Crystal structure of the cyan fluorescent protein Cerulean-S175G

Sang-wook Park, Sunghyun Kang* and Tae-Sung Yoon*

Bio-Analytical Science Major, Korea University of Science and Technology, KRIIBB Campus, 125 Gwahak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea. *Correspondence e-mail: skang@kribb.re.kr, yoonts@ust.ac.kr

Received 6 May 2016

Accepted 21 May 2016

Edited by T. C. Terwilliger, Los Alamos National Laboratory, USA

Keywords: Cerulean-S175G; enhanced cyan fluorescent protein; fluorophores; fluorescence resonance energy transfer; fluorescence lifetime imaging microscopy.

PDB reference: Cerulean-S175G, 5j2o

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

Enhanced cyan fluorescent protein (ECFP) was derived from *Aequorea victoria* green fluorescent protein (avGFP), notably with S65T/Y66W mutations. Its chromophore consists of a tripeptide comprised of Thr65, Trp66 and Gly67 (TWG) residues, while that of avGFP consists of a Ser65, Tyr66 and Gly67 (SYG) tripeptide. Cerulean and SCFP3A were derived from ECFP-S72A/H148D (a double mutation) with additional Y145A and S175G mutations, respectively, while Cerulean-S175G has both mutations (Y145A and S175G). The crystal structures of these ECFP variants at neutral pH were reported to adopt two distinct major conformations called *ECFP* and *Cerulean*. In this study, Cerulean-S175G was revealed to adopt only the *Cerulean* conformation, while Cerulean has been reported to adopt both the *ECFP* and the *Cerulean* conformations in its crystal structures. Sharing the same S175G mutation with SCFP3A, Cerulean-S175G showed a slightly increased quantum yield, like SCFP3A, but did not adopt the *ECFP* conformation adopted by SCFP3A. Detailed comparison of Cerulean-S175G and other ECFP variants revealed that the notable conformational changes in ECFP variants can be understood mainly in terms of the interaction between the Trp66 residue of the chromophore and residues 145–148 of β -strand 7.

1. Introduction

Cyan fluorescent proteins (CFPs) have been developed either by the structure-guided evolution of green fluorescent protein (GFP) obtained from *Aequorea victoria* (Goedhart *et al.*, 2012) or by the discovery of new CFPs such as Kusabira-Cyan, cmFP512, asFP499, amFP486, dsFP483 and cFP484 obtained from other organisms such as *Clavularia* sp. (Ai *et al.*, 2006). Enhanced cyan fluorescent protein (ECFP) was derived from avGFP, notably with S65T/Y66W mutations (Bae *et al.*, 2003). Its chromophore consists of a tripeptide of Thr65, Trp66 and Gly67 (TWG) residues, while that of avGFP consists of a Ser65, Tyr66 and Gly67 (SYG) tripeptide. Since the introduction of ECFP, CFPs have been widely used as donor fluorophores in fluorescence resonance energy transfer (FRET) experiments, with yellow fluorescent proteins (YFPs) as acceptor fluorophores. In this regard, ECFP variants of the reported CFPs have been developed extensively as CFP–YFP FRET pairs. Such FRET pairs can be fused and used to monitor *in vivo* processes using fluorescence lifetime imaging microscopy (FLIM; Markwardt *et al.*, 2011).

The original ECFP showed problems of moderate brightness and complex fluorescence decay kinetics, which could compromise successful applications in FRET/FLIM (Goedhart *et al.*, 2012; Markwardt *et al.*, 2011). To overcome these problems, two notable lines of ECFP variants were introduced (Watkins *et al.*, 2013). The first line resulted in Cerulean (Rizzo *et al.*, 2004), which evolved into mCerulean3 (Markwardt *et al.*, 2011), while the second line resulted in super cyan fluorescent

Table 1

Macromolecule-production information.

Source organism	<i>A. victoria</i>
DNA source	Synthetic (Bioneer Inc., Republic of Korea)
Forward primer	5'-GGGAATTCATATGGTGA-3'
Reverse primer	5'-CGGCTCGAGTTAGTGGTG-3'
Cloning vector	pBHA
Expression vector	pET-21a
Expression host	<i>E. coli</i> BL21(DE3)
Complete amino-acid sequence of the construct produced	MVSKGEELFTGVVPILVELDGDVNGHKFVSVEGEG-EGDATYGKLTCLKFICTTGKLPVPWPTLVTTLT-VQCFARYPDHMKQHDFFKSAMPEGYVQERTIF-FKDDGNYKTRAEVKFEGDTLVNRIELKGIDFK-EDGNILGHKLEYNAISDNVYITADKQKNGIKA-NFKIRHNIEDGGVQLADHYQNTPIGDGPVLL-PDNHYLSTQSALS KDPNEKRDHMLLEFVTAA-GITLGMDELYKHHHHH

protein 3A (SCFP3A; Kremers *et al.*, 2006), which evolved into mTurquoise2 (Goedhart *et al.*, 2012). As summarized in Supplementary Tables S1 and S2, both Cerulean and SCFP3A were derived from ECFP-S72A/H148D (a double mutation) with additional Y145A and S175G mutations, respectively. Hence, Cerulean-S175G has a single-residue difference from either Cerulean or SCFP3A. Because of their high degree of sequence similarity and their conserved unit-cell dimensions (Supplementary Tables S3 and S4), significant conformational differences between them would be interesting.

Two crystal structures of ECFP have been reported (Bae *et al.*, 2003), in which one crystal had conformation *A*, later identified as the *GFP* conformation (PDB entry 1oxd), and the other crystal had conformation *B*, later identified as the *ECFP* conformation (PDB entry 1oxe). Later, another crystal structure of ECFP (PDB entry 2wsn) was investigated with the help of molecular dynamics to reveal further details of the multiplicity of the conformation of ECFP (Lelimousin *et al.*, 2009). The crystal structure of Cerulean determined at pH 5.0 (PDB entry 2q57) showed a *cis* conformation of the chromophore, unlike the known ECFP structures (Malo *et al.*, 2007), while that determined at neutral pH (PDB entry 2wso) showed a *trans* conformation of the chromophore and two distinct major conformations called the *ECFP* and *Cerulean* conformations in a single crystal (Lelimousin *et al.*, 2009). In a subsequent study reported by the same group, SCFP3A (PDB entry 2ydz) was shown to adopt only the *ECFP* conformation (Goedhart *et al.*, 2012). As it shares the S175G mutation with SCFP3A and the Y145A mutation with Cerulean, structural elucidation of Cerulean-S175G may be interesting in conjunction with its corresponding spectral changes.

2. Materials and methods

2.1. Macromolecule production

The Cerulean-S175G coding sequence, which was originally synthesized in the pBHA vector (Bioneer Inc., Republic of Korea), was transferred into the pET-21a vector using the *Nde*I and *Xho*I restriction sites to insert the Cerulean-S175G coding sequence. The recombinant protein was expressed in *Escherichia coli* BL21(DE3) strain competent cells. LB

Table 2

Crystallization.

Method	Sitting-drop vapour diffusion
Plate type	24-well sitting-drop plate (Axygen)
Temperature (K)	291
Protein concentration (mg ml ⁻¹)	33
Buffer composition of protein solution	20 mM Tris-HCl pH 8.0
Composition of reservoir solution	16% PEG 8000, 100 mM MgCl ₂ , 100 mM HEPES pH 7.5
Volume and ratio of drop	3 µl, 1:1
Volume of reservoir (µl)	500

Table 3

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	Beamline 5C, PAL/PLS
Wavelength (Å)	0.9796
Temperature (K)	100
Detector	ADSC Quantum 315r
Crystal-to-detector distance (mm)	170
Rotation range per image (°)	1
Total rotation range (°)	180
Exposure time per image (s)	1.0
Space group	<i>P</i> ₂ ₁ ₂ ₁
<i>a</i> , <i>b</i> , <i>c</i> (Å)	51.14, 62.90, 70.11
α , β , γ (°)	90, 90, 90
Mosaicity (°)	0.1333–0.279
Resolution range (Å)	50.0–1.50 (1.53–1.50)
Total No. of reflections	258406
No. of unique reflections	36843
Completeness (%)	100.0 (100.0)
Multiplicity	7.0 (6.8)
$\langle I/\sigma(I) \rangle$	10.3 (2.56)
<i>R</i> _{int} ,†	0.087
Overall <i>B</i> factor from Wilson plot (Å ²)	9.4

† Estimated $R_{int} = R_{merge}[N/(N-1)]^{1/2}$, where *N* is the data multiplicity.

medium was inoculated with a 1% volume of a saturated culture of the expression strain. After growth at 310 K to an optical density of 0.6, isopropyl β -D-1-thiogalactopyranoside was added to a concentration of 1 mM and the cultures were grown for an additional 21 h at 291 K. The cells were lysed by sonication in the presence of 1× Dulbecco's phosphate-buffered saline and 10 mM β -mercaptoethanol. The His-tagged protein was purified on an ÄKTAprime/ÄKTExpress system using an Ni Sepharose excel nickel-affinity chromatography column (GE Healthcare Life Sciences) and eluted with 500 mM imidazole in the abovementioned buffer. It was further purified using a HiPrep 16/60 Sephacryl S100 gel-filtration chromatography column (GE Healthcare Life Sciences), replacing the final buffer with 20 mM Tris-HCl pH 8.0. Macromolecule-production information is summarized in Table 1.

2.2. Crystallization

The purified Cerulean-S175G protein was concentrated to 33 mg ml⁻¹ for crystallization. The concentrated protein was crystallized using the sitting-drop vapour-diffusion method at 291 K. Initially, crystallization was screened by varying the PEG 8000 concentration in the range 8–23% (w/v) in 100 mM HEPES pH 7.5, 100 mM MgCl₂ based on the crystallization conditions for Cerulean (Lelimousin *et al.*, 2009). Finally, the

crystallization was optimized at 16% (w/v) PEG 8000. Detailed crystallization conditions are summarized in Table 2.

2.3. Data collection and processing

The mounted crystal was flash-cooled in liquid nitrogen before exposure to X-rays. Cryoprotectant was not needed because the mother liquor contained 16% PEG 8000. Diffraction data collection was performed at 100 K on beamline 5C (wavelength 0.97960 Å) at Pohang Light Source (PLS) operated by Pohang Accelerator Laboratory (PAL), which was equipped with an ADSC Quantum 315r detector. *HKL-2000* v.1.97.2 (Otwinowski & Minor, 1997) was used for data processing, including indexing and scaling. Detailed data-collection and processing statistics are shown in Table 3.

2.4. Structure solution and refinement

The 1.59 Å resolution structure of SCFP3A (PDB entry 2ydz; Goedhart *et al.*, 2012) was used as a starting model for refinement with *PHENIX* v.1.10.1-2155 (Adams *et al.*, 2010). After water molecules had been removed from the starting model, rigid-body refinement of the model was tried as well as the use of *Phaser-MR*. It turned out that rigid-body refinement gave an electron-density map of sufficient quality to reveal the structural difference between the model protein (SCFP3A) and the target protein (Cerulean-S175G). Ile146 was rotated to fit the direction of electron density and Tyr145 was mutated to Ala145 using *WinCoot* (v.0.8.2; Emsley & Cowtan, 2004).

Figure 1

The overall structure of Cerulean-S175G (PDB entry 5j2o); residues are shown with ECFP residue numbering. The overall secondary structure is shown in red, yellow and green. The chromophore involving the Trp66 residue is shown in cyan. The residues mutated from ECFP in the present work (Supplementary Table S1) are shown in pink. The Ile146 and Ser147 residues which are mutated in ECFP and mCerulean3 (Supplementary Table S2) are shown in orange.

Table 4

Structure refinement.

Values in parentheses are for the outer shell.

Completeness (%)	98.2
σ Cutoff	$F > 1.34\sigma(F)$
No. of reflections, working set	36841
No. of reflections, test set	2000
Final R_{cryst}	0.169
Final R_{free}	0.188
Cruickshank DPI	0.12
No. of non-H atoms	
Total	2062
Protein	1789
Ligand	24
Water	249
Total	3836
R.m.s. deviations	
Bonds (Å)	0.008
Angles (°)	1.07
Average B factor (Å ²)	
Overall	12.13
Protein	11.2
Water	19.7
Ramachandran plot	
Favoured regions (%)	98
Additionally allowed (%)	2
Outliers (%)	0

Further refinements were made to check the disorder present in the model. The refinement of Cerulean-S175G was continued with isotropic B -factor refinement to add H atoms to the residues, including the chromophore, while that of the SCFP3A structure was stopped with anisotropic B -factor refinements without adding any H atoms.

Cerulean-S175G has a chromophore named CRF by the PDB (<http://www.rcsb.org/pdb/ligand/ligandsummary.do?hetId=CRF>) which should be treated as a ligand in *PHENIX* refinement. The program *eLBOW* (Moriarty *et al.*, 2009), as part of the *PHENIX* suite, was used to generate the restraint information for CRF, which requires that leaving atoms such as HN12, OXT and HXT are automatically removed when the ligand links to neighbouring residues. However, *ReadySet!* (the program that adds H atoms in the *phenix.refine* GUI) was not able to add the correct H atoms in this case (Moriarty, 2016; information from help@phenix-online.org). The automatic adding hydrogen option generated 'HN11' and 'HN12' between the 'N1' atom of CRF and the 'C' atom of Leu64, while the correct structure should have only one hydrogen, which should be coplanar with the peptide bond between CRF and Leu64. Likewise, the automatic adding hydrogen option generated no H atoms between CRF and Val68, while the correct structure should have one H atom which should also be coplanar to the peptide bond between CRF and Val68. Hence, we had to curate those two H atoms manually using *PyMOL* (DeLano, 2002) instead of adding them in the *phenix.refine* GUI. Water molecules without H atoms were automatically added by *PHENIX* and checked against electron density using *WinCoot* in the final stages of *PHENIX* refinement. Finally, the *phenix.pdbtools* segID field was used to correct the atomic numbering. *MolProbity* (Chen *et al.*, 2010) was used for Ramachandran analysis. All figures showing protein structures were produced using *PyMOL*

(DeLano, 2002). The structure has been submitted to the Protein Data Bank and assigned PDB entry 5j2o. Statistics of structure refinement are summarized in Table 4.

3. Results and discussion

3.1. Overview of Cerulean-S175G and related crystal structures

The overall structure of Cerulean-S175G presented in Fig. 1 shows the typical 11-stranded β -barrel structure of *avGFP* (Ormö *et al.*, 1996; PDB entry 1ema). The mutations relative to ECFP and wild-type *avGFP* are listed in Supplementary Tables S1 and S2. As demonstrated by the successful direct rigid-body refinement approach without the need for any molecular-replacement step, the orientation and packing of the Cerulean-S175G monomer crystallized in space group $P2_12_12_1$ seems to be well conserved, together with its unit-cell dimensions, across related crystal structures (Supplementary Table S3). Considering the recent report on the systematic investigation of engineered GFPs for oligomeric scaffolding (Leibly *et al.*, 2015), a systematic investigation of the monomeric packing of GFP variants and their disordered residues would seem to be worthwhile for the precision design of GFP variants in the context of crystallographic engineering.

There are very slight sequence differences between the related structures, as summarized in Supplementary Table S4. Cerulean-S175G contains a His tag which was not present in Cerulean (PDB entries 2wso and 4as8) and was not visible in the electron-density map, along with a few N-terminal and C-terminal residues. All of the residues between Lys3 and Thr230 were visible in the electron-density map. It seems that the His tag does not have any significant impact either on the crystallization or on the structure. The presented model could not avoid disorder at four residues, Ser30, Glu90, Gln184 and Leu220, which were also disordered in the three most closely related structures (PDB entries 2wso, 4as8 and 2ydz, which also had their own disordered residues, with totals of 25, 31 and seven disordered residues, respectively). As a result, the Cerulean-S175G structure has fewer disordered residues than the Cerulean and SCFP3A structures. This would seem to be reasonable because the Cerulean-S175G monomer has the most reduced surface entropy among the related structures.

3.2. Spectral properties of Cerulean-S175G and the conformation of its chromophore

Double-peaked absorption and fluorescence emission spectra and single-crystal photographs of Cerulean-S175G are shown in Fig. 2. The excitation and emission spectra were measured using diluted Cerulean-S175G protein solution, which was also used for crystallization. The quantum yields (QYs) of the Cerulean, SCFP3A and Cerulean-S175G proteins were also measured. Details of fluorescence QY measurements are given in §S1 (Supporting Information). The reported QYs for Cerulean, SCFP3A and SCFP3B are 0.49, 0.56 and 0.50, respectively (Goedhart *et al.*, 2012). Our own measurements of the QYs for these proteins as purified by us

from our 6 $\times$ His-tagged constructs were 0.42, 0.54 and 0.53, respectively (the in-house preparations of SCFP3A and SCFP3B are also named SCFP3A-K206A and Cerulean-S175G; see the footnotes to Supplementary Tables S1 and S2). The QY measurement errors for the three proteins seem to be comparable with the variation in the repeated QY measurements of ECFP (Wall *et al.*, 2015).

The chromophores of Cerulean-S175G and Cerulean at physiological pH (Lelimosin *et al.*, 2009) have a *trans* conformation, while that of Cerulean at pH 5.0 (Malo *et al.*, 2007) has a *cis* conformation. Previously, it was believed that a low pH could induce this dramatic conformational change, while the chromophore at physiological pH should maintain a *trans* conformation (Malo *et al.*, 2007; Lelimosin *et al.*, 2009). Recently, mCerulean3 at pH 4.7 (Watkins *et al.*, 2013) was

Figure 2 Fluorescence spectroscopy analysis and crystallization. (a) Fluorescence spectroscopy analysis shows double-peaked absorption and fluorescence emission spectra. Excitation and emission spectra are presented in green and blue, respectively. The excitation wavelengths (λ_{ex}) are 437 and 453 nm, while the emission wavelengths (λ_{em}) are 476 and 504 nm. (b) Crystals were grown in 16% PEG 8000, 100 mM HEPES, 100 mM MgCl_2 . Mostly, only one single crystal (indicated with a red circle) was obtained from each drop. A single crystal obtained from another drop is presented at the top.

reported to have a *trans* conformation of its chromophore. The lack of any pH-dependent spectral changes in mCerulean3 confirmed the elimination of pH-dependent chromophore isomerization (Markwardt *et al.*, 2011; Watkins *et al.*, 2013). More recently, three crystal structures of NowGFP were reported at pH 9.0 and pH 4.8 (Pletnev *et al.*, 2015). Very interestingly, the chromophore at pH 9.0 had a *cis* conformation and the chromophore at pH 4.8 had 80% *cis* conformation and 20% *trans* conformation. However, Table 3 of Pletnev *et al.* (2015) seemed to suffer from conflicting conformational assignments in terms of the application of *trans/cis* and *E/Z* nomenclature (§S2, Supporting Information). In any case, we believe that further understanding of the chromophore conformation together with its pH sensitivity and pH robustness would be desirable to serve both specific and broader applications in biology.

3.3. Structural comparison of Cerulean, SCFP3A and Cerulean-S175G

All three structures, Cerulean-S175G (PDB entry 5j2o), SCFP3A (PDB entry 2ydz) and Cerulean (PDB entries 2wso and 4as8), have a *trans* conformation of their chromophores.

Notable conformational differences between Cerulean-S175G (PDB entry 5j2o), SCFP3A (PDB entry 2ydz) and Cerulean (PDB entry 2wso) can be understood mainly in terms of the interaction between Trp66 of the chromophore and residues 145–148 of β -strand 7. This main conformational difference is highlighted graphically with the corresponding interaction distances in Fig. 3. Structural comparisons between the selected GFP variants were also performed using a root-mean-square deviation (r.m.s.d.) analysis (Supplementary Table S5). An all-C $^{\alpha}$ -atom comparison of Cerulean-S175G (PDB entry 5j2o) with Cerulean (PDB entry 2wso) and SCFP3A (PDB entry 2ydz) results in r.m.s.d.s of 0.110 and 0.126 Å, respectively. When the alternate conformations were separately calculated for Cerulean (PDB entry 2wso), C $^{\alpha}$ -atom comparison for residues 145–148 of β -strand 7 showed a clearer distinction, with r.m.s.d.s of 0.068 and 0.158 Å for the comparison of Cerulean-S175G with conformations *A* and *B* of Cerulean (PDB entry 2wso), respectively.

We tried to confirm the interaction between residue Ile146 and the chromophore (Supplementary Figs. S2 and S3). Cerulean (PDB entry 2wso) has both the *ECFP* conformation ($d_1 = 3.9$ Å, $d_2 = 5.9$ Å) and the *Cerulean* conformation ($d_1 = 6.0$ Å, $d_2 = 3.7$ Å) (Lelimosin *et al.*, 2009). SCFP3A

Figure 3

Comparison of the Cerulean, SCFP3A and Cerulean-S175G structures. Trp66 and β -strand 7 (145–148) are shown as stick models. The three structures are indicated as ribbon diagrams in (a) and (b). (a) Cerulean (PDB entry 2wso) is shown in yellow and Cerulean-S175G (PDB entry 5j2o) is shown in red. (b) SCFP3A (PDB entry 2ydz) is shown in cyan and Cerulean-S175G (PDB entry 5j2o) is shown in red. Each map was generated as described in §S3 (Supporting Information) rather than by the simple use of the Electron Density Server (Kleywegt *et al.*, 2004). (c, d) Trp66 of the chromophore and residues 145–148 of β -strand 7 in Cerulean (PDB entries 2wso and 4as8) with two alternate conformations shown as stick models. (e) The same residues in Cerulean-S175G (PDB entry 5j2o). (f) The same residues in SCFP3A (PDB entry 2ydz). In Cerulean-S175G and SCFP3A, Ile146 has a single conformation. In all three structures, Ile146 and Trp66 make van der Waals interactions. The van der Waals interaction distances classify the structural similarity. The distances d_1 and d_2 between the CZ2 atom of Trp66 and either the CG2 atom or the CD1 atom of Ile146 are shown in red and yellow, respectively. d_1 (red) and d_2 (yellow) indicate the distances between Trp66 and Ile146. The direction of Ile146 determines the *ECFP* and *Cerulean* conformations. Cerulean (PDB entry 2wso), which contains two alternate conformations in β -strand 7, has two conformations: the *ECFP* conformation ($d_1 = 3.9$ Å, $d_2 = 5.9$ Å) and the *Cerulean* conformation ($d_1 = 6.0$ Å, $d_2 = 3.7$ Å) (Lelimosin *et al.*, 2009). SCFP3A only has the *ECFP* conformation ($d_1 = 4.0$ Å, $d_2 = 5.8$ Å), while Cerulean-S175G only has the *Cerulean* conformation ($d_1 = 6.0$ Å, $d_2 = 3.8$ Å). In the case where there is a single conformation of Ile146, the Tyr/Ala145 mutation may directly affect the direction of Ile146.

Figure 3 (continued)

The interaction of Trp66 and Ile146 in Cerulean (PDB entries 2wso and 4as8), Cerulean-S175G (PDB entry 5j2o) and SCFP3A (PDB entry 2ydz) is presented as a $2mF_o - DF_c$ map contoured at the 1.0σ level in (c), (d), (e) and (f), respectively.

(PDB entry 2ydz) only has the *ECFP* conformation ($d_1 = 4.0$ Å, $d_2 = 5.8$ Å), while Cerulean-S175G (PDB entry 5j2o) only has the *Cerulean* conformation ($d_1 = 6.0$ Å, $d_2 = 3.8$ Å; Fig. 3). Where Ile146 has a single conformation, Tyr/Ala145 may directly affect the direction of Ile146.

For residues 173–175, the r.m.s.d.s of Cerulean-S175G (PDB entry 5j2o) from Cerulean (PDB entry 2wso) and SCFP3A (PDB entry 2ydz) are 0.071 Å (seven atoms) and 0.031 Å (six atoms), respectively. This confirms that Cerulean-S175G is very similar to the SCFP3A structure at residues 173–175. Asp173 of Cerulean-S175G moves outwards 0.5 Å compared

with Ser175 of Cerulean. As a result, Ser/Gly175 affects Asp173, thus changing the connecting loop between β -strands 8 and 9 (Supplementary Fig. S1). This is also consistent with the role of the S175G mutation in Venus, a YFP (Rekas *et al.*, 2002). It seems that the role of the S175G mutation is not as a crystallographic artifact because Venus crystallizes as a dimer in space group $P3_112$.

In conclusion, we attempted to determine the structural differences between Cerulean, SCFP3A and Cerulean-S175G. We hope that these findings will be an interesting contribution to the fluorescent protein research field, although further studies may be needed.

Acknowledgements

We are grateful to the staff scientists of beamline 5C at Pohang Accelerator Laboratory for help with the diffraction experiments. This research was supported by the National Space Laboratory (NSL) Program through the National Research Foundation of Korea (NRF-2013M1A3A3A02042368) funded by the Ministry of Science, ICT and Future Planning of Korea. This work was also supported by grants from KRIIBB Research Initiative Program.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Ai, H.-W., Henderson, J. N., Remington, S. J. & Campbell, R. E. (2006). *Biochem. J.* **400**, 531–540.
- Bae, J. H., Rubini, M., Jung, G., Wiegand, G., Seifert, M. H. J., Azim, M. K., Kim, J.-S., Zumbusch, A., Holak, T. A., Moroder, L., Huber, R. & Budisa, N. (2003). *J. Mol. Biol.* **328**, 1071–1081.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- DeLano, W. L. (2002). *PyMOL*. <http://www.pymol.org>.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Goedhart, J., von Stetten, D., Noirclerc-Savoye, M., Lelimosin, M., Joosen, L., Hink, M. A., van Weeren, L., Gadella, T. W. & Royant, A. (2012). *Nature Commun.* **3**, 751.
- Kleywegt, G. J., Harris, M. R., Zou, J., Taylor, T. C., Wahlby, A. & Jones, T. A. (2004). *Acta Cryst.* **D60**, 2240–2249.
- Kremers, G. J., Goedhart, J., van Munster, E. B. & Gadella, T. W. Jr (2006). *Biochemistry*, **45**, 6570–6580.
- Leibly, D. J., Arbing, M. A., Pashkov, I., DeVore, N., Waldo, G. S., Terwilliger, T. C. & Yeates, T. O. (2015). *Structure*, **23**, 1754–1768.
- Lelimosin, M., Noirclerc-Savoye, M., Lazareno-Saez, C., Paetzold, B., Le Vot, S., Chazal, R., Macheboeuf, P., Field, M. J., Bourgeois, D. & Royant, A. (2009). *Biochemistry*, **48**, 10038–10046.
- Malo, G. D., Pouwels, L. J., Wang, M., Weichsel, A., Montfort, W. R., Rizzo, M. A., Piston, D. W. & Wachter, R. M. (2007). *Biochemistry*, **46**, 9865–9873.
- Markwardt, M. L., Kremers, G.-J., Kraft, C. A., Ray, K., Cranfill, P. J. C., Wilson, K. A., Day, R. N., Wachter, R. M., Davidson, M. W. & Rizzo, M. A. (2011). *PLoS One*, **6**, e17896.
- Moriarty, N. W. (2016). Personal communication.
- Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. (2009). *Acta Cryst.* **D65**, 1074–1080.
- Ormö, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y. & Remington, S. J. (1996). *Science*, **273**, 1392–1395.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pletnev, V. Z., Pletneva, N. V., Sarkisyan, K. S., Mishin, A. S., Lukyanov, K. A., Goryacheva, E. A., Ziganshin, R. H., Dauter, Z. & Pletnev, S. (2015). *Acta Cryst.* **D71**, 1699–1707.
- Rekas, A., Alattia, J. R., Nagai, T., Miyawaki, A. & Ikura, M. (2002). *J. Biol. Chem.* **277**, 50573–50578.
- Rizzo, M. A., Springer, G. H., Granada, B. & Piston, D. W. (2004). *Nature Biotechnol.* **22**, 445–449.
- Wall, K. P., Dillon, R. & Knowles, M. K. (2015). *Biochem. Mol. Biol. Educ.* **43**, 52–59.
- Watkins, J. L., Kim, H., Markwardt, M. L., Chen, L., Fromme, R., Rizzo, M. A. & Wachter, R. M. (2013). *Acta Cryst.* **D69**, 767–773.