Eur. J. Biochem. *189*, 1 – 23 (1990)

Review

Current approaches to macromolecular crystallization

Alexander McPHERSON

Department of Biochemistry, University of California at Riverside, USA

(Received September 18, 1989) – EJB 89 1133

Given our current expertise, and the certain future developments in genetically altering organisms to produce proteins of modified structure and function, the concept of protein engineering is nearing reality. Similarly, our ability to describe and utilize protein structure and to define interactions with ligands has made possible the rational design of new drugs and pharmacological agents. Even in the absence of any intention toward applied use or value, the correlation of regulation, mechanism, and function of proteins with their detailed molecular structure has now become a primary concern of modern biochemistry and molecular biology.

At the present time, there are numerous physical-chemical approaches that yield information regarding macromolecular structure. Some of these methods, such as NMR and molecular dynamics, are becoming increasingly valuable in defining detailed protein structure, particularly for lower-molecular-mass proteins. There is, however, only one general technique that yields a detailed and precise description, in useful mathematical terms, of a macromolecule's structure, a description that can serve as a basis for drug design, and an intelligent guide for protein engineering. The method is X-ray diffraction analysis of single crystals of proteins, nucleic acids, and their complexes with one another and with conventional small molecules. Some inspirational examples of representative crystals are shown in Fig. 1 – 3.

In the past 20 years, the practice of X-ray crystallography has made enormous strides. Nearly all of the critical and time-consuming components of the technique have been improved, accelerated, and refined. X-ray crystallography today is not simply an awesome method used by physical chemists to reveal the vast beauty of macromolecular architecture; it is a practical, reliable, and relatively rapid means to obtain straightforward answers to perplexing questions.

X-ray diffraction data that once required years to obtain, can now be collected in a matter of weeks, even days in some cases. Computers of extraordinary speed and capacity are now common tools as are computer graphics systems of a versatility and cleverness that would have been unimaginable only a few years ago. Software, too, exists that is sophisticated yet friendly, flexible yet reliable, and readily available to anyone in need of it. The question, then, is where does the problem lie? What prevents the full utilization and exploitation of this enormously powerful approach.

The answer, of course, is that for application of the method to a particular macromolecule, the protein or nucleic acid must first be crystallized. Not only must crystals be grown, but they must be good quality crystals, crystals suitable for a high-resolution X-ray diffraction analysis. '*Aye, there's the rub*' as Hamlet might say, for in general, this is not an easy task. While some proteins may be trivially simple to crystallize, many others, invariably those of greatest personal interest, are elusive and stubborn [2].

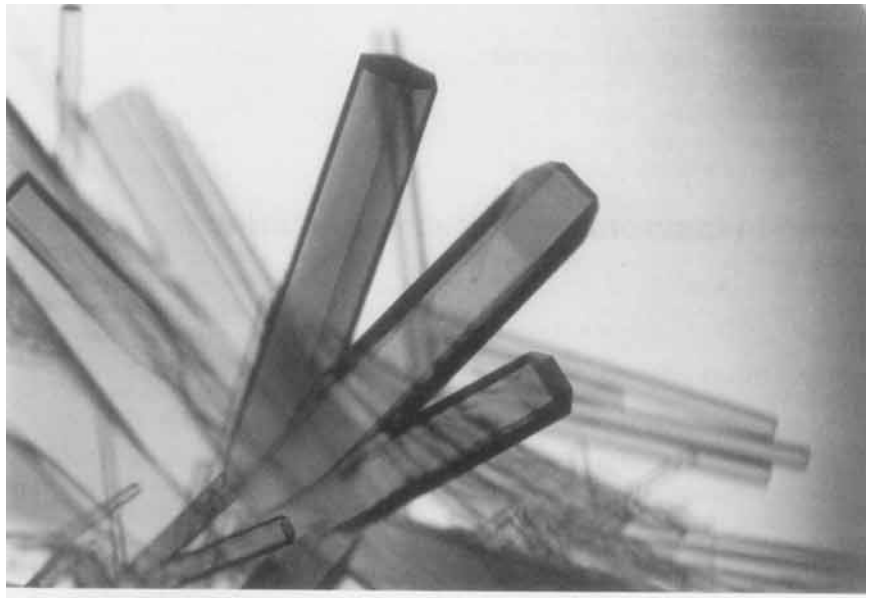
The reason that the crystallization step has become the primary obstacle to expanded structural knowledge is the necessarily empirical nature of the methods employed to overcome it [3 – 6]. Macromolecules are extremely complex physical-chemical systems whose properties vary as a function of many environmental influences such as temperature, pH, ionic strength, contaminants and solvent composition to name only a few. They are structurally dynamic, microheterogeneous, aggregating systems, and they change conformation in the presence of ligands (for a survey of protein structure and function, see [7 – 9]). Superimposed on this is the poor state of our current understanding of macromolecular crystallization phenomena and the forces that promote and maintain protein and nucleic acid crystals.

As a substitute for the precise and reasoned approaches that we commonly apply to scientific problems, we are forced, for the time being at least, to employ a strictly empirical methodology. Macromolecular crystallization is, thus, a matter of searching, as systematically as possible, the ranges of the individual parameters that impact upon crystal formation, finding a set or multiple sets of these factors that yield some kind of crystals, and then optimizing the variable sets to obtain the best possible crystals for X-ray analysis. This is done, most simply, by conducting a long series, or establishing a vast array, of crystallization trials, evaluating the results, and using information obtained to improve matters in successive rounds of trials. Because the number of variables is so large, and their ranges so broad, intelligence and intuition in designing and evaluating the individual and collective trials becomes essential.

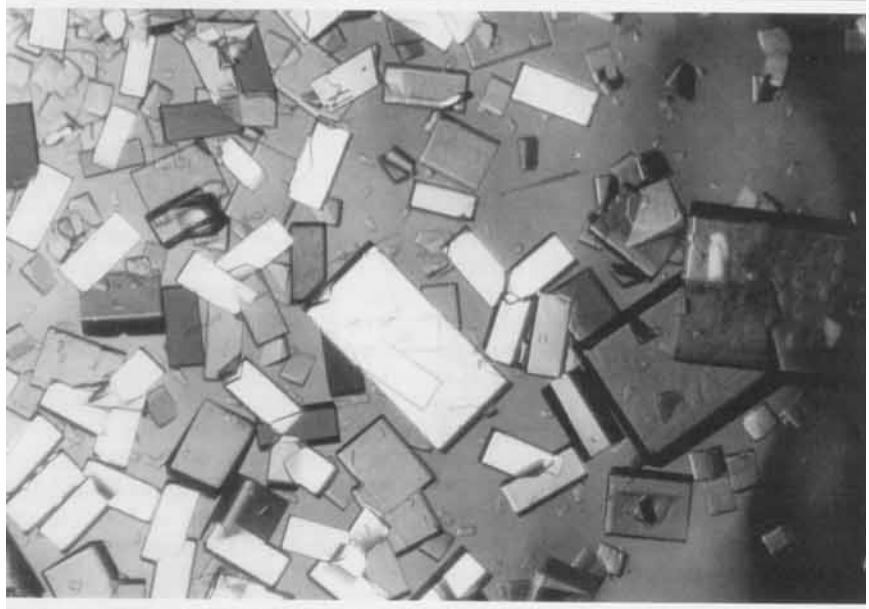
Crystals grow from supersaturated solutions

In a saturated solution, including one saturated with respect to protein, two states exist in equilibrium, the solid phase, and one consisting of molecules free in solution. At saturation, no net increase in the proportion of solid phase can accrue since it would be counterbalanced by an equivalent dissolution. Thus, crystals do not grow from a saturated solution. The system must be in a non-equlibrium, or supersaturated, state to provide the thermodynamic driving force for crystallization.

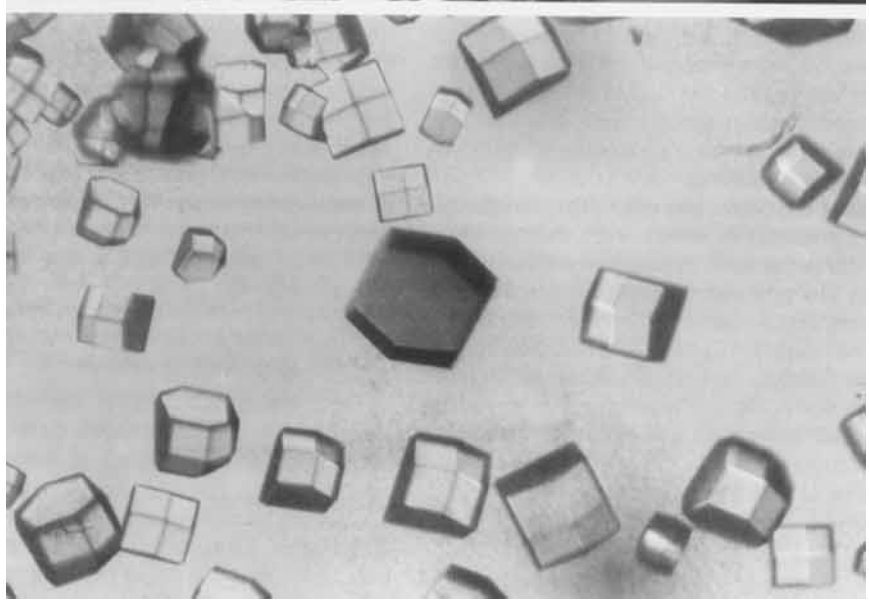
Correspondence to A. McPherson, Department of Biochemistry, University of California at Riverside, Riverside, California 92521-0129,USA



1



2



3

When the objective is to grow crystals of any compound, a solution of the molecule must by some means be transformed or brought into the supersaturated state whereby its return to equilibrium forces exclusion of solute molecules into the solid state, the crystal. If, from a saturated solution, for example, solvent is gradually withdrawn by evaporation, temperature is lowered or raised appropriately, or some other property of the system is altered, then the solubility limit will be exceeded and the solution will become supersaturated. If a solid phase is present, or introduced, then strict saturation will be reestablished as molecules leave the solvent to join the solid phase.

If no solid is present, as conditions are changed, then solute will not immediately partition into two phases, and the solution will remain in the supersaturated state. The solid state does not necessarily develop spontaneously as the saturation limit is exceeded because energy, analogous to the activation energy of a chemical reaction, is required to create the second phase, the stable nucleus of a crystal or a precipitate. Thus, a kinetic or energy barrier allows conditions to proceed further and further from equilibrium, into the zone of supersaturation. On a phase diagram [10, 11], like that seen in Fig. 4, the line indicative of saturation is also a boundary that marks the requirement for energy-requiring events to occur in order for a second phase to be established, the formation of the nucleus of a crystal or the nonspecific aggregate that characterizes a precipitate [12].

Once a stable nucleus has formed in a supersaturated solution, it will continue to grow until the system regains equilibrium. While non-equilibrium forces prevail and some degree of supersaturation exists to drive events, a crystal will grow or precipitate continue to form.

It is important to understand the significance of the term 'stable nucleus'. Many aggregates or nuclei spontaneously form once supersaturation is achieved, but most are, in general, not 'stable'. Instead of continuing to develop, they redissolve as rapidly as they form and their constituent molecules return to solution. A 'stable nucleus' is a molecular aggregate of such size and physical coherence that it will enlist new molecules into its growing surfaces faster than others are lost to solution; that is, it will continue to grow so long as the system is supersaturated.

In classical theories describing crystal growth of conventional molecules (see [13–16]), the region of supersaturation that pertains above saturation is further divided into what are termed the metastable region and the labile region [10–13], as shown in Fig. 4. By definition, stable nuclei cannot form in the metastable region just beyond saturation. If, however, a stable nucleus or solid is already present in the metastable region, then it can and will continue to grow. The labile region of greater supersaturation is discriminated from the metastable in that stable nuclei can spontaneously form. Further, because they are stable they will accumulate molecules and thus deplete the liquid phase until the system reenters the metastable, and ultimately, the saturated state.

An important point, shown graphically in Fig. 4, is that there are two regions above saturation, one of which can support crystal growth but not formation of stable nuclei, and the other which can yield nuclei as well as support growth. Now the rate of crystal growth is some function of the distance of the solution from the equilibrium position at saturation. Thus a nucleus that forms far from equilibrium and well into the labile region will grow very rapidly at first and, as the solution is depleted and moves back toward the metastable state, it will grow slower and slower. The nearer the system is to the metastable state when a stable nucleus first forms, then the slower it will proceed to mature.

It might appear that the best approach for obtaining crystals is to press the system as far into the labile region, supersaturation, as possible. There, the probability of nuclei formation is greatest, the speed of growth is greatest, and the likelihood of crystals is maximized. As the labile region is penetrated further, however, the probability of spontaneous and uncontrolled nucleation is also enhanced. Thus crystallization from solutions in the labile region far from the metastable state frequently results in extensive and uncontrolled 'showers' of crystals. By virtue of their number, none is favored and, in general, none will grow to a size suitable for X-ray diffraction studies. In addition, when crystallization is initiated from a point of high supersaturation, then initial growth is extremely rapid. Rapid growth is frequently associated with the occurrence of flaws and dislocations. Hence crystals produced from extremely saturated solutions tend to be numerous, small, and afflicted with growth defects.

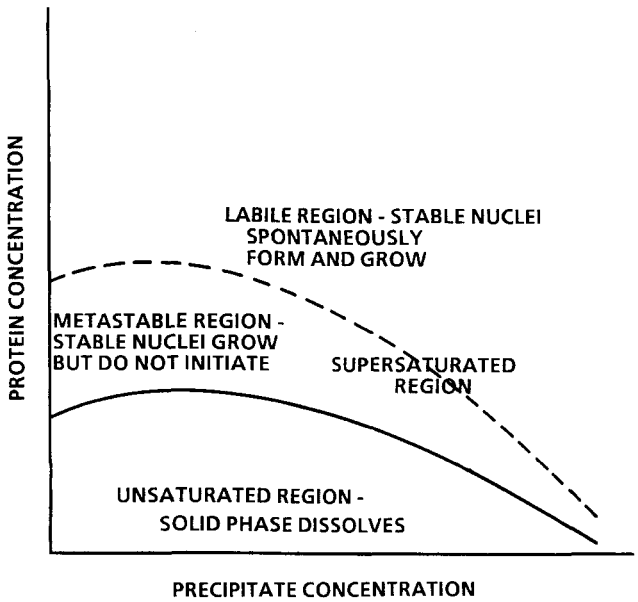
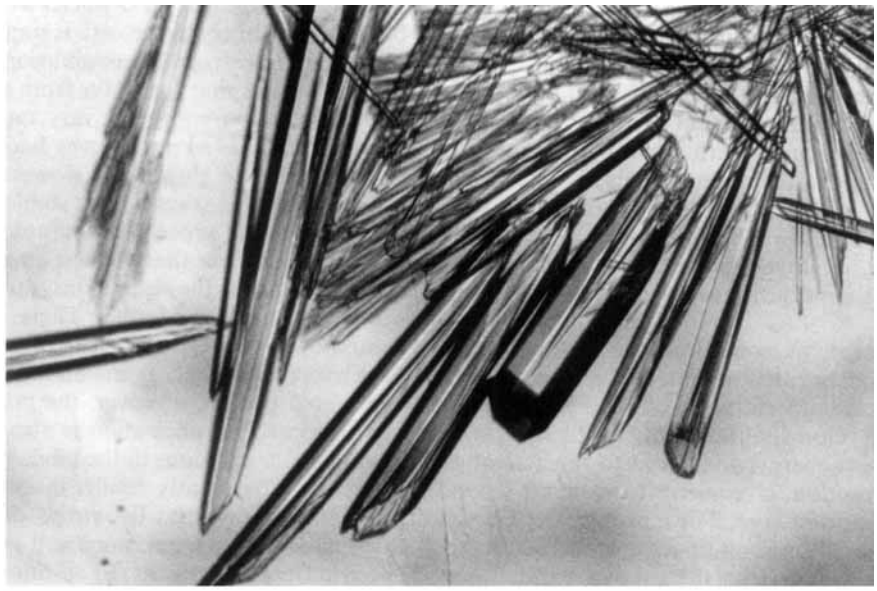


Fig. 4. *A phase diagram for a hypothetical protein showing its solubility as a function of precipitant concentration.* The solid line represents the maximum solubility or saturation curve, for the protein. Note that the protein is less soluble at very low and very high concentrations of the precipitant, corresponding to the 'salting in' and 'salting out' regions. The supersaturated region lies above the maximum solubility curve and is, in turn, demarcated by a boundary discriminating the metastable region of supersaturation from the labile region. In the labile region, crystal nuclei can both spontaneously form and grow, while in the metastable region they can only grow

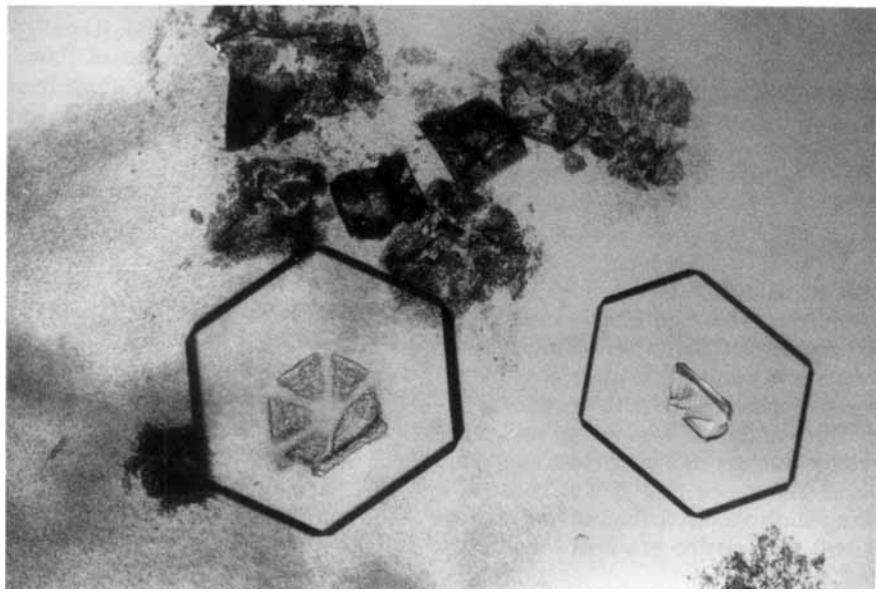
Fig. 1. *Photomicrograph of crystals of the heme-containing enzyme catalase from beef liver.* One of the earliest enzymes crystallized [1], it provides a good model for studies on protein crystal growth

Fig. 2. *Orthorhombic crystals of the major seed-storage protein from the jack bean (canavalin).* These crystals can grow to sizes of several millimeters on an edge. This protein is now being studied for its crystallization behavior in microgravity

Fig. 3. *Tetragonal crystals of hen egg lysozyme, one of the easiest proteins known to crystallize.* It has provided a source for many studies on the mechanisms of protein crystal growth



5



6

Fig. 5. *Crystals of the protein concanavalin B, first crystallized by J. B. Sumner in 1919.* These crystals are unusually stable and resistant to physical stress

Fig. 6. *Large hexagonal plates of the plant satellite tobacco mosaic virus, a protein/nucleic acid particle of over 1 MDa.* In spite of its great size, the virus crystallizes by procedures identical to those used for many proteins

In terms of the phase diagram, ideal crystal growth would begin with nuclei formed in the labile region but just beyond the metastable. There, growth would occur slowly, the solution, by depletion, would return to the metastable state where no more stable nuclei could form, and the few nuclei that had established themselves would continue to grow to maturity at a pace free of defect formation. Thus in growing crystals for X-ray diffraction analysis, one attempts, by either dehydration or alteration of physical conditions, to transport the solution into a labile, supersaturated state, but one as close as possible to the metastable phase.

Why crystals grow

The natural inclination of any system proceeding toward equilibrium is to maximize the extent of disorder, or entropy, by freeing individual constituents from physical and chemical constraint. At the same time, there is a thermodynamic requirement to minimize the free energy (or Gibbs energy) of the system. This is achieved by the formation of chemical bonds and interactions which generally provide negative free energy. Clearly the assembly of molecules into a fixed lattice severely reduces their mobility and freedom, yet crystals do form and grow.

It follows, then, that crystal nucleation and growth must be dominated by non-covalent chemical and physical bonds arising in the crystalline state that either cannot be formed in solution or are stronger than those that can. These bonds are, in fact, what hold crystals together. They are the energetically favorable intermolecular interactions that drive crystal growth in spite of the resistance to molecular constraint. From this it is clear that if one wishes to enhance the likelihood of crystal

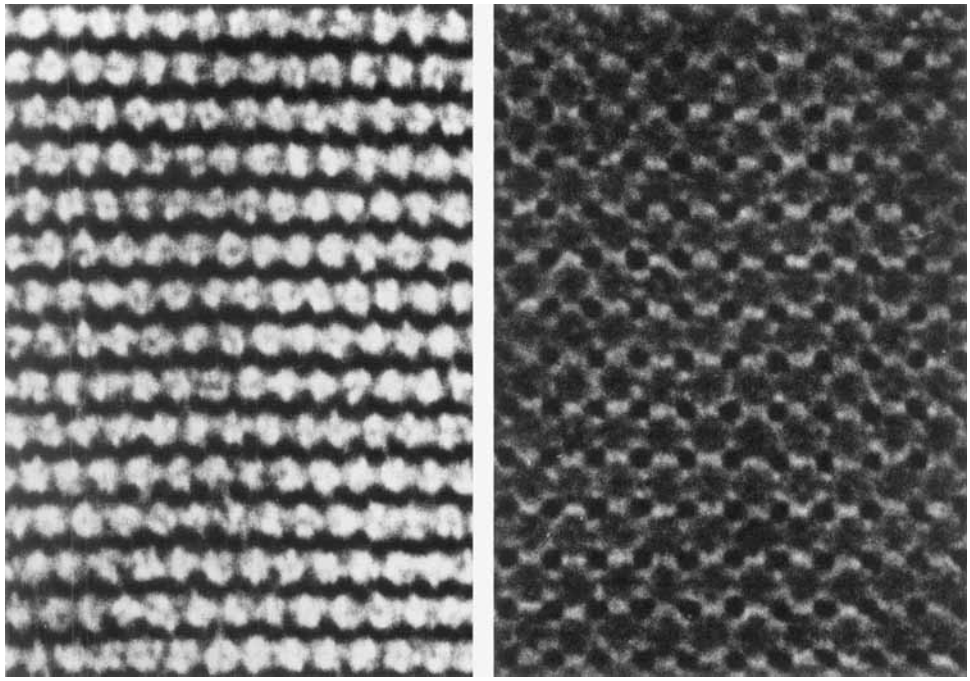


Fig. 7. *Electron micrographs of (left) a microcrystal of pig pancreas α-amylase and (right) a microcrystal of orthorhombic canavalin.* The microcrystals are negatively stained with uranyl acetate so that the light-colored areas represent protein while the darker areas reflect the presence of the stain. On the left the light-colored oval units are composed of the two molecules of α-amylase of M_r 50000 that comprise the asymmetric unit of these crystals. Note the extensive order that remains in the crystals and the clarity of the protein molecules even after dehydration and heavy metal staining. Note also the high proportion, nearly 50%, of the crystal that would be occupied by solvent, here replaced by the uranyl acetate

nuclei formation and growth, then one must do whatever is possible to ensure the greatest number of most stable interactions between the solute molecules in the solid state.

One may ask why molecules should arrange themselves into perfectly ordered and periodic crystal lattices, exemplified by those in Fig. 5, when they could equally well form random and disordered aggregates which we commonly refer to as precipitate. The answer is the same as for why solute molecules leave the solution phase at all: to form the greatest number of most stable bonds, to minimize the free energy, or free enthalpy, of the system. While precipitates represent, in general, a low-energy state for solute in equilibrium with a solution phase, crystals not precipitates are the states of lowest free energy.

A frequently noted phenomenon has been the formation of precipitate followed by its slow dissolution concomitant with the formation and growth of crystals. The converse is not observed. This is one empirical demonstration that crystals represent more favorable energy states.

Proteins present special problems for crystallographers

In principle, the crystallization of a protein, nucleic acid, or virus (like that shown in Fig. 6) is little different than the crystallization of conventional small molecules. Crystallization requires the gradual creation of a supersaturated solution of the macromolecule followed by spontaneous formation of crystal growth centers or nuclei. Once growth has commenced, emphasis shifts to maintenance of virtually invariant conditions so as to sustain continued, ordered addition of single molecules, or perhaps ordered aggregates, to surfaces of the developing crystal.

The perplexing difficulties that arise in the crystallization of macromolecules in comparison with conventional small molecules stem from the greater complexity, lability and dynamic properties of proteins and nucleic acids. The description offered above of labile and metastable regions of supersaturation are still applicable to macromolecules, but it must now be borne in mind that as conditions are adjusted to transport the solution away from equilibrium by alteration of its physical and chemical properties, the very nature of the solute molecules is changing as well. As temperature, pH, pressure or solvation are changed, so may be the conformation, charge state or size of the solute macromolecules.

In addition, proteins and nucleic acids are very sensitive to their environment and if exposed to sufficiently severe conditions may denature, degrade or randomize in a manner that ultimately precludes any hope of their forming crystals. They must be constantly maintained in a thoroughly hydrated state at or near physiological pH and temperature. Thus common methods for the crystallization of conventional molecules such as evaporation of solvent, dramatic temperature variation, or addition of strong organic solvents are unsuitable and destructive. They must be supplanted with more gentle and restricted techniques.

Properties of macromolecular crystals

Macromolecular crystals are composed of approximately 50% solvent on average, though this may vary over 25–90% depending on the particular macromolecule [17]. The protein or nucleic acid occupies the remaining volume so that the entire crystal is in many ways an ordered gel with extensive interstitial spaces through which solvent and other small molecules may freely diffuse. This is seen quite dramatically in electron micrographs of small protein crystals such as those in Fig. 7.

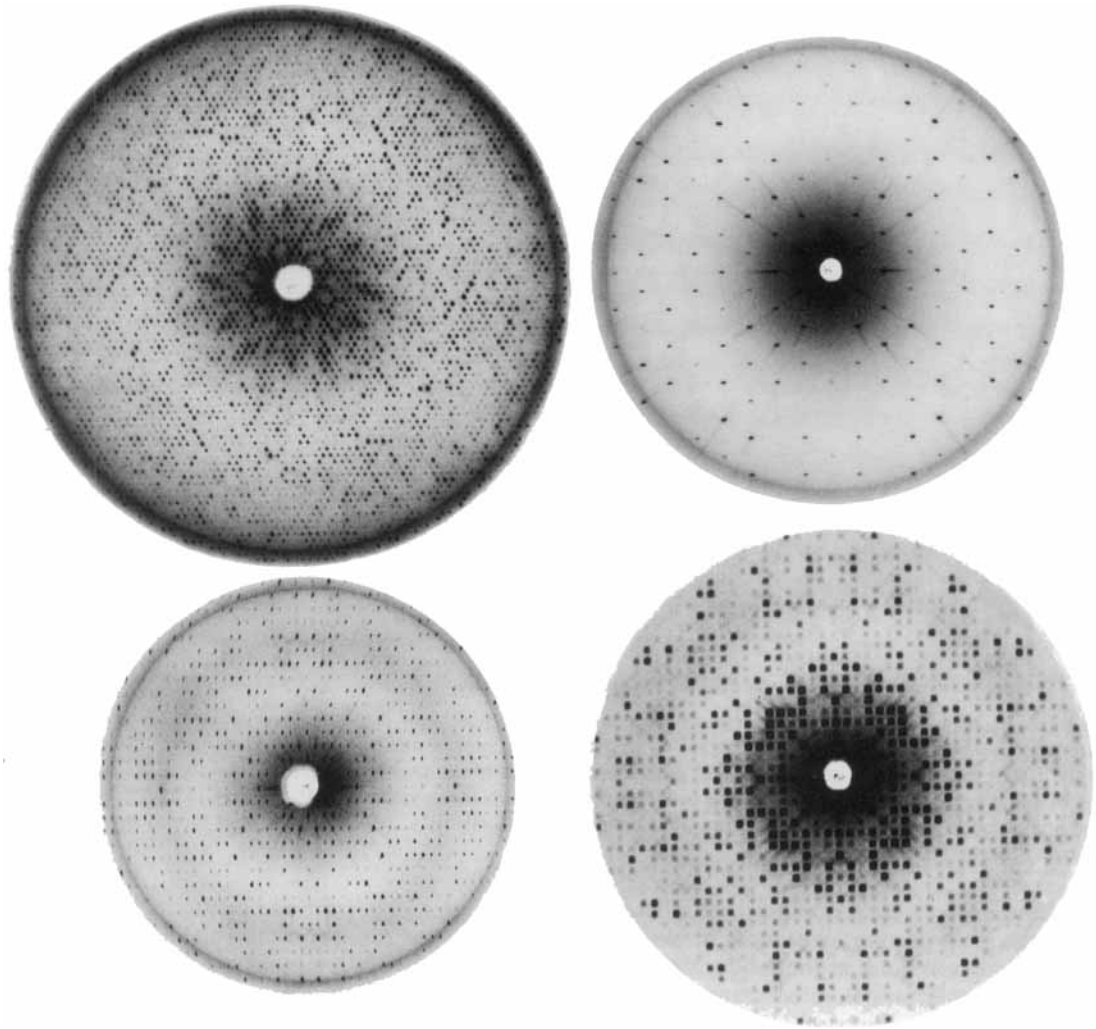


Fig. 8. *Four examples of good X-ray diffraction photographs obtained from different protein crystals.* Upper left, a hexagonal crystal of canavalin showing sixfold symmetry; upper right, a monoclinic crystal of the gene-5 DNA-unwinding protein with mm symmetry; lower left, an orthorhombic crystal of the complex between RNase A and the oligonucleotide $(dA)_4$; lower right, a tetragonal crystal of dogfish lactate dehydrogenase showing its characteristic fourfold symmetry. All of these diffraction patterns extend to a high level of resolution, and all have provided a basis for the structure determination of their constituent macromolecules

In proportion to molecular mass, the number of bonds (salt bridges, hydrogen bonds, hydrophobic interactions) that a conventional molecule forms in a crystal with its neighbors far exceeds the very few exhibited by crystalline macromolecules. Since these contacts provide the lattice interactions that maintain the integrity of the crystal, this largely explains the difference in properties between crystals of salts or small molecules, and macromolecules, as well as why it is so difficult to grow protein and nucleic acid crystals.

Because proteins are sensitive and labile macromolecules that readily loose their native structures, the only conditions that can support crystal growth are those that cause little or no perturbation of the molecular properties. Thus protein crystals, maintained within a narrow range of pH, temperature and ionic strength, must be grown from a solution to which they are tolerant. This is called the mother liquor. Because complete hydration is essential for the maintenance of structure, protein crystals are always, even during data collection, bathed in the mother liquor.

Although morphologically indistinguishable, there are important practical differences between crystals of low-molecular-mass compounds and crystals of proteins and nucleic acids. Crystals of small molecules exhibit firm lattice forces, are highly ordered, generally physically hard and brittle, easy to manipulate, usually can be exposed to air, have strong optical properties, and diffract X-rays intensely. Macromolecular crystals are by comparison usually more limited in size, are very soft and crush easily, disintegrate if allowed to dehydrate, exhibit weak optical properties and diffract X-rays poorly. Macromolecular crystals are temperature-sensitive and undergo extensive damage after prolonged exposure to radiation. In general, many crystals must be analyzed for a structure determination to be successful (for reviews of crystal structure analysis by X-ray diffraction, see [4, 18–21]), although the advent of area detectors and high intensity X-ray sources has greatly lessened this constraint in recent times.

The extent of the diffraction pattern from a crystal is directly correlated with its degree of internal order. The more extensive the pattern, or the higher the resolution to which it extends, the more uniform are the molecules in the crystal and the more precise is their periodic arrangement. The level of detail to which atomic positions can be determined by a crystal structure analysis corresponds closely with the degree of crystalline order. While conventional molecular crystals often dif-

fract almost to their theoretical limit of resolution, protein crystals by comparison are characterized by diffraction patterns of limited extent. Some better examples of diffraction patterns from protein cyrstals are shown in Fig. 8.

The liquid channels and solvent cavities that characterize macromolecular crystals are primarily responsible for the limited resolution of the diffraction patterns. Because of the relatively large spaces between adjacent molecules and the consequent weak lattice forces, every molecule in the crystal may not occupy exactly equivalent orientations and positions in the crystal but they may vary slightly from lattice point to lattice point. Furthermore, because of their structural complexity and their potential for conformational dynamics, protein molecules in a particular crystal may exhibit slight variations in the course of their polypeptide chains or the dispositions of side groups.

Although the presence of extensive solvent regions is a major contributor to the poor quality of protein crystals, it is also responsible for their value to biochemists. Because of the very high solvent content, the individual macromolecules in protein crystals are surrounded by hydration layers that maintain their structure virtually unchanged from that found in bulk solvent. As a consequence, ligand binding, enzymatic and spectroscopic characteristics and other biochemical features are essentially the same as for the native molecule in solution. In addition, the size of the solvent channels is such that conventional chemical compounds, which may be ions, ligands, substrates, coenzymes, inhibitors, drugs or other effector molecules, may be freely diffused into and out of the crystals. Crystalline enzymes, though immobilized, are completely accessible for experimentation through alteration of the surrounding mother liquor. Thus, a protein crystal can serve as a veritable ligand binding laboratory [21].

Crystallization strategy

The strategy employed to bring about crystallization is to guide the system very slowly toward a state of reduced solubility by modifying the properties of the solvent. This is accomplished by increasing the concentration of precipitating agents or by altering some physical property, such as pH. In this way, a limited degree of supersaturation is achieved. Whatever the procedure used, no effort must be spared in refining the parameters of the system, solvent and solute, to encourage and promote specific bonding interactions between molecules and to stabilize them once they have formed. This latter aspect of the problem generally depends on the chemical and physical properties of the particular protein or nucleic acid being crystallized.

In very concentrated solutions the macromolecules may aggregate as an amorphous precipitate. This result is to be avoided if possible and is indicative that supersaturation has proceeded too extensively or too swiftly. One must endeavor to approach very slowly the point of inadequate solvation and thereby allow the macromolecules sufficient opportunity to order themselves in a crystalline lattice.

The classical procedure for inducing proteins to separate from solution and produce a solid phase is to gradually increase the level of saturation of a salt. Traditionally the salt has been ammonium sulfate, but others are also in common use. Most frequently the protein separates as a precipitate, but with appropriate care, manipulation of salt concentration can be used to grow protein crystals. At the present time, in fact, this approach has probably yielded more varieties of protein crystals than any other [22].

For a specific protein, the precipitation points or solubility minima are usually critically dependent on the pH, temperature, the chemical composition of the precipitant, and the properties of both the protein and the solvent (for more extensive discussions, see [23–27]). At very low ionic strength a phenomenon known as 'salting-in' occurs in which the solubility of the protein increases as the ionic strength increases from zero (see the solubility curve in Fig. 4 for example). The physical effect that diminishes solubility at very low ionic strength is the removal of ions essential for satisfying the electrostatic requirements of the protein molecules. As these ions are removed, and in this region of low ionic strength cations are most important [23, 24], the protein molecules seek to balance their electrostatic requirements through interactions among themselves. Thus they tend to aggregate and separate from solution. Alternatively, one may say that the chemical activity of the protein is reduced at very low ionic strength.

The salting-in effect, when applied in the direction of reduced ionic strength, can itself be used as a crystallization tool. In practice, one extensively dialyzes a protein that is soluble at moderate ionic strength against distilled water. Many proteins such as catalase, concanavalin B, and a host of immunoglobulins and seed proteins have been crystallized by this means [3, 4].

As ionic strength is increased the solution again reaches a point where the solute molecules begin to separate from solvent and preferentially form self interactions that result in crystals or precipitate. The explanation for this 'salting out' phenomenon is that the salt ions and macromolecules compete for the attention of the solvent molecules, that is, water. Both the salt ions and the protein molecules require hydration layers to maintain their solubility. When competition between ions and proteins becomes sufficiently intense, the protein molecules begin to self associate in order to satisfy, by intermolecular interactions, their electrostatic requirements. Thus dehydration, or the elimination and perturbation of solvent layers around protein molecules, induces insolubility.

Just as proteins may be driven from solution at constant pH and temperature by the addition or removal of salt [26], they can similarly be crystallized or precipitated at constant ionic strength by changes in pH or temperature. This is because the electrostatic character of the macromolecule, its surface features, or its conformation may change as a function of pH, temperature and other variables as well [23]. By virtue of its ability to inhabit a range of states, proteins may exhibit a number of different solubility minima as a function of the variables, and each of these minima may afford the opportunity for crystal formation. Thus, we may distinguish the separation of protein from solution according to methods based on variation of precipitant concentration at constant pH and temperature from those based on alteration of pH, temperature or some other variable at constant precipitant concentration. The principles described here for salting-out with a true salt are not appreciably different if precipitating agents such as poly(ethylene glycol) are used instead. In practice, proteins may equally well be crystallized from solution by increasing the poly(ethylene glycol) concentration at constant pH and temperature, or at constant poly(ethylene glycol) concentration by variation of pH or temperature [5, 28].

The most common approach to crystallizing macromolecules, be they proteins or nucleic acids, is to alter gradually the characteristics of a highly concentrated protein solution to achieve a condition of limited supersaturation. As discussed above, this may be achieved by modifying some physical prop-

erty such as pH or temperature, or through equilibration with precipitating agents. The precipitating agent may be a salt such as ammonium sulfate, an organic solvent such as ethanol or methylpentanediol, or a highly soluble synthetic polymer such as poly(ethylene glycol). The three types of precipitants act by slightly different mechanisms, though all share some common properties.

In highly concentrated salt solutions competition for water exists between the salt ions and the polyionic protein molecules. The degree of competition will depend on the surface charge distribution of the protein as well. This is a function primarily of pH. Because protein molecules must bind water to remain solvated, when deprived of sufficient water by ionic competition, they are compelled to associate with other protein molecules. Aggregates may be random in nature and lead to linear and branched oligomers, and eventually to precipitate. When the process proceeds in an orderly fashion and specific chemical interactions are used in a repetitive and periodic manner to give three-dimensional aggregates, then the nuclei of crystals will form and grow.

The removal of available solvent by addition of precipitant is in principle no different than the crystallization of sea salt from tidal pools as the heat of the sun slowly drives the evaporation of water. It is a form of dehydration but without physical removal of water.

A similar effect may be achieved as well by the slow addition to the mother liquor of certain organic solvents such as ethanol or methylpentanediol. The only essential requirement for the precipitant is that at the specific temperature and pH of the experiment, the additive does not adversely effect the structure and integrity of the protein. This is often a very stringent requirement and deserves more than a little consideration. The organic solvent competes to some extent like salt for water molecules, but it also reduces the dielectric screening capacity of the intervening solvent. Reduction of the bulk dielectric increases the effective strength of the electrostatic forces that allow one protein molecule to be attracted to another.

Polymers such as poly(ethylene glycol) also serve to dehydrate proteins in solution as do salts, and they alter somewhat the dielectric properties in a manner similar to organic solvents. They produce, however, an additional important effect. Poly(ethylene glycol) perturbs the natural structure of the solvent and creates a more complex network having both water and itself as structural elements. A consequence of this restructuring of solvent is that macromolecules, particularly proteins, tend to be excluded and phase separation is promoted [29, 30].

Crystallization of macromolecules may also be accomplished by increasing the concentration of a precipitating agent to a point just below supersaturation and then adjusting the pH or temperature to reduce the solubility of the protein. Modification of pH can be accomplished very well with the vapor diffusion technique, which is described below, when volatile acids and bases such as acetic acid and ammonium hydroxide are used. This process is analogous to saturating boiling water with sugar and then cooling it to produce rock candy.

Creating the supersaturated state

Crystallization of a novel protein using any of the precipitation methods is unpredictable as a rule. Every macromolecule is unique in its physical and chemical properties because every amino acid or nucleotide sequence produces a unique three-dimensional structure having distinctive surface characteristics. Thus, lessons learned by investigation of one protein are only marginally applicable to others. This is compounded by the behavior of macromolecules which is complex owing to the variety of molecular masses and shapes, aggregate states, and polyvalent surface features that change with pH and temperature, and to their dynamic properties [7].

Because of the intricacy of the interactions between solute and solvent, and the shifting character of the protein, the methods of crystallization must usually be applied over a broad set of conditions with the objective of discovering the particular minimum (or minima) that yield crystals. In practice, one determines the precipitation points of the protein at sequential pH values with a given precipitant, repeats the procedure at different temperatures, and then examines the effects of different precipitating agents.

There are a number of devices, procedures and methods for bringing about the supersaturation of a protein solution, generally by the slow increase in concentration of some precipitant such as salt or poly(ethylene glycol). Many of these same approaches can be used as well for salting-in, modification of pH and the introduction of ligands that might alter protein solubility. These techniques have been reviewed elsewhere [3–6, 20, 31] and will not be dealt with exhaustively here. Only three of these, microdialysis, free interface diffusion, and vapor equilibration, will be described as examples of the best methods in current use. A drawing summarizing these techniques is seen in Fig. 9.

Dialysis is familiar to nearly all biochemists as a means of changing some properties of a protein-containing solution. The macromolecule solution is maintained inside a membrane casing or container having a semi-permeable membrane partition. The membrane allows, through its pores, the passage of small molecules and ions, but the pore size excludes passage of the much larger protein molecules. The vessel or dialysis tube containing the protein is submerged in a larger volume of liquid having the desired solution properties of pH, ionic strength, ligands, etc. With successive changes of the exterior solution and concomitant equilibration of small molecules and ions across the semipermeable membrane, the protein solution gradually acquires the desired properties of the exterior fluid.

Exactly this same procedure, in some manifestation or other, can and has been used to crystallize a number of proteins on a bulk scale [32, 33]. It is generally applicable on a large scale, however, only when substantial amounts of the protein are available. It has the advantage that by liquid-liquid diffusion through a semi-permeable membrane, a protein solution can be exposed to a continuum of potential crystal-producing conditions without actually altering directly the mother liquor. Diffusion through the membrane is slow and controlled. Because the rate of change of substituents in the mother liquor is proportional to the gradient of concentrations across the membrane, the nearer the system approaches equilibrium, the more slowly it changes.

This method has been adapted to much smaller amounts of protein by crystallographers who now use almost exclusively microtechniques involving no more than 5–50 µl protein solution in each trial. First described by Zeppenzauer and Zeppenzauer [34, 35] and subsequently modified and refined by numerous others, the method confines a protein solution to the interior of a glass capillary, or the microcavity of a small plexiglass button. The cavity of the button or the ends of the microcapillary tube are then closed off by a semi-permeable dialysis membrane. The whole arrangement,

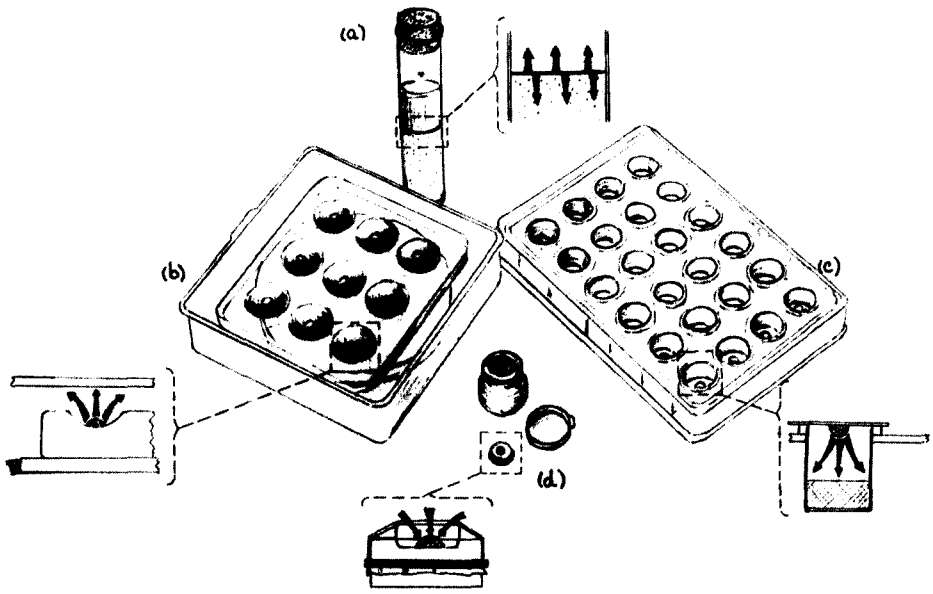


Fig. 9. *A drawing of an array of the most common microtechniques currently in use for the crystallization of macromolecules.* (a) The free interface diffusion technique; (b and c) two useful vapor diffusion methods using sitting drops on glass depression plates and hanging drops in tissue-culture plates; (d) a liquid dialysis button and a small vial which serves as the exterior liquid reservoir. All can be used with a variety of conditions and precipitating agents and each allows gradual equilibration of the protein and precipitating solutions to attain supersaturation

charged with protein solution, is then submerged in a much larger volume of an exterior liquid and the whole system kept within a closed vessel such as a test tube or vial.

If the exterior solution is at an ionic strength or pH that causes the mother liquor to become supersaturated, crystals may grow. If not, the exterior solution may be exchanged for another and the experiment continued.

The dialysis buttons, seen in Fig. 10, are particularly ingenious. Not only are they compact and easy to examine, but they have a shallow groove about their waist. After a section of wet dialysis membrane is placed over the mother-liquor-filled cavity, it can be held firmly and precisely in place by simply slipping a common rubber O ring over the top of the button and seating it in the groove.

These buttons, available from Cambridge Repetition Parts (Cambridge, UK), are now in wide use, and have proven themselves quite successful. Their cavities range in size from 5 μl to 50 μl and they can be reused many times.

A modification of the liquid-liquid diffusion method is the free interface diffusion technique [36, 37]. Here, the membrane is dispensed with completely and the mother liquor is simply layered upon a second precipitating solution in a glass tube or capillary. In some applications, the bottom solution is first frozen before the second is layered to ensure a sharp demarcation between the two.

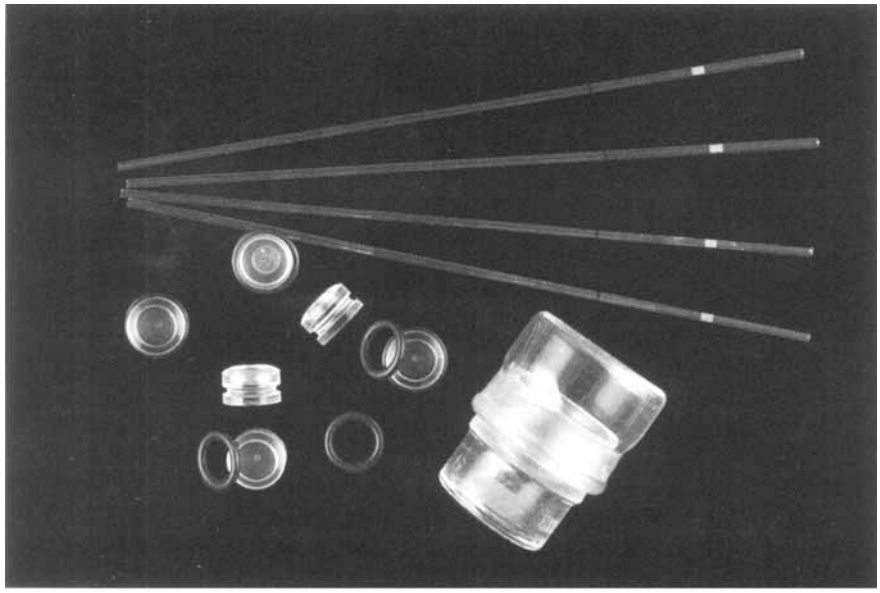
In the free interface diffusion method, direct diffusive and convective mixing at the interface generates concentration gradients that produce regions of local supersaturation. These can, in turn, yield nuclei that may grow to a size useful for diffraction analysis. Modifications of this technique are currently being planned for experiments in zero gravity aboard the space shuttle. In zero gravity, where only diffusive interchange occurs and where stable concentration gradients of precipitant and protein can be established and maintained, the method may prove to be even more successful than on earth.

Currently, the most widely used method for bringing about supersaturation in microdrops of protein mother liquid is vapor diffusion [3, 4, 31, 38]. This approach also exhibits a diversity and may be divided into those procedures that use a 'sitting drop' and those employing a 'hanging drop.' In any form, the method relies on the transport of either water or some volatile agent between a microdrop of mother liquor, generally 5 – 25 μl volume, and a much larger reservoir solution of 0.75 – 25 ml volume. Through the vapor phase, the droplet and reservoir come to equilibrium, and because the reservoir is of such larger volume, the final equilibration conditions are essentially those of the initial reservoir state. A variety of devices currently in use for protein crystal growth by vapor diffusion are shown in Fig. 11.

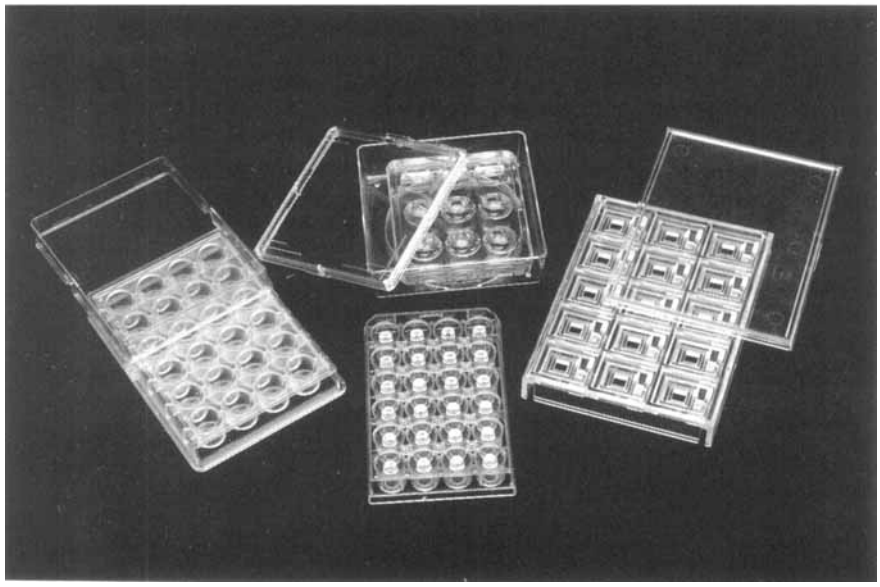
Through the vapor phase, then, water is removed slowly from the droplet of mother liquor, its pH may be changed, or volatile solvents such as ethanol may be gradually introduced. As with the liquid-liquid dialysis and diffusion methods, the procedure may be carried out at a number of different temperatures to gain advantage of that parameter as well.

According to a popular procedure, droplets of 10 – 20 μl are placed in the nine wells of depression spot plates (Corning Glass no. 7220). The samples are then sealed in transparent containers, such as Pyrex dishes or plastic boxes, which hold, in addition, reservoirs of 20 – 50 ml of the precipitating solution. The plates bearing the protein or nucleic acid samples are held off the bottom of the reservoir by the inverted half of a disposable Petri dish. Through the vapor phase, the concentration of salt or organic solvent in the reservoir equilibrates with that in the sample. In the case of salt precipitation, the droplet of mother liquor must initially contain a level of precipitant lower than the reservoir, and equilibration proceeds by distillation of water out of the droplet and into the reservoir. This holds true for nonvolatile organic solvents, such as methylpentanediol and for poly(ethylene glycol) as well. In the case of volatile precipitants, none need be added initially to the microdroplet, as distillation and equilibration proceed in the opposite direction.

This method has the advantage that it requires only small amounts of material and is ideal for screening a large number



10



11

Fig. 10. *Plexiglass dialysis buttons with the O rings than maintain the dialysis membrane in place, and a small weighing bottle that serves to hold the exterior reservoir and the button.* Some microcapillaries are seen at top in which free interface diffusion can be carried out on a microscale
Fig. 11. *The four varieties of crystallization apparatus or plates now in common use.* Left the Linbro tissue-culture plate appropriated for use in hanging-drop experiments; top, the traditional sitting-drop apparatus consisting of a Corning glass depression plate in a plastic box; bottom, the Cryschem multiwell vapor diffusion plate; right, the plate from FLO Labs for both sitting and sandwich drops. The Linbro and FLO Lab plates are used in conjunction with glass cover slips while the Cryschem plate is covered with transparent plastic tape

of conditions. The major disadvantage is that all samples in a single box must be equilibrated against the same reservoir solution. It does, however, permit some flexibility in varying conditions once the samples have been dispensed, by modification of the concentration or pH of precipitants in the reservoir. When clear plastic boxes are used, large numbers of samples can be quickly inspected for crystals under a dissecting microscope and conveniently stored.

The disadvantage of identical reservoir conditions for all samples in a single box has been overcome to a great extent with the introduction of two plastic plates specifically designed for protein crystallization. One of these, sponsored by the American Crystallographic Association and manufactured by FLO Labs, Inc., is a plastic plate having accommodation for 15 protein samples. Each chamber has a separate reservoir compartment and the mother liquor microdroplet may be either suspended from the underside of a glass cover slip as in the 'hanging drop' method or sandwiched between two glass cover slips. Sealing of the chambers from air requires silicone grease or oil between cover slips and the plastic rims of the chambers. With these plates, the optical properties are very good but equilibration tends to be slow.

A second crystallization plate [39], seen in Fig. 11, is produced by Cryschem Inc. (Riverside, CA). With these plates, the drop is sitting on the top of a clear support post that protrudes upward from a circular reservoir cavity containing the precipitating solution. The chambers can be rapidly and conveniently sealed from air by clear plastic tape pressed onto the upper surface of the plate after the reservoirs have been filled and the drops of mother liquor dispensed. Equilibration,

Table 1. *Precipitants used in macromolecular crystallization*

Salts	Organic solvents	Polymers
1. Ammonium or sodium sulfate	1. Ethanol	1. Poly(ethylene glycol) 1000, 3350, 6000, 20000
2. Lithium sulfate	2. Isopropanol	2. Jeffamine T
3. Lithium chloride	3. 1,3-Propanediol	3. Polyamine
4. Sodium or ammonium citrate	4. 2-Methyl-2,4-pentanediol	
5. Sodium or potassium phosphate	5. Dioxane	
6. Sodium or potassium or ammonium chloride	6. Acetone	
7. Sodium or ammonium acetate	7. Butanol	
8. Magnesium or calcium sulfate	8. Acetonitrile	
9. Cetyltrimethyl ammonium salts	9. Dimethyl sulfoxide	
10. Calcium chloride	10. 2,5-Hexanediol	
11. Ammonium nitrate	11. Methanol	
12. Sodium formate	12. 1,3-Butyrolactone	
	13. Poly(ethylene glycol) 400	

as with the other plate, is through the vapor phase. While the optical properties are somewhat less favorable with the Cryschem devices, they are inexpensive, convenient, compact and can be rapidly utilized in vast screens of crystallization conditions.

The 'hanging drop' procedure also uses vapor phase equilibration but, with this approach, a microdroplet of mother liquor (as small as 5 µl) is suspended from the underside of a microscope cover slip, which is then placed over a small well containing 1 ml of the precipitating solution. The wells are most conveniently supplied by disposable plastic tissue culture plates (Linbro model FB – 16 – 24-TC) that have 24 wells with rims that permit sealing by application of silicone vacuum grease or oil around the circumference. These plates provide the further advantages that they can be swiftly and easily examined under a dissecting microscope and they allow compact storage. The hanging drop technique can be used both for the optimization of conditions and for the growth of large single crystals.

While the principle of equilibration with both the 'sitting drop' and the 'hanging drop' are essentially the same, they frequently do not give the same results even though the reservoir solutions and protein solutions are identical. Presumably because of the differences in the apparatus used to achieve equilibration, the path to equilibrium is different even though the end point may be the same. In some cases there are striking differences in the degree of reproducibility, final crystal size, morphology, required time, or degree of order. These observations illustrate the important point that the pathway leading to supersaturation, the kinetics of the process, may be as important as the final point achieved.

As noted earlier, one of the most powerful techniques for producing a supersaturated protein solution is adjustment of the pH to values where the protein is substantially less soluble. This may be done in the presence of a variety of precipitants so that a spectrum of possibilities can be created whereby crystals might form. The gradual alteration of pH is particularly useful because it may be accomplished by a variety of gentle approaches that do not otherwise perturb the system or introduce unwanted effects.

Although microdialysis is probably equally suitable, more success has been achieved with the vapor diffusion method using 'sitting' microdroplets on spot plates or in one of the plastic plates available for protein crystallization. The ambient salt, effector, or buffer conditions are established prior to dispensing the microdroplets in the depressions on the plate. The pH is then slowly raised or lowered by adding a small amount of volatile acid or base to the reservoir. Diffusion of the acid or base then occurs from reservoir to sample, just as for a volatile precipitant.

If the pH is to be raised, for example, a small drop of concentrated ammonium hydroxide can be added to the reservoir; a drop of acetic acid may be used to lower it. The pH can also be gradually lowered over a period of days by simply placing a tiny chip of solid CO_2 in the reservoir. The liberated CO_2 diffuses and dissolves in the mother liquor to form weak carbonic acid.

When a specific pH end point is required, the mother liquor may be buffered with suitable compounds at that point and then moved significantly away by addition of acid or base. The microdroplets of mother liquor may then be returned to the buffer point by addition of an appropriate volatile acid or base to the reservoir.

As with pH, proteins may vary in solubility as a function of temperature, and some are quite sensitive. One can take advantage of this property with both bulk and microtechniques [40 – 42]. Many of the earliest examples of protein crystallization were based on the formation of concentrated solutions at elevated temperatures followed by slow cooling. Osborne in 1892 [43] reported the crystallization of numerous plant seed globulins by cooling relatively crude extracts from 60 °C to room temperature in the presence of varying concentrations of sodium chloride. These same procedures were followed by Bailey in 1942 [32, 33] and Vickery et al. in 1941 [44] to crystallize other proteins. More recent examples are those of glucagon [45], which is crystallized by dissolving the protein at 60 °C in appropriate buffers and cooling slowly to room temperature, insulin [46] and deoxyribonuclease [47].

If temperature change is an important consideration or the primary means for inducing crystal formation, its rate may be manipulated to some extent by enclosing the sample at elevated temperature in a Dewar flask or insulated container and then placing the container at the desired final temperature. The use of thermal insulation in this regard has been reported for insulin and has been used as well for the crystallization of numerous conventional small molecules of biological interest.

Most protein and nucleic acids are conformationally flexible or exist in several conformational equilibrium states. In addition, they may assume a substantially different conformation when they have bound coenzyme, substrate, or other ligand. Frequently a protein with bound effector may exhibit appreciably different solubility properties than the native protein. In addition, if many conformational states are available, the presence of effector may be used to select for only one of

these, thereby engendering a degree of conformity of structure and system microhomogeneity that would otherwise be absent.

The effect of ligands can be employed to induce supersaturation and crystallization in those cases where its binding to the protein produces solubility differences under a given set of ambient conditions. The effector may be slowly and gently combined with the protein, for example by dialysis, so that the resulting complex is at a supersaturating level.

The addition of ligands, substrates, and other small molecules has seen widespread use in protein crystallography, since it provides useful alternatives if the apoenzyme itself cannot be crystallized.

Precipitating agents

Protein precipitants fall into four broad categories: (a) salts, (b) organic solvents, (c) long-chain polymers and (d) low-molecular-mass polymers and non-volatile organic compounds. The first two classes are typified by ammonium sulfate and ethyl alcohol respectively, and higher polymers such as poly(ethylene glycol) 4000 are characteristic of the third. In the fourth category we might place compounds such as methylpentanediol and low-molecular-mass poly(ethylene glycol). Common members of the four groups are presented in Table 1.

As already described, salts exert their effect by dehydrating proteins through competition for water molecules. Their ability to do this is proportional to the square of the valences of the ionic species composing the salt [23, 26]. Thus, multivalent ions, particularly anions are the most efficient precipitants. Sulfates, phosphates and citrates have traditionally been employed with success.

One might think there would be little variation between different salts so long as their ionic valences were the same, or that there would be little variation with two different sulfates such as Li_2SO_4 and $(NH_4)_2SO_4$. This, however, is often not the case. In addition to salting out, which is a general dehydration effect or lowering of the chemical activity of water, there are also specific protein–ion interactions that may have other consequences. This is particularly true because of the unique polyvalent character of individual proteins, their structural complexity, and the intimate dependence of their physical properties on environmental conditions and interacting molecules. It is never sufficient, therefore, when attempting to crystallize a protein to examine only one or two salts and ignore a broader range. Changes in salt can sometimes produce crystals of varied quality, morphology, and in some cases diffraction properties.

It is usually not possible to predict the degree of saturation or molarity of a salt required for the crystallization of a particular protein without some prior knowledge of its behavior. In general, however, it is a concentration just a small percentage less than that which yields an amorphous precipitate, and this can be determined for a macromolecule under a given set of conditions using only minute amounts of material.

To determine approximately the precipitation point with a particular agent, a 10-µl droplet of a 5–15 mg/ml protein solution can be placed in a well of a depression slide and observed under a low-power light microscope as increasing amounts of saturated salt solution or organic solvent (in 1-µl or 2-µl increments) are added. If the well is sealed between additions with a cover slip, the increases can be made over a period of many hours. Indeed, the droplet should be allowed to equilibrate for 10–30 min after each addition, and longer in the neighborhood of the precipitation point. With larger amounts of material the sample may be dialyzed in standard 6.35-mm (0.25 in) celluloid tubing or in a microdialysis button against a salt solution that is incremented over a period of time until precipitation occurs.

In general, the most common organic solvents utilized have been ethanol, acetone, butanols and a few other common laboratory reagents [48, 49]. It might be noted here that organic solvents have been of more general use for the crystallization of nucleic acids, particularly tRNA and the duplex oligonucleotides. There they have been the primary means for crystal growth. This in part stems from the greater tolerance of polynucleotides to organic solvents and their polyanionic surfaces which appear to be even more sensitive to dielectric effects than are proteins.

The only general rules are that organic solvents should be used at a low temperature, at or below 0 °C, and they should be added very slowly and with good mixing. Since most are volatile, vapor diffusion techniques are equally suitable for both bulk or micro amounts. Ionic strength should, in general, be maintained as low as possible and whatever means are available should be taken to protect against denaturation.

Poly(ethylene glycol) is a polymer produced in various lengths, containing from several to many hundred monomers. It exhibits as its most conspicuous feature a regular alteration of ether oxygens and terminal glycols. In addition to its volume exclusion property, it shares some characteristics with salts that compete for water and produce dehydration, and with organic solvents which reduce the dielectric properties of the medium.

Aside from its general applicability and utility in obtaining crystals for diffraction analysis [28], poly(ethylene glycol) also has the advantage that it is most effective at minimal ionic strength and provides a low-electron-density medium. The first feature is important because it provides for higher ligand binding affinities than does a high-ionic-strength medium such as concentrated salt. As a consequence there is greater ease in obtaining isomorphous heavy-atom derivatives and in forming protein–ligand complexes for study by difference Fourier techniques. The second characteristic, a low-electron-density medium, implies a generally lower background or noise level for protein structures derived by X-ray diffraction and presumably, therefore, a more ready interpretation.

A number of protein structures have now been solved using crystals grown from poly(ethylene glycol). These confirm that the protein molecules are in as native a condition in this medium as in those traditionally used. This is perhaps even more so, since the larger-molecular-mass poly(ethylene glycols) probably do not even enter the crystals and therefore do not directly contact the interior molecules. In addition, it appears that crystals of many proteins when grown from poly(ethylene glycol) are essentially isomorphous with, and exhibit the same unit cell symmetry and dimensions as, those grown by other means.

Poly(ethylene glycol) is produced in a variety of polymer ranges. The low-molecular-mass species are oily liquids while those of M_r above 1000, at room temperature, exist as either waxy solids or powders. The size specified by the manufacturer is the mean M_r of the polymeric molecules, and the distribution about that mean may vary appreciably. Poly(ethylene glycol) in its commercial form does contain contaminants; this is particularly true of the high-molecular-mass forms such as those of M_r 15000 or 20000. These may be removed by simple purification procedures [50] or, in the case of the 20000-

M_r form by dialysis in low-pass dialysis or collodian tubes. There have been reports that repurified poly(ethylene glycol) has proved more effective [51]. Certainly the contaminants could be disadvantageous for some proteins.

All of the poly(ethylene glycol) sizes from M_r 400 to 20000 have provided protein crystals, but the most useful are those in the range 2000 – 6000. Occasionally, however, a protein can not be easily crystallized using this range but yields in the presence of polymer with M_r 400 or 20000. The sizes are generally not completely interchangeable for a given protein even within the mid-range, some producing the best-formed and largest crystals only at, say, M_r 4000 and less perfect examples at other M_r. This is a parameter which is best optimized by empirical means along with concentration and temperature.

A distinct advantage of poly(ethylene glycol) over other agents is that most proteins (but not all) crystallize within a fairly narrow range of poly(ethylene glycol) concentration, this being about 4 – 18%. In addition, the exact poly(ethylene glycol) concentration at which crystals form is rather insensitive and if one is within 2 – 3% of the optimal value some success will be achieved. With most crystallizations from high-ionic-strength solutions or from organic solvents, one must be within 1 – 2% of an optimum lying anywhere between 10 – 85% saturation. The advantage of poly(ethylene glycol) is that, when conducting a series of initial trials to determine what conditions will give crystals, one can use a fairly coarse selection of concentrations and over a rather narrow total range. This means fewer trials with a corresponding reduction in the amount of protein expended. Thus, it is well suited for particularly precious proteins of very limited availability.

The time required for crystal growth with poly(ethylene glycol) as the precipitant is also generally shorter than with ammonium sulfate or methylpentanediol but occasionally longer than required by volatile organic solvents such as ethanol. Although equilibration times will depend on the differential between starting and target concentrations, if this is no more than 3 – 4%, then crystallization may occur within a few hours or a few days. It seldom requires more than two weeks. Thus evaluation of results can be made without undue demands on patience. It should be noted, however, that protein-poly(ethylene glycol) solutions are excellent media for microbes, particularly molds, and if crystallization is being attempted at room temperature or over extended periods of time, then some retardant such as azide (commonly 0.1%) must be included in the protein solutions.

Since poly(ethylene glycol) solutions are not volatile, this precipitant must be used like salt and equilibrated with the protein by dialysis, slow mixing, or vapor equilibration. This latter approach, utilizing either 15-µl hanging drops over 0.5-ml reservoirs or multi-depression glass plates in sealed chambers, has proved the most popular. When the reservoir concentration is in the range 5 – 12%, the protein solution to be equilibrated should be at an initial concentration of about half of that, which is conveniently obtained by adding an equal volume of the reservoir to that of the protein solution. When the target poly(ethylene glycol) concentration is much higher than 12%, it is advisable to start the protein equilibrating at no more than 4 – 5% below the final value. This reduces unnecessary time lags during which the protein might denature.

Crystallization of proteins with poly(ethylene glycol) has proved most successful when the ionic strength is low and difficult when high. Good buffer conditions in the neutral range are, for example, 10 – 20 mM Tris or cacodylate buffer.

Table 2. *Factors that do or could affect protein crystal growth*

1. pH and buffer
2. Ionic strength
3. Temperature and temperature fluctuations
4. Concentration and nature of precipitant
5. Concentration of macromolecule
6. Purity of macromolecules (see Table 3 regarding Microheterogeneity)
7. Additives, effectors and ligands
8. Organism source of macromolecule
9. Substrates, coenzymes, inhibitors
10. Reducing or oxidizing environment
11. Metal and other specific ions
12. Rate of equilibration and rate of growth
13. Surfactants or detergents
14. Gravity, convection and sedimentation
15. Vibrations and sound
16. Volume of crystallization sample
17. Presence of amorphous or particulate material
18. Surfaces of crystallization vessels
19. Proteolysis
20. Contamination by microbes
21. Pressure
22. Electric and magnetic fields
23. Handling by investigator and cleanliness
24. Viscosity of mother liquor
25. Heterogeneous or expitaxial nucleating agents

If crystallization proceeds too rapidly, addition of some neutral salt may be used to slow growth and better effect crystal form. Poly(ethylene glycol) is useful over the entire pH range and over a broad temperature range and shows no anomalous effects in response to either.

Factors influencing protein crystal growth

Table 2 lists physical, chemical and biological variables that may influence to a greater or lesser extent the crystallization of proteins. The difficulty in properly arriving at a just assignment of importance for each factor is substantial for several reasons. Every protein is different in its properties and, surprisingly perhaps, this applies even to proteins that differ by no more than one or just a few amino acids. There are even cases where the identical protein prepared by different procedures or at different times may show significant variations. In addition, each factor may differ considerably in importance for individual proteins. α-Amylase and catalase, for example, are clearly sensitive to temperature change, while ovalbumin and ferritin show little, if any, variation in crystallization properties as a function of that variable.

Because each protein is unique, there are few means available to predict in advance the specific values of a variable, or sets of conditions that might be most profitably explored. Finally, the various parameters under one's control are not independent of one another and their interrelations may be complex and difficult to discern. It is, therefore, not easy to elaborate rational guidelines relating to physical factors or ingredients in the mother liquor that can increase the probability of success in crystallizing a particular protein. The specific components and conditions must be carefully deduced and refined for each individual.

As already noted, temperature may be of great importance or it may have little bearing at all. In general, it is wise to duplicate all crystallization trials and conduct parallel investigations at 4°C and at 25°C. Even if no crystals are observed

at either temperature, differences in the solubility behavior of the protein with different precipitants and with various effector molecules may give some indication as to whether temperature is likely to play an important role. If crystals are observed to grow at one temperature and not, under otherwise identical conditions, at the other, then further refinement of this variable is necessary. This is accomplished by conducting the trials under the previously successful conditions over a range of temperatures centered on the one that initially yielded crystals.

The only rules with regard to temperature seem to be that proteins in a high salt solution are usually more soluble at cold than warmer temperatures. Proteins, however, generally precipitate or crystallize from a lower concentration of poly *(ethylene glycol)*, methylpentanediol or organic solvent at cold than at warmer temperature. One must remember, however, that diffusion rates are less and equilibration occurs more slowly at cold than higher temperature, so that the times required for precipitation or crystal formation may be longer at colder temperatures.

After precipitant concentration, the next most important variable in protein crystal growth appears to be pH. This follows since the charge character of a protein and all of its attendant physical and chemical consequences are intimately dependent on the ionization state of the amino acids or chemical groups that comprise the macromolecule. Not only does the net charge on the protein change with pH, but the distribution of those charges, the dipole moment of the protein, its conformation, and in many cases its aggregation state. Thus, an investigation of the behavior of a specific protein as a function of pH is perhaps the single most essential analysis that should be carried out in attempting to crystallize the macromolecule.

As with temperature, the procedure is to first conduct multiple crystallization trials at coarse intervals over a broad pH range and then repeat the trials over a finer matrix of values in the neighborhoods of those that initially showed promise. The only limitations on the breadth of the initial range screened are the points at which the protein begins to show indications of denaturation. In refining the pH for optimal growth, it should be recalled that the difference between amorphous precipitate, microcrystals, and large single crystals may be only a ΔpH of less than 0.5 [34, 35].

In addition to adjusting pH for the optimization of crystal size, it is sometimes also useful to explore variation of pH as a means of altering the habit or morphology of a crystalline protein. This is occasionally necessary if the initial crystal form is not amenable to analysis because it grows as fine needles or flat, thin plates or demonstrates some other unfavorable tendency such as striation or twinning.

There have been virtually no systematic studies of such factors as pressure, sound, vibrations, electrical and magnetic fields, or viscosity on the rate of growth or final quality of protein crystals. Similarly, studies are only now being undertaken to evaluate the effects of gravity, convection and fluid flow on protein crystal growth, final size, and perfection [52–54]. Thus it is not possible at this time to evaluate their influence definitively.

Some useful considerations

The earliest investigators of protein crystals noted that the concentration of protein in the mother liquor should be as high as possible, 10–100 mg/ml. This is particularly true if one is attempting to grow crystals of a protein for the first time. The probability of obtaining crystals is certainly enhanced by increasing the concentration of protein. Concentration alone is sometimes sufficient to drive the system into a state of supersaturation and into the labile region where stable nuclei can form. This may not, however, be the best approach in growing large, perfect crystals once optimal conditions for all other parameters have been established.

Once conditions for nucleation and growth have been identified and the investigation of variables more or less complete, the concentration of the protein should be gradually reduced in increments to moderate the growth of the crystals. As a general rule, the largest and most perfect crystals result when the rate of accretion of molecules is slow and orderly. Reduction of concentration is an effective means for controlling this.

The time required for the appearance and growth of protein crystals is quite variable and may range from a few hours in the best of cases to several months in others. Because no truly systematic investigations have been carried out, how rapidly crystals grow once visible nuclei have formed remains in question. The rate of growth may not be reflected at all in the total amount of time required to obtain crystals adequate for analysis. This includes the time required for solvent equilibration to be achieved, for crystal nuclei to form, and for full growth to occur.

When one is screening variables to establish optimal parameters, then the practical objective is to promote crystallization at the greatest possible speed to expedite determination of most probable conditions. When optimizing and refining crystallization parameters, time itself becomes an important parameter and long periods of slow growth are generally desirable.

One caution is in order. If it is observed that a long period elapses without the formation of crystals and then, well beyond the time required for solvent equilibration to have occurred, crystals begin to appear, then some possible causes should be explored. One likelihood is that the protein has,over the long time period, undergone some physical or chemical change. It may have undergone limited proteolysis, lost a coenzyme or metal ion, or undergone a slow conformational change. By forcing this same event to occur before the crystallization trials are carried out the time required for growth may be substantially reduced. Another possibility is that the apparatus in which the crystallization experiments were carried out was leaking and that very slow evaporation occurred. Thus the final concentration of precipitant may have been appreciably higher than believed. A final possibility is change in the ambient temperature. This is particularly likely when crystallization is being carried out at room temperature and heating or air conditioning systems are switched on and off as the seasons change.

The most intriguing questions with regard to optimizing crystallization conditions concern what additional components or compounds should comprise the mother liquor in addition to solvent, protein and precipitating agent. The most probable effectors are those which maintain the protein in a single, homogeneous, and invariant state. Reducing agents such as glutathione or 2-mercaptoethanol are useful to secure sulfhydryl groups and prevent oxidation. EDTA and EGTA are good if one wishes to protect the protein from heavy or transition metal ions or the alkaline earths. Inclusion of these components may be particularly desirable when crystallization requires a long period of time to reach completion.

When crystallization is carried out at room temperature in poly(ethylene glycol) or low-ionic-strength solutions, then attention must be given to preventing the growth of microbes.

These generally secrete proteolytic enzymes that may have serious effects on the integrity of the protein under study. Inclusion of sodium azide or thymol at low levels may be necessary to discourage invasive bacteria and fungi.

Substrates, coenzymes and inhibitors often serve to fix an enzyme in a more compact and stable form. Thus a greater degree of structural homogeneity may be imparted to a population of macromolecules and a reduced level of dynamic behavior achieved by complexing the protein with a natural ligand before attempting its crystallization.

In some cases an apoprotein and its ligand complexes may be significantly different in their physical behavior and can, in terms of crystallization, be treated as almost entirely separate problems. This may permit a second or third opportunity for growing crystals if the native apoprotein appears refractile. Thus, it is worthwhile, when determining or searching for crystallization conditions, to explore complexes of the macromolecule with substrates, coenzymes, analogues and inhibitors very early. In many ways, such complexes are inherently more interesting in a biochemical sense than the apoprotein when the structure is ultimately determined.

It should be pointed out that, just as natural substrates or inhibitors are often useful, they also can have the opposite effect of obstructing crystal formation. In such cases, care must be taken to eliminate them from the mother liquor and from the purified protein before crystallization is attempted. This is exemplified by many sugar binding proteins such as lectins. Concanavalin A and *Abrus precatorius* lectin can be crystallized only with great difficulty or not at all when glucosamine or galactose, respectively, are present. Pig pancreas α-amylase can also be crystallized only after residual oligosaccharides are removed from the preparation.

Finally, it should be noted that the use of inhibitors or other ligands may sometimes be invoked to obtain a crystal form different from that grown from the native protein. When crystals of the apoprotein are poorly suited for analysis, this may provide an alternative approach.

It was noted that microbial growth frequently results in proteolysis of protein samples, something to be avoided. This, however, is not always the case. It has been shown in a number of instances [55–59] that limited and controlled proteolytic cleavage of a protein can render it crystallizable while in the native state it was not. In other cases [60], limited proteolysis results in a change of crystal form to a more suitable and useful habit. It should be emphasized that these represent examples of controlled proteolysis where the end product is an essentially homogeneous population of molecules, albeit cleaved molecules.

Proteases, it seems, occasionally trim off loose ends or degrade macromolecules to stable, compact domains. These abbreviated proteins are, as a result, more invariant, less conformationally flexible and they often form crystals more readily than the native precursor. Although one might prefer the intact protein, a partially degraded form sometimes exhibits the activity and physical properties that are of primary interest. If a molecule can undergo limited digestion, this form should also be included in the crystallization strategy.

Various metal ions have been observed to induce or contribute to the crystallization of proteins and nucleic acids. In some instances these ions were essential for activity and it was, therefore, reasonable to expect that they might aid in maintaining certain structural features of the molecule. In other cases, however, metal ions, particularly divalent metal ions of the transition series, were found that stimulated crystal growth but played no known role in the macromolecules' activity. One of the oldest examples of an animal protein being crystallized is horse spleen ferritin that forms perfect octahedra when a solution containing the protein is exposed to concentrations of Cd^{2+} ions [61]. α-Lactalbumin was similarly shown to crystallize in the presence of this ion [62] and several varieties of α-amylase crystallize spontaneously when presented with Ca^{2+} ions [63, 64]. Metal ions should be included for investigation in that class of additives which for any reason might tend to stabilize or engender conformity by specific interaction with the macromolecule.

Typical trial arrays

It is sometimes useful for those of limited experience with protein crystallization to have a flow chart or plan in advance to guide their first efforts. Similarly, it is often helpful to have a few simple objectives firmly in mind, to know where to begin. Presented in Figs 12 and 13 are general schemes for conducting crystallization trials on a protein that has not previously been crystallized. In Figs 14 and 15 are 'details' from those schemes, elaborated to show what several initial trial elements, or arrays, might typically be like.

Initially, the parameters that one wishes to establish as rapidly as possible are optimal concentration for each precipitant used, optimal pH for solubilization and crystallization, and the effect of temperature. The two precipitants that should be examined first are ammonium sulfate and poly(ethylene glycol) 4000 as representatives of salts and poly(ethylene glycol), the two major classes of precipitants in use. If quantity of protein permits than the additional two classes of organic solvents and short chain alcohols should be investigated as well. The best representatives of these latter groups are ethanol and methylpentanediol, respectively; suggestions for their use are shown in Fig. 16.

Initially, a pH range of 3.5 – 9.0 should be explored in ΔpH intervals of 0.5 but the range should be extended, abridged, or modified in appropriate cases. Generally, it is sufficient to set up two parallel sets of trials and maintain one set at 4°C and the other at 25°C. This will provide an indication of the possible influence and value of temperature as a variable.

If crystals of any sort are obtained in the first round of trials, then the coarse matrix of conditions is more finely sampled, evaluated, and in successive rounds the growth of the crystals optimized. If no crystals are obtained, ligand complexes or alternative forms of the protein are explored. If this fails, then effectors such as metal ions and detergents are introduced, and so on.

A major consideration in screening crystallization conditions is a reduction in the number of trials that must be carried out. Even in those happy cases where the quantity of protein is not a limitation, reduction of trials means less time and effort. Thus, one seeks to avoid conditions that are certain to be unprofitable. For example, if the protein is observed to precipitate rapidly at salt concentrations greater than 50% saturation, or at pH below 5.0, or at 4°C, then clearly the trials lying beyond those limits or at that temperature can be eliminated.

The entire strategy of crystallizing proteins is often a process of picking out those areas of variable space that have some chance of yielding success and intuiting those likely to produce failure. A major difficulty in this pursuit is that only a narrow range of conclusions are possible from each crystallization trial. The mother liquor (a) contains precipitate, (b) it is clear, (c) large crystals are present, or (d) microcrystals are present. This makes it rather difficult to know how close

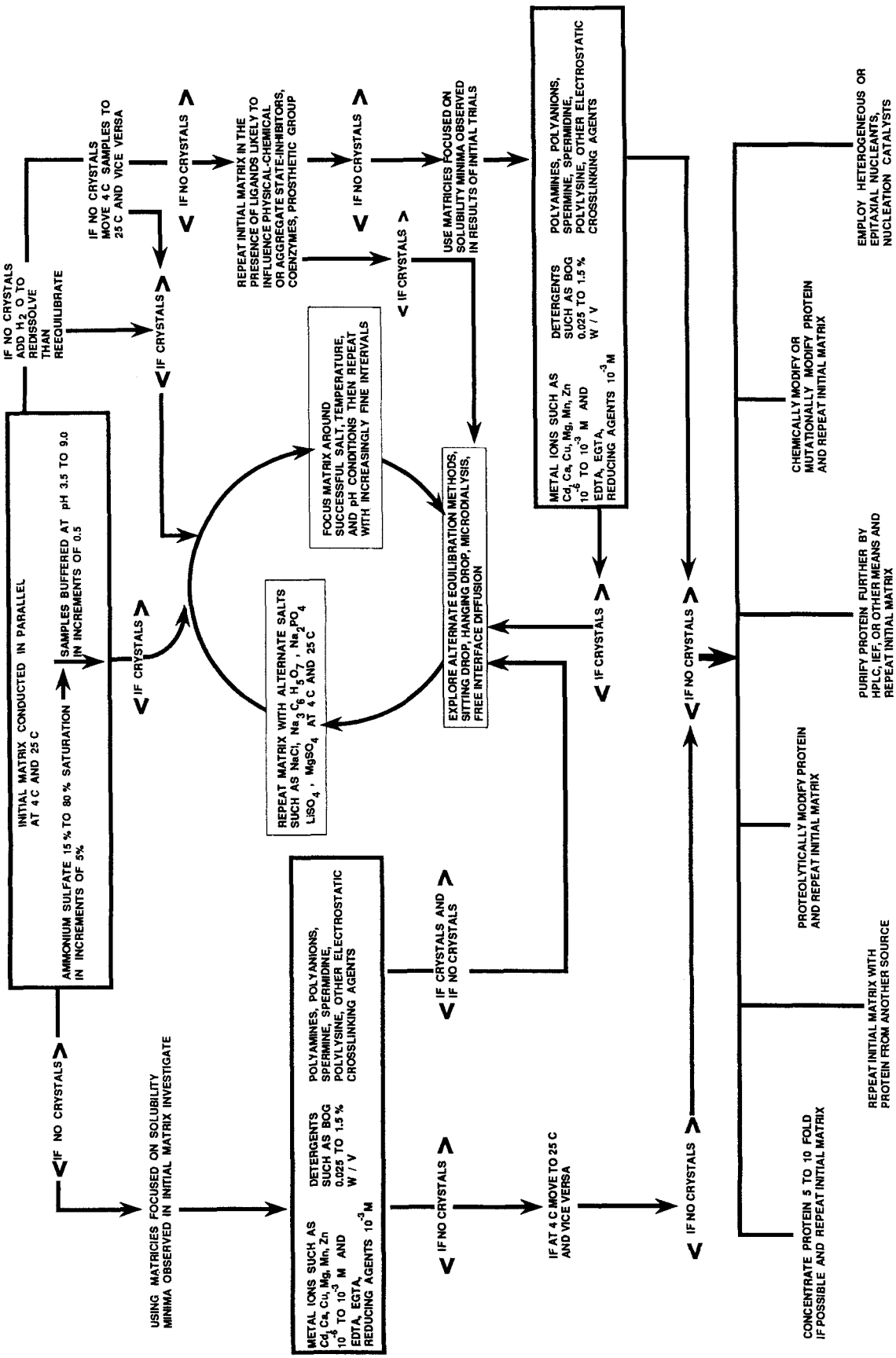


Fig. 12. *Investigation of crystallization conditions using salt.* This is a flow diagram used by the author showing the succession of variables and procedures investigated, and the order in which they are explored. Progression through the network hopefully leads to the crystallization of a protein or nucleic acid and the optimization of its growth. Other such plans of action could undoubtedly be drawn and every laboratory has its favorite variations and additions, but this diagram should serve as a guide for new experimenters. The assumption in this scheme is that ammonium sulfate will be the initial precipitate used and all other experiments will employ that, or some comparable salt. See Fig. 14 for a detailed outline of the starting matrix

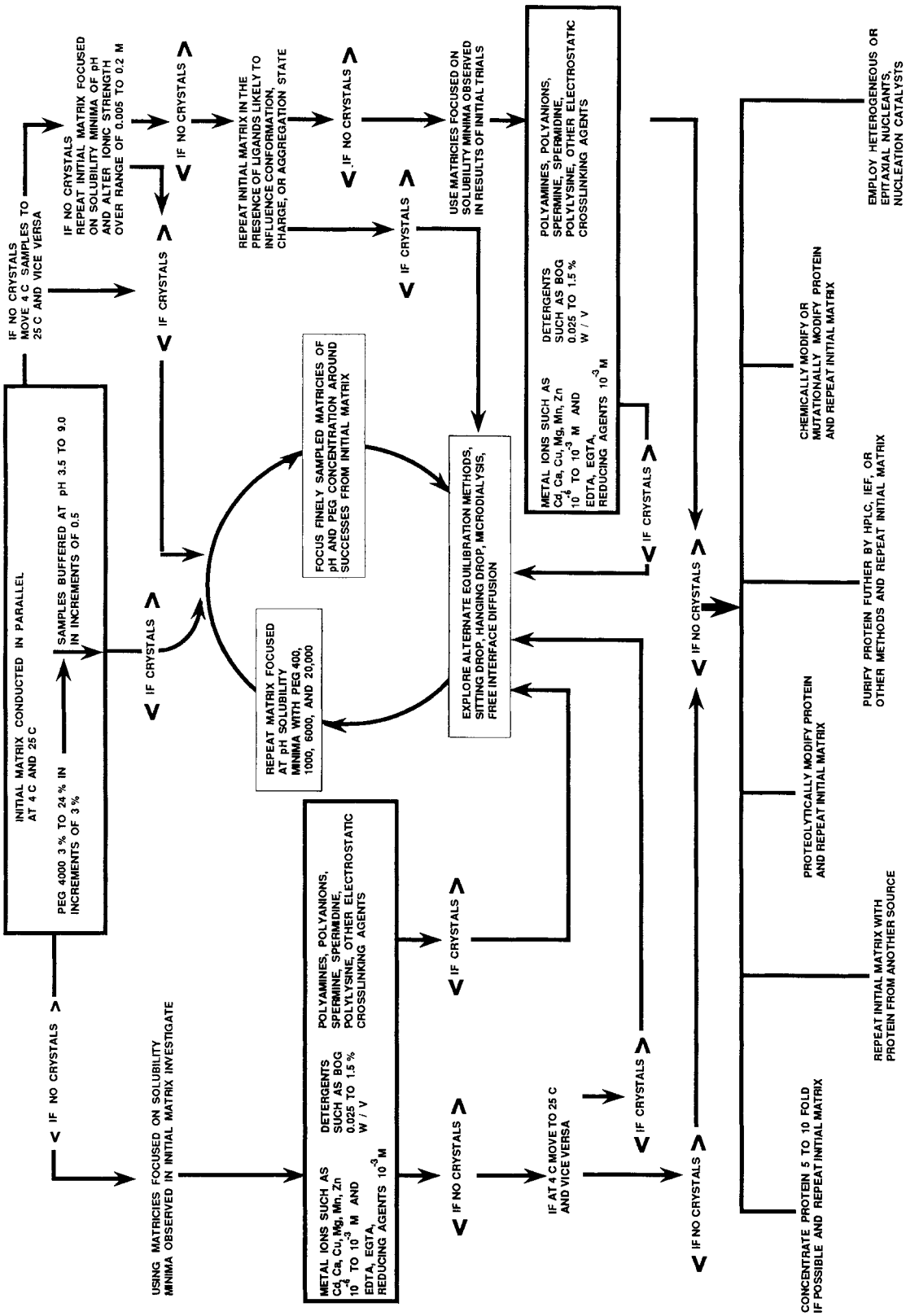


Fig. 13. *Investigation of crystallization conditions using poly(ethylene glycol)*. This is a second flow diagram, corresponding to that using salt as a precipitant (Fig. 12), but here based on poly(ethylene glycol) 4000 (PEG 4000) as the initial precipitating agent. While similar to the scheme of Fig. 12, it contains some important differences. In general, when one is attempting to grow crystals of a particular protein, the salt and poly(ethylene glycol) schemes are carried out in parallel. Often protein is limiting, and the investigator must choose between several options and decide when and how to abbreviate a specific trial matrix based on his biochemical understanding of the protein. Interpreting the results of the trials is a skill that must be developed. See Fig. 15 for a detailed outline of the starting matrix

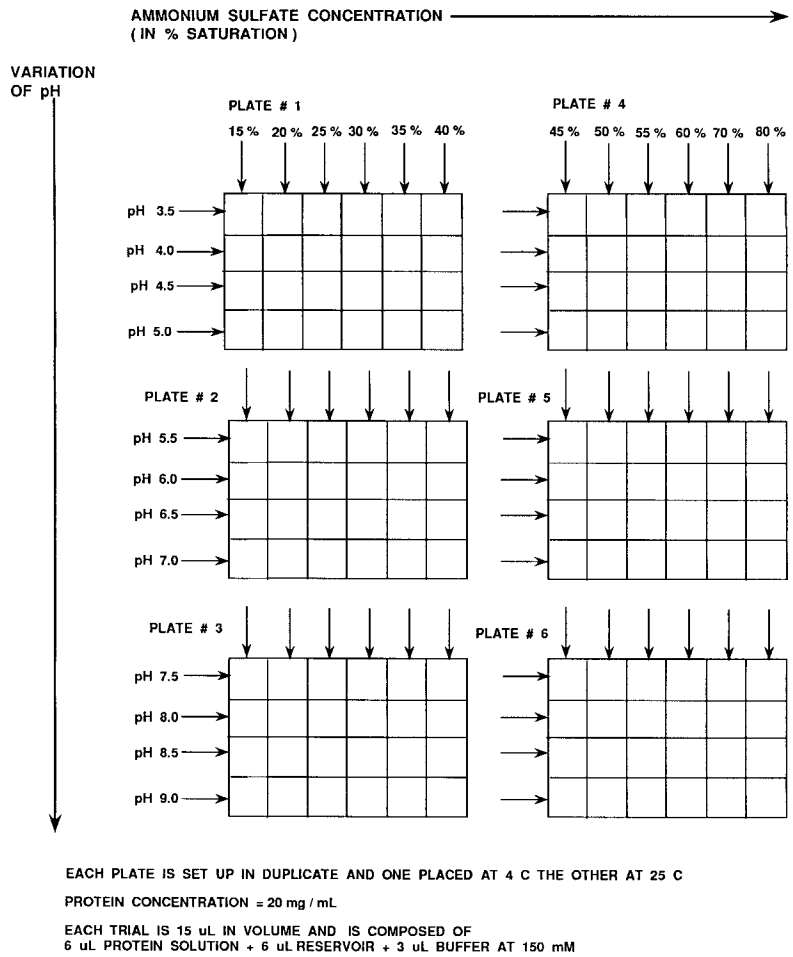


Fig. 14. *Initial salt matrix*. A detailed presentation of an initial screening matrix using ammonium sulfate as the precipitating agent and vapor diffusion by either the 'hanging drop' in a 24-well Linbro plate or the 'sitting drop' in a Cryschem 24-well vapor diffusion plate. This initial matrix investigates the effect of salt concentration and pH, generally the most important parameters, on precipitation and crystallization behavior. Equivalent matricies should be investigated at 4° and 22°C to evaluate the effect of temperature

a trial is to success unless crystals are actually present. Nevertheless, systematic approaches to the interpretation of crystallization trials are under development, and have proven useful in a number of cases [22, 65, 66].

Careful examination of precipitates formed in the mother liquor are frequently of some value. Granular precipitate, for example, sometimes is actually microcrystalline when examined under a high-power microscope; a globular or oil-like precipitate often indicates hydrophobic aggregation and suggests the use of detergents; a light, fluffy precipitate is generally a strong negative; a clear trial means a higher precipitation level is needed or another pH, and so on.

Timing is also important, and when one is carrying out initial trials it is good to examine the crystallization samples frequently, every 12 – 24 h for the first few days. In this way, conditions that cause very rapid precipitation or crystal growth can be identified. Once optimal crystal growth conditions have been precisely defined, then that is the time to lay the trials down like fine wine, in a cool, quiet place.

It is also wise to pay attention to what might be considered trivial matters. Be certain that the workplace is clean to minimize dust and microbes in the samples. When making a microdroplet, see that it is as hemispherical as possible and does not spread on the glass or plastic to yield a large surface/volume ratio. Microfilter protein samples, work quickly to avoid evaporation, do not carry on philosophical conversations while dispensing ingredients. Be alert for unusual events that may later explain anomalous results. Be patient.

The importance of protein purity and homogeneity

With regard the rate of growth of protein crystals, there are two important effects to consider: the transport of molecules to the face of a growing nucleus or crystal, and the frequency with which the molecules orient and attach themselves to the growing surface. Crystal growth rates can therefore be considered in terms of transport kinetics and attachment kinetics. For protein crystals which grow relatively slowly, transport kinetics, dependent primarily on physical forces and movements in the solution phase, is almost certainly the less important of the two, although as seen in Fig. 17, its effects are sometimes evident. There is not much doubt that the predominant limitation on the rate at which protein crystals nucleate and grow is, at least over most of the period of growth, a function of the rate of attachment.

The capture of molecules by a growing crystal surface requires, as in any multi-component chemical reaction, first, that the molecules to be incorporated have the correct orientation when they approach the crystal surface and, second, that they be in the proper chemical state to form interactions

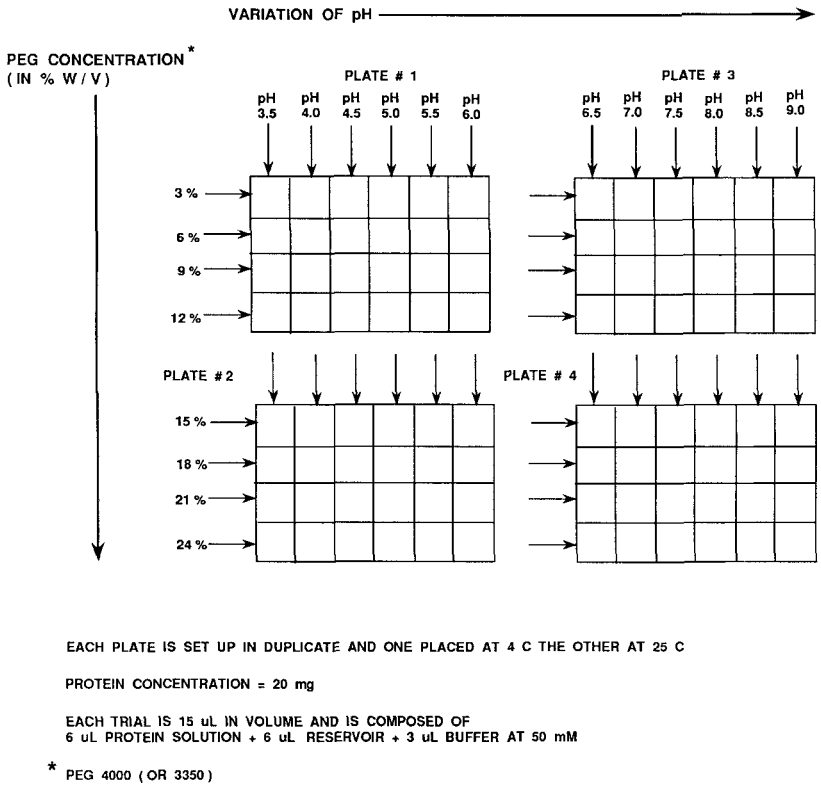


Fig. 15. *Initial poly(ethylene glycol) matrix.* A detailed plan for an initial screening matrix using poly(ethylene glycol) (PEG) the precipitating agent, but otherwise corresponding to the salt-based matrix seen in Fig. 14. Again, the matricies should be reviewed at 4° and 22°C to evaluate temperature effects, and if ligand complexes are available, these provide the basis for additional starting matricies. Either 'hanging drop' or 'sitting drop' procedures may be used in the 24-well Linbro or Cryschem plates

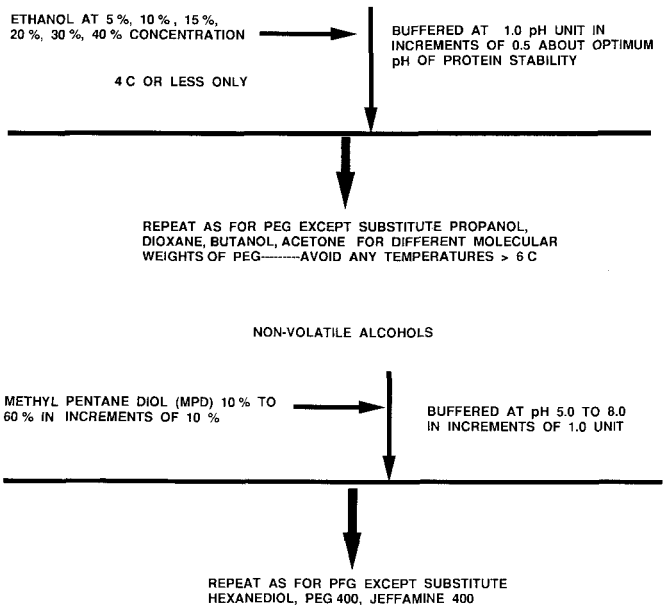


Fig. 16. *Volatile organic precipitants.* If the amount of protein is not limiting, if no success is attained with the salt or poly(ethylene glycol) approaches seen in Fig. 12 and 13, or if biochemical evidence suggests, then two additional approaches can be considered. Fundamentally, these are the same as the procedures and trials using poly(ethylene glycol) but begin and use in one case volatile organic solvents, and in the second case non-volatile reagents. These latter schemes may be particularly appropriate when the target molecule is a nucleic acid, though they have also worked well with many proteins

essential for coupling to a set of neighbors. Although there may be some things we can do to improve the statistical probability of proper orientation, there is not likely to be very much. On the other hand, we may have many opportunities to effect the frequency of attachment by enhancing the number and strength of the interactions between molecules in the lattice. We do this, for example, by optimizing the charge state of the proteins by adjusting pH, providing electrostatic crossbridges, or by minimizing the dielectric shielding between potential bonding partners by adding organic solvents such as ethanol.

Electrostatic crosslinking of protein molecules in the crystal lattice may be produced by a number of agents. This is an area of macromolecular crystallization that has been little investigated, but which the literature suggests might profitably be undertaken. We know, for example, that metal ions such as Cd^{2+} and Ca^{2+} can bridge and stabilize intermolecular contacts in crystals. This is undoubtedly the effect that causes Cd^{2+} to promote the crystallization of ferritin [61] and β-lactalbumin [67] or Ca^{2+} the crystallization of α-amylase [68]. In a similar fashion, polyamines such as spermine and spermidine have been widely used in the crystallization of nucleic acids [4], and short negatively charged oligonucleotides such as $(dA)_4$ and $(dT)_4$ were shown to be useful in promoting crystal growth of the positively charged RNase protein [69].

Certainly one major means of promoting periodic bond formation is to ensure that the population of molecules to be crystallized is as homogeneous as possible. As suggested by Table 3, this is not always straightforward. It means not only that contaminating proteins of unwanted species be elimin-

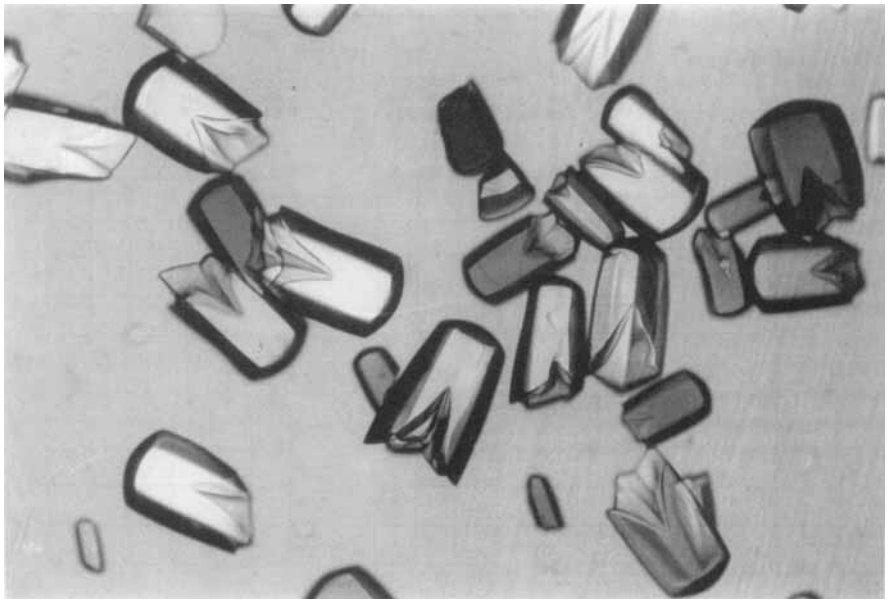


Fig. 17. *Hexagonal crystals of canavalin.* These crystals show the effects of asymmetric, rapid growth. One end of the hexagonal prisms is flat and represents the starting point of growth, the opposite end shows a deep cusp in the center arising from solute depletion during rapid growth

ated, but that within a target population all individuals assume absolute physical and chemical conformity. Because crystals have as their essential elements perfect symmetry and periodic translational relationships between molecules in the lattice, then nonuniform protein units cannot properly enter the crystal. They will not bear a proper correspondence to their neighbors. Thus, imperfect molecules will serve as inhibitors of crystal growth and bear a generally negative effect on the attachment rate. Should they enter the lattice in spite of their peculiarities, they will introduce imperfections which, by accumulation, will ultimately produce defects, dislocations, and probably termination of crystal growth.

For proteins difficult to crystallize, it is essential to take all possible measures to purify the protein free of contaminants and to do whatever is necessary to engender a state of maximum structural and chemical homogeneity. Frequently, we are misled by our standard analytical approaches, such as PAGE or IEF, into believing that a specific protein preparation is completely homogeneous. This is often illustrated for us by distinctive differences in the crystallizability of several preparations even when all analyses indicate they are identical. These imperceptible differences may be due to various degrees of microheterogeneity within preparations that lie at the margin of our ability to detect them. Table 3 lists a number of possible causes for microheterogeneity. Others could undoubtedly be added.

Table 3. *Sources of microheterogeneity*

1. Presence, absence, or variation in a bound prosthetic group, substrate, coenzyme, inhibitor, or metal ion
2. Variation in the length or composition of the carbohydrate moiety of a glycoprotein
3. Proteolytic modification of the protein during the course of isolation or crystallization
4. Oxidation of sulfhydryl groups during isolation
5. Reaction with heavy metal ions during isolation or storage
6. Presence, absence, or variation in post-translational side-chain modifications such as methylation, amidiation, and phosphorylation
7. Microheterogeneity in the amino or carboxy terminus or modification of termini
8. Variation in the aggregation or oligomer state of the protein due to association/dissociation
9. Conformational instability due to the dynamic nature of the molecule
10. Microheterogeneity due to the contribution of multiple but nonidentical genes to the coding of the protein, isozymes
11. Partial denaturation of sample
12. Genetically different animals, plants or microorganisms that make up the source of protein preparations
13. Bound lipid, nucleic acid or carbohydrate material, or substances such as detergents used in the isolation

The pronounced effects of microheterogeneity on protein crystallization have recently received much more attention from investigators than previously. Giege et al. [70] have discussed this point in detail and provided broad evidence that purification plays a crucial role in successful crystal growth. Bott et al. [71] similarly showed the pronounced beneficial effects of isoelectric focusing on an otherwise 'pure' protein.

There are occasions when even the most intense efforts to crystallize a specific protein fail in spite of the best efforts at ultra-purification and elimination of microheterogeneity. When this occurs, an alternative is to turn to a different source of the protein. Often only very small variations in amino acid sequence, as found for example between different species of organisms, is enough to produce dramatic differences in the crystallization behavior of a protein. Thus if the protein from one source proves intractable, consider another.

It might be noted that proteins manufactured in bacteria by recombinant DNA techniques appear to be especially favorable for crystallization. There are now numerous reports of such crystalline proteins in the literature. Proteins produced in this way are apparently less subject to post translational modifications and many of the other sources of microheterogeneity that characterize naturally occurring proteins. Because this technique also provides a means of amplifying the available quantity of otherwise scarce or rare proteins, it will undoubtedly play an important role in future protein crystallization strategies.

The utility of mild detergents in the crystallization of membrane proteins is now well known, and is discussed in detail elsewhere [72], but it is useful to point out here that detergents may be of value in the crystallization of otherwise soluble proteins as well [73]. Many protein molecules, particularly when they are highly concentrated and in the presence of precipitating agents such as poly(ethylene glycol) or methylpentanediol, tend to form transient and sometimes metastable non specific aggregates. The existence of a spectrum of varying sizes, shapes and charges presents problems not appreciably different from the crystallization of a protein from a heterogeneous mixture or an impure solution composed of dissimilar macromolecules. An objective in crystallizing proteins is to limit the formation of nonuniform states and reduce the population to a set of standard individuals that can form identical interactions with one another.

Indeed, evidence from inelastic light scattering experiments [65, 74, 75] suggest that the formation of nonspecific or disordered aggregates, particularly linear aggregates, may be a major obstacle to the appearance of crystals. Conditions that tend to produce a preponderance of such aggregates, therefore, are to be avoided in favor of those yielding ordered three dimensional arrangements. Many laboratories are currently investigating and developing methods to predict, even prior to the observation of microscopic crystals, which conditions favor the latter over the former.

Non-specific aggregation is primarily a consequence of hydrophobic interactions between molecules. These place few geometrical constraints on the orientations and bonding patterns between molecules that make up an oligomer. Hydrophobic contacts make proteins adhere to one another in a more or less random fashion. Hydrogen bonds and arrays of electrostatic interactions on the other hand generally demand geometrical complementarity between the protein carriers in order to form. They thereby force macromolecules to orient themselves in specific ways with respect to one another. Thus, another objective in obtaining crystals of a protein is to discourage hydrophobic interactions and to encourage those having an electrostatic basis.

A means for limiting nonspecific aggregation is the inclusion of mild,usually nonionic, detergents in the crystallization mother liquor. McPherson et al. [73] have shown that for a fairly wide range of proteins the neutral detergent octyl β-glucoside was a positive factor in obtaining crystals useful for diffraction analysis. In addition, it was demonstrated that other detergents also exhibit helpful properties in altering crystal morphology, decreasing microcrystal formation, or improving growth patterns.

Because the key to crystallizing a macromolecule successfully often lies in the procedure, means, or solvent used to solubilize it, some careful consideration should be given to this initial step. This is particularly true of membrane, lipophilic, or other proteins which, for one reason or another, are only marginally soluble in water solutions. In addition to mild detergents, for example, there are a number of chaotropic agents that can also be employed for the solubilization of proteins. These include such compounds as urea, guanidinium hydrochloride, and relatively innocuous anions such as SCN^-, ClO_4^-, I^-, Br^- and NO_3^- [76]. These compounds, even at relatively low concentrations, may serve to increase dramatically the solubility of a protein under conditions where it would otherwise be insoluble. Gradual withdrawal of the chaotrop, for example by dialysis, could then serve as a mechanism for the crystallization of the macromolecule.

Seeding

Often it is desirable to reproduce crystals of a protein previously grown where either the formation of nuclei is limiting, or spontaneous nucleation occurs at such a profound level of supersaturation that poor growth patterns result. In such cases it is desirable to induce growth in a directed fashion at low levels of supersaturation. This can sometimes be accomplished by seeding a metastable, supersaturated protein solution with crystals from earlier trials. The seeding techniques fall into two categories, those employing microcrystals as seeds and those using larger macroseeds. In both methods, the fresh solution to be seeded should be only slightly supersaturated so that controlled, slow growth will occur. The two approaches have been described in some detail by Fitzgerald [77] and by Thaller et al. [78, 79] respectively.

In the method of seeding with microcrystals, the danger is that too many nuclei will be introduced into the fresh supersaturated solution and masses of crystals will result, none of which are suitable for diffraction analysis. To overcome this, a stock solution of microcrystals is serially diluted over a very broad range. Some dilution sample in the series will, on average, have no more than one microseed per microliter. Others will have severalfold more or none at all. Then 1 µl of each sample in the series is added to fresh protein-crystallization trials under what are perceived to be the optimal conditions for growth to occur. This empirical test should, ideally, identify the correct sample to use for seeding by yielding only one or a small number of single crystals when crystal growth is completed. Seeding solutions containing too many seeds will yield additional showers of microcrystals and seeding solutions containing too low a concentration of seeds will produce nothing at all. The optimal seeding concentration as determined by the test can then be used to seed many additional samples.

The second approach to seeding involves crystals large enough to be manipulated and transferred under a microscope. Again the most important consideration is to eliminate spurious nucleation by transfer of too many seeds. Even if a single large crystal is employed, microcrystals adhering to its surface may be carried across to the fresh solution. To avoid this, it is recommended that the macroseed be thoroughly washed by passing it through a series of intermediate transfer solutions. In so doing, not only are microcrystals removed, but if the wash solutions are chosen properly, some limited dissolution of the seed may take place. This has the effect of freshening the seed crystal surfaces and promoting new growth once it is introduced into the fresh protein solution. Again, the new solution must be at least saturated with respect to protein but not extremely so in order to ensure slow and proper growth.

Seeding is frequently a useful technique for promoting nucleation of protein crystals, or initiating nucleation and growth at a lower level of supersaturation than might otherwise spontaneously occur. This can only be done, however, where crystals, even poor crystals, of the protein under investigation have previously been obtained and can be manipulated to serve as seeds. A common problem in macromolecular crystallization is inducing crystals to grow that have never previously been observed. This reflects, of course, the salient fact that the formation of stable nuclei of protein crystals is most often the single major obstacle to obtaining any crystals at all. In those cases where the immediate problem is simply growing crystals, any crystals, then attention must be focused on the nucleation problem, and any approach that might help promote nucleation should be considered.

One such technique, borrowed in part from classical small molecule crystal growth methodology, is the use of heterogeneous or epitaxial nucleants. In principle, this means the induction of growth of crystals of one substance on crystal faces of another. The classical example is galium arsenide crystals that nucleate and grow from the faces of crystals of silicon.

Because protein molecules possess chemical groups, both charged and neutral, that often readily interact with small molecules, membranes, or other surfaces, the possibility presents itself that the faces of natural and synthetic minerals might help order protein molecules at their surfaces and thereby induce the formation of ordered two dimensional arrays of the macromolecules. This ordering might occur by mechanical means due to steps and dislocations on the crystal faces or by chemical means derived from a complementarity between groups on the mineral and the protein. Such cooperation between mineral faces and nascent protein crystals might be particularly favored when the lattice dimensions of the protein unit cell are integral multiples of natural spacings in the mineral crystal.

Recently, McPherson and Schlicta [80] have shown in a series of experiments using 50 different water insoluble minerals and five different proteins that both heterogeneous nucleation and epitaxial growth of protein crystals from mineral faces do indeed occur. For each of the five proteins, certain specific sets of minerals were empirically identified that promoted nucleation and growth at earlier times and lower levels of supersaturation than occurred through spontaneous events.

A second approach to enhancing the formation of crystal nuclei has been described by Ray [81]. He introduced microdroplets of various concentrations of poly(ethylene glycol) into protein solutions that were also sufficiently high in salt concentration (approximately 50% saturated with ammonium sulfate) to support crystal growth once stable nuclei were formed. He was able to show that protein left the salt-dominated phase of the mixture and concentrated itself in the poly(ethylene glycol)-rich microdroplets, sometimes reaching effective concentrations in these droplets of several hundred milligrams/milliliter. By light microscopy techniques it was demonstrated that crystal nuclei appeared first at the surface of the droplets and then proceeded to grow into the supersaturated salt solution that surrounded them, finally reaching a terminal size appropriate for X-ray analysis. In the absence of the droplets, no crystals were ever observed to form.

These experiments are encouraging in that other, perhaps even more effective, heterogeneous precipitant/solvent systems might be found that will assist in the enhancement of crystal nucleation by what Ray refers to as 'crystallization catalysts'.

A final thought

A last word of advice regarding success. Once crystals are obtained, then that should not signal the end of the chase. Better crystals for analysis, larger crystals, a more favorable crystallographic symmetry or unit cell or crystals that diffract to a higher level of resolution might all be obtained by continued examination of conditions. The ability of a specific protein to form isomorphous heavy atom derivatives and ligand complexes is often very much dependent on the crystal lattice interactions. Thus, the search for improvements should go forth in parallel as the X-ray analysis commences.

I would like to acknowledge the capable assistance of Ms Josephine Cheung and many useful conversations with Stan Koszelak, Dave Martin, Bob Cudney, John Day, David Birdsall and Ping Ko. The writing of this review was supported in part by a grant from the National Science Foundation of the United States.

REFERENCES

1. Sumner, J. B. & Somers, G. F. (1943) *The enzymes*, Academic Press, New York.
2. McPherson, A. (1989) *Sci. Am. 260*, 62–69.
3. McPherson, A. (1976) *Methods Biochem. Anal. 23*, 249–345.
4. McPherson, A. (1982) *The preparation and analysis of protein crystals*, John Wiley and Sons, New York.
5. McPherson, A. (1985) *Methods Enzymol. 114*, 112–120.
6. McPherson, A. (1990) *Crit. Rev. Biochem.*, in the press.
7. Creighton, T. E. (1984) *Proteins: structures and molecular properties*, Freeman Co., San Francisco.
8. Schulz, G. E. & Schirmer, R. H. (1979) *Principles of protein structure*, Springer-Verlag, New York.
9. Richardson, J. (1981) *Adv. Protein Chem. 34*, 167–339.
10. Miers, H. A. & Isaac, F. (1906) *J. Chem. Soc. 89*, 413.
11. Miers, H. A. & Isaac, F. (1907) *Proc. R. Soc. Lond. A79*, 322.
12. Petrov, T. G., Trevius, E. B. & Kasatkin, A. P. (1969) *Growing crystals from solution*, Plenum, New York.
13. Buckley, H. E. (1951) *Crystal growth*, Ch. 1, pp. 1-23, John Wiley and Sons, London.
14. Feigelson, R. S. (1988) *J. Cryst. Growth 90*, 1–13.
15. Boistelle, R. & Astier, J. P. (1988) *J. Cryst. Growth 90*, 14–30.
16. Rosenberger, F. (1986) *J. Cryst. Growth 76*, 618.
17. Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491–497.
18. Glusker, J. P. & Trueblood, K. N. (1972) *Crystal structure analysis: a primer*, Oxford University Press, Oxford.
19. Stout, G. H. & Jensen, L. H. (1968) *X-ray structure determination*, MacMillan Company, New York.
20. Blundell, T. L. & Johnson, L. N. (1976) *Protein crystallography*, Academic Press, New York.
21. McPherson, A. (1987) in *Crystallography reviews* (Moore, M., ed.) vol. 1, 191–250, Gordon and Breach, London.
22. Gilliland, G. L. (1988) *J. Cryst. Growth 90*, 51–59.
23. Cohn, E. J. & Ferry, J. D. (1950) in *Proteins, amino acids and peptides* (Cohn, E. J. & Edsall, J. T., eds) Reinhold, New York.
24. Czok, R. & Bücher, Th. (1960) *Adv. Protein Chem. 15*, 315–415.
25. Northrop, J. H., Kunitz, M. & Herriott, R. M. (1948) *Crystalline enzymes*, Columbia University Press, New York.
26. Hofmeister, T. (1887) *Arch. Exp. Pathol. Pharmakol. 24*, 274.
27. Timasheff, S. N. & Arakawa, T. (1988) *J. Cryst. Growth 90*, 39–46.
28. McPherson, A. (1976) *J. Biol. Chem. 251*, 6300–6303.
29. Hermans, J. (1982) *J. Chem. Phys. 77*, 2193.
30. Lee, J. C. and Lee, L. L. Y. (1981) *J. Biol. Chem. 256*, 625–631.
31. Hirs, C. H. W., Timasheff, S. N. & Wyckoff, H. (eds) (1986) *Methods Enzymol. 114.*
32. Bailey, K. (1942) *Trans. Faraday Soc. 38*, 186.
33. Bailey, K. (1940) *Nature 145*, 934.
34. Zeppenzauer, M., Eklund, H. & Zeppenzauer, E. (1968) *Arch. Biochem. Biophys. 126*, 564.
35. Zeppenzauer, M. (1971) *Methods Enzymol. 22*, 253.
36. Salemme, F. R. (1972) *Arch. Biochem. Biophys. 151*, 533.
37. Weber, B. A. & Goodkin, P. E. (1970) *Arch. Biochem. Biophys. 141*, 489–498.
38. Hampel, A., Labanauskas, M., Conners, P. G., Kirkegard, L., RajBhandary, U. L., Sigler, P. B. & Bock, R. M. (1968) *Science 162*, 1384.
39. Morris, D., Kim, C. Y. & McPherson, A. (1989) *Biotechniques 7*, no. 5.
40. McPherson, A. (1985) *Methods Enzymol. 114*, 121–124.
41. Rosenberger, F. & Meehan, E. J. (1988) *J. Cryst. Growth 90*, 74–78.
42. Jacoby, W. B. (1968) *Anal. Biochem. 26*, 295.

43. Osborne, T. B. (1924) *The vegetable proteins*, 2nd edn, Longmans Green, London.
44. Vickery, H. B., Smith, E. L., Hubbell, R. B. & Nolan, L. S. (1941) *J. Biol. Chem. 140*, 613–624.
45. King, M. V. (1965) *J. Mol. Biol. 11*, 549.
46. Baker, E. N. & Dodson, G. (1970) *J. Mol. Biol. 54*, 605–609.
47. Kunitz, M. (1952) *J. Gen. Physol. 35*, 423.
48. King, M. V., Bello, J., Pagnatano, E. H. & Harker, D. (1962) *Acta Crystallogr. 15*, 144.
49. King, M. V., Magdoff, B. S., Adelman, M. B., & Harker, D. (1956) *Acta Crystallogr. 9*, 460.
50. Ray, W. J. & Puvathingal, J. N. (1985) *Anal. Biochem. 146*, 307.
51. Jurnak, F. (1985) *J. Mol. Biol. 185*, 215–217.
52. DeLucas, L. J., Suddath, F. L., Snyder, R., Naumann, R., Broom, M. B., Pusey, M., Yost, V., Herren, B., Carter, D., Nelson, B., Meehan, E. J., McPherson, A. & Bugg, C. E. (1986) *J. Cryst. Growth 76*, 681–693.
53. Pusey, M. L. & Nauman, R. (1986) *J. Cryst. Growth 76*, 593.
54. Pusey, M., Witherow, W. K. & Naumann, R. (1988) *J. Cryst. Growth 90*, 105–111.
55. Wyckoff, H. W., Tsernoglou, D., Hanson, A. W., Knox, J. R., Lee, B. & Richards, F. M. (1970) *J. Biol. Chem. 245*, 305.
56. McPherson, A. & Spencer, R. (1975) *Arch. Biochem. Biophys. 169*, 650–661.
57. Waller, J. P., Risler, J. L., Monteilhet, C. & Zelwer, C. (1971) *FEBS Lett. 16*, 186–188.
58. Sorensen, S. P. L. & Hoyrup, M. (1915–1917) *C. R. Trav. Lab. Carlesberg 12*, 12.
59. Solomon, A., McLaughlin, C. L., Wei, C. H. & Einstein, J. R. (1970) *J. Biol. Chem. 245*, 5289–5291.
60. Jurnak, F. A., McPherson, A., Wang, A. H. J. & Rich, A. (1980) *J. Biol. Chem. 255*, 6751–6757.
61. Granick, S. (1941) *J. Biol. Chem. 146*, 451.
62. Green, D. W. & Aschaffenburg, R. (1959) *J. Mol. Biol. 1*, 54.
63. McPherson, A. & Rich, A. (1972) *Biochim. Biophys. Acta 285*, 493–497.
64. McPherson, A. & Rich, A. (1973) *J. Ultrastruct. Res. 44*, 75–84.
65. Carter, C. W., Baldwin, E. T. & Frick, L. (1988) *J. Cryst. Growth 90*, 60–73.
66. Cox, M. J. & Weber, P. C. (1988) *J. Cryst. Growth 90*, 318–324.
67. Aschaffenburg, R., Green, D. W. & Simmons, R. M. (1965) *J. Mol. Biol. 13*, 194–201.
68. McPherson, A. & Rich, A. (1972) *Biochim. Biophys. Acta 285*, 493–497.
69. Brayer, G. D. & McPherson, A. (1981) *J. Biol. Chem. 257*, 3359–3361.
70. Giege, R., Dock, A. C., Kern, D., Lorber, B., Thierry, J. C. & Moras, D. (1986) *J. Cryst. Growth 76*, 554.
71. Bott, R. R., Navia, M. A. & Smith, J. L. (1982) *J. Biol. Chem. 257*, 9883.
72. Michael, H. (ed.) (1990) *Crit. Rev. Biochem.*, in the press.
73. McPherson, A., Koszelak, S., Axelrod, H., Day, J., Williams, R., Robinson, L., McGrath, M. & Cascio, D. (1986) *J. Biol. Chem. 261*, 1969–1975.
74. Kam, Z., Shore, H. B. & Feher, G. (1978) *J. Mol. Biol. 123*, 539.
75. Kadima, W., McPherson, A., Dunn, M. F. & Jurnak, F. A. (1990) *Biophys. J. 57*, 125–132
76. Hatefi, Y. & Hanstein, W. G. (1969) *Proc. Natl Acad. Sci. USA 62*, 1129–1136.
77. Fitzgerald, P. M. D. & Madsen, N. B. J. (1987) *J. Crystal Growth 76*, 600.
78. Thaller, C., Eicher, G., Weaver, L. H., Wilson, E., Karlsson, R., Jansonius J. N. (1985) *Methods Enzymol. 115*, 132–135.
79. Thaller, C., Weaver, L. H., Eichele, G., Wilson, E., Karlsson, R. & Jansonius, J. N. (1981) *J. Mol. Biol. 147*, 465.
80. McPherson, A. & Shlichta P. (1988) *Science 2398*, 385–387.
81. Ray, W. & Bracker, C. Jr (1986) *J. Crystal Growth 76*, 562–576.