

# Why do We Trust X-ray Crystallography?

*Andrew D Bond*

All chemists know that X-ray crystallography is the ‘gold standard’ characterisation technique: an X-ray crystal structure provides definitive proof of structure for chemical compounds. But, why do we trust X-ray crystallography in this way? This article highlights the underlying reasons: there is simply so much information in a single-crystal X-ray diffraction pattern, and, we usually know a sensible chemical structure when we see one.

## Introduction

Single-crystal X-ray diffraction is widely accepted as the most powerful characterisation technique available to chemists. For a suitable crystalline sample, the analysis provides a complete three-dimensional picture of a molecule and the geometrical features of its interactions with other molecules in the solid state. A crystal structure is resolved at the atomic level, meaning that, we can actually see the arrangement of atoms in molecules. This is astonishing when you consider that the distance between two C atoms in a C–C bond (for example) is around  $1.5 \times 10^{-10}$  m (= 1.5 Å). Any chemical sample (solid, liquid, gas) will produce a diffraction pattern. The importance of single crystals is that they are periodic in three dimensions, which leads to a regular diffraction pattern that can be interpreted in a straightforward and robust way.

## How is a Crystal Related to its Diffraction Pattern?

An X-ray diffraction pattern from a single crystal can conceptually be partitioned into the geometry of the pattern and the intensities of the diffracted beams (*Figure 1*). The geometry reflects the translational periodicity of the crystal lattice. The angles measured for a set of diffracted beams can be back-converted to yield dimensions (and orientation) of the periodic



Andrew D Bond is at the Department of Pharmacy in the University of Copenhagen, Denmark. He is a chemical crystallographer, with research interests in structures and properties of pharmaceutical solids.

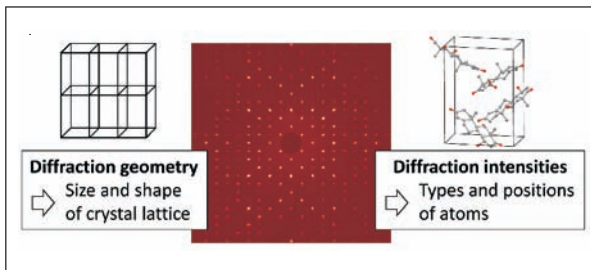
A crystal structure is resolved at the atomic level, so we can actually see the arrangement of atoms in molecules.

## Keywords

Chemical crystallography, X-ray diffraction, single crystals, structure determination.



**Figure 1.** Conceptual partitioning of an X-ray diffraction pattern according to the information that it contains.

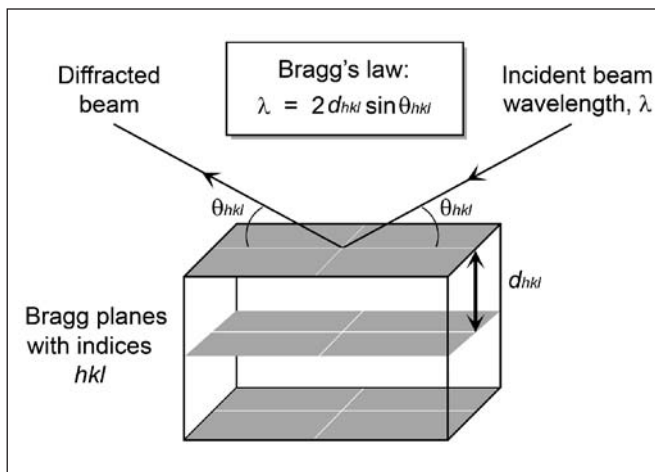


The geometry of a diffraction pattern reflects the translational periodicity of a crystal lattice. The intensities of the diffracted beams reveal the positions and types of the atoms.

lattice that produced them. The conceptual link is the famous Bragg's law, illustrated in *Figure 2*. The lattice points in a crystal are positions that are equivalent by *translation*. The great simplifying feature is that the entire crystal structure can be built up by replication of a small building block – the unit cell – according to these lattice translations. This enables the unfathomably large number of atoms in a crystal to be represented by only a relatively small number of atoms within the specified unit cell.

The positions and types of atoms within the crystal are revealed by the intensities of the diffracted beams. The underlying relationship is expressed by the structure factor equation, which is a mathematical description of the interference between the scattered X-rays:

$$F_{hkl} = \sum_{n=1}^N f_n \exp 2\pi i (hx_n + ky_n + lz_n)$$



**Figure 2.** Representation of the geometry of diffraction according to Bragg's law.



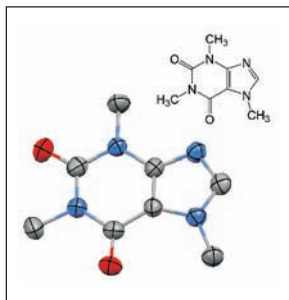
The equation is surprisingly simple (although it may not immediately look like it!). On the right-hand side, the symbol  $f_n$  represents the scattering power of the  $n$ th atom within the unit cell. It describes the atom's chemical type, and it is essentially proportional to the number of electrons. The atom has coordinates  $x_n$ ,  $y_n$ ,  $z_n$ , expressed as fractions along the edges of the specified unit cell. The symbols  $hkl$  describe the orientation of the Bragg planes (Figure 2), and each set of  $hkl$  values identifies a specific diffracted beam. The symbol  $i$  refers to the imaginary number  $\sqrt{-1}$ , which is used in mathematics to express a complex number. The summation is taken over all  $N$  atoms in the unit cell. The quantity on the left-hand side,  $F_{hkl}$ , is the structure factor. It describes the amplitude of the diffracted beam and its phase relationship to the other diffracted beams. The *intensity* of each diffracted beam, which is a quantity we can measure, is given by the square of the structure factor's amplitude.

The mathematical details of the structure factor equation are not the main issue here. The important point is how the various quantities are linked together. The structure factor equation provides an intensity value for each diffracted beam in the diffraction pattern from a set of parameters describing the crystal structure. The equation produces a value for one specific diffracted beam, and a whole set of structure factor equations is required to describe an entire diffraction pattern. To calculate each structure factor involves a summation over all atoms in the unit cell, so the position and type of every atom influences every diffracted beam. Conversely, it can be said that every diffracted beam contains information about the type and position of every atom. This is different from a spectroscopic technique, where a chemical group might produce some specific signal, and this is where X-ray crystallography gains its advantage: a typical diffraction pattern contains *thousands* of measurable diffracted intensities, while it requires only *hundreds* of parameters to describe a typical crystal structure. This gives a *data-to-parameter ratio* far in excess of any other analytical technique. Any scientist knows that multiple observations of the same parameter leads to more confident conclusions.

The key advantage of X-ray crystallography is its massive data-to-parameter ratio.

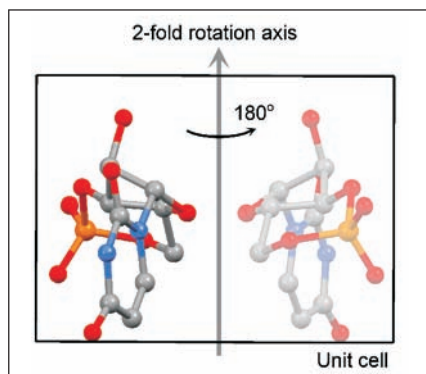
Any scientist knows that multiple observations of the same parameter leads to more confident conclusions.





**Figure 3.** A molecule of caffeine as it is commonly represented in an X-ray crystal structure. The ellipsoids representing each atom give an indication of the atom's displacement. H atoms are not shown.

**Figure 4.** Illustration of a crystal structure containing 2-fold rotation symmetry. The coordinates of the shaded atoms have a defined relationship to the coordinates of the non-shaded atoms. Therefore, parameters need to be defined only for the non-shaded atoms.



## How Many Parameters Describe a Crystal Structure?

Six parameters are required to describe a crystal's unit cell: the lengths of the three edges and the three angles between them. *Figure 1* reminds us that these parameters are obtained from the geometry of the diffraction pattern, so they are not directly involved in the fit to the measured intensities. To describe the contents of the unit cell, we need three parameters ( $x_n, y_n, z_n$ ) for the coordinates of each atom. In real life, atoms are subject to thermal vibration, which is also described by including an extra term in the structure factor equation. For each non-H atom, this usually requires six parameters, describing the size and orientation of an ellipsoid (*Figure 3*). So we require nine parameters for each non-H atom in the unit cell. H atoms have a special status, because they contain only one electron and therefore they scatter very weakly. Often, they are just placed in chemically reasonable positions with no parameters to be refined. The chemical type of each atom is expressed by  $f_n$  in the structure factor equation. This is selected from only a limited number of possibilities. It is something that can be changed, but it is not freely variable during optimisation of the crystal structure. Roughly then, we estimate  $9N$  parameters for a unit cell containing  $N$  atoms.

We have one more trick up our sleeve – crystals are highly symmetrical objects. We have already used the crystal's *translational symmetry* to define the unit cell, and we can also exploit symmetry that typically exists for the atomic positions. If a group of atoms in the unit cell are related to each other by some specified symmetry relationship, we only have to describe parameters for one atom in the group. If there is a 2-fold rotation, for example, the number of required parameters is immediately halved (*Figure 4*). The collection of symmetry relationships within a crystal is called its *space group*. A common arrangement has four molecules in the unit cell, related by symmetry so that it is only necessary to describe parameters for one of them. In this case, a chemical compound containing 10 non-H



atoms in the molecule requires fewer than 100 parameters to describe its entire crystal structure.

### How do We Decide if a Crystal Structure is Correct?

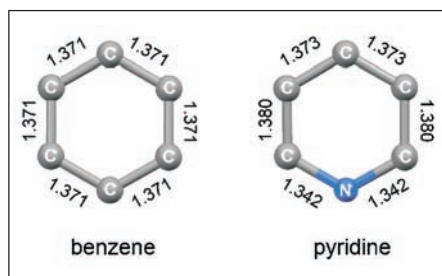
The aim of single-crystal X-ray diffraction is to produce a crystal structure that fits the measured diffraction data. The agreement between the structure and the diffraction pattern is expressed numerically by comparing the measured intensity values with those calculated using the parameters of the structure model. The overall agreement is quantified essentially as the average percentage difference between the calculated and measured values. The parameters describing the crystal structure are optimised by an iterative least-squares process to provide the best fit to all measured intensities. A typical value for the difference between calculated and measured intensities is less than 5%.

Numerical agreement is one thing. But there is another thing that chemists generally take for granted: in a crystal structure, we can *see* the molecules. One of the main reasons to trust the results of X-ray diffraction is that the results look right! Rather, it is quite easy to see if things are going wrong. For example, consider what might happen if we assign an incorrect atom type. Say a crystal contains benzene, but we mistakenly assign one atom as N so that the molecule appears to be pyridine (*Figure 5*). In the structure factor equation, the value of  $f_n$  is changed from 6 (for C) to 7 (for N) for one atom out of six. This may not have too much influence on the numerical agreement – it should get a little worse, but the model will still fit the data quite well. However, we should be worried when we interpret the result as a chemist. We know that the C–N bonds in pyridine should be shorter than the C–C bonds.

If we produce a crystal structure apparently containing a pyridine molecule where all the bonds around the ring have the same length, we should recognise our error. The point here is that we know so much about what molecules *should* look like, that we usually know a trustworthy result when we see one. If our refined parameters provide an acceptable numerical fit to the

Another good reason to trust X-ray crystallography is that the results usually look right!

**Figure 5.** Illustration of the bond lengths expected for benzene and pyridine molecules. The values are averages taken from a large number of established crystal structures. H atoms are not included.



It is possible to make mistakes, but there is no doubt that X-ray crystallography deserves its position as a chemist's most trusted analytical technique.

information-rich X-ray data set, *and* the resulting crystal structure is consistent with well-considered chemical expectations, we can trust the result.

### Is it Possible to Make a Mistake?

Yes! One problem is that experimental measurements are always affected by errors. So the clear picture that has been described above can become much cloudier in the real world. If the measured intensity values are statistically less reliable, the parameters that we fit to them become less precise. In the case of C–N vs C–C bonds (*Figure 5*), the expected length difference is something like 0.03 Å (= 3 pm). It could be that the uncertainty associated with our assessment of the bond distance is larger than the difference that we are looking for. In this case, we cannot use the bond lengths to tell us whether a molecule looks more like pyridine or benzene. With experience, we might be able to invoke other features of the structure to make a conclusion – specifically, benzene and pyridine are likely to interact differently with neighbouring molecules. But experimental errors and uncertainties can always lead to mistakes.

There *are* various pitfalls that can trap an inattentive chemical crystallographer. But many of them are so catastrophic, that the result should be obvious nonsense. For example, choosing an incorrect unit cell or space group will probably yield a structure that is entirely unrecognisable to a chemist. Some errors are more subtle, such as choosing a wrong atom type or missing some H atoms. In these cases, the established structure may not be so far from the truth, but of course the devil is in the details. Ultimately, it is the *combination* of the massive information content in a single-crystal X-ray diffraction pattern and good chemical sense that elevates X-ray crystallography to its position as the most trusted analytical technique.

### Suggested Reading

- [1] William Clegg, *Crystal Structure Determination*, Oxford Chemistry Primers No.60, Oxford University Press, New York.

#### Address for Correspondence

Andrew D Bond  
University of Copenhagen  
Department of Pharmacy  
Universitetsparken 2  
2100 Copenhagen  
Denmark.  
Email:  
andrew.bond@sund.ku.dk

