JOURNAL OF APPLIED CRYSTALLOGRAPHY

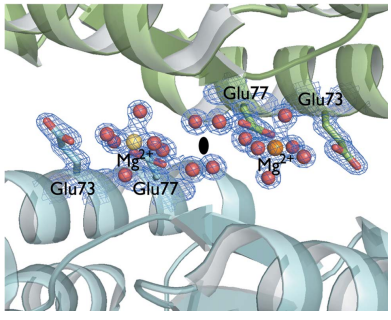
ISSN 1600-5767

Received 27 June 2017
Accepted 1 August 2017

Edited by F. Meilleur, Oak Ridge National Laboratory, USA, and North Carolina State University, USA

Keywords: crystallization screens; crystal packing; ions and buffers; crystal growth; structural biology.

Supporting information: this article has supporting information at journals.iucr.org/j



OPEN ACCESS

Berkeley Screen: a set of 96 solutions for general macromolecular crystallization

Jose H. Pereira,[a,b] Ryan P. McAndrew,[a,b] Giovani P. Tomaleri[b] and Paul D. Adams[a,b,c]*

[a]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA, [b]Joint BioEnergy Institute, Emeryville, CA 94608, USA, and [c]Department of Bioengineering, University of California, Berkeley, CA 94720, USA. *Correspondence e-mail: pdadams@lbl.gov

Using statistical analysis of the Biological Macromolecular Crystallization Database, combined with previous knowledge about crystallization reagents, a crystallization screen called the Berkeley Screen has been created. Correlating crystallization conditions and high-resolution protein structures, it is possible to better understand the influence that a particular solution has on protein crystal formation. Ions and small molecules such as buffers and precipitants used in crystallization experiments were identified in electron density maps, highlighting the role of these chemicals in protein crystal packing. The Berkeley Screen has been extensively used to crystallize target proteins from the Joint BioEnergy Institute and the Collaborative Crystallography program at the Berkeley Center for Structural Biology, contributing to several Protein Data Bank entries and related publications. The Berkeley Screen provides the crystallographic community with an efficient set of solutions for general macromolecular crystallization trials, offering a valuable alternative to the existing commercially available screens.

1. Introduction

X-ray crystallography is currently the most successful technique used to solve macromolecular structures, contributing several thousand new entries to the Protein Data Bank (PDB) every year (Berman *et al.*, 2000). Recently the PDB reached 125 000 structure entries, with approximately 90% of these entries coming from X-ray crystallography, 10% from nuclear magnetic resonance (NMR) and 1% from electron microscopy. While the number of new entries using X-ray crystallography increases every year, the number of new NMR entries has been decreasing since 2007. In contrast, new entries using cryo-electron microscopy (cryo-EM) are increasing rapidly as a result of recent technical advances enabling near atomic resolution data collection (Zhou, 2008; Baker *et al.*, 2010). In 2016, cryo-EM was responsible for 410 new entries in the PDB, almost equal to the 453 new entries from NMR, suggesting that cryo-EM will soon overtake the latter technique for solving macromolecular structures (PDB, 2017).

There are three important stages in going from the purified sample of a macromolecule(s) to a description of the structure in a scientific paper: observation of the crystal in the microscope, obtaining the first X-ray diffraction pattern, and achieving an electron density map of sufficient quality that the structure can be solved. The crystal is the critical starting point for X-ray data collection, and consequently, its properties are correlated with the quality of the electron density maps and the level of detail that can be extracted for a macromolecular

structure. Currently, a complete X-ray data set often can be obtained using synchrotron storage ring sources equipped with a pixel array detector such as the Pilatus in a matter of minutes. Crystallographic packages such as *Phenix* (Adams *et al.*, 2010; Afonine *et al.*, 2012) and the *CCP4 Suite* (Winn *et al.*, 2011) provide the structural biology community with very powerful tools for solving and completing crystal structures. However, the crystallization step still is less developed compared to X-ray data collection and structure solution.

Despite the fact that robots make it possible to efficiently screen against hundreds of crystallization solutions, crystallization is still a trial and error process, and many attempts still fail. Moreover, even after a crystal with good shape and size has been obtained, there is no guarantee that the crystal will diffract well. Therefore, understanding the agents and factors involved in the crystallization process could be crucial to obtaining a macromolecular structure. In this paper we describe a set of 96 solutions to be used for general macromolecular crystallization trials. Furthermore, we have correlated the solvent content of a set of high-resolution crystal structures to explain crystal growth in one particular solution.

Table 1
The distribution of the pH and the most common salts/precipitants among the BMCD entries and Berkeley Screen solutions.

	BMCD (%)†	Berkeley Screen solutions (%)
Distribution of pH values		
pH range 3–4	2	4
pH range 4–5	10	11
pH range 5–6	19	19
pH range 6–7	28	24
pH range 7–8	29	27
pH range 8–9	10	13
pH range 9–10	2	2
Most common salts and precipitants		
PEG (all molecular weights)	43	73
Ammonium sulfate	26	12
Sodium chloride	14	10
Magnesium chloride	9	16
2-Methyl-2,4-pentanediol	9	9
2-Propanol	3	7
Lithium sulfate	3	6

† A total of 14 372 crystal entries were used to generate the data (Tung & Gallagher, 2009).

2. Results and discussion

2.1. Designing the Berkeley Screen formulation

The development of the new protein screen started with the observation that a significant number of macromolecular structures contain bound cation/anion or small molecule ligands from the crystallization solution. In addition, the low success rate of some commercially available protein crystallization screens motived the creation of a new screen in order to optimize cost and time.

In some cases, a metal ion is important for protein activity, such as the Mg^{2+} ion that is found coordinating the phosphate groups in ATP binding proteins (Pereira, Ralston *et al.*, 2010, 2012). Although these catalytic site metals are unlikely to be involved in crystal packing interactions, they are often important in stabilizing a region of the macromolecule and therefore have an impact on the crystallization process. In other cases, the metal ion or small molecules, such as buffer components, play an important role in creating crystal contacts between the macromolecules in the lattice. Using the ions that are commonly bound to biological macromolecular structures (Mg^{2+}, Ca^{2+}, K^{+}, Na^{+}, Cl^{1-}, SO_4^{2-}, PO_4^{2-}) combined with the most successful crystallization reagents (Tung & Gallagher, 2009) and a revised set of crystallization reagents (McPherson & Gavira, 2014; Bergfors, 2008; McPherson, 2001, 1976; Ames *et al.*, 1998; Trakhanov & Quiocho, 1995; Jancarik & Kim, 1991), we created a sparse-matrix screen of 96 crystallization solutions (supplementary file S1). The occurrences of the most common salts, precipitants and pH ranges in the Biological Macromolecular Crystallization Database (BMCD; Tung & Gallagher, 2009) and Berkeley Screen solutions are described in Table 1.

In 1968, an alcohol oxidase was the first protein crystallized using a polyethylene glycol (PEG) (Janssen & Ruelius, 1968; Kirkwood *et al.*, 2015). PEGs became extensively used for crystallization after the McPherson study using PEGs of various molecular weights was published in 1976 (McPherson, 1976). In addition, a study published by Newman and collaborators using a systematic pH, anion and cation testing (PACT) with the precipitants PEG 1500, PEG 3350 and PEG 6000 contributed to wider use of the PEGs for an initial rational protein crystallization screening (Newman *et al.*, 2005). PEGs compete with the protein for water interaction, forcing the protein out of solution and consequently promoting crystal nucleation and crystal growth. Currently, PEGs (including all molecular weights) are the most successful precipitants used in crystallization, responsible for about half of all entries in the BMCD (Tung & Gallagher, 2009). PEGs therefore have the highest occurrence in the Berkeley Screen as compared to any other reagent. In the formulation are 70 solutions (73%) containing PEGs, more specifically 36 solutions with PEG 3350, nine solutions with PEG 400, nine solutions with PEG monomethyl ether (MME) 2000, five solutions with PEG MME 5000, four solutions with PEG MME 550, three solutions with PEG 4000, two solutions with PEG 1500, one solution with PEG 8000 and one solution with PEG 10 000. The significant difference between the occurrence of PEG in the BMCD (∼50%) and in the Berkeley Screen (∼70%) is due to the use of PEGs in combination with other precipitants. For example, condition D8 (solution 45 of the Berkeley Screen) has two components as precipitants: 20% PEG 3350 and 5% 2-propanol. As expected from the high occurrence among the conditions present in the Berkeley Screen formulation, the solutions containing PEG as a precipitant are the most successful for the target proteins tested from the Joint BioEnergy Institute and the Collaborative Crystallography program at the Berkeley Center for Structural Biology (Pereira *et al.*, 2017, 2016; Fallas *et al.*, 2016; Mills *et al.*, 2016; Helmich *et al.*, 2016; Javidpour *et al.*, 2014) (Fig. 1).

A comparison of 12 different salts used in crystallization solutions, including ammonium sulfate, sodium malonate, lithium sulfate, ammonium phosphate, sodium phosphate, sodium citrate, sodium acetate, sodium tartrate, magnesium sulfate, sodium chloride, ammonium formate and lithium chloride, showed that sodium malonate was the most successful salt for the 31 proteins and virus targets tested (McPherson, 2001). Small organic acids such as citrate, tartrate, acetate, malonate and formate are valuable precipitants for crystal growth (McPherson, 2001). Following this study by McPherson, we created 11 solutions specifically containing sodium malonate in the Berkeley Screen formulation, varying the concentration from 0.1 to 2.2 *M* and the pH range from 5.0 to 8.0. Moreover, 36 solutions of the screen include small organic acids (supplementary file S1).

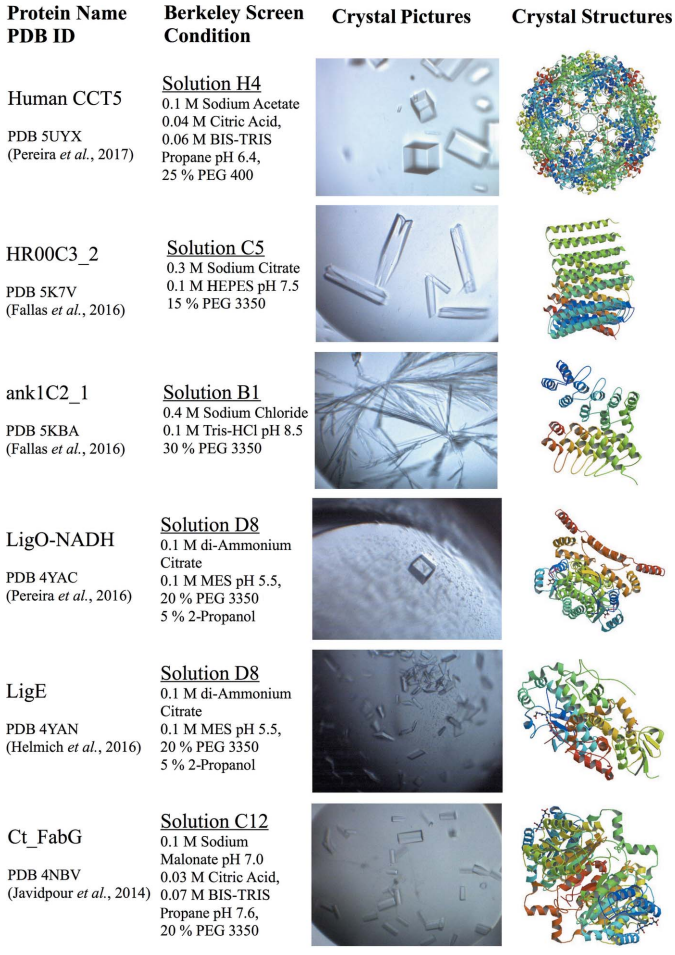


Figure 1
Examples of protein crystals successfully grown from Berkeley Screen solutions containing PEGs, which have the highest occurrence in the Berkeley Screen formulation. The human TRiC subunit CCT5 (PDB code 5uyx) crystallized in PEG 400 as precipitant, whereas the computational design proteins HR00C3_2 (PDB code 5k7v) and ank1C2_1 (PDB code 5kba), the lignin degradation enzymes LigO (PDB code 4yac) and LigE (PDB code 4yan), and a reductase involved in fatty acid biosynthesis Ct_FabG (PDB code 4nbv) crystallized in PEG 3350.

2.2. Understanding the influence of crystallization solutions on protein crystal formation

We selected four protein structures solved recently in our laboratory at high resolution (1.1–1.8 Å) to illustrate how reagents found in the Berkeley Screen influenced crystal formation.

2.2.1. Symmetrical ion coordination (magnesium – Mg^{2+}). Free magnesium (Mg^{2+}) is an essential cation for a broad range of enzymes. In mammalian cells, more than 350 enzymes require or are regulated by Mg^{2+} (Romani, 2013). Usually, Mg^{2+} is bound to the protein *via* negatively charged residues such as aspartate (Asp) or glutamate (Glu) that are involved in the common Mg^{2+} octahedral coordination (Lebbink *et al.*, 2010). However, the Mg^{2+} ion could also bind directly or through water coordination to negatively charged Asp and Glu residues that are exposed to solvent, potentially creating crystal contacts between molecules in the lattice. An example is the crystal formation of the NAD-dependent

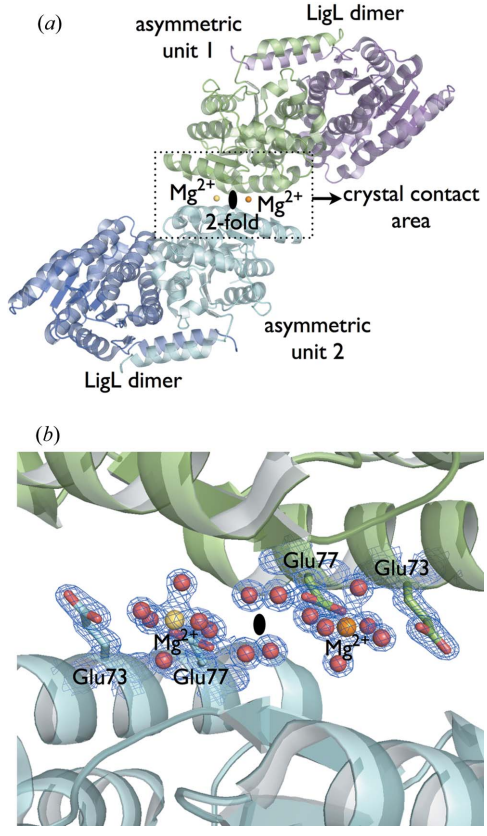


Figure 2
(*a*) The LigL structure (PDB code 4yai; Pereira *et al.*, 2016) was solved at 1.6 Å resolution with two molecules per asymmetric unit, which represents the biological dimer, showing Mg^{2+} was essential for the crystal lattice contacts. A twofold (*C*2) symmetry axis is shown with a solid oval symbol. (*b*) A $2mF_o$–DF_c electron density map for the LigL structure contoured at 1.5σ is shown in blue around the Mg^{2+} ions, the water molecules, and the side chains of Glu73 and Glu77. The octahedral coordination of Mg^{2+} is water mediated by the negative surface residues Glu73 and Glu77, generating the major crystal contact between symmetry-related molecules. Water molecules are shown as red spheres.

dehydrogenase LigL from *Sphingobium sp.* strain (Pereira *et al.*, 2016). LigL was crystallized from 0.2 *M* magnesium chloride, 0.1 *M* HEPES pH 7.5 and 25% PEG 3350. The LigL structure was solved at 1.6 Å resolution, revealing that Mg^{2+} was essential for the crystal lattice contacts. The Mg^{2+} was water coordinated and held in place by surface negative residues Glu73 and Glu77, generating the major crystal contacts between symmetry mates (Fig. 2). Statistical analysis of the BMCD showed that 10% of all the entries had magnesium chloride present in the crystallization solution. Therefore, the Berkeley Screen was designed with 16 solutions (16%) containing magnesium chloride from 0.005 to 0.6 *M* concentration.

2.2.2. The classic, ammonium sulfate – hydrogen-bond and salt bridge interactions. Historically, the most successful crystallization precipitant was ammonium sulfate. More recently, the most common precipitant has become PEG 3350 (Kirkwood *et al.*, 2015). However, analysis of the BMCD showed that ammonium sulfate is still a good precipitant candidate for crystallization trials, corresponding to 26% of all entries (Tung & Gallagher, 2009). Exploring the high-resolution crystal structure of a glutathione-dependent lyase LigG solved at 1.1 Å resolution (Pereira *et al.*, 2016), we identified a sulfate ion (SO_4^{2-}) participating in crystal contacts. The LigG was crystallized from 0.1 *M* bis-tris propane pH 7.0 and 1.5 *M* ammonium sulfate. The SO_4^{2-} ion interacts with Arg43 *via* the N main chain and NE and NH_2 atoms from the side chain. The positive residue Arg43 makes a salt bridge interaction with the LigG symmetry mate residue Glu208. The SO_4^{2-} ion makes additional contacts with the main chain N atoms of Glu208 and Lys207. The NZ atom of Lys207 makes a hydrogen bond with the carbonyl group of Arg43 (Fig. 3). Therefore, SO_4^{2-} plays a central role in stabilizing residue Arg43, creating a contact for the negatively charged residue Glu208 from a symmetry-related molecule. How the ammonium cation NH_4^+ participates in crystallization is challenging to determine owing to the difficulty of distinguishing free NH_4^+ and water molecules in the electron density map. However, it is known that NH_4^+ is the most robust cation for precipitation of hen egg white proteins from the Hofmeister series (Hofmeister, 1890; Bergfors, 2008). The NH_4^+ ions compete with the protein for water molecule interactions, forcing the protein out of solution: the salting-out effect. The capability of NH_4^+ to compete for water interactions, and the ability of SO_4^{2-} to make hydrogen bonds or salt bridges between symmetry-related molecules, make ammonium sulfate salt a reagent that must be presented in an initial protein crystallization screen. The Berkeley Screen has 12 conditions containing ammonium sulfate.

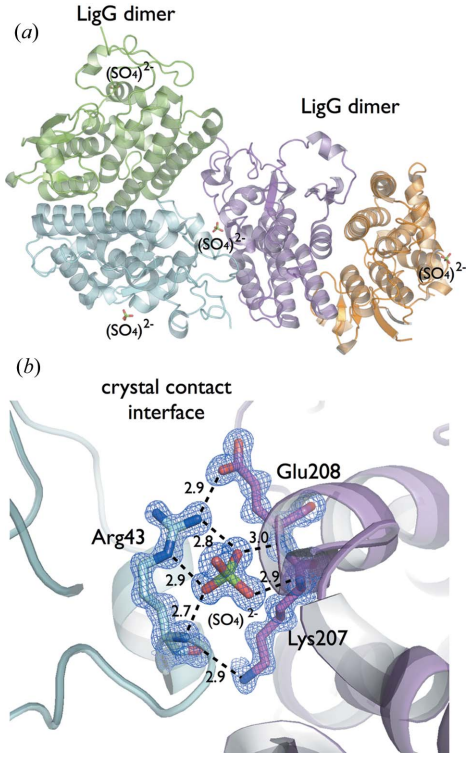


Figure 3
(*a*) Crystal packing in the dimeric crystal structure of a glutathione-dependent lyase LigG solved at 1.1 Å resolution (PDB code 4yap) (Pereira *et al.*, 2016), showing the SO_4^{2-} anion bound to the surface region participating in the crystal lattice. (*b*) Zoom in showing the LigG crystal contact with the hydrogen bonds and salt bridges observed between SO_4^{2-} and the symmetry-related copies of LigG. The SO_4^{2-} contacts Arg43 *via* the main chain N atom and NE and NH_2 from the side chain. The positive residue Arg43 makes a salt bridge interaction with the symmetry-related residue Glu208. The SO_4^{2-} ion makes additional contacts with the main chain N atoms of Glu208 and Lys207. The NZ atom of Lys207 makes a hydrogen bond with the carbonyl group of Arg43. A $2mF_o$–DF_c electron density map contoured at 1.5σ is shown in blue. Contacts are shown as broken lines and distances in ångströms.

3. Intermolecular interactions induced by cadmium – Cd^{2+}

Cadmium (Cd^{2+}) ions can shield charges located at protein surface regions, creating favorable intermolecular interactions (Trakhanov & Quiocho, 1995; Ames *et al.*, 1998). Analysis of a glycoside hydrolase Cel5A structure (Pereira, Chen *et al.*, 2010) revealed that cadmium was essential in forming 'bridges' between protein molecules involving the negatively charged residue Glu99 present at the surface of the protein (Fig. 4*a*). A similar role of cadmium in crystal formation was observed for histidine-binding protein (Yao *et al.*, 1994), dipeptide-binding protein (Nickitenko *et al.*, 1995) and leucine-specific protein (Trakhanov & Quiocho, 1995).

The Cel5A initial protein crystallization solution contained 20 m*M* of three different salts: nickel(II) chloride ($NiCl_2$), magnesium chloride ($MgCl_2$) and cadmium chloride ($CdCl_2$). When Ni^{2+} and Mg^{2+} were removed from the solution, the Cel5A protein sample still crystallized; however, when Cd^{2+} was removed, Cel5A no longer crystallized. Experiments using different concentrations of cadmium chloride were performed, and the protein could be crystallized in the presence of only 0.25 m*M* $CdCl_2$ (80 times less than the original concentration and a 1:1 molar ratio between Cel5A and $CdCl_2$). The small amount of $CdCl_2$ (0.25 m*M*) also did not drastically change the ionic strength of the crystallization solution, but it was

essential for success of crystallization experiments, indicating that Cd^{2+} was involved in the crystal packing. Moreover, the combination of Ni^{1+} and Cd^{2+} produced Cel5A crystals that belonged to the *P*1 space group. Using only Cd^{2+} produced Cel5A crystals that belonged to the $P2_1$ space group with a completely different morphology (Fig. 4*b*). The presence or absence of $MgCl_2$ did not make any difference to Cel5A crystal formation. The Cel5A crystals demonstrate that manipulating the types of metals present in a crystallization solution could affect the way a protein packs in a crystal lattice. Different kinds of crystal packing (space groups) from a single target are always welcome from the crystallization experiment since they can have a strong influence on many of the factors that ultimately result in how well a crystal diffracts.

Cadmium has been identified in 864 PDB entries (∼0.7% of the 114 424 total entries), although it is possible that there are other structures where the ion has not been correctly identified. In light of the potentially important role in crystal formation, the Berkeley Screen has two entries (∼2%) containing $CdCl_2$ to cover the possibility of some protein targets requiring Cd^{2+} for specific crystal contacts such as those seen in Cel5A crystals (Pereira *et al.*, 2010).

4. Small molecules function as bridges between symmetry mates (bis-tris propane buffer)

The combination of bis-tris propane and citric acid is often used as a buffer in crystallization solutions. The bis-tris propane molecule can form crystal bridges between symmetry-related molecules, while citric acid, which is used to adjust the pH, can also compete for water solvent, forcing the protein out of the solution to form crystals. The structure of FabG from *Cupriavidus taiwanensis* (Javidpour *et al.*, 2014), solved at 1.6 Å resolution, is a good example of bis-tris propane in crystal packing. The FabG protein was crystallized from 0.1 *M* sodium malonate pH 7.0, 0.03 *M* citric acid, 0.07 *M* bis-tris propane pH 7.6 and 20% PEG 3350. Bis-tris propane molecules were bound between the two monomers on each side of the biological unit tetramer in the crystal (Fig. 5*a*). The bis-tris propane stabilizes the charged residues Asp111 and Lys115 for one monomer and Asp107 for the other monomer. Finally, the crystal contact coordinates bis-tris propane and the symmetry mate chain using the Gln193 residue (Fig. 5*b*). Adding bis-tris propane as one of the regular buffers used in the Berkeley Screen formulation exposes the crystallization targets to this small molecule involved in crystal lattice

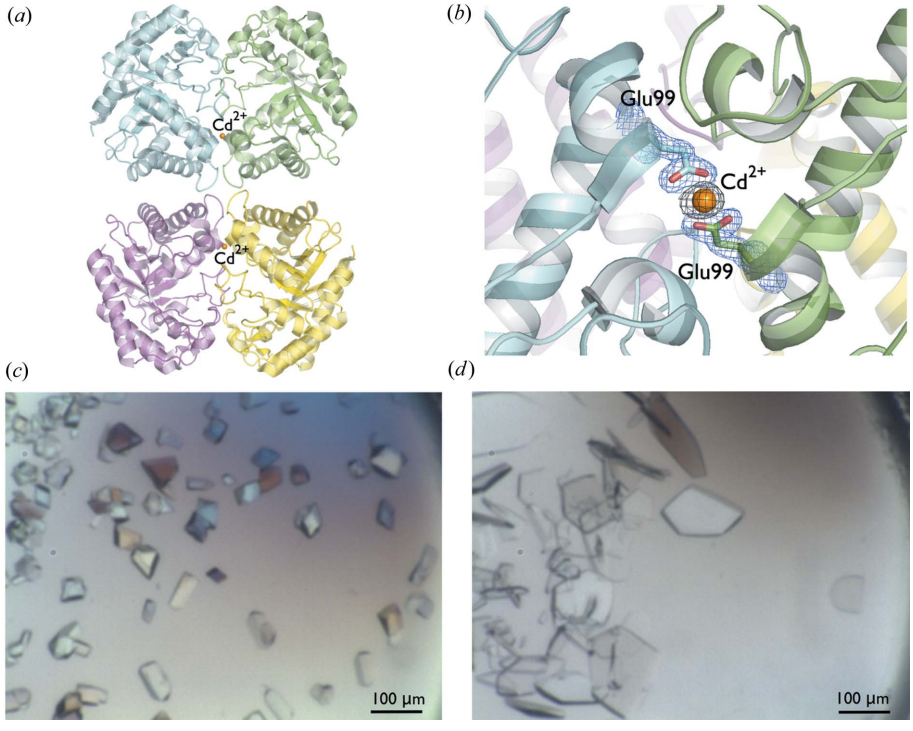


Figure 4
(*a*) The asymmetric unit contents showing four molecules of Cel5A (PDB code 3mmw; Pereira *et al.*, 2010). (*b*) The Cel5A structure showed that cadmium was essential to form 'bridges' across the interface of protein molecules through coordination by the negative charge residue Glu99 present at the surface of the protein. A $2mF_o$–DF_c electron density map contoured at 1.5σ is shown in blue around the Glu99 residues and an mF_o–DF_c electron density map contoured at 10σ when the divalent metal was omitted from the model is shown in black around Cd^{2+}. (*c*) The presence of Ni^{2+} and Cd^{2+} produced Cel5A crystals belonging to the *P*1 space group that were solved at 2.2 Å resolution. The crystallization condition of these crystals was 0.02 *M* nickel(II) chloride, 0.02 *M* magnesium chloride, 0.02 *M* cadmium chloride, 0.1 *M* sodium acetate trihydrate pH 4.5 and 16% polyethylene glycol monomethyl ether 2000. (*d*) When only Cd^{2+} was included in the crystallization solution, Cel5A crystals belonging to the $P2_1$ space group were produced with a completely different crystal morphology, which were solved at 1.8 Å resolution. The crystallization condition for the Cel5A $P2_1$ crystals was 0.02 *M* cadmium chloride, 0.1 *M* sodium acetate trihydrate pH 4.5 and 16% polyethylene glycol monomethyl ether 2000 (Pereira *et al.*, 2010).

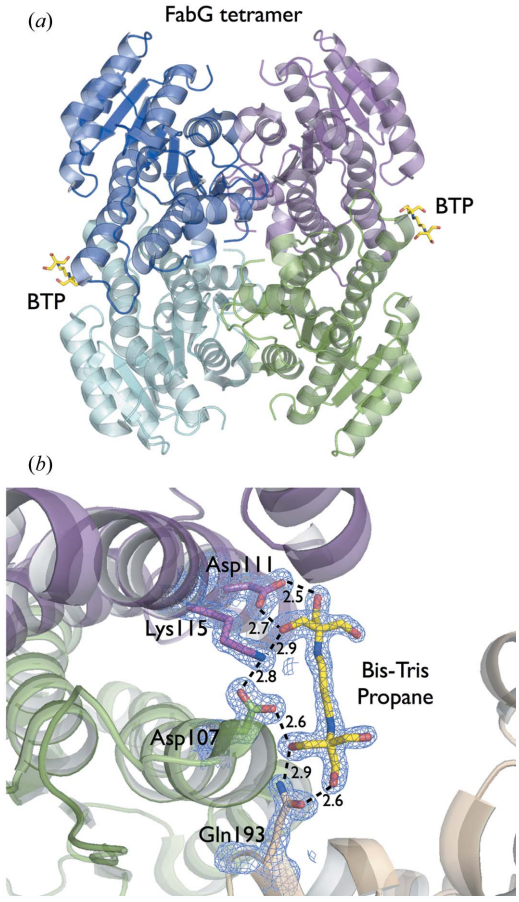


Figure 5
(*a*) The bis-tris propane molecules (BTP) bound between the two monomers on each side of the biological tetramer in the structure of FabG from *Cupriavidus taiwanensis* solved at 1.6 Å resolution (PDB code 4nbv; Javidpour *et al.*, 2014). (*b*) The crystal contact interface with a $2mF_o$–DF_c electron density map contoured at 1.2σ is shown in blue around Asp107, Asp111, Lys115, Gln193 and the bis-tris propane molecule. Bis-tris propane stabilizes the charged residues Asp111 and Lys115 in one monomer and Asp107 in the other monomer. The crystal contact is coordinated by bis-tris propane and Gln193 from the symmetry-related copy. Contacts are shown as broken lines and distances in ångströms.

formation during initial crystallization trials. The Berkeley Screen has seven solutions containing bis-tris propane as a buffer.

5. Conclusion

Initially, protein targets selected for X-ray crystallography studies depended on the availability of large quantities from natural sources. Advances in molecular biology methods have provided more freedom to macromolecular crystallographers in target selection and allowed great strides in elucidating fundamental biological questions using structural data. However, the amount of sample can often be a limiting factor for targets in modern protein X-ray crystallography; consequently the number of solutions chosen to be tested during crystallization must be done critically. Understanding the factors involved in crystallization is an essential step to selecting a good set of solutions, and the Berkeley Screen, despite limited release, has already been a valuable alternative to the commercially available screens, providing crystals for several publications in the past few years (Pereira *et al.*, 2017, 2016, 2014; Marcos *et al.*, 2017; Boyken *et al.*, 2016; Fallas *et al.*, 2016; Mills *et al.*, 2016; Eudes *et al.*, 2016; Helmich *et al.*, 2016; Javidpour *et al.*, 2014).

6. Additional Information

An Excel file containing the screen formulation, individual chemical information such as the catalog number from Hampton Research for all the stock solutions, and the calculated volumes for each stock solution necessary to make 10 ml of each Berkeley Screen condition is provided as supporting information (supplementary file S2).

The group definition for the Berkeley Screen solution has been created by Art Robbins Instruments to be used on the Scorpion Screen Builder Robot. Requests for information about the Berkeley Screen group definition can be made directly to Art Robbins (http://www.artrobbins.com/contact-support).

Acknowledgements

We are grateful to all staff at the Berkeley Center for Structural Biology at the Advanced Light Source of Lawrence Berkeley National Laboratory, in particularly Banumathi Sankaran who leads the Collaborative Crystallography Program, and Peter Zwart who was involved in data collection for some of the protein targets tested.

Funding information

Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the US Department of Energy (contract No. DE-AC02-05CH11231). The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the US Department of Energy (contract No. DE-AC02-05CH11231).

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* D**68**, 352–367.
Ames, G. F., Kreimer, D. I., Trakhanov, S., Parkin, S. & Rupp, B. (1998). *Protein Sci.* **7**, 600–604.
Baker, M. L., Zhang, J., Ludtke, S. J. & Chiu, W. (2010). *Nat. Protoc.* **5**, 1697–1708.
Bergfors, T. M. (2008). *Protein Crystallization*, 2nd ed. La Jolla: International University Line.
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.
Boyken, S. E., Chen, Z., Groves, B., Langan, R. A., Oberdorfer, G., Ford, A., Gilmore, J. M., Xu, C., DiMaio, F., Pereira, J. H.,

Sankaran, B., Seelig, G., Zwart, P. H. & Baker, D. (2016). *Science*, **352**, 680–687.
Eudes, A., Pereira, J. H., Yogiswara, S., Wang, G., Teixeira Benites, V., Baidoo, E. E., Lee, T. S., Adams, P. D., Keasling, J. D. & Loqué, D. (2016). *Plant Cell Physiol.* **57**, 568–579.
Fallas, J. A., Ueda, G., Sheffler, W., Nguyen, V., McNamara, D. E., Sankaran, B., Pereira, J. H., Parmeggiani, F., Brunette, T. J., Cascio, D., Yeates, T. R., Zwart, P. & Baker, D. (2016). *Nat. Chem.* **9**, 353–360.
Helmich, K. E., Pereira, J. H., Gall, D. L., Heins, R. A., McAndrew, R. P., Bingman, C., Deng, K., Holland, K. C., Noguera, D. R., Simmons, B. A., Sale, K. L., Ralph, J., Donohue, T. J., Adams, P. D. & Phillips, G. N. Jr (2016). *J. Biol. Chem.* **291**, 5234–5246.
Hofmeister, T. (1890). *Z. Physiol. Chem.* **14**, 165.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Janssen, F. W. & Ruelius, H. W. (1968). *Biochim. Biophys. Acta*, **151**, 330–342.
Javidpour, P., Pereira, J. H., Goh, E. B., McAndrew, R. P., Ma, S. M., Friedland, G. D., Keasling, J. D., Chhabra, S. R., Adams, P. D. & Beller, H. R. (2014). *Appl. Environ. Microbiol.* **80**, 497–505.
Kirkwood, J., Hargreaves, D., O'Keefe, S. & Wilson, J. (2015). *Acta Cryst.* F**71**, 1228–1234.
Lebbink, J. H. G., Fish, A., Reumer, A., Natrajan, G., Winterwerp, H. H. K. & Sixma, T. K. (2010). *J. Biol. Chem.* **285**, 13131–13141.
Marcos, E. *et al.* (2017). *Science*, **355**, 201–206.
McPherson, A. Jr (1976). *J. Biol. Chem.* **251**, 6300–6303.
McPherson, A. (2001). *Protein Sci.* **10**, 418–422.
McPherson, A. & Gavira, J. A. (2014). *Acta Cryst.* F**70**, 2–20.
Mills, J. H., Sheffler, W., Ener, M. E., Almhjell, P. J., Oberdorfer, G., Pereira, J. H., Parmeggiani, F., Sankaran, B., Zwart, P. H. & Baker, D. (2016). *Proc. Natl Acad. Sci. USA*, **113**, 15012–15017.
Newman, J., Egan, D., Walter, T. S., Meged, R., Berry, I., Ben Jelloul, M., Sussman, J. L., Stuart, D. I. & Perrakis, A. (2005). *Acta Cryst.* D**61**, 1426–1431.
Nickitenko, A. V., Trakhanov, S. & Quiocho, F. A. (1995). *Biochemistry*, **34**, 16585–16595.
PDB (2017). *PDB Statistics*, http://www.rcsb.org/pdb/static.do?p=general_information/pdb_statistics/.
Pereira, J. H., Chen, Z., McAndrew, R. P., Sapra, R., Chhabra, S. R., Sale, K. L., Simmons, B. A. & Adams, P. D. (2010). *J. Struct. Biol.* **172**, 372–379.
Pereira, J. H., Heins, R. A., Gall, D. L., McAndrew, R. P., Deng, K., Holland, K. C., Donohue, T. J., Noguera, D. R., Simmons, B. A., Sale, K. L., Ralph, J. & Adams, P. D. (2016). *J. Biol. Chem.* **291**, 10228–10238.
Pereira, J. H., McAndrew, R. P., Sergeeva, O. A., Ralston, C. Y., King, J. A. & Adams, P. D. (2017). *Sci. Rep.* **7**, 3673.
Pereira, J. H., Petchprayoon, C., Hoepker, A. C., Moriarty, N. W., Fink, S. J., Cecere, G., Paterson, I., Adams, P. D. & Marriott, G. (2014). *ChemMedChem*, **9**, 2286–2293.
Pereira, J. H., Ralston, C. Y., Douglas, N. R., Kumar, R., McAndrew, R. P., Knee, K. M., King, J. A., Frydman, J. & Adams, P. D. (2012). *EMBO J.* **31**, 731–740.
Pereira, J. H., Ralston, C. Y., Douglas, N. R., Meyer, D., Knee, K. M., Goulet, D. R., King, J. A., Frydman, J. & Adams, P. D. (2010). *J. Biol. Chem.* **285**, 27958–27966.
Romani, A. M. P. (2013). *Front. Biosci.* **12**, 308–331.
Trakhanov, S. & Quiocho, F. (1995). *Protein Sci.* **4**, 1914–1919.
Tung, M. & Gallagher, D. T. (2009). *Acta Cryst.* D**65**, 18–23.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yao, N., Trakhanov, S. & Quiocho, F. A. (1994). *Biochemistry*, **33**, 4769–4779.
Zhou, Z. H. (2008). *Microsc. Curr. Opin. Strut. Biol.* **18**, 218–228.