



# Orange Fluorescent Proteins: Structural Studies of LSSmOrange, PSmOrange and PSmOrange2

Sergei Pletnev<sup>1,2\*</sup>, Daria M. Shcherbakova<sup>3</sup>, Oksana M. Subach<sup>3,6</sup>, Nadya V. Pletneva<sup>5</sup>, Vladimir N. Malashkevich<sup>4</sup>, Steven C. Almo<sup>4</sup>, Zbigniew Dauter<sup>2</sup>, Vladislav V. Verkhusha<sup>3\*</sup>

**1** Leidos Biomedical Research Inc., Basic Research Program, Argonne, Illinois, United States of America, **2** Macromolecular Crystallography Laboratory, National Cancer Institute, Argonne, Illinois, United States of America, **3** Department of Anatomy and Structural Biology, Albert Einstein College of Medicine, Bronx, New York, United States of America, **4** Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York, United States of America, **5** Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russian Federation, **6** Department of Nano, Bio, Information and Cognitive Technologies, Moscow Institute of Physics and Technology, Moscow, Russian Federation

## Abstract

A structural analysis of the recently developed orange fluorescent proteins with novel phenotypes, LSSmOrange ( $\lambda_{\text{ex}}/\lambda_{\text{em}}$  at 437/572 nm), PSmOrange ( $\lambda_{\text{ex}}/\lambda_{\text{em}}$  at 548/565 nm and for photoconverted form at 636/662 nm) and PSmOrange2 ( $\lambda_{\text{ex}}/\lambda_{\text{em}}$  at 546/561 nm and for photoconverted form at 619/651 nm), is presented. The obtained crystallographic structures provide an understanding of how the ensemble of a few key mutations enabled special properties of the orange FPs. While only a single Ile161Asp mutation, enabling excited state proton transfer, is critical for LSSmOrange, other substitutions provide refinement of its special properties and an exceptional 120 nm large Stokes shift. Similarly, a single Gln64Leu mutation was sufficient to cause structural changes resulting in photoswitchability of PSmOrange, and only one additional substitution (Phe65Ile), yielding PSmOrange2, was enough to greatly decrease the energy of photoconversion and increase its efficiency of photoswitching. Fluorescence of photoconverted PSmOrange and PSmOrange2 demonstrated an unexpected bathochromic shift relative to the fluorescence of classic red FPs, such as DsRed, eqFP578 and zFP574. The structural changes associated with this fluorescence shift are of considerable value for the design of advanced far-red FPs. For this reason the chromophore transformations accompanying photoconversion of the orange FPs are discussed.

**Citation:** Pletnev S, Shcherbakova DM, Subach OM, Pletneva NV, Malashkevich VN, et al. (2014) Orange Fluorescent Proteins: Structural Studies of LSSmOrange, PSmOrange and PSmOrange2. PLoS ONE 9(6): e99136. doi:10.1371/journal.pone.0099136

**Editor:** Maria Sola, Molecular Biology Institute of Barcelona, CSIC, Spain

**Received:** March 13, 2014; **Accepted:** May 8, 2014; **Published:** June 24, 2014

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

**Data Availability:** The authors confirm that all data underlying the findings are fully available without restriction. The coordinates and structure factors for LSSmOrange, PSmOrange and PSmOrange2 have been deposited to Protein Data Bank under the accession codes 4Q7R, 4Q7T and 4Q7U.

**Funding:** Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38. The authors thank Rafael Toro for help with the crystallization screening and X29 beamline staff at NSLS for help with data collection. This work was supported in part with Federal funds from the National Cancer Institute, NIH, under contract HHSN261200800001E, the Intramural Research Program of the NIH, the NIH grants GM073913 and CA164468 (to VVV), the Albert Einstein Cancer Center grant CA013330, the New York Structural Genomics Research Center (to SCA), and by a grant from the Russian Science Foundation 14-14-00281. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does the mention of trade names, commercial products or organizations imply endorsement by the US Government. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* Email: pletnevs@mail.nih.gov (SP); vladislav.verkhusha@einstein.yu.edu (VVV)

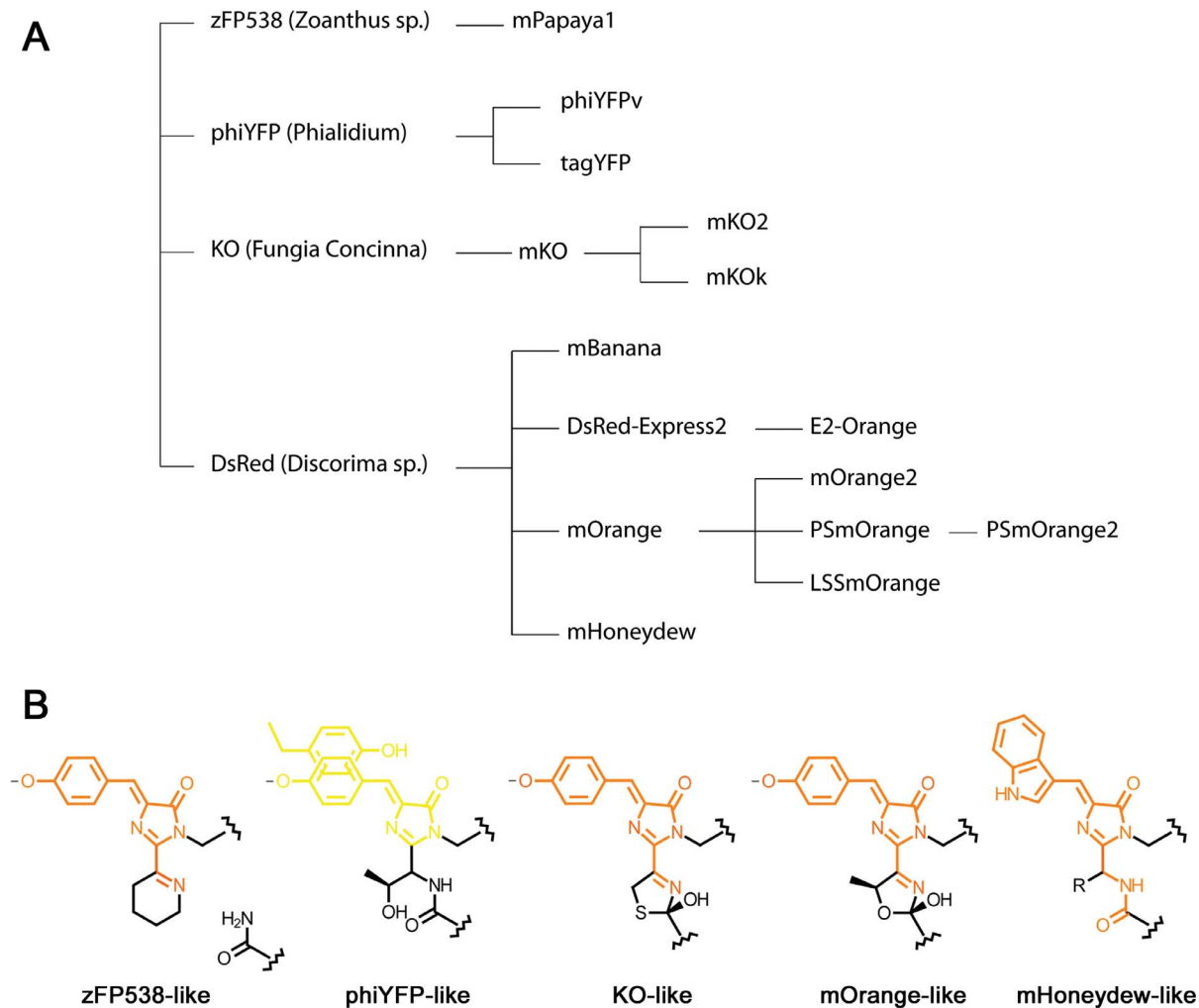
## Introduction

Fluorescent proteins (FPs) of the GFP-like family have become valuable tools for molecular biology, biochemistry, and biomedicine. Most challenging task of the FPs studies was the development of FPs with longer excitation/emission wavelength. This pursuit was motivated by advantages of red-shifted FPs, namely, lower background of cellular autofluorescence in microscopy, lower light scattering, and reduced tissue absorbance of longer wavelengths for *in vivo* imaging. In addition to common FPs, there are proteins of other phenotypes available, including FPs with a large Stokes shift (LSS) and irreversibly and reversibly photoswitchable (PS) FPs [1].

According to their emission wavelength, red-shifted FPs could be divided in the following groups: 520–540 nm yellow FPs (YFPs), 540–570 nm orange FPs (OFPs), 570–620 nm red FPs (RFPs), and >620 nm far-red FPs. Red shift of fluorescence of these FPs is predominantly achieved by extension of the

conjugated system of the chromophore and its protonation/deprotonation. The variety of spectral properties of FPs, such as excitation and emission wavelength, quantum yield, brightness, photoswitchability, Stokes shift of fluorescence, result from different chromophore structures and its interactions with surrounding amino acid residues.

OFPs fill up a spectral gap between YFPs and RFPs and enable four-color imaging together with blue, green, and far-red FPs. The spectral properties of OFPs mainly result from their chromophore structures. The chromophores of OFPs are formed by conservative -X-Y-G- tri-peptides in which X = Thr, Lys, Ser or Cys. In most cases, their orange emission is attributed to a three-ring chromophore structure with the third cycle formed by cyclization of the side chain of the X residue, extending the  $\pi$ -electron system of the chromophore [2]. Recombinant OFPs, reported so far, have been engineered from four wild-type proteins: zFP538 (*Zoanthus*



**Figure 1. Evolution of the subfamily of orange fluorescent proteins.** (A) Phylogenetic tree showing the history of the development of different orange fluorescent proteins. (B) Chemical structures of the chromophores found in orange fluorescent proteins.  
doi:10.1371/journal.pone.0099136.g001

*sp.*) [3], KO (*Fungia concinna*) [4], DsRed (*Discosoma sp.*) [5] and phiYFP (*Phialidium*) [6] (Figure 1, Table 1).

zFP538 is an obligate tetramer of low brightness with -K-Y-G- chromophore triad. Cyclization of Lys is accompanied by a cleavage of the polypeptide chain between the chromophore and the preceding amino acid residue, F65 [7–9]. Multiple rounds of directed evolution of zFP538 yielded its bright monomeric version, mPapaya1 [10], useful both as a protein tag and as a donor in orange-red Förster resonance energy transfer (FRET) pairs. The chromophore of phiYFP consists of -T-Y-G- residues. Its crystal structure shows that the side chain of Thr does not form a third cycle present in zFP538 and its yellow-orange emission was attributed to the antiparallel stacking between Y203 and the chromophore [11]. Two monomeric variants of phiYFP with improved folding, preserving this structural feature, phiYFPv and TagYFP, have been reported and described in [11]. The chromophore of KO is formed by -C-Y-G- amino acids and cyclization of Cys results in the third cycle, extending the chromophore conjugation system [12]. However, unlike in zFP538, the cyclization does not result in the cleavage of the polypeptide backbone. The monomeric form of KO with an improved folding, mKO, has been reported and characterized [4]. mKO has been further used to generate two variants with the

improved brightness, mKO2 [13] and mKOok [14]. Most prolific FP, which yielded a total of eight OFPs, is DsRed [5]. Red emission of its -Q-Y-G- chromophore was attributed to the presence of acylimine group connecting the imidazolinone ring of the chromophore with the preceding amino acid residue (F65) extending its conjugation system [5]. Variation of the DsRed chromophore tripeptide composition resulted in three OFPs: mHoneydew (-M-W-G-), mBanana (-C-Y-G-) and mOrange (-T-Y-G-) [2]. mOrange was shown to be the brightest and was used for generation of its more photostable variant mOrange2 [15]. Additionally, DsRed was used to obtain an RFP with low cytotoxicity, DsRed-Express2 [16,17], that served as a template for generation of its low cytotoxic orange variant E2-Orange [18].

Recently, three OFPs with novel phenotypes have been developed: permanently fluorescent LSSmOrange with a large Stokes shift ( $\lambda_{ex}^{max}$  437 nm,  $\lambda_{em}^{max}$  572 nm) [19] and two proteins photoswitchable from orange to far-red fluorescent states PSmOrange ( $\lambda_{ex}^{max}$  548 nm,  $\lambda_{em}^{max}$  565 nm; photoconverted form  $\lambda_{ex}^{max}$  636 nm,  $\lambda_{em}^{max}$  662 nm) [20] and PSmOrange2 ( $\lambda_{ex}^{max}$  546 nm,  $\lambda_{em}^{max}$  561 nm; photoconverted form  $\lambda_{ex}^{max}$  619 nm,  $\lambda_{em}^{max}$  651 nm) [21].

**Table 1.** Spectral and evolutionary features of orange fluorescent proteins.

| Fluorescent protein      | $\lambda_{\text{ex}}^{\text{max}}/\lambda_{\text{em}}^{\text{max}}$<br>(nm) <sup>1</sup> | $E_{\text{mol}}$<br>(M <sup>-1</sup> cm <sup>-1</sup> ) <sup>2</sup> | $\Phi_F$ <sup>3</sup> | B <sup>4</sup> | $B_{\text{rel}}^{\text{EGFP}}$ <sup>5</sup> | Reference |
|--------------------------|------------------------------------------------------------------------------------------|----------------------------------------------------------------------|-----------------------|----------------|---------------------------------------------|-----------|
| EGFP *                   | 489/509                                                                                  | 53,000                                                               | 0.6                   | 31.8           | 1                                           | [52]      |
| <i>(Zoanthus sp.)</i>    |                                                                                          |                                                                      |                       |                |                                             |           |
| zFP538                   | 528/538                                                                                  | 20,200                                                               | 0.42                  | 8.5            | 0.27                                        | [3]       |
| mPapaya1                 | 530/541                                                                                  | 43,000                                                               | 0.83                  | 35.7           | 1.12                                        | [10]      |
| <i>(Phialidium)</i>      |                                                                                          |                                                                      |                       |                |                                             |           |
| phiYFP                   | 525/537                                                                                  | 115,000                                                              | 0.6                   | 69             | 2.17                                        | [6]       |
| phiYFPv                  | 524/537                                                                                  | 101,305                                                              | 0.59                  | 59.8           | 1.88                                        | [11]      |
| TagYFP                   | 508/524                                                                                  | 20,056                                                               | 0.74                  | 14.8           | 0.47                                        | [11]      |
| <i>(Fungia concinna)</i> |                                                                                          |                                                                      |                       |                |                                             |           |
| KO                       | 548/561                                                                                  | 109,750                                                              | 0.45                  | 49.4           | 1.55                                        | [4]       |
| mKO                      | 548/559                                                                                  | 51,600                                                               | 0.6                   | 31             | 0.97                                        | [4]       |
| mKO2                     | 551/565                                                                                  | 63,800                                                               | 0.57                  | 36.4           | 1.14                                        | [13]      |
| mKOκ                     | 551/563                                                                                  | 105,000                                                              | 0.61                  | 64.1           | 2.01                                        | [14]      |
| <i>(Discosoma sp.)</i>   |                                                                                          |                                                                      |                       |                |                                             |           |
| DsRed**                  | 558/583                                                                                  | 75,000                                                               | 0.7                   | 52.5           | 1.65                                        | [5]       |
| mHoneydew                | 504/562                                                                                  | 17,000                                                               | 0.12                  | 2              | 0.06                                        | [2]       |
| mBanana                  | 540/553                                                                                  | 6,000                                                                | 0.7                   | 4.2            | 0.13                                        | [2]       |
| mOrange                  | 548/562                                                                                  | 71,000                                                               | 0.69                  | 49             | 1.54                                        | [2]       |
| mOrange2                 | 549/565                                                                                  | 58,000                                                               | 0.6                   | 34.8           | 1.09                                        | [15]      |
| LSSmOrange               | 437/572                                                                                  | 52,000                                                               | 0.45                  | 23.4           | 0.74                                        | [19]      |
| PSmOrange                | 548/565                                                                                  | 113,300                                                              | 0.51                  | 57.8           | 1.82                                        | [20]      |
| PSmOrange2               | 546/561                                                                                  | 51,000                                                               | 0.61                  | 31.1           | 0.98                                        | [21]      |
| DsRed-Express2**         | 554/591                                                                                  | 35,600                                                               | 0.42                  | 15             | 0.47                                        | [16]      |
| E2-Orange                | 540/561                                                                                  | 36,500                                                               | 0.54                  | 19.7           | 0.62                                        | [18]      |

\*The values for EGFP are given for comparison.

\*\*DsRed and DsRed-Express2 are red fluorescent proteins.

<sup>1</sup> $\lambda_{\text{ex}}^{\text{max}}/\lambda_{\text{em}}^{\text{max}}$  - excitation and emission maxima.

<sup>2</sup> $E_{\text{mol}}$  - extinction coefficient.

<sup>3</sup> $\Phi_F$  - quantum yield.

<sup>4</sup>B - brightness ( $E_{\text{mol}} \times \Phi_F$ )/1000.

<sup>5</sup> $B_{\text{rel}}^{\text{EGFP}}$  - brightness relative to EGFP (product of  $\Phi_F$  and  $E_{\text{mol}}$  compared to the brightness of EGFP (53,000 M<sup>-1</sup> cm<sup>-1</sup> × 0.60) [52]).

doi:10.1371/journal.pone.0099136.t001

LSSmOrange fills a spectral gap between green-yellow and red LSSFPs. Its brightness is five-fold larger than that of the brightest red LSSFP and is similar to that of green-yellow LSSFPs. LSSmOrange was shown to be useful for multicolor imaging, simultaneous detection of two Förster resonance energy transfer (FRET)-based biosensors using a single excitation wavelength, and single-laser fluorescence cross-correlation spectroscopy (FCCS) imaging of four FPs [19]. PSmOrange could be used for multicolor photolabeling, super-resolution photoactivated localization microscopy (PALM), and tracking of the photoswitched cell population in living animals [20]. PSmOrange2 has been shown to exhibit a higher rate and improved efficiency of photoconversion and enable a unique phenomenon, termed FRET-facilitated photoswitching [21].

Here, we report the X-ray crystal structures of LSSmOrange, PSmOrange, and PSmOrange2, focus on peculiarities of the chromophores and their immediate environments, and discuss the mechanistic basis of the unique LSS and PS properties of FPs.

## Materials and Methods

### Protein Expression and Purification

For crystallization, PCR-amplified *Bgl*II/*Eco*RI fragments encoding LSSmOrange, PSmOrange and PSmOrange2 were cloned into a pBAD/His-B vector (Invitrogen), modified by shortening the N-terminal polyhistidine tag to the MGSHHHHHHGRS-amino acid sequence. All proteins were expressed in LMG194 bacterial host (Invitrogen) grown in a minimal medium (RM) supplemented with 0.005% arabinose at 37°C for 24 h and at 25°C for 24 h upon shaking at 200 rpm. The cells were pelleted down by centrifugation, re-suspended in phosphate buffer saline, and lysed by sonication. The proteins were purified using Ni-NTA agarose (Qiagen) and size exclusion chromatography using Superdex 200 (16/60) column (GE Healthcare) followed by the dialysis against 10 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.5.

### Crystallization

For crystallization, LSSmOrange was transferred to a buffer containing 10 mM Tris-HCl, 100 mM NaCl, 2.5 mM EDTA pH 8.0 and were concentrated to 25 mg/ml. An initial search for

**Table 2.** Data collection and refinement statistics.

| Protein                              | PSmOrange         | PSmOrange2              | LSSmOrange              |
|--------------------------------------|-------------------|-------------------------|-------------------------|
| Space group                          | <i>P</i> 1        | <i>P</i> 2 <sub>1</sub> | <i>P</i> 2 <sub>1</sub> |
| Unit cell parameters                 |                   |                         |                         |
| a, b, c (Å)                          | 45.4, 50.7, 52.4  | 45.7, 43.9, 52.5        | 37.4, 107.4, 56.6       |
| α, β, γ (°)                          | 94.7, 90.0, 106.6 | 90.0, 94.4, 90.0        | 90.0, 102.2, 90.0       |
| Temperature (K)                      | 100               | 100                     | 100                     |
| Wavelength (Å)                       | 1.08              | 1.08                    | 1                       |
| Resolution (Å)                       | 30.0–1.95         | 30.0–1.30               | 30.0–1.40               |
| Total reflections                    | 87,470            | 158,987                 | 357,447                 |
| Unique reflections                   | 26,506            | 46,761                  | 84,546                  |
| Completeness <sup>a</sup> (%)        | 84.6 (56.5)       | 96.3 (84.2)             | 98.6 (97.1)             |
| I/σ(I)                               | 9.9 (7.4)         | 12.2 (3.5)              | 24.8 (1.9)              |
| R-merge <sup>b</sup>                 | 0.055 (0.176)     | 0.050 (0.361)           | 0.056 (0.628)           |
| Multiplicity                         | 3.3 (2.9)         | 3.4 (2.2)               | 4.2 (4.1)               |
| No. of protein atoms                 | 3,513             | 1,929                   | 3,812                   |
| No. of solvent atoms                 | 161               | 212                     | 301                     |
| Resolution range (Å)                 | 30.0–1.95         | 30.0–1.30               | 30.0–1.40               |
| R-work <sup>c</sup>                  | 0.2               | 0.148                   | 0.147                   |
| R-free <sup>c</sup>                  | 0.266             | 0.177                   | 0.175                   |
| R.m.s.d. bond lengths (Å)            | 0.018             | 0.014                   | 0.014                   |
| R.m.s.d. angles (°)                  | 2.15              | 1.59                    | 1.98                    |
| R.m.s.d. chirality (Å <sup>3</sup> ) | 0.13              | 0.08                    | 0.159                   |
| R.m.s.d. planarity (Å)               | 0.011             | 0.007                   | 0.01                    |
| Ramachandran statistics (%)          |                   |                         |                         |
| (for non-Gly/Pro residues)           |                   |                         |                         |
| most favorable                       | 92.7              | 93.9                    | 95                      |
| additional allowed                   | 7.3               | 6.1                     | 5                       |

<sup>a</sup>Data in parentheses are given for the outermost resolution shells: 1.98–1.95 Å for PSmOrange, 1.32–1.30 Å for PSmOrange2, and 1.45–1.40 Å for LSSmOrange.

<sup>b</sup> $R_{\text{merge}} = \sum_{hkl} \sum_j |I_j(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_j \langle I(hkl) \rangle$ , where  $I_j$  is the intensity measurement for reflection  $j$  and  $\langle I \rangle$  is the mean intensity over  $j$  reflections.

<sup>c</sup> $R_{\text{work}}/(R_{\text{free}}) = \sum |F_o(hkl) - |F_c(hkl)|| / \sum |F_o(hkl)|$ , where  $F_o$  and  $F_c$  are observed and calculated structure factors, respectively. No  $\sigma$ -cutoff was applied. 5% of the reflections were excluded from refinement and used to calculate  $R_{\text{free}}$ .

doi:10.1371/journal.pone.0099136.t002

crystallization conditions was carried out using a Mosquito Robotic Crystallization System (TTP LabTech Ltd). Large-scale crystallization was set up by the hanging drop vapor diffusion method at room temperature. The best crystals for LSSmOrange were obtained from Wizard I crystallization kit (Emerald BioSystems), condition 48 (0.2 M Zn acetate, 0.1 M acetate buffer pH 4.5, 20% PEG 1000). Suitable for data acquisition crystals of PSmOrange and PSmOrange2 were obtained using the CompAS crystallization suite (Qiagen), conditions C4 (55% MPD) and B8 (0.1 M HEPES, pH 7.5, 20% PEG 10000), respectively, in vapor diffusion sitting drop configuration.

## Data Acquisition

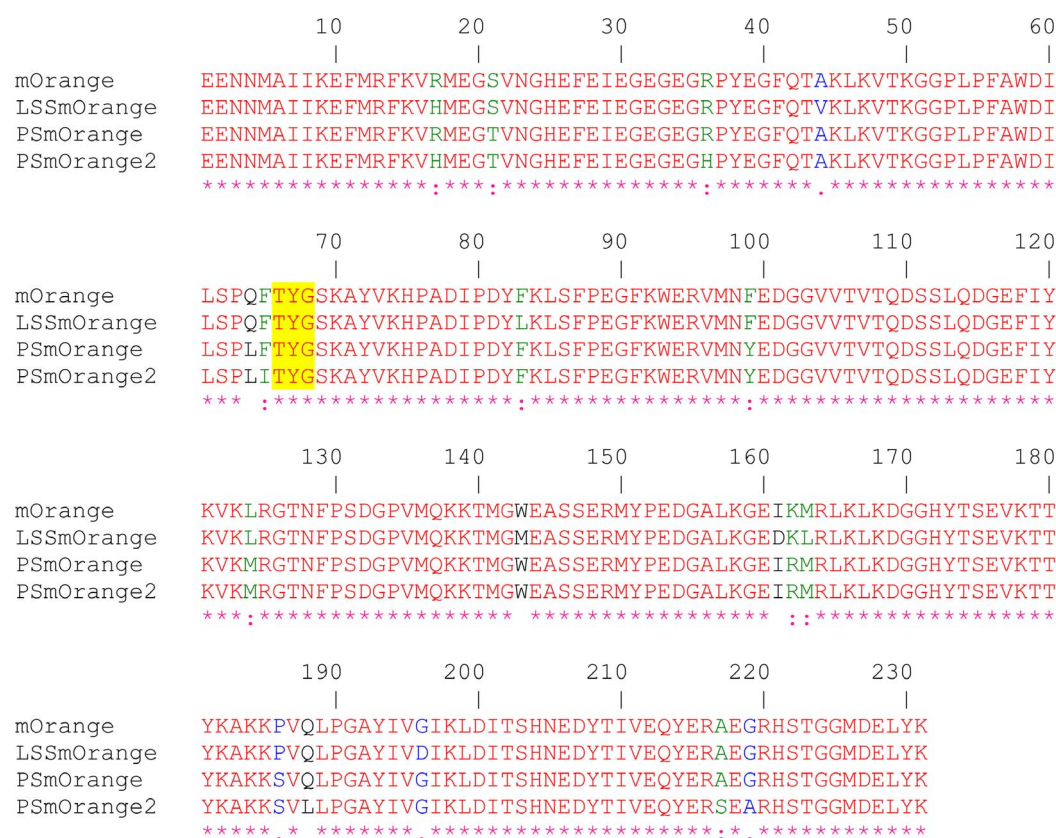
For LSSmOrange, crystals X-ray diffraction data were collected at the Advanced Photon Source on SER-CAT 22-ID beamline (Argonne National Laboratory). Diffraction intensities were registered on a MAR 300 CCD detector (Rayonix). Prior to data collection the crystals were incubated in a cryoprotecting solution consisting of 20% glycerol and 80% of well solution for 30 seconds and were flash-frozen in 100 K nitrogen stream. Cryogenic temperature was maintained by a CryoJetXL cooling device

(Oxford Cryosystems). Diffraction images were indexed, integrated and scaled with the HKL2000 software [22].

Diffraction data of the PSmOrange and PSmOrange2 crystals were collected on a Quantum 315 CCD detector (Area Detector Systems) on the X29A beamline (National Synchrotron Light Source, Brookhaven National Laboratory). PSmOrange and PSmOrange2 crystals were mounted directly from the screening trays. Prior to freezing, 20% glycerol was added to PSmOrange2 crystals as cryoprotectant. Intensities were integrated using HKL2000 and reduced to amplitudes using TRUNCATE [23,24]. Data processing statistics are given in Table 2.

## Structure Solution and Refinement

The structure of LSSmOrange was solved by molecular replacement method with MOLREP [25] using as a search model a single monomer of mOrange (PDB ID: 2H5O, [8]), excluding its chromophore. Structure refinement was performed with REFMAC [26,27] and COOT [28]. Manual structure rebuilding and addition of ordered solvent molecules were done using COOT. Structure validation was performed with COOT and PROCHECK [29].



**Figure 2. Amino acid alignment of mOrange, LSSmOrange, PSmOrange and PSmOrange2.** The chromophore-forming tri-peptides are highlighted in yellow.  
doi:10.1371/journal.pone.0099136.g002

The structures of PSmOrange and PSmOrange2 were determined by molecular replacement using PHASER [30]. Model building and refinement were performed with REFMAC and COOT. The quality of the final structure was verified with composite omit maps, and the stereochemistry was checked with MOLPROBITY [31]. The refinement statistics for LSSmOrange, PSmOrange and PSmOrange2 are given in Table 2.

## Results

LSSmOrange and PSmOrange have been engineered from mOrange and PSmOrange2 has been engineered from PSmOrange using directed molecular evolution. The resulted OFPs constitute the following mutants: mOrange/R17H/A44V/F83L/W143M/I161D/M163L/G196D (LSSmOrange), mOrange/S21T/Q64L/F99Y/L124M/K162R/P186S (PSmOrange), and PSmOrange/R17H/R36H/F65I/Q188L/A217S/G219A (PSmOrange2) (Figures 2, 3). Here and below the numbering is adopted from mOrange PDB file, 2H5O [8]. Note that not all of the listed mutations are located in the nearest chromophore environment. In fact, only three residues in LSSmOrange (M143, D161, and L163), one in PSmOrange (I64), and two in PSmOrange2 (I65 and S217) are situated close enough to the chromophore to directly affect its electronic properties (Figure 3).

## Crystal Structures of LSSmOrange, PSmOrange, and PSmOrange2

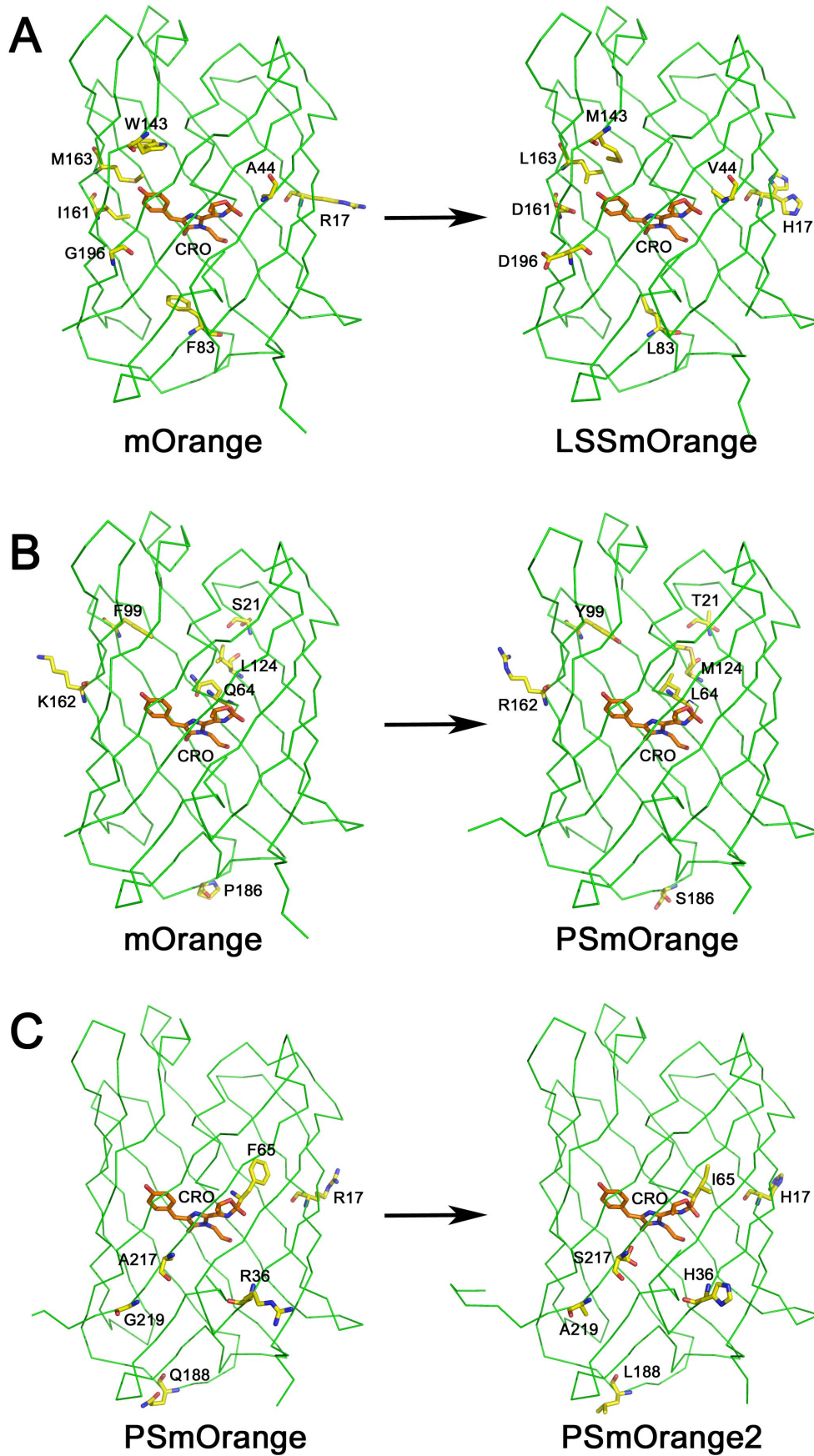
The asymmetric units of LSSmOrange and PSmOrange crystals contain two monomers oriented approximately parallel to each

other, whereas the asymmetric unit of PSmOrange2 contains a single monomer. Crystallographic symmetry operations do not complete monomers to classic GFP-like tetramers indicating a true monomeric nature of the proteins. Multiple conformations are observed for the total of 30 (6.5%) amino acid residues of LSSmOrange and for 26 (11.3%) amino acid residues of PSmOrange2. Multiple conformations for PSmOrange were not modeled due to the relatively low (1.94 Å) resolution of the diffraction data.

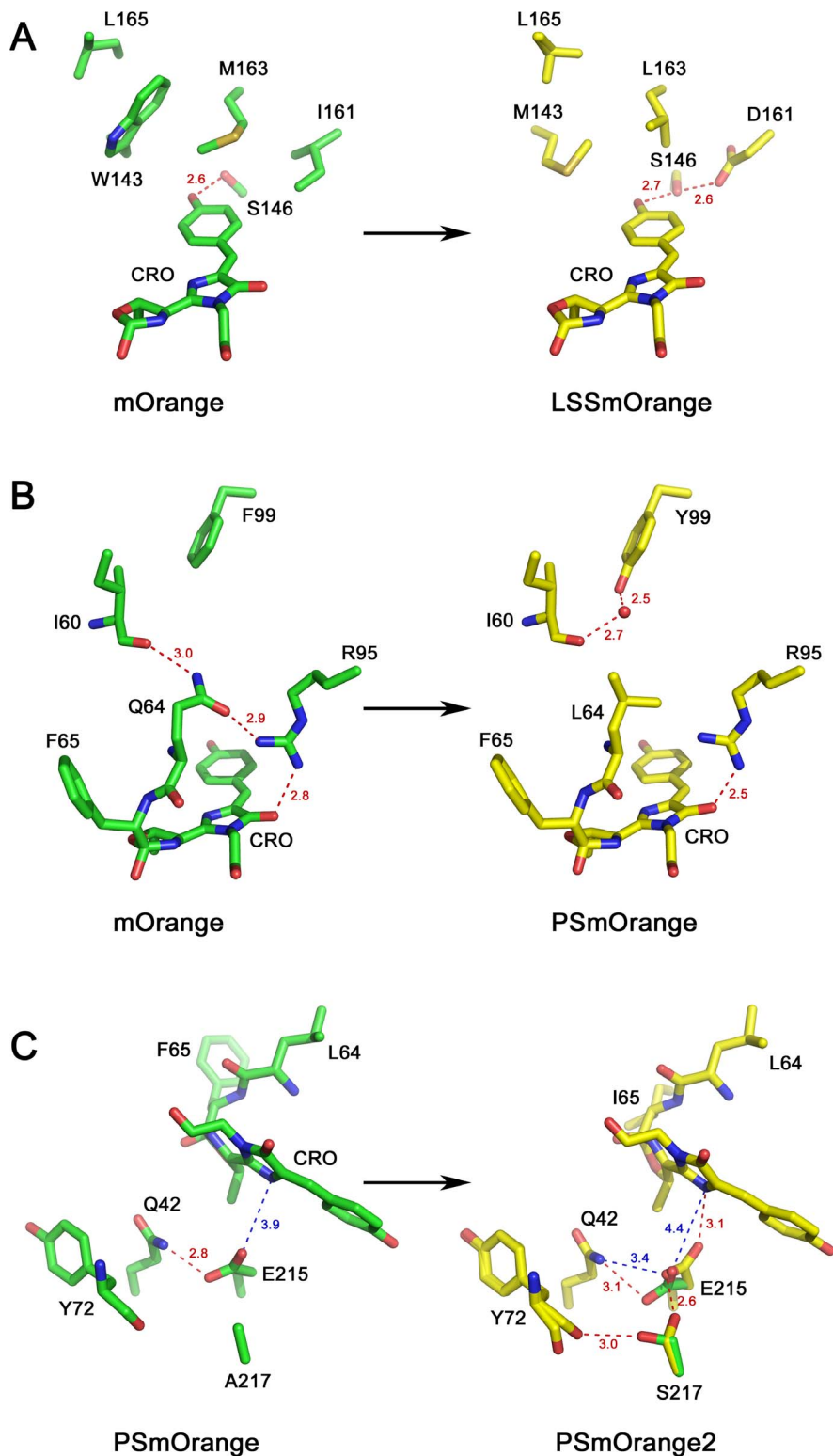
The final models of LSSmOrange, PSmOrange, and PSmOrange2 have low deviations of bond length, angles, chiral volumes and planes from ideal values indicating its high quality. In all models, over 90% of the residues are located in the most favorable regions of the Ramachandran plot with the rest located in the additionally allowed regions (Table 2). Superposition of parental mOrange with LSSmOrange, PSmOrange and PSmOrange2 by C $\alpha$  atoms resulted in root mean square deviations of 0.35 Å, 0.39 Å, and 0.38 Å, respectively, indicating close similarity of the structures.

## Chromophores of LSSmOrange, PSmOrange, and PSmOrange2 and their Environment

Electron density maps show that LSSmOrange, PSmOrange, and PSmOrange2 have the same chromophore structures as parental mOrange [8]. It consists of *p*-hydroxyphenyl, imidazolinone, and 2-hydroxy-dihydrooxazole rings. The latter one results from cyclization of T66 side chain. The chromophore adopts *cis*-conformations and in all examined FPs has a noticeable but not severe deviation from coplanarity between the adjacent *p*-



**Figure 3. Amino acid differences between the parental and successor proteins in 3D. (A)** The transformation of mOrange into LSSmOrange. **(B)** The transformation of mOrange into PSmOrange. **(C)** The transformation of PSmOrange into PSmOrange2.  
doi:10.1371/journal.pone.0099136.g003



**Figure 4. The differences in the immediate chromophore environment between the parental and successor proteins. (A)** The difference between mOrange and LSSmOrange. **(B)** The difference between mOrange and PSmOrange. **(C)** The difference between PSmOrange and PSmOrange2.  
doi:10.1371/journal.pone.0099136.g004

hydroxyphenyl and imidazolinone rings ( $\chi_1$ , the rotation around C $\alpha$ 2-C $\beta$ 2 bond, is  $\sim 5$ – $10^\circ$ ). The largest deviation from coplanarity is observed between 2-hydroxy-dihydrooxazole and imidazolinone rings of the chromophore. The C $\alpha$ 1 atom of the chromophore has partially pyramidal geometry with the sum of the valent angles around it of  $\sim 348^\circ$ . Due to a distorted geometry of C $\alpha$ 1 atom, these rings are oriented at the angle of  $\sim 30^\circ$  with respect to each other. The deviation of C $\alpha$ 1 atom from planar  $sp^2$ -geometry is presumably caused by a local steric strain between the chromophore and F65 (preceding amino acid residue).

Key substitutions in LSSmOrange responsible for its  $\sim 135$  nm large Stokes shift comprise I161D, M163L, and W143M enabling the excited state proton transfer (ESPT) [19]. These substitutions introduce noticeable changes in the nearest chromophore environment of LSSmOrange (Figure 4A). X-ray structure shows that W143M replacement provides a shift of the side chain of L165 towards protein interior, and both W143M and M163L substitutions improve the chromophore planarity relative to parental mOrange (Figure 4A). The  $\chi_1$  angles of mOrange and LSSmOrange chromophores are  $12.3^\circ$  and  $6.3^\circ$ , respectively. In both mOrange and LSSmOrange, the oxygen atom of *p*-hydroxyphenyl group of the chromophore forms an H-bond with the side chain hydroxyl of S146. Rotation of the chromophore around C $\alpha$ 2-C $\beta$ 2 bond in LSSmOrange, arising from W143M and M163L mutations, causes the change of S146 side chain conformation. The hydroxyl atom of S146 now forms two H-bonds, one with the chromophore and another with the side chain of D161 (I161 in mOrange). As a result, S146 acts as a mediator for ESPT from *p*-hydroxyphenyl ring of the chromophore to carboxyl group of D161 (Figure 4A). Four other substitutions, R17H, A44V, F83L, and G196D, are located far from the chromophore and presumably improve folding efficiency, brightness and photostability (Figure 3A).

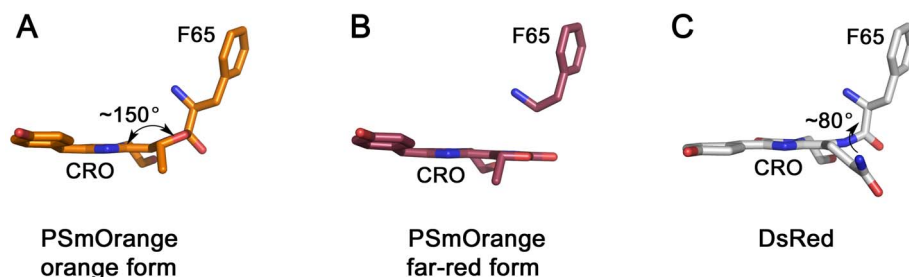
The key substitutions enabling PSmOrange photoconversion are Q64L and F99Y introduced in a process of molecular evolution via random and rational mutagenesis of parental mOrange [20]. In mOrange, Q64 forms H-bonds with guanidine group of R95 and the main chain carbonyl of I60 (Figure 4B). Substitution of Q64 with Leu disables formation of these two H-bonds (Figure 4B). The side chain carboxyl of Y99 forms new water-mediated H-bond with the main chain carbonyl of I60. These two changes result in the protein that in the presence of oxidizing agents such as potassium ferricyanide or intracellular oxidants and light could be converted into a far-red emitting form [20]. Out of six substitutions made in mOrange to obtain PSmOrange Q64L is the only one located close enough to the chromophore to cause structural changes enabling photoconversion. Five other substitutions, S21T, F99Y, L124M, K162R, and

P186S, lie far away from the chromophore, presumably affecting its folding and brightness (Figure 3B).

Analogously to PSmOrange, in PSmOrange2, position 64 is occupied by Leu pointing out at the importance of a hydrophobic residue in this position. PSmOrange2 differs from PSmOrange by total six substitutions, only two of which, F65I and A217S, are located close to the chromophore (Figure 3C). Replacement of aromatic F65 (present in both mOrange and PSmOrange) by aliphatic I65 is accompanied by a dramatic decrease of the amount and intensity of light required for an efficient photoconversion. Four other residues are located on the surface of the protein. S217 in PSmOrange2, forms a strong H-bond with catalytic E215 and both, S217 and E215, adopt two conformations (Figure 4C). In conformation “A”, the side chain of S217 is oriented away from E215 and forms a H-bond with the main chain carbonyl of Y72. This makes E215 move away from the chromophore, causes break of a H-bond between O $\epsilon$ 2 atom of E215 and N2 atom of the chromophore (4.4 Å), and provides formation of a new H-bond between the side chains of E215 and Q42. In conformation “B”, S217 points towards E215, forming a strong H-bond (2.6 Å) with its side chain and pushing E215 closer to the chromophore, facilitating H-bonding between E215 and nitrogen of imidazolinone ring (3.1 Å) that connects it with the protein matrix. In parental PSmOrange, position 217 is occupied by Ala, as a result, E215 is moved away from the chromophore forming no H-bond with it (E215-O $\epsilon$ 2 – TYG-N2 distance is 3.9 Å) (Figure 4C).

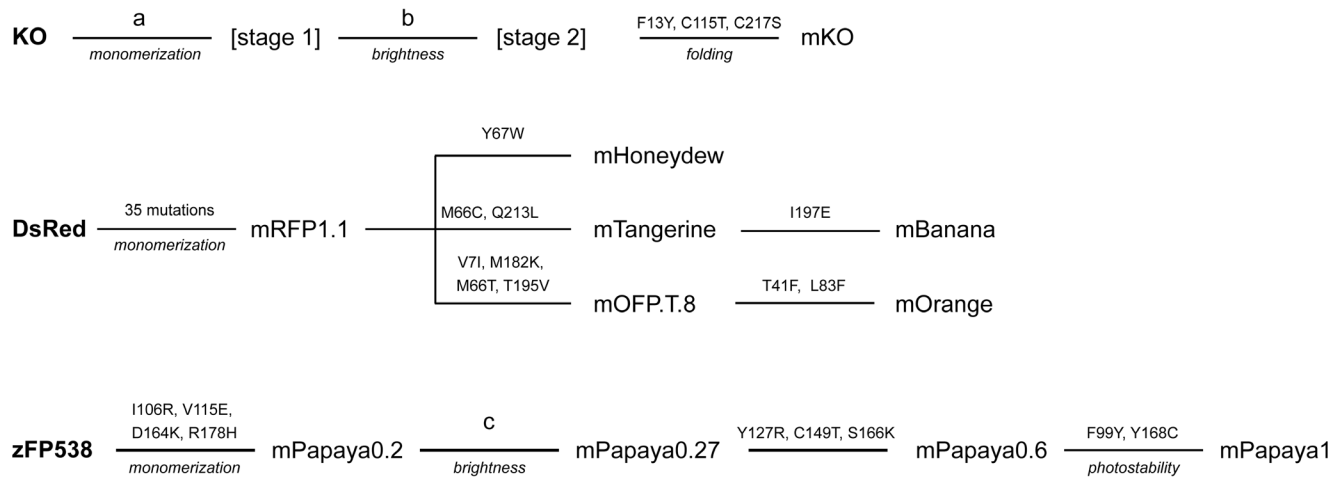
Photoconverted far-red forms of PSmOrange and PSmOrange2 are relatively unstable and undergo 50% degradation within several days after photoconversion, presumably due to the overall oxidation and degradation of the protein [20,21]. This instability made impossible obtaining the suitable crystals of the far-red forms. Photoconversion of the crystals of PSmOrange and PSmOrange2 attempted in crystallization drops with added potassium ferricyanide was also unsuccessful presumably due to a poor diffusion of the oxidant in the crystals.

In an earlier work on PSmOrange, it was shown that mass-spectrometry analysis of the photoconverted chromophore-containing peptide was in a good agreement with the mechanism of PSmOrange photoswitching involving a cleavage of the polypeptide chain between the main chain carbonyl and C $\alpha$  of F65 [20]. It has been suggested that the PSmOrange2 photoconversion is similar and may consist of a break of the polypeptide chain between the main chain carbonyl and C $\alpha$  of I65. [21]. To get an idea of the structural changes taking place in far-red forms of PSmOrange and PSmOrange2, we have modeled their structures based on the corresponding non-photoconverted proteins (Figure 5). The modeling revealed that the cleavage of C $\alpha$ -C



**Figure 5. The structures of the PSmOrange and DsRed chromophores.** (A) The structure of the orange form of the PSmOrange chromophore. (B) The modeled structure of the photoconverted far-red form of the PSmOrange chromophore. (C) The structure of the DsRed chromophore.

doi:10.1371/journal.pone.0099136.g005



a = F102S, A104S, V123T, C151S, F162Y, F193Y, and G195S

b = K11R, V25I, K32R, S55A, T62V, Q96E, E117Y, V133I, S139V, T150A, A166E, and Q190G

c = E2G, M14V, Y36N, A63P, G94E, N127Y, and M153V

**Figure 6. An evolution of orange fluorescent proteins derived from KO, DsRed and zFP538 with mutations critical to the phenotype of each variant.**

doi:10.1371/journal.pone.0099136.g006

bond of F65/I65 eliminates the strain imposed on the chromophore by internal  $\alpha$ -helix of the protein, causing C $\alpha$ 1 atom of the chromophore to adopt a true planar  $sp^2$ -geometry (Figure 5B). As a result, the oxazole ring of the chromophore moves away from F65/I65 and becomes positioned in-plane with two other rings, significantly improving the efficiency of the overall chromophore conjugation. Oxidation of the hydroxyl of 2-hydroxy-dihydrooxazole to carbonyl puts this group in plane with the rest of the chromophore; it faces (but is not embedded in) the hydrophobic pocket formed by the residues F14, V16, A44, L46, F/I65, and Y120. Appearance of C=O group extends the chromophore conjugation system, providing an additional contribution to the red shift of the fluorescence of the photoconverted PSmOrange and PSmOrange2.

## Discussion

Continuing progress in fluorescence imaging requires probes with additional colors and properties optimized for emerging imaging techniques. A typical process for the development of FPs with desired properties includes rational design followed by random mutagenesis. Rational design relies on the accumulated knowledge about the influence of the certain amino acid residues in the immediate chromophore environment on FP properties. At this stage, first key mutations responsible for the desired target properties of the FP are introduced. This minimally essential variant is then subjected to a directed molecular evolution aimed at optimization of an overall performance of FP.

Development of monomeric variant of KO started from the introduction of seven point mutations on the surface of the protein (F102S, A104S, V123T, C151S, F162Y, F193Y, and G195S) known to disrupt the tetrameric interfaces of DsRed [32]. Here and below the numbering for amino acids is adopted from original publications. Additional mutagenesis that followed found two other sets of mutations, twelve (K11R, V25I, K32R, S55A, T62V, Q96E, E117Y, V133I, S139V, T150A, A166E, and Q190G) and

three (F13Y, C115T, and C217S), to improve mKO's brightness and folding efficiency, respectively [4] (Figure 6).

The development of monomeric FPs was also started from the disruption of the interfaces of tetrameric DsRed to yield mRFP1.1 [2] that was further used as a template for the development of monomeric FPs of multiple colors [2]. Figure 6 shows the key residues providing orange emission of FPs derived from DsRed. Y67W is the key substitution converting red mRFP1.1 into orange FP, mHoneydew. Two substitutions, M66C and Q213L, are required to transform mRFP1.1 in mTangerine, which in turn can be converted into mBanana by additional I197E mutation. Four point mutations, V7I, M182K, M66T, and T195V, are required to convert mRFP1.1 in the OFP prototype, mOFP.T.8. Two additional mutations, T41F and L83F, converted mOFP.T.8 in a highly acclaimed bright OFP mOrange. In all cases, rational design has been followed by directed evolution approach to perfect newly generated FPs - enhance their brightness, folding efficiency, and photostability.

The other vivid example of rational design and directed evolution approach combination is mPapaya1 [10] (Figure 6). First, to design monomeric version of zFP538, four mutations (I106R, V115E, D164K, and R178H) were introduced at the interfaces between its subunits (mPapaya0.2). Then, to restore the brightness, lost during monomerization, the protein was subjected to four rounds of directed evolution that introduced seven additional mutations (mPapaya0.27). To further improve monomeric character of mPapaya0.27, additional three mutations, Y127R, C149T, and S166K were made resulting in truly monomeric mPapaya0.6. Finally, to improve photostability mPapaya0.6 was subjected to additional cycle of directed evolution resulting in mPapaya0.6/F99Y/Y168C designated as mPapaya1.

Bright monomeric FPs serve as templates for the development of FPs with special properties, such as FPs with a large Stokes shift, photoswitching FPs and Fluorescent timers [1]. Currently available LSSFPs include green T-Sapphire [33], yellow mAmber [34], orange LSSmOrange [19], and red LSSmKate1 [35],

LSSmKate2 [35], and mKeima1 [36]. Engineering of LSSFPs aims at providing an ESPT between Tyr hydroxyl of the chromophore and its nearest environment. The common mechanism implies that the chromophore is initially protonated. Upon excitation, the proton is transferred to a nearby proton acceptor either directly or with the help of an H-bond network. The anionic chromophore intermediate emits a photon, returns to the ground state, and becomes protonated.

In wild-type GFP, two ESPT pathways have been reported: the major one follows a well-defined chromophore-Water22-S205-E222 proton wire [37], while a less efficient one is based on a change of T203 conformation promoting proton ejection out of the  $\beta$ -barrel [38]. In the ground state, the chromophore is reprotonated by E222 acquiring a proton from the outside by the E5 entry pathway, located near the N-terminus of GFP [38]. In green T-Sapphire and yellow mAmetrine, ESPT follows the same pathway as the major one of wtGFP. In LSSmKate1, ESPT is realized as the proton shuttle between the *cis*-chromophore and E160, whereas in LSSmKate2, ESPT pathway comprises *trans*-chromophore, S158, and D160 [39]. Similarly, in mKeima, proton is transferred along the proton wire *cis*-chromophore-S146-D161 [40], [41]. All elements of the mKeima proton wire are present in LSSmOrange (*cis*-chromophore) indicating to a similar mechanism of ESPT; this suggestion has been confirmed by site-directed mutagenesis. Earlier, we have demonstrated that both LSSmOrange/S146A and LSSmOrange/D161A mutants disturb LSS phenotype and LSSmOrange/D161A displays a regular Stokes shift ( $\lambda_{ex}^{max} / \lambda_{em}^{max}$  550/565 nm) similar to that of mOrange [19]. Parental mOrange has position 146 already occupied by Ser, which is one of two essential residues required for ESPT.

PSFPs are widely used for tracking intracellular proteins, organelles, and individual cells [42], [43]. Fluorescence of PSFPs is switched from one color to another by the light of a specific wavelength. A majority of PSFPs including Dendra2 [44], mEos2 [45], Kaede [46], mKikGR [47], mClavGR2 [48] and their derivatives change their fluorescence from green to red after irradiation with a relatively phototoxic violet light (390–410 nm). The blue-to-green proteins photoswitchable with this light, PSCFP and PSCFP2, are also available [44]. Photoconversion of PSmOrange and PSmOrange2 with 480–540 nm light makes these proteins the first PSFPs efficiently photoswitched with non-phototoxic visible light [20], [21]. SDS-PAGE and mass spectrometry results demonstrated that PSmOrange photoconversion is accompanied by a cleavage of polypeptide chain between the chromophore and F65 and by oxidation of the hydroxyl of 2-hydroxy-dihydrooxazole ring of the chromophore to carbonyl [20]. This modification strongly affects the conjugation system of the chromophore - it enables the chromophore to recover its distorted planarity around C $\alpha$ 1 atom and adds C=O bond to the conjugation resulting in a substantial (~90 nm) bathochromic shift of photoconverted PSmOrange absorbance/emission bands (Figure 5B). This is not the case for the other RFPs, such as DsRed (*Discosoma* sp.) ( $\lambda_{ex}^{max} / \lambda_{em}^{max}$  558/583 nm) [5], eqFP578 (*Entacmaea quadricolor*) ( $\lambda_{ex}^{max} / \lambda_{em}^{max}$  552/578 nm) [49], or zFP574 (*Zoanthus* sp.) ( $\lambda_{ex}^{max} / \lambda_{em}^{max}$  553/574 nm) [50]: in all these RFPs, carbonyl of the amino acid preceding the chromophore is not coplanar with it (Figure 5C). In fact, for these proteins a ~80° dihedral angle between the carbonyl group and the chromophore makes the conjugation extremely inefficient. This carbonyl cannot be coplanar with the chromophore due to the bent in the central  $\alpha$ -helix embedding the chromophore. Inefficient conjugation results in a blue-shift of fluorescence of DsRed, eqFP578, and zFP574 relative to that of photoswitched PSmOrange.

It has been demonstrated that position 64 mutation is minimally required for appearance of photoswitchable properties [20]. In mOrange, Q64 forms an H-bond with R95; this H-bond is absent in PSmOrange, where Q64 is replaced with hydrophobic Leu (Figure 4B), suggesting that the absence of H-bond between residues 64 and 95 facilitates orange to far-red photoconversion of PSmOrange chromophore.

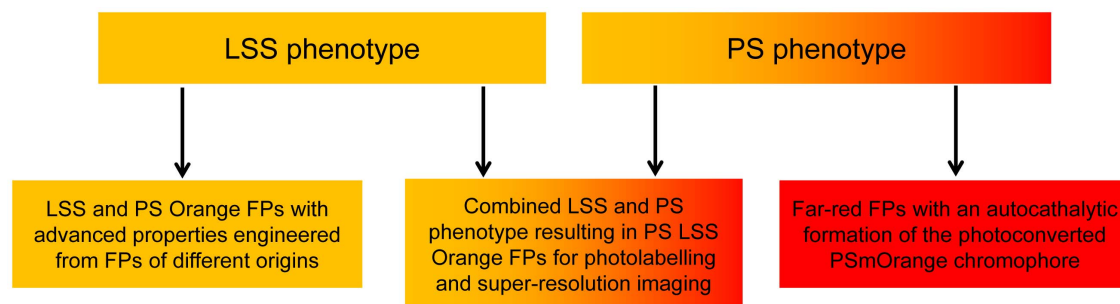
Monomeric PSmOrange2 has been developed as an improved version of PSmOrange with faster and more efficient photoswitching. Its photoconversion could be achieved with common two-photon lasers and it is, therefore, much more user-friendly than parental PSmOrange. Moreover, PSmOrange2 could be used as an acceptor in Förster resonance energy transfer with green fluorescent donors. This fact together with its high efficiency of photoconversion and fast photoswitching kinetics enabled photoswitching of PSmOrange2 via FRET from the donor FP. The brightness of orange and far-red forms of PSmOrange2 is 1.9-fold and 1.2-fold lower than that of respective PSmOrange forms (Table 1). However, its photoswitching contrast, which is 9-fold higher than that of PSmOrange, a substantially higher efficiency of PSmOrange2 photoconversion than that of the parental protein makes PSmOrange2 a better tag. The key mutation that provided dramatic improvement of PSmOrange2 photoconversion is F65L. Replacement of aromatic F65 with aliphatic I65 and decrease of the light energy required for photoswitching imply that I65 facilitates backbone cleavage and decreases the activation barrier of photoconversion. Modeling of photoconverted PSmOrange structure revealed that the residue 65 and the oxidized chromophore move away from each other after the cleavage of C $\alpha$ 1-C bond (Figure 5B). A lower energy required for a photoconversion of PSmOrange2 suggests that less bulky side chain of I65 enables an easier stabilization of detached Ile within its immediate environment than of a bulkier and more rigid F65 in PSmOrange. Substitution A217S increases quantum yield of the orange form of PSmOrange2 [21]. This is presumably caused by an increased charge separation in the chromophore, taking place as a result of a new H-bond between E215 and its N2 atom. Formation of this H-bond occurs only in the presence of S217 that makes E215 move closer to the chromophore. A similar H-bond between N2 and E215 has been described earlier for rsTagRFP where it enabled formation of a fluorescent zwitter-ionic species [51].

## Conclusions

Here we presented the structural analysis of three recently developed OFPs with novel phenotypes such as a large Stokes shift and photoswitchability. Spectroscopic characteristics provided the insights in the properties of these OFPs and extensive mutagenesis identified the residues responsible for the photochemical changes. Crystallographic structures presented here revealed how the ensemble of key mutations introduced in the parental FPs enabled those changes.

For LSSmOrange, ESPT occurs over the *cis*-chromophore-S146-D161 pathway and is caused by three key substitutions: I161D, M163L, and W143M. The W143M replacement provided a shift of the side chain of L165 towards protein interior, and both W143M and M163L improve the chromophore planarity. Rotation of the chromophore around  $\chi$ 1 angle, arising from these two mutations, causes the change of S146 side chain conformation. As a result, S146 acts as a mediator for ESPT from *p*-hydroxyphenyl ring of the chromophore to carboxyl group of D161.

For PSmOrange and PSmOrange2, Q64L mutation is minimally required for appearance of photoswitchable properties. The



**Figure 7. Possibilities for the future design of advanced orange and far-red fluorescent proteins.**  
doi:10.1371/journal.pone.0099136.g007

structural data demonstrated that this mutation disrupts an H-bond between positions 64 and R95 facilitating orange to far-red photoconversion of PSmOrange chromophore. The photoconversion itself is accompanied by cleavage of C $\alpha$ 1-C bond of F/I65, resulting in recovery of the distorted chromophore planarity, and by oxidation of OH-group of 2-hydroxy-dihydrooxazole to C=O, resulting in the extension of the chromophore conjugation system. This was the primary source of a substantial (~90 nm) bathochromic red shift of fluorescence in photoconverted PSmOrange and PSmOrange2. Substitution A217S in PSmOrange2 induces the shift of E215 closer to the chromophore and cause formation of its H-bond with N2 atom, increasing polarization of the chromophore and its quantum yield.

The detailed analysis of the chromophore structures and chromophore transformations in LSSmOrange and PSmOrange/PSmOrange2 will serve as a basis for engineering of future advanced FPs (Figure 7). First, enhanced OFP variants could be designed by introducing ESPT or photoconversion pathways in OFPs of different origins including wild-type OFPs. The key positions for mutations (Figure S1) can be used to introduce substitutions in new protein templates at the first step, followed by random mutagenesis. Second, novel OFP phenotypes can be obtained by combining different pathways of OFP chromophore transformations in a single FP. We anticipate that the LSS phenotype can be combined with PS phenotypes to engineer PS LSS OFP. This future FP should further advance the multicolor

photolabelling in wide-field fluorescence microscopy and in super-resolution imaging of live cells. Lastly, another engineering goal could be a development of a FP with the autocatalytic formation of the photoconverted far-red PSmOrange chromophore. Such permanently fluorescent far-red FP is highly desirable as a genetically-encoded probe for non-invasive deep-tissue *in vivo* imaging.

The coordinates and structure factors were deposited in the Protein Data Bank under the accession codes 4Q7R, 4Q7T and 4Q7U for LSSmOrange, PSmOrange and PSmOrange2, respectively.

## Supporting Information

**Figure S1 Amino acid alignment of orange fluorescent proteins.** The chromophore-forming tri-peptides are highlighted in yellow.  
(DOC)

## Author Contributions

Performed the experiments: SP DMS OMS NVP VNM SCA ZD VVV. Analyzed the data: SP DMS OMS NVP VNM SCA ZD VVV. Contributed reagents/materials/analysis tools: SP DMS OMS NVP VNM SCA ZD VVV. Contributed to the writing of the manuscript: SP DMS OMS NVP VNM SCA ZD VVV.

## References

- Shcherbakova DM, Subach OM, Verkhusha VV (2012) Red fluorescent proteins: advanced imaging applications and future design. *Angew Chem Int Ed Engl* 51: 10724–10738.
- Shaner NC, Campbell RE, Steinbach PA, Giepmans BN, Palmer AE, et al. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat Biotechnol* 22: 1567–1572.
- Matz MV, Fradkov AF, Labas YA, Savitsky AP, Zaraisky AG, et al. (1999) Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat Biotechnol* 17: 969–973.
- Karasawa S, Araki T, Nagai T, Mizuno H, Miyawaki A (2004) Cyan-emitting and orange-emitting fluorescent proteins as a donor/acceptor pair for fluorescence resonance energy transfer. *Biochem J* 381: 307–312.
- Yarbrough D, Wachter RM, Kallio K, Matz MV, Remington SJ (2001) Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-Å resolution. *Proc Natl Acad Sci U S A* 98: 462–467.
- Shagin DA, Barsova EV, Yanushevich YG, Fradkov AF, Lukyanov KA, et al. (2004) GFP-like proteins as ubiquitous metazoan superfamily: evolution of functional features and structural complexity. *Mol Biol Evol* 21: 841–850.
- Remington SJ, Wachter RM, Yarbrough DK, Branchaud B, Anderson DC, et al. (2005) zFP538, a yellow-fluorescent protein from *Zoanthus*, contains a novel three-ring chromophore. *Biochemistry* 44: 202–212.
- Shu X, Shaner NC, Yarbrough CA, Tsien RY, Remington SJ (2006) Novel chromophores and buried charges control color in mFruits. *Biochemistry* 45: 9639–9647.
- Pletneva NV, Pletnev SV, Chudakov DM, Tikhonova TV, Popov VO, et al. (2007) Three-dimensional structure of yellow fluorescent protein zYFP538 from *Zoanthus* sp. at the resolution 1.8 angstrom. *Bioorg Khim* 33: 421–430.
- Hoi H, Howe ES, Ding Y, Zhang W, Baird MA, et al. (2013) An engineered monomeric *Zoanthus* sp. yellow fluorescent protein. *Chem Biol* 20: 1296–1304.
- Pletneva NV, Pletnev VZ, Souslova E, Chudakov DM, Lukyanov S, et al. (2013) Yellow fluorescent protein phiYFPv (Phialidium): structure and structure-based mutagenesis. *Acta Crystallogr D Biol Crystallogr* 69: 1005–1012.
- Kikuchi A, Fukumura E, Karasawa S, Mizuno H, Miyawaki A, et al. (2008) Structural characterization of a thiazoline-containing chromophore in an orange fluorescent protein, monomeric Kusabira Orange. *Biochemistry* 47: 11573–11580.
- Sakaue-Sawano A, Kurokawa H, Morimura T, Hanyu A, Hama H, et al. (2008) Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell* 132: 487–498.
- Tsutsui H, Karasawa S, Okamura Y, Miyawaki A (2008) Improving membrane voltage measurements using FRET with new fluorescent proteins. *Nat Methods* 5: 683–685.
- Shaner NC, Lin MZ, McKeown MR, Steinbach PA, Hazelwood KL, et al. (2008) Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nat Methods* 5: 545–551.
- Strack RL, Strongin DE, Bhattacharyya D, Tao W, Berman A, et al. (2008) A noncytotoxic DsRed variant for whole-cell labeling. *Nat Methods* 5: 955–957.
- Strack RL, Hein B, Bhattacharyya D, Hell SW, Keenan RJ, et al. (2009) A rapidly maturing far-red derivative of DsRed-Express2 for whole-cell labeling. *Biochemistry* 48: 8279–8281.
- Strack RL, Bhattacharyya D, Glick BS, Keenan RJ (2009) Noncytotoxic orange and red/green derivatives of DsRed-Express2 for whole-cell labeling. *BMC Biotechnol* 9: 32.

19. Shcherbakova DM, Hink MA, Joosen L, Gadella TW, Verkhusha VV (2012) An orange fluorescent protein with a large Stokes shift for single-excitation multicolor FCCS and FRET imaging. *J Am Chem Soc* 134: 7913–7923.
20. Subach OM, Patterson GH, Ting LM, Wang Y, Condeelis JS, et al. (2011) A photoswitchable orange-to-far-red fluorescent protein, PSmOrange. *Nat Methods* 8: 771–777.
21. Subach OM, Entenberg D, Condeelis JS, Verkhusha VV (2012) A FRET-facilitated photoswitching using an orange fluorescent protein with the fast photoconversion kinetics. *J Am Chem Soc* 134: 14789–14799.
22. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276: 307–326.
23. Collaborative Computational Project n4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50: 760–763.
24. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67: 235–242.
25. Vagin A, Teplyakov A (1997) *MOLREP*: an automated program for molecular replacement. *J Appl Cryst* 30: 1022–1025.
26. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53: 240–255.
27. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67: 355–367.
28. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60: 2126–2132.
29. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) *PROCHECK*: a program to check the stereochemical quality of protein structures. *J Appl Cryst* 26: 283–291.
30. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40: 658–674.
31. Davis IW, Murray LW, Richardson JS, Richardson DC (2004) MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* 32: W615–W619.
32. Campbell RE, Tour O, Palmer AE, Steinbach PA, Baird GS, et al. (2002) A monomeric red fluorescent protein. *Proc Natl Acad Sci U S A* 99: 7877–7882.
33. Zapata-Hommer O, Griesbeck O (2003) Efficiently folding and circularly permuted variants of the Sapphire mutant of GFP. *BMC Biotechnol* 3: 5.
34. Ai HW, Hazelwood KL, Davidson MW, Campbell RE (2008) Fluorescent protein FRET pairs for ratiometric imaging of dual biosensors. *Nat Methods* 5: 401–403.
35. Piatkevich KD, Hult J, Subach OM, Wu B, Abdulla A, et al. (2010) Monomeric red fluorescent proteins with a large Stokes shift. *Proc Natl Acad Sci U S A* 107: 5369–5374.
36. Kogure T, Karasawa S, Araki T, Saito K, Kinjo M, et al. (2006) A fluorescent variant of a protein from the stony coral *Montipora* facilitates dual-color single-laser fluorescence cross-correlation spectroscopy. *Nat Biotechnol* 24: 577–581.
37. Chattoraj M, King BA, Bublitz GU, Boxer SG (1996) Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. *Proc Natl Acad Sci U S A* 93: 8362–8367.
38. Agmon N (2005) Proton pathways in green fluorescence protein. *Biophys J* 88: 2452–2461.
39. Piatkevich KD, Malashkevich VN, Almo SC, Verkhusha VV (2010) Engineering ESPT pathways based on structural analysis of LSSmKate red fluorescent proteins with large Stokes shift. *J Am Chem Soc* 132: 10762–10770.
40. Violot S, Carpentier P, Blanchoin L, Bourgeois D (2009) Reverse pH-dependence of chromophore protonation explains the large Stokes shift of the red fluorescent protein mKeima. *J Am Chem Soc* 131: 10356–10357.
41. Henderson JN, Osborn MF, Koon N, Gepshtein R, Huppert D, et al. (2009) Excited state proton transfer in the red fluorescent protein mKeima. *J Am Chem Soc* 131: 13212–13213.
42. Lippincott-Schwartz J, Patterson GH (2009) Photoactivatable fluorescent proteins for diffraction-limited and super-resolution imaging. *Trends Cell Biol* 19: 555–565.
43. Wu B, Piatkevich KD, Lionnet T, Singer RH, Verkhusha VV (2011) Modern fluorescent proteins and imaging technologies to study gene expression, nuclear localization, and dynamics. *Curr Opin Cell Biol* 23: 310–317.
44. Chudakov DM, Lukyanov S, Lukyanov KA (2007) Tracking intracellular protein movements using photoswitchable fluorescent proteins PS-CFP2 and Dendra2. *Nat Protoc* 2: 2024–2032.
45. McKinney SA, Murphy CS, Hazelwood KL, Davidson MW, Looger LL (2009) A bright and photostable photoconvertible fluorescent protein. *Nat Methods* 6: 131–133.
46. Ando R, Hama H, Yamamoto-Hino M, Mizuno H, Miyawaki A (2002) An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proc Natl Acad Sci U S A* 99: 12651–12656.
47. Habuchi S, Tsutsui H, Kochaniak AB, Miyawaki A, van Oijen AM (2008) mKikGR, a monomeric photoswitchable fluorescent protein. *PLoS One* 3: e3944.
48. Hoi H, Shaner NC, Davidson MW, Cairo CW, Wang J, et al. (2010) A monomeric photoconvertible fluorescent protein for imaging of dynamic protein localization. *J Mol Biol* 401: 776–791.
49. Pletneva NV, Pletnev VZ, Shemiakina II, Chudakov DM, Artemyev I, et al. (2011) Crystallographic study of red fluorescent protein eqFP578 and its far-red variant Katushka reveals opposite pH-induced isomerization of chromophore. *Protein Sci* 20: 1265–1274.
50. Pletneva N, Pletnev S, Tikhonova T, Popov V, Martynov V, et al. (2006) Structure of a red fluorescent protein from *Zoanthus*, zRFP574, reveals a novel chromophore. *Acta Crystallogr D Biol Crystallogr* 62: 527–532.
51. Pletnev S, Subach FV, Dauter Z, Wlodawer A, Verkhusha VV (2012) A structural basis for reversible photoswitching of absorbance spectra in red fluorescent protein rsTagRFP. *J Mol Biol* 417: 144–151.
52. Patterson GH, Knobel SM, Sharif WD, Kain SR, Piston DW (1997) Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophys J* 73: 2782–2790.

# **Orange fluorescent proteins: structural studies of LSSmOrange, PSmOrange and PSmOrange2**

Sergei Pletnev, Daria M. Shcherbakova, Oksana M. Subach, Nadya V. Pletneva, Vladimir N. Malashkevich, Steven C. Almo, Zbigniew Dauter and Vladislav V. Verkhusha

## **SUPPLEMENTARY INFORMATION**

|            | 10                                                            | 20 | 30 | 40 | 50 | 60 |
|------------|---------------------------------------------------------------|----|----|----|----|----|
|            |                                                               |    |    |    |    |    |
| zFP538     | -----MAHSKHGLKEEMTMKYHMEGCVNGHKFVITGEGIGYPFKGKQTINLCV--IEGGP  |    |    |    |    |    |
| mPapaya1   | MVSKGEGQSKHGLKEEMTVKYHMEGCVNGHKFVITGEGIGNPFGKQQTANLCV--IEGGP  |    |    |    |    |    |
| phiYFP     | -----MSSGALLFHGKIPYVVEMEGNVDGHTFSIRGKGYGDASVGKVDAQFICT---TGD  |    |    |    |    |    |
| phiYFPv    | -----GSSGALLFHGKIPYVVEMEGNVDGHTFSIRGKGYGDASVGKVDAQFICT---TGD  |    |    |    |    |    |
| TagYFP     | ----MVSKGEELFAGIVPVLIELDGDVHGHKFSVRGEGEGDADYGKLEIKFICT---TGK  |    |    |    |    |    |
| KO         | -----MSVIKPEMKMKRYFMDGSVNGHEFTVEGEGTGKPYEGHQEMTLRVTMAKGGP     |    |    |    |    |    |
| mKO        | -----MVSVIKPEMKMRYYMDGSVNGHEFTIEGEGTGRPYEGHQEMTLRVTMAKGGP     |    |    |    |    |    |
| mKO2       | -----MVSVIKPEMKMRYYMDGSVNGHEFTIEGEGTGRPYEGHQEMTLRVTMAEGGP     |    |    |    |    |    |
| mKOk       | -----MSVIKPEMKMRYYMDGSVNGHEFTIEGEGTGRPYEGHQEMTLRVTMAEGGP      |    |    |    |    |    |
| mBanana    | MVSKGEENNMAVIKEFMRFKVRMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVT--KGGP  |    |    |    |    |    |
| E2_Orange  | --MDSTEN---VIKPFMRFKVHMEGSVNGHEFEIEGEGEGKPYEGTQTAKLQVT--KGGP  |    |    |    |    |    |
| mHoneydew  | MVSKGEE---VIKEFMRFKVRMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVT--KGGP   |    |    |    |    |    |
| mOrange    | MVSKGEENNMAI KEFMRFKVRMEGSVNGHEFEIEGEGEGRPYEGFQTAKLKVT--KGGP  |    |    |    |    |    |
| mOrange2   | MVSKGEENNMAI KEFMRFKVRMEGSVNGHEFEIEGEGEGRPYEGFQTAKLKVT--KGGP  |    |    |    |    |    |
| LSSmOrange | MVSKGEENNMAI KEFMRFKVRMEGSVNGHEFEIEGEGEGRPYEGFQTVKLVKVT--KGGP |    |    |    |    |    |
| PSmOrange  | MVSKGEENNMAI KEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGFQTAKLKVT--KGGP  |    |    |    |    |    |
| PSmOrange2 | MVSKGEENNMAI KEFMRFKVHMEGTVNGHEFEIEGEGEGHPYEGFQTAKLKVT--KGGP  |    |    |    |    |    |

|            | 70                                                              | 80 | 90 | 100 | 110 | 120 |
|------------|-----------------------------------------------------------------|----|----|-----|-----|-----|
|            |                                                                 |    |    |     |     |     |
| zFP538     | LPFSEDILSAGFKYGDRI FTEYPQ--DIVDYFKNSCPAGYTWGRSFLFEDGAVCICNVDI   |    |    |     |     |     |
| mPapaya1   | LPFSEDILSPGFKYGDRI FTEYPQ--DIVDYFKNSCPAGYTWE RSYLFEDGAVCRCNVDI  |    |    |     |     |     |
| phiYFP     | VPVPWSTLVTTLT YGAQCF AKYGP--ELKDFYKSCMPEGYVQERTITFEGDGV--FKTRA  |    |    |     |     |     |
| phiYFPv    | VPVPWSTLVTTLT YGAQCF AKYGP--ELKDFYKSCMPDGYVQERTITFEGDGN--FKTRA  |    |    |     |     |     |
| TagYFP     | LPVPWPSTLVTTLT YGVQCF ARYPKHKMKNDFK SAMPEGYIQERTILFQDDGK--YKTRG |    |    |     |     |     |
| KO         | MPFSFDLVSHTF CYGHRPFTKYPE--EIPDYFKQAFPEGLSWERSLQFEDGGF--AAVSA   |    |    |     |     |     |
| mKO        | MPFAFDLVSHVFC YGHRPFTKYPE--EIPDYFKQAFPEGLSWERSLQFEDGGS--ASVSA   |    |    |     |     |     |
| mKO2       | MPFAFDLVSHVFC YGHRVFTKYPE--EIPDYFKQAFPEGLSWERSLQFEDGGS--ASVSA   |    |    |     |     |     |
| mKOk       | MPFAFDLVSHVFC YGHRVFTKYPE--EIPDYFKQAFPEGLSWERSLQFEDGGS--ASVSA   |    |    |     |     |     |
| mBanana    | LPFAWDILSPQFC YGSKAYVKHPT--GIPDYFKLSFPEGFKWERVMN FEGDGGV--VTVAQ |    |    |     |     |     |
| E2_Orange  | LPFAWDILSPQFT YGSKAYVKHPT--DIPDYKKLSFPEGFKWERVMN FEGDGGV--VTVTQ |    |    |     |     |     |
| mHoneydew  | LPFAWDILSPQFMW YGSKAYVKHPT--DIPDYKLSFPEGFKWERVMN FEGDGGV--VTVTQ |    |    |     |     |     |
| mOrange    | LPFAWDILSPQFT YGSKAYVKHPT--DIPDYKLSFPEGFKWERVMN FEGDGGV--VTVTQ  |    |    |     |     |     |
| mOrange2   | LPFAWDILSPHFT YGSKAYVKHPT--DIPDYKLSFPEGFKWERVMN YEDGGV--VTVTQ   |    |    |     |     |     |
| LSSmOrange | LPFAWDILSPQFT YGSKAYVKHPT--DIPDYKLSFPEGFKWERVMN FEGDGGV--VTVTQ  |    |    |     |     |     |
| PSmOrange  | LPFAWDILSPLFT YGSKAYVKHPT--DIPDYKLSFPEGFKWERVMN YEDGGV--VTVTQ   |    |    |     |     |     |
| PSmOrange2 | LPFAWDILSPLIT YGSKAYVKHPT--DIPDYKLSFPEGFKWERVMN YEDGGV--VTVTQ   |    |    |     |     |     |

|            | 130      | 140          | 150     | 160          | 170            | 180                 |
|------------|----------|--------------|---------|--------------|----------------|---------------------|
| zFP538     | TVSVKENC | IYHKSIFNGMNF | PADG    | PVMKK-MTTNWE | ASCEKIMPV      | PKQGILKGDVSMYLL     |
| mPapaya1   | TVSEKENC | IYHKSIFRGVNF | PADG    | PVMKK-MTTNWE | ASTEKIVPV      | PKQGILKKGKVMCLL     |
| phiYFP     | EVTFENG  | SVYNRVKLN    | QGQFKK  | DGHV         | LGNLEFNFTPHCLY | IWGDQANHGLKSAFKIMHE |
| phiYFPv    | EVTFENG  | SVYNRVKLN    | QGQFKK  | DGHV         | LGNLEFNFTPHCLY | IWGDQANHGLKSAFKICHE |
| TagYFP     | EVKFEGD  | TLVNRIELK    | GKDFKED | GNILGHKLEYS  | FNSHNVYITPDKAN | NGLEVNFKTRHN        |
| KO         | HISLRGNC | FEHKS        | KFVGVNF | PADG         | PVMQN-QSSDWE   | PSTEKITTC--DGV      |
| mKO        | HISLRGNT | FYHKS        | KFTGVNF | PADG         | PIMQN-QSVDWE   | PSTEKITAS--DGV      |
| mKO2       | HISLRGNT | FYHKS        | KFTGVNF | PADG         | PIMQN-QSVDWE   | PSTEKITAS--DGV      |
| mKOk       | HISLRGNT | FYHKS        | KFTGVNF | PADG         | PIMQN-QSVDWE   | PSTEKITAS--DGV      |
| mBanana    | DSSLQDGE | FIYKV        | KL      | RGTN         | FPSDGP         | PVMQK-KTMGWE        |
| E2_Orange  | DSSLQDGE | FIYHV        | KFIG    | VNFPSDGP     | PVMQK-KTLGWE   | PSTERLYPR--DGV      |
| mHoneydew  | DSSLQDGE | FIYKV        | KL      | RGTN         | FPSDGP         | PVMQK-KTMGWA        |
| mOrange    | DSSLQDGE | FIYKV        | KL      | RGTN         | FPSDGP         | PVMQK-KTMGWE        |
| mOrange2   | DSSLQDGE | FIYKV        | KL      | RGTN         | FPSDGP         | PVMQK-KTMGWE        |
| LSSmOrange | DSSLQDGE | FIYKV        | KL      | RGTN         | FPSDGP         | PVMQK-KTMGME        |
| PSmOrange  | DSSLQDGE | FIYKV        | KL      | RGTN         | FPSDGP         | PVMQK-KTMGWE        |
| PSmOrange2 | DSSLQDGE | FIYKV        | KL      | RGTN         | FPSDGP         | PVMQK-KTMGWE        |

|            | 190             | 200         | 210                              | 220                   | 230            | 240                |
|------------|-----------------|-------------|----------------------------------|-----------------------|----------------|--------------------|
| zFP538     | LKD--GGRYRC     | QFDTVYK     | AKSV-PSK                         | MP                    | PEWHFIQHKLL    | REDRSDAKNQKWLTEHAI |
| mPapaya1   | LKD--GGRYHC     | QFDTVYK     | AKSV-PSK                         | MP                    | PEWHFIQHKLL    | REDRSDAKNQKWLTEHAI |
| phiYFP     | ITGSKED         | FIVADHTQMNT | PIGGGPVHV                        | PEYHHITYHVTLSK        | DVTDHRDNMSL    | VE                 |
| phiYFPv    | ITGSKGD         | FIVADHTQMNT | PIGGGPVHV                        | PEYHHMSYHV            | KLSDVTDHRDNMSL | KE                 |
| TagYFP     | IEG--GGVQLADHYQ | TNVLGDGPVLI | PINHYLSYQTDISKDRNEARDH           | MVLL                  | ESVSA--        |                    |
| KO         | LAG--GGNHKC     | QFKTTYKA    | AKK-ILKMPQ                       | SHFIGHRLVRKTEGNIT---- | ELVEDA         | VAHC               |
| mKO        | LEG--GGNHKC     | QFKTTYKA    | AKK-ILKMPG                       | SHYISHRLVRKTEGNIT---- | ELVEDA         | VAHS               |
| mKO2       | LEG--GGNHKC     | QMKTTYKA    | AAKE-ILEMPGDHYIGHRLVRKTEGNIT---- | EQVEDA                | VAHS           |                    |
| mKOk       | LEG--GGNHKC     | QFKTTYKA    | AAKE-ILEMPGDHYIGHRLVRKTEGNIT---- | EQVEDA                | VAHS           |                    |
| mBanana    | LKD--GGHYSA     | ETKTTYKA    | AKK--PVQLPGAYIAGE                | KIDITSHNEDYT-IVELY    | ERA            | EGRH               |
| E2_Orange  | LKG--GGHYLV     | EFKTIYMA    | AKK--PVQLPGAYYVDSK               | LIDITSHNEDYT-VVEQY    | ERA            | EGRH               |
| mHoneydew  | LKD--GGHYDA     | EVKTTYMA    | AKK--PVQLPGAYKIDG                | KLIDITSHNEDYT-IVEQY   | ERA            | EGGH               |
| mOrange    | LKD--GGHYT      | SEVKTTYKA   | AKK--PVQLPGAYIVGI                | KLIDITSHNEDYT-IVEQY   | ERA            | EGRH               |
| mOrange2   | LKD--GGHYT      | SEVKTTYKA   | AKK--PVQLPGAYIVDI                | KLIDITSHNEDYT-IVEQY   | ERA            | EGRH               |
| LSSmOrange | LKD--GGHYT      | SEVKTTYKA   | AKK--PVQLPGAYIVDI                | KLIDITSHNEDYT-IVEQY   | ERA            | EGRH               |
| PSmOrange  | LKD--GGHYT      | SEVKTTYKA   | AKK--SVQLPGAYIVGI                | KLIDITSHNEDYT-IVEQY   | ERA            | EGRH               |
| PSmOrange2 | LKD--GGHYT      | SEVKTTYKA   | AKK--SVLLPGAYIVGI                | KLIDITSHNEDYT-IVEQY   | ERSE           | ARH                |

|            | 250                   | 260 |
|------------|-----------------------|-----|
|            |                       |     |
| zFP538     | S-----ALA-----        |     |
| mPapaya1   | D-----ELYK-----       |     |
| phiYFP     | ----CRKTYL-----       |     |
| phiYFPv    | ----CRKTYDFDAGSGDTSLS |     |
| TagYFP     | ----CSHTHGMDLYR-----  |     |
| KO         | -----                 |     |
| mKO        | -----                 |     |
| mKO2       | -----                 |     |
| mKOk       | -----                 |     |
| mBanana    | STGGMDLYK-----        |     |
| E2_Orange  | -----HLFQ-----        |     |
| mHoneydew  | STGGMDLYK-----        |     |
| mOrange    | STGGMDLYK-----        |     |
| mOrange2   | STGGMDLYK-----        |     |
| LSSmOrange | STGGMDLYK-----        |     |
| PSmOrange  | STGGMDLYK-----        |     |
| PSmOrange2 | STGGMDLYK-----        |     |

**Figure S1. Amino acid alignment of orange fluorescent proteins.** The chromophore-forming tri-peptides are highlighted in yellow.