"Turn-On" Protein Fluorescence: In Situ Formation of Cyanine Dyes

Ipek Yapici, Kin Sing Stephen Lee, Tetyana Berbasova, Meisam Nosrati, Xiaofei Jia, Chrysoula Vasileiou, Wenjing Wang, Elizabeth M. Santos, James H. Geiger,* and Babak Borhan*

Department of Chemistry, Michigan State University, East Lansing, Michigan 48824, United States

S *Supporting Information*

ABSTRACT: Protein reengineering of cellular retinoic acid binding protein II (CRABPII) has yielded a genetically addressable system, capable of binding a profluorophoric chromophore that results in fluorescent protein/chromophore complexes. These complexes exhibit far-red emission, with high quantum efficiencies and brightness and also exhibit excellent pH stability spanning the range of 2−11. In the course of this study, it became evident that single mutations of L121E and R59W were most effective in improving the fluorescent characteristics of CRABPII mutants as well as the kinetics of complex formation. The readily crystallizable nature of these proteins was invaluable to provide clues for the observed spectroscopic behavior that results from single mutation of key residues.
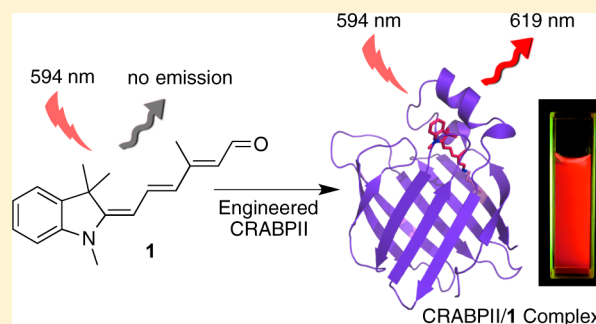
INTRODUCTION

The impact of fluorescent protein tags in molecular and cellular biology has stimulated wide interest in optimizing current systems or developing fundamentally new methodologies. The two basic strategies for fluorescent protein labeling so far employed are fluorescent proteins and site-specific chemical labeling systems (an acceptor peptide motif that couples with an exogenous chromophore), as summarized in a number of excellent reviews.[1] The use of fluorescent proteins such as GFP and its many variants has the advantage of genetic addressability, without the requirement for exogenous chromophores.[2] On the other hand, more diverse fluorophores are available with site-specific chemical labeling systems. In these methods, an organic small molecule can be conjugated to an acceptor peptide via enzymatic action [HaloTag,[3] SNAP/CLIP,[4] and LpIA acceptor peptide[5] as examples], direct binding [PYP-Tag,[6] TMP-Tag,[7] fluorogen activating proteins (FAPs)],[8] or through bioorthogonal reactions [FLAsH and ReAsH biarsenical based dyes,[9] Staudinger ligation,[10] click reactions,[11] and tetrazine-based cycloadditions[12]]. The fluorophores or fluorogenic molecules can, in principle, be altered to generate various peptide-dye reporters for a variety of applications.

Noteworthy are examples of two systems that differ from the latter two subgroups by utilizing endogenous molecules, which upon complexation with their target proteins produce fluorescent species. UnaG, isolated from a Japanese eel, belongs to the fatty acid binding protein family. Binding of bilirubin with UnaG turns on fluorescence.[13] The second example is the bacteriophytochromes that bind biliverdin, leading to fluorescence in the near IR range.[14] The photophysical characteristics of these complexes have proven optimal for in vivo imaging applications as infrared fluorescent proteins (IFPs). From the compilation of recent advances in the use of fluorescence in molecular and cellular biology, one can argue that the union of a genetically addressable system (proteins) with the flexibility in attaining the desired spectroscopic characteristics (small molecules) would provide a conceivably limitless array of protein fusion tags that could be optimized for a variety of applications.

Our interest in this area was piqued by our recent success in controlling the absorptive properties of protein-embedded chromophores, specifically the chromophore of vision, retinal.[15] The iminium-based pigment is generated with retinal and an active site Lys residue. This inspired us to consider a system that not only requires the union of a molecule and a protein for ultimate flexibility (as alluded to above) but also utilizes a nonfluorescent molecule that becomes fluorescent only upon binding and reaction with the target protein. This "turn-on" approach would have the advantage of low background since the unbound molecule is not fluorescent. The initial foray, disclosed herein, describes the pairing of a merocyanine dye precursor with a family of cellular retinoic acid binding protein II (CRABPII) mutants, exhibiting red-shifted emission (605−619 nm), with high quantum efficiencies (up to 39%) and brightness.

RESULTS AND DISCUSSION

CRABPII as the Protein Fusion Tag. CRABPII, a retinoic acid chaperone, is a member of the intracellular lipid binding protein family (iLBPs). It is a small (15.6 kDa) cytosolic protein with a relatively large binding cavity that can accommodate a diverse set of ligands[16] and is remarkably tolerant of mutations, similar to several other members of the iLBP family.[17] As such, it provides an ideal framework for applications in protein redesign.

This was demonstrated in our previous work in which CRABPII and other proteins of the iLBP family were re-engineered to generate a protonated Schiff base (PSB) with all-*trans*-retinal, analogous to rhodopsin.[15]

Choice of the Fluorophore. Cyanine dyes, which have a rich history in spectroscopy, were chosen as the target fluorophore. Various cyanine dyes have also played a central role in spectroscopic applications related to molecular biology, and are by in large innocuous and well behaved in biological systems.[18] The precursor cyanine aldehyde **1** is not fluorescent leading to negligible background fluorescence from unbound free aldehydes. Formation of the iminium yields a permanent resonating cation, resulting in a fluorophoric entity (merocyanine **2**). This push–pull system, which terminates at the nitrogen atoms at the two ends of the polyene, leads to a bathochromically shifted chromophore (ideal for biological applications) with an absorption profile distinct from its parent aldehyde (Figure 1).[19] Furthermore, the mode of fluorescence activation, namely
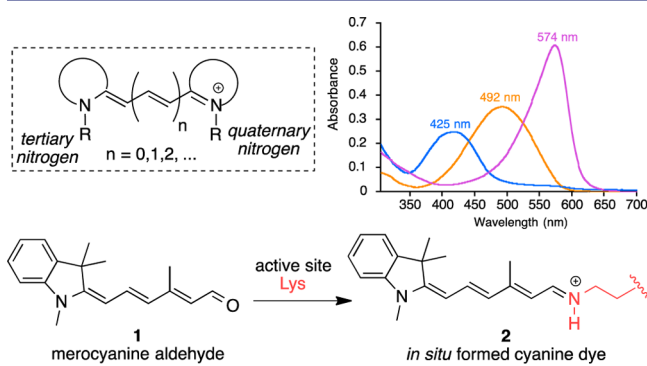

Figure 1. Structure of a typical cyanine dye (in dashed box). Merocyanine aldehyde **1** binds to an active site Lys residue to generate the red-shifted cyanine dye. UV−vis spectra of **1** (orange), Schiff base of **1** with *n*-butylamine (blue), and the protonated Schiff base of **1** with *n*-butylamine (magenta) [all in PBS].

the formation of an iminium, ideally suits our engineered protein systems. The cyanines are also attractive because the polyene tail structurally mimics the retinylidene ligands already shown to bind strongly to CRABPII mutants. Lastly, the quantum efficiency of cyanine dyes is affected environmentally. In comparison to common fluorophores, most are mildly fluorescent, probably due to their backbone flexibility. An increase in quantum efficiency is observed in viscous solvents, such as cold glycerol, that reduce torsional freedom.[20] Thus, we hypothesized that binding of the fluorophore with CRABPII could rigidify the molecule, leading to enhancement of fluorescence. We could also enhance the fluorescence through mutations that would reduce "wiggle" room in the binding pocket. This element adds yet another mode for reducing background, since only the molecule that is bound to its engineered target would yield high signal-to-noise fluorescence.

Merocyanine aldehyde **1** absorbs maximally at 492 nm in phosphate buffered saline (PBS), while the corresponding Schiff base (formed by reaction with *n*-butyl amine) blue shifts to 425 nm. Acidification results in a large red shift in the UV−vis spectrum, consistent with the formation of the iminium (λ_{max} at 574 nm, Figure 1). As expected, the resulting merocyanine-PSB **2** becomes mildly fluorescent in buffer solution ($\phi_F = 4\%$, $\lambda_{ex} = 565$ nm, $\lambda_{em} = 604$ nm), with photophysical properties similar to

sulfoindocyanines and unsymmetrical imine-based trimethine and pentamethine cyanine dye analogues.[20b,21]

CRABPII/1 Fluorescent Complexes. Incubation of non-fluorescent merocyanine aldehyde **1** with CRABPII mutants triggers the *in situ* formation of a cyanine dye and generates pigments that exhibit remarkable bathochromicity and narrow absorption bands with large extinction coefficients (Table 1). Time dependent binding of the R132K:R111L mutant (KL) with merocyanine aldehyde **1** shows a clear isosbestic point (Figure 2a), indicating homogeneous formation of the PSB. Subsequent
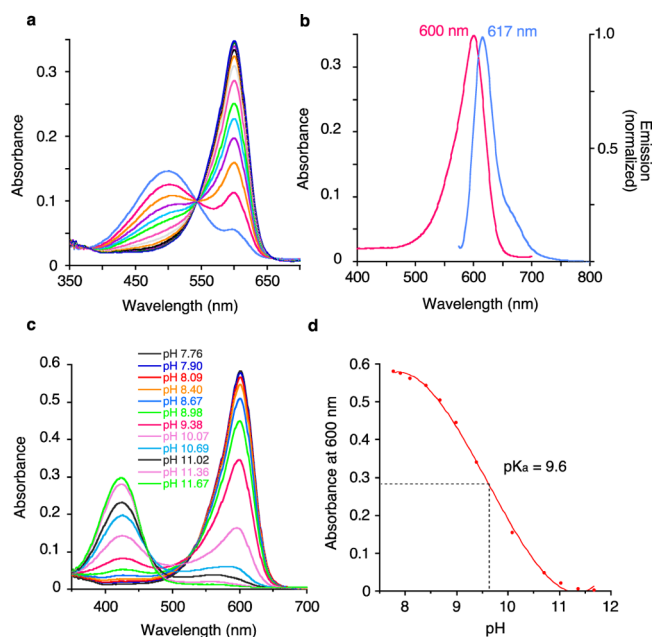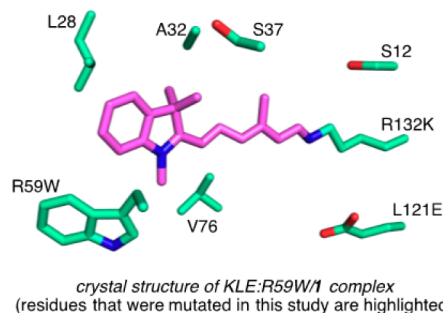

Figure 2. a. Absorption spectra, taken at 2 min intervals after addition of **1** (0.3 equiv) to KL-CRABPII (20 μM). b. Absorption and emission spectra of KL-CRABPII/**1** complex (excitation at 565 nm). c. Base titration of KL-CRABPII/**1** complex. d. Absorbance of KL-CRABPII/**1** complex as a function of pH with an apparent pK_a of 9.6.

addition of retinal does not displace the merocyanine from the binding site, demonstrating that the covalent linkage between the chromophore and the protein is stable. Excitation of the KL/**1** complex at 565 nm yields a narrow emission spectrum that peaks at 617 nm, as depicted in Figure 2b. Base titration of the iminium complex (Figure 2c) reveals a pK_a of 9.6 (Figure 2d), well above the physiological pH range.

With a brightly fluorescent protein in hand, an effort was made to regulate the absorption and emission spectra of the complex using strategies developed in our previous work.[15a,c,g] These include changing the electrostatic environment along the chromophore, altering and/or eliminating the interaction of the counteranion for the PSB, and creating a more enclosed binding pocket by introducing large hydrophobic residues at the entrance of the cavity.

The electrostatic environment in the vicinity of the PSB was altered by introducing a negatively charged residue at position 121; this mutation was required in our previous work to produce a PSB with retinal.[15c−e] Placement of a negatively charged residue near the retinylidene PSB also led to a large blue shift in absorption, as a result of stabilizing the cationic charge on the iminium nitrogen atom (less delocalization). With merocyanine as the chromophore, introduction of a glutamate (R132K:R111L:L121E, KLE) did cause a slight hypsochromic

Table 1. Spectroscopic data of CRABPII mutants complexed with 1

crystal structure of KLE:R59W/1 complex
(residues that were mutated in this study are highlighted)

	fluorescent protein	absorption λ_{max} (nm)	emission λ_{max} (nm)	quantum yield (%)[a]	ε (M⁻¹ cm⁻¹)[b]	brightness[c] (% mKate2)	brightness (% EGFP)
1	mKate2	588	633	40	62 500	100	74
2	mRFP[d]	584	607	27	50 000	54	40
3	R132K:R111L (KL)	600	617	18	111 700	80	60
4	KL:R59W	602	619	23	140 100	129	95
5	KL:L121E (KLE)	591	612	33	93 200	123	91
6	KL:L121D (KLD)	591	612	30	97 500	117	87
7	KLE:R59W	595	616	39	169 800	265	196
8	KLE:V76W	569	605	33	77 600	102	76
9	KLE:S37W	579	609	30	96 500	116	87
10	KLE:L28W	581	611	31	58 600	73	54
11	KLE:A32W	590	614	31	104 300	129	96
12	KLE:V76W:L28W	571	609	32	67 000	86	63
13	KLE:R59W:L28W	595	619	38	158 400	241	178
14	KLE:R59W:A32W	595	617	39	142 900	223	165

[a]The quantum yields of CRABPII mutants were determined based on two fluorescent standards (Oxazine-1 and Oxazine-170). [b]ε is the measured extinction coefficient for each protein complex at their λ_{max}. [c]Brightness was calculated as the product of ε and the fluorescence quantum yield (presented as % of mKate2 and EGFP brightness).[24] [d]The quantum yield of mRFP was measured under the same conditions and calculated as 27.2% (Figure S6).

shift in the absorption of the protein/chromophore complex in comparison with KL, but more significantly, improved the fluorescence quantum yield (entry 5, Table 1, see below for further discussion). Installation of an aspartate residue as the counteranion displayed similar photophysical characteristics as the KLE mutant (KLD, entry 6, Table 1). The small change in absorption for merocyanine bound protein complexes is putatively due to the more delocalized nature of the resonating cation with the cyanine dyes. As a result, the cation is not centralized on the nitrogen atom to be affected greatly by the placement of counteranion.

Cyanine dyes exhibit environmentally sensitive fluorescence properties under different conditions such as different solvent polarity, dielectric constant, ionic strength and viscosity. Previous reports have documented decreased levels of fluorescence as a function of increasing solvent polarity and vice versa.[22] In a similar fashion, changes to the binding pocket that further isolate it from the aqueous media, or reduce interactions of bound water molecules with the chromophore, could result in enhancement of fluorescence. To assess the effect of closing the binding cavity and creating a less solvent accessible binding pocket, Arg59, located at the portal of the binding site, was changed to a number of different amino acids (Table S1). Spectroscopic analyses of these mutants showed little change in the absorption or emission wavelength, although replacement with large aromatic amino acids, in particular Trp, exhibited improved quantum efficiencies in all cases (see further discussion in structure of KLE:R59W complex). They also significantly increased the extinction coefficient of the protein/chromophore complex, leading to

substantial improvement in fluorescent brightness (for an example see Table 1 entry 3 vs 4).

Further improvements in the fluorophoric properties of the merocyanine bound CRABPII complexes were focused on rigidifying the bound chromophore through packing interactions with large amino acid side chains. To this end Trp, with its large size (240 Å³ average volume, 163 Å³ van der Waals volume) and small rotomeric flexibility, was installed at various positions in the binding cavity (Table 1, entries 8−14).[23] It should be noted that the indole ring of Trp is polarizable and has a significant dipole moment, and thus could interact with the bound chromophore in ways other than purely steric. Nonetheless, no improvement in either extinction coefficient or quantum efficiency was realized when compared to the best R59W mutants. In short, KLE:R59W exhibited the largest extinction coefficient, the highest quantum yield, and therefore the highest brightness of any of the mutants tested, surpassing mKate2, one of the brightest fluorescent proteins known in this wavelength regime.[24] Notably, in comparison to the PSB of 1 with n-butylamine, the protein complex achieves nearly a 10-fold increase in quantum efficiency.

Structures of CRABPII/1 Fluorescent Complexes. In the course of this study, it became evident that single mutations of L121E and R59W were most effective in improving the fluorescent characteristics of KL-CRABPII mutants. In an effort to understand the effect of these mutations, structures of several merocyanine-bound CRABPII mutants were determined. The readily crystallizable nature of these proteins was invaluable to provide clues for the observed spectroscopic behavior that results from single mutation of key residues. As will be evident upon review of the data below, our studies indicate that the large

binding cavity and relative structural plasticity of CRABPII provides conformational freedom for the bound ligand to adopt various binding conformations. This accounts for the observed spectroscopic characteristics for different mutants.

Structure of the KL/1 Complex. Our previous efforts in designing rhodopsin protein mimics were aided through the acquisition of numerous crystal structures of CRABPII and hCRBPII bound retinal protein complexes.[15] Although rotameric differences were observed as a result of different mutations, the bound chromophore was always found in the same location within the binding pocket. Structures of merocyanine bound protein complexes defied this trend. This is immediately evident in the structure of the KL/1 complex (Figure 3a,d), where the ligand follows a trajectory that is completely distinct from that seen in previous ligand bound structures of CRABPII (for an overlay of KL/1 structure with its retinal bound counterpart, see Figure S12). With retinal, the ionone ring is located near the mouth of the binding cavity, which is in contrast to the indoline ring of merocyanine, tucked into a hydrophobic pocket deep within, featuring significant hydrophobic interactions between Leu121 and the ligand. The chromophore forms a *cis*-imine with Lys132, which is stabilized by a water-mediated hydrogen bond to Ser12. Additionally, the cyanine dye is severely twisted about the C3−C4 bond ($\psi_1 = 59.1°$), which also adopts an s-*cis* conformation to accommodate this binding regime. As previously reported, a twisted polyene chain provides an effective conduit for nonradiative decay of the excited state.[25] It is therefore not surprising that the severely twisted KL/1 complex has the lowest quantum yield among all CRABPII mutants (Table 1).

Structure of the KLE/1 Complex. A radical change in binding of merocyanine is observed upon altering the electrostatic environment at position 121. Exchange of the hydrophobic Leu121 residue with either Glu or Asp yields protein complexes with **1** that are more fluorophoric than KL/1. The structures of KLE/1 (Figure 3b) and KLD/1 complexes (Figure S14) show that the chromophore binds the active site Lys residue from a trajectory that points opposite to that observed previously. The structure of KL/1 complex illustrates the tightly packed interaction of Leu121 with the hydrophobic indoline ring of merocyanine, lying within 3.6 Å away (Figure 3a). The electrostatic change, as a result of L121E and L121D mutations, plausibly alters the latter advantageous hydrophobic interaction, leading to an alternate binding site that is more energetically favorable. In the new binding orientation, the heterocyclic ring of the ligand is projected out of the binding cavity between strand 1 and helix 2, leaving it relatively exposed to the exterior of the protein. The chromophore adopts a more planar conformation, as a result of having less structural restrictions imposed by the protein. As a result, and in contrast to the KL/1 complex, the C3−C4 dihedral angle is much more planar, while still adopting an s-*cis* conformation for both KLE/1 and KLD/1 complexes. As anticipated, the more planar KLE and KLD mutants have higher quantum yields than the severely twisted KL/1 complex (Table 1, entries 5 and 6 vs 3).

Structure of the KLE:R59W/1 Complex. Mutation of both Arg59 to Trp and Leu121 to Glu results in a third, completely distinct ligand binding mode (Figure 3c). The structure of the KLE:R59W/1 complex depicts the binding of the ligand in a fashion similar to those previously observed with retinal and retinoic acid bound CRABPII mutants (Figure 3c, for a comparison with previously determined crystal structures see Figure 3d and Figure S13). The indoline ring points toward the mouth of the binding cavity and the aldehyde end of the ligand is
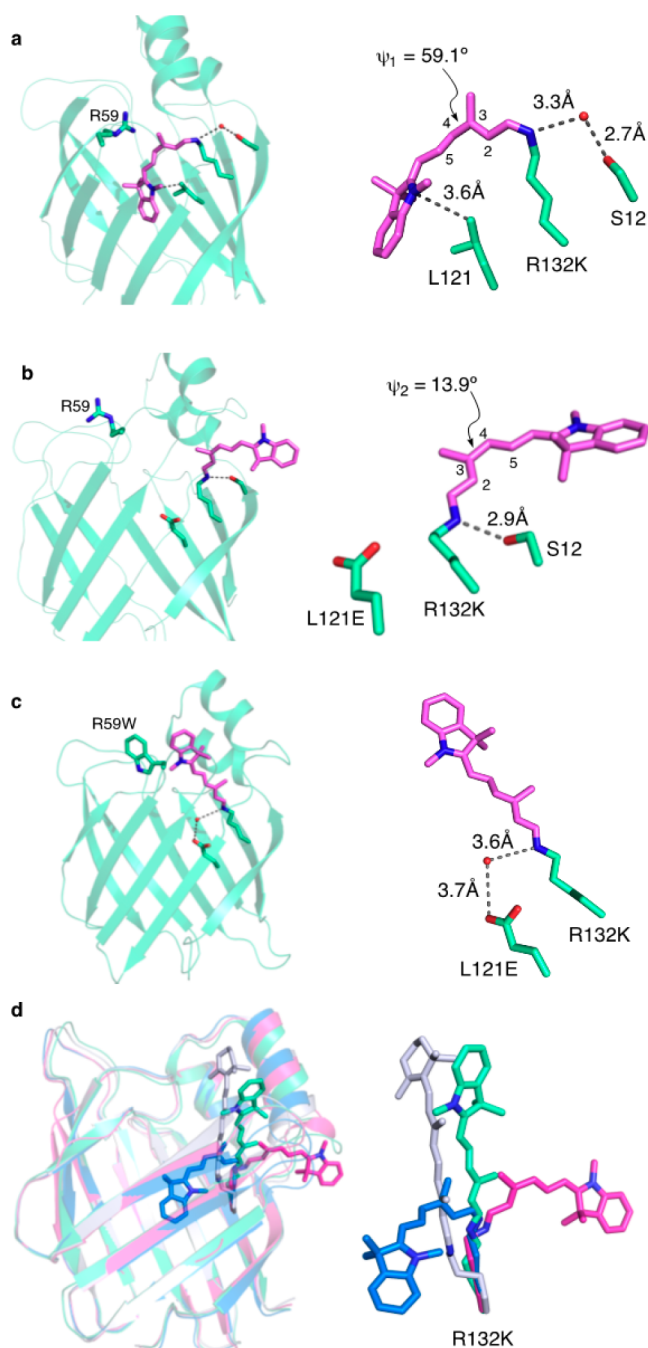

Figure 3. a. Crystal structure of KL/**1** complex (1.7 Å resolution, PDB ID: 4QGV). The enlarged image depicts the twist around the C3−C4 of the bound ligand adopting an s-*cis* conformation. b. Crystal structure of KLE/**1** complex (1.5 Å resolution, PDB ID: 4QGX) with the expanded view of binding mode. c. Crystal structure of KLE:R59W/**1** complex (2.7 Å resolution, PDB ID: 3FEP) with the exploded view of the KLE:R59W/**1** complex. d. Overlay of all three crystal structures (KL-blue, KLE-pink, and KLE:R59W-cyan) along with the crystal structure of all-*trans*-retinal bound to KLE-gray (1.2 Å resolution, PDB ID: 2G7B) illustrating the divergence in binding conformation that results from a single mutation in each case.

deeply buried in the interior of the protein. The Schiff base between the active site Lys residue, R132K, and the merocyanine aldehyde adopts a *trans*-imine geometry while the entire polymethine chain stays in an s-*trans* conformation. The installed counteranion, Glu121, interacts with the iminium via a water

molecule that resides 3.6 Å away from the nitrogen atom, yielding a complex that exhibits a pK_a of 10.5 for the iminium (Figure 3c). The conformation of the chromophore is quite linear and relatively well packed within the binding cavity with restricted ability to move, in contrast to the previous structures, where it either severely deviates from planarity (in KL) or is less well packed and more solvent exposed (as in KLE and KLD).

This mutant has superior fluorescent characteristics as compared to other mutants investigated in this study, possibly as a result of the following observations; (i) the relatively flat and well-packed conformation leads to maximal overlap of the π system, (ii) closure of the binding cavity with a large Trp residue further isolates the binding pocket and could also exclude water molecules that otherwise might weaken fluorescence through enhancing nonradiative decay of the excited state, (iii) although the crystal structure precludes $\pi-\pi$ stacking between Trp59 and the indoline ring of the bound chromophore, its close proximity (3.4 Å) could result in a favorable electrostatic interaction that further rigidifies the chromophore. Noteworthy, the replacement of Glu121 with Asp121 results in no loss in quantum efficiency and thus suggests that the shorter aspartate side chain is of sufficient length to maintain the interaction with the bound chromophore without causing a significant change in ligand geometry (KLD:R59W, ϕ_F = 38, λ_{abs} = 595 nm, λ_{em} = 615 nm).

Given all the structural information obtained, it would appear that mutation of both Leu121 to Asp or Glu, and Arg59 to Trp, is required to obtain the more orthodox ligand trajectory seen in KLE:R59W/**1**. As predicted from the crystal structure of KLE:R59W/**1** (Figure 3c), a favorable hydrophobic packing between Trp59 and the heterocycle is gained with the Trp mutation. Therefore, it appears that a combination of two conditions is necessary with merocyanine as the ligand to adopt the "normal" retinylidene-like binding trajectory. First, an anionic residue is required at position 121, which destabilizes the ligand conformation seen in KL, and provides a water-mediated interaction with the iminium in KLE:R59W. Second, placement of the Trp residue at position 59, which provides hydrophobic interactions with the merocyanine ring, leads to the relatively well-packed, flat ligand conformation seen in the KLE:R59W/**1** structure. As a result, these changes lead to the highest extinction coefficient and quantum efficiency, since the chromophore is flattened and conformationally restricted. Indeed, the mutants with the best fluorescent characteristics invariably contain the mutations at both of these positions.

Kinetics of PSB Formation. Mutations that introduced bulky residues within the binding pocket, as well as the altered binding modes of the merocyanine ligand for some mutants, could adversely affect the kinetics of PSB formation. The relative rate of mutants as compared to KL double mutant was measured by spectroscopically monitoring the conversion of the free merocyanine aldehyde **1** (absorbing at 492 nm) into its corresponding protonated Schiff base formed within the active site of the protein (absorbing higher than 570 nm, Table 2). The presence of Trp at position 59 substantially decreased the rate of PSB formation of the corresponding parent KL mutant (entry 2, Table 2). However, installation of L121D or L121E restores the original binding trajectory of the chromophore and accelerates PSB formation (entries 5 and 6, Table 2). Alternatively, installation of a counteranion (Asp and Glu) could also increase the rate of PSB formation, presumably as a result of acid catalyzed activation of the aldehydic moiety for imine formation.[15d] Interestingly, the faster rate observed for KLE:R59W as compared to KLD:R59W does suggest the role of acid catalysis,

Table 2. Relative Rate of Ligand Binding and PSB Formation[a]

	mutants	relative rate
1	KL	1
2	KL:R59W	0.389
3	KLE	0.950
4	KLD	0.243
5	KLD:R59W	0.955
6	KLE:R59W	1.809
7	KLE:V76W	3.550
8	KLE:S37W	0.431
9	KLE:L28W	0.376
10	KLE:A32W	0.428
11	KLE:V76W:L28W	1.394
12	KLE:R59W:L28W	1.160
13	KLE:R59W:A32W	0.779

[a]Kinetic measurements were performed at 23 °C with 20 μM protein and 0.3 equiv of merocyanine **1**. PSB formation was monitored by UV−vis at λ_{max} for each complex over time (see the SI for experimental details).

owing to the closer proximity of Glu to the bound aldehyde as opposed to Asp.

Most Trp mutations along the chromophore slowed imine formation, thus extending the time for full binding and complexation. This is presumably due to steric hindrance imposed by the bulky Trp residues that could interfere with either ligand entry, restrict certain binding conformations, or reduce accessibility to the active site Lys residue. An exception from this observation was the incorporation of Trp at position 76, which led to an increase in the rate of PSB formation (entries 7 and 11, Table 2). Interestingly, the increased rate of PSB formation, along with an observed blue shift of absorption for all V76W mutants (entries 8 and 12, Table 1) may suggest that these mutants are structurally more flexible and solvent accessible.

To better benchmark the current system with existing methods, detailed kinetic analysis was conducted to calculate the rate constant for CRABPII/**1** complex formation. The second-order rate constant for the reaction of merocyanine **1** with KLE:R59W was determined by measuring a series of pseudo first order rates (excess **1** at different concentrations relative to the protein), following previously reported procedures (see Figure 4a and SI for experimental details).[6b,26] KLE:R59W exhibits a
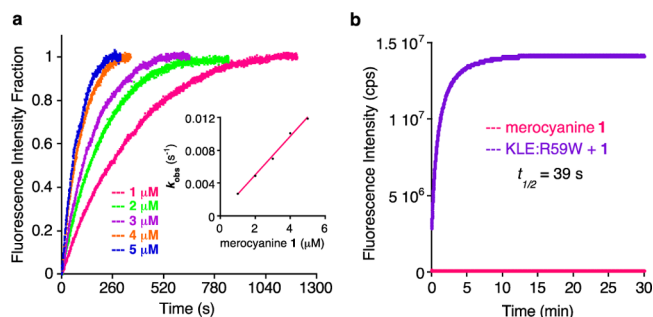

Figure 4. a. Time-course binding of KLE:R59W (100 nM) with **1** (1−5 μM) in PBS at 37 °C. k_{obs} values at different merocyanine concentrations show a linear relationship (R^2 = 0.995), yielding a k_2 (second order rate constant) of 2350 M^{-1} s^{-1}. b. Stoichiometric binding of KLE:R59W with **1** (5 μM each) in PBS at 37 °C was followed at 616 nm with excitation at 565 nm (0.5 s intervals). The data fits a second order process (R^2 = 0.996, see SI for details), with a calculated binding $t_{1/2}$ = 39 s. Merocyanine **1** does not exhibit any fluorescence, as illustrated in the graph.

remarkably high rate of PSB formation with a full complexation within 7 min (k = 2356 M^{-1} s^{-1}, Figure 4b) and a short half-life ($t_{1/2}$ = 39 s). These values compare favorably with other fast labeling methods such as PYP-Tag (k = 3950 M^{-1} s^{-1}, $t_{1/2}$ = 1.1 min, full labeling within 6 min),[6b] BGSBD/SNAP-Tag ($k \approx$ 7200 M^{-1} s^{-1}, $t_{1/2}$ = 24 s, full labeling within 3 min),[26] and bioorthogonal reactions for labeling proteins (k in the range of 10^{-4} to 10^4 M^{-1} s^{-1}),[27] all of which have similar or longer $t_{1/2}$ times.

Persistence of Fluorescence over a Broad pH Range. The fluorescence of the protein-bound merocyanine PSB depends on the protonation state of the chromophore, exhibiting maximum bathochromic shift and quantum efficiency as an iminium (Figure 5a). Figure 5b depicts the UV−vis spectra of
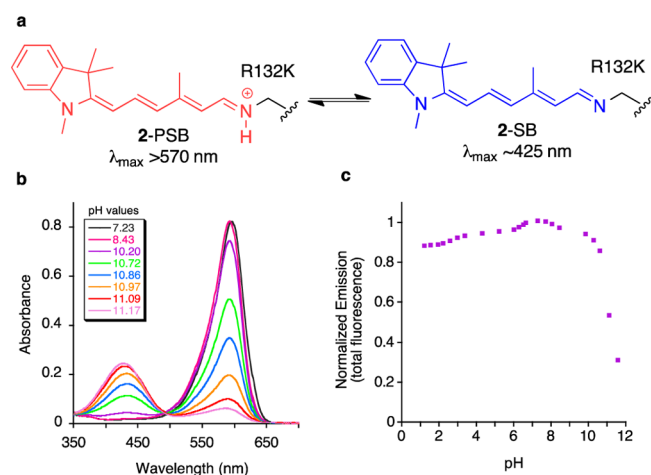
Figure 5. a. Protonated (emissive state) and unprotonated (nonemissive state) forms of protein-bound merocyanine. b. Base titration of KLE:R59W:L28W/**1** complex with NaOH in PBS, pH values are indicated. c. Total fluorescence of the KLE:R59W:L28W/**1** complex as a function of pH.

KLE:R59W:L28W/**1** complex titrated with NaOH.[28] Basification of the solution does not deprotonate the PSB until pH ∼10; further addition of base gives rise to the absorption at ∼425 nm (Schiff base absorption). During this process negligible fluorescence is lost up to pH of 10.2 (Figure 5c). The absorption spectra reveal no change in their respective wavelengths, indicating that the complex maintains its tertiary structure even under highly basic conditions (pH as high as 11). Denaturation experiments with the addition of detergent result in an absorption band at 583 nm (Figure S11), resembling that of the *n*-butyl iminium form of merocyanine aldehyde in the presence of BSA.[29]

Next, we investigated the durability of the CRABPII-merocyanine **1** complex under highly acidic conditions. Previously, we had demonstrated that some CRABPII mutants are acid resistant, maintaining their native fold at low pH levels.[15a] This stability extends to merocyanine bound CRABPII mutants, as illustrated by the KLE:R59W:L28W complex. Lowering the pH to 2.2 resulted in a small increase in absorption of the protein-chromophore complex with no change in λ_{max}. The slight change in absorption could be due to protonation of residues close to the chromophore, affecting protein−chromophore interactions. The increase of the solution acidity affects the emissive state of the protein−chromophore complex slightly (Figure 5c). Since CRABPII-merocyanine variants retain most of their fluorescence, they may find application for cellular imaging of acidic

intracellular compartments without loss of fluorescence or self-quenching issues.

Visualization of the CRABPII/1 Fluorescent Complexes in Bacteria. In a proof-of-principle experiment to show feasibility of merocyanine for cellular work, CRABPII mutants, KL (Table 1, entry 1, see images in SI, Figure S8) and KLE:R59W:L28W (Table 1, entry 13, shown in Figure 6a and
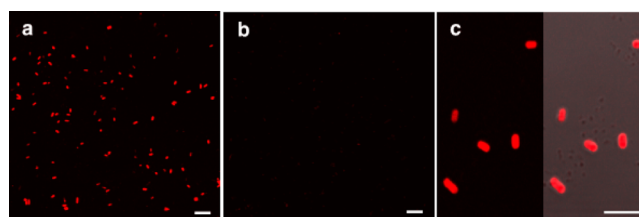
Figure 6. a. Fluorescence visualization of *E. coli* cells, transformed with a vector overexpressing KLE:R59W:L28W, incubated with **1** (594 nm excitation, 615 nm long pass filter for emission, scale bar is 10 μm). b. Control panel (nontransformed cells) shows no fluorescence after addition of **1** (scale bar is 10 μm). c. On left, fluorescence of KLE:R59W:L28W/**1** at 594 nm excitation was enlarged 5-fold (scale bar is 5 μm). On right, brightfield image was overlaid with the fluorescence of the complex.

6c), expressed in *E. coli*, were incubated with **1** and imaged. The choice of the latter two mutants was based on their differing spectral characteristics. The bacterial cells readily uptake the chromophore within a minute of addition, and instantly yield visibly colored CRABPII-merocyanine adducts (SI, Figure S7). Both CRABPII mutants exhibited intense fluorescence with excitation at 594 nm. The control cells, which were not transformed with the plasmid expressing CRABPII mutants, were devoid of nonspecific fluorescence after incubation with merocyanine aldehyde, demonstrating that background fluorescence from the unbound chromophore is minimal (Figure 6b). It should be noted that images produced in Figure 6 required less than 1 min of incubation with the profluorogenic aldehyde **1**. The CRABPII system also displays high selectivity in binding for merocyanine **1** in comparison to other Lys rich proteins such as BSA, thus reducing nonspecific binding. Uninduced cells, expressing basal levels of CRABPII mutants, also produced bright red fluorescence, indicating the ability to visualize low-level protein expression systems (SI, Figure S9).

Control *E. coli* cells that were not transformed with the vector containing CRABPII mutants showed minimal background fluorescence with **1**. This is a good indication that in the time required for maturation of CRABPII mutants with **1**, there is little nonspecific binding with off-target proteins. This was further tested in vitro, with BSA as an off-target protein capable of forming iminium bonds with **1** through reaction with its numerous Lys residues. Under identical reaction conditions, BSA reacted to yield 7% of the total fluorescence as compared to the CRABPII/merocyanine system (SI, Figure S4), showing that rapid complex formation of CRABPII with **1** overcomes the nonspecific binding processes.

Photobleaching assays of CRABPII-merocyanine complexes in *E. coli* cells were performed and compared to a monomeric red fluorescent protein, mRFP, under identical experimental conditions (see SI for details). The mutants KLE:R59W:L28W and KLE:R59W complexed with merocyanine **1** showed faster photobleaching behavior as compared to mRFP. Both mutants showed about 40% loss of fluorescence over 250 s of continuous

illumination while mRFP displayed about 20% loss during the same period (SI, Figure S10).

CONCLUSION

In summary, we have engineered CRABPII into a fluorescent protein via coupling with a nonfluorescent cyanine dye precursor. CRABPII/**1** complexes demonstrated structural variety in terms of ligand orientation and geometry that correlates well with the expected fluorescent properties. Our probe is sufficiently selective to enable live *E. coli* cell imaging. The chromophore is readily cell-permeable and well behaved in live bacteria, coupling with CRABPII variants instantaneously, and allowing visualization in bacterial cells within minutes of chromophore addition. Noteworthy, background fluorescence is minimal since the unbound aldehyde form of the chromophore has significantly different spectral properties, as compared to the covalently bound protonated Schiff base form that leads to red-shifted spectra in the protein environment. Since engineered CRABPII variants feature a modular design that can be readily adapted to new fluorophores, spanning the entire visible–infrared spectral region, ongoing work envisions constructing a library of several profluorogenic aldehydes for multicolor single molecule imaging in live cells.

ASSOCIATED CONTENT

Supporting Information

Experimental details, protein characterization data, protein crystallographic data, and supplementary figures and tables. The crystal structures of CRABPII-R132K:R111L (KL), CRABPII-R132K:R111L:L121E (KLE), CRABPII-R132K:R111L:L121D (KLD), and CRABPII-R132K:R111L:L121E:R59W (KLE:R59W) were obtained as a part of this study and deposited to the Protein Data Bank with accession codes 4QGV, 4QGX, 4QGW, and 3FEP, respectively. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors
geiger@chemistry.msu.edu
babak@chemistry.msu.edu

Notes
The authors declare no competing financial interest.

REFERENCES

(1) (a) Sunbul, M.; Yin, J. *Org. Biomol. Chem.* **2009**, *7*, 3361. (b) Jing, C. R.; Cornish, V. W. *Acc. Chem. Res.* **2011**, *44*, 784. (c) Sletten, E. M.; Bertozzi, C. R. *Angew. Chem., Int. Ed.* **2009**, *48*, 6974. (d) Rashidian, M.; Dozier, J. K.; Distefano, M. D. *Bioconjugate Chem.* **2013**, *24*, 1277. (e) Hinner, M. J.; Johnsson, K. *Curr. Opin. Biotechnol.* **2010**, *21*, 766. (f) Marks, K. M.; Nolan, G. P. *Nat. Methods* **2006**, *3*, 591. (g) Dean, K. M.; Palmer, A. E. *Nat. Chem. Biol.* **2014**, *10*, 512. (h) Jung, D.; Min, K.; Jung, J.; Jang, W.; Kwon, Y. *Mol. Biosyst.* **2013**, *9*, 862.

(2) (a) Rizzo, M. A.; Davidson, M. W.; Piston, D. W. *Cold Spring Harb. Protoc.* **2009**, *2009*, top63. (b) Shaner, N. C.; Steinbach, P. A.; Tsien, R. Y. *Nat. Methods* **2005**, *2*, 905. (c) Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. *Science* **1994**, *263*, 802. (d) Heim, R.; Cubitt, A. B.; Tsien, R. Y. *Nature* **1995**, *373*, 663. (e) Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F. G.; Cormier, M. J. *Gene* **1992**, *111*, 229. (f) Shimomura, O.; Johnson, F. H.; Saiga, Y. *J. Cell. Comp. Physiol.* **1962**, *59*, 223. (g) Chu, J.; Haynes, R. D.; Corbel, S. Y.; Li, P.; Gonzalez-Gonzalez, E.; Burg, J. S.; Ataie, N. J.; Lam, A. J.; Cranfill, P. J.; Baird, M. A.; Davidson, M. W.; Ng, H. L.; Garcia, K. C.; Contag, C. H.; Shen, K.; Blau, H. M.; Lin, M. Z. *Nat. Methods* **2014**, *11*, 572. (h) Kremers, G. J.; Gilbert, S. G.; Cranfill, P. J.; Davidson, M. W.; Piston, D. W. *J. Cell Sci.* **2011**, *124*, 157.

(3) (a) Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V. *ACS Chem. Biol.* **2008**, *3*, 373. (b) Los, G. V.; Wood, K. *Methods Mol. Biol.* **2007**, *356*, 195.

(4) (a) Gautier, A.; Juillerat, A.; Heinis, C.; Correa, I. R., Jr.; Kindermann, M.; Beaufils, F.; Johnsson, K. *Chem. Biol.* **2008**, *15*, 128. (b) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. *Nat. Biotechnol.* **2003**, *21*, 86. (c) Keppler, A.; Pick, H.; Arrivoli, C.; Vogel, H.; Johnsson, K. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9955. (d) Sun, X.; Zhang, A.; Baker, B.; Sun, L.; Howard, A.; Buswell, J.; Maurel, D.; Masharina, A.; Johnsson, K.; Noren, C. J.; Xu, M. Q.; Correa, I. R., Jr. *ChemBioChem* **2011**, *12*, 2217.

(5) (a) Uttamapinant, C.; White, K. A.; Baruah, H.; Thompson, S.; Fernandez-Suarez, M.; Puthenveetil, S.; Ting, A. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 10914. (b) Liu, D. S.; Nivon, L. G.; Richter, F.; Goldman, P. J.; Deerinck, T. J.; Yao, J. Z.; Richardson, D.; Phipps, W. S.; Ye, A. Z.; Ellisman, M. H.; Drennan, C. L.; Baker, D.; Ting, A. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, E4551.

(6) (a) Hori, Y.; Nakaki, K.; Sato, M.; Mizukami, S.; Kikuchi, K. *Angew. Chem., Int. Ed.* **2012**, *51*, 5611. (b) Hori, Y.; Norinobu, T.; Sato, M.; Arita, K.; Shirakawa, M.; Kikuchi, K. *J. Am. Chem. Soc.* **2013**, *135*, 12360.

(7) (a) Chen, Z.; Jing, C.; Gallagher, S. S.; Sheetz, M. P.; Cornish, V. W. *J. Am. Chem. Soc.* **2012**, *134*, 13692. (b) Gallagher, S. S.; Sable, J. E.; Sheetz, M. P.; Cornish, V. W. *ACS Chem. Biol.* **2009**, *4*, 547. (c) Jing, C.; Cornish, V. W. *ACS Chem. Biol.* **2013**, *8*, 1704.

(8) (a) Holleran, J.; Brown, D.; Fuhrman, M. H.; Adler, S. A.; Fisher, G. W.; Jarvik, J. W. *Cytometry A* **2010**, *77*, 776. (b) Ozhalici-Unal, H.; Pow, C. L.; Marks, S. A.; Jesper, L. D.; Silva, G. L.; Shank, N. I.; Jones, E. W.; Burnette, J. M., 3rd; Berget, P. B.; Armitage, B. A. *J. Am. Chem. Soc.* **2008**, *130*, 12620. (c) Szent-Gyorgyi, C.; Schmidt, B. F.; Creeger, Y.; Fisher, G. W.; Zakel, K. L.; Adler, S.; Fitzpatrick, J. A.; Woolford, C. A.; Yan, Q.; Vasilev, K. V.; Berget, P. B.; Bruchez, M. P.; Jarvik, J. W.; Waggoner, A. *Nat. Biotechnol.* **2008**, *26*, 235. (d) Wu, Y.; Tapia, P. H.; Fisher, G. W.; Waggoner, A. S.; Jarvik, J.; Sklar, L. A. *Cytometry A* **2013**, *83*, 220.

(9) (a) Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y. *J. Am. Chem. Soc.* **2002**, *124*, 6063. (b) Cao, H.; Xiong, Y.; Wang, T.; Chen, B.; Squier, T. C.; Mayer, M. U. *J. Am. Chem. Soc.* **2007**, *129*, 8672. (c) Griffin, B. A.; Adams, S. R.; Tsien, R. Y. *Science* **1998**, *281*, 269.

(10) (a) Hangauer, M. J.; Bertozzi, C. R. *Angew. Chem., Int. Ed.* **2008**, *47*, 2394. (b) Beatty, K. E.; Xie, F.; Wang, Q.; Tirrell, D. A. *J. Am. Chem. Soc.* **2005**, *127*, 14150. (c) Beatty, K. E.; Liu, J. C.; Xie, F.; Dieterich, D. C.; Schuman, E. M.; Wang, Q.; Tirrell, D. A. *Angew. Chem., Int. Ed.* **2006**, *45*, 7364.

(11) (a) Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 16793. (b) Yu, Z.; Pan, Y.; Wang, Z.; Wang, J.; Lin, Q. *Angew. Chem., Int. Ed.* **2012**, *51*, 10600. (c) Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 15046. (d) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192.

(12) (a) Liu, D. S.; Tangpeerachaikul, A.; Selvaraj, R.; Taylor, M. T.; Fox, J. M.; Ting, A. Y. *J. Am. Chem. Soc.* **2012**, *134*, 792. (b) Devaraj, N. K.; Weissleder, R.; Hilderbrand, S. A. *Bioconjugate Chem.* **2008**, *19*, 2297.

(c) Plass, T.; Milles, S.; Koehler, C.; Szymanski, J.; Mueller, R.; Wiessler, M.; Schultz, C.; Lemke, E. A. *Angew. Chem., Int. Ed.* **2012**, *51*, 4166. (d) Devaraj, N. K.; Hilderbrand, S.; Upadhyay, R.; Mazitschek, R.; Weissleder, R. *Angew. Chem., Int. Ed.* **2010**, *49*, 2869.

(13) Kumagai, A.; Ando, R.; Miyatake, H.; Greimel, P.; Kobayashi, T.; Hirabayashi, Y.; Shimogori, T.; Miyawaki, A. *Cell* **2013**, *153*, 1602.

(14) (a) Filonov, G. S.; Piatkevich, K. D.; Ting, L. M.; Zhang, J.; Kim, K.; Verkhusha, V. V. *Nat. Biotechnol.* **2011**, *29*, 757. (b) Shu, X.; Royant, A.; Lin, M. Z.; Aguilera, T. A.; Lev-Ram, V.; Steinbach, P. A.; Tsien, R. Y. *Science* **2009**, *324*, 804. (c) Wagner, J. R.; Zhang, J.; von Stetten, D.; Gunther, M.; Murgida, D. H.; Mroginski, M. A.; Walker, J. M.; Forest, K. T.; Hildebrandt, P.; Vierstra, R. D. *J. Biol. Chem.* **2008**, *283*, 12212. (d) Yu, D.; Gustafson, W. C.; Han, C.; Lafaye, C.; Noirclerc-Savoye, M.; Ge, W. P.; Thayer, D. A.; Huang, H.; Kornberg, T. B.; Royant, A.; Jan, L. Y.; Jan, Y. N.; Weiss, W. A.; Shu, X. *Nat. Commun.* **2014**, *5*, 3626. (e) Shcherbakova, D. M.; Verkhusha, V. V. *Nat. Methods* **2013**, *10*, 751.

(15) (a) Berbasova, T.; Nosrati, M.; Vasileiou, C.; Wang, W.; Lee, K. S.; Yapici, I.; Geiger, J. H.; Borhan, B. *J. Am. Chem. Soc.* **2013**, *135*, 16111. (b) Crist, R. M.; Vasileiou, C.; Rabago-Smith, M.; Geiger, J. H.; Borhan, B. *J. Am. Chem. Soc.* **2006**, *128*, 4522. (c) Lee, K. S. S.; Berbasova, T.; Vasileiou, C.; Jia, X.; Wang, W.; Choi, Y.; Nossoni, F.; Geiger, J. H.; Borhan, B. *ChemPlusChem.* **2012**, *77*, 273. (d) Vasileiou, C.; Vaezeslami, S.; Crist, R. M.; Rabago-Smith, M.; Geiger, J. H.; Borhan, B. *J. Am. Chem. Soc.* **2007**, *129*, 6140. (e) Vasileiou, C.; Wang, W.; Jia, X.; Lee, K. S.; Watson, C. T.; Geiger, J. H.; Borhan, B. *Proteins* **2009**, *77*, 812. (f) Wang, W.; Geiger, J. H.; Borhan, B. *Bioessays* **2014**, *36*, 65. (g) Wang, W.; Nossoni, Z.; Berbasova, T.; Watson, C. T.; Yapici, I.; Lee, K. S.; Vasileiou, C.; Geiger, J. H.; Borhan, B. *Science* **2012**, *338*, 1340.

(16) Chaudhuri, B. N.; Kleywegt, G. J.; Broutin-L'Hermite, I.; Bergfors, T.; Senn, H.; Le Motte, P.; Partouche, O.; Jones, T. A. *Acta Crystallogr. D Biol. Crystallogr.* **1999**, *55*, 1850.

(17) Li, E.; Norris, A. W. *Annu. Rev. Nutr.* **1996**, *16*, 205.

(18) (a) Escobedo, J. O.; Rusin, O.; Lim, S.; Strongin, R. M. *Curr. Opin. Chem. Biol.* **2010**, *14*, 64. (b) Levitus, M.; Ranjit, S. *Q. Rev. Biophys.* **2011**, *44*, 123. (c) Tatikolov, A. S. *J. Photochem. Photobiol. C* **2012**, *13*, 55.

(19) Hoischen, D.; Steinmuller, S.; Gartner, W.; Buss, V.; Martin, H. D. *Angew. Chem., Int. Ed.* **1997**, *36*, 1630.

(20) (a) Drexhage, K. H. In *Dye Lasers*; Schåfer, F., Ed.; Springer: Berlin, 1973; Vol. *1*, p 144. (b) Mujumdar, R. B.; Ernst, L. A.; Mujumdar, S. R.; Lewis, C. J.; Waggoner, A. S. *Bioconjugate Chem.* **1993**, *4*, 105.

(21) Meguellati, K.; Spichty, M.; Ladame, S. *Org. Lett.* **2009**, *11*, 1123.

(22) (a) Casalboni, M.; De Matteis, F.; Prosposito, P.; Quatela, A.; Sarcinelli, F. *Chem. Phys. Lett.* **2003**, *373*, 372. (b) Dai, Z. F.; Peng, B. X. *J. Chin. Chem. Soc.* **1998**, *45*, 237. (c) Muddana, H. S.; Morgan, T. T.; Adair, J. H.; Butler, P. J. *Nano Lett.* **2009**, *9*, 1559.

(23) (a) Baumann, G.; Frommel, C.; Sander, C. *Protein Eng.* **1989**, *2*, 329. (b) Najmanovich, R.; Kuttner, J.; Sobolev, V.; Edelman, M. *Proteins* **2000**, *39*, 261. (c) Richards, F. M. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151. (d) Zhao, S.; Goodsell, D. S.; Olson, A. J. *Proteins* **2001**, *43*, 271. (e) Darby, N. J.; Creighton, T. E. *Protein Structure*; Oxford University Press: Oxford, 1993.

(24) Shcherbo, D.; Murphy, C. S.; Ermakova, G. V.; Solovieva, E. A.; Chepurnykh, T. V.; Shcheglov, A. S.; Verkhusha, V. V.; Pletnev, V. Z.; Hazelwood, K. L.; Roche, P. M.; Lukyanov, S.; Zaraisky, A. G.; Davidson, M. W.; Chudakov, D. M. *Biochem. J.* **2009**, *418*, 567.

(25) (a) Cao, J. F.; Wu, T.; Hu, C.; Liu, T.; Sun, W.; Fan, J. L.; Peng, X. *J. Phys. Chem. Chem. Phys.* **2012**, *14*, 13702. (b) Sanchez-Galvez, A.; Hunt, P.; Robb, M. A.; Olivucci, M.; Vreven, T.; Schlegel, H. B. *J. Am. Chem. Soc.* **2000**, *122*, 2911. (c) Silva, G. L.; Ediz, V.; Yaron, D.; Armitage, B. A. *J. Am. Chem. Soc.* **2007**, *129*, 5710.

(26) Liu, T. K.; Hsieh, P. Y.; Zhuang, Y. D.; Hsia, C. Y.; Huang, C. L.; Lai, H. P.; Lin, H. S.; Chen, I. C.; Hsu, H. Y.; Tan, K. T. *ACS Chem. Biol.* **2014**, *9*, 2359.

(27) Lang, K.; Chin, J. W. *ACS Chem. Biol.* **2014**, *9*, 16.

(28) Although all KLE:R59W containing mutants demonstrated excellent spectrocopic characteristics, the PSB of KLE:R59W:L28W complexed with **1** exhibited the highest pK_a and thus was used for evaluating pH stability and also expression in cells (vida infra).

(29) It should be noted that although the iminium is susceptible to hydrolysis, we have found the bond between merocyanine **1** and our mutants to be remarkably stable. The characteristic wavelenght of the iminium bond was not lost during SDS treatment, although prolong incubation and SDS-PAGE analysis are not possible due to the hydrolytic nature of the Schiff base. We have also observed high stability at various pH (2−11). Nonetheless, it is prudent to note that the iminium bond can be reversibly hydrolyzed, and thus the protein−chromophore complex might not be stable to conditions such as electrophoresis or column chromatography.