Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Refined crystal structures of red and green fluorescent proteins from the button polyp *Zoanthus*

Nadezhda Pletneva,[a] Vladimir Pletnev,[a] Tamara Tikhonova,[b] Alexey A. Pakhomov,[a] Vladimir Popov,[b] Vladimir I. Martynov,[a] Alexander Wlodawer,[c] Zbigniew Dauter[d] and Sergei Pletnev[d]*

[a]Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Science, Moscow, Russia, [b]Bakh Institute of Biochemistry, Russian Academy of Science, Moscow, Russia, [c]Protein Structure Section, Macromolecular Crystallography Laboratory, National Cancer Institute at Frederick, Frederick, MD 21702, USA, and [d]Synchrotron Radiation Research Section, Macromolecular Crystallography Laboratory, National Cancer Institute, Argonne, IL 60439, USA

Correspondence e-mail: svp@ncifcrf.gov

Received 6 June 2007
Accepted 29 August 2007

PDB References: zRFP574, 2icr, r2icrsf; zGFP506, 2ojk, r2ojksf; zGFP506_N66D, 'green' form, 2pxw, r2pxwsf; zGFP506_N66D, 'red' form, 2pxs, r2pxssf.

The three-dimensional structures of the wild-type red (zRFP574) and green (zGFP506) fluorescent proteins (FP) from the button polyp *Zoanthus* have been determined at 1.51 and 2.2 Å resolution, respectively. In addition, the crystal structures of a zGFP506 variant (zGFP506_N66D) with replacement of the first chromophore-forming residue (Asn66 to Asp) have been determined in the transitional 'green' and mature 'red' states at 2.4 and 2.2 Å, respectively. The monomers of these proteins adopt the typical fold of the green fluorescent protein (GFP) family, consisting of an 11-stranded β-barrel with a chromophore embedded in the middle of an internal α-helix directed along the β-barrel axis. Post-translational modification of the chromophore-forming sequence Asn66-Tyr67-Gly68 within zGFP506 results in a typical GFP-like coplanar two-ring structure consisting of a five-membered imidazolinone heterocycle with the phenolic ring of Tyr67 in a *cis* orientation to the C^{α}—N(67) bond. A novel post-translational modification of the chromophore-forming sequence Asp66-Tyr67-Gly68 in zRFP574 expands the protein maturation beyond the green-emitting form and results in decarboxylation of the Asp66 side chain. It is suggested that electrostatic conflict between the closely spaced negatively charged side chains of the chromophore Asp66 and the proximal catalytic Glu221 is most likely to be the trigger for the chain of reactions resulting in the observed decarboxylation. The chromophore structures of wild-type zGFP506 and of its mutant zGFP506_N66D in the 'green' and 'red' states support this suggestion. The β-barrel frames of zRFP574 and zGFP506 reveal the presence of a water-filled pore leading to the chromophore Tyr67, similar to that observed previously in TurboGFP. An analysis of the residue composition at two inter-monomer interfaces in the tetrameric biological unit of zRFP574 and zGFP506, as well as of zYFP538 from the same species, has revealed a group of highly conserved residues that are apparently responsible for oligomerization. These residues present initial useful targets for rational mutagenesis aimed at designing monomeric forms of the fluorescent proteins, which are more suitable for practical applications.

1. Introduction

In recent years, native and mutant green fluorescent proteins (GFPs), as well as their homologues (here abbreviated FP), have become very useful tools in cell biology, biotechnology and medicine (Chalfie & Kain, 1998; Tsien, 1998; Haugwitz *et al.*, 2003; Zubova & Savitski, 2005; Tallini *et al.*, 2006). The extensive spectral and phylogenetic diversity of GFP-like proteins in marine organisms has been characterized during

© 2007 International Union of Crystallography
Printed in Denmark – all rights reserved

the last few years (Matz *et al.*, 1999; Labas *et al.*, 2002; Zubova *et al.*, 2003; Verkhusha & Lukyanov, 2004; Chudakov *et al.*, 2005; Remington, 2006). Using standard genetic methods, it became possible to effectively use GFP and related proteins as single or coupled FP markers for multicolour labelling of proteins, subcellular compartments and specific tissue regions. Utilization of FPs enabled the monitoring of cellular pH and concentration of Cl^-, Ca^{2+}, Cu^{2+} and Zn^{2+} ions, as well as the tracking of numerous states of the protein of interest in the cell or whole organism, such as expression, intracellular localization and trafficking, interaction with other cellular components and activity. The chromophore of an FP forms autocatalytically *in vivo* and *in vitro* from three residues, Xxx-Tyr-Gly, without the need for any cofactors or enzymes apart from molecular oxygen (Barondeau *et al.*, 2003). The water-filled pore in the protein 11-stranded β-barrel structure leading to the tyrosine of the chromophore has been suggested to serve as a path for conveyance of the oxygen (Evdokimov *et al.*, 2006).

In most cases, the post-translational modification results in a blue/green-emitting state characterized by the formation of an imidazolinone heterocycle with a *p*-hydroxybenzylidene substituent. Often, the reaction chain propagates further with the formation of an additional *N*-acylimine double bond, which extends the conjugation of the chromophore π electronic system and results in a bathochromic shift in the spectrum (Yarbrough *et al.*, 2001; Martynov *et al.*, 2003; Petersen *et al.*, 2003; Tubbs *et al.*, 2005; Pakhomov *et al.*, 2006; Pletneva *et al.*, 2006). The extensive spectral diversity of the fluorescent proteins arises mostly from variations in the chemical structure of the mature chromophore and the stereochemistry of its adjacent environment. The nature of the chromophore-forming residues is not critical to backbone cyclization. It has been shown that a colourless nonfluorescent GFP variant with a Gly-Gly-Gly chromophore sequence matures under aerobic conditions to a typical GFP post-cyclization structure (Barondeau *et al.*, 2003).

Proteins that emit red light are of particular interest. The longer wavelength extends the range of applications based on fluorescence resonance energy transfer (FRET) and helps to overcome the problem of interfering cellular auto-fluorescence, which to some extent limits the usage of the extremely popular *Aequorea victoria* GFP (avGFP). Red FPs coupled with green FPs can be used in multicolour tagging experiments, superseding the capabilities of the widely used cyan–yellow pair (Mizuno *et al.*, 2001).

The majority of wild-type GFP-like proteins form tetramers, complicating their practical use. To be especially useful tools for practical applications, the designed biomarkers should preferably exist in a monomeric form, emit in the far-red fluorescent spectral range, have high fluorescence quantum yield, be photostable and exhibit a high rate of chromophore maturation. It has always been difficult to develop variants that meet all of these criteria simultaneously. However, detailed knowledge of the structure–function relationship in this protein family provides the best basis for the rational design of improved biomarkers. It is important to know how the protein stereochemistry is able to tune the spectroscopic properties and to reveal the key residues responsible for the fold of a monomer and for protein oligomerization.

Here, we present a high-resolution structure of the red fluorescent protein zRFP574 from the button polyp *Zoanthus* sp. 2. We have now extended the resolution of the structure to 1.51 Å compared with the previously published structure at 2.4 Å (Pletneva *et al.*, 2006). We also present new crystallographic results of a study of the green fluorescent protein zGFP506, as well as of its mutant zGFP506_N66D from a related *Zoanthus* species in which Asn66, the first amino acid of the chromophore-forming triad, was substituted by aspartate. This particular selection of proteins has been made to prove our hypothesis concerning the initiating trigger of expansion of post-translational modification in zRFP574 beyond the green-emitting state. The structural results for these proteins complement our structural data for the yellow-emitting zYFP538 from *Zoanthus* sp. at a resolution extended to 1.8 Å (Pletneva *et al.*, 2007) compared with those originally published at the significantly lower resolution of 2.7 Å (Remington *et al.*, 2005).

These green, yellow and red wild-type FP variants from the button polyp *Zoanthus* (231 amino-acid residues) are characterized by Asn-Tyr-Gly, Lys-Tyr-Gly and Asp-Tyr-Gly chromophore-forming triads, respectively, at positions 66–68 and possess quite different spectral characteristics, with excitation/emission maxima at 492/506, 528/538 and 553/574 nm, respectively (Matz *et al.*, 1999; Gurskaya *et al.*, 2001; Labas *et al.*, 2002; Yanushevich *et al.*, 2003). In contrast to zRFP574 and zGFP506, the post-translational modification of the chromophore triad in zYFP538 is finalized by the formation of an additional six-membered tetrahydropyridine ring and the cleavage of the protein backbone at the $C^{\alpha}-N$ bond of Lys66. It was suggested that the energy conflict produced by the buried positive charge of the intact Lys66 side chain in the hydrophobic pocket formed by the Ile44, Leu46, Phe65, Leu204 and Leu219 side chains is most likely to be the trigger that expands the post-translational modification of zYFP538 beyond the green-emitting form (Pletneva *et al.*, 2007). Having an overall ~78% sequence identity, the selected proteins are excellent models for a structure-based search for the amino acids responsible for the transition to green-, yellow- and red-emitting states. In addition to an analysis of the stereochemistry in the chromophore area, we have also analyzed in detail the surface residues that are responsible for protein oligomerization.

2. Materials and methods

2.1. Purification and crystallization

The plasmids pQE30-zRFP574, pQE30-zGFP506 and pQE30-zGFP506_N66D were transformed into *Escherichia coli* JM109 (DE3) cells. The proteins were expressed on LB-Amp Petri plates by overnight incubation at 310 K. Cells were resuspended in 20 m*M* Tris pH 8.0 and 100 m*M* NaCl and disrupted by sonication. Supernatant clarified by centrifuga-

Table 1
Crystallographic data and refinement statistics.

Values in parentheses are for the highest resolution shell.

Protein	zRFP574	zGFP506	Green zGFP506_N66D	Red zGFP506_N66D
PDB code	2icr	2ojk	2pxw	2pxs
Crystallographic data				
Space group	$C222_1$	$P6_222$	$P6_222$	$P6_222$
Unit-cell parameters (Å)	a = 115.2, b = 146.7, c = 122.9	$a = b$ = 102.5, c = 270.9	$a = b$ = 102.2, c = 269.4	$a = b$ = 101.5, c = 271.2
Z/Z'	32/4	24/2	24/2	24/2
Estimated solvent content (%)	50	69	69	69
Temperature (K)	100	100	100	100
Wavelength (Å)	1.00	1.00	1.00	1.00
Resolution range (Å)	30.0–1.51	30.0–2.20	30.0–2.40	30.0–2.20
Total reflections measured	1145564	462772	456475	619656
Unique reflections observed				
All	162291	43725	33832	42936
Above $2\sigma(I)/3\sigma(I)$	159294/158890	35312/32982	28965/27490	36520/34577
Redundancy	7.1 (4.3)	10.4 (7.8)	13.5 (13.3)	14.4 (14.7)
$I/\sigma(I)$	30.6 (2.0)	23.1 (3.0)	30.3 (5.4)	39.1 (4.4)
R_{merge}	0.050 (0.651)	0.095 (0.741)	0.096 (0.620)	0.066 (0.663)
Completeness	99.4 (94.9)	99.9 (99.9)	99.9 (100.0)	99.8 (100.0)
Refinement statistics				
Non-H atoms in model				
Protein	7498 [4 × residues 4–231]	3616 [2 × residues 4–231]	3616 [2 × residues 4–231]	3616 [2 × residues 4–231]
Water	428	254	293	238
SO_4^{2-}	20 [4 × 5]			
R_{work}	0.178 [99% data, $F \geq 0$]	0.182 [95% data, $F \geq 0$]	0.167 [95% data, $F \geq 0$]	0.177 [94.8% data, $F \geq 0$)
R_{free}	0.225 [1% data, $F \geq 0$]	0.232 [5% data, $F \geq 0$]	0.201 [5.1% data, $F \geq 0$]	0.214 [5.2% data, $F \geq 0$]
Mean B factor [r.m.s.d.] (Å^2)				
Protein atoms				
Main chain	23.7 [1.6]	40.7 [0.8]	32.7 [0.7]	47.0 [0.8]
Side chain	28.2 [3.3]	42.5 [2.4]	34.3 [2.3]	48.8 [2.2]
Chromophore	20.5 [2.1]	37.2 [4.7]	28.6 [2.5]	39.5 [2.5]
Water	33.0	45.9	37.4	48.3
SO_4^{2-}	37.5 [3.8]			
Ramachandran statistics (for non-Gly/Pro residues) (%)				
Most favourable/additional allowed regions	100	99.5	99.5	99.5
Disallowed regions	0.0	0.5	0.5	0.5

tion was applied onto a metal-affinity column (Ni–NTA Agarose resin, Qiagen) and washed with ten column volumes of 20 m*M* Tris pH 8.0 and 100 m*M* NaCl. Target proteins were eluted with 20 m*M* Tris pH 8.0, 100 m*M* NaCl and 50 m*M* EDTA. Final purification was achieved by size-exclusion chromatography on a Superdex 75 HiLoad (16/60) column in 10 m*M* Tris pH 7.5, 100 m*M* NaCl and 2.5 m*M* EDTA. The proteins were concentrated to ~12 mg ml^{-1} in 10 kDa molecular-weight cutoff concentration units (Vivascience).

Crystals were grown by the hanging-drop vapour-diffusion method. zRFP574 crystals of intensely red colour were obtained under conditions similar to those used in our previous study (Pletneva *et al.*, 2006), except that the pH of the well solution was adjusted to 7.5 rather than 5.9. Each drop consisted of 1 µl protein solution in 10 m*M* Tris–HCl pH 7.5 and 50 m*M* NaCl mixed with an equal amount of reservoir solution. Cube-shaped crystals were obtained at 293 K from 1.4–1.6 *M* $(NH_4)_2SO_4$, 0.2 *M* potassium/sodium tartrate and 0.1 *M* Tris–HCl buffer pH 7.5. The crystal-growth process was quite slow, taking approximately five months to reach final dimensions of 0.15 × 0.15 × 0.15 mm.

The hexagonal bipyramidal crystals of the wild-type zGFP506 and of the 'red' form of the zGFP506_N66D mutant appeared in a few days at 293 K using 0.1 *M* bis-Tris propane and 1.8 *M* triammonium citrate pH 7.0 and reached final dimensions of ~0.2 × 0.2 × 0.2 mm in two weeks. After being kept at room temperature for several months, the solution of the zGFP506_N66D variant slowly changed in colour from green to red. A similar effect could be achieved in several days by incubating the protein at 310 K. To achieve the maximum rate of 'green' to 'red' transition, the solution of the mutant protein was incubated for several days at 310 K prior to crystallization. In contrast, crystals of the transitional 'green' form of the zGFP506_N66D mutant were obtained at 277 K in order to slow the maturation process. The latter crystals were obtained from 1.9 *M* sodium malonate pH 6.0, with a growth rate and final size similar to those of the crystals of wild-type zGFP506.

2.2. Structure solution and crystallographic refinement

X-ray diffraction data were collected from single crystals flash-cooled in a nitrogen stream at 100 K. Prior to cooling, the crystals were transferred into a cryoprotectant solution containing 20% glycerol and 80% reservoir solution. Data were collected on a MAR300 CCD detector at the SER-CAT

beamline 22ID (Advanced Photon Source, Argonne National Laboratory, Argonne, IL, USA). Crystallographic data were processed with *HKL*-2000 (Otwinowski & Minor, 1997).

The crystal structures of wild-type zRFP574 and zGFP506 were determined by the molecular-replacement method with *MOLREP* (Collaborative Computational Project, Number 4, 1994; Vagin & Teplyakov, 1997) using the coordinates of the monomers of the lower resolution zRFP574 structure (PDB code 2fl1; Pletneva *et al.*, 2006) as a model. The coordinates of wild-type zGFP506 were used to solve the structure of its mutant zGFP506_N66D. Structure refinement was performed with *REFMAC*5 (Murshudov *et al.*, 1997), alternating with manual revision of the model using *Coot* (Emsley & Cowtan, 2004). At the final stage of the refinement of the high-resolution zRFP574 structure, temperature factors for the non-H protein atoms were refined anisotropically with *SHELXL* (Sheldrick & Schneider, 1997).

Water molecules were located with *ARP*/*wARP* (Perrakis *et al.*, 1997). Crystallographic data and refinement statistics are presented in Table 1. Although the values of $R_{\rm merge}$ were quite high in the outermost shells of all data sets, the corresponding values of $I/\sigma(I)$ indicated that these data were still significant. In all cases, the use of complete experimental data sets for protein refinement provided the best quality of the resulting electron-density maps. Structure validation was performed with *PROCHECK* (Laskowski *et al.*, 1993). The figures were prepared with *SETOR* (Evans, 1993), *LIGPLOT*/*HBPLUS* (McDonald & Thornton, 1994; Wallace *et al.*, 1995), *PyMOL* (DeLano, 2002) and *ChemDraw* (CambridgeSoft).

3. Results and discussion

3.1. Wild-type zRFP574

3.1.1. Three-dimensional structure. The red fluorescent protein zRFP574 from the button polyp *Zoanthus* sp. 2 with chromophore-forming sequence Asp66-Tyr67-Gly68 is characterized by light excitation at $\lambda_{\rm ex}$ = 553 nm and emission at $\lambda_{\rm em}$ = 574 nm (Labas *et al.*, 2002; Yanushevich *et al.*, 2003). The asymmetric unit contains four monomers, *A*, *B*, *C* and *D*, assembled into two independent dimers, *AD* and *BC*. Furthermore, crystallographic symmetry operations create tetrameric molecules from pairs of independent dimers, resulting in the tetramers $ADA'D'$ and $BCB'C'$. The observed electron density of the final refined structure is generally of high quality, although poorer for monomer *D* than for the other three monomers in the asymmetric unit. Compared with monomers *A*, *B* and *C*, monomer *D* is characterized by an average temperature factor that is ~12 Å^2 higher and exhibits noticeably weaker density in several loop segments, comprising residues 114–115, 118–119, 155–156 and 189–194. The structure refined with 1.51 Å data is practically identical to that refined previously at 2.4 Å (Pletneva *et al.*, 2006), with an r.m.s.d. for all C^α atoms of only 0.3 Å.

The principal structural fold of the monomer of zRFP574 is an 11-stranded β-barrel, closed from both sides by loop caps, with a chromophore embedded in the middle of an internal α-helix located in the middle of the barrel and directed along its axis. Similarly to TurboGFP (Evdokimov *et al.*, 2006), the β-barrel frame of zRFP574 reveals the presence of a pore formed by the backbone of Trp145, Glu146, Pro147, His202, Lys203 and Leu204, with a chain of waters leading to the hydroxybenzylidene moiety of the chromophore (Fig. 1). Evdokimov *et al.* (2006) suggested that this pore is essential for chromophore maturation, providing access for molecular oxygen.

As a consequence of the extension of the resolution to 1.51 Å, we were able to detect alternative stable conformations for a number of side chains. A new interesting feature observed for all monomers indicates the possibility of unusually short linkages between the alternative conformations of the side chain of Phe97 and the side chains of Cys107 and Cys179 (Fig. 2). The benzyl ring of Phe97 is located at the centre of an imaginary flattened octahedron built by the side chains of three Phe residues (99, 126 and 179) and three Cys residues (105, 107 and 179) positioned symmetrically at the vertices. The side chain of Phe97 exhibits two alternative conformations, I and II, characterized by respective negative and positive values of the principal χ_1 angle. Conformational state I is mostly observed in subunits *A* and *B* (occupancy ~0.8–0.9), while state II is dominant in subunits *C* and *D* (occupancy ~0.8–0.9). The $F_{\rm o} - F_{\rm c}$ density indicates the presence of ~10–20% of the minor conformations, II for subunits *A* and *B* and I for subunits *C* and *D*. In the 2.4 Å structure determined from crystals grown at lower pH (5.9 *versus* 7.5), conformation II was only observed in subunit *D* (Pletneva *et al.*, 2006). The dominant conformations display good $2F_{\rm o} - F_{\rm c}$ electron density even at the enhanced level $\rho = 1.5\sigma$, whereas the density for the minor states is already almost absent at $\rho = 1.0\sigma$ and only displays the recognizable shape of the side chain at the rather low level $\rho < 0.7\sigma$. The side chain of Cys107 exhibits two stable alternative orientations, I and II, with respective occupancies of ~0.7 and ~0.3 for subunits *A* and *B*, and ~0.8 and

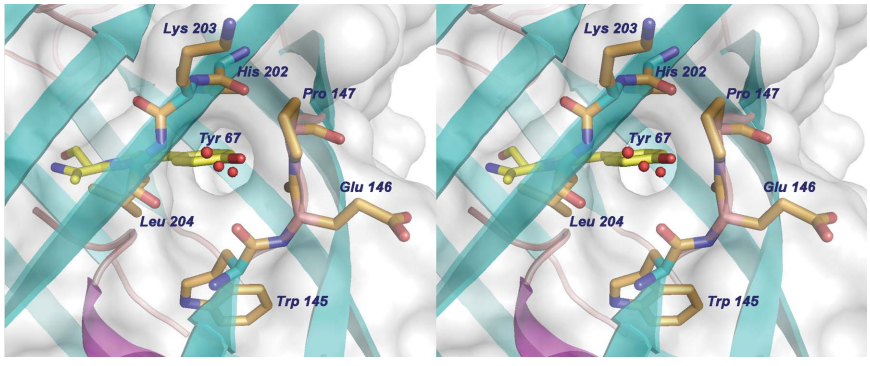

Figure 1
A stereoview of the zRFP574 β-barrel showing the pore with the chain of hydrogen-bonded waters (in red) going from the outside to the hydroxyl of the chromophore Tyr67. This figure was produced with *PyMOL* (DeLano, 2002).

~0.2 for subunits *C* and *D*, whereas the side chain of Cys179 adopts only a single conformation. All of these residues have good electron density.

An unusual short distance of single-bond character (~2.0 Å) between the C^{ε} ring C atom of Phe97 in the dominant state I and the S^{γ} sulfur of Cys107 in orientation I has been observed in subunits *A* and *B* (Fig. 2*a*). The strong dominance of state I in Phe97 coupled with the relatively high occupancy of state I in Cys107 make a corresponding cross-link highly probable. In subunits *C* and *D*, the C^{δ} ring C atom of Phe97 in the dominant conformational state II is at a short distance (~2.3 ± 0.15 Å) from the S^{γ} atom of Cys179 (Fig. 2*b*). The nature of this cross-link is intermediate between a covalent bond and a hydrogen bond.

These observed structural effects are quite unusual in character and need to be confirmed by other interdisciplinary methods. At this point, the mechanism of formation of the possible links and their functional roles are unclear. However, we should not disregard the possibility that some of these effects could arise from radiation damage, considering that cysteines are among the first residues to be affected (Burmeister, 2000; Ravelli & Garman, 2006). Similar short contacts were not detected in the previously published 2.4 Å structure based on X-ray data collected using an in-house diffractometer, either because of insufficient resolution combined with default interatomic bumping effects in *CNS* refinement or owing to less damage being caused by a weaker X-ray source (Pletneva *et al.*, 2006).

A similar Tyr–Cys cross-link possessing partial double-bond character (1.6–1.7 Å) has been observed in several oxidoreductase structures. In these enzymes, it was interpreted as a result of post-translational covalent modification of the tyrosine side chain in a self-processing reaction (Ito *et al.*, 1991; Whittaker, 2003; Schnell *et al.*, 2005). The bond is in close proximity to the substrate-binding site, suggesting its importance to the enzymatic reaction. Its removal by site-directed mutagenesis in sulfite reductase was shown to impair the catalytic activity (Schnell *et al.*, 2005).

3.1.2. Chromophore structure. Our previous study of zRFP574 with 2.4 Å data (Pletneva *et al.*, 2006) revealed an unexpected chemical structure of the internal chromophore originating from the protein sequence Asp66-Tyr67-Gly68. Maturation of the chromophore takes 3 d at ambient temperature and results in a two-

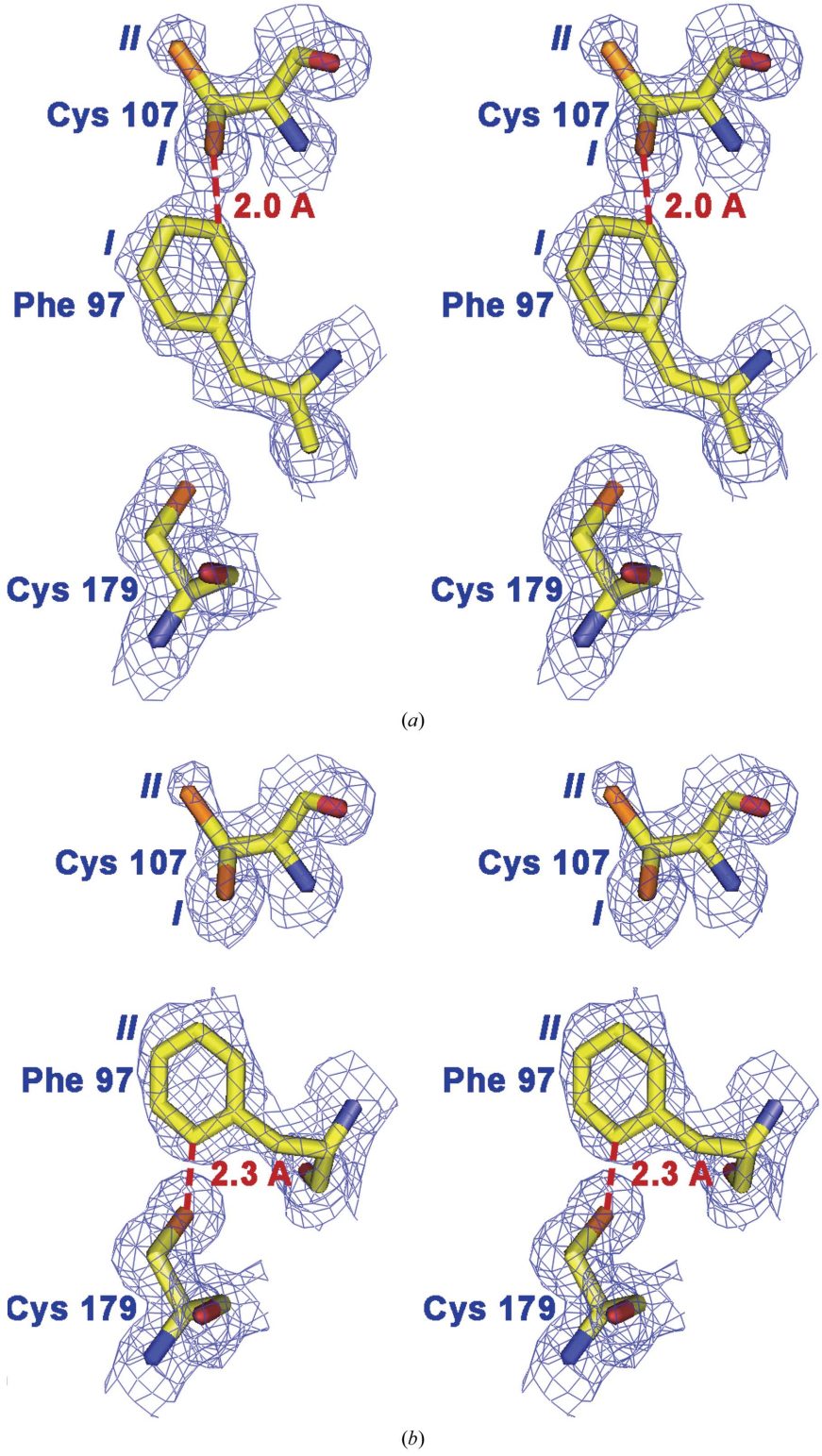

Figure 2
A stereoview of the cross-links between the side chains of Phe97 and Cys107 in subunits *A*/*B* (*a*) and of Phe97 and Cys179 in subunits *C*/*D* in a $2F_o - F_c$ electron-density map (cutoff $\rho = 1.0\sigma$). The figure was produced with *PyMOL* (DeLano, 2002) using zRFP574 monomer *B* as a template.

ring structure characterized by a decarboxylated side chain of the chromophore-forming Asp66 (the resulting residue is referred to as X66). In this study, the relatively high 1.51 Å resolution and the high quality of the X-ray data allowed us to confirm the previous observation and to present the structure in more detail. The chromophore structure is characterized by good electron density (Fig. 3*a*) and by atomic temperature factors that are below the average for the protein (Table 1). The $2F_o - F_c$ and $F_o - F_c$ electron-density maps are consistent with the modelled structure of the chromophore, with no indication of any disorder in the relatively tightly packed chromophore area. The bond lengths and angles of the chromophore of zRFP574 are similar to those of another red fluorescent protein, DsRed (Wall *et al.*, 2000; Yarbrough *et al.*, 2001; Tubbs *et al.*, 2005), and are close to the corresponding values observed in atomic resolution structures of small molecules which could be considered to be the most similar chromophore fragments. The common core of the chromophore, 2-methyl-4-benzylidene-2-imidazolin-5-one, is shared with two small-molecule analogues (CSD accession codes XACKAN and MOBKUJ in the Cambridge Database; Kurimoto *et al.*, 1999; Oshimi *et al.*, 2002) and covers almost the entire chromophore structure, from $C1^{\alpha}$ to $C3^{\alpha}$. The geometry of the essential chromophore part, $-(OC)-N-C1^{\alpha}(C1^{\beta})-C1-$, appears to be very similar to selected fragments of several amide-containing small molecules characterized by similar delocalized electron systems (CSD accession codes XATHAB and FUWSEV; Pilati, 1987; Matano *et al.*, 2000).

The stereochemical environment of the chromophore shown in Fig. 4(*a*) presumably includes the catalytic residues that participated in its formation. Among them are the nonconserved Ala63 (with position equivalent to Thr62 in GFP) and the conserved Arg95 and Glu221. The role of the catalytic residues was shown to not be vital for backbone cyclization in GFPs (Ehrig *et al.*, 1995; Barondeau *et al.*, 2003; Jung *et al.*, 2005).

Five water molecules (W) mediate hydrogen bonding between the chromophore and the functional groups of the closest amino acids. Water W1 bridges Tyr67 O^{η} and Glu146 C═O. Waters W2 and W3 connect Tyr67 C═O with Tyr91 O^{η} and with Arg70 $N^{\eta 2}$, Glu150 $O^{\varepsilon 1}$ and Tyr185 O^{η}, respectively, and W4 connects Phe65 C═O with Asn69 $N^{\delta 2}$. Water W5 is located within hydrogen-bonding distance of Arg70 $N^{\eta 2}$, Glu221 $O^{\varepsilon 1}$, Asn69 C═O and Asn69 N. Water molecules W3, W4 and W5 might be putative products of the oxidation reactions, resulting in C^{α}═C^{β} (Tyr67), C^{α}═N (X66) and (X66) C—N (Gly68) bonds, respectively, in the mature chromophore. Four of these five water molecules were also observed in the crystal structure of DsRed (Tubbs *et al.*, 2005).

According to a previously proposed hypothesis, steric interactions generated by the protein scaffold of GFP raise the precyclization-state energy, which enforces a dramatic bend in the central α-helix and drives the formation of a conventional two-ring chromophore structure (Barondeau *et al.*, 2003). Additionally, some residues located in close proximity to the chromophore might further promote its post-translational modification, which eliminates the remaining local energy conflicts and results in a shift of the spectral characteristics. The post-translational modification of the red fluorescent protein zRFP574 presumably passes through the two reaction steps illustrated in Fig. 5. The first step, catalyzed by the variable residue Ala63 and by the conserved Arg95 and Glu221, which are present in the neighbourhood of the chromophore, is common to all GFP-like proteins (Barondeau *et al.*, 2003, 2006; Wood *et al.*, 2005). This step results in an intermediate bicyclical product characterized by emission in the green spectrum range at 524 nm (Yanushevich *et al.*, 2003). The phenolic ring of Tyr67 adopts a *cis*

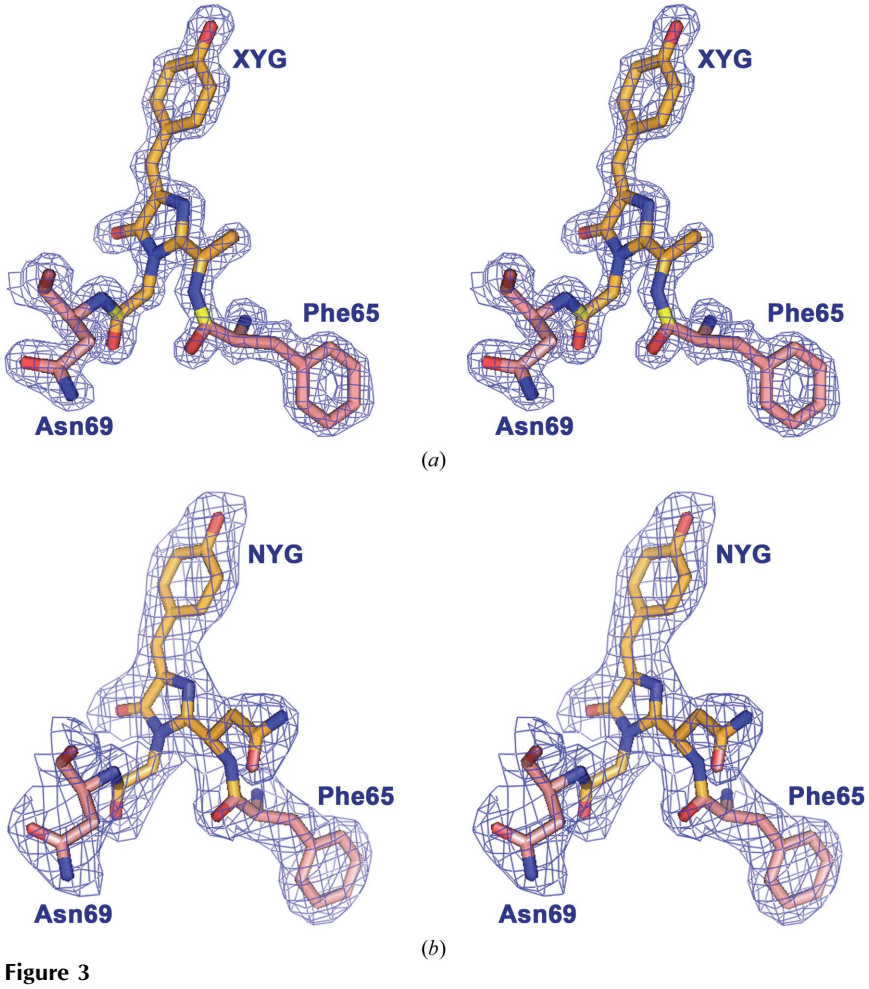

Figure 3
A stereoview of the zRFP574 (*a*) and zGFP506 (*b*) chromophores (in yellow) with the preceding and following residues (in magenta) in a $2F_o - F_c$ electron-density map (cutoff $\rho = 1.0\sigma$). This figure was produced with *PyMOL* (DeLano, 2002).

arrangement to the C^{α}—N bond with torsion angles $\chi_1 = 1.1^{\circ}$ and $\chi_2 = -3.2^{\circ}$, indicating planarity of the bicyclical system (Table 2). This arrangement, exhibited by most chromophores, is likely to be a passive consequence of a deprotonation event, resulting in the specific electronic state ultimately responsible for fluorescence. In this orientation, the phenolic ring is involved in effective stacking interactions (at 3.4 Å distance) with the parallel imidazole ring of the proximal His202. The second step, which converts the intermediate product from the 'green' to the final 'red' state, has only been observed for wild-type zRFP574. Similar to all red-emitting proteins (Tubbs *et al.*, 2005), it is likely to start from the formation of an unusual *cis*-peptide bond between residues 65 and 66 ($\omega = -8.2^{\circ}$; Table 2) and oxidation to form an acylimine. The accompanying sp^3 to sp^2 transition of the $C1^{\alpha}(66)$ hybridization (characterized by a corresponding change of the chirality angle ξ, $C1^{\alpha}$—N—$C1'$—$C1^{\beta}$, from $\sim 36^{\circ}$ to $\sim -8^{\circ}$; Table 2) results in decarboxylation of the side chain of Asp66. The additional *N*-acylimine partial double bond N═$C1^{\alpha}$ (X66) extends the conjugation of the π electronic system, resulting in a bathochromic shift in fluorescence emission to the wavelength 574 nm.

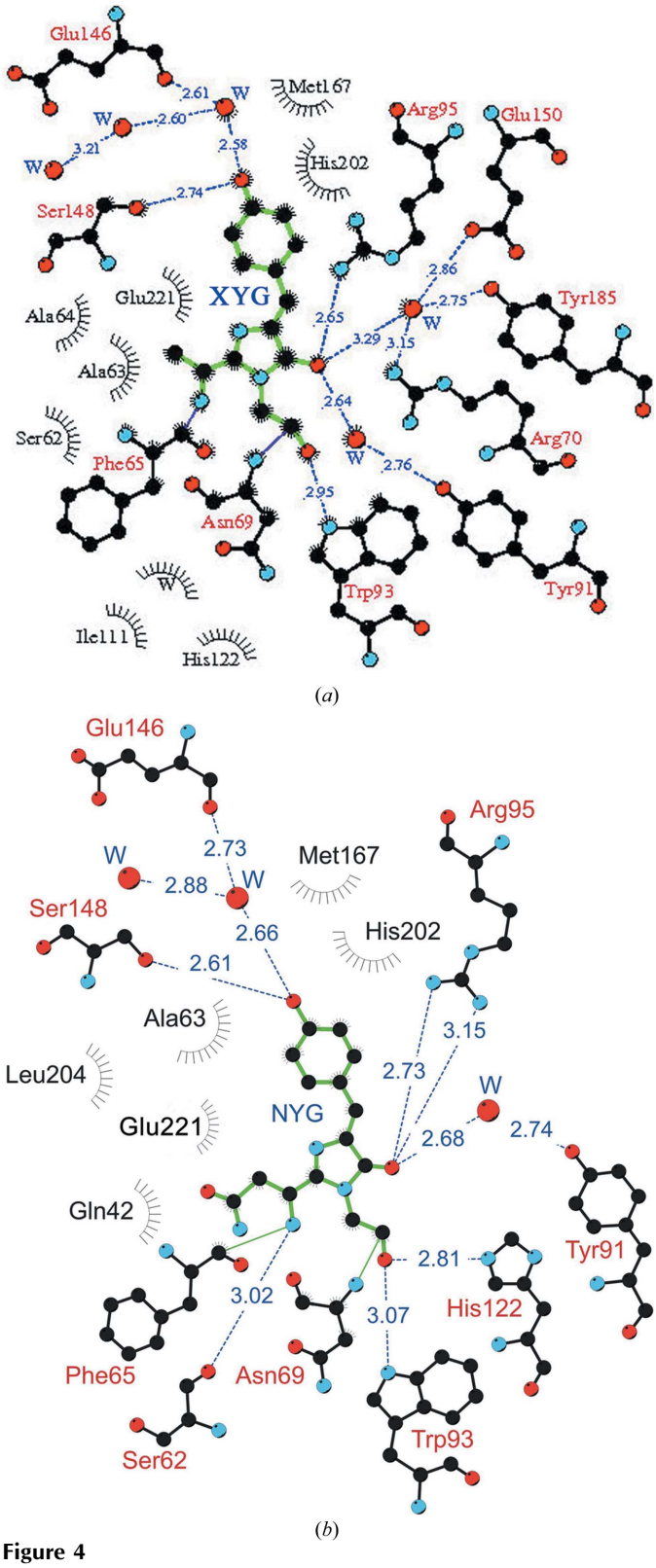

Figure 4
A schematic diagram illustrating the chromophore environment in zRFP574 (*a*) and zGFP506 (*b*). Hydrogen bonds (≤3.3 Å) are shown as blue dashed lines, waters (W) as red spheres and van der Waals contacts (≤3.9 Å) as black 'eyelashes'. This figure was produced with *LIGPLOT*/*HBPLUS* (Wallace *et al.*, 1995; McDonald & Thornton, 1994).

Table 2
Averaged values of selected torsion angles in the chromophore.

Torsion angle (°)	Averaged value (°)	R.m.s.d. (°)
ω(Phe65)	−8.2	2.4
φ(X66)	114.1	2.6
ψ(X66)	159.4	3.1
χ_1(Tyr67)	1.1	1.4
χ_2(Tyr67)	−3.2	1.3
φ(Gly68)	−95.1	1.3
ξ(X66)†	−8.1	1.4

† ξ, chirality angle $C1^{\alpha}$—N—$C1'$—$C1^{\beta}$.

Understanding of the catalytic mechanism of this process is still far from complete and requires verification by other independent methods. The autocatalytic transformation of the chromophore-forming residues to a conventional GFP two-ring structure apparently leaves the buried charge of the catalytic Glu221 side chain in a partially hydrophobic environment, which keeps this residue potentially reactive for additional steps of chemical modification under native conditions or arising from the influence of some external factors. This might be the presumptive reason for the observed decarboxylation of Glu at position 221 (zRFP574 numbering) in the green fluorescent protein avGFP (*Aequorea victoria*; van Thor *et al.*, 2002), a backbone break at the chromophore preceding the bond in asFP595 and the KFP variant (*Anemonia sulcata*; Martynov *et al.*, 2001; Wilmann, Petersen, Devenish *et al.*, 2005; Andresen *et al.*, 2005; Quillin *et al.*, 2005) and that observed in Kaede (*Trachyphyllia geoffroyi*), DendFP (*Dendronephthya* sp.) and EosFP (*Lobophyllia hemprichii*) (Mizuno *et al.*, 2003; Pakhomov *et al.*, 2004; Nienhaus *et al.*, 2005) autocatalytically or upon irradiation with UV or visible light. The replacement of the catalytic Glu by Gln in EosFP abolishes such fragmentation (Nienhaus *et al.*, 2005). Some other selectively chosen amino-acid substitutions around the chromophore can promote or suppress the extent of cleavage (Turcic *et al.*, 2006). In addition to Glu221, the proton-donor residues Gln42 and Arg70 and the adjacent

water molecule, all of which are hydrogen bonded to Glu221, might participate in the decarboxylation reaction observed in zRFP574.

We have considered several possibilities to explain the observed phenomenon of decarboxylation of Asp66 during maturation of the zRFP574 chromophore. We consider it unlikely that the effect results from radiation damage caused by the intense synchrotron source, although decarboxylation of aspartate and glutamate residues sometimes follows the irradiation of proteins (Burmeister, 2000; Ravelli & Garman, 2006). The 2.4 Å structure solved with the use of X-ray data from an in-house diffractometer source also did not show any evidence of the presence of an intact side chain for Asp66 (Pletneva *et al.*, 2006). We also excluded the mechanism of induced decarboxylation with visible light observed for Glu221 in avGFP (van Thor *et al.*, 2002), since zRFP574 is capable of maturation in the dark (Pakhomov & Martynov, 2007). The X-ray results were independently confirmed by mass-spectrometry data on the corresponding peptide containing the chromophore (Pakhomov & Martynov, 2007).

According to our hypothesis, the electrostatic conflict between the spatially close and most likely negatively charged (as the protein is in a pH 7.5 environment) side chains of the chromophore Asp66 and the proximal catalytic residue Glu221 is likely to be the initiation trigger for the chain of reactions leading to the observed decarboxylation of Asp66. In a tightly packed chromophore area, these interactions, which have a specific long-range nature, must inevitably produce a local energy jump which presumably overcomes the activation barrier for additional steps of post-translational modification. The replacement of Asp66 by the isosteric Asn in zRFP574 (Pakhomov & Martynov, 2007) removes the conflicting extra negative charge and stops the chromophore maturation of the conventional green-emitting form.

This specific role of Asp66 in chromophore formation has been evaluated by a study of the close zRFP574 homologue, zGFP506, and its mutant variant, zGFP506_N66D, with an Asn66Asp mutation.

3.2. Wild-type zGFP506 and the mutant zGFP506_N66D

3.2.1. The structure of zGFP506. The green fluorescent protein zGFP506 from the button polyp *Zoanthus* sp. is a close homologue of zRFP574, with 84% sequence identity. The chromophore of zGFP506 is formed from the sequence Asn66-Tyr67-Gly68 and this fluorescent protein has different spectral characteristics compared with zRFP574, with maxima of λ_{ex} = 492 nm and λ_{em} = 506 nm (Matz *et al.*, 1999). The final electron density of zGFP506 calculated at 2.2 Å resolution is of high quality and is in good agreement with the protein sequence for residues 4–231 of the two independent monomers (*A* and *B*) that are present in the asymmetric unit. Two independent tetramers, $AA'A''A'''$ and $BB'B''B'''$, are created by application of the crystal space-group symmetry operations. The structure of a monomer contains the same 11-stranded β-barrel as zRFP574, with the chromophore in the middle of an internal α-helix and with a similar water-filled pore formed by the backbone of the same conserved residues Trp145, Glu146, Pro147, His 202, Lys203 and Leu204 (Fig. 1).

Although this was only a medium-resolution map, its high quality allowed us to trace with confidence the alternative conformations, *A* and *B*, of the side chains of Glu100 and Cys107. In contrast to zRFP574, we found no evidence of any possible cross-links between Phe97 and Cys107 or Cys179. The observed single orientation *A* of the Phe97 side chain is characterized by a ~30° rotation of the aromatic ring around the $C^{\beta}-C^{\gamma}$ bond relative to that in zRFP574, with a shortest distance of 3.63 Å to Cys107 S^{γ} in the *A* state. The environment within ~9 Å distance of Phe97 is identical in both proteins, with the exception of the adjacent residue 98 (Leu in zGFP506 and Arg in zRFP574). Although in zGFP506 and both the 'red' and 'green' zGFP506_N66D structures Asp144 lies in the disallowed region of the Ramachandran plot, electron density clearly shows that those residues were modelled correctly. We were able to place 254 water molecules corresponding to the $\rho = 2.0\sigma$ level in the $F_o - F_c$ electron-density map.

The post-translational modification of the chromophore-forming sequence Asn66-Tyr67-Gly68 results in a conventional GFP two-ring structure consisting of a five-membered imidazolinone heterocycle having a *p*-hydroxybenzylidene substituent with the phenolic ring of Tyr67 in a *cis* orientation to the $C^{\alpha}-N(67)$ bond (Fig. 3*b*). In contrast to the decarboxylated Asp66 present in the chromophore of zRFP574, Asn66 of the mature chromophore of zGFP506 is intact, with sp^3 hybridization of the C^{α} atom, a single $C^{\alpha}-N$ bond and the *trans* configuration of the preceding peptide bond. The environment in the vicinity of the chromophores is very similar in the two proteins (Fig. 4*b*). Of a total of 37 sequence differences between zGFP506 and zRFP574, only one (Tyr16/Phe16) is close to the chromophore. In the environment of the chromophore, Tyr16 is involved in water mediation in an extensive hydrogen-bond network formed by side-chain and backbone interactions of the residues Tyr16, Gln42, Ser62, Ala63, Phe65, Asn69, Arg70, Tyr91, Arg95, Glu150, Tyr185, His202 and Glu221 (Fig. 6). Among these are the catalytic residues Glu221, Arg95 and Ala63. The proximal waters (the presumed reaction products of the maturation process) are actively involved in forming the network-mediating residue interactions. The hydrogen-bonded network interacts with the chromophore and is apparently functionally important, creating a potential proton wire in the maturation process. The three-dimensional structure of the shell surrounding the chromophore in zGFP506 is practically identical to its counterpart in zRFP574. The only exception is the side chain of Ile44 in zGFP506, which adopts a different orientation, avoiding steric conflict with the intact side chain of Asn66.

These results indicate that the replacement of Asn66 in zGFP506 by Asp66 in zRFP574 must be principally responsible for the observed significant difference in the spectral characteristics of the two proteins. To further elucidate the role of residue 66 in the formation of chromophores, we have determined the crystal structures of the zGFP506 mutant Asn66Asp in both transitional and mature states.

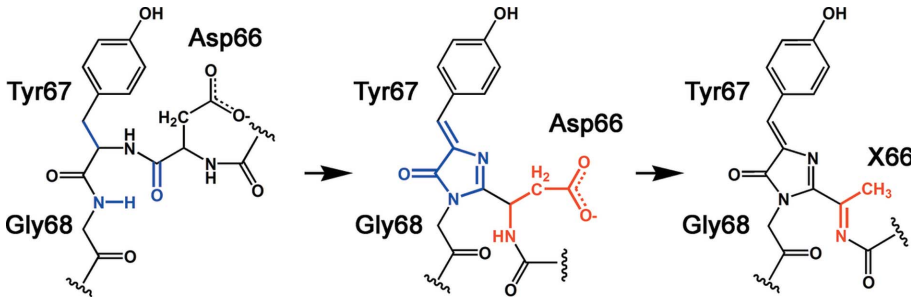

Figure 5
The reaction steps of the post-translational modification of zRFP574. The atomic groups affected by post-translational modification in the reactive paths to the 'green' and 'red' emitting forms are shown in blue and red, respectively. This figure was produced with *ChemDraw* (CambridgeSoft).

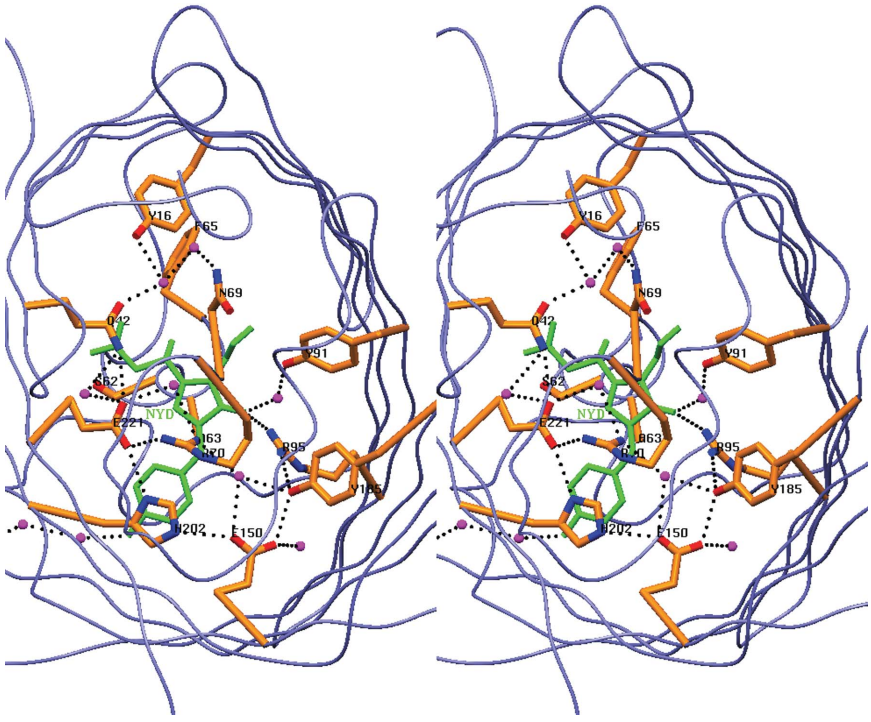

Figure 6
Stereoview of the hydrogen-bond network in the vicinity of the zGFP506 chromophore (shown in green). Mediating waters are shown as spheres (in magenta). This figure was produced with *SETOR* (Evans, 1993).

3.2.2. Chromophore structure of the mutant zGFP506_N66D. The replacement of the chromophore Asn66 by Asp in zGFP506 produces a large impact on the maturation process, expanding the post-translational modification of the chromophore beyond the green-emitting state. At room temperature, the 'green' form of zGFP506_N66D slowly (within several months) converts to the 'red' form. Increasing the temperature to 310 K strongly accelerates the transition process and results in conversion to the 'red' state within several days. The spectroscopic values of λ_{ex} = 496 nm and λ_{em} = 510 nm for the 'green' state of the N66D variant (Pakhomov & Martynov, 2007) are very close to those for the wild-type green fluorescent protein zGFP506, characterized by λ_{ex} = 492 nm and λ_{em} = 506 nm (Matz *et al.*, 1999), and the corresponding values of λ_{ex} = 555 nm and λ_{em} = 585 nm for the 'red' state of N66D (Pakhomov & Martynov, 2007) are similar to those for the wild-type red fluorescent protein zFP574, characterized by λ_{ex} = 553 nm and λ_{em} = 574 nm (Labas *et al.*, 2002; Yanushevich *et al.*, 2003).

We solved the crystal structures of both these forms, utilizing a green-emitting crystal form grown at 277 K and a red-emitting crystal form grown at room temperature after prior heating of the protein solution at 310 K. Each crystal structure shows the presence of two forms, dominant and minor, which have been modelled with different occupancies.

The dominant chromophore structure of the 'green' state of zGFP506_N66D includes an intact side chain for Asp66, with geometry corresponding to that of the wild-type zGFP506. A difference $F_o - F_c$ map reveals the presence of admixed 'red' form with a decarboxylated Asp66 side chain with ~20% occupancy. This contamination, which adds the yellowish shade to the crystal colour, apparently results from partial expansion of the chromophore maturation beyond the green-emitting form during crystallization. Crystals of the 'red' state of zGFP506_N66D contain ~70% of the dominant 'red' form, having decarboxylated Asp66 and a geometry similar to that observed in wild-type zRFP574. The remaining ~30% of the chromophore population consists of the 'green' form. As indicated by the electron density, in this state the peptide group connecting positions 65 and 66 in both the dominant and minor forms has an unusually linearized C—N—$C^{\alpha}1$ bond angle that is close to 160°. Analysis of the structures of a number of red fluorescent proteins present in the PDB also points to the linearized nature of this angle of the *cis*-peptide bond preceding the chromophore. The values of this angle vary from ~140° in zRFP574 (this work), HcRed (PDB code 1yzw; Wilmann, Petersen, Pettikiriarachchi *et al.*, 2005) and the K83M mutant of drFP583/dsRed (PDB code 2h8q; Shu *et al.*, 2006) to ~160° in DsRed (PDB code 1g7k; Yarbrough *et al.*, 2001) and eqFP611 (PDB code 1uis; Petersen *et al.*, 2003). The chemical nature of these observations is not clear. The observed variations may also reflect the differences in the restraint protocols used for structure refinement. The observed linearized C—N—$C^{\alpha}1$ angle for the 'green' contaminant in the 'red' zGFP506_N66D structure might be a

Table 3
Essential stabilizing contacts at the tetramer interfaces (Å).

Owing to the presence of twofold symmetry, the number of contacts shown has to be doubled.

Type	zRFP574	zGFP506	zYFP538
Intra-dimer contacts (IF1)			
Cluster	Val104, Ile106, Tyr127, Val129	Leu98, Val104, Ile106, Tyr127, Val129	Leu98, Val104, Ile106, Met129
Hydrogen bond	Asn130 $N^{\delta2}\cdots$Asp182 O, 3.24	Asn130 $N^{\delta2}\cdots$His94 O, 3.15	Asn127 $N^{\delta2}\cdots$Asn127 $O^{\delta1}$, 3.17
	Asn130 $N^{\delta2}\cdots$Asp182 O, 3.46	Lys186 $N^{\zeta}\cdots$Asp23 O, 3.15	Asn130 $O^{\delta1}\cdots$Lys162 N^{ζ}, 2.87
Inter-dimer contacts (IF2)			
Hydrogen bond	Gln201 $O^{\varepsilon1}\cdots$Ala229 N, 2.88	Pro147 O$\cdots$Lys151 N^{ζ}, 2.74	Ala147 O$\cdots$Lys151 N^{ζ}, 2.89
	Lys203 $N^{\zeta}\cdots$Ala231 O, 3.07	Gln201 $O^{\varepsilon1}\cdots$Ala229 N 2.85	Gln201 $O^{\varepsilon1}\cdots$Ala229 N, 2.93
	Lys203 $N^{\zeta}\cdots$Ala231 OT 3.27	Lys203 $N^{\zeta}\cdots$Pro231 OT, 2.81	Lys203 $N^{\zeta}\cdots$Ala231 OT, 3.00
Salt bridge	Glu146 $O^{\varepsilon2}\cdots$Lys151 N^{ζ}, 2.49	Glu146 $O^{\varepsilon2}\cdots$Lys151 N^{ζ}, 2.83	Glu146 $O^{\varepsilon2}\cdots$Lys151 N^{ζ}, 2.76
	Asp164 $O^{\delta1}\cdots$Arg178 $N^{\eta1}$, 3.59	Asp164 $O^{\delta1}\cdots$Arg178 $N^{\eta1}$, 3.67	Asp164 $O^{\delta1}\cdots$Arg178 $N^{\eta1}$, 3.55
	Asp164 $O^{\delta2}\cdots$Arg178 $N^{\eta1}$, 3.48	Asp164 $O^{\delta2}\cdots$Arg178 $N^{\eta1}$, 2.92	Asp164 $O^{\delta2}\cdots$Arg178 $N^{\eta1}$, 3.07
	Asp182 $O^{\delta2}\cdots$Arg178 $N^{\eta1}$, 3.36	Asp182 $O^{\delta1}\cdots$Arg178 $N^{\eta1}$, 2.92	Asp182 $O^{\delta1}\cdots$Arg178 $N^{\eta1}$, 2.83
	Asp182 $O^{\delta2}\cdots$Arg178 $N^{\eta2}$, 3.69	Asp182 $O^{\delta1}\cdots$Arg178 $N^{\eta2}$, 3.11	Asp182 $O^{\delta1}\cdots$Arg178 $N^{\eta2}$, 2.96

consequence of the locally unresolved model at medium resolution. The presence of the 'red' and 'green' forms in the zGFP506_N66D crystal structure is accompanied by two alternative orientations of the Ile44 side chain corresponding to those observed in the wild-type zRFP574 and zGFP506, respectively.

In striking contrast to zRFP574 (λ_{em} = 574 nm), which completely matures at room temperature within 3 d, the maturation process of the zGFP506_N66D mutant takes several months and is finalized by a product (λ_{em} = 585 nm) that exhibits a relative 11 nm red shift in emission spectra. The differences in the maturation and spectroscopic properties of the two proteins must be a consequence of the 36 differences in their amino-acid sequences. The most essential of these differences is likely to be the substitution of Tyr16 in zGFP506_N66D by Phe16 in zRFP574, since this is the only nonconserved residue in the vicinity of the chromophores. Involvement of Tyr16 in the hydrogen-bond network in the chromophore-adjacent environment (Fig. 6) might affect the maturation kinetics.

The combined structural data on zRFP574, zGFP506 and zGFP506_N66D highlight the critical role of Asp66 of the chromophore in the maturation process and support our conclusion that an electrostatic conflict between Asp66 and Glu221 in zRFP574 produces a strong impact on the chromophore's immediate environment and most likely initiates the chain of reactions that results in the decarboxylation of Asp66.

3.3. Tetramer packing

The crystals of the red, green and yellow (Pletneva *et al.*, 2007) fluorescent proteins from *Zoanthus* contain two independent dimers (zRFP574), two independent monomers (zGFP506) and one tetramer (zYFP538) in the asymmetric unit, respectively. Crystallographic symmetry operations transform both zRFP574 dimers and both zGFP506 monomers to corresponding pairs of independent tightly packed tetramers. Each tetrameric biological unit consists of two dimers packed at ~90° with respect to each other and exhibiting 222 symmetry. Each dimer is composed of two antiparallel monomers. The r.m.s.d. values on pairwise superposition of the zYFP538 monomer with those of zRFP574 and zGFP506 are 0.53 and 0.42 Å, respectively, for 226 equivalent C^{α} atoms, indicating a similar fold of the 11-stranded β-barrels. Superposition of the corresponding tetramers yields respective r.m.s.d. values of 1.88 and 0.65 Å for 226 × 4 C^{α} atoms, indi-

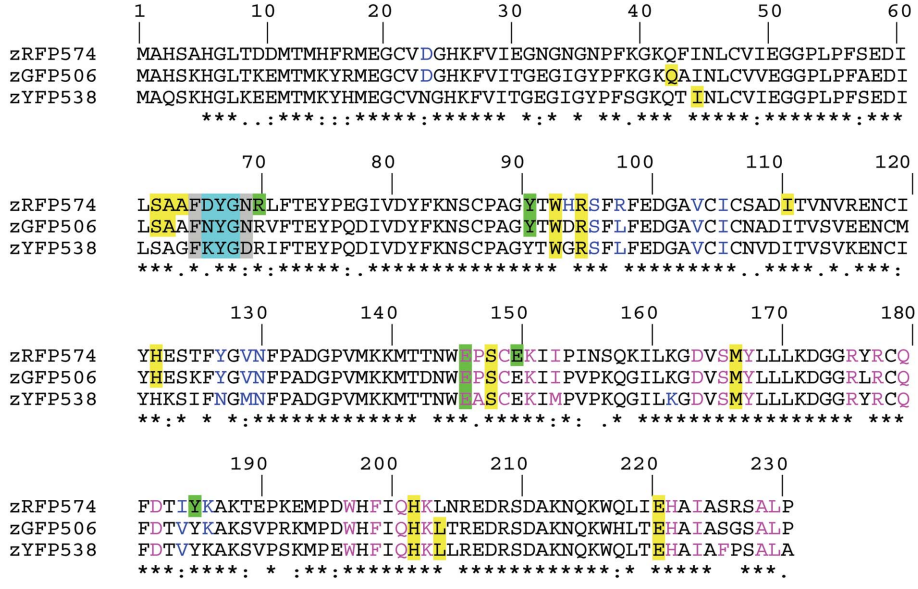

Figure 7
Sequence alignment of the red (zRFP574), green (zGFP506) and yellow (zYFP538) fluorescent proteins from *Zoanthus*. The overall sequence identity is 77.5%. Residues creating inter-monomer interfaces within a dimer and between dimers are shown in dark blue and magenta, respectively. Residues participating in hydrogen-bonding (within 3.3 Å) and van der Waals interactions (within 3.9 Å) with the chromophore directly and *via* waters are highlighted in yellow and green, respectively. The chromophore-forming triad and adjacent residues are highlighted in light blue and grey, respectively.

cating a slight difference in the tetramer packing for zRFP574. The chromophore moieties in a tetramer assembly are positioned at the vertices of an imaginary rectangle characterized by a somewhat inward orientation of the *p*-hydoxybenzylidene substituents, with shortest distances of ~34 and ~20 Å within and between dimers, respectively.

The tetrameric assemblies are characterized by two types of interface between the interacting surfaces of the monomers (Fig. 7). The first interface (IF1) is created within a dimer by residues from the antiparallel monomers and the second interface (IF2) by residues from monomers belonging to different dimers packed at ~90°. The IF1 interface is characterized by a solvent-excluded area of ~490 $Å^2$. It is mostly stabilized by the central hydrophobic cluster composed of eight to ten tightly packed side chains in proximity to the twofold symmetry axis (Table 3). The IF2 interface is more extensive, with a buried contact area of ~690 $Å^2$. Apart from a number of hydrophobic contacts, it is stabilized by a network of six hydrogen bonds and ten salt bridges, including highly specific links formed by the charged side chains of the conserved residues Glu146, Lys151, Asp164, Arg178, Asp182 and Lys203 (Table 3). Additionally, the symmetry-related Cys149 residues from both monomers in IF2 are spatially close, bringing the S atoms to a distance of ~3 Å in the alternative optimal side-chain orientations. It could be envisaged that these cysteines have the potential to form a disulfide bridge during oligomerization.

The total amino-acid composition of both intra-tetramer IF1–IF2 interfaces and the residues from the adjacent chromophore environment is presented in Fig. 7. A large number of residues from the chromophore environment and the intra-tetramer interfaces are positioned consecutively, alternating with each other. This pattern might have a functional significance, indicating the importance of the tetrameric organization for preservation of the natural fluorescent properties of the wild-type proteins.

Oligomerization of the naturally occurring fluorescent proteins is one of the factors that strongly limit their practical application for labelling proteins and subcellular structures. A number of monomeric FP variants have been generated to date, mostly by using extensive semi-random mutagenesis approaches (Campbell *et al.*, 2002; Karasawa *et al.*, 2003, 2004; Ando *et al.*, 2004; Shaner *et al.*, 2004; Wiedenmann *et al.*, 2004). Disruption of the tetrameric biological unit may produce noticeable changes in the spectral characteristics, photostability and kinetics of chromophore maturation, although the direction and the scale of these changes are not fully predictable *a priori*. As an example, a designed monomeric form of DsRed that contained 33 mutations exhibited a lower extinction coefficient, quantum yield and photostability than its tetrameric progenitor (Campbell *et al.*, 2002). On the other hand, this monomeric form matures ten times faster and exhibits an ~25 nm red shift of the excitation and emission maxima in its spectrum. The largest problem in designing monomeric FPs presumably arises from the residues responsible for the folding of a monomer. Their replacement might lead to unpredictable results, often resulting in nonfluorescent products. Detection of these residues is a difficult problem that still does not have a clear rational solution.

Cys149 and the residues listed in Table 3 present the most likely targets for rational mutagenesis that could lead to the creation of monomeric forms of the FPs discussed here. An attractive template for such a project could be the wild-type red fluorescent protein zRFP574. We expect that partial replacement of the hydrophobic residues by hydrophilic residues in the internal hydrophobic cluster at the IF1 interface, combined with mutations Cys149Ser and Lys/Arg to Glu/Asp or *vice versa* at IF2, would produce an effect sufficient to disrupt the intersubunit interfaces.

We thank Dr K. A. Lukyanov for providing us with plasmids pQE30-zFP574 and pQE30-zGFP506. We acknowledge the use of beamline 22-ID of the Southeast Regional Collaborative Access Team (SER-CAT), located at the Advanced Photon Source, Argonne National Laboratory. Use of the APS was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38. This work was supported in part by the Russian Foundation for Basic Research (grant No. 07-04-00054), a grant from the Russian Academy of Sciences for the program in Molecular and Cellular Biology (No. 200101) and by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

References

Ando, R., Mizuno, H. & Miyawaki, A. (2004). *Science*, **306**, 1370–1373.

Andresen, M., Wahl, M. C., Stiel, A. C., Grater, F., Schafer, L. V., Trowitzsch, S., Weber, G., Eggeling, C., Grubmuller, H., Hell, S. W. & Jakobs, S. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 13070–13074.

Barondeau, D. P., Cassmann, C. J., Tainer, J. A. & Getzoff, E. D. (2006). *J. Am. Chem. Soc.* **128**, 4685–4693.

Barondeau, D. P., Putnam, D. C., Kassmann, C. J., Tainer, J. A. & Getzoff, E. D. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 12111–12116.

Burmeister, W. P. (2000). *Acta Cryst.* D**56**, 328–341.

Campbell, R. E., Tour, O., Palmer, A. E., Stainbach, P. A., Baird, G. S., Zacharias, D. A. & Tsien, R. Y. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 7877–7882.

Chalfie, M. & Kain, S. (1998). *Green Fluorescent Protein: Properties, Applications and Protocols.* New York: Wiley–Liss.

Chudakov, D. M., Lukyanov, S. & Lukyanov, K. A. (2005). *Trends Biotechnol.* **23**, 605–513.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

DeLano, W. L. (2002). *PyMOL.* http://www.pymol.org.

Ehrig, T., O'Kane, D. J. & Prendergast, F. G. (1995). *FEBS Lett.* **367**, 163–166.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Evans, S. V. (1993). *J. Mol. Graph.* **11**, 134–138.

Evdokimov, A. G., Pokross, M. E., Egorov, N. S., Zaraisky, A. G., Yampolsky, I. V., Merzlyak, E. M., Shkoporov, A. N., Sander, I., Lukyanov, K. A. & Chudakov, D. M. (2006). *EMBO Rep.* **7**, 1006–1012.

Gurskaya, N. G., Savitsky, A. P., Yanushevich, Y. G., Lukyanov, S. A. & Lukyanov, K. A. (2001). *BMC Biochem.* **2**, 6.

Haugwitz, M., Dery, O., Turpin, P. & Fang, Y. (2003). *Genet. Eng. News*, **23**, 36–39.

Ito, N., Phillips, S. E., Stevens, C., Ogel, Z. B., McPherson, M. J., Keen, J. N., Yadav, K. D. & Kinowles, P. F. (1991). *Nature (London)*, **350**, 87–90.

Jung, G., Wiehler, J. & Zumbusch, A. (2005). *Biophys. J.* **88**, 1932–1947.

Karasawa, S., Araki, T., Nagai, T., Mizuno, H. & Miyawaki, A. (2004). *Biochem. J.* **381**, 307–312.

Karasawa, S., Araki, T., Yamamoto-Hino, M. & Miyawaki, A. (2003). *J. Biol. Chem.* **278**, 34167–34171.

Kurimoto, M., Subramony, P., Gurney, R. W., Lovell, S., Chmielewski, J. & Kahr, B. (1999). *J. Am. Chem. Soc.* **121**, 6952–6953.

Labas, Y. A., Gurskaya, N. G., Yanushevich, Y. G., Fradkov, A. F., Lukyanov, K. A., Lukyanov, S. A. & Matz, M. V. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 4256–4261.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

McDonald, I. K. & Thornton, J. M. (1994). *J. Mol. Biol.* **238**, 777–793.

Martynov, V. I., Maksimov, B. I., Martynova, N. Y., Pakhomov, A. A., Gurskaya, N. G. & Lukyanov, S. A. (2003). *J. Biol. Chem.* **278**, 46288–46292.

Martynov, V. I., Savitsky, A. P., Martynova, N. Y., Savitsky, P. A., Lukyanov, K. A. & Lukyanov, S. A. (2001). *J. Biol. Chem.* **276**, 21012–21016.

Matano, Y., Nomura, H. & Suzuki, H. (2000). *J. Organomet. Chem.* **611**, 89–99.

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.

Mizuno, H., Mal, T. P., Tong, K. I., Ando, R., Furuta, T., Ikura, M. & Miyawaki, A. (2003). *Mol. Cell*, **12**, 1051–1058.

Mizuno, H., Sawano, A., Eli, P., Hama, H. & Miyawaki, A. (2001). *Biochemistry*, **40**, 2502–2510.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Nienhaus, K., Nienhaus, G. U., Wiedenmann, J. & Nar, H. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 9156–9159.

Oshimi, K., Kubo, K., Kawasaki, A., Maekawa, K., Igarashi, T. & Sakurai, T. (2002). *Tetrahedron Lett.* **43**, 3291–3294.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pakhomov, A. A. & Martynov, V. I. (2007). In the press.

Pakhomov, A. A., Martynova, N. Y., Gurskaya, N. G., Balashova, T. A. & Martynov, V. I. (2004). *Biochemistry (Mosc.)*, **69**, 901–908.

Pakhomov, A. A., Pletneva, N. V., Balashova, T. A. & Martynov, V. I. (2006). *Biochemistry*, **45**, 7256–7264.

Pletneva, N., Pletnev, S., Chudakov, D., Tikhonova, T., Popov, V., Martynov, V., Wlodawer, A., Dauter, D. & Pletnev, V. (2007). *Bioorg. Khim.* **33**, 421–430.

Pletneva, N., Pletnev, S., Tikhonova, T., Popov, V., Martynov, V. & Pletnev, V. (2006). *Acta Cryst.* D**62**, 527–532.

Perrakis, A., Sixma, T. K., Wilson, K. S. & Lamzin, V. S. (1997). *Acta Cryst.* D**53**, 448–455.

Petersen, J., Wilmann, P. G., Beddoe, T., Oakley, A. J., Devenish, R. J., Prescott, M. & Rossjohn, J. (2003). *J. Biol. Chem.* **278**, 44626–44631.

Pilati, T. (1987). *Acta Cryst.* C**43**, 2432–2434.

Quillin, M. L., Anstrom, D. M., Shu, X., O'Leary, S., Kallio, K., Chudakov, D. M. & Remington, S. J. (2005). *Biochemistry*, **44**, 5774–5787.

Ravelli, R. B. G. & Garman, E. F. (2006). *Curr. Opin. Struct. Biol.* **16**, 624–629.

Remington, S. J. (2006). *Curr. Opin. Struct. Biol.* **16**, 1–8.

Remington, S. J., Wacher, R. M., Yarbrough, D. K., Branchaud, B., Anderson, D. C., Kallio, K. & Lukyanov, K. A. (2005). *Biochemistry*, **44**, 202–212.

Schnell, R., Sandalova, T., Hellman, U., Lindqvist, Y. & Schneider, G. (2005). *J. Biol. Chem.* **280**, 27319–27328.

Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E. & Tsien, R. Y. (2004). *Nature Biotechnol.* **22**, 1567–1572.

Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.

Shu, X., Shaner, N. C., Yarbrough, C. A., Tsien, R. Y. & Remington, S. J. (2006). *Biochemistry*, **45**, 9639–9647.

Tallini, Y. N. *et al.* (2006). *Proc. Natl Acad. Sci. USA*, **103**, 4753–4758.

Thor, J. J. van, Gensch, T., Hellingwer, K. J. & Johnson, L. N. (2002). *Nature Struct. Biol.* **9**, 37–41.

Tsien, R. Y. (1998). *Annu. Rev. Biochem.* **67**, 509–544.

Tubbs, J. L., Tainer, J. A. & Getzoff, E. D. (2005). *Biochemistry*, **44**, 9833–9840.

Turcic, K., Pettikiriarachchi, A., Battad, J., Wilmann, P. G., Rossjohn, J., Dove, S. G., Devenish, R. J. & Prescott, M. (2006). *Biochem. Biophys. Res. Commun.* **340**, 1139–1143.

Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

Verkhusha, V. V. & Lukyanov, K. K. (2004). *Nature Biotechnol.* **22**, 289–296.

Wall, M. A., Socolich, M. & Ranganathan, R. (2000). *Nature Struct. Biol.* **7**, 1133–1138.

Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). *Protein Eng.* **8**, 127–134.

Whittaker, J. W. (2003). *Chem. Rev.* **103**, 2347–2363.

Wiedenmann, J., Ivanchenko, S., Oswald, F., Schmitt, F., Rocker, C., Salih, A., Spindler, K. D. & Nienhaus, G. U. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 15905–15910.

Wilmann, P. G., Petersen, J., Devenish, R. J., Prescott, M. & Rossjohn, J. (2005). *J. Biol. Chem.* **280**, 2401–2404.

Wilmann, P. G., Petersen, J., Pettikiriarachchi, A., Buckle, A. M., Smith, S. C., Olsen, S., Perugini, M. A., Devenish, R. J., Prescott, M. & Rossjohn, J. (2005). *J. Mol. Biol.* **349**, 223–237.

Wood, T. I., Barondeau, D. P., Hitomi, C., Kassmann, C. J., Tainer, J. A. & Getzoff, E. D. (2005). *Biochemistry*, **44**, 16211–16220.

Yanushevich, Y. G., Gurskaya, N. G., Staroverov, D. B., Lukyanov, S. A. & Lukyanov, K. A. (2003). *Russ. Bioorg. Chem.* **29**, 325–329.

Yarbrough, D., Wachter, R. M., Kallio, K., Matz, M. V. & Remington, S. J. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 462–467.

Zubova, N. N., Bulavina, A. Y. & Savitski, A. P. (2003). *Uspekhi Biol. Khim.* **43**, 163–224.

Zubova, N. N. & Savitski, A. P. (2005). *Uspekhi Biol. Khim.* **45**, 1–66.