Organic & Biomolecular Chemistry

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: X. Song, H. Bian, C. Wang, M. Hu, N. Li and Y. Xiao, *Org. Biomol. Chem.*, 2017, DOI: 10.1039/C7OB01698K.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/obc

View Article Online
DOI: 10.1039/C7OB01698K

Journal Name

ARTICLE

Development and Applications of a Near-infrared dye-Benzylguanine Conjugate to Specifically Label SNAP-tagged Proteins

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

DOI: 10.1039/x0xx00000x

www.rsc.org/

Xinbo Song,[a,§] Hui Bian,[a,§] Chao Wang,[a] Mingyu Hu,[a] Ning Li [a] and Yi Xiao*[a]

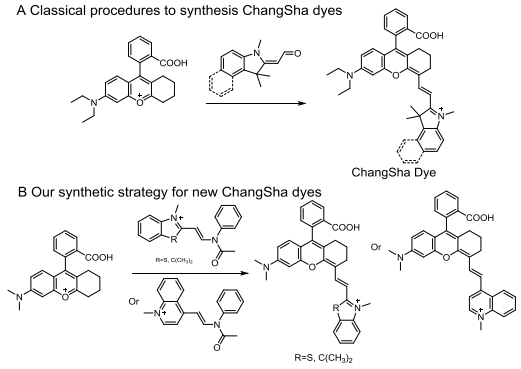
Near-infrared (NIR) fluorescent probes are advantageous over visible ones, for they can avoid the interference from the short-wavelength background emission in biological systems. However, there are a very limited number of NIR probes that can specifically label target proteins in living cells. In this work, a series of long-wavelength dyes (**N-NIR**, **S-NIR**, and **K-NIR**) analogue to the novel Changsha NIRs family are synthesized conveniently through a new approach different from previously reported one. These three dyes have similar conjugation structures but exhibit tunable photophysical properties. **N-NIR** and **S-NIR** have large extinction coefficients over 100,000, and high fluorescence quantum yields. Although NIR absorption and emission of **K-NIR** are inferior to the former two, it emits at much longer wavelength region. And all the three dyes can easy pass through the cell membranes to obtain the high-resolution NIR fluorescent images. Further, **N-NIR** is chosen as the NIR fluorophore to develop protein-labeling reagent **PYBG-D**, since it demonstrates highest fluorescence quantum yield up to 0.4 (in methanol). **PYBG-D** is efficiently synthesized through a Sonogashira coupling between bromo-substituted **N-NIR** and alkyne-substituted benzylguanine (**PYBG**). The conjugate **PYBG-D** proves to be a specific and efficient label of for O^6-alkylguanine-DNA alkyltransferase (SNAP-tag) that fused to target proteins in living cells, which contributes to high resolution NIR fluorescence images under a laser confocal microscope.

Introduction

Fluorescence labeling has made a significant contribution to the advances in cell and molecular biology.[1-9] A variety of fluorescent dyes and probes have been developed to label different subcellular organelles, such as endoplasmic reticulum,[10] lysosome,[4,11] and mitochondrion.[12] However, most of them have the absorption and emission in the visible range (400-650 nm). The fluorescent dyes or probes with absorption and emission in the near-infrared (NIR) region (650-900 nm) are relatively few. When compared to fluorescent dyes or probes with absorption and emission in the visible range, the NIR fluorescent dyes or probes have the advantages of minimum photodamage to biological samples, deep tissue penetration, and minimum interference from background auto-fluorescence by biomolecules in the living systems.[13-15]

Up to now, the common used NIR fluorophores belong to cyanine dyes, such as Cy5 and Cy7.[16-20] But their photostabilities are unsatisfactory and their fluorescence quantum yields are suboptimal.[21] Some NIR-fluorescent dyes include Si-rhodamines,[22-23] Bodipys,[24-26] and squaraines,[27] have been developed recently. Their photophysical properties of these new NIR fluorophores are enhanced to some extent. But their structures are complex and their synthetic methodologies are complicated, which make their further derivatization to obtain probes very challenging. Meanwhile, some NIR dyes with large conjugation skeletons have poor biological compatibility. Thus, the number of NIR probes is limited and their practical applicability needs to be further improved. There is still an urgent demand to develop more efficient NIR dyes. In 2011, Lin's group reported a class of novel NIR fluorophores named Changsha NIRs (scheme 1A).[28] Some of them exhibit excellent NIR fluorescence which is recommendable for biological imaging/sensing. However, so far, only a few derivatives or analogues of Changsha NIRs have

Scheme 1. Synthetic procedures for ChangSha dyes



[a] State Key Laboratory of Fine Chemicals, Dalian University of Technology, 2 Linggong Road, Dalian 116024, P.R. China E-mail: xiaoyi@dlut.edu.cn
[§] These authors contributed equally.

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

been developed by Lin or others.[28] And their applications are confined to small-molecule chemosensors.

Specifically labelling of a target protein in living cells is critical for chemical biology and cell biology. In 2003, a powerful technology, named SNAP-tag, developed by Johnsson's group, utilizes the human DNA repair protein O^6-alkyguanine-DNA alkyltransferase (AGT, known as SNAP-tag) to irreversibly transfer the alkyl group from O^6-[4-aminomethyl-benzyl] guanine derivatives, which opens up the possibility to label a single AGT fusion protein with dyes or other functional molecules.[29] From then on, there are some dye-alkyguanine conjugates that have developed and used for SNAP-tag labelling. But, to our knowledge, extremely few NIR label of SNAP-tag has been reported.[30 7, 31-32]

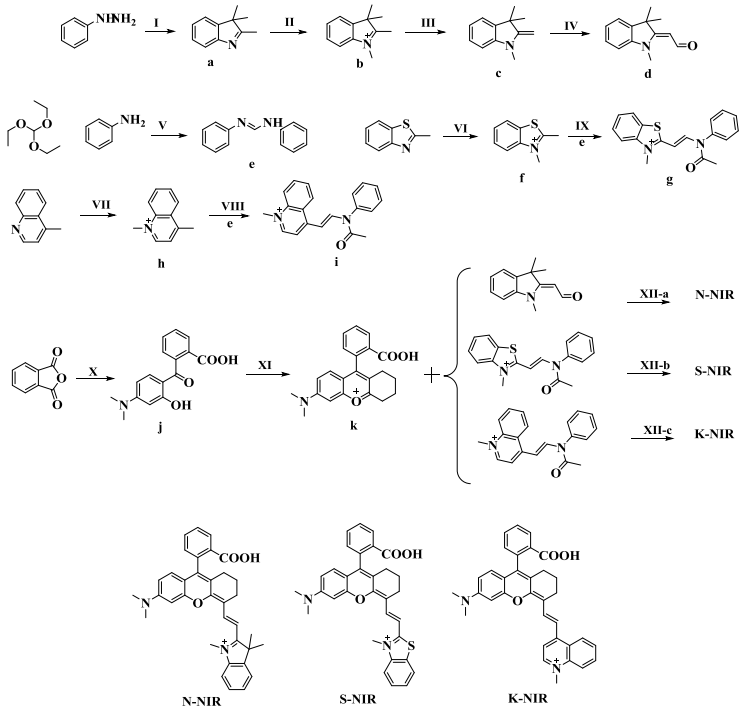
In this work, we propose a new synthetic approach to the derivatives/analogues of Changsha NIRs that is practical to obtain more candidates promising in NIR imaging (scheme 1B). The previous synthetic approach of Changsha analogues by Lin *et al*, which utilizes a readily available aldehyde, e.g. Fisher's aldehyde, to condense with another methylene intermediate activated by pyrylium, as demonstrated in Scheme 1a. However, we found that this approach might be not so practical if we hope to get other type of heterocylic Changsha analogs. This is because, as key intermediates, corresponding heterocyclic (especially the positively charged ones *e.g.* quinolinium and benzothiazolium) aldehydes are difficult to prepared. The limited source of positively charged heterocylic aldehydes could have restricted the expansion of the novel Changsha family. Here, to avoid the preparation of heterocylic aldehydes, we decide to utilize easily obtained intermediates for condensations, inspired by the syntheses of some cyanine dyes.[37-39] As shown in scheme 1b, new condensations are successful in the syntheses of different positively hetorocylic Changsha analogs. The reaction conditions are mild and the yields of new dyes are similar to Lin's approach using Fishers' aldehyde. And we also expand the application scope of Changsha NIRs to specifically protein labelling by using a derivative in the design of a substrate of genetically encoded SNAP protein tag.

Results and discussion

Design and Synthesis of NIR-fluorescence DYES: N-NIR, S-NIR, K-NIR, and PYBG-D.

It is known that the absorption and emission wavelengths of traditional rhodamine dyes are highly dependent on the substituents on the amino group. For instance, rhodamines 6G and B emit at 551 and 572nm.[33] And the absorption and emission wavelengths of traditional cyanine dyes can be tunable by using different substituents, such as indolium, thiazolium and quinolinium, this is because these moieties exert distinct electronic effects within the chromophores.[20,34-36] On the basis of this consideration, we design NIR dyes **N-NIR**, **S-NIR**, and **K-NIR**, for the sake to find out which is the best candidate for NIR imaging.

Scheme 2. Synthetic procedures for **N-NIR**, **S-NIR**, and **K-NIR**.



Condition: (I) 3-methyl-2-butanone, acetic acid, 90°C, 10h. (II) iodomethane, acetonitrile, reflux, 12h. (III) sodium hydroxide, water, 50°C, 2h. (IV) $POCl_3$, DMF, 0°C for 45min, add d, stir the mixture at 0°C for 20 min, then stir at 25°C for 12h. (V) triethyl orthoformate, acetic acid, 140°C, overnight. (VI) iodomethane, acetonitrile, reflux, 12h. (VII) iodomethane, acetonitrile, reflux, 12h. (VIII) **e**, acetic anhydride, 150°C, 0.5h. (IX) **e**, acetic anhydride, 150°C, 0.5h. (X) 3-dimethylamino phenol, toluene, under N_2, reflux, 3h; cool to 50°C; NaOH, 90°C, 6h; cool to 25°C, HCl. (XI) cyclohexanone, concentrated H_2SO_4, cool to 0°C; then add j in portions, 90°C, 1.5h; cool to 0°C, the mixture is added to water; add $HClO_4$. (XII) a: **k** and **d**, acetic anhydride, 50°C, 1h. b: **k** and **g**, CH_2Cl_2, CH_3OH, acetic anhydride, 50°C, 1h. c: **k** and **i**, CH_2Cl_2, CH_3OH, acetic anhydride, 50°C, 1h.

View Article Online
DOI: 10.1039/C7OB01698K

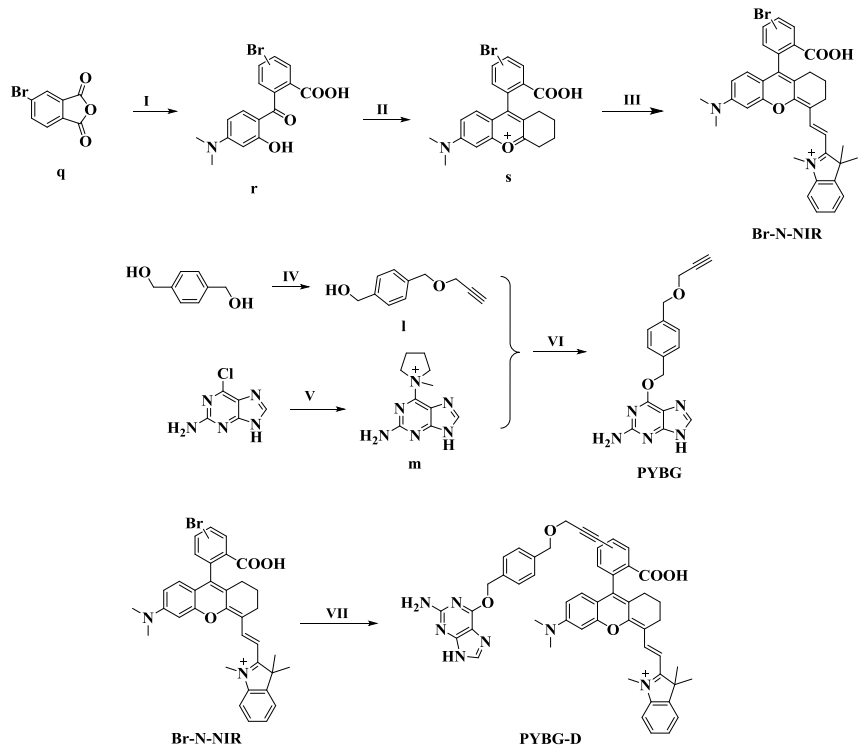
Journal Name

ARTICLE

The syntheses of **N-NIR**, **S-NIR**, and **K-NIR** are illustrated in scheme 1. Based on the retrosynthesis analysis, we design a multi-step synthetic route to **N-NIR** as outlined. Fisher's aldehyde d is synthesized according the literature,[34-36] and the starting materials o-phthalic anhydride is commercial available. Compound **j** and **k** are prepared according to the reported procedures.[4,28] Intermediate **k** condenses with aldehyde d under very mild conditions (in acetic anhydride at 50 °C within 15 min) to give **N-NIR** as a bright green solid. The synthetic approach (using to aldehyde for condensation reaction) of **N-NIR** is similar to Lin's Changsha NIRs, but we find this approach is not feasible in the preparation of **S-NIR**, and **K-NIR**. As a solution, we suggest a new approach that adopts g and i taking place of aldehyde as condensation components. Intermediate k condensates with compound g or i under very mild conditions (a mixture of CH_2Cl_2, CH_3OH, and acetic anhydride at 50°C for 1h), which give the target products **S-NIR**, and **K-NIR** with satisfactory yields. This new approach is efficient and versatile, which should be helpful to expand the diversity of analogues to Changsha NIR familiy.

In later investigation of spectral properties, we confirm that **N-NIR is** more promising than the other two dyes. So, we choose **N-NIR** as the fluorophore, and modify it by introducing a bromo group to obtain **Br-N-NIR** that can act as a key scaffold for synthesis of various NIR labels for different targets. Since in this work we want to label the cells through SNAP-tag fused proteins, a Sonogashira coupling between **Br-N-NIR** and a terminal alkyne substituted O^6-Benzylguaninethe (PYBG, synthesized according our previous work 5), which produces a novel conjugate **PYBG-D** as specific substrate and label of SNAP-tag.

Scheme 3. Synthetic procedures for PYBG-D.



Condition: (I) 3-dimethylamino phenol, toluene, under N_2, reflux, 3h; cool to 50°C; NaOH, 90°C, 6h; cool to 25°C, HCl. (II) cyclohexanone, concentrated H_2SO_4, cool to 0°C; then add j in portions, 90°C, 1.5h; cool to 0°C, the mixture is added to water; add $HClO_4$. (III) **k**, acetic anhydride, 50°C, 1h. (IV) NaH, 0°C, 15min; propargyl bromide, 0°C→25°C, 1h. (V) 1-methyl-pyrrolidin, DMF, 50°C, 24h. (VI) 4-[2-propynylmethoxy]-benzyl alcohol, NaH, DMF, 0°C, 10min, then 5 4-dimethylaminopyridine and **m** are added, 25°C, 1h. (VII) under N_2, PPh_3, $Pd[P[C_6H_5]_3]_4$, CuI, DMF and trimethylamine, 85°C, 4h

Optical Properties.

The absorption and emission profiles of **N-NIR**, **S-NIR**, and **K-NIR** in different solvents are shown in Figure 1 and the

corresponding photophysical data are recorded in Tables 1. Interestingly, **K-NIR** exhibit remarkably longer (80-90nm) absorption and emission than **N-NIR** and **S-NIR**. For instance, in CH_2Cl_2, the absorption and emission of **K-NIR** are peaked at 790 and 810 nm, respectively. The longer wavelength absorption and emission can be ascribed to the fact that K-NIR possesses a larger conjugated skeleton. If we simply define the conjugated length as the distance between eletron-donating dimethylamino nitrogen to the positively charged quaternary ammonium nitrogen, we just number the atoms and clearly know that K-NIR have longer conjugated size. The molar extinction coefficients (ε up to 100000−120000 $M^{-1}cm^{-1}$) of **N-NIR** and **S-NIR** are much larger than **K-NIR**. Also, **N-NIR** and **S-NIR** emit remarkably stronger fluorescence than **K-NIR**. Actually, the fluorescence quantum yields (Φ_f) of **N-NIR** and **S-NIR** are up to 0.4 and 0.21 in CH_3OH, which are big values among the known NIR functional dyes. The low fluorescence quantum yield of **K-NIR** reveals the non-radiative decay of excitation energy, which is possibly due to the nonplanar conjugated structure. The peri-H of quinoline should impose the steric hindrance effect to the adjacent CH=CH bond; thus, the nonplanar chromophoric system will undergo further motions (*e.g.* twist or torsion) to consume the excitation energy. Actually, the similar phenomena (weak fluorescence) had also been observed to some 4-quinoline cyanine dyes.[37,38,39] The water solubility of **N-NIR** and **S-NIR** is tested by fluorescence studies at different concentrations in ddH_2O (SI Figure 15 and16). Both of N-NIR and S-NIR are dispersive in ddH_2O at low concentration(S-NIR, <5uM) and aggregate when increase the concentration. However, the fluorescence of K-NIR is too weak to be detected accurately in water, so, it's not meaningful to study **K-NIR**'s relationship between concentration and fluorescence intensity.

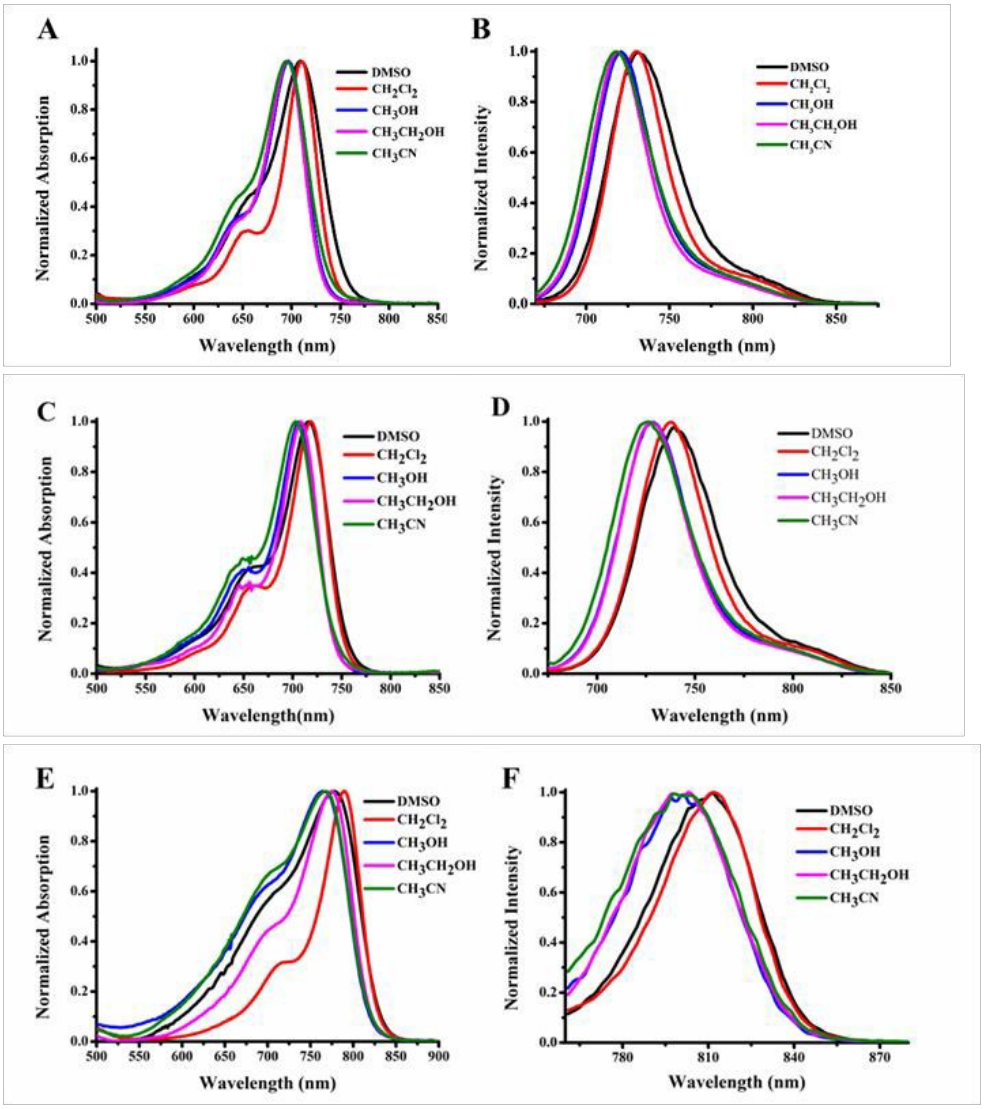


Figure 1. A: Normalized absorption spectra of **N-NIR**. B: normalized fluorescence spectra of **N-NIR**. C: Normalized absorption spectra of **S-NIR**. D: normalized fluorescence spectra of **S-NIR**. E: Normalized absorption spectra of **K-NIR**. F: normalized fluorescence spectra of **K-NIR**.

View Article Online
DOI: 10.1039/C7OB01698K

Table 1. Photophysical properties of N-NIR, S-NIR, and K-NIR.

Dyes	Solvent	λ_{abs} (nm)	ε (M^{-1} cm^{-1})	λ_{em} (nm)	Φ_f
N-NIR	CH_2Cl_2	710	108000	731	0.38[a]
	CH_3OH	697	130000	720	0.40[a]
	CH_3CH_2OH	696	129000	718	0.40[a]
	CH_3CN	695	120000	718	0.36[a]
	DMSO	710	118000	732	0.37[a]
S-NIR	CH_2Cl_2	718	119800	737	0.15[a]
	CH_3OH	707	122000	728	0.21[a]
	CH_3CH_2OH	708	120000	727	0.18[a]
	CH_3CN	703	115000	728	0.15[a]
	DMSO	717	108000	739	0.10[a]
K-NIR	CH_2Cl_2	790	35650	812	Not detected
	CH_3OH	765	65360	801	Not detected
	CH_3CH_2OH	775	60000	798	Not detected
	CH_3CN	768	28000	802	Not detected
	DMSO	778	8000	812	Not detected

[a] Reported 3, 5-bi (p–methoxy) phenyl-1, 7-bi (p-bromo) phenyl aza-BODIPY (Φ_f 0.42, in tolunene) is used as standard.[13]

Cell imaging applications of N-NIR, S-NIR, and K-NIR.

To confirm the applicability of **N-NIR**, **S-NIR**, and **K-NIR** in living cells, we attempt to label RAW cells and 7702 cells. These dyes exhibit good cell-membrane permeability and can fast stain RAW cells and 7702 cells, which is favourable for the high resolution NIR fluorescent images, as shown in Figure **2**. There is strong fluorescence in the cytoplasmic matrix and there is very weak fluorescence in the nucleus (Even no fluorescence). The place where **N-NIR** and **S-NIR** aggregated is like organelle in the cells. **K-NIR** is also used to label RAW cells. It can pass though cytomembrane to label the cells. However, its localization is not same with **N-NIR** or **S-NIR**. As is shown in Figure **3**, **K-NIR** can enter the nucleus to show strong fluorescence and there is weak fluorescence in the cytoplasmic matrix. We do not know the reason why **K-NIR** can label the nucleus, but we suspect that the structure of quinolone causes its targetability to nucleus, because some reported dyes can label nuclei which contain the structure of quinolone.[37]

Colocalization study of **N-NIR** and **S-NIR** with Rh-123 (Commercial mitochondrial probe) have been conducted. As demonstrated in Figure **4**, **N-NIR** showed red fluorescence in channel 1 (Figure **4A**). At the same time, Rh123 showed bright-green fluorescence in channel 2 (Figure **4B**). The merged image (Figure **4C**) indicates that the two channel images overlapped very well, confirming that **N-NIR** can localize in the mitochondria of living cells. The changes in the intensity profile

View Article Online
DOI: 10.1039/C7OB01698K

of linear regions of interest (ROIs) (Rh123 and **N-NIR** costaining) tends toward synchronization (Figure **4D**). From the intensity correlation plots (Figure **4E**), a high Pearson's coefficient and overlap coefficient of 0.915, is obtained. The above colocalization investigation confirmed that **N-NIR** can stain mitochondria specially. The similar result is obtained from **S-NIR** (Figure **5A**, **5B**, **5C**, **5D**, and **5E**).

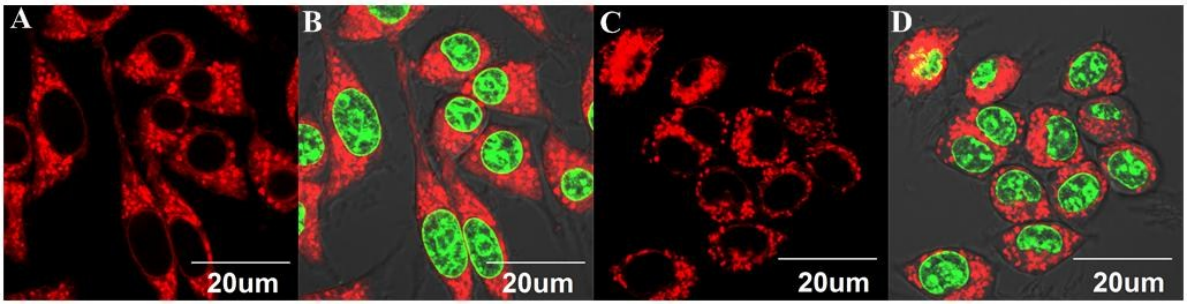


Figure 2. MCF-7 cells are costained with **N-NIR** and **S-NIR**. A: 4 μM **N-NIR** (λ_{ex} = 635 nm, λ_{em} = 650-750 nm, pseudo-color red). B: overlay of the fluorescence **N-NIR** and Hoechst 33342 (0.5 μM Hoechst 33342, λ_{ex} = 405 nm, λ_{em} = 420-480 nm, pseudo-color green). C: 6 μM **S-NIR** (λ_{ex} = 635 nm, λ_{em} = 650-750 nm) D: overlay of the fluorescence **S-NIR** and hoechst33342 (0.5 μM Hoechst 33342, λ_{ex} = 405 nm, λ_{em} = 420-480 nm).

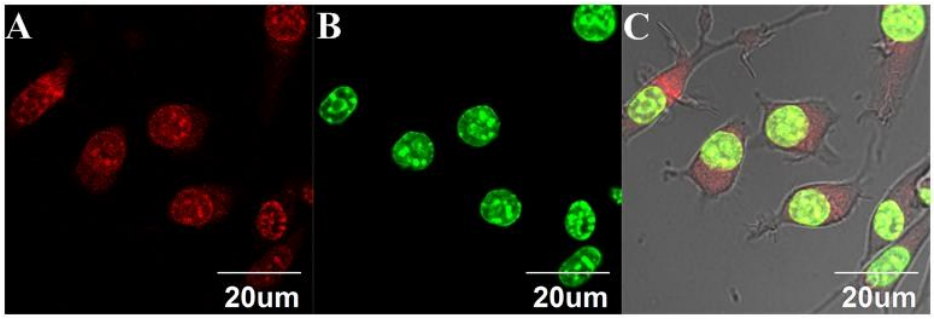


Figure 3. RAW cells are costained with (A) 8 μM **K-NIR** (λ_{ex} = 635 nm, λ_{em} = 650-750 nm) and (B) 0.5 μM Hoechst 33342, (λ_{ex} = 405 nm, λ_{em} = 420-480 nm) . C: overlay of A, B and brightfield .

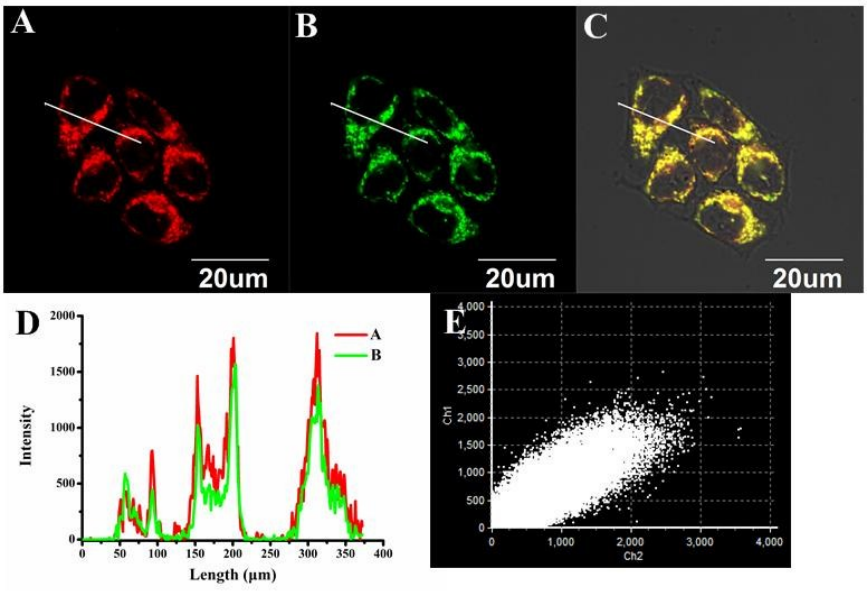


Figure 4. 7702 cells are co-stained with 4 μM **N-NIR** (A: λ_{ex} = 635 nm, λ_{em} = 650-750 nm) and 1 μM Rh-123 (B: λ_{ex} = 488 nm, λ_{em} = 500-550 nm) ; C: Overlay of A, B and brightfield; D Intensity profile of region of interest (ROI) cross 7702 cell; E: Correlation plot of **N-NIR** and Rh123.

View Article Online
DOI: 10.1039/C7OB01698K

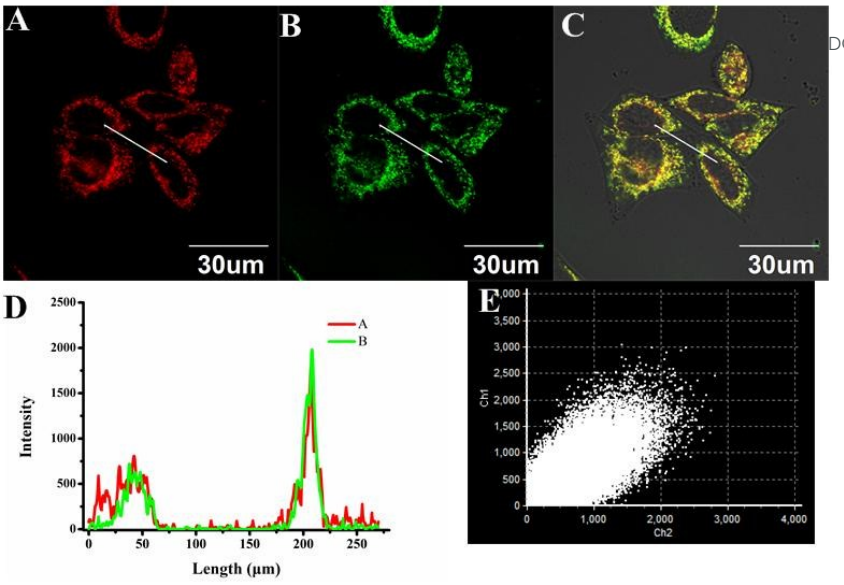


Figure 5. 7702 cells are co-stained with 4 μM **S-NIR** (A: λ_{ex} = 635 nm, λ_{em} = 650-750 nm) and 1 μM Rh-123 (B: λ_{ex} = 488 nm, λ_{em} = 500-550 nm) ; C: Overlay of A, B and brightfield; D Intensity profile of region of interest (ROI) cross 7702 cell; E: Correlation plot of **S-NIR** and Rh123.

Design PYBG-D to label SNAP-tag proteins and gene encoding cells.

To confirm the labelling reaction between **PYBG-D** and SNAP-tag proteins, the covalent labeled SNAP-tag with **PYBG-D** are analyzed by SDS-PAGE. As is shown in figure **6**, there is a strong fluorescence band at 21 kDa ascribed to SNAP-tag, which proves that **PYBG-D** labelling is successful.

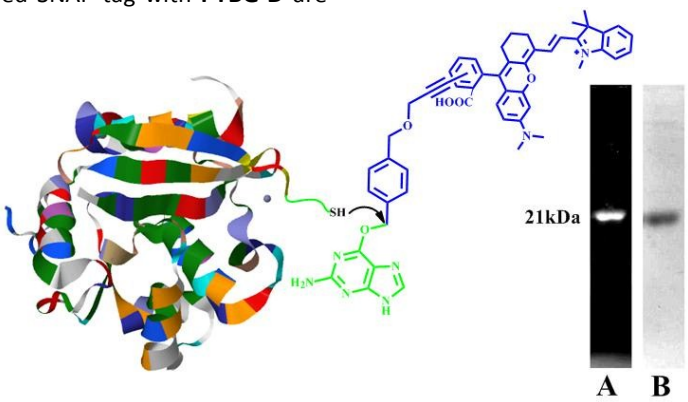


Figure 6. SDS-PAGE analysis of **PYBG-D** labelled SNAP-tag. A: .The fluorescence image of the gel; B: The coomassie blue image of the same gel.

In order to develop the applicability of **PYBG-D**, we choose the localization of the SNAP-tag to nucleus which is achieved by fusion of SNAP-tag to the C-terminus of the nuclear histone H2B proteins (pSNAPf-H2B, NEB). After transiently transfected with the plasmid, the nuclear histone H2B targeted SNAP-tag proteins are obtained in live cells. And then **PYBG-D** is used to label nuclear H2B. Fluorescent imaging is performed after washing out the un-reacted dye molecules, which rules out the influence of non-specific staining. As shown in Figure **7A** and **7B**, the transfected COS-7 cells stained with **PYBG-D** show specific labeling of the nucleus to produce a high resolution fluorescence image without staining the area outside the nucleus. Due to the incompletely transfection, a part of cells donot express H2B-SNAP, and thus they exhibit no intracellular fluorescence. This phenomenon just, from another side, reveals the high specificity of **PYBG-D** to label intracellular SNAP tag fused proteins.

Next, COS-7 cells are transfected with pSNAPf Vector that expresses diffusely in cytoplasm and nucleus. After staining with **PYBG-D**, high resolution fluorescence images are also obtained (Figure **7C**). Again, the cells without expressing SNAP to be labelled show no fluorescence due to the washing out of the unreacted **PYBG-D** (Figure **7D**).

Then we choose the cytochrome c oxidase specifically located in the mitochondrial inner membrane as the target protein for fusion with SNAP-tag. Such a mitochondria-targeting plasmid $pSNAP_f$-Cox8A is achieved by ligating the cytochrome c oxidase submit 8(Cox8A) gene fragment to the upstream of the SNAP coding sequence in the commercially available $pSNAP_f$ vector (NEB). And the recombinant $pSNAP_f$-Cox8A plasmid is stably transfected in COS-7 cells for **PYBG-D** labelling and fluorescence imaging.

View Article Online
DOI: 10.1039/C7OB01698K

Journal Name

ARTICLE

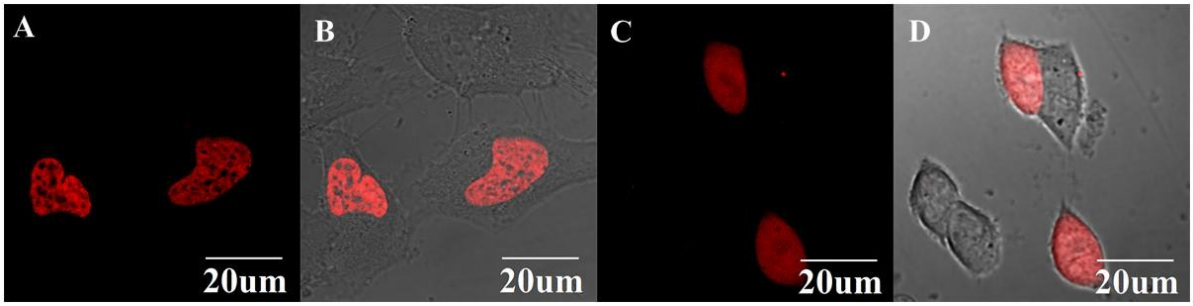


Figure 7. The transfected COS-7 cells are co-stained with 5 μM **PYBG-D** for 1 h; the cells are washed by three times, and then cells are cultured for 1 h (λ_{ex} = 635 nm, λ_{em} = 650-750 nm). A: The transfected cells express pSNAPf-H2B; B: Overlay of A and bright field; C: The transfected cells express pSNAPf; D: Overlay of C and bright field.

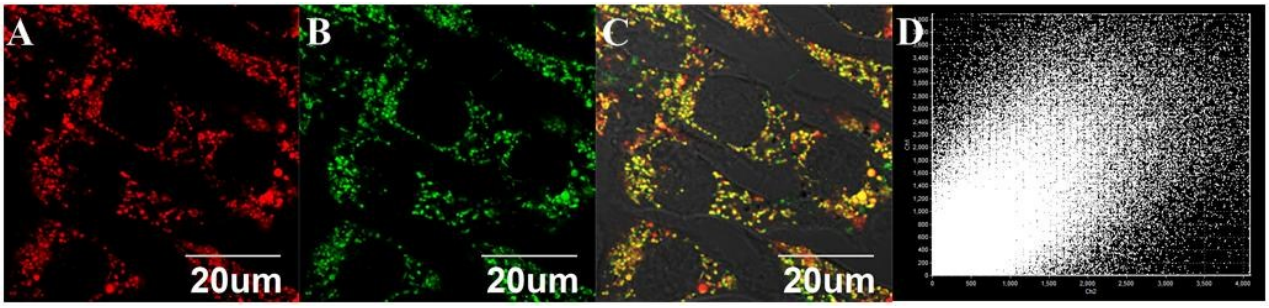


Figure 8. The transfected COS-7 cells are stained with 5 μM PYBG-D for 1 h; the cells are washed by three times, and then cells are cultured for 1 h (A:λ_{ex} = 635 nm, λ_{em} = 650-750 nm). And then the cells are stained with 0.5 μM Rh-123 for 5 min, washed three times(B: λ_{ex} = 488 nm, λ_{em} = 500-550 nm) ; C: Overlay of A, B and bright field; D: Correlation plot of **PYBG-D** and Rh123

As shown in Figure **8A**, the cells transfected with $pSNAP_f$-COX8A plasmid and stained with **PYBG-D** show excellent specific labelling of the mitochondria. To confirm the subcellular localization of the dyes, a commercially available mitochondrial dye (Rh123) is employed for the colocalization study. As shown in Figure **8**, Rh123 showed bright-green fluorescence in channel 1 (Figure **8B**). At the same time **PYBG-D** showed red fluorescence in channel 2 (Figure **8A**). The merged image (Figure **8C**) indicates that the two channel images overlapped very well, confirming that **PYBG-D** can specially react with fusion proteins expressed in the mitochondria of living cells. From the intensity correlation plots (Figure **8D**), a high Pearson's coefficient and overlap coefficient of 0.84, respectively, are obtained. And the stable covalent labeling method can be used for the dynamic process research of mitochondria even at mitochondrial membrane potential decrease and mitochondrial dysfunction, which cannot be well revealed by conventional mitochondrial trackers relying on the mitochondrial membrane potentials.

Conclusion

In conclusion, we have synthesized a series of NIR fluorescent dyes, *i.e.* **N-NIR**, **S-NIR**, and **K-NIR**. By proposing a more versatile synthetic strategy, we expand the diversity of analogues to the novel Changsha NIRs family. The new dyes exhibit tunable emission in the NIR region. For example, in dichloromethane their emission peaks are at 731, 737, and 812 nm respectively. Importantly, the fluorescent quantum yields of **N-NIR** and **S-NIR** are high, compared with most fluorescent dyes in the same spectral region. In chloromethane these values are up to 0.4 and 0.2 respectively. In addition, the dyes can readily pass through the cell membrane to stain different subcellular areas. Interestingly, **N-NIR** and **S-NIR** can target the mitochondria and their specific staining produce high resolution images of mitochondria. **K-NIR** exhibits a tendency to be localized in cell nuclei. Considering that **N-NIR** possesses the highest fluorescence quantum yield up to 0.4, an aromatic bromo is introduced to **N-NIR** to get **Br-N-NIR**, which makes the further modifications feasible. Next, using Sonogashira

View Article Online
DOI: 10.1039/C7OB01698K

coupling reaction between **Br-N-NIR** and **PYBG** (a terminal alkyne substituted O^6-benzylguanine developed by our group), a new NIR substrate of SNAP-tag, *i.e.* **PYBG-D**, is efficiently synthesized with high yield. In the cells transfected with pSNAPf-H2B, pSNAPf, and pSNAPf-Cox8A, **PYBG-D** can readily and specifically label these SNAP-tag fused proteins of interest to produce the high resolution NIR images. Thus, the NIR fluorescent dyes developed in this work demonstrate recommendable application properties. To our knowledge, the number of efficient NIR dyes is still very limited, and they have been rarely applied in genetically encoded fusion protein labelling.[7, 32] To this date, Changsha NIRs analogs/derivatives have yet been adopted in this powerful technology. And we expect that our new dyes will become more valuable tools for fluorescent imaging and sensing in cell biology.

Experimental

General methods

The 500 (^{1}H) MHz NMR and 100 (^{13}C) MHz NMR spectra are registered at room temperature on a Bruker 500 MHz spectrometer using perdeuterated solvents as internal standard. Chromatographic purification is conducted with silica gel. All solvent mixtures are given as volume/volume ratios. UV spectra are recorded on TU-1901 UV-Vis Spectrophotometer and fluorescence spectra are recorded on a F-7000 spectrometer. Relative quantum efficiencies of fluorescence of compounds are obtained by comparing the areas under the corrected emission spectrum of the test sample in diluted solvents with a reported aza-BODIPY derivative in toluene which has a quantum efficiency of 0.42 according to the literature.[1] Non-degassed, spectroscopic grade toluene and a 10 mm quartz cuvette are used. Dilute solutions ($0.01 < A < 0.05$) are used to minimize reabsorption effects. Quantum yields are determined using the equation (1):

$$\Phi_F^{(sample)} = \Phi^{(standard)} \times (Abs^{(standard)} \times F^{(sample)}) / (Abs^{(sample)} \times F^{(standard)}) \quad \text{equation (1)}$$

Where $\Phi^{(standard)}$ is the reported quantum yield of the standard, Abs is the absorbance at the excitation wavelength, F is the integrated emission spectra.

Culture cells.

7702 cells and RAW cells are obtained from Institute of Basic Medical Sciences (IBMS) of Chinese Academy of Medical Sciences (CAMS). All cell lines are maintained under standard culture conditions (atmosphere of 5% CO_2 and 95% air at 37 °C) in RPMI 1640 medium or DMEM medium, supplemented with 10% FBS (fetal bovine serum). The cells is used after the cells in the exponential phase of growth on 35 mm glass bottom culture dishes (Φ 20 mm) for 1 day.

Cells transfected.

COS-7 cells are purchased from Cell Bank of Type Culture Collection of Chinese Academy of Sciences. The cells are all maintained in Dulbecco's modified Eagle's medium (DMEM, Gibco) supplemented with 10% fetal bovine serum (FBS, Hyclone). The cells are cultured in a humidified atmosphere of 5% CO_2/95% air at 37 °C (CO_2 incubator, Thermo Scientific). Cells were transiently transfected with $pSNAP_f$ Vector, pSNAPf-H2B plasmid (NEB), or pSNAPf-COX8A plasmid (the construction was performed as we previously described26) by using Lipofectamine 2000 (Invitrogen) following the standard protocol. And stable transfected cells were established by G418 (Geneticin, Invitrogen) selection. Cells were cultured in 35 mm glass bottom dishes (Φ = 20 mm) for 24 h and used for fluorescent imaging.

PYBG-D labels the cells.

The probe PYBG-D is dissolved in DMSO (1 mM) by vortexing. The stock solution is mixed with culture medium (DMEM with 10% FBS) also by vortexing to obtain the labeling solution. Cells cultured in 35 mm glass bottom dishes are stained with the PYBG-D solution (5 μM) at 37 °C for 60 min and washed with PBS (3 × 1 mL). After incubation (1 h), the DMEM (10% FBS) is replaced with fresh medium 30 min before fluorescent imaging.

SDS-PAGE analysis of labeled SNAP-tag protein.

The method of labeling purified SNAP-tag proteins (20 μM) with PYBG-D (10 μM) is the same as that of labeling cells. Then labeled SNAP-tag proteins (10 μM, 8 μL) are analyzed by 15% SDS-PAGE. After electrophoresis, the fluorescence of gel is imaged with a UV transilluminator (GBOX-HR, SYNGENE) with an ethidium bromide filter. Then the gel is visualized by staining with Coomassie Blue.

Preparation of a

A solution of phenylhydrazine hydrochloride (30 g, 206mmol) and 3-methyl-2-butanone (27 ml, 248 mmol) in 500 ml AcOH are heated to 90°C for 10h. The mixture is distilled under reduced pressure to yield dark solid. The water (150 ml) is added and adjust pH to 10. The mixture is extracted with ether, dried with $MgSO_4$, and concentrated in vacuo to yield a red oil (24 g, 151 mmol, yields 72%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.56 (d, J = 7.6 Hz, 1H), 7.33 (dd, J = 7.5, 1.2 Hz, 1H), 7.32 – 7.29 (m, 1H), 7.22 (dd, J = 7.6, 7.1 Hz, 1H), 2.30 (s, 3H), 1.32 (s, 6H).

Preparation of b, f, and h

Methyl iodide (1.5 eq) is added to the solution of 2,3,3-trimethylindolenine, 2-methylbenzothiazole, or 4-methylquinoline in acetonitrile. And the reaction mixture is heated at reflux for 12 h. The resulting solid is filtered under vacuum, washed with Et_2O and dried to afford pure indolium **b, f,** and **h. b** ^{1}H NMR (400 MHz, DMSO) δ 7.84 – 7.79 (m, 1H), 7.75 – 7.70 (m, 1H), 7.56 – 7.48 (m, 2H), 3.88 (s, 3H), 2.67 (s, 3H), 1.43 (s, 6H). **f** ^{1}H NMR (400 MHz, DMSO) δ 8.44 (d, J = 8.1 Hz, 1H), 8.29 (d, J = 8.5 Hz, 1H), 7.92 – 7.86 (m, 1H), 7.80 (t, J = 7.7 Hz, 1H), 4.20 (s, 3H), 3.17 (s, 3H). **h**,^{1}H NMR (400 MHz, DMSO) δ 9.37 (d, J = 6.0 Hz, 1H), 8.58 – 8.47 (m, 2H), 8.28 (ddd, J = 8.6, 7.0, 1.2 Hz, 1H), 8.07 (dd, J = 8.7, 7.4 Hz, 2H), 4.59 (s, 3H), 3.01 (s, 3H).

Preparation of d[34-36]

A mixture of **b** (15 g, 50 mmol) and sodium hydroxide (2.3 g, 58 mmol) is stirred vigorously at 50°C in 100 ml water for 2 h and then extracted with dichloromethane. The organic phase is dried ($MgSO_4$) and then concentrated under reduced pressure

to obtain the product **c** which is used for the next step without further purification. $POCl_3$ (7.5 mL) is dropped to 24 ml cold DMF under stirring and the mixture is stirred below 10 °C for 45min. After adding a mixture of **c** (12.98 g) in DMF to the reaction vessel at this temperature, the reaction solution is stirred at room temperature for 10h. Then the reaction solution is poured into ice water and pH is adjust to 10 by addition of saturated sodium hydroxide solution. The resulting solution is heated to 60°C for a few minutes and cooled to room temperature. The product d is precipitated and filtered off giving 8.1 g (55%). Red crystals; mp 118 °C.

Preparation of e

A mixture of aniline (11 mL, 120 mmol), triethylorthoformate (10 mL, 60 mmol) and acetic acid (0.2 ml, 3.3 mmol) is heated at 140 ˚C overnight. The crude solid is triturated in cold n-pentane and filtered. Then it is washed with n-pentane and dried in vacuo to give the crude product as a pale yellow solid (9 g, yield: 77%) which is used for next step without further purification.

Preparation of g

2,3-Dimethylbenzothiazolium iodide (f, 6 g, 20 mmol) and e (8 g, 40 mmol) are suspended in acetic anhydride (30 mL). The suspension was stirred at 150 °C for 0.5 h. The reaction mixture is cooled to room temperature. Then the mixture is distilled under reduced pressure. Ethyl acetate is added to the residue, and the precipitate is filtered and dried in vacuo to give g as a dark red powder (4.5 g, 71%). ^{1}H NMR (400 MHz DMSO) δ 8.79 (d, J = 14.2, 1H), 8.31 (d, J = 22.0, 1H), 8.09 (d, J = 8.3, 1H), 7.80–7.52 (m, 7H), 5.68 (d, J = 13.7, 1H), 3.87 (s, 3H), 2.05 (s, 3H); HRMS (ESI) calc. for $C_{18}H_{17}N_2OS$ ($[M - I]^+$) 309.1062, found 309.1073.

Preparation of i

The method is similar to g. ^{1}H NMR (400 MHz, DMSO) δ 9.15 (d, J = 6.6 Hz, 1H), 8.88 (d, J = 14.0 Hz, 1H), 8.48 (d, J = 8.9 Hz, 1H), 8.37 – 8.33 (m, 1H), 8.27 (dd, J = 8.6, 7.3 Hz, 1H), 8.16 (s, 1H), 8.09 – 8.03 (m, 2H), 7.93 – 7.86 (m, 1H), 7.71 – 7.61 (m, 2H), 7.53 (d, J = 7.3 Hz, 1H), 6.03 (d, J = 14.0 Hz, 1H), 4.57 (s, 3H), 3.00 (s, 3H); HRMS (ESI) calc. for $C_{20}H_{19}N_2O$ ($[M - I]^+$) 303.1497, found 303.1503.

Preparation of j

A solution of 3-dimethylamino phenol (4.11 g, 30.0 mmol) and phthalic anhydride (4.66 g, 31.5 mmol) in toluene (30 mL) is refluxed under N_2 for 3 h, and cooled to 50–60 °C. Then 30 mL of 35% aqueous NaOH is added and heated at 90 °C for 6 h. The resulting mixture was poured into H_2O, acidified with HCl, and allowed to stand at room temperature for 2 h. The suspension is filtered. Then the solid is recrystallized from a mixture of water and methanol, and dried to afford the desired product (5.8 g, 69%). ^{1}H NMR (400 MHz, CDCl3) δ 8.12 (dd, J = 27.4, 8.1 Hz, 2H), 7.97 (s, 1H), 7.58 (s, 1H), 6.91 (d, J = 9.8 Hz, 1H), 6.14 – 6.03 (m, 2H), 3.06 (s, 6H).

Preparation of k[28]

Freshly cyclohexanone (6.6 mL, 63.7 mmol) is added to concentrated H_2SO_4 drop-wisely (7.0 mL) and then cooled down to 0 °C. Then, compound j (32 mmol) is added in portions with vigorous stirring. The reaction mixture is heated at 90 °C for 1.5 h, cooled down, and poured onto ice (300 g), and $HClO_4$ is added to the solutions and the resulting precipitate is filtered off and washed with cold water (100 mL). Compound k obtained as a bluish green is used for the next step without further purification.^{1}H NMR (400 MHz, $CDCl_3$:MeOD=2:1) δ 8.33 (d, J = 7.8 Hz, 1H), 7.82 (t, J = 7.5 Hz, 1H), 7.74 (t, J = 7.6 Hz, 1H), 7.26 – 7.13 (m, 3H), 6.98 (d, J = 1.9 Hz, 1H), 3.36 (s, 6H), 3.15 (t, J = 6.1 Hz, 2H), 2.32 (t, J = 5.8 Hz, 2H), 2.02 (d, J = 5.8 Hz, 2H), 1.88 – 1.79 (m, 2H).^{13}C NMR (100 MHz, $CDCl_3$:MeOD=2:1) δ 173.98 (s), 170.69 (s), 170.27 (s), 163.24 (s), 161.46 (s), 138.21 (s), 137.14 (s), 135.60 (s), 134.32 (s), 133.92 (s), 133.36 (s), 132.23 (s), 125.91 (s), 122.00 (s), 121.40 (s), 99.52 (s), 44.56 (s), 33.38 (s), 28.97 (s), 25.20 (s), 24.79 (s).

Preparation of N-NIR

Compound d (50 mg, 0.25 mmol) and compound k (130mg, 0.37mmol) are dissolved in acetic anhydride (5 mL). The mixture is heated to 50℃ for 1h. The solvent is removed under reduced pressure to give the crude product, which is purified by silica gel flash chromatography using CH_2Cl_2 to CH_2Cl_2/CH_3OH (200:1 to 20:1) as eluent to afford N-NIR. ^{1}H-NMR (400 MHz, CD_3OD) δ 8.70 (d, J = 14.4 Hz, 1H), 8.23 (d, J = 7.0 Hz, 1H), 7.74 (d, J = 4.5 Hz, 1H), 7.67 (d, J = 4.4 Hz, 1H), 7.59 – 7.50 (m, 1H), 7.43 (d, J = 7.7 Hz, 1H), 7.29 (dd, J = 24.4, 6.3 Hz, 3H), 6.77 (d, J = 12.9 Hz, 3H), 6.20 (d, J = 15.0 Hz, 1H), 3.65 (s, 3H), 3.16 (s, 6H), 2.69 (s, 2H), 2.36 (s, 2H), 2.12 (d, J = 65.5 Hz, 2H), 1.83 (s, 6H).^{13}C NMR (126 MHz, MeOD) δ 175.52 (s), 164.64 (s), 156.93 (s), 155.59 (s), 144.34 (s), 143.70 (s), 142.40 (s), 137.30 (s), 133.95 (s), 132.33 (s), 130.86 (s), 130.66 (s), 130.46 (s), 129.85 (s), 129.03 (s), 126.40 (s), 123.38 (s), 122.33 (s), 116.56 (s), 114.95 (s), 113.39 (s), 111.96 (s), 100.74 (s), 97.14 (s), 50.59 (s), 40.40 (s), 31.66 (s), 28.66 (s), 27.88 (s), 21.70 (s). m/z (FTMS+p ESI): Calcd $[M+]^+$ for $C_{35}H_{35}N_2O_3$: 531.2642, found: 531.2629.

Preparation of S-NIR

Compound **g** (200 mg, 0.46 mmol) and compound **k** (318mg, 0.92mmol) are dissolved in a mixture of CH_2Cl_2 (25 mL), CH_3OH (12 mL), and acetic anhydride (1.5 mL). The mixture is stirred for 1h at 50°C. The solvent was removed under reduced pressure to give the crude product, which is purified by silica gel flash chromatography using CH_2Cl_2 to CH_2Cl_2/CH_3OH (200:1 to 20:1) as eluent to afford **S-NIR** (120mg, yields 50%). ^{1}H-NMR (500 MHz, MeOD) δ 8.21 (d, *J* = 13.8 Hz, 1H), 7.89 (d, *J* = 5.9 Hz, 1H), 7.78 (d, *J* = 7.8 Hz, 1H), 7.45 (dt, *J* = 7.0, 6.1 Hz, 4H), 7.30 (t, *J* = 7.6 Hz, 1H), 6.99 (d, *J* = 8.7 Hz, 1H), 6.72 (d, *J* = 9.0 Hz, 1H), 6.50 (dd, *J* = 9.1, 2.3 Hz, 1H), 6.39 (d, *J* = 2.1 Hz, 1H), 6.25 (d, *J* = 13.8 Hz, 1H), 3.63 (s, 3H), 2.86 (s, 6H), 2.49 (s, 2H), 2.36 (s, 1H), 2.23 (s, 1H), 1.73 (s, 1H), 1.60 (s, 1H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 159.37 (s), 157.66 (s), 145.98 (s), 137.97 (s), 133.98 (s), 133.59 (s), 132.66 (s), 132.07 (s), 129.19 (d, *J* = 8.2 Hz), 126.28 (s), 125.31 (s), 118.26 (s), 116.36 (s), 115.76 (s), 100.03 (s), 43.89 (s), 36.91 (s), 30.49 (s), 24.36 (s). m/z (TOF LD+): Calcd $[M^+]$ for $C_{32}H_{29}N_2O_3S^+$: 521.1899, found: 521.1888.

Preparation of K-NIR

The synthetic process is similar to that of compound **S-NIR**. It is purified by silica gel flash chromatography using CH_2Cl_2 to

CH_2Cl_2/CH_3OH (200:1 to 20:1) as eluent to afford **K-NIR** (130 mg, yields 55%). ^{1}H NMR (500 MHz, $CDCl_3$:MeOD=2:1) δ 8.58 (d, *J* = 14.3 Hz, 1H), 8.39 (d, *J* = 8.4 Hz, 1H), 8.17 (s, 1H), 8.06 (d, *J* = 7.9 Hz, 1H), 7.91 (t, *J* = 7.8 Hz, 1H), 7.81 (d, *J* = 8.9 Hz, 1H), 7.69 (t, *J* = 7.7 Hz, 1H), 7.55 (d, *J* = 4.8 Hz, 1H), 7.05 (d, *J* = 7.4 Hz, 1H), 6.92 (d, *J* = 14.4 Hz, 1H), 6.77 (d, *J* = 8.9 Hz, 1H), 6.59 (s, 1H), 6.25 (d, *J* = 9.0 Hz, 1H), 4.06 (s, 3H), 2.91 (s, 6H), 2.67 (td, *J* = 15.7, 9.2 Hz, 2H), 2.45 (ddd, *J* = 89.9, 15.3, 8.2 Hz, 2H), 1.94 – 1.67 (m, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 176.54 (s), 164.08 (s), 158.74 (s), 156.43 (s), 154.75 (s), 152.61 (s), 146.58 (s), 143.02 (s), 142.19 (s), 138.16 (s), 137.34 (s), 133.65 (s), 133.19 (s), 132.63 (s), 132.21 (s), 131.67 (s), 130.85 (s), 129.25 (s), 128.93 (s), 125.61 (s), 120.98 (s), 117.80 (s), 114.04 (s), 113.61 (s), 112.47 (s), 100.58 (s), 46.06 (s), 43.68 (s), 28.81 (s), 24.51 (s). m/z (TOF LD+): Calcd $[M^+]^+$ for $C_{34}H_{31}N_2O_3^+$: 515.2335, found: 515.2322.

Preparation of l[5]

It is synthesized according to the literature procedures.[5] 1,4-benzenedimethanol (3 g, 21.71 mmol) is dissolved in 10 mL dry DMF, and NaH (1g) is added in small portions over 10 min at 0°C. The mixture is stirred for 15min. Propargyl bromide (2.32 g 19.54 mmol) is added and the mixture is stirred for 1h at 25°C. The mixture is quenched with water, extracted with CH_2Cl_2, and purified by flash column chromatography with CH_2Cl_2 to get 4-[prop-2-ynyloxymethy]-benzyl alcohol (2.68g, 15.21mmol, 70%).

Preparation of m [29]

It is synthesized according to the literature procedures.[29] 6-chloro-guanine (0.5g , 5.9 mmol)) is dissolved in 15 ml DMF at 50°C. After cooling to room temperature, 1-methyl-pyrrolidin (1.4 ml, 13.2 mmol) is added and the reaction mixture is stirred for 24 h. Then 2 ml of acetone is added to complete precipitation. The solid is filtered, washed with ether and dried in vacuo, yielding 0.55 g product (2.16 mmol, 71%).

Preparation of PYBG

4-[2-Propynylmethoxy]-benzyl alcohol (662 mg, 7.33 mmol) is dissolved in 3 mL dry DMF, and NaH (65 mg, 2.83 mmol) is added in small portions over 10 min at 0°C. And then the mixture is stirred for 15 min at 0°C. 1-(2-Amino-7H-purin-6-yl)-1-methyl-pyrrolidinium chloride (**m**, 300 mg, 1.18 mmol) and 4-dimethylaminopyridine (30 mg, 0.25 mmol) are added and the mixture is stirred for 1h at 25°C. The mixture is quenched with 0.5 mL water and purified by flash column chromatography (gradient: CH_2Cl_2/CH_3OH 40:1→15:1) to yield 255 mg **n** (0.83 mmol, 72%). mp 111.1-112.9°C. ^{1}H NMR (400 MHz, DMSO) δ 11.60 (s, 1H), 6.99 (s, 1H), 6.64 (d, J = 7.8 Hz, 2H), 6.50 (d, J = 7.8 Hz, 2H), 5.45 (s, 2H), 4.63 (s, 2H), 3.68 (s, 2H), 3.33 (s, 2H), 2.65 (s, 1H). ^{13}C NMR (100 MHz, DMSO) δ 159.6 (s) , 137.4 (s), 136.2 (s) , 128.4 (s) , 127.8 (s) , 80.1 (s) , 77.5 (s), 70.4 (s) , 66.4 (s) , 56.9 (s) . m/z (FTMS+p ESI): Calcd $[M+H]^+$ for $C_{16}H_{16}N_5O_2$: 310.1299, found : 310.1295.

Preparation of r

r is synthesized according to the literature procedures.[40]
4-Bromophthalic anhydride (q, 12 g, 53 mmol) is added to dry 120 mL toluene which contains 3-dimethylamino phenol (8.0 g, 58 mmol). The mixture is heated at 85 °C for 6 hours, then the temperature is elevated to 120 °C and kept for 5 hours. The solvent is evaporated in vacuo and the resident is dissolved in 100 mL methanol. Thionyl chloride (6 mL) was added in drops within 30 min in ice baths. Then the mixture is heated under reflux for 6 hours. After the methanol is removed, the reaction mixture is washed with a large amount of water for three times. The aqueous solution is extracted with dichloromethane (50 mL×3). The combined organic extracts are dried over $MgSO_4$ and the solvent is evaporated. The resident is purified by silica gel column chromatography using dichloromethane and n-hexane (1.5:1) as eluent to obtain pale-yellow oil. The oil is recrystalized in methanol to obtain the solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.20 (d, J = 1.9 Hz, 1H), 7.94 (d, J = 8.4 Hz, 1H), 7.76 (dd, J = 8.1, 2.0 Hz, 1H), 7.69 (dd, J = 8.4, 2.0 Hz, 1H), 7.55 (d, J = 1.9 Hz, 1H), 7.29 (d, J = 2.7 Hz, 1H), 6.90 (dd, J = 9.1, 1.3 Hz, 2H), 6.19 (d, J = 2.4 Hz, 2H), 6.11 (ddd, J = 9.3, 7.1, 2.5 Hz, 2H), 3.75 (s, 3H), 3.74 (s, 2H), 3.07 (s, 4H), 3.07 (s, 6H).^{13}C NMR (100 MHz, $CDCl_3$) δ 197.59 (S), 196.80 (S), 165.69 (S), 165.18 (S), 165.11 (S), 156.07 (S), 156.03 (S), 142.25 (S), 139.38 (S), 135.06 (S), 134.02 (S), 133.99 (S), 133.21 (S), 132.37 (S), 131.83 (S), 130.90 (S), 130.66 (S), 129.47 (S), 127.58 (S), 127.18 (S), 123.25 (S), 110.15 (S), 110.00 (S), 104.16 (S), 104.06 (S), 97.81 (S), 52.63 (S), 52.48 (S), 39.99 (S).The solid (1.5 g, 4.0 mmol) is dissolved in methanol (20 mL), 10% aq. NaOH (0.2 g, 5 mmol) and refluxed for 3 h. The methanol is evaporated. Water (30 mL) is added and the solution is acidified to pH 1.5 with 5% aq. H_2SO_4. The pale-yellow precipitate is filtered, washed with water and dried in vacuo to give **r**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.24 (d, J = 1.9 Hz, 1H), 7.97 (d, J = 8.4 Hz, 1H), 7.78 (dd, J = 8.1, 2.0 Hz, 1H), 7.69 (dd, J = 8.4, 2.0 Hz, 1H), 7.53 (d, J = 1.9 Hz, 1H), 7.26 (d, J = 8.1 Hz, 1H), 6.89 (dd, J = 9.1, 1.8 Hz, 2H), 6.20 – 6.16 (m, 2H), 6.15 – 6.07 (m, 2H), 3.07 (s, 5H), 3.07 (s, 6H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 197.29 (s), 196.46 (s), 169.22 (s), 168.58 (s), 165.23 (s), 156.18 (s), 156.15 (s), 142.75 (s), 139.82 (s), 135.67 (s), 134.14 (s), 133.96 (s), 132.59 (s), 132.42 (s), 131.09 (s), 129.63 (s), 128.04 (s), 126.46 (s), 123.24 (s), 110.12 (s), 109.96 (s), 104.27 (s), 104.18 (s), 97.80 (s), 40.00 (s).

Preparation of s[28]

The synthetic process is similar to that of compound **k**.^{1}H NMR (400 MHz, $CDCl_3$:MeOD=2:1) δ 8.45 (d, J = 1.9 Hz, 1H), 8.20 (d, J = 8.4 Hz, 1H), 7.92 (ddd, J = 25.9, 7.4, 5.5 Hz, 2H), 7.40 (d, J = 1.8 Hz, 1H), 7.24 – 7.13 (m, 5H), 6.96 (dd, J = 4.1, 2.3 Hz, 2H), 3.36 (d, J = 3.8 Hz, 12H), 3.13 (dd, J = 17.4, 11.2 Hz, 4H), 2.38 – 2.26 (m, 4H), 2.02 (s, 4H), 1.89 – 1.79 (m, 4H).^{13}C NMR (100 MHz, $CDCl_3$:MeOD=2:1) δ 174.23(s), 169.36(s), 166.42(s), 163.18(s), 161.45(s), 140.13(s), 139.92(s), 138.60(s), 137.55(s), 137.13(s), 136.93(s), 135.24(s), 134.96(s), 133.90(s), 133.67(s), 128.45(s), 125.76(s), 122.38(s), 122.21(s), 121.35(s), 110.78(s), 99.66(s), 44.71(s), 33.42(s), 29.04(s), 25.17(s), 24.75(s).

Preparation of Br-N-NIR

The synthetic process is similar to that of compound **N-NIR**. It is purified by silica gel flash chromatography using CH_2Cl_2 to CH_2Cl_2/CH_3OH (200:1 to 20:1) as eluent to afford **Br-N-NIR** (120 mg, yields 50%). ^{1}H NMR (400 MHz, CD_3OD) δ 8.71 (d, J = 10.8 Hz, 1H), 8.22 (d, J = 7.2 Hz, 1H), 7.75 (d, J = 4.6 Hz, 1H),

7.70 – 7.62 (m, 1H), 7.55 (d, J = 7.0 Hz, 1H), 7.44 (d, J = 7.0 Hz, 1H), 7.36 – 7.21 (m, 3H), 6.77 (d, J = 11.9 Hz, 3H), 6.21 (d, J = 15.3 Hz, 1H), 3.65 (s, 3H), 3.16 (s, 6H), 2.69 (s, 2H), 2.36 (s, 2H), 1.82 (s, 8H). ^{13}C NMR (126 MHz, MeOD) δ 175.15 (s), 164.75 (s), 156.95 (s), 155.57 (s), 144.30 (s), 143.52 (s), 142.21 (s), 135.73 (s), 135.44 (s), 135.01 (s), 132.05 (s), 129.96 (s), 129.32 (s), 126.45 (s), 124.20 (s), 123.41 (s), 122.56 (s), 116.80 (s), 115.03 (s), 113.54 (s), 111.89 (s), 100.55 (s), 97.22 (s), 50.55 (s), 40.77 (s), 31.91 (s), 29.07 (s), 27.98 (s), 21.71 (s). m/z (TOF LD+): Calcd $[M^+]^+$ for $C_{35}H_{34}BrN_2O_3^+$: 609.1747, 611.1727 , found : 609.1732, 611.1727.

Preparation of PYBG-D

A mixture of **Br-N-NIR** (35 mg, 0.058 mmol), PPh_3 (5 mg, 0.02 mmol), **n** (20 mg, 0.064 mmol), $Pd[P[C_6H_5]_3]_4$ (4 mg, 0.004 mmol), CuI (2 mg, 0.01 mmol), and DMF/triethylamine (1.5 mL/0.5 mL) are placed in a 10 mL round bottomed flask with a magnetic stirrer bar under a nitrogen atmosphere. The mixture is heated at 85℃ for 4 h. Then the mixture is purified by column chromatography (gradient:CH_2Cl_2/CH_3OH 40:1→8:1) to yield 35 mg **PYBG-D** (0.058 mmol, 72%). m/z (TOF-LD+): Calcd $[M^+]$ for $C_{51}H_{48}N_7O_5$: 838.3701, found : 838.3717. The resolutions of ^{1}HNMR and ^{13}CNMR spectra are not satisfactory due to the extremely poor solubility of **PYBG-D** in $CDCl_3$, CD_3OD, and DMSO-d_6 *etc*. The total purity of the two regioisomers (ratio ~1:1) measured by HPLC is 94.4%, as recorded in supporting information (Figure S12).

Acknowledgements

This work is supported by National Natural Science Foundation of China (Nos. 21174022, 21376038, 21421005, and 21576040) and National Basic Research Program of China (No. 2013CB733702).

Notes and references

1. Johnson, I., *The Histochemical Journal* **1998,** *30* (3), 123-140.937
2. Kim, H. M.; Cho, B. R., *Acc. Chem. Res.* **2009,** *42* (7), 863-872.1394
3. Ko, S.-K.; Chen, X.; Yoon, J.; Shin, I., *Chem. Soc. Rev.* **2011,** *40* (5), 2120-2130.1396
4. Song, X.; Hu, M.; Wang, C.; Xiao, Y., *RSC Advances* **2016,** *6* (73), 69641-69646.1398
5. Song, X.; Wang, C.; Han, Z.; Xu, Y.; Xiao, Y., *RSC Advances* **2015,** *5* (30), 23646-23649.1028
6. Banala, S.; Maurel, D.; Manley, S.; Johnsson, K., *ACS Chem. Biol.* **2012,** *7* (2), 289-93.25
7. Lukinavi?ius and G.; Umezawa, *Nat. Chem.* **2013,** *5* (2), 132-139.1211
8. Parthasarathy, A. and Ahn, H.-Y.; *ACS Applied Materials & Interfaces* **2010,** *2* (10), 2744-2748.1418
9. Ji, E. and Wu, D.; *Langmuir* **2010,** *26* (18), 14427-14429.1420
10. Lo, W.-S. and Kwok, W.-M.; *Inorg. Chem.* **2011,** *50* (12), 5309-5311.919
11. Yu, H. and Xiao, Y. *J. Am. Chem. Soc.* **2012,** *134* (42), 17486-17489.917
12. Neto, B. A. D. and Correa, J. R.; *RSC Advances* **2013,** *3* (16), 5291-5301.921
13. Zhang, X. and Xiao, Y.; *The Journal of Organic Chemistry* **2013,** *78* (18), 9153-9160.1408
14. Kiyose , K. and Nagano , T., *Chemistry – An Asian Journal* **2008,** *3* (3), 506-515.1410
15. Weissleder, R., A clearer vision for in vivo imaging. *Nat. Biotechnol.* **2001,** *19* (4), 316-317.1412
16. de Silva, A. P. and Gunaratne, H. Q. N.; *Chem. Rev.* **1997,** *97* (5), 1515-1566.1414
17. Nolan, E. M. and Lippard, S. J., *Chem. Rev.* **2008,** *108* (9), 3443-3480.1416
18. Que, E. L. and Domaille, D. W.; *Chem. Rev.* **2008,** *108* (5), 1517-1549.1422
19. Kobayashi, H. and Ogawa, M.; *Chem. Rev.* **2010,** *110* (5), 2620-2640.1424
20. Han, J. and Burgess, K., *Chem. Rev.* **2010,** *110* (5), 2709-2728.1428
21. Benson, R. C. and Kues, H. A., *J. Chem. Eng. Data* **1977,** *22* (4), 379-383.1426
22. Fu, M. and Xiao, Y.; *Chem. Commun.* **2008,** (15), 1780-1782.1432
23. Koide, Y. and Urano, Y.; *ACS Chem. Biol.* **2011,** *6* (6), 600-608.1245
24. Zhang, X. and Xiao, Y., *The Journal of Organic Chemistry* **2012,** *77* (1), 669-673.1438
25. Bellier, Q. and Pégaz, S.; *Org. Lett.* **2011,** *13* (1), 22-25.1440
26. Loudet, A.; Bandichhor, R.; *Org. Lett.* **2008,** *10* (21), 4771-4774.1261
27. Ahn, H.-Y. and Yao, S.; *ACS Applied Materials & Interfaces* **2012,** *4* (6), 2847-2854.1436
28. Yuan, L. and Lin, W.; *J. Am. Chem. Soc.* **2012,** *134* (2), 1200-1211.1233
29. Keppler, A. and Gendreizig, S.; *Nat. Biotechnol.* **2003,** *21* (1), 86-89.1079
30. Keppler, A. and Pick, H.; *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101* (27), 9955-9.6
31. Campos, C. and Kamiya, M.; *Dev. Dyn.* **2011,** *240* (4), 820-827.961
32. Kolmakov, K. and Hebisch, E.; *Chemistry – A European Journal* **2015,** *21* (38), 13344-13356.1215
33. Kim, H. N. and Lee, M. H.; *Chem. Soc. Rev.* **2008,** *37* (8), 1465-1472.1444
34. Hu, H. and Owens, E. A.; *J. Med. Chem.* **2015,** *58* (3), 1228-1243.1446
35. Levitz, A. and Ladani, S. T.; *Dyes and Pigments* **2014,** *105*, 238-249.1448
36. Wang, J. and Cao, W.-F.; *Dyes and Pigments* **2003,** *57* (2), 171-179.1450
37. Peng, X. and Wu, T.; *Angew. Chem. Int. Ed.* **2011,** *50* (18), 4180-4183.1452
38 David Yaron, and Bruce A. Armitage. *J. Am. Chem. Soc.* **2007**, 129 (17), 5710–5718
39 Alexandre Fu¨rstenberg and Eric Vauthey, *J. Am. Chem. Soc.* **2006**, 128 (23), 7661–7669
40. Yu, H. and Xiao, Y.; *Org. Lett.* **2012,** *14* (8), 2014-2017.923