



Article

LSSmScarlet, dCyRFP2s, dCyOFP2s and CRISPRed2s, Genetically Encoded Red Fluorescent Proteins with a Large Stokes Shift

Oksana M. Subach ¹, Anna V. Vlaskina ¹, Yuliya K. Agapova ¹, Pavel V. Dorovatovskii ¹, Alena Y. Nikolaeva ^{1,2}, Olga I. Ivashkina ^{1,3,4}, Vladimir O. Popov ^{1,2,5}, Kiryl D. Piatkevich ^{6,7,8}, Maria G. Khrenova ^{2,9}, Tatiana A. Smirnova ¹⁰, Konstantin M. Boyko ² and Fedor V. Subach ^{1,*}

¹ Complex of NBICS Technologies, National Research Center “Kurchatov Institute”, 123182 Moscow, Russia; Subach_OM@nrcki.ru (O.M.S.); vlaskina_av@nrcki.ru (A.V.V.); agapova.jk@gmail.com (Y.K.A.); Dorovatovskiy_PV@nrcki.ru (P.V.D.); Nikolaeva_AY@nrcki.ru (A.Y.N.); ivashkina_oi@nrcki.ru (O.I.I.); vpopov@fbras.ru (V.O.P.)

² Bach Institute of Biochemistry, Research Centre of Biotechnology of the Russian Academy of Sciences, 119071 Moscow, Russia; mkhrenova@lcc.chem.msu.ru (M.G.K.); kmb@inbi.ras.ru (K.M.B.)

³ Laboratory for Neurobiology of Memory, P.K. Anokhin Research Institute of Normal Physiology, 125315 Moscow, Russia

⁴ Institute for Advanced Brain Studies, M.V. Lomonosov Moscow State University, 119991 Moscow, Russia

⁵ Faculty of Biology, M.V. Lomonosov Moscow State University, 119991 Moscow, Russia

⁶ School of Life Sciences, Westlake University, Hangzhou 310024, China; kiryl.piatkevich@westlake.edu.cn

⁷ Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou 310024, China

⁸ Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou 310024, China

⁹ Department of Chemistry, Lomonosov Moscow State University, 119991 Moscow, Russia

¹⁰ Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Institute of Gene Biology, Russian Academy of Sciences, 119334 Moscow, Russia; smirnovatiana@mail.ru

* Correspondence: subach_fv@nrcki.ru; Tel.: +7-499-196-7100-3389

Citation: Subach, O.M.; Vlaskina, A.V.; Agapova, Y.K.; Dorovatovskii, P.V.; Nikolaeva, A.Y.; Ivashkina, O.I.; Popov, V.O.; Piatkevich, K.D.; Khrenova, M.G.; Smirnova, T.A.; et al. LSSmScarlet, dCyRFP2s, dCyOFP2s and CRISPRed2s, Genetically Encoded Red Fluorescent Proteins with a Large Stokes Shift. *Int. J. Mol. Sci.* **2021**, *22*, 12887. <https://doi.org/10.3390/ijms222312887>

Academic Editor: Konstantin K. Turoverov

Received: 8 September 2021

Accepted: 24 November 2021

Published: 28 November 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Genetically encoded red fluorescent proteins with a large Stokes shift (LSSRFPs) can be efficiently co-excited with common green FPs both under single- and two-photon microscopy, thus enabling dual-color imaging using a single laser. Recent progress in protein development resulted in a great variety of novel LSSRFPs; however, the selection of the right LSSRFP for a given application is hampered by the lack of a side-by-side comparison of the LSSRFPs’ performance. In this study, we employed rational design and random mutagenesis to convert conventional bright RFP mScarlet into LSSRFP, called LSSmScarlet, characterized by excitation/emission maxima at 470/598 nm. In addition, we utilized the previously reported LSSRFPs mCyRFP1, CyOFP1, and mCRISPRed as templates for directed molecular evolution to develop their optimized versions, called dCyRFP2s, dCyOFP2s and CRISPRed2s. We performed a quantitative assessment of the developed LSSRFPs and their precursors in vitro on purified proteins and compared their brightness at 488 nm excitation in the mammalian cells. The monomeric LSSmScarlet protein was successfully utilized for the confocal imaging of the structural proteins in live mammalian cells and multicolor confocal imaging in conjugation with other FPs. LSSmScarlet was successfully applied for dual-color two-photon imaging in live mammalian cells. We also solved the X-ray structure of the LSSmScarlet protein at the resolution of 1.4 Å that revealed a hydrogen bond network supporting excited-state proton transfer (ESPT). Quantum mechanics/molecular mechanics molecular dynamic simulations confirmed the ESPT mechanism of a large Stokes shift. Structure-guided mutagenesis revealed the role of R198 residue in ESPT that allowed us to generate a variant with improved pH stability. Finally, we showed that LSSmScarlet protein is not appropriate for STED microscopy as a consequence of LSSRed-to-Red photoconversion with high-power 775 nm depletion light.

Keywords: genetically encoded red fluorescent proteins; protein engineering; fluorescence imaging; large Stokes shift; LSSmScarlet; dCyRFP2s; dCyOFP2s; CRISPRed2s; red fluorescent; fluorescent protein

1. Introduction

Genetically encoded red fluorescent proteins with a large Stokes shift (LSSRFPs) can be efficiently excited by blue or cyan light while emitting red fluorescence, thus characterized by more than a 100 nm difference between excitation and emission maxima. The large Stokes shift in RFPs is induced by deprotonation of the chromophore in the excited state supported by a hydrogen bond network formed between tyrosine hydroxyl of the chromophore and side chains of the Glu or Asp amino acids located in the close vicinity. The chromophore deprotonation during excitation, also known as excited-state proton transfer (ESPT), has been studied for several LSSRFPs using structural, spectroscopic and mutagenic approaches [1]. The LSSRFPs provide an additional fluorescence color for spectrally multiplexed fluorescence microscopy and are applied for multicolor visualization of the processes in living cells [2]. Due to efficient excitation at 488 nm and 960 nm under one-photon and two-photon microscopy, respectively, LSSRFPs can be simultaneously imaged with green FPs and indicators using dual-emission under common one- and two-photon lasers [3]. Multiple LSSRFPs were reported to date and among them, CyOFP1, CyRFP1, mCyRFP1 and mCRISPRed LSSRFPs stand out in terms of molecular brightness at a 488 nm excitation (Table S1). However, a direct side-by-side comparison for these proteins is not available, thus creating a need for end-users to test several LSSRFPs to select an optimal variant for a given application. Moreover, it would be desirable to make brighter LSSRFPs using both the brightest RFP as a template, such as mScarlet, and the brightest LSSRFPs as templates, such as CyOFP1, mCyRFP1 and mCRISPRed.

In this manuscript, we report the development of a set of novel LSSRFPs characterized by efficient excitation at a 488 nm laser line. We compared the brightness of the developed LSSRFPs to their precursors in a solution on purified proteins and in the cytosol of mammalian cells. We engineered a version of the brightest RFP mScarlet, with a large Stokes shift, named LSSmScarlet. Monomeric properties of LSSmScarlet correlated with its good localization in the fusions with the most popular structural proteins in mammalian cells. LSSmScarlet was successfully applied for three-color confocal imaging and single excitation dual-emission two-photon imaging of live mammalian cells. We solved the crystal structure of LSSmScarlet with 1.4 Å resolution and performed quantum mechanics/molecular mechanics molecular dynamic (QM/MM MD) simulations and mutagenesis analyses to elucidate the role of amino acids involved in ESPT. Finally, we applied LSSmScarlet protein for STED microscopy.

2. Results and Discussion

2.1. Developing Red Fluorescent Proteins with Large Stokes Shift based on mScarlet, mCyRFP1, CyOFP1 and mCRISPRed Proteins in *E.coli*.

For the development of LSSmScarlet, the mScarlet protein was subjected to one round of rational mutagenesis at amino acid positions 147, 162 and 164 followed by eight rounds of random mutagenesis on the mutants having the LSSRed-like phenotype and screening for the highest brightness achieved upon expression in *E.coli* cells. The positions 147, 162 and 164 (corresponding to the positions 148, 165 and 167 according to EGFP enumeration) for the mScarlet directed mutagenesis were suggested according to the residues found in the key positions responsible for the LSSRed-like phenotype for LSSmKates, LSSmOrange, mKeima and other fluorescent proteins [4,5]. Three overlap libraries with mScarlet/S147D,E/I162X/M164X, mScarlet/S147X/I162D,E/M164X and mScarlet/S147X/I162X/M164D,E mutants (Figure 1) were cloned in fusion with sfGFP protein to assist the folding of mutants [6–8]. However, sfGFP was so bright that it leaked into the LSSRed channel (excitation at 480/40 nm and emission at 620/60 nm) and actually hampered the screening procedure. Despite this obstacle, we found that mScarlet/S147D/M164C, mScarlet/S147G/I164E and mScarlet/S147D mutants exhibited the LSSRed-like phenotype, i.e., had red fluorescence under blue excitation. We used a mixture of mutants as a template for the first round of random mutagenesis followed by

screening for the highest brightness in *E.coli* (or purified protein). After eight rounds of directed molecular evolution, we selected a final variant mScarlet/T74I/Y84H/K122E/W144L/S147D/F178V/K179R mutant, named LSSmScarlet, characterized by the highest molecular brightness at a 488 nm excitation and the absence of red fluorescence at a 561 nm excitation. LSSmScarlet had seven mutations relative to the mScarlet template; five and two were inside and outside of the β -barrel (Figure 1).

	1	10	20	30	40	50	60
mScarlet	MVSKGEAVIKEFMRFKVHMEGSMNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFSWD						
LSSmScarlet	MVSKGEAVIKEFMRFKVHMEGSMNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFSWD						
mCyRFP1	MVSKGEELIKENMRSKLYLEGSVNGHQFKCTHEGEGKPYEGKQTARIKVVEGGPLPFAFD						
dCyRFP2s	MVSKGEELIKE S MRSKLYLEGSVNGHQFKC E HEGEGKPYEGKQTARIKVVEGGPLPFAFD						
CyOF1	MVSKGEELIKENMRSKLYLEGSVNGHQFKCTHEGEGKPYEGKQTNRIKVVEGGPLPFAFD						
dCyOF2s	MVSKGEELIKENMRSKLYLEGSVNGHQFKCTHEGEGKPYEGKQTNRIKVVEGGPLPFAFD						
mCRISPRed	MVSKGEELIKENMRMKVMEGSVNGHQFKCTGEGEGRPYEGVQTMRIKVIIEGGPLPFAFD						
CRISPRed2s	MVSKGEELIKE S MRMKVMEGSVNGHQFKCTGEGEGRPYEGVQTMRIKVIIEGGPLPFAFD						
	61	70	80	90	100	110	120
mScarlet	ILSPQ MYG SRAFTKHPADIPDYKQSFPEGFKWERVMNFEDGGAVTVTQDTSLEDGTLI						
LSSmScarlet	ILSPQ MYG SRA F KHPADIPDY K QSFPEGFKWERVMNFEDGGAVTVTQDTSLEDGTLI						
mCyRFP1	ILATM MYG SKVFIKYPADLPDYFKQSFPEGFTWERVMVFEDGGVLTATQDTSLEDGELI						
dCyRFP2s	ILATM MYG SKVFIKYPADLPDYFKQSFPEGFTWERVMVFEDGGVLTATQDTSLEDG E LI						
CyOF1	ILATHE MYG SKVFIKYPADLPDYFKQSFPEGFTWERVMVFEDGGVLTATQDTSLEDGELI						
dCyOF2s	ILATHE MYG SKVFIKYPADLPDYFKQSFPEGFTWERVMVFEDGGVLTATQDTSLEDGELI						
mCRISPRed	ILATSE MYG SRTFIKYPADIPDFFKQSFPEGFTWERVTRYEDGGVTVTQDTSLEDGELV						
CRISPRed2s	ILATSE MYG SRTFIKYPADIPDFFKQSFPEGFTWERVTRYEDGGVTVTQDTSLEDGELV						
	121	130	140	150	160	170	180
mScarlet	YKVKLRGTNFPDPGPVMQKKTMGWEASTERLYPEDGVLKGDIKMALRLKDGGRYLADFKT						
LSSmScarlet	Y E VKLRGTNFPDPGPVMQKKTMG E AD E TERLYPEDGVLKGDIKMALRLKDGGRYLAD V RT						
mCyRFP1	YNVKLRGVNFPANGPVMQKKTTLGWEPSTETMYPADGGLEGRCDKKLVGGGHLHVNFKT						
dCyRFP2s	YNVKLRGVNFP A GPVMQKKTTLGWEPSTETMYPADGGLEGRCD K LLKLVGGGHLHVNFKT						
CyOF1	YNVKVRGVNFPANGPVMQKKTTLGWEPSTETMYPADGGLEGRCDKALKLVGGGHLHVNFKT						
dCyOF2s	YNVKVRGVNFPANGPVMQKKTTLGWEPSTETMYPADGGLEGRCDKALKLVGGGHLH V SFKT						
mCRISPRed	YNVKVRGVNFPNSGPVMQKKTGWEADTEMYPADGGLRGYLDRAKVDGGGHLHCNFVT						
CRISPRed2s	YNVK V SGVNFPNSGPVMQKKTGWEADTEMYPADGGLRGYLDRAKVDGGGHLHCNFVT						
	181	190	200	210	220	230	
mScarlet	TYKAKKPV---QMPGAYNVDRKLDITSHNEDYTVVEQYERSEGRHS--TGGMDELYK						
LSSmScarlet	TYKAKKPV---QMPGAYNVDRKLDITSHNEDYTVVEQYERSEGRHS--TGGMDELYK						
mCyRFP1	TYKSKKPV---KMPGVHYVDRRLERIEADNETYVEQYEHAVARYSNLGGGMDELYK						
dCyRFP2s	TY E SKKPV---KMPGVHYVDRRLERIEADNETYVEQYEHAVARYSNLGGGMDELYK						
CyOF1	TYKSKKPV---KMPGVHYVDRRLERIEADNETYVEQYEHAVARYSNLGGGMDELYK						
dCyOF2s	TYKSKKPV---KMPGVHYVDRRLERIEADNETYVEQYEHAVARYSNLGGGMDELYK						
mCRISPRed	TYRSKKTVDGIKMPGVHAVDHRLERIEESDNETYVVQREVAVAKYSNLGGGMDELYK						
CRISPRed2s	TYRSKKTVDGIKMPGVHAVDHRLERIEESDNETYVVQREVAVAK S NLGGGMDELYK						

Figure 1. Alignment of the amino acid sequences for the developed LSSRFPs (LSSmScarlet, dCyRFP2s and dCyOF2s) and their progenitors (mScarlet, mCyRFP1 and CyOF1). Mutations in the developed LSSRFPs are highlighted in green color. Chromophore-forming residues MYG are selected in bold. Residues buried in the β -barrel are in grey.

To enhance the brightness of CyOF1, mCyRFP1 and mCRISPRed LSSRFPs in bacterial cells, we subjected these proteins to one, four and four rounds of random mutagenesis followed by screening for the largest brightness under a fluorescence microscope,

respectively. As a result of directed molecular evolution, we selected the CyOFP1/N177S, mCyRFP1/N12S/T31I/E118G/N133Y/K165L/K183R and mCRISPRed/N12Y/R126S/N227D mutants, named dCyOFP2s, dCyRFP2s and CRISPRed2s, respectively (Figures 1 and S1). The dCyOFP2s, dCyRFP2s and CRISPRed2s proteins had one, six and three mutations vs. respective parental proteins, and all mutations found in these proteins were outside of the β -barrel (Figure 1).

2.2. In Vitro Characterization of the Purified LSSmScarlet, dCyRFP2s, dCyOFP2s and CRISPRed2s LSSRFPs

First, we characterized the spectral properties and molecular brightness of the developed and progenitor LSSRFPs (Figure 2a,b, and Table 1). LSSmScarlet, dCyRFP2s, dCyOFP2s and CRISPRed2s LSSRFPs had maxima of absorption/excitation/emission in the range of 452–508/464–516/590–598 nm, respectively (Figure 2a,b and Table 1). These maxima were close to the progenitor LSSRFPs (Figure 2a,b, and Table 1). Among developed LSSRFPs, LSSmScarlet and CRISPRed2s had the largest Stokes shifts of 128 and 126 nm, respectively (Table 1). According to the molecular brightness at the maximum of excitation, the developed LSSRFPs were ranked as following dCyOFP2s = dCyRFP2s > LSSmScarlet > CRISPRed2s (Table 1). At a 488 nm excitation, the order of the molecular brightness was similar: dCyOFP2s > dCyRFP2s > LSSmScarlet > CRISPRed2s (Table 1). The difference in brightness at a 488 nm excitation between the brightest dCyOFP2s and dimmest CRISPRed2s was 2.3-fold. Parental mCyRFP1 was 1.33-fold less bright than dCyRFP2s; however, original CyOFP1 and mCRISPRed LSSRFPs were 1.33- and 1.13-fold brighter than dCyOFP2s and CRISPRed2s, respectively. Hence, LSSmScarlet and CRISPRed2s proteins were outstanding in terms of the largest Stokes shift, but dCyOFP2s demonstrated the largest molecular brightness at a 488 nm excitation.

Table 1. In vitro properties of the LSSRFPs. ^a Data from [3]. ^b QY was determined at pH 7.40. CyOFP1 (QY = 0.76 [3]) was used as the reference standard. ^c Extinction coefficients were determined by alkaline denaturation. ^d Molecular brightness was calculated as a product of the quantum yield and extinction coefficient at maximal absorption and normalized to the brightness of EGFP that has an extinction coefficient of 56,000 M⁻¹·cm⁻¹ and quantum yield of 0.6 [2]. ^e Brightness corrected for the decreased absorption at 488 nm. ^f Half-time to bleach until 50%. One-photon photobleaching was performed under metal halide lamp on droplets of protein solution (45 μ M final protein concentration) in oil. Under identical conditions, mEGFP had photobleaching half-time of 305 \pm 38 sec. Standard deviations are shown.

Proteins	Abs, Ex/Em (nm)	QY (%) ^b	ϵ (mM ⁻¹ ·cm ⁻¹) ^c	Brightness vs. EGFP (%)		pK _a	Monomeric State	Photo- bleaching Half-Time (s) ^f	Matura- tion Half- Time (min)
				ex. at Max ^d	ex. at 488 nm ^e				
LSSmScarlet	466, 470/598	43 \pm 2	30.2 \pm 0.6	39	34	1.91 \pm 0.01; 5.78 \pm 0.06	Monomer	87 \pm 16	61
mCyRFP1	514, 520/600	61.2 \pm 0.4	30.6 \pm 0.5	55	49	5.33 \pm 0.03; 9.41 \pm 0.12	Monomer	161 \pm 26	51
dCyRFP2s	508, 516/592	58 \pm 1	42 \pm 2	73	65	5.44 \pm 0.04; 9.817 \pm 0.004	Dimer	144 \pm 20	75
CyOFP1	503, 510/594	76 ^a	45.1 \pm 0.2	102	97	5.22 \pm 0.07; >10	Dimer	223 \pm 37	25
dCyOFP2s	503, 510/592	69 \pm 7	36 \pm 4	73	73	5.29 \pm 0.03; >10	Dimer	176 \pm 30	24
mCRISPRed	452, 462/592	42 \pm 2	29.1 \pm 0.9	36	27	2.14 \pm 0.02 9.8 \pm 0.3	Monomer	93 \pm 7	354
CRISPRed2s	452, 464/590	38.2 \pm 0.2	28.7 \pm 0.9	32	24	2.23 \pm 0.03 10.0 \pm 0.1	Monomer- Tetramer	79 \pm 13	90

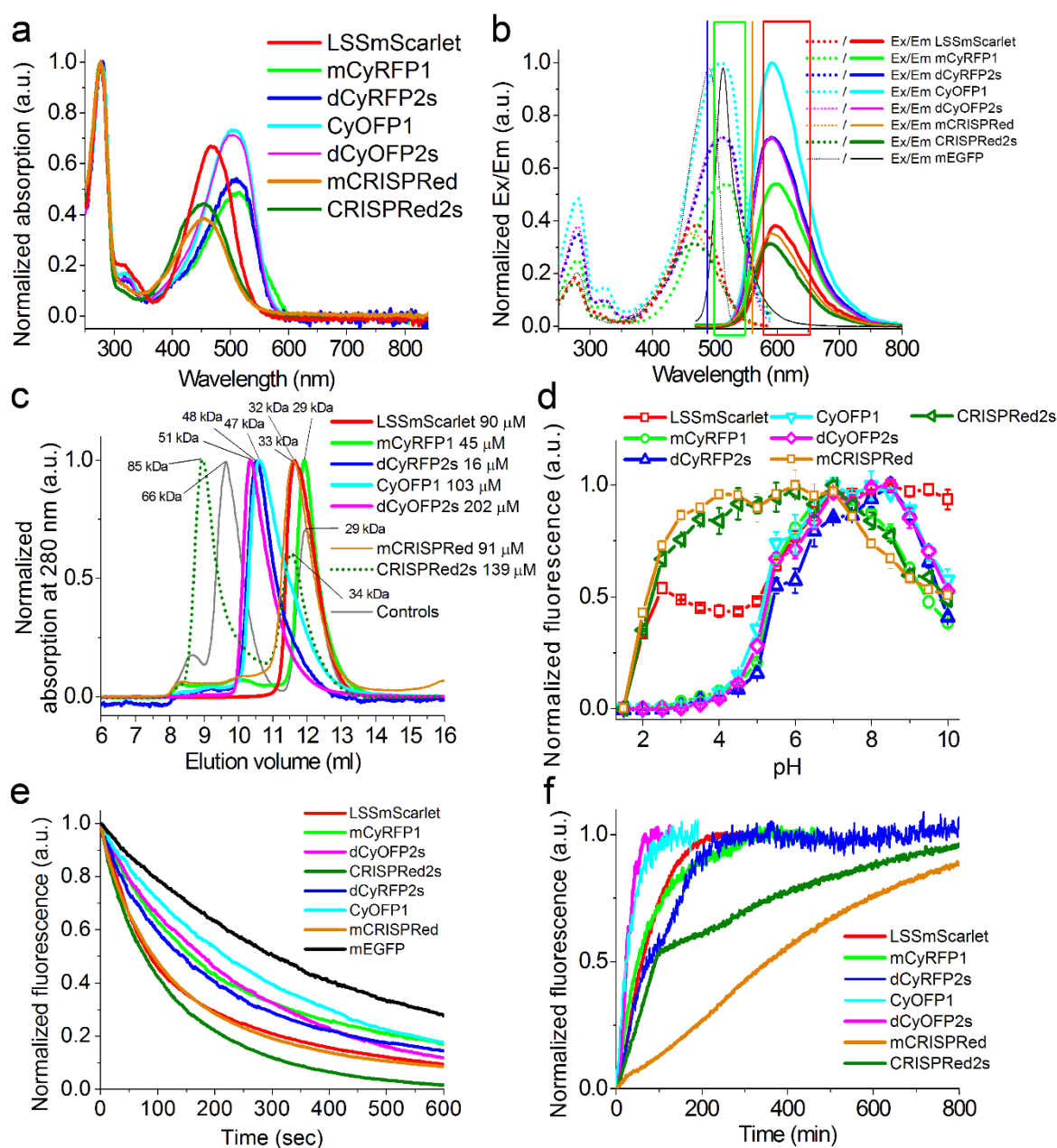


Figure 2. In vitro properties of the purified LSSRFPs proteins. **(a)** Absorption spectra for LSSRFPs proteins in PBS buffer at pH 7.40. **(b)** Excitation and emission spectra for LSSRFPs and mEGFP in PBS buffer at pH 7.40. The 488 and 561 nm excitation laser lines and green/red emission filters used for dual-color imaging are indicated as blue and orange vertical lines and green/red boxes, respectively. **(c)** Fast protein liquid chromatography of LSSRFPs proteins. LSSRFPs were eluted in 20 mM Tris-HCl (pH 7.80) and 200 mM NaCl buffer. The molecular weights of LSSRFPs were calculated from a linear regression of the dependence of logarithm of control molecular weights vs. elution volume (Figure S2). **(d)** Red fluorescence intensity for LSSRFPs as a function of pH. **(e)** Photobleaching of LSSRFPs and mEGFP proteins droplets in oil under continuous wide-field imaging using metal halide lamp (3.78 mW/cm² light power before objective lens). **(f)** Maturation of LSSRFPs in PBS buffer at pH 7.40. **(d–f)** Three-six replicates were averaged for analysis. Error bars represent the standard deviation.

The monomeric state of FP is important for protein labeling [9] and low cytotoxicity [10], so we next characterized the oligomeric state of the developed and parental LSSRFPs using fast protein liquid chromatography (FPLC) (Figures 2c and S2 and Table 1). On FPLC, LSSmScarlet, dCyRFP2s, dCyOFP2s and CRISPRRed2s eluted as a monomer, dimer, dimer and a mixture of the monomer and tetramer, respectively (Figures 2c and S2). Parental CyOFP1 was also dimer as dCyOFP2s; however, mCyRFP1 and mCRISPRRed were

monomers in contrast to dimeric dCyRFP2s and monomeric-tetrameric CRISPRed2s (Figure 2c and Table 1). Hence, among developed LSSRFPs, only LSSmScarlet is appropriate for the individual protein labeling.

Since the pH-stability of FP is important for imaging in different cellular organelles with varied pH values (5.0 in lysosomes to 7.5 in the cytosol [11]), we further characterized the dependence of the fluorescence of the developed and parental LSSRFPs on the variation of pH (Figure 2d and Table 1). Except for LSSmScarlet, all other developed LSSRFPs dCyRFP2s, dCyOFP2s and CRISPRed2s demonstrated around a 2-fold drop of their fluorescence at the increasing of the pH in the range of a pH of 7–10 with pK_a values of around 10; in turn, the fluorescence of LSSmScarlet was at maximum and did not change within this pH range (Figure 2d). Acidification resulted in the loss of the dCyRFP2s' and dCyOFP2s' fluorescence with pK_a values of 5.44 and 5.29, respectively (Figure 2d and Table 1). At acidification, the fluorescence of LSSmScarlet dropped by around 2-fold with a pK_a value of 5.78 and further diminished completely with a pK_a value of 1.91 (Figure 2d and Table 1). At acidic pH values, the fluorescence of CRISPRed2s was characterized by a pK_a value of 2.23 (Figure 2d and Table 1). The pH stabilities for the parental LSSRFPs were similar to their derivatives (Figure 2d and Table 1). The shapes of the excitation, emission and absorption spectra for LSSmScarlet protein were similar in the pH range 2.0–11.5 (Figure S3). However, at pH values above 11.5, the LSSmScarlet showed an additional RFP-like anionic form of the chromophore with absorption/excitation/emission peak maxima at 570/570/607 nm, respectively (Figure S3). The spectral properties of this RFP-like anionic form of the LSSmScarlet were similar to absorption/emission maxima for the mScarlet RFP observed at 569/598 nm, respectively [12]. Hence, among all LSSRFPs tested, LSSmScarlet demonstrated the highest pH-stability at alkaline pH values and, together with CRISPRed2s, better preserved its fluorescence at acidic pH values.

We next compared the one-photon photostability of the developed and parental LSSRFPs on purified proteins using microdroplets' suspension in oil under continuous illumination with a metal halide lamp at a 470/40 nm excitation and 63x oil objective lens (Figure 2e and Table 1). According to the measured photobleaching half-times, the developed LSSRFPs were ranked as dCyOFP2s > dCyRFP2s > LSSmScarlet > CRISPRed2s (Table 1). The difference in photostability between the most photostable dCyOFP2s and the least photostable CRISPRed2s was 2.2-fold. In turn, the photostabilities of the parental LSSRFPs were similar to photostabilities of the developed LSSRFPs (Figure 2e and Table 1). Under identical conditions, the commonly used mEGFP protein had a photostability 1.7–3.9-fold higher as compared to the developed LSSRFPs (Figure 2e and Table 1). Hence, the developed LSSRFPs demonstrated a one-photon photostability 1.7–3.9-fold lower than mEGFP, and the dCyOFP2s protein was the most photostable among them.

We further measured the maturation rate and efficiency for the engineered and original LSSRFPs at 37°C using purified proteins expressed in bacterial cells (Figure 2f and Table 1). According to the maturation half-times, the developed LSSRFPs were ranked as following dCyOFP2s < dCyRFP2s < LSSmScarlet < CRISPRed2s (Figure 2f and Table 1). The difference in maturation half-times for dCyOFP2s with the fastest maturation and CRISPRed2s with the slowest maturation was 3.8-fold. Parental CyOFP1 showed the same maturation half-time as dCyOFP2s; the mCyRFP1 progenitor matured 1.5-fold faster than dCyRFP2s, and mCRISPRed had a 3.9-fold slower maturation than CRISPRed2s (Figure 2f and Table 1). We also estimated the maturation efficiency of the LSSRFPs according to normalization of the extinction coefficient calculated at 280 nm to the extinction coefficient determined by alkaline denaturation (Table S2). All LSSRFPs tested revealed a maturation efficiency of 68% and higher. Hence, all developed LSSRFPs matured efficiently, and among all of them, dCyOFP2s exhibited the fastest maturation rate.

2.3. Brightness of the LSSmScarlet, dCyRFP2s, dCyOFP2s and CRISPRed2s LSSRFPs in Cultured Mammalian Cells

We next compared the intracellular brightness of the developed and progenitor LSSRFPs in the cytosol of live HeLa cells at a 488 nm excitation light under a confocal microscope (Figure 3 and Table 2). To account for differences in the expression levels, LSSRFPs were co-expressed with the EGFP protein via the self-cleavable P2A peptide linker. All LSSRFPs were evenly distributed in the cytosol and nucleus of the HeLa cells similarly to EGFP (Figures 3a and S4). According to the measured intracellular brightness normalized to the green fluorescence intensity of EGFP, the developed LSSRFPs were ranked as following dCyOFP2s = dCyRFP2s > LSSmScarlet > CRISPRed2s (Figure 3 and Table 2). The difference in brightness between the brightest dCyRFP2s and dimmest CRISPRed2s was 3.4-fold (Figure 3 and Table 2). Parental mCyRFP1 and CyOFP1 were 1.29- and 1.15-fold dimmer than the dCyRFP2s and dCyOFP2s proteins, respectively (Figure 3 and Table 2). The mCRISPRed progenitor did not show a statistically significant difference in brightness as compared to CRISPRed2s (Figure 3 and Table 2). We noticed a difference in brightness of LSSRFPs expressed in HeLa cells and purified from bacteria (Table 2). This difference was the largest for mCRISPRed and CyOFP1 proteins; these proteins were 2.08- and 1.93-fold brighter according to the molecular brightness determined on purified proteins as compared to the brightness in HeLa cells. The observable differences in brightness of FPs expressed in mammalian cells over bacterially expressed and purified proteins might be associated with their folding and maturation efficiency, and/or stability inside mammalian and bacterial cells [13–15]. Overall, dCyRFP2s and dCyOFP2s demonstrated the highest brightness in the cytosol of mammalian cells and were 1.29- and 1.15-fold brighter than their mCyRFP1 and CyOFP1 progenitors, respectively.

Table 2. Brightness of LSSRFPs in HeLa cells at 488 nm excitation. ^a Brightness of LSSRFPs normalized to the brightness of the LSSmScarlet protein was estimated in HeLa cells relative to EGFP in LSSRFPs-P2A-EGFP fusion using the same 488 nm excitation light and 525/50 BP and 617/73 BP emission filters for EGFP and LSSRFPs, respectively. ^b Statistical significance of the difference between the normalized brightness for LSSmScarlet and the respective LSSRFP. NA, not applicable. ^c Brightness of the purified proteins corrected for decreased absorption at 488 nm excitation from Table 1 normalized to the brightness of LSSmScarlet.

Proteins	Normalized Brightness vs. EGFP in HeLa Cells ^a (<i>p</i> Value ^b)	Normalized Brightness vs. EGFP for Pure protein ^c	Difference in Brightness between Pure Protein and HeLa Cells, Fold
LSSmScarlet	1.00 ± 0.04 (NA)	1.00	0.93
mCyRFP1	1.34 ± 0.08 (<0.0001)	1.44	1.08
dCyRFP2s	1.73 ± 0.08 (<0.0001)	1.91	1.11
CyOFP1	1.48 ± 0.08 (<0.0001)	2.85	1.93
dCyOFP2s	1.70 ± 0.11 (<0.0001)	2.15	1.26
mCRISPRed	0.48 ± 0.03 (<0.0001)	1.00	2.08
CRISPRed2s	0.51 ± 0.03 (<0.0001)	0.71	1.38

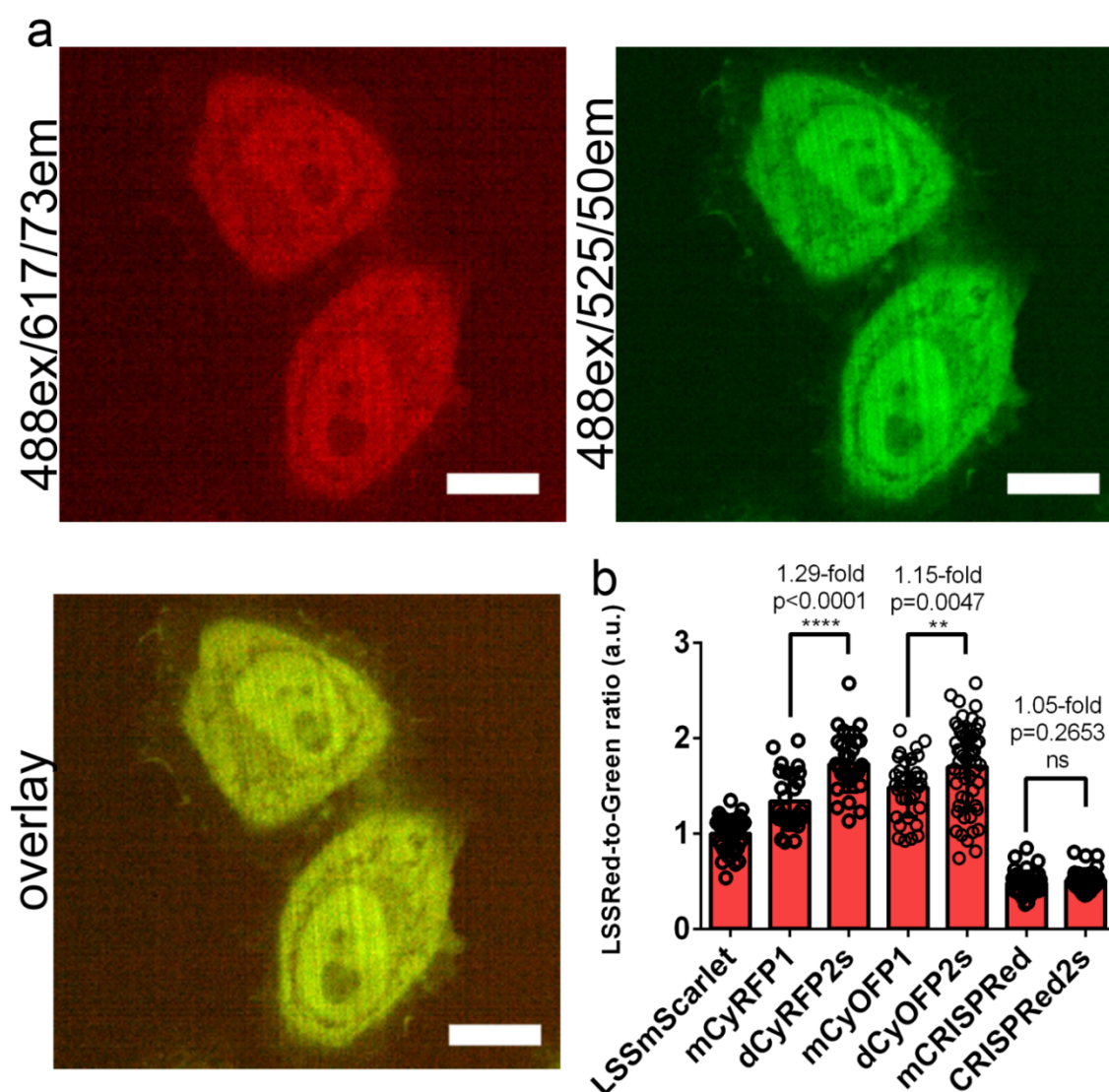


Figure 3. Comparison of the brightness of the LSSRFPs proteins in HeLa cells. (a) Confocal images of HeLa cells expressing the LSSmScarlet-P2A-EGFP protein (please, see Figure S3 for localization images for other LSSRFPs). LSSRed (488ex and 617/73em) and green (488ex and 525/50em) fluorescence channels and their overlay are shown. Scale bar, 100 μ m. (b) The averaged brightness for the LSSRFPs proteins in HeLa cells normalized to the brightness of the EGFP protein expressing in the same cells. The EGFP protein is connected to LSSRFPs via P2A-self cleavable linker. Error bars are standard deviations across twenty-eight–fifty-eight cells. *p* values show statistical difference between the respective values. Ns, not significant, *p* value is >0.05. ****, *p* value is <0.0001. **, *p* value is <0.01.

2.4. Behavior of LSSmScarlet in Mammalian Cells in Fusions with Structural Proteins

Since among the developed LSSRFPs, LSSmScarlet was the only one with monomeric behavior, we decided to test it further in mammalian cells in different protein fusions (Figure 4). LSSmScarlet demonstrated right microtubules-like localization in the fusion with β -actin, α -tubulin and vimentin cytoskeleton proteins when expressed in HeLa cells (Figure 4a). The dMito signal properly targeted LSSmScarlet into the lumen of mitochondria (Figure 4a). Fusion of LSSmScarlet with H2B histone protein localized in the nuclei of the cells and did not block cells from division (Figure 4b).

We compared the localization of LSSmScarlet with novel emiRFP703 far-red protein proposed for fusion and subcellular structural imaging [16]. In fusion with vimentin, LSSmScarlet demonstrated better localization than emiRFP703 (Figure S5), since in the case of emiRFP703, we found many cells with obvious mislocalization (Figure S5, bottom). Hence, LSSmScarlet is appropriate for the labeling of the individual proteins in mammalian cells.

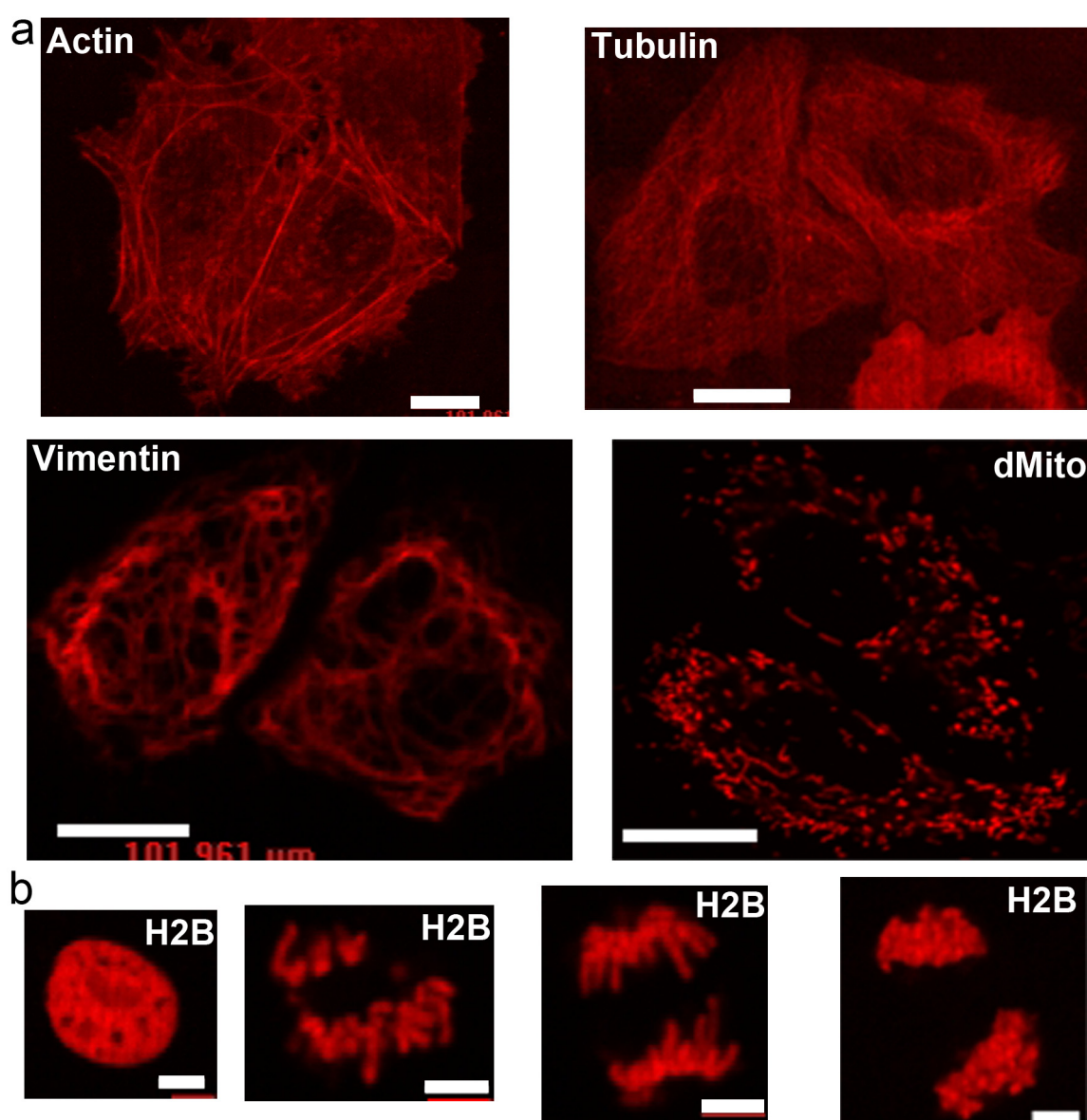


Figure 4. Localization of the LSSmScarlet protein in different fusions in mammalian cells. (a) Confocal images of HeLa cells expressing the actin-LSSmScarlet, tubulin-LSSmScarlet, vimentin-LSSmScarlet and dMito-LSSmScarlet fusions. Scale bar, 100 μ m. (b) Confocal images of non-dividing and dividing HeLa cells expressing the H2B-LSSmScarlet fusion. Scale bar, 50 μ m.

2.5. Structural Characterization of LSSmScarlet

To understand the molecular mechanism of a large Stokes shift in LSSmScarlet and the impact of mutations introduced during directed molecular evolution, we determined its crystal structure using X-ray diffraction at 1.4 Å resolution (Figure 5 and Table 3). There is one protein chain per asymmetric unit, and a contact analysis revealed that the protein is a monomer in the crystal at pH 6.5, i.e., in a highly fluorescent state (Figure 5a). LSSmScarlet has a typical β -barrel fold with the mCherry-like chromophore formed by ⁶⁷MYG⁶⁹ amino acids and positioned on the central helix of the barrel (Figure 5a,b). The chromophore also forms four direct hydrogen bonds (H-bonds) to R71, R96, D147 and E216, and several water-mediated H-bonds to Q65, Q110, T112, E145 and L200 (Figure 5c,d). The phenolic hydroxyl group of the chromophore forms two hydrogen bonds with the carboxyl group of D147 and with the buried water molecule. The negative charge of the pair D147/phenolic hydroxyl group of the chromophore is maintained by R198 that is stacked with the phenolic group of the chromophore as well as hydrogen-bonded to D147 via the solvent molecule.

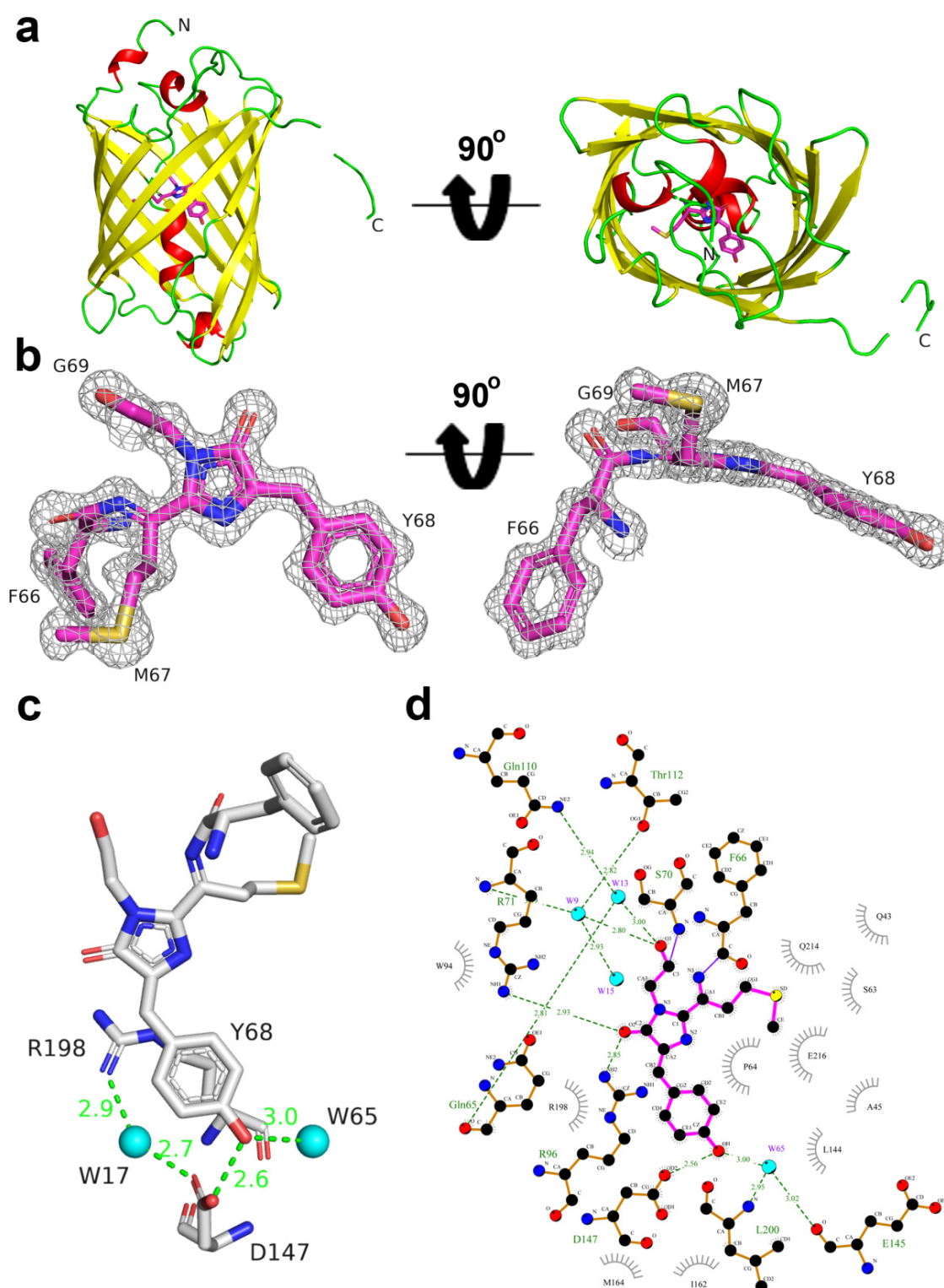


Figure 5. X-ray structure of the LSSmScarlet protein. (a) Cartoon representation of the overall LSSmScarlet structure. Chromophore, β -sheets, α -helices and disordered structures are shown in pink, yellow, red and green colors, respectively. The orientation of the panel on the right is rotated 90° around the horizontal axis with respect to that on the left. (b) The chromophore of the LSSmScarlet protein refined into the electron density map. The $2F_o - F_c$ electron density map for the LSSmScarlet chromophore structure (sticks in pink) is contoured at 1.5σ level and shown as gray mesh. The orientation of the chromophore on the right is rotated 90° around the horizontal axis with respect to that on the left. (c) Hydrogen bond network around the phenolic hydroxyl group of the LSSmScarlet chromophore. (d) The immediate environment of the LSSmScarlet chromophore.

Table 3. Data collection, processing and refinement. Values in parenthesis are for the highest-resolution shell.

Data Collection	
Diffraction Source	"Belok-RSA" Beamline, NRC "Kurchatov Institute"
Wavelength (Å)	0.75
Temperature (K)	100
Detector	CCD
Crystal-to-detector distance (mm)	120.00
Rotation range per image (°)	1.0
Total rotation range (°)	306
Space group	C2
<i>a</i> , <i>b</i> , <i>c</i> (Å)	84.52; 45.36; 59.04
α , β , γ (°)	90.0; 102.34; 90.0
Average mosaicity (°)	1.4
Unique reflections	42,242 (2137)
Resolution range (Å)	34.4–1.40 (1.42–1.40)
Completeness (%)	98.0 (100.0)
Average redundancy	6.0 (6.2)
$\langle I/\sigma(I) \rangle$	7.2 (1.7)
R _{meas} (%)	15.0 (67.2)
CC _{1/2}	99.1 (81.3)
Refinement	
<i>R</i> _{fact} (%)	16.4
<i>R</i> _{free} (%)	19.0
Bonds (Å)	0.02
Angles (°)	2.16
Ramachandran plot	
Most favored (%)	98.6
Allowed (%)	1.4
No. atoms	
Protein	1850
Water	272
Chromophore	23
Sodium ion	2
Other ligands	5
<i>B</i> -factors (Å ²)	
Protein	10.77
Water	21.34
Chromophore	7.06
Calcium ion	10.02
Other ligands	12.32

To elucidate the mechanism of the excited state proton transfer (ESPT) and a large Stokes shift in LSSmScarlet, we analyzed contacts of the phenolic hydroxyl group of the chromophore with its environment. The structure revealed that the phenolic hydroxyl group of the chromophore is engaged in direct contact with the carboxyl group of D147 with the donor-acceptor distance of 2.56 Å. Since the excitation maximum is blue-shifted by 128 nm relative to the emission maximum in LSSmScarlet compared to only a 25 nm difference between excitation and emission maxima in its predecessor mScarlet, the phenolic hydroxyl group is likely to exist in the protonated form in the ground state. Upon

excitation of the protonated chromophore in LSSmScarlet, excited-state proton transfer (ESPT) occurs, wherein the phenolic group of the neutral protonated chromophore serves as a donor of a proton to a negatively charged D147 side chain, and the formed anionic chromophore fluoresces (Figure 6a). Thus, we can explain the observed large Stokes shift in LSSmScarlet by ESPT from the chromophore to the carboxylic group of the D147 side chain (as proved below by molecular simulations). Large Stokes shifts in other LSSFs were also reported to be explained by ESPT. The resolved crystal structures for CyOFP1 (PDB:5BQL) [3] and mCRISPRed (PDB: 6XWY) [17] demonstrated that ESPT from the phenolic hydroxyl group of the chromophore to the K160 amino group in CyOFP1 (corresponding to M164 in LSSmScarlet) and the D142 carboxyl group in mCRISPRed (corresponding to D147 in LSSmScarlet) underlie large Stokes shifts in these FPs.

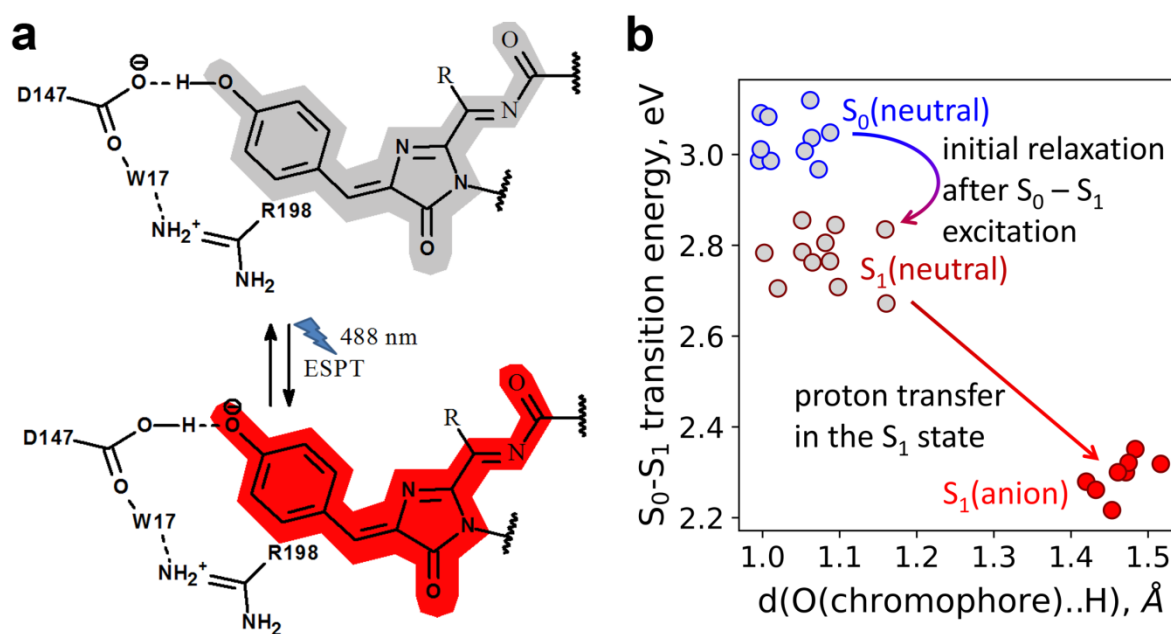


Figure 6. Suggested ESPT mechanism for LSSmScarlet. (a) The proton from the phenolic hydroxyl group of the chromophore migrates to the D147 residue after the excitation of the chromophore with 488 nm light. Upon relaxation into ground state from excited state proton comes from D147 residue back to the phenolic hydroxyl group of the chromophore. D147 residue forms water-mediated hydrogen bond with R198 residue (please, see also Figures 5c,d and 7b). (b) The ESPT mechanism obtained in the QM/MM MD simulations. Circles correspond to the selected frames from trajectories: grey circles are for the neutral state (with blue and maroon edges for the S_0 and S_1 states, respectively), and red circles are for the anionic S_1 state. The vertical axis depicts the S_0 – S_1 transition energies calculated at the QM(CAM-B3LYP/def2-SVP)/MM(CHARMM) level.

To understand the impact of mutations introduced during LSSmScarlet development, we compared chromophore planarity and the chromophore's environment for LSSmScarlet and the parental protein mScarlet (PDB: 5LK4) [12] (Figure 7). The relative positions of imidazolinone and phenolic rings in the LSSmScarlet and mScarlet chromophores are different. Two rings of the chromophore are almost perfectly in the plane in mScarlet, but phenolic ring is out of the plane in LSSmScarlet with corresponding twist and tilt angles of about 5 and 20° (according to Figure 6 in [4]). This difference in the chromophore planarity results in the higher quantum yield exhibited by mScarlet with a more planar conformation of the chromophore (QY for mScarlet is 0.7 versus 0.43 for LSSmScarlet); similarly, a decrease in quantum yield for mFruits proteins was associated with chromophore noncoplanarity [18]. Less planar conformation of the chromophore in LSSmScarlet could be a result of key mutation S147D necessary for the large Stokes shift. This mutation also led to a strong shift of the M164 side chain (RMSD between corresponding S atoms of mScarlet and LSSmScarlet is 3.64 Å) due to steric hindrance.

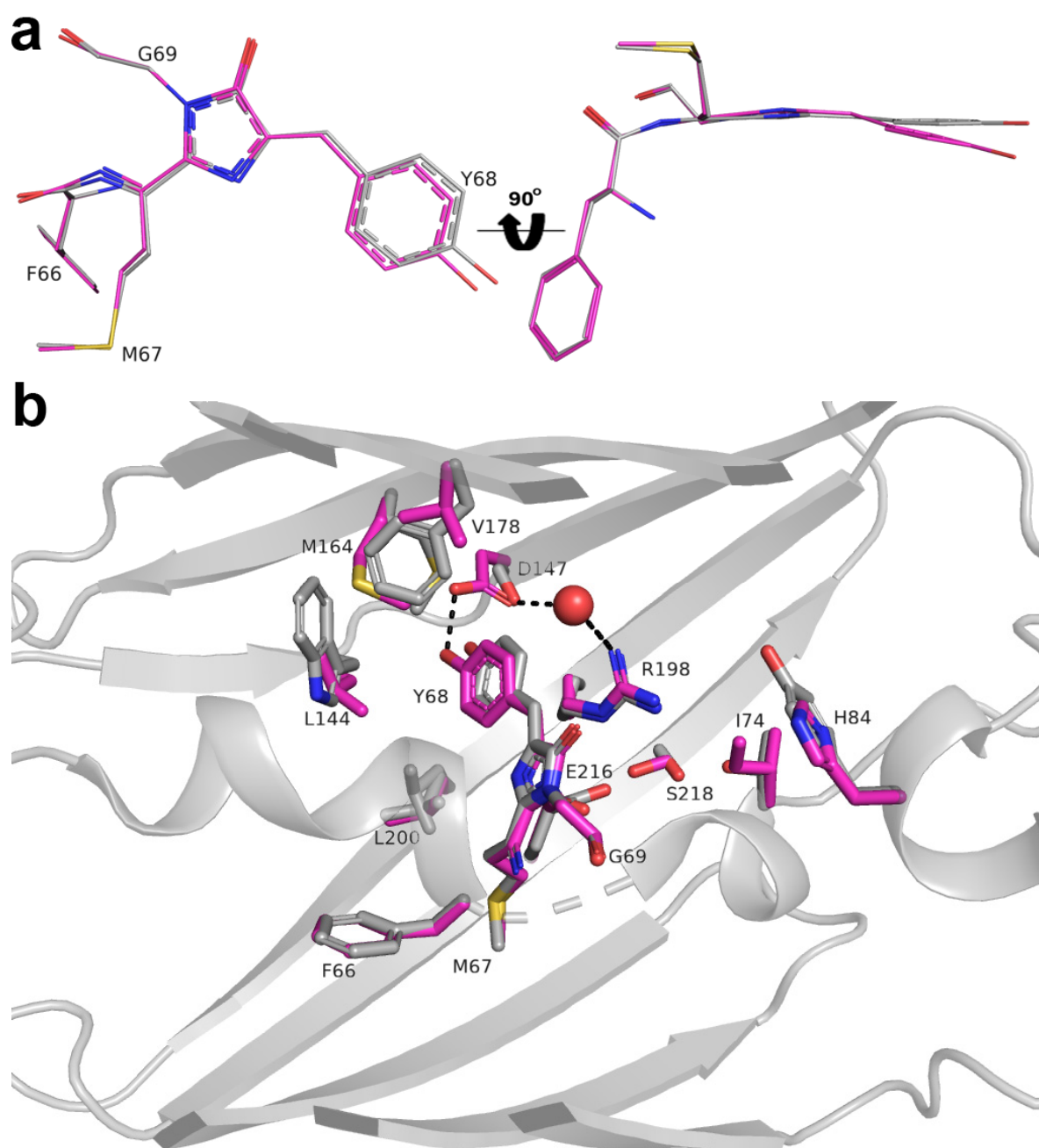


Figure 7. Structural comparison of the chromophores (a) and their immediate environments (b) for the LSSmScarlet and mScarlet (PDB: 5LK4) proteins. Water molecule (red sphere) provides hydrogen bonding (dashed lines) between D147 and R198 residues of the LSSmScarlet. D147 of LSSmScarlet makes a direct hydrogen bond (dashed line) with Y68 residue. Residues' enumeration is shown for LSSmScarlet protein. In panel (a), the orientation of the chromophore on the right is rotated 90° around the horizontal axis with respect to that on the left.

There are five mutations in LSSmScarlet compared to mScarlet that are inner to the β -barrel: S147D, W144L, F178V, T74I and Y84H. Three mutations, S147D, W144L and F178V were located at distances in the range of 6 Å away from the chromophore. Because of the strong shift of M164 caused by the S147D mutation, there is a steric clash. The new position of Met164 in LSSmScarlet required the replacement of adjacent bulky aromatic amino acid residues W144 and F178 with smaller ones—Leu and Val, respectively. The former mutation also seems to stabilize the L200 side chain (one conformation in LSSmScarlet compared to two in mScarlet), which forms van der Waals contacts with the chromophore. Mutations, T74I and Y84H, are more than 6.0 Å away from the chromophore; however, T74 substitution breaks the hydrogen bond between T74 and S218, which

in turn forms a hydrogen bond with E216. This leads to a reorientation of the S218 side chain towards E216, strengthening the hydrogen bond S218–E216 and a shift of the glutamic acid side chain towards L200 (compared to mScarlet).

2.6. Molecular Dynamics Simulations of the ESPT Mechanism in LSSmScarlet.

We performed QM/MM MD simulations based on the available X-ray structure of the LSSmScarlet protein to elucidate the molecular mechanism of the proton transfer occurring in the excited electronic state, S_1 (Figure 6b). The 20 ps ground state (S_0) QM/MM MD simulations demonstrated that the chromophore exists in the neutral state, and the D147 is negatively charged. First, we performed QM/MM MD on the first excited electronic state, S_1 , starting from the representative frame of the trajectory on the ground state, S_0 . After vertical excitation, the initial relaxation of the conjugated π -system of the neutral chromophore occurs (Figure 6b). It resulted in a small decrease (about 0.2 eV) of the energy gap between the ground and excited states. Next, the proton is transferred from the phenyl fragment of the chromophore to the carboxylate of D147, and this proton transfer results in the formation of the fluorescent state and a considerable decrease in the energy gap between the S_0 and S_1 surfaces. According to our model, the Stokes shift is ~0.7 eV that agrees with the experimental 0.57 eV and further proves the suggested ESPT mechanism. Hence, QM/MM MD simulations based on X-ray data for the LSSmScarlet protein proved the ESPT that is responsible for the large Stokes shift red fluorescence in this protein.

2.7. The Role of R198 Residue in ESPT of the LSSmScarlet Protein.

To elucidate the role of R198 residue in ESPT, we generated five mutants of LSSmScarlet containing R198E, R198H, R198I, R198K and R198Y substitutions and characterized their spectral properties and pH stabilities. LSSmScarlet with R198E and R198H substitutions preserved efficient ESPT of the chromophore at pH 7.40 according to the single absorption peaks at 480–487, corresponding to the protonated chromophore (Table 4 and Figure 8). The R198I mutation resulted in the formation of a deprotonated chromophore at pH 7.40 according to the appearance of the absorption peak at 571 nm (Table 4 and Figure 8). R198K and R198Y mutants did not efficiently form the red chromophore according to the major absorption peaks at 405 and 372 nm, respectively (Table 4 and Figure 8), so we did not characterize them further. LSSRed fluorescence of R198H, R198E and R198I LSSmScarlet mutants showed dependence on pH with pKa values of 2.403, 6.181 and 7.76, respectively (Table 4 and Figure 8). We assumed that H198 and E198 residues served as proton donor/acceptors via the W17 molecule and stabilized ESPT from D147 residue to the chromophore (Figures 5–7), but I198 did not. Positively charged H198 residue stabilized ESPT to a higher extent as compared to negatively charged E198 probably because of favorable and unfavorable electrostatic interactions with negatively charged D147 residue, respectively. According to X-ray structures' superposition, mCRISPRed LSSRed protein with high pH stability also has positively charged H196 residue in a position corresponding to position 198 in LSSmScarlet (Figure S6). However, H196 residue of mCRISPRed did not form the H-bond with D142 residue, which makes a direct H-bond to the chromophore of mCRISPRed (analogous to D147 residue in LSSmScarlet). Probably, an electrostatic interaction between positively charged H196 (R198) and negatively charged D142 (D147) is important for the high pH-stability of LSSRed fluorescence for both proteins. Hence, the residue in position 198 of LSSmScarlet may participate in ESPT to the chromophore as a proton donor/acceptor, and a positive charge in position 198 facilitates ESPT at an acidic pH values.

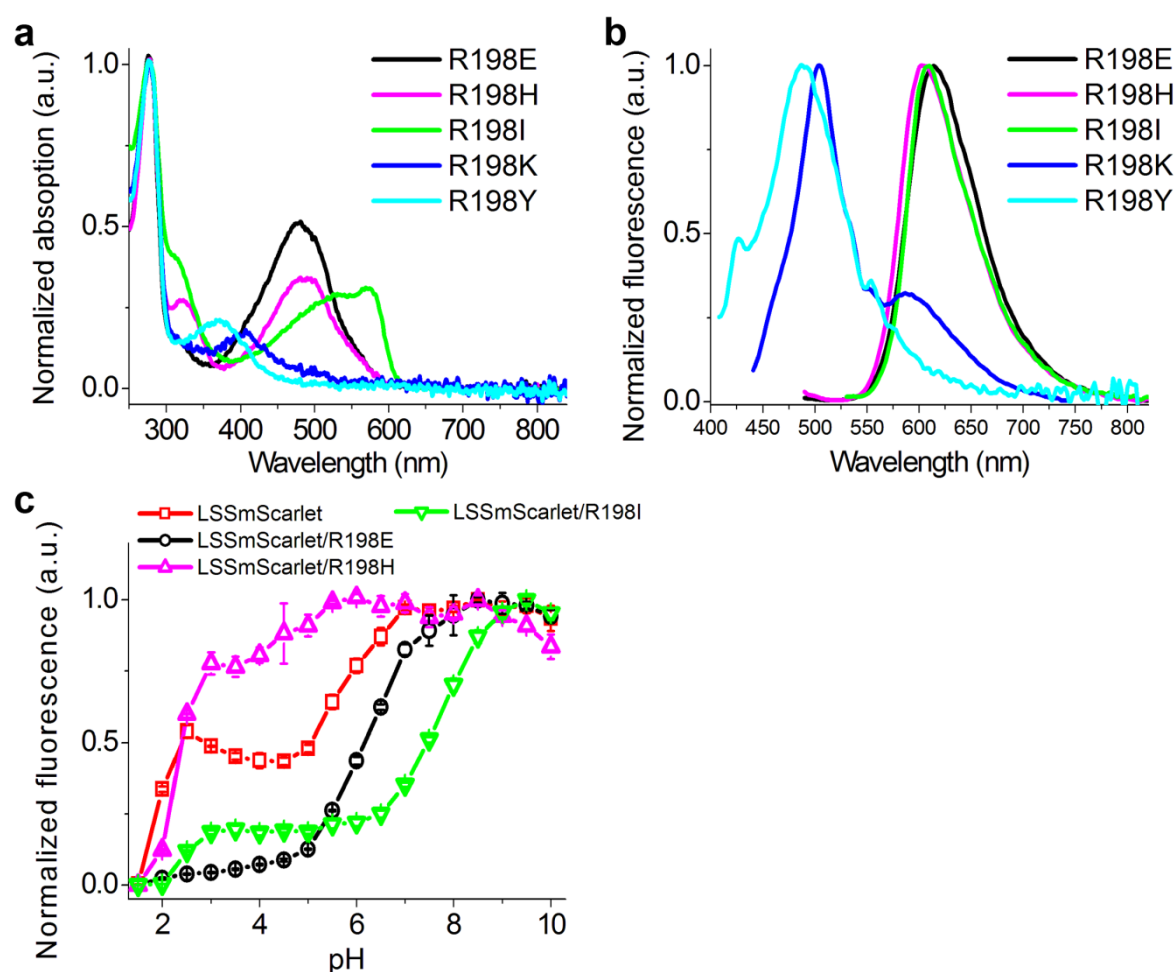


Figure 8. Spectral properties and pH stability for the purified LSSmScarlet mutants with substitutions at position 198. **(a)** Absorption spectra for LSSmScarlet mutants in PBS buffer at pH 7.40. **(b)** Emission spectra for LSSmScarlet mutants in PBS buffer at pH 7.40, excited at absorption maximum. **(c)** pH stability of LSSmScarlet and its mutants.

Table 4. In vitro properties of the LSSmScarlet and its mutants at 198 position. Minor peaks are shown in parentheses. Standard deviations are shown. NA, not applicable.

Proteins	Abs/Em (nm)	pK _a
LSSmScarlet	466/598	1.91 ± 0.01; 5.78 ± 0.06
LSSmScarlet/R198E	480/614	6.181 ± 0.006
LSSmScarlet/R198H	487/604	2.403 ± 0.008
LSSmScarlet/R198I	571 (472)/610	7.76 ± 0.03
LSSmScarlet/R198K	405/503	NA
LSSmScarlet/R198Y	372/488	NA

2.8. Two- and Three-Color Confocal Imaging with the LSSmScarlet Protein.

We next validated the utility of LSSmScarlet for multi-color confocal imaging in combination with conventional GFPs and RFPs at 488 and 561 nm excitations. First, we verified the bleed-through of LSSmScarlet fluorescence into the standard green and red channels under a confocal microscope. We expressed H2B-LSSmScarlet fusion protein in HeLa cells and imaged it in green, LSSRed and red channels (Figure 9a). We did not see a notable bleed-through of the LSSmScarlet fluorescence into both the green and red channels (Figure 9a). No bleed into the green channel was expected according to the good separation

of the LSSmScarlet emission spectra and the green emission filter (Figure 2b). We also anticipated that LSSmScarlet will not bleed into the red channel, because according to its excitation spectrum, it is excited with 561 nm light 27-fold less efficiently compared to 488 nm light (Figure 2b). Moreover, the absence of the bleed into the green channel was estimated and confirmed for other LSSRFP CyOF1 [3]. Hence, during confocal imaging, the LSSmScarlet protein does not bleed-through into the green and red channels when excited at 488 and 561 nm, respectively, and it should be applicable for combined confocal imaging together with GFP- and RFP-like proteins.

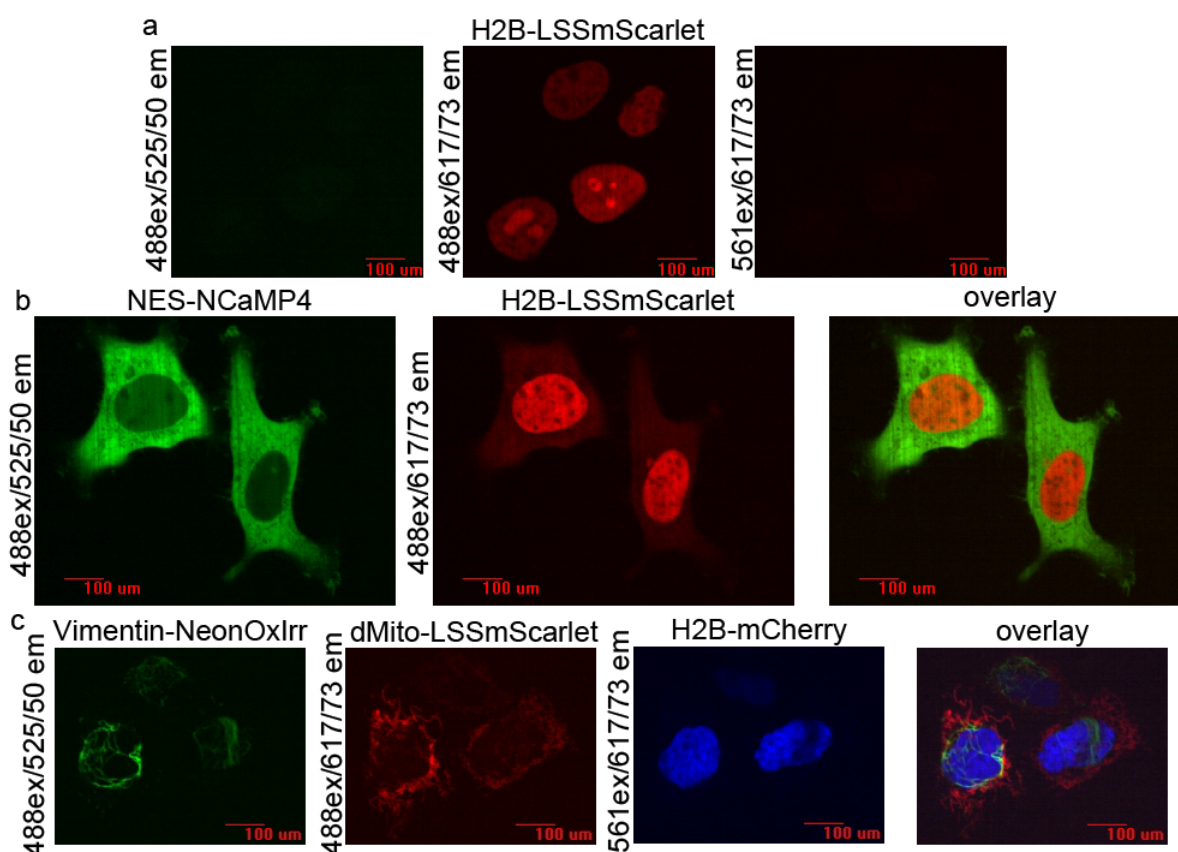


Figure 9. Two- and three-color confocal imaging with LSSmScarlet protein. (a) Confocal images of the HeLa cells expressing H2B-LSSmScarlet only in green (488ex/525/50em in green), LSSRed (488ex/617/73em in red) and red (561ex/617/73em in red) channels. We used the same exposure time and brightness/contrast adjustment levels for all channels. We do not see notable bleed-through of the LSSRFP fluorescence into both the green and red channels. (b) Two-color confocal imaging of the HeLa cells co-expressing NES (nuclear export signal)-NCaMP4 green calcium indicator and H2B-LSSmScarlet fusion LSSRFP in green (488ex/525/50em in green), LSSRed (488ex/617/73em in red) and overlaid channels using single-excitation with 488nm laser. (c) Three-color confocal imaging of the HeLa cells co-expressing the Vimentin-NeonOxIrr fusion green indicator for hydrogen peroxide, dMito-LSSmScarlet fusion LSSRFP and H2B-mCherry fusion RFP in green (488ex/525/50em in green), LSSRed (488ex/617/73em in red), red (561ex/617/73em in blue) and overlaid channels. Scale bar, 100 μm.

To demonstrate single excitation two-color emission confocal imaging with the LSSmScarlet protein and mNeonGreen-derived novel family of fluorescent calcium indicators [19], we imaged mammalian cells co-expressing the H2B-LSSmScarlet fusion and NES (nuclear export signal)-NCaMP4 green-yellow calcium indicator using a single 488 nm excitation light and standard emission filters. We transiently co-expressed the H2B-LSSmScarlet protein and NES-NCaMP4 green-yellow calcium indicator in the nucleus and cytosol of the HeLa cells, respectively. Confocal imaging with a 488 nm excitation demonstrated that the green-yellow NES-NCaMP4 indicator and LSSRed H2B-LSSmScarlet protein had anticipated localization (Figure 9b). The NES-NCaMP4 revealed a slight bleed-through into the LSSRed channel, which was not seen on the overlaid image (Figure

9b). For the green EGFP protein, the bleed-through into the LSSRed channel should be similar to that for green-yellow NCaMP4 since EGFP has a similar emission spectrum integral in the range of the 617/73 nm red emission filter as compared to the integral for the NCaMP4 calcium indicator (Figure S7). Hence, the LSSmScarlet protein is appropriate for single excitation two-color confocal imaging with GFP- and mNeonGreen-derived fluorescent proteins.

Since the LSSmScarlet protein does not bleed through into the red channel (please, see above), we next attempted to demonstrate three-color confocal microscopy with the LSSmScarlet LSSRFP, mCherry RFP and mNeonGreen-derived green-yellow NeonOxIrr indicator for hydrogen peroxide [20] using a 488/561 nm dual excitation and standard emission filters. We transiently co-expressed dMito-LSSmScarlet, H2B-mCherry and Vimentin-NeonOxIrr fusion proteins in the lumen of mitochondria, in the nucleus and in the cytoskeleton of the HeLa cells. Confocal imaging with a 488 and 561 nm excitation and dual-emission revealed an anticipated localization of the dMito-LSSmScarlet, H2B-mCherry and Vimentin-NeonOxIrr fusion proteins in HeLa cells (Figure 9c). We did not notice any bleed-through between all three channels. Hence, LSSmScarlet protein can be used for three-color confocal imaging with GFP-, mNeonGreen- and mCherry-derived fluorescent proteins.

2.9. Two-Photon Imaging of the LSSmScarlet Protein.

To assess the applicability of the LSSmScarlet protein for two-photon (2P) fluorescence microscopy, we acquired two-photon images of the HeLa cells expressing H2B-LSSmScarlet and studied the dependence of the 2P excitation of the H2B-LSSmScarlet vs. 2P excitation wavelength. Using 11% of the laser power (of the standard Ti-Sapphire laser family) of the 2P microscope at a 940 nm excitation wavelength, we successfully imaged H2B-LSSmScarlet expressed in HeLa cells (Figure 10a). The maximal fluorescence intensity of the H2B-LSSmScarlet was observed at a 900 nm two-photon excitation wavelength (Figure 10b). The maximum of the H2B-LSSmScarlet fluorescence intensity normalized to the laser power was found at 940 nm (Figure 10b). This maximum wavelength coincided with a doubled one-photon excitation maximum wavelength for the purified LSSmScarlet protein observed at 470 nm (Figure 10b and Table 1). Hence, LSSmScarlet is appropriate for two-photon fluorescence microscopy with a standard Ti-Sapphire laser family, and a 940 nm wavelength is optimal for its 2P excitation.

We next performed 2P imaging of the NES-NCaMP4 green-yellow mNeonGreen-based calcium indicator and tested the dependence of its fluorescence on the 2P-excitation wavelength. Using 10% of the laser power at 940 nm, we visualized the NES-NCaMP4 indicator transiently expressed in the cytosol of the HeLa cells (Figure 10c). The maximal fluorescence for the NES-NCaMP4 indicator was observed at 900 nm (Figure 10d). The maximum of the NES-NCaMP4 fluorescence normalized to the laser power was seen at 940 nm (Figure 10d). This maximum was the same for both NCaMP4 and LSSmScarlet proteins. Hence, the NES-NCaMP4 indicator is appropriate for 2P imaging, and 940 nm is the optimal wavelength for its 2P excitation.

To demonstrate single excitation dual-emission 2P microscopy with LSSmScarlet and the mNeonGreen-based reporter, we imaged H2B-LSSmScarlet protein and the NES-NCaMP4 indicator co-expressed in the nucleus and cytosol of the HeLa cells using 2P microscopy with a single excitation at 940 nm and standard emission filters. Using 10% laser power at a 940 nm excitation, we successfully acquired 2P images of the H2B-LSSmScarlet and NES-NCaMP4 localized in the nucleus and cytosol of the HeLa cells, respectively (Figure 10e). We noted bleed-through of the NCaMP4 fluorescence into the LSSRed channel, but this bleed was weak, and these two reporters were clearly distinguished in the superimposed image (Figure 10e). Simultaneous dual-emission 2P imaging was also earlier demonstrated for another LSSRFP CyOFF1 protein together with GFP protein and the G-CaMP6s calcium indicator in live cells in vitro and in vivo [3]. Hence,

the LSSmScarlet protein is appropriate for single excitation dual-emission 2P microscopy with GFP- and mNeonGreen-based fluorescent reporters.

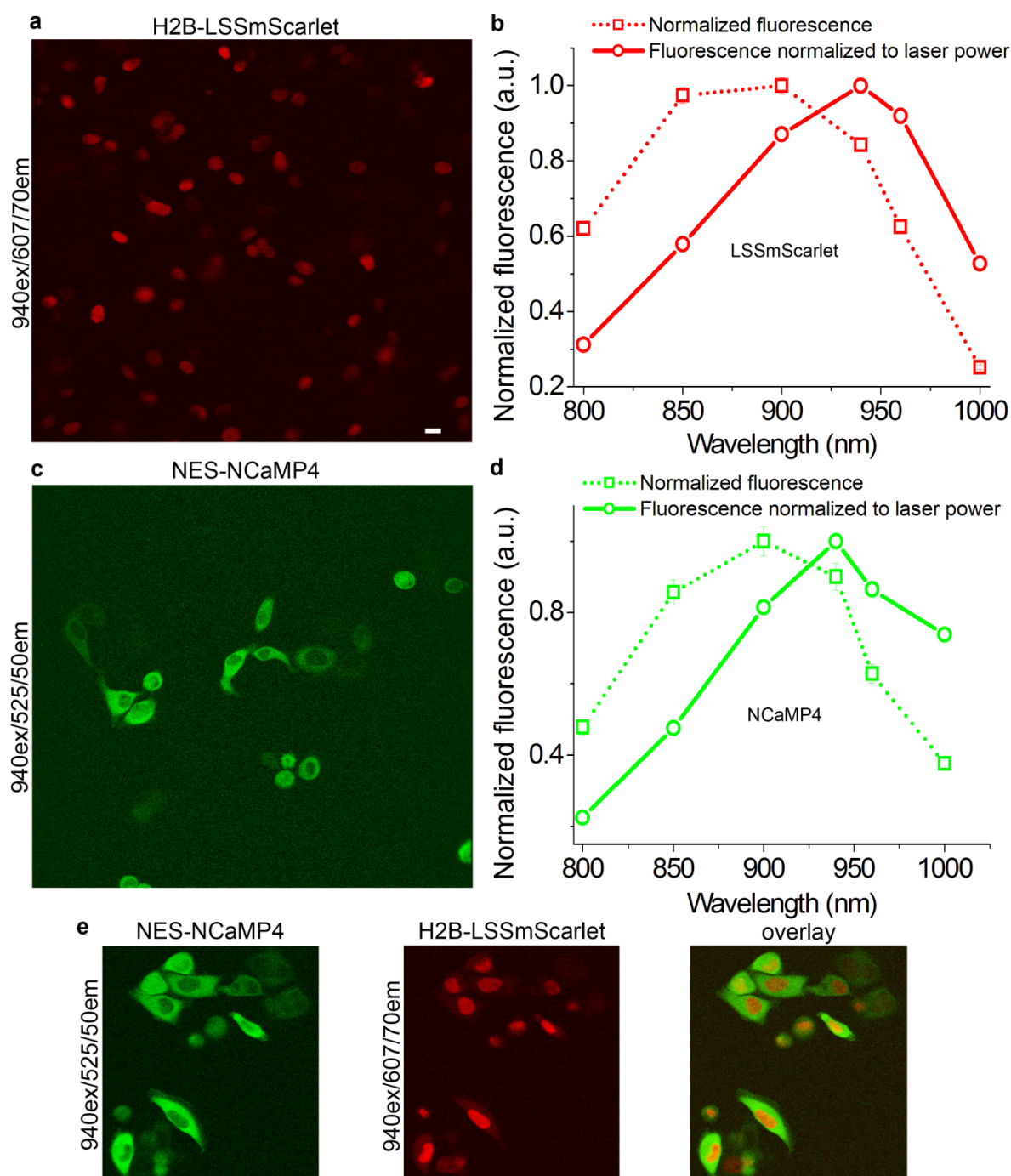


Figure 10. Simultaneous dual-emission two-photon imaging with LSSmScarlet and mNeonGreen-based calcium indicator expressed in HeLa cells. (a) Two-photon fluorescence image of the H2B-LSSmScarlet expressed in HeLa cells (940 nm excitation, 171 mW laser power; 607/70 emission filter). (b) Dependence of the H2B-LSSmScarlet fluorescence on two-photon excitation wavelength. (c) Two-photon fluorescence image of the NES-NCaMP4 green-yellow calcium indicator expressed in HeLa cells (940 nm excitation, 171 mW laser power; 525/50 emission filter). (d) Dependence of the NES-NCaMP4 fluorescence on two-photon excitation wavelength. (e) Two-photon fluorescence images of the H2B-LSSmScarlet and NES-NCaMP4 co-expressed in HeLa cells (940 nm excitation, 171 mW laser power). NES-NCaMP4 and H2B-LSSmScarlet were spectrally resolved using 525/50 and 607/70 emission filters, respectively. (a,c,e) Scale bars are the same for all images, 20 μ m.

2.10. LSSRed-to-Red Photoconversion of the LSSmScarlet Protein in the Conditions of the Super-Resolution STED Imaging

Finally, we assessed whether LSSmScarlet protein is suitable for super-resolution stimulated emission depletion (STED) microscopy similar to mKeima LSSRFP [21].

Since STED imaging is often performed on fixed samples, we studied the impact of fixation on the fluorescence of the LSSmScarlet protein expressed in mammalian cells. Using the confocal microscope, we imaged HeLa cells expressing the H2B-LSSmScarlet fusion protein before and after fixation with 4% polyformaldehyde (PFA) for 15 min at room temperature. We observed minor quenching of the fluorescence upon fixation, which was not statistically significant ($p = 0.5212$; Figure S8). Resistance of the LSSmScarlet protein to the fixation correlated with its high pH-stability.

Using fixed HeLa cells containing vimentin-LSSmScarlet protein, we could not obtain notable improvement in the resolution of the vimentin fibers using a 488 nm excitation light and both 775 nm (Figure 11a) and 595 nm (Figure 11b) depletion lasers. Depletion by the 775 nm laser did not affect image resolution in a wide range of excitation and depletion laser powers, and we did not succeed in resolution improvement. The 595 nm was probably an inappropriate wavelength for the LSSmScarlet depletion according to its emission spectrum (Figure 2b). We found that the 775 nm depletion light bleached the LSSRed fluorescence of the LSSmScarlet protein excited with 488 nm light, and new red fluorescence at 561 nm excitation appeared (Figure 11c,d).

To study photochemical transformations for the LSSmScarlet protein, we registered spectral changes accompanied by irradiation of the LSSmScarlet purified protein with a two-photon 775 nm laser, and one-photon 405 and 468 nm LED arrays. We illuminated purified LSSmScarlet protein with a 775 nm 2P laser (4010 mW power) through a 25x objective lens and found the formation of the traces of the red form with excitation/emission maxima at 570/598 nm, respectively (Figure S9a). We suggest that the same red-form is formed in the conditions of the STED imaging (Figure 11c,d). The one-photon illumination of the LSSmScarlet protein with a 405 nm LED array resulted in the formation of the green-form with an absorption peak at 412 nm and excitation/emission maxima at 426/503 nm, respectively (Figure 10b); the traces of the red-form were also registered according to the excitation/emission peaks at 550/592 nm, respectively (Figure 10b). The one-photon irradiation of the LSSmScarlet protein with the 468 nm LED array resulted in the formation of the cyan-form with an absorption peak at 401 nm and excitation/emission maxima at 422/480 nm, respectively (Figure 10c); compared to 405 nm light, we observed more traces of the red-form with an absorption at 570 nm and excitation/emission maxima at 554/598 nm, respectively (Figure 10c). LSSRed-to-Red photoswitching of the PSLSSmKate protein with a one-photon 405 nm light illumination was earlier attributed to decarboxylation of the glutamate 215 residue via the Kolbe-type reaction mechanism [22]. As compared to PSLSSmKate, the irradiation of the LSSmScarlet protein with 405 and 468 nm light resulted in the formation of the different spectral species. Hence, the LSSmScarlet protein undergoes various photochemical transformations with the formation of the green, cyan and red fluorescent products depending on the light power and the wavelength of light.

Finally, using live HeLa cells expressing vimentin-LSSmScarlet fusion, but not the fixed ones, we could perform STED imaging of the Red-photoconverted form of the LSSmScarlet protein with a 561 nm excitation and 775 nm depletion (Figure 11e). As compared to confocal imaging, using STED microscopy, we could improve resolution of vimentin fibers by 2.4-fold (Figure 11e). Overall, the LSSmScarlet protein is not a protein of choice for STED microscopy as a consequence of the photoconversion with a high-power 775 nm depletion laser.

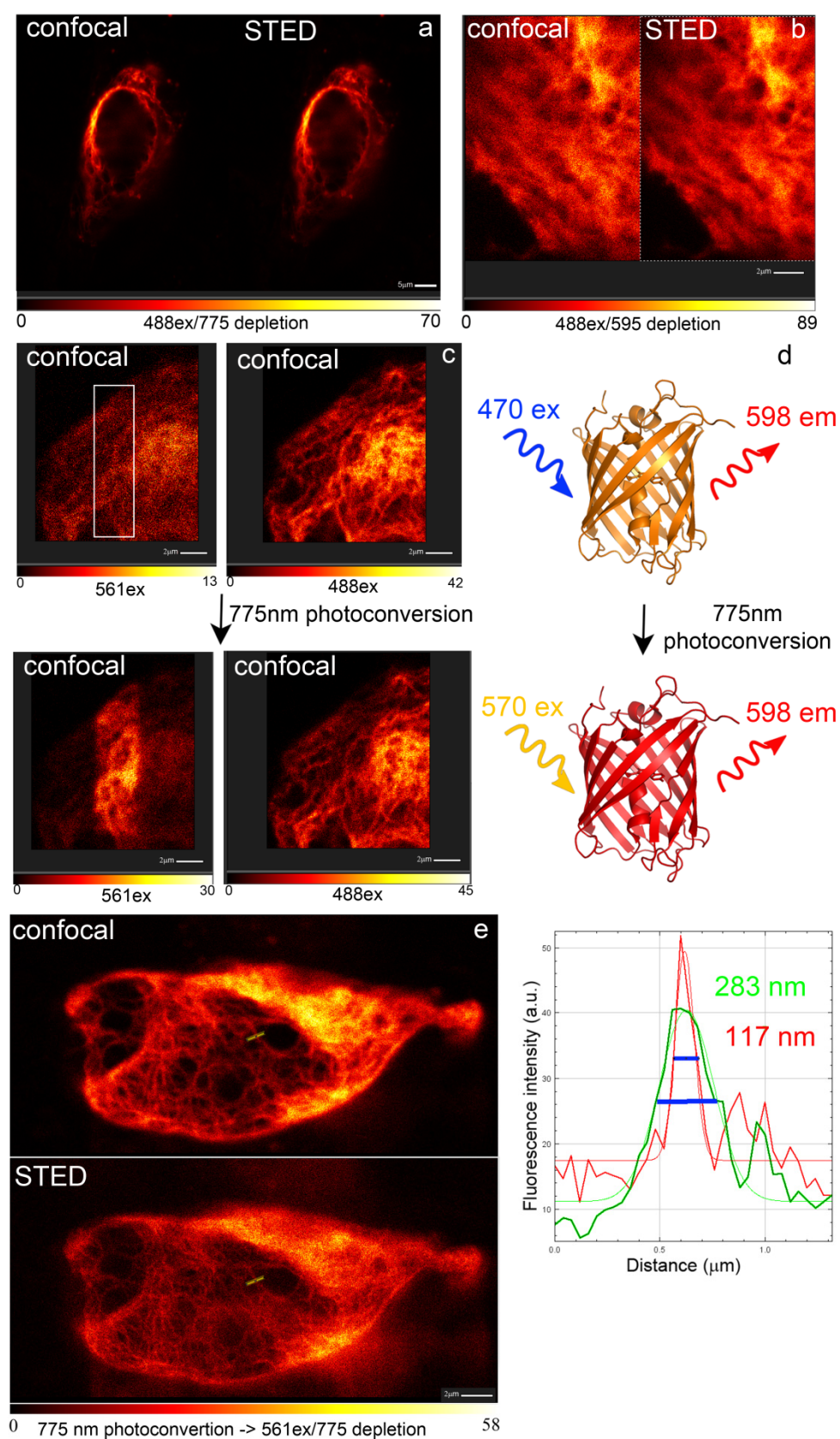


Figure 11. STED imaging of vimentin-LSSmScarlet fusion in HeLa cells. (a) Confocal and STED images of fixed cells expressing vimentin-LSSmScarlet. STED settings: 488 nm excitation—3.17% laser power; 775 nm STED—2% laser power; STED gated on delay 750 ps, width 8 ns; emission range of 560–725 nm; pixel dwell time—4 μ s in 9 repetitions. (b) Confocal and STED images of fixed cells expressing vimentin-LSSmScarlet. STED settings: 488 nm excitation—5% laser power; 595 nm

STED—100% laser power; on delay 750 ps, width 8 ns; emission range of 560–585 nm; dwell time—12 μ s; line steps—4 repetitions, 16 total lines. (c) Confocal images of fixed cells expressing vimentin-LSSmScarlet before and after photoconversion with 775 nm depletion laser (50% power) in the region shown as a white box. (d) Schematic representation of excitation maxima of the LSSmScarlet protein as a consequence of the photoconversion with 775 nm light. (e) Confocal and STED images of live cells expressing vimentin-LSSmScarlet. The line profile for the marked area shows filament with 2.4-fold improvement of resolution. Line profile was fitted with $y = a + (b - a) \exp(-(x - c)/(2 \cdot d \cdot d))$ equation. STED settings: 561 nm excitation—10.57% laser power; 775 nm STED—15% laser power; on delay 750 ps, width 8 ns; emission range of 570–765 nm; dwelltime—105 μ s; line steps—10 repetitions, 40 total lines.

3. Materials and Methods

3.1. Cloning of Bacterial Vectors, Mutagenesis and Library Screening

LSSRFP proteins were cloned into the pBAD/HisB plasmid (Invitrogen) at BglIII/EcoRI restriction sites using the mSc-BglIII/mCherry-EcoRI-r (LSSmScarlet), Fw-LSSmOrange-BglIII/mCherry-EcoRI-r (CyOFP1), Fw-LSSmOrange-BglIII/Rv-LSSmOrange-EcoRI (mCyRFP1) and Fw-BglIII-(PA)TagRFP/Rv-LSSmOrange-EcoRI (mCRISPRed) primers listed in Table S3 to express LSSRFPs proteins in BW25113 bacterial cells (kindly provided by Verkhusha V.V. from Albert Einstein College of Medicine, New York, NY, USA).

Three site-directed libraries for rational mutagenesis of the parental mScarlet protein at three positions were generated using mSc-148DE, mSc-148DE-r, mSc-165,7-X, mSc-165,7-X-r, mSc-148X, mSc-148X-r, mSc-165,7-DX, mSc-165,7-DX-r, mSc-165,7-XD and mSc-165,7-XD-r primers listed in Table S1. Assembly of the whole gene was performed using PCR with overlapping fragments [23]. Generated libraries were inserted at BglIII/EcoRI restriction sites of the pBAD/HisB-TorA-HyPer-sfGFP plasmid by swapping the HyPer gene.

Random libraries of LSSRFPs were obtained using PCR in the presence of Mn^{2+} ions (according to the Diversify PCR Random Mutagenesis Kit User Manual, Clontech, 2–3 random mutations were introduced per 1000 base pairs) and cloned at BglIII/EcoRI restriction sites of the pBAD/HisB plasmid.

Screening of bacterial libraries was performed on Petri dishes under a fluorescent microscope. Briefly, expression of the LSSRFPs on the colonies on Petri dishes was induced with 0.02% arabinose for 24 h at 37 °C. Screening of about 10,000 colonies of the bacterial library expressing LSSRFPs variants was performed on Petri dishes under fluorescent stereomicroscope Leica M205FA (Leica, Germany) equipped with the DFC310FX camera (Leica Microsystems, Germany) and mercury metal halide light source EL6000 (Leica Microsystems, Germany). LSS-red fluorescence was registered by 480/40BP excitation (62 μ W/cm² on the sample) and 620/60BP emission filters. Acquired images were analyzed using ImageJ software, and colonies having the highest brightness were picked up for further analysis on bacterial streaks on Petri dishes.

3.2. Proteins' Purification and Characterization

Proteins were expressed and purified using the pBAD/HisB arabinose-inducible system (Invitrogen) as described in reference [7]. Briefly, BW25113 bacterial cells were transformed with pBAD/HisB-LSSRFPs plasmid. For protein expression, the bacterial cultures were grown in LB medium supplemented with 0.004% arabinose and 100 μ g/mL ampicillin overnight at 37°C and 220 rpm. The cultures were centrifuged at 4648 \times g for 10 min. The cell pellets were resuspended in PBS buffer, pH 7.4, supplemented with 300 mM NaCl and 10 mM Imidazole and lysed by sonication on ice. The sonicated solution was centrifuged at 36,670 \times g at 4 °C. The proteins were further purified using Ni-NTA resin (Qiagen) followed by dialysis for 16 h against PBS buffer.

The extinction coefficients' values for the purified LSSRFPs proteins were calculated in PBS buffer, pH 7.4, using the alkaline denaturation method and assuming that the GFP-

like chromophore has the extinction coefficient of $44,000 \text{ M}^{-1} \text{ cm}^{-1}$ at 455 nm in 1M NaOH [24]. Absorption spectra were recorded using the NanoDrop 2000c Spectrophotometer (Thermo Scientific, DE, USA).

The quantum yields for purified LSSRFPs proteins excited at 460 nm were measured by a comparison of the integrated fluorescence values (in the range of 470–800 nm) in PBS buffer, pH 7.40, with the similarly integrated fluorescence values for the equally absorbing at 460 nm CyOFP1 protein (quantum yield of 0.76 [3]). Fluorescence spectra were acquired using a CM2203 spectrofluorometer (Solar, Minsk, Belarus).

pH titrations for purified LSSRFPs proteins (32–202 nM final concentration) were performed in buffers of 30 mM citric acid, 30 mM borax, 30 mM NaCl with a pH ranging from 1.5 to 10.0, using incubation for 20 min at room temperature. LSS-red fluorescence (Ex490nm/Em580–640nm) was registered using 96 well Modulus™ II Microplate Reader (Turner Biosystems, Sunnyvale, CA, USA).

Photobleaching of LSSRFPs was performed using suspensions of purified proteins in mineral oil. Briefly, the photobleaching was measured using a suspension of 1.5 μL of purified proteins dialyzed in PBS buffer, at a 45 μM concentration, in 10 μL of mineral oil placed on the cover glass. Imaging of the suspension was performed using the Zeiss Axio Imager Z2 microscope (Zeiss, Germany) equipped with a 120 W mercury short-arc lamp (LEJ, Germany), a 63×1.4 NA oil immersion objective lens (PlanApo, Zeiss, Germany), a 470/40BP excitation filter, a FT 495 beam splitter and a 605/70BP emission filter. Light power density (3.78 mW/cm^2) was measured at a rear focal plane of the objective lens using the PM100D power meter (ThorLabs, Germany) equipped with the S120VS sensor (ThorLabs, Germany). No corrections were applied to the experimental data.

Size-exclusion chromatography was performed with a Superdex™ 75 10/300 GL column using the GE AKTA Explorer (Amersham Pharmacia, UK) FPLC System.

To assess the maturation rate of LSSRFPs, 20 mL of bacterial cultures were grown in 50 mL tubes with LB medium supplemented with 100 $\mu\text{g/mL}$ ampicillin at 37 °C, 190 rpm, overnight. Next, protein expression was induced by the addition of 0.2% arabinose, and the culture was transferred into a 15 mL tube until the brim. Protein expression lasted for 3 h at 37 °C, 190 rpm at anaerobic conditions. The cultures were then centrifuged at $3500 \times g$ for 10 min at room temperature. The pellet was resuspended in 0.8 mL of PBS supplemented with 25 $\mu\text{g/mL}$ chloramphenicol (to block protein translation), transferred into a 2 mL tube, and sonicated for 40 sec, 20% power using the VCX130 Sonicator and CV18 tip (Sonics&Materials Inc., Newtown, CT, USA) on ice followed by centrifugation at $46,090 \times g$ for 4 min at 0 °C. A total of 100 μL of supernatant were mixed with 2.9 mL of PBS buffer supplemented with 25 $\mu\text{g/mL}$ of chloramphenicol (pre-warmed at 37 °C for 10 min) in a 5 mL quartz cuvette. Fluorescence kinetics was further measured using the CM2203 spectrofluorometer (Solar, Minsk, Belarus) at 37 °C.

For preparative protein purification for X-ray crystallography, the bacterial cells expressing the LSSmScarlet protein with a His-tag and Tobacco Etch Virus (TEV) protease cleavage site were harvested by centrifugation for 20 min at $5000 \times \text{rpm}$ and 4 °C using the Avanti J-E centrifuge (Beckman Coulter, USA). The pellet was further resuspended in 40 mM Tris(tris(hydroxymethyl)aminomethane)-HCl buffer, pH 7.5, supplemented with 400 mM NaCl, 10 mM Imidazole, 0.2% Triton X-100 and 1 mM PMSF (phenylmethylsulfonyl fluoride) (7 mL per 1 g of the cells) and disrupted by sonication (2 sec pulse-6-sec pause, 45% amplitude, for total time of 5 min). The crude cell extract was centrifuged for 30 min at $28,000 \times g$ and 4 °C using the Avanti J-E centrifuge (Beckman Coulter, USA). The supernatant was applied to a 5 mL Ni-NTA Superflow column (Qiagen, Hilden, Germany) equilibrated with the binding buffer (40 mM Tris-HCl, pH 7.5, containing 400 mM NaCl, 10 mM imidazole and 0.1% (v/v) Triton X-100). The column was further washed using the same binding buffer without Triton X-100 followed by a second washing using binding buffer in the absence of Triton X-100, supplemented with 40 mM imidazole. Protein elution was performed using the same binding buffer without Triton X-100 and supplemented with 300 mM imidazole. Both 1mM DTT (dithiothreitol) (final concentration) and

1mM EDTA (ethylenediaminetetraacetic acid) (final concentration) were added to the protein solution and mixed with TEV protease (1 mg per 10 mg of the protein). The final mix was dialyzed for 16 h in dialysis buffer (40 mM Tris-HCl, pH 7.8, 400 mM NaCl, 5 mM Imidazole, 2 mM BME (2-mercaptoethanol), 1 mM EDTA), at +4 °C (monitoring of the His-tag cleavage was conducted using 12% SDS-PAGE (sodium dodecyl sulphate–polyacrylamide gel electrophoresis)). After dialysis, the protein solution was applied to a Ni-NTA Superflow column (Qiagen, EU), equilibrated with dialysis buffer; TEV protease and cleaved His-tag were absorbed by the Ni-NTA Superflow column (Qiagen, EU), and the flow-through was concentrated until there was a 1 mL volume using a 10 kDa cutoff centrifugal filter device (Millipore, Burlington, MA, USA). The buffer of the concentrated protein was exchanged for the 20 mM Tris pH 7.8 buffer using the HiTrap Desalting 5 mL column (GE Healthcare, Sweden). The concentrated protein was further applied to a 1 mL ResourceQ column (GE Healthcare, Sweden) equilibrated with the same 20 mM Tris pH 7.8 buffer. Recombinant LSSmScarlet was eluted using a linear gradient from 0 to 1M NaCl; the protein was eluted by two peaks at 150 and 180 mM NaCl. Protein concentration was measured for each of the fractions individually using the Bicinchoninic Acid Protein Assay Kit (Sigma-Aldrich, Saint Louis, USA), with the BSA (bovine serum albumin) protein standard (P0914-5AMP, Sigma-Aldrich, Saint Louis, USA) solution as the standard. Total protein yield was 2.5 and 2.6 mg for the first and second fractions, respectively. The fractions were concentrated individually using a 10 kDa cutoff centrifugal filter device (Millipore, Burlington, MA, USA) until a 5 mg/mL concentration was reached and loaded onto the Superdex 75 10/300 GL column (GE Healthcare, Sweden), in a 20 mM Tris-HCl, pH 7.8, 200 mM NaCl buffer. In both cases, 11.5–13.5 mL fractions (corresponding to the monomer) were collected, and finally both fractions were combined and concentrated until a 10 mg/mL concentration for the crystallization was achieved. The total yield of the protein was 5 mg. Protein purity at each step was monitored using 12% SDS-PAGE. Chromatography was performed using ÄKTA prime plus and ÄKTA explorer 100 systems (GE Healthcare, Sweden).

3.3. Protein Crystallization

An initial crystallization screening of LSSmScarlet was performed with a robotic crystallization system (Rigaku, USA) and commercially available 96-well crystallization screens (Hampton Research and Anatrace, USA) at 15 °C using the sitting drop vapor diffusion method. The protein concentration was 10 mg/mL in the following buffer: 20 mM Tris-HCl, 200 mM NaCl, pH 7.8. Optimization of the initial conditions was performed by the hanging-drop vapor-diffusion method in 24-well VDX plates. Rod-like crystals were obtained within 2 weeks in the following conditions: 0.1M Lithium sulfate, 0.1M Ammonium acetate, 0.1 M Bis-tris pH 6.5, 23% PEG 3350.

3.4. Data Collection, Processing, Structure Solution and Refinement

LSSmScarlet crystals were briefly soaked in a 100% Paratone oil (Hampton research, USA) immediately prior to diffraction data collection and flash-frozen in liquid nitrogen. The X-ray data were collected from a single crystal at 100 K at the beamline “Belok-RSA” of the Kurchatov SNC (Moscow, Russia) [25]. The data were indexed, integrated and scaled using the Dials program [26] (Table 3). The program Pointless [27] suggested the C2 space group.

The structure was solved by the molecular replacement method using the MOLREP program [28] and the structure of the Red Fluorescent Protein mScarlet (PDB ID 5LK4) as an initial model. The refinement of the structure was carried out using the REFMAC5 program of the CCP4 suite [29]. The visual inspection of electron density maps and the manual rebuilding of the model were carried out using the COOT interactive graphics program [30]. The resolution was successively increased to 1.40 Å, and the hydrogen atoms in fixed positions, as well as anisotropic refinement, were introduced during the final

refinement cycles. In the final model, an asymmetric unit contained one independent copy of the protein of 227 residues with the chromophore together with 272 water molecules, two calcium ions and one sulfate molecule from the crystallization solution. The first residue from the N-terminal, G226 as well as the last one residue from the C-terminal part of the protein were not visible in electron density.

3.5. Structure Analysis and Validation

The visual inspection of the structure was carried out using the COOT program and the PyMOL Molecular Graphics System, Version 1.9.0.0 (Schrödinger, USA). The structure comparison and superposition were made using the PDBeFold program [31], while contacts were analyzed using the PDBePISA [32] and WHATIF software [33].

3.6. Mammalian Plasmids Construction

In order to construct the pAAV-CAG-LSSRFP-P2A-EGFP plasmids, the LSSRFP genes were PCR amplified as the KpnI-AgeI fragments, using LSSCy-KpnI/LSSCy-AgeI-r or LSSCR-KpnI/LSSCR-AgeI-r primers listed in the Table S3, and swapped with the R-GECO1 gene in the pAAV-CAG-R-GECO1-P2A-EGFP vector.

In order to construct the pLSSmScarlet-actin plasmid, the LSSmScarlet gene was PCR amplified as the NheI-HindIII fragment, using NheI-LSSmSc/LSSmSc-HindIII-r primers listed in Table S3, and swapped with the TagBFP gene in the pTagBFP-actin vector (Evrogen, Russia).

In order to construct the pLSSmScarlet-tubulin plasmid, the LSSmScarlet gene was PCR amplified as the NheI-BsrGI fragment, using NheI-LSSmSc/LSSmSc-HindIII-r primers listed in Table S3, and swapped with the TagGFP2 gene in the pTagGFP2-tubulin vector (Evrogen, Russia).

In order to construct the pAAV-CAG-H2B-LSSmScarlet plasmid, the LSSmScarlet gene was PCR amplified as the BglII-HindIII fragment, using mSc-BglII/mCherry-HindIII-r primers listed in Table S3, and swapped with the B-GECO1 gene in the pAAV-CAG-H2B-B-GECO1 vector [7].

In order to construct the pAAV-CAG-dMito-LSSmScarlet plasmid, the LSSmScarlet gene was PCR amplified as the XhoI-EcoRI fragment, using LSSmSc-XhoI/mCherry-EcoRI-r primers listed in Table S3, and swapped with the mCherry gene in the pAAV-CAG-dMito-mCherry vector.

In order to construct the pLU-CMV-vimentin-LSSmScarlet plasmid, the LSSmScarlet gene was PCR amplified as the BamHI-BsrGI fragment, using LSSmSc-BamHI/LSSmSc-XbaI-r primers listed in the Table S3, and swapped with the mCherry gene in the pLU-CMV-vimentin-NeonOxIrr vector [20].

Vimentin-emiRFP703 plasmid was purchased from Addgene (#136566).

3.7. Mammalian Live- and Fixed-Cell Imaging

Transient transfection of the HeLa Kyoto cells was performed in a 24-well format using lipofectamine reagent according to the manufacture's protocol. Cells were cultured using DMEM medium supplemented with 10% FBS, Glutamine, 50 U/mL penicillin and 50 U/mL streptomycin, at 37 degrees and 5% CO₂. HeLa cell cultures were imaged 24 h after the transient transfection using a laser spinning-disk Andor XDi Technology Revolution multi-point confocal system (Andor Technology, UK) equipped with an inverted Nikon Eclipse Ti-E/B microscope (Nikon Instruments, Japan), a 75 W mercury-xenon lamp (Hamamatsu, Japan), a 60× oil immersion objective NA 1.4 (Nikon, Japan), a 16-bit Neo sCMOS camera (Andor Technology, UK), laser module Revolution 600 (Andor Technology, UK), spinning-disk module Yokogawa CSU-W1 (Andor Technology, UK). The green and lss-red fluorescence were acquired using 80% of the 488 nm laser power (14 µW/cm² before objective lens), a confocal dichroic mirror 405/488/561/640 and filter wheel emission

filters 525/50 and 617/73, respectively. During imaging, cells were incubated at 37 degrees and 5% CO₂ using a cage incubator (Okolab, Italy).

Two-photon imaging of the cells was performed on a 33 mm dish with glass-bottom. Furthermore, 24 h after transfection in a 24-well format, cells were treated with 0.25% trypsin solution and transferred onto the 33 mm dish. Then, 24 h later, cells were imaged using a two-photon microscope system (Thorlabs, Newton, NJ, USA) equipped with the Chameleon Ultra I Ti:Sapphire laser (Coherent, Santa Clara, CA, USA), X-Cite 200DC light source (Lumen Dynamics, Canada), Galvo-Galvo Scan head and controller (Thorlabs, Newton, NJ, USA), the Model 302A amplifier (ConOptics, Danbury, CT, USA), 25× water dipping or immersion objective 1.10 NA (Nikon, Japan) and 525/50 and 607/70 emission filters.

3.8. Statistics

To estimate the significance of the difference between two values, we used the Mann–Whitney Rank Sum Test and provided *p*-values (throughout the text in the brackets) calculated for the two-tailed hypothesis. We considered the difference as significant if the *p* value was <0.05.

3.9. Molecular Modeling

The full-atom molecular model of the LSSmScarlet was built based on the X-ray structure obtained in this study. It was solvated in a rectangular water box of 80-62-68 Å³. The CHARMM27 [34,35] force field parameters for the protein and chromophore and TIP3P [36] parameters for water molecules were utilized for the preliminary equilibration of the system. We performed 1000 steps energy minimization followed by the 10 ns classical molecular dynamic (MD) run. Calculations were performed in the NAMD program package [37]. Thus, the prepared system was utilized for the molecular dynamics simulations with the combined quantum mechanics/molecular mechanics (QM/MM) potentials. The QM subsystem included the chromophore and the side chains of neighboring residues: R71, R96, D147, E149, R198, E216 and 4 water molecules of the chromophore-containing pocket. The QM subsystem was treated at the PBE0-D3/6-31G** level [38–40] on the ground electronic state energy surface, S₀. Energies and forces in the QM subsystem were calculated with the TeraChem program [41]. After equilibration of the system for 20 ps, we performed the QM/MM MD simulation on the first excited singlet state, S₁. We relied on the DFT functional CAM-B3LYP [42] with the def2-SVP basis set utilized for a similar system in ref [43]. Excited state QM/MM MD simulations were performed in the ORCA program [44]. All QM/MM MD simulations were performed with the help of the NAMD interface for the QM/MM treatment of the system [45]. All molecular dynamic simulations were performed at *p* = 1 atm and T = 300 K with the 1 fs integration time step.

3.10. . Fixation of the Cells with 4% PFA.

Before fixation, cells in a 24 well plate with glass-bottom were washed with 1 ml of DPBS buffer. Next, fixation of the cells was performed with 1 ml of 4% PFA for 15 min at room temperature. The sample was further washed with 1 mL of DPBS buffer and kept in 1 ml of DPBS buffer for further imaging.

3.11. STED Imaging

STED imaging was performed on a commercial STED microscope (STED Facility line, Abberior Instruments, Germany) equipped with 60×/1.42 N.A. The oil STED White objective lens (UPLXAPO60XO, Olympus, Japan) and standard 595 and 775 nm pulsed depletion lasers were used. The pixel size for the STED and confocal images was set to 30–40 nm. All images shown are raw data. Images were acquired using Imspector software (Max-Planck Innovation, Göttingen, Germany).

4. Conclusions

In conclusion, we developed and characterized a set of novel LSSRFPs characterized by the highest fluorescence brightness at a 488 nm excitation. LSSmScarlet was the first LSS version of the bright mScarlet red fluorescent protein (Figure 1). LSSmScarlet had medium brightness at a 488 nm excitation among a set of the developed LSSRFPs both as a purified protein (Table 1) and in the cytosol of mammalian cells (Figure 3b and Table 2). Moreover, LSSmScarlet had a moderate photostability and maturation rate compared to other LSSRFPs (Figure 2e,f and Table 1). However, LSSmScarlet was the only one among developed LSSRFPs for which fluorescence was not sensitive to alkaline changes of a pH up to 10 values (Figure 2d and Table 1); LSSmScarlet had also the highest pH-stability under acidic conditions. Furthermore, LSSmScarlet was monomeric (Figure 2c) and was successfully applied for the labeling and imaging of individual proteins in mammalian cells (Figure 4).

Since dCyRFP2s and dCyOFP2s and CRISPRed2s LSSRFPs proteins were developed from mCyRFP1, CyOFP1 and mCRISPRed LSSRFPs, respectively (Figure 1), we compared their biochemical properties to those of their parental proteins.

In contrast to monomeric mCyRFP1 LSSRFP, its derivative dCyRFP2s acquired a dimerization tendency due to the introduced mutations (Figure 2c). As compared to the mCyRFP1 progenitor, dCyRFP2s demonstrated 1.33- and 1.29-fold higher brightness at a 488 nm excitation as the purified protein (Table 1) and in mammalian cells (Figure 3b and Table 2), respectively. The dCyRFP2s variant had similar photostability and matured 1.5-fold slower compared to mCyRFP1 (Figure 2e,f and Table 1).

Similar to the CyOFP1 progenitor, dCyOFP2s was demonstrated to be a dimer (Figure 2c). dCyOFP2s had the highest brightness at a 488 nm excitation among all developed LSSRFPs both as purified proteins (Table 1) and in mammalian cells (Figure 3b and Table 2). dCyOFP2s also demonstrated the highest photostability at a 488 nm excitation and fastest maturation rate among all developed LSSRFPs (Figure 2e,f and Table 1).

In contrast to monomeric mCRISPRed LSSRFP, CRISPRed2s protein revealed a tendency to tetramerization (Figure 2c). CRISPRed2s protein had similar characteristics as mCRISPRed protein (Figure 2 and Table 1), except for a 3.9-fold faster maturation rate (Figure 2f). Similar to LSSmScarlet, CRISPRed2s had the highest pH stability at acidic conditions (Figure 2d).

Next, we solved the crystal structure of LSSmScarlet protein and suggested the mechanism of ESPT in LSSmScarlet protein, and elucidated the role of R198 residue in ESPT to the chromophore.

Finally, we demonstrated two- and three-color confocal and two-color 2P imaging with LSSmScarlet protein. We found that LSSmScarlet protein is not applicable for super-resolution STED microscopy as a consequence of the LSSRed-to-Red photoconversion with 775 nm depletion light; however, LSSRed-to-Red photoconversion was not efficient upon one-photon illumination with 405 and 468 nm light.

Overall, the new set of LSSRFPs that we present offers researchers a choice for the LSSRFP with the highest brightness, extreme pH-stability, high photostability, fast maturation or monomeric behavior, depending on the particular application. LSSmScarlet and CRISPRed2s due to their high pH-stability might be applied for the labeling of acidic organelles such as lysosomes. Monomeric behavior of LSSmScarlet allows using it for the labeling of subcellular structures, however, with limited applicability for the super-resolution STED microscopy; photoconversion of LSSmScarlet protein with 775 nm light makes it good candidate for photoactivation localization microscopy (PALM) imaging similar to the PSLSSmKate protein [22]. High brightness of dCyRFP2s and dCyOFP2s might be suitable for the labeling of cells and tissues, and for the development of the novel genetically encoded biosensors. Monomerized versions of the dCyRFP2s and dCyOFP2s LSSRFPs with high photostability might be possible candidates for super-resolution STED microscopy. Fast maturing dCyOFP2s might be useful for tracking promoter triggering and gene expression initiation.

Supplementary Materials: Supplementary Materials can be found at www.mdpi.com/article/10.3390/ijms222312887/s1.

Author Contributions: O.M.S., and F.V.S. developed the LSSmScarlet, dCyRFP2s, dCyOFP2s and CRISPRed2s fluorescent proteins. A.V.V., Y.K.A. and P.V.D. performed FPLC and purification of LSSRFPs proteins. O.M.S. and F.V.S. characterized properties of LSSRFPs in vitro, and in HeLa cells. A.Y.N. and K.M.B. crystallized LSSmScarlet and solved its structure. M.G.K. performed molecular modeling simulations. T.A.S. and F.V.S. performed STED microscopy. O.M.S., O.I.I., V.O.P., K.D.P., M.G.K., K.M.B. and F.V.S. analyzed and interpreted the data. O.M.S., K.D.P., M.G.K., K.M.B. and F.V.S. wrote the manuscript. All authors reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by an internal grant of the National Research Center Kurchatov Institute №2752 of 28.10.2021 (developing the LSSmScarlet, dCyRFP2s, dCyOFP2s and CRISPRed2s proteins), by RFBR grant №19-04-00395 to O.M.S. (confocal imaging), by RFBR grant № 18-29-13006 to M.G.K. (molecular modeling), by the Russian Science Foundation grant № 21-74-20135 to K.M.B. (crystallization, data collection and structural studies), the Ministry of Science and Higher Education of the Russian Federation grant №075-15-2019-1661 to T.A.S. (STED imaging), by RFBR grant № 20-015-00427 to O.I.I. (two-photon imaging), by start-up funding from the Foundation of Westlake University, National Natural Science Foundation of China grant 32050410298, 2020 BBRF Young Investigator Grant and MRIC Funding 103536022023 to K.D.P. The work was also supported by the Resource Centers department of the National Research Center Kurchatov Institute (imaging of bacterial cells and cultivation of mammalian cells), the supercomputer resources of the M.V. Lomonosov Moscow State University (molecular modeling simulations) and the equipment of IGB RAS core facilities.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article or supplementary material.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript or in the decision to publish the results.

Abbreviations

MDPI	Multidisciplinary Digital Publishing Institute
RFP	Red fluorescent protein
LSSRFP	Red fluorescent protein with large Stokes shift
FP	Fluorescent protein
PBS	Phosphate Buffered Saline
QY	Quantum yield
SD	Standard deviation

References

1. Piatkevich, K.D.; Malashkevich, V.N.; Almo, S.C.; Verkhusha, V.V. Engineering ESPT pathways based on structural analysis of LSSmKate red fluorescent proteins with large Stokes shift. *J. Am. Chem. Soc.* **2010**, *132*, 10762–10770. <https://doi.org/10.1021/ja101974k>.
2. 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/10.1152/physrev.00038.2009>.
3. Chu, J.; Oh, Y.; Sens, A.; Ataie, N.; Dana, H.; Macklin, J.J.; Laviv, T.; Welf, E.S.; Dean, K.M.; Zhang, F.; et al. A bright cyan-excitable orange fluorescent protein facilitates dual-emission microscopy and enhances bioluminescence imaging in vivo. *Nat. Biotechnol.* **2016**, *34*, 760–767. <https://doi.org/10.1038/nbt.3550>.
4. Subach, F.V.; Verkhusha, V.V. Chromophore transformations in red fluorescent proteins. *Chem. Rev.* **2012**, *112*, 4308–4327. <https://doi.org/10.1021/cr2001965>.
5. Subach, F.V.; Piatkevich, K.D.; Verkhusha, V.V. Directed molecular evolution to design advanced red fluorescent proteins. *Nat. Methods* **2011**, *8*, 1019–1026. <https://doi.org/10.1038/nmeth.1776>.
6. Subach, O.M.; Subach, F.V. GAF-CaMP3-sfGFP, An Enhanced Version of the Near-Infrared Genetically Encoded Positive Phytochrome-Based Calcium Indicator for the Visualization of Neuronal Activity. *Int. J. Mol. Sci.* **2020**, *21*, 6883. <https://doi.org/10.3390/ijms21186883>.

7. Subach, O.M.; Barykina, N.V.; Anokhin, K.V.; Piatkevich, K.D.; Subach, F.V. Near-Infrared Genetically Encoded Positive Calcium Indicator Based on GAF-FP Bacterial Phytochrome. *Int. J. Mol. Sci.* **2019**, *20*, 3488. <https://doi.org/10.3390/ijms20143488>.
8. Pedelacq, J.D.; Cabantous, S.; Tran, T.; Terwilliger, T.C.; Waldo, G.S. Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **2006**, *24*, 79–88. <https://doi.org/10.1038/nbt1172>.
9. Shaner, N.C.; Campbell, R.E.; Steinbach, P.A.; Giepmans, B.N.; Palmer, A.E.; Tsien, R.Y. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.* **2004**, *22*, 1567–1572.
10. Shaner, N.C.; Steinbach, P.A.; Tsien, R.Y. A guide to choosing fluorescent proteins. *Nat. Methods* **2005**, *2*, 905–909.
11. Grabe, M.; Oster, G. Regulation of organelle acidity. *J. Gen. Physiol.* **2001**, *117*, 329–344.
12. Bindels, D.S.; Haarbosch, L.; van Weeren, L.; Postma, M.; Wiese, K.E.; Mastop, M.; Aumonier, S.; Gotthard, G.; Royant, A.; Hink, M.A.; et al. mScarlet: A bright monomeric red fluorescent protein for cellular imaging. *Nat. Methods* **2017**, *14*, 53–56. <https://doi.org/10.1038/nmeth.4074>.
13. Bajar, B.T.; Wang, E.S.; Lam, A.J.; Kim, B.B.; Jacobs, C.L.; Howe, E.S.; Davidson, M.W.; Lin, M.Z.; Chu, J. Improving brightness and photostability of green and red fluorescent proteins for live cell imaging and FRET reporting. *Sci. Rep.* **2016**, *6*, 20889. <https://doi.org/10.1038/srep20889>.
14. Campbell, B.C.; Nabel, E.M.; Murdock, M.H.; Lao-Peregrin, C.; Tsoulfas, P.; Blackmore, M.G.; Lee, F.S.; Liston, C.; Morishita, H.; Petsko, G.A. mGreenLantern: A bright monomeric fluorescent protein with rapid expression and cell filling properties for neuronal imaging. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 30710–30721. <https://doi.org/10.1073/pnas.2000942117>.
15. Balleza, E.; Kim, J.M.; Cluzel, P. Systematic characterization of maturation time of fluorescent proteins in living cells. *Nat. Methods* **2018**, *15*, 47–51. <https://doi.org/10.1038/nmeth.4509>.
16. Matlashov, M.E.; Shcherbakova, D.M.; Alvelid, J.; Baloban, M.; Pennacchietti, F.; Shemetov, A.A.; Testa, I.; Verkhusha, V.V. A set of monomeric near-infrared fluorescent proteins for multicolor imaging across scales. *Nat. Commun.* **2020**, *11*, 239. <https://doi.org/10.1038/s41467-019-13897-6>.
17. Erdogan, M.; Fabritius, A.; Basquin, J.; Griesbeck, O. Targeted In Situ Protein Diversification and Intra-organelle Validation in Mammalian Cells. *Cell Chem. Biol.* **2020**, *27*, 610–621.e5. <https://doi.org/10.1016/j.chembiol.2020.02.004>.
18. Shu, X.; Shaner, N.C.; Yarbrough, C.A.; Tsien, R.Y.; Remington, S.J. Novel chromophores and buried charges control color in mFruits. *Biochemistry* **2006**, *45*, 9639–9647. <https://doi.org/10.1021/bi060773l>.
19. Subach, O.M.; Sotskov, V.P.; Plusnin, V.V.; Gruzdeva, A.M.; Barykina, N.V.; Ivashkina, O.I.; Anokhin, K.V.; Nikolaeva, A.Y.; Korzhenevskiy, D.A.; Vlaskina, A.V.; et al. Novel Genetically Encoded Bright Positive Calcium Indicator NCaMP7 Based on the mNeonGreen Fluorescent Protein. *Int. J. Mol. Sci.* **2020**, *21*, 1644. <https://doi.org/10.3390/ijms21051644>.
20. Subach, O.M.; Kunitsyna, T.A.; Mineyeva, O.A.; Lazutkin, A.A.; Bezryadnov, D.V.; Barykina, N.V.; Piatkevich, K.D.; Ermakova, Y.G.; Bilan, D.S.; Belousov, V.V.; et al. Slowly Reducible Genetically Encoded Green Fluorescent Indicator for In Vivo and Ex Vivo Visualization of Hydrogen Peroxide. *Int. J. Mol. Sci.* **2019**, *20*, 3138. <https://doi.org/10.3390/ijms20133138>.
21. Sun, N.; Yun, J.; Liu, J.; Malide, D.; Liu, C.; Rovira, I.I.; Holmstrom, K.M.; Fergusson, M.M.; Yoo, Y.H.; Combs, C.A.; et al. Measuring In Vivo Mitophagy. *Mol. Cell* **2015**, *60*, 685–696. <https://doi.org/10.1016/j.molcel.2015.10.009>.
22. Piatkevich, K.D.; English, B.P.; Malashkevich, V.N.; Xiao, H.; Almo, S.C.; Singer, R.H.; Verkhusha, V.V. Photoswitchable red fluorescent protein with a large Stokes shift. *Chem. Biol.* **2014**, *21*, 1402–1414. <https://doi.org/10.1016/j.chembiol.2014.08.010>.
23. Ho, S.N.; Hunt, H.D.; Horton, R.M.; Pullen, J.K.; Pease, L.R. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **1989**, *77*, 51–59.
24. Subach, F.V.; Patterson, G.H.; Manley, S.; Gillette, J.M.; Lippincott-Schwartz, J.; Verkhusha, V.V. Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nat. Methods* **2009**, *6*, 153–159.
25. Svetogorov, R.; Dorovatovskii, P.; Lazarenko, V. Belok/XSA Diffraction Beamline for Studying Crystalline Samples at Kurchatov Synchrotron Radiation Source. *Cryst. Res. Technol.* **2020**, *55*, 1900184.
26. Winter, G.; Waterman, D.G.; Parkhurst, J.M.; Brewster, A.S.; Gildea, R.J.; Gerstel, M.; Fuentes-Montero, L.; Vollmar, M.; Michels-Clark, T.; Young, I.D.; et al. DIALS: Implementation and evaluation of a new integration package. *Acta Crystallogr. Sect. D Struct. Biol.* **2018**, *74*, 85–97. <https://doi.org/10.1107/S2059798317017235>.
27. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *62* (Pt 1), 72–82. <https://doi.org/10.1107/S0907444905036693>.
28. Vagin, A.; Teplyakov, A. MOLREP: An automated program for molecular replacement. *J. Appl. Crystallogr.* **1997**, *30*, 1022–1025.
29. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1994**, *50* (Pt 5), 760–763. <https://doi.org/10.1107/S0907444994003112>.
30. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60* (Pt 12 Pt 1), 2126–2132. <https://doi.org/10.1107/S0907444904019158>.
31. Krissinel, E.; Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60* (Pt 12 Pt 1), 2256–2268. <https://doi.org/10.1107/S0907444904026460>.
32. Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007**, *372*, 774–797. <https://doi.org/10.1016/j.jmb.2007.05.022>.
33. Vriend, G. WHAT IF: A molecular modeling and drug design program. *J. Mol. Graph.* **1990**, *8*, 52–56, 29. [https://doi.org/10.1016/0263-7855\(90\)80070-v](https://doi.org/10.1016/0263-7855(90)80070-v).

34. Best, R.B.; Zhu, X.; Shim, J.; Lopes, P.E.; Mittal, J.; Feig, M.; Mackerell, A.D., Jr. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273. <https://doi.org/10.1021/ct300400x>.
35. Denning, E.J.; Priyakumar, U.D.; Nilsson, L.; Mackerell, A.D., Jr. Impact of 2'-hydroxyl sampling on the conformational properties of RNA: Update of the CHARMM all-atom additive force field for RNA. *J. Comput. Chem.* **2011**, *32*, 1929–1943. <https://doi.org/10.1002/jcc.21777>.
36. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935. <https://doi.org/10.1063/1.445869>.
37. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802. <https://doi.org/10.1002/jcc.20289>.
38. Francel, M.M.; Pietro, W.J.; Hehre, W.J.; Binkley, J.S.; Gordon, M.S.; DeFrees, D.J.; Pople, J.A. Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **1982**, *77*, 3654–3665. <https://doi.org/10.1063/1.444267>.
39. Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170. <https://doi.org/10.1063/1.478522>.
40. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **2010**, *132*, 154104. <https://doi.org/10.1063/1.3382344>.
41. Seritan, S.; Bannwarth, C.; Fales, B.S.; Hohenstein, E.G.; Isborn, C.M.; Kokkila-Schumacher, S.I.L.; Li, X.; Liu, F.; Luehr, N.; Snyder, J.W.; Song, C.; et al. TeraChem: A graphical processing unit-accelerated electronic structure package for large-scale ab initio molecular dynamics. *WIREs Comput. Mol. Sci.* **2020**, *11*, e1494. <https://doi.org/10.1002/wcms.1494>.
42. Yanai, T.; Tew, D.; Handy, N. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57. <https://doi.org/10.1016/j.cplett.2004.06.011>.
43. Chen, F.; Zeng, Q.; Zhuang, W.; Liang, W. Characterizing the Structures, Spectra, and Energy Landscapes Involved in the Excited-State Proton Transfer Process of Red Fluorescent Protein LSSmKate1. *J. Phys. Chem. B* **2016**, *120*, 9833–9842. <https://doi.org/10.1021/acs.jpcc.6b04708>.
44. Neese, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73–78. <https://doi.org/10.1002/wcms.81>.
45. Melo, M.C.R.; Bernardi, R.C.; Rudack, T.; Scheurer, M.; Riplinger, C.; Phillips, J.C.; Maia, J.D.C.; Rocha, G.B.; Ribeiro, J.V.; Stone, J.E.; et al. NAMD goes quantum: An integrative suite for hybrid simulations. *Nat. Methods* **2018**, *15*, 351–354. <https://doi.org/10.1038/nmeth.4638>.