Review

x Ray crystallography

M S Smyth, J H J Martin

Abstract
***x* Ray crystallography is currently the most favoured technique for structure determination of proteins and biological macromolecules. Increasingly, those interested in all branches of the biological sciences require structural information to shed light on previously unanswered questions. Furthermore, the availability of a protein structure can provide a more detailed focus for future research. The extension of the technique to systems such as viruses, immune complexes, and protein–nucleic acid complexes serves only to widen the appeal of crystallography. Structure based drug design, site directed mutagenesis, elucidation of enzyme mechanisms, and specificity of protein–ligand interactions are just a few of the areas in which *x* ray crystallography has provided clarification.**
(*J Clin Pathol: Mol Pathol* 2000;**53**:8–14)

The aim of *x* ray crystallography is to obtain a three dimensional molecular structure from a crystal. A purified sample at high concentration is crystallised and the crystals are exposed to an *x* ray beam. The resulting diffraction patterns can then be processed, initially to yield information about the crystal packing symmetry and the size of the repeating unit that forms the crystal. This is obtained from the pattern of the diffraction spots. The intensities of the spots can be used to determine the "structure factors" from which a map of the electron density can be calculated. Various methods can be used to improve the quality of this map until it is of sufficient clarity to permit the building of the molecular structure using the protein sequence. The resulting structure is then refined to fit the map more accurately and to adopt a thermodynamically favoured conformation.

It is beyond the scope of this review to provide a complete manual for everything from crystallisation to model building. More detailed reviews can be found elsewhere.[1–3]

Protein crystallisation
To perform protein crystallography, a reliable source of protein must be available, together with a purification/concentration protocol that will yield high quality, homogeneous, soluble material.

The growth of protein crystals of sufficient quality for structure determination is, without doubt, the rate limiting step in most protein crystallographic work, and is the least well understood. The principle of crystallisation, whether of macromolecules or salts (unfortunately!) is to take a solution of the sample at high concentration and induce it to come out of solution; if this happens too fast then precipitation will occur, but under the correct conditions crystals will grow.[4–6] The elucidation of these conditions determines the rate limiting step and indeed whether or not the project will be possible. Many projects prove not to be possible because of the inability to crystallise the protein. The magnitude of the problem can be understood when one considers the variables[7 8]: the choice of precipitant, its concentration, the buffer, its pH, the protein concentration, the temperature, the crystallisation technique, and the possible inclusion of additives. Essentially, initial experiments will be based on a trial and error procedure, which aims to cover as wide a range as possible of the variables as is practical. Commercially available "crystal screen" packages are often used at this stage. Each one usually consists of 50 solutions varying widely in precipitant, buffer, pH, and salt, known as a sparse matrix.[9] These can then be set up using the techniques of sitting drop vapour diffusion, hanging drop vapour diffusion (fig 1),[10] and possibly dialysis,[11] usually at both room temperature and 4°C. In this way, many of the variables can be covered easily, and one or more might even yield
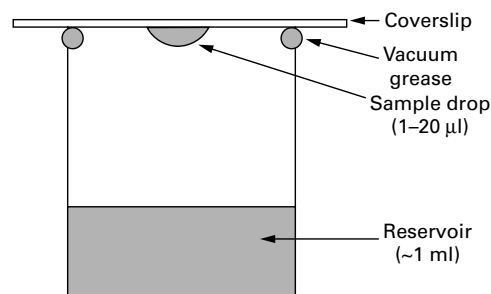

Figure 1   Hanging drop vapour diffusion is usually set up with a drop of 1–20 μl suspended from a glass coverslip over the reservoir solution. The drop is a 50/50 (vol/vol) mixture of the protein solution and the reservoir solution. Hence, the vapour pressure of water around the drop is greater than that over the reservoir. The pressure gradient across the vapour space leads to a net loss of water in the drop. Sitting drop vapour diffusion differs only in that the drop sits in a depression on a specially constructed raised platform within the diffusion chamber.

Department of Biochemistry, University of Leicester, Leicester, LE1 7RH, UK
M S Smyth

Division of Biomedical Sciences, School of Health Sciences, University of Wolverhampton, Wolverhampton, WV1 1DJ, UK
J H J Martin
Correspondence to:
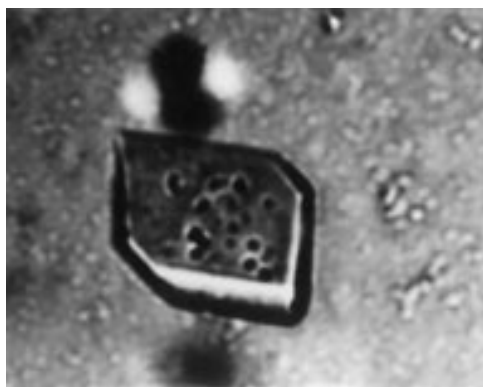Dr Martin
email: J.Martin@wlv.ac.uk

Figure 2 A crystal of a bovine picornavirus[16] measuring 0.2 mm in the longest dimension grown, using ammonium sulphate as precipitant and using microdialysis apparatus. These crystals take 48 hours to grow at room temperature, but are very fragile because the virus particles in the crystal are roughly spherical and 300 Å in diameter, making the solvent content very high and the particle contacts very low. This is common with many protein crystals and great care must be taken during manipulation.

crystals of sufficient quality to proceed to the next step. However, it is more usual at this stage to see various combinations of the following: nothing, precipitation, showers of microcrystals (which often resemble a precipitate), or a few very tiny crystals. Obtaining either of the last two results is encouraging because it indicates that the macromolecular subunit might possess sufficient inherent structural order or symmetry to make crystallisation possible. In general, those proteins that are glycosylated[12]
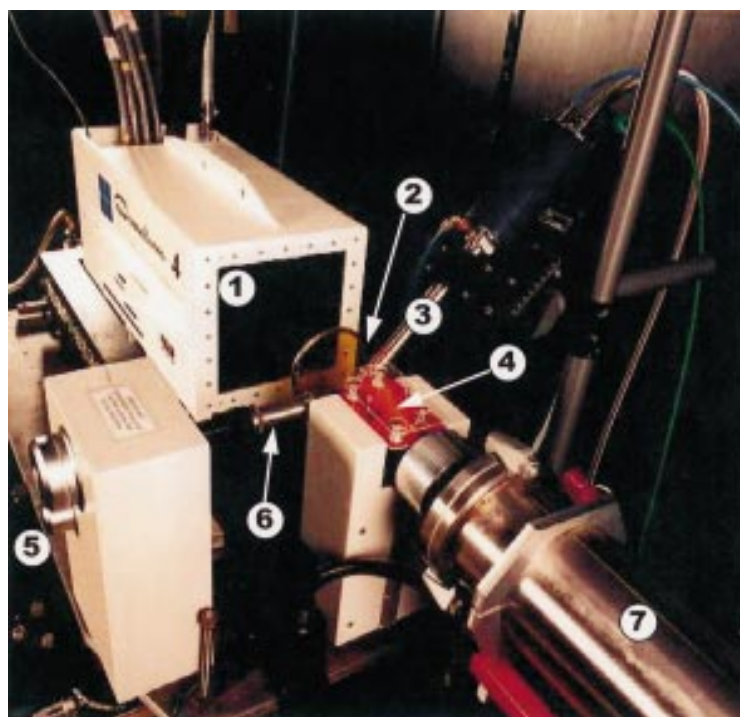

Figure 3 A typical synchrotron data collection station: PX 9.6 at Daresbury, Cheshire, UK. A charged coupled device detector (1) is mounted on a motorised system, which enables the "crystal to detector" distance to be altered. This is protected from the direct beam by the backstop, a small lead pellet mounted on plastic film supported by a ring (2). Frozen crystals are cooled by a stream of nitrogen at 100 K from the nozzle (3). The diameter of the beam incident on the crystal is adjusted manually using two horizontal and two vertical slits on the panel (4). The spindle (6) on which the crystal is mounted can be rotated manually (5) to ensure perfect alignment. The complete assembly is housed within a lead walled "hutch" and the x ray beamline (7) runs tangentially from the storage ring. Photograph courtesy of M Papiz, CLRC Daresbury Laboratory.

or contain flexible or less conformationally constrained domains are difficult to crystallise, whereas even very large complexes of high symmetry, such as many viruses, will crystallise.[13 14] Various techniques can be used to improve crystal size. These include seeding,[15] alteration of protein concentration, or alteration of temperature. For diffraction analysis, protein crystals are usually required to be a minimum of 0.1 mm in the longest dimension, to provide a sufficient volume of crystal lattice that can be exposed to the beam (fig 2).

Before proceeding, it must be verified that the crystals contain the desired macromolecule as opposed to one of the salts from the precipitant or buffer. This can be done by sacrificing some crystals to polyacrylamide gel electrophoresis analysis, staining, or a test x ray diffraction exposure.

Optical setup

The x rays can be generated from accelerating electrons in a synchrotron storage ring or from electrons striking a copper anode. In the former case, a single x ray wavelength is usually selected by absorption of unwanted wavelengths in a procedure known as monochromation. In the latter case, this is not necessary because one predominant wavelength is produced. From either source, the x rays must be focused into a beam and then collimated to ensure that the beam is parallel and, as far as possible, lacking "crossfire". x Ray beams are usually collimated with sets of adjustable slits to 0.1–0.3 mm diameter.

The crystal is mounted in this beam and adjusted carefully on a device known as a goniometer head to ensure that it remains in the beam as the spindle on which it is mounted rotates (fig 3). The only remaining object between the crystal and the x ray detector is the backstop. This is usually a small lead pellet suspended in the path of the direct beam to prevent this intense source reaching the detector and either damaging it or causing over exposure of the central region.

Diffraction analysis

Having obtained crystals of suitable size, which have been confirmed to contain the macromolecular subunit of interest, the next step is to analyse their x ray diffraction behaviour. Trial exposures of new crystals can be performed on laboratory "in house" x ray generators or at synchrotron sources. The latter have the advantage of an extremely intense x ray beam with high quality optics, which allows much shorter exposure times and a higher signal to noise ratio of the diffraction image; therefore, these are often favoured for more challenging crystallographic problems.[17 18] In addition, one also has the choice of exposing the crystals mounted in a capillary tube at room temperature, or mounted frozen in a small loop in a stream of liquid nitrogen at 100 K.[19] Data collection from frozen crystals has the advantage that it may be possible to collect a complete data set from a single crystal, whereas capillary mounting is often a simple way to

determine the diffraction characteristics of a new crystal.

Until quite recently, *x* ray diffraction images were collected on conventional *x* ray film, which was developed and fixed using normal photographic techniques. Within the last decade, this has been superseded by the imaging plate,[20] which is at least 10 times more sensitive than *x* ray film, and which within a few minutes of the completion of the exposure reads out a digitised image to the controlling workstation. However, even these are now being replaced by detectors using charged coupled device (CCD) technology.[21 22] Such CCD detectors offer the further advantage of readout times in the order of seconds rather than minutes, which can appreciably cut the time for collection of a complete data set, especially when one considers that the exposure times at many synchrotron beamlines are less than one minute. This compares very favourably with film; exposure times of 30–40 minutes at synchrotrons and of many hours at in house generators were not uncommon. Such long exposure times, coupled with the tedious process of developing and fixing the films in a dark room and optically scanning to digital form before processing could begin, made data collection with *x* ray film a very labourious process.

Before the first exposure, the distance from the crystal to the detector is calculated and adjusted to allow for collection of diffracted spots, usually up to a maximum of 1.5–3.0 Å resolution. The resolution of spots collected on the detector increases as the diffracted angle increases. Hence, the highest resolution will be at the edge of the detector, and if one determines the diffracted angle required, the distance of the detector from the crystal can be adjusted accordingly.

Awaiting the first diffraction image from a new crystal is an exciting event, but what information can be gained from this image?

First, one must confirm that the diffraction extends to sufficient resolution to make structure determination to near atomic detail feasible. Essentially, this involves being able, visually, to detect well ordered arrays of spots towards the edge of the diffraction image. Many of the image display programmes incorporate an algorithm to determine the resolution of a particular spot. In general, spots beyond 3 Å are required: a carbon–carbon bond is approximately 1.5 Å, but a resolution of close to 3 Å is sufficient to be able to detect the amino acid side chains in the electron density map. The diffraction image becomes weaker at high resolution, limited ultimately by how ordered the molecular subunit is. Hence, a compromise must be made between increased resolution and decreased diffraction quality: often the cut off point can only be determined when the data are processed. Second, one may determine the unit cell dimensions, the crystal system, and the space group. The unit cell is the smallest repeating unit that makes up the crystal. Its dimensions are given as three lengths: a, b, and c; and three angles: α, β, and γ (which are usually omitted if 90°). The dimensions of the unit cell determine the spot

Table 1   The seven crystal systems

Crystal system	Conditions
Triclinic	None
Monoclinic	$\alpha = \gamma = 90°$
Orthorhombic	$\alpha = \beta = \gamma = 90°$
Tetragonal	$a = b; \alpha = \beta = \gamma = 90°$
Trigonal	$a = b; \alpha = \beta = 90°; \gamma = 120°$
Hexagonal	$a = b; \alpha = \beta = 90°; \gamma = 120°$
Cubic	$a = b = c; \alpha = \beta = \gamma = 90°$

spacing on the diffraction image: it is a reciprocal relation, so the larger the cell the more spots present for each unit area. The shape, whether cube, parallelepiped, or whatever, determines the crystal system, seven of which exist (table 1). The space group is provided by the symmetry of the diffraction pattern. It allows the packing of the molecules into the crystal lattice to be determined. A total of 230 space groups exist, but not all of these are permitted for proteins because of the chirality of amino acids.

Data collection
The quantity of data required and strategy for data collection[23] for a structure determination depend on several conditions.

(1) Crystallographic symmetry: the amount of symmetry present in the crystal system and space group. With a high symmetry crystal system—for example, a cubic, one needs only to collect diffraction data through as little as 35°. Conversely, in a lower symmetry crystal system, such as a monoclinic, data might need to be collected through 180°.

(2) Non-crystallographic symmetry (NCS): the amount of symmetry present in the asymmetric unit; that is, the equal particles in the unit cell related by symmetry operations. A particle such as a virus, composed of many identical subunits, has a high level of NCS, as much as 30-fold or even 60-fold.[14] In cases such as these, high quality structures can be obtained from data sets that are far from 100% complete because of compensation by averaging (see below). At the other extreme, a monomeric protein might exhibit no NCS and therefore a more complete data set will be required.

(3) The availability of molecular replacement (see below). If a sufficiently similar structure has already been solved, it can be used as a starting model and any gaps in the new data set may be "filled in". If not, it will be necessary to collect at least one more data set, this time from a heavy atom derivative or derivatives, the data sets being referred to as "native" and "derivative", respectively.

(4) The upper resolution limit required. This will usually be determined by the quality of the crystal itself and in practice most data sets are collected to the upper limit of useful diffraction. However, the amount of diffraction data increases exponentially with resolution. Although it affects the total amount of data required, this parameter does not change the angle through which data must be collected.

Initial diffraction analysis (above) will have given an indication of the spot spacing on the
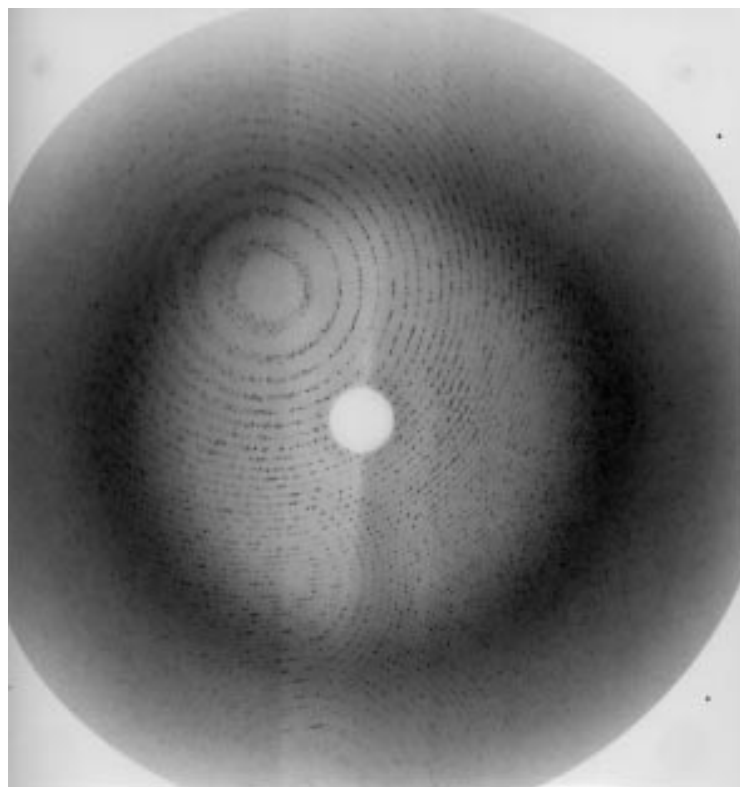
Figure 4   x Ray diffraction photograph taken at PX7.2, Daresbury Laboratory, Cheshire, UK, of bovine enterovirus[16] with 0.5° oscillation range and maximum resolution at the edge of the film of 2.8 Å. The x ray wavelength was 1.488 Å.

detector at the desired "crystal to detector" distance. This may be used to determine the oscillation range, or ϕ range, of each exposure. The crystal is rotated on a spindle while exposed in the *x* ray beam. This rotation axis is perpendicular to the beam and has the effect of sweeping more crystal lattice planes through the beam so that the maximum amount of data can be recorded for each image. For large unit cells, generally upwards of 300 Å, as little as
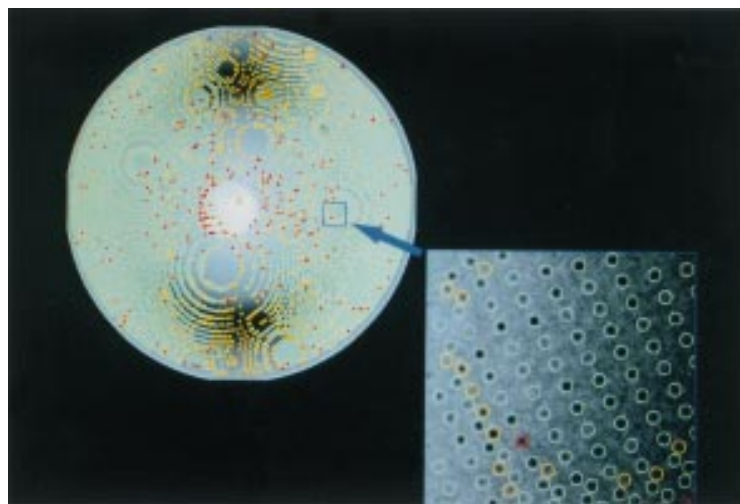

Figure 5   Processing of x ray diffraction data using the program DENZO[26] for autoindexing. The program searches for peaks of intensity on the image and, using the "crystal to detector" distance and the wavelength, determines the unit cell dimensions and crystal system. It then calculates a prediction of what the image will look like at this crystal orientation and superimposes this on the real image. The enlarged portion shows the coloured circles of the predicted spots surrounding real spots. The user refines various parameters until the fit is optimised and then the program measures the intensities within the circles.

0.25° might be possible, whereas for smaller unit cells, up to 2° might be collected while still avoiding spot overlap.

In an ideal situation, a complete data set will be collected from a single crystal. This greatly facilitates the scaling process (see below), but in practice is often not possible because of radiation damage to crystals. Such damage is caused by heating and the generation of free radicals; hence, radiation damage can continue for some time after the *x* ray exposure is complete.

Modern detectors, such as those using CCD technology, have very short read out times. Taken together with the short exposures possible on the brightest synchrotron beamlines, this facilitates the collection of a data set in a matter of hours. Moreover, the entire process is automated, with the data collection software controlling the opening and closing of the *x* ray shutter, the rotation of the spindle, the read out of the detector, and even the numbering of the images.

Figure 4 shows an *x* ray diffraction image from a virus crystal, similar to that in fig 2, taken on conventional *x* ray film. The close spacing of the spots is a result of the large unit cell, with each dimension in excess of 350 Å. The concentric rings, or lunes, of spots are the result of the diffracted rays being emitted in cones from the crystal lattice. The general appearance of the diffraction pattern is the result of the crystallographic symmetry, in this case a monoclinic system, whereas the spacing of the spots is dependent only on the unit cell dimensions. It is the variation in the intensities of each of the spots that contains the structural information and which is extracted during the data processing (see below). The thickness of the lunes, or total number of spots on the image, can be controlled by the oscillation range of the exposure: this being a compromise between collecting the maximum amount of data possible for each image and the onset of spot overlap. The unexposed region at the centre is the shadow of the backstop and the darker ring just over halfway out is the solvent ring, resulting from *x* ray scattering from disordered solvent molecules.

Data processing

The processing of the diffraction data is mathematically complex. Fortunately, however, well established algorithms are available in many software packages and program suites. Their existence enables the relative newcomer to process data and calculate an electron density map with only minimal guidance and with even more minimal mathematical knowledge. However, like any other procedure, an in depth understanding is beneficial if and when data processing problems occur.

The first step in the processing involves the determination of the crystal system and of the unit cell dimensions as accurately as possible. In addition, at this stage we determine the orientation of the crystal in the beam.[24] This is usually carried out on the first diffraction image because it is usually the best quality image, and with the knowledge of the

oscillation ranges one can calculate the subsequent crystal orientations. When the cell and orientation are known, indexing can be carried out.[25] In this, each spot on the image is assigned an index, quoted as three integers: *h*, *k*, and *l*. Computer programs for autoindexing do this by calculating a prediction of what the diffraction image will look like from the cell dimensions and orientation, then attempting to fit the real image with the predicted one (fig 5). This is an interactive process, so that the user can check for accuracy, which helps to avoid possible errors as a result of mis-indexing.

The next stage of the data processing is the measurement of the intensities of the spots. Protein crystals diffract weakly because they are composed mainly of light atoms and they have large unit cells. The larger the volume of the protein crystal, the stronger its diffraction. Intensities of diffracted spots vary as a result of both the amplitude of the diffracted waves and their phase relation. These factors cannot be deconvoluted at this stage; accordingly, the accuracy of the measurement of the intensities is of paramount importance. One of the programs used most frequently in protein crystallography is DENZO,[26] which performs both autoindexing and intensity measurements.

A scale factor must be allocated so that the intensities of all the images in the data set can be related. The first image is usually allocated a scale factor of 1 and all the subsequent images will be scaled up to this. At this stage, we see the benefit of the collection of the complete data set from a single crystal because only one reference image is required for the allocation of the scale factor. If data have been collected from more than one crystal, possibly from more than one *x* ray source, data subsets from each crystal will have to be scaled separately and then these batches scaled together in blocks, the best batch then being given a scale factor of 1. Careful monitoring of the scaling statistics allows spots, or even whole images, to be rejected or reprocessed to preserve the quality of the overall data set. SCALEPACK is a widely used program for this step.[26]

The output from scaling is a computer file that contains the index of each spot and its measured intensity. This file must be sorted so that the spots are listed in numerical order according to index; numerous programs exist for this purpose, a commonly used one being SortMTZ of the CCP4 suite.[27]

Determination of the amplitudes
The intensity of any diffracted spot is the result of the diffracted waves incident on the detector at that point. Therefore, the intensity will be determined by the amplitude of those waves and by the phase difference, expressed as an angle, between them. A phase difference of zero results in constructive interference, whereas a phase difference of 180° results in completely destructive interference. The determination of the amplitudes is mathematically simpler than that of the phases (see below). Various computer programs exist to calculate amplitudes, for example Truncate of the CCP4 suite.[27] Usually, these take the square root of the intensity, with negative values being set to zero, and outputting amplitudes accordingly.

Solution of the phase problem
As discussed above, the measured intensity of a diffracted spot is a function of the amplitude of the reflection and the phase angle between the diffracted waves. From a knowledge of the amplitude and the phase we can determine a parameter known as the structure factor from which the arrangement of the atoms in the unit cell can be calculated. We have already seen how the amplitudes can be found. The phase angle cannot normally be determined directly in the case of protein crystals and so must be found in an indirect way. The two most frequently used methods are isomorphous replacement and molecular replacement.

THE ISOMORPHOUS REPLACEMENT METHOD
This method[28–30] is normally used in cases where no closely related structure is available and requires at least two data sets: one native set from the protein crystal alone and at least one derivative set from the protein crystal with attached heavy atoms. In practice, the protein crystal can be soaked in a solution of a heavy atom salt, such as mercury, platinum, or gold. The object of this exercise is to incorporate and attach one or a few heavy atoms to the protein molecule while not appreciably altering either the conformation of the protein or the unit cell dimensions. Such perfect isomorphism rarely occurs but a small degree of non-isomorphism is tolerable. When compared, the differences between these data sets results solely from the heavy atoms and, therefore, their positions in the protein molecules can be determined. Computer refinement of the heavy atom parameters is carried out[31] and these parameters can be used as a starting point to determine the protein phase angles. Together with the amplitudes, these can be used to calculate the structure factors, which can then be subjected to rounds of refinement (see below).

THE MOLECULAR REPLACEMENT METHOD
This method[32] is the most rapid and most frequently used when a very closely related protein structure is available. This might either be when there is close homology of the amino acid sequence—for example, an enzyme derived from two different sources, or if two structures are expected to be similar, for example two closely related viruses. It might even involve the construction of a special model, taking the most closely homologous sections from various related proteins, in an attempt to mimic more closely the new molecule.[14] The method involves the crystallographic calculation in reverse: structure factors from the known coordinate file, and the subsequent "borrowing" of the phases from the known model structure and their application to the new data set to calculate the new structure factor. Hence, there will always be a certain amount of bias towards the model in the initial structure factor calculations. This bias will not be present if the isomorphous replacement
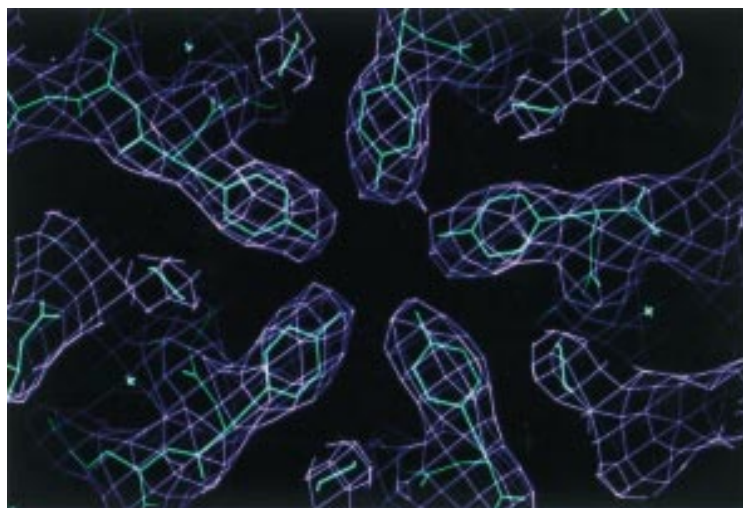
Figure 6   A portion of an electron density map calculated by molecular replacement to 3 Å resolution. The photograph represents a slice through a virus structure[14] at the point of fivefold non-crystallographic symmetry. Five tyrosine residues from symmetry related subunits can be seen in their entirety, but the complete model has been built into the density. This natural fivefold symmetry axis also serves as a centre for crystallographic averaging. Because this particular virus has 12 such points, none of which lie along crystallographic symmetry axes, 60-fold averaging was possible.

method is used. Before the phases can be applied, the model structure must be placed into the unit cell in exactly the same position and orientation as the new protein molecule. This is carried out initially by a rotation function,[33 34] which rotates the model data to fit the new structure accurately, and secondly by a translation function,[35 36] which moves this reoriented data through the unit cell to fit the position of the new molecule most accurately. Once our model has been thus placed its position and orientation can be improved through cycles of refinement.[37] The phases from the model structure factors may then be applied and, with the new amplitudes, a new set of structure factors calculated and refined (see below). A commonly used molecular replacement program package is AmoRe.[38]

Calculation of an electron density map
Now that we have the amplitudes and phases we can calculate the structure factors, invariably using the fast fourier transform (FFT) method.[39] The resulting electron density map will form the three dimensional contours into which the protein structure will be built. Each of the unit cell edges is divided into spacings of a few Ångstroms. The spacing determines the

Atom	10	N	ALA	1	15	171.409	–3.184	136.111	1.00	98.33
Atom	11	CA	ALA	1	15	171.071	–4.402	136.876	1.00	95.40
Atom	12	C	ALA	1	15	172.191	–5.456	136.991	1.00	92.03
Atom	13	O	ALA	1	15	172.885	–5.753	136.007	1.00	91.20
Atom	14	CB	ALA	1	15	169.782	–5.047	136.315	1.00	94.99

Figure 7   The format of a protein data bank (PDB) file. Structural data may be downloaded from the Brookhaven protein data bank in files containing the following information: the first column indicates that the line contains atom data rather than—for example, a remark; the second column is the atom number; the third, the atom type, CA representing an α carbon, and so on; the fourth, the residue type in three letter code; the fifth, the chain identifier, if the structure has more than one subunit; the sixth, the residue number; the seventh, eighth, and ninth, the coordinates of the atom within the unit cell; the tenth, the occupancy—usually assigned to 1.00; the eleventh, the B factor, a measure of how much an atom vibrates around its equilibrium position.

quality of the map detail and the speed of the calculation. This creates a three dimensional grid within the unit cell and the electron density is calculated at each of the grid points. This computation may be greatly accelerated by applying limits to the size of the grid, according to where our protein molecule is located within the unit cell. The setting of these limits also allows the map to be calculated for the protein molecule only, and not the surrounding solvent, which has its density set to zero. The limits, referred to as the envelope, can be adjusted further—for example, to exclude adjacent protein molecules in the crystal lattice. The choice of the grid spacing will affect the quality of the map and the speed of the computation. At this stage, it is useful to monitor the results, not only by examination of the numerical output, but by checking the quality of the map using a computer graphics program such as "O".[40]

Refinement and model building
The quality of the electron density map may be improved by refinement.[37] For example, it might be possible to use molecular averaging if more than one identical molecule or subunit is contained within the asymmetric unit. The geometrical operation that connects the subunits, the NCS, may be used to average the electron density of the subunits in a series of cycles in which the phases are continually improved. These are checked by comparing the observed structure factors with the calculated ones, and the discrepancy is expressed as a percentage known as the R factor. The phasing power is proportional to the square root of the NCS.[41]

When an electron density map (fig 6) appears by inspection to have sufficient quality to permit reliable location of amino acids then model building can commence. When inspecting initially for this quality, we are checking that the density can be traced almost continuously from one end of the protein molecule to the other. Small breaks in this continuity are often encountered—for example, as a result of localised regions of disorder, but should not pose a problem for model building as long as there is no ambiguity in the route that the protein chain takes. We are also checking that the quality and resolution permit reasonably reliable identification of the classes of amino acid side chain. Model building then involves the use of a computer graphics program to display the map and, with the aid of the protein sequence, the insertion of each residue. Often, we find that the start of the electron density does not correspond with the N-terminal of the protein. This is caused by conformational disorder of the terminal residues. If this is the case, then care must be taken to commence building at the correct point on the protein sequence. However, during model building an error in the sequence is sometimes detected.

The structure that has been built is output as a file in the format of the protein data bank, known as a PDB file, which has the format shown in fig 7. PDB files can be downloaded from the Brookhaven database and viewed,

even on a desktop computer, using a program such as RASMOL.[42]

Further refinement of the manually built model is carried out by computer programs that minimise the energy of the conformation of the protein using a dictionary of data on bond lengths and angles, while still abiding by the constraints of the calculated map. Cycles of refinement continue with successive improvement of the model and of the map until convergence—that is, the point of no further improvement—is reached. XPLOR[37] is frequently used for refinement, and again the quality of the model is expressed with an R factor.[43]

1 Blundell TL, Johnson LN. *Protein crystallography*. London: Academic Press, 1976.
2 Carter CW, Sweet RM, eds. Macromolecular crystallography, part A. *Methods Enzymol* 1997;**276**.
3 Carter CW, Sweet RM, eds. Macromolecular Crystallography, part B. *Methods Enzymol* 1997;**277**.
4 Luft JR, Arakali SV, Kirisits J, *et al*. A macromolecular crystallization procedure employing diffusion cells of varying depths as reservoirs to taylor the time course of equilibration in hanging drop and sitting drop vapour diffusion and microdialysis experiments. *Journal of Applied Crystallography* 1994;**27**:443–53.
5 Wilson LJ, Bray TL, Suddath FL. Crystallization of proteins by dynamic control of evaporation. *Journal of Crystal Growth* 1991;**110**:142–7.
6 Gernert KN, Smith R, Carter D. A simple apparatus for controlling nucleation and size in protein crystal growth. *Anal Biochem* 1988;**168**:141–7.
7 Carter CW, Carter CW. Protein crystallization using incomplete factorial experiments. *J Biol Chem* 1979;**254**:12219–23.
8 Carter CW, Baldwin ET, Frick L. Statistical design of experiments for protein crystal growth and the use of a precrystallisation assay. *Journal of Crystal Growth* 1988;**90**:60–73.
9 Jancarik J, Kim S-H. Sparse-matrix sampling—a screening method for crystallisation of proteins. *Journal of Applied Crystallography* 1991;**24**:409–11.
10 Hampel A, Labanauskas M, Connors PG, *et al*. Single crystals of transfer RNA from formylmethionine and phenylalanine transfer RNAs. *Science* 1968;**162**:1384–7.
11 Thomas DH, Rob A, Rice DW. A novel dialysis procedure for the crystallisation of proteins. *Protein Eng* 1989;**2**:489–91.
12 Song L, Gouaux JE. Membrane protein crystallisation: application of sparse matrices to the α-hemolysin heptamer. *Methods Enzymol* 1997;**276**:60–74.
13 Rossmann MG, Arnold E, Erickson JW, *et al*. Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature* 1985;**317**:145–53.
14 Smyth M, Tate J, Hoey E, *et al*. Implications for viral uncoating from the structure of bovine enterovirus. *Nat Struct Biol* 1995;**2**:224–31.
15 Fitzgerald PMD, Madsen NB. Improvement of limit of diffraction and useful X-ray lifetime of crystals of glycogen debranching enzyme. *Journal of Crystal Growth* 1986;**76**:600–6.
16 Smyth M, Fry E, Hoey E, *et al*. Preliminary crystallographic analysis of bovine enterovirus. *J Mol Biol* 1993;**231**:930–2.
17 Helliwell JR. *Macromolecular crystallography with synchrotron radiation*. Cambridge: Cambridge University Press, 1992.
18 Ealick S, Walter R. Synchrotron beamlines for macromolecular crystallography. *Curr Opin Struct Biol* 1993;**3**:725–36.
19 Hope H. Crystallography of biological macromolecules at ultra-low temperature. *Annual Review of Biophysics and Biophysical Chemistry* 1990;**19**:107–26.
20 Miyahara J, Takahashi K, Amemiya Y, *et al*. A new type of X-ray area detector utilising LASER stimulated luminescence. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1986;**A246**:572–8.
21 Moy J-P. A 200 mm input field, 5–80 keV detector based on an X-ray image intensifier and CCD camera. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment Methods* 1994;**A348**:641–4.
22 Gruner SM. X-ray detectors for macromolecular crystallography. *Curr Opin Struct Biol* 1994;**4**:765–9.
23 Dauter Z. Data collection strategy. *Methods Enzymol* 1997;**276**:326–44.
24 Leslie A. Data collection and processing. In: Sawyer L, Isaac N, Bailey S, eds. *Proceedings of the CCP4 study weekend*. Warrington, UK: SERC Daresbury Laboratory, 1993:44–51.
25 Kabsch W. Automatic indexing of rotation diffraction patterns. *Journal of Applied Crystallography* 1988;**21**:67–71.
26 Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in the oscillation mode. *Methods Enzymol* 1997;**276**:307–26.
27 Collaborative Computing Project Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D* 1994;**D50**:760–3.
28 Otwinowski Z. Isomorphous replacement and anomalous scattering. In: Wolf W, Evans PR, Leslie AGW, eds. *Proceedings of the CCP4 study weekend*. Warrington, UK: SERC Daresbury Laboratory, 1991:80–6.
29 Dickerson RE, Kendrew JC, Strandberg BE. The phase problem and isomorphous replacement methods in protein structures. In: Pepinsky R, Robertson JM, Speakman JC, eds. *Computing methods and the phase problem in X-ray crystal analysis*. Oxford: Pergamon Press, 1961:236–51.
30 Hengming K. Overview of isomorphous replacement phasing. *Methods Enzymol* 1997;**276**:448–61.
31 Terwilliger TC, Eisenberg D. Isomorphous replacement—effects of errors on the phase probability-distribution. *Acta Crystallogr A* 1987;**A43**:6–13.
32 Rossmann MG, ed. *The molecular replacement method*. New York: Gordon and Breach, 1972.
33 Rossmann MG, Blow DM. The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallogr* 1962;**15**:24–31
34 Rossmann MG. The molecular replacement method. *Acta Crystallogr A* 1990;**A46**:73–82.
35 Taylor CA, Morley KA. An improved method for determining the relative positions of molecules. *Acta Crystallogr* 1959;**12**:101–5.
36 Fujinaga M, Read RJ. Experiences with a new translation-function program. *Journal of Applied Crystallography* 1987;**20**:517–21.
37 Brunger AT. *X-plor manual*. New Haven: Yale University Press, 1992.
38 Navaza J, Saludjian P. AMoRe: an automated molecular replacement program package. *Methods Enzymol* 1997;**276**:581–94.
39 Ten Eyck LF. Crystallographic fast fourier transforms. *Acta Crystallogr A* 1973;**A29**:183–91.
40 Jones TA, Zou J-Y, Cowan SW, *et al*. Improved methods of building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 1991;**A47**:110–19.
41 Arnold E, Rossmann MG. Effect of errors, redundancy and solvent content in the molecular replacement procedure for the structure determination of biological macromolecules. *Proc Natl Acad Sci USA* 1986;**83**:5489–93.
42 Molinaro M, Grossman G. UCB enhanced RASMOL version 2.6. MultiCHEM Facility, University of California, Berkeley. © 1995, 1996 UC Regents/ModularCHEM Consortium.
43 Brunger AT, Kuriyan J, Karplus M. Crystallographic R-factor refinement by molecular dynamics. *Science* 1988;**235**:472–5.