Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Optimization of buffer solutions for protein crystallization

Rajendrakumar A. Gosavi,[a] Timothy C. Mueser[b] and Constance A. Schall[a]*

[a]Department of Chemical Engineering, The University of Toledo, Toledo, OH 43606, USA, and [b]Department of Chemistry, The University of Toledo, Toledo, OH 43606, USA

Correspondence e-mail: constance.schall@utoledo.edu

Received 2 November 2007
Accepted 11 February 2008

Increasing the solubility of protein stock solutions to above that in a standard chromatography buffer (50 m*M* Tris–HCl pH 7.5, 100 m*M* NaCl) led to an increase in the number of crystallization conditions for ten globular proteins subjected to two crystal screens: the Index and Precipitant/Precipitant–Additive (P/PA) Screens. Solubility enhancement of protein stock solutions was achieved through screening and selection of buffer components to formulate an optimal buffer. Relative improvements in solubility were estimated through protection against the precipitation of protein by polyethylene glycol 8000. Proteins with limited solubility improvement in optimal buffer showed an enhancement in solubility on addition of glycerol. Maximum solubility was then determined by the concentration of optimized solutions until precipitate formed. The supernatant concentration then provided an estimate of the upper limit of protein solubility. This 'solubility' estimate is used to specify the initial concentration of the protein used in the screening experiments and is an important step in successful crystallization. Buffer optimization and establishment of initial protein concentration for crystal screening based on solubility estimates provides a methodology for improved crystal screening results.

1. Introduction

The number of candidate proteins for expression and structure determination from known genomic sequences is increasing at a rapid rate. In order to cope with this development, automation of crystallization is taking place with equal rapidity (Bard *et al.*, 2004). These developments have resulted in nanolitre-volume crystallization with hundreds to thousands of combinatorial crystal screening conditions (Zheng *et al.*, 2005; Luft *et al.*, 2001). However, obtaining diffraction-quality protein crystals is often a bottleneck in the structure determination of a protein molecule using X-ray or neutron crystallography. The success rate of screening in producing a diffraction-quality crystal is often low. Based on a recent review of 100 expressed proteins, about 9% of these proteins were crystallized after extensive screening (Chayen & Saridakis, 2002). Strategies that improve crystal screening success rates may lead to new structures and decrease the quantity of protein required for screening.

In combinatorial crystallization screening, proteins are often subjected to a sparse-matrix screen containing a set of finite crystallization conditions with varying salt, buffer, pH, precipitants and other chemical components for a few selected

temperatures (Cudney *et al.*, 1994). Further optimization of chemical constituents, protein concentration, temperature and other parameters is performed around conditions leading to microcrystals or precipitates to produce diffraction-quality crystals. Because of the vast number of parameters, the formulation of trial screens is often derived from macromolecular databases of past successful crystallization trials (Gilliland *et al.*, 1994; Berman *et al.*, 2000; Jancarik *et al.*, 2004). This approach has led to the design of screens that are specific for the crystallization of globular proteins, protein–nucleic acid complexes or membrane proteins.

The success of protein crystallization using these screens can depend on the chemical constituents of the protein stock solution or buffer in addition to the chemical components of the crystal screens. Each crystal screen solution may contain salts, buffers, alcohols and other additives or precipitants which differ from those in the protein stock solution. Thus, the crystallization of a protein molecule is dependent on the behavior of the molecules in the presence of up to five to six different chemical components. In previous studies, we have focused on selecting pH and salts for the protein stock solution prior to crystallization screening using a simple solubility screen (Collins *et al.*, 2004) in order to identify optimal buffers and to quantitate the improvement in protein solubility leading to improvement in protein crystallization results. This differs from the qualitative approach of Collins and coworkers, in which optimal buffers for proteins were selected by inspection of crystal screens (Collins *et al.*, 2005). Our hypothesis is that by enhancing the solubility of a protein in buffer to a judicious extent, crystallization is favored over the formation of amorphous precipitate or clear solutions. Protein crystal nucleation data analyzed by classical nucleation theory indicate that increasing the solubility lowers the interfacial energy between the crystal and the solution and the barrier to nucleation (Kulkarni & Zukoski, 2002). This suggests that excessive solubility enhancement can lead to near-zero barriers to nucleation, thus favoring the formation of amorphous precipitate; a high barrier can lead to excessively long lag times before crystal formation and on practical time scales leads to clear solutions (Durbin & Feher, 1996).

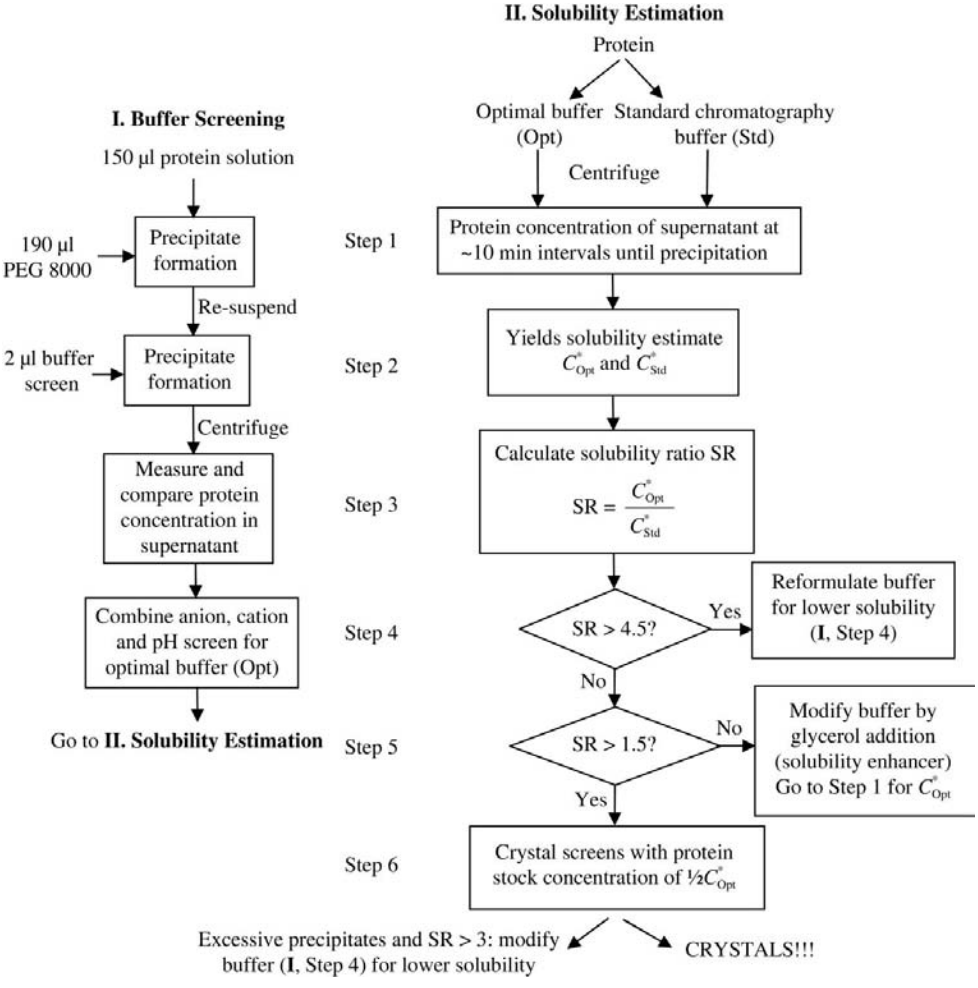

Figure 1
I. Buffer screening (Izaac *et al.*, 2006). Step 1: precipitate protein in PEG 8000. Step 2: resuspend and mix with buffer-screen components or water. Centrifuge sample and (step 3) measure protein concentration. Step 4: components that yield a protein concentration in supernatant greater than water are combined to form 'optimal buffer'. II. Solubility estimation. Step 1: concentrate protein to precipitate in standard chromatography and optimal buffers using centrifugal concentrators. Step 2: measure protein supernatant concentration for the solubility estimate. Step 3: calculate the protein solubility ratio SR. Step 4: if SR > 4.5, reformulate the optimal buffer by selecting a different combination of components from part I step 4 to produce a lower solubility. Step 5: if SR < 1.5, modify optimal buffer by adding a solubility enhancer (*e.g.* glycerol). Step 6: if 1.5 < SR < 4.5, prepare protein stock in optimal (or modified) buffer at half concentration and use in crystallization screens. If excessive precipitates and high SR ($\gtrsim$3) result, modify the buffer for lower protein solubility (part I step 4).

We routinely use a protein crystallization strategy consisting of four steps: (i) buffer selection and solubility assessment, (ii) sparse-matrix screening, (iii) grid search optimization and (iv) addition of polyalcohol cryoprotectants to the optimized conditions. In the first step, the individual salt and buffer components are selected for their ability to protect the protein from precipitation by PEG 8000 (Izaac *et al.*, 2006). Since common counter-ions are used, the 'optimal' ions and buffer components, defined as those ions and buffer components maximizing protein solubility in PEG 8000 compared with other standard

buffer components, are then combined to formulate an 'optimal buffer'. The protein in optimal buffer is then concentrated until precipitate forms. The protein concentration in the supernatant is a relative estimation of the 'maximum solubility' of the protein. In the second step, the protein in optimal buffer is subjected to sparse-matrix screens with the initial protein stock solution concentration set to one half of the estimated maximum solubility in the PEG 8000 solution supernatant. Selected screen conditions (producing either precipitate or microcrystals) are then subjected to fine grid screening with the local expansion of each condition. In the final step, the proteins are subjected to an additive screen in which optimized conditions are used with an additional coarse screen containing polyalcohol cryoprotectants such as glycerol, 2-methyl-2,4-pentanediol (MPD), ethylene glycol, propylene glycol or glucose. To complete this, water is replaced with the polyalcohol in the well solution in 5%(*v*/*v*) intervals from 0% to 20%. Subjecting a protein to such a polyalcohol screen has two advantages. Since some of these cryoprotectants improve protein solubility, nucleation events are minimized, with the concomitant formation of larger and fewer crystals. When the optimized conditions produce only crystal showers or crystalline precipitate, a protein solubility increase reduces the high supersaturation required for nucleation and may improve crystal-growth conditions. The second advantage of a polyalcohol screen is the ease of facilitating cryocooling of the protein crystals, as one of the components of the buffer contains a cryoprotectant. In these studies we have used one such cryoprotectant, glycerol, to improve protein solubility further when optimal buffer produced limited solubility improvement.

In a previous study, a solubility screen was implemented to determine optimal buffers for a set of ten globular proteins (Izaac *et al.*, 2006). The solubility of the proteins was increased in comparison to a standard chromatography buffer by using optimal buffers for seven of the ten proteins, but with only a marginal solubility improvement for two of the proteins. The results of the study showed an overall improvement in the quality of the crystals, as indicated by an increase in the number of large crystals using optimal buffer for a majority of the proteins. However, the combined total number of positive crystal hits in the screens showed only a slight increase for the optimal buffer. In our current study, we focus on the five proteins from the previous study which showed the least improvement. Here, we expand the previous optimal buffer studies (Izaac *et al.*, 2006) by increasing protein solubility further to formulate a 'modified buffer' with the addition of glycerol to those proteins which exhibited little, no or excessive improvement in solubility compared with the standard buffer. We show how the addition of glycerol improves the statistics of the number of positive crystal hits. The effect of changes in the initial protein concentration on crystal screen results is also assessed, with estimated protein solubility from precipitated solutions providing a means for selection of initial protein concentration for crystallization screening. A strategy for buffer selection and modification is formulated based on our findings (Fig. 1).

2. Materials and methods

2.1. Preparation of proteins

The proteins, which were chosen based on availability and known crystallizability, were purchased as lyophilized powders or as concentrated solutions from Sigma [catalase (C40), subtilisin (P5380), thaumatin (T7638), α-lactalbumin (L5385), trypsin (T1426), pepsin (P7012), ovalbumin (A5503) and myoglobin (M0630)] or from Hampton Research [xylanase (HR7-106) and D-xylose isomerase (HR7-102)]. Proteins were prepared as described in previous studies (Izaac *et al.*, 2006), with the exception of trypsin, for which a different protease inhibitor, benzamidine hydrochloride (BZD–HCl, Sigma–Aldrich), was added to prevent degradation in place of the AEBSF used in previous studies.

2.2. Solubility screening and optimal buffer determination

A summary of the steps in solubility and buffer screening is presented in Fig. 1. A set of cations (with a common anion), anions (with a common cation) and buffers of differing pH comprises the solubility screen (see supplemental material[1]). As described previously (Izaac *et al.*, 2006), the proteins were partially precipitated in 20%(*w*/*v*) polyethylene glycol (PEG 8000) in water, resuspended and combined with buffer-screen components. The protein remaining in solution in individual components of the buffer solubility screen was then compared with the amount of soluble protein partially precipitated in water alone. The anion and cation salts and buffer pH producing greater solubility than in water were combined and are referred to as the optimal buffer for that protein. Protein solutions were prepared in a standard chromatography buffer (50 m*M* Tris–HCl pH 7.5, 100 m*M* NaCl) as well as in optimal buffer (50 m*M* buffer, 100 m*M* salt). The maximum solubility of the protein in both these buffers was obtained by concentrating small samples in centrifugal concentrators (Amicon, Microcon YM-10, 10 000 Da molecular-weight cutoff; 0.5 ml capacity; Izaac *et al.*, 2006). The supernatant was sampled at regular intervals and examined for precipitate formation. An estimate of 'solubility' was obtained by concentrating protein solutions to the point at which precipitate formed. The ExPasy Swiss-Prot *ProtParam* tool was used to calculate the extinction coefficients (Gasteiger *et al.*, 2003). The protein supernatant concentration was measured using UV–Vis absorption.

The optimal buffer for each of the proteins comprised of salts and buffer components as described above. The solubility in the optimal buffer was compared with that in the standard buffer (50 m*M* Tris–HCl pH 7.5 and 100 m*M* NaCl). The ratio of the solubility of the protein in optimal buffer (Opt) to that in standard buffer (Std) yields the solubility ratio SR. For those proteins that showed an SR of $\leq$1.4 in previous studies (Izaac *et al.*, 2006), optimal buffers were adjusted by addition of glycerol to increase the SR to a value greater than 1.5 and less than 4.5. For those proteins with an SR greater than 3.5

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: BW5222). Services for accessing this material are described at the back of the journal.

Table 1
The estimated protein solubilities for the ten test proteins determined using centrifugal concentration of proteins until precipitate formation (Fig. 1).

The solubility ratio, SR, is defined as the ratio of solubility in optimal (Opt) or modified (Mod) buffer to that in standard chromatography (Std) buffer (50 m*M* Tris–HCl, 100 m*M* NaCl pH 7.5). Optimal buffer was comprised of 50 m*M* buffer and 100 m*M* salt (or no salt).

	Protein (source, calculated pI, MW in kDa)		Buffer	Solubility (mg ml^{-1}) Std	Opt (SR)	Mod (SR)
1	Catalase (bovine, 6.4, 57.6)	Opt	Na TAPS pH 8.5, trisodium citrate	11	40 (3.6)	19 (1.7)†
		Mod†	Na TAPS pH 8.5, 0.5% glycerol			
2	Subtilisin Carlsberg (*Bacillus licheniformis*, 6.6, 27.3)	Opt	Na TAPS pH 8.5, 10 m*M* $CaCl_2$	17	52 (3.1)	
3	Thaumatin (*Thaumatococcus danellii*, 8.5, 22.2)	Opt	Na PIPES pH 6.5, no salt	17	50 (2.9)	
4	Xylanase (*Tricoderma longribrahiatum*, 9.0, 21)	Opt	Na TAPS pH 8.5, sodium formate	28	76 (2.7)	
5	α-Lactalbumin (bovine, 4.8, 14.2)	Opt	Na TAPS pH 8.5, trisodium citrate	22	41 (1.9)	
6	D-Xylose isomerase (*Streptomyces rubiginosus*, 5.0, 43.3)	Opt	Na TAPS pH 8.5, sodium cacodylate pH 7.0	144	178 (1.2)	254 (1.8)†
		Mod†	Opt + 10% glycerol			
7	Trypsin (bovine, 8.7, 23.3)	Opt	Na TAPS pH 8.5, no salt	46	50 (1.1)	
		Opt‡	50 m*M* Na MES pH 7.5, 100 m*M* $CaCl_2$	13‡	55 (4.2)‡	
8	Pepsin (porcine, 3.4, 34.6)	Opt	Na TAPS pH 8.5, no salt	66	58 (0.9)	174 (2.6)†
		Mod†	Opt + 10% glycerol			
9	Ovalbumin (chicken, 5.2, 42.8)	Opt	Na HEPES pH 7.5, no salt	148	120 (0.8)	316 (2.1)†
		Mod†	Opt + 10% glycerol			
10	Myoglobin (equine, 7.4, 17)	Opt	Na TAPS pH 8.5, no salt	122	86 (0.7)	220 (1.8)†
		Mod†	Opt + 10% glycerol			

† Data for proteins in modified buffer from this work. ‡ Data for trypsin with BZD from this work. Other data are from Izaac *et al.* (2006).

and a large fraction of precipitated or phase-separated crystal screens (*e.g.* catalase), the buffer was reformulated to produce a more limited SR. In the case of catalase, the new buffer-screen components were comprised of 50 m*M* Na TAPS pH 8.5 with 0.5%(*v*/*v*) glycerol. Buffers with added glycerol are referred to as modified buffers (Mod).

2.3. Crystal screening and scoring

Two crystallization screens were used for proteins in standard, optimal or modified buffers: the commercially available sparse-matrix Index Screen (Hampton Research) and the Precipitant/Precipitant–Additive Screen (P/PA Screen) designed for this study (Izaac *et al.*, 2006). The crystal screens were set up at ambient temperature in 96-well three-drop Corning trays (Corning 3555, Hampton Research) using protein solutions at a concentration of half the maximum solubility in the standard optimal or modified buffers. All the crystallization trays were set up using a Honeybee 963 robot (Genomic Solutions, Irvine, California, USA) to dispense the screen solution into the reservoir as well as to dispense the crystallization drops onto the shelves (2 µl drops: 1 µl well solution + 1 µl protein solution). For the studies using protein stock solutions at low concentrations in optimal (LowOpt) or modified (LowMod) buffers, the initial protein concentration in the drop was matched to that used for protein in standard buffer.

After one week at ambient temperature, drops were imaged using a Rhombix digital imager (Data-Centric Automation, Nashville, Tennessee, USA) using both bright-field and polarized-light exposure in order to ensure proper interpretation. The drops were scored into two categories by visual inspection of images: negative results (clear drops, precipitate and phase separation) and positive results (microcrystals and large crystals).

3. Results

3.1. Protein solubility

The solubility of the proteins in the optimal buffers was increased over that in standard chromatography buffer (50 m*M* Tris–HCl pH 7.5 and 100 m*M* NaCl) for seven of the ten proteins, but with only marginal solubility improvement for two of the proteins in previous studies (Izaac *et al.*, 2006). The ratio of the solubility of the protein in optimal buffer to that in standard buffer yields the solubility ratio SR. For those proteins that showed an SR of $\leq$1.4 or of greater than 3.5 in the previous study, buffers were adjusted by the addition of glycerol to increase the SR to a value greater than 1.5 and less than 4.5. These buffers with added glycerol are referred to as modified buffer (Mod). For trypsin, a new optimal buffer was formulated with a change in the inhibitor to BZD.

The solubilities of proteins in standard buffer (Std; 50 m*M* Tris–HCl pH 7.5 and 100 m*M* NaCl), optimal buffer (Opt; 50 m*M* buffer, 100 m*M* salt) and modified buffer (Mod; optimal buffer plus added glycerol) are summarized in Table 1. The ratio of the solubility of protein in optimal buffer (Opt) or modified buffer (Mod) to that in standard buffer (Std) yields the solubility ratio SR.

The addition of glycerol to the optimal buffer resulted in increased protein solubility. The maximum solubility of D-xylose isomerase increased to 254 mg ml^{-1} (SR = 1.76) from 178 mg ml^{-1} (SR = 1.24). For pepsin, SR increased to 2.64 from 0.88. For ovalbumin the improved ratio was 2.14 compared with 0.81 and for myoglobin this ratio increased to 1.8 from 0.71. Standard and optimal buffer studies on trypsin

were carried out using the inhibitor benzamidine hydrochloride (BZD–HCl), which was more stable than the AEBSF used in previous studies (Izaac *et al.*, 2006). A new buffer with 50 m*M* Na MES pH 7.5, 100 m*M* $CaCl_2$ was obtained. Solubility measurements in the standard buffer and the new optimal buffer resulted in values of 13 and 55 mg ml^{-1}, respectively, an enhancement ratio (SR) of 4.2. Since this ratio is greater than 1.4, the optimal buffer required no further modification.

Catalase exhibited a large increase in solubility from standard buffer to optimal buffer (SR 3.64) in previous studies. A large fraction of previous screens contained precipitate and phase separation using optimal buffer compared with standard buffer (Izaac *et al.*, 2006). This may have been the result of excessive solubility enhancement. The optimal buffer for catalase from solubility-screening studies was formulated to produce a more limited SR using 50 m*M* Na TAPS pH 8.5 with 0.5%(*v*/*v*) glycerol. A change of buffer salts and pH alone resulted in a solubility below that of standard buffer. Addition of glycerol was required for modest solubility improvement. The solubility for catalase was improved from 11 mg ml^{-1} in standard buffer to 19 mg ml^{-1} in modified buffer (SR of 1.73).

3.2. Crystal screening results for proteins in optimal/modified buffer

Crystal screening results in modified buffer (Mod) are compared with those obtained in standard (Std) and optimal buffer (Opt) for the Index Screen and the Precipitant/Precipitant–Additive (P/PA) Screen in Table 2. Each crystallization drop was scored after a one-week incubation at 293 K as either clear (C), precipitate formation (ppt) or phase separation (Ph), collectively termed as negative hits, or as microcrystals (μx) or large crystals (lgx), collectively denoted as positive (pos) hits.

The total number of positive crystallization results for trypsin, pepsin, ovalbumin, myoglobin and catalase in Table 2 improved when modified buffer was used as opposed to standard or optimal buffers for the Index and P/PA Screens. D-Xylose isomerase was an exception and showed a reduced number of total positive hits with both screens in modified buffer. There was a small increase in the number of large crystals for D-xylose isomerase when modified buffer was used compared with that obtained in standard buffer. With the use of optimal or modified buffer for D-xylose isomerase, the number of drops exhibiting phase separation and precipitate decreased. Trypsin showed an increase in the total number of microcrystals and large crystals with both screens in optimal buffer compared with those obtained in standard buffer. Large crystals were seen in modified buffer that were absent in previous trials with optimal buffer. A decrease in the number of drops with precipitate was seen when optimal buffer was used in Index Screen. However, the number of drops with precipitates and phase separation increased in the P/PA Screen when optimal buffer was used. Pepsin exhibited an increase in the number of drops with microcrystals in modified buffer compared with that obtained in standard and optimal buffers. Additionally, a large crystal was seen in Index Screen which was absent in trials with standard or optimal buffers. As the buffer was altered from standard to optimal to modified buffer, the number of drops with precipitate and phase separation increased. For ovalbumin, no large crystals were observed when modified buffer was used, but there was an increase in the number of drops with microcrystals compared

Table 2
Ten test proteins subjected to two 96-condition crystallization screens: the Index Screen (Hampton Research) and the Precipitant/Precipitant–Additive Screen (P/PA; Izaac *et al.*, 2006).

For each protein, three or four buffer solutions were prepared: standard (Std), optimal (Opt), Modified (Mod) buffer and Mod or Opt buffer at low protein concentration (LowMod or LowOpt). Multi-drop trays were incubated at 293 K. Images of drops were visually examined and classified as clear (C), precipitate (ppt), phase separation (Ph), microcrystalline material (μx) or large crystals (lgx). The last two are positive (pos) results. The total number of drops in each category is given.

	Index Screen						P/PA Screen					
	C	ppt	Ph	μx	lgx	pos	C	ppt	Ph	μx	lgx	pos
Catalase												
Std	55	2	9	**9**	**21**	30	60	2	17	**14**	**3**	17
Opt	32	53	10	**0**	**1**	1	49	4	13	**19**	**0**	19
Mod†	22	35	4	**26**	**9**	35	60	17	1	**13**	**1**	14
LowMod†	28	61	1	**3**	**3**	6	33	56	0	**7**	**0**	7
Subtilisin												
Std	85	3	8	**0**	**0**	0	90	0	5	**1**	**0**	1
Opt	68	9	18	**1**	**0**	1	76	7	8	**5**	**0**	5
LowOpt†	66	27	0	**2**	**1**	3	78	17	0	**1**	**0**	1
Thaumatin												
Std	92	0	3	**0**	**1**	1	96	0	0	**0**	**0**	0
Opt	74	7	2	**2**	**11**	13	92	1	0	**2**	**1**	3
LowOpt†	61	18	2	**13**	**2**	15	74	6	0	**11**	**5**	16
Xylanase												
Std	29	3	1	**62**	**1**	63	63	17	11	**5**	**0**	5
Opt	62	4	9	**8**	**13**	21	74	8	4	**7**	**3**	10
LowOpt†	37	9	0	**50**	**0**	50	75	8	0	**13**	**0**	13
α-Lactalbumin												
Std	86	4	4	**1**	**1**	2	87	3	5	**1**	**0**	1
Opt	81	3	7	**3**	**2**	5	84	2	6	**3**	**1**	4
LowOpt†	81	7	0	**7**	**1**	8	61	26	0	**7**	**2**	9
D-Xylose isomerase												
Std	56	13	4	**7**	**16**	23	12	30	41	**11**	**2**	13
Opt	42	10	4	**13**	**27**	40	76	4	4	**7**	**5**	12
Mod†	60	14	0	**5**	**17**	22	49	14	25	**3**	**5**	8
LowMod†	54	22	0	**4**	**16**	20	52	35	0	**5**	**4**	9
Trypsin												
Std†	67	19	3	**7**	**0**	7	79	10	4	**3**	**0**	3
Opt	65	14	9	**8**	**0**	8	93	1	0	**2**	**0**	2
Opt†‡	59	16	3	**6**	**12**	18	58	16	8	**6**	**8**	14
LowOpt	73	6	1	**4**	**12**	16	77	9	3	**2**	**5**	7
Pepsin												
Std	83	4	6	**3**	**0**	3	86	1	7	**2**	**0**	2
Opt	64	14	4	**14**	**0**	14	81	5	6	**4**	**0**	4
Mod†	45	22	2	**26**	**1**	27	60	10	16	**9**	**0**	9
LowMod†	76	15	1	**4**	**0**	4	59	35	0	**2**	**0**	2
Ovalbumin												
Std	59	7	29	**1**	**0**	1	72	13	11	**0**	**0**	0
Opt	53	4	36	**2**	**1**	3	33	34	28	**1**	**0**	1
Mod†	60	17	4	**15**	**0**	15	37	31	25	**3**	**0**	3
LowMod†	53	27	13	**3**	**0**	3	57	30	1	**7**	**1**	8
Myoglobin												
Std	65	11	16	**4**	**0**	4	45	17	29	**5**	**0**	5
Opt	69	3	17	**7**	**0**	7	34	22	36	**4**	**0**	4
Mod†	78	6	0	**12**	**0**	12	54	25	10	**7**	**0**	7
LowMod†	72	19	0	**5**	**0**	5	51	41	0	**3**	**1**	4

† Data from this work. Other data are from Izaac *et al.* (2006). ‡ Trypsin with BZD inhibitor.

with that obtained in standard and optimal buffers. An increase in the total number of precipitates was seen with modified buffer. However, the number of drops with phase separations decreased in modified buffer in Index Screen. Myoglobin also exhibited an increase in the total number of positive hits in modified buffer compared with those obtained in standard and optimal buffers with both screens. For myoglobin, a reduction in the total number of drops with precipitate plus phase separation was seen when modified buffer was used.

For catalase, where the modified buffer was formulated to obtain a lower SR, an increase in the total number of positive hits (combined screens) was observed when compared with those obtained in standard or optimal buffer. The total number of drops (both screens) with precipitates plus phase separation was reduced when modified buffer was used instead of optimal buffer. In optimal buffer, where very significant solubility enhancement was observed, precipitate formation was favored, but as the SR was lowered (modified buffer) precipitate formation was suppressed.

For subtilisin, thaumatin, xylanase and α-lactalbumin, the optimal buffer from previous studies required no modification (SR > 1.4). Subtilisin showed an increase in the total number of positive hits in both screens with optimal buffer compared with standard buffer. An increase in the number of drops with precipitate and phase separation was also observed in both screens with optimal buffer. An increase in the number of microcrystals and large crystals was observed for thaumatin in both screens with optimal buffer compared with standard buffer. Xylanase showed a reduced number of drops with microcrystals in Index Screen when optimal buffer was used compared with standard buffer. However, the number of drops containing large crystals was greater with optimal buffer. In the P/PA Screen, an increase in the number of positive hits was observed for optimal buffer compared with standard buffer. α-Lactalbumin exhibited an increase in the total number of microcrystals and large crystals in both screens for optimal buffer compared with standard buffer.

When optimal buffer was used in our previous studies (Izaac *et al.*, 2006), the number of positive hits improved for eight of the ten proteins tested (the exceptions were catalase and xylanase). Addition of glycerol to the optimal buffer of five proteins and the change in the trypsin inhibitor (optimal buffer with BZD) improved the solubility ratio further. As a result, an improvement was seen in the total number of positive hits over that obtained in previous studies for five of the six proteins. A statistical analysis of the results is presented in §3.4.

3.3. Selection of initial protein concentration

The protein concentration used in stock solutions prepared for crystal screening experiments is an important factor in crystal formation. Most screens are designed with excess precipitating agent in order to ensure supersaturation. With excessively low protein concentrations, the addition of precipitants to the screening solution cannot generate an adequate level of supersaturation (a protein concentration greater than equilibrium solubility) to nucleate crystals. If protein concentration is excessively high, crystal nucleation is rapid, with the formation of microcrystals or amorphous precipitate. In our screening experiments, the concentration of the protein stock solution used in screens was set to approximately half of the solubility of its precipitate in that buffer (standard, optimal or modified). Because of the differences in the maximum solubilities of proteins in standard and optimal or modified buffers, the initial protein concentration was lower using standard buffer compared with optimal or modified buffers. We sought to answer the following questions. If we deviate from half of the maximum solubility in the protein buffer, are crystal screening results adversely affected? Does the estimated protein-solubility measurement in the protein buffer give guidance for selection of an appropriate protein concentration for successful screening?

A set of crystallization experiments were conducted in which the initial protein concentrations in all drops was held constant at one quarter of the solubility of each protein in standard buffer. For example, the solubility of catalase in standard buffer is 11 mg ml^{-1}. When setting up trays for catalase in standard buffer, the stock solution was 5.5 mg ml^{-1}. The protein stock solution is mixed with the crystal screen to produce a drop with an initial protein concentration of 2.25 mg ml^{-1} (one quarter of the solubility in the standard buffer). The solubility of catalase in modified buffer is 19 mg ml^{-1}. For the LowMod studies, a protein stock solution of 5.5 mg ml^{-1} of catalase in modified buffer was used to match the initial concentration to that in standard buffer. LowOpt protein stocks were similarly prepared. Therefore, the initial protein concentrations in the drops with modified (LowMod) or optimal buffer (LowOpt) were matched to the initial concentration in the standard buffer. D-Xylose isomerase, pepsin, ovalbumin, myoglobin and catalase were prepared in modified buffer (with added glycerol) and trypsin modified with the BZD inhibitor. The remaining proteins were prepared in optimal buffer. The crystal screening results are summarized in Table 3, where they are indicated by either LowMod or LowOpt depending on whether modified or optimal buffer was used.

In the studies with optimal or modified buffer at low protein concentration (Table 2), pepsin, ovalbumin, myoglobin, trypsin, subtilisin, thaumatin and α-lactalbumin showed an increase in the combined positive hits with Index Screen compared with the standard buffer. Xylanase, D-xylose isomerase and catalase exhibited reduced positive results in Index Screen compared with those in standard buffer. With the P/PA Screen, ovalbumin, trypsin, thaumatin, xylanase and α-lactalbumin showed a greater number of combined positive hits compared with standard buffer; pepsin and subtilisin showed the same number of positive hits. Two proteins, catalase and D-xylose isomerase, showed a decrease in combined positive hits compared with standard buffer. Trypsin, subtilisin and thaumatin showed an increase in the number of large crystals compared with standard buffer in the Index Screen. In the P/PA Screen, D-xylose isomerase, ovalbumin, myoglobin,

Table 3
Crystallization screening results (Table 2) for proteins prepared in standard buffer statistically compared with those for proteins prepared (*a*) in optimal buffer (data from Izaac *et al.*, 2006), (*b*) in modified/optimal buffer with improved solubility ratios and (*c*) at low concentration (matching concentration used for standard buffer) in modified or optimal buffer (LowMod or LowOpt).

The means, standard deviations and *p*-values from the Mann–Whitney rank sum test are given for comparisons of positive hits, microcrystals, large crystals, clear drops or precipitates. *p*-values of 0.10 or less indicate significant differences in the comparison groups; these comparisons are highlighted in bold. These values are based on the comparison of average results for the ten proteins in each screen.

(*a*) Standard (Std) and optimal (Opt) buffer (Izaac *et al.*, 2006).

	Index Screen			P/PA Screen		
	Average (± standard deviation)			Average (± standard deviation)		
	Std buffer	Opt Buffer	*p*-value	Std buffer	Opt Buffer	*p*-value
Positive hits	13.4 (20.2)	12.3 (12.0)	0.18	4.5 (5.9)	6.4 (5.6)	0.12
Microcrystals	8.7 (18.9)	5.8 (4.9)	0.18	4.0 (4.8)	5.4 (5.2)	0.13
Large crystals	4.0 (7.7)	5.6 (8.9)	0.21	0.5 (1.1)	1.0 (1.7)	0.25
Clear drops	66.8 (19.5)	61.0 (14.8)	0.30	70.5 (26.5)	69.2 (22.4)	0.37
Precipitate	6.3 (5.3)	12.1 (14.9)	0.15	8.4 (10.3)	8.8 (10.8)	0.25

(*b*) Std and modified/optimal (Mod/Opt) buffer.

	Index Screen			P/PA Screen		
	Average (± standard deviation)			Average (± standard deviation)		
	Std buffer	Mod/Opt Buffer	*p*-value	Std buffer	Mod/Opt Buffer	*p*-value
Positive hits	13.4 (20.8)	16.9 (10.0)	0.11	**4.7 (5.8)**	**7.7 (4.1)**	0.04†
Microcrystals	**9.4 (18.7)**	**10.4 (9.3)**	0.10†	**4.2 (4.8)**	**5.8 (3.4)**	0.08†
Large crystals	4.0 (7.7)	6.5 (6.5)	0.17	**0.5 (1.1)**	**1.9 (2.7)**	0.08†
Clear drops	67.7 (19.3)	60.9 (17.3)	0.25	69.0 (25.5)	64.4 (16.8)	0.17
Precipitate	**6.6 (5.9)**	**13.3 (9.8)**	0.04†	9.3 (9.9)	13.1 (9.6)	0.18

(*c*) Opt buffer and Std and Opt/Mod at low protein concentrations (LowOpt Or LowMod)

	Index Screen			P/PA Screen		
	Average (± standard deviation)			Average (± standard deviation)		
	Std buffer	LowMod/LowOpt buffer	*p*-value	Std buffer	LowMod/LowOpt buffer	*p*-value
Positive hits	13.4 (20.2)	13.0 (14.3)	0.16	**4.7 (5.8)**	**7.6 (4.6)**	0.06†
Microcrystals	9.4 (18.7)	9.5 (14.5)	0.23	4.2 (4.7)	5.8 (3.9)	0.12
Large crystals	4.0 (7.70	3.5 (5.7)	0.35	**0.5 (1.1)**	**1.8 (2.1)**	0.07†
Clear drops	67.7 (19.3)	60.1 (17.3)	0.15	**69.0 (25.4)**	**61.7 (14.5)**	0.09†
Precipitate	**6.6 (5.9)**	**21.1 (15.9)**	0.003†	**9.3 (9.9)**	**26.3 (16.2)**	0.01†

† Statistically significant *p*-value.

trypsin, thaumatin and α-lactalbumin showed an improved number of drops with large crystals compared with those in standard buffer. All proteins (except trypsin) showed an increased total number of precipitates (combined screens). However, the number of drops with phase separation was lower compared with standard buffer in both screens.

When crystallization experiments were set up with modified or optimal buffer at concentrations approximating that of standard buffer (LowMod or LowOpt in Table 2), six proteins showed an increase the number of the combined positive hits (total of both screens) compared with those obtained in standard buffer. However, when the comparison group is the optimal or modified buffer with higher protein concentration of half maximum solubility (Opt or Mod), seven of the ten proteins exhibited a decrease in total positive hits using low protein concentration (LowMod or LowOpt). This indicates that solubility enhancement in the protein buffer combined with proper selection of protein stock solution concentration provides the best screening results. The screening results are analyzed in the following sections to ascertain whether the differences between comparison groups are statistically significant.

3.4. Statistical analysis of crystallization results

3.4.1. Buffer selection. A comparison of the mean number of crystallization hits for the ten proteins in each screen using standard *versus* optimal buffer for previous studies (Izaac *et al.*, 2006) was evaluated using a nonparametric Mann–Whitney rank sum test (Rosner, 2000) with *Sigma Stat* 3.1 software (Systat). Means were calculated and compared for positive hits, microcrystals, large crystals, clear drops and precipitates in each screen separately (Table 3*a*). A statistical comparison was also made of the mean number of crystallization hits for standard buffer *versus* the combination of optimal buffer (trypsin, subtilisin, thaumatin, xylanase and α-lactalbumin) and modified buffer (catalase, D-xylose isomerase, pepsin, ovalbumin and myoglobin) in Table 3(*b*). The null hypothesis of the single-tailed Mann–Whitney comparison test is that the mean of one group is less than or equal to that of the comparison group. The results of this test provide a probability value (*p*-value) defined as the probability of wrongly rejecting the null hypothesis. Small *p*-values suggest that the mean for one group is greater than that of the comparison group and that differences between the two groups are significant. Considering the stochastic nature of protein crystallization experiments, a significance level of 90% (a *p*-value of 0.10) was chosen. The *p*-values, averages and standard deviations for crystal screening results for the ten proteins prepared in standard buffer, optimal buffer or a combination of modified and optimal buffers are shown in Table 3 for the Index and P/PA Screens.

Comparing data from our previous work (Izaac *et al.*, 2006) between standard and optimal buffers (Table 3*a*) in the P/PA Screen shows that the optimal buffer produced a greater number of positive hits compared with the standard buffer (*p*-value of 0.12), although the differences are not statistically significant. With Index Screen, the average number of positive hits was lower for optimal buffer compared with standard buffer. However, these differences were not statistically

significant. With buffer modification to improve the solubility ratio (SR, Table 1) comparisons were made of the number of crystallization hits for standard *versus* modified/optimal buffers (Table 3*b*). The number of positive hits was significantly greater for modified/optimal buffers than for standard buffer in both the P/PA (p-value = 0.038) and Index Screens (p-value = 0.106). With standard *versus* modified/optimal buffer, the number of microcrystal hits are significantly greater in both P/PA (p-value = 0.081) and Index Screens (p-value = 0.099) when modified/optimal buffers are used. However, large crystal hits were significantly greater only with P/PA Screen (p-value = 0.075) in optimal/modified buffer *versus* standard buffer. The average number of precipitated drops obtained with modified/optimal buffer was significantly greater than with standard buffer in Index Screen (p-value = 0.035).

By further improving the solubility of proteins with a limited or excessive SR through buffer modification, a statistically greater number of microcrystal hits over that of standard buffer was achieved. An significant increase in positive hits for the P/PA Screen and the number of precipitates with Index Screen for optimal/modified buffer is observed. For screening experiments leading to no crystal hits at all, precipitates are often used as a starting point for further screening since adjustment of the precipitating agents may result in the formation of crystals.

3.4.2. Initial protein concentration. A comparison of the mean number of positive hits and other screening results for the ten proteins using standard buffer *versus* modified/optimal buffer with matched protein concentrations, LowMod/LowOpt, was made using a Mann–Whitney Rank Sum test as described in §3.4.1. The results are presented in Table 3(*c*). With reduced protein concentrations, the mean numbers of positive hits with Index Screen are approximately equal. However, the number of positive hits in the P/PA Screen is significantly greater (p-value = 0.061) for LowMod/LowOpt buffer *versus* standard buffer. In the P/PA Screen the average number of microcrystals is greater with a low p-value (0.121), whereas the average number of large crystals is significantly greater (p-value = 0.065) in LowMod/LowOpt buffer *versus* standard buffer. A significantly greater number of drops with precipitates were seen with both Index (p-value = 0.003) and P/PA Screens (p-value = 0.009) for LowMod/LowOpt buffer *versus* standard buffer. Solubility enhancement alone, without adjustment to the initial protein concentration, appears to result in only modest improvement in crystal hits. This is in contrast to results with an initial concentration of half the protein solubility (Table 3*b*), where improvements in screening results are statistically significant.

4. Discussion

The selection of protein stock buffer components was optimized in an attempt to increase 'solubility' over that of a standard chromatography buffer. Other methods for increasing the stability of proteins and hence improving the probability of obtaining diffraction-quality crystals includes specific surface mutations such as glutamine to aspartic acid or alanine (Mateja *et al.*, 2002), However, this often requires structural information, which is lacking for a previously uncrystallized protein. Since measurement of the solubility of protein crystals is also not possible for a previously uncrystallized protein, an estimate of solubility in a particular solvent or buffer can be made from the protein in its precipitated form. The precipitated protein is not the thermodynamically stable solid state and the solubility of precipitates will be greater than that of the crystal form. Despite this, the solubility of the precipitated protein can provide insight into the relative effects of buffer or solvent formulation on protein solubility. In the results reported here, for buffers formulated to yield an increase in solubility over that of the standard buffer (solubility ratios greater than 1.6), an increase in positive crystallization hits was achieved (Tables 2 and 3). Solubility enhancement was achieved through buffer optimization of anion, cation and pH selection, supplemented where necessary with the addition of glycerol to further increase solubility.

Solubility enhancement may be effective in increasing the success of crystal screening, owing in part to enhancement of crystal nucleation. Addition of components that increase the solubility of crystallizing proteins have been found to decrease nucleation induction time at fixed values of protein supersaturation (Kulkarni & Zukoski, 2002). Supersaturation is often expressed as the ratio of protein bulk-solution concentration to equilibrium solubility. The addition of glycerol and ethylene glycol to hen egg-white lysozyme with sodium chloride precipitant resulted in an increase in lysozyme solubility and a decrease in nucleation-induction times (Kulkarni & Zukoski, 2002). A similar trend was observed by Lu and coworkers (Lu *et al.*, 2002, 2003) with the addition of glycerol and dimethyl sulfoxide to lysozyme solutions, resulting in increased solubility and a decrease in the supersaturation required for spontaneous nucleation.

The rate expression for nucleation from classical nucleation theory can be written as

$$J = J_0 \exp\left[\frac{-B}{(\ln S)^2}\right] \quad \text{with} \quad B = 4\left(\frac{\pi}{3}\right)^3 (\overline{\gamma})^3, \qquad (1)$$

where the prefactor J_0 is proportional to the concentration of the crystallizing units and the diffusion of those units to the crystal surface. The exponential factor B is proportional to the dimensionless crystal–solute interfacial tension, $\overline{\gamma} = (\gamma d^2/kT)$, where d is the molecular diameter, k is Boltzmann's constant and T is the absolute temperature. S is defined as the relative supersaturation, often expressed as the ratio of bulk solute (protein in our case) to equilibrium concentration. The exponential term can be viewed as an 'activation barrier' to crystal nucleation. In analyzing their nucleation-induction data using classical nucleation theory, Kulkarni & Zukoski (2002) concluded that solubility enhancement reduced the crystal–solute interfacial tension, effectively lowering the barrier to crystal nucleation. An increase in solubility also

results in an increase in protein concentration at a fixed value of S, with a concomitant increase in crystal nucleation rates.

The interfacial energy γ and supersaturation S constitute two important terms in the energy barrier to nucleation. At constant supersaturation, a decrease in γ lowers the nucleation barrier, thus facilitating crystal nucleation. However, lowering this barrier to values close to zero through excessive solubility enhancement can lead to the formation of amorphous precipitates rather than crystals (Durbin & Feher, 1996). With catalase, adjustment of the solubility ratio SR using modified (SR = 1.7) *versus* optimal (S = 3.6) buffer reduced precipitate formation and increased the number of positive crystal hits.

Significant differences in crystal screening results were found for protein stock solutions prepared in standard chromatography buffer compared with proteins prepared in optimal/modified buffers with SR greater than 1.6 for the ten proteins studied with the Index or P/PA Screens. These results infer that improving the protein solubility in stock solutions is an important parameter in searching for crystallization conditions through screening experiments. The SR ratio is neither a measure of protein solubility in the crystal form nor a measure of solubility in the composite protein stock/screen solutions. However, the solubility enhancement in protein stock may effectively act to enhance crystal nucleation in the screen solutions.

Determination of estimated solubility helps to set an appropriate protein concentration for crystal screening experiments. The best screening results were those with initial protein stock solution concentrations of half the maximum solubility. When stock solution concentration in optimal/modified buffers was decreased (LowOpt/LowMod in Tables 2 and 3) to that of standard buffer, the number of positive crystal hits decreased. Thus, solubility screening may also allow the approximation of initial protein concentration for screening protein crystallization conditions.

Buffer choice for protein stock solutions is an important factor for successful crystallization screening. Buffer screening, combined with an estimate of protein solubility from precipitated solutions, provides a methodology for protein stock solution preparation, leading to improved screening results (Fig. 1).

References

Bard, J., Ercolani, K., Svenson, K., Olland, A. & Somers, W. (2004). *Methods*, **34**, 329–347.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.

Chayen, N. E. & Saridakis, E. (2002). *Acta Cryst.* D**58**, 921–927.

Collins, B., Stevens, R. C. & Page, R. (2005). *Acta Cryst.* F**61**, 1035–1038.

Collins, B. K., Tomanicek, S. J., Lyamicheva, N., Kaiser, M. W. & Mueser, T. C. (2004). *Acta Cryst.* D**60**, 1674–1678.

Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y. & McPherson, A. (1994). *Acta Cryst.* D**50**, 414–423.

Durbin, S. D. & Feher, G. (1996). *Annu. Rev. Phys. Chem.* **47**, 171–204.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D. & Bariocj, A. (2003). *Nucleic Acids Res.* **31**, 3784–3788.

Gilliland, G. L., Tung, M., Blakeslee, D. M. & Ladner, J. E. (1994). *Acta Cryst.* D**50**, 408–413.

Izaac, A., Schall, C. A. & Mueser, T. C. (2006). *Acta Cryst.* D**62**, 833–842.

Jancarik, J., Pufan, R., Hong, C., Kim, S.-H. & Kim, R. (2004). *Acta Cryst.* D**60**, 1670–1673.

Kulkarni, A. M. & Zukoski, C. F. (2002). *Langmuir*, **18**, 3090–3099.

Lu, J., Wang, X.-J. & Ching, C.-B. (2002). *Prog. Cryst. Growth Charact.* **45**, 201–217.

Lu, J., Wang, X.-J. & Ching, C.-B. (2003). *Cryst. Growth Des.* **3**, 83–87.

Luft, J. R., Wolfley, J., Jurisica, I., Glasgow, J., Fortier, S. & DeTitta, G. T. (2001). *J. Cryst. Growth*, **232**, 591–595.

Mateja, A., Devedjiev, Y., Krowarsch, D., Longenecker, K., Dauter, Z., Otlewski, J. & Derewenda, Z. S. (2002). *Acta Cryst.* D**58**, 1983–1991.

Rosner, B. (2000). *Fundamentals of Biostatistics*, pp. 331–347. Pacific Grove, USA: Duxbury Press.

Zheng, B., Gerdts, C. J. & Ismagilov, R. F. (2005). *Curr. Opin. Struct. Biol.* **15**, 548–555.