International Tables for Crystallography (2012). Vol. F, Section 4.1.1, pp. 99–100.

PART 4. CRYSTALLIZATION

Chapter 4.1. General methods

C. Sauter, B. Lorber, A. McPherson and R. Giegé

4.1.1. Introduction

4.1.1.1. Prologue

Macromolecular crystals are much softer than salt crystals. They contain ∼50% of solvent on average, with values ranging from as little as 25 up to 90%. A direct consequence is that, although morphologically indistinguishable, crystals of macromolecules differ in many respects from crystals of low-molecular-mass compounds. While the latter exhibit firm lattice forces, are highly ordered, generally physically hard and brittle, easy to manipulate, can usually be exposed to air, have strong optical properties and diffract X-rays intensely, crystals of macromolecules are, by comparison, smaller in size, they crush easily, disintegrate if allowed to dehydrate, exhibit weak optical properties and diffract X-rays poorly. They are temperature sensitive and undergo extensive damage after prolonged exposure to radiation.

Proteins or nucleic acids build up a crystalline scaffold, which may be imagined as an ordered gel with extensive interstitial spaces through which small molecules can diffuse freely. In proportion to molecular mass, large macromolecules establish far fewer packing interactions than do small molecules inside crystalline lattices. Since these contacts are responsible for the integrity of the crystals, this largely explains the differences in properties between the two types of crystals. Thus, liquid channels and solvent cavities are directly responsible for the generally poor diffraction properties of macromolecular crystals. Owing to the large spaces between adjacent molecules and the related weak lattice forces, every molecule in the crystal may not occupy exactly equivalent orientations and positions. Furthermore, because of their structural complexity and their conformational dynamics, macromolecules in a given crystal form may exhibit slight variations in their folding patterns or dispositions of side chains.

However, high solvent content is not as negative as it might appear at first glance. It allows maintenance of the macromolecular structures virtually unchanged from those 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 dimensions of solvent channels are such that conventional chemical compounds, such as ions and heavy atoms, substrates or other ligands, may be freely diffused into and out of the crystals. Thus, many crystalline enzymes, though immobilized, are completely accessible for experimentation through alteration of the surrounding mother liquor.

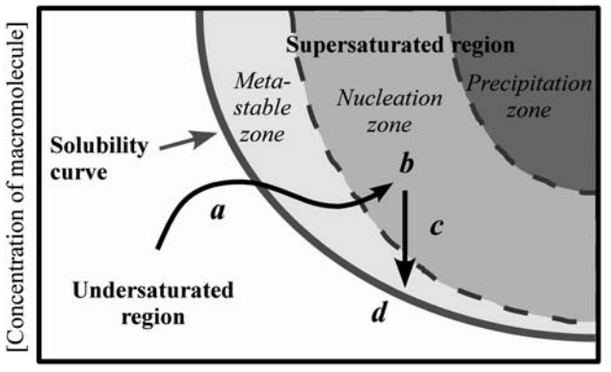
The intrinsic instability of most macromolecules requires that conditions suitable for crystal growth are those that do not perturb their molecular properties. This explains why crystals must be grown from solutions compatible with the target macromolecules, *i.e.* within a narrow range of pH, temperature or ionic strength. Finally, because hydration is essential for the maintenance of structure, crystals of macromolecules must be always bathed in the mother liquor, even during data collection (except in the practice of cryocrystallography).

4.1.1.2. Crystallization principles

The crystallization of biological macromolecules has often been considered unpredictable, although it obeys the same principles as that of small molecules (Giegé *et al.*, 1995; McPherson *et al.*, 1995; Rosenberger, 1996; Chernov, 2003; McPherson & Giegé, 2007). It is, similarly, a multiparametric process. The difference compared with conventional crystal growth arises from the biochemical and biophysical properties of proteins or nucleic acids, and crystallization methods must take into account these features. The methods described below apply for most proteins, large RNAs, multimacromolecular complexes and viruses (for small oligonucleotides or peptides, crystallization by dialysis is not appropriate). For hydrophobic membrane proteins, special techniques are required (see Chapter 4.2).

Crystallization proceeds from macromolecules in solution that 'aggregate' upon entering a supersaturated state and eventually undergo a phase transition. This leads to nuclei formation and ultimately to crystals that grow by different mechanisms. Supersaturation is the driving force of crystallization and is defined as the ratio $[C]/[S]$, where $[C]$ and $[S]$ are the initial concentration of the macromolecule and its final concentration at saturation, *i.e.* its solubility. Nucleation is homogeneous when nuclei form in the bulk of the solution, but heterogeneous when they preferentially form on walls of crystallization vessels, on solid particles (dust, aggregates, seeds), or on the surface of existing crystals. Unlike most conventional crystals, protein crystals are, in general, not initiated from seeds, but are nucleated *ab initio* at high levels of supersaturation that can reach 200 to 1000% (in what follows and for simplicity, the generic name 'protein' is used for macromolecule). It is this high degree of supersaturation that, in large part, distinguishes protein crystal formation from that of conventional crystals. That is, once a stable nucleus has formed, it subsequently grows under very unfavourable conditions of excessive supersaturation. Distant from the metastable zone, where ordered growth could occur, crystals rapidly accumulate nutrient molecules, as well as impurities. They also concomitantly accumulate statistical disorder and a high frequency of defects that exceeds those observed for most conventional crystals.

The different stages of crystallization (*i.e.* pre-nucleation, nucleation, growth, cessation of growth) can be visualized in a phase diagram (Fig. 4.1.1.1). In short, phase diagrams are divided into undersaturated regions (where proteins are soluble) and supersaturated regions (where protein crystals nucleate and grow) delimited by the solubility curve. The supersaturated region is thermodynamically out of equilibrium and can be divided into three kinetically dependent domains: a precipitation domain (at extreme supersaturation) where macromolecules rapidly separate from solution in a solid state either amorphous or microcrystalline, a domain (at lower supersaturation) where



Biochemical parameters
Purity of the macromolecule
Ligands, inhibitors, effectors
Aggregation state of the macromolecule
Post-translational modifications
Source of macromolecule
Proteolysis or hydrolysis
Chemical modifications
Genetic variants
Inherent symmetry of the macromolecule
Stability of the macromolecule
Isoelectric point
History of the sample

Chemical parameters
pH
Crystallant type
Crystallant concentration
Ionic strength
Specific ion effects
Supersaturation
Reductive or oxidative environment
Concentration of the macromolecule
Metal ions
Crosslinkers or polyions
Detergents, surfactants or amphophiles
Non-macromolecular impurities

Physical parameters
Temperature variation
Contact surfaces
Methodology or approach to equilibrium
Gravity
Pressure
Time
Vibrations, sound or mechanical perturbations
Electrostatic or magnetic fields
Dielectric properties of the medium
Viscosity of the medium
Rate of equilibration
Homogeneous or heterogeneous nucleants

Figure 4.1.1.1
Crystallization is a multiparametric process under the control of a great variety of biochemical, chemical and physical parameters. The crystal grower can play with them to drive a macromolecule from the soluble towards the crystalline state. In the top left, a typical two-dimensional phase diagram illustrates how this can be achieved. (*a*) By modifying two parameters (concentrations of macromolecule and crystallant), the system will move from the undersaturated region where the macromolecule is soluble into the supersaturated region beyond the solubility curve where it will try to escape from the solution. The prenucleation trajectory essentially depends on the crystallization method (see Fig. 4.1.3.1). (*b*) As soon as they enter the supersaturation region, macromolecules will tend to aggregate. However, the system needs to cross an energy barrier to produce stable molecular assemblies, the nuclei, and this only happens in the so-called nucleation zone. (*c*) When stable nuclei are formed, they will capture more macromolecular entities from the mother liquor and produce three-dimensional crystals. (*d*) Crystals will grow until the system comes back to the solubility curve, crossing the metastable zone. Growth stops and crystals are in dynamic equilibrium with the mother liquor. When the system is driven to high supersaturation, macromolecules may rather produce amorphous or microcrystalline precipitates than useful monocrystals. Note that the limits between zones of the supersaturated zone (dashed curve) move with time.

nucleation occurs spontaneously and a metastable domain (at low supersaturation) where nucleation does not occur spontaneously but where crystals grow. This domain is favourable for seeding. The wisdom of the crystal grower will be to take advantage of an overall understanding of phase diagrams for designing crystallization strategies and selecting favourable solvent conditions (Sauter, Lorber *et al.*, 1999; Asherie, 2004).

4.1.2. Main parameters that affect crystallization of macromolecules

4.1.2.1. Crystallizing agents

Crystallizing agents (also referred to as crystallants or precipitants) are defined as those chemical species that are able to precipitate a soluble macromolecule and hence to bring it into a supersaturated state needed for its crystallization. They fall into four categories: (i) salts, (ii) organic solvents, (iii) long-chain polymers, and (iv) low-molecular-mass polymers and non-volatile organic compounds (McPherson, 1990). The first two classes are typified by ammonium sulfate and ethanol; higher polymers such as PEGs (polyethylene glycols) of molecular weight (M_r) > 1000, are characteristic of the third. In the fourth are placed compounds such as MPD (2-methyl-2,4-pentanediol) and low-molecular-mass PEGs.

Salts. Salts exert their effects by dehydrating proteins through competition for water molecules (Green & Hughes, 1955). Their ability to do this is roughly proportional to the square of the valences of the ionic species composing the salt. Thus, multivalent ions, particularly anions, are the most efficient. One might think there would be little variation between different salts, so long as their ionic valences were the same, or between two different sulfates, such as lithium and ammonium sulfates. This, however, is often not the case. In addition to salting out, which is a dehydration effect, or lowering of the chemical activity of water, there are specific protein–ion interactions that have other consequences (Riès-Kautt & Ducruix, 1991; Ducruix & Giegé, 1999). This is true because of the polyvalent character of individual proteins, their structural complexity, and the dependence of their physical properties on environmental conditions and interacting molecules. Therefore, restricting the search to one or two salts is never sufficient when attempting to crystallize a protein. Changing the salt may produce crystals of varied quality, morphology and diffraction properties.

It is usually not possible to predict the molarity of a salt required to crystallize a particular protein without some prior knowledge of its behaviour. In general, the concentration is just a small percentage less than that yielding an amorphous precipitate. To determine the precipitation point with a given crystallant, a 10 µl droplet of a 5–15 mg ml^{-1} protein solution is placed in the well of a depression slide and observed under a microscope as increasing amounts of salt solution or organic solvent (in l–2 µl increments) are added. If the well is sealed between additions with a coverslip, the increases can be made over a period of many hours. Indeed, the droplet should equilibrate 10–30 min after each addition, and longer in the neighbourhood of the precipitation point.

Widely used and the most successful salt to grow protein crystals is ammonium sulfate. A variety of other salts included in crystallization screens yield success as well. Sodium malonate, not systematically present in screens, deserves particular attention, since it gave the highest success rate (better than ammonium sulfate) in a comparative study designed to identify the most effective salts for protein crystallization (McPherson, 2001). Following this example, other organic salts have been added to the repertoire (McPherson & Cudney, 2006).

Organic solvents. The most common organic solvents (volatile compounds) utilized are ethanol, methanol, isopropanol, acetone and dioxane. They have been frequently used for crystallizing nucleic acids, particularly tRNAs and duplex oligonucleotides