Acta Cryst. (1996). D**52**, 647–654

A Local Approximation to Supersaturation Affords a Useful Coordinate Transformation for the Study of Crystal Growth

CHARLES W. CARTER JR

Department of Biochemistry and Biophysics, CB 7260 University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7260, USA. E-mail: carter@med.unc.edu

(*Received 1 November 1995; accepted 31 January 1996*)

Abstract

Two of the most important experimental variables in the search for appropriate crystallization conditions are the initial concentrations of macromolecule and crystallizing agent. Previously, it has been suggested that the coordinate transformation {[crystallizing agent], [macromolecule]} → {[macromolecule] × [crystallizing agent], [macromolecule]} be used to sample crystal growth conditions. Here, it is shown that this transformation is a special case of a generally applicable transformation. The initial supersaturation can be represented locally by a rectangular hyperbola involving multiples of the product of macromolecule and crystallizing agent concentrations. The coordinate system for the solubility diagram, ([crystallizing agent] *versus* [macromolecule]), can thus be transformed analytically to an alternative coordinate system in which the independent variables are local approximations to the initial supersaturation and the reservoir of soluble macromolecule available to feed a growing seed. In the new coordinate system the 'nucleation zone' is 'orthogonalized', so it can be sampled efficiently on a rectangular grid, with greater assurance that experiments will give rise to crystals. Moreover, since these new coordinate directions segregate fundamental effects on nucleation from effects on growth, using them in experimental designs should improve data analysis for response surface experiments.

1. Introduction

Many crystal growth processes involve a single targeted set of macromolecular and crystallizing agent concentrations to be achieved by vapor-phase or dialysis equilibration. Different choices for these targeted values can cause considerable variation in crystal growth, so they are critical parameters for optimization.

In our efforts to fit quadratic polynomial models to a sample of experimental measurements of crystal growth properties chosen for that purpose (Carter & Yin, 1994), we found that expressly sampling different values of the supersaturation improved the overall performance of these 'response surface' experiments. This paper presents a conceptual justification for our use of the product, [macromolecule] × [crystallizing agent], to represent the supersaturation, together with some of the practical advantages of transforming the solubility diagram coordinates in this manner.

2. Solubility phase diagrams and the nucleation zone

The solubility or phase diagram representing the [macromolecule] × [crystallizing agent] plane, Fig. 1(*a*), is central to the study of macromolecular crystal growth. Considerable evidence exists to show that changing the initial concentrations of either species can substantially alter a crystallogenesis experiment (for example, see Feher & Kam, 1985; Ataka & Tanaka, 1986; Ataka & Michihiko, 1988; Scarborough, 1994; Thibault, Langowski & Leberman, 1992). The two concentrations impact crystal growth *via* their effects on nucleation and growth kinetics (Rosenberger & Meehan, 1988). They can also influence, and are sensitive to, equilibration rates (Luft, Arakali, Kirisits, Kalenik, Wawrzak, Cody, Pangborn & DeTitta, 1994; Luft & DeTitta, 1996). The temporal evolution of a system's position on the solubility diagram therefore represents succinctly the essential thermodynamic and kinetic influences on crystal growth (see, for example, Ducruix & Giegé, 1992).

Many authors (Kam, Shore & Feher, 1978; Boistelle & Astier, 1988; Ducruix & Riès-Kautt, 1990; Mikol & Giegé, 1992; Riès-Kautt & Ducruix, 1992) subdivide the region above the solubility curve into three 'zones'. Points that spontaneously give rise to crystals lie inside the shaded region in Fig. 1(*a*). Between this 'nucleation zone' and the solubility curve the concentration of macromolecule is too low to nucleate crystal growth. The solution is nevertheless supersaturated and hence, 'metastable' with respect to growth on existing nuclei. Above the nucleation zone, at very high supersaturations, nucleation is so rapid that only microcrystalline or amorphous precipitation occurs.

2.1. *Sampling the nucleation zone*

Once conditions have been found that stabilize ordered lattice interactions, the search for optimal crystallization conditions can often be cast in terms of finding appropriate initial and final positions relative to the nucleation zone. Unless seeding is used to circumvent the need for *in situ* nucleation, this is the region that must

be sampled in the design for an optimization experiment.

In view of its importance, surprisingly little attention has focused on how to sample the 'nucleation zone' experimentally. Most approaches to sampling rely on a rudimentary rectangular grid (Cox & Weber, 1988). We found this approach to be inadequate when we began the routine use of response surface analysis (Box, Hunter & Hunter, 1978; Carter & Yin, 1994) to define optimum conditions for growing crystals for X-ray diffraction analysis. Response surface experiments (Box, Hunter & Hunter, 1978) produce a functional representation, or surface, to represent the behavior of a system, *i.e.* its response, as input conditions change. They rely on having results from all or most of the experiments in a carefully chosen design (Hardin & Sloane, 1993) and fitting them to a quadratic polynomial model.

Response surface designs emphasize the problematic nature of sampling the irregularly curved nucleation zone. An essential feature of the design matrices in such experiments is that, to ensure the most accurate parameters, most experimental sampling points are set as far as practical from the suspected optimum. That strategy defines as closely as possible the curvature of the response within the sampled domain of the surface. However, it also heightens the risk that sampling points will lie outside the nucleation zone (see the sampling points in Fig. 1(*a*); the black crosses, or roughly half of them, will not produce crystals). From a practical standpoint, it is therefore critical to distribute experiments within the sausage-shaped figure in Fig. 1(*a*) so that each one will actually produce crystals. Variation of the two concentrations on a rectangular grid in the traditional representation will inevitably waste experiments, as illustrated by the black crosses outside the nucleation zone in Fig. 1(*a*).

2.2. *Ridge analysis and the choice of independent variables*

Our first response surfaces revealed the sampling problem in a second rather different light. It is always useful, if possible, to choose variables with linearly independent effects on crystallogenesis. Combinations of variables for which this is not true give rise to 'ridges', along which the surface achieves the same maximum value, rather than to unique optima.

Although the response surfaces themselves are multi-dimensional functions, two-dimensional 'slices' can be represented graphically by choosing constant values for all but two of the independent variables, and plotting the resulting values of the function. This representation is called a 'level surface', where 'level' refers to the constant values of the remaining variables. A recurrent feature of our early response surfaces was a ridge in one or more of the level surfaces involving [protein] at constant [crystallizing agent] (Fig. 2). In these plots, the optimum value of the dependent variable, plotted along the z axis, has a more or less constant value for different combinations of the two independent variables.

The significance of ridges can be appreciated by transforming to 'canonical coordinates' in which one of the new axes lies along the ridge (Box, Hunter & Hunter, 1978). This procedure demonstrates that the two variable dimensions giving rise to a ridge actually comprise a single, more fundamental dimension, and are therefore, in a sense, redundant. The ridges shown in Fig. 2 probably arose because macromolecular solubility at constant

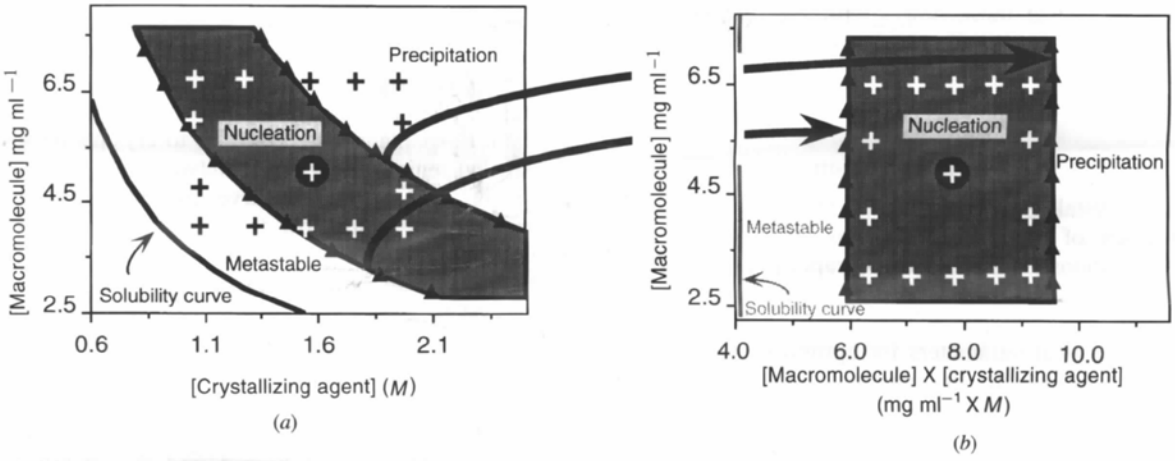

Fig. 1. Coordinate transformation facilitating sampling in the problematic [crystallizing agent] × [macromolecule] plane. (*a*) Here, the solubility curve is approximated by a rectangular hyperbola in the neighborhood of the circled point, where crystals are suspected to be optimal. In order for experiments in a response surface design to produce scores, they must lie inside the curved, shaded region. Sampling on a rectangular grid in this coordinate system places only the white crosses inside the 'nucleation zone'. (*b*) The same region re-plotted in a new coordinate system, with axes [macromolecule] × [crystallizing agent] and [macromolecule]. There is a one-to-one mapping of experimental points between the two coordinate systems, but in (*b*), the nucleation zone appears as a rectangle, so judicious sampling on a rectangular grid places all experiments inside the nucleation zone. The two coordinate directions in (*b*) also approximate the supersaturation and reservoir size, respectively (see text).

ionic strength is temperature dependent (Jakoby, 1968). Hence, the same initial supersaturation can be achieved at a given salt concentration by many different combinations of [macromolecule] and temperature. We therefore suspected that the fundamental variable whose optimum was achieved at the top of the ridges was in fact supersaturation.

The presence of ridges in level-surface plots like those in Fig. 2 led us, in this manner, to recognize that solubility might be approximated by a rectangular hyperbola, $C_S \simeq$ a constant/[crystallizing agent]. The product, [macromolecule] × [crystallizing agent], might therefore behave approximately like the initial supersaturation and serve as a new independent variable (Carter & Yin, 1994).

The one-to-one coordinate transformation,

$$\{x,y\} = \{[\text{crystallizing agent}],[\text{macromolecule}]\} \rightarrow$$

$$\{x',y'\} = \{[\text{macromolecule}] \times [\text{crystallizing agent}],[\text{macromolecule}]\},$$

illustrated in Fig. 1(*b*) has two interesting consequences. First, the new coordinate directions represent the fundamental parameters that act on nucleation and subsequent crystal growth, respectively. Second, experiments can be distributed on the rectangular representation of the nucleation zone in Fig. 1(*b*) according to a simple two-dimensional grid, compatible with Hardin–Sloane response surface design matrices (Hardin & Sloane, 1993), thereby minimizing the risk of experiments not producing scores because they fell outside the nucleation zone. Our initial success with this procedure (Carter & Yin, 1994) motivated the following demonstration of its significance and generality.

2.3. *Basis vectors for the geometric resolution of nucleation and growth*

The obvious remark that the nucleation zone is two dimensional has non-trivial practical and conceptual implications. Experiments targeted to different parts of the nucleation zone produce different kinds of crystals. We have noted, in particular (Carter & Yin, 1994), that certain regions of the nucleation zone correspond to 'stationary points' where crystals grow to an optimal volume, with optimal isometry, and with a minimum of secondary nucleation. Neither the existence nor the location of such stationary points are obvious without careful experimentation. In one example, the stationary point proved to be at protein and precipitant concentrations previously thought to lie in the precipitation zone. Moreover, moving away from such stationary points by changing either the macromolecular or crystallizing agent concentrations can lead to smaller, badly formed and multiple crystals. Thus, finding stationary points involves at least a two-dimensional search.

What is the physical significance of this two dimensionality? Qualitatively, crystal growth occurs within the nucleation zone because the macromolecular concentration, C_T, exceeds its equilibrium solubility, C_S, sufficiently that either homogeneous nucleation of crystal growth or heterogeneous nucleation on contaminating surfaces occurs at a reasonable rate (Boistelle & Astier, 1988; Weber, 1991). Once formed, however, a seed can be nourished to different extents, depending on the reservoir of supersaturated protein. The supersaturation and reservoir represent the two fundamental dimensions of the nucleation zone.

Quantitatively, this two dimensionality of the nucleation zone is rooted in a potentially confusing mathematical curiosity: the linear independence of the different formulations of the supersaturation. Supersaturation is alternatively defined by the ratio, C_T/C_S, by the reservoir itself, $(C_T - C_S)$, or by the ratio of the reservoir to the solubility, $(C_T - C_S)/C_S$, (Boistelle & Astier, 1988; Weber, 1991).

The supersaturation ratio is a unitless intensive quantity related to the differential chemical potential of a supersaturated solution, $\Delta\mu = k_B T \ln(C_T/C_S)$. At low supersaturation ratios, < 0.15, the logarithm can be approximated by the third representation, $(C_T - C_S)/C_S$. On

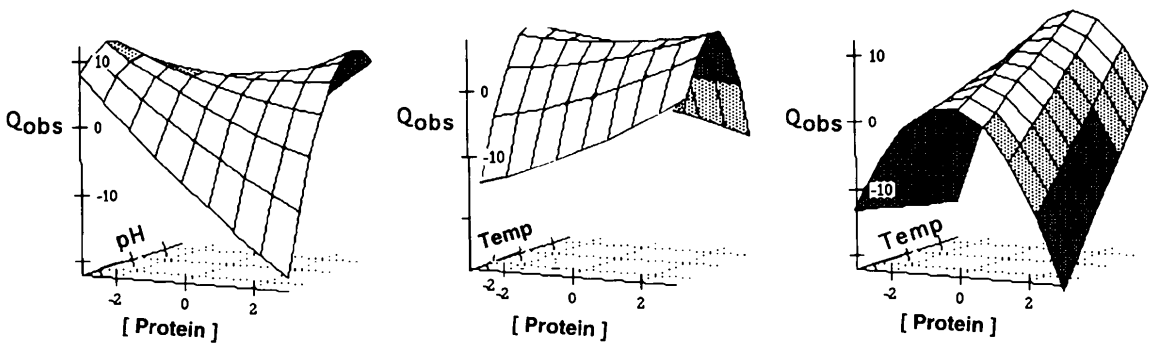

Fig. 2. Level surfaces (Carter & Yin, 1994) for different response surfaces based on data for tryptophanyl-tRNA synthetase crystals and illustrating ridges involving the [macromolecule]. Scores, Q_{obs}, were based on a quality scale introduced by Carter & Carter (1979). (*a*) and (*b*) Level surfaces for triclinic crystals of ligand-free enzyme. (*c*) Level surface for the tetragonal form grown with ATP and tryptophan.

the other hand, the concentration difference, $(C_T - C_S)$, has the units of concentration. Although these representations are widely believed to be interconvertable, they are not. In fact, they can be seen to be linearly independent vectors that span the nucleation zone in Fig 1(*a*). Supersaturation increases along a direction perpendicular to the solubility curve, while the reservoir is parallel to the y axis. As we shall see, they form natural basis vectors for the nucleation zone, the former determining the nucleation rate and the latter the subsequent growth rate onto a nucleus.

2.3.1. *Nucleation.* The activation free energy, and hence the rate of nucleation, depend on the supersaturation ratio, $[\text{macromolecule}_{(\text{Total})}] \div [\text{macromolecule}_{(\text{Equilibrium})}] = C_T/C_S$. This ratio governs nucleation kinetics *via* a power law characteristic of multimolecular chemical reactions, rate $\propto (C_T/C_S)^{\nu}$, where ν represents the number of molecules in the critical nucleus (Hofrichter, Ross & Eaton, 1974, 1976; Feher & Kam, 1985; Ataka & Tanaka, 1986).

A potential source of confusion is the apparent discrepancy between the power law expression and that found in physical texts on crystal growth (Mutaftschiev, 1993), which give the nucleation rate as $J = AC_T \exp\{-B/[\ln(C_T/C_S)]^2\}$. In fact, the two expressions are alternative derivations based on Gibbs' capillarity approximation, $\Delta G^* = -(V/\Omega)k_BT[\ln(C_T/C_S)] + S\gamma$, where V is the volume of the nucleus, S is its surface area, γ is the free energy required to create a unit area of surface interface, and Ω is the volume of a molecule in the interior of the nucleus (Boistelle & Astier, 1988). The power law dependence arises from the fact that V/Ω is ν, the number of molecules inside the nucleus, so $\Delta G^*/k_BT = \gamma\ln(C_T/C_S) + S\gamma/k_BT$.

The alternative expression, $\Delta G^* = -B/[\ln(C_T/C_S)]^2$, results from assuming a spherical nucleus, solving for and substituting the critical radius, $r^* = (2\Omega\gamma)/[k_BT\ln(C_T/C_S)]$. The two expressions are therefore equivalent if the critical radius is constant with respect to supersaturation, which is approximately true for several common systems (Hofrichter, Ross & Eaton, 1974, 1976; Ataka & Tanaka, 1986). The expression favored by physicists accomodates the variation of critical radius with supersaturation. However, it obscures the simple and chemically intuitive power law dependence of the nucleation rate on the supersaturation ratio and the number of subunits in the critical nucleus, without suggesting any compensating physical intuition.

The rate at which a supersaturated solution will generate nuclei is approximately constant along curves parallel to the solubility curve and within the nucleation zone. Moreover, this functional dependence is the same for both homogeneous and heterogeneous nucleation, the only difference being that the presence of heterogeneous nucleants lowers the threshold supersaturation at which nucleation occurs (Boistelle & Astier, 1988).

2.3.2. *Growth.* Once formed, a nucleus grows by an accumulation of soluble particles (monomers or aggregates) that is governed in part by their rates of diffusion (Smoluchowski, 1917) and in part by the kinetics of adsorption to the crystal surface (Rosenberger, Muschol, Thomas & Vekilov, 1996). The solution in the immediate vicinity of a growing crystal is in equilibrium with it, and therefore has a macromolecular concentration given by the solubility. Far away, the macromolecular concentration is higher, and initially equal to C_T. Diffusion toward the growing crystal is therefore driven by the gradient, which, upon the initial nucleation, is given by the concentration difference, $(C_T - C_S)$. The extent of growth may also be determined largely by the concentration difference, which also represents the available reservoir of supersaturated protein to add to a growing crystal. This reservoir can change by orders of magnitude for experiments performed at the same supersaturation, but at different pH values or crystallizing agent concentrations (Boistelle & Astier, 1988).

2.4. *An intuitively satisfying local coordinate system*

To summarize the argument thus far, we can capture the essential two dimensionality of the nucleation zone by choosing local basis vectors related directly to determinants of nucleation and growth. The two alternative definitions of supersaturation have distinct effects on crystal growth and provide natural basis vectors for resolving nucleation from subsequent growth. In this context, it seems preferable to reserve 'supersaturation' for the ratio, C_T/C_S, in order to preserve the distinction between its effect on the activation free energy for nucleation and use another term, 'reservoir', to denote $(C_T - C_S)$, whose impact is largely on growth.

Such a coordinate system cannot be applied globally. A pre-exponential contribution to the initiation rate law is proportional to the total macromolecular concentration, and the size of the critical nucleus may depend on the supersaturation. So the nucleation rate is not exactly constant for each supersaturation, but increases at higher solubilities (Boistelle & Astier, 1988). Moreover, specific ion-pairing interactions (Riès-Kautt & Ducruix, 1991) can occur with the macromolecule, so the species that crystallize and hence their response surfaces, may change as ionic strength is increased. However, near a suspected optimum it is possible to transform the conventional coordinates of the solubility diagram to represent instead the parameters that govern nucleation and growth, giving an intuitively satisfying local coordinate system.

3. A local approximation to supersaturation

How equivalent is the product, [crystallizing agent] × [macromolecule] to supersaturation? There is of course a heuristic analogy between this expression

and that for the solubility-product constant in inorganic chemistry (Nernst, 1889). In that case, supersaturation formally corresponds to stoichiometric combinations of anion and cation concentrations in excess of the solubility product. Stoichiometric equivalence of macromolecule and crystallizing agent is unlikely in macromolecular crystals, but this analogy suggests that values of [crystallizing agent] × [macromolecule] in supersaturated macromolecular solutions bear a similar relation to their solubility curves.

For a single precipitant, solubility, C_S, is often described by the empirical Cohn–Green formula (Arakawa & Timasheff, 1986),

$$\ln C_S = \beta - K_S \cdot m$$
$$C_S = \exp(\beta - K_S \cdot m) = A\exp(-K_S \cdot m), \quad (1)$$

where β and K_S are constants, m represents [crystallizing agent], and $A = \exp(\beta)$ represents the limiting value of C_S at $m = 0$. This classical expression for the solubility curve as an exponential decay is semi-empirical; it arises from the thermodynamic dependence of the chemical potential of the macromolecule, *i.e.*, the logarithm of its concentration or activity, on the concentration of the crystallizing agent. Experimental systems deviate to a greater or lesser extent from this expected behavior, so in general this expression, too, is only valid locally. Our motivation was to capture this exponential decay behavior in an alternative functional form more amenable to experimental design requirements.

We seek a locally equivalent expression with the functional form of a rectangular hyperbola,

$$C_S = C_0/(m - m_{T_r}), \quad (2)$$

where C_0 and m_{T_r} are constants providing, respectively, the multiplicative factor and translation necessary for the two functions to intersect, with the same slope, at the arbitrary point, ($x_0 = m_0$, $y_0 = C_{S,0}$). Equating (1) and (2) and their first derivatives provides simultaneous equations for C_0 and m_{T_r}, giving,

$$C_0 = (A/K_S)\exp(-K_S m_0) \text{ and } m_{T_r} = m_0 - (1/K_S). \quad (3)$$

The arbitrary point on the solubility curve, around which the approximation is centered, is, ($x_0 = m_0$ = [crystallizing agent]$_0$, $y_0 = C_{S,0} = C_0 K_S$). These relationships are illustrated in Fig. 3(*a*) for a simulated case and in Fig. 3(*b*) for a published solubility curve for lysozyme (Ducruix & Riès-Kautt, 1990).

Values have been tabulated for K_S for a number of crystallizing agents (Arakawa & Timasheff, 1986), and an estimate for the maximum macromolecular solubility, A, is usually available from experience during purification or elsewhere. (2) can therefore be used, in the neighborhood of the suspected optimum ([crystallizing agent] = m_0), to represent the supersaturation by,

$$S = [\text{macromolecule}]/C_S$$
$$= ([\text{macromolecule}]/C_0)(m - m_{T_r}). \quad (4)$$

The same approximate form can be used for curves of constant supersaturation, passing through multiples of $C_0 K_S$.

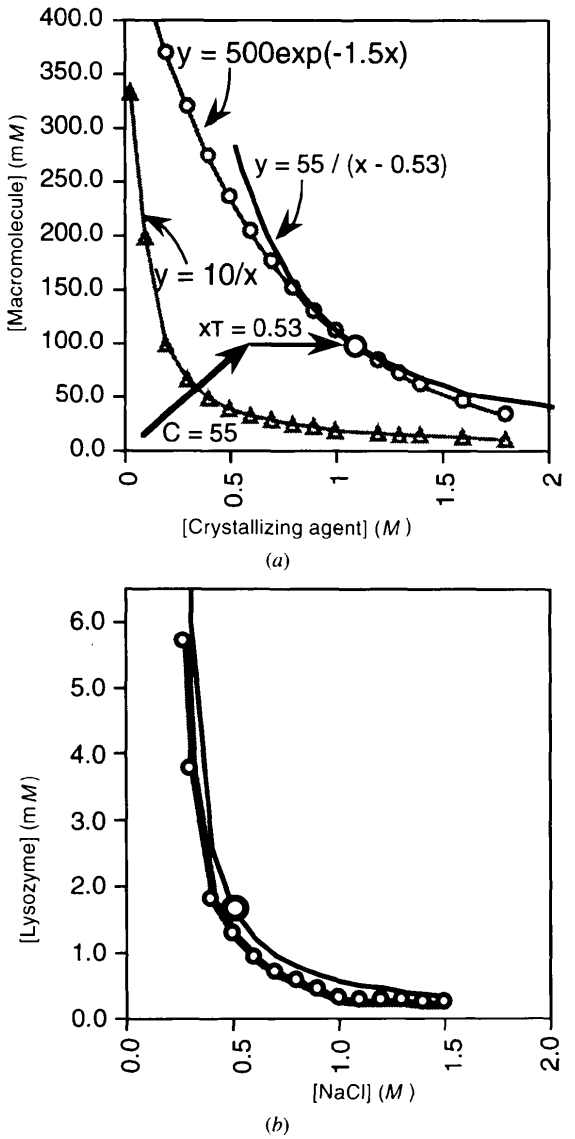

Fig. 3. The local hyperbolic approximation to the solubility curve. (*a*) The rectangular hyperbola, $y = 1/x$ is scaled by C_0 and translated along x by $-X_{T_r}$ such that both y and y' agree with the solubility curve at the point, (x_0, $C_{S,0}$). (*b*) An example approximating solubility data for lysozyme in Riès-Kautt & Ducruix (1992). The approximation is centered at the point, $x_0 = 0.5$ *M*, $C_{S,0} = 1.64$ m*M*; the translation, $x_{T_r} = 1/KS = 0.20.275$, and the constant, $C_0 = 0.453$. The lack of fit results from the fact that the lysozyme data do not obey a single exponential relationship, and the fit was selected to represent a range of salt concentrations spanning the bimodal lysozyme data.

The local approximation to C_S also provides an approximation to the reservoir,

$$\begin{aligned}\mathcal{R} &= C_T - [C_0/(m - m_{T_r})] \\ &= [\text{macromolecule}] \\ &\quad - \{C_0/([\text{crystallizing agent}] - m_{T_r})\}. \end{aligned} \quad (5)$$

These local approximations lead directly to the general coordinate transformation, $\{[\text{crystallizing agent}], [\text{macromolecule}]\} \rightarrow \{\mathcal{S}, \mathcal{R}\}$. The special case illustrated in Fig. 1(*b*), where $m_T \ll m$, [or $m_0 \simeq (1/K_S)$], obtains in the neighborhood of the concentration of crystallizing agent required to reduce the solubility to $\exp(-1) = 0.37$ of its value in buffer, and is thus not unreasonable for a crystal growth experiment. Under these conditions the approximation reduces to the expression proposed originally (Carter & Yin, 1994),

$$\mathcal{S} = ([\text{macromolecule}][\text{crystallizing agent}])/C_0,$$

$$\mathcal{R} = [\text{macromolecule}] - (C_0/[\text{crystallizing agent}]). \quad (6)$$

The increasing use of binary mixtures of crystallizing agents, such as a salt and a nonionic polymer (Jancarik & Kim, 1991), raises the question of how to generalize (4). For such cases one can use (4) to express independent hyperbolic approximations for each crystallizing reagent. An alternative model for supersaturation may be approximated by generalizing the solubility formula. Assuming to a first approximation that several crystallizing agents act independently, *i.e.* that $\ln C_S = \beta - \sum_{i=1}^{n} K_S^i \cdot M_S^i$ gives,

$$\begin{aligned}\mathcal{S} &= [\text{macromolecule}] \cdot \prod_{i=1}^{n} \exp(K_S^i \cdot M_S^i) \\ &\simeq [\text{macromolecule}] \cdot \prod_{i=1}^{n} (1 + K_S^i \cdot M_S^i). \end{aligned} \quad (7)$$

Use of these approximations requires a knowledge of the coefficients, K_S^i. The best available compilation of these parameters (Arakawa & Timasheff, 1986) suggests that for salts, they depend somewhat on both the nature of the protein and the solution pH. Both effects probably relate to differences in net charge (Carbonnaux, Riès-Kautt & Ducruix, 1995). Coefficients for different molecular weights of polyethylene glycol appear to satisfy a linear relationship: $K_S = 0.017 \times M_r\text{PEG} - 5.28$, $r^2 = 0.97$, for data from several different proteins. This relationship, obtained from Table III of Arakawa & Timasheff (1986), indicates that for molal concentrations of all crystallizing agents, the salting-out effectiveness of low molecular weight PEG reagents (PEG 200 and 400) is comparable to that of salts like NaCl and $NaCH_3COO$, for which $K_S \simeq 10$, whereas that of higher polymers (PEG 6000 and 8000) can be more than an order of magnitude greater: $K_S \simeq 100$. Further study of these relationships and careful measurement of more K_S values is certainly warranted.

4. Implementation

Although the variables, $\mathcal{S}$ and $\mathcal{R}$ seem unfamiliar at first, designing experiments using them is greatly facilitated by the use of spreadsheets, programmed to carry out the necessary algebra on a set of input parameters and produce a description for each experiment. *EXCEL* spreadsheets programmed for Hardin–Sloane response-surface designs involving four and five variables are available from the author at carter@med.unc.edu or *via* the World Wide Web at URL http://russell.med.unc.edu/~carter/designs.

Response surface experiments using these and other variables can be carried out in a variety of different situations, including dialysis, vapor diffusion, and microbatch equilibrations. In cases where stationary points can be identified that correspond to optima, their coordinates provide recipes for improved and more reproducible crystal growth, as well as credible evidence for potentially interesting synergistic and non-linear effects.

5. Conclusions

Producing appropriate crystals for structure determination often requires both initial screening (Carter & Carter, 1979; Carter, Baldwin & Frick, 1988; Carter, 1990, 1992; Jancarik & Kim, 1991; Doudna, Grosshans, Gooding & Kundrot, 1993; Kundrot, 1996; Gouaux, 1996) and subsequent optimization. Screening relies on uniform, randomized sampling to identify combinations of variables that give rise to crystalline lattices. Optimization adjusts conditions identified by screening to improve the crystals. Reaching these two goals involves quite different strategies (Carter & Yin, 1994).

In many cases, in particular those where crystals grow by *in situ* nucleation and a fixed targeted equilibrium, optimization becomes a problem in sampling the nucleation zone for the best initial combination of [crystallizing agent] and [macromolecule] at various pH values, temperatures, ligand concentrations and so on. In these cases, the local coordinate transformation described here can be very helpful. Response surface experiments provide a coherent way to study simultaneously relationships between a measurable result, such as crystal shape or volume, and a local region of several input variables, such as protein concentration, crystallizing agent concentration, temperature, and so on, that influence that result. They are, however, sensitive to the loss of crucial information whenever any of the ensemble of experiments fails to produce crystals.

Relationships in the preceding paragraphs are summarized for reference in Table 1. The coordinate transformation amounts to a local orthogonalization of the solubility phase diagram with the advantages associated with such transformations. They facilitate sampling by minimum-prediction variance designs on a rectangular grid. Crystal growth experiments properly defined in the

Table 1. *Local orthogonalization of the solubility phase diagram*

Coordinate system	X variable	Y variable	Comments
Solubility diagram (Natural coordinates)	m = [crystallizing agent]	C_T = [macromolecule]	
Local orthogonalization	$\mathcal{S}$, Initial Supersaturation	$\mathcal{R}$, Reservoir	
(Fundamental coordinates)	$\frac{C_T}{C_0} \times (m - mT_r)$	$C_T - \frac{C_0}{(m - m_{T_r})}$	Improved searching/sampling properties
Constants	$m_{T_r} = \left(m_0 - \frac{1}{K_S}\right)$	$C_0 = \frac{A}{K_S}\exp(-K_S m_0)$	Estimate from published values of K_S
	K_S = est. slope of solubility curve	$A = \exp(\beta)$ = y-intercept of solubility curve	or solubility curve
Equivalence point	$x_0 = m_0$ = [crystallizing agent]$_0$	$y_0 = C_{S,0} = C_0 K_S$.	Center designs here for optimization.
Condition for $m_T = 0$	$m_0 \simeq \frac{1}{K_S}$	$C_0 = \frac{A}{K_S}\exp(-1) = 0.37^* \frac{A}{K_S}$	Special case described by Carter & Yin (1994)

local orthogonal system will be more efficient because few, if any experiments will lie outside the nucleation zone. Moreover, since the search directions, $\mathcal{S}$ and $\mathcal{R}$, correspond to physico-chemical determinants of nucleation and growth rates, respectively, such experiments should do better at locating optimal or stationary points for production of crystals for X-ray diffraction.

This discussion emphasizes the fundamental importance of the solubility curve in dictating the course of crystal growth and the search for optimal conditions. Studies of the solubility diagram involve exhaustive equilibrium measurements (Ducruix & Riès-Kautt, 1990; Riès-Kautt & Ducruix, 1992) and are therefore prohibitive in most practical cases. Rough estimates of the solubility parameters, β and K_S, can be obtained from plots of two or three values of ln C_S *versus* [crystallizing agent]. Sampling of the nucleation zone can then be carried out efficiently by response surface experiments, without the need for detailed measurements of the solubility curve itself.

I thank J. Hermans, L. Jiang, M. Hasson, H. Ke, M. Ataka, and M. Riès-Kautt for discussions and critical comments on the manuscript, and NASA for financial support. The work of R. Boistelle and comments from a referee were very helpful in highlighting differences between physical and chemical models for nucleation.

References

Arakawa, T. & Timasheff, S. N. (1986). *Methods Enzymol.* **114**, 49–77.

Ataka, M. & Michihiko, A. (1988). *J. Cryst. Growth*, **90**, 86–93.

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

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

Box, G. E. P., Hunter, W. G. & Hunter, J. S. (1978). *Statistics for Experimenters*, Vol. pp. Pages. New York: Wiley Interscience.

Carbonnaux, C., Riès-Kautt, M. & Ducruix, A. (1995). *Protein Sci.* **4**, 2123–2128

Carter, C. W. Jr (1990). *Methods Companion Methods Enzymol.* **1**, 12–24.

Carter, C. W. Jr (1992). *Design of Crystallization Experiments and Protocols*, in *Crystallization of Proteins and Nucleic Acids, A Practical Approach*, edited by R. Giegé & A. Ducruix, pp. 47–71. Oxford: IRL Press/Oxford University Press.

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

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

Carter, C. W. Jr & Yin, Y. (1994). *Acta Cryst.* D**50**, 572–590.

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

Doudna, J. A., Grosshans, C., Gooding, A. & Kundrot, C. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 7829–7833.

Ducruix, A. & Giegé, R. (1992). *Methods of Crystallization*, in *Crystallization of Nucleic Acids and Proteins*, edited by A. Ducruix & R. Giegé, pp. 73–98. Oxford University Press.

Ducruix, A. F. & Riès-Kautt, M. (1990). *Methods Companion Methods Enzymol.* **1**, 25–30.

Feher, G. & Kam, Z. (1985). *Nucleation and Growth of Protein Crystals: General Principals and Assays*, in *Diffraction Methods for Biological Macromolecules*, Vol. 114, edited by H. Wyckoff, C. H. W. Hirs & S. N. Timasheff, pp. 77–111. Orlando, FL: Academic Press.

Gouaux, E. (1996). *Methods Enzymol.* In the press.

Hardin, R. H. & Sloane, N. J. A. (1993). *J. Stat. Plan. Inf.* **37**, 339–369.

Hofrichter, J., Ross, P. D. & Eaton, W. A. (1974). *Proc. Natl Acad. Sci. USA*, **71**, 4864–4868.

Hofrichter, J., Ross, P. D. & Eaton, W. A. (1976). *Proc. Natl Acad. Sci. USA*, **73**, 3035–3039.

Jakoby, W. B. (1968). *Anal. Biochem.* **26**, 295–298.

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.

Kam, Z., Shore, H. B. & Feher, G. (1978). *J. Mol. Biol.* **132**, 539–555.

Kundrot, C. E. (1996). *Methods Enzymol.* In the press.

Luft, J. R., Arakali, S. V., Kirisits, M. J., Kalenik, J., Wawrzak, I., Cody, V., Pangborn, W. A. & DeTitta, G. T. (1994). *J. Appl. Cryst.* **27**, 443–452.

Luft, J. R. & DeTitta, G. T. (1996). *Methods Enzymol.* Submitted.

Mikol, V. & Giegé, R. (1992). *The Physical Chemistry of Protein Crystallization*, in *Crystallization of Proteins and Nucleic Acids, A Practical Approach*, edited by A. Ducruix & R. Giegé, pp. 219–239. Oxford: IRL Press/Oxford University Press.

Mutaftschiev, B. (1993). *Nucleation Theory*, in *Handbook of Crystal Growth*, Vol. 1*a*, edited by D. T. J. Hurle, pp. 187–247. Amsterdam: North Holland.

Nernst, W. (1889). *Z. Physik. Chem.* **4**, 379.

Riès-Kautt, M. & Ducruix, A. (1991). *J. Cryst. Growth*, **110**, 20–25.

Riès-Kautt, M. & Ducruix, A. (1992). *Phase Diagrams*, in *Crystallization of Nucleic Acids and Proteins: A Practical Approach*, edited by A. Ducruix & R. Giegé, pp. 195–218. Oxford: IRL Press/Oxford University Press.

Rosenberger, F. & Meehan, E. J. (1988). *J. Cryst. Growth*, **90**, 74–78.

Rosenberger, F., Muschol, M., Thomas, B. R. & Vekilov, P. G. (1996). *J. Cryst. Growth*. In the press.

Scarborough, G. (1994). *Acta Cryst.* D**50**, 643–649.

Smoluchowski, M. V. (1917). *Z. Physik. Chem.* **92**, 129.

Thibault, F., Langowski, J. & Leberman, R. (1992). *J. Cryst. Growth*, **122**, 50–59.

Weber, P. C. (1991). *Adv. Protein Chem.* **41**, 1–36.