with a Student-*t* distribution with 5 degrees of freedom (dashed line). Notice that the tails of the Student-t distribution extend further than those of the Gaussian distribution. A Student-t distribution with more than about 15 degrees of freedom is very similar to a Gaussian.

11. Concluding remarks

According to *Wikipedia*, it was Mark Twain, or possibly Benjamin Disraeli, who remarked that 'There are three kinds of lies: lies, damned lies, and statistics'. It certainly seems remarkable that data such as those shown in Fig. 3b can be used to assign absolute structure with the absolute certainty alluded to in Section 9. However, the two-way Bayesian test can be viewed as asking whether Fig. 3b is more like Fig. 7a (the line $y = x$) or Fig. 7b (the line $y = -x$). If these are the only two possibilities, the choice is fairly clear: Fig. 3b is more like Fig. 7a than Fig. 7b, and this is the origin of the level of certainty. In fact, data in the centre of Fig 3b have very little influence on the path of the line of best fit, and when this is taken into account the choice becomes even clearer.[16] The validity of the underlying assumption of enantiopurity is clearly critical, and experimental methods for demonstrating this are discussed in ref. [27] of this memorial issue of *Tetrahedron Asymmetry*.

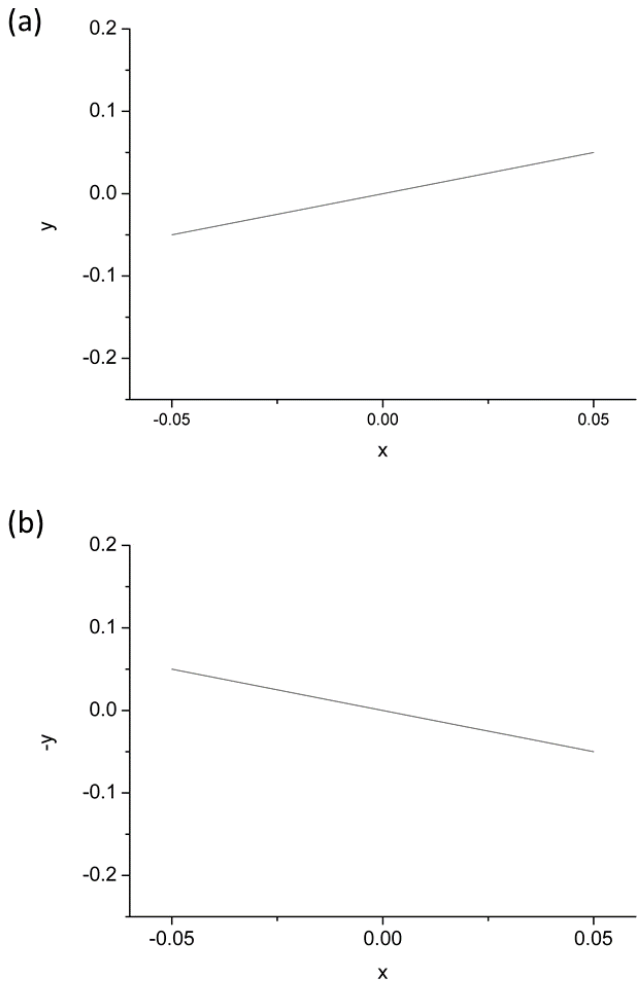


Figure 7: The lines (a) $y = x$ and (b) $y = -x$.

The reason that conventional least squares refinement of x performs so poorly in comparison to the difference, quotient or Bayesian methods is a question which is still the subject of research. Use of weighting schemes in structure refinement which attempt to account for model errors, such those which arise from the use of spherical atom scattering factors, as well as random measurement errors have been shown to degrade precision when applied to the methods of Section 8.[20] Overall absolute structure is quite a subtle detail, and diffraction data are much less sensitive to it than they are to atomic positions and displacement parameters. It is only when the data which are sensitive to absolute structure are 'separated-out' from rest that the underlying precision becomes apparent.[16]

The influence of different reflection intensities on the value of the Flack (or any other) parameter can be calculated using *leverage analysis*.[28] Not all reflections contribute equally, and in fact there are usually only a relatively small number that are really influential. Fig. 8 shows an example for *L*-alanine with data measured using Cu *K*α radiation, where the

quantities T or T^2 which measure the influence of a reflection on the Flack parameter, are plotted against various other quantities.

Fig. 8a shows T^2 plotted against sinθ/λ, which measures resolution, with the highest resolution data being on the right hand side of the plot. It is clear that the data at high resolution are most influential. The reason for this can be seen from Table 1: because the non-resonant scattering factor decreases with resolution, while the resonant scattering factors do not, the contribution of resonant effects is greatest at high resolution. Fig. 8*b*, where T^2 is plotted against $|F_{obs}|$, shows that the most influential reflections are quite weak, but Fig. 8*c*, a plot of T against $D_{model}(hkl)/u[D_{obs}(hkl)]$, shows that the most influential data are those for which the Bijvoet signal is most above background, illustrating a conclusion reached by Rabinovich and Hope.[29] The demands of Figs 8b and c seem somewhat contradictory, and this is perhaps why relatively few data have a really decisive influence on x.

The weakness of the resonant signal means that it is necessary to take considerable care when collecting diffraction data if the intention is to assign absolute configuration. Specifically, the plots in Fig. 8 show that precise absolute structure determination requires weak high-resolution data to be measured as carefully as possible. The subject of optimising data collection procedures for absolute structure determination could form a paper in its own right, but recommendations would be likely to include the following:

i. Collecting data to as high a resolution as an instrument will allow, to d = 0.84 Å or higher with Cu $K\alpha$ radiation.
ii. Collecting data at low temperature, experience suggests that between 100 and 150 K is a suitable range.
iii. Collecting data to high redundancy, so equivalent reflection intensities are measured many times to improve statistics; redundancies averaging 5 and between 8 and 35 were used in refs. [30] and [16], respectively. Experience suggests that a value at the lower end of this range, 6-8, depending on the time available, should be suitable for most samples.
iv. Collecting as complete as possible set of Bijvoet pairs.
v. Collecting data using a high-quality crystalline sample.

Fortunately, stable low-temperature devices and modern diffractometer software should enable criteria i-iv to be met routinely, and a survey of current practice is available in this memorial issue in ref.[31]. Number (v) is probably the most important, and depends on optimising crystal growth, which can be a substantial experimental challenge. Note that since it is necessary to make separate observations of the intensities of each reflection in a

Bijvoet pair, absolute configuration can only be determined using single-crystal diffraction data, and not from powder diffraction data.

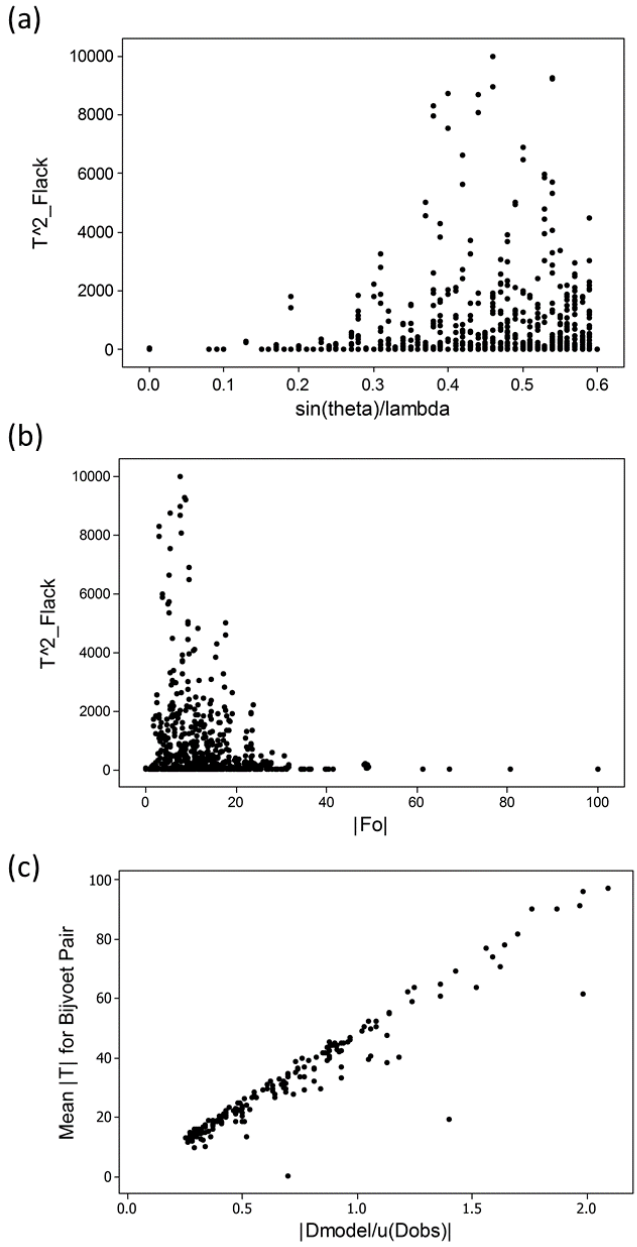


Figure 8: Leverage analysis for *L*-alanine,[28a] showing variation of T^2 for the Flack parameter with (a) resolution and (b) intensity. (c) shows how the 'measurability' of a Bijvoet difference (expressed as the calculated difference divided by the uncertainty of the observed difference) determines the influence of a Bijvoet pair. The quantity *T* measures the influence of a measurement of a specific parameter, here the Flack parameter.

Given suitable data, the methods outlined here have been shown to provide clear indications of absolute structure even for hydrocarbons. In general it is advisable to collect diffraction data on such materials with Cu *K*α radiation, but an illustration of the power of the new methods has been given by Escudero-Adán, Benet-Buchholz and Ballester,[30] who successfully determined a series of light-atom absolute structures using Mo *K*α radiation. The values of *FRIEDIF* these materials ranged from just 5.6 to 7.1, and the work required

collection of data to prodigiously high resolution, yet standard uncertainties below 0.1 were obtained in 42 out of 44 cases.

Readers who wish to try the methods described for themselves can obtain test data-sets from the electronic supplementary materials of refs [16] and [30].

Acknowledgements

This paper is dedicated to the memory of Dr. Howard Flack whose contributions to the field of absolute structure determination, indeed of small-molecule crystallography in general, can hardly be over-stated. He was always extremely supportive of colleagues wishing to try or discuss new ideas, being able to draw on a vast experience of practical, theoretical and mathematical crystallography. He is sorely missed.

It is a pleasure to acknowledge a long-standing collaborations with Dr. Alan Coehlo (Coelho Software), Professor Richard Cooper (Oxford), Dr. Trixie Wagner (Novatis) and Dr. David Watkin (Oxford) and Edinburgh students Thomas Heathcote, Murray MacLeod, Paul McGovern and Peter Wood. The methods of Section 8 have been included in the programs XPREP and SHELXL by Professor George Sheldrick (University of Göttingen), which has greatly expanded the number and range of materials to which the methods have been applied. The author would also like to thank Dr David Watkin and Professor Ton Spek (Utrecht University) for their comments on the manuscript, and Professor Ethan Merritt of the Biomolecular Structure Center at The University of Washington and Professor Ton Spek for advice on the origins of the terms *Bijvoet pair* and *Friedel pair*. The author is also very grateful to Professor Spek for his advice on the pronunciation of *Bijvoet*, which, it turns out, he had been innocently mangling for years.

Many of the ideas which led to the methods described in Sections 8 and 10 emerged from research projects funded by EPSRC in the apparently remote area of high-pressure crystallography.

Appendix: Summary of symbols used

hkl: The Miller indices of a reflection, sometimes abbreviated to the vector symbol **h**. The Miller indices, which are integers, can be interpreted as the coordinates of a reflection spot in a diffraction pattern, or as a description of the crystal planes from which the reflection is scattered. A data-set will usually contain hundreds or thousands of reflections, each with its own values of *h*, *k* and *l.*

d(*hkl*): The distance between the crystal planes from which a reflection with indices *hkl* is scattered. The resolution of a data set is expressed by the minimum value of *d* for which

reflection intensities have been measured. An alternative to *d* is sinθ/λ (=1/2*d* by Bragg's law), where θ is the Bragg angle and λ the X-ray wavelength. The resolution of a data set can also be expressed by the maximum value of sinθ/λ for which data have been measured. The units of *d* are usually expressed in Å, those of sinθ/λ in $Å^{-1}$.

x: The Flack parameter. See Section 4.

u: The standard uncertainty of the Flack parameter. Some texts and papers use the symbol σ for this quantity. Older papers refer to standard uncertainties as *estimated standard deviations* or e.s.d.s.

y: The Hooft parameter. This used to refer to the value of x when it is determined using Bayesian probability methods (Section 9).

T: The influence of an observation of the refined value of a parameter.

Structure factors:

$F(hkl)$: The structure factor of a reflection with Miller indices *hkl*. The structure factor is a complex quantity (*i.e.* it has real and imaginary components), encoding both the aplitude and the phase of a scattered X-ray. Only the magnitude of $F(hkl)$, $|F(hkl)|$ is measured in a diffraction experiment, where the intensity of a spot is proportional to $|F(hkl)|^2$. The phases cannot be measured: this is the famous *phase problem*.

$|F_{obs}(hkl)|^2$: The observed value of $|F(hkl)|^2$ determined from an experimentally measured diffraction pattern.

$|F_{calc}(hkl)|^2$: The value of $|F(hkl)|^2$ calculated from a structural model, consisting of a set of atom types, coordinates, displacement parameters *etc.* When absolute structure is considered, the value of $|F_{calc}(hkl)|^2$ with $x = 0$ is referred to as $|F_{single}(hkl)|^2$. When *x* has a general, possibly non-zero, value $|F_{calc}(hkl)|^2$ is referred to as $|F_{twin}(hkl)|^2$. The origin of the terms 'single' and 'twin' in this context comes from Flack's method for competitive refinement of one absolute structure against its opposite, which treats the sample as an inversion twin. See Section 4. Note that both $|F_{single}(hkl)|^2$ and $|F_{twin}(hkl)|^2$ are calculated from a refinement model.

Bijvoet differences:

$D(hkl)$: The Bijvoet difference between $|F(hkl)|^2$ and $|F(\bar{h}\bar{k}\bar{l})|^2$ or a symmetry equivalent. The value of *D* measured from the diffraction pattern is

$D_{obs}(hkl) = |F_{obs}(hkl)|^2 - |F_{obs}(\bar{h}\bar{k}\bar{l})|^2$ (Equ. 3).

The value of $D(hkl)$ calculated from a single-domain refinement model, *i.e.* with $x = 0$, is

$D_{single}(hkl) = |F_{single}(hkl)|^2 - |F_{single}(\bar{h}\bar{k}\bar{l})|^2$

For a model structure with a general value of x

$D_{model}(hkl) = (1\text{-}2x)D_{single}(hkl)$ (Equ. 5)

Note that $D_{single}(hkl)$ and $D_{model}(hkl)$ are calculated from a refinement model.

Quotients

$Q(hkl)$: The Bijvoet difference between $|F(hkl)|^2$ and $|F(\bar{h}\bar{k}\bar{l})|^2$ or a symmetry equivalent divided by the corresponding sum of $|F(hkl)|^2$ and $|F(\bar{h}\bar{k}\bar{l})|^2$ (or a symmetry equivalent).

The value of Q measured from the diffraction pattern is

$$Q_{obs}(hkl) = \frac{|F_{obs}(hkl)|^2 - |F_{obs}(\bar{h}\bar{k}\bar{l})|^2}{|F_{obs}(hkl)|^2 + |F_{obs}(\bar{h}\bar{k}\bar{l})|^2} = \frac{D_{obs}(hkl)}{2A_{obs}(hkl)} \quad \text{(Equ.6)}$$

where $A_{obs}(hkl) = (|F(hkl)|^2 + |F(\bar{h}\bar{k}\bar{l})|^2)/2$.

$Q_{single}(hkl)$ and $Q_{model}(hkl)$, which are calculated from a refinement model, are defined in an analogous way to $D_{single}(hkl)$ and $D_{model}(hkl)$.

Probability formulae

p(A|B): The probability of *A* given *B*. Probabilities have values between 0 and 1.

P2 test: The probability, determined using Bayesian methods, that a structure has the proposed absolute structure of the refinement model, assuming that this and its opposite are the only possibilities available (i.e. the material being studied is enantiopure).

P3 test: The probability, determined using Bayesian methods, that a structure has the proposed absolute structure of the refinement model, assuming that this, a 50:50 inversion twin, and the inverted structure are the only possibilities available.

Gaussian probability density function (pdf): A probability density function defined as

$$f(z \mid \mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\mu - z)^2}{2\sigma^2}\right)$$

where μ is the mean of the distribution and σ^2 its variance. In Hooft's method, described in Section 9, μ is associated with $D_{single}(hkl)$, z with $D_{obs}(hkl)$ and σ with $u[D_{obs}(hkl)]$. A unit Gaussian pdf has μ = 0 and σ = 1. Experimental data are usually assumed to have Gaussian distribution of random errors on account of the central limit theorem, which states that when a measurement is subject to many sources of random error, the overall error distribution is Gaussian. The Gaussian distribution is also often referred to as the *normal distribution*.

Student-t probability density function: A probability density function defined as

$$p(z \mid \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{z^2}{\nu}\right)^{-\frac{(\nu+1)}{2}}$$

where ν is the number of degrees of freedom. The Γ function is closely related to a factorial, but can also be defined for non-integral values. This function is applied in Hooft's method (Section 9) because it has convenient mathematical properties rather than for some fundamental physical reason.

References

1. Flack, H. D.; Bernardinelli, G., The use of X-ray crystallography to determine absolute configuration. *Chirality* **2008,** *20* (5), 681-690.
2. Merritt, E. A. X-ray Anomalous Scattering. http://skuld.bmsc.washington.edu/scatter/AS_index.html (accessed 16th August 2017).
3. Creagh, D. C., X-ray Dispersion Corrections. In *International Tables for Crystallography Volume C*, Prince, E., Ed. Kluwer Academic Publishers: Dordrecht, Boston, London, 2004; pp 255-257.
4. Leffingwell, J. C. Chirality & Odour Perception: Steroid Urine Type Odorants & Sandalwood Type Odorants. http://www.leffingwell.com/chirality/steroid_%26_sandalwood.htm (accessed 16th August 2017).
5. Flack, H. D., Chiral and Achiral Crystal Structures. *Helvetica Chimica Acta* **2003,** *86* (4), 905-921.
6. Bernardinelli, G.; Flack, H. D., Least-squares absolute-structure refinement. Practical experience and ancillary calculations. *Acta Crystallographica Section A* **1985,** *41* (5), 500-511.
7. Friedel, G., Sur les symétries cristallines que peut révéler la diffraction des rayons Röntgen. (Crystalline symmetry as revealed by the diffraction Röntgen rays). *Compt. Rend.* **1914,** *157*, 1533-6;158,130-1.
8. (a) Peerdeman, A. F.; Bijvoet, J. M., The indexing of reflexions in investigations involving the use of the anomalous scattering effect. *Acta Crystallographica* **1956,** *9* (12), 1012-1015; (b) Ramaseshan, S., The Use of Anomalous Scattering in Crystal Structure Analysis. In *Advanced Methods of Crystallography*, Ramachandran, G. M., Ed. Academic Press: London, New York, 1964; pp 67-95; (c) Vaidya, S. N.; Ramaseshan, S., Some Procedures in the Determination of the Absolute configuration of Crystals In *Crystallography and crystal perfection*, Ramachandran, G. N., Ed. Academic Press London, New York, 1963; pp 243-255.
9. Vijayan, M.; Ramaseshan, S., Isomorphous Replacement and Anomolous Scattering. In *International Tables for Crystallography Volume B: Reciprocal Space*, Shmueli, U., Ed. Kluwer Academic Publishers: Dordrecht, Boston and London, 2001; pp 264-275.
10. (a) Bijvoet, J. M.; Peerdeman, A. F.; van Bommel, A. J., Determination of the Absolute Configuration of Optically Active Compounds by Means of X-Rays. *Nature* **1951,** *168* (4268), 271-272; (b) Lutz, M.; Schreurs, A. M. M., Was Bijvoet right? Sodium rubidium (+)-tartrate tetrahydrate revisited. *Acta Crystallographica Section C* **2008,** *64* (8), m296-m299.
11. Forvo Media SL How to pronounce Bijvoet. https://nl.forvo.com/word/bijvoet (accessed 17th August 2017).
12. Flack, H. D., On enantiomorph-polarity estimation. *Acta Cryst.* **1983,** *A39* (6), 876-81.

13. Flack, H. D.; Bernardinelli, G., Reporting and evaluating absolute-structure and absolute-configuration determinations. *J. Appl. Cryst.* **2000,** *33* (4), 1143-1148.
14. Flack, H. D.; Shmueli, U., The mean-square Friedel intensity difference in P1 with a centrosymmetric substructure. *Acta Cryst.* **2007,** *A63* (3), 257-265.
15. Flack, H. D.; Bernardinelli, G., Applications and properties of the Bijvoet intensity ratio. *Acta Cryst.* **2008,** *A64* (4), 484-493.
16. Parsons, S.; Flack, H. D.; Wagner, T., Use of intensity quotients and differences in absolute structure refinement. *Acta Crystallographica Section B* **2013,** *69* (3), 249-259.
17. Thompson, A. L.; Watkin, D. J., CRYSTALS enhancements: Absolute structure determination. *J. Appl. Cryst.* **2011,** *44* (5), 1017-1022.
18. Sheldrick, G. M., Crystal structure refinement with SHELXL. *Acta Crystallographica Section C* **2015,** *71* (1), 3-8.
19. Coelho, A. A.; Evans, J.; Evans, I.; Kern, A.; Parsons, S., The TOPAS symbolic computation system. *Powder Diffraction* **2011,** *26* (S1), S22-S25.
20. Watkin, D. J.; Cooper, R. I., Why Direct and Post-refinement Determinations of Absolute Structure May Give Different Results. *Acta Cryst. Sect. B* **2016**, 661-683.
21. Hooft, R. W. W.; Straver, L. H.; Spek, A. L., Determination of absolute structure using Bayesian statistics on Bijvoet differences. *J. Appl. Cryst.* **2008,** *41* (1), 96-103.
22. Spek, A., L., Structure validation in chemical crystallography. *Acta Cryst.* **2009,** *D65* (2), 148-155.
23. Betteridge, P. W.; Carruthers, J. R.; Cooper, R. I.; Prout, K.; Watkin, D. J., CRYSTALS version 12: software for guided crystal structure analysis. *J. Appl. Cryst.* **2003,** *36*, 1487.
24. Abrahams, S. C.; Keve, E. T., Normal probability plot analysis of error in measured and derived quantities and standard deviations. *Acta Cryst.* **1971,** *A27*, 157-165.
25. Hooft, R. W. W.; Straver, L. H.; Spek, A. L., Using the t-distribution to improve the absolute structure assignment with likelihood calculations. *J. Appl. Cryst.* **2010,** *43* (4), 665-668.
26. Hooft, R. W. W.; Straver, L. H.; Spek, A. L., Probability plots based on Student's t-distribution. *Acta Cryst.* **2009,** *A65*, 319-321.
27. Tranter, G., Place-holder for George Tranter's paper in this special issue. *Tetrahedron: Asymmetry* **2017,** *??* (??), ??-??
28. (a) Parsons, S.; Wagner, T.; Presly, O.; Wood, P. A.; Cooper, R. I., Applications of leverage analysis in structure refinement. *J. Appl. Cryst.* **2012,** *45* (3), 417-429; (b) Prince, E., *Mathematical Techniques in Crystallography and Materials Science*. 2 ed.; Springer: Berlin, 2004; p 224 pp.
29. Rabinovich, D.; Hope, H., Absolute configuration determination by anomalous scattering of light atoms. Have the limits been reached? *Acta Crystallographica Section A* **1980,** *36* (4), 670-678.
30. Escudero-Adan, E. C.; Benet-Buchholz, J.; Ballester, P., The use of Mo K[alpha] radiation in the assignment of the absolute configuration of light-atom molecules; the importance of high-resolution data. *Acta Crystallographica Section B* **2014,** *70* (4), 660-668.
31. Linden, A., Place holder for Tony Linden's paper in this issue. *Tetrahedron: Asymmetry* **2017,** *?* (?), ??-??