Crystal structures of green fluorescent protein with the unnatural amino acid 4-nitro-L-phenylalanine

Nicole Maurici, Nicole Savidge, Byung Uk Lee, Scott H. Brewer and
Christine M. Phillips-Piro*

Department of Chemistry, Franklin and Marshall College, PO Box 3003, Lancaster, PA 17604, USA. *Correspondence
e-mail: cpiro@fandm.edu

Received 21 June 2018

Accepted 17 August 2018

Edited by F. T. Tsai, Baylor College of Medicine,
Houston, USA

Keywords: 4-nitro-L-phenylalanine; unnatural
amino acids; noncanonical amino acids; green
fluorescent protein.

PDB references: sfGFP, with Asp133 mutated to
4-nitro-L-phenylalanine, 6dq0; with Asn149
mutated to 4-nitro-L-phenylalanine, 6dq1

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

The X-ray crystal structures of two superfolder green fluorescent protein (sfGFP) constructs containing a genetically incorporated spectroscopic reporter unnatural amino acid, 4-nitro-L-phenylalanine (pNO₂F), at two unique sites in the protein have been determined. Amber codon-suppression methodology was used to site-specifically incorporate pNO₂F at a solvent-accessible (Asp133) and a partially buried (Asn149) site in sfGFP. The Asp133pNO₂F sfGFP construct crystallized with two molecules per asymmetric unit in space group *P*₃₂21 and the crystal structure was refined to 2.05 Å resolution. Crystals of Asn149pNO₂F sfGFP contained one molecule of sfGFP per asymmetric unit in space group *P*₄₁22 and the structure was refined to 1.60 Å resolution. The alignment of Asp133pNO₂F or Asn149pNO₂F sfGFP with wild-type sfGFP resulted in small root-mean-square deviations, illustrating that these residues do not significantly alter the protein structure and supporting the use of pNO₂F as an effective spectroscopic reporter of local protein structure and dynamics.

1. Introduction

The ease of incorporation of a wide variety of unnatural amino acids (UAAs) site-specifically into proteins provides scientists with the opportunity to expand the tools with which they study local protein structure and dynamics. Spectroscopic reporter-containing UAAs represent one subclass of UAAs that can be utilized to study these features in proteins. UAAs containing vibrational reporters, fluorescent groups and NMR probes have all been genetically incorporated into protein systems (Schultz *et al.*, 2006; Jackson *et al.*, 2007; Cellitti *et al.*, 2008; Miyake-Stoner *et al.*, 2010; Ye *et al.*, 2010; Smith *et al.*, 2011; Thielges *et al.*, 2011; Bazewicz *et al.*, 2012, 2013; Adhikary *et al.*, 2014; Ma *et al.*, 2015; Tookmanian, Fenlon *et al.*, 2015; Tookmanian, Phillips-Piro *et al.*, 2015; Dippel *et al.*, 2016; Hartley *et al.*, 2016; Slocum & Webb, 2016). In order to be effective probes of protein structure, the probes must have both a specific spectroscopic signal sensitive to the local environment and be minimally perturbative to the local protein environment under investigation.

The spectroscopic reporter-containing UAA 4-nitro-L-phenylalanine (pNO₂F; Fig. 1) has been shown to act as both a fluorescence quencher (Taki *et al.*, 2002; Tsao *et al.*, 2006) and a vibrational reporter (Smith *et al.*, 2011). Two earlier studies incorporating pNO₂F into either streptavidin or the leucine zipper illustrated that pNO₂F acts as a quencher of tryptophan fluorescence in a distance-dependent manner (Taki *et al.*, 2002; Tsao *et al.*, 2006). Studies using difference Fourier-transform

Table 1

Macromolecule-production information.

	Asp133pNO ₂ F sfGFP	Asn149pNO ₂ F sfGFP
Source organism	<i>Aequorea victoria</i>	<i>Aequorea victoria</i>
Expression vectors†	pBAD_sfGFP_D133TAG and pDULE_pNO ₂ F	pBAD_sfGFP_N149TAG and pDULE_pNO ₂ F
Expression host	<i>E. coli</i> DH10B	<i>E. coli</i> DH10B
Complete amino-acid sequence of the construct produced‡	MVSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKL TLKFICTTGKLPVWPPTLVTTT (CRO) VQCFSTRYPDHMKRH DFFKSAMPEGYVQERTISFKDDGYTKTRAEVKFEGDTLVNR IELKGIDFKE (PPN) GNILGHKLEYNFSHNYYITADKQKN GIKANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNH LSTQSVLSKDPNEKRDHMLVLEFVTAAGITHGMDELYKGSH HHHHH	MVSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKL TLKFICTTGKLPVWPPTLVTTT (CRO) VQCFSTRYPDHMKRH DFFKSAMPEGYVQERTISFKDDGYTKTRAEVKFEGDTLVNR IELKGIDFKEDGNILGHKLEYNFSH (PPN) VYITADKQKN GIKANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNH LSTQSVLSKDPNEKRDHMLVLEFVTAAGITHGMDELYKGSH HHHHH

† Expression vectors have been described in previous publications (Smith *et al.*, 2011; Dippel *et al.*, 2016). ‡ These are the full construct sequences that were expressed and purified. Before crystallization the protein was treated with trypsin, removing the last eight amino acids (GSHHHHHH). The second amino acid in the sequence, valine, is an added amino acid resulting from the cloning vector used. We have numbered the residues to be consistent with previous crystal structures, resulting in the first methionine being residue 0 and the valine residue 1. CRO refers to the post-translationally formed chromophore consisting of residues Thr65-Tyr66-Gly67. PPN refers to the translationally incorporated 4-nitro-L-phenylalanine (pNO₂F) residue.

infrared (FTIR) spectroscopy with isotopically labeled p¹⁵NO₂F-containing and naturally abundant p¹⁴NO₂F-containing superfolder green fluorescent protein (sfGFP) allowed the elimination of vibrational absorptions other than the nitro group and allowed the location of the nitro symmetric stretching frequency to be observed (Smith *et al.*, 2011) for both the Asp133pNO₂F and Asn149pNO₂F sfGFP constructs. Once identified, the specific nitro symmetric stretching frequency was correlated with reference spectra of p¹⁴NO₂F in aqueous solvent or tetrahydrofuran (THF), which mimicked solvated and buried sites on proteins, respectively. These studies were completed on Asp133pNO₂F sfGFP and Asn149pNO₂F sfGFP (or sites 134 and 150 according to the numbering used in Smith *et al.*, 2011) and correlate with the structures that we report here.

While crystal structures have been reported for some spectroscopic reporter UAAs genetically incorporated into proteins (Tookmanian *et al.*, 2015; Dippel *et al.*, 2016; Kearney *et al.*, 2018), to date no structures that contain a genetically incorporated pNO₂F residue have been reported in the Protein Data Bank (PDB; Berman *et al.*, 2000). Earlier work with another vibrational reporter, 4-cyano-L-phenylalanine (pCNF; Fig. 1), illustrated that the incorporation of pCNF had a minimal impact on the structures of either GFP or heme

nitric oxide and/or oxygen-binding protein (H-NOX; Dippel *et al.*, 2016; Kearney *et al.*, 2018). The lack of significant UAA-induced structural perturbation is foundational to the effectiveness of a specific UAA as a tool to study protein structure. Here, we report the crystal structures of Asp133pNO₂F sfGFP and Asn149pNO₂F sfGFP, the first structures in the PDB with the UAA pNO₂F genetically incorporated into a protein *via* the amber codon-suppression methodology.

2. Materials and methods

2.1. Macromolecule production

The preparation of pBAD plasmids containing the gene for superfolder GFP (sfGFP) with a TAG site incorporated at position 133 or 149 has been described previously (Smith *et al.*, 2011; Dippel *et al.*, 2016) and the pDULE_pNO₂F plasmid containing the engineered, orthogonal tRNA synthetase selected for pNO₂F was used as described previously (generated by Ryan Mehl, Oregon State University; Smith *et al.*, 2011; Table 1). Both plasmids were co-transformed into *Escherichia coli* DH10B cells. Starter cultures of 5 ml of non-inducing medium were inoculated with the co-transformed cells and were grown overnight at 37°C with shaking at 250 rev min⁻¹. After 15–20 h, 250 ml of autoinduction medium was inoculated with 2.5 ml starter culture and pNO₂F (Peptech) was added to a final concentration of 1 mM (Studier, 2005; Hammill *et al.*, 2007). Following incubation at 37°C with shaking at 250 rev min⁻¹ for 24–30 h, the cells were collected using centrifugation and stored at –80°C. To obtain the purified protein, the cells were thawed on ice, sonicated and the supernatant was added to TALON resin similar to previous procedures (Miyake-Stoner *et al.*, 2010; Bazewicz *et al.*, 2012; Tookmanian, Phillips-Piro *et al.*, 2015; Dippel *et al.*, 2016) to obtain the histidine-tagged construct (Table 1). Following elution from the TALON resin, the protein was desalted over a PD10 column into 20 mM HEPES pH 7.5 and the histidine tag was removed by adding trypsin protease as described previously (Pédrelacq *et al.*, 2006; Dippel *et al.*, 2016). Following the removal of the histidine tag, the protein was

Figure 1

The structures of two spectroscopic reporter UAAs: 4-nitro-L-phenylalanine (pNO₂F) and 4-cyano-L-phenylalanine (pCNF).

Table 2
Crystallization.

	Asp133pNO ₂ F sfGFP	Asn149pNO ₂ F sfGFP
Method	Sitting-drop vapor diffusion	Sitting-drop vapor diffusion
Plate type	Crychem sitting-drop plate	Crychem sitting-drop plate
Temperature	Room temperature (~296 K)	Room temperature (~296 K)
Protein concentration (mg ml ⁻¹)	36	42
Buffer composition of protein solution	20 mM HEPES pH 7.5	20 mM HEPES pH 7.5
Composition of reservoir solution	0.2 M ammonium citrate tribasic pH 7.0, 18% PEG 3350	1.2 M malic acid
Volume and ratio of drop	1 µl protein solution and 1 µl reservoir solution	1 µl protein solution and 1 µl reservoir solution
Volume of reservoir (µl)	500	500

Table 3
Data-collection and processing statistics for Asp133pNO₂F sfGFP and Asn149pNO₂F sfGFP.

Values in parentheses are for the outer shell.

	Asp133pNO ₂ F sfGFP	Asn149pNO ₂ F sfGFP
Diffraction source	24-ID-C, APS	24-ID-E, APS
Wavelength (Å)	0.979	0.979
Temperature (K)	100	100
Detector	Pilatus 6MF	Quantum 315 CCD
Crystal-to-detector distance (mm)	300	200
Rotation range per image (°)	0.2	1.0
Total rotation range (°)	120	180
Exposure time per image (s)	0.2	1.0
Space group	<i>P</i> ₃ ₂ ₁	<i>P</i> ₄ ₁ ₂
<i>a</i> , <i>b</i> , <i>c</i> (Å)	97.72, 97.72, 154.87	55.30, 55.30, 166.53
α , β , γ (°)	90, 90, 120	90, 90, 90
Mosaicity (°)	0.30	0.15
Resolution range (Å)	50–2.05 (2.09–2.05)	50–1.60 (1.63–1.60)
No. of unique reflections	53525 (2616)	35157 (1706)
Completeness (%)	98.5 (97.7)	99.4 (98.8)
Multiplicity	6.6 (6.2)	13.9 (13.6)
$\langle I/\sigma(I) \rangle$	21.36 (1.56)†	36.83 (2.33)
<i>CC</i> _{1/2}	99.5 (74.5)	99.3 (82.0)
<i>R</i> _{int}	3.2 (41.1)	2.8 (30.8)
Overall <i>B</i> factor from Wilson plot (Å ²)	37.45	21.13

† The decision to cut the data at 2.05 Å resolution was made because of the high *CC*_{1/2}, multiplicity and completeness even though $\langle I/\sigma(I) \rangle < 2$ at this resolution. The data were collected on a PILATUS 6MF detector and each frame resulted from a 0.2 s exposure, resulting in lower intensities but also less environmental noise. $\langle I/\sigma(I) \rangle > 2$ at resolutions below 2.12 Å. ‡ *R*_{int} is the redundancy-independent merging *R* factor.

concentrated and stored in 20 mM HEPES pH 7.5 buffer. Both protein constructs had a vibrant green colour resulting from a properly formed chromophore.

2.2. Crystallization

The purified pNO₂F mutants of sfGFP were concentrated to between 35 and 45 mg ml⁻¹ and screened against a number of crystal screens in 96-well sitting-drop crystal trays. These initial crystallization screens contained 50 µl of a precipitation condition per reservoir and the sitting drops consisted of 1 µl protein solution and 1 µl precipitation solution. Initial crystallization hits were optimized in Crychem sitting-drop plates and the final crystallization conditions are summarized in Table 2.

2.3. Data collection and processing

X-ray diffraction data were collected on the Northeastern Collaborative Access Team (NE-CAT) beamlines (either

24-ID-C or 24-ID-E) at the Advanced Photon Source (APS), Argonne National Laboratory (Table 3). Data for Asp133pNO₂F sfGFP were processed in space group *P*₃₂₁ to 2.05 Å resolution using *HKL*-2000 (Otwinowski & Minor, 1997). The data for Asn149pNO₂F sfGFP were processed in space group *P*₄₁₂ to 1.60 Å resolution using *HKL*-2000 (Otwinowski & Minor, 1997).

2.4. Structure solution and refinement

Initial molecular replacement was completed using the wild-type sfGFP model (PDB entry 2b3p; Pédelacq *et al.*, 2006) with solvent molecules removed, *B* factors set to 20 Å² and the residue at the location of pNO₂F incorporation mutated to an alanine (D133A or N149A). Molecular replacement was completed using *Phaser* (Storoni *et al.*, 2004; McCoy *et al.*, 2007) in *PHENIX* (Adams *et al.*, 2010) for both structures. Molecular replacement with the Asp133pNO₂F sfGFP data identified two molecules in the asymmetric unit with a translation-function *Z*-score (TFZ score) of 7777.4 and a log-likelihood gain (LLG) of 88.4. By contrast, the Asn149pNO₂F sfGFP structure contained only one molecule in the asymmetric unit, and the TFZ score and LLG following molecular replacement in *Phaser* were 5381.9 and 71.7, respectively.

Following molecular replacement, the Asp133pNO₂F sfGFP structure was automatically refined in *PHENIX* (Adams *et al.*, 2010) alternating with manual refinement in *Coot* (Emsley & Cowtan, 2004). The initial 2*F*_o – *F*_c electron density and *F*_o – *F*_c electron density illustrated the presence of the pNO₂F residue (Fig. 2*b*). After the replacement of D133A with the pNO₂F residue and further refinement, water molecules, a sodium ion and ethylene glycol molecules were added, resulting in a final *R* and *R*_{free} of 18.50 and 21.89%, respectively (Table 4, Figs. 2*a* and 2*c*).

The first round of refinement in *PHENIX* (Adams *et al.*, 2010) following molecular replacement with the N149A sfGFP model resulted in clear electron density for the pNO₂F residue (Fig. 3*b*). The pNO₂F residue was added to this site, replacing N149A, and the refinement continued, alternating automatic refinement in *PHENIX* and manual refinement in *Coot* (Emsley & Cowtan, 2004). The parameter files for the chromophore (ligand ID CRO) and pNO₂F (ligand ID PPN) were refined from the starting parameters available in the Protein Data Bank in *Grade* (Smart *et al.*, 2011). Following initial adjustments to the backbone and side chains, water molecules,

Table 4

Structure solution and refinement of the Asp133pNO₂F sfGFP and Asn149pNO₂F sfGFP crystal structures.

	Asp133pNO ₂ F sfGFP	Asn149pNO ₂ F sfGFP
Resolution range (Å)	50–2.05	50–1.60
Final R_{cryst} (%)	18.80	16.05
Final R_{free} (5% of data) (%)	21.89	19.05
R.m.s. deviations		
Bonds (Å)	0.010	0.009
Angles (°)	1.026	1.005
No. of atoms		
Protein	3685	1931
4-Nitro-L-phenylalanine (PPN)	28	14
Sodium ion(s)	1	1
Ethylene glycol	16	
Water	313	272
Overall	4015	2204
Average B factors for non-H atoms (Å ²)		
Protein	43.37	23.88
4-Nitro-L-phenylalanine (PPN)	60.96	19.41
Sodium ion(s)	45.10	31.63
Ethylene glycol	54.16	
Water	49.32	34.89
Overall	43.87	25.24
Ramachandran plot		
Most favored (%)	97.05	99.11
Allowed (%)	2.95	0.89

a sodium ion and eventually riding H atoms were added to the 1.60 Å resolution structure. The final R and R_{free} for the

Asn149pNO₂F sfGFP structure were 16.05 and 19.05%, respectively (Table 4, Figs. 3*a* and 3*c*).

3. Results and discussion

The initial electron densities in the areas of the D133A or N149A sites in the sfGFP crystal structures reported here illustrate that the 4-nitro-L-phenylalanine was successfully site-specifically incorporated into sfGFP (Figs. 2*b* and 3*b*). In addition, the electron density for the refined pNO₂F residues incorporated at either the Asp133 or Asn149 sites illustrate good $2F_o - F_c$ electron density for the full residue including the NO₂ group, suggesting that the pNO₂F did not undergo photodegradation during protein expression, purification, crystallization and data collection (Figs. 2*c* and 3*c*).

To accurately report on the local environment and/or dynamics in a protein, a reporter UAA must not cause any significant structural perturbation. The Asp133pNO₂F sfGFP structure was aligned with both wild-type sfGFP [root-mean-square deviation (r.m.s.d.) of 0.229 Å for 195 C $^{\alpha}$ atoms] and Asp133pCNF sfGFP (PDB entry 5dpg; r.m.s.d. of 0.189 Å for 192 C $^{\alpha}$ atoms; Dippel *et al.*, 2016), illustrating minimal structural alteration upon pNO₂F incorporation at the Asp133 site (Fig. 4*a*). The phenyl ring in Asp133pNO₂F sfGFP is in the same orientation as the phenyl ring in the Asp133pCNF sfGFP crystal structure, rotated approximately 180° from the side

Figure 2

Crystal structure of Asp133pNO₂F sfGFP (PDB entry 6dq0). The overall structure is shown in cartoon representation in (a) with the Asp133pNO₂F residue and chromophore shown in stick representation. The initial $2F_o - F_c$ map at 1.0σ and $F_o - F_c$ map at $+3.0\sigma$ and -3.0σ around the D133A site are shown in (b) in blue, green and red mesh, respectively. (c) The final $2F_o - F_c$ map at 1.0σ is shown in blue mesh around the pNO₂F modeled at the Asp133 site. Figures were generated in PyMOL (Schrödinger).

chain of Asp133 in wild-type sfGFP (Fig. 4a). This side-chain rotation is not concerning as the backbone aligns well with the wild-type sfGFP Asp133 residue and its location on a solvent-

accessible loop suggests that in solution any residue at site 133 (including Asp133 in wild-type sfGFP) is likely to be experiencing a number of solvent-exposed conformations. Finally,

Figure 3

Crystal structure of Asn149pNO₂F sfGFP (PDB entry 6dq1). The overall structure is shown in cartoon representation in (a) with the Asn149pNO₂F residue and chromophore shown in stick representation. The initial $2F_o - F_c$ map at 1.0σ and $F_o - F_c$ map at $+3.0\sigma$ and -3.0σ around the N149A site are shown in (b) in blue, green and red mesh, respectively. (c) The final $2F_o - F_c$ map at 1.0σ is shown in blue mesh around the pNO₂F modeled at the Asn149 site. Figures were generated in *PyMOL* (Schrödinger).

Figure 4

Structure alignments of pNO₂F-containing crystal structures with wild-type sfGFP and the corresponding pCNF-containing crystal structure. In (a) the Asp133pNO₂F sfGFP structure (PDB entry 6dq0) is shown in green with the Asp133pCNF sfGFP structure (PDB entry 5dpg) in orange and wild-type sfGFP (PDB entry 2b3p) in gray. The residues at the Asp133 site for each of these structures are shown in stick representation and are colored by atom type. In (b) the Asn149pNO₂F sfGFP structure (PDB entry 6dq1) is shown in pink with the Asn149pCNF sfGFP structure (PDB entry 5dph) in teal and wild-type sfGFP (PDB entry 2b3p) in gray. The residues at the Asn149 site in each of these structures are shown in stick representation and are coloured by atom type. Figures were generated in *PyMOL* (Schrödinger).

the location of Asp133pNO₂F is consistent with previously published IR results suggesting that this site is solvated (Smith *et al.*, 2011).

The Asn149pNO₂F sfGFP structure was aligned with wild-type sfGFP (r.m.s.d. of 0.180 Å for 190 C^α atoms) and Asn149pCNF sfGFP (PDB entry 5dph; r.m.s.d. of 0.218 Å for 198 C^α atoms; Fig. 4b). These alignments illustrated that the backbone, C^β, C^γ and δ position atoms overlaid well for the Asn149pNO₂F, Ala149pCNF and Asn149 residues (Fig. 4b). The Asn149pNO₂F residue is partially buried from bulk solvent by neighbouring side chains along the side of the β-barrel, which is consistent with the observed IR frequency for the nitro stretching frequency reported for Asn149pNO₂F sfGFP (Smith *et al.*, 2011).

The two structures reported here represent the first reported X-ray crystal structures with pNO₂F incorporated into a protein structure using the amber codon-suppression technology. Previous structures containing the pNO₂F residue contained pNO₂F on a synthesized short peptide fragment (4–8 residues; Rose *et al.*, 1996; Lizak *et al.*, 2011; Aleem *et al.*, 2016; Napiórkowska *et al.*, 2017) or a semi-synthetic protein in which the pNO₂F was synthesized in a peptide that was later ligated to an expressed protein fragment (Baril *et al.*, 2017). The Asp133pNO₂F sfGFP and Asn149pNO₂F sfGFP crystal structures illustrate clear density for the pNO₂F residue at the specified position, indicating that pNO₂F can be site-specifically incorporated to produce a stable pNO₂F group. Further, these structures illustrate that the incorporation of pNO₂F at the Asp133 or Asn149 sites in sfGFP does not cause a significant structural change, supporting the use of pNO₂F as a spectroscopic probe for various protein systems.

Acknowledgements

We thank Lisa Mertzman for obtaining supplies and Ryan Mehl (Oregon State University) for the pDULE and pBAD vectors.

Funding information

The following funding is acknowledged: Camille and Henry Dreyfus Foundation (award No. TH-15-009 to Scott H. Brewer). Data for both structures were collected at the Northeastern Collaborative Access Team (NE-CAT) beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P41 GM103403). The PILATUS 6MF detector on the 24-ID-C beamline is funded by grant S10RR029205 from the National Center of Research Resources of the National Institutes of Health. This research used resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.

- Adhikary, R., Zimmermann, J., Dawson, P. E. & Romesberg, F. E. (2014). *ChemPhysChem*, **15**, 849–853.
- Aleem, S., Georgiou, G., Kleiner, R. E., Guja, K., Craddock, B. P., Lyczek, A., Chan, A. I., Garcia-Diaz, M., Miller, W. T., Liu, D. R. & Seeliger, M. A. (2016). *Cell Chem. Biol.* **23**, 1103–1112.
- Baril, S. A., Koenig, A. L., Krone, M. W., Albanese, K. I., He, C. Q., Lee, G. Y., Houk, K. N., Waters, M. L. & Brustad, E. M. (2017). *J. Am. Chem. Soc.* **139**, 17253–17256.
- Bazewicz, C. G., Lipkin, J. S., Smith, E. E., Liskov, M. T. & Brewer, S. H. (2012). *J. Phys. Chem. B*, **116**, 10824–10831.
- Bazewicz, C. G., Liskov, M. T., Hines, K. J. & Brewer, S. H. (2013). *J. Phys. Chem. B*, **117**, 8987–8993.
- 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.
- Cellitti, S. E. *et al.* (2008). *J. Am. Chem. Soc.* **130**, 9268–9281.
- Dippel, A. B., Olenginski, G. M., Maurici, N., Liskov, M. T., Brewer, S. H. & Phillips-Piro, C. M. (2016). *Acta Cryst.* **D72**, 121–130.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Hammill, J. T., Miyake-Stoner, S., Hazen, J. L., Jackson, J. C. & Mehl, R. A. (2007). *Nature Protoc.* **2**, 2601–2607.
- Hartley, A. M., Worthy, H. L., Reddington, S. C., Rizkallah, P. J. & Jones, D. D. (2016). *Chem. Sci.* **7**, 6484–6491.
- Jackson, J. C., Hammill, J. T. & Mehl, R. A. (2007). *J. Am. Chem. Soc.* **129**, 1160–1166.
- Kearney, C., Olenginski, L. T., Hirn, T. D., Fowler, G. D., Tariq, D., Brewer, S. H. & Phillips-Piro, C. M. (2018). *RSC Adv.* **8**, 13503–13512.
- Lizak, C., Gerber, S., Numao, S., Aebi, M. & Locher, K. P. (2011). *Nature (London)*, **474**, 350–355.
- Ma, J., Pazos, I. M., Zhang, W., Culik, R. M. & Gai, F. (2015). *Annu. Rev. Phys. Chem.* **66**, 357–377.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Miyake-Stoner, S. J., Refakis, C. A., Hammill, J. T., Lusich, H., Hazen, J. L., Deiters, A. & Mehl, R. A. (2010). *Biochemistry*, **49**, 1667–1677.
- Napiórkowska, M., Boilevin, J., Sovdat, T., Darbre, T., Raymond, J.-L., Aebi, M. & Locher, K. P. (2017). *Nature Struct. Mol. Biol.* **24**, 1100–1106.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pédelaq, J.-D., Cabantous, S., Tran, T., Terwilliger, T. C. & Waldo, G. S. (2006). *Nature Biotechnol.* **24**, 79–88.
- Rose, R. B., Craik, C. S., Douglas, N. L. & Stroud, R. M. (1996). *Biochemistry*, **35**, 12933–12944.
- Schultz, K. C., Supekova, L., Ryu, Y., Xie, J., Perera, R. & Schultz, P. G. (2006). *J. Am. Chem. Soc.* **128**, 13984–13985.
- Slocum, J. D. & Webb, L. J. (2016). *J. Am. Chem. Soc.* **138**, 6561–6570.
- Smart, O. S., Womack, T. O., Sharff, A., Flensburg, C., Keller, P., Paciorek, W., Vornrhein, C. & Bricogne, G. (2011). *Grade*, v1.1.1. Global Phasing Ltd, Cambridge, UK.
- Smith, E. E., Linderman, B. Y., Luskin, A. C. & Brewer, S. H. (2011). *J. Phys. Chem. B*, **115**, 2380–2385.
- Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* **D60**, 432–438.
- Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.
- Taki, M., Hoshaka, T., Murakami, H., Taira, K. & Sisido, M. (2002). *J. Am. Chem. Soc.* **124**, 14586–14590.
- Thielges, M. C., Axup, J. Y., Wong, D., Lee, H. S., Chung, J. K., Schultz, P. G. & Fayer, M. D. (2011). *J. Phys. Chem. B*, **115**, 11294–11304.
- Tookmanian, E. M., Fenlon, E. E. & Brewer, S. H. (2015). *RSC Adv.* **5**, 1274–1281.
- Tookmanian, E. M., Phillips-Piro, C. M., Fenlon, E. E. & Brewer, S. H. (2015). *Chemistry*, **21**, 19096–19103.
- Tsao, M.-L., Summerer, D., Ryu, Y. & Schultz, P. G. (2006). *J. Am. Chem. Soc.* **128**, 4572–4573.
- Ye, S., Zaitseva, E., Caltabiano, G., Schertler, G. F. X., Sakmar, T. P., Deupi, X. & Vogel, R. (2010). *Nature (London)*, **464**, 1386–1389.