The role of flexibility and molecular shape in the crystallization of proteins by surface mutagenesis

Yancho D. Devedjiev*

Department of Anesthesiology, University of Virginia Medical Center, 1215 Lee Street, PO Box 800634, Charlottesville, VA 22908-0634, USA. *Correspondence e-mail: idd3c@virginia.edu

Received 28 September 2014

Accepted 21 December 2014

Keywords: protein crystallization; static disorder; conformational entropy; surface mutagenesis; molecular shape.

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

Proteins are dynamic systems and interact with their environment. The analysis of crystal contacts in the most accurately determined protein structures ($d < 1.5$ Å) reveals that in contrast to current views, static disorder and high side-chain entropy are common in the crystal contact area. These observations challenge the validity of the theory that presumes that the occurrence of well ordered patches of side chains at the surface is an essential prerequisite for a successful crystallization event. The present paper provides evidence in support of the approach for understanding protein crystallization as a process dependent on multiple factors, each with its relative contribution, rather than a phenomenon driven by a few dominant physicochemical characteristics. The role of the molecular shape as a factor in the crystallization of proteins by surface mutagenesis is discussed.

1. Introduction

Three methods are used to obtain high-precision structural information about biological macromolecules: X-ray crystallography, NMR and electron microscopy. In spite of the advances in the latter two techniques, the first one provides about 90% of the structures deposited in the PDB. High-quality diffracting crystals constitute a prerequisite for a structural investigation. The preparation of crystals of biological macromolecules requires extensive screening and optimization of the resultant crystallization conditions (Ducruix & Giegé, 1992; Bergfors, 1999). Therefore, it is a time-consuming and labor-consuming procedure and forms a bottleneck in structural investigation.

It is thought that water provides a critical contribution to the free energy of the crystallization of proteins (Vekilov, 2003) and may be the driving force for the phase transition from the liquid to the condensed phase. A negative entropic factor resulting from the ordering of protein molecules in a crystal lattice, where they lose their rotational and translational degrees of freedom, is compensated for by the positive contribution in entropy resulting from the migration of water molecules from the hydration shell to the bulk solvent. At ambient temperature the molecules of the unbound solvent explore translational and rotational degrees of freedom, which creates a reservoir of entropy in the bulk solvent and is believed to drive functional motions in the protein molecules (Frauenfelder *et al.*, 2009).

Traditionally, protein crystallization is considered to be a knowledge-based and empirical method (McPherson, 1985). To approach crystallization on a more rational basis, an attempt has been made to account for some surface properties

of the protein molecules in crystallization, for instance for the conformational flexibility of amino-acid side chains, such as glutamate, lysine and others, exploring a concept called surface-entropy reduction (SER; Derewenda, 2004; Derewenda & Vekilov, 2006). SER uses statistical methods for justification by the analysis of crystal contacts in a selection of protein structures downloaded from the PDB (Cieřlik & Derewenda, 2009).

A dramatic increase in the number of available protein structures arising from the Protein Structure Initiative has created a wealth of structural information that has been investigated using automated approaches (Price *et al.*, 2009; Cieřlik & Derewenda, 2009). These searches assume that the existence of well ordered patches on the protein surface is an essential factor for crystallization, and that reduction of the conformational entropy of some surface-exposed amino-acid

side chains increases the propensity for crystallization. In a study based on a selection of 175 protein structures deposited in the PDB, a search for correlation between physical properties of a particular amino-acid side chain and its involvement in crystal contacts has been made (Price *et al.*, 2009).

Recent studies go beyond statistical methods and use an atomistic approach that reveals more complex interactions in the crystal contact area (Fusco & Charbonneau, 2014; Fusco *et al.*, 2014). Two different mechanisms that drive protein crystallization are suggested, one driven by SER and another by electrostatics (Fusco *et al.*, 2014).

Data that relate the shape of a protein molecule to crystallization propensity with the use of surface mutagenesis are obscure. In the present paper, data describing static disorder (SD) and molecular shape as a factor in protein crystallization are presented. Analysis of selected protein structures crystallized with the use of SER demonstrates that in surface mutagenesis changes in the shape of the protein surface have non-uniform character and relate to successful crystallization events.

2. Materials and methods

The theoretical approach used in the present study is based on the unified model for protein dynamics (Frauenfelder *et al.*, 2009). The solvent hydration shell that surrounds a protein is an independent moiety, fluctuating over the surface of the protein molecule, that provides a framework of donor/acceptor atoms. The reservoir of entropy in the hydration shell induces entropy at the surface of the protein and drives functional motions in the matrix.

Protein structures of interest were downloaded from the PDB (Berman *et al.*, 2002). Comparison between wild-type and mutated structures of RhoGDI inhibitor was performed on polypeptide segments encompassing residues 67–202 with no gaps.

Stereochemical analysis of the protein structures and symmetry-related partners was performed with *Coot* (Emsley & Cowtan, 2004). Electron density for the protein molecules under investigation was calculated with the Electron Density Server at Uppsala University (Kleywegt *et al.*, 2004). In the present paper, the scale of Avbelj & Fele (1998) was used for side-chain conformational entropy (SCE) and a $-1.5 \text{ kcal mol}^{-1}$ cutoff was used to define a residue with SCE (an SCE residue). *AREAMOL* (Winn *et al.*, 2011) was used for solvent-accessible surface area (SASA) calculations with a probe radius of 1.4 Å . *SUPERPOSE* (Winn *et al.*, 2011) was used for superimposition of homologous structures and the calculation of r.m.s.d.. A minimum cutoff of 2 Å^2 was applied for a surface-exposed (SE) residue. Fractions were calculated as $\text{SD}_{\text{cryst}}/\text{SD}_{\text{se}}$ and $\text{SCE}_{\text{cryst}}/\text{SCE}_{\text{se}}$, where SD_{cryst} and $\text{SCE}_{\text{cryst}}$ are for residues found in crystal contacts, whereas SD_{se} and SCE_{se} are the corresponding values of SD and SCE of surface-exposed residues in the molecule. Flexible loops were defined as connecting polypeptide segments lacking hydrogen bonds between main-chain atoms and donor/acceptor atoms. Illustrations were prepared with *PyMOL* (DeLano, 2002).

Figure 1

(a) Stacking interactions in HEWL (PDB entry 2vb1; Wang *et al.*, 2007). Arg14 (yellow bonds) forms crystal contacts with the symmetry-related Arg61^{sym} (cyan bonds). The guanido groups form multiple van der Waals interactions. The nitrogen forms hydrogen bonds. Note that not all of the interactions occurring between symmetry-related residues are shown. (b) In neutrophil cytosol factor 1 (PDB entry 1kq6; M. Wahl, H. Delbrueck, H. Oschkinat & U. Heinemann, unpublished work) Lys16 (yellow bonds) adopts correlated alternative conformations with the symmetry-related Glu46^{sym} (cyan bonds). Salt bridges connect alternating NZ atoms with the symmetry-related Glu46^{sym} (cyan bonds) also found in alternative conformations. The rest of the side chain of Lys16 forms multiple van der Waals contacts with the symmetry-related Arg121^{sym} (cyan bonds), which is found in a single conformation.

3. Results and discussion

3.1. Crystal contacts tolerate high side-chain conformational entropy and static disorder

It has been argued before that the occurrence of patches of well ordered amino-acid side chains constitutes an essential prerequisite for the formation of intermolecular lattice contacts. Another finding relates to the importance of non-polar residues for successful crystallization events (Cieřlik & Derewenda, 2009; Price *et al.*, 2009). The data presented in Supplementary Table S1, based on a selection of 105 unambiguously monomeric protein structures refined to high resolution ($d < 1.5$ Å), do not corroborate those hypotheses. The majority of surface-exposed SCE residues (57%) were found in the contact area, although their share compared with all SE residues was 36%. The data presented here strongly correlate with previous findings made using the same methodology as used in the present study, counting the number of occurrences by residue type in crystal contacts (Dasgupta *et al.*, 1997). Dasgupta *et al.* (1997) found that the top five residues at crystal contacts were Lys, Glu, Arg, Ser and Asp, four of which are charged and four of which have SCE (Avbelj & Fele, 1998). Surprisingly, the occurrence of SD in crystal contacts was found to be present to an even greater extent. 62% of surface-exposed residues with SD (SD residues) were involved in crystal contacts, although their share compared with all SE residues was 16%. The latter observation means that if a protein molecule contains SD residues, they will be 3.9 times more likely to be found in a crystal contact compared with ordered residues.

Two types of stereochemical patterns characterized SD in the contact area. The first was where symmetry-related mates display conformers that move in a correlated manner (Fig. 1*a*). In hen egg-white lysozyme (HEWL) Arg14 oscillates between

two conformations, *A* and *B*, with occupancies of 0.6 and 0.4, respectively. The symmetry-related Arg61^{sym} is located at a van der Waals distance from Arg14 and moves in a similar fashion. The difference is that the symmetry-related *B* conformer swings further away from the *A* conformer compared with Arg14, and the occupancies are 0.73 and 0.27 for the *A* and *B* conformers, respectively. In a subset of residues exhibiting correlated disorder illustrated in Fig. 1(*b*), a single atom or a fraction of the side chain adopts an alternative conformation, while the rest is well ordered.

In the second pattern symmetry-related residues move in an independent mode (Fig. 2). Glu50 in rubredoxin displays two conformations. The higher populated conformer *A*, with an occupancy of 0.61, forms three hydrogen bonds to a symmetry-related molecule. One is between OE1 and the main-chain N atom of Ser25^{sym}. On the other hand, the OG atom of Ser25^{sym} oscillates between three conformers, two of which are hydrogen-bonded to OE1 and OE2 of Glu50. The second conformer of Glu50, with an occupancy of 0.39, stacks with the side chain of Lys7 from its own molecule at a distance of about 4.0 Å, which is indicative of multiple van der Waals contacts between both side chains.

In summary, the three examples of SD presented above seem to constitute a mechanism utilized by the crystal lattice to accommodate the disorder, and explains the tolerance of the crystal contacts to SCE residues as given in Supplementary Table S1.

3.2. A subset of SCE residues favor the contact area

Price and coworkers observed a higher than expected occurrence of Arg, Gln and Asn residues in crystal contacts despite their high SCE values (Price *et al.*, 2009). They hypothesized that the weaker effect of Arg and Gln in opposing crystallization may be related to the chemical nature of their side chains, which is similar to widely used protein denaturants such as guanidium and urea that interact strongly with proteins in aqueous solutions. The observation made by Price and coworkers does not correlate with SER theory (Derewenda, 2004) and the data reported in Cieřlik & Derewenda (2009) that discuss the role of polarity in intermolecular contacts in protein crystals, as well as the higher propensity for crystallization of so-called small residues (Gly, Ser, Ala and Thr).

In the present study, a subset of SCE residues, namely Gln and Arg, with a combined occurrence in proteins of 8.4 (Baud & Karlin, 1999), was compared with the group of small residues (Cieřlik & Derewenda, 2009), with a combined occurrence in proteins of 28.2 (Supplementary Table S2). Surprisingly, Gln and Arg were seven times more likely to occur in the crystal contact area compared with the group of small residues, which were 1.7 times more likely to form crystal contacts. Gln and Arg are 3.4-fold less abundant in proteins compared with the Gly, Ser, Ala and Thr group (Baud & Karlin, 1999). To normalize the statistical probabilities for both groups, SCE and small residues, based on their combined frequencies of occurrence in proteins, a factor of 3.4 was

Figure 2

An illustration that depicts noncorrelated disorder in the structure of rubredoxin (PDB entry 1yk4; Bönisch *et al.*, 2005). Alternative conformers of Glu50 (yellow bonds) form hydrogen bonds to the symmetry-related S25^{sym} (cyan bonds), which is also found in alternative conformations. In the second conformer Glu50 makes a van der Waals contact (4.0 Å) with its own Lys7. The electron-density map is contoured at the 1σ level.

applied to the occurrences of Gln and Arg in crystal contacts. As a result, the propensities of the Gln and Arg subset exceeded those of the Gly, Ser, Ala and Thr group by about 14-fold, a value significantly higher than that reported by Price *et al.* (2009). However, this seems to be consistent with the observation made by Dasgupta *et al.* (1997), who recommend Gln and Arg as good candidates for surface mutations based on their occurrence in crystal contacts. The potential of Lys→Arg mutations has been explored before (Czepas *et al.*, 2004) and it was demonstrated that the resultant mutant significantly improved the quality of crystals compared with the wild-type protein. Interestingly, the mutated Arg199 in the RhoGDI inhibitor was found at a van der Waals distance from a symmetrically related Arg199^{sym}, interacting in a manner reminiscent to that illustrated in Fig. 1. There are no statistical data available for Lys→Gln mutations and, obviously, as follows from the results reported above, such a mutation has a potential that is still underestimated.

3.3. Side-chain conformational entropy may be advantageous in the formation of cohesive crystal contacts

It is generally accepted that SCE is a factor that negatively affects the propensity of a group of amino acids to form crystal contacts. The results reported in Supplementary Tables S1 and S2 vividly demonstrate that the nature of the intermolecular protein crystal contacts has a more complex character than described in SER (Derewenda, 2004; Derewenda & Vekilov, 2006; Cieřlik & Derewenda, 2009). These results corroborate a view promoted in recent publications (Fusco & Charbonneau, 2014; Fusco *et al.*, 2014) and invite caution in their interpretation. Multiple factors are involved in the formation of protein crystal contacts, rather than a few dominant characteristics (McPherson, 1985; McPherson & Gavira, 2014).

Interestingly, SCE may be also an advantage in the formation of crystal contacts. Combined factors such as the large size of the side chain, its chemical nature and the ability to adopt many different conformations (SCE) characterizes Gln and Arg as versatile binders, as shown in Supplementary Table S2 and Fig. 1 as well as in Supplementary Figs. S1 and S2. As a consequence, these two residues do not belong to the group in which SCE has a negative impact on crystallization propensity.

3.4. Static disorder is driven largely by surface motions rather than by matrix flexibility

SER theory presumes that flexible residues are located on flexible loops (Cieřlik & Derewenda, 2009) and do not favor the formation of crystal contacts. This presumption implies that static disorder is a phenomenon originating in the core of the protein molecule rather than at the surface. The results presented in Supplementary Table S1 show that only 19% of SD residues are actually located on flexible loops, and therefore their SD is driven by the matrix, whereas the majority of SD residues are located on structured elements of the protein; consequently, their flexibility is a surface phenomenon. The latter observation is consistent with the unified model for

protein dynamics (Frauenfelder *et al.*, 2009) and data for the occurrence of SD in contact areas.

3.5. Molecular shape as a factor in crystallization of proteins by surface mutagenesis

A. I. Kitaigorodski first noted the role of the molecular shape in the crystallization of organic molecules (Kitaigorodski, 1961). The crystals of organic compounds are assembled by the governance of the Aufbau principle: molecules assemble in closely packed one-dimensional strings, which form closely packed two-dimensional layers; finally, closely packed layers form the three-dimensional structure of the crystals. Kitaigorodski's Aufbau principle is widely used for prediction of the symmetry of crystals and in crystal engineering (Ganguly & Desiraju, 2010).

Globular proteins have a much more complicated molecular shape, consisting of many 'hills' and 'depressions' at the surface. In proteins, the concept of molecular shape has mostly been discussed with respect to the development of synthetic proteins with surfaces that are conducive to self-assembly (for a review, see Lai *et al.*, 2012).

Despite success in understanding the mechanism of crystal growth in proteins and viruses (Vekilov, 2012; McPherson & Kuznetsov, 2011) and the basic physics of the protein condensed phase (Matsuura & Chernov, 2003), the crystallization of proteins still remains an empirical method. The success rate even at a high-throughput line for soluble to crystalline proteins is ~17% (Chayen & Saridakis, 2008).

Extensive efforts made recently to use surface mutagenesis as a tool to address the crystallization bottleneck (for reviews, see Dale *et al.*, 2003; Derewenda, 2011) yielded a large number of mutant structures. Those can be used as a database to better understand the consequences of surface mutations for the crystallization behavior of mutated proteins. It is intuitively clear that crystallization is determined by the microscopic nature of the protein surface. Surface mutants such as Lys→Ala and Glu→Ala substitutions in patches of 1–3 residues where large and charged side chains are replaced by small and nonpolar side chains constitute a significant change in molecular shape and the microscopic nature of the protein surface. Therefore, one can argue that reduction of the SCE is not the sole factor that contributes to the crystallization propensity of surface-mutated proteins.

3.5.1. A non-uniform relationship between surface mutations and molecular shape. The globular domain of human RhoGDI has been used as a model system to justify the validity of SER theory in the crystallization of proteins (Longenecker *et al.*, 2001; Cooper *et al.*, 2007; Cieřlik & Derewenda, 2009). This approach to crystallization achieves an overall success rate roughly estimated at about 15% (Derewenda, 2011).

To better understand the impact of molecular shape on the crystallization behavior of mutated proteins, a study was conducted that correlated changes in the molecular shape with the crystallization behavior of the protein. Molecular shape was quantitatively represented by the size of the SASA.

Changes in the protein structures were determined by the superimposition of multiple copies of mutated proteins in the asymmetric unit and by comparison of the mutants with the wild-type RhoGDI inhibitor. The results are given in Supplementary Table S3. The changes in SASA of the individual residues subject to surface mutagenesis are presented in Supplementary Table S4.

The first and foremost conclusion is that the relationships between surface mutations and molecular shape are non-uniform and are not tenable in an analysis that would account only for the relative changes of SASA in the mutated patch of residues.

In the E154,155A mutant (PDB entry 1kmt; Mateja *et al.*, 2002) there is a loss of 93 Å² owing to the replacement of glutamate with alanine compared with the wild-type protein, whereas the total SASA for the molecule increases by 5 Å² (Supplementary Tables S3 and S4). To the same end, an identical mutation but grown in a different crystallization medium (PDB entry 2jhw; Cooper *et al.*, 2007) results in a loss of SASA in mutated residues of 98 Å², but a gain of 51 Å² total SASA. Substitution of two glutamates with alanines, one of which has a different sequence number (E155,157A; PDB entry 2jhw; Cooper *et al.*, 2007), shows the opposite trend: a loss of SASA in the mutated patch of 73 Å² is accompanied by a loss in total SASA for the molecule of 28 Å². Note the changes in the geometries of the resulting crystal lattices and the contents of the asymmetric units compared with the wild-type RhoGDI inhibitor (Supplementary Table S3).

In a Lys→Ala series of mutants differences in molecular shape between the wild-type protein and mutants were found to an even larger extent. The K135,138,141A;F196I mutant (PDB entry 1fso; Longenecker *et al.*, 2001) exhibits a loss in SASA in the mutated patch of 174 Å² and a loss in total SASA equal to 73 Å², whereas a triple Lys→Ala mutant in the same patch has a closely similar loss in mutated residues of 172 Å² but gains 71 Å² in total SASA. Surprisingly, the largest change in the shape of the RhoGDI inhibitor was found for the single Lys→Ala mutation (PDB entry 1ft0; Longenecker *et al.*, 2001). The mutated residue becomes buried, resulting in a loss of 40 Å² compared with the wild-type protein, whereas the SASA of the mutated molecule decreases by 459 Å², a value equal to 6% of the surface of the molecule.

The asymmetric unit of the triple Lys→Tyr mutant (PDB entry 2bxw; Cooper *et al.*, 2007) contains two copies of the mutated protein. Comparison between mutated Tyr residues shows differences in the SASA of as large as 45% (Supplementary Table S4).

In modeling studies aimed at understanding the principles that govern protein crystallization, molecules are represented by hard spheres with fixed patches conducive to crystallization (Chang *et al.*, 2004; Fusco & Charbonneau, 2014). The present study reveals the 'soft' nature of the patches. Even when placed in identical solvent environments in the asymmetric unit, surface-exposed residues (patches) adopt different conformations with changes in the SASA of up to 45%, an indication of the dynamic nature of the patches (Supplementary Table S4).

An explanation for the non-uniform relationship may be sought in the dynamic nature of the protein surface. The unified model for protein dynamics defines a protein molecule as a picomachine with exquisite structural machinery that moves nearly without friction driven by the reservoir of entropy located in the solvent (Frauenfelder *et al.*, 2009). On the other hand, the bulk solvent forms clustered formations depending on the contents (Muscumeci *et al.*, 2012); therefore, its transient structure will vary depending on the composition. To complete the factors one must add the pH of the solvent, which will affect the charge distribution at the surface of the protein and will have an impact on how side chains interact with each other when forced together by movements of the solvent. In the crystal, the equilibrium conformations of the protein molecules in the asymmetric unit will be affected by crystal-packing forces as well. Taking all of these factors together, one may explain the origins of the non-uniform relationships illustrated in Supplementary Tables S3 and S4 founded on the dynamic shape of protein molecules modulated in the crystal by the crystallization medium and the packing forces.

The present study provides evidence in support of an approach considering the relative contributions of multiple factors that govern protein crystallization (McPherson, 1985; McPherson & Gavira, 2014), each with a share that may vary in a particular crystallization event, rather than surface entropy or hydrophobicity dominating in crystal formation (Cieřlik & Derewenda, 2009).

3.6. The relevance of the use of the scales for SCE in protein crystallization

It is generally accepted that the major force opposing protein folding is conformational entropy (Doig & Sternberg, 1995). In the unfolded state, protein residues probe a huge number of side-chain conformations by exploring rotamer space. Once the protein is folded, the rotameric degrees of freedom of the side chains are restricted to one in the core of the protein or one or a few at the surface of the protein. In this aspect, side-chain conformational entropy is one of the physical characteristics that relates mostly to the entropy, rather than the enthalpy, of protein folding. Quantitative evaluation of side-chain conformational entropy is achieved with the use of the available scales of SCE (Doig & Sternberg, 1995; Avbelj & Fele, 1998). Avbelj and Fele analyzed the role of the hydrophobic effect, side-chain conformational entropy and electrostatics and reached the conclusion that short-range main-chain electrostatic interactions dominate in the folding of the secondary structure. The role of hydrophobicity and side-chain conformational entropy was found to be small (Avbelj & Fele, 1998). Nascent proteins are synthesized in ribosomes under physiological conditions with the molarity of salts in the medium not exceeding 0.1 M. A review of the formulations of the most popular crystallization screens shows that predominant numbers of formulations contain a non-physiological concentration of salts and additives such as organic solvents. The presence of organic solvents in the

crystallization medium will change the dielectric properties of the solvent and a high salt concentration will shield interaction between charged groups. As a result, electrostatic interactions at the protein surface will be altered. The presence of a high concentration of organic solvents will also affect the ionization behavior of the charged side chains. All of these factors taken together raise the question as to whether the values for SCE calculated for a model protein-folding problem will be the same and accurately describe the SCE for a phenomenon which occurs in conditions that extend far beyond the physiological level. One indication of a possible inaccuracy of the scales for SCE in protein crystallization is the large discrepancy between the data for the crystallization propensity of Gln and Arg obtained by linear regression (Price *et al.*, 2009) and by occurrence at crystal contacts (Dasgupta *et al.*, 1997 and the present study). The second indication is the tolerance of crystal contacts to the whole group of SCE residues as indicated by their occurrence in the crystal contact area (Supplementary Table S1).

4. Conclusions

The occurrence of SD and high SCE in the area of protein crystal contacts may be a more common phenomenon than has previously been thought. The present study demonstrates that even in the most perfectly ordered crystals that yield structures diffracting to the best available resolution, the occurrence of well ordered patches that mediate crystal contacts is not an essential factor conducive to successful crystallization events.

This study supports the approach to the understanding of protein crystallization that accounts for multiple factors that govern the nucleation and the growth of protein crystals. A successful crystallization event is associated with the relative contributions of a plethora of physicochemical parameters, each with a relative contribution, that determine the crystallization outcome. The data reported here seem to be related to the relative share of the chemical nature of the amino-acid side chains involved in crystal contacts, the microscopic properties and the shape of the protein surface.

The present paper provides evidence in support of a strategy for the crystallization of proteins with the use of surface mutagenesis that is not limited to SCE residues. The immediate practical outcome of the present work is that virtually any surface-exposed residue can be considered to be a target to create a patch conducive to crystallization. A recent publication (Fusco *et al.*, 2014) provides some guidance in the selection of the target residues.

Acknowledgements

Careful proofreading of the manuscript by Dr Boris Atanasov and suggestions for improvement are gratefully acknowledged. The diligent help provided by Ms Margarita Dimova in the preparation of the manuscript for publication is highly appreciated. The author is also indebted to the anonymous reviewers and the editors of *Acta Crystallographica Section F*:

Structural Biology Communications for their help in bringing the manuscript into an acceptable shape for publication.

References

- Avbelj, F. & Fele, L. (1998). *J. Mol. Biol.* **279**, 665–684.
- Baud, F. & Karlin, S. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 12494–12499.
- Bergfors, T. (1999). *Protein Crystallization: Techniques, Strategies and Tips*. La Jolla: International University Line.
- Berman, H. M. *et al.* (2002). *Acta Cryst.* **D58**, 899–907.
- Bönisch, H., Schmidt, C. L., Bianco, P. & Ladenstein, R. (2005). *Acta Cryst.* **D61**, 990–1004.
- Chang, J., Lenhoff, A. M. & Sandler, S. I. (2004). *J. Chem. Phys.* **120**, 3003.
- Chayen, N. & Saridakis, E. (2008). *Nature Methods*, **5**, 147–153.
- Cieřlik, M. & Derewenda, Z. S. (2009). *Acta Cryst.* **D65**, 500–509.
- Cooper, D. R., Boczek, T., Grelewska, K., Pinkowska, M., Sikorska, M., Zawadzki, M. & Derewenda, Z. (2007). *Acta Cryst.* **D63**, 636–645.
- Czepas, J., Devedjiev, Y., Krowarsch, D., Derewenda, U., Otlewski, J. & Derewenda, Z. S. (2004). *Acta Cryst.* **D60**, 275–280.
- Dale, G. E., Oefner, C. & D'Arcy, A. (2003). *J. Struct. Biol.* **142**, 88–97.
- Dasgupta, S., Iyer, G. H., Bryant, S. H., Lawrence, C. E. & Bell, J. A. (1997). *Proteins*, **28**, 494–514.
- DeLano, W. L. (2002). *PyMOL*. <http://www.pymol.org>.
- Derewenda, Z. S. (2004). *Structure*, **12**, 529–535.
- Derewenda, Z. S. (2011). *Acta Cryst.* **D67**, 243–248.
- Derewenda, Z. S. & Vekilov, P. G. (2006). *Acta Cryst.* **D62**, 116–124.
- Doig, A. J. & Sternberg, M. J. (1995). *Protein Sci.* **4**, 2247–2251.
- Ducruix, A. & Giegé, R. (1992). *Crystallization of Nucleic Acids and Proteins: A Practical Approach*. Oxford: IRL Press.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Frauenfelder, H., Chen, G., Berendzen, J., Fenimore, P. W., Jansson, H., McMahon, B. H., Stroe, I. R., Swenson, J. & Young, R. D. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 5129–5134.
- Fusco, D., Barnum, T. J., Bruno, A. E., Luft, J. R., Snell, E. H., Mukherjee, S. & Charbonneau, P. (2014). *PLoS One*, **9**, e101123.
- Fusco, D. & Charbonneau, P. (2014). *J. Phys. Chem. B*, **118**, 8034–8041.
- Ganguly, P. & Desiraju, G. R. (2010). *CrystEngComm*, **12**, 817–833.
- Kitaigorodski, A. I. (1961). *Organic Chemical Crystallography*, pp. 65–112. New York: Consultants Bureau.
- Kleywegt, G. J., Harris, M. R., Zou, J., Taylor, T. C., Wählby, A. & Jones, T. A. (2004). *Acta Cryst.* **D60**, 2240–2249.
- Lai, Y.-T., King, P. N. & Yeates, T. O. (2012). *Trends Cell Biol.* **22**, 653–661.
- Longenecker, K. L., Garrard, S. M., Sheffield, P. J. & Derewenda, Z. S. (2001). *Acta Cryst.* **D57**, 679–688.
- Mateja, A., Devedjiev, Y., Krowarsch, D., Longenecker, K., Dauter, Z., Otlewski, J. & Derewenda, Z. S. (2002). *Acta Cryst.* **D58**, 1983–1991.
- Matsuura, Y. & Chernov, A. A. (2003). *Acta Cryst.* **D59**, 1347–1356.
- McPherson, A. (1985). *Methods Enzymol.* **114**, 112–120.
- McPherson, A. & Gavira, J. A. (2014). *Acta Cryst.* **F70**, 2–20.
- McPherson, A. & Kuznetsov, Y. G. (2011). *Methods Mol. Biol.* **736**, 171–195.
- Muscumeci, F., Grasso, R., Lanzano, L., Scordino, A., Triglia, A., Tudisco, S. & Gulino, M. (2012). *J. Biol. Phys.* **38**, 181–195.
- Price, W. N. II *et al.* (2009). *Nature Biotechnol.* **27**, 51–57.
- Vekilov, P. G. (2003). *Methods Enzymol.* **368**, 84–105.
- Vekilov, P. (2012). *J. Phys. Condens. Matter*, **24**, 1–16.
- Wang, J., Dauter, M., Alkire, R., Joachimiak, A. & Dauter, Z. (2007). *Acta Cryst.* **D63**, 1254–1268.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.