to uncouple nucleation and growth. Seeding techniques fall into two categories that employ either macroseeds (Thaller *et al.*, 1985) or microcrystals as seeds (Stura *et al.*, 1999; Bergfors, 2003). In both cases, the solution to be seeded should be only slightly supersaturated so that controlled growth can occur.

When seeding with crystals large enough to be manipulated under a microscope, the most important consideration is to eliminate spurious nucleation by transfer of too many seeds. This drawback may be overcome using laser tweezers, a technique that permits non-mechanical, *in situ* manipulation of individual seeds as small as 1 μm (Bancel *et al.*, 1998). Even if a single large crystal is employed, microcrystals adhering to its surface may be carried across to the fresh solution. To avoid this, the macroseed is washed by passing it through a series of intermediate transfer solutions. In doing so, not only are microcrystals removed, but if the wash solutions are chosen properly, some limited dissolution of the seed surface may take place. This has the effect of freshening the seed-crystal surfaces and promoting new growth once it is introduced into the new protein solution. Note that crystals of homologous macromolecules can serve as seeds (Thaller *et al.*, 1985).

In the second approach with microcrystals, the danger is that too many nuclei will be introduced into the fresh supersaturated solution, and masses of crystals will result. 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 ml; others will have several times more, or none at all. An aliquot (∼1 μl) of each sample in the series is then added to fresh crystallization trials. This empirical test, ideally, identifies the correct sample to use for seeding by yielding only one or a small number of single crystals when crystal growth is completed. Microseeds can be introduced into crystallization trials at any stage of microbatch or vapour-diffusion experiments (D'Arcy, MacSweeney & Haber, 2003; D'Arcy *et al.*, 2004) and this process can be automated (D'Arcy *et al.*, 2007; Newman *et al.*, 2008; Khurshid *et al.*, 2010).

4.1.5. From the macromolecule to perfect crystals: the physics view

Each of the four stages in crystallization (prenucleation, nucleation, growth and cessation of growth) can be monitored by specific physical techniques. Although systematic characterization of crystallization is usually not carried out in practice, characterization of individual steps and measurement of the physical properties of crystals obtained under various conditions may help in the design of appropriate experimental conditions to reproducibly obtain crystals of a desired quality (*e.g.* of larger size, improved morphology, increased resolution or greater perfection).

4.1.5.1. Prenucleation and nucleation

DLS relies on the scattering of monochromatic light by aggregates or particles moving in solution. Since diffusivity of the particles is related to their size by the Stokes–Einstein equation, measurement of diffusion coefficients can be translated into hydrodynamic radii. By making measurements as a function of scattering angle, information regarding aggregate shape can also be obtained. For single-component systems, the method for determining the size of macromolecules, viruses and larger particles up to a few μm is straightforward. For polydisperse and concentrated systems, the problem is more complex, but with the use of auto-correlation functions and advances in signal detection, DLS provides good estimates of aggregate size distribution and supplies a diagnostic tool for quality control of proteins and optimization of crystallization conditions. Experimental necessity encouraged the design of dedicated instruments for protein crystallization, combining *e.g.* imaging of crystals and DLS analysis within crystallization droplets (Dierks *et al.*, 2008).

In biocrystallogenesis, investigations based on light scattering have been useful in detecting nucleation prior to the appearance of crystals observable under the light microscope, that is, in understanding prenucleation and nucleation. Many studies have been carried out with lysozyme as the model (Kam *et al.*, 1978; Durbin & Feher, 1996), though not exclusively, and they have been developed with two objectives. One is to analyse the kinetics and the distribution of molecular aggregate sizes as a function of supersaturation. The idea is to understand the nature of prenuclear clusters that form in solution and how they transform into crystal nuclei (Kam *et al.*, 1978; Georgalis *et al.*, 1993; Malkin & McPherson, 1994). Such a quantitative approach has sought to define the underlying kinetic and thermodynamic parameters that govern the nucleation process.

A more practical objective is to use light-scattering methods to predict which combinations of crystallants, additives and physical parameters are most likely to lead to the nucleation and growth of crystals (Mikol, Hirsch & Giegé, 1990; Ferré-D'Amaré & Burley, 1997; Borgstahl, 2007; Wilson, 2003; Niesen *et al.*, 2008). A major goal here is to reduce the number of empirical trials. The analyses depend on the likelihood that precipitates are usually linear, branched and extended in shape, since they represent a kind of random polymerization process (Kam *et al.*, 1978). Aggregates leading to nuclei, on the other hand, tend to be more globular and three dimensional in form. Thus, mother liquors that indicate a nascent precipitate can be identified as a failure, while those that have the character of globular aggregates hold promise for further exploration and refinement. Other analyses have been based on discrimination between polydisperse and monodisperse protein solutions, which suggests that polydispersity hampers crystallization, while monodispersity favours it (Mikol, Hirsch & Giegé, 1990).

A more quantitative approach is based on measurement of the second virial coefficient B2, which serves as a predictor of the type of interaction between macromolecules occurring in solution. Using static light scattering, it was found that mother liquors that invariably yield crystals have second virial coefficients that fall within a narrow range of small negative values. Correlations between the associative properties of proteins in solution, their solubility and the B2 coefficient were highlighted (*e.g.* George *et al.*, 1997; Wilson, 2003), and seem to be a general feature. This is a powerful diagnostic of crystallization conditions.

Related methods, such as fluorescence spectroscopy (Crosio & Jullien, 1992; Forsythe *et al.*, 2006), osmotic pressure (Bonneté *et al.*, 1997; Neal *et al.*, 1999), small-angle X-ray scattering (Finet *et al.*, 1998) and small-angle neutron scattering (Ebel *et al.*, 1999; Gripon *et al.*, 1997; Minezaki *et al.*, 1996; Vidal *et al.*, 1998), were used to investigate specific aspects of protein interactions under precrystallization conditions and produced, in several instances, complementary answers to those from light-scattering studies.

4.1.5.2. Growth and cessation of growth

A number of microscopies and other optical methods can be used for studying the crystal growth of proteins (Van Driessche, Otalora, Sazaki *et al.*, 2008). These are time-lapse video microscopy with polarized light, Schlieren and phase-contrast micro-

scopy, Mach–Zehnder and phase-shift Mach–Zehnder interferometry, Michelson interferometry, electron microscopy (EM), atomic force microscopy (AFM) and laser confocal microscopy with differential interference contrast microscopy. Each of these methods provides complementary data which, in combination, have yielded answers to many relevant questions.

Time-lapse video microscopy has been used to measure growth rates (Koszelak & McPherson, 1988; Lorber & Giegé, 1992; Zhu *et al.*, 2001). It was valuable in revealing unexpected phenomena, such as capture and incorporation of microcrystals by larger crystals, contact effects, consequences of sedimentation, flexibility of thin crystals, fluctuations in growth rates and initiation of twinning (Koszelak *et al.*, 1991). Optical microscopy and interferometric methods gave information on concentration gradients that appear as a consequence of incorporation of molecules into the solid state. These methods, however, suffer from rather shallow response dependence with respect to protein concentration. This can be overcome by the introduction of phase-shift methods, as has been successfully achieved in the case of Mach–Zehnder interferometry. With this technique, gradients of protein concentration have been mapped in the mother liquor and around growing crystals. Classical Mach–Zehnder interferometry has been used to monitor diffusion kinetics and supersaturation levels during crystallization, as was done in dialysis setups (Snell *et al.*, 1996) or in counter-diffusion crystal growth cells (Garcia-Ruiz *et al.*, 1999). Using laser confocal microscopy combined with differential interference contrast microscopy it was possible to visualize dislocations in protein crystals during growth (Sazaki *et al.*, 2005). Single-molecule visualization techniques gave access to direct observation of the diffusion of individual fluorescence-labelled protein molecules at an interface between a solution and a protein crystal (Sazaki *et al.*, 2008).

Michelson interferometry can be used for direct growth measurements on crystal surfaces (Komatsu *et al.*, 1993). It depends on the interference of light waves from the bottom surface of a crystal growing from a reflective substrate and the top surface, which is developing and, therefore, changes as a function of time with regard to its topological features. Because growth of a crystal surface is generally dominated by unique growth centres produced by dislocations or two-dimensional nuclei, the surfaces and the resultant interferograms change in a regular and periodic manner. Changes in the interferometric fringes with time provide accurate measures of the tangential and normal growth rates of a crystal (Vekilov *et al.*, 1993; Kuznetsov *et al.*, 1995; Kurihara *et al.*, 1996). From these data, one can determine the surface free energy and the kinetic coefficients that underlie the crystallization process.

EM (Durbin & Feher, 1990) and especially AFM (McPherson *et al.*, 2004) are powerful techniques for the investigation of crystallization mechanisms and their associated kinetics. The power of AFM lies in its ability to investigate crystal surfaces *in situ*, while they are still developing, thus permitting one to directly visualize, over time, the growth and change of a crystal face at near-nanometre resolution. The method is particularly useful in delineating the growth mechanisms involved, identifying dislocations, recording the kinetics of the changes and directly revealing impurity effects on the growth of protein crystals (Konnert *et al.*, 1994; Malkin *et al.*, 1996; Nakada *et al.*, 1999) (Fig. 4.1.5.1). AFM was also applied for the visualization of growth characteristics of crystals made of viruses (Malkin *et al.*, 1995) and RNA (Ng, Kuznetsov *et al.*, 1997). A noteworthy outcome of such studies was the sensitivity of growth to minor temperature changes. A variation of 2–3 °C was sufficient to transform the growth mechanism of yeast $tRNA^{Phe}$ from spiral screw dislocation growth at low supersaturation to two-dimensional island formation at high supersaturation (Ng, Kuznetsov *et al.*, 1997).

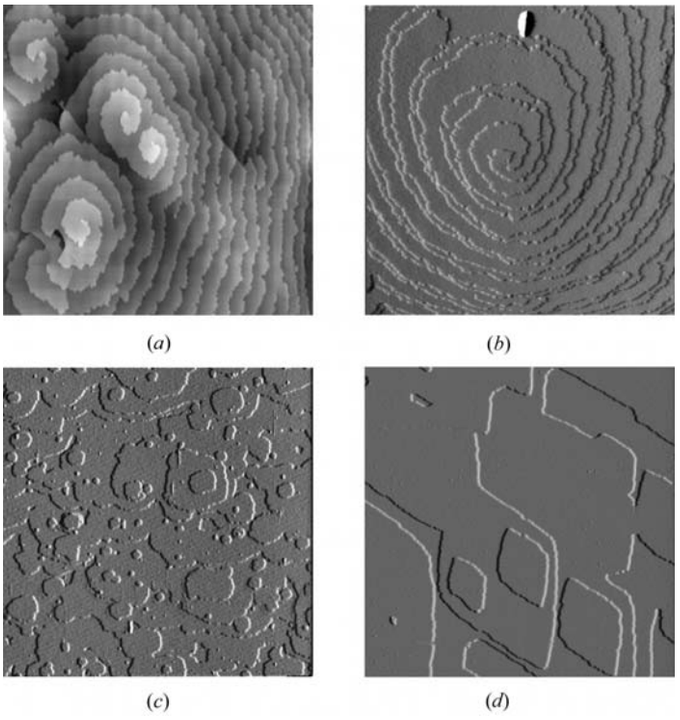

Figure 4.1.5.1
Growth mechanisms and visualization of protein crystal surfaces by AFM. In (*a*) and (*b*) are images of screw dislocations on the surfaces of crystals of the proteins canavalin and trypsin, respectively. The scan areas are 10 μm^2 in (*a*) and 30 μm^2 in (*b*). Screw dislocation growth predominates at low supersaturation. In (*c*) and (*d*) are examples of crystal growth by the formation of two-dimensional islands on the surfaces of crystals of the proteins thaumatin and glucose isomerase, respectively. The scan areas are 20 μm^2 in (*c*) and 11 μm^2 in (*d*). Growth by two-dimensional island formation and spread dominates at higher supersaturation.

4.1.5.3. Uncoupling nucleation and growth, and the constant-growth regime

The preparation of high-quality protein crystals should preferentially occur at lowest supersaturation and under a constant-growth regime. Achieving this aim with conventional crystallization methods is *a priori* not easy since growth of crystals is accompanied by a decrease of supersaturation in the crystallization medium (see the trajectory of crystal/solution equilibration in a phase diagram, Figs. 4.1.1.1 and 4.1.3.1). The implication is that it is possible to change the growth regime during the course of the crystallization process, as could be seen by AFM (Ng, Kuznetsov *et al.*, 1997). Such a change will perturb crystal formation and probably accounts for the frequently observed non-reproducibility in diffraction properties of protein crystals. Using flow cells with a constant supply of fresh protein may help to overcome this difficulty. Separating the nucleation and growth phases is another alternative (Chayen, 2005). This can be straightforwardly done by seeding procedures (see Section 4.1.4.4.).

4.1.5.4. Crystal perfection

The ultimate objective of structural biologists is to analyse crystals of high perfection, in other words, with a minimum of

defects, disorder, impurity incorporation and internal stress. Such imperfections can be visualized using laser confocal microscopy combined with differential interference contrast microscopy (Iimura *et al.*, 2005; Sazaki *et al.*, 2005). They can also be evaluated by the resolution limit of diffraction, which expresses the average disorder of the molecules in the crystal lattice. Wilson plots provide good illustrations of diffraction quality for protein crystals. Other sources of disorder, such as the mosaic structure of the crystal, may strongly influence the quality of the diffraction data. They are responsible for increases in the diffuse background scatter and a broadening of diffraction intensities. These defects are difficult to monitor with precision, and dedicated techniques and instruments are required for accurate analysis.

Mosaicity can be defined experimentally by X-ray rocking width measurements. An overall diagnostic of crystal quality can be obtained by X-ray diffraction topography. Both techniques have been refined with lysozyme as a test case and were used for comparative analysis of crystals grown under different conditions, both on earth and in microgravity. For lysozyme and thaumatin, improvement of the mosaicity, as revealed by decreased rocking widths measured with synchrotron radiation, was observed for the microgravity-grown crystals (Snell *et al.*, 1995; Ng, Lorber *et al.*, 1997; Lorber, Sauter, Robert *et al.*, 1999).

An illustration of mosaic block character in a lysozyme crystal was provided by X-ray topography (Fourme *et al.*, 1995). Comparison of earth- and microgravity-grown lysozyme crystals showed a high density of defects in the earth control crystals, while in the microgravity case, several discrete regions were visible (Stojanoff *et al.*, 1996). X-ray topographs have also been used to compare crystal polymorphs (Izumi *et al.*, 1996), to monitor temperature-controlled growth of tetragonal lysozyme crystals (Stojanoff *et al.*, 1997), to study the effects of solution variations during growth on crystal perfection (Dobrianov *et al.*, 1998; Otálora *et al.*, 1999), to compare crystals grown in solution and in agarose gel (Lorber, Sauter, Ng *et al.*, 1999), and to map defects in the bulk of protein crystals (Hu *et al.*, 2001).

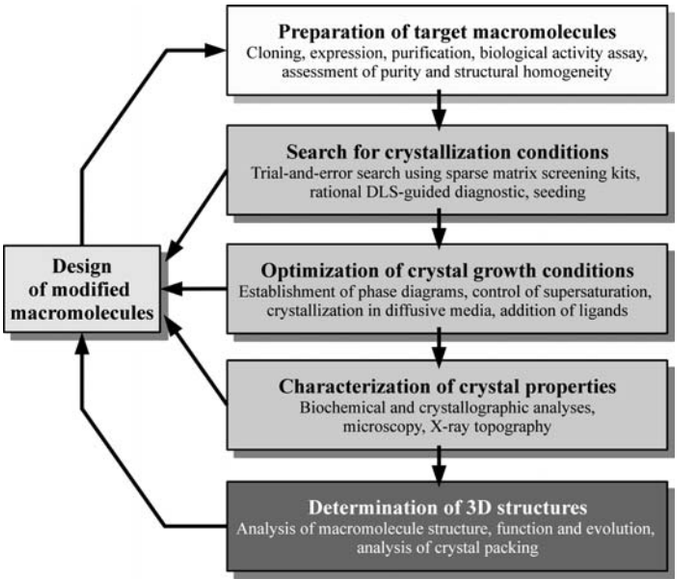

Figure 4.1.6.1
From the target molecule to its three-dimensional structure: a flowchart for a structure detemination. The figure illustrates the different steps the crystal grower will go through to access three-dimensional information. Ideally, the trajectory will be linear, crystallization remaining the major bottleneck in the whole process. If one gets stuck at any stage, the difficulties can often be circumvented by modifying the target molecule, either genetically or (bio)chemically, by improving its purity/homogeneity, or by adding ligands to stabilize its structure.

4.1.6. How to crystallize a new macromolecule: the structural biology view

Fig. 4.1.6.1 schematizes the progress of structural biology projects and highlights the pivotal role of the crystallization step. Despite tremendous progress in macromolecule and crystal preparation, crystallization remains the most serious bottleneck. Various strategies have been proposed to overcome the difficulty (*e.g.* D'Arcy, 1994; Ducruix & Giegé, 1999; Kundrot, 2004; McPherson, 2004; DeLucas *et al.*, 2005; Pusey *et al.*, 2005; Hughes & Ng, 2007; Li *et al.*, 2007; Chayen & Saridakis, 2008). The choice is largely dependent on the type of project [functional biology with well defined protein(s) or structural genomics with large ensembles of known and unknown proteins] and thus on non-exclusive rational, incomplete factorial screening, or empirical approaches.

4.1.6.1. How to start and how to choose what screening kits to start with

In practice, the very first questions posed when one begins to think about crystallization assays are: (i) which crystallization kit to start with, (ii) what to do if no crystals grow and (iii) what happens when one gets crystals. The answer to the first question is in fact simplest. One should start with a screen that does not unreasonably tax the supply of protein but which explores the widest volume of crystallization space, *i.e.* samples the largest number of crystallants and crystallant concentrations over the largest range of pH. There are many kits on the market using 96-sample trials that accomplish this well. If protein is severely limited, then there are 48-sample kits that do the job satisfactorily. In case of no success one should examine the resulting pattern of precipitates, phase separations, clear drops and assorted odd accumulations, and try to divine what might be the best option to try next. Good advice is to check how others have crystallized similar proteins and to try complexes of the protein with its physiological ligands and effectors.

If, on the other hand, one is fortunate and the first kit does indeed yield crystals, but perhaps of insufficient size or quality, or of troubling morphology, then crystallization must be optimized to get better ones. One should also check whether it is really the macromolecule that crystallized and not a ligand or any component of the crystallization mixture. Optimization means varying the chemical and physical parameters of the reagent mix that yielded the crystals and searching the crystallization parameter space by small increments away from the starting point. There is still a component of art and mystery in science, and this is one instance where its appreciation is paramount.

As a practical example, one can consider the 'Silver bullet' concept, which employs a limited set of fundamental crystallization conditions combined with a broad screen of potentially useful small-molecule additives (McPherson & Cudney, 2006), with the hypothesis that some of them might establish intermolecular noncovalent cross-links (Larson *et al.*, 2008). In the 'Uppsala' concept, a crystallization project entails two distinct steps: screening and optimization. The aim of the initial phase is to screen the parameters affecting crystallization as broadly as possible. If promising conditions are found, these are optimized with other protocols including grid and additive screens, seeding, and manipulation of drop kinetics (Bergfors, 2007).