



ACADEMIC
PRESS

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Journal of Structural Biology 142 (2003) 170–179

Journal of
Structural
Biology

www.elsevier.com/locate/jysbi

A deliberate approach to screening for initial crystallization conditions of biological macromolecules

Joseph R. Luft,^{a,b,*} Robert J. Collins,^a Nancy A. Fehrman,^a Angela M. Lauricella,^a
Christina K. Veatch,^a and George T. DeTitta^{a,b,1}

^a *The Hauptman–Woodward Institute, 73 High Street, Buffalo, New York 14203, USA*

^b *Department of Structural Biology, State University of New York at Buffalo, 73 High Street, Buffalo, New York 14203, USA*

Received 5 February 2003

Abstract

A method to rationally predict crystallization conditions for a previously uncrystallized macromolecule has not yet been developed. One way around this problem is to determine initial crystallization conditions by casting a wide net, surveying a large number of chemical and physical conditions to locate crystallization leads. A facility that executes the rapid survey of crystallization lead conditions is described in detail. Results and guidelines for the initial screening of crystallization conditions, applicable to both manual and robotic setups, are discussed.

© 2003 Published by Elsevier Science (USA).

Keywords: Crystallization; Macromolecule; High throughput; Microbatch

1. Introduction

Through advancements in molecular biology and the efforts of structural genomics projects, unprecedented numbers of biological macromolecules are being produced in milligram quantities (Structural Genomics Supplement, 2000). X-ray crystallography is the prevalent method of determining the three-dimensional structure of biological macromolecules. Therefore, as the number of samples available in quantities sufficient for structure solution continues to increase, X-ray crystallography will continue to play an ever-increasing, critical role in the understanding of molecular level interactions. Despite numerous advancements in the field, crystallographic structure determination remains a linear process. Several steps must be completed in succession in order to obtain a three-dimensional molecular structure. Often identified as a bottleneck along this pathway is crystallization of the target molecule (Chayen, 2002; Kuhn et al., 2002). Thus far, a method for de

novo prediction of crystallization conditions for biological macromolecules has not been developed.

Although the design of an all-encompassing crystallization experiment that would do an exhaustive, fine screen of all known parameters could be developed, implementation of such an experiment would be entirely impractical. The idealized experiment could potentially produce optimized crystallization conditions for nearly any sample. However, setting up the required number of experiments would require a lifetime to complete for each macromolecule under investigation and an almost limitless supply of sample to be crystallized. The more practical approach: an intelligent design that separates the screening and optimization stages of the crystallization process.

Broken down to the simplest of concepts, crystallization can be thought of as a two-stage process: (1) identify lead conditions for crystallization of the molecule and (2) optimize lead conditions to produce X-ray diffraction-quality crystals. During the first stage a broad net is cast to identify the chemical and physical conditions that will produce outcomes suitable for optimization. We will discuss our efforts to improve the efficiency of the crystallization process by rapid

* Corresponding author. Fax: 1-716-852-6086.

E-mail address: detitta@hwi.buffalo.edu (J.R. Luft).

¹ Also corresponding author.

identification of lead conditions using high-throughput screening (HTS) techniques.

There are two questions that should be asked before setting up even a single crystallization experiment: (1) Is the sample stable and in a solution state that is likely to be amenable to crystallization? (2) What physical and chemical conditions will be explored with the crystallization experiments?

After addressing these initial concerns, the number and type of experiments used to match the sample against a variety of chemical and physical conditions has to be decided upon. When screening for lead conditions, sampling a wide variety of crystallization cocktail conditions is beneficial. Sampling a larger number of cocktails is more likely to produce not a *single* lead condition, but *numerous* lead conditions. Assuming crystallization parameters are optimized from an assortment of chemical variables, the crystals produced are likely to exhibit polymorphism and have diverse physical properties. This will be favorable when addressing soaking, cryoprotection and diffraction issues: it is desirable to provide more than a single choice for optimization. Differences in diffraction quality and ease of optimization will be evident for crystals grown under different chemical and physical conditions. For these reasons, using an approach with an ample number of crystallization cocktails may be advantageous.

The method used in the search for crystallization conditions must exhibit dual efficacy: efficient in both time and volume requirements. The less time delay between a sample's purification and its combination with crystallization cocktails, the less likely it is that problems with homogeneity, aggregation state, or activity loss will plague the crystallization trials. Reducing the volume of the sample will have an obvious benefit, requiring less of the sample to be purified for the initial crystallization trials.

After setup of the crystallization experiments is complete, the results must be recorded and analyzed. Manual inspection and recording of the results in a notebook is one option. Images of the outcomes of a crystallization experiment can be tremendously descriptive both for novice and for experienced investigators. Adding images to the pages of a notebook can be very advantageous, creating a permanent record that is frozen in time, of the state of the experiment. If imaging is used as a method to record the outcomes of the experiments, analysis of those results must be completed before any advancement in the determination of the crystal structure. However, simply setting up and recording a multitude of experiments will not advance any field of study. A single experiment is no better than a great number of experiments until analysis and follow-up on the results take place. Manual image analysis is still the most reliable method available to resolve the outcomes of crystallization experiments (Abola, 2002). Enhancements to manual scoring, through software that

allows more effortless viewing of the outcomes, or through image analysis software that enriches images with respect to desirability of outcome are both helpful. Outsourcing of the image analysis to allow parallel processing by many investigators is another alternative.

The ability to communicate the results of the screening experiments both intra- and interlaboratory offers another potential roadblock. Correspondence between groups that are separated geographically can present a great difficulty without adequate, reliable, clear methods of communication to share both the data and the results. The use of a database for intralaboratory communication is critical, eliminating repeated entry of information that contributes to the propagation of error. A robust computer infrastructure capable of reliably handling large volumes of data with fail-over systems and backups is imperative for success.

A concluding, fundamental ingredient for the development of a laboratory capable of handling a large number of samples on an unvarying schedule is system redundancy. It is absolutely imperative to maintain backup systems in a laboratory of this nature. A single point of failure is unacceptable and will lead to downtime that cannot be allowed to occur. Samples of biological macromolecules are for the most part inherently unstable; a delay in setting up crystallization experiments is undesirable. If the recording process for the experiments is hindered in any manner, it is impossible to simply go back and rerecord the outcomes, as they change over time. Redundant systems are critical for both the mechanical systems and the personnel of a HTS crystallization laboratory.

2. Description of the HTS crystallization conduit

The technical details of the rapid survey of crystallization lead conditions are based upon the formerly described guidelines. Although there is not a refined process to predict the crystallization parameters of a previously uncrystallized macromolecule, a logical and deliberate progression can take place to determine these conditions in a sensible and efficient manner.

2.1. Sample preparation

The importance of preparing a sample for crystallization trials cannot be overemphasized. It has long been recognized that if a sample suffers from batch-to-batch variation, aggregation, or microheterogeneity, it will likely prove very difficult to crystallize (Giegé et al., 1986). Time spent preparing a sample so that it appears monodisperse by light scattering analysis, remains stable under storage conditions, and does not suffer from loss of activity will save countless hours and lessen frustration when searching for lead and optimized

crystallization conditions. The time required to set up crystallization experiments is minimal compared with the time spent analyzing and trying to follow up on the results of those experiments.

Optimizing crystallization leads for a protein sample that is constantly changing is very counterproductive. Batch consistency should always be considered a top priority. Batch-to-batch variation in a sample will lead to results that are almost impossible to optimize.

Our studies with one group of collaborators, and published reports (Zulauf and D'Arcy, 1992), show a high likelihood of determining lead conditions for crystallization when a sample is prepared in a solution environment that contains a single aggregation state by light scattering analysis. This is perhaps the single best criterion to use as a prescreening measure of the likelihood that a sample will crystallize. It is wise to place at least some effort into ensuring that a single aggregation state exists for a sample about to undergo crystallization trials. After the desired monodisperse solution state is achieved, crystallization trials and outcomes will have a much higher probability of success.

2.2. Crystallization cocktail selection

Why could it be advantageous to set up 1536 crystallization screening experiments? Setting up more conditions means you can sample a wider variety of chemical conditions. The choice of crystallization screening solutions is largely historical. A search to determine all of the best salts and polymers for the crystallization of biological macromolecules has never occurred. Chemicals that are suitable for the crystallization of biological macromolecules are revealed in the literature as they are discovered. The options to push established chemical boundaries for crystallization, increase the probability of locating initial crystallization conditions, and determine multiple lead conditions all exist when 1536 solution conditions are available for crystallization trials. That said, several cases exist in which a single chemical cocktail from the 1536 screened provided a lead condition suitable for optimization. We have had numerous cases in which a single sample produced 50 or up to 300 lead conditions for optimization screening from the 1536 solution screen. Multiple leads translate to multiple chemical conditions. Crystals of biological macromolecules produced from different chemical environments will exhibit differences in diffraction quality, suitability for cryopreservation, and suitability for soaking experiments (solvent channel volume). It can be a great benefit to have a variety of solution conditions to pursue in optimization trials.

The screening solutions used in our laboratory can be broken down into three main categories: (1) salt, buffer (36 salts at three concentrations combined with eight buffers); (2) PEG, salt, buffer (eight PEGs at two con-

centrations, combined with 36 salts and eight buffers); and (3) commercially available screening solutions. Crystallization cocktails from categories 1 and 2 were originally created using an incomplete factorial design (Carter, 1990). The cocktails are updated annually, replacing solutions that have proven unstable or otherwise difficult to use in the laboratory. Cocktails are also replaced when they are not performing well in the search for crystallization lead conditions. The screening solutions are in a constant state of evolution with the goal to produce stable cocktails with a high probability of providing lead conditions for crystallization of biological macromolecules.

2.3. Crystallization experiment setup

Microbatch-under-oil (MUO) is a particularly efficient crystallization method, well suited to determining lead conditions (Chayen et al., 1992). The method has been successfully used to crystallize both soluble and membrane-bound proteins (Chayen, 1999). A MUO experiment is elegant in its simplicity. Oil surrounds an aqueous solution that contains: (1) crystallization cocktail (to limit the solubility of the sample to be crystallized) and (2) concentrated solution of a biological macromolecule. To conduct a MUO experiment one simply has to choose a suitable container and add oil, crystallization cocktail, and protein solution (Fig. 1). There is a distinct difference in the approach to supersaturation between batch and vapor diffusion methods (Fig. 2). This difference in the approach to supersaturation explains findings in the literature that describe groupings in the discovery of lead conditions for crystallization when using a batch versus a vapor diffusion approach to the problem (Stewart et al., 1996).

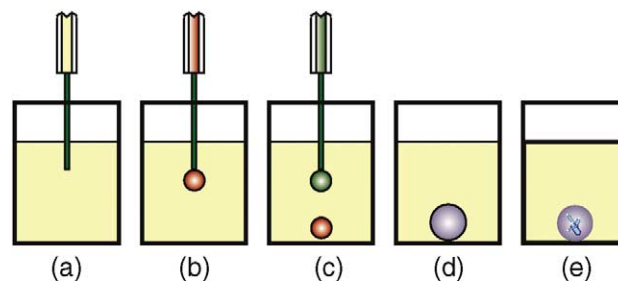


Fig. 1. A microbatch-under-oil crystallization experiment is broken down to its most basic steps: (a) a suitable container is selected and oil is added to the container (the type of oil can be selected or mixed to regulate the rate of water transport from the aqueous experiment drop; paraffin oil allows little water transport, silicon oils allow water transport); (b) crystallization cocktail solution is delivered to the oil (the cocktail solution should ideally decrease the solubility of the sample being crystallized, driving it to a solution state of supersaturation); (c) a solution of the sample to be crystallized is delivered to the oil; (d) the cocktail and sample drops merge (a centrifugation step can aid this process); (e) under ideal conditions a crystal forms in the aqueous drop protected by an oil barrier.

The MUO approach offers several distinct advantages that include the following.

2.3.1. Small sample volumes

Batch experiments are particularly efficient in their use of solutions to set up crystallization experiments. The MUO experiments use a comparatively small volume of crystallization cocktail solution, generally requiring a volume of cocktail on scale with that of the sample to be crystallized. This contrasts with the typical vapor diffusion approach that will require a significantly larger volume of cocktail solution to be placed in a reservoir to invoke the controlled dehydration of an experiment drop. Another favorable situation is the presence of oil as a container for the experiment. The oil acts as a wick to remove small-volume drops from delivery tips. This allows the routine delivery of solution volumes as low

as 200 nl. There is no rush to deliver solutions to avoid dehydration. Once solutions are delivered to the oil they are protected from significant dehydration. This allows plates containing only oil and a small volume of cocktail solution to be stored for a period of time prior to the addition of the sample to be crystallized. The small-volume sample and cocktail drops can be merged under the oil using a simple centrifugation step.

2.3.2. Minimal manipulation

There is only one part required to set up a MUO crystallization experiment, namely a container. The oil, cocktail, and sample are all delivered to that container. There are no additional pieces that demand manipulation; the experiment does not depend upon a sealing procedure for its success, as the oil protects the aqueous drop from dehydration.

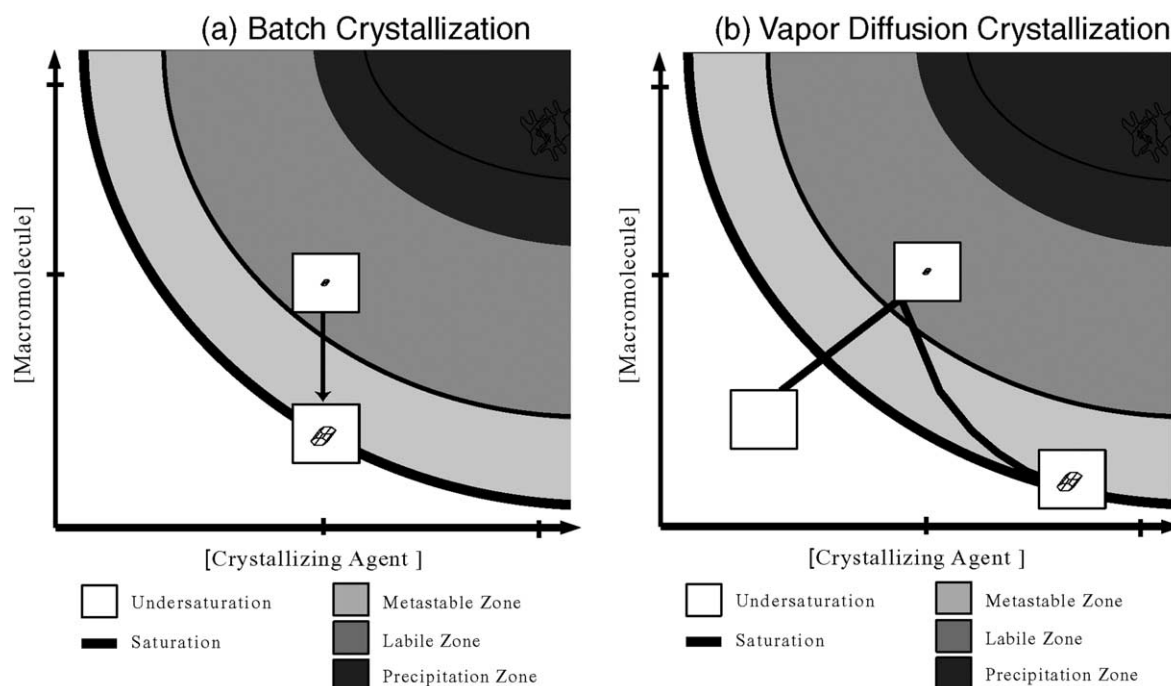


Fig. 2. The difference between batch and vapor diffusion crystallization experiments comes down to the approach to supersaturation. (a) Batch experiments are stagnant. The starting conditions in the experiment drop will not change unless a component precipitates from the solution. A batch crystallization experiment starts off with both cocktail solution and sample to be crystallized at a concentration of $2\times$. After the samples are combined (assuming for this example in a 1:1 volume ratio) the final concentration of both cocktail solution and sample in the experiment drop will be at $1\times$. This $1\times$ concentration will have to fall in the labile region of supersaturation if no outside influence triggers crystal nucleation and subsequent crystal growth. For the experiment to successfully produce crystals the concentration of the cocktail and sample must fall into a narrow range of solubility. (b) Vapor diffusion crystallization experiments follow a different, and distinct, path to supersaturation. The beginning of a vapor diffusion experiment is identical to a batch experiment. Both cocktail solution and sample to be crystallized begin at a concentration of $1\times$. After the two solutions are combined in a 1:1 volume ratio (for this example) the final concentration of both cocktail solution and sample in the experiment drop will be at $1/2\times$. The experiment drop will equilibrate against a reservoir solution that is typically much larger in volume than the experiment drop. The reservoir solution volume dominates the endpoint concentration in the experiment drop. As the vapor pressure of water over the experiment drop (high) equilibrates with the vapor pressure of water over the reservoir solution (low) water will be transported through the vapor phase to the reservoir solution. The loss of water from the experiment drop will decrease the volume of the drop. This decrease in drop volume will lead to a relative increase in the concentration of any nonvolatile components in the experiment drop. The experiment solution will (if conditions permit) achieve continuously increasing states of supersaturation until either vapor phase equilibration with the reservoir solution is reached or the chemical components of the experiment drop precipitate from solution. If solution conditions permit crystallization, the macromolecule solution concentration will decrease, as macromolecules become part of the growing crystal. Eventually, a state of equilibration is achieved in the drop between the solid crystal phase and the surrounding liquid.

2.3.3. Robust setups

A MUO experiment is inherently stable. The experiment drops will not be destroyed by physically jarring the crystallization tray and temperature fluctuation will not produce condensation. This allows the experiments to be moved to a variety of incubator environments with no negative effects caused by limitations or incompatibilities of the technique. The oil also acts to stabilize the experiment drop from short-term temperature fluctuations, normalizing the thermal environment of the experiment drop. This effect is similar to that used by manufacturers of temperature monitoring devices for incubators, inserting the thermometer into a vial containing oil to reduce temperature fluctuation on the monitor.

2.3.4. Adaptable to a variety of storage solutions

Almost any container can be used to hold MUO experiments. An answer to the packaging question should consider the ease of labeling and storage and ease of setup and have a shape that is amenable to viewing the results. A convenient format for storage of MUO experiments is a 96-well footprint plate. An enormous selection of plates are available with a multitude of well shapes and volumes, they are standardized to uniform footprints, manufactured from a variety of materials, and widely available.

2.3.5. Easily scaled up to a high-throughput model

Using an incarnation of mother–daughter plate technology that is very common in other areas of research, the MUO experiment is easily adapted to a HTS model. Through progression from a low- to a high-density plate format, it is possible to efficiently set up thousands of crystallization experiments in a reliable, repetitive manner using well-established protocols. Source to destination plate delivery protocols are easy to implement with MUO. Mother plates of cocktail solutions can be prepared in advance and delivered to experiment plates containing oil as needed. The small sample volume requirements and compactness of containers that are suitable for MUO make ultra-HTS of crystallization conditions a readily achievable goal.

2.3.6. Control of evaporation rates

Paraffin oil has low water permeability and will allow very little evaporation of water from the experiment drop. Silicon oil has higher water permeability and will allow rapid transport of water from an experiment drop, leading to its dehydration. By selecting a combination of paraffin and silicon oils, the rate of evaporation of water from an experiment drop can be regulated (D'Arcy et al., 1996). This regulation is relative to the plate storage environment. If the experiments are stored in a humid environment, evaporation will be slower; if stored in a dry environment, evaporation will be more rapid. Like-

wise, if stored at low temperature, the vapor pressure of water in the experiment drop will decrease, limiting evaporation from the drop; whereas storage at a higher temperature will increase the vapor pressure of water in the drop and lead to more rapid water transport.

3. Setting up the experiments

The procedure for setting up the crystallization experiments is very standardized. The process begins well in advance of the experiments. Collaborators are contacted to establish the number of samples that will be processed during a particular run. After calculating the number of samples preparation for a run will begin and plates set up will progress in three stages. We employ an Apogent Discoveries/Robbins Scientific Tango liquid handling system (LHS) to deliver (1) paraffin oil, (2) 1536 unique crystallization cocktail solutions, and (3) sample to be crystallized to a 1536-well microassay plate.

The Tango (Fig. 3) is a robotic LHS that is well suited to the setup of MUO crystallization experiments. Some considerations that led to the decision to use the Tango LHS were the following.

3.1. Syringe bank-based delivery system

The Tango uses a variety of exchangeable syringe banks (1, 12, 96, and 384) to deliver solutions. This is critical in an application in which both time constraints and sample volumes need to be optimized for various stages of the delivery process. Increasing the number of syringes reduces the number of deliveries that have to be made to reformat a given plate, decreasing the time required to complete a given operation. This increase in syringes comes at a cost, increased dead-volume loss (associated with each syringe) and a greater number of mother-plate wells that must be manually filled (single-well mother plates are not a viable option for sample solution due to increased volume requirements).

3.2. Permanent versus disposable tips

An issue that requires a great deal of consideration in choosing a LHS for setting up crystallization experiments is what type of tips should be used. Disposable plastic tips provide a measure of security. They are unused, eliminating any risk of contamination caused by inadequate cleaning protocols. However, the number of tips used in a HTS laboratory can quickly reach unmanageable levels. The added difficulty of delivering volumes as low as 200 nl with disposable tips is another issue that must be addressed. After consideration of these factors a system featuring permanent syringes rather than disposable tips was chosen. The use of syringes comes at a price. The Teflon tips that form the seal



Fig. 3. The Apogent Discoveries, Robbins Scientific Tango liquid handling system is very well suited for delivery of solutions to set up 1536-well plates containing microbatch-under-oil crystallization experiments. The LHS is made up of an exchangeable delivery head. The head can hold an array of 1, 12, 96, or 384 syringes. The syringes work in unison to aspirate or dispense solutions from a source to a destination plate. The plates are held in position on an *xy* translation stage that moves beneath the syringe bank head. The entire system is controlled through a GUI that permits the user to define all aspects of the delivery protocol.

within each syringe must be changed on a monthly basis with our pattern of use. Great care must be taken to ensure there is no carryover of even trace amounts of contamination when using a system that does not incorporate disposable plastic tips. Cleaning protocols that are currently in place are exhaustive to make certain no chance of cross-contamination occurs.

3.3. Speed

The Tango system's rate of stage translation produces relatively fast delivery times even with reduced numbers of syringes. Using a 12-syringe bank head for protein delivery the system will aspirate from 12 mother-plate wells containing the protein solution and make 128 identical deliveries to a 1536-well experiment plate containing oil and cocktail solutions. The system completes this reformatting from a 12-well protein array into

a 1536-well plate array in 8 min. Speed and accuracy during this stage of the process are critical. Variation in delivery volumes of the sample solution will lead to results that are difficult to reproduce. Slow delivery times increase the likelihood of sample integrity issues before crystallization trials can begin.

3.4. Precision and accuracy

The capability to deliver solutions with varying viscosity is an important consideration in the setup of crystallization experiments. The 1536 crystallization cocktails were measured and found to have viscosity values ranging from 1 cP to as high as 653 cP at 25 °C. To put this in perspective, the viscosity of heavy machine oil is 233 cP at 20 °C. For a delivery system to be capable of handling solutions with such a range of viscosities certain flexibilities in the delivery protocols are required. The Tango allows the rates of both descent and ascent of the syringe needles from the solution to be controlled. Also, through software control of the plunger speed in the syringes, the Tango has a user-defined rate of aspiration and delivery of solutions. The LHS has been used in our laboratory to deliver solutions of widely varying physical parameters ranging in volume from 200 nl to 30 μ l.

The *xy* translation stage of the Tango contains 12-plate position holders. The translation stage moves underneath the syringe banks with a reported positional accuracy of $\pm 5 \mu$ m. This exceptional accuracy keeps the syringes lined up with well positions even on a 1536-well plate without the need for constant, time-consuming system alignment.

3.5. Automated delivery and cleaning cycles

The 12-position stage on the Tango allows automated delivery and wash cycles. Placing a 96-well plate containing a 12-well array of protein solution and a 1536-well plate containing oil and cocktail solutions on the stage is required to set up a MUO experiment. The additional plate positions on the stage permit automation of the wash cycle for the LHS between protein deliveries. By placing trays of wash solutions, waste containers, and blotter paper on the stage the wash cycle becomes fully automated. This makes the setup of experiment plates much less user-intensive, providing time to perform more intellectual tasks.

Preparation of the MUO experiment plates is a multistage process requiring the delivery of (1) oil, (2) crystallization cocktail, and (3) sample at separate times.

3.5.1. Oil delivery

The oil delivery requires reformatting of a single-well mother plate containing paraffin oil to an empty 1536-well microassay plate (Greiner 783101). During this

portion of the setup full use is made of the multiple plate positions on the Tango stage. The LHS is programmed to aspirate oil from a single-well mother plate and make multiple identical deliveries of 5 μ l of oil to each well of the experiment plates. Dead-volume losses are not an issue, making a bank of 96 or 384 syringes preferable for this type of delivery.

The goal at this stage is to set up replicate plates with minimal deviation in delivery volume and with minimal human intervention in as short a time as possible. The scheduling for oil delivery is very accommodating. Plates can be prepared far in advance of the crystallization experiments. Plates are often prepared with oil in a *burst mode*, creating a stockpile of oil-filled plates. If paraffin oil is used there will be little if any oil creep, providing stable storage. Silicon-based oils and combinations of silicon and paraffin oil will tend to creep out of the polystyrene plates due to surface tension effects. This property can cause uneven volumes of oil in each well and difficulties with imaging and lead to an undesirable, oily laboratory environment. The benefit of using a silicon oil mix (more rapid dehydration of the experiment drop) can be significant (D'Arcy et al., 2003) and for many researchers this offsets the negative effects.

3.5.2. Crystallization cocktail: solution delivery

The import of a mother–daughter plate experiment design from other areas of research was critical for the successful application of HTS to crystallization in our laboratory. In no area is that more evident than in the delivery of cocktail solutions. Cocktails are initially prepared in 10-ml volumes and stored in scintillation vials. The 1536 cocktail solutions are manually transferred to 16 unique 96-well deep-well mother plates. Subsequent reformatting takes place using an automated LHS. Sixteen 96-well plates are reformatted to four 384-well plates using a 96-syringe bank head on the LHS. The 384-well mother plates are sealed and stored at -20°C for up to 3 months. Scheduling for this process, like that of the oil delivery, is very flexible. Interweaving of the cocktails will occur as they are reformatted from 16 96-well plates to 4 384-well plates; the syringes are spaced to deliver to every other reservoir on the 384-well plates. The rationale behind this reformatting is the decrease in the number of mother plates required to make deliveries of cocktails to the 1536-well experiment plate. Although it would be plausible to use 16 96-well mother plates and deliver them directly to the 1536-well experiment plate, there is an increased likelihood of human or machine error occurring during the reformatting process. An intermediate reformatting step from a 96- to a 384-well mother plate has proven much more efficient.

Cocktail solutions now have to be reformatted from the four 384-well mother plates to the experiment plates containing oil-filled reservoirs. This reformatting step

will interweave the cocktail solutions. A 384-syringe bank head is used for this stage of the delivery process. Each of the four cocktail mother plates will be used as source plates in succession. Experiment plates will have cocktail solution delivered to them in clusters. All of the experiment plates will receive cocktail solutions from the first mother plate; the LHS will go through a cleaning cycle, followed by the delivery of cocktail solutions from the second mother plate, until all cocktails from all four source plates have been delivered. Managing the deliveries in this manner minimizes cleaning cycles, human intervention, and the opportunity for a human- or machine-based error to occur. Each plate is centrifuged after cocktail delivery to make certain the drop is at the bottom of the well, under the paraffin oil to protect it from dehydration.

Experiment plates are typically set up with 200-nl aliquots of cocktail solution. This volume is manageable by the LHS, large enough to produce results that can be viewed by users and (if using equal volumes of protein and cocktail solution) small enough to permit 1536 experiments to be set up with $<400\mu\text{l}$ of protein stock solution.

At this stage the plates have a limited shelf life. Plates can be stored for up to 2 weeks at 4°C or up to 1 month at -20°C . Storage of the plates containing oil and cocktail at higher temperatures or for longer periods of time leads to an unacceptable amount of water loss from the cocktail solutions; producing salt crystals in a number of the cocktail drops.

3.5.3. Sample: solution delivery

The final stage of setting up the crystallization experiment is the delivery of the sample under investigation. For this step we need to optimize two competing fields: dead-volume loss and time of delivery. By increasing the number of syringes used in the LHS we can decrease the robot's delivery time with a subsequent increase in both the dead-volume loss and the time required for manual pipetting to array the protein into a source plate. Reducing the number of syringes used in the LHS will reduce dead-volume loss and increase delivery time by the LHS. Our original LHS used a 96-syringe bank head for delivery of the protein solutions. Manual expression of the 96-well mother plate proved tedious. A 12-syringe bank head has proven much more effective for this stage of the process. The manual pipetting is minimal (12 wells of a 96-well plate), sample requirements have decreased from $600\mu\text{l}$ using a 96-syringe bank head to $400\mu\text{l}$ using the 12-syringe bank head. The time required by the LHS for delivery of the sample solution is 8 min per plate, comparable to the delivery time with the 96-syringe bank head on the older LHS. Any time gained by the LHS will be offset by the time required to manually format a 96-well mother plate. The protein delivery is accomplished in 128

identical deliveries of 200-nl volume from the 12-syringe bank head. Fifty microliters of sample is retrieved from the mother plate after the experiment's completion. The Tango continues with an automated cleaning cycle after delivery of the protein solution to the experiment plate. The system, as currently configured, is capable of delivering protein to two 1536-well experiment plates per hour with minimal user intervention.

4. Recording the results

Although the MUO experiments arrayed in a 1536-well plate are undeniably efficient in their construction, they are not ideal candidates for manual observation under a low-power light microscope. It is difficult to discern the location of a particular well on a plate without intense calculation. The reformatting of the cocktail solutions as they are delivered to the experiment plate adds further difficulty to the process. The process of manual inspection is an arduous task for a single 1536-well plate. To keep up with 100 such plates a month would be impossible. For this reason automated methods for recording the outcomes of the experiments were required.

Two plate-imaging systems are in place in the HTS crystallization laboratory to record the results of the experiment plates over a 1-month time course. Each imaging table has a capacity of 28 1536-well experiment plates. Every plate contains a single biological macromolecular combined with 1536 unique crystallization cocktails. This corresponds to a capacity of 43 008 experiments on each imaging system (times 2 systems equals 86 016 experiments in all). The images are recorded at a rate of 2 plates per hour per table.

The imaging tables consist of a 30-in. *xy* translation stage controlled by stepper motors. Software was written in-house (Spot) to control the reader tables. The software accepts unique plate identifier codes as input for each position on the table. The user selects the plate orientation (upside down or right-side up) and defines three corner positions on each 1536-well plate using a joystick with software assistance (Fig. 4). Images are recorded in TIFF format using a stationary 12-bit, cooled, black and white, 1.3-megapixel, IEEE1394 digital camera and transferred through a gigabit Ethernet connection to a fileserver for further processing.

Experiment plates are recorded during the following times:

- (1) just prior to the addition of the macromolecule,
- (2) immediately after the addition of the macromolecule solution,
- (3) 1 day after setup,
- (4) 3 days after setup,
- (5) 1 week after setup,
- (6) 2 weeks after setup, and
- (7) 1 month after setup.

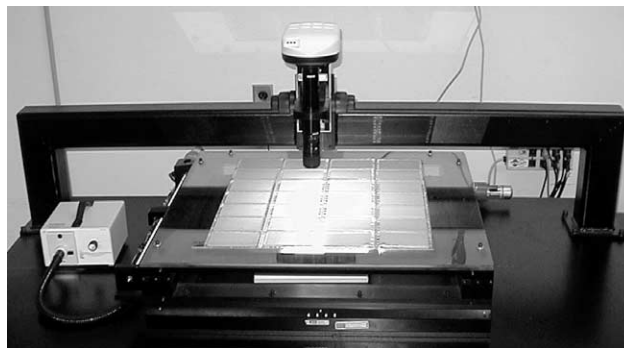


Fig. 4. The plate-imaging tables consist of an *xy* translation stage controlled by stepper motors. Each table has a maximum capacity of 28 1536-well plates. Images are recorded in TIFF format using a 12-bit, cooled, black and white, 1.3-megapixel digital camera that is positioned above the plates. A telecentric zoom lens on the camera views directly down the center of each well. Lighting for the system comes from a light pad with IR filter to protect the samples from heating.

Recording the plate just prior to the addition of sample often proves valuable during analysis of the results. If crystals form in the cocktail solution prior to the addition of sample, or a small piece of debris is present in the drop, these sources can be delineated from the outcomes that occur after the addition of sample. This information can prove critical when trying to determine the composition of a crystal when only image data are available.

5. Computer resources for data storage and distribution

Significant and necessary upgrades have taken place in the computer resources for the laboratory. At its inception the laboratory was able to function with a single PC. Since that time the number of samples, number of images, and volume of data generated have dramatically increased the requirements of the computer resources.

A two-tier database was written in-house using a PHP front end and a MySQL back end. User interface takes place through a standard Web browser. This was an important consideration for us. We envision access to the database by users for data entry on samples being submitted for crystallization screening. This access should not require the outside user to purchase special software. Using a standard Web browser interface was the only reasonable solution to this potential difficulty. The database integrates all information in the laboratory. These data include sample receipt, user information, sample information, imaging, and repackaging of image data, as well as distribution and storage. A secure, robust database is an absolute requirement for the HTS laboratory. The tie-in with all other systems is critical and reduces the chance of incorrect data entry through reduction. Data only have to be entered into the system once.

A fail-over machine has been installed to take snapshots of the 0.55 TB RAID5 files server, twice a day. This allows files that were mistakenly deleted from the RAID array to be retrieved for up to 30 days. The drive array connected to the files server is protected by its very nature with the presence of a parity drive. This parity drive allows data to be reconstructed in the event that one of the nine hard drives making up the array fails. The data are also backed up on a 30-tape 1.8 GB AME tape library system every 6 h. Archive tapes are written after imaging is completed on 100 experiment plates. This level of protection is critical; if image data are lost it is impossible to reobtain the image. Images are a time-based recording of the outcomes of the crystallization experiments. The experiments change over time, making it impossible to go back and record a plate's 2-week image once that time has passed. An optical disc recorder with thermal printer provides automated production of CDs containing image data for distribution to investigators. A second system outfitted with a DVD recorder provides a redundant system and allows us to record TIFF images for in-house backup and distribution to collaborators who are developing image-processing algorithms.

An ftp server allows images from plate recordings to be available to users within 24 h of being recorded. Every user of the facility is also sent a CD-ROM for each sample submitted for HTS. The CD-ROM contains all of the recordings of the sample and the image viewing software, MacroScope.

A dual AMD Athlon 1900+ computer server has also been added to the system. The server is used to run the image-processing algorithms. IBM personal computers are used and networked throughout the laboratory to run all of the robotics systems.

6. Analyze the results

Ideally, outcomes of the results from the crystallization screening experiments would be reviewed and promising crystallization leads separated from the bulk of the outcomes like wheat from chaff. A reliable real-time image processing software approach is not yet available. The authoritative reviewer of images captured to describe the outcome of a crystallization experiment is still a human expert.

Easing the burden on a human reviewer and setting up the images so they display the maximum amount of information were goals in the development of MacroScope. MacroScope is a software package that is distributed to users of the HTS crystallization facility along with packaged image results from the 1536-cocktail crystallization. The software takes RAR-compressed JPEG images and displays them in 16 groups of 96 thumbnail images (Fig. 5). The thumbnail images are arrayed in a logical design that relates back to the cocktail arrangement in the original mother plates. The images are displayed in groups that are related by



Fig. 5. MacroScope is software designed to view the results of the 1536 microbatch-under-oil crystallization experiments. The software displays 16 groups of 96 thumbnail images. When an image is selected it will expand to full size and provide information on the time the image was recorded and the chemical information about the cocktail solution.

cocktail components to maximize visual information from related experiments. By selecting a thumbnail image for closer inspection, the full-sized image is displayed. The full-sized image displays a time–date stamp and well codes along with chemical information about the cocktail. MacroScope also provides the option of selecting particularly interesting outcomes using a checkbox. After reviewing the images, the selected data can be saved to a file for review at a later time. The selected images can be viewed over the history of the recording process, viewed as a group, or saved in a Web page format. Crystals as small as 15 μm can be clearly distinguished in the JPEG images.

7. Discussion

The HTS crystallization laboratory is capable of processing 100 samples per month. To date, after a 3-year period, 2.1 million experiments have been performed and the results recorded in the form of 12 million digital images. More than 270 researchers provided samples of 1300 proteins and nucleic acids for these experiments. The data generated from these experiments are stored in a database with the goal of developing a way to rationally predict crystallization conditions for previously uncrystallized macromolecules.

The design of a laboratory capable of annually setting up and recording the results of a few million crystallization experiments could not proceed without a great deal of consideration and common sense. The same rules that hold true for anyone attempting to crystallize a biological macromolecule will hold true for a HTS crystallization laboratory. All aspects of the experiment must be considered and broken down to determine how to use a HTS approach in a manner that will be effective. It is not enough to simply do things different from, or faster than, someone else, you have to show some advantage to an approach. In the case of HTS, we feel there are several true advantages that are not available to an investigator working with standard manual protocols. The speed at which crystallization cocktails can be combined with freshly prepared protein solutions provides a HTS laboratory with a distinct advantage, with less time for the sample to decay prior to the completion of crystallization trials. Another distinct advantage HTS offers is the ability to set up 1536 unique crystallization cocktails with each sample. Although it may be argued that a smaller number of cocktails is sufficient, it is difficult to argue the benefit of having an assortment of chemical conditions that will produce crystals of a sample under study. Unique crystal growth environments will likely produce crystals with dissimilar physical properties. These crystal forms

will have varying degrees of suitability for different tasks. A HTS approach allows researchers the benefit of temporal flexibility in viewing outcomes of experiments. By reviewing image data collected as the experiment progresses, a nonsubjective record exists that allows comparison and monitoring of outcomes of the crystallization experiments at any time. Finally, we believe these records and data will permit data mining and discovery of information about the growth of biological macromolecule crystals that may help to further rationalize a process that creates an unnecessary bottleneck in X-ray crystal structure determination.

Acknowledgments

This work was supported in part by NASA Grants NAG8-1594, NAG8-1839, and NCC8-232; through National Institutes of Health NIH GM-64655 (Structural Genomics of Pathogenic Protozoa) and NIH GM-62413 (Northeast Structural Genomics); and through The John R. Oishei Foundation, The Cummings Foundation, and The Western New York Foundation.

References

- Abola, E., 2002. Syrrx's High Throughput Protein Structure Determination Pipeline—Reshaping the Structure Based Drug Design Landscape. 60th Pittsburgh Diffraction Conference, Pittsburgh, PA.
- Carter Jr., C.W., 1990. Efficient factorial designs and the analysis of macromolecular crystal growth conditions. *Methods: Companion Methods Enzymol.* 1, 12.
- Chayen, N.E., 1999. Crystallization with oils: a new dimension in macromolecular crystal growth. *J. Cryst. Growth* 196, 434.
- Chayen, N.E., 2002. Tackling the bottleneck of protein crystallization in the post-genomic era. *Trends Biotechnol.* 20, 98.
- Chayen, N.E., Shaw Stewart, P.D., Blow, D.M., 1992. Microbatch crystallization under oil—a new technique allowing many small-volume crystallization trials. *J. Cryst. Growth* 122, 176.
- D'Arcy, A., Elmore, C., Stihle, M., Johnston, J.E., 1996. A novel approach to crystallising proteins under oil. *J. Cryst. Growth* 168, 175.
- D'Arcy, A., Sweeney, A.M., Stihle, M., Habner, A., 2003. The advantages of using a modified microbatch method for rapid screening of protein crystallization conditions. *Acta. Cryst. D* 59, 396.
- Giegé, R., Dock, A.C., Kern, D., Lorber, B., Thierry, J.C., Moras, D., 1986. The role of purification in the crystallization of proteins and nucleic acids. *J. Cryst. Growth* 76, 554.
- Kuhn, P., Wilson, K., Patch, M.G., Stevens, R.C., 2002. The genesis of high-throughput structure-based drug discovery using protein crystallography. *Curr. Opin. Chem. Biol.* 6, 704.
- Stewart, P.S., Baldock, P., Mills, V., 1996. A comparison of microbatch and vapour diffusion for initial screening of crystallization conditions. *J. Cryst. Growth* 168, 170.
- Structural Genomics Supplement, *Nat. Struct. Biol.* (2000).
- Zulauf, M., D'Arcy, A., 1992. Light scattering of proteins as a criterion for crystallization. *J. Cryst. Growth* 122, 102.