

[4] X-Ray Diffraction of Protein Crystals

By F. H. C. CRICK

Introduction

The X-ray study of protein crystals is a difficult and highly specialized field. Eventually, when the structure of a few proteins has been unraveled, the results will be of vital interest to enzymologists, since they should give information about the spacial arrangement of the "active site" of the enzyme. But meanwhile X-ray methods will have only a secondary interest for enzymologists.

The equipment required for X-ray work is expensive and will normally not be available in an enzymological laboratory. This chapter has therefore been written on the assumption that the X-ray work will be carried out in collaboration with a professional crystallographer. It aims to show both the enzymologist and the X-ray crystallographer—who may not be familiar with the X-ray work on proteins^{1,2}—the kind of information that can usefully be obtained in a fairly short time. A knowledge of normal crystallographic terms and techniques³ is therefore assumed, but all the special techniques required are described in detail.

The most useful information that can be obtained from an X-ray examination of protein crystals is a rather good value for the molecular weight. If crystals of a reasonable size are available, a figure correct to a few per cent can be obtained in a few days. This method has been neglected in the past, and enzymologists might well find it helpful. Occasionally it can be shown that two proteins from different sources have very similar structures, and more detailed studies may, in favorable cases, yield some information on the shape of the molecule.

Identification

It might be thought that the very detailed X-ray picture of a protein crystal would be like a set of fingerprints and that it would provide a good method of identifying a protein. This is not so, for two reasons. First, the same protein may form, under slightly different conditions, quite different crystals, having radically different unit cells. Polymorphism is

¹ B. W. Low, in "The Proteins" (Neurath and Bailey, eds.), Vol. 1, Part A, p. 235. Academic Press, New York, 1953.

² J. C. Kendrew, in "Progress in Biophysics" (Butler and Randall, eds.), Vol. 4, p. 244. Pergamon Press, London, 1954.

³ See, for example, C. W. Bunn, "Chemical Crystallography." Oxford, New York, 1945.

the rule rather than the exception in protein crystals. With present techniques there is no method of deciding from the X-ray pattern that such different crystals contain the same protein. Second, proteins which are known to be similar, but are in fact different from the point of view of the protein chemist, can form identical crystals and give substantially identical X-ray pictures.⁴ This is because the X-ray picture depends mainly on the broad architecture of the protein molecule, and small changes, such as the substitution of one amino acid for another, make such very slight changes to the X-ray intensities that in practice they can hardly be detected.

It should be realized, however, that if two "different" proteins (e.g., from different sources) are found to give identical crystals, having very similar X-ray patterns, it is certain that the broad features of the structure of the two proteins are very similar. Such a case occurs, for example, among the myoglobins, in which it is found that crystals obtained from porpoise myoglobin and from three different species of whale give essentially the same X-ray diagrams.² Thus in special circumstances protein crystals can give limited information about the identity of proteins.

Crystallization

Nothing will be said about the search for the conditions under which crystals can be obtained from amorphous protein, as the protein chemist is familiar with this problem, and there appear to be no general rules to act as guides. It will be assumed that crystals of some sort have already been obtained.

Protein crystals must be a certain size, however, if they are to be studied conveniently by X-rays. The optimum is about 0.3 to 0.5 mm. in all directions, though preliminary work can sometimes be done on crystals as small as 0.1 mm., or even smaller. The protein chemist is usually satisfied if his crystals are large enough to be seen in the microscope, so it will often be necessary as a first step to grow larger crystals, especially as crystals as large as 1 mm. are required for very accurate molecular weight determination. There are many ways of doing this, but in essence they all consist in growing as few crystals as possible, and in growing these few rather slowly, so that fresh crystal nuclei cannot easily form.

Proteins vary considerably: some may form quite large crystals without any special precautions, whereas others produce such small crystals that considerable effort must be expended to get them large enough. An example of the first is chick lysozyme; with this enzyme from an amorphous precipitate large crystals will often form overnight. Trypsin inhibitor (from bovine pancreas), on the other hand, requires some weeks

⁴ M. F. Perutz, A. M. Liquori, and F. Eirich, *Nature* **167**, 929 (1951).

to form adequate crystals. The difference may depend on whether or not there is any marked difference in the solubility of the amorphous and the crystalline material.

If fairly large quantities of the protein are available, it is simplest to make up a series of test tubes in which the protein concentration or the solvent composition is varied in very small steps from the conditions under which small crystals are formed. Those tubes, which initially show no crystals, may produce quite large crystals if left for a few days. If little protein is available a better method is to adjust the solvent so that from a concentrated protein solution a very slight precipitate is obtained. This should be carefully filtered to give a saturated solution free from crystal nuclei, and then very slowly concentrated further by evaporation. This can conveniently be done by placing a shallow layer of the solution, say a few millimeters thick, in the bottom of a flat-bottomed tube. The tube is then sealed with a rubber bung containing a length of capillary (about 2 cm. long and 1 mm. internal diameter) to restrict the rate of evaporation, placed in a constant temperature desiccator, and left for some days or even weeks. If the crystals produced are too small the procedure should be repeated with a finer capillary tube or with a slightly less concentrated solution.

Some workers favor the use of seed crystals. Large crystals of ribonuclease have been successfully grown by adding to the lyophilized protein powder a minute amount of crushed dried crystals to act as seeds and then adding solvent. The smaller the amount of seeds, the larger are the resulting crystals.

Other methods are: the slow concentration of the protein solution by forcing it through a semipermeable membrane; a controlled and gradual change of temperature (if the solubility varies with temperature); or slow salting out by diffusing more salt into the solution through a semipermeable membrane (if alcohol forms part of the solvent, it may be allowed to diffuse in through the vapor phase). It may even be possible to dispense with a membrane and merely place two layers of liquid (one of which contains the protein and the other alcohol or salts) one above the other, the lighter on top of the denser. Crystals may form on the walls of the test tube near the interface.

Since these processes are all slow, trouble may be encountered from bacterial growth or from molds, unless the solvent is one which resists such contamination, such as strong salt solutions or alcohol-water mixtures. It may be necessary to pass the protein solution through a bacterial filter and work under sterile conditions.

It sometimes happens that, although crystals grow readily, they do so in an unfavorable habit, such as long thin needles or, even worse, as

hedgehogs of needles. A protein crystal can always be cut, so extreme length is no handicap as long as the smallest dimension is not too small, which is unfortunately often the case when needles form. Whereas the problem of increasing the size of the crystal can be tackled in a rational manner, that of altering the habit cannot be, and the only method is to try empirical alterations in the conditions of crystallization, by varying the pH, by altering the salt used, or the type of alcohol if from alcohol-water mixtures, or by adding small amounts of heavy metals. These may, of course, not only alter the habit but produce a new type of crystal. It is for this reason that the conditions under which the crystallization is carried out should always be carefully reported.

The Nature of Protein Crystals

Protein crystals differ from almost all other crystals in containing within themselves a large amount of the solvent from which they were grown, much of which is still in the liquid state. Typically about half the volume of the crystal is solvent. If such crystals are exposed to the air, liquid evaporates and the crystals shrink. This usually causes considerable deterioration of the optical properties of the crystal and the quality of the X-ray pattern. The crystals are often deformed and may split. Such crystals, although still containing some water, are referred to as "dry," and the original crystals, as grown, as "wet." If the humidity is changed in a controlled manner it is sometimes found^{5,6} that between the wet and the dry stages one or more "hydration stages" or "shrinkage stages" exist, each of which, like the wet stage, has well-defined cell dimensions and a typical X-ray pattern. The dimensions of the dry stage, which is usually rather disordered, are less precisely defined than the wet stage and may vary a little, depending on the exact method of drying.

Because of the liquid state of the solvent molecules inside the crystal it is often possible for small molecules to diffuse into the crystal without disturbing the arrangement of the protein molecules to any appreciable extent. The salt concentration of the crystal may be changed in this way, or, if the crystal is grown from alcohol-water mixtures, the concentration of the alcohol may be altered, or a different alcohol allowed to penetrate the crystal. Even comparatively large molecules, like dyes, can sometimes diffuse into protein crystals, and it would not be surprising if an enzyme inhibitor could be made to combine with the enzyme while the latter was in the form of a crystal. Naturally it may happen that such a combination will change the unit cell somewhat, but it is quite possible that the unit cell dimensions will hardly alter at all.

⁵ M. F. Perutz, *Trans. Faraday Soc.* **42B**, 187 (1946).

⁶ H. E. Huxley and J. C. Kendrew, *Acta Cryst.* **6**, 76 (1953).

Optical Examination

The optical examination must be done on the wet crystals immersed in the mother liquor, as the optical quality of the dry crystals is usually very poor. This restriction makes some of the usual optical measurements very awkward, and in practice the only properties measured are those which can be studied with reasonable convenience. It is often difficult, for example, to view a rodlike crystal end on. The infrared and ultraviolet dichroism of protein crystals have been measured in a few cases, but these techniques are outside the scope of this discussion. If the protein is conjugated with a molecule which absorbs in the visible region, the crystal should be examined for dichroism.⁷

Mounting Crystals

To obtain X-ray pictures of a wet protein crystal it is obviously necessary to mount the crystal in an enclosed space. This is usually done by placing it in a thin-walled capillary, made of low-absorption glass, which is then sealed at both ends. Such capillaries, typically 1½ mm. in diameter and several centimeters long, can be made by heating a suitable glass tube in a flame and then pulling it out very rapidly. For accurate work they should have very thin walls (say 20 μ), but for a preliminary examination thicker walls can be tolerated. Suitable capillaries can be obtained commercially from Paul Raebiger, Berlin-Spandau (British Sector) Franzstrasse 43, Germany.

Some workers find it an advantage to coat the inside of the capillary with a hydrophobic film, as this makes for a cleaner mount. It also reduces any effects due to the alkalinity of the glass. A technique for doing this has been worked out by King.⁸ To commercial Desicote add 5% water-saturated ether, and dip the capillary into this mixture. This step produces partial polymerization on the glass surface. Then bake in an oven (about 150°) to remove the solvents, and complete the polymerization by immersing the capillary in hot water (100°) for a few minutes. The advantage of this technique is that it removes all traces of HCl, which otherwise may produce a low pH.

The crystal is mounted with as little mother liquor in actual contact with it as possible, but a few drops of mother liquor are included elsewhere in the capillary to maintain the correct humidity. The crystal will usually cling to the wall of the capillary of its own accord. The mounting of such crystals is tricky, but the technique can be mastered in a day. The usual method is to insert the crystal, inside a drop of mother liquor, into the capillary, and then manoeuvre the drop until the crystal is in the

⁷ M. F. Perutz, *Acta Cryst.* **6**, 689 (1953).

⁸ M. V. King, *Acta Cryst.* **7**, 601 (1954).

desired region. The excess liquid is then withdrawn (for example, with a special fine pipet), and the crystal is gently pushed with a fine glass fiber into the required orientation on the capillary wall. Finally a further drop of mother liquor is introduced at a distance from the crystal, and the tube is sealed at both ends, care being taken to avoid heating the crystal or the drop of mother liquor in the process. Some workers use black wax (picien) for this purpose. This is unsuitable if there is alcohol in the mother liquor, and in such cases the capillary may have to be sealed with a flame. It is sometimes easier to seal off one end of the capillary before starting to mount the crystal. Whatever method is used, the seal must be perfect or the crystal will dry out.

Each experimenter has his own favorite variation on this basic method.^{8,9} As protein crystals are often rather soft (though this varies from protein to protein), considerable care must be taken in handling them, but, although the initial efforts are apt to be discouraging, a little experience soon overcomes this.

Dry crystals can usually be mounted in the more customary manner by sticking them gently to a fine glass rod. They are often less fragile than wet crystals. Alternatively a crystal may be mounted wet in a capillary, which is left unsealed so that the crystal dries in position, though there is the danger that such crystals sometimes drop off the capillary walls after drying.

It is often an advantage to study the same crystal in the wet and the dry state, as this may help to decide which axes of the dry cell correspond to those of the wet. After the studies on the wet crystal have been completed the capillary can be gently cracked, the crystal realigned if necessary, and X-ray pictures taken at intervals until the crystal has become dry and shrinks no further. Typically this will take some hours, or even a day or so.

X-Ray Technique

Beyond the problem of mounting, the X-ray techniques employed are fairly standard. As the unit cell is large, much longer exposures are necessary than for smaller organic molecules. Reflections rarely extend to spacings shorter than 2 Å., whereas cell dimensions greater than 100 Å. are quite common. The region of interest of the reciprocal lattice is therefore nearer the origin than is customary. For this reason a precession camera is a very favorable tool for studying protein crystals. The experimenter should use a fine collimator, ideally little bigger than the size of the crystal, and as small a backstop as possible. A very good arrangement

⁸ J. Boyes-Watson, E. Davidson, and M. F. Perutz, *Proc. Roy. Soc. A* **191**, 83 (1947).

consists of a hanging backstop, with some method of adjusting its position, placed 2 cm. in front of the film, as this reduces air scattering near the backstop. Certainly an effort should be made to observe the lowest orders of the diffraction pattern, as important information may otherwise be missed.

If only very small crystals are available it may pay, for preliminary work, to reduce the film distance from, say, 6 cm. to as little as 2 cm., as this will decrease the exposure time considerably. It will then usually be necessary to use a very fine collimator, or alternatively a lead pinhole, of about the same diameter as the crystal, mounted immediately in front of the crystal. If this pinhole is tapered (so that it acts both as a pinhole and a guard pinhole combined), diffraction from the lead itself can be avoided. Such pinholes can easily be produced with a little care. The backstop must be correspondingly reduced in size. Naturally the small film distance reduces the accuracy with which the cell dimensions can be measured, but the space group and the approximate dimensions can be obtained without prohibitively long exposure times.

Protein crystals, especially when wet, deteriorate with long exposure to X-rays. This deterioration takes the form of a fading away of the X-ray intensities. Typically it occurs after exposure times of about 100 or 200 hours.

In obtaining the molecular weight from a *dry* crystal the accuracy of the method is usually limited by the precision with which the X-ray spacings can be measured, and the experimenter should adjust his technique to try to obtain the greatest possible accuracy. Owing to the disorder in such crystals this cannot be very high, but $\pm 1/2\%$ on cell dimensions is not impossible in favorable cases, especially if several crystals are measured.

The X-Ray Picture

There are two features of protein X-ray pictures, arising out of the small reciprocal spacings, which are worth mentioning although their interpretation is straightforward. A "still," taken with the X-ray beam parallel to an axis of the unit cell, will show, on a flat film, concentric circles of reflections. These are clearly due to the intersection of successive planes of the reciprocal lattice with the sphere of reflection, and from the diameters of such circles the length of that particular axis of the unit cell can be computed. The circles show up because of the dense packing of these planes with reciprocal lattice points. If the axis is mis-set a few degrees the circles become eccentric, and the degree of this eccentricity can be used to estimate the degree of mis-setting. This is the standard

method for aligning protein crystals, after a preliminary optical alignment. The final adjustment may be confirmed by an oscillation or precession photo of 2° or 3° amplitude.

The second feature is that a nonprimitive lattice can often be recognized at a glance from inspection of the X-ray photograph, say a 5° oscillation, because the reciprocal cell being so small, its arrangement can be visualized, and it is often unnecessary to index the reflections and use the well-known rules for systematic absences. In their place the following equivalent rules can be followed:

A body-centered reciprocal cell corresponds to an all-face-centered real cell.

An all-face-centered reciprocal cell corresponds to a body-centered real cell.

A side-centered reciprocal cell corresponds to a side-centered real cell.

Space Group Determination

The space group of a protein can usually be determined with a minimum of ambiguity, since the polypeptide chain—the backbone of the protein structure—contains many asymmetric carbon atoms, all of which are believed to occur only in the L-configuration. Because of this, mirror planes, glide planes, and centers of symmetry are impossible. It is therefore well worth while to look for the systematic absences which distinguish screw axes from rotation axes, and the preliminary study of a protein crystal should not be regarded as complete until this has been done and the space group determined. It may not be possible to do this so convincingly for the dry crystal, but in all cases investigated so far the space group of the dry cell appears to be the same as that of the wet. At least one case of pseudosymmetry¹⁰ of the wet crystal has been found (in which the real symmetry was P2₁ but the pseudosymmetry P2₂2₁), and this possibility should be watched for. For this reason it is useful to take precession pictures of the principal zones of the wet crystal, since any pseudosymmetry or other striking features of the X-ray pattern can often be seen at a glance.

From the space group and the unit cell volume the volume of the asymmetric unit can be obtained. If the approximate molecular weight¹¹ (M) of the protein is known (say, to within $\pm 30\%$) from other methods, the number of molecules in the asymmetric unit can be calculated. If the molecular weight of the asymmetric protein unit is denoted by A , the volume of the dry unit cell by V (units A.³), and the number of asymmet-

¹⁰ J. C. Kendrew and I. F. Trotter, *Acta Cryst.* **7**, 347 (1954).

¹¹ The term molecular weight is used throughout in the loose sense commonly employed.

ric units in the unit cell by n , we may write

$$A = \frac{1}{k} \cdot \frac{V}{n}$$

It is found empirically that k is never far from 1.5, the observed values lying between 1.35 and 1.75. This rule is possible because most proteins have about the same density, and about the same degree of hydration in the dry cell. With $k = 1.5$, therefore, the number of molecules in the asymmetric unit is the nearest integer, or reciprocal integer, to A/M . The term reciprocal integer is included to cover the case where the asymmetric unit is, say, one-half or one-third the "molecular weight." For example, in horse hemoglobin the molecular weight normally observed in solution is about 67,000, but the asymmetric unit in the usual crystals (space group C2) is half this,⁹ showing clearly that the "molecule" consists of two essentially identical halves related by a diad axis. Note that such an internal symmetry of the molecule is not necessarily shown by the space group. Another type of horse hemoglobin crystal⁹ shows an asymmetric unit of 67,000, and there are numerous cases where the asymmetric unit of a protein crystal consists of two or more molecules.

The value of $k = 1.5$ will not, of course, apply to a protein containing a substantial quantity of a nonprotein component, such as ferritin. In such cases more precise methods must be employed, and the density and nonprotein content of the crystals must be measured.

The Implications of Protein Axes

In considering the relationship between the molecule in solution and the molecule in the crystal the difference in the implication of screw axes and rotation axes should be realized. In a space group without rotation axes and with one molecule in the asymmetric unit there is no particular association of the molecules of the crystal into *small* groups. One would not expect, therefore, to find the molecules in the mother liquor forming dimers, for example. If the space group contains a *rotation* axis, however, say a twofold axis, there is a very real sense in which the molecules of the structure are related in pairs, and it would not be surprising if dimers were found in the corresponding mother liquor. Insulin provides a very pretty example of this.^{12,13} The usual crystals have the space group R3, which has a threefold axis, and under rather similar pH conditions the molecule is found to have a molecular weight not far from three times the molecular weight of the asymmetric unit of 12,000. At low pH different forms of the crystal are found, one of which has the space group P2₁2₁2₁ which

¹² H. Gutfreund, *Biochem J.* **50**, 564 (1952).

¹³ For a full discussion and references, see J. T. Edsall, in "The Proteins" (Neurath and Bailey, eds.), Vol. 1, Part B, p. 717. Academic Press, New York, 1953.

contains only screw axes, and it is therefore not surprising that at low pH the molecular weight in solution is found to be 12,000. Thus the space group of the crystal may give a very strong hint as to which condition in solution will show an unaggregated molecule. Notice here, as mentioned earlier, that the fact that the asymmetric unit is 12,000 gives no evidence as to whether the "true" molecular weight of insulin is half this.

Molecular Weight Determination

It will be realized from what has been already said that the molecular weight as determined from the crystal alone may be a multiple of the "true" molecular weight and a multiple or a submultiple of the molecular weight found in solution. This fact will not be referred to again, and in what follows the term "molecular weight" should be read as "crystal molecular weight."

A rough value of the molecular weight can be obtained from the volume of the dry unit cell by the empirical rule described earlier and with $k = 1.5$. Note that this rule seems to apply irrespective of the amount of salt or alcohol in the wet crystal. This rule should be accurate to -10 to $+20\%$ and will often be within -5 to $+10\%$. There are, in fact, few other methods of molecular weight determination which will give a value as accurate as this from a single simple experiment.

To obtain a better value it is necessary to measure the density of the crystal when it is in some well-defined state. This will allow the mass of the unit cell to be obtained from the observed volume. From this figure for the mass must be subtracted the mass of the nonprotein components of the unit cell—namely, the volatile material, such as water or alcohol, and the nonvolatile nonprotein material (such as salt) remaining after drying. The former is found by measuring the percentage loss of weight in taking the crystal from the original state (in which the density can be measured) to the vacuum-dried condition. The latter must then be found by chemical methods. What remains after subtracting these corrections is considered to be the molecular weight of the protein.

There are three obvious choices for the "well-defined" state of the crystal: the wet crystal, the air-dried crystal, and the vacuum-dried crystal. Experience has shown¹⁴ that the last of these is not suitable. The effect of vacuum drying is to remove the solvent from the spaces between the protein molecules and to leave holes. If the density of such a crystal is determined (by immersion in a fluid), the fluid may enter these holes, and the observed density may not correspond to the density of the unit cell but be more nearly that of the protein molecule itself. One must therefore start from either the wet crystal or the air-dried crystal.

¹⁴ B. W. Low and F. M. Richards, *J. Am. Chem. Soc.* **76**, 2511 (1954).

At the moment of writing it is not clear which of these methods is the better for the enzymologist. In theory the wet crystal should yield a more accurate figure, but in practice the technique is more difficult, and this may perhaps lead to errors. On the other hand, the air-dried method can be carried out with rather small crystals, whereas the wet-crystal method requires large crystals. Both methods will therefore be described. They are followed by detailed accounts of the techniques of density determination and composition measurements.

Air-Dried Crystals

Air-dried crystals are not precisely defined for two reasons. First, the dimensions of the unit cell and the degree of disorder usually depend somewhat on the method of drying. Slow drying usually produces a more ordered crystal. Second, the cell dimensions and the water content vary with the humidity, though not very greatly—certainly much less than they do for wet crystals. Reasonable care should therefore be taken to keep the humidity the same for the different sets of measurements—on the cell dimensions, on the density, and on the nonprotein content. It is convenient to measure the cell dimensions first and to carry out the other measurements at the humidity of the atmosphere at the time of the X-ray measurements. It should be not far from 40%.

The rate of drying of a crystal can be reduced by obvious methods, such as restricting the entry of the dry air. A less obvious way is to immerse the wet crystals in xylene.⁹ The water of the crystal is only slightly miscible with xylene, and because of this it diffuses out of the crystal at a very reduced rate.

The perfection of an air-dried crystal depends partly on the nature of the crystal. A given protein may have one unit cell which, when dried, still gives X-ray reflections at spacings of 4 Å., whereas a second type of unit cell of the same protein may under similar conditions give no reflections beyond 7 Å. (the latter amount of disorder is the more customary for dry crystals). Thus, if several forms of crystal exist, it is worth while trying them all.

High concentrations of salt are usually a disadvantage as, on drying, the salt in the crystal tends to become saturated and crystallize out separately. This not only disrupts the crystal somewhat but introduces complications in the other determinations. Thus the protein should be crystallized with as small a quantity of salt as possible. Alternatively the crystals may be grown under one set of conditions, but measured under another. For example, Low and Richards¹⁴ in their studies on dimeralbumin and β -lactoglobulin crystals grew their crystals from three-component systems containing protein, water, and either salt or alcohol.

After the crystals had grown they washed them thoroughly with a saturated aqueous solution of the protein before any measurements were made. Analysis showed that less than 0.5% salt or alcohol was left in the crystal.

Wet Crystals

The unit cell dimensions of a wet crystal can be measured with considerable accuracy, as can their density. The main difficulty comes in measuring the amount of solvent in the crystal. On the one hand, the crystal must be weighed without excess mother liquor clinging to it; on the other hand, it must not be allowed to dry. The wet crystal is exceedingly sensitive to humidity changes, and some solvent is lost very easily.

It is therefore almost impossible to measure the nonprotein content of the wet cell by using a mass of small crystals. The measurement must be made on large individual crystals. The details are described in a later section.

In certain cases it may be possible to calculate the protein content of the unit cell. This happens when the solvent has only one component, water. If one assumes that the apparent specific volume of the protein in the crystal is equal to the partial specific volume of the protein in dilute solution (which can be measured in the usual way) and that the water in the crystal has the same partial specific volume as free water, namely unity, then one can easily calculate the value of w , the grams of water in the crystal per gram of dry protein, from the formula of Adair and Adair:¹⁵

$$w = \frac{(D_p - D)}{(D - D_1)} \cdot \frac{D_1}{D_p}$$

where D = density of crystal.

D_1 = density of water.

D_p = reciprocal of the partial specific volume of the protein.

This version of the formula, the derivation of which is straightforward, is usually, but not always, correct for a wet crystal.¹⁴ It does not give correct results for an air-dried crystal.¹⁴ It should in any case be used with discretion.

The formula cannot be followed when the solvent has more than one component—for example, when it is a strong salt solution or an alcohol-water mixture. Measurements have shown⁵ that the concentration of salt solution inside a protein crystal is always less than that of the mother liquor, as if some of the water of the crystal was bound to the protein and not available to the salt.

¹⁵ G. S. Adair and M. E. Adair, *Proc. Roy. Soc. B* **120**, 422 (1936).

Density Measurements

There are three methods of density measurement: flotation, gradient column, and microbalance. The microbalance method,^{16,17} in which the apparent mass of the crystal is measured while it is immersed in its mother liquor, and the volume of the crystal computed from photographs, demands an instrument which, though not too difficult to make, will not normally be available to an enzymologist. It will therefore not be described further.

In the other two methods—flotation and gradient column—the protein crystal is immersed in a liquid which is immiscible with the liquid in the crystal, and whose density is the same as that of the crystal. In the flotation method the density of the liquid is adjusted by trial and error until the crystals neither float nor sink decisively. In the gradient column, which consists of a long vertical liquid column, the density of which increases steadily from top to bottom, the crystals fall to a level where the density of the liquid is the same as their own. In both these methods centrifugation may be used to make the method more sensitive or more rapid.

The flotation method requires little special comment. A series of tubes may be prepared containing mixtures of slightly different density in a regular sequence. Crystals are added to each, and the density is considered to be midway between those of the two adjacent tubes in which the crystals just sink and just float. Care should be taken to prevent evaporation of the liquids if they are volatile, and to avoid excessive temperature gradients which may set up convection currents. Alternatively the density of a single tube containing the crystals may be adjusted in steps until the experimenter judges that the crystals are very close to equilibrium. For wet crystals the precautions mentioned in the description of the gradient column should be followed—large crystals should be used, they should be wiped as quickly as possible, and the measurements taken within a few minutes.

The gradient column is based on that originally devised by Linderstrøm-Lang,¹⁸ and modified for rapid though less accurate work (0.1%) by Jacobson and Linderstrøm-Lang.¹⁹ As is well known, such columns are prepared by placing two liquids of different density in a long vertical column, the heavier below the lighter. A linear gradient of density develops near the interface. Manipulation of a plunger-type stirrer in a vertical tube can extend the gradient over the greater part of the column.

¹⁶ B. W. Low and F. M. Richards, *Nature* **170**, 412 (1952).

¹⁷ F. M. Richards, *Rev. Sci. Instr.* **24**, 1029 (1953).

¹⁸ K. Linderstrøm-Lang, *Nature* **139**, 713 (1937).

¹⁹ C. F. Jacobson and K. Linderstrøm-Lang, *Acta Physiol. Scand.* **2**, 149 (1940).

Since extreme accuracy is not required, a reasonably high temperature gradient is acceptable, and therefore no special precautions need be taken concerning accurate temperature control and vibration-free mounting as would be required for a more sensitive column. A column formed in this way is surprisingly stable and will maintain its gradient virtually unchanged for many months.

A suitable column for general use has been described by Low and Richards.²⁰ The column was 20 cm. high, made of bromobenzene and xylene, and covered a density range from 1.15 to 1.35. The position of the crystal could be observed to within 0.5 mm. from a millimeter rule supported alongside the column. The gradient was calibrated by means of drops of about 0.01 ml. and of known specific gravity, of potassium phosphate solutions, with density increments of 0.002.

Low and Richards²⁰ also used gradient columns set up in 10-ml. centrifuge tubes and spun for 1 or 2 minutes at 2500 r.p.m. During this time the heating was negligible. The useful range of length, developed by stirring, was about 5 cm. They did not assume that the density gradient was linear but bracketed the crystal position with two calibration drops. As the amount of material was small, the entire column was discarded after each measurement.

For proteins Low and Richards²⁰ used mainly 10-cm. columns of bromobenzene-kerosene or bromobenzene-xylene at 24°. Centrifugation proved unnecessary for the larger protein crystals. For wet crystals the column used had a density range of 1.10 to 1.20. The density of the crystals was measured after a standard time of 5 minutes. The large crystals reached an equilibrium position in about 3 minutes, but smaller crystals took up to 5 minutes.

It was found¹⁴ to make no observable difference during the first 10 minutes whether the column was made up with dry components or water-saturated components. After that time, however, crystals in both types of column showed a perceptible increase of density due to loss of water to column components. Within 24 hours β -lactoglobulin crystals in the dry column had attained densities of about 1.26, whereas those in the water-saturated ones gave values of about 1.16 to 1.17.

To prepare the crystal for the column the excess mother liquor was removed by the method described in the next section, and the time taken from the moment the crystal left its mother liquor to the moment that it was dropped in the column was recorded. For β -lactoglobulin it was found²⁰ that the lowest density was obtained with the largest crystal and with drying times of approximately 25 seconds. They²⁰ state: "Comparing crystals of different sizes (e.g. weighing 2.0 to 0.2 mg.) it was evident that

²⁰ B. W. Low and F. M. Richards, *J. Am. Chem. Soc.* **74**, 1660 (1952).

the larger the crystals used, the more accurate the measurement. This is clearly due to the greater surface area for water loss per unit weight of the smaller crystal. When we compared the measurements on two crystals of approximately the same weight (2.0 mg.), dried during different lengths of time (20 and 60 seconds), it was evident that a slight increase in drying time materially affects the observed density."

If the wet protein crystals contain alcohol it is necessary to plot the density of the crystal against time and extrapolate backward, as the alcohol is taken up by the liquids in the column. This will introduce some uncertainty into the value obtained for the density, but it is unlikely to be sufficient to influence the molecular weight value to any great extent.

For air-dried crystals Low and Richards²⁰ used mainly 10-ml. columns with a density range of 1.24 to 1.34, making up the column with air-dried components. The observed density is not a function of the time of the operations as it is with the wet crystals, and the measurements are therefore simpler.

Composition Measurements

In these experiments it is necessary to measure the percentage loss of weight in passing from the "well-defined" state to the vacuum dried state. This presents no special difficulties for air-dried crystals, but for wet crystals special precautions must be taken.¹⁴

A wet crystal should be removed from its mother liquor on a small spatula and placed on a piece of hard filter paper previously damped with a little of the mother liquor. It should then be turned over on several of its faces with a probe and quickly transferred to a weighing pan where its weight should be measured as a function of time. The value for the wet weight should be obtained by extrapolating to zero time, taken as the moment when the crystal was placed on the filter paper.

Although the exact value of the dry weight depends on the precise conditions used for drying, the variation appears to be small. For example, for β -lactoglobulin Low and Richards¹⁴ find that the dry weight after 24 hours at 25° over P_2O_5 is only about 1% greater than that at 100° over P_2O_5 . Thus the exact conditions of drying are of minor importance. The weight of the crystal after vacuum drying should be measured as quickly as possible and recorded as a function of time for 2 or 3 minutes. Extrapolation to zero time gives the dry weight. This correction is usually found to be very small.

Because of the necessity for wiping the crystals the experiments on the wet crystals must be done on large individual crystals. For air-dried crystals there is no such restriction, and there seems to be no reason

why a mass of small crystals should not be used, provided little amorphous material is present.

The correction for nonvolatile nonprotein material most obviously depends on its chemical nature. It will therefore not be discussed here. As already stated the salt concentration of the crystals should be reduced to as low a value as possible. There seems to be no case recorded in the literature in which an attempt has been made to obtain an accurate molecular weight from a protein crystal containing a large amount of salt. Air-dried lysozyme chloride, which contains only a small amount, was found²¹ to have 0.5% Na and 3.15% Cl, and a correction was applied to the molecular weight for the bound chloride, and for the effect of the NaCl, which formed small crystals on the surface of the protein crystal, on the observed density. These corrections were small, but in the case of a protein crystallized from, say, 2 M (NH₄)₂SO₄, the correction might be considerable. There would seem to be no reason, however, why the wet crystal method should not be reasonably accurate, provided there is a good micromethod for estimating the nonvolatile nonprotein component. The accuracy of the molecular weight found with the air-dried crystals might be considerably reduced, however, owing to the type of effect described above for lysozyme chloride.

Summary of Molecular Weight Methods

We can summarize the methods for determining molecular weight by classifying them according to the accuracy the experimenter wishes to obtain.

Accuracy ±15%

- Obtain the space group of the wet unit cell.
- Obtain the dimensions of the air-dried unit cell.
- Use the empirical rule to calculate the molecular weight.

Accuracy ±5%

- Obtain the space group of the wet unit cell.
- Obtain the dimensions of the air-dried unit cell as accurately as possible.
- Measure the density of the air-dried crystals (small crystals may be used).
- Measure the percentage loss of weight on vacuum drying (a mass of small crystals may be used).
- Measure the percentage mass of the nonvolatile nonprotein component (a mass of small crystals may be used).

²¹ K. J. Palmer, M. Ballantyne, and J. A. Galvin, *J. Am. Chem. Soc.* **70**, 906 (1948).

The accuracy will depend mainly on the precision with which the cell dimensions can be measured, unless the crystal contains much salt, which may reduce the accuracy further.

Accuracy as High as Possible

- Obtain the space group and cell dimensions of the wet unit cell.
- Measure the density of single, large, wet crystals.
- Measure the percentage loss of weight on drying of single, large, wet crystals.
- Measure the percentage mass of the nonvolatile nonprotein component.

The accuracy will depend mainly on the accuracy of the last two measurements.

If the solvent is water it is sometimes possible to perform only the first two operations and find the water content from the formula of Adair and Adair, but this should not be done if the highest accuracy is required.

Further Work

Further work may be undertaken for two reasons: first, to find the arrangement of the molecules in the crystal and the general shape of the molecule; second, to discover whether there are any features of the crystal which suggest that it should be studied more intensively.

It is not possible here to discuss the information, often rather limited, which may be obtained from a study of Patterson projections. Fortunately an excellent review, in nontechnical language, has been published by Kendrew,² and this should be read as a preliminary to any further studies on the protein crystal, as should the very comprehensive review by Low.¹

The Shape of the Molecule

X-Rays will "see" the shape of the molecule only if there is a good contrast between the average electron density of the protein and that of the solvent. The average electron density of protein is usually about 0.43 electrons/Å.³, whereas that of hydrated protein is rather lower, say 0.41 or 0.40. The electron density of water is 0.33, electrons/Å.³, whereas that of 4 M (saturated) ammonium sulfate is about 0.41. Alcohol-water mixtures are less dense than water. Thus if the solvent is water, an alcohol-water mixture, or a salt solution of low concentration, its electron density will be well below that of the hydrated protein, and the very low-order X-ray reflections may give information about the shape of the

protein. If the solvent is a strong salt solution, its electron density may be rather close to that of the hydrated protein, and the very low orders are likely to be weak or absent. It may happen that for *very strong* salt solutions, or for salts of high molecular weight, the electron density of the solvent may exceed that of the hydrated protein by so much that the very low orders are strong because of the contrast the other way round. Such cases are not common, however.

Even if the salt has reduced the contrast between the salt solution and the protein it is possible to obtain valuable information by measuring how the low-order intensities alter with changes of salt concentration. This method has been utilized for hemoglobin by Bragg and Perutz.²²

Thus, if the electron density of the solvent is low, it is worth while to measure *all* the very low orders (spacings > 20 Å.) in the reciprocal lattice of the *wet* crystal—there will not usually be very many of them. If it is found that only two or three of the reflections are outstandingly strong it may be reasonable to calculate a very low resolution Fourier from these terms alone. Such a Fourier is likely to give information on the arrangement of the molecules in the cell, and also possibly something about the shape of the molecule. Thus the very great strength of the 001 and the 110 reflections in normal wet salt-free horse methemoglobin (space group C2) immediately gives an idea of the general arrangement of the molecules.²² It is advisable in such cases to measure the absolute intensities of the strong reflections, as this makes the interpretation more certain. More detailed studies on hemoglobin,²² with data from various shrinkage stages, have made it possible to obtain the general shape of the hemoglobin molecule at low resolution.

In this work on hemoglobin the packing arrangements in various other crystals of hemoglobin were considered,²³ and helped to re-enforce the argument for the shape. The possibility of using packing arguments should therefore be kept in mind. For example, if there are sufficient polymorphic forms of the crystal it is not improbable that the *smallest* diameter of the molecule is about the size of the *smallest* (primitive) cell dimension found in any of the unit cells, but such arguments should be considered with caution.

Points to Look For

The following features are ones which might make it worth while to study the crystals further.

1. A very small unit cell, or one with one dimension very small (<20 Å.). A small unit cell will arise only with a protein of low molecular

²² W. L. Bragg and M. F. Perutz, *Acta Cryst.* **5**, 277 (1952).

²³ W. L. Bragg and M. F. Perutz, *Acta Cryst.* **5**, 323 (1952).

weight (say under 20,000) and with one molecule in the asymmetric unit. So far no protein crystal has been found with a *wet* dimension under 28 Å.

2. Indications that in the shrinkage stages the molecules remain parallel from one stage to another. This may be suggested by the constancy of two of the cell dimensions, as in horse methemoglobin (C2). This enabled Bragg and Perutz²⁴ to link together the signs of groups of reflections.

3. A unit cell with a very high percentage of solvent (>70%). Again it may be possible to use this fact to link signs together.

4. A unit cell with a few outstandingly strong reflections. If in the very low orders, this may suggest something about the arrangement and the shape of the molecules; if in the 9-A. or 5-A. regions, something about the internal structure of the molecule may be indicated.

5. A unit cell with striking features in its Patterson projections, in particular ones which suggest a hexagonal packing of vector rods. (See the discussion by Kendrew.²⁵)

6. Heavy atom isomorphous replacement. This is the most important possibility, and will be explained in more detail.

It now seems clear that for any worth-while progress to be made in the structure determination of a protein it must be possible to find two essentially identical unit cells, and one cell should contain a rather heavy atom. This has been brilliantly achieved by Green *et al.*²⁵ for the case of hemoglobin by reacting one SH group (of each asymmetric unit of 34,000 molecular weight) with one molecule of parachloromercuribenzoate. This mercury derivative has cell dimensions which are identical with those of the normal hemoglobin. The heavy atom can be loosely thought of as a sort of known reference point in the cell. It has thus not only been possible to find the position of the mercury atoms but also to obtain a picture of one projection of the molecule, though so far only with limited resolution.

The requirements of this method should be carefully noted. The atom must be heavy, or it will be swamped by the protein. As a rough rule an atom lighter than iodine is unlikely to be of use, except perhaps for a very small protein. Then the heavy atom must be attached to the *same site* in each protein molecule and be present in a high percentage of them. If a particular site is occupied in only half the molecules the heavy atom appears to the X-rays to be of only half the atomic number and therefore too light. It is not necessarily an advantage to attach more than one heavy atom per asymmetric unit, as the extra ones may only

²⁴ W. L. Bragg and M. F. Perutz, *Proc. Roy. Soc.* **A213**, 425 (1952).

²⁵ D. W. Green, V. M. Ingram, and M. F. Perutz, *Proc. Roy. Soc.* **A225**, 287 (1954).

confuse the situation. A number of lighter atoms (say bromine) will not be acceptable instead of a heavy atom except possibly in the case where they are bound together in a group of a definite shape. Finally, the unit cells of the protein with and without the heavy atom should have identical dimensions.

It can be seen that the chemical requirement of attaching one heavy atom, or at the most two or three, with high efficiency onto a few specific sites is a formidable one. Any protein to which it can be done becomes a very favorable object for X-ray studies. It may be possible, for example, to find the general shape of such a protein molecule rather quickly. It should not be forgotten that the active site of an enzyme offers great possibilities as a unique site of attachment, and that *an enzyme inhibitor containing a heavy atom* might rather easily fulfill the above requirements. If such a case comes to the knowledge of an enzymologist he should certainly bring it to the attention of a protein crystallographer.

Program of Work

In this section is set out a typical program of a preliminary examination as it might be made by a protein crystallographer. It is not suggested that the whole of this program should always be carried out, but it may serve as a guide to general practice.

1. Optical examination of the wet crystals.
2. Cell dimensions and space group of the wet cell and the air-dried cell.
3. Density measurements (as discussed).
4. Composition measurements (as discussed).
5. Precession pictures of three projections of the wet cell: Spacings out to 7 Å. or perhaps to 3 Å.
6. Calculation of the Pattersons of these projections.
7. Preliminary exploration of shrinkage stages.
8. Preliminary examination of the intensities of the dry cell.
9. Measurement of the three-dimensional low orders of the wet cell (in cases where the solvent has a low electron density).

Acknowledgments

I should like to thank Drs. Barbara W. Low and F. M. Richards for allowing me to read their manuscript prior to publication, and my colleagues at the Protein Structure Project for helpful criticism.