Progress in powder and surface X-ray diffraction crystallography

Jordi Rius*, Xavier Torrelles and Carles Miravitlles
Departament de Cristal·lografia i Química de l’Estat Sòlid. Institut de Ciència de Materials de Barcelona (CSIC), Bellaterra

Abstract

One of the most important scientific achievements of the XX century has been without any doubt the discovery of X-ray diffraction by crystalline matter and the subsequent use of this process to determine the crystal structures of the materials. This field matured at the end of the century with the almost automated solution of the crystal structures of small compounds, allowing various new applications of X-ray diffraction methods to emerge. Among the most relevant ones for Materials Sciences one may mention the solution of crystal structures from X-ray powder diffraction intensity data or the determination of atomic arrangements in the surfaces of substrates and, eventually, of the corresponding deposits. Since these advances are closely related to a recent reinterpretation of the traditional direct methods of solving crystal structures, a short introduction to them will be given first. Then we deal with the modification of such methods to cope with the particularities of these two applications and, finally, some examples will be described.

Keywords: Powder diffraction, surface crystallography, atomic structure solution, direct methods

Nowadays, the solution of crystal structures from X-ray diffraction data is a well standardized technique with important implications in most fields of science as, for example, in Materials Sciences where it provides the basic features of the materials, thus allowing better description and frequently an improvement in the knowledge of their properties. In Biochemistry it provides detailed knowledge of compounds like proteins, nucleotides, and enzymes etc. In the specific case of Biochemistry, the principal difficulty in solving atomic structures is the problem of huge unit cells containing thousands of atoms. In Materials Sciences, however, the compounds possess smaller unit cells but, since they are normally polycrystalline, the diffraction experiment gives only incomplete intensity data sets.

In Crystallography, the term direct methods denotes the methods capable of solving the atomic arrangement in a crystal structure directly from the measured intensities of the reflections using as additional information the approximate unit cell atomic content. Until the last decade most of the advances in the field of direct methods concentrated on the...
development of the methods themselves, i.e. the intensity data used for these developments corresponded to ideal single crystal diffraction experiments. More recently, however, developments like the direct methods modulus sum function have helped to extend the applicability of such methods to less favourable situations like powder diffractionmetry or the crystallography of reconstructed surfaces. The following section gives a brief introduction to these powerful methods.

**Modern direct methods**

Direct methods take advantage of two general properties of the periodic electron density distribution $\rho$ in the crystal, namely its positivity and its peakness (atomicity). Consequently, as the first requirement for their application, direct methods need intensity data up to atomic resolution. Should this not be available, supplementary information is necessary (for example, information about the internal geometry of the molecular fragment).

As is well-known, the electron density in a point $\mathbf{r}$ of the crystal can be expressed in the form of a Fourier series in which the coefficient $F(\mathbf{h})$ is called the structure factor of reflection $\mathbf{h}$:

$$\rho(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{h}} |F(\mathbf{h})| \exp(-i2\pi\mathbf{hr}) \quad (1)$$

The structure factor is a complex quantity, $F(\mathbf{h}) = |F(\mathbf{h})| \exp(i\phi(\mathbf{h}))$. Its modulus $|F(\mathbf{h})|$ is related to the measured intensity and can be easily derived from it. If both modulus and phase value $\phi(\mathbf{h})$ were known (at least for the strongest reflections), it would be quite straightforward to calculate the electron density distribution from the above expression. Unfortunately, the phase information is lost during a conventional diffraction experiment. This constitutes the so-called ‘phase problem’ in Crystallography. In other words Direct methods are necessary to find the values of the collectivity $\Phi$ of phases of the large structure factors. The general strategy for finding the values of $\Phi$ consists of giving initially random values to the collectivity and then introducing them into a suitable phase refinement function which incorporates or makes use of all the a priori known information (positivity, atomicity and structure factor moduli derived from measurement). After some refinement cycles, the values of $\Phi$ maximizing this function are assumed to be the true ones and are used to compute the Fourier map which will show the atomic peaks. To be sure that the correct solution is found, multiple phase refinement trials with different sets of initial phase values are performed. The solutions of these trials are ranked according to the final value of the phase refinement function (see Table 1).

The critical point, however, is how to combine all the a priori known information to give an effective and robust phase refinement function. Experience has shown that this can be achieved by introducing the ‘squared structure’ concept. In analogy to the true electron density distribution $\rho$, the squared distribution $\rho^2$ can also be written in the form of a Fourier series. If the corresponding Fourier coefficients are denoted by $G(\mathbf{h}) = |G(\mathbf{h})| \exp[i\psi(\mathbf{h})]$, then the modulus $|G(\mathbf{h})|$ can be computed in terms of $\Phi$ by means of the expression

$$|G(\mathbf{h}, \Phi)| = K \sum_{\mathbf{h}'} |F(\mathbf{h}')| |F(\mathbf{h} - \mathbf{h}')| \cos[\Phi(-\mathbf{h}) + \psi(\mathbf{h}') + \psi(\mathbf{h} - \mathbf{h}')] \quad (2)$$

which uses the experimentally known large moduli of reflections $\mathbf{h}'$ and $\mathbf{h} - \mathbf{h}'$ in an active way. The atomicity condition and the positivity criterion are incorporated by forcing $\rho$ and $\rho^2$ to be similar, or in other words, by making, for the strong

---

**Table 1. Schematic description of the crystal structure determination process as implemented in the 2000 version of the direct methods program XLENS [25].** Firstly, NSET phase refinement trials are performed by assigning initially random values to the collectivity $\Phi$ of phases of the structure factors of the large reflections. The phases are then refined by maximizing the direct methods modulus sum function $S(\Phi)$ [21]. In general, a value NSET=100 suffices for solving small structures (up to approx. 200 atoms).
reflections, the phases values $\phi(h)$ and $\phi(h)$ equal and the values $i\mathbf{G}(h)$ and $i\mathbf{G}(h)$ proportional.

The knowledge that the structure is made of atoms can also be used to decompose $i\mathbf{G}(h)$ as a function of the interatomic vectors $\mathbf{r}_i$. In this way, all the zero vectors (those of type $j = k$) can be grouped into a single term $<i\mathbf{G}>$. For an equal atom structure belonging to space group $P1$ with $N$ atoms in the unit cell and with corrected scattering factor dependence, it holds that

$$i\mathbf{G}(h) = K' \sum_{j<k} \cos 2\pi \mathbf{r}_j \cdot \mathbf{r}_k$$

For structure factors with large moduli $|i\mathbf{G}(h)|$, the double sum of cosinus cosine terms will be a large positive quantity while for small structure factors the corresponding sum will be large and negative. If only these two subsets of reflections are considered one can define function

$$Q(\Phi) = \sum_{h} \left[ |G(h,\Phi)| - <G> \right]$$

which has a large positive maximum for the correct phase values. $Q(\Phi)$ is part of the more powerful phase refinement function called modulus sum function $S(\Phi)$ which results from combining $Q(\Phi)$ with the residual $R(\Phi)$.

$$R(\Phi) = \sum_{h} \left[ G(h) - |G(h)| \right]$$

Hence, the final expression for $S(\Phi)$ is

$$S(\Phi) = Q(\Phi) - R(\Phi) = 2 \sum_{h} \left[ |G(h)| - <G> \right] |G(h)|$$

$S(\Phi)$ is extremely effective and powerful and, for small structures, the top ranked solutions are almost always the correct ones. These phase values are then used to compute the final Fourier map. Physically, maximization of $S(\Phi)$ is approximately equivalent to maximization of the integral

$$\int P'(\mathbf{u}) \cdot P(\mathbf{u},\Phi) \cdot dV$$

with $P'$ being an observed Patterson-type function with removed origin peak and $R(\Phi)$ the corresponding calculated one.

**Structure solution from powder data**

The number of structures solved from powder data has increased spectacularly recently. This increase is a direct consequence of the availability of more complete and reliable sets of extracted intensity data, which is largely due to (i) the use of strictly monochromatic radiation, coming either from synchrotron sources or from laboratory diffractometers equipped with curved incident-beam monochromators, (ii) the development of two-dimensional CCD detectors and image-plates, and (iii) the extraction of integrated intensities with whole-pattern profile-fitting procedures which incorporate more accurate corrections and peak profile descriptions [11], [29]. In spite of this considerable progress, there are some limiting factors which are either inherent to the crystal structure, e.g. unit cell size and crystal symmetry, or difficult to control, e.g. sample quality. In general, the best patterns are those obtained from highly crystalline materials measured with synchrotron radiation. Due to the narrow peak widths, accidental peak overlap is then minimized and, for relatively small structures, the intensities of most reflections up to atomic resolution can be extracted. In practice, however, the crystallinity of most powder samples for which no sufficiently large single crystals can be grown is rather poor. Consequently, this ideal situation is often not attained, and the incidence of accidental overlap rapidly increases with the Bragg angle.

In particular, for structures in which the Laue group is a subgroup of the lattice group, systematic overlap is also present. This means that the $d$-spacings of multiple symmetry-independent reflections coincide, and, therefore, it is not longer possible to estimate individual integrated intensities. In any case, one «cold comfort» of the presence of systematic overlap is that the amount of accidentally overlapping multiplets is reduced. Recently, [15] and [26] have demonstrated by solving the crystal structure of the dominant crystalline phase of the high alumina cement that the direct methods modulus sum function can be successfully employed with intensity data obtained by equipping the global intensity of each multiplet (Fig. 1). This result will have important implications not only for powder diffraction but also for other techniques with similar ‘twinning’ problems.

**Fig. 1. Perspective view along the c axis of the crystal structure of the dominant crystalline phase (CAH10 in cement notation) of the hydrated high-alumina cement [15].** The crystallographic data are $a=16.387$, $c=8.279(9)$, $Z=6$, $V=1925(1)$, space group $P6_3/m$, formula $[\text{CaAl}_2\text{(OH)}_6\text{(H}_2\text{O})] \cdot 1.84\text{H}_2\text{O}$. The powder pattern used for solving the structure was collected in reflection geometry on a high resolution powder diffractometer equipped with a rotating anode and a primary curved Ge(111) monochromator.
The first procedure for extracting integrated intensities which uses the metric as a constraint was developed in 1981 [17]. However, it is the two-stage LeBail method which is becoming more popular. Unlike Pawley’s method, here the integrated intensities and the profile parameters are refined separately. In stage one, the integrated intensities are estimated by introducing the approximate profile parameters in a recursive expression, like the following employed in the program AJUST [24]:

$$I(h)_{\text{rec}} = I(h)_{\text{meas}} + c(h) \cdot \sum w(i) \cdot \Omega(h,i) \cdot [y_c(i) - y_b(i)]$$  \hspace{1cm} (8)

where the summation extends over the i points of the pattern for which the value of the profile function $\Omega$ of reflection $h$ is not negligible, and where $y_c(i)$, $y_b(i)$, and $w(i)$ are, respectively, the net observed and calculated counts and the weight associated with the measurement at $\Theta_i$. To increase the stability of the extraction process, the additional condition $l(h)_{\text{rec}} \simeq 0$ is introduced. In stage two, the profile parameters are refined by Rietveld least-squares while keeping the values of the integrated intensities fixed. This two-stage process is repeated until convergence is achieved. In general, two reflections can be resolved if their angular distance is greater than 0.45 times the respective FWHM (full width at half maximum). The functions more commonly used to describe the peak functions are lorentzians, gaussians or the convolution of both, depending on the type of radiation, the nature of the sample and the instrumental geometry.

For patterns with slight accidental peak overlap, the simplest way to estimate the intensities of the overlapping reflections is to equidistribute the total intensity of the multiplet. More sophisticated approaches are based on the squared Patterson function or on maximum entropy methods, e.g. [4], [5], [8], [16]. All these methods have been applied with variable degrees of success and, in general, work better for high quality powder patterns.

Prior to commenting on the two principal strategies of overcoming the accidental peak overlap problem, a classification of the structure types according to the effective resolution of their powder patterns is in order. Effectively, in powder diffraction the successful application of direct methods depends on the quality of the diffraction data and on the chemical composition of the structure to be solved. The information regarding these two points can be summarized in the following parameters:

(i) The effective resolution of the data set, which is defined as the $d$-spacing ($d_{\text{min}}$) of the outer resolution shell for which most reflections have reliable integrated intensities.

(ii) The minimum separation $t$ between dominant scatterers. The two limiting cases are the purely organic compound ($t = 1.45 \text{ Å}$) and the inorganic compounds containing heavy atoms ($t > 3 \text{ Å}$).

In general, direct methods will be successful only if the effective resolution of the intensity data satisfies the condition $d_{\text{min}} \leq 0.7 \cdot t$ [22].

Obviously, the corresponding direct methods solution will principally reveal the partial structure of the dominant scatterers. The cases represented by inorganic and zeolitic compounds are very illustrative. While for organic structures, a $d_{\text{min}}$ value of 1 Å is required, for zeolitic compounds where the SiO$_4$ units act to some extent as dominant scatterers, $d_{\text{min}}$ values in the order of 2 Å can be tolerated (the minimum separation between Si atoms is approximately 3.1 Å). In any case, the availability of extraction methods supplying intensities up to $d_{\text{min}}$ resolution is necessary for the application of direct methods. In general, the smaller the value of $d_{\text{min}}$, the better they work. However, due to the changing aspect of the Fourier maps at different resolutions, their correct interpretation is not always simple, e.g. in the case of zeolites, the SiO$_4$ unit appears as a single peak at moderate resolution. Hence, according to the previous expression, the following conclusions can be drawn:

(i) Direct methods cannot be applied to organic compounds if only laboratory X-ray powder diffraction data are available. For these compounds, use of synchrotron radiation is mandatory. An alternative solution is the use of Patterson search methods which will work well even at low resolution [20].

(ii) Inorganic compounds represent (at least theoretically) the most favourable case, since the positions of the heavy atoms can be found at relatively low resolution. For coordination compounds, the rest of the structure formed by unresolved but defined molecular fragments (e.g. C5 or C6 rings) can be modeled using the pseudo-atom approach [6].

**How to overcome the accidental peak overlap**

According to the foregoing analysis, it is evident that inorganic and organic compounds require different peak overlap treatments and hence will be handled separately:

**Inorganic compounds**

As has been shown recently, it is possible to overcome the peak overlap problem up to a $d_{\text{min}}$ resolution either by simultaneously refining the phase values and the partitioning coefficients during the structure solution process, or by using textured samples of the compound under study, thus taking advantage of the latest instrumental developments (synchrotron radiation and 2D CCD detectors). Note that both methods are complementary and quite general, i.e. they can be applied to compounds with uncertain cell contents and with large atomic disorder, a situation often found when studying complex minerals with limited crystallinity. The simultaneous refinement of phases and partition coefficients has been already applied to the solution of the mineral tinstite [27].

When peak overlap becomes more serious, i.e. when for
many peaks the shortest interpeak separation is less than 0.3-0.5 times the FWHM (full width at half-maximum), alternative strategies are required (Table 2.). The principal method is the combined minimization of powder pattern differences and potential energy by simulated annealing. Recently, [9], [19], [3] have applied related procedures to the solution and refinement of complex inorganic structures. In all these methods, the refined variables are defined in direct space, so that no extraction of integrated intensities is necessary. However, in order to be successful, they require, firstly, the unit cell contents to be accurately known and, secondly, the absence of significant disorder in the structure. Considering the steady increase in computing power, these methods are expected to play an increasingly important role in the future.

Finally, it is worth mentioning that direct methods can be also assisted by Fourier recycling procedures, including topological searches. This strategy has been specifically developed for zeolite-like compounds by [13] and [14] and has solved a number of previously unknown structures.

**Organic compounds**

For these, a different philosophy is followed. Since there is no possibility of solving the overlap problem up to \( d_{\text{min}} \) resolution and since at least part of the geometry of the molecule and its connectivity is known, the structure solution is carried out with variables defined in real space, namely the Euler angles giving the correct orientation of the molecule, the coordinates in the unit cell of the local origin of the molecule and, finally, when necessary, the refined torsion angles. Thus, the whole powder pattern can be calculated at each stage of the refinement using this limited number of variables and the overlap problem is avoided. The function to be minimized is the residual \( R(\Phi) \)

\[
R(\Phi) = \sum_h |F(h) - |F_h(\Phi)||^2
\]

with \( \Phi \) being the variables defined in real space [7].

This is a very active research field with important economic implications. These methods can be applied to the structure solution of the various polymorphs of pharmaceutical compounds and drug substances which do not grow as single crystals.

**Advances in surface X-ray diffraction crystallography**

At first glance, the surface of a crystal may appear very smooth. However, its heterogeneity becomes evident when observed in more detail with an electron microscope. The surface exhibits steps of variable height separating domains made up of atomic planes in which multiple point defects can be detected. These domains can extend over several atomic rows depending, of course, on the preparation of the sample. Fortunately, due to the large area illuminated in an X-ray diffraction experiment, most of the surface defects are averaged out and only the ideal atomic arrangement can be derived from the measured intensities. This is in contrast to electron diffraction where smaller areas are considered. The modification of the atomic geometry at the surface or interface causes their electronic properties to be different than in the bulk. Besides the electronic properties, both the atomic
structure and composition also determine to a large extent the chemical reactivity of the surface.

Normally, atoms in the vicinity of the surface undergo different types of rearrangements, the two most common ones being:

a) Surface reconstruction: If \(a\) and \(b\) denote the basis vectors of the unit cell of the bulk structure which are parallel to the surface, then the two lattice vectors \(a\) and \(b\) of the reconstructed surface cell are of the type \(a = m_1a + m_2b\) and \(b = n_1a + n_2b\) with \(m_1, m_2, n_1, n_2 = 0, \pm 1, \pm 2, \ldots\). This is shown in Fig. 2 for the \(c(4x2)\) reconstruction of the \(Ge(001)\) surface [10]. The symbol \(c(4x2)\) indicates that the cell is centered and that the coefficients of \(a\) and \(b\) are \(m_1=4, m_2=0, n_1=0, n_2=2\), respectively. The unit cell of the reconstructed surface is always less symmetrical than the surface unit cell of the bulk structure.

b) Surface relaxation: Here, both the topology as well as the interatomic distances of the more external layers remain approximately the same as for the bulk \((m_1=1, m_2=0, n_1=0, n_2=1)\). The principal modifications are in the separations between the upper layers which can change slightly.

There are a great variety of techniques giving information about the atomic structure of surfaces. One of the most popular is the low-energy electron diffraction technique (LEED). This technique is affected by electronic multiple scattering and requires extensive computer simulations, which often leads to ambiguous results. In contrast, the application of the well-known X-ray diffraction techniques to surface science is relatively recent. Since the scattered intensity from a surface is approximately \(10^4\) orders of magnitude lower than that from a bulk crystal, a very intense X-ray beam is required. The solution to this limitation arrived with the advent of the second and third generation synchrotron radiation sources which provide much more intense beams than rotating-anodes.

Even though the X-ray diffraction study of surface structures relies on well-established concepts first developed for the analysis of bulk structures, there are significant differences. In the case of a bulk crystal, for example, diffraction occurs only if the momentum transfer \(q\) of the X-ray beam satisfies the three Laue conditions, i.e. it is given by a point \((h,k,l)\) in reciprocal space where \(h, k,\) and \(l\) are three integers (Miller indices) describing the reflection. For surfaces, the third Laue condition is relaxed due to the truncation of the surface, and may take any real value. Consequently, diffraction will take place along rods in reciprocal space. The termination of the crystal by the surface gives rise to intensity tails of the reciprocal lattice points (normal to the surface). The tails connect the various reciprocal lattice points [1], [2], [28]. The total intensity in a crystal truncation rod (CTR) is the result of the interference between bulk and surface contributions. These rods can be measured with a 6-circle diffractometer. Fig. 3 gives a schematic view of the 6-circle diffractometer normally used for surface structure determination, and Fig. 4 indicates the reciprocal space geometry achieved with this diffractometer.

In Surface Crystallography, the sample preparation as well as the measurements must be performed in ultra high vacuum (UHV) conditions coupled to the X-ray diffractometer. Nowadays, these UHV chambers have become very sophisticated by using a rotary feedthrough which couples the motions of the diffractometer to the sample in vacuum [12].
the determination of the atomic structure is normally accomplished by direct interpretation of 2D difference Patterson maps. These maps are computed with the intensities of the reconstruction in-plane reflections only, i.e. those reflections with no contribution from the bulk ($l \neq 0$). Once interpreted, it is necessary to extend the derived structure model to 3D by using the full data set (in-plane plus out-of-plane data). Very often, the interpretation of such 2D Patterson maps is not trivial, especially for reconstructions with large unit cells. In such cases, direct methods offer an alternative to the other existing methods.

For the application of these methods, which are the methods of choice for solving the atomic structure of bulk crystals, the intensity data set has to meet two conditions:

a) **Atomicity condition:** The experimental resolution of the diffraction data must be high enough to produce separate atomic peaks in the corresponding Fourier maps. Since there is in principle no problem in measuring reconstruction reflections up to a suitable resolution, this first requirement is generally satisfied.

b) **Positivity condition:** The data set must be measured without systematic omission of any set of reflections, or—what amounts to the same thing—the structure factor moduli of all Fourier terms involved in the calculation of the electron-density $p$ with expression (1) must be available. Unfortunately, the integer order reflections possessing contributions of the bulk and the surface are problematic and, at least for the moment, must be excluded from direct methods. This systematic omission of Fourier terms in the summation (1) results in a $p$ distribution containing both positive and negative regions. Hence, the positivity criterion is violated and $p$ and $p^2$ are no longer similar. To distinguish between this partial electron density distribution and the true one, the former is called difference electron density distribution ($\delta$). In general, knowledge of $\delta$ implies the solution of $p$.

The positivity of $p$ (combined with the atomicity) is mainly used in (2) to express the $|G(h)|$ in terms of the measured $|F(h)|$ and of the phases $\phi(h)$. Since the positivity no longer holds for $\delta$, one can assume that $\delta$ resembles its cube $\delta^3$ [23,30]. Let the structure factors $C(h)$ of $\delta^3$ now be introduced. Then the observed $|C(h)|$ are approximately proportional to the measured $|F(h)|$ and the $|C(h,\phi)|$ can be expressed in terms of the $|F(h)|$ and their respective phases with an expression equivalent to (2), the principal difference being an additional summation.

Consequently, by replacing the $G$ by the $C$ in (6), a direct methods sum function results, which is capable of working with partial data sets.

An example of the application of this procedure to the solution of reconstructed surfaces is illustrated in Fig. 5. This figure shows the projected $\delta$ distribution ($\delta$-map) of the C60/Au(110)-p(6x5) reconstruction [18]. Fullerene molecules are placed hexagonally onto the surface at two different levels. These levels are given as dashed lines forming distorted (deeper level) or undistorted (topmost level) hexagonal patterns. The distorted hexagons are larger than the undistorted and, in addition, form calyx-shaped units in which the Fullerene molecules can be accommodated.
Acknowledgements

This work was supported by the Spanish Ministerio de Ciencia y Tecnología (project PB98-0483) and by the Direcció General de Recerca of the Catalan Government (grant SGR-00460, 1995).

References

About the authors

Prof. Miravitlles is a Dr. of Pharmacy and very soon after ending his studies became interested in Crystallography and more specifically in the application of X-ray diffraction methods to the solution of crystal structures with pharmacological interest. He was one of the principal people responsible for the implementation of the direct methods program MULTAN in the Dept. of Crystallography of the University of Barcelona in close collaboration with Prof. Germain at the University of Louvain. Besides his numerous scientific contributions, he has been very active in the planning of the scientific policies of the CSIC. He is a member of the Real Academia de Ciències i Arts de Barcelona and of the European Academy. In 1988, he was instrumental in the creation of the Institut de Ciència de Materials de Barcelona (CSIC) of which he is currently the Director.

Jordi Rius studied Geology at the Universitat de Barcelona and completed his doctorate in Natural Sciences at the Marburg Universität in Germany (1980). He has made numerous contributions to the field of Crystallography and Mineralogy, and has been responsible for the X-ray laboratory of the Instituto Nacional de Seguridad e Higiene en el Trabajo for eight years. He became a full-time [MP5]researcher of the Institut de Ciència de Materials in 1987. His principal research field is the development of new X-ray diffraction methods for structure solution from single crystal and powder data as well as for the microstructure characterization of materials.

Xavier Torrelles is Dr. of Physics (Universitat Autònoma de Barcelona, 1993). During a three-year stay at the Institut Laue-Langevin he became familiar with the techniques of studying disordered materials by neutron diffraction. After four years of postdoctoral work at the surface diffraction synchrotron beam line of E.S.R.F, he became a member of the Institut de Ciencia de Materials de Barcelona. Currently, he is a full-time researcher at this institute and his principal research field is in Surface Crystallography.