Electronic Supplementary Information (ESI)

A New Chemical Approach for Proximity Labelling of Chromatin-associated RNAs and Proteins with Visible Light Irradiation

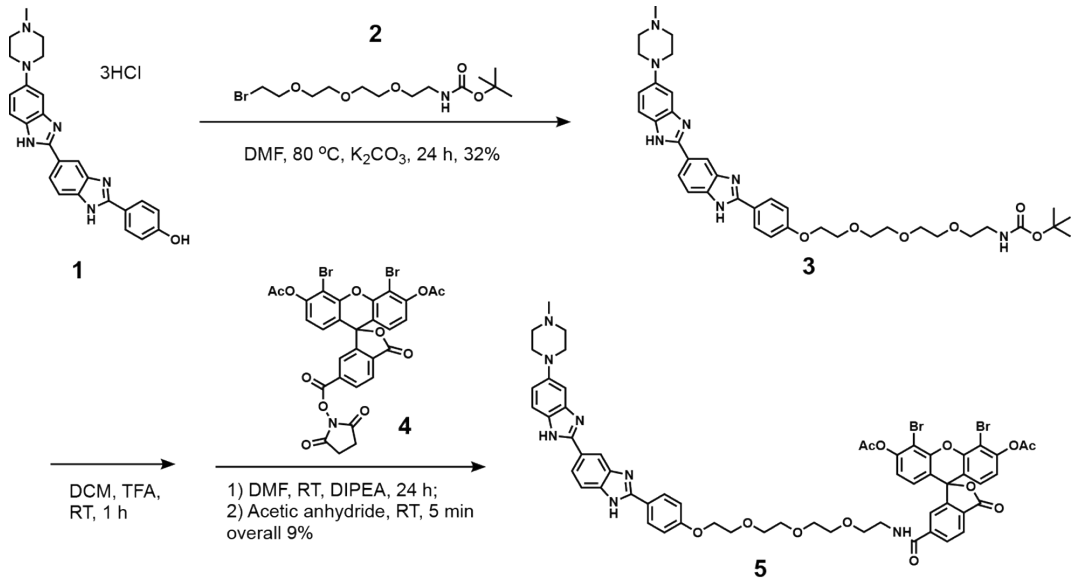
Lan Li,[‡a] Jiying Liang,[‡a] Hao Luo,[a] K. Ming Tam,[a] Edmund C. M. Tse[*a] and Ying Li[*a]

[a] Department of Chemistry, The University of Hong Kong, Pokfulam Road, HK SAR
[‡] These authors contributed equally.
Corresponding author email: ecmtse@hku.hk, yingli0e@hku.hk

1. General Procedures

All reagents were purchased from commercial suppliers and were of analytical grade and used without further purification unless otherwise noted. Reaction progress was monitored by thin-layer chromatography on EMD 60 F254 plates, visualized with UV light and $KMnO_4$. Compounds were purified via flash column chromatography using Sorbent Technologies 60 Å 230 x 400 mesh silica gel. Anhydrous solvents acetonitrile (MeCN), dichloromethane (DCM), methanol (MeOH), tetrahydrofuran (THF), and dimethylformamide (DMF) were degassed and dried over molecular sieves (3 Å). Trifluoroacetic acid (TFA) and *N*,*N*-diisopropylethylamine (DIPEA) were used directly after purchase. All reaction vessels were flame dried prior to use. NMR spectra were acquired with Bruker Advanced spectrometers. All spectra were acquired at 298 K. ^{1}H-NMR spectra were acquired at 500 MHz. ^{13}C-NMR spectra were acquired at 125 MHz. Chemical shifts are reported in ppm relative to residual non-deuterated NMR solvent, and coupling constants (J) are provided in Hz. All NMR spectra were analyzed using iNMR reader software. Low-resolution (LR) and high-resolution (HR) electrospray ionization (ESI) mass spectra (MS) were collected on a micromass 70S-250 spectrometer or an ABI/Sciex QStar Mass Spectrometer. Absolute emission quantum yields of solutions were recorded on a Hamamatsu Quantaurus-QY Absolute PL quantum yields measurement system C11347.

2. Synthesis



Compounds **2**[1] and **4**[2] were prepared according to reported procedures.

Compound **3**.
To 2.5 mL of dry DMF was added **1** (100 mg, 187 μmol), **2** (73 mg, 206 μmol, 1.1 eq), and potassium carbonate (129 mg, 935 μmol, 5 eq). The mixture was stirred for 24 h at 80 ºC. The reaction mixture was diluted with DCM and filtered. The solvent was removed via a rotary evaporator and the crude was purified by gradient column chromatography [NH_3(aq):MeOH:DCM = 0:1:49 to 1:9:90 to 1:19:80] to afford 42 mg (32%) of the product as a pale yellow solid. ^{1}H NMR (CD_3OD) δ 8.266 (s, 1H), 8.07 (d, *J* = *9*, 2H), 7.97 (m, 2H), 7.71 (d, *J* = *8*, 1H), 7.51 (d, *J* = *9*, 1H), 7.23 (d, *J* = *8*, 1H), 7.15 (s, 1H), 7.12 (d, *J* = 9, 2H), 7.05 (dd, J = 9&2.5, 1H), 6.95 (m, 1H), 4.23 (m, 2H), 3.88 (m, 2H), 3.73 (m, 2H), 3.68 (m, 2H), 3.65 (m, 2H), 3.60 (m, 2H), 3.50 (m, 2H), 3.24-3.20 (m, 6H), 2.69 (m, 4H), 2.39 (s, 3H), 1.45 (s, 9H). ^{13}C NMR (CD_3OD) δ 162.21, 158.36, 155.12, 153.63, 149.53, 132.52, 129.51, 125.63, 123.05, 122.41, 116.27, 116.05, 115.79, 71.72, 71.58, 71.56, 71.47, 71.24, 71.01, 70.66, 68.69, 56.14, 51.71, 49.85, 46.09, 45.66, 41.28, 28.77. LRMS Calcd for $C_{38}H_{50}N_7O_6$ $[M+H]^+$ 700.4, found 700.4 $[M+H]^+$.

Compound **5**.
To 0.5 mL of dry DCM was added **3** (20 mg, 29 μmol) and 0.5 mL of anhydrous TFA. The reaction mixture was stirred at room temperature for 1 h. DCM and TFA were removed by a rotary evaporator. The residue was resuspended in DCM and the solvent was removed by a rotary evaporator. The process was repeated for additional three times. The crude was resuspended in 0.9 mL of dry DMF and 0.1 mL of DIPEA. The solution was mixed with **4** (20 mg, 28 μmol). The mixture was stirred at room temperature for 24 h and was added 0.2 mL of acetic anhydride. The solution was stirred for 5 min at 80 ºC. The mixture was concentrated down via a rotary evaporator and dilute with DCM. This procedure was repeated 5 times. The crude product was dissolved in ethyl acetate and washed with saturated aqueous ammonium chloride solution. The organic layer was dried over

anhydrous $MgSO_4$, filtered and concentrated down via a rotary evaporator. The brown viscous liquid was purified by gradient column chromatography (Hx:DCM:MeOH = 90:10:0 to 0:100:0 to 0:90:10) to afford 3 mg (9% over three steps) of the product as a yellow solid. ^{1}H NMR ($CDCl_3$:CD_3OD=1:1) δ 8.43 (s, 1H), 8.20 (d, *J* = *8*, 1H), 8.09 (m, 4H), 8.00 (d, *J* = *7.5*, 1H), 7.75 (m, 3H), 7.59 (d, *J* = *8.5*, 1H), 7.22 (s, 1H), 7.05 (d, *J* = 10.5, 2H), 6.99 (d, *J* = *9*, 2H), 6.95 (d, *J* = *8.5*, 2H), 6.85 (d, *J* = *9*, 2H), 4.11 (m, 2H), 3.8-3.3 (m, 14H), 2.89 (s, 4H), 2.40 (s, 10H). ^{13}C NMR ($CDCl_3$:CD_3OD=1:1) δ 168.28, 168.21, 165.79, 162.31, 162.01, 161.74, 161.46, 160.98, 152.62, 150.66, 150.57, 148.78, 147.94, 147.69, 141.64, 130.14, 128.80, 127.42, 127.13, 125.57, 122.84, 121.77, 121.13, 120.26, 119.68, 117.92, 117.44, 116.81, 115.59, 115.17, 115.10, 114.90, 113.27, 106.53, 101.18, 81.58, 70.61, 70.50, 70.27, 69.96, 69.47, 67.45, 53.53, 47.81, 46.00, 43.31, 42.24, 40.19, 31.89, 29.65, 29.32, 20.62, 14.02, 11.04, 8.38. HRMS Calcd for $C_{58}H_{54}Br_2N_7O_{12}$ $[M+H]^+$ 1200.2186, found 1200.2170 $[M+H]^+$.

3. Biochemical Methods

Cell lines and culture conditions

HEK293T cells were cultured in DMEM supplemented with 10% FBS, 1% penicillin, and streptomycin and grown at 37 ºC, 5% CO_2.

Live cell fluorescence imaging

Cells were seeded in poly-D-lysine (50 μg/ml) treated ibidi 8-well plates to further minimize HEK293T cells peeling off. Once cells reached 80% confluence, cells were incubated with 10 μM HoeDBF for 60 min in HBSS and washed twice with HBSS for 20 min each. Propargyl amine (PA, 2 mM) was added right before live cell imaging to represent labeling conditions.

Light source for irradiation

Low-power blue light: 14 W, ~480 nm. High-power green light: 35 W, ~530 nm. Cells were sandwiched between two green lamps to afford 70 W during irradiation.

Study biomolecule labeling via fluorescence confocal imaging after CuAAC.

The cell culture dishes and glass cover slips were coated with poly-D-lysine (50 μg/mL) for 6 h at 37 ºC and washed three times with autoclaved water to remove the excessive amount of poly-D-lysine. HEK cells were seeded at equal density and grew to 70% confluence. Cells were incubated with HoeDBF (10 μM in HBSS from 5 mM stock in DMSO) for 60 min at 37 ºC, 5% CO_2 and replaced with fresh full media for 20 min twice. Cells were then incubated with 2 mM PA in fresh HBSS media for 3 min at 37 ºC, 5% CO_2, irradiated for 3 min with high-power green light at room temperature. Cells were washed twice with DPBS, fixed and permeabilized for 30 min at room temperature with 3.7% paraformaldehyde and 0.1% Triton-X100. Cells were then washed three times (7 min/each) on orbital shaker with DPBS, blocked with BSA (1 mg/mL in DPBS, 0.45% NaCl and 0.025% NaN_3) for 30 min at room temperature, washed twice for 5 min with DPBS, and incubated with 50 μL of click solution (1 mM $CuSO_4$, 4 mM THPTA ligand, 10 mM sodium ascorbate, and 15 μM Cy5 azide) for 1 h at 37 ºC in the dark. Cells were washed three times for 5 min on an orbital shaker with DPBS-0.1% Triton-X100 and once with DPBS. Cells were stained with Hoechst 33342 (1:2000, Trihydrochloride, Trihydrate - 10 mg/mL solution in water from Thermo Fisher Scientific) for 3 min, washed three times with DPBS for 5 min and mounted using VectaShield (Vector Labs). Slides were imaged via fluorescence confocal microscopy using a 63x oil immersion objective on a Leica microscope.

Photo-oxidation labeling *in situ* for RNA or protein analysis

HEK293T cells grew to 95% confluence and media was removed. Cells were incubated with HoeDBF (10 μM in HBSS from 5 mM stock in DMSO) for 60 min at 37 ºC, 5% CO_2 and replaced with fresh full media for 20 min twice. Cells were then incubated with 2 mM PA in fresh HBSS media for 3 min at 37 ºC, 5% CO_2, and irradiated for 3 min with high-power green light at room temperature. Cells were washed with HBSS and collected into 1.5 mL eppendorf tube with HBSS. Cell pellets were centrifuged down at room temperature at 300 g for 4 min and the supernatant was decanted.

Protein isolation and biotinylation via CuAAC.

Cell pellets (50 μL) were resuspended in 800 μL of RIPA buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate). Proteinase inhibitor cocktail (8 μL) and 4 μL RQ1 RNase-free DNAse (4 U/reaction) were added and the

mixture was incubated at 37°C for 5 min at 1000 rpm. The cell lysate was sonicated at 20% amplitude for 1 s on and 1 s off for total sonication of 4 s. BCA assay was performed to measure total protein concentration and adjust samples to 2 μg/μL.

Click reactions were prepared using 300 μg of protein, 10 mM biotin azide to a final concentration of 1 mM, 46 mM THPTA to a final concentration of 4 mM, fresh 60 mM sodium ascorbate to a final concentration of 10 mM, and 24 mM $CuSO_4$ to a final concentration of 1 mM. The reactions were incubated in dark at 37 ºC for 1 h. Protein was purified by Zepa column according to the manufacturer's procedure. BCA assay was performed to measure total protein concentration and adjust samples to 1 μg/μL.

Protein in-gel fluorescence after cell fractionation.

Cell pellets (50 μL) were resuspended in 400 μL of cell lysis buffer (50 mM Tris buffer, pH 7, 150 mM NaCl, 0.1% Triton X-100, supplemented with 4 μL protease inhibitor) and incubated on ice for 15 min. Nucleus was collected by centrifuging at room temperature for 10 min with 1300 g. The supernatant was saved as cytoplasm fraction. The nucleus pellet was washed once with DPBS at room temperature for 5 min with 1300 g and resuspended in 400 μL RIPA buffer supplemented with 4 μL protease inhibitor. The suspension was incubated on ice for 15 min and sonicated at 20% amplitude for 1 s on and 1 s off for total sonication of 4 s. The mixture was added 4 μL RQ1 RNase-free DNAse (4 U/reaction) and incubated at 37 ºC for 5 min with 1000 rpm. BCA assay was performed to measure total protein concentration and adjust samples to 2 μg/μL. Similar CuAAC reactions were carried out except that Cy5 azide (1 mM stock and 0.1 mM final conc.) was used instead of biotin azide. Without further purification, equal amount of loading buffer was added to the purified protein and heated at 95 ºC for 5 min. Samples were loaded (10 μg protein) onto 4-12% SDS PAGE gel, run at 200 V for 25-30 min (until the dye runs off) and imaged on a ChemiDoc MP imaging system (Bio-Rad). The recovered gel was incubated with PageBlue™ Protein Staining Solution for 30 min and washed with deionized water several times for loading control.

Protein enrichment and western blot.

Pierce™ Streptavidin Magnetic Beads (20 μL for 600 μg biotinylated protein) were washed with RIPA buffer twice and incubated with protein samples at room temperature for 1 h with end-to-end rotation. Samples were put on magnetic stand and supernatant was removed to a new tube labeled as flow through. The magnetic beads were washed with high salt wash buffer (4 M NaCl + 0.2% Tween 20 + 100 mM Tris pH 7.0 + 10 mM EDTA) four times, 0.1% SDS buffer once and DPBS buffer twice. The enriched protein was eluded off beads with elusion buffer (50 mM Tris pH 7.0, 10 mM EDTA, 1% SDS, 15 mM D-Biotin) at 95 ºC for 5 min. For western blot analysis, input and enriched proteins were boiled at 95 ºC for 5 min in 1X protein loading buffer. Equal volume of samples was loaded (10 μL protein) onto a 4-12% SDS PAGE gel (Bio-Rad, catalog# 456-1094) and run at 200 V for appropriate time. Protein was transferred to nitrocellulose membrane (Bio-Rad, catalog# 1620112), using a Trans-Blot Turbo Transfer System (Bio-Rad). The membrane was

blocked overnight with blocking buffer (0.5% w/v non-fat dry milk, 0.1% (v/v) Tween 20 in DPBS (DPBST)) and then washed twice (5 min/each) on an orbital shaker with DPBST buffer. The membrane was incubated with the primary antibody in blocking buffer for 3 h at room temperature on an orbital shaker and then washed twice (5 min/each) on orbital shaker with DPBST buffer. The membrane was incubated with secondary antibody (anti-rabbit IgG conjugated with Horseradish peroxidase (HRP)) in blocking buffer for 1 h at RT on an orbital shaker and washed twice for 5 min with DPBST buffer. The membrane was incubated in Pierce ECL Western Blotting Substrate (Thermo Scientific, catalog# 32106) and imaged on a ChemiDoc MP imaging system (Bio-Rad).

RNA isolation and biotinylation via CuAAC.

After *in situ* photo-oxidation labeling, total cellular RNA was harvested using Trizol (2 mL for 50 μL cell pellets). Click reactions were prepared using 30 μg of total RNA, 10 mM biotin azide to a final concentration of 1 mM, 46 mM THPTA to a final concentration of 4 mM, fresh 60 mM sodium ascorbate to a final concentration of 10 mM, and 24 mM $CuSO_4$ to a final concentration of 1 mM. The reactions were incubated on an orbital shaker (500 rpm) at room temperature for 30 min. The reactions were purified by Qiagen RNAeasy Mini Purification kit according to the manufacturer's procedure, and RNA was eluted in 20 μL of nuclease-free water.

HRP-streptavidin dot blotting

Equal volume of purified INPUT (equal conc.) and ENRICHED RNA was loaded onto Hybond-N+ membrane (GE Healthcare, Genesee Scientific, catalog# 83-378) and UV-crosslinked to the membrane (Stratalinker UV crosslinker, 2500 μJ/m^2). The membrane was blocked followed by incubation with Pierce high sensitivity streptavidin-HRP (Thermo Scientific, catalog# 21130), washed twice in a 1:10 solution of blocking buffer for 20 min each and twice in Tris-saline buffer for 5 min each. It was then incubated in Pierce ECL Western Blotting Substrate (Thermo Scientific, catalog# 32106) and imaged on a ChemiDoc MP imaging system (Bio-Rad). The recovered membrane was incubated with methylene blue stain (0.04% w/v, 0.3 M sodium acetate) for 30 min and washed with deionized water several times for loading control.

RT-qPCR for biotinylated RNA post enrichment

Dynabeads™ MyOne™ Streptavidin C1 Beads (10 μL for 20 μg biotinylated RNA) were washed with wash buffer A (100 mM NaOH + 50 mM NaCl), washed with wash buffer B (100 mM NaCl), washed with binding buffer (1 M NaCl + 0.2% Tween 20 + 100 mM Tris pH 7.0 + 10 mM EDTA) and resuspended in same volume of binding buffer as original volume. Washed C1 beads were incubated with RNA samples at room temperature for 1 h with end-to-end rotation. Samples were put on a magnetic stand and the supernatant was removed to a new tube labeled as flow through. The magnetic beads were washed with high salt wash buffer (4 M NaCl + 0.2% Tween 20 + 100 mM Tris pH 7.0 + 10 mM EDTA) four times, 50 ºC high salt wash buffer once and DPBS buffer twice. The enriched RNA was eluded off the beads with elusion buffer (50 mM Tris pH 7.0, 10 mM EDTA, 1% SDS, 15 mM D-Biotin, supplemented with 1 μL protease K, 1 μL RnaseOut, and 1 μL 1 M DTT,

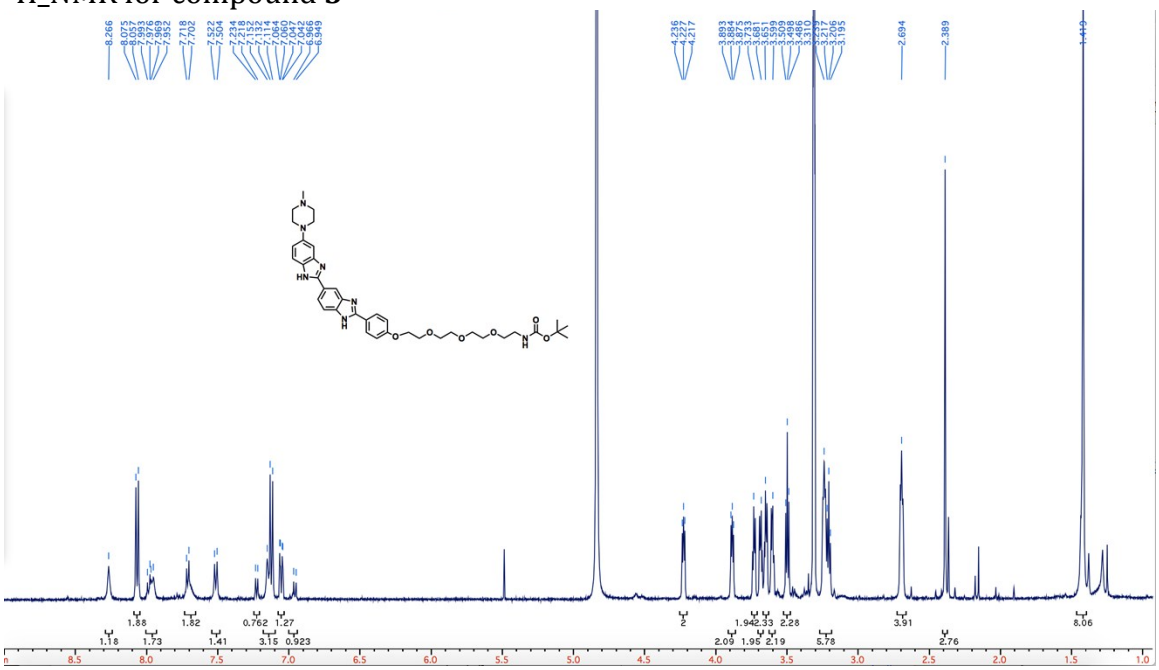
60 μL for each sample) at 42 ºC for 60 min. The RNA was extracted with Trizol (240 μL) and processed accordingly.

Enriched RNA samples were treated with RQ1 RNase-free DNAse (4 U/reaction) at 37 ºC for 15 min, purified by isopropanol precipitation, and resuspended in 8 μL nuclease-free water. 1 μL of biotinylated RNA from each sample was subjected to cDNA synthesis using PrimeScript™ Reverse Transcriptase from Clonetech/Takara (catalog# 2680B) with both random hexamer and oligoT primer in a 10 μL reaction. cDNA was diluted to 20 μL and semi-quantitative SYBR Green PCR was performed using reported primers. Per 12.5 μL semi-quantitative SYBR Green PCR analysis, 1 μL of diluted INPUT and ENRICHED were subjected to RT-qPCR using SYBR® Advantage® qPCR Premix from Clonetech/Takara (catalog# 639676) and Biorad CFX connect™ real time system. The fold of enrichment was calculated against a no-DBF negative control: $2^{((Ct_{ENRICH\text{-}noDBF}-Ct_{INPUT\text{-}noDBF})-(Ct_{ENRICH}-Ct_{INPUT}))}$.

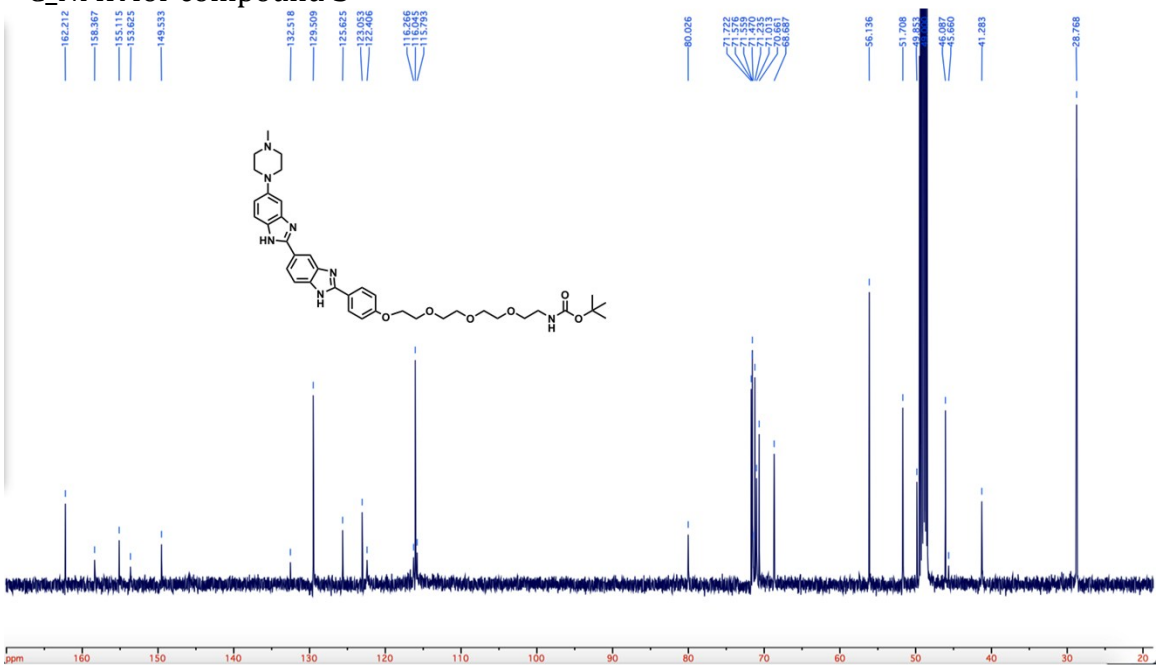
4. Materials Characterization

NMR spectra

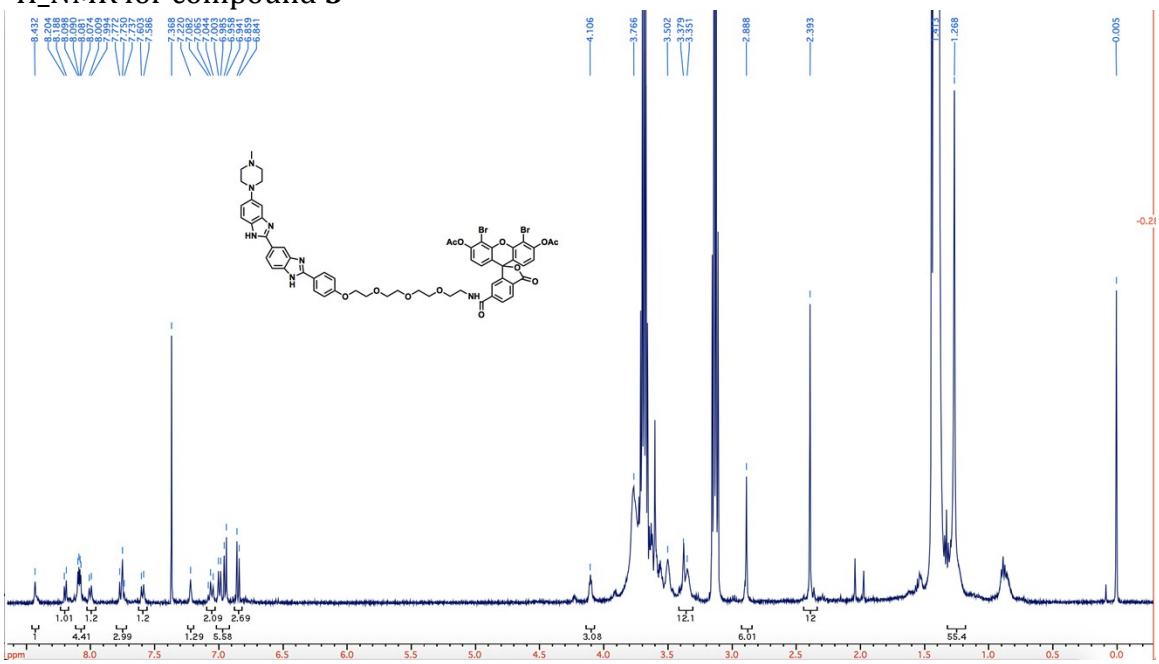
^{1}H_NMR for compound **3**



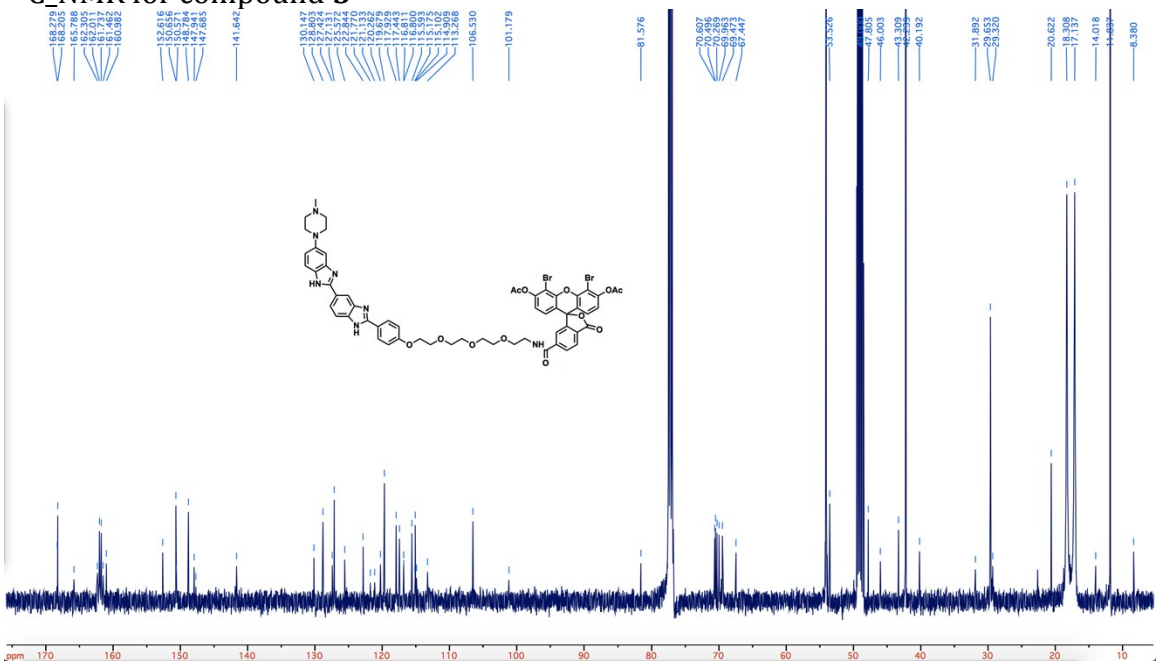
^{13}C_NMR for compound **3**



^{1}H_NMR for compound **5**



^{13}C_NMR for compound **5**



5. Supporting Figures

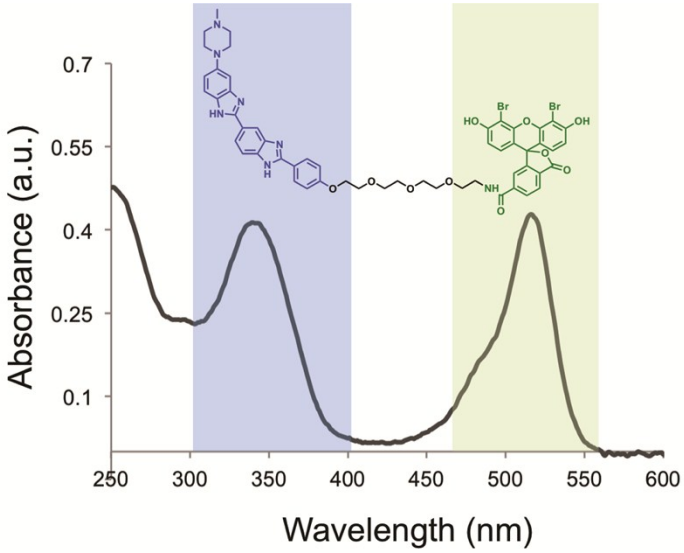


Fig. S1. UV-vis spectrum of deprotected HoeDBF.

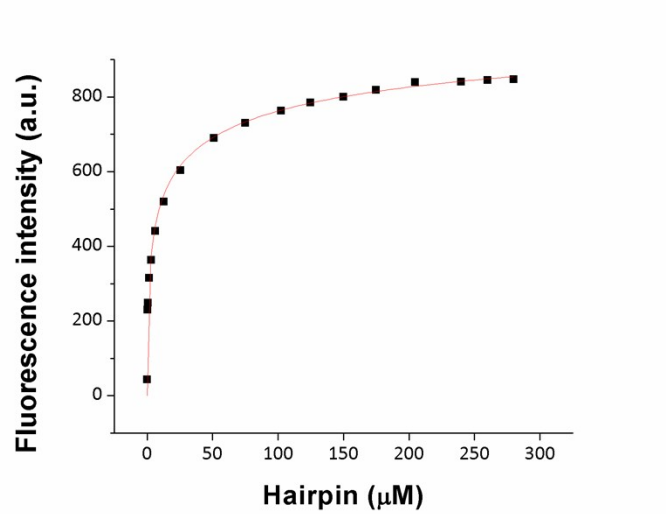


Fig. S2. Titration of 0.2 μM HoeDBF probe against varying concentrations of hairpin DNA with ex 500 nm and em 530 nm. Fitting the curve in OriginPro 8 using Growth/Sigmoidal-Hill function suggested that HoeDBF bound to dsDNA with a K_D of 15 μM, which is similar to other reported bisbenzimide derivatives. 28-bp hairpin DNA has a sequence of CGC GAA TTC GCG TTT TCG CGA ATT CGC G.[3] The quantum yield of DNA-bound HoeDBF was measured to be 25%.

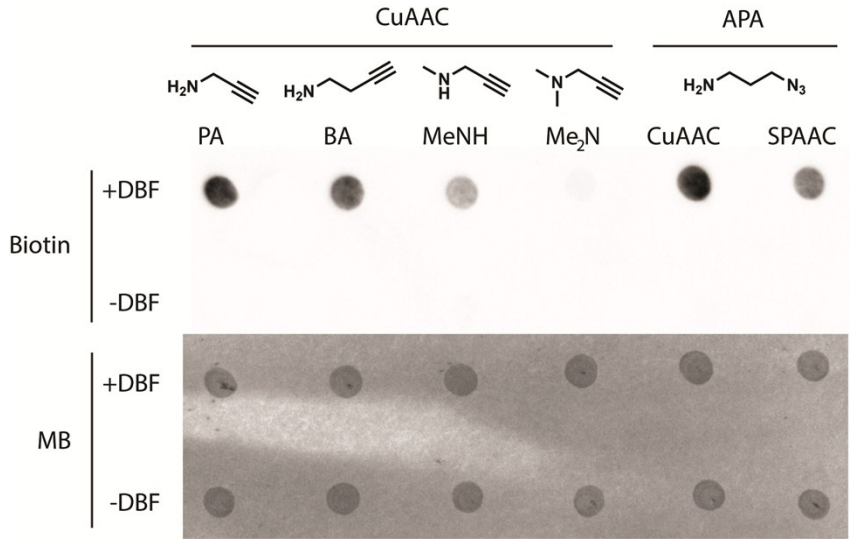


Fig. S3. Dot blot analysis for RNA *in vitro* labeling. PA: propargyl amine; BA: but-3-yn-1-amine; MeNH: *N*-methylprop-2-yn-1-amine; Me2N: *N*,*N*-dimethylprop-2-yn-1-amine; APA: 3-azidopropan-1-amine.

We explored different alkyne-containing amine species for RNA labelling. The capability of different amine compounds to label RNA was assayed by *in vitro* RNA labelling and dot blot analysis. The strength of the biotin signal in descending order is as follows: PA ≈ APA(CuAAC) ≈ BA > APA(SPAAC) > MeNH > Me_2N. The result supported the previous suggestion that amine reacted with the oxidized guanosine through nucleophilic addition. Under CuAAC condition, PA afforded similarly strong signal as APA while SPAAC showed lower reactivity compared with CuAAC.

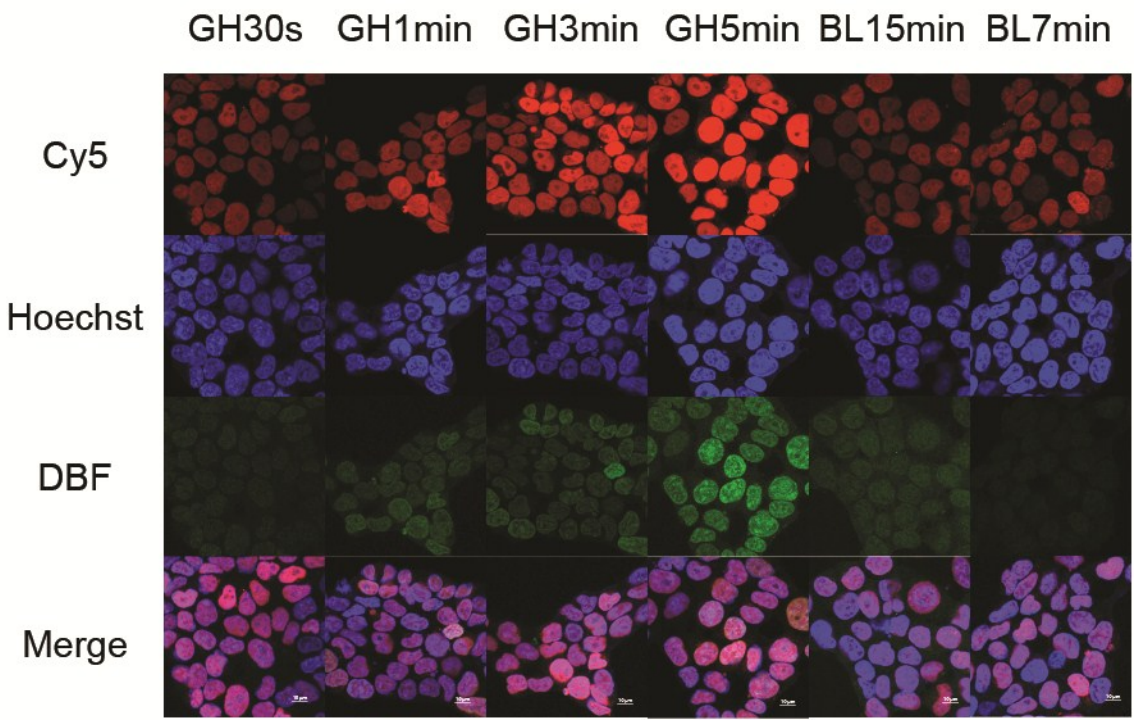


Fig. S4. Optimizing the light source and irradiation time for the generation of singlet oxygen and the subsequent labeling of biomolecules via fluorescence confocal imaging. GH30s: green light – high power – 30s irradiation; GH1min: green light – high power – 1 min irradiation; GH3min: green light – high power – 3 min irradiation; GH3min: green light – high power – 5 min irradiation; BL15min: blue light – low power – 15 min irradiation; BL7min: blue light – low power – 7 min irradiation; Cy5: ~670 nm emission channel; Hoechst: ~450 nm emission channel; DBF: ~520 nm emission channel; Merge: combined channels.

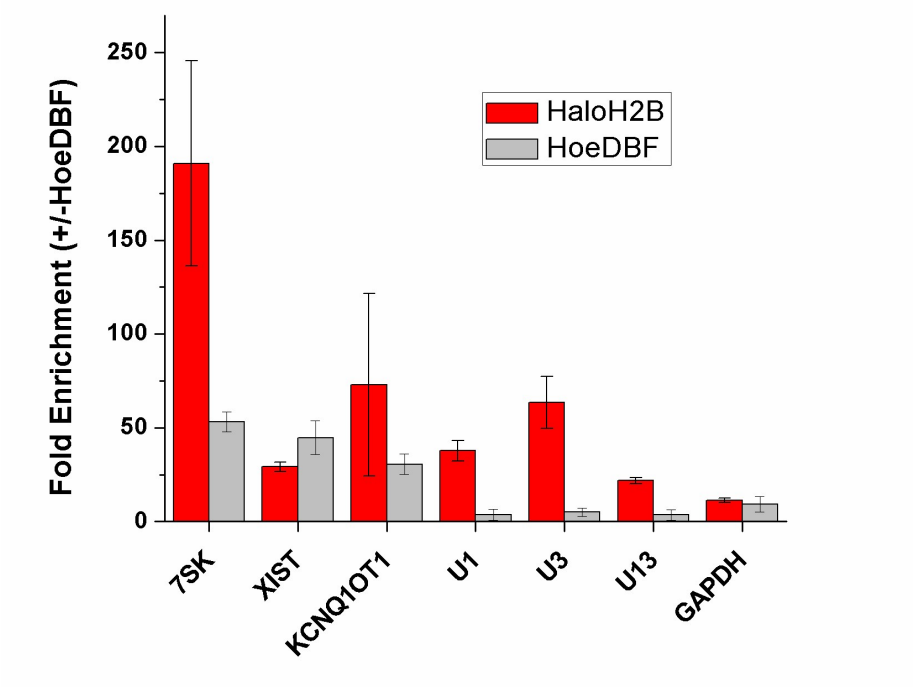


Fig. S5. RT-qPCR comparative analysis of RNAs tagged by HaloH2B and HoeDBF. Enrichment was calculated against a negative control with the ΔΔCt method, in biological duplicates.

6. References

1 R. C. Oslund and M. H. Gelb, *Biochemistry*, 2012, **51**, 8617-8626.
2 Y. Li, M. B. Aggarwal, K. Nguyen, K. Ke and R. C. Spitale, *ACS Chem. Biol.*, 2017, **12**, 2709-2714.
3 G. Lukinavičius, C. Blaukopf, E. Pershagen, A. Schena, L. Reymond, E. Derivery, M. Gonzalez-Gaitan, E. D'Este, S. W. Hell, D. Wolfram Gerlich and K. Johnsson, *Nat. Comm.*, 2015, **6**, 8497