Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

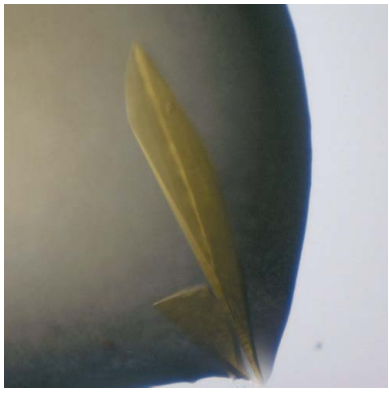
Expression, purification, crystallization and preliminary X-ray analysis of eCGP123, an extremely stable monomeric green fluorescent protein with reversible photoswitching properties

Craig Don Paul,[a] Daouda A. K. Traore,[b] Emma Byres,[b] Jamie Rossjohn,[a] Rodney J. Devenish,[a] Csaba Kiss,[c] Andrew Bradbury,[c] Matthew C. J. Wilce[b]* and Mark Prescott[a]*

[a]Department of Biochemistry and Molecular Biology, School of Biomedical Science, Monash University, Clayton Campus, Melbourne, Victoria 3800, Australia, [b]The Protein Crystallography Unit, School of Biomedical Sciences, Monash University, Clayton Campus, Melbourne, Victoria 3800, Australia, and [c]Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

Correspondence e-mail: matthew.wilce@monash.edu, mark.prescott@monash.edu

Received 25 May 2011
Accepted 13 July 2011



Enhanced consensus green protein variant 123 (eCGP123) is an extremely thermostable green fluorescent protein (GFP) that exhibits useful negative reversible photoswitching properties. eCGP123 was derived by the application of both a consensus engineering approach and a recursive evolutionary process. Diffraction-quality crystals of recombinant eCGP123 were obtained by the hanging-drop vapour-diffusion method using PEG 3350 as the precipitant. The eCGP123 crystal diffracted X-rays to 2.10 Å resolution. The data were indexed in space group $P1$, with unit-cell parameters a = 74.63, b = 75.38, c = 84.51 Å, α = 90.96, β = 89.92, γ = 104.03°. The Matthews coefficient (V_M = 2.26 Å^3 Da^{-1}) and a solvent content of 46% indicated that the asymmetric unit contained eight eCGP123 molecules.

1. Introduction

*av*GFP cloned from the bioluminescent jellyfish *Aequorea victoria* (Chalfie *et al.*, 1994) represents the first identified member of a GFP-like family of proteins that have now become established as indispensable tools for use in a myriad of molecular cell-biology applications (Tsien, 1998; Griesbeck *et al.*, 2001; Mazzola *et al.*, 2006; Ishii *et al.*, 2007; Stepanenko *et al.*, 2008). GFP-like proteins consist of an 11-stranded β-barrel threaded by an α-helix from which the chromophore is suspended (Ormö *et al.*, 1996). The 'β-can' structure of these proteins serves to promote the post-translational formation of the tripeptide chromophore and forms a matrix of interactions with the chromophore, resulting in their many varied and useful optical properties, one of which is photoswitching (Lukyanov *et al.*, 2005).

Considerable effort has been invested in the search for variants with new and improved optical properties. As a result of extensive engineering, aided by structural data, variants are available with fluorescence emissions that collectively cover the visible spectrum, extending into the near-infrared (Olenych *et al.*, 2007). Other variants are able to undergo reversible or irreversible photoswitching and are useful for super-resolution microscopy applications. Many GFP-like proteins serve as platforms for the engineering of new biosensors for monitoring key cellular parameters.

In order to further broaden the applications of GFP-like proteins, approaches have been adopted to engineer variants that have significantly increased thermal stability (Pédelacq *et al.*, 2006; Dai *et al.*, 2007; Cava *et al.*, 2008; Kiss *et al.*, 2009). A recent study reported that the popular enhanced GFP variant was nonfluorescent when expressed in the extremophile *Thermus thermophilus*, whereas the stable superfolder GFP was highly fluorescent (Cava *et al.*, 2008). Such proteins represent structural platforms for the development of probes for use in harsh industrial environments or as rigid scaffolding for grafting new functionalities in order to generate novel biosensors (Belousov *et al.*, 2006; Souslova *et al.*, 2007).

eCGP123 was developed as part of a separately published study using a two-step process (Kiss *et al.*, 2009). Firstly, using the green fluorescent protein monomeric Azami-Green (mAG; GenBank accession No. AB108447; Karasawa *et al.*, 2003) as a guide, a consensus green protein (CGP) was designed by aligning the amino-acid sequences of 31 different fluorescent proteins (Dai *et al.*, 2007). Subsequently, CGP was subjected to an iterative program of evolution, resulting in an extremely thermostable variant eCGP123, the

```
                1         10        20        30        40        50        60
mAG     M- - - - SVIKPEMKIKLCMRGTVNGHNFVIEGEGKGNPYEGTQILDLNVTEGAPLPFAYDI
eCGP123 MGAHASVIKPEMKIKLRMEGAVNGHKFVIEGEGIGKPYEGTQTLDLTVKEGAPLPFSYDI

               61        70        80        90       100       110       120
mAG     LTTVFQYGNRAFTKYPADIQDYFKQTFPEGYHWERSMTYEDQGICIATSNISMRGDCFFY
eCGP123 LTPAFQYGNRAFTKYPKDIPDYFKQAFPEGYSWERSMTYEDQGICIATSDITMEGDCFFY

              121       130       140       150       160       170       180
mAG     DIRFDGTNFPPNGPVMQKKTLKWEPSTEKMYVRDGVLKGDVNMALLLEGGGHYRCDFKTT
eCGP123 KIRFDGTNFPPNGPVMQKKTLKWEPSTEKMYVRDGVLKGDVNMALLLEGGGHYRCDFKTT

              181       190       200       210       220       230       240
mAG     YKAKKDVRLPDAHKVDHRIEILKHDKDYNKVKLYENAVARYSMLPSQAKASGKPIPNPLLGL
eCGP123 YKAKKDVRLPDAHEVDHRIEILSHDKDYNKVRLYEHAEARYSMLPSQAKASGKPIPNPLLGL

                250
mAG     DST
eCGP123 DSTHHHHHH
```

Figure 1
Amino-acid sequence alignment of eCGP123 with mAG. The amino-acid sequences of eCGP123 and mAG are shown aligned. Shading indicates residues that differ between the two sequences. eCGP123 is shown here with the C-terminal 6×His tag introduced to facilitate purification; percentage identities were calculated for eCGP123 without the tag. This figure was generated by *Kalign* (Lassmann & Sonnhammer, 2005).

fluorescence emission properties of which remain intact after exposure to high temperature for extended periods. Compared with both mAG and CGP, which became nonfluorescent after 14 h incubation at 353 K, eCGP123 retained 80% of its original fluorescence emission (Kiss *et al.*, 2009). An amino-acid sequence alignment of mAG and eCGP123, highlighting differences between the proteins, is shown in Fig. 1. Unlike mAG, the protein that was used to guide the consensus engineering, eCGP123 possesses useful reversible photoswitching capabilities similar to those of Dronpa, a well characterized photoswitching fluorescent protein for which the X-ray crystal structure is known (Wilmann *et al.*, 2006). Here, we describe the production, purification, crystallization and preliminary X-ray analysis of eCGP123.

2. Materials and methods

2.1. Production and purification of eCGP123

An expression vector encoding eCGP123 (pETCK3:eCGP123) with a noncleavable C-terminal 6×His tag has been described previously (Kiss *et al.*, 2009). Freshly transformed *Escherichia coli* Nova Blue DE3 cells (Novagen) were cultured overnight in flasks containing 10 ml LB medium at 310 K with shaking in an orbital shaker. The following day, cell cultures were diluted to an OD_{600} of 0.6 in 250 ml fresh growth medium contained in flasks and protein expression was induced for 6 h at 310 K with shaking by the addition of IPTG to a final concentration of 0.5 m*M*. Cells were harvested by centrifugation (8 000*g*, 15 min, 277 K), resuspended in 40 ml 20 m*M* Tris–HCl pH 8.0 and 300 m*M* NaCl (buffer *A*) and disrupted by sonication on ice using a Soniprep 150 sonicator equipped with a 4 mm diameter probe (Sanyo) with eight cycles consisting of 30 s sonication followed by 30 s rest. Lysates were centrifuged (10 000*g*, 45 min, 277 K) and the supernatant was used to saturate an Ni–NTA column (2 ml; Qiagen) equilibrated with buffer *A*. The column was washed with ten column volumes of buffer *A* containing 10 m*M* imidazole and eCGP123 protein was eluted with buffer *A* containing 500 m*M* imidazole. Fractions containing eCGP123 protein were pooled and loaded onto a HiLoad 16/60 Superdex 200 gel-filtration column (GE Healthcare) equilibrated with buffer *A* and calibrated in a separate run using native molecular-mass marker proteins. Fractions containing eCGP123 were pooled and concentrated to ∼10 mg ml^{-1} using centrifugal concentrators (10 kDa cutoff; Amicon Ultra, Millipore). Purified eCGP123 protein was used for crystallization experiments without removal of the 6×His tag.

2.2. Crystallization, data collection and preliminary X-ray analysis

Crystallization screens were performed using commercially available kits including JCSG+ (Molecular Dimensions), PEG/Ion (Hampton Research), Basic Crystallography Kit (Sigma–Aldrich), Index (Hampton Research) and PACT (Molecular Dimensions). Trials were performed at 293 K using the sitting-drop vapour-diffusion method. 100 nl each of eCGP123 protein solution (10 mg ml^{-1} in buffer *A*) and reservoir solution were dispensed into 96-well Intelli-Plates (Art Robbins) using a Honeybee liquid-dispensing system (Digilab Inc.). Plate-like crystals formed after approximately 1–2 d under several different conditions including (i) 0.1 *M* Bis-Tris pH 5.5, 25%(*w*/*v*) PEG 3350 and 0.2 *M* NaCl; (ii) 0.1 *M* Tris–HCl pH 8.0 and 5%(*w*/*v*) PEG 600; and (iii) 0.1 *M* HEPES pH 7.5, 25%(*w*/*v*) PEG 3350 and 0.2 *M* NaCl. All crystallization experiments were performed at 293 K in a purpose-designed temperature-controlled facility.

The results of these initial screens were used to design experiments for further optimization of both the precipitant concentration [15–27.5%(*w*/*v*) PEG 3350] and the pH (Tris–HCl pH 7–8.5). Initial pH values for the reservoir buffer were selected to be above the pK_a of the chromophore (pH 6.0). These trials were performed using the hanging-drop vapour-diffusion method. Equal volumes of protein solution (10 mg ml^{-1} in buffer *A*) and reservoir buffer were

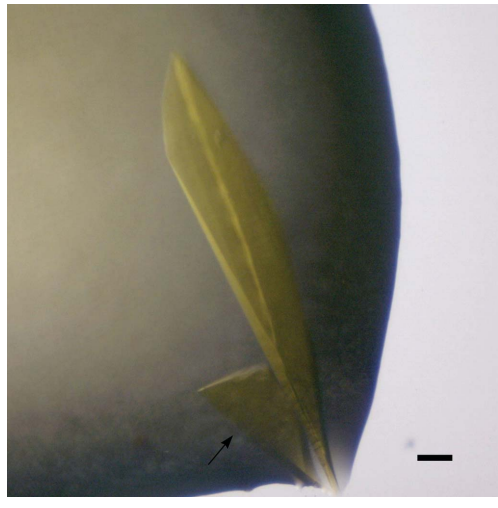


Figure 2
Diffraction-quality crystals of eCGP123. The arrow indicates the position of the single crystal used for diffraction studies. The scale bar represents 1 mm.

Table 1
Summary of data-collection statistics for eCGP123.

Values in parentheses are for the highest resolution shell.

No. of crystals	1
X-ray source	Rigaku RU-3HBR rotating anode
Temperature (K)	100
Wavelength (Å)	1.542
Detector	Rigaku R-AXIS IV^{++}
Crystal-to-detector distance (mm)	166.2
Rotation range per image (°)	1
Total rotation range (°)	360
Exposure time per image (s)	120
Resolution range (Å)	29.85–2.10 (2.21–2.10)
Space group	*P*1
Unit-cell parameters (Å, °)	a = 74.6, b = 75.4, c = 84.5, α = 91.0, β = 89.8, γ = 104.0
Total No. of measured intensities	195400
Unique reflections	98461 (14033)
Multiplicity	2.0 (2.0)
Mean $I/\sigma(I)$	16.6 (2.3)
Completeness (%)	94.4 (91.9)
$R_{\rm merge}$† (%)	0.042 (0.292)
$R_{\rm meas}$ (%)	0.059 (0.412)
$R_{\rm p.i.m.}$ (%)	0.042 (0.292)
Overall B factor from Wilson plot (Å^2)	30.68

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

combined in a final volume of 2 µl. VDXm plates (Hampton Research) were used containing 0.5 ml reservoir solution in each well.

The results of other experiments designed to screen for the effects of various salts, including NaCl, $MgCl_2$, KI and sodium acetate, indicated that substituting NaCl for $MgCl_2$ significantly improved the quality of the crystals and produced crystals with smoother and better defined faces. The best diffracting crystals were obtained with a reservoir composition of 0.2 *M* Tris–HCl pH 8.0, 22.5%(*w*/*v*) PEG 3350 and 0.2 *M* $MgCl_2$. Crystals of eCGP123 appeared green/yellow under white light (Fig. 2) and highly green-fluorescent when viewed under a fluorescence microscope.

Selected crystals were mounted in loops and dipped in the cryoprotectant perfluoropolyether oil (PFO-X175/08, Hampton Research) for 1 min before vitrification in a nitrogen-gas stream maintained at 100 K. Diffraction data were collected in-house using a Rigaku R-AXIS IV^{++} detector. 360 images were collected using a wavelength of 1.542 Å, a 1° oscillation range and an exposure time of 120 s per image. The data were indexed and scaled using the *HKL*-2000 program package (Otwinowski & Minor, 1997). eCGP123 crystals retained their green/yellow colour after irradiation; when redissolved, they showed optical spectra identical to those of the starting protein preparation (data not shown).

3. Results and discussion

After the lysis of cells expressing eCGP123 and centrifugation, green-coloured protein was quantitatively released into the supernatant, indicating that the protein was soluble. Protein eluted from Ni–NTA chromatography medium was applied onto a calibrated HiLoad 16/60 Superdex 200 gel-filtration column equilibrated with buffer *A*. Most of the eCGP123 protein eluted as a single species at a position that corresponded closely to the expected size of the intact expressed protein (28 569 Da). This result indicated that the protein behaves as a monomer in solution, as previously reported (Kiss *et al.*, 2009). The protein was concentrated to 10 mg ml^{-1} in buffer *A* and used in crystallization trials without removal of the C-terminal 6×His tag.

Green/yellow-coloured crystals of eCGP123 were obtained by the hanging-drop vapour-diffusion method and typically grew to final dimensions of 0.3 × 0.4 × 0.04 mm within 1.5 d at 293 K (Fig. 2). The best crystal diffracted X-rays to a resolution of 2.10 Å. The data-collection statistics are summarized in Table 1. The Matthews coefficient ($V_{\rm M}$ = 2.26 Å^3 Da^{-1}; Matthews, 1968) indicated that the asymmetric unit contained eight eCGP123 molecules, with a solvent content of 46%. The structure of eCGP123 will be solved by molecular replacement using the coordinates of a structure of mAG (89% identity excluding the 6×His tag; Fig. 1; PDB entry 3adf; Ebisawa *et al.*, 2009). In other experiments designed to characterize its optical properties, we have established that unlike the protein used to guide its design (mAG), eCGP123 possesses photoswitching properties similar to those of Dronpa (data not shown), a well characterized green photoswitching protein for which the crystal structure is known (Ando *et al.*, 2004; Wilmann *et al.*, 2006; Andresen *et al.*, 2007). eCGP123 has 78% identity to Dronpa (excluding the 6×His tag). A crystal structure of eCGP123 will be important in order to determine the molecular basis for the different properties of this extremely thermostable fluorescent protein compared with those of mAG.

GFP-like proteins have proven to be useful in a wide range of biotechnological applications and form the basis of numerous biosensors. Photoswitching GFP-like proteins such as eCGP123 will guide the design of a new generation of biosensors for studies in extreme environments where existing fluorescent proteins are unsuitable.

References

Ando, R., Mizuno, H. & Miyawaki, A. (2004). *Science*, **306**, 1370–1373.
Andresen, M., Stiel, A. C., Trowitzsch, S., Weber, G., Eggeling, C., Wahl, M. C., Hell, S. W. & Jakobs, S. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 13005–13009.
Belousov, V. V., Fradkov, A. F., Lukyanov, K. A., Staroverov, D. B., Shakhbazov, K. S., Terskikh, A. V. & Lukyanov, S. (2006). *Nature Methods*, **3**, 281–286.
Cava, F., de Pedro, M. A., Blas-Galindo, E., Waldo, G. S., Westblade, L. F. & Berenguer, J. (2008). *Environ. Microbiol.* **10**, 605–613.
Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. (1994). *Science*, **263**, 802–805.
Dai, M., Fisher, H. E., Temirov, J., Kiss, C., Phipps, M. E., Pavlik, P., Werner, J. H. & Bradbury, A. R. (2007). *Protein Eng. Des. Sel.* **20**, 69–79.
Ebisawa, T., Yamamura, A., Kameda, Y., Hayakawa, K., Nagata, K. & Tanokura, M. (2009). *Acta Cryst.* F**65**, 1292–1295.
Griesbeck, O., Baird, G. S., Campbell, R. E., Zacharias, D. A. & Tsien, R. Y. (2001). *J. Biol. Chem.* **276**, 29188–29194.
Ishii, M., Kunimura, J. S., Jeng, H. T., Penna, T. C. & Cholewa, O. (2007). *Appl. Biochem. Biotechnol.* **137–140**, 555–571.
Karasawa, S., Araki, T., Yamamoto-Hino, M. & Miyawaki, A. (2003). *J. Biol. Chem.* **278**, 34167–34171.
Kiss, C., Temirov, J., Chasteen, L., Waldo, G. S. & Bradbury, A. R. (2009). *Protein Eng. Des. Sel.* **22**, 313–323.
Lassmann, T. & Sonnhammer, E. L. (2005). *BMC Bioinformatics*, **6**, 298.
Lukyanov, K. A., Chudakov, D. M., Lukyanov, S. & Verkhusha, V. V. (2005). *Nature Rev. Mol. Cell Biol.* **6**, 885–891.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mazzola, P. G., Ishii, M., Chau, E., Cholewa, O. & Penna, T. C. (2006). *Biotechnol. Prog.* **22**, 1702–1707.
Olenych, S. G., Claxton, N. S., Ottenberg, G. K. & Davidson, M. W. (2007). *Current Protocols in Cell Biology*, ch. 21, unit 21.5. New York: Wiley.
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.
Pédelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C. & Waldo, G. S. (2006). *Nature Biotechnol.* **24**, 79–88.
Souslova, E. A., Belousov, V. V., Lock, J. G., Strömblad, S., Kasparov, S., Bolshakov, A. P., Pinelis, V. G., Labas, Y. A., Lukyanov, S., Mayr, L. M. & Chudakov, D. M. (2007). *BMC Biotechnol.* **7**, 37.
Stepanenko, O. V., Verkhusha, V. V., Kuznetsova, I. M., Uversky, V. N. & Turoverov, K. K. (2008). *Curr. Protein Pept. Sci.* **9**, 338–369.
Tsien, R. Y. (1998). *Annu. Rev. Biochem.* **67**, 509–544.
Wilmann, P. G., Turcic, K., Battad, J. M., Wilce, M. C., Devenish, R. J., Prescott, M. & Rossjohn, J. (2006). *J. Mol. Biol.* **364**, 213–224.