



SUBJECT AREAS:

BIOLOGICAL
FLUORESCENCE

PROTEIN DESIGN

BIOPHYSICAL CHEMISTRY

FLUORESCENCE SPECTROMETRY

Extended Stokes Shift in Fluorescent Proteins: Chromophore–Protein Interactions in a Near-Infrared TagRFP675 Variant

Kiryil D. Piatkevich¹, Vladimir N. Malashkevich², Kateryna S. Morozova¹, Nicolai A. Nemkovich³, Steven C. Almo² & Vladislav V. Verkhusha¹

¹Department of Anatomy and Structural Biology and Gruss-Lipper Biophotonics Center, ²Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, United States, ³B. I. Stepanov Institute of Physics, National Academy of Sciences of Belarus, 68 Independence Avenue, Minsk 220072, Belarus.

Received
23 April 2013Accepted
29 April 2013Published
15 May 2013

Correspondence and requests for materials should be addressed to V.V.V. (vladislav.verkhusha@einstein.yu.edu)

Most GFP-like fluorescent proteins exhibit small Stokes shifts (10–45 nm) due to rigidity of the chromophore environment that excludes non-fluorescent relaxation to a ground state. An unusual near-infrared derivative of the red fluorescent protein mKate, named TagRFP675, exhibits the Stokes shift, which is 30 nm extended comparing to that of the parental protein. In physiological conditions, TagRFP675 absorbs at 598 nm and emits at 675 nm that makes it the most red-shifted protein of the GFP-like protein family. In addition, its emission maximum strongly depends on the excitation wavelength. Structures of TagRFP675 revealed the common DsRed-like chromophore, which, however, interacts with the protein matrix via an extensive network of hydrogen bonds capable of large flexibility. Based on the spectroscopic, biochemical, and structural analysis we suggest that the rearrangement of the hydrogen bond interactions between the chromophore and the protein matrix is responsible for the TagRFP675 spectral properties.

GFP-like fluorescent proteins (FPs) are indispensable imaging tools for all areas of biomedical research^{1–3}. Three dimensional structures of GFP-like proteins are highly conserved, consisting of a beta-barrel formed by about 220–240 amino acids. A chromophore is buried inside the barrel, shielded by the protein matrix from stochastic interactions with solvent molecules. The rigid environment provided by protein scaffold prevents thermal isomerization and nonfluorescent relaxation of the chromophore. A wide variety of FP spectral phenotypes originate from two major contributing factors: the chemical structure of the chromophore, and interactions occurring between the chromophore, both in the ground and excited states, and its immediate environment.

Based on the chemical structure, which to a large extent determines the spectral properties of FPs, the chromophores can be classified into several groups⁴. Most of the red and far-red FPs contain so called DsRed-like chromophores⁵, which can exist in either neutral or anionic states. Neutral DsRed-like chromophores absorb blue-cyan light and emit green-yellow, whereas the anionic forms possess excitation and emission maxima at ~560–580 and 570–610 nm, respectively^{4,6,7}. However, fluorescence spectra can be significantly perturbed by changes in the immediate chromophore environment. Spectroscopic studies combined with high resolution crystal structures revealed the interactions responsible for the bathochromic shift of fluorescence in several red and far-red FPs.

Among the most common modifications of the chromophore environment in far-red FPs is the introduction of a hydrogen bond between the chromophore and its immediate environment. An important example is the hydrogen bond between the N-acylimine oxygen of the DsRed-like chromophore and a water molecule or a side chain of an amino acid (Figure 1A). This type of interaction has been observed in mNeptune⁸, eqFP650 (Ref. 9,10), eqFP670 (Ref. 9,10), mRojoA¹¹, mRouge¹¹, which possess water-mediated hydrogen bonds. In mPlum¹² and its variant mPlum/E16Q¹³, the N-acylimine oxygen of the chromophore forms direct hydrogen bonds with the side chain functionalities of Glu16 and Gln16, respectively¹³. A hydrogen bond between the protonated Glu215 carboxyl group and the imidazolinone ring nitrogen was proposed to account for the

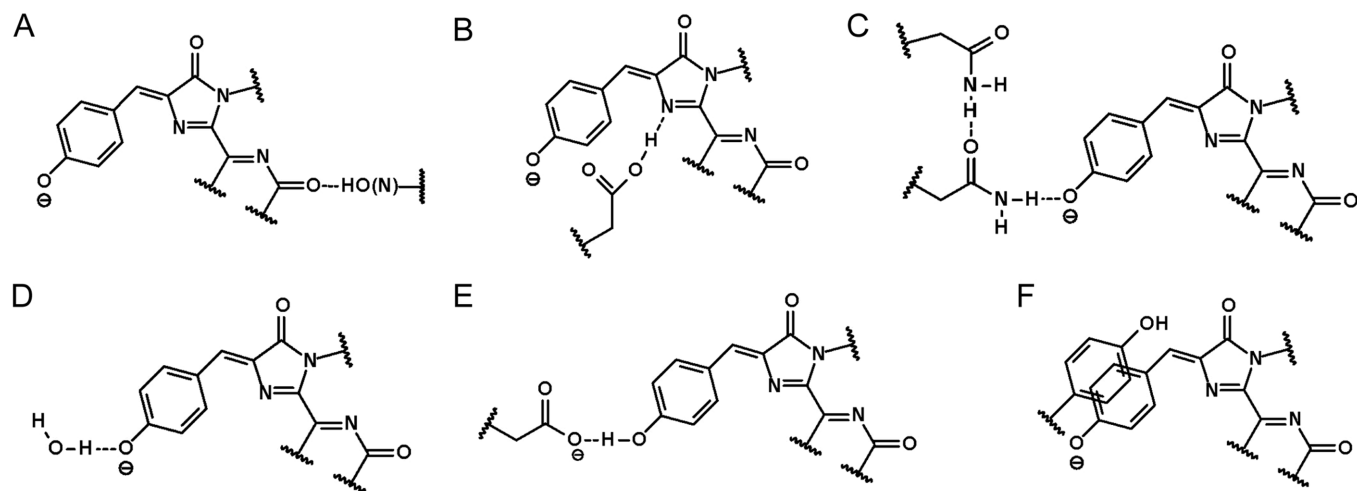


Figure 1 | Interactions between the DsRed-like chromophore and its immediate environment showed to account for a bathochromic shift in the red and far-red FPs. (A) A hydrogen bond between the N-acylimine oxygen and a water molecule or a side chain of an amino acid found in mNeptune, eqFP650, eqFP670, mRojoA, mRouge, mPlum, and mPlum/E16Q. (B) A hydrogen bond between the protonated Glu215 carboxyl group and the imidazolinone ring nitrogen found in mCherry and Strawberry. (C) A network of hydrogen bonds between the phenolate group of the chromophore and amides of Asn143 and Asn158 found in eqFP670. (D) A hydrogen bond between the phenolate groups of the chromophore and a water molecule found in GmKate and eqFP670. (E) A hydrogen bond between the *p*-hydroxyphenyl group of the chromophore and the glutamate carboxyl group found in LSSmKate1. (F) A π - π -stacking interaction between the phenolate group of the chromophore and the *p*-hydroxyphenyl group of Tyr found in mRojoA.

red-shifted fluorescence of mCherry and mStrawberry (Figure 1B)¹⁴. Termination of this hydrogen bond upon the carboxyl group deprotonation resulted in a 15 nm hypsochromic shift.

Other interactions attributed to bathochromic shifts in the GFP-like proteins occur in the proximity to the *p*-hydroxyphenyl site of the chromophore. For example, in eqFP670, the side chains of Asn143 and Asn158 are positioned in the plane of the chromophore and create an extensive network of hydrogen bonds between the chromophore and the protein matrix (Figure 1C)¹⁰. These interactions were shown to cause an almost 20 nm red shift of excitation/emission of eqFP670 compared to its precursor. In GmKate, at pH 10.0 the *p*-hydroxyphenyl group of the chromophore forms only one water-mediated hydrogen bond with the protein scaffold (Figure 1D)⁷. The mechanism of bathochromic shift of GmKate could be similar to that reported for mNeptune, with the only difference being that a water molecule interacts with the N-acylimine oxygen in the latter case. One additional example of chromophore-protein interactions strongly influencing spectral properties is an excited state proton transfer (ESPT) relay¹⁵ found in the large Stokes shift red FPs^{6,16,17}. Several different proton pathways supporting ESPT in GFP-like proteins have been identified, the shortest one being a hydrogen bond between the *p*-hydroxyphenyl group of the chromophore and the glutamate carboxyl group found in LSSmKate1 (Figure 1E)^{6,18}.

Other than hydrogen bond interactions resulting in the bathochromic shift are π - π -stacking interactions. Phe197 in mRojoA and mRojoB, which participates in a π - π -stacking interaction with the chromophore phenolate group, similar to that observed in YFP¹⁹, was demonstrated to red-shift both excitation and emission wavelengths (Figure 1F)¹¹. The same interaction has been suggested to contribute to the far-red fluorescence in mGrapes⁸, E2-Crimson²⁰, and TagRFP657 (Ref. 21). Apparently, more than one mechanism of bathochromic shift described above can occur in a single protein. For instance, the interactions shown in Figure 1A, C and D were found in eqFP670, and the interactions shown in Figure 1A and F were found in mRojoA and mRojoB.

Most GFP-like FPs are characterized by a small Stokes shift of 10–45 nm and the independence of emission maximum on excitation wavelength within the absorption band according to Vavilov's law. Typically FPs exhibit no or very moderate dependence of fluorescence

excitation and emission maxima on pH. Here we report the biochemical, spectroscopic, and structural analysis of a near-infrared FP variant, named TagRFP675, which exhibits unusual spectral characteristics. The TagRFP675 variant was generated by means of directed molecular evolution²² of mKate²³. At physiological pH, TagRFP675 has fluorescence excitation and emission maxima at 598 nm and 675 nm, respectively, that makes it the most red-shifted protein of the GFP-like family. Along with near-infrared fluorescence and an extended Stokes shift of almost 77 nm at neutral pH, the excitation and emission maxima of TagRFP675 are characterized by strong pH dependence. In addition, its emission maximum is strongly dependent on the excitation wavelength (more than 10 nm shift at room temperature). The X-ray structures of TagRFP675, determined at pH 4.5 and 8.0 at resolutions of 2.3 Å, revealed an extensive network of hydrogen bond interactions between the chromophore and the protein matrix. We performed analysis of the time-resolved spectroscopic and biochemical properties of TagRFP675 and its mutants to address a possible role of the revealed chromophore-protein interactions in the near-infrared emission, extended Stokes shift, and pH-dependence of TagRFP675 fluorescence. We next studied inhomogeneous broadening of electronic spectra of TagRFP675, which appears in the dependence of the emission maximum on the excitation wavelength known as a red-edge effect^{24–26}. Lastly, we demonstrated the use of TagRFP675 as a protein tag in live mammalian cells and in dual-color cell imaging with another near-infrared FP developed from bacterial phytochrome, iRFP²⁷.

Results

Characterization of a near-infrared derivative of mKate. During molecular evolution of a red FP mKate²³, in attempt to develop further far-red shifted probes with improved brightness, we found a near-infrared variant with unusual spectral characteristics. It exhibited strong dependence of emission maximum on the excitation wavelength and pH of buffer and its fluorescence is characterized by an extended Stokes shift of almost 77 nm. The selected variant, named TagRFP675, contained seven amino acid substitutions as compared to the parental mKate, namely M41Q, F80W, S143N, L147M, S158N, D159Y, and N173S (Supporting Figure 1). We decided to study the spectral and biochemical properties of TagRFP675 in more detail.

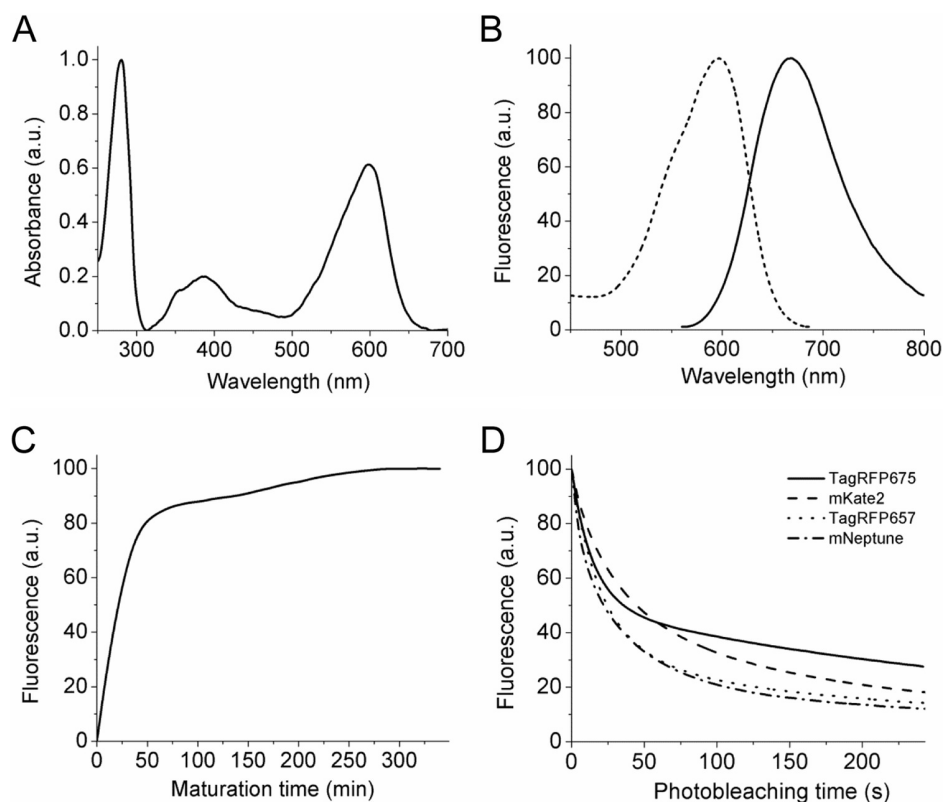


Figure 2 | Properties of purified TagRFP675 at pH 7.4. (A) Absorbance spectrum. (B) Fluorescence excitation (dashed line) and emission (solid line) spectra. (C) Maturation kinetics at 37°C. (D) Photobleaching kinetics of TagRFP675 (solid line), mKate2 (dashed line), TagRFP657 (dotted line), and mNeptune (dashed-dotted line). The curves were normalized to absorbance spectra and extinction coefficients of the proteins, spectrum of an arc lamp, and transmission of a photobleaching filter.

First we characterized the TagRFP675 purified protein at pH 7.4 in PBS. The absorbance spectrum demonstrated a first absorption singlet band with maximum at 598 nm and a second absorption singlet band S2–S0 at 360–390 nm, which is typical for all red FPs (Figure 2A). Excitation of TagRFP675 at 598 nm yielded fluorescence with maximum at 675 nm (Figure 2B and Table 1). The molar extinction coefficient for TagRFP675 measured using alkaline

denaturation assay was $46,000 \text{ M}^{-1}\text{cm}^{-1}$, with quantum yield of 0.08, resulting in a brightness that is 15% of mKate2 (Table 1)²⁸. TagRFP675 had rapid chromophore maturation at 37°C with maturation half-time of 25 min (Figure 2C). To determine photostability of TagRFP675, we performed photobleaching experiments using wild-field illumination. The normalized photostability for TagRFP675 was 1.4-fold and 1.6-fold higher than that for

Table 1 | Properties of TagRFP675 and far-red FPs with emission maxima above 630 nm

Protein	Excitation/emission maxima, nm	Quantum yield	Extinction coefficient, $\text{M}^{-1} \text{cm}^{-1}$	Brightness relative to mKate2, %	pKa	Photo-bleaching half-time, s	Maturation half-time at 37°C, min	Stokes shift, nm
mPlum	590/649	0.10 ^a	22,000 ^a	9	5.5 ^c	42 ^b	96 ^c	59
		0.10 ^c	29,300 ^c	12		31 ^c		
mRojoA	597/633	0.02	48,000	3.8	5.4	ND	ND	36
E2-Crimson	605/646	0.12	58,500	28	4.5	26	26	41
mKate	588/635	0.28 ^b	31,500 ^b	35	6.5 ^a	71 ^b	75 ^a	47
mKate2	588/633	0.40 ^b	62,500 ^b	100	6.5 ^c	23 ^c	48 ^c	45
		0.39 ^c	56,400 ^c	88	6.5 ^d	44 ^d		
GmKate ^e	598/646	0.045	25,800	4.6	9.5	ND	ND	48
mNeptune	600/650	0.20	67,000 ^f	54	5.4 ^f	22	35	50
			59,000	47	5.7			
eqFP650	592/650	0.24	65,000	62	5.7	ND	ND	58
eqFP670	605/670	0.06	70,000	17	4.5	ND	ND	65
TagRFP657	611/657	0.26	34,000	35	5.0	25	125	46
TagRFP675	598/675	0.08	46,000	15	5.7	35	25	77

^aData from ref. 23.

^bData from ref. 28.

^cData from ref. 20.

^dMeasured for mKate2 in our lab using the same conditions as for TagRFP675.

^eMeasured at pH 9–11, data from ref. 7.

^fData from ref. 8. ND, not determined.



TagRFP657 and mNeptune, respectively, and 1.25-fold lower than that for enhanced mKate variant, mKate2 (Figure 2D).

To test the suitability of TagRFP675 as a fluorescent tag for cellular proteins we fused it to β -actin, paxillin, keratin, histone 2B, vimentin, and α -actinin. The fusion constructs were expressed in HeLa cells and imaged under epifluorescence wide-field microscope. All fusions incorporated well into the endogenous structures, indicating that TagRFP675 does not disrupt the function of the fusion partners (Figure 3A). To verify the utility of TagRFP675 as a partner for dual-color imaging in conjugation with more red shifted near-infrared FPs derived from bacterial phytochromes²⁹, the histone 2B-TagRFP675 fusion construct was co-expressed with the bacteriophytochrome derived near-infrared iRFP²⁷ targeted to mitochondria in live HeLa cells (Figure 3B). Two sets of filters allowed visualization of TagRFP675 with iRFP without cross-talk between the channels, for the first time demonstrating a possibility of the dual-color imaging in the near-infrared range of spectrum using genetically-encoded FPs only.

Structural studies of TagRFP675. To explore the molecular origin of the far-red fluorescence of TagRFP675, we determined its crystal structures at two different pH values. The crystal structures of TagRFP675 in the low-fluorescent state (pH 4.5) and the high-fluorescent state (pH 8.0), were determined at 2.3 Å resolution. Diffraction from the TagRFP675 crystals is consistent with the space group P6₁22 with unit cell dimensions $a = b = 105.9$ Å, $c = 218.0$ Å, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$. The crystal asymmetric unit contains two protein chains with the principal axes of each β -barrel at about 120° relative to each other, when viewed along crystallographic c -axis. Electron density is well defined for the residues 3-229 of the subunit A, and residues 3-224 of the subunit B. A *cis*-configuration of the chromophore hydroxyphenyl ring predominates in all molecules of the asymmetric units. The hydroxyphenyl group of the chromophore has an exclusively *cis*-configuration in the A molecule at both pH values; and a mixture of *cis*- and *trans*-configurations in the

molecule B (Supporting Figure 2). Some surrounding side chains also adopt multiple conformations to accommodate chromophore movements. Superposition of the high and low fluorescent state structures resulted in an r.m.s. deviation of only 0.25 Å (0.61 Å) for C α (all) atoms, which is at the level of overall coordinate error at this resolution. The positions of the side chains in the chromophore environments of TagRFP675 also superimpose well with the parental protein mKate, with the exception of residues introduced during mutagenesis (rmsd 0.54 (1.07) Å for C α (all) atoms).

The refined electron density maps for both structures revealed a conventional DsRed-like chromophore formed by the Met63-Tyr64-Gly65 tripeptide. However, the density looks slightly bulkier than one would assume for completely planar conjugated bond system, suggesting traces of immature chromophore (with sp^3 geometry at the C α of Met63), the occupancy of which is difficult to quantify at the current resolution. The *cis*-chromophores of TagRFP675 at both pHs, exhibit noticeable but not severe deviations from the coplanarity between the hydroxyphenyl and imidazolidinone rings. For the *cis*-chromophore in high fluorescent state, the values of χ_1 and χ_2 torsion angles (i.e., around the C α -C β and C β -C γ bonds of tyrosine) are $-3 \pm 5^\circ$ and $-30 \pm 5^\circ$ (average over two subunits), respectively. For the *trans*-chromophore in the B monomer torsion angles are $167 \pm 5^\circ$ and $12 \pm 5^\circ$. For the *cis*-chromophore in low fluorescent state, these angles are $1 \pm 5^\circ$ and $-35 \pm 5^\circ$, respectively. For the *trans*-chromophore in the B monomer torsion angles are $162 \pm 5^\circ$ and $28 \pm 5^\circ$.

Detailed comparison of the structures for TagRFP675 with that of the parental mKate revealed the presence of novel interactions within the active site of TagRFP675. One set of the interactions increased the polarity of the N-acylimine binding site of the chromophore due to introduction of the several hydrogen bonds. One of the hydrogen bonds is formed between the side chain of Gln41 and the oxygen of N-acylimine group of the chromophore. This hydrogen bond was refined to distances of 2.9 Å at pH 8.0, and 2.8 Å at pH 4.5. The amide group of Gln41 buried inside of the β -barrel of TagRFP675 is also positioned to hydrogen bond to the hydroxyl of Ser28 (Figure 4A). An additional hydrogen bond, with length 3.1 Å, is formed between the oxygen of N-acylimine group and a water molecule. The hydrogen bond network formed by the N-acylimine group of the chromophore, the Gln41 and Ser28 side chains and the water molecule is practically identical in all subunits at both pHs. This is the first instance in which the formation of two hydrogen bonds with the chromophore N-acylimine group has been reported for any red FP.

Other unique interactions involve the side of the hydroxyphenyl moiety of the chromophore and side chains of Asn143, Asn158 and Arg197. The hydrogen networks formed between these amino acids and the chromophore in low and high fluorescent states are quite distinct (Figure 4B–E). Moreover, side chains of Asn143, Asn158 and Arg193 adopt different conformations in the A and B subunits. In the high fluorescent state in chain A, the hydrogen bond between the chromophore and the Asn143 side chain is mediated by a water molecule. However, in chain B, this hydrogen bond is absent, disrupting the interaction of the phenolate group of the chromophore with the protein matrix. In chain B, the amide group of Asn158 forms a rather short hydrogen bond (2.2 Å) with the tyrosine hydroxyl group of the *trans*-chromophore. In both subunits, the Arg197 side chain adopted two conformations (A and B, with occupancies 0.6 and 0.4, respectively; Figure 4B and C; Figure 4B and C). In the A conformation, Arg197 forms hydrogen bonds with Asn143 and Asn158, which is not the case for the B conformation. For both conformations, the Arg197 guanidinium group is located about 3.5 Å away from the tyrosine ring of the chromophore and does not seem to participate in a cation- π interaction with the chromophore. In the low-fluorescent state, the phenolate group of the *cis*- and *trans*-chromophore directly forms strong hydrogen bonds with the side chains of Asn143 and Asn158, respectively (Figure 4D and E). In the A

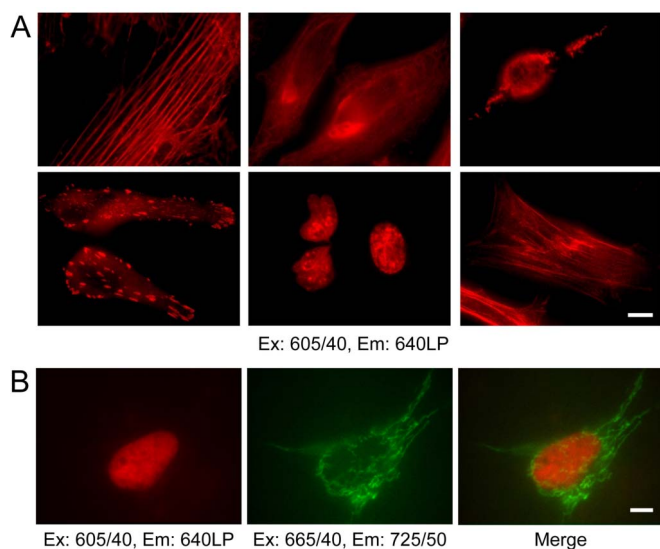


Figure 3 | Near-infrared fluorescence microscopy of live HeLa cells expressing fusion proteins. (A) Cells (from left to right) transfected with the TagRFP675- β -actin, keratin-TagRFP675, vimentin-TagRFP675, paxillin-TagRFP675, histone-2B-TagRFP675, and α -actinin-TagRFP675 are shown. (B) The histone-2B-TagRFP675 fusion protein was co-expressed with the iRFP protein targeted to mitochondria. Images of the TagRFP675 fusions (red color) were acquired using 605/40 excitation and 640LP emission filters, and images of iRFP labeled mitochondria (green pseudocolor) were acquired using 665/40 excitation and 725/50 emission filters. Scale bars, 10 μ m.

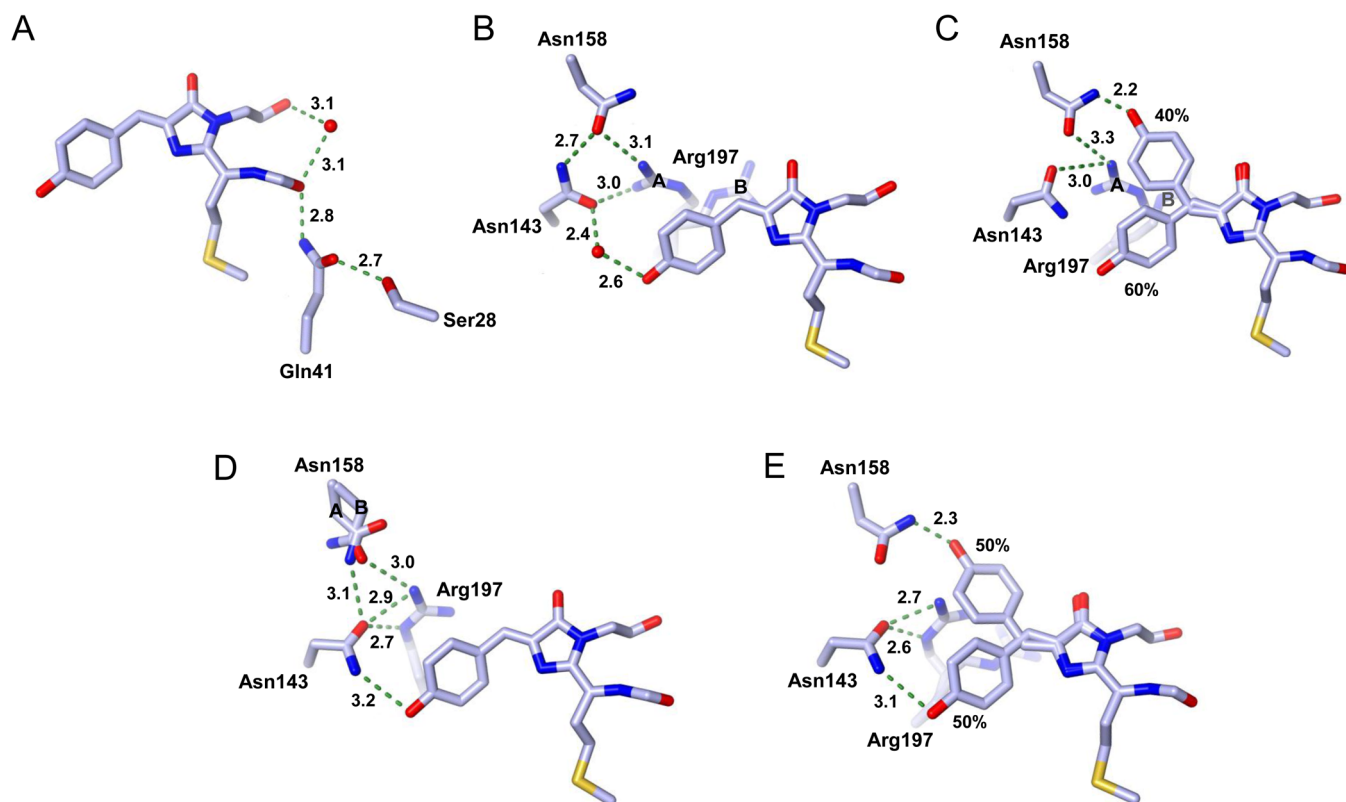


Figure 4 | Molecular structures of the TagRFP675 chromophore and its immediate environments at different pH values. (A) Subunit A, pH 8.0. The positions of the chromophore's N-acylimine group, Ser28, and Gln41 are practically identical in all subunits. (B) Subunit A, pH 8.0. (C) Subunit B, pH 8.0. (D) Subunit A, pH 4.5. (E) Subunit B, pH 4.5. Hydrogen bonds are represented as dashed green lines with the lengths indicated in angstroms, atoms are colored by atom type, and water molecules are shown as red spheres. The occupancy of the *cis*- and *trans*-chromophore was calculated during structure refinement.

subunit, the Asn158 side chain adopts two alternative conformations with occupancies of 0.5/0.5, forming similar hydrogen bonds with Asn143 and Arg197. The Arg197 guanidinium group adopts a conformation different from that of Arg197 in the high fluorescent state. Notably, the main alternations in the structures of TagRFP675 upon pH changing occur around the tyrosine group of the chromophore and involve rearrangements of the hydrogen networks and repositioning of the Arg197 side chain.

Effect of pH on spectral properties of TagRFP675. Since the structures of TagRFP675 were determined at different pH values, we examined the pH-dependence of its spectral properties. Absorbance

and fluorescence of the TagRFP675 protein over pHs ranging from 2.5 to 11.0 were measured at room temperature (Figure 5). The absorption and fluorescence spectra of TagRFP675 were sensitive to changes in pH value. At pH < 5.0, the protein exhibited strong absorption centered at 460 nm, which corresponds to the neutral DsRed-like chromophore (Figure 4A). Excitation at 460 nm yielded weak fluorescence characterized by a broad band with maxima at 616 nm (Supporting Figure 3A). The minor peak of absorption at pH < 5 had a maximum around 570 nm, which corresponds to the 644 nm emission maxima (Supporting Figure 3B). At pH > 5, the absorbance spectra featured a major band with maxima around 580–602 nm corresponding to the anionic DsRed-like chromophore

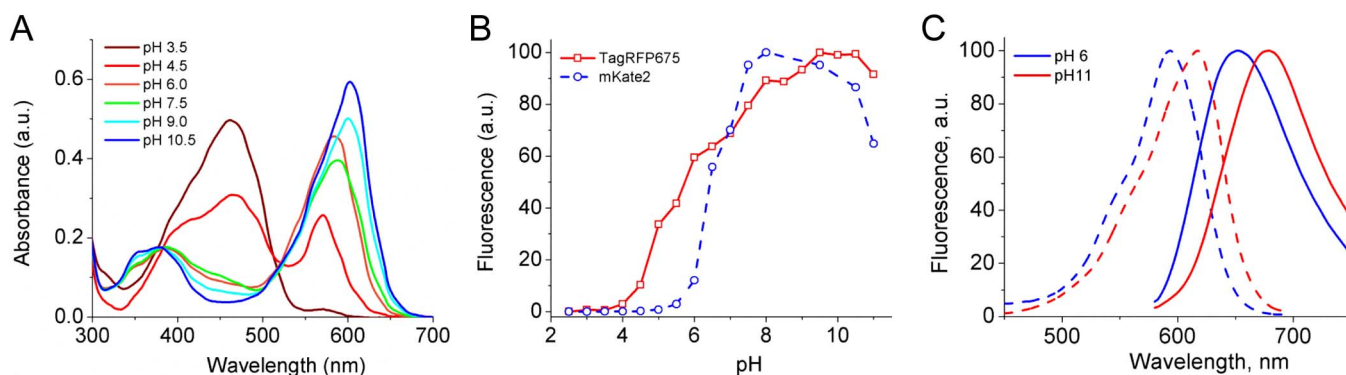


Figure 5 | Effect of pH on the spectral properties of TagRFP675. (A) Absorbance spectra at pH 3.5 (brown line), pH 4.5 (red line), pH 6.0 (orange line), pH 7.5 (green line), pH 9.0 (cyan line), and pH 10.5 (blue line). (B) Equilibrium pH dependence for TagRFP675 (solid red line) and mKate2 (dashed blue line) fluorescence. (C) Fluorescence excitation (dashed line) and emission (solid line) spectra at pH 6.0 (blue line) and pH 11.0 (red line).



(Figure 5A). It should be noted that increasing pH of the TagRFP675 samples resulted in substantial red-shift of the absorbance peak of the anionic chromophore from 570 nm at pH 3.5 to 602 nm at pH 11.0. The fluorescence of TagRFP675 excited at 580 nm was pH stable with an apparent pKa value of 5.7, compared to a pKa of 6.5 for mKate2 (Figure 5B). The intensity of fluorescence emission of the anionic chromophore was maximal at pH 9.5–10.5 and slowly dropped at lower pH values, probably because of the chromophore protonation. Along with absorbance peaks, both fluorescence maxima of excitation and emission spectra of the anionic form substantially shifted to the red upon increasing pH of the buffer (Figure 5C). Similar pH-dependent shifts in the absorbance and fluorescence spectra were also observed for other far-red FPs, such as mPlum¹² and HcRed³⁰. In contrast, the absorbance and emission spectra of some red FPs, such as mCherry and mStrawberry, exhibit substantial blue shifts with increasing pH¹⁴.

Mutagenesis and spectroscopic analysis of TagRFP675. To examine the influence of the residues involved in the interactions with *p*-hydroxyphenyl and N-acylimine sites of the chromophore on far-red fluorescence of TagRFP675, we performed site-directed mutagenesis and compared spectral characteristics of the mutants with those of TagRFP675. The spectral characteristics of the TagRFP675 variants were measured at pH 7.4 in PBS buffer.

First, we introduced Gln41Met and Gln41Pro mutations to understand the role of the hydrogen bond between the N-acylimine group of the chromophore and amide group of Gln41. The Met side chain is incapable of forming a hydrogen bond with the acylimine oxygen of the chromophore. In turn, we suggest that the Gln41Pro substitution can alter the local geometry and creates a cavity near the acylimine site. This cavity is likely to be filled with a water molecule that forms a hydrogen bond to the acylimine oxygen. The TagRFP675/Gln41Met/Pro variants had similar absorbance and excitation maxima as compared to those of the parental protein, whereas their emission maxima exhibited significant blue-shift on 33 and 20 nm, respectively (Figure 5A, B). Both variants had increased quantum yields (Table 2). In contrast to Gln41Met, the Gln41Pro mutation exhibited a substantially decreased extinction coefficient, possibly due to inefficient protein folding caused by limited flexibility of the Pro side chain (Table 2). Thus, both Gln41Met and Gln41Pro mainly affected the Stokes shift of TagRFP675, decreasing it on 29 and 17 nm, respectively (Table 2).

An additional set of mutants was generated to test the influence of Asn143 and Asn158 on the spectroscopic properties of TagRFP675. The Asn143Ser mutation resulted in a dramatic hypsochromic shift of both absorbance and fluorescence bands (Figure 5C, D). In addition, the Asn143Ser mutation results in appearance of additional short wavelength shoulders at both absorbance and fluorescence bands (Figures 5C, D). The quantum yield of TagRFP675/N143S was about 5-fold higher than that of original TagRFP675 (Table 2). Interestingly, the substitution of Asn158 by Lys did not result in any significant alternations of the spectral properties of TagRFP675. The optical

characteristics of TagRFP675/N143K are practically indistinguishable from those of the parental TagRFP675 (Figures 5C, D and Table 2).

In Table 2, the fluorescence lifetimes for TagRFP675, its Q41M, Q41P, N143S, N158K single point mutants, and mKate are presented (see also Supporting Figure 4). The values of fluorescence lifetime for these FPs are well correlated with their quantum yields (Table 2). This behavior can indicate that non-radiative decay of their excited states does not involve any specific photochemical reactions. For TagRFP675/N143S the fluorescence lifetimes were measured at the wavelengths 530, 590, and 650 nm because its instantaneous fluorescence spectra have three peaks (Supporting Figure 4D). The peaks at 530 and 650 nm may be attributed to the fluorescence of neutral and anionic forms of the chromophore, respectively. Elucidation of the nature of the 590 nm peak will require additional analysis.

The instantaneous fluorescence spectra of mKate, TagRFP675 and TagRFP675/Q41M underwent a very minor shift to the blue in a nanosecond time scale (Supporting Figures 4A, B, F). This indicates that their fluorescence spectra are homogeneous on a nanosecond time scale. In contrast, the time-resolved emission spectra of TagRFP675/Q41P and TagRFP675/N143S shifted to the shorter wavelengths in time (Supporting Figures 4D and E). The blue shift of the TagRFP675/Q41P and TagRFP675/N143S emission spectra can indicate evidence of hydrogen bonds between the chromophore and water molecules in the excited state. In case of TagRFP675/Q41P this explanation is in agreement with the suggested above alternation of the TagRFP675 chromophore environment upon introduction of the Q41P mutation.

Temperature dependence and red-edge excitation effects. The steady-state fluorescence spectra of TagRFP675 and its parental protein mKate were measured between room temperature and 77 K in pH 7.4 buffer. In addition, we recorded the fluorescence spectra of TagRFP675/Q41M/P under the same conditions. Upon lowering temperature, the steady-state emission maximum for TagRFP675 shifted from 675 nm at 298 K to 642 nm at 77 K resulting in a 29 nm decrease of Stokes shift (Figure 7A). The TagRFP675/Q41M/P variant also demonstrated a substantial hypsochromic shift of emission upon freezing and their Stokes shift decreased by 17 and 26 nm, respectively (Figure 7B, C). In contrast, mKate exhibited a small shift of about 10 nm in emission maximum under the same conditions (Figure 7D), an effect that is similar to that seen in the conventional GFP-like proteins³¹. In addition, the fluorescence intensity was increased substantially for all the proteins upon freezing (Supporting Figure 5). It was observed that the peaks of the excitation spectra did not demonstrate any significant shift with decreasing temperature for the TagRFP675, TagRFP675/Q41P and mKate proteins. The only exception is TagRFP675/Q41M which exhibited a 9 nm blue shift of its excitation maximum. At 77 K the excitation spectra of the proteins had elements of the vibrational structure at the short wavelength slope, also observed for other chromophores at lowered temperatures. The Stokes shift of TagRFP675 and its mutant TagRFP675/Q41P exhibited strong

Table 2 | Spectroscopic properties of TagRFP675, its mutants, and parental mKate

Protein	Absorbance maximum, nm	Excitation/emission maxima, nm	Quantum yield	Extinction coefficient, M ⁻¹ cm ⁻¹	Fluorescence lifetime, ns	Stokes shift, nm
TagRFP675	598	598/675	0.08	46,000	0.9 (1.0)	77
TagRFP675/Q41M	595	595/643	0.18	46,600	1.4 (1.0)	48
TagRFP675/Q41P	595	595/655	0.12	14,600	1.1 (1.0)	60
TagRFP675/N143S	575	578/590 and 633	0.38	45,000	2.5 (1.0); 3.2 (1.0); 0.8 (0.44), 2.1 (0.56)	55
TagRFP675/N158 K	594	599/670	0.08	45,200	0.8 (1.0)	71
mKate	588	588/635	0.28	31,500	2.5 (1.0)	47

The fluorescence life-time for mKate, TagRFP675, and its Q41M, Q41P, N143S, N158K single-point mutants were measured at wavelengths of 630, 675, 645, 660, and 670 nm, respectively. For N143S the fluorescence life-times were measured at 530, 590, and 650 nm (Supporting Figure 4D). The pre-exponential factors for the life-time fitting are indicated in brackets.

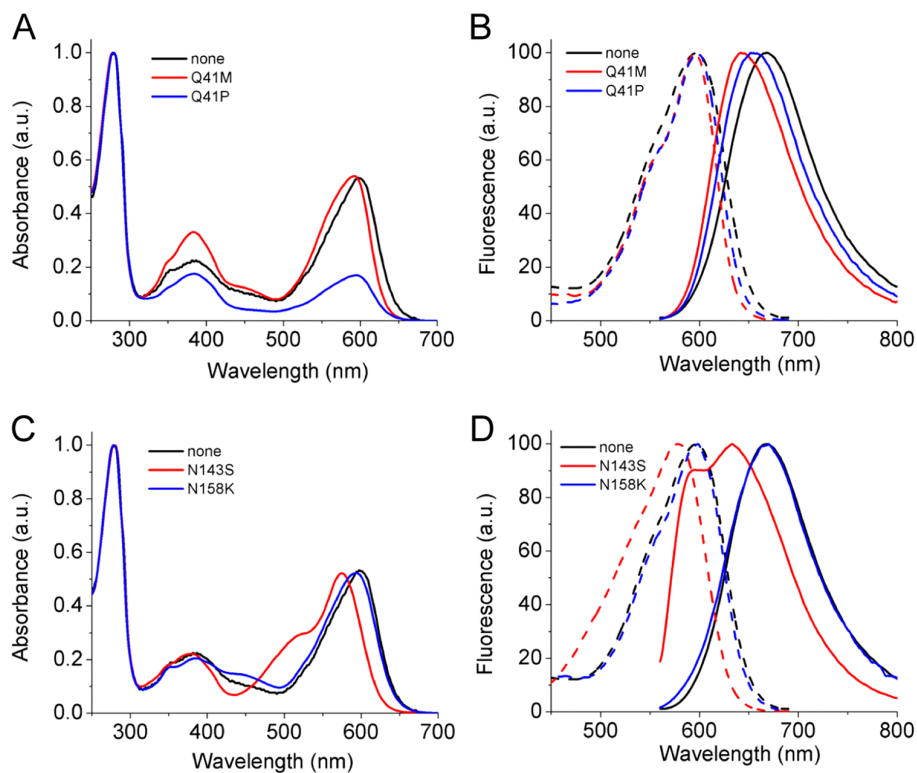


Figure 6 | Spectral properties of TagRFP675 and its mutants at pH 7.4. (A) Absorbance spectra of TagRFP675 (black line), TagRFP675/Q41M (red line), and TagRFP675/Q41P (blue line). (B) Fluorescence excitation (dashed line) and emission (solid line) spectra of TagRFP675 (black line), TagRFP675/Q41M (red line), and TagRFP675/Q41P (blue line). (C) Absorbance spectra of TagRFP675 (black line), TagRFP675/N143S (red line), and TagRFP675/N158K (blue line). (D) Fluorescence excitation (dashed line) and emission (solid line) spectra of TagRFP675 (black line), TagRFP675/N143S (red line), and TagRFP675/N158K (blue line).

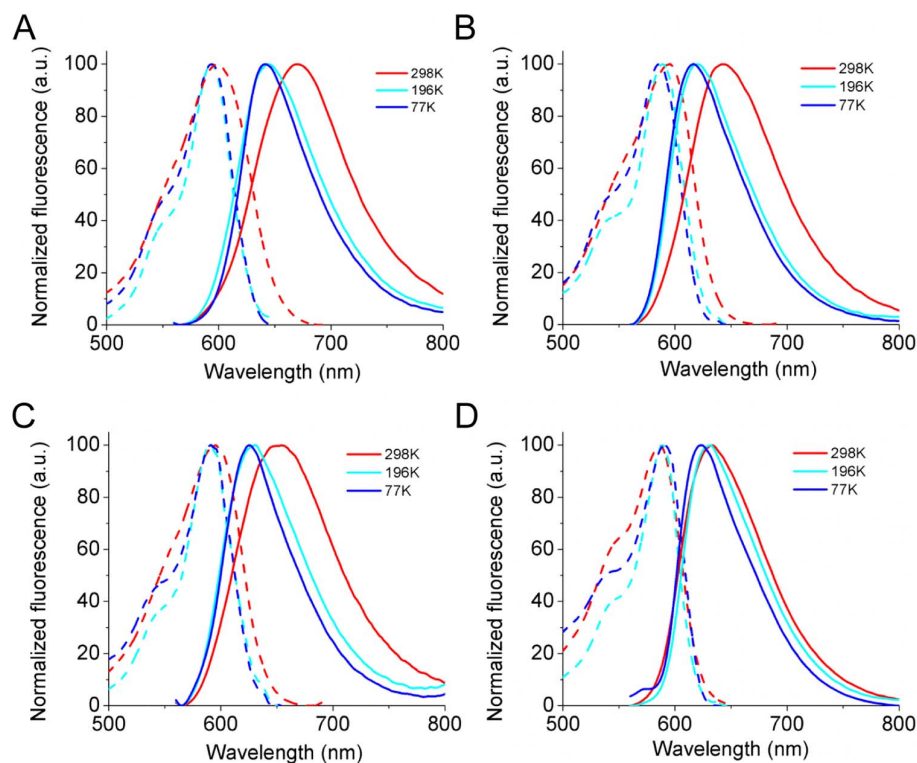


Figure 7 | Effect of temperature on fluorescence spectra of TagRFP675, its variants, and parental mKate. Fluorescence excitation (dashed line) and emission (solid line) spectra of (A) TagRFP675, (B) TagRFP675/Q44M, (C) TagRFP675/Q44P, and (D) mKate at 298 K (red line), 196 K (cyan line), and 77 K (blue line) are shown.



Table 3 | Spectroscopic properties of TagRFP675, its several mutants, and mKate at different temperatures

Protein	298 K			196 K			77 K		
	Ex (nm)	Em (nm)	Stokes shift (nm)	Ex (nm)	Em (nm)	Stokes shift (nm)	Ex (nm)	Em (nm)	Stokes shift (nm)
TagRFP675	598	675	77	596	645	49	594	642	48
TagRFP675/Q41M	595	643	48	589	621	32	586	617	31
TagRFP675/Q41P	595	655	60	592	630	38	591	625	34
mKate	588	633	45	589	632	43	591	623	32

temperature-dependent behavior (Table 3). Similar changes in the fluorescence spectra upon freezing were observed for mPlum and its variant mPlum/E16Q¹³. The emission spectrum of red FPs possessing a large Stokes shift due to ESPT, which also demonstrates strong temperature dependence^{6,17}.

The temperature dependence of the TagRFP675 its mutants fluorescence spectra revealed that their emission maxima also depended on the wavelength of excitation, especially, at long-wavelength edge of absorption spectrum (Figure 8A, B). It is significant that the shape of the fluorescence spectra of TagRFP675 did not change with red-edge excitation (Supporting Figure 6). The magnitude of the red-edge effect depended on temperature, as at 77 K it was larger for TagRFP675 than for TagRFP675/Q41M, which was 33 nm versus 17 nm, respectively.

Discussion

TagRFP675 has a 10 nm red shift in excitation and a 40 nm red shift in emission maxima at pH 7.4 relative to its parental protein mKate. Along with the most red-shifted emission among known GFP-like proteins, TagRFP675 also possesses an extended Stokes shift, about 77 nm. Most GFP-like FPs have small Stokes shifts (10–40 nm), and this has been attributed to the rigidity of the chromophore's environment necessary to exclude non-fluorescent relaxation to a ground state. Increasing flexibility of the region in protein solvating the chromophore may lead to a solvation response in the excited state, which results in a substantial enhancement of the Stokes shift^{12,13}. Based on the X-ray analysis of the TagRFP675 structures and spectroscopic properties, two groups of hydrogen-bond interactions in the chromophore environment appear to be responsible for the observed near-infrared fluorescence and the extended Stokes shift of TagRFP675.

First, in both crystal structures of TagRFP675, the N-acylimine oxygen of the chromophore forms two hydrogen-bonds, one with a water molecule and another with the amide group of the Gln41 side chain (Figure 4A). Very similar hydrogen bonds were reported to

cause bathochromic shifts of fluorescence in other far-red FPs due to stabilization of the chromophore conjugated π -electron system in the excited state (Figure 1A)^{8,10,11,13}. For example, in case of mPlum³² the hydrogen bond between Glu16 carboxylate and the N-acylimine oxygen of the chromophore was shown to be responsible for the solvation response (Table 1)^{12,13}. The solvent response in Plum results in a substantial dynamic Stokes shift producing far-red emission. Substitution of Glu16 with any other amino acid or lowering temperature substantially decreased the Stokes shift of mPlum^{12,13}. Similar to mPlum, site-specific mutagenesis of TagRFP675 confirmed the essential role of the hydrogen bond between the N-acylimine oxygen of the chromophore and the protein matrix, and the quality of the hydrogen bond donor for the near-infrared emission (Table 2). The conclusion on the importance of the quality of the H-bond for near-infrared fluorescence emission was drawn from the observation that TagRFP675/Q41P, capable of forming a hydrogen bond with the N-acylimine oxygen indirectly via a buried water molecule, as suggested above, had an emission maximum 20 nm blue shifted compared to that of TagRFP675. In addition, TagRFP675 and its Q41M,P mutants exhibited the temperature dependence of Stokes shifts very similar to that of mPlum and its respective variants (Table 3)^{12,13}.

Second, an extensive hydrogen-bonding network composed of the p-hydroxyphenyl group of the chromophore, Asn143, Asn158, and a water molecule increases the interaction between the chromophore and the protein matrix. The bathochromic shift of fluorescence in eqFP670 was also attributed to the Asn143 and Asn158 residues, creating a similar hydrogen bond network around the p-hydroxyphenyl group of the chromophore¹⁰. In the A subunit of TagRFP675, the phenolate group of the chromophore forms a hydrogen bond with Asn143 via a water molecule. This kind of hydrogen bond (i.e., water mediated) can add some flexibility to the region of the protein solvating the chromophore, and therefore contribute to the extended Stokes shift of TagRFP675. Substitution of Asn143 with Ser most likely disrupts this hydrogen bond, resulting in significantly altered spectral properties (Table 2 and Figure 6C, D). According to the

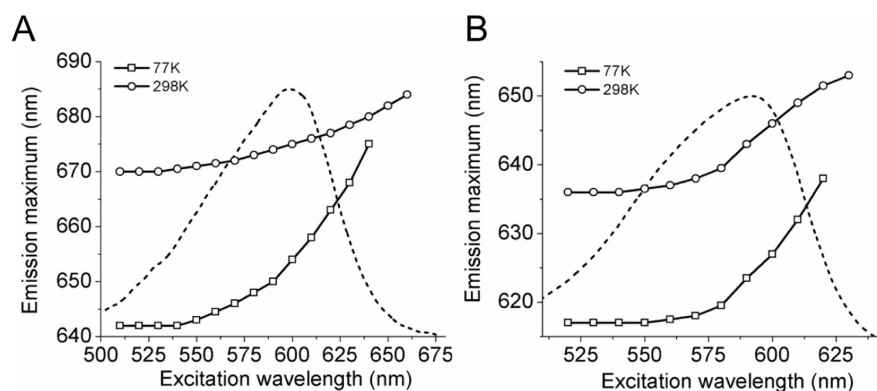


Figure 8 | Dependence of fluorescence maximum on excitation wavelength at 77 K and 298 K for TagRFP675, and TagRFP675/Q41M. (A) Maximum of the emission maximum versus the excitation wavelength of TagRFP675 at 77 K (squares) and 298 K (circles) superimposed on its absorbance spectrum at 298 K (dashed line). (B) Maximum of the emission maximum versus the excitation wavelength of TagRFP675/Q41M at 77 K (squares) and 298 K (circles) superimposed on its absorbance spectrum at 298 K (dashed line).



crystal structures of mKate³³, Neptune⁸ and eqFP650 (ref. 10) the Ser residue at the 143 position always forms a direct hydrogen bond with the chromophore in its *cis*-configuration and does not support extensive hydrogen bonding around the phenolate group of the chromophore. Interestingly, the Asn158Lys substitution had very minor impact on the spectral properties of TagRFP675 (Table 2 and Figure 6C, D). This might be due to the ability of the Lys side chain to form several hydrogen bonds, which supports the existing hydrogen bond network around the chromophore.

The structural analysis of TagRFP675 in its low and high fluorescent states can provide an explanation for the pH-dependent spectral changes. We suggest that the pH-dependence of the TagRFP675 fluorescence and absorbance involves two processes: (i) protonation-deprotonation of the chromophore *p*-hydroxyphenyl group and (ii) isomerization of the side chains of the Asn143, Asn158 and Arg197 resulting in the rearrangements of hydrogen bonds around the *p*-hydroxyphenyl site of the chromophore. It should be noted that the structural data for TagRFP675 do not provide clear evidence for the pH-inducible *cis-trans* chromophore isomerization, as is commonly observed for other red and far-red FPs^{7,30,33}. As mentioned in the Methods part of the manuscript, the crystal of TagRFP675 at pH 8.0 was obtained by stepwise replacement of the solution in the crystal wells with solution containing 3.0 M sodium chloride and 0.1 M Tris, pH 8.5. The color change of the crystal was observed within two minutes, however, longer soaking times were not possible due to gradual decomposition of crystals. The effective pH change can be less than expected due to the high ionic strength of the mother liquor (3.0 M sodium chloride and 0.8 M lithium sulfate) making buffer capacity insufficient. Additionally, the time of buffer exchange at room temperature could be insufficient for reaching the equilibrium between the *cis*- and *trans*-conformations of the chromophore.

We observed very little increase in the total amount of the *cis*-chromophore in the TagRFP675 crystal at pH 8.0 vs. pH 4.5. However, considerable pH-dependent structural alterations in TagRFP675 occur around the tyrosine site of the chromophore, including rearrangements of hydrogen network and pronounced repositioning of the Arg197 side chain. The number and length of the hydrogen bonds formed by the phenolate group of the chromophore and the side chains of Asn143, Asn158, and Arg197 are different for pH 4.5 and pH 8. We suggest that the number and strength of hydrogen bonds are responsible for the additional red shift of excitation and emission spectra with increasing pH due to lowering of the excitation energy by accumulating electron density around the chromophore (Figure 5C).

The slightly sloping curve of the equilibrium pH dependence for TagRFP675 suggests that a multiple step process occurs in response to pH alternation (Figure 5B). In addition to protonation-deprotonation of the chromophore phenolate group, protonation-deprotonation and conformational reorganization of amino acids side chains in the immediate vicinity of the chromophore may be involved.

The results described above unambiguously testify to inhomogeneous broadening of the electronic spectra of TagRFP675. Inhomogeneous broadening in condensed media is caused by fluctuations of energy of intermolecular interactions between a fluorophore and surrounding molecules or their segments³⁴. As a result, the ground and excited electronic states of the chromophore represent a continuum of sublevels, the energies of which depend on the position and orientation of the nearest neighboring molecules or their segments. The inhomogeneous broadening of the electronic spectra is caused mainly by dipole-dipole interactions between chromophore and its environment. The extent of the broadening is determined by the change of the fluorophore dipole moment on excitation, the number and the values of solvent dipoles, and the energies of intermolecular interactions^{26,34}. It was also shown that hydrogen bonding can significantly increase inhomogeneous broadening of electronic spectra of fluorophores^{35,36}. The half-width of an inhomogeneous

broadening function, which describes the distribution of fluorescent centers over 0–0 (pure) electronic transition, is 500–800 cm⁻¹ for polar solutions and is comparable with vibrational broadening²⁶.

In solutions and biological systems, inhomogeneous broadening manifests itself as a dependence of the maximum of fluorescence spectra on the excitation wavelengths. This phenomenon is also known as a red-edge effect^{24–26}. When fluorescence is excited at the long-wavelength shoulder of the absorption spectrum, the emission maximum starts to depend on excitation frequency. The spectral heterogeneity of fluorophores in condensed media is responsible for a number of the spectroscopic effects. For example, during freezing of media, the absorption and fluorescence spectra shift to shorter wavelengths, because spontaneous emission results from non-equilibrium, high energy sublevels. Such blue shifting is larger for emission than for absorption and results in the decrease of the Stokes shift. In addition, there is an increase in fluorescence intensity due to the drawing together of the excited and ground electronic states²⁵. All listed effects take place for TagRFP675 and its mutants at a greater extent than for mKate (Table 3, Figure 6, and Supporting Figure 4), providing solid evidence that inhomogeneous broadening of their electronic spectra takes place in the chromophore's immediate environment.

The dependence of the emission maxima on the excitation wavelength for TagRFP675 and TagRFP675/Q41M at room and liquid nitrogen temperatures is shown in Figure 8. The emission maximum exhibited noticeable red shift through the range of excitation wavelengths, starting even before the absorbance band maximum. The observed dependence can be caused by overlapping of the effect of ground state heterogeneity with the red-edge effect²⁵. The ground state heterogeneity arises from the presence of two fluorescent chromophore forms with close but different absorbance maxima. Indeed, the structure of TagRFP675 revealed the chromophore in two configurations *cis* and *trans*, which are typically characterized by different absorbance and fluorescence maxima⁶.

The change of dipole moment of the chromophore upon transition to the excited state, caused by the charge transfer from donor to acceptor groups³⁷, results in the dynamic Stokes shift of fluorescence spectrum³⁸. For the red and far-red FPs the changes in dipole moment upon chromophore excitation were determined to be from 5 to 6 debye¹². In fact, the time-dependent Stokes shift is the consequence of relaxation of the fluorescent centers over the continuum of configurational sub-levels after excitation. It was shown that for Plum the time of the dynamic Stokes shift is about one hundred picoseconds¹². From our measurements (Supporting Figure 4), it follows that the dynamic Stokes shifts for TagRFP675 and TagRFP675/Q41M are outside the time resolution (0.1 ns) of the set-up used and likely lies in the picosecond time scale. Hence, the fluorescence emitted during dynamic Stokes shift does not distinctly contribute to steady-state fluorescence. This mechanism explains why the position of the emission spectra for both proteins at 298 K is more red shifted than after freezing (T = 77 K) (Figure 7C, D). The bathochromic shift of emission essentially depends on the shape of a red shoulder of absorption band, which is strongly individual for fluorophores. One may easily excite TagRFP675 to result in fluorescence emission longer than 680 nm (Figure 8A).

In this paper we report the near-infrared derivative of mKate, called TagRFP675, which exhibits the substantially extended Stokes shift and the most red-shifted emission among the GFP-like protein family. Our data indicate the complex contribution of several factors in determining the spectral properties of TagRFP675. The structural, biochemical and spectroscopic analysis of TagRFP675 revealed key interactions responsible for the fluorescence red-shift and Stokes shift extension, such as the hydrogen bonds between the protein matrix and the N-acylimine oxygen and phenolate group of the DsRed-like chromophore. The mechanism of the Stokes shift extension described here is different from that revealed in other GFP-like



large Stokes shift red FPs, such as LSSmKate⁶ and mKeima¹⁷. In these FPs the large Stokes shift is caused by ESPT from the chromophore to the carboxylic group of the amino acid side chain, thus changing the protonation state of the chromophore. In contrast, fluorescence excitation of TagRFP675 is not accompanied by changes in the chromophore chemical structure but rather the rearrangement or reorientation of dipole-dipole interactions between the chromophore and the protein matrix. Together with the studies on other far-red FPs, our findings provide the solid basis for rational design strategies to induce further red-shifts in FPs of the GFP-like family. These future FP variants will allow multicolor imaging of mammalian cells, tissues, and, possibly, whole animals in the near-infrared part of spectrum. The atomic coordinates and structure factors of TagRFP675 have been deposited in the Protein Data Bank (entry 4KGE for pH 4.5 and entry 4KGF for pH 8.0).

Methods

Mutagenesis, expression, and protein purification. The gene encoding the far-red FP mKate was PCR (polymerase chain reaction) amplified as a *Bam*HI-*Hind*III fragment and inserted into the *Bgl*II-*Hind*III sites of the pBAD/His-B vector (Invitrogen). Site-specific mutations were introduced using QuikChange Mutagenesis kit (Stratagene). Random mutagenesis was performed with a GeneMorph II Random Mutagenesis Kit (Stratagene) using conditions that resulted in a mutation frequency of up to 16 mutations per 1,000 base pairs. After mutagenesis, a mixture of the mutants was electroporated into LMG194 host cells (Invitrogen). Directed molecular evolution of mKate to obtain a red-shifted FP was performed as described²¹.

For spectroscopic studies, mKate and its mutants with polyhistidine tag were expressed in LMG194 bacterial cells grown in RM medium supplemented with 0.002% arabinose overnight at 37°C and then purified using a Ni-NTA agarose (Qiagen).

For crystallization, PCR-amplified *Bgl*II/*Eco*RI fragments encoding TagRFP675 were cloned into a pBAD/His-B vector (Invitrogen), modified by shortening the N-terminal polyhistidine tag to the MGSHHHHHGRS-amino acid sequence. TagRFP675 was expressed in LMG194 bacterial host (Invitrogen) in RM 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 NaH₂PO₄, 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 NaH₂PO₄, pH 7.5.

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). The 2 nm slits on excitation was used to measure dependence of emission maximum on the excitation wavelength. For absorbance measurements, a Hitachi U-2000 spectrophotometer was used. LB/agar Petri dishes with fluorescent colonies 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 NaH₂PO₄, 300 mM NaCl for pH 5.0–11.0).

Photobleaching measurements with purified proteins were performed using microdroplets of freshly purified FPs at 1.0 mg/ml in PBS under mineral oil using an Olympus IX81 inverted microscope equipped with a 200 W Me-Ha arc light source, 100 × 1.4 NA oil immersion lens, and standard 605/40 nm excitation filter (Chroma). Single drops were found at low light levels (0.15% ND filter), the ND filter removed, and the protein drops photobleached. The procedure took into account the spectral output of the source, transmission profile of the excitation filter and dichroic mirror, and absorbance spectra of the proteins.

Time-resolved fluorescence measurements were performed on a laser spectrofluorimeter. Briefly, the set-up consisted of a pulsed TEA nitrogen laser providing excitation at 337.1 nm by pulses having the FWHM of about 1.5 ns, double grating monochromator (dispersion 0.65 nm) for emission wavelength selection, fast-response (1.0 GHz bandwidth) photomultiplier, and boxcar integrator. The spectrofluorimeter provided measurements of instantaneous fluorescence spectra with the time gate of 0.1 ns and fluorescence decay curves for the wavelength of choice. This method was employed for the analysis of fluorescence decay curves for conventional one- three-exponential fitting based on iterative deconvolution routine. In the method, the “true” fluorescence decay $F(t)$ is presented as a sum of exponentials

$$F(t) = \sum_{i=1}^N f_i \exp(-t/\tau_i).$$

The number of exponents $N = 1 \dots 3$ and the decay analysis procedure for both the pre-exponential factors f_i and fluorescence lifetimes were varied to provide the optimum fitting to experimental data.

Protein crystallization. Prior to crystallization, TagRFP675 was dialyzed against 10 mM NaH₂PO₄, pH 7.5, and concentrated to 24 mg/ml using Centricon-3000 microconcentrators. Diffraction quality crystals were grown using the sitting drop vapor diffusion method by mixing 1 μl of protein and 1 μl of reservoir solution and equilibrating samples against corresponding reservoir solution. Reservoir solution contained 3.0 M sodium chloride and 0.1 M sodium acetate, pH 4.5. Crystals reached maximal size in one week. Diffraction from the TagRFP675 crystals was consistent with the space group P6₂22 with unit cell dimensions $a = b = 72.18 \text{ \AA}$, $c = 226.56 \text{ \AA}$, $\alpha = \beta = \gamma = 90^\circ$. Each crystal form contains two protein chains in the asymmetric unit. Crystals of TagRFP675 near its fluorescence optimum were obtained by stepwise replacement of the solution in the crystal wells, with solution containing 3.0 M sodium chloride and 0.1 M Tris, pH 8.5. Color change was observed within two minutes; longer soaks were not possible due to gradual decomposition of crystals.

X-ray diffraction data collection and crystallographic refinement. Crystals of TagRFP675 with dimensions $0.2 \times 0.2 \times 0.3 \text{ mm}^3$ were mounted in cryo-loops directly from the crystallization droplet and flash-cooled in liquid nitrogen. Prior to freezing, each crystal droplet was supplemented with 0.8 M lithium sulfate as cryo-protectant. Diffraction data were collected on a Quantum 315 CCD detector (Area Detector Systems Corporation, Poway) with 1.08 Å wavelength radiation on the X29A beamline (National Synchrotron Light Source, Brookhaven). Intensities were integrated using the program HKL2000 and reduced to amplitudes using the program SCALEPACK2MTZ (see Supporting Table 1 for statistics)^{39,40}. The structures were determined using the molecular replacement method with PHASER⁴¹. Model building and refinement were performed with the programs REFMAC and COOT^{40,42}. The quality of the final structures was verified with composite omit maps, and the stereochemistry was checked with the program MOLPROBITY⁴³. The LSQKAB and SSM algorithms were used for structural superimpositions^{40,44}. Structural figures were prepared using PyMOL (Schrödinger).

Mammalian plasmids and cell culture. To generate vectors encoding fusion constructs, the PCR-amplified TagRFP675 gene fragments and the respective mTagBFP⁴⁵ fusion constructs were digested with specific enzymes and ligated together. To prepare TagRFP675 N-terminal fusions, the following digests were performed: chicken paxillin with *Bam*HI and *Not*I, human keratin with *Eco*RI and *Not*I, human histone-2B with *Bam*HI and *Not*I, human vimentin with *Bam*HI and *Not*I, and human α -actinin with *Eco*RI and *Not*I. To prepare a pTagRFP657- β -actin plasmid, the TagRFP675 gene was PCR amplified as a *Nhe*I-*Bgl*II fragment and swapped with the EGFP gene in the pEGFP- β -actin plasmid (Clontech). HeLa cell line was grown in DMEM containing 10% FBS, penicillin-streptomycin and 2 mM glutamine (HyClone). Plasmid transfections were performed using an Effectene reagent (Qiagen).

1. Wu, B., Piatkevich, K. D., Lionnet, T., Singer, R. H. & Verkhusha, V. V. Modern fluorescent proteins and imaging technologies to study gene expression, nuclear localization, and dynamics. *Curr Opin Cell Biol* **23**, 310–317 (2011).
2. Piatkevich, K. D. & Verkhusha, V. V. Guide to red fluorescent proteins and biosensors for flow cytometry. *Methods Cell Biol* **102**, 431–461 (2011).
3. Chudakov, D. M., Matz, M. V., Lukyanov, S. & Lukyanov, K. A. Fluorescent proteins and their applications in imaging living cells and tissues. *Physiol Rev* **90**, 1103–1163 (2010).
4. Subach, F. V. & Verkhusha, V. V. Chromophore transformations in red fluorescent proteins. *Chem Rev* **112**, 4308–4327 (2012).
5. Piatkevich, K. D. & Verkhusha, V. V. Advances in engineering of fluorescent proteins and photoactivatable proteins with red emission. *Curr Opin Chem Biol* **14**, 23–29 (2010).
6. 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* **132**, 10762–10770 (2010).
7. Wang, Q. *et al.* Molecular mechanism of a green-shifted, pH-dependent red fluorescent protein mKate variant. *PLoS one* **6**, e23513 (2011).
8. Lin, M. Z. *et al.* Autofluorescent proteins with excitation in the optical window for intravital imaging in mammals. *Chemistry & biology* **16**, 1169–1179 (2009).
9. Shcherbo, D. *et al.* Near-infrared fluorescent proteins. *Nat Methods* **7**, 827–829 (2010).
10. Pletnev, S. *et al.* Structural basis for bathochromic shift of fluorescence in far-red fluorescent proteins eqFP650 and eqFP670. *Acta crystallographica. Section D, Biological crystallography* **68**, 1088–1097 (2012).
11. Chica, R. A., Moore, M. M., Allen, B. D. & Mayo, S. L. Generation of longer emission wavelength red fluorescent proteins using computationally designed libraries. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 20257–20262 (2010).
12. Abbyad, P., Childs, W., Shi, X. & Boxer, S. G. Dynamic Stokes shift in green fluorescent protein variants. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 20189–20194 (2007).
13. Shu, X., Wang, L., Colip, L., Kallio, K. & Remington, S. J. Unique interactions between the chromophore and glutamate 16 lead to far-red emission in a red fluorescent protein. *Protein science: a publication of the Protein Society* **18**, 460–466 (2009).



14. Shu, X., Shaner, N. C., Yarbrough, C. A., Tsien, R. Y. & Remington, S. J. Novel chromophores and buried charges control color in mFruits. *Biochemistry* **45**, 9639–9647 (2006).
15. Meech, S. R. Excited state reactions in fluorescent proteins. *Chemical Society reviews* **38**, 2922–2934 (2009).
16. Violot, S., Carpentier, P., Blanchoin, L. & Bourgeois, D. Reverse pH-dependence of chromophore protonation explains the large Stokes shift of the red fluorescent protein mKeima. *Journal of the American Chemical Society* **131**, 10356–10357 (2009).
17. Henderson, J. N. *et al.* Excited state proton transfer in the red fluorescent protein mKeima. *Journal of the American Chemical Society* **131**, 13212–13213 (2009).
18. Piatkevich, K. D. *et al.* Monomeric red fluorescent proteins with a large Stokes shift. *Proc Natl Acad Sci U S A* **107**, 5369–5374 (2010).
19. Wachter, R. M., Elsliger, M. A., Kallio, K., Hanson, G. T. & Remington, S. J. Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein. *Structure* **6**, 1267–1277 (1998).
20. Strack, R. L. *et al.* A rapidly maturing far-red derivative of DsRed-Express2 for whole-cell labeling. *Biochemistry* **48**, 8279–8281 (2009).
21. Morozova, K. S. *et al.* Far-red fluorescent protein excitable with red lasers for flow cytometry and superresolution STED nanoscopy. *Biophys J* **99**, L13–15 (2010).
22. Subach, F. V., Piatkevich, K. D. & Verkhusha, V. V. Directed molecular evolution to design advanced red fluorescent proteins. *Nat Methods* **8**, 1019–1026 (2011).
23. Shcherbo, D. *et al.* Bright far-red fluorescent protein for whole-body imaging. *Nat Methods* **4**, 741–746 (2007).
24. Mukherjee, S. & Chattopadhyay, A. Wavelength-selective fluorescence as a novel tool to study organization and dynamics in complex biological systems. *Journal of Fluorescence* **5**, 237–246 (1995).
25. Demchenko, A. P. Site-selective Red-Edge effects. *Methods in enzymology* **450**, 59–78 (2008).
26. Demchenko, A. P. The red-edge effects: 30 years of exploration. *Luminescence: the journal of biological and chemical luminescence* **17**, 19–42 (2002).
27. Filonov, G. S. *et al.* Bright and stable near-infrared fluorescent protein for in vivo imaging. *Nat Biotechnol* **29**, 757–761 (2011).
28. Shcherbo, D. *et al.* Far-red fluorescent tags for protein imaging in living tissues. *Biochem J* **418**, 567–574 (2009).
29. Piatkevich, K. D., Subach, F. V. & Verkhusha, V. V. Engineering of bacterial phytochromes for near-infrared imaging, sensing, and light-control in mammals. *Chemical Society reviews* (2013).
30. Mudalige, K. *et al.* Photophysics of the red chromophore of HcRed: evidence for cis-trans isomerization and protonation-state changes. *The journal of physical chemistry. B* **114**, 4678–4685 (2010).
31. Creemers, T. M., Lock, A. J., Subramaniam, V., Jovin, T. M. & Volker, S. Three photoconvertible forms of green fluorescent protein identified by spectral hole-burning. *Nature structural biology* **6**, 557–560 (1999).
32. Wang, L., Jackson, W. C., Steinbach, P. A. & Tsien, R. Y. Evolution of new nonantibody proteins via iterative somatic hypermutation. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 16745–16749 (2004).
33. Pletnev, S. *et al.* A crystallographic study of bright far-red fluorescent protein mKate reveals pH-induced cis-trans isomerization of the chromophore. *The Journal of biological chemistry* **283**, 28980–28987 (2008).
34. Nemkovich, N. A., Rubinov, A. N. & Tomin, V. I. Inhomogeneous Broadening of Electronic Spectra of Dye Molecules in Solutions. *Topics in Fluorescence Spectroscopy* **2**, 367–428 (2002).
35. Bushuk, B. A., Rubinov, A. N. & Stupak, A. P. Inhomogeneous broadening of spectra of dye solutions due to intermolecular hydrogen bonding. *Journal of Applied Spectroscopy* **47**, 1251–1254 (1987).
36. Webber, N. M. & Meech, S. R. Electronic spectroscopy and solvatochromism in the chromophore of GFP and the Y66F mutant. *Photochemical & photobiological sciences* **6**, 976–981 (2007).
37. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, Springer (2006).
38. Nemkovich, N. A., Detert, H. & Schmitt, V. Localized excitation effect on dipole moments of oligophenylenevinyls in their excited Franck-Condon state. *Chemical Physics* **378**, 37–41 (2010).
39. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* **276**, 307–326 (1997).
40. Collaborative Computational Project, N. The CCP4 suite: programs for protein crystallography. *Acta crystallographica. Section D, Biological crystallography* **50**, 760–763 (1994).
41. Storoni, L. C., McCoy, A. J. & Read, R. J. Likelihood-enhanced fast rotation functions. *Acta crystallographica. Section D, Biological crystallography* **60**, 432–438 (2004).
42. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* **60**, 2126–2132 (2004).
43. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica. Section D, Biological crystallography* **66**, 12–21 (2010).
44. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta crystallographica. Section D, Biological crystallography* **60**, 2256–2268 (2004).
45. Subach, O. M. *et al.* Conversion of red fluorescent protein into a bright blue probe. *Chem Biol* **15**, 1116–1124 (2008).

Acknowledgments

We thank Jinghang Zhang and Lydia Tesfa for assistance with flow cytometry, Andrei Sobchuk for time-resolved measurements, and Rafael Toro for setting up crystallization screens. This work was supported by grants GM073913, CA164468, CA013330, and EB013571 from the National Institutes of Health.

Author contributions

K.D.P. and K.S.M. developed and characterized the protein in mammalian cells. K.D.P. characterized the protein *in vitro*. V.N.M. and S.C.A. solved the crystal structures of the protein. N.A.N. performed the time-resolved experiments. V.V.V. designed the overall project and together with K.D.P. and N.A.N. planned experiments, discussed data, and wrote the manuscript.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors declare no competing financial interests.

License: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>

How to cite this article: Piatkevich, K.D. *et al.* Extended Stokes Shift in Fluorescent Proteins: Chromophore-Protein Interactions in a Near-Infrared TagRFP675 Variant. *Sci. Rep.* **3**, 1847; DOI:10.1038/srep01847 (2013).