J|A|C|S
ARTICLES

Published on Web 07/20/2010

Engineering ESPT Pathways Based on Structural Analysis of LSSmKate Red Fluorescent Proteins with Large Stokes Shift

Kiryl D. Piatkevich,[†] Vladimir N. Malashkevich,[‡] Steven C. Almo,[‡] and Vladislav V. Verkhusha*,[†]

Department of Anatomy and Structural Biology, Gruss-Lipper Biophotonics Center, and Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461

Received March 8, 2010; E-mail: vladislav.verkhusha@einstein.yu.edu

Abstract: LSSmKate1 and LSSmKate2 are monomeric red fluorescent proteins (RFPs) with large Stokes shifts (LSSs), which allows for efficient separation of absorbance and emission maxima, as well as for excitation with conventional two-photon laser sources. These LSSmKates differ by a single amino acid substitution at position 160 and exhibit absorbance maxima around 460 nm, corresponding to a neutral DsRed-like chromophore. However, excitation at 460 nm leads to fluorescence emission above 600 nm. Structures of LSSmKate1 and LSSmKate2, determined at resolutions of 2.0 and 1.5 Å, respectively, revealed that the predominant DsRed-chromophore configurations are cis for LSSmKate1 but trans for LSSmKate2. Crystallographic and mutagenesis analyses, as well as isotope and temperature dependences, suggest that an excited-state proton transfer (ESPT) is responsible for the LSSs observed in LSSmKates. Hydrogen bonding between the chromophore hydroxyl and Glu160 in LSSmKate1 and a proton relay involving the chromophore tyrosine hydroxyl, Ser158, and the Asp160 carboxylate in LSSmKate2 represent the putative ESPT pathways. Comparisons with mKeima LSS RFP suggest that similar proton relays could be engineered in other FPs. Accordingly, we mutated positions 158 and 160 in several conventional red-shifted FPs, including mNeptune, mCherry, mStrawberry, mOrange, and mKO, and the resulting FP variants exhibited LSS fluorescence emission in a wide range of wavelengths from 560 to 640 nm. These data suggest that different chromophores formed by distinct tripeptides in different environments can be rationally modified to yield RFPs with novel photochemical properties.

1. Introduction

On the basis of the far-red fluorescent protein (FP) mKate,[1] two red FP (RFP) variants with large Stokes shifts (LSSs), designated LSSmKate1 and LSSmKate2, were recently developed.[2] LSSmKate1 and LSSmKate2 differ from each other by a single amino acid substitution at position 160 (Figure 1) and exhibit excitation/emission maxima at 463/624 and 460/605 nm, respectively (Table 1). (The term of "large Stokes shift" is accepted nomenclature in the field of fluorescent proteins to denote fluorescent proteins exhibiting excitation and emission maxima that differ greater than ~100 nm. This definition, however, differs from that used in photochemistry (http://goldbook.iupac.org/S06031.html). To avoid confusion among readers from the field of biology, who are the major users of the LSSmKate proteins, we will continue to use the nomenclature appearing in previous publications in the field.[3−5]) The LSSs in the fluorescence emission make them highly attractive tools for multicolor two-photon in vivo imaging. Among all RFPs, the LSSmKate proteins exhibit the highest pH stability of red fluorescence reported to date. Moreover, they possess enhanced photostability, as well as rapid and efficient chromophore maturation. However, the mechanistic basis for their LSS fluorescence emission is unknown. Structural and spectroscopic studies of mKeima RFP,[3] which also exhibits the LSS, revealed a hydrogen-bonding network capable of supporting an excited-state proton transfer (ESPT) from the chromophore hydroxyphenyl donor to the side chain of Asp158. This arrangement provides a possible ESPT pathway consistent with the LSS properties of mKeima.[4,5]

A simple ESPT model also explains the key photophysical features of wild-type green FP (wtGFP) at room temperature.[6] According to this model, the two absorption bands of wtGFP with maxima at 396 and 476 nm can be attributed to the neutral A and anionic B ground-state forms of the chromophore,

[†] Department of Anatomy and Structural Biology, and Gruss-Lipper Biophotonics Center.
[‡] Department of Biochemistry.

(1) Shcherbo, D.; Merzlyak, E. M.; Chepurnykh, T. V.; Fradkov, A. F.; Ermakova, G. V.; Solovieva, E. A.; Lukyanov, K. A.; Bogdanova, E. A.; Zaraisky, A. G.; Lukyanov, S.; Chudakov, D. M. *Nat. Methods* **2007**, *4*, 741–746.
(2) Piatkevich, K. D.; Hulit, J.; Subach, O. M.; Wu, B.; Abdulla, A.; Segall, J. E.; Verkhusha, V. V. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 5369–5374.
(3) Kogure, T.; Karasawa, S.; Araki, T.; Saito, K.; Kinjo, M.; Miyawaki, A. *Nat. Biotechnol.* **2006**, *24*, 577–581.
(4) Violot, S.; Carpentier, P.; Blanchoin, L.; Bourgeois, D. *J. Am. Chem. Soc.* **2009**, *131*, 10356–10357.
(5) Henderson, J. N.; Osborn, M. F.; Koon, N.; Gepshtein, R.; Huppert, D.; Remington, S. J. *J. Am. Chem. Soc.* **2009**, *131*, 13212–13213.
(6) Chattoraj, M.; King, B. A.; Bublitz, G. U.; Boxer, S. G. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8362–8367.

 10.1021/ja101974k

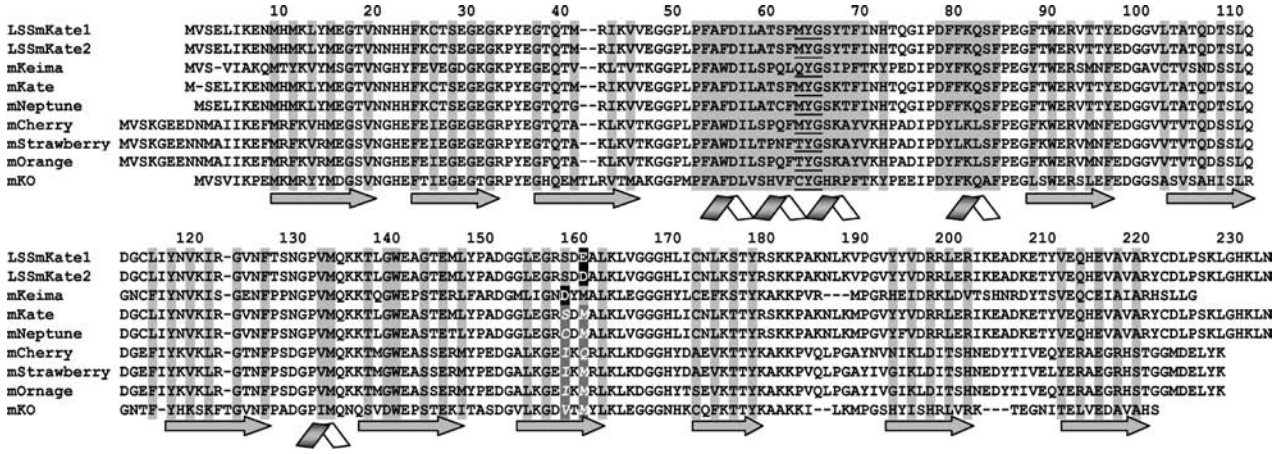

Figure 1. Amino acid sequence alignment of LSSmKate1 and LSSmKate2 with mKeima, mKate, mNeptune, mCherry, mStrawberry, mOrange, and mKO. Residues facing the interior of the protein's β-barrel are highlighted in light gray; α-helices are shown as ribbons; β-strands are shown as arrows. The chromophore-forming residues are underlined. The proton acceptor amino acid residues in LSSmKate1, LSSmKate2, and mKeima are shown in white on black. Mutated amino acids are shown in white italic on dark gray. The alignment numbering follows that of LSSmKate proteins.

Table 1. Spectroscopic Properties for the Mutants of LSSmKate1 and LSSmKate2

protein	excitation maximum, nm	emission maximum, nm	quantum yield	pK_a
LSSmKate1	463	624	0.08	3.2
LSSmKate1/S158A	463	624	0.09	3.5
LSSmKate2	460	605	0.17	2.7
LSSmKate2/S158A	463	624	0.08	3.4
LSSmKate2/S158T	463	627	0.12	3.2
LSSmKate2/S158M	463	627	0.08	3.6

respectively (Supporting Information Figure 1). Excitation of either the A or the B form results in green fluorescence from the anionic state. When excited at 396 nm, the excited state of the neutral chromophore A* rapidly converts via proton transfer to an intermediate excited species denoted I*. This species is an anionic form that subsequently emits green fluorescence and rarely undergoes further evolution to B*. Spectroscopic and structural studies have revealed that residues in the immediate vicinity of the chromophore adopt different conformations in the B and I forms.[7] After emission of a green photon, I* decays to ground state I, which converts to the ground state A, completing the wtGFP photocycle. Both A* decay and I* formation are slowed by replacing exchangeable protons with deuterons, suggesting that an ESPT process connects these states.[6,8] The ESPT rates also exhibit a strong temperature dependence. In most cases, the quantum yield of the green emission, arising upon excitation of band A, decreases as the temperature is lowered, and blue emission of the A* state can be observed in the steady-state fluorescence spectrum.[9] The X-ray structures of wtGFP and its mutants, which exhibit ESPT, revealed hydrogen-bond networks involving amino acid side chains and waters that can support proton movements in the excited state between the chromophore's tyrosyl hydroxyl group and proton acceptor groups.[10−12] Substitutions of the residues in these putative proton wires, as well as point mutations in the immediate chromophore environment, can lead to increased sensitivity to pH and formation of new ESPT pathways.

In the present study, we have addressed a possible role of ESPT in the LSS of the two LSSmKate RFPs. We have determined the X-ray structures of LSSmKate1 and LSSmKate2 at resolutions of 2.0 and 1.5 Å, respectively. These structures revealed possible ESPT pathways involving a hydrogen bond between the chromophore hydroxyl and Glu160 in LSSmKate1, and a hydrogen-bonding network composed of the chromophore tyrosine hydroxyl, Ser158, and Asp160 in LSSmKate2. Additional spectroscopic and biochemical studies of LSSmKates and their mutants confirmed that ESPT is responsible for the LSS in these RFPs. We next applied this structure-based knowledge to rationally modify several conventional orange FPs (OFPs) and RFPs into variants exhibiting LSS fluorescence emission.

2. Experimental Procedures

2.1. Mutagenesis, Expression, and Protein Purification. Site-specific mutations were introduced using overlapping PCR or QuikChange Mutagenesis kit (Stratagene). After mutagenesis, a pool of mutants was electroporated into the *E. coli* bacterial host LMG194 (Invitrogen) and grown overnight on LB/agar/ampicillin Petri dishes supplemented with 0.02% (w/v) L-arabinose at 37 °C. For spectroscopic studies, colonies of interest were cultured

(7) Palm, G. J.; Zdanov, A.; Gaitanaris, G. A.; Stauber, R.; Pavlakis, G. N.; Wlodawer, A. *Nat. Struct. Biol.* **1997**, *4*, 361–365.
(8) Lossau, H.; Kummer, A.; Heinecke, R.; Pöllinger-Dummer, F.; Kompa, C.; Bieser, G.; Jonsson, T.; Silva, C. M.; Yang, M. M.; Youvan, D. C.; Michel-Beyerle, M. E. *Chem. Phys.* **1996**, *213*, 1–16.
(9) McAnaney, T. B.; Shi, X.; Abbyad, P.; Jung, H.; Remington, S. J.; Boxer, S. G. *Biochemistry* **2005**, *44*, 8701–8711.
(10) Brejc, K.; Sixma, T. K.; Kitts, P. A.; Kain, S. R.; Tsien, R. Y.; Ormö, M.; Remington, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2306–2311.
(11) Hanson, G. T.; McAnaney, T. B.; Park, E. S.; Rendell, M. E.; Yarbrough, D. K.; Chu, S.; Xi, L.; Boxer, S. G.; Montrose, M. H.; Remington, S. J. *Biochemistry* **2002**, *41*, 15477–15488.
(12) Shu, X.; Kallio, K.; Shi, X.; Abbyad, P.; Kanchanawong, P.; Childs, W.; Boxer, S. G.; Remington, S. J. *Biochemistry* **2007**, *46*, 12005–12013.

overnight in 10 mL of LB/ampicillin with 0.2% (w/v) L-arabinose to induce protein expression. Bacterial cells were then lysed with B-PER II (Pierce).

For crystallization, PCR-amplified *Bgl*II/*Eco*RI fragments encoding LSSmKate1 and LSSmKate2 were cloned into a pBAD/His-B vector (Invitrogen), modified by shortening the N-terminal polyhistidine tag to the MGSHHHHHHGRS-amino acid sequence. LSSmKate1 and LSSmKate2 were expressed in LMG194 bacterial host (Invitrogen) in LB medium supplemented with 0.005% arabinose for 24 h at 37 °C. The culture was centrifuged at 5000 rpm at 4 °C for 15 min. The cell pellet was resuspended in 50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0 buffer and lysed by sonication on ice. The recombinant protein was purified using Ni-NTA agarose (Qiagen) followed by the dialysis against 10 mM NaH_2PO_4, pH 7.5.

2.2. Spectroscopic and Biochemical Procedures. Excitation and emission spectra were recorded with a FluoroMax-3 spectrofluorometer (Jobin Yvon) and a SpectraMax-M2 plate reader (Molecular Devices). For absorbance measurements, a Hitachi U-2000 spectrophotometer was used. LB/agar Petri dishes were screened using a Leica MZ16F fluorescence stereomicroscope equipped with the standard blue, green, and red filter sets (Chroma). pH titrations were performed using a series of buffers (100 mM NaOAc, 300 mM NaCl for pH 2.5−5.0, and 100 mM NaH_2PO_4, 300 mM NaCl for pH 5.0−11.0).

In the LSSmKate samples, exchangeable protons were replaced with deuterium by repeated dilution of the protein with phosphate buffered saline (PBS)/glycerol (1:1 by v/v) buffer in D_2O (buffer pH's were not corrected for the isotope effect) followed by concentration by ultracentrifugation.

2.3. Protein Crystallization. Diffraction quality crystals were grown by sitting drop vapor diffusion by mixing 1 μL of protein (LSS−mKate1 and LSS−mKate2 concentrations were 21 and 25 mg/mL, respectively, in 10 mM NaH_2PO_4, pH 7.5) with 1 μL of reservoir solution and equilibrating the samples against the corresponding reservoir solution. For LSSmKate1, the reservoir solution contained 20% PEG 1000, 0.1 M phosphate-citrate, pH 5.6, 0.2 M Li_2SO_4, and for LSSmKate2, the reservoir solution contained 25% PEG 3350, 0.1 M Na-acetate, pH 4.5.

2.4. X-ray Diffraction Data Collection and Crystallographic Refinement. Crystals of both LSSmKate1 and LSSmKate2, with dimensions 0.2 × 0.2 × 0.3 mm^3, were mounted in cryo-loops directly from the crystallization droplet and flash-cooled in liquid nitrogen. Prior to freezing, a final concentration of 20% glycerol was added (as a cryo-protectant) to the droplets. Diffraction data were recorded on a Quantum 315 CCD detector (Area Detector Systems Corp., Poway) with 1.08 Å wavelength radiation on the X29A beamline (National Synchrotron Light Source, Brookhaven). Intensities were integrated using the HKL2000 program and reduced to amplitudes using the TRUNCATE program (see Supporting Information Table 1 for statistics).[13,14] Structures were determined by molecular replacement with PHASER.[15] Model building and refinement were performed with the programs COOT and REFMAC, respectively.[13,16] The quality of the final structures was verified with composite omit maps, and stereochemistry was checked with the programs WHATCHECK[17] and PROCHEK.[18] LSQKAB and SSM algorithms were used for structural superpositions.[14,19]

3. Results

3.1. Crystal Structures of the LSSmKate Proteins. To explore the molecular origin of the LSS, we determined crystal structures of LSSmKate1 and LSSmKate2 near their fluorescence maxima of pH 5.6 and 4.5, respectively.[2] Diffraction from the LSSmKate1 crystals is consistent with the space group $P4_3$ with unit cell dimensions $a = b = 72.18$ Å, $c = 226.56$ Å, $\alpha = \beta = \gamma = 90°$. Diffraction from the LSSmKate2 crystals is consistent with the space group $P2_1$, with unit cell dimentions $a = 71.70$ Å, $b = 49.45$ Å, $c = 127.17$ Å, $\alpha = \gamma = 90°$, $\beta = 93.13°$. Each crystal form contains four protein chains in the asymmetric unit (Supporting Information Table 1). Electron densities of the chromophores (formed by the Met63−Tyr64−Gly65 tripeptide) are consistent with that of the common DsRed-like chromophore.[20] In the LSSmKate1 structure, a cis configuration of the chromophore hydroxyphenyl ring predominates in all four molecules in the asymmetric unit. Electron density also indicates the presence of less than 40% of the chromophore in a trans configuration in two of the four molecules. In contrast, in the LSSmKate2 structure, the trans configuration of the chromophore hydroxyphenyl ring predominates, with a minor population (~10%) of a cis isomer present in two of the four molecules.

The cis and trans chromophores in LSSmKate1 and LSSmKate2, respectively, exhibit noticeable deviations from the coplanarity between the hydroxyphenyl and imidazolidinone rings. For the cis chromophore in LSSmKate1, the values of $\chi1$ and $\chi2$ torsion angles (i.e., around the Cα=Cβ and Cβ−Cγ bonds of tyrosine) are 6 ± 5° and −19 ± 5° (average over four subunits), respectively (Figure 2A). For the trans chromophore of LSSmKate2, these angles are 175 ± 5° and 24 ± 5°, respectively (Figure 2B).

3.2. Proton Relays in the LSSmKate Proteins. Despite being crystallized at different pH values, the position of the imidazolinone rings of the chromophores and the spatial arrangement of the side chains in the two LSSmKate structures are nearly identical (the average root mean-square deviation is 0.36 Å for all Cα atoms). However, the phenolic rings occupy slightly different positions. Determination of the detailed structures revealed two different proton relays. In LSSmKate1, the tyrosine hydroxyl group of the cis chromophore forms a hydrogen bond with the side chain of Glu160, which is hydrogen bonded to Ser158 and a water molecule (Figure 2A). In the LSSmKate2 structure, the hydroxyphenyl group of the trans chromophore forms a hydrogen bond with Ser158, which is hydrogen bonded to the carboxylate of Asp160 and a water molecule (Figure 2B). Additionally, the trans hydroxyphenolic ring is stabilized by a water-mediated hydrogen bond with Ser176 and by stacking interactions with the guanidinum group of Arg197. The main chain conformation between Ser158 and Asp160 appears to be well suited for a proton transfer. Structures of both LSSmKates revealed hydrogen-bonded water chains extending from the Asp/Glu160 side chain carboxylates to the exterior of the β-barrels. It is possible that the substantial difference between the wavelengths of the emission maxima and the quantum yields of LSSmKate1 and LSSmKate2 (Table 1) is caused by the different configurations of their chromophores and environments

(13) Otwinowski, W.; Minor, F. *Methods Enzymol.* **1997**, *276*, 307–326.
(14) Storoni, L. C.; McCoy, A. J.; Read, R. J. *Acta Crystallogr., Sect. D* **2004**, *60*, 432–438.
(15) Collaborative Computational Project, Number 4. *Acta Crystallogr., Sect. D* **1994**, *50*, 760−763.
(16) Emsley, P.; Cowtan, K. *Acta Crystallogr., Sect. D* **2004**, *60*, 2126–2132.
(17) Hooft, R. W.; Vriend, G.; Sander, C.; Abola, E. E. *Nature* **1996**, *381*, 272.
(18) Laskowski, R.; MacArthur, M.; Moss, D.; Thornton, J. *J. Appl. Crystallogr.* **1993**, *26*, 283–291.
(19) Krissinel, E.; Henrick, K. *Acta Crystallogr., Sect. D* **2004**, *60*, 2256–2268.
(20) Yarbrough, D.; Wachter, R. M.; Kallio, K.; Matz, M. V.; Remington, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 462–467.

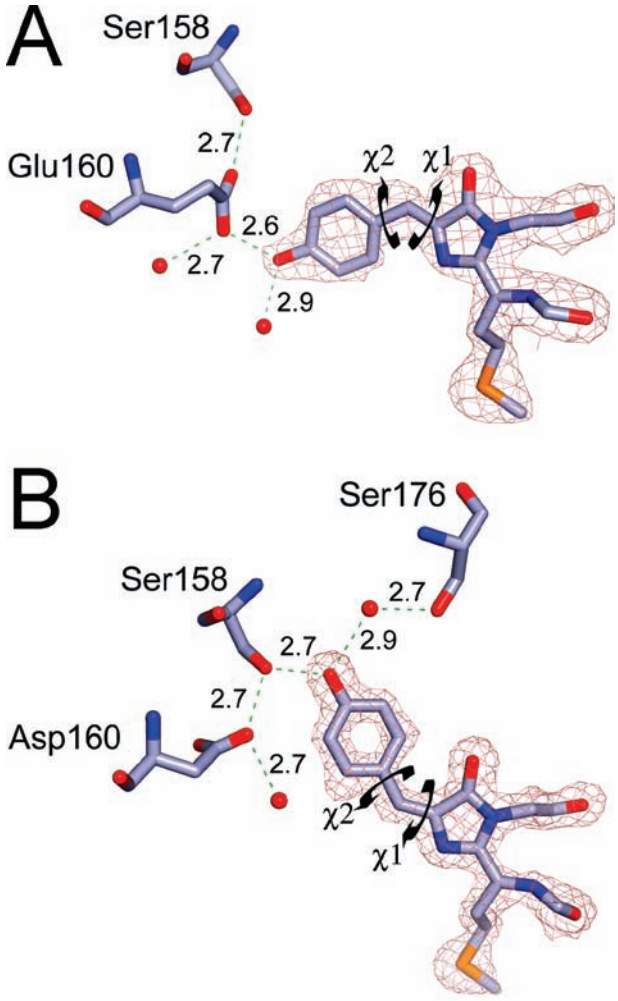

Figure 2. Molecular structures of the LSSmKate1 (A) and LSSmKate2 (B) chromophores and their immediate environments, superimposed over the corresponding $2F_o - F_c$ electron density (contoured at 1σ). Hydrogen bonds are represented as dashed green lines with the lengths indicated in angstroms; atoms are colored by atom type; water molecules are shown as red spheres. The arrows indicate χ_1 and χ_2 torsion angles.

of their phenolic rings, such as those observed in the case of eqFP611 mutants.[21]

3.3. Effect of pH on Spectral Properties. Absorption spectra of the LSSmKate proteins at pH's of 3, 5, 7, 9, and 11 were measured at room temperature (Figure 3). At pH's lower than 11, both proteins exhibit strong absorption centered at 460 nm, which we denoted as an A band. Unlike others fluorescent proteins with ESPT, the absorption maxima, as well as the spectral area of the A bands of LSSmKates, do not change significantly over the wide range of pH's spanning from 3 to 9. Fluorescence emission is maximal at pH 4.0−6.0 and drops at more acidic pH values (Supporting Information Figure 2). At pH 11, a new red-shifted absorbance peak appears at 579 nm, which we denoted as a B band. Excitation at the B band results in a red emission with a maximum at 628 nm, which is the same for both LSSmKate proteins (Figure 3). Interestingly, shifting LSSmKate samples from pH 11 back to pH 7 resulted in a disappearance of the B absorbance band and a complete restoration of the A band, suggesting that the LSSmKate proteins remain intact at pH 11.

3.4. Isotope and Temperature Dependences. To further define whether ESPT is responsible for the LSS of LSSmKates, we examined isotope and temperature dependencies of the steady-state emission spectra. In contrast to mKeima[4] and dual emission variants of GFP,[11] LSSmKate proteins have rather stable spectroscopic characteristics over a wide pH range (Supporting Information Figure 2). Therefore, the fluorescence spectra of LSSmKates were measured both in H_2O and in D_2O at pH and pD values near 7.5, both at room and at liquid nitrogen temperatures (Figure 4). Replacing exchangeable protons with deuterons at the same pH resulted in small emission red-shifts at room temperature. At low temperature, the emission maxima of LSSmKate1 and LSSmKate2 demonstrated blue shifts of 18 and 10 nm, respectively. Deuteration also caused a larger enhancement to the high-energy parts of the spectra. The enhancement of the blue shoulder at low temperature in the case of LSSmKate2 was about 2-fold greater than that of LSSmKate1. Similar deuterium exchange effects have been previously observed for mKeima[5] and several GFP variants exhibiting ESPT.[22]

The mutation of residues directly involved in the proton relay chain usually affects ESPT dynamics[23] and spectroscopic properties.[11,12] To test the role of Ser158 in the proposed proton transfer pathways, we performed site-directed mutagenesis and compared spectral characteristics of the mutants with those of LSSmKate2. We introduced Ser158Ala/Met mutations, which cannot support proton transfer, and a Ser158Thr mutation, which might partially disrupt the proton wire. The LSSmKate2/Ser158Ala/Thr/Met variants had substantially decreased quantum yields and red-shifted emission spectra as compared to those of the parental protein (Table 1). Moreover, pH stability of the mutants decreased as well. These substitutions in LSSmKate2 possibly caused destabilization of the trans configuration of the chromophore, shifting it toward the cis configuration, with an associated rewiring of the proton pathway.

3.5. Conversion of Conventional FPs into LSS Variants. LSSmKates and mKeima are the only LSS RFPs described to date. However, whereas LSSmKates exhibit an intact backbone, the main chain of mKeima is broken at positions between Leu62 and Gln63 and immediately preceding the chromophore. In addition, the first residues in the chromophore-forming tripeptides differ in LSSmKates and mKeima (MYG and QYG, respectively). Last, LLSmKates share only ~50% amino acid homology with mKeima. Despite these differences, it is likely that all of these proteins utilize ESPT to achieve the observed LSS; however, each utilizes a different proton relay (Figure 5). In LSSmKates and mKeima, the ultimate proton acceptor is an acidic side chain at either residue 158 or 160 that forms a hydrogen bond, directly or via a side chain of a serine side chain, with the hydroxyphenyl group of chromophores.

We reasoned that analogous ESPT pathways could be induced in other OFPs and RFPs with tyrosine-based chromophores by introducing Asp or Glu into the immediate chromophore environment. To test this hypothesis, we selected five new FPs, mNeptune,[24] mCherry,[25] mStrawberry,[25] mOrange,[25] and

(21) Nienhaus, K.; Nar, H.; Heilker, R.; Wiedenmann, J.; Nienhaus, G. U. *J. Am. Chem. Soc.* **2008**, *130*, 12578–12579.
(22) Shi, X.; Abbyad, P.; Shu, X.; Kallio, K.; Kanchanawong, P.; Childs, W.; Remington, S. J.; Boxer, S. G. *Biochemistry* **2007**, *46*, 12014–12025.
(23) Shu, X.; Leiderman, P.; Gepshtein, R.; Smith, N. R.; Kallio, K.; Huppert, D.; Remington, S. J. *Protein Sci.* **2007**, *16*, 2703–2710.
(24) Lin, M. Z.; McKeown, M. R.; Ng, H.-L.; Aguilera, T. A.; Shaner, N. C.; Campbell, R. E.; Adams, S. R.; Gross, L. A.; Ma, W.; Alber, T.; Tsien, R. Y. *Chem. Biol.* **2009**, *16*, 1169–79.

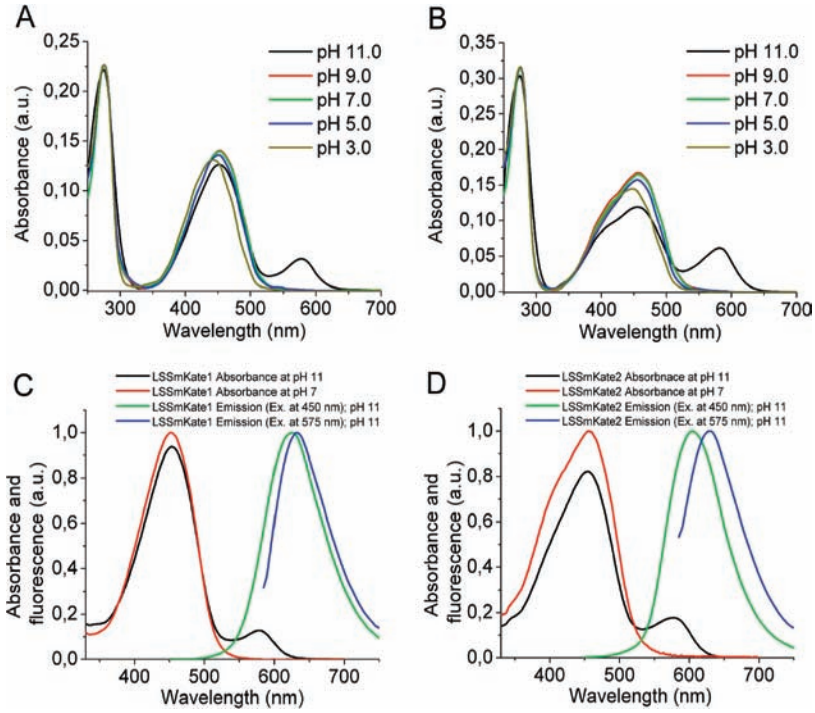

Figure 3. Absorbance and fluorescence spectra of LSSmKate1 (A,C) and LSSmKate2 (B,D) at different pH values (see also Table 1 and Supporting Information Figure 2). In panels C and D, the absorbance spectra at pH 7 were measured after the equilibration of protein samples from pH 11 back to the neutral pH value.

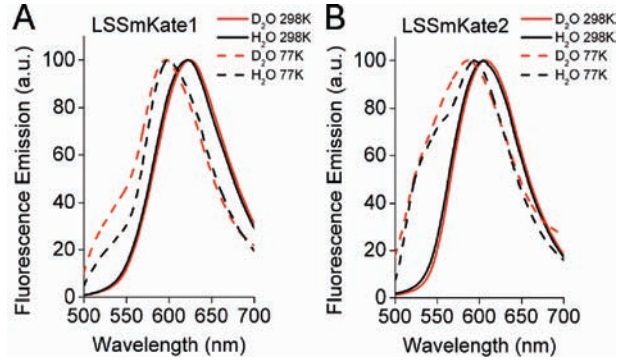

Figure 4. Fluorescence emission spectra of LSSmKate1 (A) and LSSmKate2 (B) in two different solvents at 77 and 298 K. The protein samples were excited at 450 nm. The spectra are normalized to 100%.

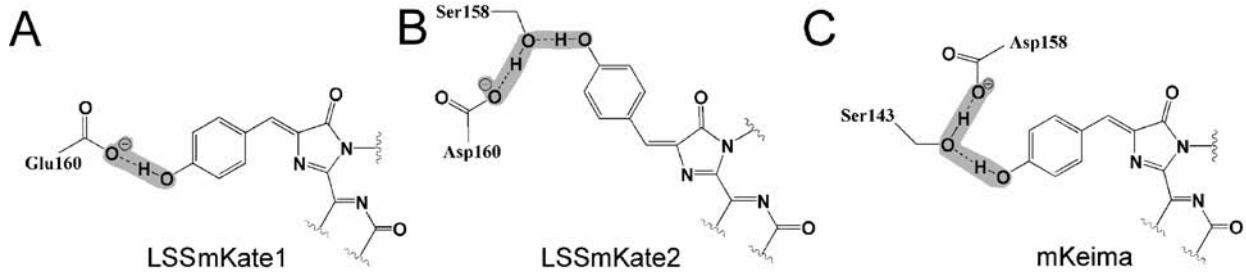

Figure 5. Proton relays of LSSmKate1 (A), LSSmKate2 (B), and mKeima (C) are shown for the chromophore and the residues involved in ESPT (hydrogen bonds are highlighted in gray).

mKO,[26] as well as the parental mKate, to introduce site-specific mutations that might enable ESPT. Our primary choice of positions for amino acid substitutions was based on the sequence alignment between LSSmKates and selected proteins (Figure 1).

To convert FPs into LSS variants, we constructed two sets of libraries for each protein. In one set, the codons in position 158 and 160 of selected FPs were randomized by using a primer

(25) Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N.; Palmer, A. E.; Tsien, R. Y. *Nat. Biotechnol.* **2004**, *22*, 1567–1572.
(26) Karasawa, S.; Araki, T.; Nagai, T.; Mizuno, H.; Miyawaki, A. *Biochem. J.* **2004**, *381*, 307–312.

Table 2. Spectroscopic Properties of the LSS Mutants of Conventional Orange and Red FPs

protein	mutations	excitation maximum, nm	emission maximum, nm
mKate	none	588	635
	160D	449	607
	160E	451	624
	158G, 160E	451	623
	158D	450	623
mNeptune	none	600	650
	158D	460	640
	158G, 160E	467	633
	158S, 160E	466	628
mCherry	none	587	610
	158E, 160A	457	610
	158S, 160E	456	610
mStrawberry	none	574	596
	158D	438	573
	158D, 160 L	438	572
mOrange	none	548	562
	158D, 160G	433	560
	158D, 160A	434	560
	158D, 160 L	441	563
	158E, 160A	436	561
	158A, 160E	441	561
mKO	none	548	559
	158E	439	560
	158D, 160 L	438	559

with a codon GAN, encoding Asp or Glu in position 158 and a degenerate codon NNS in position 160 (N is any nucleotide, S is T or C). In the other set of libraries, the codons encoding position 158 and 160 were randomized by using a primer with a degenerate codon NNS in position 158 and GAN codon in position 160. We selected clones that (i) exhibited the brightest orange or red fluorescence when excited with 436/20 nm light and (ii) had the minimal fluorescence when excited with 540/30 or 570/30 nm light for OFP and RFP mutants, respectively. Sequencing revealed mutants that encoded proteins with one or two amino acid substitutions at positions 158 and/or 160 (Table 2). Fluorescence excitation spectra of the selected variants indicated that even at neutral pH the majority of the protein species had a protonated chromophore (Figure 6). The LSS mutants of mKate and mNeptune exhibited a slight blue-shift in emission maxima. In the case of mCherry, mOrange, and mKO LSS variants, the fluorescence maxima were identical to those of the parental proteins. The mStrawberry mutants demonstrated a substantial blue shift of fluorescence. Fluorescent emission spectra also revealed a noticeable admixture of green emitting species in the mCherry LSS variants and mKate/158D mutant, which possibly had the GFP-like chromophore.

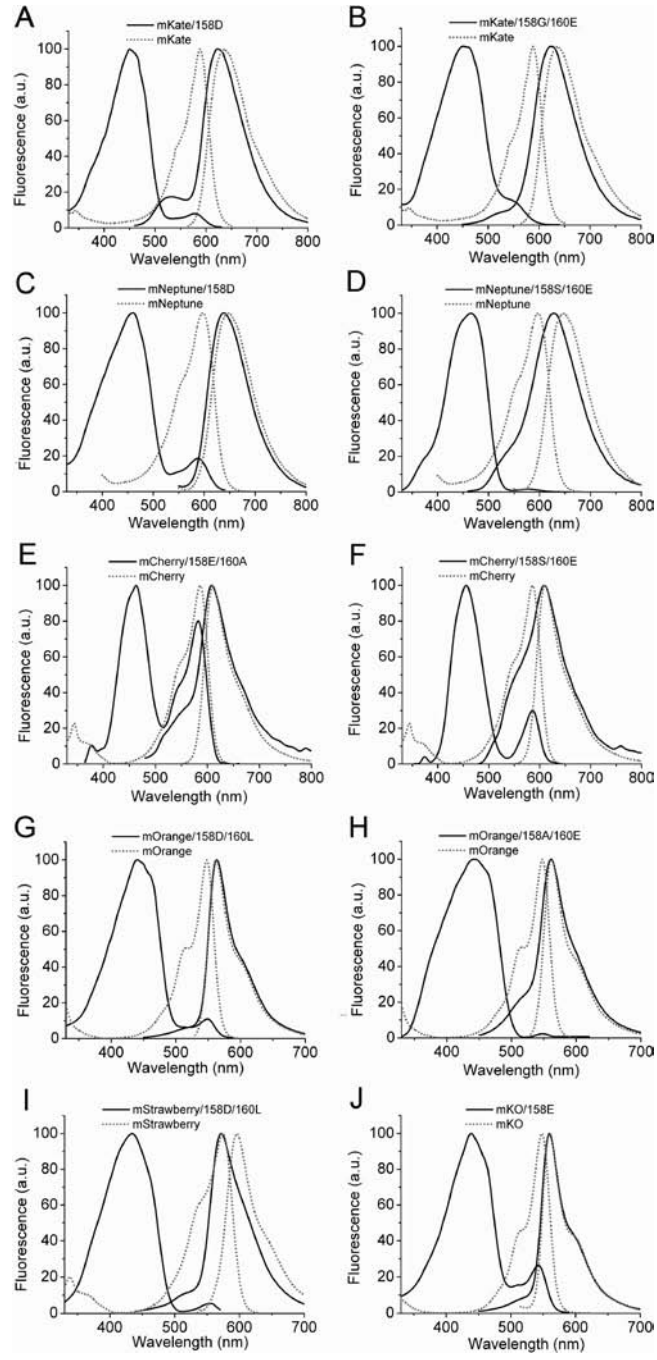

Figure 6. Fluorescence excitation and emission spectra are shown for mKate/158D (solid line) and mKate (dotted line) (A); mKate/158G/160E (solid line) and mKate (dotted line) (B); mNeptune/158D (solid line) and mNeptune (dotted line) (C); mNeptune/158S/160E (solid line) and mNeptune (dotted line); mCherry/158E/160A (solid line) and mCherry (dotted line) (E); mCherry/158S/160E (solid line) and mCherry (dotted line) (F); mOrange 158D/160 L (solid line) and mOrange (dotted line) (G); mOrange/158A/160E (solid line) and mOrange (dotted line) (H); mStrawberry/158D/160 L (solid line) and mStrawberry (dotted line) (I); and mKO/158E (solid line) and mKO (dotted line) (J).

4. Discussion

Several FPs with extended Stokes shifts have been recently developed. They include mPlum,[27] mKeima,[3] and LSSmKates.[2] In the case of mPlum, which was developed using somatic hypermutation and cell sorting to select for such mutations, the extended Stokes shift is associated with solvent reorganization.[28] In other FPs, however, the large Stokes shift is due to ESPT.

The chromophore environments of LSSmKate1 and LSSmKate2 exhibit high structural similarity, and the interactions between the chromophores and residues in their immediate proximity are also rather similar. The major structural difference is in the configurations of the chromophores, cis in LSSmKate1 and trans in LSSmKate2, which explains the difference in their fluorescence maxima. A single Glu160Asp substitution in LSSmKate1 favors the trans chromophore configuration and rewires the proton relay via Ser158. A replacement of Ser158 with Met, Thr, or Ala in LSSmKate2 (Table 1) converts the chromophore

(27) Wang, L.; Jackson, W. C.; Steinbach, P. A.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16745–9.
(28) Abbyad, P.; Childs, W.; Shi, X.; Boxer, S. G. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 20189–20194.

to its cis configuration. Crystallographic analysis of the LSSmKate2 structure suggests that the side chain of Asp160 may form a water-mediated hydrogen bond with the hydroxyphenyl group of the cis chromophore.

4.1. pH-Dependent Spectral Changes. Crystallographic analysis of LSSmKates provides an explanation for their pH-dependent spectral changes (Figure 3 and Supporting Information Figure 2). In LSSmKate1, a well-ordered hydrogen bond is formed between the chromophore and Glu160, while in LSSmKate2, Asp160 indirectly interacts with the chromophore via the side chain of Ser158. The Glu/Asp160 side chains can donate or accept a proton. It is expected that in the chromophore ground state, the pK_a of the Glu/Asp160 carboxylate is lower than the pK_a of the Tyr64 side chain hydroxyl,[29] and, consequently, these chromophores can be stabilized in a neutral form. In the excited state, the hydroxyphenol group of the chromophore is more acidic than the Glu/Asp160 residues, allowing for the latter residues to act as proton acceptors.[29] Reducing pH below 3.5 leads to the protonation of the carboxylate groups of the acceptors and an associated decrease of red fluorescence (Supporting Information Figure 2). The absorbance, however, does not change significantly (Figure 3). Thus, it is reasonable to assign the apparent pK_a values of LSSmKates to the direct titration of the carboxylate groups at position 160.

In the absorbance spectra of LSSmKates at pH 11, a red-shifted B band appears at 579 nm, which can be attributed to the deprotonated form of the chromophore (Figure 3). The maxima of the B bands as well as the maxima of fluorescence emission upon excitation of these bands (~630 nm) are the same for LSSmKate1 and LSSmKate2. We suggest that the B species for both LSSmKate proteins are identical and possibly share the cis configuration of the chromophore. The latter suggestion is supported by the observation that their emission maxima are close to the emission maximum of 625 nm of the A (protonated) species of the LSSmKate1 chromophore found in its cis configuration. Moreover, the parental mKate protein, which possesses a cis chromophore configuration, exhibits a fluorescence emission maximum at 633 nm.[30] Last, all mutants of LSSmKates at position 158 have the fluorescence maxima at 624−627 nm (Table 1).

A simple three-state chromophore model adequately explains the pH-dependent behavior of LSSmKates (Supporting Information Figure 3). At pH below 3.5, both the chromophore and the carboxylate groups of Glu/Asp160 are protonated, and ESPT does not occur. In the pH range of 3.5−10, the chromphore is protonated, the carboxylate groups of Glu/Asp160 are deprotonated, and efficient ESPT results in the strong LSS red fluorescence. At pH 11 and possibly above, an anionic form of the chromophore gives rise to conventional (small) Stokes shift red fluorescence. This behavior suggests that deprotonation of a neutral chromophore is energetically unfavorable, possibly due to the close proximity of two negatively charged groups (i.e., the carboxylic group of Glu160/Asp160 and the phenolic group of the chromophores). Moreover, the absorbance spectra of LSSmKates are reversibly influenced by pH (Figure 3C,D). Thus, we suggest that band A corresponds to the neutral state and band B corresponds to the anionic state of the chromophores in LSSmKates.

LSSmKates exhibit substantially greater pH stability of LSS fluorescence than mKeima.[2,4] A single Ser143Ala substitution in the vicinity of the hydroxyl group of the chromophore dramatically increased the pH stability of mKeima.[4] On the other hand, during our screening for LSSmKates,[2] we found a LSSmKate1/Y67K/G143S mutant that had a substantially lower pH stability, with an apparent pK_a value of 6.6 (Supporting Information Figure 2). Despite these substitutions, the LSS fluorescence emission is preserved in the mKeima/S143A and LSSmKate1/Y67K/G143S mutants. These data suggest that the Ala and Gly residues at position 143 in mKeima and LSSmKates, respectively, may destabilize the anionic form of the chromophore and preserve LSS fluorescence emission from the protonated state at low pH. Therefore, position 143 also plays an important role in the pH-dependent fluorescence properties of LSSmKates and mKeima.

4.2. Proposed Photocycle for LSSmKate Proteins. In LSSmKate1, a hydrogen bond between the chromophore hydroxyl and Glu160 enables ESPT from the chromophore donor to the carboxylate acceptor and may be responsible for the large Stokes shift (Figure 5A). A similar proton relay has been observed in the wtGFP/S65T/H148D (residues corresponding to positions 62 and 143 in LSSmKates) LSS green fluorescent variant, involving the chromophore and the carboxylate of Asp148.[12] In the case of LSSmKate2, a proton wire consisting of the tyrosine chromophore hydroxyl→Ser158→Asp160 could serve as a possible ESPT pathway (Figure 5B), which is similar to that observed for mKeima[5] (Figure 5C) and asFP499.[31] By analogy with wtGFP, we suggest a possible photocycle for LSSmKates. Figure 7 illustrates the key events of the model for photophysical and photochemical processes occurring in LSSmKates at neutral pH at room temperature. When excited at ~460 nm, the excited state of neutral chromophore A* rapidly decays, involving proton transfer to another excited species, which we denoted as I* for LSSmKate1 and LSSmKate2. Subsequently, the excited anionic I* forms emit red fluorescence and decay to ground anionic I states, which further convert to the ground state A, completing the photocycle. The B form, described above for GFP, also exists in LSSmKates and corresponds to the anionic cis chromophore, which is only observed at very alkali pH.

Our spectroscopic data are in agreement with the proposed photocycles for LSSmKates. The steady-state emission spectra of LSSmKates demonstrate a strong dependence on temperature and deuterium substitution, consistent with ESPT (Figure 4). As it is known that the magnitude of the high-energy shoulder is proportional to the lifetimes of intermediate states, we propose that the high-energy shoulders of the LSSmKate emission spectra at 77 K may correspond to the slowed ESPT of A* (Figure 7). In addition, the observed enhancement of the high-energy shoulders after the exchange of protons with heavier deuterons at 77 K suggests that A* decay to I* is possibly coupled to proton motion. The differences between the emission spectrum profiles of LSSmKate1 and LSSmKate2 at 77 K in both H_2O and D_2O (Figure 4) can be explained by the distinct proposed proton relays (Figure 5A,B). It is believed that protons in such "wires" exhibit close to synchronous behavior during ESPT and that solvent isotope effects increase with increasing number of contributing modes (or steps).[23,32] The more groups that are involved in the proton wire, the greater are the expected

(29) Wiehler, J.; Jung, G.; Seebacher, C.; Zumbusch, A.; Steipe, B. *ChemBioChem* **2003**, *4*, 1164–1171.
(30) Pletnev, S.; Shcherbo, D.; Chudakov, D. M.; Pletneva, N.; Merzlyak, E. M.; Wlodawer, A.; Dauter, Z.; Pletnev, V. *J. Biol. Chem.* **2008**, *283*, 28980–28987.
(31) Nienhaus, K.; Renzi, F.; Vallone, B.; Wiedenmann, J.; Nienhaus, G. U. *Biophys. J.* **2006**, *91*, 4210–4220.
(32) van Thor, J. J. *Chem. Soc. Rev.* **2009**, *38*, 2935–50.

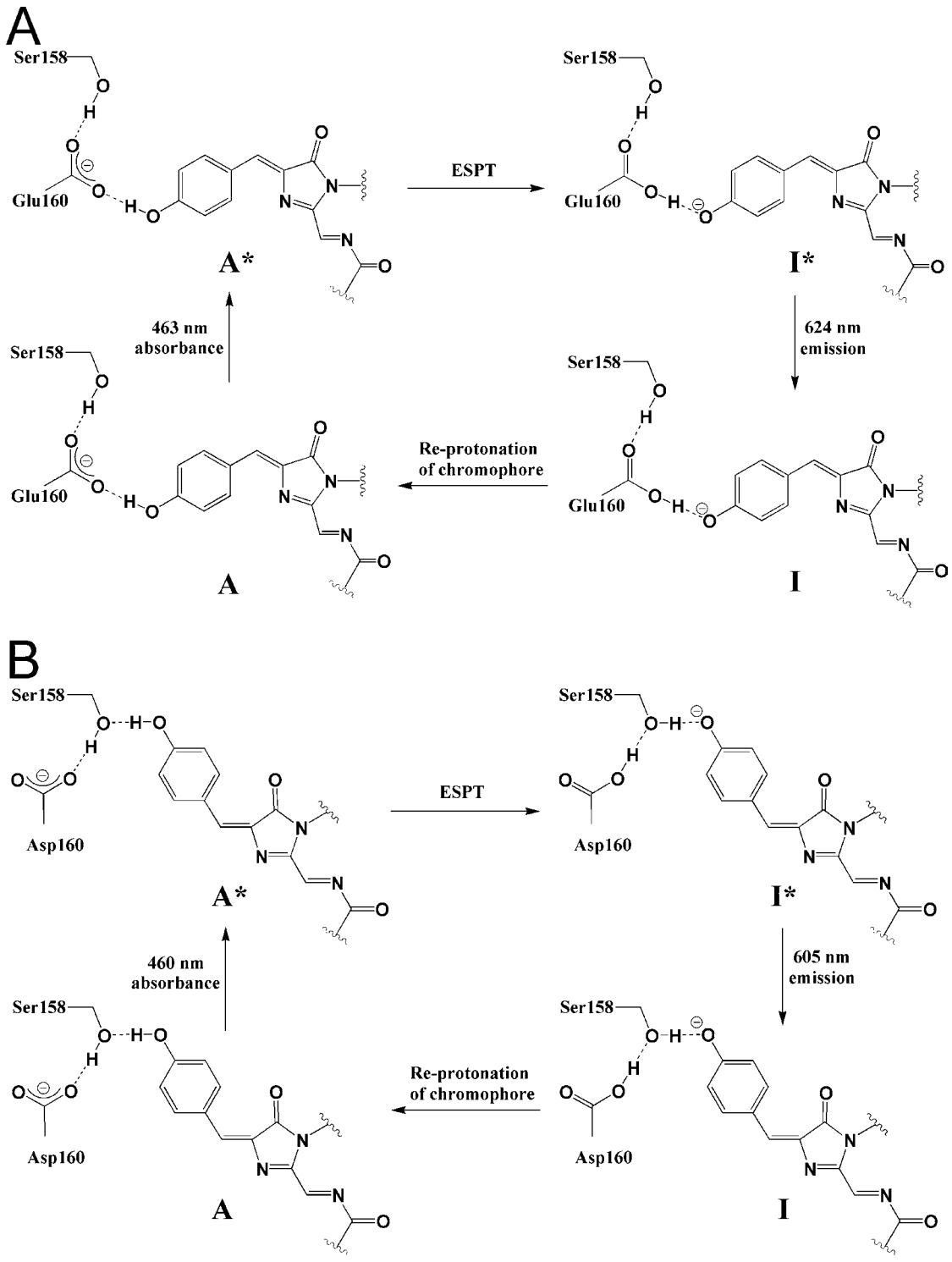

Figure 7. Proposed photocycles for LSSmKate1 (panel A) and LSSmKate2 (panel B) at neutral pH and room temperature. The excited state (compound A*) of the neutral chromophore, which is photoexcited from the ground state (compound A), is transformed after ESPT into the intermediate excited state denoted as the compound I*. After fluorescence emission, proton transfer from Glu160 (LSSmKate1) or Asp160 (LSSmKate2) to the chromophore regenerates the ground state (compound A) of the neutral chromophore.

effects of low temperature and isotope exchange. We reason that Glu160 in LSSmKate1 is protonated via a one-step proton-transfer reaction, while for LSSmKate2 the protonation of Asp160 is a coupled two-step process. Consistent with these principles, the effects of the isotopic exchange and the decrease of temperature are greater for LSSmKate2 than for LSSmKate1.

Our mutagenesis data also suggest that the proton relays are different in the two LSSmKate proteins. While Ser158 in LSSmKate2 is involved in the proton wire, it is not critical for ESPT. Indeed, the Ser158Ala/Thr/Met substitutions preserved ESPT but via an alternative proton relay, which is supported by the substantial changes in the fluorescence excitation and emission maxima (Table 1). In contrast, the Ser158Ala mutation in LSSmKate1 did not result in any noticeable changes of the spectroscopic properties, suggesting that Ser158 is not involved in ESPT pathway in this protein.

4.3. Properties of LSS Mutants of Conventional FPs. Given the close proximity of positions 158 and 160 to the hydroxyl group of the trans or cis chromophores, respectively, we hypothesized that the introduction of proton acceptors at these positions might lead to formation of a proton relay and consequently support ESPT in other FPs. Specifically, the introduction of Asp or Glu at these positions could lead to formation of a hydrogen bond with the phenolate oxygen and favor stabilization of the neutral chromophore, potentially supporting ESPT and associated LSS of fluorescence emission in the resulting mutants. Indeed, site-specific mutagenesis resulted in the conversion of several different OFPs and RFPs into the LSS variants (Table 2). The basis of the observed LSS fluorescence properties is possibly similar to that described for LSSmKates; however, the precise proton relays may differ from those of LSSmKates.

Because of the similarity of the emission maxima exhibited by the majority of LSS mutants and their respective parental FPs (Figure 6), we suggest that the emitting excited-state chromophore isoforms of the LSS mutants are identical to those present in their parental proteins (i.e., the introduced mutations did not cause chromophore cis−trans isomerization). Substantial blue shift in the emission of the mStrawberry LSS mutants as compared to mStrawberry (Table 2, Figure 6I) may be caused by isomerization of the cis chromophore of parental mStrawberry into a trans configuration in the LSS mutants. The observed small blue shoulders in the emission spectra of the LSS variants may be attributed to the presence of the GFP-like chromophores that appear after mutagenesis. Although beyond the scope of the current study, we believe that further improvement of the existing orange and far-red LSS variants will result in excellent intracellular probes with two spectrally resolvable LSS colors.

5. Summary

In this study, the crystal structures of the LSS RFPs, LSSmKate1 and LSSmKate2, were determined at high resolution. Crystallographic analysis, in combination with temperature- and hydrogen/deuterium isotope-dependencies of the fluorescence emission, suggested that ESPT is responsible for the LSS of both LSSmKate proteins. Despite nearly complete sequence identity, LSSmKate1 and LSSmKate2 possess different configurations of the DsRed-like chromophore and distinct ESPT pathways. The insights provided by the structural analyses of the LSSmKates and mKeima suggest that analogous proton relays can be designed in other red-shifted FPs. On the basis of this prediction, we generated LSS variants of several conventional orange and red FPs. Thus, different chromophores formed by distinct tripeptides in different environments can accommodate for related ESPT pathways. These data suggest that it is possible to rationally engineer proton wires through mutagenesis of the immediate chromophore environment to generate FPs with novel photophysical and photochemical properties.

The atomic coordinates and structure factors have been deposited in the Protein Data Bank (entry 3NT9 for LSSmKate1 and entry 3NT3 for LSSmKate2).

Acknowledgment. This work was supported by the grant GM073913 from the National Institutes of Health to V.V.V. and by the Albert Einstein Cancer Center.

Supporting Information Available: Photocycle of wtGFP; fluorescence of LSSmKate1, LSSmKate2, and LSSmKate1/Y67K/G143S at different pH's; proposed three-state model of LSSmKate1 and LSSmKate2 chromophores at different pH's; and X-ray data collection and refinement statistics for LSSmKate1 and LSSmKate2. This material is available free of charge via Internet at http://pubs.acs.org.

JA101974K