Determination of the Phase Diagram for Soluble and Membrane Proteins

Sameer Talreja, Sarah L. Perry, Sudipto Guha, Venkateswarlu Bhamidi, Charles F. Zukoski,* and Paul J. A. Kenis*

Department of Chemical & Biomolecular Engineering, University of Illinois at Urbana–Champaign, Urbana, Illinois 61801

Received: December 13, 2009; Revised Manuscript Received: February 23, 2010

Methods to efficiently determine the phase behavior of novel proteins have the potential to significantly benefit structural biology efforts. Here, we present protocols to determine both the solubility boundary and the supersolubility boundary for protein/precipitant systems using an evaporation-based crystallization platform. This strategy takes advantage of the well-defined rates of evaporation that occur in this platform to determine the state of the droplet at any point in time without relying on an equilibrium-based end point. The dynamic nature of this method efficiently traverses phase space along a known path, such that a solubility diagram can be mapped out for both soluble and membrane proteins while using a smaller amount of protein than what is typically used in optimization screens. Furthermore, a variation on this method can be used to decouple crystal nucleation and growth events, so fewer and larger crystals can be obtained within a given droplet. The latter protocol can be used to rescue a crystallization trial where showers of tiny crystals were observed. We validated both of the protocols to determine the phase behavior and the protocol to optimize crystal quality using the soluble proteins lysozyme and ribonuclease A as well as the membrane protein bacteriorhodopsin.

1. Introduction

Structural biology efforts are often hampered by challenges related to protein expression, purification, stability, and/or crystallization.1 Here, we present an experimental approach which has the potential to improve the rate of success associated with obtaining crystals of sufficient quality for high resolution X-ray structure determination.2 Such screening efforts can be viewed as a search through parameter space for regions where phase boundaries exist between an undersaturated solution and a saturated solution from which a crystal can grow. However, for newly expressed and isolated proteins, the extent of this parameter space is undefined and methods to predict a priori where such phase transitions occur do not exist.3 As a result, such screens for suitable conditions can be time and material intensive. Once identified, crystallization conditions typically need to be optimized with respect to crystal quality, so that high resolution structural information can be obtained through X-ray analysis.1,2 Knowledge of protein phase diagrams is also invaluable for the process of crystal quality optimization.

A phase diagram provides information about crystal nucleation and growth conditions as a function of protein concentration and thermodynamic parameters such as temperature, pH, and/or precipitant concentration. The portions of the phase diagram associated with crystallization can be divided into three regions. The solubility boundary separates the undersaturated region, where any crystals present would dissolve, from the metastable zone, where protein in solution is supersaturated and crystals are able to grow. The supersolubility boundary separates the metastable zone from the labile zone, where the level of protein supersaturation is high enough to induce the nucleation of new crystals. Thus, for a crystal to nucleate and grow, the solution needs to reach the labile zone first, after which it can continue to grow in either the metastable or the labile zone.

A variety of methods to estimate the solubility boundary for proteins have been reported.4–12 Equilibrium solubility can be determined by probing the relationship between a crystalline solid and the corresponding saturated solution. Typically, this is accomplished by either dissolving crystals into an undersaturated solution until saturation is reached or exposing crystals to an oversaturated solution and allowing crystal growth to bring the system to equilibrium. The concentration of protein in the equilibrated mother liquor can then be assayed to determine a point on the solubility curve. The various techniques that have developed around these methods have typically suffered from the need for substantial quantities of protein,5–6,12 the time to complete an experiment,5,7,9,12 and/or the need for highly specialized equipment.5,8 As a result, structural genomics laboratories typically forego determination of a phase diagram, in favor of a much less-informed “optimization” screen of phase space. The limited supply of novel proteins outside of structural genomics laboratories has meant that the phase behavior of only a handful of, mostly commercially available, proteins has been reported in detail, including lysozyme,5–22 thaumatin,23 cytochrome c oxidase,24 insulin,25 chymotrypsinogen,10 glucose isomerase,26 xylose isomerase,27 canavalin,28 concanavalin A,29 collagenses,30 photosynthetic reaction center,31 horse serum albumin,8 ovalbumin,23,24 hemoglobin,13,15 carboxypeptidase G2,6,34 bovine pancreatic trypsin inhibitor,35 and ribonuclease A.32 More efficient methods to determine protein solubility that consume only small amounts of precious protein sample are needed, not only to accelerate structural biology efforts but also to improve the scientific understanding of protein phase behavior. The ability to rationally predict phase behavior would dramatically reduce the need for massive screening efforts.

Though solubility is the more common thermodynamic descriptor associated with crystallization, obtaining information about the supersolubility boundary for proteins is much more typical. In an optimization screen, this boundary can be approximated by examining which conditions led to clear drops
(the drop could be under- or supersaturated, but no nuclei have formed) and which conditions resulted in the formation of crystals. However, the exact location of this supersolubility boundary depends on the method and path by which supersaturation is achieved.

Methods for determining the solubility and supersolubility curves typically require large quantities of protein because of either the scale at which measurements must be made, as in microcolumn measurements to locate the solubility boundary or the large number of experiments needed, as with microbatch screens, to determine the supersolubility boundary. A key advantage of the microbatch method is the fact that the composition of each droplet is known and that it does not vary over time unless a phase change occurs. However, mapping out phase behavior with the microbatch method inherently requires selecting discrete experiments to cover phase space (an approach which we will refer to as pixilation), thus introducing a trade-off between the number of trials and the quality of the data. Because of these various difficulties, methods for characterizing the phase behavior of proteins that are more efficient in time and sample utilization as well as their sampling of phase space are needed.

Once information about the phase behavior of a protein is known, the crystallization experiment can be optimized rationally, for example, to overcome the common problems of excess nucleation leading to many tiny crystals, and poor crystal quality. These two issues can be avoided by separating the nucleation and growth stages: the droplet needs to traverse a path so it first experiences the high levels of supersaturation needed for nucleation (labile zone) and then experiences lower levels of supersaturation (metastable zone) allowing for slower, more ordered crystal growth. The various previously reported strategies to decouple crystal nucleation and growth typically involve preparation of a highly supersaturated droplet followed by a dilution of the droplet to lower levels of supersaturation or the use of heterogeneous nucleants. However, without prior knowledge of the phase diagram, finding adequate conditions to decouple nucleation and growth becomes a screening exercise of its own.

In prior work, we have introduced a controlled evaporation-based crystallization method in which the sample equilibrates with the ambient atmosphere rather than with a reservoir solution (Figure 1). The rate of evaporation is controlled by varying the cross-sectional area or the length of the evaporation channel that connects the crystallization chamber with the ambient. The major advantages of this controlled evaporation approach over traditional vapor diffusion or microbatch methods are that (i) a phase change is guaranteed in each experiment, since the end point is not limited by the equilibrium imposed by the solution present in the crystallization reservoir, (ii) the kinetics of the experiment can be controlled through variation of the rate of evaporation, (iii) the concentration of the sample at any point in time can be calculated on the basis of the known, constant evaporation rate and known initial conditions, and (iv) the experiment can be paused at any time once promising conditions have been reached. We have used this platform, for example, to selectively crystallize polymorphs of glycine and to determine a lower limit to the supersolubility boundary (critical supersaturation).

In this paper, we demonstrate the utility of our evaporation-based crystallization platform (i) to construct the phase diagram of a given protein/precipitant system and (ii) to optimize crystal quality, in a much more efficient manner than traditional vapor diffusion or microbatch methods. Evaporation allows us to sample phase space by having a limited number of droplets (typically five different conditions repeated three times) of known composition traverse different paths, with each experiment giving a point on the solubility and supersolubility curves. This protocol eliminates the need for extensive pixilation (i.e., many experiments) to fully sample the phase space and thus reduces the amount of protein needed. Furthermore, once the solubility and supersolubility boundaries have been determined in this way, a dilution protocol can be used in these same evaporation-based crystallization platforms to decouple nucleation and growth to optimize crystal quality. We validate these protocols using the model protein/precipitant combinations of lysozyme/NaCl, ribonuclease A/NaCl, and bacteriorhodopsin/NaH2PO4.

2. Experimental Methods

2.1. Evaporation-Based Crystallization Platform with Automated Data Acquisition. The evaporation-based crystallization platform used here connects the crystallization droplet with the ambient atmosphere rather than with a reservoir solution (Figure 1). The rate of evaporation is controlled by varying the cross-sectional area or the length of the evaporation channel that connects the crystallization chamber with the ambient.
with the outside environment through a channel of length $L$ and cross-sectional area $A$, which allows for gradual evaporation of the solvent at a diffusion-limited rate. As characterized previously, the volumetric rate of evaporation of solvent from the droplet scales with the humidity or pressure difference $A\Delta P$ between the vapor phase around the drop and the outside environment, and the dimensions of the evaporation channel, specifically the ratio of $A$ over $L$.\textsuperscript{37,51–53}

$$J = \Delta P \left( \frac{A}{L} \right)$$

Thus, the rate of evaporation can be regulated independently from the composition or volume of the droplet itself by altering the dimensions of the channel. This rate is taken to be essentially constant because the effect of changes in the activity coefficient of the droplet with increasing concentration can be neglected over the time scale of a crystallization experiment.\textsuperscript{51}

The platforms used here were machined from poly-ether-ether-ketone (PEEK) with 16 crystallization compartments, each connected to the external environment by an evaporation channel (Figure 1b). The evaporation channels used here had cross-sectional areas $A$ ranging from 0.3 to 1.76 mm$^2$ and a length $L$ of 7 mm. The evaporation rate for each of the different channel dimensions was determined by measuring the time for complete evaporation of both drops of pure water and drops containing different solutes (proteins and/or precipitants). For the solutions used in these experiments (e.g., buffered salt solutions and soluble proteins), the drying times for these different solutions differed by less than 2% from that of pure water, indicating that the decrease in water activity at the end of the drying process has a negligible effect on the overall drying time.\textsuperscript{57,51} However, additives such as PEGs and detergents are known to affect the rate of evaporation,\textsuperscript{54} necessitating separate calibration. More information on the design, fabrication, and use of these crystallization platforms can be found in the Supporting Information.

Crystallization experiments in these 16-compartment plates were tracked using an automatic, computer controlled imaging system (Figure 1a) comprised of an optical microscope (Leica Z16 APO) equipped with an autozoom lens (Leica 10447176), a digital camera (Leica DFC280), and a motorized X–Y stage (Semprex KL66). This setup uses the Advanced Acquisition capability of Image Pro Plus (Media Cybernetics) to monitor 48 wells in autonomous fashion by sequentially moving from well to well and capturing and storing multiple images at different focal planes for each drop.

Using this setup, we determined the nucleation time for each experiment with an error of ±15 min. We define the nucleation time as the time elapsed between the start of the experiment and the first appearance of a crystal of discernible size (∼5 μm) in the evaporating drop. The time needed for a crystal to nucleate and grow to this discernible size can be neglected owing to relatively rapid crystal growth compared to typical times of nucleation.\textsuperscript{55,56}

### 2.2. Protein and Precipitant Solutions

Hen egg white lysozyme (Sigma) was dissolved in 100 mM sodium acetate buffer at pH 4.6 to obtain a protein stock solution of concentration $C_P = 60$ mg/mL, with an initial precipitant concentration $C_{S0}$ of NaCl ranging from 0.2 to 1.0 M. The buffer solution was prepared using acetic acid and sodium acetate (Fisher Scientific). Lysozyme concentrations were determined by UV absorbance measurements (Hewlett-Packard 8453 UV—vis spectrophotometer) at 280 nm using an extinction coefficient of 2.64 mL/(mg·cm).\textsuperscript{57}

Ribonuclease A (RNase A, R-5500, Sigma) from bovine pancreas was dissolved in 100 mM acetate buffer at pH 4.5 to obtain a protein stock solution of concentration $C_P = 100$ mg/mL with an initial precipitant concentration $C_{S0}$ of NaCl ranging from 1.0 to 3.0 M. RNase A concentrations were determined by absorbance measurements using an extinction coefficient of 0.70 mL/(mg·cm).\textsuperscript{58}

A stock solution of $C_P = 24.5$ mg/mL bacteriorhodopsin solubilized in 25 mM NaH$_2$PO$_4$ buffer of pH 5.5 (EMD Chemicals) with 1.2 w/v % β-octyl-d-glucoside (Anatrace) was prepared from a culture of Halobacterium salinarium.\textsuperscript{59,60} The NaH$_2$PO$_4$ buffer in this instance also serves as the precipitant. Different initial precipitant concentrations, $C_{S0} = 0.025$–1.015 M, were obtained by dilution with a solution of 2.5 M buffer at pH 5.5 with 1.2 w/v % β-octyl-d-glucoside. Protein concentration was determined by absorbance measurements at 550 nm using an extinction coefficient of 58 000 M$^{-1}$ cm$^{-1.}$\textsuperscript{61}

Prior to setting up a crystallization experiment, both the protein and precipitant solutions were filtered through 0.02 μm (Anotop 25, Whatman), 0.1 μm (Ultrafree-MC, Millipore), or 0.22 μm (Steriflip, Millipore) filters. Dilutions of the protein stock solutions with their respective buffers yielded solutions of the same $C_P/C_S$ ratio but different initial protein and precipitant concentrations. All experiments were carried out in a climate controlled laboratory at room temperature (23 ± 1 °C) and constant relative humidity (30 ± 2%).

Reservoir solutions for all of the proteins consisted of an equivalent buffer solution with varying concentrations of the precipitant and no protein present. Thus, for lysozyme and RNase A, solutions of varying NaCl concentration in a corresponding acetate buffer were used. The reservoir solution used with bacteriorhodopsin involved varying concentrations of a buffered NaH$_2$PO$_4$ solution along with 1.2 w/v % β-octyl-d-glucoside.

### 2.3. Crystallization Experiments

The crystallization experiments for lysozyme and RNase A were performed using the 16-well evaporation-based crystallization platforms shown in Figure 1. Crystal Clear tape (Hampton Research) was used to seal the bottom of each of the individual evaporation compartments. This tape provides transparency for viewing the crystals and can easily be removed and replaced during the course of an experiment. Five μL droplets of the prepared protein solutions were pipetted onto silanized glass coverslips (Round, $D = 18$ mm, Hampton Research) and then immediately inverted to cap and seal (high vacuum grease along the edges) an individual evaporation compartment. More information on the crystallization experiments can be found in the Supporting Information.

After crystals appeared, approximately 100 μL of reservoir solution was introduced into the compartment through the evaporation channel via a syringe as part of the protocol to determine the solubility boundary (see section 3 below). The outlets of the evaporation channels were then sealed using Crystal Clear tape. The volume of the reservoir solution was chosen to be at least 100 μL to avoid error associated with changes in concentration of the reservoir solution upon equilibration of the droplet. For each iteration of the protocol (section 3), the reservoir solutions were exchanged by first removing the Crystal Clear tape from the bottom and sides of the evaporation platform, blotting away the remaining liquid, and then resealing the bottom of the platform with fresh tape. New reservoir solution can then be introduced through the evaporation
channels as before. The concentration of the reservoir solution was changed in steps of 0.02 M.

Crystallization experiments for bacteriorhodopsin were performed using a 96-well vapor diffusion hanging drop setup (screw top hanging drop crystallization plate, Molecular Dimensions) to demonstrate the compatibility of the protocols reported here with this more commonly employed crystallization method. The wells in the crystallization tray used are sealed by a screw cap which incorporates the coverslip. The screw top allows for repeated access to the wells without affecting the seal on the chamber. Three µL droplets were pipetted into the recess on these screw top coverslips which were then suspended over a reservoir solution of at least 100 µL. The reservoir concentration was changed over time first to concentrate and then to dilute the drops. The concentrations of the reservoir solutions were changed by the addition of a more dilute or concentrated solution rather than by iterative replacement. The concentration of the salt solution was changed in steps of 0.025–0.1 M.

To determine a solubility boundary of a given protein/precipitant system, the protocol explained below (section 3) was performed for four to eight droplets of different initial composition, with each condition replicated multiple times (N = 3–6). The lysozyme and RNase A solubility data was well-behaved. The error in each point on the solubility curve was taken to be equal to the change in the reservoir concentration in the iterative process explained below. For the bacteriorhodopsin solubility data, error bars were calculated based off of a 95% confidence interval. The error bars associated with the supersolubility boundary data of lysozyme and RNase A were calculated using a 95% confidence interval based on the uncertainties associated with the slope and intercept resulting from a linear least-squares fit of the data.

3. Protocols, Experimental Results, and Discussion

Structure determination efforts of novel proteins involves first identification of crystal forming conditions, followed by optimization of the crystallization procedure to produce high quality crystals suitable for X-ray analysis. Especially the identification of these “initial hits” is a significant bottleneck in structural biology efforts. While we report here specifically on protocols to map out the phase behavior for a given crystal-forming condition, the evaporation-based platforms used in this work can also be used effectively to screen for initial hits, as we discussed previously.\textsuperscript{52} Exactly as in traditional initial crystallization screening approaches, a wide range of phase space can be screened by starting with an array of droplets of different initial composition. However, the evaporation-based approach allows for each droplet to sample a wider range of concentrations than is screened per droplet in equilibrium-limited approaches. To record this additional information, the droplets need to be monitored continuously (e.g., every 30 min) during the course of an evaporation-based experiment, as enabled by the automated data acquisition setup (Figure 1). A first screen using evaporation-based platforms would hopefully yield some initial hits, as well as preliminary information with respect to phase behavior around those initial hits so that subsequent crystallization optimization experiments can be chosen more rationally.

Following identification of a crystal forming condition (i.e., an initial hit), a three-step process is used to map out both the supersolubility and solubility boundaries for the protein as a function of precipitant concentration. Step 1: Use evaporation to concentrate the droplet until crystals form. This provides an initial estimate of the supersolubility boundary. Step 2: Use reverse vapor diffusion to slowly rehydrate the droplet until the crystals present fully dissolve, thus determining the solubility boundary. Step 3: Repeat step 1 for the same droplet composition at different rates of supersaturation either by starting with dilutions of the original droplet (same C/P but different initial concentrations C\textsubscript{S0}, C\textsubscript{P0}) or by using different, slower rates of evaporation (different A/L ratio). This allows for refinement of the location of the supersolubility curve to identify a critical supersaturation boundary above which nucleation is expected to occur essentially instantaneously.\textsuperscript{57} These three steps are discussed in sections 3.1, 3.2, and 3.3 below. A variation on this method can also be used to decouple nucleation and growth events to rescue overnucleated samples and/or improve crystal quality, as will be discussed in section 3.4.

3.1. Step 1: Estimation of the Supersolubility Boundary.

Starting from a specified initial condition (C\textsubscript{S0}, C\textsubscript{P0} at point 1 in Figure 2b), evaporation causes the concentration of protein and precipitant in the droplet to increase linearly along the line connecting the initial condition with the origin (line of constant C/P = C\textsubscript{S0}/C\textsubscript{P0}).\textsuperscript{51} Because of the absence of a reservoir-associated equilibrium state, as in vapor diffusion crystallization experiments, the concentration in the droplet will continue to increase over time until a phase change such as crystal growth occurs. The use of automated image acquisition at regular time intervals coupled with a known constant rate of evaporation from the device allows for calculation of both the protein and precipitant concentrations at any time. Given that the time it takes for a crystal to nucleate and grow to a discernible size is negligible,\textsuperscript{55,56} the time at which a crystal is first observed can be used to calculate the protein and precipitant concentrations in the droplet at this time (C\textsubscript{S2}, C\textsubscript{P2} in Figure 2b). The state of this droplet thus provides an estimate for the location of the supersolubility boundary which is dependent on the rate of evaporation used. Additional points on the supersolubility curve are determined by performing additional experiments starting with droplets of different protein to precipitant ratio (C\textsubscript{S}/C\textsubscript{P}), as shown in Figure 2c.

Step 1 can also be performed using a traditional vapor diffusion platform.\textsuperscript{2} The reservoir concentration can be increased to drive the sample to adequately high levels of supersaturation for crystals to form. Data acquisition at regular intervals, as needed with the evaporation-based crystallization, is not necessary when using the vapor-diffusion platform. However, the vapor-diffusion approach is potentially much slower, more labor intensive, and potentially inaccurate because of the need to change the reservoir concentrations repeatedly and the time needed for equilibration.

3.2. Step 2: Determination of the Solubility Boundary.

Once crystals have formed in a droplet (Figure 2b, point 2), a reservoir solution is introduced into the chamber, and further evaporation is arrested by sealing the evaporation channel. Contrary to traditional vapor diffusion, the goal of the reservoir solution here is to slowly rehydrate the drop. Thus, a solution with precipitant concentration that is lower than the concentration at point 2 is introduced (C\textsubscript{S3} < C\textsubscript{S2}) and then the droplet slowly equilibrates with the reservoir. The volume of the liquid in the reservoir is much larger than the volume of the protein droplet such that, when the reservoir and drop finally reach equilibrium, the precipitant concentration in the droplet will be equal to C\textsubscript{S3}. The concentration of protein in the solution, while unknown due to portioning of protein between the solution and the crystal, will equilibrate over time to a point on or close to the solubility curve (Figure 2, point 3). Note that crystal growth also could stop before the equilibrium condition at point 3 is reached due to accumulation of surface defects or other factors.
and solubility curves. With droplets of different initial protein/precipitant concentration ratio (1), the droplet first crossed the solubility boundary. (c) Repeating the process of the droplet to move along the solubility boundary, gradually dissolving the reservoir is filled with a precipitant solution of a lower concentration supersolubility boundary into the labile zone. Evaporation is arrested, and the protein in solution will decrease as the crystal grows until the solubility boundary or at worst resulting in the precipitation of the protein in a noncrystalline form. Between these two scenarios, the first is preferable because of the need to identify solubility and supersolubility boundaries presented here has some key benefits, most notably the fact that a single droplet gradually traverses a path through phase space (as opposed to discrete points, i.e., pixilation), and that the concentration of the droplet can be calculated at any point along this path. Researchers who are interested in obtaining phase information, however, may not have easy access to these evaporation-based vapor diffusion crystallization platforms, although it will be more labor intensive than the evaporation-based approach.

Figure 2. (a) Schematic depiction of the process to determine the solubility boundary of a protein/precipitant system by evaporation (step A) followed by an iterative reverse vapor diffusion process (steps B–D). (b) Depiction of the process from part a in a generalized protein/precipitant phase diagram. From the initial state (point 1), evaporation causes the droplet to increase in concentration along the straight line connecting point 1 to the origin (constant $C_P/C_S$). Crystals nucleate and are observed at point 2 after the droplet has traversed the metastable zone and crossed the supersolubility boundary into the labile zone. Evaporation is arrested, and the reservoir is filled with a precipitant solution of a lower concentration than point 2. As the droplet equilibrates with this reservoir, the concentration of protein in solution will decrease as the crystal grows until the solubility boundary is reached (point 3). Subsequent iterations of this process cause the droplet to move along the solubility boundary, gradually dissolving the protein crystal (point 4). The point at which the protein crystal has dissolved completely (point 5) represents the point where the initial path of the droplet first crossed the solubility boundary. (c) Repeating the process with droplets of different initial protein/precipitant concentration ratio ($C_P/C_S$) allows for determination of additional points on the supersolubility and solubility curves.

However, the subsequent steps of the protocol will dissolve the crystal, thus removing these defects and allowing the droplet to reach the solubility boundary. During equilibration, we monitor the droplet and we assume it to be at equilibrium when crystal size does not change anymore. The rate at which equilibrium is reached depends on the nature of the protein/precipitant system. For our systems, we have observed that equilibrium can typically be reached in less than 3 days.

Next, the reservoir solution is replaced with a more dilute solution ($C_{S4} < C_{S3}$) and the process of equilibration is repeated. With each exchange of reservoir solution, the crystals in the droplet will decrease in size, until the crystals dissolve completely (Figure 2, point 5). At this point, the protein concentration can be calculated because the $C_P/C_S$ ratio for the droplet is known, and the salt concentration $C_S$ in the droplet must be equal to the salt concentration in the reservoir, $C_{S3}$. The calculated $C_P$ and the known $C_S = C_{S3}$ determine a point on the solubility boundary. Repeating this process with droplets of different $C_P/C_S$ ratios will provide additional points on the solubility curve (Figure 2c). We used this method to obtain the solubility boundary for the lysozyme/NaCl (Figure 3a) and RNase A/NaCl (Figure 4a) systems. In the case of lysozyme, solubility data is available from the literature, which is in good agreement with our data (Figure 3a).

In a typical experiment, a first estimate of the supersolubility boundary can be obtained within 2–4 days. The time needed to determine the solubility boundary depends on the number of reservoir concentration iterations and the distance that needs to be traversed through phase space to get to the point on the final solubility boundary, i.e., from point 2 to point 5 in Figure 2b. Thus, a full determination of a phase diagram takes on the order of 2 weeks.

Compared to the existing, microbatch- or vapor-diffusion-based methods, the evaporation-based method to determine the solubility and supersolubility boundaries presented here has some key benefits, most notably the fact that a single droplet gradually traverses a path through phase space (as opposed to discrete points, i.e., pixilation), and that the concentration of the droplet can be calculated at any point along this path. Researchers who are interested in obtaining phase information, however, may not have easy access to these evaporation-based well plates and/or an automated imaging system. Fortunately, the above-described protocols to determine the solubility and supersolubility boundary are also amenable to more conventional vapor diffusion crystallization platforms, although it will be more labor intensive than the evaporation-based approach.

The major difference between the vapor diffusion and evaporation-based methods for determining phase behavior occurs during step 1. Whereas a phase change is guaranteed in the evaporation-based approach, the choice of reservoir solution in vapor diffusion affects the end point of the trial. Furthermore, automated imaging of directed evaporation trials provides a reasonable estimate of the state of the droplet when crystals first appear. Because the chance of a vapor diffusion trial precisely hitting an unknown end point is very low, we anticipate two more likely scenarios: (i) the initial choice of reservoir solution is inadequate to drive the formation of crystals and several iterations are necessary to concentrate the droplet to the point where crystals form, or (ii) the initial reservoir solution results in an overshoot of the supersolubility condition, thus requiring at best additional iterations of step 2 to find the solubility boundary or at worst resulting in the precipitation of the protein in a noncrystalline form. Between these two scenarios, the first is preferable because of the need to identify...
crystal forming conditions as accurately as possible. In fact, small changes in reservoir precipitant concentration of the droplet with multiple changes of the reservoir solution would be the ideal way to locate the supersolubility boundary using vapor diffusion. However, this strategy would be very labor intensive and also slower because of the relatively low driving force for evaporation resulting from only small changes in the reservoir concentration. Figure 5 shows the results of this vapor diffusion method for the determination of the solubility boundary for the bacteriorhodopsin/NaH2PO4 system.

3.3. Step 3: Determination of the Critical Supersolubility Boundary. In prior work, we have observed that crystals form at progressively lower levels of supersaturation as the rate of supersaturation is decreased. This trend ends at a critical supersaturation where the level at which crystals are observed becomes independent of the rate of supersaturation.37 This lower boundary of the labile zone occurs when the rate of increase in supersaturation is much smaller than the rate of formation of critical nuclei. To determine a point on the critical supersolubility boundary, we determined the nucleation time for a series of droplets, each with the same $C_P/C_S$ ratio, for a range of different initial protein concentrations. From our previous work, we know that the nucleation time and the initial protein concentration exhibit a linear relationship for slow rates of supersaturation.37 Direct control over the evaporation rate, and thus over the supersaturation rate, as enabled by the platforms used here, allows determination of the critical supersolubility boundary with ease. Indeed, we observe this linear relationship between nucleation time and initial protein concentration for both systems studied here: lysozyme/NaCl (Figure 3b) and RNase A/NaCl (Figure 4b). By extrapolating this line to the point of zero nucleation time, we identify the $C_P$ of the point...
on the critical supersolubility boundary. The $C_5$ of the point can then be calculated from the known $C_P/C_S$ ratio. Repeating this process with droplets of different $C_P/C_S$ ratios provides additional points on the supersolubility curve. We used this method to obtain the critical supersolubility boundaries for the lysozyme/NaCl and RNase A/NaCl systems, as shown in Figures 3a and 4a.

Traditional methods for determining the supersolubility boundary, such as a grid microbatch experiment, utilize a large number of discrete trials at specific $C_P$ and $C_S$ concentrations to map out the phase behavior of a protein. The main advantage of the microbatch method for determining the supersolubility boundary is that the state of each droplet does not change over time and therefore does not need to be determined. However, the strength of the microbatch method is also its weakness in terms of efficient sample use because many more experiments are needed to cover the phase space than the number of experiments required in the evaporation-based method reported here. Other methods such as vapor diffusion, free interface diffusion, or dialysis are able to dynamically sample a larger amount of phase space per experiment but require equilibration before the state of the drop (i.e., the protein and precipitant concentration) can be determined, oftentimes through separate analysis. In contrast to these traditional methods to determine the supersolubility boundary, the evaporation-based method used here can easily sample phase space using only 15 or so droplets (i.e., five conditions repeated three times). Starting from a set of arbitrary initial conditions, the full extent of a phase diagram can easily be traversed and the state of each droplet (i.e., protein concentration, the presence of nuclei or not) can be determined accurately at nearly any point in time. Unlike several of the presently used techniques, this evaporation-based method samples a wide range of conditions in each experiment.

3.4. Protocol for Improving Crystal Quality. Once the phase diagram of a certain protein/precipitant system has been determined (see section 3.3), a modification of the evaporation-based protocols can be used to decouple crystal nucleation and growth, thereby effectively reducing the number of nuclei and improving crystal size/quality (Figure 6). First, a droplet is concentrated by evaporation until it reaches the labile zone where nuclei form (from point 1 to point 2, Figure 6b). As soon as crystals are observed, evaporation is arrested and a reservoir solution of a concentration less than the concentration of the droplet at point 2 is introduced ($C_{S3} < C_{S2}$). Equilibration of the droplet with this reservoir solution causes the dissolution of many of the nuclei that had formed previously, thus decreasing the number of small crystals present in solution. Once equilibrium at this reservoir concentration has been reached (point 3, Figure 6b), the reservoir is drained and the drop is allowed to concentrate again by evaporation (step C, Figure 6b). As long as the drop remains in the metastable region, further evaporation will only drive growth of the few existing crystals.

Figure 5. Solubility boundary for the membrane protein bacteriorhodopsin vs NaH$_2$PO$_4$ obtained using the same protocol as before but with a common vapor diffusion platform. The solid curve was fit through our experimental data to provide a guide for the eye. Data obtained at 23 °C.

Figure 6. (a) Schematic depiction of the method to decouple nucleation and growth events during evaporation-based protein crystallization (step A) followed by reverse vapor diffusion to dissolve some nuclei (step B), and then a second evaporation step to drive crystal growth of the remaining nuclei (step C). (b) Graphical depiction of the process from part a shown in reference to a generalized protein/precipitant phase diagram. From the initial state (point 1), evaporation causes the droplet to increase in concentration along the straight line connecting point 1 to the origin. Crystals nucleate and are observed at point 2 after the droplet has crossed the supersolubility boundary to enter the labile zone. Evaporation is arrested, and the reservoir is filled with a precipitant solution of a concentration between point 2 and where the droplet originally crossed the solubility boundary. As the droplet equilibrates with this reservoir, the number of nuclei present in the droplet will decrease until the solubility boundary is reached (point 3). The well is then opened to evaporation again to allow the few remaining crystal nuclei to grow.
First, while the traditional vapor diffusion strategies have the potential to “rescue” otherwise clear droplets by driving the droplet past its equilibrium end point and hopefully into the labile zone, droplets cannot be rescued from the common occurrence of excessive nucleation using previously reported approaches. In contrast, the method to decouple crystal nucleation and growth reported here allows experiments that exhibit excessive nucleation to be rescued, as exemplified by our results with bacteriorhodopsin (Figure 7c).

Third, the evaporation-based protocol introduced here provides control not only over the end point of each experiment but also over the rate of increase in supersaturation while the droplet resides in the metastable zone, which can have a
profound effect on the number, quality, and size of crystals.\textsuperscript{51} While traditional methods such as vapor diffusion provide some control over the rate of supersaturation in a crystallizing droplet, precise control over these rates is difficult because they vary non-linearly with time.\textsuperscript{54,65} Control over the rate of supersaturation in the evaporation-based platforms here is straightforward, and is linear with time, thus allowing for more ordered growth of crystals.

4. Conclusions

While screening for crystallization conditions is a relatively high throughput exercise, the platform and methods presented here provide a way to make more efficient use of each droplet in a trial. Specifically, we introduced and validated protocols for (i) the determination of protein/precipitant phase diagrams and (ii) the decoupling of crystal nucleation and growth to enhance crystal size and quality. The first set of protocols provides an efficient method for estimating the solubility and for (i) the determination of protein/precipitant phase diagrams in a trial. Specifically, we introduced and validated protocols (ii) the decoupling of crystal nucleation and growth to afford fully automatic filling and operation would further improve crystallization efforts from an art to a science.

Acknowledgment. This work was funded through the NIH Roadmap for Medical Research through grant R21 GM075930-01, as well as NIH grant R01 GM086727, and a NIH Kirschstein Predoctoral Fellowship from the National Institute of Biomedical Imaging and Bioengineering (F31 EB008330). We thank Dr. T. Baybert, Prof. S. Sligar, Dr. D. Berthold, and Prof. C. Rienstra for the bacteriorhodopsin samples.

Supporting Information Available: Details on the design, fabrication, and operation of the evaporation-based crystallization platform, including an AutoCad drawing of the platform, a detailed description of the automated imaging setup, calibration of evaporation rates, and a detailed explanation of the common y-intercept in nucleation vs initial protein concentration plots, are provided. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Chayen, N. E.; Saridakis, E. \textit{2008}, 5, 147–153.
(2) McPherson, A. \textit{Crystallization of Biological Macromolecules}; CSHL Press: 1999.
(3) Hansen, C.; Quake, S. R. \textit{Curr. Opin. Struct. Biol.} \textit{2003}, 13, 538–544.
(4) Green, A. A. J. \textit{Biol. Chem.} \textit{1931}, 93, 495–516.
(5) Green, A. A. J. \textit{Biol. Chem.} \textit{1932}, 95, 47–66.
(6) Haire, L. F.; Blow, D. M. J. \textit{Cryst. Growth} \textit{2001}, 232, 17–20.
(7) Szazki, G.; Kurilaha, K.; Nakada, T.; Miyashita, S.; Komatsu, H. J. \textit{Cryst. Growth} \textit{1996}, 169, 355–360.
(8) Rosenberger, F.; Howard, S. B.; Sowers, J. W.; Nyce, T. A. J. \textit{Cryst. Growth} \textit{1993}, 129, 1–12.
(9) Guilloteau, J. P.; Rieskautt, M. M.; Ducruix, A. F. J. \textit{Cryst. Growth} \textit{1992}, 122, 223–230.
(10) Ciacoppio, E.; Munson, S.; Pusey, M. L. J. \textit{Cryst. Growth} \textit{1991}, 110, 66–71.
(11) Ciacoppio, E.; Pusey, M. L. J. \textit{Cryst. Growth} \textit{1991}, 114, 286–292.
(12) Ries-Kautt, M. M.; Ducruix, A. F. J. \textit{Biol. Chem.} \textit{1989}, 264, 745–748.
(13) Aldabaibeh, N.; Jones, M. J.; Myerson, A. S.; Ulrich, J. \textit{Cryst. Growth Des.} \textit{2009}, 9, 3313–3317.
(14) Annunziata, O.; Payne, A.; Wang, Y. J. \textit{Am. Chem. Soc.} \textit{2008}, 130, 13347–13352.
(15) Ataka, M.; Asai, M. J. \textit{Cryst. Growth} \textit{1988}, 90, 86–93.
(16) Ducruix, A. F.; Ries-Kautt, M. M. Methods \textit{1990}, 1, 25–30.
(17) Ewing, F.; Forsythe, E.; Pusey, M. Acta \textit{Crystalogr., Sect. D} \textit{1994}, 50, 424–428.
(18) Forsythe, E. L.; Pusey, M. L. J. \textit{Cryst. Growth} \textit{1996}, 168, 112–117.
(19) Forsythe, E. L.; Judge, R. A.; Pusey, M. L. J. \textit{Cryst. Eng. Data} \textit{1999}, 44, 637–640.
(20) Howard, S. B.; Twigg, P. J.; Baird, J. K.; Meehan, E. J. \textit{Cryst. Growth} \textit{1988}, 90, 94–104.
(21) Pusey, M. L.; Gernert, K. J. \textit{Cryst. Growth} \textit{1988}, 88, 419–424.
(22) Pusey, M. L.; Munson, S. J. \textit{Cryst. Growth} \textit{1991}, 113, 385–389.
(23) Asherie, N.; Ginsberg, C.; Blass, S.; Greenbaum, A.; Knafo, S. \textit{J. Cryst. Growth} \textit{2008}, 311, 1815–1817.
(24) Ataka, M.; Shinzawaitho, K.; Yoshikawa, S. J. \textit{Cryst. Growth} \textit{1992}, 122, 60–65.
(25) Bergeron, L.; Filobelo, L. F.; Galkin, O.; Vekilov, P. G. \textit{Biophys. J.} \textit{2003}, 85, 3935–3942.
(26) Chayen, N.; Akins, J.; Campbellsmith, S.; Blow, D. M. \textit{J. Cryst. Growth} \textit{1988}, 90, 112–116.
(27) Vaalanta, A.; Uotila, S.; Leisola, M.; Visuri, K. J. \textit{Cryst. Growth} \textit{2003}, 257, 403–411.
(28) DeMattei, R. C.; Feigelson, R. S. J. \textit{Cryst. Growth} \textit{1991}, 110, 34–40.
(29) Mikol, V.; Gieie, R. J. \textit{Cryst. Growth} \textit{1989}, 97, 324–332.
(30) Carbonnaux, C.; Rieskautt, M.; Ducruix, A. Protein \textit{Sci.} \textit{1995}, 4, 2123–2128.
(31) Guacher, J.; Ries-Kautt, M.; ReissHusson, F.; Ducruix, A. FEBS \textit{Lett.} \textit{1997}, 401, 113–116.
(32) Dumetz, A. C.; Chockla, A. M.; Kaler, E. W.; Lenhoff, A. M. \textit{Cryst. Growth Des.} \textit{2009}, 9, 682–691.
(33) Judge, R. A.; Johns, M. R.; White, E. T. \textit{J. Chem. Eng. Data} \textit{1996}, 41, 422–424.
(34) Saridakis, E. G.; Stewart, P. D. S.; Lloyd, L. F.; Blow, D. M. \textit{Acta Crystallogr., Sect. D} \textit{1994}, 50, 293–297.
(35) Veesler, S.; Ferte, N.; Costes, M. S.; Czjzek, M.; Astier, J. P. \textit{Cryst. Growth Des.} \textit{2004}, 4, 1137–1141.
(36) Saridakis, E.; Chayen, N. E. \textit{Biophys. J.} \textit{2003}, 84, 1218–1222.
