Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The role of bias in crystallization conditions in automated microseeding

Franz J. St John, Bo Feng and Edwin Pozharski*

Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, Baltimore 21201, USA

Correspondence e-mail: epozhars@rx.umaryland.edu

Received 7 May 2008
Accepted 29 September 2008

The automated microseeding protocol previously proposed by D'Arcy *et al.* [D'Arcy *et al.* (2007), *Acta Cryst*. D**63**, 550–554] includes an inherent chemical shift in all conditions in a sparse-matrix screen. This results from the uniform addition of mother liquor used to stabilize the microcrystalline seed stock, which is usually the mother liquor that resulted in successful crystal growth. It was found that there is an overlap in identified crystallization leads between addition of seed and addition of the mother liquor used to stabilize the seed. This indicates that additional crystallization 'hits' primarily (but not exclusively) arise from changes in the drop composition, not the introduction of seeds. A complementary automated microseeding protocol is proposed as a novel approach for biasing drops in a screen with many of the most popular precipitants.

1. Introduction

Protein crystallization remains a bottleneck in protein crystallography. Numerous improvements have been made to the most popular technique of vapor-diffusion protein crystallization, beginning with the introduction of the concept of sparse-matrix screening (Jancarik & Kim, 1991). Novel screens have been introduced over the years based on the success rates of low- and high-throughput protein-crystallization efforts (Cudney *et al.*, 1994; Scott *et al.*, 1995; Berger *et al.*, 1996; Garman & Mitchell, 1996; Valjakka *et al.*, 2000; Page *et al.*, 2003; McPherson & Cudney, 2006). Introduction of automated crystallization screening greatly improves the chance of successful crystallization of a protein by significantly reducing the amount of material (protein) required for screening and by simplifying the crystallization setup, which allows a scientist in pursuit of a crystal structure to try a great variety of conditions.

One such novel approach has recently been proposed by D'Arcy and coworkers (D'Arcy *et al.*, 2007; Walter *et al.*, 2008) and is based on the classical idea of seeding (Bergfors, 2003) taken to a new level in the era of high-throughput robotic systems in crystallography. In traditional seeding, crystalline material in mother liquor is sheared to form tiny crystalline seeds. Some dilution of this seed stock is then added to a pre-equilibrated drop of protein and mother liquor that was prepared with somewhat reduced concentrations of precipitant. The idea is that although no nucleation occurs owing to the reduced precipitant concentration, crystal growth is still possible and thus the introduced seeds will grow into large single crystals. In the automated microseeding method proposed by D'Arcy and coworkers, the traditional sparse-matrix

© 2008 International Union of Crystallography
Printed in Singapore – all rights reserved

crystallization screen is set up by mixing the reservoir solution, the protein solution and the seed solution in a 2:3:1 ratio. It was observed that such screens show an increase in crystal-positive conditions.

When using this automated microseeding approach, the conditions in every drop of the screen are not only different from the original screen but are also different when compared with the intent of the traditional microseeding method. One-sixth of the drop's initial content is the mother liquor used to stabilize the seed stock. For instance, if the original conditions from which seeds were obtained contained 2.4 *M* ammonium sulfate, then every condition of the 'seeded screen' will have 0.4 *M* (one-sixth volume) of this salt as an initial concentration (0.8 *M* approximate final concentration). The concentration of the precipitant and other components of the reservoir solution will also be modified. For instance, if the screen condition contains 24% PEG 3350, the final concentration of PEG 3350 when equilibrated may be as low as 16%. This is because the final composition is defined by the equality of the osmotic pressure between the drop and the reservoir solution. Therefore, accurate prediction of the final composition is thus quite complicated and the numbers above should be considered as estimates assuming that the final volume of the drop is the same as in the absence of microseeding. It is conceivable that this modification of crystallization conditions itself is the reason behind the increase in the number of observed crystallization 'hits' since every condition in the screen is thus 'biased' towards the successful crystallization condition used to produce seeds. Here, we report our verification of this explanation of the observed phenomenon. We demonstrate that in the majority of cases using the original mother liquor without seeds gives similar results.

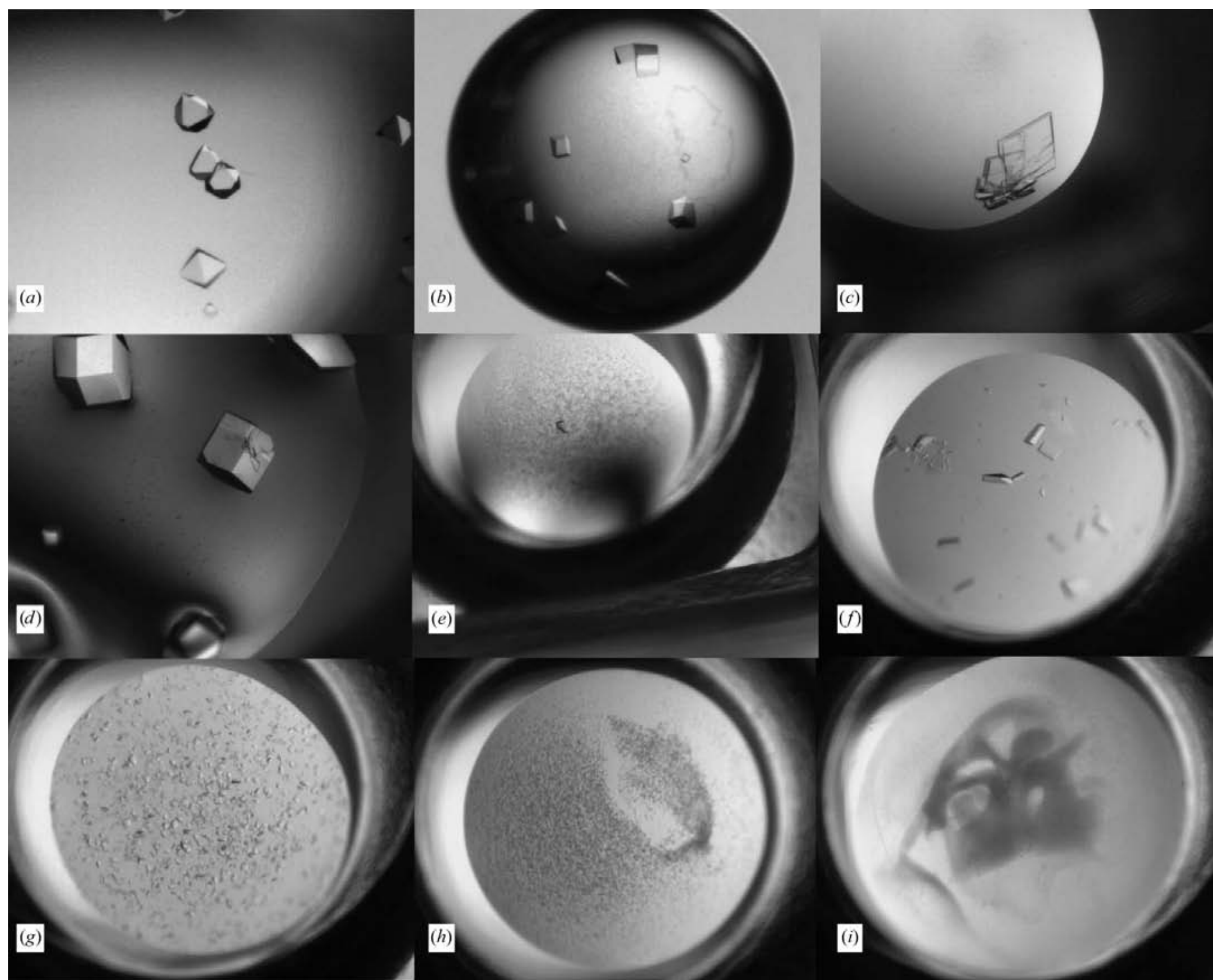

Figure 1
Crystals of (*a*) *T. maritima* glycerol dehydrogenase (*Tm*GldA), (*b*) *E. coli* glycerol dehydrogenase (*Ec*GldA), (*c*) *Paenibacillus* sp. JDR-2 xylanase A1 (XynACD), (*d*) hen egg-white lysozyme (HEWL) and (*e*) *M. musculus* (mouse) histidine triad nucleotide-binding protein (HINT1). Crystals such as these were used to generate crystal seed stocks as described in §2. Representative images of results used to judge the outcome include single crystal (*f*), crystalline shower (*g*), crystalline precipitate (*h*) and precipitate (*i*).

2. Materials and methods

2.1. Proteins used for study

Thermotoga maritima glycerol dehydrogenase (*Tm*GldA; PDB code 1kq3) was expressed in *Escherichia coli* cells provided as a gift by the Joint Center of Structural Genomics (Lesley *et al.*, 2002). The protein was purified by a combination of metal-affinity and size-exclusion chromatography. The final protein preparation used for crystal screens contained 10 mg ml^{-1} *Tm*GldA in 20 m*M* Tris pH 7.5 and 150 m*M* NaCl (TBS). The crystals used to prepare the seed stock were grown by the sitting-drop vapor-diffusion method under previously reported conditions (35% MPD, 0.1 *M* potassium/sodium phosphate pH 6.2; Fig. 1*a*; Lesley *et al.*, 2002).

E. coli glycerol dehydrogenase (*Ec*GldA) cloning, expression and purification procedures will be described in detail elsewhere. Briefly, recombinant protein with a C-terminal octa-His tag was expressed in *E. coli* BL21 (DE3) cells and purified by a combination of metal-affinity and size-exclusion chromatography. The final protein preparation used for crystal screens contained 10 mg ml^{-1} *Ec*GldA in TBS with 20 μ*M* $ZnSO_4$ and either 1 m*M* NAD^+ or 5 m*M* glycerol. The *Ec*GldA preparation with NAD^+ gave crystals in 1.2 *M* sodium formate, 0.1 *M* sodium acetate pH 4.2 by hanging-drop vapor diffusion and these were used to prepare the seed stock (Fig. 1*b*).

The catalytic module of Xyn10A (XynACD) from *Paenibacillus* sp. strain JDR-2 was overexpressed as previously described (St John *et al.*, 2006) and purified by metal-affinity chromatography. The His tag was removed by thrombin protease cleavage and the preparation was then dialyzed and further purified using Mono-Q anion-exchange chromatography. Details will be published elsewhere. Purified XynACD was then dialyzed into 10 m*M* Tris pH 7.5 and concentrated to 10 mg ml^{-1} for crystal screens. Nextal Classic II condition No. 84 [0.2 *M* $MgCl_2$, 0.1 *M* HEPES pH 7.5, 25%(*w*/*v*) PEG 3350] resulted in successful crystal growth and the condition was subsequently refined. The crystals used to prepare the seed stock for this work were grown in 18% PEG 3350 containing 0.1 *M* HEPES pH 7.3 and 0.1 *M* $MgCl_2$ (Fig. 1*c*).

Hen egg-white lysozyme (HEWL) was purchased and crystallized using the industry-standard tetragonal lysozyme recipe (Hampton Research). Briefly, HEWL was dissolved at 50 mg ml^{-1} in 50 m*M* sodium acetate pH 4.8 and mixed with an equal volume of 8% NaCl with 0.1 *M* sodium acetate pH 4.8 as precipitant in sitting-drop vapor-diffusion trays. Crystals grown in this way were used to prepare the seed stock used in this work (Fig. 1*d*).

Mouse histidine-triad nucleotide-binding protein (HINT1) was a gift from the laboratory of Dr Jia Bei Wang. The HINT1 construct contained an N-terminal His tag that was used for metal-affinity column purification. A factor Xa site-specific protease-recognition site designed into the expression construct allowed His-tag cleavage with factor Xa and the affinity tag was removed from the HINT1 preparation by successive dialysis into 50 m*M* Tris pH 7.4 containing 0.2 *M* NaCl. HINT1 was concentrated to 8 mg ml^{-1} for crystal screens and the inhibitor AICAR (5-aminoimidazole-4-carboxamide ribonucleotide) was added to a final concentration of 300 μ*M* with a final HINT1 concentration of 7.2 mg ml^{-1}. Nextal Classic II condition No. 6 (0.1 *M* Tris pH 8.5, 2 *M* ammonium sulfate) resulted in single crystals and these were used to make the seed stock (Fig. 1*e*).

2.2. Crystal screens

Crystallization trays were set up using 96-reservoir three-sample-drop IntelliPlates from Art Robbins (Sunnyvale, California, USA) with an OryxNano crystallization robot from Douglas Instruments. A 70 μl volume of screening solutions from the Nextal Classics Suite (Qiagen, Valencia, California, USA) was placed in each reservoir. In the three crystallization drops, 200 nl reservoir solution was mixed with 300 nl protein solution (as prepared above) and either 100 nl seed stock, 100 nl seed-stock solution without seeds or 100 nl reservoir solution (as a control). To prepare the seed stock, the crystallization drop containing crystals (*e.g.* Figs. 1*a*–1*e*) was combined with 50 μl mother liquor (not containing the cofactor NAD^+, the substrate glycerol or the inhibitor AICAR) and processed with the Seed Bead Kit (Hampton Research, Aliso Viejo, California, USA). This concentrated seed stock was further diluted 100 times with mother liquor (as above) and directly used for seeding. The six protein samples analyzed in this study include (i) *Tm*GldA seeded with *Tm*GldA seed, (ii) *Ec*GldA with $ZnSO_4$ and NAD^+ seeded with *Ec*GldA seed, (iii) *Ec*GldA with $ZnSO_4$ and glycerol seeded with *Ec*GldA seed, (iv) XynACD seeded with XynACD seed, (v) HEWL seeded with HEWL seed and (vi) HINT1 with 300 μ*M* AICAR seeded with HINT1 seed. The two different conditions analyzed using *Ec*GldA (glycerol or NAD^+) were both seeded with *Ec*GldA seed originating from crystal growth containing NAD^+ not glycerol. It must be emphasized that the same protein sample was used in all three drops (unaltered, seeded and supplemented with mother liquor) and therefore any change in crystallization behavior arising from the presence of ligand/substrate had a uniform effect. For all proteins included in this study, similar crystals to those used to make the seeds have also led to structure determination.

2.3. Screen analysis

Analysis of the plates was performed after about one month of incubation at 294 K and was performed subjectively by a single person. A single crystal, a crystalline shower and a crystalline precipitate were all considered to be crystal-positive observations. We present the data as absolute numbers resulting from a simple summation of the various observations (Table 1). It is thought that this analysis would offer an average contribution from six different crystallization screens. Various percentage values are also presented in Table 1 and each is defined in the table legend. As a control to identify possible salt crystals, a tray was set up with buffers in the exact seeding arrangement described for the actual experiment. Only in a few instances (primarily in the MPD-

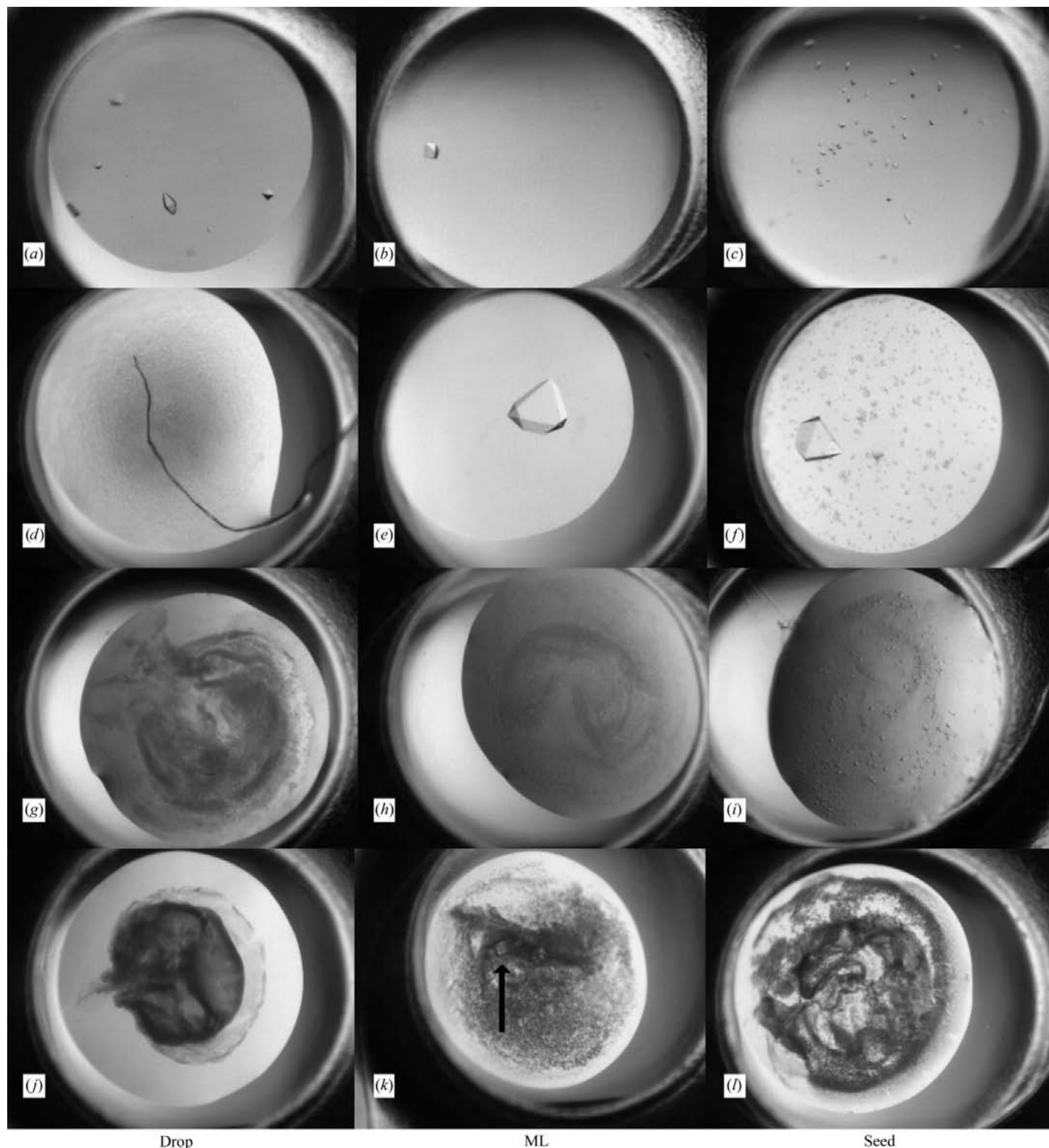

Figure 2
Observed crystal-positive scenario examples. Results are shown for Drop (unaltered conditions), ML (mother liquor with no seed added) and Seed (seed) conditions. See §2 for the description of seed-stock preparation and the content of the mother liquor in each case. (*a*)–(*c*) *T. maritima* glycerol dehydrogenase (*Tm*GldA) in Nextal Classic Suite condition No. 48 (4 *M* sodium formate), with crystals observed in all conditions. (*d*)–(*f*) *E. coli* glycerol dehydrogenase (*Ec*GldA) (with glycerol, not NAD^+) in Nextal Classic Suite condition No. 46 (0.1 *M* Na HEPES pH 7.5, 0.8 *M* sodium phosphate, 0.8 *M* potassium phosphate), showing similar results for the mother liquor and seeded conditions. (*g*)–(*i*) *Ec*GldA in Nextal Classic Suite condition No. 85 [0.2 *M* ammonium sulfate, 0.1 *M* sodium acetate, pH 4.6, 25%(*w*/*v*) PEG 4000], with a crystal shower observed only in the seeded condition. (*j*)–(*l*) *Ec*GldA in Nextal Classic Suite condition No. 72 [0.2 *M* ammonium sulfate, 30%(*w*/*v*) PEG 8000] with a single crystal observed only in the mother-liquor condition (see arrow in *k*).

Table 1
Tabulated results for the seeding-analysis screens.

Positive results include observations of single crystals, crystal showers and crystalline precipitates. ML stands for mother liquor. A dash, a 'Yes' or a 'No' indicate the requirements needed to be included in the described category, with the dash identifying a result that was not applicable to the question asked.

No.	Screen analysis	Crystals were in: Original condition	ML	Seed	Total	Percentage
1	Crystals in the unaltered conditions	Yes	—	—	82	14†
2	New crystal conditions with ML	No	Yes	—	29	6‡
3	Lost crystal conditions with ML	Yes	No	—	30	37§
4	New crystal conditions with seeding	No	—	Yes	24	5‡
5	Lost crystal conditions with seeding	Yes	—	No	20	24§
6	Crystals in ML but not seed	—	Yes	No	16	15¶
7	Crystals in seed but not ML	—	No	Yes	23	22¶
8	Crystals in both ML and seed	—	Yes	Yes	65	63¶
9	Total crystals in ML conditions	—	Yes	—	81	80††
10	Total crystals in seed conditions	—	—	Yes	88	74‡‡

† Taken from the ratio of No. 1 to the total number of conditions for all proteins ($6 \times 96 = 576$). ‡ Taken from the ratio of Nos. 2 and 4 to the total number of conditions for all proteins ($6 \times 96 = 576$) less No. 1. § Taken from the ratio of Nos. 3 and 5 to No. 1. ¶ Taken from the ratio of Nos. 6, 7 or 8 to the sum of Nos. 6, 7 and 8. †† Percentage of mother-liquor crystal conditions that also have crystals from seed (No. 8 to No. 9). ‡‡ Percentage of seed crystal conditions that also have crystals from mother liquor (No. 8 to No. 10).

based *Tm*GldA analysis and the HINT1–AICAR cocrystallization) was it determined that salt-crystal formation led to possibly ambiguous results. The appropriate observations were corrected to exclude the false positives.

3. Results

Crystals similar to those used to prepare the seed stocks of *Tm*GldA, *Ec*GldA, XynACD, HEWL and HINT1 are presented in Figs. 1(*a*)–1(*e*), respectively. Also shown are examples of the conditions as subjectively analyzed, including single crystals (Fig. 1*f*), a crystal shower (Fig. 1*g*), crystalline precipitate (Fig. 1*h*) and an example of a precipitate condition (Fig. 1*i*). A crystalline precipitate is differentiated from a normal precipitate by the color. Precipitates are commonly brown or black and crystalline precipitates are generally clear. The control plates identified three possible salt-crystal conditions for *Tm*GldA and three for HINT1. For *Tm*GldA this was not surprising since the seeding process introduced an initial concentration of 16.6 m*M* phosphate, which favors salt-crystal formation in the presence of divalent cations. The occurrence of salt crystals in three of the HINT1 screen conditions may have resulted from the presence of the AICAR ligand.

From our study we can identify several conditions for each possible scenario (Fig. 2). In Figs. 2(*a*)–2(*c*) crystals were observed in every drop for this *Tm*GldA condition. This was not uncommon for this protein. Figs. 2(*d*)–2(*f*) show an example of *Ec*GldA (with glycerol) resulting in a single perfect crystal in the mother-liquor and seeded drops. In the background of the seeded drop it is evident that the seed did not dissolve but may have grown. This scenario presents a clear example of when mother-liquor addition, not seed, seems to be the key reagent in crystallization. Figs. 2(*g*)–2(*i*) show a crystalline shower in the seeded drop only for *Ec*GldA. The crystals did not grow to any significant size but clearly resulted from seeding. The last scenario (Figs. 2*j*–2*l*) presents a condition for *Ec*GldA in which a crystal is observed only in the mother-liquor addition (an arrow indicates the crystal in Fig. 2*k*).

From the screening of the combined 576 (6×96) unaltered conditions, 82 (14%) resulted in crystal-positive hits (Table 1). Addition of mother liquor increased the new hits by 29 (6%) over the original conditions but also resulted in a loss of 30 (37%) of the original hits. The seeding results were similar, with 24 (5%) new hits and 20 (24%) lost from the original hits. There were a total of 104 conditions that resulted in crystallization with seeding and/or addition of mother liquor. Of these, 65 (63%) produced crystals with both mother-liquor and seed addition. Most importantly, out of the 88 positive crystallization conditions in the presence of seed (this includes both new conditions and successful repeats of the original screen), in 65 instances just the addition of the mother liquor led to similar results. Thus, we have found in our experiments that there is at least a 74% frequency of simultaneous occurrence of crystallization in the drops supplemented with seed stock and mother liquor (Table 1).

Qualitative assessment of the results from the various screens suggests that they could be grouped into three categories. The microseeding procedure with *Tm*GldA and HEWL resulted in a high occurrence of crystal growth with both seed addition and mother-liquor addition. With these two proteins, in 75% of all conditions where crystals were observed with seeding, mother-liquor addition also produced crystals. Only a few new conditions were identified using the microseeding procedure with XynACD and *Ec*GldA with its cofactor NAD^+; for *Ec*GldA with glycerol and HINT1 with AICAR nearly insignificant increases were observed.

4. Discussion

In this work, the role of seed-stock mother liquor in microseeding as described by D'Arcy and coworkers was analyzed (D'Arcy *et al.*, 2007). Five different proteins were used, with one (*Ec*GldA) being applied in two different screening conditions, to give a total of six unique screens. Most importantly, the original crystallization conditions used to produce microseeds were vastly different and covered a broad range of precipitants: salts, polyethylene glycols and organic solvents. While *Ec*GldA and *Tm*GldA share 53% sequence identity, they crystallize under very different conditions and can therefore be considered to be completely different proteins from the crystallization point of view.

The question raised by this work concerns the mechanism by which the automated microseeding methodology (D'Arcy

et al., 2007) increases the number of crystallization conditions obtained from a screen. It was originally proposed that it is the introduction of seeds that is responsible. This essentially means that many more conditions in a standard screen than are usually detected allow protein-crystal growth and it is the nucleation event which is missing. The evidence we present suggests that an alternative reason for the increased likelihood of crystallization is the substantial shift of the conditions in the mother liquor when seeds are introduced together with the stabilizing solution. Since such a stabilizing solution is always very similar to the condition which produced crystals, it is not surprising that a bias introduced into the screen favors crystallization. In theory, the robotic seeding protocol indeed violates one of the requirements of microseeding: seeds should be introduced into a pre-equilibrated mixture of protein and mother liquor because otherwise they may dissolve. While it is quite possible that in some conditions seeds may be preserved because of the high concentration of precipitant in the reservoir solution, a sixfold dilution of the seed-stabilizing solution is inherent to the protocol and it is therefore likely that microcrystalline seeds will dissolve quite often.

In our screens, for the majority of drops where 'seed-induced' crystallization was observed similar results were obtained when the mother liquor of the seed stock (the original crystal-growth mother liquor) was used instead. In these cases we can be quite confident that it was the change in the conditions in the drop and not the introduction of seeds which led to crystallization. Sometimes we observed that the seeded drop did not produce crystals while that supplemented with mother liquor did. We also observed that in some cases seeded drops produced many more and smaller crystals, indicating that seeds may be stable under such conditions. In some cases, however, addition of mother liquor did not yield crystals when seeding did. This emphasizes that it is quite possible that under certain screening conditions seeds may be preserved. It is unknown whether such a procedure will lead to novel crystal forms.

'Crystallizability' varied greatly among the proteins used in this study. From our analysis, *Tm*GldA displayed robust crystallization characteristics. The original crystals were derived from an MPD-based condition (Lesley *et al.*, 2002) and screening of the same protein preparation with the Nextal MPD Suite resulted in a large number of crystal observations (data not shown). From this, we deduce that *Tm*GldA displays robust MPD-dependent crystallization characteristics and thus it made the most significant contribution (together with HEWL) to the tabulated data. The seed and mother liquor (MPD-based) added to each condition of the Nextal Classic Suite provided a key component for successful crystallization of *Tm*GldA. At the other extreme, the original screens for the cocrystallization of the HINT1–AICAR complex and of *Ec*GldA with its substrate glycerol did not seem robust. Owing to this imbalance in crystallizability, it is important to realise that the final numbers are substantially influenced by the more robustly crystallizing proteins *Tm*GldA and HEWL.

Our findings do not reduce the value of the method proposed by D'Arcy and coworkers. Even if on some (and potentially a majority of) occasions the additional crystallization conditions simply lead to the same crystal form, the discovery of alternatives is invaluable in a case when improvement of the quality of diffraction is sought. We propose, as has been pointed out previously (Majeed *et al.*, 2003), that modification of an existing sparse-matrix screen by the addition of a second precipitant may increase the likelihood of crystallization. Since two major precipitants (ammonium sulfate and polyethylene glycols of various lengths) are used in a significant fraction of all successful protein crystallizations, one can definitely consider such modified screens (*e.g.* adding 2 *M* ammonium sulfate to every drop in lieu of the 'seeds') as a way to improve chances of crystallization. Expanded use of simple crystallization robots and multiple-well crystallization plates may greatly expand the range of conditions each sparse-matrix screen can offer.

We thank James Smith of Douglas Instruments for valuable discussions, Dr Jia Bei Wang for contribution of the HINT1 protein and the Joint Center for Structural Genomics for making available the protein-expression vector for *T. maritima* glycerol dehydrogenase.

References

Berger, I., Kang, C., Sinha, N., Wolters, M. & Rich, A. (1996). *Acta Cryst.* D**52**, 465–468.
Bergfors, T. (2003). *J. Struct. Biol.* **142**, 66–76.
Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y. & McPherson, A. (1994). *Acta Cryst.* D**50**, 414–423.
D'Arcy, A., Villard, F. & Marsh, M. (2007). *Acta Cryst.* D**63**, 550–554.
Garman, E. F. & Mitchell, E. P. (1996). *J. Appl. Cryst.* **29**, 584–587.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Lesley, S. A. *et al.* (2002). *Proc. Natl Acad. Sci. USA*, **99**, 11664–11669.
Majeed, S., Ofek, G., Belachew, A., Huang, C.-C., Zhou, T. & Kwong, P. D. (2003). *Structure*, **11**, 1061–1070.
McPherson, A. & Cudney, B. (2006). *J. Struct. Biol.* **156**, 387–406.
Page, R., Grzechnik, S. K., Canaves, J. M., Spraggon, G., Kreusch, A., Kuhn, P., Stevens, R. C. & Lesley, S. A. (2003). *Acta Cryst.* D**59**, 1028–1037.
St John, F. J., Rice, J. D. & Preston, J. F. (2006). *Appl. Environ. Microbiol.* **72**, 1496–1506.
Scott, W. G., Finch, J. T., Grenfell, R., Fogg, J., Smith, T., Gait, M. J. & Klug, A. (1995). *J. Mol. Biol.* **250**, 327–332.
Valjakka, J., Hemminki, A., Teerinen, T., Takkinen, K. & Rouvinen, J. (2000). *Acta Cryst.* D**56**, 218–221.
Walter, T. S., Mancini, E. J., Kadlec, J., Graham, S. C., Assenberg, R., Ren, J., Sainsbury, S., Owens, R. J., Stuart, D. I., Grimes, J. M. & Harlos, K. (2008). *Acta Cryst.* F**64**, 14–18.