



**HAL**  
open science

# The Glowing Panoply of Fluorogen-based Markers for Advanced Bioimaging

Arnaud Gautier

► **To cite this version:**

Arnaud Gautier. The Glowing Panoply of Fluorogen-based Markers for Advanced Bioimaging. Optogenetics: Light-driven Actuators and Light-emitting Sensors in Cell Biology, pp.41-62, 2018, 10.1039/9781788013284-00041 . hal-02425264

**HAL Id: hal-02425264**

**<https://hal.science/hal-02425264>**

Submitted on 30 Dec 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **The glowing panoply of fluorogen-based markers for advanced bioimaging**

Arnaud Gautier<sup>1,2,\*</sup>

<sup>1</sup> École Normale Supérieure, PSL Research University, UPMC Univ Paris 06, CNRS, Département de Chimie, PASTEUR, 24 rue Lhomond, 75005 Paris, France

<sup>2</sup> Sorbonne Universités, UPMC Univ Paris 06, ENS, CNRS, PASTEUR, 75005 Paris, France

Correspondence should be addressed to: [arnaud.gautier@ens.fr](mailto:arnaud.gautier@ens.fr)

## **Abstract**

Biological Sciences rely nowadays extensively on imaging tools to decipher the complexity of living organisms. Fluorescence microscopy allows the study of biological processes with an unprecedented temporal and spatial resolution. The revolution in fluorescence imaging has been the development of a large toolbox of fluorescent proteins able to reveal the abundance, position and dynamics of proteins. The fluorescence toolbox has been recently expanded with innovative reporters enabling to visualize proteins, and other biomolecules such as RNA, in new ways. These innovative reporters are bipartite systems composed of a genetically encoded tag forming a fluorescent complex with a small organic fluorogenic chromophore (also called fluorogens). This chapter is a user-oriented presentation of some of the most mature fluorogen-based markers available to biologists.

## 4.1 Introduction

Living systems are controlled by dynamic biological events tightly orchestrated in space and time. Our understanding of the inner workings of these complex machines deeply relies on our ability to observe how their constituents organize and interact. Nowadays, researchers can use a wide range of imaging modalities (e.g. optical microscopy, electron microscopy, mass-spectrometry imaging and optoacoustic imaging) to dissect the behavior of biological systems at various temporal and spatial scales. Among them, optical microscopy has spread in most biology labs, and has become unavoidable for addressing major questions in e.g. cancer biology, immunology, or for deciphering e.g. embryo development or brain function. The widespread adoption of optical microscopy results from the important efforts made to engineer microscopes that enable imaging with high speed, sensitivity and resolution while minimizing the side effects of the applied light; the most advanced fluorescence microscopes allow today the observation of dynamic systems in 3D, at subsecond resolution and at the diffraction limit or below.

Optical microscopy has encountered such a large success because of the concomitant advances in imaging probes able to fluorescently label tissues, cells and molecules with high selectivity. The main breakthrough for research imaging has been without question the discovery and development of the green fluorescent protein (GFP)<sup>1</sup>. For the first time, one could see and observe a broad range of specific proteins and cells in live specimens through straightforward genetic tagging techniques. The development of various color variants allowed furthermore effective multicolor imaging, and the design of various

classes of biosensors<sup>2-4</sup>. One remarkable application of fluorescent proteins is the mapping of the brain connectivity using the Brainbow technology, which allows one to mark (and track) individual neurons with several hundreds of hues using stochastic expression of multiple color variants<sup>5</sup>. Fluorescent proteins have also been essential for the development of localization-based super-resolution microscopy techniques: the discovery that some fluorescent proteins can switch from a dark to a bright state under light excitation has provided a unique way to generate subsets of isolated emitting fluorophores whose positions can be determined with subdiffraction accuracy<sup>6</sup>. The significance and importance of GFP-like fluorescent proteins for research in Life Sciences was acknowledged by awarding the Nobel Prize of Chemistry to Osamu Shimomura, Martin Chalfie and Roger Tsien in 2008 for the “discovery and development of the green fluorescent protein, GFP”.

By pushing the boundaries to more or more sophisticated and challenging observations, investigators have observed some limitations of GFP-like fluorescent proteins<sup>4,7,8</sup>. First, GFP-like fluorescent proteins are weakly fluorescent in low-oxygen environment; the full maturation of their chromophore includes cyclization, dehydration and oxidation of a triplet of amino acids (Ser–Tyr– Gly at position 65–67 in GFP), and strictly depends on molecular oxygen as cofactor. Second, GFP-like fluorescent proteins fluoresce tens of minutes (up to several hours) after folding because of the slow maturation of their chromophore, preventing e.g. real-time monitoring of protein synthesis. Third, GFP-like fluorescent proteins are rather large proteins (25–30 kDa) and some of them have tendency to oligomerize, which may lead to dysfunctional fusion proteins. Last, some

GFP-like fluorescent proteins display confounding photophysics such as photoswitching, kindling or dark state conversion, which may complicate the quantitative analysis of some experiments.

Over the last decade, alternative markers have been developed in order to push the limit of what is seeable. In this chapter, I focus on key developments relying on genetically encoded protein (or nucleic acid) tags forming fluorescent complex with small organic fluorogenic chromophores (also called fluorogens). Because of their fluorogenic properties, fluorogens are only fluorescent when bound to their cognate complementary tag, and are otherwise dark when free. Such fluorogenic labeling allows selective background-free imaging even in presence of an excess of free fluorogen, opening great prospects for imaging in complex samples such as tissues and whole organisms. Fluorogenic response usually results from changes in fluorescence quantum yield, spectral position or chromophore absorption coefficient induced by the change of environment undergone by the fluorogen upon binding.

This chapter is divided into two parts in function of whether the bipartite fluorescent markers incorporate a *natural* fluorogenic chromophore or a *synthetic* one. This *natural* vs. *synthetic* distinction is explained by the distinct pros and cons of the two approaches. Hijacking natural fluorogenic chromophores solves the issue of the delivery, as the chromophores are endogenously present in cells. Moreover, several classes of natural chromophore-binding proteins can serve as starting point to design fluorescent markers, facilitating thus engineering. These advantages are however counterbalanced by, first,

the small number of natural chromophores displaying fluorogenic properties, which can limit the potential engineering space, and, secondly, by the fact that the diversion of endogenous chromophores from their natural functions may engender cellular and physiological stress. On the other hand, synthetic fluorogenic chromophores have the clear advantage of being tailored with various spectral and photo-physical properties by molecular engineering, enabling to address biological questions with the molecular diversity offered by modern chemistry. Moreover, relying on the addition of an exogenous synthetic chromophore allows additional labeling refinement: as fluorescence is fully controlled by the applied concentration of fluorogen, semi-synthetic fluorogen-based markers enable on-demand applications in which fluorescence is desired only at a specific time or at a given density, opening great prospects for the design of innovative labeling protocols for advanced multiplexed and super-resolution imaging.

## **4.2 Fluorogen-Based Markers Engineered from Natural Photoreceptors**

In introduction of his Nobel Lecture, Roger Tsien told how he considered back in the eighties, before the first use of GFP as fluorescent marker, phycobiliproteins as potential genetically encoded fluorescent markers<sup>9</sup>. Phycobiliproteins are a class of photoreceptors containing a phycocyanobilin (PCB) chromophore, which act as light-harvesting antennae in the photosynthetic system of blue-green algae and cyanobacteria. Phycobiliproteins were known to fluoresce under visible light, making them attractive for imaging applications. Tsien told however how he rapidly realized that the complex biosynthesis

and assembly of phycocyanobilin chromophore would limit the general applicability of such system, and thus abandoned the idea.

Progress in structural biology enabled in the last three decades to solve the structures of a large collection of photoreceptors, while developments in molecular biology provided scientists with various techniques to easily modify protein sequence. These developments allowed to push further the idea of turning chromophore-binding proteins into fluorescent proteins. Nature provides a large collection of chromophore-binding proteins, mainly photoreceptors, in which an endogenous chromophore (e.g. bilins, flavin, retinal, coumaric acid) is bound to the protein matrix<sup>10</sup>. In natural photoreceptors, the chromophore reacts to light illumination by e.g. photoreduction or photoisomerization, which induces a conformational change and initiates a signaling cascade. Most photoreceptors evolved to maximize the efficacy of these photocycles, and are thus weakly fluorescent. As fluorescence is a competing mechanism for dissipating light energy, introducing variations within the backbone of photoreceptors can impair their photocycle and increase thus their fluorescence properties. Figure 4.1 and Table 4.1 present the natural fluorogens and engineered reporters described in section 4.2.

#### 4.2.1 Flavin-Binding Cyan-Green Fluorescent Proteins

The idea of reformatting natural photoreceptors was used to transform small light, oxygen, and voltage (LOV) sensing domains – a class of blue-light photoreceptors found in plants, algae and bacteria – into cyan-green fluorescent proteins<sup>11-13</sup>. LOV proteins associate

with the ubiquitous cofactor flavin mononucleotide (FMN). Blue-light photoreception is believed to involve the reversible formation of a covalent adduct between FMN and a conserved cysteine within the FMN-binding pocket upon blue light illumination. The conformational change associated with the adduct formation activates then downstream signaling. Flavin-based Fluorescent Proteins (FbFP) were engineered by replacing the conserved cysteine of bacterial LOV domains by an alanine in order to inhibit the natural photocycle and reduce the fluorescence quenching of FMN<sup>14</sup>. Similarly the fluorescent protein iLOV was generated from the LOV2 domain of *Arabidopsis thaliana* phototropin 2<sup>15,16</sup>.

FbFPs gained particular attention because of their potential use as alternative to GFP in anaerobic conditions. Because fluorescence results only from the binding of FMN, an abundant cofactor in cells, FbFPs are fluorescent regardless of the level of oxygen. FbFPs proved to surpass GFP-like fluorescent proteins for reporting on protein expression in absence of oxygen in bacteria<sup>14,17,18</sup>, fungi<sup>19</sup> and mammalian cells<sup>20</sup>. FbFPs allowed in particular the study of host-pathogen interactions under physiologically relevant anaerobic conditions<sup>21,22</sup>. FbFP was moreover used to design a ratiometric Förster resonance energy transfer (FRET) oxygen sensor by fusing it to the yellow fluorescent protein (YFP)<sup>23</sup>. Oxygen is required for maturation of YFP (playing the role of FRET acceptor), but not for that of FbFP (the FRET donor). Consequently, the FRET efficiency directly gives the oxygen level: high FRET efficiency indicates high oxygen level, while low FRET efficiency means low oxygen level. FbFPs also proved to overcome GFP-like fluorescent proteins because of their smaller size: 12-16 kDa in average instead

of 25-30 kDa for GFP-like fluorescent proteins. iLOV was shown for instance to be better suited than GFP to label viruses and monitor plant infection because smaller genetic load maintain higher infectivity<sup>15</sup>.

Engineering of FbFBs further allowed the development of reporters that not only fluoresce but also generate singlet oxygen. The LOV2 domain of *Arabidopsis thaliana* phototropin 2 was transformed into MiniSOG (Mini Singlet Oxygen Generator), a fluorescent reporter generating high level of singlet oxygen and other reactive oxygen species (ROS) upon blue light illumination<sup>24</sup>. Because of this singular property, MiniSOG was used to promote local photooxidation of diaminobenzidine into osmiophilic insoluble polymer, which can be then stained with osmium for high resolution imaging by electron microscopy. As miniSOG also fluoresces green light upon light excitation, MiniSOG fusions can be observed both in fluorescence and electron microscopy, allowing thus correlative imaging. The ability of MiniSOG to generate ROS locally upon light illumination was also used to (i) promote cell death in cancer cells<sup>25</sup> and in neuronal cells in *C. elegans*<sup>26</sup>, (ii) silence genetically specified synapses by chromophore assisted light-inactivation (CALI) of synaptic proteins<sup>27</sup>, and (iii) map protein proximity in large protein complexes<sup>28</sup>.

#### 4.2.2 Biliverdin-Binding Far-Red and Infrared Fluorescent Proteins

Infrared fluorescent proteins have been a long time goal in the field of probe design because of their great potential for deep-tissue and whole body imaging. Infrared light scatters much less than visible light into tissue, and the window 650-900 nm is transparent

(hemoglobin, water and lipids do not absorb at these wavelength)<sup>29</sup>, enabling to image tissues at an unprecedented depth. The highest maximal emission wavelength found among red GFP-like fluorescent proteins is however only 650 nm (for mPlum and mNeptune) and is intrinsically limited by the structure of the chromophore.

The first infrared fluorescent protein IFP1.4 (with maximal emission at 708 nm) was engineered from the N-terminal PAS and GAF domains of *Deinococcus radiodurans* bacteriophytochrome<sup>30</sup>. This biliprotein photoreceptor, which incorporates biliverdin IX $\alpha$  (hereafter called biliverdin) as cofactor, regulates pigment synthesis to protect the bacterium from intense visible light. Biliverdin is covalently attached to the apo-receptor via a thioether bond with a nearby cysteine side chain through a self-catalytic process. Naturally, bacteriophytochromes sense light by reversible cis-trans photoisomerization of the C15=C16 double bond of biliverdin. Restricting the conformational freedom of biliverdin by introducing mutations that prevent photoisomerization to occur enabled to increase fluorescence significantly. Although biliverdin is ubiquitously present in mammals, optimal labeling of IFP1.4 in mammalian cells and mice required exogenous supply of biliverdin. Coexpression of heme-oxygenase (HO1) in charge of biliverdin synthesis proved to boost biliverdin levels, enabling to image the optimized IFP2.0 in *Drosophila* neurons (which contains low levels of biliverdin) and in mouse brain tumors without exogenous addition of biliverdin<sup>31</sup>.

The first infrared fluorescent protein that does not require addition of exogenous biliverdin was developed using a truncated version of the bacteriophytochrome *RpBphP2* from

*Rhodopseudomonas palustris*<sup>32</sup>. The higher affinity for biliverdin explains the higher performance of this improved near infrared fluorescent protein (iRFP). Engineering of iRFP variants with emission wavelengths going from 670 nm to 720 further expanded the color range available, opening new opportunities for multiplexed whole-body imaging<sup>33</sup>.

A limitation of the above mentioned proteins is however that they are either dimeric (iRFP) or form dimers at high concentrations (IFP1.4 and IFP2.0)<sup>34</sup>. Fully monomeric IFP (mIFP) was engineering from a monomeric truncated bacteriophytochrome from *Bradyrhizobium*<sup>34</sup>. mIFP proved to correctly label proteins in mammalian cells, flies and zebrafish. Brighter monomeric iRFPs (miRFPs) with emission wavelength going from 670 nm to 710 were next engineered from bacteriophytochrome *RpBphP1*<sup>35</sup>. miRFPs proved to be well suited for widefield and structure-illumination microscopy (SIM).

The family of near infrared fluorescent proteins was further extended by engineering the allophycocyanin  $\alpha$ -subunit from a cyanobacterial phycobiliprotein instead of a bacterial phytochrome<sup>36</sup>. Cyanobacterial phycobiliproteins are normally functionalized with phycocyanobilin (PCB) by a specific lyase. To develop a useful fluorescent protein, the protein was evolved to be self-sufficient (i.e. to not require any lyase) and to covalently bind biliverdin instead of PCB. The engineering process ultimately gave a bright protein designated small ultra-red fluorescent protein (smURFP). smURFP is a homodimer of 15 kDa subunits. It has an exceptional absorption coefficient ( $\epsilon = 180,000 \text{ M}^{-1}\text{cm}^{-1}$ ) and a modest fluorescence quantum yield ( $\phi = 0.18$ ), making it the brightest far-red/near-infrared fluorescent protein. Although the brightest fluorescence in cells is obtained with

addition of exogenous biliverdin (or the more cell-permeant biliverdin methyl ester), expression of smURFP was efficiently imaged in HT1080 tumor rodent xenographs in mice even without the addition of exogenous biliverdin<sup>36</sup>.

The great potential of infrared fluorescent proteins for *in vivo* imaging led to the development of various biosensors to sense specific substances/activities or protein-protein interactions. Most phytochrome-based biosensors are split- or insertion-based biosensors that rely on the reconstitution of the PAS-GAF assembly. A near-infrared split reporter for the detection of protein-protein interaction *in vivo* was designed by separating the PAS and GAF domains of iRFP<sup>37</sup>. The two domains only efficiently complement when in close proximity. Proximity-induced complementation promotes then biliverdin binding and thus fluorescence. Split iRFP allowed the visualization of known interactions in mice, however (i), as all biliverdin-based fluorescent proteins, the covalent attachment of biliverdin was slow, which led to a slow fluorescence maturation (several hours) in cells and (ii) the assembly was irreversible and dimeric, which prevented the study of dynamic processes or complex stoichiometry. A reversible infrared split system was obtained from IFP1.4. Split IFP1.4 allowed the monitoring in real-time of the disruption of protein-protein interactions in mammalian cells and yeast<sup>38</sup>. Split IFP1.4 is however less bright than split iRFP, which may limit its use for whole-body imaging. Recently, Verkhusha and coworkers obtained truly monomeric split reporters using miRFPs, which should allow the efficient screening of novel protein-protein interactions<sup>35</sup>. By designing split systems with distinct colors but sharing one split fragment, they were furthermore able to distinguish interactions of one given protein with two alternative partners<sup>35</sup>.

Apart from split-biosensors, infrared fluorescent proteins were also used to design protease biosensors<sup>39</sup>. IFP1.4 was modified so that protease activity promotes biliverdin attachment and therefore fluorescence. The uncleaved sensor (iProtease) is non-fluorescent because the cysteine involved in biliverdin attachment cannot react with biliverdin because of physical displacement. Proteolytic cleavage frees the cysteine, which returns in the biliverdin binding site, promoting thus fluorescence. This approach allowed the development of a caspase-3 sensor (iCasper) for the visualization of apoptosis<sup>39</sup>. iCasper enabled to study the spatiotemporal coordination between cell apoptosis and embryonic morphogenesis in *Drosophila*, and revealed the dynamics of apoptosis during tumorigenesis in the brain of *Drosophila*. Such design opens exciting prospects for *in vivo* biosensing, and demonstrates how infrared fluorescent proteins can allow the construction of fluorogenic biosensors for visualizing the spatiotemporal dynamics of cell signaling *in vivo*.

#### 4.2.3 Bilirubin-Binding Green Fluorescent Proteins

Although flavin- and biliverdin-based fluorescent proteins are man-made reporters obtained by engineering natural photoreceptors, natural evolution has also generated fluorescent proteins incorporating natural fluorogenic chromophore. A natural fluorogen-based green fluorescent protein was discovered in Japanese eel muscles<sup>40</sup>. The fluorescence properties of this small monomeric protein, called UnaG, result from the non-covalent, high affinity and specific binding of the fluorogenic bilirubin. Fluorescent holo

UnaG forms efficiently in mammalian cells because bilirubin, an endogenous catabolic product of heme, is present at high concentration in animals. Exogenous supply of bilirubin allowed the use of UnaG in organisms such as bacteria that do not produce bilirubin<sup>40</sup>.

UnaG displays the same advantages as FbFPs: UnaG is (i) twice smaller than GFP-like fluorescent proteins, (ii) almost instantaneously fluorescent upon chromophore binding, and (iii) well suited to visualize fusion proteins under anaerobic conditions (because the fluorescence maturation process is fully independent of molecular oxygen)<sup>40</sup>. UnaG is furthermore 5-fold brighter than FbFPs, making it one of the brightest alternatives to GFP. The oxygen independence of UnaG fluorescence was used to design genetically encoded hypoxia sensors for light microscopy<sup>41</sup>. In these sensors, the expression of destabilized versions of UnaG was under the control of hypoxia-responsive promoters. These sensors proved to be highly effective to visualize hypoxia in tumors, and allowed to reveal strong heterogeneity in tumor hypoxia at the cellular level.

### **4.3 Semi-synthetic fluorogen-based markers**

The modular nature of fluorogen-based markers enables a priori to tune the chromophore by molecular engineering, and address biological questions with the molecular diversity offered by modern chemistry. Synthetic fluorogenic chromophores with various spectral and physico-chemical properties have been mobilized for engineering new imaging probes<sup>42</sup>. Semi-synthetic fluorogen-based markers usually use design principles based

on conformational locking (of e.g. molecular rotors) or ground-state isomerization (of e.g. silicon rhodamines) to achieve fluorogenic response. In the case of molecular rotors, internal rotation disrupts the dye planarity and is a source of non-radiative relaxation; when bound to complementary receptors, the rotation is blocked and the fluorescence of these fluorogens is strongly enhanced. In the case of silicon rhodamines, the dye adopts a non-fluorescent spirolactone form in polar solvent, and a fluorescent zwitterionic open form at the vicinity of proteins because of the local drop of polarity. Compared to the fluorogen-based markers described in section 4.2, which rely on endogenous natural fluorogens, semi-synthetic markers allow additional labeling refinement since fluorescence can be fully controlled by the applied concentration of the exogenous fluorogen. This feature allows the development of on-demand applications in which fluorescence is desired only at a specific time or at a given density, opening great prospects for the design of innovative labeling protocols for advanced multiplexed and super-resolution imaging. Figure 4.2 and Table 4.2 present the synthetic fluorogens and engineered reporters described in section 4.3.

#### 4.3.1 Semi-synthetic fluorogen-based protein markers

##### *Fluorogen-Activating Proteins (FAPs)*

Activating the fluorescence of a fluorogen by molecular recognition requires the development of complementary protein receptors. Fluorogen-Activating Proteins (FAP) that generate fluorescence through immobilization of fluorogenic molecular rotors were first engineered from 25-30 kDa single-chain antibodies (scFvs). The screening of yeast-

displayed libraries of human scFvs by Fluorescence-Activated Cell Sorting (FACS) allowed the isolation of FAPs that bind non-covalently with high affinity modified variants of Malachite Green (MG) and Thiazole Orange (TO), two well-known fluorogenic molecular rotors that strongly fluoresce in constrained environment<sup>43</sup>. The generated systems fluoresce far-red light (for MG complex) and green-yellow light (for TO complex) with brightness levels as good as those encountered in GFP-like fluorescent proteins. To be fully active, the original FAPs needed however non-reducing environments to allow the formation of internal disulfide bond, limiting thus their use to the cell surface and secretory pathway<sup>43</sup>. Recently, disulfide-free FAPs were engineered for protein labeling in various reducing compartments, including the cytosol, using cell-permeant MG-ester<sup>44,45</sup>. Additional engineering efforts enabled to expand the available chromatic palette from the blue to the far-red edge of the visible spectrum<sup>46,47</sup>. Interestingly, FAP's fluorogens can be rendered poorly membrane-permeant by adding electronic charge. This property was used to selectively label cell-surface proteins<sup>48,49</sup>. Add-and-read protocols were developed to study the endocytosis and recycling of FAP-tagged receptors<sup>50</sup>. Pulse-chase with two fluorogens of different colors and different permeability properties further allowed the quantitative study of receptor recycling upon agonist activation<sup>51</sup>.

FAPs proved to display great potentials for super-resolution microscopy and single molecule tracking. Far-red MG-based FAPs are highly photostable, which allowed live cell imaging with stimulated emission depletion (STED) nanoscopy in mammalian cells and bacteria<sup>52,53</sup>. Furthermore, because the investigator can control at will fluorogen

concentration, it is possible to label only a subset of proteins independently of their expression level. This property was used for tracking single receptors on cell surface<sup>54</sup>, and for generating stochastic binding-based blinking in order to generate sparse subsets of emitters for the reconstruction of images of FAP-tagged proteins with sub-diffraction resolution<sup>55</sup>.

The FAP technology allowed also the design of photosensitizer molecules able to produce reactive oxygen species (ROS) upon illumination for applications such as inactivation through chromophore-assisted light inactivation (CALI) of directly linked proteins, targeted cell damage or cellular ablation. A genetically encoded FAP that binds a heavy-atom-containing fluorogenic MG dye forms an active photosensitizer that produces efficiently singlet oxygen when activated by near-infrared light<sup>56</sup>. Unlike MiniSOG presented in paragraph 4.2.1 that is constitutively active, this FAP-based photosensitizer can be activated on-demand by addition of the fluorogenic dye. Interestingly, FAP-based photosensitizers display near-infrared excitation and emission, which provides a new spectral range for photosensitization and opens thus great prospects for imaging, protein and cell manipulation, and cellular ablation in whole organisms. Beyond the use for protein inactivation, FAP-based photosensitizers enabled CALI of proteins, targeted cell killing, and targeted lineage ablation in zebrafish<sup>56</sup>.

### *Self-labeling tags*

Semi-synthetic fluorogen-based markers were also obtained by exploiting site-specific labeling systems such as SNAP-tag/CLIP-tag and Halo-tag, which react covalently with

specific substrates bearing chemical probes. SNAP-tag is a 20 kDa protein evolved from the human DNA repair protein *O*<sup>6</sup>-alkylguanine-DNA alkyltransferase (AGT)<sup>57,58</sup>. SNAP-tag transfers the functionalized benzyl group of *O*<sup>6</sup>-benzylguanine (BG) derivatives to its active site cysteine, thus allowing irreversible covalent labeling of fusion proteins. SNAP-tag accepts a broad variety of chemical functionalities on BG, making it one of the most versatile tags currently available<sup>59</sup>. CLIP-tag, an engineered variant of SNAP-tag, reacts selectively with *O*<sup>2</sup>-benzylcytosine (BC) substrates instead of BG<sup>60</sup>. Halo-tag is a 33 kDa protein engineered from a bacterial haloalkane dehydrogenase that covalently binds chloroalkane ligands<sup>61</sup>.

The use of fluorogenic chromophores instead of permanent fluorophores allowed the development of labeling protocols with no-washing steps, as free unreacted substrates do not fluoresce. Beyond increasing contrast, the use of fluorogenic substrates allowed to get rid of the need for extensive washing steps increasing thus the temporal resolution. Silicon-rhodamine (SiR) derivatives were used to design fluorogenic substrates for SNAP-tag, CLIP-tag and Halo-tag<sup>62</sup>. The fluorogenic response of SiR relies on ground-state isomerization that breaks the dye conjugation: in aqueous solution, SiR adopt mainly a closed UV absorbing spirolactone form, while it undergoes ring opening in less polar environment such as protein vicinity. The open zwitterionic form absorbs at 640-650 nm and fluoresces in the far-red at 660-670 nm. SiR-based substrates proved to be highly efficient to label SNAP-, CLIP- and Halo-tagged proteins in various organelles of living cells within 30-60 minutes with no significant background<sup>62</sup>. Because of their excellent spectroscopic properties, SiR substrates allowed live-cell super-resolution microscopy of

biological structures using GSDIL (ground-state depletion followed by individual molecule return)<sup>63</sup> and dSTORM (direct stochastic optical reconstruction microscopy)<sup>64</sup> by relying on the stochastic blinking of these conventional fluorophores.

Recently near-infrared SiR analogs with emission and excitation wavelength maxima at 680 and 715 nm were reported, allowing multicolor super-resolution imaging<sup>65</sup>. Moreover, SiR brightness and photostability were improved by incorporating azetidine four-member ring. The resulting dye, Janelia Fluor 646 (JF<sub>646</sub>), was shown to be an efficient label for imaging SNAP- and Halo-tagged proteins by conventional and super-resolution microscopies<sup>66</sup>. Refinement and extension of this strategy allowed the development of fluorogenic labels with excitation ranging from orange (JF<sub>585</sub>) to red (JF<sub>635</sub>)<sup>67</sup>. These dyes are cell-permeant and exhibit very high fluorogenicity, which allowed protein labeling in neural tissues in explants and in *Drosophila* larva<sup>67</sup>, and opens great prospects for imaging deep structures in whole animals.

#### *Fluorescence-Activating and Absorption Shifting Tag*

Fluorescence-activating and absorption-shifting tag (FAST) is a small protein tag of 14 kDa that binds and switches on the fluorescence of hydroxybenzylidene rhodanine (HBR) analogs through conformational locking<sup>68</sup>. FAST was evolved from the apo photoactive yellow protein (apo-PYP) by directed devolution using yeast display and FACS sorting. The fluorogenicity of HBR analogs originates from their push-pull structure composed of an electron-donating phenol ring conjugated with an electron-withdrawing rhodanine heterocycle. In solution, these fluorogens dissipate light energy non-radiatively through

internal rotation or cis-trans isomerization. Binding to FAST locks the fluorogen, which slows down non-radiative decay and strongly increases fluorescence. Upon binding, the fluorogen undergoes also a 80 nm red-shift in absorption due to a selective deprotonation reaction. Consequently, the free fluorogen barely absorbs at the wavelength used for exciting the bipartite complex, further enhancing the fluorogenic response.

FAST proved to be highly effective to fluorescently label proteins in living cells (bacteria, yeast, mammalian cells) in a large number of organelles and subcellular localizations<sup>68</sup>. Full labeling is achieved within few seconds after fluorogen addition in living cells. Provided that the fluorogen is present, FAST is fluorescent instantaneously after folding because of fast binding kinetics, which allows to follow fast processes in near real-time. Because HBR analogs are highly cell-permeant, efficient labeling of FAST were observed in multicellular organisms such as zebrafish embryo, opening great prospect for in vivo imaging.

FAST distinguishes itself from other fluorogen-based markers because fluorogen binding is non-covalent, highly dynamic and fully reversible (because of a high dissociation rate constant). Fluorogen washing allows one to reverse labeling and switch off fluorescence within few seconds in cells, making FAST a fluorescence switch that can be efficiently switched on or off at will by addition or removal of fluorogen. The rapid exchange dynamics further allows efficient fluorogen renewal, which reduces the apparent photobleaching rate<sup>69</sup>. Finally, because of its fast exchange dynamics, FAST behaves as

a blinking fluorophore at the single molecule level. Such spontaneous stochastic blinking might find very interesting applications for super-resolution microscopy in live cells.

Ongoing efforts are further testing the ability to expand the spectral properties of FAST. Recently, FAST emission color was extended to the orange and red by modifying the structure of its complementary fluorogen<sup>70</sup>. The ability to make FAST fluoresce green-yellow, orange, or red light by a simple change of fluorogen enables one to adapt the color of FAST to the experimental spectral constraints without the need for recloning the tag, providing an experimental versatility not encountered with GFP-like fluorescent proteins. The ability to dynamically swap color by exchanging fluorogens was furthermore used as a unique kinetic signature to selectively image FAST in spectrally crowded environments. By evaluating the degree of temporal anticorrelation of the green and red fluorescence signals upon color swapping, FAST-tagged proteins could be selectively detected in cells already tagged with green and red reporters, illustrating the general potential of non-covalent fluorogenic reporters for the development of new innovative imaging methods for advanced biological imaging.

#### 4.3.2 Semi-synthetic fluorogen-based RNA markers

The idea of forming a fluorescent marker by association of a genetically encoded module and a fluorogenic chromophore is a very general idea that goes well beyond the labeling of proteins. Fluorogenic labeling was used to expand fluorescent labeling to more diverse cellular molecules such as RNA. Various studies showed that fluorescence could be

generated by engineered RNA aptamers binding selectively fluorogenic chromophores<sup>71-78</sup>, paving the road towards new ways to image RNA in living cells<sup>79-81</sup>.

Efficient RNA imaging in live cells was made possible with the development of an RNA aptamer mimic of GFP named Spinach<sup>82,83</sup>, and its optimized versions Spinach 2<sup>84</sup> and Broccoli<sup>85</sup>. These engineered RNA aptamers form fluorescent complexes with analogs of the fluorogenic 3,5-difluoro-4-hydroxybenzylidene imidazolidinone (DFHBI). DFHBI is related to the chromophore of GFP known to be only fluorescent when encased inside the GFP barrel. Likewise DFHBI only fluoresces when bound to Spinach RNA aptamers<sup>86</sup>, allowing to imaged Spinach-tagged RNA in living cells with high contrast.

Spinach aptamers proved to be highly efficient to design biosensors by coupling them with aptameric sensing units. Sensors, in which the binding of a given analyte (e.g. metabolite or protein) promotes fluorogen binding and activation through conformational coupling, enabled to visualize the dynamics of the analyte levels in bacteria<sup>87-90</sup>, further demonstrating the great potential of such fluorogen-based markers for the design of sensors able to sense the abundance, distribution and flux of intracellular molecules. Broccoli was also used to design a fluorimetric assay to measure the activity of RNA-modifying enzymes in cells<sup>91</sup>. Broccoli was modified to contain *N*<sup>6</sup>-methyladenosine, a prevalent mRNA base modification. Methylation renders Broccoli non-fluorescent; fluorescence can be recovered by action of RNA demethylases. This approach allowed the development of high-throughput screens for inhibitors of the RNA demethylase fat mass and obesity-associated protein (FTO).

RNA mimics of red fluorescent proteins (RFP) were obtained by similar engineering strategies. These fluorogen-based RNA markers, named corn, orange broccoli and red broccoli, bind 3,5-difluoro-4-hydroxybenzylidene-imidazolinone-2-oxime (DFHO), a fluorogen resembling the chromophore found in RFP<sup>92,93</sup>. Corn, Orange Broccoli and Red Broccoli form fluorescent complexes with DFHO displaying red-shifted emissions (respectively 545 nm, 562 nm and 582 nm) with respect to Spinach. Corn:DFHO showed high photostability, unlike Spinach and Broccoli that undergo fast reversible photobleaching<sup>94,95</sup>, allowing quantitative fluorescence imaging of mTOR-dependent Pol III transcription.

#### **4.4 Concluding remarks**

Fluorogen-based markers allow more and more sophisticated and challenging observations in living cells and organisms. They are very attractive alternatives of canonical fluorescent proteins because they open exciting new possibilities for whole body imaging, biosensor design, multiplex imaging, and high-resolution imaging. Fluorogen-based markers have not shown yet their full potentials, and it is a safe bet to say that the coming years will see further exciting and unpredictable advances.

## **Acknowledgments**

The European Research Council (ERC-2016-CoG-724705 FLUOSWITCH), France BioImaging (ANR-10-INBS-04) and the Equipex Morphoscope 2 (ANR-11-EQPX-0029) have supported this work.

## References

- 1 R. Y. Tsien, *Annu Rev Biochem*, 1998, **67**, 509–544.
- 2 N. Shaner, P. Steinbach and R. Tsien, *Nat. Meth.*, 2005, **2**, 905–909.
- 3 D. M. Chudakov, M. V. Matz, S. Lukyanov and K. A. Lukyanov, *Physiol Rev*, 2010, **90**, 1103–1163.
- 4 E. A. Rodriguez, R. E. Campbell, J. Y. Lin, M. Z. Lin, A. Miyawaki, A. E. Palmer, X. Shu, J. Zhang and R. Y. Tsien, *Trends Biochem Sci*, 2017, **42**, 111–129.
- 5 J. Livet, T. A. Weissman, H. Kang, R. W. Draft, J. Lu, R. A. Bennis, J. R. Sanes and J. W. Lichtman, *Nature*, 2007, **450**, 56–62.
- 6 E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz and H. F. Hess, *Science*, 2006, **313**, 1642–1645.
- 7 S. J. Remington, *Curr Opin Struct Biol*, 2006, **16**, 714–721.
- 8 M. Zimmer, *Chem Rev*, 2002, **102**, 759–782.
- 9 R. Y. Tsien, *Angew Chem Int Ed*, 2009, **48**, 5612–5626.
- 10 M. A. van der Horst and K. J. Hellingwerf, *Acc Chem Res*, 2004, **37**, 13–20.
- 11 A. Mukherjee and C. M. Schroeder, *Curr Opin Biotechnol*, 2015, **31**, 16–23.
- 12 T. Drepper, T. Gensch and M. Pohl, *Photochem Photobiol Sci*, 2013, **12**, 1125–1134.
- 13 M. Wingen, J. Potzkei, S. Endres, G. Casini, C. Rupprecht, C. Fahlke, U. Krauß, K.-E. Jaeger, T. Drepper and T. Gensch, *Photochem Photobiol Sci*, 2014, **13**, 875–9.
- 14 T. Drepper, T. Eggert, F. Circolone, A. Heck, U. Krauß, J.-K. Guterl, M. Wendorff, A. Losi, W. Gärtner and K.-E. Jaeger, *Nat Biotechnol*, 2007, **25**, 443–445.

- 15 S. Chapman, C. Faulkner, E. Kaiserli, C. Garcia-Mata, E. I. Savenkov, A. G. Roberts, K. J. Oparka and J. M. Christie, *Proc Natl Acad Sci USA*, 2008, **105**, 20038–20043.
- 16 J. M. Christie, K. Hitomi, A. S. Arvai, K. A. Hartfield, M. Mettlen, A. J. Pratt, J. A. Tainer and E. D. Getzoff, *J Biol Chem*, 2012, **287**, 22295–22304.
- 17 T. Drepper, R. Huber, A. Heck, F. Circolone, A. K. Hillmer, J. Buchs and K. E. Jaeger, *Appl Environ Microbiol*, 2010, **76**, 5990–5994.
- 18 A. M. Buckley, C. Jukes, D. Candlish, J. J. Irvine, J. Spencer, R. P. Fagan, A. J. Roe, J. M. Christie, N. F. Fairweather and G. R. Douce, *Sci. Rep.*, 2016, **6**, 23463.
- 19 D. Tielker, I. Eichhof, K. E. Jaeger and J. F. Ernst, *Eukaryotic Cell*, 2009, **8**, 913–915.
- 20 J. Walter, S. Hausmann, T. Drepper, M. Puls, T. Eggert and M. Dihn e, *PLoS ONE*, 2012, **7**, e43921.
- 21 L. A. Lobo, C. J. Smith and E. R. Rocha, *FEMS Microbiology Letters*, 2011, **317**, 67–74.
- 22 C. H. Choi, J. V. DeGuzman, R. J. Lamont and  . Yilmaz, *PLoS ONE*, 2011, **6**, e18499–8.
- 23 J. Potzkei, M. Kunze, T. Drepper, T. Gensch, K.-E. Jaeger and J. B uchs, *BMC Biol*, 2012, **10**, 28.
- 24 X. Shu, V. Lev-Ram, T. J. Deerinck, Y. Qi, E. B. Ramko, M. W. Davidson, Y. Jin, M. H. Ellisman and R. Y. Tsien, *PLoS Biol*, 2011, **9**, e1001041.
- 25 A. P. Ryumina, E. O. Serebrovskaya, M. V. Shirmanova, L. B. Snopova, M. M. Kuznetsova, I. V. Turchin, N. I. Ignatova, N. V. Klementieva, A. F. Fradkov, B. E.

- Shakhov, E. V. Zagaynova, K. A. Lukyanov and S. A. Lukyanov, *BBA - General Subjects*, 2013, **1830**, 5059–5067.
- 26 Y. B. Qi, E. J. Garren, X. Shu, R. Y. Tsien and Y. Jin, *Proc Natl Acad Sci USA*, 2012, **109**, 7499–7504.
- 27 J. Y. Lin, S. B. Sann, K. Zhou, S. Nabavi, C. D. Proulx, R. Malinow, Y. Jin and R. Y. Tsien, *Neuron*, 2013, **79**, 241–253.
- 28 T.-L. To, M. J. Fadul and X. Shu, *Nat Commun*, 2014, **5**, 4072.
- 29 V. Marx, *Nat Meth*, 2014, **11**, 717–720.
- 30 X. Shu, A. Royant, M. Z. Lin, T. A. Aguilera, V. Lev-Ram, P. A. Steinbach and R. Y. Tsien, *Science*, 2009, **324**, 804–807.
- 31 D. Yu, W. C. Gustafson, C. Han, C. E. L. Lafaye, M. Noirclerc-Savoye, W.-P. Ge, D. A. Thayer, H. Huang, T. B. Kornberg, A. Royant, L. Y. Jan, Y. N. Jan, W. A. Weiss and X. Shu, *Nat Commun*, 2014, **5**, 1–7.
- 32 G. S. Filonov, K. D. Piatkevich, L.-M. Ting, J. Zhang, K. Kim and V. V. Verkhusha, *Nat Biotechnol*, 2011, **29**, 759–763.
- 33 D. M. Shcherbakova and V. V. Verkhusha, *Nat Meth*, 2013, **10**, 751–754.
- 34 D. Yu, M. A. Baird, J. R. Allen, E. S. Howe, M. P. Klassen, A. Reade, K. Makhijani, Y. Song, S. Liu, Z. Murthy, S.-Q. Zhang, O. D. Weiner, T. B. Kornberg, Y. N. Jan, M. W. Davidson and X. Shu, *Nat. Meth.*, 2015, **12**, 763–765.
- 35 D. M. Shcherbakova, M. Baloban, A. V. Emelyanov, M. Brenowitz, P. Guo and V. V. Verkhusha, *Nat Commun*, 2016, **7**, 12405.
- 36 E. A. Rodriguez, G. N. Tran, L. A. Gross, J. L. Crisp, X. Shu, J. Y. Lin and R. Y. Tsien, *Nat. Meth.*, 2016, **13**, 763–769.

- 37 G. S. Filonov and V. V. Verkhusha, *Chem Biol*, 2013, **20**, 1078–1086.
- 38 E. Tchekanda, D. Sivanesan and S. W. Michnick, *Nat. Meth.*, 2014, **11**, 641–644.
- 39 T.-L. To, B. J. Piggott, K. Makhijani, D. Yu, Y. N. Jan and X. Shu, *Proc Natl Acad Sci USA*, 2015, **112**, 3338–3343.
- 40 A. Kumagai, R. Ando, H. Miyatake, P. Greimel, T. Kobayashi, Y. Hirabayashi, T. Shimogori and A. Miyawaki, *Cell*, 2013, **153**, 1602–1611.
- 41 R. Erapanedi, V. V. Belousov, M. Schaffers and F. Kiefer, *EMBO J*, 2016, **35**, 102–113.
- 42 A. S. Klymchenko, *Acc Chem Res*, 2017, **50**, 366–375.
- 43 C. Szent-Gyorgyi, B. A. Schmidt, Y. Creeger, G. W. Fisher, K. L. Zakel, S. Adler, J. A. J. Fitzpatrick, C. A. Woolford, Q. Yan, K. V. Vasilev, P. B. Berget, M. P. Bruchez, J. W. Jarvik and A. Waggoner, *Nat Biotechnol*, 2008, **26**, 235–240.
- 44 B. P. Yates, M. A. Peck and P. B. Berget, *Mol Biotechnol*, 2012, **54**, 829–841.
- 45 C. A. Telmer, R. Verma, H. Teng, S. Andreko, L. Law and M. P. Bruchez, *ACS Chem Biol*, 2015, **10**, 1239–1246.
- 46 H. Ozhalici-Unal, C. L. Pow, S. A. Marks, L. D. Jesper, G. L. Silva, N. I. Shank, E. W. Jones, J. M. Burnette, P. B. Berget and B. A. Armitage, *J. Am. Chem. Soc.*, 2008, **130**, 12620–12621.
- 47 K. J. Zanotti, G. L. Silva, Y. Creeger, K. L. Robertson, A. S. Waggoner, P. B. Berget and B. A. Armitage, *Org Biomol Chem*, 2011, **9**, 1012–1020.
- 48 J. Holleran, D. Brown, M. H. Fuhrman, S. A. Adler, G. W. Fisher and J. W. Jarvik, *Cytometry A*, 2010, **77**, 776–782.
- 49 G. W. Fisher, S. A. Adler, M. H. Fuhrman, A. S. Waggoner, M. P. Bruchez and J.

- W. Jarvik, *J Biomol Screen*, 2010, **15**, 703–709.
- 50 Q. Yan, B. F. Schmidt, L. A. Perkins, M. Naganbabu, S. Saurabh, S. K. Andreko and M. P. Bruchez, *Org Biomol Chem*, 2015, **13**, 2078–2086.
- 51 C. P. Pratt, J. He, Y. Wang, A. L. Barth and M. P. Bruchez, *Bioconjug Chem*, 2015, **26**, 1963–1971.
- 52 J. A. J. Fitzpatrick, Q. Yan, J. J. Sieber, M. Dyba, U. Schwarz, C. Szent-Gyorgyi, C. A. Woolford, P. B. Berget, A. S. Waggoner and M. P. Bruchez, *Bioconjug Chem*, 2009, **20**, 1843–1847.
- 53 S. Saurabh, A. M. Perez, C. J. Comerchi, L. Shapiro and W. E. Moerner, *J Am Chem Soc*, 2016, **138**, 10398–10401.
- 54 S. L. Schwartz, Q. Yan, C. A. Telmer, K. A. Lidke, M. P. Bruchez and D. S. Lidke, *ACS Chem Biol*, 2015, **10**, 539–546.
- 55 Q. Yan, S. L. Schwartz, S. Maji, F. Huang, C. Szent-Gyorgyi, D. S. Lidke, K. A. Lidke and M. P. Bruchez, *ChemPhysChem*, 2014, **15**, 687–695.
- 56 J. He, Y. Wang, M. A. Missinato, E. Onuoha, L. A. Perkins, S. C. Watkins, C. M. St Croix, M. Tsang and M. P. Bruchez, *Nat Meth*, 2016, **13**, 263–268.
- 57 A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel and K. Johnsson, *Nat Biotechnol*, 2003, **21**, 86–89.
- 58 A. Keppler, H. Pick, C. Arrivoli, H. Vogel and K. Johnsson, *Proc Natl Acad Sci USA*, 2004, **101**, 9955–9959.
- 59 A. Gautier, K. Johnsson and H. O'Hare, *AGT/SNAP-Tag: A Versatile Tag for Covalent Protein Labeling*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2008.

- 60 A. Gautier, A. Juillerat, C. Heinis, I. R. Corrêa Jr., M. Kindermann, F. Beaufile and K. Johnsson, *Chem Biol*, 2008, **15**, 128–136.
- 61 G. Los, L. Encell, M. McDougall, D. Hartzell, N. Karassina, C. Zimprich, M. Wood, R. Learish, R. Ohana, M. Urh, D. Simpson, J. Mendez, K. Zimmerman, P. Otto, G. Vidugiris, J. Zhu, A. Darzins, D. Klaubert, R. Bulleit and K. Wood, *ACS Chem Biol*, 2008, **3**, 373–382.
- 62 G. Lukinavicius, K. Umezawa, N. Olivier, A. Honigmann, G. Yang, T. Plass, V. Mueller, L. Reymond, I. R. Corrêa Jr, Z.-G. Luo, C. Schultz, E. A. Lemke, P. Heppenstall, C. Eggeling, S. Manley and K. Johnsson, *Nat Chem*, 2013, **5**, 132–139.
- 63 J. Fölling, M. Bossi, H. Bock, R. Medda, C. A. Wurm, B. Hein, S. Jakobs, C. Eggeling and S. W. Hell, *Nat Meth*, 2008, **5**, 943–945.
- 64 M. Heilemann, S. van de Linde, M. Schüttpelz, R. Kasper, B. Seefeldt, A. Mukherjee, P. Tinnefeld and M. Sauer, *Angew Chem Int Ed*, 2008, **47**, 6172–6176.
- 65 G. Lukinavicius, L. Reymond, K. Umezawa, O. Sallin, E. D'Este, F. Göttfert, H. Ta, S. W. Hell, Y. Urano and K. Johnsson, *J. Am. Chem. Soc.*, 2016, **138**, 9365–9368.
- 66 J. B. Grimm, B. P. English, J. Chen, J. P. Slaughter, Z. Zhang, A. Revyakin, R. Patel, J. J. Macklin, D. Normanno, R. H. Singer, T. Lionnet and L. D. Lavis, *Nat. Meth.*, 2015, **12**, 244–250.
- 67 J. B. Grimm, A. K. Muthusamy, Y. Liang, T. A. Brown, W. C. Lemon, R. Patel, R. Lu, J. J. Macklin, P. J. Keller, N. Ji and L. D. Lavis, *Nat. Meth.*, 2017, **14**, 987–994.
- 68 M.-A. Plamont, E. Billon-Denis, S. Maurin, C. Gauron, F. M. Pimenta, C. G. Specht, J. Shi, J. Querard, B. Pan, J. Rossignol, K. Moncoq, N. Morellet, M. Volovitch, E.

- Lescop, Y. Chen, A. Triller, S. Vriza, T. Le Saux, L. Jullien and A. Gautier, *Proc Natl Acad Sci USA*, 2016, **113**, 497–502.
- 69 F. M. Pimenta, G. Chiappetta, T. Le Saux, J. Vinh, L. Jullien and A. Gautier, *Sci. Rep.*, 2017, **7**, 12316.
- 70 C. Li, M.-A. Plamont, H. L. Sladitschek, V. Rodrigues, I. Aujard, P. Neveu, T. Le Saux, L. Jullien and A. Gautier, *Chem Sci*, 2017, **8**, 5598–5605.
- 71 J. R. Babendure, S. R. Adams and R. Y. Tsien, *J Am Chem Soc*, 2003, **125**, 14716–14717.
- 72 M. N. Stojanovic and D. M. Kolpashchikov, *J Am Chem Soc*, 2004, **126**, 9266–9270.
- 73 B. A. Sparano and K. Koide, *J Am Chem Soc*, 2005, **127**, 14954–14955.
- 74 S. Sando, A. Narita, M. Hayami and Y. Aoyama, *Chem Commun (Camb)*, 2008, 3858–4.
- 75 T. P. Constantin, G. L. Silva, K. L. Robertson, T. P. Hamilton, K. Fague, A. S. Waggoner and B. A. Armitage, *Org Lett*, 2008, **10**, 1561–1564.
- 76 R. Pei, J. Rothman, Y. Xie and M. N. Stojanovic, *Nucleic Acids Res.*, 2009, **37**, e59–e59.
- 77 J. Lee, K. H. Lee, J. Jeon, A. Dragulescu-Andrasi, F. Xiao and J. Rao, *ACS Chem Biol*, 2010, **5**, 1065–1074.
- 78 E. V. Dolgosheina, S. C. Y. Jeng, S. S. S. Panchapakesan, R. Cojocar, P. S. K. Chen, P. D. Wilson, N. Hawkins, P. A. Wiggins and P. J. Unrau, *ACS Chem Biol*, 2014, **9**, 2412–2420.
- 79 S. Tyagi, *Nat. Meth.*, 2009, **6**, 331–338.

- 80 B. A. Armitage, *Curr Opin Chem Biol*, 2011, **15**, 806–812.
- 81 J. Ouellet, *Front. Chem.*, 2016, **4**, e144–12.
- 82 J. S. Paige, K. Y. Wu and S. R. Jaffrey, *Science*, 2011, **333**, 642–646.
- 83 W. Song, R. L. Strack, N. Svensen and S. R. Jaffrey, 2014, **136**, 1198–1201.
- 84 R. L. Strack, M. D. Disney and S. R. Jaffrey, *Nat Meth*, 2013.
- 85 G. S. Filonov, J. D. Moon, N. Svensen and S. R. Jaffrey, 2014, **136**, 16299–16308.
- 86 K. D. Warner, M. C. Chen, W. Song, R. L. Strack, A. Thorn, S. R. Jaffrey and A. R. Ferré-D'Amaré, *Nat Struct Mol Biol*, 2014, **21**, 658–663.
- 87 J. S. Paige, T. Nguyen-Duc, W. Song and S. R. Jaffrey, *Science*, 2012, **335**, 1194–1194.
- 88 W. Song, R. L. Strack and S. R. Jaffrey, *Nat Meth*, 2013, **10**, 873–875.
- 89 M. You, J. L. Litke and S. R. Jaffrey, *Proc Natl Acad Sci USA*, 2015, **112**, E2756–65.
- 90 C. A. Kellenberger, S. C. Wilson, J. Sales-Lee and M. C. Hammond, *J Am Chem Soc*, 2013, **135**, 4906–4909.
- 91 N. Svensen and S. R. Jaffrey, *Cell Chem Biol*, 2016, **23**, 415–425.
- 92 W. Song, G. S. Filonov, H. Kim, M. Hirsch, X. Li, J. D. Moon and S. R. Jaffrey, *Nat Chem Biol*, 2017, **13**, 1187–1194.
- 93 K. D. Warner, L. Sjekloća, W. Song, G. S. Filonov, S. R. Jaffrey and A. R. Ferré-D'Amaré, *Nat Chem Biol*, 2017, **13**, 1195–1201.
- 94 P. Wang, J. Querard, S. Maurin, S. S. Nath, T. Le Saux, A. Gautier and L. Jullien, *Chem Sci*, 2013, **4**, 2865–2873.
- 95 K. Y. Han, B. J. Leslie, J. Fei, J. Zhang and T. Ha, *J. Am. Chem. Soc.*, 2013, **135**,

19033–19038.

## Table captions

Table 4.1 Fluorogen-based markers engineered from natural photoreceptors. Abbreviations are as follows:  $\lambda_{\text{abs}}$ , wavelength of maximal absorption;  $\lambda_{\text{em}}$ , wavelength of maximal emission;  $\epsilon$ , molar absorption coefficient at  $\lambda_{\text{abs}}$ ;  $\phi$ , fluorescence quantum yield. Structures of the fluorogens are given on Figure 4.1.

Table 4.2 Semi-synthetic fluorogen-based protein markers. Abbreviations are as follows:  $\lambda_{\text{abs}}$ , wavelength of maximal absorption;  $\lambda_{\text{em}}$ , wavelength of maximal emission;  $\epsilon$ , molar absorption coefficient at  $\lambda_{\text{abs}}$ ;  $\phi$ , fluorescence quantum yield. Structures of the fluorogens are given on Figure 4.2.

## Figure captions

Figure 4.1 Natural fluorogens found in fluorogen-based markers. Physico-chemical properties of the corresponding fluorescent markers are given in Table 4.1.

Figure 4.2 Synthetic fluorogens found in fluorogen-based markers. Physico-chemical properties of the corresponding fluorescent markers are given in Table 4.2.

**Table 4.1**

Tag	Fluorogen	Binding mode	Oligomeric state	$\lambda_{\text{abs}}$ (nm)	$\lambda_{\text{em}}$ (nm)	$\epsilon$ ( $\text{M}^{-1}\text{cm}^{-1}$ )	$\phi$ (%)	Ref.
BsFbFP	FMN	Non-covalent	Dimer	449	495	13,900	39	13
EcFbFP	FMN	Non-covalent	Dimer	448	496	14,500	44	13
PpFbFP	FMN	Non-covalent	Dimer	450	496	13,900	27	13
iLOV	FMN	Non-covalent	Monomer	447	497		44	15
phiLOV2.1	FMN	Non-covalent	Monomer	450	497		20	13
miniSOG	FMN	Non-covalent	Monomer	447	497	14,200	41	13
IFP1.4	Biliverdin	Covalent	Dimer	684	708	92,000	7.0	30
iRFP	Biliverdin	Covalent	Dimer	692	713	105,000	5.9	32
iRFP670	Biliverdin	Covalent	Dimer	643	670	114,000	11.1	33
iRFP682	Biliverdin	Covalent	Dimer	663	682	90,000	11.3	33
iRFP702	Biliverdin	Covalent	Dimer	673	702	93,000	8.2	33
iRFP720	Biliverdin	Covalent	Dimer	702	720	96,000	6.0	33
IFP2.0	Biliverdin	Covalent	Dimer	690	711	98,000	8.1	31
mIFP	Biliverdin	Covalent	Monomer	683	705	82,000	8.4	34
miRFP670	Biliverdin	Covalent	Monomer	642	670	87,400	14	35
miRFP703	Biliverdin	Covalent	Monomer	674	703	90,900	8.6	35
miRFP709	Biliverdin	Covalent	Monomer	683	709	78,400	5.4	35
smURFP	Biliverdin	Covalent	Dimer	642	670	180,000	18	36
UnaG	Bilirubin	Non-covalent	Monomer	498	527	77,300	51	40

**Table 4.2**

Tag	Fluorogen	Binding mode	Oligomeric state	$\lambda_{\text{abs}}$ (nm)	$\lambda_{\text{em}}$ (nm)	$\epsilon$ ( $\text{M}^{-1}\text{cm}^{-1}$ )	$\phi$ (%)	Ref.
FAP HL1.01	Thiazole Orange	Non-covalent	Monomer	509	530	60,000	47	43
FAP H6	Malachite Green	Non-covalent	Monomer	635	656	105,000	25	43
Self-labeling proteins	SiR650	Covalent	Monomer	645	661	100,000	39	62
Self-labeling proteins	SiR700	Covalent	Monomer	687	716	100,000		65
Self-labeling proteins	JF646	Covalent	Monomer	646	664	152,000	54	67
Self-labeling proteins	JF585	Covalent	Monomer	585	609	156,000	78	67
Self-labeling proteins	JF635	Covalent	Monomer	635	652	167,000	56	67
FAST	HMBR	Non-covalent	Monomer	481	540	45,000	23	70
FAST	HBR-3,5DM	Non-covalent	Monomer	499	562	48,000	49	70
FAST	HBR-3,5DOM	Non-covalent	Monomer	518	600	39,000	31	70
Spinach2	DFHBI	Non-covalent		447	501	22,000	72	83
Spinach2	DFHBI-1T	Non-covalent		482	505	31,000	94	83
Broccoli	DFHBI-1T	Non-covalent		472	507	29,600	94	85
Corn	DFHO	Non-covalent		505	545	29,000	25	92
Orange Broccoli	DFHO	Non-covalent		513	562	34,000	28	92
Red Broccoli	DFHO	Non-covalent		518	582	35,000	34	92



