ISSN 1068-1620, Russian Journal of Bioorganic Chemistry, 2007, Vol. 33, No. 4, pp. 390–398. © Pleiades Publishing, Inc., 2007.
Original Russian Text © N.V. Pletneva, S.V. Pletnev, D.M. Chudakov, T.V. Tikhonova, V.O. Popov, V.I. Martynov, A. Wlodawer, Z. Dauter, V.Z. Pletnev, 2007, published in Bioorganicheskaya Khimiya, 2007, Vol. 33, No. 4, pp. 421–430.

Three-Dimensional Structure of Yellow Fluorescent Protein zYFP538 from *Zoanthus* sp. at the Resolution 1.8 Å

N. V. Pletneva[a,1], S. V. Pletnev[c], D. M. Chudakov[a], T. V. Tikhonova[b], V. O. Popov[b], V. I. Martynov[a], A. Wlodawer[d], Z. Dauter[c], and V. Z. Pletnev[a]

[a] *Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya 16/10, Moscow, 117997 Russia*

[b] *Bach Institute of Biochemistry, Russian Academy of Sciences, Leninskii pr. 33, Moscow, 119071 Russia*

[c] *Synchrotron Radiation Research Section, Laboratory of Macromolecular Crystallography, National Cancer Institute, Argonne, IL 60439, USA*

[d] *Protein Structure Section, Laboratory of Macromolecular Crystallography, National Cancer Institute, Frederick, MD 21702, USA*

Received January 18, 2007; in final form, January 22, 2007

Abstract—The three-dimensional structure of yellow fluorescent proteins zYFP538 (zFP538) from the button polyp *Zoanthus* sp. was determined at a resolution of 1.8 Å by X-ray analysis. The monomer of zYFP538 adopts a structure characteristic of the green fluorescent protein (GFP) family, a β-barrel formed from 11 antiparallel β segments and one internal α helix with a chromophore embedded into it. Like the TurboGFP, the β-barrel of zYFP538 contains a water-filled pore leading to the chromophore Tyr67 residue, which presumably provides access of molecular oxygen necessary for the maturation process. The post-translational modification of the chromophore-forming triad Lys66-Tyr67-Gly68 results in a tricyclic structure consisting of a five-membered imidazolinone ring, a phenol ring of the Tyr67 residue, and an additional six-membered tetrahydropyridine ring. The chromophore formation is completed by cleavage of the protein backbone at the C^{α}–N bond of Lys66. It was suggested that the energy conflict between 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 the most probable trigger that induces the transformation of the bicyclic green form to the tricyclic yellow form. A stereochemical analysis of the contacting surfaces at the intratetramer interfaces helped reveal a group of conserved key residues responsible for the oligomerization. Along with others, these residues should be taken into account in designing monomeric forms suitable for practical application as markers of proteins and cell organelles.

Key words: GFP-like proteins, crystal structure, chromophore structure, tetramer structure, intersubunit interfaces; yellow fluorescent protein, Zoanthus sp.

DOI: 10.1134/S1068162007040048

INTRODUCTION

Representatives of the family of green fluorescent proteins (GFP, *M* 25–30 kDa) possess a unique ability to generate fluorescence in live organisms. They cover a wide range of colors in their emission spectra, from cyan to red (λ_{em} = 442–645 nanometers). The GFP-like proteins find a wide application in molecular and cellular biology, biotechnology, and some medicinal areas as tools for labeling proteins, cellular compartments, cells, and tissues as markers that allow one to monitor in vivo the expression, localization, movement, and interactions of proteins. On their basis, sensors are created that monitor pH, concentration of Ca^{2+}, Cu^{2+}, Zn^{2+}, and Cl^{-} ions, cAMP, and activity of cellular kinases [1–7]. Such genetically encoded tools have a number of advantages over the chemically synthesized fluorophores. The chimerical constructs help label any protein of interest for a researcher, and also direct the fluorescent protein to any cellular compartment. The use of tissue-specific promoters allows staining the tissues of interest at a certain stage of development.

The tertiary structure of fluorescent proteins (FPs) has the form of a β-barrel with chromophore situated in its center. The chromophore is protected from the environment, which makes FPs not phototoxic and their properties independent of the environment parameters, such as, for example, pH. The chromophore is formed as a result of autocatalytic posttranslational modification of three amino acid residues of the protein molecule X-Tyr-Gly achieved with the help of molecular oxygen without participation of any cofactor or

[1] Corresponding author; phone +7 (495) 330-7510; e-mail: nadand@mail.ru.

enzyme [8]. In most cases, the posttranslational modification is completed by the formation of standard green form, whose chromophore is a planar bicyclic system of conjugated double bonds containing a five-membered imidazolinone cycle and a phenolic ring of the Tyr residue [2, 4]. In some cases, the process does not come to its end at this stage. An additional modification of an oxidative character proceeds; it is frequently accompanied by the formation of acylimine bond at the first residue of the chromophore-forming triad, which results in the expansion of the conjugated π-electronic system of chromophore and a related shift of the absorption and emission maxima in the long-wave area of spectrum. In this manner, the formation of yellow and red forms FPs with emission ranges from 528 to 590 nm proceeds [9–11]. The quaternary structure of the majority of GFP-like proteins of wild type has a tetramer organization [4, 5].

The practical application of FPs requires in most cases the design of monomer mutant variants possessing a high quantum yield, a high rate of chromophore maturing, and photostability. In this connection, it is important to understand how the special features of protein stereochemistry determine the characteristics of color spectrum and establish the key residues responsible for the formation of monomer structure and the protein oligomerization. The knowledge of the mechanism of chromophore formation, its structure, and amino acid environment allow the design by mutagenesis of the improved variants of the proteins that meet the required criteria.

In this work, we present the results of X-ray research of yellow fluorescent protein zYFP538 (λ_{exc} = 525 nm, λ_{em} = 538 nm [12]) isolated from a marine polyp *Zoanthus* sp. at a resolution 1.8 Å (spatial group $P2_12_12_1$). The work continues the research of another crystal form ($P3_121$) of this protein at lower resolution 2.7 Å [9]. The importance of design of monomeric FP forms for practical application prompted us to pay a particular attention to the stereochemical organization of intersubunit interfaces in the tetramer structure.

RESULTS AND DISCUSSION

The Structural Organization of Subunit

Monomer Structure. The asymmetric part of crystal cell of zYFP538 contains one tetramer. A high quality of its electronic density map allowed us to unequivocally describe the residues 6–231 for all four subunits (A, B, C, and D). The average size of the temperature *B*-factor for the D subunit is higher by 3.5 Å than that of the other subunits, which correlates with its lower electronic density. Water molecules (502) with the level of electronic density $\rho \geq 2.0\sigma$ form a network of hydrogen bonds with the functional groups of the tetramer amino acid residues. The topology of the studied structure of zYFP538 (spatial group $P2_12_12_1$) at the resolution of 1.8 Å is similar to that determined earlier for the same protein in another crystal form ($P3_121$) at the resolution of 2.7 Å [9] and is characterized by the size of root-mean-square deviation (RMSD) equal to 0.44 Å for all equivalent C^α atoms.

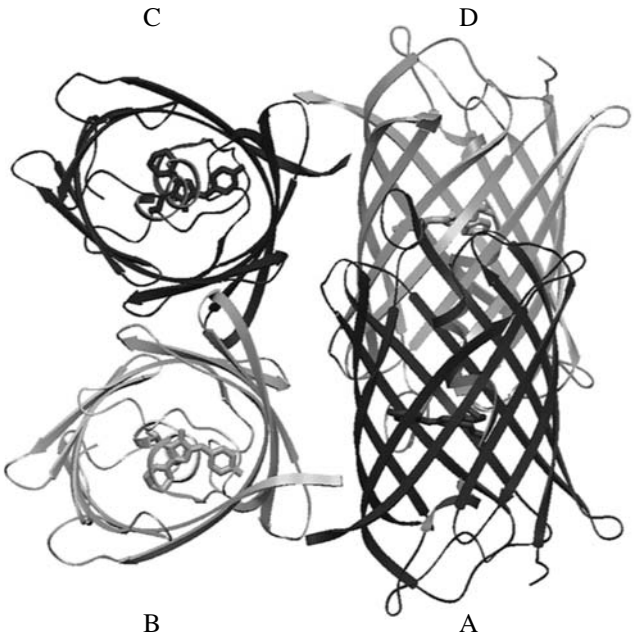

Fig. 1. Three-dimensional structure of tetramer zYFP538.

The monomer zYFP538 structure has the standard form of a barrel closed at both flat ends; it consists of 11 antiparallel β segments and one β helix in the center of which chromophore is located (Fig. 1). Like the green TurboGFP [13], the zYFP538 β-barrel exhibits the presence of the water-filled pore leading to the phenolic cycle of the Tyr residue of chromophore. The pore is formed by the atoms of the main chain of residues Trp145, Glu146, Lys203, Leu204 and Ala147. A similar pore is present in the red fluorescent protein zRFP574 from a related species *Zoanthus* sp2 (PDB ID 2FL1 [11]). Evdokimov et al. [13] hypothesized that the formation of such a pore in the FB structure is due to the necessity of providing of oxygen access at the chromophore maturing.

Chromophore structure. The chromophore of yellow fluorescent protein zYFP538 is formed from the amino acid triad Lys66-Tyr67-Gly68 and represents a planar tricyclic system of conjugated double bonds (Figs. 2, 3). The unicity of its structure consists in that the zYFP538 chromophore contains an additional six-membered tetrahydropyridine cycle besides the five-membered imidazolinone and phenolic (Tyr67) cycles characteristic of the standard bicyclic green form [2, 4]. Like in the green form, the phenolic ring of Tyr67 is located in the plane of chromophore in *cis*-orientation to the C^α–N (Tyr67) bond. Such a conformation is characteristic of the majority of known FPs of the GFP family (but not chromoproteins, whose quantum yield of

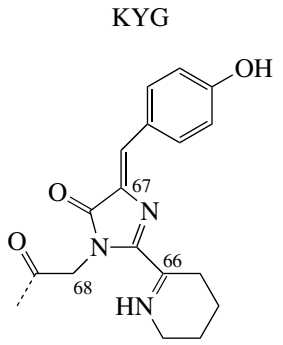

Fig. 2. Chemical formula of zYFP538 chromophore.

fluorescence is extremely low) [14–16]. In the structure of zYFP538, the Tyr67 phenolic ring participates in effective stacking-interaction (at the distance of ~3.7 Å) with the imidazole cycle of His202 located in parallel. The nearest stereochemical environment of the chromophore (Fig. 4) presumably includes the main catalytic residues that first of all participate in the formation of intermediate bicyclic structure of the green form. The variable residue Ala63 and conservative residues Arg95 and Glu221 are related to them [8]. In accordance to the suggested mechanism of chromophore formation [8, 17], the positively charged side chain of Arg95 creates an increased acidity of N atom of the Gly68 residue and/or the C^{α} atom of Tyr67, initiating corresponding deprotonation reactions. The oxygen atoms of the Ala63 carbonyl group and the side chain carboxyl group of the Glu221 residue presumably serve as acceptors of protons of one or both deprotonation reactions. The deprotonation at the N atom of the Gly68 residue favors the formation of five-membered imidazolinone cycle by the nucleophilic attack of the unshared electron pair on the carbonyl carbon of Lys66. The third six-membered heterocycle, according to the Remington hypothesis [9], is autocatalytically formed as a result of the transimination reaction of the N^{ζ} atom of the side chain of Lys66 with the reactive acylimine bond C^{α}=N of the chromophore transition state during its maturation. This reaction is accompanied by the cleavage of the protein backbone at the C^{α}–N of Lys66 residue with the formation of terminal carboxamide group at Phe65. In this case, the amino and carbonyl groups of carboxamide become located at the distances of hydrogen bond with the Gly68 carbonyl group (2.97 Å) and the N^{ζ} atom of the heterocyclic rings of Lys66 (2.82 Å), which points out to the protonated form of N^{ζ}. The C^{α} and N^{ζ} atoms of the additional chromophore heterocycle adopt the sp^2 hybridization and form a multiple bond C^{α}=N^{ζ} (d = 1.33 Å). This bond and two adjacent bonds, unlike the other heterocycle bonds, have a planar arrangement. The C^{α}=N^{ζ} bond is arranged in a plane of chromophore (torsion angle ψ[C–C^{α}] ~ 16°), which leads to the expansion of the conjugated system of double bonds and, consequently, to a shift of spectral characteristics toward long-waves. The observable bathochromic shift of excitation and emission maxima relative to the green fluorescent protein from the same source (*Zoanthus* sp.) are ~32 nm [12].

Note that five possible structures of the zFP538 chromophore have been considered in [18]. Four of them did not contradict the NMR data for chromopeptides isolated from zFP538. The authors thought that the chromophore structure (**V**) (Fig. 3 [18]) was less

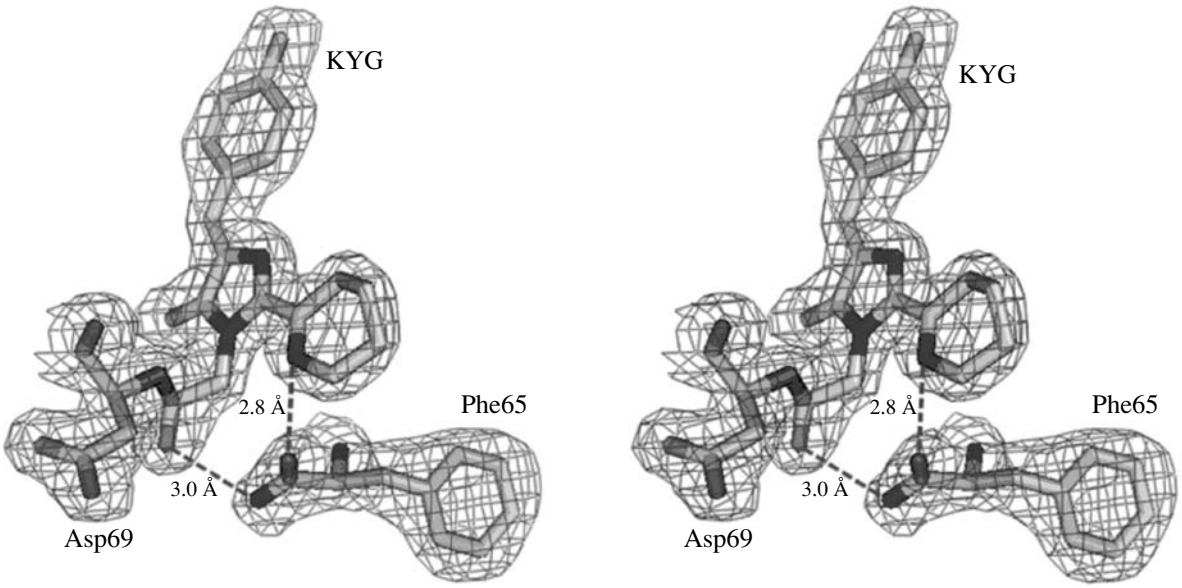

Fig. 3. 3D view of the chromophore structure in electron density 2Fo–Fc ($\rho \geq 1.0\sigma$).

probable, because it contradicts the NMR data for the chromopeptide. However, it coincides with the structure observable in this work for the crystal state of mature protein. Apparently, the unique spatial structure of the β-barrel provides the preset sequence of reactions resulting in the observable chromophore structure in mature protein.

The residues responsible for the formation of the yellow form. According to the data of mutagenesis [12], the Lys66 residue of the chromophore-forming sequence of zYFP538 plays the key role in the appearance of yellow fluorescence. The other two residues from the nearest chromophore environment, Asp69 and Gly64, are only partly involved in the process of the green form transition into the yellow form. The first plays probably a catalytic role, and the second provides the required conformational mobility of the backbone. The importance of these residues was confirmed by the creation of a mutant variant of green fluorescent protein from the same organism (λ_{em} = 506 nm) with three amino acid replacements (Asn66Lys/Asn69Asp/Ala64Gly), which was characterized by yellow fluorescence at λ_{em} = 537 nm [12].

Five water molecules in the environment of the zYFP538 chromophore form a network of hydrogen bonds mediating the chromophore interaction with the nearest amino acid residue. Two water molecules form direct chromophore interactions with chromophoreii (shown in Fig. 4). The third water molecule forms strong hydrogen bonds with the side chains of Gln42 and Tyr16 and a weak hydrogen bond (3.3 Å) with the carbonyl of terminal carboxamide group of Phe65. The fourth and fifth water molecules connect the side chains of Arg70, Glu150, and Tyr185 and Arg70, Glu221, and Asp69 (N), respectively.

The now accepted hypothesis [8] claims that the formation of three-dimensional architecture of a β-barrel creates a steric tension leading to a strong bend of the central α helix initiating the chain of chromophore formation reactions. In addition to the steric factor, a number of the residues from the nearest environment of chromophore form the specific spatial arrangement of functional groups that accelerates the process of chromophore maturing. After the formation of green bicyclic form, some residues from the nearest environment of chromophore can initiate the continuation of posttranslational modification.

Interestingly, the cyan FP amFP486 (λ_{em} = 486 nm) from *Anemonia majano* [19] with the chromophore-forming sequence Lys-Tyr-Gly identical to that of zYFP538 protein has (unlike zYFP538) the bicyclic structure characteristic of the green form, and the side chain of Lys residue does not undergo any changes at the chromophore maturing. The common identity of amino acid sequences (49%) is accompanied by a difference in the nearest environment of chromophores of these proteins that consists only in two amino acid replacements. The cyan protein amFP486 has the Ser and Asn residues instead of Ile44 and Asp69 characteristic of the yellow protein zYFP538.

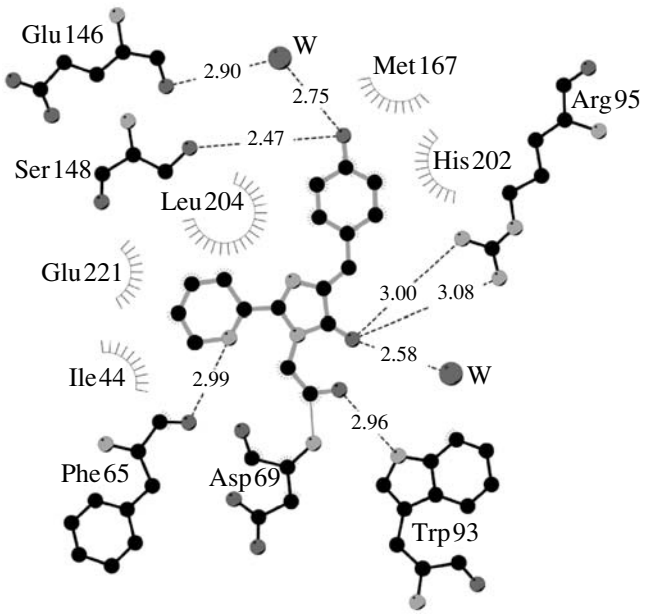

Fig. 4. Amino acid environment of chromophore. Hydrogen bonds (≤3.3 Å) are shown by dashed lines and van der Waals contacts (≤3.9 Å), by flagella.

The negatively charged side chain of residue Asp69 forms hydrogen bonds with the side chains of Tyr91 and His122 residues in the zYFP538 structure. It was shown that this residue plays an important role in transition of the green form into the yellow one [12]. A comparison of amino acid sequences and superposition of spatial structures of amFP486 and zYFP538 proteins showed that, in the immature form of zYFP538, the positively charged amino group of the natural side chain of Lys66 residue of the chromophore-forming sequence should by analogy occupy the hydrophobic cavity generated by the side chains of the Ile44, Leu46, Phe65, Leu204, and Leu219 residues. We believe that such arrangement of the charge in a hydrophobic environment leads to a severe energy conflict and induces an additional posttranslational modification leading to the formation of yellow chromophore. The forces responsible for the process initiation most likely have an enthalpy–entropy nature like the forces stabilizing a hydrophobic nucleus in proteins with water displacement [20]. In amFP486, the side chain of Ser occupying position 44 in the hydrophobic pocket forms a hydrogen bond with the intact side chain of the chromophore-forming Lys residue, partially neutralizing its charge.

Mutant variants. Only two amino acid residues, Ile44 and Asp69, are different in the environment of chromophore of the yellow zYFP538 protein and cyan protein amFP486 bearing a standard green chromophore. Therefore, we presumed that it is exactly the conflict of Ile44 with Lys66 that plays an essential role in the initiation of the formation of an additional six-

membered cycle. To test the hypothesis, we carried out the mutagenesis of zYFP538 protein with Ala substituted for Ile in position 44. In fact, a much smaller Ala in position 44 led to essentially less effective maturing of yellow chromophore. The absorption spectra showed that yellow form is dominating in the completely mature zYFP538 protein (green form is also present always) [12], whereas already green form of the chromophore slightly dominated in the mutant variant zYFP538 Ile44Ala. A stereochemical analysis showed that the observable effect of stabilization of the green form is only partially connected with the reduction of steric hindrances. Reduction of the power conflict of the charge of Lys66 side chain with the hydrophobic environment is most likely caused by a decrease in free energy owing to the reduced hydrophobicity of the above mentioned hydrophobic cavity. It has earlier been shown that the mutagenesis of zYFP538 at the second important residue, Asp69Asn, led to similar results (domination of the green chromophore form) [12].

We tested the cumulative contribution of Ile44 and Asp69 residues into the maturation of the zYFP538 yellow form using its double mutant Ile44Ala, Asp69Asn. The simultaneous replacement of the two residues has led to the full break in the formation of yellow chromophore without, at the same time, any disturbance in the mechanism of formation of standard green chromophore. The resulting mutant protein is characterized by a rather rapid maturation, bright green fluorescence (λ_{exc} = 496 nm and λ_{em} = 506 nm) and the only absorption peak at 492 nm. No absorption peak at 528 nm corresponding to the yellow form is observed even after 72-h incubation of the protein.

As mentioned above, a Ser residue is located in position 44 of the cyan fluorescent protein amFP486. Its side chain forms a hydrogen bond with the intact side chain of Lys66 residue. We assumed that the mutagenesis Ile44Ser of the zYFP538 protein will lead to a similar interaction of these residues and to the Lys66 stabilization, which would hinder the formation of additional heterocycle. Actually, an analysis of properties of the resulting mutant variant showed that Ser residue in position 44 to even greater degree than Ala interferes with the formation of yellow chromophore form. Dominates green and, which is especially interesting, there is the cyan form characterized by the absorption at 465 nm. It is highly probable that the cyan form of the mutant FP zYFP538_Ile44Ser corresponds to the chromophore whose Lys66 residue forms hydrogen bond with Ser44 similarly to the chromophore of amFP486 protein. Thus, in addition to the actual key residue Lys66 forming an additional chromophore heterocycle and also Gly64 that provides the necessary conformational mobility of the backbone, there are two more residues, Ile44 and Asp69, in the nearest environment of the zYFP538 chromophore that play an essential role in reaction of formation of an additional six-membered heterocycle.

Another example of a strong electrostatic effect in the nearest environment of chromophore represents the red FP zRFP574 from *Zoanthus* sp2 [11]. In this protein, the drawn together negatively charged side chains of Asp66 residue of chromophore and the catalytic Glu221 residue lead to a strong electrostatic conflict that initiates an additional chain of reactions of post-translational modification with the transition of the green form into the final red form. This transition most probably begins from the formation of unusual *cis*-peptide bond between the Phe65 and Asp66 residues accompanied by the formation of acylimine bond $C1^{\alpha}(66)$=N, which is characteristic of the majority of red FBs [21]. The subsequent change of hybridization of $C1^{\alpha}(66)$ atom from sp^3 to sp^2 is completed by the decarboxylation of Asp66 side chain.

The autocatalytic transformation of the chromophore-forming triad into the standard bicyclic green form of chromophore presumably leaves the charge of the side chain catalytic residue Glu221 not completely balanced. In this case, Glu221 potentially remains reactive for the subsequent additional stages of posttranslational modification under the native conditions or under the action of external factors. This is the possible cause of the observable decarboxylation of Glu221 residue (according to the numbering of zYFP538) at the photoactivation of green FP avGFP (*Auequorea victoria*) [22], and also break of the backbone before the chromophore in asFP595 (*Anemonia sulcata*) [15, 23] and in EosFP (*Lobophyllia hemprichii*) [24] under the influence of UV or visible light of a certain wavelength. The replacement of catalytic Glu residue by Gln in EosFP results in the absence of backbone fragmentation [24].

Tetramer Structure

The zYFP538 biological unit is a tetramer (dot symmetry *222*) consisting of two dimers packed at an angle of ~90° to each other (Fig. 1). Each dimer is formed by two antiparallel monomers. In tetramer, two types of interfaces are realized between the contacting surfaces of monomer subunits (Fig. 5). The first interface (IF1) is formed between two monomers inside the dimer, the second (IF2), between the monomers of two different dimers. IF1 is characterized by the contact area protected from the water environment of ~490 Å^2. It is mainly stabilized by the central hydrophobic cluster formed by the firmly packed side chains of two groups of residues: (1) L98, V104, I106, and M129 and (2) L98', V104', I106', and M129', adjoining to noncrystallographic axis of the second order. An additional contribution to the stabilization brings also a double set of hydrogen bonds connected by the elements of crystallographic symmetry A147(O)····K151(NZ) (2.89 Å), Q201(OE1)····A229(N) (2.93 Å), K203(NZ)····A231(OT) (3.00 Å) (designations of atoms are given according to the accepted nomenclature of the International Bank of Protein Structures, PDB).

Interface IF2 has a more developed contacting surface, ~690 Å^2. It is stabilized by a network of six hydrogen bonds

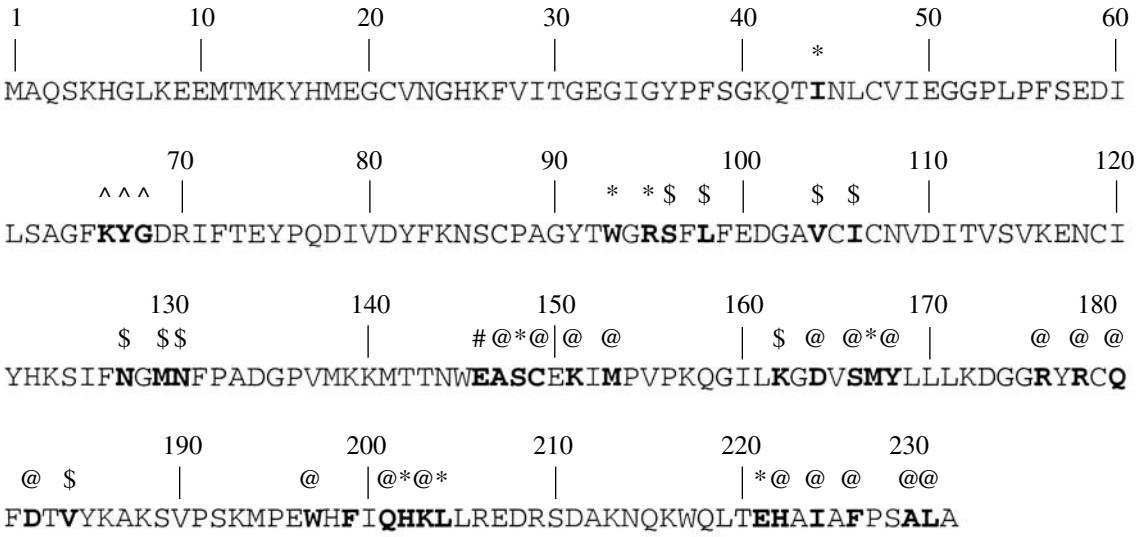

Fig. 5. Amino acid sequence of zYFP538. Chromophore-forming triad is marked by ^; residues interacting with chromophore via hydrogen bonds and van der Waals contacts, by *; interacting through water, by #; residues forming contact surfaces of interfaces IF1, by $; and interfaces IF2, by @.

and ten salt bridges formed by the side chains of residues of two identical groups, connected by an axis of symmetry of the second order: A147(O)···K151(NZ) (2.89 Å), Q201(OE1)···A229(N) (2.93 Å), K203(NZ)···A231(OT) (3.00 Å), E146(OE2)···K151(NZ) (2.76 Å), D164(OD1)···R178(NH1) (3.55 Å), D164(OD2)···R178(NH1) (3.07 Å), D182(OD1)···R178(NH1) (2.83 Å), D182(OD1)···R178(NH2) (2.96 Å). In addition, the side chains of the closely located residues Cys149 of both monomers are drawn together to the distance of ~3 Å in the case of another alternative orientation, which points out to a potential opportunity of formation of disulfide bonds in the process of oligomerization. Note that the stereochemical composition of interfaces in the zYFP538 structure is identical to that of the red FP zRFP574 from a related organism *Zoanthus* sp2 [11].

The marked residues on the interfaces IF1 and IF2 represent the most important group of the residues responsible for oligomerization. Along with other residues, they should be taken into account at the design of monomeric forms of FPs for the practical use as markers of proteins and cellular organelles. One should expect that a partial replacement of the hydrophobic residues of the central cluster IF1 by hydrophilic residues together with mutations Cys149 by Ser and Lys/Arg by Glu/Asp (or Glu/Asp by Lys/Arg) on IF2 would lead to the destruction of interfaces. The total amino acid structure in interfaces IF1 and IF2, and also the residues forming the nearest environment of the chromophore are represented in Fig. 5. It is interesting that a significant number of residues from both groups is located in the primary structure successively in an alternating order. It is not excluded that such an arrangement has the functional importance since it provides the spectral properties of protein of wild type due to the tetrameric organization.

The oligomerization of natural fluorescent proteins represents a serious biotechnological problem that significantly limits their application as protein markers and the development of genetically encoded fluorescent sensors. Some monomeric variants were designed with the use of a combination of the directed and accidental mutagenesis methods [25–30]. However, the attempts of intervention in stereochemistry of interfaces of the tetrameric biological unit are frequently accompanied by significant changes in spectral characteristics, photostability, and chromophore maturing kinetics. The direction and scale of these changes is difficult to predict. For example, the monomeric form DsRed designed on the basis of 33 mutational replacements possesses lower values of molar extinction coefficient, quantum yield, and photostability than its tetrameric precursor [25]. When designing monomeric variants with required characteristics, a substantial problem arises probably due to the residues responsible for the local formation of spatial structure. Along with other replacements, their substitution can lead to unpredictable results, frequently to the loss of fluorescent properties.

The studies of three-dimensional protein structures at a high resolution provide the important basis for the successful development of monomeric FP variants with preferable characteristics. Protein zYFP538 possesses unique spectral characteristics, and its monomeric variant would be of a wide demand for the technologies of multicolor labeling and resonance transfer of fluorescence energy (FRET). We hope that the results of this work would be useful to understand the regularities of

The experimental x-ray data and statistics of the crystallographic refinement of zYFP538

Yellow fluorescent protein zYFP538	
Experimental parameters	
Spatial group	$P2_12_12_1$ ($Z = 4$)
Cell parameters, Å	$a = 87.4$, $b = 106.5$, $c = 115.7$
Tetramer/asymm. part of cell	1
Water content, %	52.9
Temperature, K	100
Wavelength, Å	1.00
Resolution area, Å	30.0–1.8
Total number of measured reflexes	338587
A number of independent reflexes	92110
above $2\sigma(I)/3\sigma(I)$	68889/62490
Excess	3.7 (2.1)*
$I/\sigma(I)$	11.9 (1.4)
R_{merge}	0.064 (0.491)*
Set completeness	91.1 (53.1)*
Refinement statistics	
A number of nonhydrogen atoms	
Protein	7171 [4 × (6–231) residues]
Water	497
R_{work}	0.184 (99% of data, $F \geq 0$)
R_{free}	0.240 (1% of data, $F \geq 0$)
Average B factor/(RMSD) (Å^2)	
Protein atoms	
backbone	30.0 (0.8)
side chain	32.4 (2.3)
Chromophore	33.0 (4.0)
Water	37.0
MS deviation	
Bond lengths, Å	0.019
Valent angles, degree	1.8
Torsion angles [period 360/n] degree: n = 1/2/3/4	6.5/36.7/15.6/20.0
deviation from planarity, Å	0.008
Chiral volumes, Å^3	0.15
Statistics of deviation of φ–ϕ angles on the Ramachandran map (%) (excluding Gly/Pro residues	
the most advantageous areas	93.6
additional advantageous areas	6.2
allowed advantageous areas	0.1

Note: * The values in parentheses are given for the last resolution layer 1.86–1.80 Å.

formation of zYFP538 yellow chromophore and to develop its monomeric variant.

EXPERIMENTAL

Mutagenesis, expression, and isolation. The pQE30-zFP538 was transformed into the *Escherichia coli* JM109 (DE3) strain. Protein was expressed on Petri dishes with a LB agar by incubation at 37°C overnight. The dishes were filled with two layers of a firm LB agar with ampicillin additive (100 mg/ml). The bottom layer additionally contained 1 mM IPTG. The time delay of IPTG diffusion into the top layer of a LB agar provided the preset delay of the induction expression. Cells were resuspended in the buffer, pH 8.0, containing 20 mM Tris and 100 mM NaCl and destructed by sonication. The cellular suspension was clarified by centrifugation (15000 rpm for 5 min). The protein from supernatant was purified by metal affinity chromatography on a column (Ni-NTA Agarose resin, Qiagen). The column was washed with 10 volumes of 20 mM Tris buffer, pH 8.0, containing 100 mM NaCl. Protein was eluted with 20 mM Tris buffer, pH 8.0, containing 100 mM NaCl and 50 mM EDTA. The last stage of purification stage was carried out by the method of gel filtration on a Superdex 75 Hi-Load (16/60) column in 10 mM Tris buffer, pH 7.5, containing 100 mM NaCl and 2.5 mM EDTA.

The site-directed mutagenesis was carried out by the method described in [31] with the use of primers containing the corresponding point replacements. For the expression in *E. coli*, the PCR-amplified fragment encoding a fluorescent protein was cloned into a vector pQE30 (Qiagen) using the BamHI/HindIII restriction sites. The proteins bearing a polyhistidine sequence on the *N*-terminus (encoded in the expression vector pQE30 used) were expressed in *E. coli* (XL1-Blue strain, Invitrogen). Bacterial cultures centrifuged, the precipitate was resuspended in 20 mM Tris-HCl buffer, pH 7.4, with 100 mM NaCl, and sonicated. The recombinant proteins were purified on a metal-affinity resin TALON (Clontech).

Absorption spectra were registered on a Beckman DU520 UV/VIS spectrophotometer.

Crystallization and structure determination. Crystals zYFP538 were obtained by a method of diffusion through gaseous state in a variant of a hanging drop from a solution containing 0.085 M Na-Hepes, pH 7.5, 17% PEG 4000, 8.5% isopropanol, and 15% glycerol. Platelike crystals reached the maximal size of $0.1 \times 0.2 \times 0.4$ mm^3 after three weeks.

X-Ray diffraction data were collected from monocrystals at the temperature of liquid nitrogen, 100 K, with the use of MAR300 CCD detector on a SER-CAT 22-ID station (Advanced Photon Source, Argonne, United States). The processing of experimental data was carried out by a program complex HKL2000 [32].

The structure of zYFP538 was determined by the method of molecular replacement with the help of MOLREP program [33] within the structure of complex CCP4 [34] with use as model of coordinates of crystal structure zYFP538 determined at a resolution of 2.7 Å (PDB ID : 1XA9 [9]). The structure refinement was carried out with the use of program REFMAC5 [35] alternated by a manual correction of model in electronic density by program COOT [36]. The position of water molecules in the structure was determined with the help of program ARP/wARP [37]. The crystallographic parameters and the results of the crystallographic refinement of structure are given in the table. An analysis of geometrical parameters of structure was carried out with the help of program PROCHECK [38]. Figures were plotted by the use of programs SETOR [39], LIGPLOT/HBPLUS [40, 41], PYMOL [42], and ChemDraw [43].

Coordinates of atoms and experimental structural factors are deposited in the Protein Data Bank under the identification code 2OGR.

ACKNOWLEDGMENTS

We are grateful to Dr Sci (Chemistry) K.A. Luk'yanov (IBCh RAS) for a kind gift of plasmid pQE300-zFP538.

This work was supported by the Russian Foundation of Basic Research (project nos. 06-04-48196 and 07-04-00054), and the Program Molecular and Cellular Biology of the Russian Academy of Sciences. D.M. Chudakov was supported by the grant of the President of Russian Federation IE-8236.2006.4 and the Fund of Supporting the Domestic Science.

REFERENCES

1. Chalfie, M. and Kain, S., *GFP Green Fluorescent Protein. Properties, Applications and Protocols*, New York: Wiley-Liss, 1998.
2. Tsien, R.Y., *Annu. Rev. Biochem.*, 1998, vol. 67, pp. 509–544.
3. Haugwitz, M., Dery, O., Turpin, P., and Fang, Y., *Genet. Eng. News*, 2003, vol. 23, pp. 36–39.
4. Zubova, N.N., Bulavina, A.Yu., and Savitskii, A.P., *Usp. Biol. Khim.*, 2003, vol. 43, pp. 163–224.
5. Verkhusha, V.V. and Lukyanov, K.K., *Nat. Biotechnol.*, 2004, vol. 22, pp. 289–296.
6. Zubova, N.N. and Savitskii, A.P., *Usp. Biol. Khim.*, 2005, vol. 45, pp. 1–66.
7. Tallini, Y.N., Ohkura, M., Choi, B.R., Ji, G., Imoto, K., Doran, R., Lee, J., Plan, P., Wilson, J., Xin, H.B., Sanbe, A., Gulick, J., Mathai, J., Robbins, J., Salama, G., Nakai, J., and Kotlikoff, M.I., *Proc. Natl. Acad. Sci. USA*, 2006, vol. 103, pp. 4753–4758.
8. Barondeau, D.P., Putnam, D.C., Kassmann, C.J., Tainer, J.A., and Getzoff, E.D., *Proc. Natl. Acad. Sci. USA*, 2003, vol. 100, pp. 12 111–12 116.
9. Remington, S.J., Wacher, R.M., Yarbrough, D.K., Branchaud, B., Anderson, D.C., Kallio, K., and Lukyanov, K.A., *Biochemistry*, 2005, vol. 44, pp. 202–212.
10. Barondeau, D.P., Cassmann, C.J., Tainer, J.A., and Getzoff, E.D., *J. Am. Chem. Soc.*, 2006, vol. 128, pp. 4685–4693.
11. Pletneva, N., Pletnev, S., Tikhonova, T., Popov, V., Martynov, V., and Pletnev, V., *Acta Crystallogr.*, 2006, vol. D62, pp. 527–532.
12. Gurskaya, N.G., Savitsky, A.P., Yanushevich, Y.G., Lukyanov, S.A., and Lukyanov, K.A., *BMC Biochemistry*, 2001, vol. 2, p. 6..
13. 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., and Chudakov, D.M., *EMBO Rep.*, 2006, vol. 7, pp. 1006–1012.
14. Chudakov, D.M., Feofanov, A.V., Mudrik, N.N., Lukyanov, S., and Lukyanov, K.A., *J. Biol. Chem.*, 2003, vol. 278, pp. 7215–7219.
15. Andresen, M., Wahl, M.C., Stiel, A.C., Grater, F., Schafer, L.V., Trowitzsch, S., Weber, G., Eggeling, C., Grubmuller, H., Hell, S.W., and Jakobs, S., *Proc. Natl. Acad. Sci. USA*, 2005, vol. 102, pp. 13070–13074.
16. Remington, S.J., *Curr. Opin. Struct. Biol.*, 2006, vol. 16, pp. 1–8.
17. Wood, T.I., Barondeau, P., Hitomi, C., Kassmann, C.J., Tainer, J.A., and Getzoff, E.D., *Biochemistry*, 2005, vol. 44, pp. 16 211–16 220.
18. Zagranichny, V.E., Rudenko, N.V., Gorokhovatsky, A.Yu., Zakharov, M.V., Shenkarev, Z.O., Balashova, T.A., and Arseniev, A.S., *Biochemistry*, 2004, vol. 43, pp. 4764–4772.
19. Henderson, J.N. and Remington, S.J., *Proc. Natl. Acad. Sci. USA*, 2005, vol. 102, pp. 12 712–12 717.
20. Vondrasek, J., Bendova, L., Klusak, V., and Hobza, P., *J. Am. Chem. Soc.*, 2005, vol. 127, pp. 2615–2619.
21. Tubbs, J.L., Tainer, J.A., and Getzoff, E.D., *Biochemistry,* 2005, vol. 44, pp. 9833–9840.
22. van Thor, J.J., Gensch, T., Hellingwer, K.J., and Johnson, L.N., *Nat. Struct. Biol.*, 2002, vol. 9, pp. 37–41.
23. Quillin, M.L., Anstrom, D.M., Shu, X., O'Leary, S., Kallio, K., Chudakov, D.M., and Remington, S.J., *Biochemistry,* 2005, vol. 44, pp. 5774–5787.
24. Nienhaus, K., Nienhaus, G.U., Wiedenmann, J., and Nar, H., *Proc. Natl. Acad. Sci. USA*, 2005, vol. 102, pp. 9156–9159.
25. Campbell, R.E., Tour, O., Palmer, A.E., Stainbach, P.A., Baird, G.S., Zacharias, D.A., and Tsien, R.Y., *Proc. Natl. Acad. Sci. USA*, 2002, vol. 99, pp. 7877–7882.
26. Karasawa, S., Araki, T., Yamamoto-Hino, M., and Miyawaki, A., *J. Biol. Chem.*, 2003, vol. 278, pp. 34167–34171.
27. Karasawa, S., Araki, T., Nagai, T., Mizuno, H., and Miyawaki, A., *Biochem. J.*, 2004, vol. 381, pp. 307–312.
28. Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N., Palmer, A.E., and Tsien, R.Y., *Nature Biotechnol.*, 2004, vol. 22, pp. 1567–1572.
29. Ando, R., Mizuno, H., and Miyawaki, A., *Science*, 2004, vol. 306, pp. 1370–1373.
30. Wiedenmann, J., Ivanchenko, S., Oswald, F., Schmitt, F., Rocker, C., Salih, A., Spindler, K.D., and Nienhaus, G.U.,

Proc. Natl. Acad. Sci. USA, 2004, vol. 101, pp. 15905–15910.

31. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R., *Gene*, 1989, vol. 77, pp. 51–59.
32. Otwinowski, Z. and Minor, W., *Methods Enzymol.*, 1997, vol. 276, pp. 307–326.
33. Vagin, A. and Teplyakov, A., *J. Appl. Crystallogr.*, 1997, vol. 30, pp. 1022–1025.
34. Collaborative Computational Project Number 4, *Acta. Cryst.*, 1994, vol. D50, Part 5, pp. 760–763.
35. Murshudov, G.N., Vagin, A.A., and Dodson, E.J., *Acta Crystallogr.*, 1997, vol. D53, pp. 240–255.
36. Emsley, P. and Cowtan, K., *Acta Crystallogr.*, 2004, vol. D60, pp. 2126–2132.
37. Perrakis, A., Sixma, T.K., Wilson, K.S., and Lamzin, V.S., *Acta Crystallogr.*, 1997, vol. D53, pp. 448–455.
38. Laskowski, R.A., MacArthur M.W., Moss D.S., Thornton J.M, *J. Appl. Crystallogr.*, 1993, vol. 26, pp. 283–291.
39. Evans, S.V., *J. Mol. Graphics*, 1993, vol. 11, pp. 134–138.
40. Wallace, A.C., Laskowski, R.A., and Thornton, J.M., *Protein Eng.*, 1995, vol. 8, pp. 127–134.
41. McDonald, I.K. and Thornton, J.M., *J. Mol. Biol.*, 1994, vol. 238, pp. 777–793.
42. DeLano, W.L., in *http://www.pymol.org*, 2002.
43. Cambridge Soft Corporation, http://www.camsoft.com.