

X-Ray Crystallography as an Optical Technique

Adapted from a contribution to the Encyclopedia
of Optical Engineering by Robert M. Sweet
18 November 2002
Biology Department, Brookhaven Nat'l Laboratory
Upton, NY 11973

X-Ray diffraction crystallography has become a major tool for chemistry and biochemistry. It can easily provide detailed structures of molecules of a few atoms, or, with more difficulty, molecular systems of a hundred thousand atoms. It is a tool used in physics and materials science to examine structural features of simple materials in exquisite detail.

X-Ray diffraction is an optical technique in the sense that electromagnetic radiation is used to create an image of an object. We will try in this section to emphasize the similarities and differences between the techniques. In the x-ray case the objects are submicroscopic, and the x-ray light has a wavelength very much shorter than visible light. The interactions between the light and the object to be imaged are similar in both visible and x-ray cases – light is scattered from the electrons in the object. The similarities end when the scattered light is to be combined to form the image; the x-ray worker does not have the lenses or sharply curved mirrors possessed by the optician working with visible light to reconstruct the image. Instead the phase of the scattered rays must be deduced, and then the image must be calculated by Fourier synthesis.

The relation between x-ray crystallography and simple optics can be visualized in Figure 1. In the case of lens optics, one can think of rays of light scattering from points on an object and being collected and directed by the lens to form a magnified image of that point. Alternately, one can realize that all rays scattering in each

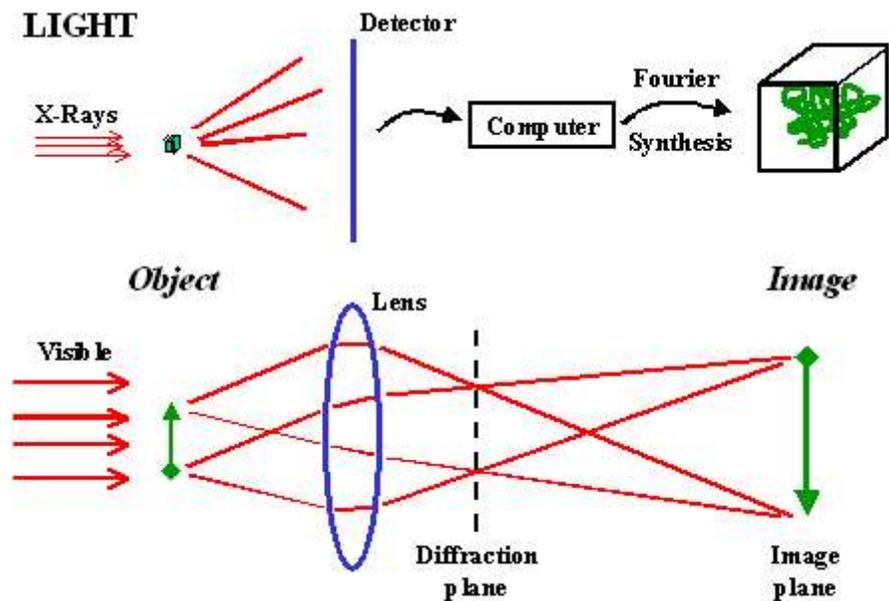


Figure 1

direction are steered by the lens to meet at a point on a plane (the diffraction plane) where they will interfere. These diffracted rays then diverge and continue to the image plane, where they interfere again with all rays coming from a single point on the object but from different directions,

In a classical treatise that still provides a primary text for diffraction physics after 50 years, James (1948) explains diffraction by simple crystals by a method that is similar in principle to one used originally by von Laue. Consider a space lattice of identical point scatterers. When a plane wave sweeps through this lattice, the parallel rays might interact with two of the point scatterers as in Figure 2. The unit vectors \mathbf{s}_0 are perpendicular to the plane wave and are parallel to its direction of incidence. One might observe the radiation scattered from points A_1 and A_2 at a location much farther away than the distance between the two points, in the direction of the unit vector \mathbf{s} . Now one can evaluate the phase difference between these two scattered rays observed at this distant point. The first step is to calculate the path-length difference. Referring to Figure 2, one can see clearly that this difference will be $A_1N - A_2M$. One can use vector notation to

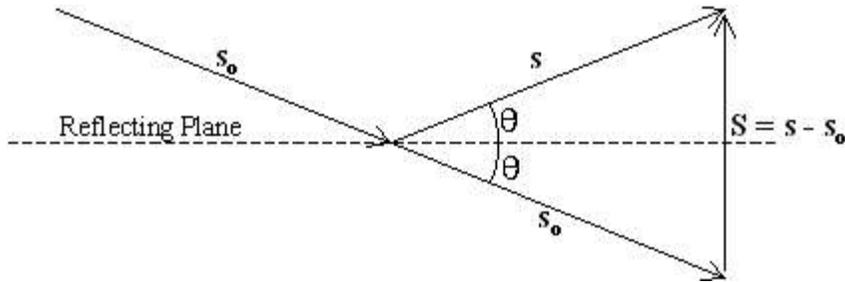


Figure 3

represent these two distances: A_2M equals the scalar product between \mathbf{s}_0 and \mathbf{r} : $\mathbf{r} \cdot \mathbf{s}_0$. A_1N is that between \mathbf{s} and \mathbf{r} : $\mathbf{r} \cdot \mathbf{s}$. Therefore the overall path-length difference is

$$\delta = \mathbf{r} \cdot \mathbf{s} - \mathbf{r} \cdot \mathbf{s}_0 = \mathbf{r} \cdot (\mathbf{s} - \mathbf{s}_0). \quad (1)$$

This vector $(\mathbf{s} - \mathbf{s}_0)$ has a simple geometric interpretation (see Figure 3). One can see that the vector $\mathbf{S} = (\mathbf{s} - \mathbf{s}_0)$ is perpendicular to a plane that would reflect the incident vector \mathbf{s}_0 into the scattered vector \mathbf{s} . Also, by setting the incidence and reflection angles to θ , and since \mathbf{s} and \mathbf{s}_0 are unit vectors, we get that

$$|\mathbf{s} - \mathbf{s}_0| = 2\sin\theta. \quad (2)$$

We can return to our expression for the path-length difference to calculate the phase difference between the two waves scattered from A_1 and A_2 . This phase is that difference δ (eq. 1) scaled around the circle by the wavelength. That is, the phase difference ϕ is:

$$\phi = 2\pi\delta / \lambda = 2\pi\mathbf{r} \cdot (\mathbf{s} - \mathbf{s}_0) / \lambda = 2\pi\mathbf{r} \cdot \mathbf{S} / \lambda = 2\pi\mathbf{r} \cdot \mathbf{R} \quad (3)$$

where \mathbf{R} is now a dimensionless vector parallel to $\mathbf{s} - \mathbf{s}_0$ and scaled by the wavelength.

We use the convention that represents a wave with a complex number: the amplitude of the wave, A_0 , is the modulus of the complex number, and the phase is the arctangent of the ratio of the imaginary to real parts.

$$A = A_0 (\cos\phi + i\sin\phi) = A_0 e^{i\phi} \quad (4)$$

In general, the strength of scattering of x-rays from matter is proportional to the number of electrons in the volume doing the scattering. The phased, scattered amplitude that results when a wave approaching in direction \mathbf{s}_0 is scattered in direction \mathbf{s} from a volume in space at location \mathbf{r} is $\rho(\mathbf{r}) e^{2\pi i \mathbf{r} \cdot \mathbf{R}}$, where $\rho(\mathbf{r})$ is the density of electrons at that point. When there is a finite volume of matter causing the scattering, we can integrate this expression over all of that volume to give the total amplitude of scattering, including phase interference among all the scattering volumes:

$$T(\mathbf{R}) = \int \rho(\mathbf{r}) e^{2\pi i \mathbf{r} \cdot \mathbf{R}} d\mathbf{r} \quad (5)$$

We recognize this as being the Fourier transform of electron density, and know therefore that the inverse transform:

$$\rho(\mathbf{r}) = \int T(\mathbf{R}) e^{-2\pi i \mathbf{r} \cdot \mathbf{R}} d\mathbf{R} \quad (6)$$

converts scattering amplitude back into electron density. The essence of crystallography lies in the problem that the phased scattering amplitude itself cannot be measured in general. Instead when radiation is measured, what one can observe is the *intensity* of the radiation, which is the square of the amplitude. In this measurement, the phase of scattering is lost, and it is therefore impossible directly to evaluate the inverse transform in (6) to reconstruct the original object.

It will pay to explore further the properties of the scattering amplitude given in (5) to see what it shows about the scattering, not from an undefined volume, but from a crystal. Let us build up a crystal of parallelepipeds of matter, repeating N_1 times in the vector \mathbf{a} direction to form a row, this row of units repeated N_2 times in the \mathbf{b} direction to form a plane, and N_3 times in the \mathbf{c} direction to form a solid. Now let

$$F(\mathbf{R}) = \int \rho(\mathbf{r}) e^{2\pi i \mathbf{r} \cdot \mathbf{R}} d\mathbf{r} \quad (7)$$

be the amplitude (transform) of the original unit. When the unit is translated by \mathbf{a} , this transform becomes $F(\mathbf{R}) \cdot e^{2\pi i \mathbf{a} \cdot \mathbf{R}}$. Therefore the transform of the entire row of N_1 repeated units is

$$F(\mathbf{R}) \cdot \sum_{n=0}^{N_1-1} e^{2\pi i n \mathbf{a} \cdot \mathbf{R}} = F(\mathbf{R}) G_a(\mathbf{R}) \quad (8)$$

$G_a(\mathbf{R})$ is the transform of a row of N_1 points, each of unit scattering power and spaced \mathbf{a} along a line (James (1948) pg 4 and 616):

$$G_{\mathbf{a}}(\mathbf{R}) = e^{\pi i(N_1-1)\mathbf{a}\cdot\mathbf{R}} \frac{\sin(\pi N_1 \mathbf{a}\cdot\mathbf{R})}{\sin(\pi \mathbf{a}\cdot\mathbf{R})} \quad (9)$$

If N_1 is large, this function $G_a(R)$ has a magnitude of N_1 at values of $\mathbf{a}\cdot\mathbf{R}$ that are integers, and insignificant values anywhere else. One can carry this out for the other two directions to give a final transform for the crystal of

$$T(\mathbf{R}) = F(\mathbf{R})G_a(\mathbf{R})G_b(\mathbf{R})G_c(\mathbf{R}) = F(\mathbf{R})G(\mathbf{R}). \quad (10)$$

Again, F is the transform of the smallest repeating unit; G is the transform of the lattice alone.

It's worth pausing a moment to examine the relationship between \mathbf{R} and the lattice translations \mathbf{a} , \mathbf{b} , and \mathbf{c} . Recall that $\mathbf{R} = \mathbf{S}/\lambda$ is perpendicular to a plane that would reflect \mathbf{s}_0 into \mathbf{s} . In the case of the translation \mathbf{c} , the reflecting plane is made up by the lattice translations \mathbf{a} and \mathbf{b} (see Figure 4). Let us say that in this case $\mathbf{R} = \mathbf{c}^*$, a vector that goes in the same general direction as \mathbf{c} , but is perpendicular to the \mathbf{ab} plane. We also require that, for (9) to be obeyed, $\mathbf{c}\cdot\mathbf{c}^*$ must equal an integer; a suitable integer is 1. One can see that a way to construct a \mathbf{c} , \mathbf{c}^* pair with this property is to write

$$\mathbf{c}^* = \frac{\mathbf{a} \times \mathbf{b}}{\mathbf{a} \times \mathbf{b} \cdot \mathbf{c}}. \quad (11)$$

And of course one can generate \mathbf{a}^* and \mathbf{b}^* in a symmetrical way. These vectors \mathbf{a}^* , \mathbf{b}^* , and \mathbf{c}^* represent a basis set of vectors in a space with dimensions of reciprocal length: a reciprocal space. The properties of this space are especially useful. If we write that

$$\mathbf{R} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^* \quad (12)$$

we see that we've defined a reciprocal space, with integer indices (h, k, l) , that include all of the values of \mathbf{R} for which (9) and (10) are obeyed. In particular, for the planes generated by the \mathbf{c} translation of Figure 4, the \mathbf{R} vector that describes it is represented by the set of indices $(h,k,l) = (0,0,1)$.

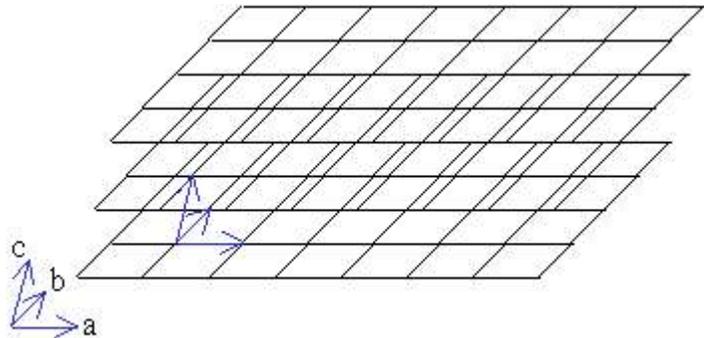


Figure 4

Remember that in Figure 3 we had that the diffraction vector \mathbf{S} lies perpendicular to a reflecting plane (the relationship between \mathbf{S} and \mathbf{R} is shown in (3)). The relationship between these reflecting planes and the indices h, k, l is shown in Figure 5, for a two-dimensional case where only h, k are shown. Note that the h or k index is essentially the number of times that the set of planes cuts the

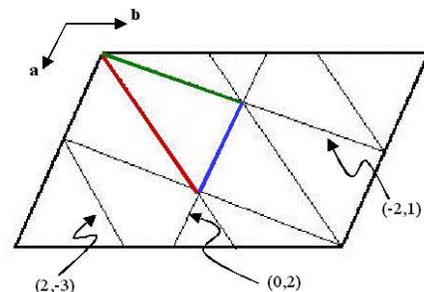


Figure 5

a or **b** axes, respectively. We call these planes “reflecting planes” or “Bragg planes” interchangeably.

A final benefit of this treatment is that the magnitude of the vector **R** equals 1/d, where d is the perpendicular spacing between the planes to which it is perpendicular. If we substitute this and **R** = **S**/λ into (2), we get the classical form of Braggs’ Law:

$$\lambda = 2d\sin\theta. \quad (13)$$

There are other subtle issues here. One is that the lattice function G has a value of $N_1N_2N_3$, where these are the numbers of lattice translations in each direction. This represents the amplification of the transform of the original repeating unit that results from there being a crystal. A second is that when the numbers of lattice repeats are not so large (a few tens rather than many thousands or more), the G function has some breadth at each value of **R** and (12) isn’t true **only** with integer indices.

Reconstruction of the image and phase determination.

The expression for electron density in (6) can be restructured to represent not an integration over space, but a summing over all of the vectors **R** in the reciprocal space implied in (12). This expression

$$\rho(x, y, z) = \sum_{h,k,l} F_{h,k,l} \left\{ \cos[2\pi i(hx+ky+lz)] + \sin[2\pi i(hx+ky+lz)] \right\} \quad (14)$$

is known as a Fourier synthesis, and it depends on complete knowledge of all the terms, the most difficult of which is the phase of the complex term F, the structure factor. Since this phase is lost in measurement of the amplitude of diffraction, it must be determined in some way. To perform this synthesis is the primary scientifically useful result of the diffraction experiment since the density of electrons tells the experimenter where to find the atoms. Typically the work is done by modern Fast Fourier Transform (FFT) methods.

There are several ways of determining the phase of each reflection in crystallography. The classic method, employed by Bragg and his compatriots, was the educated guess. A large number of structures were solved simply by intelligent interpretation of the diffraction patterns that were obtained from crystals; symmetry, overall dimensions of the repeating unit, and chemical composition contributed to the solution.

More recently (1955) Patterson realized that the Fourier transform of the *intensity* of diffraction (the square of the scattering amplitude) was an autocorrelation function of the actual structure. That is, the three dimensional map that resulted from such a calculation (in equation 6, letting **T**(**r**) be +F²) represented all of the *vectors* among atoms in the crystal structure. The next generation of crystallographers solved a series of structures by the use of the Patterson Method.

It was an exquisite exercise in three-dimensional problem solving. A feature of crystals that is entertaining and useful to consider is their symmetry. The combined symmetry elements of a three dimensional crystal form a “group,” in the sense that there is *closure, identity, inverse, etc.* among these elements.

A contemporary of Patterson, David Harker, noticed that the symmetry of a crystal might provide particular vectors that would lead to a direct measure of coordinates of individual atoms. For example, if the crystal’s symmetry included an inversion center, there would be equivalent atoms at positions (x, y, z) and $(-x, -y, -z)$. The vector between these atoms would be found at $(2x, 2y, 2z)$ in the Patterson-function map, indicating the original atomic position. In another example, if there were a two-fold rotation axis parallel to the *c* axis in a crystal, atoms at (x, y, z) would find equivalents at $(-x, -y, z)$. The vector between these two would give $(2x, 2y, 0)$. Thus vectors found at $(u, v, 0)$ could indicate $(2x, 2y)$ positions for individual atoms.

Throughout the 1950s and 1960s, a slight variation of this method was used to solve many structures in the size range of 5-40 independent atoms, where there was at least one atom that was much heavier than all the others. Examples might be organic molecules with one or more metal ions, or sulfur, chlorine, or bromine atoms bound. This was to use the Patterson function to locate the heavy atom(s), use the structure-factor expression to calculate the structure factors for just these atoms, then to use the phases from these structure factors in calculation of a complete Fourier synthesis. Since often the heavy atoms contribute most the phases, the remaining atoms would appear in an electron density map that was calculated with the complete amplitudes. In the macromolecular community, there is a range of algorithms, embedded in computer software, that can solve quite complicated Patterson functions quite accurately.

During the 1950s and ‘60s a range of workers, especially Sayre, Hauptmann, and the Karles, were working to find ways to use the diffraction intensities directly to determine crystal structures, without resorting to use of heavy atoms. They made use of two basic principles. The first was that certain sets of Bragg planes could be combined to form what was known as a “structure invariant.” In particular, it was straightforward to show that the sum of phases of three Bragg planes that form a closed triangle is invariant to the choice of the origin of the crystallographic unit cell. Secondly, in the case where, firstly the crystal is comprised of discreet atoms and, secondly, all three of the diffraction intensities (structure factors) from these Bragg planes are large, this sum of three phases is near to zero. It’s not necessarily so easy to see what this means. Firstly, notice that the three different Bragg planes we’ve chosen to illustrate the plane indices in Figure 5 form a closed, colored triangle at the upper left of the parallelogram. Then notice that the indices of the three sets of planes in Figure 5 sum to zero. This particular set of reflections is called a “triplet,” for obvious reasons. This condition,

$$h + k + l = 0 \quad (15)$$

defines the structural invariant wherein the sum of phases $\phi_h + \phi_k + \phi_l = \text{const}$. It’s not so hard to show that this is true: multiply three Fs and the three phases end up in a sum. Finally, it’s not so hard to see that if the three Fs are large, the sum of phases should be near zero, as follows. If

the only three atoms in the unit cell were at the corners of the colored triangle in Figure 5, firstly, all three structure factors would be large since all atoms lie only on the planes, and secondly, since the atoms are *on* the planes, the phases would be zero. One can see that it makes sense that this sum of phases might be constant. If one moves an atom from one vertex of the colored triangle in Figure 5 to the next along the green line (the -2,1 plane), one can see that the other two phases, for the red and blue sets of planes, would shift smoothly by $+2\pi$ and -2π respectively, keeping the total constant.

The basic scheme for “direct” phase determination is to find the largest structure factors (in practice a modified structure factor is used), to find all the triplet relations (15) among these reflections, to assign a few phases arbitrarily to define the origin of the unit cell, and then to propagate these phases through all the relations (15). The refinements of this simple scheme that make the method practical involve methods to deal with phase sums not being precisely zero, and ways to determine phases when all strong reflections are not related by (15).

Crystallographic methods also can be applied to quite large structures: proteins, viruses, protein/nucleic-acid complexes. The methods described above won't work. In the case of the “heavy atom” method, there simply isn't enough residual scattering power in a few heavy atoms to provide phases that are close to being correct. In the case of the direct methods, a necessity is that the diffraction represent atomic resolution of atoms, and this is not usually realized with large molecules like these. The principal method used instead is some variant of a technique known as “isomorphous replacement.” In this method, crystals of the macromolecule are prepared in the normal way, and then with a small number of heavy atoms bound to the macromolecule in chemically specific and reproducible places, and in a way that does not disrupt the crystallinity of the specimen. The two crystal forms must be “isomorphous,” having the same form except for the binding of a few heavy atoms. The way phase information is extracted is that diffraction

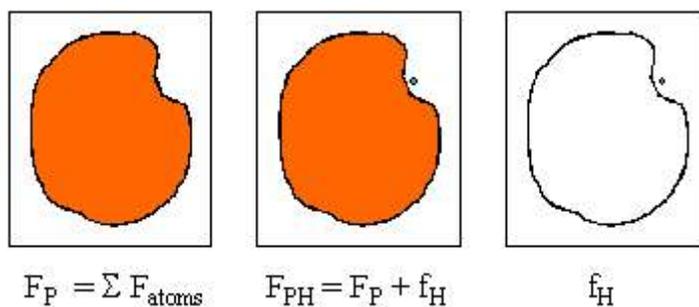


Figure 6

measurements are made on both kinds of crystal: without (called the “parent” structure) and with the heavy atoms bound (termed the “derivative” structure).

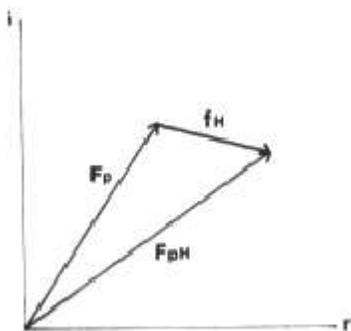


Figure 7

This situation is represented in Figure 6. If the two structures on the left are truly isomorphous, the structure factors of the derivative (in the middle in Figure 6) are simply the sum of the structure factors of the parent plus the structure factor of the heavy atoms alone. The thrust of the phasing method is somehow to determine the positions of the

small number of heavy atoms, and then to use the phase information that comes from this to infer phases of the parent structure. We find that, for the purposes of this method we can approximate $|F_H|$ with $|F_{PH} - F_P|$. This often suffices for us to solve for the positions of the heavy atoms as if this were a small-molecule structure.

The true situation (in the complex plane) is represented in Figure 7. Each structure factor has a phase, and they sum as vectors in the complex plane. Once the heavy atom positions are known, the value of f_H is known also. Since only the amplitudes for the other two structure factors can be known, they must be estimated by ones making an attempt to close the triangle shown in this figure. The construction devised by Max Perutz for this purpose is shown in Figure 8. Here one uses the exact equation $F_P = F_{PH} + (-)f_H$ and constructs the equation graphically. The circle with diameter F_H is centered at the origin. The one with diameter F_{PH} is centered at the end of $-f_H$. Therefore, the vector sum is at the two intersections shown at H and G. One of these two has to be the correct phase. Typically, the way this ambiguity is resolved is by ones doing the experiment again – double or multiple isomorphous replacement will give a reasonably accurate estimate of the true phase, and will lead to an accurate Fourier synthesis and a model that represents the true structure of the macromolecule.

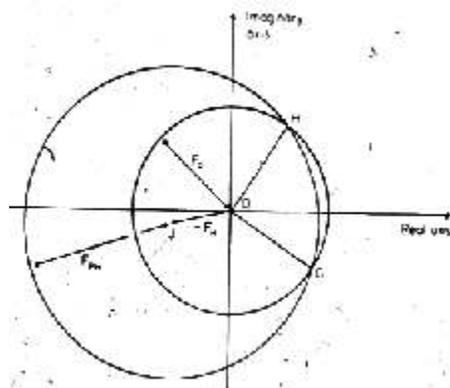


Figure 8

Since 1980 there has been a marriage of high-energy physics and x-ray studies that has provided extraordinary x-ray sources for diffraction studies. These are the electron synchrotron sources, wherein the radiation that is produced when electrons are induced to turn the corners in these particle accelerators is used for x-ray diffraction experiments. These x-rays have numerous properties that make them especially useful for this work: they are well collimated, they often are intense, and that they are polychromatic. The individual energy or wavelength bands can be selected by use of a perfect single crystal used in Bragg diffraction. See a representative diffraction image from a synchrotron source in Figure 9.

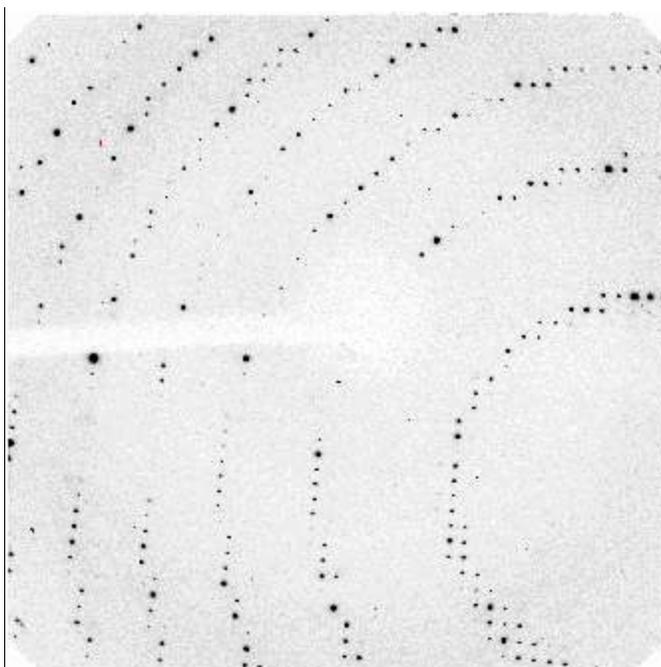


Figure 9: X-Ray diffraction image taken at the National Synchrotron Light Source from a single crystal of hen egg-white lysozyme. The crystal was rotated about 1 deg to get the pattern; the exposure was about 1/2 minute.

The tunability or wavelength selection of the x-rays can be especially useful for macromolecular crystallography.

If the x-ray source can be tuned to an atomic absorption edge of the heavy atom used for phase determination, the atom can be made to resonate with the radiation, and there are accompanying shifts in the amplitude and phase of the scattering from that atom. These shifts can be used in the same way as multiple isomorphous replacement to resolve phase ambiguity and produce an accurate phase. At the time of this writing, a reasonably large fraction of the new macromolecular structures solved depend on not only synchrotron radiation, but also the use of this resonant (also called anomalous) effect.

The usefulness of crystallographic results:

In recent years, the sort of structural details that result from crystallographic studies are taken as “givens” in many fields of research. Because modern x-ray sources, detectors, and mathematical methods allow it, one can expect to know the structure of a ~50 atom compound within a few hours of having grown a crystal of it. Materials scientists and the semiconductor industry use x-ray diffraction as a tool to probe microscopic details of structure. Pharmaceutical companies have developed sizable research groups, firstly to explore new macromolecular drug targets, like enzymes, and then to test the binding of putative drugs in substantial screening efforts. The effectiveness of several treatments for human disease depend on the use of these methods. Much more fundamental questions are answered with some effort by employment of the method. The structure of the ribosome and its subunits has been determined by fairly conventional application of the sort of crystallographic methods, but with the subject being a particle of 150,000 atoms. The position is known of virtually every atom in that structure.

It is truly a golden age for x-ray diffraction methods. The modern synchrotron sources provide x-rays with unparalleled brightness. Area-sensitive x-ray detectors, virtually TV cameras for x-rays, allow rapid and accurate collection of the diffraction data. The brightness of the synchrotrons can provide beams in the 1 μ m size range to probe individual components of computer chips. The tunability and brightness allows the solving of macromolecular structures in ways that could become as routine as those employed in solving of small molecules.

International recognition of the value of diffraction methods

Numerous Nobel Prizes have been given that relate directly to crystallography or to diffraction from ordered but non-crystalline materials:

1914	Max von Laue	Physics	Original observation and explanation of x-ray diffraction from crystals
1915	William Bragg, Lawrence Bragg	Physics	First solving of atomic structure of crystals
1954	Linus Carl Pauling	Chemistry	Understanding of the chemical bond, derived in part from crystal structures

1962	Francis Harry Compton Crick, James Dewey Watson, Maurice Hugh Frederick Wilkins	Medicine	Elucidation of structure of DNA from diffraction from fibres of the molecule
1962	Max Ferdinand Perutz, John Cowdery Kendrew	Chemistry	First determination of the structure of proteins
1964	Dorothy Crowfoot Hodgkin	Chemistry	The structure of vitamin B12 by x-ray crystallography
1976	William N. Lipscomb	Chemistry	Studies on the structure of boranes illuminating problems of chemical bonding
1982	Aaron Klug	Chemistry	Development of crystallographic electron microscopy
1985	Herbert A. Hauptman, Jerome Karle	Chemistry	Development of direct methods for crystal-structure determination
1988	Johann Deisenhofer, Robert Huber, Hartmut Michel	Chemistry	Determination of the structure of a membrane-bound protein
1994	Bertram N. Brockhouse, Clifford G. Shull	Physics	Development of methods for diffraction of neutrons
1997	Paul D. Boyer, John E. Walker, Jens C. Skou	Chemistry	Elucidation of the enzymatic mechanism underlying the synthesis of adenosine triphosphate (ATP)

References:

Ewald, P.P. editor. (1962) Fifty Years of X-Ray Diffraction. International Union of Crystallography, Utrecht, The Netherlands.

James, R.W. (1948) The Optical Principles of the Diffraction of X-Rays. Cornell Univ. Press, Ithaca.

Friedrich W., Knipping P., and Laue, M. (1912) Interferenz-Erscheinungen bei Röntgenstrahlen, Proceedings of the Bavarian Academy of Sciences, pg. 303-322.

Laue, M. (1912) Eine quantitative Prüfung der Theorie für die Interferenzerscheinungen bei Röntgenstrahlen, Proceedings of the Bavarian Academy of Sciences, pg. 363-373.