Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Initial evaluations of the reproducibility of vapor-diffusion crystallization

Janet Newman,[a]* Jian Xu[b] and Michael C. Willis[b]

[a]Molecular and Health Technologies, CSIRO, 343 Royal Parade, Parkville, VIC 3052, Australia, and [b]Rigaku Automation, 5999 Avenida Encinas, Suite 150, Carlsbad, CA 92008, USA

Correspondence e-mail: janet.newman@csiro.au

Received 22 April 2007
Accepted 27 May 2007

Experiments were set up to test how the crystallization drop size affects the crystallization process; in the test cases studied, increasing the drop size led to increasing numbers of crystals. Other data produced from a high-throughput automation-system run were analyzed in order to gauge the effect of replication on the success of crystallization screening. With over 40-fold multiplicity, lysozyme was found to crystallize in over half of the conditions in a standard 96-condition screen. However, despite the fact that industry-standard lysozyme was used in our tests, it was rare that we obtained crystals reproducibly; this suggests that replication whilst screening might improve the success rate of macromolecular crystallization.

1. Introduction

Crystallization of macromolecules is generally a two-step process consisting of a screening step, in which the molecule of interest is tested against a bank of crystallization conditions (or cocktails), followed by cycles of optimization, in which any promising result from the screening is refined in an attempt to produce a well diffracting crystal. The chemical space used in protein crystallization – that is, the possible combinations of different chemicals and concentrations of those chemicals which may be used to crystallize a macromolecule – is large: generally, a crystallization condition contains three to six chemicals at various concentrations chosen from a list of about 350 chemical species which have been used with some success in the past (Peat *et al.*, 2005). Furthermore, a number of experimental choices have to be made for each crystallization trial: what temperature to try, what drop size to use, what technique to apply *etc.* Crystallization space can be considered to be the convolution of chemical space with this experimental space. Protein supply is almost always limiting and ensures that it is impossible to perform extensive screening of crystallization space. Various strategies have been developed to try to best search crystallization space, including grid searches, random searches and sparse-matrix searches (Kundrot, 2004; Carter & Carter, 1979; Rupp & Wang, 2004). All of these approaches are used to obtain initial hit(s) that can be carried forward into optimization. However, despite the body of work suggesting rational strategies for crystallization screening, most experimentalists simply try whatever commercial screens they happen to have in the laboratory already or that have been recommended to them by a persuasive colleague.

Furthermore, it is known that many protein crystals (even those that are optically perfect) do not have sufficient order to produce the high-resolution diffraction pattern required to determine the structure of the constituent macromolecule, so

Table 1
Crystallization cocktails used in the 96-fold redundancy experiment.

The block position number from the Hampton HT screen is given in parentheses after the condition number. The crystal count on the last inspection is shown for both the small (200 nl total initial volume; red in Figs. 1 and 2) and the large (800 nl total initial volume; blue in Figs. 1 and 2) drops.

			Crystal count (day 13)	
Cocktail	Chemicals	Hampton HT condition	Small drops (200 nl)	Large drops (800 nl)
1	0.2 *M* sodium citrate, 0.1 *M* sodium HEPES pH 7.5, 30%(*v*/*v*) 2-methyl-2,4-pentanediol	5 (A5)	60	93
2	0.1 *M* sodium HEPES pH 7.5, 0.8 *M* NaH_2PO_4, 0.8 *M* KH_2PO_4	35 (C11)	41	81
3	0.1 *M* HEPES pH 7.5, 2 *M* ammonium formate	81 (G9)	0	13
4	0.1 *M* Tris pH 8.5, 25%(*v*/*v*) *tert*-butanol	88 (H4)	96	96

it is useful to have more than one hit from the screen to potentially move forward into cycles of optimization.

Most screening strategies put forward discuss how to scan crystallization space efficiently (Segelke, 2001) or how to glean information about the macromolecule that will limit screening space in some way (Hennessy *et al.*, 2000; Charles *et al.*, 2006; Meining, 2006). These discussions appear to assume that if a screening condition exists under which a crystal could form then a crystal will form or, to put it in a different way, that protein crystallization is a reliable process. A recent paper (Newman, 2005) in which crystallization droplets were equilibrated against different reservoirs suggested that this is not true at all and showed that a solution of 'dirty lysozyme' [a mixture of 25 mg ml^{-1} hen egg-white lysozyme (HEWL) and 5 mg ml^{-1} bovine serum albumen (BSA)] crystallized on average in only two of four identical 200 nl total initial volume droplets.

The advent of automation in the process of protein crystallization has freed the process from the limitations of hand pipetting: liquid-handling robots are now available that allow the deposition of sub-microlitre crystallization droplets. Droplets of size 100–200 nl are routinely used in many laboratories, as these small droplets allow many more crystallization conditions to be tested than would be feasible with the larger droplets obtained using hand-pipetting techniques. Another advantage of nanolitre crystallization is that a smaller droplet used in a vapor-diffusion experiment will equilibrate more rapidly with the reservoir than will a larger drop, allowing the crystallizer to read out the results of a crystallization experiment more rapidly. Further benefits of nanolitre crystallization experiments have been reported in the literature (Carter *et al.*, 2005), including better diffraction from crystals grown in smaller drops.

The aim of this paper is to consider factors that might influence protein crystallization screening strategies, other than the often-discussed number of initial screening experiments (Segelke, 2001) and which conditions should be tested in the initial screening (Newman *et al.*, 2005). Crystallization requires two processes to occur: the formation of an initial ordered cluster of molecules (nucleation) and the subsequent addition of molecules to that nucleus (crystal growth). As nucleation is known to be a stochastic event that is dependent on both time and volume (Bodenstaff *et al.*, 2002), we wondered whether the use of sub-microlitre droplets for screening crystallization conditions would have any impact (besides those mentioned above) on the results of that screening. We were also interested in seeing the effect that high levels of replication would have on a crystallization experiment.

2. Methods

2.1. 96-fold redundancy experiment

Four crystallization cocktails that were known to produce crystals of dirty lysozyme when equilibrated against 1.5 *M* sodium chloride were selected from previous work on alternative reservoirs (Newman, 2005). The conditions selected are shown in Table 1. These conditions were selected to encompass a variety of crystallization chemicals from *tert*-butanol to phosphate salts.

For each of the four conditions 192 droplets were set up, 96 with a total initial volume of 200 nl and 96 with a total initial volume of 800 nl. All 192 droplets were set up on a single crystallization plate (Corning CrystalEx 384 Flat Bottom Plate, catalog No. 3775) and consisted of the same pre-mixed protein cocktail. The drop solution was made by mixing a dirty lysozyme solution and the crystallization condition in an Eppendorf tube in equal volumes and then using a 10 µl syringe fitted with a ratchet device (1701N, PB-600 from the Hamilton Company) to dispense the droplets. This was performed to ensure that the droplets were as similar as possible: they should only vary by the error in the dispensing device. The reservoir wells had been pre-filled with 50 µl 1.5 *M* sodium chloride. The drops were arrayed in alternating columns on the crystallization plate in order to even out any edge effects that might occur.

The initial drop size of 200 nl was chosen as this is the smallest dispense volume from the 10 µl syringe with the ratchet attachment; 800 nl was chosen as the large droplet size as droplets larger than this have a distinct tendency to wick to the edges of the subwell, changing the shape of the droplet and thus changing the kinetics of equilibration of the system. The 96-fold redundancy in the experiment allows us to assume with a 95% confidence level a confidence interval of ± 0.1 ($Z_{a/2} = 1.96$, $p = q = 0.5$). Thus, if the count of crystals in two separate experiments differs by more than 10%, the results are statistically significant at the 95% level.

The plates were all equilibrated at room temperature and were imaged on setup and on days 1, 2, 4, 5, 6, 8, 12 and 13 after setup with a Minstrel II imaging system from Robo-Design International (now Rigaku Automation). Each image was inspected manually (by one person) and a positive result

was recorded if one or more crystals were subjectively observed in the drop image.

2.2. Standard screen experiment

An opportunity for further investigation into the stochastic nature of crystallization was afforded by the commissioning runs of a suite of crystallization automation robots (CrystalMation) built by Rigaku Automation for The Scripps Research Institute (TSRI)/The Joint Center for Structural Genomics (JCSG). In this process, 75 µl of the 96-condition Crystal Screen HT crystallization screen (Hampton Research catalog No. HR2-130) was dispensed into the reservoir wells of low-profile crystallization plates (Greiner catalog No. 609 711). Droplets consisting of 50% of the reservoir solution and 50% 25 mg ml^{-1} lysozyme (Sigma L7651) in water were dispensed with a Phoenix RE (a modified Art Robbins Phoenix liquid-dispensing robot) into the sample subwells of the plates. A total of 56 plates were set up: 28 where the total initial drop volume was 200 nl and a further 28 where the total

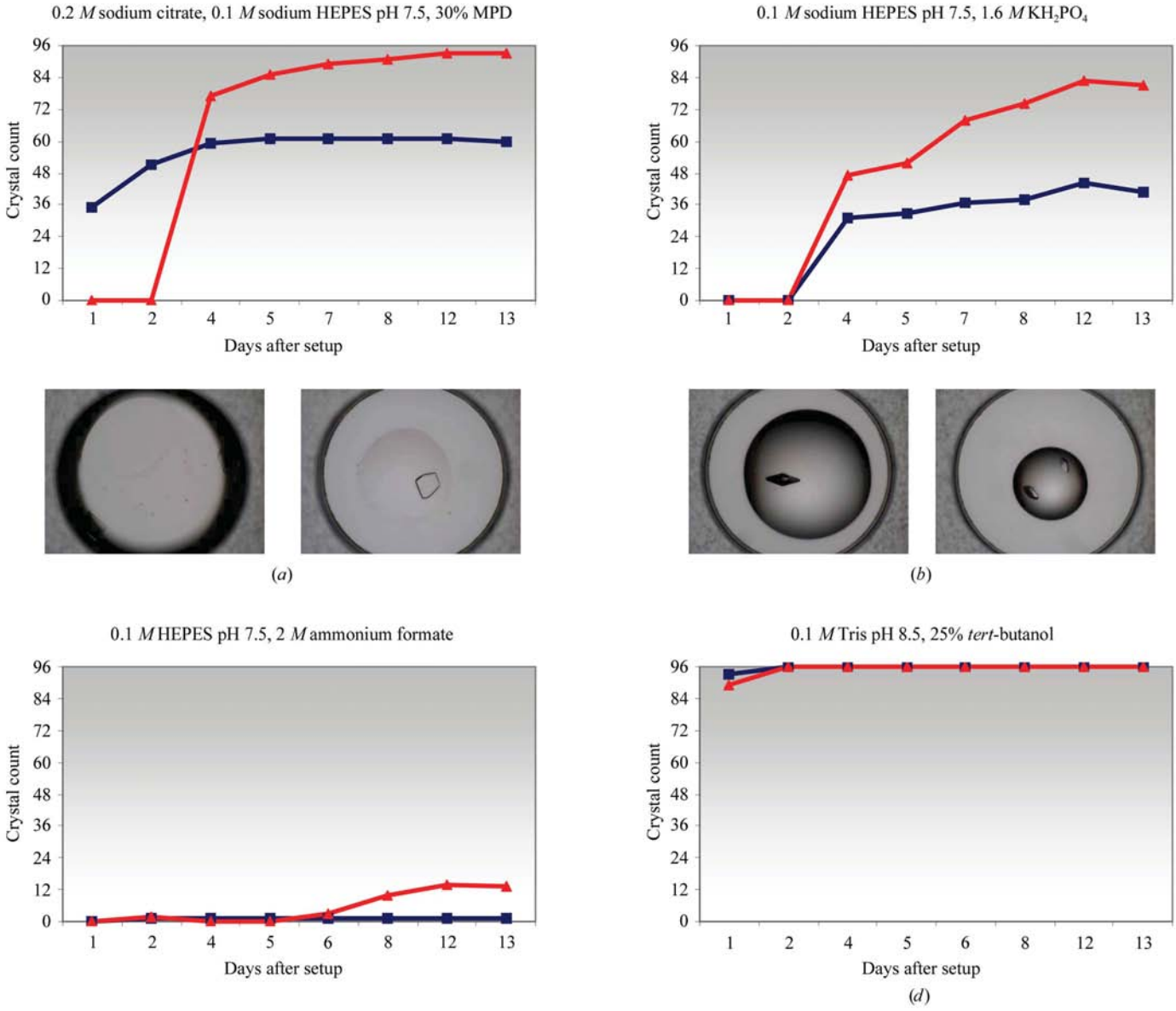

Figure 1
For each of the four conditions selected for use in the 96-fold redundancy experiment, the 'crystal count' (that is, the number of drops containing visible crystals) was plotted for each inspection of the experimental plates. The crystal count for the large (800 nl total initial volume) drops is plotted in red and the crystal count for the small (200 nl total initial volume) drops is plotted in blue. A representative drop image showing the crystals obtained for each condition is shown: the large drop image is shown on the left and the small drop image on the right. (*a*) Cocktail 1, (*b*) cocktail 2, (*c*) cocktail 3, (*d*) cocktail 4.

initial drop volume was 100 nl. These plates were sealed automatically with an AbGene 300 plate sealer; half were stored and imaged with a Rigaku Minstrel HT imager at 293 K and the other half were stored and imaged with another Rigaku Minstrel HT imager at 277 K. Each drop size/temperature/condition combination (henceforth called a trial) was thus set up 14 times. The reservoir solutions were obtained from Hampton Research and were all from a single production batch. The same lysozyme solution was used for all the drops. No machine settings were altered during this experiment. The 56 plates were imaged on either day 3 or day 4 after being set up. As we wished to compare droplets at roughly the same point of equilibration, only those plates that were imaged on day 3 (that is, on the third day after being set up, within the window of 48–72 h after setup) were selected for comparison: this gave 23 plates at 277 K (11 plates set up with 50 + 50 nl drops and 12 plates set up with 100 + 100 nl drops) and 21 plates at 293 K (seven plates set up with 50 + 50 nl drops and 14 plates set up with 100 + 100 nl drops). Each image was examined manually and a positive result was recorded if one or more crystals were observed in the drop image.

3. Results and discussion

3.1. 96-fold redundancy experiment

For each 192-well plate set up with dirty lysozyme, the number of drops with crystals was counted for each inspection. The results were plotted and are shown in Fig. 1.

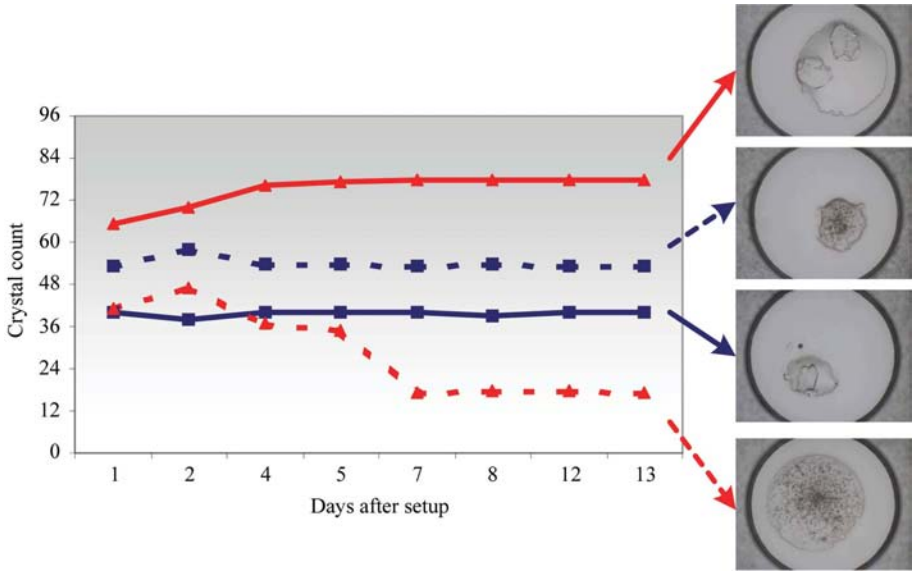

Figure 2
For the 96-fold redundancy experiment using the crystallization cocktail 25% *tert*-butanol, 0.1 *M* Tris pH 8.5 all drops showed crystals after 2 d of equilibration. However, two different crystal forms, a microcrystal form and a macrocrystal form, were seen for this condition. The small (200 nl total initial volume) drops are plotted in blue and the large (800 nl total initial volume) drops are plotted in red. The dashed lines map the growth of the microcrystal form and the solid lines the growth of the macrocrystal form. The drop images on the right-hand side show (from the top moving down) macrocrystals in a large drop, microcrystals in a small drop, macrocrystals in a small drop and microcrystals in a large drop.

For cocktail 1, the count of crystals shows that the smaller drops produced crystals even on the first day after setup. At the end of the 13 d experiment, 60 of the small drops and 93 of the large drops contained crystals. None of the large drops stayed centered in the subwell; all wicked to the side of the subwell. This is shown in Fig. 1(*a*).

With cocktail 2, both the small and the large drops required 2 d equilibration before crystals were found. After 13 d, 41 of the small drops and 81 of the large drops contained crystals (Fig. 1*b*).

Cocktail 3 proved to be a difficult crystallization condition: after 13 d, none of the 96 small drops showed any crystals and only 13 of the larger drops contained crystals (Fig. 1*c*). Thus, this condition gave crystals in less than 7% of the droplets after almost two weeks incubation. The condition seemed to be a robust crystallization condition for dirty lysozyme from the 2005 paper on alternative reservoirs, producing crystals in six of the 16 droplets set up in that experiment or closer to 38% of the drops (Newman, 2005). In the discussion below of the standard screen experiment, this condition (81) proves to be one of the most robust conditions in the Hampton HT screen for producing lysozyme crystals. The standard screen experiment used 25 mg ml^{-1} HEWL dissolved in water as the protein sample, whereas this duplication study used dirty lysozyme as the protein sample. The dirty lysozyme was made by deliberately contaminating a 25 mg ml^{-1} HEWL solution with BSA. The robustness of this crystallization condition could be a function of the purity of the HEWL solution. It may be that the dirty lysozyme made up for the alternative reservoir study had less contaminating BSA than the dirty lysozyme made up for the present study and that this resulted in the very different reliability of this crystallization condition in these two experiments.

On initial inspection, cocktail 4 was a highly successful condition irrespective of the initial drop size, as from the second day after setup all 192 drops contained crystals. Further examination of this plate revealed that this cocktail produced two distinct crystal forms: a microcrystalline form and large single crystals. The distribution of microcrystals and macrocrystals showed a statistically significant variation dependent on the initial drop size, with the larger drops growing more of the larger crystals (Fig. 2). Early inspections of the large droplets showed that micro and large crystals could occur simultaneously in the same droplet; over the time course of the experiment, the number of large drops with microcrystals in them decreased. The microcrystals were thus either redissolving or being incorporated by some type of Ostwald ripening into the large crystal

form. This suggests that the two crystal forms were both lysozyme: however, it could be that one of the forms was HEWL and the other was BSA. There are literature precedents for obtaining different crystal forms from the same crystallization cocktail (Pohl *et al.*, 1998).

For each of these four very different crystallization conditions with dirty lysozyme as the protein, the drop size made a statistically significant difference in the number of drops found that contained crystals. In all cases, after allowing time for initial equilibration, the larger drop size was more likely to produce a protein crystal than the smaller drop size.

3.2. Standard screen experiment

For each of the 96 cocktails used in this experiment, a count of the number of drops containing crystals in the large (200 nl initial volume) and small (100 nl initial volume) drops was made for both temperatures. A trial was considered to be successful if any of the replicate drops for that trial contained crystals. The numbers of successful conditions for each trial were normalized and graphed (Fig. 3). Of the 96 conditions in Crystal Screen HT, 52 produced crystals in one or more of the duplicate trials. The 277 K trials were noticeably better (more conditions gave crystals) than the 293 K trials; this agrees with previously published work that shows that lysozyme nucleates more reliably at lower temperatures (Galkin & Vekilov, 2000). Tables 2 and 3 give more details of these 384 trials.

The standard screen trials show that crystallization of a 25 mg ml^{-1} HEWL solution depends of a number of factors, including the formulation of the crystallization condition, the incubation temperature and the drop volume. 15 of the 96 conditions in Crystal Screen HT crystallized exclusively in the cold and seven crystallized exclusively at 293 K. 18 crystallized only in small drops and 12 only in large drops. 11 of the 96 conditions were successful at both temperatures in both sized drops: these were conditions 25, 27, 29, 35, 41, 48, 49, 57, 77, 81 and 82 (C1, C3, C5, C11, D5, D12, E1, E9, G5, G9 and G10, respectively).

Only four of the successful trials showed complete reproducibility after 3 d; that is, crystals in every drop set up for that trial. These were conditions 29, 41 and 79 (block positions C5, D5 and G7) at 277 K and condition 57 (block position E9) at 293 K. Interestingly, all of the trials showing complete reproducibility involved large drops, even though the smaller drops should equilibrate more rapidly than the larger drops. Furthermore, there were three of these conditions in the cold compared with one at room temperature, even though equilibration would be more rapid at 293 K. It could be that all of these conditions produce supersaturated HEWL solutions immediately and that further equilibration is unnecessary.

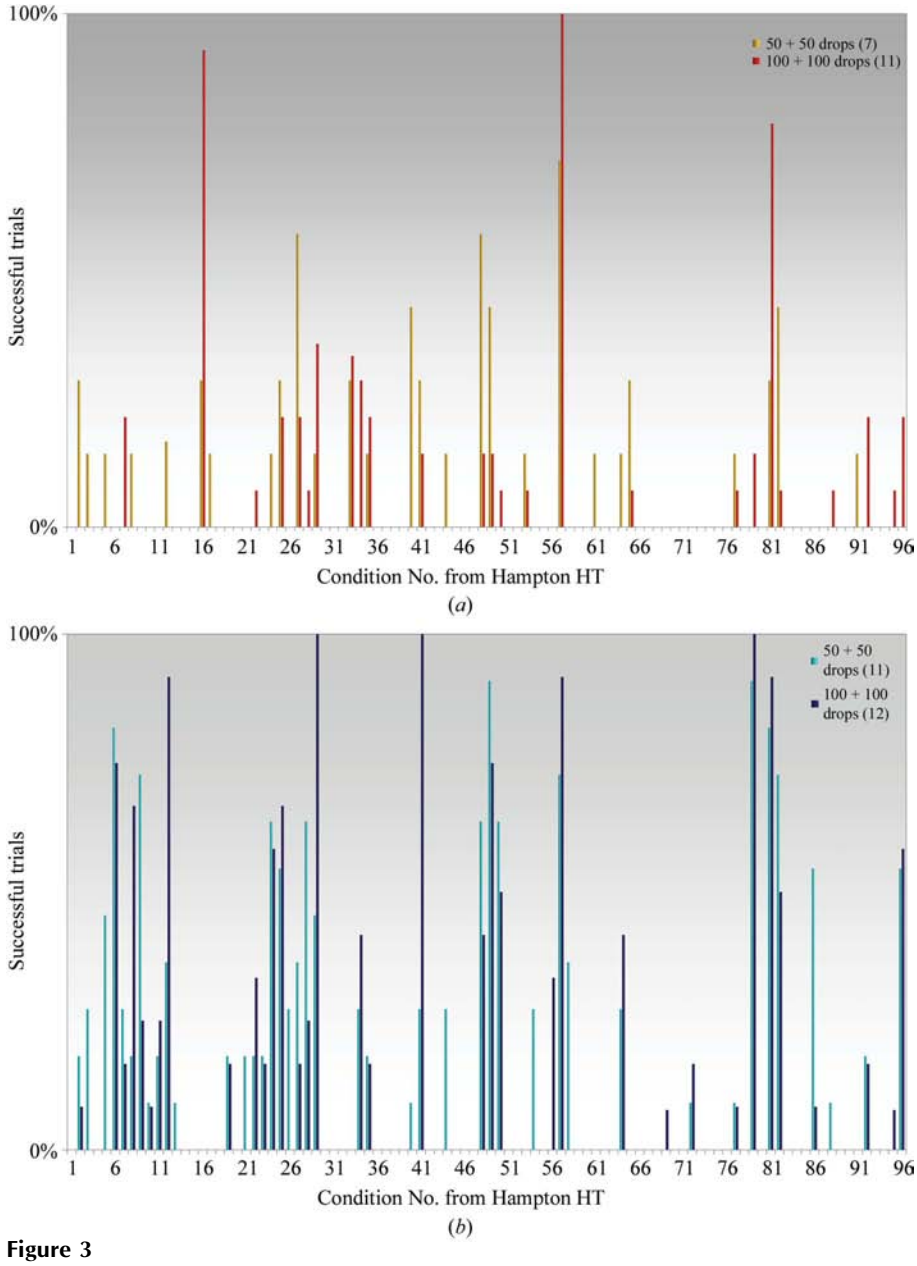

Figure 3
The crystallization success rate of lysozyme using Crystal Screen HT. The success rate is the number of drops within a trial containing one or more crystals divided by the total number of drops for that trial, shown as a percentage. (*a*) plots the results of crystallization of lysozyme using the Crystal Screen HT at 293 K and (*b*) plots the results of crystallization of lysozyme using the same screen at 277 K. Each condition (along the *x* axis) may show two bars: the left-hand bar (yellow in *a*, cyan in *b*) shows the results for small (50 + 50 nl) drops and the right-hand bar (red in *a*, blue in *b*) shows the results for large (100 + 100 nl) drops.

Table 2
Results of the standard screen trials.

The count of the number of conditions from the Hampton HT screen that produced crystals in one or more of the replicate drops for each trial is given.

Temperature (K)	Small drops only	Large drops only	Both drop sizes	Either drop size	This temperature only
277	10	3	32	45	15
293	12	10	15	37	7
277 + 293	18	12	11	52	

The most reproducible conditions in the warm were 16, 57 and 79 (B4, E9 and G9). In the cold, the reproducible conditions were 12, 29, 41, 57, 79 and 81 (A12, C5, D5, E9, G7 and G9). Only conditions 57 and 81 (E9, G9) seem to be temperature-independent; condition 57 is 2 *M* NaCl, 0.1 *M* sodium acetate pH 4.6 or about double the 'standard' lysozyme crystallization condition (Hampton Research product insert guide for lysozyme HR7-110). This condition should be perfect for crystallizing HEWL in the initial setup and equilibration should just push the concentration of salt too high. We see this in the images (data not shown): most of the crystals produced by this condition are the 'sea-urchin' masses of fine crystals indicative of too much sodium chloride in the crystallization solution. The other 'universal' lysozyme crystallization condition, G9, consists of 2 *M* ammonium formate, 0.1 *M* HEPES pH 7.5. As discussed above, the reliability of this condition seems to be highly sensitive to the purity of the lysozyme solution (or at least to the amount of contaminating BSA, as suggested by the previous experiment).

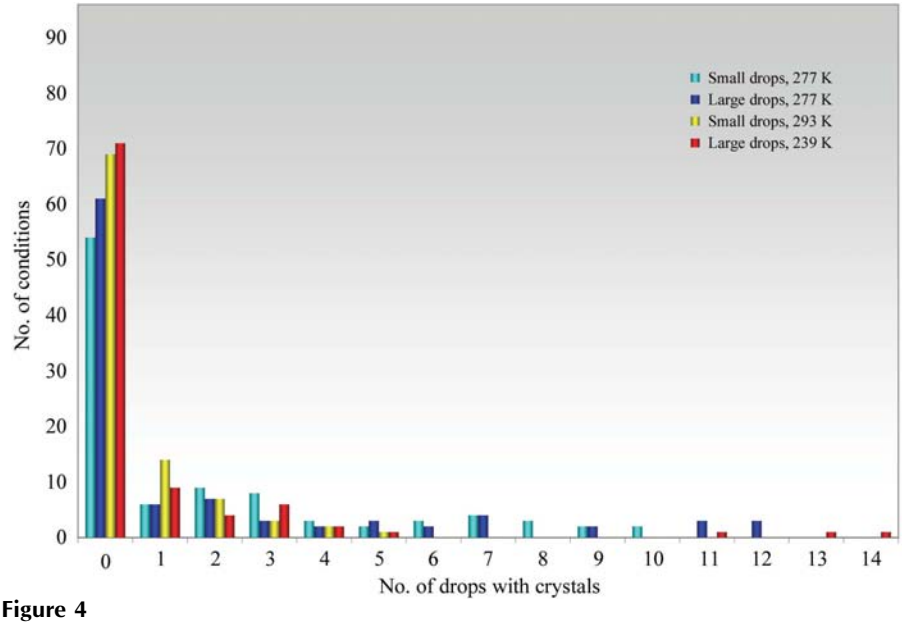

Figure 4
A histogram where for each drop-size/temperature combination (colored according the scheme in Fig. 3) the count of the number of conditions giving hits is plotted. Most of the conditions give no hits and are found at '0' on the *x* axis. Each of the four temperature/drop-size combinations shows a preponderance of conditions that produce crystals in only one of the duplicate drops. Notice that the large drops at 277 K show higher reproducibility than other temperature/drop-size combinations. As there is only 12-fold redundancy in the large/cold drops, there are no bars for these trials at positions '13' and '14' on the *x* axis of this graph.

Table 3
Summary of results for the standard screen trials.

The number in the 'Redundancy' column shows the number of trials that were imaged 3 d after setup and thus used in this analysis. The number in the 'Successful' column counts how many of the 96 conditions in the Hampton HT screen produced visible crystals after 3 d. The 'Average reproducibility' column shows the average success rate for the conditions at that temperature and drop size that were successful. The number in parentheses is the standard deviation for that trial. For the small drops at 277 K, if a condition was successful, it was successful in about 40% of the duplicate drops that were set up.

Temperature/drop size	Redundancy	Successful	Average reproducibility (%)
277 K, 50 + 50 nl drops	11	42	39 (25)
277 K, 100 + 100 nl drops	12	35	43 (32)
277 K, both drop sizes	23	45	
293 K, 50 + 50 nl drops	7	27	27 (16)
293 K, 100 + 100 nl drops	14	25	25 (26)
293 K, both drop sizes	21	37	

It is likely that some of the successful trials would show complete reproducibility given a longer equilibration time. As these data are parasitic data from a commissioning experiment, we have only the single time point (3 d after setup) for these trials and so questions about the time course of the crystallization trials go unanswered in this study. That being said, two possible causes of the imperfect reproducibility of our trials could be incomplete equilibration of the drops and/or insufficient nucleation in the drops. As Table 2 shows, there are 18 conditions that produce crystals only in small drops. It seems reasonable to suspect that the larger drops containing these conditions would also have produced crystals, but had not equilibrated sufficiently in 3 d. There were 12 conditions that produced crystals only in large drops: this suggests a failure of nucleation rather than a problem with incomplete equilibration. Of course, it could be that the solubility of lysozyme is such that these 12 conditions could have produced crystals earlier in the experiment and the equilibration of the small drops was so rapid that crystals did not have time to form before the drops became too concentrated to support crystallogenesis.

Table 3 shows the average success rate of each of the four temperature/drop-size trials. This number is the average percentage of drops containing crystals, given that the condition is a successful one. All of these averages are less that 50%, suggesting that for crystallization of our HEWL solution we really did need to set up a large (greater than five) number of replications to find many of our successful conditions. Fig. 4 shows these data graphically: most of the successful conditions showed crystals in fewer than four drops. We can

also see from this graph that there is less difference between the number of unsuccessful conditions in large and small drops at room temperature than there is in the cold: this suggests that the equilibration in the cold lags that at room temperature, consistent with general belief.

Many crystallographers can recall cases where extensive screening was required to find the crystal that enabled the project to move forward. In these cases the extensive screening may be either (or both) of two things: it could be widening the range of chemical or experimental space being tested or it could be expanding the degree of redundancy in the screen. Many commercial screening conditions are similar: for example, the Personal Structure Screen (Molecular Dimensions), The Classics Suite (Qiagen) and Crystal Screen HT (Hampton Research) all encompass the same conditions, based on the original Jancarik and Kim screen (Jancarik & Kim, 1991). This implicit redundancy is rarely mentioned, in part owing to the difficulties in determining when two crystallization conditions are the same.

We did not subject any of the crystals from the successful trials to diffraction analysis, so we cannot comment on the diffraction quality of the crystals produced by the different trials.

4. Conclusions

The present work shows that although the size of the drops set up makes a statistically significantly difference to the outcome of a lysozyme crystallization experiment, it does not answer the question of the optimal drop size for protein crystallization in general. The optimal drop size is certainly dependent on the protein sample, as proteins that have a higher barrier to nucleation are more likely to produce crystals in a larger volume, whereas smaller drops might be appropriate for proteins that tend to over-nucleate. More importantly, these experiments suggest that the kinetic barrier to crystallization is much more influential than suspected. Even for the easily crystallized protein lysozyme, which this work shows crystallizes in a vast range of conditions (over half of the Crystal Screen HT), a large amount of redundancy in the crystallization trials is required to determine this. This study also shows that the reproducibility of a crystallization experiment is dependent on the crystallization condition as well as the protein sample. Further work is clearly needed to develop an optimal screening strategy for a novel protein, knowing that we have to balance redundancy against diversity against temperature, given a limited protein supply.

Thanks to Marc Pusey for suggesting this line of investigation and the team at JCSG for their input into the automation suite that produced the duplication data. We are grateful to Tom Peat and Mike Newman for helpful and enthusiastic discussions, and to Mike Lawrence and Rob Townley for critical reading of the manuscript.

References

Bodenstaff, E. R., Hoedemaeker, F. J., Kuil, M. E., de Vrind, H. P. & Abrahams, J. P. (2002). *Acta Cryst.* D**58**, 1901–1906.
Carter, C. W. Jr & Carter, C. W. (1979). *J. Biol. Chem.* **254**, 12219–12223.
Carter, D. C., Rhodes, P., McRee, D. E., Tari, L. W., Dougan, D. R., Snell, G., Abola, E. & Stevens, R. C. (2005). *J. Appl. Cryst.* **38**, 87–90.
Charles, M., Veesler, S. & Bonneté, F. (2006). *Acta Cryst.* D**62**, 1311–1318.
Galkin, O. & Vekilov, P. G. (2000). *J. Am. Chem. Soc.* **122**, 156–163.
Hennessy, D., Buchanan, B., Subramanian, D., Wilkosz, P. A. & Rosenberg, J. M. (2000). *Acta Cryst.* D**56**, 817–827.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kundrot, C. E. (2004). *Cell. Mol. Life Sci.* **61**, 525–536.
Meining, W. (2006). *J. Appl. Cryst.* **39**, 759–766.
Newman, J. (2005). *Acta Cryst.* D**61**, 490–493.
Newman, J., Egan, D., Walter, T. S., Meged, R., Berry, I., Ben Jelloul, M., Sussman, J. L., Stuart, D. I. & Perrakis, A. (2005). *Acta Cryst.* D**61**, 1426–1431.
Peat, T. S., Christopher, J. A. & Newman, J. (2005). *Acta Cryst.* D**61**, 1662–1669.
Pohl, E., Holmes, R. K. & Hol, W. G. J. (1998). *J. Biol. Chem.* **273**, 22420–22427.
Rupp, B. & Wang, J. (2004). *Methods*, **34**, 390–407.
Segelke, B. W. (2001). *J. Cryst. Growth*, **232**, 553–562.