ISSN 1063-7745, Crystallography Reports, 2018, Vol. 63, No. 2, pp. 191–195. © Pleiades Publishing, Inc., 2018.
Original Russian Text © A.A. Simanovskaya, T.V. Ivashina, Y.S. Zeifman, T.V. Fateeva, M.V. Krukova, A.N. Popov, G.S. Kachalova, T.V. Rakitina, 2018, published in Kristallografiya, 2018, Vol. 63, No. 2, pp. 208–212.

STRUCTURE OF MACROMOLECULAR COMPOUNDS

First Russian Crystallographic Congress

Mutation L232H Promotes Chromophore Maturation of EGFP-Based Fluorescent Fusion Proteins

A. A. Simanovskaya[a], T. V. Ivashina[b], Y. S. Zeifman[a], T. V. Fateeva[a], M. V. Krukova[a], A. N. Popov[c], G. S. Kachalova[a,d], and T. V. Rakitina[a,e,*]

[a] *National Research Centre "Kurchatov Institute," Moscow, 123098 Russia*
[b] *Skryabin Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences, Pushchino, Moscow Region, 142290, Russia*
[c] *European Synchrotron Radiation Facility, Grenoble, 38000 France*
[d] *Federal Research Centre "Fundamentals of Biotechnology," Russian Academy of Sciences, Moscow, 119071 Russia*
[e] *Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, 117997 Russia*
**e-mail: taniarakitina@yahoo.com*

Received May 23, 2017

Abstract—The L232H mutant of the enhanced green fluorescent protein (EGFP) was expressed and crystallized. An X-ray diffraction data set was collected from the crystals to 1.53 Å resolution. An analysis of the three-dimensional structure revealed a stacking interaction between the amino-acid residues H78 and H232, which contributes to the fastening of the C-terminal region of the protein in the vicinity of the chromophore and influences chromophore maturation of hybrid fluorescent proteins produced by fusion of the target proteins with the C-terminus of EGFP. This hypothesis was experimentally confirmed by investigating chromophore maturation of the hybrid proteins fused to the N- and C-termini of EGFP and EGFP-L232H.

DOI: 10.1134/S1063774518020256

INTRODUCTION

The use of fluorescent fusion proteins is a powerful tool for imaging and investigating living cells, primarily, the intracellular transport and localization of individual proteins. Currently, fluorescent proteins derived from different organisms and their mutant forms are used as fluorescent biological labels [1]. Meanwhile, the enhanced green fluorescent protein (EGFP) [2], which was constructed by site-directed mutagenesis based on the green fluorescent protein from the jellyfish *A. victoria* (GFP) [3], remains the most popular tool. The green fluorescent protein has a rather low fluorescence intensity, and its spectrum exhibits the main absorption peak in the short-wavelength range (<400 nm), in which UV light lethal to cells has to be used. In order to improve the spectral characteristics of GFP, substitutions of amino-acid residues forming the chromophore, F64L and S65T, were selected [4]. The mutant variant of GFP (EGFP) has only one excitation peak at a wavelength of 488 nm and high photostability.

In order to use EGFP in different fields of molecular and cell biology, commercial vectors, including the most commonly employed pEGFP-N1,2,3 and pEGFP-C1,2,3 vectors, were constructed. These vectors allow engineering and expression of target proteins fused to the N- or C-termini of EGFP (http://www.clontech.com/US/Support/Applications/Using_Fluorescent_Proteins/EGFP_Vectors). Many years of experience applying these vectors showed that, regardless of the expression system used, the production of target fusion proteins is almost always accompanied by a decrease or even the loss of fluorescence of the EGFP chromophore apparently due to inefficient chromophore maturation.

The goal of this study was to find factors responsible for a decrease in the efficiency of chromophore maturation of EGFP-containing chimeric proteins. It should be noted that, apart from the F64L and S65T substitutions introduced in order to improve the fluorescence properties of the chromophore, the EGFP amino-acid sequence has the H232L mutation. We made the reverse mutation L232H by site-directed

Table 1. Nucleotide sequences of the oligonucleotide primers used for cloning and mutagenesis

Primer name	Nucleotide sequence 5' → 3'
EGFP_L-Hf	GGGATCACTCaCGGCATGGACGAG
EGFP_L-Hr	CTCGTCCATGCCGtGAGTGATCCC
EGFPmut_f	TTTTTCCATGGTGAGCAAGGGCGAG
EGFPmut_r	CCTTGAATTCTGCCTTGTACAG
EGFPmut_f-1	ATAGAATTCATGGTGAGCAAGGG
EGFPmut_r-1	TTTTTCTCGAGTTACTTGTACAGCTCGTCC

mutagenesis and determined the crystal structure of the EGFP-L232H protein. The analysis of the three-dimensional structure of this protein revealed a stacking interaction between the residues H232 and H78, which contributes to the fastening of the C-terminal region of the protein in the vicinity of the chromophore. This stacking interaction may be one of the factors promoting chromophore maturation, particularly of hybrid proteins, in which the target proteins are fused to the C-terminus of EGFP. This hypothesis was experimentally confirmed by investigating chromophore maturation of hybrid proteins fused to the N- and C-termini of EGFP and EGFP-L232H.

MATERIALS AND METHODS

Site-directed mutagenesis. The commercial plasmid pEGFP-C1 [4] was used as a template for site-directed mutagenesis. The polymerase chain reaction (PCR) was performed using the Kod Hot start DNA polymerase (Novagen, USA) and the EGFP_L-Hf and EGFP_L-Hr primers (Table 1) containing required nucleotide substitutions in the corresponding codons. The PCR product was treated with DpnI restriction endonuclease for one hour at 37°C to remove the starting methylated DNA and then used for the transformation of competent *E. coli* DH5α cells. The presence of the mutation was confirmed by sequencing the pEGFP-L232H plasmid.

Preparation of plasmid constructs for synthesis of hybrid proteins fused to EGFP and EGFP-L232H. Recombinant constructs were prepared by standard molecular biology techniques [5] using the pHisTEV plasmid [6] and the gene encoding the HU protein from *M. Gallisepticum (huMgal)* [7–9]. The coding region of *huMgal* flanked by EcoRI restriction sites, which allowed additional insertion of the genes for fluorescent proteins in the 5' or 3' regions of *huMgal*, was prepared by PCR and then inserted into the pHisTEV plasmid at the NcoI and XhoI sites. The bacterial expression of the HU protein fused to EGFP will enable one to study in vivo the influence of stress conditions on the degree of condensation of bacterial genome.

The pHis-EGFP-L232H_HUMgal and pHis-EGFP_HUMgal constructs were generated by PCR on the pEGFP-L232H and pEGFP-C1 plasmid templates using the EGFPmut_f and EGFPmut_r primers (Table 1). The PCR products were treated with NcoI and EcoRI restriction enzymes and ligated with the pHis-HUMgal plasmid, which was treated with the same restriction enzymes. The ligation mixture was used for the transformation of competent *E. coli* MachI cells. The isolated plasmids were employed for the transformation of *E. coli* BL21(DE3) cells.

The pHis-HUMgal_EGFP-L232H and pHis-HUMgal_EGFP constructs were generated in a similar fashion but using the EGFPmut_f-1 and EGFPmut_r-1 primers (Table 1) and the EcoRI and XhoI restriction sites.

Study of chromophore maturation of fusion proteins by SDS-PAGE electrophoresis. To evaluate the efficiency of chromophore maturation of EGFP- and EGFP-L232H-containing fusion proteins, we compared the fluorescence levels of cell lysates separated by SDS-PAGE electrophoresis according to Laemmli [10]. Producer strain cells were incubated at 37°C until OD_{600} reached 0.8. Then isopropyl β-D-1-thiogalactopyranoside was added as the inducing agent to the final concentration of 1 mM, the incubation was performed at 20°C for a specified period of time, the cells were centrifuged at 5000 rpm for 15 min, and the supernatant was removed. The precipitates were lysed

Table 2. X-ray diffraction data collection and structure refinement statistics for the EGFP-L232H protein

Data-collection statistics	
Space group	*C*121
a, *b*, *c*, Å; α, β, γ, deg	84.59, 67.93, 47.27; 90.00, 118.66, 90.00
Resolution limits, Å	12.50–1.53
Number of measured reflections	34978
Completeness, %	98.9 (12.50–1.53) 99.5 (1.58–1.53)
Redundancy	3.4 (12.50–1.53) 3.3 (1.58–1.53)
R_{merge}	0.07 (12.50–1.53) 0.51 (1.58–1.53)
$\langle I/\sigma(I)\rangle$	8.6 (12.50–1.53) 2.0 (1.58–1.53)
Structure refinement statistics	
R, R_{free}	0.17 (12.50–1.53) 0.21 (1.58–1.53)
Number of reflections	35411
R_{free}, number of reflections in the test set	1733 (5.2%)
Wilson B factor, Å^2	16.1
Correlation coefficient between F_o and F_c	0.97
Total number of atoms in the model	2264
Number of protein atoms	1935
Number of atoms of water molecules	293
Number of atoms of PEG chains	36
Average B factor, Å^2	
Polypeptide backbone atoms	17.6
Polypeptide side-chain atoms	21.5
All atoms	19.5
Atoms of water molecules	35.5
Atoms of PEG	41.3
Ramachandran statistics, %	
Residues in allowed regions	98.7
Residues in generously allowed regions	1.3

in SDS-PAGE sample buffer containing 0.5% SDS and were loaded onto 12% polyacrylamide gel without heating. After electrophoresis the gels were imaged under UV light on a Molecular Imager ChemiDoc XRS System (Bio Rad, USA), and the protein bands were visualized by staining with Coomassie G-250 (Sigma, USA) [11]. Chromophore maturation was evidenced by the appearance of a protein band, which became fluorescent under UV light.

Crystallization of EGFP-L232H, the structure solution and refinement. The recombinant protein was expressed and purified according to a procedure described in [6–8]. The screening of crystallization conditions was performed using Hampton Research crystal screen kits (Index, Crystal Screen, PEGRx, and PEG_ION). Crystals were obtained in the following conditions: reservoir solution no. 69 (F9) from PEGRx (20 vol % 2-propanol, 0.1 MES monohydrate, pH 6.0, 20 wt % polyethylene glycol monomethyl ether 2000). The X-ray diffraction data set was collected from the crystals at 1.53 Å resolution at the beamline ID23-1 (ESRF, Grenoble). The structure was solved by the molecular replacement method using the EGFP structure as the starting model. The structure was refined to R-fact = 0.17 (R-fact free = 0.21) with the REFMAC program. The X-ray data collection and structure refinement statistics for the EGFP-L232H protein are given in Table 2. The protein structure was deposited in the Protein Data Bank (PDB ID 5N9O).

RESULTS AND DISCUSSION

It is known that in fluorescent proteins, in particular in EGFP [12], the chromophore is protected from fluorescence quenching by bulk solvent and molecular oxygen because it is held by two small α-helices near the center of the inner cavity. The latter is a β-barrel formed by eleven antiparallel β-sheets and capped on the top and bottom ends with loop regions that connect the β-sheets (Fig. 1). In the hybrid protein, the EGFP fold may be somewhat destabilized, resulting in a decrease in the efficiency of chromophore maturation and, correspondingly, in a decrease in fluorescence intensity. Meanwhile, appropriate amino-acid substitutions can exert a stabilizing effect and promote chromophore maturation of EGFP.

An analysis of the three-dimensional structure of the EGFP-L232H mutant produced in this study revealed a contact between the imidazole rings of the residues H232 and H78 (Fig. 2a). The H232–H78 stacking interaction stabilizes the mutual arrangement of unstructured regions of the EGFP polypeptide chain (residues 75–78 and 229–232) and the adjoining eleventh β-sheet (residues 217–228) in the nearest

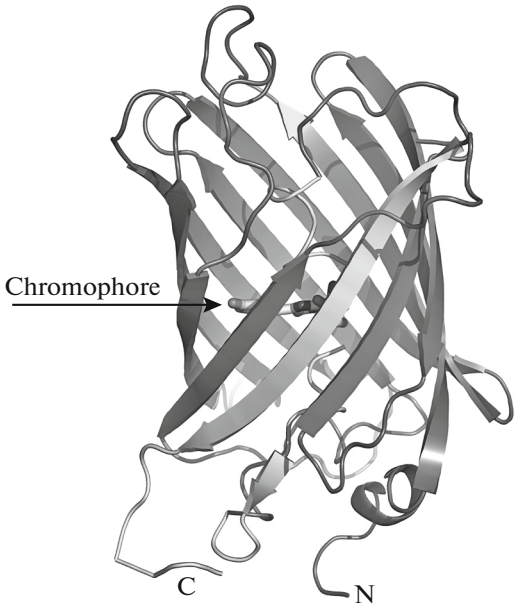

Fig. 1. Three-dimensional structure of the EGFP protein: 11 β-sheets that form a β-barrel with a chromophore in its center.

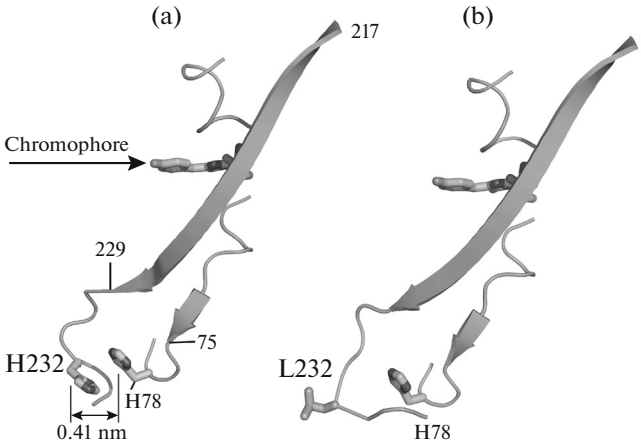

Fig. 2. (a) Fragment of the three-dimensional structure of EGFP-L232H depicting the stacking interaction between the residues H78 and H232. The secondary-structure elements of the protein and residues discussed in the paper are denoted. (b) The same fragment of EGFP prior to mutation is shown for comparison.

environment of the chromophore. Stabilization of the C-terminal region of EGFP may promote chromophore maturation of fluorescent fusion proteins.

The positive effect of the L232H mutation on chromophore maturation was confirmed in a model experiment aimed to compare the fluorescence intensity of fluorescent fusion proteins containing EGFP and the EGFP-L232H mutant (Fig. 3). The fluorescent fusion proteins were expressed in *E. coli* cells. Total cell lysates were subjected to electrophoresis under denaturing conditions. As can be seen in the electrophoretic patterns shown in Fig. 3a, despite the same level of synthesis of recombinant proteins, the target proteins fused to the C-terminus of the fluorescent protein exhibited fluorescence only if the samples contained EGFP-L232H, whereas there is no such difference for the hybrid proteins, in which the target proteins are fused to the N-terminus of the fluorescent protein (Fig. 3b).

Therefore, our data confirmed that the replacement of leucine by histidine at position 232 has a positive effect on chromophore maturation of fluorescent fusion proteins.

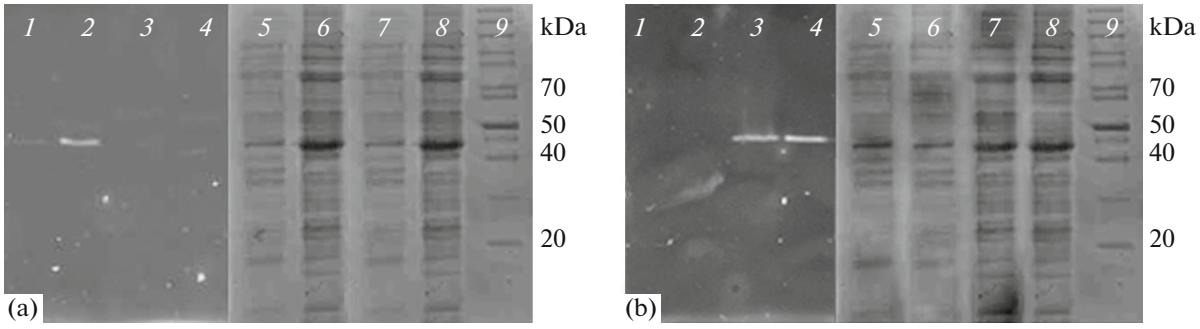

Fig. 3. Comparison of the fluorescence intensity of fluorescent fusion proteins in SDS-PAGE. UV-induced fluorescence on the right side of the gel and staining of the gel with Coomassie G-250, which demonstrates similar levels of synthesis of the proteins under consideration, on the left side of the gel. Lane 9 is the marker PageRuler™ Unstained Protein Ladder #SM0661, Fermentas; (a) the target protein is fused to the C-terminus of EGFP-L232H (*1*, *2*, *5*, *6*) or EGFP (*3*, *4*, *7*, *8*). Cell lysates were loaded onto SDS-PAGE 30 min (*1*, *3*, *5*, *7*) and 3 h (*2*, *4*, *6*, *8*) after the induction; (b) the target protein is fused to the N-terminus of EGFP (*1*, *3*, *5*, *7*) or EGFP-L232H (*2*, *4*, *6*, *8*). Cell lysates were loaded onto SDS-PAGE 30 min (*1*, *2*, *5*, *6*) and 3 h (*3*, *4*, *7*, *8*) after the induction.

ACKNOWLEDGMENTS

The study was supported by the Ministry of Education and Science of the Russian Federation (unique identifier of the project RFMEFI61614X0003) in part of X-ray diffraction data collection and the structure solution and refinement and by the Russian Science Foundation (agreement No. 15-14-00063) in part of genetic manipulations with the *huMgal* gene and its bacterial expression.

REFERENCES

1. R. N. Day and M. W. Davidson, Chem. Soc. Rev. **38**, 4 (2009).
2. G. Zhang, V. Gurtu, and S. R. Kain, Biochem. Biophys. Res. Commun. **227**, 707 (1996).
3. D. C. Prasher, V. K. Eckenrode, W. W. Ward, et al., Gene **111**, 229 (1992).
4. B. P. Cormack, R. H. Valdivia, and S. Falkow, Gene **173**, 33 (1996).
5. M. R. Green and J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Fourth Edition) (Cold Spring Harbor Laboratory Press, 2012).
6. V. Timofeev, E. Slutskaya, M. Gorbacheva, et al., Acta Crystallogr. F **71**, 951 (2015).
7. A. Yu. Nikolaeva, V. I. Timofeev, K. M. Boiko, et al., Crystallogr. Rep. **60**, 880 (2015).
8. D. A. Altukhov, Yu. K. Agapova, A. V. Vlaskina, et al., Vestn. Mosk. Univ. **57** (4), 226 (2016).
9. D. A. Altukhov, A. A. Talyzina, Y. K. Agapova, et al., J. Biomol. Struct. Dyn. (2016); doi 10.1080/07391102.2016.1264893
10. U. K. Laemmli, Nature **227**, 680 (1970).
11. L. Syrovy and Z. J. Hodny, Chromatography **569**, 175 (1991).
12. M. Ormö, A. B. Cubitt, K. Kallio, et al., Science **273**, 1392 (1996).

Translated by T. Safonova