Development of bioorthogonal fluorogenic reporters for biological imaging

Chenge Li

THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres
PSL Research University

Préparée à l' École Normale Supérieure

Development of bioorthogonal fluorogenic reporters for biological imaging

Ecole doctorale n°388

Chimie Physique et Chimie Analytique de Paris Centre

Spécialité Chimie-Physique

COMPOSITION DU JURY :

Mme. BLANCHARD-DESCE Mireille
Université de Bordeaux, Président du jury

M. GAUTIER Arnaud
Ecole Normale Supérieure, Directeur de thèse

M. JULLIEN Ludovic
Ecole Normale Supérieure, Invité

M. KLYMCHENKO Andrey
Université de Strasbourg, Rapporteur

M. WINSSINGER Nicolas
Université de Genève, Rapporteur

Soutenue par Chenge LI
le 04 10 2017

Dirigée par **Arnaud GAUTIER**

Département de CHIMIE ENS

PSL RESEARCH UNIVERSITY PARIS

Acknowledgement

I would like to extend my great gratitude to my supervisors, Dr. Arnaud Gautier, for his instructive advice, constant encouragement and guidance. Doing a PhD is always hard and full of unexpected setbacks, but his bright smile, positive attitude and solid support gave me the strength to keep going and overcome the difficulties. His consistent instruction contributes greatly to complete and modify this thesis.

I sincerely thank Dr. Mireille Blanchard-Desce, Dr. Andrey Klymchenko and Pr. Nicolas Winssinger, members of the jury who agreed to review my thesis and gave me the chance to present them my PhD work.

I want to give my gratitude to Pr. Ludovic Jullien for all those interesting discussion and useful suggestions for my PhD projects. It is such an honor to have him in the jury to witness this important moment.

I would like to thank all the members in our lovely group, Marie-Aude Plamont, Alison G. Tebo, Martha Zoumpoulaki, Frederico Milheiro Pimenta, Quentin Delacour for their kindness, their help and all the happy times we spent together.

I would like to acknowledge the financial grant from the Region Ile-de-France in the framework of DIM NanoK for supporting my PhD.

Special thanks to everyone who helped and contributed in the development of my projects. Thank Dr Thomas Le Saux for his help and instruction in my physical experimental part. Thank Annie Munier for the cell sorting. Thank Dr Isabelle Aujard for her kind help and warm care to me when I worked in the synthesis laboratory.

I am deeply indebted to all the other members in our laboratory, Jean-Bernard Baudin, Raja Chouket, Alexandra Colin, Marina Garcia-Jove Navarro, Louise Bonnemay, Geoffrey Brun, Fabrice Dalier, Remi Ducasse, Zoher Gueroui. Louise Hespel, Eliane Ipendey, Shunichi Kashida, Didier Law-Hine, Emmanuelle Marie, Karim Ounoughi, Jerome Querard, Jorge Royes Mir, Pierre Toxé, Christophe Tribet, Wei-An Wang, Xiaojiang Xie, Drazen Zanchi and Ruikang Zhang, for their warm welcome, their kindness and all the help they have given to me.

I want to thank all my friends for their strong support and our wonderful friendship. Thank Joelle and Abdelatif for accompanying me since the beginning of my life in France and for all the joy we shared together. Thank the family of my little angle Yvette for all the surprises and bright color added to my life. Special thanks to the group of Yuanqishaonv and the group of Meishaonv. Having you all in my life is an invaluable gift.

Thank my parents for loving me and supporting me. Because of you I always see the nice side of the world. Love you, dad and mom.

Last but definitely not least, I want to thank my husband who is also my lifelong soulmate. You make me brave because I know you will always be there behind my back. Being together with you I am improving myself all the time since in your eyes I see a better me. Love you, my dear.

Contents

Chapter I General Introduction

Bio-imaging, which makes the invisible visible, is the most direct method to reveal complicated and mysterious biological activities in living body. There are two general imaging modalities, clinical modalities, such as magnetic resonance imaging (MRI), X-ray computed tomography (CT), etc. and research modalities, like all kinds of optical microscopes, mass spectrometry imaging, etc.,[1] which have significant contribution to the development of basic science in all directions.

Live cell imaging, which unveils cellular events and enables to observe biomolecular processes in real time, has been applied to a wide range of biological researches. Development of novel reporters and research microscopy techniques has revolutionized the way of visualizing biological activities in cells and accelerate the pace of discovery. Proteins or other biomolecule as the crucial elements of many processes, are normally invisible and usually need to be visualized and tracked in space and time to help biologists to understand their functions. Currently, fluorescence is one of the most widely applied methods of investigation due to its high sensitivity and low experimental cost.

I-1 Fluorescence

Fluorescence is an emission of light from a molecule that has absorbed photons. This phenomenon has been first observed in 1565 by Nicolas Monardes, a Spanish physician and botanist. The term of fluorescence was introduced in 1852 by George Gabriel Stokes.[2-4]

I-1.1 The principle of fluorescence

The physical principle of fluorescence can be explained by the Perrin-Jablonski diagram (Figure 1.1). The singlet fundamental electronic state is denoted as S_0, singlet excited electronic states S_1, S_2…and the triplet states T_1, T_2…Each electronic state contains a range of vibrational and rotational energy levels. After the absorption of photon, electron from S_0 is promoted to the high energy electronic state S_1. The excited electron is not stable and prone to lose the extra energy to go back to the stable state S_0 through various de-excitation processes, such as internal conversion (IC), intersystem crossing (ISC), fluorescence and

phosphorescence. Fluorescence is the emission of photons accompanying the relaxation from S1 to S0.[2]

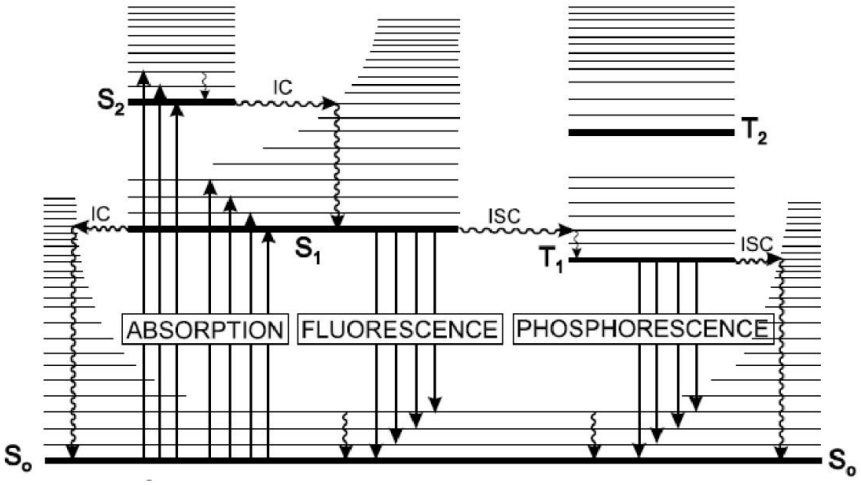


Figure 1.1 Perrin-Jablonski diagram. S_0: singlet fundamental electronic state, S_1, S_2: singlet excited electronic states, T_1, T_2: triplet electronic states, IC: internal conversion, ISC: intersystem crossing. Adapted from Valeur.[2]

I-1.2 Absorption, excitation and emission spectra

The absorption of the energy of a photon is an all or none phenomenon. There is no absorption if the photon contains insufficient energy for an electronic transition. On the contrary, absorption occurs when the photon contains enough energy and the excess energy is usually converted into vibrational and rotational energy. Consequently, the absorption of photon by fluorescent molecules can only occur with incident light of specific wavelengths known as absorption bands.

In general, fluorescent molecules undergo electronic transitions with radiation having wavelengths ranging from the ultraviolet to the visible regions of the electromagnetic spectrum. Absorption spectra of fluorescent molecules can be recorded by scanning absorbance $A(\lambda)$ (a factor that demonstrates the efficiency of absorption) at certain wavelength λ with ultraviolet-visible spectroscopy and their absorbance usually follows the Beer-Lambert Law.

$$\mathbf{A}(\boldsymbol{\lambda}) = \mathbf{log}\,\frac{\boldsymbol{I_0(\lambda)}}{\boldsymbol{I(\lambda)}} = \boldsymbol{\varepsilon(\lambda) l c}$$

Where $I_0(\lambda)$ and $I(\lambda)$ are the intensities of the light received and transited by the absorbing sample; $\varepsilon(\lambda)$ is the molar absorption coefficient (in L mol^{-1} cm^{-1}), l is the absorption path length (in cm) and c is the concentration (in mol L^{-1}) of absorbing sample.[2]

Excitation spectra of most fluorescent molecules are identical to their absorption spectra. Generally, the differences of vibrational levels are similar in the ground and excited states, so that the excitation and emission spectra are usually symmetrical. The wavelength of emitted light is longer than the incident light because of the energy loss of vibrational relaxation in the excited state. The difference between the maximum wavelength of excitation and emission spectra is termed as the Stokes shift (Figure 1.2).[2]

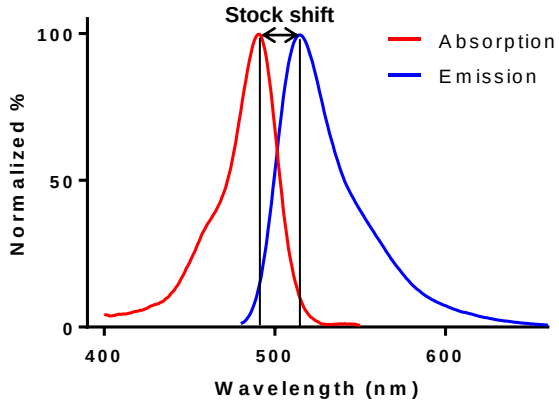


Figure 1.2 Absorption and emission spectra of fluorescein. Absorption and emission spectra are symmetrical. The gap between the maximum wavelength of absorption and emission spectra is the Stokes shift.

Excitation and emission spectra can be recorded by scanning fluorescence intensity in a selected spectrum region with a fluorometer at a fixed emission or excitation wavelength.[5] Through recording excitation and emission spectra of fluorescent reporters, one can learn their spectral properties and obtain indispensable information for the spectral region they can be applied in cell imaging. The color of light that we usually speak of can be quantitively described as range of wavelength (Figure 1.3).

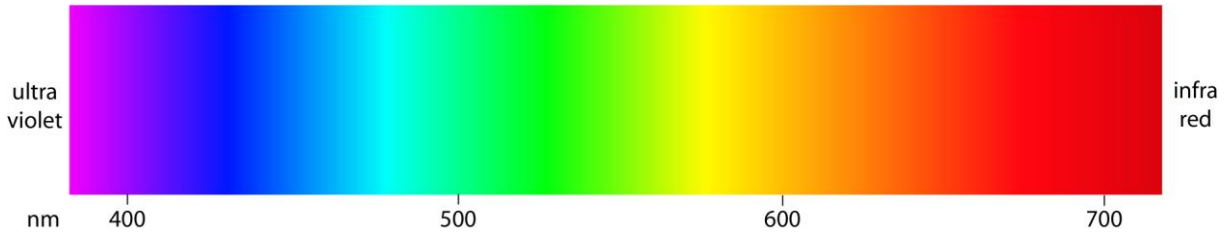


Figure 1.3 Color regions of visible light. violet: 400-455nm, blue: 455-492nm, green:492-577nm, yellow: 577-597nm, orange: 597-620nm, red: 620-700nm.

I-1.3 Fluorescence quantum yield, lifetime and brightness

The amount of fluorescence emitted by a fluorescent reporter depends on many factors.[2,6] Among them, the fluorescence quantum yield and lifetime are crucial for characterizing fluorescence performance of fluorescent reporter.

Fluorescence quantum yield (Φ) is the ratio of the number of photons emitted to the number of photons absorbed by fluorescent species, which reflects its efficiency of fluorescence emission and is given by:

$$\Phi = \frac{k_r}{k_{nr} + k_r}$$

where k_r is the rate constant for radiative deactivation from S_1 to S_0 through fluorescence emission, and k_{nr} is the rate constant for non-radiative deactivation from S_1 to S_0.

The lifetime τ of the excited state S_1 is the average time the molecule spends in the S_1 before returning to S_0 and is given by:

$$\tau = \frac{1}{k_{nr} + k_r}$$

Usually, one measured the relative quantum yield of fluorescent reporter by comparison with standards whose quantum yields are known, such as fluorescein (0.95±0.03 in 0.1M NaOH), Rhodamine 6G (0.94 in ethanol), etc.[6]

The relative quantum yield of fluorescent reporter is calculated by:

$$\Phi = \Phi_s \frac{1 - 10^{-A_s(\lambda)}}{1 - 10^{-A(\lambda)}} \frac{D}{D_s} (\frac{n}{n_s})^2$$

where the subscript *s* stands for the standard, A(λ) is the absorbance at the excitation wavelength λ, *D* is the integrated intensity of emission, and *n* is the refractive index of the solvent.

The brightness (B) is an important parameter for estimating the fluorescence performance of a fluorescent sample, which is determined by the molar absorption coefficient and the fluorescence quantum yield. Generally, brighter fluorescent reporters increase the signal-to-noise ratio. Moreover, the weaker intensity of light for exciting brighter fluorescent reporters decreases intervention and phototoxicity for hosts.[7]

Fluorescence techniques have many impressive advantages compare to other analytical and detective methods, explaining the wide applications of fluorescence in various domains.[2,8] Notably, the development of all kinds of fluorescence microscopies has been boosting spectacularly the ability of live-cell imaging.

I-1.4 Advantages of fluorescence-based investigative technologies

Fluorescence techniques have many impressive advantages compare to other analytical and detective methods, explaining the wide applications of fluorescence in various domains.[2,8]

High sensitivity. Fluorometry is up to 1000 times more sensitive than spectrophotometry.[2] With proper dyes and experimental instruments, the absolute sensitivity of fluorescence detection can reach the limit of single molecules.[6,8] Instead of obtaining average value of ensemble fluorophores in the observing region, behavior and performance of single fluorophore can be recorded by single-molecule fluorescence microscopy, which can provide additional information on the dynamics of fluorescently labeled biomolecule. [6,9,10]

High temporal and spatial resolution. The response of fluorescence techniques could be as fast as 10^{-8}-10^{-10}s, which is limited by the fluorescence lifetime and the experimental conditions.[8] The spatial resolution is restricted by the diffraction limit of light. Many novel fluorescence super-resolution microscopy technologies have been developed to overcome the light diffraction and to reach nanometric spatial resolution.[11-13]

I-2 Fluorescent reporters

Molecules that can generate fluorescence after absorbing photons are known as fluorophores. Specific fluorophores are needed to track movement of biomolecules for various biological researches. Fluorophores can be endogenous, such as some aromatic amino acids, fluorescent proteins, etc., which naturally exist in organisms and can be studied directly through fluorescence imaging techniques. However, the number of endogenous fluorophores is limited, which restricts application of fluorescence imaging. To conquer this, diversiform exogenous fluorescent probes have been developed and introduced into live cell for biological research.

I-2.1 Organic fluorescent probes

An important landmark in the field of fluorescent probes has been the development of organic fluorescent dyes. They can have some affinity for certain biomolecules or metabolites, and thus visualize locations of biomolecules and reveal biological processes in cells. Innumerous molecular probes with various properties has been synthesized and applied to a large numbers of biological studies.[14,15] Two key advantages of the organic fluorescent probes are the small size of these molecules, which will not disturb the natural function of labeled elements, and their marvelous versatility which enables to easily adjust their properties to the need of the experiment through chemical modulations. However, the fact that they show low labeling specificity has encouraged researchers to continue developing novel fluorescent probes for the selective fluorescent labeling of biomolecules, which tremendously boosted imaging techniques based on fluorescence and accelerated the progress of biological research. A substantive leap is the discovery and application of the green fluorescent protein (GFP).

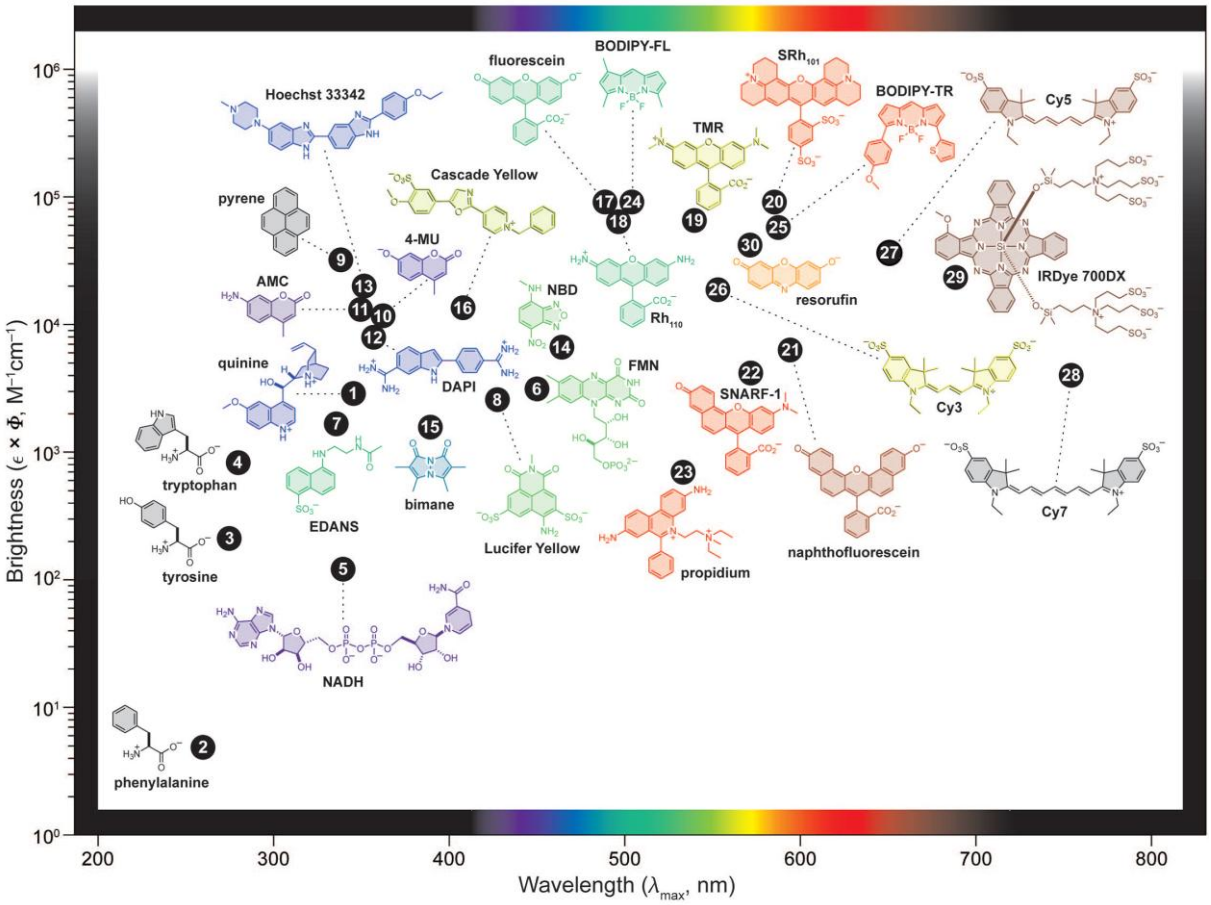


Figure 1.3 Common organic fluorescent dyes with various spectral properties. Adapted from Raines *et al.*[15]

I-2.2 Autofluorescent proteins (AFPs)

GFP is a companion protein of the chemiluminescent protein Aequorin, discovered by Shimomura *et al.* from the jellyfish *Aequorea Victoria*.[16] Thirty years later, the first cloning of the gene of GFP[17] and the initial experiments of expressing GFP in other organisms[18] opened the new era of bio-imaging. In most applications, AFP works as a fluorescent tag that is genetically fused to a protein of interest (POI), allowing visualizing specifically the fusion protein (Figure 1.4). The high labeling specificity provided by AFPs makes them indispensable in biological research. Since the success of GFP, many GFP-like AFPs with various colors have been developed to cover the spectral needs in multicolor imaging experiments (Figure 1.5).[19-25]

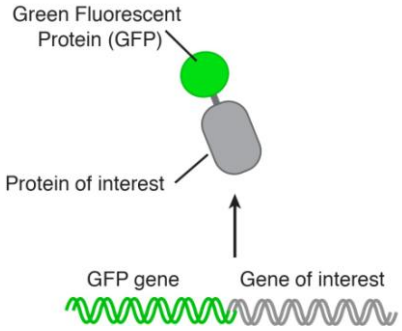


Figure 1.4 Genetic fusion of GFP to a POI enables specific fluorescent protein labeling.

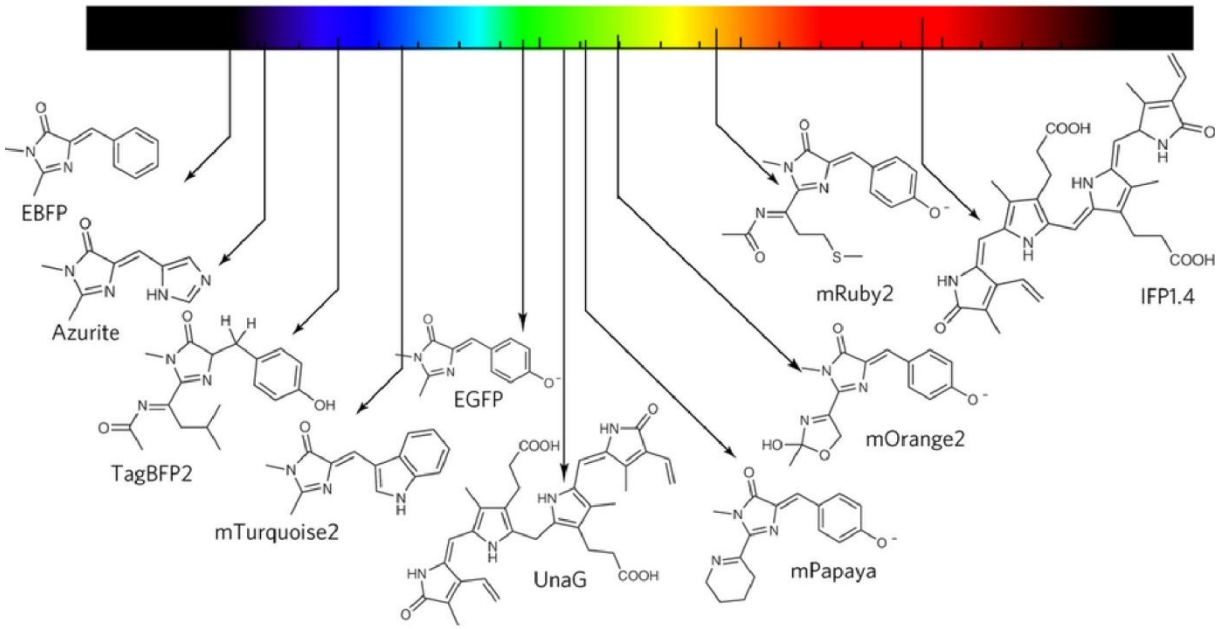


Figure 1.5 Examples of chromophores within AFPs covering the whole visible spectrum. Adapted from Palmer *et al.*[25]

As one of the most successful fluorescent reporters, AFPs have been widely applied in various biological applications (Figure 1.6).

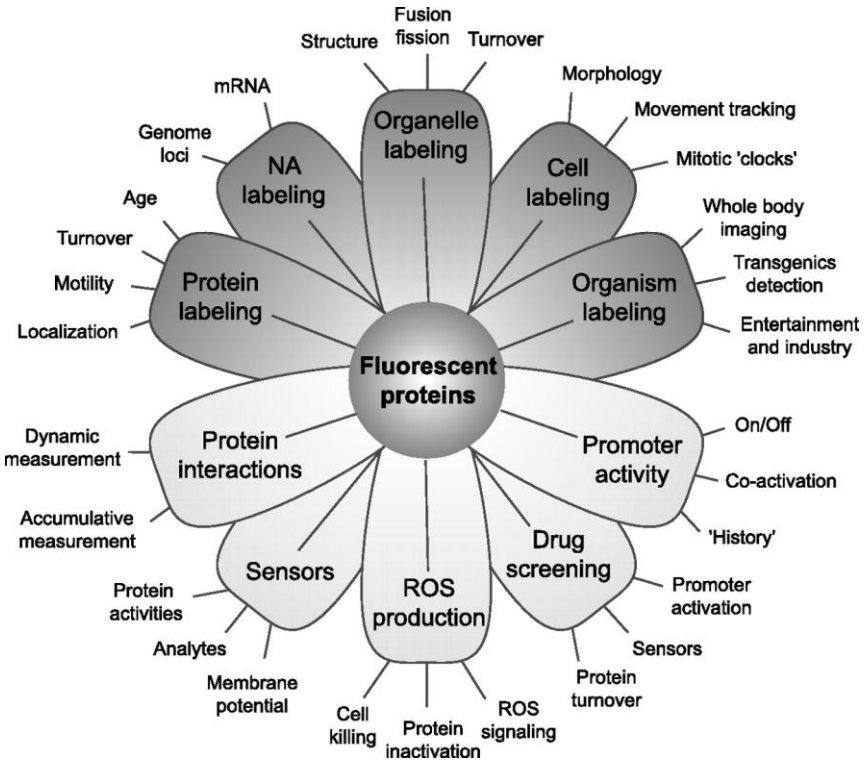


Figure 1.6 Biological applications of AFPs in scientific research. Adapted from Lukyanov *et al.*[7]

I-2.2.1 Applications of AFPs

Labeling

The first type of application is labeling. Genetic manipulations allow to fuse AFP to a protein of interest or a binding domain for a given biomolecule, which enables to obtain information about location, movement, function of biomolecules and can provide us with information on biological constituents (organelle, subcellular structure, cell, organism, tissue, etc.).

Apart from conventional labeling, AFPs have been used to develop novel techniques for biological detection and analysis.

Photobleaching techniques

Photobleaching is a common and undesirable phenomenon during the observation of fluorophore-labeled biomolecule through light irradiation. However, applying photobleaching as a disturbance factor can provide information about the mobility of biomolecules, which is important to understand their activities and functions in cell.[7,26,27]

The most obvious strategy is fluorescence recovery after photobleaching (FRAP). At first, a small region of interest (ROI) is irradiated by intense light to bleach AFPs in this region. Then the rate of fluorescence recovery in this region is measured, which reflects the rate of migration of unbleached AFPs from nearby regions, so that one can estimate mobility of AFP-labeled biomolecules.

There is another related strategy, fluorescence loss after photobleaching (FLIP). Instead of bleaching ROI, one can bleach a region nearby ROI and measure the rate of fluorescence loss of ROI, which avoid influence of possible photodamage in the observed region.

Actually, Photobleaching technique suffers from some inevitable disadvantage: phototoxicity caused by intense light irradiation; photobleaching occuring during the measurement of fluorescence recovery (loss); influence of AFP influx to ROI during bleaching. However, it has been well established because it is simple and straightforward to operate and generally it can provide valuable information on the mobility of biomolecules.

Förster resonance energy transfer (FRET)

FRET is an important technique to study protein-protein interactions. Two AFPs are chosen as a pair of donor-acceptor and are genetically encoded with two proteins of interest. The information about interaction of two proteins of interest is obtained by recording FRET between AFPs. To be specific, if two proteins interact, corresponding fusion AFPs get closer and thus the energy can be transferred from the excited donor fluorophore to the acceptor fluorophore through nonradiative dipole-dipole coupling (Figure 1.7). [7,28]

A suitable AFPs pair with high FRET efficiency should have several key elements: overlap of donor emission spectrum and acceptor absorption spectrum; gap between the donor absorption spectrum and acceptor emission spectrum; significant and appropriate brightness and high photostability. Based on this, various combinations of AFPs pairs have been developed for enriching FRET techniques. [7]

Bimolecular fluorescence complementation (BiFC)

Another common technique to study protein-protein interaction is BiFC or split fluorescent protein (Split-FP). Split-FP is designed as two separate complementary fragments of an AFP fused to two proteins of interest. The two fragments do not associate unless the two protein of interest interact (Figure 1.7). The reconstitution of AFP generates fluorescence so one can

detect protein-protein interaction with the appearance of fluorescence. Different split systems have been developed from various AFPs, enlarging significantly the applications of this technique.[7,29]

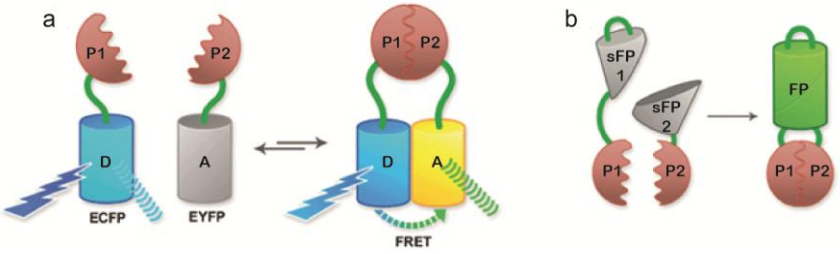


Figure 1.7 AFP-based techniques for studying protein-protein interaction. a) FRET: ECFP (donor) transfers its energy to EYFP (acceptor) when two fusion proteins interact. **b) BiFC:** Interaction between two fusion proteins enables the reconstitution of AFP. P: protein, D: donor, A: acceptor, ECFP: enhanced cyan fluorescent protein, EYFP: enhanced yellow fluorescent protein. Adapted from Frommer *et al*.[30]

I-2.2.2 Limitations of AFPs

Compared to organic fluorescent probes, AFPs have some inevitable disadvantages: (1) The relative large size of AFP may hinder the normal function of the fusion protein. (2) The spectral properties of AFPs are still limited such as the brightness and the photostability. (3) AFPs fluoresce only after the formation of the fluorophore. The multi-step maturation (Figure 1.8) usually requires the presence of oxygen and takes tens of minutes or hours,[31,32] which prevents their applications in anoxic conditions or for studying fast biological processes.

Although AFPs with improved brightness,[33] better photostability[34] or faster maturation[35] have been developed, we still expect further improvement that could combine the high specificity of AFP and versatility of small fluorescent molecules in the same fluorescent probe system. The most impressive developments are the site-specific labeling techniques, in which peptidic tags are used as genetic anchors reacting specifically with synthetic probes.

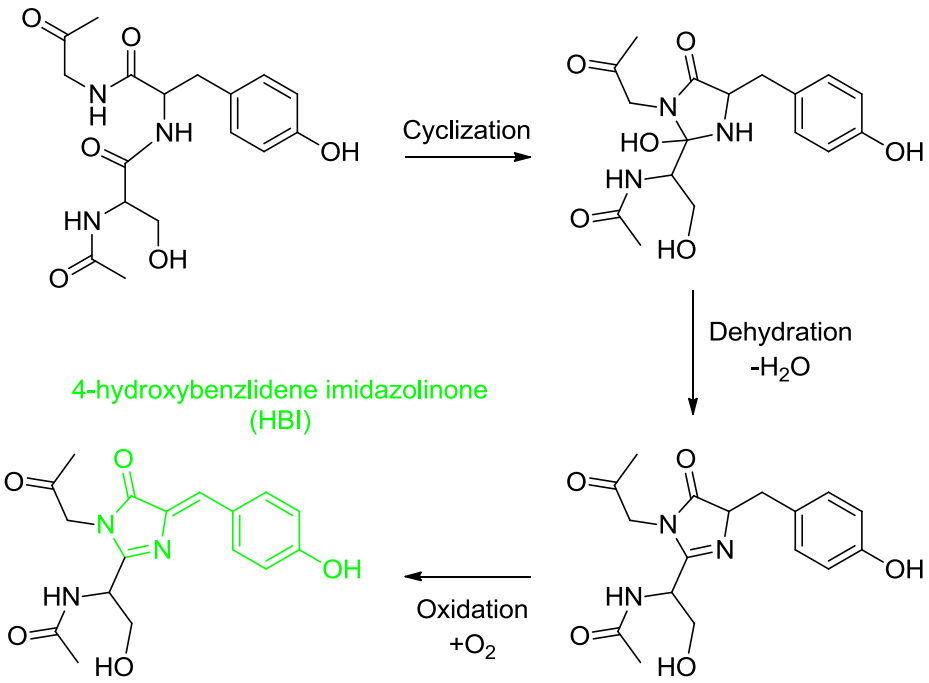


Figure 1.8 The formation of the chromophore HBI in the binding barrier of GFP. This multi-step maturation process requires the presence of O_2 and takes tens of minutes or hours.

I-2.3 Site-specific labeling techniques

I-2.3.1 Biarsenical- tetracysteine tagging system

The first tag developed for site-specific protein labeling was the tetracysteine motif, a short peptide usually composed of six amino acids CCPGCC that can bind with biarsenical dyes.[36] The principle of the binding is the well-known interaction of arsenic with thiol groups. The first developed biarsenical dye, FlAsH, is a fluorescein derivative with two As(III) substituents, which is weakly fluorescent as an adduct with two 1,2-ethanedithiol (EDT) but becomes green fluorescent after binding with the tetracysteine tag (Figure 1.9).[37,38] Analogs have been developed for expending the spectral properties.[36,39] ReAsH (Figure 1.9) as the most useful one is a derivative of the red fluorophore resorufin and could be applied in pulse-chase imaging with FlAsH.[39] Although the small size of the peptide tag and the various biarsenical dyes makes it one of the best established site-specific labeling systems, it still needs to be improved in many aspects like the signal-to-background ratio, the binding

specificity, the need of antidotes to reduce the toxicity and the requirement of a reducing environment.[36]

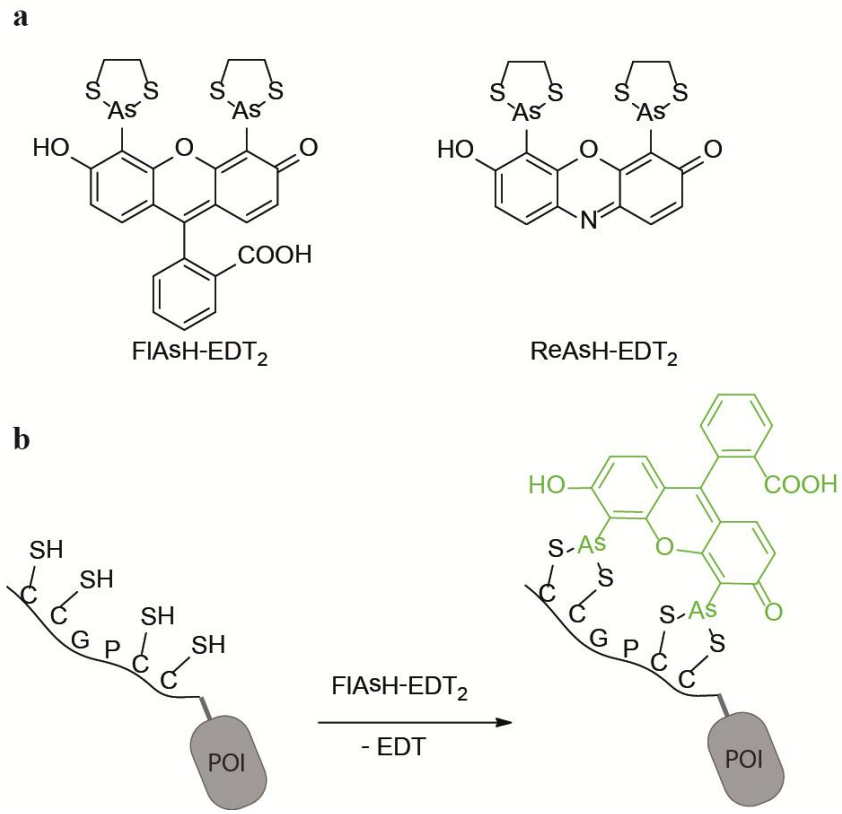


Figure 1.9 a) Chemical stucture of biarsenical substrates FlAsH-EDT$_2$ and ReAsH-EDT$_2$. b) Schematic illustration of the biarsenical- tetracysteine tagging system. FlAsH as an adduct with two 1,2-ethanedithiol (EDT) is weakly fluorescent but becomes fluorescent after interacting with thiol groups of the tetracysteine tag POI: protein of interest.

Since the development of biarsenical- tetracysteine tagging system, there have been many efforts to create new site-specific labeling systems with better properties. SNAP-tag, CLIP-tag and Halo-tag are three outstanding representatives which are all commercially available now.

I-2.3.2 Self-labeling tags: SNAP-tag, CLIP-tag, Halo-tag

SNAP-tag is a mutant of the human DNA repair protein O^6-alkylguanine-DNA alkyltransferase (AGT) developed by the group of Johnsson.[40 - 43] It can transfer the functionalized benzyl group of the O^6-benzylguanine (BG) derivative to its active site cysteine and thus achieves covalent and irreversible labeling of fusion protein (Figure 1.10).

The versatility of the developed BG derivatives enables to easily adapt SNAP-tag functionalities to different conditions for visualizing proteins *in vitro* and *in vivo*.[44-46]

To meet the need for multiplexed imaging, the group of Johnsson developed CLIP-tag[47], which shows substrate specificity orthogonal to SNAP-tag. CLIP-tag is an engineered mutant of AGT that works as SNAP-tag but reacts specifically with O^2-benzylcytosine (BC) derivatives (Figure 1.10). These two tags have been used to label two different proteins in living cells for simultaneous pulse-chase experiments[47] and also used to cross-link interacting proteins to detect and characterize protein-protein interactions[48].

The Halo-tag is engineered from a bacterial halogen dehydrogenase. The carbon halogen bond of the functionalized chloroalkane derivative could be cleft by the Halo-tag and thus the fluorescent ligand covalently binds with the Halo-tag fusion protein (Figure 1.10).[49]

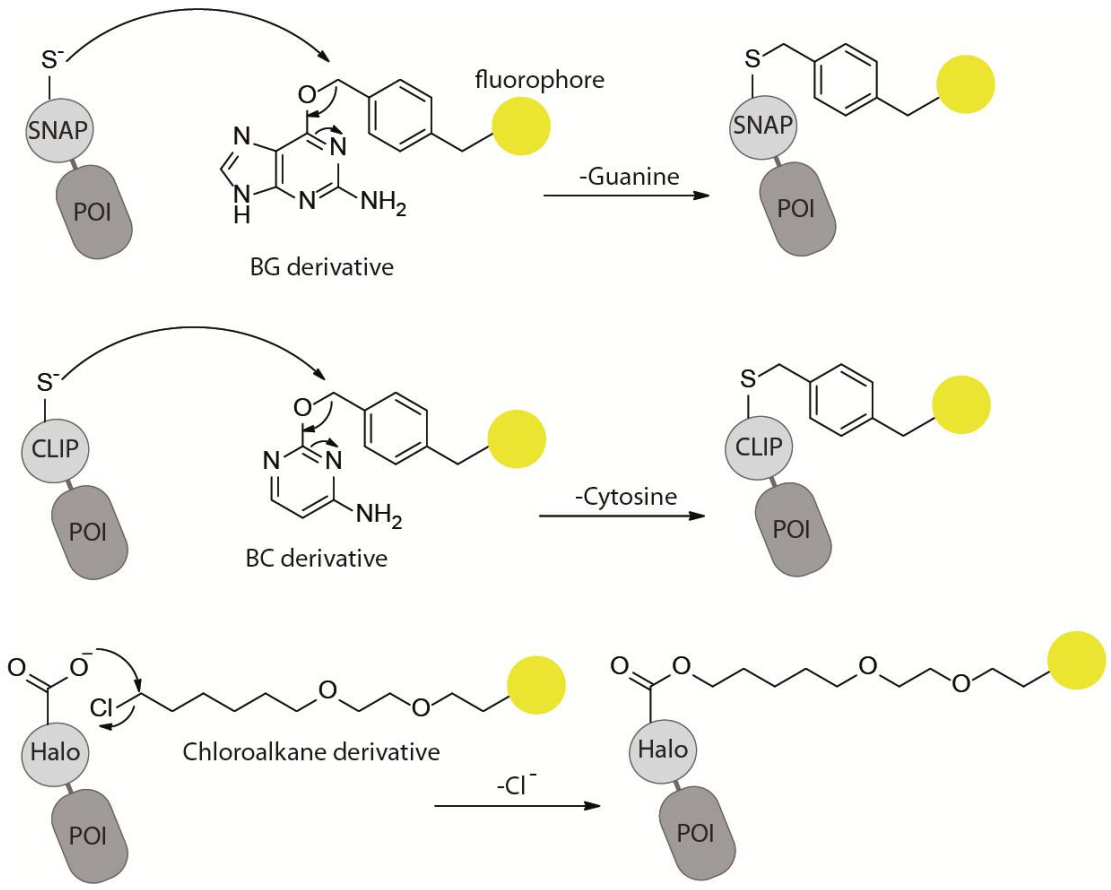


Figure 1.10 General mechanism of site-specific labeling systems. SNAP-tag, CLIP-tag and Halo-tag. POI: Protein of interest.

The labeling specificity and the easy chemical modification of the substrates make these self-labeling systems very attractive. However the imaging of all these systems can show fluorescence background because of the excess of substrates and nonspecific binding. Therefore extensive washing steps are requested before imaging, which may limit the temporal resolution of the experiments.

I-2.3.3 Fluorogenic site-specific labeling systems

To improve the imaging contrast, there have been some developments of fluorogenic substrates for these self-labeling systems. Fluorogenic chromophores (so-called fluorogens) are non-fluorescent by their own but become strongly fluorescent only upon specific binding to a given target. The application of fluorogens minimizes unspecific fluorescence background in cells as free fluorogens are almost invisible even in excess[50] and therefore saves the need for washing steps. The simplified labeling protocol and the high imaging contrast have recently placed fluorogen-based reporters to the central spot of fluorescence imaging.

Komatsu *et al.* presented some fluorescence activation-coupled protein labeling (FAPL) probes for the SNAP-tag protein based on Förster resonance energy transfer (FRET) principle.[51] They designed SNAP-tag labeling probes with different fluorophores at the benzyl moiety and a quencher at the guanine moiety via a linker attached at the C-8 position of the guanine. Before labeling, these substrates show almost no fluorescence because of the high FRET efficiency between the fluorophore and the quencher. However, after reaction with SNAP-tag, fluorescence rapidly appears as the fluorophore is separated from the quencher leaving with the guanine moiety (Figure 1.11). A fluorescence activation ratio of over 300 is observed after labeling, achieving high contrast imaging and avoiding the need of washout process. Komatsu *et al* showed that the fluorogen-based SNAP-tag system could be used to study real-time protein activities with high temporal resolution.[51]

Sun *et al*. further developed intramolecularly quenched fluorogenic probes for $SNAP_f$, a faster labeling variant of SNAP-tag. $SNAP_f$ displayed up to tenfold increase in activity towards benzylguanine substrates compared to SNAP26m, a commercially available version of SNAP-tag. By exploring different combinations of fluorophore/quencher pairs, they obtained optimal

intramolecularly quenched substrates with fluorescence emission across the visible spectrum.[52]

Except for these intramolecularly quenched fluorogenic probes, the group of Johnsson developed near-infrared fluorogenic silicon-rhodamine (SiR) probes for SNAP-tag, CLIP-tag and Halo-tag (Figure 1.11).[53] These SiR probes are environment-sensitive SiR-carboxyl derivatives. In relative polar environment, the carboxylate group of these probes can attack the electrophilic center of the heterocycle forming the non-fluorescent spirolactone. However the coupling to protein tags favours the formation of the open fluorescent zwitterions. The opening of the cyclized chromophore is the important mechanism for the fluorogenic response because of a large change of the absorption properties between the open and closed states (Figure 1.11).[54] These SiR probes enable imaging fusion proteins with high sensitivity at 640-650 nm, a wavelength range where cellular autofluorescence is weak and phototoxicity is minimal.[53]

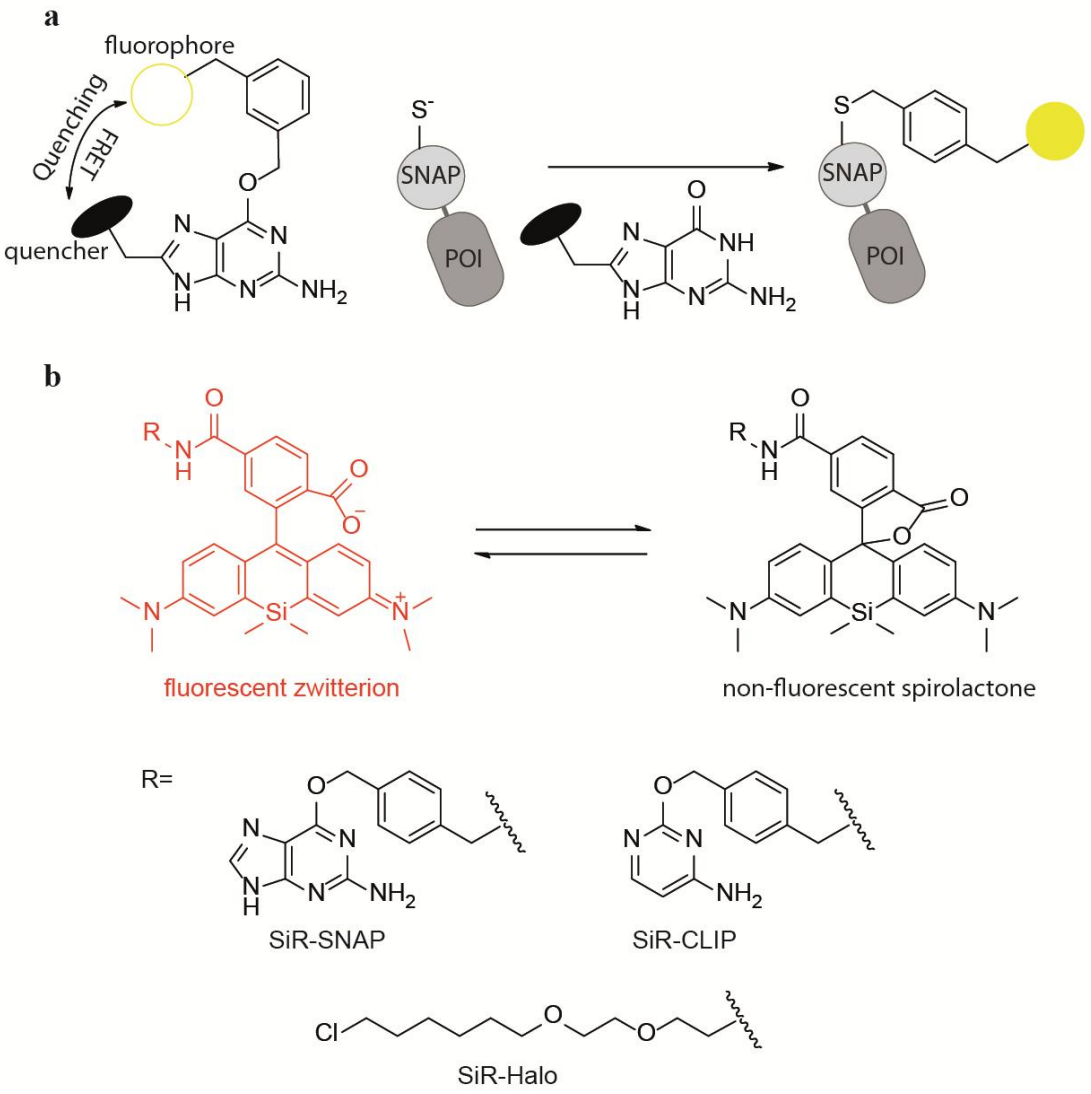


Figure 1.11 Fluorogenic site-specific labeling systems. a) Principle of intramolecularly quenched fluorogenic probes for SNAP-tag. Before labeling, fluorescence is quenched because of the high FRET efficiency between the fluorophore and the quencher. After reaction with SNAP-tag protein, the quencher leaves with the guanine moiety and fluorescence appears rapidly. **b) Structures and mechanism of near-infrared fluorogenic SiR probes**. SiR probes are sensitive to the environment. In relative polar environment, the carboxylate group of SiR probes attacks the electrophilic center of the heterocycle forming the non-fluorescent spirolactone. However coupling to protein tags favours the formation of the open fluorescent zwitterions**.** POI: protein of interest.

These fluorogenic site-specific labeling systems eliminate the high fluorescence background, the common disadvantage for chemical labeling and save the time for washing steps. However, the relatively large sizes of SNAP-tag (20kDa), CLIP-tag (20kDa) and Halo-tag

(34kDa) and the long time for complete labeling still may hinder their applications in some cases. Actually, there have been many efforts to combine different fluorogens with different small genetically encoded tags to create novel fluorogen based specific labeling systems with high temporal resolution.

I-2.4 Fluorogen-activating proteins (FAPs)

Fluorogen-activating proteins (FAPs) engineered from libraries of human single-chain antibodies (scFvs) bind non-covalently fluorogenic derivatives of malachite green (MG) and thiazole orange (TO) with nanomolar affinity and generate huge enhancement of their red and green fluorescence, respectively.[55] MG and TO are well-known molecular rotors that are poorly fluorescent in solution because of internal rotation; however in constrained environments, the internal rotations are slowed down, which strongly enhance their fluorescence intensities.[54] FAPs were shown to be suitable to label proteins without the need of wash-out process. Cell impermeant MG-2p and TO1-2p enable labeling of only the FAP-tagged proteins at the plasma membrane, while the cell permeant fluorogen MG-ester allows lighting up proteins in the secretory pathway within several minutes after addition (Figure 1.12).[55]

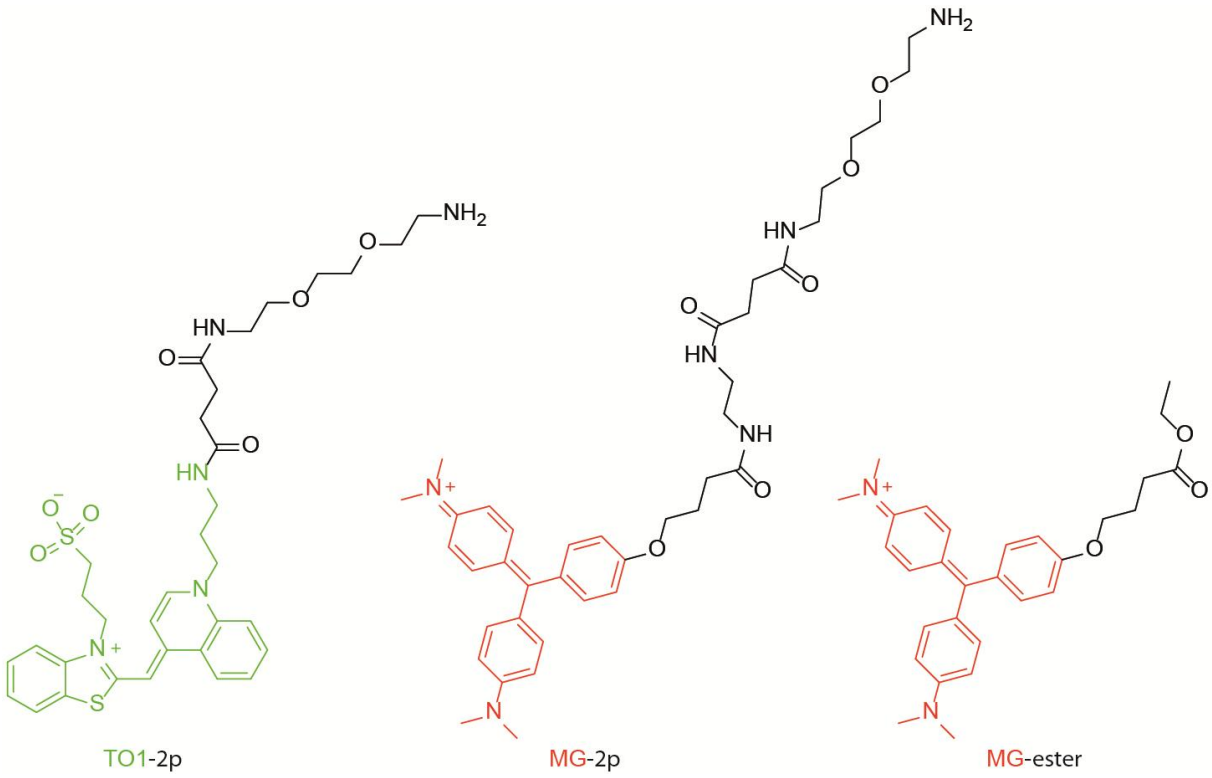


Figure 1.12 Structures of fluorogens for FAPs.

Various fluorogens have been developed to bind the single scFv protein K7 which shows huge versatility and accept different fluorogenic dyes.[56,57] Dimethylindole Red (DIR) and its analogues (unsymmetrical cyanines) were developed to bind the FAP K7 and expand the spectral properties of FAPs by giving complexes with emission wavelengths ranging from the blue (450 nm) to near-IR (750 nm) (Figure 1.13).[56] These dyes were further modified (the α-CN-DIR and the fluorinated TO (Figure 1.13)) to improve the photostability of the FAP system.[57] The following development of blue FAP systems based on blue fluorogenic cyanine dyes (OTB-SO_3) (Figure 1.13)[58] and the far-red FAP systems based on the ester of malachite green dyes (MGe) (Figure 1.13)[59] makes emission colors of FAPs cover almost the whole visible spectrum, which is versatile for multiplexed imaging.

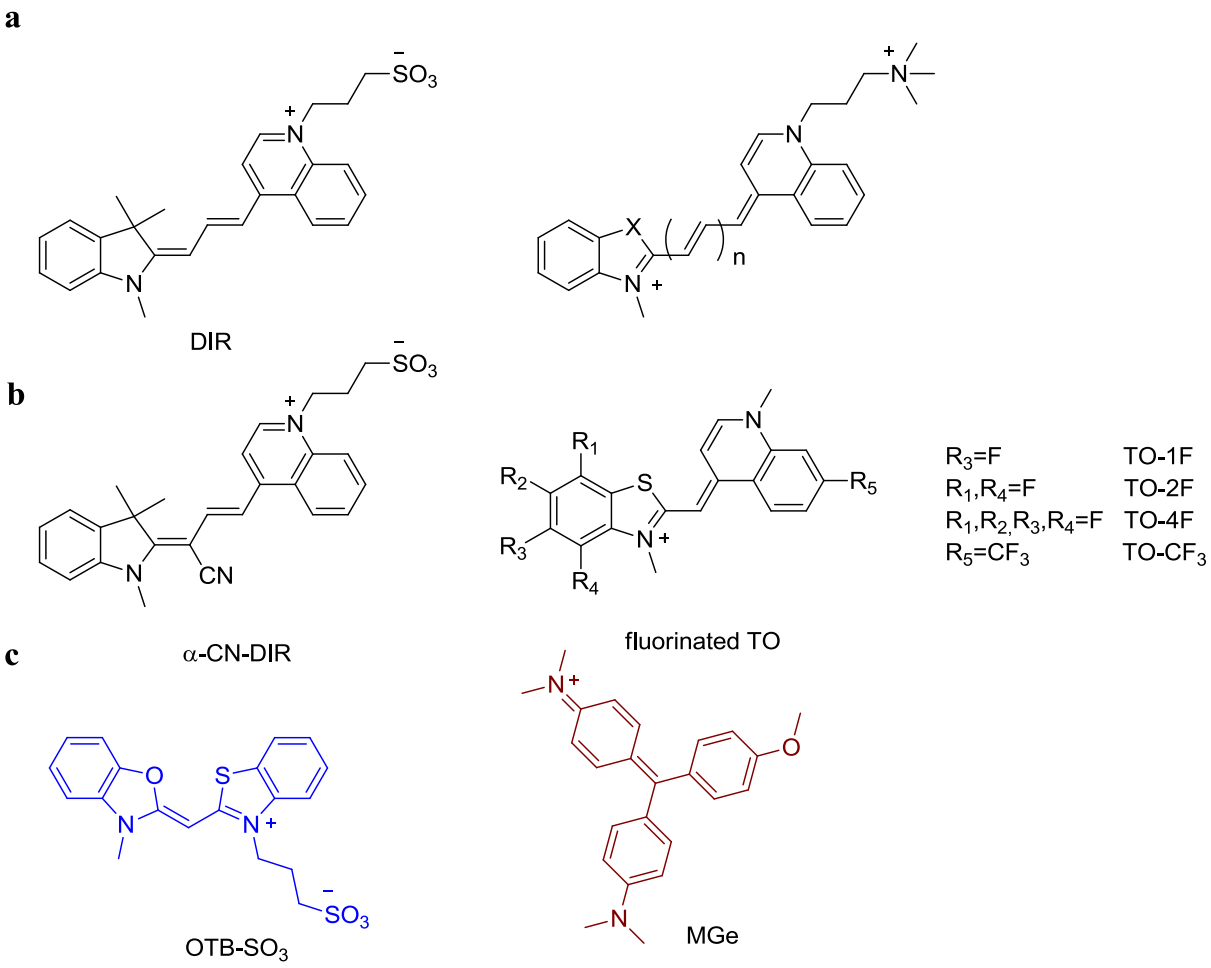


Figure 1.13 Development of fluorogens for FAPs. a) The DIR based fluorogens which could bind K7 and generate FAP systems with various emission wavelengths. b) Structures of α-CN-DIR and fluorinated TO which improve the photostability of FAP system. c) Structures of fluorogens for blue and far-red FAP systems.

However, the fact that single-chain antibodies contain internal disulfide bonds that form only in non-reducing environment limits their applications to non-reducing environments.[55] To overcome this limitation, disulfide-free variants FAPs has been developed, which enables the labeling of FAP systems in various environment in live cells.[59]

I-2.5 PYP-tag

The group of Kikuchi has developed the photoactive yellow protein (PYP)-based fluorogenic protein labeling systems[60-65] by using the small (14kDa) PYP discovered from the purple bacteria *Halorhodospira halophila*[66] as protein tag. The fluorogens they developed, coumarin thioester and cinnamic acid thioester derivatives, covalently bind to the residue Cys69 of the PYP-tag through transthioesterification and generate fluorescence (Figure 1.14).[60-62]

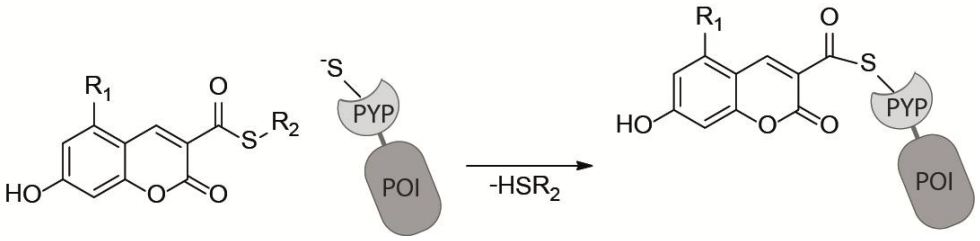


Figure 1.14 The principle of PYP-tag labeling system. Coumarin thioester derivatives react with PYP through transthioesterification. POI: Protein of Interest.

The first fluorogen FCTP for the PYP-tag used a quenching mechanism based on intramolecular association between fluorescein and the coumarin moiety (Figure 1.15). However, the complete labeling required more than 24 h, which limited the temporal resolution of the method.[60]

To accelerate the labeling process, the group of Kikuchi evaluated a new fluorogenic probe FCANB consisting of a cinnamic acid thioester as PYP-tag binding ligand and nitrobenzene as fluorescence quencher (Figure 1.15). The cinnamic acid thioester reduced the intramolecular interaction between the ligand and fluorophore moieties. Furthermore the introduction of the quencher made the fluorophore preferably interact with the nitrobenzene instead of the ligand, which reduced the steric hindrance around the ligand and thus enhanced the labeling kinetic. The cell-impermeable FCANB could be used to image cell-membrane proteins with high contrast.[61]

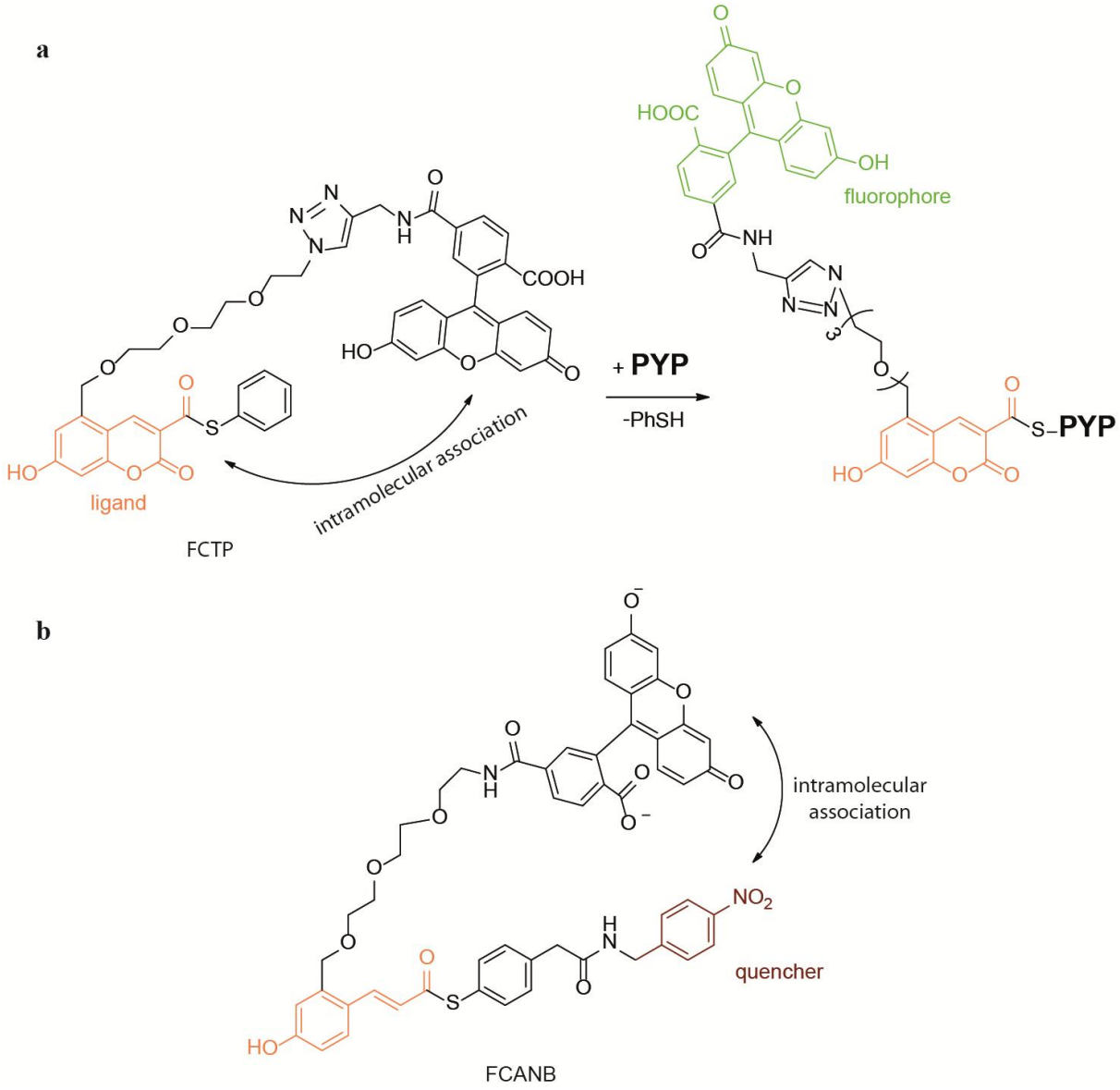


Figure 1.15 The intramolecular interaction based fluorogen for PYP-tag. a) Principle of fluorogenic labeling of PYP-tag with FCTP. b) Structure of FCANB, a new fluorogen for PYP-tag to accelerate the labeling. The cinnamic acid thioester and the introduction of the quencher reduced the intramolecular interaction between ligand and fluorophore moieties and thus enhanced the labeling kinetic.

Kikuchi and co-workers also developed new fluorogens to meet the need of cell permeant probes to image intracellular proteins. TMBDMA and CMBDMA are 7-dimethylaminocoumarin thioester derivatives that are environmentally sensitive (Figure 1.16). In polar environments, they show no fluorescence, while in the relatively low polar protein cavity, they display enhanced fluorescence. They could be used to completely label PYP-

tagged intracellular protein and membrane protein in a short incubation time, especially TMBDMA (<6 min).[62]

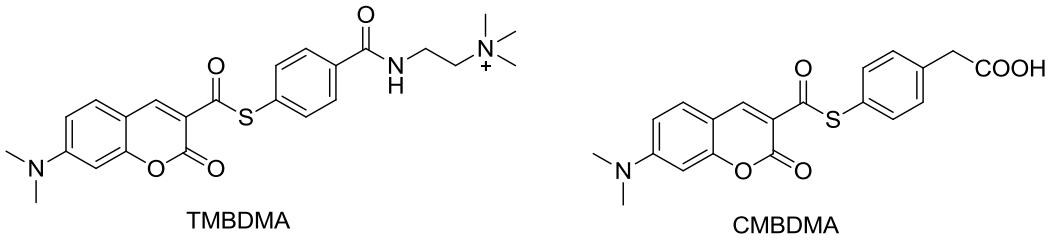


Figure 1.16 Structures of environment-sensitive cell permeant fluorogens for the labeling of PYP-tagged intracellular proteins.

To further improve the properties of PYP-tag labeling system, they created mutants of the PYP-tag and chemically modified fluorogens by considering the electrostatic interactions of the protein tag with the fluorogen and the pK_A of the leaving group.[63,64] Fluorogens that could generate different colors have also been developed for multicolor imaging.[65] The small size of the PYP-tag and the quick wash-free imaging ability make PYP-tag very attractive for fluorescence imaging in living cells.

I-2.6 Cellular retinoic acid binding protein II (CRABPII)

The cellular retinoic acid binding protein II (CRABPII) is a binding and transport protein that binds cellular retinoic acids and thus transport them inside the cell cytoplasm.[67-69] Its small size (15.6 kDa), relatively large binding pocket that could binds with various synthetic retinoids (Figure 1.17) and tolerance of mutations makes it a promising protein scaffold for the engineering of fluorogenic fluorescent protein tag.[70] The group of Borhan has developed a novel fluorogenic protein labeling system by reengineering CRABPII. A non-fluorescent precursor merocyanine aldehyde reacts with the active site Lys residue of CRABPII-tag to generate the fluorescent protonated Schiff base which also displays a huge red-shift of absorption and a significant increase of extinction coefficient compare to the precursor merocyanine aldehyde (Figure 1.18). The chromophore is planar and restricted in the protein cavity, which enhance its fluorescence.[70].

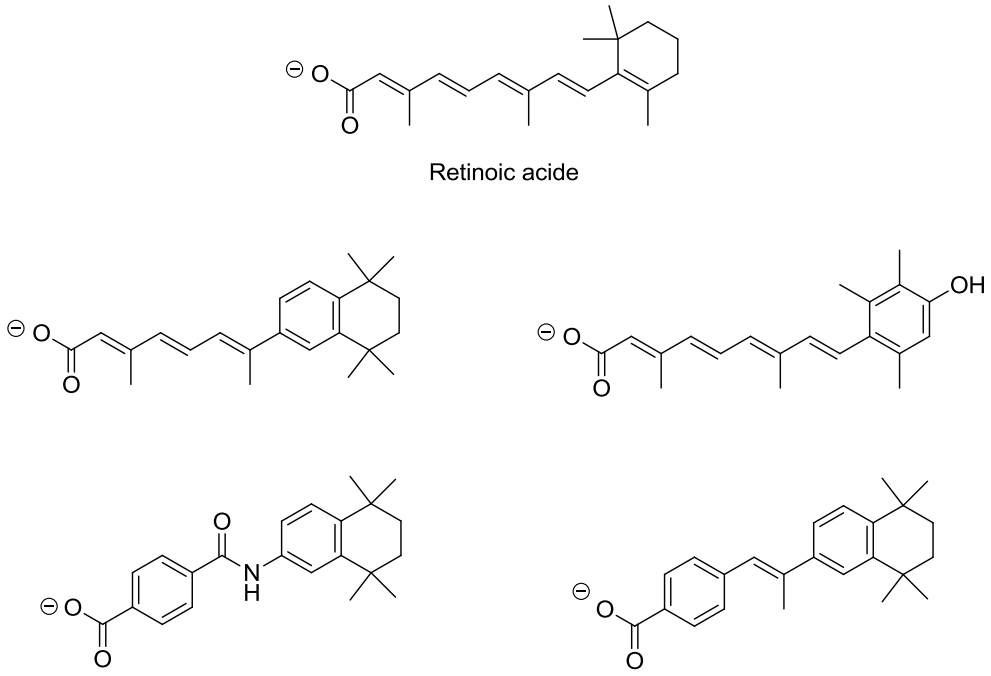


Figure 1.17 Structures of retinoic acids and synthetic retinoids that shows binding affinity with CRABPII.

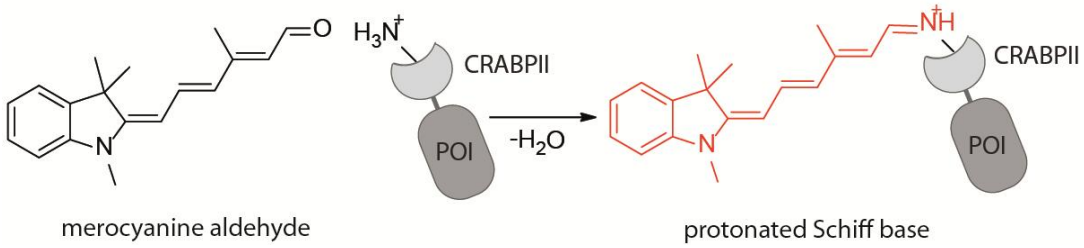


Figure 1.18 Fluorogenic labeling of CRABPII-tag with merocyanine aldehyde.

I-2.7 Spinach

Except for visualizing proteins, fluorogenic labeling has also been applied to other genetically encoded biomolecules, such as RNA. There have been many developments of fluorogen-activating RNA aptamer.[71-76] However the first fluorogen-based RNA aptamer that could efficiently image RNA in living cells is the Spinach system (Figure 1.19) which has been developed by the group of Jaffrey.[77] Spinach relies on the binding and activation of analogs of HBI, the chromophore of GFP. As presented before, GFP becomes fluorescent only after the formation of HBI which is a well-known fluorogenic chromophore. HBI is only fluorescent

when encased in GFP. Denaturation of GFP leads to a loss of fluorescence.[77] Inspired by the GFP, the group of Jaffrey designed and synthesized several fluorogenic analogs of HBI and selected RNA aptamers that could bind specifically with the fluorogen from a library of random RNA molecules.[77] To improve the properties of Spinach, they developed Spinach2, which shows enhanced folding and thermal stability[78] and they selected Broccoli through a fluorescence activated cell sorting (FACS) based directed evolution approach (Figure 1.20)[79].

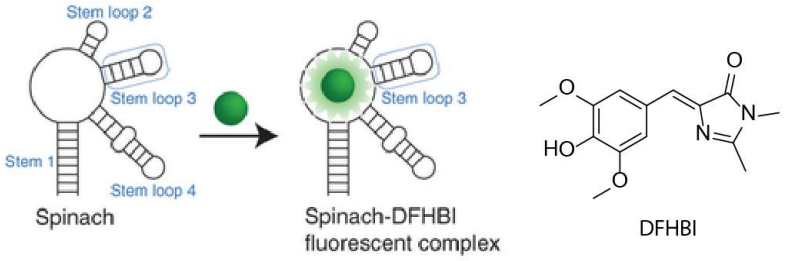


Figure 1.19 The Spinach labeling system. The non-fluorescent DFHBI reacts with the RNA aptamer Spinach and forms a fluorescent complex. Adapted from Jaffrey *et al.*[80]

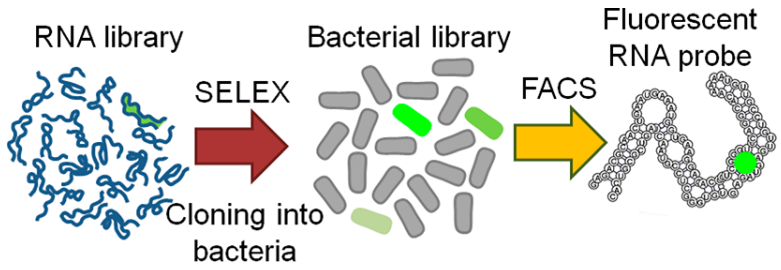


Figure 1.20 General procedure for isolating fluorescent "light up" aptamers. Millions of aptamers were expressed in Escherichia coli. and selected through FACS (Fluorescence Activated Cell Sorting) in the presence of potential fluorophores. SELEX (Systematic Evolution of Ligands by Exponential Enrichment) Adapted from Jaffrey *et al.*[79]

I-2.8 Fluorescence-Activating and absorption-Shifting Tag (FAST)

Our group recently developed a novel fluorogenic protein labeling system called Fluorescence-Activating and absorption-Shifting Tag (FAST).[81] The fluorogenic dye 4-hydroxy-3-methylbenzylidene rhodanine HMBR binds with FAST non-covalently and

generates a bright green-yellow fluorescent complex (Figure 1.21). In addition to fluorescence activation as generally observed with fluorogen-based labeling systems, HMBR also displays an absorption red shift after binding with FAST which enables high contrast imaging by distinguishing free and bound fluorogen via the choice of the excitation wavelength. This unique fluorogen activation mechanism implying two spectroscopic changes, fluorescence quantum yield increase and absorption red shift, ensures thus selective labeling and high imaging contrast.

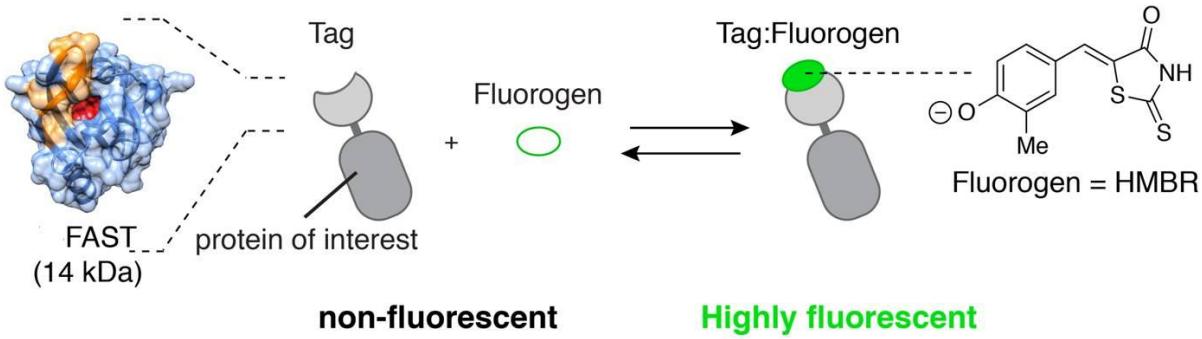


Figure 1.21 Principle of FAST.

The development of HMBR was inspired by GFP. HMBR is an analog of HBI and is based on a push-pull structure. The electron-donating phenol group interacts with electron-withdrawing rhodanine group through a conjugated system. HMBR (pK_a=8.4) stays almost all in protonated form at pH 7.4 showing no fluorescence. However, HMBR binds with FAST in its deprotonated form. The binding induced deprotonation of the phenol group generates phenolate with higher electron-donating effect, which explains the absorption red shift. Meanwhile, constrained environment in the protein cavity restrains the conformational freedom of HMBR and thus activates its fluorescence.

FAST is an engineered variant of PYP. The chromophore of PYP, parahydroxycinnamoyl (HC), binds to its residue Cys69 and shares structure similarity with HMBR. This is an important reason why PYP was chosen as protein scaffold to develop FAST. The engineering of FAST used a fluorescence activating cell sorting (FACS)-based directed evolution strategy (Figure 1.22). At first, libraries of PYP mutants were built by saturation mutagenesis of positions in the chromophore binding sites. Yeast-displayed libraries expressing PYP mutants were incubated with the fluorogen and sorted by high throughput FACS.

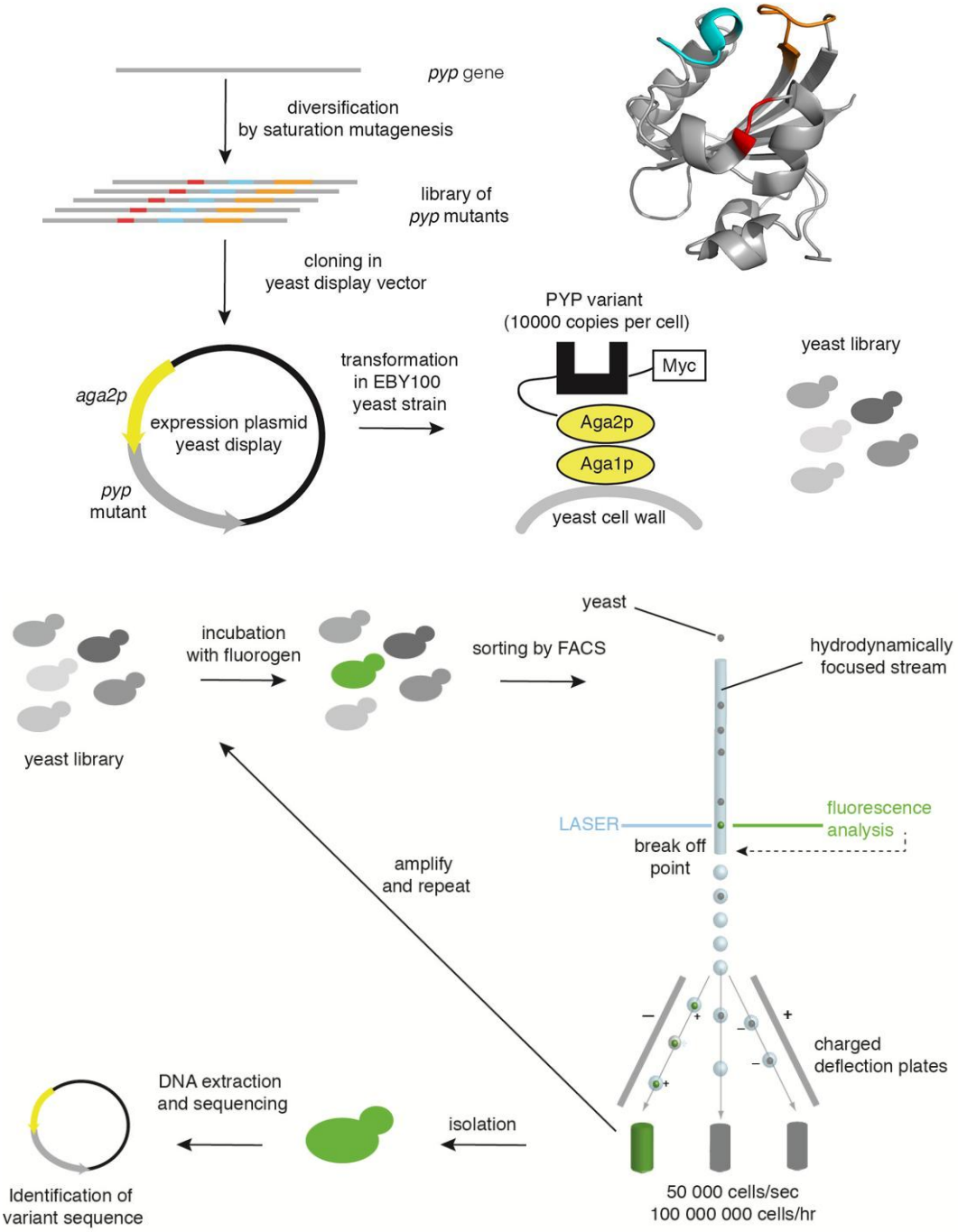


Figure 1.22 Engineering of FAST. Libraries of PYP mutants are built by saturation mutagenesis of loops (the loop 94-101 is given in orange) within the chromophore binding site. Through surface display technology, libraries of yeast cells expressing protein variants on their surface are obtained. Yeast libraries incubated with the fluorogen are sorted by high throughput FACS. Iterative FACS cycles enable to enrich the population of fluorogen-activating variants.

Compared to other fluorogen-based labeling systems, FAST has many attractive advantages: The protein tag is monomeric and small (14 kDa), which minimize its hindrance for normal

activities of fusion protein; FAST works as soon as it is folded and the labeling of FAST fusion protein occurs almost immediately after addition of fluorogen, which enables its application for real-time imaging; The binding between FAST and its fluorogen is reversible and shows rapid exchange dynamics, which makes labeling of FAST highly tunable. It has been proved that FAST can be rapidly switched on and off by addition and removal of HMBR and this process is repeatable. The controllable labeling opens new opportunities for multiplexed imaging. These advantages make FAST a very promising protein tag.

I-3 Objectives

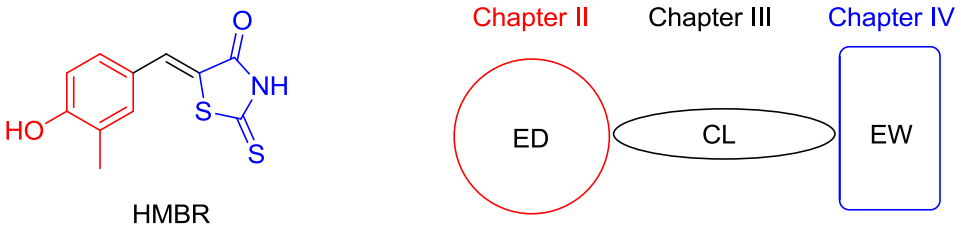


Figure1.23 Chemical structure of HMBR (left). HMBR is based on a push-pull system. Electron-donating phenol interacts with electron-withdrawing rhodanine through a conjugated linker. **General structure of HMBR and the arrangement of this thesis.** HMBR analogues are fully engineered by modifying its electron-donating part (See Chapter II), conjugated linker (See Chapter III) or electron-withdrawing part (See Chapter IV) for developing novel FAST approaches. ED: electron-donating part, CL: conjugated linker, EW: electron-withdrawing part.

My PhD aimed at extending the possibilities offered by FAST through the engineering of new fluorogens and additional protein engineering. The first step was the development of novel fluorogens. HMBR is a push-pull system. The electron-donating phenol group interacts with the electron-withdrawing rhodanine group through a conjugated system. Molecules with such push-pull structure are known as environment-sensitive dyes[54]. Spectral properties of these molecules are determined by the electron-donating or electron-withdrawing abilities of their electron donors or acceptors and the length of the electronic conjugation. On the basis of this push-pull structure, we engineered HMBR analogues by modifying either their electron-donating or electron-withdrawing groups or the length of the conjugation paths (Figure1.23).

Modification of the electron-donating group allowed us to generate fluorogens binding FAST with red-shifted emission (Chapter II). Increase of the conjugation and variation of the electron-withdrawing group enabled us to generate fluorogens with far-red emission when bound to engineered variants of FAST (Chapter III). Finally, introduction of charged groups on the heterocycle of HMBR allowed us to render HMBR membrane-impermeant for the selective labeling of membrane proteins (Chapter IV).

Based on this chapter, we have written a relative review.

Fluorogenic labeling strategies for biological imaging.

International Journal of Molecular Sciences 18, 1473-1483 (2017).

C. Li, A. G. Tebo & A. Gautier

I-4 References

[1] Weissleder R., Nahrendorf M., Advancing biomedical imaging, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, 112, 14424-14428.

[2] Valeur B., Molecular Fluorescence: Principles and Applications, *Wiley-VCH*, 2001.

[3] Valeur B., Berberan-Santos M.N., A brief history of fluorescence and phosphorescence before the emergence of quantum theory, *J. Chem. Educ.*, 2011, **88**, 731-738.

[4] Stokes G.G., *Philos. Trans.*, 1852, **142**, 463-562.

[5] Reichman J., Handbook of optical filters for fluorescence microscopy, *Chroma Technology Corp.*, 1998.

[6] Lakowicz J.R., Principles of Fluorescence Spectroscopy, 3rd edition, *Springe*, 2006.

[7] Chudakov D.M., Matz M.V., Lukyanov S., Lukyanov K.A., Fluorescent proteins and their applications in imaging living cells and tissues, *Physiol. Rev.* , 2010, **90**, 1103-1163.

[8] Demchenko A.P., Introduction to fluorescence sensing, *Springer Science + Business Media B.V.*, 2009.

[9] Moerner W.E., Shechtman Y., Wang Q., Single-molecule spectroscopy and imaging over the decades, *Faraday Discuss.*, 2015, **184**, 9-36.

[10] Kusumi A., Tsunoyama T.A., Hirosawa K.M., Kasai R.S., Fujiwara T.K., Tracking single molecules at work in living cells, *Nat. Chem. Biol.*, 2014, **10**, 524-532.

[11] Hell S.W., Far-field optical nanoscopy, *Science*, 2007, **316**, 1153-1158.

[12] Betzig E., Patterson G.H., Sougrat R., Lindwasser O.W., Olenych S., Bonifacino J.S. Davidson M.W., Lippincott-Schwartz J., Hess H.F., Imaging intracellular fluorescent protein at nanometer resolution, *Science*, 2006, **313**, 1642-1645.

[13] Gahlmann A., Moerner W.E., Exploring bavterial cell biology with single-molecule tracking and super-resolution imaging, *Nature Rev. Microbiol.*, 2014,**12**, 9-22.

[14] Lavis L.D., Raines R.T., Bright building blocks for chemical biology, *ACS Chem. Biol.*, 2014, **9**, 855-866.

[15] Lavis L.D., Raines R.T., Bright ideas for chemical biology, *ACS Chem. Biol.*, 2008, **3**, 142-155.

[16] Shimomura O., Johnson F.H., Saiga Y., Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea, *J. Cell. Comp. Physiol.* 1962, **59**, 223-239.

[17] Prasher D.C., Eckenrode V.K., Ward W.W., Prendergast F.G., Cormier, M.J., Primary structure of the Aequorea victoria green-fluorescent protein, *Gene*, 1992, **111**, 229-233.

[18] Chalfie M., Tu Y., Euskirchen G., Ward W.W., Prasher D.C., Green fluorescent protein as a marker for gene expression, *Science*, 1994, **263**, 802-805.

[19] Giepmans B.N.G., Adams S.R., Ellisman M.H., Tsien R.Y., The fluorescent toolbox for assessing protein location and function, *Science*, 2006, **312**, 217-224.

[20] Tsien R.Y., The green fluorescent protein, *Annu. Rev.Biochem.*, 1998, **67**, 509-554.

[21] Chudakov D.M., Lukyanov S., Lukyanov K.A., Fluorescent protein as a toolkit for *in vivo* imaging, *Trends Biotechnol.*, 2005, **23**, 605-613.

[22] Shaner N.C., Patterson G.H., Davidson M.W., Advances in fluorescent protein technology, *J. Cell Sci.*, 2007, **120**, 4247-4260

[23] Shaner N.C., Campbell R.E., Steinbach P.A., Giepmans B.N.G., Palmer A.E., Tsien R.Y., Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein, *Nat. Biotech.*, 2004, **22**, 1567-1572.

[24] Wang L., Jackson W.C., Steinbach P.A., Tsien R.Y., Evolution of new nonantibody proteins via iterative somatic hypermutation, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 16745-16749.

[25] Dean K.M., Palmer A.E., Advances in fluorescence labeling strategies for dynamic cellular imaging, *Nat. Chem. Biol.*, 2014, **10**, 512-523.

[26] Lippincott-Schwartz J., Altan-Bonnet N., Patterson G.H., Photobleaching and photoactivation: following protein dynamics in living cells, 2003, *Nat. Cell Biol.*, **Suppl.**, S7-S14.

[27] Day R.N., Davidson M.W., The fluorescent protein palette: tools for cellular imaging, 2009 , *Chem. Soc. Rev.*, **38**, 2887-2921.

[28] Jares-Erijman E.A., Jovin T.M., FRET imaging , *Nat. Biotechnol.*, 2003, **21**, 1387-1395.

[29] Kerppola T.K., Bimolecular fluorescence complementation: visualization of molecular interactions in living cells. *Methods Cell Biol.*, 2008, **85**, 431-470.

[30] Frommer W.B, Davidsonb M.W., Campbell R.E., Genetically encoded biosensors based on engineered fluorescent proteins, *Chem. Soc. Rev.*, 2009, **38**, 2833–2841

[31] Cubitt A.B., Heim R., Adams S.R., Boyd A.E., Gross L.A., Tsien R.Y., Understanding, improving and using green fluorescent proteins, *Trends Biochem. Sci.*, 1995, **20**, 448-545

[32] Heim R., Prasher D.C., Tsien R.Y., Wavelength mutations and posttranslational autoxidation of green fluorescent protein, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 12501-12504

[33] Heim R., Cubitt A.B., Tsien R.Y., Improved green fluorescence, *Nature*, 1995, **373**, 663-664.

[34] Shaner N.C., Lin M.Z., McKeown M.R., Steinbach P.A., Hazelwood K.L., Davidson M.W., Roger Y Tsien R.Y., Improving the photostability of bright monomeric orange and red fluorescent proteins, *Nat. Methods*, 2008, **5**, 545-551.

[35] Evdokimov A.G., Pokross M.E., Egorov N.S., Zaraisky A.G., Yampolsky I.V., Merzlyak E.M., Shkoporov A.N., Sander I., KonstantinA. Lukyanov K.A., Chudakov D.M., Structural basis for the fast maturation of Arthropoda green fluorescent protein, *EMBO Rep.*, 2006, **7**, 1006–1012.

[36] Adams S.R., Campbell R.E., Gross L.A., Martin B.R., Walkup G.K., Yao Y., Llopis J., Tsien R.Y., New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications, *J. Am. Chem. Soc.*, 2002, **124**, 6063-6076.

[37] Griffin B.A., Adams S.R., Tsien R.Y., Specific covalent labeling of recombinant protein molecules inside live cells, *Science*, 1998, **281**, 269-272

[38] Griffin B.A., Adams S.R., JONES J., Tsien R.Y., Fluorescent labeling of recombinant proteins in living cells with FlAsH, *Methods Enzymol.*, 2000, **327**, 565-578.

[39] Gaietta G., Deerinck T.J., Adams S.R., Bouwer J., Tour O., Laird D.W., Sosinsky G.E., Tsien R.Y., Ellisman M.H., Multicolor and electron microscopic imaging of connexin trafficking, *Science*, 2002, **296**, 503-507.

[40] Keppler A., Gendreizig S., Gronemeyer T., Pick H., Vogel H., Johnsson K., A general method for the covalent labeling of fusion proteins with small molecules *in vivo*, *Nat. Biotechnol*, 2003, **2**, 186-189.

[41] Keppler A., Pick H., Arrivoli C., Vogel H., Johnsson K., Labeling of fusion proteins with synthetic fluorophores in live cells, *Proc. Natl. Acad. Sci. U.S.A.*, 2004, **101**, 9955-9959.

[42] Gronemeyer T., Chidley C., Juillerat A., Heinis C., Johnsson K., Directed evolution of O^6-alkylguanine-DNA alkyltransferase for applications in protein labeling, *Protein Eng. Des. Select.*, 2006, **19**, 309-316.

[43] Juillerat A., Heinis C., Sielaff I., Barnikow J., Jaccard H., Kunz B., Terskikh A., Johnsson K., Engineering substrate specificity of O^6-alkylguanine-DNA alkyltransferase for specific protein labeling in living cells, *ChemBioChem*, 2005, **6**, 1263-1269.

[44] Kindermann M., Sielaff I., Johnsson K., Synthesis and characterization of bifunctional probes for the specific labeling of fusion proteins, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 2725–2728.

[45] Keppler A., Kindermann M., Gendreizig S., Pick H., Vogel H., Johnsson K., Labeling of fusion proteins of O^6-alkylguanine-DNA alkyltransferase with small molecules in vivo and in vitro, *Methods*, 2004, **32**, 437–444.

[46] Gronemeyer T., Godin G., Johnsson K., Adding value to fusion proteins through covalent labelling, *Curr. Opin. Biotechnol.*, 2005, **16**, 453–458.

[47] Gautier A., Juillerat A., Heinis C., Correa I. R., Kindermann M., Beaufils F., Johnsson K., An engineered protein tag for multiprotein labeling in living cells *Chem. Biol.*, 2008, **15**, 128–136.

[48] Gautier A., Nakata E., Lukinavicius G., Tan K.T., Johnsson K., Selective cross-linking of interacting proteins using self-labeling tags, *J. Am. Chem. Soc.*, 2009, **131**, 17954–17962.

[49] Los G.V., Encell L.P., McDougall M.G., Hartzell D.D., Karassina N., Zimprich C., Wood M.G., Learish R., Ohane 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., HaloTag: a novel protein labeling technology for cell imaging and protein analysis, *ACS Chem. Biol.*, 2008, **3**, 373-382.

[50] Jullien L., Gautier A., Fluorogen-based reporters for fluorescence imaging: a review, *Methods Appl. Fluoresc.*, 2015, **3**, 1-12.

[51] Komatsu T., Johnsson K., Okuno H., Bito H., Inoue T., Nagano T., Urano Y., J. Real-time measurements of protein dynamics using fluorescence activation-coupled protein labeling method, *J. Am. Chem. Soc.*, 2011, **133**, 6745-6751.

[52] Sun X., Zhang A., Baker B., Sun L., Howard A., Buswell J., Maurel D., Masharina A., Johnsson K., Noren C.J., Xu M.-Q., Corrêa , Jr. I.R., Development of SNAP-tag fluorogenic probes for wash-free fluorescence imaging, *ChemBioChem*, 2011, **12**, 2217-2226.

[53] Lukinavičius G., Umezawa K., Olivier N., Honigmann A., Yang G., Plass T., Mueller V., Reymond L., Corrêa, Jr. I.R., Luo Z.-G., Schultz C., Lemke E.A., Heppenstall P., Eggeling C., Manley S. and Johnsson K., A near-infrared fluorophore for live-cell superresolution microscopy of cellular proteins, *Nat. Chem.*, 2013, **5**, 132-139

[54] Klymchenko A.S., Solvatochromic and fluorogenic dyes as environment-sensitive probes: design and biological applications, *Acc. Chem. Res.*, 2017, **50**, 366-375.

[55] 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., Fluorogen-activating single-chain antibodies for imaging cell surface proteins, *Nat. Biotechnol.*, 2008, **26**, 235-240.

[56] 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., A rainbow of fluoromodules: A promiscuous scFv protein binds to and activates a diverse set of fluorogenic cyanine dyes, *J. Am. Chem. Soc.*, 2008, **130**, 12620-12621.

[57] Shank N.I., Zanotti K.J., Lanni F., Berget P.B., Armitage B.A., Enhanced photostability of genetically encodable fluoromodules based on fluorogenic cyanine dyes and a promiscuous protein partner, *J. Am. Chem. Soc.*, 2009, **131**, 12960-12969.

[58] Zanotti K.J., Silva G.L., Creeger Y., Robertson K.L., Waggoner A.S., Berget P.B., Armitage B.A., Blue fluorescent dye-protein complexes based on fluorogenic cyanine dyes and single chain antibody fragments, *Org. Biomol. Chem.*, 2011, **9**, 1012-1020.

[59] Telmer C.A., Verma R., Teng H., Andreko S., Law L., Bruchez M.P., Rapid, specific, no-wash, far-red fluorogen activation in subcellular compartments by targeted fluorogen activating proteins, *ACS Chem. Biol.*, 2015, **10**, 1239-1246.

[60] Hori Y., Ueno H., Mizukami S., Kikuchi K., Photoactive yellow protein-based protein labeling system with turn-on fluorescence intensity, *J. Am. Chem. Soc.*, 2009, **131**, 16610-16611.

[61] Hori Y., Nakaki K., Sato M., Mizukami S., Kikuchi K., Development of protein-labeling probes with a redesigned fluorogenic switch based on intramolecular association for no-wash live-cell imaging, *Angew. Chem. Int. Ed. Engl.*, 2012, **51**, 5611-5614.

[62] Hori Y., Norinobu T., Sato M., Arita K., Mizukami S., Kikuchi K., Development of fluorogenic probes for quick no-wash live-cell imaging of intracellular proteins, *J. Am. Chem. Soc.*, 2013, **135**, 12360-12365.

[63] Hori Y., Hirayama S., Sato M. Kikuchi K., Redesign of a fluorogenic labeling system to improve surface charge, brightness, and binding kinetics for imaging the functional localization of bromodomains, *Angew. Chem. Int. Ed. Engl.*, 2015, **54**, 14368-14371.

[64] Kamikawa Y., Hori Y., Yamashit K., Jin L., Hirayama S., Standleyb D.M., Kikuchi K., Design of a protein tag and fluorogenic probe with modular structure for live-cell imaging of intracellular proteins, *Chem. Sci.*, 2016, **7**, 308-314.

[65] Hirayama S., Hori Y., Benedek Z., Suzuki T., Kikuchi K., Fluorogenic probes reveal a role of GLUT4 N-glycosylation in intracellular trafficking, *Nat. Chem. Biol.*, 2016, **12**, 853-859.

[66] Meyer T.E., Isolation and characterization of soluble cytochromes, ferredoxins and other chromophoric proteins from the halophilic phototrophic bacterium *Ectothiorhodospira halophila*, *Biopchim. Biophys. Acta*, 1985, **806**, 175-183.

[67] Kleywegt G.J., Bergfors T., Senn H., Le Motte P., Gsell B., Shudo K. Jones T.A., Crystal structures of cellular retinoic acid binding proteins I and II in complex with all-*trans*-retinoic acid and a synthetic retinoid, *Structure*, 1994, **2**, 1241-1258.

[68] Li E., Norris A.W., Structure/ function of cytoplasmic vitamin A-binding proteins, *Annu. Rev. Nutr.*, 1996, **16**, 205-234.

[69] Chaudhuri B.N., Kleywegt G.J., Broutin-L'Hermite I., Bergfors T., Senn H., Le Motte P., Partouche O., Jones T.A., Structures of cellular retinoic acid binding proteins I and II in complex with synthetic retinoids, *Acta Cryst.*, 1999, **D55**, 1850-1857.

[70] Yapici I., Lee K.S.S., Berbasova T., Nosrati M., Jia X., Vasileiou C., Wang W., Santos E.M., Geiger J.H., Borhan B., 'Turn-on' protein fluorescence: *in situ* formation of cyanine dyes, *J. Am. Chem. Soc.*, 2015, **137**, 1073-1080.

[71] Babendure J.R., Adams S.R., Tsien R.Y., Aptamers switch on fluorescence of triphenyl-methane dyes, *J. Am. Chem. Soc.*, 2003, **125**, 14716-14717.

[72] Sparano B.A., Koide K., A strategy for the Development of small-molecule-based sensors that strongly fluoresce when bound to a specific RNA, *J. Am. Chem. Soc.*, 2005, **127**, 14954-14955.

[73] Sando S., Narita A., Hayami M., Aoyama Y., Transcription monitoring using fused RNA with a dye-binding light-up aptamer as a tag: a blue fluorescent RNA, *Chem. Commun.*, 2008, **33**, 3858-3860

[74] Constantin T.P., Silva G.L., Robertson K.L., Hamilton T.P., Fague K., Waggoner A.S., Armitage B.A., Synthesis of new fluorogenic cyanine dyes and incorporation into RNA fluoromodules, *Org. Lett.*, 2008, **10**, 1561-1564.

[75] Armitage B.A., Imaging of RNA in live cells, *Curr. Opin. Chem. Biol.*, 2011, **15**, 806-812.

[76] Rath A.K., Rentmeister A., Genetically encoded tools for RNA imaging in living cells, *Curr. Opin. Biotechnol.*, 2015, **31**, 42-49.

[77] Paige J.S., Wu K.Y., Jaffrey S.R., RNA mimics of green fluorescent protein, *Science*, 2011, **333**, 642-646.

[78] Strack R.L., Disney M.D., Jaffrey S.R., A superfolding Spinach2 reveals the dynamic nature of trinucleotide repeat–containing RNA, *Nat. Methods*, 2013, **10**, 1219-1227.

[79] Filonov G.S., Moon J.D., Svensen N., Jaffrey S.R., Broccoli: Rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution, *J. Am. Chem. Soc.*, 2014, **136**, 16299-16308.

[80] Strack R.L., Song W., Jaffrey S.R., Using Spinach-based sensors for fluorescence imaging of intracellular metabolites and proteins in living bacteria, *Nat. Protoc.*, 2014, **9**, 146-155.

[81] Plamont M.-A., Billon-Denis E., Maurin S., Gauron C., Pimenta F. M., Specht C. G., Shi J., Querard J., Pan B., Rossignol J., Moncoq K., Morellet N., Volovitch M., Lescop E., Chen Y., Triller A., Vriz S., Le Saux T., Jullien L., Gautier A., Small fluorescence-activating and

absorption-shifting tag for tunable protein imaging in vivo, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 497-502.

Chapter II Expansion of the spectral properties of FAST

This *chapter is based on the following article:*

Dynamic multi-color protein labeling in living cells.
Chemical Science (2017) just accepted DOI: 10.1039/C7SC01364G.
C. Li, M.-A. Plamont, H. Sladitschek, V. Rodrigues, I. Aujard, P. Neveu, T. Le Saux, L. Jullien & A. Gautier

II-1 Presentation of article 1

FAST is a promising fluorogen-based probe for multiplexed imaging with a lot of attractive advantages. To meet requirements for multiplexed imaging, we need to engineer new systems with various properties. A general strategy for the development of protein labeling systems is to extend their spectral properties and finally cover the whole visible spectrum. The marvelous versatility of molecules-based protein labeling systems enables to easily adjust properties of molecules to the need of the experiment through chemical modulations. So we tried to develop HMBR analogs that could selectively bind with FAST, preserve the main features of FAST but expend its spectral properties. Red-emitting probes are particularly attractive for biological research because the use of longer wavelengths reduces cellular autofluorescence, improves imaging contrast and decreases phototoxicity. One part of my PhD project was focused on the development of red-emitting FAST system. It was shown previously that FAST could accept various substituents on the aromatic ring. Moreover, aromatic substituents are a priori key determinants of the fluorescence efficiency, as they can play a major role in immobilizing the fluorogen within FAST cavity. In this chapter, we explored the role of substituents on the phenol ring of HMBR in varying spectral properties and binding affinities of FAST:fluorogen complexes. New fluorogens with conserved binding ability with FAST and showing different spectral properties are reported.

We designed and synthesized a collection of HMBR analogs with various aromatic substituents in different positions of the phenol ring (Figure 2.1). New fluorogens were obtained by condensation of the starting aldehydes and rhodanine in water or refluxing

ethanol in presence of piperidine. Their abilities to bind FAST were verified by characterizing their spectral properties alone or in the presence of FAST. We also tested their binding affinity by measuring thermodynamic dissociation constants (K_D). As expected, we observed fluorescence activation and absorption red shift in the presence of FAST, which demonstrated that they all kept the main features of the FAST approach. These FAST:fluorogen complexes displayed various spectroscopic behaviors.

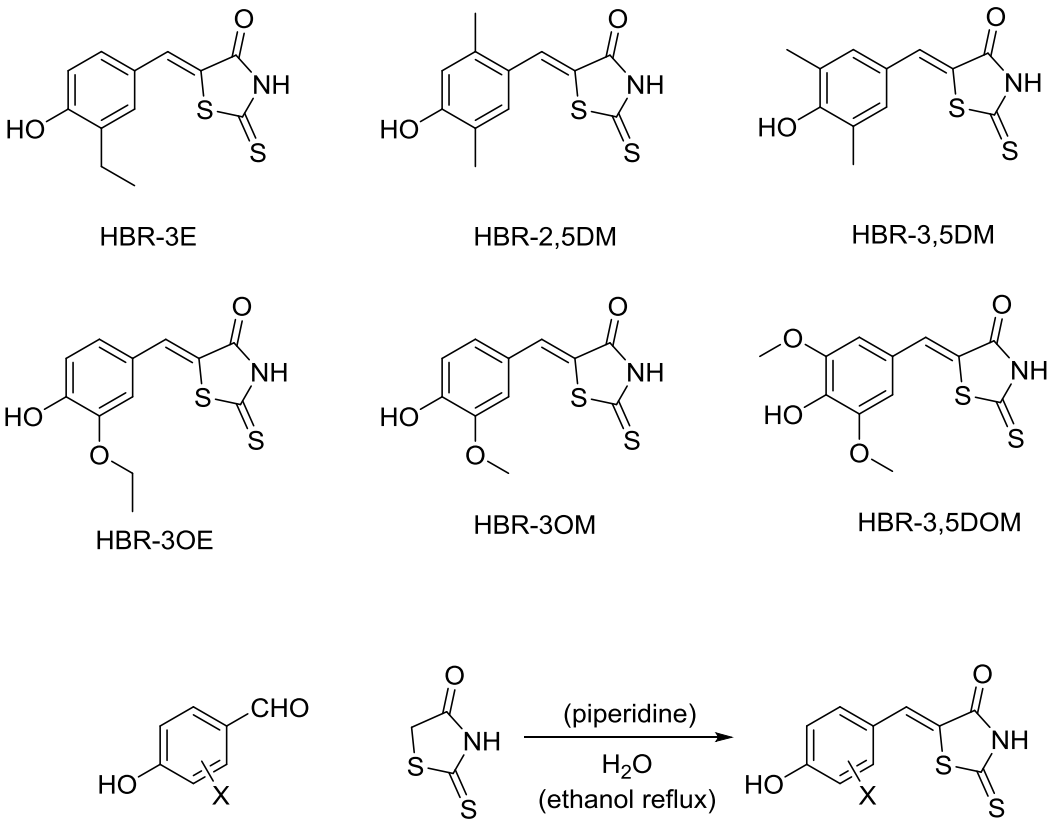


Figure 2.1 Structures of new fluorogens for FAST and general strategy for the synthesis of new fluorogens.

Compared to HMBR, HBR-3E in which an ethyl group replaces the methyl group of HMBR binds to FAST with higher affinity and forms complex displaying almost identical absorption and emission properties but reduced fluorescence brightness (See Article 1-Table 1). As brightness is directly related to the immobilization of the fluorogen within the protein cavity, we next examined the effect of an additional methyl group either in ortho-position or meta-position of the phenolic hydroxyl group. HBR-2,5DM and HBR-3,5DM bind tightly to FAST and form complexes displaying a 15-20 nm red shift in absorption and a 10-20 nm red shift in

emission with improved fluorescence brightness compared to FAST:HMBR (See Article 1-Table 1). Of particular interest, FAST:HBR-3,5DM is more than two times brighter than FAST:HMBR, reaching fluorescence efficiency of enhanced green fluorescent protein (EGFP).

To further shift absorption and emission to the red, we next replaced the methyl group of HMBR with alkoxy groups (methoxy or ethoxy group), which are common electron donors. HBR-3OM and HBR-3OE bind to FAST with affinities comparable to HMBR and form complexes displaying a 15 nm red shift in absorption and a 20 nm red-shift in emission. While FAST:HBR-3OE was less bright than FAST:HMBR, FAST:HBR-3OM exhibited greater brightness (See Article 1-Table 1). We next tested the introduction of a second methoxy group in ortho-position of the phenolic hydroxyl group. The resulting complex, FAST:HBR-3,5DOM, displayed a 40-nm red shift in absorption and a 60-nm red shift emission, and improved fluorescence brightness compared to FAST:HMBR (See Article 1-Table 1). These spectral modifications shift the absorption to the green and the emission to the orange-red, providing thus a new color for fluorescence imaging. In terms of fluorescence performance, FAST:HBR-3,5DOM was as bright as the red fluorescent protein mCherry.

In sum, we obtained four new fluorogens (HBR-3OM, HBR-2,5DM, HBR-3,5DM and HBR-3,5DOM) enabling to tune the fluorescence color of FAST from green-yellow to orange and red with higher brightness. Their pK_a are between 8.3-8.7 (See Article 1-Table S1) ensuring that they stay almost fully protonated at pH 7.4 with no fluorescence when free in solution. Moreover the range of K_D (~0.1-1µM) (See Article 1-Table 1) keeps their capabilities of reversible labeling with FAST-tagged proteins in living cells, while ensuring selective labeling.

We next verified their performance for labeling protein in living cells. Cytotoxicity testing proved that they were non-toxic for mammalian cells (See Article 1- Figure S2). Then we treated HeLa cells expressing FAST in different cell localizations (cytoplasm, nucleus or mitochondria) with the new fluorogens (HBR-3OM, HBR-3,5DM or HBR-3,5DOM). We observed specific fluorescence in the expected localizations through confocal microscopy (See Article 1-Figure 2). Their abilities of reversible labeling were examined by labeling HeLa cells expressing nuclear FAST repeatedly with HBR-3OM, HBR-3,5DM or HBR-3,5DOM solutions and then reversing the labeling by removing the fluorogen-containing solutions followed by a wash with pH 7.4 phosphate buffer saline (PBS) (See

Article 1-Figure 3). We observed that labeling of FAST could be switched on and off by addition and withdrawing of fluorogen and that this process was repeatable with various fluorogens which demonstrating the high tunability of the FAST technology.

The discovery of red-emitting FAST: HBR-3,5DOM makes the application of FAST system more versatile because now researchers have additional color options for different experimental conditions. In the presence of red fluorescent probes we can achieve multiplexed imaging by applying HMBR to label FAST-tagged proteins. On the contrary, in the presence of green fluorescent probes we can choose HBR-3,5DOM to label FAST-tagged proteins (See Article 1-Figure 2).

Another outstanding feature of FAST: HBR-3,5DOM complex is its high Stokes shift (82nm). We succeeded to distinguish FAST fusion protein with HBR-3,5DOM from EGFP-tagged protein by using a single excitation source (488nm) and two separate emission regions (See Article 1-Figure S3), which is significant for multiplexed imaging of living cells because one single excitation source simplifies imaging protocols and reduces phototoxicity.

Furthermore the absorption and emission difference between FAST:HMBR and FAST:HBR-3,5DOM is large enough to distinguish them as two different colors through microscopy (See Article 1-Figure 3). Moreover, we examined and proved that it was possible to switch the color (red-green) of FAST dynamically by changing the fluorogen solution (HBR-3,5DOM - HMBR) (See Article 1-Figure 4a). Using this unique feature, we developped an approach to extract the signal of FAST in the presence of both green and red fluorescent proteins by using only the green and the red channels, which was not evident for common fluorescent reporters because spectral discrimination is impossible in these conditions. The principle of this strategy is that the fluorescence signals of FAST:HMBR and FAST:HBR-3,5DOM are temporally anti-correlated during the dynamic color switching whereas fluorescence signals from fluorescent proteins stay stable. By measuring the degree of anti-correlation of the green and red fluorescence signals during the dynamic color switching, we could extract signals from FAST. To clearly show the performance of this strategy, we chose three separate cell localizations (nucleus, mitochondria and membrane) for expressing FAST, mCherry and EGFP. The localizations of mCherry and EGFP were determined by common imaging with red and green channels at the beginning. We firstly labeled FAST with HBR-3,5DOM and then replaced the medium with an HMBR solution recording a time-lapse video during the color switching. After analyzing time-lapse photos

and calculating the degree of anti-correlation of the green and red fluorescence signals, the correct localization of FAST was obtained (Article 1-Figure 4b). This experiments showed that dynamic multi-color protein labeling strategy could allow overcoming the limitation of observation channels, opening new opportunities for multiplexed imaging.

The labeling of FAST-tagged proteins with HMBR or HBR-3,5DOM is not orthogonal because they share the same protein tag. Ongoing projects in our team aim at selecting FAST variants that display orthogonal selectivity and eventually better spectral properties. Directed evolution approaches based on yeast display technology and Fluorescence-Activated Cell Sorting (FACS) are used to isolate variants from yeast-displayed FAST variant libraries (See also Chapter III). We are expecting FAST variants that could bind specifically and orthogonally with HMBR and HBR-3,5DOM with an improved brightness.

To further expand the spectral properties of FAST to the far-red edge of the visible spectrum, we also engineered new fluorogens and protein tags, which will be discussed in the Chapter III.

II-2 Article 1: Dynamic multi-color protein labeling in living cells

Chemical Science

EDGE ARTICLE

View Article Online
View Journal

Cite this: DOI: 10.1039/c7sc01364g

Dynamic multicolor protein labeling in living cells†

Chenge Li,[ab] Marie-Aude Plamont,[ab] Hanna L. Sladitschek,[c] Vanessa Rodrigues,[ab] Isabelle Aujard,[ab] Pierre Neveu,[c] Thomas Le Saux,[ab] Ludovic Jullien*[ab] and Arnaud Gautier *[ab]

Received 27th March 2017
Accepted 20th May 2017

DOI: 10.1039/c7sc01364g

rsc.li/chemical-science

Yellow Fluorescence-Activating and absorption-Shifting Tag (Y-FAST, hereafter called FAST) is a 14 kDa protein tag giving a bright green-yellow fluorescent complex upon interaction with the fluorogenic dye 4-hydroxy-3-methylbenzylidene rhodanine (HMBR). Here, we report a collection of fluorogens enabling tuning of the fluorescence color of FAST from green-yellow to orange and red. Beyond allowing the multicolor imaging of FAST-tagged proteins in live cells, these fluorogens enable dynamic color switching because of FAST's reversible labeling. This unprecedented behavior allows for selective detection of FAST-tagged proteins in cells expressing both green and red fluorescent species through two-color cross-correlation, opening up exciting prospects to overcome spectral crowding and push the frontiers of multiplexed imaging.

Introduction

Imaging the dynamics of proteins in living cells is essential for deciphering biological processes. A common strategy for imaging proteins is to fuse them to peptide or protein sequences that provide fluorescence, such as autofluorescent proteins[1,2] or self-labeling tags[3–5] (such as a SNAP-tag[6,7]/CLIP-tag[8] and Halo-tag[9]) that can be labeled specifically with chemical probes. A related strategy is to use protein tags that can generate fluorescence upon the interaction and activation of fluorogenic compounds (so-called fluorogens).[10,11] Fluorogen-activating proteins (FAPs) that are able to specifically bind to fluorogens, such as malachite green and thiazole orange, and activate their fluorescence have been developed from single-chain antibodies.[12,13] Because fluorogens exhibit very low fluorescence background in cells, proteins can be imaged with high contrast without the need for washing the fluorogen excess.

The concept of fluorogenic reporters was recently pushed one step further with the design of a tunable genetically encodable reporter called Yellow Fluorescence-Activating and absorption-Shifting Tag (Y-FAST, hereafter called FAST) that not only can be switched on by the addition of a fluorogenic ligand, but can also be switched off rapidly by the removal of the fluorogen through washing.[14] This fluorescence on/off switch behavior was made possible by designing a system displaying dynamic reversible fluorogen binding, characterized in particular by a high off-rate constant enabling rapid complex dissociation upon fluorogen removal. FAST is a 14 kDa variant of the photoactive yellow protein (PYP) engineered to specifically bind and activate the fluorescence of the fluorogenic dye HMBR. HMBR undergoes two spectroscopic changes upon binding, ensuring high labelling selectivity. While unbound HMBR mainly deexcites from the excited state through non-radiative processes, bound fluorogen undergoes a very large increase of the fluorescence quantum yield because of immobilization within the FAST cavity. Moreover, FAST binds HMBR in its anionic form, minor form in solution at pH 7.4, leading to an apparent 80 nm absorption red shift. FAST is thus an acronym that echoes back to its fast exchange dynamics and stands for Fluorescence-Activating and absorption-Shifting Tag, highlighting the two spectroscopic changes ensuring labeling selectivity.

The complex FAST:HMBR fluoresces green-yellow light (λ_{em} = 540 nm) upon blue light excitation. In order to provide investigators with the possibility to tune the spectral properties of FAST with their imaging conditions, we sought alternative fluorogens able to bind FAST and form complexes with novel fluorescence properties. We explored if changing the structure of HMBR could shift both the excitation and emission of the tag:fluorogen complex to the red edge of the visible spectrum (Fig. 1A). We report a collection of fluorogen analogs with various spectral properties. We present in particular HBR-3,5DOM (4-hydroxy-3,5-dimethoxybenzylidene rhodanine), a new fluorogen that forms a tight complex with FAST that fluoresces red light under green light excitation. This allows for the choice of the FAST color by choosing the appropriate fluorogen, providing an experimental versatility that is not

[a]École Normale Supérieure, PSL Research University, UPMC Univ Paris 06, CNRS, Département de Chimie, PASTEUR, 24 rue Lhomond, 75005 Paris, France
[b]Sorbonne Universités, UPMC Univ Paris 06, ENS, CNRS, PASTEUR, 75005 Paris, France. E-mail: ludovic.jullien@ens.fr; arnaud.gautier@ens.fr
[c]Cell Biology and Biophysics Unit, European Molecular Biology Laboratory, Meyerhofstr. 1, D-69117 Heidelberg, Germany

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c7sc01364g

View Article Online
Chemical Science Edge Article

encountered with autofluorescent proteins. We also show that we could exploit dynamic fluorogen binding to switch the color from red to green or green to red by simply changing the fluorogen used. The ability to dynamically switch color allowed us to further develop a new strategy to selectively image FAST-tagged proteins in a spectrally crowded environment. The spectral overlap between fluorophores limits most experiments to the simultaneous observation of three or four targets, thus preventing highly multiplexed imaging. Here, we show that one can overcome spectral crowding using kinetic filtering based on dynamic color switching and two-color cross-correlation.

Results and discussion

The HMBR fluorogen is composed of an electron-donating phenol ring conjugated to an electron-withdrawing rhodanine heterocycle. This push–pull structure deexcites non-radiatively in solution but relaxes radiatively to the ground state within the FAST cavity. The substituents on the aromatic ring are thus *a priori* key determinants of the fluorescence efficiency, as they can play a major role in the (im)mobility of the fluorogen within the FAST cavity. In addition, the spectral properties of the push–pull systems, such as HMBR and its analogs, are directly related to the electron donating or withdrawing ability of the donor and acceptor groups, as they affect the energy of the molecular orbitals. Changing the electron-donating properties of the phenol moiety through adequate substitution can thus affect the spectroscopic properties. In particular, here we can anticipate that the stronger the electron donating ability of the donor part, the more red-shifted the absorption and emission. We thus examined the role of the substituents on the aromatic ring. Fig. 1B gives the structure of the different analogs obtained by condensation of the rhodanine heterocycle with appropriate aromatic aldehydes. The described fluorogens bound tightly to FAST with dissociation constants below 1 μM, enabling to anticipate efficient and selective FAST labeling (Table 1). In addition, all the fluorogenic analogs underwent a very large fluorescence increase (from 200 up to 600 fold) and absorption red-shift (between 80 and 110 nm) upon FAST binding (Table 1). This latter absorption red-shift is due to the binding-induced deprotonation of the fluorogens, which are all mainly protonated in solution at physiological pH (Table S1 and Fig. S1†).

We observed various spectroscopic behaviors. HBR-2,5DM, HBR-3,5DM, HBR-3OM and HBR-3OE formed complexes with ~15–20 nm red-shifted absorption (λ_{abs} from 494 to 499 nm) and ~10–20 nm red-shifted emission (λ_{em} from 552 to 562 nm) relative to FAST:HMBR (see also Fig. 1C and D). Of particular interest, FAST:HBR-3,5DM exhibits a fluorescence quantum yield of 49%, and an overall brightness 2.5-fold higher than the original FAST:HMBR complex. The greatest spectral shift was observed with HBR-3,5DOM, which bears two electron-donating methoxy substituents in the *ortho* position of the hydroxyl group. The resulting complex, FAST:HBR-3,5DOM, displayed a 40 nm red-shifted absorption (λ_{abs} = 518 nm) and a 60 nm red-shifted emission (λ_{em} = 600 nm) (see also Fig. 1E). These spectral modifications shift the absorption to the green and the emission to the red, thus providing a new color for the fluorescence imaging. Remarkably, FAST:HBR-3,5DOM is as bright as the red fluorescent protein mCherry.

We next examined if our new fluorogenic analogs could efficiently label FAST-tagged proteins in living cells. We first verified that they were non-toxic for mammalian cells (Fig. S2†). Then, we expressed in HeLa cells (i) FAST, (ii) FAST fused to histone H2B for nuclear localization, and (iii) FAST fused to the mitochondrial targeting sequence (MTS) from subunit VIII of human cytochrome c oxidase. The cells were treated with media

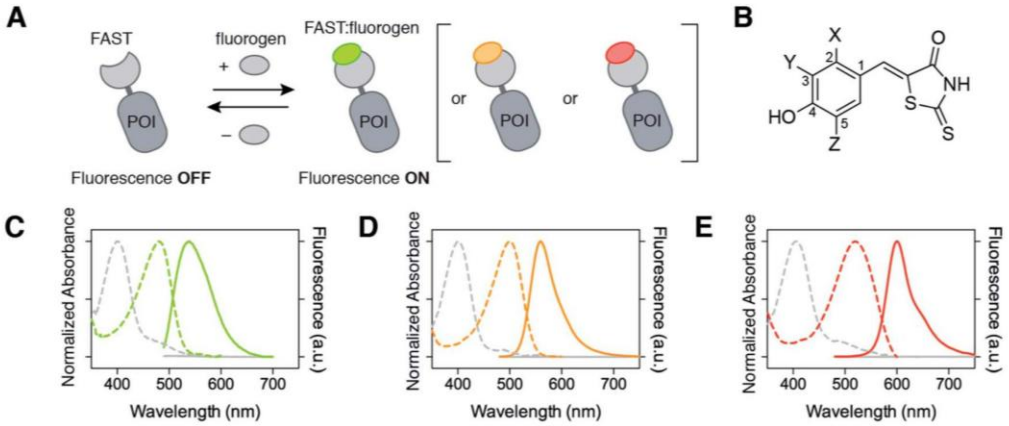


Fig. 1 Multicolor imaging of FAST-tagged proteins. (A) The fluorescence color of FAST can be adapted by changing the fluorogen (POI: protein of interest). (B) Structures of the fluorogens discussed in this study: HMBR (X = H, Y = Me, Z = H), HBR-3E (X = H, Y = Et, Z = H), HBR-2,5DM (X = Me, Y = H, Z = Me), HBR-3,5DM (X = H, Y = Me, Z = Me), HBR-3OM (X = H, Y = OMe, Z = H), HBR-3OE (X = H, Y = OEt, Z = H), and HBR-3,5DOM (X = H, Y = OMe, Z = OMe). (C–E) The absorption (dashed line, left axis) and emission (solid line, right axis) spectra of the free fluorogens (grey line) and FAST:fluorogen complexes (colored lines) in pH 7.4 phosphate buffer saline at 25 °C ((C) HMBR, (D) HBR-3,5DM, and (E) HBR-3,5DOM). Note that the shoulders at 480 nm on the absorption spectra of the free fluorogens correspond to anionic fluorogens present in low amounts at pH 7.4 (see also Fig. S1 and Table S1†).

Table 1 Physicochemical properties of FAST:fluorogen complexes in PBS pH 7.4. The abbreviations are as follows: λ_{abs}, the wavelength of maximal absorption; λ_{em}, the wavelength of maximal emission; ε, the molar absorption coefficient at λ_{abs}; $\Delta\lambda_{abs} = \lambda_{abs,bound} - \lambda_{abs,unbound}$, the absorption red-shift upon FAST binding; ϕ, the fluorescence quantum yield; $F_{bound}/F_{unbound}$, the fluorescence activation upon FAST binding; brightness $= \varepsilon \times \phi$; K_D, the thermodynamic dissociation constant. The spectroscopic data of mCherry[15] are given for comparison

Complex	λ_{abs} (nm)	$\Delta\lambda_{abs}$ (nm)	λ_{em} (nm)	ε (mM^{-1} cm^{-1})	ϕ (%)	$F_{bound}/F_{unbound}$	Brightness	K_D (μM)
FAST:HMBR	481	80	540	45	23	300	10 300	0.13
FAST:HBR-3E	481	81	543	50	13	300	6500	0.08
FAST:HBR-2,5DM	494	88	552	50	29	430	14 500	0.008
FAST:HBR-3,5DM	499	98	562	48	49	660	23 500	0.08
FAST:HBR-3OM	494	91	561	40	36	350	14 400	0.31
FAST:HBR-3OE	497	94	562	41	17	230	7000	0.27
FAST:HBR-3,5DOM	518	112	600	39	31	220	12 000	0.97
mCherry	587		610	72	22		15 800	

containing 5 μM of fluorogen and imaged by confocal microscopy (Fig. 2A). We visualized the various FAST fusions in their expected localization, in agreement with the specific fluorogenic labeling.

Given that the fluorescence color of FAST can be either green-yellow (with HMBR) or red (with HBR-3,5DOM), we next showed that FAST could be imaged in the presence of either the red fluorescent protein mCherry or the green fluorescent protein EGFP by adding the complementary fluorogen. We expressed FAST fused to H2B for nuclear targeting together with either mCherry or EGFP fused to lyn N-terminal domain (lyn) for membrane anchoring. The cells co-expressing FAST and mCherry fusions were treated with media containing 5 μM HMBR, while the cells co-expressing FAST and EGFP fusions were treated with media containing 5 μM of HBR-3,5DOM. By using a dual 488/543 nm excitation and variable spectral detection, FAST:HMBR could be separated spectrally from mCherry, while FAST:HBR-3,5DOM could be distinguished from EGFP (Fig. 2B), demonstrating the versatility of FAST and the possibility to easily adapt its color to the spectral conditions of the experiment. An additional interesting feature of FAST:HBR-3,5DOM is its high Stokes shift (82 nm), which allowed us to image the FAST-tagged proteins together with the EGFP fusions using a single 488 nm excitation source and two separate spectral regions (Fig. S3†), opening up interesting prospects for live cell imaging as single excitation simplifies the imaging protocols, results in a gain of temporal resolution and reduces the phototoxicity.

A unique feature that distinguishes FAST from other labeling techniques is the possibility to rapidly reverse the labeling by washing the fluorogen away. We thus verified that we kept this property when using the new fluorogenic analogs. We expressed FAST fused to histone H2B in mammalian cells. Iterative labeling/washing with the new fluorogens allowed us to demonstrate that FAST conserved its ability to be switched on and off at will by fluorogen addition and removal (Fig. 3A). Moreover, as HMBR and HBR-3,5DOM give complexes of different fluorescence colors, FAST could be switched from one color to the other by washing away one fluorogen and adding the other one (Fig. 3B). The color switch could be done furthermore in a dynamic fashion in a few seconds simply by changing the medium content (Fig. 3C and Movie S1†).

The ability to dynamically switch color by fluorogen exchange provides FAST with a unique signature that can be exploited to push multiplexed imaging. Modern biology needs to be able to observe multiple targets (>10–50) to study biological processes in all their complexity.[16] Despite the large collection of probes and biosensors available, distinguishing a large number of labels remains one of the biggest challenges in imaging. A major obstacle to multiple observations is the spectral overlap between fluorophores, which limits most experiments to the observation of three or four targets. New approaches for highly multiplexed imaging relying on parameters other than spectral discrimination are thus essential if one wants to be able to better analyze multiparametric cellular processes. We thus explored if the unique signature provided by dynamic color switching could provide a solution for overcoming spectral crowding. Upon dynamic color switching, HMBR and HBR-3,5DOM are fluorescent in a temporally anti-correlated manner, as they both bind to the same tag. Kinetic information has been previously exploited to image selectively reactive fluorescent species in the presence of a high fluorescence background.[17–19] We thus anticipated that measuring the degree of anti-correlation of the green and red fluorescence signals should enable the extraction of FAST contribution in the presence of the green and red fluorescent species – even though spectral discrimination is impossible in these conditions – exploiting the fact that only the green and red signals originating from FAST would interconvert (Fig. 4A). To evaluate this idea, we used time-lapse confocal microscopy to image HeLa cells co-expressing FAST (fused to histone H2B), mCherry (fused to a mitochondrial targeting sequence (MTS)) and EGFP (fused to lyn for membrane anchoring) upon the dynamic fluorogen exchange (Fig. 4B, S4A and B and Movie S2†). The cross-correlation of the temporal evolution of the green and red fluorescence signals allowed us to selectively extract the contribution of H2B-FAST and fully eliminate the signals of mitochondrial mCherry and membrane-targeted EGFP (Fig. 4B). Note that we chose different localizations for FAST, mCherry and EGFP to assign the proteins without any ambiguity in this proof-of-principle experiment, but the approach is

View Article Online
Chemical Science Edge Article

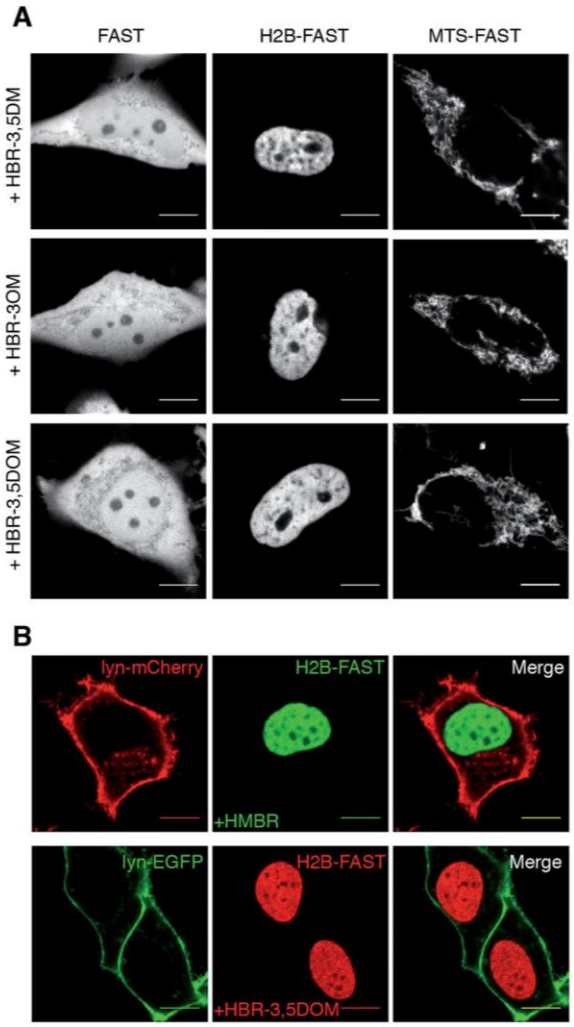


Fig. 2 Multicolor labeling of the FAST-tagged proteins in live cells. (A) Confocal micrographs of HeLa cells expressing cytoplasmic FAST, nuclear H2B-FAST and mitochondrial MTS-FAST and labeled with 5 μM HBR-3,5DM, HBR-3OM and HBR-3,5DOM (HBR-3OM, HBR-3,5DM: Ex/Em 488/493–797 nm; HBR-3,5DOM: Ex/Em 514/519–797 nm). (B) Confocal micrographs of live HeLa cells co-expressing lyn-mCherry/H2B-FAST (top) and lyn-EGFP/H2B-FAST (bottom) labeled with 5 μM HMBR or HBR-3,5DOM, respectively (green channel Ex/Em 488/493–575 nm; red channel Ex/Em 543/578–797 nm). The scale bars are 10 μm.

generic and can of course be expanded to co-localized proteins. This experiment demonstrates that dynamic color switching and two-color cross-correlation allow unveiling of the presence of FAST in a spectrally crowded environment containing both green and red fluorescent species.

Conclusion

This study presents fluorogenic dyes expanding the spectral properties of FAST to the red edge of the visible spectrum. Originally, FAST emitted green-yellow light upon blue light excitation when using HMBR as a fluorogenic partner. Here, we show that FAST can fluoresce red light upon green light excitation using HBR-3,5DOM instead of HMBR. With FAST, it is thus now possible to adapt the tag color to the spectral conditions of the experiment without changing the protein tag, providing an experimental versatility that is not encountered with autofluorescent proteins. In addition, the ability to change the color by changing the dye renders it possible to switch the color in a single experiment.[20,21] In this work, we show that the fast exchange dynamics characterizing FAST and the high permeability of the designed fluorogens allow color swapping on the second timescale. We used this unique kinetic signature to develop a method for the selective imaging of intracellular FAST-tagged proteins in spectrally crowded environments. A major obstacle to multiple observations is the spectral overlap between fluorophores, which limits most experiments to the observation of three or four targets. By measuring the degree of temporal anti-correlation of the two fluorescent states of FAST during dynamic color switching, we successfully imaged FAST-tagged proteins in the presence of green and red fluorescent species despite spectral crowding. Imaging three proteins with only two observation channels is a promising step towards overcoming spectral crowding and opens up exciting prospects for pushing the frontiers of multiplexed imaging.

Materials and methods

Chemical synthesis—general information

Commercially available reagents were used as the starting materials without further purification. The NMR spectra were recorded on an AC Bruker spectrometer at 300 MHz for ^{1}H NMR and 75 MHz for ^{13}C NMR; the chemical shifts are reported in ppm with protonated solvent as an internal reference in ^{1}H, $CHCl_3$ in $CDCl_3$ 7.26 ppm, CHD_2SOCD_3 in CD_3SOCD_3 2.50 ppm; ^{13}C, $^{13}CDCl_3$ in $CDCl_3$ 77.0 ppm, $^{13}CD_3SOCD_3$ in CD_3SOCD_3 39.5 ppm; the coupling constants J are given in Hz. Mass spectroscopy (chemical ionization or high resolution) was performed by the Service de Spectrométrie de Masse de Chimie ParisTech and the Institut de Chimie Organique et Analytique de l'Université d'Orléans. Column chromatography was performed on silica gel 60 (0.040–0.063 nm) Merck. Analytical thin-layer chromatography (TLC) was conducted on Merck silica gel 60 F254 precoated plates.

Synthesis of 4-hydroxy-3-ethylbenzaldehyde

To a solution of 2-ethylphenol (6.1 g, 50 mmol) in 10% aqueous sodium hydroxide (80 mL, 200 mmol), trichloromethane (15.0 g, 125 mmol) was added dropwise at 60 °C over 1 h, and then the reaction mixture was stirred for 2 h at 60 °C. After cooling, the mixture was neutralized by an aqueous solution of hydrochloric acid and extracted with dichloromethane. The combined organic layers were washed with brine, dried over magnesium sulfate, and concentrated under reduced pressure. The residue was purified by flash chromatography on silica gel with cyclohexane/ethylacetate (7.5/2.5, v/v) to yield the desired

View Article Online
Edge Article Chemical Science

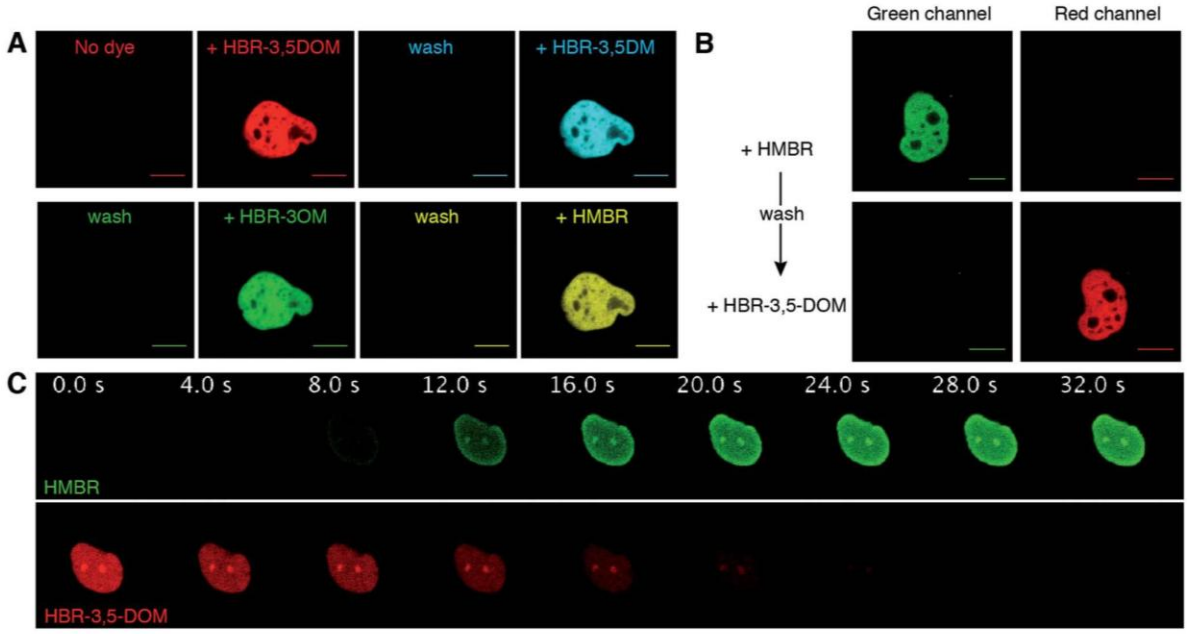


Fig. 3 Reversible FAST labeling. (A) Confocal micrographs of a HeLa cell expressing nuclear H2B-FAST sequentially labeled with HBR-3,5DOM (red), HBR-3,5DM (cyan), HBR-3OM (green) and HMBR (yellow) (HBR-3OM, HBR-3,5DM, HMBR: Ex/Em 488/493–797 nm; HBR-3,5DOM: Ex/Em 514/519–797 nm). The labeling was reversed by removing the fluorogen-containing solutions followed by a wash with pH 7.4 PBS. (B) Confocal micrographs of a HeLa cell expressing nuclear H2B-FAST sequentially labeled with HMBR and HBR-3,5DOM (green channel: Ex/Em 488/493–550 nm; red channel: Ex/Em 543/600–797 nm). The color switch was performed by first washing away the excess of HMBR with pH 7.4 PBS, and then adding HBR-3,5DOM solution. (C) Time series of a live HeLa cell expressing H2B-FAST labeled with HBR-3,5DOM upon the replacement of the medium with a HMBR solution (single excitation at 488 nm; HMBR emission channel 493–538 nm; HBR-3,5-DOM emission channel 649–797 nm). HMBR was added at $t = 0$ s. See also Movie S1.† (A–C) The concentration of the fluorogen solutions was 5 μM. The cells were grown in a minifluidic channel enabling easy solution replacement. The scale bars are 10 μm.

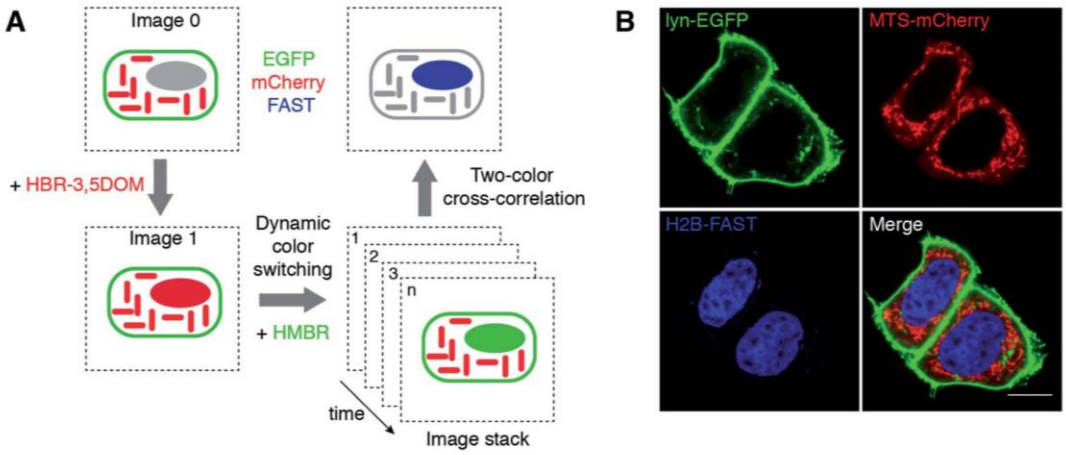


Fig. 4 Selective imaging. (A) Measuring the degree of anti-correlation of the green and red fluorescence signals upon dynamic color switching by two-color cross-correlation is an effective method to quantitatively extract the signal of FAST in a spectrally crowded environment containing green and red fluorescent species. (B) The micrographs show live HeLa cells co-expressing three proteins: lyn-EGFP, MTS-mCherry and H2B-FAST. The lyn-EGFP (green) and MTS-mCherry (red) images are confocal micrographs obtained in the absence of fluorogen, while the H2B-FAST (blue) image results from two-color cross-correlation upon the HBR-3,5DOM-to-HMBR exchange (green channel Ex/Em 488/493–575 nm; red channel Ex/Em 543/578–797 nm). Fig. S4 and Movie S2† show the HBR-3,5DOM-to-HMBR exchange. The scale bars are 10 μm.

4-hydroxy-3-ethylbenzaldehyde (1.1 g, 15% yield) as a grey pink solid. ^{1}H NMR (300 MHz, $CDCl_3$, δ in ppm): 9.84 (s, 1H), 7.72 (s, 1H), 7.65 (d, J = 8.1 Hz, 1H), 6.92 (d, J = 8.1 Hz, 1H), 6.51 (s, 1H), 2.70 (q, J = 7.5 Hz, 2H), 1.27 (t, J = 7.5 Hz, 3H); ^{13}C NMR (75 MHz, $CDCl_3$, δ in ppm): 192.3, 160.5, 131.5, 131.4, 130.6, 129.4, 115.6, 22.8, 13.5; MS (ESI): m/z 149.2 $[M - H]^-$, calcd mass for $[C_9H_9O_2]^-$: 149.1; HRMS (ESI): m/z 149.0608 $[M - H]^-$, calcd mass for $[C_9H_9O_2]^-$: 149.0603.

General protocol for the synthesis of HBR-3OM and HBR-2,5DM

A solution containing rhodanine (133 mg, 1.0 mmol) and substituted 4-hydroxy-benzaldehyde (1.0 mmol) in 40 mL of water was stirred at 90 °C for 7 days. After cooling to 4 °C and standing overnight, the precipitate was filtered through a glass filter and the crude solid was washed with water and ethanol and dried over P_2O_5, to give the desired product.

(*Z*)-5-(4-Hydroxy-3-methoxybenzylidene)-2-thioxo-1,3-thiazolidin-4-one (HBR-3OM). Yellow powder (52%). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 13.68 (s, 1H), 10.08 (s, 1H), 7.57 (s, 1H), 7.15 (s, 1H), 7.08 (d, J = 8.1 Hz, 1H), 6.94 (d, J = 8.1 Hz, 1H), 3.83 (s, 3H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 195.5, 169.5, 150.0, 148.1, 132.8, 125.1, 124.4, 121.2, 116.4, 114.4, 55.7; MS (ESI): m/z 266.2 $[M - H]^-$, calcd mass for $[C_{11}H_8NO_3S_2]^-$: 266.0; HRMS (ESI): m/z 265.9951 $[M - H]^-$, calcd mass for $[C_{11}H_8NO_3S_2]^-$: 265.9946.

(*Z*)-5-(4-Hydroxy-2,5-dimethylbenzylidene)-2-thioxo-1,3-thiazolidin-4-one (HBR-2,5DM). Orange powder (38%). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 13.70 (s, 1H), 10.22 (s, 1H), 7.64 (s, 1H), 7.10 (s, 1H), 6.75 (s, 1H), 2.33 (s, 3H), 2.14 (s, 3H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 196.3, 169.8, 158.9, 140.1, 131.1, 130.0, 123.4, 122.9, 122.0, 117.7, 19.6, 16.1; MS (ESI): m/z 264.1 $[M - H]^-$, calcd mass for $[C_{12}H_{10}NO_2S_2]^-$: 264.0; HRMS (ESI): m/z 264.0158 $[M - H]^-$, calcd mass for $[C_{12}H_{10}NO_2S_2]^-$: 264.0153.

General protocol for the synthesis of HBR-3E, HBR-3OE, HBR-3,5DM and HBR-3,5DOM

To a stirred solution of rhodanine (1.5 mmol, 1.0 eq.) and substituted 4-hydroxy-benzaldehyde (1.56 mmol, 1.1 eq.) in 4.5 mL of absolute ethanol, piperidine (1.5 mmol, 1.0 eq.) was added. The solution was stirred at room temperature for 16 h and then neutralized by an aqueous solution of hydrochloric acid. After cooling to 4 °C and standing for 2 h, the precipitate was filtered through a glass filter and the crude solid was washed with water and ethanol and dried over P_2O_5, to give the desired product.

(*Z*)-5-(3-Ethyl-4-hydroxybenzylidene)-2-thioxothiazolidin-4-one (HBR-3E). Yellow powder (70%). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 13.67 (s, 1H), 10.35 (s, 1H), 7.54 (s, 1H), 7.30 (m, 2H), 6.94 (d, J = 8.1 Hz, 1H), 2.57 (q, J = 7.5 Hz, 2H), 1.15 (t, J = 7.5 Hz, 3H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 195.5, 169.4, 158.3, 132.8, 132.0, 131.3, 130.6, 124.0, 120.5, 115.8, 22.5, 13.7; MS (ESI): m/z 264.3 $[M - H]^-$, calcd mass for $[C_{12}H_{10}NO_2S_2]^-$: 264.0; HRMS (ESI): m/z 264.0157 $[M - H]^-$, calcd mass for $[C_{12}H_{10}NO_2S_2]^-$: 264.0153.

(*Z*)-5-(3-Ethoxy-4-hydroxybenzylidene)-2-thioxothiazolidin-4-one (HBR-3OE). Orange powder (74%). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 13.69 (s, 1H), 10.00 (s, 1H), 7.55 (s, 1H), 7.12 (d, J = 1.8 Hz, 1H), 7.07 (dd, J = 1.8 Hz, 8.4 Hz, 1H), 6.95 (d, J = 8.1 Hz, 1H), 4.08 (q, J = 6.9 Hz, 2H), 1.36 (t, J = 6.9 Hz, 3H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 195.4, 169.4, 150.2, 147.3, 132.8, 125.1, 124.4, 121.0, 116.4, 115.4, 63.9, 14.6; MS (ESI): m/z 280.1 $[M - H]^-$, calcd mass for $[C_{12}H_{10}NO_3S_2]^-$: 280.0; HRMS (ESI): m/z 280.0108 $[M - H]^-$, calcd mass for $[C_{12}H_{10}NO_3S_2]^-$: 280.0102.

(*Z*)-5-(4-Hydroxy-3,5-dimethylbenzylidene)-2-thioxothiazolidin-4-one (HBR-3,5DM). Yellow powder (73%). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 13.67 (s, 1H), 9.26 (s, 1H), 7.45 (s, 1H), 7.17 (s, 2H), 2.21 (s, 6H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 195.6, 169.5, 156.6, 132.6, 131.5 (2C), 125.4 (2C), 124.0, 120.8, 16.6 (2C); MS (ESI): m/z 264.3 $[M - H]^-$, calcd mass for $[C_{12}H_{10}NO_2S_2]^-$: 264.0; HRMS (ESI): m/z 264.0158 $[M - H]^-$, calcd mass for $[C_{12}H_{10}NO_2S_2]^-$: 264.0153.

(*Z*)-5-(4-Hydroxy-3,5-dimethoxybenzylidene)-2-thioxothiazolidin-4-one (HBR-3,5DOM). Brown powder (72%). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 13.71 (s, 1H), 9.46 (s, 1H), 7.57 (s, 1H), 6.88 (s, 2H), 3.83 (s, 6H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 195.3, 169.3, 148.3 (2C), 139.2, 133.0, 123.2, 121.4, 108.5 (2C), 56.1 (2C); MS (ESI): m/z 296.1 $[M - H]^-$, calcd mass for $[C_{12}H_{10}NO_4S_2]^-$: 296.0; HRMS (ESI): m/z 296.0056 $[M - H]^-$, calcd mass for $[C_{12}H_{10}NO_4S_2]^-$: 296.0051.

Physical chemistry experiments

pH measurements were performed on a standard pH meter PHM210 Radiometer Analytical (calibrated with aqueous buffers at pH 4 and 7 or 10) with a Crison 5208 Electrode (Barcelona, Spain). The UV/Vis absorption spectra were recorded with 1 cm × 1 cm quartz cuvettes (Hellma) on a diode array UV/Vis spectrophotometer (Evolution array, Thermo Scientific). The corrected fluorescence spectra upon one-photon excitation were recorded with a Photon Technology International QuantaMaster QM-1 spectrofluorimeter (PTI, Monmouth Junction, NJ) equipped with a Peltier cell holder (TLC50, Quantum Northwest, Shoreline, WA). The overall emission quantum yields after one-photon excitation ϕ were determined as previously described.[14] The affinity constants were determined by spectrofluorometric titration using a Spark 10M plate reader (Tecan) following the protocols previously described.[14]

Molecular biology

FAST is a variant of the photoactive yellow protein (PYP) containing the mutations C69G, Y94W, T95M, F96I, D97P, Y98T, Q99S, M100R, and T101G. The construction of plasmid pAG87 for the bacterial expression of FAST fused to a His-tag was previously described.[14] The construction of the plasmids pAG104, pAG109 and pAG106 for the mammalian expression of FAST, FAST fused to zebrafish H2B (H2B-FAST), and FAST fused to the mitochondrial targeting sequence (MTS) from subunit VIII of human cytochrome c oxidase (MTS-FAST) was previously described.[14] Plasmid pAG28 for the mammalian expression of EGFP fused to lyn N-terminal domain (lyn) for inner membrane

anchoring was previously described.[18] The plasmids CV2307 and CV2326 for the mammalian expression of mCherry fused to the membrane anchoring (lyn) peptide from the mouse tyrosine-protein kinase lyn and to the N-terminal 25 amino acid mitochondrial targeting signal (MTS) of COX4 from *S. cerevisiae* were assembled using MXS chaining by sequentially inserting a CMV promoter, and the sequences of the lyn (MGCIKSKRKDVEN) or MTS (MLSLRQSIRFFKPATRTLCSSRYLL) peptide tags in a plasmid containing the mCherry coding sequence coupled to the bovine growth hormone poly A signal.

Bacterial expression and protein purification

pAG87 was transformed in Rosetta(DE3)pLysS *E. coli* (New England Biolabs). The cells were grown at 37 °C in Lysogeny Broth (LB) medium complemented with 50 μg mL^{-1} kanamycin and 34 μg mL^{-1} chloramphenicol to $OD_{600\ nm}$ 0.6. Expression was induced for 4 h by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. The cells were harvested by centrifugation (6000 × *g* for 15 min at 4 °C) and frozen. The cell pellet was resuspended in lysis buffer (phosphate buffer 50 mM, NaCl 150 mM, $MgCl_2$ 2.5 mM, protease inhibitor, DNase, pH 7.4) and sonicated (5 min at 20% of amplitude). The lysate was incubated for 2 h at 4 °C to allow DNA digestion by DNase. The cellular fragments were removed by centrifugation (15 000 × *g* for 1 h at 4 °C). The supernatant was incubated overnight at 4 °C under gentle agitation with Ni-NTA agarose beads in phosphate buffered saline (PBS) (sodium phosphate 50 mM, NaCl 150 mM, pH 7.4) complemented with 10 mM imidazole. The beads were washed with 20 volumes of PBS containing 20 mM imidazole, and with 5 volumes of PBS complemented with 40 mM imidazole. His-tagged proteins were eluted with 5 volumes of PBS complemented with 0.5 M imidazole, followed by dialysis with PBS.

Mammalian cell culture

The HeLa cells were cultured in DMEM supplemented with phenol red, Glutamax I, 10% (v/v) fetal calf serum (FCS), and 1% penicillin–streptomycin at 37 °C within a 5% CO_2 atmosphere. For microscopic imaging, the cells were seeded in μDish or μSlide IBIDI (Biovalley) coated with poly-L-lysine. The cells were transiently transfected using Genejuice (Merck) according to the manufacturer's protocol. The cells were washed with PBS, and treated with DMEM media (without serum and phenol red) containing the fluorogens at the indicated concentration. The cells were imaged directly without washing.

Cell viability assay

The HeLa cells were treated with DMEM media containing the fluorogens at the indicated concentrations for the indicated times. The cell viability was evaluated by fluorescence microscopy using the LIVE/DEAD® viability/cytotoxicity assay kit (Molecular Probes, Life Technologies) following the manufacturer's protocol.

Fluorescence microscopy

The confocal micrographs were acquired on a Zeiss LSM 710 Laser Scanning Microscope equipped with a Plan Apochromat 63×/1.4 NA oil immersion objective. ZEN software was used to collect the data. The images were analyzed with Image J.

Two-color cross-correlation analysis

The two-color cross-correlated images were obtained as follows. A movie was acquired in the green and red (g: green and r: red) fluorescence channels over the evolution duration observed after adding HMBR fluorogen to the cellular medium. The cross-correlated image was then computed from the movie as $\mathcal{I}[\sqrt{\sum_i (p_i^g - \bar{p}^g)(p_i^r - \bar{p}^r)}]$ where $\mathcal{I}$ takes the imaginary part of the FAST concentration proportional square root function, i denotes the frame number, p_i^g (p_i^r) is the fluorescence intensity of pixel p in the green (red) channel of the image i, and $\bar{p}^g$ ($\bar{p}^r$) denotes the average fluorescence intensity of pixel p over the whole movie duration in the green (red) channel. Eventually a median filter (3 × 3 $pixel^2$) was applied to yield the final image.

Acknowledgements

This work has been supported by the Agence National de la Recherche (ANR-14-CE09-0002-01), a PhD grant (C. L.) from the Region Ile-de-France in the framework of DIM NanoK, PSL Research University (project IMRESOV), France BioImaging and the Equipex Morphoscope 2.

References

1 R. Y. Tsien, *Annu. Rev. Biochem.*, 1998, **67**, 509–544.
2 N. Shaner, P. Steinbach and R. Tsien, *Nat. Methods*, 2005, **2**, 905–909.
3 H. O'Hare, K. Johnsson and A. Gautier, *Curr. Opin. Struct. Biol.*, 2007, **17**, 488–494.
4 M. J. Hinner and K. Johnsson, *Curr. Opin. Biotechnol.*, 2010, **21**, 766–776.
5 K. M. Dean and A. E. Palmer, *Nat. Chem. Biol.*, 2014, **10**, 512–523.
6 A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel and K. Johnsson, *Nat. Biotechnol.*, 2003, **21**, 86–89.
7 A. Keppler, H. Pick, C. Arrivoli, H. Vogel and K. Johnsson, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 9955–9959.
8 A. Gautier, A. Juillerat, C. Heinis, I. R. Corrêa Jr, M. Kindermann, F. Beaufils and K. Johnsson, *Chem. Biol.*, 2008, **15**, 128–136.
9 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.
10 M. P. Bruchez, *Curr. Opin. Chem. Biol.*, 2015, **27**, 18–23.
11 L. Jullien and A. Gautier, *Methods Appl. Fluoresc.*, 2015, **3**, 042007.

12 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.
13 C. A. Telmer, R. Verma, H. Teng, S. Andreko, L. Law and M. P. Bruchez, *ACS Chem. Biol.*, 2015, **10**, 1239–1246.
14 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. Vriz, T. Le Saux, L. Jullien and A. Gautier, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 497–502.
15 N. C. Shaner, R. E. Campbell, P. A. Steinbach, B. N. G. Giepmans, A. E. Palmer and R. Y. Tsien, *Nat. Biotechnol.*, 2004, **22**, 1567–1572.
16 R. Weissleder and M. Nahrendorf, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 14424–14428.
17 G. Marriott, S. Mao, T. Sakata, J. Ran, D. K. Jackson, C. Petchprayoon, T. J. Gomez, E. Warp, O. Tulyathan, H. L. Aaron, E. Y. Isacoff and Y. Yan, *Proc.Natl. Acad. Sci. U. S. A.*, 2008, **105**, 17789–17794.
18 J. Querard, T.-Z. Markus, M.-A. Plamont, C. Gauron, P. Wang, A. Espagne, M. Volovitch, S. Vriz, V. Croquette, A. Gautier, T. Le Saux and L. Jullien, *Angew. Chem., Int. Ed.*, 2015, **54**, 2633–2637.
19 J. Querard, T. Le Saux, A. Gautier, D. Alcor, V. Croquette, A. Lemarchand, C. Gosse and L. Jullien, *ChemPhysChem*, 2016, **17**, 1396–1413.
20 G. W. Fisher, M. H. Fuhrman, S. A. Adler, C. Szent-Gyorgyi, A. S. Waggoner and J. W. Jarvik, *J. Biomol. Screening*, 2014, **19**, 1220–1226.
21 C. P. Pratt, J. He, Y. Wang, A. L. Barth and M. P. Bruchez, *Bioconjugate Chem.*, 2015, **26**, 1963–1971.

SUPPORTING INFORMATION

Dynamic multi-color protein labeling in living cells

Chenge Li[a,b], Marie-Aude Plamont[a,b], Hanna L. Sladitschek[c], Vanessa Rodrigues[a,b], Isabelle Aujard[a,b], Pierre Neveu[c], Thomas Le Saux[a,b], Ludovic Jullien[a,b,*] & Arnaud Gautier[a,b,*]

[a] École Normale Supérieure, PSL Research University, UPMC Univ Paris 06, CNRS, Département de Chimie, PASTEUR, 24 rue Lhomond, 75005 Paris, France

[b] Sorbonne Universités, UPMC Univ Paris 06, ENS, CNRS, PASTEUR, 75005 Paris, France

[c] Cell Biology and Biophysics Unit, European Molecular Biology Laboratory, Meyerhofstr. 1, D-69117 Heidelberg

* Correspondence should be addressed to: ludovic.jullien@ens.fr and arnaud.gautier@ens.fr

LEGENDS OF SUPPLEMENTARY MOVIES

Movie S1. Dynamic color switching. Confocal time-lapse of a live HeLa cell expressing H2B-FAST labeled with 5 μM of HBR-3,5DOM upon replacement of the medium with solution containing 5 μM of HMBR (Single excitation at 488 nm; HMBR emission channel 493-538 nm; HBR-3,5-DOM emission channel 649-797 nm). HMBR was added at t = 0 s. Cells were grown in a minifluidic channel enabling easy solution replacement. See also **Figure 3C**.

Movie S2. Dynamic color switching in a spectrally crowded environment. Confocal time-lapse of live HeLa cells co-expressing lyn-EGFP (membrane), MTS-mCherry (mitochondria) and H2B-FAST (nucleus) upon fluorogen exchange (Green channel Ex/Em 488/493-575 nm; Red channel Ex/Em 543/578-797 nm). Cells were initially stained with 5 μM HBR-3,5DOM. Fluorogen exchange was induced by addition of an excess of HMBR at t = 0 s. The final concentrations of HBR-3,5DOM and HMBR were respectively 0.83 μM and 4.2 μM. See also **Figure S4**.

SUPPLEMENTARY FIGURES

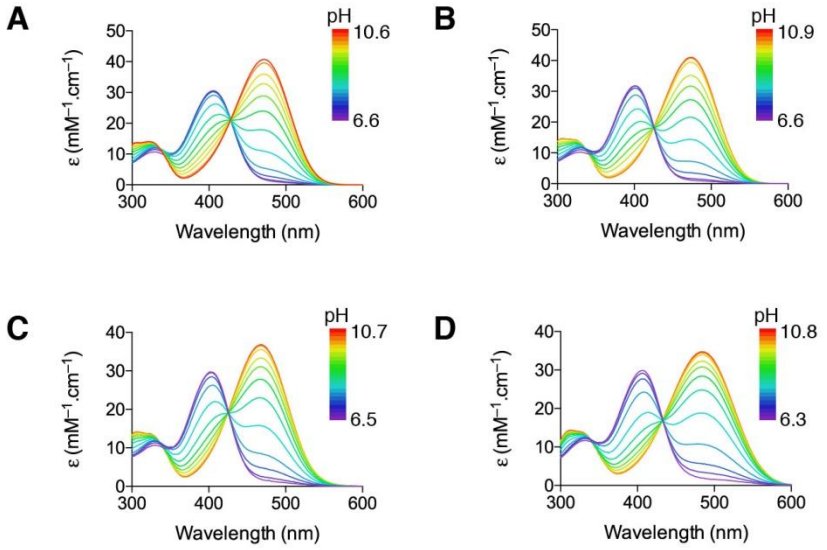


Figure S1. Absorption spectra of HBR-2,5DM (A), HBR-3,5DM (B), HBR-3OM (C) and HBR-3,5DOM (D) in solution in function of pH. The spectra were recorded in 0.04 M Britton–Robinson buffer (0.1 M ionic strength) at 25°C.

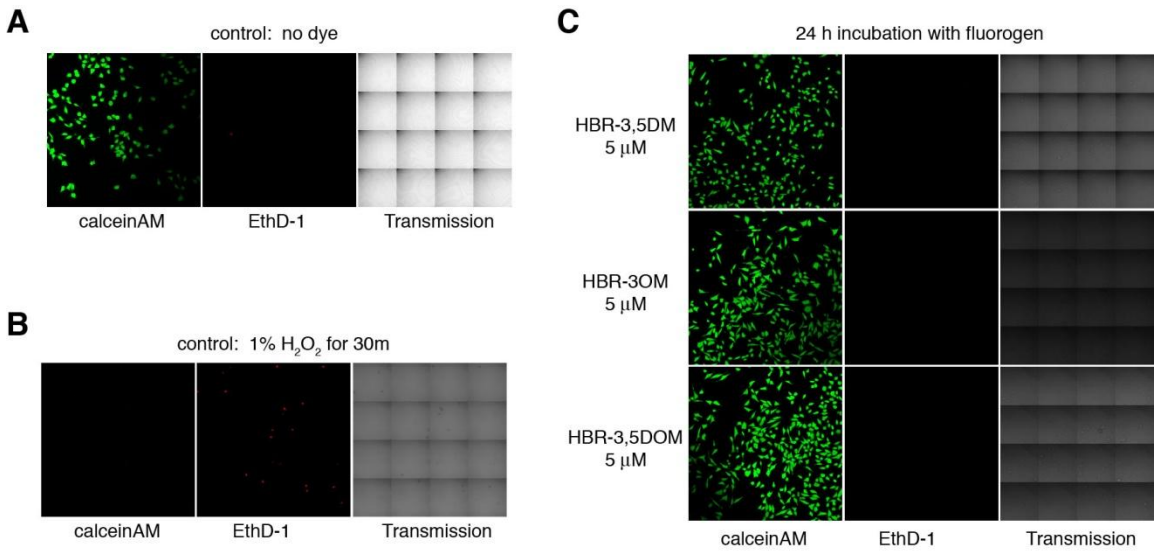


Figure S2. Viability assay of HeLa cells incubated for 24 h with solutions of HBR-3,5DM, HBR-3OM and HBR-3,5DOM at 5 μM. Cell viability was tested by using calceinAM and EthD1 (LIVE/DEAD® viability/cytotoxicity assay kit). CalceinAM is a cell-permeant profluorophore cleaved by intracellular esterases releasing the green fluorescent polyanionic calcein in live cells. EthD1 (Ethidium homodimer 1) is a non cell-permeant nucleic acid red fluorescent stain that enters only cells with damaged membranes and undergoes a fluorescence enhancement upon binding to nucleic acids, thereby producing a bright red fluorescence in dead cells. Control experiments with HeLa cells non-incubated with dye (**A**) or incubated for 30 min with 1% hydrogen peroxyde (**B**) are shown. Cell fluorescence was evaluated by confocal microscopy. The experiment in (**C**) shows that HBR-3,5DM, HBR-3OM and HBR-3,5DOM are non-toxic for HeLa cells at the concentrations used for imaging.

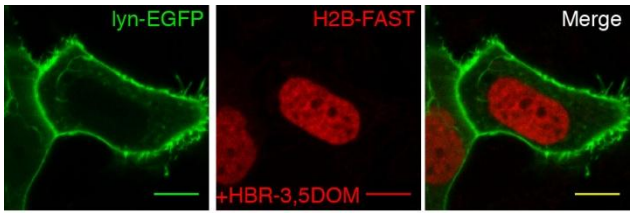


Figure S3. Dual-color imaging with a single excitation. Confocal micrographs of live HeLa cells co-expressing lyn-EGFP and H2B-FAST labeled with 5 μM HBR-3,5DOM upon single excitation at 488 nm (Green emission channel 493-538 nm; Red emission channel 600-797 nm). Scale bars 10 μm.

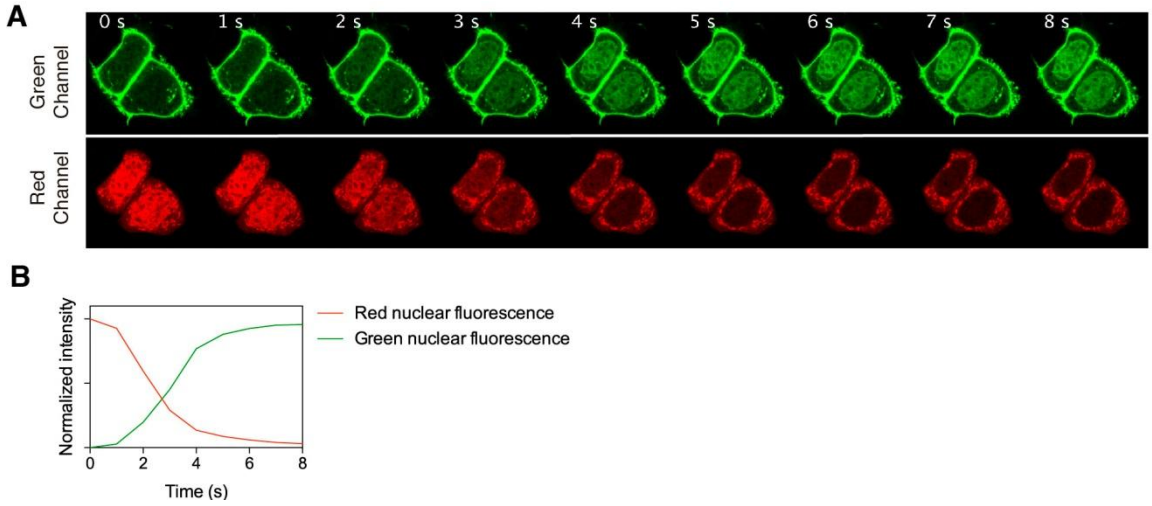


Figure S4. Dynamic color switching in spectrally crowded environment. (A) Time series of live HeLa cells co-expressing lyn-EGFP (membrane), MTS-mCherry (mitochondria) and H2B-FAST (nucleus) upon fluorogen exchange (Green channel Ex/Em 488/493-575 nm; Red channel Ex/Em 543/578-797 nm). Cells were initially stained with 5 μM HBR-3,5DOM. Fluorogen exchange was induced by addition of an excess of HMBR at t = 0 s. The final concentrations of HBR-3,5DOM and HMBR were respectively 0.83 μM and 4.2 μM. See also **Movie S2**. **(B)** Temporal evolution of the green and red nuclear fluorescence intensities.

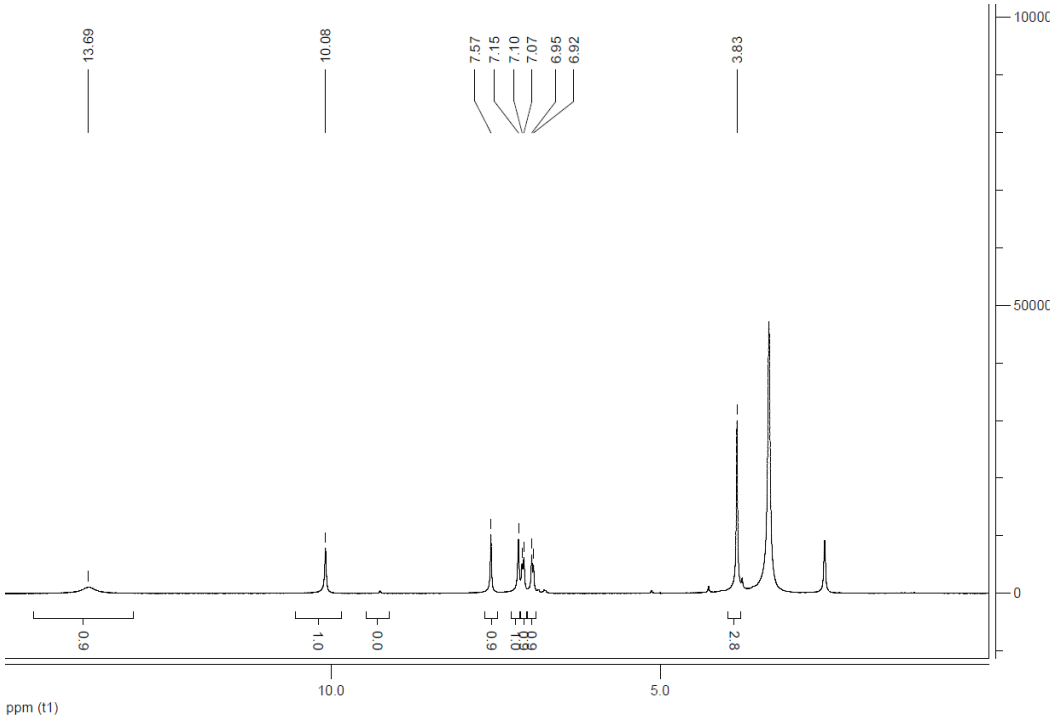
SUPPLEMENTARY TABLES

Table S1. Physico-chemical properties of HMBR, HBR-3,5DM, HBR-2,5DM, HBR-3OM, HBR-3,5DOM in aqueous solutions. Abbreviations are as follows : pK_A, acidity constant ; $\lambda_{abs,neutral}$, wavelength of maximal absorption of the protonated state; $\lambda_{abs,anionic}$, wavelength of maximal absorption of the anionic state.

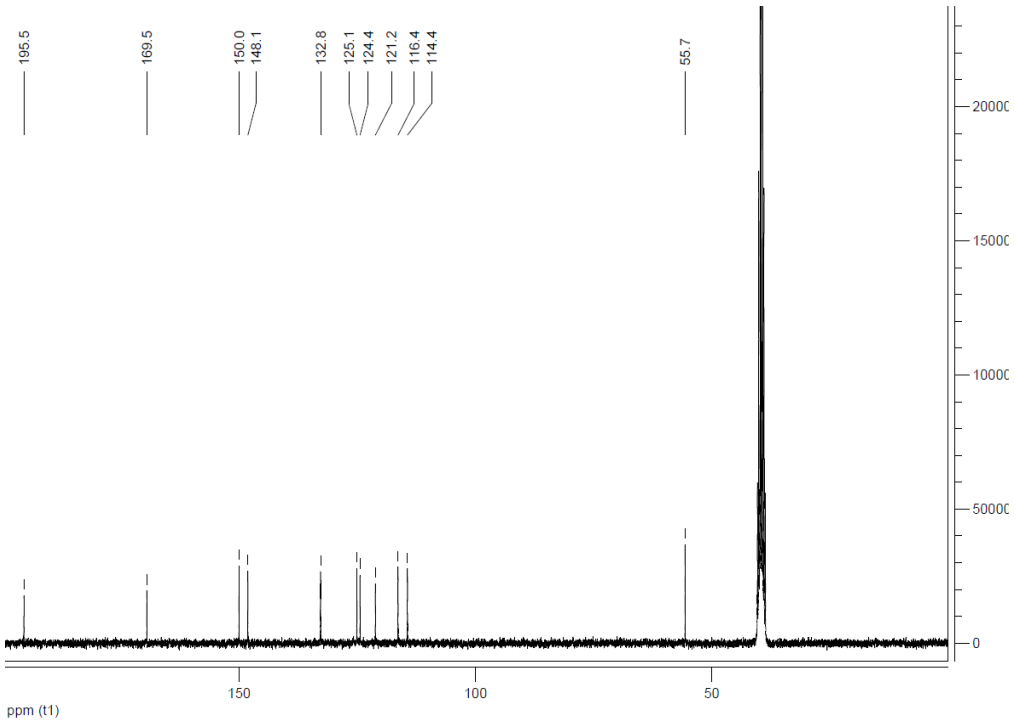
Fluorogen	pK_A	$\lambda_{abs,neutral}$ (nm)	$\lambda_{abs,anionic}$ (nm)
HMBR	8.7	401	461
HBR-3,5DM	8.7	401	473
HBR-2,5DM	8.7	406	472
HBR-3OM	8.3	403	468
HBR-3,5DOM	8.3	407	484

(Z)-5-(4-hydroxy-3-methoxybenzylidene)-2-thioxo-1,3-thiazolidin-4-one (HBR-3OM)

^{1}H

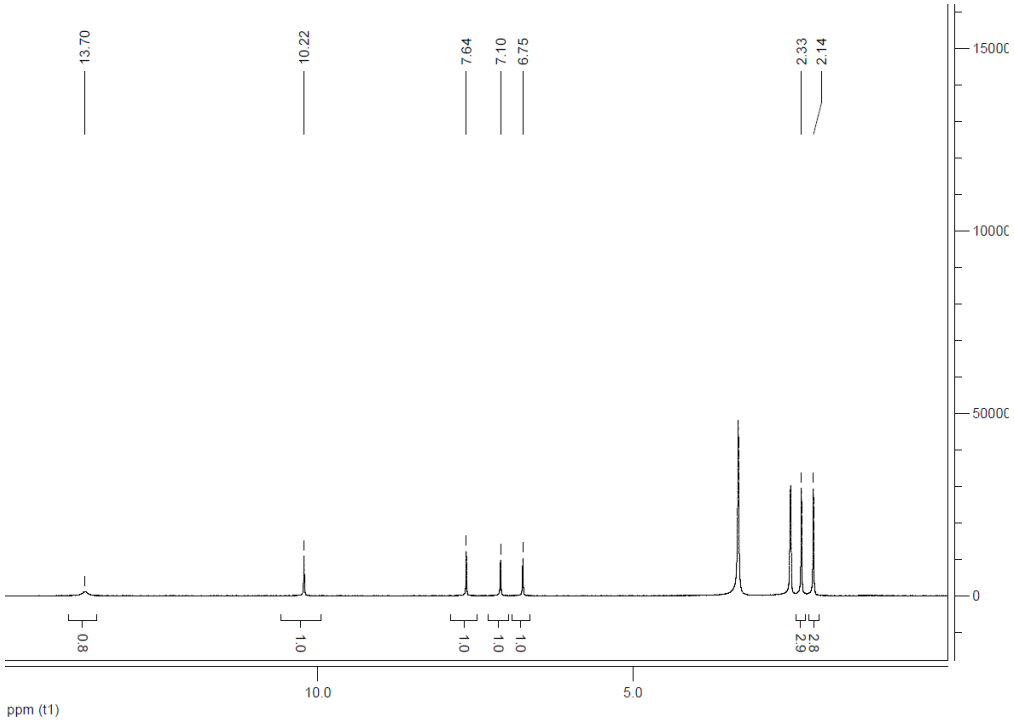


^{13}C

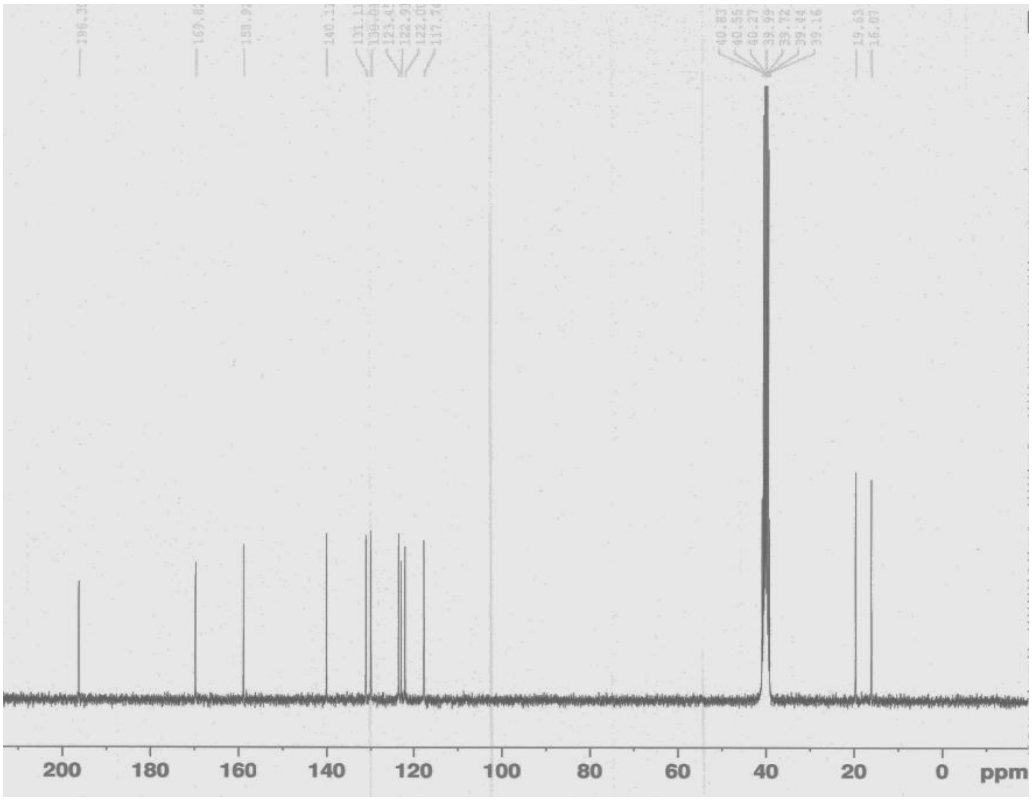


(Z)-5-(4-hydroxy-2, 5-dimethylbenzylidene)-2-thioxo-1,3-thiazolidin-4-one (HBR-2,5DM)

^{1}H

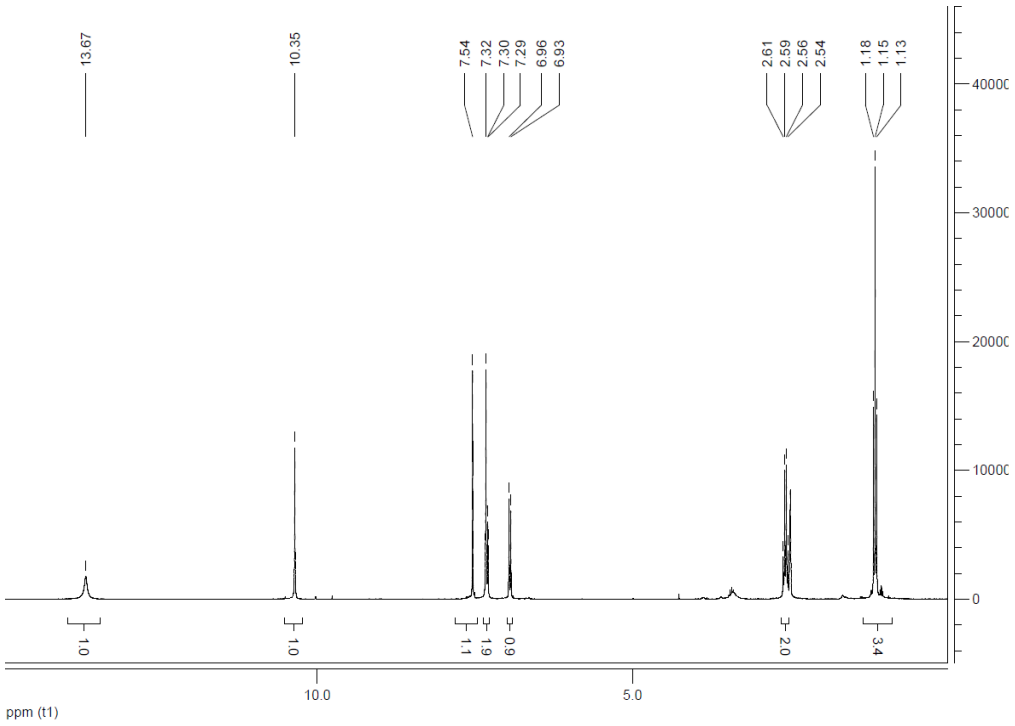


^{13}C

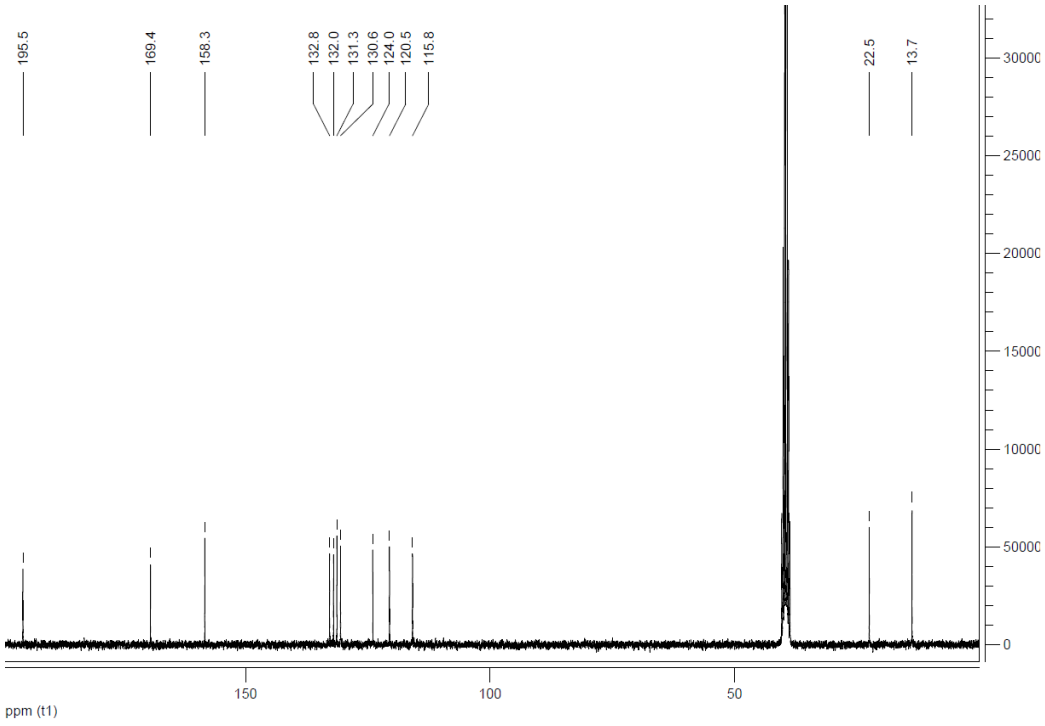


(Z)-5-(3-ethyl-4-hydroxybenzylidene)-2-thioxothiazolidin-4-one (HBR-3E)

^{1}H

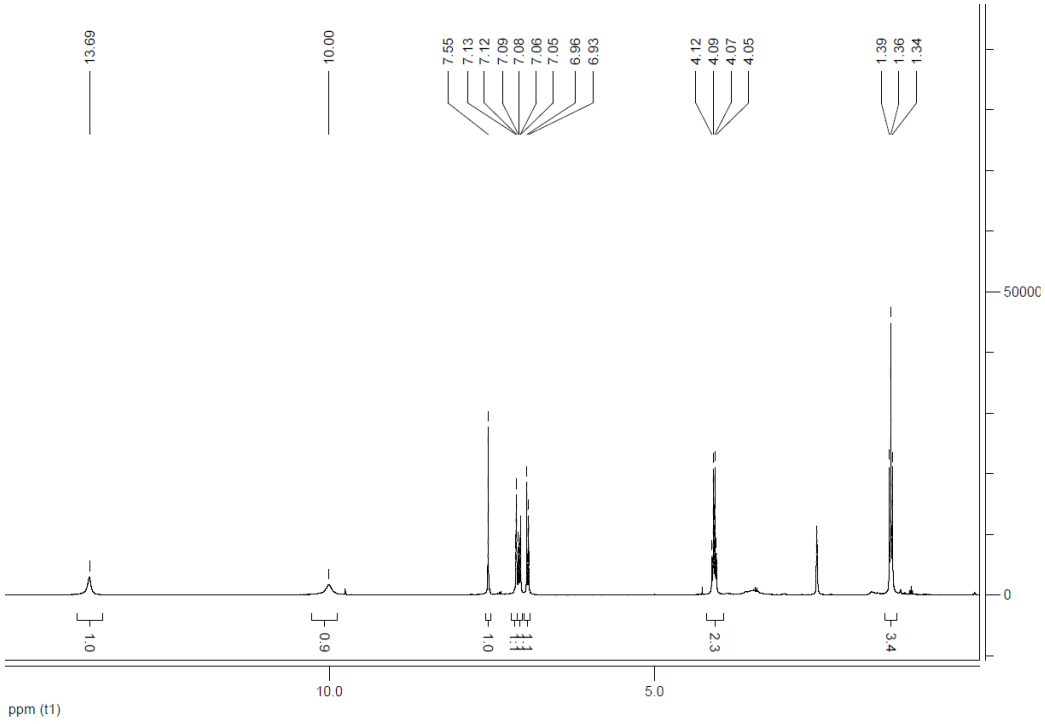


^{13}C

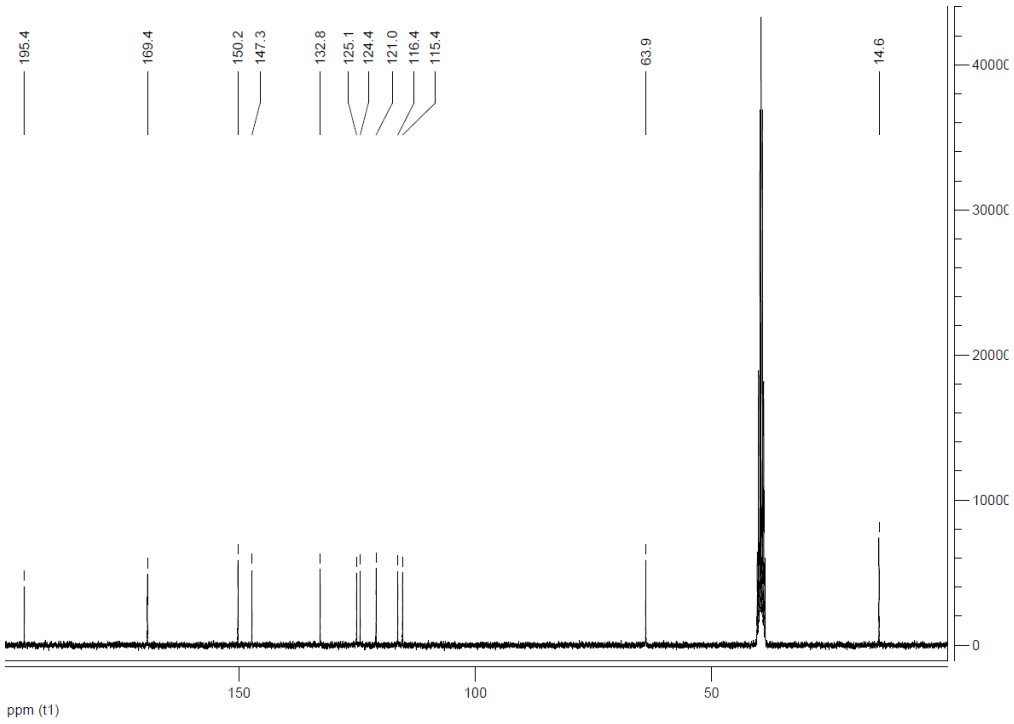


(Z)-5-(3-ethoxy-4-hydroxybenzylidene)-2-thioxothiazolidin-4-one (HBR-3OE)

^{1}H

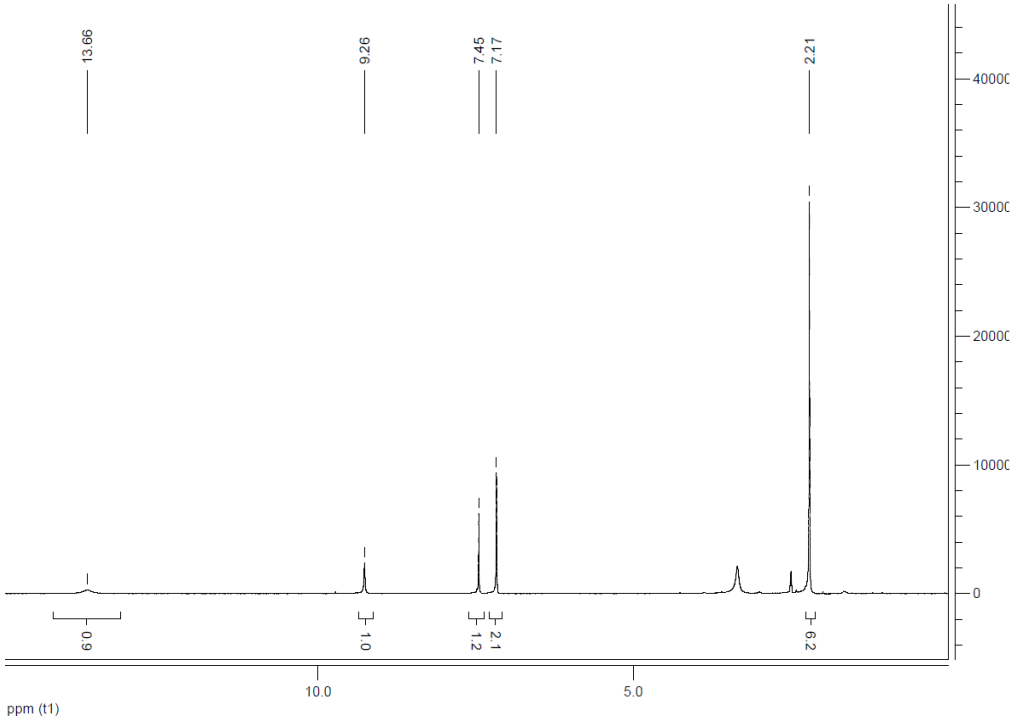


^{13}C

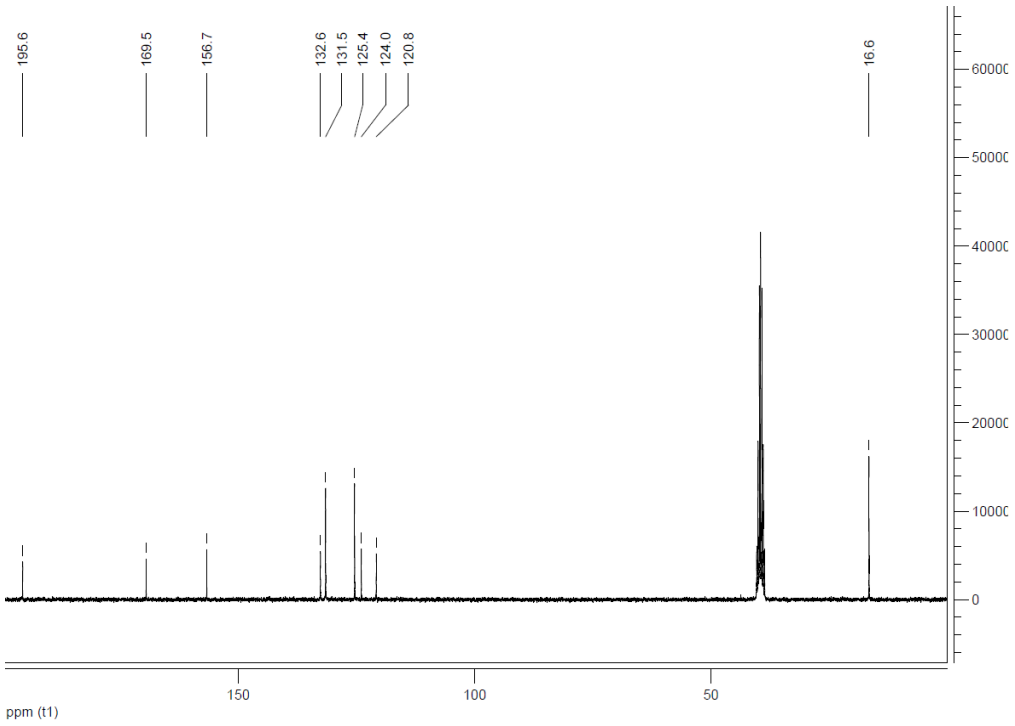


(Z)-5-(4-hydroxy-3,5-dimethylbenzylidene)-2-thioxothiazolidin-4-one (HBR-3,5DM)

^{1}H

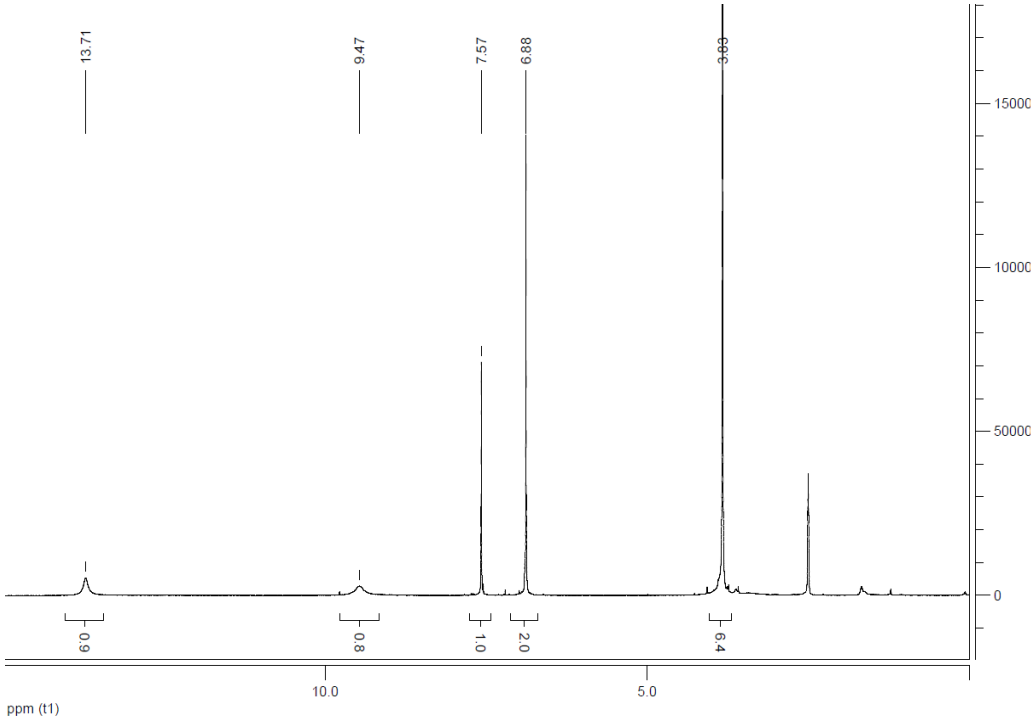


^{13}C

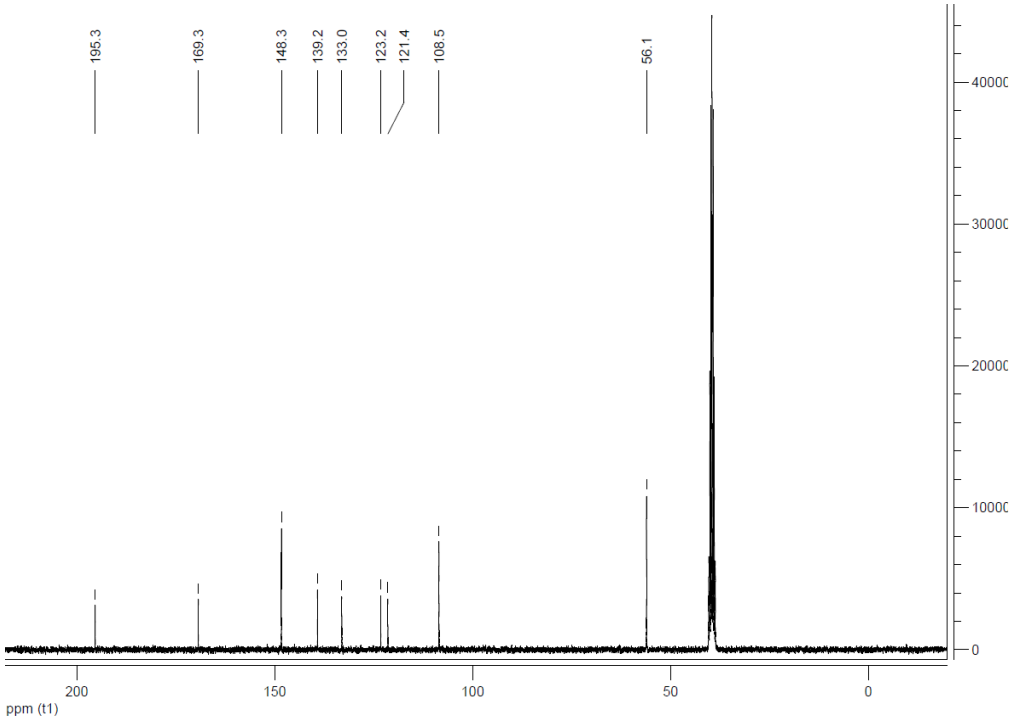


(Z)-5-(4-hydroxy-3,5-dimethoxybenzylidene)-2-thioxothiazolidin-4-one (HBR-3,5DOM)

^{1}H



^{13}C



Chapter III Development of red emitting FAST

Red-emitting fluorescent reporters have always been attractive to biologists for several reasons: there is almost no cellular autofluorescence in that spectral region; longer wavelength excitation reduces phototoxicity for biological organisms; longer wavelength ensures effective light penetration for deep imaging of tissues.

For developing red-emitting FAST, we have engineered a collection of HMBR analogues with various substituents on the phenol ring, which provides us with novel orange or red emitting FAST:fluorogen complexes with better fluorescence properties (See Chapter II). The red fluorogen HBR-3,5DOM is currently under use in the lab to evolve orthogonal FAST variants that will enable multi-protein labeling in live cells. To further expand the spectral properties of FAST systems to the red edge of the visible spectrum, we decided to engineer red-emitting fluorogens (Section III-1, Article 2) and to evolve new protein tags binding them (Section III-2).

III-1 Molecular engineering of red emitting fluorogens

This section is based on the following article:

Design and characterization of red fluorogenic push-pull chromophores holding great potential for biosensing and imaging.
Organic and Biomolecular Chemistry 14, 9253–9261 (2016).
C. Li, M.-A. Plamont, I. Aujard, T. Le Saux, L. Jullien & A. Gautier

III-1.1 Presentation of article 2

The fluorogens developed for FAST are based on a push-pull conjugated structure, which consist in an aromatic electron-donating part and a heterocyclic electron-withdrawing terminal group interacting through a conjugated system. For push-pull fluorophores, effects of electron-donating and electron-withdrawing character and the extension of the conjugated system influence their spectral properties. We designed two series of putative fluorogens by combining two electron donors, the 4-hydroxyphenyl (series 1) or the 4-dimethylamino-

phenyl group (series 2) with various heterocyclic electron acceptors through longer conjugated system (with an additional double bond compared to previous fluorogens) (Figure 3.1) for expending furthest their spectral properties to red. Apart from engineering red fluorogens for FAST approach, we have demonstrated in this article the general idea and method to develop and characterize potential red fluorogens in this article.

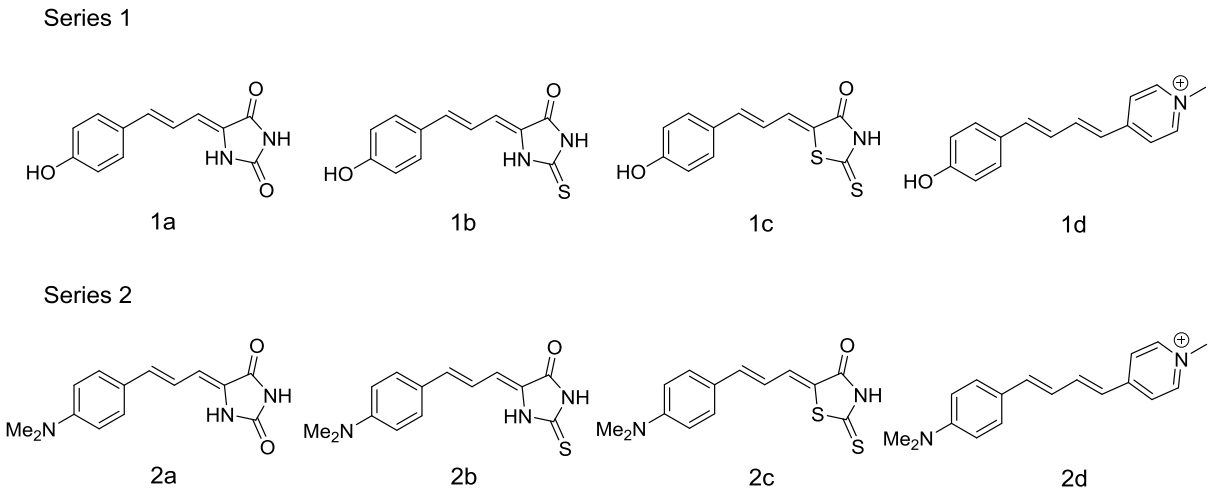


Figure 3.1 Structures of the two series of putative fluorogens.

New fluorogens were synthesized by condensation of the starting aldehydes and heterocycles in refluxing ethanol in presence of amine base (See Article 2–Figure 2). For series 2, the starting aldehyde is commercially available while the starting aldehyde for series 1 was obtained from a methoxymethyl (MOM)-protected 4-hydroxybenzaldehyde which underwent a Wittig oxyprenylation followed by acid hydrolysis (See Article 2–Figure 2). Next, we characterized their spectral properties (absorption and fluorescence emission) in aqueous buffer and in different conditions mimicking environment changes in protein cavity to see if they possess potential fluorogenic properties.

Chromophores of series 1 have two forms, protonated (phenol group) and deprotonated (phenolate group) form. Their pK_a are between 8.8-8.9 (See Article 2-Table 1 and Figure 3) ensuring that they stay almost fully protonated at pH 7.4. The phenolate displays higher electron-donating effect than a phenol group leading to a red-shift of the absorption. These acido-basic properties can be used to increase the fluorogenic response, since binding to a complementary cavity stabilizing the negatively charged phenolate could result in a large red-

shift in absorption at physiological pH, as previously demonstrated in the case of FAST and the HBR family[1] (See also Chapters I and II). We tested their absorption and emission properties at pH 6 (fully protonated) and pH 10 (fully deprotonated) to see the differences between these two forms. Deprotonation of phenol group induced 20-65 nm red shift in absorption and 55-95 nm red shift in emission, shifting the absorption from ultra violet or violet to violet or blue and the emission from yellow or orange to orange or red (See Article 2-Table 1 and Figure 4). The two forms displayed weak fluorescence (fluorescence quantum yield < 2%, see Article 2-Table 1). To test their fluorogenic response upon conformational locking, we used aqueous buffer at pH 10 complemented with 40% glycerol. Glycerol increases solution viscosity and thus restricts the motion of chromophores, mimicking environment changes in protein cavities. We observed about two-fold increase of fluorescence quantum yield (See Article 2-Table 1), which proved their capability of fluorescence enhancement upon conformational restriction. From our previous work[1], this behavior allowed us to anticipate larger fluorogenic responses upon binding within a tailored protein cavity. Chromophores of series 2 absorbed violet or blue light and fluoresced weakly in red (fluorescence quantum yield < 1%) in aqueous buffer at pH 7.4 (See Article 2-Table 2 and Figure 5). In viscous buffer (in presence of 40% glycerol), they also displayed about two-fold increase of fluorescence quantum yield (See Article 2-Table 2). Moreover, solvato(fluo)chromic properties tests revealed their negative solvatochromism properties (See Article 2-Figure 6), which could also be used to obtain absorption red shift upon entering low-polarity protein cavities. In summary, the two series of chromophores are all potential red fluorogens for FAST approach because they display fluorescence activation upon motion restriction and potential of red shift in absorption through binding-induced ionization of auxochromic group or negative solvatochromism. In particular, these fluorogens exhibit high Stoke shifts from 150 nm to 260 nm, opening great perspectives for multiplexed imaging with a single excitation source.

Before engineering protein scaffolds enabling to bind and activate these new fluorogens, we studied the behavior of these putative fluorogens in mammalian cells by fluorescence microscopy. Appropriate fluorogens are supposed to exhibit no toxicity for cells and barely fluoresce in culture medium and cells. Cytotoxicity test in HeLa cells allowed the identification of the concentration range ensuring no toxicity (See Article 2- Table 3). When imaging HeLa cells incubated with culture media containing 1 μM of dye by fluorescence microscopy, we observed two trends in function of the series used. Fluorogens of series 1

were invisible in cells as desired (See Article 2-Figure 7), while the fluorogens of series 2 fluorescently labeled specific organelles in cells, as verified by colocalization experiments with commercial organelle trackers (See Article 2-Figure 7 and Figure 8). Although this latter series is not suitable for the development of new FAST systems, it could be applied as specific organelle trackers.

Then we tested the binding abilities of the fluorogens in series 1 with FAST. Except for **1c** which displayed a little fluorescence activation, the other fluorogens did not show any detectable binding (by measurement of absorption or fluorescence) to FAST. The larger size of the new fluorogens compared to HMBR is very likely the main raison preventing the binding to FAST. Therefore, we considered engineering complementary protein tags able to accept the newly developed fluorogens.

III-1.2 Article 2: Design and characterization of red fluorogenic push-pull chromophores holding great potential for bioimaging and biosensing

Organic & Biomolecular Chemistry

PAPER

View Article Online
View Journal | View Issue

Cite this: *Org. Biomol. Chem.*, 2016, **14**, 9253

Design and characterization of red fluorogenic push–pull chromophores holding great potential for bioimaging and biosensing†

Chenge Li,[a,b] Marie-Aude Plamont,[a,b] Isabelle Aujard,[a,b] Thomas Le Saux,[a,b] Ludovic Jullien*[a,b] and Arnaud Gautier*[a,b]

Received 28th July 2016,
Accepted 2nd September 2016

DOI: 10.1039/c6ob01612j

www.rsc.org/obc

Fluorogenic chromophores have been used recently for fluorescence reporting and biosensing. Their ability to turn on upon specific interaction with a given target has been exploited in particular for the design of fluorogen-based reporters enabling biomolecule labeling and imaging. In this paper, we report the development and exhaustive characterization of a new family of red fluorogenic push–pull chromophores, holding great potential for the development of fluorogen-based reporters or intracellular fluorogenic markers. The proposed methodology is generic and should find general applicability in the discovery of new fluorogenic dyes suitable for the design of fluorogen-based reporters and biosensors.

Introduction

Chemical Biology aims at discovering powerful molecules to address biological problems that would otherwise remain elusive. The development of chemical biology tools to observe or manipulate biological processes often needs the pairing of tailor-made synthetic molecules with engineered genetically encodable biomolecules in order to obtain selective and controlled output functions. The developments of advanced biomolecular labeling tools—composed of genetically targetable tags engineered to specifically interact with synthetic functional ligands—nicely illustrate the power of hybridizing synthetic and genetically encodable molecules.[1–3]

In the context of fluorescence labeling, turn-on systems exploiting fluorogenic compounds (so-called fluorogens) have recently gained popularity. The use of fluorogenic dyes—that light up when they interact with their complementary target—easily provides high contrast without the need for washing dye excess.[4,5] In addition, the requirement for fluorogen supply opens great opportunities for on-demand applications where the fluorescence is desired only at a given time or at a given density.

One of the current limitations in the development of efficient fluorogen-based reporters is the small number of fluorogenic compounds eligible for their design.[5] The exploration and discovery of new families of fluorogenic dyes is thus indispensable. In this context, several spectroscopic changes have been envisioned to generate a fluorogenic effect upon biomolecular interactions.[4,5] Given that brightness is a convolution of fluorescence quantum yield and absorption coefficient, any changes of one of these two parameters upon target binding can give a fluorogenic behavior. The fluorescence quantum yield of some dyes can increase upon motion restriction. These dyes do not fluoresce when free in solution because they deexcite from the excited state mainly through non-radiative processes involving internal molecular motions; however, when they are immobilized inside a complementary cavity, these non-radiative processes are impaired and fluorescence becomes the main relaxation pathway, which gives rise to very bright complexes. Examples of such fluorogens are (i) DAPI or Hoechst dyes used to stain DNA, (ii) thiazole orange and malachite green used to label proteins fused to fluorogen-activating proteins (FAPs),[6–9] (iii) the chromophore of the green fluorescent protein (GFP) and its analogs used for the development of RNA mimics of GFP,[10–14] or (iv) merocyanines used to label CRABPII.[15] Increase in fluorescence quantum yield can also be induced by polarity changes as in the case of 7-dimethylaminocoumarin derivatives that are specifically activated within non-polar environments. This behavior has been used for instance to design fluorogenic substrates for the specific covalent labeling of PYP-tag.[16] A second way to obtain a fluorogenic effect is to play with the absorption coefficient. Far-red fluorescent silicon-rhodamines display for instance fluorogenic behaviour because of their polarity-dependent interconversion between a spiro closed form with low absorbance at 650 nm and a zwitterionnic opened form strongly absorbing at 650 nm.[17] Switching from the non-absorbing to

[a]*École Normale Supérieure, PSL Research University, UPMC Univ Paris 06, CNRS, Département de Chimie, PASTEUR, 24 rue Lhomond, 75005 Paris, France. E-mail: ludovic.jullien@ens.fr, arnaud.gautier@ens.fr*
[b]*Sorbonne Universités, UPMC Univ Paris 06, ENS, CNRS, PASTEUR, 75005 Paris, France*

†Electronic supplementary information (ESI) available. See DOI: 10.1039/c6ob01612j

the absorbing form activates thus fluorescence. This singular behavior was used to develop fluorogenic probes for the covalent labeling of SNAP-tag/CLIP-tag/Halo-tag[17] and for fluorescence imaging of actin and tubulin in living cells.[18] A change in absorption coefficient at a given wavelength can also result from a chromophore absorption shift induced by a change of environment. Fluorogenic labeling can hence be obtained using solvatochromic dyes displaying large changes in maximal absorption wavelength upon binding hydrophobic biomolecular cavities. Alternatively, absorption shift can also result from ionization of an auxochromic group. This spectroscopic change has been recently coupled with binding-induced fluorescence quantum yield increase to design a highly specific fluorogen-based reporter called Y-FAST (Yellow Fluorescence-Activating and Absorption-Shifting Tag),[19] which relies on the reversible non-covalent binding of 4-hydroxybenzylidene rhodanine (HBR) or 4-hydroxy-3-methyl-benzylidene rhodanine (HMBR). HBR and HMBR are non-fluorescent in solution but strongly fluoresce yellow light when immobilized within Y-FAST, a small protein of 14 kDa engineered from the photoactive yellow protein (PYP); in addition, HBR and HMBR are protonated in solution because of a high pK_a, but are deprotonated within Y-FAST, leading to a 80 nm red-shift in absorption (because of the stronger electron-donation of the phenolate). This second spectroscopic modification enables to further increase the fluorogenic effect by choosing the excitation wavelength so that free or nonspecifically bound fluorogen does not contribute to the fluorescence signal, ensuring hence high imaging contrast.

In order to extend the collection of Fluorescence-Activating and absorption-Shifting Tags (FAST) and obtain a palette of various colours, we sought for red emitting fluorogenic chromophores (Fig. 1a), which could display both fluorescence quantum yield increase and absorption red shift upon binding well-design complementary tags. We first targeted dyes with high potential of presenting fluorescence quantum yield increase upon motion restriction (Fig. 1c). Second, to obtain binding-induced absorption red-shift, we sought for (i) dyes bearing an auxochromic group ionizable upon binding (altering thus fluorogen's photophysics), or (ii) negative solvatochromic dyes whose absorption could red-shift within hydrophobic protein cavities (Fig. 1c). Given the successful use of push–pull fluorogenic dyes in the design of fluorogen-based reporters,[10,15,19] we studied two series of fluorogen candidates composed of an electron-donating aromatic group (either 4-hydroxyphenyl or 4-dimethylamino-phenyl) conjugated *via* two double bonds to heterocyclic acceptors of various electron-withdrawing strengths (Fig. 1b). The extended π-electrons delocalization pathway allowed pushing the fluorescence of these donor–acceptor conjugated systems to the red edge of the visible spectrum. We evidenced their potential as fluorogens by a complete characterization of their photophysical properties (absorption and fluorescence emission) in conditions mimicking environment changes that might occur within biomolecule cavities such as immobilization, ionization or polarity drop. We further assayed their toxicity in mammalian cells, and evaluated their potential usability as fluorogenic probes by fluorescence microscopy.

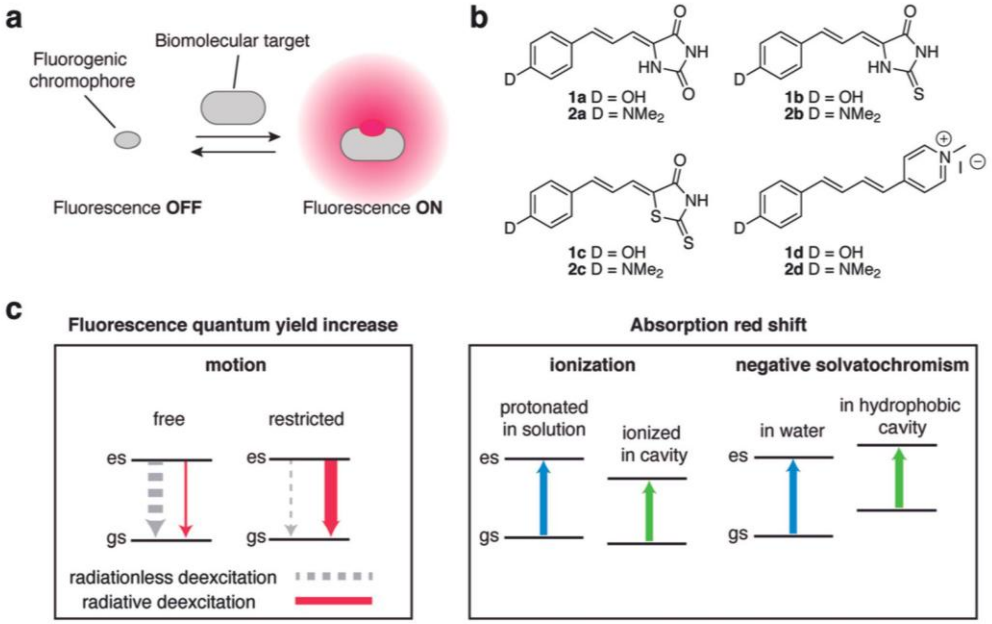


Fig. 1 Red fluorogenic chromophores for reporting and biosensing. (a) General fluorogen activation principle. (b) Chemical structures of the red fluorogenic compounds **1a–d** and **2a–d** developed in this study. (c) Highly specific fluorogen activation upon binding in a complementary biomolecular cavity can result from the combination of various spectroscopic changes such as fluorescence quantum yield increase (through motion restriction) and absorption red shift (through specific ionization or polarity change).

View Article Online
Organic & Biomolecular Chemistry Paper

Results

Design and synthesis

We recently reported the fluorogenic properties of HBR (4-hydroxybenzylidene rhodanine).[19] HBR is composed of an electron-donating phenol moiety conjugated to an electron-withdrawing rhodanine head group. When deprotonated, HBR absorbs at 450 nm and emits around 545 nm.[19] This push–pull compound is related to the fluorogenic 4-hydroxybenzylidene-5-imidazolinone (HBI) found in the green fluorescent protein (GFP),[20] in which the electron withdrawing head is an imidazolinone instead of a rhodanine. HBI is known to fluoresce only when immobilized in the GFP beta-barrel; the absence of fluorescence observed for the bare HBI in solution correlates with accessible radiationless decay channels along the *cis–trans* photoisomerization path.

To push the emission wavelength further to the red, we shifted absorption by extending the π-electron conjugation with an additional double bond between the donor and the acceptor. As the Stokes shift depends on the change of the dipole moment between the ground state and excited states, we also varied both the donor and the acceptor in order to extract structure-relationship trends. As donor, we examined (i) the 4-hydroxyphenyl, which can exist in a protonated and deprotonated state, the latter being more π-electron donor, and (ii) the 4-dimethylaminophenyl, a strong π-electron donor (Fig. 1b). As acceptors, we tested the hydantoin (or 2,4-imidazolidinedione), the thiohydantoin (or 2-thioxo-4-imidazolidinone), the rhodanine (or 2-thioxo-4-thiazolidinone) and the methyl-pyridinium, in order to explore a palette of electron-withdrawing strengths (Fig. 1b).

The push–pull compounds **1a–d** and **2a–d** were obtained by condensation of the aldehydes **1** and **2** with either hydantoin **a**, thiohydantoin **b**, rhodanine **c** or methyl-pyridinium **d** in refluxing ethanol in presence of an appropriate amine base (Fig. 2a, see Materials and methods for details). The aldehyde **2** was commercially available. The aldehyde **1**, on the other hand, was prepared by homologation of the 4-hydroxy-benzaldehyde (Fig. 2b). After efficient protection of the hydroxyl group with a methoxymethyl (MOM) acetal group, the protected 4-hydroxybenzaldehyde underwent a Wittig olefination with tributyl(1,3-dioxolan-2-ylmethyl) phosphonium bromide in presence of sodium hydride (NaH) in tetrahydrofuran (THF). Final acidolysis of the Wittig product provided the aldehyde **1** with high yield.

Absorption and emission properties

Given that a phenolate is a better π-electron donor than a phenol, the protonated and deprotonated states of the compounds **1a–d** exhibit different absorption properties (Fig. 3 and Table 1). In aqueous solutions, the protonated states of **1b–d** absorb mainly violet light, while their deprotonated states absorb blue light (Fig. 3 and Table 1). The absorption of **1a** is blue-shifted compared to **1b–d** (Fig. 3 and Table 1). **1a** absorbs mainly ultraviolet light when protonated and violet

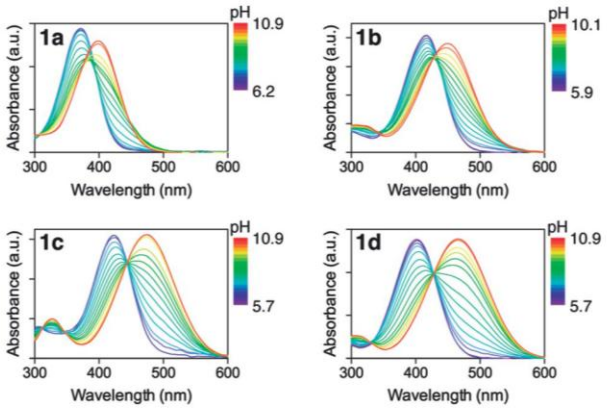


Fig. 3 Absorption spectra of **1a–d** at various pH. The spectra were recorded in 0.01 M Britton–Robinson buffer[23] at 25 °C. The evolution of absorption in function of pH allowed determining the acidity constant.

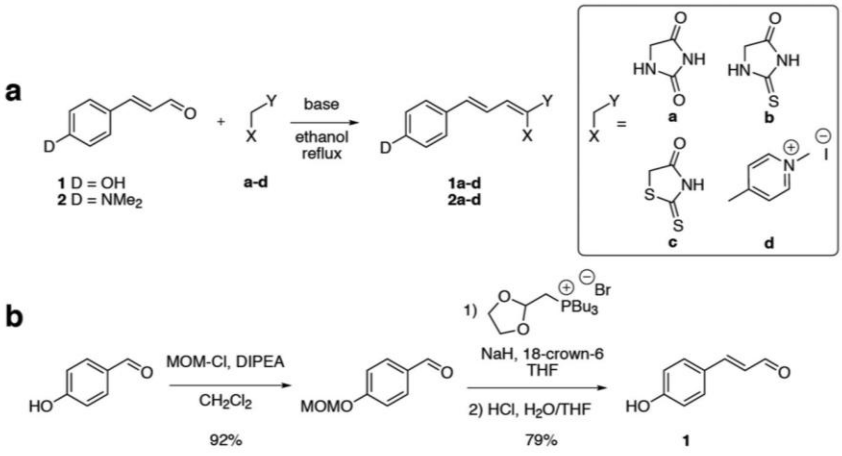


Fig. 2 Chemical synthesis. (a) General strategy for the syntheses of **1a–d** and **2a–d**. (b) Synthesis of aldehyde **1**.

Table 1 Physicochemical properties of compounds **1a–d** at pH 6 and pH 10 (±40% glycerol). Abbreviations are as follows: pK_a, acidity constant; λ_{abs}, wavelength of maximal absorption; λ_{em}, wavelength of maximal emission; ε, molar absorption coefficient at λ_{abs}; ϕ, fluorescence quantum yield

		pH 6		pH 10		pH 10 + 40% glycerol	
	pK_a	λ_{abs} [ε] [nm][mM^{-1} cm^{-1}]	λ_{em} [ϕ] [nm][%]	λ_{abs} [ε] [nm][mM^{-1} cm^{-1}]	λ_{em} [ϕ] [nm][%]	λ_{abs} [ε] [nm][mM^{-1} cm^{-1}]	λ_{em} [ϕ] [nm][%]
1a	8.9	370 [40]	520 [0.4]	394 [32]	590 [0.5]	396 [31]	595 [1.0]
1b	8.9	415 [33]	580 [0.1]	450 [30]	635 [0.5]	455 [29]	635 [1.2]
1c	8.9	425 [36]	595 [0.1]	471 [33]	655 [1.1]	481 [34]	659 [2.4]
1d	8.8	400 [35]	595 [1.8]	465 [30]	690 [2.0]	481 [29]	695 [5.3]

light when deprotonated. Deprotonation of **1a**, **1b**, **1c** and **1d** induces a red shift of about 20, 35, 55 and 65 nm, respectively. Spectrophotometric titrations showed that compounds **1a–d** exhibit low acidity constant (pK_a ~ 8.8–8.9) (Fig. 3 and Table 1). Compounds **1a–d** are thus protonated at physiological pHs allowing to envision inducing an absorption red-shift —which can participate to fluorogen activation—by specific ionization within an engineered protein tag possessing a network of residues able to stabilized the negative charge (*vide supra*). Our photophysical study further showed that protonated **1a–d** displayed high Stokes shift (150–200 nm) and fluoresced, albeit very weakly, between 520 and 590 nm (Table 1). Deprotonated **1a–d** also exhibited high Stokes shift (180–230 nm) and a palette of emission wavelengths in the orange/red, going from 590 nm (for **1a**) to 690 nm (for **1d**) (Table 1 and Fig. 4). The high Stokes shifts of **1a–d** open great prospects for multiplexed fluorescence observation using a single excitation wavelength.

Photophysical analysis of the series **2a–d** in aqueous solution at pH 7.4 showed that compounds **2a–d** absorb blue to violet light and fluoresced, albeit very weakly, in the red, with emission wavelengths going from 630 nm to 710 nm (Table 2 and Fig. 5). Moreover, **2a–d** exhibited ultra high Stokes shifts from 210 to 260 nm. In order to see whether binding within biomolecules could change the absorption and emission of **2a–d** because of polarity drop, we examined the effect of

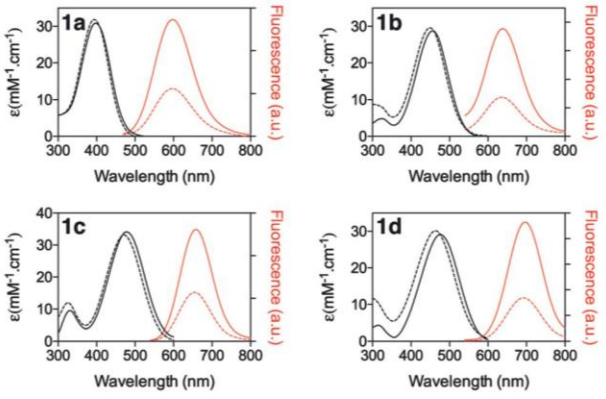


Fig. 4 Absorption (black line) and emission (red line) spectra of **1a–d** at pH 10 in presence (solid line) or absence (dotted line) of 40% glycerol. The spectra were recorded in 0.01 M Britton–Robinson buffer[23] at 25 °C.

Table 2 Photophysical properties of compounds **2a–d** at pH 7.4 (±40% glycerol). Abbreviations are as follows: λ_{abs}, wavelength of maximal absorption; λ_{em}, wavelength of maximal emission; ε, molar absorption coefficient at λ_{abs}; ϕ, fluorescence quantum yield

	pH 7.4		pH 7.4 + 40% glycerol	
	λ_{abs} [ε] [nm][mM^{-1} cm^{-1}]	λ_{em} [ϕ] [nm][%]	λ_{abs} [ε] [nm][mM^{-1} cm^{-1}]	λ_{em} [ϕ] [nm][%]
2a	401 [21]	631 [0.9]	415 [19]	630 [1.7]
2b	461 [32]	667 [0.7]	474 [34]	669 [1.8]
2c	461 [29]	672 [1.0]	474 [33]	676 [2.5]
2d	447 [31]	709 [1.0]	470 [30]	709 [3.5]

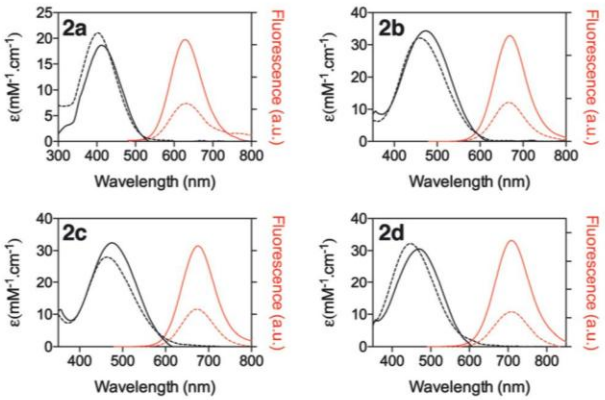


Fig. 5 Absorption (black line) and emission (red line) spectra of **2a–d** at pH 7.4 in presence (solid line) or absence (dotted line) of 40% glycerol. The spectra were recorded in 0.01 M Britton–Robinson buffer[23] at 25 °C.

solvent polarity on their absorption and emission spectra. We observed that a drop in polarity led to a significant red shift in absorption and an important blue shift in emission (Fig. 6). These observations enable to await strong chromatic modifications—both in absorption and emission—upon specific interaction with complementary protein cavities of lower polarity.

Next, we asked if the fluorescence properties of **1a–d** and **2a–d** were sensitive to motion restriction. As mentioned above, compounds **1a–d** and **2a–d** barely fluoresce in aqueous solutions (see also fluorescence quantum yields reported in Tables 1 and 2). However, these compounds belong to the same family as HBI and HBR/HMBR, which are known to fluoresce only when immobilized in GFP and Y-FAST[19] respectively. In

View Article Online
Organic & Biomolecular Chemistry Paper

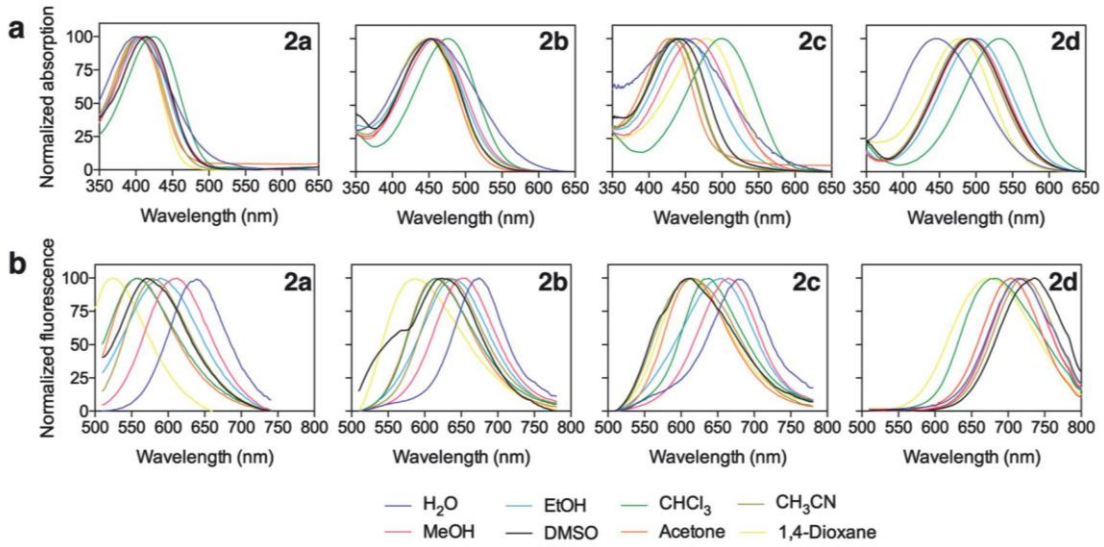


Fig. 6 Solvato(fluo)chromic properties of **2a–d**. Normalized absorption (a) and emission (b) spectra of **2a–d** in various solvents. The spectra were recorded at 25 °C.

the case of such fluorogens, the effect of motion restriction on fluorescence quantum yield can be evidenced easily by observing fluorescence enhancement in viscous solutions.[19] To evaluate whether **1a–d** and **2a–d** hold potential as fluorogens, we thus examined how the fluorescence quantum yield evolved upon solution viscosity increase. As viscous media, we used aqueous buffer complemented with 40% glycerol. This allowed us to obtain a buffer fourfold viscous than water.[21] The fluorescence quantum yields of compounds **1a–d** (deprotonated) and **2a–d** increased two- or three-fold in viscous medium (Tables 1 and 2, Fig. 4 and 5). These results are in agreement and comparable with what we previously reported for the fluorogenic HBR/HMBR,[19] enabling to anticipate large fluorescence exaltation upon immobilization within complementary protein cavities.

Table 3 Cellular toxicity of **1a–d** and **2a–d**. HeLa cells were treated with media containing the fluorogenic dyes at various concentrations. Cell viability was tested after 5 h and 24 h of incubation. The table provides the degree of toxicity. No cell death (–): dead cells <1%, low cell death (+): 1% <dead cells <20%, significant cell death (++): dead cells >50%

	Cell death after 5 h	Cell death after 24 h		
	10 μM	0.1 μM	1 μM	10 μM
1a	–	–	–	+
1b	–	–	–	+
1c	–	–	–	–
1d	–	–	–	–
2a	–	–	+	++
2b	–	–	+	++
2c	–	–	–	–
2d	–	–	–	–

Cellular evaluation

To be suitable for the design of fluorogen-based reporters, fluorogenic dyes must ideally exhibit no toxicity for cells and give no fluorescence background in cells in absence of their targets. We therefore examined the behaviour of **1a–d** and **2a–d** in mammalian cells, focusing on toxicity and non-specific fluorogen activation. The toxicity of **1a–d** and **2a–d** were evaluated in HeLa cells. Cells were treated with media containing the dyes and cell viability was determined after 5 h or 24 h (Table 3). We observed no cell death after 5 h of incubation with 10 μM of dye, no matter the dye used. However, we observed detectable cell death after 24 h of incubation with high concentrations (10 μM) of **2a** and **2b**, and, in a much lesser extent, of **1a** and **1b** (Table 3). These results showed that our fluorogenic compounds did not exhibit significant toxicity in the micromolar or submicromolar range usually used for labeling fluorogen-activating tags.

We next asked by fluorescence microscopy if **1a–d** and **2a–d** underwent non-specific fluorescence activation in living mammalian cells. We treated HeLa cells with media containing 1 μM dye. We observed two different trends. While compounds **1a–d** exhibited no fluorescence background (Fig. 7), compounds **2a–d** did fluorescently label intracellular organelles or structures. Compounds **2a**, **2b**, and **2c** specifically stained intracellular membranes (Fig. 7). Colocalization studies showed that the fluorescently labelled structures were very likely membranes of the endoplasmic reticulum (ER) (Fig. 8). Activation of **2a–c** can result from immobilization within membranes. Compound **2d**, on the other hand, fluorescently labeled mitochondria (Fig. 7 and 8), in accordance with previous studies reporting that pyridinium-containing compounds tend to accumulate in mitochondria because of their positive charge.[22] Our results demonstrated that only **1a–d**

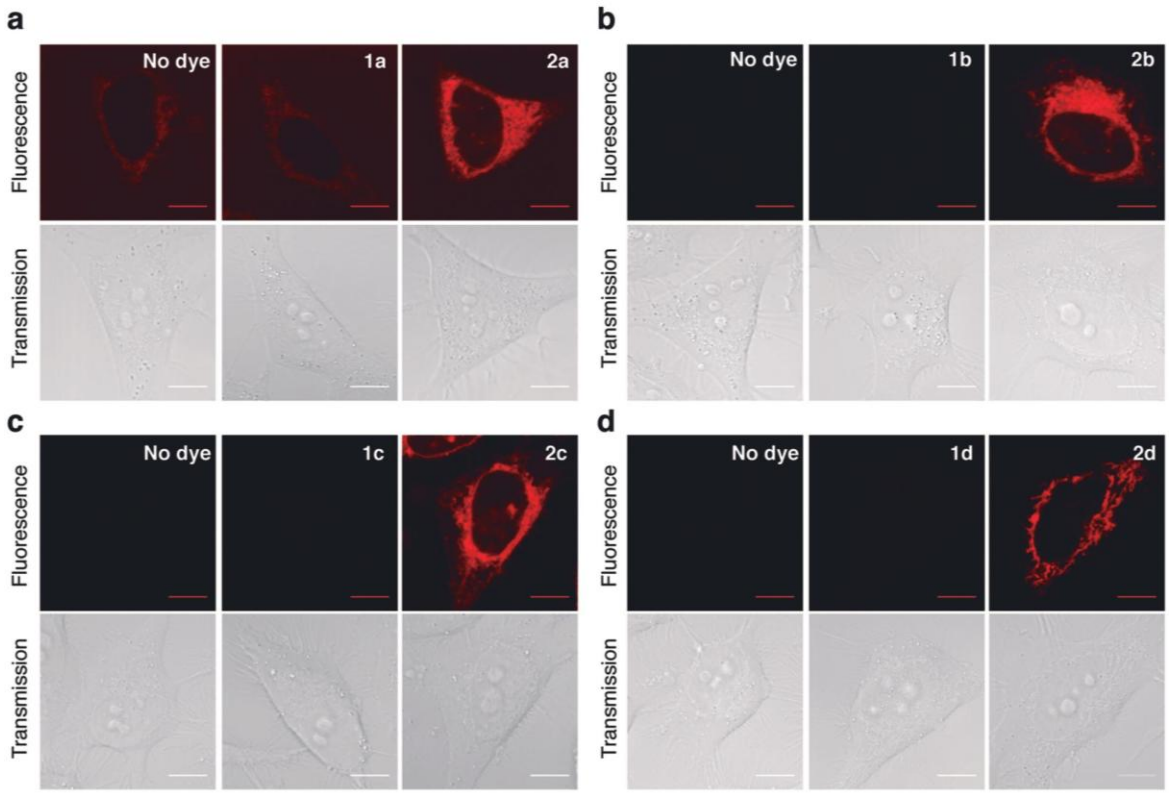


Fig. 7 Cellular analysis. Confocal micrographs of HeLa cells incubated with 1 μM of dye. (a–d) Side-by-side images were recorded using the same settings for direct comparison of the fluorescence intensities. (a) Ex/Em 405 nm/454–797 nm. (b–d) Ex/Em 488 nm/493–797 nm. Scale bars 10 μm.

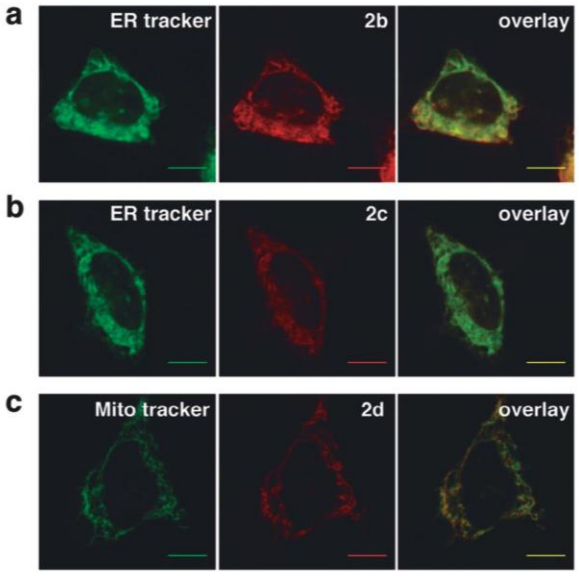


Fig. 8 Co-localization studies. Confocal micrographs of HeLa cells co-incubated with 1 μM of **2b** (a), **2c** (b), **2d** (c) and either an ER tracker or a Mitochondrial tracker (Mito tracker). Imaging conditions: (a, b) ER tracker Ex/Em 405 nm/454–797 nm, **2b,c** Ex/Em 488 nm/493–797 nm; (c) Mito tracker Ex/Em 543 nm/578–797 nm, **2d** Ex/Em 458 nm/465–593 nm. Scale bars 10 μm.

were well-suited for designing new fluorogen-based reporters as they did not undergo, unlike **2a–d**, any fluorescence activation in cells. However, even though **2a–d** are not usable for the design of fluorogen-based reporters, they demonstrate that this series of dyes can undergo strong fluorogenic activation upon motion restriction. In addition, **2a–d** could find interesting applications as intracellular fluorogenic markers. Indeed, compared to commercially available mitochondrial or ER trackers, which requires long incubation time and washing steps, staining of mitochondria and the ER with **2d** and **2a–c** is straightforward and instantaneous.

Discussion

Fluorogenic dyes (so-called fluorogens) are chromophores that are non-fluorescent by their own but give strongly fluorescent complex when bound to a complementary receptor, allowing high contrast imaging. This behavior has been recently used to develop fluorogen-based reporters composed of genetically encodable protein tags that bind specifically fluorogenic dyes and activate their fluorescence.[4,5] Recently, a new concept of fluorogen-based reporters has been proposed with the development of Y-FAST (Yellow Fluorescence-Activating and absorption-Shifting Tag).[19] Y-FAST's name highlights the two

View Article Online
Organic & Biomolecular Chemistry Paper

spectroscopic modifications used for fluorogen activation: upon binding, the fluorogen undergoes both an increase of the fluorescence quantum yield (because of motion restriction) and an absorption red shift (because of ionization of an auxochromic group). This second spectroscopic modification allows increasing the specificity of the fluorogen activation and thus the imaging contrast, as free or nonspecifically bound fluorogen does not contribute to the fluorescence signal.

With the goal in mind of developing Fluorescence-Activating and absorption-Shifting Tags (FASTs) emitting in the red, we studied two series of putative red fluorogenic dyes. The compounds **1a–d** and **2a–d** are push–pull compounds with an extended π-electron conjugation. They fluoresce, albeit very weakly, orange or red light in solution and possess high Stokes shifts, opening great prospects for multiplexed fluorescence observation using a single excitation wavelength. To be suitable for FASTs design, chromophores have to meet several criteria: (i) undergo fluorescence quantum yield exaltation upon motion restriction and (ii) hold potential for absorption red-shift upon binding. This latter property can arise either from ionization of an auxochromic group (as observed in the case of HBR and HMBR within Y-FAST) or from polarity change. The two series **1a–d** and **2a–d** presented in this paper respond positively to these criteria. They all exhibit increase of fluorescence quantum yield upon motion restriction. Moreover, compounds **1a–d** are protonated at physiological pHs and undergo significant absorption red-shift upon deprotonation, holding thus great potential for absorption red shift upon interaction with engineered tags complementary to their deprotonated state. Compounds **2a–d** do not have ionisable auxochromic group but they display negative solvatochromism, which could also be used to obtain absorption red shift upon binding engineered tags with low-polarity cavities.

Within the overall process of designing fluorogen-based reporter, the protein engineering that follows fluorogen design is highly time-consuming and expensive. Therefore, choosing (and thus evaluating) correctly the fluorogen is essential. Among the criteria a fluorogen must fulfill, showing no toxicity and no fluorescence activation in cells are primordial. Our cell viability experiment enabled to determine the range of concentrations usable for each fluorogen, and will thus allow us to define upper values for reporter:fluorogen affinities. Our cellular evaluation further allowed us to determine that only **1a–d** were suitable for designing new fluorogen-based reporters as they do not undergo, unlike **2a–d**, any significant fluorescence activation in cells. Although **2a–d** are not ideal for the design of fluorogen-based reporters, their ability to specifically stain the endoplasmic reticulum and mitochondria could find applications in intracellular imaging.

Conclusion

In conclusion, we present in this paper the development and full characterization of new red fluorogenic dyes holding great potential for various applications in fluorescence bioimaging. The design of red-emitting fluorogen-based reporters through engineering of complementary activating protein tags is a current research direction of our laboratory. The framework/methodology used in this study is generic and should find general applicability in the discovery of new fluorogenic dyes suitable for the design of fluorogen-based reporters and biosensors.

Materials and methods

Chemical synthesis

General information. Commercially available reagents were used as starting materials without further purification. NMR spectra were recorded on a AC Bruker spectrometer at 300 MHz for ^{1}H and 75 MHz for ^{13}C; chemical shifts are reported in ppm with protonated solvent as internal reference ^{1}H, CHD_2COCD_3 in CD_3COCD_3 2.05 ppm, CHD_2SOCD_3 in CD_3SOCD_3 2.50 ppm; ^{13}C, $^{13}CD_3COCD_3$ in CD_3COCD_3 29.84 ppm, $^{13}CD_3SOCD_3$ in CD_3SOCD_3 39.52 ppm; coupling constants J are given in Hz. Mass spectra (chemical ionization and electronic impact with NH_3, or high resolution) were performed by the Service de Spectrométrie de Masse de Chimie ParisTech or the Institut de Chimie Organique et Analytique de l'Université d'Orléans. Column chromatography was performed on silica gel 60 (0.040–0.063 nm) Merck. Analytical thin-layer chromatography (TLC) was conducted on Merck silica gel 60 F254 precoated plates.

Synthesis of 4-(methoxymethoxy)benzaldehyde. To a stirred solution of 4-hydroxybenzaldehyde (610 mg, 5.0 mmol) in anhydrous dichloromethane (6 mL) at 0 °C was added diisopropylethylamine (1292 mg, 10.0 mmol) and methoxymethyl chloride (0.57 mL, 7.5 mmol). The mixture was stirred overnight at room temperature. The reaction was quenched with a saturated aqueous solution of ammonium chloride, and the aqueous layer was extracted with dichloromethane. The combined organic layers were washed with brine and dried over magnesium sulfate. After concentrating *in vacuo*, the residue was purified by flash chromatography on silica gel with cyclohexane/ethylacetate (8/2, v/v) to yield the desired 4-(methoxymethoxy) benzaldehyde (762 mg, 92%) as a colorless oil. ^{1}H NMR (300 MHz, CD_3COCD_3, δ in ppm): 9.92 (s, 1H), 7.88 (d, J = 9.0 Hz, 2H), 7.21 (d, J = 8.7 Hz, 2H), 5.33 (s, 2H), 3.46 (s, 3H); ^{13}C NMR (75 MHz, CD_3COCD_3, δ in ppm): 191.3, 163.1, 132.4 (2C), 131.9, 117.2 (2C), 94.9, 55.4.

Synthesis of 4-hydroxycinnamaldehyde (1). To a stirred solution of sodium hydride (60% dispersion in mineral oil) (360 mg, 9.0 mmol) and tributyl (1,3-dioxolan-2-ylmethyl) phosphonium bromide (3323 mg, 9.0 mmol) in tetrahydrofurane (20 mL) was added 4-(methoxymethoxy)benzaldehyde (748 mg, 4.5 mmol). The solution was stirred overnight at room temperature in the presence of a catalytic amount of 18-crown-6. After addition of water, the reaction mixture was extracted with dichloromethane and the combined organic layers were washed with brine. The residue concentrated *in vacuo* was dissolved in tetrahydrofuran (10 mL). An aqueous

solution of hydrochloric acid was added and the solution was stirred until complete hydrolysis. The reaction mixture was extracted with dichloromethane and the combined organic layers were washed with brine and dried over magnesium sulfate. After concentrating *in vacuo*, the residue was purified by flash chromatography on silica gel with cyclohexane/ethyl-acetate (8/2, v/v) to yield the desired 4-hydroxycinnamaldehyde 1 (526 mg, 79%) as a yellow flocculent solid. ^{1}H NMR (300 MHz, CD_3COCD_3, δ in ppm): 9.64 (d, J = 7.8 Hz, 1H), 9.04 (s, 1H), 7.60 (m, 3H), 6.94 (d, J = 8.7 Hz, 2H), 6.61 (dd, J = 7.8, 15.9 Hz 1H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 193.9, 161.3, 153.7, 131.6 (2C), 127.0, 126.8, 116.9 (2C); MS (CI, NH_3): m/z 149.07 $[M + H]^+$, calcd mass for $[C_9H_9O_2]^+$: 149.06; HRMS (ESI): m/z 149.0598 $[M + H]^+$, calcd mass for $[C_9H_9O_2]^+$: 149.0603.

General protocol for the synthesis of 1a–d and 2a–d

To a stirred solution of aldehyde (1 or 2; 1 mmol) and heterocycle (a or b or c or d; 1 mmol) in ethanol (10 mL) was added the amine base (0.09 mL of 2-aminoethanol for **1a**, **1b**, **1d**, **2a**; 12 mg of 4-(4-methylamino)-pyridine for **1c**, **2b**, **2c**; 2 drops of piperidine for **2d**). The solution was stirred at reflux for 20 h. After cooling to 4 °C, the precipitate was filtered and the crude solid was washed with water and then with ethanol to give the desired product.

(*Z*)-5-((*E*)-3-(4-Hydroxyphenyl)allylidene)imidazolidine-2,4-dione (1a). Yellow powder (6%). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 10.95 (s, 1H), 10.56 (s, 1H), 9.87 (s, 1H), 7.34 (d, J = 8.7 Hz, 2H), 7.06 (dd, J = 11.7, 15.3 Hz, 1H), 6.81 (m, 3H), 6.221 (d, J = 12 Hz, 1H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 164.7, 158.2, 154.5, 137.2, 128.3 (2C), 128.2, 127.9, 119.3, 115.8 (2C), 109.8; MS (ESI): m/z 229.2 $[M - H]^-$, calcd mass for $[C_{12}H_9N_2O_3]^-$: 229.1; HRMS (ESI): m/z 231.0762 $[M + H]^+$, calcd mass for $[C_{12}H_{11}N_2O_3]^+$: 231.0770.

(*Z*)-5-((*E*)-3-(4-Hydroxyphenyl)allylidene)-2-thioxoimidazolidin-4-one (1b). Orange red solid (22%). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 12.17 (s, 2H), 9.95 (s, 1H), 7.40 (d, J = 8.4, 2H), 7.32 (dd, J = 12.0, 15.3 Hz, 1H), 6.99 (d, J = 15.3, 1H), 6.79 (d, J = 8.4, 2H), 6.34 (d, J = 12.0 Hz, 1H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 176.2, 164.8, 158.7, 140.0, 128.9 (2C), 128.1, 127.8, 119.4, 115.9 (2C), 113.4; MS (ESI): m/z 245.3 $[M - H]^+$, calcd mass for $[C_{12}H_9N_2O_2S]^-$: 245.0; HRMS (ESI): m/z 247.0540 $[M + H]^+$, calcd mass for $[C_{12}H_{11}N_2O_2S]^+$: 247.0541.

(*Z*)-5-((*E*)-3-(4-Hydroxyphenyl)allylidene)-2-thioxothiazolidin-4-one (1c). Orange solid (46%). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 13.45 (s, 1H), 7.51 (d, J = 9.0 Hz, 2H), 7.32 (d, J = 11.4 Hz, 1H), 7.22 (d, J = 15.0 Hz, 1H), 6.68 (m, 3H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 195.1, 168.7, 159.7, 145.6, 133.2, 130.2 (2C), 126.9, 124.6, 120.5, 116.0 (2C); MS (CI, NH_3): m/z 264.08 $[M + H]^+$, calcd mass for $[C_{12}H_{10}NO_2S_2]^+$: 264.02; HRMS (ESI): m/z 264.0149 $[M + H]^+$, calcd mass for $[C_{12}H_{10}NO_2S_2]^+$: 264.0153.

4-((1*E*,3*E*)-4-(4-Hydroxyphenyl)buta-1,3-dien-1-yl)-1-methylpyridin-1-ium iodide (1d). Brown solid (21%). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 8.72 (d, J = 6.6 Hz, 2H), 8.05 (d, J = 6.9 Hz, 2H), 7.78 (ddd, J = 15.6, 7.5 Hz, 2.4 Hz 1H), 7.46 (d, J = 8.7 Hz, 2H), 7.03 (m, 2H), 6.78 (m, 3H), 4.19 (s, 3H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 159.4, 152.6, 144.7 (2C), 142.7, 141.2, 129.4 (2C), 127.0, 124.8, 124.6, 122.8 (2C), 116.0 (2C), 46.7; MS (ESI): m/z 238.2 $[M']^+$, calcd mass for $[C_{16}H_{16}NO]^+$: 238.1; HRMS (ESI): m/z 238.1228 $[M']^+$, calcd mass for $[C_{16}H_{16}NO]^+$: 238.1226.

(*Z*)-5-((*E*)-3-(4-(Dimethylamino)phenyl)allylidene) imidazolidine-2,4-dione (2a). Red powder (42%). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 10.70 (s, 2H), 7.35 (d, J = 8.7 Hz, 2H), 7.01 (dd, J = 11.7, 15.3 Hz, 1H), 6.80 (d, J = 15.3 Hz, 1H), 6.72 (d, J = 8.7 Hz, 2H), 6.22 (d, J = 11.7 Hz, 1H), 2.95 (s, 6H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 164.8, 154.6, 150.5, 137.8, 128.1 (2C), 127.4, 124.5, 117.6, 112.4 (2C), 110.5, 39.8 (2C); MS (CI, NH_3): m/z 258.17 $[M + H]^+$, calcd mass for $[C_{14}H_{16}N_3O_2]^+$: 258.12; HRMS (ESI): m/z 258.1249 $[M + H]^+$, calcd mass for $[C_{14}H_{16}N_3O_2]^+$: 258.1243.

(*Z*)-5-((*E*)-3-(4-(Dimethylamino)phenyl)allylidene)-2-thioxoimidazolidin-4-one (2b). Dark powder (55%). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 12.18 (s, 1H), 12.07 (s, 1H), 7.40 (d, J = 8.4 Hz, 2H), 7.26 (m, 1H), 6.97 (d, J = 15.0 Hz, 1H), 6.73 (d, J = 8.4 Hz, 2H), 6.35 (d, J = 11.7 Hz, 1H), 2.97 (s, 6H).^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 175.5, 164.7, 150.9, 141.0, 128.7 (2C), 127.0, 124.3, 117.6, 114.4, 112.1 (2C), 39.8 (2C); MS (CI, NH_3): m/z 274.13 $[M + H]^+$, calcd mass for $[C_{14}H_{16}N_3OS]^+$: 274.10; HRMS (ESI): m/z 274.1009 $[M + H]^+$, calcd mass for $[C_{14}H_{16}N_3OS]^+$: 274.1014.

(*Z*)-5-((*E*)-3-(4-(Dimethylamino)phenyl)allylidene)-2-thioxothiazolidin-4-one (2c). Dark powder (42%). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 13.45 (s, 1H), 7.51 (d, J = 9.0 Hz, 2H), 7.32 (d, J = 11.4 Hz, 1H), 7.22 (d, J = 15.0 Hz, 1H), 6.68 (m, 3H), 2.99 (s, 6H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 194.7, 168.6, 151.6, 146.6, 133.9, 130.0 (2C), 123.1, 122.3, 118.3, 111.9 (2C), 39.7 (2C); MS (CI, NH_3): m/z 291.10 $[M + H]^+$, calcd mass for $[C_{14}H_{15}N_2OS_2]^+$: 291.06; HRMS (ESI): m/z 291.0623 $[M + H]^+$, calcd mass for $[C_{14}H_{15}N_2OS_2]^+$: 291.0626.

4-((1*E*,3*E*)-4-(4-(Dimethylamino)phenyl)buta-1,3-dien-1-yl)-1-methylpyridin-1-ium iodide (2d). Dark powder (41%). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 8.69 (d, J = 6.8 Hz, 2H), 8.02 (d, J = 6.9 Hz, 2H), 7.79 (dt, J = 15.4, 5.0 Hz, 1H), 7.47 (d, J = 8.8 Hz, 2H), 7.01 (d, J = 5.1 Hz, 2H), 6.72 (m, 3H), 4.18 (s, 3H), 2.98 (s, 6H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 152.7, 151.1, 144.5 (2C), 143.2, 142.0, 129.1 (2C), 123.5, 123.3, 123.1, 122.4 (2C), 112.0 (2C), 46.5, 39.8 (2C); MS (ESI): m/z 265.2 $[M']^+$, calcd mass for $[C_{18}H_{21}N_2]^+$: 265.2; HRMS (ESI): m/z 265.1696 $[M']^+$, calcd mass for $[C_{18}H_{21}N_2]^+$: 265.1699.

Physico-chemical experiments

pH measurements were performed on a standard pH meter PHM210 Radiometer Analytical (calibrated with aqueous buffers at pH 4 and 7 or 10) with a Crison 5208 Electrode (Barcelona, Spain). UV/Vis absorption spectra were recorded in 1 cm × 1 cm quartz cuvettes (Hellma) on a diode array UV/Vis spectrophotometer (Evolution array, Thermo Scientific). Corrected fluorescence spectra upon one-photon excitation

were recorded with a Photon Technology International QuantaMaster QM-1 spectrofluorimeter (PTI, Monmouth Junction, NJ) equipped with a Peltier cell holder (TLC50, Quantum Northwest, Shoreline, WA). The overall emission quantum yields after one-photon excitation ϕ were determined as previously described.[19]

Fluorescence microscopy

Confocal micrographs were acquired on a Zeiss LSM 710 Laser Scanning Microscope equipped with a Plan Apochromat 63×/1.4 NA oil immersion objective. ZEN software was used to collect the data. Images were analyzed with Image J.

Mammalian cell culture

HeLa cells were cultured in DMEM supplemented with phenol red, Glutamax I, 10% (vol/vol) fetal calf serum (FCS), and 1% penicillin–streptomycin at 37 °C within a 5% CO_2 atmosphere. For microscopic imaging, cells were seeded in µSlide IBIDI (Biovalley) coated with poly-L-lysine. Cells were washed with PBS, incubated in DMEM (without phenol red) complemented with **1a–d** and **2a–d** at the indicated concentration, and imaged directly without washing. For the co-localization study, HeLa cells were prestained with ER-Tracker™ Blue-White DPX (Molecular Probes, Life Technologies) or MitoTracker® Red CMXRos (Molecular Probes, Life Technologies) following the manufacturer's protocol.

Toxicity cellular assay

HeLa cells were incubated with **1a–d** and **2a–d** solutions at the indicated concentrations for various durations. Cell viability was assayed by fluorescence microscopy using the LIVE/DEAD® viability/cytotoxicity assay kit (Molecular Probes, Life Technologies) following the manufacturer's protocol.

Acknowledgements

This work has been supported by the Agence National de la Recherche (ANR-14-CE09-0002-01), a PhD grant (C. L.) from Region Ile-de-France in the framework of DIM NanoK, PSL Research University (project IMRESOV), France BioImaging and the Equipex Morphoscope 2.

Notes and references

1 H. O'Hare, K. Johnsson and A. Gautier, *Curr. Opin. Struct. Biol.*, 2007, **17**, 488–494.
2 M. J. Hinner and K. Johnsson, *Curr. Opin. Biotechnol.*, 2010, **21**, 766–776.
3 K. M. Dean and A. E. Palmer, *Nat. Chem. Biol.*, 2014, **10**, 512–523.
4 M. P. Bruchez, *Curr. Opin. Chem. Biol.*, 2015, **27**, 18–23.
5 L. Jullien and A. Gautier, *Methods Appl. Fluoresc.*, 2015, **3**, 042007.
6 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.
7 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.
8 C. A. Telmer, R. Verma, H. Teng, S. Andreko, L. Law and M. P. Bruchez, *ACS Chem. Biol.*, 2015, **10**, 1239–1246.
9 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.
10 J. S. Paige, K. Y. Wu and S. R. Jaffrey, *Science*, 2011, **333**, 642–646.
11 R. L. Strack, M. D. Disney and S. R. Jaffrey, *Nat. Methods*, 2013, **10**, 1219–1224.
12 P. Wang, J. Querard, S. Maurin, S. S. Nath, T. Le Saux, A. Gautier and L. Jullien, *Chem. Sci.*, 2013, **4**, 2865–2873.
13 K. Y. Han, B. J. Leslie, J. Fei, J. Zhang and T. Ha, *J. Am. Chem. Soc.*, 2013, **135**, 19033–19038.
14 W. Song, R. L. Strack, N. Svensen and S. R. Jaffrey, *J. Am. Chem. Soc.*, 2014, **136**, 1198–1201.
15 I. Yapici, K. S. S. Lee, T. Berbasova, M. Nosrati, X. Jia, C. Vasileiou, W. Wang, E. M. Santos, J. H. Geiger and B. Borhan, *J. Am. Chem. Soc.*, 2015, **137**, 1073–1080.
16 Y. Hori, T. Norinobu, M. Sato, K. Arita, M. Shirakawa and K. Kikuchi, *J. Am. Chem. Soc.*, 2013, **135**, 12360–12365.
17 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.
18 G. Lukinavicius, L. Reymond, E. D'Este, A. Masharina, F. Göttfert, H. Ta, A. Güther, M. Fournier, S. Rizzo, H. Waldmann, C. Blaukopf, C. Sommer, D. W. Gerlich, H.-D. Arndt, S. W. Hell and K. Johnsson, *Nat. Methods*, 2014, **11**, 731–733.
19 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. Vriz, T. Le Saux, L. Jullien and A. Gautier, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 497–502.
20 R. Y. Tsien, *Annu. Rev. Biochem.*, 1998, **67**, 509–544.
21 N.-S. Cheng, *Ind. Eng. Chem. Res.*, 2008, **47**, 3285–3288.
22 G. R. G. Rosania, J. W. J. Lee, L. L. Ding, H.-S. H. Yoon and Y.-T. Y. Chang, *J. Am. Chem. Soc.*, 2003, **125**, 1130–1131.
23 J. A. Coch-Frugoni, *Gazz. Chim. Ital.*, 1957, **87**, 403–407.

Supporting Information

Design and characterization of red fluorogenic push-pull chromophores holding great potential for bioimaging and biosensing

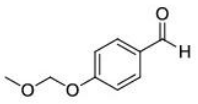
Chenge Li[1,2], Marie-Aude Plamont[1,2], Isabelle Aujard[1,2], Thomas Le Saux[1,2], Ludovic Jullien[1,2,*] & Arnaud Gautier[1,2,*]

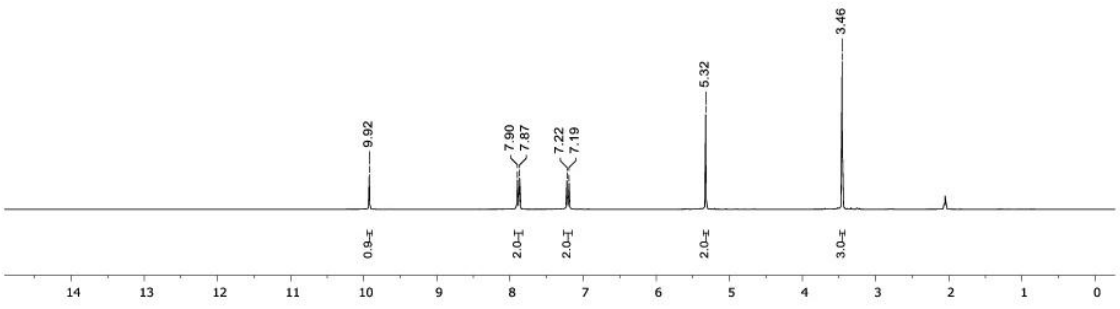
[1] École Normale Supérieure, PSL Research University, UPMC Univ Paris 06, CNRS, Département de Chimie, PASTEUR, 24 rue Lhomond, 75005 Paris, France

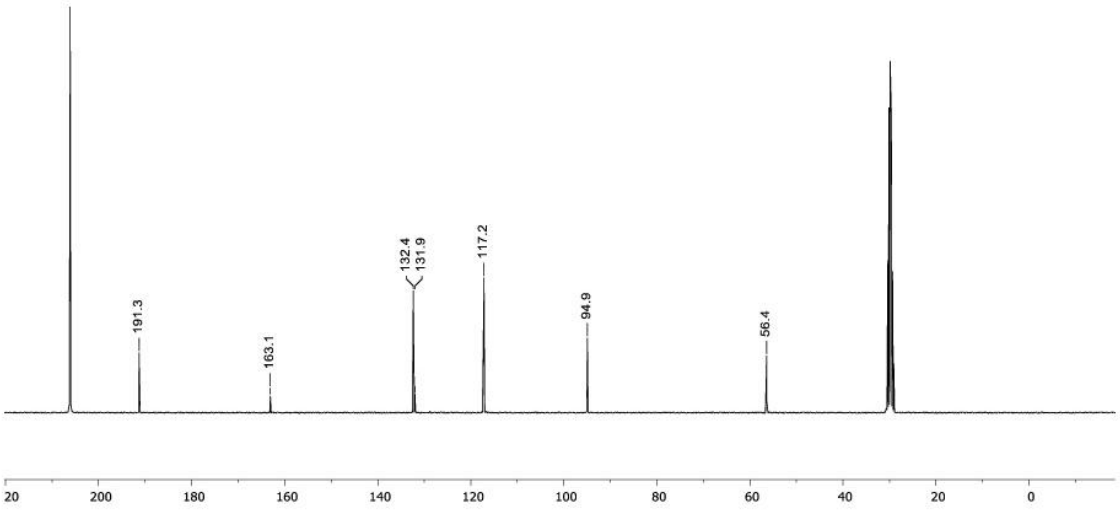
[2] Sorbonne Universités, UPMC Univ Paris 06, ENS, CNRS, PASTEUR, 75005 Paris, France

This file contains the NMR spectra of all the compounds described in this paper. For each compounds are given the ^{1}H NMR spectrum (Top spectrum) and the ^{13}C NMR spectrum (Bottom spectrum). NMR spectra were recorded on a AC Bruker spectrometer at 300 MHz for ^{1}H and 75 MHz for ^{13}C; chemical shifts are reported in ppm with protonated solvent as internal reference ^{1}H, CHD_2COCD_3 in CD_3COCD_3 2.05 ppm, CHD_2SOCD_3 in CD_3SOCD_3 2.50 ppm; ^{13}C, $^{13}CD_3COCD_3$ in CD_3COCD_3 29.84 ppm, $^{13}CD_3SOCD_3$ in CD_3SOCD_3 39.52 ppm.

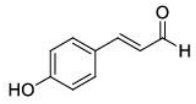
4-(methoxymethoxy)benzaldehyde

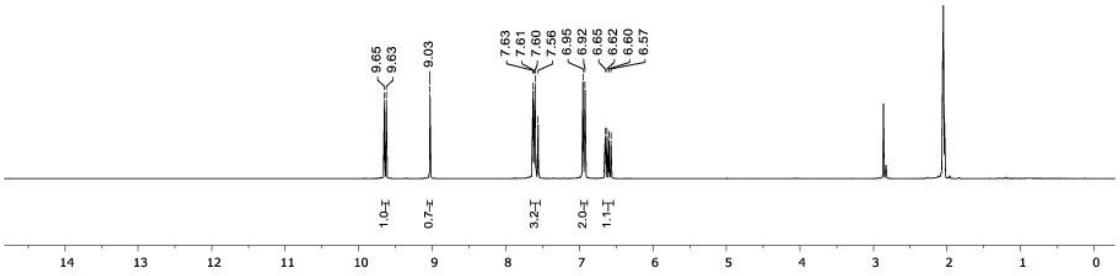


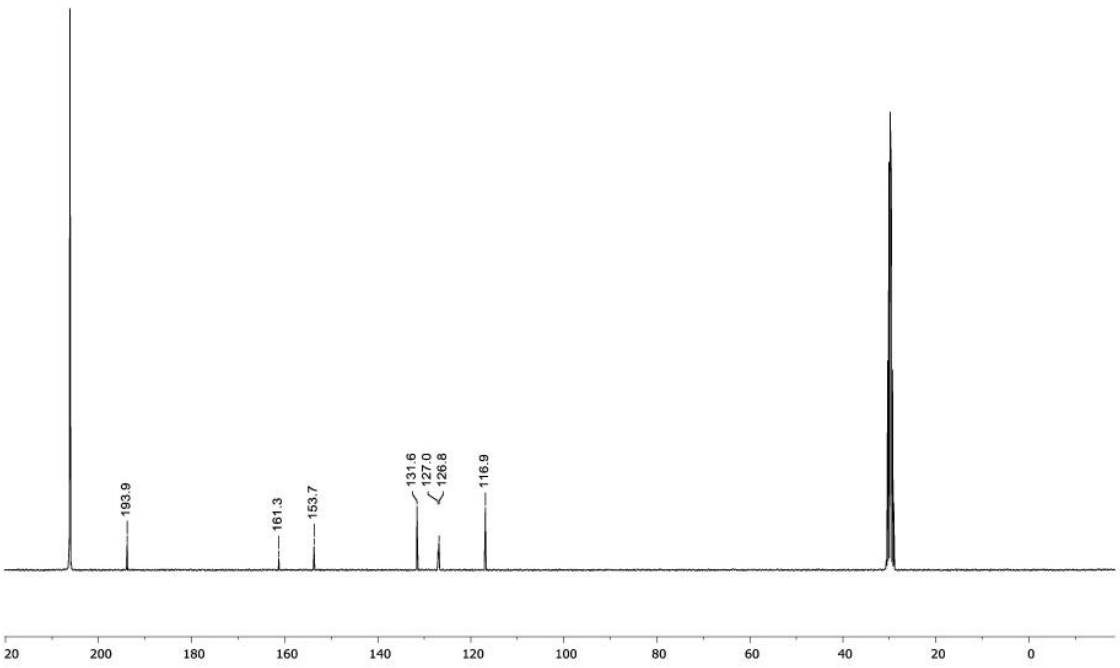




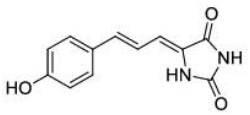
4-hydroxycinnamaldehyde (1)

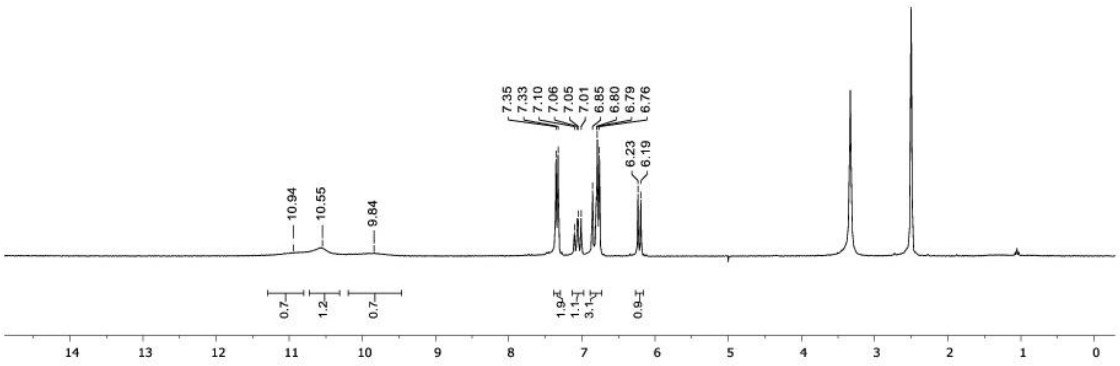


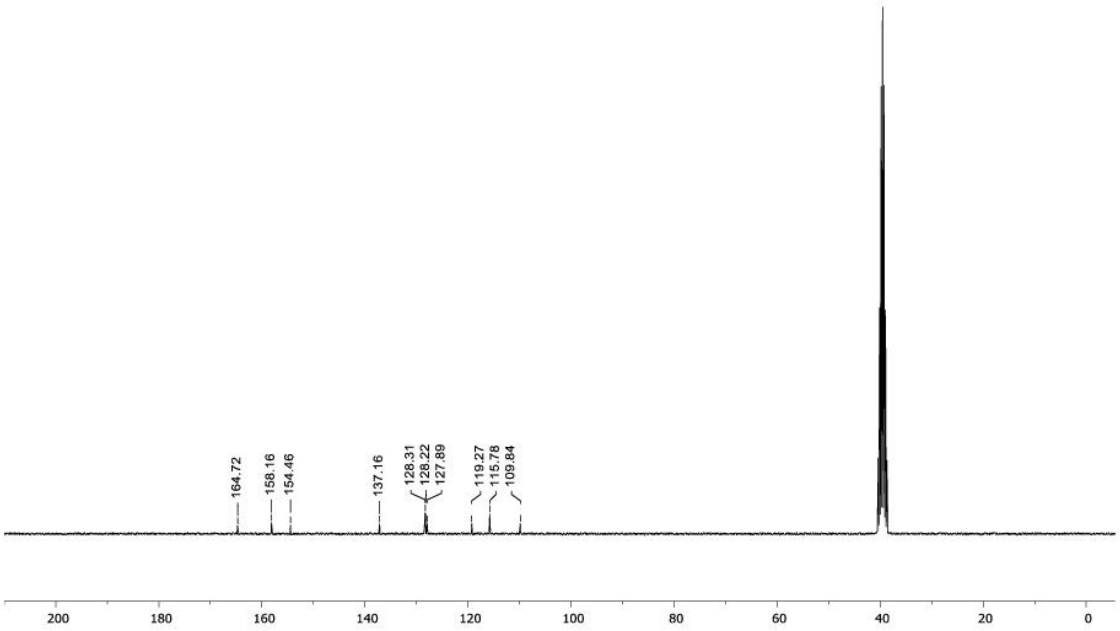




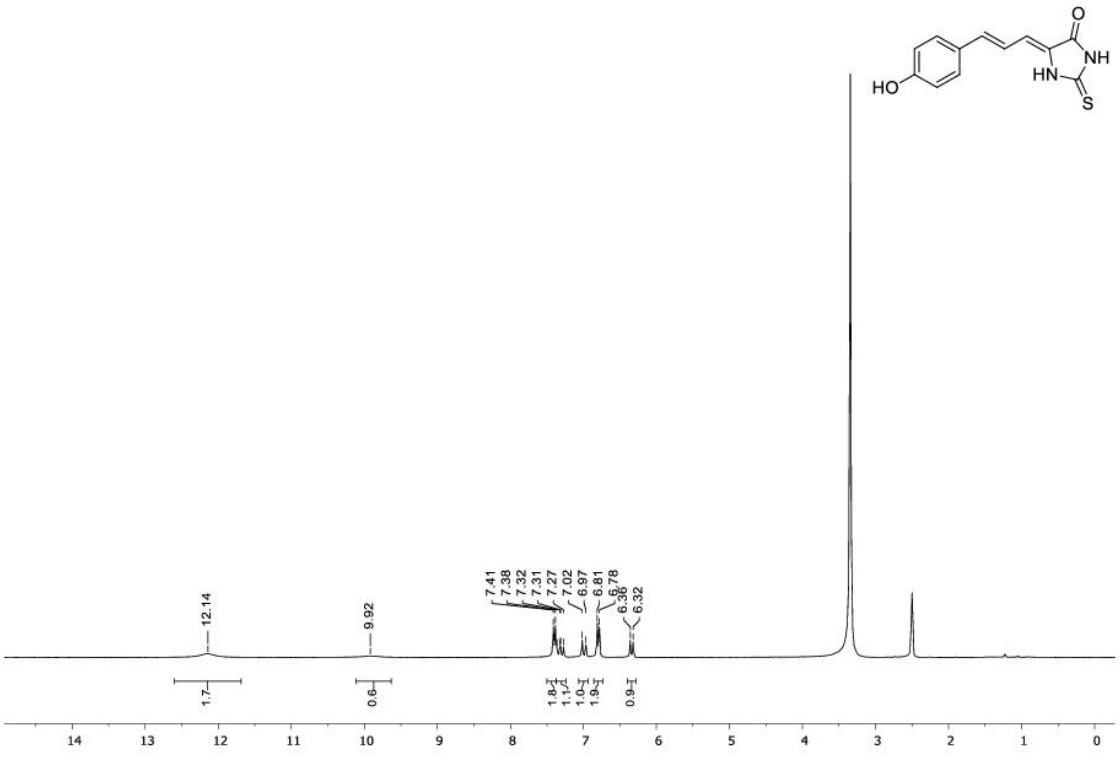
(Z)-5-((E)-3-(4-hydroxyphenyl)allylidene)imidazolidine-2,4-dione (1a)

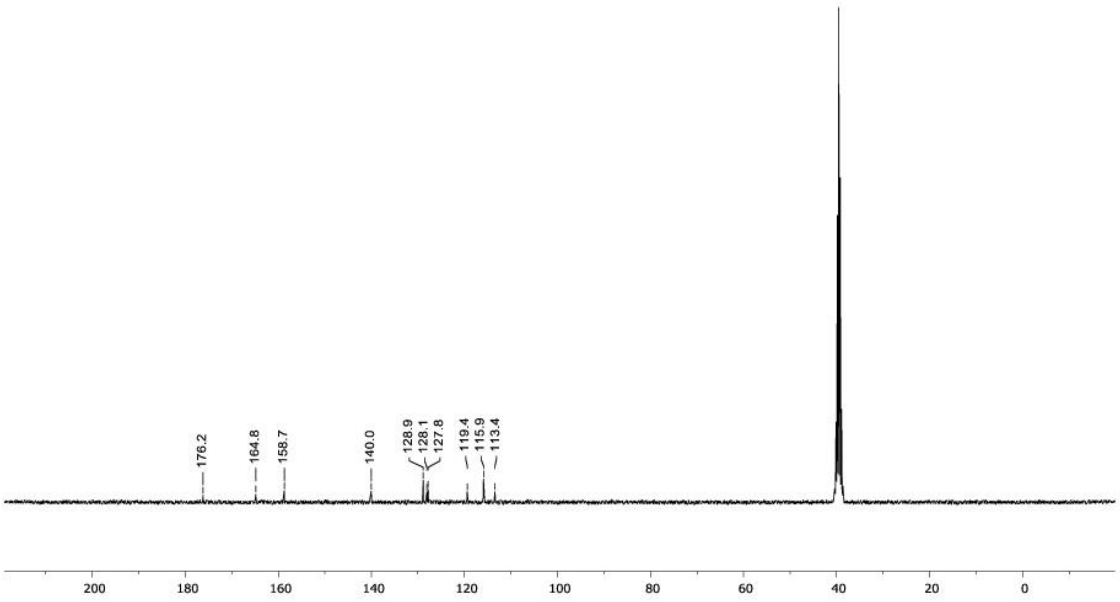




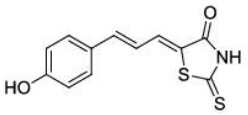


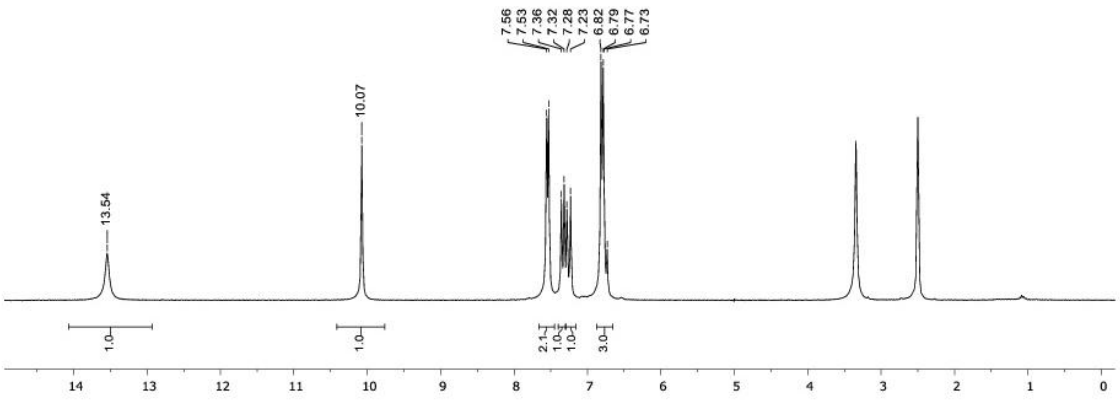
(Z)-5-((E)-3-(4-hydroxyphenyl)allylidene)-2-thioxoimidazolidin-4-one (1b)

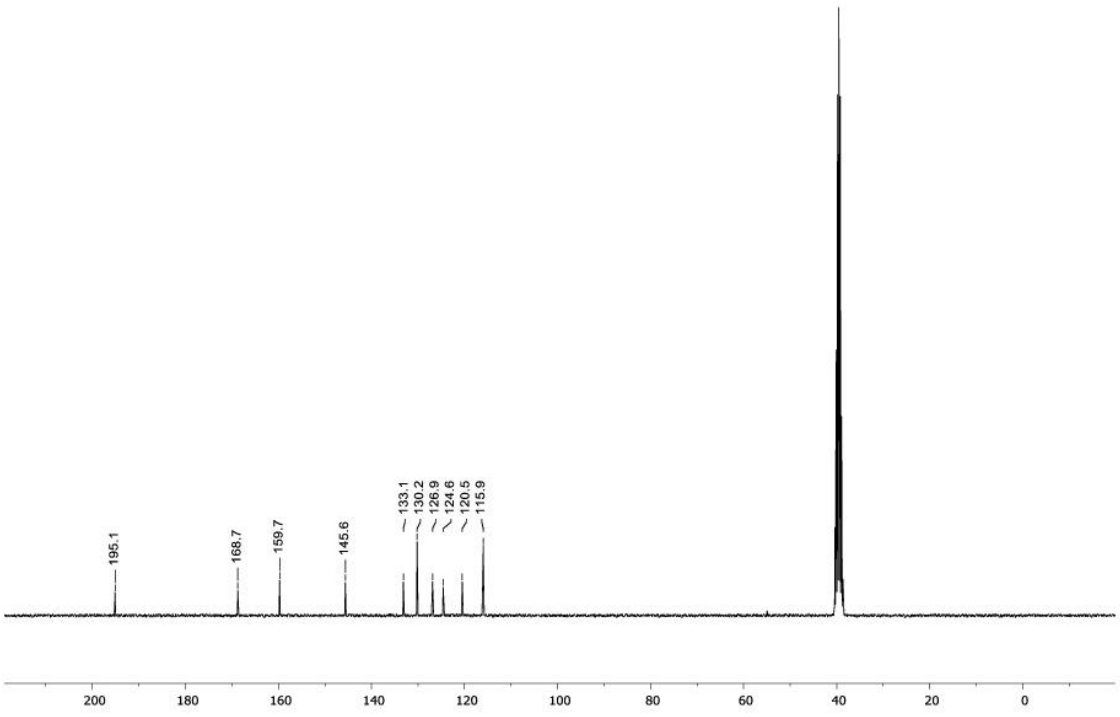




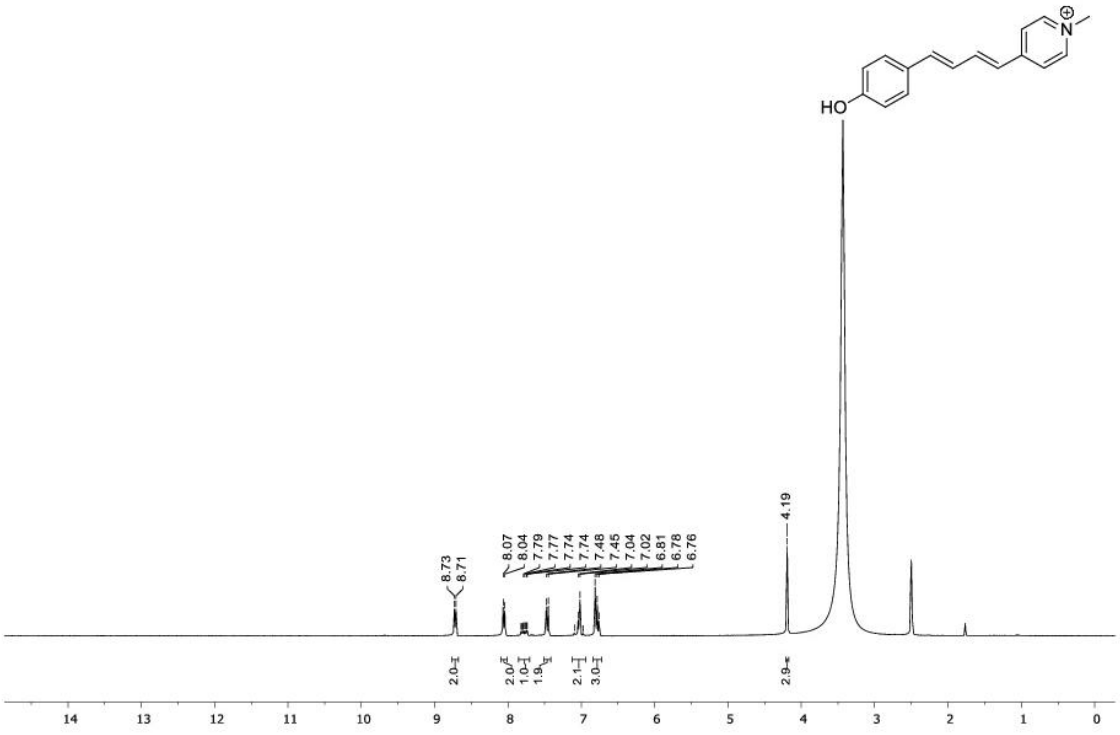
(Z)-5-((E)-3-(4-hydroxyphenyl)allylidene)-2-thioxothiazolidin-4-one (1c)

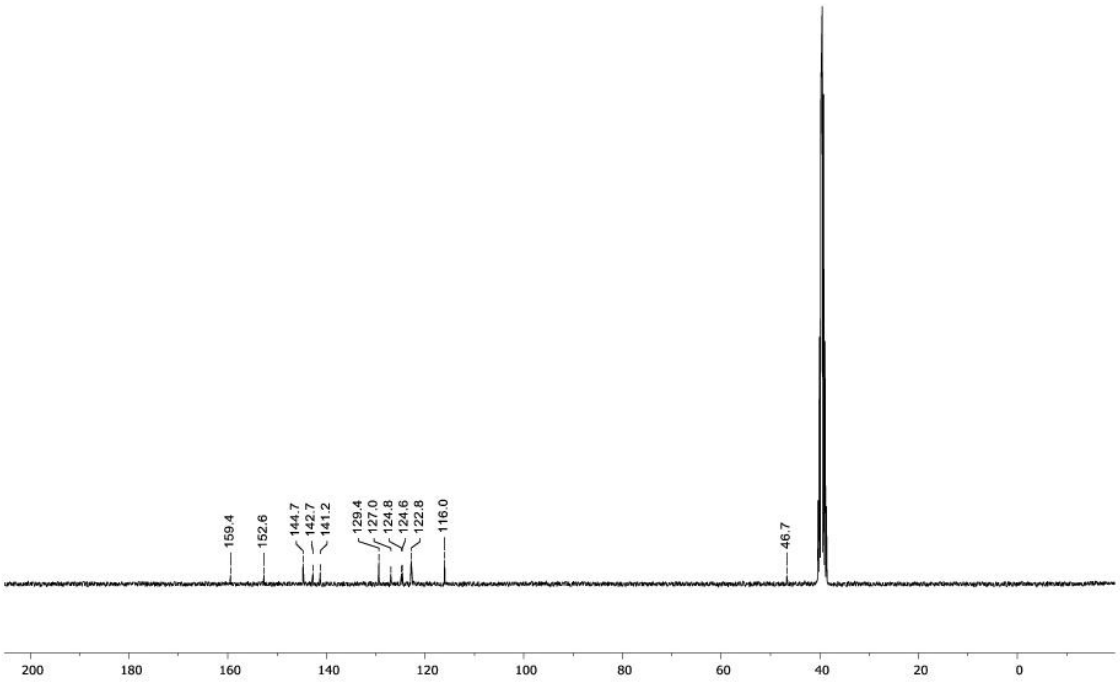




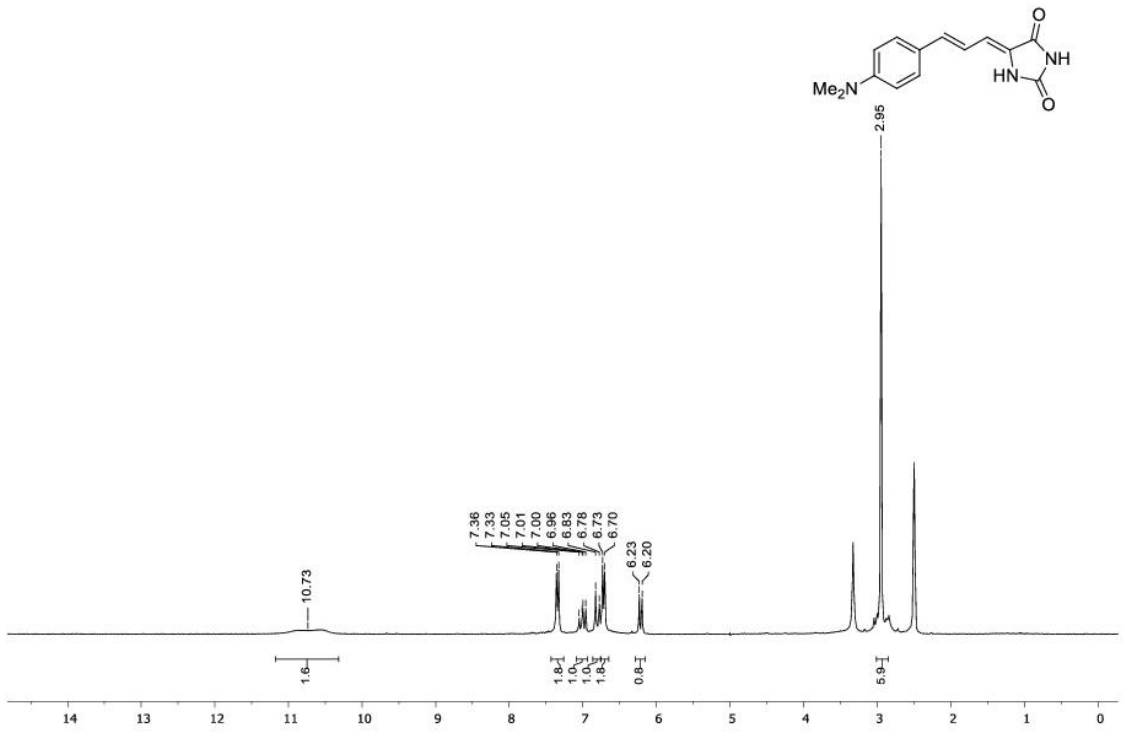


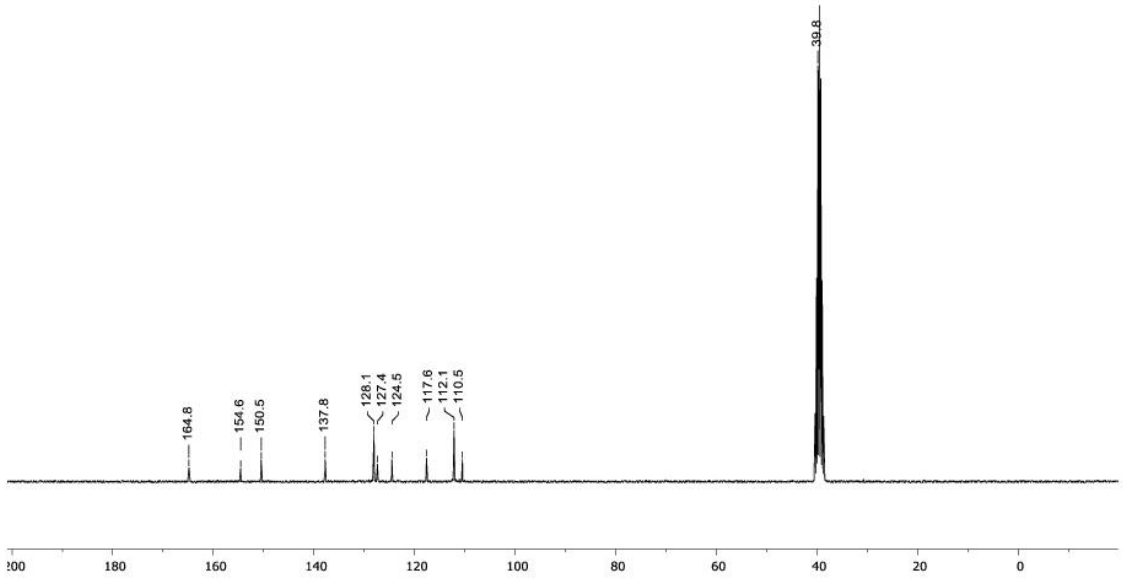
4-((1E,3E)-4-(4-hydroxyphenyl)buta-1,3-dien-1-yl)-1-methylpyridin-1-ium iodide (1d)



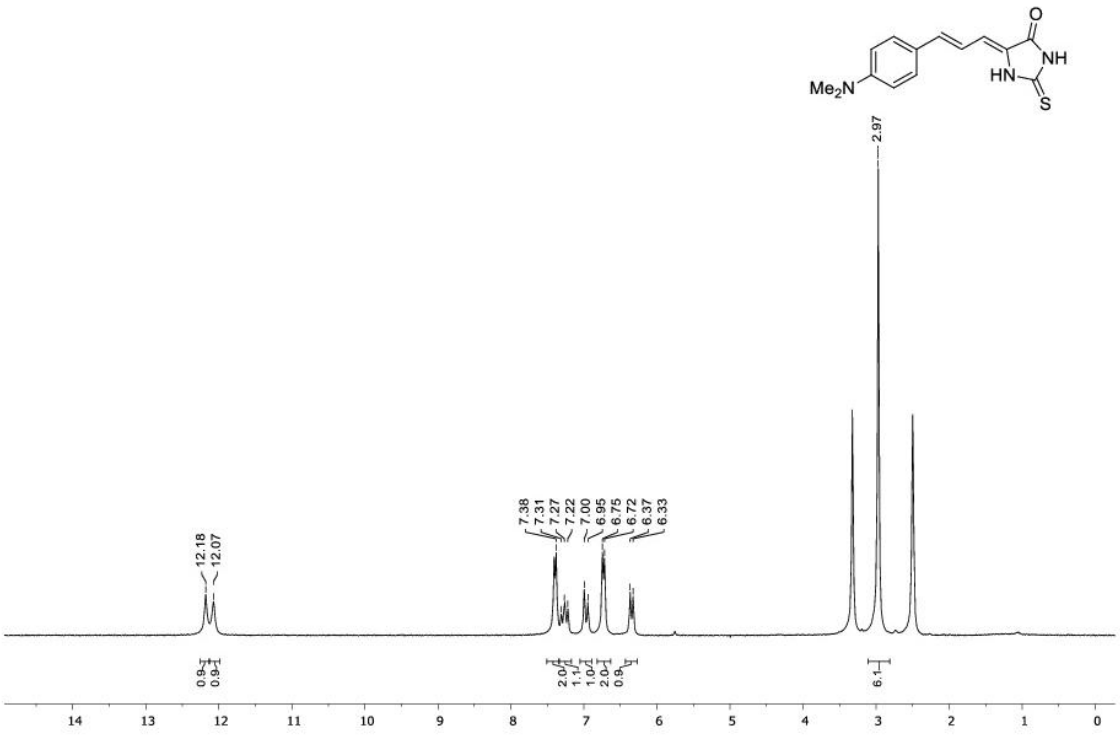


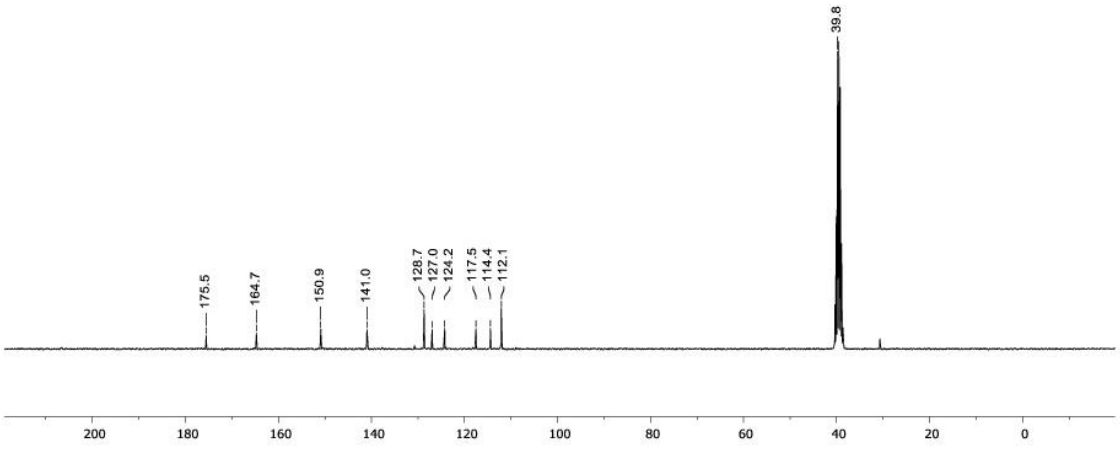
(Z)-5-((E)-3-(4-(dimethylamino)phenyl)allylidene)imidazolidine-2,4-dione (2a)



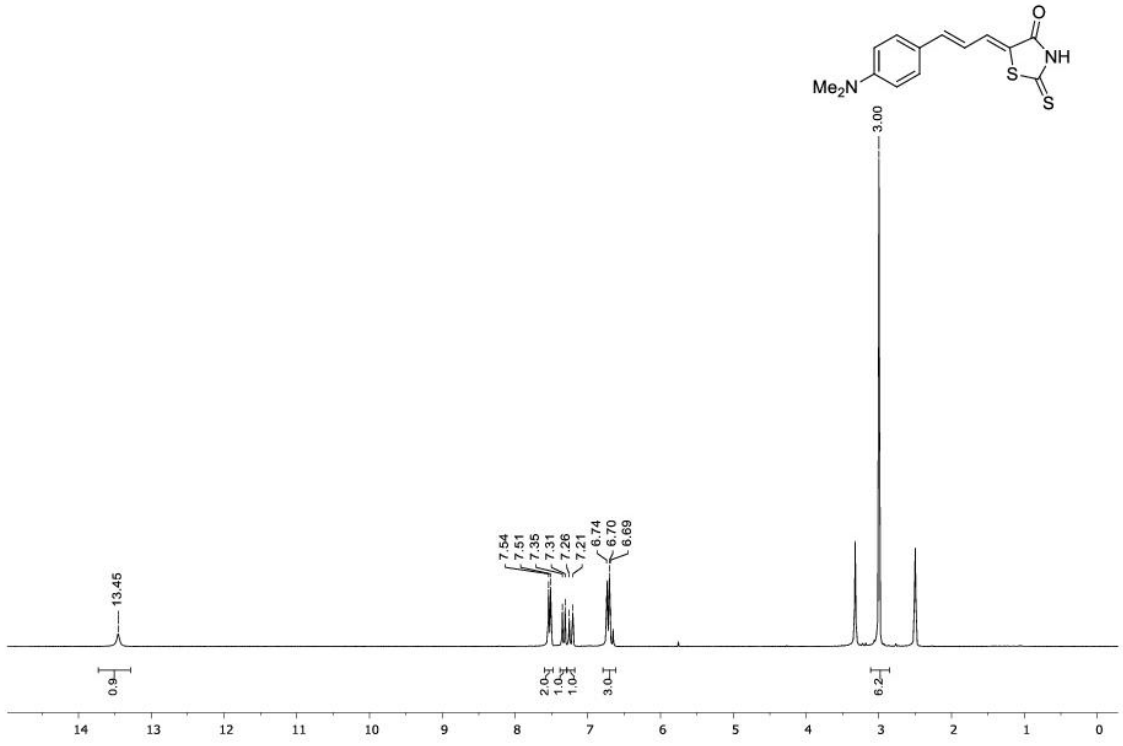


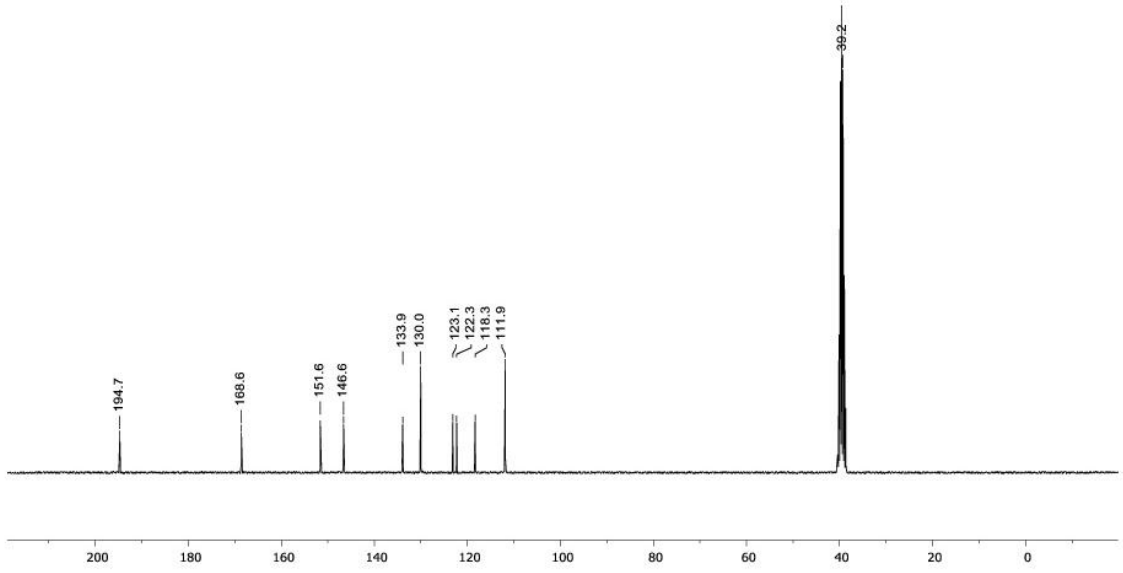
(Z)-5-((E)-3-(4-(dimethylamino)phenyl)allylidene)-2-thioxoimidazolidin-4-one (2b)



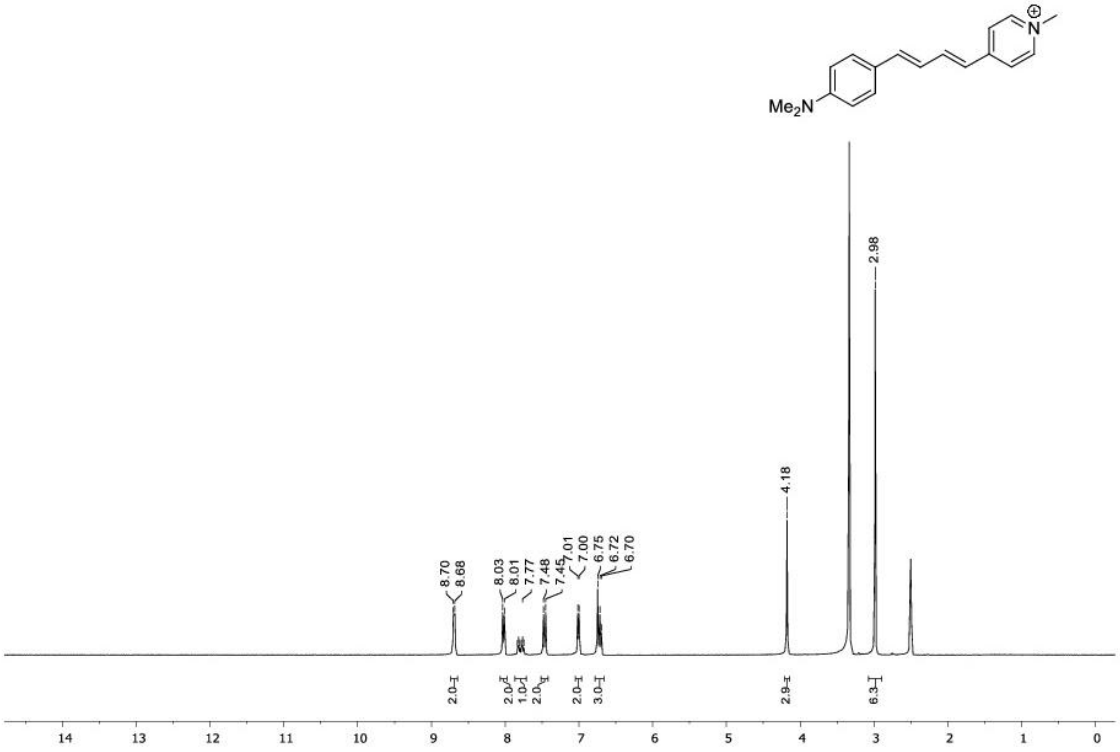


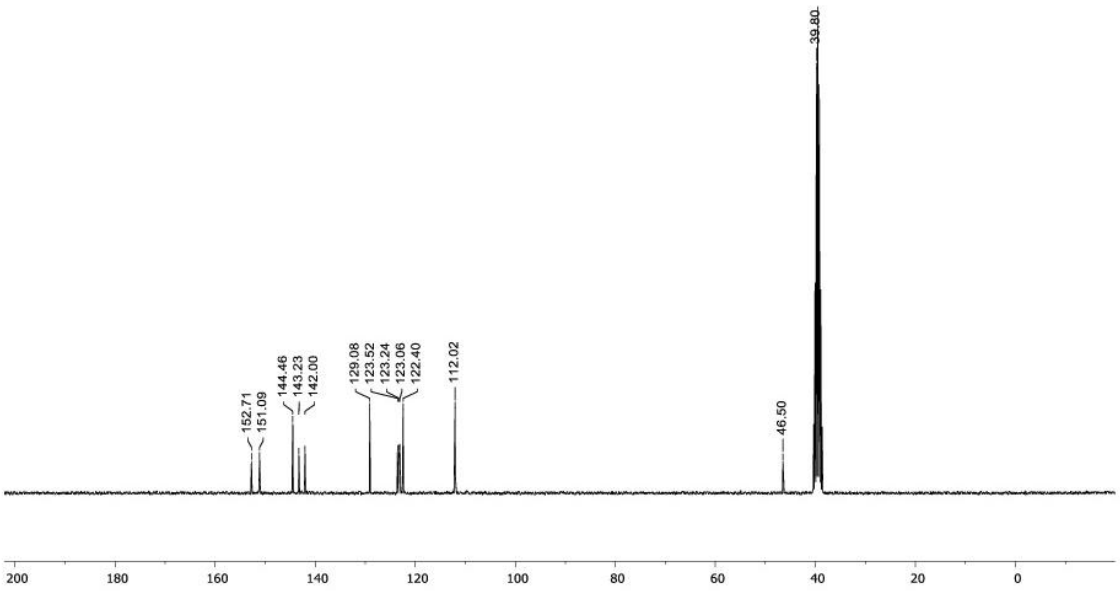
(Z)-5-((E)-3-(4-(dimethylamino)phenyl)allylidene)-2-thioxothiazolidin-4-one (2c)





4-((1E,3E)-4-(4-(dimethylamino)phenyl)buta-1,3-dien-1-yl)-1-methylpyridin-1-ium iodide (2d)





III-2 Directed evolution of protein tags binding and activating red-emitting fluorogens

III-2.1 Introduction

We have presented above the general method to develop red-emitting fluorogens and obtained several potential red-emitting fluorogens for fluorogenic labeling. In this part, the general idea for engineering protein tags able to bind and activate some of these new fluorogens, will be introduced.

Protein engineering is commonly conducted by two strategies, rational protein design and directed evolution. Rational protein design aims at generating desired variants through site-directed mutagenesis of original protein gene. This strategy requires full knowledge of the structure and function of the original protein, which is important to design mutants with expected properties. However, influence of structural modifications on functional changes is difficult to predict. Another strategy is directed evolution for which desired protein variant will be selected from large protein libraries generated by mutagenesis of the original protein gene. Compared to rational design, directed evolution does not need to understand the original protein in detail and the large numbers of mutants increase the chance to obtain protein variants with desired properties. However, screening out rare desired mutants from mutant libraries takes a lot of time and effort. High throughput selection methods are thus required for accomplishing directed evolution efficiently.

FAST was originally engineered by directed evolution. The high throughput method we chose to select desired variants is the fluorescence-activated cell sorting (FACS). FACS is a cell separating instrument based on fluorescence which was developed for sorting specific cell types from cell mixture. Cell sample should be prepared as fluorescently stained single cell suspension. Under certain pressure, cell suspension is pressed into flow chamber where cells are enclosed by sheath fluid (simple buffer under pressure), leading to a single-cell stream. This stream of cells passes through detection lasers and desired cells emitting fluorescence upon the excitation of laser light were detected by the detector. At the break off point, the stream of cells will be charged accordingly and be partitioned into droplets. Droplets are then separated by passing within an electrical field. Sorting up to 50,000 cells per second, FACS allows one to achieve high throughput cell analysis. As the master factor of the sorting, the

fluorescence of cells is usually generated by immunofluorescent labeling with fluorescent dye-tagged antibodies or fluorescent stains. We used the yeast surface display technique to build yeast libraries expressing protein variants on the surface, so desired protein variants could bind with fluorogens in solution and activate their fluorescence. The yeast surface display method was established by the group of Wittrup[2] and has been widely applied for efficient protein engineering by FACS.[3] Generally, protein variants are fused to one subunit of a-agglutinin (Aga2p) that is tethered to yeast cell wall by linking with Aga1p (the other subunit of a-agglutinin) through two disulfide bonds (Figure 3.2). Usually the hemagglutinin (HA) or c-Myc epitope is fused to the N or C terminal of fusion protein as the site for immunolabeling (Figure 3.2).

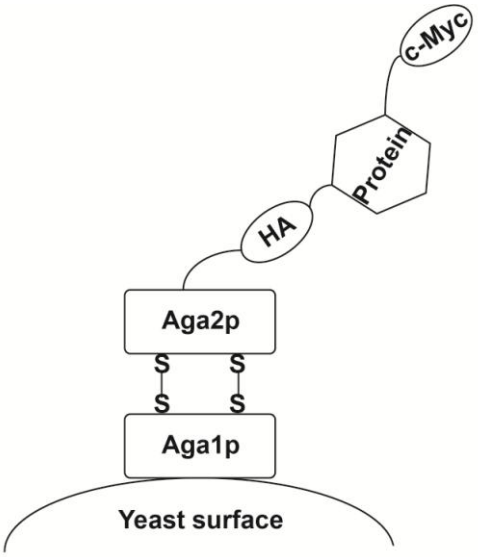


Figure 3.2 Schema of yeast surface display. Protein variants are expressed on the surface of yeast cell as Aga2p fusion protein. Aga2p is attached to yeast cell wall by linking with Aga1p through two disulfide bonds. Fused to the N or C terminal of fusion protein, HA or c-Myc epitope acts as the site of immunolabeling for detecting protein expression. Inspired by Chao *et al.*[4]

Before starting to screen complementary protein tags for putative fluorogens, we needed to identify protein scaffolds containing cavity that could possibly accept the fluorogens and activate their fluorescence. PYP had originally been chosen as protein scaffold to develop FAST because its chromophore parahydroxycinnamoyl (HC) shared structure similarity with

HMBR. In Section III-1, the red-emitting fluorogens we developed were analogues of HMBR with an additional double bond. So firstly we tested their binding abilities with FAST. Except for **1c** (4-hydroxyphenyl)allylidene-rhodanine (HPAR), which showed a little fluorescence activation by interacting with FAST, the other fluorogens did not exhibit any binding affinity with FAST. We thus decided to focus first on the evolution of FAST variants able to bind the fluorogenic HPAR and to activate its red fluorescence. We built a library of FAST variants by random mutagenesis using error-prone PCR. Error-prone PCR is a common technique for random mutagenesis, which generates mutations during PCR process by using error prone DNA polymerases. We were able to identify from this libraries FAST variants able to bind and activate the red fluorescence of HPAR.

III-2.2 Results and discussions

Construction of the library of FAST variants.

Our library of FAST variants was generated by random mutagenesis of FAST gene by error-prone PCR. A mutation rate of 3.8 nucleotides/gene was obtained after mixing PCR reactions with different error rates varied by adjusting the amount of template gene. The PCR product was cloned into the pCTCON2 plasmid from Wittrup's lab, allowing the expression of the FAST variants in fusion to Aga2p. The resulting plasmids were transformed into DH10B *E. coli* cells with a yield of 10^6 transformants. DNA was extracted and then transformed into EBY100 yeast strain using a large-scale high-efficiency transformation protocol[5], leading to 2×10^6 transformants. In sum, we constructed a FAST variant library containing about 10^6 FAST variants.

Choice of HPAR for screening protein tag.

HPAR was considered to be the most promising for engineering red-emitting FASTs. Its physicochemical properties in aqueous buffer and conditions mimicking environmental changes in protein cavity showed that it could be possible to preserve fluorescence activation and absorption red shift by binding with complementary protein tag (See Section III-1). Moreover, its structure shares the most similarity with that of HMBR, allowing us to envision evolving the binding properties of FAST. We showed that it was the only one among the different tested fluorogens that exhibited a binding ability albeit weak with FAST. In aqueous

buffer, its deprotonated form fluoresces weakly red light (maximal emission wavelength: 655nm), demonstrating great potential for the generation of red-emitting FASTs. It was furthermore shown to be non-toxic for mammalian cells even at high concentration (See Article 2-Table 3) and not to undergo any non-specific fluorescence activation in living cells (See Article 2-Figure 7). For all these reasons, we chose HPAR as fluorogen candidate to engineer new protein tags for the development of red-emitting FASTs.

Screening of new protein tags

We screened the yeast surface-displayed FAST library in the presence of 10 μM HPAR by high-throughput FACS. The level of protein expression on the yeast surface was normalized by immunolabeling the expressed proteins with an Alexa647-conjugated antibody before cell sorting. According to the spectral properties of HPAR in different conditions, the prospective fluorogen:protein tag complexes were expected to absorb blue-green light and fluoresce red. We thus excited yeast cells incubated in HPAR solution with 488 nm laser and detected and collected fluorogen-activating cell populations that fluoresced in the region of 620±29 nm by FACS. The fluorescence of Alexa 647 was detected in the condition of Ex 640 nm/Em 671±30 nm. For the first round of FACS we collected about 5% of the yeast population to make sure that enough positive yeast cells were collected for enriching fluorogen-activating population. We reduced gradually the collecting percentage of yeast population (second round: 1%, third round: 0.03%). From the fourth round, we observed clearly the presence of a fluorogen-activating population (Figure 3.3). We reduce the collecting percentage of population to 0.015% for accumulating fluorogen-activating population efficiently. After 5 rounds of FACS we obtained a fluorogen-activating population. Individual clones were isolated and tested by flow cytometry. Among 24 selected clones, 23 clones displayed fluorescence in the presence of HPAR. Genes of the 23 clones were extracted and sequenced. Sequencing revealed only 4 clones: clone 1 (appeared 19 times), clone 2, clone 3 (appeared 2 times) and clone 4 (Figure 3.4).

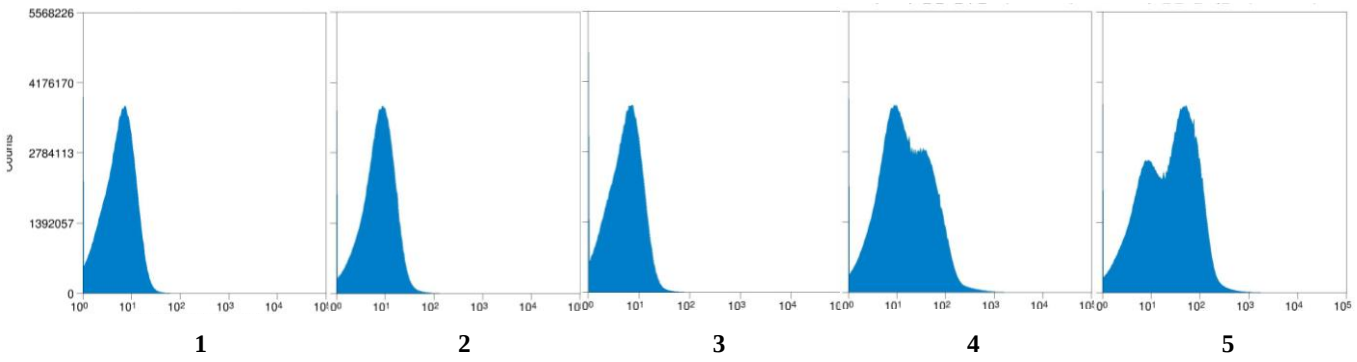


Figure 3.3 Graphic data of 5 FACS rounds for engineering protein tag in presence of HPAR.
Histograms of the distribution of population exhibiting different fluorescence intensity in the condition Ex 488 nm/Em 620±29 nm.

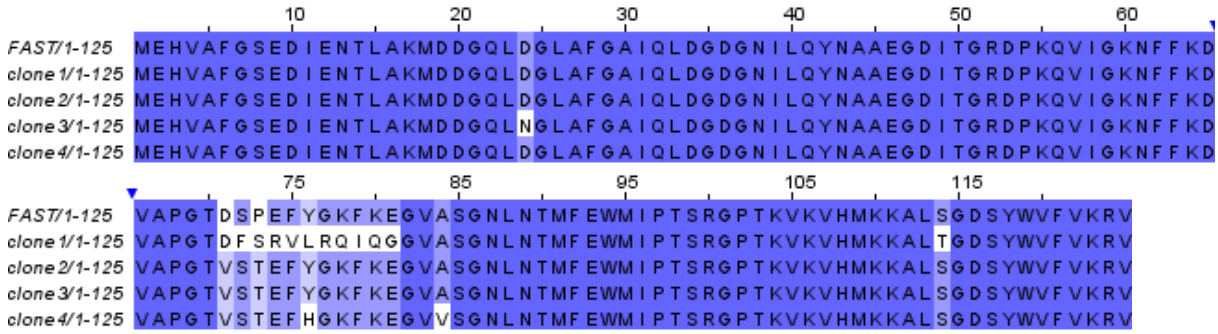


Figure 3.4 Sequence alignment of the four selected clones with FAST

Physicochemical characterization

The selected variants were expressed in *Escherichia coli* and purified by affinity chromatography. So far, we tested the spectral properties of two protein variants:HPAR complexes, with clone 1and clone 3. Unlike FAST which showed little binding affinity with HPAR, these protein variants could bind HPAR more tightly and formed red-emitting complexes preserving the features of fluorescence activation and absorption red shift upon binding (Figure 3.5, Table 3.1).

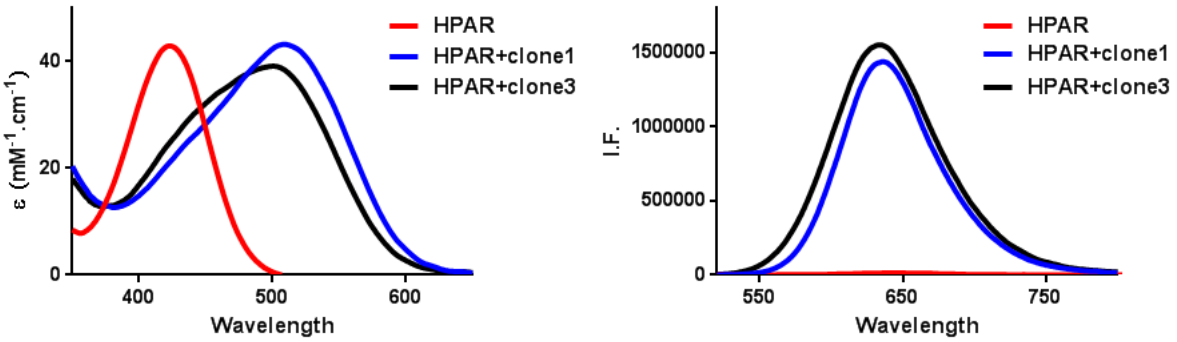


Figure 3.5 Absorption (left) and emission (right) spectra of free HPAR and protein variants:HPAR complexes in PBS pH 7.4 at 25 °C.

Complex	$\Delta\lambda_{abs}$ nm	$F_{bound}/F_{unbound}$
HPAR:FAST	75	10
HPAR:clone1	84	100
HPAR:clone3	76	100

Table 3.1 Protein variants:HPAR complexes preserve the features of fluorescence activation and absorption red shift upon binding. The abbreviations are as follows: $\Delta\lambda_{abs} = \lambda_{abs,bound} - \lambda_{abs,unbound}$, the absorption red-shift upon FAST binding; $F_{bound}/F_{unbound}$, the fluorescence activation upon FAST binding.

Compared to FAST:HBR, protein variants:HPAR complexes displayed 34-42nm red shift in absorption and a 108nm red shift in emission (Figure 3.6, Table 3.2), which verified the bathochromic effect induced by the elongation of the conjugated system. Moreover, the new complexes displayed 9-17 nm blue shift in absorption and 34-36nm red shift in emission compared to the previous red-emitting FAST:HBR-3,5DOM, generating a real red-emitting (λ_{em}~635 nm) fluorescent complexes with a huge Stokes shift (~130nm) (Figure 3.6, Table 3.2) which is attractive for multicolor imaging with a single excitation source.

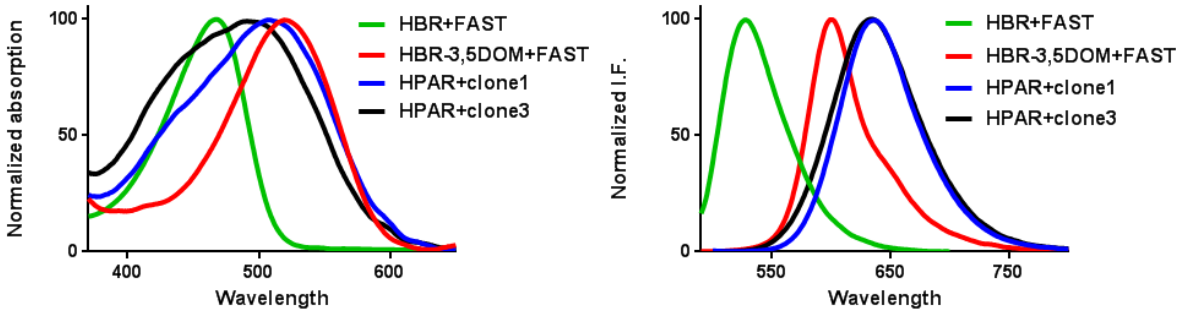


Figure 3.6 Comparison of absorption and emission of new protein variants:HPAR complexes with FAST:HBR and FAST:HBR-3,5DOM. The spectra were recorded in PBS pH7.4 at 25 °C.

Complex	λ_{abs} nm	λ_{em} nm	ε $mM^{-1}cm^{-1}$	ϕ %	Brightness	Stokes shift nm
FAST:HBR	467	527	44	9	4,000	60
FAST:HBR-3,5DOM	518	600	39	31	12,000	82
Clone1:HPAR	509	636	40	6	2,400	127
Clone3:HPAR	501	634	36	8	2,900	133

Table 3.2 Physicochemical properties protein variants:HPAR complexes in PBS pH 7.4. The abbreviations are as follows: λ_{abs}, the wavelength of maximal absorption; λ_{em}, the wavelength of maximal emission; ε, the molar absorption coefficient at λ_{abs}; ϕ, the fluorescence quantum yield. The spectroscopic data of FAST:HBR and FAST: HBR-3,5DOM are given for comparison.

III-2.3 Conclusion and perspective

We have selected 4 FAST variants activating fluorescence of HPAR from the FAST variant library. Two protein variants were purified and spectral properties of variant:HPAR complexes were characterized. Actually, redder fluorescent complexes which displayed 35nm red shift in emission compared to the previous red-emitting FAST:HBR-3,5DOM were obtained with a higher Stokes shift (~130nm), which showed great potential for multicolor imaging with a single excitation source.

The spectral differences between FAST:HBR and new red-emitting complexes proved that the strategy of elongating the conjugated system of HBR worked for extending spectral properties

to the red. We envision to engineer HPAR analogues to further shift their spectral properties to red and improve the binding affinity and brightness of these new red-emitting complexes.

The other two variants will later be expressed and then spectral properties of FAST variant:HPAR complexes will be characterized. Variants which are able to activate specifically the red fluorescence of HPAR preserving unique feature of FAST approach will be identified and compared with the previous characterized two complexes. Next the performance of the best red-emitting complex will be tested in living cells.

III-2.4 Materials and Methods

Yeast display library construction

The gene of FAST was randomly mutagenized by error-prone PCR using the GeneMorph II Random Mutagenesis Kit (Agilent). The error rate of the PCR was varied by using either 1 or 10 ng template gene. The mixed PCR product was digested with *Nhe* I and *BamH* I, and then ligated in pCTCON2 using *Nhe* I / *BamH* I restriction sites. Large-scale transformation into DH10B *E. coli* cells was performed by electroporation. The DNA was then maxiprepped and retransformed into EBY100 yeast strain using a large-scale high-efficiency transformation protocol[5].

Yeast cell sorting

Yeast library (about 1×10^9 cells) was grown overnight (30°C, 280 rpm) in 1 L of SD (20 g/L dextrose, 6.7 g/L yeast nitrogen base, 1.92 g/L yeast synthetic dropout without tryptophane, 7.44 g/L NaH_2PO_4 and 10.2 g/L Na_2HPO_4-$7H_2O$, 1% penicillin-streptomycin 10,000 U/mL). Yeast culture was diluted to OD_{600} 1 in 1L of SD and grown (30°C, 280 rpm) until OD_{600} 2-5. 5×10^9 cells yeast cells were then collected and grown for 36 h (23°C, 280 rpm) in 1L SG (20 g/L galactose, 2 g/L dextrose, 6.7 g/L yeast nitrogen base, 1.92 g/L yeast synthetic dropout without tryptophane, 7.44 g/L NaH_2PO_4 and 10.2 g/L Na_2HPO_4-$7H_2O$, 1% penicillin-streptomycin 10,000 U/mL). 5×10^8 induced cells were then pelleted by centrigugation (25°C, 3 min, 2,500 g), washed with 10 mL DPBS-BSA (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4, 1 g/L bovine serum albumin, pH 7.4), and incubated for 30 min at room temperature in 200 µL of 1/250 primary antibody chicken antic- Myc IgY (Life

Technologies) solution in DPBS-BSA. Cells were then washed with 10 mL DPBS-BSA, and incubated in 200 µL of 1/100 secondary antibody Alexa Fluor® 647–goat anti-rabbit IgG (Life Technologies) solution in DPBS-BSA for 20 min on ice. After washing with DPBS, cells were incubated in 10 mL DPBS supplemented with 10 µM HPAR, and sorted on a MoFlo™ Astrios Cell Sorter (Beckman Coulter) equipped with a 488 nm and a 640 nm laser. The sorted cells were collected in SD, grown overnight (30°C, 240 rpm) and spread on SD plates (SD supplemented with 182 g/L sorbitol, 15 g/L agar). Plates were incubated for 60 h at 30°C. The cell lawn was collected in SD supplemented with 30% glycerol, aliquoted and frozen or directly used in the next round.

Molecular Biology The plasmids pAG291 and pAG293 driving bacterial and in vitro expression of clone 1and clone 3, respectively, with an N-terminal His-tag under the control of a T7 promoter were obtained by inserting the gene encoding clone 1 or clone 3 between *Nhe* I and *Xho* I restriction sites in the pET28a vector.

Bacterial expression and protein purification Expression vectors were transformed in Rosetta(DE3)pLysS *E. coli* (New England Biolabs). Cells were grown at 37°C in Lysogeny Broth (LB) medium complemented with 50 µg/ml kanamycin and 34 µg/ml chloramphenicol to OD_{600nm} 0.6. Expression was induced for 4 h by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cells were harvested by centrifugation (4,300 × g for 20 min at 4°C) and frozen. The cell pellet was resuspended in lysis buffer (phosphate buffer 50 mM, NaCl 150 mM, $MgCl_2$ 2.5 mM, protease inhibitor, DNase, pH 7.4) and sonicated (5 min at 20 % of amplitude). The lysate was incubated for 2 h at 4 °C to allow DNA digestion by DNase. Cellular fragments were removed by centrifugation (9,300 × g for 1h at 4°C). The supernatant was incubated overnight at 4°C under gentle agitation with Ni-NTA agarose beads in phosphate buffered saline (PBS) (sodium phosphate 50 mM, NaCl 150 mM, pH 7.4) complemented with 10 mM Imidazole. Beads were washed with 20 volumes of PBS containing 20 mM Imidazole, and with 5 volumes of PBS complemented with 40 mM Imidazole. His-tagged proteins were eluted with 5 volumes of PBS complemented with 0.5 M Imidazole, followed by dialysis with PBS.

Physical Chemistry experiments pH measurements were performed on a standard pH meter PHM210 Radiometer Analytical (calibrated with aqueous buffers at pH 4 and 7 or 10) with a Crison 5208 Electrode (Barcelona, Spain). UV/Vis absorption spectra were recorded in 1 cm

× 1 cm quartz cuvettes (Hellma) on a diode array UV/Vis spectrophotometer (Evolution array, Thermo Scientific). Corrected fluorescence spectra upon one-photon excitation were recorded with a Photon Technology International QuantaMaster QM-1 spectrofluorimeter (PTI, Monmouth Junction, NJ) equipped with a Peltier cell holder (TLC50, Quantum Northwest, Shoreline, WA). The overall emission quantum yields after one-photon excitation ϕ were determined as previously described[1].

III-2.5 Reference

[1] Plamont M.-A., Billon-Denis E., Maurin S., Gauron C., Pimenta F. M., Specht C. G., Shi J., Querard J., Pan B., Rossignol J., Moncoq K., Morellet N., Volovitch M., Lescop E., Chen Y., Triller A., Vriz S., Le Saux T., Jullien L., Gautier A., Small fluorescence-activating and absorption-shifting tag for tunable protein imaging in vivo, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 497-502.

[2] Boder E.T., Wittrup K.D., Yeast surface display for screening combinatorial polypeptide libraries, *Nat. Biotech.*, 1997, **15**, 553-557.

[3] Gai S.A., Wittrup K.D., Yeast surface display for protein engineering and characterization, *Curr. Opin. Struct. Biol.*, 2007, **17**, 467-473.

[4] Chao G., Lau W.L., Hackel B.J., Sazinsky S.L., Lippow S.M., Wittrup K.D., Isolating and engineering human antibodies using yeast surface display, *Nat. Protoc.*, 2006, **1**, 755-768.

[5] Gietz R.D., Schiestl R.H., Large-scale high-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method, *Nat Protoc*, 2007, **2**, 38–41.

Chapter IV Development of cell impermeant fluorogens for the selective imaging of cell surface proteins

This part has been the object of a patent:

Membrane-impermeant fluorogenic chromophores.
Application number : EP17305591.4
A. Gautier, L. Jullien, C. Li, F. Perez

IV-1 Introduction

Cell surface proteins are important for biological research because they play indispensable roles in various biological activities. Specific labeling of cell surface proteins for studying their surface expression, transportation and function is highly required. Fluorogen-based specific labeling systems have significant potential for specifically labeling cell surface protein because one can easily adapt their methods to experimental condition by simply engineering cell impermeant fluorogens. Nowadays the discovery of small molecules able to restore trafficking defects induced by genetic mutations has become highly important in drug design[1]. Specific surface labeling methods have been applied to test the ability of a small molecule to restore normal traffic. However, spatial information of the protein of interest is usually obtained through microscopy and image analysis, limiting thus screening throughput. Detecting methods without the need for imaging, such as spectrofluorimetry or flow cytometry that allow high throughput analysis, are thus indispensable[2].

Fluorogens used in fluorogen-based specific labeling system generally cross the cell membrane and enter into the cell through a spontaneous process named diffusion. The rate of simple diffusion of a molecule across the membrane depends on its size and its hydrophobicity. Ordinarily, small hydrophobic molecules diffuse across the cell membrane easily; polar but small and uncharged molecules can also cross the membrane through diffusion; large polar molecules or ions are not cell-permeant. In sum, potential cell impermeant fluorogens should be polar, charged or large in size.

Previous fluorogens for FAST are all highly membrane-permeant. In order to extend the properties of FAST and provide investigators with the possibility of selectively detecting FAST-tagged cell-surface proteins, we sought for novel fluorogens able to form fluorescent complexes with FAST, but incapable of crossing the plasma membrane. To reduce the ability to penetrate the plasma membrane and improve cell exclusion, we considered functionalizing previous HMBR analogues with a carboxylate, a negatively charged group well known to reduce plasma membrane penetration. We have presented in Chapter II the general method to tune spectral properties of FAST approaches by engineering HMBR analogues with various substituents on their phenol rings. It was shown that aromatic substituents are one of determinants for binding with FAST and for spectral properties of formed FAST:fluorogen complexes, as they can play a major role in immobilizing the fluorogen within FAST cavity. Here we tried to develop cell impermeant fluorogens by modifying the rhodanine group of HMBR. An acetate acid was added on the NH position of the rhodanine, expected to create cell impermeant fluorogen and preserve its binding ability with FAST. At pH 7.4, the acetate acid will be deprotonated, staying as acetate with a negative charge. Moreover it has been shown that HBRAA (Figure 4.1) displayed binding affinity with FAST in the initial research for brighter FAST fluorogens[3]. The results has also shown that FAST:HBRAA complex was not as bright as desired. In this chapter, we verified that the strategy of adding negative charged acetate on the rhodanine group was feasible to generate cell impermeant fluorogen for FAST and also tried to improve the brightness of cell impermeant FASTs by fluorogen engineering through the same strategy presented in Chapter II. We present in particular HBRAA-3E exhibiting good binding affinity with FAST and forming fluorescent complex with brightness acceptable for cell imaging. This new membrane-impermeant fluoorgen has been applied to selectively image surface-exposed FAST-tagged proteins. It opens great prospects to efficiently obtain spatial information on FAST-tagged proteins by using spectrofluorimetry, which is promising for efficiently screening drugs to treat trafficking related diseases.

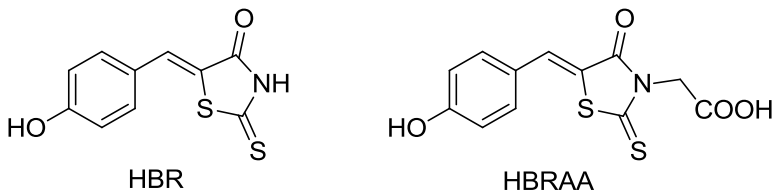


Figure 4.1 Chemical structure of HBR and HBRAA

IV-2 Results and discussion

Fluorogen design

HMBR is mainly neutral at physiological pH (the phenol pK_A is 8.7) and highly membrane-permeant[3]. To reduce its ability to penetrate the plasma membrane and improve its cell exclusion, we functionized its rhodanine head with an acetate acid group which would be deprotonated and stayed as negative charged acetate at pH 7.4. Commercially available rhodanine acetic acid was condensed with 4-hydroxy-methylbenzaldehyde to provide HBRAA-3M (Figure 4.2), an HMBR analog bearing an additional acetate acid group on the rhodanine head. The resulting FAST:HBRAA-3M complex had a similar brightness (Table 4.1) as the FAST:HBR complex (the first FAST:fluorogen complex before the development of FAST:HMBR, brightness 4000). A simple cell permeability test has been done by imaging yeast cells expressing Aga2p-FAST at the surface and HeLa cells expressing FAST-mCherry in the cytoplasm incubated with HBR or HBRAA-3M (Figure 4.3). HBR labeled not only Aga2p-FAST at the yeast cell surface but also Aga2p-FAST in the process of transportation to cell surface, while HBRAA-3M seemed only labeled Aga2p-FAST at the yeast cell surface (Figure 4.3). For HeLa cells expressing FAST in the cytoplasm, the fluorescence from FAST was only observed in the presence of HBR (Figure 4.3). Even after 30 minutes incubation with HBRAA-3M, there was no fluorescence signal from FAST (Figure 4.3), which demonstrated that HBRAA-3M could not enter into cells to label FAST expressed in the cytoplasm of HeLa cells. Theses cell experiments proved that the strategy of introducing a negative charged acetate group on the rhodanine head worked well to design cell impermeant HMBR analogues for FAST and the brightness of FAST:HBRAA-3M was reasonable for imaging FAST-tagged protein in living cells through microscopy.

We next tried to improve brightness of the membrane-impermeant FAST systems by engineering the fluorogen structure. The brightness of FAST:fluorogen complexes depend on the ability of the protein cavity to block the fluorogen in the appropriate emissive conformation. We presented previously that the substituents on the aromatic ring were key determinants for the affinity and brightness of FAST:fluorogen. We thus designed and synthesized a collection of HBRAA-3M analogues bearing different substituents on the phenol ring through condensation reaction of the starting aldehydes and rhodanine acetic acid (Figure 4.2).

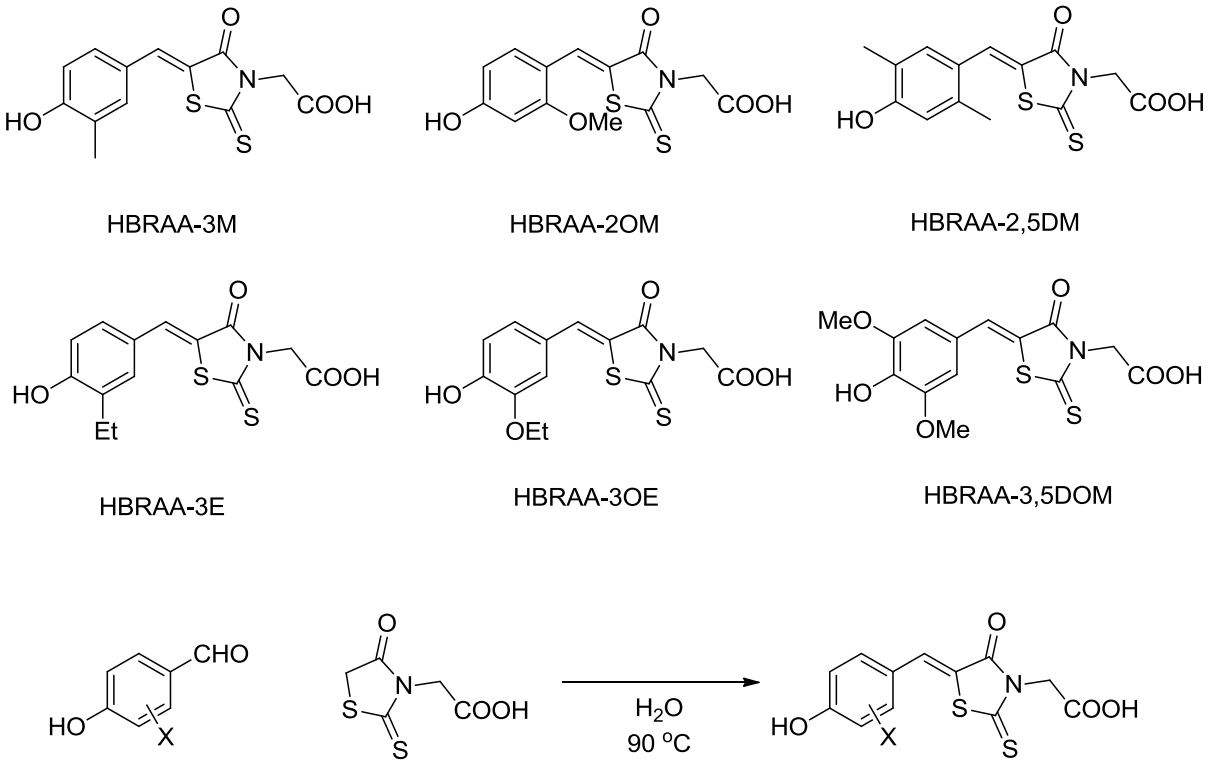


Figure 4.2 Chemical Structure of cell impermeant fluorogens for FAST and general strategy for the synthesis.

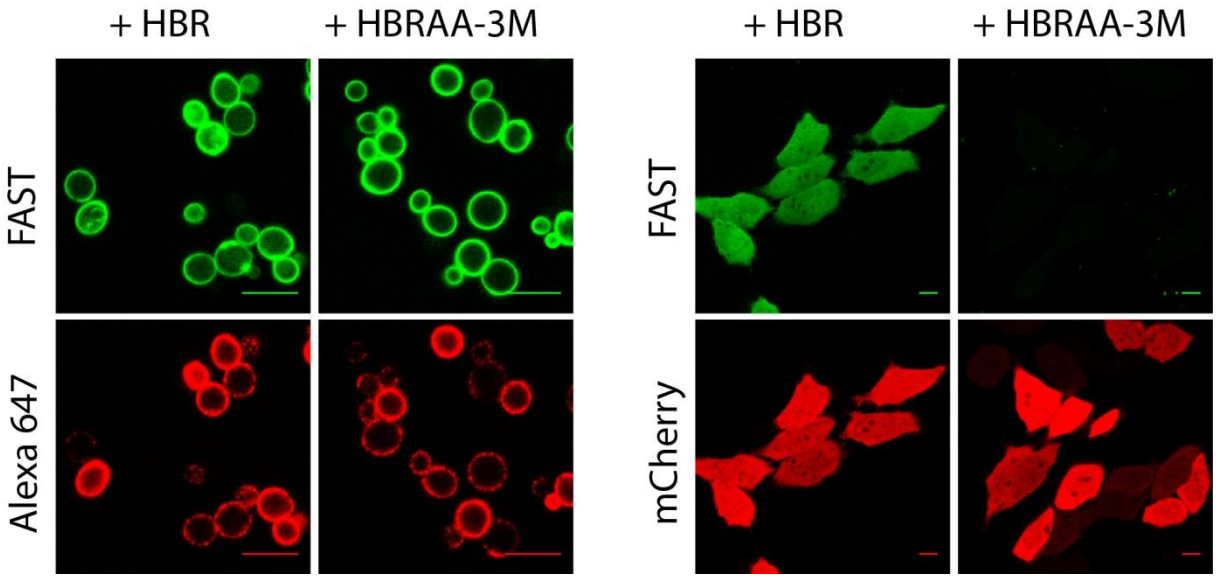


Figure 4.3 Confocal micrographs (left) of yeast cells expressing Aga2p- FAST incubated with 20 µM HBR or HBRAA-3M. The Aga2p fusion protein was immunolabeled with an Alexa647-

conjugated antibody; Confocal micrographs (right) of mammalian cells expressing mCherry-FAST in the cytoplasm labeled with 20 µM HBR or HBRAA-3M. Scale bars 10 µm.

Fluorogen engineering and physicochemical characterization

We tested the binding abilities of the membrane-impermeant fluorogens with FAST and spectral properties of free fluorogens and FAST:fluorogen complexes. Actually they displayed various binding affinities with FAST but preserving the features of fluorescence activation and absorption red shift upon binding to FAST.

Compared to HMBR, HBRAA-3M bound Y-FAST 50-fold less tightly and formed a fluorescent complex displaying a 20 nm red shift in absorption and emission with three-fold lower brightness (Table 4.1). The loss of affinity and brightness resulted very likely from steric clashes between the added acetate group and FAST.

Complex	λ_{abs} nm	λ_{em} nm	ε $mM^{-1}cm^{-1}$	ϕ %	Brightness	K_D µM
FAST:HMBR	481	540	45	23	10,300	0.13
FAST:HBRAA-3M	502	557	56	6	3,400	6.4
FAST:HBRAA-2OM	497	537	55	5	2,800	0.48
FAST:HBRAA-2,5DM	519	566	63	5	3,200	5.4
FAST:HBRAA-3E	505	559	61	8	4,900	1.3
FAST:HBRAA-3OE	525	575	62	6	3,700	1.7
FAST:HBRAA-3,5DOM	532	606				>20

Table 4.1 Physico-chemical properties of FAST:fluorogen complexes in PBS pH 7.4.
Abbreviations are as follows: λ_{abs}, wavelength of maximal absorption; λ_{em}, wavelength of maximal emission; ε, molar absorption coefficient at λ_{abs}; ϕ, fluorescence quantum yield; K_D affinity constant.

Switching the methyl group in position 3 to a methoxy group in position 2 of the phenol ring resulted in a compound HBRAA-2OM that bound FAST with higher affinity (Table 4.1). However, the resulting complex exhibited reduced fluorescence brightness (Table 4.1). On the other hand, adding an additional methyl group in meta-position of the phenolic hydroxyl group of HBRAA-3M gave a compound HBRAA-2,5DM resembling HBRAA-3M in terms of affinity and brightness (Table 4.1). Compared to FAST:HBRAA-3M, FAST:HBRAA-2,5DM exhibited a 20nm red shift in absorption and a 10nm red shift in emission (Table 4.1).

HBRAA-3,5DOM with two methoxy groups in ortho-positions of the phenolic hydroxyl group bound FAST with really low affinity (K_D > 20uM), which was not suitable for FAST. However, FAST:HBRAA-3,5DM exhibited a 50 nm red shift in absorption and a 65 nm red shift in emission compared to FAST:HMBR, absorbing green light and fluorescing red. Although it is not a good fluorogen for FAST, it displayed significant potential to engineer red-emitting membrane-impermeant FAST system.

The brightest FAST:fluorogen complex was eventually obtained with HBRAA-3E, in which the methyl was replaced by an ethyl group. HBRAA-3E bound FAST with affinities around 1 μM. The resulting complex FAST:HBRAA-3E was 50 % as bright as FAST:HMBR (Table 4.1), reaching brightness reasonable for cell imaging. Note that FAST:HBRAA-3E exhibited also a 25 nm red shift in absorption and a 20 nm red shift in emission compared to FAST:HMBR (Table 4.1). Interestingly, replacing the ethyl group by an ethoxy group (HBRAA-3OE) further red-shifted the absorption and the emission. FAST:HBRAA-3OE exhibited a 45 nm red shift in absorption and a 35 nm red shift in emission compared to FAST:HMBR.

To sum up, molecular engineering allowed us to increase the affinity and the fluorescence properties of the membrane-impermeant fluorogens. However, there is still some room for improvement both in terms of brightness and affinity. In consequence, new protein tags should be engineered in the future to optimize the system.

Labeling of surface exposed FAST-tagged proteins in live cell with HBRAA-3E

We next evaluated the ability of HBRAA-3E to label cell-surface exposed FAST-tagged proteins in living cells. We expressed transiently in HeLa cells FAST fused at the N-terminus to the murine Ig κ-chain leader sequence, which directs the protein to the secretory pathway, and at the C-terminus to the platelet derived growth factor receptor (PDGFR) transmembrane domain, which anchors the protein to the plasma membrane, displaying FAST on the extracellular side. Transfected cells were treated with media containing 5 μM of HMBR or HBRAA-3E. Confocal microscopy allowed us to show that HMBR labeled both cell-surface exposed FAST and intracellular FAST still present within the secretory pathway, while HBRAA-3E labeled selectively cell-surface proteins (Figure 4.4), in agreement with

HBRAA-3E being membrane impermeant. Addition of membrane-permeant HMBR to cells pre-labeled with HBRAA-3E allowed visualizing intracellular FAST within the secretory pathway (Figure 4.4), further demonstrating that HBRAA-3E did not cross the plasma membrane. Even after two hours incubation with media containing 5 μM HBRAA-3E, there is no intracellular labeling (Figure 4.4) observed, allowing anticipating the use of HBRAA-3E for long-term studies. Note that we also verified that long exposure to HBRAA-3E (> 24 h) was non-toxic for mammalian cells (Figure 4.5).

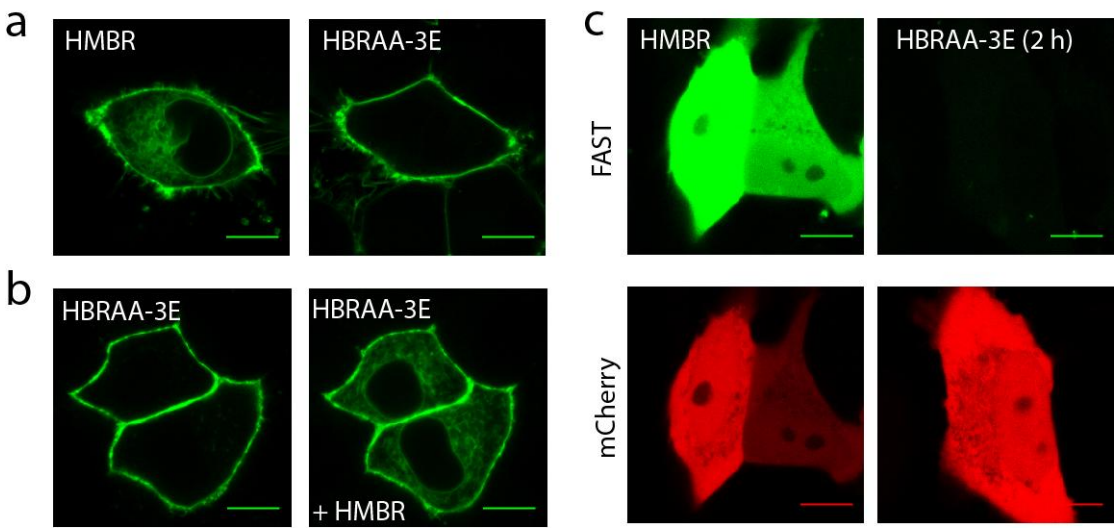


Figure 4.4 Selective labeling of cell-surface FAST-tagged proteins with membrane-impermeant fluorogens. (a-b) Confocal micrographs of HeLa cells expressing FAST on the extracellular side. (a) Cells were imaged after 5 minutes of treatment with 5 μM HMBR or HBRAA-3E. While HMBR revealed membrane and intracellular proteins, HBRAA-3E revealed only surface-exposed proteins. (b) Subsequent addition of cell permeant HMBR and immediate imaging (< 10 s) revealed proteins within the secretory pathway. (c) Confocal micrographs of HeLa cells expressing FAST fused to the red fluorescent protein mCherry treated with 5 μM HMBR for 15 seconds or 5 μM HBRAA-3E for 2 hours. Side-by-side images were recorded using the same settings for direct comparison of the fluorescence intensities. (a-c) Scale bars 10 μm.

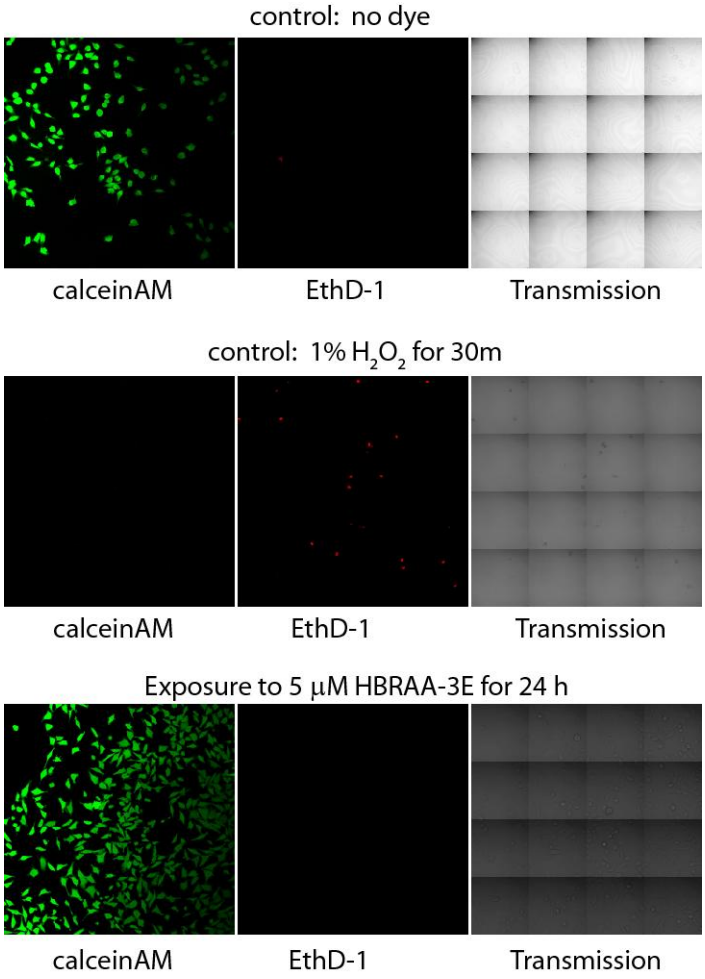


Figure 4.5 Viability assay of HeLa cells incubated for 24 h with solutions of HBRAA-3E at 5 μM. Cell viability was tested by using calceinAM and EthD1 (LIVE/DEAD® viability/cytotoxicity assay kit). CalceinAM is a cell-permeant profluorophore cleaved by intracellular esterases releasing the green fluorescent polyanionic calcein in live cells. EthD1 (Ethidium homodimer 1) is a non cell-permeant nucleic acid red fluorescent stain that enters only cells with damaged membranes and undergoes a fluorescence enhancement upon binding to nucleic acids, thereby producing a bright red fluorescence in dead cells. Control experiments with HeLa cells non-incubated with dye (the first row) or incubated for 30 min with 1% hydrogen peroxyde (the second row) are shown. Cell fluorescence was evaluated by confocal microscopy. The experiment (in the last row) shows that HBRAA-3E are non-toxic for HeLa cells at the concentrations used for imaging.

Screening of new protein tags for HBRAA-3,5DOM

The red-shifted spectral properties of HBRAA-3,5DOM inspired us to engineer an optimal protein tag for HBRAA-3,5DOM for multi-color cell-surface protein labeling. With the same

methodology used in Section III-2, we selected FAST variants that could tightly bind with HBRAA-3,5DOM and form complexes displaying bright red emission from the library of FAST variants (See Section III-2). We screened our library of yeast surface-displayed FAST variants in the presence of 5 μM HBRAA-3,5DOM by high throughput FACS (Figure 4.6). We selected yeast cells incubated with HBRAA-3,5DOM in the condition of Ex 561 nm/Em 614±20 nm by FACS. The fluorescence of Alexa 647 was detected in the condition of Ex 640 nm/Em 671±30 nm. For the first round of FACS we collected about 5% yeast population to make sure that we collected enough positive yeast cells for enriching fluorogen-activating population. We reduced gradually the collecting percentage of yeast population (second round: 4%, third round: 0.4%). From the fourth round, we observed the existing of a fluorogen-activating cell population. We reduce the collecting percentage of cell population to 0.05% for accumulating fluorogen-activating population efficiently. After 5 FACS rounds we obtained a fluorogen-activating cell population. Individual clones were isolated and tested by flow cytometry. Among 24 selected clones, 22 clones displayed fluorescence activation in the presence of HBRAA-3,5DOM. The genes of the 22 clones were extracted and sequenced. The sequencing revealed five single clones: clone 1 (appeared 16 times) clone 2, clone 3 (appeared 2 times) clone 4 and clone 5 (appeared 2 times) (Figure 4.7).

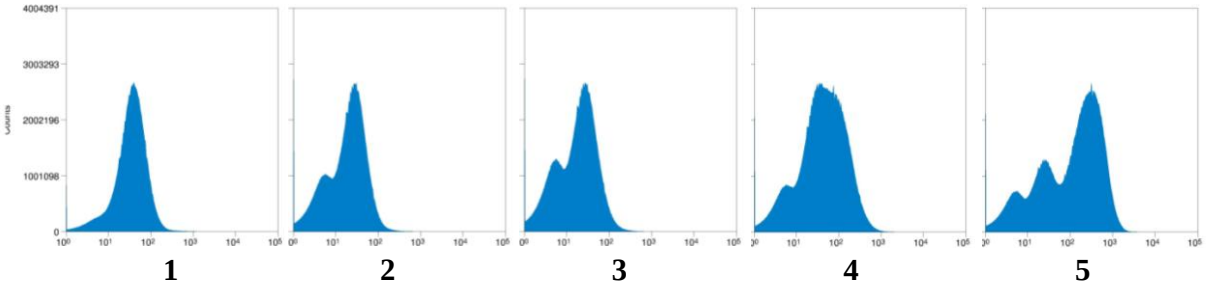


Figure 4.6 Graphic data of 5 FACS rounds for engineering protein tag in presence of HBRAA-3,5DOM. Histograms of the distribution of population exhibiting different fluorescence intensity in the condition Ex 561 nm/Em 614±20 nm.

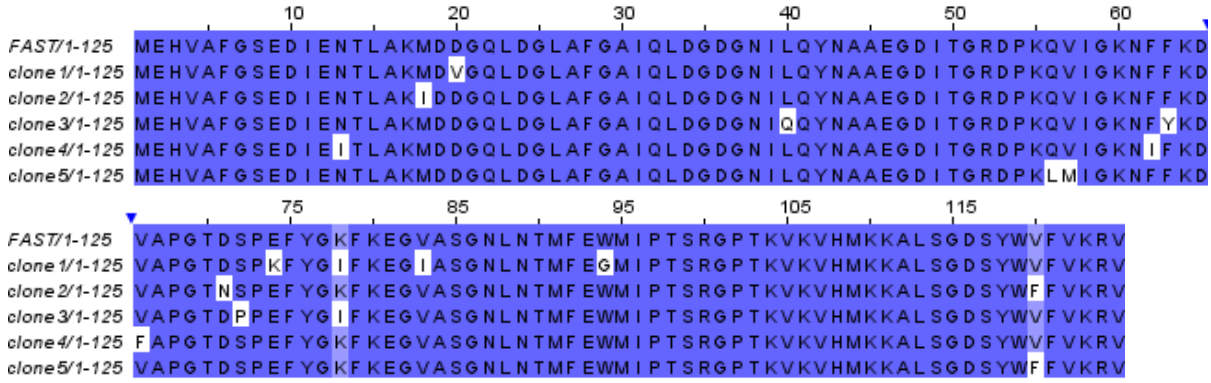


Figure 4.7 Sequence alignment of the five selected clones with FAST

Physicochemical characterization

The selected variants were expressed in *Escherichia coli* and purified by affinity chromatography. So far, we tested the spectral properties of three protein variants:HBRAA-3,5DOM complexes, including clone1, 2 and 5. Unlike FAST, these protein variants seemed to be capable of binding tightly with HBRAA-3,5DOM and formed red-emitting complexes preserving the features of fluorescence activation and absorption red shift upon binding (Figure 4.8, Table 4.2).

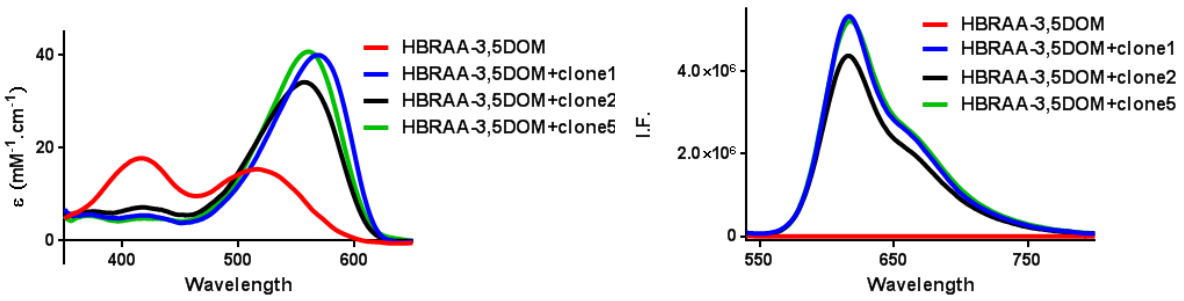


Figure 4.8 Absorption (left) and emission (right) spectra of free HBRAA-3,5DOM and protein variants:HBRAA-3,5DOM complexes in PBS pH 7.4 at 25 °C.

Complex	$\Delta\lambda_{abs}$ nm	$F_{bound}/F_{unbound}$
HBRAA-3,5DOM:FAST	116	5
HBRAA-3,5DOM:clone1	150	200
HBRAA-3,5DOM:clone2	140	200
HBRAA-3,5DOM:clone5	144	230

Table 4.2 Protein variants:HBRAA-3,5DOM complexes preserve the features of fluorescence activation and absorption red shift upon binding. The abbreviations are as follows: $\Delta\lambda_{abs}= \lambda_{abs,bound} - \lambda_{abs,unbound}$, the absorption red-shift upon FAST binding; $F_{bound}/F_{unbound}$, the fluorescence activation upon FAST binding.

Compared to FAST:HBRAA-3E, protein variants:HBRAA-3,5DOM complexes displayed 51-61nm red shift in absorption and a 56nm red shift in emission with improved brightness (Figure 4.9, Table 4.3), which generated a new color for the fluorescence imaging of membrane- impermeant FAST systems.

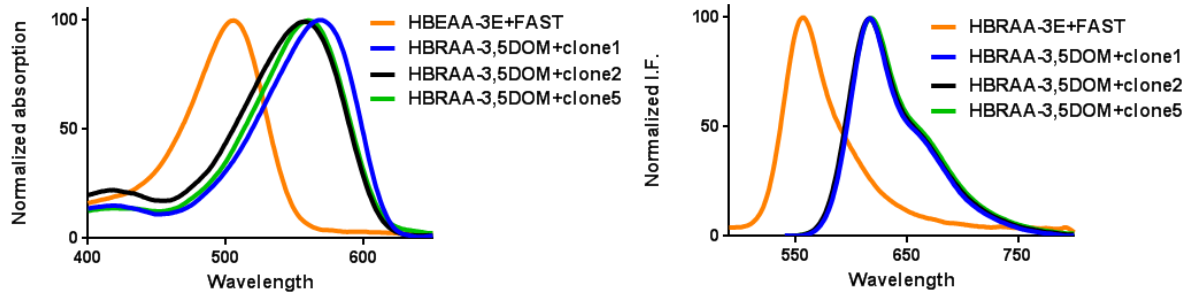


Figure 4.9 Comparison of absorption and emission of new protein variants: HBRAA-3,5DOM complexes with FAST: HBRAA-3E. The spectra were recorded in PBS pH7.4 at 25 °C.

Complex	λ_{abs} nm	λ_{em} nm	ε $mM^{-1}cm^{-1}$	ϕ %	Brightness
FAST: HBRAA-3E	505	559	61	8	4,900
Clone1: HBRAA-3,5DOM	566	615	40	17	6,800
Clone2: HBRAA-3,5DOM	556	614	35	13	4,500
Clone5: HBRAA-3,5DOM	560	616	41	15	6,100

Table 4.3 Physicochemical properties protein variants:HBRAA-3,5DOM complexes in PBS pH 7.4. The abbreviations are as follows: λ_{abs}, the wavelength of maximal absorption; λ_{em}, the wavelength

of maximal emission; ε, the molar absorption coefficient at λ_{abs}; ϕ, the fluorescence quantum yield. The spectroscopic data of FAST: HBRAA-3E is given for comparison.

IV-3 Conclusion and perspective

In conclusion, we developed a collection of membrane-impermeant fluorogens that bind FAST and give fluorescent complex with various spectral properties. Their inability to cross the plasma membrane enables to selectively image surface exposed FAST-tagged proteins, without labeling intracellular proteins present in the secretory pathway. It shows potential to obtain spatial information on FAST-tagged proteins without the need of microscopy and image analysis by using fluorescence-based techniques such as spectrofluorimetry and flow cytometry, opening great prospects for the design of high throughput screening assays for the discovery of drugs to treat trafficking related diseases.

The promising applications of cell-impermant FAST approach motivate us to extend its spectral properties. We are selecting FAST variants that could bind new cell impermeant fluorogens forming fluorescent complexes with different colors and improved brightness from the library of FAST variants. After 5 rounds of FACS, 5 FAST variants displaying fluorogen-activation in the presence of HBRAA-3,5DOM have been selected. Three protein variants were purified and spectral properties of variant:HPAR complexes were characterized. Actually, bright red-emitting membrane-impermeant complexes which displayed a ~55nm red shift in absorption (λ_{abs} ~560 nm) and a 56 nm red shift in emission (λ_{em} ~615 nm) compared to the previous FAST:HBRAA-3E were obtained, which made the application of cell-impermeant FAST systems more versatile.

The other two variants will later be expressed and then spectral properties of FAST variant: HBRAA-3,5DOM complexes will be characterized. Variants which are able to activate specifically the red fluorescence of HBRAA-3,5DOM preserving unique feature of FAST approach will be identified and compared with the three previous characterized complexes. Next the performance of the best red-emitting cell-impermeant complex will be tested in living cells.

IV-4 Materials and Methods

Chemical synthesis — general information Commercially available reagents were used as starting materials without further purification. NMR spectra were recorded on a AC Bruker spectrometer at 300 MHz for ^{1}H and 75 MHz for ^{13}C; chemical shifts are reported in ppm with protonated solvent as internal reference ^{1}H, $CHCl_3$ in $CDCl_3$ 7.26 ppm, CHD_2SOCD_3 in CD_3SOCD_3 2.50 ppm; ^{13}C, $^{13}CDCl_3$ in $CDCl_3$ 77.0 ppm, $^{13}CD_3SOCD_3$ in CD_3SOCD_3 39.52 ppm; coupling constants *J* are given in Hz. Mass spectra (chemical ionization or high resolution) were performed by the Service de Spectrométrie de Masse de Chimie ParisTech and the Institut de Chimie Organique et Analytique de l'Université d'ORléans. Column chromatography was performed on silica gel 60 (0.040-0.063 nm) Merck. Analytical thin-layer chromatography (TLC) was conducted on Merck silica gel 60 F254 precoated plates.

Synthesis of 4-hydroxy-3-ethylbenzaldehyde To a solution of 2-ethylphenol (6.1 g, 50 mmol) in 10% aqueous sodium hydroxide (80 mL, 200 mmol) was added trichloromethane (15.0 g, 125 mmol) dropwise at 60°C over 1 h, and then the reaction mixture was stirred for 2 h at 60 °C. After cooling, the mixture was neutralized by an aqueous solution of hydrochloric acid and extracted with dichloromethane. The combined organic layers were washed with brine, dried over magnesium sulfate, and concentrated under reduced pressure. The residue was purified by flash chromatography on silica gel with cyclohexane/ethylacetate (7.5/2.5, v/v) to yield the desired 4-hydroxy-3-ethylbenzaldehyde (1.1 g, 15% yield) as a gray pink solid. ^{1}H NMR (300 MHz, $CDCl_3$, δ in ppm): 9.84 (s, 1H), 7.72 (s, 1H), 7.65 (d, *J* = 8.1 Hz, 1H), 6.92 (d, *J* = 8.1 Hz, 1H), 6.51 (s, 1H), 2.70 (q, *J* = 7.5 Hz, 2H), 1.27 (t, *J* = 7.5 Hz, 3H); ^{13}C NMR (75 MHz, $CDCl_3$, δ in ppm): 192.3, 160.5, 131.5, 131.4, 130.6, 129.4, 115.6, 22.8, 13.5; MS (ESI): m/z 149.2$[M-H]^-$, calcd mass for $[C_9H_9O_2]^-$: 149.1; HRMS (ESI): m/z 149.0608 $[M-H]^-$, calcd mass for $[C_9H_9O_2]^-$: 149.0603.

General protocol for the synthesis of the membrane-impermeant fluorogens A solution containing rhodanine-3-acetic acid (191 mg, 1.0 mmol) and the substituted 4-hydroxy-benzaldehyde (1.0 mmol) in 40 mL of water was stirred at 90°C for 7 days. After cooling to 4°C and standing overnight, the precipitate was filtered through a glass filter and the crude solid was washed with water, ethanol and dried over P_2O_5, to give the desired product.

(Z)-2-(5-(4-hydroxy-3-methylbenzylidene)-4-oxo-2-thioxothiazolidin-3-yl)acetic acid

(HBRAA-3M) Yellow powder (80 %). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 10.50 (s, 1H), 7.74 (s, 1H), 7.40 (s, 1H), 7.38 (d, *J* = 8.1 Hz, 1H), 6.96 (d, *J* = 8.1 Hz, 1H), 4.72 (s, 2H), 2.18 (s, 3H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 193.2, 167.4, 166.5, 159.3, 134.9, 134.1, 131.1, 125.7, 123.8, 116.8, 115.8, 45.0, 15.9; MS (ESI): m/z 308.2$[M-H]^-$, calcd mass for $[C_{13}H_{10}NO_4S_2]^-$: 308.0; HRMS (ESI): m/z 310.0202 $[M+H]^+$, calcd mass for $[C_{13}H_{12}NO_4S_2]^+$: 310.0208.

(Z)-2-(5-(4-hydroxy-2-methoxybenzylidene)-4-oxo-2-thioxothiazolidin-3-yl)acetic acid (HBRAA-2OM) Orange powder (74 %). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 10.68 (s, 1H), 7.92 (s, 1H), 7.33 (d, *J* = 8.7 Hz, 1H), 6.57 (dd, *J* = 8.4, 2.1 Hz, 1H), 6.53 (d, *J* = 2.1 Hz, 1H), 4.71 (s, 2H), 3.87 (s, 3H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 193.6, 167.4, 166.7, 163.4, 160.6, 132.6, 129.9, 116.6, 112.8, 109.2, 99.4, 55.7, 45.0; MS (ESI): m/z 324.2$[M-H]^-$, calcd mass for $[C_{13}H_{10}NO_5S_2]^-$: 324.0; HRMS (ESI): m/z 324.0007 $[M-H]^-$, calcd mass for $[C_{13}H_{10}NO_5S_2]^-$: 324.0000.

(Z)-2-(5-(4-hydroxy-2, 5-dimethylbenzylidene)-4-oxo-2-thioxothiazolidin-3-yl) acetic acid (HBRAA-2,5DM) Orange powder (59 %).^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 13.43(s, 1H), 10.36 (s, 1H), 7.86 (s, 1H), 7.17 (s, 1H), 6.77 (s, 1H), 4.72 (s, 2H), 2.36 (s, 3H), 2.15 (s, 3H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 193.5, 167.4, 166.4, 159.0, 140.2, 131.8, 130.9, 123.3, 122.4, 117.8, 117.4, 44.9, 19.2, 15.6; MS (ESI): m/z 322.2 $[M-H]^-$, calcd mass for $[C_{14}H_{12}NO_4S_2]^-$: 322.0; HRMS (ESI): m/z 322.0214 $[M-H]^-$, calcd mass for $[C_{14}H_{12}NO_4S_2]^-$: 322.0208.

(Z)-2-(5-(4-hydroxy-3-ethylbenzylidene)-4-oxo-2-thioxothiazolidin-3-yl)acetic acid (HBRAA-3E) Yellow powder (62 %). ^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 13.44 (s, 1H), 10.49 (s, 1H), 7.77 (s, 1H), 7.41 (s, 1H), 7.39 (d, *J* = 7.8 Hz, 1H), 6.97 (d, *J* = 7.8 Hz, 1H), 4.72 (s, 2H), 2.59 (q, *J* = 7.2 Hz, 2H), 1.17 (t, *J* = 7.2 Hz, 3H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm): 193.1, 167.4, 166.5, 158.9, 135.0, 132.5, 131.5, 131.0, 123.9, 116.8, 116.0, 45.0, 22.5, 13.7; MS (ESI): m/z 322.2 $[M-H]^-$, calcd mass for $[C_{14}H_{12}NO_4S_2]^-$: 322.0; HRMS (ESI): m/z 322.0212 $[M-H]^-$, calcd mass for $[C_{14}H_{12}NO_4S_2]^-$: 322.0208.

(Z)-2-(5-(4-hydroxy-3-ethoxybenzylidene)-4-oxo-2-thioxothiazolidin-3-yl)acetic acid (HBRAA-3OE) Yellow powder (59 %).^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 13.43 (s, 1H), 10.13 (s, 1H), 7.79 (s, 1H), 7.20 (s, 1H), 7.16 (d, *J* = 8.4 Hz, 1H), 6.97 (d, *J* = 8.4 Hz,

1H), 4.73 (s, 2H), 4.10 (q, *J* = 6.9 Hz, 2H), 1.37 (t, *J* = 6.9Hz, 3H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm) : 193.0, 167.3, 166.4, 150.7, 147.3, 135.0, 125.5, 124.3, 117.2, 116.5, 115.7, 64.0, 45.0, 14.6; MS (ESI): m/z 338.3 $[M-H]^-$, calcd mass for $[C_{14}H_{12}NO_5S_2]^-$: 338.0; HRMS (ESI): m/z 338.0162 $[M-H]^-$, calcd mass for $[C_{14}H_{12}NO_5S_2]^-$: 338.0157.

(Z)-2-(5-(4-hydroxy-3,5-dimethoxybenzylidene)-4-oxo-2-thioxothiazolidin-3-yl)acetic acid (HBRAA-3,5DOM) Orange powder (59 %).^{1}H NMR (300 MHz, CD_3SOCD_3, δ in ppm): 9.63 (s, 1H), 7.82 (s, 1H), 6.97 (s, 2H), 4.73 (s, 2H), 3.85 (s, 6H); ^{13}C NMR (75 MHz, CD_3SOCD_3, δ in ppm) : 193.4, 167.8, 166.8, 148.8(2C), 140.2, 135.7, 123.6, 118.1, 109.3(2C), 56.6(2C), 45.4; MS (ESI): m/z 354.5 $[M-H]^-$, calcd mass for $[C_{14}H_{12}NO_6S_2]^-$: 354.0.

Physical Chemistry experiments pH measurements were performed on a standard pH meter PHM210 Radiometer Analytical (calibrated with aqueous buffers at pH 4 and 7 or 10) with a Crison 5208 Electrode (Barcelona, Spain). UV/Vis absorption spectra were recorded in 1 cm × 1 cm quartz cuvettes (Hellma) on a diode array UV/Vis spectrophotometer (Evolution array, Thermo Scientific). Corrected fluorescence spectra upon one-photon excitation were recorded with a Photon Technology International QuantaMaster QM-1 spectrofluorimeter (PTI, Monmouth Junction, NJ) equipped with a Peltier cell holder (TLC50, Quantum Northwest, Shoreline, WA). The overall emission quantum yields after one-photon excitation ϕ were determined as previously described[3]. Affinity constants were determined by spectrofluorometric titration using a Spark 10M plate reader (Tecan) following protocols previously described[3].

Molecular Biology The plasmid pAG211 for mammalian expression of FAST (fused also to mCherry as additional transfection marker) at the outer plasma membrane was obtained by inserting the sequence coding the protein of interest within the pDisplay plasmid (TermoFisher) using *Sal* I and *Bgl* II restriction sites. The plasmid pAG96 for cytoplasmic expression of FAST fused to mCherry and the plasmid pAG 87 for bacterial expression of FAST were previously reported[3]. The plasmids pAG286, pAG287 and pAG290 driving bacterial and in vitro expression of clone 1, clone 2 and clone 5, respectively, with an N-terminal His-tag under the control of a T7 promoter were obtained by inserting the gene encoding clone 1, clone 2 or clone 5 between *Nhe* I and *Xho* I restriction sites in the pET28a vector.

Bacterial expression and protein purification Expression vectors were transformed in Rosetta(DE3)pLysS *E. coli* (New England Biolabs). Cells were grown at 37°C in Lysogeny Broth (LB) medium complemented with 50 µg/ml kanamycin and 34 µg/ml chloramphenicol to OD_{600nm} 0.6. Expression was induced for 4 h by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cells were harvested by centrifugation (4,300 × g for 15 min at 4°C) and frozen. The cell pellet was resuspended in lysis buffer (phosphate buffer 50 mM, NaCl 150 mM, $MgCl_2$ 2.5 mM, protease inhibitor, DNase, pH 7.4) and sonicated (5 min at 20 % of amplitude). The lysate was incubated for 2 h at 4 °C to allow DNA digestion by DNase. Cellular fragments were removed by centrifugation (9,300 × g for 1h at 4°C). The supernatant was incubated overnight at 4°C under gentle agitation with Ni-NTA agarose beads in phosphate buffered saline (PBS) (sodium phosphate 50 mM, NaCl 150 mM, pH 7.4) complemented with 10 mM Imidazole. Beads were washed with 20 volumes of PBS containing 20 mM Imidazole, and with 5 volumes of PBS complemented with 40 mM Imidazole. His-tagged proteins were eluted with 5 volumes of PBS complemented with 0.5 M Imidazole, followed by dialysis with PBS.

Mammalian Cell Culture HeLa cells were cultured in DMEM supplemented with phenol red, Glutamax I, 10% (vol/vol) fetal calf serum (FCS), and 1% penicillin-streptomycin at 37 °C within a 5% CO_2 atmosphere. For microscopic imaging, cells were seeded in µDish or µSlide IBIDI (Biovalley) coated with poly-L-lysine. Cells were transiently transfected using Genejuice (Merck) according to the manufacturer's protocol. Before imaging, cells were washed with PBS, and treated with DMEM media (without serum and phenol red) containing the fluorogens at the indicated concentration. Cells were imaged directly without washing.

Cell viability assay HeLa cells were treated with DMEM media containing the fluorogens at the indicated concentrations for the indicated times. Cell viability was evaluated by fluorescence microscopy using the LIVE/DEAD® viability/cytotoxicity assay kit (Molecular Probes, Life Technologies) following the manufacter's protocol.

Fluorescence Analysis Confocal micrographs were acquired on a Zeiss LSM 710 Laser Scanning Microscope equipped with a Plan Apochromat 63×/ 1.4 NA oil immersion objective. ZEN software was used to collect the data. Images were analyzed with Image J. Flow cytometry analyses were performed on an Accuri C6 cytometer (BD Biosciences).

Yeast cell sorting Yeast library (about 1×10^9 cells) was grown overnight (30°C, 280 rpm) in 1 L of SD (20 g/L dextrose, 6.7 g/L yeast nitrogen base, 1.92 g/L yeast synthetic dropout without tryptophane, 7.44 g/L NaH_2PO_4 and 10.2 g/L Na_2HPO_4-$7H_2O$, 1% penicillin-streptomycin 10,000 U/mL). Yeast culture was diluted to OD_{600} 1 in 1L of SD and grown (30°C, 280 rpm) until OD_{600} 2-5. 5×10^9 cells yeast cells were then collected and grown for 36 h (23°C, 280 rpm) in 1L SG (20 g/L galactose, 2 g/L dextrose, 6.7 g/L yeast nitrogen base, 1.92 g/L yeast synthetic dropout without tryptophane, 7.44 g/L NaH_2PO_4 and 10.2 g/L Na_2HPO_4-$7H_2O$, 1% penicillin-streptomycin 10,000 U/mL). 5×10^8 induced cells were then pelleted by centrigugation (25°C, 3 min, 2,500 g), washed with 10 mL DPBS-BSA (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4, 1 g/L bovine serum albumin, pH 7.4), and incubated for 30 min at room temperature in 200 µL of 1/250 primary antibody chicken antic- Myc IgY (Life Technologies) solution in DPBS-BSA. Cells were then washed with 10 mL DPBS-BSA, and incubated in 200 µL of 1/100 secondary antibody Alexa Fluor® 647–goat anti-rabbit IgG (Life Technologies) solution in DPBS-BSA for 20 min on ice. After washing with DPBS, cells were incubated in 10 mL DPBS supplemented with HBRAA-3,5DOM, and sorted on a MoFlo™ Astrios Cell Sorter (Beckman Coulter) equipped with a 488 nm and a 640 nm laser. The sorted cells were collected in SD, grown overnight (30°C, 240 rpm) and spread on SD plates (SD supplemented with 182 g/L sorbitol, 15 g/L agar). Plates were incubated for 60 h at 30°C. The cell lawn was collected in SD supplemented with 30% glycerol, aliquoted and frozen or directly used in the next round.

IV-5 Reference

[1] Hung M.C., Link W., Protein localization in disease and therapy, *J. Cell Sci.*, 2011, **124**, 3381-3392.

[2] Chia P.Z.C., Ramdzan Y.M., Houghton F.J., Hatters D.M., Gleeson P.A., *Traffic*, 2014, **15**, 572-582.

[3] Plamont M.-A., Billon-Denis E., Maurin S., Gauron C., Pimenta F. M., Specht C. G., Shi J., Querard J., Pan B., Rossignol J., Moncoq K., Morellet N., Volovitch M., Lescop E., Chen Y., Triller A., Vriz S., Le Saux T., Jullien L., Gautier A., Small fluorescence-activating and

absorption-shifting tag for tunable protein imaging in vivo, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 497-502.

Chapter V General discussion

Fluorogen-based systems for specific protein labeling play a more and more significant role in live cell imaging because they enable one to achieve high imaging contrast and to adapt experimental requirement by simple chemical modification of the fluorogens. The development of FAST has expanded the application of fluorogen-based specific labeling systems with a lot of impressive advantages. A unique feature of FAST is its highly tunable labeling. Its fluorescence can be switched on or off rapidly by addition or removal of the fluorogen, which opens new opportunities for multiplexed imaging. My PhD aimed to enlarge the FAST family by developing novel systems with different properties for multiplexed imaging. This thesis presents general strategies used for engineering hybrid systems composed of a fluorogenic chromophore and a protein tag. As this type of system contains two parts, the molecule and the genetically encoded tag, the engineering of new systems obviously has two directions, chemical modification of the molecule and engineering of the tag. For chemical modification, one can change functional groups to tune its spectral properties or modify it with specific groups to achieve new functional properties.

V-1 Development of new fluorogens

The fluorogen HMBR for FAST is based on a push-pull system. An electron-donating phenol ring contacts with the electron-withdrawing rhodanine ring through a conjugated system. It deexcites non-radiatively in solution but relaxes radiatively to the ground state in FAST protein cavity. Based on this push-pull system, we have developed several series of HMBR analogues by changing substituents on the phenol ring, by elongating the conjugated system or by modifying the rhodanine head to design novel fluorogens suitable for protein labeling. I have explored the role of substituents on the phenol ring of HMBR for varying spectral and affinity properties of FAST:fluorogen complexes. Aromatic substituents are a priori key determinants of the fluorescence properties, as they can play a major role in tuning

absorption/emission properties and immobilizing the fluorogen within FAST cavity. New fluorogens preserving the binding ability with FAST and displaying different spectral properties are reported. Notably, orange-emitting FAST:HBR-3,5DM is 2.5-fold brighter than FAST-HMBR, reaching the brightness level of enhanced green fluorescent protein. Furthermore, a red-emitting FAST:fluorogen complex displaying comparable brightness as the red fluorescent protein mCherry was obtained using HBR-3,5DOM as fluorogen. It expands the application of FAST approach because one can adapt the color of FAST to different spectral conditions. Particularly, the possibility to perform dynamic color switching (red-green) by changing the fluorogen solution (HBR-3,5DOM - HMBR) opens a new dimension for multiplexed imaging. One can use the kinetic signature provided by the color exchange to extract the signal of FAST in the presence of both green and red fluorescent proteins, even though spectral crowding prevents spectral discrimination.

As the red-emitting fluorescent reporters interests biologists specially, I next elongated the conjugation system of the fluorogen by adding an additional double bound between the phenol ring and the electron-withdrawing heterocycle to develop far-red-emitting FAST systems. As their relatively large size prevented their binding to FAST, I selected FAST variants able to bind HPAR (one of the far-red fluorogens) and activated its far-red fluorescence (*vide infra*). New protein tags that can bind tightly with new red-emitting fluorogens and specifically activate their fluorescence are thus required.

Finally, I developed cell-impermeant fluorogens by adding a negative charged acetate group in the rhodanine head of HMBR. After additional engineering of the phenol group, HBRAA-3E was chosen as the fluorogen for selective imaging of FAST-tagged cell-surface proteins in living cells because its superior fluorogenic performance and affinity. HBRAA-3E enabled to selectively image surface exposed FAST-tagged proteins, without labeling intracellular proteins present in the secretory pathway. This approach displays the potential to obtain spatial information on FAST-tagged proteins by using fluorescence-based techniques such as spectrofluorimetry and flow cytometry, opening great prospects for the design of high

throughput screening assays for the discovery of drugs to treat trafficking related diseases.

V-2 Selection of new protein tags

We used directed evolution as the strategy for selecting new protein tags. Yeast surface display and FACS were applied to achieve high throughput selection. The possible protein scaffolds should contain reasonable cavities for binding fluorogens and activating their fluorescence. The protein library applied in this thesis for selecting new protein tag was a library of FAST variants constructed by random mutagenesis through error-prone PCR. This protein scaffold obviously shows binding affinity with HMBR analogues and was perfect to engineer variants with improved properties. From this library, I have selected novel protein tags activating the fluorescence of the far-red emitting HPAR and the membrane-impermeant red-emitting HBRAA-3,5DOM.

The described library of FAST variants library brings a lot of possibilities for developping new FAST systems. In addition to the two applications mentioned above, protein engineering for improving the spectral and binding properties of FAST:HMBR, FAST:HBR-3,5DOM and FAST:HBRAA-3E is already on-going in the lab. Multiplexed imaging motivates in particular the engineering of FAST variants displaying orthogonal fluorogen binding specificities to achieve orthogonal labeling.

In conclusion, with (i) its unique advantages for live cell imaging, (ii) the possibility of easily engineering new fluorogens through molecular engineering, and (iii) the robustness of the protein evolution protocol, FAST has become a really promising fluorogen-based labeling system, opening new opportunities for various applications in biological imaging, including multiplexed imaging and cellular biosensing.

Résumé

L'étude de la dynamique des protéines est essentielle pour comprendre les processus biologiques. Notre laboratoire a développé une nouvelle classe de protéines fluorescentes semi-synthétiques, appelée *Fluorescence-Activating and absorption-Shifting Tag* (FAST). Cette thèse de doctorat présente le développement de nouveaux systèmes FAST avec diverses propriétés pour l'imagerie multiplexée. Nous avons développé une série de fluorogènes permettant de modifier la couleur de FAST de vert-jaune à orange et rouge. Au delà de l'application de l'imagerie multi-couleurs, ces fluorogènes permettant un échange dynamique des couleurs grâce à la liaison réversible de FAST, ouvrant de nouvelles perspectives pour le développement de méthodes d'imagerie sélective reposant sur la dynamique de systèmes réactifs. Pour étendre davantage les propriétés spectrales de FAST vers le rouge lointain, nous avons développé une nouvelle série de fluorogènes rouges, pour lesquels nous avons sélectionné par une stratégie d'évolution dirigée basée sur le yeast display et la cytométrie en flux de nouveaux tags protéiques capables d'interagir avec ces fluorogènes et d'activer leur fluorescence. Nous avons enfin développé de nouveaux fluorogènes capables de former des complexes fluorescents avec FAST, mais incapables de traverser la membrane plasmique, ce qui permet de détecter sélectivement les protéines membranaires.

Mots Clés

marqueur fluorogénique, fluorescence, marquage protéique, imagerie biologique

Abstract

Studying protein activities could help us to understand the complex mechanisms controlling cells and organisms. Our laboratory recently developed Fluorescence-Activating and absorption-Shifting Tag (FAST), a small fluorogen-based reporter enabling to fluorescently label fusion proteins in living cells. My PhD thesis presents the developments of new FAST systems with various properties for multiplexed imaging. We report a collection of fluorogens enabling to tune the fluorescence color of FAST from green-yellow to orange and red. Beyond allowing multicolor imaging of FAST-tagged proteins in live cells, these fluorogens enable dynamic color switching because of FAST's reversible labeling, opening great prospects for the design of selective imaging methods relying on dynamic systems. In order to further expand the spectral properties of FAST to red, we also designed and developed a library of red fluorogenic dyes, for which we engineered specific protein binders by applying a directed evolution strategy based on the yeast display technology and high-throughput fluorescence activating cell sorting (FACS). We finally developed novel fluorogens able to form fluorescent complexes with FAST, but incapable of crossing the plasma membrane, which makes it possible to selectively detect FAST-tagged cell-surface proteins.

Keywords

fluorogenic probes, fluorescence, protein labeling, biological imaging.