Letter

# Combining Wittig olefination with photo-assisted domino reaction to distinguish 5-formylcytosine from 5-formyluracil

Qian Zhou, Kun Li, Lingling Li, Kang-Kang Yu, Hong Zhang, Lei Shi, Hao Chen, and Xiaoqi Yu

*Anal. Chem.*, **Just Accepted Manuscript** • DOI: 10.1021/acs.analchem.9b02499 • Publication Date (Web): 19 Jul 2019

# Combining Wittig olefination with photo-assisted domino reaction to distinguish 5-formylcytosine from 5-formyluracil

Qian Zhou,[1] Kun Li,[1,*] Ling-Ling Li,[1] Kang-Kang Yu,[1] Hong Zhang,[1] Lei Shi,[1] Hao Chen[2] and Xiao-Qi Yu[1,*]

[1] Laboratory of Green Chemistry and Technology (Ministry of Education), College of Chemistry, Sichuan University, Chengdu 610064, P. R. China

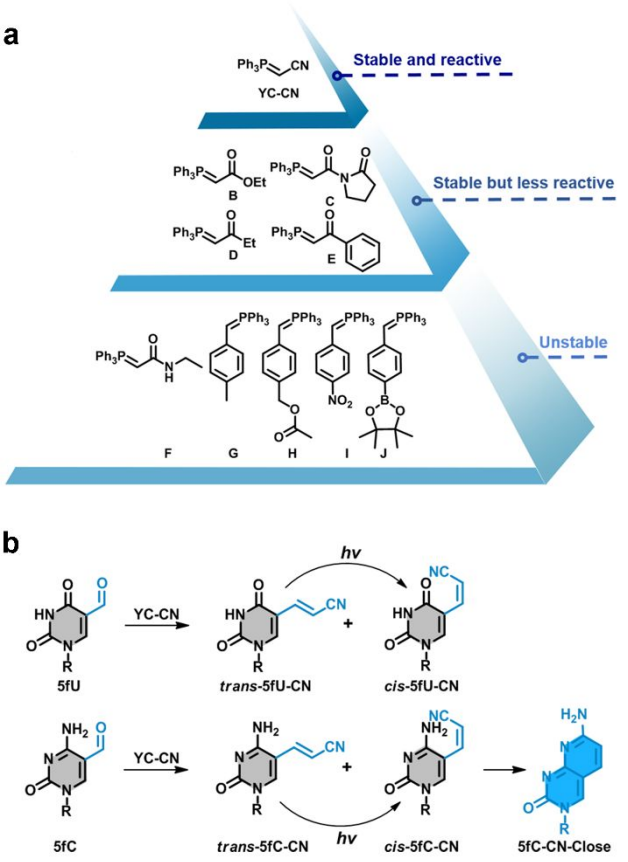
[2] Department of Pharmaceutical Sciences, College of Pharmacy, University of Tennessee Health Science Centre Memphis, Tennessee 38163, the United States

**ABSTRACT:** In view of the important epigenetic functions of 5-formylcytosine (5fC), the development of quantitative detection methods for 5fC is a long-standing issue. In this regard, how to distinguish 5fC from 5-formyluracil to achieve higher accuracy is particularly difficult because the latter one is more reactive. Herein, we reported a phosphorus ylide, **YC-CN**, and introduced a triple domino reaction to fluorescently switch on 5fC with excellent selectivity, which also enable us to quantify 5fC mutations induced by γ-irradiation. This Wittig-initiated covalent labelling strategy provide a novel strategy for qualitative and quantitative detection of 5fC.

Cytosine methylation at C5-positon produces 5-methylcytosine (5mC) is the most common epigenetic modification in mammalian DNA.[1] 5mC can be further oxidized to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) stepwise by Ten-eleven translocation (Tet) family enzymes, leading to an active DNA demethylation pathway. Recent reports revealed that 5fC is not only an important intermediate in this demethylation process, but also acts as a stable epigenetic modification[2,3] associated with gene regulation, cell differentiation and some diseases.[4] In order to gain a deeper insight of the epigenetic functions of 5fC, it is necessary to develop highly robust, accurate and sensitive methods for 5fC.

Owing to inherent sensitivity and selectivity, LC-MS$^n$ has been widely used in the analysis of rare nucleosides. However, quantifications of 5fC is challenging due to its extremely low abundance in vivo[5,6] (for example, 0.002% of all cytosine residues in mouse ESCs), low ionization efficiency as well as the multiple interferences from the sample matrix. In comparison with the time-consuming, expensive and isotope-labelled internal standard needed LC-MS$^n$, fluorescence-based detection methods are more economical, faster, and easier to operate.[7-10] 5fC can be fluorescently switch on by amine,[11] hydroxylamine,[12-15] hydrazine[16,17] and hydrazide[18] derivatives. Nevertheless, such Schiff-based reactions have their inherent drawbacks, for example, hydrazone linkage at neutral pH suffers from poor reactivity and the formed *C=N* is susceptible to hydrolysis over time. More importantly, 5-formyluracil (5fU), an oxidation product of thymine (T), is more reactive than 5fC, resulting in most fluorescent reagents preferentially reacting with 5fU.[18,19] Thereby, it has always been a daunting task that how to make use of the slight structural difference between 5fC and 5fU to design reaction-based fluorescent probes to selectively light on 5fC. In 2016, Höbartner and co-workers first explored an aldol-type condensation toward 5fC/5fU with 2,3,3-trimethylindole derivatives,[20] in which a breakthrough C=C was constructed. Moreover, based on the inherently distinct electronic properties of 5fC and 5fU, the corresponding hemicyanine (Hcy)-like fluorescent nucleosides have distinct excitation/emission maxima, allowing the two to be distinguished. However, the labelling of 5fC and 5fU needs to be done under different conditions and the labelling reactions are incomplete. Similar to Friedländer synthesis to form quinoline derivatives, Yi[21,22] and Zhou[23,24] reported a class of -$CH_2$- reagents that can highly selectively condense with 5fC bearing a characteristic 2-aminobenzaldehyde structure to yield fluorescent intramolecular-cyclized nucleobases, whereas 5fU and abasic sites (AP) without this structure do not interfere. They achieved 5fC-specific recognition from the point of reaction mechanism, however, the reaction time (10~24 h) is a little long. Therefore, developing simple but efficient method to distinguish 5fC from 5fU, and realizing its quantitative detection in short time is highly needed.

Not long ago, for the first time, our group introduced Wittig reaction into the design of reaction-based fluorescent probes and successfully constructed conjugated C=C for ratiometrically detecting 5fU, selectively labelling 5fU-modified DNA and imaging intracellular 5fU.[25] However, **YU**'s reactivity is so poor that it takes 48 h when incubated with 5fU, while 5fC remains intact. We, therefore, put a




**Scheme 1.** (a) Chemical structures of Wittig reagents; (b) Derivatization and fluorescent labelling strategy of aldehyde pyrimidines.

lot of effort into developing more reactive Wittig reagents and trying to achieve 5fC specificity. In this work, a series of phosphorus ylides are synthesized and fortunately, the commercially available **YC-CN** stands out. Both 5fC and 5fU can be completely modified by the star via a Wittig olefination pathway in 0.5~1 h. Due to the *E*-preference of **YC-CN**, the major products are *trans*-**5fC-CN** and *trans*-**5fU-CN,** respectively**,** but they are not fluorescent. To solve this problem, we innovatively utilized UV irradiation to convert *trans*-**5fC-CN** to *cis*-**5fC-CN**, which further undergo intramolecular cyclization with exocyclic 4-$NH_2$ to produce a highly fluorescent **5fC-CN-Close**. The nucleoside-derived fluorophore possesses an extremely high quantum yield, excellent photostability and good water solubility. However, **5fU-CN** cannot undergo such conversion due to the lack of vital 4-$NH_2$ element, which allowed us to fluorescently switch on 5fC with high selectivity (Scheme 1). It is worth noting that the yield of Wittig reaction and photocyclization are almost 100%. Moreover, to the best of our knowledge, the time-consuming in our fluorescent labelling strategy is the shortest. Eventually, based on this entirely new Wittig olefination-involved domino reaction, **YC-CN** was successfully applied into qualitatively and quantitatively detect 5fC mutations caused by γ-irradiation.

As we all known, phosphorus ylides (P-ylides) can be divided into stabilized, semi-stabilized and non-stabilized types depending on the substituents at the methylene carbon. In general, an electron-withdrawing group corresponds to the first class. In order to find a stable and highly reactive Wittig reagent for effectively labelling 5fC, we constructed a library of commonly used P-ylides, compounds **A**-**J**. The varied electron-withdrawing capabilities of the substituents on the methylene carbon-α may lead to different stabilities and reactivities. All of phosphonium salts were readily synthesized in 1-2 steps. However, when treated with a base to prepared the respective phosphoranes, compound **F**-**J** deteriorated totally, only **A**-**E** were alkalinized successfully and obtained in high yields. Next, to determine the Wittig reactivity of compounds **A**-**E** toward formyl pyrimidines, 5fU nucleobase was firstly selected as the substrate for the preliminary model reactions. According to our long-time $^1$H NMR monitoring (Figure S1), compound **C**, **D**, **E** did not react with 5fU at 25 °C in DMSO-$d_6$, even though they coexisted for 24 h. Besides, compound **E** seems to decompose 5 h later. As for compound **B**, a weak characteristic doublet for the generated *CH=CH* between δ=5.5~7.5 ppm slowly appeared at 3 h, and significant formyl-quenching was observed at 24 h. In contrast, after mixed with 5fU, the spectra of compound **A** showed distinct double bond signals within 15 minutes and the Wittig reaction was almost complete within 3 h. Based on the above results, we conclude that compound **A**, renamed as **YC-CN**, is most reactive in our Wittig reagent library and maybe a candidate for 5fC labelling under certain conditions.

To verify the feasibility of our speculation, 5fC was incubated with **YC-CN** in methanol without any additives or catalysts. It is delightful that the reactions proceeded very well and a white powder which proved to be **5fC-CN** adduct was obtained in good yield (Scheme S3). For better derivatization efficiency of 5fC, we further optimized our experimental conditions, such as reaction temperature, reaction time and the molar ratio of **YC-CN**. Temperature was first investigated, our results demonstrated that the maximum peak area of **5fC-CN** can be achieved at 60 °C (Figure S2a). As for derivatization time, Figure S2b tells 0.5 h, or at most 1 h, is sufficient. We then changed the molar ratios of reagent-to-5fC from 2:1 to 150:1 to optimize the dosage of **YC-CN**. As shown in Figure S2c, the peak area of 5fC derivative reaches a plateau at 20 eq. It is worth mentioning that since 5fU harbour a similar 5-formyl group and is more reactive than 5fC, theoretically, under the conditions optimized for 5fC, 5fU can also be effectively labelled. Taken together, after reacting with **YC-CN** at 60 °C for 1 h， more than 99% 5fC was converted to the corresponding nucleoside adducts (Figure S2d), indicating high derivatization efficiency was achieved. It is worth mentioning that due to the introduction of a more hydrophobic acrylonitrile structure, both 5fC and 5fU harvest significantly extended retention times and higher separation resolution after labelling with **YC-CN** (Figure S3). According to the previous reports,[26,27] these significantly altered HPLC properties may improve the detection sensitivity of

5fC/5fU in chemical derivatization-combined LC-MS$^n$ analysis, which we will discuss in detail in the near future.

During the above condition optimization process, we accidentally discovered a minor blue fluorescent product ($\lambda_{ex}/\lambda_{em}$=345/410 nm, Figure S4) in the reaction system of 5fC and **YC-CN**, but not in the reaction system of 5fU and **YC-CN**. Considering this unique phenomenon may enable us to selectively fluorescently switch on 5fC, we carried out detailed structural identification of this blue fluorescent product. $^1$H NMR, $^{13}$C NMR, and HRMS together revealed it is **5fC-CN-Close**. A possible mechanism for the generation of this cyclized nucleoside is intramolecular cycloaddition of exocyclic 4-$NH_2$ to -CN in *cis*-**5fC-CN**, where -$NH_2$ is closer to -CN. The optimized molecular structures and total energies of *trans*-**5fC-CN**, and **5fC-CN-Close** obtained from DFT-B3LYP/6-31G (d,p) calculation were presented in Figure S5. The *cis–trans* energy differences of stable conformers for **5fC-CN** is 3.33 kJ/mol, while the energy of **5fC-CN-Close** is 76.06 kJ/mol lower than that of *cis*-**5fC-CN**, which supports our proposed intramolecular cycloaddition pathway of *cis*-**5fC-CN** to **5fC-CN-Close**. However, **YC-CN** belongs to stabilized ylides and (E)-selectivity is preferred. That is, *trans*-**5fC-CN** is the main product in which -CN is far away from -$NH_2$, making intramolecular cyclization difficult. In this case, more attention should be paid to the mutual transformation between *trans-cis* isomers, and then the low conversion rate of 5fC to fluorescent **5fC-CN-Close** can be solved. It was reported that irradiation is an effective strategy for converting *trans*-isomers to *cis*-isomers. Immediately, we examined the photoisomerization of *trans*-**5fC-CN** in $D_2O$ and monitored by $^1$H NMR. As shown in Figure 1, a rapid conversion of *trans*- to cyclized-product with almost 100% yield was observed upon UV irradiation. Why we didn't monitor the characteristic peaks of inter

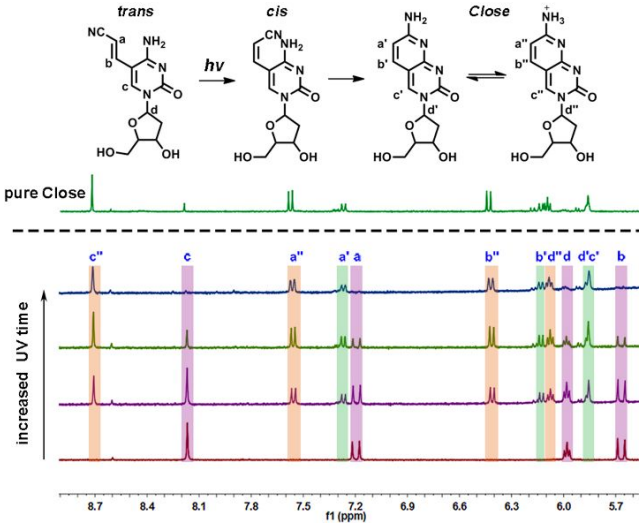



**Figure 1.** $^1$H NMR monitoring of *trans*-**5fC-CN** to **5fC-CN-Close** conversion in $D_2O$ under UV irradiation.

mediate *cis*-**5fC-CN** can be attributed to its short lifetime, and once produced, it will rapidly undergo intramolecular cycloaddition to produce **5fC-CN-Close** in $D_2O$.

In order to optimize the irradiation time, we performed a photocatalytic reaction of 0.1 mM *trans*-**5fC-CN** in PBS with a UV lamp at room temperature. The results are summarized in Figure S6. As the irradiation time prolonged, the fluorescence emission at 410 nm increases sharply during the first 1.5 h and then reach a plateau. Simultaneously, in UV-vis spectra (Figure S7), the characteristic absorption band centred at 273 nm belonged to *trans*-**5fC-CN** dropped gradually and a weak shoulder peak at 350 nm appeared, indicating a light-assisted transformation from *trans*-**5fC-CN** to **5fC-CN-Close**. TLC also showed a complete conversion after 1.5 h irradiation (Figure S8); plus 1 h for Wittig olefination, **YC-CN** take a total time of 2.5 h to fluorescently switch on 5fC, which is still much faster than previous reports. In addition, photobleaching was not observed even after continuous irradiation for 4.5 h, indicating the photostability of **5fC-CN-Close** was excellent. Moreover, its absolute fluorescence quantum yield was measured to be 0.87, even higher than quinine sulphate and fluorescein. Based on the above results, we successfully developed a novel triple domino reaction strategy, including Wittig olefination, photo-assisted *trans*-to-*cis* isomerization and intramolecular cycloaddition, to achieve fast fluorescent labelling of 5fC with high selectivity and almost 100% yield.

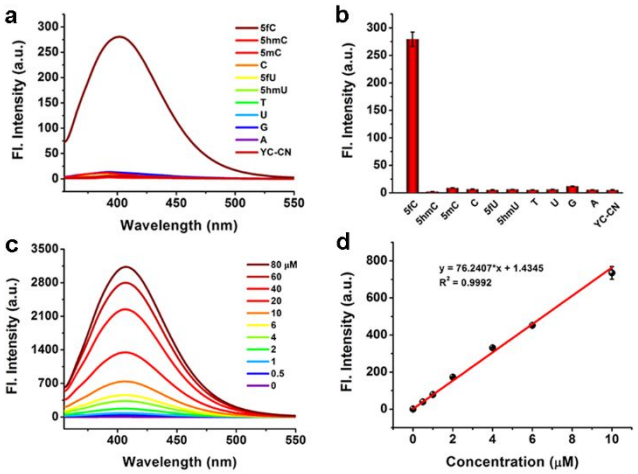



**Figure 2.** (a) Fluorescence spectra and (b) emission intensity of **YC-CN** after incubated with 5fC and other interfering species, including A, G, C, T, U, 5mC, 5hmC, 5hmU and 5fU (5 μM). (c) Fluorescence titration spectra and (d) linear relationship of emission intensity toward 5fC. $\lambda_{ex}$ /$\lambda e_m$=345/410 nm.

To further test the specificity of **YC-CN** for 5fC, we treated canonical deoxynucleosides and their modifications with **YC-CN** under optimized conditions, followed by UV irradiation. The data (Figure 2a and b) shows only 5fC triggered a remarkable fluorescent enhancement of 54-fold at 410 nm ($\lambda_{ex}$ = 345 nm), whereas other potential interferences, even 5fU, induced no obvious spectral changes of **YC-CN**, indicating probe **YC-CN** has superior selectivity for 5fC. Meanwhile, this desirable selectivity also indirectly confirms our proposed

triple domino reaction initiated by Wittig olefination. Subsequently, fluorescence titration experiments were conducted. As depicted in Figure 2c, the emission peak centred at 410 nm belonging to **5fC-CN-Close** gradually increased and the fluorescence intensity exhibited an excellent linear relationship with the concentration range of 0-10 μM (Figure 2d). Additionally, the corresponding detection limit (3σ/slope) was calculated to be 35.7 nM, which indicates the high sensitivity of **YC-CN** toward 5fC.

Recently, Wagner *et al*[28] have reported that 5mC exposed to ionizing radiation would produce a variety of oxidation analogs, including 5fC. Encouraged by the above findings, we continue to investigate whether **YC-CN** is capable of detecting 5fC mutations caused by γ-rays. Typically, an aerated aqueous solution of 5mC (10 mM), prior to being bubbled with oxygen for 1 h, was exposed to a $^{60}$Co γ-source at a dose rate of 16.7 Gy/min. Aliquots were taken from this solution at different time points to obtain testing samples with a series of irradiation doses. We next performed our domino reaction derivation with **YC-CN**, scanned their fluorescence spectra, and calculated the 5fC lesions using the emission intensity at 410 nm. As can be seen from Figure 3b, the yield of 5fC was proportional to the dose of γ-rays with a formation rate of (0.20 5fC/$10^6$ nucleosides)/Gy. This data is in reasonable agreement with (0.45 5fC/$10^6$ nucleosides)/Gy reported by Wagner *et al.*, who delivered the γ-ray at a completely different dose rate of 1.2 Gy/min. Taken together, the above results demonstrate that our independently developed fluorescence detection method based on the Wittig-initiated domino reactions is reliable for quantifying 5fC.

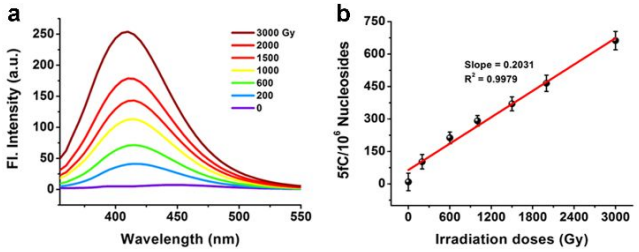



**Figure 3.** (a) Fluorescence spectra of γ-irradiated 5mC after treatment with **YC-CN**; (b) Quantification of 5fC mutations in γ-irradiated 5mC at different irradiation doses (0–3000 Gy). $\lambda_{ex}$ /$\lambda_{em}$=345/410 nm.

In summary, after screening from a wide variety of P-ylides, for the first time, we developed an effectively derivatization method for 5fC using a commercially available Wittig reagent **YC-CN**. Then, through detailed analysis and identification of the obtained products, we propose a photo-assisted triple domino reaction strategy to achieve rapid qualitative and quantitative detection of 5fC. Specifically, upon UV irradiation, the resulting Wittig olefin, nonfluorescent *trans*-**5fC-CN**, further undergoes *trans*-to-*cis* isomerization and intramolecular cycloaddition of exocyclic 4-$NH_2$ to -CN to afford highly fluorescent nucleoside **5fC-CN-Close**, making it feasible for us to distinguish 5fC from 5fU and accurately quantify 5fC lesions caused by γ-irradiation with significant selectivity and sensitivity. The high reactivity of **YC-CN** and the efficient photocatalytic strategy make it the fastest method for 5fC selective-labelling so far. What's more, our study provides a new idea for distinguishing 5fC and 5fU from the perspective of reaction mechanism and also broadened the application of Wittig reaction in the field of chemical biology.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.
Experimental details, NMR, ESI-MS and additional figures (PDF)

## AUTHOR INFORMATION

### Corresponding Author

*E-mail: kli@scu.edu.cn.
*E-mail: xqyu@scu.edu.cn.

### Notes

The authors declare no competing financial interests.

## ACKNOWLEDGMENT

This work was financially supported by National Natural Science Foundation of China (No.: 21572147 and 21877082). We also thank the Comprehensive Training Platform of Specialized Laboratory, College of Chemistry, Sichuan University for sample analysis.

## REFERENCES

(1) Wu, X.; Zhang, Y. *Nat. Rev. Genet.* **2017,** *18*, 517.
(2) Bachman, M.; Uribe-Lewis, S.; Yang, X.; Burgess, H. E.; Iurlaro, M.; Reik, W.; Murrell, A.; Balasubramanian, S. *Nat. Chem. Biol.* **2015,** *11*, 555.
(3) Su, M.; Kirchner, A.; Stazzoni, S.; Mueller, M.; Wagner, M.; Schroeder, A.; Carell, T. *Angew. Chem. Int. Ed.* **2016,** *55*, 11797.
(4) Berney, M.; McGouran, J. F. *Nat. Rev. Chem.* **2018,** *2*, 332.
(5) Pfaffeneder, T.; Hackner, B.; Truss, M.; Muenzel, M.; Mueller, M.; Deiml, C. A.; Hagemeier, C.; Carell, T. *Angew. Chem. Int. Ed.* **2011,** *50*, 7008.
(6) Wagner, M.; Steinbacher, J.; Kraus, T. F. J.; Michalakis, S.; Hackner, B.; Pfaffeneder, T.; Perera, A.; Mueller, M.; Giese, A.; Kretzschmar, H. A.; Carell, T. *Angew. Chem. Int. Ed.* **2015,** *54*, 12511.
(7) Shi, L.; Li, K.; Li, L.-L.; Chen, S. Y.; Li, M.-Y.; Zhou, Q.; Wang, N.; Yu, X.-Q. *Chem. Sci.* **2018,** *9*, 8969.
(8) Pan, S.-L.; Li, K.; Li, L.-L.; Li, M.-Y.; Shi, L.; Liu, Y.-H.; Yu, X.-Q. *Chem. Commun.* **2018,** *54*, 4955.
(9) Li, K.; Hou, J.-T.; Yang, J.; Yu, X.-Q. *Chem. Commun.* **2017,** *53*, 5539.
(10) Li, M.; Cui, P.; Li, K.; Feng, J.; Zou, M.; Yu, X. *Chin. Chem. Lett.* **2018,** *29*, 992.
(11) Hu, J. L.; Xing, X. W.; Xu, X. W.; Wu, F.; Guo, P.; Yan, S. Y.; Xu, Z. H.; Xu, J. H.; Weng, X. C.; Zhou, X. *Chem. Eur. J.* **2013,** *19*, 5836.
(12) Wang, S. R.; Song, Y. Y.; Wei, L.; Liu, C. X.; Fu, B. S.; Wang, J. Q.; Yang, X. R.; Liu, Y. N.; Liu, S. M.; Tian, T.; Zhou, X. *J. Am. Chem. Soc.* **2017,** *139*, 16903.
(13) Guo, P.; Yan, S.; Hu, J.; Xing, X.; Wang, C.; Xu, X.; Qiu, X.; Ma, W.; Lu, C.; Weng, X.; Zhou, X. *Org. Lett.* **2013,** *15*, 3266.
(14) Hong, T.; Wang, T.; Guo, P.; Xing, X.; Ding, F.; Chen, Y.; Wu, J.; Ma, J.; Wu, F.; Zhou, X. *Anal. Chem.* **2013,** *85*, 10797.
(15) Wang, S.-R.; Wang, J.-Q.; Fu, B.-S.; Chen, K.; Xiong, W.; Wei, L.; Qing, G.; Tian, T.; Zhou, X. *J. Am. Chem. Soc.* **2018,** *140*, 15842.
(16) Xu, L.; Chen, Y. C.; Chong, J.; Fin, A.; McCoy, L. S.; Xu, J.; Zhang, C.; Wang, D. *Angew. Chem. Int. Ed.* **2014,** *53*, 11223.

(17) Xu, L.; Chen, Y. C.; Nakajima, S.; Chong, J.; Wang, L.; Lan, L.; Zhang, C.; Wang, D. *Chem. Sci.* **2014,** *5*, 567.
(18) Hardisty, R. E.; Kawasaki, F.; Sahakyan, A. B.; Balasubramanian, S. *J. Am. Chem. Soc.* **2015,** *137*, 9270.
(19) Liu, C.; Luo, X.; Chen, Y.; Wu, F.; Yang, W.; Wang, Y.; Zhang, X.; Zou, G.; Zhou, X. *Anal. Chem.* **2018,** *90*, 14616.
(20) Samanta, B.; Seikowski, J.; Hobartner, C. *Angew. Chem. Int. Ed.* **2016,** *55*, 1912.
(21) Xia, B.; Han, D.; Lu, X.; Sun, Z.; Zhou, A.; Yin, Q.; Zeng, H.; Liu, M.; Jiang, X.; Xie, W.; He, C.; Yi, C. *Nat. Methods* **2015,** *12*, 1047.
(22) Zhu, C. X.; Gao, Y.; Guo, H. S.; Xia, B.; Song, J. H.; Wu, X. L.; Zeng, H.; Kee, K.; Tang, F. C.; Yi, C. Q. *Cell Stem Cell* **2017**, 20, 720.
(23) Liu, C.; Wang, Y.; Yang, W.; Wu, F.; Zeng, W.; Chen, Z.; Huang, J.; Zou, G.; Zhang, X.; Wang, S.; Weng, X.; Wu, Z.; Zhou, Y.; Zhou, X. *Chem. Sci.* **2017,** *8*, 7443.
(24) Wang, Y.; Liu, C.; Zhang, X.; Yang, W.; Wu, F.; Zou, G.; Weng, X.; Zhou, X. *Chem. Sci.* **2018,** *9*, 3723.
(25) Zhou, Q.; Li, K.; Liu, Y.-H.; Li, L.-L.; Yu, K.-K.; Zhang, H.; Yu, X.-Q. *Chem. Commun.* **2018,** *54*, 13722.
(26) Jiang, H.-P.; Liu, T.; Guo, N.; Yu, L.; Yuan, B.-F.; Feng, Y.-Q. *Anal. Chim. Acta* **2017,** *981*, 1.
(27) Tang, Y.; Xiong, J.; Jiang, H. P.; Zheng, S. J.; Feng, Y. Q.; Yuan, B. F. *Anal. Chem.* **2014,** *86*, 7764.
(28) Madugundu, G. S.; Cadet, J.; Wagner, J. R. *Nucleic Acids Res.* **2014,** *42*, 7450.

For Table of Contents Only