



**HAL**  
open science

# Bright dimerized fluorogenic probes for imaging nucleic acids and proteins

Tkhe Kyong Fam

► **To cite this version:**

Tkhe Kyong Fam. Bright dimerized fluorogenic probes for imaging nucleic acids and proteins. Organic chemistry. Université de Strasbourg, 2019. English. NNT : 2019STRAF066 . tel-03510149

**HAL Id: tel-03510149**

**<https://theses.hal.science/tel-03510149v1>**

Submitted on 4 Jan 2022

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

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

**ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES (ED 222)**

**UMR 7021 - Laboratoire de Bioimagerie et Pathologies**

**THÈSE** présentée par :

**Kyong Tkhe FAM**

soutenue le : **13 décembre 2019**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Chimie/ Biologie

**Bright dimerized fluorogenic probes for  
imaging nucleic acids and proteins**

**THÈSE dirigée par :**

**M. KLYMCHENKO Andrey**

Directeur de recherche CNRS, Université de Strasbourg

**RAPPORTEURS :**

**Mme. DUCA Maria**

**M. GAUTIER Arnaud**

Chargé de recherches CNRS, Université Côte d'Azur, Nice  
Professeur, Sorbonne Université, Paris

---

**AUTRES MEMBRES DU JURY :**

**M. WAGNER Alain**

Directeur de recherche CNRS, Université de Strasbourg

**ENCADRANT :**

**M. COLLOT Mayeul**

Chargé de recherches CNRS, Université de Strasbourg





# Content

<b>Acknowledgments</b> .....	5
<b>List of abbreviations</b> .....	8
<b>1. Introduction</b> .....	10
<b>1.1. Fluorescence imaging approaches based on fluorogenic dyes</b> .....	10
1.1.1. <i>Photoinduced electron transfer and intramolecular charge transfer.</i> .....	12
1.1.2. <i>Förster resonance energy transfer</i> .....	13
1.1.3. <i>Aggregation-induced emission</i> .....	14
1.1.4. <i>Aggregation-caused quenching</i> .....	14
<b>2. A dimerization-based fluorogenic dye-aptamer modules for RNA imaging in mammalian cells</b> .....	18
<b>2.1. RNA Imaging</b> .....	18
2.1.1. <i>Hybridization methods</i> .....	19
2.1.2. <i>Chemical labelling approach</i> .....	19
2.1.3. <i>Enzymatic labelling approach</i> .....	19
2.1.4. <i>RNA-binding proteins</i> .....	20
2.1.5. <i>CRISPR-Cas9</i> .....	20
<b>2.2. RNA aptamers as genetically encoded imaging tags</b> .....	21
2.2.1. <i>Hydroxybenzylidene imidazolinone (HBI)-binding aptamers</i> .....	24
2.2.2. <i>Cyanine-binding aptamers</i> .....	27
2.2.3. <i>Mango aptamers</i> .....	28
2.2.4. <i>Malachite Green- and rhodamine-binding aptamers.</i> .....	30
2.2.5. <i>Non-fluorogen-binding aptamers</i> .....	35
<b>2.3. Gemini-561/o-Coral fluoromodule as a bright genetically encodable system for live-cell RNA imaging (article 1)</b> .....	42
<b>2.4. Gemini-505 and Gemini-647 fluorogens for SELEX-based RNA directed evolution (unpublished)</b> .....	89
2.4.1. <i>Design and synthesis Gemini-647</i> .....	89
2.4.2. <i>Gemini-647 exhibited double fluorogenic properties</i> .....	90
2.4.3. <i>Generation of Gemini-647 2.0, spectral properties and interaction with diSiRapt</i> .....	93
2.4.4. <i>Design and synthesis Gemini-505 and its spectral properties.</i> .....	95
<b>3. Self-labelling fluorogenic tags for protein imaging</b> .....	97

<b>3.1. Overview on self-labeling fluorogenic tags for protein imaging.....</b>	<b>97</b>
3.1.1. <i>Tags based on covalent modification .....</i>	97
3.1.2. <i>Fluorogen-binding tags.....</i>	99
<b>3.2. Gemini-584 as a fluorogen for dimeric fluorophore-binding peptide TR-512 (unpublished) .....</b>	<b>101</b>
3.2.1. <i>Design and synthesis of Gemini-584-alkyne and its spectral properties.</i>	101
3.2.2. <i>Photophysical studies of interaction of Gemini-584-alkyne with TR512 peptides.....</i>	103
<b>4. Fluorescence imaging of biotin receptors in cancer cells.....</b>	<b>105</b>
4.1. <b>Overview of fluorescent probes for imaging cellular biotin receptors..</b>	<b>106</b>
4.2. <b>Probing cell surface biotin receptors with rationally designed fluorogenic dimerized squaraines (manuscript, article 2) .....</b>	<b>109</b>
<b>5. General conclusion and perspectives .....</b>	<b>131</b>
<b>6. Materials and methods.....</b>	<b>133</b>
6.1. <i>Materials .....</i>	133
6.2. <i>Synthesis .....</i>	133
6.2.1. <i>Synthesis of Si-Rhodamine derivarive.....</i>	133
6.2.2. <i>Synthesis of Rhodamine derivatives.....</i>	134
6.2.3. <i>Synthesis of Gemini family of fluorogens.....</i>	136
6.2.4. <i>Synthesis of control compounds.....</i>	144
6.2.5. <i>Synthesis of biotinylated dimerized squaraines.....</i>	146
6.3. <i>Spectroscopy .....</i>	153
6.4. <i>Microscopy and cellular studies. ....</i>	154
<b>7. Annexe.....</b>	<b>157</b>
<b>Fluorogenic probes for imaging lipid droplets in cells and tissues .....</b>	<b>157</b>
<b>8. Résumé de thèse (version française) .....</b>	<b>197</b>
<b>References .....</b>	<b>209</b>
<b>List of publications.....</b>	<b>220</b>

## **Acknowledgments**

This work was performed in Nanochemistry and Bioimaging group (Laboratory of Bioimaging and Pathologies, UMR7021) and received financial support from Agence Nationale de la Recherche (BrightRiboProbes, ANR-16-CE11-0010-01/02).

Firstly, I would like to acknowledge jury members for their kind acceptance to evaluate my work. I am humbled to be evaluated by high-rank experts in chemical biology field. This PhD research work was significantly inspired by nature, in particular by widely used dimerization process. Dimerization approach was also applied to my everyday work in the lab, in particular I had a unique opportunity to work with two great supervisors and mentors: Dr. Mayeul COLLOT and Dr. Andrey KLYMCHENKO. I feel extremely fortunate for the chance to work with two incredible scientists who guided and extremely supported me throughout my PhD journey and beyond. Their passion for science and constant flow of new ideas kept me inspired all these 3 years helping me to develop my own ideas for research proposals. I have learned so many things from chemistry to biology from them! Not only are they great scientists, but also great mentors and colleagues! They always welcomed my ideas/suggestions/opinion, encouraged me to learn new techniques etc. Especially, I would like to express my enormous gratitude to Mayeul since I am his first PhD student. It was not easy sometimes to supervise me and I do not know how it turned out to him, but personally for me, it turned out beyond what I could imagined for my PhD! I am so grateful for all the opportunities, support and critiques from both my supervisors that helped me to grow on professional as well as on personal level. On top of that, they created a healthy environment to work in which is extremely important for well-being. I believe our work together will continue to fruitfully elaborate and amplify like our ultrabright fluorogenic probes. Indeed, sometimes two is better than one. *Merci maître Mayeul pour votre soutien!*

Dimerization does not end there. We would not be able to successfully accomplish our work on fluorogenic aptamers without fruitful collaboration with Farah BOUHEDDA and Dr. Michael RYCKELYNCK who selected and characterized o-Coral aptamer. Thank you for introducing me to the world of RNA SELEX and microfluidics. It was a pleasure to work together these years. Also, I would like to acknowledge our long-standing collaborator Dr. Lydia DANGLOT for our fruitful work on SMCy probes and many others.

I cannot thank enough previous and current members of Nanochemistry and Bioimaging group for such a friendly and supportive environment where you feel appreciated and valued. It was heartwarming to know that I could always find a help whenever I needed from every member of the lab. As a diverse lab, I learned so many things whether it was work related or just cultural exchange. Also, not only I could learn things, but also I had opportunities to share some of my experiences and knowledge. It was such a great pleasure to work and be around you, guys! I am not highlighting particular names as I am very grateful to everyone equally ☺ Thank you so much for sharing so many precious moments and cakes together!

Our lab is very lucky to have Dr. Guy DUPORTAIL who is usually one of the first person a newcomer will encounter. Thank you, Guy, for welcoming newcomers and navigating us in French system. I am very thankful to Marlyse WERNET and Ingrid BARTHEL for their incredible efficient administrative work and willingness to help! I am also thankful to previous and current members of group Mely for the opportunity to meet such brilliant scientists and nice people, especially Dr. Eleonore REAL and Dr. Emmanuel BOUTANT for giving a flask course on RNA biology, Dr. Nicolas HUMBERT for always helping with technical problems.

I would never have thought in a million years that I would ever experience sense of social community, especially of a Ukrainian one. I am so fortunate to meet many incredible people and my very good Ukrainian friends in Strasbourg. Thank you so much for your enormous support and making 'ma vie Strarbourgeoise' so comfortable and heartwarming! Дякую за те, що ви є! Саме ви змусили мене полюбити Страсбург.

I am not sure I would determinedly pursue PhD if I didn't have experience at EPFL. My time in Lausanne, in particular at EPFL, was an eye-opening experience. It was the first place where I met so many inspiring scientists and got opportunity to see cutting edge world-class science. Special acknowledgement to Prof. Vassily HATZIMANIKATIS for his support and navigating me through my EPFL journey; Prof. Kai JONHSSON for his support and insightful Chemical Biology course; Prof. Jerome WASER and Dr. Cormac GAHAN for boosting my confidence as the result of their feedback for my 'first year candidacy exam'. The special place is for my colleagues and good friends from LCBIM: Tamara, Arkadiy, Pavel, Hacer, Grisha, Anzhela, Alma and others. Thank you so much for inspiring me to do science, for your incredible

support, trainings, help and unforgettable moments during my time at EPFL! I learned so many valuable things from you! Even though it did not work out at EPFL, thank you, Prof. Elena DUBIKOVSKAYA (GOUN), for the opportunity to work in LCBIM meeting this amazing group of people and for my favorite frenchy Charlotte who moved to Strasbourg for her PhD as well!

Separate paragraph goes to amazing women in science I met! Gender equality is still far from common practice in science. I am proud working in France where this issue is addressed towards improving gender balance on all levels of organization. Throughout my journey I met so many talented and inspiring female scientists who are vocal about their positions. Thank you for raising public awareness about gender equality misbalance. I am humbled to call some of those women my friends: Julia and Anya, Tamara, Anzhela, Hacer and Alma, Nina, Lesia, Katya, Caterina, Sophie, Sylvie, Marie, Anne, Farah, my favorite frenchy Charlotte and many many others. Thank you for your valuable representation of women in science!

Back in Ukraine I worked at Enamine LTD in the 407/143 lab. Thank you a lot, guys, for fun time there! So many things to remember. Also, thanks to ex-colleagues at Curplyx (Macrochem) for valuable experience in industrial synthetic chemistry. Thanks to Natural Products chemistry group (Taras Shevchenko Kyiv National University) for their support and encouragement, especially my former advisor Dr. Olga KHILYA. Kyiv uni is a great place to study chemistry. I cannot forget my group mates from Taras Shevchenko Kyiv National University! Thank you for all the ups and downs of uni we went through together.

Va tất nhiên Cừơng muốn gửi lời cảm ơn cho cả gia đình ở Việt Nam, cộng đồng Việt Nam tại Ukraina. Cảm ơn bố mẹ Long Xuân và em gái Linh vì luôn luôn bên cạnh Cừơng, mong, tin và giúp đỡ. Cừơng rất hạnh phúc và may mắn vì có gia đình như vậy!

Thank you, all my friends from all over the world, for your support and belief! Sometimes it just strikes me how people can believe in you more than you believe in yourself. Thank you everyone for your collective contribution in better version of me.

## List of abbreviations

μIVC	microfluidic-assisted <i>in vitro</i> compartmentalization
ACQ	aggregation-caused quenching
AGT	O <sup>6</sup> -alkylguanine-DNA alkyltransferase
AIE	aggregation-induced emission
BG	O <sup>6</sup> -benzylguanine
BHQ1	black hole quencher
Boc	<i>tert</i> -butoxycarbonyl
BODIPY	boron-dipyrromethene
BR	biotin receptor
BSA	bovine serum albumin
Cbl	cobalamin
CR	carborhodamine
CRABP II	cellular retinoic acid binding protein II
CRISPR	clusters of regularly interspaced short palindromic repeats
Cas	CRISPR associated protein
CuAAC	copper-catalyzed azide-alkyne cycloaddition
DCM	dichloromethane
DFHBI	3,5-difluoro-4-hydroxybenzylidene imidazolinone
DFHO	3,5-difluoro-4-hydroxybenzylidene imidazolinone-2-oxime
DIEA	<i>N,N</i> -diisopropylethylamine, Hünig's base
DIR	dimethylindole red
DMF	dimethylformamide
DMHBI	3,5-dimethoxy-4-hydroxybenzylidene imidazolinone
DN	dinitroanilin
DTT	dithiothreitol
<i>E.coli</i>	<i>Escherichia coli</i>
eDHFR	<i>E.coli</i> dihydrofolate reductase
EGFR	Epidermal growth factor receptor
FACS	fluorescence-activated cell sorting
FAST	Fluorescence-Activating and absorption-Shifting
FBS	fetal bovine serum
FCS	Fluorescence correlation spectroscopy
FISH	fluorescence in situ hybridization
FP	fluorescent protein
FRET	Förster resonance energy transfer
GBL	gamma-butyrolactone
g-Coral	green Coral
GFP	green fluorescent protein
GMC	gemcitabine
GSH	glutathione
HATU	Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium
HBC	cyanophenylacetonitrile
HBI	hydroxybenzylidene imidazolinone
HBR-3,5DOM	4-hydroxy-3,5-dimethoxybenzylidene rhodanine

hCA	human carbonic anhydrase I
hCG	human chorionic gonadotrophin
HJ	Holliday junction
HMBR	4-hydroxy-3-methylbenzylidene-rhodanine
HPLC	high-performance liquid chromatography
ICT	intramolecular charge transfer
IEDDA	inverse-electron demand Diels–Alder
$K_d$	dissociation constant
LD	lipid droplet
LPA	lysophosphatidic acid
LSS	large Stokes shifts
MCP	MS2 coat protein
MGA	malachite green aptamer
MPP	<i>N</i> -( <i>p</i> -methoxyphenyl)piperazine
NIR	near-infrared
NMR	nuclear magnetic resonance
o-Coral	orange Coral
PBV	blue violet
PBVF	patent blue VF
PeT	photoinduced electron transfer
POI	protein of interest
QY	quantum yield
R110	rhodamine 110
RBP	RNA-binding protein
r-Coral	red Coral
ROS	reactive-oxygen species
RT-PCR	reverse transcription polymerase chain reaction
scFvs	single-chain antibodies
SDS	sodium dodecyl sulfate
SELEX	Systematic Evolution of Ligands by EXponential enrichment
SiR	silicon rhodamine
SMCy	StatoMerocyanine
SMVT	Sodium dependent multivitamin transporter
SPAAC	strainpromoted azide-alkyne cycloaddition
SRB	sulforhodamine B
SRB101	sulforhodamine 101
TFA	trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
TMP	2,4-diamino-5-(3,4,5-trimethoxybenzyl) pyrimidine
TMR	tetramethylrhodamine
TO	thiazole orange
TR	Texas Red, sulforhodamine 101
XyCyFF	xylene cyanole FF



## **1. Introduction**

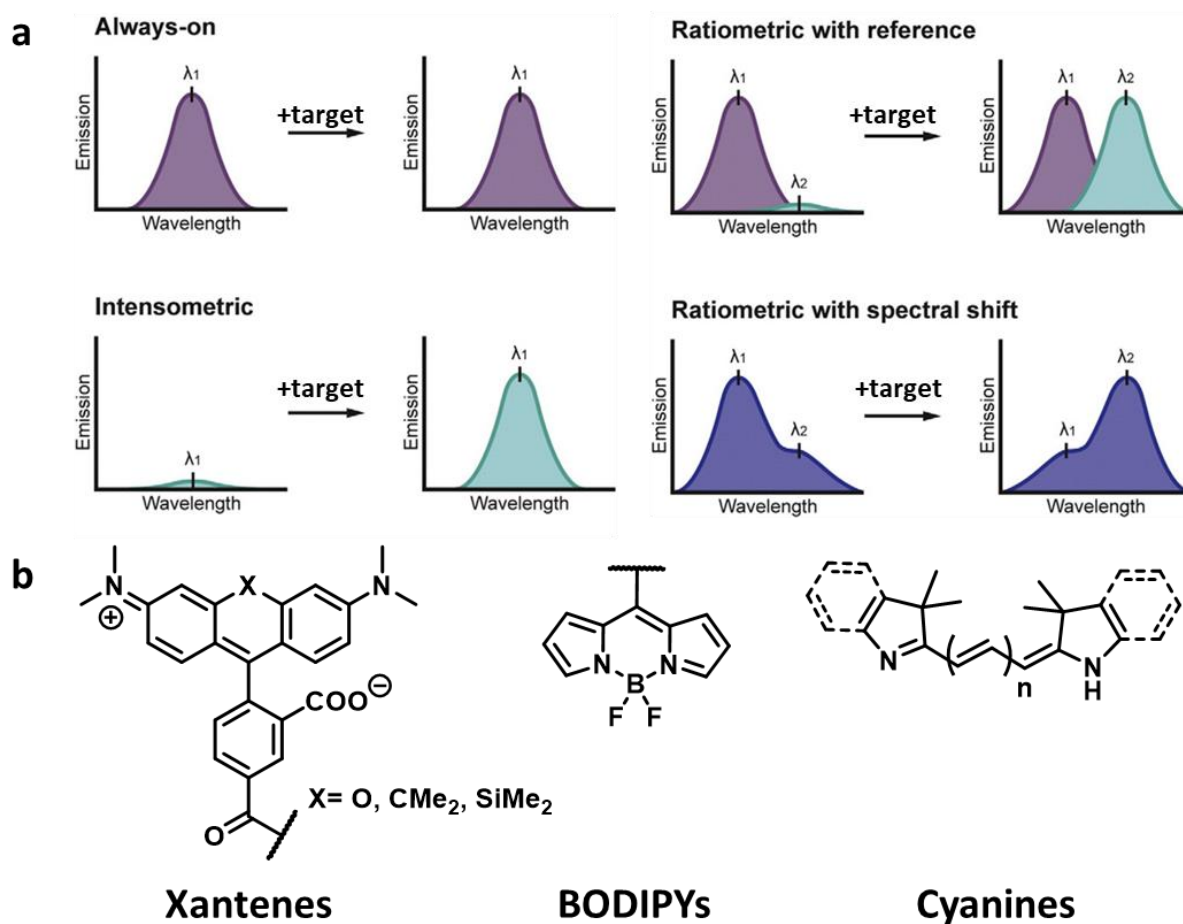
### **1.1. Fluorescence imaging approaches based on fluorogenic dyes**

Fluorescence is the process of emission of a photon by a molecule from its singlet excited state after its excitation by light.<sup>1</sup> Fluorescence molecular imaging techniques have enabled us to better understand the complexity of living systems and to reveal the mechanisms of their functionalities under both physiological and pathological conditions. As an imaging modality, fluorescence provides the specific spatial and temporal information about the biological event of interest.<sup>2</sup> From a practical point of view, fluorescence imaging techniques are relatively inexpensive, user-friendly to handle, and avoid exposure to radioactivity risk of other molecular imaging modalities such as positron emission tomography. Owing to its sensitivity, fluorescence molecular imaging is commonly used in pre-clinical research allowing non-invasive qualitative and quantitative visualization of bioprocesses from subcellular level of the living systems up to whole body of small animals.<sup>2,3</sup>

Fluorescence imaging requires utilization of fluorophores such as endogenous molecules (e.g., tryptophan, hemoglobin), genetically expressed fluorescent proteins (e.g., green fluorescent protein, GFP), or exogenous molecular fluorophores and fluorescent nanoparticles. Fluorescent molecules are usually coupled to a biomolecule of interest with the aim to study its biological behaviour.<sup>2</sup> Another strategy is to integrate a fluorophore to a recognition unit (e.g. receptor ligand, antibody etc.) to interact with a biological target. The major usefulness of fluorescent probes for molecular imaging is their tunable responsivity to their environment.<sup>2,4</sup> This feature pushed the research towards extensive development of fluorogenic probes emitting photons in response to a biological target.<sup>2,4,5</sup>

There are several established strategies in designing fluorescent probes.<sup>2</sup> Traditionally, a simple fluorescent label attachment to its target was used to navigate a researcher like a light beacon in the cellular environment, so called always-on probes. Always-on fluorescent probes rely on their selective accumulation in target in order to provide specificity.<sup>5</sup> For example, this strategy is commonly used to fluorescently label drug so that it can report the biodistribution but not bioavailability of the encapsulated drug. For many drugs such a chemical modification can alter its therapeutic effect. To overcome this problem, chemists developed a number of cleavable linkers that release a drug from a fluorescent cargo.<sup>2,6</sup> In addition, some

fluorophores can non-specifically bind to non-targeted biomolecules, thus challenging for the discrimination of the target signal from false-positive signal. For this reason, the dyes activating fluorescence in response to an interaction with the target can improve the signal-to-background fluorescence, so called fluorogenic probes or fluorogens.<sup>2,4</sup> There are two general design approaches to achieve activatable fluorescence: intensometric or ratiometric fluorogenic response (Figure 1a).<sup>2</sup> In intensometric approach the fluorogen is not emissive unless it interacts with its target to produce fluorescence response (Figure 1a). This approach is also known as the 'off-on' fluorescence imaging approach. The main advantage of 'off-on' systems is its low background fluorescence since its quenched form is not emissive providing a better contrast of images. Unlike intensometric systems, in ratiometric systems, the fluorophore is usually fluorescent, but its interaction with the biological target changes its fluorescence wavelength.<sup>4</sup> As the result, it enables to monitor the change the ratio of emission intensities at two wavelengths that correspond to inactive and active form (Figure 1a).<sup>7</sup> Ratiometric response is convenient methodology to study distribution of the probe and the biological process of interest at same time since it produces the emission of a different color once it reaches the target and the ratio response is independent of instrumental parameters and probe concentration.



**Figure 1.** (a) Different approaches for fluorescent probes. Comparison of emission spectra of always-on, intensometric and ratiometric probes. Adapted from ref.<sup>2</sup> (b) Chemical structures of typical fluorophores commonly used in chemical biology.

Despite efforts in developments of novel fluorophores, most fluorescent agents are generally derivatives of well-known dyes, like xanthene dyes (fluorescein, rhodamines), cyanines and BODIPY dyes (Figure 1b). The choice of the fluorophore is usually dependent on the final application of the fluorescent probe and desired photophysical and photochemical properties (*i.e.* photostability, brightness etc.).<sup>5</sup>

In this chapter, designing fluorogenic probes to develop fluorescence imaging approaches for biological application will be briefly discussed.

### 1.1.1. Photoinduced electron transfer and intramolecular charge transfer.

Photoinduced electron transfer (PeT) is the process of the electron transfer from excited state of donor molecule to a ground state of an acceptor molecule.<sup>1</sup> This

mechanism is distinct from Dexter interactions where the donor and acceptor inter exchange electrons. PeT can occur through molecular collisions (dynamic quenching) or through molecular complex formation (static quenching). During the static quenching the complex forms transient state. This type of interaction is also known as excited dimer (excimer) in the case where the donor and acceptor are the same species or excited complex (exiplex) if they are of the different origin. After PeT takes place, the recomposition of the transferred electrons occurs so that the starting ground state is restored. Since the efficient electron transfer is distance dependent, PeT is also known as contact quenching and the close proximity of donor and acceptor is in the range of  $\sim 0.7$  nm.<sup>1</sup>

If excitation lead to repartition of charge within the same molecule, the process is called intramolecular charge transfer (ICT).<sup>1</sup> This photophysical principle is widely used in development of push-pull fluorophores.<sup>4</sup> In a typical push-pull fluorophore the repartition of the electron density from a donor to an acceptor is environment dependent, i.e. the efficient electron transfer is favorable in polar solvents.<sup>4</sup>

Owing to its sensitivity to surroundings, in our group push-pull dyes found their application to study biophysical properties of cellular membranes, in particular membrane polarity, hydration, viscosity etc. In addition, their usefulness was shown in the investigation of biomolecular interactions.<sup>4,8-12</sup>

### *1.1.2. Förster resonance energy transfer*

Förster resonance energy transfer (FRET), is the non-radiative energy transfer from an excited state of the donor to a ground state of the acceptor due to their dipole-dipole coupling.<sup>1</sup> Generally, the donor emission should have efficient spectral overlap with the acceptor absorption in FRET systems. The efficiency of this energy transfer is highly dependent on the distance between donor and acceptor and proportional to  $\sim 1/R^6$ , where R is the Förster radius ( $\sim 5$  nm).<sup>1</sup> FRET efficiency can be measured by monitoring the change in the ratio between fluorescence intensity of the donor and acceptor. Because of the proximity dependence of the FRET pair, measurements of the FRET efficiency found a wide application to study intercommunication of the fluorescently labelled biomolecules at certain distances.<sup>13</sup>

### 1.1.3. Aggregation-induced emission

Fluorogenic molecules with aggregation-induced emission (AIE) properties (also known as AIEgens) are poorly fluorescent in monomolecular state, but highly emissive in collectively aggregated state, owing to the restriction of intramolecular motions (RIM) and prohibition of non-radiative decay.<sup>14</sup> AIEgens have been actively designed as fluorescence turn-on probes for biosensing through controllable de-aggregation/aggregation processes.<sup>15</sup>

### 1.1.4. Aggregation-caused quenching

Aggregation-caused quenching (ACQ) is caused by strong hydrophobic effects or  $\pi$ - $\pi$  stacking of fluorophores and is commonly observed as non-linearity in fluorescence intensity at high concentrations.<sup>1</sup> This phenomenon is common for poorly soluble dyes (typically hydrophobic fluorophores in aqueous media) since its planar aromatic systems provoke formation of non-radiative excimers/exciples.<sup>1</sup> Aggregation is thermodynamically favorable process and is challenging to disrupt.<sup>1</sup> Therefore, ACQ can generate an intensometric fluorogenic response if the initially quenched species are spatially separated upon interaction with the biological target (Figure 2).<sup>16</sup> Both intermolecular and intramolecular aggregation processes can be used to design ACQ-driven probes.

Intermolecular self-aggregation (Figure 2a) is common for all fluorophores in concentrated solutions, but especially for very lipophilic dyes such as squaraines, cyanines, BODIPYs etc.<sup>16</sup> These dyes were exemplified in a number of applications. Chen *et al.* synthesized chemosensors based on symmetrical and unsymmetrical squaraines to detect  $\text{Hg}^{2+}$  ions in water (Figure 2a).<sup>17,18</sup> These sensors self-assemble in water forming *H*-aggregates with characteristic blue-shifted absorption peak at ~ 550 nm. Once they were bound to  $\text{Hg}^{2+}$  the coordination arms promoted steric hindrance and thus render disaggregation. As the response, the monomolecular absorption peak at 644 nm appeared as well as the fluorescence enhanced.

Zhai *et al.* used hydrophobic BODIPY as fluorescent reporter in the development of a probe for the illicit date rape drug-gamma-butyrolactone (GBL) named Green Date (Figure 2a).<sup>19</sup> Green Date formed fluorescence-quenched particles in aqueous solution. Upon increasing concentration of GBL, it entered into the stacked

nanoparticles promoting disassembly and decrease of large nanoparticles in the solution, hence the fluorescence of Green Date was turned on.<sup>19</sup>

Hamachi and colleagues introduced a self-quenched fluorogenic probes with tethered ligand to study protein interaction, in particular human carbonic anhydrase I (hCA), avidin and trypsin.<sup>20</sup> This work highlighted the importance of the fluorophore choice: hydrophobic BODIPY-based probe showed a better self-assembly thus producing less fluorescence in the inactive form unlike hydrophilic tetramethylrhodamine (TMR)–based probe. To promote self-aggregation of hydrophilic dyes such as fluorescein and rhodamine, Hamachi group proposed an introduction of hydrophobic module between fluorophore and detection modality.<sup>21</sup> These recognition-driven disassembly of a ligand-tethered fluorogenic probes enabled visualization of overexpressed cell-surface folate receptor and hypoxia-inducible membrane-bound carbonic anhydrases on cancer cells.<sup>21</sup>

In our group, amphiphilic probe NR12S (Figure 2a) for lipid membrane was developed.<sup>8</sup> It formed of nonfluorescent micellar *H*-aggregates in aqueous media. Binding to membranes provoked disassembly in monomolecular fluorescent form with increase of quantum yield from 0.002 to 0.5. Additionally, an analogue of NR12S was developed with controllable disassembly by a reductive stimuli producing fluorogenic response.<sup>22</sup> Moreover, a family of multicolor ultrabright fluorogenic membrane probes based on cyanines was developed namely MemBright.<sup>23</sup> MemBright probes aggregated in self-quenched particles in aqueous solutions and disassembled upon binding to membranes with fluorescent response.<sup>23</sup> Our group also explored squaraine tendency to ACQ resulting in development of SQ-PEG-CBT<sup>24</sup> for oxytocin receptor imaging and dSQ12S<sup>25</sup> for membrane imaging. Disaggregation of cyanine dyes was used in the recognition of mixed G-quadruplex structures in human telomeres.<sup>26</sup> The probe formed *J*-aggregates in water, but disassembled upon strong interaction with a specific antiparallel G-quadruplex without diagonal loop.<sup>26</sup> Recently, we have developed a family of multicolor fluorogenic probes for tracking lipid droplets (LDs) in cells and tissues.<sup>27</sup> This family of new dyes was based on merocyanine fluorophores and a dioxaborine barbiturate moiety, named StatoMerocyanines (SMCy, figure 2a). SMCy dyes are highly lipophilic and neutrally charged, hence formed self-quenched nanoparticles in aqueous environment. Once in cellular environment SMCy accumulated in neutral lipophilic organelles such as LDs revealing its fluorescence.<sup>27</sup>

Historically, the first example of intramolecular dimerized ACQ-probe is the commercially available DNA probe YOYO-1 (Figure 2b).<sup>28</sup> Its design is based on dimerization of the dye covalently conjugated through a linker.<sup>28</sup> YOYO-1 formed ground-state dimer with spectral characteristics that can be explained by exciton theory.<sup>1</sup> Dimers were non-emissive in water and responded as fluorescence enhancement upon binding to DNA.<sup>28</sup>

One of the earliest example in dimerized-dye probes was shown back in 1994 with system comprising an epitope of human chorionic gonadotrophin (hCG) labeled with fluorescein and tetramethylrhodamine at N- and C-terminus respectively.<sup>29</sup> When free in solution, the dyes aggregated forming intramolecular quenched heterodimer, but upon binding to target antibody, the probe dissociated with the fluorescence enhancement response as FRET signal.

Significant efforts in development of fluorogenic dimers for protease sensing were made by Packard *et al.* They developed a series of probes targeting the serine protease elastase activity.<sup>30–33</sup> A typical protease activity-based probe leveraged NorFES (recognition unit) and dual-labels on peptide-terminus with fluorophores (Figure 2b).<sup>30–33</sup> One of the first probes was labeled with tetramethylrhodamine, TMR-NorFES-TMR.<sup>30</sup> Addition of prolines provided additional conformational constraints, thus promoted spatial proximity of two fluorophores to quench their fluorescence forming *H*-aggregates. Once NorFES-sequence is cleaved by serine protease, dyes were spatially separated with fluorescence enhancement response. Interestingly, authors showcased the advantage of homodimerization strategy over heterodimerization as well as addition of flexible linker (6-aminohexanoic acid) promoted intramolecular dimerization.<sup>32,33</sup>

Recently, others and we expanded this concept. Knemeyer *et al.* developed a hairpin-structured oligonucleotide that is labeled with two identical fluorescent dyes (MR 121) that are able to form quenched *H*-type dimers while the hairpin is closed (Figure 2b).<sup>34</sup> Upon hybridization to the target DNA, the dyes are separated and a 12-fold increase of the fluorescence intensity occurs. The probe was used for the specific detection of *Mycobacterium xenopi* in a model system. Our group reported a probe to image oxytocin receptor based on dimeric squaraine dye attached to carbetocin (Figure 2b).<sup>35</sup> In aqueous media it showed quenched fluorescence due to formation of intramolecular *H*-aggregate, where its fluorescence revealed upon binding to the receptor and





development of fluorogen for dimeric peptide-tag and development of the dimerized probe for sensitive imaging of cancerous biotin receptors.

## **2. A dimerization-based fluorogenic dye-aptamer modules for RNA imaging in mammalian cells**

### **2.1. RNA Imaging**

RNA is an important biomolecule in a variety of biological processes such as gene regulation and protein translation.<sup>38</sup> The synthesis of RNA molecules is the result of the transcription of DNA by the cell machinery. The RNA pool is quite heterogeneous and the relative abundance of each is dependent on the physiological state of the cell.<sup>39</sup> RNAs are synthesized during transcription process by RNA-polymerases in the nucleus. After maturation, the majority of RNAs are exported from the nucleus to cytoplasm *via* nuclear pores, but some RNAs can be redirected back to the nucleus. Further processing can include RNA transport and specific localization to perform its destined function in response to developmental cell stage, environmental signaling or perturbation.<sup>40</sup>

Traditional methods to investigate RNA biology involve total RNA extraction from cell lysates with consequential *in vitro* analysis using biochemical techniques, in particular reverse transcription polymerase chain reaction (RT-PCR) and the northern blot.<sup>40</sup> RT-PCR is based on reverse transcription of RNA into complementary DNA and its further amplification. While RT-PCR is primarily used to quantify gene expression level, the northern blot is used to detect and separate RNAs by gel electrophoresis.<sup>40</sup> Despite being powerful techniques, RT-PCR and northern blot provide only snapshot of the indirect RNA processing and vague idea on its distribution in cells.

On the other hand, imaging technologies can bring new insights on RNA role in the intracellular biological processes. For instance, mRNAs play a key role in gene expression and control. Therefore, imaging the transcription process, spotting the localization, studying the mRNA transport and its further translation into the protein can provide a bigger picture of gene regulation processing and dynamics.<sup>41</sup> Unfortunately, naturally occurring RNAs do not have any fluorescent means to be compatible for imaging. For this reason, for decades scientists have been working on development of imaging probes to visualize RNAs.<sup>42</sup>

This chapter will briefly describe the advantages and pitfalls of current methods to image sequence-specific RNAs using fluorescence as a read-out, the potential of new methods that are under development.

### *2.1.1. Hybridization methods*

Historically, fluorescence in situ hybridization (FISH) was the first method to label cellular RNA in the late 1960s.<sup>43</sup> The method is based on interaction of chemically labelled oligonucleotide sequence upon hybridization with the previously determined sequence of the targeted RNA (Figure 3.1). FISH is highly specific to the endogenous RNA of interest and does not require manipulation to induce expression of genetically encoded tags in cells. However, its multistep protocol, in particular cell fixation and permeabilization, can alter their native localization. Moreover, FISH provides only the snap shot of the RNA bioprocessing.

### *2.1.2. Chemical labelling approach*

Recent attempts to visualize RNA is the chemical covalent labelling approach (Figure 3.2).<sup>44</sup> Introduction of small chemical tag minimizes the potential disruption of RNA functionality and folding, in particular it is important for small RNAs.<sup>44</sup> The main strategies for covalent RNA modification are (i) direct synthesis of fluorescent nucleotide derivatives and further its insertion into RNA backbone during transcription; (ii) introduction of small reactive group into nucleotide that can react with respective modified fluorophore by bioorthogonal chemistry after transcription.<sup>44</sup>

### *2.1.3. Enzymatic labelling approach*

Alternative approach, complementary to chemical labelling, is the chemo-enzymatic labelling.<sup>45</sup> It allows the introduction of non-canonical nucleotides bearing reactive bioorthogonal handle *via* cell machinery using RNA polymerases (Figure 3.2).<sup>45</sup> Following enzyme-mediated installation of the chemical reporter, chemical labelling takes place *via* widely used bioorthogonal reactions like copper-catalyzed azide-alkyne cycloaddition (CuAAC), strainpromoted azide-alkyne cycloaddition (SPAAC), the inverse-electron demand Diels–Alder reaction (IEDDA) and the photoclick reaction. Tandem of chemo-enzymatic labelling strategy allows site-specific RNA modification.<sup>45</sup>

A drawback of incorporating modified nucleotides during transcription is that it is difficult to differentiate different subtypes of RNA.<sup>44</sup>

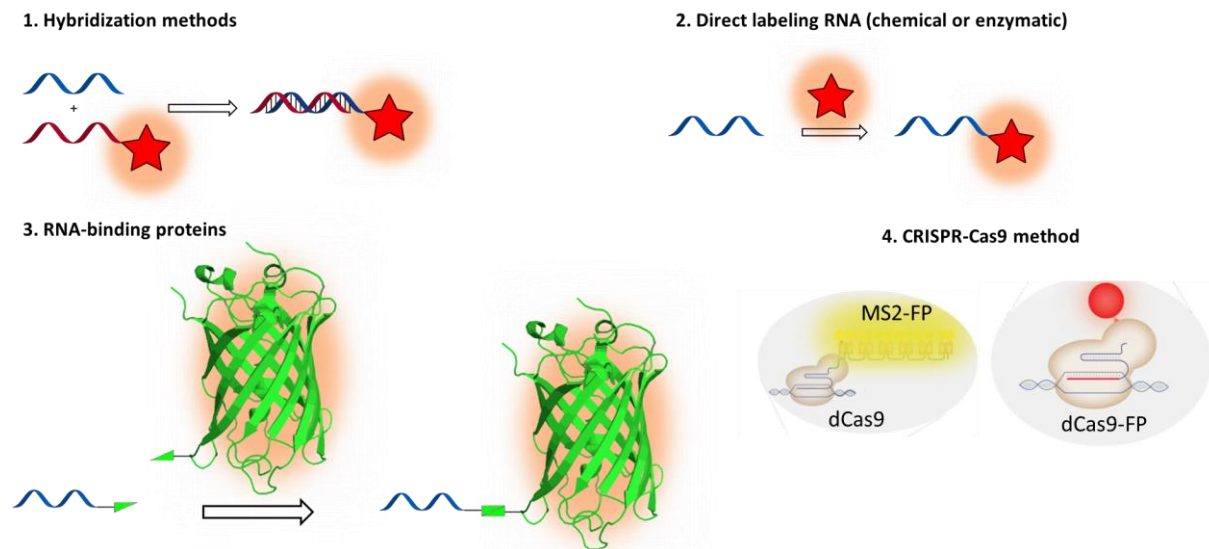
#### 2.1.4. RNA-binding proteins

RNAs can be tagged with a certain nucleotide sequence that can be specifically recognized by protein motifs (Figure 3.3).<sup>46</sup> The most commonly used RNA-binding proteins (RBPs) is bacteriophage derived MS2 coat protein (MCP, 14.4 kDa) that is able to selectively bind to MS2 nucleotide binding sequence.<sup>47</sup> This RNA motif can be genetically labeled to 3'UTR untranslated region of targeted RNA. MCP can be fused with fluorescent protein like GFP so that the system can be used in fluorescence imaging. Owing to FPs limited brightness, these systems require multimerization of a tag (typically 24 tandem repeats) to achieve single-transcript sensitivity.<sup>46</sup> Additionally, to reduce background signal, a split version was developed so that the fluorescence is switched on upon binding of split fragments of FP.<sup>48</sup> The size of FPs is generally large, especially in comparison to small RNAs. This may burden the movement and perturb the function of the target RNA.

#### 2.1.5. CRISPR-Cas9

The discovery of the CRISPR-Cas9 (clusters of regularly interspaced short palindromic repeats, CRISPR associated protein 9) system has tremendously changed the field of biotechnology providing high precision nucleic acids editing control.<sup>49</sup> While primarily CRISPR-Cas9 system was originally developed for gene editing its potential is far beyond. The recent approaches demonstrated its versatility and potential in gene regulation, epigenome editing and live-cell chromatin imaging.<sup>49,50</sup> Advances in development of Cas9 nucleases variants, in particular discovery of catalytically inactive Cas9 (dCas9), enabled to create fluorescent protein-fused platforms for imaging repetitive targeted nucleic acid sequences (Figure 3.4). Until recently, most work has focused on Cas9 systems that target DNA. Recently Nelles *et al.* reported optimized *S. pyogenes*' Cas9 (spCas9) system to target RNA in live cells by adding a small part of DNA that create a duplex with the target RNA thus providing the required PAM binding site.<sup>51</sup> The resulting RNA-targeting Cas9 (RCas9) system can be used to selectively label target RNA with fluorescent proteins, thus allowing to image endogenous RNA in live cells.<sup>51</sup> In 2016, Abudayyeh *et al.* reported the discovery the

first native RNA-targeting single-subunit Cas protein (Cas13a) and showed that it can be programmed to bind a target RNA.<sup>52</sup> Alternatively, a single guide RNA (sgRNA) was re-engineered to contain MS2 motifs that bind to the bacteriophage MCP and label these motifs with fluorescently tagged MCPs.<sup>53</sup>



**Figure 3.** Current state-of-art RNA labeling technologies. (1) RNA of interest (blue strand) is imaged upon hybridization with antisense sequence synthetically synthesized with fluorophore; (2) RNA of interest (blue strand) is directly labeled *via* incorporation of fluorescent nucleoside or small reactive chemical tag. Alternatively, labelling can be facilitated *via* enzymatic incorporation of reactive nucleoside; (3) RNA of interest (blue strand) can be fused with MS2 motif which binds to its cognate bacteriophage protein tagged with fluorescent protein; (4) Fusion of FP to dCas9 or incorporation of MS2 motif into sgRNA enables precise RNA imaging using CRISPR technology. Images for CRISPR-Cas9 method were adapted from ref. <sup>49</sup>

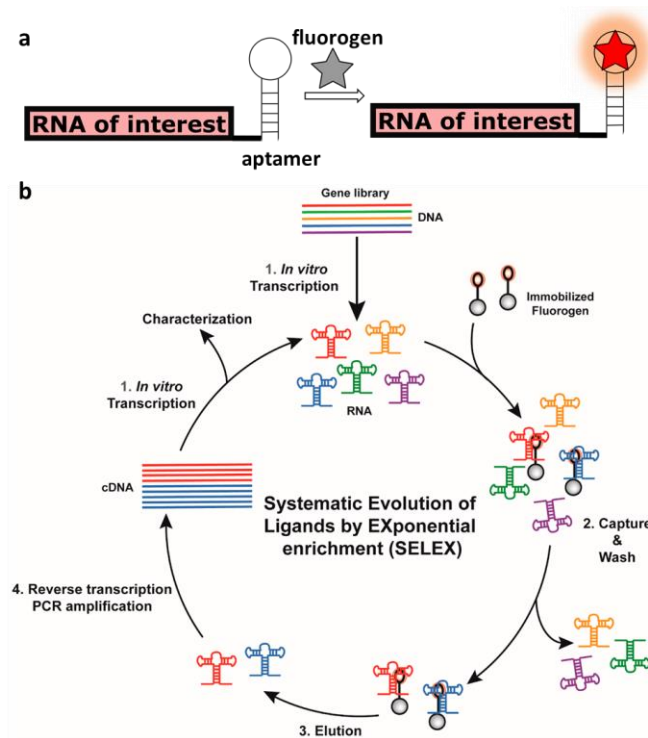
## 2.2. RNA aptamers as genetically encoded imaging tags

Though techniques described above made significant discoveries in RNA cellular biology, there are some drawbacks limiting their wider scope of application. For example, molecular beacons (hybridization method) require invasive exogenous injection of the probe into live cells limiting the throughput of the assay. Hybridization methods provide only snapshots of RNA cellular evolution since it requires fixation of cells and complex protocol.<sup>46</sup> RNA-binding protein based systems enabled non-invasive imaging of endogenous RNA in live cell. RBPs proved to be compatible in

single molecule super resolution imaging shining light on complex molecular RNA machinery.<sup>42,46</sup> However, the system demands maturation of GFP and its bulky size makes it unsuitable to tag small RNAs.

RNA aptamers are short oligonucleotides that recognize its target with extremely high selectivity.<sup>54</sup> Aptamers are attractive alternative to antibodies for targeted recognition of a molecule of interest. Selection process for aptamers has a number of advantages over antibody selection techniques: (i) flexibility of selection strategies; (ii) relative low cost; (iii) cell/animal-free *in vitro* production. Additionally, aptamers are relatively small in size, thermally stable, low-toxic, structurally modular for chemical transformations and desired functionalities.<sup>55</sup> These characteristics make aptamers appealing as functional modules in biosensors. Despite these compelling features, aptamers are currently not widely used.<sup>40,56</sup>

Conventional aptamer isolation technology uses Systematic Evolution of Ligands by EXponential enrichment (SELEX) approach.<sup>56</sup> SELEX is conceptually driven by repetitive stepwise directed evolution of RNA (or DNA) library that are selected to bind the targeted molecule. Current SELEX technologies can handle a pool of sequence diversity up to  $10^{15}$  sequences within one round of experiments. Main steps of SELEX process are described in Figure 4b.<sup>55</sup> SELEX starts with a diverse library of single-stranded nucleic acid random sequences which is displayed to a target of interest. Then, bound complexes are collected while unbound is discharged. Next, elution step allows to extract RNA (or DNA) binders which are amplified by PCR (polymerase chain reaction) for the further analysis. The generated population of sequences is used for the next round of selection. The cycle is repeated until the desired affinity; specificity or other characteristics are achieved.<sup>55,56</sup> Interestingly, a target can be not only small molecules, but also proteins<sup>57</sup> or even more complex systems like cells<sup>58,59</sup>.



**Figure 4.** Fluorogenic aptamer approach in RNA imaging. **(a)** Principle of aptamer-based technology. RNA of interest is fused with aptamer that selective binds fluorogen with switch-on fluorescence response; **(b)** *In vitro* selection cycle using SELEX. Each round follows four main steps. A DNA gene library is *in vitro* transcribed (Step 1) and mixed with target-loaded (fluorogen) beads. Weak binders are discharged by washing (selection pressure, Step 2) prior to recovering binding RNAs during the elution step (Step 3). Finally, RNAs are converted into cDNAs (Step 4), later used to start a new round of selection. Figure 4b is taken from ref. <sup>55</sup>

Aptamers provide a new alternative RNA imaging approach (Figure 4a). Several RNA aptamers which specifically recognize fluorogenic dyes or fluorophores have been developed by SELEX technology.<sup>60–62</sup> Fluorogenic aptamers are advantageous as RNA aptamers are highly modular and the structure can be rapidly optimized by SELEX.<sup>55,63,64</sup> On the other hand, fluorogens can be chemically evolved to gain improved photophysical properties retaining or even increasing binding capacity to cognate aptamer.<sup>65</sup> Additionally, fluorogens are low in molecular weight with tunable cell permeability.

In this chapter, the overview on established fluorogenic dye – aptamer systems will be discussed. The table summarizing their properties is provided at the end of this section.

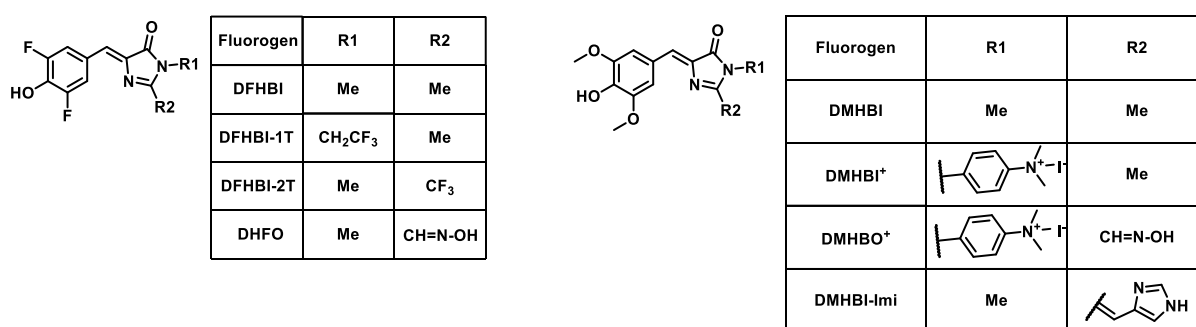
### 2.2.1. Hydroxybenzylidene imidazolinone (HBI)-binding aptamers

In 2011, an aptamer named Spinach has been developed and boosted RNA live mammalian cell imaging.<sup>61</sup> Its cognate fluorogenic dye (fluorogen) 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI, figure 5) is structurally derived from the chromophore of GFP active site.<sup>61</sup> Once Spinach tightly binds DFHBI the fluorescent signal is produced, closely mimicking the green emission of eGFP.<sup>61</sup> Importantly, DFHBI is exclusively in the phenolate form due to electron-withdrawing effect of two fluorine atoms in aromatic ring (Figure 5). It is known that phenolate form in eGFP exhibits a higher extinction coefficient, therefore contributing to its increased brightness.<sup>61</sup> To show capacity of Spinach to tag RNAs in live cells, Paige *et al.* genetically expressed Spinach fused to the 3' end of 5S ribosomal RNA (rRNA) in HEK293T cells. Green fluorescence of 5S-Spinach was distributed throughout the cell in the same way as 5S, while sucrose treatment induced the formation of RNA granules and accumulation of 5S-Spinach in them.

Using systemic mutagenesis, the same group generated Spinach2 with improved thermostability and folding capacity.<sup>66</sup> Spinach2 was applied to image the localization of CGG repeat-containing RNAs in COS-7 cells. Their studies showed that the RNA component of the nuclear foci is highly dynamic and undergoes morphological changes during cell division.

Using chemical diversity approach, the Jaffrey group generated fluorogens based on DFHBI scaffold.<sup>65</sup> The introduction of trifluoroethyl moiety at N1 position of the imidazolone ring resulted in DFHBI-1T fluorogen, while the introduction of trifluoromethyl group at C2 position resulted in DFHBI-2T (Figure 5). When bound to Spinach2 DFHBI-1T exhibited a 35 nm red shift in excitation maximum, slight red shift in emission maximum and increased brightness.<sup>65</sup> At the same time, DFHBI-2T/Spinach2 showed 53 nm red shift in excitation and 22 nm red shift in emission, but with a reduced brightness. Interestingly, DFHBI-1T reduced the background fluorescence in the microscopy experiments while DFHBI-2T allowed Spinach2 imaging using YFP filter (an excitation bandpass filter  $500 \pm 10$  nm light, a dichroic mirror at 515 nm and emission filter  $535 \pm 15$  nm light). This 'plug-and-play' approach allowed modulating the spectral properties of Spinach2 aptamer by simply switching fluorogens. Moreover, this approach allowed on-demand spectral property change within a single experiment by washing one fluorogen and replacing to another.

In 2014 Jeffrey group selected an aptamer named Broccoli which showed a better cellular performance in comparison to Spinach2.<sup>63</sup> Identification and selection of Broccoli was possible due to novelty in SELEX analysis approach. The fundamental problem of SELEX for screening RNA aptamer-fluorogen pairs is that the selection process is based only on binding event, but not on the activation of fluorescence. Coupling SELEX to fluorescence-activated cell sorting (FACS) allowed direct visualization of expressed RNAs in bacteria and further selection based on the fluorescence induction resulted in isolation of Broccoli.<sup>63</sup> Broccoli is shorter, more thermostable, performs better in cells. It does not require tRNA scaffold like Spinach to properly fold in cellular environment.<sup>63</sup>



**Figure 5.** Chemical structures of fluorogens based on HBI scaffold.

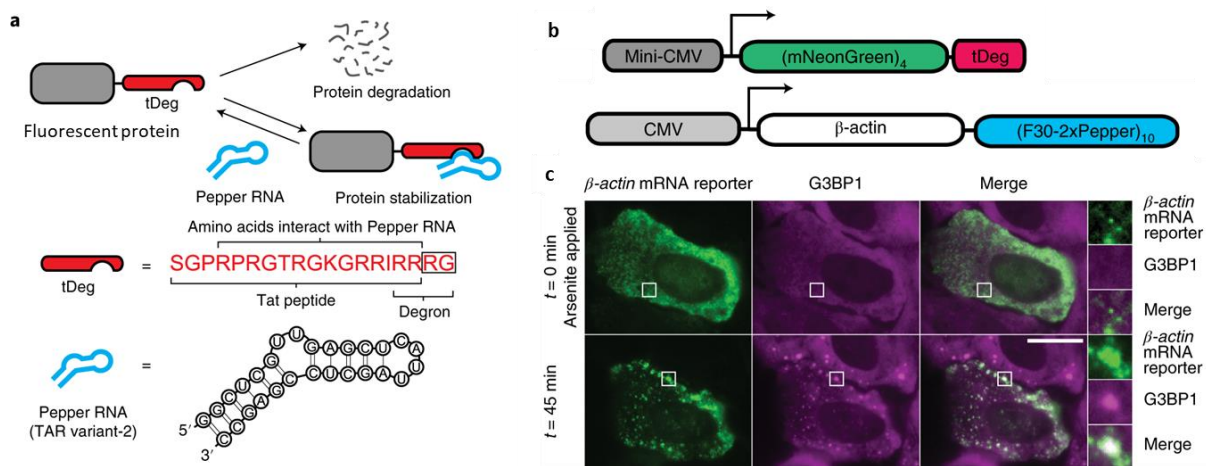
Recently, the same group has developed an aptamer named Corn that recognizes its cognate fluorogen DFHO (3,5-difluoro-4-hydroxybenzylidene imidazolinone-2-oxime) (Figure 5).<sup>67</sup> Upon binding to Corn, DFHO emitted yellow fluorescence ( $\lambda_{exc} = 518$  nm;  $\lambda_{em} = 545$  nm). Improved photostability compared to Spinach and Broccoli and reduced background fluorescence enabled quantification of Pol III RNA transcription activity under modulation of mTOR inhibitors.<sup>67</sup>

Since the screening method plays very important role in analysis of fluorogen-aptamer libraries, Ryckelynck and colleagues took innovative high-throughput analytical approach based on microfluidic-assisted *in vitro* compartmentalization ( $\mu$ IVC).<sup>64</sup> Using  $\mu$ IVC methodology, the authors focused on applying selection pressure to identify salt-independent and temperature resistant variant of Spinach.<sup>64</sup> Indeed, powerful microfluidic-assisted *in vitro* evolution allowed isolation of an aptamer named iSpinach.<sup>64</sup> Interestingly, iSpinach also had a better binding affinity to DFHBI than Spinach2 (0.92  $\mu$ M for iSpinach and 1.45  $\mu$ M for Spinach2).<sup>64</sup>



Explorations in chemical diversity of 3,5-dimethoxy-substituted HBI (DMHBI, figure 5) derivatives by Höbartner group led to the discovery of cationic analogues, DMHBI<sup>+</sup> and DMHBO<sup>+</sup> (Figure 5).<sup>68</sup> The authors aimed at developing a new fluoromodule with large Stokes shifts (LSS). LSS allows reducing the reabsorption of the emitted light, application of the module as FRET donor for far-red fluorescence reporter and possibility to tune the color emission. Höbartner and colleagues turned their attention to the 13-2<sup>61</sup> aptamer originally reported to light up fluorescence of DMHBI. The directed evolution of the 13-2<sup>61</sup> aptamer resulted in identification of a multicolor LSS aptamer named Chili.<sup>68</sup> The Chili RNA aptamer in complex with DMHBI<sup>+</sup> and DMHBO<sup>+</sup> imitated green and red LSS fluorescent proteins like LSSmOrange and LSSmKate2. Chili was the first example of HBI-binding aptamers recognizing cationic HBI-derivatives.<sup>68</sup> Moreover, DMHBO<sup>+</sup>/Chili fluoromodule is the most red-shifted wavelength-emitting (592 nm) among the HBI-binding aptamers.

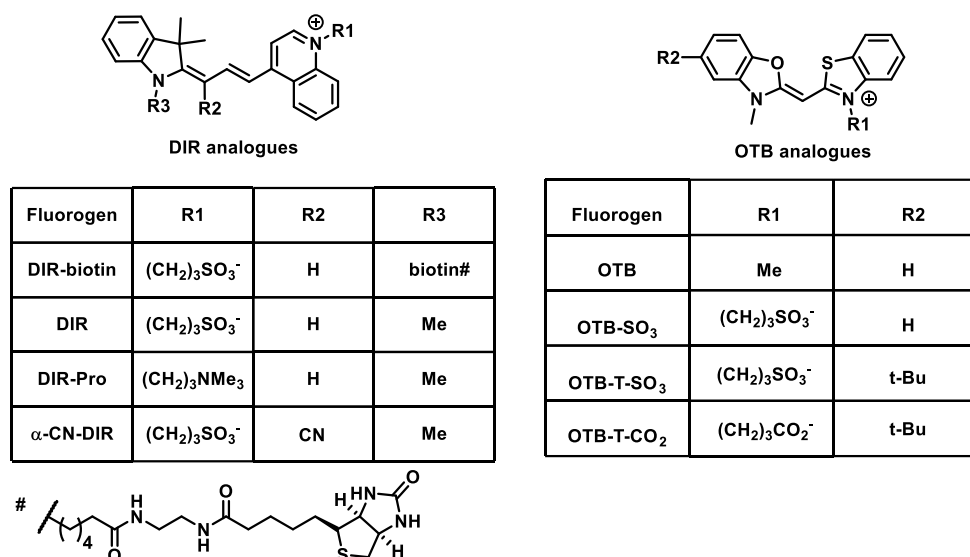
One of the major problem with fluorogenic RNA aptamers is that their low expression in mammalian cells limits their detection. Litke and Jaffrey have very recently addressed this issue by developing the Tornado expression system (Twister-optimized RNA for durable overexpression).<sup>69</sup> Tornado is based on circularization of 5' and 3' ends of RNA of interest enabling improved stability in cellular environment. Tornado significantly improved stability of Broccoli<sup>63</sup> and Corn<sup>67</sup> aptamers expanding their application in a range of cell lines (HepG2, HeLa, COS-7) and metabolite imaging.<sup>69</sup> Almost at the same time, Jaffrey group revised aptamer-based technology and turned it to fully genetically encoded system by developing Pepper-aptamer binding to proteins to turn on their fluorescence that otherwise are not emissive (Figure 6).<sup>70</sup> Though primarily Pepper-aptamer was developed to transform fluorescent proteins into fluorogenic, its potential is beyond that, e.g. conditional control of protein expression.



**Figure 6.** Imaging Pepper-tagged mRNA. **(a)** Scheme demonstrating a Pepper RNA-regulated protein destabilization domain, tDeg. When Pepper was fused to RNA of interest, it stabilized FP fused to tDeg in cells allowing imaging RNA target; **(b)** DNA plasmid constructs used for imaging  $\beta$ -actin mRNA in living cells; **(c)** Cells coexpressing this  $\beta$ -actin mRNA reporter and a green fluorogenic protein, (mNeonGreen)<sub>4</sub>-tDeg, were imaged before and 45 min after arsenite (500  $\mu$ M) treatment to induce stress granules. White boxes correspond to regions magnified on the right. Scale bar, 20  $\mu$ m. Figure was adapted from ref.<sup>70</sup>

### 2.2.2. Cyanine-binding aptamers

Armitage and co-workers exploited fluorogenic properties of unsymmetrical cyanines (Figure 7) to select an RNA aptamer.<sup>71</sup> The diminished fluorescence of such cyanines is caused by twisted rotation around methine bridge in excited state, therefore leading to nonradiative energy dissipation. In particular, they targeted dimethylindole red (DIR, figure 7) dye since its emission is in the red region of spectra and its bulky structure as well as the anionic tail decreased specificity to DNA. After 15 rounds of selection 32 RNA aptamers increased the fluorescence of DIR.<sup>71</sup> Further analysis led to discovery of the DIR-Apt1 sequence as a promising candidate.<sup>71</sup> Though authors did not study the DIR/DIR-Apt1 pair behavior in cell models, they showed the importance of rational design of fluorogens and the power of *in vitro* selection to develop fluorogen-RNA aptamer fluromodules.



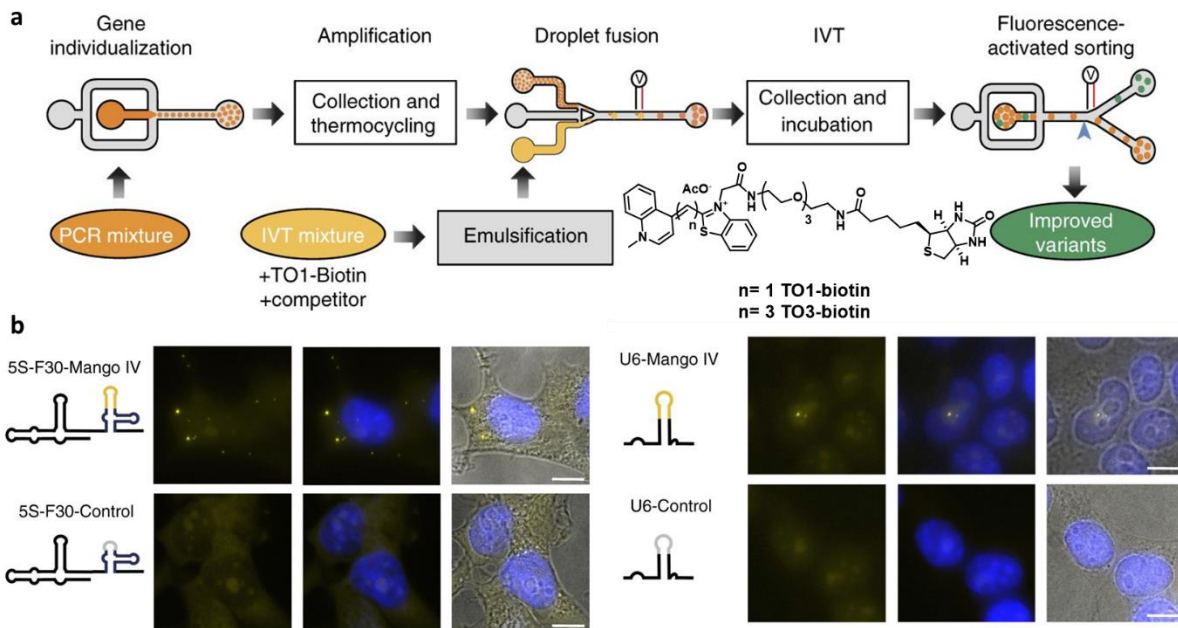
**Figure 7.** Chemical structures of DIR and OTB fluorogens for DIR-Apt1 and DIR2s-Apt aptamers.

In 2017, the same group selected a promiscuous DIR2s-Apt aptamer capable of binding DIR fluorogenic dye and its analogues with a red emission, as well as oxazole thiazole blue (OTB) analogues with blue emission (Figure 7).<sup>72</sup> To direct the fluoromodule to a specific target, the authors developed a bifunctional aptamer by fusing DIR2s-Apt to anti-EGFR RNA aptamer (EGFR-DIR2s). This bifunctional aptamer EGFR-DIR2s allowed to discriminate high and low level of expression of EGFR receptors in A431 and HEK293 cells respectively.

### 2.2.3. Mango aptamers

In addition to GFP-mimicking family aptamers, an aptamer named Mango with impressively high affinity ( $K_d=3.2$  nM, table 1) to thiazole orange (TO1, figure 8a) was developed.<sup>73</sup> TO1 was previously known to bind DNA quadruplex structures.<sup>74</sup> TO1 cyanine is a cationic dye usually associated to a tosylate as counterion. To reduce specificity towards DNA, counterion exchange from tosylate to acetate was performed and resulted in TO1-biotin molecule (Figure 8a). Once bound, Mango enhanced fluorescence of TO1-biotin 1100-folds. Formation of the complex caused ~10 nm red shift in excitation and emission in comparison to unbound TO1-biotin ( $\lambda_{exc} = 510$  nm;  $\lambda_{em} = 535$  nm). To explore substrate specificity of Mango, Dolgosheina *et al.* synthesized red-shifted thiazole orange analogue TO3-biotin (Figure 8a).<sup>73</sup> Indeed, Mango was able to recognize TO3-biotin with  $\lambda_{exc} = 637$  nm;  $\lambda_{em} = 658$  nm. It is worth

mentioning that TO1-biotin/Mango fluoromodule has the highest reported affinity with  $K_d = 3.2$  nM, and its analogue TO3-biotin/Mango has slightly higher value  $K_d = 6-8$  nM. Overexpression of Mango in BL21 (DE3) *E. coli* cells and treatment with TO1-biotin enabled detection of the complex by flow cytometry.<sup>73</sup> Expression of incorporated Mango to 6S transcriptional RNA and further RNA extraction allowed monitoring 6S RNA binding to RNA polymerase. Additionally, TO1-biotin/Mango fluoromodule was visualized in *C. elegans* worm after injection of the complex.<sup>73</sup>



**Figure 8.** (a) Experimental scheme for microfluidic-assisted fluorescence screening. Ovals and boxes represent on- and off-chip steps, respectively. Three microfluidic devices were used for gene individualization in 2.5 pL droplets containing PCR mixture; after thermocycling, fusing each PCR droplet with a droplet containing an in vitro transcription (IVT) mixture supplemented with TO1-Biotin and competitor and, after incubation, analysing the fluorescence profile of each droplet and sorting them accordingly; (b) Cellular imaging of Mango IV tagged RNAs (5S, U6) stained with TO1-Biotin (yellow) and Hoechst 33258 (blue); construct diagrams shown as RNA-Mango (yellow) or non-fluorescent control RNA (grey), F30 folding scaffold (blue) and remaining RNA sequence (black). Figure was adapted from ref. <sup>75</sup>

Recently, Autour *et al.* have isolated Mango II, III and IV with improved fluorescence properties, binding affinity and salt dependence compared to previously reported Mango.<sup>75</sup> Development of the new aptamers was possible due to application of new screening approach of droplet-based microfluidic-assisted *in vitro*

compartmentalization coupled to SELEX (Figure 8a). Improved properties of new Mango aptamers allowed fluorescence imaging of small non-coding 5S and U6 RNAs in mammalian HEK293T live cells (Figure 8b).

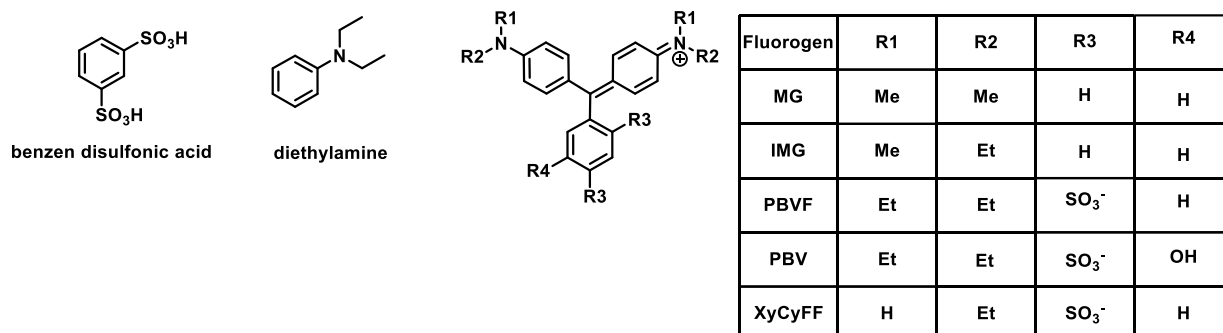
Despite improved performance, Mango family possesses disadvantages like potassium-dependent parallel-stranded G-quadruplex structure that tends to misfold in eukaryotic cells.<sup>76,77</sup>

#### 2.2.4. Malachite Green- and rhodamine-binding aptamers.

Historically, malachite green aptamer (MGA) was the first example of the fluorogenic aptamers reported. Initial purpose of its development was to control site-specific RNA cleavage caused by generation of reactive oxygen species under laser illumination.<sup>78</sup> Later, Tsien and colleagues tuned MGA to RNA-imaging tag system.<sup>60</sup> Malachite green (MG, figure 9) dye is fluorogenic since its free vibrational rotation causes deexcitation, while in viscous environment the motion is restricted causing planar structure and thus efficient energy transfer to illuminate the fluorescence. Therefore, Babendure *et al.* showed that MGA stabilized planar form of MG dye to increase its fluorescence up to 2360 folds.<sup>60</sup> Interestingly, MGA was shown to be very selective to MG since it did not bind triphenylmethane dyes with sulfonate groups.<sup>60</sup> For a very long time, MG/MGA module was shown to be inefficient in bacterial and mammalian cells until recently Kim and colleagues optimized MGA scaffold and multimerized MGA to image cellular RNAs.<sup>79</sup> Shortcoming like phototoxic effect of MG prevents wide usage of MG/MGA pair, but it has boosted the field of fluorogen-aptamer research to develop new systems.

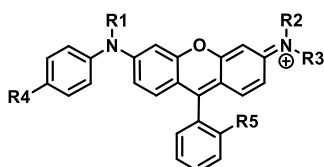
Complementary to MGA at the same time, the Wilson group has identified sulforhodamine B (SRB) – binding aptamer named SRB-2.<sup>80</sup> From the pool of  $5 \times 10^{14}$  RNA molecules with 114 nucleotides in length and randomized in 72 positions, two clones were isolated to bind SRB (SRB-1 and SRB-2 aptamers). Detailed structure analysis revealed SRB-2 aptamer had a better ligand-binding specificity than SRB-1 aptamer. Surprisingly, in competition assay either SRB precursors (benzene disulfonic acid, diethylaniline, figure 9) and xylene cyanole FF (XyCyFF, SRB analogue without oxygen in xanthene system, figure 9) did not significantly disturb SRB/SRB-2 complex suggesting their poor affinity to SRB-2 aptamer. Structure-activity relationship studies led to conclusion of synergic effect of planarity of xanthene system and the presence

of negative charge in the lower ring of the dye for efficient dye-aptamer interaction.<sup>80</sup> Within the same work, orthogonal to SRB-2, fluorescein-binding aptamer was also identified, FB-1 aptamer.<sup>80</sup> *In vitro* dual-labeling test showed that SRB-2 and FB-1 can be efficiently discriminated due to their selectivity towards corresponding dye. Additionally to MGA characterization, Tsien and colleagues conducted the study of SRB-2 substrate specificity.<sup>60</sup> They showed that SRB-2 can also bind SRB analogues, patent blue VF (PBVF) and patent blue violet (PBV) (figure 9).



**Figure 9.** Chemical structure of Malachite green (MG) and its analogues.

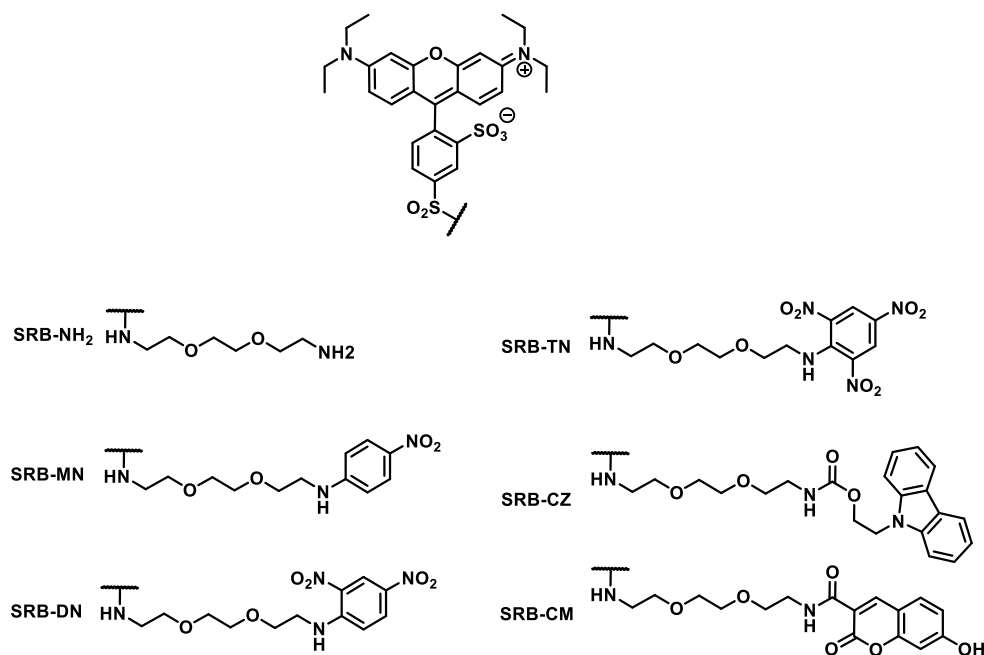
In 2010, Rao group designed and synthesized a PeT-quenched SRB variant, ASR-1 probe (Figure 10).<sup>81</sup> A non-fluorescent ASR-1 probe design included modification of alkyl moiety in dialkylamino group in xanthene ring to aryl moiety. It was hypothesized that the fluorescence of ASR-1 probe would be activated upon binding to RNA. From the pool of 100mer RNA sequences containing 60 randomly mutated positions and after 3 rounds of SELEX, 35 colonies of *E. coli* bacteria expressing RNA aptamers showed fluorescent signal.<sup>81</sup> Further analysis led to identification of Apt10L aptamer.<sup>81</sup> To improve binding affinity a systematic truncation and mutation study was performed. Surprisingly, evolution of Apt10L aptamer led to significant loss of the affinity to ASR-1 probe, therefore a chemical approach in designing ASR-1 analogues was used. Systematic substitution of groups led to ASR-7 probe (Figure 10) with significantly improved binding affinity ( $K_d = 1.2 \mu\text{M}$ ) as well as fluorogenic properties.



Fluorogen	R1	R2	R3	R4	R5
ASR-1	Me	Me		H	SO <sub>3</sub> <sup>-</sup>
ASR-2	Me	Et	Et	H	SO <sub>3</sub> <sup>-</sup>
ASR-3	Me	Et	Et	H	SO <sub>3</sub> <sup>-</sup>
ASR-4	Me	Me		OMe	SO <sub>3</sub> <sup>-</sup>
ASR-5	Me	Me		H	SO <sub>3</sub> <sup>-</sup>
ASR-6	Me	Et	Et	H	SO <sub>2</sub> NMe <sub>2</sub>
ASR-7	Me	Me			SO <sub>3</sub> <sup>-</sup>

**Figure 10.** Chemical structures of SRB analogues for Apt10L aptamer.

In 2013, Jaschke group rationally designed the fluorogenic molecule by conjugating a quencher to SRB dye.<sup>62</sup> Close proximity of a quencher and SRB-dye forced the formation of the non-fluorescent complex due to  $\pi$ -stacking of aromatic systems. The fluorescence is restored upon binding to a fluorophore-binding aptamer SRB-2.<sup>62</sup> Screening a number of quenchers including *p*-nitrobenzylamine (MN), dinitroaniline (DN), trinitroaniline (TN), carbazole (CZ), and 7-hydroxycoumarin-3-carboxylic acid (CM) led to a finding of DN as the best contact quencher (Figure 11). After characterization of spectral properties, SRB-DN/SRB-2 fluoromodule was further evaluated in capacity to report *in vitro* RNA transcription under T7-promoter activity. Also, SRB-2 aptamer with tRNA scaffold was genetically expressed in bacteria and induced orange fluorescence after brief treatment with SRB-DN molecule.



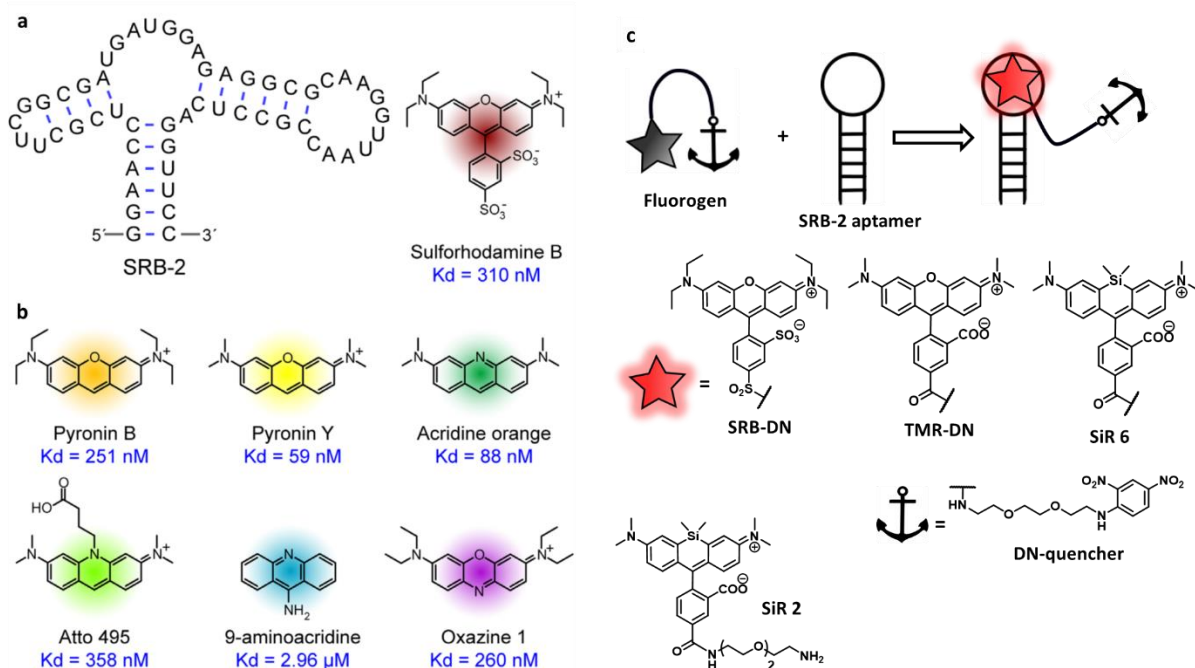
**Figure 11.** Structures of SRB-quencher conjugates.

Further studies of Jaschke group dealt with a systematic screening of the influence of fluorophore variation and recognition pattern by SRB-2 aptamer.<sup>82</sup> Taking into account the importance of the xanthene system and the sulfonate group in the structure of the to be recognized by SRB-2 aptamer, Sunbul and Jaschke analyzed a panel of dyes (Pyronin B, Pyronin Y, Acridine orange, ATTO495, 9-aminoacridine, Oxazine) (figure 12b).<sup>82</sup> They found out that SRB-2 aptamer could tolerate dye structure variability.<sup>82</sup> Interestingly, despite previously reported importance of preserving sulfonate in lower aromatic ring of SRB, unexpectedly, they showed a negligible role of sulfonate in binding. Moreover, replacement of diethylamino to dimethylamino groups on xanthene system led to increase of the SRB-2 aptamer-binding affinity, while their removal led to dramatic loss of the affinity. This study has identified a new candidate based on 5-carboxytetramethylrhodamine fluorophore (TMR). To turn the fluorophore to fluorogenic, TMR was conjugated to DN-quencher (TMR-DN, figure 12c). TMR-DN was shown to better perform than SRB-DN in bacterial and mammalian cells expressing multimerized SRB-2 aptamer fused to 5S rRNA and embedded in a tRNA scaffold.

Recently, Wirth *et al.* reported a silicon rhodamine (SiR)-binding aptamer, SiRA, compatible for super-resolution fluorescence RNA imaging in live cells.<sup>83</sup> SiR is a particularly attractive fluorophore since it is bright, photostable and emits in the far-red region ( $\lambda_{exc} = 649 \text{ nm}$ ;  $\lambda_{em} = 666 \text{ nm}$ ). Its fluorogenic behavior is caused by



environment-dependent intramolecular spirocyclization. In particular, SiR exhibited equilibrium between fluorescent (zwitterion) form in polar media and poorly fluorescent form (cyclic spirolacton) in apolar environment (Figure 18a). This set of properties made SiR a promising candidate for SELEX. After 14 rounds of in vitro selection from a RNA pool of  $2 \times 10^{15}$  and a standard length of 103 nucleotides (nt), 21 sequences were isolated to increase SiR fluorescence. Detailed bioinformatics analysis revealed a three-way junction structure. It was also shown that truncation of initial sequence of 103 nt to 50 nt conserved the aptamer capacity to bind SiR.<sup>83</sup> Interestingly, a short PEG-linker of SiR2 (Figure 12c) played a role in binding affinity to SiRA aptamer.<sup>83</sup> SiR2/SiRA complex was reported to be sensitive to temperature variabilities, but retained its ~80% fluorescence at 37°C which is usually used during live-cell imaging. Additionally, its signal remained at the same level within physiological range of  $Mg^{2+}$  concentrations. Interestingly, SiRA aptamer possessed promiscuity property capable of binding TMR and carborhodamine (CR), but with a considerably lower affinity ( $K_d$  to TMR 3.4  $\mu M$ , to CR 0.9  $\mu M$ , to SiR2 0.43  $\mu M$ ). To gain turn-on fluorescence effect, Wirth *et al.* synthesized a dual-fluorogenic probe SiR6 (Figure 12c) by attaching DN-contact quencher to SiR.<sup>83</sup> Though the fluorescence enhancement of SiR6 reached 17-folds, this modification significantly lost the binding affinity to SiRA ( $K_d$  to SiR6 1  $\mu M$ ) since a higher energy was required to dissociate fluorophore and the quencher. Since SiR2/SiRA showed the optimal properties, it was chosen for further validation. SiRA was embedded in tRNA scaffold under T7 promoter control and visualized in *E. coli* cells. Additionally, multimerized version (5 repeats) of SiRA could report *gfp* mRNA when it was inserted into 3'UTR untranslated region. SiR2/SiRA system was the first example of aptamer-based technology compatible for super-resolution RNA imaging.



**Figure 12.** (a) Predicted secondary structure of SRB-2, chemical structure of its cognate dye SRB and its dissociation constant (Kd). SRB-2 is a promiscuous RNA aptamer and can bind to xanthene dyes; (b) Chemical structures of substrates for SRB-2 aptamer. Adopted from ref. <sup>82</sup> (c) Chemical structures of contact-mediated quenching probes (SRB-DN, TMR-DN, SiR6) for SRB-2 aptamer and SiR2 as a fluorogen for SiRA aptamer.

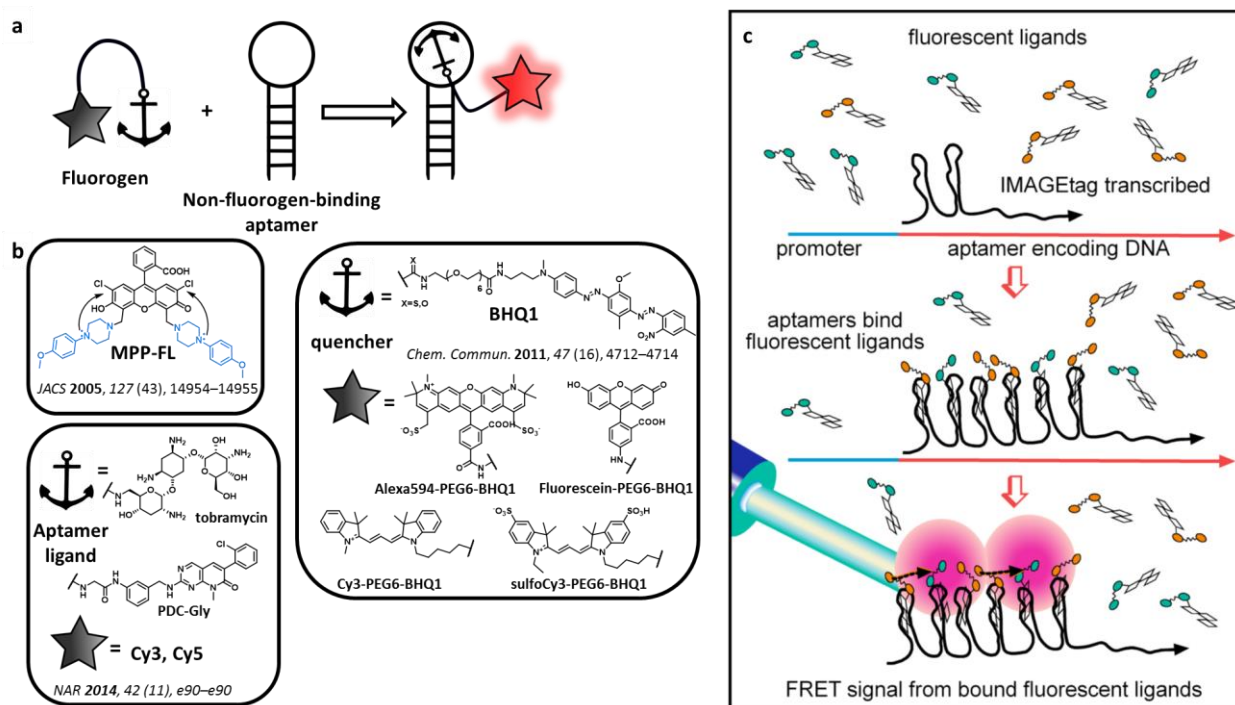
### 2.2.5. Non-fluorogen-binding aptamers

Photoinduced Electron Transfer (PeT) process is commonly used for designing fluorogenic biosensors.<sup>84</sup> Typically, such sensors consist of quencher as donor of electrons. Sparano and Koide reasoned if RNA aptamer can be selected towards binding the quencher so that it would restore the emission of the pro-fluorescent molecule.<sup>85,86</sup> The main advantage of such fluorogenic aptamers is that the system would be highly modular since it is independent of the fluorophore of choice. To this end, the group has isolated the aptamer named Apt1 after 8 rounds of selection.<sup>86</sup> Apt1 was able to bind a quencher *N*-(*p*-methoxyphenyl)piperazine (MPP). Two MPP groups were conjugated to dichlorofluorescein to yield fluorogenic molecule MPP-FL (Figure 13b).<sup>86</sup> Apt1 aptamer increased the fluorescence of MPP-FL probe by 13-folds (1  $\mu$ M of fluorogen and 100  $\mu$ M of the aptamer). Even though MPP-FL/Apt1 was not shown to be compatible in cellular application, this work was the first proof-of-concept

example of alternative strategy in designing PeT fluorogenic molecules and selection of aptamers binding quenchers for RNA imaging.

In 2011, Uesugi and colleagues selected another quencher-binding aptamer A1.<sup>87</sup> The aim of the work was to develop an aptamer capable to bind a black hole quencher (BHQ1, figure 13b). For this reason, three BHQ1-fluorophore conjugates were synthesized using Alexa594, sulfoCy3 and fluorescein dyes (Figure 13b).<sup>87</sup> After 13 rounds of RNA selection, A1-A4 sequences binding BHQ1 were isolated. Detailed analysis of A1-A4 variants properties (affinity to BHQ1-fluorophore conjugates and turn-on effect) revealed A1 as the best candidate. Interestingly, fluorogenic molecules responded differently once bound to A1 aptamer, in particular the emission of quenched sulfoCy3-PEG6-BHQ1 (0.3  $\mu$ M) was restored 11.3-fold by A1 (100  $\mu$ M) while the signal of Alexa594-PEG6-BHQ1 and fluorescein-PEG6-BHQ1 was enhanced only 6.3- and 2.5-fold, respectively. Alexa594-PEG6-BHQ1 and sulfoCy3-PEG6-BHQ1 were further tested to evaluate their capacity to monitor *in vitro* transcript levels under T7 promoter activity. Thus, both conjugates showed fluorescence increase in respect to the time-dependent A1 aptamer expression level.<sup>87</sup> Moreover, Alexa594-PEG6-BHQ1 enabled detection of GFP-coding mRNA transcripts in cell extracts.<sup>87</sup> Later on, the same group improved the cell permeability of BHQ-fluorophore by removing sulfonate groups in a cyanine (from sulfoCy3 to Cy3, figure 13b).<sup>88</sup> This probe enabled the detection of  $\beta$ -actin mRNA labeled with RT-Apt.<sup>88</sup>

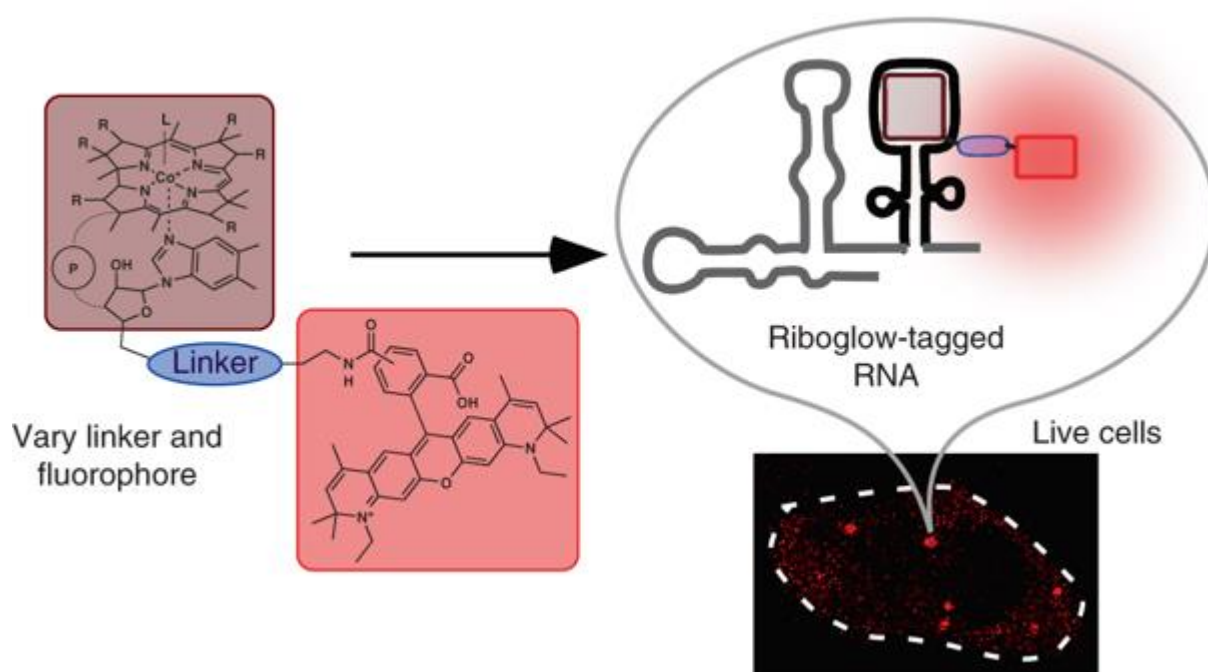
In 2014 Shin *et al.* has proposed an alternative RNA reporter system for live-cell imaging named IMAGEtag (Intracellular MultiAptamer GENetic tag).<sup>89</sup> The reporters utilized strings of RNA aptamers (tobramycin, neomycin and PDC) expressed from GAL1-promoter.<sup>89</sup> The authors aimed to create a FRET system by coupling aptamer ligands to a fluorophore (Cy3 and Cy5).<sup>89</sup> IMAGEtag was validated in yeast.<sup>89</sup> After galactose-induced activation of GAL1-promoter the transcribed aptamers could bind their cognate fluorescent ligands resulting in FRET signal. IMAGEtag system found a very interesting application in evaluation heterogeneous Pol II promoter activity in living cells.<sup>89</sup> To show this, three promoters were used, GAL1, ACT1 and ADH1.<sup>89</sup>



**Figure 13.** Non-fluorogen binding aptamers. **(a)** Aptamer selectively binds to its cognate substrate (PeT quencher or promoter ligand); **(b)** Chemical structures of probes for non-fluorogen-binding aptamers; **(c)** Schematic diagram of the IMAGETag system. IMAGETags (Intracellular Multiaptamer GENetic tags) are tandemly repeated aptamers that are transcribed in the cells as an mRNA construct from a promoter of choice. IMAGETags bind fluorescently labeled ligands that are labeled with one of the two members of a FRET pair and the bound ligands are visualized by FRET. Adopted from <sup>89</sup>.

In 2018, Braselmann *et al.* turned to bacterial riboswitch RNA capable of binding to its natural ligand (cobalamin, Cbl) and developed multicolor RNA imaging platform namely Riboglow (Figure 14).<sup>90</sup> It is known that Cbl can efficiently quench fluorescence of conjugated dyes.<sup>90</sup> Taking advantage of natural properties, a series of Cbl-derivatives coupled to fluorophores (fluorescein, ATTO488, ATTO590, ATTO633 and Cy5) were developed.<sup>90</sup> Taking into consideration the effect of the fluorophore and the tethered linker between Cbl and the dye, the systemic analysis showed that green emitting dyes tended to be more efficiently quenched, while the increasing the linker length promoted the fluorescence dequenching producing higher background signal. Cbl-5×PEG-ATTO590, Cbl-4×Gly-ATTO590 and Cbl-Cy5 were subjected to further cellular validation. Riboglow was tested in imaging the localization of  $\beta$ -actin mRNA

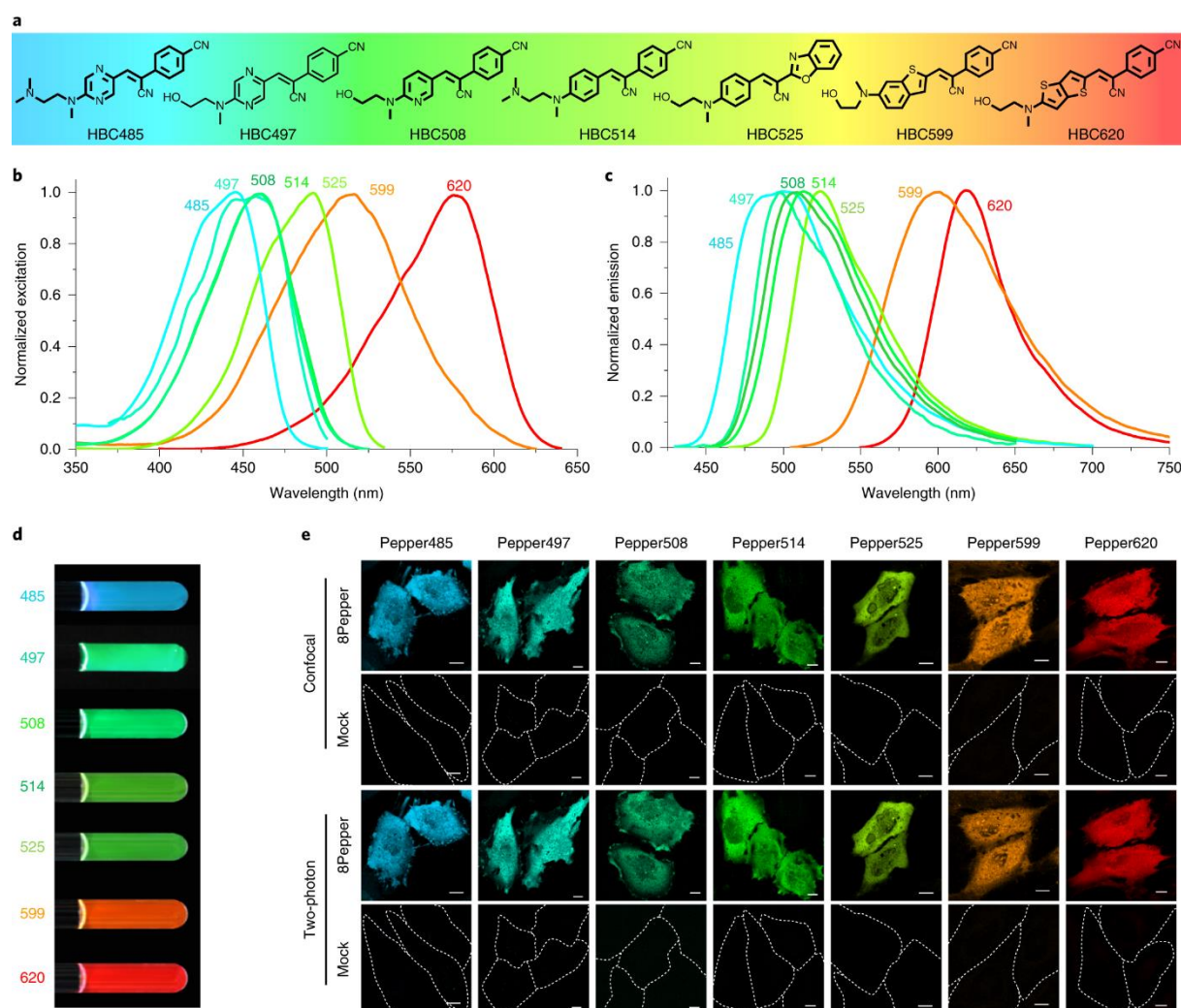
(encoded as *ACBT*) within stress granules in arsenite-treated live cells.<sup>90</sup> Indeed, Cbl-Cy5 or Cbl-4x Gly-ATTO590 signal appeared in stress granules, but only when at least four consecutive tags of A were used. Additionally, Riboglow was used to visualize U1 small noncoding RNA in U-bodies of live cells. Despite its superior performance in comparison to Broccoli aptamer and MS2 system, it posed some drawbacks.<sup>90</sup> Firstly, *in vitro* characterization showed a weak activated fluorescence in comparison to Broccoli.<sup>90</sup> Secondly, Cbl-fluorophore conjugates are poorly permeant, thus required bead-loading for the delivery of the fluorogen.<sup>90</sup> Additionally, authors mentioned the natural occurrence of Cbl-binding proteins in cells, therefore raising awareness of false-positive fluorescent signal from Riboglow.<sup>90</sup> Nevertheless, Riboglow was shown to have a good imaging contrast (3-4 folds fluorescence enhancement) in cellular experiments and useful as orthogonal modular approach for RNA imaging in mammalian cells.



**Figure 14.** Riboglow utilized bacterial riboswitch RNA binding Cbl. Cbl conjugated to a dye quenched its fluorescence, but the emission was restored once Cbl moiety was bound to the riboswitch. Figure is taken from graphical abstract of ref.<sup>90</sup>

Chen, Zhang, Su *et al.* have very recently reported a new very powerful RNA aptamer named Pepper.<sup>91</sup> Pepper exhibited promiscuity capable of binding the cyanophenyl-acetonitrile analogues (HBC, figure 15a). HBC are typical twisted ICT molecules non-fluorescent in aqueous solutions, but emissive in rigidified environment like Pepper

RNA aptamer or constraining PMMA glass.<sup>91</sup> By expanding  $\pi$ -conjugation of HBC structure, they synthesized a palette of HBC probes (Figure 15a) for Pepper aptamer spanning their emission from cyan to red upon binding (Figure 15b). Impressively, HBC/Pepper system had a very broad range of applications, notably diverse RNA targets, usability in different cell lines, compatibility with 2-photon excitation and super-resolution microscopy.<sup>91</sup> Pepper enabled real-time, high-throughput quantification of RNA level in live cells, tracking of protein–RNA tethering and imaging genomic loci with CRISPR display.<sup>91</sup> Off the note, as previously mentioned in section 2.2.1 Jaffrey group developed the aptamer<sup>70</sup> activating fluorescence of proteins once bound. Surprisingly, the aptamer held the same name Pepper, as the papers were published approximately at the same time in different high-rank journals.



**Figure 15.** Expanding spectral range with HBC/Pepper module. (a) Structures of HBC analogues; (b) Excitation and (c) emission spectra of Pepper complexed with HBCs; (d) Photograph of Pepper–fluorogen complexes (10  $\mu$ M) illuminated with UV light; (e)



Imaging of HeLa cells expressing 8xPepper (8 repeats) and incubated with various HBCs. Cells transfected with the empty vector were used as controls. Scale bars, 10  $\mu\text{m}$ . Figure was taken from ref. <sup>91</sup>

Properties of described fluorogen/aptamer modules are summarized in the table 1.

**Table 1.** Properties of RNA-based fluorogenic modules.

Fluorogen/aptamer	$\lambda_{\text{ex/em}}$ max (nm)	$\epsilon$ ( $\text{M}^{-1}\cdot\text{cm}^{-1}$ )	QY	Brightness <sup>a</sup>	$K_d$ (nM)	Ref.
<b><i>HBI-binding aptamers</i></b>						
DFHBI/Spinach	469/501	24,300	0.72	17,500	540	61
DFHBI-1T/Spinach2	482/505	31,000	0.94	29,100	560	65
DFHBI-2T/Spinach2	500/523	29,000	0.12	3,480	1,300	65
DFHBI-1T/Broccoli	472/507	29,600	0.94	27,800	360	63
DFHO/Corn	518/545	29,000	0.25	7,250	70	67
DFHBI/iSpinach	442/503	n.r.	n.r.	n.r.	920	64
DMHBI/Chili	400/537	~20,000	0.08	1,600	570	68
DMHBI <sup>+</sup> /Chili	413/542	~26,000	0.40	10,400	63	68
DMHBI-Imi/Chili	463/545,594*	~8,000	0.08	640	71	68
DMHBO <sup>+</sup> /Chili	456/592	~25,000	0.10	2,500	12	68
<b><i>Cyanine-binding aptamers</i></b>						
DIR/DIR-Apt1	602/650	n.r.	n.r.	n.r.	86	71
DIR/DIR2s-Apt	600/646	150,000	0.26	39,000	708	72
DIR-Pro/DIR2s-Apt	600/658	164,000	0.33	54,000	313	72
OTB-SO <sub>3</sub> /DIR2s-Apt	380/421	73,000	0.51	37,000	739	72
OTB-T-SO <sub>3</sub> /DIR2s-Apt	380/421	71,000	0.17	12,000	1193	72
<b><i>Mango aptamers</i></b>						
TO1-biotin/ Mango	510/535	77,500	0.14	11,000	3.2	73
TO3-biotin/ Mango	637/658	9,300	n.r.	n.r.	6	73
TO1-biotin/ Mango II	510/535	n.r.	0.56	n.r.	0.7	75
TO1-biotin/ Mango III	510/535	~77,800	0.56	43,000	5.6	75
TO1-biotin/ Mango IV	510/535	~76,000	0.42	32,000	11.1	75
<b><i>Malachite Green- and rhodamine-binding aptamers</i></b>						
MG/MGA	630/650	150,000	0.19	28,500	117	60,79
IMG/MGA	n.r./~670	n.r.	0.32	n.r.	666	60

PBVF/SRB-2	n.r./~675	n.r.	0.03	n.r.	86,000	60
PBV/SRB-2	n.r./~665	n.r.	0.04	n.r.	23,000	60
SRB/SRB-2	565/584	n.r.	0.19	n.r.	310	78
ASR-1/Apt10L	555/610	75,000	0.65	48,750	39,100	81
SRB-DN/SRB-2	579/596	85,200	0.65	55,380	1,340	62
TMR-DN/SRB-2	564/587	90,500	0.33	29,865	35	82
SiR2/SiRA	649/662	86,000	0.98	84,280	430	83
<b>Non-fluorogen-binding aptamers</b>						
Alexa594-PEG6-BHQ1/A1	570/590-695	n.r.	n.r.	n.r.	4,700	87
Cy3-PEG6-BHQ1/A1	520/540-645	n.r.	n.r.	n.r.	n.r.	87
Fluorescein-PEG6-BHQ1/A1	470/490-645	n.r.	n.r.	n.r.	n.r.	87
Cy3-tobramycin	~540/570	n.r.	n.r.	n.r.	1,400	89
Cy5-tobramycin	~620/680	n.r.	n.r.	n.r.	1,000	89
Cbl-5xPEG-Atto590/A (Riboglow)	594/624	120,000	0.31	37,200	34	90
Cbl-5xPEG-Atto590/D (Riboglow)	594/624	120,000	0.31	37,200	3	90
Cbl-Cy5/A (Riboglow)	640/656-800	271,000	0.26	70,460	n.r.	90
Cbl-Cy5/D (Riboglow)	640/656-800	271,000	0.25	67,750	n.r.	90
Cbl-4xGly-Atto590/A (Riboglow)	594/624	120,000	0.55	66,000	n.r.	90
Cbl-4xGly-Atto590/D (Riboglow)	594/624	120,000	0.45	54,000	n.r.	90
<b>Pepper aptamer</b>						
HBC485/Pepper	443/485	49,100	0.42	20,622	8	91
HBC497/Pepper	435/497	54,700	0.57	31,179	6.7	91
HBC508/Pepper	458/508	42,500	0.30	12,750	27	91
HBC514/Pepper	458/514	44,100	0.45	19,845	12	91
HBC525/Pepper	491/525	74,100	0.70	51,870	3.8	91
HBC530/Pepper	485/530	65,300	0.66	43,098	3.5	91
HBC599/Pepper	515/599	54,400	0.43	23,392	18	91
HBC620/Pepper	577/620	100,000	0.58	58,000	6.1	91

<sup>a</sup> brightness calculated as  $\epsilon \times QY$ .



A major challenge with RNA aptamers is that they typically function poorly in cells. This can be due to susceptibility to RNA degradation or due to poor intracellular folding. Impaired folding can be due to many causes, including the presence of competing folding pathways, as well as thermal instability, or dependence on ion concentrations that are not normally found in cells.<sup>55</sup> Additionally, lack of brightness, cell-permeability and specificity of fluorogens create a gap in technologies routine application.

### **2.3. Gemini-561/o-Coral fluoromodule as a bright genetically encodable system for live-cell RNA imaging (article 1)**

Despite great advancements in development of fluorogenic aptamers, established RNA aptamer-fluorogen systems generally suffer from weak brightness as only one fluorophore is turned on per aptamer-binding event. Moreover, lack of the dye affinity to the aptamer, dye cytotoxicity and background fluorescence make them less attractive to routine use in laboratory. There is still a strong need for developing new rationally designed fluorogens and selecting new RNA aptamers for RNA labeling.

Recently, our group showed that dimerized dyes can be used as very bright and sensitive fluorogenic protein sensors that respond to a biological event of interest.<sup>35</sup> Herein, we expanded this concept to develop Gemini-561 (G561), a fluorogenic SRB dimer-dye based on dimerization-caused quenching approach. We teamed up with Farah BOUHEDA and Michael RYCKELYNCK (IBMC, University of Strasbourg) who isolated and characterized RNA aptamer named orange Coral (o-Coral). Gemini-561 can be selectively recognized by its cognitive RNA-aptamer named o-Coral. G561/o-Coral system provided a selective and efficient RNA-tag with enhanced brightness and stability compared to previous fluorogen-binding aptamers. The results are described and published in Nature Chemical Biology, 2019 (article 1).<sup>92</sup>

Overall, the ability to target RNA directly with genetically encoded tools is a milestone for RNA-based biotechnology. New RNA-targeting programmable systems will revolutionize our ability to study endogenous RNAs in living systems.

# A dimerization-based fluorogenic dye-aptamer module for RNA imaging in live cells

Farah Bouhedda<sup>1,3</sup>, Kyong Tkhe Fam<sup>2,3</sup>, Mayeul Collot<sup>2\*</sup>, Alexis Autour<sup>1</sup>, Stefano Marzi<sup>1</sup>, Andrey Klymchenko<sup>2\*</sup> and Michael Ryckelynck<sup>1\*</sup>

**Live-cell imaging of RNA has remained a challenge because of the lack of naturally fluorescent RNAs. Recently developed RNA aptamers that can light-up small fluorogenic dyes could overcome this limitation, but they still suffer from poor brightness and photostability. Here, we propose the concept of a cell-permeable fluorogenic dimer of self-quenched sulforhodamine B dyes (Gemini-561) and the corresponding dimerized aptamer (o-Coral) that can drastically enhance performance of the current RNA imaging method. The improved brightness and photostability, together with high affinity of this complex, allowed direct fluorescence imaging in live mammalian cells of RNA polymerase III transcription products as well as messenger RNAs labeled with a single copy of the aptamer; that is, without tag multimerization. The developed fluorogenic module enables fast and sensitive detection of RNA inside live cells, while the proposed design concept opens the route to new generation of ultrabright RNA probes.**

Cells constantly adapt their content to their needs, to changing environment or to pre-determined cell cycles and differentiation programs by tuning their gene expression landscape. Moreover, live-cell imaging experiments have demonstrated that significant cell-to-cell variation in gene expression occurs even within populations of isogenic cells in the same environment<sup>1</sup>. Currently, imaging of gene expression in live cells relies mainly on proteins genetically modified with either fluorescent proteins<sup>2</sup> or tags for specific chemical labeling<sup>3</sup>. Since no naturally fluorescent RNA has been discovered yet, extensive efforts have been devoted to the development of RNA imaging technologies<sup>4</sup>.

An early breakthrough came with the use of RNA-binding proteins (RBP) fused to fluorescence proteins (FPs)<sup>5,6</sup>. In these completely genetically encoded systems, an array (tens of repeats) of the RNA RBP-binding motif is incorporated into the 3' untranslated region (UTR) of the target messenger RNA (mRNA). Co-expressing the RBP-FP-coding gene in the same cell allowed for tracking mRNA (on decoration by FP), which has become a reference method for collecting data on gene expression and RNA trafficking<sup>7</sup>. Substantial simplification is possible by using RNA-based fluorogenic modules<sup>8</sup> in which bulky FPs are substituted by small fluorogens<sup>9</sup>; that is, dyes fluorescing on interaction with a target (bio)molecule<sup>10</sup>. In this scheme, the fluorogen is activated on specific binding to a light-up RNA aptamer<sup>11</sup>.

The capacity of RNA to light-up fluorogenic dyes was first established with an aptamer activating Malachite Green<sup>12</sup>, a dye known for its phototoxicity. Yet, the real breakthrough<sup>8,13</sup> came with the introduction of the cell-permeable and nontoxic green fluorescent protein (GFP)-mimicking fluorogen, 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI), together with its activating aptamer Spinach<sup>14</sup>, later followed by improved aptamers (Spinach2 (ref. 15), Broccoli<sup>16</sup>) and fluorogen derivatives<sup>17</sup>. Unfortunately, these systems suffered from limited brightness and photostability compromising the detection of low-abundant RNAs and extended imaging time<sup>18</sup>. Substantial gain in photostability and brightness was achieved using fluorogens based on organic dyes (cyanines<sup>19–22</sup>

and rhodamines<sup>23–25</sup>), including those operating by photoinduced electron transfer (PET)<sup>26</sup> or Förster resonance energy transfer (FRET)<sup>20,24,27</sup> mechanisms. For instance, conjugates of sulforhodamine B (SRB) dye with dinitroaniline PET quencher recover their fluorescence on association with an aptamer binding the SRB<sup>24,25,28</sup> (for example, SRB-2 aptamer) or the dinitroaniline<sup>23</sup> moiety. An alternative strategy, which could significantly improve fluorogen brightness, is to use dye homo-dimers that self-quench in aqueous solution but become fluorescent on dimer opening after binding to the target biomolecule, a concept that yielded probes for detecting ligand-receptor interaction<sup>29</sup> or DNA hybridization<sup>30</sup>.

Brightness and photostability also rely on the aptamer itself as illustrated by Corn<sup>31,32</sup> pointing the key role of the selection strategy used. Light-up RNA aptamers are usually isolated by systematic evolution of a ligand by exponential enrichment (SELEX)<sup>33,34</sup>, which allows selecting aptamers with high affinity and selectivity for their target, as exemplified by Mango RNA, a light-up aptamer binding its fluorogen (TO1-biotin) with nM affinity<sup>21</sup>. Yet, since no selection pressure is applied for the fluorogenic capacity, these aptamers often suboptimally activate their fluorogen. This limitation can be overcome by functional screening<sup>35–37</sup> such as the microfluidic-assisted in vitro compartmentalization ( $\mu$ IVC)<sup>38</sup> we recently used to identify brighter mutants of Spinach<sup>36</sup> and Mango<sup>37</sup>.

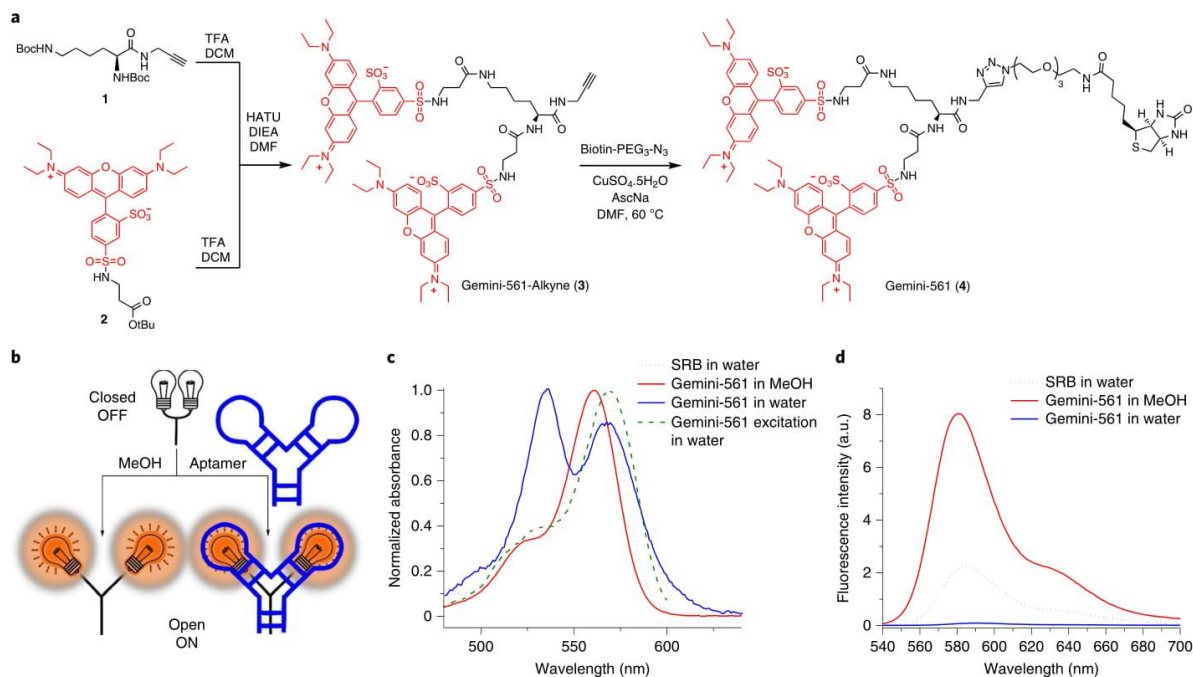
Here, we propose a new concept for preparation of a bright and photostable fluorogen for RNA imaging in cells by exploiting dimerization-induced self-quenching of SRB dyes, which yielded the fluorogen Gemini-561. Following a new selection scheme combining SELEX and  $\mu$ IVC we developed o-Coral, an aptamer of unprecedented compact dimeric structure, able to form a high affinity, bright and photostable complex with Gemini-561, enabling live-cell imaging of RNAs labeled with a single copy of the aptamer.

## Results

**Design and synthesis of Gemini-561 fluorogen.** We intended to develop a red emitting fluorogen efficiently excited with a common laser (532–561 nm), allowing multicolor imaging in combination

<sup>1</sup>Université de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, UPR 9002, Strasbourg, France. <sup>2</sup>Université de Strasbourg, Laboratoire de Bioimagerie et Pathologies, UMR 7021 CNRS, Illkirch, France. <sup>3</sup>These authors contributed equally: Farah Bouhedda, Kyong Tkhe Fam. \*e-mail: mayeul.collot@unistra.fr; andrey.klymchenko@unistra.fr; m.ryckelynck@unistra.fr





**Fig. 1 | Design, synthesis and fluorogenicity of Gemini-561.** **a**, Synthesis of Gemini-561. **b**, Concept of fluorogenic response of Gemini-561 to environment (organic solvent) and aptamer. **c**, Absorption and excitation spectra of Gemini-561 (200 nM) in water and methanol and SRB (200 nM) in water. Results were identical in  $n=3$  independent experiments. **d**, Fluorescence spectra of Gemini-561 (200 nM) in water and methanol and SRB (200 nM) in water. Values are expressed in arbitrary units (a.u.). Results were similar in  $n=3$  independent experiments.

with enhanced GFP (eGFP)-tagged proteins. Rhodamine fluorophores such SRB fulfill this requirement and possess numerous advantages. First, due to their tendency to form H-aggregates<sup>39</sup> and their ability to be quenched by different systems (for example, spiro-lactamization, PET), rhodamines constitute efficient platforms to develop reliable fluorogenic sensors<sup>40,41</sup>. SRB bears two sulfonate groups and on functionalization becomes zwitterionic. Compared to cationic rhodamines or noncharged fluorophores, this feature should increase the polarity of the molecule, thus enhancing its water solubility and preventing nonspecific interactions in biological media. Finally, SRB features optimal photophysical properties including elevated quantum yield, good photostability, sharp emission peak and high molar absorption coefficient ( $\sim 100,000 \text{ M}^{-1} \text{ cm}^{-1}$ ). Gemini-561 was designed to promote the dimerization-induced self-quenching of two SRBs. For this purpose, lysine was chosen as a connector to provide a small distance between the dyes and thus ensure efficient  $\pi$ -stacking on dimerization. Lysine (**1**) and SRB (**2**) derivatives were deprotected and coupled to lead to Gemini-561 alkyne (**3**) (Fig. 1a). The latter was clicked to biotin-PEG- $\text{N}_3$  (the biotin moiety was later used to immobilize the dye during selection experiments) to yield Gemini-561 (**4**).

**Spectroscopic properties of Gemini-561.** Gemini-561 fluorogenicity was first assessed by spectroscopy. In water, Gemini-561 displayed weak fluorescence intensity with a quantum yield value of 0.01. Moreover, a blue shifted band (530 nm) appeared in the absorption spectrum indicating the formation of dimeric H-aggregate (Fig. 1c), in line with earlier report on the squaraine dimer<sup>39</sup>. Additionally, excitation spectrum showed that this band did not correspond to emissive specie (Fig. 1d), thus confirming a dimerization-induced quenching phenomenon. However, on solubilization in methanol, the dimer opened and Gemini-561 displayed absorption and

emission spectra similar to free SRB (Fig. 1c,d) along with an impressive increase in the quantum yield value (0.31, Table 1). In a second step, the nonspecific opening of the dimer was evaluated, revealing that Gemini-561 conserves its quenched form in various physiological media and is not involved in nonspecific interactions with proteins and lipoproteins (Supplementary Fig. 1). Altogether, these experiments demonstrate that Gemini-561 is an effective fluorogenic molecule compatible with biological media, making it a promising candidate for selection of the RNA aptamer.

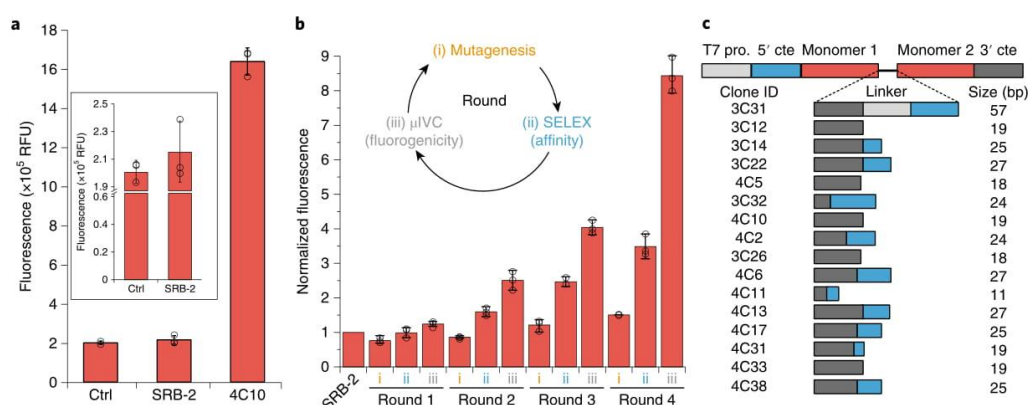
**Isolation of Gemini-561 lighting-up aptamers.** We first studied the SRB-2 aptamer, reported to specifically interact with sulforhodamine B<sup>28</sup> and its derivatives<sup>24</sup>. However, it was a poor activator of Gemini-561 fluorescence (Fig. 2a and Supplementary Fig. 2). We assumed that this weak effect might be attributed to inhibition of dye-aptamer interaction by the dye-dye dimerization.

We therefore started an in vitro evolution of SRB-2 using a strategy combining SELEX and  $\mu\text{IVC}$  (Fig. 2b and Supplementary Fig. 3) to isolate RNAs endowed with both high affinity and high fluorogenic capacity. SRB-2 mutant library ( $\sim 3.4$  mutations per gene) was first generated by error-prone PCR and subjected to a first round of SELEX during which RNAs were challenged to bind bead-immobilized Gemini-561. On stringent wash, bound aptamers were recovered and reverse transcribed into complementary DNAs carrying T7 RNA polymerase promoter. Resulting genes were then subjected to a round of  $\mu\text{IVC}$  screening<sup>37</sup>. Genes encoding fluorogenic RNAs were finally recovered and used to prime a new round of error-prone PCR. A total of four such evolution cycles (mutagenesis, SELEX,  $\mu\text{IVC}$ ) were performed and allowed to gradually improve the average fluorogenic capacity of the population (Fig. 2b), where  $\sim 40\%$  of the obtained sequences were significantly more fluorogenic than the parental SRB-2 aptamer. Out of the 19 best mutants, 16 also

**Table 1 | Spectral and biochemical properties of Gemini-561 alone or in complex with SRB-2 or o-Coral aptamers**

	$\lambda_{\text{Abs max}}$ (nm)	$\epsilon$ ( $\text{M}^{-1}\text{cm}^{-1}$ )	$\lambda_{\text{Em max}}$ (nm)	QY	Fluorescence enhancement	$K_d$ (nM)	Relative brightness
eGFP	490	39,200	508	0.68	-	-	1.00
Buffer	535	100,000	590	0.02	1.0	-	0.08
MeOH	561	162,000	580	0.31	10.3	-	1.88
SRB-2	567	98,000	597	0.04	1.2	$441 \pm 167.5$	0.15
o-Coral	580	141,000	596	0.58	12.8	$73 \pm 1.5$	3.07

Measures were performed in selection buffer (40 mM phosphate buffer pH 7.5, 100 mM KCl, 1 mM MgCl<sub>2</sub>, and 0.05% Tween-20). For each condition, the maximum absorption ( $\lambda_{\text{Abs max}}$ ) and the maximum fluorescence emission ( $\lambda_{\text{Em max}}$ ) wavelength of the fluorogen is given together with its molar extinction coefficient ( $\epsilon$ ). Fluorescence quantum yield (QY) and enhancement were obtained at an excitation wavelength of 530 nm. Dissociation constant ( $K_d$ ) values are the mean of  $n=3$  independent experiments  $\pm 1$  s.d.



**Fig. 2 | Isolation of Gemini-561 lighting-up aptamers by in vitro evolution.** **a**, Gemini-561 activation capacity of the parental SRB-2 and the evolved 4C10 variant. 500 nM of RNAs were incubated with 50 nM of Gemini-561 and the fluorescence was measured at  $\lambda$  excitation/emission (ex/em) = 560/600 nm. The values are expressed in relative fluorescence units (RFU) and are the mean of  $n=3$  independent experiments and the error bars correspond to  $\pm 1$  s.d. **b**, Monitoring of the evolution process. For each round, the enriched library was transcribed in vitro in the presence of 100 nM of Gemini-561 and the fluorescence monitored. The fluorescence apparition rate was computed for each library and normalized to that of the parental SRB-2 aptamer. The inset schematizes the different steps (i, ii and iii) of an evolution round. The values are the mean of  $n=3$  independent experiments, each measurement being shown as an open circle. The error bars correspond to  $\pm 1$  s.d. **c**, Schematic representation of genes coding for the 16 dimerized variants found among the 19 best aptamers at the end of the evolution process. For each variant, the width and the color of the box, respectively, inform on linker length (numerical value given on the right) in base pairs (bp) and the nature of the sequence (light gray, T7 promoter; blue, 5' constant; dark gray, 3' constant). Red boxes correspond to SRB-2-derived core. The clone ID refers to the round of selection from which the clone was extracted (first number) and the clone number assigned during the final screening.

displayed a doubled size resulting from complete duplication (dimerization) of SRB-2 sequence on recombination at a variable position between the 5' and 3' constant regions of two aptamers (Fig. 2c and Supplementary Fig. 4). The exact mechanism of this spontaneous recombination will require a dedicated study. In addition to this duplication, each optimized variant displayed 1–6 point mutations concentrated on P1 and P2 regions of the SRB-2 (Supplementary Fig. 4), while leaving region J2/3, P3 and L3 largely intact, in agreement with their proposed involvement in SRB recognition<sup>28</sup>.

Among the different clones, 4C10 had a remarkably high activation capacity by forming a complex with Gemini-561 that was an order of magnitude more fluorescent than with SRB-2 (Fig. 2a). The 4C10 improvement correlates with a notable increase of affinity ( $\sim 73 \pm 1.5$  nM and  $441 \pm 167$  nM for 4C10 and SRB-2, respectively, see Table 1 and Supplementary Table 1). By further engineering 4C10, we successfully reduced the 19-nucleotide long linker spacing the repeats down to six nucleotides while preserving intact fluorogenicity (Fig. 3a). Further reducing this length down to three nucleotides made the aptamer sensitive to the linker sequence. Therefore, we pursued the study of a six-nucleotide long linker 4C10

derivative, named 'o-Coral' (Supplementary Table 2). The duplication event by itself accounts only partly for the high performances of o-Coral, since a molecule made of two wild-type SRB-2 modules displays only 6% of o-Coral fluorescence (Fig. 3b). Progressive reimplantation of o-Coral mutations showed that all the mutations contributed to o-Coral function, with the double mutant  $U_{25}C/A_{36}G$  (a mutation found in a third of the 19 best mutants, Supplementary Fig. 4) having the predominant effect. Simple introduction of  $U_{25}C/A_{36}G$  mutation into SRB-2 did not yield any improvement of the monomer, indicating a synergic effect of the mutation in the dimer (Fig. 3b). Furthermore,  $U_{25}C/A_{36}G$  had a higher effect when present in the 5' monomer, whereas introducing it in both monomers did not further improve the aptamer.

**o-Coral structural characterization.** The role of the RNA module duplication was further studied at a structural level. Enzymatic structural probing characterization<sup>42</sup> (Supplementary Fig. 5) was in agreement with an overall conservation of the structure initially proposed for the SRB-2 aptamer made of three stems and three unpaired regions<sup>28</sup>. Yet, two RNA folding models could account



and RNA nucleobases. Additionally, we showed that the interaction of Gemini-561 (100 nM) with o-Coral (300 nM) was fast, reaching a plateau after 400 s (Supplementary Fig. 10e). The quantum yield of the complex reached a higher value than Gemini-561 in MeOH (Table 1), showing that the dye is well confined within the aptamer. Moreover, estimating the brightness (that is, molar extinction coefficient  $\times$  quantum yield) of the complex indicated that a single copy of o-Coral associated with Gemini-561 is more than three-fold brighter than eGFP (Table 1), making it a very bright aptamer-based module<sup>8</sup>. Moreover, fluorescence correlation spectroscopy (FCS) confirmed high brightness of the Gemini-561–o-Coral complex (at least 1.14-fold brighter than single tetramethyl rhodamine, a close analog of SRB), suggesting a 1:1 complex composed of a single molecule of Gemini-561 bound to a single copy of o-Coral (Supplementary Table 3). We finally investigated the effect of physiological media on the Gemini-561–o-Coral complex and found that, whereas 10% fetal bovine serum (FBS) slightly interferes with the aptamer–fluorogen interaction (Supplementary Fig. 13), SRB-2 aptamer, BSA or DNA did not challenge the fluorescence of the complex, highlighting the compatibility of the Gemini-561–o-Coral module with a complex cellular environment.

**Imaging Gemini-561–o-Coral in live cells.** After ruling out potential cytotoxicity of Gemini-561 (Supplementary Fig. 14), we validated the functionality of the Gemini-561–o-Coral complex in live HeLa cells by microinjecting the preformed complex either directly into their nucleus or into their cytoplasm of living HeLa cells (Supplementary Fig. 15a). In both cases, an intense red fluorescence was readily observed in the presence of the module, whereas the injection of Gemini-561 alone or mixed with SRB-2 aptamer did not yield a significant fluorescence signal, confirming the specific activation of Gemini-561 by o-Coral and the suitability of the system for live cell applications. Cell permeability of Gemini-561 was also assessed by incubating HeLa cells with 200 nM fluorogen before washing them and microinjecting o-Coral along with Dextran-Alexa-647 conjugate. Both o-Coral and SRB-2 were successfully injected in the cells as attested by Alexa-647 fluorescence, but cytosolic fluorescence of Gemini-561 was observed only in the presence of o-Coral (Supplementary Fig. 15b), demonstrating both cell permeability of the fluorogen and its capacity to detect o-Coral inside live cells.

We next evaluated the performances of our system with aptamers synthesized by the cell machinery. o-Coral was inserted into a F30 scaffold<sup>44</sup> and expressed from a U6 truncated promoter<sup>45</sup> allowing the gene to be transcribed by RNA polymerase III (Pol III). The plasmid also contained an eGFP-coding region produced from a RNA polymerase II (Pol II) promoter and used for identifying transfected green fluorescent cells. Taking advantage of Gemini-561 cell permeability, we expressed o-Coral gene in live HeLa and human embryonic kidney- (HEK)293T cells and incubated them with 200 nM Gemini-561 for 5 min before imaging. Remarkably, green fluorescent transfected cells also displayed intense red (Gemini-561) fluorescence, especially in the nucleus (Fig. 4, Supplementary Fig. 16a) as expected for U6-driven transcripts, confirming that o-Coral preserves RNA sub-cellular localization. Moreover, the sensitivity of this red fluorescence to actinomycin D treatment, a known inhibitor of RNA polymerases, confirmed the requirement of the RNA component. During these live-cell imaging experiments, we noticed cell-to-cell variations in the fluorescence intensity of the Gemini-561–o-Coral (Supplementary Fig. 17), suggesting different transcription states as reported recently with a different aptamer<sup>9</sup>. Taken together, these data show that, due to elevated brightness and affinity of the Gemini-561–o-Coral complex, a single copy of Pol III-expressed o-Coral aptamer is sufficient for RNA detection in living mammalian cells.

To better understand Gemini-561 cell permeability, we monitored Gemini-561–o-Coral fluorescence in cells and found that

internalization kinetics accelerated while increasing Gemini-561 concentration (200 versus 50 nM), reaching a plateau in <10 min with both HeLa and HEK293T cells (Supplementary Fig. 18). Yet, further increasing Gemini-561 concentration to 800 nM led to significant background and dye aggregation (Supplementary Fig. 19). Moreover, since Gemini-561 internalization was still observed at 4°C in various cell types (Supplementary Fig. 20), the fluorogen likely reaches the cytosol by direct diffusion through the cell plasma membrane rather than by endocytosis.

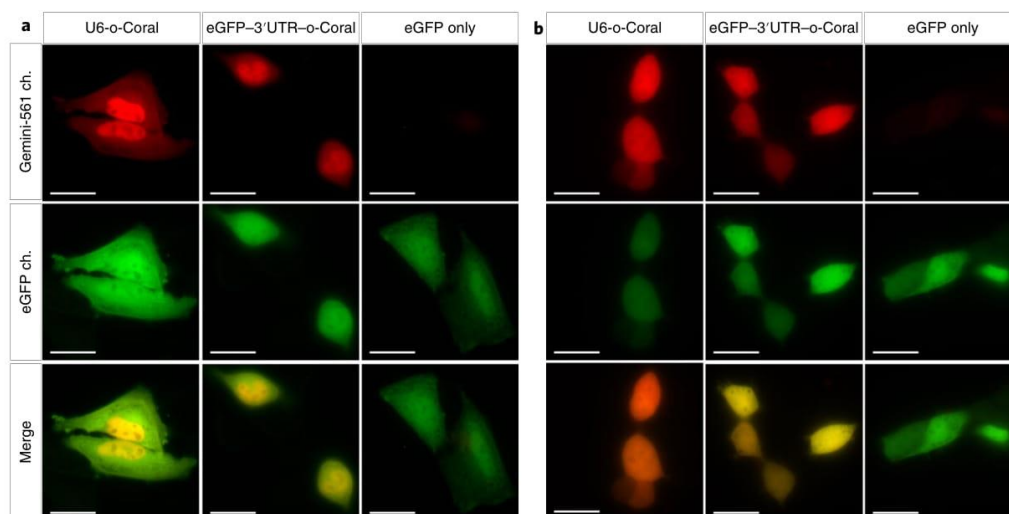
To extend our technology to Pol II transcripts, we inserted o-Coral-F30 directly into the 3' UTR of eGFP gene carried on the plasmid. We found that labeling eGFP mRNA with a single copy of o-Coral was sufficient to detect an intense and homogeneous red fluorescence in eGFP transfected cells (Fig. 4 and Supplementary Fig. 16b), while not substantially altering eGFP mRNA level (Supplementary Fig. 21). The mRNA detection with the Gemini-561–o-Coral module showed its efficiency in HeLa, HEK293T and U87 cells (Supplementary Figs. 20 and 22), revealing fine difference in mRNA distribution. Indeed, in HEK cells the signal of mRNA was evenly distributed between nucleus and cytosol, similar to eGFP, whereas in HeLa cells the mRNA signal was slightly stronger in the nucleus (Fig. 4). Additionally, 5S RNA was also successfully imaged in all three studied cell lines (Supplementary Fig. 23), showing universality of our RNA detection approach. Therefore, an mRNA labeled with a single copy of a o-Coral was visualized without the need for extensive image processing<sup>46</sup>.

The eGFP fluorescence signal correlated well with that of Gemini-561–o-Coral in *egfp3'UTR-o-Coral* expressing HeLa cells (Supplementary Fig. 24). As the protein expression correlates to some extent with corresponding mRNA expression<sup>47</sup>, we concluded that Gemini-561 signal reflects semi-quantitatively the abundance of the corresponding mRNA. The same correlation was also observed with cells in which o-Coral and eGFP were produced from genes colocalized on the same plasmid and controlled by constitutive promoters (U6 and CMV, respectively).

**Comparison of Gemini-561–o-Coral to other systems.** To further demonstrate the advantage of our new fluorogenic system, we systematically compared its performances with other systems using commercially available fluorogens: Broccoli<sup>16</sup>, Mango III<sup>37</sup> and Corn<sup>31</sup>. To properly compare our data with those reported in the literature, aptamers were embedded into the RNA scaffold as they were previously characterized in (that is, transfer RNA scaffold for Corn<sup>31</sup> and F30 scaffold for Broccoli<sup>16</sup>, Mango III<sup>37</sup> and o-Coral). Photostability was first assessed in cuvettes where the entire solution of a fluorogenic dye (0.2  $\mu$ M) and aptamer (1  $\mu$ M) was irradiated with a laser light (Fig. 5a). Moreover, to ensure that each system absorbed similar numbers of photons, the applied irradiation power density (irradiance) was inversely proportional to the molar absorption coefficient of the corresponding dye at the excitation wavelength used. DFHBI-1T–Broccoli complex showed very poor photostability as its intensity vanished within <1 s, as reported before<sup>31,37</sup>. Significantly higher photostability was observed for DFHO–Corn and TO1–biotin–Mango III, as their emission changed to a much lower extent within 200 s of irradiation, in line with earlier studies<sup>31,37</sup>. Emission intensity of Gemini-561–o-Coral did not show any change in the fluorescence intensity in these conditions, indicating that it is substantially more photostable than the three other systems. Similar experiments performed with red fluorescent protein (RFP) and mCherry, also revealed a much higher photostability of Gemini-561–o-Coral (Supplementary Fig. 25), highlighting its advantage over approaches employing FPs (for example, the MS2 system).

Second, we compared brightness of the complexes at the single-molecule level using FCS (the Mango system was not characterized because of incompatibility with our setup). Single-molecule





**Fig. 4 | Live-cell imaging of o-Coral expressed from Pol II and Pol III promoters.** **a, b,** Live cell imaging of HeLa (**a**) and HEK293T (**b**) cells expressing U6-o-Coral, the *egfp* mRNA eGFP-3'UTR-o-Coral, or eGFP only. Cells were incubated with Gemini-561 (200 nM) for 5 min before imaging. Hoechst was used to stain the nucleus ( $5 \mu\text{g ml}^{-1}$ ). The images were acquired using a 500 ms exposure time. Pictures taken in Gemini-561 channel (Gemini-561 ch.; ex, 550 nm; em,  $595 \pm 40$  nm) are shown in red and those taken in eGFP channel (eGFP ch.; ex, 470 nm; em,  $531 \pm 40$  nm) are shown in green. Results were similar in  $n = 3$  independent experiments. Scale bars, 30  $\mu\text{m}$ .

brightness of DFHO-Corn and DFHBI-1T-Broccoli were, correspondingly, 0.62 and 0.18 (Supplementary Table 3) with respect to that of the reference dye fluorescein (at pH 9), confirming the higher brightness of DFHO-Corn<sup>31</sup>. On the other hand, the single molecule brightness of Gemini-561-o-Coral was 1.14 that of the reference dye tetramethyl rhodamine in water. Taking into account that fluorescein and TMR exhibit similar brightness in our FCS setup, the new fluorogenic module, Gemini-561-o-Coral, is approximately two-fold brighter than DFHO-Corn in the single molecule experiments. To compare brightness of our module with MS2-based technology<sup>5,6</sup>, we measured FCS for representative fluorescent proteins used in MS2 labeling. Our module was 19-fold brighter than RFP and seven-fold brighter than mCherry and eGFP, explaining why our technology does not require multimerization and allows mRNA detection using a single copy of the aptamer tag. Nevertheless, we have to admit that multimerized tandem of 30 copies of these proteins, used in MS2 method, is expected to be brighter than a single copy of our module.

Finally, we microinjected the different fluorogen-aptamer complexes into live cells right before taking images (Fig. 5b–d) and measured the signal to background noise ratio. Gemini-561-o-Coral system provided images with a signal to background noise value that was substantially higher than those obtained with other studied systems (Fig. 5b). Image analysis also showed that fluorescence intensity in the cells decayed much slower for Gemini-561-o-Coral (Fig. 5c,d), confirming superior photostability of our system. In an alternative experiment, cells were incubated in the presence of the fluorogen followed by microinjection of the corresponding aptamer. This experiment further confirmed that Gemini-561-o-Coral was significantly more photostable than reference aptamer-based systems (Supplementary Fig. 26), allowing >20 s continuous imaging with only a minor loss of fluorescence intensity. Remarkably, Gemini-561-o-Coral fluorescence can still be detected on several hours of constant illumination (Supplementary Fig. 27).

## Discussion

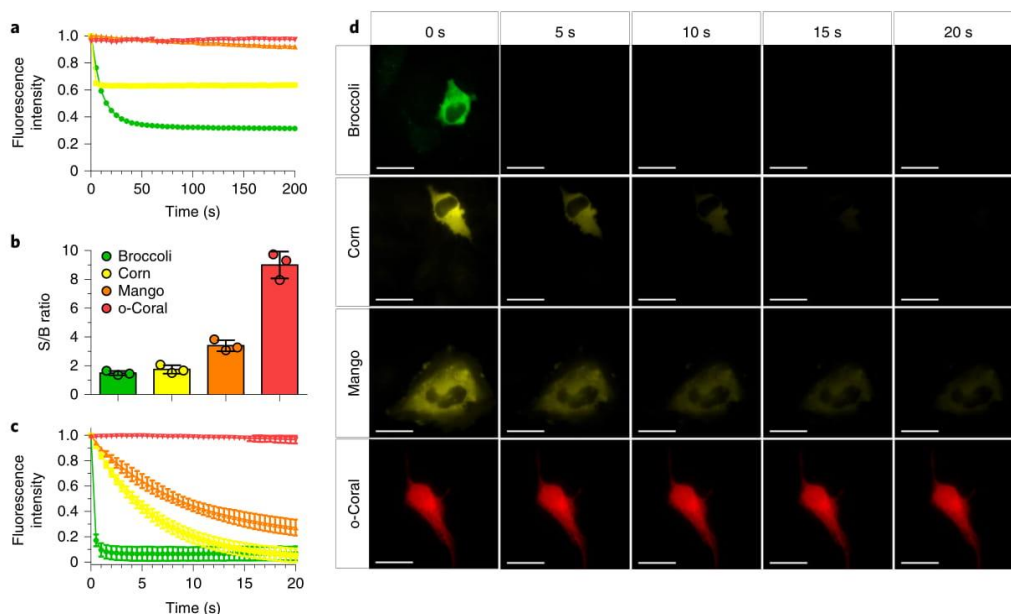
Light-up aptamers gained their niche as powerful genetically encoded RNA imaging tools. Due to high modularity of nucleic

acids and available SELEX methodologies, a palette of fluorogen-aptamer systems was discovered to shine light on complex cell machinery. Unfortunately, the limited brightness and photostability of aptamer-based fluorogenic modules developed so far have narrowed their broad application.

In this work we developed and characterized Gemini-561-o-Coral, a new RNA-based fluorogenic module displaying high brightness, affinity and photostability making it a particularly bright module<sup>8</sup>. We reached this goal by combining two key innovations. First, the fluorogen Gemini-561 was designed based on two copies of the bright and photostable sulforhodamine B dye that self-quenches in water and is able to enter the live cells in a few minutes. This quenching mode is important since, on activation, both fluorophores become strongly fluorescent making such a probe brighter compared with other monomeric probes. Second, we developed the light-up RNA aptamer o-Coral using an integrated *in vitro* evolution strategy combining rounds of mutagenesis and SELEX in tandem with  $\mu\text{IVC}$  screening. In this scheme, the SELEX step allows isolation of RNAs with a high affinity for Gemini-561, whereas the  $\mu\text{IVC}$  isolates the most fluorogenic sequences. The obtained o-Coral efficiently opens the self-quenched dimer Gemini-561, thus activating its fluorescence. Using the same overall strategy, it should be possible to further expand RNA imaging toolbox by developing new orthogonal modules made of dimeric (Gemini) fluorophores of any desired color while selecting new specific Coral aptamers.

The intertwined dimerized structure of o-Coral may also be advantageous for the future engineering of aptamers. Indeed, both L2 and L2' loops together with the linker region are highly amenable to sequence modification and are attractive sites for inserting other sequences (for example, sensing aptamers). By doing so, o-Coral could be converted into complex multi-inputs logic gates or biosensors.

o-Coral is readily expressed in mammalian cells where it forms a bright complex with Gemini-561 that is otherwise not activated by cell components. Similar to the Malachite Green<sup>48</sup> and DIR2s aptamers<sup>49</sup>, o-Coral does not seem to possess the G-quadruplex organization shared by most of the other structurally characterized light-up



**Fig. 5 | Comparative analysis of photostability by fluorescence microscopy and spectroscopy. a**, Photostability of G561-o-Coral ( $0.2\ \mu\text{M}/1\ \mu\text{M}$ ) compared to Broccoli + DFHBI-1T ( $0.2\ \mu\text{M}/1\ \mu\text{M}$ ), Corn + DFHO ( $0.2\ \mu\text{M}/1\ \mu\text{M}$ ), Mango + TO1-Biotin ( $0.2\ \mu\text{M}/1\ \mu\text{M}$ ). Each complex was excited at the same molar extinction coefficient value:  $30,000\ \text{M}^{-1}\ \text{cm}^{-1}$ . Broccoli, Corn and Mango were excited using 488 nm laser ( $7.75$ ,  $11$  and  $10\ \text{mW}\ \text{cm}^{-2}$ , respectively) and o-Coral was excited using a 532 nm laser ( $7\ \text{mW}\ \text{cm}^{-2}$ ). Fluorescence intensity was monitored at 507 nm for Broccoli, 545 nm for Corn, 535 nm for Mango and 596 nm for o-Coral. Fluorescence values were normalized to that of the first acquired image. **b-d**, Photostability and signal to background noise ratio measurement in live HeLa cells. In vitro transcribed and purified aptamers were preincubated with respective fluorogenic dyes for 10 min in selection buffer to form complex. Complexes were microinjected in live HeLa cells using  $5\ \mu\text{M}$  dye and  $20\ \mu\text{M}$  aptamer concentration. Microinjection parameters:  $P_i = 90\ \text{hPa}$ ;  $T_i = 0.3\ \text{s}$  and  $P_c = 10\ \text{hPa}$ . Consecutive images were acquired, each using a 500 ms exposure time. The excitation power was adjusted for the fluoromolecules to absorb similar numbers of photons. Broccoli (ex, 470 nm; em,  $475 \pm 50\ \text{nm}$ ); Corn (ex, 470 nm; em,  $531 \pm 40\ \text{nm}$ ); Mango (ex, 470 nm; em,  $531 \pm 40\ \text{nm}$ ); Coral (ex, 550 nm; em,  $595 \pm 40\ \text{nm}$ ). **b**, Signal to background noise (S/B) ratio of the first acquired image depicting the brightness of the system and the quality of obtain images. Signal to background noise ratios were calculated from fluorescence intensity values extracted from images using same region of interest from  $n = 3$  independent injections. The value of each measurement is shown as a colored dot. The error bars correspond to  $\pm 1\ \text{s.d.}$  **c**, Fluorescence intensity decay curves over the time. The fluorescence values were normalized to that of the first acquired image. Data represent average values  $\pm 1\ \text{s.d.}$  extracted from images from  $n = 3$  independent experiments. **d**, Representative micrographs taken during the experiment. Results were similar in  $n = 3$  independent experiments. Scale bars:  $30\ \mu\text{m}$ .

aptamers<sup>43</sup>. This is of particular interest when considering a recent report indicating that most RNA G-quadruplex domains would be kept globally unfolded in mammalian cells<sup>50</sup>, suggesting that G-quadruplex-based RNA may not be optimal for use in living cells.

Direct comparison of Gemini-561-o-Coral with the most representative aptamer-based fluorogenic modules in live cells showed clear advantages of the new module in terms of brightness and photostability. Moreover, its substantially higher brightness (7–20-fold according to FCS) and photostability compared to fluorescent proteins explain the capacity to detect RNA using a single copy of the aptamer tag, whereas in the case of the MS2-eGFP module<sup>5,6</sup> a tandem of ~30 copies should be used. Imaging of RNA by integrating just a single copy of the aptamer, achieved with Gemini-561-o-Coral, has remained a challenge so far in this technology. However, since dedicated experiments are yet to be performed, it is too early to conclude whether the single molecule resolution can be reached without tag multimerization. Overall, the Gemini-561-o-Coral system notably strengthens the toolbox for RNA imaging and shows a new direction in the development of ultrabright fluorogenic aptamer-based modules.

#### Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and

associated accession codes are available at <https://doi.org/10.1038/s41589-019-0381-8>.

Received: 20 October 2018; Accepted: 22 August 2019;

Published online: 21 October 2019

#### References

- Levine, J. H., Lin, Y. & Elowitz, M. B. Functional roles of pulsing in genetic circuits. *Science* **342**, 1193–1200 (2013).
- Giepmans, B. N., Adams, S. R., Ellisman, M. H. & Tsien, R. Y. The fluorescent toolbox for assessing protein location and function. *Science* **312**, 217–224 (2006).
- Xue, L., Karpenko, I. A., Hiblot, J. & Johnsson, K. Imaging and manipulating proteins in live cells through covalent labeling. *Nat. Chem. Biol.* **11**, 917–923 (2015).
- Dean, K. M. & Palmer, A. E. Advances in fluorescence labeling strategies for dynamic cellular imaging. *Nat. Chem. Biol.* **10**, 512–523 (2014).
- Bertrand, E. et al. Localization of ASH1 mRNA particles in living yeast. *Mol. Cell* **2**, 437–445 (1998).
- Tutucci, E. et al. An improved MS2 system for accurate reporting of the mRNA life cycle. *Nat. Methods* **15**, 81–89 (2018).
- Buxbaum, A. R., Haimovich, G. & Singer, R. H. In the right place at the right time: visualizing and understanding mRNA localization. *Nat. Rev. Mol. Cell Biol.* **16**, 95–109 (2015).
- Bouhedda, F., Autour, A. & Ryckelynck, M. Light-Up RNA aptamers and their cognate fluorogens: from their development to their applications. *Int. J. Mol. Sci.* **19**, 44 (2018).



9. Li, C., Tebo, A. G. & Gautier, A. Fluorogenic labeling strategies for biological imaging. *Int. J. Mol. Sci.* **18**, 1473 (2017).
10. Klymchenko, A. S. Solvatochromic and fluorogenic dyes as environment-sensitive probes: design and biological applications. *Acc. Chem. Res.* **50**, 366–375 (2017).
11. Zhang, J. et al. Tandem spinach array for mrna imaging in living bacterial cells. *Sci. Rep.* **5**, 17295 (2015).
12. Babendure, J. R., Adams, S. R. & Tsien, R. Y. Aptamers switch on fluorescence of triphenylmethane dyes. *J. Am. Chem. Soc.* **125**, 14716–14717 (2003).
13. You, M. & Jaffrey, S. R. Structure and mechanism of RNA mimics of green fluorescent protein. *Annu. Rev. Biophys.* **44**, 187–206 (2015).
14. Paige, J. S., Wu, K. Y. & Jaffrey, S. R. RNA mimics of green fluorescent protein. *Science* **333**, 642–646 (2011).
15. Strack, R. L., Disney, M. D. & Jaffrey, S. R. A superfolder spinach2 reveals the dynamic nature of trinucleotide repeat-containing RNA. *Nat. Methods* **10**, 1219–1224 (2013).
16. Filonov, G. S., Moon, J. D., Svendsen, 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.* **136**, 16299–16308 (2014).
17. Song, W., Strack, R. L., Svendsen, N. & Jaffrey, S. R. Plug-and-Play fluorophores extend the spectral properties of spinach. *J. Am. Chem. Soc.* **136**, 1198–1201 (2014).
18. Han, K. Y., Leslie, B. J., Fei, J., Zhang, J. & Ha, T. Understanding the photophysics of the spinach-DFHBI RNA aptamer-fluorogen complex to improve live-cell RNA imaging. *J. Am. Chem. Soc.* **135**, 19033–19038 (2013).
19. Constantin, T. P. et al. Synthesis of new fluorogenic cyanine dyes and incorporation into RNA fluoromodules. *Org. Lett.* **10**, 1561–1564 (2008).
20. Murata, A., Sato, S., Kawazoe, Y. & Uesugi, M. Small-molecule fluorescent probes for specific RNA targets. *Chem. Commun.* **47**, 4712–4714 (2011).
21. Dolgoshina, E. V. et al. RNA mango aptamer-fluorophore: a bright, high-affinity complex for RNA labeling and tracking. *ACS Chem. Biol.* **9**, 2412–2420 (2014).
22. Tan, X. et al. Fluoromodules consisting of a promiscuous RNA aptamer and red or blue fluorogenic cyanine dyes: selection, characterization, and bioimaging. *J. Am. Chem. Soc.* **139**, 9001–9009 (2017).
23. Arora, A., Sunbul, M. & Jaschke, A. Dual-colour imaging of RNAs using quencher- and fluorophore-binding aptamers. *Nucleic Acids Res.* **43**, e144 (2015).
24. Sunbul, M. & Jaschke, A. Contact-mediated quenching for RNA imaging in bacteria with a fluorophore-binding aptamer. *Angew. Chem. Int. Ed. Engl.* **52**, 13401–13404 (2013).
25. Sunbul, M. & Jaschke, A. SRB-2: a promiscuous rainbow aptamer for live-cell RNA imaging. *Nucleic Acids Res.* **46**, e110 (2018).
26. 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.* **127**, 14954–14955 (2005).
27. Shin, I. et al. Live-cell imaging of Pol II promoter activity to monitor gene expression with RNA IMAGETag reporters. *Nucleic Acids Res.* **42**, e90 (2014).
28. Holeman, L. A., Robinson, S. L., Szostak, J. W. & Wilson, C. Isolation and characterization of fluorophore-binding RNA aptamers. *Fold. Des.* **3**, 423–431 (1998).
29. Karpenko, I. A. et al. Fluorogenic squaraine dimers with polarity-sensitive folding as bright far-red probes for background-free bioimaging. *J. Am. Chem. Soc.* **137**, 405–412 (2015).
30. Okamoto, A. ECHO probes: a concept of fluorescence control for practical nucleic acid sensing. *Chem. Soc. Rev.* **40**, 5815–5828 (2011).
31. Song, W. et al. Imaging RNA polymerase III transcription using a photostable RNA-fluorophore complex. *Nat. Chem. Biol.* **13**, 1187–1194 (2017).
32. Warner, K. D. et al. A homodimer interface without base pairs in an RNA mimic of red fluorescent protein. *Nat. Chem. Biol.* **13**, 1195–1201 (2017).
33. Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822 (1990).
34. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–510 (1990).
35. Gotrik, M. et al. Direct selection of fluorescence-enhancing RNA aptamers. *J. Am. Chem. Soc.* **140**, 3583–3591 (2018).
36. Autour, A., Westhof, E. & Ryckelynck, M. iSpinach: a fluorogenic RNA aptamer optimized for in vitro applications. *Nucleic Acids Res.* **44**, 2491–2500 (2016).
37. Autour, A. et al. Fluorogenic RNA mango aptamers for imaging small non-coding RNAs in mammalian cells. *Nat. Commun.* **9**, 656 (2018).
38. Ryckelynck, M. et al. Using droplet-based microfluidics to improve the catalytic properties of RNA under multiple-turnover conditions. *RNA* **21**, 458–469 (2015).
39. Shulov, I. et al. Fluorinated counterion-enhanced emission of rhodamine aggregates: ultrabright nanoparticles for bioimaging and light-harvesting. *Nanoscale* **7**, 18198–18210 (2015).
40. Collot, M. et al. Calcium rubies: a family of red-emitting functionalizable indicators suitable for two-photon Ca<sup>2+</sup> imaging. *J. Am. Chem. Soc.* **134**, 14923–14931 (2012).
41. Despras, G. et al. H-Rubies, a new family of red emitting fluorescent pH sensors for living cells. *Chem. Sci.* **6**, 5928–5937 (2015).
42. Duval, M. et al. in *RNA Structure and Folding: Biophysical Techniques and Prediction Methods* (eds Klostermeier, D. & Hammann, C.) 29–50 (De Gruyter, 2013).
43. Trachman, R. J. 3rd, Truong, L. & Ferre-D'Amare, A. R. Structural principles of fluorescent RNA aptamers. *Trends Pharm. Sci.* **38**, 928–939 (2017).
44. Filonov, G. S., Kam, C. W., Song, W. & Jaffrey, S. R. In-gel imaging of RNA processing using Broccoli reveals optimal aptamer expression strategies. *Chem. Biol.* **22**, 649–660 (2015).
45. Good, P. D. et al. Expression of small, therapeutic RNAs in human cell nuclei. *Gene Ther.* **4**, 45–54 (1997).
46. Guet, D. et al. Combining spinach-tagged RNA and gene localization to image gene expression in live yeast. *Nat. Commun.* **6**, 8882 (2015).
47. Liu, Y., Beyer, A. & Aebersold, R. On the dependency of cellular protein levels on mRNA abundance. *Cell* **165**, 535–550 (2016).
48. Baugh, C., Grate, D. & Wilson, C. 2.8 Å crystal structure of the malachite green aptamer. *J. Mol. Biol.* **301**, 117–128 (2000).
49. Shelke, S. A. et al. Structural basis for activation of fluorogenic dyes by an RNA aptamer lacking a G-quadruplex motif. *Nat. Commun.* **9**, 4542 (2018).
50. Guo, J. U. & Bartel, D. P. RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science* **353**, aaf5371 (2016).

#### Acknowledgements

We thank R. Leblay, A.-C. Helfer, E. Gutzwiller and G. Lieber for technical assistance as well as P. Romby and I. Caldelari for fruitful scientific discussion. We thank O. Kucherak for synthesis of intermediate 2, E. Real for scientific discussions and J. Valanciunaite and N. Humbert for the technical assistance. This work received financial support from Agence Nationale de la Recherche (BrightRiboProbes, grant no. ANR-16-CE11-0010-01/02 to M.R. and A.S.K.) and BacNet (grant no. ANR-10-BINF\_02\_02 to M.R.) and European Research Council ERC Consolidator grant BrightSens (no. 648528 to A.S.K.). This work has also been published under the framework of the LabEx (no. ANR-10-LABX-0036\_NETRNA to M.R.) and benefits from a funding from the state managed by the French National Research Agency as part of the Investments for the Future Program. It was also supported by the Université de Strasbourg and the Centre National de la Recherche Scientifique.

#### Author contributions

A.K., M.C. and M.R. proposed the concept of this study. M.C. synthesized Gemini-561. F.B. and K.T.F. performed main part of experimental work and data analysis with help and supervision of M.R., M.C. and A.K. F.B. and M.R. developed the aptamer o-Coral, including its variants and characterized their structure with the help of S.M. A.A. and F.B. characterized the contribution of the different selected mutations. K.T.F., M.C., F.B. and A.K. characterized Gemini-561–o-Coral fluorescence properties in solution. F.B. generated plasmids encoding o-Coral. K.T.F. realized all cellular studies and all fluorescence imaging of the cells, with help of A.K. and M.C.

#### Competing interests

F.B., K.T.F., M.C., A.K., M.R., the University of Strasbourg and the CNRS have filed a patent application covering the technology presented in this manuscript.

#### Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41589-019-0381-8>.

Correspondence and requests for materials should be addressed to M.C., A.K. or M.R.

Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019



## Methods

**Synthetic procedures.** Synthetic procedures of Gemini-561-alkyne (3) and Gemini-561 (4) are given in the Supplementary Note available online.

**Optical spectroscopy.** Optical spectroscopy was performed with Milli-Q water (Millipore). All the solvents were spectroscopic grade. Absorption and emission spectra were recorded on a Cary 4000 Scan ultraviolet-visible spectrophotometer (Varian) and a FluoroMax-4 spectrofluorometer (Horiba Jobin Yvon) equipped with a thermostated cell compartment, respectively. For standard recording of fluorescence spectra, the emission was collected 10 nm after the excitation wavelength. All the spectra were corrected from wavelength-dependent response of the detector and measured at room temperature (25 °C). Absorbance values of all solutions were systematically below 0.1 at the maximum. Quantum yields were determined using a reference dye (Rhodamine B in water).

**Gene library generation.** The sequence coding for the SRB aptamer<sup>28</sup> was flanked with constant regions at 5' (GGGAGACAGCTAGAGTAC) and 3' (GACACGAGCAGTGTAC) ends to allow DNA amplification and RNA reverse transcription. Mutant libraries were generated by error-prone PCR by subjecting 10 fmol of DNA (diluted in 200 µg ml<sup>-1</sup> of yeast total RNA solution (Ambion)) to four amplification cycles in the presence of Fwd (CTTTAATACGACTACTATAGGGAGACAGCTAGAGTAC) and Rev (GACACGAGCAGTGTAC) primers, as well as nucleotide s (JBS dNTPMutagenesis Kit, Jena Bioscience) as described before<sup>28</sup>. Then, 1 ng of PCR products was amplified in a second PCR mixture containing 10 pmol of each primer (Fwd and Rev), 0.2 mM of each dNTPs, 5 U of PhireII (Fermentas) and the corresponding buffer (Fermentas). The mixture was thermocycled starting with an initial step of denaturation of 30 s at 95 °C followed by 25 cycles of 5 s at 95 °C and 30 s at 60 °C. The PCR products were purified following the Wizard SV Gel and PCR Clean-up System (Promega) kit instructions and the quantity of DNA was determined by NanoDrop measurement.

**In vitro transcription and RNA purification.** Genes coding for aptamers were PCR amplified with the same procedure used before (25 cycles of PCR using PhireII enzyme). Next, 1 µg of PCR products was in vitro transcribed in 500 µl of mixture containing 2 mM of each NTP (Larova), 25 mM MgCl<sub>2</sub>, 44 mM Tris-HCl pH 8.0 (at 25 °C), 5 mM DTT, 1 mM Spermidine and 17.5 µg ml<sup>-1</sup> T7 RNA polymerase (prepared in the laboratory). After 3 h of incubation at 37 °C, 1,000 units of Baseline-Zero DNase (Epicentre) and the corresponding buffer were added to the mixture and a second incubation of 1 h at 37 °C was performed. RNA was then recovered by phenol extraction. In vitro transcribed RNA was purified using ion exchange chromatography (FastFlow DEAE sepharose, GE Healthcare) by loading the RNA in and washing the resin with bind/wash buffer (50 mM NaCl, 50 mM Tris-HCl pH 7.5 and 10% Glycerol) and eluting it with elution buffer (600 mM NaCl and 50 mM Tris-HCl pH 7.5). Alternatively, RNA was gel purified by ethanol precipitating transcription mixture and dissolving the pellet into loading buffer (0.05% bromophenol blue, 20% glycerol, TBE 1x, 8 M urea). The solution was then loaded on a 12% denaturing 8 M urea acrylamide/bisacrylamide gel. The piece of gel containing RNA was identified by ultraviolet shadowing, and the RNA electroeluted as described before<sup>28</sup>. Eluted RNA was then ethanol precipitated, the washed pellets were dissolved in diethylpyrocarbonate-treated water and quantified with Nanodrop (Thermo Scientific).

**SELEX.** Here, 100 µl of streptavidin-agarose beads (Sigma-Aldrich) were washed with 200 µl of activation buffer (100 mM NaOH, 50 mM NaCl). The beads were then centrifuged 5 min at 5,000g and room temperature, and the supernatant was removed by pipetting. This procedure was repeated with 200 µl of pre-wash buffer (40 mM potassium phosphate buffer pH 7.5, 100 mM KCl, 1 mM MgCl<sub>2</sub> and 0.05% Tween 20) and finally 200 µl of wash buffer (pre-wash buffer supplemented with 1 mg ml<sup>-1</sup> BSA, New England Biolabs), 0.1 mg ml<sup>-1</sup> sodium heparin (Sigma-Aldrich) and 200 µg ml<sup>-1</sup> yeast total RNA (Ambion). The resin was loaded into a cartridge (Plastic small column CS-20 ABT) previously equilibrated by an overnight incubation with wash buffer at 4 °C. Then, 500 µl of wash buffer supplemented with 10 nmol of biotinylated Gemini-561 dye were added on the beads at a controlled flow-rate of 10 ml h<sup>-1</sup> using a syringe pump (PhD 2000, Harvard Apparatus). Afterward, the unbound fluorophore was washed away by 20 ml of wash buffer (20 ml h<sup>-1</sup>). Next, ~50 µg of purified RNA was introduced in 250 µl of ultra-pure diethylpyrocarbonate-treated water and renatured by 2 min at 85 °C followed by 5 min at 25 °C. Then, 250 µl of twice concentrated washing buffer was added and the mixture was infused through the Gemini-561 substituted resin at a flow-rate of 1.5 ml h<sup>-1</sup>. Unbound RNAs were eliminated per 20 ml of wash buffer infused at 20 ml h<sup>-1</sup>. This initial wash was followed by three additional washes of 15 ml of buffer while reducing the ionic strength (respectively, 100, 10 and 1 mM KCl). The selection pressure was further increased during the last round by introducing 5 µM of free Gemini-561 dye during the last wash of the column. The beads were then collected using a Pasteur pipette, centrifuged and placed in 100 µl of elution buffer (95% formamide and 25 mM EDTA). After 2 min of heating at 90 °C, the beads were centrifuged, the supernatant was recovered and the RNA precipitated as above. RNA was pelleted, washed and resuspended

in 50 µl of 2 µM Rev primer solution. The mixture was heated for 2 min at 85 °C and cooled at 25 °C for 5 min. Next, 50 µl of reaction mixture containing 0.5 mM of each dNTP, 400 U of reverse transcriptase (Maxima H Minus, ThermoFisher) and the corresponding 2x concentrated buffer were added and the mixture and incubated for 1 h at 55 °C. The resulting cDNA was then extracted with a mixture of phenol:chloroform:isoamyl alcohol 25:24:1 (Roth) and precipitated. cDNA was recovered by 30 min of centrifugation at 21,000g and 4 °C, washed, dried, resuspended in 250 µl of PhireII PCR mixture and amplified by PCR as described above.

**Droplet-based microfluidics.** Microfluidic chips were made of polydimethylsiloxane as described in ref.<sup>28</sup>.

**Droplet digital PCR.** DNA mutant libraries were diluted in 200 µg ml<sup>-1</sup> yeast total RNA solution (Ambion) down to roughly eight template DNA molecules per picoliter. Then, 1 µl of this dilution was then introduced in 100 µl of PCR mixture containing 0.2 µM of each primer (Fwd and Rev), 0.2 mM of each dNTPs, 20 µM of coumarin, 0.1% Pluronic F68 (Sigma), 5 U of PhireII enzyme (Fermentas) and the corresponding buffer (Fermentas). The mixture was loaded in a length of polytetrafluoroethylene tubing and infused into droplet generator microfluidic chip where it was dispersed in 2.5 pl droplets (production rate of ~12,000 droplets s<sup>-1</sup>) carried by Novec 7500 fluorinated oil (3 M) supplemented with 3% of a fluorosurfactant (proprietary synthesis). Droplet production frequency was monitored in real time using an optical device and software developed by the team<sup>28</sup> and used to determined droplet volume. Then, 2.5 pl droplets were generated by adjusting pump flowrates (MFCS, Fluigent). The emulsion was collected in 0.2 ml tubes and subjected to an initial denaturation step of 2 min at 98 °C followed by 30 cycles of 10 s at 98 °C, 30 s at 55 °C and 30 s at 72 °C. Droplets were then reinjected into a droplet fusion microfluidic device.

**Droplet fusion.** PCR droplets were reinjected and spaced into a fusion device at a rate of ~1,500 droplets s<sup>-1</sup>. Each PCR droplet was then synchronized with a 16 pl in vitro transcription (IVT) droplet containing 2 mM each nucleoside triphosphate (NTP, Larova), 25 mM MgCl<sub>2</sub>, 44 mM Tris-HCl pH 8.0 (at 25 °C), 5 mM DTT, 1 mM Spermidine, 0.1% of Pluronic F68 (Sigma), 1 µg of pyrophosphatase (Roche), 500 nM Gemini-561, 1 µM coumarin acetate (Sigma) and 17.5 µg ml<sup>-1</sup> T7 RNA polymerase (prepared in the laboratory). IVT mixture was loaded in a length of polytetrafluoroethylene tubing and kept on ice during all experiment. PCR droplets were spaced and IVT droplets produced using a single stream of Novec 7500 fluorinated oil (3 M) supplemented with 2% (w/w) of fluorinated. Flow rates (MFCS, Fluigent) were adjusted to generate 16 pl IVT droplets and maximize synchronization of one PCR droplet with one IVT droplet. Pairs of droplets were then fused with an alternating current field (400 V at 30 kHz) and the resulting emulsion was collected off-chip and incubated for 2 h at 37 °C.

**Droplet analysis and sort.** The emulsion was finally reinjected into an analysis and sorting microfluidic device at a frequency of ~150 droplets s<sup>-1</sup> and spaced with a stream of surfactant-free Novec 7500 fluorinated oil (3 M). The orange fluorescence (Gemini-561 in complex with the aptamer) of each droplet was analyzed and the 1–2% most orange fluorescent droplets were sorted. The gated droplets were deflected into collecting channel by applying alternating current fields (1,000 V, 30 kHz) and collected into a 1.5 ml tube. Sorted droplets were recovered from the collection tubing by flushing 200 µl of Novec fluorinated oil (3 M). Next, 100 µl of 1H, 1H, 2H, 2H-perfluoro-1-octanol (Sigma-Aldrich) and 200 µl of 200 µg ml<sup>-1</sup> yeast total RNA solution (Ambion) were then added and the droplets broken by vortexing the mixture. DNA-containing aqueous phase was recovered, and the DNA recovered by PCR.

**TA cloning.** The DNA contained in the libraries obtained after the last two rounds of µIVC screening were PCR-amplified as described above but using the DreamTaq enzyme and buffer (Fermentas) instead of PhireII (Fermentas). PCR products were purified (Wizard SV Gel and PCR Clean-Up System kit (Promega)) and inserted into the cloning vector of the insTAclone PCR Cloning kit (Thermo Scientific) following the manufacturer's recommendations by overnight ligation at 4 °C. ElectroTEN Blue bacteria (Agilent) were transformed by electroporation and plated onto a 2YT/agar/ampicillin (100 µg ml<sup>-1</sup>) plate.

Individual colonies were used to seed 20 µl of PhireII PCR mixture (see above) while the rest of the colony was introduced in 3 ml of liquid medium 2YT/ampicillin (100 µg ml<sup>-1</sup>) overnight at 37 °C under agitation. On thermocycling, 1 µl of PCR product was added to 18 µl of in vitro transcription mixture (see above) supplemented with 100 nM of Gemini-561. The mixture was then incubated at 37 °C in a thermocycler (Stratagene Mx3005P, Agilent Technologies) and the fluorescence (excitation/emission (ex/em) = 575/602 nm) was monitored for 2 h. Finally, plasmid DNA was extracted from bacteria of interest using the GenElute Plasmid Miniprep kit (ThermoFisher) and sequenced by the Sanger method (GATC Biotech).

**Real-time IVT measurements.** PCR products of each selection cycle were purified (Wizard SV Gel and PCR Clean-up System (Promega)) kit and quantified



(NanoDrop). Then, 50 ng of pure DNA was introduced into 38  $\mu$ l of *in vitro* transcription mixture (see above) supplemented with 100 nM of Gemini-561. This mixture was then incubated at 37 °C in a real-time thermocycler (Stratagene Mx3005P, Agilent Technologies) and the fluorescence was monitored as above.

**Fluorescence measurement on purified RNA.** Here, 1  $\mu$ M purified RNA was heated for 1 min at 90 °C and cooled at 4 °C for 1 min. The solution was then supplemented with 1 volume of a twice concentrated mixture containing 80 mM potassium phosphate buffer pH 7.5, 2 mM MgCl<sub>2</sub>, 0.1% Tween 20, 200 mM of salt (KCl, NaCl, LiCl or CsCl) and 100 nM Gemini-561. The mixture was then incubated for 10 min at 25 °C before measuring the fluorescence at 25 °C on a real-time thermocycler (ex/em = 575/602 nm, Mx3005P, Agilent) or on microplate reader (ex/em = 560/600 nm, SpectraMax iD3, Molecular Devices).

**RNA probing.** Here, 20  $\mu$ g of RNA were first dephosphorylated for 20 min at 37 °C using 1 U FastAP (Fermentas) per  $\mu$ g of RNA. On phenol/chloroform extraction and RNA precipitation, dephosphorylated RNA was 5' labeled by incubating 5  $\mu$ g of dephosphorylated RNA with 50  $\mu$ Ci of [<sup>32</sup>P] $\gamma$ -ATP and 10 U of T4 polynucleotide kinase, with T4 PNK 1 $\times$  buffer in final volume of 15  $\mu$ l for 1 h at 37 °C before the phenol/chloroform was extracted, precipitated and pelleted. Labeled RNA was then gel-purified and eluted from the gel by an overnight incubation at 4 °C and gently mixing in RNA Elution Buffer (500 mM of ammonium acetate and 1 mM of EDTA). Eluted radiolabeled RNA was extracted by phenol/chloroform treatment, precipitated in ethanol and pelleted as described above. The RNA is resuspended in nuclease-free water. The specific activity (c.p.m.  $\mu$ l<sup>-1</sup>) was calculated by measuring 1  $\mu$ l in a radioactivity counter Multi-Purpose Scintillator Counter (Beckman) by Cerenkov counting. Labeled RNA (200,000 c.p.m.) was renatured by heating it for 1 min at 90 °C then 1 min on ice and then preincubated at 20 °C for 15 min in a buffer containing 20 mM of this-HCl pH 7.5, 1 mM of MgCl<sub>2</sub>, and 150 mM of KCl. Next, 1  $\mu$ g of total RNA was then added to the preparation and RNAs were incubated with T1 enzyme (0.25, 0.5 and 1 U), T2 enzyme (0.0125, 0.025 and 0.05 U) and V1 enzyme (0.001, 0.002 and 0.004 U) for 5 min at 20 °C (T1 and V1) or 10 min at 20 °C (T2) or water (control lane). The same amount of digested product was loaded on a 10% denaturing 8 M urea acrylamide/bisacrylamide gel in parallel to an alkaline hydrolysis ladder and a denaturing T1 as described previously<sup>12</sup>. The radiolabeled RNAs were then visualized on autoradiographic film.

**Affinity measurements.** To measure the dissociation constant K<sub>d</sub>, the concentration of renatured and purified RNAs was progressively increased from 2.45 nM to 40  $\mu$ M for SRB-2 and from 3.9 nM to 4  $\mu$ M for 4C10 and o-Coral aptamer with 100 nM (for SRB-2) or 25 nM (for 4C10 and o-Coral) of Gemini-561 in 40 mM potassium phosphate buffer pH 7.5, 100 mM KCl, 1 mM MgCl<sub>2</sub> and 0.05% Tween 20. The fluorescence was measured on a microplate reader (ex/em = 580/620 nm, SpectraMax iD3, Molecular Devices).

**Expression vectors design.** The sequences coding for o-Coral or 20 nucleotides from the Broccoli aptamer (control) were introduced downstream a U6 promoter into a F30-scaffold (Supplementary Table 2) contained into a pUC57 vector (Proteogenix) via a restriction (SbfI)/ligation step. The entire sequences (pU6\_o-Coral\_F30 or pU6\_Ctrl\_F30) were then introduced into an eGFP-N1 vector (Clontech) using AflIII restriction sites. Alternatively, o-Coral-F30 or Ctrl-F30 sequences were introduced directly in the 3' UTR of the eGFP coding sequence under the control of a CMV promoter by a restriction (MfeI)/ligation step.

**RNA extraction and reverse transcription followed by quantitative PCR (RT-qPCR).** Total RNA was extracted from HeLa cells (transfected with plasmid carrying *gfp* of *gfp*-o-Coral genes) using the Monarch Total RNA Miniprep kit (New England Biolabs) according to the manufacturer's instructions. Then, 4  $\mu$ g of total RNA were used for DNase digestion using double-stranded DNase enzyme (Thermo Scientific). cDNA was then synthesized using Maxima minus H reverse

transcriptase (Thermo Scientific) following the manufacturer's instructions and specific cDNAs were then quantified by qPCR: 1  $\mu$ l of cDNA was introduced into 19  $\mu$ l of Ssofast evagreen qPCR mixture (Biorad) supplemented in eGFP specific primers or into 19  $\mu$ l of Ssoadvanced qPCR mixture (Biorad) supplemented with S18 or glyceraldehyde 3-phosphate dehydrogenase-specific primers and Taqman probes (Biorad). Real-time PCR was carried out on a CFX96 Touch System (Biorad) starting with an initial denaturation step of 2 min at 95 °C followed by 40 cycles of 5 s at 95 °C and 30 s at 60 °C. Reactions were run in triplicate in three independent experiments. The housekeeping genes *GAPDH* and *S18* were used as internal controls to normalize the variability in expression levels.

**Cell culture and transfection.** HeLa (ATCC CCL-2) and HEK293T (ATCC CRL-3216) cells were grown in Dulbecco's Modified Eagle Medium without phenol red (DMEM, Gibco-Invitrogen) supplemented with 10% FBS (Lonza), 1% L-glutamine (Sigma-Aldrich) and 1% antibiotic solution (penicillin/streptomycin, Sigma-Aldrich) at 37 °C in humidified atmosphere containing 5% CO<sub>2</sub>. U87MG (ATCC HTB-14) were grown in Minimum Essential Medium (MEM, Gibco-Invitrogen) supplemented with 10% FBS, 1% ultra-glutamine (Gibco-Invitrogen) and 1% antibiotic solution. RNA-coding constructs were transfected directly into a 35 mm glass-bottomed imaging dish (IBiDi) using the FuGene HD (Promega) transfecting agent following the recommended manufacturer's protocol. Imaging experiments were performed 16–48 h post-transfection.

**Cellular imaging.** Cells were seeded onto a 35 mm glass-bottomed imaging dish (IBiDi) at a density of 3–5  $\times$  10<sup>4</sup> cells per well 48 h before the microscopy measurement. Then, 16–24 h before imaging, cells were transfected with corresponding DNA plasmid. For imaging, the culture medium was removed and the attached cells were washed with Opti-MEM (Gibco-Invitrogen). Next, the cells were incubated in Opti-MEM with Hoechst (5  $\mu$ g ml<sup>-1</sup>) to stain the nuclei and in the presence of Gemini-561 dye (0.2  $\mu$ M) for 5 min, the living cells were washed twice with Opti-MEM and visualized in Opti-MEM or were fixed in 4% PFA in PBS for 5 min before being washed twice in PBS. The images were acquired in epi-fluorescence mode with a Nikon Ti-E inverted microscope, equipped with CFI Plan Apo  $\times$ 60 oil (numerical aperture = 1.4) objective, and a Hamamatsu Orca Flash 4 sCMOS camera. The acquisition settings were: Hoechst (ex, 395 nm; em, 510  $\pm$  42 nm), eGFP (ex, 470 nm; em, 531  $\pm$  40 nm), G561-o-Coral complex (ex, 550 nm; em, 595  $\pm$  40 nm) and Alexa-647 (ex, 638 nm; em, LP 647 nm). The images were recorded using NIS Elements and then processed with Icy software.

**Microinjection experiments.** Cells were seeded onto a 35 mm glass-bottomed imaging dish (IBiDi) at a density of 1  $\times$  10<sup>5</sup> cells per well 24 h before the microscopy measurement. For imaging, the culture medium was removed and the attached cells were washed with Opti-MEM (Gibco-Invitrogen). Next, the cells were incubated in Opti-MEM with Hoechst (5  $\mu$ g ml<sup>-1</sup>) to stain the nuclei. *In vitro* transcribed and purified aptamers were preincubated with respective fluorogen for 10 min in selection buffer to form a complex at the corresponding concentrations indicated in the figures. Microinjection parameters: Pi = 90 hPa; Ti = 0.3 s and Pc = 10 hPa. The images were acquired in epi-fluorescence mode with a Nikon Ti-E inverted microscope, equipped with CFI Plan Apo  $\times$ 60 oil (numerical aperture = 1.4) objective and a Hamamatsu Orca Flash 4 sCMOS camera. The acquisition settings were: Hoechst (ex, 395 nm; em, 475  $\pm$  50 nm), Broccoli (ex, 470 nm; em, 531  $\pm$  40 nm); Corn (ex, 470 nm; em, 531  $\pm$  40 nm); Mango (ex, 470 nm; em, 531  $\pm$  40 nm) and Coral (ex, 550 nm; em, 595  $\pm$  40 nm). The images were recorded using NIS Elements and then processed with Icy software.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

The data supporting the findings of this study are available within the paper and its Supplementary Information file.

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- n/a Confirmed
- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
  - A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
  - The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
  - A description of all covariates tested
  - A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
  - A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
  - For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
  - For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
  - For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
  - Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

- Data collection:
- Data analysis:

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

### Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The experiments were performed in at least three independent replicates.
Data exclusions	Dead cells were excluded from analysis. No other data were excluded.
Replication	All the experiments were readily reproduced and the replication numbers are provided in the caption of corresponding figures.
Randomization	Cells from randomly chosen fields were used for analysis.
Blinding	The investigators were not blinded to group allocation.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involvement in the study	n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines	<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	Cell lines were obtained from the American Type Culture Collection (ATCC) through UMR7021 (University of Strasbourg) cell lines bank. HEK293T (ATCC-CRL-11268), HeLa (ATCC-CCL-2) and U87MG (ATCC® HTB-14™)
Authentication	The cell lines were not authenticated prior to use.
Mycoplasma contamination	The cell lines were not tested for mycoplasma contamination prior to use. However, no signs of mycoplasma were observed during cell imaging experiments when Hoechst 33258 was used.
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	No commonly misidentified cell lines were used.

In the format provided by the authors and unedited.

## A dimerization-based fluorogenic dye-aptamer module for RNA imaging in live cells

Farah Bouhedda<sup>1,3</sup>, Kyong Tkhe Fam<sup>2,3</sup>, Mayeul Collot<sup>2\*</sup>, Alexis Autour<sup>1</sup>, Stefano Marzi<sup>1</sup>,  
Andrey Klymchenko<sup>2\*</sup> and Michael Ryckelynck<sup>1\*</sup>

<sup>1</sup>Université de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, UPR 9002, Strasbourg, France. <sup>2</sup>Université de Strasbourg, Laboratoire de Bioimagerie et Pathologies, UMR 7021 CNRS, Illkirch, France. <sup>3</sup>These authors contributed equally: Farah Bouhedda, Kyong Tkhe Fam. \*e-mail: mayeul.collot@unistra.fr; andrey.klymchenko@unistra.fr; m.ryckelynck@unistra.fr



## **Supplementary Tables and Figures for the manuscript:**

### **A dimerization-based fluorogenic dye-aptamer module for RNA imaging in live cells**

Farah Bouhedda<sup>1,3</sup>, Kyong Tkhe Fam<sup>2,3</sup>, Mayeul Collot<sup>2\*</sup>, Alexis Autour<sup>1</sup>, Stefano Marzi<sup>1</sup>, Andrey Klymchenko<sup>2,4\*</sup> & Michael Ryckelynck<sup>1,4\*</sup>

1. Université de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, UPR 9002, F-67000 Strasbourg, France

2. Nanochemistry and Bioimaging group, Laboratoire de Bioimagerie et Pathologies, CNRS UMR 7021, Université de Strasbourg, 67401 Illkirch, France

3. These authors equally contributed to this work

4. Co-last authors

\* mayeul.collot@unistra.fr; andrey.klymchenko@unistra.fr; m.ryckelynck@unistra.fr

**Supplementary Table 1** | Affinity between RNA aptamers and Sulforhodamine B, Gemini-561 or its alkyne derivative.

<b>Dye/RNA couple</b>	<b><math>K_D</math> affinity constant</b>
Sulforhodamine B/SRB2 <sup>a</sup>	310 ± 60.0 nM
Sulforhodamine B/o-Coral	> 400 nM
Gemini-561/SRB2	441 ± 167.5 nM
Gemini-561/o-Coral	73 ± 1.5 nM
Gemini-561-alkyne/o-Coral	188 ± 2.5 nM

<sup>a</sup> As reported in Holeman, L.A., Robinson, S.L., Szostak, J.W. & Wilson, C. Isolation and characterization of fluorophore-binding RNA aptamers. *Fold Des* **3**, 423-31 (1998).  $K_D$  values are the mean of n = 3 independent experiments ± 1 SD.

**Supplementary Table 2** | Sequence of o-Coral constructs.

Name	Sequence
o-Coral <sup>a</sup>	5' <b>GGGAGACAGCTAGAGTAC</b> AGAACCCCGCTTCGGCGGTGATGGAGAGGCGCAAGGTTAACCGCC TCAGGTTCCGGTGACGGGGCCTCGCTTCGGCGATGATGGAGAGGCGCAAGGTTAACCGCCTCAGG TTCT <b>GACACGAGCACAGTGTAC</b> 3'
o-Coral in F30 scaffold a, b	5' <i>GTGCTCGCTTCGGCAGCACATATACTAGTCGACTTGCCATGTGTATGTGGGCCTGCAGGGGGA</i> <b>GACAGCTAGAGTAC</b> AGAACCCCGCTTCGGCGGTGATGGAGAGGCGCAAGGTTAACCGCCTCAGGT TCCGGTGACGGGGCCTCGCTTCGGCGATGATGGAGAGGCGCAAGGTTAACCGCCTCAGGTTCT <b>GA</b> <b>CACGAGCACAGTGTACC</b> TGCAGGCCACATACTCTGATGATCCTTCGGGATCATTTCATGGCAAT <i>CTAGAGCGGACTTCGGTCCGCTTTT3'</i>

<sup>a</sup> The sequence of 5' and 3' extensions are shown in bold. <sup>b</sup> The sequence of the F30 scaffold is underlined. The sequence of 27 first nucleotides of U6 promoter (5' end) and of the U6 terminator (3' end) are italicized.

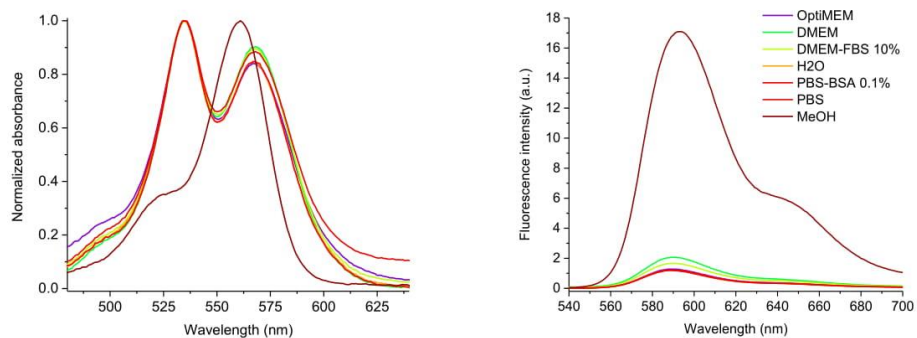


**Supplementary Table 3** | Fluorescence Correlation Spectroscopy (FCS) analysis of the fluoromolecules.<sup>a</sup>

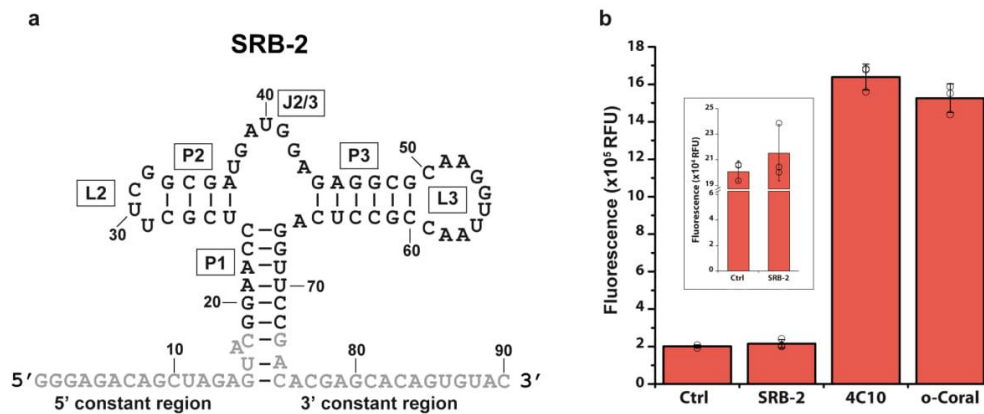
Fluoromodule/ reference dye	n	$\tau_{\text{corr}}$ , ms	c, nM	Size, nm	Brightness
Tetramethylrhodamine	11.64	0.032	50	1	1
o-Coral/Gemini-651	43.22	0.22	185.71	6.86	1.14*
RFP	18.2	0.077	84.42	1.53	0.06
mCherry	4.09	0.102	17.54	3.45	0.16
Fluorescein	13.03	0.026	50	1	1
Broccoli/ DFHBI-1T	14.16	0.16	54.35	6.27	0.18
Corn/ DFHO	54.7	0.063	209.94	2.56	0.62
eGFP	20.2	0.075	75.47	2.59	0.16

<sup>a</sup>n-number of emissive species per excitation volume;  $\tau_{\text{corr}}$  - correlation time; size - hydrodynamic diameter of fluorescent specie; brightness with respect to one molecule of the standard (rhodamine B for o-Coral, fluorescein for Broccoli and Corn). Mango was excluded from this study since its photophysical properties did not fit the analysis settings. Concentrations of fluorogen and corresponding aptamer were systematically 200 nM and 1  $\mu$ M, respectively.

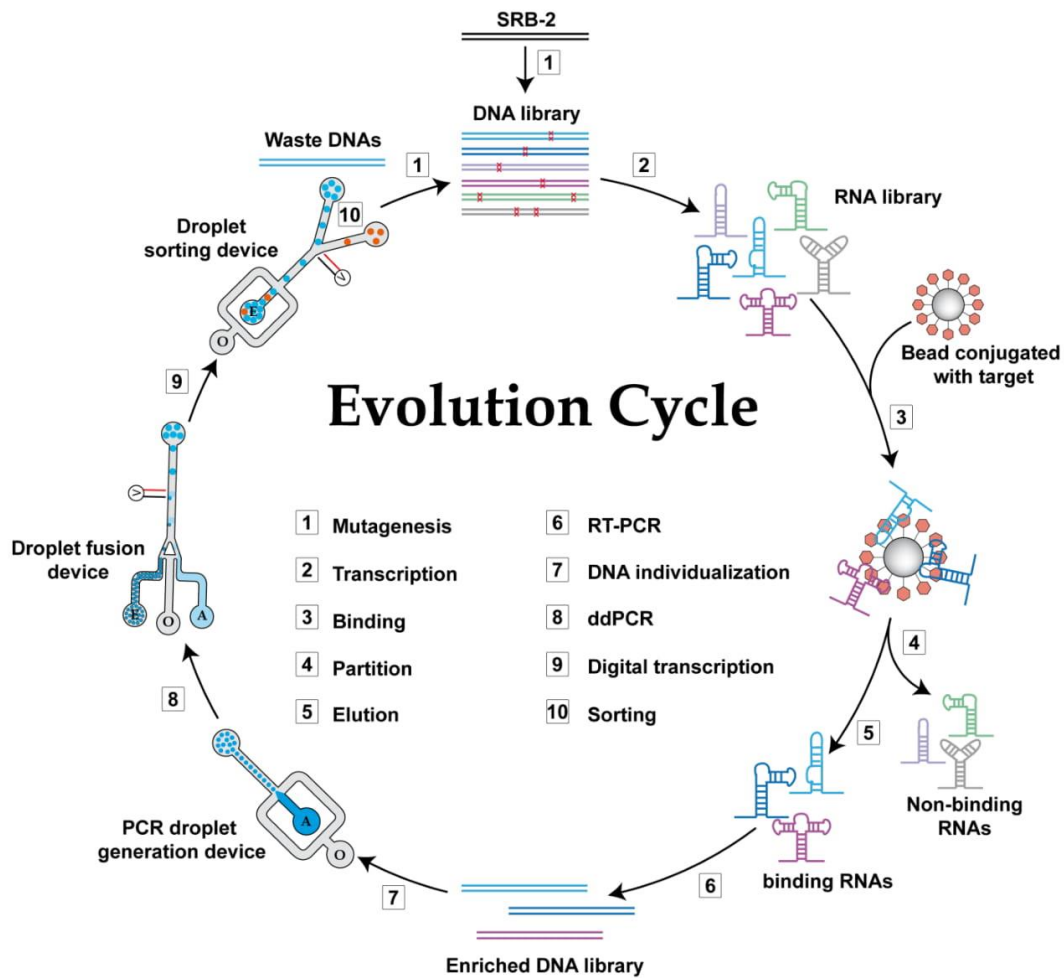
\*This value corresponds to 532 nm excitation, which is less efficient for excitation than that for the TMR standard, because of the red shifted absorption of the former.



**Supplementary Figure 1** | Normalized absorption and emission spectra of Gemini-561 in different aqueous media mimicking cellular environments. Results were found similar in  $n = 3$  independent experiments.

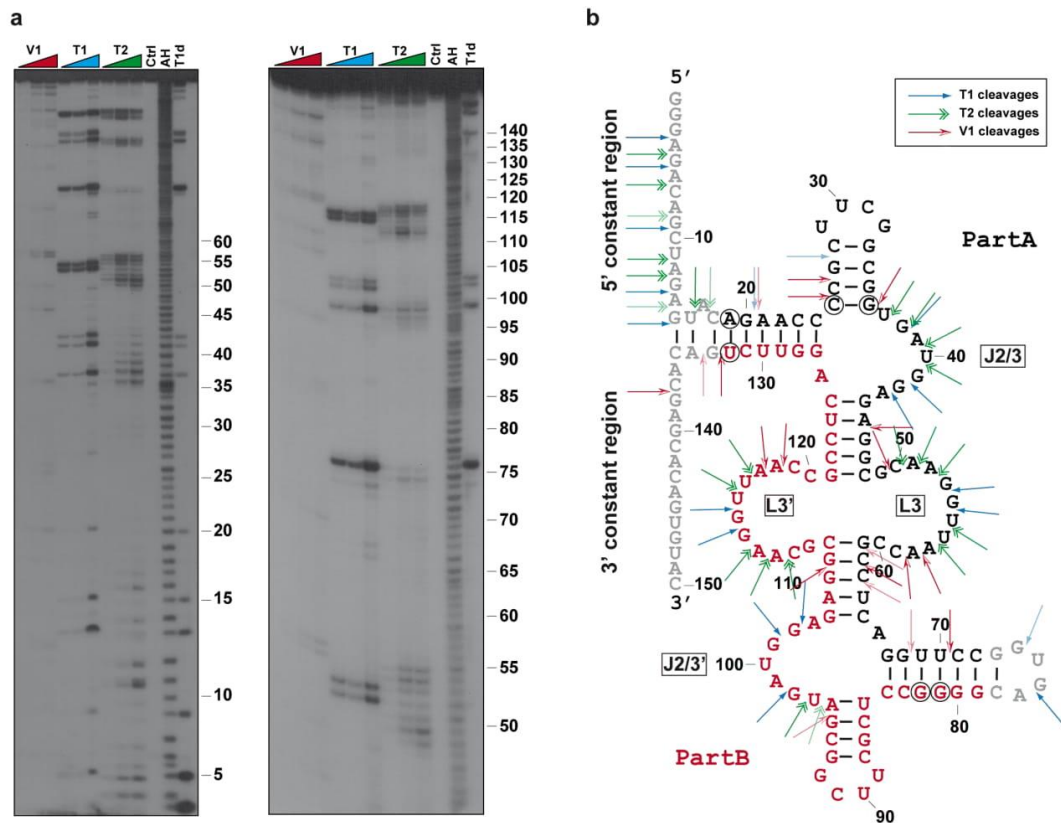


**Supplementary Figure 2 | Gemini-561 activation by SRB-2 aptamer and its derivatives.** (a) Secondary structure model of SRB-2 aptamer as originally proposed in Holeman, L.A., Robinson, S.L., Szostak, J.W. & Wilson, C. Isolation and characterization of fluorophore-binding RNA aptamers. *Fold Des* **3**, 423-31 (1998). Paired regions (P1, P2 and P3) are distinguished from Loop (L2 and L3) and Junction (J2/3) regions. Constant sequence regions appended for RT-PCR amplification purposes are shown in gray. (b) Fluorogenic capacity of SRB-2 and its evolved forms 4C10 and o-Coral. 500 nM of RNAs were incubated with 50 nM of Gemini-561 and the fluorescence was measured at  $\lambda_{ex/em} = 560/600$  nm. The values are the mean of  $n = 3$  independent experiments, each measurement being shown as an open circle. The error bars correspond to  $\pm 1$  standard deviation.

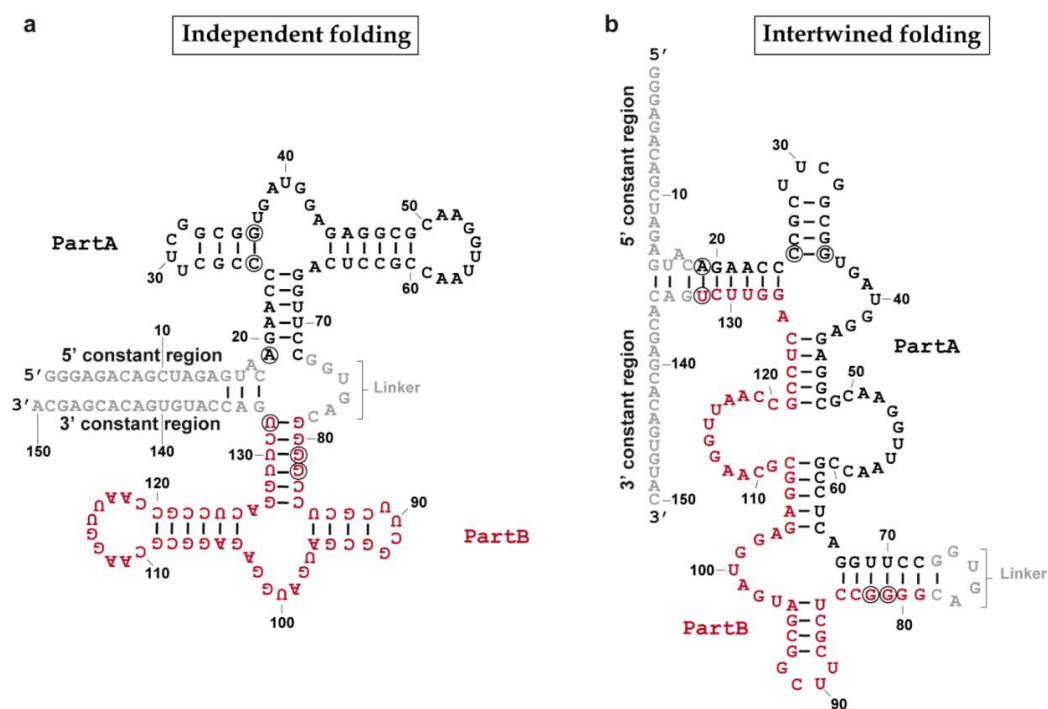


**Supplementary Figure 3 | Overall *in vitro* evolution strategy.** Each round of evolution cycle consisted of 10 main steps. SRB-2 was used as a template for error-prone PCR (step 1) to create a DNA mutant library that was *in vitro* transcribed (step 2). Resulting RNAs were then selected for their binding capacity *via* a SELEX (Systematic Evolution of Ligands by Exponential enrichment) approach (steps 3, 4 and 5) prior to being screened for their light-up capacity using  $\mu$ IVC (steps 7, 8, 9 and 10). For steps performed in microfluidic chips, Oil (O) and Aqueous phase (A) inlets are labeled together with inlets and outlets where Emulsion (E) were respectively reinjected and collected. Finally, the enriched pool was reamplified by an error-prone PCR (Step 1) before re-entering the whole process again. 4C10 was obtained after 4 rounds of this evolution cycle.



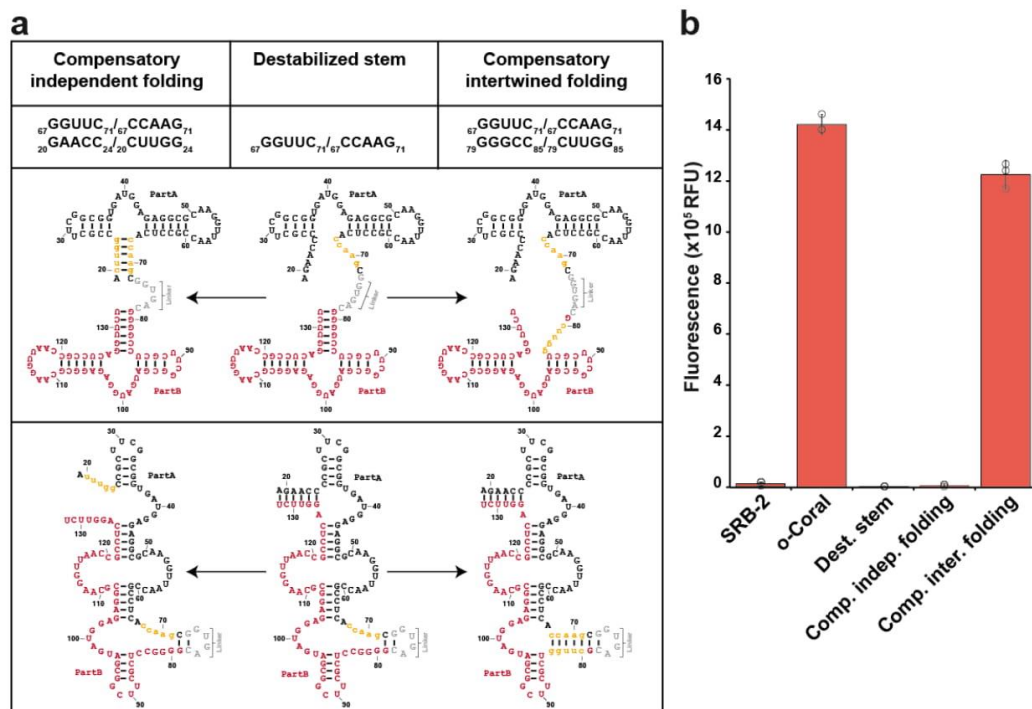


**Supplementary Figure 5 | Probing of o-Coral secondary structure.** (a) Probing experiment. Radioactively labeled o-Coral RNA was subjected to digestion by V1, T1 or T2 nucleases prior to analyzing the digestion products on 10 % polyacrylamide denaturing gels. The increased concentration of the enzymes is schematized by the colored triangles (V1: 0.001 U/ $\mu$ L – 0.002 U/ $\mu$ L – 0.004 U/ $\mu$ L, T1: 0.25 U/ $\mu$ L – 0.5 U/ $\mu$ L – 1 U/ $\mu$ L, T2: 0.0125 U/ $\mu$ L – 0.025 U/ $\mu$ L – 0.05 U/ $\mu$ L). Ctrl lane corresponds to an enzyme-free experiment, AH stands for Alkaline Hydrolysis in which o-Coral was statistically hydrolyzed, dT1 stands for denaturing T1 cleavages. The numbers on the right refer to o-Coral nucleotides. (b) Secondary structure model of o-Coral. T1, T2 and V1 cleavage sites are indicated respectively by the blue, green and red arrows. SRB-2 derived monomers are shown in black or red (Part A and B), whereas constant regions and linker are shown in gray. Acquired mutations found to contribute to o-Coral function are circled in black.



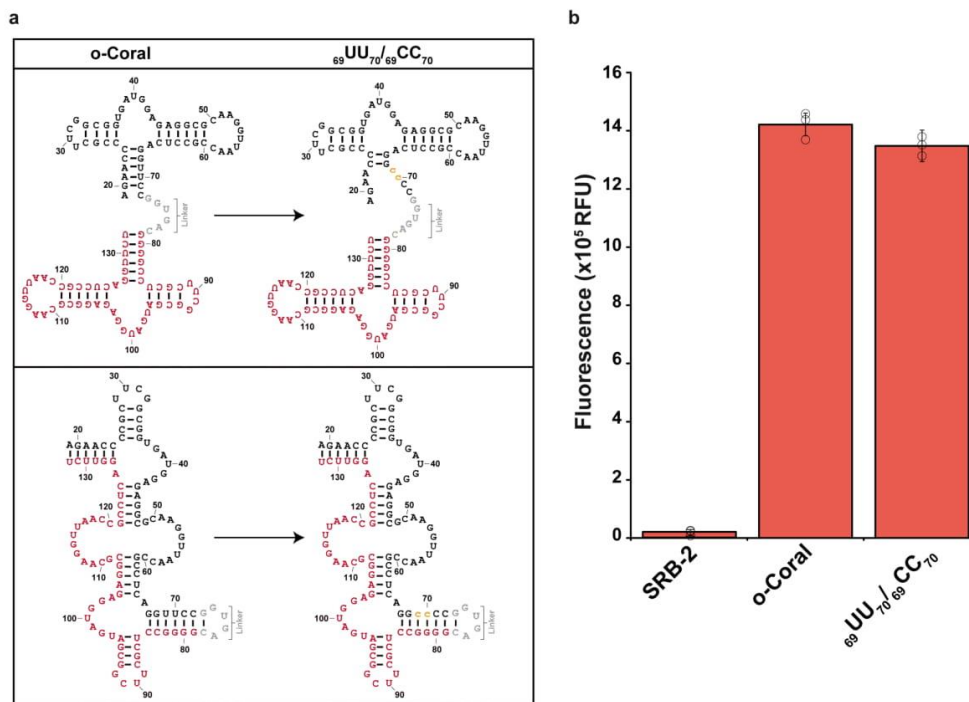
**Supplementary Figure 6 | Secondary structure models of o-Coral aptamer.** (a) Independent folding model. In this model, each SRB-2-derived monomer adopts an independent folding and closely resemble the original SRB-2 molecule associated by single stranded linker region. (b) Intertwined folding model. In this model, both each SRB-2 derived monomers fold on each other and form an intertwined structure. On both models, SRB-2 derived monomers are shown in black or red (Part A and B), whereas constant regions and linker are shown in gray. Acquired mutations found to contribute to o-Coral function are shown as circled in black.



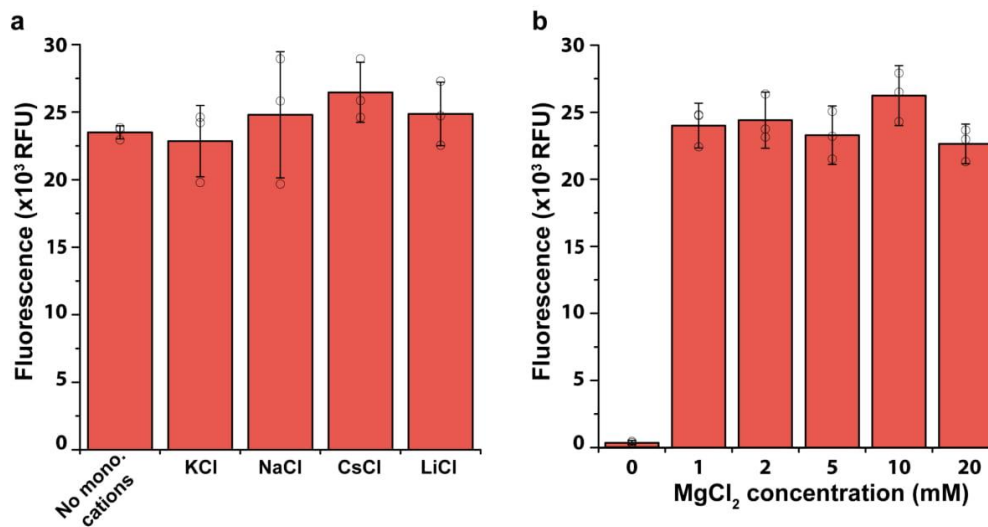


**Supplementary Figure 7 | Refinement of structural model using P1 compensatory mutants. (a)** Three mutants were generated: a destabilized mutant (Destabilized stem:  ${}_{67}\text{GGUUC}_{71}/{}_{67}\text{CCAAG}_{71}$ ) of o-Coral and two potentially compensatory mutants; the first one based on the independent folding model ( ${}_{67}\text{GGUUC}_{71}/{}_{67}\text{CCAAG}_{71-20}\text{GAACC}_{24}/{}_{20}\text{CUUGG}_{24}$ ) and the second one based on the intertwined folding model ( ${}_{67}\text{GGUUC}_{71}/{}_{67}\text{CCAAG}_{71-79}\text{GGGCC}_{85}/{}_{79}\text{CUUGG}_{85}$ ). SRB-2 derived sequences (Part A and B) are shown in black or red whereas the linker sequence is shown in grey. Implemented mutations described before are shown in orange. **(b)** Impact of implemented mutations on o-Coral aptamer fluorogenicity. 500 nM of RNAs were incubated with 50 nM of Gemini-561 and the fluorescence was measured at  $\lambda_{\text{ex/em}} = 560/600$  nm. The values are the mean of  $n = 3$  independent experiments, each measurement being shown as an open circle. The error bars correspond to  $\pm 1$  standard deviation.

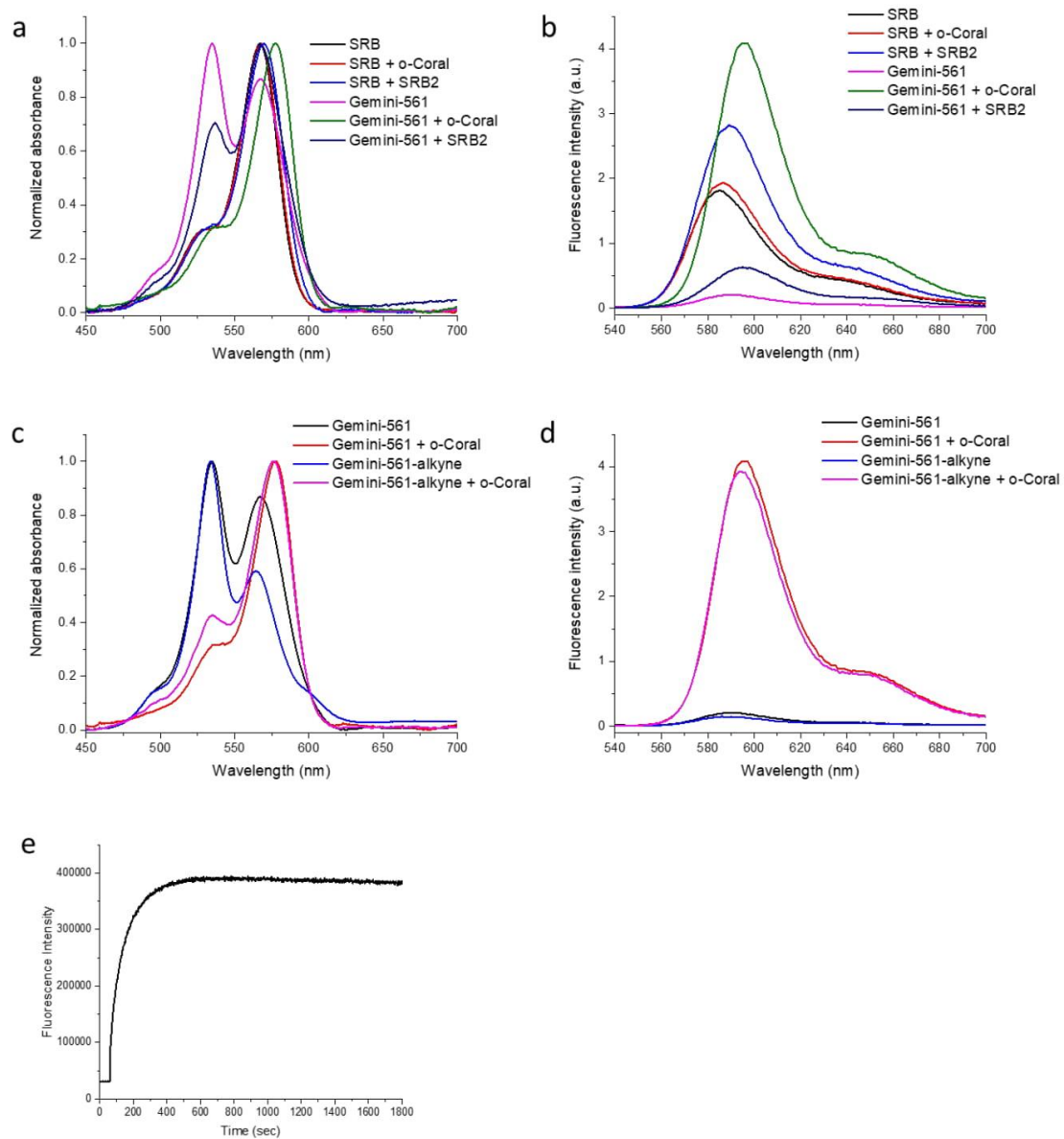




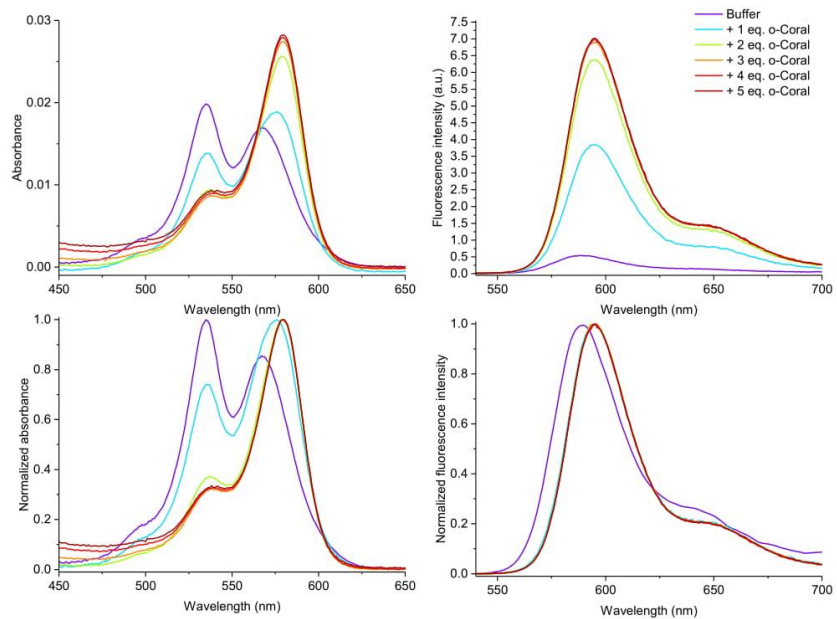
**Supplementary Figure 8 | Refinement of structural model using CC double mutant.** (a) Representation of o-Coral and the <sup>69</sup>UU<sub>70</sub>/<sup>69</sup>CC<sub>70</sub> o-Coral double mutant according to the independent folding model (upper part) and the intertwined folding model (lower part). SRB-2 derived sequences (Part A and B) are shown in black or red whereas the linker sequence is shown in grey. Implemented mutations described below is shown in orange. (b) Impact of the implemented mutation on o-Coral aptamer fluorogenicity. 500 nM of RNAs were incubated with 50 nM of Gemini-561 and the fluorescence was measured at  $\lambda_{ex/em} = 560/600$  nm. The values are the mean of  $n = 3$  independent experiments, each measurement being shown as an open circle. The error bars correspond to  $\pm 1$  standard deviation.



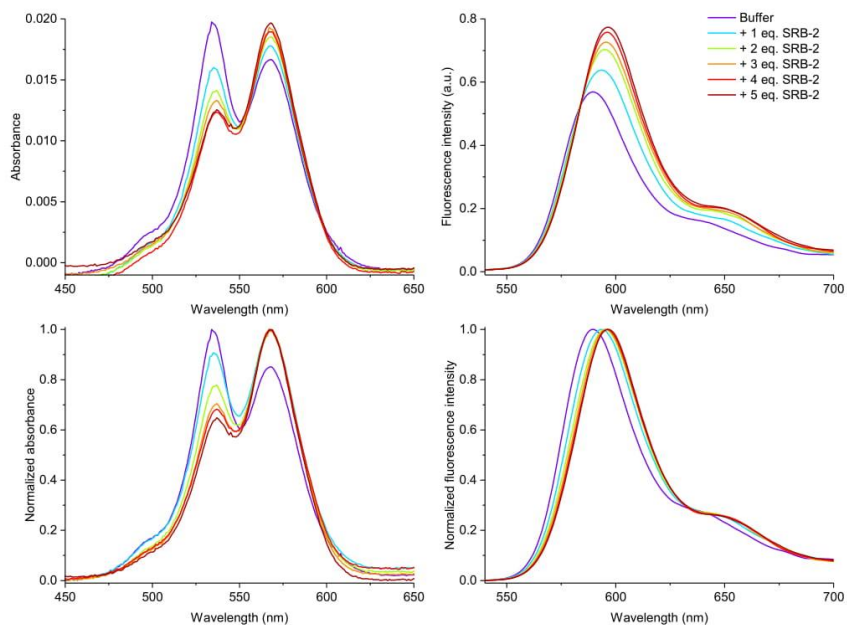
**Supplementary Figure 9 | Salt dependency of Gemini561/o-Coral module.** (a) Monovalent ions dependency of o-Coral. o-Coral RNA and Gemini-561 were mixed in a solution containing 40 mM Phosphate buffer pH7.5, 100 mM KCl or NaCl or CsCl or LiCl or in the absence of monovalent cations, 2 mM MgCl<sub>2</sub> and 0.05% Tween-20 and the fluorogenic capacity was measured. (b) Magnesium dependency of o-Coral. o-Coral RNA and Gemini-561 were mixed in a solution containing 40 mM Phosphate buffer pH7.5, 100 mM KCl, the indicated concentration of MgCl<sub>2</sub> and 0.05% Tween-20 and the fluorogenic capacity was measured. For both condition (a and b), 500 nM of RNAs were incubated with 50 nM of Gemini-561 and the fluorescence was measured at  $\lambda_{ex/em} = 560/600$  nm. The values are the mean of  $n = 3$  independent experiments, each measurement being shown as an open circle. The error bars correspond to  $\pm 1$  standard deviation.



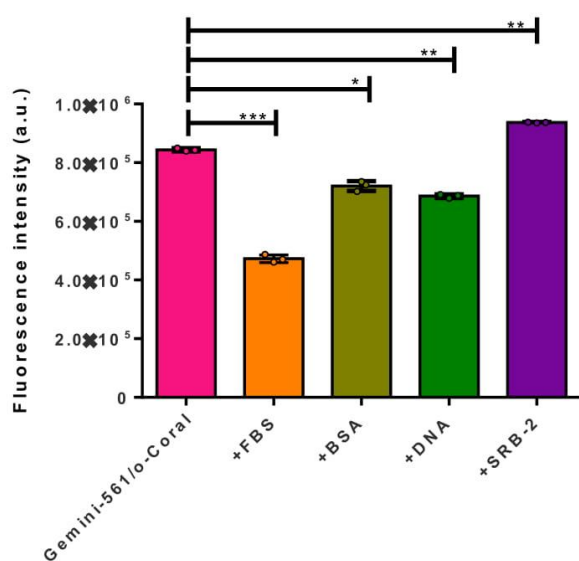
**Supplementary Figure 10** | (a) Normalized absorption and (b) emission spectra of Gemini-561 or SRB (200 nM) in absence and in the presence of different RNA aptamers (600 nM). (c) Normalized absorption and (d) emission spectra of Gemini-561 or Gemini-561-alkyne (200 nM) in absence and in the presence of o-Coral (600 nM). Results were found similar in  $n = 3$  independent experiments. (e) Kinetics of interaction between Gemini-561 (100 nM) and o-Coral (300 nM) measured as emission intensity at 586 nm as a functional of time. Excitation wavelength was 530 nm in all cases.



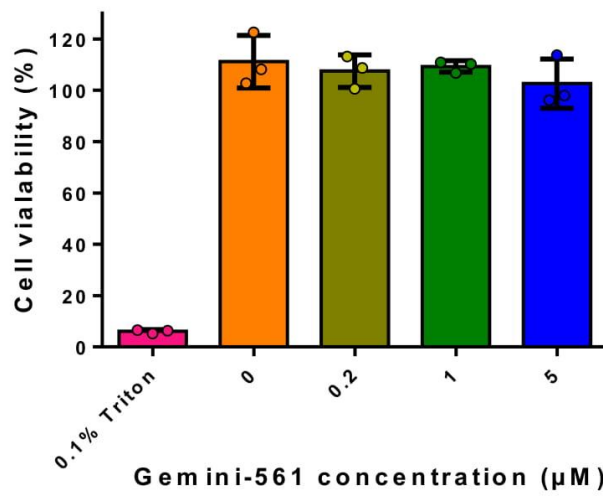
**Supplementary Figure 11** | Absorption and emission spectra of Gemini-561 (200 nM) in the presence of increasing concentrations (equivalents, eq.) of o-Coral. Excitation wavelength was 530 nm. Measurements were performed once.



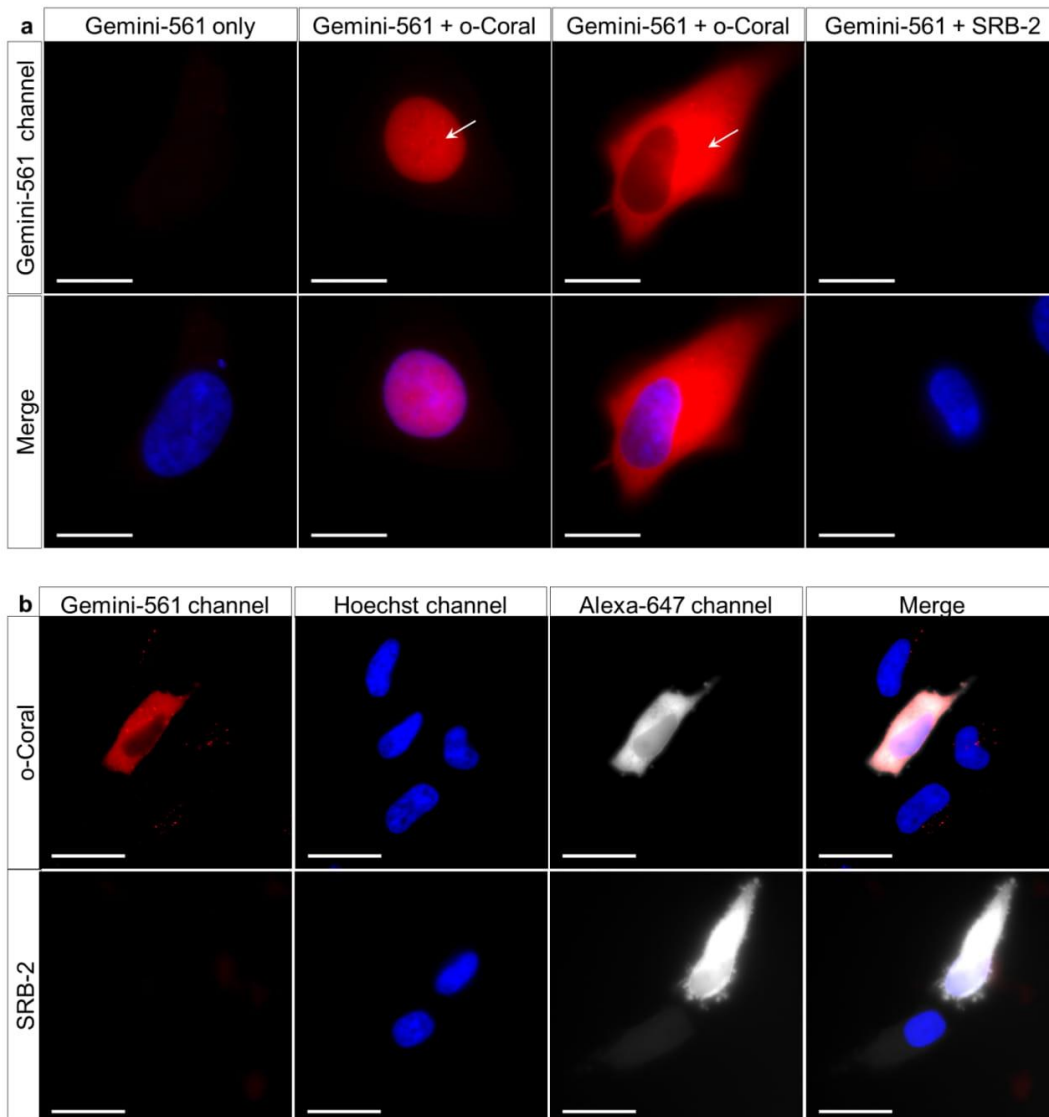
**Supplementary Figure 12** | Absorption and emission spectra of Gemini-561 (200 nM) in the presence of increasing concentration (equivalents, eq.) of SRB-2 aptamer. Excitation wavelength was 530 nm. Measurements were performed once.



**Supplementary Figure 13** | Effect of biomolecules and biological medium on the fluorescence intensity of Gemini-561/o-Coral (1/1 molar ratio) complex at 0.2  $\mu\text{M}$  concentration. After Gemini-561/o-Coral complex was formed, the mixture was incubated with the corresponding biomolecule (BSA 10 mg/mL, non-targeted DNA 50  $\mu\text{M}$  or SRB-2 aptamer 0.2  $\mu\text{M}$ ) or biological medium (FBS 10%) for 15 min and the fluorescence was recorded at 596 nm. Excitation wavelength was 530 nm. The values are the mean of  $n = 3$  independent experiments, each measurement being shown as a colored dot. The error bars correspond to  $\pm 1$  standard deviation. Statistical significance based on ANOVA analysis: \*  $p > 0.05$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.001$ .

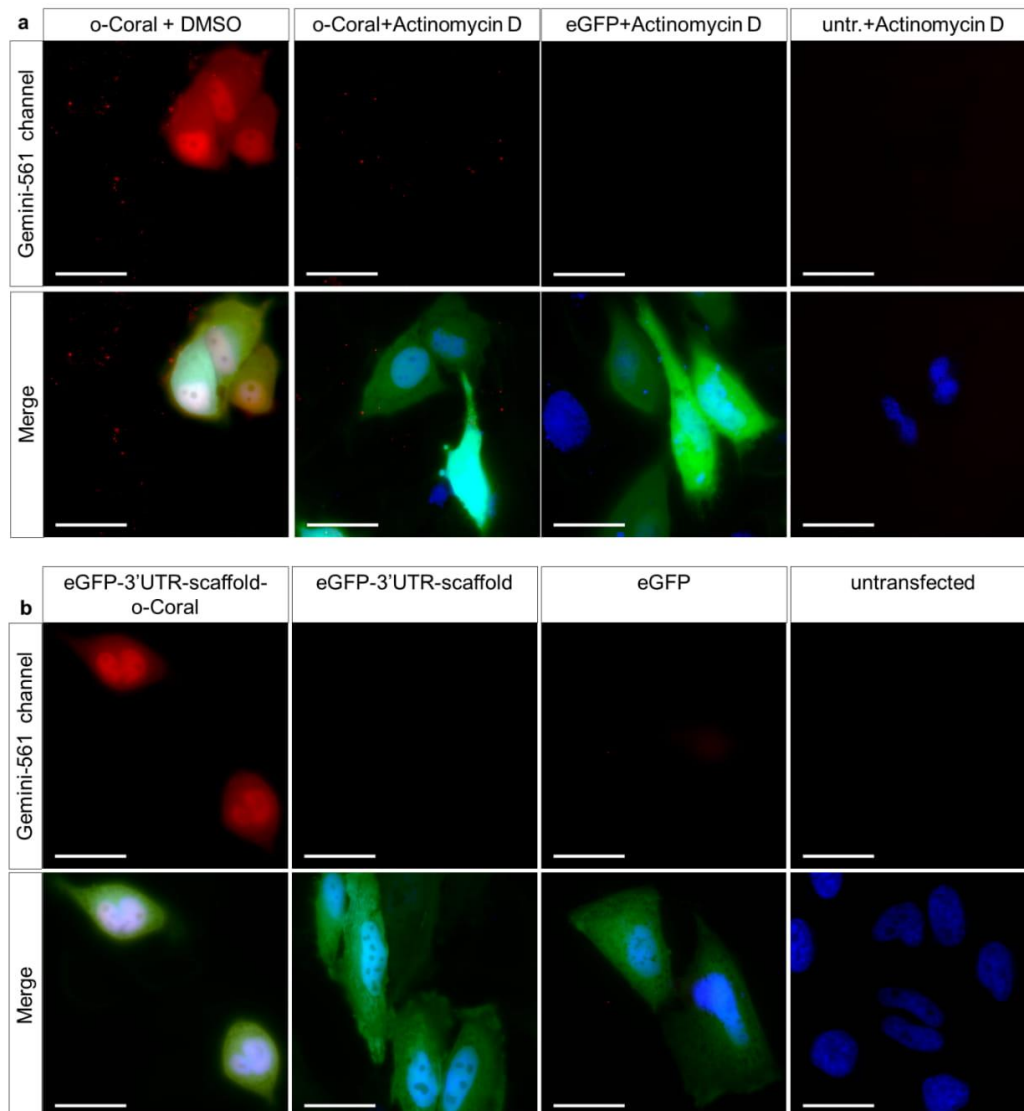


**Supplementary Figure 14 | Cytotoxicity assay of Gemini-561.** HeLa cells were incubated with various concentration of Gemini-561 and their viability was assessed after 24 hours using MTT test. An incubation with 0.1% Triton X100 was used as negative control. The values are the mean of  $n = 3$  independent experiments, each measurement being shown as a colored dot. The error bars correspond to  $\pm 1$  standard deviation.



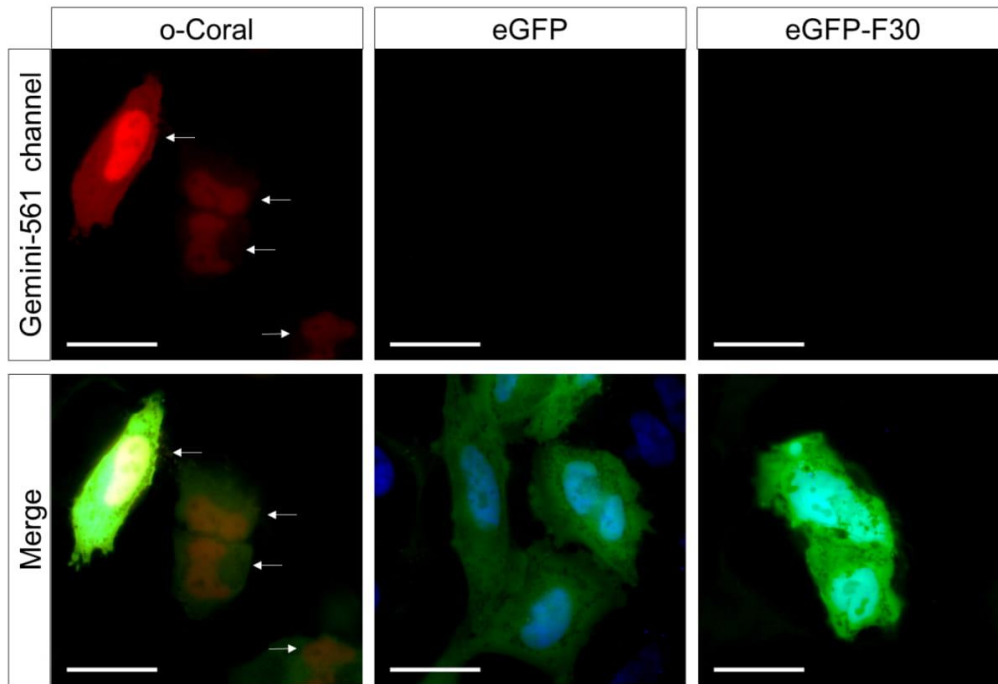
**Supplementary Figure 15 | Microinjection in HeLa cells.** (a) Microinjection of Gemini-561 ( $1 \mu\text{M}$ ) alone (in cytosol), complex of Gemini-561/o-Coral ( $1 \mu\text{M}$ ) or Gemini-561/SRB-2 ( $1 \mu\text{M}$ ) (in cytosol). Arrows show that Gemini-561/o-Coral complex was microinjected into either nucleus or cytosol. Scale bar is  $20 \mu\text{m}$ . (b) Microinjection of o-Coral ( $52 \mu\text{M}$ ) or SRB-2 ( $52 \mu\text{M}$ ) with Dextran-Alexa-647 conjugate ( $10 \mu\text{M}$ ) in cells pre-treated with Gemini-561 ( $200 \text{ nM}$ ) for 5 min. Microinjection parameters:  $P_i=90 \text{ [hPa]}$ ;  $T_i=0.3 \text{ [s]}$ ;  $P_c=10 \text{ [hPa]}$ . The nucleus was stained with Hoechst ( $5 \mu\text{g/mL}$ ). The images were acquired using a 10s exposure time. Gemini-561 in red (ex:  $550 \text{ nm}$ , em:  $595 \pm 40 \text{ nm}$ ), Hoechst in blue (ex:  $395 \text{ nm}$ , em:  $510 \pm 42 \text{ nm}$ ) and Alexa-647 in white (ex:  $638 \text{ nm}$ , em:  $810 \pm 90 \text{ nm}$ ). Scale bar is  $30 \mu\text{m}$ . Results were found similar in  $n = 3$  independent experiments.





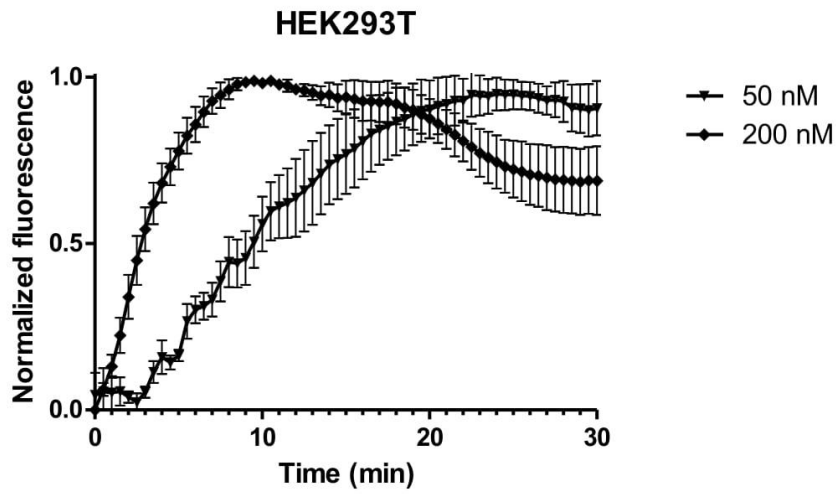
**Supplementary Figure 16 | Live-cell imaging of o-Coral expressed from pol. II and pol. III promoter.**

(a) Live cell imaging of HeLa cells expressing o-Coral from the U6-promoter in the absence and presence of Actinomycin D, cells expressing eGFP only and untransfected (untr.) cells treated with Actinomycin D. (b) Live cell imaging of HeLa cells expressing the gfp mRNA labelled with single copy of o-Coral. Cells expressing the gfp mRNA with or without scaffold inserted and untransfected cells were used as negative controls. (a-b) Cells were incubated with Gemini-561 (200 nM) for 5 min before imaging. Hoechst was used to stain the nucleus (5  $\mu$ g/mL). The images were acquired using a 500 ms exposure time. Gemini-561 in red (ex: 550 nm, em: 595 $\pm$ 40 nm), Hoechst in blue (ex: 395 nm, em: 510 $\pm$ 42 nm) and eGFP in green (ex: 470 nm, em: 531 $\pm$ 40 nm). Results were found similar in n = 3 independent experiments. Scale bar is 30 $\mu$ m.

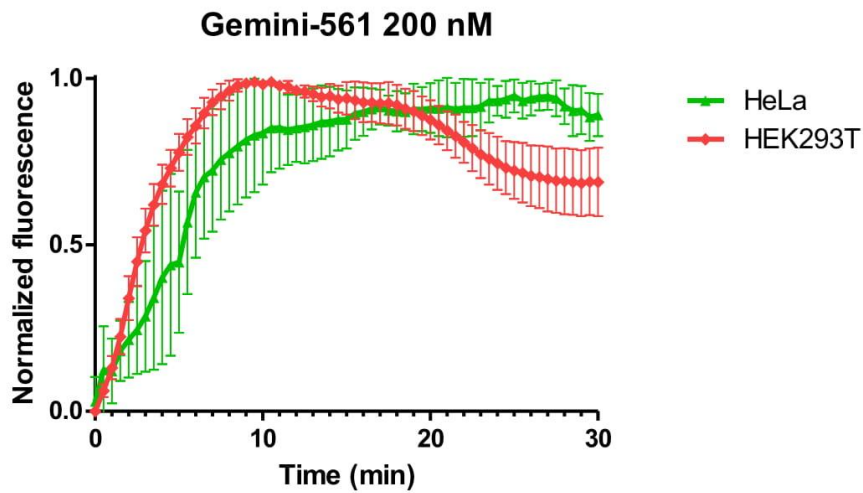


**Supplementary Figure 17 | Live-cell imaging of transfected HeLa cells expressing eGFP and o-Coral, eGFP only or eGFP and F30 scaffold only.** Top panel shows Gemini-561 channel only. Bottom channel shows merged all channels. Cells were incubated with Gemini-561 (200 nM) for 5 min. White arrows on the images depict the correlation between expression of eGFP and o-Coral as well as the different transcription states of cells. The images were acquired using a 500 ms exposure time. Gemini-561 in red (ex: 550 nm, em: 595±40 nm), Hoechst in blue (ex: 395 nm, em: 510±42 nm) and eGFP in green (ex: 470 nm, em: 531±40 nm). Results were found similar in n = 3 independent experiments. Scale bar is 30µm.

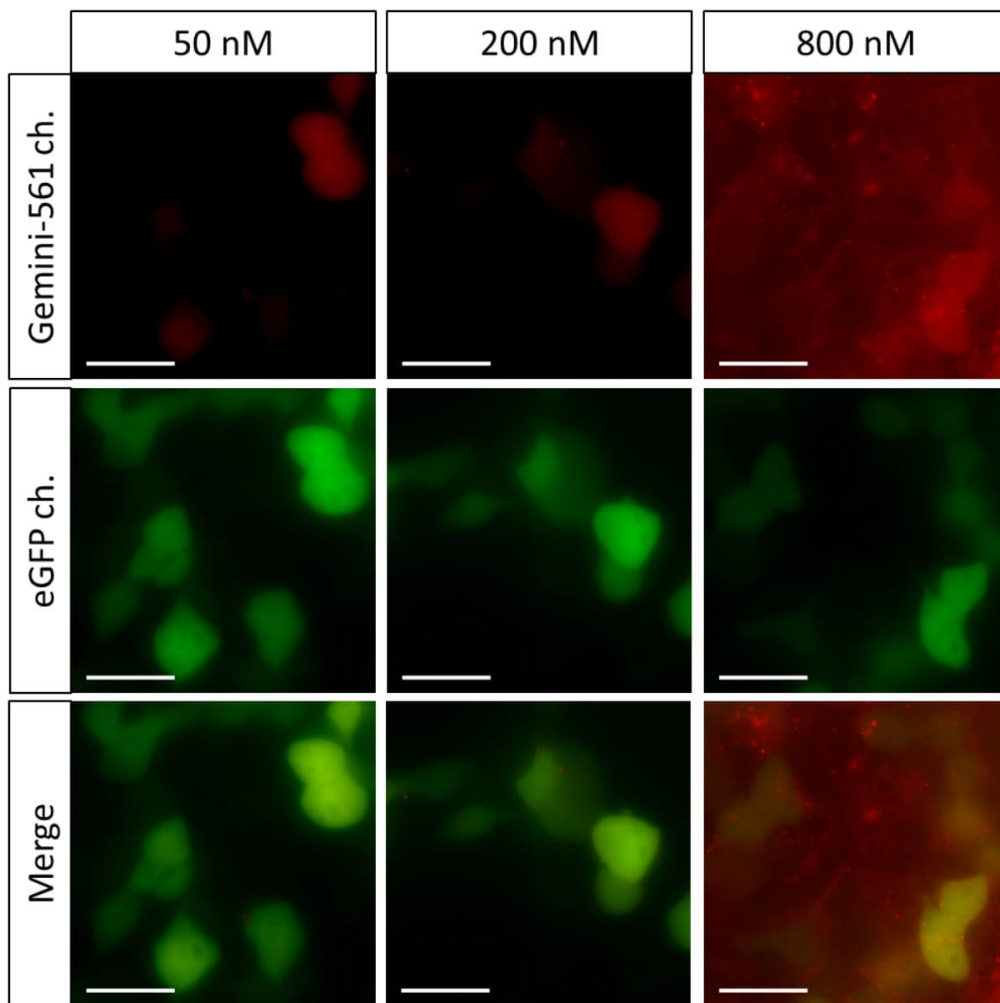
a



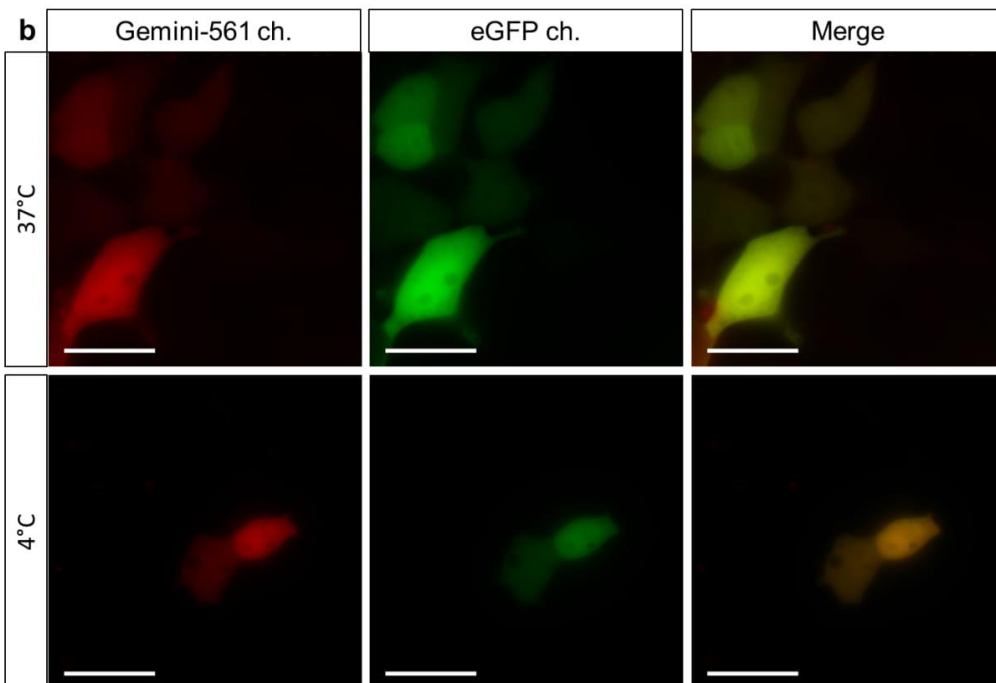
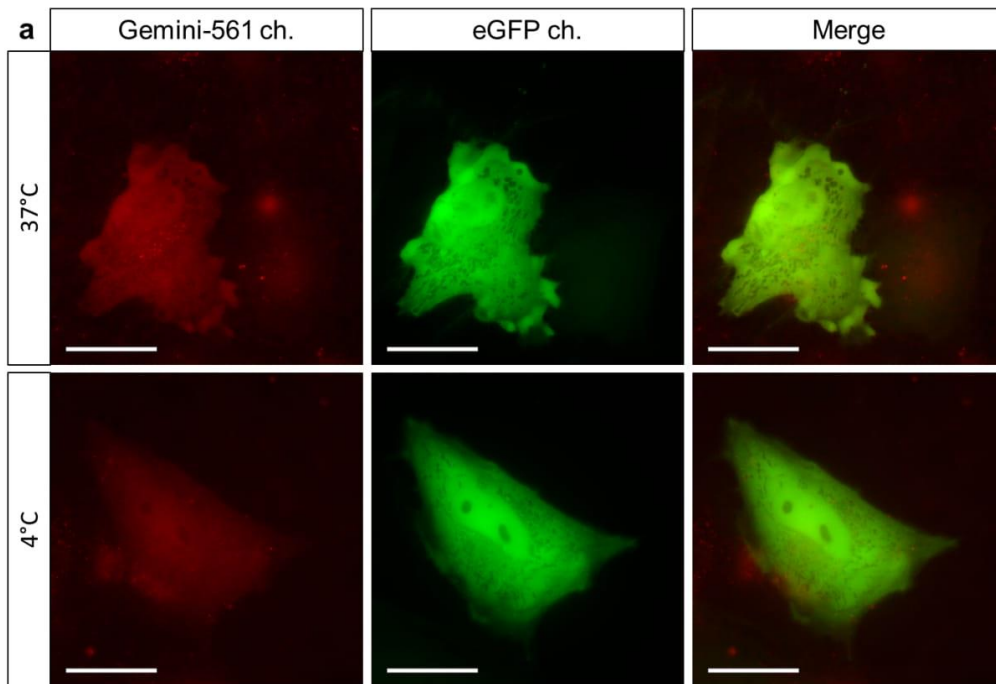
b

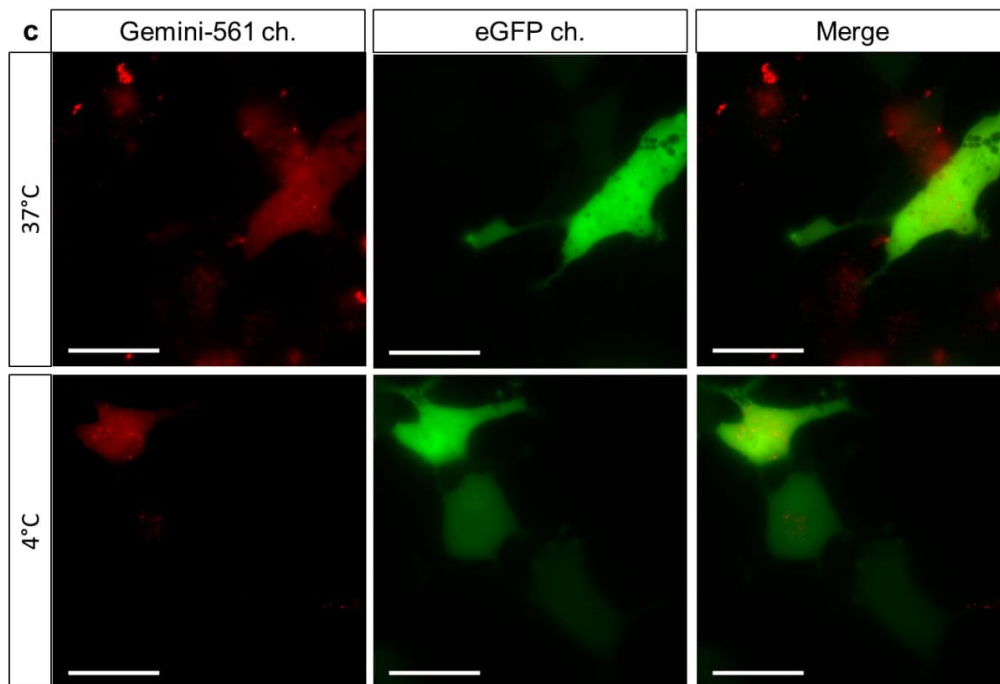


**Supplementary Figure 18** | Kinetic studies of Gemini-561 uptake in *U6-o-Coral*-expressing live cells. (a) Concentration effect (50 nM and 200 nM) on Gemini-561 uptake in *U6-o-Coral*-expressing HEK293T cells. (b) Cell line effect (HeLa, HEK293T) on Gemini-561 (200 nM) uptake. Data represent average values of Gemini-561/*o-Coral* fluorescence intensity normalized to the maximum  $\pm$  1 S.D. extracted from images (n=6). Experiments were independently performed at least 3 times.



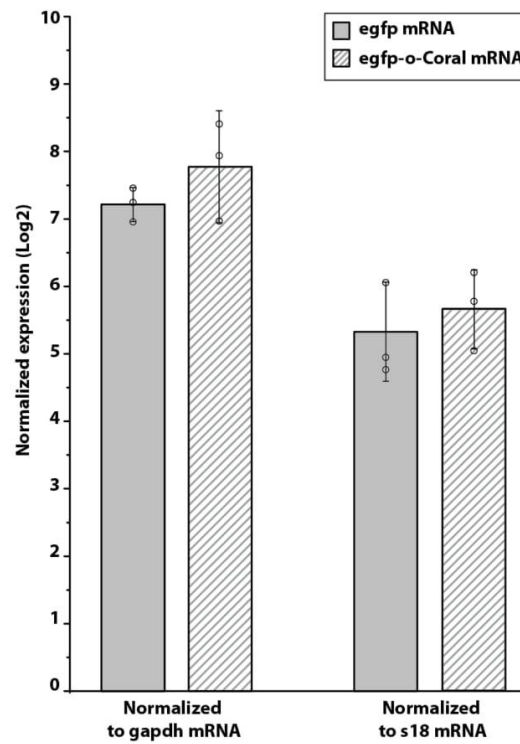
**Supplementary Figure 19** | Live cell imaging of HEK293T cells expressing o-Coral from the U6-promoter, incubated for 5 min with Gemini-561 at concentrations 50 nM, 200 nM, 800 nM. The images were acquired using a 500 ms exposure time. Gemini-561 in red (ex: 550 nm, em: 595±40 nm) and eGFP in green (ex: 470 nm, em: 531±40 nm). Results were found similar in n = 3 independent experiments. Scale bar is 30  $\mu$ m.



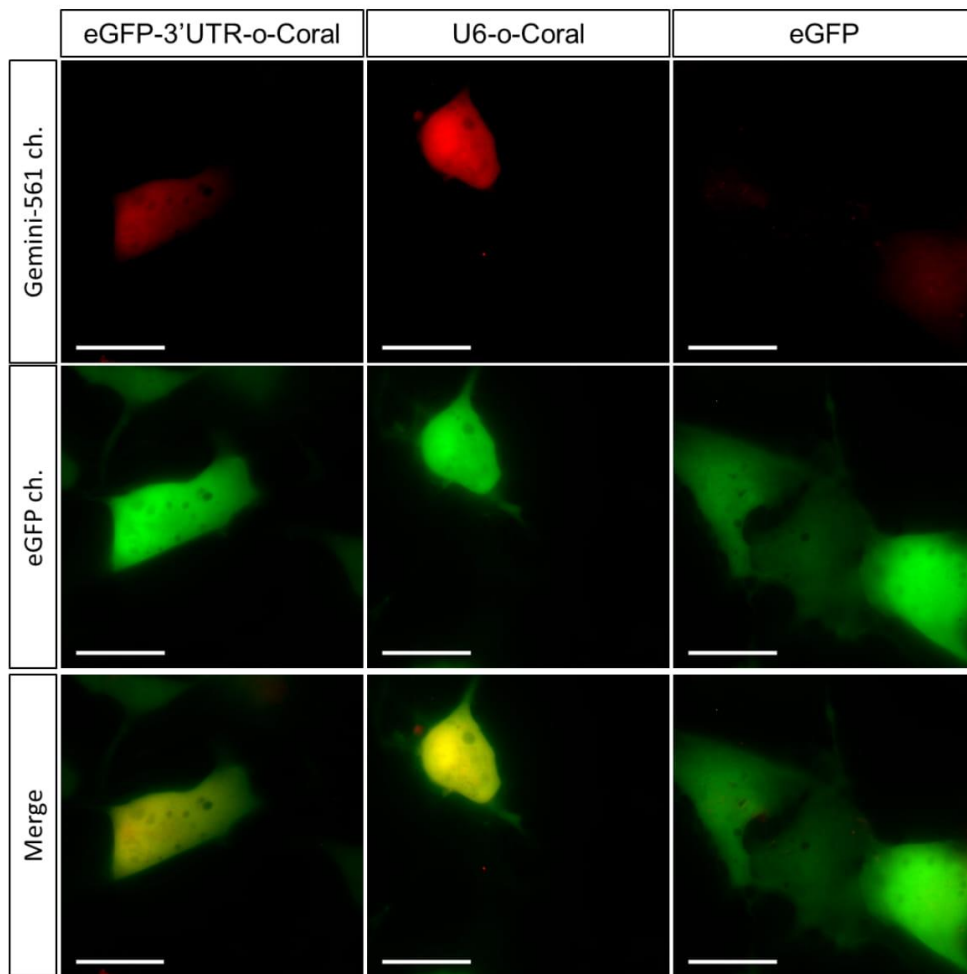


**Supplementary Figure 20** | Effect of temperature on Gemini-561 internalization. Fluorescence imaging of HeLa (a), HEK293T (b) and U87 (c) cells expressing the U6-o-Coral, incubated for 5 min with Gemini-561 (200 nM) at 4 °C or 37 °C. (a-c) The images were acquired using a 500 ms exposure time. Gemini-561 in red (ex: 550 nm, em: 595±40 nm) and eGFP in green (ex: 470 nm, em: 531±40 nm). Results were found similar in n = 3 independent experiments. Scale bar is 30µm.

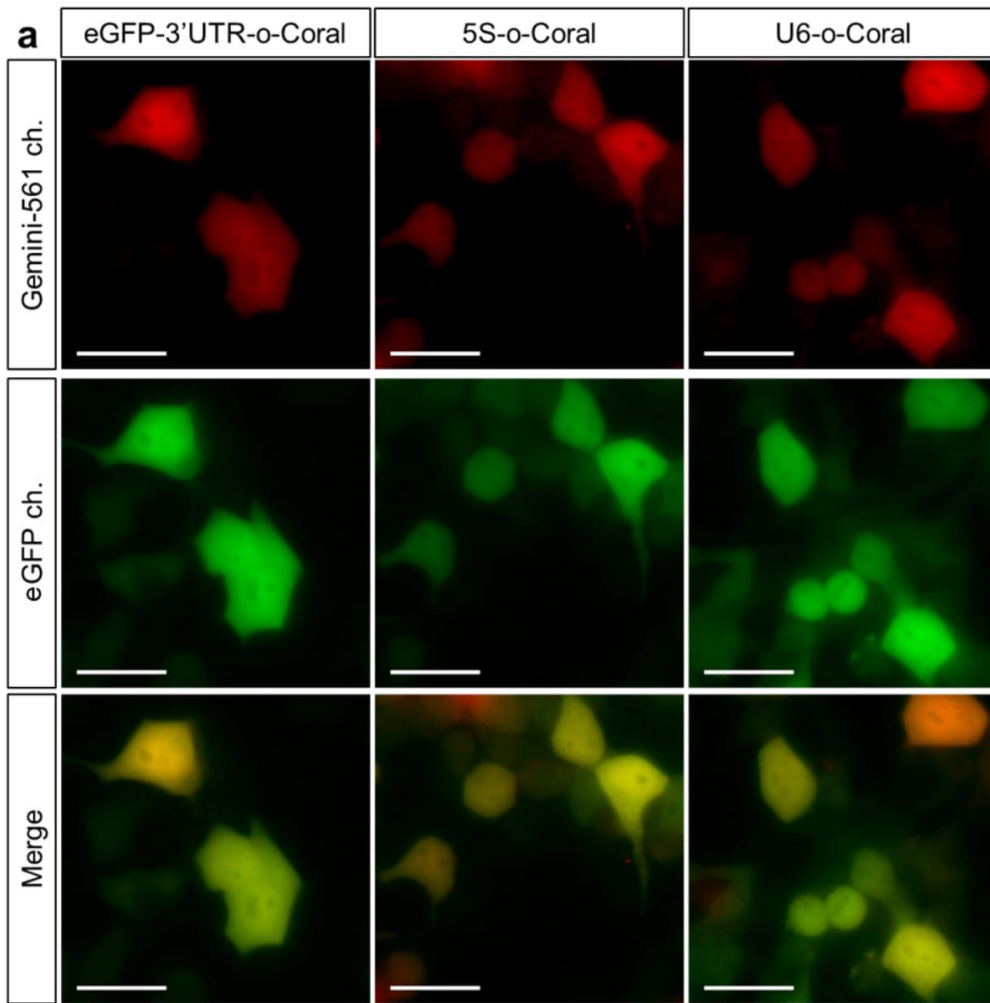


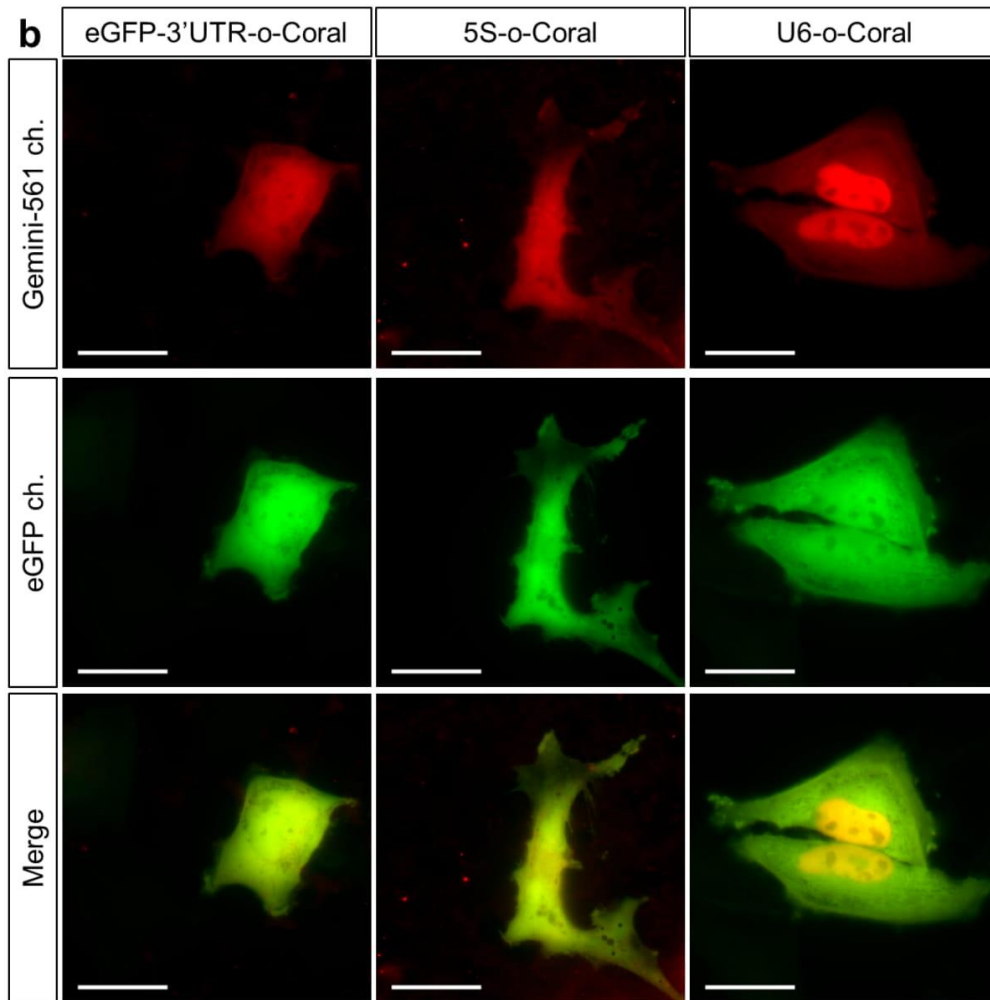


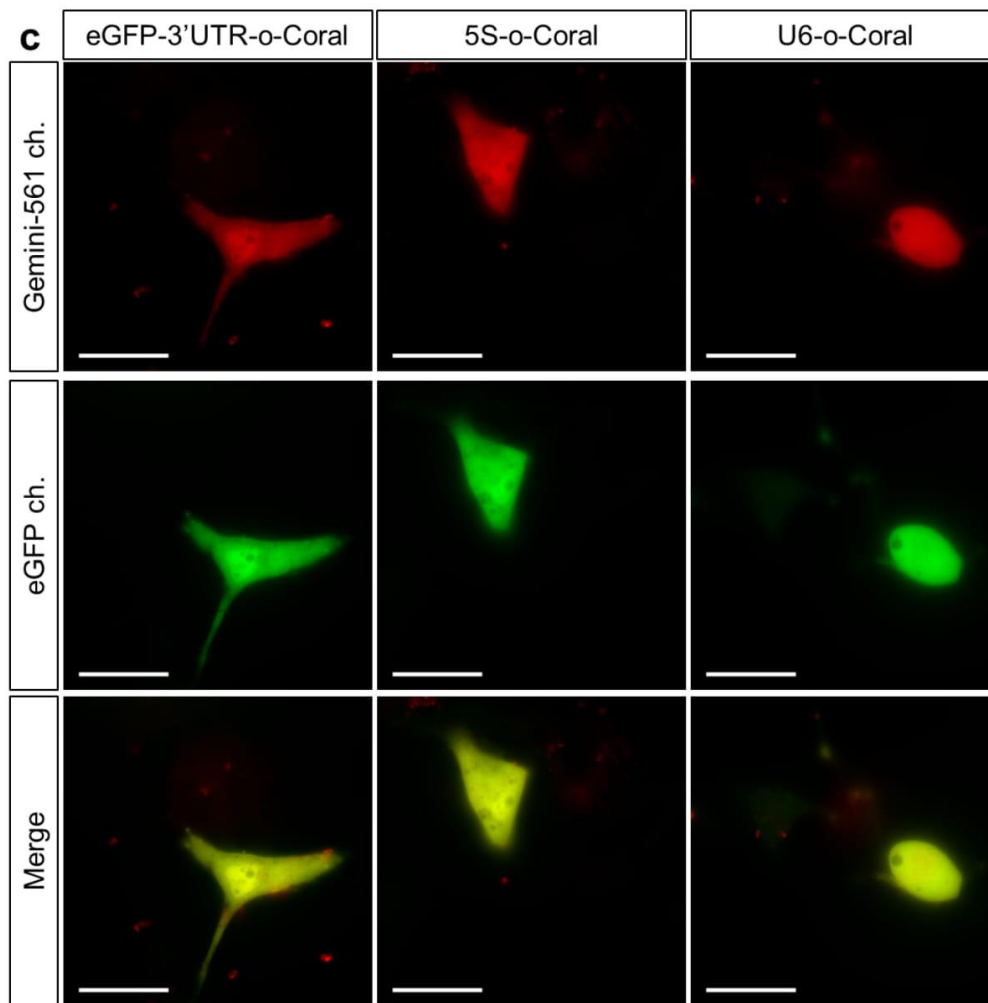
**Supplementary Figure 21 | Effect of o-Coral on egfp mRNA level in the cells.** Cells were transfected with a plasmid carrying a eGFP-coding gene containing (egfp-o-Coral) or not (egfp) o-Coral in the 3' untranslated region. Upon growth, cells were harvested and their content in gfp or gfp-o-Coral mRNA was quantified and normalized to that of two internal references: gapdh or s18 mRNA. The values are the mean of n = 3 independent measurements performed on three different batches of cells and the error bars correspond to  $\pm$  one standard deviation.



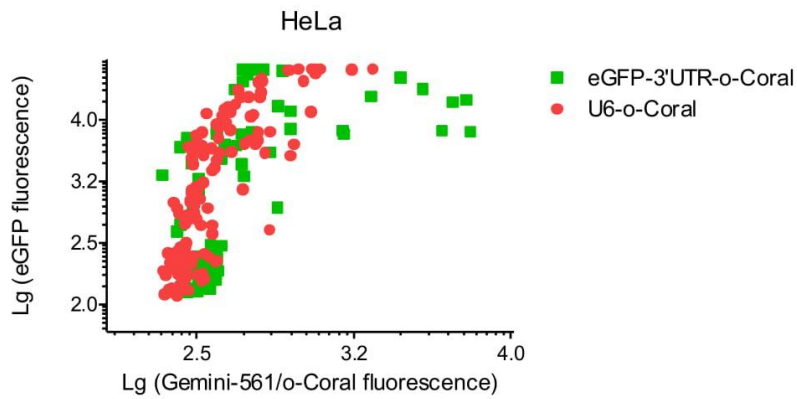
**Supplementary Figure 22** | Fluorescence imaging of U87MG cells expressing the egfp mRNA labelled with a single copy of o-Coral in the 3'-untranslated region (3' UTR-o-Coral), o-Coral from the U6-promoter or eGFP only. Results were found similar in n = 3 independent experiments. Scale bar is 30 $\mu$ m.







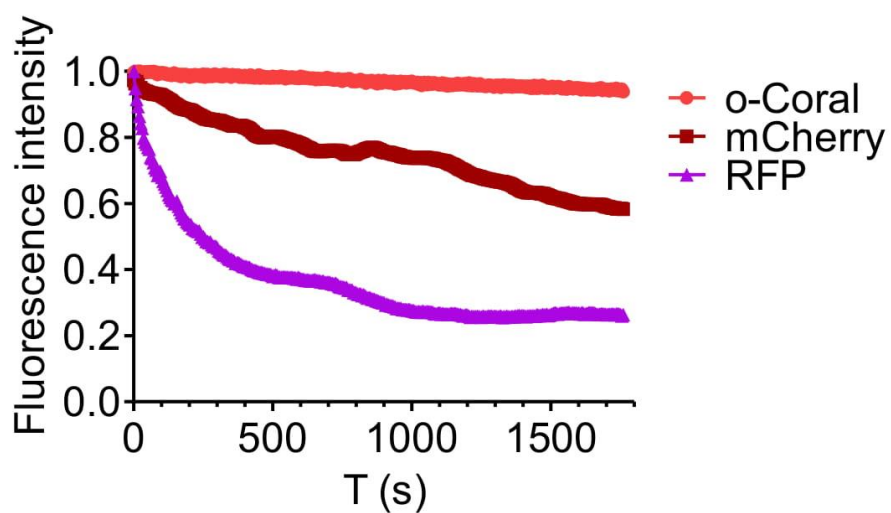
**Supplementary Figure 23** | Live cell imaging of three different RNA constructs in three cell lines. HEK293T (a) HeLa (b) and U87MG (c) cells expressing the eGFP-3'UTR-o-Coral, 5S-o-Coral or U6-o-Coral, incubated for 5 min with Gemini-561 (200 nM). The images were acquired using a 500 ms exposure time. Gemini-561 in red (ex: 550 nm, em: 595±40 nm) and eGFP in green (ex: 470 nm, em: 531±40 nm). Results were found similar in n = 3 independent experiments. Scale bar is 30µm.

**a****b**

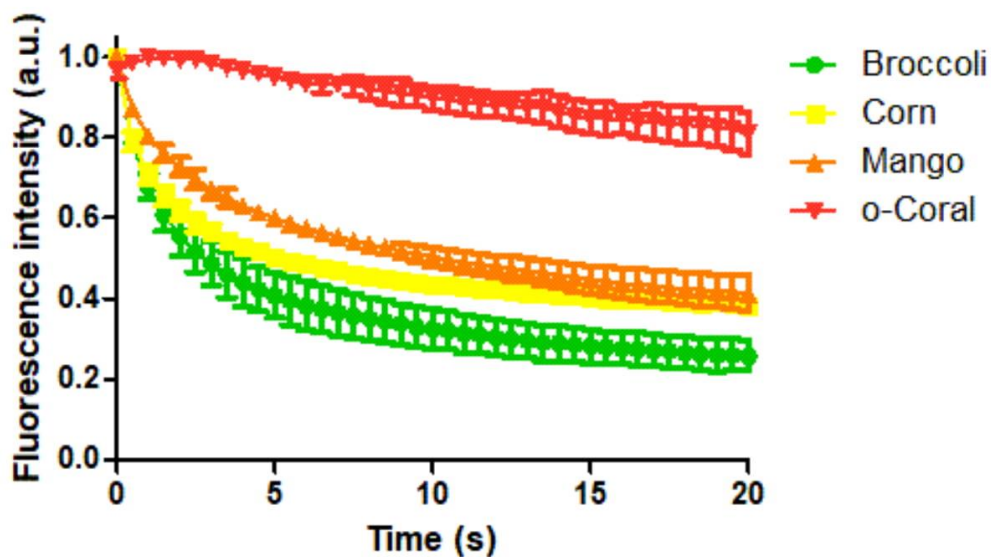
Parameter	o-Coral	egfp_3'UTR_o-Coral
Number of XY Pairs	131	82
Pearson r	0.7996	0.5781
95% confidence interval	0.7279 to 0.8540	0.4129 to 0.7065
P value (two-tailed)	< 0.0001	< 0.0001
P value summary	****	****
Is the correlation significant? (alpha=0.05)	Yes	Yes
R square	0.6394	0.3342

**Supplementary Figure 24 |** (a) Correlation between fluorescence signal from eGFP with that from Gemini-561 in a single cell analysis of U6-o-Coral and egfp-3'UTR-o-Coral expressing HeLa cells. Each point represents the fluorescence signals from a single cell. (b) The table shows the results of the correlation analysis. Inbuilt two-tailed XY correlation analysis in GraphPad Prism was used with confidence interval 95% and data assumed to be sampled from Gaussian population (Pearson). The number of cells for o-Coral and eGFP\_3'UTR\_o-Coral were 131 and 82 respectively.

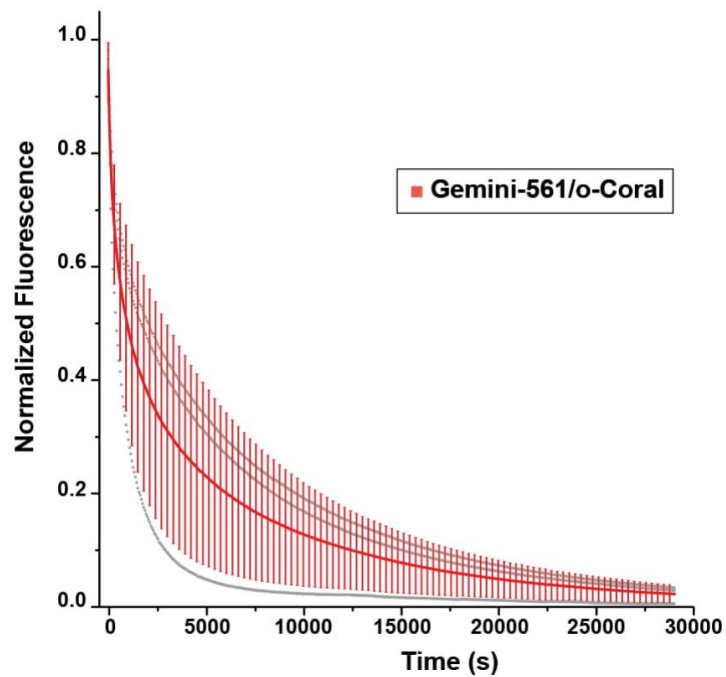




**Supplementary Figure 25** | Comparative analysis of photostability by spectroscopy. Photostability of Gemini-561/o-Coral ( $0.2 \mu\text{M}/1 \mu\text{M}$ ) compared to mCherry ( $0.2 \mu\text{M}$ ) and RFP ( $0.2 \mu\text{M}$ ). Each system was excited at the same molar extinction coefficient value:  $30,000 \text{ M}^{-1} \text{ cm}^{-1}$ . o-Coral, mCherry and RFP were excited using 532 nm laser ( $7 \text{ mW cm}^{-2}$ ,  $18 \text{ mW cm}^{-2}$ ,  $11 \text{ mW cm}^{-2}$  respectively). Fluorescence intensity was monitored at 596 nm for o-Coral, 610 nm for mCherry and 584 nm for RFP. Experiment was repeated 2 times. Representative graph from 1 experiment is displayed.



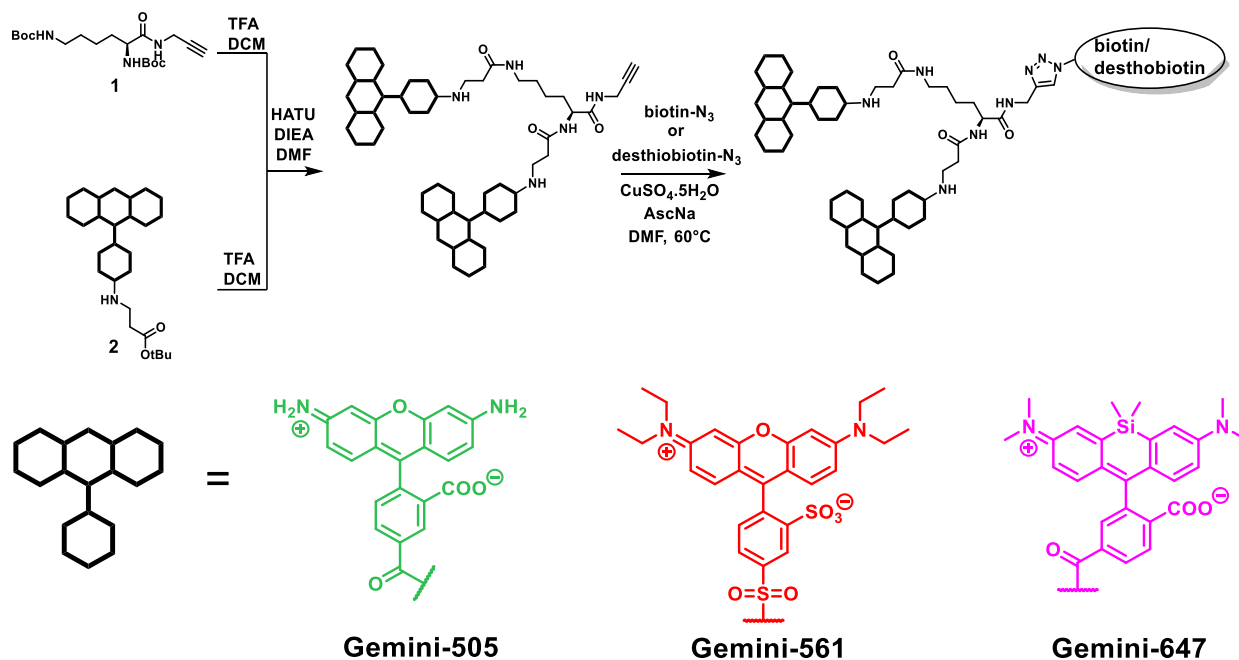
**Supplementary Figure 26 | Photostability of aptamer-dye couples in live HeLa cells.** Broccoli/DFHBI-1T, Corn/DFHO, Mango/TO1-biotin and o-Coral/Gemini-561 photostability was assessed by fluorescence microscopy. Cells were preincubated with corresponding dye (10  $\mu$ M DFHBI-1T, 10  $\mu$ M DFHO, 0.2  $\mu$ M TO1-biotin for 30 min and 0.2  $\mu$ M Gemini-561 for 5 min). Aptamers were microinjected in live HeLa cells at 20  $\mu$ M concentration. Microinjection parameters:  $P_i=90$  [hPa];  $T_i=0.3$  [s];  $P_c=10$  [hPa]. Consecutive images were acquired, each using a 500-ms exposure time. Broccoli (ex: 470 nm, em:  $475\pm 50$  nm); Corn (ex: 470 nm, em:  $531\pm 40$  nm); Mango (ex: 470 nm, em:  $531\pm 40$  nm); Coral (ex: 550 nm, em:  $595\pm 40$  nm). The excitation power was adjusted to reach similar emission intensity. Fluorescence intensity values were extracted using same region of interest from 3 independent injections. The values are mean  $\pm$  S.D (n=3).



**Supplementary Figure 27 | Photostability of Gemini-561/o-Coral over extensive constant illumination.** A mixture of Gemini-561/o-Coral ( $1 \mu\text{M}/2 \mu\text{M}$ ) was prepared and individualized into water-in-oil droplets to prevent unwanted exchange of complexes between illuminated and non-illuminated areas as described before in Fernandez-Millan, P., Autour, A., Ennifar, E., Westhof, E. & Ryckelynck, M. Crystal structure and fluorescence properties of the iSpinach aptamer in complex with DFHBI. *RNA* **23**, 1788-1795 (2017). The emulsion was then exposed to a constant illumination wavelength (575 nm) at the maximum intensity of the light source (Spectra X, Lumencor), and the emitted fluorescence ( $625 \pm 50 \text{ nm}$ ) was collected by an Orca-Flash IV camera for 500 ms every 100 ms with  $\times 40$  objective (numerical aperture (NA) 0.45). Individual measurements are shown in gray, whereas the mean values of the 3 independent experiments are shown in red. The error bars correspond to  $\pm 1$  standard deviation.

## 2.4. Gemini-505 and Gemini-647 fluorogens for SELEX-based RNA directed evolution (unpublished)

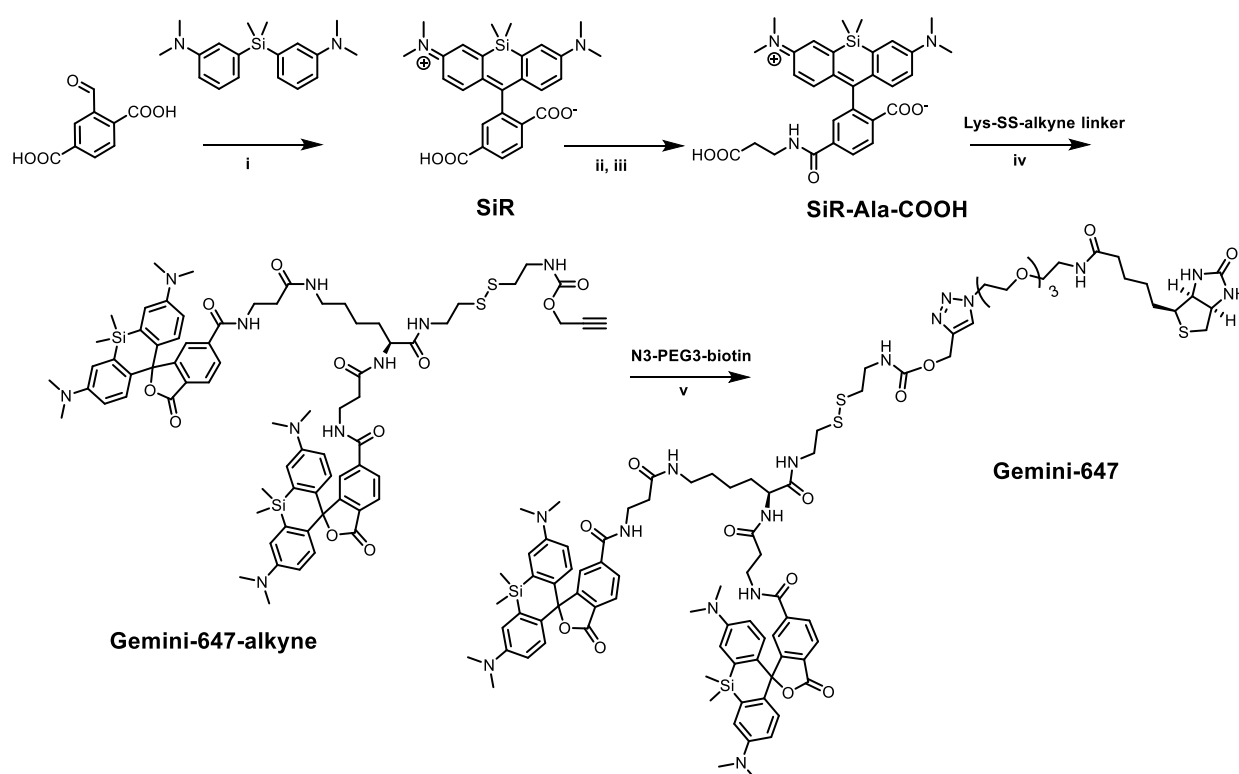
In this project, we aimed to expand the toolbox of fluorogenic aptamers based on our dimerization approach to new orthogonal systems, in particular aptamers activating rhodamines with tunable emission color. To reach this goal, we designed and synthesized dimerized probes based on rhodamine 110 (R110, operating in the green region) and Si-rhodamine (SiR, operating in far-red region) (Figure 16).



**Figure 16.** General scheme of synthesis of Gemini fluorogens.

### 2.4.1. Design and synthesis Gemini-647

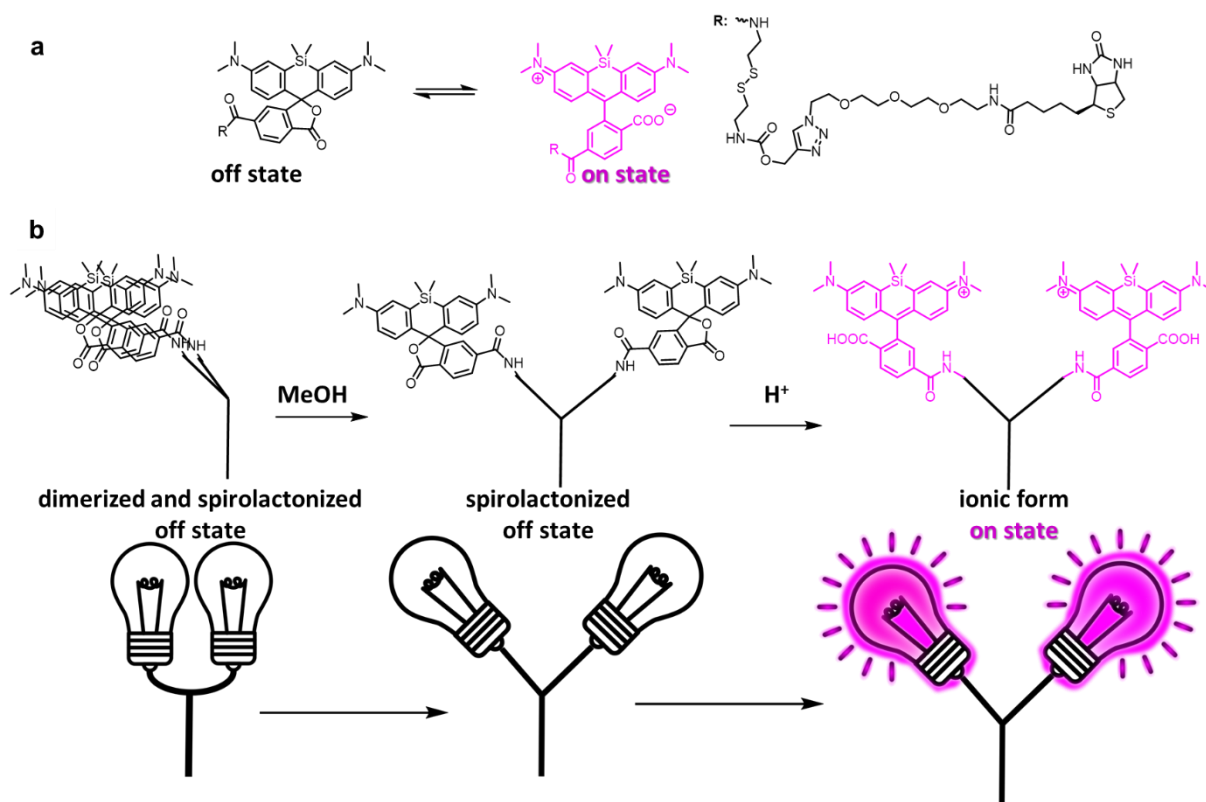
To synthesize Gemini-647, we first developed a new scalable protocol to synthesize SiR using CuBr<sub>2</sub> as catalyst and 2,2,2-trifluoroethanol (TFE) as solvent (Figure 17, step i and Scheme 1, Materials and methods). SiR was coupled to β-Ala tert-butyl ester. After hydrolysis, it was coupled to a Lys-SS-alkyne and clicked to N<sub>3</sub>-PEG3-biotin. Introduction of disulfide bridge aimed at simplifying elution process during SELEX selection by using reducing agent like dithiothreitol (DTT).



**Figure 17.** Scheme of Gemini-647 synthesis. (i)  $\text{CuBr}_2$ , TFE,  $80^\circ\text{C}$ , 24 h (7%); (ii)  $\beta$ -Ala-COOt-Bu, HATU, DIEA, DMF, r.t., 18 h (34%); (iii) TFA, DCM, r.t., 1 h; (iv) HATU, DIEA, DMF, r.t., 18 h (36%); (v) N<sub>3</sub>-PEG3-biotin,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , DMF, r.t. (44%).

#### 2.4.2. Gemini-647 exhibited double fluorogenic properties

Since its discovery in 2013<sup>93</sup>, SiR attracted a lot attention in cellular imaging because it possesses superior optical properties to rhodamines, notably high fluorescence quantum yields, high extinction coefficients and photostability.<sup>94</sup> Additionally, SiR is far-red emitting and a fluorogenic dye. Its fluorogenic property is based on equilibrium between non-fluorescent spirolactonized uncharged form and fluorescent zwitterionic form (Figure 18a).<sup>93</sup> Interestingly, when SiR is coupled to targeted ligand, the equilibrium is shifted towards the formation non-fluorescent form, but the fluorescence is switched on once the probe is bound to a target (protein, DNA).<sup>94</sup> Spirolactonization of SiR is highly dependent of environment thus difficult to control. We reasoned that dimerization of SiR can affect the property of the dye in a more modular way. To show the benefit of dimerization approach, we also synthesized the monomeric variant of SiR, SiR-SS-biotin (Figure 18a, section 5.2.4 Materials and methods), and performed side-by-side comparisons.

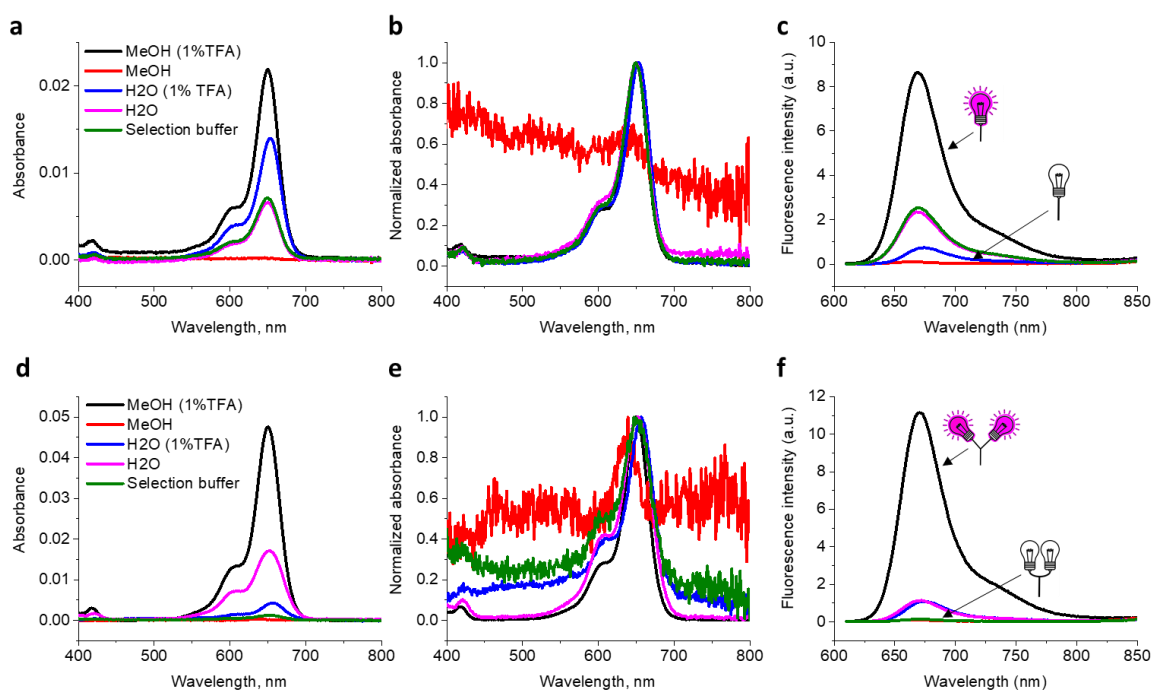


**Figure 18.** Double fluorogenicity of Gemini-647. (a) Equilibrium between OFF and ON states of SiR via intramolecular spirolactonization; (b) Proposed mechanism of double fluorogenic nature of Gemini-647. In organic solvent like MeOH dimer disassembles, but SiR molecule remains spirolactonized (OFF state) because of a lack of acidity. Upon addition of TFA the SiR equilibrium is shifted to ionic form (on state).

We presumed that close proximity of lipophilic SiR in the dimerized molecule will provoke self-quenching by both intramolecular aggregation and formation of neutral spirolactone (Figure 18b). To confirm our hypothesis, we performed a series of photophysical experiments. Firstly, we checked the fluorogenic response Gemini-647 and SiR-SS-biotin from water to MeOH (Figure 19c, f). To our surprise, both molecules were completely quenched in pure MeOH, but fluorescent at certain degree in pure water. This is attributed to the fact that water is more acidic than MeOH,<sup>95</sup> therefore in water protonation of carboxylate of SiR pushed equilibrium towards formation of open fluorescent form. To promote formation of fluorescent (ionic) form, we added 1% trifluoroacetic acid (TFA) to the solutions of the fluorogen in corresponding solvent and observed formation of fully fluorescent form in MeOH (1% TFA) and negligibly fluorescent form in water (1% TFA). Interestingly, Gemini-647 exhibited lower fluorescence in pure water, water (1% TFA) and selection buffer in comparison to



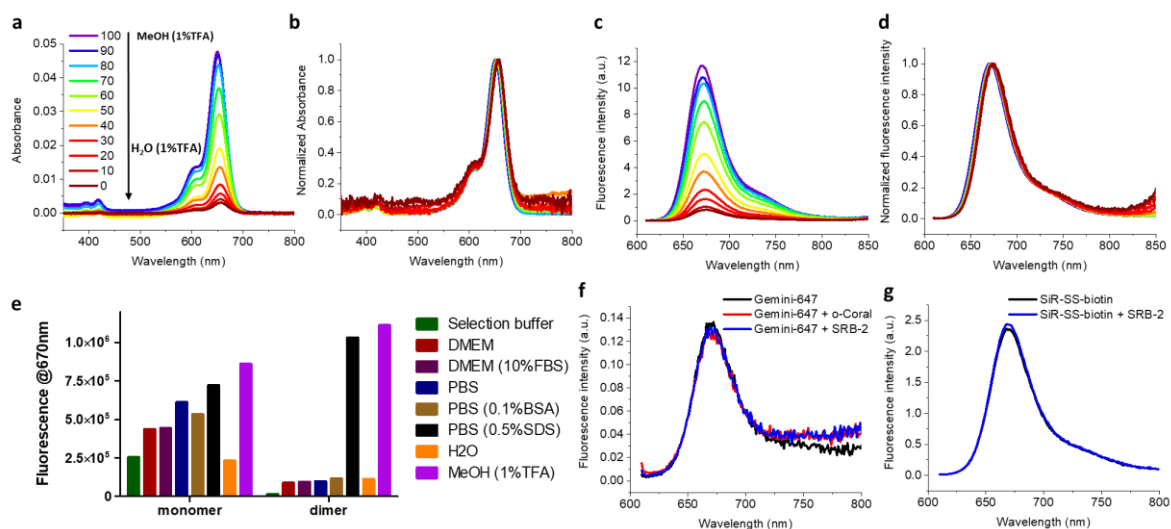
monomeric SiR due to dimerization-caused self-quenching (Figure 19c, f). We also noticed appearance of small blue-shifted shoulder in absorbance spectra of Gemini-647 (Figure 19e) suggesting formation of the unimolecular intramolecular aggregate. Unlike Gemini-561 (chapter 2.3.1), Gemini-647 did not form profound shoulder of *H*-aggregate because upon dimerization the polarity SiR microenvironment changed causing spirolactonization and thus less absorbing, non-emissive form.



**Figure 19.** Photophysical properties of SiR-SS-biotin (top panel) and Gemini-647 (bottom panel) in various conditions. Absorbance (a), normalized absorbance (b) and fluorescence spectra (c) of SiR-SS-biotin. Absorbance (d), normalized absorbance (e) and fluorescence spectra (f) of Gemini-647.

We performed a detailed analysis to study double fluorogenic process in the mixture of MeOH-H<sub>2</sub>O (Figure 20a-d). To control acidity of the system, we supplemented solvents mixture with 1% TFA. Upon increase of the water content in the mixture, absorbance and fluorescence intensity gradually decreased (Figure 20a, c). The fluorescence increase from H<sub>2</sub>O (1% TFA) to MeOH (1% TFA) was moderate 10.2-fold, but from selection buffer MeOH (1% TFA) it was 77-fold. Additionally, we studied Gemini-647 optical properties in modular bioenvironment (with FBS, BSA). Results showed that monomeric SiR tended to respond to biomolecules with switching ON its fluorescence, while Gemini-647 (dimer) retained its quenched form (Figure 20e). We also confirmed positive response of dyes to detergent sodium dodecyl sulfate (SDS)

as reported in the literature<sup>93</sup> (Figure 20e). We also tested binding capacity of Gemini-647 and SiR-SS-biotin to o-Coral and SRB-2 aptamers. Both Gemini-647 and SiR-SS-biotin did not produce fluorescence once mixed with o-Coral or SRB-2 aptamers (Figure 20f, g). Taking together, Gemini-647 was the promising candidate for SELEX-based selection of functional RNA aptamer.

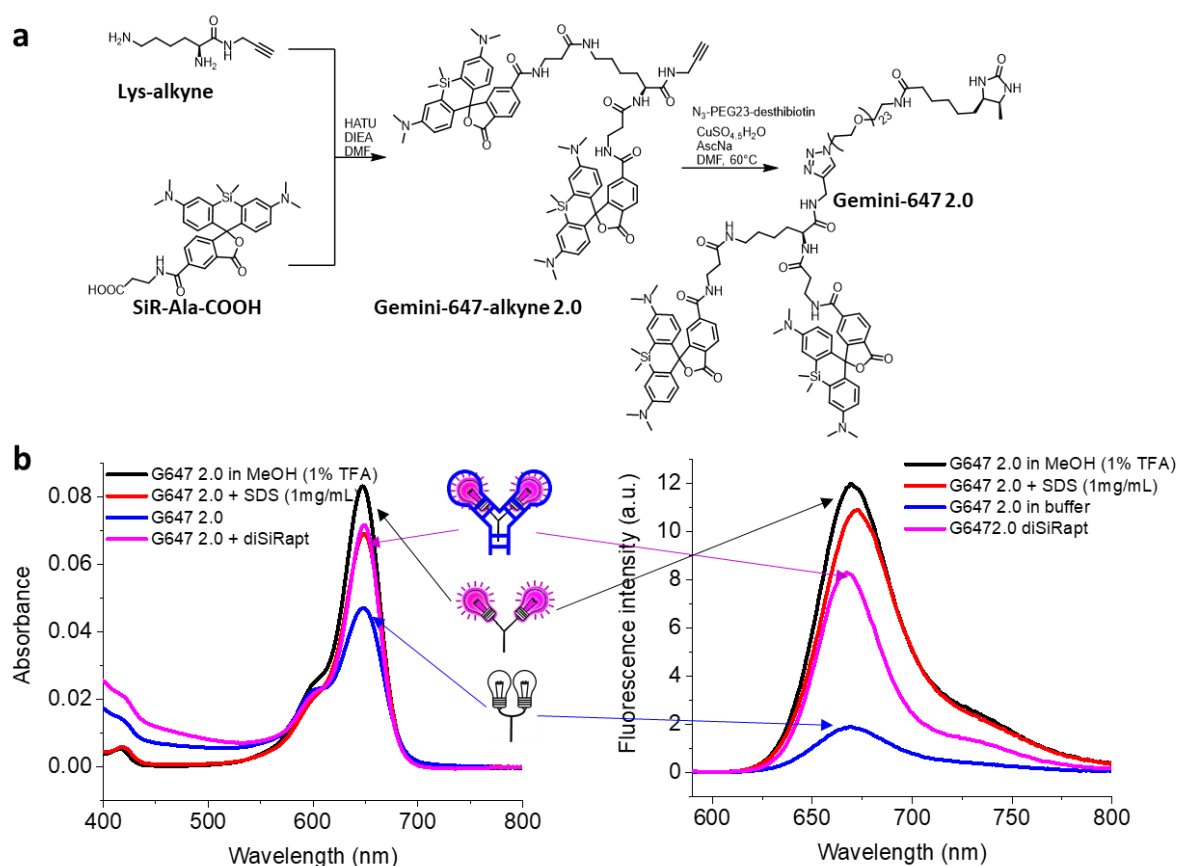


**Figure 20.** Spectral properties of SiR-based fluorogenic probes. (a) Absorbance, normalized absorbance (b), emission (c) and normalized emission (d) spectra of Gemini-647 (200 nM) in the gradient mixtures of MeOH-water (1% TFA) (100 to 0); (e) Comparison of fluorescence of SiR-SS-biotin (monomer, 200 nM) and Gemini-647 (dimer, 200 nM) in buffers with or without proteins; (f) Emission spectra of Gemini-647 (200 nM) in the absence or presence of RNA aptamers (o-Coral, SRB-2, 600 nM); (g) Emission spectra of SiR-SS-biotin (200 nM) in the absence or presence of SRB-2 aptamer (600 nM).

#### 2.4.3. Generation of Gemini-647 2.0, spectral properties and interaction with diSiRapt.

During n-rounds of RNA selection, Farah BOUHEDA (RYCKELYNCK lab, IBMC, University of Strasbourg) found that Gemini-647 produced a non-specific fluorescence once it is compartmentalized in oil droplets during microfluidic-assisted SELEX. We assumed that this signal is caused by probe aggregation-disaggregation equilibrium in the oil droplet environment. Therefore, we synthesized Gemini-647 2.0 by introducing PEG23 linker between dye-dimer and ligand (desthiobiotin) to promote retention of unimolecular water-soluble form. Additionally, biotin was replaced by desthiobiotin as a bead-loading ligand and the reducible S-S bridge was removed (Figure 21a). We used desthiobiotin instead of biotin since it has a lower affinity to streptavidin, therefore it

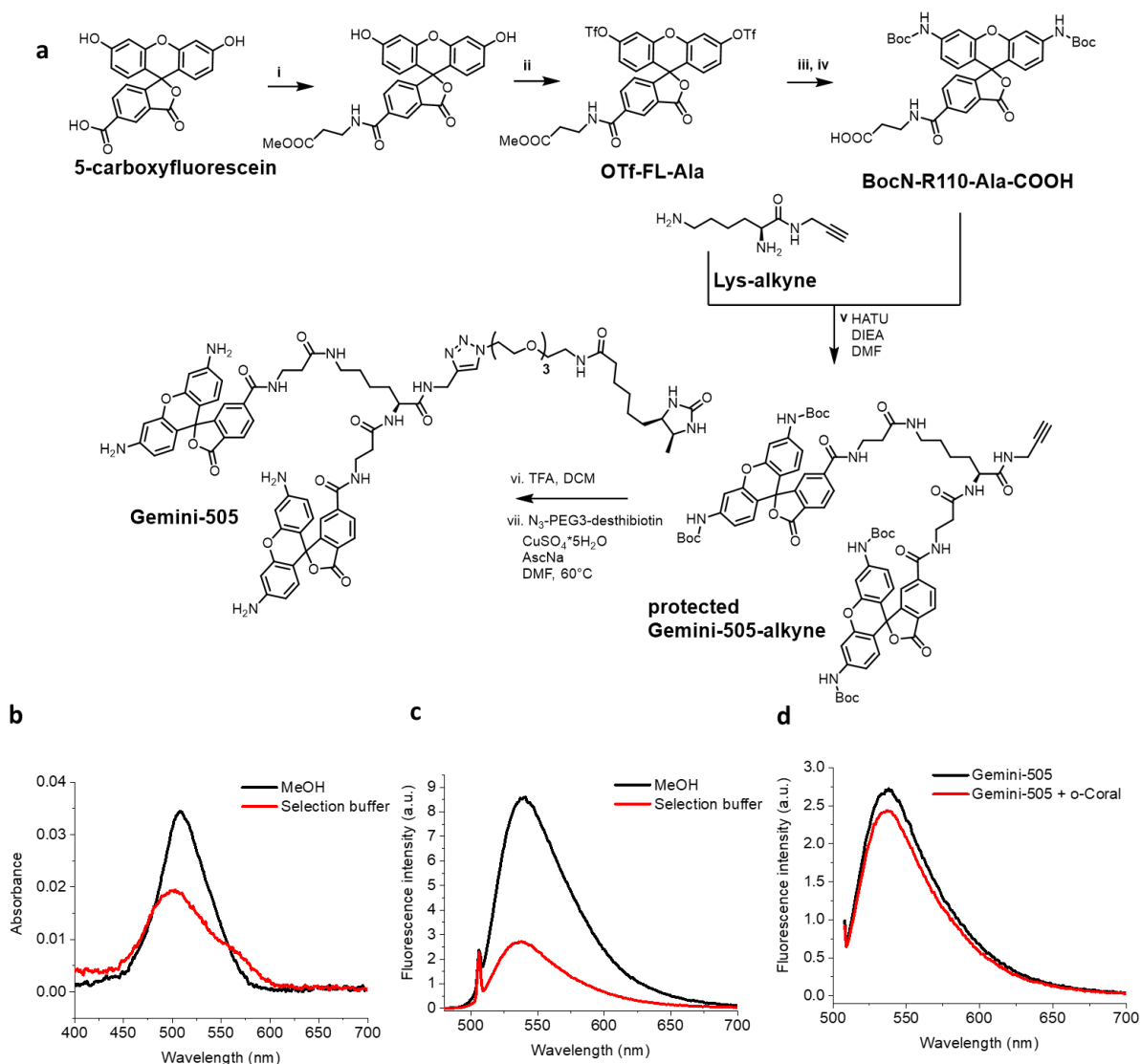
would ease the elution process during RNA SELEX. Gemini-647 2.0 showed fluorescence enhancement of ~8.5 times from selection buffer to MeOH (1% TFA) and could reveal its fluorescence in buffer supplemented with SDS (1 mg/mL) (Figure 21b). Farah BOUHEDDA and Michael RYCKELYNCK (IBMC, University of Strasbourg) designed and generated prototype of red-Coral by dimerization SiRA<sup>83</sup> aptamer, named diSiRapt. *In vitro* transcribed diSiRapt aptamer was capable to turn on fluorescence of Gemini-647 2.0 reaching ~80% of max fluorescence in MeOH (1%TFA) (Figure 21b). Judging by absorbance spectra from figure 21b, diSiRapt was not fully bound to Gemini-647 2.0. Additionally, isolated and purified diSiRapt did not respond to Gemini647 2.0 suggesting that further directed evolution of diSiRapt might improve properties of the system. Microfluidic-assisted SELEX-based directed evolution of diSiRapt is currently in progress in RYCKELYNCK lab. Additionally, a monomeric probe SiR-SS-biotin was used as a control compound for RNA library enrichment in early rounds of SELEX as well as in negative selections in later rounds.



**Figure 21.** Gemini-647 2.0. (a) Scheme of synthesis of Gemini-647 2.0; (b) Spectra of Gemini-647 2.0 (0.2  $\mu$ M) in MeOH (1% TFA), in the presence of SDS (1 mg/mL), in buffer and in the presence of *in vitro* transcribed diSiRapt.

#### 2.4.4. Design and synthesis Gemini-505 and its spectral properties.

We aimed to develop bright photostable green-emitting fluorogen compatible with standard 488 nm laser. To synthesize Gemini-505, firstly, we synthesized R110 precursor with  $\beta$ -Ala tail to include a degree of flexibility in a final dimeric probe. 5-Carboxyfluorescein was coupled to  $\beta$ -Ala methyl ester, then triflated following reported protocol<sup>96</sup> (Figure 22a). Triflated fluorescein ester derivative (OTf-FL-Ala) was used in Buchwald-Hartwig cross-coupling reaction with Boc-NH<sub>2</sub> to obtain protected R110 derivative (BocN-R110-Ala, figure 22a). BocN-R110-Ala was hydrolyzed and coupled to Lys-alkyne. The coupled compound was further clicked to N<sub>3</sub>-PEG3-desthiobiotin to obtain Gemini-505 (Figure 22a). As previously mentioned, replacement of biotin with desthiobiotin was aimed to ease the elution from deposit platform during SELEX process. Surprisingly, Gemini-505 did not show formation of blue-shifted *H*-aggregate shoulder in absorbance spectra (Figure 22b) and showed relatively modest fluorescence enhancement of 3.2-fold from selection buffer to MeOH (Figure 22c), presumably due to spirolactonization of rhodamine upon dimerization like in case of Gemini-647. Importantly, Gemini-505 did not produce positive response when o-Coral aptamer was added (Figure 22d). These data indicated the selectivity of o-Coral to bind Gemini-561 only and made Gemini-505 as a promising candidate to select green Coral (g-Coral) aptamer. A control compound (R110-desthiobiotin, monomeric probe) of Gemini-505 was also synthesized (Material and methods, section 5.2.4).



**Figure 22.** Gemini-505. **(a)** Scheme of synthesis of Gemini-505: (i)  $\beta$ -Ala-OMe-HCl, HATU, DIEA, DMF, r.t., 18h (43%); (ii) Py,  $\text{TiF}_2\text{O}$ , DCM,  $0^\circ\text{C}$  to r.t., 4 h (28%); (iii)  $\text{BocNH}_2$ ,  $\text{Pd}_2\text{dba}_3$ , Xantphos,  $\text{Cs}_2\text{CO}_3$ , dioxane,  $100^\circ\text{C}$ , 18 h (30%); (iv) KOH (4 M MeOH), MeOH, r.t., 5.5 h; **(b)** Absorbance and **(c)** fluorescence spectra of Gemini-505 in MeOH and selection buffer; **(d)** Fluorescence spectra of Gemini-505 in the absence or presence of o-Coral aptamer.

In summary, within this project we developed Gemini-647 and Gemini-505 aiming to select their cognate RNA aptamers (r-Coral and g-Coral). This work has an important potential to expand the toolbox of fluorogenic aptamers with orthogonal systems spanning their fluorescence from green to far-red. Orthogonal Gemini/Coral pairs will enable to label and image several RNA targets simultaneously.

### 3. Self-labelling fluorogenic tags for protein imaging

#### 3.1. Overview on self-labeling fluorogenic tags for protein imaging

Imaging proteins in cells is one of the major technological advancement allowing to study their functionalities in native environment.<sup>97</sup> In order to visualize proteins by fluorescence, it is necessary to introduce a site-specific imaging reporter to a protein of interest (POI). The discovery of the genetically-encoded fluorescent proteins has revolutionized the field of protein imaging.<sup>97</sup> Additionally, other approaches using self-labeling protein tags<sup>98,99</sup> have drastically expanded a tool box of protein-imaging methods providing flexibility of synthetic fluorophore of choice. Typical self-labeling system requires two steps: (i) genetically-encoded fusion of the tag to POI and (ii) chemical reaction of fluorogenic ligand-tethered conjugate with tag.

In this chapter, a brief overview on recent advancements of protein labeling based on genetically-encoded tags capable of selective recognition of small molecules will be outlined.

##### 3.1.1. Tags based on covalent modification

Covalent labeling of proteins is based on expression of tag-fused POI followed by selective irreversible chemical reaction of the tag with its cognate fluorogenic substrate (Figure 23a).<sup>94</sup> Fluorescence response is generated upon the reaction. The widely used and commercially available is SNAP-tag developed in Johnsson lab pioneering the field of semisynthetic protein-based biosensors.<sup>100</sup> SNAP-tag is a small protein (20 kDa) evolved from the human DNA repair protein O<sup>6</sup>-alkylguanine-DNA alkyltransferase (AGT).<sup>100</sup> Functionalized O<sup>6</sup>-benzylguanine (BG) derivatives (SNAP-ligand, figure 23b) alkylate cysteine in the active site of SNAP-tag providing irreversible fusion.<sup>100</sup> The major advantage of SNAP-technology is that a wide range of BG-substrates can be used allowing variation and flexibility of fluorophore choice.<sup>100</sup> Gautier *et al.* evolved SNAP-tag and engineered CLIP-tag.<sup>101</sup> CLIP-tag selectively reacts with O<sup>2</sup>-benzylcytosine (BC) conjugates (CLIP-ligand, figure 23b) making this system orthogonal to SNAP-tag.<sup>101</sup>

At the same time, Los *et al.* developed Halo-tag<sup>102</sup> (33 kDa) derived from bacterial enzyme haloalkane dehalogenase from *Rhodococcus rhodochrous*. Chloride-derivatives (Halo-ligand, figure 23b) selectively alkylate dehalogene providing covalent



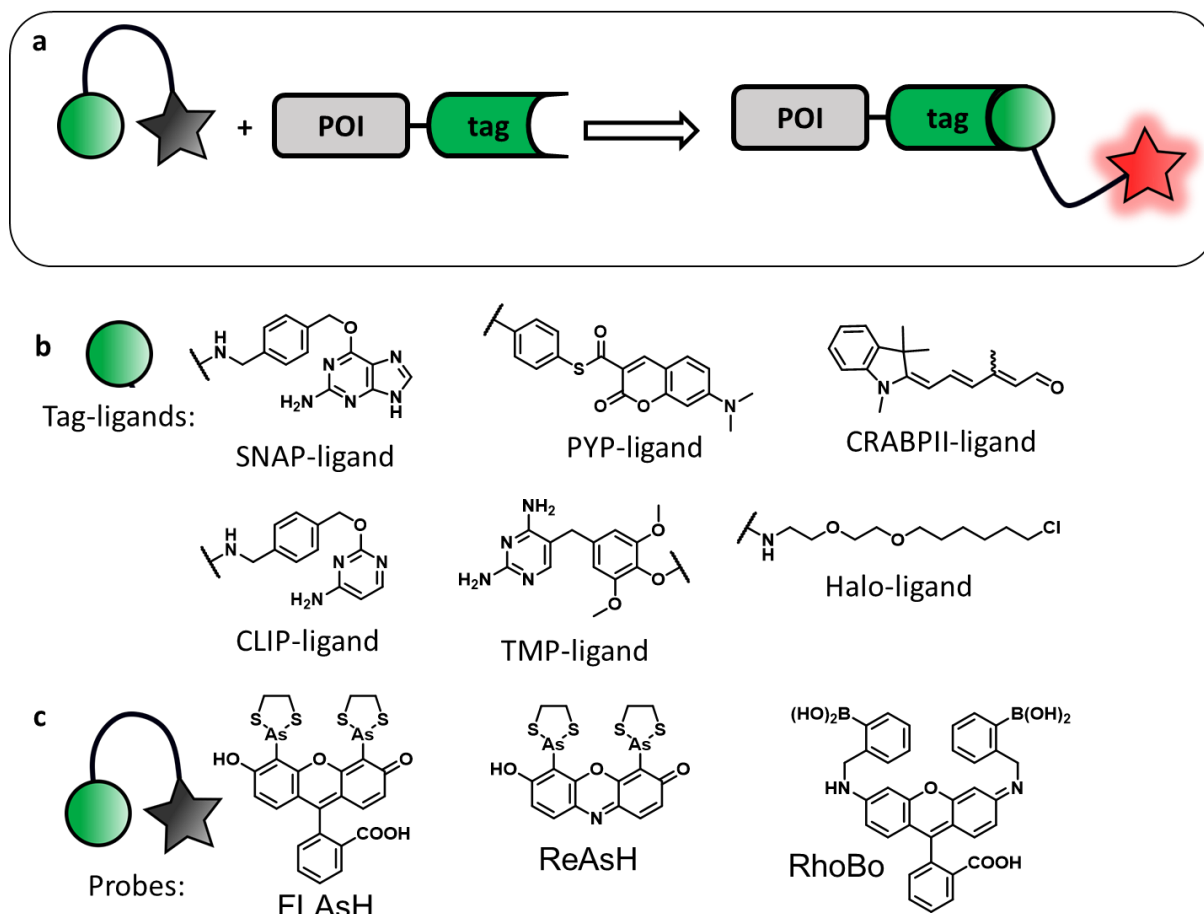
labeling.<sup>102</sup> Halo-tag found different applications ranging from protein isolation to imaging.

Hori *et al.* developed photoactive yellow protein (PYP<sup>103–105</sup>) tag derived from purple bacteria. PYP is a small protein (14 kDa) attaching covalently and selectively derivatives of 7-hydroxycoumarin-3-carboxylic acid and 4-hydroxycinnamic acid (PYP-ligand, figure 23b) *via* transthioesterification with Cys69 of the active site.<sup>103</sup> Slow reaction rate of first generation of PYP-probes pushed the development of 7-dimethylaminocoumarin with accelerated kinetics.<sup>104</sup> This allowed imaging of localization of methyl-CpG-binding domain which is known to bind DNA containing 5-methylcytosine and DNA methylation in live cells.<sup>105</sup>

Cornish group reported that 2,4-diamino-5-(3,4,5-trimethoxybenzyl) pyrimidine (trimethoprim or TMP, figure 23b) derivatives can be used to selectively label *E. coli* dihydrofolate reductase (eDHFR, ~ 18 kDa) fusion proteins in cells.<sup>106,107</sup>

Another small protein tag (15 kDa), cellular retinoic acid binding protein II (CRABP II), was recently developed by Borhan and colleagues.<sup>108</sup> Merocyanine aldehyde (CRABP II, figure 23b) selectively reacts with Lys from active site of CRABP II to produce fluorescence. Directed evolution of CRABP II by Arnold group resulted in development of Mero6 mutant that emit bright and NIR fluorescence upon covalent linkage to merocyanine aldehyde.<sup>109</sup>

It is worth mentioning pioneering works with the shortest peptide-tags of tetracysteine or tetraserine motifs. Tsien group was the first to introduce organobisarsenic acid thioesters that form covalent bonds with peptide-tags containing a tetracysteine core CCxxCC, most commonly CCGPCC.<sup>110</sup> FLAsH (Figure 23c) increased its fluorescence approximately 50,000-fold after reaction with cysteine motifs.<sup>110</sup> Improvement of FLAsH probe resulted in development of ReAsH,<sup>111</sup> CHoXAsH,<sup>111</sup> CrAsH<sup>112</sup> and Cy3As<sup>113</sup>. Alternatively, Schepartz group developed rhodamine-derived bisboronic acid (RhoBo, figure 23c) selectively labeling tetraserine motif SSPGSS.<sup>114</sup> Introduction of tetracysteine-motif alter proper disulphide-bond formation in POI, thus affecting its functionality. Additionally, tetraserine motif is widely found in many proteins limiting its scope of application.<sup>98</sup> Nevertheless, these technologies dramatically boosted the development of other tags.



**Figure 23.** Self-labeling fluorogenic tags based on covalent modification. (a) Labeling proteins via genetic fusion of tag to POI followed by covalent modification through fluorophore-ligand conjugate; (b) Examples of tag ligands; (c) Examples of pioneer reactive probes for short peptide motifs.

### 3.1.2. Fluorogen-binding tags

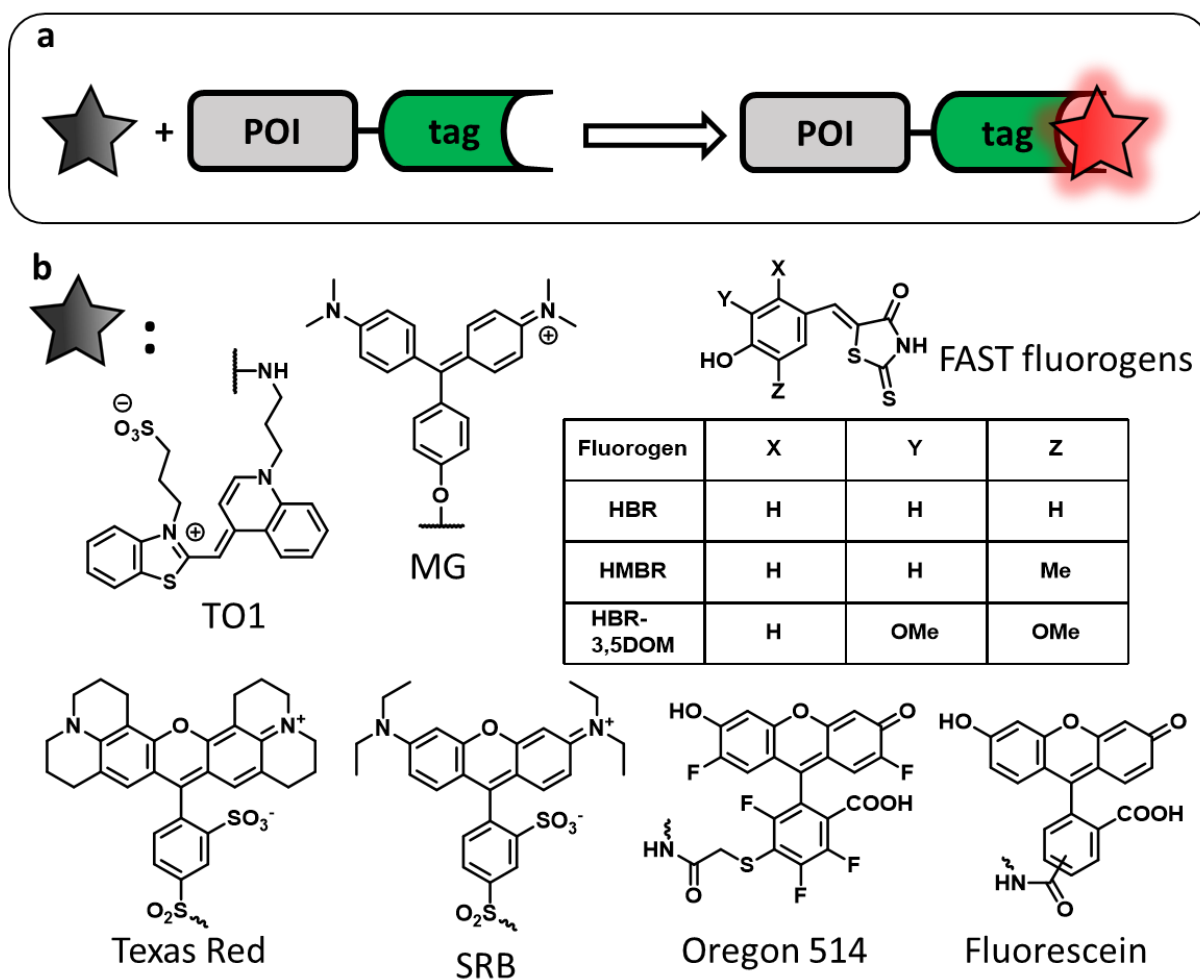
Protein tags have substantial size and, therefore, prone to perturb the functionality and mobility of POI.<sup>98</sup> To circumvent this obstacle, small peptide tags have been developed.<sup>98</sup> Another strategy for protein labeling is to introduce a tag enabling specific reversible binding its cognate ligand such as fluorogenic molecule (Figure 24a).

Orthogonally to Malachite Green aptamer<sup>60</sup> and Mango<sup>73</sup> mentioned in section 2.2.4 and 2.2.3, Szent-Gyorgyi *et al.* developed single-chain antibodies (scFvs), named fluorogen-activating proteins (FAPs), capable of binding to dyes with fluorescence enhancement response.<sup>115</sup> FAPs were selected to activate fluorescence of TO1 and MG (Figure 24b). Armitage group showed that FAPs are promiscuous and can be selected to bind fluorogenic cyanine dyes.<sup>116,117</sup> Despite known phototoxicity of MG dye, FAP-tag found a wide scope of applications since activated fluorescence emitted

in the NIR region.<sup>118–120</sup> Bruchez group tuned FAP properties for on-demand photoactivation of sensitizer enabling protein inactivation, targeted cell killing and rapid targeted lineage ablation in living larval and adult zebrafish.<sup>118</sup> Additionally, it was shown that FAPs were compatible for super-resolution imaging in live bacterial cells and promising for *in vivo* tumor targeting.<sup>119,120</sup>

In 2015, Gautier group introduced Yellow Fluorescence-Activating and absorption-Shifting Tag (Y-FAST) evolved from PYP-tag.<sup>121</sup> Unlike PYP-tag, Y-FAST could reversibly bind hydroxybenzylidene-rhodanine (HBR) or 4-hydroxy-3-methylbenzylidene-rhodanine (HMBR) fluorogens (FAST fluorogens, figure 24b).<sup>121</sup> This feature enabled to modulate fluorescence from ON to OFF state by simple addition or discharge of the fluorogen.<sup>121</sup> Later, the same group showed that FAST-tag is promiscuous and can recognize other fluorogens, 4-hydroxy-3,5-dimethoxybenzylidene rhodanine (HBR-3,5DOM).<sup>122</sup> Interestingly, once bound to HBR-3,5DOM fluorogen, FAST produced red-shifted emission, different from the yellow like emission in complex with HMBR fluorogen.<sup>122</sup> This allowed experimental versatility and on-demand color switch simply by replacing fluorogens, but using the same tag.<sup>122</sup> Further efforts from the group resulted in development of circularly permuted FAST (cpFAST) compatible for sensing  $\text{Ca}^{2+}$  fluxes in living cells,<sup>123</sup> dimerization of iFAST improved its spectral properties,<sup>124</sup> and FAST was compatible for super-resolution imaging.<sup>125</sup> Additionally, Tebo and Gautier very recently have reported a split version of FAST-tag (splitFAST).<sup>126</sup> SplitFAST allowed real-time dynamic observation of protein formation, assembly and dissociation in various cellular compartments (cytosol, nucleus, plasma membrane).<sup>126</sup>

Rozinov and Nolan were the first to introduce the concept of fluorophore-binding peptides termed fluorettes.<sup>127</sup> From a library of 12-mer peptides expressed on bacteriophages they selected and identified a panel of peptides binding sulforhodamine 101 (also known as Texas Red, TR), sulforhodamine B (SRB, Rhodamine Red), Oregon green 514 and fluorescein dyes (Figure 24b). Directed evolution of TR-binding peptide resulted in development of TR-512 mutant compatible for genetically encoded  $\text{Ca}^{2+}$  sensor.<sup>128</sup> Recently, Sunbul *et al.* converted TR-512 into a reactive tag (ReacTR-tag) by adding cysteine residue, thus enabling covalent linkage of Texas Red dye *via* N- $\alpha$ -chloroacetamide.<sup>129</sup>



**Figure 24.** Self-labeling fluorogenic tags based on reversible fluorogen-binding. (a) Fluorescence activation of fluorogen upon binding to tag-fused POI; (b) Examples of fluorogens for FAP-tag (MG, TO1), FAST-tag (HBR, HMBR, HBR-3,5DOM) and fluorettes (Texas Red, SRB, Oregon 514, fluorescein).

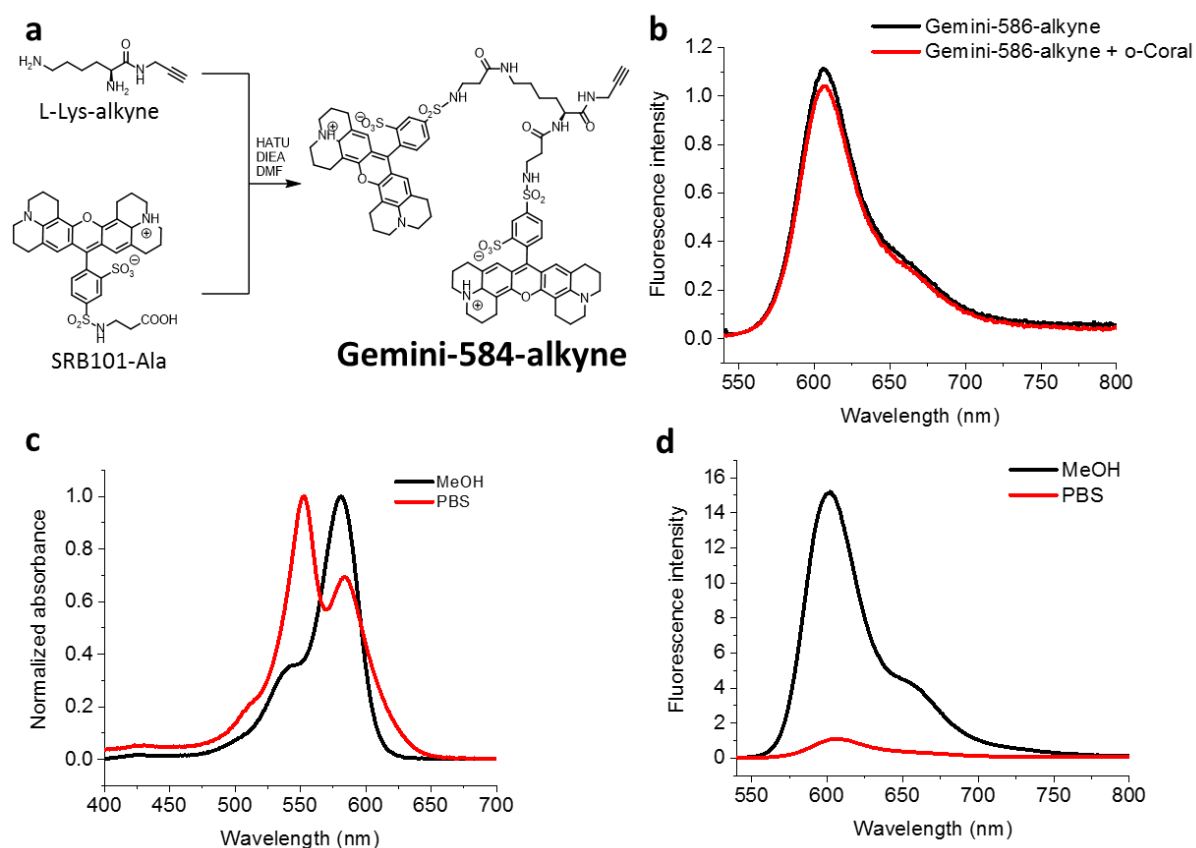
### 3.2. Gemini-584 as a fluorogen for dimeric fluorophore-binding peptide TR-512 (unpublished)

Inspired by the dimerization approach in the RNA part of the work, we aimed to develop a fluorogenic probe based on dimerized sulforhodamine 101 (SRB101) lighting-up its fluorescence upon binding to dimeric TR512-peptide.

#### 3.2.1. Design and synthesis of Gemini-584-alkyne and its spectral properties.

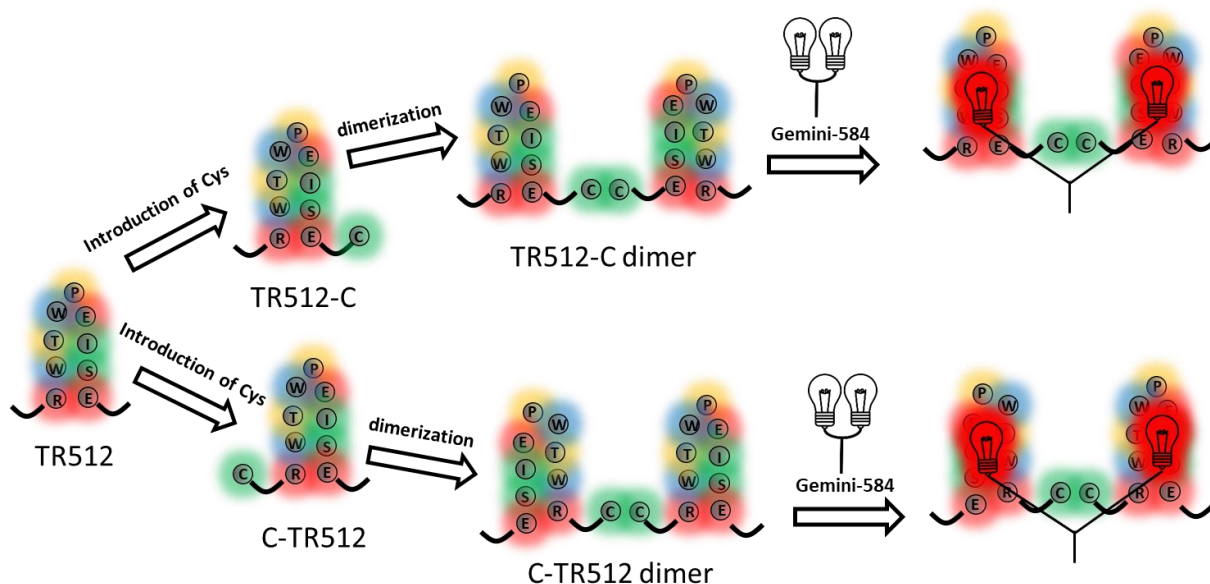
SRB101 was transformed to sulfochloride by oxalyl chloride, coupled to  $\beta$ -Ala tert-butyl ester (SRB101-Ala, section 5.2.2 Materials and methods). After hydrolysis, SRB101-Ala was coupled to L-Lys-alkyne to obtain Gemini-584-alkyne (Figure 25a).

Photophysical studies showed that Gemini-584-alkyne exhibited typical dimerization-caused quenching through formation of unimolecular *H*-aggregate in aqueous medium, evident by the blue-shifted band in the absorption spectrum (Figure 25c). It exhibited fluorescence enhancement of 13.6-fold from aqueous media to MeOH (Figure 22d), absorbance maximum at 584 nm in MeOH and 553/584 nm in PBS, emission maximum at 602 nm in MeOH and 605 nm in PBS. As we previously showed, biotin did not affect binding affinity of o-Coral to its partner fluorogen Gemini-561-alkyne<sup>92</sup>, we also verified if o-Coral can turn on fluorescence of Gemini-584-alkyne. As shown on figure 25 b, o-Coral was not able to recognize Gemini-584 proving its selectivity towards Gemini-561 and Gemini-561-alkyne.



**Figure 25.** Gemini-584-alkyne. (a) Synthesis of Gemini-584-alkyne; (b) Fluorescence spectra of Gemini-584-alkyne (200 nM) in the absence and presence of o-Coral aptamer (600 nM); absorption (c) and emission (d) spectra of Gemini-584-alkyne in MeOH and PBS.

3.2.2. *Photophysical studies of interaction of Gemini-584-alkyne with TR512 peptides.* Dimerization is natural process in protein biology leading to a range of new functionalities.<sup>37</sup> We also showed that during directed evolution of SRB-2 RNA aptamer, the sequence was dimerized and evolved resulting in identification of o-Coral RNA aptamer with improved properties.<sup>92</sup> Thus, we questioned if dimerization of TR-512 peptide recognizing SRB 101 dye could lead to a new peptide-tag with improved features (Figure 26). From commercial vendor (ProteoGenix, Schiltigheim, France), we purchased 3 variants of TR512-peptides (native TR512 peptide sequence<sup>128</sup>, TR512 with Cys at N- or C-terminus: C-TR512 and TR512-C). Introduction of Cys (C) was aimed to be a reactive moiety for dimerization *via* spontaneous naturally occurring oxidization process. The peptides were solubilized in mixture of DMSO-water and left shaking at 30°C until dimerization through cysteins occurred. The dimerization reaction was monitored by HPLC.

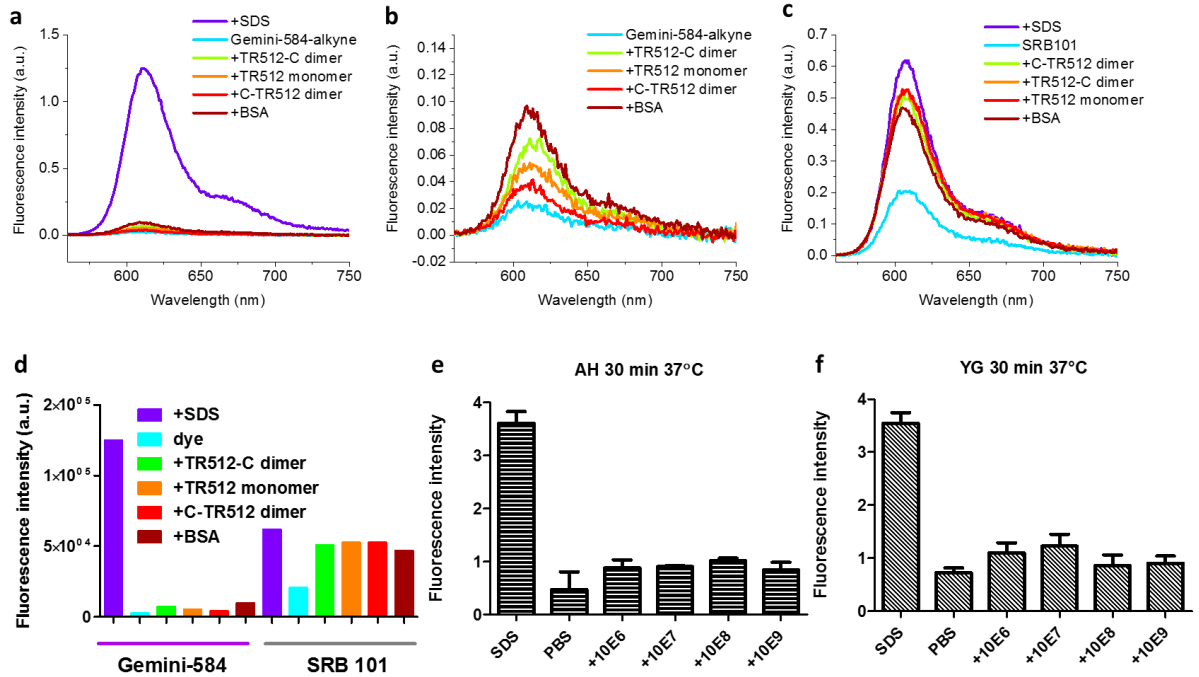


**Figure 26.** Chemical dimerization of TR512-peptides.

We further pursued binding studies of Gemini-584-alkyne with peptides and systematically compared with monomeric SRB101 dye. As a positive control (with maximum fluorescence intensity), emission of dyes in 0.5% SDS buffer solution was used (Figure 27a, c and d). As expected, Gemini-586-alkyne exhibited 48.6 times fluorescence turn-on from PBS to SDS-solution, while SRB 101 fluorescence increased only 3 times (Figure 27a, c). We also noticed 7 nm red shift of emission maximum from SDS solution to PBS for both dyes. After incubation of dyes (0.1  $\mu$ M)



with peptides (10  $\mu$ M C-TR512 dimer, TR512-C dimer or TR512 monomer) for 30 min at 37°C, fluorescence was recorded. We also used BSA protein as a control (non-specific signal). In accordance to the previously reported results<sup>128</sup>, we observed ~2.5 times fluorescence increase of SRB101 dye in complex with all 3 variants of TR-512 peptides, but fluorescence was also increased 2.3 times when SRB101 was mixed with BSA suggesting non-specific binding (Figure 27c). Unlike SRB101, Gemini-584-alkyne showed different values of fluorescence enhancement once mixed with peptides (2.8 times for TR512-C dimer, 2.1 for TR512 monomer and 1.6 for C-TR512 dimer), but also 3.7 times fluorescence increase for BSA protein (Figure 27b, d). Interestingly, Gemini-584-alkyne emission peak was red-shifted when bound to peptides (8 nm shift for TR512-C dimer, 15 nm for TR512 monomer and 10 nm for C-TR512 dimer), but only 6 nm peak red-shift for BSA protein. Emission peak of SRB101 bound to peptides or BSA was red-shifted only 3 nm for all cases. This data suggest that Gemini-584-alkyne was better confined in folded dimeric TR512-peptides (presumably due to better affinity) even though with very modest values of activated fluorescence. Additionally, we confirmed that Gemini-584-alkyne negligibly produced non-specific fluorescence when incubated with 2 yeast strains: (i) AH were *Saccharomyces cerevisiae* strain AH109; (ii) YG were YHGX13 strain. These 2 strains are intended to be used for further directed evolution of dimeric TR-512 peptides.



**Figure 27.** Binding studies of Gemini-584-alkyne with peptides assessed by fluorescence spectroscopy. **(a)** Emission spectra of Gemini-584-alkyne in the absence or presence of SDS, peptides/proteins; **(b)** Zoomed emission spectra of a; **(c)** Emission spectra of SRB101 in the absence or presence of SDS, peptides/proteins; **(d)** Comparison of intensity at the maximum of fluorescence of Gemini-584-alkyne and SRB101 in response to peptides/proteins; **(e)** Gemini-584-alkyne (0.1  $\mu\text{M}$ ) response to incubation with yeast (AH and YG). Fluorescence of the probe in SDS was used as a maximum of the signal (positive control) and fluorescence in PBS as a background signal (negative control). A concentration ranges from  $10^6$  to  $10^9$  of yeast cells per mL was tested.

Altogether, Gemini-584-alkyne is a promising candidate to develop dimeric TR-512-peptide-based tag since the maximum turn-on effect is  $\sim 50$  times (Figure 27d). We are currently collaborating with Dr. Eleonore REAL (UMR 7021, University of Strasbourg) for directed evolution of TR-512 peptides with Gemini-584 to generate genetically-encodable peptide-tag for live-cell imaging.

#### 4. Fluorescence imaging of biotin receptors in cancer cells.

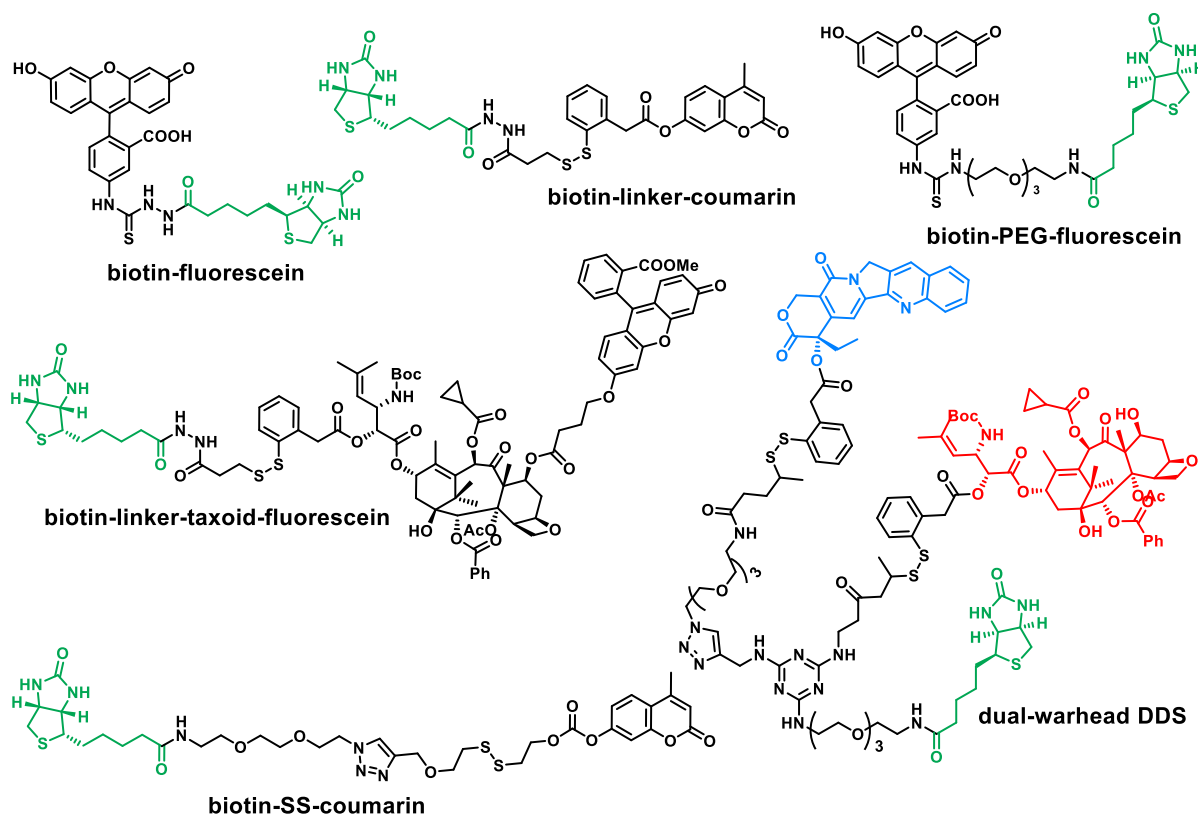
Biotin is an important food supplement involved in cellular carbohydrate, amino acid and lipid metabolism.<sup>130</sup> There is an evidence that cancerous cells overexpress biotin receptors (BRs), thus tend to consume more biotin in comparison to normal cells.<sup>131</sup>

Therefore, BRs are potentially useful cancer biomarkers for tumor diagnostics. While biotin receptors have been attractive target as the drug delivery system to tumour, there is lack of robust imaging probes for clinical diagnostic of BRs in cancerous cells and evaluation of new targeted therapeutics.<sup>131,132</sup>

#### **4.1. Overview of fluorescent probes for imaging cellular biotin receptors.**

In this chapter, a brief discussion on recent efforts to develop fluorescent probes for BRs will be outlined.

Chen *et al.* reported biotin-fluorescein and biotin-linker-coumarin for a mechanistic study on tumor-targeting drug delivery system (DDS) by receptor-mediated internalization (Figure 28).<sup>133</sup> The introduction of self-immolative disulfide linker in biotin-linker-coumarin (Figure 28) was aimed to evaluate the release of the fluorophore triggered by endogenous thiols like glutathione (GSH). This model closely mimicked the drug-release system, biotin-linker-taxoid-fluorescein, designed in the study (Figure 28). Both probes showed cell-specificity against three cell lines in the following order L1210FR>L1210>WI38, which correlated with BRs expression level. Later, Ojima's group developed dual-warhead DDS (Figure 28).<sup>134</sup> It is based on 1,3,5-triazine core with three side arms: biotin and toxic drugs (SB-T-1214 in red and camptothecin in blue, figure 28). In addition to its antitumor activity, camptothecin served as a fluorescent unit as well, enabling evaluation of the prodrug internalization. In the same study, redesign of biotinylated fluorescein yielded in more water-soluble variant biotin-PEG-fluorescein (Figure 28). Biotin-PEG-fluorescein was assessed in a screening of BR-expression level in a number of human breast cancer cell lines, MX-1, MCF-7, LCC6-WT, LCC6-MDR, MDA-MB 231, and SkBr3, which had not been reported previously.



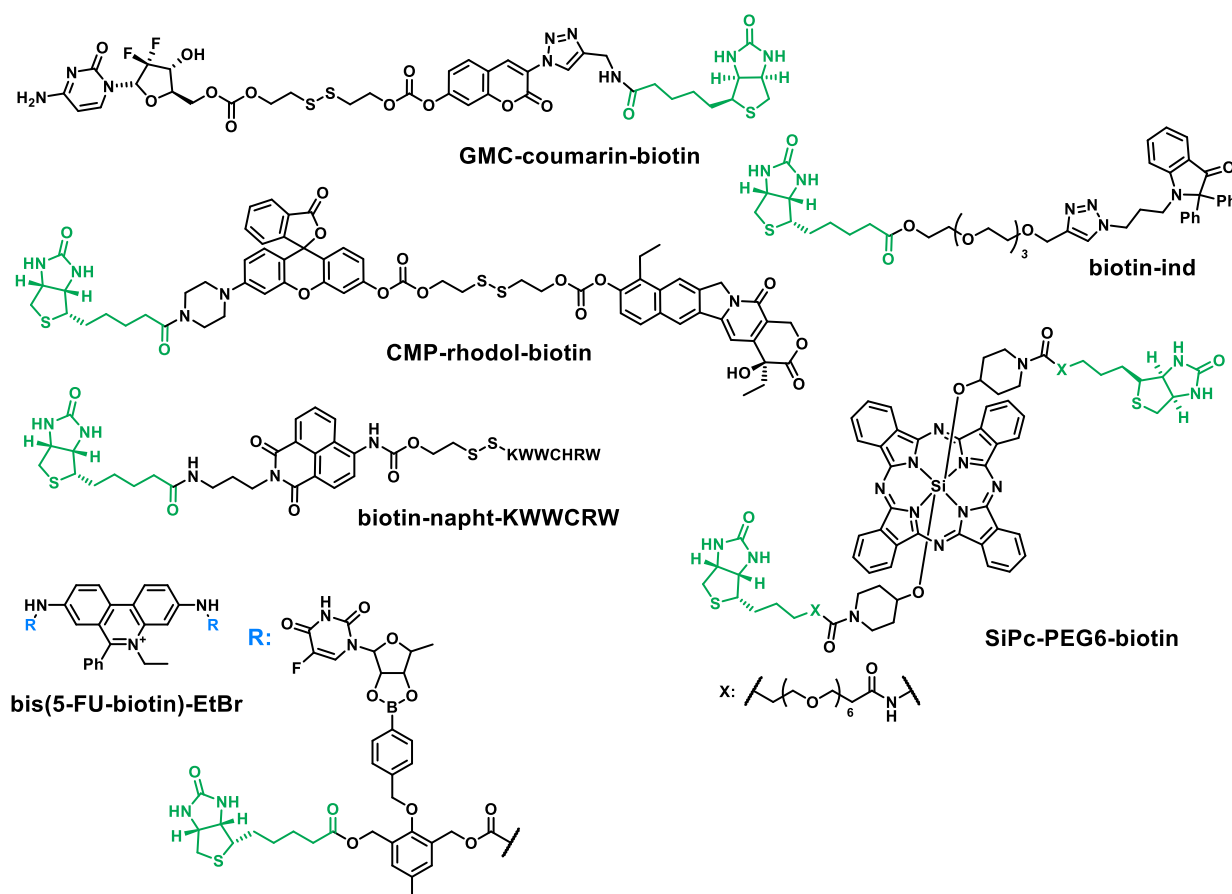
**Figure 28.** Examples of biotin-conjugates for studying tumor-targeting DDS.

Kim group reported a series of single molecular imaging and therapeutic molecules. Biotin-SS-coumarin was designed to detect intracellular thiols in living cancer cells upon disulfide bond cleavage (Figure 28).<sup>135</sup> It showed the selectivity to cancer A549 cells over WI38. Maiti *et al.* developed a cancer-targeting theranostic DDS that contains a gemcitabine (GMC) as chemotherapeutic drug and a coumarin moiety (signaling unit) which is coupled with biotin as a targeting ligand (GMC-coumarin-biotin, figure 29).<sup>136</sup> Confocal microscopy experiments showed that prodrug GMC-coumarin-biotin went to BR-positive tumor cells A549, but not to BR-negative WI38 cells. In the presence of thiols in cancer cells, the disulfide bond cleaved, followed by thiolactonization and release of GMC.

Then, the same group expanded the palette of theranostic agents with CMP-rhodol-biotin.<sup>137</sup> CMP-rhodol-biotin consists of camptothecin analog and biotinylated peperazine-rhodol (Figure 29).<sup>137</sup> The liberation of the drug underwent using S-S bridge reduction by intercellular thiols. Evaluation of cellular uptake showed selectivity towards cancer cells (HeLa and A549). The prodrug showed an anti-proliferation activity against HeLa and an antitumor effect in xenograft murine model.<sup>137</sup> Later on, Kim and colleagues showed that biotinylation of Holliday junction (HJ) inhibitor

peptide2 (KWWCRW) increased its bioavailability and metabolic stability.<sup>138</sup> To assess this study, naphthalimide with disulfide bridge was used as signaling unit. The targetability of biotin-napht-KWWCRW (Figure 29) was tested in cancer HepG2 (BR-positive) and WI38 (BR-negative) cells.

In 2014, Kim and co-workers reported antitumor prodrug activated by reactive-oxygen species (ROS).<sup>139</sup> The prodrug (bis(5-FU-biotin)-EtBr, figure 29) consisted of two biotin-targeting units, two 5'-deoxy-5-fluorouridine molecules (5-FU, prochemotherapeutic), two ROS-sensitive boronate triggers, and an ethidium bromide as fluorescence reporter.<sup>139</sup> Upon activation by ROS, the cascade of the reactions released active drug with the fluorescence response of ethidium bromide (EtBr) as the sign of apoptosis. *In vivo* and *ex vivo* imaging data revealed the specific uptake of prodrug at tumor regions over other organs.



**Figure 29.** Examples of biotin-conjugates for studying tumor-targeting DDS.

Lin group developed bis-biotinylated Si-phthalocyanines (SiPc-biotin<sup>140</sup> and SiPc-PEG6-biotin<sup>141</sup>, figure 29) for targeted photodynamic therapy. SiPc-biotin<sup>140</sup> selectively accumulated in BR-positive (HeLa) cells and then killed them after the irradiation. The

effect was reduced when cells were pretreated with biotin showing specificity of the probe. To increase water-solubility and prevent aggregation, SiPc-PEG6-biotin<sup>141</sup> was developed. SiPc-PEG6-biotin selectively accumulated in tumor tissues in mice thus preventing further progression of its growth.

Recently, Pal *et al.* developed BR-selective biotinylated solvatochromic probe based on indoline and tetraethylene glycol linker (biotin-ind, figure 25).<sup>142</sup> Interestingly, the probe allowed visualization of 3D spheroid and was compatible in two-photon excitation (2PE) microscopy.

Alternatively, Jin *et al.* proposed a dual-emissive nanoprobe based on quantum dot nanohybrid consisting of silica-encapsulated red-emitting QD (rQD@uSiO<sub>2</sub>, used as the “core”), green-emitting QDs (gQDs, used as “satellites”) and the shell decorated with biotins. The nanoprobe enabled semiquantitatively assessed level of BR-expression in HCT116, MDA-MB-231, SKBr3 and HepG2 cells.<sup>143</sup>

#### **4.2. Probing cell surface biotin receptors with rationally designed fluorogenic dimerized squaraines (manuscript, article 2)**

Despite the progress in biotin-probes, no fluorescent probes operating in far-red region have been reported. Our group<sup>24,25,35</sup> and others<sup>17,18</sup> showed that a very bright fluorophore like squaraine is particularly interesting in probe development because of its superior photophysical properties.

In this project, we aimed to develop a selective fluorogenic probe based on dimerization-caused quenched squaraines for selective BR-imaging. To reach this goal, we undertook a systematic analysis of relation between the bio-probe structure and its fluorogenic properties. To this end, we designed and synthesized dimeric squaraines with rigid L-lysine linker or a flexible PEG8-bridge. We attached biotin as the ligand enabling to visualize cell-surface biotin receptors (BRs) in cancer cells.

Our structure-properties relationship study revealed that a small and relatively rigid L-lysine served as a better connector in dimer-design in comparison to PEG8-linker. L-lysine-based dimer promoted retention of unimolecular quenched form while preventing unspecific interaction in cellular environment. We developed Sq<sub>2</sub>B enabling us to probe cell-surface biotin receptor in cancerous cells.

## Probing cell surface biotin receptors with rationally designed fluorogenic dimeric squaraines

Kyong T. Fam,<sup>a</sup> Mayeul Collot<sup>\*a</sup> and Andrey S. Klymchenko<sup>\*a</sup>

Received 00th January 20xx,  
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Fluorogenic probes enable imaging biomolecular targets with high sensitivity and minimal background to noise ratio in non-wash conditions. Here, we focused on the molecular design of dimeric squaraines that in aqueous media form self-quenched intramolecular H-aggregates, but turn on their fluorescence in apolar environment due to opening of the dimer. Our structure-properties study revealed that a relatively short L-lysine linker served as a better connector to dimerize squaraine dyes in comparison to PEG7-linker. Additional small PEG-linkers increase water solubility, promote retention of the unimolecular form while preventing nonspecific interactions in biological environments. The best selected fluorogenic dimer conjugated to biotin was successfully applied for imaging biotin receptors, which are overexpressed in cancer cells. It was shown that the new probe presents certain specificity to sodium-dependent multivitamin transporter (SMVT) showing competitive displacement with its targets, such as biotin, lipoic acid or sodium pantothenate. The obtained results provide guidelines for development of new dimerization-based fluorogenic probes and propose new bright tools for imaging biotin receptors.

### Introduction

Biotin is an essential vitamin playing its role in cellular carbohydrate, amino acid and lipid metabolism. Unlike bacteria, mammalian cell machinery does not produce biotin, therefore biotin is supplemented exogenously. There is an evidence that expression of biotin receptors (BRs) is correlated with cancer. While biotin receptors have been attractive target as the drug delivery system to tumour, there is lack of robust imaging probes for clinical diagnostic of BRs in cancerous cells and evaluation of new targeted therapeutics.<sup>1,2</sup>

Fluorogenic probes are particularly adapted for deciphering biological processes with background-free imaging.<sup>3,4</sup> Although environment-sensitive biotin probes have been developed for BRs imaging operating in the visible region, detection of BRs at low concentration requires superior brightness. Moreover, there is a particular demand in development of probes operating in far-red and near-infrared (NIR) region to image deeper in the tissues and potentially *in vivo*.

Squaraines are particularly interesting for probe development since they have exceptional molar extinction coefficient ( $\sim 300,000 \text{ M}^{-1}\text{cm}^{-1}$ ) and their absorption and emission are in the far-red to NIR window.<sup>5</sup> In aqueous environment, due to high lipophilic nature squaraines tend to self-aggregate forming non-fluorescent species, phenomenon known as aggregation-caused quenching (ACQ).<sup>4</sup> ACQ is caused

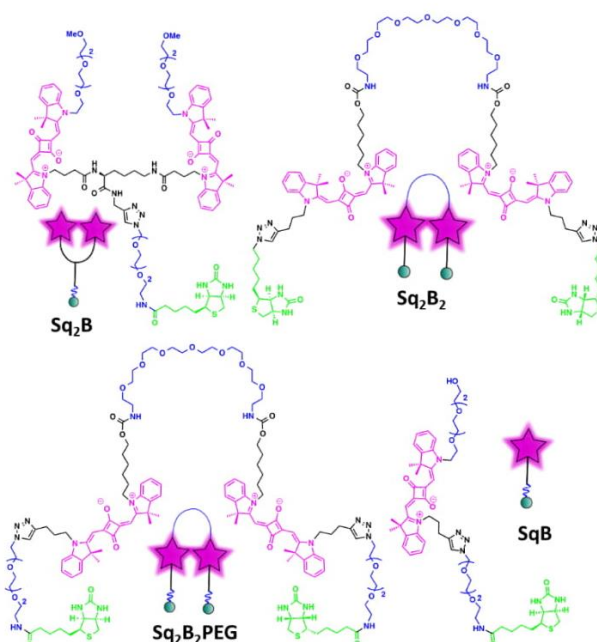


Fig. 1. Chemical structures of biotinylated fluorogenic squaraines.

by strong hydrophobic effects or  $\pi$ - $\pi$  stacking of fluorophores and is commonly observed as non-linearity in fluorescence intensity at high concentrations.<sup>6</sup> This phenomenon is common for poorly soluble dyes (typically hydrophobic fluorophores in aqueous media) since its planar aromatic systems provoke formation of non-radiative excimers/exciples.<sup>6</sup> Aggregation is thermodynamically favorable process and is challenging to

<sup>a</sup> Nanochemistry and Bioimaging group, Laboratoire de Bioimagerie et Pathologies, CNRS UMR 7021, Université de Strasbourg, Faculté de Pharmacie, 67401 Illkirch, France.

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x



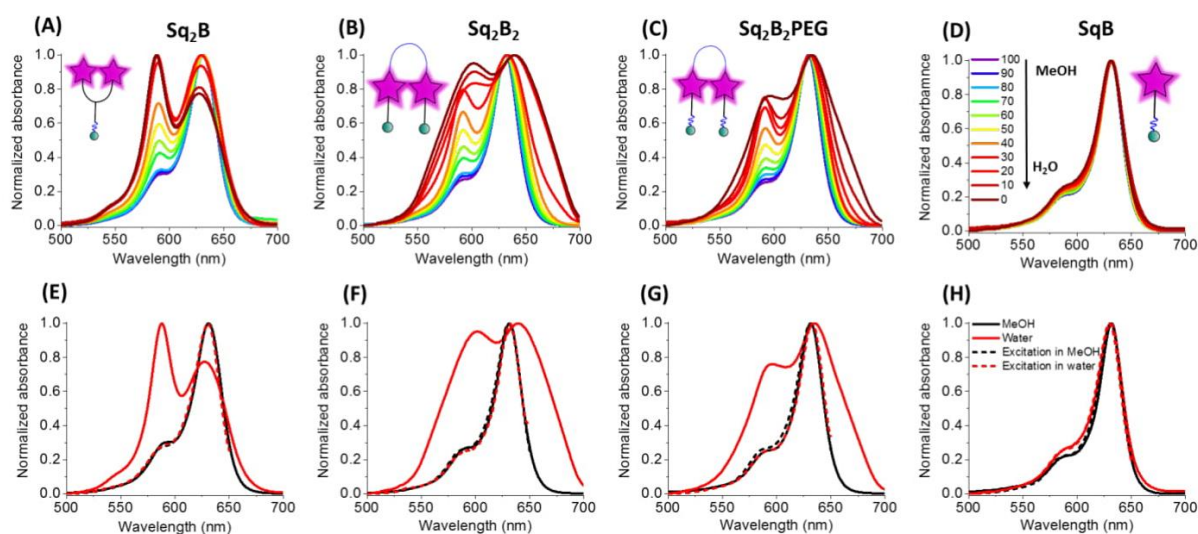


Fig. 2. Photo-physical properties of biotinylated squaraine probes. Normalized absorbance spectra of Sq<sub>2</sub>B (A), Sq<sub>2</sub>B<sub>2</sub> (B), Sq<sub>2</sub>B<sub>2</sub>PEG (C) and SqB (D) in the gradient mixture of MeOH-H<sub>2</sub>O (from 100 to 0). Normalized absorbance and excitation spectra in MeOH and water of Sq<sub>2</sub>B (E), Sq<sub>2</sub>B<sub>2</sub> (F), Sq<sub>2</sub>B<sub>2</sub>PEG (G) and SqB (H). Concentration of probes was fixed to 0.2 μM.

disrupt.<sup>6</sup> Therefore, ACQ can generate an intensometric fluorogenic response if the initially quenched species are spatially separated upon interaction with the biological target.<sup>7</sup> Both intermolecular and intramolecular aggregation processes can be used to design ACQ-driven probes. Previously, in order to obtain fluorogenic probes for plasma membranes, we designed amphiphilic squaraines that assemble in water into non-fluorescent ACQ-particles, but upon binding to cellular membranes disassembled, activating fluorescence.<sup>8</sup> Additionally, we combined two squaraine units into one molecule, which folded in aqueous media into intramolecular quenched dimer. Attachment of a receptor-specific ligand to a dimeric squaraine resulted in fluorogenic probe that turned on its fluorescence on binding specifically to membrane receptor oxytocin.<sup>9</sup>

In this work, we rationally designed and synthesized four variants of biotinylated fluorogenic squaraines. We studied structure-properties relationship as well as evaluated lead candidate's capacity and specificity to image cellular BRs in a bright and specific manner.

## Results and discussion

### Synthesis of biotinylated squaraine dimers

Recently, molecular aggregation approach has attracted attention in development of ACQ-driven novel fluorogenic probes.<sup>7</sup> While self-assembly of dyes in aqueous solution is thermodynamically favourable process, it is dependent on the concentration and the structure of the dye. Therefore, preparation of fluorophore dimers is an attractive approach, because in these systems the aggregation state depends only the environment, but not on the probe concentration.<sup>4</sup>

We wondered if designing different type of dimeric squaraines can tune their fluorescence turn-on response and

whether attachment of one or two ligands (biotins) can affect the interaction with BRs. To this end, we designed and synthesised two dimers of squaraines connected with a flexible PEG7-bridge (Sq<sub>2</sub>B<sub>2</sub>, Sq<sub>2</sub>B<sub>2</sub>PEG, fig. 1). PEG7-diamine was coupled to activated acid **5** resulting in **6** with two propargyl handles, which were further clicked to N<sub>3</sub>-biotin or N<sub>3</sub>-PEG3-biotin (Scheme 2 SI). We assumed that introduction of additional short PEG-linker between biotin and squaraine unit in Sq<sub>2</sub>B<sub>2</sub>PEG would increase the solubility of the probe in aqueous solutions as well as add the degree of flexibility to dynamically fold and prevent nonspecific binding to biomolecules. We also redesigned previously reported squaraine dimer<sup>9</sup> and synthesised a shorter more rigid dimer (Sq<sub>2</sub>B, fig. 1), based on L-lysine linker with one biotin ligand, and monomeric dye analogue (SqB, fig. 1).

### Fluorogenic properties of squaraine dimers

To evaluate the fluorogenic nature of the dimers, we studied their photophysical properties at a fixed concentration (0.2 μM) in the mixture of water with gradual increase of organic solvent content (MeOH). All four probes exhibited absorption (Fig.2) and emission (Fig.4) spectra in MeOH typical for monomeric fluorophore, characterized by narrow bands with maxima at ~630 and ~640 nm, respectively. With increase in the fraction of water in methanol, the monomeric squaraine SqB did not show any significant changes in the absorption spectra, indicating that at 0.2 μM it did not aggregate (Fig.2D). However, some decrease in fluorescence intensity (Fig.4D) in aqueous compositions was observed, probably caused by water quenching.<sup>10</sup> The absence of aggregation of monomeric SqB can be attributed to high degree of PEGylation that ensures its good water solubility. On the other hand, all dimers showed formation of blue-shifted band at higher water fractions (Fig.2A-C). This new band was absent in the fluorescence excitation spectra of all dimers (Fig.2E-G), superposed with

those in methanol. This indicates that this blue shifted absorption band is non-emissive. According to our previous studies,<sup>9</sup> this new band can be assigned to intramolecular *H*-aggregates of squaraines.

Table 1 Photo-physical properties of biotinylated squaraine probes

		$\Phi$	$\epsilon$ ( $M^{-1}cm^{-1}$ )	Brightness	Relative brightness $s^a$	$\Phi$ enh.
<b>Sq<sub>2</sub>B</b>	MeOH	0.15	684,000	104,170	1.21	46.60
	H <sub>2</sub> O	0.003	684,000	2,235	0.026	
<b>Sq<sub>2</sub>B<sub>2</sub>PEG</b>	MeOH	0.25	684,000	167,470	1.95	24.55
	H <sub>2</sub> O	0.01	684,000	6,820	0.080	
<b>Sq<sub>2</sub>B<sub>2</sub></b>	MeOH	0.18	684,000	124,334	1.45	149.54
	H <sub>2</sub> O	0.0012	684,000	831	0.0097	
<b>SqB</b>	MeOH	0.13	342,000	45,052	0.53	3.86
	H <sub>2</sub> O	0.034	342,000	11,672	0.136	

<sup>a</sup> Relative brightness to DiD in methanol ( $\Phi = 0.33$ ).<sup>11</sup>

However, at the highest water fractions and in neat water, the shape of the absorption spectra of dimers depended on their chemical structure (Fig.2A-C). Indeed, **Sq<sub>2</sub>B** showed typical sharp blue shifted *H*-aggregate absorption band centred at 588 nm and a peak at 632 nm (Fig.2A), whereas for PEG-based dimers displayed broad absorption spectra (Fig.2B and C). We assumed that the difference originates from their aggregation behaviour: intra-, intermolecular aggregation or the synergy of both phenomena.

To decipher the nature of the aggregates, we studied the absorption spectra of the squaraines at different concentrations

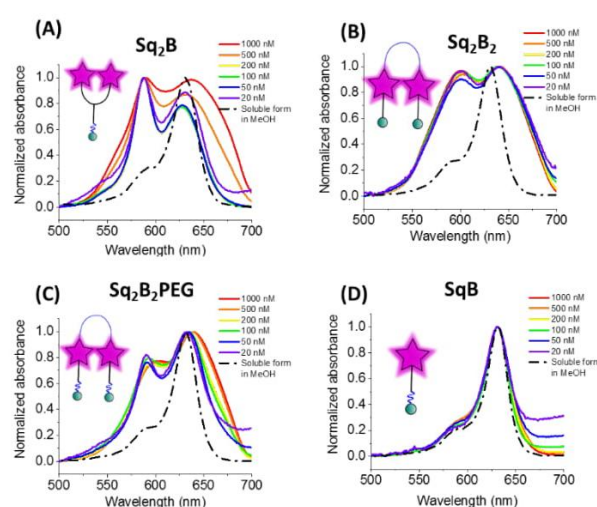


Fig. 3. Absorbance spectra at different concentrations in water of Sq<sub>2</sub>B (A), Sq<sub>2</sub>B<sub>2</sub> (B), Sq<sub>2</sub>B<sub>2</sub>PEG (C) and SqB (D). For comparison normalized absorbance spectra of probes are shown in black dashed line.

(20 nM to 1  $\mu$ M) (Fig.3). Lysine-based dimer **Sq<sub>2</sub>B** displayed sharp absorption bands that did not change their relative intensity up to 200 nM (Fig.3A). This result confirms the intramolecular nature of the form *H*-aggregate, which should not be concentration dependent. However, at higher concentrations, the bands of **Sq<sub>2</sub>B** broadened, indicating a formation of more complex (intermolecular) aggregates. In case

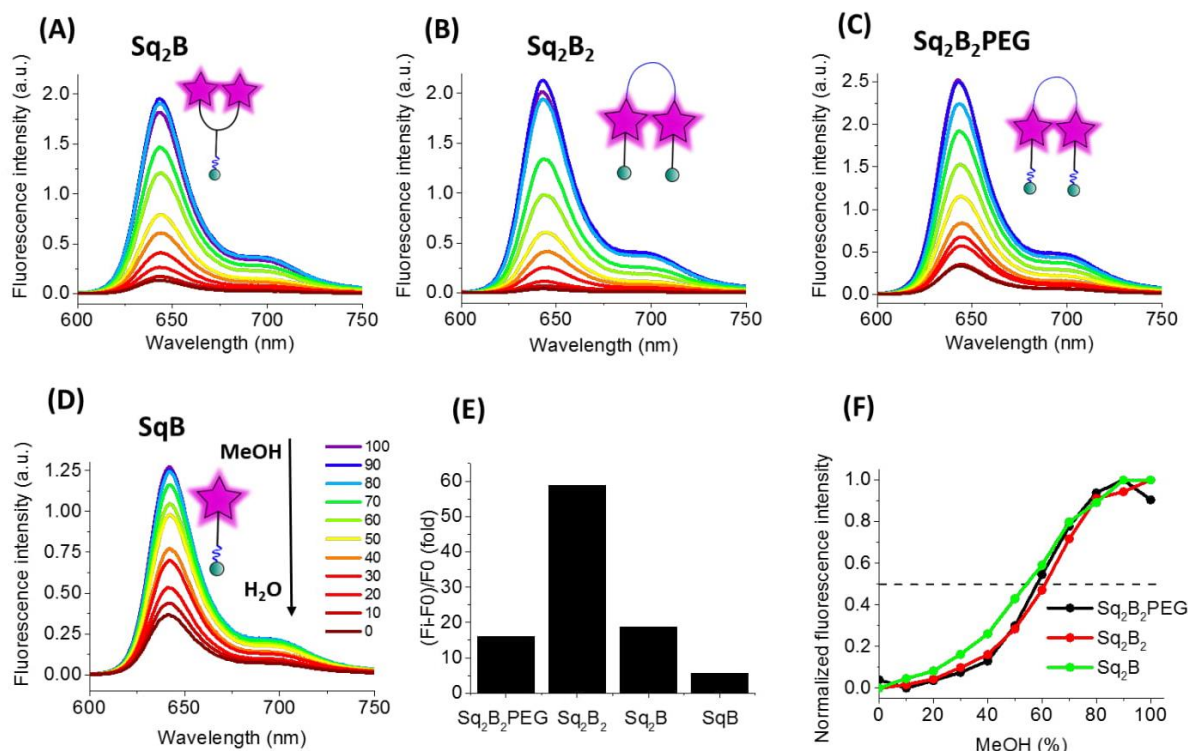


Fig. 4. Photo-physical properties of biotinylated squaraine probes. Fluorescence spectra of Sq<sub>2</sub>B (A), Sq<sub>2</sub>B<sub>2</sub> (B), Sq<sub>2</sub>B<sub>2</sub>PEG (C) and SqB (D) in the gradient mixture of MeOH-H<sub>2</sub>O (from 100 to 0). Analysis of normalized fluorescence maxima in the gradient mixture of MeOH-H<sub>2</sub>O (from 100 to 0) of Sq<sub>2</sub>B, Sq<sub>2</sub>B<sub>2</sub>, Sq<sub>2</sub>B<sub>2</sub>PEG and SqB. (F) Fluorescence enhancement plot from water to MeOH of Sq<sub>2</sub>B, Sq<sub>2</sub>B<sub>2</sub>, Sq<sub>2</sub>B<sub>2</sub>PEG and SqB. Concentration of probes was fixed to 0.2  $\mu$ M.



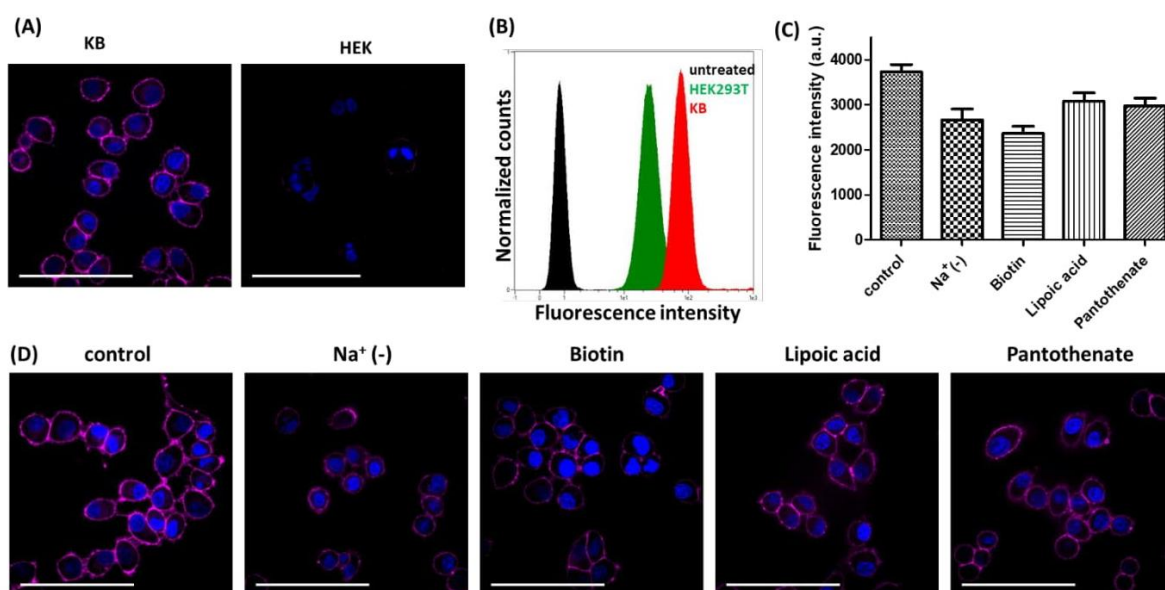


Fig. 3. Fluorescence microscopy images. (A) Live cell fluorescence imaging of Sq<sub>2</sub>B in BR-positive (KB) and BR-negative (HEK293T) cells. (B) Flow cytometry analysis of KB and HEK293T cell treated with Sq<sub>2</sub>B. Untreated KB cells were used as negative control. (C) Analysis of fluorescence intensities from images (D). Bars represent mean value  $\pm$  S.D. (n=6). (D) Fluorescence imaging of Sq<sub>2</sub>B in KB cells in Na<sup>+</sup>-depleted HBSS, pretreated with biotin (100  $\mu$ M), lipoic acid (100  $\mu$ M) or pantothenate (100  $\mu$ M) compared to control. Scale

of bisbiotinylated **Sq<sub>2</sub>B<sub>2</sub>PEG** at low concentration, the bands were also relatively sharp and concentration-independent, although the *H*-aggregate band were less intense (Fig.3C). Thus, similar to **Sq<sub>2</sub>B**, **Sq<sub>2</sub>B<sub>2</sub>PEG** can also form intramolecular *H*-aggregate, but its fraction is lower probably because its squaraine units have higher degree of freedom in case of PEG7-based linker vs lysine linker. At concentrations above 100 nM **Sq<sub>2</sub>B<sub>2</sub>PEG** showed band broadening probably due to intermolecular aggregation. By contrast, bisbiotinylated dimer without PEG3 units, **Sq<sub>2</sub>B<sub>2</sub>**, displayed broad absorption spectra at all studied concentrations (Fig.3B), which is probably caused by lipophilic nature of the molecule that favour the intermolecular aggregation even at low concentration. Unlike dimeric variants, monomeric **SqB** kept its soluble non-aggregated form throughout the studied concentration range resulting in linear increase of the fluorescence with increase of the concentration (Fig.S13E).

The observed solvent-dependent *H*-aggregation of the dimers resulted in a fluorogenic behaviour of the dimers when the fraction of methanol in water increased (Fig.4A-C). Indeed, all these dimers showed >10-fold fluorescence increase from water to methanol with the strongest change for **Sq<sub>2</sub>B<sub>2</sub>**, whereas the monomeric **SqB** showed only minor intensity increase (Fig. 4E). The normalized titration curves revealed that the intensity increase for **Sq<sub>2</sub>B** was continuous starting from low methanol fraction, while for other dimers, it was sigmoidal with the strong increase only at higher methanol fraction (Fig. 4F). This difference in the sensitivity to methanol correlates with different behaviour aggregation behaviour of the dyes. Indeed, at the low methanol content (*i.e.* high water fractions with 0.2  $\mu$ M dye) only **Sq<sub>2</sub>B** is present in form of mono-molecular *H*-aggregate (Fig.2A). The latter is probably more sensitive to low fractions of methanol that opens this monomolecular dimer,

compared to the other two dyes that form in these conditions the intermolecular aggregates (Fig.2B, C).

In the next step, we evaluated potential nonspecific interactions of the biotin conjugates with biological components: bovine serum albumin (BSA), fetal bovine serum (FBS), lipid vesicles composed of DOPC and biotin. To this end we measured fluorescence intensity of the dye, which was expected to increase in case of nonspecific interaction with these components, and compared it to that of activated form of dyes in methanol. To our surprise, each probe responded differently to these model systems (Table 2). Probes **Sq<sub>2</sub>B** and **Sq<sub>2</sub>B<sub>2</sub>PEG** showed lower degree of nonspecific fluorescent turn-on response in comparison to more lipophilic **Sq<sub>2</sub>B<sub>2</sub>**. Generally, all dimers strong fluorescence increase in the presence of BSA (31-40% of open form, fig. S14). Surprisingly, **SqB** was more than twice more fluorescent in the presence of BSA in comparison to its emission in MeOH. Such a behaviour is probably caused by binding of the dye into the hydrophobic pocket of the albumin, where the intramolecular rotation of the molecules is restricted, favouring strong emission of the dye. Importantly, in FBS, which is the biological medium rich in proteins and lipoproteins, **Sq<sub>2</sub>B<sub>2</sub>** and **Sq<sub>2</sub>B** showed significantly lower fluorescence intensity compared to **Sq<sub>2</sub>B<sub>2</sub>PEG** probably because higher water solubility of the former decreased their nonspecific interactions. As expected, the presence of free biotin did not contribute to unspecific fluorescence. In case of DOPC vesicles, **Sq<sub>2</sub>B<sub>2</sub>** showed the weakest fluorescence, probably of its strong aggregation.

Table 2 Unspecific turn-on fluorescence by interaction with biomolecules

% of max F.I.	Sq <sub>2</sub> B <sub>2</sub> PEG	Sq <sub>2</sub> B <sub>2</sub>	Sq <sub>2</sub> B	SqB
MeOH	100	100	100	100
H <sub>2</sub> O	13.30	2.067	7.47	28.93

PBS	6.96	2.060	2.098	16.99
BSA	40.075	31.17	38.56	243.50
FBS	18.36	8.20	9.084	127.91
Avidin	6.83	2.31	1.98	21.18
Streptavidin	7.075	2.56	1.51	11.75
Biotin	7.25	2.24	4.56	19.74
DOPC	21.33	7.81	17.32	91.35

#### Imaging of biotin receptors with Sq<sub>2</sub>B

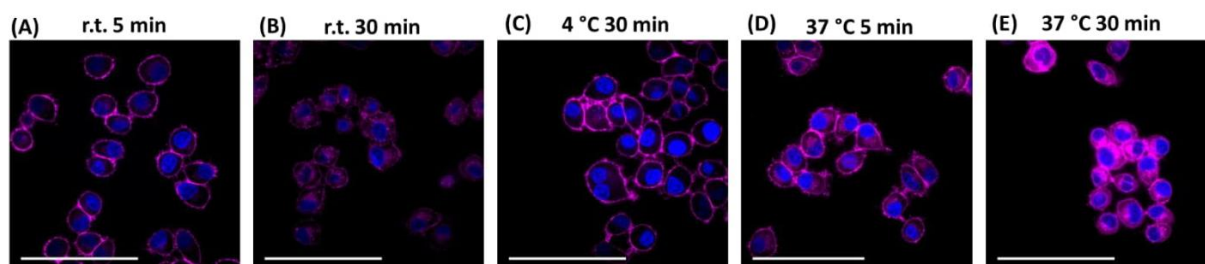


Fig. 4. Fluorescence microscopy images of KB cells treated with Sq<sub>2</sub>B at different conditions. (A) Incubation for 5 min at room temperature (r.t.). (B) Incubation for 30 min at r.t. (C) Incubation for 30 min at 4 °C. (D) Incubation for 5 min at 37 °C. (E) Incubation for 5 min at 37 °C. Scale bar, 30 μm.

After characterisation of spectral properties, we pursued cellular studies. We chose Sq<sub>2</sub>B probe since it showed low unspecific fluorescence with biomolecules as well as retained its water soluble quenched monomolecular form in range 20 to 200 nM. Cancer cells, like KB (HeLa cells derivatives) tend to highly consume biotin as vital supplement in comparison to other cells suggesting expression of BRs.<sup>1</sup> Therefore, we used KB cells as cellular model to evaluate performance of Sq<sub>2</sub>B and HEK293T cells were used as BR-negative cell line.<sup>1</sup> The biotinylated Sq<sub>2</sub>B showed low toxicity in KB cells in the range of imaging concentrations (Fig.S15). Freshly prepared solution in HBSS (200 nM) of the biotinylated fluorogenic dyes were briefly incubated with cells for 5 min at room temperature and imaged without washing. In KB cells, Sq<sub>2</sub>B readily produced distinctive fluorescence labelling of the cell surface, whereas not signal of the dye was observed in case of HEK293T cells (Fig. 5A). We additionally evaluated specificity of Sq<sub>2</sub>B by flow cytometry. The population Sq<sub>2</sub>B-treated KB cells exhibited significantly higher brightness compared to (HEK293T) in the cell-by-cell analysis (Fig. 5B). This result confirmed our observations by microscopy, suggesting that Sq<sub>2</sub>B distinguishes KB cells from HEK293T probably because the former is BR-positive.

Next, we evaluated the specificity of Sq<sub>2</sub>B to BRs in KB cells. Sodium-dependent multivitamin transporter (SMVT) is well characterized BR in cancer cells to deliver biotin.<sup>1</sup> Except for biotin transport, SMVT is also known to be vital for translocation of essential cofactors like lipoic acid and pantothenate.<sup>12</sup> To evaluate the targetability of Sq<sub>2</sub>B to SMVT, we performed a set of competition experiments (Fig.5D). Firstly, to evaluate dependence of the staining on the presence of extracellular Na<sup>+</sup> ions, we incubated KB cells with Sq<sub>2</sub>B in Na<sup>+</sup>-deprived HBSS buffer (Na<sup>+</sup> was replaced with choline). The results showed the decrease in Sq<sub>2</sub>B fluorescence when Na<sup>+</sup>-deprived media was used indicating a strong effect of sodium ions on the Sq<sub>2</sub>B cellular binding (Fig. 5D). Secondly, KB cells were pre-treated with competitors (biotin, lipoic acid or sodium pantothenate) to

saturate SMVT and thus reduce its availability for Sq<sub>2</sub>B. Further brief incubation with Sq<sub>2</sub>B showed the decrease in the fluorescence in competitor-treated KB cells in comparison to untreated ones (Fig. 5D). These results suggest that Sq<sub>2</sub>B displays certain specificity for SMVT.

We further studied the internalisation of the probe. When cells were incubated with Sq<sub>2</sub>B at room temperature for 30 min, fluorescence was observed in the cytoplasm (Fig.6B), particularly in the endosomes (Video S1). Incubation at 37 °C

increased the fluorescence inside the cells, indicating faster dynamic interaction with BRs and higher uptake rate (Fig.6D, E). However, when KB cells were incubated with probes at 4 °C for 30 min (Fig.6C) only the cell surface pattern was detected, suggesting the occurrence of binding event. Absence of the intracellular signal suggested that the dye does not internalize passively into the cells, while the active transport is inhibited by the low temperature. This set of experiments demonstrated that biotin probe Sq<sub>2</sub>B can readily report the BRs on the cell surface and the internalization of the molecules is caused by active transport.

#### Materials and methods

**Synthesis.** The synthesis, protocols, characterizations, and spectra are described in the Supporting Information. NMR spectra were recorded on a Bruker Avance III 400 MHz spectrometer. Mass spectra were obtained using an Agilent QTOF 6520 mass spectrometer.

**Spectroscopy.** Absorption spectra were recorded on a Cary 5000 spectrophotometer (Varian) and fluorescence spectra on a F55 (Edinburgh Instruments) spectrofluorometer. The fluorescence signal was corrected for the lamp intensity fluctuations and for the wavelength-dependent sensitivity of the detector. Relative fluorescence quantum yields were measured using DiD in methanol<sup>11</sup> (QY = 0.33) as standard.

**Microscopy imaging.** Cells were grown at 37 °C in humidified atmosphere containing 5% CO<sub>2</sub>: KB cells (ATCC® CCL-17) in minimum essential medium (MEM, Gibco-Invitrogen) with 10% fetal bovine serum (FBS, Lonza), 1% non-essential amino acids (Gibco-Invitrogen), 1% MEM vitamin solution (Gibco-Invitrogen), 1% L-Glutamine (Sigma Aldrich) and 0.1% antibiotic solution (gentamicin, Sigma-Aldrich); HEK293T (ATCC® CRL-3216™) in Dulbecco's Modified Eagle Medium without phenol red (DMEM, Gibco-Invitrogen) supplemented with 10% fetal bovine serum (FBS, Lonza), 1% L-Glutamine (Sigma Aldrich)



and 1% antibiotic solution (Penicillin-Streptomycin, Sigma-Aldrich). Cells were seeded onto a chambered coverglass (IBiDi®) 24 h before the microscopy measurement. For imaging, the culture medium was removed, the attached cells were washed with Hank's Balanced Salt Solution (HBSS, Gibco-Invitrogen) and incubated with solutions of biotinylated probes (200 nM). In competition experiment, KB cells were pretreated with biotin (100 µM) for 30 min prior to incubation with corresponding biotinylated probe. Images were taken with Nikon Ti-E inverted microscope, equipped with CFI Plan Apo × 60 oil (NA = 1.4) objective, X-Light spinning disk module (CresOptics) and a Hamamatsu Orca Flash 4 sCMOS camera, was used. The microscopy settings were: Hoechst (ex. 405 nm, em. 510±42 nm), Squaraine (ex. 638 nm, em. 705±36 nm). The images were recorded using NIS Elements and then processed with Icy.

**Flow cytometry.** Cells were grown at 37°C in humidified atmosphere containing 5% CO<sub>2</sub> in 25 cm<sup>2</sup> (Nunc™ EasYFlask, ThermoFisher). On the day of the analysis, the cells were washed and harvested. The cell suspension (3×10<sup>5</sup> cells/mL) was incubated with corresponding biotinylated probe (200 nM) for 5 min at room temperature and analyzed immediately using flow cytometry (MACSQuant VYB, Miltenyi Biotec).

**Cytotoxicity assay.** Cytotoxicity assay of the dyes was quantified by the MTT assay (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide). A total of 1×10<sup>4</sup> KB cells/well were seeded in a 96-well plate 24 h prior to the cytotoxicity assay in growth medium and were incubated in a 5% CO<sub>2</sub> incubator at 37°C. After medium removal, an amount of 100 µL DMEM containing 5 µM, 1 µM or 0.2 µM of was added to the KB cell and incubated for 24 h at 37°C (5% CO<sub>2</sub>). As control, for each 96-well plate, the cells were incubated with DMEM containing the same percentage of DMSO (0,5% v/v) as the solution with the tested dyes. After 24h of dye incubation, the medium was replaced by 100 µL of a mix containing DMEM + MTT solution (diluted in PBS beforehand) and the cells were incubated for 4 h at 37°C. Then, 75 µL of the mix was replaced by 50 µL of DMSO (100%) and gently shaken for 15 min at room temperature in order to dissolve the insoluble purple formazan reduced in living cells. The absorbance at 540 nm was measured (absorbance of the dyes at 540 nm were taken into account). Each concentration of dye was tested in sextuplicate and the data were shown as the mean value plus a standard deviation (±SD). For each concentration, we calculated the percentage of cell viability in reference of the control DMEM+ 0.5% DMSO.

## Conclusions

In conclusion, we designed and synthesized four fluorogenic probes. We focused on the molecular design of dimeric squaraines that enabled tuning the fluorogenic properties while minimizing nonspecific interaction with biomolecules and reducing the background fluorescence.

We made a systematic analysis of structure effect on fluorogenic properties of dimeric squaraines. While designing dimeric fluorogenic dyes some factors should be taken in consideration: (i) close proximity of fluorophores to each other

to form aggregated monomolecular species in aqueous environment; (ii) hydrophilic linkers to prevent nonspecific binding; (iii) effect of the ligand on inter-/intramolecular interaction. Our structure-properties study revealed that a relatively short L-lysine linker served as a better connector in dimer design in comparison to PEG7-linker. Additional small PEG-linkers increase water solubility, promote retention of the unimolecular form while preventing nonspecific interactions in cellular environment.

We focused on **Sq<sub>2</sub>B** as the most promising candidate to report BRs in microscopy and flow cytometry analysis. We showed that **Sq<sub>2</sub>B** underwent active translocation into the cancerous KB (HeLa derivative) cells. We proved that **Sq<sub>2</sub>B** displays some specificity to sodium-dependent multivitamin transporter (SMVT)<sup>1</sup> in a set of experiments. In particular, **Sq<sub>2</sub>B** transport was sodium-dependent and inhibited by native vitamins like biotin, lipoic acid and pantothenic acid.

We believe that this study will provide useful guidelines in designing fluorogenic dimerized probes for bioimaging. Additionally, specific **Sq<sub>2</sub>B** probe will contribute in deciphering the biological role of BRs and built new strategies for evaluation of the new therapeutics.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work received financial support from Agence Nationale de la Recherche (BrightRiboProbes, ANR-16-CE11-0010-01/02) and ERC Consolidator grant (BrightSens, 648528).

## Notes and references

- 1 W. X. Ren, J. Han, S. Uhm, Y. J. Jang, C. Kang, J.-H. Kim and J. S. Kim, *Chem. Commun.*, 2015, **51**, 10403–10418.
- 2 M. Gao, F. Yu, C. Lv, J. Choo and L. Chen, *Chem. Soc. Rev.*, 2017, **46**, 2237–2271.
- 3 H. Zhu, J. Fan, J. Du and X. Peng, *Acc. Chem. Res.*, 2016, **49**, 2115–2126.
- 4 A. S. Klymchenko, *Acc. Chem. Res.*, 2017, **50**, 366–375.
- 5 J. O. Escobedo, O. Rusin, S. Lim and R. M. Strongin, *Current Opinion in Chemical Biology*, 2010, **14**, 64–70.
- 6 J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer US, 3rd edn., 2006.
- 7 D. Zhai, W. Xu, L. Zhang and Y.-T. Chang, *Chem. Soc. Rev.*, 2014, **43**, 2402–2411.
- 8 M. Collot, R. Kreder, A. L. Tatarets, L. D. Patsenker, Y. Mely and A. S. Klymchenko, *Chem. Commun.*, 2015, **51**, 17136–17139.
- 9 I. A. Karpenko, M. Collot, L. Richert, C. Valencia, P. Villa, Y. Mély, M. Hibert, D. Bonnet and A. S. Klymchenko, *Journal of the American Chemical Society*, 2015, **137**, 405–412.
- 10 G. E. Dobretsov, T. I. Syrejschikova and N. V. Smolina, *Biophysics*, 2014, **59**, 183–188.
- 11 I. Texier, M. Goutayer, A. Da Silva, L. Guyon, N. Djaker, V. Josserand, E. Neumann, J. Bibette and F. Vinet, *J. Biomed. Opt.*, 2009, **14**, 054005-054005–11.

## Probing cell surface biotin receptors with rationally designed fluorogenic dimeric squaraines

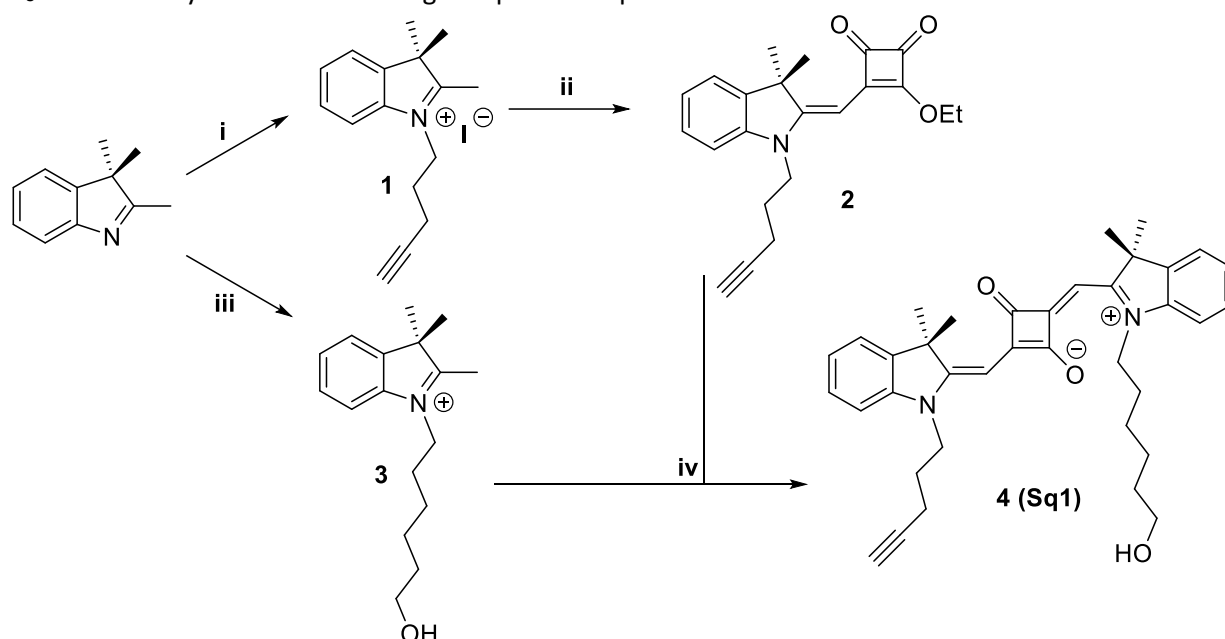
Kyong T. Fam,<sup>a</sup> Mayeul Collot\*<sup>a</sup> and Andrey Klymchenko\*<sup>a</sup>

<sup>a</sup> Nanochemistry and Bioimaging group, Laboratoire de Bioimagerie et Pathologies, CNRS UMR 7021, Université de Strasbourg, Faculté de Pharmacie, 67401 Illkirch, France.

### Supporting information

PEG8-diamine was synthesized as described elsewhere<sup>144</sup>.

N<sub>3</sub>-biotin was synthesized according to a published protocol.<sup>145</sup>



**Scheme 1.** Synthesis of Sq1. i) 5-chloro-1-pentyne, KI, CH<sub>3</sub>CN, 85°C, 48h (20%); ii) Diethyl squarate, Et<sub>3</sub>N, EtOH, 80°C, 6h (36%); iii) 6-chloro-1-hexanol, KI, CH<sub>3</sub>CN, 110°C, 12h (12%); (iv) pyridine, 125°C, 18h (31%)

Compound **1** was synthesized according to a published protocol.<sup>3</sup>

**Synthesis of compound 2.** Compound **1** (1.2 g, 3.4 mmol, 1 eq), diethyl squarate (578 mg, 3.4 mmol, 1 eq) and Et<sub>3</sub>N (g, mmol, 2.7 eq) were dissolved in EtOH (20 mL). The mixture was refluxed overnight. The solvents were then removed under vacuum and the crude product was purified by column chromatography with Heptane/EtOAc (7/3) as eluent. Yellow solid was obtained, yield 36%.

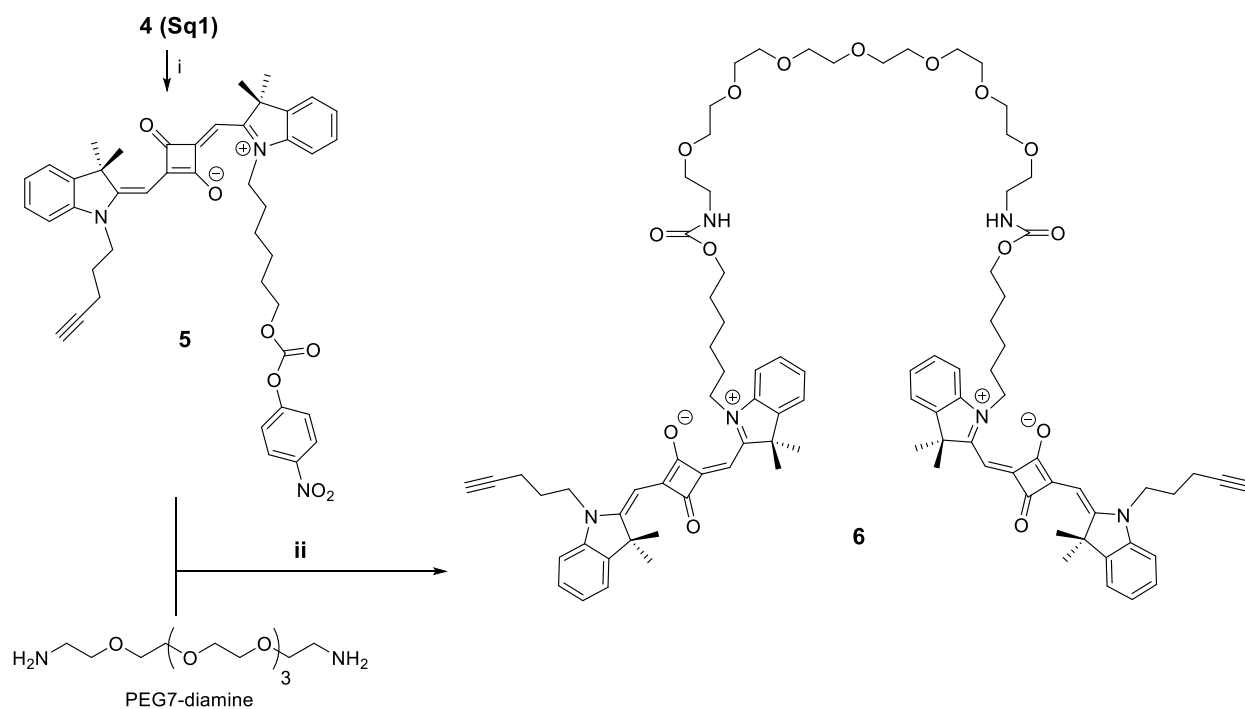
<sup>1</sup>H NMR (400 MHz, Chloroform-d)  $\delta$  7.25 – 7.16 (m, 2H), 7.01 (td,  $J = 7.4, 0.9$  Hz, 1H), 6.93 – 6.87 (m, 1H), 5.41 (s, 1H), 4.82 (q,  $J = 7.1$  Hz, 2H), 3.91 (t,  $J = 7.4$  Hz, 2H), 2.26 (td,  $J = 6.7, 2.6$  Hz, 2H), 2.06 (t,  $J = 2.6$  Hz, 1H), 1.91 (p,  $J = 6.9$  Hz, 2H), 1.56 (s, 6H), 1.46 (t,  $J = 7.1$  Hz, 3H). <sup>13</sup>C NMR (101 MHz, Chloroform-d)  $\delta$  192.35, 187.81, 187.72, 173.87, 168.23, 142.60, 140.83, 127.82, 122.75, 122.02, 108.32, 82.69, 81.57, 70.09, 69.93, 47.94, 41.54, 27.06, 25.11, 16.13, 15.92. HRMS (ES<sup>+</sup>) Calc. for C<sub>22</sub>H<sub>23</sub>NO<sub>3</sub> [M+H]<sup>+</sup> 350.1678, found 350.1757.

**Synthesis of compound 3.** 2,3,3-trimethylindolenine (1 eq, 9.54g), 6-chloro-1-hexanol (1 eq, 8.16 g) and potassium iodide (2.2 eq, 21.9 g) were dissolved in CH<sub>3</sub>CN (75mL), the reaction was heated up to 110°C under stirring overnight. The solvent was then removed under vacuum and the crude product was dissolved in a small amount of acetone. The mixture is then precipitated in cold ether, three times to obtain the maximum of product. Pink crystals were obtained, yield 12%.

<sup>1</sup>H NMR (400 MHz, Chloroform-d) δ 7.74 – 7.64 (m, 1H), 7.59 – 7.46 (m, 4H), 5.24 (s, 1H), 3.54 (t, J = 5.9 Hz, 2H), 3.07 (s, 3H), 1.93 (t, J = 7.6 Hz, 2H), 1.59 (s, 6H), 1.55-1.42 (m, 8H). <sup>13</sup>C NMR (101 MHz, Chloroform-d) δ 195.69, 141.65, 140.94, 130.16, 129.65, 123.35, 115.62, 61.71, 49.88, 32.00, 28.01, 26.31, 25.26, 23.44, 23.26, 17.13.

**Synthesis of 4 (Sq1).** Compound 3 (244 mg, 0.63 mmol, 1.1 eq) and compound 2 (200 mg, 0.573 mmol, 1 eq) were dissolved in 5 mL of pyridine. After 5 h the solution turns green. The solvent was removed under vacuum and the crude product was purified by column chromatography with EtOAc/MeOH (95/5) as eluent. Blue-green powder was obtained, yield 31%.

<sup>1</sup>H NMR (400 MHz, Chloroform-d) δ 7.28 (d, J = 1.2 Hz, 1H), 7.26 (d, J = 1.1 Hz, 2H), 7.24 (dd, J = 2.5, 1.2 Hz, 2H), 7.22 (d, J = 1.3 Hz, 2H), 7.20 (d, J = 1.2 Hz, 1H), 7.09 – 7.00 (m, 5H), 6.93 (s, 1H), 6.91 (s, 1H), 5.92 (s, 2H), 3.93 (s, 4H), 3.61 (t, J = 6.1 Hz, 4H), 2.26 (d, J = 2.7 Hz, 2H), 1.95 (s, 3H), 1.79 – 1.72 (m, 3H), 1.69 (d, J = 3.9 Hz, 27H), 1.58 – 1.49 (m, 3H), 1.45 (dh, J = 9.0, 2.3, 1.8 Hz, 6H). <sup>13</sup>C NMR (126 MHz, Chloroform-d) δ 179.62, 170.54, 169.73, 142.47, 142.26, 127.85, 127.83, 123.91, 123.71, 122.32, 122.29, 109.49, 109.31, 86.78, 82.71, 70.21, 62.40, 49.37, 49.21, 43.21, 42.23, 32.43, 27.14, 26.98, 26.25, 26.24, 25.62, 24.93, 16.11. HRMS (ES+) Calc. for C<sub>37</sub>H<sub>42</sub>N<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup> 563.3195, found 563.3265.



**Scheme 2.** Synthesis of 6. i) pyridine, DCM, r.t., 3 h (61%); ii) DIEA, DMF, 60°C, 18h (21%);

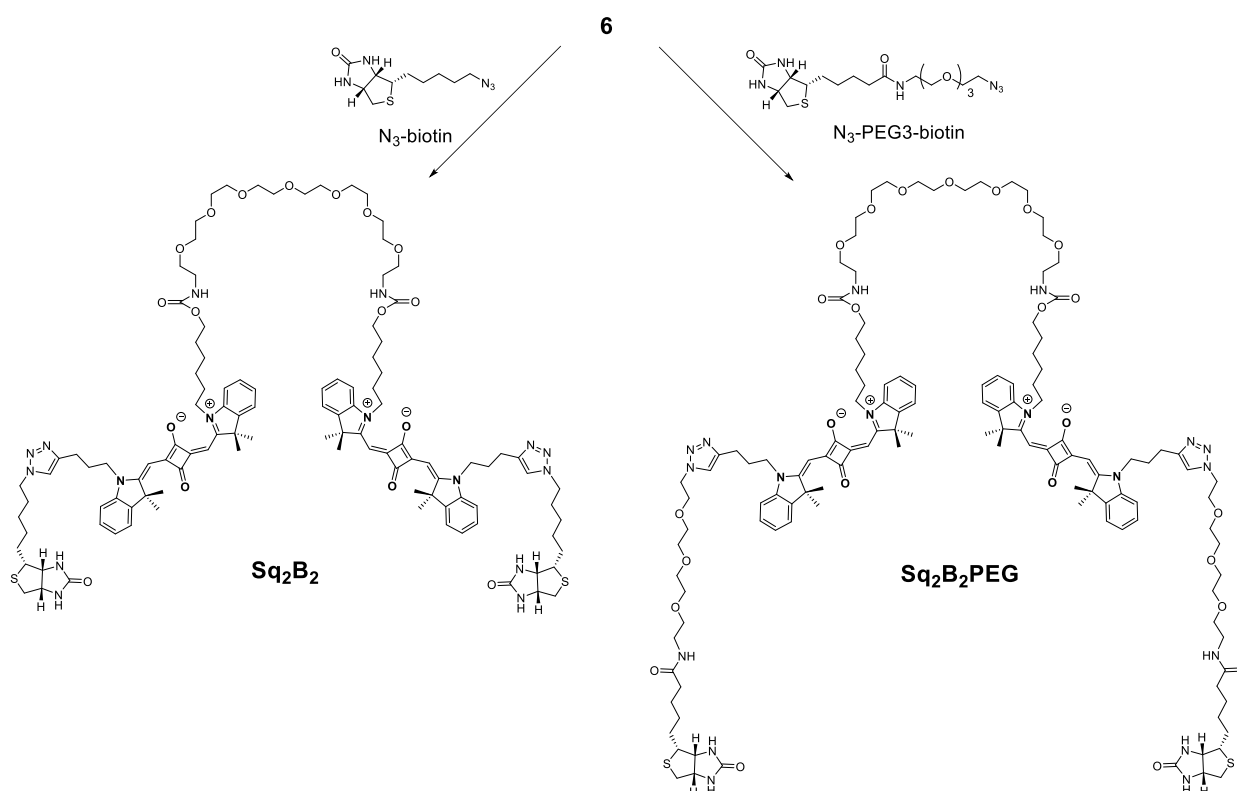
**Synthesis of compound 5.** Compound 4 (260 mg, 0.462 mmol, 1 eq) was dissolved in DCM, followed by pyridine (112 μL, 1.386 mmol, 3 eq) addition and 4-nitrophenyl chloroformate (186 mg, 0.924 mmol, 2 eq). Reaction was left stirring for 3 hours at room temperature. The solvent was evaporated and the residue was purified on column EtOAc/DCM (3/7). Yield, 61%.



$^1\text{H}$  NMR (400 MHz, Chloroform- $d$ )  $\delta$  8.21 (d, 2H), 7.32-7.26 (m, 6H), 7.1-7 (m, 3H), 6.9 (d, 2H), 5.89 (s, 2H), 4.22 (t, 2H), 4.07-4.04 (m, 4H), 2.3 (m, 2H), 2.09 (s, 1H), 1.99 (m, 2H), 1.8 (m, 4H), 1.77 (s, 12H), 1.55 (m, 4H), 1.44 (m, 4H), (d,  $J = 3.9$  Hz, 27H), 1.58 – 1.49 (m, 3H), 1.45 (dh,  $J = 9.0, 2.3, 1.8$  Hz, 6H).  $^{13}\text{C}$  NMR (100 MHz, Chloroform- $d$ )  $\delta$  127.88, 125.29, 123.82, 109.31 86.65, 69.28, 49.42, 28.43, 27.18, 27.11, 26.99, 25.67, 25.57, 16.17. HRMS (ES+) Calc. for  $\text{C}_{44}\text{H}_{45}\text{N}_3\text{O}_7$   $[\text{M}+\text{H}]^+$  728.3258, found 728.3327.

**Synthesis of 6.** Compound **5** (60 mg, 2eq) and PEG8-diamine (15 mg, 1 eq) were dissolved in DMF and DIEA (26 mg, 5 eq) was added. The mixture was left to stir overnight at  $60^\circ\text{C}$ . The solvent was removed under vacuum and the crude was dissolved in DCM. The solution was washed with concentrated solution of  $\text{NaHCO}_3$ , dried over  $\text{MgSO}_4$  and concentrated. The residue was purified by column chromatography with DCM/MeOH (97/3) as eluent. Yield 21%.

$^1\text{H}$  NMR (500 MHz, Chloroform- $d$ )  $\delta$  7.29-7.27 (m, 8H), 7.09-7.02 (m, 6H), 6.91 (d, 2H), 5.88 (s, 4H), 3.98 (m, 12H), 3.57 (t, 28H), 3.29 (t, 4H), 2.29 (m, 4H), 1.98 (m, 6H), 1.75 (s, 24H), 1.54 (m, 4H), 1.36 (m, 8H).  $^{13}\text{C}$  NMR (126 MHz, Chloroform- $d$ )  $\delta$  180.26, 179.38, 170.43, 169.69, 156.83, 142.36, 127.85, 123.88, 122.36, 109.46, 86.68, 70.29, 64.63, 49.43, 43.67, 42.24, 40.78, 28.94, 27.2, 27.06, 26.76, 25.77, 16.17. HRMS (ES+) Calc. for  $\text{C}_{92}\text{H}_{116}\text{N}_6\text{O}_{15}$   $[\text{M}+\text{H}]^+$  1544.8499, found 1545.8536



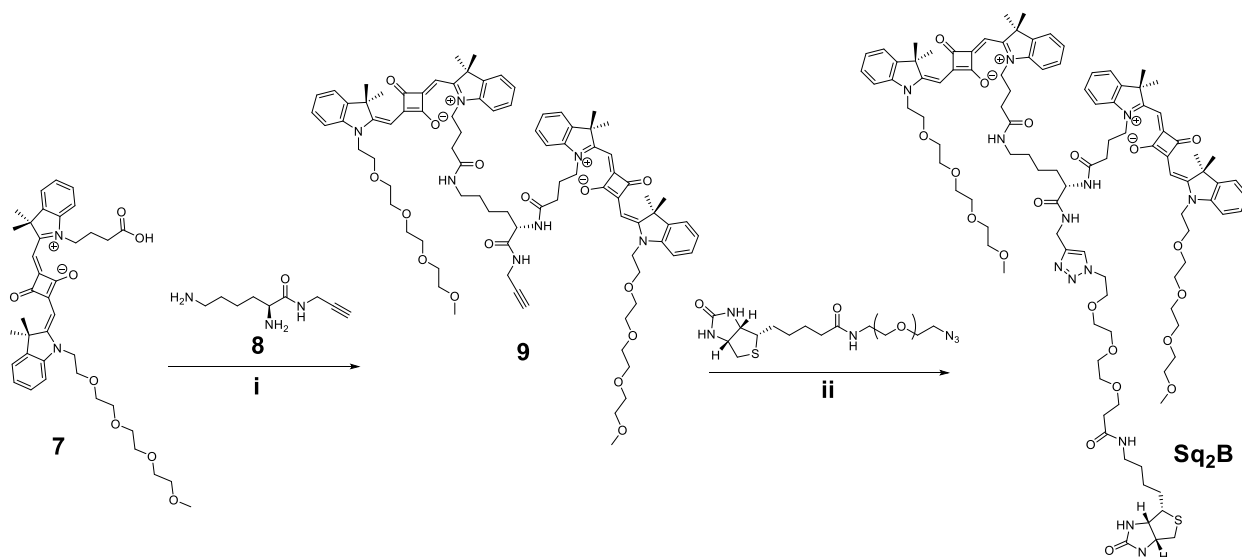
**Scheme 2.** Synthesis of  $\text{Sq}_2\text{B}_2$  and  $\text{Sq}_2\text{B}_2\text{PEG}$ .  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , sodium ascorbate, DMF/water (3/1),  $60^\circ\text{C}$ , 18h

#### General protocol for synthesis of $\text{Sq}_2\text{B}_2$ and $\text{Sq}_2\text{B}_2\text{PEG}$ .

**6** (1eq) and  $\text{N}_3\text{-biotin}$  (2.4 eq) or  $\text{N}_3\text{-PEG3-biotin}$  (2.4 eq) were dissolved in a mixture DMF/water (3/1).  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$  and sodium ascorbate were dissolved in an eppendorf with water and the solution was mixed until the mixture turn yellow. The content of the eppendorf was then added to the mixture and the reaction was stirred at  $60^\circ\text{C}$  overnight. The solvent was then removed under vacuum, the crude product was dissolved in DCM, washed with 0.05 M EDTA solution, dried over  $\text{MgSO}_4$  and concentrated. The crude was purified by column chromatography in the gradient of eluent DCM/MeOH (9/1 to 7/3) to yield blue syrup (44% for  $\text{Sq}_2\text{B}_2$ , 70% for  $\text{Sq}_2\text{B}_2\text{PEG}$ )

**Sq<sub>2</sub>B<sub>2</sub>** <sup>1</sup>H NMR (400 MHz, Chloroform-d) δ 7.54 (s, 2H), 7.28 (dd, J = 7.8, 6.6 Hz, 5H), 7.23 (dd, J = 7.7, 1.3 Hz, 3H), 7.07 (q, J = 7.1 Hz, 4H), 6.94 (dd, J = 16.5, 7.9 Hz, 4H), 5.98 (s, 2H), 5.87 (d, J = 11.3 Hz, 4H), 5.48 (d, J = 8.7 Hz, 3H), 4.45 – 4.35 (m, 2H), 4.32 (t, J = 7.2 Hz, 4H), 4.19 (dd, J = 8.5, 4.5 Hz, 2H), 4.04 – 3.84 (m, 14H), 3.55 (d, J = 7.4 Hz, 27H), 3.47 (t, J = 5.2 Hz, 4H), 3.41 (s, 1H), 3.28 (q, J = 5.6 Hz, 4H), 3.06 – 2.98 (m, 2H), 2.83 – 2.73 (m, 7H), 2.64 (d, J = 12.7 Hz, 2H), 2.22 – 2.04 (m, 14H), 1.84 (q, J = 7.3 Hz, 4H), 1.70 (d, J = 2.4 Hz, 26H), 1.60 – 1.48 (m, 7H), 1.35 (ddt, J = 24.5, 16.7, 7.2 Hz, 18H). <sup>13</sup>C NMR (101 MHz, Chloroform-d) δ 178.92, 170.15, 127.92, 127.85, 123.82, 122.32, 122.24, 121.55, 109.69, 109.46, 86.59, 70.51, 70.48, 70.23, 70.16, 64.62, 62.01, 60.14, 55.61, 50.12, 49.30, 43.62, 42.89, 40.76, 40.51, 29.93, 28.91, 28.47, 28.38, 27.10, 27.05, 26.99, 26.70, 26.41, 26.29, 25.73, 22.92. HRMS (ES+) Calc. for C<sub>112</sub>H<sub>150</sub>N<sub>16</sub>O<sub>14</sub>S<sub>2</sub> [M+H]<sup>+</sup> 2056.0806, found 2056.0832

**Sq<sub>2</sub>B<sub>2</sub>PEG**: <sup>1</sup>H NMR (500 MHz, Chloroform-d) δ 7.92 (s, 2H), 7.46 (t, 4H), 7.39-7.35 (q, 4H), 7.27-7.19 (m, 8H), 5.99 (d, 4H), 4.594 (t, 4H), 4.035 (t, 4H), 3.62-3.59 (m, 34H), 3.56-3.48 (m, 16H), 3.34 (p, 16H), 3.28 (t, 4H), 3.18 (p, 4H), 2.93-2.29 (m, 6H), 2.69 (d, 2H), 2.25-2.18 (m, 8H), 1.96 (s, 64H), 1.85-1.83 (m, 4H), 1.76 (d, 24H), 1.73-1.58 (m, 10H), 1.45-1.38 (p, 5H). <sup>13</sup>C NMR (126 MHz, Chloroform-d) δ 184.57, 180.42, 177.20, 176.56, 176.11, 172.51, 172.00, 166.12, 159.34, 143.72, 143.42, 143.26, 129.42, 129.41, 125.65, 125.50, 124.44, 123.51, 111.63, 111.46, 87.36, 87.22, 71.64, 71.56, 71.52, 71.42, 71.37, 71.28, 70.69, 70.52, 65.87, 63.47, 61.73, 57.11, 51.46, 50.74, 50.62, 44.79, 44.12, 41.73, 41.19, 40.46, 36.85, 30.88, 30.23, 29.90, 29.84, 29.63, 28.31, 27.84, 27.73, 27.57, 27.47, 26.96, 23.80, 23.71, 23.67. HRMS (ES+) Calc. for C<sub>128</sub>H<sub>180</sub>N<sub>18</sub>O<sub>25</sub>S<sub>2</sub> [M+H]<sup>+</sup> 2433.2808, found 2433.2624



**Scheme 3.** Synthesis of Sq<sub>2</sub>B. i) HATU, DIEA, DMF, r.t., 1h (50%); ii) CuSO<sub>4</sub>·5H<sub>2</sub>O, sodium ascorbate, DMF/water (3/1), 60°C, 18h (60%).

Compound **7** and **8** were synthesized according to a published protocols.<sup>35,92</sup>

**Synthesis of compound 9.** To a solution of **7** (40 mg, 0.06 mmol, 2 eq) and **8** (12 mg, 0.03 mmol, 1 eq) in DMF (3 mL) was added HATU (54 mg, 0.14 mmol, 1.2 eq) followed by DIEA (62 μL, 0.36 mmol, 12 eq). After 1 h the solvents were evaporated and the crude was first purified by column chromatography on silica gel (DCM/MeOH: 8/2). Yield 50%.

<sup>1</sup>H NMR (500 MHz, Chloroform-d) δ 10.96 (s, 1H), 7.96 (s, 1H), 7.53 (s, 2H), 7.32 – 7.26 (m, 0H), 7.28 – 7.20 (m, 5H), 7.21 (d, J = 1.6 Hz, 2H), 7.22 – 7.16 (m, 1H), 7.09 (d, J = 7.9 Hz, 1H), 7.10 – 7.00 (m, 7H), 6.93 (d, J = 14.6 Hz, 1H), 5.93 (d, J = 8.2 Hz, 2H), 5.86 (s, 2H), 4.48 (q, J = 7.4 Hz, 1H), 4.17 (s, 5H), 4.04 (s, 3H), 3.92 (ddd, J = 8.8, 5.4, 2.5 Hz, 2H), 3.77 (td, J = 5.9, 1.7 Hz, 4H), 3.62 (p, J = 6.7 Hz, 2H), 3.55 – 3.41 (m, 22H), 3.28 (s, 5H), 3.21 (dq, J = 13.9, 6.8 Hz, 2H), 3.04 (q, J = 7.4 Hz, 2H), 2.42 (t, J = 6.7 Hz, 2H), 2.37 (t, J = 6.9 Hz, 2H), 2.11 – 1.96 (m, 2H), 1.79 (t, J = 7.2 Hz, 1H), 1.66 (d, J = 12.7 Hz, 25H), 1.51 (dq, J = 21.9, 7.7, 7.3 Hz, 2H), 1.42 – 1.27 (m, 17H), 1.20 (d, J = 15.5 Hz, 6H), 0.80 (q, J = 10.5, 8.6 Hz, 1H). <sup>13</sup>C

NMR (126 MHz, Chloroform-d)  $\delta$  172.32, 142.94, 142.06, 128.09, 127.72, 124.12, 123.81, 123.73, 122.25, 122.02, 110.24, 110.09, 109.85, 79.70, 71.84, 71.82, 71.17, 71.05, 70.54, 70.39, 70.37, 67.65, 59.01, 53.58, 49.49, 49.43, 49.23, 43.91, 43.17, 42.95, 41.96, 38.76, 32.83, 32.42, 30.31, 29.69, 29.03, 28.61, 27.12, 26.95, 22.80, 22.61, 22.52, 18.57, 17.45, 11.84. Calc. for  $C_{87}H_{109}N_7O_{15}$   $[M]^+$  1492.7982, found 1491.7989.

Synthesis of **Sq<sub>2</sub>B** and **SqB** was done following general protocol for synthesis of **Sq<sub>2</sub>B<sub>2</sub>** and **Sq<sub>2</sub>B<sub>2</sub>PEG**.

**Sq<sub>2</sub>B**:  $^1H$  NMR (400 MHz, Chloroform-d)  $\delta$  10.04 – 9.99 (m, 15H), 8.04 (d,  $J$  = 6.8 Hz, 1H), 7.94 (s, 1H), 7.72 (d,  $J$  = 7.3 Hz, 1H), 7.52 (d,  $J$  = 6.0 Hz, 1H), 7.40 (s, 1H), 7.32 (td,  $J$  = 11.7, 9.6, 5.2 Hz, 9H), 7.18 (qd,  $J$  = 8.9, 7.7, 3.9 Hz, 9H), 6.52 (s, 3H), 6.07 (s, 1H), 6.02 (s, 1H), 5.94 (d,  $J$  = 14.5 Hz, 2H), 4.62 – 4.55 (m, 1H), 4.59 – 4.47 (m, 4H), 4.39 (dd,  $J$  = 8.0, 4.8 Hz, 2H), 4.25 (d,  $J$  = 5.6 Hz, 4H), 4.07 (t,  $J$  = 8.1 Hz, 4H), 3.95 (s, 1H), 3.84 (d,  $J$  = 5.4 Hz, 6H), 3.53 (ddt,  $J$  = 10.0, 6.6, 4.1 Hz, 35H), 3.46 – 3.38 (m, 2H), 3.34 (s, 6H), 3.18 (dq,  $J$  = 18.7, 6.7 Hz, 3H), 2.90 (dd,  $J$  = 13.2, 4.8 Hz, 1H), 2.76 (d,  $J$  = 13.0 Hz, 1H), 2.52 (t,  $J$  = 7.3 Hz, 4H), 2.24 (t,  $J$  = 7.5 Hz, 2H), 2.08 (q,  $J$  = 8.7, 7.8 Hz, 5H), 1.79 (s, 1H), 1.70 (s, 17H), 1.62 (t,  $J$  = 7.4 Hz, 1H), 1.57 – 1.49 (m, 1H), 1.46 – 1.38 (m, 1H), 1.37 (s, 3H).  $^{13}C$  NMR (126 MHz, Chloroform-d)  $\delta$  175.53, 172.86, 142.02, 141.86, 141.58, 128.04, 125.35, 125.19, 124.90, 122.20, 122.09, 114.23, 111.19, 111.11, 110.90, 110.65, 71.79, 71.00, 70.98, 70.51, 70.49, 70.41, 70.37, 70.24, 70.00, 69.27, 68.80, 67.75, 62.49, 60.92, 58.91, 55.41, 51.03, 49.90, 49.85, 49.67, 49.65, 44.43, 43.38, 40.30, 39.67, 39.43, 35.16, 34.04, 32.44, 32.10, 31.20, 28.23, 27.92, 27.80, 26.54, 26.47, 26.43, 25.40, 23.35, 22.71. HRMS (ES+) Calc. for  $C_{105}H_{141}N_{13}O_{20}S$   $[M+H+Na]^+$  1960.0137, found 1960.0110

**SqB**:  $^1H$  NMR (400 MHz, Methanol-d<sub>4</sub>)  $\delta$  8.16 – 8.08 (m, 2H), 7.78 (s, 1H), 7.35 – 7.27 (m, 4H), 7.25 – 7.15 (m, 3H), 7.08 (ddt,  $J$  = 8.8, 7.8, 1.4 Hz, 3H), 5.92 (s, 1H), 5.83 (s, 1H), 4.50 – 4.44 (m, 2H), 4.37 – 4.33 (m, 1H), 4.28 – 4.22 (m, 4H), 4.16 (dd,  $J$  = 7.9, 4.4 Hz, 1H), 4.08 (td,  $J$  = 6.8, 3.0 Hz, 2H), 3.84 – 3.75 (m, 4H), 3.61 – 3.53 (m, 2H), 3.50 – 3.43 (m, 6H), 3.36 (t,  $J$  = 5.4 Hz, 2H), 3.21 (dt,  $J$  = 3.3, 1.6 Hz, 6H), 2.82 – 2.72 (m, 3H), 2.08 (dt,  $J$  = 14.7, 7.3 Hz, 4H), 1.63 (d,  $J$  = 4.2 Hz, 12H), 1.49 (dd,  $J$  = 14.7, 7.2 Hz, 2H), 1.30 (q,  $J$  = 7.5 Hz, 2H).  $^{13}C$  NMR (101 MHz, Methanol-d<sub>4</sub>)  $\delta$  183.05, 179.81 – 173.30 (m), 171.04 (d,  $J$  = 87.9 Hz), 164.63, 155.68, 152.55, 146.22, 146.14 – 141.27 (m), 127.76 (d,  $J$  = 23.7 Hz), 124.87, 123.97 (d,  $J$  = 5.0 Hz), 122.88, 110.38 (d,  $J$  = 84.5 Hz), 85.77, 73.13 – 63.93 (m), 61.08 (d,  $J$  = 175.8 Hz), 55.59, 50.76 – 48.90 (m), 43.22 (d,  $J$  = 118.9 Hz), 39.30 (d,  $J$  = 72.7 Hz), 35.33, 28.23 (d,  $J$  = 27.2 Hz), 26.31, 25.73 (d,  $J$  = 60.7 Hz), 22.29. HRMS (ES+) Calc. for  $C_{57}H_{78}N_8O_{11}S$   $[M+Na]^+$  1105.5511, found 1105.5420.

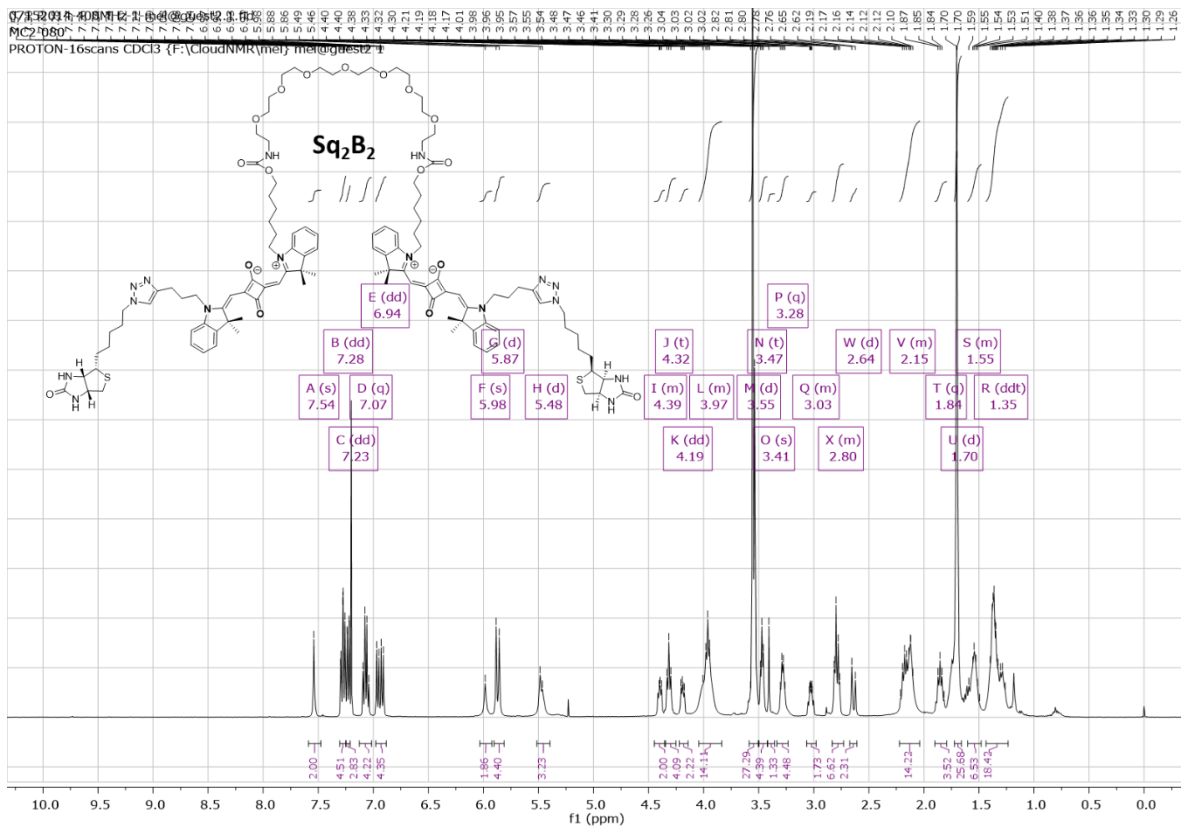


Fig. S30. <sup>1</sup>H NMR spectrum of Sq<sub>2</sub>B<sub>2</sub>

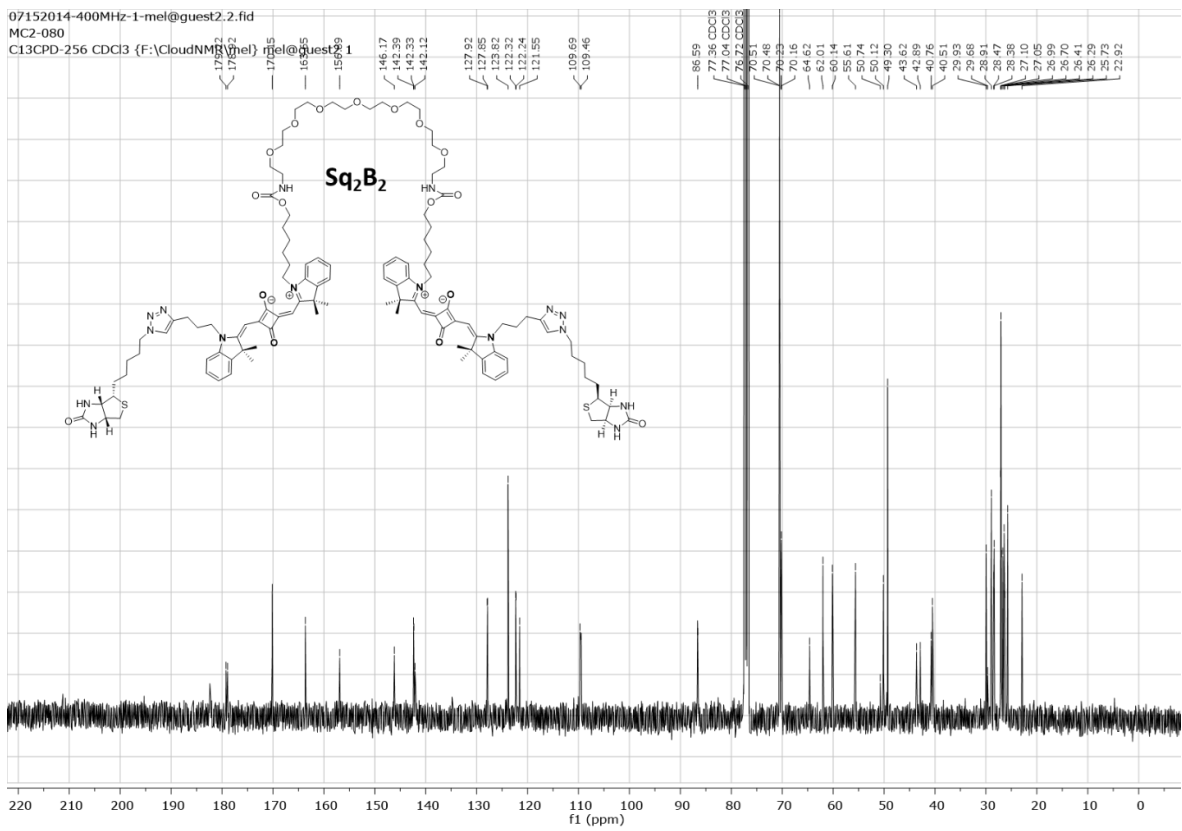


Fig. S31. <sup>13</sup>C NMR spectrum of Sq<sub>2</sub>B<sub>2</sub>

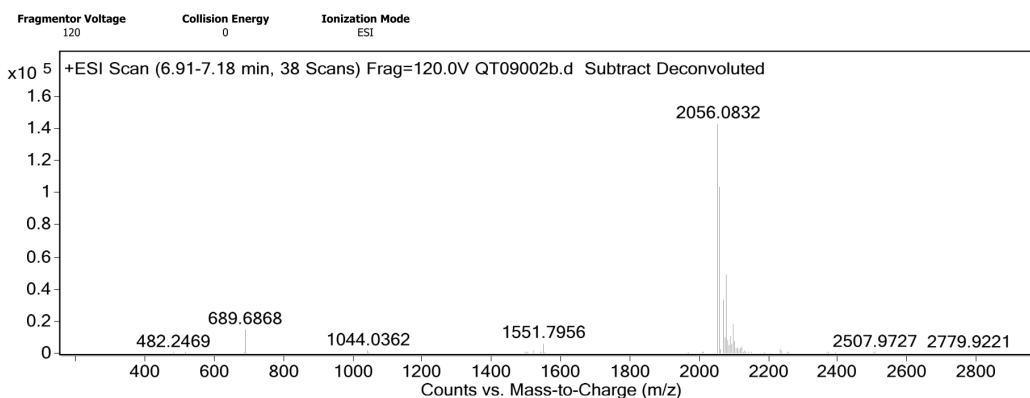
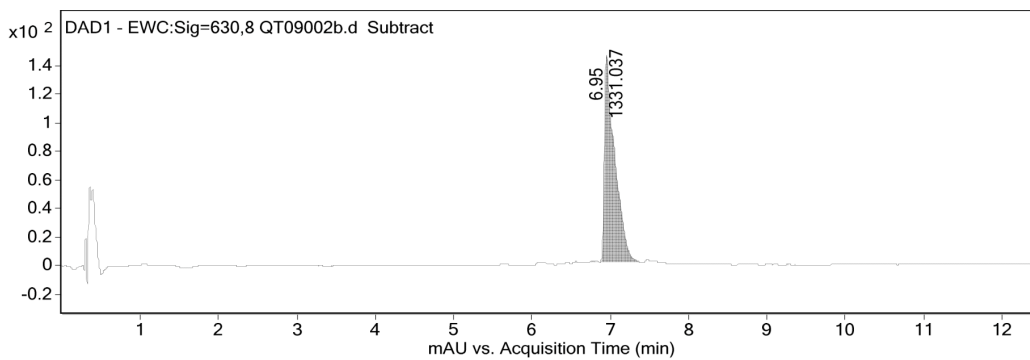


Fig. S32. HP-LC chromatogram and HR-MS spectrum of Sq<sub>2</sub>B<sub>2</sub>

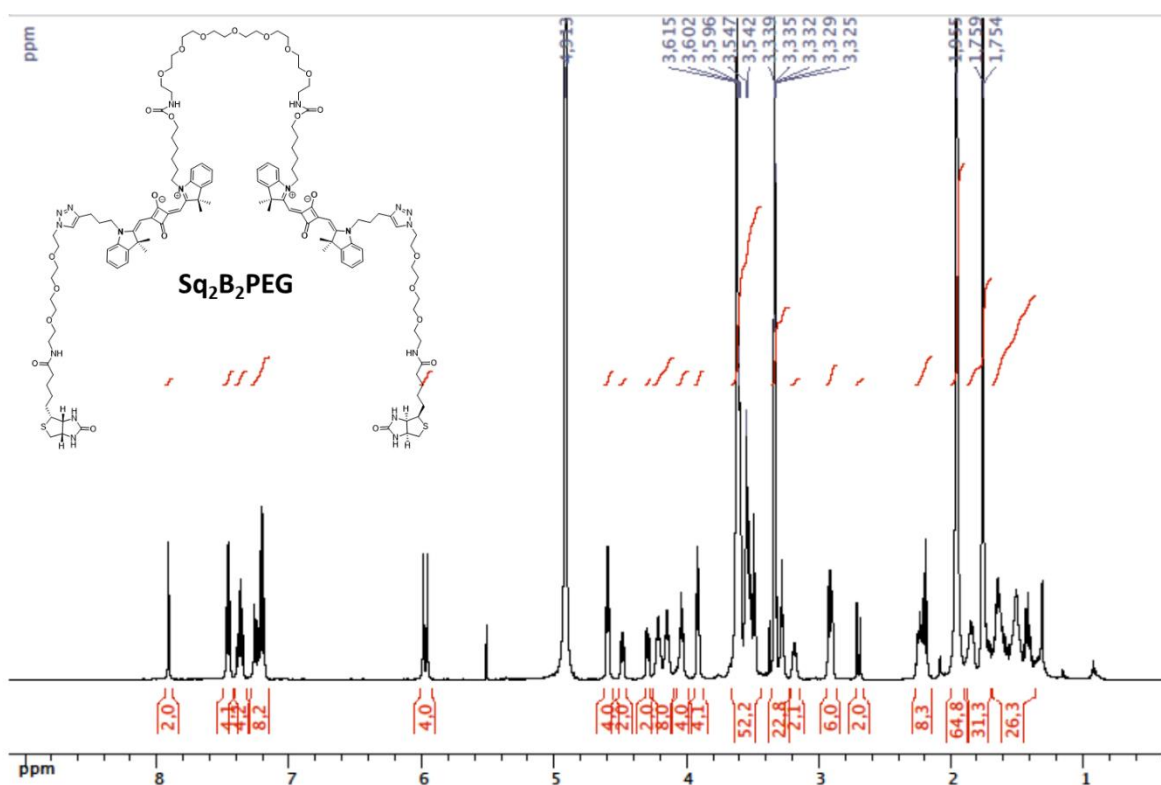


Fig. S33. <sup>1</sup>H NMR spectrum of Sq<sub>2</sub>B<sub>2</sub>PEG

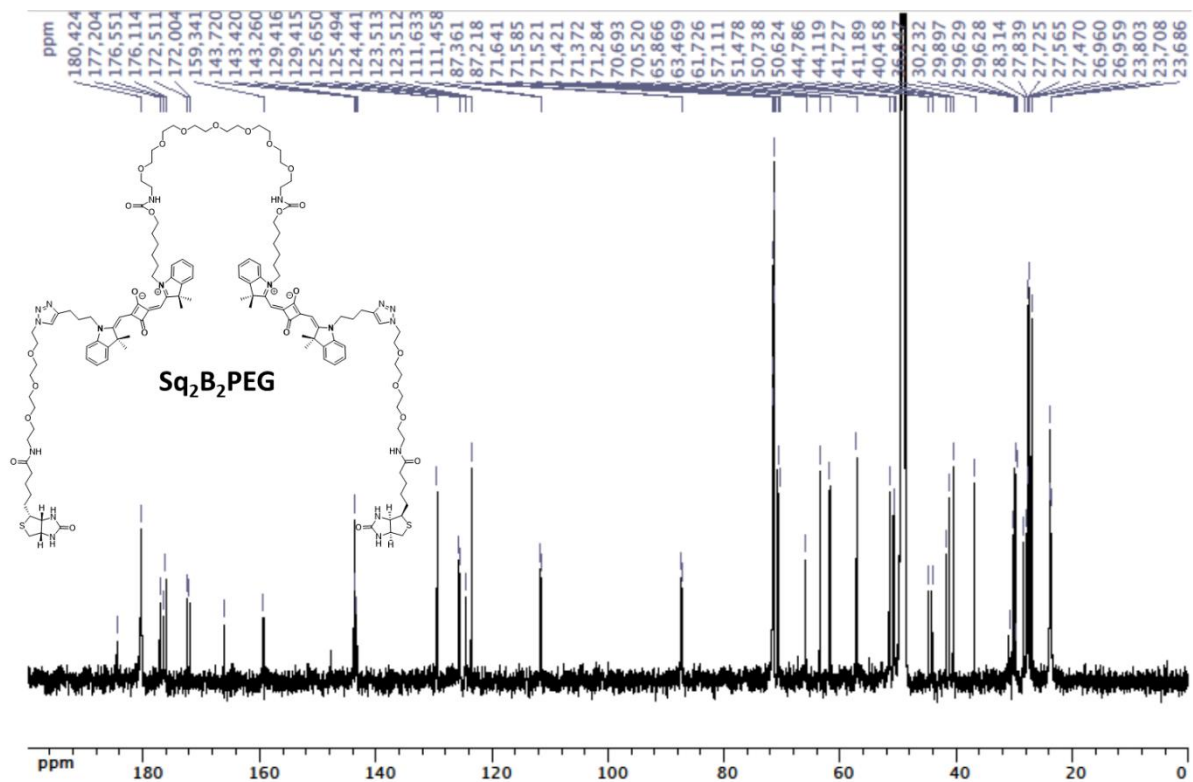


Fig. S34.  $^{13}C$  NMR spectrum of  $Sq_2B_2PEG$

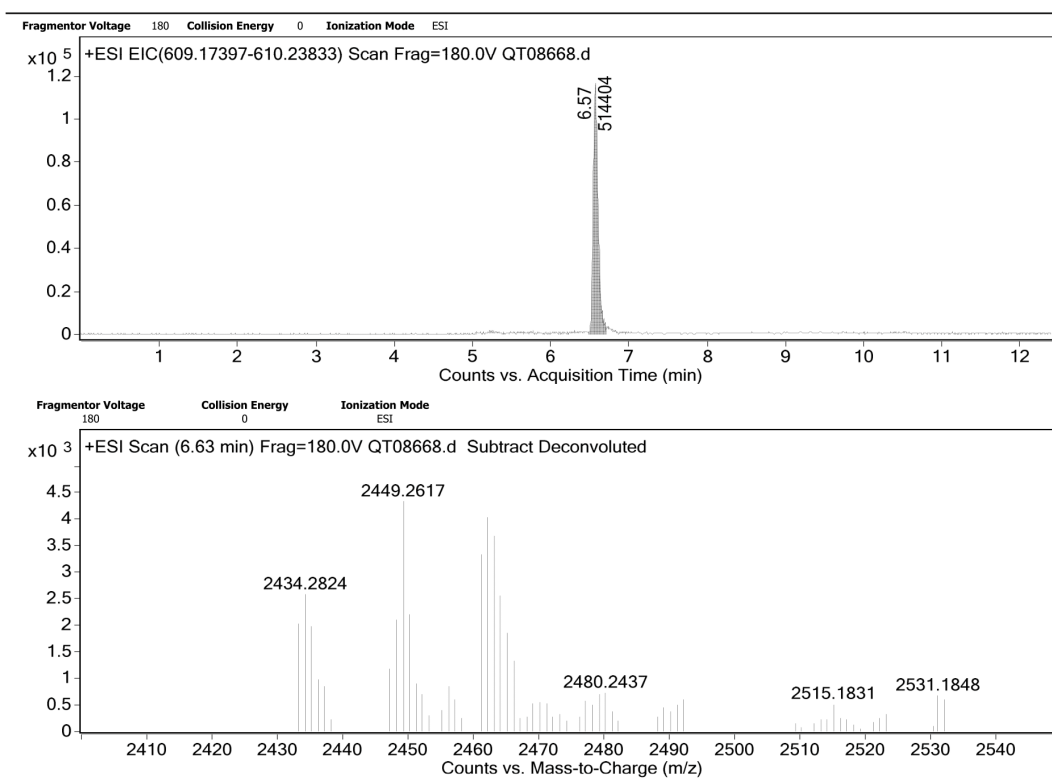


Fig. S35. . HP-LC chromatogram and HR-MS spectrum of  $Sq_2B_2PEG$



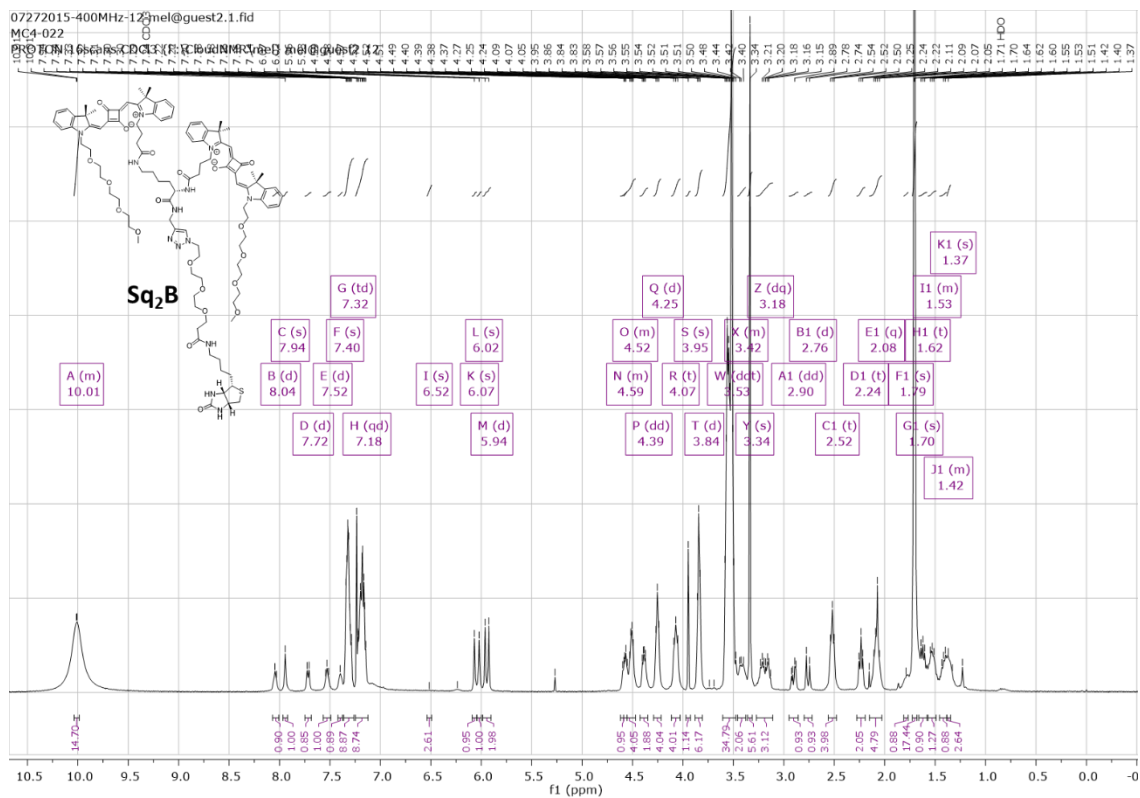


Fig. S36. <sup>1</sup>H NMR spectrum of Sq<sub>2</sub>B

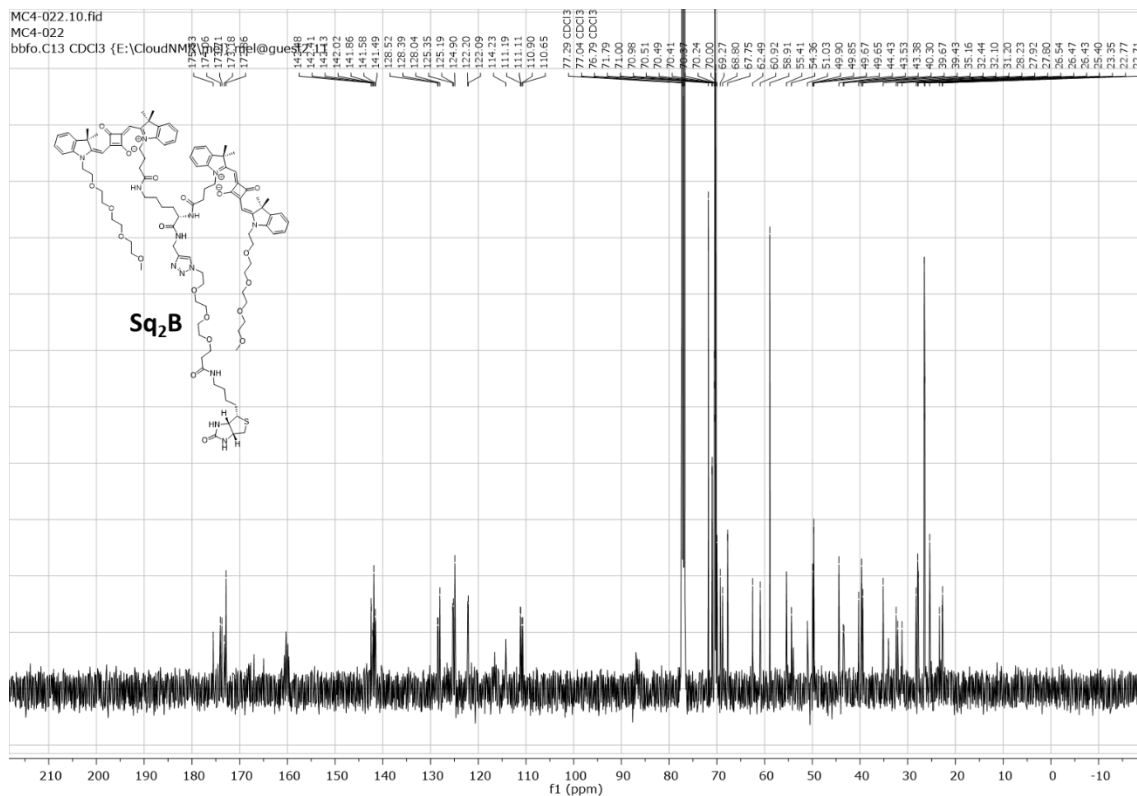


Fig. S37. <sup>13</sup>C NMR spectrum of Sq<sub>2</sub>B

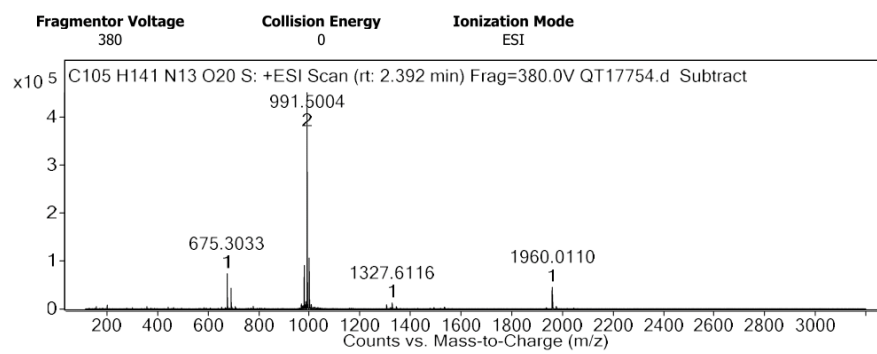


Fig. S38. HR-MS spectrum of Sq<sub>2</sub>B

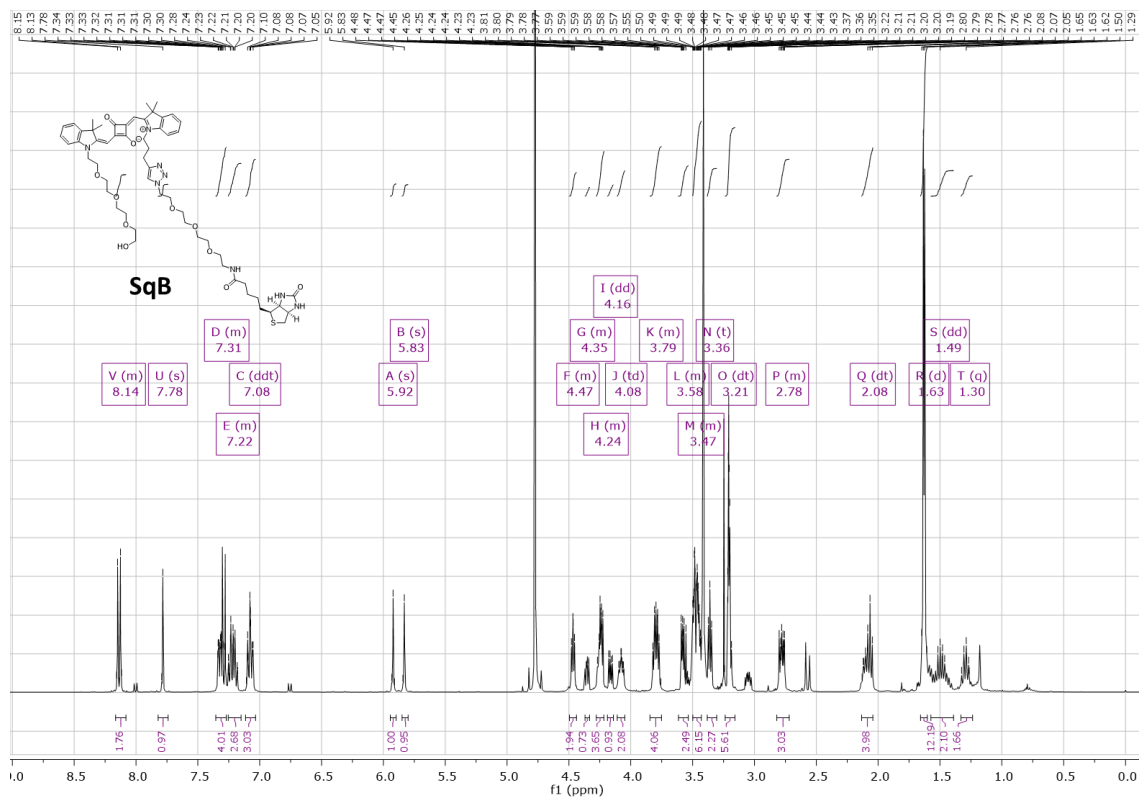


Fig. S39. <sup>1</sup>H NMR spectrum of SqB

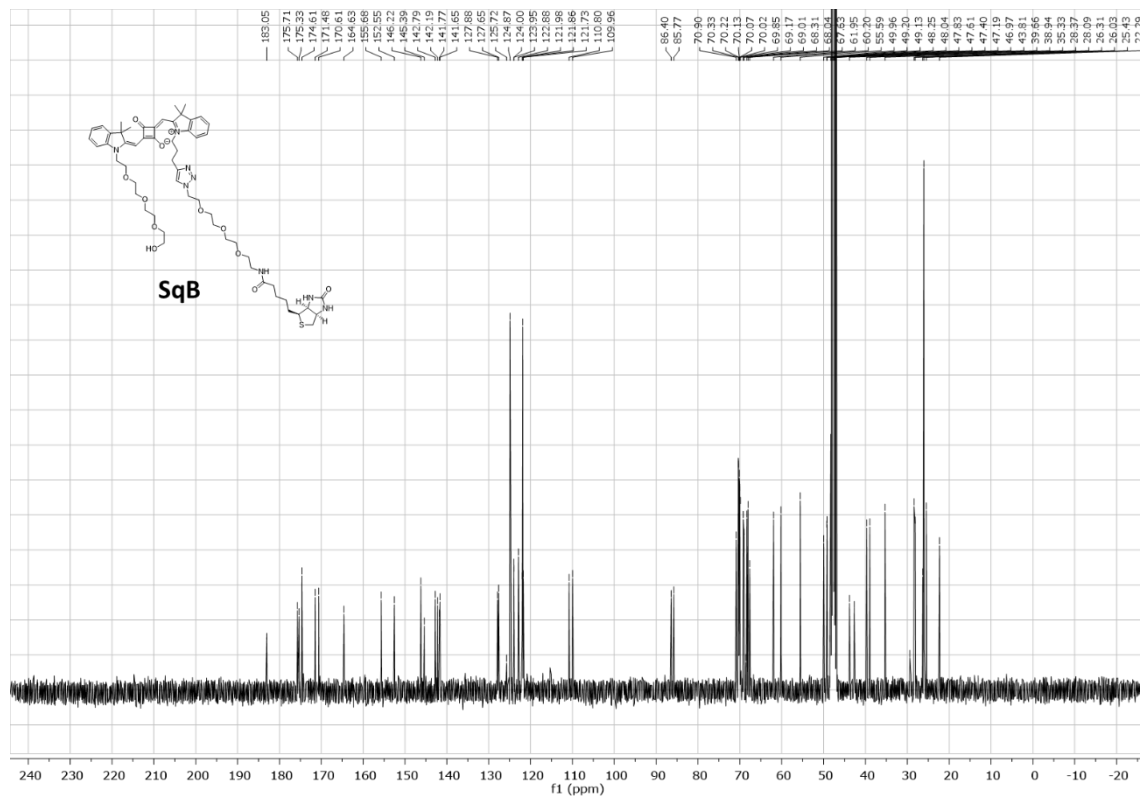


Fig. 40. <sup>13</sup>C NMR spectrum of SqB

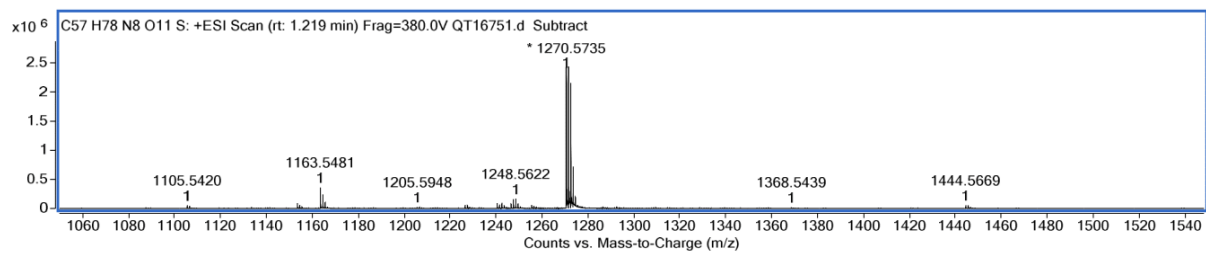


Fig. 41. HR-MS spectrum of SqB

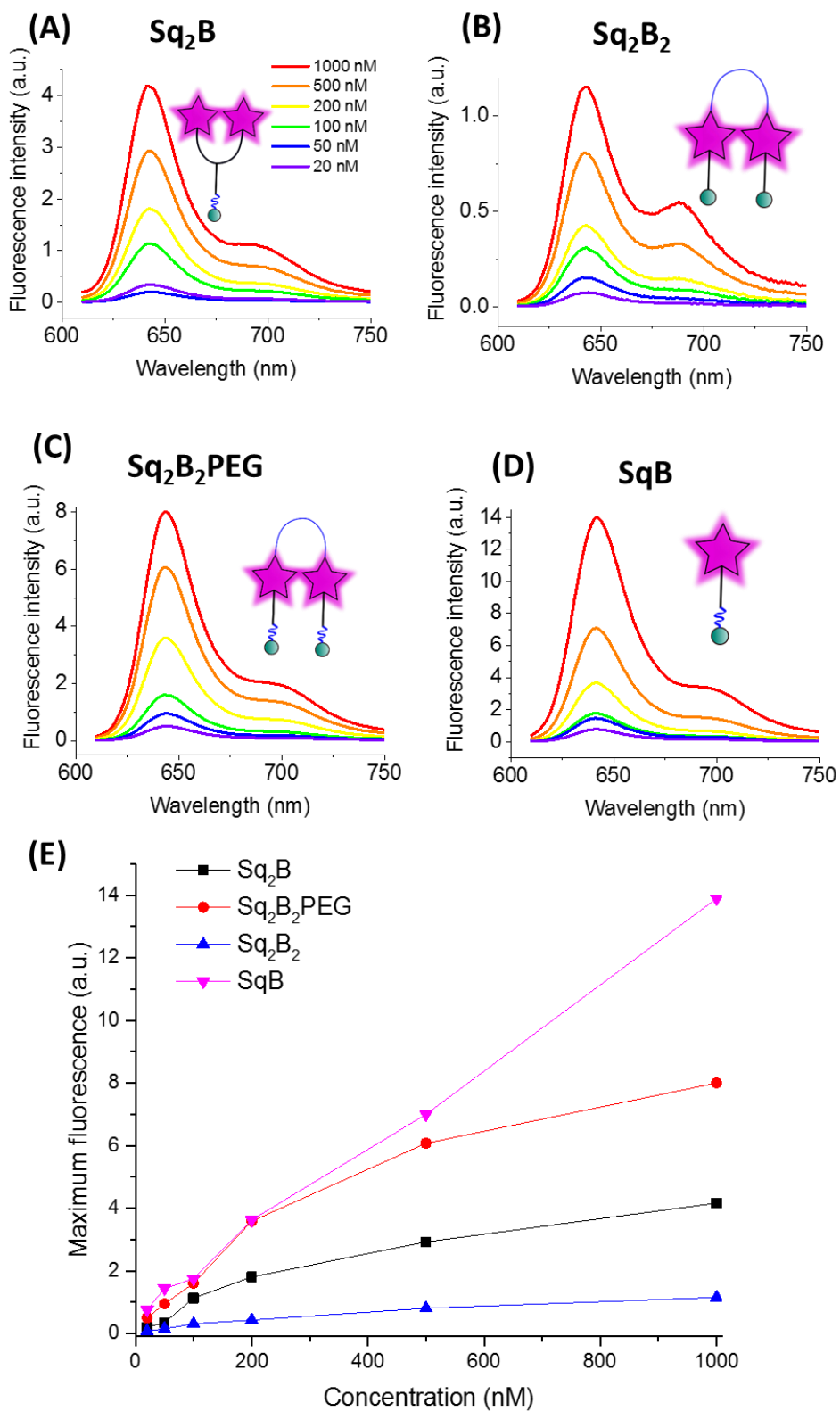


Fig. S42. Fluorescence spectra of Sq<sub>2</sub>B (A), Sq<sub>2</sub>B<sub>2</sub> (B), Sq<sub>2</sub>B<sub>2</sub>PEG (C) and SqB (D) at different concentrations in water. (E) Plot of fluorescence maxima from A-D of Sq<sub>2</sub>B, Sq<sub>2</sub>B<sub>2</sub>, Sq<sub>2</sub>B<sub>2</sub>PEG and SqB.

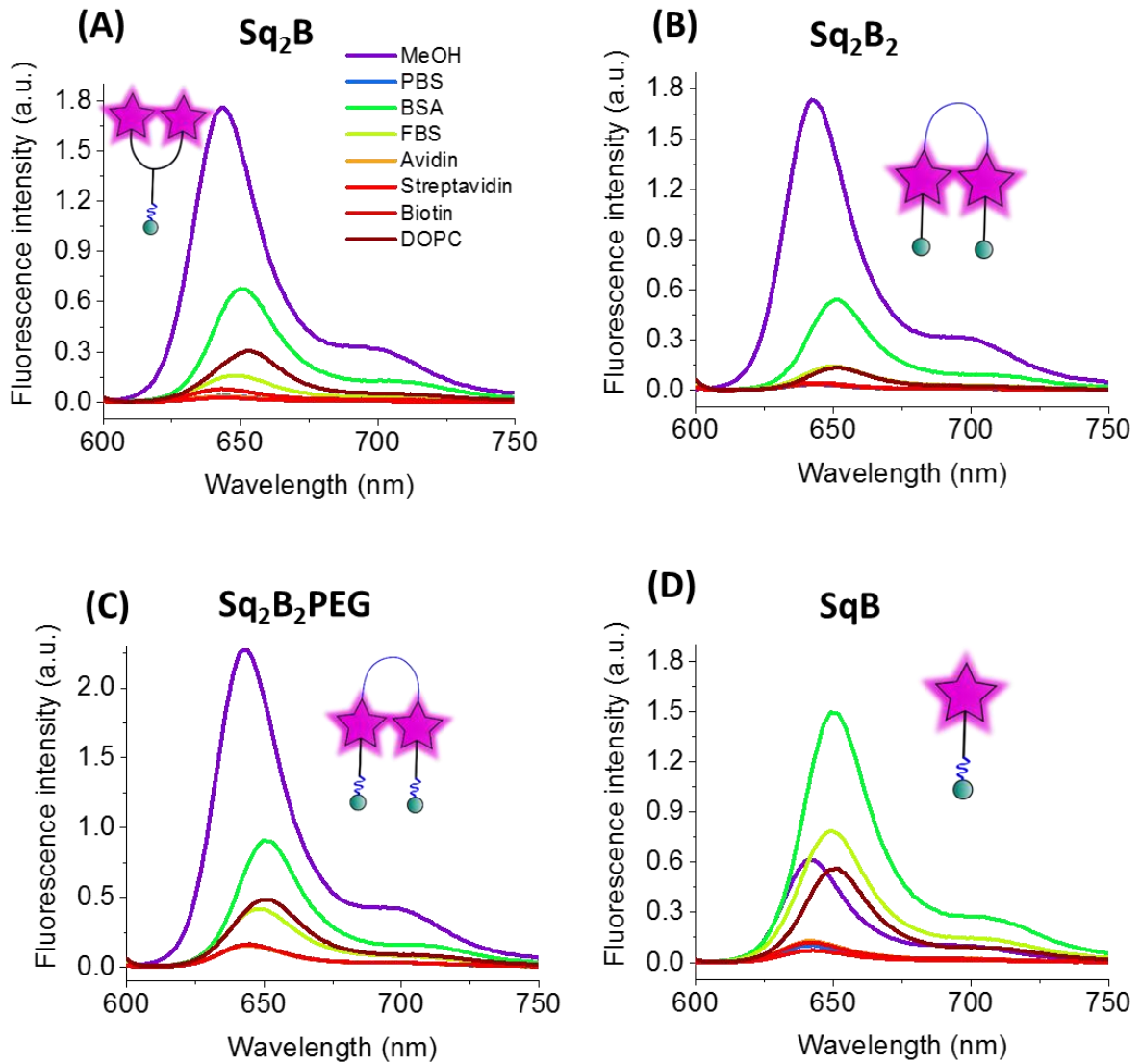


Fig. S43. Fluorescence spectra of 0.2  $\mu$ M of Sq<sub>2</sub>B (A), Sq<sub>2</sub>B<sub>2</sub> (B), Sq<sub>2</sub>B<sub>2</sub>PEG (C) and SqB (D) in MeOH, PBS, in the presence of BSA (0.1 mg/mL), FBS (0.1 mg/mL), avidin (100 nM), streptavidin (100 nM), biotin (100  $\mu$ M) or DOPC (20  $\mu$ M).



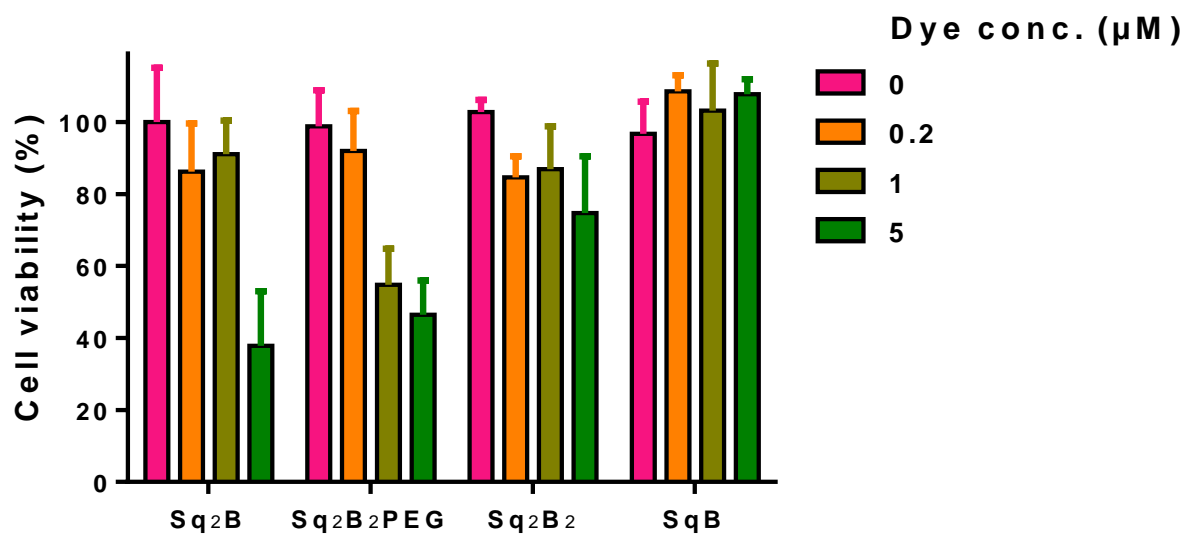


Fig. S44. Cytotoxicity of the probes after 24h incubation with KB cells determined by the MTT assay.

## 5. General conclusion and perspectives

Dimerization is common natural process in protein biology<sup>37</sup> enabling to unleash new and expand available pallet of functionalities, but there are few reported examples of dimerization approach in chemical biology for fluorescence imaging. Inspired by nature, we applied dimerization approach to chemical biology to expand toolbox of dimerized fluorogenic probes. Thus, within this PhD work we developed bright fluorogenic probes for selective imaging of live-cell biological events. To accomplish this goal, we applied intramolecular self-aggregation strategy in probe design, in particular homodimerization of fluorophores *via* short linkage and tethered targeted ligand.

First, we developed a family of fluorogens based on rhodamine scaffolds, spanning their fluorescence from green to far-red region, named Gemini (Gemini-505, Gemini-561, Gemini-647). We demonstrated that Geminis are bright fluorogens that did not produce unspecific switch-on fluorescence in complex biological media, thus making them promising candidates for SELEX-based selection. In collaboration with RYCKELYNCK group, we found cognate RNA aptamers for Gemini fluorogens. With Gemini-561 fluorogen, a selective RNA aptamer (o-Coral) was isolated using microfluidics-assisted SELEX. We fully characterized properties of Gemini-561/o-Coral couple and validated its selectivity *in vitro*. o-Coral is genetically encodable and can label small RNAs (U6, 5S) and mRNA (*egfp*). Gemini-561/o-Coral fluoromodule enabled imaging o-Coral-tagged RNAs in several cell lines showing heterogeneous transcriptional activity as well as heterogeneous RNA expression levels. Directed evolution of RNA aptamers for Gemini-505 and Gemini-647 2.0 is currently ongoing. Though expanding the toolbox of fluorogenic aptamers with orthogonal emission colors will enable to label and image several RNA targets simultaneously, the potential of technology is beyond that. For example, the intertwined dimerized structure of o-Coral may also be advantageous for the future engineering of aptamers for metabolite biosensing.<sup>92</sup> Additionally, Gemini-561/o-Coral can be coupled with CRISPR/Cas system by inserting o-Coral aptamer into single guide RNA scaffold. This approach will enable imaging dynamics of genes, promoters, enhancers and various genomic elements during cell development.<sup>49,146</sup> Another potential of Gemini-561/o-Coral technology can be explored in the field of RNA regulation. In particular, there are RNA-based regulators such as the riboswitches as part of mRNAs able to respond to small

metabolites, tRNA, secondary messengers, ions, vitamins or amino acids.<sup>147</sup> Optimization of o-Coral structure and correct insertion in RNA domain can potentially transform o-Coral as regulatory unit that will be active upon binding to Gemini-561. This approach can potentially be used for on-demand transcription/translation RNA activation/control while pinpointing the commencement of the process by mean of activated fluorescence of Gemini-561.

Secondly, we developed Gemini-584-alkyne capable of binding to dimerized TR-512 peptide. Though dimerized TR-512 peptides could only modestly activate Gemini-584-alkyne's fluorescence, we demonstrated that Gemini-584-alkyne is a promising candidate to develop evolved dimeric TR-512-peptide-based tag for protein imaging. The project is currently ongoing in collaboration with Dr. Eleonore REAL (UMR 7021, University of Strasbourg).

Lastly, we performed a systematic analysis of structure-activity relationship in a design of a dimerization probe for biotin receptors. Our study revealed that a small relatively rigid linker based on natural amino acid (L-lysine) served as a better connector in dimer design in comparison to flexible synthetic PEG8-linker. L-lysine-based dimer promoted the retention of unimolecular quenched form while preventing unspecific interaction in cellular environment. This strategy enabled us to develop Sq<sub>2</sub>B probe based on biotinylated dimerized squaraine. We proved that Sq<sub>2</sub>B is sensitive to biotin receptors (BR) providing selective discrimination of cancerous cells expressing BRs on cell surface. Additionally, Sq<sub>2</sub>B could report different BRs activity states modulated by natural ligands or culturing conditions (*e.g.* temperature). Sq<sub>2</sub>B can potentially be used to develop high-throughput screening assay to evaluate new therapeutics targeting BRs in cancer cells.<sup>132</sup>

In summary, this PhD work proved that dimerization-caused quenching is a very attractive approach in fluorogenic probe design for chemical biology. The main advantage of our approach is being modular while providing bright activated fluorescence response of probes to biological event of interest. We showed its broad applicability in the field of semisynthetic RNA/peptide-based technologies and targeted protein imaging. We envision that the further expansion of dimerized fluorogenic probes will significantly contribute to understanding the complexity of biological processes in living systems.

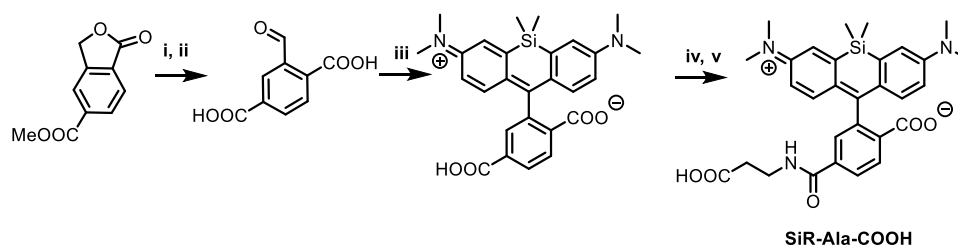
## 6. Materials and methods

### 6.1. Materials

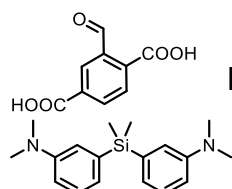
Starting materials for synthesis were purchased from Alfa Aesar, Sigma Aldrich, TCI Europe or abcr GmbH and used as received unless stated otherwise. The solvents were of analytical grade. NMR spectra were recorded on a Bruker Avance III 400 MHz and 500 MHz spectrometers. Mass spectra were obtained using an Agilent Q-TOF 6520 mass spectrometer. MilliQ-water (Millipore) was used in all experiments.

### 6.2. Synthesis

#### 6.2.1. Synthesis of Si-Rhodamine derivative

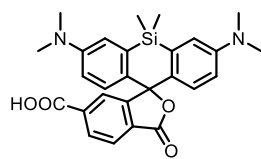


**Scheme 1.** Synthesis of SiR-Ala-COOH. (i) NBS, AIBN, DCE, reflux, 3 h; (ii) aq. NaOH, THF/H<sub>2</sub>O, reflux, 2 h (48%, after 2 steps); (iii) (N,N'-dimethyl)dimethyldiphenylsilane, CuBr<sub>2</sub>, TFE, 80°C, 24 h (7%); (iv) β-Ala-COOt-Bu, HATU, DIEA, DMF, r.t., 18 h (34%); (v) TFA, DCM, r.t., 1 h.

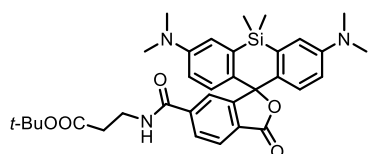


**2-Formyl-terephthalic acid.** Synthesis was followed by protocol from literature.<sup>148</sup>

(N,N'-dimethyl)dimethyldiphenylsilane was kindly provided by Dr. Bohdan Andreiuk.

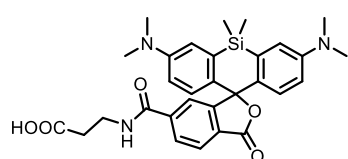


**SiR.** The diarylsilane (260 mg, 0.87 mmole, 0.5 eq), benzaldehyde (338 mg, 1.74 mmole, 1 eq), copper bromide (97 mg, 0.44 mmole, 0.25 eq) were dissolved in TFE and heated at 80°C for 24 h. After cooling to room temperature, the solvent was concentrated *in vacuo*, and dark blue residue solid was purified by column chromatography on silica gel (DCM:MeOH= 9:1) to obtain dark blue solid. Yield 7%. LCMS, NMR characterization was in accordance with literature.<sup>93</sup>



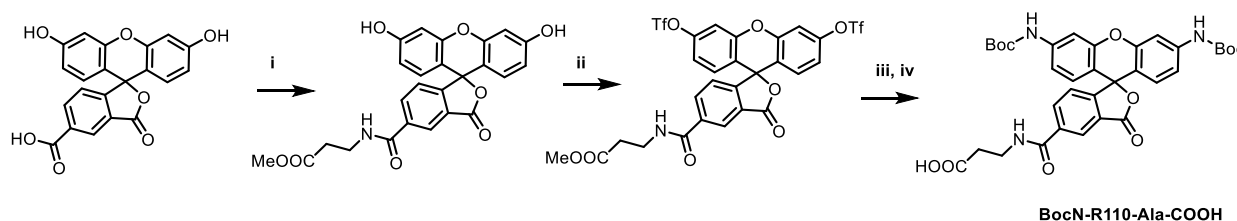
**SiR-Ala-COOt-Bu (FTK047).** To a solution of SiR (51 mg, 0.11 mmole, 1 eq) and β-alanine *t*-butyl ester hydrochloride (32 mg, 0.13 mmole, 1.2 eq) in DMF (15 mL) was added HATU (72.7 mg, 0.14 mmole, 1.3 eq) followed by DIEA (77 μL, 0.33 mmole, 3 eq). After 18 h the solvent was evaporated, dissolved in DCM, washed with water

and brine. Organic phase was collected, dried over  $\text{MgSO}_4$  and concentrated. The crude was purified on chromatography column on silica gel ( $\text{DCM}:\text{MeOH} = 9:1$ ) to obtain dark blue solid. Yield 30 mg (34%).  $^1\text{H NMR}$  (400 MHz,  $\text{Methanol-}d_4$ )  $\delta$  8.02 (s, 2H), 7.65 (s, 1H), 7.05 (t,  $J = 2.5$  Hz, 2H), 6.71 (dd,  $J = 8.9, 1.9$  Hz, 2H), 6.63 (dt,  $J = 8.9, 2.5$  Hz, 2H), 3.61 – 3.53 (m, 2H), 3.33 (d,  $J = 3.0$  Hz, 2H), 2.53 (dt,  $J = 7.9, 3.9$  Hz, 2H), 1.38 (d,  $J = 2.0$  Hz, 9H), 1.30 (s, 0H), 0.67 (d,  $J = 2.0$  Hz, 3H), 0.57 (d,  $J = 2.0$  Hz, 3H).  $^{13}\text{C NMR}$  (101 MHz,  $\text{Methanol-}d_4$ )  $\delta$  171.29, 170.65, 167.22, 155.51, 149.76, 140.11, 136.44, 130.73, 129.62 – 125.89 (m), 125.19, 123.09, 116.39, 113.56, 93.11, 80.53, 39.07, 35.88, 34.59, 26.93, -1.06, -2.55. HRMS (ESI+), calc. for  $\text{C}_{34}\text{H}_{41}\text{N}_3\text{O}_5\text{Si}$   $[\text{M}+\text{H}]^+$  600.2815, found 600.2884.

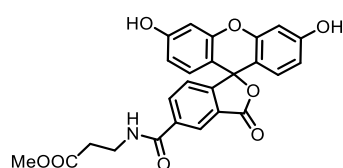


**SiR-Ala-COOH.** SiR-Ala-COO $t$ -Bu (FTK47) was dissolved in DCM, cooled on ice bath and TFA was slowly added. After 1 h of stirring at room temperature, reaction mixture was evaporated to dryness and used without further purification.

### 6.2.2. Synthesis of Rhodamine derivatives

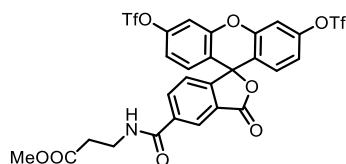


**Scheme 2.** Synthesis of BocN-R110-Ala-COOH (i)  $\beta$ -Ala-OMe $\cdot$ HCl, HATU, DIEA, DMF, r.t., 18h (43%); (ii) Py,  $\text{Tf}_2\text{O}$ , DCM,  $0^\circ\text{C}$  to r.t., 4 h (28%); (iii) BocNH,  $\text{Pd}_2\text{dba}_3$ , Xantphos,  $\text{Cs}_2\text{CO}_3$ , dioxane,  $100^\circ\text{C}$ , 18 h (30%); (iv) KOH (4 M MeOH), MeOH, r.t., 5.5 h.

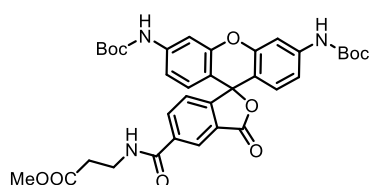


**Fluorescein-Ala (FTK120).** 5-Carboxyfluorescein (1 g, 2.67 mmole, 1 eq) was dissolved in DMF, followed by addition of DIEA (3.71 mL, 21.39 mmole, 8 eq), HATU (1.32 g, 3.47 mmole, 1.3 eq) and  $\beta$ -alanine methyl ester hydrochloride (0.45 g, 3.2 mmole, 1.2 eq). After 18 h stirring at room temperature the solvent was evaporated, dissolved in DCM, washed with brine (2x). Organic phase was collected, dried over  $\text{MgSO}_4$  and concentrated. The crude was purified on chromatography column on silica gel ( $\text{DCM}:\text{MeOH} = 8:2$ ) to obtain 530 mg of a yellow powder. Yield 43%.  $^1\text{H NMR}$  (400 MHz,  $\text{Methanol-}d_4$ )  $\delta$  8.29 (s, 1H), 8.06 (dd,  $J = 8.0, 1.9$  Hz, 1H), 7.18 (d,  $J = 8.1$  Hz, 1H), 6.58 (d,  $J = 2.1$  Hz, 2H), 6.56 – 6.38 (m, 4H), 3.64 – 3.51 (m,

6H), 3.24 (s, 1H), 3.08 (qd,  $J = 7.5, 2.4$  Hz, 1H), 2.59 (t,  $J = 6.7$  Hz, 2H), 1.23 (dp,  $J = 6.7, 3.9, 3.1$  Hz, 9H).  $^{13}\text{C}$  NMR (101 MHz, Methanol- $d_4$ )  $\delta$  172.49, 160.05, 152.66, 136.37, 134.05, 128.75, 124.32, 123.47, 112.35, 109.49, 102.27, 54.49, 50.91, 47.65, 47.44, 47.22, 47.01, 42.43, 35.83, 33.20, 11.73. HRMS (ESI+), calc. for  $\text{C}_{25}\text{H}_{19}\text{NO}_8$   $[\text{M}+\text{H}]^+$  462.1111, found 462.1184.



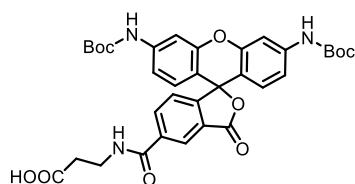
**Protected Fluorescein-Ala (FTK122).** Synthesis was followed according to a described protocol.<sup>96</sup> In brief, fluorescein-Ala FTK120 (530 mg, 1.15 mmole, 1 eq) was suspended in  $\text{CH}_2\text{Cl}_2$  (30 mL) and cooled to  $0^\circ\text{C}$ . Pyridine (0.91 mL, 9.2 mmole, 8.0 eq) and trifluoromethanesulfonic anhydride (0.95 mL, 4.6 mmole, 4.0 eq) were added, and the ice bath was removed. The reaction was stirred at room temperature for 4 h. It was subsequently diluted with water and extracted with  $\text{CH}_2\text{Cl}_2$  (2x). The combined organic extracts were washed (saturated  $\text{CuSO}_4$ , brine), dried ( $\text{MgSO}_4$ ), filtered, and concentrated *in vacuo*. Flash chromatography on silica gel (Heptane:EtOAc = 4:6) afforded 213 mg of product as a colorless foam. Yield 28%.  $^1\text{H}$  NMR (400 MHz, Chloroform- $d$ )  $\delta$  8.39 (dd,  $J = 1.6, 0.7$  Hz, 1H), 8.19 (dd,  $J = 8.0, 1.6$  Hz, 1H), 7.31 (d,  $J = 2.5$  Hz, 2H), 7.27 (dd,  $J = 8.0, 0.8$  Hz, 2H), 7.16 (t,  $J = 5.9$  Hz, 1H), 7.04 (dd,  $J = 8.8, 2.5$  Hz, 2H), 6.94 (d,  $J = 8.8$  Hz, 2H), 3.82 – 3.74 (m, 2H), 3.74 (s, 3H), 3.48 (s, 10H), 2.74 – 2.66 (m, 2H), 2.17 (s, 1H), 1.36 (s, 5H).  $^{13}\text{C}$  NMR (101 MHz, Chloroform- $d$ )  $\delta$  173.34, 167.56, 165.18, 154.50, 151.28, 150.40, 137.53, 135.09, 129.84, 126.04, 124.26, 123.91, 118.67, 117.83, 117.08, 110.87, 80.24, 52.02, 50.83, 35.67, 33.44. HRMS (ESI+), calc. for  $\text{C}_{27}\text{H}_{17}\text{F}_6\text{NO}_{12}\text{S}$   $[\text{M}+\text{H}]^+$  725.0096, found 725.977.



**BocN-R110-Ala (FTK123).** Synthesis was followed according to a described protocol.<sup>96</sup> In brief, a flask was charged with protected fluorescein-Ala FTK122 (200 mg, 0.28 mmole, 1 eq), tert-butyl carbamate (150 mg, 0.66 mmole, 2.4 eq),  $\text{Pd}_2\text{dba}_3$  (49 mg, 0.03 mmole, 0.1 eq), Xantphos (92.5 mg, 0.09 mmole, 0.3 eq), and  $\text{Cs}_2\text{CO}_3$  (486 mg, 0.78 mmole, 2.8 eq). The flask was sealed and evacuated/backfilled with nitrogen (3x). Dioxane (2 mL) was added, and the reaction was flushed again with nitrogen (3x). The reaction was stirred at  $100^\circ\text{C}$  for 18 h. It was then cooled to room temperature, filtered through celite with DCM, and evaporated. The residue was purified by silica gel chromatography (Heptane:EtOAc =

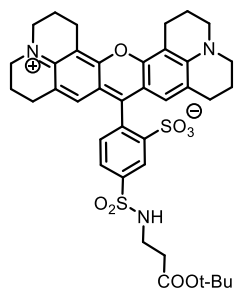


4:6) to afford 55 mg as a colorless solid. Yield 30%.  $^1\text{H}$  NMR (400 MHz, Chloroform-*d*)  $\delta$  8.35 (d,  $J$  = 1.6 Hz, 1H), 8.11 (dd,  $J$  = 8.0, 1.6 Hz, 1H), 7.47 (s, 2H), 7.31 – 7.25 (m, 1H), 7.25 – 7.19 (m, 1H), 7.14 (dt,  $J$  = 8.1, 2.1 Hz, 1H), 6.96 (s, 1H), 6.93 – 6.82 (m, 2H), 6.60 (dt,  $J$  = 8.7, 1.9 Hz, 2H), 3.76 (td,  $J$  = 6.3, 3.2 Hz, 2H), 3.71 (d,  $J$  = 1.6 Hz, 3H), 3.47 (d,  $J$  = 1.3 Hz, 1H), 2.69 (t,  $J$  = 6.0 Hz, 2H), 1.50 (d,  $J$  = 1.3 Hz, 16H), 1.33 – 1.20 (m, 4H), 0.91 – 0.83 (m, 2H).  $^{13}\text{C}$  NMR (101 MHz, Chloroform-*d*)  $\delta$  173.09, 155.75, 152.42, 152.39, 151.86, 140.91, 140.89, 136.48, 134.43, 128.35, 126.84, 124.40, 123.32, 114.27, 112.33, 112.30, 106.20, 81.16, 51.95, 35.73, 33.60, 31.86, 29.67, 28.28, 22.66, 14.07. HRMS (ESI+), calc. for  $\text{C}_{35}\text{H}_{37}\text{N}_3\text{O}_{10}$   $[\text{M}+\text{H}]^+$  660.2479.



**BocN-R110-Ala-COOH (FTK127).** FTK123 (55 mg, 0.085 mmole, 1 eq) was dissolved in 5 mL MeOH following by addition of 4M KOH solution in MeOH (1 mL). The reaction was left to stir for 5.5 h at room temperature. Yield 50 mg

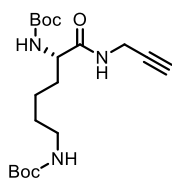
(93%). The product was used without further purification.



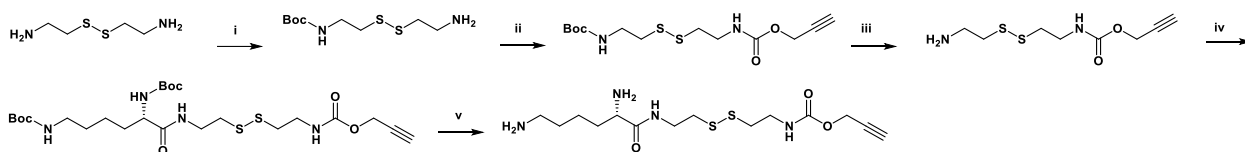
**SRB101-Ala (FTK2-02).** To a stirred suspension of sulforhodamine 101 acid (50 mg, 0.082 mmole, 1 eq) in 3 mL DCM was added, sequentially, oxalyl chloride (50  $\mu\text{L}$ , 0.49 mmole, 6 eq) slowly and drop of DMF as catalyst. The resulting mixture was stirred at r.t. for 6 h then the reaction was concentrated in vacuo to dryness. The residue was dissolved in DCM, cooled to 0-5°C followed by addition

of  $\beta$ -alanine t-butyl ester hydrochloride (20.48 mg, 0.098 mmole, 1.2 eq) and trimethylamine (79  $\mu\text{L}$ , 0.49 mmole, 6 eq). After stirring at room temperature for 18 h the reaction mixture was concentrated and purified on flash column chromatography on silica gel using gradient DCM:MeOH from 95:5 to 85:15.  $^1\text{H}$  NMR (400 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  8.63 (s, 1H), 8.06 (d,  $J$  = 8.0 Hz, 1H), 7.39 (d,  $J$  = 7.8 Hz, 1H), 6.62 (s, 2H), 4.55 (s, 4H), 3.50 (dt,  $J$  = 19.2, 5.8 Hz, 10H), 3.24 (t,  $J$  = 6.7 Hz, 2H), 3.06 (d,  $J$  = 6.8 Hz, 4H), 2.68 (dt,  $J$  = 13.9, 6.5 Hz, 5H), 2.48 (s, 1H), 2.08 (t,  $J$  = 6.1 Hz, 5H), 1.92 (q,  $J$  = 6.4 Hz, 5H), 1.45 (s, 10H), 1.28 (d,  $J$  = 8.5 Hz, 6H).

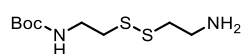
### 6.2.3. Synthesis of Gemini family of fluorogens



**Lysine-alkyne-linker, compound 1.**<sup>1</sup> To a mixture of di-boc-L-lysine dicyclohexylamine salt (1.53 g, 2.89 mmol, 1 eq) and propargylamine (277  $\mu$ L, 4.33 mmol, 1.5 eq) in dry DMF (20 mL) was added HATU (1.32 g, 3.46 mmol, 1.2 eq) followed by DIEA (1.5 mL, 8.67 mmol, 3 eq). After 1 h, the solvents were evaporated and the product was extracted with EtOAc and washed with water (2 times) and brine. The organic phase was dried over anhydrous  $MgSO_4$ , filtered and evaporated. The crude was purified by column chromatography on silica gel (DCM:MeOH = 98:2) to obtain 822 mg of compound 1 (74% yield) as a white solid. The NMR was in accordance with the literature<sup>149</sup>.

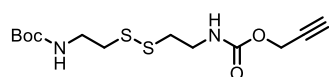


Scheme 3. Synthesis of Lysine-SS-alkyne (i)  $Boc_2O$ , TEA, MeOH, r.t. (35%); propargyl chloroformate, TEA, DCM, 5°C to r.t. (43%); (iii) TFA, DCM, 5°C to r.t.; (iv) boc-Lys-OH, HATU, DIEA, DMF, r.t. (40%); (v) TFA, DCM, 5°C to r.t.



**Boc-cystamine (FTK025).** Cystamine dihydrochloride (10 g, 1 eq) was suspended in 500 mL of methanol, and triethylamine (20 mL, 3.2 eq) was added.

Then, di-tert-butyl dicarbonate (4.85 g, 0.5 eq) was added to the solution under stirring, and the mixture was allowed to stir for 20 min at room temperature. After that, the solvent was evaporated to obtain a white solid, which was treated with 1 M  $NaH_2PO_4$  (200 mL, pH 4.2) and extracted with ether 3 times to remove the residual di-t-Boc-cystamine. The aqueous solution was basified to pH 9 using NaOH and extracted with 100 mL of ethyl acetate 5 times. The organic phase was collected and dried with anhydrous  $MgSO_4$ , filtered, and evaporated to obtain a yellowish oil-like product 3.6 g. Yield 35%. The  $^1H$  NMR was in accordance with the literature.<sup>150</sup>

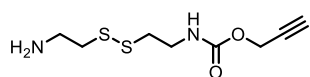


**Boc-cystamine-PPG (FTK027).** Boc-cystamine FTK025 (6.12 g, 19.72 mmole, 1 eq) and triethylamine (8 mL, 23.66 mmole, 1.2 eq.)

was dissolved in DCM and cooled to 0-5°C. Propargyl chloroformate (1.97 mL, 23.66 mmole, 1.2 eq) was dropwisely added to the mixture. The reaction mixture was left stirring at room temperature for 18 h. Solvent was evaporated and the crude was purified by column chromatography on silica gel (Heptane/EtOAc= 5/5).

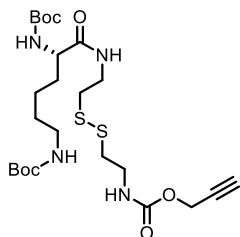
<sup>1</sup> Synthesized by Dr. Mayeul Collot

Yield, 43%.  $^1\text{H}$  NMR (400 MHz, Chloroform-*d*)  $\delta$  5.21 (s, 1H), 4.75 – 4.55 (m, 2H), 4.40 (s, 5H), 3.33 – 2.97 (m, 9H), 2.53 (dt,  $J$  = 13.7, 6.2 Hz, 8H), 2.19 (s, 2H), 1.17 (s, 9H).  $^{13}\text{C}$  NMR (101 MHz, Chloroform-*d*)  $\delta$  155.85, 155.50, 79.67, 78.21, 74.65, 52.55, 39.82, 39.39, 38.26, 28.40. HRMS (ESI+), calc. for  $\text{C}_{69}\text{H}_{85}\text{N}_9\text{O}_{15}\text{S}_4$   $[\text{M}+\text{Na}]^+$  357.1021, found 357.004.



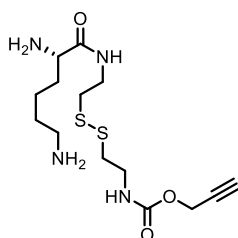
**Cystamine-PPG (FTK028).** Boc-cystamine-PPG FTK027 (1.265 g, 4.076 mmole, 1 eq) was dissolved in 50 mL DCM and

cooled on ice bath. Trifluoroacetic acid (5 mL) was slowly added and left to warm up to room temperature. The reaction mixture was left stirring for 2 h. Solvent was evaporated.  $^1\text{H}$  NMR (400 MHz, Chloroform-*d*)  $\delta$  4.69 (d,  $J$  = 14.4 Hz, 2H), 3.52 (dq,  $J$  = 25.8, 6.0 Hz, 6H), 3.06 (d,  $J$  = 6.6 Hz, 2H), 2.87 (t,  $J$  = 6.4 Hz, 3H), 2.52 (s, 1H).

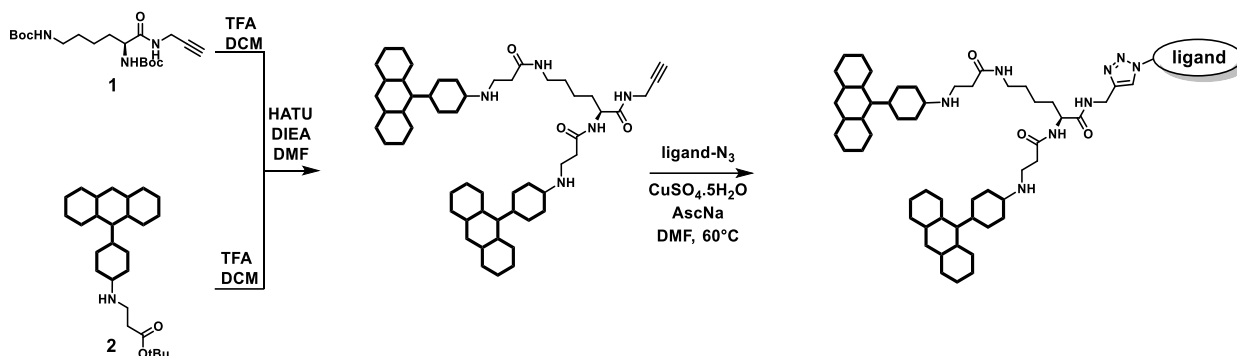


**Boc-Lysine-SS-alkyne (FTK031).** To a solution of cystamine-PPG FTK028 (2.8 g, 8.08 mmole, 1 eq) and DiBoc-Lys(OH) (1.31 g, 8.08 mmole, 1 eq) in DMF (15 mL) was added HATU (1.72 g, 9.69 mmole, 1.2 eq) followed by DIEA (2.9 mL, 12.12 mmole, 1.5 eq). After 18 h the solvent was evaporated. The residue was dissolved

in ethyl acetate, washed with water (2x), 15% citric acid solution (2x), water,  $\text{NaHCO}_3$  solution (2x) and brine (2x). Organic phase was collected and concentrated. The crude was purified by column chromatography on silica gel (Heptane/EtOAc= 4/6). Yield, 40%.  $^1\text{H}$  NMR (400 MHz, Chloroform-*d*)  $\delta$  6.72 (s, 1H), 5.48 (s, 1H), 5.23 (s, 1H), 4.70 (s, 2H), 4.62 (s, 1H), 4.10 (s, 1H), 3.55 (d,  $J$  = 28.1 Hz, 4H), 3.11 (s, 2H), 2.81 (s, 4H), 2.49 (s, 1H), 1.88 (s, 1H), 1.63 (s, 2H), 1.60 (s, 2H), 1.51 (s, 3H), 1.44 (d,  $J$  = 2.2 Hz, 18H), 1.37 (d,  $J$  = 7.3 Hz, 3H), 1.25 (s, 2H), 0.90 (s, 1H). HRMS (ESI+), calc. for  $\text{C}_{24}\text{H}_{42}\text{N}_4\text{O}_7\text{S}_2$   $[\text{M}+\text{H}]^+$  563.2495, found 563.161.



**Lysine-SS-alkyne.** Boc-Lysine-SS-alkyne was dissolved in DCM and cooled on ice bath followed by flow addition of TFA. Reaction was allowed to stir for 1 h. Solvents were evaporated to dryness and the product was used without further purification.



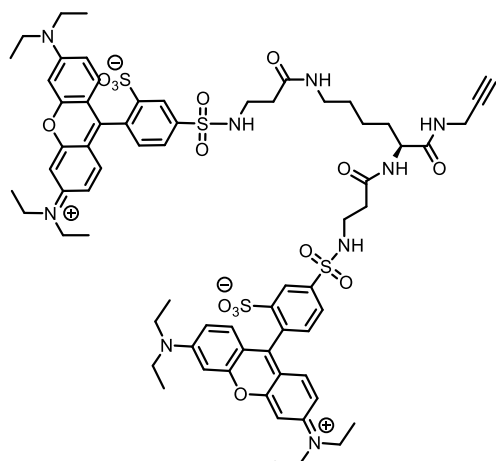
**Scheme 4.** General scheme of synthesis of Gemini fluorogens.

### **General protocol for synthesis of Gemini-alkynes.**

To a solution of deprotected Lysine-SS-alkyne or Lysine-alkyne (2 eq) and deprotected dye-Ala-COOH (1 eq) in DMF (3 mL) was added HATU (1.2 eq) followed by DIEA (3-6 eq). After 18 h the solvents were evaporated and the crude was first purified by column chromatography on silica gel, preparative TLC or by reverse phase column chromatography (C-18 column, MeCN/Water: 20/80 to 100/0 over 30 minutes).

### **General protocol of 'click' reaction.**

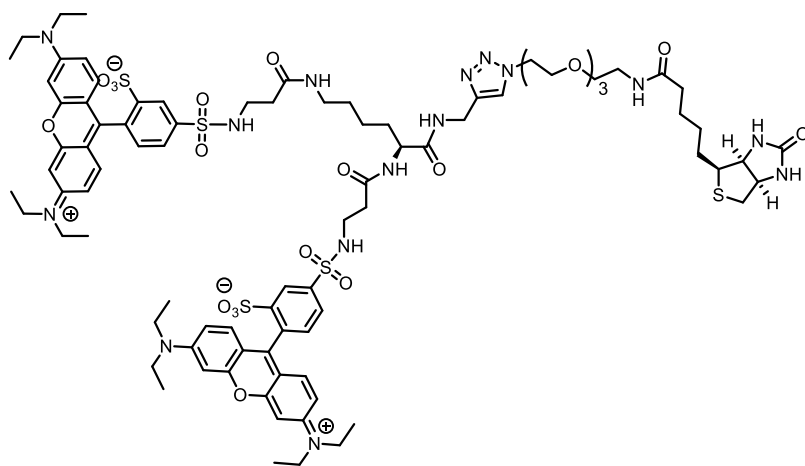
To a solution of Gemini-alkyne (1 eq) and biotin-PEG<sub>3</sub>-N<sub>3</sub>, desthiobiotin-PEG<sub>3</sub>-N<sub>3</sub> or desthiobiotin-PEG<sub>23</sub>-N<sub>3</sub> (1.5 eq) in DMF (2 mL) was added a mixture of CuSO<sub>4</sub>·5H<sub>2</sub>O (1 eq) and sodium ascorbate (1.2 eq) in water (0.2 mL). The solution was allowed to stir at room temperature or 60°C overnight. The solvents were evaporated and the product was purified by preparative TLC or reverse phase chromatography.



**Gemini-561-alkyne**<sup>2</sup> (48% yield) as dark violet syrup. *R*<sub>f</sub> = 0.27 (DCM/MeOH: 9/1). <sup>1</sup>H-NMR (400 MHz, CHCl<sub>3</sub>/MeOD): δ 8.66 (s, 2H, H Ar), 8.04 (dt, *J* = 8.0, 1.6 Hz, 2H, H Ar), 7.28 (dd, *J* = 8.0, 0.7 Hz, 2H, H Ar), 7.19 (dd, *J* = 9.5, 2.7 Hz, 4H, H Ar), 6.87-6.84 (m, 4H, H Ar), 6.71 (s, 4H, H Ar), 4.30-4.26 (m, 1H, H<sub>α</sub> lysine), 3.95-3.85 (m, 2H, CH<sub>2</sub>), 3.55-3.53 (m, 16H, CH<sub>2</sub> Et), 3.33-3.25 (m, 8H, 4 CH<sub>2</sub>), 2.46-2.43 (m, 2H), 2.35-2.33 (m, 2H), 2.18 (t, *J* = 2.4 Hz, 1H, C≡CH), 1.45-1.41 (m, 4H, 2 CH<sub>2</sub>), 1.28 (t, *J* = 7.1 Hz, 24H, CH<sub>3</sub> Et).

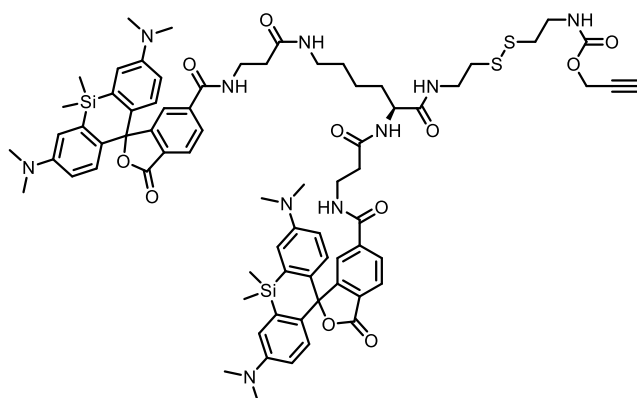
<sup>2</sup> Synthesized by Dr. Mayeul Collot

$^{13}\text{C}$ -NMR (126 MHz,  $\text{CHCl}_3/\text{MeOD}$ ): Note that some peaks are doubled due to rotamers formed with amide bonds.  $^{13}\text{C}$ -NMR (126 MHz,  $\text{CHCl}_3/\text{MeOD}$ ):  $\delta$  157.8 (C Ar), 157.2 (C Ar), 155.5 (C Ar), 146.4 (C Ar), 142.1 (C Ar), 141.9 (C Ar), 133.8 (C Ar), 133.7 (C Ar), 132.9 (C Ar), 130.4 (C Ar), 128.0 (C Ar), 127.9 (C Ar), 126.8 (C Ar), 126.7 (C Ar), 114.1 (C Ar), 113.6 (C Ar), 95.7 (C Ar), 79.4 (C alkyne), 71.0 (C alkyne), 53.6 ( $\text{C}\alpha$ ), 45.8 ( $\text{CH}_2$  Et), 39.6, 39.4, 38.8, 35.8, 35.5, 31.1, 28.7, 28.3, 22.7, 12.4 ( $\text{CH}_3$  Et). HRMS (ESI+), calc. for  $\text{C}_{69}\text{H}_{85}\text{N}_9\text{O}_{15}\text{S}_4$  [ $\text{M}+2\text{H}$ ] $^+$  703.7518, found 703.7506.



**Gemini-561<sup>3</sup>**. Synthesis was followed according to the general protocol of 'click' reaction. Dark violet syrup. Yield 49%.  $R_f$  = 0.23 (DCM/MeOH: 85/15).  $^1\text{H}$  NMR (400 MHz, MeOD):  $\delta$  8.66-8.64 (m, 2H, H Ar), 8.14-8.10

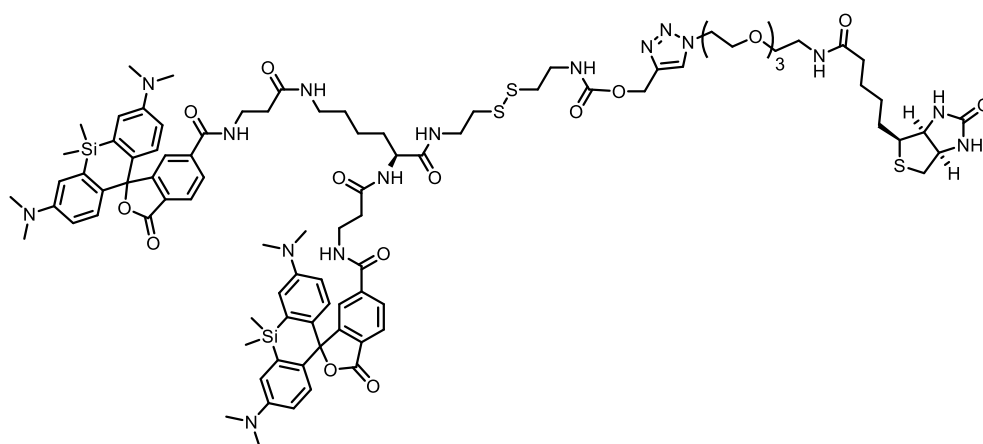
(m, 2H, H Ar), 7.94-7.91 (m, 1H, H triazol), 7.53-7.50 (m, 2H, H Ar), 7.17-7.13 (m, 4H, H Ar), 7.03-6.99 (m, 4H, H Ar), 6.96-6.92 (m, 4H, H Ar). The rest of the spectrum is difficult to assign probably due to rotamers and the weak amount of product. HPLC traces are provided as a proof of the purity of Gemini-561. HPLC: Zorbax SB-C18, particule size 1.8  $\mu\text{m}$  (Agilent), ACN/Water (0.05% formic acid) 2/98 to 100/0 in 8 min, 0.5 mL/min. HRMS (ESI+), calc. for  $\text{C}_{87}\text{H}_{118}\text{N}_{15}\text{O}_{20}\text{S}_5$  [ $\text{M}+3\text{H}$ ] $^+$  617.5755, found 617.5737.



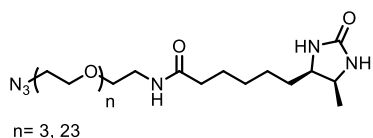
**Gemini-647-SS-alkyne (FTK055)**. Synthesis was followed according to the general protocol for Gemini-alkyne. Purification by preparative TLC (DCM:MeOH = 95:5). Yield 36%.  $^1\text{H}$  NMR (500 MHz, Methanol- $d_4$ )  $\delta$  7.90 (s, 3H), 7.86 (s, 3H), 7.81 (s, 5H), 7.62 (s, 2H), 7.54 (s, 2H), 6.92 (s,

<sup>3</sup> Synthesized by Dr. Mayeul Collot

10H), 6.56 (s, 11H), 6.47 (s, 11H), 4.50 (s, 9H), 4.07 (s, 3H), 3.57 (s, 3H), 3.46 (s, 9H), 2.92 – 2.84 (m, 11H), 2.77 – 2.71 (m, 6H), 2.59 (d,  $J = 21.4$  Hz, 11H), 2.41 (s, 5H), 2.33 (s, 5H), 1.58 (s, 3H), 1.44 (s, 3H), 1.19 (d,  $J = 51.7$  Hz, 17H), 0.55 (s, 6H), 0.54 (s, 7H), 0.44 (s, 12H).  $^{13}\text{C}$  NMR (126 MHz, Methanol- $d_4$ )  $\delta$  174.63, 173.01, 169.74, 158.89, 157.70, 152.19, 142.45, 139.26, 133.30, 131.29, 127.68, 125.99, 125.68, 118.98, 115.93, 55.96, 54.23, 42.23, 41.55, 40.73, 39.56, 38.96, 37.57, 33.44, 30.99, 25.02, 1.42. HRMS (ESI+), calc. for  $\text{C}_{74}\text{H}_{88}\text{N}_{10}\text{O}_{11}\text{S}_2\text{Si}_2$   $[\text{M}]^+$  1412.5614, found 1412.5651.

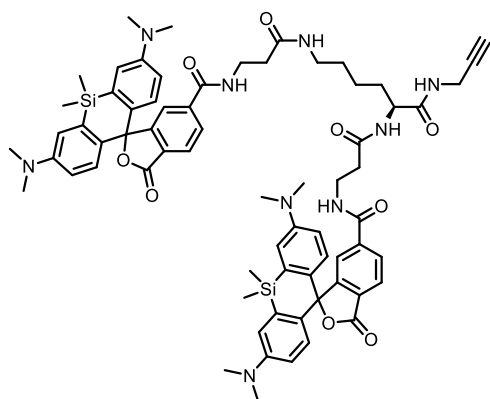


**Gemini-647 1.0 (FTK057).** Synthesis was followed according to the general protocol for click reaction. Purification by preparative TLC (DCM:MeOH = 85:15). Yield 44%.  $^1\text{H}$  NMR (500 MHz, Methanol- $d_4$ )  $\delta$  7.93 (s, 1H), 7.91 (s, 4H), 7.87 (s, 2H), 7.85 (s, 1H), 7.84 (s, 1H), 7.83 (s, 4H), 7.62 (s, 2H), 7.55 (s, 2H), 6.93 (s, 10H), 6.57 (d,  $J = 18.0$  Hz, 11H), 6.47 (d,  $J = 13.5$  Hz, 9H), 5.39 (s, 1H), 5.00 (s, 4H), 4.85 (s, 2H), 4.81 (s, 6H), 4.73 (s, 12H), 4.49 (s, 1H), 4.42 (s, 5H), 4.32 (s, 3H), 4.14 (s, 3H), 4.07 (s, 3H), 3.72 (s, 5H), 3.56 (s, 3H), 3.45 (s, 29H), 3.39 (s, 7H), 3.35 (q,  $J = 1.7$  Hz, 1H), 3.29 (dd,  $J = 3.9, 2.1$  Hz, 3H), 3.25 (s, 8H), 2.87 (s, 8H), 2.79 (s, 6H), 2.77 – 2.73 (m, 3H), 2.60 (s, 12H), 2.41 (s, 5H), 2.33 (s, 5H), 2.07 (s, 6H), 1.50 (d,  $J = 67.5$  Hz, 19H), 1.18 (d,  $J = 65.9$  Hz, 21H), 0.55 (s, 6H), 0.55 (s, 6H), 0.44 (s, 13H).  $^{13}\text{C}$  NMR (126 MHz, Methanol- $d_4$ )  $\delta$  174.64, 173.06, 172.71, 172.16, 170.51, 167.26, 155.21, 149.72, 143.04, 136.79, 136.50, 130.85, 128.81, 127.83, 125.22, 123.53, 123.23, 116.51, 113.34, 70.13, 68.86, 61.93, 60.18, 55.57, 53.51, 47.93, 39.09, 38.27, 37.55, 36.51, 35.34, 35.11, 30.99, 28.54, 28.08, 25.43, 22.58. HRMS (ESI+), calc. for  $\text{C}_{92}\text{H}_{120}\text{N}_{16}\text{O}_{16}\text{S}_3\text{Si}_2$   $[\text{M}]^+$  1856.7769, found 1856.7719.



**Desthiobiotin-PEGn-N<sub>3</sub> (n=3 FTK132; n=23 FTK116).**

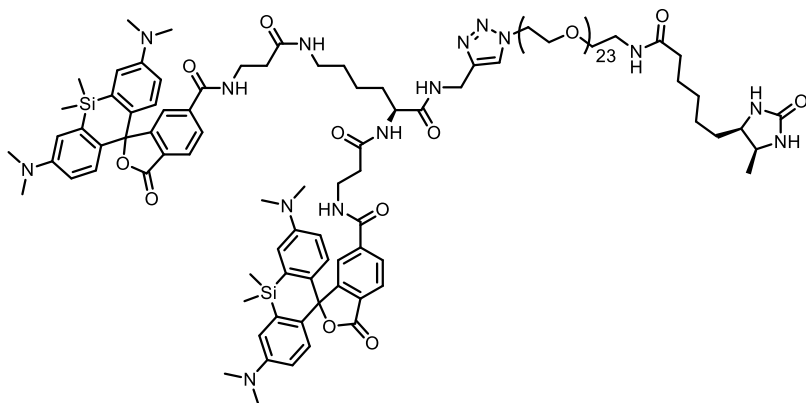
Desthiobiotin (1.2 eq), N<sub>3</sub>-PEGn-NH<sub>2</sub> (1 eq), HATU (1.5 eq), DIEA (3 eq) were dissolved in DMF. The mixture was allowed to stir for 18 h. Solvent was evaporated; the residue was dissolved in DCM, washed with water and brine. Organic layer was collected, dried over MgSO<sub>4</sub> and concentrated. The crude was purified by column chromatography on silica gel or reverse phase column chromatography. Desthiobiotin-PEG<sub>3</sub>-N<sub>3</sub>: <sup>1</sup>H NMR (400 MHz, Chloroform-d) δ 3.80 – 3.66 (m, 8H), 3.63 (d, *J* = 5.3 Hz, 4H), 3.38 (dt, *J* = 11.2, 4.8 Hz, 1H), 3.18 (tq, *J* = 8.4, 5.1, 4.5 Hz, 10H), 2.98 (d, *J* = 17.1 Hz, 6H), 2.34 (t, *J* = 7.5 Hz, 1H), 1.62 (dd, *J* = 10.1, 4.4 Hz, 2H), 1.25 (s, 11H), 0.87 (t, *J* = 6.7 Hz, 3H). Desthiobiotin-PEG<sub>23</sub>-N<sub>3</sub>: HRMS (ESI+), calc. for C<sub>58</sub>H<sub>114</sub>N<sub>6</sub>O<sub>25</sub> [M+H]<sup>+</sup> 1295.7834, found 1295.8000.



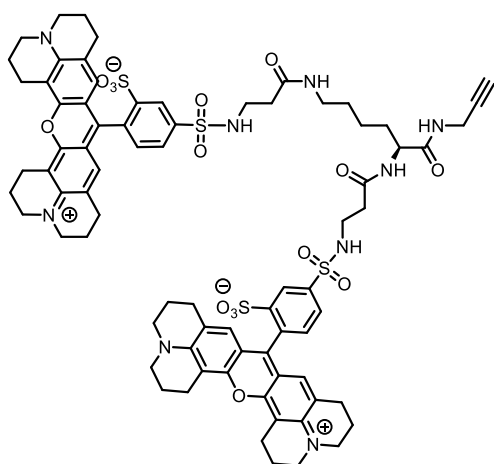
**Gemini-647-alkyne (FTK113).**

Synthesis was followed according to the general protocol for Gemini-alkyne. Purification by preparative TLC (DCM:MeOH = 9:1). Yield 20%. <sup>1</sup>H NMR (400 MHz, Methanol-*d*<sub>4</sub>) δ 7.93 – 7.78 (m, 2H), 7.80 (d, *J* = 1.3 Hz, 1H), 7.61 (t, *J* = 1.1 Hz, 1H), 7.58 – 7.51 (m, 1H), 6.99 – 6.87 (m, 4H), 6.62 – 6.49 (m, 3H), 6.42 (td, *J* = 9.1, 2.9 Hz, 3H), 4.76 (s, 3H), 4.07 (dd, *J* = 9.1, 4.9 Hz, 1H), 3.75 – 3.66 (m, 2H), 3.56 (ddt, *J* = 16.8, 13.3, 6.6 Hz, 1H), 3.49 – 3.36 (m, 2H), 3.08 (q, *J* = 7.4 Hz, 1H), 2.85 (s, 5H), 2.80 (s, 2H), 2.89 – 2.76 (m, 20H), 2.41 (s, 2H), 2.39 (q, *J* = 3.4, 2.6 Hz, 2H), 2.30 (q, *J* = 6.8, 5.8 Hz, 2H), 1.57 – 1.46 (m, 1H), 1.45 – 1.36 (m, 0H), 1.26 (s, 14H), 1.24 (dd, *J* = 7.0, 3.4 Hz, 7H), 1.22 – 1.09 (m, 4H), 1.09 (q, *J* = 4.6 Hz, 1H), 0.78 (td, *J* = 7.0, 4.6 Hz, 1H), 0.53 (d, *J* = 3.0 Hz, 5H), 0.44 (s, 1H), 0.42 (s, 4H). <sup>13</sup>C NMR (101 MHz, Methanol-*d*<sub>4</sub>) δ 174.58, 173.00, 169.61, 157.35, 152.16, 138.91, 131.22, 130.32, 125.91, 125.63, 118.91, 115.93, 95.93, 73.35, 56.87, 55.66, 41.52, 40.73, 38.92, 37.61, 33.41, 31.79, 30.92, 30.52, 24.92, 1.53. HRMS (ESI+), calc. for C<sub>69</sub>H<sub>79</sub>N<sub>9</sub>O<sub>9</sub>Si<sub>2</sub> [M]<sup>+</sup> 1234.5539, found 1234.5614.

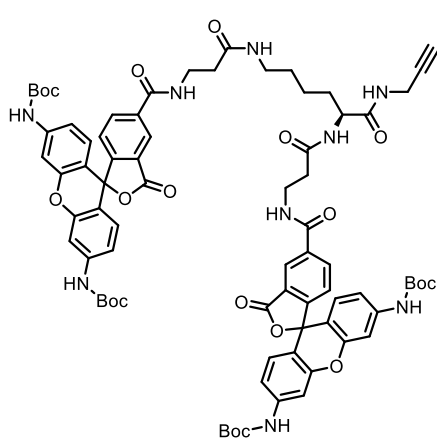




**Gemini-647 2.0 (FTK125).** Synthesis was followed according to the general protocol for click reaction. Purification by reverse phase chromatography (C-18 column, gradient MeCN to water). HRMS (ESI+), calc. for  $C_{127}H_{193}N_{15}O_{34}Si_2$   $[M]^+$  2528.3373, found 2528.3282.

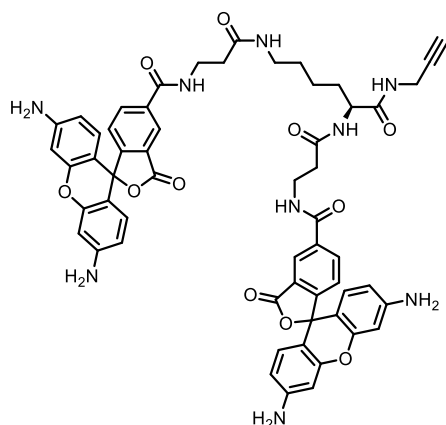


**Gemini-584-alkyne (FTK2-03).** Synthesis was followed by general protocol for Gemini-alkyne. Purification by preparative TLC (DCM:MeOH = 9:1). Yield 38%. HRMS (ESI+), calc. for  $C_{77}H_{83}N_9O_{15}S_4$   $[M+H]^+$  751.7445, found 751.7513.

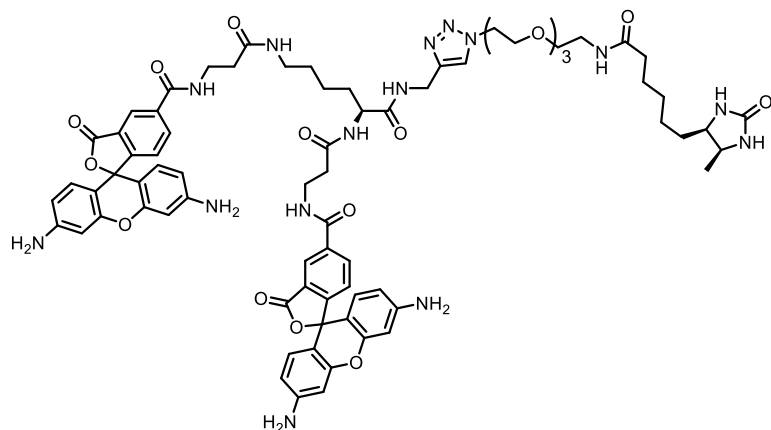


**Protected Gemini-505-alkyne (FTK128).** Synthesis was followed according to the general protocol for Gemini-alkyne. Product was purified by preparative TLC (DCM:MeOH = 92:8,  $R_f$  ~0.25). Yield, 33%.  $^1H$  NMR (500 MHz, Chloroform- $d$ )  $\delta$  8.40 (d,  $J$  = 12.1 Hz, 1H), 8.10 (s, 1H), 7.45 (s, 2H), 6.96 (s, 1H), 6.80 (s, 1H), 6.49 – 6.33 (m, 2H), 4.42 (s, 0H), 4.06 – 3.92 (m, 0H), 3.77 (s, 1H), 3.70 (s, 0H), 3.59 (hept,  $J$  = 6.6 Hz, 3H), 3.01 (q,  $J$  = 7.4 Hz, 3H), 2.50 (d,  $J$  = 44.9 Hz, 2H), 1.43 – 1.37 (m, 14H), 1.32 (s, 7H), 1.30 (s, 8H).  $^{13}C$  NMR (126 MHz, Chloroform- $d$ )  $\delta$  172.34, 169.25, 166.24, 161.82, 161.55, 152.73, 151.84, 141.23, 134.79, 128.24, 126.66, 124.11, 114.51, 112.00, 106.17, 80.91, 53.61, 41.98, 36.94,

29.69, 28.26, 18.88, 11.87. HRMS (ESI+), calc. for  $C_{77}H_{83}N_9O_{19}$   $[M+H]^+$  1438.5805, found 1438.5848.



**Gemini-505-alkyne (FTK131).** FTK128 was dissolved in 1 mL DCM. The flask was cooled in ice bath and 1 mL TFA was slowly added. Ice bath was removed and the reaction was left to stir for 2 h at room temperature. Gradual colour change was observed from pale yellow to dark orange. Solvents were evaporated to dryness and the product was further used without purification.

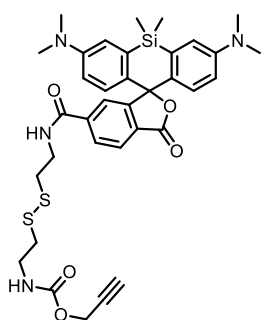


**Gemini-505 (FTK134).**

Synthesis was followed according to the general protocol for click reaction. Product was purified by reverse phase chromatography with eluent gradient (MeCN to water).

HRMS (ESI+), calc. for  $C_{75}H_{85}N_{15}O_{16}$   $[M+H]^+$  1452.6299, found 1452.65.

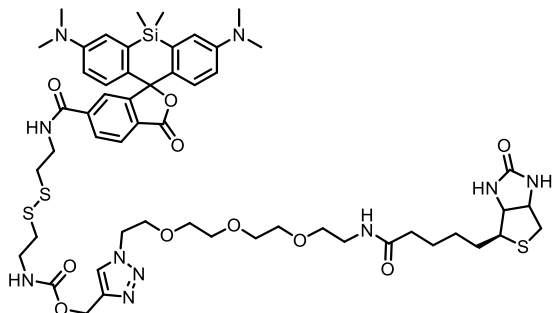
#### 6.2.4. Synthesis of control compounds



**SiR-SS-alkyne (FTK049).** SiR (35 mg, 1 eq) and DIEA (80  $\mu$ L, 6 eq) were dissolved in DMF. HATU (33.8 mg, 1.2 eq) and cystamine-PPG (20.8 mg, 1.2 eq) were sequentially added to the mixture. The reaction was left to stir for 18 h. Solvent was evaporated; the residue was dissolved in DCM, washed with water and brine. Organic layer was collected, dried over  $MgSO_4$  and concentrated.

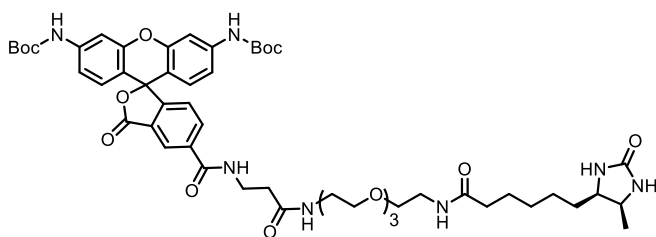
The crude was purified by preparative TLC (DCM:MeOH = 95:5). Yield 63%.  $^1H$  NMR (400 MHz, Methanol- $d_4$ )  $\delta$  8.07 – 7.98 (m, 2H), 7.70 (d,  $J$  = 8.6 Hz, 1H), 7.04 (d,  $J$  = 3.1 Hz, 2H), 6.77 – 6.68 (m, 2H), 6.62 (d,  $J$  = 9.0 Hz, 2H), 4.66 (s, 1H), 4.55 (d,  $J$  = 2.9 Hz, 1H), 3.70 – 3.59 (m, 2H), 3.39 – 3.29 (m, 5H), 2.89 (dd,  $J$  = 7.7, 4.4 Hz, 3H), 2.86 – 2.80 (m, 1H), 2.80 – 2.74 (m, 2H), 1.23 (t,  $J$  = 6.9 Hz, 1H), 0.67 (d,

$J = 2.4$  Hz, 2H), 0.57 (d,  $J = 2.4$  Hz, 3H).  $^{13}\text{C}$  NMR (101 MHz, Methanol- $d_4$ )  $\delta$  170.63, 167.40, 155.28, 149.77, 140.02, 136.56, 130.78, 128.24 (d,  $J = 74.2$  Hz), 125.26, 123.26, 116.46, 113.51, 78.07, 74.46, 51.75, 39.10, 37.59, -1.78 (d,  $J = 157.6$  Hz). HRMS (ESI+), calc. for  $\text{C}_{35}\text{H}_{40}\text{N}_4\text{O}_5\text{S}_2\text{Si}$   $[\text{M}]^+$  688.2209, found 688.2204.



**SiR-SS-biotin (FTK052).** Synthesis was followed according to the general protocol for click reaction. Product was purified by preparative TLC (DCM:MeOH = 85:15). Yield 14%.  $^1\text{H}$  NMR (500 MHz, Methanol- $d_4$ )  $\delta$  7.98 – 7.88 (m, 1H), 7.62 (s, 1H), 7.00 –

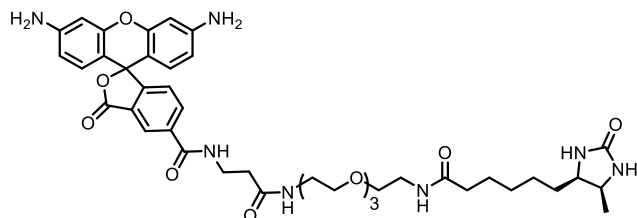
6.91 (m, 2H), 6.60 (d,  $J = 8.9$  Hz, 1H), 6.53 (dd,  $J = 8.9, 2.9$  Hz, 1H), 4.95 (s, 1H), 4.46 – 4.36 (m, 1H), 4.34 (ddd,  $J = 7.9, 5.0, 0.9$  Hz, 0H), 4.15 (dd,  $J = 7.9, 4.4$  Hz, 1H), 3.73 (t,  $J = 5.0$  Hz, 1H), 3.59 – 3.48 (m, 2H), 3.39 (t,  $J = 5.5$  Hz, 1H), 3.29 – 3.24 (m, 1H), 3.11 – 3.02 (m, 1H), 2.68 (t,  $J = 6.6$  Hz, 1H), 2.58 (s, 0H), 2.55 (s, 0H), 2.07 (t,  $J = 7.7$  Hz, 1H), 1.65 – 1.40 (m, 2H), 1.30 (h,  $J = 8.6, 8.0$  Hz, 1H), 1.23 – 1.16 (m, 2H), 0.55 (s, 1H), 0.45 (s, 1H).  $^{13}\text{C}$  NMR (101 MHz, Methanol- $d_4$ )  $\delta$  170.63, 167.40, 155.28, 149.77, 140.02, 136.56, 130.78, 128.24 (d,  $J = 74.2$  Hz), 125.26, 123.26, 116.46, 113.51, 78.07, 74.46, 51.75, 39.10, 37.59, -1.78 (d,  $J = 157.6$  Hz). HRMS (ESI+), calc. for  $\text{C}_{53}\text{H}_{72}\text{N}_{10}\text{O}_{10}\text{S}_3\text{Si}$   $[\text{M}]^+$  1132.4364, found 1132.4352.



**Protected R110-desthiobiotin (FTK150).** Synthesis was followed according to a typical coupling chemistry (see SiR-SS-alkyne). Yield 30%.  $^1\text{H}$  NMR (400 MHz, Chloroform- $d$ )  $\delta$  8.48 (s, 1H), 8.38 (t,  $J = 5.5$  Hz, 1H), 8.14 (d,  $J = 8.1$  Hz, 1H), 7.62 (s, 3H), 7.50

(s, 2H), 7.18 (s, 2H), 7.10 (dd,  $J = 13.4, 6.9$  Hz, 2H), 6.98 – 6.91 (m, 2H), 6.79 (t,  $J = 5.6$  Hz, 1H), 6.56 (d,  $J = 8.5$  Hz, 2H), 5.90 (s, 1H), 5.30 (s, 1H), 3.82 – 3.67 (m, 4H), 3.65 – 3.53 (m, 5H), 3.52 (dd,  $J = 9.5, 3.4$  Hz, 2H), 3.52 (s, 6H), 3.47 (s, 3H), 3.52 – 3.41 (m, 1H), 3.28 (dq,  $J = 12.0, 6.0$  Hz, 4H), 2.67 (s, 3H), 2.55 (t,  $J = 6.1$  Hz, 2H), 2.11 (t,  $J = 7.4$  Hz, 2H), 1.76 – 1.65 (m, 4H), 1.59 – 1.51 (m, 4H), 1.47 (s, 15H), 1.36 (d,  $J = 10.4$  Hz, 1H), 1.27 (d,  $J = 7.7$  Hz, 1H), 1.21 (d,  $J = 16.6$  Hz, 2H), 1.04 (d,  $J = 6.4$  Hz, 3H).  $^{13}\text{C}$  NMR (101 MHz, Chloroform- $d$ )  $\delta$  171.90, 165.93, 163.96, 152.67, 151.86,

141.26, 134.50, 128.22, 124.16, 123.95, 114.43, 112.16, 106.19, 80.92, 70.38, 70.32, 69.93, 69.91, 69.63, 56.02, 51.42, 37.65, 37.43, 36.75, 36.18, 35.35, 29.53, 29.10, 28.90, 28.81, 28.30, 25.89, 25.41, 15.66. HRMS (ESI+), calc. for C<sub>54</sub>H<sub>73</sub>N<sub>7</sub>O<sub>14</sub> [M+Na]<sup>+</sup> 1066.5216, found 1066.5102.

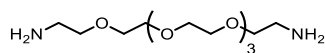


**R110-desthiobiotin (FTK154).**

Synthesis was followed according to a typical Boc-deprotection chemistry (see Gemini-505-alkyne). Yield 100%. HRMS (ESI+), calc. for C<sub>44</sub>H<sub>57</sub>N<sub>7</sub>O<sub>10</sub>

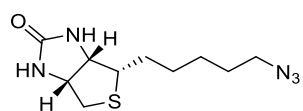
[M]<sup>+</sup> 843.4167, found 843.4145.

**6.2.5. Synthesis of biotinylated dimerized squaraines**

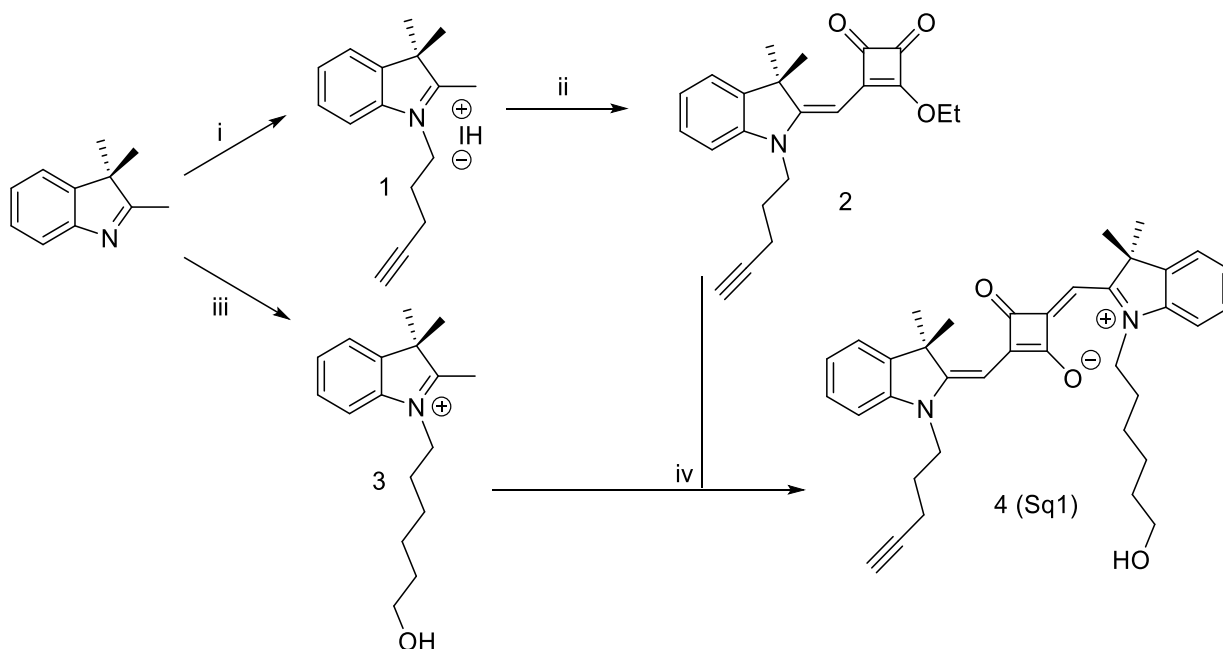


**PEG8-diamine.**<sup>4</sup>

Synthesized PEG8-diamine was in accordance with the literature<sup>144</sup>.



**N<sub>3</sub>-biotin.**<sup>4</sup> Synthesized according to a published protocol.<sup>145</sup>



<sup>4</sup> Synthesized by Dr. Mayeul Collot

**Scheme 5.** Synthesis of Sq1. i) 5-chloro-1-pentyne, KI, CH<sub>3</sub>CN, 85°C, 48 h (20%); ii) Diethyl squarate, Et<sub>3</sub>N, EtOH, 80°C, 6h (36%); iii) 6-chloro-1-hexanol, KI, CH<sub>3</sub>CN, 110°C, 12h (12%); (iv) pyridine, 125°C, 18h (31%)

Compound **1** was synthesized according to a published protocol.<sup>25</sup>

**Synthesis of compound 2.**<sup>5</sup> Compound **1** (1.2 g, 3.4 mmol, 1 eq), diethyl squarate (578 mg, 3.4 mmol, 1 eq) and Et<sub>3</sub>N (g, mmol, 2.7 eq) were dissolved in EtOH (20 mL). The mixture was refluxed overnight. The solvents were then removed under vacuum and the crude product was purified by column chromatography with Heptane/EtOAc (7/3) as eluent. A Yellow solid was obtained, yield 36%.

<sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ 7.25 – 7.16 (m, 2H), 7.01 (td, *J* = 7.4, 0.9 Hz, 1H), 6.93 – 6.87 (m, 1H), 5.41 (s, 1H), 4.82 (q, *J* = 7.1 Hz, 2H), 3.91 (t, *J* = 7.4 Hz, 2H), 2.26 (td, *J* = 6.7, 2.6 Hz, 2H), 2.06 (t, *J* = 2.6 Hz, 1H), 1.91 (p, *J* = 6.9 Hz, 2H), 1.56 (s, 6H), 1.46 (t, *J* = 7.1 Hz, 3H). <sup>13</sup>C NMR (101 MHz, Chloroform-*d*) δ 192.35, 187.81, 187.72, 173.87, 168.23, 142.60, 140.83, 127.82, 122.75, 122.02, 108.32, 82.69, 81.57, 70.09, 69.93, 47.94, 41.54, 27.06, 25.11, 16.13, 15.92. HRMS (ESI+) Calc. for C<sub>22</sub>H<sub>23</sub>NO<sub>3</sub> [M+H]<sup>+</sup> 350.1678, found 350.1757.

**Synthesis of compound 3.**<sup>5</sup> 2,3,3-trimethylindolenine (1 eq, 9.54 g), 6-chloro-1-hexanol (1 eq, 8.16 g) and potassium iodide (2.2 eq, 21.9 g) were dissolved in CH<sub>3</sub>CN (75mL), the reaction was heated up to 110°C under stirring overnight. The solvent was then removed under vacuum and the crude product was dissolved in a small amount of acetone. The mixture is then precipitated in cold ether, three times to obtain the maximum of product. Pink crystals were obtained, yield 12%.

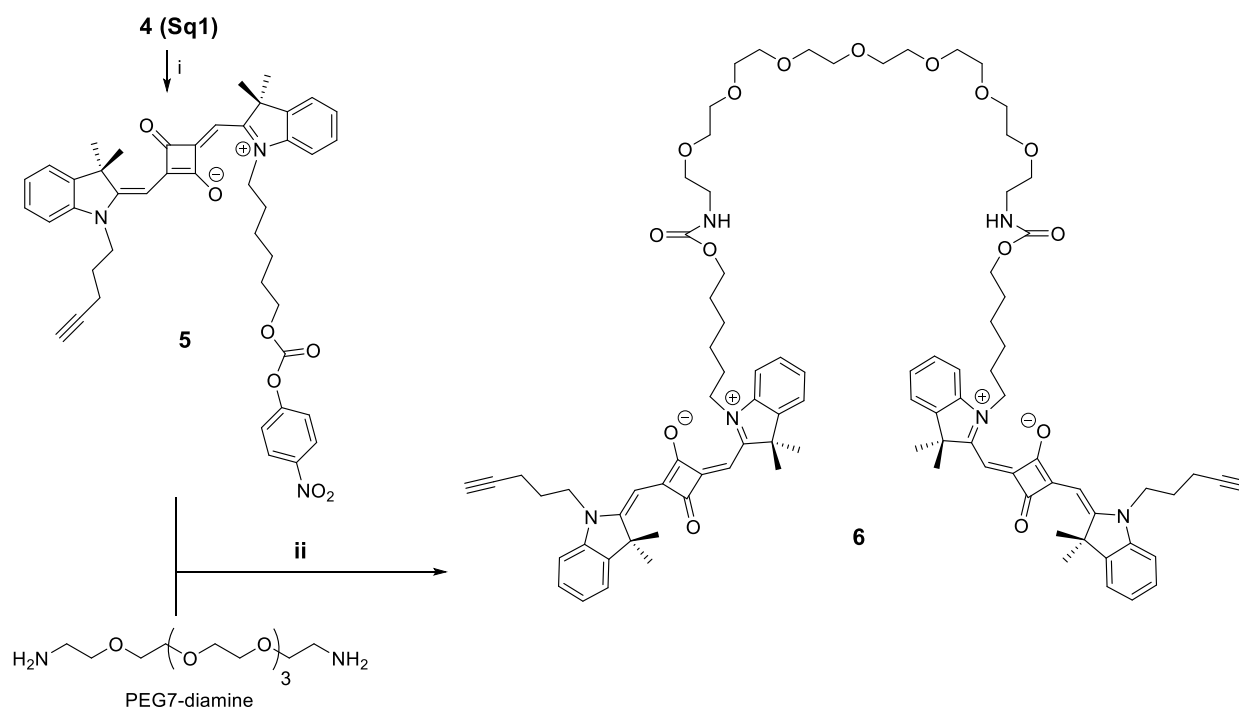
<sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ 7.74 – 7.64 (m, 1H), 7.59 – 7.46 (m, 4H), 5.24 (s, 1H), 3.54 (t, *J* = 5.9 Hz, 2H), 3.07 (s, 3H), 1.93 (t, *J* = 7.6 Hz, 2H), 1.59 (s, 6H), 1.55-1.42 (m, 8H). <sup>13</sup>C NMR (101 MHz, Chloroform-*d*) δ 195.69, 141.65, 140.94, 130.16, 129.65, 123.35, 115.62, 61.71, 49.88, 32.00, 28.01, 26.31, 25.26, 23.44, 23.26, 17.13.

**Synthesis of 4 (Sq1).**<sup>5</sup> Compound **3** (244 mg, 0.63 mmol, 1.1 eq) and compound **2** (200 mg, 0.573 mmol, 1 eq) were dissolved in 5 mL of pyridine. After 5 h the solution turned green. The solvent was removed under vacuum and the crude product was purified by column chromatography with EtOAc/MeOH (95/5) as eluent. Blue-green powder was obtained, yield 31%.

---

<sup>5</sup> Synthesized by Dr. Mayeul Collot

$^1\text{H}$  NMR (400 MHz, Chloroform-*d*)  $\delta$  7.28 (d,  $J$  = 1.2 Hz, 1H), 7.26 (d,  $J$  = 1.1 Hz, 2H), 7.24 (dd,  $J$  = 2.5, 1.2 Hz, 2H), 7.22 (d,  $J$  = 1.3 Hz, 2H), 7.20 (d,  $J$  = 1.2 Hz, 1H), 7.09 – 7.00 (m, 5H), 6.93 (s, 1H), 6.91 (s, 1H), 5.92 (s, 2H), 3.93 (s, 4H), 3.61 (t,  $J$  = 6.1 Hz, 4H), 2.26 (d,  $J$  = 2.7 Hz, 2H), 1.95 (s, 3H), 1.79 – 1.72 (m, 3H), 1.69 (d,  $J$  = 3.9 Hz, 27H), 1.58 – 1.49 (m, 3H), 1.45 (dh,  $J$  = 9.0, 2.3, 1.8 Hz, 6H).  $^{13}\text{C}$  NMR (126 MHz, Chloroform-*d*)  $\delta$  179.62, 170.54, 169.73, 142.47, 142.26, 127.85, 127.83, 123.91, 123.71, 122.32, 122.29, 109.49, 109.31, 86.78, 82.71, 70.21, 62.40, 49.37, 49.21, 43.21, 42.23, 32.43, 27.14, 26.98, 26.25, 26.24, 25.62, 24.93, 16.11. HRMS (ESI+) Calc. for  $\text{C}_{37}\text{H}_{42}\text{N}_2\text{O}_3$   $[\text{M}+\text{H}]^+$  563.3195, found 563.3265.



**Scheme 6.** Synthesis of 6. i) 4-nitrophenyl chloroformate, pyridine, DCM, r.t., 3 h (61%); ii) DIEA, DMF, 60°C, 16 h (21%).

**Synthesis of compound 5.**<sup>6</sup> Compound 4 (260 mg, 0.462 mmol, 1 eq) was dissolved in DCM, followed by pyridine (112  $\mu\text{L}$ , 1.386 mmol, 3 eq) addition and 4-nitrophenyl chloroformate (186 mg, 0.924 mmol, 2 eq). Reaction was left stirring for 3 hours at

<sup>6</sup> Synthesized by Dr. Remy Creder

room temperature. The solvent was evaporated and the residue was purified on column EtOAc/DCM (3/7). Yield, 61%.

$^1\text{H}$  NMR (400 MHz, Chloroform-*d*)  $\delta$  8.21 (d, 2H), 7.32-7.26 (m, 6H), 7.1-7 (m, 3H), 6.9 (d, 2H), 5.89 (s, 2H), 4.22 (t, 2H), 4.07-4.04 (m, 4H), 2.3 (m, 2H), 2.09 (s, 1H), 1.99 (m, 2H), 1.8 (m, 4H), 1.77 (s, 12H), 1.55 (m, 4H), 1.44 (m, 4H), (d,  $J = 3.9$  Hz, 27H), 1.58 – 1.49 (m, 3H), 1.45 (dh,  $J = 9.0, 2.3, 1.8$  Hz, 6H).  $^{13}\text{C}$  NMR (100 MHz, Chloroform-*d*)  $\delta$  127.88, 125.29, 123.82, 109.31 86.65, 69.28, 49.42, 28.43, 27.18, 27.11, 26.99, 25.67, 25.57, 16.17. HRMS (ESI+) Calc. for  $\text{C}_{44}\text{H}_{45}\text{N}_3\text{O}_7$   $[\text{M}+\text{H}]^+$  728.3258, found 728.3327.

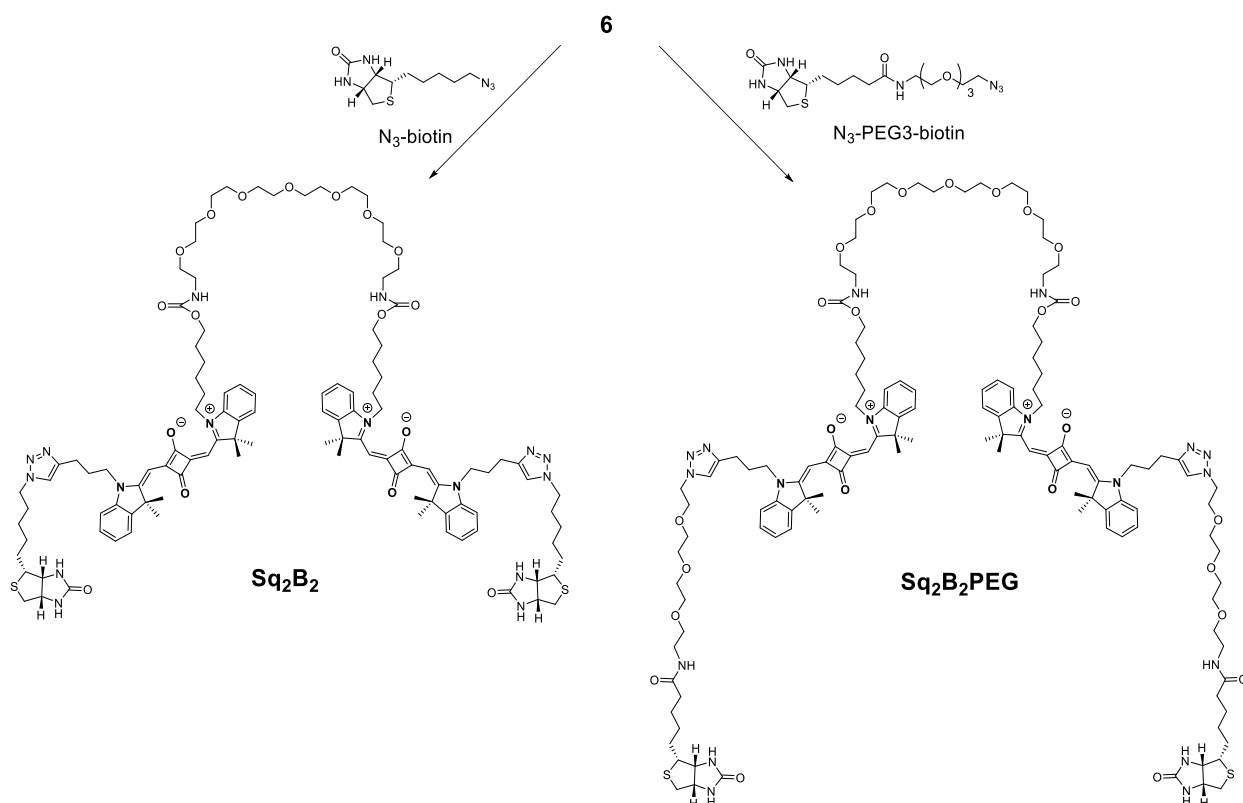
**Synthesis of 6.**<sup>7</sup> Compound **5** (60 mg, 2eq) and PEG8-diamine (15 mg, 1 eq) were dissolved in DMF and DIEA (26 mg, 5 eq) was added. The mixture was left to stir overnight at 60°C. The solvent was removed under vacuum and the crude was dissolved in DCM. The solution was washed with concentrated solution of  $\text{NaHCO}_3$ , dried over  $\text{MgSO}_4$  and concentrated. The residue was purified by column chromatography with DCM/MeOH (97/3) as eluent. Yield 21%.

$^1\text{H}$  NMR (500 MHz, Chloroform-*d*)  $\delta$  7.29-7.27 (m, 8H), 7.09-7.02 (m, 6H), 6.91 (d, 2H), 5.88 (s, 4H), 3.98 (m, 12H), 3.57 (t, 28H), 3.29 (t, 4H), 2.29 (m, 4H), 1.98 (m, 6H), 1.75 (s, 24H), 1.54 (m, 4H), 1.36 (m, 8H).  $^{13}\text{C}$  NMR (126 MHz, Chloroform-*d*)  $\delta$  180.26, 179.38, 170.43, 169.69, 156.83, 142.36, 127.85, 123.88, 122.36, 109.46, 86.68, 70.29, 64.63, 49.43, 43.67, 42.24, 40.78, 28.94, 27.2, 27.06, 26.76, 25.77, 16.17. HRMS (ESI+) Calc. for  $\text{C}_{92}\text{H}_{116}\text{N}_6\text{O}_{15}$   $[\text{M}+\text{H}]^+$  1544.8499, found 1545.8536.

---

<sup>7</sup> Synthesized by Dr. Mayeul Collot





**Scheme 7.** Synthesis of Sq<sub>2</sub>B<sub>2</sub> and Sq<sub>2</sub>B<sub>2</sub>PEG. CuSO<sub>4</sub>·5H<sub>2</sub>O, sodium ascorbate, DMF/water (3/1), 60°C, 16 h.

### General protocol for synthesis of Sq<sub>2</sub>B<sub>2</sub> and Sq<sub>2</sub>B<sub>2</sub>PEG.

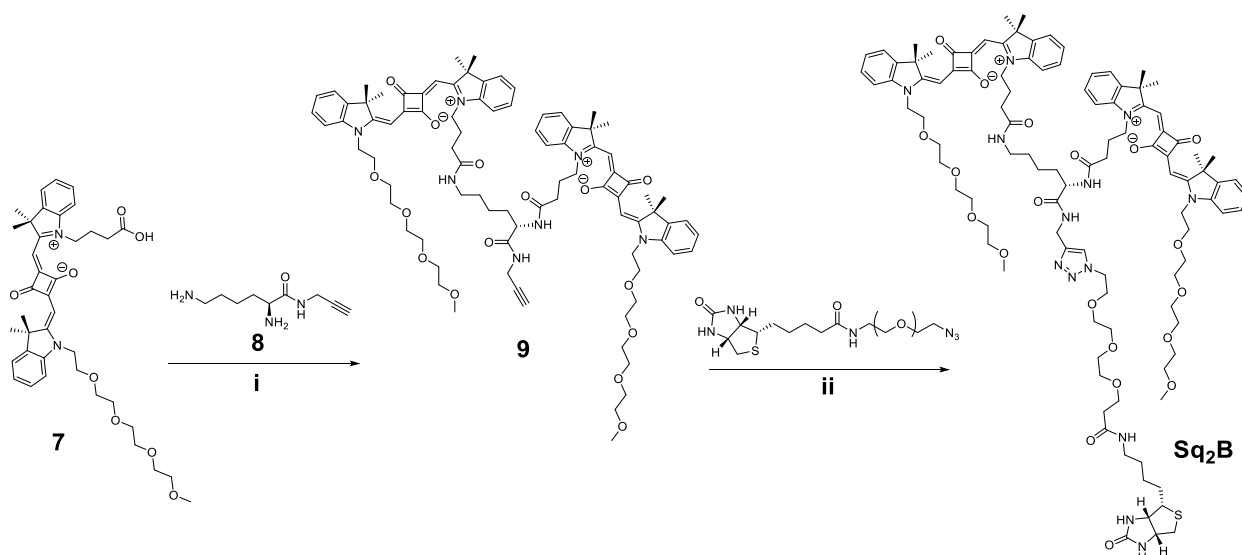
**6** (1eq) and N<sub>3</sub>-biotin (2.4 eq) or N<sub>3</sub>-PEG3-biotin (2.4 eq) were dissolved in a mixture DMF/water (3/1). CuSO<sub>4</sub>·5H<sub>2</sub>O and sodium ascorbate were dissolved in an eppendorf with water and the solution was mixed until the mixture turn yellow. The content of the eppendorf was then added to the mixture and the reaction was stirred at 60°C overnight. The solvent was then removed under vacuum; the crude product was dissolved in DCM, washed with 0.05 M EDTA solution, dried over MgSO<sub>4</sub> and concentrated. The crude was purified by column chromatography in the gradient of eluent DCM/MeOH (9/1 to 7/3) to yield blue syrup (44% for Sq<sub>2</sub>B<sub>2</sub>, 70% for Sq<sub>2</sub>B<sub>2</sub>PEG)

**Sq<sub>2</sub>B<sub>2</sub>**<sup>8</sup>: <sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ 7.54 (s, 1H), 7.32 – 7.17 (m, 4H), 7.07 (q, *J* = 7.1 Hz, 2H), 6.94 (dd, *J* = 16.5, 7.9 Hz, 2H), 5.98 (s, 1H), 5.87 (d, *J* = 11.3 Hz, 2H), 5.48 (d, *J* = 8.7 Hz, 2H), 4.40 (dd, *J* = 8.0, 4.8 Hz, 1H), 4.32 (t, *J* = 7.3 Hz, 2H), 4.19 (dd, *J* = 8.2, 4.8 Hz, 1H), 4.03 – 3.92 (m, 4H), 3.55 (d, *J* = 7.4 Hz, 12H), 3.47 (t, *J* = 5.2 Hz, 2H), 3.41 (s, 1H), 3.28 (q, *J* = 5.6 Hz, 2H), 3.03 (ddd, *J* = 8.4, 6.1, 4.4 Hz, 1H), 2.79

<sup>8</sup> Synthesized by Dr. Mayeul Collot

(q,  $J = 6.4, 5.4$  Hz, 3H), 2.64 (d,  $J = 12.7$  Hz, 1H), 2.23 – 2.10 (m, 6H), 1.85 (p,  $J = 7.5$  Hz, 2H), 1.70 (d,  $J = 2.4$  Hz, 13H), 1.60 (q,  $J = 7.5$  Hz, 1H), 1.42 – 1.28 (m, 6H), 1.29 (s, 3H), 1.18 (d,  $J = 5.9$  Hz, 1H).  $^{13}\text{C}$  NMR (101 MHz, Chloroform- $d$ )  $\delta$  178.92, 170.15, 127.92, 127.85, 123.82, 122.32, 122.24, 121.55, 109.69, 109.46, 86.59, 70.51, 70.48, 70.23, 70.16, 64.62, 62.01, 60.14, 55.61, 50.12, 49.30, 43.62, 42.89, 40.76, 40.51, 29.93, 28.91, 28.47, 28.38, 27.10, 27.05, 26.99, 26.70, 26.41, 26.29, 25.73, 22.92. HRMS (ESI+) Calc. for  $\text{C}_{112}\text{H}_{150}\text{N}_{16}\text{O}_{14}\text{S}_2$   $[\text{M}+\text{H}]^+$  2056.0806, found 2056.0832

**Sq<sub>2</sub>B<sub>2</sub>PEG<sup>9</sup>**:  $^1\text{H}$  NMR (500 MHz, Chloroform- $d$ )  $\delta$  7.92 (s, 2H), 7.46 (t, 4H), 7.39-7.35 (q, 4H), 7.27-7.19 (m, 8H), 5.99 (d, 4H), 4.594 (t, 4H), 4.035 (t, 4H), 3.62-3.59 (m, 34H), 3.56-3.48 (m, 16H), 3.34 (p, 16H), 3.28 (t, 4H), 3.18 (p, 4H), 2.93-2.29 (m, 6H), 2.69 (d, 2H), 2.25-2.18 (m, 8H), 1.96 (s, 64H), 1.85-1.83 (m, 4H), 1.76 (d, 24H), 1.73-1.58 (m, 10H), 1.45-1.38 (p, 5H).  $^{13}\text{C}$  NMR (126 MHz, Chloroform- $d$ )  $\delta$  184.57, 180.42, 177.20, 176.56, 176.11, 172.51, 172.00, 166.12, 159.34, 143.72, 143.42, 143.26, 129.42, 129.41, 125.65, 125.50, 124.44, 123.51, 111.63, 111.46, 87.36, 87.22, 71.64, 71.56, 71.52, 71.42, 71.37, 71.28, 70.69, 70.52, 65.87, 63.47, 61.73, 57.11, 51.46, 50.74, 50.62, 44.79, 44.12, 41.73, 41.19, 40.46, 36.85, 30.88, 30.23, 29.90, 29.84, 29.63, 28.31, 27.84, 27.73, 27.57, 27.47, 26.96, 23.80, 23.71, 23.67. HRMS (ESI+) Calc. for  $\text{C}_{128}\text{H}_{180}\text{N}_{18}\text{O}_{25}\text{S}_2$   $[\text{M}+\text{H}]^+$  2433.2808, found 2433.2624.



**Scheme 8.** Synthesis of Sq<sub>2</sub>B. i) HATU, DIEA, DMF, r.t., 1h (50%); ii) CuSO<sub>4</sub>·5H<sub>2</sub>O, sodium ascorbate, DMF/water (3/1), 60°C, 18 h (60%).

Compound **7** and **8** were synthesized according to published protocols.<sup>35,92</sup>

**Synthesis of compound 9<sup>9</sup>.** To a solution of **7** (40 mg, 0.06 mmol, 2 eq) and **8** (12 mg, 0.03 mmol, 1 eq) in DMF (3 mL) was added HATU (54 mg, 0.14 mmol, 1.2 eq) followed by DIEA (62  $\mu$ L, 0.36 mmol, 12 eq). After 1 h the solvents were evaporated and the crude was first purified by column chromatography on silica gel (DCM/MeOH: 8/2). Yield 50%.

<sup>1</sup>H NMR (500 MHz, Chloroform-*d*)  $\delta$  10.96 (s, 1H), 7.96 (s, 1H), 7.53 (s, 2H), 7.32 – 7.26 (m, 0H), 7.28 – 7.20 (m, 5H), 7.21 (d, *J* = 1.6 Hz, 2H), 7.22 – 7.16 (m, 1H), 7.09 (d, *J* = 7.9 Hz, 1H), 7.10 – 7.00 (m, 7H), 6.93 (d, *J* = 14.6 Hz, 1H), 5.93 (d, *J* = 8.2 Hz, 2H), 5.86 (s, 2H), 4.48 (q, *J* = 7.4 Hz, 1H), 4.17 (s, 5H), 4.04 (s, 3H), 3.92 (ddd, *J* = 8.8, 5.4, 2.5 Hz, 2H), 3.77 (td, *J* = 5.9, 1.7 Hz, 4H), 3.62 (p, *J* = 6.7 Hz, 2H), 3.55 – 3.41 (m, 22H), 3.28 (s, 5H), 3.21 (dq, *J* = 13.9, 6.8 Hz, 2H), 3.04 (q, *J* = 7.4 Hz, 2H), 2.42 (t, *J* = 6.7 Hz, 2H), 2.37 (t, *J* = 6.9 Hz, 2H), 2.11 – 1.96 (m, 2H), 1.79 (t, *J* = 7.2 Hz, 1H), 1.66 (d, *J* = 12.7 Hz, 25H), 1.51 (dq, *J* = 21.9, 7.7, 7.3 Hz, 2H), 1.42 – 1.27 (m, 17H), 1.20 (d, *J* = 15.5 Hz, 6H), 0.80 (q, *J* = 10.5, 8.6 Hz, 1H). <sup>13</sup>C NMR (126 MHz, Chloroform-*d*)  $\delta$  172.32, 142.94, 142.06, 128.09, 127.72, 124.12, 123.81, 123.73, 122.25, 122.02, 110.24, 110.09, 109.85, 79.70, 71.84, 71.82, 71.17, 71.05, 70.54, 70.39, 70.37, 67.65, 59.01, 53.58, 49.49, 49.43, 49.23, 43.91, 43.17, 42.95, 41.96, 38.76, 32.83, 32.42, 30.31, 29.69, 29.03, 28.61, 27.12, 26.95, 22.80, 22.61, 22.52, 18.57, 17.45, 11.84. HRMS (ESI+) Calc. for C<sub>87</sub>H<sub>109</sub>N<sub>7</sub>O<sub>15</sub> [M]<sup>+</sup> 1492.7982, found 1491.7989.

Synthesis of **Sq<sub>2</sub>B<sup>10</sup>** and **SqB<sup>10</sup>** was followed according to the general protocol for synthesis of **Sq<sub>2</sub>B<sub>2</sub><sup>10</sup>** and **Sq<sub>2</sub>B<sub>2</sub>PEG**.

**Sq<sub>2</sub>B<sup>10</sup>**: <sup>1</sup>H NMR (400 MHz, Chloroform-*d*)  $\delta$  10.04 – 9.99 (m, 15H), 8.04 (d, *J* = 6.8 Hz, 1H), 7.94 (s, 1H), 7.72 (d, *J* = 7.3 Hz, 1H), 7.52 (d, *J* = 6.0 Hz, 1H), 7.40 (s, 1H), 7.32 (td, *J* = 11.7, 9.6, 5.2 Hz, 9H), 7.18 (qd, *J* = 8.9, 7.7, 3.9 Hz, 9H), 6.52 (s, 3H), 6.07 (s, 1H), 6.02 (s, 1H), 5.94 (d, *J* = 14.5 Hz, 2H), 4.62 – 4.55 (m, 1H), 4.59 – 4.47 (m, 4H), 4.39 (dd, *J* = 8.0, 4.8 Hz, 2H), 4.25 (d, *J* = 5.6 Hz, 4H), 4.07 (t, *J* = 8.1 Hz, 4H), 3.95 (s, 1H), 3.84 (d, *J* = 5.4 Hz, 6H), 3.53 (ddt, *J* = 10.0, 6.6, 4.1 Hz, 35H), 3.46 – 3.38 (m, 2H), 3.34 (s, 6H), 3.18 (dq, *J* = 18.7, 6.7 Hz, 3H), 2.90 (dd, *J* = 13.2, 4.8 Hz, 1H), 2.76 (d, *J* = 13.0 Hz, 1H), 2.52 (t, *J* = 7.3 Hz, 4H), 2.24 (t, *J* = 7.5 Hz, 2H), 2.08

---

<sup>10</sup> Synthesized by Dr. Mayeul Collot

(q,  $J = 8.7, 7.8$  Hz, 5H), 1.79 (s, 1H), 1.70 (s, 17H), 1.62 (t,  $J = 7.4$  Hz, 1H), 1.57 – 1.49 (m, 1H), 1.46 – 1.38 (m, 1H), 1.37 (s, 3H).  $^{13}\text{C}$  NMR (126 MHz, Chloroform- $d$ )  $\delta$  175.53, 172.86, 142.02, 141.86, 141.58, 128.04, 125.35, 125.19, 124.90, 122.20, 122.09, 114.23, 111.19, 111.11, 110.90, 110.65, 71.79, 71.00, 70.98, 70.51, 70.49, 70.41, 70.37, 70.24, 70.00, 69.27, 68.80, 67.75, 62.49, 60.92, 58.91, 55.41, 51.03, 49.90, 49.85, 49.67, 49.65, 44.43, 43.38, 40.30, 39.67, 39.43, 35.16, 34.04, 32.44, 32.10, 31.20, 28.23, 27.92, 27.80, 26.54, 26.47, 26.43, 25.40, 23.35, 22.71. HRMS (ESI+) Calc. for  $\text{C}_{105}\text{H}_{141}\text{N}_{13}\text{O}_{20}\text{S}$   $[\text{M}+\text{H}+\text{Na}]^+$  1960.0137, found 1960.0110.

**SqB:**  $^1\text{H}$  NMR (400 MHz, Methanol- $d_4$ )  $\delta$  8.16 – 8.08 (m, 2H), 7.78 (s, 1H), 7.35 – 7.27 (m, 4H), 7.25 – 7.15 (m, 3H), 7.08 (ddt,  $J = 8.8, 7.8, 1.4$  Hz, 3H), 5.92 (s, 1H), 5.83 (s, 1H), 4.50 – 4.44 (m, 2H), 4.37 – 4.33 (m, 1H), 4.28 – 4.22 (m, 4H), 4.16 (dd,  $J = 7.9, 4.4$  Hz, 1H), 4.08 (td,  $J = 6.8, 3.0$  Hz, 2H), 3.84 – 3.75 (m, 4H), 3.61 – 3.53 (m, 2H), 3.50 – 3.43 (m, 6H), 3.36 (t,  $J = 5.4$  Hz, 2H), 3.21 (dt,  $J = 3.3, 1.6$  Hz, 6H), 2.82 – 2.72 (m, 3H), 2.08 (dt,  $J = 14.7, 7.3$  Hz, 4H), 1.63 (d,  $J = 4.2$  Hz, 12H), 1.49 (dd,  $J = 14.7, 7.2$  Hz, 2H), 1.30 (q,  $J = 7.5$  Hz, 2H).  $^{13}\text{C}$  NMR (101 MHz, Methanol- $d_4$ )  $\delta$  183.05, 179.81 – 173.30 (m), 171.04 (d,  $J = 87.9$  Hz), 164.63, 155.68, 152.55, 146.22, 146.14 – 141.27 (m), 127.76 (d,  $J = 23.7$  Hz), 124.87, 123.97 (d,  $J = 5.0$  Hz), 122.88, 110.38 (d,  $J = 84.5$  Hz), 85.77, 73.13 – 63.93 (m), 61.08 (d,  $J = 175.8$  Hz), 55.59, 50.76 – 48.90 (m), 43.22 (d,  $J = 118.9$  Hz), 39.30 (d,  $J = 72.7$  Hz), 35.33, 28.23 (d,  $J = 27.2$  Hz), 26.31, 25.73 (d,  $J = 60.7$  Hz), 22.29. HRMS (ESI+) Calc. for  $\text{C}_{57}\text{H}_{78}\text{N}_8\text{O}_{11}\text{S}$   $[\text{M}+\text{Na}]^+$  1105.5511, found 1105.5420.

### 6.3. Spectroscopy

Absorption and emission spectra were recorded on a Cary 4000 or 5000 spectrophotometer (Varian) and on a FluoroMax-4 spectrofluorometer (Horiba Jobin Yvon) or FS5 (Edinburgh Instruments) spectrofluorometer respectively. The water used for optical spectroscopy was Milli-Q water (Millipore®). All the solvents were spectroscopic grade. For standard recording of fluorescence spectra, the excitation wavelength was chosen to excite the sample optimally. The fluorescence spectra were corrected for detector response and lamp fluctuations. Relative fluorescence quantum yields were measured using appropriate reference dye.

**Fluorescence correlation spectroscopy, FCS (project ‘Gemini-561/o-Coral fluoromodule’).** FCS measurements were acquired on a home-build setup based on

a Nikon Ti-E inverted microscope as described in <sup>151</sup>. For sample excitation, we used the module of a diode laser (Oxxius, France) guided through the fiber optics and focused inside the sample with a water immersion objective (100×, NA = 1.2). For the fluorescence signal detection, a fibered Avalanche Photodiode (APD, PerkinElmer, Fremont, CA) was used. Before the APD, the fluorescence signal was filtered through a 405/488/532/635 nm BrightLine® quad-edge laser-grade dichroic (Semrock, NY) and respective filter (Chroma Corp., Rockingham, VT). The fluorescence signal was further processed by an ALV7002/USB digital correlator (ALV, Langen, Germany). The measurements were carried out in a 96-well plate, using a 200-μL volume per well. The focal spot was set about 20 μm above the coverslip. To calculate the size and concentration of fluorogen-aptamer modules or of fluorescent proteins (FPs), 50 nM solution of corresponding reference dyes was used (fluorescein (TRIS, pH 9) and tetramethylrhodamine in water). For FCS data analysis, we used the PyCorrFit<sup>152</sup> free software. The size of the fluorescent species (FSs, fluoromolecules or FPs) were estimated by calculating hydrodynamic diameter (d) as:

$d(\text{FSc}) = \tau_{\text{corr}}(\text{FSc}) / \tau_{\text{corr}}(\text{TMR}) \times d(\text{ref.})$ , where  $d(\text{ref.})$  is 1 nm, a hydrodynamic diameter of the reference;  $\tau_{\text{corr}}$  - correlation time.

Concentration (C) of FSs was calculated from the number of species (n) by:

$C(\text{FSs}) = n(\text{FSs}) / n(\text{ref.}) \times C(\text{ref.})$ , where  $C(\text{ref.})$  is 50 nM, concentration of the reference.

#### 6.4. Microscopy and cellular studies.

**Cell culture.** HeLa (ATCC® CCL-2™) and HEK293T (ATCC® CRL-3216™) cells were grown in Dulbecco's Modified Eagle Medium without phenol red (DMEM, Gibco-Invitrogen) supplemented with 10% fetal bovine serum (FBS, Lonza), 1% L-Glutamine (Sigma Aldrich) and 1% antibiotic solution (Penicillin-Streptomycin, Sigma-Aldrich) at 37°C in humidified atmosphere containing 5% CO<sub>2</sub>. U87MG (ATCC® HTB-14™) were grown in Minimum Essential Medium (MEM, Gibco-Invitrogen) supplemented with 10% FBS, 1% Ultra-Glutamine (Gibco-Invitrogen) and 1% antibiotic solution. KB cells (ATCC® CCL-17) were grown in minimum essential medium (MEM, Gibco-Invitrogen) with 10% FBS (Lonza), 1% non-essential amino acids (Gibco-Invitrogen), 1% MEM vitamin solution (Gibco-Invitrogen), 1% L-Glutamine (Sigma Aldrich) and 0.1% antibiotic solution (gentamicin, Sigma-Aldrich) at 37°C in humidified atmosphere containing 5% CO<sub>2</sub>.

**Transient transfection.** RNA-coding constructs were transfected directly into a 35 mm glass-bottomed imaging dish (IBiDi®) using FuGene HD (Promega) transfecting agent following recommended manufacturer protocol. Imaging experiments were performed between 16-48 h post-transfection.

**Cytotoxicity assay.** Cytotoxicity assay of the dyes was quantified by the MTT assay (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide). A total of  $1 \times 10^4$  cells/well were seeded in a 96-well plate 24 h prior to the cytotoxicity assay in growth medium and were incubated in a 5% CO<sub>2</sub> incubator at 37°C. After medium removal, an amount of 100 µL DMEM containing corresponding concentration of dye was added to cells and incubated for indicated time at 37°C (5% CO<sub>2</sub>). As control, for each 96-well plate, the cells were incubated with DMEM containing the same percentage of DMSO (0,5% v/v) as the solution with the tested dyes. After dye incubation, the medium was replaced by 100 µL of a mix containing DMEM + MTT solution (diluted in PBS beforehand) and the cells were incubated for 4 h at 37°C. Then, 75 µL of the mix was replaced by 50 µL of DMSO (100%) and gently shaken for 15 min at room temperature in order to dissolve the insoluble purple formazan reduced in living cells. The absorbance at 540 nm was measured (absorbances of the dyes at 540 nm were taken into account). Each concentration of dye was tested in sextuplicate and the data were shown as the mean value plus a standard deviation ( $\pm$ SD). For each concentration, we calculated the percentage of cell viability in reference of the control DMEM+ 0.5% DMSO.

**Microscopy imaging (project 'Gemini-561/o-Coral fluoromodule').** Cells were seeded onto a 35 mm glass-bottomed imaging dish (IBiDi®) at a density of  $3-5 \times 10^4$  cells/well 48 h before the microscopy measurement. 16-24 h prior to imaging cells were transfected with corresponding DNA plasmid. For imaging, the culture medium was removed and the attached cells were washed with Opti-MEM (Gibco–Invitrogen). Next, the cells were incubated in Opti-MEM with Hoechst (5 µg/mL) to stain the nuclei and in the presence of Gemini-561 dye (0.2 µM) for 5 min, the living cells were washed twice with Opti-MEM and visualized in Opti-MEM or were fixed in 4% PFA in PBS for 5 minutes before being washed twice in PBS. The images were acquired in epifluorescence mode with a Nikon Ti-E inverted microscope, equipped with CFI Plan Apo  $\times$  60 oil (NA = 1.4) objective, and a Hamamatsu Orca Flash 4 sCMOS camera. The acquisition settings were: Hoechst (ex. 395 nm, em. 510 $\pm$ 42 nm), eGFP (ex: 470 nm,

em: 531±40 nm), G561/o-Coral complex (ex: 550 nm, em: 595±40 nm) and Alexa-647 (ex: 638 nm, em: LP 647 nm). The images were recorded using NIS Elements and then processed with Icy software.

**Microscopy imaging (project 'Squaraine dimers')**. Cells were seeded onto a chambered coverglass (IBiDi®) 24 h before the microscopy measurement. For imaging, the culture medium was removed, the attached cells were washed with Hank's Balanced Salt Solution (HBSS, Gibco-Invitrogen) and incubated with solutions of biotinylated probes (200 nM). In competition experiment, KB cells were pretreated with biotin (100 µM) for 30 min prior to incubation with corresponding biotinylated probe. Images were taken with Nikon Ti-E inverted microscope, equipped with CFI Plan Apo × 60 oil (NA = 1.4) objective, X-Light spinning disk module (CresOptics) and a Hamamatsu Orca Flash 4 sCMOS camera, was used. The microscopy settings were: Hoechst (ex. 405 nm, em. 510±42 nm), Squaraine (ex. 638 nm, em. 705±36 nm). The images were recorded using NIS Elements and then processed with Icy.

**Flow cytometry.** Cells were grown at 37° C in humidified atmosphere containing 5% CO<sub>2</sub> in 25cm<sup>2</sup> (Nunc™ EasYFlask, ThermoFisher). On the day of the analysis, the cells were washed and harvested. The cell suspension (3x10<sup>5</sup> cells/mL) was incubated with corresponding biotinylated probe (200 nM) for 5 min at room temperature and analyzed immediately using flow cytometry (MACSQuant VYB, Miltenyi Biotec).



## **7. Annexe**

### **Fluorogenic probes for imaging lipid droplets in cells and tissues**

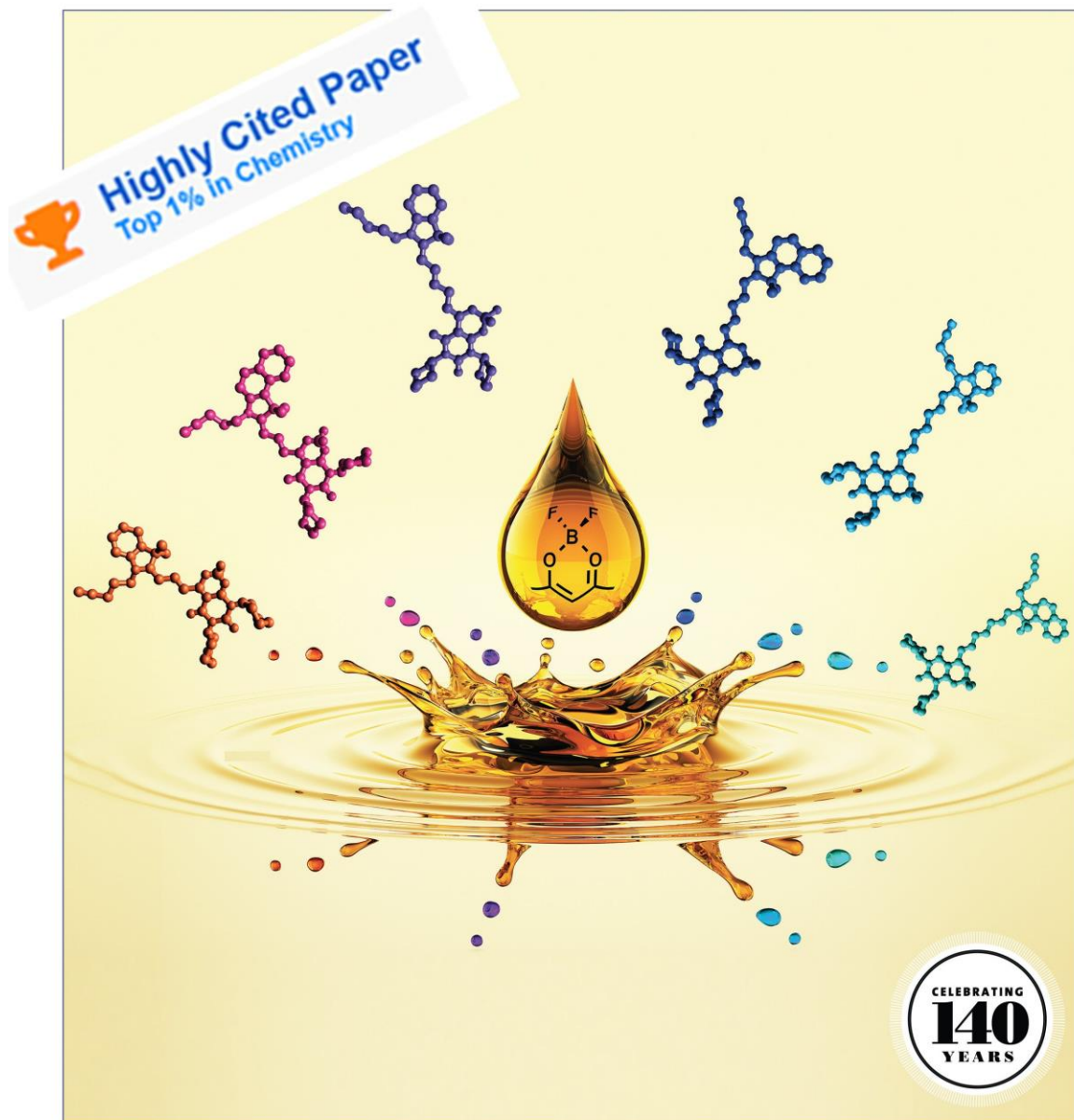
Apart from my main projects, I also contributed significantly to other research directions in our group, in particular, to fluorogenic probes for lipid droplets (LDs) imaging. As mentioned in chapter 1.1.4, we recently developed Statamerocyanine (SMCy) dyes as ultrabright and fluorogenic probes for multicolor imaging of LDs in cells and tissues.<sup>27</sup> In particular, I performed some spectroscopy studies revealing the mechanism of fluorogenicity of SMCy and cellular studies showing mechanism of uptake and specificity of SMCy towards LDs. The highlight of my work in this project was deciphering cell-cell communication in form of LDs exchange labelled with SMCy of two orthogonal colors.

Additionally, we summarized recent progress in the field of LDs fluorescent probes in published review<sup>153</sup>, where I was the major contributor. I also contributed (cytotoxicity experiments) to development of blue-emitting fluorogenic molecule (DAF<sup>154</sup>) probing LDs polarity and heterogeneity in cells.

April 25, 2018  
Volume 140  
Number 16  
pubs.acs.org/JACS

# J | A | C | S

JOURNAL OF THE AMERICAN CHEMICAL SOCIETY



 ACS Publications  
Most Trusted. Most Cited. Most Read.

[www.acs.org](http://www.acs.org)

## Ultrabright and Fluorogenic Probes for Multicolor Imaging and Tracking of Lipid Droplets in Cells and Tissues

Mayeul Collot,<sup>\*,†</sup> Tkhe Kyong Fam,<sup>†</sup> Pichandi Ashokkumar,<sup>†</sup> Orestis Faklaris,<sup>‡</sup> Thierry Galli,<sup>§,||</sup> Lydia Danglot,<sup>§,||</sup> and Andrey S. Klymchenko<sup>†</sup>

<sup>†</sup>Laboratoire de Biophotonique et Pharmacologie, CNRS UMR 7213, Université de Strasbourg, Faculté de Pharmacie, 74, Route du Rhin, 67401 Illkirch, France

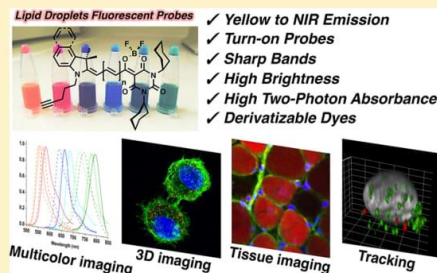
<sup>‡</sup>ImagoSeine Core Facility, Institut Jacques Monod, Université Paris Diderot/CNRS, UMR 7593, 15 Rue Hélène Brion, 75205 Paris CEDEX 13, France

<sup>§</sup>INSERM U894, Centre de Psychiatrie et Neurosciences, “Membrane Traffic in Health and Diseased Brain” Team, 102-108 Rue de la Santé, 75014 Paris, France

<sup>||</sup>Université Paris Descartes, 75014 Paris, France

### Supporting Information

**ABSTRACT:** Lipid droplets (LDs) are intracellular lipid-rich organelles that regulate the storage of neutral lipids and were recently found to be involved in many physiological processes, metabolic disorders, and diseases including obesity, diabetes, and cancers. Herein we present a family of new fluorogenic merocyanine fluorophores based on an indolenine moiety and a dioxaborine barbiturate derivative. These so-called StatoMerocyanines (SMCy) fluoresce from yellow to the near-infrared (NIR) in oil with an impressive fluorescence enhancement compared to aqueous media. Additionally, SMCy display remarkably high molar extinction coefficients (up to 390 000 M<sup>-1</sup> cm<sup>-1</sup>) and high quantum yield values (up to 100%). All the members of this new family specifically stain the LDs in live cells with very low background noise. Unlike Nile Red, a well-known lipid droplet marker, SMCy dyes possess narrow absorption and emission bands in the visible, thus allowing multicolor imaging. SMCy proved to be compatible with fixation and led to high-quality 3D images of lipid droplets in cells and tissues. Their high brightness allowed efficient tissue imaging of adipocytes and circulating LDs. Moreover their remarkably high two-photon absorption cross-section, especially SMCy5.5 (up to 13 300 GM), as well as their capacity to efficiently fluoresce in the NIR region led to two-photon multicolor tissue imaging (liver). Taking advantage of the available color palette, lipid droplet exchange between cells was tracked and imaged, thus demonstrating intercellular communication.



### INTRODUCTION

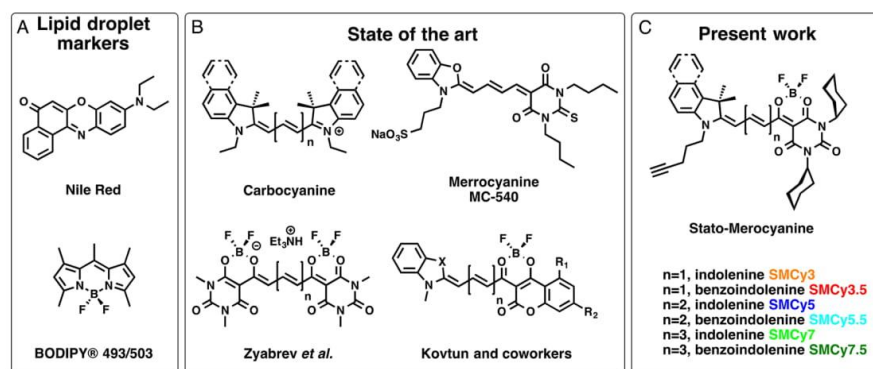
Lipid droplets (LDs), also known as adiposomes or lipid bodies, are intracellular lipid-rich organelles that regulate the storage of neutral lipids including triglycerides and cholesterol esters.<sup>1,2</sup> Although LDs are mainly found in adipose tissue, nearly all cells are capable of storing lipids in these reservoirs, as this ability to store such a source of metabolic energy is crucial for survival.<sup>3,4</sup> It has been shown that cells other than adipocytes can form LDs as a response to a stress.<sup>4</sup> In recent years, LDs have drawn considerable attention, as they were found to be involved in many physiological processes including membrane synthesis and trafficking,<sup>5,6</sup> protein degradation,<sup>7</sup> but also inflammation<sup>8</sup> and pathologies including obesity, diabetes, and atherosclerosis<sup>9</sup> as well as viral replication<sup>10</sup> and cancer.<sup>11–13</sup> Therefore, imaging and tracking LDs is of prior importance. The rise of fluorescence imaging techniques stimulated the development of efficient molecular fluorescent probes specific for various physiological phenomena or able to specifically stain organelles including mitochondria, the nucleus,

and the plasma membrane.<sup>14–17</sup> In particular, one should mention fluorogenic probes that turn on their emission only after binding to the biological target, allowing background-free bioimaging.<sup>18–21</sup> However, examples of probes for LDs are still limited. Nile Red is a commercially available solvatochromic dye (Figure 1A) that was reported to stain LDs<sup>22</sup> and is the most commonly used LD marker. However, Nile Red suffers from several drawbacks: (i) nonspecific staining of other organelles (particularly lipid membranes) and (ii) broad absorption and emission spectra that limit the possibilities for multicolor imaging. Although recent efforts have been made to develop new LD probes based on various fluorophores including BODIPY (496/503,<sup>23</sup> 540<sup>24</sup>), AIEgens (aggregation induced emission fluorophores),<sup>25</sup> azafluorenone,<sup>26</sup> and benzothiadiazole,<sup>27</sup> only one example of a far-red-emitting LD marker was recently published.<sup>28</sup> However, the latter, based

Received: December 21, 2017

Published: February 15, 2018





**Figure 1.** (A) Commonly used LD probes. (B) Structure of various cyanines including the pioneer work on dioxaborine cyanines. (C) The six members of the StatoMerocyanines family.

on AIEgen, suffers from a broad emission spectrum emitting in both the far-red and near-infrared (NIR) regions and was used at high concentrations (10  $\mu$ M) probably due to its low quantum yield (0.05). In addition to emitting at the far-red and NIR wavelengths, the probe should be excitable at an optical window suitable for optimal tissue penetration (670–910 nm), which can be achieved from the two-photon excitable fluorescent probes with high two-photon-absorption (TPA) cross-section.<sup>29</sup> Therefore, bright and selective LD probes, compatible in multicolor imaging and operating in the far-red and NIR region, are in demand. Cyanines are one of the most famous and widely used fluorophores. Their emission color spans from the visible range to the NIR and constitute a platform for biomolecule labeling<sup>30</sup> and for the development of chemosensors for bioimaging.<sup>31,32</sup> Among cyanines one can differentiate the cationic carbocyanines and the noncharged merocyanines (Figure 1B). Carbocyanines are bright fluorophores, but their positive charge can be detrimental for designing specific fluorescent probes, as it has a tendency to provoke attraction toward negatively charged biomolecules (proteins and biomembranes) by opposite charge attractions. Merocyanines circumvent this problem, as the core fluorophore generally does not display any charges. Unfortunately, in many cases merocyanines are nonphotostable. Conversely, BODIPY are generally neutral, bright, and stable fluorophores due to their planar structures stabilized by a difluoroboron complex that reduces the rotations and vibrations. Consequently, fluorophores including merocyanines stabilized with a rigid boron bridge (dioxaborine complex) are drawing continuous attention in the field of bioimaging. Indeed these fluorophores served in cellular imaging either in molecular form<sup>33–35</sup> or incorporated in nanoparticles,<sup>36–38</sup> as well as for imaging the nucleus,<sup>39</sup> sensing oxygen<sup>40,41</sup> or DNA,<sup>42,43</sup> and imaging ligand receptor interactions.<sup>44</sup>

Barbiturate-based merocyanines have been known for decades. MC-540 (Figure 1B), a fluorescent probe for lipid membranes,<sup>45</sup> is a figurehead of this family. Although it is a bright red emitting fluorophore, it suffers from a low photostability and high phototoxicity. The dioxaborine barbituryl moiety was first introduced by Zybrev et al. and served to develop bright negatively charged dyes (Figure 1B).<sup>46</sup> In parallel, fluorophores incorporating dioxaborine complexes and indolenine moieties were intensively studied by Kovtun and co-workers (Figure 1B).<sup>47–49</sup>

Herein, we developed a family of new merocyanine fluorophores based on an indolenine moiety and a dioxaborine barbiturate derivative separated by a polymethyne chain of various lengths (Figure 1). These so-called StatoMerocyanines (SMCy) have a fluorescence spanning from yellow to the NIR and display a bright fluorescence due to high molar extinction coefficients and high quantum yield values. SMCy are fluorogenic, as they are nonemissive in aqueous media and stain the lipid droplets of cells in a bright and selective manner. Additionally, due to their narrow absorption and emission bands, SMCy showed their efficiency in multicolor imaging compared to Nile Red and successfully served for multicolor tissue imaging. SMCy dyes were shown to possess remarkably high two-photon absorption cross-sections (up to more than 10 000 GM) and were successfully used in multicolor two-photon excitation (TPE) tissue imaging. Finally we used SMCy probes to track two populations of LDs, and we demonstrated that LD exchange occurs between cells.

## ■ RESULTS AND DISCUSSION

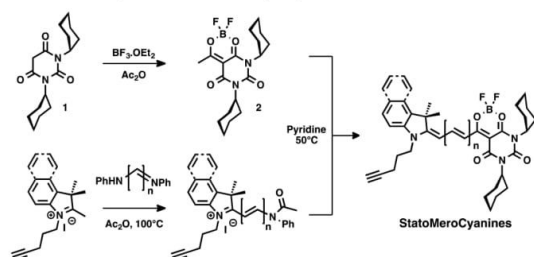
**Synthesis.** Herein we introduce the SMCy fluorophores that arose from the stabilization of a barbiturate-based merocyanine by a dioxaborine bridge that fixes (*stativus* in latin) a degree of rotation of the fluorophore and therefore enhances its brightness (Figure 1C).

SMCy dyes are composed of (1) an indolenine moiety bearing a pentynyl tail that reinforces the hydrophobicity and allows the opportunity to functionalize the dye by copper(I)-catalyzed alkyne–azide cycloaddition (CuAAC), if needed, and (2) a dioxaborine barbituric moiety bearing two cyclohexyl rings that were chosen for their hydrophobicity and bulkiness that can help in preventing formation of nonemissive H-aggregates by  $\pi$ -stacking.<sup>50–52</sup>

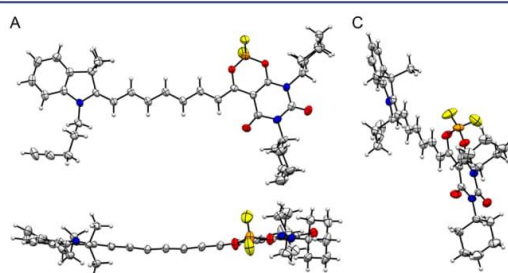
The synthesis involved the acetylation of dicyclohexylbarbituric acid **1** in the presence of  $\text{BF}_3 \cdot \text{OEt}_2$  to form the dioxaborine barbituryl **2** (Scheme 1). The latter was then condensed on hemicyanines with various polymethyne chain lengths in pyridine to provide the SMCy dyes.

The same nomenclature as the one introduced by Ernst et al. for the carbocyanines<sup>53</sup> was used. The suffix number indicates the number of carbons separating the indole ring from the dioxaborine barbituryl and the “.S” depicts the use of the extended benzoindolenine moiety instead of indolenine (Figure 1C). The obtained fluorophores were characterized by <sup>1</sup>H, <sup>13</sup>C,

## Scheme 1. Synthesis of SMCy Dyes

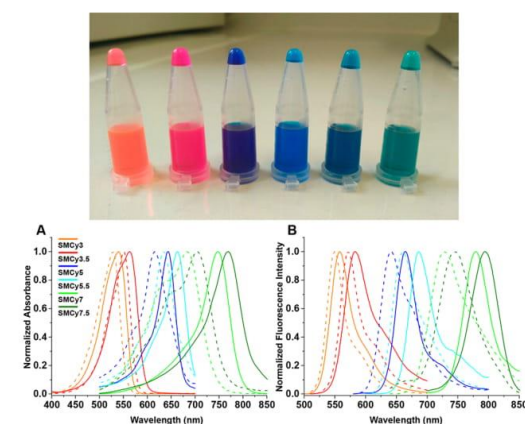


and  $^{19}\text{F}$  NMR as well as high-resolution mass spectrometry. Additionally, the structure of SMCy7 was confirmed by X-ray diffraction and showed a planar structure with minimal twist of the polymethine chain as well as the bulkiness induced by the cyclohexyl moieties (Figure 2).



**Figure 2.** ORTEP representation of SMCy7 obtained by X-ray diffraction of a single crystal. (A) Side view, (B) top view, and (C) diagonal view. For clarity, heteroatoms were colored: Boron (orange), nitrogen (blue), oxygen (red), and fluorine (yellow).

**Spectroscopic Studies.** The photophysical properties of SMCy dyes were first studied in 10 various organic solvents with increasing polarity from toluene to DMSO and are summarized in Table 1 (all spectra are available in the Supporting Information). First, the SMCy mostly displayed narrow absorption and emission spectra and their absorption wavelengths span from 526 to 770 nm and their emission wavelengths from 550 to 794 nm (Figure 3). The use of the benzindolenine instead of the indolenine moiety, giving rise to the “.5” markers, typically provoked a  $\sim 20$  nm bathochromic



**Figure 3.** Top: Picture of the SMCy in solution (200  $\mu\text{M}$  in DMSO). From left to right: SMCy3, SMCy3.5, SMCy5, SMCy5.5, SMCy7, and SMCy7.5. Bottom: Normalized absorption spectra (A) and normalized emission spectra (B) of SMCy3 to SMCy7.5 in toluene (dashed lines) and in DMSO (solid lines). Concentrations were 1  $\mu\text{M}$  for SMCy3, SMCy3.5, SMCy7, and SMCy7.5 and 200 nM for SMCy5 and SMCy5.5. Excitations were 490 nm for SMCy3 and SMCy3.5, 560 nm for SMCy5, 590 nm for SMCy5.5, 630 nm for SMCy7, and 630 and 690 nm for SMCy7.5.

shift in both absorption and emission spectra. All SMCy exhibited positive solvatochromism, as their absorption and emission spectra were significantly red-shifted in polar solvents. Moreover, the solvatochromic effect, as well as the Stokes shift, increases with the length of the polymethine chain. Indeed, whereas SMCy3 displayed only a 8 nm bathochromic emission shift from toluene to DMSO, SMCy7 displayed a 52 nm shift (Table 1). Weak solvatochromism for the dyes with shorter conjugation (SMCy3 and SMCy3.5) can be explained by their merocyanine type of electronic distribution, where the electronic charge of indolenine nitrogen is transferred to the barbiturate unit in the ground state without further strong change in the excited state. However, for dyes exhibiting longer conjugation this ground-state charge transfer is probably less pronounced, so that they undergo charge transfer in the excited state. The latter increases the dipole moment on electronic excitation, leading to significant solvatochromism in absorption and fluorescence, in line with the earlier report on other

**Table 1.** Photophysical Properties of SMCy Dyes in Toluene and DMSO<sup>a</sup>

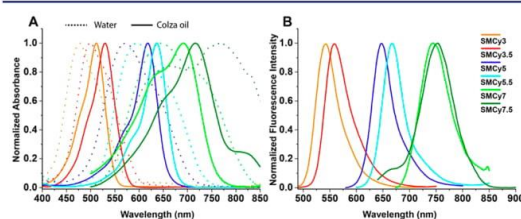
probe	solvent	$\lambda_{\text{Abs}}$ (nm)	$\epsilon$ ( $\text{M}^{-1} \text{cm}^{-1}$ )	$\text{fwhm}_{\text{Abs}}$ (nm)	$\lambda_{\text{Em}}$ (nm)	$\text{fwhm}_{\text{Em}}$ (nm)	Stokes shift	QY ( $\phi$ )	brightness ( $\epsilon \times \phi$ )
SMCy3	toluene	528	94 000	52	550	33	22	0.3	28 200
	DMSO	540	77 300	57	558	40	18	0.37	28 600
SMCy3.5	toluene	551	124 000	51	572	37	21	0.68	84 320
	DMSO	563	110 000	64	583	56	20	0.18	19 800
SMCy5	toluene	615	203 000	61	642	63	27	0.35	71 000
	DMSO	644	363 000	42	665	39	21	0.42	152 400
SMCy5.5	toluene	634	168 000	52	661	53	27	0.65	109 200
	DMSO	663	241 200	50	687	44	24	0.26	62 700
SMCy7	toluene	683	109 000	114	727	84	44	0.21	22 890
	DMSO	748	231 000	67	779	49	31	0.33	76 230
SMCy7.5	toluene	705	89 000	116	746	67	41	0.18	16 020
	DMSO	770	183 000	78	794	53	24	0.18	32 900

<sup>a</sup> $\epsilon$  is the molar extinction coefficient, fwhm is the full width at half-maximum, and the Stokes shifts are expressed in nm.



asymmetrical cyanine dyes.<sup>54</sup> SMCy3 and SMCy3.5 also slightly differ from the other SMCy. First, whereas hypochromism was observed with increasing solvent polarity, the other SMCy displayed hyperchromism (Figures and Tables S1 to S6). Then, whereas other SMCy have shortened full widths at half-maximum (fwhm) with increasing solvent polarity, SMCy3 and SMCy3.5 did not display significant changes in their spectra. In addition, the brightness of SMCy is highly dependent on the nature of the solvent; they generally display high molar extinction coefficients ranging from 77 000 to an impressive value of 394 000 M<sup>-1</sup> cm<sup>-1</sup>, and their quantum yield values vary from 0.07 to 0.76 (see S1). Among the SMCy dyes, SMCy5 was found to be the brightest, with a maximum brightness of 227 500 M<sup>-1</sup> cm<sup>-1</sup> in dichloromethane (DCM), competing with the squaraine dyes<sup>55</sup> and thus constituting one of the brightest small molecules to date.<sup>56,57</sup> As mentioned earlier, the photostability of merocyanine can be a limitation for their use in bioimaging; therefore the photostability of SMCy dyes was evaluated. SMCy3 and SMCy5 were continuously irradiated in a cuvette for 1 h in dioxane. Their photostability was monitored and compared to fluorophores with similar spectral properties: MC-540, a barbiturate-based merocyanine, and DID, a far-red-emitting carbocyanine 5,<sup>2,3</sup> respectively. The results showed that both SMCy3 and SMCy5 displayed higher photostability compared to the merocyanine and the carbocyanine (Figure S7).

In the light of the high lipophilic nature of the SMCy dyes, their photophysical properties were measured in oils—colza oil (mainly composed of unsaturated long chain triglycerides) and in Labrafac (a medium chain triglyceride containing oil)—as well as in water. The results showed first that, despite their high lipophilic nature, SMCy can be suspended in water: their significantly broadened absorption spectra suggest that they actually form soluble aggregates (Figure 4A and Figures S12



**Figure 4.** Normalized absorption spectra of SMCy dyes (1  $\mu$ M) in water (dashed lines) and in colza oil (solid lines) (A) and their normalized emission spectra in oil (B). Excitation wavelength was 480 nm for SMCy3 and SMCy3.5, 560 nm for SMCy5, 590 nm for SMCy5.5, and 630 nm for SMCy7 and SMCy7.5.

and S13). Consequently, SMCy are nonemissive in water (Figure 4B). In oils, SMCy dyes display high brightness similar to or higher than those obtained in nonpolar solvents, thus providing impressive fluorescence enhancements (up to 1700-fold) compared to water (Table 2). It is also noteworthy that, in these lipophilic environments, all the SMCy dyes are efficiently excited with the commonly available lasers (488, 560, and 630 nm). Remarkably, SMCy5 and SMCy5.5 exhibit the highest brightness, especially in Labrafac, reaching high quantum yields of 0.74 and 1.00, respectively. Higher quantum yields of SMCy in oils compared to organic solvents could be linked to additional restriction of intramolecular rotation in viscous media, in line with a previous report on a push–pull

dioxaborine derivative.<sup>44</sup> The photostability in oil of SMCy3 and SMCy5 was compared to commercially available LD markers, namely, BODIPY 493/503 and Nile Red. Whereas BODIPY 493/503 and SMCy5 lost 15% fluorescence intensity after 1 h of irradiation in the cuvette, SMCy3 displayed high photostability similar to Nile Red (Figure S8). According to the present results, SMCy appeared as suitable candidates to probe neutral lipophilic environments such as LDs in cells, using fluorescence microscopy techniques.

## CELLULAR EXPERIMENTS

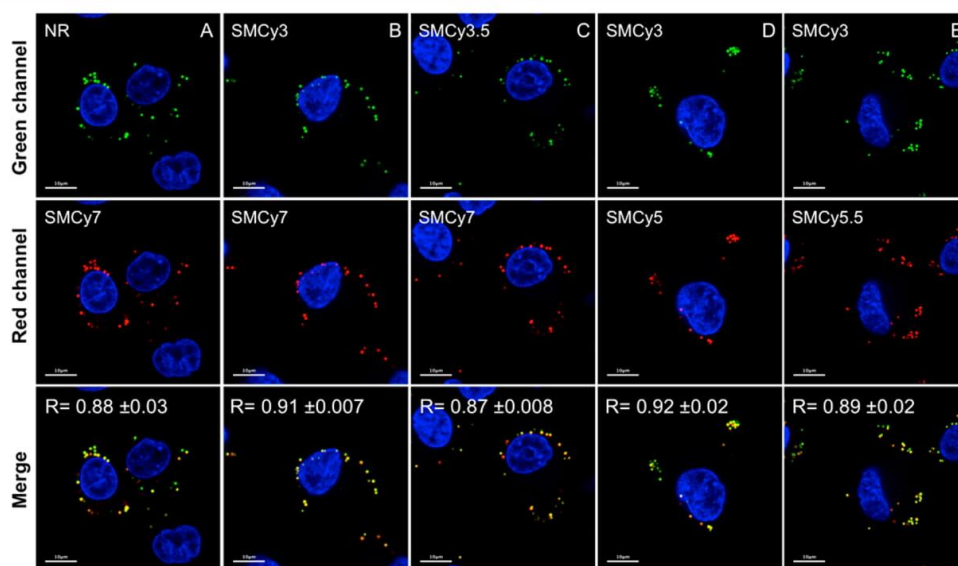
**Localization of SMCy.** On the basis of their chemical properties described above, SMCy were incubated with KB cells (HeLa cells derivatives), and the fluorescence was monitored over the time. After several minutes of incubation in a starvation medium (Hank's balanced salt solution, HBSS), a selective and intense fluorescence was observed in vesicles slowly moving in the cytoplasm of the cells (see Movie 1). The slow motion of these vesicles as well as their size and spherical shape led us to think that it was LDs. In order to confirm our expectations, a colocalization experiment was performed using Nile Red as a reference LD marker. As mentioned, Nile Red is a solvatochromic fluorophore with broad emission spectra and therefore can hardly be used for colocalization experiments with green- and red-emitting fluorophores due to crosstalk between the imaging channels. Thus, after checking that crosstalk did not occur in the NIR channel (excitation at 635 nm) (Figure S9), SMCy7 that emits in the NIR (730 nm) was first colocalized with Nile Red (Figure 5A). The results confirmed a clear colocalization between Nile Red and SMCy7. Consequently SMCy7 was used in combination with SMCy3 and SMCy3.5 (Figure 5B and C), and SMCy3 was then used to colocalize with SMCy5 and SMCy5.5 (Figure 5D and E). These experiments confirmed that the SMCy dyes selectively mark the LDs. In order to investigate the mechanism of staining, images have been taken at various times of incubation on live and fixed cells. The results showed a fluorescence enhancement in the LDs over time in both cells (see Figure S10). Consequently, we assume that SMCy accumulate in the LDs in a passive manner. The SMCy were also compared to BODIPY 493/503. Although SMCy dyes require incubation at 37 °C, BODIPY 493/503 immediately stained the LDs after addition (Figure S11). After investigating, we discovered that BODIPY 493/503 did not aggregate in water at concentrations as high as 1  $\mu$ M with a quantum yield of 1.0. In sharp contrast, the turn-on SMCy LD sensors are nonfluorescent in water due to aggregation in the form of soluble nanoparticles even at 100 nM concentration (Figures S12 and S13). Additionally, BODIPY 493/503 was shown to be less selective (Figures S14 and S15) and less photostable than SMCy3 (Figure S16). At this stage of our work, cytotoxicity studies (MTT assay) were performed, and the results showed that at concentrations up to 5  $\mu$ M the SMCy did not lead to any significant cytotoxicity (Figure S17).

**3D Imaging.** 3D imaging with confocal microscopy is normally obtained by reconstruction of Z-stack images, and therefore it involves multiple laser scans of the sample leading to the photobleaching of the fluorophores. Therefore, high brightness and photostability of the fluorescent markers are required to obtain good-quality 3D images. Here, SMCy allowed 3D imaging of the LDs, revealing their subcellular localization in KB cells (Figure 6). In these experiments a 21  $\mu$ m Z-stack was performed using 70 slices with three averaged

Table 2. Photophysical Properties of StatoMerocyanines in Triglyceride Oils<sup>a</sup>

	solvent	$\lambda_{\text{Abs}}/\lambda_{\text{Em}}$ (nm)	Stokes shift (nm)	$\epsilon$ ( $\text{M}^{-1} \text{cm}^{-1}$ )	fwhm Abs/Em (nm)	QY ( $\phi$ )	brightness ( $\epsilon \times \phi$ )	$\Delta F/F_0$
SMCy3	colza oil	512/541	29	82 900	54/49	0.21	17 409	423
	Labrafac	511/542	31	87 300	55/51	0.25	21 825	515
SMCy3.5	colza oil	530/559	29	100 000	57/55	0.40	40 000	674
	Labrafac	529/557	28	108 000	57/58	0.54	58 320	827
SMCy5	colza oil	618/648	30	256 000	54/46	0.60	153 600	1088
	Labrafac	617/644	27	270 000	54/49	0.74	199 800	1700
SMCy5.5	colza oil	638/662	24	169 000	52/42	0.74	125 060	1535
	Labrafac	636/660	24	194 000	52/43	1.00	194 000	1535
SMCy7	colza oil	692/744	52	44 000	134/75	0.42	18 480	55
	Labrafac	687/735	48	44 000	137/84	0.44	19 360	79
SMCy7.5	colza oil	716/753	37	103 000	125/71	0.19	19 570	203
	Labrafac	712/743	31	83 000	132/74	0.2	16 600	145

<sup>a</sup>Concentration of the dyes was 1  $\mu\text{M}$ .  $\Delta F/F_0$  is the fluorescence enhancement between oil and water.  $F_0$  is the maximum fluorescence intensity in water.



**Figure 5.** Laser scanning confocal microscopy of live KB cells incubated for 3 h in opti-MEM in the presence of fluorophores (500 nM). NR is Nile Red. Hoechst (5  $\mu\text{g}/\text{mL}$ ) was used to stain the nucleus (blue). Green channel: excitation at 488 nm, emission collected: 500–600 nm. Red channel (middle images): excitation at 635 nm, emission collected: 645–800 nm. The merged images correspond to the combined green and red channels. The displayed Pearson's correlation coefficient, which denotes the goodness of colocalization, is the average of at least five different images. Scale bar is 10  $\mu\text{m}$ .

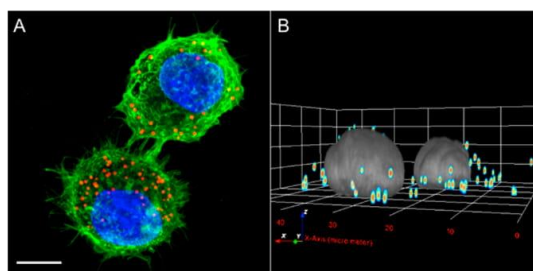
acquisitions per slice, representing 210 scans per image. The 3D images displayed clear and bright spots (see Movie 2). Although SMCy7 and SMCy7.5 emit in the NIR region, where the sensitivity of the microscope's detectors is lower, they led to excellent 3D imaging due to the efficient excitation with the 635 nm laser (Figure S18).

**Multicolor Imaging.** Multicolor imaging provides important information and allows revealing different cellular processes on the same image. For this reason, it is of great importance to have available fluorescent markers with various excitation and emission wavelengths combined with narrow emission bands, thus allowing their concomitant use. A lambda scan experiment with the help of a laser scanning confocal microscope allowed us to obtain *in situ* emission spectra of Nile Red and the SMCy in the cell lipid droplets (Figure 7). It is noteworthy that unlike Nile Red, which was bright enough in

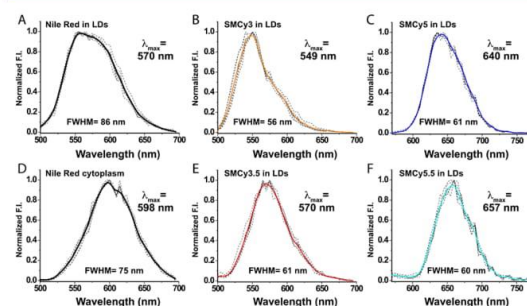
the cytoplasm to provide a lambda scan emission spectrum (Figure 7D), SMCy dyes, due to their selective LD staining, did not provide any spectrum. The obtained spectra were similar to those obtained in oil using a fluorometer, confirming the localization of the dyes in a nonpolar lipophilic environment. In cell lipid droplets the SMCy displayed significantly narrower spectra than Nile Red ( $\lambda_{\text{Em max}} = 570 \text{ nm}$ ; fwhm = 86 nm).

For these reasons, we assumed that, unlike Nile Red, SMCy might be efficient in multicolor imaging. First, according to the obtained spectroscopic data in oil, we show that unlike Nile Red, SMCy3 can be used as a green LD marker and in combination with red-emitting fluorophores (Figure 8). Although both Nile Red and SMCy3 are efficiently excited with a 488 nm laser (70% and 62%, respectively), SMCy3 displays much narrower absorption and emission bands (fwhm: Abs 54 nm, Em: 49 nm), with a blue-shifted emission ( $\lambda_{\text{Em max}}$

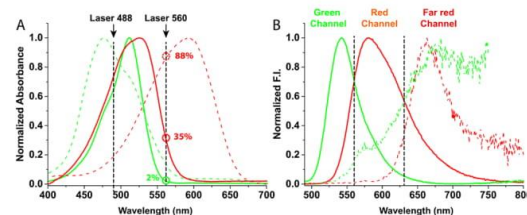




**Figure 6.** Images of KB cells incubated with SMCy5 (500 nm) then fixed (PFA 4%), obtained by a 21  $\mu\text{m}$  depth Z-stack (300 nm stepwise) showing the repartition of the LDs within the cells. (A) Maximum intensity projection image displaying the LDs (red), the nucleus (blue), and the plasma membrane (green). (B) 3D image displaying the LDs (colored spots) and the nucleus (gray). The nucleus was stained with 5  $\mu\text{g}/\text{mL}$  Hoechst; the plasma membrane, with MemBright-488 (200 nM). Incubation with SMCy5 was done in opti-MEM for 3 h. Scale bar is 10  $\mu\text{m}$ .



**Figure 7.** Normalized emission spectra of Nile Red in LDs (A) and in the cytoplasm (D) and normalized emission spectra of SMCy dyes in LDs obtained by lambda scan (5 nm steps) with a laser confocal microscope using the 488 nm laser (for Nile Red, SMCy3, and SMCy3.5) and the 560 nm laser (for SMCy5 and SMCy5.5). Dashed lines correspond to five independent measurements (normalized), which were averaged and smoothed (solid lines).



**Figure 8.** Normalized absorption (A) and emission (B) spectra of SMCy3 (green lines) and Nile Red (red lines) in colza oil (solid line) and in water (dashed lines). Concentration was 1  $\mu\text{M}$ , and excitation was 480 nm.

541 nm), compared to Nile Red ( $\lambda_{\text{Em max}}$  581 nm, fwhm: Abs 80 nm, Em 84 nm). Moreover, Nile Red is also efficiently excited with the 560 nm laser (35% in oil), whereas SMCy3 is only excited at 2%. Finally, due to its solvatochromism, Nile Red when irradiated at 560 nm in water is efficiently excited (88%, Figure 8). Consequently, when Nile Red localizes in more polar cellular environments in the cytoplasm, its

excitation and emission spectra shift to the red, thus leading to crosstalk in the red and far-red channels.

As a green-emitting LD probe was available (BODIPY 493/503),<sup>23</sup> we then wanted to show that SMCy3.5 could be used as a red LD marker in multicolor imaging. To this aim, KB cells were incubated with (1) Hoechst, a blue marker for the nucleus; (2) MitoTracker Green, a green marker for mitochondria,<sup>23</sup> and (3) Nile Red or SMCy3.5, as a red-emitting LD markers. Prior to the imaging, dSQ12S was used as a far-red plasma membrane marker.<sup>55</sup> The obtained multicolor imaging showed that whereas Nile Red displayed distinct crosstalk in the green channel (LDs) and in the far-red channel (cytoplasm), SMCy3.5, due to its sharp absorption and emission bands and its weak solvatochromism, displayed no crosstalk, thus leading to clear multicolor imaging (Figure 9). This set of experiments showed that, unlike Nile Red, SMCy3 and SMCy3.5 can be used as green and red lipid droplet markers, respectively, with no crosstalk in fluorescent microscopy imaging.

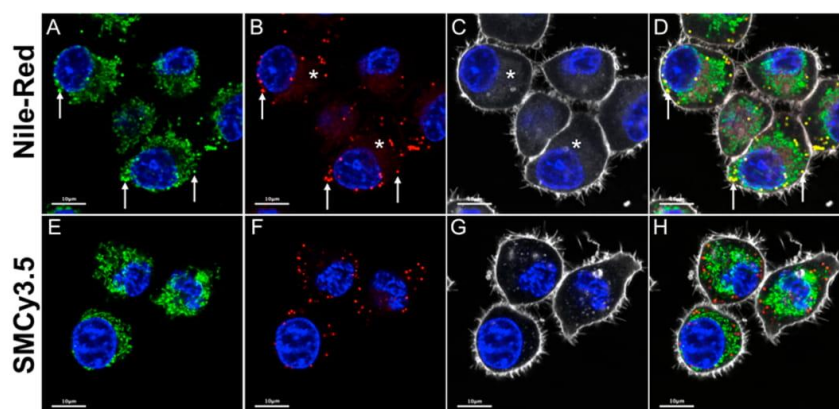
## TISSUE IMAGING

**Monophoton Imaging.** Taking advantage of the brightness and specificity of the SMCy dyes, we investigated their ability to reveal lipid-rich structures/organelles in multicolor tissue imaging. For this purpose white adipose tissue of a mouse was incubated in the presence of Hoechst (nuclear staining), MemBright-488 (plasma membrane staining), and SMCy3.5. The tissue was then imaged by fluorescence microscopy (Figure 10). The obtained images allowed for 3D imaging (see Movie 3) as well as a projection that clearly shows that SMCy3.5 selectively stained the fat reservoirs of adipocytes (Figure 10B) but also revealed the presence of dots in the blood vessel that could be identified as low density lipoprotein (LDL) or circulatory LDs (Figure 10D).

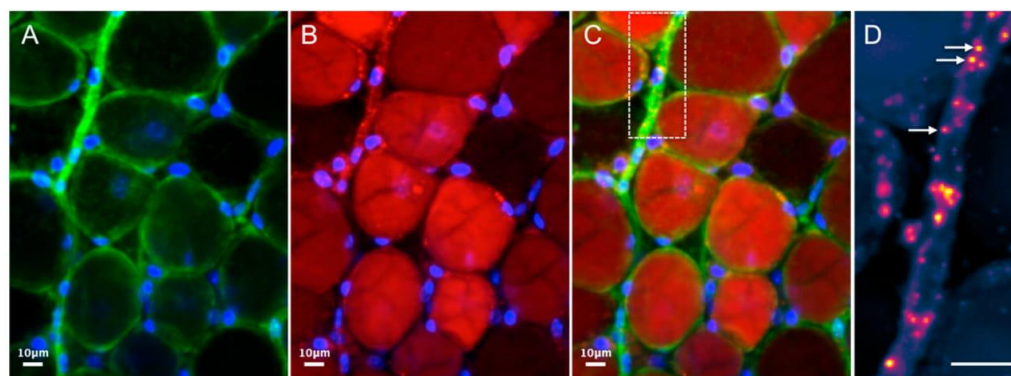
**TPE Imaging.** Two-photon excitation microscopy imaging has received considerable attention as an advanced optical technique in the field of bioimaging for several reasons.<sup>58–60</sup> First, TPE microscopy inherently provides three-dimensional sectioning since the excitation occurs only at the focal point of the sample. In addition, due to its localized excitation with a NIR laser, it reduces the photobleaching, the phototoxicity, and the cell autofluorescence. It also ensures deeper tissue penetration, which enables experiments on thicker live samples. For these reasons, there is a high demand for bright fluorescent probes with high TPE cross-section values, leading to ongoing efforts to design such fluorophores.<sup>61</sup> To this aim, we first measured the TPA cross-section of SMCy dyes in a cuvette. The two-photon nature of the absorption process was verified from the quadratic dependence of the observed emission intensity vs excitation power (Figures S19 to S22).

The results showed that although SMCy7 and SMCy7.5 did not display TPE below 900 nm, the remaining dyes displayed higher TPA cross-section values (Figure 11) when compared to the known LDs probes.<sup>29</sup> While SMCy3 only possessed a modest TPA cross-section (178 GM at 690 nm), SMCy3.5 reaches values as high as 2400 GM (at 760 nm). Finally, SMCy5 and SMCy5.5 displayed impressive TPA cross-section values, 6250 GM at 740 nm and 13 330 GM at 770 nm, respectively.

Interestingly, SMCy5.5 displays a very high cross-section values of 10 400 GM at the commonly used 820 nm wavelength (used for rhodamine-based probes). Combined with its exceptional brightness in lipidic environments, this probe



**Figure 9.** Laser scanning confocal microscopy images of live KB cells incubated 3 h with Hoechst ( $5 \mu\text{g}/\text{mL}$ , blue color), MitoTracker Green (200 nM), Nile Red, or SMCy3.5 (500 nM). Plasma membrane was stained prior to imaging with dSQ12S, a far-red-emitting plasma membrane probe (gray color). Green channel: A and E (ex: 488 nm, em: 498–540 nm); red channel: B and F (ex: 561 nm, em: 570–640 nm); far-red channel: C and G (ex: 632 nm, em: 640–750 nm). (D and H) Merging of A, B, C and E, F, G, respectively. White arrows and white stars indicate the crosstalk of Nile Red in the green and the far-red channels, respectively. Scale bar is  $10 \mu\text{m}$ .



**Figure 10.** Images of adipocytes in mouse adipose tissue ( $25 \mu\text{m}$  depth) obtained by a maximum projection of a Z-stack (100 slices of 250 nm depth each). The tissue was incubated overnight at  $4 \text{ }^\circ\text{C}$  with Hoechst ( $5 \mu\text{g}/\text{mL}$ , nuclei in blue color), (A) MemBright-488 ( $5 \mu\text{M}$ , plasma membrane in green color), and (B) SMCy3.5 ( $5 \mu\text{M}$ , lipid domains, red color). (C) Merged channels. (D) Zoomed-in region of interest in C showing the lipid-rich vesicles (white arrows) in the blood vessel. Scale bar is  $10 \mu\text{m}$ .

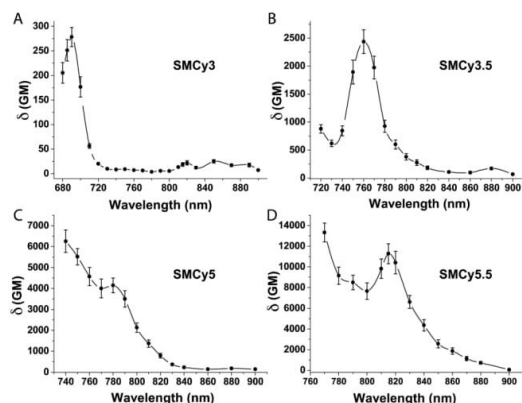
constitutes one of the brightest two-photon-excitable fluorophores to date.

In light of the exceptional TPE brightness of SMCy, we mapped the lipophilic environment in tissue using multicolor TPE imaging. A mouse liver slice was incubated with Hoechst and SMCy5.5. Using only one excitation wavelength (810 nm), we obtained a  $52 \mu\text{m}$  depth 3D image where the nuclei, the lipid droplets, and the collagen fibers (using second-harmonic generation<sup>62,63</sup>) were clearly distinguished (Figure 12). The image revealed large lipid-rich vesicles that are not present in the collagen-rich region.

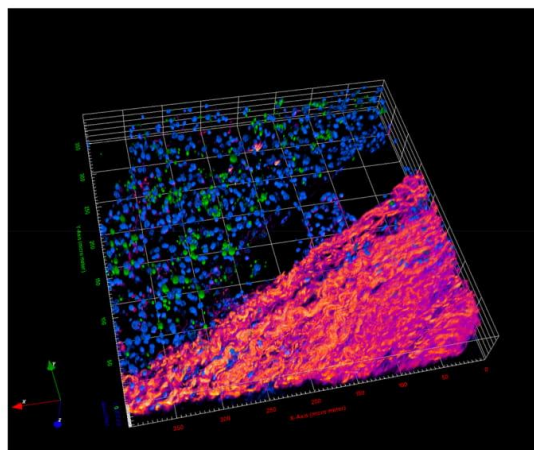
**Lipid Droplet Exchange.** With multicolor LD markers in hand it becomes possible to monitor the fate of different LD populations in the same experiment. First, as a control experiment, a batch of KB cells was incubated with the same concentration of SMCy3.5 and SMCy5.5 for 3 h. The images showed a strong colocalization of the dyes within the same LDs, showing nondifferent uptake of the dyes (Figure 13A). It is noteworthy that although almost no green LDs were

detected, some red LDs were observed. This might be attributed to the Förster resonance energy transfer (FRET) phenomenon occurring in some LDs and driven by the close proximity between the donor SMCy3.5 and the acceptor SMCy5.5. Second, two batches of KB cells were independently incubated with SMCy3.5 and SMCy5.5 for 3 h. After several washing steps to avoid cross contamination, the cells were trypsinized, mixed, plated, and cocultured. After 12 h, two distinct populations of cells could be identified by the color of their lipid droplets with no apparent exchange of LDs between cells showing no colocalization of the dyes (Figure 13B). This tends to prove that the dyes are confined in the LDs with no leakage out of the cells. After 48 h, the same sample of cells was visualized. Remarkably, within some cells, LDs of two distinct colors were observed (Figure 13C and D), suggesting LDs exchange between cells. In these cells the colocalization of the two dyes was also minimal, proving that the dyes were not exchanged within the LDs of the same cell. These experiments illustrate that KB cells are able to actively exchange LDs or that





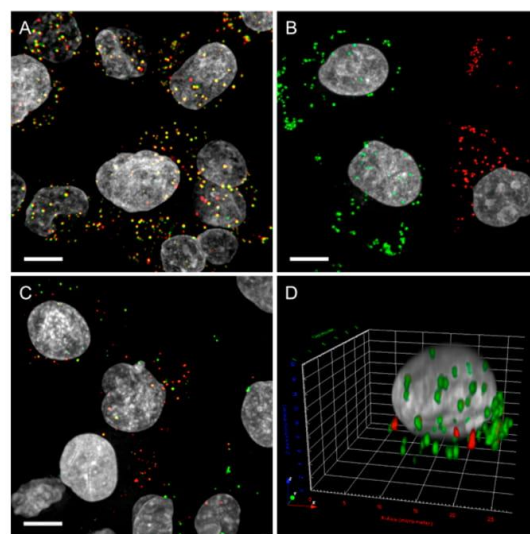
**Figure 11.** Two-photon-excitation cross-section spectra of SMCy dyes in DMSO. The reference used for measurements was rhodamine B. GM is Goepfert–Mayer unit.



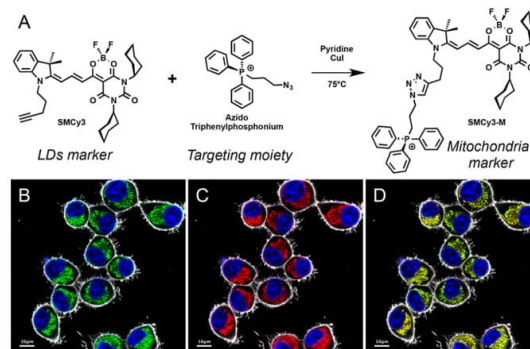
**Figure 12.** Two-photon excitation 3D imaging (52  $\mu\text{m}$  depth) of a mouse liver slice, incubated with Hoechst (5  $\mu\text{g}/\text{mL}$ ) and SMCy5.5 (5  $\mu\text{M}$ ). The nuclei are displayed in blue, the lipid droplets in green, and the collagen fibers in fire color. The two-photon excitation wavelength was 810 nm.

cells are able to take up LDs released from dead cells. Owing to their photostability, brightness, and ability to remain in the LDs with no leakage, SMCy dyes constitute a robust tool to stain and track LDs in bioimaging experiments.

**Functionalization of SMCy.** SMCy dyes bear an alkyne side arm for further functionalization by CuAAC click chemistry. Indeed, owing to their interesting physicochemical properties (neutral fluorophores, high brightness, narrow absorption and emission bands), SMCy can serve as a platform to develop various selective markers or sensors. As an example, SMCy3 was involved in a click chemistry reaction with an azido triphenylphosphonium cation that is known to target mitochondria (Figure 14A).<sup>64</sup> The obtained SMCy3-M was then incubated for 1 h at 200 nM in the presence of KB cells. After checking that crosstalk did not occur (Figure S23), the colocalization experiment using MitoTracker Red led to a measured Pearson correlation coefficient of 0.96, thus proving



**Figure 13.** Images of live KB cells obtained by a maximum projection of a 20  $\mu\text{m}$  depth Z-stack. (A) Co-incubation of SMCy3.5 (1  $\mu\text{M}$ ) and SMCy5.5 (1  $\mu\text{M}$ ) for 3 h then imaged after 18 h. (B) Coculture of KB cells incubated with SMCy3.5 and SMCy5.5 separately after 18 h. (C) Coculture of KB cells incubated with SMCy3.5 and SMCy5.5 separately after 48 h. (D) 3D view of a KB cell displaying LDs from another cell. Scale bar is 10  $\mu\text{m}$ .



**Figure 14.** Synthesis of SMCy3-M by CuAAC click chemistry (A). Laser scanning confocal microscopy of live KB cells incubated for 1 h with Hoechst (5  $\mu\text{g}/\text{mg}$ , blue color) and 200 nM of SMCy3-M (B) and MitoTracker Red CMXRos (C). (D) Merged images. The plasma membrane was stained with dSQ12S (20 nM) prior to imaging. Excitation wavelength and emission range for imaging in nm: Hoechst (405, 415–470), SMCy3-M (488, 498–560), MitoTracker Red (561, 571–620), and dSQ12S (635, 645–700). Scale bar is 10  $\mu\text{m}$ .

that SMCy3 was successfully converted into an efficient and selective mitochondria marker (Figure 14).

## CONCLUSION

In this work we designed and synthesized a new family of dioxaborine-containing merocyanines, named StatoMerocyanines. These fluorophores cover the visible spectrum and the NIR region with high brightness and photostability. SMCy were found to be soluble in both water and oil and exhibit an

impressive fluorescence enhancement between these two environments. When incubated with cells, SMCy selectively stain the LDs in a bright and background-free manner, thus leading to high-quality images including multicolor 3D images. Owing to their sharp absorption and emission bands, SMCy were shown to be efficient for multicolor imaging (four colors) more particularly in the green and red channels, where Nile Red leads to crosstalk. SMCy served in multicolor tissue imaging, where it successfully probed the lipophilic environment in adipocytes as well as revealed circulating lipidic vesicles in the blood vessel. SMCy3 to SMCy5.5 are two-photon excitable, displaying impressive cross-section values up to 13 300 GM, making SMCy5.5 one of the brightest TPE fluorophores to date. These properties allowed obtaining multicolor TPE 3D-images revealing the distribution of lipidic vesicles in liver tissue. Benefiting from the different available colors of these LDs marker, we used them to track two populations of LDs and showed that LD exchange occurs between cells. Finally, we demonstrated that, using their alkyne group, SMCy can be derivatized by click chemistry in order to obtain fluorescent probes with different targeting specificity. This study should encourage the continuation in the development of new dioxaborine merocyanines for bioimaging.

## MATERIALS AND METHODS

**Synthesis.** All starting materials for synthesis were purchased from Alfa Aesar, Sigma-Aldrich, or TCI Europe and used as received unless stated otherwise. NMR spectra were recorded on a Bruker Avance III 400 MHz spectrometer. Mass spectra were obtained using an Agilent Q-TOF 6520 mass spectrometer. Synthesis of all new compounds is described in the Supporting Information.

**Spectroscopy.** The water used for spectroscopy was Milli-Q water (Millipore), and all the solvents were spectral grade. Colza oil, mainly composed of unsaturated lipids ~93%, was obtained from Lesieur (Asnières-sur-Seine, France). Labrafac (Labrafac Lipophile WL 1349) is composed of medium-chain triglycerides and caprylic/capric triglyceride and was obtained from Gattefossé (Saint Priest, France). Absorption and emission spectra were recorded on a Cary 400 Scan ultraviolet-visible spectrophotometer (Varian) and a FluoroMax-4 spectrofluorometer (Horiba Jobin Yvon) equipped with a thermostated cell compartment, respectively. For standard recording of fluorescence spectra, the emission was collected 10 nm after the excitation wavelength. All the spectra were corrected from the wavelength-dependent response of the detector. Quantum yields were determined by comparison with a reference according to their excitation and emission wavelengths: rhodamine 6G in water<sup>65</sup> was used as the reference for SMCy3 and SMCy3.5, DID in MeOH<sup>66</sup> for SMCy5 and SMCy5.5, and rhodamine 800 in EtOH<sup>67</sup> for SMCy7. For SMCy7.5, rhodamine 800 and indocyanine green in MeOH<sup>68</sup> were used depending on the solvent.

**TPE Measurements.** Two-photon-absorption cross-section measurements were performed using rhodamine B in methanol as a calibration standard according to the method of Webb et al.<sup>69</sup> Two-photon excitation was provided by an InSight DS+ laser (Spectra Physics) with a pulse duration of 120 fs. The laser was focused with an achromatic lens ( $f = 2$  cm) in a cuvette containing the dye (SMCy3: 6.04  $\mu$ M, SMCy3.5: 2.08  $\mu$ M, SMCy5: 1.06  $\mu$ M, SMCy5.5: 285 nM) in DMSO, and the spectra were recorded with a fibered spectrometer (Avantes) by collecting the fluorescence emission at 90° with a 20× Olympus objective.

**X-ray Crystallography.** Single crystals of SMCy7 were obtained by a solvent exchange method using DCM as solvent and cyclohexane as antisolvent.<sup>70</sup> X-ray diffraction data collection was carried out on a Bruker APEX II DUO Kappa-CCD diffractometer equipped with an Oxford Cryosystem liquid N<sub>2</sub> device, using Mo K $\alpha$  radiation ( $\lambda = 0.71073$  Å). The crystal–detector distance was 38 mm. The cell parameters were determined (APEX2 software)<sup>71</sup> from reflections

taken from three sets of 12 frames, each at 10 s exposure. The structure was solved by direct methods using the program SHELXS-2013.<sup>72</sup> The refinement and all further calculations were carried out using SHELXL-2013.<sup>73</sup> The H atoms were included in calculated positions and treated as riding atoms using SHELXL default parameters. The non-H atoms were refined anisotropically, using weighted full-matrix least-squares on  $F^2$ . A semiempirical absorption correction was applied using SADABS in APEX2;<sup>71</sup> transmission factors:  $T_{\min}/T_{\max} = 0.6697/0.7456$ .

**Confocal Imaging.** KB cells (ATCC CCL-17) were grown in minimum essential medium (MEM, Gibco-Invitrogen) with 10% fetal bovine serum (FBS, Lonza), 1% nonessential amino acids (Gibco-Invitrogen), 1% MEM vitamin solution (Gibco-Invitrogen), 1% L-glutamine (Sigma-Aldrich), and 0.1% antibiotic solution (gentamicin, Sigma-Aldrich) at 37 °C in a humidified atmosphere containing 5% CO<sub>2</sub>. Cells were seeded onto a chambered coverglass (IBiDi) at a density of  $5 \times 10^4$  cells/well 24 h before the microscopy measurement. For imaging, the culture medium was removed and the attached cells were washed with Opti-MEM (Gibco-Invitrogen). Next, the cells were incubated in Opti-MEM with Hoechst (5  $\mu$ g/mL) to stain the nuclei and in the presence of Nile Red or SMCy dyes (0.5 to 1  $\mu$ M) for 3 h. The living cells were washed three times with HBSS and visualized in HBSS or were fixed in 4% paraformaldehyde (PFA) in HBSS for 5 min before being washed three times in HBSS. The images were processed with Icy software.<sup>74</sup> The plasma membrane was stained with dsQ12S (200 nM) prior to imaging.<sup>55</sup> The Pearson's correlation coefficient was measured with the Colocalization Studio plugin<sup>75</sup> using Icy software.

**Cytotoxicity Assay.** The cytotoxicity assay of the SMCy dyes was quantified by the MTT assay (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide). A total of  $1 \times 10^4$  KB cells/well were seeded in a 96-well plate 24 h prior to the cytotoxicity assay in Dulbecco's modified Eagle's medium (Gibco Lifetechnologies, DMEM) complemented with 10% FBS, gentamicin (100  $\mu$ g/mL), L-glutamine (2 mM), nonessential amino acids (1 mM), and MEM vitamin solution (1%) and were incubated in a 5% CO<sub>2</sub> incubator at 37 °C. After medium removal, 100  $\mu$ L of DMEM containing 5, 1, or 0.2  $\mu$ M SMCy (SMCy3, SMCy3.5, SMCy5, SMCy5.5, SMCy7, and SMCy7.5) was added to the KB cells and incubated for 3 h at 37 °C (5% CO<sub>2</sub>). As a control, for each 96-well plate, the cells were incubated with DMEM containing the same percentage of DMSO (0.5% v/v) as the solution with the tested dyes or with Triton 1% as a positive control of cytotoxicity. After 3 h of dye incubation, the medium was replaced by 100  $\mu$ L of a mix containing DMEM + MTT solution (diluted in phosphate-buffered saline (PBS) beforehand), and the cells were incubated for 4 h at 37 °C. Then, 75  $\mu$ L of the mix was replaced by 50  $\mu$ L of DMSO (100%) and gently shaken for 15 min at room temperature in order to dissolve the insoluble purple formazan reduced in living cells. The absorbance at 540 nm was measured (absorbances of the dyes at 540 nm were taken into account). Each concentration of dye was tested in sextuplicate in three independent assays. For each concentration, we calculated the percentage of cell viability in reference to the control DMEM + 0.5% DMSO.

**Tissue Imaging.** C57BL6/J mice were maintained on a 12 h light–dark cycle with *ad libitum* access to food and water. All animal work was conducted following a protocol approved by the ethical committee. Adult C57BL6 mice were euthanized by CO<sub>2</sub> administration. Liver and fat pads were immediately dissected and washed in PBS at 4 °C. The liver was sliced into 1 mm slices on a Leica vibratome. Tissues were placed overnight at 4 °C in 1 mL of freshly prepared solution of dye. Adipose tissue: SMCy3.5: 5  $\mu$ M, MemBright-488, 5  $\mu$ M, and Hoechst: 5  $\mu$ g/mL in PBS. Liver tissue: SMCy5.5: 5  $\mu$ M and Hoechst: 5  $\mu$ g/mL in PBS. The tissue samples were washed three times with PBS before being placed in a homemade glass chamber, allowing the imaging on both sides of the slides. Briefly a 1 cm diameter hole was breached in the glass slide using a diamond drill bit. The 1 mm tissue slice was inserted in the hole, covered by a coverslip (170 nm, #1.5) on each slide, and sealed with Picodent Twinsil. The adipose tissue slices were imaged with a Leica DMi8 microscope equipped with a spinning disc CSU-X1. The microscopy



settings were as follows: Hoechst (ex 405 nm, em 450 ± 60), MemBright-488 (ex 488 nm, em 510 ± 23), SMCy3.5 (ex 561 nm, em 590 ± 35). The images were acquired with a 0.37 pixel width in XY and 0.5 μm in Z with a PL APO 20×/0.75 oil immersion Leica objective. The images were processed with Icy software. For two-photon imaging see below.

**Two-Photon Imaging.** Two-photon fluorescence microscopy imaging was performed on a TriMScope II (LaVision BioTec, Bielefeld, Germany) built on an Olympus BX51 microscope. The excitation source was a tunable InSight laser (Spectra Physics) (range 690–1300 nm, repetition rate 80 MHz). For our experiments the excitation wavelength was fixed at 810 nm. Each line was scanned twice (line average 2) with a speed of 0.4 μs per pixel using a 20× Plan Apo NA:1 water immersion objective, with a working distance of 2 mm (Carl Zeiss, Germany). A Z-stack of 66 images was acquired with a Z-step of 0.8 μm for a total depth of 52 μm. The fluorescence signal was detected backward with sensitive gallium arsenide phosphide detectors (GaAsp). We used bandpass filters for the detection, more precisely for the collagen fiber and the SHG signal at 400–410 nm, the Hoechst at 435–485 nm, and the SMCy5.5 at 655–685 nm. The images were processed with Icy software.

**Lipid Droplet Exchange Experiments.** KB cells were seeded on T-25 flasks at a density of  $1.5 \times 10^6$  cells before the day of the LD staining. The next day the medium was removed and the adhered cells were rinsed with PBS two times. Then, the cells were separately incubated with SMCy3.5 or SMCy5.5 (1 μM in OptiMEM 0.5% DMSO) at 37 °C, 90% humidity, 5% CO<sub>2</sub> for 3 h. After incubation the media containing the dyes were discharged, and KB cells were rinsed with PBS two times, growth medium to ensure a proper removal of the excess dyes, and PBS again. Cells were trypsinised and centrifuged at 15 000 rpm for 3 min, and the pellets of cells were resuspended in growth medium. The suspensions of cells were centrifuged again, and cells were resuspended in the growth medium. Both fractions of cells were thoroughly mixed together in the same amount ( $\sim 6 \times 10^6$  in total). The cells were seeded onto a chambered coverglass (IBiDi) at a density of  $1 \times 10^5$  cells/well. Prior to imaging, cells were incubated with Hoechst (5 μg/mL) for 15 min for nucleus staining. The cells were then fixed with PFA 4% for 5 min and washed three times with HBSS. Images were taken after 18 and 48 h post cell adherence with spinning-disk microscopy. For this purpose, a Nikon Ti-E inverted microscope, equipped with CFI Plan Apo 60× oil (NA = 1.4) objective, an X-Light spinning disk module (CresOptics), and a Hamamatsu Orca Flash 4 sCMOS camera, was used. The microscopy settings were as follows: Hoechst (ex 405 nm, em 510 ± 42 nm), SMCy3.5 (ex 532 nm, em 600 ± 25 nm), SMCy5.5 (ex 638 nm, em 705 ± 36 nm). The images were recorded using NIS Elements and then processed with Icy software.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.7b12817.

Synthesis, protocols, characterizations, and spectra (PDF)

Movies (ZIP)

X-ray crystallographic data (CIF)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*mayeul.collot@unistra.fr

### ORCID

Mayeul Collot: 0000-0002-8673-1730

Andrey S. Klymchenko: 0000-0002-2423-830X

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank Romain Vauchelles for his assistance at the PIQ imaging platform and Bogdan Andreiuk for providing us with compound **1**. We thank the national research group GDR Imabio for promoting interdisciplinary meetings. We acknowledge the Service de Chimie Analytique (SCA) for LC-MS and RMN analyses. We also acknowledge the ImagoSeine core facility of the Institut Jacques Monod, members of IBiSA and the France-BioImaging infrastructure, and supported by ANR grant (ANR-10-INBS-04). This work was supported by ERC Consolidator grant BrightSens 648528, and ANR BrightRiboProbes (ANR-16-CE11-0010) and ANR grant (ANR-10-INBS-04).

## ■ REFERENCES


- (1) Farese, R. V.; Walther, T. C. *Cell* **2009**, *139*, 855–860.
- (2) Martin, S.; Parton, R. G. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 373–378.
- (3) Walther, T. C.; Farese, R. V., Jr. *Annu. Rev. Biochem.* **2012**, *81*, 687–714.
- (4) Thiam, A. R.; Farese, R. V., Jr.; Walther, T. C. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 775–786.
- (5) Zehmer, J. K.; Huang, Y.; Peng, G.; Pu, J.; Anderson, R. G. W.; Liu, P. *Proteomics* **2009**, *9*, 914–921.
- (6) Blom, T.; Somerharju, P.; Ikonen, E. *Cold Spring Harbor Perspect. Biol.* **2011**, *3*, a00471310.1101/cshperspect.a004713.
- (7) Olzmann, J. A.; Richter, C. M.; Kopito, R. R. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 1345–1350.
- (8) Bozza, P. T.; Viola, J. P. B. *Prostaglandins, Leukotrienes Essent. Fatty Acids* **2010**, *82*, 243–250.
- (9) Kraemer, N.; Farese, R. V.; Walther, T. C. *EMBO Mol. Med.* **2013**, *5*, 973–983.
- (10) Herker, E.; Harris, C.; Hernandez, C.; Carpentier, A.; Kaehlecke, K.; Rosenberg, A. R.; Farese, R. V.; Ott, M. *Nat. Med.* **2010**, *16*, 1295–1298.
- (11) Tirinato, L.; Pagliari, F.; Limongi, T.; Marini, M.; Falqui, A.; Seco, J.; Candeloro, P.; Liberale, C.; Di Fabrizio, E. *Stem Cells Int.* **2017**, *2017*, 1.
- (12) Liu, Q.; Luo, Q.; Halim, A.; Song, G. *Cancer Lett.* **2017**, *401*, 39–45.
- (13) Abramczyk, H.; Surmacki, J.; Kopeć, M.; Olejnik, A. K.; Lubecka-Pietruszewska, K.; Fabianowska-Majewska, K. *Analyst* **2015**, *140*, 2224–2235.
- (14) Lavis, L. D. *Annu. Rev. Biochem.* **2017**, *86*, 825–843.
- (15) Zhu, H.; Fan, J.; Du, J.; Peng, X. *Acc. Chem. Res.* **2016**, *49*, 2115–2126.
- (16) Yang, Y.; Zhao, Q.; Feng, W.; Li, F. *Chem. Rev.* **2013**, *113*, 192–270.
- (17) Kobayashi, H.; Ogawa, M.; Alford, R.; Choyke, P. L.; Urano, Y. *Chem. Rev.* **2010**, *110*, 2620–2640.
- (18) Nadler, A.; Schultz, C. *Angew. Chem., Int. Ed.* **2013**, *52*, 2408–2410.
- (19) Li, X.; Gao, X.; Shi, W.; Ma, H. *Chem. Rev.* **2014**, *114*, 590–659.
- (20) Su, D.; Teoh, C. L.; Wang, L.; Liu, X.; Chang, Y.-T. *Chem. Soc. Rev.* **2017**, *46*, 4833–4844.
- (21) Klymchenko, A. S. *Acc. Chem. Res.* **2017**, *50*, 366–375.
- (22) Greenspan, P.; Mayer, E. P.; Fowler, S. D. *J. Cell Biol.* **1985**, *100*, 965–973.
- (23) Life Technologies <https://www.thermofisher.com/>.
- (24) Spandl, J.; White, D. J.; Peychl, J.; Thiele, C. *Traffic* **2009**, *10*, 1579–1584.
- (25) Wang, Z.; Gui, C.; Zhao, E.; Wang, J.; Li, X.; Qin, A.; Zhao, Z.; Yu, Z.; Tang, B. Z. *ACS Appl. Mater. Interfaces* **2016**, *8*, 10193–10200.
- (26) Sharma, A.; Umar, S.; Kar, P.; Singh, K.; Sachdev, M.; Goel, A. *Analyst* **2015**, *141*, 137–143.
- (27) Appelqvist, H.; Stranius, K.; Börjesson, K.; Nilsson, K. P. R.; Dryager, C. *Bioconjugate Chem.* **2017**, *28*, 1363–1370.

- (28) Kang, M.; Gu, X.; Kwok, R. T. K.; Leung, C. W. T.; Lam, J. W. Y.; Li, F.; Tang, B. Z. *Chem. Commun.* **2016**, 52, S957–S960.
- (29) Jiang, M.; Gu, X.; Lam, J. W. Y.; Zhang, Y.; Kwok, R. T. K.; Wong, K. S.; Tang, B. Z. *Chem. Sci.* **2017**, 8, 5440–5446.
- (30) Gonçalves, M. S. T. *Chem. Rev.* **2009**, 109, 190–212.
- (31) Sun, W.; Guo, S.; Hu, C.; Fan, J.; Peng, X. *Chem. Rev.* **2016**, 116, 7768–7817.
- (32) Chan, J.; Dodani, S. C.; Chang, C. J. *Nat. Chem.* **2012**, 4, 973–984.
- (33) Kamada, K.; Namikawa, T.; Senatore, S.; Matthews, C.; Lenne, P.-F.; Maury, O.; Andraud, C.; Ponce-Vargas, M.; Le Guennic, B.; Jacquemin, D.; Agbo, P.; An, D. D.; Gauny, S. S.; Liu, X.; Abergel, R. J.; Fages, F.; D'Aléo, A. *Chem. - Eur. J.* **2016**, 22, 5219–5232.
- (34) Kim, E.; Felouat, A.; Zaborova, E.; Ribierre, J.-C.; Wu, J. W.; Senatore, S.; Matthews, C.; Lenne, P.-F.; Baffert, C.; Karapetyan, A.; Giorgi, M.; Jacquemin, D.; Ponce-Vargas, M.; Guennic, B. L.; Fages, F.; D'Aléo, A. *Org. Biomol. Chem.* **2016**, 14, 1311–1324.
- (35) Zhou, Y.; Chen, Y.-Z.; Cao, J.-H.; Yang, Q.-Z.; Wu, L.-Z.; Tung, C.-H.; Wu, D.-Y. *Dyes Pigm.* **2015**, 112, 162–169.
- (36) Pfister, A.; Zhang, G.; Zareno, J.; Horwitz, A. F.; Fraser, C. L. *ACS Nano* **2008**, 2, 1252–1258.
- (37) Contreras, J.; Xie, J.; Chen, Y. J.; Pei, H.; Zhang, G.; Fraser, C. L.; Hamm-Alvarez, S. F. *ACS Nano* **2010**, 4, 2735–2747.
- (38) Kerr, C.; DeRosa, C. A.; Daly, M. L.; Zhang, H.; Palmer, G. M.; Fraser, C. L. *Biomacromolecules* **2017**, 18, 551–561.
- (39) Pitter, D. R.; Brown, A.; Baker, J.; Wilson, J. *Org. Biomol. Chem.* **2015**, 13, 9477–9484.
- (40) DeRosa, C. A.; Samonina-Kosicka, J.; Fan, Z.; Hendargo, H. C.; Weitzel, D. H.; Palmer, G. M.; Fraser, C. L. *Macromolecules* **2015**, 48, 2967–2977.
- (41) DeRosa, C. A.; Seaman, S. A.; Mathew, A. S.; Gorick, C. M.; Fan, Z.; Demas, J. N.; Peirce, S. M.; Fraser, C. L. *ACS Sens.* **2016**, 1, 1366–1373.
- (42) Wilson, J. N.; Wigenius, J.; Pitter, D. R. G.; Qiu, Y.; Abrahamsson, M.; Westerlund, F. *J. Phys. Chem. B* **2013**, 117, 12000–12006.
- (43) Pitter, D. R. G.; Wigenius, J.; Brown, A. S.; Baker, J. D.; Westerlund, F.; Wilson, J. N. *Org. Lett.* **2013**, 15, 1330–1333.
- (44) Karpenko, I. A.; Niko, Y.; Yakubovskiy, V. P.; Gerasov, A. O.; Bonnet, D.; Kovtun, Y. P.; Klymchenko, A. S. *J. Mater. Chem. C* **2016**, 4, 3002–3009.
- (45) Easton, T. G.; Valinsky, J. E.; Reich, E. *Cell* **1978**, 13, 475–486.
- (46) Zybrev, K.; Doroshenko, A.; Mikitenko, E.; Slominskii, Y.; Tolmachev, A. *Eur. J. Org. Chem.* **2008**, 2008, 1550–1558.
- (47) Gerasov, A. O.; Shandura, M. P.; Kovtun, Y. P. *Dyes Pigm.* **2008**, 77, 598–607.
- (48) Gerasov, A. O.; Shandura, M. P.; Kovtun, Y. P. *Dyes Pigm.* **2008**, 79, 252–258.
- (49) Gerasov, A. O.; Zybrev, K. V.; Shandura, M. P.; Kovtun, Y. P. *Dyes Pigm.* **2011**, 89, 76–85.
- (50) Gieseck, R. L.; Mukhopadhyay, S.; Risko, C.; Marder, S. R.; Brédas, J.-L. *Chem. Mater.* **2014**, 26, 6439–6447.
- (51) Frischmann, P. D.; Würthner, F. *Org. Lett.* **2013**, 15, 4674–4677.
- (52) Ozdemir, T.; Atilgan, S.; Kutuk, I.; Yildirim, L. T.; Tulek, A.; Bayindir, M.; Akkaya, E. U. *Org. Lett.* **2009**, 11, 2105–2107.
- (53) Ernst, L. A.; Gupta, R. K.; Mujumdar, R. B.; Waggoner, A. S. *Cytometry* **1989**, 10, 3–10.
- (54) Padilha, L.; Webster, S.; Przhonska, O.; Hu, H.; Peceli, D.; Rosch, J.; Bondar, M.; Gerasov, A.; Kovtun, Y.; Shandura, M.; Kachkovski, A.; Hagan, D.; Stryland, E. W. V. *J. Mater. Chem.* **2009**, 19, 7503–7513.
- (55) Collot, M.; Kreder, R.; Tatars, A. L.; Patsenker, L. D.; Mely, Y.; Klymchenko, A. S. *Chem. Commun.* **2015**, 51, 17136–17139.
- (56) Lavis, L. D. *Biochemistry* **2017**, 56, 5165–5170.
- (57) Lavis, L. D.; Raines, R. T. *ACS Chem. Biol.* **2014**, 9, 855–866.
- (58) Helmchen, F.; Denk, W. *Nat. Methods* **2005**, 2, 932–940.
- (59) Svoboda, K.; Yasuda, R. *Neuron* **2006**, 50, 823–839.
- (60) Piston, D. W. *Trends Cell Biol.* **1999**, 9, 66–69.
- (61) Yao, S.; Belfield, K. D. *Eur. J. Org. Chem.* **2012**, 2012, 3199–3217.
- (62) Mohler, W.; Millard, A. C.; Campagnola, P. J. *Methods* **2003**, 29, 97–109.
- (63) Chen, X.; Nadiarynkh, O.; Plotnikov, S.; Campagnola, P. J. *Nat. Protoc.* **2012**, 7, 654–669.
- (64) Smith, R. A. J.; Porteous, C. M.; Gane, A. M.; Murphy, M. P. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, 100, 5407–5412.
- (65) Magde, D.; Rojas, G. E.; Seybold, P. G. *Photochem. Photobiol.* **1999**, 70, 737–744.
- (66) Texier, I.; Goutayer, M.; Da Silva, A.; Guyon, L.; Djaker, N.; Josserand, V.; Neumann, E.; Bibette, J.; Vinet, F. *J. Biomed. Opt.* **2009**, 14, 054005–11.
- (67) Alessi, A.; Salvalaggio, M.; Ruzzon, G. *J. Lumin.* **2013**, 134, 385–389.
- (68) Benson, R.; Kues, H. *Phys. Med. Biol.* **1978**, 23, 159–163.
- (69) Albota, M. A.; Xu, C.; Webb, W. W. *Appl. Opt.* **1998**, 37, 7352–7356.
- (70) Spingler, B.; Schnidrig, S.; Todorova, T.; Wild, F. *CrystEngComm* **2012**, 14, 751–757.
- (71) M86-E01078 APEX2 User Manual; Bruker AXS Inc.: Madison, WI, USA, 2006.
- (72) Sheldrick, G. M. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1990**, 46, 467–473.
- (73) Sheldrick, G. M. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **2008**, 64, 112–122.
- (74) Chaumont, F. de; Dallongeville, S.; Chenouard, N.; Hervé, N.; Pop, S.; Provoost, T.; Meas-Yedid, V.; Pankajakshan, P.; Lecomte, T.; Montagner, Y. L.; Lagache, T.; Dufour, A.; Olivo-Marin, J.-C. *Nat. Methods* **2012**, 9, 690.
- (75) Lagache, T.; Sauvonnnet, N.; Danglot, L.; Olivo-Marin, J.-C. *Cytometry, Part A* **2015**, 87, S68–S79.



Review

# Recent Advances in Fluorescent Probes for Lipid Droplets

Tkhe Kyong Fam, Andrey S. Klymchenko and Mayeul Collot \* 

Nanochemistry and Bioimaging Group, Laboratoire de Bioimagerie et Pathologies, CNRS UMR 7021, Université de Strasbourg, Faculté de Pharmacie, 67401 Illkirch, France; tkhe-kyong.fam@unistra.fr (T.K.F.); andrey.klymchenko@unistra.fr (A.S.K.)

\* Correspondence: mayeul.collot@unistra.fr; Tel.: +33-3-6885-4266

Received: 19 August 2018; Accepted: 14 September 2018; Published: 18 September 2018



**Abstract:** Lipid droplets (LDs) are organelles that serve as the storage of intracellular neutral lipids. LDs regulate many physiological processes. They recently attracted attention after extensive studies showed their involvement in metabolic disorders and diseases such as obesity, diabetes, and cancer. Therefore, it is of the highest importance to have reliable imaging tools. In this review, we focus on recent advances in the development of selective fluorescent probes for LDs. Their photophysical properties are described, and their advantages and drawbacks in fluorescence imaging are discussed. At last, we review the reported applications using these probes including two-photon excitation, *in vivo* and tissue imaging, as well as LDs tracking.

**Keywords:** lipid droplets; fluorescent probes; bioimaging

## 1. Introduction

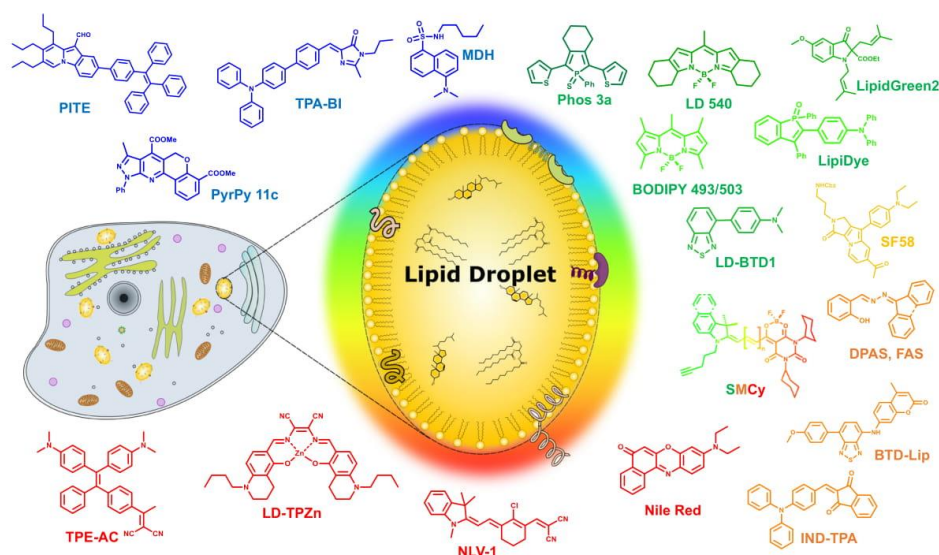
Lipid droplets (LDs) are the lipid storage cellular organelles consisting of a neutral lipid core, mainly containing cholesterol esters and triglycerides [1,2] covered by a phospholipid monolayer shell where a number of associated proteins were recently found (Figure 1). In particular, these include the PAT protein family, with PAT standing for Perilipin (PLIN), Adipocyte differentiation-related protein (ADRP; also called as adipophilin), and TIP47 (tail-interacting protein of 47 kDa) proteins [3]. LDs can primarily be found not only in the adipose tissue of eukaryotic organisms, but also in other types of cells, since LDs play a very important role in the regulation of cellular lipid metabolism [3].

Recently, LDs received considerable attention as they were shown to be involved in different cellular processes such as membrane formation, trafficking [4], and protein–protein interaction [5]. Moreover, the misregulation of LD functions can cause metabolic disorders such as obesity, fatty liver diseases, diabetes, atherosclerosis, and others [6].

In order to understand its role in different metabolic disorders and get insights on their interactions with other cellular compartments, it is of the highest importance to have reliable imaging tools for LDs. The visualization of cellular LDs can be carried out by label-free imaging techniques, such as conventional transmission light microscopy [7], direct organelle mass spectrometry [8], Raman microscopy [9], and coherent anti-Stokes Raman scattering microscopy [10]. These advanced techniques allowed us to study the biophysics of cellular LDs, but they required complex sample preparations and data analysis. In addition, the mentioned methodologies usually require cells to be fixed or LDs extraction, omitting the possibility of studying the real-time dynamics of LDs in the cellular environment. On the other hand, the fluorescence imaging techniques have proved to be powerful methods to study complex biological processes due to their high sensitivity [11]. It allows precisely imaging the localization, distribution, and biophysical properties of many intracellular compartments such as mitochondria, nucleus, plasma membrane, etc. [12–14].



In this review, we focus on recent advances in the field of selective LDs fluorescence imaging. In particular, the advantages and disadvantages of developed fluorescent probes to study LDs biology are discussed, as well as emerging probes that enable LD tracking in cells and tissues.



**Figure 1.** Schematic representation of a lipid droplet and the various multicolor liquid droplets (LDs) probes discussed in this review.

## 2. Common Methods and Probes in LDs Fluorescence Bioimaging

LD biology research is greatly assisted by chemical tools, allowing easy-to-handle manipulations to study intracellular bioprocesses in live cells. These compounds allowed scientists to image and track LDs under both physiological and pathophysiological conditions [3,15]. In this part, we will present the immunostaining targets and the commonly used molecular lipophilic dyes to identify the LDs in cells.

### 2.1. Fluorescence Immunostaining

The immunofluorescence staining technique is a golden standard biochemical method to detect target molecules in laboratory practice. The expression of PAT proteins in LDs permits visualizing LDs using the immunofluorescence technique, in particular if adipophilin is targeted, which is usually abundant in LDs [16]. Despite being a powerful and reliable method that is applicable to staining LDs, careful considerations should be taken when immunofluorescence staining is used. Common protocols require the fixation of cells. Methods of fixation can have a dramatic impact on LDs' morphology. In addition, some LD-associated proteins are weakly detected after the permeabilization step [16]. Thus, the detection of cellular LDs is more practical and feasible when fluorescent small molecules are used, as they stain the lipid core of native LDs. Moreover, since immunostaining requires fixation, it does not allow live cell studies and the tracking of LDs.

### 2.2. Diazo Dyes

Sudan III [17] and Oil Red O [18] (Figure 2) are diazo dyes that have been used as LD stains for decades, since they enable visualizing LDs using conventional bright-field (transmitted light) microscopy. Moreover, both dyes emit a red fluorescence, permitting the use of fluorescence microscopy to analyze LDs. The downfall of those dyes is their poor solubility, thus requiring the usage of ethanol

or isopropanol for staining. Such invasive staining conditions cause the disruption of native LDs and occasional fusion, despite cells being fixed [19,20].

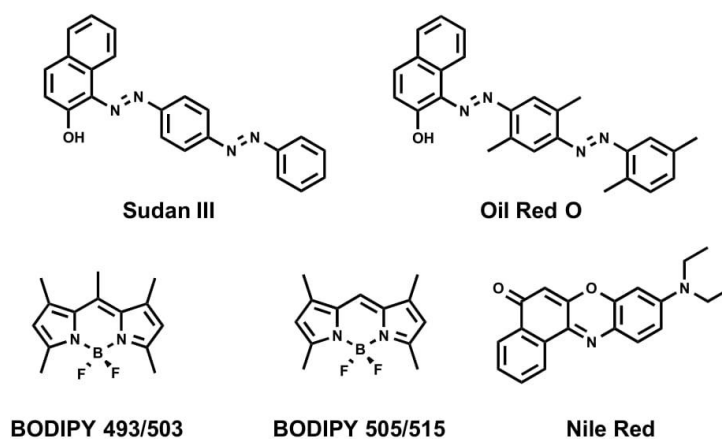


Figure 2. Chemical structures of commonly used fluorescent dyes.

### 2.3. Nile Red

Nile Red (Figure 2) is a commercially available fluorogenic dye that is commonly used to stain LDs in cells [21]. Nile Red exhibits solvatochromic properties fluorescing in most organic solvents and a lipidic environment, while its fluorescence is diminished in aqueous media [21]. Even though Nile Red is an easy-to-handle fluorescent LD marker, there are some limitations to be taken into consideration: (i) non-specific labeling of cellular lipid organelles, in particular intracellular membranes, and (ii) Nile red has broad absorption and emission spectra that lead to cross-talk in the red channel (typically: excitation between 530–560 nm), thus making it non-suitable for multicolor imaging with different dyes or fluorescent proteins.

### 2.4. BODIPY 493/503

Another commonly used commercial fluorescent dye is BODIPY 493/503 (Figure 2) [22]. A structurally close analogue of BODIPY 493/503, namely BODIPY 505/515, is also known to stain neutral lipids [22]. Although BODIPY 493/503 rapidly and reliably labels LDs with bright green fluorescence, it possesses limitations. First, it tends to produce a background signal in imaging due to its non-fluorogenic nature. Indeed, we recently showed that BODIPY 493/503 displays a quantitative quantum yield in aqueous media [23]. Then, we also showed that BODIPY 493/503 has limited photostability [23]. At last, BODIPY 493/503 displays a small Stokes shift (the difference between excitation maximum and emission maximum), which can cause cross-talk between the excitation source and the fluorescence emission. However, BODIPYs were shown to be more selective to LDs than Nile Red, and immediately stain LDs due to a better cell permeation [24,25].

Even though Oil Red O, Nile Red, and BODIPY 493/503 proved to be useful chemical tools to visualize LDs, their utility faces the previously mentioned limitations (specificity, cross-talk, photostability, small Stokes shifts, and background noise). Moreover, the analysis should be done with high precautions; appropriate controls are necessary to be performed to discriminate false positive fluorescence signals [19]. Consequently, these limitations led to the development of new LD fluorescent probes.

### 3. Recent Developments in LDs' Selective Fluorescent Probes

Recent efforts have been made to develop new probes based on various lipophilic fluorophores expanding the available toolbox to probe LDs. In particular, the smart design of probes improved organelle specificity, photophysical properties, cell permeability, cell toxicity, etc. In this review, we tried to compile an exhaustive list of fluorescent probes targeting LDs spanning their fluorescence from blue to near-infrared (NIR) region. For clarity, we grouped the fluorescent probes by their reported emission color in the less polar solvent as followed: blue ( $\lambda_{em} < 500$  nm), green ( $\lambda_{em} = 500\text{--}550$  nm), orange ( $\lambda_{em} = 550\text{--}600$  nm), and red to NIR ( $\lambda_{em} > 600$  nm).

#### 3.1. Blue Emitting LD Probes ( $\lambda_{em} < 500$ nm)

The color availability of LD markers provides the flexibility of setting multicolor fluorescence imaging. The gap in the blue region was recently filled by the development of blue-emitting LD probes.

To this end, Yang et al. reported the utility of commercially available fluorophore monodansylpentane (MDH) as a blue marker of LDs (Figure 3) [26]. MDH was challenged to be compatible in multicolor one-photon excitation (1PE) imaging. It was also compatible with two-photon excitation (2PE) imaging using 760-nm excitation.

Inspired by the successful utility of MDH as a blue LD marker, two series of dyes were developed based on dihydrochromenopyrrolopyridine and the  $\alpha$ -pyrazolopyridine core (Figure 3. PyrPy 10d, PyrPy 11c) as potential blue LD markers [27]. The dyes showed solvatochromic behavior. Bathochromic shifts were observed with an increasing polarity of solvents indicating a greater dipole moment in the excited state in comparison to the ground state of the molecules. This feature proved that the molecules undergo intramolecular charge transfer (ICT) and are associated with positive solvatochromism. It was noted that the introduction of the ester group into the pyridine core improved ICT and cell permeability. Moreover, the introduction of the second ester group dramatically increased selectivity toward LDs.

PITE [28] arose from the combination of pyroindole (PI) and tetraphenylethylene (TE) (Figure 3). PI emits strong green fluorescence in organic solvents, while its fluorescence is quenched in aqueous solution due to poor solubility and aggregation. On the other hand, TE is known to be an aggregation-induced emission (AIE) molecule [29]. Coupling PI and TE resulted in a molecule with new properties. Authors have reported PITE capacity to specifically detect LDs. The absorption spectra of PITE displays two bands at 320 and 420 nm as contributions of both moieties TE and PI respectively, while it emits at 490 nm and does not depend on excitation wavelength. Using *Saccharomyces cerevisiae* (as the yeast cell model) and HeLa cells (as the mammalian cell model), Sk et al. showed the compatibility of PITE as a blue fluorescent LD probe [28].

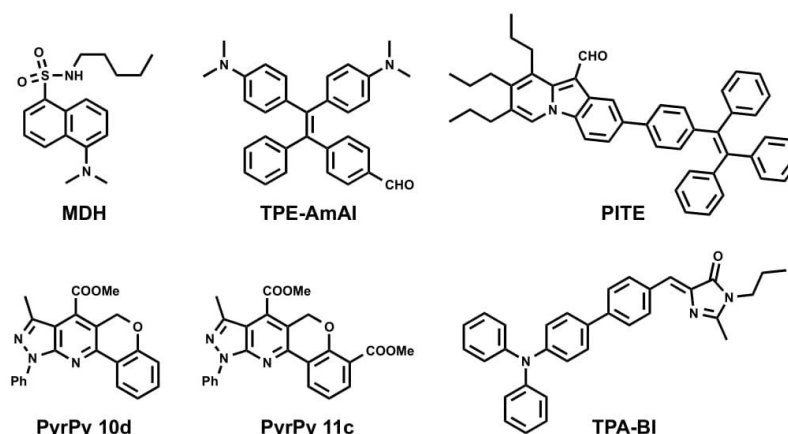


Figure 3. Chemical structures of blue emitting LD probes.



Other blue AIEgens that are specific for LDs were developed in the Tang group and named TPE-AmAl [30] and TPA-BI (Figure 3) [31]. TPE-AmAl consists of TE as an AIE-unit emitting in the blue region, the alkylamino group as an electron donor, and carbonyl as an electron acceptor to promote the intramolecular charge transfer (ICT) process. Although TPE-AmAl emits at 610 nm as aggregated form in aqueous conditions, it exhibits blue emission once in cellular LDs. This difference is attributed to the non-polar environment of LDs and solvatochromic properties of the dye (blue shift in less polar solvents). The LD-targeting property was confirmed by costaining with Nile Red [30]. TPA-BI bears triphenylamine (TPA) moiety as an electron donor and as a common unit for enhancing two-photon absorption (2PA) properties. On the other hand, benzylidene imidazolone (BI) was shown to be AIE-active. TPA-BI exhibited solvatochromic properties emitting from blue to red when the solvent changed from n-hexane (447 nm) to acetonitrile (619 nm), which was nearly covering the full visible spectrum. Authors have shown that TPA-BI can perform both in 1PE and 2PE fluorescence cell imaging. In 1PE imaging, TPA-BI also showed an excellent resistance to photobleaching in addition to high LD specificity. Since TPA-BI shows a 2PA cross-section ( $\delta_{2PA}$ ) value of up to 213 GM at 840 nm, the evaluation of its compatibility in 2PE cell imaging was also conducted. The result showed that TPA-BI performed better in 2PE in comparison to 1PE imaging.

Even though reported blue probes successfully stained intracellular LDs, they share common limitations such as limited brightness (due to low  $\epsilon$ ), a high background signal due to cell autofluorescence, and the considerable phototoxicity of ultraviolet/violet excitation.

### 3.2. Green Emitting LD Probes ( $\lambda_{em} = 500\text{--}550\text{ nm}$ )

Anh et al. developed green-emitting LDs probes LipidGreen ( $\lambda_{ex/em}$  in PBS = 485/515 nm) and LipidGreen2 ( $\lambda_{ex/em}$  in PBS = 456/534 nm) (Figure 4) [32,33]. Both dyes were proved to stain LDs in co-localization experiments with LD-associated protein (Periplin) and red emitting LipidTOX<sup>TM</sup> (neutral lipid).

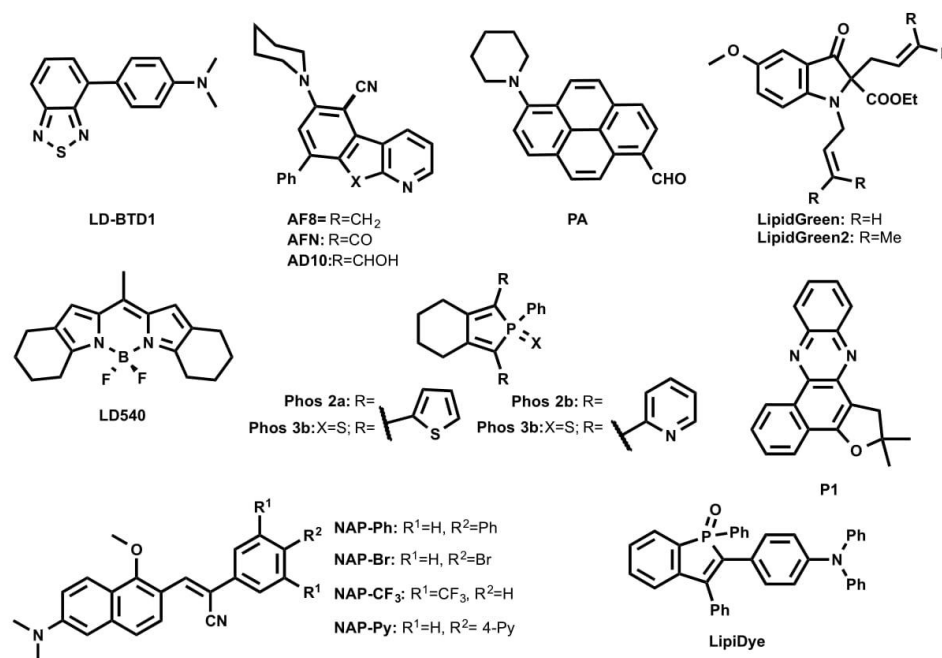


Figure 4. Chemical structures of green-emitting LD probes.

A BODIPY-based lipophilic dye LD540 was introduced for LD imaging (Figure 4) [34]. In sunflower oil, LD540 fluoresced at 545 nm. Well-defined images were obtained when LDs were costained with LD540 and LD-protein ancient ubiquitous protein 1 (AUP1). It is noteworthy that, unlike BODIPY 493/503 and due to its red-shifted emission wavelength, LD540 is prone to cause cross-talk between green and red fluorescence channels in microscopy.

Another family of fluorogenic dyes that have been reported to stain LDs are based on azafluorenes (AF8, AF10) and azafluorenones (AFN) (Figure 4) [35]. In DMSO, AF8 and AF10 showed absorption maxima at 375 nm and 352 nm, respectively, and fluorescence maxima at 479 nm and 477 nm, respectively. Interestingly, in DMSO, AFN exhibited absorption maximum at 432 nm and a fluorescence maximum at 592 nm with a Stokes shift of 160 nm. AF8 and AF10 did not exhibit solvatochromism, while AFN showed solvatochromic shifts, increasing polarity from cyclohexane ( $\lambda_{em} = 510$  nm) to DMSO ( $\lambda_{em} = 592$  nm). The solvatochromic nature of AFN clearly indicates that the molecule undergoes ICT process. Even though AF8 and AF10 showed interesting photophysical properties, only AFN was able to permeate cells and nicely co-localized with Nile Red, showing its specificity to LDs [35].

A family of two-photon excitable naphthalene-based LD probes abbreviated as NAP AIEgens (NAP-Ph, NAP-Br, NAP-CF<sub>3</sub>, and NAP-Py, as shown in Figure 4) was recently introduced [36]. AIEgens NAP probes have been reported to fluoresce at 523–540 nm in an aggregated state in aqueous solution. The NAP AIEgens exhibited a large Stokes shift (>110 nm) and good 2PA cross-section ( $\delta_{2PA}$ ) (45–100 GM at 860 nm), and were thus compatible for 2PE fluorescence live cell and tissue imaging. The ability to specifically stain LDs at very low concentrations (50 nM) within a short time (~15 min incubation) makes NAP probes very attractive to use in the LDs' cell and tissue imaging by the both 1PE and 2PE fluorescence microscopy.

Recently, Appelqvist et al. introduced a new green LD selective emitter based on a benzothiadiazole (BTD) fluorophore named LD-BTD1 (Figure 4) [37]. BTD was cross-coupled to electron-donating dimethylaminophenyl, resulting in a push-pull fluorophore. In hexane, LD-BTD1 exhibits an absorption maximum at 420 nm ( $\epsilon = 7100$  M<sup>-1</sup> cm<sup>-1</sup>) combined with a fluorescence maximum at 511 nm. LD-BTD1 exhibits an impressive solvatochromic shift with increasing solvent polarity from hexane ( $\lambda_{em}$  max = 511 nm) to DMSO ( $\lambda_{em}$  max = 759 nm). Since the excitation and emission spectra of LD-BTD1 overlap with Nile Red due to its broad peaks, it was challenging to proceed with co-localization experiments in a fluorescence imaging experiment.

A polarity-sensitive benzophosphole oxide-based fluorophore was recently developed as an LD marker under commercial name LipiDye (Figure 4) [38,39]. In toluene, it absorbed at 415 nm and emitted at 528 nm, as the ICT fluorophore LipiDye exhibited solvatochromic properties. Interestingly, once LipiDye stained LDs, its emission maximum was between 521–530 nm. LipiDye also labeled other cellular compartments, but with a red-shifted emission ( $\lambda_{em} = 556$ –565 nm). Due to these spectral differences, it was possible to obtain deconvoluted images to discriminate the hydrophobic environment of LDs from other cellular domains.

A similar approach was employed by Niko et al. using a push-pull pyrene probe PA (Figure 4) [40]. This dye features absorption around 430 nm, with a good extinction coefficient in apolar media ( $\epsilon = 25,000$  M<sup>-1</sup> cm<sup>-1</sup>), a quantum yield close to unity, and emission ranging between 480–600 nm as a function of solvent polarity. Its cell permeability enabled the efficient labeling of different lipid structures inside the cells, showing the brightest signal from LDs. Due to the strong solvatochromism of PA, lipid droplets were clearly distinguished by their characteristic emission color, using fluorescence ratiometric imaging that was obtained as a ratio of red channel > 550 nm to green channel < 550 nm. In this case, LDs appeared as the most apolar intracellular lipid structures.

A series of non-fused pyridyl and thienyl-substituted phospholes was proposed as alternative green LD stains (Phos, as shown in Figure 4) [41]. It was shown that the introduction of thienyl groups to the phosphole caused shifts to the red region both in excitation and emission spectra compared to pyridyl groups. In DMSO, Phos 2b and Phos 3b exhibited excitation maxima at 377 nm and 374 nm and emission maxima at 554 nm and 550 nm, whereas Phos 2a and Phos 3a displayed excitation maxima

at longer wavelengths, 439 nm and 429 nm, and emission maxima at shorter wavelengths, 470 nm and 512 nm, respectively. It was also shown that thienyl-derived phospholes (Phos 2a and Phos 3a) outperformed the pyridyl-derivatives in fluorescence LD cell staining.

Exploiting natural products as intermediates, Moliner et al. designed the phenazine P1 (Figure 4) as an LD-specific probe [42]. P1 was reported to be environment-sensitive, and could report viscosity. In dioxane, P1 absorbed at 428 nm and fluoresced at 500 nm, while it formed weakly emissive aggregates in water. P1 co-localized with Nile Red in HeLa cells.

### 3.3. Orange Emitting LD Probes ( $550 \text{ nm} < \lambda_{em} < 600 \text{ nm}$ )

A series of fluorogenic-dyes based on Seoul-Fluor (SF) core was reported to target cellular LDs (Figure 5) [43,44]. By introducing the electron-donating diethyl amino group to the SF core, the molecule was tuned to be fluorogenic, undergoing an ICT process. As consequence, a promising candidate to probe LDs was found to be SF44. The latter fluoresces in a solvatochromic manner, spanning its emission from 561 nm to 624 nm (from diethyl ether to acetonitrile). The authors confirmed that SF44 indeed stains LDs by immunostaining LD-associated membrane proteins—adipose differentiation-related protein (ADRP) and perilipin—in the differentiated 3T3-L1 cells and by co-localization experiments using Nile Red [43]. A further rational design to increase the lipophilicity of the molecule resulted in the discovery of the promising candidate SF58 [44]. The improved analogue was tested together with SF44 to probe the LDs in microalgae [44].

Tang et al. recently introduced FAS and DPAS as AIEgens reporting LDs in cells (Figure 5) [45]. They underwent an excited-state intramolecular proton transfer (ESIPT) process, giving large Stokes shift and enhanced orange and yellow emissions due to their keto forms' fluorescence in the aggregated form. FAS and DPAS showed two emission peaks in solutions. Emission from the short-wavelength region corresponds to enol forms of the molecules, while long-wavelength emission is accredited to keto forms ( $\lambda_{em} = 430/600 \text{ nm}$  for FAS;  $\lambda_{em} = 425/550 \text{ nm}$  for DPAS). FAS and DPAS were nicely co-localized with BODIPY 493/503, proving their LD targeting.

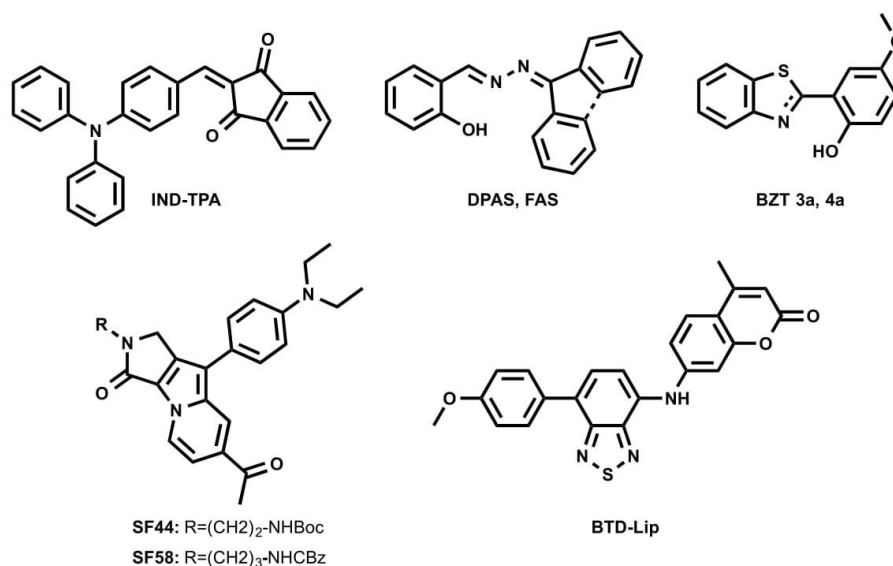


Figure 5. Chemical structures of orange-emitting LD probes.

The previously mentioned green-emitting BTD fluorophore was finely tuned into a BTD–coumarin hybrid resulted to discovery of new red-emitting probe, BTD-Lip (Figure 5), for LDs imaging [46].



BTD-Lip consists of BTD bearing a coumarin moiety in order to extend the  $\pi$ -conjugation. BTD is known to be a good electron acceptor; therefore, the p-methoxyphenyl group was linked to BTD as an electron donor to facilitate the ICT process in an excited state. ICT molecule BTD-Lip fluorescence is solvent-dependent, ranging from 585 nm to 643 nm (from hexane to DMSO). BTD-Lip can mark LDs in Caco-2 cells, together with BODIPY 493/503. Interestingly, BTD-Lip can also report LDs in worms (*Caenorhabditis elegans*) [46]. It is important to note that BTD-Lip displays broad emission bands and is prone to cross-talk in imaging.

Inspired by the interesting properties of blue LDs probe TPA-BI (Figure 3) [31], Tang et al. attempted to tune the spectral properties of the molecule by changing the electron-accepting moiety from benzylidene imidazolone to indane-1,3-dione, resulting in a new orange-emitting IND-TPA (Figure 5) [47]. In THF, IND-TPA absorbed at 478 nm and emitted at 594 nm. IND-TPA displayed atypical AIE behavior: increasing water fraction up to 70% in THF/water solution resulted in a 13.8-fold fluorescence intensity drop, but the further addition of water up to 90% led to a 20.4-fold fluorescence increase. The authors explained that the drop of IND-TPA fluorescence is caused by a twisted intramolecular charge transfer (TICT) effect [48], while the further fluorescence enhancement is credited to aggregation, which caused poor solubility in the water [47]. The co-localization of IND-TPA with standard LD marker BODIPY 493/503 proved the probe specificity. Unlike TPA-BI, IND-TPA gave high signal-to-background ratios. IND-TPA showed a moderate two-photon absorption cross-section of 119 GM upon excitation at 920 nm.

The first example of the photoactivatable AIE probe for LD imaging was recently reported by the same group (BZT 3a Figure 5) [49]. 2-(2-hydroxyphenyl)-benzothiazoline (BZT 3a) can be generated in situ from its disulfide precursor. Such a synthetic route simplified the availability of the LD probe and made it easy to use. BZT 3a underwent photooxidation under illumination at 365 nm to generate an oxidized form of BZT 4a. BZT 4a showed an emission peak at 570 nm in the solid state, while it was absorbed at 365 nm. Further investigations showed that in cells, BZT 3a co-localized with BODIPY 493/503. Interestingly, the photoactivation of BZT 3a could occur under two-photon illumination at 780 nm.

Despite the limited brightness compared to commercial BODIPY 493/503, the green and orange LD markers mentioned above were shown to be chemically accessible, cell permeable, and organelle-specific. In addition, some probes have large Stokes shifts, and are two-photon excitable. On the other hand, their broad absorption and emission spectra can lead to fluorescence cross-talk in the red channel, making the utility of those dyes challenging in multicolor fluorescence imaging.

#### 3.4. Red to Near-Infrared Emitting LD Probes ( $\lambda_{em} > 600$ nm)

Fluorescent dyes operating in the red region are in high demand due to their compatibility with the widely used eGFP protein labeling and with biomedical diagnostics techniques. The predominant advantages of red fluorophores are: (i) low background signal from the autofluorescence of biological samples, (ii) efficient excitation of dyes in thicker tissue samples, and (iii) reduced light scattering.

A red-emitting LD probe was proposed based on a luminescent Zn (II) complex with a salen ligand named LD-TPZn (Figure 6) [50]. In DMSO, LD-TPZn showed a sharp absorption band centered at 390 nm with a shoulder at 440 nm assigned to  $\pi$ - $\pi^*$  transition, and a low-energy band centered at 599 nm. It displayed an intense red emission at 637 nm and a quantum yield of 0.44. LD-TPZn displayed fluorogenic behavior, since it formed non-emissive aggregates in PBS, while there was fivefold fluorescence increase in oil-in-water emulsion. LD-TPZn was proved to stain LDs in cells by co-localization with BODIPY 493/503 and immunolabeled LD-associated protein perilipin-1 in cells. LD-TPZn was reported to undergo endocytic cellular uptake, which is an energy-dependent process [50]. The authors claimed that this property prevents non-specific interaction with other cell organelles, thus reducing background fluorescence. LD-TPZn exhibited a two-photon absorption cross-section ( $\delta$ ) of ca. 110 GM at 880 nm, and was used to image LDs in adipose tissue.

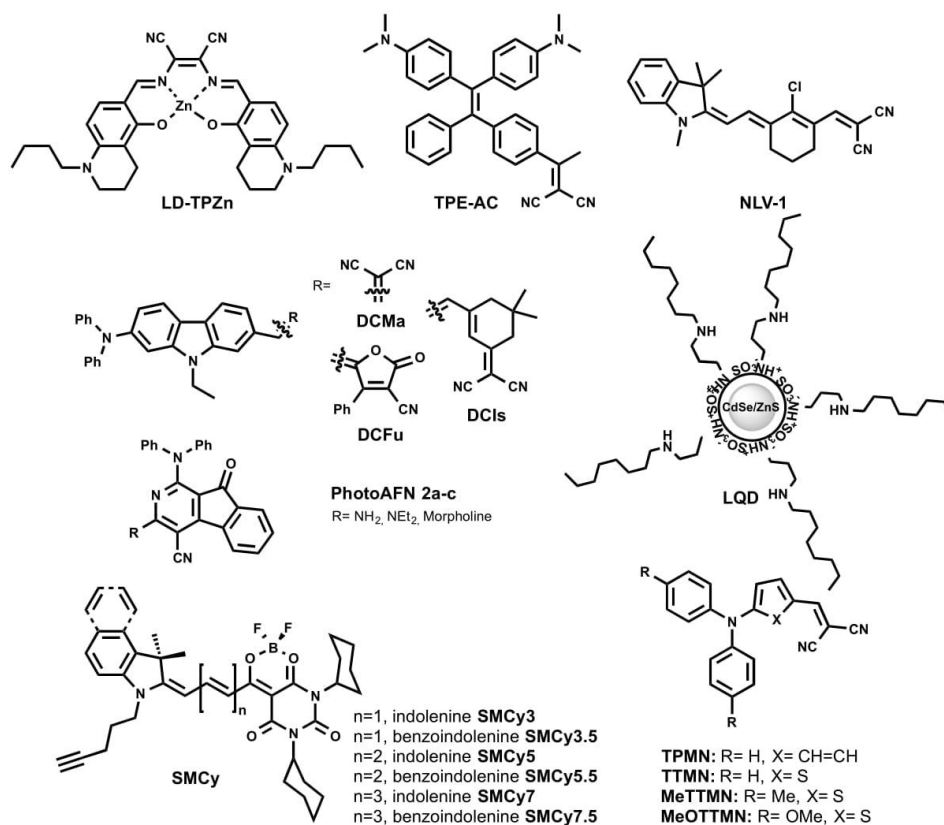


Figure 6. Chemical structures of red-emitting LD probes.

Recently, LQD (Figure 6) was reported as a quantum dot-based fluorescent probe that is selective for LDs [51]. LQD consisted of a red-emitting CdSe quantum dot core capped with a ZnS shell, and covered with polyacrylate coated with octylamine tails. LQD in colloidal solution showed a narrow emission peak at ~600 nm. As a quantum dot, LQD has a wide excitation window; thus, blue excitation was used for the imaging of LDs. LQD entered cell via lipid-raft endocytosis and co-localized with Nile Red. The lipid-raft endocytosis mechanism of LQD uptake minimized its trafficking to the endosome/lysosome. As the authors mentioned, the designed probe can be challenging to use *in vivo*, since they can accumulate in organs due to the larger size compared to the required size for the clearance.

The exploration of 2-azafluorenone core by Tang et al. resulted in the development of the photoactivatable LD-specific probes [52]. The probes are based on dihydro-2-azafluorenones that could undergo photooxidative reaction to form AIEgens 2-azafluorenones (PhotoAFN 2a-c, as shown in Figure 6). They spanned their maximum fluorescence from 610 nm to 624 nm in THF solution, and from 571 nm to 620 nm in a solid state. As TICT molecules, PhotoAFN 2a-c exhibited solvatochromic properties (blue shift with the decrease of solvent polarity). As expected, in LDs, the probes emitted closer to the orange region of the wavelength. The photoactivation property allowed sequentially activating the fluorescence in the LD of selected cells. In addition, the authors showed the difference in the number of LDs in lung cancer cells and normal lung cells.

Further efforts of Tang et al. resulted in a NIR AIEgen-based LD probe to stain LDs (TPE-AC Figure 6) [53]. By simple modification of the TE core with another electron-accepting group,

malononitrile (instead of carbonyl in TPE-AmAl), the spectral properties of TE were tuned from blue emission to the NIR region, which was an AIEgen named TPE-AC. TPE-AC aggregates in water and shows a NIR emission at 780 nm, with an absolute quantum yield of 5% in the solid state. Surprisingly, unlike TPE-AmAl, there was no blue-shifted emission for TPE-AC once it stains cellular LDs, as it preserved its emission in the NIR region similar to in aqueous solution.

The redesign of the triphenylamine core led to the discovery of red/NIR AIEgens, TPMN, TTMN, MeTTMN, and MeOTTMN (Figure 6) [54]. Ranging their emission from 648 nm to 719 nm in the solid state, these AIEgens exhibited interesting characteristics such as synthetically one-pot accessible molecules, large Stokes shifts, and bright emissions. They were shown to specifically stain LDs in a co-localization experiment with lipophilic BODIPY 493/503. The reported AIEgens demonstrated the capacity to produce reactive oxygen species, and thus open the venue to be utilized as photodynamic therapy sensitizers. In addition, TPMN, TTMN, and MeTTMN enabled fluorescence imaging in live zebrafish embryos.

The same group recently reported carbazole-bridged push-pull NIR AIEgens named DCMa, DCIs, and DCFu (Figure 6) [55]. These luminogenic molecules exhibited polarity-dependent solvatochromism. In polar solvent, they are non-emissive due to the TICT effect. In aggregated form, their emission ranged from 665 nm to 775 nm. DCMa, DCIs, and DCFu stained LDs suggested by the co-localization with BODIPY 505/515. All three dyes were two-photon excitable and showed moderate photosensitizing properties to generate reactive oxygen species.

One of the latest approaches to developing a NIR LD probe resulted in the discovery of NLV-1 (Figure 6) [56]. NLV-1 is a merocyanine composed of an indolenine part and a malonitrile moiety (Figure 6). The fluorogenic property of the molecule is caused by free rotation motion in low viscous media and stabilization of the planar emissive form in highly viscous media. Owing to such a property, NLV-1 can report viscosity changes in LDs. The probe displayed a red shift in absorption from 644 nm to 680 nm (from methanol to glycerol), while weak emission at 704 nm in methanol enhanced and shifted to 719 nm in glycerol. The turn-on fluorescence effect was 13.77-fold. Even though NLV-1 can target LDs in cells, it fluoresces only when native LDs environment viscosity is increased by the treatment of viscosity modulators (Monesin, Nystatin) [56]. This property is indeed useful to study the viscosity changes of LDs, but makes it impractical to use NLV-1 as an LD label under physiological conditions.

### 3.5. Families of LD Probes with Tunable Emission Color

Recently available LipidTox™ Neutral Lipid Stains by ThermoFisher were developed to characterize the intracellular accumulation of neutral lipids [57]. Unfortunately, their structural information has not yet been disclosed. LipidTox™ Neutral Lipid Stains are available in three colors: green ( $\lambda_{x/em} = 495/505$ ), red ( $\lambda_{x/em} = 577/609$ ), and deep red ( $\lambda_{x/em} = 637/655$ ). The series of LDs stains was reported to be more specific to LDs in comparison to Nile Red. The drawback is the compatibility for fixed cells only.

In 2018, a family of six fluorogenic dyes named StatoMerocyanines (SMCy) was introduced (Figure 6) [23]. SMCy exhibited impressive performance as fluorescent LD-targeting probes (Figure 5) [23]. The structure of SMCy is composed of barbiturate-based merocyanine with a dioxaborine bridge. Its design included: (i) two cyclohexyl rings to increase the molecule hydrophobicity and bulkiness that can help in preventing non-specific staining; (ii) the non-charged nature favoring attraction to a neutral lipid-rich environment; (iii) the length of the polymethine chain and extension benzoindolenine moiety instead of indolenine to provide the tunability of the photophysical properties; (iv) a dioxaborine bridge to fix a degree of rotation of the fluorophore and therefore enhances its brightness, moreover; and (v) a pentynyl tail to give the opportunity to functionalize the dye via bioorthogonal chemistry.

In oil, the SMCy mostly displayed narrow absorption and emission spectra and spanned their fluorescence from yellow to the NIR (from 541 nm to 753 nm). The brightness of SMCy was



solvent-dependent; they generally displayed high molar extinction coefficients up to an impressive value of  $394,000 \text{ M}^{-1} \cdot \text{cm}^{-1}$  and quantum yields of 100%. In the light of the high lipophilic nature of the SMCy dyes, their photophysical properties were measured in water and colza oil (which was mainly composed of unsaturated long chain triglycerides). Remarkably, SMCy fluorogenic dyes showed up to 1700-fold fluorescence enhancement from water to oil, which is significantly larger compared to other reported LD probes. The fluorogenic behavior could be assigned to a combination of different factors. First, the considerable ICT character of these dyes makes them more fluorescent in less polar and viscous media. Second, in water, they undergo aggregation-caused quenching, which makes them non-fluorescent in aqueous media. This unprecedented light-up effect of SMCy dyes provided a background-free fluorescence imaging of LDs. SMCy co-localized well with both Nile Red and BODIPY 493/503. Interestingly, orthogonally colored SMCy were able to cross-co-localize as well. While IND-TPA can be used for tracking the dynamic motions of LDs in live cells, whether IND-TPA is compatible in multicolor multiorganelle tracking such as SMCy was not shown.

SMCy 3 was shown to be an efficient "green" LD marker [23]. Due to its sharp excitation and emission and its high brightness ( $\epsilon = 87,000 \text{ M}^{-1} \text{ cm}^{-1}$ ,  $\phi = 0.25$  in oil), SMCy 3 provides a robust fluorescence signal from cellular LDs in the green channel, minimizing cross-talk in the channels of other colors.

We showed that four out of six members SMCys displayed remarkably high 2PA cross-section values. Notably, SMCy 5.5 displayed 10,400 GM at 820 nm. Combined to its quantum yield close to unity in lipid-rich environments, this probe is, to our best knowledge, the brightest 2PE fluorophore [23]. Using SMCy 5.5, we were able to map LDs in mouse adipose tissue in multicolor 2PE fluorescence imaging, providing high quality images. A reconstruction of a 52- $\mu\text{m}$  depth 3D 2PE image revealed clear localization of the nuclei, the LDs, as well as the collagen fibers (using second harmonic generation). Owing to their sharp absorption and emission peaks and color availability, SMCy were shown to be perfectly compatible with multicolor imaging in cells and tissues.

#### 4. Application of Fluorescent LD Probes

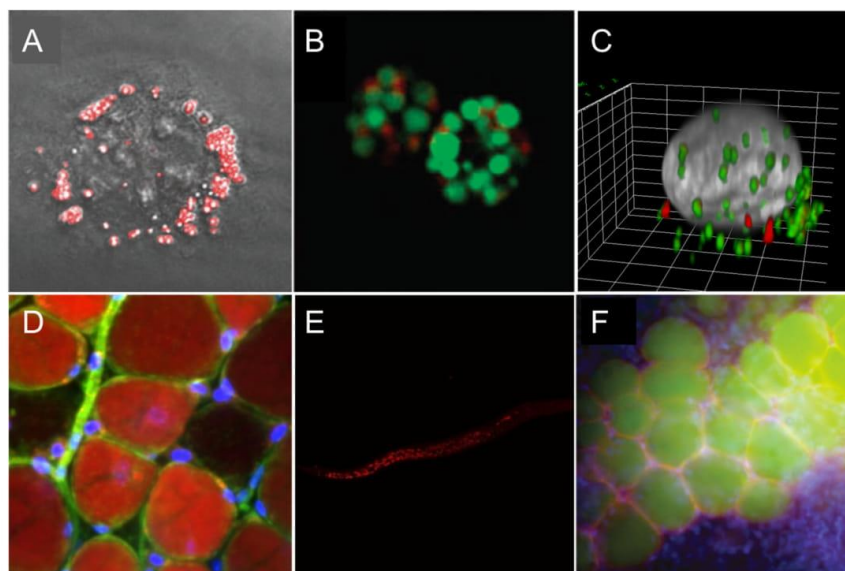
For a long time, LDs were considered simple neutral lipids reservoirs [1,2]. Until recently, its important role as a dynamic organelle in many cellular bioprocesses, in particular metabolic bioprocesses, was reevaluated [58]. Conventional lipid dyes (Nile Red, BODIPY 493/503) permitted shining light on the elucidation of fundamental LD-associated biological processes such as the lipid biosynthesis and further storage in LDs, as well as its association with endoplasmic reticulum (ER) and communication with other cellular compartments [3,15,59]. Also, BODIPY 493/503 was able to sense nuclear LDs [60] and revealed the LDs' biogenesis in hepatitis C-infected cells [61]. Interestingly, the sensitivity of lipophilic dyes is sufficient for monitoring the formation and degradation of LDs under stress conditions in mammalian cells [62] and bacterial pathogens [63,64]. Lipophilic BODIPY 665/676 found its utility to study the dynamics of milk LDs' growth during mice lactation, revealing oxytocin-dependent LDs secretion [65].

Since lipid uptake by monocytes induces inflammation and this is directly connected to an elevated LD content, Nile Red served as a platform in the development of a quantitative imaging-based analysis of lipid accumulation [66]. Recently BioTek<sup>®</sup> has developed a phenotypic-based assay using lipophilic dyes (Nile Red and BODIPY 493/503) to probe LDs formation in hepatic HepG2 cells associated with non-alcoholic fatty liver disease (NAFLD) caused by fatty acids overload [67]. However, Nile Red was also able to detect alterations in neutral LD content in circulating monocytes derived from patients [66] and report LDs NAFLD-associated overproduction [67].

Emerged fluorescent chemical LDs probes enabled new discoveries in LDs research. One can mention efforts to apply them in biofuels studies (Figure 7B) [30,44]. Microalgae attracted considerable attention in biofuel research as an alternative energy source, since it is highly rich in lipid content. Lipid-abundant reservoirs in these organisms correlate with biomass productivity. SF44 [43], SF58 [44], and TPE-AmAl [30] were shown to robustly stain the algal LDs with high specificity to find high

biofuel algae species. In addition, SF44 was subjected to develop a high-throughput screening (HTS) assay to find a small molecule modulator of intracellular LDs. Impressively, SF44-based screening assay led to the discovery of P8B05 that promotes LD formation in cells [43].

Some LDs probes found applications in in vivo imaging. BTD-Lip was able to selectively stain LDs in the worm, displaying a clear-cut distribution of LDs inside the organism, but not in the posterior region of the body, which is known to lack lipid-rich organelles (Figure 7E). The green-emitting LipidGreen and LipidGreen2 were successfully applied to image fat depots in zebrafish [32,33]. LipidGreen was able to monitor lipid synthesis and mobilization during the fasting and feeding cycles of zebrafish (Figure 7F). Also, LipidGreen potential was explored in drug screening for diacylglycerol acyltransferase 1 (DGAT1) inhibitors. DGAT1 is an enzyme that is responsible for the esterification of 1,2-diacylglycerol to form triglycerides, which are a main component of the LDs core; therefore, DGAT1 is a promising therapeutic target for metabolic regulations. The NIR emissive NLV-1 probe was applied to detect viscosity changes in zebrafish under the treatment of viscosity modulators and in living mice under Lipopolysaccharide-induced inflammation conditions [56]. Indeed, NLV-1 can nicely report viscosity alterations, but its non-specificity in complex living organisms may complicate the data analysis.



**Figure 7.** Representative fluorescence microscopy images of LDs. (A) Two-photon excited fluorescence image of LDs labeled with IND-TPA in HCC827 cells [47]. (B) LDs of green algae stained by SF44 (in green); fluorescence from chlorophyll (in red) [44]. (C) Three-dimensional (3D) view of a KB (HeLa derivative) cell displaying exchanged LDs from another cell (SMCy 3.5 in green and SMCy 5.5 in red) [23]. (D) Image of adipocytes in mouse adipose tissue (25  $\mu$ m depth) obtained by a maximum projection of a Z-stack (100 slices of 250-nm depth each). LDs were stained with SMCy 3.5 (in red), nuclei in blue, and plasma membrane in green [23]. (E) Images of LDs stained with BTD-Lip (in red) in *C. elegans* (worm) [46]. (F) LDs imaging with LipidGreen (in green) in vertebral adipose tissues of zebrafish (nucleus in blue, LD-associated protein in red) [32]. Images (A,B,F) were reproduced with permission from The Royal Society of Chemistry. Images (C,D,E) were reproduced with permission from the American Chemical Society.

Two-photon excitation (2PE) microscopy became widely used as an advanced optical imaging technique for biomedical research [68]. 2PE imaging, in addition to reducing cell auto fluorescence,

permits the long-term monitoring of cellular network dynamics while providing high spatio-temporal resolution in both cells and deep tissue samples. For these reasons, the inevitable demand resulted in the development of 2PE-fluorescent probes for LDs research. Hence, MDH ( $\delta_{2PA}$  is not provided) [26], TPA-BI ( $\delta_{2PA} = 213$  GM at 840 nm) [31], LD-TPZn ( $\delta_{2PA} = 110$  GM at 880 nm) [50], IND-TPA ( $\delta_{2PA} = 119$  GM at 920 nm) [47] (Figure 7A), and the family of SMCy dyes ( $\delta_{2PA} = 178\text{--}13,330$  GM) [23] were developed as 2PE LD fluorescent probes. Most of these dyes were able to image LDs under fatty acids-stressed conditions, the brightest to date two-photon excitable SMCy 5.5 was able to map LDs in mouse liver tissue under physiological conditions. Due to its extreme brightness and near-infrared emission, the further utilization of SMCy dyes in 2PE imaging of tissues or in vivo living animals imaging will facilitate discoveries in the field of LDs.

SMCy 3.5 was successfully used to stain large LDs in adipocytes as well as small circulating ones in the blood vessels (Figure 7D). Additionally, the color palette availability of LDs markers provides the opportunity to track the LDs of different populations within the same set of experiments (Figure 7C). Thus, SMCy dyes enabled demonstrating the occurrence of time-dependent cell-to-cell LDs exchange: it was shown that exchange took place after 48 h of intercellular interaction, which is not the case after 4 h or 18 h [23].

Table 1 sums up the photophysical properties of the reviewed LD probes as well as their applications.

**Table 1.** Photophysical properties of LDs probes and their applications.

Dye	$\lambda_{ex}/\lambda_{em}$ , nm	$\epsilon$ , $M^{-1}cm^{-1}$	QY	Excitation Mode	Application	Reference
<b>Commonly used</b>						
Nile Red	553/635 (MeOH)	44,000	n.r.	1P	Cells	[21]
BODIPY 493/503	493/503 (MeOH)	90,000	n.r.	1P	Cells	[22]
<b>Blue</b>						
PITE	313 <sup>a</sup> /487 (MeCN)	n.r.	0.48 (MeCN)	1P	Mammalian and bacterial cells	[28]
MDH	405/570 (MeOH)	n.r.	n.r.	1P, 2P ( $\delta_{2PA}$ n.r.)	Cells	[26]
TPA-BI	447/619 (MeCN)	34,000	0.22 (water 70%)	1P, 2P ( $\delta_{2PA} = 213$ GM at 840 nm)	1PE and 2PE imaging in cells	[31]
TPE-AmAl	400/470 (THF)	n.r.	0.22 (solid state)	1P	Cells and algae	[30]
PyrPy 10d	356/449 (CHCl <sub>3</sub> )	12,400	0.322 (CHCl <sub>3</sub> )	1P	Cells	[27]
PyrPy 11c	344/447 (CHCl <sub>3</sub> )	13,200	0.029 (CHCl <sub>3</sub> )	1P	Cells	[27]
<b>Green</b>						
BODIPY 505/515	505/515 (MeOH)	94,000	n.r.	1P	Cells	[22]
LipidGreen	485/515 (PBS)	n.r.	n.r.	1P	See footnote 1	[32]
LipidGreen2	456/534 (PBS)	n.r.	n.r.	1P	See footnote 2	[33]
LD540	540/545.5 (oil)	n.r.	n.r.	1P	Multicolor imaging in cells	[34]
LD-BTD1	420/511 (hexane)	7100	0.66 (hexane)	1P	Cells	[37]
AF8	380/479 (DMSO)	n.r.	0.31 (DMSO)	1P	Cells	[35]
AFN	428/592 (DMSO)	n.r.	0.17 (DMSO)	1P	Cells	[35]
AF10	356/477 (DMSO)	n.r.	0.18 (DMSO)	1P	Cells	[35]
NAP-Ph	409/523 (water)	n.r.	0.018 (water)	1P, 2P ( $\delta_{2PA} = 100$ GM at 860 nm)	1PE and 2PE imaging in cells and tissues	[36]
NAP-Br	409/525 (water)	n.r.	0.014 (water)	1P, 2P ( $\delta_{2PA} = -50$ GM at 860 nm)	1PE and 2PE imaging in cells and tissues	[36]
NAP-CF <sub>3</sub>	425/560 (water)	n.r.	0.016 (water)	1P, 2P ( $\delta_{2PA} = -50$ GM at 860 nm)	1PE and 2PE imaging in cells and tissues	[36]
NAP-Py	413/541 (water)	n.r.	0.015 (water)	1P, 2P ( $\delta_{2PA} = 45$ GM at 860 nm)	1PE and 2PE imaging in cells and tissues	[36]
LipiDye	415 (toluene)	18,700	0.94 (toluene)	1P	Cells	[38,39]
PA	434/521 (toluene)	25,000	0.95 (toluene)	1P, 2P ( $\delta_{2PA} = 35$ GM at 820 nm)	Lipid order in cells	[40]
Phos 2a	439/554 (DMSO)	n.r.	n.r.	1P	Cells	[41]
Phos 3a	429/512 (DMSO)	n.r.	n.r.	1P	Cells	[41]
Phos 2b	437/554 (DMSO)	n.r.	n.r.	1P	Not applicable for cells	[41]
Phos 3b	374/550 (DMSO)	n.r.	n.r.	1P	Not applicable for cells	[41]
P1	428/500 (Dioxane)	n.r.	0.22	1P	Cells	[42]



Table 1. Cont.

Dye	$\lambda_{ex}/\lambda_{em}$ , nm	$\epsilon$ , M <sup>-1</sup> cm <sup>-1</sup>	QY	Excitation Mode	Application	Reference
<b>Orange</b>						
SF44	455/626 (MeCN)	n.r.	0.09 (MeCN)	1P	Cells, HTS for LD modulator	[43]
SF58	440/623 (MeCN)	n.r.	0.09 (MeCN)	1P	Cells and algae	[44]
FAS	322/600 (water) <sup>b</sup>	n.r.	0.021 (solid state)	1P	Cells	[45]
DPAS	301/550 (water) <sup>b</sup>	n.r.	0.03 (solid state)	1P	Cells	[45]
BTD-Lip	455/624 (MeCN)	7586	0.2 (MeCN)	1P	Cells and worms	[46]
IND-TPA	478/594 (THF)	n.r.	0.069 (THF)	1P, 2P ( $\delta_{2PA}$ = 119 GM at 920 nm)	1PE and 2PE imaging in cells	[47]
BZT 3a	309/576 (water)	n.r.	n.r.	1P, 2P at 780 nm	1PE and 2PE imaging in cells	[49]
BZT 4a	-/570 (solid)	n.r.	0.4 (solid state)	1P, 2P at 780 nm	1PE and 2PE imaging in cells	[49]
<b>Red to NIR</b>						
LD-TPZn	599/630 (DMSO)	10,600	0.44 (DMSO)	1P, 2P ( $\delta_{2PA}$ = 110 GM at 880 nm)	1PE and 2PE imaging in cells	[50]
LQD	495/600 (colloidal)	n.r.	n.r.	1P	Cells	[51]
PhotoAFN 2a	409/624 (THF)	n.r.	0.006 (THF)	1P	Cells	[52]
PhotoAFN 2b	402/617 (THF)	n.r.	0.005 (THF)	1P	Cells	[52]
PhotoAFN 2c	400/610 (THF)	n.r.	0.008 (THF)	1P	Cells	[52]
TPE-AC	455/724 (THF)	n.r.	0.05 (solid state)	1P	Cells	[53]
TPMN	441/635 (MeCN)	n.r.	0.0021 (MeCN)	1P	Cells and zebrafish	[54]
TTMN	483/664 (MeCN)	n.r.	0.0032 (MeCN)	1P	Cells and zebrafish	[54]
MeTTMN	441/635 (MeCN)	n.r.	0.0021 (MeCN)	1P	Cells and zebrafish	[54]
MeOTTMN	499/- (MeCN)	n.r.	n.r.	1P	Cells	[54]
DCMa	478/665 (water)	n.r.	0.296 (solid state)	1P, 2P ( $\delta_{2PA}$ = 394 GM at 940 nm)	1PE and 2PE imaging in cells	[54]
DCIs	510/709 (water)	n.r.	0.135 (solid state)	1P, 2P ( $\delta_{2PA}$ = 548 GM at 980 nm)	1PE and 2PE imaging in cells	[55]
DCFu	538/755 (water)	n.r.	0.017 (solid state)	1P, 2P ( $\delta_{2PA}$ = 887 GM at 1020 nm)	1PE and 2PE imaging in cells	[55]
v NLV-1	680/719 (glycerol)	n.r.	0.204 (glycerol)	1P	See footnote 3	[56]
<b>Multicolor</b>						
<i>LipidTox</i> <sup>TM</sup>					Analysis of steatosis	[57]
green	495/505	n.r.	n.r.	1P		[57]
red	577/609	n.r.	n.r.	1P		[57]
deep red	637/655	n.r.	n.r.	1P		[57]
<i>SMCy family</i>					See footnote 4	[23]
SMCy3	512/541 (oil)	82,900	0.21 (oil)	1P, 2P ( $\delta_{2PA}$ = 178 GM at 690 nm, DMSO)		[23]
SMCy3.5	530/559 (oil)	100,000	0.4 (oil)	1P, 2P ( $\delta_{2PA}$ = 2400 GM at 760 nm, DMSO)		[23]
SMCy5	618/648 (oil)	256,000	0.6 (oil)	1P, 2P ( $\delta_{2PA}$ = 6250 GM at 740 nm, DMSO)		[23]
SMCy5.5	638/662 (oil)	169,000	0.74 (oil)	1P, 2P ( $\delta_{2PA}$ = 13330 GM at 770 nm and 10400 GM at 820 nm, DMSO)		[23]
SMCy7	692/744 (oil)	44,000	0.42 (oil)	1P		[23]
SMCy7.5	716/753 (oil)	103,000	0.19 (oil)	1P		[23]

n.r. not reported. <sup>a</sup> For tetraphenylethylene (TE) region; <sup>b</sup> Emission of keto form. Footnote 1: Imaging in cells and zebrafish. Monitoring lipid synthesis and mobilization during fasting and feeding cycles. Drug screening for diacylglycerol acyltransferase 1 (DGAT1) inhibitor. Footnote 2: Imaging in cells and zebrafish. Early identification of hepatosteatosis. Footnote 3: Viscosity changes in cellular LDs, zebrafish, and mice. Footnote 4: Multicolor imaging of cells and tissues. LD tracking and intracellular exchange.

## 5. Conclusions and Perspectives

As outlined above, small-molecule LD-targeting fluorescent probes have already contributed in LDs research, revealing their niche role in cellular metabolism. Significant progress has already been made in the development of new LD-specific probes, including: (1) the range of available emission colors was expanded, and now span from blue to NIR; (2) efficient LDs probes that are compatible with 2PE are now available; (3) the selectivity toward LDs was improved; (4) some

photophysical properties were enhanced such as: higher brightness, photostability, and Stokes shifts. Most of described probes exhibit solvatochromic and fluorogenic properties, which can originate from different photophysical mechanisms, such as ICT, TICT, AIE, and aggregation-caused quenching [14]. First, the dyes are based on ICT processes that have a strong electron-donating group and good electron-accepting moiety (PyrPy probes [27], Seoul Fluor probes [43,44], azafluorenone [35], and probes based on benzothiadiazole core [37,46]. The other mechanism involves aggregation-caused quenching (ACQ), where the molecules, being aggregated and non-fluorescent in aqueous media, become highly emissive after binding lipid droplets in molecular form. This mechanism concerns the SMCy family and many other lipophilic LD probes. The third distinguished group utilizes two photophysical fluorescence-activating principals, twisted ICT (TICT) and AIE (the probes based on tetraphenylethylene core as an AIE-driving force [30,31,47,49,53,55]. The AIEgens are usually weakly emissive in organic solvent solutions, but they are strongly fluorescent in an aggregated state in aqueous media (as nanoparticles or solid film). While developed AIE-based bioprobes demonstrated their utility to stain LDs, they usually displayed blue-shifted emission in LDs in comparison to the emission in water, except for TPE-AC, DCMa, DCIs, and DCFu which preserved its NIR fluorescence [53,55]. The fluorescent probes described here have enabled the visualization and detection of LDs in cells, tissues, and living animals.

We anticipate that continued progress will be made to develop new LD-specific probes, such as probes that can be used in multimodal imaging techniques to increase the credibility of the obtained results. For instance, the introduction of radioactive isotopes to the core of reported LD-specific probes would open the opportunity to use a positron emission tomography (PET) imaging technique in addition to the fluorescence. In particular, the introduction of the radioisotope  $^{18}\text{F}$  would be significant, since it has longer lifetime in comparison to commonly used  $^3\text{H}$  and  $^{14}\text{C}$ , and is already widely used in clinics (for example, fludeoxyglucose ( $^{18}\text{F}$ ) is used for cancer diagnostics). Emerging advances in stimulated Raman scattering (SRS) imaging technology allowed the visualization of alkyne-tagged small molecules in live cells [69]. The alkyne tag exhibits a defined strong stretching bond signal of  $\text{C}\equiv\text{C}$  at  $\sim 2100\text{ cm}^{-1}$ ; therefore, SMCy dyes can potentially be used in SRS imaging for cellular LDs in addition to fluorescence. Interestingly, the nitrile  $\text{C}\equiv\text{N}$  bond has a characteristic peak at  $\sim 2200\text{ cm}^{-1}$ ; thus, fluorophores such as NLV-1, LD-TPZn, TPE-AC, DCMa, DCIs, TPMN, TTMN, MeTTMN, and MeOTTMN might have a potential in SRS imaging techniques. Thus, the application of the single probe for LDs in multimodal imaging settings will improve the fidelity of the acquired data.

In further efforts, it is essential to establish stronger collaborations between chemists and biologists for tuning the LD research, not only as tool-making oriented, but also as a *bona fide* solution that is oriented to better understanding the LD role in biological processes.

**Author Contributions:** T.K.F.; Writing-Original Draft Preparation, T.K.F., A.S.K. and M.C.; Writing-Review & Editing; A.S.K. and M.C.; Supervision.

**Funding:** This research was funded by ERC Consolidator grant BrightSens 648528, ANR BrightRiboProbes (ANR-16-CE11-0010).

**Acknowledgments:** The authors would like to thank Hisashi Satoh for the invitation to publish in the special issue of Materials: "Fluorescent Sensors for Selective Detection". The authors also thank The Royal Society of Chemistry and the American Chemical Society for the permission for the preproduction of the images in Figure 7.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Martin, S.; Parton, R.G. Opinion: Lipid Droplets: A Unified View of a Dynamic Organelle. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 373–378. [[CrossRef](#)] [[PubMed](#)]
2. Farese, R.V.; Walther, T.C. Lipid Droplets Finally Get a Little R-E-S-P-E-C-T. *Cell* **2009**, *139*, 855–860. [[CrossRef](#)] [[PubMed](#)]
3. Thiam, A.R.; Farese, R.V., Jr.; Walther, T.C. The Biophysics and Cell Biology of Lipid Droplets. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 775–786. [[CrossRef](#)] [[PubMed](#)]

4. Blom, T.; Somerharju, P.; Ikonen, E. Synthesis and Biosynthetic Trafficking of Membrane Lipids. *CSH Perspect. Biol.* **2011**, *3*, a004713. [CrossRef] [PubMed]
5. Olzmann, J.A.; Richter, C.M.; Kopito, R.R. Spatial Regulation of UBXD8 and P97/VCP Controls ATGL-Mediated Lipid Droplet Turnover. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1345–1350. [CrossRef] [PubMed]
6. Onal, G.; Kutlu, O.; Gozuacik, D.; Dokmeci Emre, S. Lipid Droplets in Health and Disease. *Lipids Health Dis.* **2017**, *16*, 128. [CrossRef] [PubMed]
7. Fujimoto, T.; Ohsaki, Y.; Suzuki, M.; Cheng, J. Chapter 13—Imaging Lipid Droplets by Electron Microscopy. In *Methods in Cell Biology*; Yang, H., Li, P., Eds.; Lipid Droplets; Academic Press: Cambridge, MA, USA, 2013; Volume 116, pp. 227–251.
8. Horn, P.J.; Ledbetter, N.R.; James, C.N.; Hoffman, W.D.; Case, C.R.; Verbeck, G.F.; Chapman, K.D. Visualization of Lipid Droplet Composition by Direct Organelle Mass Spectrometry. *J. Biol. Chem.* **2011**, *286*, 3298–3306. [CrossRef] [PubMed]
9. Abramczyk, H.; Surmacki, J.; Kopeć, M.; Olejnik, A.K.; Lubecka-Pietruszewska, K.; Fabianowska-Majewska, K. The Role of Lipid Droplets and Adipocytes in Cancer. Raman Imaging of Cell Cultures: MCF10A, MCF7, and MDA-MB-231 Compared to Adipocytes in Cancerous Human Breast Tissue. *Analyst* **2015**, *140*, 2224–2235. [CrossRef] [PubMed]
10. Kim, K.; Lee, S.; Yoon, J.; Heo, J.; Choi, C.; Park, Y. Three-Dimensional Label-Free Imaging and Quantification of Lipid Droplets in Live Hepatocytes. *Sci. Rep.* **2016**, *6*, 36815. [CrossRef] [PubMed]
11. Lakowicz, J.R. *Principles of Fluorescence Spectroscopy*, 3rd ed.; Springer: New York City, NY, USA, 2006.
12. Lavis, L.D. Teaching Old Dyes New Tricks: Biological Probes Built from Fluoresceins and Rhodamines. *Annu. Rev. Biochem.* **2017**, *86*, 825–843. [CrossRef] [PubMed]
13. Zhu, H.; Fan, J.; Du, J.; Peng, X. Fluorescent Probes for Sensing and Imaging within Specific Cellular Organelles. *Acc. Chem. Res.* **2016**, *49*, 2115–2126. [CrossRef] [PubMed]
14. Klymchenko, A.S. Solvatochromic and Fluorogenic Dyes as Environment-Sensitive Probes: Design and Biological Applications. *Acc. Chem. Res.* **2017**, *50*, 366–375. [CrossRef] [PubMed]
15. Kuerschner, L.; Moessinger, C.; Thiele, C. Imaging of Lipid Biosynthesis: How a Neutral Lipid Enters Lipid Droplets. *Traffic* **2007**, *9*, 338–352. [CrossRef] [PubMed]
16. Listenberger, L.L.; Brown, D.A. Fluorescent Detection of Lipid Droplets and Associated Proteins. *Curr. Protoc. Cell Biol.* **2007**, *35*, 24. [CrossRef] [PubMed]
17. Aoki, T.; Hagiwara, H.; Fujimoto, T. Peculiar Distribution of Fodrin in Fat-Storing Cells. *Exp. Cell Res.* **1997**, *234*, 313–320. [CrossRef] [PubMed]
18. Koopman, R.; Schaart, G.; Hesselink, M.K. Optimisation of Oil Red O Staining Permits Combination with Immunofluorescence and Automated Quantification of Lipids. *Histochem. Cell. Biol.* **2001**, *116*, 63–68. [PubMed]
19. Ohsaki, Y.; Shinohara, Y.; Suzuki, M.; Fujimoto, T. A Pitfall in Using BODIPY Dyes to Label Lipid Droplets for Fluorescence Microscopy. *Histochem. Cell. Biol.* **2010**, *133*, 477–480. [CrossRef] [PubMed]
20. Fukumoto, S.; Fujimoto, T. Deformation of Lipid Droplets in Fixed Samples. *Histochem. Cell. Biol.* **2002**, *118*, 423–428. [CrossRef] [PubMed]
21. Greenspan, P.; Mayer, E.P.; Fowler, S.D. Nile Red: A Selective Fluorescent Stain for Intracellular Lipid Droplets. *J. Cell Biol.* **1985**, *100*, 965–973. [CrossRef] [PubMed]
22. Thermo Fisher Scientific. BODIPY 493/503. Available online: <https://www.thermofisher.com/order/catalog/product/D3922> (accessed on 26 July 2018).
23. Collot, M.; Fam, T.K.; Ashokkumar, P.; Faklaris, O.; Galli, T.; Danglot, L.; Klymchenko, A.S. Ultrabright and Fluorogenic Probes for Multicolor Imaging and Tracking of Lipid Droplets in Cells and Tissues. *J. Am. Chem. Soc.* **2018**, *140*, 5401–5411. [CrossRef] [PubMed]
24. Gocze, P.M.; Freeman, D.A. Factors Underlying the Variability of Lipid Droplet Fluorescence in MA-10 Leydig Tumor Cells. *Cytometry* **1994**, *17*, 151–158. [CrossRef] [PubMed]
25. Koreivienė, J. Microalgae Lipid Staining with Fluorescent BODIPY Dye. In *SpringerLink; Methods in Molecular Biology*; Humana Press: New York City, NY, USA, 2017; pp. 1–7.
26. Yang, H.-J.; Hsu, C.-L.; Yang, J.-Y.; Yang, W.Y. Monodansylpentane as a Blue-Fluorescent Lipid-Droplet Marker for Multi-Color Live-Cell Imaging. *PLoS ONE* **2012**, *7*, e32693. [CrossRef] [PubMed]



27. Becerra-Ruiz, M.; Vargas, V.; Jara, P.; Tirapegui, C.; Carrasco, C.; Nuñez, M.; Lezana, N.; Galdámez, A.; Vilches-Herrera, M. Blue-Fluorescent Probes for Lipid Droplets Based on Dihydrochromeno-Fused Pyrazolo- and Pyrrolopyridines. *Eur. J. Org. Chem.* **2018**, *34*, 4795–4801. [CrossRef]
28. Sk, B.; Thakre, P.K.; Tomar, R.S.; Patra, A. A Pyridoindole-Based Multifunctional Bioprobe: PH-Induced Fluorescence Switching and Specific Targeting of Lipid Droplets. *Chem. Asian J.* **2017**, *12*, 2501–2509. [CrossRef] [PubMed]
29. Mei, J.; Leung, N.L.C.; Kwok, R.T.K.; Lam, J.W.Y.; Tang, B.Z. Aggregation-Induced Emission: Together We Shine, United We Soar! *Chem. Rev.* **2015**, *115*, 11718–11940. [CrossRef] [PubMed]
30. Wang, E.; Zhao, E.; Hong, Y.; Lam, J.W.Y.; Tang, B.Z. A Highly Selective AIE Fluorogen for Lipid Droplet Imaging in Live Cells and Green Algae. *J. Mater. Chem. B* **2014**, *2*, 2013–2019. [CrossRef]
31. Jiang, M.; Gu, X.; Lam, J.W.Y.; Zhang, Y.; Kwok, R.T.K.; Wong, K.S.; Tang, B.Z. Two-Photon AIE Bio-Probe with Large Stokes Shift for Specific Imaging of Lipid Droplets. *Chem. Sci.* **2017**, *8*, 5440–5446. [CrossRef] [PubMed]
32. Lee, J.H.; So, J.-H.; Jeon, J.H.; Choi, E.B.; Lee, Y.-R.; Chang, Y.-T.; Kim, C.-H.; Bae, M.A.; Ahn, J.H. Synthesis of a New Fluorescent Small Molecule Probe and Its Use for in Vivolipid Imaging. *Chem. Commun.* **2011**, *47*, 7500–7502. [CrossRef] [PubMed]
33. Chun, H.-S.; Jeon, J.H.; Pagire, H.S.; Lee, J.H.; Chung, H.-C.; Park, M.J.; So, J.-H.; Ryu, J.-H.; Kim, C.-H.; Ahn, J.H.; et al. Synthesis of LipidGreen2 and Its Application in Lipid and Fatty Liver Imaging. *Mol. Biosyst.* **2013**, *9*, 630–633. [CrossRef] [PubMed]
34. Spandl, J.; White, D.J.; Peychl, J.; Thiele, C. Live Cell Multicolor Imaging of Lipid Droplets with a New Dye, LD540. *Traffic* **2009**, *10*, 1579–1584. [CrossRef] [PubMed]
35. Sharma, A.; Umar, S.; Kar, P.; Singh, K.; Sachdev, M.; Goel, A. A New Type of Biocompatible Fluorescent Probe AFN for Fixed and Live Cell Imaging of Intracellular Lipid Droplets. *Analyst* **2015**, *141*, 137–143. [CrossRef] [PubMed]
36. Niu, G.; Zhang, R.; Kwong, J.P.C.; Lam, J.W.Y.; Chen, C.; Wang, J.; Chen, Y.; Feng, X.; Kwok, R.T.K.; Sung, H.H.-Y.; et al. Specific Two-Photon Imaging of Live Cellular and Deep-Tissue Lipid Droplets by Lipophilic AIEgens at Ultralow Concentration. *Chem. Mater.* **2018**, *30*, 4778–4787. [CrossRef]
37. Appelqvist, H.; Stranius, K.; Börjesson, K.; Nilsson, K.P.R.; Dyrager, C. Specific Imaging of Intracellular Lipid Droplets Using a Benzothiadiazole Derivative with Solvatochromic Properties. *Bioconjug. Chem.* **2017**, *28*, 1363–1370. [CrossRef] [PubMed]
38. Yamaguchi, E.; Wang, C.; Fukazawa, A.; Taki, M.; Sato, Y.; Sasaki, T.; Ueda, M.; Sasaki, N.; Higashiyama, T.; Yamaguchi, S. Environment-Sensitive Fluorescent Probe: A Benzophosphole Oxide with an Electron-Donating Substituent. *Angew. Chem. Int. Edit.* **2015**, *127*, 4622–4626. [CrossRef]
39. High Sensitive Lipid Droplets Imaging Fluorescent Dye | LipiDye | Funakoshi Co., Ltd.: Tokyo, Japan. Available online: [https://www.funakoshi.co.jp/exports\\_contents/80682](https://www.funakoshi.co.jp/exports_contents/80682) (accessed on 30 July 2018).
40. Niko, Y.; Didier, P.; Mely, Y.; Konishi, G.; Klymchenko, A.S. Bright and Photostable Push-Pull Pyrene Dye Visualizes Lipid Order Variation between Plasma and Intracellular Membranes. *Sci. Rep.* **2016**, *6*, 18870. [CrossRef] [PubMed]
41. Öberg, E.; Appelqvist, H.; Nilsson, K.P.R. Non-Fused Phospholes as Fluorescent Probes for Imaging of Lipid Droplets in Living Cells. *Front. Chem.* **2017**, *5*. [CrossRef] [PubMed]
42. De Moliner, F.; King, A.; Dias, G.G.; de Lima, G.F.; de Simone, C.A.; da Silva Júnior, E.N.; Vendrell, M. Quinone-Derived  $\pi$ -Extended Phenazines as New Fluorogenic Probes for Live-Cell Imaging of Lipid Droplets. *Front. Chem.* **2018**, *6*. [CrossRef] [PubMed]
43. Kim, E.; Lee, S.; Park, S.B. A Seoul-Fluor-Based Bioprobe for Lipid Droplets and Its Application in Image-Based High Throughput Screening. *Chem. Commun.* **2012**, *48*, 2331–2333. [CrossRef] [PubMed]
44. Lee, Y.; Na, S.; Lee, S.; Jeon, N.L.; Park, S.B. Optimization of Seoul-Fluor-Based Lipid Droplet Bioprobes and Their Application in Microalgae for Bio-Fuel Study. *Mol. Biosyst.* **2013**, *9*, 952–956. [CrossRef] [PubMed]
45. Wang, Z.; Gui, C.; Zhao, E.; Wang, J.; Li, X.; Qin, A.; Zhao, Z.; Yu, Z.; Tang, B.Z. Specific Fluorescence Probes for Lipid Droplets Based on Simple AIEgens. *ACS Appl. Mater. Interfaces* **2016**, *8*, 10193–10200. [CrossRef] [PubMed]

46. Mota, A.A.R.; Correa, J.R.; de Andrade, L.P.; Assumpção, J.A.F.; de Souza Cintra, G.A.; Freitas-Junior, L.H.; da Silva, W.A.; de Oliveira, H.C.B.; Neto, B.A.D. From Live Cells to *Caenorhabditis Elegans*: Selective Staining and Quantification of Lipid Structures Using a Fluorescent Hybrid Benzothiadiazole Derivative. *ACS Omega* **2018**, *3*, 3874–3881. [CrossRef] [PubMed]
47. Gao, M.; Su, H.; Li, S.; Lin, Y.; Ling, X.; Qin, A.; Tang, B.Z. An Easily Accessible Aggregation-Induced Emission Probe for Lipid Droplet-Specific Imaging and Movement Tracking. *Chem. Commun.* **2017**, *53*, 921–924. [CrossRef] [PubMed]
48. Grabowski, Z.R.; Rotkiewicz, K.; Rettig, W. Structural Changes Accompanying Intramolecular Electron Transfer: Focus on Twisted Intramolecular Charge-Transfer States and Structures. *Chem. Rev.* **2003**, *103*, 3899–4032. [CrossRef] [PubMed]
49. Li, S.; Ling, X.; Lin, Y.; Qin, A.; Gao, M.; Tang, B.Z. In Situ Generation of Photoactivatable Aggregation-Induced Emission Probes for Organelle-Specific Imaging. *Chem. Sci.* **2018**, *9*, 5730–5735. [CrossRef] [PubMed]
50. Tang, J.; Zhang, Y.; Yin, H.-Y.; Xu, G.; Zhang, J.-L. Precise Labeling and Tracking of Lipid Droplets in Adipocytes Using a Luminescent ZnSalen Complex. *Chem. Asian J.* **2017**, *12*, 2533–2538. [CrossRef] [PubMed]
51. Mandal, S.; Jana, N.R. Quantum Dot-Based Designed Nanoprobe for Imaging Lipid Droplet. *J. Phys. Chem. C* **2017**, *121*, 23727–23735. [CrossRef]
52. Gao, M.; Su, H.; Lin, Y.; Ling, X.; Li, S.; Qin, A.; Tang, B.Z. Photoactivatable Aggregation-Induced Emission Probes for Lipid Droplets-Specific Live Cell Imaging. *Chem. Sci.* **2017**, *8*, 1763–1768. [CrossRef] [PubMed]
53. Kang, M.; Gu, X.; Kwok, R.T.K.; Leung, C.W.T.; Lam, J.W.Y.; Li, F.; Tang, B.Z. A Near-Infrared AIEgen for Specific Imaging of Lipid Droplets. *Chem. Commun.* **2016**, *52*, 5957–5960. [CrossRef] [PubMed]
54. Wang, D.; Su, H.; Kwok, R.T.K.; Shan, G.; Leung, A.C.S.; Lee, M.M.S.; Sung, H.H.Y.; Williams, I.D.; Lam, J.W.Y.; Tang, B.Z. Facile Synthesis of Red/NIR AIE Luminogens with Simple Structures, Bright Emissions, and High Photostabilities, and Their Applications for Specific Imaging of Lipid Droplets and Image-Guided Photodynamic Therapy. *Adv. Funct. Mater.* **2017**, *27*, 1704039. [CrossRef]
55. Zheng, Z.; Zhang, T.; Liu, H.; Chen, Y.; Kwok, R.T.K.; Ma, C.; Zhang, P.; Sung, H.H.Y.; Williams, I.D.; Lam, J.W.Y.; et al. Bright Near-Infrared Aggregation-Induced Emission Luminogens with Strong Two-Photon Absorption, Excellent Organelle Specificity, and Efficient Photodynamic Therapy Potential. *ACS Nano* **2018**, *12*, 8145–8159. [CrossRef] [PubMed]
56. Guo, R.; Yin, J.; Ma, Y.; Li, G.; Wang, Q.; Lin, W. A Novel NIR Probe for Detection of Viscosity in Cellular Lipid Droplets, Zebra Fishes and Living Mice. *Sensor. Actuators B-Chem.* **2018**, *271*, 321–328. [CrossRef]
57. Thermo Fisher Scientific, LipidTOX Neutral Lipid Stain. Available online: <https://www.thermofisher.com/order/catalog/product/H34475> (accessed on 27 July 2018).
58. Fujimoto, T.; Ohsaki, Y.; Cheng, J.; Suzuki, M.; Shinohara, Y. Lipid Droplets: A Classic Organelle with New Outfits. *Histochem. Cell. Biol.* **2008**, *130*, 263–279. [CrossRef] [PubMed]
59. Jacquier, N.; Choudhary, V.; Mari, M.; Toulmay, A.; Reggiori, F.; Schneider, R. Lipid Droplets Are Functionally Connected to the Endoplasmic Reticulum in *Saccharomyces Cerevisiae*. *J. Cell. Sci.* **2011**, *124*, 2424–2437. [CrossRef] [PubMed]
60. Layerenza, J.P.; González, P.; García de Bravo, M.M.; Polo, M.P.; Sisti, M.S.; Ves-Losada, A. Nuclear Lipid Droplets: A Novel Nuclear Domain. *BBA-Mol. Cell Biol. Lipids* **2013**, *1831*, 327–340. [CrossRef] [PubMed]
61. Nevo-Yassaf, I.; Lovelle, M.; Nahmias, Y.; Hirschberg, K.; Sklan, E.H. Live Cell Imaging and Analysis of Lipid Droplets Biogenesis in Hepatitis C Virus Infected Cells. *Methods* **2017**, *127*, 30–36. [CrossRef] [PubMed]
62. Tuohetahunttila, M.; Molenaar, M.R.; Spee, B.; Brouwers, J.F.; Wubbolts, R.; Houweling, M.; Yan, C.; Du, H.; VanderVen, B.C.; Vaandrager, A.B.; et al. Lysosome-Mediated Degradation of a Distinct Pool of Lipid Droplets during Hepatic Stellate Cell Activation. *J. Biol. Chem.* **2017**, *292*, 12436–12448. [CrossRef] [PubMed]
63. Wang, J.; Guo, X.; Li, L.; Qiu, H.; Zhang, Z.; Wang, Y.; Sun, G. Application of the Fluorescent Dye BODIPY in the Study of Lipid Dynamics of the Rice Blast Fungus *Magnaporthe Oryzae*. *Molecules* **2018**, *23*, 1594. [CrossRef] [PubMed]
64. Knight, M.; Braverman, J.; Asfaha, K.; Gronert, K.; Stanley, S. Lipid Droplet Formation in *Mycobacterium Tuberculosis* Infected Macrophages Requires IFN- $\gamma$ /HIF-1 $\alpha$  Signaling and Supports Host Defense. *PLOS Pathog.* **2018**, *14*, e1006874. [CrossRef] [PubMed]
65. Masedunskas, A.; Chen, Y.; Stussman, R.; Weigert, R.; Mather, I.H.; Nusrat, A. Kinetics of Milk Lipid Droplet Transport, Growth, and Secretion Revealed by Intravital Imaging: Lipid Droplet Release Is Intermittently Stimulated by Oxytocin. *MBoC* **2017**, *28*, 935–946. [CrossRef] [PubMed]

66. Schnitzler, J.G.; Moens, S.J.B.; Tiessens, F.; Bakker, G.J.; Dallinga-Thie, G.M.; Groen, A.K.; Nieuwdorp, M.; Stroes, E.S.G.; Kroon, J. Nile Red Quantifier: A Novel and Quantitative Tool to Study Lipid Accumulation in Patient-Derived Circulating Monocytes Using Confocal Microscopy. *J. Lipid. Res.* **2017**, *58*, 2210–2219. [[CrossRef](#)] [[PubMed](#)]
67. Held, P. Lipid Accumulation in HepG2 Cells Exposed to Free Fatty Acids. Available online: <https://www.biotek.com/resources/application-notes/lipid-accumulation-in-hepg2-cells-exposed-to-free-fatty-acids/> (accessed on 14 August 2017).
68. Helmchen, F.; Denk, W. Deep Tissue Two-Photon Microscopy. *Nat. Methods* **2005**, *2*, 932–940. [[CrossRef](#)] [[PubMed](#)]
69. Wei, L.; Hu, F.; Shen, Y.; Chen, Z.; Yu, Y.; Lin, C.-C.; Wang, M.C.; Min, W. Live-Cell Imaging of Alkyne-Tagged Small Biomolecules by Stimulated Raman Scattering. *Nat. Methods* **2014**, *11*, 410–412. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

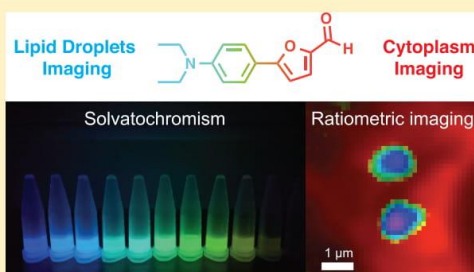


## Probing Polarity and Heterogeneity of Lipid Droplets in Live Cells Using a Push–Pull Fluorophore

Mayeul Collot,<sup>\*,†,‡</sup> Sophie Bou,<sup>†</sup> Tkhe Kyong Fam,<sup>†</sup> Ludovic Richert,<sup>†</sup> Yves Mély,<sup>†,§</sup> Lydia Danglot,<sup>‡,§</sup> and Andrey S. Klymchenko<sup>†,§</sup><sup>†</sup>Laboratoire de Biophotonique et Pathologies, Faculté de Pharmacie, UMR 7021 CNRS, Université de Strasbourg, 74, Route du Rhin, 67401 Cedex, Illkirch, France<sup>‡</sup>Institut Jacques Monod, University Paris Diderot, Sorbonne Paris Cité, CNRS UMR 7592, 75013 Paris, France<sup>§</sup>Membrane Traffic in Healthy and Diseased Brain, INSERM U894, Institute of Psychiatry and Neuroscience of Paris, 102 rue de la Santé, 75 014 Paris, France

## Supporting Information

**ABSTRACT:** Lipid droplets (LDs) are organelles composed of a lipid core surrounded by a phospholipid monolayer. Lately, LDs have attracted considerable attention due to recent studies demonstrating their role in a variety of physiological processes as well as diseases. Herein we synthesized a push–pull molecule named DAF (Dimethyl Aniline Furaldehyde) that possesses a strong positive solvatochromism in emission of 119 nm from toluene to methanol. Its impressive fluorogenic properties from water to oil (2000-fold) as well as its high quantum yields (up to 0.97) led us to investigate its ability to sense the distribution of polarity in live cells by fluorescence ratiometric imaging. When added to live cells and excited at 405 nm, DAF immediately and brightly stain lipid droplets using a blue channel (410–500 nm) and cytoplasm in a red channel (500–600 nm). DAF also proved to be compatible with fixation thus allowing 3D imaging of LDs in their cytoplasm environment. Taking advantage of DAF emission in two distinct channels, ratiometric imaging was successfully performed and led to the polarity mapping of the cell unraveling some heterogeneity in polarity within LDs of the same cell.



Lipid droplets (LDs) are organelles found in almost all organisms from prokaryotes to humans.<sup>1,2</sup> They are composed of an oily core containing metabolic lipids, mainly triglycerides and cholesterol esters, coated with a phospholipid monolayer membrane in which proteins are anchored.<sup>3,4</sup> Much more than simple lipid storages playing a role in metabolism,<sup>5–7</sup> LDs were recently found to be involved in various physiological processes including membrane trafficking,<sup>8,9</sup> protein storage,<sup>10</sup> and inflammation.<sup>11</sup> Additionally, LDs have been shown to play an important role in diseases<sup>12</sup> such as neurodegeneration,<sup>13</sup> obesity, diabetes, and atherosclerosis,<sup>14</sup> viral replication,<sup>15</sup> and cancer.<sup>16–18</sup> Although several conserved functions have been reported in the literature,<sup>19</sup> recent studies tend to show that LDs' number, size, and composition can greatly vary.<sup>20–22</sup>

Therefore, imaging of LDs and studying their dynamic composition is of prior importance to understand their involvement in biological mechanisms. Recently, considerable efforts have been made to develop efficient molecular fluorescent probes for bioimaging.<sup>23,24</sup> Among them, fluorogenic probes offer background free images due to the fact that they only emit once they have reached their target.<sup>25–28</sup>

Although efficient probes are available in multiple colors for different organelles,<sup>29</sup> LDs probes must be improved as the

most commonly used ones, namely BODIPY 493/503<sup>29</sup> and Nile Red,<sup>30</sup> showed some limitations in long-term multicolor imaging due to their limited photostability and broad emission peak, respectively.<sup>31</sup>

To address this issue, continuous efforts are being made to develop new efficient LDs probes<sup>32</sup> based on push–pull<sup>33</sup> or aggregation induced emission (AIE) fluorophores.<sup>34–36</sup> We also recently introduced a family of efficient fluorogenic LD markers emitting from yellow to near-infrared called StatoMerocyanines (SMCy) that successfully served in cell and tissue imaging using mono- or two-photon excitation imaging.<sup>31</sup> In order to complete the available palette of color, blue emitting LDs markers have recently been reported.<sup>37–40</sup> Although they showed good selectivity and despite an example where two-photon excitation was used,<sup>35</sup> their brightness in imaging is quite limited due to their UV shifted excitation spectra and quite low molar absorption coefficients, thus leading to low excitation efficiency using the common 405 nm laser.

Received: September 14, 2018

Accepted: December 28, 2018

Published: December 28, 2018

Except for one example where a near-infrared fluorophore served to visualize the most viscous LDs,<sup>41</sup> LDs probes being able to report their environment such as composition or polarity using fluorescence imaging have not yet been reported. Solvatochromic dyes display different emission wavelengths depending on their environment's polarity.<sup>42</sup> By means of chemical modifications certain fluorophores like Nile Red,<sup>43</sup> pyrene,<sup>44</sup> and fluorene<sup>45</sup> were successfully used to selectively probe polarity of the plasma membrane and/or intracellular membranes, respectively, using ratiometric imaging.

Herein we developed a push–pull fluorophore, Dimethyl Aniline Furaldehyde (DAF), that displayed positive solvatochromic properties in emission. In live or fixed cells imaging, DAF is efficiently excited with the widely used 405 nm laser and stained LDs leading to background free images due to its impressive fluorogenic effect from water to oil. Although DAF showed a lower photostability compared to BODIPY 493/503 and thus was not suitable for long-term live imaging, it was successfully used in multicolor imaging with no cross talk as it is only excited by the 405 nm laser. Taking advantage of its solvatochromism, ratiometric imaging has been performed unraveling differences of polarity in the LDs population.

## ■ EXPERIMENTAL SECTION

**Synthesis.** All starting materials for synthesis were purchased from Alfa Aesar, Sigma-Aldrich, or TCI Europe and used as received unless stated otherwise. NMR spectra were recorded on a Bruker Avance III 400 MHz spectrometer. Mass spectra were obtained using an Agilent Q-TOF 6520 mass spectrometer. Synthesis and characterization of DAF is described in the [Supporting Information](#).

**Spectroscopy.** The water used for spectroscopy was Milli-Q water (Millipore), and all the solvents were spectroscopic grade. Labrafac (Labrafac Lipophile WL 1349) is composed of medium chain triglycerides, caprylic/capric triglyceride, and was obtained from Gattefossé (Saint Priest, France). Absorption and emission spectra were recorded on a Cary 400 Scan ultraviolet–visible spectrophotometer (Varian) and a FluoroMax-4 spectrofluorometer (Horiba Jobin Yvon) equipped with a thermostated cell compartment, respectively. For standard recording of fluorescence spectra, the emission was collected 10 nm after the excitation wavelength. All the spectra were corrected from a wavelength-dependent response of the detector. Quantum yields were determined by comparison with coumarine 7-(diethylamino) 3-carboxylic acid in water ( $\phi = 0.03$ ) as a reference.<sup>46</sup> Time-resolved fluorescence measurements were performed with the time-correlated, single-photon counting technique using the excitation pulses at 405 nm provided by a pulse-picked supercontinuum laser (EXR-20, NKT Photonics) equipped with an UV extension (SuperK EXTEND-UV). The emission was collected through a polarizer set at the magic angle and an 8 nm band-pass monochromator (H10, Jobin-Yvon) at the maximum of fluorophore emission. The instrumental response function was recorded with a polished aluminum reflector, and its full-width at half-maximum was 50 ps. Time-resolved intensity data were treated with an analysis using a Matlab homemade software based on Tcspcfit<sup>47</sup> and the Maximum Entropy Method (MEM).<sup>48</sup> In all cases, the  $\chi^2$  values were close to 1, and the weighted residuals as well as the autocorrelation of the residuals were distributed randomly around 0, indicating an optimal fit.

**Cytotoxicity.** Cytotoxicity of DAF was evaluated by MTT assay (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide). A total of  $1 \times 10^4$  KB cells/well were seeded in a 96-well plate 18 h prior to the experiment in Dulbecco's Modified Eagle Medium (Gibco Lifetechnologies - DMEM) complemented with 10% fetal bovine serum, gentamicin (100  $\mu\text{g}/\text{mL}$ ), L-glutamine (2 mM), nonessential amino acids (1 mM), and MEM vitamin solution (1%) and were incubated in a 5%  $\text{CO}_2$  incubator at 37 °C. After medium removal, an amount of 100  $\mu\text{L}$  of DMEM containing 20  $\mu\text{M}$ , 5  $\mu\text{M}$ , 1  $\mu\text{M}$ , or 0.2  $\mu\text{M}$  of DAF was added to the KB cells and incubated for 1 h at 37 °C (5%  $\text{CO}_2$ ). As a control, the cells were incubated with DMEM containing the same percentage of DMSO (0.5% v/v) as the solution of DAF or with triton 0.1% as a positive control of cytotoxicity. After 1 h of dye incubation, the medium was replaced by 100  $\mu\text{L}$  of MTT solution in DMEM, and the cells were incubated for 3 h at 37 °C. Then, 75  $\mu\text{L}$  of the solution was replaced by 50  $\mu\text{L}$  of DMSO (100%) and gently shaken for 15 min at room temperature. The absorbance at 540 nm was measured. Each concentration of dye was tested in sextuplicate in 3 independent assays. For each concentration, we calculated the percentage of cell viability in reference to the control DMEM + 0.5% DMSO as 0  $\mu\text{M}$  of DAF.

**Confocal Imaging.** KB cells (ATCC CCL-17) were grown in minimum essential medium (MEM, Gibco-Invitrogen) with 10% fetal bovine serum (FBS, Lonza), 1% nonessential amino acids (Gibco-Invitrogen), 1% MEM vitamin solution (Gibco-Invitrogen), 1% L-glutamine (Sigma-Aldrich), and 0.1% antibiotic solution (gentamicin, Sigma-Aldrich) at 37 °C in a humidified atmosphere containing 5%  $\text{CO}_2$ . Cells were seeded onto a chambered coverglass (IBiDi) at a density of  $5 \times 10^4$  cells/well 24 h before the microscopy measurement. For imaging, the culture medium was removed, and the attached cells were washed with Opti-MEM (Gibco-Invitrogen). The live cells were stained with DAF (2  $\mu\text{M}$ ) prior to imaging and were imaged without any washing steps. Fixed cells were fixed with PFA (4%) and were washed three times with HBSS and visualized in HBSS. The images were processed with Icy software.<sup>49</sup> The plasma membrane was stained with dsQ12S (200 nM) prior to imaging.<sup>50</sup> Mitotracker Green FM was purchased from ThermoFisher scientific.<sup>29</sup> The correlation studies were performed on several cells from different acquisitions. Lipid droplets were segmented with spot detector plugin (scale 2, sensitivity 40). Colocalization analysis was performed using Icy SODA plugin (Standard Object Distance Analysis)<sup>51</sup> that gives both coupling proportion, colocalization map, and LD size. Ratiometric images were obtained using Intensity Color Version 2.5 plugin in ImageJ software.

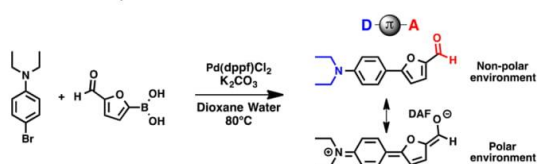
## ■ RESULTS AND DISCUSSION

**Synthesis.** As a push–pull molecule, DAF was designed as a donor- $\pi$ -acceptor system where the aniline's nitrogen is the electron donor moiety, the aniline and furan constitute the  $\pi$ -conjugated system, and the aldehyde acts as an electron acceptor (Scheme 1). DAF was synthesized in one step by Suzuki coupling between 4-bromodiethylaniline and furaldehyde boronic acid and was fully characterized by <sup>1</sup>H and <sup>13</sup>C NMR as well as high-resolution mass spectroscopy ([Supporting Information](#)).

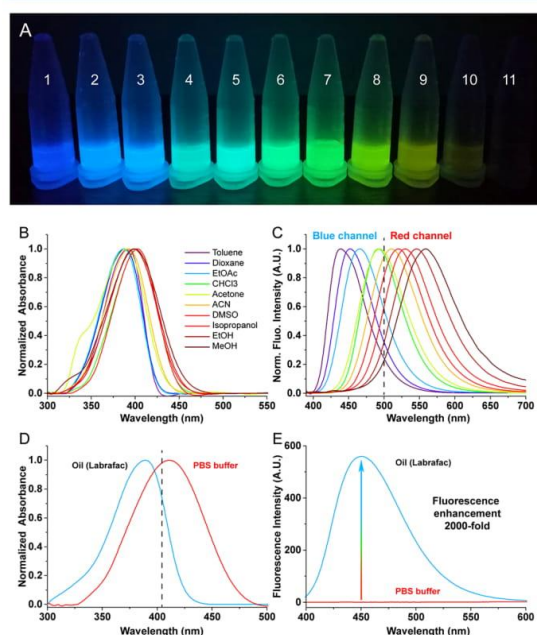
**Spectroscopic Studies.** The photophysical properties of DAF were then studied in various organic solvents with increasing polarity. When excited by a UV lamp (365 nm) the



Scheme 1. Synthesis of DAF



solutions of DAF emitted light in the visible range from deep blue to yellow (Figure 1A). Although absorption spectra



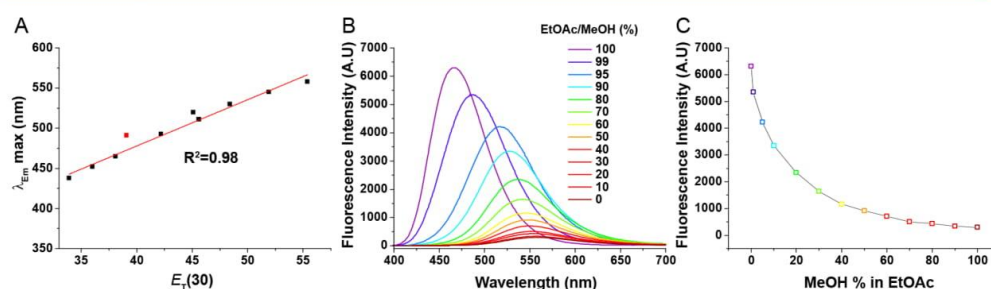
**Figure 1.** (A) Picture of DAF solutions in various solvents under UV lamp excitation (365 nm): toluene (1), dioxane (2), EtOAc (3),  $\text{CHCl}_3$  (4), acetone (5), ACN (6), DMSO (7), isopropyl alcohol (8), EtOH (9), MeOH (10), and PBS (11). Normalized absorbance (B) and emission spectra (C) of DAF in various organic solvents with increasing polarity. Normalized absorbance (D) and emission spectra (E) of DAF in PBS buffer and in oil (Labrafac) showing the impressive fluorogenic effect. Dashed line in (D) represents the 405 nm laser excitation. Concentration of DAF was  $1 \mu\text{M}$ . Excitation wavelength was 380 nm.

showed only a slight bathochromic shift (Figure 1B), emission spectra displayed a strong positive solvatochromism from toluene (439 nm) to methanol (558 nm) representing a 119 nm bathochromic shift (Figure 1C) which is in accordance with structurally close reported fluorophores.<sup>52,53</sup> Additionally, DAF showed a linear relationship between its maximum emission wavelength and the solvent polarity with a slope of 5.7 nm per  $E_T(30)$  unit (Figure 2A). Interestingly, whereas DAF is efficiently excited by the commonly used 405 nm laser (lowest: 70% efficiency), it is virtually nonexcitable by common light sources dedicated to the green channels, typically 488 nm. These features are important as regardless its environment, DAF will only emit when excited with a short wavelength laser thus avoiding cross talk phenomenon in

multicolor imaging. Importantly, DAF possesses quite high molar absorption coefficients ( $\sim 40,000 \text{ M}^{-1}\cdot\text{cm}^{-1}$ ) and can reach impressive quantum yields up to 0.97 (Table 1), making it a relatively bright blue LDs probe.<sup>32</sup> When solubilized in PBS, DAF displays a broader spectrum than in organic solvents (Figure 1D) with a significantly reduced molar absorption coefficient (see Table 1).

Consequently DAF was found to be nonemissive in water (Figure 1A and 1E). As lipid droplets are mainly composed of triglyceride, we studied the behavior of DAF in Labrafac, an oil composed of medium chain triglycerides. In this oil DAF efficiently absorbs light at 390 nm ( $\epsilon = 43,000 \text{ M}^{-1}\cdot\text{cm}^{-1}$ ) and emits in the blue region at 450 nm with a good quantum yield (0.74), thus representing a 2000-fold fluorescence enhancement from PBS to oil. In order to illustrate the environment sensitivity of DAF, it was placed in various EtOAc/MeOH ratios. From EtOAc to MeOH, emission spectra quickly shift to the red (Figure 2B) with an exponential decay of fluorescence intensity (Figure 2C). It is noteworthy that only 10% of MeOH was sufficient to decrease its maximum fluorescence intensity by half (Figure 2C). As lipid media contained in the LDs can be highly viscous,<sup>41</sup> the influence of this parameter on DAF was assessed using methanol–glycerol mixtures with increasing viscosity (Figure S1). From a low viscosity value (0.6 mPa·s) to a high one (1317 mPa·s), the fluorescence enhancement was only 1.6-fold showing a rather low sensitivity of DAF toward the viscosity. Altogether these results show that, in addition to typical twisted intramolecular charge transfer (TICT) observed for push–pull fluorophores,<sup>54</sup> DAF could be quenched in polar protic media by other mechanisms, such as electron transfer and H-bonding with the solvent.<sup>55</sup> Lifetime measurements were in line with these hypotheses as the obtained values ranged from 1.71 to 3.02 ns in aprotic solvents with increasing polarity, whereas in ethanol a lower value of 1.29 ns was obtained (Figure S2). Overall, efficient fluorogenic properties displaying an impressive fluorescence enhancement from PBS to oil combined with a high environment sensitivity make DAF a promising candidate as blue fluorescent LDs probe for multicolor imaging and for sensing LDs polarity.

**Cellular Imaging.** As a first experiment we checked the cytotoxicity of DAF and showed that it did not affect the cell viability after 1 h incubation at concentrations up to  $20 \mu\text{M}$  (Figure S3). DAF was then added to live KB cells (HeLa cells derivatives) at a concentration of  $2 \mu\text{M}$ . Laser scanning confocal imaging revealed that DAF immediately and brightly stained LDs (Figure 3A) using a 405 nm laser source and collecting the emission in a “blue channel” (415–480 nm). Additionally, using the same excitation source and collecting the emitted light in a second “red channel” (500–600 nm), DAF stained the cytoplasm (Figure 3B). These experiments showed that DAF possesses both advantages of BODIPY 493/503 and Nile Red without their drawbacks. Indeed similarly to BODIPY 493/503, DAF immediately penetrates the cells and stains LDs. However, we recently showed that BODIPY 493/503 has a quantitative quantum yield in water,<sup>31</sup> whereas DAF is nonemissive (Figure 1E) thus leading to images with significantly reduced background noise compared to BODIPY 493/503. As solvatochromic dyes, DAF and Nile Red can both report LDs and cytoplasmic membranes using two emission channels (Figure 3C). However, due to its blue-shifted excitation, DAF is only excitable with the 405 nm laser, whereas Nile Red, due to its red-shifted and broad absorption



**Figure 2.** Solvatochromic and fluorogenic properties of DAF. Linear relationship between the maximum emission wavelength and the solvent's polarity (A). Chloroform was excluded from the fit as other parameters might interfere.<sup>56</sup> Emission spectra of DAF in various EtOAc/MeOH ratios (B) and its exponential decay of fluorescence intensity denoting fluorogenic properties (C). Concentration of DAF was 1  $\mu\text{M}$ . Excitation wavelength was 380 nm.

**Table 1. Photophysical Properties of DAF (1  $\mu\text{M}$ ) in Various Solvents**

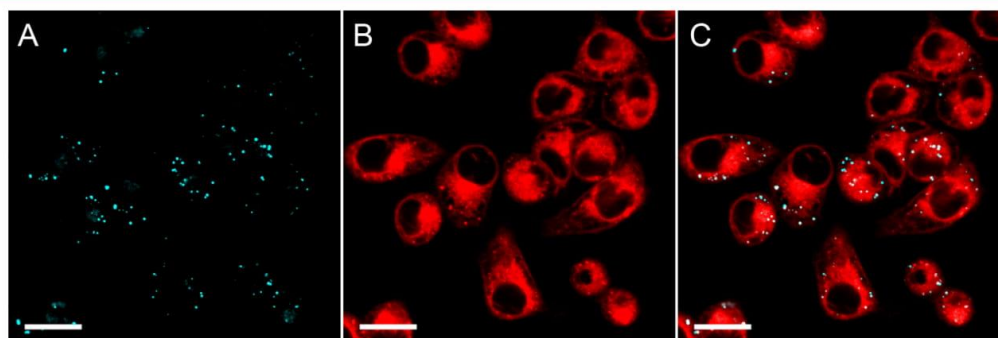
solvent	$E_T(30)^a$	$\lambda_{\text{Abs max}}^b$ (nm)	$\lambda_{\text{Em max}}^b$ (nm)	$\epsilon$ ( $\text{M}^{-1}\cdot\text{cm}^{-1}$ )	$\phi^b$	fwhm <sup>c</sup> (nm)
toluene	33.9	392	439	36,000	0.79	62
oil <sup>d</sup>	N/A	390	450	43,000	0.74	72
dioxan	36	388	452	40,000	0.19	71
EtOAc	38.1	386	465	44,000	0.17	67
$\text{CHCl}_3$	39.1	403	491	40,500	0.85	66
acetone	42.2	389	493	40,000	0.64	75
ACN	45.6	392	511	39,500	0.75	79
DMSO	45.1	403	520	34,500	0.97	79
isopropyl alcohol	48.4	398	530	37,500	0.57	80
EtOH	51.9	398	545	39,000	0.26	87
MeOH	55.4	400	558	37,500	0.06	94
PBS	63.1	411	N/A	19,500	N/A	N/A

<sup>a</sup> $E_T(30)$  is the molar transition energy expressed in  $\text{kcal}\cdot\text{mol}^{-1}$  and is a solvent polarity parameter.<sup>42</sup> <sup>b</sup>Quantum yields were measured using coumarin 7-(diethylamino)-3-carboxylic acid in water ( $\phi = 0.03$ ) as a reference.<sup>46</sup> <sup>c</sup>Full width at half-maximum emission spectrum. <sup>d</sup>Labrafac was used (medium chain triglycerides).

and emission spectra, is prone to cross talk.<sup>31</sup> In situ emission spectra of DAF in various regions of the cells were then performed and showed that two main peaks could be observed (Figure S4). In LDs, DAF emits at 440 nm with a quite sharp

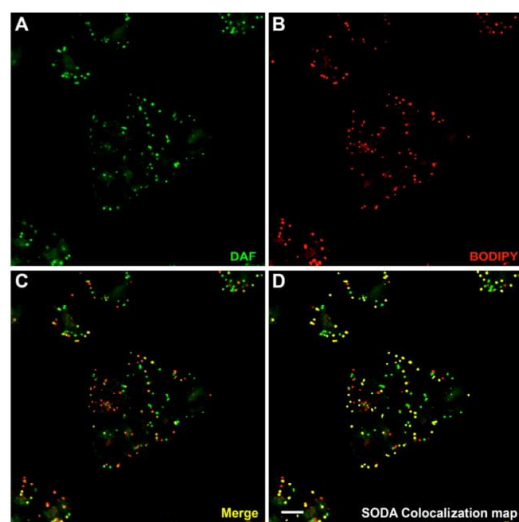
band (fwhm = 56 nm), showing that it is in a very low polar environment (polarity similar to toluene, see Table 1). In the cytoplasm, DAF emits in a broader manner with a peak at 510 nm. In order to assign this peak more precisely, we placed DAF in an aqueous solution of a protein (BSA). The results showed that DAF interacts with BSA and emits at 518 nm, a value close to the one measured *in situ* in the cytoplasm (Figure S5). This similarity led us to think that DAF might interact with the hydrophobic sites of proteins of the cytoplasm. Interestingly, when cells were imaged at different times of incubation, it was noteworthy that whereas the LDs could be detected with no apparent loss of signal, the staining of the cytoplasm in the red channel decreased over the time probably denoting a slow accumulation of the dye from the cytoplasm to the LDs (Figure S6).

At this point, after verifying that cross talk did not occur (Figure S7), we checked whether DAF colocalized with BODIPY 493/503 (Figure 4). First, it was noteworthy that DAF and BODIPY 493/503 displayed significant difference in their intensity within the LDs (Figure 4C). Conventional Pearson coefficient analysis gave a value of  $0.6995 \pm 0.0154$  indicating rather good, though not complete correlation between DAF and BDP intensity (Figure S8). In order to further study this observation, we used a recently developed colocalization technique based on mass center colocalization (see details in the SI).<sup>51</sup> 59% of DAF positive LDs ( $n = 755$ ) were found to be associated with BODIPY 493/503, and



**Figure 3.** Laser scanning confocal images of live KB cells stained with DAF (2  $\mu\text{M}$ ). (A) Blue channel:  $\lambda_{\text{Ex}} = 405$  nm,  $E_m$ : 415–480 nm. (B) Red channel:  $\lambda_{\text{Ex}} = 405$  nm,  $E_m$ : 500–600 nm. (C) is the merge of the blue and red channels. DAF was added prior to imaging. No washing steps were performed. Scale bar is 20  $\mu\text{m}$ .



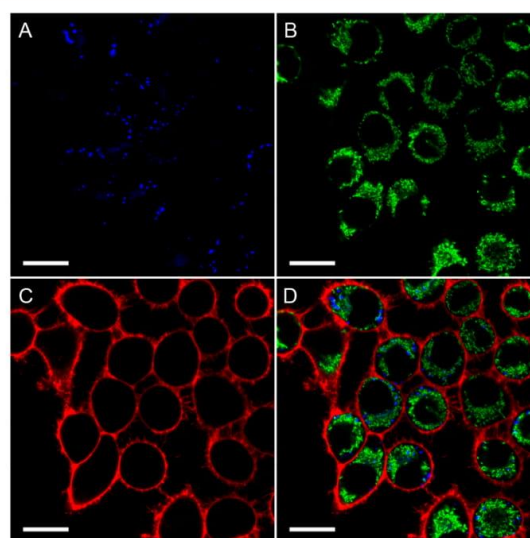


**Figure 4.** Colocalization experiments. KB cells were stained with DAF ( $2 \mu\text{M}$ ,  $\lambda_{\text{Ex}} = 405 \text{ nm}$ ,  $E_{\text{m}}: 415\text{--}480 \text{ nm}$ ) (A) and BODIPY 493/503 ( $200 \text{ nM}$ ,  $\lambda_{\text{Ex}} = 488 \text{ nm}$ ,  $E_{\text{m}}: 495\text{--}550 \text{ nm}$ ) (B). (C) is the merge of A and B. (D) is the composite picture of the segmented fluorescent spots identified with Icy SODA colocalization analysis plugin. Localization of DAF spots was analyzed with respect to BODIPY staining. Green spots are the DAF isolated spots found without BODIPY, yellow spots are the DAF spots associated with BODIPY staining, red ones are the BODIPY spots found alone. Scale bar is  $10 \mu\text{m}$ .

conversely 62% of BODIPY 493/503 positive LD ( $n = 729$ ) were associated with DAF. This association was statistically highly significant ( $p$  value  $10^{-216}$ ) and is depicted as yellow spots on Figure 4D. Very interestingly, this study showed that DAF stained LDs that BODIPY 493/503 could not detect and *vice versa* (Figure 4D, green and red spots). These data tend to show that DAF and BODIPY 493/503 are complementary LDs markers and might be used in combination in order to unravel a larger population of LDs. Moreover we showed that LDs only positive for DAF or BODIPY have the same pixel size distribution (Figure S9), whereas LDs positive for both markers (DAF and BODIPY) are bigger ( $p < 0.0001$ ) than those labeled with only one marker (see Figure S9). These results question the universality of BODIPY 493/503 as a LDs marker and provide new evidence of LDs' heterogeneity.<sup>57,58</sup>

Taking advantage of these features, DAF was first involved in multicolor cellular imaging (Figure 5). LDs were stained with DAF and excited at 405 nm, and the emission was only collected in the blue channel. Mitochondria were stained with Mitotracker green FM using 488 nm excitation and displayed no cross talk with DAF as the latter is not excitable at this wavelength (see Figure 1B). Finally, the plasma membrane was stained with dSQ12S<sup>50</sup> using 635 nm excitation. While these imaging experiments were performed without any washing step, the different organelles were distinctly revealed by its corresponding color with virtually no background signal (Figure 5D).

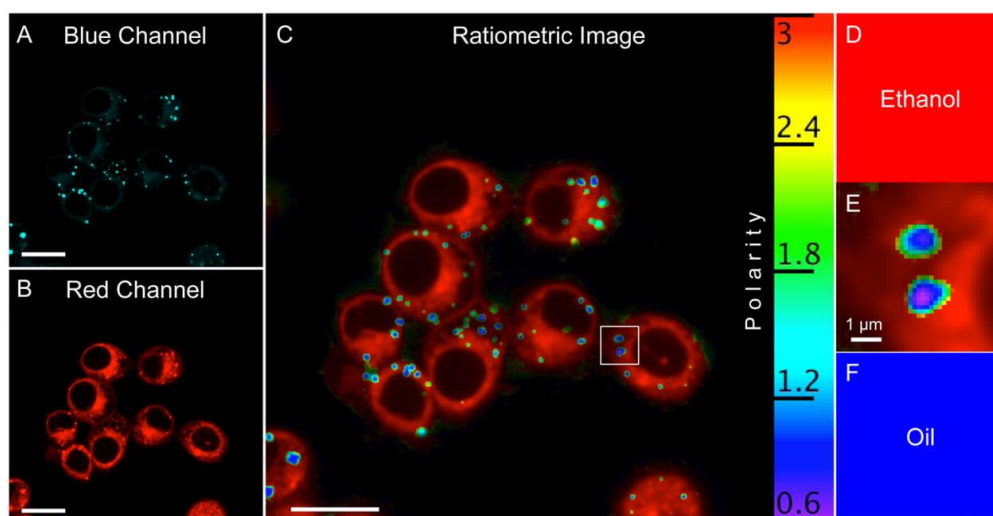
DAF was then challenged in fixed cells. KB cells were stained with DAF and subsequently fixed. The cells were then imaged by laser scanning confocal microscopy, and images were acquired at increasing heights (Z-Stack experiment) using 405



**Figure 5.** Multicolor imaging of live KB cells using DAF as a blue emitting LDs marker. A is the blue channel: DAF ( $2 \mu\text{M}$ ),  $\lambda_{\text{Ex}} = 405 \text{ nm}$ ,  $E_{\text{m}}: 415\text{--}480 \text{ nm}$ . B is the green channel: Mito tracker Green FM ( $200 \text{ nM}$ ),  $\lambda_{\text{Ex}} = 488 \text{ nm}$ ,  $E_{\text{m}}: 495\text{--}550 \text{ nm}$ . C is the red channel: dSQ12S ( $200 \text{ nM}$ ),  $\lambda_{\text{Ex}} = 635 \text{ nm}$ ,  $E_{\text{m}}: 645\text{--}750 \text{ nm}$ . D is the merge of the 3 channels. Unlike the Mito tracker, which was incubated for 1 h, DAF and dSQ12S were added just before imaging. No washing steps were performed. Scale bar is  $20 \mu\text{m}$ .

nm laser as single excitation source and collecting the fluorescence signals in the two distinct channels (blue: 415–480 nm and red: 500–600 nm). Although the photostability of DAF was found to be lower compared to BODIPY 493/503 (Figure S10) and fixation seemed to lower the brightness of DAF compared to live cell imaging, it was shown to be sufficiently bright and photostable over the required 67 successive acquisitions to provide good quality images (Figure S11). Consequently, maximum projection of the obtained images as well as 3D reconstruction unraveled LDs in their cytoplasmic environment (Figure S12). These results showed that, owing to its solvatochromism, DAF enables imaging LDs and the cytoplasm with only one excitation source without blocking the remaining channels that use other excitation sources such as 488, 530, 560, and 630 nm.

Finally, taking advantage of the solvatochromism properties of DAF, we investigated its ability to probe the difference in polarity within the LDs using ratiometric imaging. To do so, the signal of the red channel (Figure 6B) was divided by the signal of the blue channel (Figure 6A) thus providing a ratiometric image that maps the polarity of the cytoplasm and the LDs (Figure 6C). These images indicated that the cytoplasm displayed the highest polarity (ratio value 3.0) with a high homogeneity. As a comparison, a solution of DAF in pure ethanol was imaged using the same parameters and provided a similar ratio color indicating that the labeled cytoplasm exhibits rather high polarity that could be assigned to water exposed binding sites of intracellular proteins (Figure 6D). Conversely, LDs displayed significantly lower polarities with an increasing gradient from their core to the surface. Remarkably, the core composed of triglycerides and cholesterol esters displayed the lowest polarity (ratio value between



**Figure 6.** Laser scanning confocal images of KB cells in the presence of DAF ( $2 \mu\text{M}$ ). (A) Image obtained in the blue channel:  $\lambda_{\text{Ex}} = 405 \text{ nm}$ ,  $E_{\text{m}}$ : 415–480 nm. (B) Image obtained in the red channel:  $\lambda_{\text{Ex}} = 405 \text{ nm}$ ,  $E_{\text{m}}$ : 500–600 nm. (C) Ratiometric image obtained from the ratio between the red and the blue channels, mapping the polarity distribution in live cells. The inset in C was zoomed in E showing two LDs with heterogeneous polarity, and D and F are ratiometric images using the same parameters performed with DAF ( $10 \mu\text{M}$ ) in pure ethanol and pure oil (Labrafac), respectively. Scale bar is  $20 \mu\text{m}$ .

0.6 and 1) comparable to the one obtained when DAF in pure oil was imaged (Figure 6F). The difference between the core and the surface of LDs could be explained, on one hand, by polarity gradient within the LDs, where the water exposed membrane should be much more polar than the lipid core. On the other hand, the size of droplets is close to the diffraction-limited resolution of the confocal microscope. Therefore, we could also explain higher values of the red/blue ratio at the edges of LDs as a superposition of signal coming from the lipid core (low ratio values) and the cytoplasm (high ratio values). Dedicated super-resolution microscopy studies using stimulated emission depletion (STED), structured illumination microscopy (SIM),<sup>59</sup> or direct stochastic optical reconstruction microscopy (dSTORM)<sup>60</sup> will be needed in order to better understand the distribution of polarity within the same LD. Moreover, Figure 6E revealed a heterogeneity in polarity observed in the center of LDs as one displays the ratio value close to the oil of medium-chain triglycerides (Labrafac, Figure 6F), and for the second one it is even lower (0.6), probably denoting a difference in the lipid composition in the core of these LDs.

## CONCLUSION

We herein synthesized and characterized a push–pull fluorophore, DAF, that displays a 119 nm positive solvatochromism in emission. Using a single excitation source, DAF was shown to immediately stain the LDs in live cells using a blue emission channel and the cytoplasm using a red emission channel. Colocalization experiments showed that DAF unraveled LDs that could not be detected by BODIPY 493/503 and *vice versa*, thus showing that DAF is complementary to BODIPY 493/503. Due to its blue-shifted excitation spectra, DAF proved to be compatible with multicolor fluorescence imaging using green and red fluorescent markers. Additionally, due to its fluorogenic

properties DAF does not require any washing steps and provided images displaying virtually no background signal. Taking advantage of its solvatochromism and using a red-shifted fluorescent channel, DAF provided ratiometric images leading to the polarity mapping of the cytoplasm and LDs. DAF also showed some variation in polarity between the core and the surface of LDs; however, a more dedicated study using super-resolution imaging will be needed to confirm this observation. This feature might be of importance to study the variation LDs membrane composition especially during interaction with other organelles and fusion. Interestingly, we demonstrated that, within the LDs population, their lipid core displayed various polarities probably denoting heterogeneity in their composition. Studying the composition of metabolic lipids by imaging might accelerate LDs studies and open the way to a better comprehension of not only metabolic disorders and associated diseases such as atherosclerosis<sup>14</sup> but also Alzheimer's disease and epilepsy.<sup>61</sup> In conclusion, DAF is a blue emitting LDs marker suitable for multicolor live and fixed cells fluorescence imaging and that can also efficiently probe the heterogeneity in polarity of LDs. This work contributes to a new field of research namely fluorescent and environment-sensitive LDs probes that consist not only in staining LDs but also in sensing their environment like their polarity, their composition, or their viscosity.<sup>41</sup> Further works are ongoing in our laboratory to find a correlation between the polarity of LDs and their metabolic lipid composition.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.8b04218.

Synthesis of DAF, protocols, characterizations including NMR and mass spectra, cytotoxicity assay, additional



spectroscopic data, and imaging experiments and controls (PDF)

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: mayeul.collet@unistra.fr.

### ORCID

Mayeul Collet: 0000-0002-8673-1730

Yves Mély: 0000-0001-7328-8269

Andrey S. Klymchenko: 0000-0002-2423-830X

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank Romain Vauchelles for his assistance at the PIQ imaging platform and for the Ratio Intensity Color Version 2.5 plugin. We acknowledge the Service de Chimie Analytique (SCA) for LC-MS and RMN analyses. This work was supported by ERC Consolidator grant BrightSens 648528, ANR BrightRiboProbes (ANR-16-CE11-0010).

## REFERENCES

- (1) Murphy, D. J. *Protoplasma* **2012**, *249* (3), 541–585.
- (2) Waltermann, M.; Steinbuchel, A. *J. Bacteriol.* **2005**, *187* (11), 3607–3619.
- (3) Farese, R. V.; Walther, T. C. *Cell* **2009**, *139* (5), 855–860.
- (4) Martin, S.; Parton, R. G. *Nat. Rev. Mol. Cell Biol.* **2006**, *7* (5), 373–378.
- (5) Walther, T. C.; Farese, R. V. F., Jr. *Annu. Rev. Biochem.* **2012**, *81* (1), 687–714.
- (6) Thiam, A. R.; Jr Farese, R. V.; Walther, T. C. *Nat. Rev. Mol. Cell Biol.* **2013**, *14* (12), 775–786.
- (7) Cao, H.; Gerhold, K.; Mayers, J. R.; Wiest, M. M.; Watkins, S. M.; Hotamisligil, G. S. *Cell* **2008**, *134* (6), 933–944.
- (8) Zehmer, J. K.; Huang, Y.; Peng, G.; Pu, J.; Anderson, R. G. W.; Liu, P. *Proteomics* **2009**, *9* (4), 914–921.
- (9) Blom, T.; Somerharju, P.; Ikonen, E. *Cold Spring Harbor Perspect. Biol.* **2011**, *3* (8), a004713.
- (10) Li, Z.; Thiel, K.; Thul, P. J.; Beller, M.; Kühnlein, R. P.; Welte, M. A. *Curr. Biol.* **2012**, *22* (22), 2104–2113.
- (11) Bozza, P. T.; Viola, J. P. B. *Prostaglandins, Leukotrienes Essent. Fatty Acids* **2010**, *82* (4–6), 243–250.
- (12) Welte, M. A. *Curr. Biol.* **2015**, *25* (11), R470–481.
- (13) Liu, L.; Zhang, K.; Sandoval, H.; Yamamoto, S.; Jaiswal, M.; Sanz, E.; Li, Z.; Hui, J.; Graham, B. H.; Quintana, A.; et al. *Cell* **2015**, *160* (1–2), 177–190.
- (14) Kraemer, N.; Farese, R. V.; Walther, T. C. *EMBO Mol. Med.* **2013**, *5* (7), 973–983.
- (15) Herker, E.; Harris, C.; Hernandez, C.; Carpentier, A.; Kaehlcke, K.; Rosenberg, A. R.; Farese, R. V.; Ott, M. *Nat. Med.* **2010**, *16* (11), 1295–1298.
- (16) Tirinato, L.; Pagliari, F.; Limongi, T.; Marini, M.; Falqui, A.; Seco, J.; Candeloro, P.; Liberale, C.; Di Fabrizio, E. *Stem Cells Int.* **2017**, *2017*, 1656053.
- (17) Liu, Q.; Luo, Q.; Halim, A.; Song, G. *Cancer Lett.* **2017**, *401* (Supplement C), 39–45.
- (18) Abramczyk, H.; Surmacki, J.; Kopeć, M.; Olejnik, A. K.; Lubecka-Pietruszewska, K.; Fabianowska-Majewska, K. *Analyst* **2015**, *140* (7), 2224–2235.
- (19) Zhang, C.; Liu, P. *Protein Cell* **2017**, *8* (11), 796–800.
- (20) Thiam, A. R.; Beller, M. *J. Cell Sci.* **2017**, *130* (2), 315–324.
- (21) Bader, C. A.; Shandala, T.; Carter, E. A.; Ivask, A.; Guinan, T.; Hickey, S. M.; Werrett, M. V.; Wright, P. J.; Simpson, P. V.; Stagni, S.; et al. *PLoS One* **2016**, *11* (8), No. e0161557.
- (22) Ghosh, C.; Nandi, S.; Bhattacharyya, K. *Chem. Phys. Lett.* **2017**, *670*, 27–31.
- (23) Lavis, L. D. *Annu. Rev. Biochem.* **2017**, *86* (1), 825–843.
- (24) Yang, Z.; Cao, J.; He, Y.; Yang, J. H.; Kim, T.; Peng, X.; Kim, J. S. *Chem. Soc. Rev.* **2014**, *43* (13), 4563–4601.
- (25) Nadler, A.; Schultz, C. *Angew. Chem., Int. Ed.* **2013**, *52* (9), 2408–2410.
- (26) Li, X.; Gao, X.; Shi, W.; Ma, H. *Chem. Rev.* **2014**, *114* (1), 590–659.
- (27) Su, D.; Teoh, C. L.; Wang, L.; Liu, X.; Chang, Y.-T. *Chem. Soc. Rev.* **2017**, *46* (16), 4833–4844.
- (28) Klymchenko, A. S. *Acc. Chem. Res.* **2017**, *50* (2), 366–375.
- (29) Life Technologies. <https://www.thermofisher.com/> (accessed Jan 3, 2019).
- (30) Greenspan, P.; Mayer, E. P.; Fowler, S. D. *J. Cell Biol.* **1985**, *100* (3), 965–973.
- (31) Collet, M.; Fam, T. K.; Ashokkumar, P.; Faklaris, O.; Galli, T.; Danglot, L.; Klymchenko, A. S. *J. Am. Chem. Soc.* **2018**, *140* (16), 5401–5411.
- (32) Fam, T.; Klymchenko, A.; Collet, M. *Materials* **2018**, *11* (9), 1768.
- (33) Appelqvist, H.; Stranius, K.; Börjesson, K.; Nilsson, K. P. R.; Dyrager, C. *Bioconjugate Chem.* **2017**, *28* (5), 1363–1370.
- (34) Zheng, Z.; Zhang, T.; Liu, H.; Chen, Y.; Kwok, R. T. K.; Ma, C.; Zhang, P.; Sung, H. H. Y.; Williams, I. D.; Lam, J. W. Y.; et al. *ACS Nano* **2018**, *12* (8), 8145–8159.
- (35) Jiang, M.; Gu, X.; Lam, J. W. Y.; Zhang, Y.; Kwok, R. T. K.; Wong, K. S.; Tang, B. Z. *Chem. Sci.* **2017**, *8* (8), 5440–5446.
- (36) Wang, Z.; Gui, C.; Zhao, E.; Wang, J.; Li, X.; Qin, A.; Zhao, Z.; Yu, Z.; Tang, B. Z. *ACS Appl. Mater. Interfaces* **2016**, *8* (16), 10193–10200.
- (37) Kuntam, S.; Puskás, L. G.; Ayaydin, F. *Plant Cell Rep.* **2015**, *34* (4), 655–665.
- (38) Becerra-Ruiz, M.; Vargas, V.; Jara, P.; Tirapegui, C.; Carrasco, C.; Nuñez, M.; Lezana, N.; Galdámez, A.; Vilches-Herrera, M. *Eur. J. Org. Chem.* **2018**, *2018* (34), 4795–4801.
- (39) Yang, H.-J.; Hsu, C.-L.; Yang, J.-Y.; Yang, W. Y. *PLoS One* **2012**, *7* (3), e32693.
- (40) Sk, B.; Thakre, P. K.; Tomar, R. S.; Patra, A. *Chem. - Asian J.* **2017**, *12* (18), 2501–2509.
- (41) Guo, R.; Yin, J.; Ma, Y.; Li, G.; Wang, Q.; Lin, W. *Sens. Actuators, B* **2018**, *271*, 321–328.
- (42) Reichardt, C. *Chem. Rev.* **1994**, *94* (8), 2319–2358.
- (43) Kucherak, O. A.; Oncul, S.; Darwich, Z.; Yushchenko, D. A.; Arntz, Y.; Didier, P.; Mély, Y.; Klymchenko, A. S. *J. Am. Chem. Soc.* **2010**, *132* (13), 4907–4916.
- (44) Niko, Y.; Didier, P.; Mely, Y.; Konishi, G.; Klymchenko, A. S. *Sci. Rep.* **2016**, *6* (1), 18870.
- (45) Shaya, J.; Collet, M.; Benailly, F.; Mahmoud, N.; Mély, Y.; Michel, B. Y.; Klymchenko, A. S.; Burger, A. *ACS Chem. Biol.* **2017**, *12* (12), 3022.
- (46) Chatterjee, A.; Seth, D. *Photochem. Photobiol.* **2013**, *89* (2), 280–293.
- (47) Enderlein, J.; Erdmann, R. *Opt. Commun.* **1997**, *134* (1), 371–378.
- (48) Smith, D. A.; McKenzie, G.; Jones, A. C.; Smith, T. A. *Methods Appl. Fluoresc.* **2017**, *5* (4), No. 042001.
- (49) de Chaumont, F. de; Dallongeville, S.; Chenouard, N.; Hervé, N.; Pop, S.; Provoost, T.; Meas-Yedid, V.; Pankajakshan, P.; Lecomte, T.; Montagner, Y. L.; et al. *Nat. Methods* **2012**, *9* (7), 690–696.
- (50) Collet, M.; Kreder, R.; Tatars, A. L.; Patsenker, L. D.; Mely, Y.; Klymchenko, A. S. *Chem. Commun.* **2015**, *51* (96), 17136–17139.
- (51) Lagache, T.; Grassart, A.; Dallongeville, S.; Faklaris, O.; Sauvonnnet, N.; Dufour, A.; Danglot, L.; Olivo-Marin, J.-C. *Nat. Commun.* **2018**, *9* (1), 698.
- (52) Ando, Y.; Homma, Y.; Hiruta, Y.; Citterio, D.; Suzuki, K. *Dyes Pigm.* **2009**, *83* (2), 198–206.
- (53) Kucherak, O. A.; Didier, P.; Mély, Y.; Klymchenko, A. S. *J. Phys. Chem. Lett.* **2010**, *1* (3), 616–620.

- (54) Sasaki, S.; Drummen, G. P. C.; Konishi, G. *J. Mater. Chem. C* **2016**, *4* (14), 2731–2743.
- (55) Kucherak, O. A.; Richert, L.; Mély, Y.; Klymchenko, A. S. *Phys. Chem. Chem. Phys.* **2012**, *14* (7), 2292–2300.
- (56) Ooyama, Y.; Asada, R.; Inoue, S.; Komaguchi, K.; Imae, I.; Harima, Y. *New J. Chem.* **2009**, *33* (11), 2311–2316.
- (57) Herms, A.; Bosch, M.; Ariotti, N.; Reddy, B. J. N.; Fajardo, A.; Fernández-Vidal, A.; Alvarez-Guaita, A.; Fernández-Rojo, M. A.; Rentero, C.; Tebar, F.; et al. *Curr. Biol.* **2013**, *23* (15), 1489–1496.
- (58) Zhang, S.; Wang, Y.; Cui, L.; Deng, Y.; Xu, S.; Yu, J.; Cichello, S.; Serrero, G.; Ying, Y.; Liu, P. *Sci. Rep.* **2016**, *6* (1), 29539.
- (59) Zheng, X.; Zhu, W.; Ni, F.; Ai, H.; Yang, C. *Sens. Actuators, B* **2018**, *255*, 3148–3154.
- (60) Eggert, D.; Rösch, K.; Reimer, R.; Herker, E. *PLoS One* **2014**, *9* (7), No. e102511.
- (61) Welte, M. A. *Curr. Biol.* **2015**, *25* (11), R470–481.

## 8. Résumé de thèse (version française)

### **Sondes fluorogènes dimérisés brillantes pour l'imagerie d'acide nucléique et de protéines**

#### 1) Introduction

Les techniques d'imagerie par fluorescence permettent de mieux comprendre la complexité des systèmes vivants et de révéler les mécanismes de leurs fonctionnalités dans des conditions physiologiques et pathologiques. Le principal avantage des sondes fluorescentes pour l'imagerie moléculaire est leur réponsivité vis-à-vis de leur environnement. Cette caractéristique implique le développement de sondes fluorogènes sélectives de cibles biologiques.<sup>2</sup>

Les sondes fluorogènes sensibles à l'environnement peuvent être basées sur différents mécanismes, tels que le transfert de charge / proton, le changement de conformation, etc.<sup>4</sup> Une stratégie alternative pour concevoir des biosondes fluorogéniques plus brillantes est la stratégie d'agrégation. La désactivation induite par agrégation (ACQ) est à l'origine de fortes interactions hydrophobes des fluorophores et est généralement observée à des concentrations élevées.<sup>1</sup> Ce phénomène est commun pour les fluorophores peu solubles dans les milieux aqueux, généralement les fluorophores hydrophobes (squaraines, cyanines, BODIPY). L'agrégation est thermodynamiquement favorable et difficile à éviter. Par conséquent, l'ACQ peut générer une réponse fluorogénique si les espèces initialement désactivées sont spatialement séparées lors de l'interaction avec la cible biologique.<sup>16</sup> Notre équipe a montré que la dimérisation de squaraines conjugués de manière covalente par un espaceur et avec un ligand permettait d'imager les récepteurs d'oxytocine dans des cellules vivantes.<sup>35</sup>

Dans ce travail, nous avons fait une analyse systématique de l'effet de la structure sur les propriétés fluorogènes des dimères de fluorophores pour la bio-imagerie sélective. Nous avons élargi le nombre de biosondes basées sur la stratégie de désactivation induite par la dimérisation pour la détection et l'imagerie de l'ARN et de récepteurs de la biotine.

## 2) Résultats et discussions

### ***2.1 Un module de fluorogène-aptamère basé sur la dimérisation pour l'imagerie ciblée d'ARN dans des cellules de mammifère***

L'ARN est une biomolécule importante dans la régulation des gènes et la traduction des protéines. Pour visualiser la dynamique et la localisation des ARN intracellulaires, de nombreuses approches ont été proposées. Les techniques traditionnelles d'imagerie par fluorescence pour la visualisation l'ARN sont basées sur des méthodes d'hybridation et les protéines de liaison à l'ARN (RBP).<sup>42,46,55</sup> Bien que ces techniques aient mené à des découvertes importantes en biologie cellulaire, elles possèdent un certain nombre d'inconvénients limitant leur champ d'application.

Récemment, une nouvelle approche d'imagerie ARN alternative a été introduite à l'aide d'aptamères d'ARN. Les aptamères d'ARN sont de courts oligonucléotides qui reconnaissent leur cible avec une sélectivité extrêmement élevée. Plusieurs aptamères d'ARN qui reconnaissent spécifiquement des fluorogènes ou fluorophores ont été développés à l'aide de l'évolution systématique de ligands par enrichissements exponentiels (SELEX).<sup>55</sup> La stratégie des aptamères fluorogènes est une approche très attrayante, car les aptamères d'ARN sont très modulables et leur structure peut être rapidement optimisée par SELEX.<sup>55,64</sup> Cependant, les systèmes d'ARN aptamère-fluorogènes établis ont montré des performances limitées pour l'imagerie des ARN en cellules vivantes. L'absence d'affinité du fluorogène pour l'aptamère, la cytotoxicité du fluorogène et la fluorescence de base et non spécifique rendent, en pratique, leur utilisation très limitée en laboratoire. Récemment, un aptamère appelé 'Spinach' a été mis au point et a révolutionné l'imagerie d'ARN en cellules de mammifère vivantes.<sup>61</sup> 'Spinach' se lie spécifiquement au colorant fluorogène, le 3,5-difluoro-4-hydroxybenzylidène imidazolinone (DFHBI, dérivé du fluorophore GFP) pour produire une forte fluorescence verte.<sup>61</sup> 'Spinach' et le colorant DFHI ont été optimisés pour obtenir des aptamères plus stables, plus petits et plus brillants.<sup>63,67,73</sup> En outre, un aptamère 'Mango', présentant une affinité extrêmement élevée pour le thiazole orange, a été mis au point.<sup>73</sup> Récemment, un aptamère SRB-2 (liant la SulfoRhodamine B) se liant également à des fluorogènes quenchés par contact.<sup>62</sup> Ces

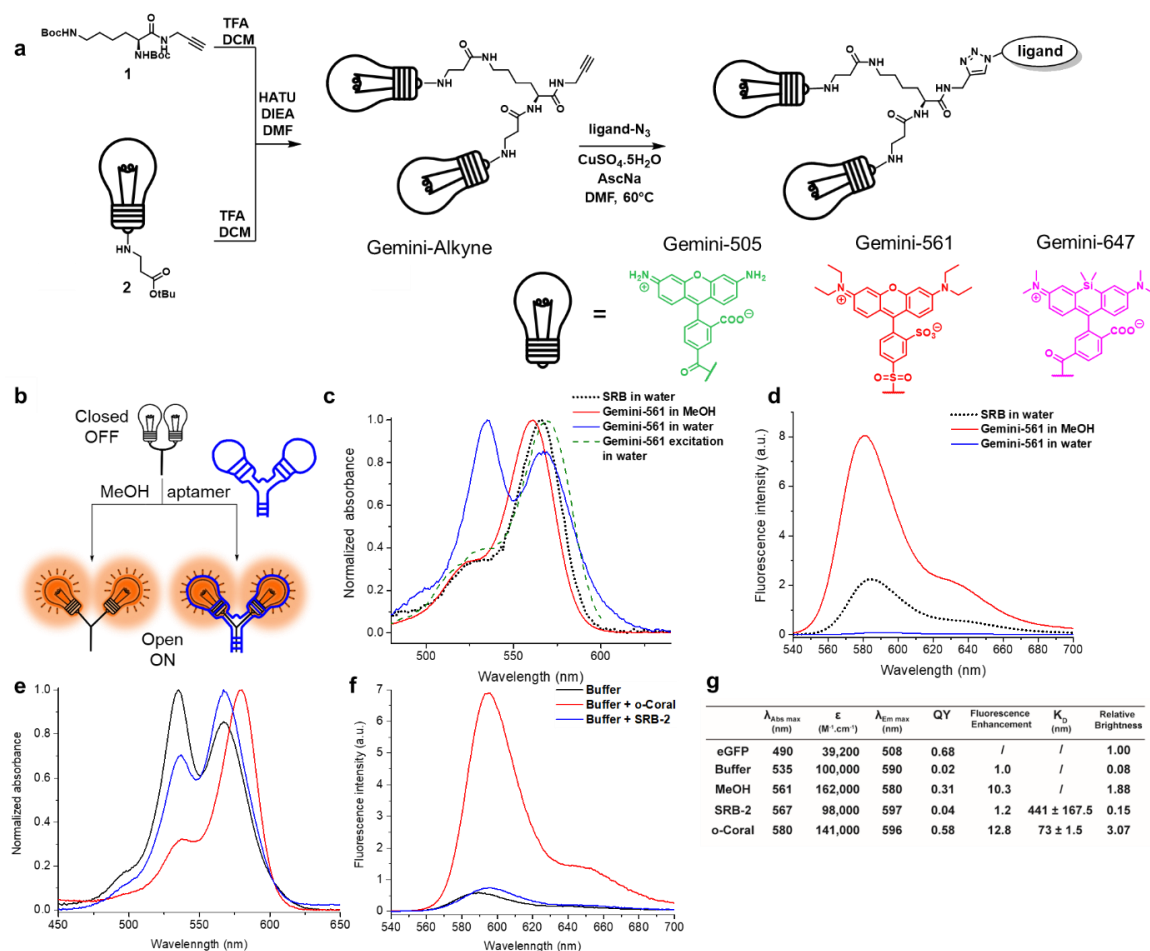
systèmes sont efficaces mais souffrent généralement d'une faible brillance car un seul fluorophore est activé par l'aptamère.

Dans ce projet, nous avons pour objectif de développer un fluorogène avec son aptamère correspondant, basé sur une stratégie de dimérisation du fluorogène (sulforhodamine B).

#### *Conception et synthèse de Gemini-561*

Les rhodamines sont des fluorophores couramment utilisés en imagerie de fluorescence. En raison de leur tendance à former des H-agrégats<sup>155</sup> et de leur aptitude à être neutralisés par différents systèmes, les rhodamines constituent des plateformes efficaces pour développer des capteurs fluorogéniques fiables. Le SRB présente des propriétés photophysiques optimales, notamment un rendement quantique élevé, une bonne photostabilité et un haut coefficient d'absorption molaire ( $\sim 100\,000\text{ M}^{-1}\cdot\text{cm}^{-1}$ ). Gemini-561 a été conçu pour promouvoir l'auto-extinction induite par la dimérisation de deux SRB. À cette fin, la lysine a été choisie comme connecteur afin de fournir une faible distance entre les rhodamines et de garantir ainsi un empilement (stacking) efficace lors de la dimérisation. Les dérivés de lysine (1) et de SRB (2) ont été déprotégés et couplés pour obtenir Gemini-561-alkyne (Fig.1 a). Ce dernier a été cliqué sur la biotine-PEG-N<sub>3</sub> pour donner Gemini-561. La biotine a été introduite pour fonctionnaliser les billes recouvertes de streptavidine pour l'étape de SELEX.





**Figure 1.** Conception, synthèse et fluorogénicité de Gemini-561. (a) Synthèse de Gemini-561. (b) Concept de réponse fluorogénique de Gemini-561 à l'environnement (solvant organique) et à l'aptamère. (c) Spectres d'absorption et d'excitation de Gemini-561 (200 nM) dans l'eau et le méthanol et de SRB (200 nM) dans l'eau. (d) Spectres de fluorescence de Gemini-561 (200 nM) dans l'eau et méthanol et de SRB (200 nM) dans l'eau. (e-g) Propriétés spectrales et biochimiques de Gemini-561 seul ou en complexe avec des aptamères SRB-2 ou o-Coral. Les mesures ont été effectuées dans un tampon de sélection (tampon phosphate pH7.5, 100 mM KCl, 1mM MgCl<sub>2</sub> and 0.05% Tween-20).

### Propriétés spectrales

Dans l'eau, Gemini-561 présente une faible intensité de fluorescence avec une valeur de rendement quantique de 0,01. De plus, une bande décalée dans le bleu (530 nm) est apparue dans le spectre d'absorption, ce qui indique la formation d'un agrégat dimérique de type H (Fig. 1b), conformément au rapport précédent sur les dimères de squaraine.<sup>35</sup> De plus, le spectre d'excitation a montré que cette bande ne

correspondait pas à une espèce émissive (Fig. 1c), confirmant ainsi un phénomène de désactivation induit par dimérisation. Cependant, lors de la solubilisation dans le méthanol, le dimère s'ouvre et Gemini-561 présente des spectres d'absorption et d'émission similaires à ceux du SRB libre, ainsi qu'une augmentation impressionnante de la valeur de rendement quantique (0,31, fig. 2e). De plus, nous avons constaté que Gemini-561 n'était pas impliqué dans des interactions non spécifiques avec des protéines et des lipoprotéines pouvant provoquer une activation non souhaitée du dimère. Globalement, ces expériences démontrent que Gemini-561 constitue une molécule fluorogène efficace compatible avec les milieux biologiques, ce qui en fait un candidat prometteur pour la sélection d'aptamère d'ARN.

#### *Isolement d'aptamères fluorogènes Gemini-561 par évolution in vitro*

En utilisant SELEX et la compartimentation in vitro assistée par microfluidique ( $\mu$ IVC)<sup>75</sup>, nos collaborateurs (Farah BOUHEDDA and Michael RYCKELYNCK) ont identifié un nouvel aptamère d'ARN fluorogène, nommé o-Coral.

#### *Caractérisation in vitro du module Gemini-561 / o-Coral*

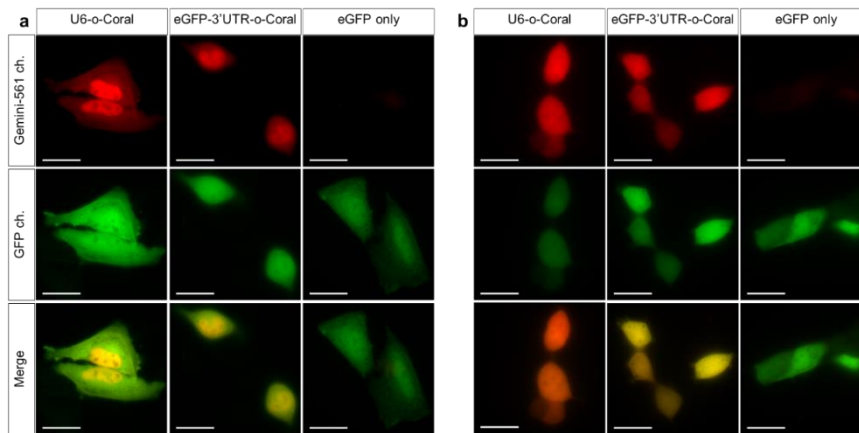
Avec l'aptamère sélectionné o-Coral, nous avons d'abord effectué des études spectroscopiques. Lorsque Gemini-561 a été mis en présence d'o-Coral, la bande d'absorption décalée dans le bleu dans les spectres d'absorption a rapidement diminué pour afficher les spectres de SRB sous forme moléculaire. Dans le même temps, l'intensité de la fluorescence a été multipliée par 13, montrant une dissociation du dimère SRB provoquée par l'interaction de o-coral avec Gemini-561. Ces données montrent que Gemini-561 est confiné dans l'aptamère o-Coral, où les deux fragments de rhodamine sont bien séparés, empêchant ainsi l'auto-extinction et conduisant ainsi à un rendement quantique accru. Comme contrôle, nous avons vérifié que Gemini-561 se liait faiblement à l'aptamère SRB-2 originale, ce qui donnait un effet de fluorescence d'activation de 1,3 fois seulement (Fig. 1f-g). Globalement, cela suggère que la reconnaissance de Gemini-561 par o-Coral est rapide, spécifique et influencée de manière négligeable par d'autres biomolécules.

#### *Imagerie Gemini-561/o-Coral dans des cellules vivantes*

Comme Gemini-561 n'a pas montré de cytotoxicité, nous avons appliqué le module Gemini-561/o-Coral dans des cellules HeLa vivantes en micro-injectant le complexe préformé soit directement dans le noyau, soit dans le cytoplasme de cellules HeLa

vivantes. Dans les deux cas, une fluorescence rouge intense a été facilement observée en présence du module, alors que l'injection de Gemini-561 seul et en mélange avec l'aptamère SRB-2 n'a pas produit de fluorescence significative, confirmant que la fluorescence de Gemini-561 est spécifiquement induite par o-Coral et que ce module est bien adapté aux applications de cellules vivantes. Nous avons également évalué la perméabilité cellulaire de Gemini-561 en incubant des cellules HeLa avec du *Gemini-561* à 200 nM. Après lavage, l'aptamère o-Coral ou SRB-2 a été injecté dans les cellules. La fluorescence de Gemini-561 dans le cytosol n'a été observée qu'en présence de o-Coral, suggérant à la fois la perméabilité cellulaire du fluorogène et sa capacité à détecter le corail o à l'intérieur de cellules vivantes.

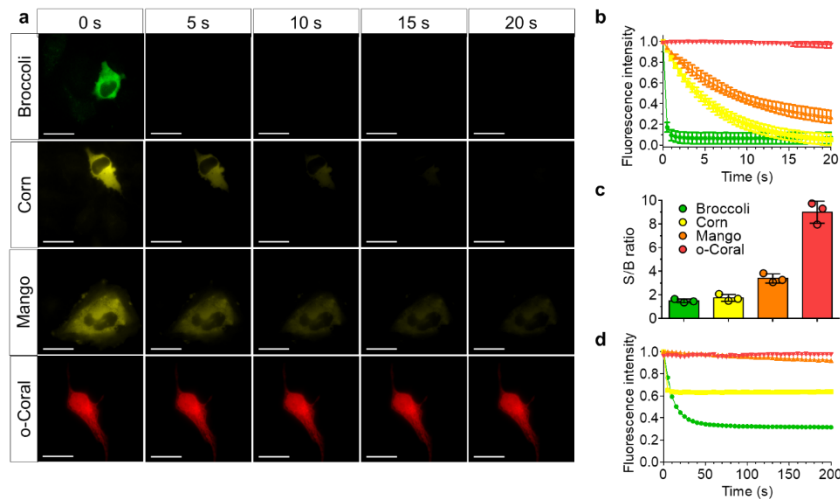
Nous avons ensuite évalué les performances de notre module fluorogène avec des aptamères synthétisés *in situ* par la machinerie cellulaire. À cette fin, Farah BOUHEDDA et Michael RYCKELYNCK ont conçu un plasmide codant pour U6, eGFP marqués avec o-Coral. Profitant de la perméabilité des cellules envers Gemini-561, nous avons effectué une expérience en exprimant le gène o-Coral dans des cellules HeLa et HEK293T vivantes, puis nous avons brièvement incubé les cellules avec Gemini-561. Comme prévu, les cellules vertes transfectées par la GFP présentaient également une fluorescence rouge intense (Gemini-561) dans le noyau et dans le cytoplasme (Fig. 2). Dans l'ensemble, ces données montrent que, en raison de la brillance et de l'affinité élevées du complexe Gemini-561 / o-Coral, une seule copie d'aptamère o-Coral est suffisante pour la détection l'ARN d'intérêt dans des cellules de mammifère vivantes en utilisant Gemini-561. Il convient de souligner que, à notre connaissance, c'est la première fois qu'un ARNm marqué d'une seule copie d'un aptamère fluorogène peut être visualisé sans nécessiter un traitement approfondi de l'image.



**Figure 2.** Imagerie sur cellules vivantes de o-Coral exprimée à partir des promoteur pol. II et pol. III. Imagerie de cellules vivantes HeLa (a) et HEK293T (b) exprimant U6-o-Coral, l'ARNm d'egfp eGFP-3'UTR-o-Coral ou eGFP uniquement. Les cellules ont été incubées avec Gemini-561 (200 nM) pendant 5 min avant l'imagerie. Les images ont été acquises en utilisant un temps d'exposition de 500 ms. Gemini-561 en rouge (ex: 550 nm, em: 595 ± 40 nm), Hoechst en bleu (ex: 395 nm, em: 510 ± 42 nm) et eGFP en vert (ex: 470 nm, em: 531 ± 40 nm). La barre d'échelle est de 30 µm.

#### *Comparaison de Gemini-561 / o-Coral à d'autres modules d'aptamère*

Afin de démontrer l'intérêt de ce nouveau système, nous avons systématiquement comparé le système G561 / o-Coral aux aptamères fluorogènes disponibles, à savoir Broccoli<sup>63</sup>, Mango III<sup>75</sup> et Corn<sup>67</sup>. Avant les expériences sur cellules, nous avons évalué la photostabilité des complexes en cuvette avec des concentrations fixes de fluorogène (0,2 µM) et d'aptamère (1 µM) et une intensité de puissance laser fixée (Fig.3d). L'intensité d'émission de G561/o-Coral n'a montré aucun changement, alors que les trois autres systèmes d'aptamères ont montré une perte d'intensité, la plus significative pour Broccoli et la moins prononcée pour Mango (Fig. 3). Ensuite, les complexes fluorogène/aptamère ont été préformés puis micro-injectés dans des cellules vivantes. Juste après l'injection, les images ont été prises au fil du temps. Alors que Broccoli, Corn et Mango fournissent des images avec des valeurs Signal/Bruit <4, notre système fournit des images de meilleure qualité avec une valeur S/B supérieure à 8. De plus, G561/o-Coral peut être visualisé pendant plus de 20 s, affichant une photostabilité accrue par rapport aux trois autres systèmes. La photostabilité en imagerie suit cet ordre : o-Coral >> Mango > Mais >> Brocoli, conformément aux données obtenues en cuvette.



**Figure 3.** Analyse comparative de la photostabilité par microscopie à fluorescence et par spectroscopie. (a) Mesure de photostabilité dans des cellules HeLa vivantes. Les aptamères ont été préincubés avec les fluorogènes respectifs pendant 10 min dans un tampon de sélection pour former un complexe. Les complexes ont été microinjectés en utilisant 5  $\mu\text{M}$  de fluorogène et une concentration d'aptamère de 20  $\mu\text{M}$ . Des images consécutives ont été acquises, chacune utilisant un temps d'exposition de 500 ms. La puissance d'excitation a été ajustée pour que les fluoromodules absorbent une quantité similaire de photons. Broccoli (ex: 470 nm, em:  $475 \pm 50$  nm); Corn (ex: 470 nm, em:  $531 \pm 40$  nm); Mango (ex: 470 nm, em:  $531 \pm 40$  nm); Coral (ex: 550 nm, em:  $595 \pm 40$  nm). La barre d'échelle est de 30  $\mu\text{m}$ . (b) Courbes de décroissance de l'intensité de la fluorescence dans le temps. Les données représentent les valeurs moyennes  $\pm$ D.S. extraites d'images de 3 expériences indépendantes ( $n = 3$ ). (c) Rapport signal sur bruit de la première image acquise à partir d'une image décrivant la brillance du système et la qualité des images obtenues. La valeur de chaque mesure est indiquée par un point coloré. Les barres d'erreur correspondent à deviation standard. (d) Photostabilité de G561/o-Coral par rapport à Broccoli/DFHBI-1T, Corn/DFHO, Mango/TO1-biotine. La concentration a été fixée à 0,2/1  $\mu\text{M}$  (fluorogène/aptamère). Chaque complexe était excité à la même valeur du coefficient d'extinction molaire. L'intensité de la fluorescence a été mesurée à 507 nm pour Broccoli, à 545 nm pour Corn, à 535 nm pour Mango et à 596 nm pour o-Coral.



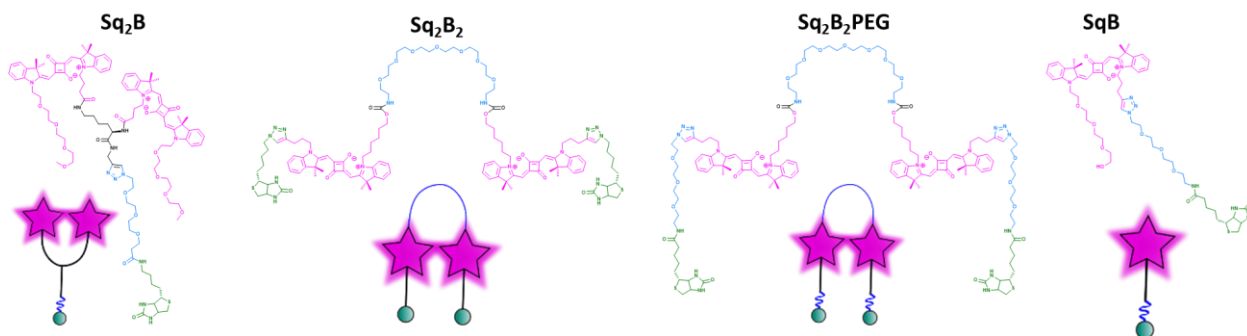
## **2.2. Détection de récepteurs de surface à la biotine grâce à la conception rationnelle de dimères de squaraines fluorogènes**

La biotine est une vitamine essentielle jouant son rôle dans le métabolisme cellulaire des glucides, des acides aminés et des lipides. Contrairement aux bactéries, la machinerie cellulaire des mammifères ne produit pas de biotine; par conséquent, la biotine est complétée de manière exogène. Il existe des preuves que l'expression des récepteurs de la biotine (BR) est corrélée au cancer. Bien que les récepteurs de la biotine constituent une cible attrayante en tant que système d'administration du médicament contre la tumeur, il manque des sondes d'imagerie robustes pour le diagnostic clinique des RSB dans les cellules cancéreuses et l'évaluation de nouveaux traitements ciblés.<sup>131,132</sup> Bien que des sondes biotinylées sensibles à l'environnement et fonctionnant dans la région visible aient été développées pour l'imagerie des BR, la détection des BR à faible concentration nécessite une brillance accrue. De plus, il existe une demande particulière dans le développement de sondes opérant dans les régions dans rouge lointain et le proche infrarouge (NIR) pour une imagerie plus profonde dans les tissus et potentiellement *in vivo*.

Dans ce projet, nous avons rationnellement conçu et synthétisé quatre variantes de squaraines fluorogènes biotinylées. Nous avons étudié la relation structure-propriétés de ces sondes et évalué la capacité et la spécificité du meilleur candidat à détecter les BR cellulaires par imagerie de manière brillante et spécifique.

### *Conception de dimères de squaraine biotinylés*

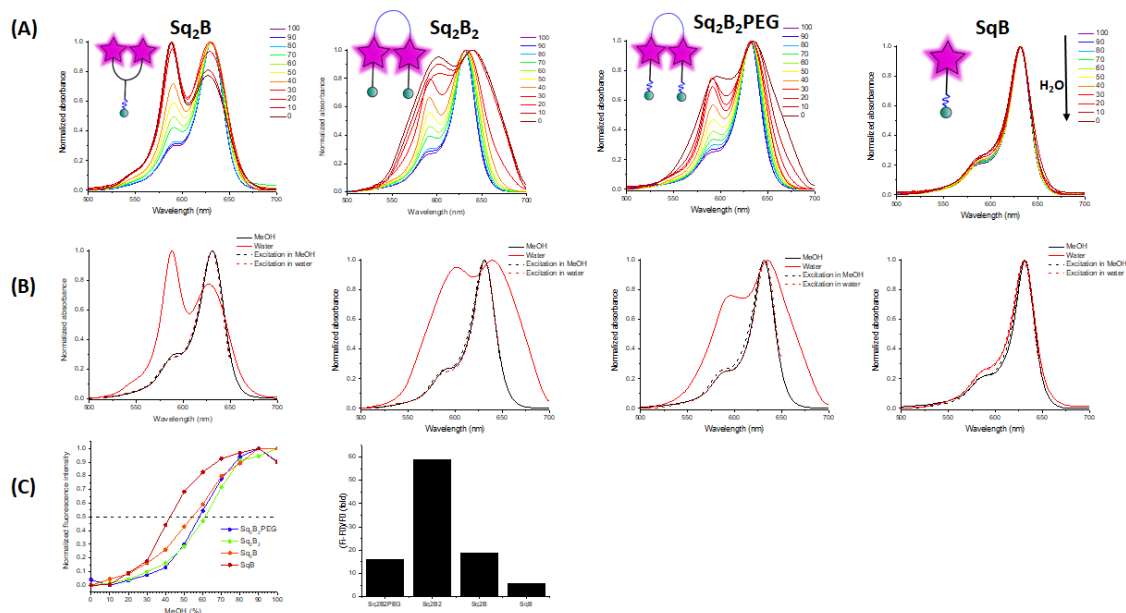
Nous nous sommes posé la question de savoir si la conception de différents types de squaraines dimériques pouvait servir à ajuster leur efficacité fluorogénique. Pour ce faire, nous avons conçu et synthétisé deux dimères avec des liens PEG8 flexibles (Sq<sub>2</sub>B<sub>2</sub>, Sq<sub>2</sub>B<sub>2</sub>PEG). Nous avons également conçu et synthétisé un dimère plus court et plus rigide basé sur le lien L-lysine (Sq<sub>2</sub>B) ainsi qu'un colorant analogue monomérique (SqB) (Fig.4).



**Figure 4.** Structures chimiques des squaraines fluorogéniques biotinylées

### Propriétés fluorogéniques des dimères de squaraine

Pour évaluer le caractère fluorogène des dimères, nous avons étudié leurs propriétés photophysiques en augmentant la proportion de méthanol dans l'eau. Les quatre sondes ont présenté des spectres d'absorption similaires dans le MeOH avec un maximum d'absorbance d'environ 630 nm et une émission à environ 640 nm. Contrairement au monomère SqB, tous les dimères ont montré la formation d'une épaule décalés dans le bleu dans de forte proportion d'eau, probablement en raison de la formation d'un agrégat intramoléculaire de type H (Fig. 5). Cependant, les dimères à base de PEG8 présentaient de larges spectres d'absorption dans l'eau, ce qui peut s'expliquer par une agrégation intermoléculaire supplémentaire (Fig. 5).

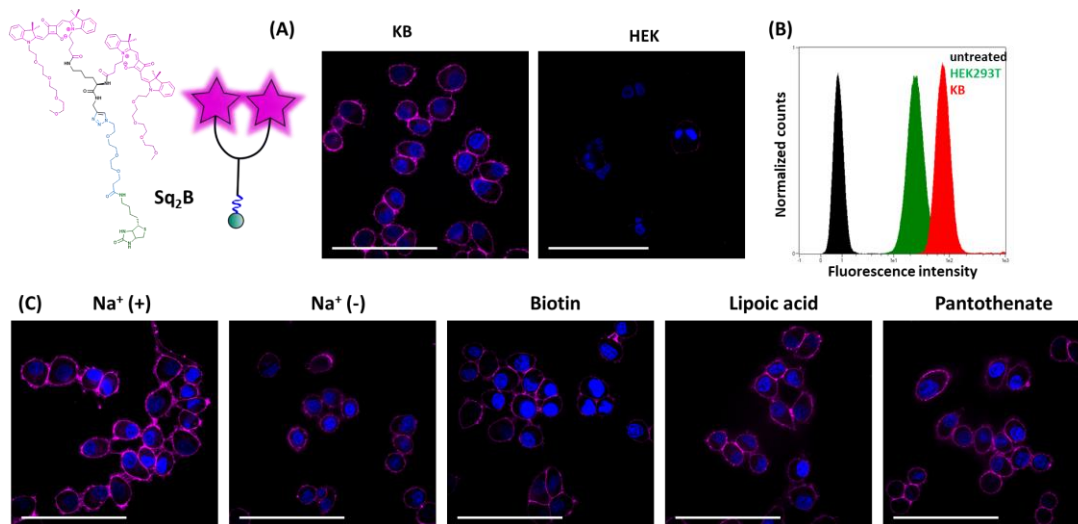


**Figure 5.** Propriétés photophysiques des sondes squaraines biotinylées (**Sq<sub>2</sub>B**, **Sq<sub>2</sub>B<sub>2</sub>**, **Sq<sub>2</sub>B<sub>2</sub>PEG** and **SqB**).

### *Imagerie des récepteurs de la biotine avec Sq<sub>2</sub>B*

Nous avons choisi la sonde Sq<sub>2</sub>B pour une validation ultérieure car elle présentait de faible interaction non spécifiques en présence de biomolécules et conservait sa forme monomoléculaire soluble dans une plage de concentration de 20 à 200 nM. Les cellules cancéreuses, telles que les KB (dérivés des cellules HeLa) ont tendance à consommer beaucoup de biotine comme complément vital par rapport à d'autres cellules, suggérant l'expression de BR.<sup>131</sup> Nous avons donc utilisé les cellules KB comme modèle cellulaire pour évaluer les performances de la sonde Sq<sub>2</sub>B et les cellules HEK293T (ligné cellulaire BR-négative) comme contrôle négatif.<sup>131</sup> Le Sq<sub>2</sub>B biotinylé a montré une faible cytotoxicité dans la gamme de concentrations utilisée pour l'imagerie. Lors d'une brève incubation sans étape de lavage, Sq<sub>2</sub>B produit facilement une réponse de fluorescence avec un marquage distinctif de la surface cellulaire des cellules KB, mais pas dans des cellules HEK293T dépourvues de BR (Fig.6A). Nous avons également évalué la spécificité de Sq<sub>2</sub>B vis-à-vis des BR par cytométrie en flux. La différence d'intensité de fluorescence a été démontrée entre les cellules BR-positives (KB) traitées avec Sq<sub>2</sub>B et les cellules BR-négatives (HEK293T) dans l'analyse cellule par cellule (Fig. 6B).

Ensuite, nous avons évalué la spécificité de Sq<sub>2</sub>B par rapport aux BR. Le transporteur multivitamines dépendant du sodium (SMVT) est un transporteur de biotine bien caractérisé dans les cellules KB (dérivés de HeLa).<sup>131</sup> A part la biotine, SMVT transfère également des cofacteurs essentiels comme l'acide lipoïque et le panthothénate.<sup>156</sup> Premièrement, afin d'évaluer la dépendance aux ions sodium, nous avons incubé des cellules KB avec Sq<sub>2</sub>B dans un tampon HBSS dépourvu de Na<sup>+</sup>. Les résultats ont montré la diminution de la fluorescence de Sq<sub>2</sub>B lorsque le milieu déprivé de Na<sup>+</sup> était utilisé, ce qui indique un fort effet du potentiel sodique sur l'absorption cellulaire de Sq<sub>2</sub>B (Fig. 6C). Deuxièmement, les cellules KB ont été prétraitées avec des inhibiteurs (biotine, acide lipoïque ou pantothénate de sodium) pour saturer les transporteurs SMVT et réduire ainsi leur activité. Une brève incubation supplémentaire avec la molécule Sq<sub>2</sub>B a montré la diminution de la fluorescence dans les cellules KB traitées par un compétiteur par rapport aux cellules KB non traitées (Fig. 6C). Ces résultats suggèrent que Sq<sub>2</sub>B est majoritairement pris en charge par les récepteurs SMVT.



**Figure 6.** Imagerie par fluorescence. (A) Imagerie par fluorescence de **Sq<sub>2</sub>B** sur cellules vivantes BR-positive (KB) and BR-négative (HEK293T). (B) Analyse par la cytométrie en flux des cellules KB et HEK293T traitées avec **Sq<sub>2</sub>B**. Cellules KB non traitées ont été utilisées comme contrôle. (C) Imagerie par fluorescence de **Sq<sub>2</sub>B** sur les cellules KB dans le milieu HBSS déprivé de Na<sup>+</sup>, prétraitées avec biotine (100 μM), acide lipoïque (100 μM) ou pantothénate de sodium (100 μM). La barre d'échelle est de 30 μm.

### 3) Conclusion générale

Dans ce travail, nous avons montré que la stratégie de désactivation induite par la dimérisation constitue une plate-forme pour la conception de sondes fluorogènes brillantes fournissant une imagerie spécifique d'événements biologiques.

Dans la première partie, nous avons synthétisé et caractérisé une sonde fluorogène à base de rhodamines appelée Gemini-561. Nous avons illustré la stratégie de dimérisation dans le développement de Gemini-561/o-Coral, un nouveau module fluorogène à base d'ARN destiné à l'imagerie de séquences ARN dans des cellules vivantes. Ce système est basé sur le colorant SRB Gemini-561 dimérique qui active sa fluorescence lors de la reconnaissance de l'aptamère bivalent o-Coral, obtenue par un processus de sélection. Nous avons démontré que l'aptamère o-Coral exprimé génétiquement dans des cellules peut être détecté par le fluorogène Gemini-561 après son entrée spontanée dans des cellules vivantes. Gemini-561 affiche une brillance, une affinité et une photostabilité élevées, ce qui en fait, à notre connaissance, l'un des modules les plus brillants décrits à ce jour dans la littérature<sup>55</sup>. Ce système étoffe la

boîte à outils des systèmes aptamères fluorogènes pour l'imagerie de séquences ARN, ce qui permet de les identifier pour élucider leur rôle dans l'évolution cellulaire. Dans la seconde partie, nous avons conçu et synthétisé quatre sondes fluorogènes pour l'imagerie des récepteurs de la biotine dans des cellules vivantes, en nous concentrant sur la conception moléculaire de dimères de squaraines. Nous avons montré que les squaraines dimérisés biotinylés permettaient de modifier les propriétés fluorogéniques tout en minimisant les interactions non spécifiques avec les biomolécules et de réduire la fluorescence basale dans le contexte cellulaire. Nous avons montré que le candidat le plus prometteur, Sq<sub>2</sub>B, était soumis à une translocation active dans les cellules cancéreuses KB. Nous avons prouvé que Sq<sub>2</sub>B reconnaît majoritairement le transporteur multivitaminique dépendant du sodium (SMVT)<sup>131</sup> dans un ensemble d'expériences. En particulier, le transport de Sq<sub>2</sub>B est dépendant du sodium et est inhibé par des vitamines naturelles telles que la biotine, l'acide lipoïque et l'acide pantothénique. Nous pensons que la sonde spécifique Sq<sub>2</sub>B contribuera à la compréhension du rôle biologique des récepteurs de la biotine et à mettre au point de nouvelles stratégies d'évaluation de nouveaux traitements contre le cancer.

Dans l'ensemble, les présentes études fournissent des informations importantes sur la conception de sondes dimérisées fluorogènes pour l'imagerie de processus biologiques complexes.

## References

- (1) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 3rd ed.; Springer US, 2006.
- (2) Luby, B. M.; Charron, D. M.; MacLaughlin, C. M.; Zheng, G. Activatable Fluorescence: From Small Molecule to Nanoparticle. *Advanced Drug Delivery Reviews* **2017**, *113*, 97–121. <https://doi.org/10.1016/j.addr.2016.08.010>.
- (3) Leblond, F.; Davis, S. C.; Valdés, P. A.; Pogue, B. W. Pre-Clinical Whole-Body Fluorescence Imaging: Review of Instruments, Methods and Applications. *Journal of Photochemistry and Photobiology B: Biology* **2010**, *98* (1), 77–94. <https://doi.org/https://doi.org/10.1016/j.jphotobiol.2009.11.007>.
- (4) Klymchenko, A. S. Solvatochromic and Fluorogenic Dyes as Environment-Sensitive Probes: Design and Biological Applications. *Acc. Chem. Res.* **2017**, *50* (2), 366–375. <https://doi.org/10.1021/acs.accounts.6b00517>.
- (5) Kobayashi, H.; Longmire, M. R.; Ogawa, M.; Choyke, P. L. Rational Chemical Design of the next Generation of Molecular Imaging Probes Based on Physics and Biology: Mixing Modalities, Colors and Signals. *Chem. Soc. Rev.* **2011**, *40* (9), 4626–4648. <https://doi.org/10.1039/C1CS15077D>.



- (6) Bargh, J. D.; Isidro-Llobet, A.; Parker, J. S.; Spring, D. R. Cleavable Linkers in Antibody–Drug Conjugates. *Chem. Soc. Rev.* **2019**, *48* (16), 4361–4374. <https://doi.org/10.1039/C8CS00676H>.
- (7) Demchenko, A. P. Practical Aspects of Wavelength Ratiometry in the Studies of Intermolecular Interactions. *Journal of Molecular Structure* **2014**, *1077*, 51–67. <https://doi.org/https://doi.org/10.1016/j.molstruc.2013.11.045>.
- (8) Kucherak, O. A.; Oncul, S.; Darwich, Z.; Yushchenko, D. A.; Arntz, Y.; Didier, P.; Mély, Y.; Klymchenko, A. S. Switchable Nile Red-Based Probe for Cholesterol and Lipid Order at the Outer Leaflet of Biomembranes. *J. Am. Chem. Soc.* **2010**, *132* (13), 4907–4916. <https://doi.org/10.1021/ja100351w>.
- (9) Niko, Y.; Didier, P.; Mely, Y.; Konishi, G.; Klymchenko, A. S. Bright and Photostable Push-Pull Pyrene Dye Visualizes Lipid Order Variation between Plasma and Intracellular Membranes. *Scientific Reports* **2016**, *6*, 18870. <https://doi.org/10.1038/srep18870>.
- (10) Karpenko, I. A.; Niko, Y.; Yakubovskiy, V. P.; Gerasov, A. O.; Bonnet, D.; Kovtun, Y. P.; Klymchenko, A. S. Push–Pull Dioxaborine as Fluorescent Molecular Rotor: Far-Red Fluorogenic Probe for Ligand–Receptor Interactions. *J. Mater. Chem. C* **2016**, *4* (14), 3002–3009. <https://doi.org/10.1039/C5TC03411F>.
- (11) Shaya, J.; Collot, M.; Bénailly, F.; Mahmoud, N.; Mély, Y.; Michel, B. Y.; Klymchenko, A. S.; Burger, A. Turn-on Fluorene Push–Pull Probes with High Brightness and Photostability for Visualizing Lipid Order in Biomembranes. *ACS Chem. Biol.* **2017**, *12* (12), 3022–3030. <https://doi.org/10.1021/acscchembio.7b00658>.
- (12) Danylchuk, D. I.; Moon, S.; Xu, K.; Klymchenko, A. S. Switchable Solvatochromic Probes for Live-Cell Super-Resolution Imaging of Plasma Membrane Organization. *Angewandte Chemie International Edition* **0** (0). <https://doi.org/10.1002/anie.201907690>.
- (13) Yuan, L.; Lin, W.; Zheng, K.; Zhu, S. FRET-Based Small-Molecule Fluorescent Probes: Rational Design and Bioimaging Applications. *Acc. Chem. Res.* **2013**, *46* (7), 1462–1473. <https://doi.org/10.1021/ar300273v>.
- (14) Mei, J.; Leung, N. L. C.; Kwok, R. T. K.; Lam, J. W. Y.; Tang, B. Z. Aggregation-Induced Emission: Together We Shine, United We Soar! *Chem. Rev.* **2015**, *115* (21), 11718–11940. <https://doi.org/10.1021/acs.chemrev.5b00263>.
- (15) Ding, D.; Li, K.; Liu, B.; Tang, B. Z. Bioprobes Based on AIE Fluorogens. *Acc. Chem. Res.* **2013**, *46* (11), 2441–2453. <https://doi.org/10.1021/ar3003464>.
- (16) Zhai, D.; Xu, W.; Zhang, L.; Chang, Y.-T. The Role of “Disaggregation” in Optical Probe Development. *Chem. Soc. Rev.* **2014**, *43* (8), 2402–2411. <https://doi.org/10.1039/C3CS60368G>.
- (17) Chen, C.; Wang, R.; Guo, L.; Fu, N.; Dong, H.; Yuan, Y. A Squaraine-Based Colorimetric and “Turn on” Fluorescent Sensor for Selective Detection of Hg<sup>2+</sup> in an Aqueous Medium. *Org. Lett.* **2011**, *13* (5), 1162–1165. <https://doi.org/10.1021/ol200024g>.
- (18) Chen, C.; Dong, H.; Chen, Y.; Guo, L.; Wang, Z.; Sun, J.-J.; Fu, N. Dual-Mode Unsymmetrical Squaraine-Based Sensor for Selective Detection of Hg<sup>2+</sup> in Aqueous Media. *Org. Biomol. Chem.* **2011**, *9* (23), 8195–8201. <https://doi.org/10.1039/C1OB06519J>.
- (19) Zhai, D.; Agrawalla, B. K.; Eng, P. S. F.; Lee, S.-C.; Xu, W.; Chang, Y.-T. Development of a Fluorescent Sensor for an Illicit Date Rape Drug – GBL. *Chem. Commun.* **2013**, *49* (55), 6170–6172. <https://doi.org/10.1039/C3CC43153C>.
- (20) Mizusawa, K.; Ishida, Y.; Takaoka, Y.; Miyagawa, M.; Tsukiji, S.; Hamachi, I. Disassembly-Driven Turn-On Fluorescent Nanoprobes for Selective Protein Detection. *J. Am. Chem. Soc.* **2010**, *132* (21), 7291–7293. <https://doi.org/10.1021/ja101879g>.
- (21) Mizusawa, K.; Takaoka, Y.; Hamachi, I. Specific Cell Surface Protein Imaging by Extended Self-Assembling Fluorescent Turn-on Nanoprobes. *J. Am. Chem. Soc.* **2012**, *134* (32), 13386–13395. <https://doi.org/10.1021/ja304239g>.

- (22) Niko, Y.; Arntz, Y.; Mely, Y.; Konishi, G.; Klymchenko, A. S. Disassembly-Driven Fluorescence Turn-on of Polymerized Micelles by Reductive Stimuli in Living Cells. *Chemistry – A European Journal* **2014**, *20* (50), 16473–16477. <https://doi.org/10.1002/chem.201405040>.
- (23) Collot, M.; Ashokkumar, P.; Anton, H.; Boutant, E.; Faklaris, O.; Galli, T.; Mély, Y.; Danglot, L.; Klymchenko, A. S. MemBright: A Family of Fluorescent Membrane Probes for Advanced Cellular Imaging and Neuroscience. *Cell Chemical Biology* **2019**, *26* (4), 600–614.e7. <https://doi.org/10.1016/j.chembiol.2019.01.009>.
- (24) Karpenko, I. A.; Klymchenko, A. S.; Gioria, S.; Kreder, R.; Shulov, I.; Villa, P.; Mély, Y.; Hibert, M.; Bonnet, D. Squaraine as a Bright, Stable and Environment-Sensitive Far-Red Label for Receptor-Specific Cellular Imaging. *Chem. Commun.* **2015**, *51* (14), 2960–2963. <https://doi.org/10.1039/C4CC09113B>.
- (25) Collot, M.; Kreder, R.; Tatarets, A. L.; Patsenker, L. D.; Mely, Y.; Klymchenko, A. S. Bright Fluorogenic Squaraines with Tuned Cell Entry for Selective Imaging of Plasma Membrane vs. Endoplasmic Reticulum. *Chem. Commun.* **2015**, *51* (96), 17136–17139. <https://doi.org/10.1039/C5CC06094J>.
- (26) Yang, Q.; Xiang, J.; Yang, S.; Zhou, Q.; Li, Q.; Tang, Y.; Xu, G. Verification of Specific G-Quadruplex Structure by Using a Novel Cyanine Dye Supramolecular Assembly: I. Recognizing Mixed G-Quadruplex in Human Telomeres. *Chem. Commun.* **2009**, No. 9, 1103–1105. <https://doi.org/10.1039/B820101C>.
- (27) Collot, M.; Fam, T. K.; Ashokkumar, P.; Faklaris, O.; Galli, T.; Danglot, L.; Klymchenko, A. S. Ultrabright and Fluorogenic Probes for Multicolor Imaging and Tracking of Lipid Droplets in Cells and Tissues. *J. Am. Chem. Soc.* **2018**, *140* (16), 5401–5411. <https://doi.org/10.1021/jacs.7b12817>.
- (28) Okamoto, A. ECHO Probes: A Concept of Fluorescence Control for Practical Nucleic Acid Sensing. *Chem. Soc. Rev.* **2011**, *40* (12), 5815–5828. <https://doi.org/10.1039/C1CS15025A>.
- (29) Wei, A.-P.; Blumenthal, D. K.; Herron, J. N. Antibody-Mediated Fluorescence Enhancement Based on Shifting the Intramolecular Dimer–Monomer Equilibrium of Fluorescent Dyes. *Anal. Chem.* **1994**, *66* (9), 1500–1506. <https://doi.org/10.1021/ac00081a023>.
- (30) Packard, B. Z.; Topygin, D. D.; Komoriya, A.; Brand, L. Profluorescent Protease Substrates: Intramolecular Dimers Described by the Exciton Model. *Proc Natl Acad Sci USA* **1996**, *93* (21), 11640. <https://doi.org/10.1073/pnas.93.21.11640>.
- (31) Packard, B. Z.; Komoriya, A.; Topygin, D. D.; Brand, L. Structural Characteristics of Fluorophores That Form Intramolecular H-Type Dimers in a Protease Substrate. *J. Phys. Chem. B* **1997**, *101* (25), 5070–5074. <https://doi.org/10.1021/jp9702210>.
- (32) Packard, B. Z.; Topygin, D. D.; Komoriya, A.; Brand, L. Characterization of Fluorescence Quenching in Bifluorophoric Protease Substrates. *Biophysical Chemistry* **1997**, *67* (1), 167–176. [https://doi.org/10.1016/S0301-4622\(97\)00036-7](https://doi.org/10.1016/S0301-4622(97)00036-7).
- (33) Packard, B. Z.; Komoriya, A.; Nanda, V.; Brand, L. Intramolecular Excitonic Dimers in Protease Substrates: Modification of the Backbone Moiety To Probe the H-Dimer Structure. *J. Phys. Chem. B* **1998**, *102* (10), 1820–1827. <https://doi.org/10.1021/jp973419t>.
- (34) Knemeyer, J.-P.; Marmé, N.; Häfner, B.; Habl, G.; Schäfer, G.; Müller, M.; Nolte, O.; Sauer, M.; Wolfrum, J. Self-Quenching DNA Probes Based on Dye Dimerization for Identification of Mycobacteria. *International Journal of Environmental Analytical Chemistry* **2005**, *85* (9–11), 625–637. <https://doi.org/10.1080/03067310500146094>.
- (35) Karpenko, I. A.; Collot, M.; Richert, L.; Valencia, C.; Villa, P.; Mély, Y.; Hibert, M.; Bonnet, D.; Klymchenko, A. S. Fluorogenic Squaraine Dimers with Polarity-Sensitive Folding As Bright Far-Red Probes for Background-Free Bioimaging. *Journal of the American Chemical Society* **2015**, *137* (1), 405–412. <https://doi.org/10.1021/ja5111267>.
- (36) Yao, D.; Lin, Z.; Wu, J. Near-Infrared Fluorogenic Probes with Polarity-Sensitive Emission for in Vivo Imaging of an Ovarian Cancer Biomarker. *ACS Appl. Mater. Interfaces* **2016**, *8* (9), 5847–5856. <https://doi.org/10.1021/acsami.5b11826>.

- (37) Marianayagam, N. J.; Sunde, M.; Matthews, J. M. The Power of Two: Protein Dimerization in Biology. *Trends in Biochemical Sciences* **2004**, *29* (11), 618–625. <https://doi.org/https://doi.org/10.1016/j.tibs.2004.09.006>.
- (38) RNA. *Wikipedia*; 2019.
- (39) Beckman, W.; Vuist, I. M.; Kempe, H.; Verschure, P. J. Cell-to-Cell Transcription Variability as Measured by Single-Molecule RNA FISH to Detect Epigenetic State Switching. In *Epigenome Editing: Methods and Protocols*; Jeltsch, A., Rots, M. G., Eds.; Springer New York: New York, NY, 2018; pp 385–393. [https://doi.org/10.1007/978-1-4939-7774-1\\_21](https://doi.org/10.1007/978-1-4939-7774-1_21).
- (40) Irina V. Novikova. New Ideas for in Vivo Detection of RNA. In *Biosensors*; Kirill A. Afonin, Ed.; Neocles B. Leontis ED1 - Pier Andrea Serra, Series Ed.; IntechOpen: Rijeka, 2010; p Ch. 8. <https://doi.org/10.5772/7207>.
- (41) Tutucci, E.; Livingston, N. M.; Singer, R. H.; Wu, B. Imaging mRNA In Vivo, from Birth to Death. *Annual Review of Biophysics* **2018**, *47* (1), null. <https://doi.org/10.1146/annurev-biophys-070317-033037>.
- (42) Xia, Y.; Zhang, R.; Wang, Z.; Tian, J.; Chen, X. Recent Advances in High-Performance Fluorescent and Bioluminescent RNA Imaging Probes. *Chem. Soc. Rev.* **2017**. <https://doi.org/10.1039/C6CS00675B>.
- (43) Rudkin, G. T.; Stollar, B. D. High Resolution Detection of DNA–RNA Hybrids in Situ by Indirect Immunofluorescence. *Nature* **1977**, *265* (5593), 472–473. <https://doi.org/10.1038/265472a0>.
- (44) Holstein, J. M.; Rentmeister, A. Current Covalent Modification Methods for Detecting RNA in Fixed and Living Cells. *Methods* **2016**, *98*, 18–25. <https://doi.org/10.1016/j.ymeth.2015.11.016>.
- (45) Anhäuser, L.; Rentmeister, A. Enzyme-Mediated Tagging of RNA. *Current Opinion in Biotechnology* **2017**, *48*, 69–76. <https://doi.org/10.1016/j.copbio.2017.03.013>.
- (46) Dean, K. M.; Palmer, A. E. Advances in Fluorescence Labeling Strategies for Dynamic Cellular Imaging. *Nat Chem Biol* **2014**, *10* (7), 512–523. <https://doi.org/10.1038/nchembio.1556>.
- (47) Tyagi, S. Imaging Intracellular RNA Distribution and Dynamics in Living Cells. *Nat Meth* **2009**, *6* (5), 331–338. <https://doi.org/10.1038/nmeth.1321>.
- (48) Ozawa, T.; Natori, Y.; Sato, M.; Umezawa, Y. Imaging Dynamics of Endogenous Mitochondrial RNA in Single Living Cells. *Nature Methods* **2007**, *4* (5), 413–419. <https://doi.org/10.1038/nmeth1030>.
- (49) Adli, M. The CRISPR Tool Kit for Genome Editing and Beyond. *Nature Communications* **2018**, *9* (1), 1911. <https://doi.org/10.1038/s41467-018-04252-2>.
- (50) Chen, B.; Gilbert, L. A.; Cimini, B. A.; Schnitzbauer, J.; Zhang, W.; Li, G.-W.; Park, J.; Blackburn, E. H.; Weissman, J. S.; Qi, L. S.; et al. Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. *Cell* **2013**, *155* (7), 1479–1491. <https://doi.org/10.1016/j.cell.2013.12.001>.
- (51) Nelles, D. A.; Fang, M. Y.; O’Connell, M. R.; Xu, J. L.; Markmiller, S. J.; Doudna, J. A.; Yeo, G. W. Programmable RNA Tracking in Live Cells with CRISPR/Cas9. *Cell* **2016**, *165* (2), 488–496. <https://doi.org/10.1016/j.cell.2016.02.054>.
- (52) Rauch, S.; Dickinson, B. C. Programmable RNA Binding Proteins for Imaging and Therapeutics. *Biochemistry* **2018**, *57* (4), 363–364. <https://doi.org/10.1021/acs.biochem.7b01101>.
- (53) Qin, P.; Parlak, M.; Kuscu, C.; Bandaria, J.; Mir, M.; Szlachta, K.; Singh, R.; Darzacq, X.; Yildiz, A.; Adli, M. Live Cell Imaging of Low- and Non-Repetitive Chromosome Loci Using CRISPR-Cas9. *Nature Communications* **2017**, *8*, 14725.
- (54) Ellington, A. D.; Szostak, J. W. In Vitro Selection of RNA Molecules That Bind Specific Ligands. *Nature* **1990**, *346* (6287), 818–822. <https://doi.org/10.1038/346818a0>.

- (55) Bouhedda, F.; Autour, A.; Ryckelynck, M. Light-Up RNA Aptamers and Their Cognate Fluorogens: From Their Development to Their Applications. *International Journal of Molecular Sciences* **2017**, *19* (1), 44. <https://doi.org/10.3390/ijms19010044>.
- (56) Gotrik, M. R.; Feagin, T. A.; Csordas, A. T.; Nakamoto, M. A.; Soh, H. T. Advancements in Aptamer Discovery Technologies. *Acc. Chem. Res.* **2016**, *49* (9), 1903–1910. <https://doi.org/10.1021/acs.accounts.6b00283>.
- (57) Taylor, A. I.; Holliger, P. Directed Evolution of Artificial Enzymes (XNAzymes) from Diverse Repertoires of Synthetic Genetic Polymers. *Nature Protocols* **2015**, *10* (10), 1625–1642. <https://doi.org/10.1038/nprot.2015.104>.
- (58) Mayer, G.; Ahmed, M.-S. L.; Dolf, A.; Endl, E.; Knolle, P. A.; Famulok, M. Fluorescence-Activated Cell Sorting for Aptamer SELEX with Cell Mixtures. *Nature Protocols* **2010**, *5* (12), 1993–2004. <https://doi.org/10.1038/nprot.2010.163>.
- (59) Sefah, K.; Shangguan, D.; Xiong, X.; O'Donoghue, M. B.; Tan, W. Development of DNA Aptamers Using Cell-SELEX. *Nature Protocols* **2010**, *5* (6), 1169–1185. <https://doi.org/10.1038/nprot.2010.66>.
- (60) Babendure, J. R.; Adams, S. R.; Tsieng, R. Y. Aptamers Switch on Fluorescence of Triphenylmethane Dyes. *J. Am. Chem. Soc.* **2003**, *125* (48), 14716–14717. <https://doi.org/10.1021/ja037994o>.
- (61) Paige, J. S.; Wu, K. Y.; Jaffrey, S. R. RNA Mimics of Green Fluorescent Protein. *Science* **2011**, *333* (6042), 642–646. <https://doi.org/10.1126/science.1207339>.
- (62) Sunbul Murat; Jäschke Andres. Contact-Mediated Quenching for RNA Imaging in Bacteria with a Fluorophore-Binding Aptamer. *Angewandte Chemie International Edition* **2013**, *52* (50), 13401–13404. <https://doi.org/10.1002/anie.201306622>.
- (63) Filonov, G. S.; Moon, J. D.; Svendsen, 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* (46), 16299–16308. <https://doi.org/10.1021/ja508478x>.
- (64) Autour, A.; Westhof, E.; Ryckelynck, M. iSpinach: A Fluorogenic RNA Aptamer Optimized for in Vitro Applications. *Nucleic Acids Res* **2016**, *44* (6), 2491–2500. <https://doi.org/10.1093/nar/gkw083>.
- (65) Song, W.; Strack, R. L.; Svendsen, N.; Jaffrey, S. R. Plug-and-Play Fluorophores Extend the Spectral Properties of Spinach. *J. Am. Chem. Soc.* **2014**, *136* (4), 1198–1201. <https://doi.org/10.1021/ja410819x>.
- (66) Strack, R. L.; Disney, M. D.; Jaffrey, S. R. A Superfolding Spinach2 Reveals the Dynamic Nature of Trinucleotide Repeat-Containing RNA. *Nature Methods* **2013**, *10* (12), 1219–1224. <https://doi.org/10.1038/nmeth.2701>.
- (67) Song, W.; Filonov, G. S.; Kim, H.; Hirsch, M.; Li, X.; Moon, J. D.; Jaffrey, S. R. Imaging RNA Polymerase III Transcription Using a Photostable RNA-Fluorophore Complex. *Nature Chemical Biology* **2017**, *13* (11), 1187. <https://doi.org/10.1038/nchembio.2477>.
- (68) Steinmetzger, C.; Palanisamy, N.; Gore, K. R.; Höbartner, C. A Multicolor Large Stokes Shift Fluorogen-Activating RNA Aptamer with Cationic Chromophores. *Chemistry – A European Journal* **2019**, *25* (8), 1931–1935. <https://doi.org/10.1002/chem.201805882>.
- (69) Litke, J. L.; Jaffrey, S. R. Highly Efficient Expression of Circular RNA Aptamers in Cells Using Autocatalytic Transcripts. *Nature Biotechnology* **2019**, *37* (6), 667–675. <https://doi.org/10.1038/s41587-019-0090-6>.
- (70) Wu, J.; Zaccara, S.; Khuperkar, D.; Kim, H.; Tanenbaum, M. E.; Jaffrey, S. R. Live Imaging of mRNA Using RNA-Stabilized Fluorogenic Proteins. *Nature Methods* **2019**, *16* (9), 862–865. <https://doi.org/10.1038/s41592-019-0531-7>.
- (71) 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* (8), 1561–1564. <https://doi.org/10.1021/ol702920e>.

- (72) Tan, X.; Constantin, T. P.; Sloane, K. L.; Waggoner, A. S.; Bruchez, M. P.; Armitage, B. A. Fluoromodules Consisting of a Promiscuous RNA Aptamer and Red or Blue Fluorogenic Cyanine Dyes: Selection, Characterization, and Bioimaging. *J. Am. Chem. Soc.* **2017**, *139* (26), 9001–9009. <https://doi.org/10.1021/jacs.7b04211>.
- (73) Dolgosheina, E. V.; Jeng, S. C. Y.; Panchapakesan, S. S. S.; Cojocar, R.; Chen, P. S. K.; Wilson, P. D.; Hawkins, N.; Wiggins, P. A.; Unrau, P. J. RNA Mango Aptamer-Fluorophore: A Bright, High-Affinity Complex for RNA Labeling and Tracking. *ACS Chem. Biol.* **2014**, *9* (10), 2412–2420. <https://doi.org/10.1021/cb500499x>.
- (74) Nygren, J.; Svanvik, N.; Kubista, M. The Interactions between the Fluorescent Dye Thiazole Orange and DNA. *Biopolymers* **1998**, *46* (1), 39–51. [https://doi.org/10.1002/\(SICI\)1097-0282\(199807\)46:1<39::AID-BIP4>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1097-0282(199807)46:1<39::AID-BIP4>3.0.CO;2-Z).
- (75) Autour, A.; Jeng, S. C. Y.; Cawte, A. D.; Abdolhazadeh, A.; Galli, A.; Panchapakesan, S. S. S.; Rueda, D.; Ryckelynck, M.; Unrau, P. J. Fluorogenic RNA Mango Aptamers for Imaging Small Non-Coding RNAs in Mammalian Cells. *Nature Communications* **2018**, *9* (1), 656. <https://doi.org/10.1038/s41467-018-02993-8>.
- (76) Trachman, R. J.; Truong, L.; Ferré-D'Amaré, A. R. Structural Principles of Fluorescent RNA Aptamers. *Trends in Pharmacological Sciences* **2017**, *38* (10), 928–939. <https://doi.org/10.1016/j.tips.2017.06.007>.
- (77) Guo, J. U.; Bartel, D. P. RNA G-Quadruplexes Are Globally Unfolded in Eukaryotic Cells and Depleted in Bacteria. *Science* **2016**, *353* (6306), aaf5371. <https://doi.org/10.1126/science.aaf5371>.
- (78) Grate, D.; Wilson, C. Laser-Mediated, Site-Specific Inactivation of RNA Transcripts. *Proc Natl Acad Sci USA* **1999**, *96* (11), 6131. <https://doi.org/10.1073/pnas.96.11.6131>.
- (79) Yerramilli, V. S.; Kim, K. H. Labeling RNAs in Live Cells Using Malachite Green Aptamer Scaffolds as Fluorescent Probes. *ACS Synth. Biol.* **2018**, *7* (3), 758–766. <https://doi.org/10.1021/acssynbio.7b00237>.
- (80) Holeman, L. A.; Robinson, S. L.; Szostak, J. W.; Wilson, C. Isolation and Characterization of Fluorophore-Binding RNA Aptamers. *Folding and Design* **1998**, *3* (6), 423–431. [https://doi.org/10.1016/S1359-0278\(98\)00059-5](https://doi.org/10.1016/S1359-0278(98)00059-5).
- (81) Lee, J.; Lee, K. H.; Jeon, J.; Dragulescu-Andrasi, A.; Xiao, F.; Rao, J. Combining SELEX Screening and Rational Design to Develop Light-Up Fluorophore–RNA Aptamer Pairs for RNA Tagging. *ACS Chem. Biol.* **2010**, *5* (11), 1065–1074. <https://doi.org/10.1021/cb1001894>.
- (82) Sunbul, M.; Jäschke, A. SRB-2: A Promiscuous Rainbow Aptamer for Live-Cell RNA Imaging. *Nucleic Acids Research* **2018**, *46* (18), e110–e110. <https://doi.org/10.1093/nar/gky543>.
- (83) Wirth, R.; Gao, P.; Nienhaus, G. U.; Sunbul, M.; Jäschke, A. SiRA: A Silicon Rhodamine-Binding Aptamer for Live-Cell Super-Resolution RNA Imaging. *J. Am. Chem. Soc.* **2019**, *141* (18), 7562–7571. <https://doi.org/10.1021/jacs.9b02697>.
- (84) Daly, B.; Ling, J.; de Silva, A. P. Current Developments in Fluorescent PET (Photoinduced Electron Transfer) Sensors and Switches. *Chem. Soc. Rev.* **2015**, *44* (13), 4203–4211. <https://doi.org/10.1039/C4CS00334A>.
- (85) Sparano, B. A.; Koide, K. Fluorescent Sensors for Specific RNA: A General Paradigm Using Chemistry and Combinatorial Biology. *J. Am. Chem. Soc.* **2007**, *129* (15), 4785–4794. <https://doi.org/10.1021/ja070111z>.
- (86) 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* (43), 14954–14955. <https://doi.org/10.1021/ja0530319>.
- (87) Murata, A.; Sato, S.; Kawazoe, Y.; Uesugi, M. Small-Molecule Fluorescent Probes for Specific RNA Targets. *Chem. Commun.* **2011**, *47* (16), 4712–4714. <https://doi.org/10.1039/C1CC10393H>.
- (88) Sato Shin-ichi; Watanabe Mizuki; Katsuda Yousuke; Murata Asako; Wang Dan Ohtan; Uesugi Motonari. Live-Cell Imaging of Endogenous MRNAs with a Small Molecule. *Angewandte*



- Chemie International Edition* **2014**, *54* (6), 1855–1858.  
<https://doi.org/10.1002/anie.201410339>.
- (89) Shin, I.; Ray, J.; Gupta, V.; Ilgu, M.; Beasley, J.; Bendickson, L.; Mehanovic, S.; Kraus, G. A.; Nilsen-Hamilton, M. Live-Cell Imaging of Pol II Promoter Activity to Monitor Gene Expression with RNA IMAGETag Reporters. *Nucleic Acids Res* **2014**, *42* (11), e90–e90.  
<https://doi.org/10.1093/nar/gku297>.
- (90) Braselmann, E.; Wierzba, A. J.; Polaski, J. T.; Chromiński, M.; Holmes, Z. E.; Hung, S.-T.; Batan, D.; Wheeler, J. R.; Parker, R.; Jimenez, R.; et al. A Multicolor Riboswitch-Based Platform for Imaging of RNA in Live Mammalian Cells. *Nature Chemical Biology* **2018**, *14* (10), 964–971.  
<https://doi.org/10.1038/s41589-018-0103-7>.
- (91) Chen, X.; Zhang, D.; Su, N.; Bao, B.; Xie, X.; Zuo, F.; Yang, L.; Wang, H.; Jiang, L.; Lin, Q.; et al. Visualizing RNA Dynamics in Live Cells with Bright and Stable Fluorescent RNAs. *Nature Biotechnology* **2019**. <https://doi.org/10.1038/s41587-019-0249-1>.
- (92) Bouhedda, F.; Fam, K. T.; Collot, M.; Autour, A.; Marzi, S.; Klymchenko, A.; Ryckelynck, M. A Dimerization-Based Fluorogenic Dye-Aptamer Module for RNA Imaging in Live Cells. *Nature Chemical Biology* **2020**, *16* (1), 69–76. <https://doi.org/10.1038/s41589-019-0381-8>.
- (93) Lukinavičius, G.; Umezawa, K.; Olivier, N.; Honigmann, A.; Yang, G.; Plass, T.; Mueller, V.; Reymond, L.; Corrêa Jr, I. R.; Luo, Z.-G.; et al. A Near-Infrared Fluorophore for Live-Cell Super-Resolution Microscopy of Cellular Proteins. *Nat Chem* **2013**, *5* (2), 132–139.  
<https://doi.org/10.1038/nchem.1546>.
- (94) Xue, L.; Karpenko, I. A.; Hiblot, J.; Johnsson, K. Imaging and Manipulating Proteins in Live Cells through Covalent Labeling. *Nat Chem Biol* **2015**, *11* (12), 917–923.  
<https://doi.org/10.1038/nchembio.1959>.
- (95) Rived, F.; Canals, I.; Bosch, E.; Rosés, M. Acidity in Methanol–Water. *Analytica Chimica Acta* **2001**, *439* (2), 315–333. [https://doi.org/https://doi.org/10.1016/S0003-2670\(01\)01046-7](https://doi.org/https://doi.org/10.1016/S0003-2670(01)01046-7).
- (96) Grimm, J. B.; Lavis, L. D. Synthesis of Rhodamines from Fluoresceins Using Pd-Catalyzed C–N Cross-Coupling. *Org. Lett.* **2011**, *13* (24), 6354–6357. <https://doi.org/10.1021/ol202618t>.
- (97) Rodriguez, E. A.; Campbell, R. E.; Lin, J. Y.; Lin, M. Z.; Miyawaki, A.; Palmer, A. E.; Shu, X.; Zhang, J.; Tsien, R. Y. The Growing and Glowing Toolbox of Fluorescent and Photoactive Proteins. *Trends in Biochemical Sciences* **2017**, *42* (2), 111–129.  
<https://doi.org/10.1016/j.tibs.2016.09.010>.
- (98) Lotze, J.; Reinhardt, U.; Seitz, O.; Beck-Sickinger, A. G. Peptide-Tags for Site-Specific Protein Labelling in Vitro and in Vivo. *Mol. Biosyst.* **2016**, *12* (6), 1731–1745.  
<https://doi.org/10.1039/C6MB00023A>.
- (99) Li, C.; Tebo, A. G.; Gautier, A. Fluorogenic Labeling Strategies for Biological Imaging. *International Journal of Molecular Sciences* **2017**, *18* (7), 1473.  
<https://doi.org/10.3390/ijms18071473>.
- (100) 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. *Nature Biotechnology* **2003**, *21* (1), 86–89. <https://doi.org/10.1038/nbt765>.
- (101) Gautier, A.; Juillerat, A.; Heinis, C.; Corrêa, I. R.; Kindermann, M.; Beaufile, F.; Johnsson, K. An Engineered Protein Tag for Multiprotein Labeling in Living Cells. *Chemistry & Biology* **2008**, *15* (2), 128–136. <https://doi.org/https://doi.org/10.1016/j.chembiol.2008.01.007>.
- (102) Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; et al. HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chem. Biol.* **2008**, *3* (6), 373–382.  
<https://doi.org/10.1021/cb800025k>.
- (103) 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* (46), 16610–16611. <https://doi.org/10.1021/ja904800k>.

- (104) 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. *Angewandte Chemie International Edition* **2012**, *51* (23), 5611–5614. <https://doi.org/10.1002/anie.201200867>.
- (105) Hori, Y.; Norinobu, T.; Sato, M.; Arita, K.; Shirakawa, M.; Kikuchi, K. Development of Fluorogenic Probes for Quick No-Wash Live-Cell Imaging of Intracellular Proteins. *J. Am. Chem. Soc.* **2013**, *135* (33), 12360–12365. <https://doi.org/10.1021/ja405745v>.
- (106) Miller, L. W.; Cai, Y.; Sheetz, M. P.; Cornish, V. W. In Vivo Protein Labeling with Trimethoprim Conjugates: A Flexible Chemical Tag. *Nature Methods* **2005**, *2* (4), 255–257. <https://doi.org/10.1038/nmeth749>.
- (107) Jing, C.; Cornish, V. W. A Fluorogenic TMP-Tag for High Signal-to-Background Intracellular Live Cell Imaging. *ACS Chem. Biol.* **2013**, *8* (8), 1704–1712. <https://doi.org/10.1021/cb300657r>.
- (108) 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* (3), 1073–1080. <https://doi.org/10.1021/ja506376j>.
- (109) Herwig, L.; Rice, A. J.; Bedbrook, C. N.; Zhang, R. K.; Lignell, A.; Cahn, J. K. B.; Renata, H.; Dodani, S. C.; Cho, I.; Cai, L.; et al. Directed Evolution of a Bright Near-Infrared Fluorescent Rhodopsin Using a Synthetic Chromophore. *Cell Chemical Biology* **2017**, *24* (3), 415–425. <https://doi.org/https://doi.org/10.1016/j.chembiol.2017.02.008>.
- (110) Griffin, B. A.; Adams, S. R.; Tsien, R. Y. Specific Covalent Labeling of Recombinant Protein Molecules Inside Live Cells. *Science* **1998**, *281* (5374), 269–272. <https://doi.org/10.1126/science.281.5374.269>.
- (111) 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* (21), 6063–6076. <https://doi.org/10.1021/ja017687n>.
- (112) Cao, H.; Chen, B.; Squier, T. C.; Mayer, M. U. CrAsH: A Biarsenical Multi-Use Affinity Probe with Low Non-Specific Fluorescence. *Chem. Commun.* **2006**, No. 24, 2601–2603. <https://doi.org/10.1039/B602699K>.
- (113) Cao, H.; Xiong, Y.; Wang, T.; Chen, B.; Squier, T. C.; Mayer, M. U. A Red Cy3-Based Biarsenical Fluorescent Probe Targeted to a Complementary Binding Peptide. *J. Am. Chem. Soc.* **2007**, *129* (28), 8672–8673. <https://doi.org/10.1021/ja070003c>.
- (114) Halo, T. L.; Appelbaum, J.; Hobert, E. M.; Balkin, D. M.; Schepartz, A. Selective Recognition of Protein Tetraserine Motifs with a Cell-Permeable, Pro-Fluorescent Bis-Boronic Acid. *J. Am. Chem. Soc.* **2009**, *131* (2), 438–439. <https://doi.org/10.1021/ja807872s>.
- (115) Szent-Gyorgyi, C.; Schmidt, B. F.; Creeger, Y.; Fisher, G. W.; Zakel, K. L.; Adler, S.; Fitzpatrick, J. A. J.; Woolford, C. A.; Yan, Q.; Vasilev, K. V.; et al. Fluorogen-Activating Single-Chain Antibodies for Imaging Cell Surface Proteins. *Nature Biotechnology* **2007**, *26*, 235.
- (116) Hayriye Özhalici-Ünal; Pow, C. L.; Marks, S. A.; Jesper, L. D.; Silva, G. L.; Shank, N. I.; Jones, E. W.; Burnette, J. M.; 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* (38), 12620–12621. <https://doi.org/10.1021/ja805042p>.
- (117) 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* (4), 1012–1020. <https://doi.org/10.1039/C0OB00444H>.
- (118) He, J.; Wang, Y.; Missinato, M. A.; Onuoha, E.; Perkins, L. A.; Watkins, S. C.; St Croix, C. M.; Tsang, M.; Bruchez, M. P. A Genetically Targetable Near-Infrared Photosensitizer. *Nature Methods* **2016**, *13*, 263.

- (119) Saurabh, S.; Perez, A. M.; Comerchi, C. J.; Shapiro, L.; Moerner, W. E. Super-Resolution Imaging of Live Bacteria Cells Using a Genetically Directed, Highly Photostable Fluoromodule. *J. Am. Chem. Soc.* **2016**, *138* (33), 10398–10401. <https://doi.org/10.1021/jacs.6b05943>.
- (120) Wang, Y.; Ballou, B.; Schmidt, B. F.; Andreko, S.; St. Croix, C. M.; Watkins, S. C.; Bruchez, M. P. Affibody-Targeted Fluorogen Activating Protein for in Vivo Tumor Imaging. *Chem. Commun.* **2017**, 53 (12), 2001–2004. <https://doi.org/10.1039/C6CC09137G>.
- (121) Plamont, M.-A.; Billon-Denis, E.; Maurin, S.; Gauron, C.; Pimenta, F. M.; Specht, C. G.; Shi, J.; Quérard, J.; Pan, B.; Rossignol, J.; et al. Small Fluorescence-Activating and Absorption-Shifting Tag for Tunable Protein Imaging in Vivo. *Proceedings of the National Academy of Sciences* **2016**, *113* (3), 497–502. <https://doi.org/10.1073/pnas.1513094113>.
- (122) Li, C.; Plamont, M.-A.; Sladitschek, H. L.; Rodrigues, V.; Aujard, I.; Neveu, P.; Saux, T. L.; Jullien, L.; Gautier, A. Dynamic Multicolor Protein Labeling in Living Cells. *Chem. Sci.* **2017**. <https://doi.org/10.1039/C7SC01364G>.
- (123) Tebo, A. G.; Pimenta, F. M.; Zoumpoulaki, M.; Kikuti, C.; Sirkia, H.; Plamont, M.-A.; Houdusse, A.; Gautier, A. Circularly Permuted Fluorogenic Proteins for the Design of Modular Biosensors. *ACS Chem. Biol.* **2018**, *13* (9), 2392–2397. <https://doi.org/10.1021/acscchembio.8b00417>.
- (124) Tebo, A. G.; Pimenta, F. M.; Zhang, Y.; Gautier, A. Improved Chemical-Genetic Fluorescent Markers for Live Cell Microscopy. *Biochemistry* **2018**, *57* (39), 5648–5653. <https://doi.org/10.1021/acs.biochem.8b00649>.
- (125) Smith, E. M.; Gautier, A.; Puchner, E. M. Single-Molecule Localization Microscopy with the Fluorescence-Activating and Absorption-Shifting Tag (FAST) System. *ACS Chem. Biol.* **2019**, *14* (6), 1115–1120. <https://doi.org/10.1021/acscchembio.9b00149>.
- (126) Tebo, A. G.; Gautier, A. A Split Fluorescent Reporter with Rapid and Reversible Complementation. *Nature Communications* **2019**, *10* (1), 2822. <https://doi.org/10.1038/s41467-019-10855-0>.
- (127) Rozinov, M. N.; Nolan, G. P. Evolution of Peptides That Modulate the Spectral Qualities of Bound, Small-Molecule Fluorophores. *Chemistry & Biology* **1998**, *5* (12), 713–728. [https://doi.org/https://doi.org/10.1016/S1074-5521\(98\)90664-0](https://doi.org/https://doi.org/10.1016/S1074-5521(98)90664-0).
- (128) Marks, K. M.; Rosinov, M.; Nolan, G. P. In Vivo Targeting of Organic Calcium Sensors via Genetically Selected Peptides. *Chemistry & Biology* **2004**, *11* (3), 347–356. <https://doi.org/10.1016/j.chembiol.2004.03.004>.
- (129) Sunbul, M.; Nacheva, L.; Jäschke, A. Proximity-Induced Covalent Labeling of Proteins with a Reactive Fluorophore-Binding Peptide Tag. *Bioconjugate Chem.* **2015**, *26* (8), 1466–1469. <https://doi.org/10.1021/acs.bioconjchem.5b00304>.
- (130) Said, H. M. Biotin: Biochemical, Physiological and Clinical Aspects. In *Water Soluble Vitamins: Clinical Research and Future Application*; Stanger, O., Ed.; Springer Netherlands: Dordrecht, 2012; pp 1–19. [https://doi.org/10.1007/978-94-007-2199-9\\_1](https://doi.org/10.1007/978-94-007-2199-9_1).
- (131) Ren, W. X.; Han, J.; Uhm, S.; Jang, Y. J.; Kang, C.; Kim, J.-H.; Kim, J. S. Recent Development of Biotin Conjugation in Biological Imaging, Sensing, and Target Delivery. *Chem. Commun.* **2015**, 51 (52), 10403–10418. <https://doi.org/10.1039/C5CC03075G>.
- (132) Gao, M.; Yu, F.; Lv, C.; Choo, J.; Chen, L. Fluorescent Chemical Probes for Accurate Tumor Diagnosis and Targeting Therapy. *Chem. Soc. Rev.* **2017**, *46* (8), 2237–2271. <https://doi.org/10.1039/C6CS00908E>.
- (133) Chen, S.; Zhao, X.; Chen, J.; Chen, J.; Kuznetsova, L.; Wong, S. S.; Ojima, I. Mechanism-Based Tumor-Targeting Drug Delivery System. Validation of Efficient Vitamin Receptor-Mediated Endocytosis and Drug Release. *Bioconjugate Chem.* **2010**, *21* (5), 979–987. <https://doi.org/10.1021/bc9005656>.
- (134) Vineberg, J. G.; Zuniga, E. S.; Kamath, A.; Chen, Y.-J.; Seitz, J. D.; Ojima, I. Design, Synthesis, and Biological Evaluations of Tumor-Targeting Dual-Warhead Conjugates for a Taxoid–

- Camptothecin Combination Chemotherapy. *J. Med. Chem.* **2014**, *57* (13), 5777–5791. <https://doi.org/10.1021/jm500631u>.
- (135) Jung, D.; Maiti, S.; Lee, J. H.; Lee, J. H.; Kim, J. S. Rational Design of Biotin–Disulfide–Coumarin Conjugates: A Cancer Targeted Thiol Probe and Bioimaging. *Chem. Commun.* **2014**, *50* (23), 3044–3047. <https://doi.org/10.1039/C3CC49790A>.
- (136) Maiti, S.; Park, N.; Han, J. H.; Jeon, H. M.; Lee, J. H.; Bhuniya, S.; Kang, C.; Kim, J. S. Gemcitabine–Coumarin–Biotin Conjugates: A Target Specific Theranostic Anticancer Prodrug. *J. Am. Chem. Soc.* **2013**, *135* (11), 4567–4572. <https://doi.org/10.1021/ja401350x>.
- (137) Bhuniya, S.; Maiti, S.; Kim, E.-J.; Lee, H.; Sessler, J. L.; Hong, K. S.; Kim, J. S. An Activatable Theranostic for Targeted Cancer Therapy and Imaging. *Angewandte Chemie International Edition* **2014**, *53* (17), 4469–4474. <https://doi.org/10.1002/anie.201311133>.
- (138) Kim, T.; Jeon, H. M.; Le, H. T.; Kim, T. W.; Kang, C.; Kim, J. S. A Biotin-Guided Fluorescent-Peptide Drug Delivery System for Cancer Treatment. *Chem. Commun.* **2014**, *50* (57), 7690–7693. <https://doi.org/10.1039/C4CC02878C>.
- (139) Kumar, R.; Han, J.; Lim, H.-J.; Ren, W. X.; Lim, J.-Y.; Kim, J.-H.; Kim, J. S. Mitochondrial Induced and Self-Monitored Intrinsic Apoptosis by Antitumor Theranostic Prodrug: In Vivo Imaging and Precise Cancer Treatment. *J. Am. Chem. Soc.* **2014**, *136* (51), 17836–17843. <https://doi.org/10.1021/ja510421q>.
- (140) Li, K.; Qiu, L.; Liu, Q.; Lv, G.; Zhao, X.; Wang, S.; Lin, J. Conjugate of Biotin with Silicon(IV) Phthalocyanine for Tumor-Targeting Photodynamic Therapy. *Journal of Photochemistry and Photobiology B: Biology* **2017**, *174*, 243–250. <https://doi.org/https://doi.org/10.1016/j.jphotobiol.2017.08.003>.
- (141) Li, K.; Dong, W.; Liu, Q.; Lv, G.; Xie, M.; Sun, X.; Qiu, L.; Lin, J. A Biotin Receptor-Targeted Silicon(IV) Phthalocyanine for in Vivo Tumor Imaging and Photodynamic Therapy. *Journal of Photochemistry and Photobiology B: Biology* **2019**, *190*, 1–7. <https://doi.org/10.1016/j.jphotobiol.2018.09.001>.
- (142) Pal, K.; Sharma, A.; Koner, A. L. Synthesis of Two-Photon Active Tricomponent Fluorescent Probe for Distinguishment of Biotin Receptor Positive and Negative Cells and Imaging 3D-Spheroid. *Org. Lett.* **2018**, *20* (20), 6425–6429. <https://doi.org/10.1021/acs.orglett.8b02748>.
- (143) Jin, H.; Jin, Q.; Liang, Z.; Liu, Y.; Qu, X.; Sun, Q. Quantum Dot Based Fluorescent Traffic Light Nanoprobe for Specific Imaging of Avidin-Type Biotin Receptor and Differentiation of Cancer Cells. *Anal. Chem.* **2019**, *91* (14), 8958–8965. <https://doi.org/10.1021/acs.analchem.9b00924>.
- (144) Anderson, J. P.; Reynolds, B. L.; Baum, K.; Williams, J. G. Fluorescent Structural DNA Nanoballs Functionalized with Phosphate-Linked Nucleotide Triphosphates. *Nano Lett.* **2010**, *10* (3), 788–792. <https://doi.org/10.1021/nl9039718>.
- (145) Soares da Costa, T. P.; Tieu, W.; Yap, M. Y.; Zvarec, O.; Bell, J. M.; Turnidge, J. D.; Wallace, J. C.; Booker, G. W.; Wilce, M. C. J.; Abell, A. D.; et al. Biotin Analogues with Antibacterial Activity Are Potent Inhibitors of Biotin Protein Ligase. *ACS Med. Chem. Lett.* **2012**, *3* (6), 509–514. <https://doi.org/10.1021/ml300106p>.
- (146) Ma, H.; Tu, L.-C.; Naseri, A.; Chung, Y.-C.; Grunwald, D.; Zhang, S.; Pederson, T. CRISPR-Sirius: RNA Scaffolds for Signal Amplification in Genome Imaging. *Nature Methods* **2018**, *15* (11), 928. <https://doi.org/10.1038/s41592-018-0174-0>.
- (147) Serganov, A.; Nudler, E. A Decade of Riboswitches. *Cell* **2013**, *152* (1), 17–24. <https://doi.org/10.1016/j.cell.2012.12.024>.
- (148) Dwight, S. J.; Levin, S. Scalable Regioselective Synthesis of Rhodamine Dyes. *Organic Letters* **2016**, *18* (20), 5316–5319. <https://doi.org/10.1021/acs.orglett.6b02635>.
- (149) Lu, D.; Hossain, M. D.; Jia, Z.; Monteiro, M. J. One-Pot Orthogonal Copper-Catalyzed Synthesis and Self-Assembly of L-Lysine-Decorated Polymeric Dendrimers. *Macromolecules* **2015**, *48* (6), 1688–1702. <https://doi.org/10.1021/acs.macromol.5b00195>.

- (150) Duro-Castano, A.; England, R. M.; Razola, D.; Romero, E.; Oteo-Vives, M.; Morcillo, M. A.; Vicent, M. J. Well-Defined Star-Shaped Polyglutamates with Improved Pharmacokinetic Profiles As Excellent Candidates for Biomedical Applications. *Mol. Pharmaceutics* **2015**, *12* (10), 3639–3649. <https://doi.org/10.1021/acs.molpharmaceut.5b00358>.
- (151) Trofymchuk, K.; Valanciunaite, J.; Andreiuk, B.; Reisch, A.; Collot, M.; Klymchenko, A. S. BODIPY-Loaded Polymer Nanoparticles: Chemical Structure of Cargo Defines Leakage from Nanocarrier in Living Cells. *J. Mater. Chem. B* **2019**, *7* (34), 5199–5210. <https://doi.org/10.1039/C8TB02781A>.
- (152) Müller, P.; Schwille, P.; Weidemann, T. PyCorrFit—Generic Data Evaluation for Fluorescence Correlation Spectroscopy. *Bioinformatics* **2014**, *30* (17), 2532–2533. <https://doi.org/10.1093/bioinformatics/btu328>.
- (153) Fam, T. K.; Klymchenko, A. S.; Collot, M. Recent Advances in Fluorescent Probes for Lipid Droplets. *Materials* **2018**, *11* (9), 1768. <https://doi.org/10.3390/ma11091768>.
- (154) Collot, M.; Bou, S.; Fam, T. K.; Richert, L.; Mély, Y.; Danglot, L.; Klymchenko, A. S. Probing Polarity and Heterogeneity of Lipid Droplets in Live Cells Using a Push–Pull Fluorophore. *Anal. Chem.* **2019**, *91* (3), 1928–1935. <https://doi.org/10.1021/acs.analchem.8b04218>.
- (155) Shulov, I.; Oncul, S.; Reisch, A.; Arntz, Y.; Collot, M.; Mely, Y.; Klymchenko, A. S. Fluorinated Counterion-Enhanced Emission of Rhodamine Aggregates: Ultrabright Nanoparticles for Bioimaging and Light-Harvesting. *Nanoscale* **2015**, *7* (43), 18198–18210. <https://doi.org/10.1039/C5NR04955E>.
- (156) Dutt Vadlapudi, A.; Krishna Vadlapatla, R.; K. Mitra, A. Sodium Dependent Multivitamin Transporter (SMVT): A Potential Target for Drug Delivery. *Current Drug Targets* **2012**, *13* (7), 994–1003. <https://doi.org/10.2174/138945012800675650>.



## List of publications

# = equal contribution; \* = correspondence

1. Collot, M.\*; Boutant, E.; **Fam, K.T.**; Danglot, L.; Klymchenko, A. S. Molecular Tuning of Styryl Dyes Leads to Versatile and Efficient Plasma Membrane Probes for Cell and Tissue Imaging. *Bioconjugate Chem.* 2020 (in press). DOI: 10.1021/acs.bioconjchem.0c00023
2. Bouhedda, F.#; **Fam, T. K.**#; Collot, M.\*; Autour A.; Marzi, S.; Klymchenko, A.S.\*; Ryckelynck, M.\* A Dimerization-Based Fluorogenic Dye-Aptamer Module for RNA Imaging in Live Cells. *Nat. Chem. Bio.* 2020, 16 (1), 69–76. DOI: 10.1038/s41589-019-0381-8
  - **Highlighted in** *Un marqueur fluorescent pour pister l'ARN cible* [www.recherche.unistra.fr/index.php?id=31168](http://www.recherche.unistra.fr/index.php?id=31168)
3. Wang, X.; Anton, N.\*; Ashokkumar, P.; Anton, H.; **Fam, T. K.**; Vandamme, T. F.; Klymchenko, A. S.\*; Collot, M.\* Optimizing the Fluorescence Properties of Nanoemulsions for Single Particle Tracking in Live Cells. *ACS Appl. Mater. Interfaces* 2019, 11 (14), 13079–13090. DOI: 10.1021/acsami.8b22297
4. Collot, M.\*; Bou, S.; **Fam, T. K.**; Richert, L.; Mély, Y.; Danglot, L.; Klymchenko, A. S. Probing Polarity and Heterogeneity of Lipid Droplets in Live Cells Using a Push–Pull Fluorophore. *Anal. Chem.* 2019, 91 (3), 1928–1935. DOI: 10.1021/acs.analchem.8b04218
5. **Fam, T.K.**; Klymchenko, A.; Collot, M.\* Recent Advances in Fluorescent Probes for Lipid Droplets. *Materials* 2018, 11 (9), 1768. DOI: 10.3390/ma11091768
6. Collot, M.\*; **Fam, T. K.**; Ashokkumar, P.; Faklaris, O.; Galli, T.; Danglot, L.; Klymchenko, A. S. Ultrabright and Fluorogenic Probes for Multicolor Imaging and Tracking of Lipid Droplets in Cells and Tissues. *J. Am. Chem. Soc.* 2018, 140 (16), 5401–5411. DOI: 10.1021/jacs.7b12817
  - **Highlighted in** Spotlights on Recent JACS Publications. *J. Am. Chem. Soc.* 2018, 140 (16), 5321–5321. DOI: 10.1021/jacs.8b03930
  - **Highlighted in** 2018 UNE ANNÉE AVEC LE CNRS EN ALSACE, p.10.
7. Artamonov, O.S.; Bulda, T.; **Fam, T.K.**; Slobodyanyuk, E.Y.; Volochnyuk, D.M.; Grygorenko, O.O.\* A stereolibrary of conformationally restricted amino acids based on pyrrolidinyl/piperidinylloxazole motifs. *Heterocycl. Commun.* 2015, 21, 391–395. DOI: 10.1515/hc-2015-0137
8. Collot, M.\*; Schild, J.; **Fam, T. K.**; Menadi, F.; Bouchaala R.; Klymchenko, A.S.\* *ACS Nano* 2020 (submitted)
9. **Fam, K.T.**; Collot, M.\*; Klymchenko, A.S.\* Probing cell surface biotin receptors with rationally designed fluorogenic dimerized squaraines. 2020 (to be submitted)



## Patents

1. Ryckelynck M.; Bouhedda F.; Klymchenko A.; Collot M.; **Fam K.T.** 'Fluorescent compounds for detection of nucleic acids'. European patent application number EP19305806.2

# Bright dimerized fluorogenic probes for imaging nucleic acids and proteins

## Résumé

La dimérisation est un processus courant en biologie des protéines qui permet d'accéder à de nouvelles fonctionnalités. Toutefois, il existe peu d'exemples de stratégie de dimérisation en biologie chimique notamment pour l'imagerie par fluorescence. Inspirés par la nature, nous avons utilisé une approche de dimérisation afin d'enrichir la panoplie de sondes fluorogènes pour l'imagerie sélective d'événements biologiques en cellules vivantes. La stratégie repose sur une homodimérisation du fluorophore via une liaison courte et connecté à un ligand cible. Dans un premier temps, nous avons développé une famille de fluorogènes à base de rhodamines, dont l'émission de fluorescence va du vert au rouge lointain, nommée Gemini. Nous avons développé en collaboration, un fluoromodule photostable et brillant, Gemini-561/o-Coral, pour l'imagerie d'ARN dans les cellules vivantes. D'autre part, grâce à une analyse systématique de la relation structure-activité dans la conception de sondes dimérisées, nous avons développé des sondes pour la discrimination sélective des cellules cancéreuses exprimant des récepteurs membranaires à la biotine. Le quenching par dimérisation est une approche très attrayante dans la conception de sondes fluorogènes pour la biologie chimique. Tout en étant modulaire, elle fournit une réponse brillante et spécifique de la sonde à un événement biologique d'intérêt. Nous avons montré sa large applicabilité dans le domaine des technologies semi-synthétiques à base d'ARN et de peptide ainsi que pour l'imagerie ciblée de protéines. L'expansion des sondes fluorogènes dimérisées contribuera de manière significative à la compréhension de la complexité des processus biologiques dans les systèmes vivants.

**Mots-clés:** fluorogen, ARN, fluorescence, quenching par dimérisation, récepteurs de la biotine, Transporteur multivitaminique dépendant du sodium

## Abstract

Although dimerization is common process in protein biology enabling to discover new functionalities, there are few examples of dimerization strategy in chemical biology for fluorescence imaging. Inspired by nature, we used dimerization approach to expand the toolbox of bright fluorogenic probes for selective imaging of live-cell biological events. The strategy was based on intramolecular self-aggregation of dyes, in particular homodimerization of the fluorophore *via* short linkage and tethered targeted ligand. We developed a family of fluorogens based of rhodamine scaffolds spanning their fluorescence from green to far-red region, named Gemini. In collaboration, we developed a bright photostable fluoromodule, Gemini-651/o-Coral, for RNA live-cell imaging. Systematic analysis of structure-activity relationship in dimerization probe design enabled the development of the probe for selective discrimination of cancer cells expressing biotin receptors on the cell surface. Dimerization-caused quenching is a very attractive approach in fluorogenic probe design for chemical biology. While being modular, it provides bright activated fluorescence response of probes to biological event of interest. We showed its broad applicability in field of semisynthetic RNA/peptide-based technologies and targeted protein imaging. Further expansion of dimerized fluorogenic probes will significantly contribute to understanding the complexity of biological processes in living systems.

**Keywords:** fluorogen, RNA, aptamer, fluorescence, dimerization-caused quenching, biotin receptors, Sodium dependent multivitamin transporter