Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Comparison of three commercial sparse-matrix crystallization screens

Jong Wei Wooh,[a,b] Richard D. Kidd,[b]† Jennifer L. Martin[a,b,c] and Bostjan Kobe[a,b,c]*

[a]Department of Biochemistry and Molecular Biology, University of Queensland, Brisbane, Queensland 4072, Australia, [b]Institute for Molecular Bioscience and ARC Special Research Centre for Functional and Applied Genomics, University of Queensland, Brisbane, Queensland 4072, Australia, and [c]Cooperative Research Centre for Chronic Inflammatory Diseases, University of Queensland, Brisbane, Queensland 4072, Australia

† Present address: Inpharmatica Ltd, 60 Charlotte Street, London W1T 2NU, England.

Correspondence e-mail: b.kobe@mailbox.uq.edu.au

Received 11 November 2002
Accepted 4 February 2003

Sparse-matrix sampling using commercially available crystallization screen kits has become the most popular way of determining the preliminary crystallization conditions for macromolecules. In this study, the efficiency of three commercial screening kits, Crystal Screen and Crystal Screen 2 (Hampton Research), Wizard Screens I and II (Emerald BioStructures) and Personal Structure Screens 1 and 2 (Molecular Dimensions), has been compared using a set of 19 diverse proteins. 18 proteins yielded crystals using at least one crystallization screen. Surprisingly, Crystal Screens and Personal Structure Screens showed dramatically different results, although most of the crystallization formulations are identical as listed by the manufacturers. Higher molecular weight polyethylene glycols and mixed precipitants were found to be the most effective precipitants in this study.

1. Introduction

Crystallization remains one of the major bottlenecks in macromolecular structure determination by X-ray crystallography. The search for conditions suitable for crystal growth remains largely empirical, particularly screening for initial crystallization conditions. Once initial crystals have been obtained, the conditions can usually be optimized by a more systematic approach.

Because crystallization is affected by many variables and the amount of protein is usually limited, sampling of the multidimensional condition space can be challenging. Various approaches have been proposed for this task. While sampling methods such as the incomplete factorial approach (Carter & Carter, 1979), orthogonal arrays (Kingston *et al.*, 1994) and reverse screening (Stura *et al.*, 1994) offer statistically superior sampling of various parameters or a more methodical route to obtaining preliminary crystals, the method of sparse-matrix screening has arguably become the most popular approach for initial crystallization screening. In an early application of sparse-matrix sampling, a set of 50 crystallization solutions was proposed based on known or published crystallization conditions for various proteins (Jancarik & Kim, 1991). Many variations of sparse-matrix screens have subsequently been developed in various laboratories and the popularity of such screens has increased through the availability of commercial kits (*e.g.* Crystal Screen, Hampton Research). The sparse-matrix approach was originally suggested to be well suited to automation (Jancarik & Kim, 1991) and, accordingly, this approach has been adopted by structural genomics initiatives (Burley, 2000).

Here, we posed the question whether three popular commercial screens (Crystal Screen and Crystal Screen 2, Hampton Research; Wizard Screens I and II, Emerald BioStructures; Personal Structure Screens 1 and 2, Molecular Dimensions) are similarly effective in crystallizing a set of 19 diverse proteins. Surprisingly, we found that the Hampton Research Crystal Screens and Molecular Dimensions Personal Structure Screens yielded quite different results, although most of the formulations are identical as listed by the manufacturers. Our study shows some trends in the efficiency of individual crystallization components in the crystallization of proteins and will help in formulating even more efficient crystallization screens.

2. Experimental methods

The following proteins were used without further purification: equine myoglobin, hen egg-white lysozyme, bovine catalase, rabbit phosphorylase B, porcine pepsin, bovine α-lactalbumin, bovine trypsin, human haemoglobin, *Bacillus licheniformis* subtilisin Carlsberg, bovine ribonuclease A, porcine elastase (all obtained from Sigma–Aldrich), *Trichoderma longibrachiatum* xylanase and *Streptomyces rubiginosus* glucose isomerase (both obtained from Hampton Research). Porcine ribonuclease inhibitor (Kobe & Deisenhofer, 1993), mouse importin-α (Teh *et al.*, 1999), feline immunodeficiency virus gp36 (residues 652–784)–maltose-binding protein (MBP)

Table 1
Proteins used, sources, various properties and the outcomes of crystallization in the three screens (WS, Wizard Screens I and II; CS, Crystal Screen and Crystal Screen 2; PSS, Personal Structure Screens 1 and 2).

Protein	Biological source	Calculated pI†	Theoretical molecular mass (kDa)	WS	CS	PSS
Myoglobin	Horse skeletal muscle	7.4	17.0		Yes	Yes
Lysozyme	Hen egg white	9.4	15.2	Yes	Yes	Yes
Catalase	Bovine liver	6.4	57.6	Yes	Yes	Yes
Xylanase	*T. longibrachiatum*	9.0	20.7	Yes	Yes	Yes
Phosphorylase B	Rabbit muscle	6.8	97.2	Yes	Yes	Yes
Pepsin	Porcine gastric mucosa	4.0	41.4			Yes
Chk2–MBP	Human	5.6	80.1			Yes
Glucose isomerase	*S. rubiginosus*	3.0	43.2	Yes	Yes	Yes
α-Lactalbumin	Bovine liver	4.9	16.2	Yes	Yes	Yes
Trypsin	Bovine pancreas	4.7	26.3	Yes	Yes	Yes
Haemoglobin	Human	7.8	62.0			
Subtilisin Carlsberg	*B. licheniformis*	6.6	27.2	Yes		Yes
Ribonuclease A	Bovine pancreas	8.9	16.4	Yes	Yes	Yes
Ribonuclease inhibitor	Porcine	4.8	49.0	Yes	Yes	−‡
Elastase	Porcine pancreas	8.4	28.8	Yes	Yes	Yes
DsbG	*E. coli*	7.4	27.0	Yes	Yes	Yes
Dun1-FHA	*S. cerevisiae*	8.8	16.9			Yes
Importin-α	Mouse	5.8	63.6		Yes	Yes
gp36–MBP	Feline immunodeficiency virus	6.5	56.4	Yes		Yes

† The pI was calculated using the ExPASy Moelcular Biology Server (http://www.expasy.ch) (Bjellqvist *et al.*, 1993). ‡ Not analysed owing to limited amount of protein.

chimera (Y. M. Qi, P. Poumbourios & B. Kobe, unpublished work), *Saccharomyces cerevisiae* Dun1p (residues 19–159) (Hammet *et al.*, 2000), human Chk2 (residues 214–543)–maltose-binding protein chimera (H. T. Lee & B. Kobe, unpublished work) and *Escherichia coli* DsbG (B. Heras, S. Raina & J. L. Martin, unpublished work) were purified in our laboratories (Table 1). The proteins were used at a concentration of 10 mg ml^{-1} (dissolved or dialysed in 25 m*M* Tris–HCl pH 7.0), except for ribonuclease inhibitor (27 mg ml^{-1} in 20 m*M* HEPES pH 7.0), importin-α [11 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl, 2 m*M* dithiothreitol (DTT)], gp36–MBP (14 mg ml^{-1} in 100 m*M* Tris–HCl pH 6.5, 300 m*M* NaCl), Dun1p (12 mg ml^{-1} in 20 m*M* imidazole pH 7.0 and 15 m*M* β-mercaptoethanol), Chk2–MBP (8.6 mg ml^{-1} in 50 m*M* Tris pH 7.5, 0.5 *M* NaCl, 10% glycerol, 5 m*M* DTT) and DsbG (10 mg ml^{-1} in 25 m*M* HEPES pH 6.7, 150 m*M* NaCl, 5 m*M* DTT). Trypsin and subtilisin were inhibited with 4.2 and 3.7 m*M* Pefabloc SC inhibitor (Roche), respectively.

The hanging-drop vapour-diffusion method was used for crystallization. 1 µl of protein solution was combined with 1 µl of reservoir solution on the sticky side of 3M PCR tape (using an eight-channel pipette), the tape was inverted over a Falcon flat-bottom 96-well plate containing 100 µl of reservoir solution per well and the plate was incubated at 289 K. For each protein, three plates were set up, each plate containing 96 conditions corresponding to Crystal Screen (CS1) and Crystal Screen 2 (CS2) (Hampton Research; both screens are referred to as CS), Wizard Screens I (WS1) and II (WS2) (Emerald BioStructures; both screens are referred to as WS) and Personal Structure Screens 1 (PSS1) and 2 (PSS2) (Molecular Dimensions; both screens are referred to as PSS), respectively (owing to a limited amount of protein, PSS was not completed for ribonuclease inhibitor). Formulation 1 in CS2, WS2 and PSS2 is numbered 49 in this setup. To accommodate all formulations on a 96-well plate, conditions 49 and 50 of CS1 and PSS1 were omitted. Conditions 1 and 2 from PSS1 were substituted by conditions 49 and 50 from PSS2, respectively. To mimic a typical crystallization screening experiment using a novel protein, one drop was set up per condition and protein; the reader should note that this poses limitations on the statistical validity of any conclusion drawn. All the formulations are listed in Supplementary Table S1[1].

The drops were inspected using a stereo dissecting microscope (Leica MZ75) immediately after set-up and after one week; the final results were compiled after 1–2 months. Needles, plates and three-dimensional crystals were considered as successful crystallizations, regardless of the size of the crystals. The diffraction quality of the crystals was not assessed. Because not all crystals were tested for protein content, it is possible (although unlikely) that some crystals considered to be protein crystals did not contain protein. The formulations resulting in successful crystallizations are listed in Supplementary Table S2.

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: gr2328). Details for accessing these data are described at the back of the journal.

The pH of the formulations in CS and PSS was measured using an Ionode pH meter (TPS, Brisbane, Australia) at 289 K (Supplementary Table S3).

3. Results and discussion

The proteins we used in our test set were chosen to represent a diverse set of proteins from a diverse set of organisms (ranging from bacteria and viruses to humans), with diverse molecular weights and pI values and with diverse functions (ranging from enzymes through proteins involved in protein–protein interactions and cofactor-binding proteins to recombinant fusion proteins) (Table 1). Most of these proteins have previously been shown to be crystallizable under some set of conditions. In our analysis, we obtained crystals for 18 out of 19 proteins using at least one of the screen conditions; this observation highlights the high efficiency of the commercial screens in obtaining preliminary crystals (Supplementary Fig. S1). The study also showed that the crystallization technique using vapour diffusion with 96-well plates and the placement of drops directly onto the tape, a technique amenable to high-throughput crystallization, is an effective crystallization method.

The highest success rate in crystallizing our set of proteins was achieved with PSS (17 out of 18 proteins; for CS and WS, 13 out of 19 proteins produced crystals). Ten proteins crystallized in all three screens. Many proteins crystallized using numerous different formulations, with the highest number recorded for glucose isomerase (63 conditions; Supplementary Table S2). The only protein that failed to crystallize was haemoglobin; successful crystallizations have previously been reported for this protein (Perutz, 1968) and we conclude that its crystallization may require higher concentrations of protein, different temperatures or other conditions not sampled in this study [a standard concentration of 10 mg ml^{-1} (Jancarik & Kim, 1991) was used for most proteins in this study].

The most successful formulation overall was No. 14 from CS [28% polyethyleneglycol (PEG) 400, 100 m*M* HEPES pH 7.5 and 200 m*M* calcium chloride], which produced crystals for ten (53%) of the proteins (Fig. 1). Another formulation produced crystals for eight different proteins, two formulations produced crystals for seven proteins and three formulations produced crystals for six proteins (Fig. 1). The most effective formulations in CS are

consistent with a histogram of successful crystallizations presented in the Hampton Research Catalogue. Two proteins were crystallized uniquely with only one condition. No crystals were obtained with 31, 40 and 44% of the formulations in PSS, CS and WS, respectively.

One unexpected result that emerged from this study was the difference in efficiency between CS and PSS. Most crystallization formulations in these two screens are identical as listed by the manufacturers. Figs. 1(*d*) and 1(*e*) show a comparison of the two screens after we matched the conditions (the order of the formulations is different in the two screens; Supplementary Table S3). In 38 cases, the solutions in PSS were more successful in producing crystals than CS, while the opposite was the case for 26 formulations. For example, the PSS equivalent of the most successful formulation in CS (No. 14, ten successful crystallizations) produced only one successful crystallization; conversely, the CS equivalent of the most successful formulation in PSS (No. 46, eight successful crystallizations) produced only one successful crystallization. Crystal nucleation is a chance event and some fluctuation in the results is unavoidable; however, nucleation in virtually all drops producing crystals was frequent enough to prevent large fluctuations. Some of the differences could therefore result from the source of the chemicals and the procedures used by the companies in preparing the formulations. For example, Hampton Research prepare the buffers by adjusting the pH with hydrochloric acid or sodium hydroxide, whereas Molecular Dimensions use glacial acetic acid to adjust the pH (personal communication). 22 matched formulations had pH values which differed by more than half a pH unit between the two screens, with the highest discrepancy reaching almost 5 pH units (Supplementary Table S3). Many of the discrepancies in pH correspond to formulations that use no buffer or specific chemicals such as Jeffamine M-600. The pH differences and similar but non-identical formulations explain some but not all of the cases where only one screen produced crystals.

We also analysed the results to deduce the effects of individual components of the crystallization formulations. Because this is not a systematic analysis of individual components and other components present may modulate any effects of an individual component, inferences must be made cautiously. The most effective precipitant class corresponded to polymers (7.8% success over all formulations), compared with salts (5.0%), non-volatile and volatile organic solvents (5.5 and 3.2%, respectively) and other precipitants (7.6%) (Supplementary Table S3 and Fig. S2). The success of the

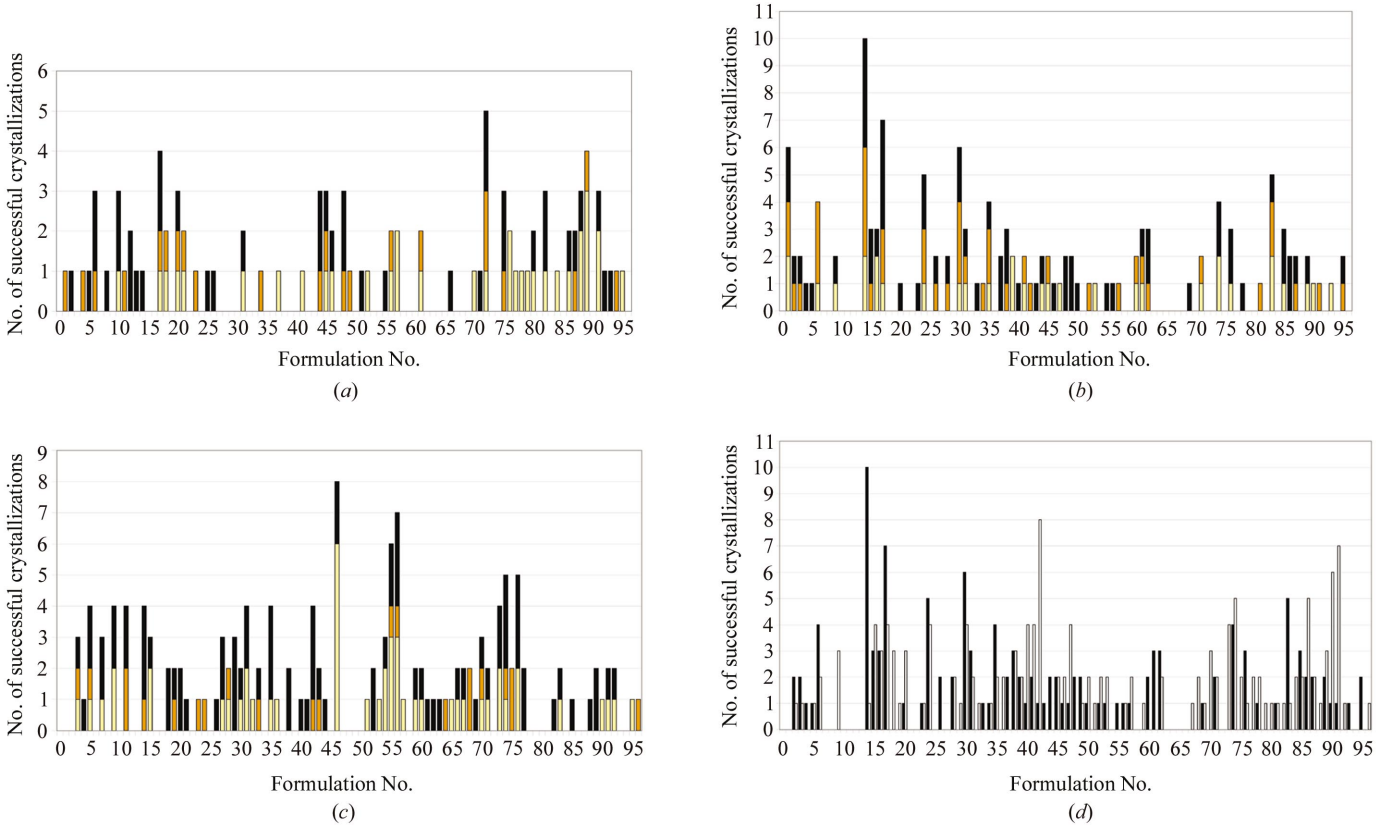

Figure 1
(*a*) Histogram showing the number of successful crystallizations using WS formulations. Needles, yellow; rods, orange; three-dimensional crystals, black. (*b*) The results of CS, shown as in Fig. 1(*a*). (*c*) The results of PSS, shown as in Fig. 1(*a*). (*d*) Comparison of CS and PSS. Equivalent formulations were matched and the number of successful crystallizations is shown for each equivalent condition side by side (CS, black; PSS, white; see Supplementary Table S3 for the matched conditions). The numbers correspond to the order of formulations in CS. CS conditions 50, 64, 67 and 94 are similar but not identical in PSS. CS conditions 1 and 10 do not have a match in PSS in our setup and the number of successful crystallizations is therefore not shown in the histogram for these conditions (*i.e.* the numbers are set to zero). (*e*) Comparison of CS and PSS, shown as a Venn diagram. Only identical conditions were considered in the comparison. CS, number of crystallizations unique to CS; PSS, number of crystallizations unique to PSS; cross-section, number of crystallizations in both CS and PSS using the identical formulations.

'other' precipitant class entirely arises from the 'combined precipitants', most of which contain at least one polymer precipitant. The precipitants imidazole, urea, polyethyleneimine and polyvinylpyrolidone yielded no crystals, while Jeffamine M-600 produced only one successful crystallization (however, these precipitants are poorly represented among the formulations). No significant trends are observed among individual salts (see also a recent study by McPherson, 2001) and organic solvents; on the other hand, there is a clear trend of increasing success rates with increasing PEG molecular weights, showing maxima at PEG 10 000 and PEG MME 5000 (jointly the most effective precipitants, with a success rate of 18.9%; Fig. 2). The results suggest that there may be an optimal molecular weight of PEG that balances the effects of lowering the dielectric constant of water and reducing protein solubility and of inducing crystal growth. Of the salts, the most successful precipitant was lithium sulfate (success rate 10.8%) and of the organic solvents, surprisingly, it was dioxane (success rate 8.1%). Three combined precipitants had a success rate of 13.5%; two of these contained PEG 4000 and 2-propanol, while the other was a combination of ammonium sulfate and PEG 400. These results suggest that PEGs and the less well known PEG monomethylethers (MMEs; Brzozowski & Tolley, 1994) are the most effective precipitants and that it is also very beneficial to explore the combined precipitants, perhaps through the effects of 'hard' and 'soft' precipitants (Huang *et al.*, 1999).

Successful crystallization as a function of pH shows two maxima at pH 5.5 and 8.0 (Supplementary Fig. S3). The success rate decreases between these two points, with a minimum at pH 6.2. Owing to a non-random association of the pH values with particular crystallization components and the uneven sampling distribution of the pH values, the significance of this observation is unclear; however, it is in line with previous observations (McPherson, 1999). There does not appear to be any correlation between crystallization success and the calculated pI values of the proteins (Table 1).

Another way to assess the effects of different components is to analyse the least successful formulations. For example, heavy irregular precipitation could point to conditions where proteins (at the concentrations used here) were either dramatically above the solubility limit or were severely destabilized. The analysis shows that most conditions frequently inducing heavy precipitation contain zinc acetate or another divalent ion-containing salt, have a pH of 4.6 or lower or contain PEG (with molecular weight 4000 or higher) at a concentration of 30%. These observations suggest that low pH values and divalent ions may be detrimental to the stability of many proteins and that concentrations of $\geq$30% of PEGs with molecular weights of $\geq$4000 may cause the proteins to substantially exceed the solubility limit.

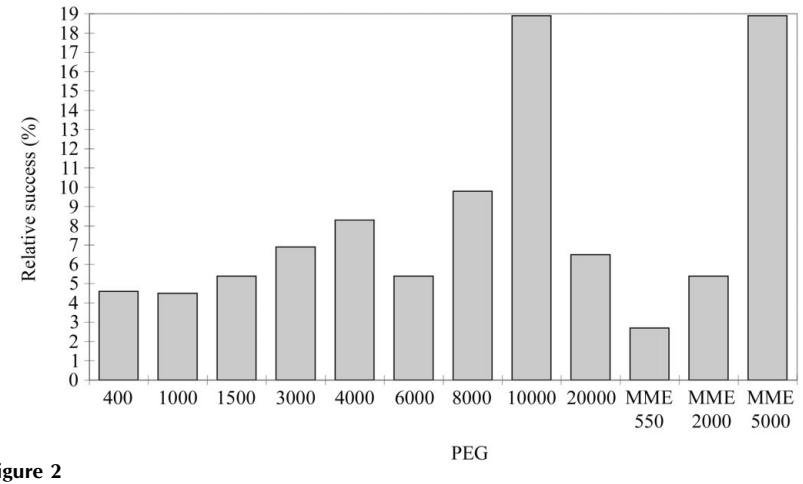

Figure 2
Relative success (needles, plates and three-dimensional crystals were considered to be successful crystallizations) for various PEGs, calculated over all the screens and all 19 proteins (5376 conditions).

We consider the observed trends are likely to be representative of proteins in general, despite the limited number of proteins used in this study. However, more comprehensive information should soon emerge from analyses of the large numbers of crystallization experiments performed by structural genomics initiatives.

4. Conclusions

All three commercially available screens used in our analysis were shown to be highly successful in yielding preliminary crystallization conditions for the diverse proteins in our test set. Certain formulations show very high success rates and our analysis should help to formulate new sparse-matrix screens that are even more efficient than those presently available. The unexpected difference in the efficiency of CS and PSS emphasizes the importance of the preparation, purity and quality of crystallization components in inducing crystal growth and the delicacy of protein crystallization. It also suggests that it is worthwhile setting up both screens, despite the fact that most formulations are identical as listed by the manufacturers. The success rates of individual components are consistent with trends observed in other studies and suggest higher molecular-weight PEGs and PEG MMEs to be particularly useful crystallization agents.

We thank Begona Heras, Huang-Teck Lee, Ying Mei Qi, Douglas Smyth and Trazel Teh for contributing proteins for this study and also Douglas Smyth for help with crystallization, Tom Peat and Janet Newman for advice on the crystallization setup and members of our laboratoriess for valuable discussions. JLM is an ARC Senior Research Fellow and BK was a Wellcome Senior Research Fellow in Medical Science in Australia and is an NHMRC Senior Research Fellow.

References

Bjellqvist, B., Hughes, G. J., Pasquali, C., Paquet, N., Ravier, F., Sanchez, J. C., Frutiger, S. & Hochstrasser, D. (1993). *Electrophoresis*, **14**, 1023–1031.
Brzozowski, A. M. & Tolley, S. P. (1994). *Acta Cryst.* D**50**, 466–468.
Burley, S. K. (2000). *Nature Struct. Biol.* **7**, *Suppl.*, 932–934.
Carter, C. W. Jr & Carter, C. W. (1979). *J. Biol. Chem.* **254**, 12219–12223.
Hammet, A., Pike, B. L., Mitchelhill, K. I., Teh, T., Kobe, B., House, C. M., Kemp, B. E. & Heierhorst, J. (2000). *FEBS Lett.* **471**, 141–146.
Huang, Q. Q., Teng, M. K. & Niu, L. W. (1999). *Acta Cryst.* D**55**, 1444–1448.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kingston, R. L., Baker, H. M. & Baker, E. N. (1994). *Acta Cryst.* D**50**, 429–440.
Kobe, B. & Deisenhofer, J. (1993). *J. Mol. Biol.* **231**, 137–140.
McPherson, A. (1999). *Crystallization of Biological Macromolecules.* Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
McPherson, A. (2001). *Protein Sci.* **10**, 418–422.
Perutz, M. F. (1968). *J. Cryst. Growth*, **2**, 54–56.
Stura, E. A., Satterthwait, A. C., Calvo, J. C., Kaslow, D. C. & Wilson, I. A. (1994). *Acta Cryst.* D**50**, 448–455.
Teh, T., Tiganis, T. & Kobe, B. (1999). *Acta Cryst.* D**55**, 561–563.