Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary crystallographic analysis of a novel orange fluorescent protein from the *Cnidaria* tube anemone *Cerianthus* sp.

Denis T. M. Ip,[a] Siu-Hong Chan,[a] Mark D. Allen,[b] Mark Bycroft,[b] David C. C. Wan[a] and Kam-Bo Wong[a]*

[a]Department of Biochemistry, The Chinese University of Hong Kong, Hong Kong SAR, People's Republic of China, and [b]MRC Centre for Protein Engineering, Hills Road, Cambridge CB2 2QH, England

Correspondence e-mail: kbwong@cuhk.edu.hk

Received 3 October 2003
Accepted 25 November 2003

A novel orange fluorescent protein, with excitation and emission maxima at 548 and 565 nm, respectively, from the *Cnidaria* tube anemone *Cerianthus* sp. has been cloned and overexpressed in *Escherichia coli*. The orange fluorescent protein has been crystallized by the sitting-drop vapour-diffusion method at 290 K using polyethylene glycol 3350 as a precipitant. A complete set of diffraction data was collected to 2.0 Å resolution at 100 K. The crystals belong to the space group $R3$, with hexagonal unit-cell parameters $a = b = 216.947$, $c = 51.839$ Å. There are four protein molecules in the asymmetric unit, giving a Matthews coefficient of 2.3 Å^3 Da^{-1} and a solvent content of 47%.

1. Introduction

In this paper, we report the expression, purification, crystallization and preliminary X-ray diffraction analysis of a novel orange fluorescent protein (OFP) from the *Cnidaria* tube anemone *Cerianthus* sp. Fluorescent proteins, most notably the green fluorescent protein (avGFP) from the jellyfish *Aequorea victoria* (Prasher *et al.*, 1992), have become a popular tool for studying gene expression, subcellular localization of protein and protein–protein interaction in living cells (Gerdes & Kaether, 1996; Tsien, 1998; Lippincott-Schwartz & Patterson, 2003). Fluorescent proteins that emit a wide spectrum of colours are desirable because they allow multi-colour tracking of several proteins simultaneously and facilitate applications based on fluorescence resonance-energy transfer (Boute *et al.*, 2002). Variants of avGFP that emit blue, cyan and yellow fluorescence are available, but none with emission that extends to the red region of the spectrum (Ehrig *et al.*, 1995; Delagrave *et al.*, 1995). Longer wavelength fluorescence has the advantages of better tissue penetration and better spectral separation from the green autofluorescence observed in living cells. The first native red fluorescent protein DsRed, which has excitation and emission maxima at 558 and 583 nm, respectively, was cloned from the non-bioluminescent *Discosoma* coral (Matz *et al.*, 1999). Recently, we have isolated an orange fluorescent protein from the tentacles of the *Cnidaria* tube anemone *Cerianthus* sp. This protein has excitation and emission maxima at 548 and 565 nm, respectively, which fall nicely within the spectral gap between existing yellow and red fluorescent proteins.

Structure determination of OFP will provide the basis for a better understanding of its biophysical properties and for future protein-engineering efforts. As a first step toward structure determination, we have obtained crystals of OFP, from which diffraction data have been collected to 2.0 Å. Molecular-replacement calculations have yielded a clear solution that suggests the presence of four protein molecules in the asymmetric unit.

2. Expression and purification of OFP

The full-length coding sequence of OFP (GenBank Accession No. AY296063, with C48S and F190L double mutations) was cloned into the pET3a plasmid (Novagen) using the *Nde*I and *Bam*HI restriction sites. The expression plasmid was transformed into *Escherichia coli* C41 strain (Miroux & Walker, 1996) for overexpression. The bacterial culture was grown in M9ZB medium containing 0.1 mg ml^{-1} ampicillin at 310 K until the absorbance at 600 nm reached 0.4. At this point, isopropyl-β-D-thiogalactopyranose was added to a concentration of 0.2 mM to induce protein expression. Cells were harvested after 16 h and lysed by sonication in buffer A (20 mM Tris pH 8.8) with 0.5 mM phenylmethanesulfonyl fluoride. After sonication, 0.1%(v/v) β-mercaptoethanol was added to prevent oxidation of the protein. The lysate was centrifuged at 30 000g for 30 min. The supernatant was loaded onto a HiTrap Q column (Amersham Biosciences) pre-equilibrated with buffer A. The protein was eluted with 0.1 M NaCl in buffer A. After dialysis against buffer A, the protein solution was loaded onto a MonoQ 10/10 (Amersham Biosciences) pre-equilibrated with buffer A. The protein was eluted using a linear gradient of 0–0.5 M NaCl in buffer A over a volume of

Table 1
Data-processing statistics for OFP.

Values in parentheses are for the highest resolution shell (2.11–2.00 Å).

Resolution (Å)	25.7–2.00
No. of measurements	551186
No. of unique reflections	61279
Multiplicity	9.0 (8.7)
Completeness (%)	99.7 (99.7)
$R_{\rm merge}$† (%)	9.5 (29.9)
Mean $I/\sigma(I)$	18.5 (5.7)

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

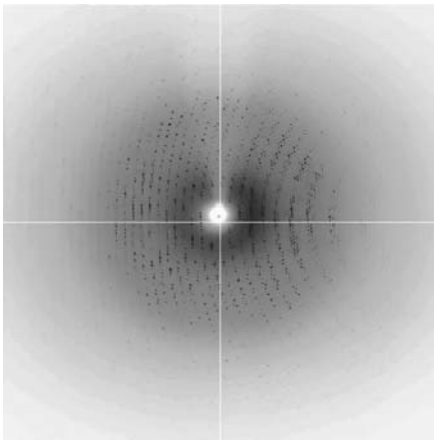

Figure 1
A diffraction image of OFP collected at beamline 14.1, Daresbury SRS, UK. The edge of the frame corresponds to 2.0 Å resolution.

128 ml. The purified protein was dialyzed against 20 m*M* Tris, 20 m*M* NaCl pH 8.8 and concentrated to 5 mg ml^{-1} by ultrafiltration. The yield of protein was typically ∼20 mg per litre of bacterial culture.

3. Crystallization and X-ray analysis

Initial screening for crystallization conditions was performed with the Index crystallization kit from Hampton Research using the hanging-drop vapour-diffusion method at 290 K. Crystals (many of them small crystal clusters) appeared in 29 of the initial 96 solutions. Since many of the solutions that grew crystals contained polyethylene glycol (PEG) 3350 as a precipitant, the crystallization conditions were further screened using the PEG/Ion kit from Hampton Research. Crystals for diffraction data collection were grown using the sitting-drop vapour-diffusion method at 290 K in conditions based on condition No. 3 of the PEG/Ion kit. Protein (1 µl at 5 mg ml^{-1}) was mixed with 1 µl of reservoir solution containing 0.1 *M* NH_4F, 10%(*w*/*v*) PEG 3350 and equilibrated over 100 µl of reservoir solution. Trigonal crystals grew in 24 h to maximum dimensions of 0.1 × 0.1 × 0.2 mm. Cryoprotection was achieved by soaking crystals in mother-liquor solution containing 20%(*v*/*v*) glycerol for 1 min at room temperature. The crystals were loop-mounted and flash-frozen in liquid nitrogen before transfer into the cryostream at 100 K. X-ray diffraction data were collected at a wavelength of 1.488 Å at beamline 14.1 of the Daresbury Synchrotron Radiation Source using an ADSC Quantum-4R CCD detector. A total of 360 images were collected from a single crystal in 1° oscillation steps, with a crystal-to-detector distance of 100 mm and an exposure time of 60 s per image (Fig. 1).

The diffraction data were indexed, integrated, scaled and merged using the programs *MOSFLM*, *SCALA* and *TRUNCATE* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994; Table 1). The crystals belong to the rhombohedral space group *R*3, with unit-cell parameters $a = b = 216.947$, $c = 51.839$ Å in the hexagonal lattice setting. With four protein molecules per asymmetric unit, the Matthews coefficient of the crystal is 2.3 Å^3 Da^{-1} and the solvent content is 47%.

Molecular replacement was performed with the program *MOLREP* (Collaborative Computational Project, Number 4, 1994) using all data in the resolution range 25.7–3.5 Å. A polyserine model of the blue coral pigment protein (PDB code 1mou; Prescott *et al.*, 2003), which shares 38% sequence identity with OFP, was used as the search model. Cross-rotation function calculation did not give a clear solution (Table 2*a*). In fact, peaks ranked third, fifth, sixth and tenth in the cross-rotation function calculation were found to be the correct solution in the subsequent translation-function calculation (Table 2*b*). Each monomer in the asymmetric unit was located sequentially and each translation function was calculated with the coordinates of previously found monomers fixed. A solution with four protein molecules per asymmetric unit was obtained. The correlation coefficient and *R* factor of the solution were 0.347 and 0.509, compared with values of 0.293 and 0.530 for the highest noise peak (Table 2*b*). No unfavourable molecular contacts were observed in the crystal packing. Model building and structure refinement are now under way.

Table 2
Molecular-replacement data.

(*a*) Cross-rotation function.

Peak No.	α (°)	β (°)	γ (°)	Peak height (σ)
1	117.73	31.15	352.83	5.35
2	64.14	148.10	174.92	5.25
3	30.83	90.38	153.04	5.22
4	48.96	90.23	144.45	5.16
5	29.40	91.21	332.69	4.99
6	88.92	88.99	301.55	4.97
7	10.34	27.73	345.62	4.78
8	74.09	89.19	129.44	4.73
9	105.17	90.84	309.65	4.71
10	89.38	90.46	121.77	4.70

(*b*) Translation function. $X_{\rm frac}$, $Y_{\rm frac}$ and $Z_{\rm frac}$ are fractional Cartesian coordinates. Values in parentheses correspond to the highest noise peak.

Monomer	Peak No.	$X_{\rm frac}$	$Y_{\rm frac}$	$Z_{\rm frac}$	*R* factor	Correlation coefficient
1st	6	0.559	0.310	0.000	0.578 (0.581)	0.144 (0.135)
2nd	3	0.183	0.992	0.703	0.555 (0.552)	0.226 (0.215)
3rd	5	0.847	0.188	0.832	0.531 (0.553)	0.290 (0.227)
4th	10	0.084	0.644	0.203	0.509 (0.530)	0.347 (0.293)

We thank Professor Alan Fersht for his generous support during K-BW's sabbatical visit to Cambridge. This work was partially supported by a Strategic Grant (SRP0102) from the Chinese University of Hong Kong.

References

Boute, N., Jockers, R. & Issad, T. (2002). *Trends Pharmacol. Sci.* **23**, 351–354.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Delagrave, S., Hawtin, R. E., Silva, C. M., Yang, M. M. & Youvan, D. C. (1995). *Biotechnology*, **13**, 151–154.
Ehrig, T., O'Kane, D. J. & Prendergast, F. G. (1995). *FEBS Lett.* **367**, 163–166.
Gerdes, H. H. & Kaether, C. (1996). *FEBS Lett.* **389**, 44–47.
Lippincott-Schwartz, J. & Patterson, G. H. (2003). *Science*, **300**, 87–91.
Matz, M. V., Fradkov, A. F., Labas, Y. A., Savitsky, A. P., Zaraisky, A. G., Markelov, M. L. & Lukyanov, S. A. (1999). *Nature Biotechnol.* **17**, 969–973.
Miroux, B. & Walker, J. E. (1996). *J. Mol. Biol.* **260**, 289–290.
Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G. & Cormier, M. J. (1992). *Gene*, **111**, 229–233.
Prescott, M., Ling, M., Beddoe, T., Oakley, A. J., Dove, S., Hoegh-Guldberg, O., Devenish, R. J. & Rossjohn, J. (2003). *Structure*, **11**, 275–284.
Tsien, R. Y. (1998). *Annu. Rev. Biochem.* **67**, 509–544.