Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

A preliminary solubility screen used to improve crystallization trials: crystallization and preliminary X-ray structure determination of *Aeropyrum pernix* flap endonuclease-1

Brandon K. Collins,[a] Stephen J. Tomanicek,[a] Natasha Lyamicheva,[b] Michael W. Kaiser[b] and Timothy C. Mueser[a]*

[a]Department of Chemistry, The University of Toledo, Toledo, OH 43606, USA, and [b]Third Wave Technologies Inc., Madison, WI 53719, USA

Correspondence e-mail: timothy.mueser@utoledo.edu

Received 21 October 2003
Accepted 23 July 2004

Crystallization of protein and protein complexes is a multi-parametric problem that involves the investigation of a vast number of physical and chemical conditions. The buffers, salts and additives used to prepare the protein will be present in every crystallization condition. It is imperative that these conditions be defined prior to crystal screening since they will have a ubiquitous involvement in the crystal-growth experiments. This study involves the crystallization and preliminary analysis of the flap endonuclease-1 (FEN-1) DNA-repair enzyme from the crenarchaeal organism *Aeropyrum pernix* (*Ape*). *Ape* FEN-1 protein in a standard chromatography buffer had only a modest solubility and minimal success in crystallization trials. Using an ion/pH solubility screen, it was possible to dramatically increase the maximum solubility of the protein. The solubility-optimized protein produced large diffraction-quality crystals under multiple conditions in which the non-optimized protein produced only precipitate. Only minor adjustments of the conditions were required to produce single diffraction-quality crystals. The native *Ape* FEN-1 crystals diffract to 1.4 Å resolution and belong to space group $P6_1$, with unit-cell parameters $a = b = 92.8$, $c = 80.9$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$.

1. Introduction

Crystallization experiments require control of the rates at which macromolecules are subjected to supersaturating conditions. A phase diagram is typically used to describe saturation limits and an experimental trajectory defines the conditions witnessed by the macromolecular solute during equilibration (reviewed in McPherson, 1999). Supersaturating conditions are obtained by increasing the concentrations of precipitating agents. By definition, a precipitating agent is any component that affects the solubility of the macromolecule. The profile of the boundary between unsaturated and supersaturated states, the saturation level, is defined by the macromolecule's preference between solution and solid states. It is clear that this boundary is in part defined by the chemical components used to initially solubilize the macromolecule. The majority of purification protocols utilize standard salts and buffers, typically Tris–HCl and NaCl, chosen on the basis of expense. It is then generally assumed that crystallization screening can then be used to discover the conditions for stable supersaturation, ordered assembly and subsequent crystal growth. The initial solvent, ubiquitous in all crystallization trials, may impose a ceiling on the supersaturation level, perhaps even collapsing the metastable region of the phase diagram. In certain cases, where the macromolecules are soluble at low ionic strength, as is the case for hemoglobin and lysozyme, the effect of the initial solvent is minimal. However, when a more typical protein is studied, substantial salting-in may be required. These conditions, unless optimized, may mask crystallization results.

We utilize a preliminary solubility screen for the purpose of optimizing the salt cations and anions and the buffer and pH in order to prepare customized solvents that raise the ceiling of the saturation levels at an overall lower ionic strength. The presumption is that we are selecting between amorphous precipitate and ordered assembly. Maximizing solubility is in effect masking amorphous precipitation and the resultant exclusion of water. An ordered assembly, *i.e.* a crystal lattice, of a macromolecule incorporates a large fraction of solvent. A protein crystal with 50% solvent content is in fact at a concentration of $\sim$500 mg ml^{-1}. With this in mind, perhaps it is best to consider a macromolecular crystal as the ultimate in solubility. Another consideration is the rate of formation of amorphous precipitation *versus* the rate of ordered assembly. One can envision that the rate of precipitation is rapid in comparison to nucleation and crystal growth. If conditions are found under which interactions of amorphous precipitation are eliminated, then the super-

saturation state can exist over a longer period of time, promoting the probability of nucleation events. With these assumptions made, an optimal solvent is defined as one which maximizes solubility at the lowest ionic strength. One anticipates that these conditions will enhance the positive results found in sparse-matrix, systematic and incomplete factorial crystallization trials.

In our studies, we incorporate a solubility screen to improve crystallization results. Here, we report the crystallization of a crenarchaeal flap endonuclease-1 (FEN-1) in which the solubility screen had a profound effect on our success. FEN-1s are members of the rad2/rad27 family of DNA-repair enzymes (Harrington & Lieber, 1994*b*) which includes the prokaryotic polymerase-associated 5′ to 3′ exonucleases, the eukaryotic flap endonucleases (DNase IV), yeast rad2 and rad27 and the human XPG protein (Harrington & Lieber, 1994*a*). To date, the X-ray crystal structures of six enzymes in this family have been determined: one prokaryotic source, the 5′ to 3′ exonuclease domain of *Thermus aquaticus* polymerase (PDB code 1taq; 2.40 Å resolution; Kim *et al.*, 1995), two from bacteriophages, the T4 RNase H (PDB code 1tfr; 2.1 Å resolution; Mueser *et al.*, 1996) and the T5 5′ to 3′ exonuclease (PDB code 1exn; 2.50 Å resolution; Ceska *et al.*, 1996, Garforth *et al.*, 1999), and three from euryarchaeal organisms, *Pyrococcus furiosus* FEN-1 (PDB code 1b43; 2.00 Å resolution; Hosfield *et al.*, 1998), *Methanococcus jannaschii* FEN-1 (PDB codes 1a76 and 1a77; 2.00 Å resolution; Hwang *et al.*, 1998) and *P. horikoshii* (Pho) FEN-1 (PDB code 1mc8; 3.10 Å resolution; Matsui *et al.*, 2002). The preferred substrate of the FEN-1 enzymes is branched DNA that has both a short 5′ ssDNA flap and downstream duplex DNA (Harrington & Lieber, 1994*a*). Here, we report the crystallization of the first crenarchaeal flap endonuclease-1, with diffraction to the highest resolution reported for this family of enzymes.

2. Materials and methods

2.1. Expression and purification of *Aeropyrum pernix* FEN-1

A. pernix (*Ape*) FEN1 protein was expressed and purified at Third Wave Technologies Inc. Substantial expression of soluble protein was obtained from IPTG induction in a bacterial host (*Escherichia coli* BL21). Bacterial cell pellets were resuspended in lysis buffer (10 ml Tris–HCl pH 7.5, 100 m*M* NaCl, 2 m*M* EDTA) with 10 mg hen egg-white lysozyme and incubated for 15 min at 277 K. Deoxycholic acid [200 µl, 10%(*w*/*v*) solution] was added and the solution was sonicated (2 min, 80% power) and centrifuged (14 000*g*, 15 min, 277 K). The supernatant was decanted and then heated (340 K, 1 h) in order to denature endogenous bacterial proteins. Polyethyleneimine [0.25 ml, 10%(*v*/*v*) solution] was added and the solution was incubated (30 min, 277 K) and centrifuged (14 000*g*, 15 min, 277 K), with the protein remaining in the clarified supernatant. The protein was precipitated by the addition of ammonium sulfate (0.476 g ml^{-1}) and the solution was incubated (30 min, 277 K) then centrifuged (14 000*g*, 15 min, 277 K) and the supernatant discarded. The pellet was resuspended in buffer *A* (5 ml, 50 m*M* Tris–HCl pH 8.0, 1 m*M* EDTA) and dialyzed (buffer *A*, overnight, 277 K). The protein was run on a heparin HPLC column and the peak that eluted at 0.5–0.8 *M* NaCl was collected. The pooled fractions were dialyzed (buffer *A*, 277 K, overnight), followed by a second dialysis in a storage buffer [buffer *A* with 50%(*v*/*v*) glycerol, 277 K, overnight] and stored at 193 K.

2.2. Solubility screen

A sample of purified protein (5 mg) was dialyzed in deionized water (277 K, overnight) to produce a flocculent precipitate. The precipitate was resuspended, immediately aliquoted (20 samples in 1.5 ml centrifuge tubes) and then centrifuged (20 000*g*, 5 min). The supernatant was collected from the samples and tested for protein remaining in solution. Using four protein pellets, a series of Good buffers (Good *et al.*, 1966; 10 µl per sample, 50 m*M* each, Na MES pH 5.6, Na PIPES pH 6.5, Na HEPES pH 7.5, Na TAPS pH 8.5) were tested to determine the pH profile of solubility. Using six protein pellets, a series of chloride salts were tested to determine the best cation (10 µl per sample, 100 m*M* each, NH_4Cl, NaCl, KCl, LiCl, $MgCl_2$, $CaCl_2$). Using six protein pellets, a series of sodium salts were tested to determine the best anion (10 µl per sample, 100 m*M* each, sodium formate, sodium acetate, sodium cacodylate, Na_2SO_4, Na_2HPO_4, disodium hydrogen citrate). The pellets were resuspended in the added test solutions, incubated (10 min, ambient temperature) and centrifuged (5 min, 20 000*g*). The individual samples were tested for protein in the supernatant (2 µl each, Bio-Rad Protein Assay, 5 min incubation, 595 nm absorbance). The remaining four protein pellets were used to test combinations of the best cation, anion and pH.

2.3. Crystallization and data collection

Crystallization trials were conducted twice, first using the protein dialyzed into a standard solvent (7 mg ml^{-1}, 50 m*M* Tris–HCl, 150 m*M* NaCl) and a second time using the results from the solubility screen (10 mg ml^{-1}, 25 m*M* Na PIPES pH 6.5, 50 m*M* disodium hydrogen citrate, 50 m*M* KCl, 50 m*M* NH_4Cl). Both solutions were subjected to sparse-matrix crystal screens [1 + 1 µl hanging drop, Crystal Screens 1 and 2 and Natrix from Hampton Research, Laguna Niguel, CA, USA, Wizard1, Wizard2, Cryo1 and Cryo2 from DeCode Genetics and the Ion Screen (Mueser *et al.*, 2000)] at 277 and 294 K. The 'Ion Screen' is the precursor to the PEG/Ion Screen (Bob Cudney, Hampton Research, personal communication). Crystals from the solvent-optimized screen were obtained in 2 days [20%(*w*/*v*) PEG 4000, 200 m*M* sodium formate, 100 m*M* Na MES pH 5.6, 294 K]. Diffraction-quality crystals were produced from a hanging-drop gradient expansion [4 + 4 µl, 10–14%(*w*/*v*) PEG 4000, 200 m*M* sodium formate and 100 m*M* Na MES pH 5.6, 294 K].

For data collection, the crystals were placed momentarily in a substitute mother liquor containing a cryoprotectant [25%(*v*/*v*) 2-methyl-2,4-pentanediol, 12%(*w*/*v*) PEG 4000, 25 m*M* Na PIPES pH 6.5, 100 m*M* Na MES pH 5.6, 200 m*M* sodium formate, 50 m*M* disodium hydrogen citrate, 50 m*M* KCl, 50 m*M* NH_4Cl] and flash-frozen in liquid nitrogen. Several native data sets were collected at BioCARS 14-BMC (Argonne National Laboratories, Advanced Photon Source, Chicago, IL, USA) using an ADSC Quantum 4 CCD detector. The best data set, reported here, diffracted to 1.4 Å resolution. The high-resolution data were collected first (0.9 Å wavelength, crystal-to-detector distance 120 mm, 0.5° oscillations, 15 s exposures, 150 frames, 100 K) followed by a low-resolution data collection (0.9 Å wavelength, crystal-to-detector distance 170 mm, 1.0° oscillations, 10 s exposures, 100 frames, 100 K). This low-resolution pass proved to have substantial rejections owing to oversaturation and a third pass was collected (0.9 Å wavelength, crystal-to-detector distance 170 mm, 1.0° oscillations, 2 s exposures, 100 frames, 100 K). The data were integrated using *DENZO* and merged using *SCALEPACK* (Otwinowski & Minor, 1997).

3. Results and discussion

Recombinant flap endonuclease-1 from the crenarchaeal *A. pernix* (*Ape* FEN-1) was expressed in *E. coli* and purified to homogeneity. *A. pernix* is an aerobic hyperthermophilic ocean-vent organism with an optimal growth temperature approaching 373 K (Sako *et al.*, 1996). Purified *Ape* FEN-1 was stable in differential scanning calorimetry to above 403 K, the limits of the instrument (Collins, unpublished data). The purification involved heating the crude extract to 340 K for 1 h in order to denature endogenous proteins and centrifugation to remove denatured material, followed by purification in one step on a heparin Sepharose column. The protein had limited solubility ($\sim$3 mg ml^{-1} at 277 K, $\sim$9 mg ml^{-1} at 294 K) in the storage buffer. A solubility screen was employed to maximize the solubility of the protein (Table 1). The use of common counterions allowed a direct comparison of anions and cations in the screen. The results clearly indicate that the optimal buffer is Na PIPES pH 6.5, the best anion is citrate and the best cation is magnesium. In a solution of dipotassium hydrogen citrate (50 m*M*, pH 6.5) with $MgCl_2$ (10 m*M*), the maximum solubility of *Ape* FEN-1 was a remarkable 106 mg ml^{-1} at 294 K.

Divalent cations are essential for nuclease activity, in which two magnesium ions are found in the active site (Mueser *et al.*, 1996). Our interest is to solve the structure with and without active-site metals. Under the assumption that a metal-free crystal form could be soaked with metals to obtain both structures, we decided to attempt crystallization trials using a metal-free composite buffer containing the addition of disodium hydrogen citrate and KCl. The protein in this buffer had a maximum solubility of $\sim$21 mg ml^{-1} and crystal trials were conducted at 10 mg ml^{-1} (48% of the maximum). For comparison, we also conducted crystallization trials of the non-optimized protein in the standard buffer (7 mg ml^{-1}, 78% of the maximum). In the standard solvent, the protein solubility was similar to that of the storage buffer ($\sim$9 mg ml^{-1} at 294 K). We obtained large crystals from the optimized buffer directly from crystal screens with only marginal results for protein in standard buffer (Fig. 1). A summary of the comparison of the two crystal trials is presented in Fig. 2. Very small crystals and crystalline material were noted in a few screen conditions of protein in standard buffer. The results using optimized buffer were substantially better. There were no results where crystals grew in standard buffer but not in optimized buffer. The optimized buffer has far fewer clear and precipitate results, with a dramatic increase in the number of crystalline results, including the large crystals shown (Fig. 1).

Linear-gradient expansions of the best crystallization condition obtained for the optimized protein produced diffraction-quality single crystals (Fig. 3) which diffracted to 1.4 Å resolution. The data were processed using *DENZO*/*SCALEPACK*

Table 1
Solubility-screen results for *Ape* FEN-1.

Samples of precipitated protein were solubilized by the individual components using a set of common cations (Na^+), common anions (Cl^-) and buffers to allow direct comparison of effectiveness. Concentrations of proteins are given (mg ml^{-1}); the best anion is citrate and the best cations are Mg^{2+} and Ca^{2+}. A composite solution of Na PIPES pH 65, disodium hydrogen citrate and KCl improved the solubility to 21 mg ml^{-1}. A combination of the best anion and best cation (dipotassium hydrogen citrate/$MgCl_2$) improved the solubility to 106 mg ml^{-1}.

Supernatant	Concentration (mg ml^{-1})
Na MES pH 5.6	0.39
Na PIPES pH 6.5	1.64
Na HEPES pH 7.5	0
Na TAPS pH 8.5	0
NH_4Cl	0.60
NaCl	0.92
KCl	1.63
LiCl	1.51
$MgCl_2$	8.50
$CaCl_2$	8.90
Sodium formate	1.36
Sodium acetate	0.64
Sodium cacodylate	0
Na_2SO_4	5.93
Na_2HPO_4	1.06
Disodium hydrogen citrate	8.73
Na PIPES/disodium hydrogen citrate	21.0
Dipotassium hydrogen citrate/$MgCl_2$	106.0

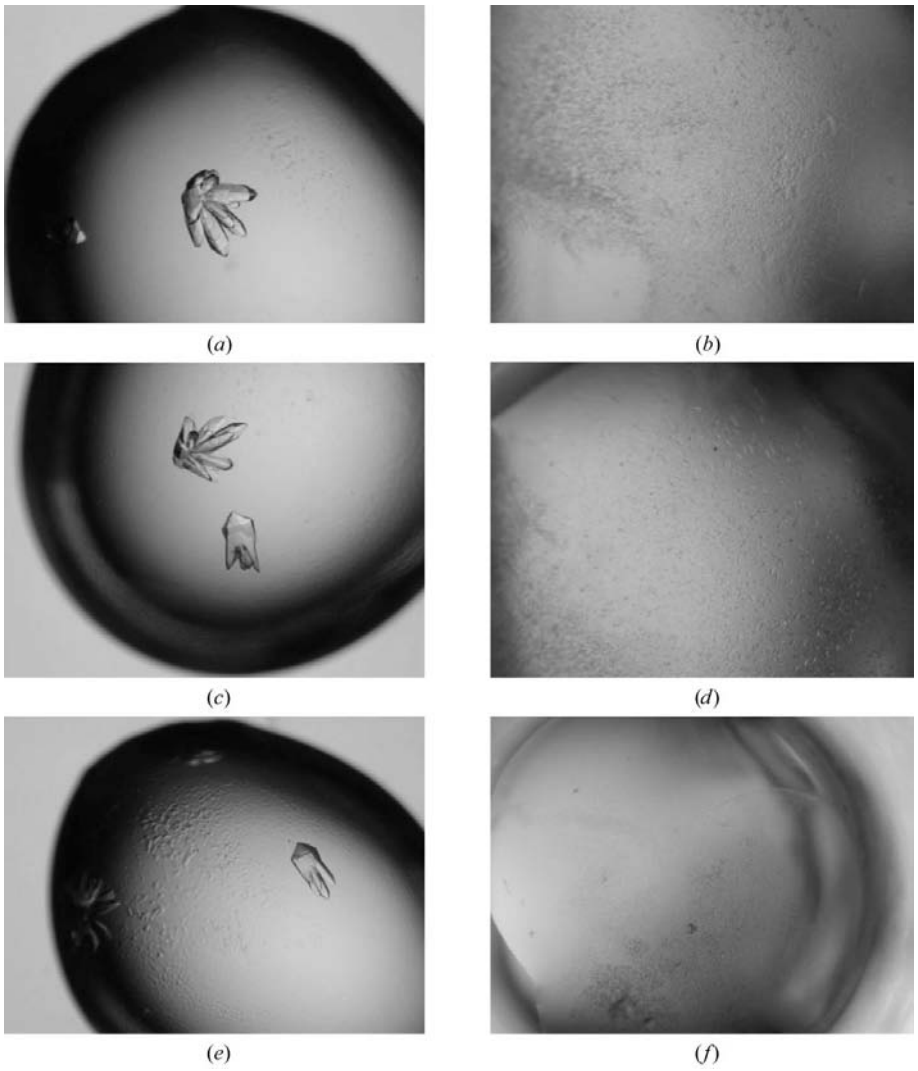

Figure 1
Large crystals obtained in sparse-matrix crystallization trials using optimized protein solvent (*a*, *c* and *e*) are shown paired with the results obtained using a standard solvent (*b*, *d* and *f*) under the same conditions. The best result obtained from the crystal trials in the standard solvent is shown in (*f*).

and a summary of the statistics is presented in Table 2. The crystals belong to the hexagonal space group $P6_1$, with systematic absences along the 00*l* axis consistent with a sixfold screw. Calculation of the Matthews coefficient (V_M = 2.51 Å^3 Da^{-1}; MW = 40.1 kDa) indicates the presence of one molecule per asymmetric unit (Matthews, 1968). The initial phasing was solved by molecular replacement using *AMoRe* (Navaza & Saludjian, 1997) with the FEN-1 structure from *P. furiosus* FEN-1 (PDB code 1b43) as the search model. The space group was confirmed by comparison of the molecular-replacement results using space groups $P6$ (CC = 14.0%, R factor = 54.2%), $P6_1$ (CC = 41.8%, R factor = 48.6%) and $P6_5$ (CC = 23.3%, R factor = 54.4%).

Our interest is in the analysis of the active site of the enzyme. However, metal-soaking experiments on this crystal form have been completely unsuccessful. Attempts to soak crystals in $MgCl_2$ (10, 25 and 200 m*M*) and crystals grown in the presence of $MgCl_2$ (200 m*M*) have not produced evidence of binding. The formation of magnesium citrate could interfere with chelation of the divalent cations by the protein (Pearce, 1980). However, crystals grown with excess magnesium (200 m*M* $MgCl_2$) still do not display metal bound in the active sites. At this point, we must assume that a conformational change associated with metal binding inhibits incorporation into the crystal lattice. A stabilization of the metal-free conformation by the crystal lattice and a relatively weak binding constant in the absence of the DNA substrate could account for the exclusion of metals. We are planning to perform crystallization trials of a citrate-free magnesium-optimized solvent (50 m*M* Tris–HCl pH 7.5, 50 m*M* NH_4Cl, 10 m*M* $MgCl_2$) in an effort to find a metal-bound crystal form.

We have observed a significant correlation between maximizing the solubility of a protein with positive results in crystal screens. By enhancing the solubility of a protein, we have observed an increase in the likelihood that a protein will crystallize.

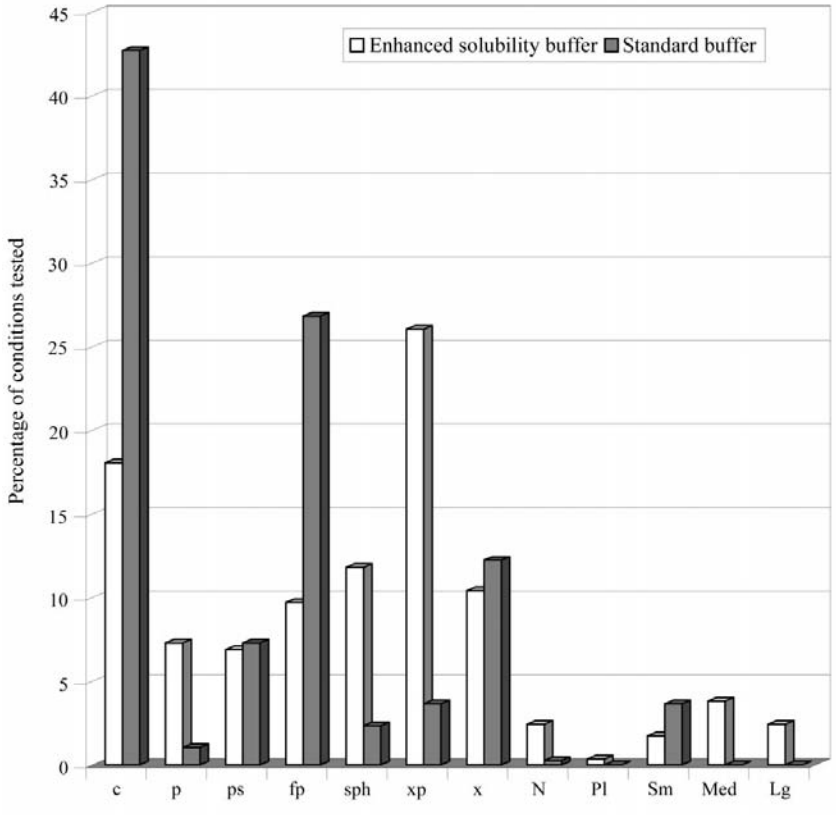

Figure 2
Crystallization results of sparse-matrix crystallization trials were documented for one protein in two different solutions: one in an enhanced buffer formulated based on the solubility screen (white bars) and a second in a standard chromatography buffer (grey bars). Columns report the percentage of the results of the 768 conditions tested for each protein as: c, clear; p, precipitate; ps, phase separation; fp, flocculent precipitate; sph, spherulites; xp, crystalline precipitates; x, microcrystals; N, needle crystals; Pl, plate crystals; Sm, small three-dimensional crystals (0.01–0.05 mm); Med, medium three-dimensional crystals (0.05–0.2 mm); Lg, large three-dimensional crystals (>0.2 mm).

Table 2
Data-collection summary.

Values in parentheses are for the highest resolution shell (1.45–1.40 Å).

Resolution (Å)	45.0–1.4
Wavelength (Å)	0.9 (APS BioCARS-14BMC)
Observed reflections	925349
Unique reflections	74361
Completeness (%)	93.9 (79.7)
Average $I/\sigma(I)$	43.68 (4.46)
R_{merge}† (%)	5.3 (22.3)
Redundancy	12.4 (3.4)
Space group	$P6_1$
Unit-cell parameters (Å, °)	$a = b = 92.8$, $c = 80.9$, $\alpha = \beta = 90$, $\gamma = 120$

† $R_{merge} = \sum |I - \langle I \rangle| / \sum I$.

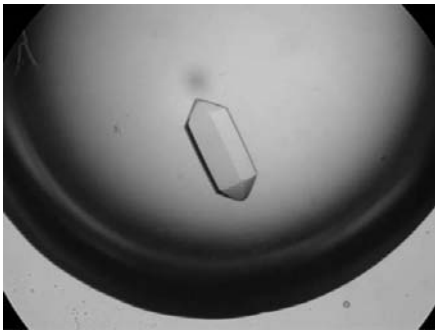

Figure 3
A diffraction-quality crystal of *Ape* FEN-1 grown using hanging-drop vapour diffusion at 296 K against a well solution of 8%(*w*/*v*) PEG 4000, 100 m*M* MES pH 5.6 and 200 m*M* sodium formate has crystal dimensions of 0.6 × 0.2 × 0.2 mm.

Each protein has unique requirements for solubility and crystallization. In the proper salt and buffer, non-specific aggregation can be limited, enhancing the probability of ordered nucleation and subsequent crystal growth. For our protein, we were able to determine based on results from our solubility screen that the presence of citrate is essential for the growth of large diffraction-quality crystals. The general applicability of this approach is not yet proven. Obviously, many proteins would not recover from deionizing precipitation. We have tested the applicability of the solubility screen to standard proteins with positive results along with alternative methods for solubility determination (manuscript in preparation). We anticipate this method will have general applicability.

We wish to thank the staff at APS BioCARS for their assistance in data collection. This work was supported by Third Wave Technologies, Madison, WI, USA and the University of Toledo, Toledo, OH, USA. SJT was supported by the Ohio Affiliate of the American Heart Association predoctoral fellowship No. AHA0315174B.

References

Ceska, T. A., Sayers, J. R., Stier, G. & Suck, D. (1996). *Nature (London)*, **382**, 90–93.

Garforth, S. J., Ceska, T. A., Suck, D. & Sayers, J. R. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 38–43.

Good, N. E., Winget, G. D., Winter, W., Connolly, T. N., Izawa, S. & Singh, R. M. M. (1966). *Biochemistry*, **5**, 467–477.

Harrington, J. J. & Lieber, M. R. (1994*a*). *Genes Dev.* **8**, 1344–1355.

Harrington, J. J. & Lieber, M. R. (1994*b*). *EMBO J.* **13**, 1235–1246.

Hosfield, D. J., Mol, C. D., Shen, B. & Tainer, J. A. (1998). *Cell*, **95**, 135–146.

Hwang, K. Y., Baek, K., Kim, H. Y. & Cho, Y. (1998). *Nature Struct. Biol.* **5**, 707–713.

Kim, Y., Eom, S. H., Wang, J., Lee, D. S., Suh, S. W. & Steitz, T. A. (1995). *Nature (London)*, **376**, 612–616.

McPherson, A. (1999). *Crystallization of Biological Macromolecules.* New York: Cold Spring Harbor Laboratory Press.

Matsui, E., Musti, K. V., Abe, J., Yamasaki, K., Matsui, I. & Harata, K. (2002). *J. Biol. Chem.* **277**, 37840–37847.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Mueser, T. C., Nossal, N. G. & Hyde, C. C. (1996). *Cell*, **85**, 1101–1112.

Mueser, T. C., Rogers, P. H. & Arnone, A. (2000). *Biochemistry*, **39**, 15353–15364.

Navaza, J. & Saludjian, P. (1997). *Methods Enzymol.* **276**, 581–594.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pearce, K. N. (1980). *Aust. J. Chem.* **33**, 1511–1517.

Sako, Y., Nomura, N., Uchida, A., Ishida, Y., Morii, H., Koga, Y., Hoaki, T. & Maruyama, T. (1996). *Int. J. Syst. Bacteriol.* **46**, 1070–1077.