united nations educational, scientific and cultural organization

international atomic energy agency

the
**abdus salam**
international centre for theoretical physics

SMR: 1098/2

*WORKSHOP ON THE STRUCTURE OF BIOLOGICAL MACROMOLECULES*

( 16 - 27 March 1998)

---

# *"Principles of Biological Macromolecules Crystallisation"*

presented by:

**Doriano LAMBA**
Istituto di Strutturistica Chimica "G. Giacomello"
C.N.R. - Sezione di Trieste
Area Science Park, Strada Statale 14, Km. 163.5
I-34012 Trieste, Italy
and
International Centre for Genetic Engineering and Biotechnology
Area Science Park, Padriciano 99
I-34012 Trieste, Italy

---

These are preliminary lecture notes, intended only for distribution to participants.

# Principles of Biological Macromolecules Crystallisation

*Doriano Lamba*

Istituto di Strutturistica Chimica "G. Giacomello"
C.N.R - Sezione di Trieste
Area Science Park, Strada Statale 14, Km. 163.5
I-34012 TRIESTE, Italy

&

International Centre for Genetic Engineering and Biotechnology
Area Science Park, Padriciano 99
I-34012 TRIESTE, Italy

*"The important thing in science is not so much to obtain new facts*

*as to discover new ways of thinking about them"*

Sir William Lawrence Bragg

## Crystallisation and Structural Biology

Crystal growth, which is a very old activity, has always intrigued mankind, and many philosophers and scientists have compared it with the biological process of reproduction, and it has even been speculated that the duplication of genetic material would occur through crystallisation-like mechanisms. Through the years crystallisation of macromolecules has wandered in and out of the protein chemist's domain. Before the advent of high technology - chromatography, electrophoresis, and especially, effective methods for structure determination -crystallisation played an essential role in the study of proteins because of its importance to purification. Thus, the first crystal growers were protein chemists. The first reports on protein crystals were published more than a century ago (1853) when Funke, Hünefeld, Lehman, Teichman, and others crystallised haemoglobin from the blood of various invertebrates and vertebrates. This was followed by the crystallisation of hen-egg albumin and a series of plant proteins. In 1926 Summer reported the crystallisation of urease from jack beans, soon followed by Northrop who crystallised pepsin and a series of other proteolytic enzymes. Another scientific achievement arose in 1935 when Stanley crystallised Tobacco Mosaic Virus. This situation has changed almost completely during the past 50 years because the analytical tools necessary to verify the purity become much more generally applicable and easier to use than crystallisation. As enzymologists and protein chemists grew independent of crystallisation , structural molecular biologists and, in particular, protein crystallographers became increasingly dependent on it.

The availability of cloned genes has established a powerful partnership between protein chemistry and structural molecular biology. Once again, there are important reasons for protein chemists and enzymologists to crystallise the macromolecules they study: answering the questions posed by new proteins and nucleic acids depends critically on knowledge of their three-dimensional structures; X-ray diffraction technology is common place and structure determination is no longer necessarily a bottleneck; and finally, important advances have been made in the area of protein and nucleic acid crystal growth itself. In this new environment it is hoped that enzymologists, protein chemists, and molecular biologists interested in the structure and function of a favourite gene product will reassume the central position their forbears once enjoyed with respect to crystallisation and will undertake the crystallisation themselves.

As Bränden & Tooze recently observed: "Recombinant DNA techniques have provided tools for the rapid determination of DNA sequences and, by inference, the amino acid sequences of proteins from structural genes. The number of such sequences is now increasing almost exponentially, but by themselves these sequences tell little more about the biology of the system than a New York City telephone directory tells about the function and marvels of that city" (*Introduction to Protein Structure*, Garland Publ., Inc., New York, 1991) .

However, rapid developments in preparation of crystals of macromolecules and in experimental techniques for structure analysis and refinement have led to a revolution in Structural Biology. The accelerating growth of the Brookhaven Protein Databank and the growing number of entries in the Biological Macromolecule Crystallisation Database are two measures of the usefulness of these approaches.

**Table I**. *Protein Data Bank entries loaded on February 25, 1998.*

**7163 released Atomic Co-ordinate Entries**

| **Molecule Type** | | **Experimental technique** |
|---|---|---|
| 6339 | proteins, peptides, and viruses | 5850 X-ray diffraction |
| 284 | protein/nucleic acid complexes | 1133 NMR |
| 528 | nucleic acids | 180 Theoretical Modelling |
| 12 | carbohydrates | |

**Count By Experiment**

| | **Diffraction** | **NMR** | **Model** |
|---|---|---|---|
| proteins, peptides, and viruses | 5261 | 922 | 156 |
| protein/nucleic acid complexes | 233 | 42 | 9 |
| nucleic acids | 346 | 167 | 15 |
| carbohydrates | 10 | 2 | 0 |

These factors have contributed significantly to an enormous increase in the number of laboratories performing structural studies of macromolecules to atomic resolution. Advances include:

- Recombinant DNA techniques permit production of almost any protein or nucleic acid in large amounts.
- Better X-ray detectors combined with interactive computer graphics and improved structure determination and refinement methods have yielded better quality and faster structure determinations.
- Synchrotron radiation, permit the use of extremely tiny crystals, Multiple Wavelength Anomalous Dispersion (MAD) phasing, and time-resolved studies *via* Laue techniques.
- NMR methods permit structure determination of macromolecules in solution.
- Electron microscopy (EM) techniques, for obtaining high-resolution structures.

Crystallisation of a macromolecule need not to be a tedious or fruitless exercise. In fact, for many who attempt crystallisation the experience is quite the contrary. There is a considerable evidence that a substantial majority of proteins will crystallise readily under fairly standard conditions. Familiar recipes often do not work with new proteins, provided that the protein is homogeneous, that it is sufficiently concentrated, and that scrupulous attention is paid to the cleanliness of the surfaces used for crystal growth and to the purity and concentration of all reagents. Unquestionably, the growth of macromolecular crystals is an acquired knack, however, and nearly everyone who has tried to grow crystals has at least one unhappy memory.

Until recently, crystallisation of macromolecules was rather empirical, and because of its unpredictability and frequent irreproducibility, it has long been considered an 'art' rather than a 'science'. It is only in the last decade that a real need has emerged to better understand and to rationalise the crystallisation of biological macromolecules.

Crystallisation is the greatest challenge to massive protein structure determination. Robotics permit systematic screening of solution conditions and survey of crystallisation 'parameter space'. Fluid physics specialists are entering the field *via* space research where microgravity eliminated convection and sedimentation. New ideas have led to controlled convection and 'containerless' growth on earth. One goal of microgravity work is to understand why bigger, lower mosaicity and/or better resolution crystals are obtained with 20% of the proteins tested. In a recent small molecule crystal growth report, crystals grown in microgravity have greatly reduced disorder and diffuse scattering and Bragg peaks enhancement. Since many protein crystals exhibit diffuse scattering its reduction

could enhance diffraction quality. Crystals perfection achievable today seemed inconceivable just ten years ago.

What characterises biological macromolecular crystals from small molecular crystals? In terms of morphology, one finds with macromolecular crystals the same diversity as for small molecular crystals. In terms of crystal size, however, biological macromolecular crystals are rather small, with volumes rarely exceeding 10 $mm^3$, and thus they have to be examined under a binocular microscope. Except for special usage, such as neutron diffraction, this is not too severe a limitation. Among the most striking differences between the two families of crystals are the poor mechanical properties and the high content of solvent of macromolecular crystals. These crystals are always extremely fragile and are very sensitive to external conditions. This fragility is a consequence both of the weak interactions between macromolecules within the crystal lattices and of the high solvent content (from 20% to more than 80%) in these crystals. For that reason, macromolecular crystals have to be kept in a solvent-saturated environment, otherwise dehydration will lead to crystal cracking and destruction. The high solvent content, however, has useful consequences because solvent channels permit diffusion of small molecules, a property used for the preparation of isomorphous heavy-atom derivatives needed to solve the structures. Other characteristic properties of macromolecular crystals are their rather weak optical birefringence under polarised light: colours may be intense for large crystals but less bright than for salt crystals (isotropic cubic crystals or amorphous material will not be birefringent). Also, because the building blocks composing macromolecules are enantiomers (L-amino acids in proteins - except in the case of some natural peptides- and D-sugars in nucleic acids) macromolecules will not crystallise in space group with inversion symmetries. Accordingly, out of the 230 possible space groups, macromolecules only crystallise in the 65 space groups without such inversion. Macromolecular crystals are also characterised by large unit cells with dimensions that can reach up to 1000Å for virus crystals. From a practical point of view, it is important to remember that crystal morphology is not synonymous with crystal quality. Therefore the final diagnostic of the suitability of a crystal for structural studies will always be the quality of the diffraction pattern which reveals its internal order, as is reflected at first glance by the so-called 'resolution' parameter.

## From Empirism to Rationality

The first major breakthrough towards better and easier crystallisations was the development, in the 1960s, of crystallisation micromethods (e.g. dialysis and vapour phase diffusion). It was promoted by structural projects on macromolecules reluctant to crystallise easily and available in limited amounts. Further significant improvements came

from the discovery of specific properties of additives to be included in the crystallisation solvents, such as polyamines and non-ionic-detergents which gave the clue for crystallising nucleic acids or membrane proteins. More recently, the perception of the importance of purity for growing better crystals was an important achievement in the field and the adequate choice of the biological material was an important determinant for the success of many crystallisations. With the ribosome, for instance, the preparation of homogeneous particles from halophilic or thermophilic bacteria, instead of from mesophilic bacteria, considerably improved crystal quality. Also, the ease of synthesising oligonucleotides with the automated methods, or the development of genetic engineering technologies for over-expression of proteins, explains the increasing number of crystallised DNA fragments or rare proteins.

Biocrystallisation, like any crystallisation, is a multiparametric process, involving the three classical steps of nucleation, growth, and cessation of growth. What makes crystal growth of biological macromolecules different is, firstly, the much larger number of parameters than those involved in small molecule crystal growth and, secondly, the peculiar physico-chemical properties of these compounds.

For instance, their optimal stability in aqueous media is restricted to a rather narrow temperature and pH range. But the main difference from small molecules crystal growth is the conformational flexibility and chemical versatility of macromolecules, and their consequent greater sensitivity to external conditions. This complexity is the main reason, why systematic investigations were not undertaken earlier.

**Table II**. *Parameters affecting the crystallisation of macromolecules.*

Intrinsic physico-chemical parameters:

-supersaturation
-temperature
-time
-ionic strength and purity of the chemicals
-diffusion and convection
-volume and geometry of samples and set-ups
-solid particles, wall and interface effects
-density and viscosity effects
-pressure, electric filed, and magnetic fields
-vibration and sound
-sequence of events

Biochemical and biophysical parameters:

-sensitivity of conformations to physical parameters
-binding of ligands
-specific additives
-related with properties of macromolecules
-ageing of samples

Biological parameters:

-rarity of most biological macromolecules
-biological sources and physiological state of organisms or cells
-bacterial contaminants

Purity of macromolecules:

-macromolecular contaminants
-sequence heterogeneities
-conformational heterogeneities
-batch effects

Furthermore, the importance of some parameters, such as the geometry of crystallisation vessels or the biological origin of macromolecules, had not been recognised. It is only in the last few years that the hierarchy of parameters have been perceived. A practical consequence of this new perception was the development of statistical methods to screen crystallisation conditions. For a rational design of growth conditions, however, physical and biological parameters have to be controlled.

Because macromolecules are extracted from complex biological mixtures, purification plays an extremely important role in crystallogenesis. Purity, however, is not an absolute requirement since crystals of macromolecules can sometimes be obtained from mixtures, but such crystals are mostly small or grow as polycrystalline masses, are not well shaped, and are of bad diffraction quality, and thus cannot be used for diffraction studies. However, crystallisation of macromolecules from mixtures may be used as a tool for purification. For the purpose of X-ray crystallography, high-quality monocrystals of appreciable size (0.2mm at least for the dimension of a face) are needed. It is common belief that poor purity is the most common cause of unsuccessful crystallisation, and for crystallogenesis the purity requirements of macromolecules have to be higher than in other fields of molecular biology. Purity has to be of 'crystallographic grade': the macromolecules not only have to be pure in terms of lack of contaminants, they have also to be conformationally 'pure'. Denaturated macromolecules, or macromolecules with structural microheterogeneities, adversely affect crystal growth more than do unrelated molecules, especially when structural heterogeneities concern domains involved in crystal packing. On the other hand presence of microquantities of proteases (or nucleases) can alter the structure of the macromolecules during storage or the rather long time needed for crystallisation. As a consequence, when starting a crystallisation project one has to be primarily concerned with purification methodologies and to take all precautions against protease and nuclease action. To have reproducible results, the physiological state of cells should be controlled, because protease or nuclease levels may vary as well as the balance between cellular components. As a general rule, batches of macromolecules should not be mixed and crystallisation experiments should be conducted on fresh material so that ageing phenomena are limited.

To grow crystals of any compound, molecules have to be brought in a supersaturated, thermodynamically unstable state, which may develop in a crystalline or amorphous phase when it returns to equilibrium. Supersaturation can be achieved by slow evaporation of the solvent or by varying parameters. The recent use of pressure as a parameter is of note. From this it follows that knowledge of macromolecular solubilities is a prerequisite for controlling crystallisation conditions. However, the theoretical background underlying solubility is still controversial, especially regarding salt effects, so that solubility data almost always originates from experimental determinations. Recently,

quantitative methods permitting such determinations on small protein samples were described. The main output was the experimental demonstration of the complexity of solubility behaviours, emphasising the importance of phase-diagram determinations for a rational design of crystal growth.

As to the nature of the salt used to reach supersaturation, one can wonder why ammonium sulphate is so frequently chosen by crystal growers. This usage is in fact incidental and results from the practice of biochemists for salting-out proteins. Indeed many others salts can be employed, but their effectiveness for inducing crystallisation is variable. The practical consequence is that protein supersaturation can be reached or changed in a large concentration range of protein and salt, provided that adequate salts are used.

Because proteins and nucleic acids require defined pH and ionic strength for stability and function, biomacromolecule crystals have to be grown from chemically rather complex aqueous solutions. Crystallisation starts by a nucleation phase (i.e. the formation of the first ordered aggregates) which is followed by a growth phase. Nucleation conditions are sometimes difficult to reproduce, and thus seeding procedures with pre-formed crystalline material should not be overlooked as in many cases represent the only method to obtain reproducible results. It should be noticed that nucleation requires a greater supersaturation than growth, and that crystallisation rates increase when supersaturation increases. Thus nucleation growth should be uncoupled, which is almost never done consciously but which occurs sometimes under uncontrolled laboratory conditions. From a practical point of view, interface or wall effects as well as shape and volume of drops can affect nucleation or growth, and consequently the geometry of crystallisation chambers or drops has to be defined.

Cessation of growth can have several causes. apart from trivial ones, like depletion of the macromolecules from crystallising media, it can result from growth defects, poisoning of the faces, or ageing of the molecules. Better control growth conditions, in particular of the flow of the molecules around the crystals, may in some cases overcome the drawbacks as was shown in microgravity experiments.

Only a few attempts have been published describing the active control of crystallisation experiments. In general, the history of experiments is not well known, because crystal growers do not monitor parameters. This is especially the case for temperature, which is almost never known with accuracy, even if experiments are conducted in thermostated cabinets. Also the kinetics of events are practically never monitored.

In conclusion, the rational approach for prompt crystallisations will demand a synergy between biochemically- and physically-directed research and usage of

automated methods for control of nucleation and growth, as well as rapid preparation of high-quality monocrystals.

## References


–A. McPherson (1982). *Preparation and Analysis of Protein Crystals*. New York: John Wiley.

–A. McPherson (1990). *"Current Approaches to Macromolecular Crystallization"* in *Eur. J. Biochem.* , **189**, 1-23.

–C.W. Carter Jr. (1990). Editor. *Protein and Nucleic Acid Crystallisation. In METHODS. A Companion to Methods in Enzymology*, Vol.1 San Diego: Academic Press.

–P.C. Weber (1991). *"Physical Principles of Protein Crystallisation"* in *Advances In Protein Chemistry*, **41**, 1-36.

–A. Ducruix & R. Giege' (1992). Editors. *Crystallisation of Nucleic Acids and Proteins, A Practical Approach*. Oxford: IRL Press/Oxford Univ. Press.

–J.P. Glusker (1994). Editor. *Biological Crystallography* in *Acta Crystallographica* **D50**, 337-666.

–A. McPherson, A.J. Malkin and Y.G. Kuznetsov (1995). *"The Science of Macromolecular Crystallisation"* in *Structure*, **3**, 759-768.

–G.L. Gilliland & J.E. Ladner (1996). *"Crystallisation of Biological Macromolecules for X-ray Diffraction Studies"* in *Current Opinion in Structural Biology*, **6**, 595-603.

–A. McPherson (1997). *"Recent Advances in the Microgravity Crystallisation of Biological Macromolecules"* in *Trends in Biotechnology*, **15**, 197-237.

–C. Ostermeier & H. Michel (1997). *"Crystallisation of Membrane Proteins"* in *Current Opinion in Structural Biology*, **7**, 697-701.


## Web Sites of Interest

**–Protein Database**
http://www.pdb.bnl.gov/

**–Nucleic Acid Database**
http://ndb-mirror-2.rutgers.edu/

**–Biological Macromolecule Crystallisation Database and the NASA Archive for Protein Crystal Growth Data**
http://ibm4.carb.nist.gov:4400/

**–Crystallisation Research Tools**
http://www.hamptonresearch.com/

Eur. J. Biochem. *189*, 1 – 23 (1990)
© FEBS 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.

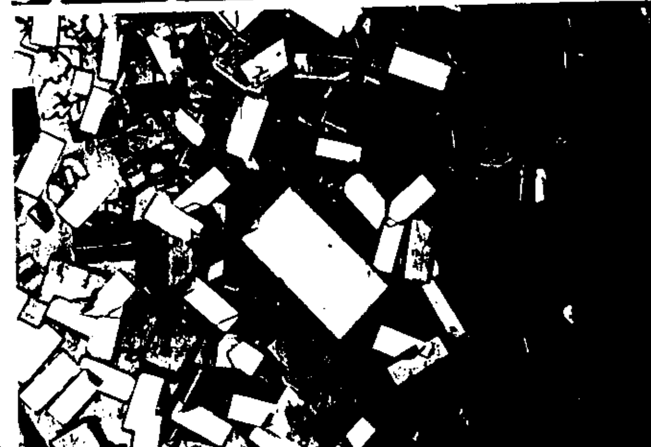
As a substitute for the precise and reasoned approaches that we commonly apply to scientific problems, we are forced, for the time being at least, to employ a strictly empirical methodology. Macromolecular crystallization is, thus, a matter of searching, as systematically as possible, the ranges of the 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*

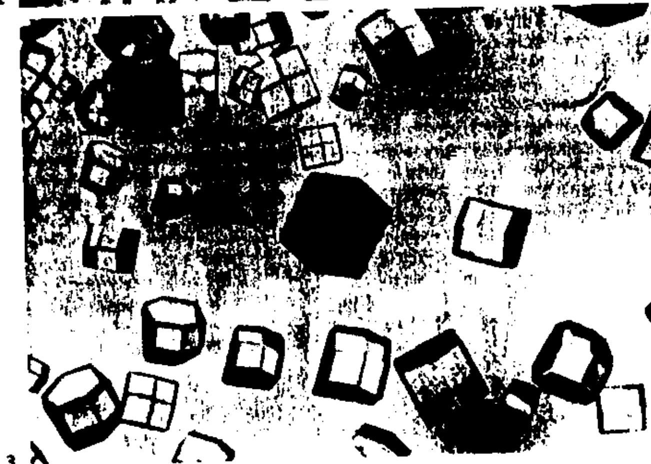
In a saturated solution, including one saturated with respect to protein, two states exist in equilibrium, the solid phase, and one consisting of molecules free in solution. At saturation, no net increase in the proportion of solid phase can accrue since it would be counterbalanced by an equivalent dissolution.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

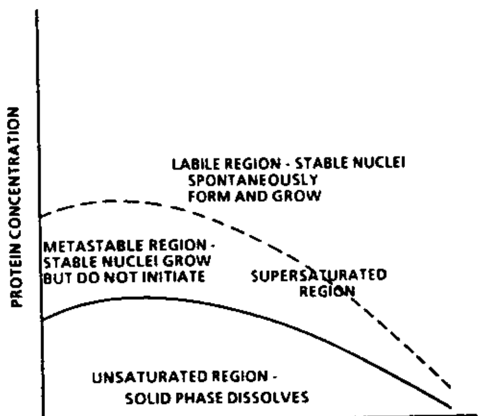



Fig. 4. *A phase diagram for a hypothetical protein showing its solubility as a function of precipitant concentration.* The solid line represents the maximum solubility or saturation curve, for the protein. Note that the protein is less soluble at very low and very high concentrations of the precipitant, 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

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

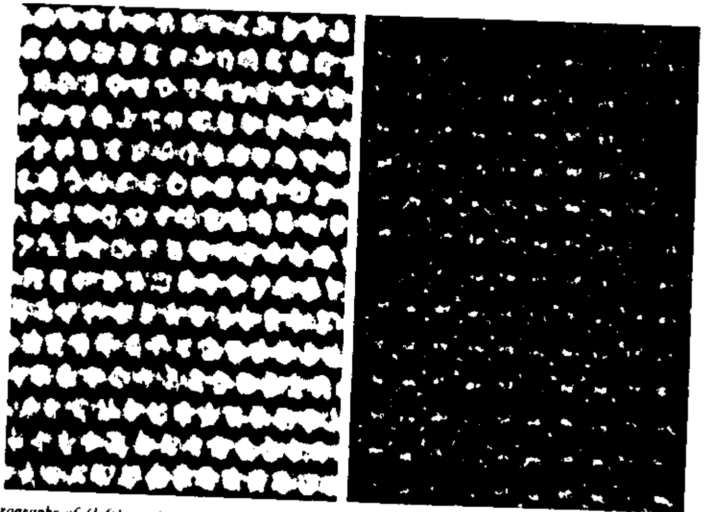


Fig. 7. *Electron micrographs of (left) a microcrystal of pig pancreas α-amylase and (right) a microcrystal of orthorhombic canavalin.* Th microcrystals are negatively stained with uranyl acetate so that the light-colored areas represent protein while the darker areas reflect th presence of the stain.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

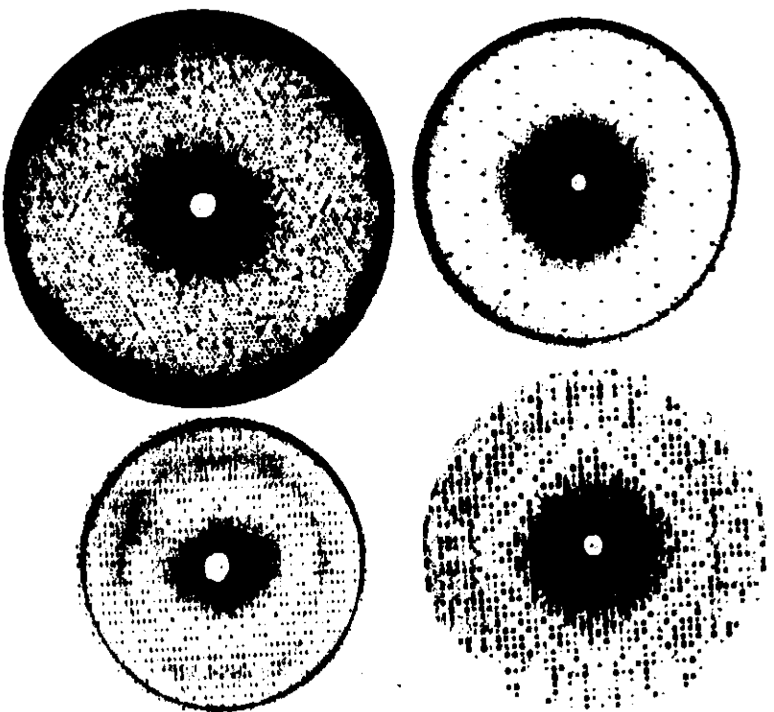


Fig. 8. *Four examples of good X-ray diffraction photographs obtained from different protein crystals.* Upper left, a hexagonal crystal of canavalin showing sixfold symmetry; upper right, a monoclinic crystal of the gene-5 DNA-unwinding protein with mm symmetry; lower left, an orthorhombic crystal of 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 diffract 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,

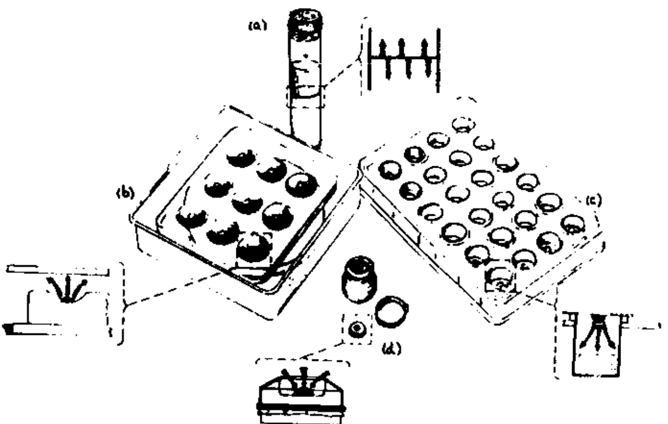



Fig. 9. *A drawing of an array of the most common microtechniques currently in use for the crystallization of macromolecules.* (a) The free interface diffusion technique; (b and c) two useful vapor diffusion methods using sitting drops on glass depression plates and hanging drops in tissue culture plates; (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

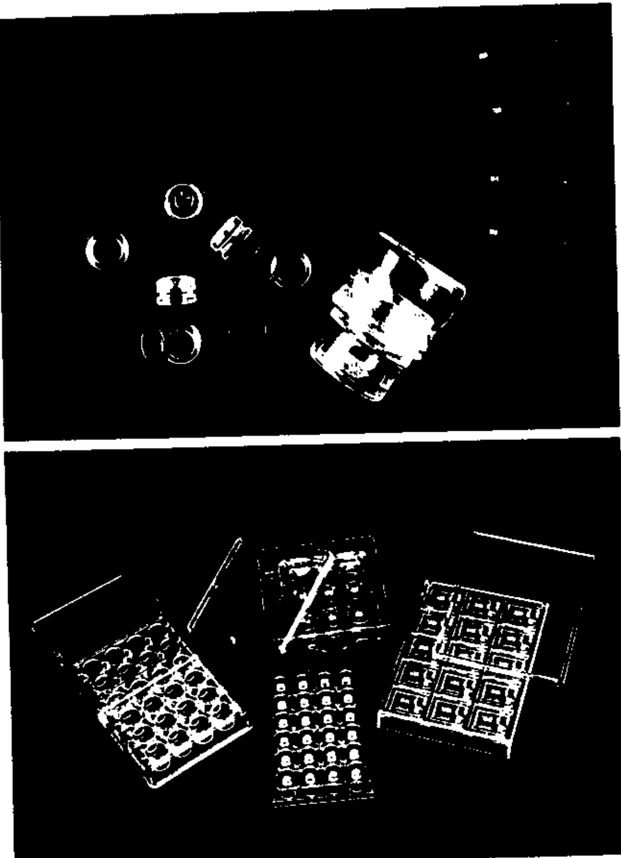


Fig. 10. *Plexiglass dialysis buttons with the O rings than maintain the dialysis membrane in place, and a small weighing bottle that serves to hold the exterior reservoir and the button.* Some microcapillaries are seen at top in which free interface diffusion can be carried out on a microscale

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 I |
| 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. If crystallization proceeds too rapidly, addition of some neutral salt may be used to slow growth and better effect crysta[l] 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.

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

### *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 jus[t] 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 $\Delta pH$ 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 $\Delta$pH 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

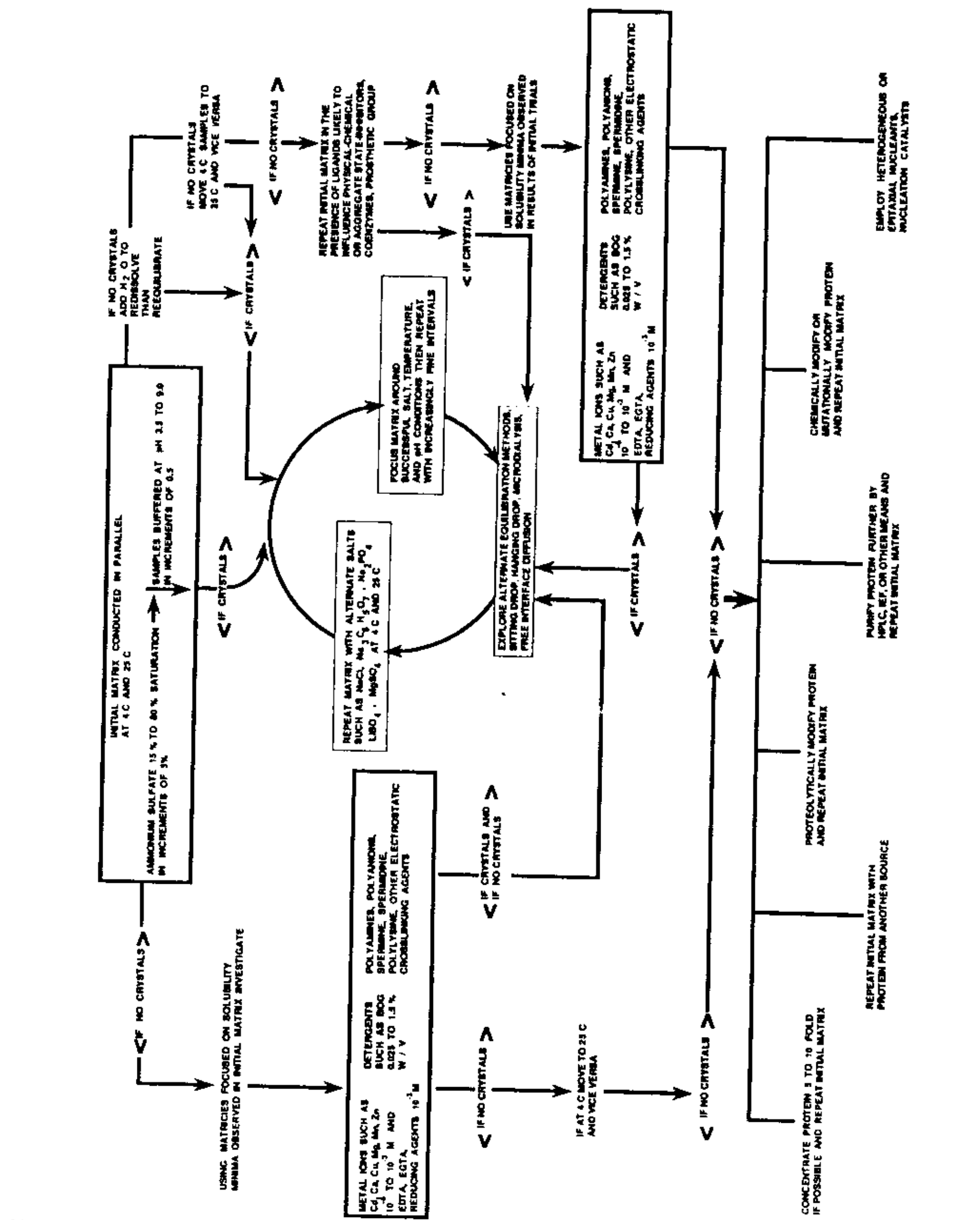



Fig. 12. *Investigation of crystallization conditions using salt.* This is a flow diagram used by the author showing the succession of variables and procedures investigated, and the order in which they are explored. Progression through the network hopefully leads to the crystallization of a protein or 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

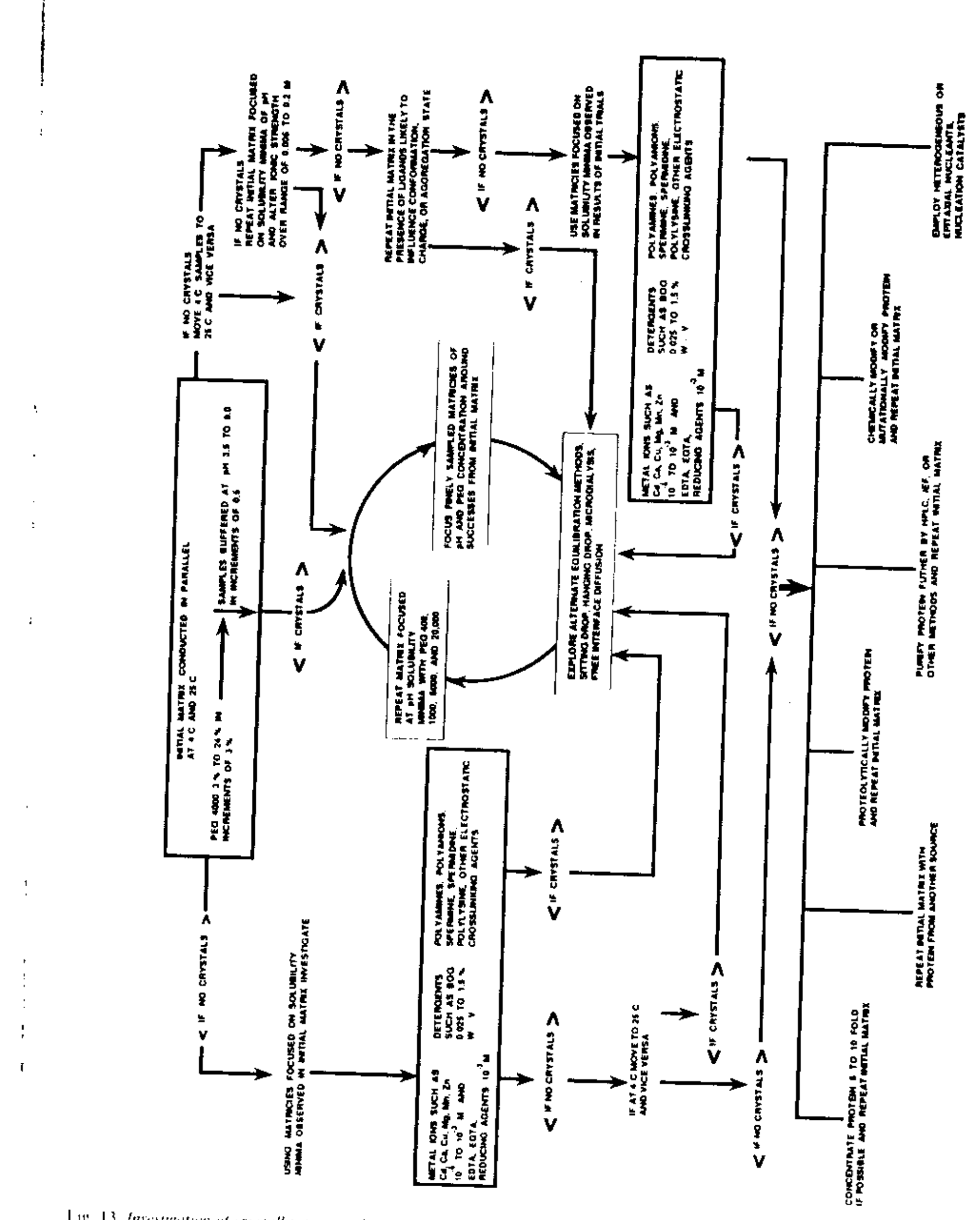



Fig. 13. *Investigation of crystallization conditions using poly(ethylene glycol).* This is a second flow diagram, corresponding to that using salt as a precipitant (Fig. 12), but here based on poly(ethylene glycol) 4000 (PEG 4000) as the initial precipitating agent. While similar to the scheme of Fig. 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

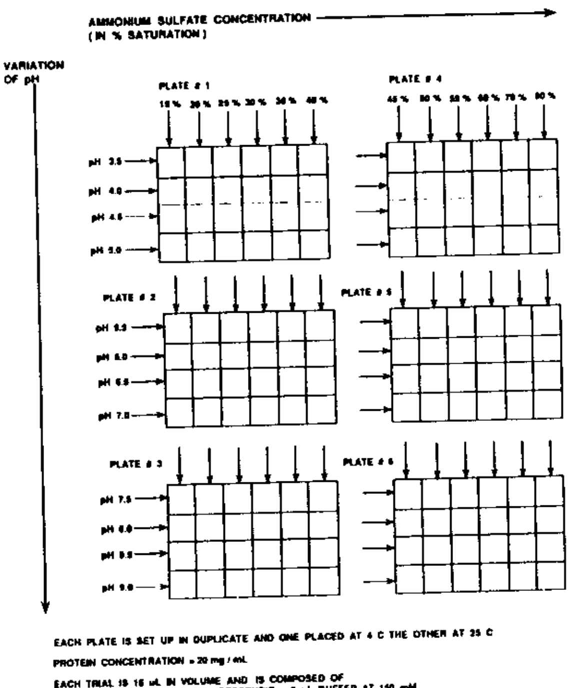



Fig. 14. *Initial salt matrix*. A detailed presentation of an initial screening matrix using ammonium sulfate as the precipitating agent and vapor diffusion by either the 'hanging drop' in a 24-well Linbro plate or the 'sitting drop' in a Cryschem 24-well vapor diffusion plate. This initial matrix investigates 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



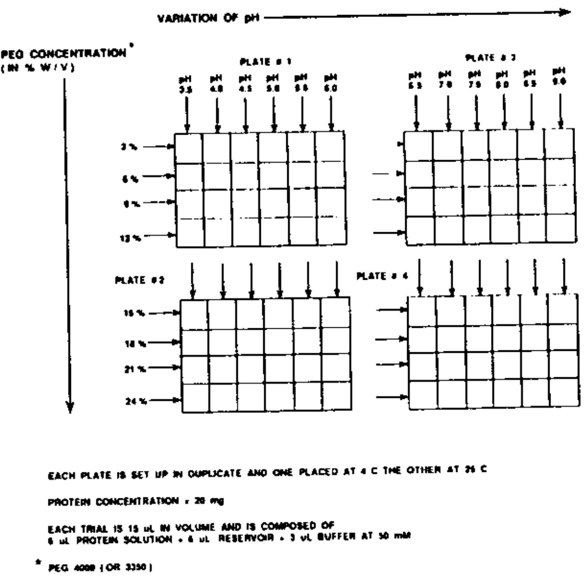



Fig. 15. *Initial poly(ethylene glycol) matrix*. A detailed plan for an initial screening matrix using poly(ethylene glycol) (PEG) the precipitating agent, but otherwise corresponding to the salt-based matrix seen in Fig. 14. Again, the matricies should be reviewed at 4° and 22°C to evaluate temperature 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

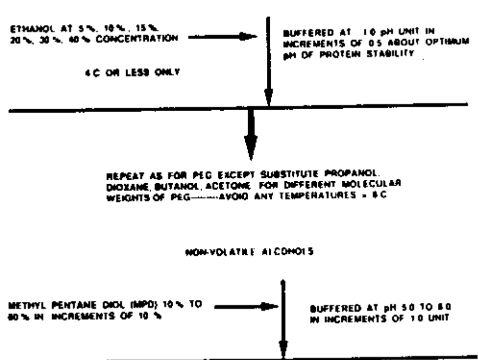



Fig. 16. *Volatile organic precipitants*. If the amount of protein is not limiting, if no success is attained with the salt or poly(ethylene glycol) approaches seen in Fig. 12 and 13, or if biochemical evidence suggests, then two additional approaches can be considered. Fundamentally, these are the same 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 polyethylene 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 $\beta$-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 protei[n] previously grown where either the formation of nuclei i[s] limiting, or spontaneous nucleation occurs at such a profoun[d] level of supersaturation that poor growth patterns result. I[n] such cases it is desirable to induce growth in a directed fashio[n] at low levels of supersaturation. This can sometimes be ac[-] complished by seeding a metastable, supersaturated protei[n] solution with crystals from earlier trials. The seeding tech[-] niques fall into two categories, those employing microcrystal[s] 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* *239B*, 385–387.
81. Ray, W. & Bracker, C. Jr (1986) *J. Crystal Growth* *76*, 562–576.

### Lysozyme and α-Lactalbumin: Structure, Function, and Interrelationships

HUGH A. MCKENZIE AND FREDERICK H. WHITE, JR.


I. Introduction . . . 174
II. Early History . . . 176
III. Some Aspects of the Occurrence, Isolation, and Characterization of Lysozyme and α-Lactalbumin . . . 181
IV. Three-Dimensional Structure of Lysozyme . . . 192
V. Three-Dimensional Structure of α-Lactalbumin . . . 206
VI. Comparative Binding of Metal Ions in Lysozyme and α-Lactalbumin . . . 214
VII. Amino Acid Composition and Sequence Homologies in Lysozyme and α-Lactalbumin . . . 224
VIII. Galactosyltransferase and the Lactose Synthase System . . . 250
IX. Some Additional Physical, Chemical, and Biological Comparisons between Lysozyme and α-Lactalbumin . . . 259
X. Evolutionary Origins of Lysozyme and α-Lactalbumin . . . 276
XI. Conclusions and the Future . . . 293
References . . . 299
Note Added in Proof . . . 315

AUTHOR INDEX . . . 317

SUBJECT INDEX . . . 337


# PHYSICAL PRINCIPLES OF PROTEIN CRYSTALLIZATION

By PATRICIA C. WEBER

Central Research & Development Department, E. I. du Pont de Nemours and Co., Inc., Wilmington, Delaware 19880


I. Introduction . . . 1
II. Stages of Crystal Growth . . . 3
III. Driving Forces for Crystal Growth . . . 3
IV. Nucleation . . . 6
A. Homogeneous Nucleation . . . 6
B. Heterogeneous Nucleation . . . 7
C. Nucleation Rate . . . 7
D. Experimental Determination of Nucleation Conditions . . . 12
V. Crystal Growth Mechanisms . . . 13
A. Transport-Controlled Growth . . . 13
B. Growth Controlled by Surface Kinetics . . . 13
C. Measurements of Crystal Growth Rates . . . 14
D. Transport Phenomena in Protein Crystal Growth . . . 15
E. Role of Molecular Preassociation in Nucleation and Crystal Growth . . . 16
VI. Competition between Nucleation and Growth . . . 17
VII. Cessation of Growth and Crystal Disorder . . . 19
VIII. Crystallization Methods . . . 20
A. Batch Method . . . 20
B. Dialysis . . . 20
C. Vapor Diffusion . . . 21
D. Temperature Shift . . . 21
E. Achieving Different Conditions for Nucleation and Growth . . . 22
F. Free Interface Diffusion . . . 22
G. Seeding . . . 23
IX. Protein Purity . . . 23
X. Searching for Crystallization Conditions . . . 25
XI. New Developments in Protein Crystallization . . . 26
A. Crystallization in Microgravity . . . 26
B. Automated Crystallization . . . 30
XII. Summary Remarks . . . 32
References . . . 33


## I. INTRODUCTION

Protein crystals are three-dimensionally ordered arrays of biological macromolecules. Although the dimensions of these crystals that sparkle and polarize light are measured in only tenths of millimeters, their ability to diffract X-rays provides the experimental data needed to image



ADVANCES IN PROTEIN CHEMISTRY, Vol. 41

Copyright © 1991 by Academic Press, Inc. All rights of reproduction in any form reserved.

biological structures at atomic resolution. The detail that can be resolved by X-ray crystallography depends on the degree of molecular and lattice ordering in the crystal. In the absence of well-ordered crystals, X-ray studies at atomic resolution are impossible. Protein crystals are also used in neutron diffraction studies and a variety of optical and magnetic resonance spectroscopies. Here, the molecular order of the crystal enhances the directional resolving power of experimental methods that include Mössbauer, electron spin resonance, circular dichroism, and Raman spectroscopies. In an emerging technology, the ordered assembly of biological macromolecules is envisioned to allow the construction of new biomaterials (Furuno and Sasabe, 1985).

Proteins are crystallized from aqueous solutions using methods that have been extensively studied for simpler molecules and salts (Rosenberger, 1986; Feigelson, 1988). Despite similar underlying physical principles, protein and small-molecule crystallizations differ in many respects. Unlike the crystallization of simpler molecules, in which solvent is effectively excluded from the crystal, substantial numbers of solvent molecules are immobilized and become ordered at protein lattice contacts, although otherwise protein crystals have large cavities containing essentially liquid water.

An important feature of protein crystal growth experiments is the need to carry out crystallization trials with very small quantities of scarce and expensive materials. When experiments are carried out in such small volumes (typically, 5–100 $\mu l$), it becomes difficult to define and control solution properties. The situation becomes particularly complicated when vapor diffusion or other nonequilibrium approaches to crystal growth are used, as these produce different and changing conditions throughout the small volumes involved.

This article reviews recent work on various aspects of the physical chemistry of protein crystal growth. Several books and reviews treat experimental and technical aspects of this area (e.g., Blundell and John- perimental and technical aspects of this area (e.g., Blundell and Johnson, 1976; McPherson, 1982, 1989; Michel, 1983; Sheshadri and Vankatappa, 1983; Matsuura, 1985; Wyckoff *et al.*, 1985; Garavito *et al.*, 1986; Ollis and White, 1989; see also the Proceedings of the First and Second International Conferences on Protein Crystal Growth [*J. Cryst. Growth* **76**, 535–718 (1986), *J. Cryst. Growth* **90**, 1–374 (1988), and Carter, 1990]. The objective here is to relate physical conceptions of how protein crystals grow in order to understand and improve existing crysprotein crystals grow in order to understand and improve existing crystallization methods. The ultimate practical goal is to allow the easy crystallization of targeted proteins in order to realize the potential utility of structural knowledge in protein engineering and drug design.



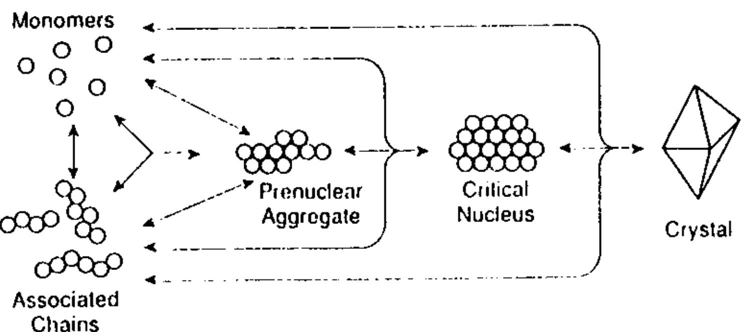



FIG. 1. Reversible molecular association reactions involved in the assembly of crystals. Monomers initially combine into small aggregates (here, called chains). The association of monomers into chains leads to the formation of prenuclear aggregates that continue to grow by further addition of monomers or chains. The partition of molecules into monomers, chains, and prenuclear aggregates is called a quasiequilibrium state (Kam *et al.*, 1978). When sufficient molecules associate in three dimensions, a thermodynamically stable critical nucleus is formed. The addition of monomers and/or chains to critical nuclei eventually leads to the formation of macroscopic crystals.

## II. Stages of Crystal Growth

Crystallization is a complex multiequilibrium process (Fig. 1). The three stages of crystallization common to all molecules are nucleation, crystal growth, and cessation of growth. During nucleation enough molecules associate in three dimensions to form a thermodynamically stable aggregate. These nuclei provide surfaces suitable for crystal growth. Crystal growth ceases when the solution is sufficiently depleted of protein molecules, deformation-induced strain destabilizes the lattice, or the growing crystal faces become poisoned by impurities.

## III. Driving Forces for Crystal Growth

Crystals form in supersaturated solutions in which the solute concentration exceeds its solution solubility. Supersaturation is usually expressed as either of the ratios $c/c_s$ or $(c - c_s)/c_s$, where $c$ is the concentration of solute before crystallization and $c_s$ is the solute equilibrium saturation concentration. Supersaturated solutions are thermodynamically metastable. Equilibrium can be restored by reducing the solute concentration through precipitation or formation of nuclei and subsequent crystal growth. The supersaturation requirements for nucleation and

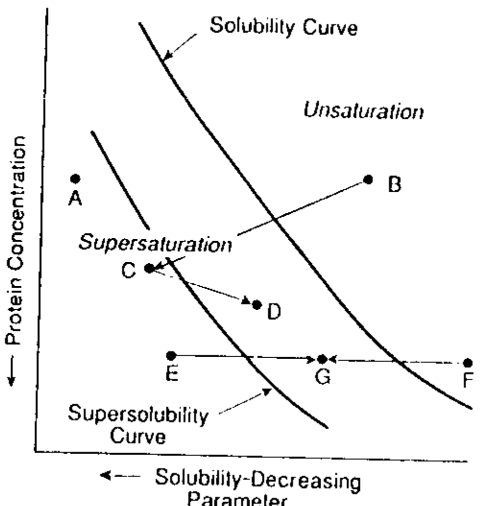



FIG. 2. A hypothetical protein solubility graph showing the changes in supersaturation for commonly used protein crystallization methods. Protein concentration is plotted as a function of a parameter that decreases protein solubility. The solubility curve divides the solubility graph into supersaturated and unsaturated regions. Supersaturated solutions support crystal growth, with increased rates observed at higher supersaturation levels (smaller abscissa values). At supersaturation levels greater than the supersolubility curve, homogeneous nucleation occurs (after Feigelson, 1988). Point A shows the supersaturation level of a batch crystallization experiment in which the protein solution is mixed with precipitating agents to achieve supersaturation and then left unchanged. The change in protein supersaturation during typical vapor diffusion experiments is shown by the line from B to C. Solutions are unsaturated on setup (B). During equilibration the solution enters the supersaturated region (C). Nuclei form when the supersaturation exceeds the supersolubility curve. If the supersaturation is then lowered by moving from C to D, only larger stable nuclei remain to support crystal growth. In a free interface diffusion experiment, when the protein and precipitant solutions are first layered, molecules at the protein–precipitant interface are sufficiently supersaturated to spontaneously nucleate (E). The remaining protein solution is unsaturated (F). On equilibration, the entire protein solution is supersaturated (G).

growth are different (Fig. 2). For a given solute, spontaneous nucleation occurs at high supersaturation, whereas lower supersaturation will support growth of a seed crystal, but not spontaneous nucleation. At concentrations below saturation, crystals dissolve.

Protein crystal growth involves the incorporation of a complex unit into an existing lattice. The growth unit usually includes the covalent polypeptide chain, water molecules that are integral components of the folded protein structure, and additional water molecules and solvent ions that may become immobilized at crystal lattice contacts. Direct interactions between protein molecules are relatively tenuous in most protein crystals that have been examined in detail (Frey *et al.*, 1988; Salemme *et al.*, 1988). Typically, water molecules that become immobilized during crystal formation serve to fill irregular gaps that occur between molecules at lattice contacts. Occasionally, intermolecular salt linkages (Baker, 1988; Dreusicke *et al.*, 1988) or counterions (Sheriff *et al.*, 1987) form electrostatic interactions at crystal contacts to stabilize the lattice structure.



Crystals are entropically destabilized, owing to both the loss of rigid-body molecular translational and rotational degrees of freedom and the immobilization of surface loops that may be flexible in solution, but become ordered at lattice contacts (Finzel and Salemme, 1985; Sheriff *et al.*, 1985; Salemme *et al.*, 1988). Although some immobilization may be a natural consequence of packing objects that tend generally to have loops on their surfaces, loop flexibility may more easily accommodate minor structural changes that facilitate incorporation of the protein into a crystalline lattice (Salemme *et al.*, 1988). Although losses of molecular entropy make unfavorable contributions to the stabilization free energies of lattice formation, some of this stabilization can be recovered due to the appearance of lattice vibrational modes, evidence for which is seen from some protein crystal studies (Finzel and Salemme, 1986; Caspar *et al.*, 1988).

Proteins are generally induced to crystallize by adding agents that either alter their surface charges, or perturb the interactions between the protein and bulk solvent water to promote associations that lead to crystallization. While many proteins crystallize near their isoelectric points in low ionic strength solutions (Blundell and Johnson, 1976), it is more common to use organic molecules, polymers, or salts at high concentrations to promote crystal growth. Most "precipitants" change the chemical potential of the protein in solution and act by affecting the partition of water between the protein and the precipitant (Timasheff and Arakawa, 1988). The protein usually has a higher affinity for water than it has for the precipitant. The preferential interaction with water creates a precipitant-poor layer near the protein surface (Fig. 3). Formation of the exclusion layer is thermodynamically unfavorable. Protein association is favored because it decreases the area of the precipitant-poor layer near the protein surface (Fig. 3). Conversely, the additives may also stabilize protein structure by concentrating water near the protein surface and possibly favoring more compact structural organizations (Arakawa and Timasheff, 1984). Random-chain polymers (e.g., polyethylene glycol) are also frequently used to promote crystallization. These polymers act to preferentially hydrate the protein through excluded volume effects,

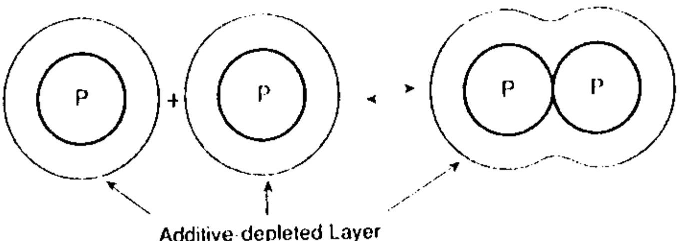



FIG. 3. Preferential protein hydration in the presence of precipitating agents used in crystallization experiments. When high concentrations of salts are used as precipitants, a precipitant-poor layer forms near the protein (P) surface due to a higher affinity of the protein for water than for the precipitant. Other precipitants (e.g., polyethylene glycol polymers) induce formation of a similar precipitant-depleted region near the protein by solvent exclusion effects. In either case formation of the precipitant-depleted layer is energetically unfavorable. Consequently, the overall effect of precipitants is to promote molecular associations that decrease the total protein surface area exposed to solvent. After Timasheff and Arakawa (1988).

whereby the extended chain polymer and its entrained water are excluded from the area near the protein (Arakawa and Timasheff, 1985). Similar to the situation with salts, this system becomes energetically more favorable when the protein molecules associate to minimize unfavorable surface tension effects.

## IV. Nucleation

### A. *Homogeneous Nucleation*

The smallest stable unit of a crystal is the nucleus. Nuclei are formed by either homogeneous or heterogeneous nucleation. Homogeneous nucleation is the spontaneous formation of solute nuclei in a supersaturated solution. In the absence of external changes, the force for spontaneous nucleation arises from fluctuations in solution. The energy required to form stable nuclei from monomeric species in solution is the sum of opposing free-energy terms. With increasing incorporation of molecules, the nucleus becomes more stable as favorable intermolecular contacts form in the three-dimensional lattice. However, formation of the nuclear surface produces an energetically unfavorable surface tension contribution. The incremental increase in surface tension on the addition of molecules to the nucleus becomes smaller as nuclei become



larger. Consequently, beyond a critical size, the energetically favorable volume term becomes dominant and nuclei are stable. The activation energy of nucleation decreases at higher supersaturation and increased temperatures. The critical nuclear size (i.e., the number of molecules needed to form a stable nucleus) also decreases with increasing supersaturation (Kam *et al.*, 1978; Boistelle and Astier, 1988).

### B. *Heterogeneous Nucleation*

Heterogeneous nucleation is the formation of solute nuclei on foreign substrates such as dust particles or surface irregularities in the container. The activation energy for heterogeneous nucleation is less than that for homogeneous nucleation, due to an attraction between the solute and the nucleant, so that heterogeneous nucleation occurs at lower supersaturation. Frequently, crystals grow on foreign nucleation sites and never appear in the bulk solution. Although the nucleation activation energy is lower for heterogeneous nucleation, the critical dimensions are similar for nuclei formed by heterogeneous and homogeneous nucleation mechanisms (Boistelle and Astier, 1988). Although heterogeneous nucleation of protein crystals frequently occurs accidentally, systematic studies of nucleation on mineral substrates demonstrate successful protein crystal growth (McPherson and Shlichta, 1988) (Fig. 4A). Several different minerals were tested with each of several proteins. Usually, a given protein crystallized on only a subset of minerals, indicating that heterogeneous nucleation involved some fairly specific interaction between the protein molecules and the mineral surface. Nucleation and crystal growth occurred more rapidly and at lower supersaturation than in the absence of the mineral nucleant. Interestingly, epitaxial growth was observed for lysozyme on the mineral apophyllite (McPherson and Shlichta, 1988). In this case a surface of the apophyllite crystal presents a two-dimensional lattice repeat that is a nearly exact fraction of the lysozyme cell dimensions.

### C. *Nucleation Rate*

Crystal nucleation rates, expressed as the number of nuclei formed per unit volume per unit time, increase with protein solubility. Higher solubility leads to increased molecular encounters in solution and reduced levels of supersaturation required for spontaneous nucleation. Nucleation rates typically show a high-power dependence on protein supersaturation, and so empirically increase rapidly above a critical value

of protein supersaturation (Boistelle and Astier, 1988). At supersaturation levels below the critical value, the slow nucleation rates lead to long times between attaining supersaturation and achieving nucleation. The slight difference between supersaturation levels where nucleation is very slow and very fast makes the nucleation rate difficult to control experimentally. The exponential dependence of nucleation rate on supersaturation has been studied in detail for sickle cell hemoglobin (hemoglobin S) (Hofrichter *et al.*, 1974, 1976) and tetragonal lysozyme (Ataka and Tanaka, 1986). In both cases the elapsed time prior to crystal appearance depended on a high power of supersaturation. For hemoglobin S, nucleation rates were also significantly faster at higher temperatures

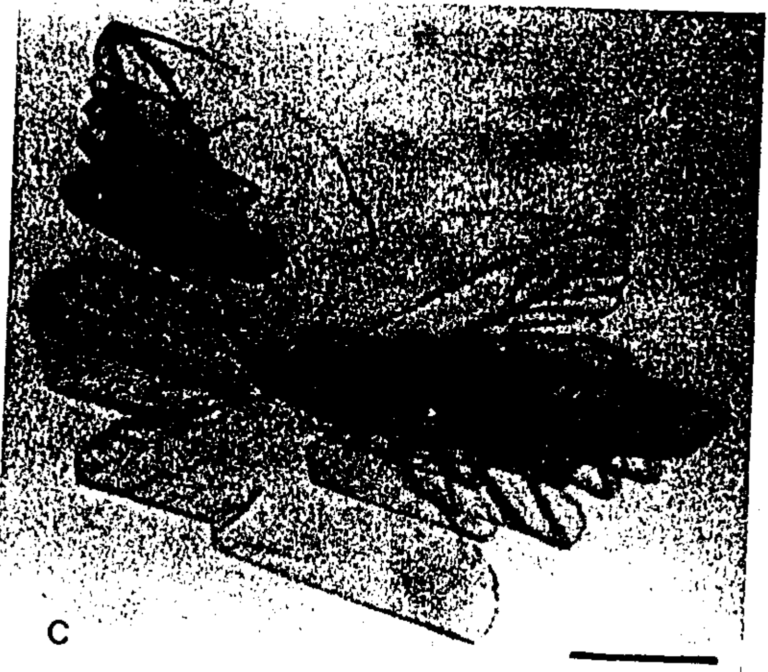
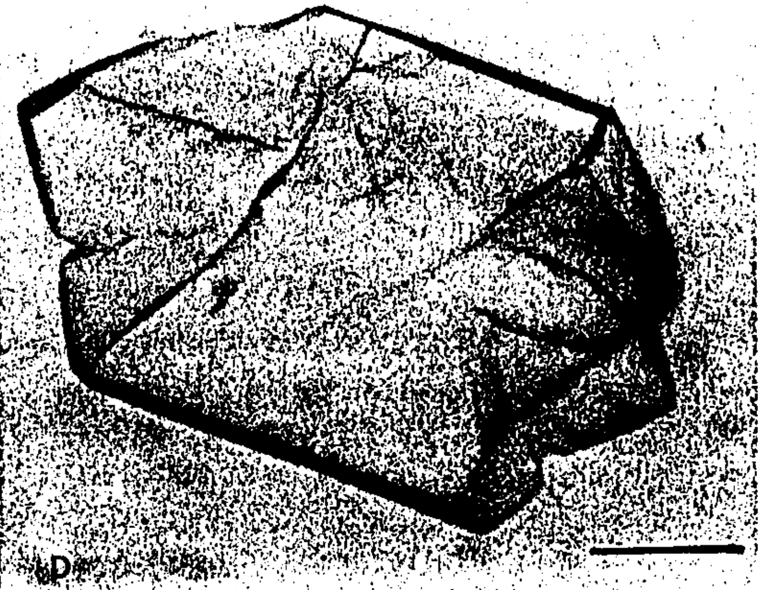



FIG. 4. Crystal photographs. (A) The heterogeneous nucleation of *Streptomyces avidinii* streptavidin crystals on the mineral biotite. (B) Streptavidin crystals are shown growing at the surface of a hanging drop. (C and D) *Pseudomonas indigofera* isocitrate lyase crystals grown (C) on earth and (D) in microgravity aboard the STS-26 space shuttle. (E–H) Crystals grown using an automated pipetting device are shown. Crystals of recombinant human interleukin 1$\beta$. [(E) Gilliland *et al.* (1987); D. B. Carter *et al.* (1988)] and apostreptavidin [(F) Pahler *et al.* (1987)] were reproduced from conditions reported in the literature. Crystallization conditions for (G) *E. coli* ketol-acid reductoisomerase and (H) a Fab fragment of a monoclonal antibody to angiotensin were found using successive automated grid searches (Cox and Weber, 1988). Bar (A–H): 0.1 mm.



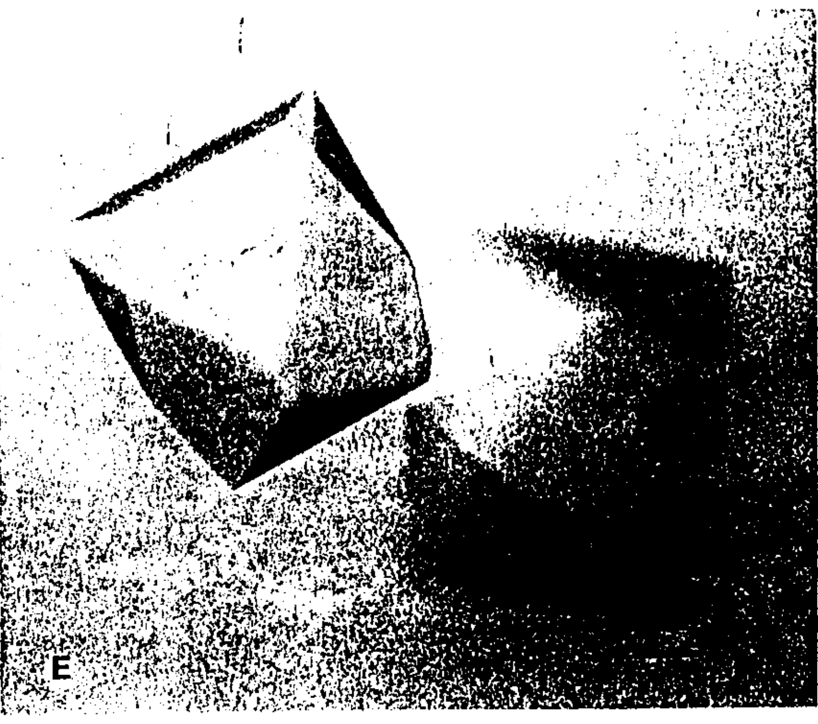
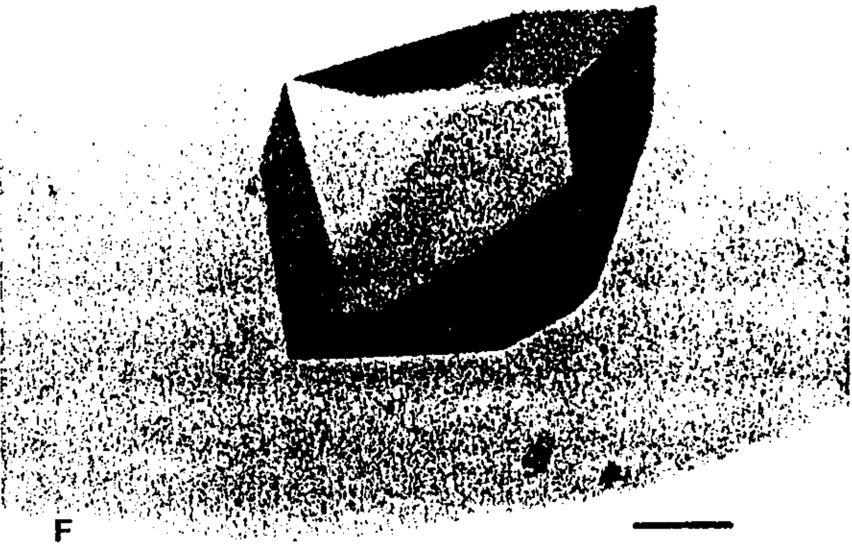




FIG. 4D and E. See legend on p. 8



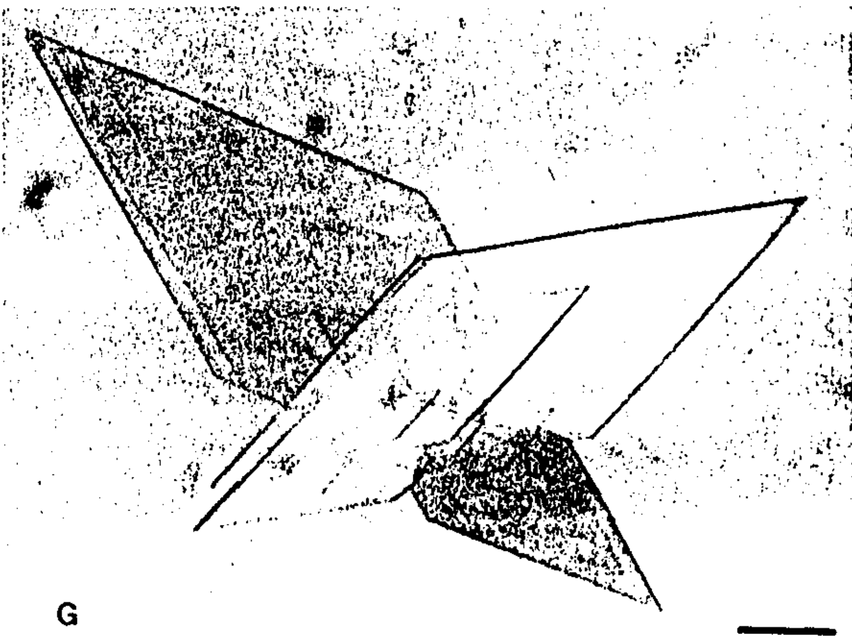




FIG. 4F and G. See legend on p. 8

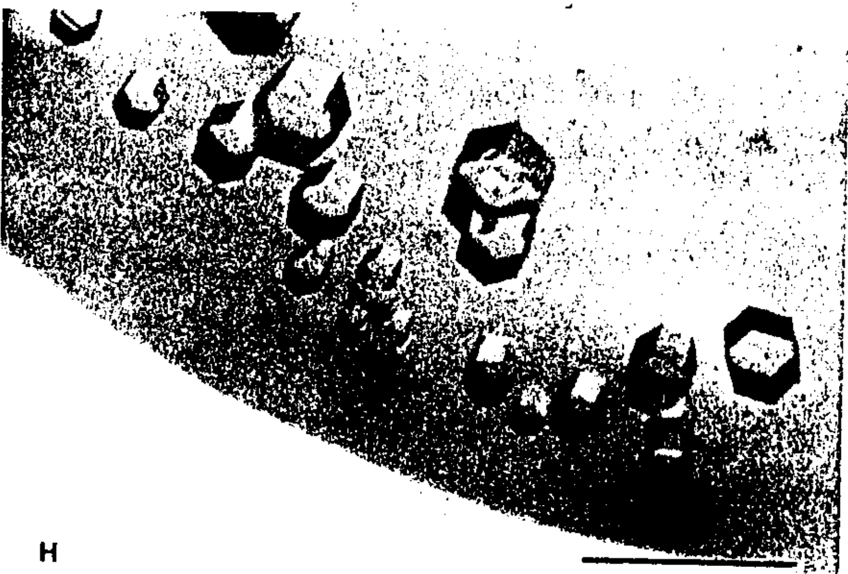



FIG. 4H. See legend on p. 8.

(Hofrichter *et al.*, 1976), again reflecting the dependence of the nucleation rate on the frequency of molecular encounters.

### D. *Experimental Determination of Nucleation Conditions*

Experimental methods for the early detection of nuclei formation using solution light-scattering measurements were described by Kam *et al.* (1978). A principal difficulty is to distinguish between the formation of three-dimensional nuclei and amorphous aggregates at early stages of protein association. Kam *et al.* (1978) developed a two-parameter model to discriminate between these processes, based on the idea that the number and distribution of intermolecular contacts formed by nuclei and precipitates differ. Nuclei are compact, each molecule making several three-dimensional intermolecular contacts, while precipitates form more extended chain networks that are larger and less dense. This variation in size and density gives rise to different signals in dynamic light-scattering measurements as a function of protein concentration and allows discrimination between the two aggregation states prior to macroscopic crystal formation (Kam *et al.*, 1978; Feher and Kam, 1985). This



approach has been used to study nuclei formation of lysozyme (Kam *et al.*, 1978; Baldwin *et al.*, 1986) and phosphoglucomutase (C. W. Carter *et al.*, 1988).

## V. CRYSTAL GROWTH MECHANISMS

### A. *Transport-Controlled Growth*

Crystal growth rates depend potentially on both the transport rates of solution molecules to the crystal surfaces and their rate of incorporation after they have arrived. Models of crystal growth (Fiddis *et al.*, 1979; Davey, 1986; Boistelle and Astier, 1988) have been developed that distinguish between transport and surface-ordering events as factors that control growth rates. In the transport-limited growth model, growth rates reflect the frequency with which molecules reach the crystal surface. Although many growth experiments show evidence for a depletion region around growing crystals (Kam *et al.*, 1978; Pusey *et al.*, 1988), as described in Section V.D, most studies suggest that surface effects are rate limiting in protein crystal growth.

### B. *Growth Controlled by Surface Kinetics*

Experimental studies of lysozyme (Fiddis *et al.*, 1979; Pusey *et al.*, 1986), insulin (Schlichtkrull, 1957; Fiddis *et al.*, 1979), and canavalin (DeMattei and Feigelson, 1989) suggest that events occurring at the lattice surface are rate limiting in protein crystal growth. Important effects occurring at the lattice surface include the formation of favorable growth sites (usually some form of irregularity) and molecular attachment to these sites. For rough crystal surfaces, where many growth sites exist or the energy to create them is low, growth proceeds at relatively low values of supersaturation. Growth on smooth surfaces, where the energy to create a growth site is high and may necessitate the formation of surface nuclei, requires relatively higher supersaturation levels. In more detailed terms the growth models include

1. A rough surface model in which molecular incorporation is favored at many vacant sites in the nascent lattice layer. In this case the growth rate is

$$G = k_1(c - c_S)$$

where $k_1$ depends on the surface binding energy and the mass transfer rate to the crystal surface, $c$ is the protein solution concentration, and $c_S$ is the protein solubility.

2. A screw dislocation model that predicts preferred growth along a defined dislocation. Here,

$$G = k_2 c(\ln c/c_S)^2$$

where $k_2$ similarly depends on characteristics of the growing crystal face.

3. A surface nucleation model in which attachment sites exist as molecular clusters that, like nuclei, must reach a critical size to be stable and support subsequent crystal growth. Here,

$$G = k_3 c^{1/3}(\ln c/c_S)^{5/6} \exp[-\pi\gamma^2/3(k_B T)^2 \ln c/c_S]$$

where $k_3$ is a function of the crystal face, $\gamma$ is the excess free energy of a molecule with unsatisfied lattice interactions at the edge of a growth site, $k_B$ is Boltzmann's constant, and $T$ is the temperature.

Because these models predict various dependencies of the crystal growth rate on solution supersaturation, growth mechanisms can be distinguished experimentally (Schlichtkrull, 1957; Fiddis *et al.*, 1979; Durbin and Feher, 1986; Pusey *et al.*, 1986; DeMattei and Feigelson, 1989). Most of these studies suggest that surface nucleation and screw dislocation models most accurately describe protein crystal growth kinetics, although a detailed study of hen egg white lysozyme (Durbin and Feher, 1986) shows different mechanisms at low and high protein supersaturation levels. At low supersaturation different growth rates were observed on equivalent faces of tetragonal lysozyme crystals. This result suggested that preferred growth occurs at a very small number of local surface defects, presumably introduced at random to account for unequal growth rates of equivalent crystal faces. At higher supersaturation equivalent faces of lysozyme crystals grow to similar size. In this case the growth rate dependence on supersaturation follows the two-dimensional nucleation model.

## C. *Measurements of Crystal Growth Rates*

Growth rates on the order of $10^{-8}$ cm/sec have been measured for several protein crystals (Fiddis *et al.*, 1979; Pusey *et al.*, 1986; DeMattei and Feigelson, 1989). In general, crystals grow faster at increasing levels of supersaturation, and except for small crystals (i.e., $<10\ \mu m$), growth rates appear to be independent of size (Schlichtkrull, 1957). In a comparison of protein and small molecule crystal growth rates from solution, [illegible]



Feigelson (1988) found that their surface linear growth rates are comparable. When the number of molecules added per unit time is considered, protein crystals do grow more slowly. However, because protein molecules are much larger, fewer are needed to achieve the linear growth rates observed for small molecules.

## D. *Transport Phenomena in Protein Crystal Growth*

Sustained crystal growth requires that solute molecules continually reach the crystal surface. Transport of molecules can occur by both diffusion and convection. Because the solution near the growing crystal surface is depleted of solute as the crystal grows (Fig. 5), gravity can act on the density difference between the solute-depleted layer and the bulk solvent to produce convection currents. Although many studies of crystal growth describe the transport phenomenon responsible for mass transfer to growing crystals as "diffusion," theoretical arguments suggest that buoyant or, in small volumes, surface tension-driven convection actually dominates simple Fick's law diffusion in determining the

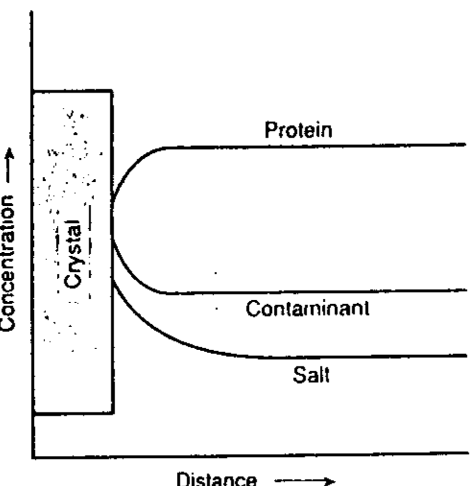



FIG. 5. Protein, protein contaminant, and salt concentrations near the surface of a growing crystal. As molecules add to the crystal, the solution near the crystal is depleted of protein. The exclusion of protein contaminants and salts increases their effective concentration near the crystal surface. The relative shapes of the concentration curves depend on the molecular diffusivity, with more rapidly moving molecules such as salts having wider concentration gradients. The concentration profiles shown are expected to occur in the absence of convection. Convection currents caused by the differences in solution density [illegible]

rates of mass transfer to growing crystals (Rosenberger, 1986). Indeed, buoyancy-induced convection currents around growing protein crystals have been directly observed using schlieren optics that allow visualization of variations in the solution-refractive index (Pusey *et al.*, 1988). Pusey *et al.* (1986) suggest from their data on lysozyme that convection due to density gradients near the growing crystal is sufficient to prevent diffusion-limited crystal growth. Flow of solute also influences protein crystal morphology. Preliminary experiments show that relative growth rates of human serum albumin crystal faces depend on crystal orientation in a flowing solution (Broom *et al.*, 1988). Based on the observed changes in relative sizes of crystal faces, these authors speculate that unfavorable habits such as needles could be improved by oriented seeding.

### *E. Role of Molecular Preassociation in Nucleation and Crystal Growth*

Models for protein crystal formation follow those for small molecules and assume that crystal nuclei form and grow by the association or addition of solute monomers. Nevertheless, several lines of evidence suggest that aggregates may participate in some aspects of protein crystal nucleation and growth. It is a common experience that the onset of nucleation and crystal growth are delayed for long periods of time after suitable supersaturating conditions exist. Kam *et al.* (1978) suggest that this preequilibrium state is characterized by the formation of various molecular aggregates prior to the eventual formation of stable nuclei. In fact, solution studies by Banerjee *et al.* (1975) show that lysozyme self-associates into indefinite head-to-tail polymers under conditions similar to those used for crystallization experiments. Analysis of molecular interactions in several lysozyme crystal forms (Salemme *et al.*, 1988) showed that the polymorphs could all be assembled from a common subset of linear molecular chains (Fig. 6). Since the crystals and polymers form under similar conditions (Banerjee *et al.*, 1975), it is possible that chains observed in solution represent preaggregates that associate to form crystal nuclei. Such "sequential" mechanisms might provide easy formation routes for protein crystal nuclei where the molecules are only tenuously connected in the three-dimensional lattice.

Whether crystals can grow by the addition of molecular aggregates (as opposed to single molecules) to the crystal faces is less clear. The observation that the predominant growth mechanisms for many crystals involve two-dimensional surface nucleation (Schlichtkrull, 1957; Fiddis *et al.*, 1979; Durbin and Feher, 1986; Pusey *et al.*, 1986) could be the result of molecular chain association to an otherwise smooth and com-



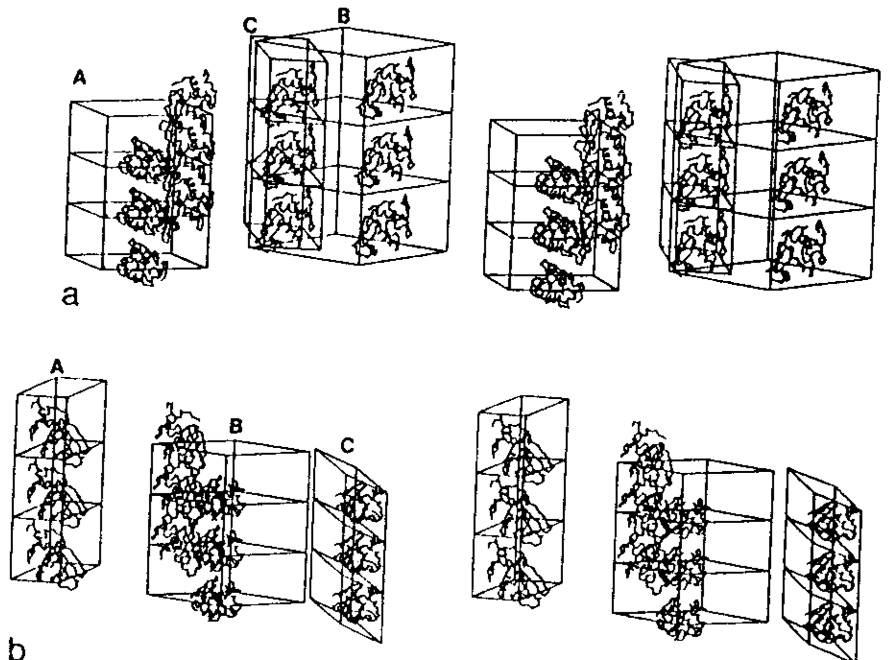



FIG. 6. Stereoscopic views of common molecular chains in three crystal forms of hen egg white lysozyme. (a) Views of monoclinic (A), tetragonal (B), and triclinic (C) cells, illustrating the recurrence of the chain corresponding to the triclinic *c* axis. (b) The triclinic *c*-axis chain (A) aligned with one subunit of the dimer in the asymmetric unit of the monoclinic cell (B). (C) The triclinic *a*-axis array oriented with the other subunit of the monoclinic asymmetric unit. From Salemme *et al.* (1988).

pleted lattice plane. In this case growth rates could depend on the polymeric species distribution in complicated ways. For lysozyme the heat of formation for the self-associated chains (Banerjee *et al.*, 1975) is −6.4 kcal/mol, a value comparable to measured heats of crystallization for tetragonal lysozyme [−17.2 kcal/mol (Ataka and Asai, 1988), and −25.1 kcal/mol (Takizawa and Hayashi, 1976)], although these apparently vary somewhat, depending on solution conditions.

## VI. Competition between Nucleation and Growth

Rates of crystal nucleation and growth generally have different dependencies on protein supersaturation, and additionally vary substantially for different protein–precipitant systems. This can lead to a variety of unexpected behaviors in crystallization experiments.

rate is typically an exponential function of supersaturation, ranging from values of 3–4 for lysozyme (Ataka and Tanaka, 1986) to 35 for hemoglobin S (Hofrichter *et al.*, 1976). Growth rates, in contrast, can have a variety of functional forms, although supersaturation levels that support nucleation also generally promote rapid growth.

The difference in the dependence of nucleation and growth rates on supersaturation has important implications for experiments in which the desired result is the formation of a few large crystals. For example, in systems in which the nucleation rate depends on a relatively low power of supersaturation, narrow ranges of supersaturation exist that favor growth over nucleation (Ataka and Tanaka, 1986). For these systems nucleation is likely to dominate at low and high supersaturations, while crystal growth is favored at intermediate supersaturation. This model is supported by data from several systematic studies of crystal size and number as a function of supersaturation. For rabbit muscle aldolase (Heidner, 1978) and hen egg white lysozyme (Ataka and Tanaka, 1986), more crystals formed at high and low supersaturation and fewer were observed at intermediate supersaturation levels. In studies of crystal size as a function of protein concentration (Heidner, 1978; Ataka and Tanaka, 1986; Betts *et al.*, 1989), final crystal size depended critically on the supersaturation ratio, the ratio of initial protein concentration to protein solubility. The linear dimensions of lysozyme crystals, for example, could be increased from 0.6 mm to 1.0 mm, using solutions in which the supersaturation ratio varied from 2.5 to 4 (Ataka and Tanaka, 1986) (Fig. 7). Betts *et al.* (1989) also obtained large crystals (1.0 × 0.6

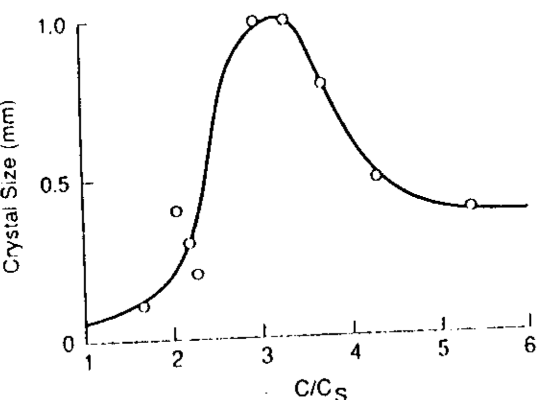



FIG. 7. The relationship between final size of hen egg white lysozyme crystals and the degree of supersaturation. The largest crystals grew when the ratio of initial lysozyme concentration ($c$) to its solubility ($c_s$) ranged between 2.5 and 4.0. From Ataka and Tanaka (1986).



× 0.4 mm) of cytidine deaminase over a narrow range of protein concentrations. At protein concentrations above and below the optimum, only precipitate formed or else the solution remained clear.

## VII. Cessation of Growth and Crystal Disorder

Perhaps the most practical, yet mysterious, aspect of protein crystallization concerns the causes of growth cessation. It has frequently been observed that some crystals do not grow beyond a certain size, even in the presence of excess protein. There are several possible reasons for this effect. Probably the most common, in view of the complexity and chemical sensitivity of protein molecules, is the gradual poisoning of the crystal surfaces by defective molecules that themselves attach to the crystal lattice, but do not support subsequent growth. If the defective molecules bind more weakly than native molecules, then they tend to be excluded from the lattice until the native molecules are nearly depleted from the solution (Fig. 5). At this stage, owing to the higher solution concentration of the defective molecules, their addition to the lattice dominates and the surfaces become poisoned toward further growth. For example, crystallization of hemoglobin C is inhibited by the addition of hemoglobins A and F (Hirsch *et al.*, 1988).

Although it is apparent that progressive poisoning of crystals is possible (and probably underlies the strain–defect cessation model described below), crystal growth can occasionally be reinitiated by changes in the surrounding protein solution (Young *et al.*, 1988), suggesting the localization of the poisoning molecules at the crystal surface. Similarly, macroseeding experiments (Section VIII,G) are usually initiated with an etching step that presumably removes defective surface molecules that would otherwise poison the crystal surfaces toward further growth.

An alternative mechanism that can lead to growth cessation is the introduction of crystal strain or cumulative defects into the lattice as the crystal grows. The concept was physically realized by Kam *et al.* (1978), who halved a lysozyme crystal and found that each half then grew to the size of the original. It was suggested that this reflected the continuous accretion of defects while the crystal grew, so that finally the addition of new molecules to the (defective) surface lattice became unfavorable. Although the incorporated defects could be either molecular or structural (resulting, say, from too rapid growth), the observation that the crystal halves regrew to the original size suggests that the defects were structural, since they were propagated into the newly reformed lattices. Structural defects might be expected to be more common in rapidly grown crystals, as suggested by many experiments showing that growth

lar conditions, larger crystals of hen egg white lysozyme were grown in hanging-drop experiments when supersaturation was achieved slowly at controlled rates (Gernert *et al.*, 1988).

Although the effects of incorporating defective molecules into the growing crystal lattice have been introduced in the context of growth cessation, it is clear that situations can exist in which the crystal can continue to grow macroscopically, even though it incorporates defects that destroy the long-range molecular order in the lattice. The disappointing result is a crystal that typically looks good, may be strongly birefringent, but does not usefully diffract X-rays. The variety of diffraction effects observed—ranging from no diffraction, through resolution-limited or anisotropic diffraction patterns, to disordered patterns in which still exposures look like precession photographs—illustrates the many types and spatial scales of lattice disorder that can occur (Harburn *et al.*, 1975).

## VIII. Crystallization Methods

The commonly used protein crystallization methods achieve and maintain supersaturation in several ways (Fig. 2). Several articles and books that describe methods used to grow protein crystals are referred to in Section I. The objective in this section is to briefly review methods as they relate to the phenomena described above. Examples of proteins crystallized by each method are given. More complete listings of crystallized proteins are compiled in McPherson, 1982 and Gilliland and Bickham, 1990.

### A. *Batch Method*

The simplest technique used to grow protein crystals is the batch method in which the protein is mixed with salts or other precipitants to achieve supersaturation (Fig. 2), and the vessel is sealed and set aside until crystals appear. Frequently, the supersaturation point required to induce nucleation is empirically determined by observing the onset of transient turbidity as powdered salt is progressively added to the solution. Crystals of hen egg white lysozyme used for most systematic studies of protein crystallization are grown by batch methods (Blundell and Johnson, 1976). Mouse pancreatic ribonuclease (Perry and Palmer, 1988) and the biotin operon repressor (Brennan *et al.*, 1989) represent recent examples of use of the batch method.

### B. *Dialysis*

In the dialysis method protein solution is retained by a dialysis membrane which maintains the solution at a constant concentration while al-



lowing equilibration with a surrounding solution. Although the method can be used for any of the usual solvent perturbation approaches that involve added salts or small organic molecules to reduce protein solubility, the method is uniquely suited to the formation of crystals that are induced to crystallize at low ionic strengths. Occasionally, nucleation may be transiently induced at the dialysis solution boundary, where the membrane serves as a site for heterogeneous nucleation. Examples of the use of this method include the crystallizations of hexokinase (Steitz, 1971) and anthranilate phosphoribosyltransferase (Edwards *et al.*, 1988).

### C. *Vapor Diffusion*

Vapor diffusion methods are among those most commonly used for protein crystallization because they readily lend themselves to the use of 5- to 50-$\mu l$ solution volumes. Typically, a hanging or sitting drop, containing a solution of protein plus precipitant at subsaturating concentrations of protein, is equilibrated against a larger reservoir of solution containing precipitant or another dehydrating agent. After sealing in a closed vessel, the solutions equilibrate to achieve supersaturating concentrations of protein and thereby induce crystallization in the drop.

Theoretical models of hanging-drop experiments suggest that vapor equilibration at the droplet surface is sufficiently rapid to produce transient concentration gradients in the droplet (see, e.g., Yonath *et al.*, 1982) (Fig. 4B) that might induce homogeneous nucleation (Fehribach and Rosenberger, 1989). Experimental measurements indicate final water equilibration times of 36–80 hr that depend on drop size and geometry and on reservoir precipitant concentrations (Fowlis *et al.*, 1988; Mikol *et al.*, 1989). When ammonium sulfate solutions are used in the reservoir, protein solution pH also rapidly equilibrates, owing to the low vapor pressure of ammonia (Mikol *et al.*, 1989). Many additional variations are possible, particularly in systems including organic solvents in which diffusion both to and from the drop and the reservoir becomes a possibility. Controlled pH changes can be made by vapor equilibration of volatile organic acids or bases (McPherson and Spencer, 1975).

### D. *Temperature Shift*

Proteins that are near their supersaturation points in concentrated salt solutions can frequently be induced to crystallize by changing the temperature. This phenomenon, which has been used for the fractional purification of proteins (Jacoby, 1968), has found only scattered application in protein crystallization. Nevertheless, given the relative ease of precise temperature regulation, methods based on temperature alteration de-

serve more thorough investigation (Rosenberger and Meehan, 1988). Clear applications include the introduction of temperature shifts to achieve transient nucleating conditions or to reduce the number of nuclei formed on initial solution supersaturation. Examples of proteins crystallized using temperature shift methodology include bovine neurophysin II (Rose *et al.*, 1988) and elongation factor Tu (Lippmann *et al.*, 1988).

### *E. Achieving Different Conditions for Nucleation and Growth*

The objective of most protein crystallization experiments is to obtain a few large crystals. As outlined in Sections IV,C and VI, two of the major obstacles to controlled protein crystal growth are the extreme sensitivity of nucleation rate to supersaturation conditions and the necessity for higher supersaturations to promote nucleation than are needed for growth (Fig. 2). An inherent shortcoming of many crystallization methods is that they depend on similar conditions both to promote nucleation and to support growth. A frequent result is either no crystals or the formation of many small crystals. However, alternative approaches have been developed that attempt to individually optimize nucleation and growth conditions.

### *F. Free Interface Diffusion*

Crystallization using free interface diffusion represents an attempt to achieve transient highly supersaturating conditions required for nucleation, followed by relaxation to conditions of lower supersaturation required for growth, within a single experimental setup (Salemme, 1972). In this method, a protein solution is layered over a precipitant solution. Initially, molecules at the liquid–liquid interface achieve high supersaturation, while the remainder maintain bulk conditions of the protein layer. The high supersaturating conditions at the interface promote nuclei formation. As the liquids diffuse, the high protein supersaturation initially achieved at the interface decreases. At equilibrium, when the precipitant and protein solutions are mixed completely, the entire protein solution is supersaturated (Fig. 2). Ideally, smaller nuclei dissolve at the lower levels of protein supersaturation and only the larger nuclei continue to grow. Cytochrome *c*′ (Weber and Salemme, 1977), phospholipase $A_2$ (Dijkstra *et al.*, 1978), and adenylate kinase (Althoff *et al.*, 1988) are among the proteins that have been crystallized by free interface diffusion.



### *G. Seeding*

Seeding is a method that physically separates the processes of crystal nucleation and growth, so that conditions for crystal growth can be independently optimized. Seeding from crushed crystals can be done by using a fine glass rod to transfer nuclei from a stabilizing solution to a growth-promoting solution. An alternate microseeding method involves first passing a hair through a solution containing crystals or particles too small to be definitively identified as crystals, and then streaking the hair through the protein solution to be seeded (Stura and Wilson, 1990; Leung *et al.*, 1989). Systematic methods of microseeding have been investigated (Fitzgerald and Madsen, 1986), as it is frequently difficult to control the number of seeds transferred via the glass rod or by streaking.

Individual crystals having dimensions as small as 0.01 mm can be grown to larger sizes by macroseeding (Thaller *et al.*, 1981). In this procedure a single crystal is repeatedly transferred to a fresh protein solution after crystal growth ceases. Before transfer to the new protein solution, the crystal is washed and its surface is etched by partial dissolution in a solution of low supersaturation. The need for etching suggests that crystal terminal size is caused by poisoning the crystal surface with impurities.

## IX. Protein Purity

It has long been recognized that protein purity plays a critical role in crystallization. Many investigators assert that if a protein fails to crystallize, or crystallization is irreproducible, the protein sample is simply lacking in sufficient purity (Anderson *et al.*, 1988; Giege *et al.*, 1988). As outlined in Section VII, impurities structurally similar to the solute are most likely to poison crystal growth or otherwise disrupt crystalline order. Both the availability of newer separation methods and the necessity to improve crystal quality in order to obtain key structural information have motivated detailed studies of how protein heterogeneity affects crystal growth.

Protein contaminants can occur as natural isoforms or can arise during purification. Adventitious proteolysis and cysteine oxidation are probably the most common sources of microheterogeneity that occur during isolation (Lorber *et al.*, 1987). This has frequently motivated the inclusion of protease inhibitors and/or reducing agents in crystallization solutions, as well as during purification. In many cases modifications that produce molecular heterogeneity are reflected in enzyme activity. For

example, it was noted that only the most active preparations of ribosome subunits would form useful crystals (Yonath *et al.*, 1982).

In many cases proteins which are otherwise intractable can be crystallized in fragments or as truncated forms. However, the chemical or enzymatic methods to cleave the molecules are a frequent source of product heterogeneity. For example, Fab fragments are liberated from intact immunoglobin molecules by endoproteolytic cleavage. The protease treatment often produces molecules having variability in the location of the cleavage site and extraneous nicks elsewhere in the molecule. The resultant Fabs are similar in structure, but differ in their isoelectric points. Purification of a single Fab isozyme by ion-exchange chromatography (Cygler *et al.*, 1987; Boodhoo *et al.*, 1988; Orbell *et al.*, 1988), isoelectric focusing (Bott *et al.*, 1982), or chromatofocusing (Prasad *et al.*, 1988), followed by crystallization of the isoelectrically pure Fab fragment, has been shown to dramatically improve crystal quality. Similar results have also been reported as a result of separating isozymes of other biological macromolecules (Bott *et al.*, 1982; Spangler and Westbrook, 1989).

The complex consequences of cysteine oxidation were thoroughly detailed in a study of 4-hydroxybenzoate 3-monooxygenase prompted by an initial failure to reproduce the original crystal form (Van Der Laan *et al.*, 1989). This work showed that enzyme crystallization was sensitive to the oxidation state of a single cysteine residue. Cysteine-116, located at the molecular surface, is also situated at a lattice contact near Cys-116 from an adjacent molecule. Consequently, crystallization of an enzyme having the cysteine oxidized to sulfonic acid is inhibited by the requirement to juxtapose two large negatively charged groups at a lattice contact occupied by reduced cysteines in the native crystals. It was additionally noted that only slight molecular movements could result in intermolecular disulfide formation. However, 4-hydroxybenzoate 3-monooxygenase preparations contaminated with covalently linked aggregates form poorly ordered crystals, indicating the disruptive effect of this small contact modification on crystal lattice ordering.

Although protein microheterogeneity usually disrupts crystal formation, it can occasionally promote crystallization. For example, crystals of *Escherichia coli* single-stranded binding protein contain a 1:1 mixture of intact and proteolyzed protein (Ollis *et al.*, 1983). While crystallization experiments were initially conducted with intact protein, crystals grew only when enough molecules to form the mixed crystals had been cleaved by contaminating proteases.

Recombinant DNA technology has had an enormous impact on crystallography because it has made naturally scarce proteins plentiful. Moreover, it has largely eliminated proteolysis as a source of microheterogeneity in protein fragments, since these can now be produced genetically. A recent example of crystals grown from a genetically engineered fragment is the Klenow fragment of DNA polymerase (Brick *et al.*, 1983). Nevertheless, biotechnological production methods can also introduce contaminants, particularly when the host organism expresses a protein similar to the foreign gene being overexpressed. In the expression of *Bacillus stearothermophilus* tryptophanyl-tRNA synthetase (trpS; tryptophan–tRNA ligase) in *E. coli*, low levels of *E. coli* trpS enzyme copurified with the cloned *trpS* and inhibited crystal growth. Only when the *E. coli* gene was deleted did the overexpressed *B. stearothermophilus* enzyme crystallize successfully (Carter, 1988). Many of the naturally scarce proteins that are typically overexpressed in *E. coli* functionally bind nucleic acids. Microheterogeneity often arises from nonspecific binding of host cell DNA or RNA, necessitating special isolation procedures to produce uncontaminated protein (Ruff *et al.*, 1988).



## X. Searching for Crystallization Conditions

Although an extensive and expanding database of protein crystallization conditions reveals some trends in the uses of techniques and precipitants (Gilliland and Davies, 1984; Gilliland, 1988; Gilliland and Bickham, 1990), it is not yet possible to predict the conditions required to crystallize a protein from its other physical properties. Lacking any predictive scheme, the crystallization of a new protein is usually attacked using a more or less ad hoc approach based on the previous experiences of an investigator. Most often, crystallization attempts must be made with limited amounts of material, leaving the experimenter with the problem of searching a potentially large parameter space with a limited number of experiments. Despite, or perhaps owing to, these limitations, there have been two systematic approaches to searching for protein crystallization conditions.

The first method to be described was an incomplete factorial approach (Carter and Carter, 1979). This is basically a method that, given a matrix of compositional components and their concentrations, defines how to sample the variables with a minimum number of experiments. Using statistical methods to analyze the results, it is possible to identify variables that are correlated, and in later stages to concentrate on their variation to optimize crystallization conditions. This method has been

implemented using microdialysis cells (Blundell and Johnson, 1976) and used to optimize crystal growth conditions for tryptophan–tRNA ligase (Carter and Carter, 1979) and cytidine deaminase (Betts *et al.*, 1989).

An alternative approach, suited particularly to using laboratory robotics, is the successive automated grid search (SAGS) method (Cox and Weber, 1988). The method has been implemented with the hanging-drop crystallization method (McPherson, 1982) and involves the systematic variation of two major crystallization parameters, pH and precipitant concentration, with provisions to vary two others. The variation of solution pH and precipitant concentration effectively varies molecular charge as a function of protein supersaturation in searching for suitable crystallization conditions. The coarse grid is initially used for sampling. Once initial crystals are obtained, the increments of the grid are reduced in the vicinity of the initial successful experiments to optimize crystallization conditions (Fig. 8). The method has been successfully used to crystallize several proteins (Cox and Weber, 1987; Weber *et al.*, 1987).

Irrespective of whether crystallization conditions are found by systematic or trial-and-error methods, the process first involves locating some point in the parameter space of possible conditions where the protein will crystallize. Many experiments suggest that the range of parameters over which a given protein will form some sort of crystal is reasonably large (Cox and Weber, 1988; Weber, 1990). In contrast, the parameter space defined by the optimal conditions where crystals suitable for X-ray diffraction studies are grown can be much smaller. Studies of crystal terminal size and the dependence of crystal size on the supersaturation ratio demonstrate that the largest crystals grow within a narrow range of conditions (Heidner, 1978; Ataka and Tanaka, 1986; Betts *et al.*, 1989) (Fig. 7). A typical result is that observed with haloalkane dehalogenase (Rozeboom *et al.*, 1988): Crystals form in ammonium sulfate concentrations greater than 60% saturation over the pH range 5.4–7.2. However, the best crystals form under the more restrictive conditions of 64% saturation and pH 6.3 ± 0.1.

## XI. New Developments in Protein Crystallization

### A. *Crystallization in Microgravity*

Space exploration offers a unique opportunity to test the effects of gravity on protein crystal growth (Morita, 1985; Bugg, 1986; DeLucas *et al.*, 1986; Drenth *et al.*, 1987). Although it may not be obvious, the growth of millimeter-sized protein crystals in microliter volumes is af-

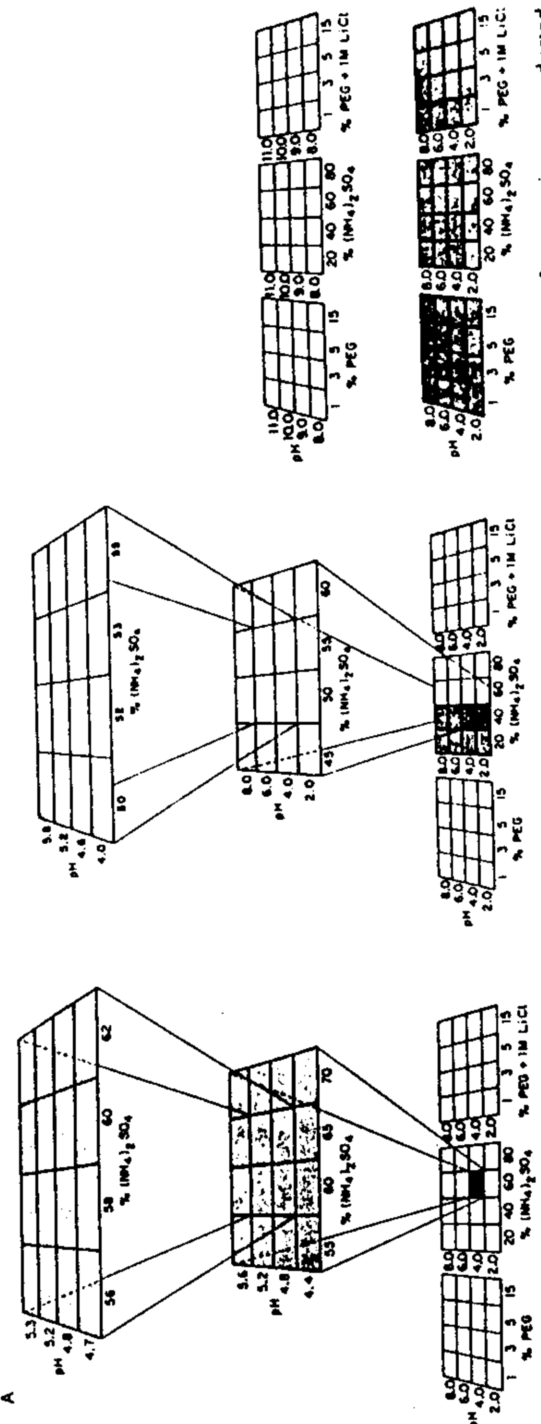



Fig. 8. Approach to protein crystallization using successive automated grid searches. (A) The bottom rows show the first experiments conducted and typical results using this approach. Three trays each containing 16 droplets are set up. Using the citrate/phosphate buffer system, the pH of the solution is varied from 2.4 to 7.8. A different type of precipitating agent is used in each tray. These are polyethylene glycol (PEG) from a stock solution of 25% (w/v) PEG 8000, ammonium sulfate from a saturated stock solution, and a PEG–salt mixture using the 25% PEG solution above and adding 1 *M* LiCl to all droplets by dilution from a 10 *M* LiCl stock solution. (Left) The shaded square indicates a crystal-containing droplet. The top two figures in this panel show a typical experimental strategy of expanding and overlapping conditions about those yielding crystals in the first wide-screen experiments. (Middle) Occasionally, the distinction between conditions that do and do not cause precipitation is clear, as indicated by shading. In this case successive grids explore the region along the precipitation border. (Right) If precipitation occurs in all droplets, the three broad-screen experiments are repeated using quite different conditions (e.g., higher pH, an alternate temperature or protein concentration, or the addition of cofactors). (B) Successive automated grid searches were applied to the crystallization of a streptavidin–biotin complex. (Top) The initial wide-screen experiment in which the pH of the solution in the columns varied from 2.4 to 7.8 with the citrate/phosphate buffer system and the polyethylene glycol (PEG) concentration ranged from 1% to 15% in the row. Droplets were photographed about 2 weeks after beginning the crystallization experiment. Crystals grew overnight at pH 7.8, 15% PEG. These crystals turned brown within 1 week. Crystals that grew more slowly [illegible] (Bottom) An expansion of crystallization conditions about those producing crystals in

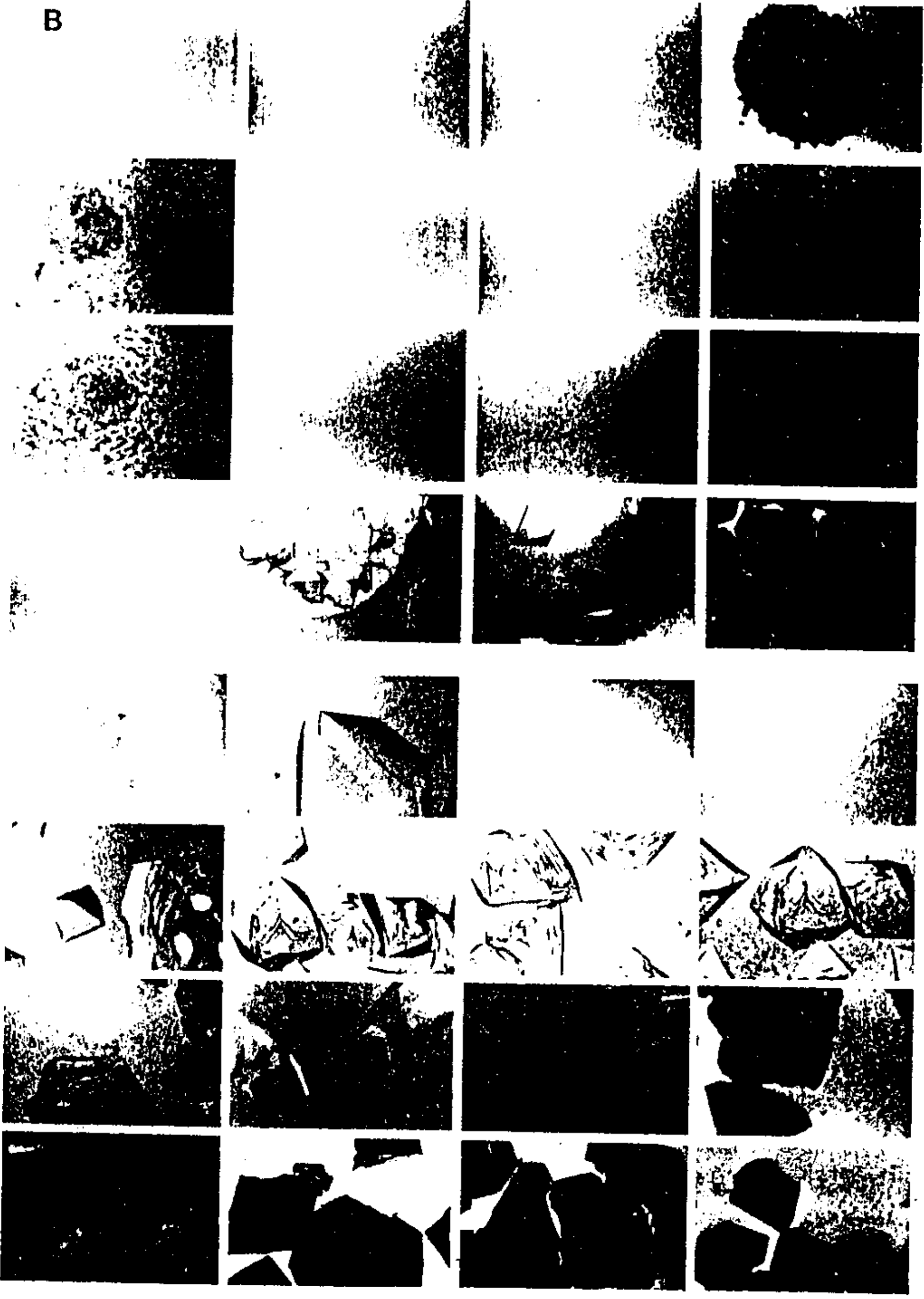



FIG. 8B. See legend on p. 27.



fected by gravity in several ways. Protein crystals are more dense than the bulk solution, so that an immediate advantage to crystallization in microgravity is the elimination of crystal sedimentation. Instead of falling to the bottom of the crystallization solution, where fused aggregates can form, crystals grow suspended in solution at the site of nucleation. As a result larger single crystals with more perfect habits are frequently observed in microgravity crystallizations (DeLucas *et al.*, 1986).

Density gradients are established at several stages in the crystallization process (Fig. 5). As molecules attach to the growing crystal surface, the solution near the crystal is depleted of solute and becomes less dense than the bulk solution. Under the influence of gravity, such density differences result in convection currents. However, in microgravity, solutions with different densities are not subject to convection, so that solutions mix with less turbulence (Littke and John, 1984) and equilibration between solutions is much slower (DeLucas *et al.*, 1986).

Solution turbulence could affect several stages of protein crystal growth. Littke and John (1986) suggested that the rapid onset of $\beta$-galactosidase crystallization observed on Spacelab 1 was attributable to a lack of convective mixing turbulence that otherwise disrupted prenuclear molecular complexes in terrestrial control experiments. In terrestrial control experiments isocitrate lyase crystallizes with a dendritic habit (Fig. 4C) that results from local concentration fluctuations at the growing crystal surface, which cause unusually rapid growth along principal crystal axes (Langer, 1989). A more regular habit was obtained in microgravity (Fig. 4D), consistent with the expected elimination of convection-induced concentration fluctuations at the crystal surfaces. Growth in the quiescent microgravity environment may also improve the internal order of protein crystals independent of increases in crystal size (DeLucas *et al.*, 1989). Taken together, the initial data on protein crystal growth suggest that the mechanisms for introduction of lattice defects frequently associated with turbulent growth affect both crystal order and terminal size.

Some of the advantages of crystallization in microgravity can be achieved by crystallization in gels (Robert and Lefaucheux, 1988). For example, crystals remain suspended in the gel, although the gel matrix is sufficiently flexible to allow crystal growth. Convection currents are attenuated so that nucleation is reduced and mass transfer occurs primarily by diffusion. Entrapment of foreign particles reduces heterogeneous nucleation, and initial studies suggest that homogeneous nucleation is restricted to the largest pores of the gel containing enough solute molecules to form a critical nucleus. Hen egg white lysozyme and por-

cine trypsin have been crystallized in 1% (v/v) tetramethoxysilane and 0.4% agar gels (Robert and Lefaucheux, 1988).

### *B. Automated Crystallization*

Laboratory automation has been adapted for protein crystallization. Setting up a crystallization experiment involves several liquid-handling operations—including dilution, mixing, and dispensing of multicomponent solutions—that are readily automated. Automation reduces manual labor and increases reproducibility by reducing errors and improving the accuracy of solution delivery. Hanging- and sitting-drop crystallizations have been automated to varying extents, ranging from liquid handling for setup (Kelders *et al.*, 1987; Morris *et al.*, 1989) to mixing reservoir solutions from concentrated stocks and combining them in the crystallization droplet (Cox and Weber, 1987) (Fig. 9) to attempts at total automation, in which robotics are additionally used to grease, flip, and seal coverslips on the individual vapor diffusion wells

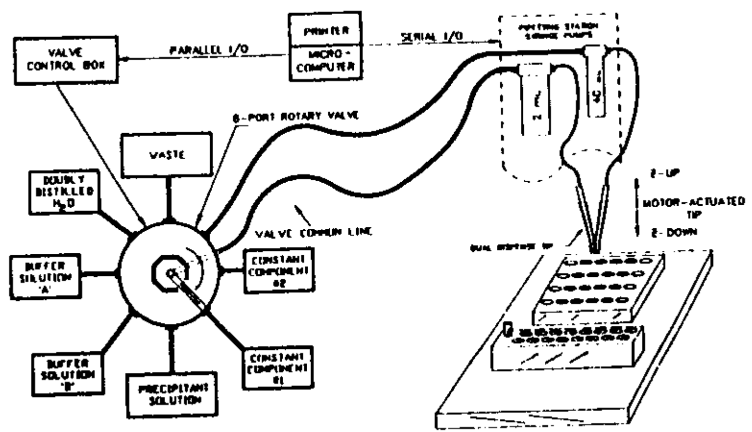



**Fig. 9.** An automated system for protein crystallization. This system, composed of a programmable pipetting station, a rotary valve, and a computer, is designed to manipulate liquids automatically for protein crystallization experiments by the hanging-drop technique. Stock solutions are placed at ports of the rotary valve shown at the left. Liquids are dispensed into the wells of the crystallization tray by the 2-ml syringe of the pipetting station. The dispense tip of the pipetting station moves vertically while the crystallization plate moves in the horizontal plane below it. Once the wells are filled, a droplet from each well is placed on the coverslip using the 40-μl transfer syringe (shown on the right). A vial of protein solution is placed on the coverslip holder (lower right) after all well solutions have been transferred to the coverslips. In the final step the protein solution is added to the droplets of well solution again using the 10-μl transfer syringe.



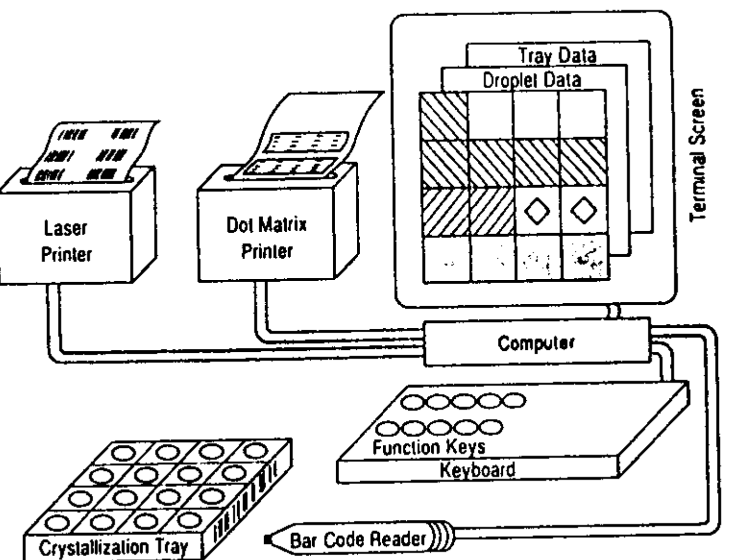



**Fig. 10.** Components of a data acquisition system for recording the results of crystallization experiments. Sixteen experiments are conducted in a 4 × 4 array on a tray. Each time the crystallization tray is examined visually under the microscope, the bar code label on the tray is first scanned with the bar code reader and a 4 × 4 grid appears on the computer terminal. A rating scheme of 10 descriptive comments is used to evaluate experimental results. Droplet ratings are entered into the computer using the 10 function keys, each corresponding to one of the ratings. When the observation for an individual droplet is entered, the corresponding grid position on the terminal display is filled by a color corresponding to the rating. The grid space is filled with a different color for each rating, rather than a number, to aid in the recognition of input data and to decrease data entry errors. Droplet data, including pH, the concentrations of precipitating agent and additives, and previous ratings, can be displayed for the entire grid by depressing a specified key. During the data entry procedure, the complete description of the tray is also accessible from the database. After the data for all 16 droplets have been entered, they are stored in the computer memory. A hard copy of the ratings is printed on a sticker that can be peeled from the backing and placed in the laboratory notebook. Results of other analyses of the database and the bar codes are printed on the laser printer.

(Ward *et al.*, 1988; Jones *et al.*, 1987). A crystallization plate, on which the protein solution is sandwiched between glass plates, was designed for the automated visual inspection of crystallization experiments (Jones *et al.*, 1987). Photographs of crystals first produced by an automated instrument together with some reproduced from the literature are shown in Fig. 4E–H. As a complement to the automated setup of crystallization experiments, a database system for recording crystallization results (Fig. 10) has been developed to facilitate data acquisition and to aid in the design of subsequent experiments.

## XII. Summary Remarks

Despite reports of nearly 100 new protein crystal forms in 1988, crystallization remains an important obstacle to many structural studies of biologically important molecules (DeLucas and Bugg, 1987; Giege *et al.*, 1988). Although much remains to be learned, research in this area reveals several recurrent themes, outlined here.

1. Proteins typically appear to be able to form crystals over a fairly large range of solution conditions. However, conditions required to produce large ordered crystals occupy a smaller fraction of the total crystallization parameter space. This observation is important in designing strategies to search for crystal growth conditions.

2. Stochastic events are important in protein crystallization. Solution fluctuations provide the driving force for homogeneous nucleation in supersaturated solutions. Effects are amplified in the small solution volumes used in vapor diffusion experiments, where high surface area–volume ratios produce transient concentration gradients (Fehribach and Rosenberger, 1989). Similarly, temperature and vibration are important factors in controlling crystal nucleation and growth (Feher and Kam, 1985).

3. Solution parameters change during crystal growth. The bulk protein concentration decreases as crystals grow, while the concentration of impurities increases. The growing crystals produce concentration gradients in solution. At the same time, electrostatic surface properties of the crystal can alter the activity of charged components in solution (Rosenberger, 1986; Young *et al.*, 1988).

4. Many crystallization reports emphasize the need to use pure proteins to ensure crystal reproducibility. The application of recombinant DNA technology to the production of truncated gene products promises to alleviate many of the difficulties associated with purifying protein fragments produced with proteolytic enzymes.

5. Changes in a single experimental parameter can simultaneously influence several aspects of a crystallization experiment. For example, temperature changes affect protein solubility, rates of nucleation and growth, and equilibration of the experimental apparatus. The interaction of parameters makes it difficult to design experiments to isolate individual effects and likewise complicates the interpretation of experimental results.

6. Many lines of evidence suggest that molecular preassociation may be important for protein nucleation and growth.



## Acknowledgments

The author thanks R. C. DeMattei, R. S. Feigelson, R. Giege, A. McPherson, and D. Ollis for providing manuscript preprints prior to publication and F. R. Salemme for critical discussions.

## References


Althoff, S., Zambrowicz, B., Liang, P., Glaser, M., and Phillips, G. N., Jr. (1988). *J. Mol. Biol.* **199,** 665–666.

Anderson, W. F., Boodhoo, A., and Mol, C. D. (1988). *J. Cryst. Growth* **90,** 153–159.

Arakawa, T., and Timasheff, S. N. (1984). *Biochemistry* **23,** 5912–5923.

Arakawa, T., and Timasheff, S. N. (1985). *Biochemistry* **24,** 6756–6762.

Ataka, M., and Asai, M. (1988). *J. Cryst. Growth* **90,** 86–93.

Ataka, M., and Tanaka, S. (1986). *Biopolymers* **25,** 337–350.

Baker, E. N. (1988). *J. Mol. Biol.* **203,** 1071–1095.

Baldwin, E. T., Crumley, K. V., and Carter, C. W., Jr. (1986). *Biophys. J.* **49,** 47–48.

Banerjee, S. K., Pogolotti, A., Jr., and Rupley, J. A. (1975). *J. Biol. Chem.* **250,** 8260–8266.

Betts, L., Frick, L., Wolfenden, R., and Carter, C. W., Jr. (1989). *J. Biol. Chem.* **264,** 6737–6740.

Blundell, T. L., and Johnson, L. N. (1976). "Protein Crystallography," pp. 59–81. Academic Press, New York.

Boistelle, R., and Astier, J. P. (1988). *J. Cryst. Growth* **90,** 14–30.

Boodhoo, A., Mol, C. D., Lee, J. S., and Anderson, W. F. (1988). *J. Biol. Chem.* **263,** 18578–18581.

Bott, R. R., Navia, M. A., and Smith, J. L. (1982). *J. Biol. Chem.* **257,** 9883–9886.

Brennan, R. G., Vasu, S., Matthews, B. W., and Otsuka, A. J. (1989). *J. Biol. Chem.* **264,** 5.

Brick, P., Ollis, D., and Steitz, T. A. (1983). *J. Mol. Biol.* **166,** 453–456.

Broom, M. B. H., Witherow, W. K., Snyder, R. S., and Carter, D. C. (1988). *J. Cryst. Growth* **90,** 130–135.

Bugg, C. E. (1986). *J. Cryst. Growth* **76,** 535–544.

Carter, C. W., Jr. (1988). *J. Cryst. Growth* **90,** 168–179.

Carter, C. W., Jr., ed. (1990). *Methods: Companion Meth. Enzymol.* 1.

Carter, C. W., Jr., and Carter, C. W. (1979). *J. Biol. Chem.* **254,** 12219–12223.

Carter, C. W., Jr., Baldwin, E. T., and Frick, L. (1988). *J. Cryst. Growth* **90,** 60–73.

Carter, D. B., Curry, K. A., Tomich, C.-S. C., Yem, A. W., Deibel, M. R., Tracey, D. E., Paslay, J. W., Carter, J. B., Theriault, N. Y., Harris, P. K. W., Reardon, I. M., Zurcher-Neely, H. A., Heinrikson, R. L., Clancy, L. L., Muchmore, S. W., Watenpaugh, K. D., and Einspahr, H. M. (1988). *Proteins: Struct., Funct. Genet.* **3,** 121–129.

Caspar, D. L. D., Clarage, J., Salunke, D. M., and Clarage, M. (1988). *Nature (London)* **332,** 659–662.

Cox, M. J., and Weber, P. C. (1987). *J. Appl. Crystallogr.* **20,** 366–373.

Cox, M. J., and Weber, P. C. (1988). *J. Cryst. Growth* **90,** 318–324.

Cygler, M., Boodhoo, A., Lee, J. S., and Anderson, W. F. (1987). *J. Biol. Chem.* **262,** 643–648.

Davey, R. J. (1986). *J. Cryst. Growth* **76,** 637–644.

DeLucas, L. J., and Bugg, C. E. (1987). *TIBTECH* **5,** 188–193.

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., and Bugg, C. E. (1986). *J. Cryst. Growth* **76**, 681–693.
DeLucas, L. J., Smith, C. D., Smith, H. W., Senadhi, V., Sendhi, S. E., Ealick, S. E., Carter, D. C., Snyder, R. S., Weber, P. C., Salemme, F. R., Ohlendorf, D. H., Einspahr, H. M., Clancy, L. L., Navia, M. A., McKeever, B. M., Nagabhushan, T. L., Nelson, G., McPherson, A., Koszelak, S., Taylor, G., Stammers, D., Powell, K., Darby, G., and Bugg, C. E. (1989). *Science* **246**, 651–654.
DeMattei, R. C., and Feigelson, R. S. (1989). *J. Cryst. Growth* **97**, 333–336.
Dijkstra, B. W., Drenth, J., Kalk, K. H., and Vandermaelen, P. J. (1978). *J. Mol. Biol.* **124**, 53–60.
Drenth, J., Helliwell, J. R., and Littke, W. (1987). *In* "Fluid Sciences and Materials Science in Space" (H. U. Walter, ed.), pp. 451–476. Springer-Verlag, Berlin and New York.
Dreusicke, D., Karplus, P. A., and Schulz, G. E. (1988). *J. Mol. Biol.* **199**, 359–371.
Durbin, S. D., and Feher, G. (1986). *J. Cryst. Growth* **76**, 583–592.
Edwards, S. L., Kraut, J., Xuong, N., Ashford, V., Halloran, T. P., and Mills, S. E. (1988). *J. Mol. Biol.* **203**, 523–524.
Feher, G., and Kam, Z. (1985). *In* "Methods in Enzymology" (H. W. Wyckoff, C. H. W. Hirs, and S. N. Timasheff, eds.), Vol. 114, pp. 77–112. Academic Press, Orlando, Florida.
Fehribach, J. D., and Rosenberger, F. (1989). *J. Cryst. Growth* **94**, 6–14.
Feigelson, R. S. (1988). *J. Cryst. Growth* **90**, 1–13.
Fiddis, R. W., Longman, R. A., and Calvert, P. D. (1979). *J. Chem. Soc., Faraday Trans.* **75**, 2753–2761.
Finzel, B. C., and Salemme, F. R. (1985). *Nature (London)* **315**, 686–688.
Finzel, B. C., and Salemme, F. R. (1986). *Biophys. J.* **49**, 73–76.
Fitzgerald, P. M. D., and Madsen, N. B. (1986). *J. Cryst. Growth* **76**, 600–606.
Fowlis, W. W., DeLucas, L. J., Twigg, P. J., Howard, S. B., Meehan, E. J., Jr., and Baird, J. K. (1988). *J. Cryst. Growth* **90**, 117–129.
Frey, M., Genovesio-Taverne, J.-C., and Fontecilla-Camps, J. C. (1988). *J. Cryst. Growth* **90**, 245–258.
Furuno, T., and Sasabe, H. (1985). *Kino Zairyo* **5**, 77–81.
Garavito, R. M., Markovic-Housley, Z., and Jenkins, J. A. (1986). *J. Cryst. Growth* **76**, 701–709.
Gernert, K. M., Smith, R., and Carter, D. C. (1988). *Anal. Biochem.* **168**, 141–147.
Giege, R., Moras, D., Ruff, M., Thierry, J. C., Hirsch, E., Rodeau, J. L., and Mikol, V. (1988). *Proc. Int. Biotechnol. Symp., 8th* pp. 1–11.
Gilliland, G. L. (1988). *J. Cryst. Growth* **90**, 51–59.
Gilliland, G. L., and Bickham, D. (1990). *Methods: Companion Meth. Enzymol.* **1**, 6–11.
Gilliland, G. L., and Davies, D. R. (1984). *In* "Methods in Enzymology" (W. B. Jakoby, ed.), Vol. 104, pp. 370–381. Academic Press, Orlando, Florida.
Gilliland, G. L., Winborne, E. L., Masui, Y., and Hirai, Y. (1987). *J. Biol. Chem.* **262**, 12323–12324.
Harburn, G., Taylor, C. A., and Welberry, T. R. (1975). "Atlas of Optical Transforms." Cornell Univ. Press, Ithaca, New York.
Heidner, E. (1978). *J. Cryst. Growth* **44**, 139–144.
Hirsch, R. E., Lin, M. J., and Nagel, R. L. (1988). *J. Biol. Chem.* **263**, 5936–5939.
Hofrichter, J., Ross, P. D., and Eaton, W. A. (1974). *Proc. Natl. Acad. Sci. U.S.A.* **71**, 4864–4868.
Hofrichter, J., Ross, P. D., and Eaton, W. A. (1976). *Proc. Natl. Acad. Sci. U.S.A.* **73**, 3035–3039.



Jacoby, W. B. (1968). *Anal. Biochem.* **26**, 295–298.
Jones, N. D., Decter, J. B., Swartzenderber, J. K., and Landis, P. L. (1987). *Am. Crystallogr. Assoc. Meet., March 15–20, Austin, Texas* Abstr. H-4.
Kam, Z., Shore, H. B., and Feher, G. (1978). *J. Mol. Biol.* **123**, 539–555.
Kelders, H. A., Kalk, K. H., Gros, P., and Hol, W. G. J. (1987). *Protein Eng.* **1**, 301–303.
Langer, J. S. (1989). *Science* **243**, 1150–1156.
Leung, C. J., Nall, B. T., and Brayer, G. D. (1989). *J. Mol. Biol.* **206**, 783–785.
Lippmann, C., Betzel, C., Dauter, Z., Wilson, K., and Erdmann, V. A. (1988). *FEBS Lett.* **240**, 139–142.
Littke, W., and John, C. (1984). *Science* **225**, 203–204.
Littke, W., and John, C. (1986). *J. Cryst. Growth* **76**, 663–672.
Lorber, B., Kern, D., Mejdoub, H., Boulanger, Y., Reinbolt, J., and Giege, R. (1987). *Eur. J. Biochem.* **165**, 409–417.
Matsuura, Y. (1985). *Nippon Kessho Seicho Gakkaishi* **12**, 106–116.
McPherson, A. (1982). "The Preparation and Analysis of Protein Crystals." Wiley, New York.
McPherson, A. (1990). *Eur. J. Biochem.* **189**, 1–23.
McPherson, A., and Shlichta, P. (1988). *Science* **239**, 385–387.
McPherson, A., and Spencer, R. (1975). *Arch. Biochem. Biophys.* **169**, 650–661.
Michel, H. (1983). *Trends Biochem. Sci. (Pers. Ed.)* **8**, 56–59.
Mikol, V., Rodeau, J.-L., and Giege, R. (1989). *J. Appl. Crystallogr.* **22**, 155–161.
Morita, Y. (1985). *Pharm. Tech. Jpn.* **1**, 991–995.
Morris, D. W., Kim, C. Y., and McPherson, A. (1989). *Biotechniques* **7**, 522–527.
Ollis, D., and White, S. (1990). *In* "Methods in Enzymology" (M. P. Deutscher, ed.), Vol. 182, pp. 646–659.
Ollis, D. L., Brick, P., Abdel-Meguid, S. S., Murthy, K., Chase, J. W., and Steitz, T. A. (1983). *J. Mol. Biol.* **170**, 797–801.
Orbell, J. D., Guddat, L. W., Machin, K. J., and Isaacs, N. W. (1988). *Anal. Biochem.* **170**, 390–392.
Pahler, A., Hendrickson, W. A., Gawinowicz Kolks, M. A., Argarana, C. E., and Cantor, C. E. (1987). *J. Biol. Chem.* **262**, 13933–13937.
Perry, A. L., and Palmer, R. A. (1988). *J. Mol. Biol.* **201**, 659–660.
Prasad, L., Vandonselaar, M., Lee, J. S., and Delbaere, L. T. J. (1988). *J. Biol. Chem.* **263**, 2571–2574.
Pusey, M. L., Snyder, R. S., and Naumann, R. (1986). *J. Biol. Chem.* **261**, 6524–6529.
Pusey, M., Witherow, W., and Naumann, R. (1988). *J. Cryst. Growth* **90**, 105–111.
Robert, M. C., and Lefaucheux, F. (1988). *J. Cryst. Growth* **90**, 358–367.
Rose, J. P., Yang, D., Yoo, C, S., Sax, M., Breslow, E., and Wang, B.-C. (1988). *Eur. J. Biochem.* **174**, 145–147.
Rosenberger, F. (1986). *J. Cryst. Growth* **76**, 618–636.
Rosenberger, F., and Meehan, E. J. (1988). *J. Cryst. Growth* **90**, 74–78.
Rozeboom, H. J., Kingma, J., Janssen, D. B., and Dijkstra, B. W. (1988). *J. Mol. Biol.* **200**, 611–612.
Ruff, M., Cavarelli, J., Mikol, V., Lorber, B., Mitschler, A., Giege, R., Thierry, J. C., and Moras, D. (1988). *J. Mol. Biol.* **201**, 235–236.
Salemme, F. R. (1972). *Arch. Biochem. Biophys.* **151**, 533–539.
Salemme, F. R., Genieser, L., Finzel, B. C., Hilmer, R. M., and Wendoloski, J. J. (1988). *J. Cryst. Growth* **90**, 273–282.
Schlichtkrull, J. (1957). *Acta Chem. Scand.* **11**, 439–460.

Sheriff, S., Hendrickson, W. A., Stenkamp, R. E., Sieker, L. C., and Jensen, L. H. (1985). *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1104–1107.
Sheriff, S., Hendrickson, W. A., and Smith, J. L. (1987). *J. Mol. Biol.* **197**, 273–296.
Sheshadri, B. S., and Venkatappa, M. P. (1983). *J. Sci. Ind. Res.* **42**, 615–627.
Spangler, B. D., and Westbrook, E. M. (1989). *Biochemistry* **28**, 1333–1340.
Steitz, T. A. (1971). *J. Mol. Biol.* **61**, 695–700.
Stura, E. A., and Wilson, I. A. (1990). *Methods: Companion Meth. Enzymol.* **1**, 38–49.
Takizawa, T., and Hayashi, S. (1976). *J. Phys. Soc. Jpn.* **40**, 299–300.
Thaller, C., Weaver, L. H., Eichele, G., Wilson, E., Karlsson, R., and Jansonius, J. N. (1981). *J. Mol. Biol.* **147**, 465–469.
Timasheff, S. N., and Arakawa, T. (1988). *J. Cryst. Growth* **90**, 39–46.
Van Der Laan, J. M., Swarte, M. B. A., Groendijk, H., Hol, W. G. J., and Drenth, J. (1989). *Eur. J. Biochem.* **179**, 715–724.
Ward, K. B., Perozzo, M. A., and Zuk, W. M. (1988). *J. Cryst. Growth* **90**, 325–339.
Weber, P. C. (1990). *Methods: Companion Meth. Enzymol.* **1**, 31–37.
Weber, P. C., and Salemme, F. R. (1977). *J. Mol. Biol.* **117**, 815–820.
Weber, P. C., Cox, M. J., Salemme, F. R., and Ohlendorf, D. H. (1987). *J. Biol. Chem.* **262**, 12728–12729.
Wyckoff, H. W., Hirs, C. H. W., and Timasheff, S. N. (eds.) (1985). "Methods in Enzymology," Vol. 114. Academic Press, Orlando, Florida.
Yonath, A., Mussig, J., and Wittmann, H. G. (1982). *J. Cell Biochem.* **19**, 145–155.
Young, C. C., DeMattei, R. C., Feigelson, R. S., and Tiller, W. A. (1988). *J. Cryst. Growth* **90**, 79–85.

# PROTEIN HYDRATION AND FUNCTION

By JOHN A. RUPLEY* and GIORGIO CARERI†

*Department of Biochemistry, University of Arizona, Tucson, Arizona 85716
†Dipartimento di Fisica, Universita di Roma I, Rome 00185, Italy

I. Introduction . . . 38
II. Thermodynamics . . . 41
A. Powders and Films . . . 41
B. Denaturation . . . 52
C. Nonfreezing Water . . . 54
D. Solutions . . . 56
III. Dynamics . . . 61
A. Dielectric Relaxation . . . 61
B. Percolation Model . . . 69
C. Resonance . . . 71
D. Hydrogen Exchange . . . 80
E. Spectroscopy . . . 84
F. Enzyme Activity . . . 91
G. Other Measurements . . . 95
IV. Structure . . . 99
A. Diffraction . . . 99
B. Spectroscopy . . . 107
V. Computer Simulation . . . 112
A. Molecular Dynamics . . . 112
B. Monte Carlo Simulations . . . 115
C. Accessible Surface and Thermodynamics of Hydration . . . 117
D. Other Simulations . . . 120
VI. Picture of Protein Hydration . . . 122
A. Fully Hydrated Protein . . . 126
B. Hydration Process . . . 131
C. Bound Water . . . 137
D. Hydration and Conformation . . . 139
VII. Hydration and Function . . . 141
A. Folding . . . 142
B. Chemistry of Transition States . . . 143
C. Protonic Conduction and Percolation . . . 145
D. Substrate Binding . . . 146
E. Water Networks . . . 147
F. Fluctuations and Protein Motions . . . 148
G. Fluctuations and Catalysis . . . 148
H. Membranes . . . 149
I. Complex Systems . . . 150
VIII. Conclusion . . . 152
Appendix: Percolation Theory . . . 154
A. Introduction . . . 154



Copyright © 1991 by Academic Press, Inc.
All rights of reproduction in any form reserved.

occur during the course of the reaction. More structures will certainly shed light on the mechanism of action of these proteins but structures alone will not provide all the answers. The challenge for the future remains to be able to combine biochemical and structural data to understand how this type of protein does its work.

*Acknowledgements:* I thank T Jardetzky and J Widom for their comments and suggestions.

**References**


1. Kornberg, A. & Baker, T.A. (1992). *DNA Replication.* W.H. Freeman & Co, New York.
2. Hartfull, G.F. & Grindley, N.D.F. (1988). Resolvases and DNA invertases: a family of enzymes active in site-specific recombination. In *Genetic Recombination.* (Kucherpalati, R. & Smith, G.R., eds), pp. 357–396, American Society for Microbiology, Washington, DC.
3. Sherrat, D. (1989). Tn3 and related transposable elements: site-specific recombination and transposition. In *Mobile DNA.* (Berg, D.E. & Howe, M.M., eds), pp. 163–184, American Society for Microbiology, Washington, DC.
4. Hartfull, G.F., Noble, S.M. & Grindley, N.D.F. (1987). The γδ resolvase induces an unusual DNA structure at the recombinational crossover point. *Cell* **49**, 103–110.
5. Stark, W.M. & Boocock, M.R. (1994). The linkage of a knotting reaction catalyzed by Tn3 resolvase. *J. Mol. Biol.* **239**, 25–36.
6. Rice, P.A. & Steitz, T.A. (1994). Model for a DNA-mediated synaptic complex suggested by crystal packing of γδ resolvase subunits. *EMBO J.* **13**, 1514–1524.
7. Abdel-Meguid, S.S., Grindley, N.D.F., Smyth Templeton, N. & Steitz, T.A. (1984). Cleavage of the site-specific recombination protein γδ resolvase: the smaller of two fragments binds DNA specifically. *Proc. Natl. Acad. Sci. USA* **81**, 2001–2005.
8. Sanderson, M.R., *et al.*, & Steitz, T.A. (1990). The crystal structure of the catalytic domain of the site-specific recombination enzyme γδ resolvase at 2.7 Å resolution. *Cell* **63**, 1323–1329.
9. Rice, P.A. & Steitz, T.A. (1994). Refinement of γδ resolvase reveals a strikingly flexible molecule. *Structure* **2**, 371–384.
10. Yang, W. & Steitz, T.A. (1995). Crystal structure of the site-specific recombinase γδ resolvase complexed with a 34 base pair cleavage site. *Cell* **82**, 193–208.
11. Feng, J.-A., Johnson, R.C. & Dickerson, R.E. (1994). Hin recombinase bound to DNA: the origin of specificity in major and minor groove interactions. *Science* **263**, 348–355.
12. Ellenberger, T.E., Brandl, C.J., Struhl, K. & Harrison, S.C. (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α helices: crystal structure of the protein–DNA complex. *Cell* **71**, 1223–1237.
13. Glover, J.N.M. & Harrison, S.C. (1995). Crystal structure of the heterodimeric bZip transcription factor c-Fos–c-Jun bound to DNA. *Nature* **373**, 257–261.
14. Schultz, S.C., Shields, G.C. & Steitz, T.A. (1991). Crystal structure of a CAP–DNA complex: the DNA is bent by 90°. *Science* **253**, 1001–1007.
15. Kim, Y., Geiger, J.H., Hahn, S. & Sigler, P.B. (1993). Crystal structure of a yeast TBP/TATA-box complex. *Nature* **365**, 512–520.
16. Kim, J.L., Nikolov, D.B. & Burley, S.K. (1993). Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365**, 520–527.


Alfonso Mondragón, Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, 2153 Sheridan Road, Evanston, IL 60208-3500, USA.

A McPherson, AJ Malkin and YG Kuznetsov 

# The science of macromolecular crystallization

**Structure** 15 August 1995, 3:759–768

For the X-ray crystallographer, the future of his field is, and has always been, determined by what he is able to crystallize. This is now particularly true as a result of the extraordinary improvement in X-ray detectors and sources [1], the advent of computers of amazing speed, and the development of programs that are both friendly and efficient [2]. The critical component, trailing other technology, is the growth of crystals of macromolecules having sufficient size and quality to permit X-ray analysis.

Crystallization, however, has become less problematic. With synchrotron sources [3], sensitive, fast detectors, cryogenic techniques that eliminate radiation damage [4], and more powerful phasing tools, both the number and the size of crystals required for the analysis has decreased: crystals of 30–50 μm may soon be adequate, frozen crystals often provide a complete data set, and molecular replacement can allow determination of entire structures from this one set. Equally importantly, crystallization procedures, reagents, and diagnostic tools are now available that greatly improve the probability of success for both experienced crystallographers and interested biochemists.

Persistent obstacles to even more rapid progress, however, are our poor perception of the crystallization process as it applies to macromolecules; a vague concept of the nature of these crystals; and a marginal understanding of the solutal and material effects that determine their ultimate size and quality. Compared with our understanding of conventional small molecule crystallization, we are currently at a very primitive stage. Fortunately, this is rapidly changing as new and powerful methods are applied to the investigation of macromolecular crystal growth. It is our intention here to highlight some of the more innovative of these approaches, and to show how these are beginning to provide us with both an improved qualitative picture of the phenomenon of macromolecular crystallization and a quantitative foundation for its modeling.

## Properties of macromolecular crystals and their growth

It may be useful to review, firstly, some of the unique properties of protein, nucleic acid, and virus crystals (for examples of protein and virus crystals see Fig. 1). Macromolecular crystals are relatively small in comparison with conventional crystals, rarely exceeding a millimeter on an edge. Because only one stereoisomer of a biological macromolecule exists in nature, they do not form crystals possessing inversion symmetry and, therefore, generally exhibit simple shapes that lack the polyhedral character of small molecule crystals. They are extremely fragile, often crushing at the touch; they degrade when outside of a narrow range of temperature, ionic strength, or pH; they generally exhibit weak optical properties; and they diffract X-rays to resolutions far short of the theoretical limit. The reason for most of these character deficiencies is that macromolecular crystals incorporate large amounts of solvent in their lattices, ranging from about 30%, at the lower limit, to 90% or more, in the most extreme cases.

Two other crucial differences between macromolecular and conventional crystal growth have important practical consequences. The first is that macromolecular crystals are usually nucleated at extremely high levels of supersaturation, often several hundred to a thousand percent. Small molecule crystals, on the other hand, usually nucleate at only a few percent supersaturation. Virtually every quantitative aspect of crystal growth is a direct function of supersaturation [5]. Although high levels of supersaturation may be essential to promote nucleation of macromolecular crystals, in general it is far from ideal for growth, and the many problems observed for these crystals attest to this. Furthermore, supersaturated macromolecular solutions produce, in addition to crystal nuclei, alternative solid states that we refer to collectively as amorphous precipitates. Unlike conventional systems, therefore, competition between crystals and precipitate exists at both the nucleation and growth stages and this competition is particularly acute because it is promoted by high levels of supersaturation. Because amorphous precipitates are kinetically favored, though of higher energy state, they tend to dominate the solid phase and inhibit or preclude crystal formation.

Given the complexities that beset macromolecular crystallization, can we reasonably expect this process to resemble that of conventional molecules? Evidence suggests that the answer is, in principle, yes, but in practice, no. It appears that the fundamental mechanisms and pathways of macromolecular crystal growth are the same as for conventional crystals [6,7], but the magnitudes of the underlying kinetic and thermodynamic parameters that govern the process differ dramatically [7–10].

Our traditional image of protein crystal growth appears remarkably naive, both in the light of new data, and following everyday observations of crystals in the laboratory, especially those crystals that exaggerate certain growth problems (Fig. 2). Macromolecular crystals, like conventional crystals, are neither monolithic solids of perfectly organized units, nor precise arrangements of minute blocks in an ordered 'mosaic'. More like the crust of the earth, crystals contain ordered strata, layers, pockets, and deposits of molecules arranged with more or less global order. They contain inclusions, disordered water molecules, precipitant, ions, a range of impurities, defects and

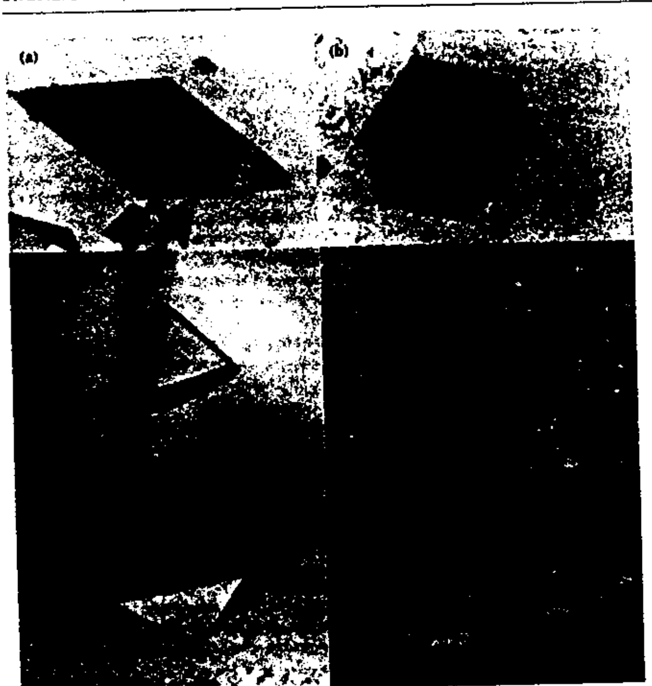


**Fig. 1.** An array of protein and virus crystals that exemplify the objectives of crystal growth researchers. **(a)** A tetragonal crystal of the sweet protein thaumatin. **(b)** An orthorhombic crystal of beef liver catalase. **(c)** A cubic crystal of satellite tobacco mosaic virus (STMV) nucleated on the wall of its quartz vessel. **(d)** An orthorhombic crystal of the same virus. **(e)** A hexagonal bipyramid of turnip yellow mosaic virus (TYMV). **(f)** A triclinic crystal of an intact mouse monoclonal antibody. The resolution limits of the diffraction patterns from these crystals vary from 1.7 Å for thaumatin and orthorhombic STMV, to about 2.7 Å for catalase and the antibody, 3.0 Å for TYMV, and to only 4.0 Å for cubic STMV.

dislocations of many sorts, and many molecules spatially inconsistent with the dominant lattice. This is important because it is the ensemble of all of these that determines the ultimate size, quality, and diffraction properties of the crystals.

Crystallization from solution [8–13], like the formation of ice from water, represents a phase change. Macromolecular crystallization utilizes three principles. The first involves perturbing the relationship between the macromolecules and solution components (water molecules and ions); the second involves altering the structure of the solvent so that macromolecules are less well accommodated, thereby promoting phase separation; and the third involves enhancing the number and strength of favorable interactions between individual macromolecules.

In the non-equilibrium state of supersaturation — where supersaturation is defined as $\sigma = \ln(c/s)$ ($c$ is concentration and $\sigma$ is the equilibrium solubility) — molecules are continuously associating to form clusters and aggregates of unknown order, whose size may tacitly be defined by a radius R. Molecules that are free in solution are continually recruited into a potential nucleus, while others dissociate. The free energy ($\delta F$) for formation of a spherical crystalline aggregate of radius R is

$$\delta F = -4\pi R^3 kT\sigma/3\Omega + 4\pi R^2\alpha$$

where $\Omega$ is the volume of a protein monomer in the crystal, $\sigma$ is the solution supersaturation, $k$ is the Boltzman constant, T is the temperature and $\alpha$ is the interfacial free energy between a crystal nucleus and the bulk solution [4]. The interfacial free energy represents the amount of work that is necessary in order to create a unit of surface of the new crystalline phase. As illustrated in Figure 3, if at some point R exceeds some critical value $R_c$ [$R_c = 4\Omega\alpha/kT\sigma$ and therefore the activation energy $\delta F_c = 16\pi\Omega^2\alpha^3/3(kT\sigma)^2$] then it will be energetically favorable for the aggregate to accumulate new molecules more rapidly than it loses old. A crystal nucleus will then be born and growth will proceed [14]. The higher $\sigma$ is, the greater the probability that molecules will be gained rather than lost to the aggregate, and the smaller is $R_c$. The energy barrier $\delta F_c$ to achievement of critical nuclear size $R_c$, therefore, is supersaturation dependent [5].

The nature of the nucleus and the process by which $R_c$ is attained are largely a mystery. It is not known whether the critical nucleus is initially ordered or assumes order through a restructuring, nor whether it forms by coalescence of arbitrary subnuclear clusters or by strict monomer or oligomer addition, or by all of these at once.

We have some reasonable hypotheses regarding differences between crystal nuclei and the aggregates that evolve to form amorphous material. In crystal nuclei, the



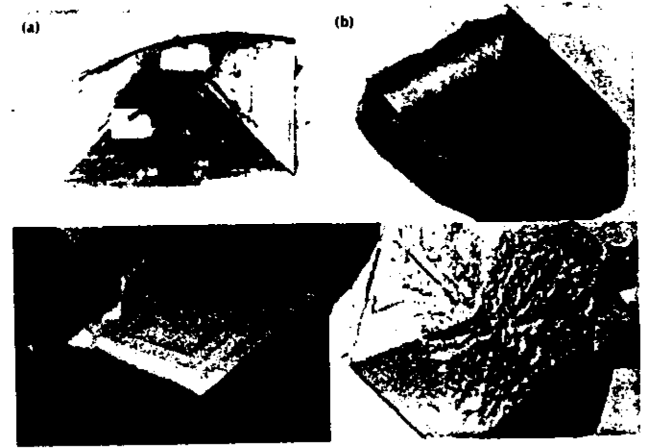


**Fig. 2.** Four crystals that permit us to see directly some of the substructure characteristic of most macromolecular crystals. **(a)** An orthorhombic STMV crystal which has begun to degrade, revealing the layers, strata, secondary nuclei, and many fault lines obscured in the intact crystal. **(b)** A rhombohedral canavalin crystal with a large macrostep, seen near the top edge of an isolated platelike macrostep near the center of the top surface, and many microcrystals embedded in, or growing from, the interior of the crystal. **(c)** Symmetric growth frames on the major face of an alpha amylase crystal, comparable to the annual rings of a tree stump, provide a record of the temperature fluctuations that occurred during its growth. Overlapping, arrowhead-shaped macrosteps can also be seen on the long, prismatic crystal from which it has nucleated. These two alpha amylase crystals have quite different unit cells and symmetries. **(d)** A rhombohedral canavalin crystal, back etched to reveal the substructure of the interior, viewed along the unique threefold axis.

macromolecules tend to make interactions with one another in all three directions in a periodic manner, whereas amorphous precipitates tend toward linear arrangements with arbitrary branching. Studies based on quasi-elastic light scattering (QELS) suggest that crystal nucleation and the formation of amorphous precipitates are distinct and can be discriminated in advance of the appearance of visible crystals [15–21]. Some investigators postulate transition from an amorphous, or fractile state into a crystal nucleus [22]. Other work suggests that conditions propitious for crystal growth may be predicted from properties of the mother liquor which are identifiable by QELS or static light scattering [17,19,20].

One model, consistent with the data, is that macromolecules initially associate through hydrophobic, non-specific interactions to produce a quasi-stable fluid aggregate analogous to the molten globule theorized for nascent polypeptides [23]. Reorientation and rearrangement then occurs in the unique internal environment of the fluid aggregate, accompanied by the formation of more geometrically rigorous hydrogen and electrostatic bonds. The result is a three-dimensionally ordered core, again analogous to a folded protein. For such a phase transition, $R_c$ may have a more complex meaning than we currently ascribe to it, specifying, possibly, the initial size of the disordered aggregate, the size of an eventually self-propagating, ordered core, or some combination thereof.

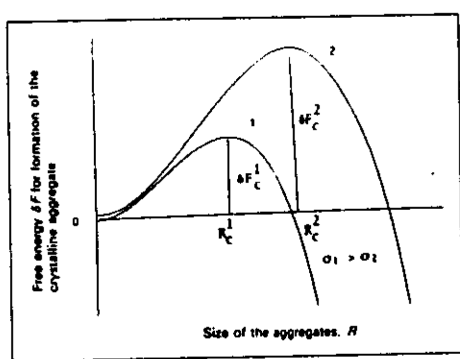



**Fig. 3.** Dependencies of $\delta F$ on R for two values of solution supersaturation $\sigma$ ($\sigma_1$ and $\sigma_2$). (Reproduced from [26], with permission.)

#### Techniques for studying nucleation events

Estimates of $R_c$ as a function of $\sigma$ have now been obtained for both proteins and spherical viruses, using QELS [18,24–26]. QELS can provide $\sigma$-dependent, quantifiable pictures of the aggregation pathways that lead ultimately to the formation of critical nuclei. In addition, QELS allows measurement of the growth rates of nuclei until they reach a size visible with a light microscope. From plots such as Figure 4a it is possible to calculate estimates of $R_c$ as a function of $\sigma$. A comparison with values measured for conventional, well characterized crystals, such as potassium dihydrogen phosphate (KDP) [27], shows that, for macromolecules, critical nuclei are of roughly similar dimensions, but one to two orders of magnitude less in terms of the number of particles incorporated. Estimates can also be obtained of the activation energy required for critical nucleus formation, $\delta F_c$ (Fig. 5a).

#### Novel techniques for studying the mechanisms of crystal growth

*Atomic force microscopy (AFM)*

Once a critical nucleus has formed, by what mechanisms does growth occur? What external factors promote and affect growth, and what are the kinetic and thermodynamic parameters that characterize the process? These have been addressed both qualitatively and quantitatively

using a variety of novel techniques. One of the most important to emerge is atomic force microscopy (AFM) [28–30]. Here, a delicately balanced stylus scans slightly above the surface of an object, in this case a growing protein or virus crystal, in such a way that it can detect and record variations in the height of the surface. It can operate in most typical mother liquors and has a spatial resolution of less than a nanometer, which enables it to resolve individual virus particles in the lattice of a crystal [31]. It can readily visualize the formation and development of steps and other features found on growing crystal surfaces.

An advantage of AFM is that repeated scanning over an area encompassing only a few nanometers to several microns, at one to two minute intervals, can be maintained with high precision for many days, and on crystals actively growing in their mother liquor. The images are, therefore, obtained *in situ* and without perturbation of the growing crystal.

Figures 6–9 illustrate the use of AFM in delineating the mechanisms by which macromolecular crystals grow.

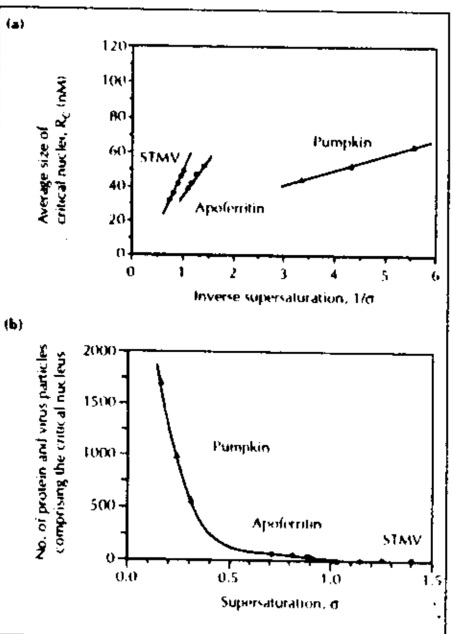



**Fig. 4.** **(a)** Dependencies of $R_c$ on inverse supersaturation, $1/\sigma$, for pumpkin globulin, apoferritin and STMV crystallization. Corresponding values for α (and molar surface energies) are: pumpkin globulin, $\alpha=6.1\times10^{-9}$ J cm$^{-2}$ (3.3 kJ mol$^{-1}$); apoferritin, $\alpha=2.7\times10^{-9}$ J cm$^{-2}$ (7.5 kJ mol$^{-1}$), STMV, $\alpha=1.8\times10^{-9}$ J cm$^{-2}$ (9.2 kJ mol$^{-1}$) **(b)** The supersaturation dependency of the number of protein and virus particles comprising the critical nucleus. (Reproduced from [26], with permission.)

**Fig. 5.** **(a)** Free energy δF versus R for formation of the pumpkin globulin crystalline aggregate at solution supersaturations σ: (1) 0.5; (2) 0.3; (3) 0.1; (4) 0.03. **(b)** The supersaturation dependency of the potential barriers to the formation of the critical nuclei for pumpkin globulin, apoferritin and STMV crystallization. (Reproduced from [26], with permission.)

These images provide evidence that macromolecular crystals do indeed grow, as do conventional crystals, by two-dimensional (2D) nucleation on surfaces (Fig. 6), by the development of spiral and multiple spiral dislocations (Fig. 7), and indeed by some novel mechanisms not previously observed for conventional crystal growth (Figs 8,9). What gives rise to the dislocations and imperfections that promote growth? They may appear when errors occur in the incorporation of new molecules into a growing step edge, or by adsorption and incorporation of a variety of impurities ranging from macromolecular aggregates to contaminating proteins or dust particles. In fact, they may be attributed to any event that produces a discontinuity in an otherwise ordered growth.

By repeated scanning, the propagation of steps, the expansion of spirals, and the formation of new surfaces are recorded. Rates of step advancement (tangential growth velocity) and sizes of critical nuclei can be measured as a function of σ, and from these, thermodynamic and kinetic parameters calculated. One such parameter, the interfacial free energy (α) of a growing crystal (its most fundamental thermodynamic parameter), provides a measure of the work required to create a unit of new surface area on the crystal. The α parameter is important because it has implications for the relative chemical environments and interactions that a molecule experiences in solution and in the crystal. From AFM studies [32,33],



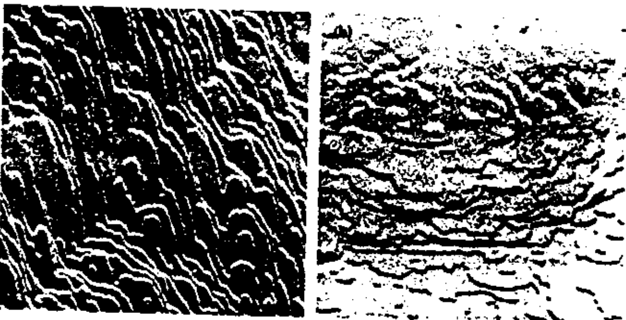


**Fig. 6.** *In situ* atomic force microscopy (AFM) images of the surfaces of **(a)** 5×5 μm and **(b)** 8×8 μm of growing thaumatin crystals, showing seas of growth steps, and the appearance of two-dimensional (2D) nuclei or islands. 2D nuclei larger than a critical size $R_c$ give rise to steps and are the sources of growth for these crystals. 2D nuclei of a size less than $R_c$ redissolve. Quantitative analysis of this phenomenon provides a good estimate of $R_c$. From AFM investigations of a number of different protein crystals, including lysozyme and catalase, it appears that growth by 2D nucleation may be the predominant mechanism for generating growth steps in macromolecular crystals, in contrast to most conventional crystals, where this mode of growth occurs only infrequently.

and also from QELS [18,24–26], α was determined for satellite tobacco mosaic virus (STMV) and canavalin to be in the range $(1.8–60)\times10^{-9}$ J cm$^{-2}$. This is about one to two orders of magnitude less than for molecules recruited from solution into a conventional crystal. When one considers the surface area of a macromolecule or virus, however, then the total energy required to incorporate it into a crystal is about the same as for a small molecule.

The low value of α for macromolecular crystals confirms, as we might have suspected, that the environments of proteins, nucleic acids, and viruses in their crystals are not, from an energetic standpoint, appreciably different from those when they are free in solution. It also suggests that the hydration shells surrounding macromolecules in solution may be changed very little upon entering the crystal.

*Michaelson interferometry*

The fundamental kinetic parameter governing crystal growth is the kinetic coefficient β. This is a measure of the kinetics of incorporation of a molecule into the crystal and is determined by the combined kinetics of transport, absorption, surface diffusion, reorientation and any other components of incorporation. The slowest of these is rate limiting and dominates the value of β. Although estimates of β have been obtained for several macromolecules using AFM [33], the most accurate approach involves the application of Michaelson interferometry to growing crystals. This method, pioneered by Chernov and his colleagues in Russia [27], is based on the interference of a wavefront of light, reflected from a growing crystal face, with a reference beam. Phase shifts occur in the reflected beam because of height variations on the crystal surface, which are caused by growth hillocks and 2D islands such as those seen in Figures 6–9. Because, as with AFM, the measurements are carried out *in situ*, as the crystal grows, a continuous series of interferograms such as those shown in Figure 10 are produced.

In the interferograms, the distance between successive dark (or successive light) fringes corresponds to a height difference on the crystal face equal to λ/2n, where λ and n are the wavelength of the light employed and the refractive index of the protein solution, respectively. The density of the fringes around any growth center is a quantitative measure of the slope of the growth hillock. The radial velocity of the fringes, measurable in time sequences, specifies the velocity of advancing step edges, known as the tangential growth velocity. As noted above for AFM studies, from the tangential velocity one can calculate β.

By choosing a point, at the growth center or otherwise, and recording the period of change in intensity there, the normal growth rate can be determined. From the tangential growth rates, β can be calculated, and indeed has been determined [32,34,35] for crystals of the protein canavalin and STMV. Values of β, in the range of $3.2–5.3\times10^{-4}$ cm sec$^{-1}$ for canavalin to $4.0–8.0\times10^{-4}$ cm sec$^{-1}$ for STMV, were deduced and, like α, found to be nearly two orders of magnitude less than for conventional crystals [33]. This indicates that the low growth rates of macromolecular crystals in comparison with inorganic crystals must be due to a greatly lowered probability of incorporation of individual molecules into the growing crystal.

*Mach-Zehnder interferometry*

A somewhat different kind of interferometry, Mach-Zehnder interferometry, is being used to investigate the immediate solutal environment of growing crystals. This is useful for quantifying concentration gradients that form around growing crystals as a consequence of the nutrient molecules becoming incorporated into the crystal. In very convection-free systems, as are found in microgravity, gels, or narrow capillaries, these 'depletion zones' may dramatically effect local supersaturation and, therefore, the overall kinetics and ultimate products of the crystal growth process.

Mach-Zehnder interferometry, which takes advantage of the variation in refractive index of a solution as a function of its macromolecule concentration [36–38], is being used to determine the relative contributions of

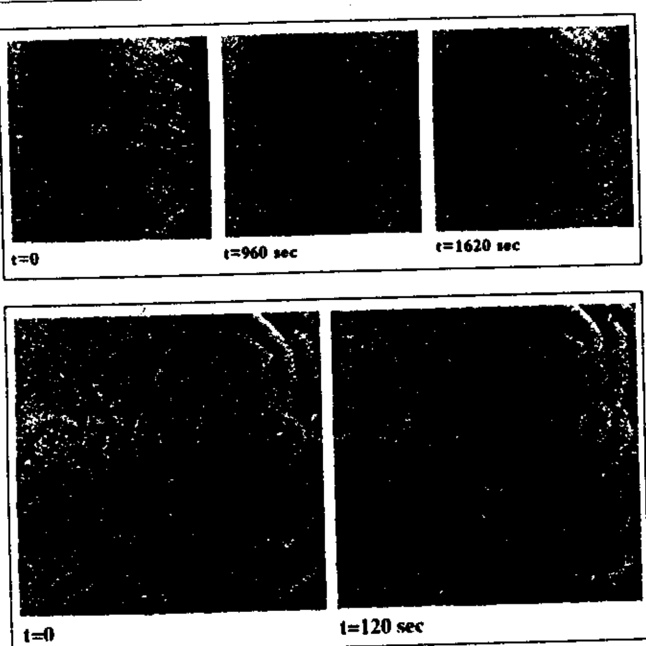



**Fig. 7.** Sequential, *in situ* atomic force microscopy (AFM) images of a 20μm × 20μm area on the surface of a growing rhombohedral canavalin crystal showing the development with time of a double spiral dislocation. A fortunate outcome of AFM studies of macromolecular crystals growing *in situ* was that the time scale of growth was much slower than that for conventional crystals, thus permitting visualization of phenomena that could not otherwise be observed. In many respects, macromolecules provide model systems superior to traditional systems for the quantitative study of crystal growth.

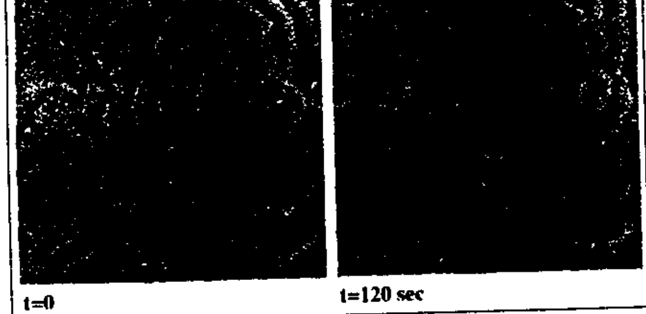



**Fig. 8.** Sequential, *in situ* atomic force microscopy (AFM) images of a 10μm × 10μm area on the surface of a growing rhombohedral canavalin crystal. Two neighboring, single screw dislocations, one right handed and one left handed, and a linear defect in the upper left corner, are shown. Progressive interaction of the steps generated by the two spirals drives growth on the surface. The edges of the growth steps here are very rough, in contrast to the smooth steps that characterize most small molecule crystals. This roughness indicates the presence of a high level of impurities in the mother liquor, which are able to incorporate into the crystal.

transport processes, both convective and diffusive, to the kinetics of growth [39,40]. This may prove to be a particularly important approach for defining the changes in fluid dynamics that occur in the absence of gravity and are responsible for the observations of altered, and generally improved, quality of macromolecule crystals grown in space [41–43].

Macromolecular crystal growth is an intricate and dynamic process which is constantly changing in character. This makes it both complex and intriguing, comparable with the changing weather patterns above the earth, or the waves and currents that roil the sea. There are usually, on a crystal face, weak centers, in addition to the strong growth centers that dominate growth and these compete with the formation of 2D nuclei. In the sequence of Figure 10, the strengths of different growth centers change with time and these centers compete with one another; heights of hillocks change; some growth centers divide into complex multihillock centers and others coalesce; steps advance, bunch, coalesce, give rise to new centers, etc. It is hardly surprising that the resultant crystal is not the perfect solid our simple models suggest.

The more realistic image of macromolecular crystal growth provided by these methods may, hopefully, contribute to a better understanding of the mosaicity and indicate the ultimate determinants of diffraction resolution, radiation sensitivity, mechanical and chemical stability, the distribution of intensities, and diffuse scatter. As demonstrated by these techniques, macromolecular crystals are a vast array of dislocations, layers and planes, step bunches, irregularities, incorporated impurities, inclusions, and regions of perfect order. Thus, when we speak of disorder in crystals, caution is advised. Disorder is not simply a measure of 'molecular wobble' at abstract lattice points.

### Conditions for macromolecular crystal growth

The information now accumulating is impressive in its extent and precision, particularly in terms of its revelation of the underlying mechanisms and processes of macromolecular crystal growth. Even so, the X-ray crystallographer must still deal with the practical questions of growing crystals for analysis [44–49]. While the physics and chemistry will ultimately lead to improved crystallization methods and practices, the diffractionist needs immediate material results. The wave of interest in macromolecular crystal growth, accompanying the introduction of the techniques described above, has, however, already had a dramatic impact in everyday laboratory crystallography.

Over the years, a vast number of reports and observations, without much form or even clear purpose, have entered the literature [50]. Attempts were made to sort, classify [51,52], and shape *ad hoc* experiments into more rational approaches [49,53–56]; to gather, collate and



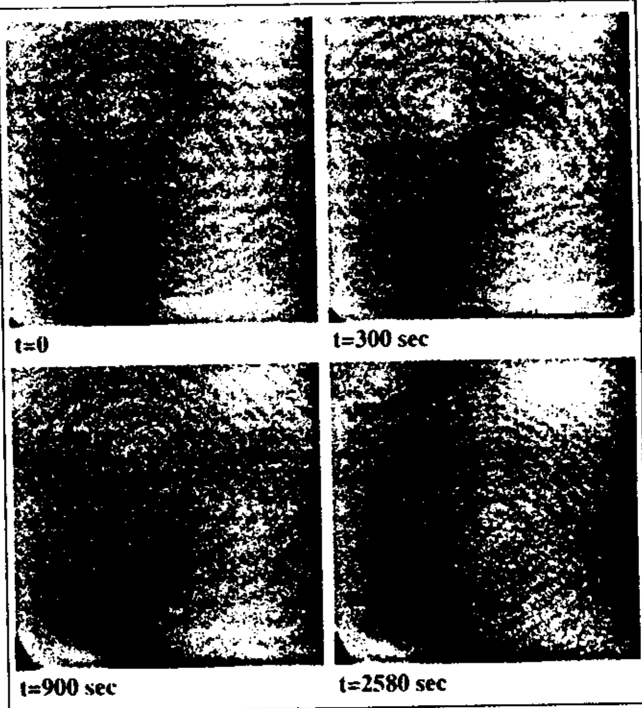



**Fig. 9.** Sequential, *in situ* atomic force microscopy (AFM) images of a 30μm × 30μm area of the surface of a growing rhombohedral canavalin crystal. Initially, two right-handed, double-spiral dislocations coexist adjacent to one another, but these compete over time. Eventually one growth spiral becomes dominant but in turn gives rise to a new, minor spiral dislocation (upper left corner of the last frame). For rhombohedral canavalin crystals, unlike most of the other macromolecular crystals studied, step generation by development of spiral dislocations appears to be the predominant growth mechanism. In canavalin, the density of dislocations and defects is very high ($10^5$–$10^6$/cm$^2$), nearly four orders of magnitude higher than for most conventional crystals. This high defect density may explain the limited diffraction quality and resolution of many macromolecular crystals.

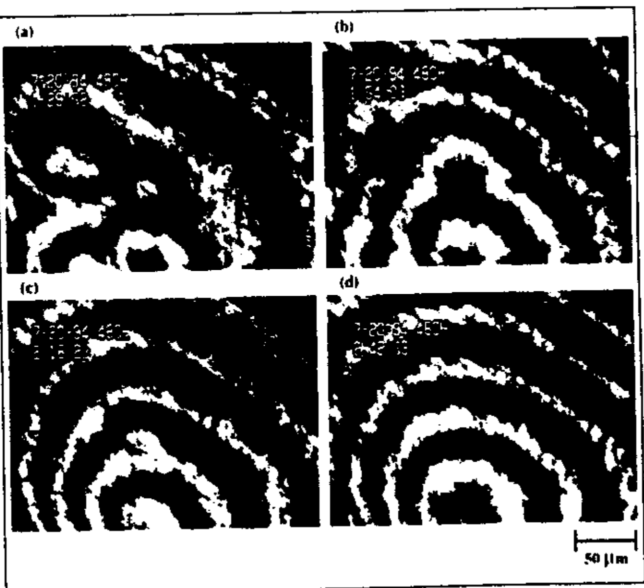



**Fig. 10.** A series of Michaelson interferograms of a growing canavalin crystal showing the patterns of fringes arising from actively developing spiral dislocations on the crystal surfaces. Fringes are produced by variations in surface height, and their motion by growth of the various sources with time. **(a)** Four independent, but closely coupled, growth centers compete in generating growth steps. As is typical with such complex growth centers, stronger dislocations eventually overwhelm the weaker, and ultimately all are consolidated into a single, strong growth center **(b–d)**. Quantitative values for the thermodynamic and kinetic parameters can be precisely calculated from the time-dependent phenomena seen here, and this has, in fact, been done for several systems [32].

systematize the experiences of thousands of investigators into rationally constructed databases [57,58]; and finally to extract, from those bases, arrays of conditions that seemed the most likely to yield protein and nucleic-acid crystals. An example of the kinds of general information available from such bases is found in Figure 11.

These sets of conditions form the foundation for the crystal screens [59–66], now commercially available (Hampton Research; Protein Solutions), that have virtually revolutionized the search for effective crystallization conditions. Although some important macromolecules continue to slip through these still imperfect nets, the number of successes is impressive. Improved screening techniques alone are, to a great extent, responsible for the explosion of new crystals, grown not only by crystallographers, but by the myriad of biochemists and molecular biologists whose need for structural information is acute. The search for novel, and perhaps even better, precipitants and conditions for crystal growth continues, and we can reasonably expect even more useful developments in this area.

With arrays of high probability conditions that need to be tested against no more than a few milligrams of protein or nucleic acid, attention has focused on the macromolecules themselves. Three characteristics have clearly come forward as being important to their crystallization: purity; stability; solubility. Failure in crystallization can probably be attributed, in general, to one or more of these.

*Purity, stability and solubility*

Homogeneity has always been recognized as important, though not always essential, for successful crystallization. The trick is to recognize when it is the crucial factor and then to direct one's efforts there. Important advances have been made with regard to purification. Not only are there now more powerful techniques and instruments available [46,47], but X-ray crystallography can now benefit to an extraordinary degree from the refinements in recombinant DNA technology and the development of high-yield expression systems [67–69]. With the introduction of endogenous purification tools like the histidine tag or glutathione sulfur transferase (GST) chimers [70], an even greater availability of ultrapure, 'crystallizable proteins' seems to be on the horizon. Recombinant systems provide us not only with pure, native proteins, but also with a way to introduce mutations and thereby create almost limitless alternatives for crystallization, should the native molecule prove unyielding. They also offer systematic approaches to the intelligent engineering of solubility properties and stability.

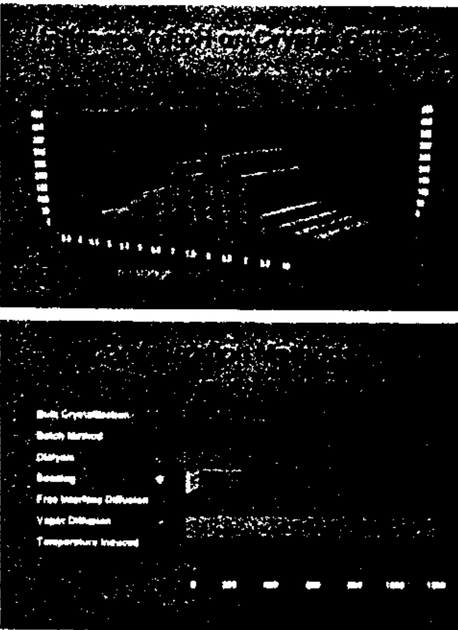



**Fig. 11.** Using information from crystallization databases, such as that assembled at the National Institute of Standards and Technology [51], analyses of methodologies, parameters, crystal properties, and reagents can readily be carried out. This database, for example, now has well over a thousand entries, and it is invaluable for the development of new crystallization screens and matrices, which are used for identifying the most probable initial conditions, and for discovering clues as to how to proceed with the crystallization process. Crystallization screens based on these databases are currently driving the success of crystallography. **(a)** The number of macromolecules crystallized against pH range. **(b)** The number of macromolecules crystallized for each precipitant type. **(c)** The number of macromolecules crystallized by each crystallization method. **(d)** The number of macromolecules crystallized for temperatures ranging from –10 °C to 100 °C.

The importance of chemical and conformational stability of the target macromolecule too, has also been recognized and given serious attention for a long time. Ligands are used to lock discrete conformations; disulfides (introduced by mutation), or chemical modifications are made to reduce mobility [71]; proteins and ribosomes from thermophilic organisms are used for their natural stability; and proteolytic or mutational truncations are made to remove naturally flexible domains or elements [46]. We are only beginning to understand how to manipulate conformation to enhance crystallization properties [72], and this area also promises much for the future, particularly for those difficult cases such as antibodies, multiple-domain enzymes and multimolecular complexes.

Perhaps the principle least recognized, but now receiving the most attention, is solubility, meaning here not only dissolution of macromolecules in a liquid, but their disaggregation into a monodisperse molecular form. Indeed, we have come to recognize aggregation and polydispersity as principal reasons why many macromolecules cannot be crystallized. This includes not only membrane and lipophilic proteins [73], but also macromolecules that have a strong propensity to aggregate, such as antibodies [74], hormone polypeptides, and proteins whose purpose is ordered aggregation (such as those from viral capsids).

Extending the observations that non-ionic detergents promote the growth of hydrophilic [75] as well as membrane proteins [76,77], light scattering studies have confirmed the salient importance of monodispersity in allowing macromolecular crystallization [17–19,63,78]. The current consensus among those involved in crystallization research is that if by some means monodispersity could be insured, then there would be a >80% chance of obtaining crystals. To promote monodispersity, approaches used involve incorporating detergents in crystallization trials and crystallizing at low protein concentrations, higher temperatures, or in the presence of chaotropic agents. This is an area that continues to be under active investigation.

Additional factors have emerged both from recent crystal growth research, and from the empirical results generated by the current revolution in crystallography. Among these factors are the importance of surfaces and their promotion of heterogeneous nucleation [79–82]; the value of unique environments such as microgravity [41–43,83], gels [84–86] and thin capillaries [87]; the broad range of polymeric precipitants found to be useful in growing crystals [88–89]; and new approaches to solving the problem of persistent microcrystals [90].

In a sense, our current assault on the problem of macromolecular crystallization has two fronts. As it did in the past, but now with far more efficiency, the empirical approach continues to provide new crystals, new information for databases, and new clues about how to proceed. Now, however, very powerful physical methods are being brought to bear, and their contributions, at the practical level, are only beginning to be appreciated.

*Acknowledgements*: Thanks to Bob Cudney, Hampton Research, Laguna Hills, CA, USA for kindly providing Figure 11. The research described here was supported by grants and contracts from the National Aeronautics and Space Administration.

## References


1. Pflugrath, J.W. (1992). Developments in X-ray detectors. *Curr. Opin. Struct. Biol.* **2**, 811–815.
2. Finzel, B.C. (1993). Software for macromolecular crystallography: a user's overview. *Curr. Opin. Struct. Biol.* **3**, 741–747.
3. Helliwell, J.R. (1992). *Macromolecular Crystallography with Synchrotron Radiation*, Cambridge University Press, Cambridge.
4. Kotlke, T. & Stalke, D. (1994). Crystal handling at low temperatures. *J. Appl. Crystallogr.* **26**, 615–619.
5. Chernov, A.A. (1984). Crystal Growth. In *Modern Crystallography III*. Springer-Verlag, Berlin.
6. Buckley, H.E. (1951). *Crystal Growth*. John Wiley and Sons, London.
7. Burton, W.K., Cabrera, N. & Frank, F.C. (1951). The growth of crystals and the equilibrium structure of their surfaces. *Philos. Trans. R. Soc. Lond. A* **243**, pp. 299.
8. Feher, G. (1986). Mechanisms of nucleation and growth of protein crystals. *J. Cryst. Growth* **76**, 545–546.
9. Feigelson, R.S. (1988). The relevance of small molecule crystal growth theories and techniques to the growth of biological macromolecules. *J. Cryst. Growth* **90**, 1–13.
10. Boistelle, R. & Astier, J.P. (1988). Crystallization mechanisms in solution. *J. Cryst. Growth* **90**, 14–30.
11. Rosenberger, F. (1986). Inorganic and protein crystal growth — similarities and differences. *J. Cryst. Growth* **76**, 618–636.
12. Arakawa, T. & Timasheff, S.N. (1985). Theory of protein solubility. *Methods Enzymol.* **114**, 49–77.
13. Kaivarainen, A.I. (1985). *Solvent-Dependent Flexibility of Proteins and Principles of Their Function*. D. Reidel Publishing Company, Dordrecht.
14. Malkin, A.J. & McPherson, A. (1993). Crystallization of pumpkin seed globulin: growth and dissolution kinetics. *J. Cryst. Growth* **133**, 29–37.
15. Feher, G. & Kam, Z. (1985). Nucleation and growth of protein crystals: general principles and assays. *Methods Enzymol.* 114, 77–112.
16. Kam, F., Shore, H.B. & Feher, G. (1978). On the crystallization of proteins. *J. Mol. Biol.* **123**, 539–555.
17. Baldwin, E.T., Crumly, K.V. & Carter, C.W. (1986). Practical, rapid screening of protein crystallization conditions by dynamic light scattering. *Biophys. J.* **49**, pp. 47.
18. Malkin, A.J., Cheung, J. & McPherson, A. (1993). Crystallization of satellite tobacco mosaic virus I. Nucleation Phenomena. *J. Cryst. Growth* **126**, 544–554.
19. Mikol, V., Hirsh, E. & Giegé, R. (1990). Diagnostic of precipitant for biomacromolecule crystallization by quasi-elastic light scattering. *J. Mol. Biol.* **213**, 187–195.
20. George, A. & Wilson, W.W. (1994). Predicting protein crystallization from a dilute solution property. *Acta Crystallogr. D* **50**, 361–365.
21. Wilson, W.W. (1993). Monitoring crystallization experiments using dynamic light scattering, assaying and monitoring protein crystallization in solution. In *Methods, a Companion to Methods in Enzymology*. (Carter, C.W., ed), Vol. 1, pp. 110–117. Academic Press, New York.
22. Georgalis, Y., Zouni, A., Eberstein, W. & Saenger, W. (1993). Formation dynamics of protein precrystallization fractal clusters. *J. Cryst. Growth* **126**, 245–260.
23. Kuntz, I.D., Jr. (1984). Stability and dynamics of globular proteins. In *The Protein Folding Problem*. (Wetlaufer, D.B., ed), pp. 65–86, AAAS Selected Symposia Series.
24. Malkin, A.J. & McPherson, A. (1993). Crystallization of satellite tobacco mosaic virus II. Postnucleation events. *J. Cryst. Growth* **126**, 555–564.
25. Malkin, A.J. & McPherson, A. (1993). Light scattering investigations of protein and virus crystal growth: ferritin, apoferritin and satellite tobacco mosaic virus. *J. Cryst. Growth* **128**, 1232–1235.
26. Malkin, A. & McPherson, A. (1994). Light scattering investigations of nucleation processes and kinetics of crystallization in macromolecular systems. *Acta Crystallogr. D* **50**, 385–395.
27. Chernov, A.A., Rashkovich, L.N., Smol'skii, I.L., Kuznetsov, Y.G., Mkrtchyan, A.A. & Malkin, A.J. (1988). Growth of KDP crystals from solution. In *Growth of Crystals*. (Givargizov, E.I. & Grinberg, S.A., eds), Vol. 15, pp. 41–91. Consultants Bureau, NY London.

28. Konnert, J.H., D'Antonio, P. & Ward, K.B. (1994). Observation of growth steps, spiral dislocations and molecular packing on the surface of lysozyme crystals with the atomic force microscope. *Acta Crystallogr.* D **50**, 603–613.
29. Gratz, J., Manne, S. & Hansma, P.K. (1991). Atomic force microscopy of atomic-scale ledges and etch pits formed during dissolution of quartz. *Science* **251**, 1343–1346.
30. Durbin, S.D. & Carlson, W.E. (1992). Lysozyme crystal growth studied by atomic force microscopy. *J. Cryst. Growth* **122**, 71–79.
31. Morris, V.J. (1994). Biological Applications of Scanning Probe Microscopies. *Prog. Biophys. Mol. Biol.* **61**, 131–185.
32. Kuznetsov, Y.G., Malkin, A.J., Greenwood, A. & McPherson, A. (1995). Michelson interferometric studies of protein and virus crystal. *J. Cryst. Growth*, in press.
33. Land, L.A., Malkin, A.J., Kuznetsov, Y.G., McPherson, A. & DeYoreo, J.J. (1995). Mechanisms of protein and virus crystal growth: an atomic force microscopy study of Canavalin and STMV crystallization. *J. Cryst. Growth*, in press.
34. Kuznetsov, Y.G., Malkin, A.J., Greenwood, A. & McPherson, A. (1995). Interferometric studies of growth kinetics and surface morphology in macromolecular crystal growth: Canavalin, thaumatin and turnip yellow mosaic virus. *J. Struct. Biol.* **14**, 184–196.
35. Vekilov, P.G., Ataka, M. & Katsura, T. (1992). Laser Michelson interferometry investigation of protein crystal growth. *J. Cryst. Growth* **130**, 317–320.
36. Francon, M. (1965). *Optical Interferometry*. Academic Press, New York.
37. Merzkirch, W.F. (1974). *Flow Visualization*. Academic Press, New York.
38. Shlichta, P.J. (1986). Feasibility of mapping solution properties during the growth of protein crystals. *J. Cryst. Growth* **76**, 656–662.
39. Grant, M.L. & Saville, D.A. (1991). The role of transport phenomena in protein crystal growth. *J. Cryst. Growth* **108**, 8–18.
40. Pusey, M., Witherow, W.K. & Naumann, R. (1988). Preliminary investigations into solutal flow about growing tetragonal lysozyme crystals. *J. Cryst. Growth* **90**, 105–111.
41. Day, J. & McPherson, A. (1992). Macromolecular crystal growth experiments on International Microgravity Laboratory-1. *Protein Sci.* **1**, 1254–1268.
42. Koszelak, S., Day, J., Leja, C., Cudney, R. & McPherson, A. (1995). Protein and virus crystal growth on international microgravity laboratory-2. *Biophys. J.*, **69**, 13–19.
43. DeLucas, L.J., *et al.*, & Darby, G. (1989). Protein crystal growth in microgravity. *Science* **246**, 651–654.
44. Carter, C.W., Jr. (1990). *Protein and Nucleic Acid Crystallization Methods: A Companion to Methods in Enzymology*. Academic Press, Orlando.
45. Ducruix, A. & Giège, R. (1992). *Crystallization of Nucleic Acids and Proteins, A Practical Approach*, IRL Press, UK.
46. Giegé, R., Lorber, B. & Theobald-Dietrich, A. (1994). Crystallogenesis of biological macromolecules: facts and perspectives. *Acta Crystallogr.* D **50**, 339–350.
47. Boyer, R.F. (1993). *Modern Experimental Biochemistry*. (2nd edn), Benjamin/Cummings Inc., New York.
48. Hurle, D.T.J. (1993). *Handbook of Crystal Growth: Part 2a. Basic Techniques*. North Holland.
49. McPherson, A. (1982). *Preparation and Analysis of Protein Crystals*. John Wiley Co., New York.
50. McPherson, A. (1991). A brief history of protein crystal growth. *J. Cryst. Growth* **110**, 1–10.
51. Bailey, K. (1942). The growth of large single crystals of proteins. *Trans. Faraday Soc.* **38**, 186–190.
52. Reichert, E.T. & Brown, A.P. (1909). *The differentiation and specificity of corresponding proteins and other vital substances in relation to biological classification and organic evolution: the crystallography of hemoglobins*. Carnegie Institute, Washington, DC.
53. Weber, P.C. (1991). Physical principles of protein crystallization. *Adv. Protein Chem.* **41**, 1–36.
54. Carter, C.W., Jr. & Carter, C.W. (1979). Protein crystallization using incomplete factorial experiments. *J. Biol. Chem.* **254**, 12219–12223.
55. Carter, C.W., Jr., Baldwin, E.T. & Frick, L. (1988). Statistical design of experiment for protein crystal growth and use of precrystallization assay. *J. Cryst. Growth* **90**, 60–73.
56. McPherson, A. (1976). The growth and preliminary investigation of protein and nucleic acid crystals by X-ray diffraction techniques. In *Methods of Biochemical Analysis* (David Glick, ed.), **23**, pp. 249–345. Academic Press, New York.
57. Gilliland, G.L., Tung, M., Blakeslee, D.M. & Ladner, J.E. (1994). Biological macromolecule crystallization database, version 3.0: new features, data and the NASA Archive for protein crystal growth data. *Acta Crystallogr.* D **50**, 408–413.
58. Hassell, A.M., Harrocks, T.J., Dashman, E.H. and Mistry, A. (1994). Two distinct approaches to crystallization results-recording databases. *Acta Crystallogr.* D **50**, 459–465.
59. Jancarik, J. & Kim, S.H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Crystallogr.* **24**, 409–411.
60. Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y. & McPherson, A. (1994). Screening and optimization strategies for macromolecular crystal growth. *Acta Crystallogr.* D **50**, 414–423.
61. McPherson, A. (1992). Two approaches to the rapid screening of crystallization conditions. *J. Cryst. Growth* **122**, 161–167.
62. Stura, E.A., Satterthwait, A.C., Calvo, J.C., Kaslow, D.C. & Wilson, I.A. (1994). Reverse screening. *Acta Crystallogr.* D **50**, 448–455.
63. D'Arcy, A. (1994). Crystallizing proteins — a rational approach? *Acta Crystallogr.* D **50**, 469–471.
64. Shaw Stewart, P.D. & Khimasia, M. (1994). Predispensed gradient matrices — a new rapid method of finding crystallization conditions. *Acta Crystallogr.* D **50**, 441–442.
65. Harris, L.J., Skaletsky, E. & McPherson, A. (1995). Crystallization of intact monoclonal antibodies. *Proteins*, in press.
66. McPherson, A. (1990). Current approaches to macromolecular crystallization. *Eur. J. Biochem.* **189**, 1–23.
67. Rees, A.R., Sternberg, M.J.E. & Wetzel, R. (1992). *Protein Engineering, A Practical Approach*, IRL Press, Oxford; New York.
68. Ago, H., *et al.*, & Xuong, N.H. (1994). Improved crystals of the toxic protein MAP by protein engineering towards the host specificity. *Acta Crystallogr.* D **50**, 355–360.
69. Nilsson, B., Forsberg, G., Moks, T., Hartmanis, M. & Uhlén, M. Fusion proteins in biotechnology. (1992). *Curr. Opin. Struct. Biol.* **2**, 569–575.
70. Privé, G.G., Verner, G.E., Weitzman, C., Zen, K.H., Eisenberg, D. & Kaback, H.R. (1994). Fusion proteins as tools for crystallization: the lactose permease from *Escherichia coli*. *Acta Crystallogr.* D **50**, 375–379.
71. Rayment, I., Rypniewski, W.R., Schmidt-base, K. & Smith, R. (1993). 3-Dimensional structure of myosin subfragment-1 — a molecular motor. *Science* **261**, 50–58.
72. Dyda, F.A.B., Hickman, T.M., Jenkins, A., Engelman, R., Craigie, J. & Davies, D.R. (1981). Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science* **266**, 1981–1986.
73. Michel, H. (1990). *Crystallization of Membrane Proteins*. CRC Press, Orlando.
74. Harris, L.J., Larson, S.B., Hassel, K.W., Day, J., Greenwood, A. & McPherson, A. (1992). The three-dimensional structure of an intact monoclonal antibody for canine lymphoma. *Nature* **360**, 369–372.
75. McPherson, A., *et al.*, & Cascio, D. (1986). An experiment regarding crystallization of soluble proteins in the presence of β-octyl-glucoside. *J. Biol. Chem.* **261**, 1969–1975.
76. Michel, H. (1990). General and practical aspects of membrane protein crystallization. In *Crystallization of Membrane Proteins*. (Michel, H., ed), pp. 73–88, CRC Press, Orlando.
77. Garavito, R.M. (1990). Crystallizing membrane proteins — experiments on different systems. In *Crystallization of Membrane Proteins*, (Michel, H., ed), pp. 89–106, CRC Press, Orlando.
78. Veesler, S., Marcq, S., Lafont, S., Astier, J. P. & Boistelle, R. (1994). Influence of polydispersity on protein crystallization: a quasi-elastic light-scattering study applied to α-amylase. *Acta Crystallogr.* D **50**, 355–360.
79. McPherson, A. & Shlichta, P. (1988). Heterogeneous and epitaxial nucleation of protein crystals on mineral surfaces. *Science* **239**, 385–387.
80. Wood, S.P., Janes, R.W., Sweeney, E. & Palmer, R.A. (1992). Crystallization of bovine pancreatic polypeptide in the presence of crystalline diglycidyl ether of bisphenol A (Araldite): epitaxy or covalent nucleation? *J. Cryst. Growth* **122**, 204–207.
81. Leung, C.J., Nall, B.T. & Brayer, G.D. (1989). Crystallization of yeast Iso-2-Cytochrome C using a novel hair seeding technique. *J. Mol. Biol.* **206**, pp. 783.
82. Ray, W.J., Jr. & Bracker, C.E. (1986). Polyethylene glycol: catalytic effect on the crystallization of phosphoglucomutase at high salt concentration. *J. Cryst. Growth* **76**, 562–576.
83. Littke, W. & John, C. (1984). Protein single crystal growth under microgravity. *Science* **225**, 203–204.
84. Patel, S., Cudney, R. & McPherson, A. (1994). Crystallization of macromolecules in silica gels. *Acta Crystallogr.* D **50**, 479–483.
85. Thiessen, K.J. (1994). The use of two novel methods to grow protein crystals by microdialysis and vapor diffusion in an agarose gel. *Acta Crystallogr.* D **50**, 491–495.
86. Robert, M.C., Bernard, Y. & Lefaucheux, F. (1994). Study of nucleation — related phenomena in lysozyme solutions. Application to gel growth. *Acta Crystallogr.* D **50**, 496–503.
87. Garcia-Ruiz, J.M. & Moreno, A. (1994). Investigations on protein crystal growth by the gel acupuncture method. *Acta Crystallogr.* D **50**, 484–490.
88. Brzozowski, A.M. & Tolley, S.P. (1994). Poly(ethylene) glycol monomethyl ethers — an alternative to poly(ethylene) glycols in protein crystallization. *Acta Crystallogr.* D **50**, 466–468.
89. Patel, S., Cudney, B. & McPherson, A. (1995). Polymeric precipitants for the crystallization of macromolecules. *Biochem. Biophy. Res. Commun.* **207**, 819–828.
90. McPherson, A. (1995). Increasing the size of microcrystals by fine sampling of pH limits. *J. Appl. Crystallogr.* **28**, 362–365.


Alexander McPherson, Alexander J Malkin and Yurii G Kuznetsov, Department of Biochemistry, University of California, Riverside, CA 92521, USA.

# Crystal structure of the superantigen enterotoxin C2 from *Staphylococcus aureus* reveals a zinc-binding site

**Anastassios C Papageorgiou[1], K Ravi Acharya[1*], Robert Shapiro[2], Edward F Passalacqua[1], Rossalyn D Brehm[3] and Howard S Tranter[3]**

[1]School of Biology and Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY, UK, [2]Center for Biochemical and Biophysical Sciences and Medicine and Department of Pathology, Harvard Medical School, Boston, MA 02115, USA and [3]Developmental Production Department, Production Division, Centre for Applied Microbiology and Research, Porton Down, Salisbury SP4 0JG, UK

**Background:** *Staphylococcus aureus* enterotoxin C2 (SEC2) belongs to a family of proteins, termed 'superantigens', that form complexes with class II MHC molecules enabling them to activate a substantial number of T cells. Although superantigens seem to act by a common mechanism, they vary in many of their specific interactions and biological properties. Comparison of the structure of SEC2 with those of two other superantigens — staphylococcal enterotoxin B (SEB) and toxic shock syndrome toxin-1 (TSST-1) — may provide insight into their mode of action.

**Results:** The crystal structure of SEC2 has been determined at 2.0 Å resolution. The overall topology of the molecule resembles that of SEB and TSST-1, and the regions corresponding to the MHC class II and T-cell receptor binding sites on SEB are quite similar in SEC2. A unique feature of SEC2 is the presence of a zinc ion located in a solvent-exposed region at the interface between the two domains of the molecule. The zinc ion is coordinated to Asp83, His118, His122 and Asp9* (from the neighbouring molecule in the crystal lattice). Atomic absorption spectrometry demonstrates that zinc is also bound to SEC2 in solution.

**Conclusions:** SEC2 appears to be capable of binding to MHC class II molecules in much the same manner as SEB. However, structure–function studies have suggested an alternative binding mode that involves a different site on the toxin. The zinc ion of SEC2 lies within this region and thus may be important for complex formation, for example by acting as a bridge between the two molecules. Other possible roles for the metal cation, including a catalytic one, are also considered.

**Structure** 15 August 1995, 3:769–779

Key words: enterotoxin, *Staphylococcus aureus*, superantigen, X-ray crystallography, zinc binding

## Introduction

*Staphylococcus aureus* and *Streptococcus pyogenes* produce a group of related pyrogenic toxins [1] including the staphylococcal enterotoxins (SEs) A, B, C1, C2, C3, D, and E [2,3], toxic shock syndrome toxin-1 (TSST-1 [4]), exotoxins A and B [5], and the streptococcal exotoxins A, B and C [6]. These toxins cause a number of illnesses, including toxic shock syndrome and scarlet fever [7]. The SEs, in particular, are known to be involved in causing the emesis and diarrhoea associated with staphylococcal food poisoning [8].

The pyrogenic toxins have been shown to function as 'superantigens' [9]. Like classical antigens, they bind to MHC class II molecules [10–12] and subsequently form ternary complexes with receptors on T cells, thereby stimulating proliferation and increased cytokine production by these cells [7,9,13]. However, superantigens differ from conventional antigens in that they are not subject to proteolytic processing prior to MHC class II binding and they interact primarily with regions outside the antigen-binding groove. Moreover, superantigen-mediated T-cell activation is achieved predominantly through interaction with the β-variable region ($V_\beta$) of the T-cell receptor (TcR) [14,15], whereas ordinary peptide antigens bind to several variable elements of the TcR. As a consequence, superantigens stimulate a much larger number of T cells than do other antigens. Despite the common mechanism shared by all superantigens, individual members of this family differ in their interactions with MHC class II molecules and TcRs and in some of their physiological effects. X-ray crystallographic studies on SEB [16], TSST-1 [17,18], and their complexes with an MHC class II molecule (HLA-DR1) [19,20] have recently begun to shed light on the structural basis for these functional differences.

Most of the available data on bacterial superantigens have been obtained with SEA, SEB and TSST-1. SED and SEE are only produced in small quantities, thus making structural studies difficult. However, SECs have been purified to homogeneity and have been crystallized [21,22]. The three SEC subtypes differ in amino acid sequence at only 10 residues [23], but nonetheless exhibit differences in their interactions with T cells [9]. In order to further understand the molecular basis of the superantigenic behaviour of SEs, we have determined the structure of SEC2 at 2.0 Å resolution using X-ray crystallography. The core topology of the molecule is similar to that reported for SEB [16] and TSST-1 [17,18], with which it shares 63% and ~24% sequence identity, respectively. However, SEC2 contains a novel

*Corresponding author.

# Crystallization of biological macromolecules for X-ray diffraction studies

Gary L Gilliland* and Jane E Ladner†

Advances in the crystallization of biological macromolecules have come not only from the application of biochemical, molecular biological and immunological principles and techniques, but also from continued efforts to understand the crystallization process. Developments in crystallization methodologies, protocols, and reagents are also facilitating crystallization efforts.

**Addresses**
Center for Advanced Research in Biotechnology of the Maryland Biotechnology Institute and the National Institute of Standards and Technology, 9600 Gudelsky Drive, Rockville, MD 20850, USA
*e-mail: gary@ibm3.carb.nist.gov
†e-mail: jane@iris8.carb.nist.gov

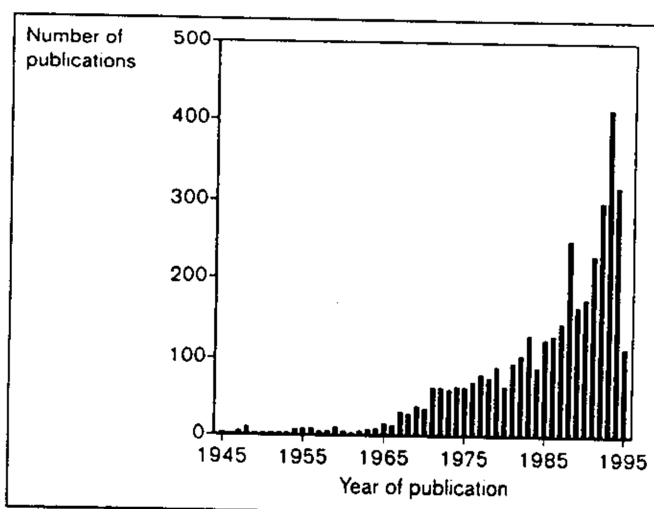
**Current Opinion in Structural Biology** 1996, 6:595–603

© Current Biology Ltd ISSN 0959-440X

**Abbreviations**

| | |
|---|---|
| 2D | two-dimensional |
| 3D | three-dimensional |
| PEG | polyethylene glycol |

**Figure 1**




The total number of publications in Version 3.0 of the Biological Macromolecules Crystallization Database [2] for each year of publication. The number of references is directly correlated with the number of crystal entries contained in the database.

## Introduction

The three-dimensional structure determination of a biological macromolecule by X-ray crystallographic techniques requires the production of large single crystals. Establishing the correct conditions for the crystallization of a particular biological macromolecule is an empirical process that typically uses techniques and reagents that have proven successful in other cases. The accelerating growth of the Brookhaven Protein Data Bank [1] and the growing number of entries in the Biological Macromolecule Crystallization Database [2] (see Fig. 1) are two measures of the usefulness of these approaches.

Despite the many successes of crystallization studies, however, our knowledge of macromolecular crystal-growth processes remains rather limited. Recently, various experimental techniques have been newly applied to the investigation of the crystallization process, furthering our understanding of how biological macromolecules crystallize (for recent reviews, see [3•,4•]).

This article reviews recent results from studies directed towards obtaining crystals of biological macromolecules for diffraction experiments, and from studies aimed at understanding many aspects of the crystal-growth process. Many recent advances in crystallization have come from the application of biochemical, molecular biological, and immunological principles and techniques. The importance of macromolecular purity and solubility remains the focus of crystal-growth studies, and the application of a variety of techniques has advanced our understanding of crystal nucleation and growth. Recent crystal-growth studies in space have continued the trend towards investigating the influence of gravity on crystal quality and the crystallization process. In addition, modifications of standard methods and strategies for crystallization, together with new crystallization techniques, offer novel ways to approach the crystallization of a particular biological macromolecule. Though the focus of this review is the crystallization of soluble biological macromolecules, the principles and techniques described here are applicable to membrane proteins also (for a recent review, see [5]).

## Sample selection, purity, and solubility properties

In the early years of protein crystallography, it was repeatedly discovered that changing the biological source or limited proteolysis of a macromolecule used in the crystallization experiments would often produce (or improve) crystals suitable for diffraction studies. Thus, it was discovered that small changes in the amino acid sequence are often sufficient to influence favorably the crystallization behavior of a protein. Systematic protease treatment of proteins still provides a route towards improving the diffraction quality of crystals [6].

### Site-directed mutagenesis

The powerful tools of recombinant DNA technology provide the means to make specific changes in a

protein sequence to improve its solubility and hence crystallizability. Success in the crystallization and structure determination of HIV integrase [7] provides perhaps the best example. Attempts at crystallizing this protein in many laboratories were carried out before anyone had employed a systematic approach of improving the crystallizability by changing hydrophobic residues to hydrophilic residues. This has been extended in studies of a protein–nucleic acid complex by combining alterations of the protein sequence with the variation of the nucleic acid sequence of an RNA hairpin to optimize the production and diffraction quality of crystals [8].

Site-directed mutagenesis also offers a mechanism of investigating the role of specific residues in crystal lattice contacts and of introducing new interactions at lattice contact points. For example, a lattice contact in crystals of glutathione reductase from *Escherichia coli* was strengthened by the addition of a new interaction introduced by a double mutation, Ala86His and Ala90Tyr [9]. The diffraction quality of the crystal was not improved but the crystals of the variant grew 40 times faster than the wild-type crystals. Another recent example is the alteration of the two-dimensional (2D) crystallization behavior of horse apoferritin [10]. Two residues, Asp84 and Gln86, known to interact with $Cd^{2+}$ ions essential for crystal-growth, were replaced with serine. Two new 2D crystal forms that grow in conditions independent of $Cd^{2+}$ ions were observed.

#### Fusion proteins

Several investigators have proposed the use of fusion protein systems to aid in difficult peptide or protein crystallization problems [11–13]. Privé and co-workers [11] have proposed the fusion of carrier proteins that can be introduced into an internal position of a target protein to increase the solubility and hence its crystallizability. The carrier protein requirements include: solubility, a single compact domain, crystallizability, N- and C-termini close together on the surface of the structure, no disulfide bonds, easily cloned and expressed, larger than the target protein, a measurable enzymatic activity or color, and purification by affinity chromatography. Privé and co-workers [11] introduced the *E. coli* cytochrome $b_{562}$ into lactose permease, a membrane protein: however, the method should prove useful for other proteins with low solubility or stability.

Two other studies involving crystallization and structure determination of fusion proteins used other carrier proteins, chicken egg-white lysozyme [12] and *Shistosoma japonicum* glutathione S-transferase [13]. The initial phases in the crystal structures were obtained by the molecular-replacement technique. Both peptides, the human fibrinogen γ-chain C-terminal segment (residues 398–411) and a hexapeptide conserved epitope of HIV-1 gp41, were fused to the C-terminus of the carrier protein, lysozyme and glutathione S-transferase, respectively. In both cases, the peptide conformations are stabilized not only by interactions with the carrier protein but also by interactions with symmetry-related carrier molecules. Although, in both cases, the question remains whether or not the conformation of the peptide is similar to that found in the natural protein, this technique holds the promise of providing structural information for this class of small, difficult-to-crystallize structures.

#### Complexes with antibody fragments

Since the report of the first high-resolution crystal structure of an antigen–antibody complex, an Fab bound to hen egg-white lysozyme [14], many complexes of antigens and antibody fragments have followed, revealing the structural principals that govern antibody–antigen interactions. This immunological strategy has also yielded structures of biological macromolecules that have otherwise proved difficult to crystallize. The use of Fab and Fv in the crystallization and the structure determination of biological macromolecules and viruses has been reviewed recently [15•]. The combination of the immunoglobulin fragment and the antigen will often drastically change the solubility properties and stability of the antigen, making it more amenable to crystallization. A novel study that successfully crystallized a complex of chicken egg-white cystatin and a recombinant Fab fragment was enhanced by generating three variant Fabs with differing constant domains [16]. The antibody fragment structure is often easily determined by molecular-replacement methods and can provide the initial phases that leads directly to a complete structure solution of the complex.

#### Macromolecule purity

The effects of purity on the crystallization of chicken egg-white lysozyme continue to be investigated [17–19]. These studies support much earlier work that indicated that the purer the biological macromolecule the better the chance of producing suitable crystals for diffraction studies and of reproducing the crystals from preparation to preparation of the macromolecule.

Analysis of impurities in commercial preparations [17,19] detected varying amounts of contamination by other proteins present in chicken egg white. Dynamic light-scattering measurements indicate aggregation of lysozyme and contaminating proteins occurs, and it was speculated that this may lead to heterogeneous nucleation leading to the formation of ill-shaped microcrystals [17].

Ewing and co-workers [19] identified in commercial preparations of lysozyme three classes of impurities that influence the crystallization behavior: contaminating proteins, small molecules, and heterogeneous forms of the protein. Of these, the heterogeneous forms of lysozme appeared to have the greatest perturbing influence on crystallization behavior.

A new study on the growth of chicken egg-white lysozyme contaminated with turkey egg-white lysozyme, a protein with 95% sequence identity, shows crystal nucleation inhibition and morphology changes [18]. This study contrasts growth studies in solution with those in agarose and silica gels. A higher concentration of contaminant is tolerated in the gel experiments than in the solution experiments. The findings show that lower concentrations of contaminants affect nucleation whereas higher concentrations influence crystal-growth morphology.

### Aggregation state and solubility

The crystallizability of a biological macromolecule is dependent upon its solubility properties. Static and dynamic light-scattering measurements have been used to examine the aggregation state and solubility of biological macromolecules [20]. A correlation between the presence of a monodisperse species of biological macromolecules in the crystallization solution and crystal-growth has been observed. Recent studies of lysozyme solutions showing that the hydrodynamic interactions of the protein decrease significantly with increasing salt concentration agree with earlier studies [21]. As the salt concentration is increased, the interactions between protein molecules change from repulsive to attractive. These new studies suggest that salt ions are binding to the surface of the protein, or that their presence changes the dielectric constant of the medium.

The effects of different anions on the solubility of an acidic protein *Hypoderma lineatum* collagenase, which has a pI of 4.1, was measured at pH 7.2 [22•]. The anions of ammonium salts were ranked as $HPO_4^{2-}/HPO_4^{-} > SO_4^{2-} > citrate^{3-}/citrate^{2-} >> Cl^{-}$ in their ability to decrease the solubility of this protein. This is in agreement with the ancient studies of Hofmeister [23], who found similar results with the ability of anions to precipitate chicken egg-white proteins, but in complete contrast with findings for basic proteins such as lysozyme [24]. If these findings can be generalized, the iosoelectric point of the biological macromolecule may dictate the choice of salts for crystallization trials.

A comparison of the solubility of chicken egg-white lysozyme in solutions with either $H_2O$ or $^2H_2O$ has produced interesting results [25•]. The studies show that lysozyme is 1.3 times more soluble in $^2H_2O$ than in $H_2O$. The higher solubility, and thus the higher supersaturation, may prove useful in crystal-growth studies if this observation holds true for other biological macromolecules.

## Fundamentals of crystal growth

The crystal-growth process for biological macromolecules is no different from that for other substances. The process can be divided into discrete stages, nucleation, crystal-growth, and cessation of growth. Each of these stages is dependent upon the solution properties of the biological macromolecule and the state of the system.

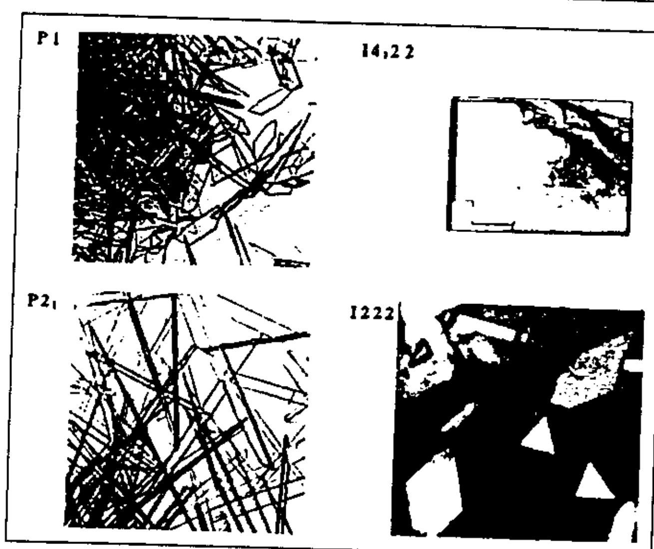
The study of aspects of the crystal-growth process has become a focus of many laboratories throughout the world. This work has established the general principle that the conditions that favor nucleation are different from those required to sustain crystal-growth.

### Nucleation and epitaxial growth

Two common problems are associated with nucleation: either no nucleation occurs, and hence no crystals are produced; or too many nuclei form, producing a shower of microcrystals that may never grow large enough for diffraction studies. The use of heterogeneous nucleants to initiate nucleation and the use of filtration to reduce nucleation are two approaches to solving these problems.

Recent work has highlighted a promising method of growing crystals by epitaxial nucleation on a charged lipid layer [26]. In the original application of this method, biotinated lipid layers were used to produce 2D crystals of streptavidin that nucleated the growth of three-dimensional (3D) crystals of streptavidin [27]. The approach was then broadened to show that charged lipids could be used instead of lipids with a bound ligand (see Fig. 2) [28••].

**Figure 2**




The morphology of four crystal forms of apostreptavidin grown in hanging drops. The $I4_122$ crystal shown was grown on a coverslip using biotinylated lipids to nucleate growth. The pictures were all taken at the same magnification; the width of each picture is equivalent to 2 mm. Reproduced with permission from [28••].

This method is, in some respects, an extension of the growth of protein crystals by epitaxial nucleation on the surfaces of minerals [29]. In this case, non-specific electrostatic interactions of the protein at the lipid surface result in the concentration of the protein in that region and thus an increase in the possibility of nucleation.

The crystals grow very rapidly and at lower protein concentrations than those normally used. For the growth of streptavidin crystals, the protein concentration can be as low as 10 μM [28••].

In another experiment with streptavidin, 2D crystals were grown using a monolayer of an iminodiacetate-Cu(II) lipid [30••]. The surface His87 was shown to bind to the Cu(II) ion. Mutants His87Ala and His127Cys were used to demonstrate the specificity of the binding. It was proposed that the method would be general for proteins with solvent-exposed histidine residues and could be expanded to include proteins that were specifically mutated to display surface histidines. The further expansion to altering the conditions to allow 3D crystal-growth seems imminent.

Several techniques for the control of nucleation have been proposed [31,32]. In these studies, filtration of the protein solution was found to reduce nucleation. By varying the exclusion size of the filters, it was shown that the number of nuclei decreases with decrease in pore size. Nucleation was also controlled by placing the droplets under parafilm oil, reducing the contact region of the supporting vessel, and by changing the temperature once nucleation has occurred to favor crystal-growth [31]. The use of gelled media in nucleating studies reduced the nucleation events, whereas the aging of a protein solution was shown to have little or no effect on nucleation [32].

#### Crystal growth kinetics

Techniques for studying the kinetics of crystal-growth have been reviewed recently by McPherson and co-workers [3•]. Simple visual microscopy, Michelson and Mach–Zehnder interferometry, and atomic force microscopy are all being successfully employed to study crystal-growth kinetics and other phenomena. Visual microscopy has been employed to measure growth rates of tetragonal chicken egg-white lysozyme [33,34] and jack bean concanavalin A [35].

The use of Michelson interferometry, which has also been applied to tetragonal lysozyme, is providing insight into the crystal-growth process [36•,37–39]. In these studies, the growth of a particular crystal face, such as the (101) face of tetragonal lysozyme, is measured optically by observing the changing interference pattern as a function of time while varying a crystallization parameter. The crystal-growth mechanism of tetragonal lysozyme was shown to change at a critical saturation point [36•]. These studies have led to recommendations for conditions for growing larger and more perfect crystals by using macromolecules of the highest purity, by growing crystals at elevated temperatures, and by growing crystals at a precipitant concentration at the low end of the range of concentrations that maintain supersaturation.

Further experimental and modeling studies [37,38] have shown that the crystal-growth interface attains a convex shape that increases with crystal size and growth rate with pure protein. In contrast, the crystal-growth interface attains a concave shape when impurities are present. This effect is also dependent on crystal size, indicating its relationship to transport-induced impurity non-uniformities at the crystal-growth interface. Further studies on the influence of impurities on the crystal-growth process of tetragonal lysozyme have shown that the presence of $< 1\%$ protein impurities perceptibly alters the growth kinetics and compositional uniformity of the crystals [39].

#### Gravity

The presence of the gravitational field and its influence of the movements of molecules in solution is a constant component of all laboratory crystallization attempts. Studies carried out in space since the early 1980s on the US NASA Space Shuttles, the Russian *MIR* Space Station and other unmanned platforms have been investigating the influence of gravity in macromolecular crystal-growth (for a recent review, see [40]). In several instances, crystals grown in the early studies in space did diffract to measurably higher resolution than those grown in the laboratory, but in other cases no improvement was evident. New reports document improvements in the diffraction quality for both virus and protein crystals [41–43]. Even though the crystals sometimes diffract to higher resolution than the earth-grown counterparts [41], the crystallization process remains subject to all of the other earth-bound variables, including the choice of method, i.e. the vapor diffusion method versus liquid–liquid diffusion method [42]. In one recent example, crystallization studies of apocrustacyanin C1 grown in space and in the laboratory indicate that the crystals grown in microgravity diffract to higher resolution than those grown in the laboratory control experiments in the same reaction vessels [43]; however, the space-grown crystals did not diffract as well as those obtained using another technique.

The diffraction quality or limiting resolution of the data from a crystal is an important parameter for assessing the quality of a single crystal, but another measure of perfection is the mosaicity. Snell and co-workers [44•] have shown that the mosaicity of crystals of lysozyme grown on two different space flights offered a four-fold improvement over the mosaicity of crystals grown in control laboratory experiments. This difference in mosaic spread provides further evidence that the force of gravity is influencing the crystal-growth process in the laboratory, at least partly through gravity-induced convection and altered solution-transport properties.

## Crystallization techniques and reagents

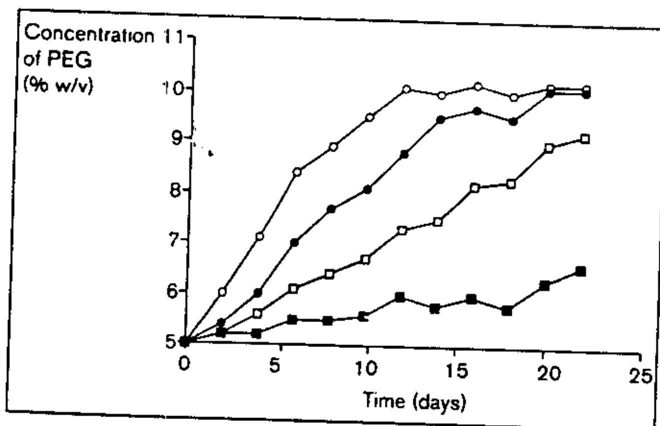
The methods used in macromolecular crystal-growth are continually being refined. The most frequently used method is vapor diffusion with either hanging or sitting

drops. Here, we review reports of attempts to overcome specific problems ranging from how to induce more proteins to crystallize to how to prevent crystals from sticking to the coverslips used to support the drops.

### Vapor diffusion and polyethylene glycol

Detailed studies of the vapor-diffusion technique and the effect of polyethylene glycol (PEG) on water–vapor equilibration have yielded some quite surprising results [45•]. Using PEG8000, for example, in a 24 µl droplet at an initial concentration of 5.0% (w/v) with twice the concentration in the reservoir, 12 weeks are required to reach equilibrium at 293 K. The addition of 200 mM sodium chloride to the droplet and reservoir decreases the equilibration time to 2 weeks (see Fig. 3). The results are a consequence of the non-ideality of aqueous PEG solutions. Understanding these interactions can add yet another set of options when planning and refining crystallization experiments.

**Figure 3**




Equilibration of vapor-diffusion experiments containing PEG solutions with NaCl. The concentration of PEG in the droplet is plotted as a function of time for equilibrations involving PEG solutions with varying quantities of NaCl. The equilibrations were carried out at 293 K using 24 µl sitting droplets with 1 ml reservoirs. All reservoir concentrations were 10%(w/v) PEG8000, plus (■) 0 mM, (□) 100 mM, (●) 200 mM, and (○) 400 mM NaCl. The initial droplet concentrations were half that of the reservoir. Each point is the mean of six observations. Adapted with permission from [45•].

Other studies have focused on determining the concentration of salt that exhibits an equivalent vapor pressure as a given concentration of PEG. The lesson is that relatively low salt concentrations are equivalent to relatively high PEG concentrations [46•]. For example, 45 mM ammonium sulfate, 60 mM sodium chloride or 100 mM magnesium sulfate heptahydrate in the reservoir all have the same vapor pressure of water as 15% (w/v) PEG8000 in the droplet.

In yet another series of experiments, the residual air pressure in the vapor chamber for PEG solutions was studied [47]. It was found that a decrease in pressure leads to an increase in the rate of equilibration, and it was suggested that appropriate alteration of the pressure in the vapor chamber might enable the time course of the crystallization to be controlled.

One problem encountered with sandwich drops or sitting drops is that the crystals can sometimes adhere quite strongly to the support surface. In response to this problem, the plug drop has been suggested [48]. A 1 cm length of 4 mm siliconized glass tubing is epoxyed to a coverslip, and a macromolecule solution is then placed in the tube, forming a plug of liquid. This is then equilibrated against a reservoir as usual but the crystals fall to the liquid–air interface as they grow instead of onto the drop-support surface.

### Shaped crystals

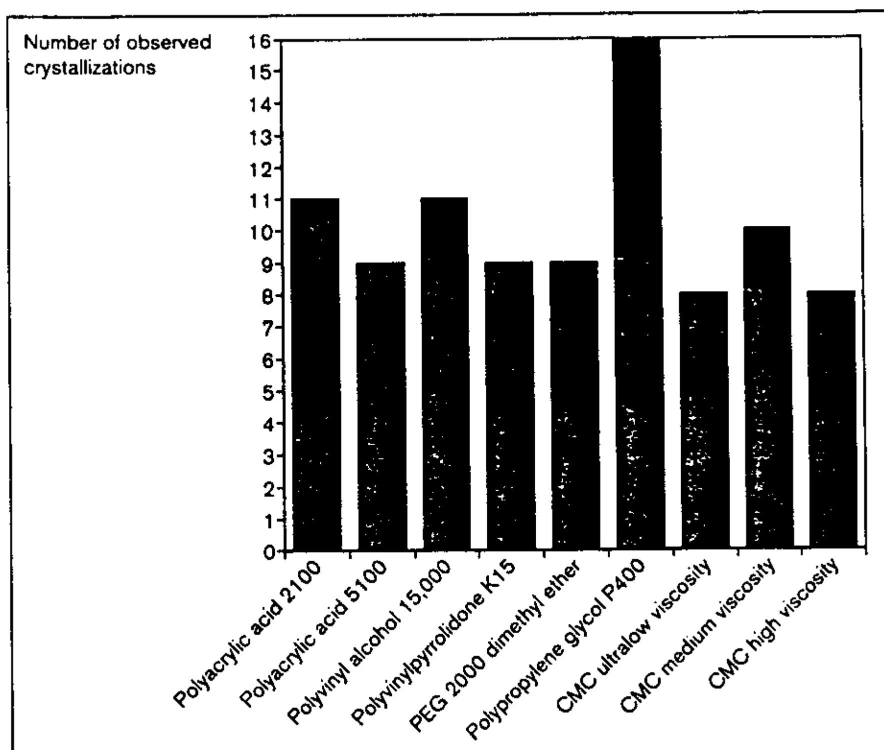
Crystallography can be applied to the study of kinetics of macromolecular reactions by conducting experiments where substrates are diffused through the crystal lattice. Growing crystals that completely fill the cross-section of a capillary has been suggested as a way of simplifying these experiments [49•]. Gel-acupuncture under isothermal conditions is used to limit the number of crystals to less than three per capillary; then the crystals are allowed to grow and assume the morphology of the inside of the capillary. Using this technique, chicken egg-white lysozyme crystals with circular cross-sections of 0.2 and 0.5 mm diameter have been grown and shown to diffract as well as crystals grown by other methods.

### Crystallization reagents

The search continues for a magic crystallizing agent. No one has claimed to have found it, but some new compounds have been offered for consideration when screening for crystals. Nine different commercially available soluble polymers have been tested with a series of biologically relevant molecules that includes 24 proteins, three viruses and two small biologically relevant molecules [50••]. With a coarse screen of crystallization conditions, all three viruses, both small molecules and 14 of the proteins crystallized. The success rate and compounds are shown in Figure 4.

Glycerol is often used in crystallization mixtures, as an additive to the macromolecule solution, a crystallizing agent, and a cryoprotectant. In a review of glycerol and other polyols by Sousa [51], it is pointed out that glucose, sorbitol, sucrose, glycerol and most other polyols increase protein thermal-denaturation temperatures. Consequently, it is postulated that the inclusion of the agents in the crystallization mix can help suppress conformational flexibility and stabilize the protein against denaturation.

**Figure 4**




Crystallization of proteins with novel soluble polymers. The graph records the number of proteins successfully crystallized (out of a battery of 24 proteins) for each of nine different soluble polymers. Reproduced with permission from [50••].

In another report of crystallization additives, short-chain phospholipids were used to lower the solubility of crambin and retard crystal-growth in the normally fastest growing direction [52•]. For crambin, the most effective additive, phosphotidylcholine, has a size comparable to the proposed binding site. In this case, the inclusion of phospholipids reduced the protein concentration needed for crystallization by 30–60-fold.

## Crystallization strategies

As mentioned in the introduction, the search for crystallization conditions is an empirical process that relies on our knowledge of what has worked in the past for other biological macromolecules. Surprisingly, no generally accepted strategy has emerged, even though many suggestions have been put forth [53], with the exception perhaps of the implementation of a fast screen procedure [54] (see below). The search for solution conditions that produce crystals suitable for diffraction studies requires that parameters such as pH, protein concentration, temperature, ionic strength, etc. be sampled over a wide range of values. Once crystal-growth conditions are found, the parameters are optimized to produce crystals of suitable size and quality for diffraction studies. Recent suggestions for sampling and optimizing conditions found in this multidimensional space [55–57], along with new fast screens [54,58], have been put forth.

### Sampling techniques

Carter and Carter [59] proposed the use of the incomplete factorial method to reduce the number of experiments that one needed to carry out in the crystallization discovery process. This method was subsequently refined to develop a more systematic method [60]. Recently, a similar procedure that has had measurable success has been reported [55]. The method starts with the initiation of experiments based on random sampling of crystallization parameters followed by a quantitative assessment of the results. The parameters are then assigned weights based on the results. Further experiments are then based on this weighted parameter set. This is repeated until crystals suitable for diffraction experiments are obtained.

In an interesting study, Carter and co-workers [56] employed quantitative analysis of full-factorial crystallization experiments to analyze the perturbation of a protein structure, that of tryptophanyl-tRNA synthetase. The study reveals that crystallization parameters including ligands alter the protein conformation, resulting in changes in crystallization behavior. Indeed, different stages of catalysis are trapped in different crystal forms depending on the environment and hence the conformation of the protein.

After microcrystals of a biological macromolecule are discovered, and the optimization of parameters such as

precipitant, pH, temperature, etc., does not produce crystals, the specific sampling of pH in small increments has been found to produce crystals suitable for diffraction studies [57]. This procedure samples pH at intervals of 0.05 units over the range of pH that microcrystals are observed. Five different sets of experiments using five different buffers over the pH range are recommended.

### Fast screening

Many recent successes in the crystallization of biological macromolecules are the results of implementing the fast-screen technique first popularized by Jancarik and Kim [54]. Fast screens use premixed solutions that have frequently produced crystals for setting up crystallization experiments. A new fast screen for RNA was devised and used in combination with a second screen that varied the RNA sequence in hammerhead RNA constructs [58]. The RNA screen employs salts, organic reagents and several different molecular weight PEGs as precipitants, with a variety of salts as additives. One of six hammerhead RNA constructs crystallized in a form suitable for diffraction studies.

## Conclusions

The continued demand for structural information of biological macromolecules has maintained the necessity for obtaining suitable crystals for diffraction studies. Indeed, the study of biological processes is offering up new challenges to structural biologists. Crystals of not only single components but also of macromolecular assemblies are required to discover the structural basis for function. The complexes include those composed of only proteins, and those composed of proteins and nucleic acids.

The fundamental problem of finding crystallization conditions for a new macromolecule or macromolecular assembly has not changed, but the tools for manipulating the systems that are under investigation to produce materials that crystallize have improved. The many successes in growing suitable crystals for diffraction studies coupled with the application of novel techniques have increased our understanding of the crystallization process. Yet further ingenuity (and luck) are often required to produce X-ray quality crystals of a specific material. Additional studies that will increase our understanding of the fundamental processes of crystal-growth are needed until a coherent picture of the processes of crystal-growth is obtained.

## Acknowledgement

The authors thank Travis Gallagher for his critical review of the manuscript.

## References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rogers JR, Kennard O, Shimanouchi T, Tasumi M: **The Protein Data Bank: a computer-based archival file for macromolecular structures.** *J Mol Biol* 1977, 112:535–542.

2. Gilliland GL, Tung M, Blakeslee DM, Ladner J: **The biological macromolecule crystallization database, version 3.0: new features, data, and the NASA archive for protein crystal growth data.** *Acta Crystallogr [D]* 1994, 50:408–413.

3. McPherson A, Malkin AJ, Kuznetsov YG: **The science of**
• **macromolecular crystallization.** *Structure* 1995, 3:759–768.
The authors review the crystal-growth process and provide a description of the application of novel techniques that are being applied to the study of the crystal-growth process.

4. Giegé R, Drenth J, Ducruix A, McPherson A, Saenger W:
• **Crystallogenesis of biological macromolecules, biological, microgravity and other physico-chemical aspects.** *Prog Crystal Growth Charact* 1995, 30:237–281.
This work provides a comprehensive review of what is known about the crystal-growth process of biological macromolecules.

5. Sowadski JM: **Crystallization of membrane proteins.** *Curr Opin Struct Biol* 1994, 4:761–764.

6. Chitarra V, Souchon H, Spinelli S, Juy M, Beguin P, Alzari PM: **Multiple crystal forms of endoglucanase celD: signal peptide residues modulate lattice formation.** *J Mol Biol* 1995, 248:225–232.

7. Dyda FAB, Hickman TM, Jenkins A, Engelman R, Craigie J, Davies DR: **Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases.** *Science* 1994, 255:1981–1986.

8. Oubridge C, Ito I, Teo C-H, Fearnley I, Nagai K: **Crystallisation of RNA–protein complexes II. The application of protein engineering for crystallisation of the U1A protein–RNA complex.** *J Mol Biol* 1995, 249:409–423.

9. Mittl PRE, Berry A, Scrutton NS, Perham RN, Schulz GE: **A designed mutant of the enzyme glutathione reductase shortens the crystallization time by a factor of forty.** *Acta Crystallogr [D]* 1994, 50:228–231.

10. Takeda S, Yoshimura H, Endo S, Takahashi T, Nagayama K: **Control of crystal forms of apoferritin by site-directed mutagenesis.** *Proteins* 1995, 23:548–556.

11. Privé GG, Verner GE, Weitzman C, Zen KH, Eisenberg D, Kaback HR: **Fusion proteins as tools for crystallization: the lactose permease from *Escherichia coli*.** *Acta Crystallogr [D]* 1994, 50:375–379.

12. Donahue JP, Patel H, Anderson WF, Hawiger J: **Three-dimensional structure of the platelet integrin recognition segment of the fibrinogen chain obtained by carrier protein-driven crystallization.** *Proc Natl Acad Sci USA* 1994, 91:12178–12182.

13. Carter DC, Ruker F, Ho, JX, Lim K, Keeling K, Gilliland G, Ji X: **Fusion proteins as alternate crystallization paths to difficult structure problems.** *Protein Peptide Lett* 1994, 1:175–178.

14. Amit AG, Mariuzza RA, Phillips SEV, Poljak RJ: **The three-dimensional structure of an antigen–antibody complex at 2.8 Å resolution.** *Science* 1986, 233:755–757.

15. Kovari LC, Momany C, Rossmann MG: **The use of antibody**
• **fragments for crystallization and structure determinations.** *Structure* 1995, 3:1291–1293.
The authors review the use of Fv and Fab antibody fragments as tools to aid in the growth of crystals of biologically important macromolecules that are difficult to crystallize. The use of the fragments in determination of phases is also discussed.

16. Schiweck W, Skerra A: **Fermenter production of an artificial Fab fragment, rationally designed for the antigen cystatin, and its optimized crystallization through constant domain shuffling.** *Proteins* 1995, 23:561–565.

17. Skouri M, Lorber B, Giegé R, Munch J-P, Candau JS: **Effect of macromolecular impurities on lysozyme solubility and crystallizability: dynamic light scattering, phase diagram, and crystal growth studies.** *J Cryst Growth* 1995, 152:209–220.

18. Provost K, Robert MC: **Crystal growth of lysozyme in media contaminated by parent molecules: influence of gelled media.** *J Cryst Growth* 1995, 156:112–120.

19. Ewing FL, Forsythe EL, Van der Woerd M, Pusey ML: **Effects of purification on the crystallization of lysozyme.** *J Cryst Growth* 1996, 160:389–397.

20. Ferré-dAnaré AR, Burley S: **Use of dynamic light scattering to assess crystallizability of macromolecules and macromolecular assemblies.** *Structure* 1994, 2:357–359.

21. Muschol M, Rosenberger F: **Interactions in undersaturated and supersaturated lysozyme solutions: static and dynamic light scattering results.** *J Chem Phys* 1995, 103:10424–10432.

22. • Carbonnaux C, Ries-Kautt M, Ducruix A: **Relative effectiveness of various anions on the solubility of acidic *Hypoderma lineatum* collagenase at pH 7.2.** *Protein Sci* 1995, 4:2123–2128.
This paper describes the effects of anions on decreasing the solubility of the acidic protein *H. lineatum* collagenase. The authors contrast these studies with those of a basic protein lysozyme, for which the results were reversed. Thus, they propose that the probability of growing crystals with a particular salt depends upon whether the macromolecule is basic or acidic.

23. Hofmeister F: **Zur Lehre on der Wikung der Salz.** *Arch Exp Pathol Pharmakol (Leipzig)* 1888, 24:223–230.

24. Riès-Kautt M, Ducruix A: **Relative effectiveness of various ions on the solubility and crystal growth of lysozyme.** *J Biol Chem* 1989, 264:745–748.

25. • Broutin I, Riès-Kautt M, Ducruix A: **Lysozyme solubility in $H_2O$ and $D_2O$ solutions as a function of sodium chloride concentration.** *J Appl Crystallogr* 1995, 28:614–617.
The solubility of lysozyme in $D_2O$ is found to be 1.3 times greater than that measured in $H_2O$. If this can be generalized, the growth of crystals in $D_2O$ may prove helpful for biological macromolecules with poor solubility.

26. Darst SA, Edwards AM: **Epitaxial growth of protein crystals from two-dimensional crystals on lipid layers.** *Curr Opin Struct Biol* 1995, 5:640–644.

27. Edwards AM, Darst SA, Hemming SA, Li Y, Kornberg RD: **Epitaxial growth of protein crystals on lipid.** *Struct Biol* 1994, 1:195–197.

28. •• Hemming SA, Bochkarev A, Darst SA, Kornberg RD, Ala P, Yang DSC, Edwards AM: **The mechanism of protein crystal growth from lipid layers.** *J Mol Biol* 1995, 246:308–316.
The use of lipid layers to facilitate the growth of 3D crystals is investigated, and it is shown that the formation of 2D crystals is not a necessary precondition.

29. McPherson A, Schlichta P: **Heterogeneous and epitaxial nucleation of protein crystals on mineral surfaces.** *Science* 1988, 239:385–387.

30. •• Frey W, Schief WR Jr, Pack DW, Chen C-T, Chilkoti A, Stayton P, Vogel V, Arnold FH: **Two-dimensional protein crystallization via metal-ion coordination by naturally occurring surface histidines.** *Proc Natl Acad Sci USA* 1996, 93:4937–4941.
The technique of epitaxial growth of crystals in 2D using a lipid layer with metal ions would appear to be the forerunner of a similar technique for 3D crystallization.

31. Blow DM, Chayen NE, Lloyd LF, Saridakis E: **Control of nucleation of protein crystals.** *Protein Sci* 1994, 3:1638–1643.

32. Hirschler J, Charon M-H, Fontecilla-Camps JC: **The effects of filtration on protein nucleation in different growth media.** *Protein Sci* 1995, 4:2573–2577.

33. Li M, Nadarajah A, Pusey ML: **Modeling the growth rates of tetragonal lysozyme crystals.** *J Cryst Growth* 1995, 156:121–132.

34. Grant ML, Saville DA: **Long-term studies on tetragonal lysozyme crystals grown in quiescent and forced convection environments.** *J Cryst Growth* 1995, 153:42–54.

35. Moré S, Saenger W: **Growth mechanism of concanavalin A crystals.** *J Cryst Growth* 1995, 153:35–41.

36. • Vekilov PG, Ataka M, Katsura T: **Growth processes of protein crystals revealed by laser Michelson interferometry investigation.** *Acta Crystallogr [D]* 1995, 51:207–219.
This article provides a description of the application of laser Michelson interferometry to monitoring the crystal growth of tetragonal lysozyme. The method allows the measurement of crystal growth *in situ* and assessment of various surface features.

37. Vekilov PG, Monaco LA, Rosenberger F: **Facet morphology response to nonuniformities in nutrient and impurity supply. I. Experiments and interpretation.** *J Cryst Growth* 1995, 156:267–278.

38. Lin H, Vekilov PG, Rosenberger F: **Facet morphology to nonuniformities in nutrient and impurity supply. II. Numerical simulations.** *J Cryst Growth* 1996, 158:552–559.

39. Vekilov PG, Rosenberger F: **Dependence of lysozyme growth kinetics on step sources and impurites.** *J Cryst Growth* 1996, 158:540–551.

40. Stoddard BL, Farber GK, Strong RK: **The facts and fancy of microgravity protein crystallization.** *Biotech Genet Engr Rev* 1993, 11:57–77.

41. Koszelak S, Day J, Leja C, Cudney, R, McPherson A: **Protein and virus crystal growth on International Microgravity Laboratory-2.** *Biophys J* 1995, 69:13–19.

42. Borisova SN, Birnbaum GI, Rose DR, Evans SV: **Experiments in microgravity: a comparison of crystals of carbohydrate-binding Fab grown on the ground, on space shuttle *Discovery*, and on space station *MIR*.** *Acta Crystallogr [D]* 1996, 52:267–271.

43. Chayen NE, Gordon EJ, Zagalsky PF: **Crystallization of apocrustacyanin C1 on the International Microgravity Laboratory (IML-2) mission.** *Acta Crystallogr [D]* 1996, 52:156–159.

44. • Snell EH, Weisgerber S, Helliwell JR, Weckert E, Hölzer K, Schroer K: **Improvements in lysozyme protein crystal perfection through microgravity growth.** *Acta Crystallogr [D]* 1995, 51:1099–1102.
Laue spot size is used to measure the mosaicity of tetragonal lysozyme crystals grown in space and on earth.

45. • Luft JR, DeTitta GT: **Chaperone salts, polyethylene glycol and rates of equilibration in vapor-diffusion crystallization.** *Acta Crystallogr [D]* 1995, 51:780–785.
This work provides useful insight into the inclusion of PEG and salt in the crystallization solution.

46. • Arakali SV, Luft JR, DeTitta GT: **Non-ideality of aqueous solutions of polyethylene glycol: consequences for its use as a macromolecular crystallizing agent in vapor-diffusion.** *Acta Crystallogr [D]* 1995, 51:772–779.
PEG is a common component of the crystallization solutions of biological macromolecules. Understanding its properties can make crystallization a more rational science.

47. DeTitta GT, Luft JR: **Rate of water equilibration in vapor-diffusion crystallization: dependence on the residual pressure of air in the vapor.** *Acta Crystallogr [D]* 1995, 51:786–791.

48. Strickland CL, Puchalski R, Savvides SN, Karplus PA: **Overexpression of *Crithidia fasciculata* trypanothione reductase and crystallization using a novel geometry.** *Acta Crystallogr [D]* 1995, 51:337–341.

49. • Garcia-Ruiz JM, Moreno A, Parraga A, Coll M: **Shaped protein single crystals.** *Acta Crystallogr [D]* 1995, 51:278–271.
The formation of crystals with the morphology of the crystallization capillary opens the way for kinetic studies of enzyme crystals.

50. •• Patel S, Cudney B, McPherson A: **Polymeric precipitants for the crystallization of macromolecules.** *Biochem Biophys Res Commun* 1995, 205:819–828.
Nine commercially available water soluble polymers are offered for inclusion in the crystal-growth screening process.

51. Sousa R: **Use of glycerol, polyols and other protein structure stabilizing agents in protein crystallization.** *Acta Crystallogr [D]* 1995, 51:271–277.

52. • Markman O, Roh C, Roberts MF, Teeter MM: **Designed additives for controlled growth of crystals of phospholipid interacting proteins: short chain phospholipids.** *J Cryst Growth* 1996, 160:382–388.
It is demonstrated that the rate of growth of crystal faces can be controlled by binding small molecules to the growth surface.

53. Gilliland GL, Bickham DM: **The Biological Macromolecule Crystallization Database: a tool for developing crystallization strategies.** *METHODS: A Companion to Methods Enzymol* 1990, 1:6–11.

54. Jancarik J, Kim S-H: **Sparse matrix sampling: a screening method for crystallization of proteins.** *J Appl Crystallogr* 1991, 24:409–411.

55. Shieh H-S, Stallings WC, Stevens AM, Stegeman RA: **Using sampling techniques in protein crystallization.** *Acta Crystallogr [D]* 1995, 51:305–310.

56. Carter CW Jr, Doublie S, Coleman DE: **Quantitative analysis of crystal growth. Tryptophanyl-tRNA synthetase crystal polymophism and its relationship to catalysis.** *J Mol Biol* 1994, 238:346–365.

57. McPherson A: **Increasing the size of microcrystals by fine sampling of pH limits.** *J Appl Crystallogr* 1995, 28:362–365.

58. Scott WG, Finch JT, Grenfell R, Fogg J, Smith T, Gait MJ, Klug A: **Rapid crystallization of chemically synthesized hammerhead RNAs using a double screening procedure.** *J Mol Biol* 1995, 250:327–332.

59. Carter CW Jr, Carter CW: **Protein crystallization using incomplete factorial experiments.** *J Biol Chem* 1979, 254:12219–12223.

60. Carter CW Jr: **Efficient factorial designs and the analysis of macromolecular crystal growth conditions.** *METHODS: A Companion to Methods Enzymol* 1990, 1:12–24.

# Crystallization of membrane proteins

Christian Ostermeier* and Hartmut Michel†

Five new membrane protein structures have been determined since 1995 using X-ray crystallography: bacterial light-harvesting complex; bacterial and mitochondrial cytochrome c oxidases; mitochondrial $bc_1$ complex; and α-hemolysin. These successes are partly based on advances in the crystallization procedures for integral membrane proteins. Variation of the size of the detergent micelle and/or increasing the size of the polar surface of the membrane protein is the most important route to well-ordered membrane protein crystals. The use of bicontinuous lipidic cubic phases also appears to be promising.

**Addresses**
*Department of Molecular Biophysics and Biochemistry, Yale University, Bass Center 433, Whitney Avenue, New Haven, CT 06520-8114, USA; e-mail: osti@laplace.csb.yale.edu
†Max-Planck-Institut für Biophysik, Abteilung für Molekulare Membranbiologie, Heinrich-Hoffmann-Strasse 7, 60528 Frankfurt/Main, Germany; e-mail: michel@mpibp-frankfurt.mpg.de

**Current Opinion in Structural Biology** 1997, 7:697–701

http://biomednet.com/elecref/0959440X00700697

© Current Biology Ltd ISSN 0959-440X

## Introduction

Up until now, the 3D structures of about 8,000 biological macromolecules, mainly soluble proteins, have been determined using X-ray crystallography. During the past few years, the availability of recombinant DNA technology to produce and to tailor the protein of interest has substantially contributed to the exponential growth in the number of new protein structures published each year. In contrast, the number of known membrane protein structures is still below 20. This fact is remarkable considering that close to 40% of the 6,000 gene products encoded by the genome of baker's yeast are expected to be integral membrane proteins. These numbers not only underscore the importance of membrane proteins, but also emphasize the enormous biochemical and structural work that remains to be done in the field of membrane proteins.

Membrane proteins are difficult to handle; the difficulties reside in the amphipathic nature of their surface. They possess a hydrophobic surface where they are in contact with the alkyl chains of the lipids, and they possess a polar surface where they are in contact with the aqueous phases on both sides of the membrane or with the polar headgroups of the lipids. In order to solubilize and to purify membrane proteins one has to add a vast excess of detergents—amphiphilic molecules that form micelles above their critical micellar concentration. The detergent micelles take up the membrane proteins and cover the hydrophobic surface of the membrane protein with their alkyl chains in a belt-like manner. The polar headgroups of the detergents face the aqueous environment.

Any crystallization strategy has to take into account the amphipathic nature of the surface of membrane proteins. Essentially, there are two possibilities for arranging membrane proteins in the form of 3D crystals [1].

First, one can try to form 2D crystals in the plane of the membrane, and then stack these membranes in an ordered way. These crystals are called 'type I'.

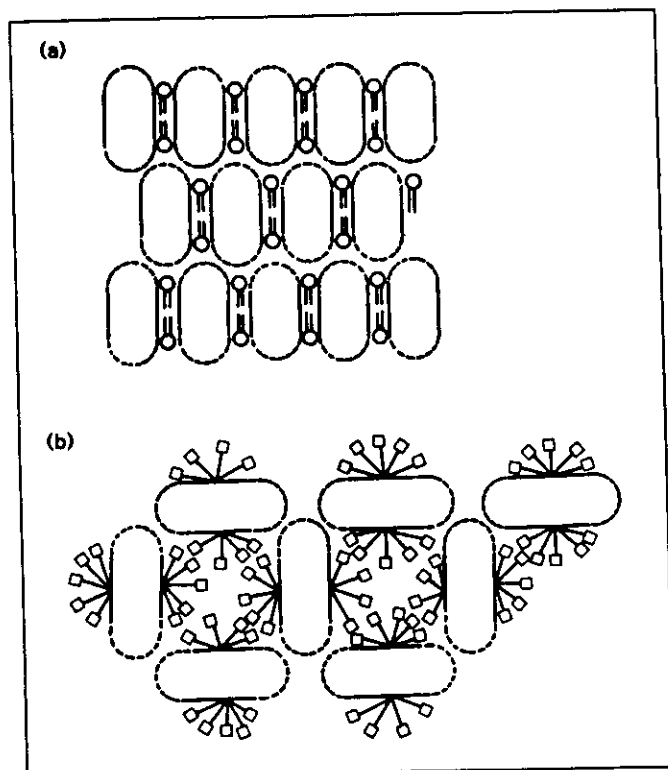
Second, one can try to crystallize membrane proteins within their detergent micelle. The crystal lattice will be established via polar contacts of the polar surface parts of the membrane protein extending out of the detergent micelles. In this case, the crystallization procedures are very similar to those for soluble proteins. For these 'type II' crystals, detergents with relatively small polar headgroups should be used in order not to cover too much of the membrane proteins polar surface.

Both types of membrane protein crystals are schematically represented in Figure 1. Mixed types of crystals appear to be possible; however, the overwhelming majority of membrane protein crystals belong to type II.

Since 1985, when the structure of the bacterial photosynthetic reaction center was presented as the first membrane protein, detailed structures from only seven families of integral membrane proteins have been published (Table 1; for reviews, see [2]). The past two years, however, have seen an enormous increase in the number of newly solved membrane protein structures, including two bacterial light-harvesting complexes [3,4], bacterial [5] and mitochondrial [6,7] cytochrome *c* oxidase, mitochondrial $bc_1$ complex [8], and α-hemolysin [9]. All integral membrane proteins crystallized so far are either pigmented or eubacterial outer membrane proteins. The latter proteins contain only β strands in their membrane-spanning section and are of extraordinary stability.

We discuss the recent advances in membrane protein crystallization. Clearly, obtaining well-ordered crystals is the bottle-neck of membrane protein structure determination. The problem of crystallizing membrane proteins cannot be reduced to the issue of which screening method or crystallization set up is to be used. Rather, thorough biochemical work and intensive protein characterization, in combination with comprehensive screening for the most suited detergent, may be the most efficient strategy to cope with the difficulties of membrane protein crystallization.

**Figure 1**




The two basic types of membrane protein crystals. **(a)** Type I: stacks of membranes contain 2D crystalline membrane proteins, which are then ordered in the third dimension. **(b)** Type II: a membrane protein is crystallized with detergents bound to its hydrophobic surface. The polar surface portion of the membrane protein is indicated by broken lines; lipids are indicated by spheres with two alkyl chains attached; detergents are indicated by squares with one alkyl chain attached. Reproduced with permission from [1].

The first step on the way to the structure of a membrane protein is to obtain a sufficient amount of pure and homogeneous protein. The second step is to find the one detergent needed to obtain well-ordered crystals for crystallization. In fact, this is the most critical step for crystallization; unfortunately, it is also the most error-prone step. Interestingly, finding the optimal crystallization conditions does not seem to be a bottle-neck. With respect to precipitating agents and pH, all membrane protein crystals obtained so far have been obtained under quite standard crystallization conditions. If the biochemist and the crystal-grower has done his/her job well, data collection and structure determination of a membrane protein are as straightforward as for nonmembrane proteins. Currently, flash cooling of membrane protein crystals is often used to enhance crystal stability in the X-ray beam [10] or for trapping reaction intermediates; however, establishing cryoconditions for membrane protein crystals seems to be much more difficult than for soluble proteins. This problem may be due to the presence of detergent micelles in membrane protein crystals.

## Most important: the wet-lab biochemistry

For crystallization trials, up to 100 mg of pure protein must be isolated. Soluble proteins can often be obtained by overexpression of the gene or cDNA, combined with the use of affinity tags for detection and purification. Refolding from inclusion bodies sometimes works well. Engineering membrane proteins for crystallization is possible in principle [11] but less helpful, as a sufficient level of overexpression rarely can be achieved. In all published cases, membrane protein crystals have been grown from proteins isolated from natural sources. In nature, mainly photosynthetic membrane proteins and those from bioenergetics are abundant, which explains why these membrane proteins are the best characterized structurally. The majority of membrane proteins in the cell are present at only very low levels. Up until now, there seems to be no general way to obtain large quantities of functional membrane proteins using recombinant DNA techniques [12].

Recently, however, a strategy for the overproduction of membrane proteins, which are usually lethal to their host cells, has been published [13••]. This strategy involves usage of selected *Escherichia coli* strains and the bacteriophage T7 RNA polymerase system for the the overproduction of a number of membrane proteins. In these strains, membrane proteins are formed in large amounts as inclusion bodies. Protein yields in the range of 100 mg per liter of bacterial cell culture have been reported. Unfortunately, the refolding of membrane proteins from inclusion bodies is mostly an unsolved problem and is one of the main challenges for the future.

**Table 1**

**Membrane protein familes for which crystal structures exist.**

| Membrane protein family | Resolution (Å) | Pigmented | β sheet |
|---|---|---|---|
| Photosynthetic reaction centers | 2.3 | Yes | No |
| Porins | 1.8 | No | Yes |
| Light harvesting complexes | 2.4 | Yes | No |
| Cytochrome c oxidases | 2.7 | Yes | No |
| α-hemolysin | 1.9 | No | Yes |
| Cytochrome $bc_1$ complex | 3.0 | Yes | No |
| Prostaglandin $H_2$ synthase | 3.5 | Yes | No |

## Detergents: expensive soaps

Since the early years of membrane protein crystallization, choosing the right detergent has been the key to success. Well-ordered crystals of the photosynthetic reaction center from the purple bacterium *Rhodopseudomonas viridis* could only be grown using N,N-dimethyl dodecylamine-n-oxide as detergent. Even use of the decyl homolog did not lead to crystals. Recent experiences confirm this observation. The cytochrome *c* oxidases provide illustrative examples. Crystallization attempts with the cytochrome *c* oxidase from bovine beef heart mitochondria continued in Yoshikawa's laboratory for about twenty years, and crystals have been obtained in a number of different detergents [14]; however, only the use of n-decyl-β-D-maltoside ($C_{10}$-maltoside) — a mild, well-known detergent — has yielded well-ordered crystals.

Cytochrome *c* oxidase from the soil bacterium *Paracoccus denitrificans* is another typical example. For the purification and crystallization of the four-subunit complex, only detergents of the maltoside-type can be used. All other detergents remove subunits III and IV leaving an active complex consisting of subunits I and II. Only n-dodecyl-β-D-maltoside ($C_{12}$-maltoside) leads to the formation of well-ordered crystals of the four-subunit oxidase as an complex with an Fv fragment [15]. Recently, the catalytically active two-subunit complex could be crystallized, again with the help of an antibody Fv fragment (see also below) in different detergents. Originally, crystals were grown using the $C_{12}$-maltoside, but these diffracted to only about 8 Å. Crystals grown with hexaethylene glycol monododecyl ether ($C_{12}E_6$) showed the same poor diffraction quality. With the $C_{10}$-maltoside, no crystals could be obtained at all. Recently, the $C_{11}$-maltoside became also commercially available. Crystals grown in this detergent diffract to better than 2.6 Å resolution (C Ostermeier, A Harrenga, U Ermler, H Michel, unpublished data). Similar crystals can be grown with cyclohexyl-hexyl-β-D-maltoside (CYMAL-6), but not with cyclohexyl-pentyl-β-D-maltoside (CYMAL-5). Cyclohexyl-heptyl-β-D-maltoside (CYMAL-7) is not yet commercially available.

These results show that even small chemical differences in the detergent can cause essential differences in the crystallization behaviour of these detergent–membrane protein complexes. The conclusion has to be drawn that more efforts should be put into screening various detergents for crystallization than into the variation of other parameters. A major problem may be the high costs of many detergents. The optimal way to cope with this hindrance is to purify the protein using a rather inexpensive detergent such as Triton X-100 or N,N-dimethyldodecylamine-N-oxide and then to exchange the detergent prior to the crystallization attempts. One should keep in mind that it may be difficult to control the completeness of the detergent exchange. In our opinion, the simplest and most efficient method for a complete exchange is ion exchange chromatography, or another method in which the membrane protein is bound to column materials and can be washed with a large amount of buffer containing the new detergent without being eluted from the column. Gel filtration or exchange by ultrafiltration is not recommended if one requires a complete exchange.

It would be helpful if a continuous set of alkyl chain lengths were commercially available for many detergent headgroups, for example, the $C_xE_y$-detergents are available only with an even number of C atoms. One should also keep in mind that mixtures of detergents often may be useful. Finally, a need still exists for new classes of detergents.

An alternative to the classic detergents may be the so-called 'amphipols' — polymers that can potentially keep membrane proteins in aqueous solution [16•]. These possess a strongly hydrophilic backbone that is decorated with hydrophobic sidechains, resulting in an amphiphilic structure. So far, amphipols have not yet been used for crystallization, but they might be useful in the future.

## Crystallization: finding the needle (detergent?) in the haystack

For the time being, most promise lies with trying to obtain a type II crystal. This approach has the advantage that the membrane protein surrounded by its belt-like detergent micelle can be treated as an ordinary soluble protein, and standard crystallization procedures can be used (for general reviews, see [1,17–20]). Most membrane protein crystals have been obtained using standard precipitants like polyethyleneglycols or salts (ammonium sulfate, potassium phosphate). The vapour-diffusion method with sitting drops is most frequently applied to achieve supersaturation of membrane proteins.

As outlined above, the choice of the detergent is the most important factor apart from the stability and homogeneity of the protein. This is understandable because the detergent micelle has to fit optimally into the crystal lattice of the protein. Attractive, polar interactions between neighboring detergent micelles appear to be helpful and to contribute to the stability of the crystal lattice. Such contacts cannot occur when the detergents have a rather short alkyl chain, thus explaining why crystals are sometimes obtained only with longer homolog of the same detergent type. That attractive interactions between detergent micelles play a role is also indicated by the fact that crystallization often occurs close to conditions in which where phase separation into a detergent-rich and a detergent-depleted phase occurs. This phase separation is caused by attractive interactions between detergent micelles [21]. In the case of the bacterial cytochrome *c* oxidase crystal, formation normally starts at the physical boundary between the detergent-rich and the detergent-depleted phase.

Detergent micelles can be made smaller by adding small amphiphilic molecules such as heptane-1,2,3-triol [1,22,23]. This approach has been successful in the case of bacterial photosynthetic reaction centers and light-harvesting complexes [4], for which rather harsh detergents with small polar headgroups can be used. It is unsuccessful when rather mild detergents, such as the alkyl maltosides, are required.

## The trick with the complex

Instead of trying to get a smaller detergent micelle, one can try to increase the surface area of the hydrophilic portion of the membrane protein. Binding a soluble protein to the membrane protein under investigation is one possibility for extending the polar regions. This strategy has been used successfully twice. The four- and two-subunit bacterial cytochrome *c* oxidases have been crystallized as a cocomplex with an Fv fragment of a monoclonal antibody ([5]; C Ostermeier, A Harrenga, U Ermler, H Michel, unpublished data). The crystallization conditions, the space group and the crystal packing of both complexes are completely different. In both crystal structures, the Fv fragment plays an essential role in forming the well-ordered crystal lattice. Another advantage of using engineered Fv fragments for cocrystallization is the possibility of using an affinity tag engineered to the antibody fragment for the rapid isolation of the whole membrane protein–antibody complex [24]. Thus, an affinity tag for purification of the membrane protein can be used even if genetic engineering of the membrane protein itself is not possible. In the case of the two-subunit oxidase, isolation has been simplified by this strategy. Starting with crude membranes, crystallization trials can be set up within six hours after starting purification (C Ostermeier, A Harrenga, U Ermler, H Michel, unpublished data). Producing the Fv fragments may be a labour-intensive and often cumbersome procedure; however, for many important membrane proteins well-characterized hybridoma cell lines are already available.

## Use of bicontinuous lipidic cubic phases

When mixed with aqueous solvents, some lipids form a bicontinuous cubic phase, in which the lipids are arranged in a curved, continuous 3D bilayer. Landau and Rosenbusch [25••] have succeeded to incorporate monomeric bacteriorhodopsin prepared from purple membranes in such a bilayer, and to use this as a matrix for crystallization. The idea is that the protein can diffuse in the bilayer, but it is also able to form 3D contacts. Landau and Rosenbusch have been able to demonstrate that bacteriorhodopsin forms small, but well-ordered 3D crystals. The X-ray data obtained from the most well-ordered crystal form indicate that the same 2D crystal lattice is formed that is observed in the native purple membrane. These membranes appear to be stacked and well ordered in the third dimension; therefore, the crystals belong to type I. It is to be hoped that this method can also be used for membrane proteins that do not have a strong tendency to form 2D crystalline arrays spontaneously. In particular, this method appears to be the only chance for membrane proteins that are unstable in detergent micelles or in the absence of added lipids.

## Conclusions

The picture that emerges is that the membrane proteins tend to form the crystal lattice; the crystal lattice that forms is strongly influenced by the polar headgroup of the detergent. Sometimes, for example in the case of the photosynthetic reaction center from the purple bacterium *R. viridis*, the headgroups are involved by forming critical protein/headgroup/protein contacts (CRD Lancaster, H Michel, unpublished data). Often, the length of the alkyl chain of the detergent has to be optimized in order to get a well-ordered crystal. A possible reason for this is that polar interactions between neighboring detergent micelles are needed to stabilize the protein crystal lattice.

The recent advances in structural membrane protein research raise some hope that crystallography of membrane proteins will be no longer a wallflower in the field of structural biology but will become a powerful tool for understanding essential functions of membrane proteins, such as cell–cell communication via hormones or neurotransmitters, transport across membranes or energy conversion. The prerequisite for membrane protein crystallography — membrane protein crystallization — is still far away from being straightforward or routine. Two of the most important problems to be solved in the near future are the overproduction of functional membrane proteins in their native membrane environments, and the refolding of recombinant membrane proteins from inclusion bodies. Patience and many long-term grants are necessary before we can state that membrane protein crystallography is no longer in its infancy.

## Acknowledgements

We are grateful to Bryan Sutton for reading the manuscript.

## References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest


1. Michel H: **Crystallization of membrane proteins.** *Trends Biochem Sci* 1983, 8:56-59.
2. Walker JE, Saraste M (Eds): **Membrane proteins.** *Curr Opin Struct Biol* 1996, 6:457-510.
3. McDermott G, Prince SM, Freer AA, Hawthornthwaite-Lawless AM, Papiz MZ, Cogdell RJ, Isaacs NW: **Crystal structure of an integral membrane light-harvesting complex from photosynthetic bacteria.** *Nature* 1995, 374:517-521.
4. Koepke J, Hu X, Muenke C, Schulten K, Michel H: **The crystal structure of the light-harvesting complex II (B800-850) from *Rhodospirillum molischianum*.** *Structure* 1996, 4:581-597.
5. Iwata S, Ostermeier C, Ludwig B, Michel H: **Structure at 2.8 Å resolution of cytochrome *c* oxidase from *Paracoccus denitrificans*.** *Nature* 1995, 376:660-669.

6. Tsukihara T, Aoyama H, Yamashita E, Tomizaki T, Yamaguchi H, Shinzawa-Itoh K, Nakashima R, Yaona R, Yoshikawa S: **Structures of metal sites of oxidized bovine heart cytochrome c oxidase at 2.8 Å.** *Science* 1995, **269**:1069-1074.

7. Tsukihara T, Aoyama H, Yamashita E, Tomizaki T, Yamaguchi H, Shinzawa-Itoh K, Nakashima R, Yaono R, Yoshikawa S: **The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 Å.** *Science* 1996, **272**:1136-1144.

8. Yu C-A, Xia J-Z, Kachurin AM, Yu L, Xia D, Kim H, Deisenhofer J: **Crystal structure of beef heart mitochondrial cytochrome-$bc_1$ complex.** *Science* 1997, **277**:60-66.

9. Song L, Hobaugh MR, Shustak C, Cheley S, Bayley H, Gouaux JE: **Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore.** *Science* 1996, **274**:1859-1866.

10. Lee JW, Chan M, Law TV, Kwon HJ, Jap BK: **Preliminary cryocrystallographic study of the mitochondrial cytochrome $bc_1$ complex: improved crystallization and flash-cooling of a large membrane protein.** *J Mol Biol* 1995, **252**:15-19.

11. Popot J-L, Saraste M: **Engineering membrane proteins.** *Curr Opin Biotechnol* 1995, **6**:394-402.

12. Griesshammer R, Tate CG: **Overexpression of integral membrane proteins for structural studies.** *Quart Rev Biophys* 1995, **28**:315-422.

13. Miroux B, Walker JE: **Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels.** *J Mol Biol* 1996,
•• **260**:289-298.

The overproduction of membrane proteins in an *Escherichia coli*, BL21(DE3), T7 RNA polymerase expression system is investigated. Mutant host strains are selected that produce the protein as inclusion bodies at very high levels. Seven membrane proteins can be expressed in levels up to 100 mg per liter of culture.

14. Shinzawa-Itoh K, Ueda H, Yoshikawa S, Aoyama H, Yamashita E, Tsukihara T: **Effects of ethyleneglycol chain length of dodecyl polyethyleneglycol monoether on the crystallization of bovine cytochrome c oxidase.** *J Mol Biol* 1995, **246**:572-575.

15. Ostermeier C, Iwata S, Ludwig B, Michel H: **Fv fragment-mediated crystallization of the membrane protein bacterial cytochrome c oxidase.** *Nat Struct Biol* 1995, **2**:842-846.

16. Tribe C, Audebert R, Popot J-L: **Amphipols: polymers that keep membrane proteins soluble in aqueous solutions.** *Proc Natl Acad Sci USA* 1996, **93**:15047-15050.
•

A new class of surfactants, the amphipols, is developed. This class of surfactants is able to keep membrane proteins soluble in aqueous solutions. Because of their size heterogeneity, however, they may not be useful for crystallization. Further developments may improve these amphipols so that they can be a useful alternative to detergents in the future.

17. Garavito RM, Markovic-Housley Z, Jenkins JA: **The growth and characterization of membrane protein crystals.** *J Cryst Growth* 1986, **76**:701-709.

18. Kühlbrandt W: **Three-dimensional crystallization of membrane proteins.** *Quart Rev Biophys* 1988, **21**:429-477.

19. Michel H (Ed): *Crystallization of Membrane Proteins.* Boca Raton, FL: CRC Press; 1991.

20. Garavito RM, Picot D, Loll PJ: **Strategies for crystallizing membrane proteins.** *J Bioenerg Biomembr* 1996, **28**:13-27.

21. Zulauf H: **Detergent Phenomena in membrane protein crystallization.** In *Crystallization of Membrane Proteins.* Edited by Michel H. Boca Raton, FL: CRC Press; 1991:53-72.

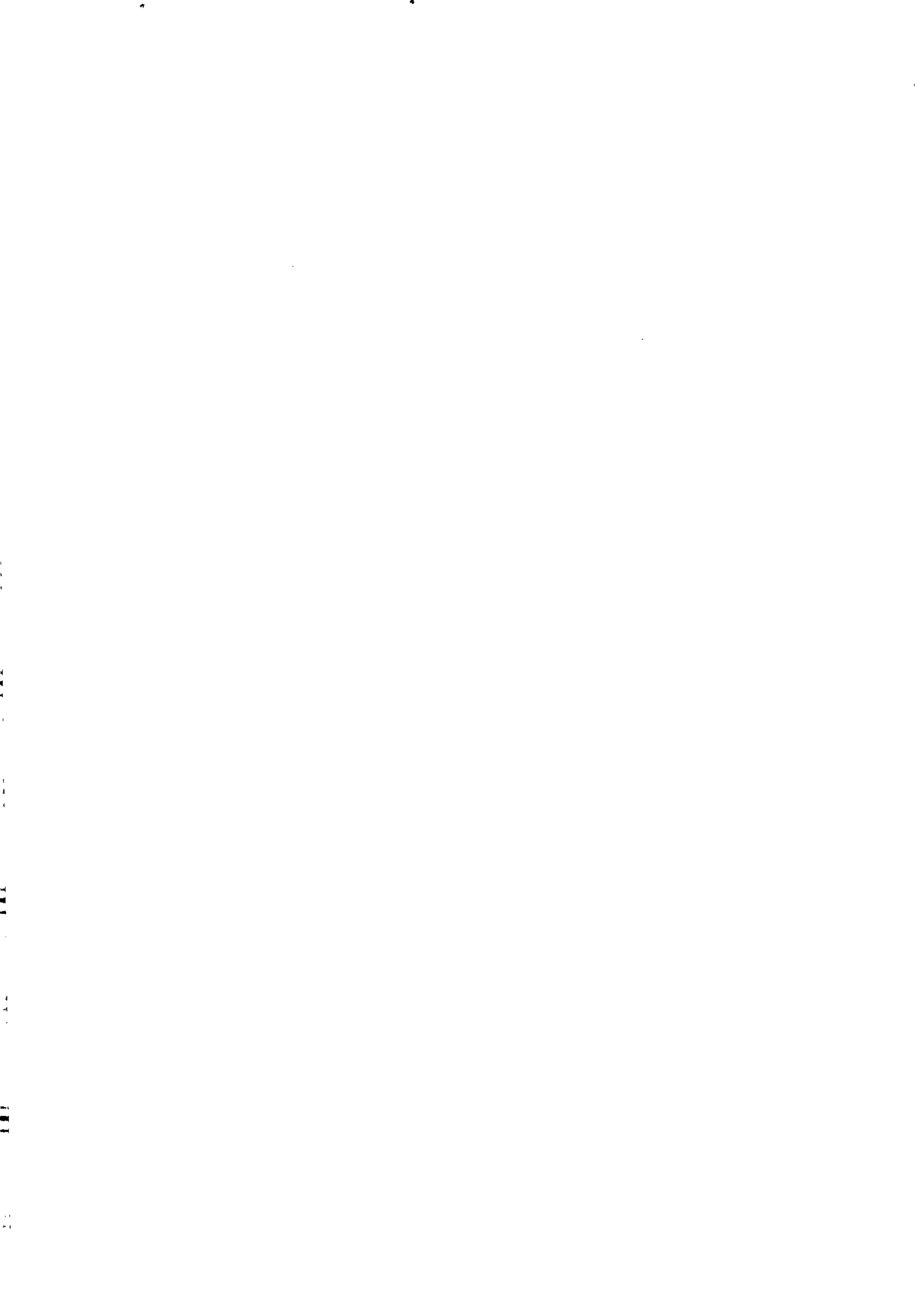
22. Timmins PA, Hauk J, Wacker T, Welte W: **The influence of heptane-1,1,3-triol on size and shape of LDAO micelles.** *FEBS Lett* 1991, **280**:115-120.

23. Gast P, Hemelrijk P, Hoff AJ: **Determination of the number of detergent molecules associated with the reaction center protein isolated from the photosynthetic bacterium *Rhodopseudomonas viridis*.** *FEBS Lett* 1994, **337**:39-42.

24. Kleymann G, Ostermeier C, Ludwig B, Skerra A, Michel H: **Engineered Fv fragments as a tool for the one-step purification of integral multisubunit membrane protein complexes.** *Biotechnology* 1995, **13**:155-160.

25. Landau EM, Rosenbusch JP: **Lipidic cubic phases: a novel concept for the crystallization of membrane proteins.** *Proc Natl Acad Sci USA* 1996, **93**:14532-14535.
••

Bacteriorhodopsin is crystallized using a system consisting of lipids, which form bicontinuous cubic bilayer phases, salts, water, and protein. Bacteriorhodopsin prepared in octylglucoside is incorporated into the lipids. The lipidic phase should act as a sink for the detergent molecules so that the bacteriorhodopsin molecules diffuse in the membrane system of the lipids without their detergent micelle. The lipid phases are suggested to provide nucleation sites and support crystal growth by lateral diffusion of the protein molecules in the bilayer membrane.



*trends in*

TIBTECH June 1997, Volume 15 No. 6 (161), pp. 197–237 ISSN 0167-7799

*Growing crystals in space*

plastics

# Recent advances in the microgravity crystallization of biological macromolecules

Microgravity experiments concerning the crystallization of biological macromolecules (reviewed in detail in Ref. 1) were initiated in the early to mid-1980s[2–5]. The hope was that crystals of improved quality for X-ray diffraction analyses might be obtained, like those seen in Fig. 1, as a consequence of the elimination of sedimentation and the convective flow inherent in crystallization solutions. Indeed, earlier experiments on conventional crystals, both on the ground and in space, had suggested that this might be the outcome[6–9].

Experiments since then have been dogged by persistent skepticism in a substantial segment of the crystallographic research community[10,11]. Doubt was appropriate and justified for several reasons for example, there was no persuasive explanation of why the elimination of convection and sedimentation should necessarily produce more-ordered or larger macromolecular crystals. Control experiments were difficult to conceive and execute on the ground. There were, at least initially, no accepted quantitative criteria for evaluating relative crystal quality, and many of the early experiments were impossible to reproduce and often poorly documented.

The obvious difficulties of carrying out sensitive, automated experiments in a space-flight environment were compounded by other factors – infrequent flight opportunities, few committed investigators, excessive expectations and overly enthusiastic appraisals based on qualitative data. Chronic technical limitations were imposed by the available instruments all being large in mass and volume, thereby limiting flight access, but at the same time executing relatively few crystallization experiments. While crystallization analyses in conventional laboratories might entail hundreds or even thousands of individual trials, the Cryostat[12] allowed only 14, and the most widely used device, the VDA[13], allowed at most 60 per unit.

Several years ago, investigations of macromolecular crystallization in microgravity split into two paths. The objective of the first was to produce high-quality crystals for biotechnology and research applications, such as X-ray diffraction analysis. The goal of the second line

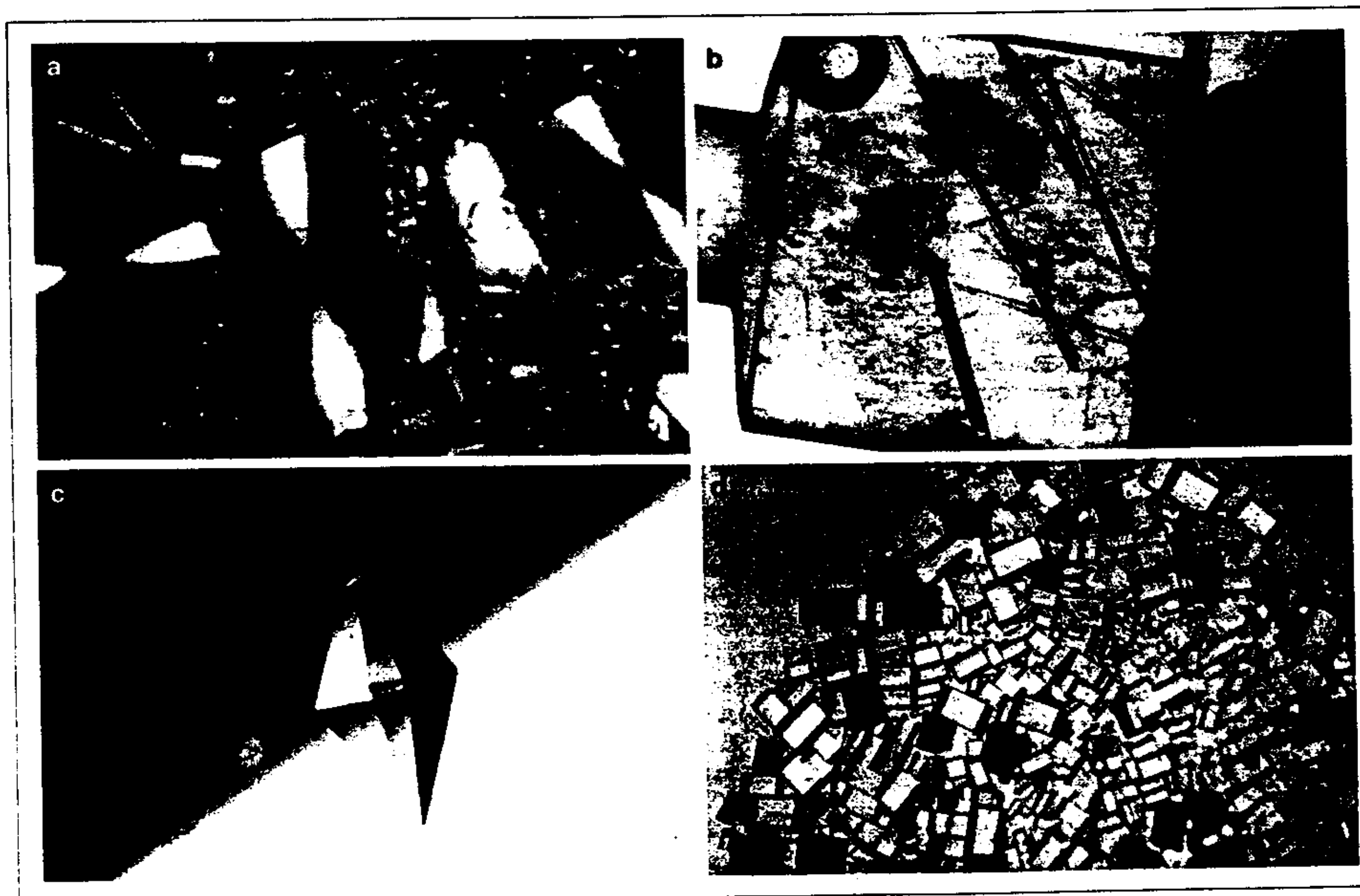



**Figure 1**

Examples of macromolecular crystals grown in space using four different crystallization devices. (**a**) A large orthorhombic crystal of satellite tobacco mosaic virus grown in the Cryostat[12] on the International Microgravity Laboratory-1. (**b**) Hexagonal crystals of the plant seed protein concanavalin B, nearly 5 mm in length, grown in the hand-held diffusion test cell on the US Microgravity Laboratory-2. (**c**) A single perfect crystal, about 1.5 mm from apex to apex, of the sweet protein thaumatin grown in the European-designed Advanced Protein Crystallization Facility[26]. (**d**) A mass of orthorhombic crystals of the seed storage protein canavalin grown in a Dewar by the flash-frozen, biphasic method on the Russian space station, Mir[29].

Copyright © 1997, Elsevier Science Ltd. All rights reserved. 0167 – 7799/97/$17.00. PII: S0167-7799(97)01041-X

of investigation was a definition and description, in a quantitative sense, of the mechanisms by which the quality of crystals was improved in microgravity. Understanding and, in the end, controlling the physics of the process was the objective. This second interest was ably supported by extensive ground-based research using a variety of sophisticated techniques such as static and quasielastic light scattering[14–18], interferometry[19,20], atomic force microscopy[21–25], and X-ray diffraction analyses[5,12,26,27].

### The superior quality of space-grown macromolecular crystals

The confluence of results from these two streams has significantly altered the prevailing circumstances and attitudes. Persuasive explanations for the observed improvements in size and quality of macromolecular crystals grown in microgravity have emerged (Fig. 2), and a convincing theoretical framework now exists for understanding the phenomena involved. Physical methods such as interferometry and atomic force microscopy have revealed the unsuspected variety, structure and number of dislocations and defects inherent in macromolecular crystals. These arrays of defects, which provide the key to the improvement attained in microgravity, have been shown to be far more complex, extensive and dense, by several orders of magnitude, than those commonly associated with conventional small-molecule crystals[21,29]. Thus macromolecular crystals are more sensitive to the unusually high degrees of supersaturation at which they are usually grown, and to the mass transport mechanisms responsible for bringing nutrient to their growing surfaces. The self-regulating nature of protein crystallization in microgravity, through the establishment of local concentration gradients of reduced supersaturation, explains why the diffusive transport that predominates in space produces a significant difference in ultimate crystal quality[1].

Another major development is that reliable, quantitative analyses of crystal quality based on X-ray diffraction properties (Fig. 3) have augmented and even replaced earlier subjective criteria for comparing crystals.

### The flash-frozen-samples–Dewar approach for multiple crystallizations

Other developments of a more practical nature are also beginning to

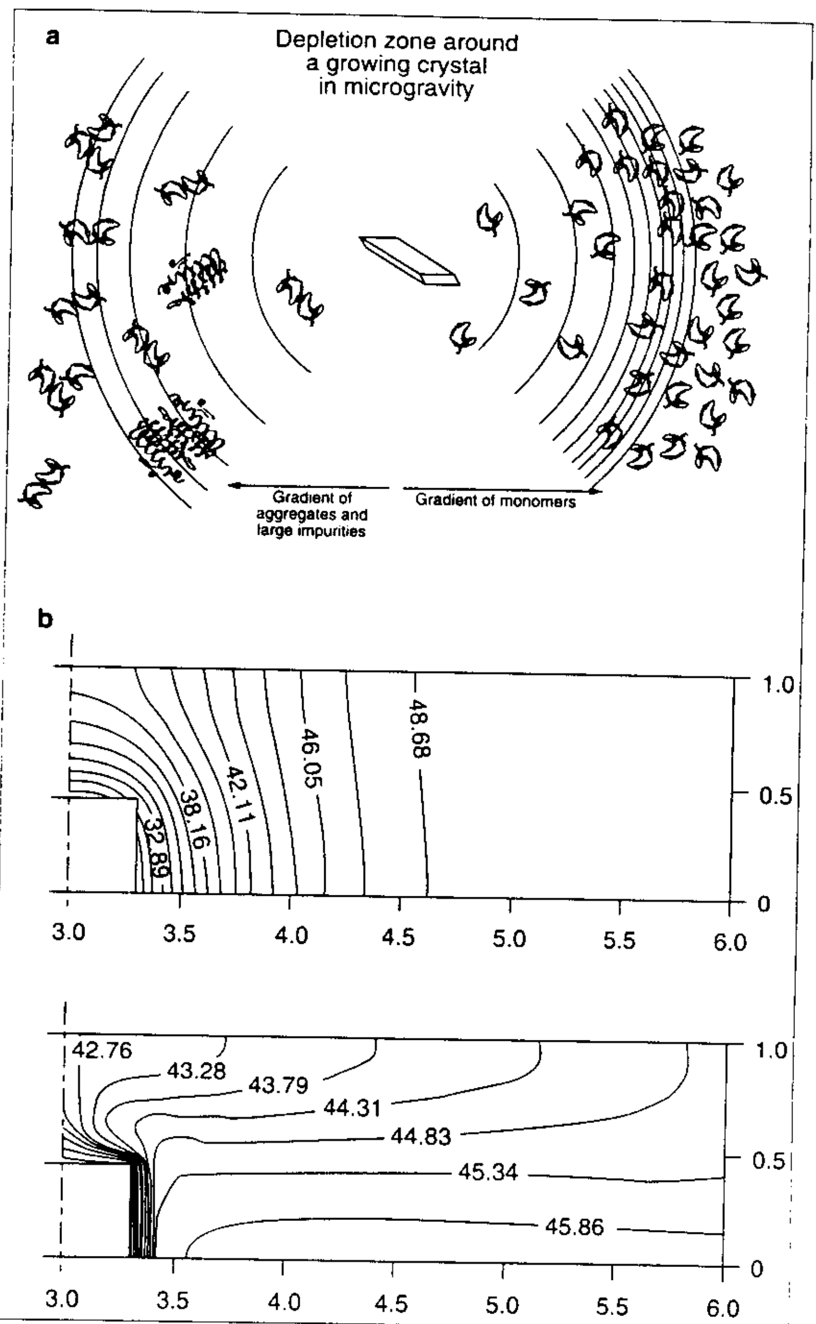



**Figure 2**

**(a)** The hypothesis underlying the improvement of macromolecular crystal growth in space. As a crystal grows from solution, recruitment of molecules from the surrounding medium produces a concentration gradient that extends from the surfaces. Thus the concentration of nutrient molecules near the crystal is less than that of the bulk solution and the crystal experiences an environment of reduced supersaturation. In microgravity, because there is no convective mixing, the depletion zones are quasistable. The nutrient molecules (on the right) diffuse very slowly because of their large size, thus prolonging and extending the effect. Large impurities, which include other proteins and aggregates of the nutrient molecules (on the left), diffuse even more slowly than monomers; thus the depletion zones serve as a kind of 'diffusion filter' to protect the growing crystal from incorporating some harmful impurities. **(b)** Two mathematical models by Rosenberger and colleagues[28] of the concentration gradients predicted to form around growing crystals in a microgravity environment. The plots show that there can be as much as a 40% reduction in macromolecule concentration between the crystal surfaces and the bulk solution.

increase the attractiveness of crystallization in space for the biotechnology community. Serious obstacles to the exploitation of space for commercially oriented macromolecular crystal growth have been the limited

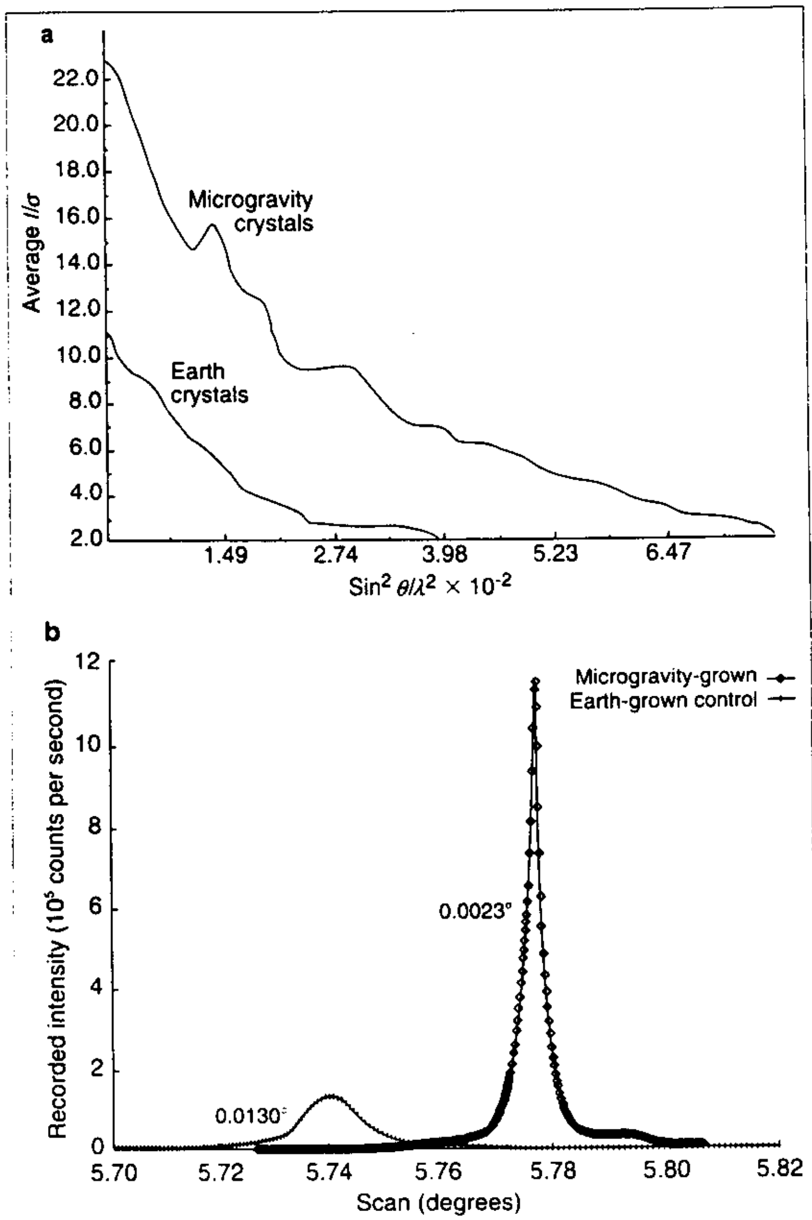



**Figure 3**

Two types of X-ray diffraction analysis that provide evidence for the superior quality of macromolecular crystals grown in a microgravity environment. (**a**) Intensity/sigma ($I/\sigma$) ratios for crystals of satellite tobacco mosaic virus grown in space and in a conventional laboratory on earth[12]. Favorable ratios for the crystals grown in space extend to significantly higher resolution and demonstrate a consistently higher value over the entire range, ultimately providing nearly 50% more X-ray diffraction intensities. Such comparisons have now been obtained for a number of different protein and virus crystals. (**b**) A comparison of the rocking curves for crystals of hen-egg-white lysozyme grown both on earth and in microgravity[27]. The much sharper peak (half-height widths for each peak are given on the graph) for the microgravity-grown crystal clearly indicates a substantial reduction in the mosaic spread, a highly desirable characteristic.

number of samples that could be accommodated on a mission, limitations on sample volumes and, therefore, numbers of crystals, and the expense involved. These impediments have prevented the screening of large matrices of conditions applied to large numbers of proteins to establish growth optima in microgravity. However, a novel approach that largely overcomes many of the problems of scale and economy has recently been described and shown to be successful. This relies on the use of flash-frozen crystallization samples carried into space in a Dewar charged with a small amount of liquid nitrogen.

Details of this methodology, and results of the experiments, have been presented elsewhere[30], but its simplicity, convenience and thrift deserve mention here. Samples of virtually any volume, from a few microliters to several milliliters, can be generated at any time before flight. They are composed of two phases, which were initially liquids – one the macromolecule solution and the other a precipitant solution – that have been sequentially injected or poured into a tygon tube and flash-frozen in liquid nitrogen so that a direct interface exists between the two frozen phases. In this form the inert samples can be stored indefinitely in a freezer until transferred to the flight Dewar and subsequently transported, still frozen, into space. Upon thawing in the microgravity environment, a direct liquid–liquid interface is re-established between the two phases, diffusion of precipitant into the protein solution occurs, and crystallization proceeds by free interface diffusion.

The Dewar has no moving parts, hence no electronic or mechanical overhead, and its entire volume can be used to hold samples. No astronaut intervention is required, and the process is entirely passive. For screening purposes, using 20–40 μl samples in small-bore tygon tubes, about 15 000 different crystallization samples could be carried in the standard NASA Dewar (GN/2) now in use. As described elsewhere[30], the method is amenable to a variety of gradient or multiphase sample configurations in addition to the simple biphasic mode.

The flash-frozen-samples–Dewar experiments have, to this point, been carried out only on the Russian space station, Mir. Dewar units filled with samples are carried to Mir on each space-shuttle–Mir rendezvous mission and exchanged for one from the previous mission. Thus experiments have lasted between 3 and 5 months. In the future, however, Dewars with thermal control will be flown, and these can be employed on shorter (8 to 16 day) US shuttle missions. Dewar-based systems are also likely to be available on the International Space Station beginning in about 2002. In the meantime, the principal investigators invite any interested scientists, from the public or private sector, to provide samples. For at least the immediate future these can be flown at no cost to the co-participants, following toxicological

approval by NASA, whenever flight opportunities arise.

## The future

The size of the research community has increased, and with it the diversity of ideas. Opportunities for flight experiments have expanded, along with the variety of instruments with which to conduct crystallization experiments (also reviewed in Ref. 1).

There are currently a number of powerful observational systems under development in the USA, Europe and Japan. These will be deployed on the International Space Station, where they will form the core facilities for the investigation of macromolecular crystallization in space. Those studies will extend and refine our understanding of the physical principles governing microgravity crystal growth and will help identify the properties of macromolecules likely to benefit most from crystallization in microgravity.

## References

1 McPherson, A. (1996) *Crystallogr. Rev.* 6, 157–308
2 Littke, W. and John, C. (1984) *Science* 225, 203–204
3 Littke, W. and John, C. (1986) *J. Cryst. Growth* 76, 663–672
4 DeLucas, L. J. *et al.* (1986) *J. Cryst. Growth* 76, 681–693
5 DeLucas, L. J. *et al.* (1989) *Science* 246, 651–654
6 Yoo, H-D., Wilson, W. R., Lal, R. and Troliner, J. D. (1988) *J. Cryst. Growth* 92, 101–117
7 Wilcox, W. R. (1983) *J. Cryst. Growth* 65, 133
8 Rindone, G. E., ed. (1982) *Materials Processing in the Reduced Gravity Environment of Space*, North-Holland
9 Kroes, R. L. and Reiss, D. (1984) *J. Cryst. Growth* 69, 414–420
10 Kolata, G. (1985) *Science* 229, 370
11 Leberman, R. (1985) *Science* 230, 373
12 Day, J. and McPherson, A. (1992) *Protein Sci.* 1, 1254–1268
13 DeLucas, L. J. *et al.* (1994) *J. Cryst. Growth* 135, 183–195
14 Mikol, V., Vincendon, P., Eriani, G. and Giegé, R. (1991) *J. Cryst. Growth* 110, 195–200
15 Wilson, W. W. (1990) in *Methods: a Companion to Methods in Enzymology* (Vol. 1), pp. 110–117, Academic Press
16 Malkin, A. and McPherson, A. (1994) *Acta Crystallogr. D* 50, 385–395
17 Malkin, A. J., Cheung, J. and McPherson, A. (1993) *J. Cryst. Growth* 126, 544–555
18 Mikol, V., Hirsh, E. and Giegé, R. (1990) *J. Mol. Biol.* 213, 187–195
19 Kuznetsov, Y. G., Malkin, A. J., Greenwood, A. and McPherson, A. (1995) *J. Cryst. Growth* 166, 913–918
20 Vekilov, P. G., Ataka, M. and Katsura, T. (1992) *J. Cryst. Growth* 130, 317–320
21 Hillner, P. E., Manne, S., Hansma, P. K. and Gratz, A. J. (1993) *Faraday Discuss.* 95, 191–197
22 Durbin, S. D., Carlson, W. E. and Saros, M. T. (1993) *J. Phys. D* 26, B128–B132
23 Konnert, J. H., D'Antonio, P. and Ward, K. B. (1994) *Acta Crystallogr. D* 50, 603–613
24 McPherson, A., Malkin, A., Kuznetsov, Y. G. and Koszelak, S. (1995) *J. Cryst. Growth* 168, 74–92
25 Malkin, A. J., Kuznetsov, Y. G., Land, T. A., DeYoreo, J. J. and McPherson, A. (1995) *Nat. Struct. Biol.* 2, 956–959
26 Koszelak, S., Day, J., Leja, C., Cudney, R. and McPherson, A. (1995) *Biophys. J.* 69, 13–19
27 Snell, E. H., Weisgerber, S. and Helliwell, J. R. (1995) *Acta Crystalllogr. D* 51, 1099–1102
28 Lin, H., Rosenberger, F., Alexander, J. I. D. and Nadarajah, A. (1995) *J. Cryst. Growth* 151, 153–162
29 Malkin, A. J., Kuznetsov, Y. G. and McPherson, A. (1996) *J. Struct. Biol.* 117, 124–137
30 Koszelak, S., Leja, C. and McPherson, A. (1996) *Biotechnol. Bioeng.* 52, 449–458

**Alexander McPherson**

*University of California, Riverside, Department of Biochemistry, Riverside, CA 92521, USA.*

*Meeting report*

# The dawn of the post-genome era, seen from the ocean front

The Human Genome Project is in the sixth of its projected 15 years. The large-scale production of DNA sequence data from human, nematode, microbial and plant genomes is in full swing. New strategies of whole-genome sequencing are emerging, and the challenges and opportunities of the post-genome era are unfolding. The landscape of gene expression and functional studies is being rapidly changed by the newly developed technologies. These are the messages conveyed from the recent, and largest, annual genome-sequencing-community gathering, the Eighth International Genome Sequencing and Analysis Conference, held at the picturesque ocean front of Hilton Head Island, South Carolina, USA during the stormy days of 5–8 October 1996[1]. One indication of the ever-expanding effort and interest is the size of the conference. This year there were over 900 participants, over 50% more than last year. Heavy commercial sponsorship and vendor presence has been one of the distinct features of this conference, illustrating the interest and involvement of industry in genome research.

## Production-sequencing progress

After the much-publicized advocation of production sequencing by several key players in the genome community, including Bob Waterston and John Sulston, 1996 was the first year that the Genome Project went into large-scale sequencing-data production. Several genome projects have made significant headway. The *Caenorhabditis elegans* genome project conducted by the Sanger Centre (Hinxton, UK) and the Washington University Genome Center (St. Louis, USA) is still leading in data production, with 58 Mb of the 100 Mb genome sequenced. Nearly 2 Mb of sequence from *Drosophila* chromosome 2 have been completed by the Lawrence Berkeley Genome Center (Berkeley, USA). However, most of these sequenced regions still contain a significant number of gaps, and completion remains a major undertaking.

Certainly, the main thrust is still in the human genome, and there were relevant presentations from a number of groups. The Sanger Centre has been sequencing regions on chromosomes 22 (in collaboration with the Washington University Genome Center), 6 and X (web site: http://www.sanger.ac.uk). The Washington University Genome Center has produced 380 000 expressed sequence tags (ESTs), nearly 200 000 of them unique, and 24 Mb of genomic sequence from chromosomes 22, 7, 5 and X



Copyright © 1997, Elsevier Science Ltd. All rights reserved. 0167 – 7799/97/$17.00. PII: S0167-7799(97)01046-9