



Amino acid residue at the 165th position tunes EYFP chromophore maturation. A structure-based design



Nadya V. Pletneva^{a,1}, Eugene G. Maksimov^{b,1}, Elena A. Protasova^b, Anastasia V. Mamontova^c, Tatiana R. Simonyan^c, Rustam H. Ziganshin^a, Konstantin A. Lukyanov^c, Liya Muslinkina^d, Sergei Pletnev^e, Alexey M. Bogdanov^{a,*}, Vladimir Z. Pletnev^{a,*}

^a Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Moscow 117997, Russia

^b Faculty of Biology, Lomonosov Moscow State University, 119992 Moscow, Russia

^c Center of Life Sciences, Skolkovo Institute of Science and Technology, Moscow 121205, Russia

^d Structural Biology Section, Research Technologies Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA

^e Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA

ARTICLE INFO

Article history:

Received 19 February 2021

Received in revised form 6 May 2021

Accepted 9 May 2021

Available online 11 May 2021

Keywords:

Fluorescent proteins

EYFP

Chromophore maturation

X-ray structure

Femtosecond spectroscopy

Excitation energy transfer

Tryptophan fluorescence

Structure-guided mutagenesis

ABSTRACT

For the whole GFP family, a few cases, when a single mutation in the chromophore environment strongly inhibits maturation, were described. Here we study EYFP-F165G – a variant of the enhanced yellow fluorescent protein – obtained by a single F165G replacement, and demonstrated multiple fluorescent states represented by the minor emission peaks in blue and yellow ranges (~470 and ~530 nm), and the major peak at ~330 nm. The latter has been assigned to tryptophan fluorescence, quenched due to excitation energy transfer to the mature chromophore in the parental EYFP protein. EYFP-F165G crystal structure revealed two general independent routes of post-translational chemistry, resulting in two main states of the polypeptide chain with the intact chromophore forming triad (~85%) and mature chromophore (~15%). Our experiments thus highlighted important stereochemical role of the 165th position strongly affecting spectral characteristics of the protein. On the basis of the determined EYFP-F165G three-dimensional structure, new variants with ~2-fold improved brightness were engineered.

© 2021 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The GFP-like fluorescent proteins (FPs) have become an important noninvasive tool for visualization and monitoring of the bio-

chemical processes within cells or whole organisms, and the range of their application is continuously expanding [1–3]. A simple fact that a genetically encoded fluorophore can be used in almost any expression system without special requirements has laid a basis for the elaboration of a bundle of techniques used today in virtually every life science research field. The potential of GFP is provided primarily by its self-sufficiency as a product of genetic expression [4]. The GFP chromophore is formed autocatalytically from the three of its own amino acid residues by a multistage pathway called maturation [5–7].

In avGFP, the first fluorescent protein isolated from *Aequorea victoria* jellyfish, the chromophore is matured from the serine-tyrosine-glycine amino acid triad, and its structure serves a prototype for the GFP-type family of chromophores – bicyclic compounds originated from the tripeptide -X65-Tyr66-Gly67- motif (where X is any amino acid residue, and the number indicates a residue position in the polypeptide chain according to the amino acid numbering of the ancestral avGFP). It was shown that the chromophore maturation process, which requires no external factors except the molecular oxygen, includes consecutive cyclization,

Abbreviations: EYFP, enhanced yellow fluorescent protein; GFP, green fluorescent protein; Phe (F), phenylalanine; Gly (G), glycine; Trp (W), tryptophan; FP, fluorescent protein; avGFP, *Aequorea victoria* green fluorescent protein; Tyr (Y), tyrosine; sfGFP, superfolder GFP; ESET, excited-state electron transfer; EGFP, enhanced green fluorescent protein; GYG, glycine-tyrosine-glycine; Arg (R), arginine; Ala (A), alanine; EET, excitation energy transfer; Val (V), valine; His (H), histidine; Ser (S), serine; EC, extinction coefficient; FQY, fluorescence quantum yield; DTT, dithiothreitol; PBS, phosphate buffered saline; Glu (E), glutamic acid; FLIM, fluorescence lifetime imaging microscopy; FRET, Förster resonance energy transfer; REACh, resonance energy-accepting chromoprotein; Gln (Q), glutamine; Leu (L), leucine; Asn (R), asparagine; IVA-cloning, in vivo assembly cloning; PCR, polymerase chain reaction; FTIR, Fourier-transform infrared (spectroscopy).

* Corresponding authors at: Department of biophotonics (both), Laboratory of genetically encoded molecular tools (A.M.B.), Laboratory of of X-ray study (V.Z.P.).

E-mail addresses: noobissat@ya.ru (A.M. Bogdanov), vzpletnev@gmail.com (V.Z. Pletnev).

¹ These authors contributed equally to the study.

<https://doi.org/10.1016/j.csbj.2021.05.017>

2001-0370/© 2021 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

dehydration, and oxidation, with the latter reaction being rate-limiting one [6–10]. An alternative order of reactions leading to chromophore formation (cyclization-oxidation-dehydration) was also proposed [11]. Despite an extensive study of maturation chemistry by both instrumental and theoretical methods, there are still the details of this process remaining poorly understood. Although the drastic effect of mutations at 96 and 222 positions, originally occupied by arginine and glutamate, respectively, on the maturation kinetics has been repeatedly described [12,13], these residues are not strictly necessary for successful chromophore formation [14–16]. Thus, the exact sequence of events, catalysis mechanism at the different maturation stages, and the role of 96th and 222nd residues in the reaction are debatable to date [17].

The rate of the GFP chromophore maturation can be measured by different methods, giving slightly divergent results for the same fluorescent protein [8,13,18,19]. In addition, this value substantially varies among the fluorescent proteins family members [20]. Usually, the green-yellow fluorescent proteins demonstrate the maturation rates within the minutes-to-hours range [8,10,18,20]; though the mutants with the dramatically impaired maturation (with the rate constants up to several months) were also engineered [12,21].

Apparently, the chromophore maturation should not be considered as a completely isolated post-translational phenomenon. There is evidence that the protein folding and the chromophore maturation are interconnected and interdependent processes. First, the chromophore amino acid environment formed during folding is of importance in lowering both the entropic and enthalpic barriers to cyclization, the first maturation step [17]. In other words, the chromophore-forming triad is prepared for maturation by the neighboring amino acids environment within the beta-barrel protein fold similar to a substrate activation in the enzyme active site [17]. Second, the measurements on the sfGFP (superfolder GFP) – avGFP-derived mutant with improved folding – showed hysteresis in the unfolding and refolding equilibrium titration curves, thus revealing a rough energy route of folding [21]. Meanwhile for the mutants with weakly maturing chromophores, the corresponding curves were shown to be superimposable. These facts could point to the mature chromophore as a cause for the energy hindrance in GFP unfolding/refolding [21–22]. Thus, chromophore formation is closely related to the architecture of the entire protein molecule and understanding the role of the chromophore amino acid environment is a key condition for design of new improved FPs.

In this contribution, we reveal an important role of amino acid substituent at position 165 in the chromophore maturation of the enhanced yellow fluorescent protein (EYFP). In EYFP, amino acid residue at the 165th position directly contacts with the chromophore and is natively occupied by a phenylalanine [15,23]. Earlier, participation of the 165th residue in the excited-state electron transfer (ESET) in EGFP/EYFP [24] and in the EGFP brightness tuning [25] were proposed. We therefore have targeted mutagenesis to the EYFP 165th position aiming at the deeper analysis of its role in the fluorescent protein behavior. We have found that the EYFP-F165G variant (amino acid sequence is shown in Fig. 1) demonstrates unexpectedly low fluorescence intensity in a yellow spectral region (520–530 nm) typical for the parental EYFP, while emission in a near-UV range (~330 nm) was significantly enhanced. To understand the drastic changes in the EYFP-F165G spectral behavior compared to its ancestor, we carried out a study combining mutagenesis, time-resolved fluorescence spectroscopy, and X-ray crystallography. These efforts allowed to clarify the function of residues at position 165 in EYFP-family and to perform a structure-based design of the promising probes with improved brightness.

2. Materials and methods

2.1. Site-directed mutagenesis and protein expression

The EYFP mutants were generated using IVA-cloning PCR technique [26] with the following oligonucleotide set containing the appropriate substitutions (primer sequences are shown in Suppl. methods). pQE30 plasmid vector backbone (Qiagen) and *E. coli* XL1 Blue strain (Evrogen) were used for the DNA constructs assembly.

FP variants, cloned into the pQE30 vector (Qiagen) with a 6His tag at the N-terminus, were expressed in *E. coli* XL1 Blue strain (Evrogen). The proteins were isolated by ultrasonic cell lysis, purified using TALON metal-affinity resin (Clontech), eluted from the adsorbing material by treatment with 100 mM imidazole (pH 8) and then desalted by ultrafiltration with Amicon® Ultra-0.5 10 K (Merck) filters. All isolation and chromatography procedures were performed in PBS buffer pH 7.4 (Gibco). The same solution was further used for the specimen storage and spectroscopy measurements.

2.2. Spectroscopy and fluorescence brightness evaluation

For absorbance and fluorescence excitation-emission spectra measurements, a Cary 100 UV/VIS spectrophotometer and a Cary Eclipse fluorescence spectrophotometer (Varian) were used. Fluorescence brightness was evaluated as a product of molar extinction coefficient by quantum yield multiplication. Measurements on all native proteins were carried out in phosphate buffered saline (PBS, pH 7.4, Gibco). For molar extinction coefficient determination, we relied on measuring mature chromophore concentration. EYFP and its mutants were alkali-denatured in 1 M NaOH. Under these conditions, GFP-like chromophore is known to absorb at 447 nm with an extinction coefficient of $44,000 \text{ M}^{-1}\text{cm}^{-1}$ [27]. Based on the absorption of the native and alkali-denatured proteins, molar extinction coefficients for the native states were calculated. For determination of the quantum yield, the areas under fluorescence emission spectra of the mutants were compared with equally absorbing EYFP (quantum yield 0.61).

2.3. Time-resolved emission spectroscopy

Fluorescence was excited by pulses at 262 nm obtained from the 4th harmonic of a Yb femtosecond laser (TEMA-150 and AFsG-A, Avesta Project LTD., Moscow, Russia), driven at an 80 MHz repetition rate, delivering 150 fs pulses to the sample. Laser power was adjusted by neutral density filters and was typically 10 mW. In other series of experiments, excitation was performed at 405 nm by a 26 picosecond laser (InTop, Russia), driven at a 25 MHz repetition rate. Fluorescence of the samples in the 300–400 nm range was collected perpendicular to the excitation beam path via a collimation lens coupled to the optical fiber. Fluorescence was detected by a cooled ultra-fast single-photon counting detector (HPM-100-07C, Becker&Hickl, Germany) with low dark count rate (~10 counts per second), coupled to a monochromator (ML-44, Solar, Belarus). In order to obtain time-resolved fluorescence spectra, decay kinetics were measured at different wavelengths with identical integration times (10 s). The temperature of the sample was stabilized at 25 °C by a Peltier-controlled cuvette holder Qpod 2e (Quantum Northwest, USA) with a magnetic stirrer. Changes of fluorescence intensity, lifetime and spectrum were processed using the SPCImage 8.1 (<https://www.becker-hickl.com/products/category/software/>, Becker&Hickl, Germany) software package. To obtain the time-integrated fluorescence intensity and steady-state emission spectra, the num-

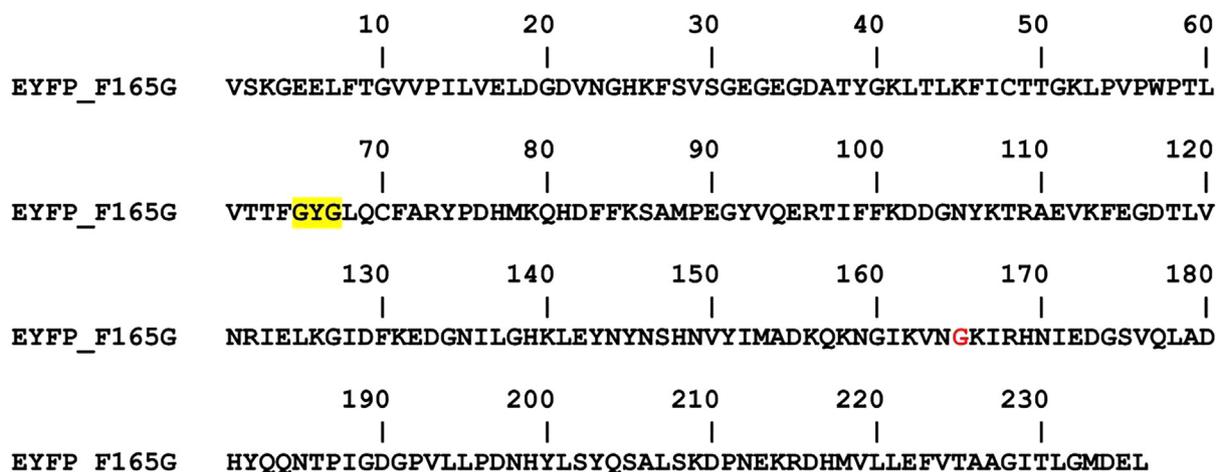


Fig. 1. Amino acid sequence of EYFP-F165G. The chromophore triad GYG is highlighted in yellow and the residue G165 is shown in red.

ber of photons in each time channel of individual fluorescence decay histograms was summed up after background noise correction (subtraction of dark offset measured in advance). All calculations were performed using Origin Pro 2015 (<https://www.originlab.com/origin>, OriginLab Corporation, USA). All experiments were repeated at least three times.

2.4. Fourier-transform infrared (FTIR) spectroscopy

IR-spectra were obtained at 4 cm^{-1} resolution with 25 scans using a MIRacle single reflection horizontal ATR with KRS-5 diamond (PIKE Technologies, USA) coupled to a Spectrum Two FTIR spectrometer (Perkin Elmer, Germany) equipped with lithium tantalate MIR detector.

2.5. Crystallization, X-ray data collection and structure determination.

Crystals of EYFP-F165G suitable for X-ray study were grown at $20\text{ }^{\circ}\text{C}$ by a hanging drop vapor diffusion method. For crystallization, the protein was transferred to a solution of 5.7 mg/ml in 20 mM Tris pH 7.5, containing 200 mM NaCl. The protein solution ($2\text{ }\mu\text{l}$) was mixed with $2\text{ }\mu\text{l}$ of a 1.44 M $(\text{NH}_4)_2\text{SO}_4$, 60 mM Bicine pH 9.0 reservoir solution. The crystals appeared in two weeks. X-ray diffraction data were collected from single crystals flash-cooled in a 100 K nitrogen stream. Prior to cooling, the crystals of EYFP-F165G were transferred to a cryo-protecting solution containing 20% glycerol and 80% reservoir solution. Data were collected at 2.2 \AA resolution with a MAR300 CCD detector at the SER-CAT beamline 22-ID (Advanced Photon Source, Argonne National Laboratory, Argonne, IL) and were processed with HKL2000 [28].

Crystal structures of EYFP-F165G were determined by the molecular replacement method with *MOLREP* [29], using as a search model the coordinates of the variant of EYFP [30] (PDB ID: 3V3D). Crystallographic refinement was performed with *REFMAC5* [31], alternating with manual revision of the model with *COOT* [32]. Location of water molecules and structure validation were performed with *COOT*. The figures were prepared with *PYMO*L [33]. Crystallographic data and structure refinement statistics are given in Table S1. The coordinates and structure factors of EYFP-F165G were deposited in the Protein Data Bank under accession code 6ZQO.

2.6. Chromophore maturation measurement

The maturation rate measurement protocol is based on a fluorescence appearance recording upon fluorescent protein renatura-

tion [20]. The purified proteins (0.2 mg/ml) were denatured by heating at $70\text{ }^{\circ}\text{C}$ for 3 min in 7 M urea, 1 mM DTT, 5 mM sodium dithionite. This procedure provides both a polypeptide chain unfolding and a reversion of the rate-limiting oxidation step of chromophore formation. To start renaturation, the denatured proteins were dissolved 1:100 in 1 ml PBS pH 7.4 (Gibco), and then subjected to a single-wavelength fluorimetry monitoring (kinetics were recorded at 515 nm). Fluorescence curves reaching plateau were normalized to a maximum value, and the time to reach a half-maximum was taken as a chromophore maturation time.

3. Results

3.1. Spectral features

Surprisingly, solution of EYFP-F165G mutant showed no intense yellow-green color and fluorescence. Instead of the main $\sim 510\text{ nm}$ absorption band typical for the EYFP and close homologs, its absorbance spectrum was characterized by a 279 nm peak, and a minor peak at 514 nm (Fig. 2A). Fluorescence spectra of EYFP-F165G mutant revealed excitation maxima (two of which correspond to the absorption peaks, and one additional at 401 nm , not visible as a separate peak in the absorbance spectrum of the mutant) and 3 emission bands peaking at 328 nm (major band), ~ 465 and $\sim 530\text{ nm}$ (minor bands) (Fig. 2B). While $\text{ex}514/\text{em}530\text{ nm}$ fluorescence should be, most likely, attributed to the mature EYFP chromophore (GFP-like chromophore π -stacked with the tyrosine-203 residue [24]), origin of the remaining UV and blue bands had to be established. UV-violet spectral form analysis was of a particular interest since its specific excitation/emission wavelengths could be potentially used in fluorescence imaging in a separate spectral channel.

Interestingly, in a freshly isolated protein, the long-wave absorption/fluorescence is weakly expressed, while in the protein that has been stored for 2 weeks, the contribution of these bands increases substantially (Fig. S1). We connected such phenomenon with the gradual slow chromophore maturation, a similar behavior was described for the GFP R96A mutants [12,22].

3.2. UV fluorescence origin

Typically, fluorescent proteins have intense absorption in the $260\text{--}320\text{ nm}$ region, and do not fluoresce in the region shorter than 425 nm [34]. As we noted above, EYFP-F165G is characterized by a strong emission in the UV region ($\sim 330\text{ nm}$), which is not characteristic for any previously characterized modifications of a fluores-

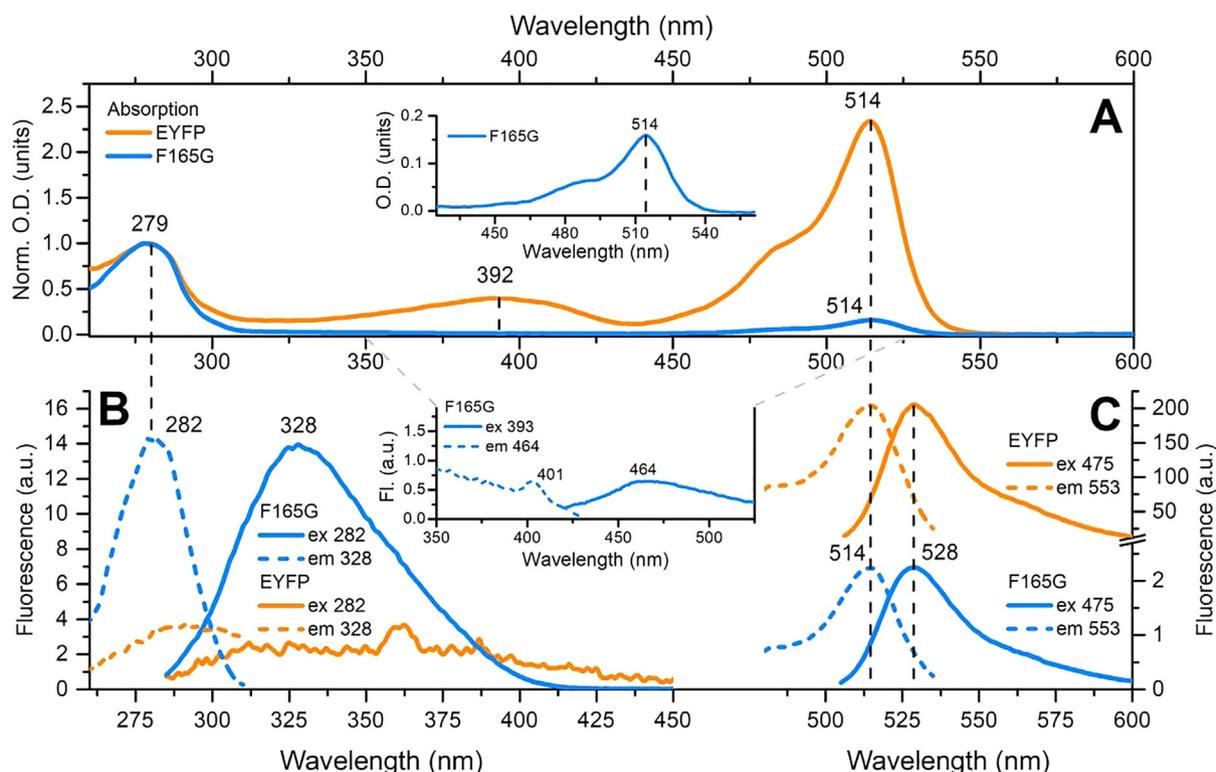


Fig. 2. (A) Absorption spectra of EYFP (orange line) versus EYFP-F165G (blue line). In EYFP, mature chromophore absorbs in the two spectral regions, ~ 400 nm (minor neutral form) and ~ 510 nm (major anionic form). For EYFP-F165G, enlarged absorbance peak of the mature chromophore is shown in a separate box. (B, C) Fluorescence excitation (dashed lines) and emission (solid lines) spectra of EYFP-F165G (in blue) and EYFP (in orange). The data recorded within the different spectral ranges are shown separately (UV region – panel B, green-yellow region – panel C, violet-blue region – the box between B and C). Excitation and emission maxima are marked above the peaks. The wavelengths used for spectra recordings are indicated in the legend. Fluorescence spectra were recorded using equal detector gain adjustments, monochromator slits, and corrected to the detector spectral sensitivity values and power of the excitation light source. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

cent protein chromophore. We thus supposed that the observed UV emission is provided by W57, the only tryptophan residue in this protein. Although EYFP has 9 tyrosine residues, most of them are exposed into the solvent and thus must be significantly quenched, while the indole group of tryptophan is located in the hydrophobic core of the protein approximately 14.5 \AA away from the chromophore, which is close enough for the efficient excitation energy transfer which might be the reason of low fluorescence intensity of EYFP with mature chromophore in the UV region [15,23,35]. In order to test our hypothesis and to assess coupling between the only tryptophan (W57) residue and the chromophore at the different stages of its maturation, we used a femtosecond UV-light source and measured time-resolved emission spectra of EYFP, EYFP-F165G and EYFP-R96A (Fig. 3A–C). The latter protein was designed to provide us with a control sample, where the maturation is inhibited *a priori* (see [12,22]). As we mentioned in Introduction, R96A substitution was shown to dramatically affect the GFP chromophore maturation rate [12,22,36–37]. Notably, slow maturation of the diverse GFP R96A mutants was accompanied by a gradual decrease in the Trp57 fluorescence [22].

Examination of the emission in the UV region revealed that the shape of the spectrum of the fluorophore (which we assign to W57) is largely similar in EYFP and its mutants and dominated by the main component with the maximum at approximately 330 nm. In EYFP, we found an additional minor component with emission at ~ 390 nm (average fluorescence lifetime about 2.5 ns) which was completely absent in mutants (Fig. 3A). Lifetime of the main component with 330 nm emission maximum was equal to 270 ps in EYFP, while in the mutants it was about 1.7 ns, which explains observed differences in the intensity of steady-state fluorescence. We assume that the reason of 6.4-fold

reduction of W57 fluorescence lifetime in EYFP is in the excitation energy transfer (EET) to the mature chromophore. Indeed, inspection of the fluorescence decay of EYFP (and mutants) chromophore at 530 nm upon the direct excitation in the visible region (Fig. 3J) reveals that accumulation of its excited state occurs with about 280 ps delay when excited in the UV, which proves that W57 transfers excitation energy to chromophore. Considering the lifetimes of W57 fluorescence we can estimate the efficiency of EET about 84.4%.

Since spectra and fluorescence decays of yellow states are similar in EYFP and EYFP-F165G, we assume that maturation of the chromophore in EYFP-F165G variant is still possible; however, its fraction is low. Assuming that (i) the efficiency of EET in such EYFP-F165G with mature chromophore is equivalent to that in EYFP, (ii) the number of tryptophans and chromophores is equal in ‘mature’ proteins, and considering the lifetimes of chromophores and observed ratios of visible range to UV fluorescence, we can estimate the fraction of EYFP-F165G with mature chromophore of about 16.3%. Yield of the mature chromophore in EYFP-R96A was below detection limit.

Examination of EYFP, EYFP-F165G and EYFP-R96A fluorescence under 405 nm excitation revealed significant contribution of rapidly decaying state of the chromophore with a lifetime shorter than 50 ps to emission at around 440–480 nm (Fig. 3D–H and K). Although this component is present in EYFP, its relative yield is higher in mutants with low fraction of mature chromophore (Fig. 3H). We assume that this component might be related either to an enolate maturation intermediate [6] or the protonated form of the mature chromophore [18,38,39]. To assess possible contribution of the neutral (protonated) state of EYFP’s mature chromophore ($\lambda_{\text{abs}} \sim 400$ nm) to EET from W57, we conducted similar

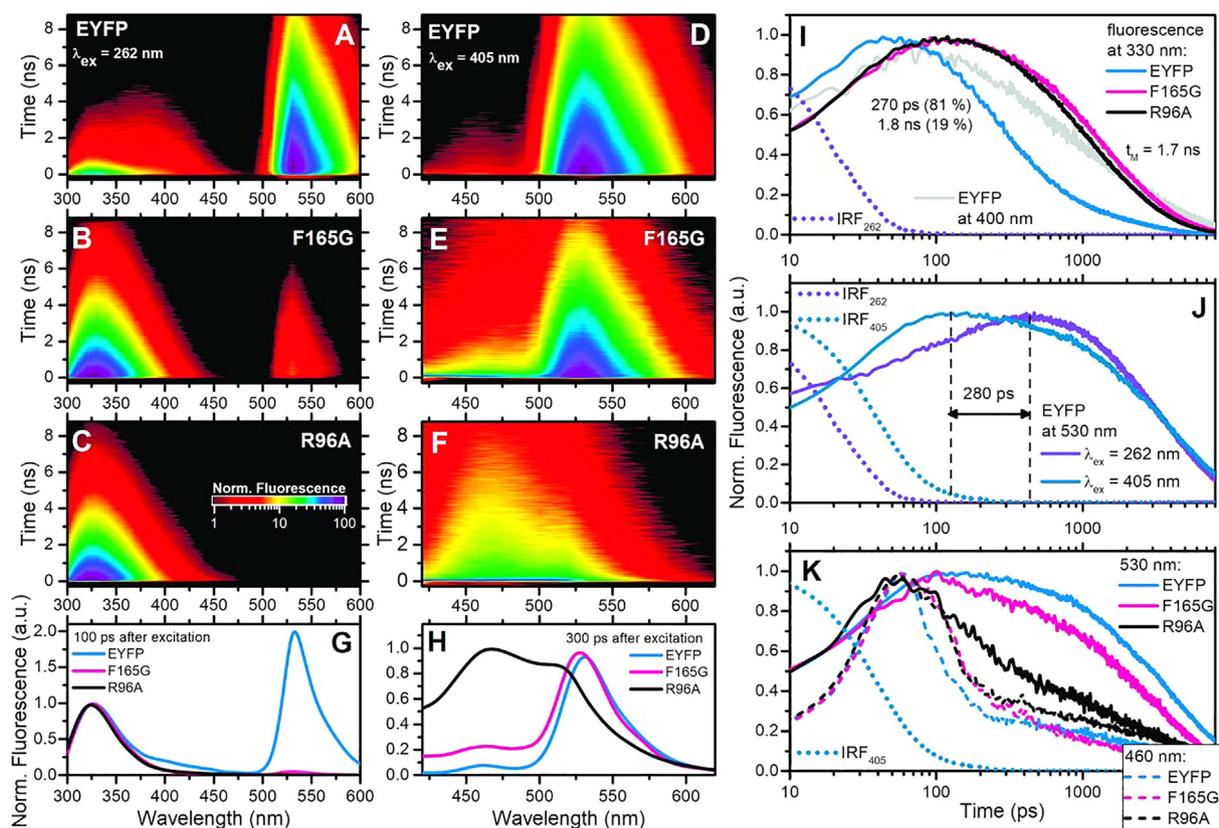


Fig. 3. Time-resolved emission spectra of EYFP and its mutants. Distribution of fluorescence intensity of EYFP, F165G and R96A over wavelength and time after excitation by 262 (A–C) and 405 nm (D–F) laser pulses represented as color-coded images. For all samples fluorescence intensity was normalized to fit the same logarithmic color scale (from 1 to 100 relative units, see panel C). Characteristic spectra at 100 ps after excitation by 262 nm laser (G) and 300 ps after excitation by 405 nm (H) obtained as cross-sections of time-resolved spectra. Spectra in (G) were normalized at 330 nm. (I) – normalized time-courses of EYFP, F165G and R96A fluorescence in the UV region. Dotted line shows instrumental response of the detector to 150 fs laser pulse (full IRF is shown in Fig. S2). Numbers indicate characteristic lifetimes of the components. (J) – comparison of EYFP's chromophore kinetics obtained under direct (405 nm, blue line) and indirect (262 nm, violet line) excitation. Dotted lines show instrumental response of the detector to corresponding light sources. Arrow indicates the difference between the moments when concentration of the excited states reaches maximal level. (K) – normalized time-courses of EYFP, F165G and R96A fluorescence at 530 nm (solid lines) and at 460 nm (dashed lines). Dotted line shows instrumental response of the detector to 26 ps pulse of 405 nm laser. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

set of measurements at the higher pH (pH 9), where the fraction of protonated chromophore is expected to be significantly lower. While we found some decrease of tryptophan fluorescence lifetime in EYFP-F165G (~1.6 ns at pH 9 versus ~1.8 ns at pH 7.4, see Fig. S3), which well corresponds to a higher absorption of EYFP-F165G anionic chromophore at pH 9 (Fig. S1B), its value in EYFP remained essentially unchanged (Fig. S3L, M).

Since poor chromophore maturation could predict unsatisfactory protein folding [22], we also examined the native and thermally denatured EYFP and EYFP-F165G by Fourier-transform infrared (FTIR) spectroscopy (Fig. S4). In contrast to native EYFP, both denatured proteins as well as native F165G showed an intense band at ~1675 cm^{-1} characteristic of α -helical and poorly structured polypeptides (Fig. S4). These data indicate that EYFP-F165G demonstrates the signatures of poor folding.

To better understand the structural impact of the F165G substitution on the EYFP chromophore environment, and most significantly, to reveal the structural determinants of low chromophore maturation efficiency, we have crystallized the EYFP-F165G mutant and determined its three dimensional structure by X-ray method (PDB_ID 6ZQO, Supplementary: Table 1S).

3.3. The crystal structure of EYFP-F165G

In crystalline state, EYFP-F165G forms a dimer with a contact area of ~2,866 Å² [2] at the inter-monomer interfaces stabilized

by ten direct H-bonds, a number of the water-mediated H-bonds, and hydrophobic interactions (Table 2S). Dimerization in the crystalline state suggests a tendency of the protein to oligomerization that is in good agreement with parental EYFP being a weak dimer (PDB ID: 3V3D). The principal structural fold of the EYFP-F165G monomer, shared with other members of the GFP family, is an 11-stranded β -barrel with loop caps from both sides (Fig. 4A). Comparison of EYFP-F165G with EYFP showed a nearly identical overall fold and a significant difference concentrated in the chromophore area (Fig. 4B). A distinctive feature of the EYFP-F165G structure is the presence of an intact chromophore forming G65-Y66-G67 tripeptide. Modeling the intact chromophore tripeptide in the 2Fo-Fc electron density map leaves a substantial shapeless Fo-Fc electron density adjacent to the position of the intact tripeptide (Fig. 4C). Based on the absorbance and fluorescence spectra of EYFP-F165G, the latter electron density was ascribed to mature EYFP chromophore that resulted from classic posttranslational modification of GYG tripeptide, comprising cyclization-dehydration-oxidation steps [40]. Unfortunately, the quality of the electron density does not permit reliable estimation of the occupancies of intact GYG and the chromophore based on X-ray data. The fluorescence spectra recorded at pH 7.4 indicate the presence of 16.3% of the mature chromophore. The crystals of EYFP-F165G were obtained for a protein that, after 3 weeks at pH 7.4 in PBS, was crystallized at pH 9 for two weeks. The absorbance spectra similar to the crystallization conditions (protein being for 3 weeks at pH

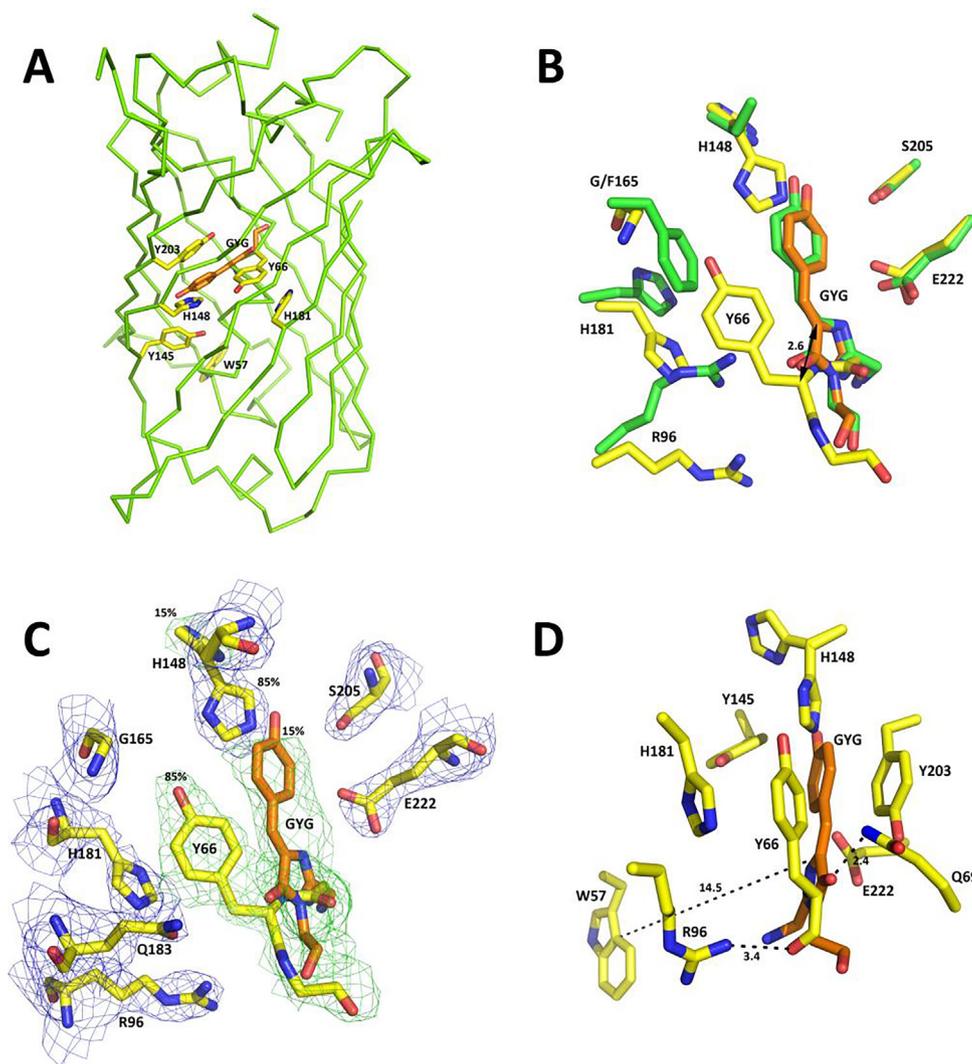


Fig. 4. Crystal structure EYFP-F165G. (A) The principal structural fold of the β -barrel and the key amino acid residues. (B) Comparison of the immediate chromophore environment by superposition of EYFP-F165G and EYFP (shown in green). (C) $2Fo-Fc$ electron density (blue, 1σ level) and $Fo-Fc$ omit map (green, 2σ level) for the nearest chromophore environment of the mature chromophore (brown) and intact GYG tripeptide (yellow). (D) The mutual arrangement of W57 and the chromophore. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

7.5 and 3 days at pH 9.0) do not show any appreciable increase of the mature chromophore fraction of EYFP-F165G. Thus, to reflect the two states of the chromophore-forming tripeptide, we set the occupancies of intact GYG and matured chromophore as 0.85 and 0.15, based on fluorescence measurements at pH 7.4. The structure of the mature chromophore was modeled using the coordinates of EYFP (PDB ID: 3V3D), whereas intact GYG was directly fit in the $2Fo-Fc$ electron density. The $C\alpha$ atom of intact Tyr66 and $C\alpha^2$ atom of the chromophore are separated by 2.6 Å. Superposition of EYFP and EYFP-F165G structures further exposed the change in the conformations of R96, H181, and H148 in the immediate chromophore environment (Fig. 4B). The change in the conformations of these residues originates from a push by the side chain of intact Y66 that in EYFP-F165G takes the place of the F165 side chain. For His148, we observe two alternative conformations, the major one, corresponding to the presence of intact G65-Y66-G67 tripeptide and, the minor, corresponding to mature GYG chromophore and also seen in EYFP. For R96 and H181, on the other hand, electron density corresponds strictly to conformations different from those observed in EYFP. In addition to positions 96 and 222, filled with catalytic arginine and glutamate residues, the structure revealed a key site at position 165 near the Y66. We found this site signifi-

cant due to the steric effect of branched residues at this position on chromophore maturation (see Table 1).

3.4. Structure-guided mutagenesis

Structural analysis both clarified an impact of the 165th site and allowed recognition of an important role of the 148th amino acid position in the manifestation of the EYFP-F165G properties. In the crystal structure, glycine at the first site is located near the intact Y66 residue, and histidine at the second one is near the Y66 of the mature chromophore. Based on their stereochemical environment, we have proposed two groups of point mutations that affect the thermodynamic equilibrium between the mature and immature states of chromophore. (i) The first group of amino acid substituents at site 165 included residues (A, V, H and W) with side chain that sterically suppress the immature state in favor of a mature one, suggesting more efficient maturation and, as a consequence, an improvement of the protein's photophysical characteristics; (ii) The second group, represented by the serine residue at positions 165 and 148, presumably stabilizes the immature and mature states by H-bonding to the corresponding intact and chromophore Y66 residues.

Table 1

Optical and physicochemical properties of EYFP, EYFP-F165G, and the mutants designed on a basis of X-ray structure of the latter.

Fluorescent protein	$\lambda_{ex}/\lambda_{em}$, nm	EC, $M^{-1}cm^{-1}$	FQY	B, %*	A.M.T.**, days/%	T, s***
EYFP	514/526	84,000	0.61	100	1/100	1233
F165G	514/529	n/d	n/d	n/d	n/d	[2522]
F165A	514/527	124,000	0.66	160	7/10	574
F165V	515/527	123,000	0.69	164	4/6	763
F165S	n/d	n/d	n/d	n/d	n/d	n/d
F165H	513/526	140,000	0.8	220	4/14	619
F165W	512/526	55,500	0.32	35	7/24	482
H148S/F165H	506/524	135,000	0.73	192	1/100	908

B – relative brightness, A.M.T. – apparent maturation time in bacteria, T – maturation time for purified protein. *Relative brightness was calculated as a product of the molar extinction coefficient and the fluorescence quantum yield, and reported relative to the brightness of EYFP.

**Apparent maturation time was determined in the course of a quantitative fluorescence analysis of bacterial colonies. We presented these data as pairs of numbers separated by slash; first one indicates days required to reach a maximum fluorescence intensity ($\lambda_{em} \sim 530$ nm), second one shows the relative intensity (%) measured 24 h post transformation and compared to the maximum intensity reached within the observation period.

***Determined from the fluorescence recovery kinetics (Fig. S6) after protein denaturation under reducing conditions followed by a renaturation in PBS pH 7.4 (see Materials and Methods for details).

To test the hypotheses mentioned, we performed site-directed mutagenesis of the EYFP, isolated the mutants, and analyzed their spectral characteristics *in vitro* (see Table 1, Fig. S5). Indeed, some of the substitutions introduced promoted maturation by shifting the equilibrium towards a fully mature chromophore emitting in a yellow spectral range. The most of mutations weakly influenced fluorescence excitation/emission maxima wavelengths (compared to that of EYFP), all found in $\sim 515/525$ nm region. Mutants carrying 165A, 165V and 165H substitutions showed fluorescence brightness comparable or even higher to that of the EYFP (Table 1). The most remarkable of them – 165H – having an extinction coefficient of $140,000 M^{-1}cm^{-1}$ and the fluorescence quantum yield of 0.8 can be included into the chart of the top-10 brightest fluorescent proteins ever described [34]. The effect of 165W substitution on brightness recovery was found to be distinctly weaker, while 165S mutant appeared to be almost non-fluorescent (similarly to its cousin, the EYFP-F165G). We suppose that introduction of bulky W165 led to a local steric hindrance accompanied by a significant decrease in yellow fluorescence. Most probably, the loss of maturation ability in the 165S variant is due to the stabilization of the immature chromophore state by H-bonding to intact Y66. The H148S mutant was also found to be non-fluorescent. Moreover, we detected its severe aggregation and precipitation upon isolation and metal-affinity purification. One can suppose that S148 introduction led to the alterations in folding of this protein variant. We also engineered a double mutant – 148S/165H – hoping to strengthen histidine-165 impact on the chromophore maturation by introducing an additional H-bond to stabilize the mature chromophore. The 165H/148S variant demonstrated a ~ 10 nm hypsochromic shift in absorbance and fluorescence excitation maxima and fluorescence brightness similar to that of the 165H mutant (Table 1).

Noticeably, most of the 165th position mutants demonstrated a significant delay in their fluorescence manifestation upon expression in bacteria compared to that of EYFP (see Table 1, “apparent maturation time” column). Even the brightest variants (165A, V and H) used to reach the maximum fluorescence intensity in bacterial colonies by only the 4th–7th day of expression. We supposed that the reason for such behavior is the decreased chromophore maturation rates caused by the mutations. In 148S/165H double mutant, fluorescence in bacterial cells appeared faster than in 165H but still with some delay compared to the EYFP. Surprisingly, the chromophore maturation measurements carried out *in vitro*, upon renaturation of the isolated proteins, revealed that the variants with high intrinsic brightness showing slow fluorescence appearance in bacteria recovered faster than the parental EYFP (see Table 1, “maturation time” column). Therefore, we would con-

nect the delays in a fluorescence appearance of the mutants in bacteria rather with peculiarities of the maturation process in a particular *E.coli* strain [41], or with early stage folding complications. The latter assumption found support in fluorescence microscopy imaging of EYFP-F165H expressed in mammalian cells, where it showed a mediocre brightness (~ 2.5 times lower than EYFP) when visualized 24 h post transfection (Fig. S7).

4. Discussion

In the fluorescent proteins, amino acid environment plays a crucial role in the chromophore properties determination. Amino acid residues come to proximity with the chromophore triad upon folding, and form the environment providing its accompaniment at all life-stages, from the very maturation start to photodestruction in light-dependent reactions. This circumstance underlies the contrast between fluorescent proteins and small organic fluorophores. Although for the latter, the environment (featured by a solvent chemical nature) can also play an essential role, it is not a matter for designing, being often dictated by experimental conditions. In the fluorescent proteins’ engineering, conversely, attempts to tune and improve fluorophore characteristics are mainly connected with the chromophore environment modification [8,42,43]. Thus, the earliest achievements in the GFP-family development, represented by the fluorescence brightness improvement due to changes in the chromophore ground-state protonation [8], obtaining new spectral forms [44] and photoactivatable variants [45], were associated with the modification of the chromophore environment rather than affecting its own structure. EYFP – one of the conventionally used fluorescent proteins – represents a nice example of the GFP-like chromophore spectral tuning mediated by the interaction with the neighbor Y203 residue [45–46]. Several important studies were devoted to the role of two other residues, R96 and E222, in the maturation, protonation state, and spectral tuning of the GFP chromophore [12,13,47]. Here we show that a single insertion of the different substituents at the EYFP 165th position can give a variety of outcomes – from weakly UV-fluorescent protein with severely impaired maturation (165G, 165S) to extremely bright variants (165H or 165H/148S) that outperform the original EYFP, show similar characteristics to LanYFPs and almost twofold higher brightness than such yellow-emitting probes as phiYFP, mCitrine, YPet [34]. Although phenylalanine-165 was previously targeted for mutagenesis aimed at the parental protein fluorescence alteration (for example, when creating a FLIM-FRET acceptor REACH [48]), and even studied systematically by the saturation mutagenesis in OPT-GFP, a close EYFP homologue

[49], 165th position mutations have never been described as dramatically affecting the protein characteristics. Thus, in the latter case, F165 position was described as tolerating all mutations. Probably, Y203 makes 165th position significantly more sensitive to the substituent residue nature.

A single replacement F165G in EYFP (chromophore forming triad Gly65-Tyr66-Gly67) had a significant influence on the chromophore maturation and spectral properties of the protein. In the early days on GFP structural research, a vast number of works explored similar changes arising from single point mutations of highly conserved residues in the nearest chromophore environment and the chromophore itself. Their effect ranged from complete cessation of the chromophore maturation to highly precise change of the fluorescence corresponding to different steps of the chromophore maturation [36,50–52]. A distinctive feature of the EYFP-F165G structure is the presence of two species at the space typically occupied by the chromophore. Only ~ 15% of GYG tripeptide matured into the chromophore, the remaining 85% of the tripeptide remained intact. According to Barondeau et al. [40], the driving force of the chromophore maturation is a tight-turn conformation of a highly distorted central helix enforced by the rigid scaffold of the protein β -barrel. Cyclization of the tripeptide shortens alpha helix and releasing its steric strain. The structure of EYFP-F165G revealed that the absence of the bulky side chain of F165 provides a free space for accommodation of the phenolic ring of the intact Y66. It relaxes the kinked conformation of the central α -helix required for maturation of the chromophore. In the absence of the side chain of G165, the amide nitrogen of G67 is no longer pushed to closely approach the carbonyl carbon of the residue 65 enabling a nucleophilic attack. As a result, the chromophore maturation relies on a statistical probability of a favorable alignment of the residues 67 and 65 to undergo cyclization and subsequent maturation of a small fraction of the EYFP-F165G chromophore. In addition, it initiates the switching of the H148 side chain to a new dominant orientation that in mature protein is occupied by the hydroxyl group of the chromophore. The second less populated orientation (~15%) of the H148 in EYFP-F165G, observed as the dominant conformation in the EYFP precursor, favors the state of the polypeptide chain carrying the mature chromophore.

The deficiency of the chromophore maturation in EYFP-F165G promotes an intense fluorescence of its only tryptophan residue (W57), which manifestation in the parental protein is “masked” by the excitation energy transfer (EET) to the mature chromophore. Such nonradiative intra-protein energy transfer system could have been an important evolutionary advantage for the marine animals naturally expressing fluorescent proteins. Indeed, UV-violet sunlight component appropriate for tryptophan excitation can penetrate ocean water as deep as 40 m [53–54], thus providing a natural way to recruit protein fluorescence for intragroup communication, prey attraction, predator repelling, and other biological functions. However, in terms of the biological photosensitivity tryptophan fluorescence itself is suboptimal in both spectral range and intensity. One might suppose that the animals which elaborated protein variants with the UV excitation and FRET-based emission in the visible range got a sufficient advantage in a natural selection.

5. Conclusions

Here, we presented the results of spectroscopy, biochemical and X-ray studies of the gene-engineered variant EYFP-F165G (chromophore forming triad G65-Y66-G67) demonstrating multiple fluorescent states (minor ones in yellow and blue, and a major – in UV-violet spectral region) resulting from the competing routes of post-translational modification. Time-resolved fluorescence spectroscopy allowed connecting EYFP-F165G UV-violet

emission with an impaired chromophore maturation and attributing it to a fluorescence of own tryptophan residue (W57). In the parental EYFP, tryptophan fluorescence was shown to be quenched significantly by an excitation energy transfer (EET) to the mature chromophore reaching an efficiency of ~ 85%. In turn, the minor (~15%) yellow fluorescent form of the EYFP-F165G, being examined in both steady-state and time-resolved modes, displayed a typical behavior of the mature anionic EYFP chromophore. The nature of the blue fluorescent state remained unclear.

The crystal structure revealed two main structural states of the protein – the major one, with an intact polypeptide chain and the minor one, with a mature chromophore. Site 165 adjacent to intact Y66 was found to be one of the most critical in the post-translational process through its proposed steric effect on the activation barrier of the reaction. Modification of this site has been shown to have a significant stereochemical effect on the spectral properties of the protein. Mutations G165A/V/H resulted in complete restoration of yellow fluorescence with brightness comparable or even higher than that of the parental protein EYFP. The brightness of the most successful variant of the EYFP-F165H exceeded that of the EYFP by ~2.2 times.

Funding sources

This project was supported by the Russian Foundation for Basic Research (fluorescence spectroscopy and genetic engineering – grant 19-04-00845 and X-ray study – grant 19-04-00107), and the Intramural Research Program of the Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health. Diffraction experiments were carried out at synchrotron beamline 22-ID of the Southeast Regional Collaborative Access Team (SER-CAT), located at the Advanced Photon Source, Argonne National Laboratory. Use of the APS was supported by the U.S. Department of Energy, Office of Science, and Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U. S. Government. The crystallization experiments were supported by the Russian State Ministry of Education and Science (registration number AAAA-A19-119042590107-1).

CRedit authorship contribution statement

Nadya V. Pletneva: Investigation, Formal analysis, Software. **Eugene G. Maksimov:** Methodology, Validation, Investigation, Visualization. **Elena A. Protasova:** Formal analysis, Investigation. **Anastasia V. Mamontova:** Visualization, Investigation. **Tatiana R. Simonyan:** Investigation. **Rustam H. Ziganshin:** Resources, Methodology. **Konstantin A. Lukyanov:** Validation, Supervision. **Liya Muslinkina:** Methodology, Data curation. **Sergei Pletnev:** Resources, Investigation, Funding acquisition. **Alexey M. Bogdanov:** Conceptualization, Validation, Investigation, Project administration, Funding acquisition. **Vladimir Z. Pletnev:** Methodology, Formal analysis, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We sincerely thank Dr. Kyril Solntsev for the valuable discussion on the fluorescent proteins spectroscopy and Dr. Rita Chertkova for her kind help with the high-content protein isolation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2021.05.017>.

References

- Chudakov, D. M.; Matz, M. V.; Lukyanov, S.; Lukyanov, K. A. Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. *Physiol Rev* 2010, 90, 1103–1163. <https://doi.org/doi:10.1152/physrev.00038.2009>.
- Lam A, St-Pierre F, Gong Y, Marshall JD, Baird MA, McKeown MR, et al. Improving FRET dynamic range with bright green and red fluorescent proteins. *Nat Methods* 2013;9(10):1005–12. <https://doi.org/10.1021/jacs.9b07152>.
- Wu, B.; Piatkevich, K. D.; Lionnet, T.; Singer, R. H. Modern Fluorescent Proteins and Imaging Technologies to Study Gene Expression, Nuclear Localization, and Dynamics. *Curr Opin Cell Biol*. 2011, 23 (3), 310–317. <https://doi.org/doi:10.1016/j.ccb.2010.12.004>.
- Chalfie M, Tu Y, Euskirchen G, Ward W, Prasher D. Green fluorescent protein as a marker for gene expression. *Science* 1994;263(5148):802–5. <https://doi.org/10.1126/science.8303295>.
- Cubitt AB, Heim R, Adams SR, Boyd AE, Gross LA, Tsien RY. Understanding, improving and using green fluorescent proteins. *Trends Biochem Sci* 1995;20(11):448–55. [https://doi.org/10.1016/S0968-0004\(00\)89099-4](https://doi.org/10.1016/S0968-0004(00)89099-4).
- Barondeau DP, Tainer JA, Getzoff ED. Structural Evidence for an Enolate Intermediate in GFP Fluorophore Biosynthesis. *J Am Chem Soc* 2006;128(10):3166–8. <https://doi.org/10.1021/ja0552693>.
- Pletneva, N. V.; Pletnev, V. Z.; Lukyanov, K. A.; Gurskaya, N. G.; Goryacheva, E. A.; Martynov, V. I.; Wlodawer, A.; Dauter, Z.; Pletnev, S. Structural Evidence for a Dehydrated Intermediate in Green Fluorescent Protein Chromophore Biosynthesis. *J Biol Chem* 2010, 285 (21), 15978–15984. <https://doi.org/doi:10.1074/jbc.M109.092320>.
- Heim R, Cubitt AB, Tsien RY. Improved Green Fluorescence. *Nature* 1995;373(6516):663–4. <https://doi.org/10.1038/373663b0>.
- Grigorenko BL, Krylov AI, Nemukhin AV. Molecular modeling clarifies the mechanism of chromophore maturation in the green fluorescent protein. *J Am Chem Soc* 2017;139(30):10239–49. <https://doi.org/10.1021/jacs.7b00676>.
- Barondeau DP, Kassmann CJ, Tainer JA, Getzoff ED. The case of the missing ring: radical cleavage of a carbon–carbon bond and implications for GFP chromophore biosynthesis. *J Am Chem Soc* 2007;129(11):3118–26. <https://doi.org/10.1021/ja063983u>.
- Zhang L, Patel HN, Lappe JW, Wachter RM. Reaction progress of chromophore biogenesis in green fluorescent protein. *J Am Chem Soc* 2006;128(14):4766–72. <https://doi.org/10.1021/ja0580439>.
- Barondeau DP, Putnam CD, Kassmann CJ, Tainer JA, Getzoff ED. Mechanism and Energetics of Green Fluorescent Protein Chromophore Synthesis Revealed by Trapped Intermediate Structures. *Proc Natl Acad Sci* 2003;100(21):12111–6. <https://doi.org/10.1073/pnas.2133463100>.
- Sniegowski, J. A.; Lappe, J. W.; Patel, H. N.; Huffman, H. A.; Wachter, R. M. Base Catalysis of Chromophore Formation in Arg96 and Glu222 Variants of Green Fluorescent Protein. *J Biol Chem*. 2005, 280 (28), 26248–26255. <https://doi.org/doi:10.1074/jbc.M412327200>.
- Ehrig, T.; O’Kane, D. J.; Prendergast, F. G. Green-Fluorescent Protein Mutants with Altered Fluorescence Excitation Spectra. *FEBS Lett* 1995, 367, 163–166. [https://doi.org/doi:10.1016/0014-5793\(95\)00557-p](https://doi.org/doi:10.1016/0014-5793(95)00557-p).
- Wachter, R. M.; Elsliger, M.-A.; Kallio, K.; Hanson, G. T.; Remington, S. J. Structural Basis of Spectral Shifts in the Yellow-Emission Variants of Green Fluorescent Protein. *Structure* 1998, 6(10), 1267–1277. [https://doi.org/doi:10.1016/s0969-2126\(98\)00127-0](https://doi.org/doi:10.1016/s0969-2126(98)00127-0).
- Jung, G.; Wiehler, J.; Zumbusch, A. The Photophysics of Green Fluorescent Protein: Influence of the Key Amino Acids at Positions 65, 203, and 222. *Biophys. J.* 2005, 88, 1932–1947. <https://doi.org/doi:10.1529/biophysj.104.044412>.
- Craggs, T. D. Green Fluorescent Protein: Structure, Folding and Chromophore Maturation. *Chem Soc Rev*. 2009, 38, 2865–2875. <https://doi.org/doi:10.1039/b903641p>.
- Heim, R.; Prasher, D. C.; Tsien, R. Y. Wavelength Mutations and Posttranslational Autoxidation of Green Fluorescent Protein. *Proc Natl Acad Sci USA* 1994, 91, 12501–12504. <https://doi.org/doi:10.1073/pnas.91.26.12501>.
- Reid BG, Flynn GC. Chromophore formation in green fluorescent protein. *Biochemistry* 1997;36(22):6786–91. <https://doi.org/10.1021/bi970281w>.
- Balleza, E.; Kim, J. M.; Cluzel, P. Systematic Characterization of Maturation Time of Fluorescent Proteins in Living Cells. *Nat Methods* 2018, 15 (1), 47–59. <https://doi.org/doi:10.1038/nmeth.4509>.
- Andrews BT, Schoenfish AR, Roy M, Waldo G, Jennings PA. The rough energy landscape of superfolder GFP is linked to the chromophore. *J Mol Biol* 2007;373(2):476–90. <https://doi.org/10.1016/j.jmb.2007.07.071>.
- Huang J, Craggs TD, Christodoulou J, Jackson SE. Stable intermediate states and high energy barriers in the unfolding of GFP. *J Mol Biol* 2005;370:356–71. <https://doi.org/10.1016/j.jmb.2007.04.039>.
- Ormo M, Cubitt AB, Kallio K, Gross LA, Tsien RY, Remington SJ. Crystal structure of the aequorea victoria green fluorescent protein. *Science* 1996;273(5280):1392–5. <https://doi.org/10.1126/science.273.5280.1392>.
- Bogdanov AM, Acharya A, Titelmayer AV, Mamontova AV, Bravaya KB, Kolomeisky AB, et al. Turning on and off Photoinduced Electron transfer in fluorescent proteins by π -stacking, halide binding, and Tyr145 mutations. *J Am Chem Soc* 2016;138(14):4807–17. <https://doi.org/10.1021/jacs.6b00092>.
- Mamontova AV, Solovvey ID, Savitsky AP, Shakhov AM, Lukyanov KA, Bogdanov AM, et al. GFP with Subnanosecond Fluorescence Lifetime. *Sci Rep* 2018;8(13224):1–5. <https://doi.org/10.1038/s41598-018-31687-w>.
- García-Nafraía J. IVA cloning: a single-tube universal cloning system exploiting bacterial in vivo assembly. *Sci Rep* 2016;6(27459):1–12. <https://doi.org/10.1038/srep27459>.
- Chudakov DM, Verkhusha VV, Staroverov DB, Souslova EA, Lukyanov S, Lukyanov KA. Photoswitchable cyan fluorescent protein for protein tracking. *Nat Biotechnol* 2004;22(11):1435–9. <https://doi.org/10.1038/nbt1025>.
- Otwinowski, Z.; Minor, W. Processing of X-Ray Diffraction Data Collected in Oscillation Mode. In *Methods in Enzymology*; Elsevier, 1997; Vol. 276, pp 307–326. [https://doi.org/10.1016/S0076-6879\(97\)70666-X](https://doi.org/10.1016/S0076-6879(97)70666-X).
- Vagin A, Teplyakov A. Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* 2010;66(1):22–5. <https://doi.org/10.1107/S0907444909042589>.
- Meulenaere E, Nguyen Bich N, De Wergifosse M, Van Hecke K, Van Meervelt L, Vanderleyden J, et al. Clays improving the second-order nonlinear optical response of fluorescent proteins: the symmetry argument. *J Am Chem Soc* 2013;135(10):4061–9.
- Murshudov GN, Lebedev AA, Pannu NS, Nicholls RA, Winn MD, Long F, et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 2011;67(4):355–67. <https://doi.org/10.1107/S0907444911001314>.
- Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of coot. *Acta Crystallogr D Biol Crystallogr* 2010;66(4):486–501. <https://doi.org/10.1107/S0907444910007493>.
- DeLano W.L. The PyMOL Molecular Graphics System // <http://www.pymol.org> 2002.
- Interactive Chart of FP Properties. Available online: <https://www.fpbase.org/chart> (accessed Nov 25, 2020).
- Prasher DC, Eckenrode VK, Ward WW, Prendergast FG, Cormier MJ. Primary structure of the aequorea victoria green-fluorescent protein. *Gene* 1992;111(2):229–33. [https://doi.org/10.1016/0378-1119\(92\)90691-b](https://doi.org/10.1016/0378-1119(92)90691-b).
- Barondeau DP, Kassmann CJ, Tainer JA, Getzoff ED. Understanding GFP chromophore biosynthesis: controlling backbone cyclization and modifying post-translational chemistry. *Biochemistry* 2005;44(6):1960–70. <https://doi.org/10.1021/bi0479205>.
- Wood TI, Barondeau DP, Hitomi C, Kassmann CJ, Tainer JA, Getzoff ED. Defining the role of arginine 96 in green fluorescent protein fluorophore biosynthesis. *Biochemistry* 2005;44(49):16211–20. <https://doi.org/10.1021/bi051388j>.
- Chattoraj M, King BA, Bublitz GU, Boxer SG. Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. *Proc Natl Acad Sci USA* 1996;93:8362–7. <https://doi.org/10.1073/pnas.93.16.8362>.
- Shu X, Leiderman P, Gepshtein R, Smith NR, Kallio K, Huppert D, et al. An alternative excited-state proton transfer pathway in green fluorescent protein variant S205V. *Protein Sci* 2007;16:2703–10. <https://doi.org/10.1110/ps.073112007>.
- Barondeau DP, Kassmann CJ, Tainer JA, Getzoff ED. Understanding GFP posttranslational chemistry: structures of designed variants that achieve backbone fragmentation, hydrolysis, and decarboxylation. *J Am Chem Soc* 2006;128(14):4685–93. <https://doi.org/10.1021/ja056635l>.
- Hebisch E, Knebel J, Landsberg J, Frey E, Leisner M. High variation of fluorescence protein maturation times in closely related escherichia coli strains. *PLoS ONE* 2013;8(10):. <https://doi.org/10.1371/journal.pone.0075991>.
- Zhang G, Gurtu V, Kain SR. An enhanced green fluorescent protein allows sensitive detection of gene transfer in mammalian cells. *Biochem Biophys Res Commun* 1996;227(3):707–11. <https://doi.org/10.1006/bbrc.1996.1573>.
- Tsien RY. The green fluorescent protein. *Annu Rev Biochem* 1998;67:509–44. <https://doi.org/10.1146/annurev.biochem.67.1.509>.
- Kremers G-J, Goedhart J, van Munster EB, Gadella TWJ. Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET Fo’rster radius. *Biochemistry* 2006;45(21):6570–80. <https://doi.org/10.1021/bi0516273>.
- Henderson JN, Gepshtein R, Heenan JR, Kallio K, Huppert D, Remington SJ. Structure and mechanism of the photoactivatable green fluorescent protein. *J Am Chem Soc* 2009;131(12):4176–7. <https://doi.org/10.1021/ja808851n>.
- Jayaraman S, Haggie P, Wachter RM, Remington SJ, Verkmann AS. Mechanism and cellular applications of a green fluorescent protein-based halide sensor. *J Biol Chem* 2000;275(9):6047–50. <https://doi.org/10.1074/jbc.275.9.6047>.

- [47] Lin C-Y, Romei MG, Oltrogge LM, Mathews IJ, Boxer SG. Unified model for photophysical and electro-optical properties of green fluorescent proteins. *J Am Chem Soc* 2019;141(38):15250–65. <https://doi.org/10.1021/jacs.9b07152>.
- [48] Ganesan S, Ameer-beg SM, Ng TTC, Vojnovic B, Wouters FS. A dark yellow fluorescent protein (YFP)-based resonance energy-accepting chromoprotein (REACH) for forster resonance energy transfer with GFP. *Proc Natl Acad Sci* 2006;103(11):4089–94. <https://doi.org/10.1073/pnas.0509922103>.
- [49] Banerjee S, Schenkelberg CD, Jordan TB, Reimertz JM, Crone EE, Crone DE, et al. Mispacking and the fitness landscape of the green fluorescent protein chromophore milieu. *Biochemistry* 2017;56(5):736–47. <https://doi.org/10.1021/acs.biochem.6b00800>.
- [50] Martynov VI, Savitsky AP, Martynova NY, Savitsky PA, Lukyanov KA, Lukyanov SA. Alternative cyclization in GFP-like proteins family. *J Biol Chem* 2001;276(24):21012–6. <https://doi.org/10.1074/jbc.M100500200>.
- [51] Bravaya KB, Subach OM, Korovina N, Verkhusha VV, Krylov AI. An insight into the common mechanism of the chromophore formation in the red fluorescent proteins: the elusive blue intermediate revealed. *J Am Chem Soc* 2013;134(5):2807–14. <https://doi.org/10.1021/ja2114568>.
- [52] Muslinkina L, Pletnev VZ, Pletneva NV, Ruchkin DA, Kolesov DV, Bogdanov AM, et al. Two independent routes of post-translational chemistry in fluorescent protein fusionred. *Int J Biol Macromol* 2020;155:551–9. <https://doi.org/10.1016/j.ijbiomac.2020.03.244>.
- [53] Dattilo AM, Decembrini F, Bracchini L, Focardi S, Mazzuoli S, Rossi C. Penetration of solar radiation into the waters of messina strait (Italy). *Ann Chim* 2005;95(3–4):177–84. <https://doi.org/10.1002/adic.200590020>.
- [54] Jerlov NG. Ultra-Violet Radiation in the Sea. *Nature* 1950;166(4211):111–2. <https://doi.org/10.1038/166111a0>.