Derivation of results

Anthony Linden

Department of Chemistry, University of Zurich, Switzerland
anthony.linden@chem.uzh.ch

1 Introduction

When you have finished the refinement and validation of your structure, the result looks nice and you have a pretty picture on your computer screen, you can sit back and enjoy the moment. The structure has been determined and probably most of the hard work has been done, but really you are only half-way home. Now we have to analyse and interpret the results and answer some questions. What does the structure tell me? Why did I determine the structure in the first place, and has, or can that question be answered? Is the structure providing (interesting) information beyond my initial question and can I benefit from looking at that information? Is the structure really correct? Is it reliable? What can I deduce and what can't I deduce? How can I present the results? All of these questions fall under the topics of derivation, analysis, interpretation and presentations of results, which form the basis of this and some of the following lectures.

At the completion of the structure refinement, the crystal structure determination yields essentially the atomic positional coordinates, atomic displacement parameters (ADP's), unit cell parameters and symmetry information. These are the primary data obtained from the experimental data, which consist of diffraction intensities and their distribution in reciprocal space. The primary data are not particularly meaningful to most people; we are more interested in the information that can be derived from the primary data. The secondary data consist of parameters like bond lengths, bond angles, torsion angles, deviations from planes, angles between planes, ring puckering, and intra- and intermolecular interactions, such as hydrogen bonding. Such parameters are much more interesting and enlightening, so it is important to understand how these parameters are derived and how we can judge the relevance and significance of these parameters.

The refinement delivers not only the primary parameters, but also an estimated error for each parameter. In statistics, this estimated error is usually known as the estimated standard deviation, but in crystallography, we refer to these terms as standard uncertainties (s.u.). The presence of s.u.s enables us to assess the reliability and significance of any parameter or the differences between parameters. Therefore, it is important that any calculation carries the s.u.s through to the secondary data.

2 Errors, accuracy, precision, statistics

2.1 Random and systematic errors

Errors affecting experimental measurements fit into two categories: *random errors* and *systematic errors*. If we measure some quantity experimentally, e.g. a bond length or a structure factor amplitude, our observation will inevitably suffer from some sort of error.

Uncertainties or random errors are introduced by random fluctuations; these usually cannot be avoided, but can be minimized by considering and conducting all aspects of the experiment carefully. Random errors are assumed to be normally distributed, so

that their mean effect is zero and therefore do not bias the parameter being determined. The true value of the parameter will be approached the more times the experiment is repeated. A random error might be in filling a 1 litre beaker to the 250 ml mark and judging the correct amount of liquid by eye. You are unlikely to get it perfect on any one occasion, but if 20 people fill the beaker once each and the liquid is collected, you will not be far away from 5.00 litres at the end and the likely deviation can be estimated statistically.

On the other hand, systematic errors cannot be treated by any general statistical theory. Systematic errors cause measurements to deviate from their true values because of some physical effect, which we might or might not be aware of. If the beaker in the above example had been calibrated incorrectly in the factory so that the 250 ml mark should have been labelled 200 ml, you will have a large systematic error in every filling of the beaker, and an even bigger error in the total accumulated volume after 20 fillings.

When measuring X-ray diffraction intensities random errors might arise from the random fluctuation of the power and flux of an X-ray source, a low-temperature device, or in the cooling cycle of a CCD chip, and systematic errors from the influences of absorption, crystal misalignment or the models and methods used in structure determination (*e.g.* incorrect atomic scattering factors, inappropriate ADPs, wrong space group symmetry).

2.2 Precision and accuracy

It is important to distinguish between *precision* and *accuracy*. Precision concerns the *reproducibility* of a measurement. If you repeat it several times, what is the spread around the mean value? The reproducibility is indicated by the parameter's s.u. In contrast, accuracy indicates how well the measurement agrees with the true value of the parameter we are interested in. The incorrectly marked beaker has awful accuracy, even though we may fill it repeatedly with great precision.

Random errors affect the precision, but usually not the accuracy of a measurement. Systematic errors might or might not affect precision, but will almost certainly affect accuracy. Thus, reducing random errors will give smaller s.u.s in the final results and reducing systematic errors will give final results which are closer to their 'true' values. High precision alone is not necessarily an indication of a good (correct) result. The final results are affected not only by the data themselves, but also by how these data are used. The choice of least-squares refinement parameters, the method of refinement, and the weighting scheme all have an effect on the geometry we are trying to determine.

Random errors are treated by statistics and the precision is measured by s.u.s. Systematic errors cannot be treated by such a general theory, and each source of error must be identified and its effect modelled by consideration of its physical nature, or better yet, avoided in the first place. Therefore, as a crystallographer, one needs to have an understanding of all the potential sources of systematic errors in a crystal structure determination, how to detect them, remove or minimise them, and/or correct for them. Sources of errors, particularly systematic errors, can be absorption, extinction, thermal diffuse scattering, poor instrument alignment, resonant scattering (refinement of the wrong absolute structure), constraints and restraints (if inappropriate), atomic scattering factors (incorrect element assignment), incorrect or pseudo-symmetry, untreated disorder, or just plain wrong structures (*e.g.* von Schnering & Vu, 1983).

In crystallography we quote the s.u. in parentheses, for example 1.520(4) Å for a bond length. The figure in parentheses refers to the last quoted decimal place, and in this example the s.u on our measurement of 1.540 Å is 0.004 Å; a measurement of 1.54(4) Å is

ten times less precise. Usually we do not round up until the value in parentheses exceeds 19; i.e. we use 1.541(17), but 1.541(22) would be written as 1.54(2).
Statistics find application throughout data collection, data reduction, structure analysis and the derivation of results, and are intricately involved in the treatment of random errors: Prince (1982, 1994), Daly *et al*. (1995), Cowan (1998), Hamilton (1964).

2.3 Probability distributions

A frequency distribution or population is a set of observations or a function which describes the frequency with which different values are found for some measurable quantity, or the probability of finding these values for the quantity. Examples might be the number or people belonging to certain height ranges, the frequency of a certain bond length for a particular type of C–C bond, the variation of $(|Fo|^2-|Fc|^2)/\sigma(|Fc|^2)$ taken from the refinement of the crystal structure, *etc. (Fig. 1).* There are several main types of distributions: Binomial, Poisson and Normal (Gaussian) distributions. While the Poisson distribution is important for counting statistics of X-rays, the Normal distribution is the most important one in science and, in our derivation of crystallographic results, it is the one that is applicable when considering errors and uncertainties.

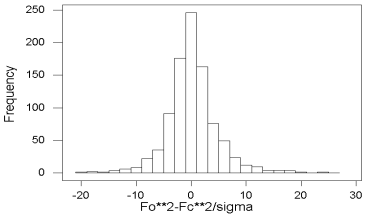

Figure 1. The variation of $(|Fo|^2-|Fc|^2)/\sigma(|Fc|^2)$ taken from the refinement of a crystal structure.

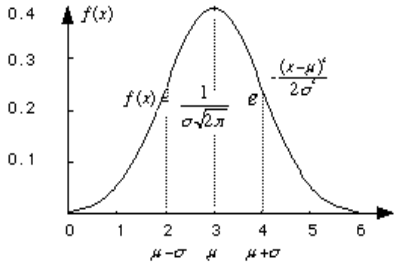

Figure 2. The normal distribution calculated with a mean of 3 and a standard deviation of 1. The characteristics of a normal distribution are that 68.3% of a population lies within ±1σ of the mean, 95.4% lies within ±2σ and 99.7% lies within ±3σ.

The normal distribution can be considered as a limiting case of the binomial distribution. The binomial distribution is described by N trials, each with a probability p, but in the case of a range of values of $(|Fo|^2-|Fc|^2)/\sigma(|Fc|^2)$, or bond lengths, what should n be? Values of the mean and standard deviation are readily obtained from a data set, however, and the normal distribution is characterised by these quantities. In any value where there can be measurement error, like a weight or a pH, you can make a histogram of the results of your trials. If you do enough trials, you will start to see a normal distribution, which has a symmetrical bell shape. If you measure many similar things, which have differences caused by a random variation, those results will be in a normal distribution. The main characteristics of a normal distribution are shown in Fig. 2.

The normal distribution is implicitly assumed in routine crystallography. The 3σ rule for testing the significance of different bond distances is based on the fact that 99.7% of a normal distribution lies within ±3σ of the mean. Therefore any value that lies more than 3σ from the mean is probably a significantly different value. Any value within 3σ is probably part of that population. Since it is widely believed that crystallographic s.u.s are

over optimistic (too small), particularly as they assume only random errors, only a difference greater than 3σ is often taken as a significant difference (*i.e.* there is an extremely low probability that they belong to the same normal distribution), while smaller differences are probably not significant.

2.4. Sampling a population

In order to estimate the mean weight of the people in a town, without having to ask every person, we might take a random sample and calculate its mean weight, but how do we know how reliable this mean is? In science, and crystallography we often need to measure a quantity and estimate the reliability of the measurement. Suppose we make N separate measurements of a parameter x. The measured values $x_1 \ldots x_N$ are a sample from all the possible measureme0nts we could make. For sufficiently large N, a consequence of the Central Limit Theorem is that the mean of our N sample values is normally distributed with the same mean as that of the total population and with variance σ^2/N where σ^2 is the variance of the parent population. It turns out that, if we have taken only one sample of measurements, which we frequently do (we usually determine a structure from only one crystal of the substance), then the best estimate we have for the population mean is our own sample mean $\bar{x}$, and the best estimate of the population variance σ^2 is related to the variance of our own sample, s^2, by $\sigma = \frac{N}{N-1}s^2$, where $s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$.
The estimate of the variance of the mean $\bar{x}$ of the sample is thus
$\sigma^2(\bar{x}) = \frac{1}{N(N-1)}\sum_{i=1}^{N}(x_i - \bar{x})^2 = \frac{s^2}{N}$.
and this is a measure of the reliability or confidence with which we can use the sample mean as a true mean of the population, *i.e.* as the true value of the quantity we are trying to measure. Note that the variance of the mean depends on the spread of individual sample points x_i and inversely on the size of the sample N. Thus, repeating the measurement of a quantity increases its precision.

2.5. Standard uncertainties in crystallographic results

In crystallography, we do not usually determine a structure many times in order to obtain mean values of the atomic parameters and (good) estimates of the variances of these parameters. Yet from a *single experiment* we obtain both parameters and their s.u.s (Schwarzenbach *et al.*, 1989). This is possible because our experimentally measured data (diffraction intensities) greatly outnumber the parameters to be derived. The problem is thus over-determined and in such cases the *estimated* s.u.s. can be obtained from a single set of data. The s.u. is a measure of the precision or statistical reliability of this value; it is our best estimate of the variation we might find for this parameter if the whole experiment was repeated many times. Consequently it is unwise to report a parameter without its s.u.

In a structure determination, the quantity minimized in the least-squares refinement is $\sum_{i=1}^{N} w_i \Delta_i^2$ where Δ_i is usually either $|F_O|_i - |F|_i$ or $(|F_O|^2)_i - (|F_C|^2)_i$ for each of the N reflections with weight w_i. The s.u.s of the refined parameters depend on the minimized function, the number of data, N, the number of refined parameters, P, and on

the diagonal elements of the inverse least squares matrix, $\mathbf{A}^{-1}$, where p_j is the jth of the P parameters:

$$\sigma(p_j) = \left((\mathbf{A}^{-1})_{jj} \sum_{i=1}^{N} w_i \Delta_i^{\,2} / (N - P) \right)^{1/2}$$

Don't worry too much about this expression. The key point is that low s.u. values (high precision) result from a combination of good agreement between the observed and calculated data (good quality data plus a well developed model), which yields a small numerator in the above equation and a large excess of data over parameters (large denominator). Even the best crystal will not give precise results if a sufficiently extensive data set has not been collected. Conversely, extensive data are not much help if the crystal was poor or the model is inadequately developed (*e.g.* untreated disorder).

3 Geometric parameters

As already stated, it is important as far as possible to report s.u.s with any geometric parameter derived from the primary data. Therefore, any file with atomic coordinates, but no s.u.s (SHELXL .res file) is unsuitable for calculating geometric parameters properly. The molecular geometry depends not only on the atomic coordinates, but also on the unit cell parameters, and they too have s.u.s. Thus, the calculation of the s.u.s of any derived geometric parameters must include an allowance for the s.u.s of the unit cell parameters; usually taken care of in commonly used software. Ideally, use the refinement program itself to do the calculations, if possible, or use the CIF as input to, for example, PLATON (Spek, 2009, 2015). MERCURY (Macrae et al., 2012) does not generate s.u.s on calculated parameters. Keep in mind the need to use the variance-covariance matrix whenever possible in order to properly estimate the s.u.s. This matrix is only available during a refinement run – *i.e.* bond lengths and angles in a CIF created by SHELXL are calculated using the variance-covariance matrix, but parameters in a CIF or other output created by PLATON using only the s.u.s on the atomic coordinates and unit cell parameters are not; in the latter case, the s.u.s may be underestimated. The effect is usually small, but might be significant when symmetry related atoms, atoms on special positions, or atoms involved in restraints are involved, especially for H-atoms.

PLATON will calculate just about any geometric parameter you need, mostly with the inbuilt defaults (use CALC ALL or any of the CALC XXX buttons), but the defaults can be changed by using command line entries, such as '*calc coord 4.2*', to calculate the geometry out to 4.2 Å about any hetero or metal atom (default distance is 3.6 Å). PLATON can also do many calculations by using the various menu options in the ORTEP or PLUTON view screens and clicking on the desired atoms.

The CALC XXX instructions in PLATON generate a comprehensive listing in the .lis file, including relevant useful references, and you should familiarize yourself with the contents of the file and the sequence of information when CALC ALL is used.

A frequent question when interpreting a crystal structure is whether a particular bond length or angle differs significantly from another bond length or angle, or from some standard value. Unless we make some assumptions about uncertainties, we cannot deduce anything useful at all. The basic assumptions we make are (i) that each of our derived results is an *unbiased* estimate of its true value (*i.e.* we assume that we have an *accurate* result, either free from any significant effects of systematic errors, or corrected for them),

and (ii) that the s.u. is a true estimate of the precision of the result and a measure of the variation we would expect to find if we determined the structure many times; such a variation is expected to follow a normal distribution.

The above assumptions, together with a knowledge of the properties of a normal distribution, allow us to work out the significance or otherwise of the difference between two parameters. To compare two bond lengths (or any other parameters), we imagine that the two values are equal and examine, statistically, if the observed difference between the bond lengths is likely. If the bond lengths were equal, we could measure them many times and the differences between these lengths would be subject to random error. We would obtain a set of observed differences normally distributed about a mean of zero with some standard uncertainty. In practice, we only have *one* measurement for the difference, together with an estimate of the standard uncertainty, obtained by combining the s.u.s of the two bond lengths. The question now becomes, what is the probability that the observed difference could be as far as it is from the true value of zero? As stated earlier, the characteristics of a normal distribution apply and we can use the 3σ rule.

How do we obtain the s.u. of a difference between two bond lengths A and B (or any other parameters)? We usually only have the variances to hand (square of the s.u.s) and no covariances are available, which is equivalent to assuming zero correlation among the parameters, *i.e.* that they are completely independent. This becomes simply $\sigma^2(A - B) = \sigma^2(A) + \sigma^2(B)$ and this is the form generally used. As an example, consider two bonds of length 1.540(3) and 1.570(4) Å. Are they significantly different? The observed difference is 0.030 Å and the s.u. of the difference is $(0.003^2 + 0.004^2)^{1/2} = 0.005$ Å. Therefore, the difference is 6 and this would certainly appear to be significant. Conversely, bond lengths of 1.540(8) and 1.570(9) Å are not significantly different and any attempts to discuss them as being different in length are unfounded. These examples clearly show the importance of reducing random errors, as well as eliminating systematic errors, as far as possible.

The comparison of a particular bond length with some 'standard value' is similar, except that deriving the s.u. for the standard value may be difficult and is sometimes assumed to have a zero s.u, so the s.u. of the difference between the observed and standard value is just the s.u. of the observed value itself. Note, however, that it would be folly to use this argument to compare a bond length from your structure with a single (and therefore clearly not standard) bond length taken from the literature, where the latter had been reported without any mention of its s.u.

One way of comparing two structures, or the conformations of two molecules in the same or different structures, is to attempt to superimpose the two molecules or the packing thereof as closely as possible on one other. MERCURY can fit molecules from within the same structure or different structures. PLATON can fit symmetry-independent molecules in the same structure. Try AUTO-MOL-FIT or '*fit 1 2*' on the command line (fit 1 3 or fit 2 3 also works if there are three independent molecules) or using the 'fit-by-click' option (under the third option menu of ORTEP). The footnote on the graphic indicates whether or not one molecule needed to be inverted to obtain the fit and you should consider what this implies about the independent molecules given the space group symmetry.

4 Thermal and rigid body motion

Although the major interest in a structure determination usually centres on the geometry derived from the atomic positions, the primary results also include the atomic displacement parameters (ADPs) or 'thermal parameters'. These can describe not only the

time-averaged temperature-dependent movement of the atoms about their mean equilibrium positions (*dynamic disorder*), but also their random distribution over different sets of equilibrium positions from one unit cell to another, thereby representing a deviation from perfect periodicity in the crystal (*static disorder*) which is not great enough to be resolved into distinct alternative atom sites. See Dunitz *et al*. (1988).

The interpretation and analysis of ADPs is not often undertaken. One reason is that various systematic errors in the data (*e.g.* absorption), inappropriate refinement weights and poor aspects of the structural model used in the refinement all tend to affect the ADPs far more than the atomic positions (fortunately!). Thus the ADPs of a structure are often regarded as a sort of 'error dustbin' capable of mopping up all sorts of evils, and their physical significance is questionable unless the experimental work is of high quality.

One effect of thermal vibration is to produce an apparent shrinkage in molecular dimensions. Analysis of this effect and correction for it are possible only in certain cases. If a molecule has only small internal vibrations, both bond stretching and angle deformations, compared with its movement as a whole about its mean position in a crystal structure, then it can be treated approximately as a rigid body. In this case, the movements of the individual atoms are not independent and so the U^{ij} parameters of the atoms must be consistent with the overall molecular motion. PLATON can perform a TLS (torsion, libration, screw) analysis and lists corrections for bond lengths within the molecule; these depend only on the librational tensor components. Many molecules cannot be regarded as even approximately rigid, it may be possible to treat certain groups of atoms within the molecule as rigid bodies and make corrections within those groups. One can test whether a molecule, or part of a molecule, might be regarded as a rigid body (Hirshfeld, 1976).

ADPs of the atoms are strongly temperature dependent and they can be reduced significantly by carrying out data collection at a lower temperature; highly recommended for all work. Low temperature data usually give greater precision in atomic positions, more reliable molecular geometry, and an opportunity to assess and distinguish between dynamic and static disorder: the former, but probably not the latter, will be reduced at lower temperature. In addition, low temperature enhances the diffracted intensities (because the reduction in thermal motion means that atoms act more like point sources), particularly at higher angles, so the benefits of low temperature work, particularly for weakly diffracting samples should not be underestimated.

5 References

Cowan, G. (1998). *Statistical Data Analysis.* Oxford: Oxford University Press.

Daly, F., Hand, D. J., Jones, M. C., Lunn, A. D. &. McConway, K. J. (1995). *Elements of Statistics.* Pearson Education.

Dunitz, J. D., Maverick, E. F. & Trueblood, K. N. (1988). *Angew. Chem. Int. Ed. Engl.* **27**, 880-895.

Hamilton, W. C. (1964). *Statistics in Physical Science*. New York: Ronald Press.

Hirshfeld, F. L. (1976). *Acta Cryst.* A**32**, 239-244.

Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., Pidcock, E., Rodriguez-Monge, L., Taylor, R., van de Streek, J. & Wood, P. A. (2008). J. Appl. Cryst. 41, 466–470.

Prince, E. (1982). *Mathematical Techniques in Crystallography and Materials Science*. New York: Springer.

Prince, E. (1994). *Mathematical Techniques in Crystallography and Materials Science*, 2nd Ed. New York: Springer.

Schnering, H. G. von & Vu, D. (1983). *Angew. Chem. Int. Ed. Engl.* **22**, 408.

Schwarzenbach, D., Abrahams, S. C., Flack, H. D., Gonschorek, W., Hahn, Th., Huml, K., Marsh, R. E., Prince, E., Robertson, B. E., Rollett J. S. & Wilson, A. J. C. (1989). *Acta Cryst.* A**45**, 63-75.
Spek, A. L. (2009). *Acta Cryst*. D**65**, 148–155; http://www.cryst.chem.uu.nl/spek/platon/
Spek, A. L. (2015). *Acta Cryst*. C**71**, 9-18.
Taylor, R. & Kennard, O. (1983). *Acta Cryst.* B**39**, 517-525.