Supplementary Materials for

Ultrastructural visualization of chromatin in cancer pathogenesis using a simple small-molecule fluorescent probe

Jianquan Xu, Xuejiao Sun, Kwangho Kim, Rhonda M. Brand, Douglas Hartman, Hongqiang Ma, Randall E. Brand, Mingfeng Bai, Yang Liu*

*Corresponding author. Email: liuy@pitt.edu

Published 4 March 2022, Sci. Adv. 8, eabm8293 (2022)
DOI: 10.1126/sciadv.abm8293

This PDF file includes:

Supplementary Methods
Figs. S1 to S8
Supplementary Methods

Synthesis of Hoechst-Cy5

Materials
Solvents were obtained from either an MBraun MB-SPS solvent system or freshly distilled (tetrahydrofuran was distilled from sodium-benzophenone; toluene was distilled from calcium hydride and used immediately; dimethyl sulfoxide was distilled from calcium hydride and stored over 4 Å molecular sieves). Commercial reagents were used as received. The molarity of n-butyllithium solutions was determined by titration using diphenylacetic acid as an indicator (average of three determinations).

Reagent abbreviations
DMF: N,N-dimethylformamide
DMSO: dimethyl sulfoxide
TFA: trifluoroacetic acid.
DCM: Dichloromethane
DIPEA: N,N-diisopropylethylamine
HBTU: O-(benzotriazol-1-yl)-N,N,N’,N’-tetramethyluronium hexafluorophosphate.

Instrumentation
Semi-preparative reverse phase HPLC was conducted on a Waters HPLC system using a Phenomenex Luna 5 μm C18(2) 100 Å Axia 250 × 10.00 mm column or preparative reverse phase HPLC (Gilson) using a Phenomenex Luna column (100 Å, 50 × 21.20 mm, 5 μm C18) with UV/Vis detection. Infrared spectra were obtained as thin films on NaCl plates using a Thermo Electron IR100 series instrument and are reported in terms of frequency of absorption (cm⁻¹). 1H NMR spectra were recorded on Bruker 400, 500, or 600 MHz spectrometers and are reported relative to deuterated solvent signals. Data for 1H NMR spectra are reported as follows: chemical shift (δ ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, p = pentet, m = multiplet, br = broad, app = apparent), coupling constants (Hz), and integration. 13C NMR spectra were recorded on Bruker 100, 125, or 150 MHz spectrometers and are reported relative to deuterated solvent signals. LC/MS was conducted and recorded on an Agilent Technologies 6130 Quadrupole instrument. High-resolution mass spectra were obtained from the Department of Chemistry and Biochemistry, University of Notre Dame using either a JEOL AX505HA or JEOL LMS-GCmate mass spectrometer or by the Vanderbilt University Center for Neuroscience Drug Discovery (VCNDD) on a Micromass –Q-Tof API-US mass spectrometer.

Compound preparation

Compound 1. To a solution of 4-(4-methylpiperazin-1-yl)benzene-1,2-diamine (0.3 g, 0.85 mmol) in nitrobenzene (8.5 mL) was added N-(4-formyl-2-nitrophenyl)acetamide (0.18 g, 0.85 mmol), and stirred for 24 h at 130 °C. The reaction mixture was cooled to room temperature and excess amounts of hexanes were added. The precipitate was filtered, washed with hexanes, and dried. The crude product was purified by ISCO column chromatography (gradient elution from EtOAc to 9:1 EtOAc/MeOH solvent mixture with 0.1% triethylamine) to provide brown solid product (0.29 g, 85%). 1H NMR (400 MHz, CDCl₃) δ 10.81 (s, 1 H), 10.71 (s, 1 H), 8.79 (s, 1 H), 8.54 (d, J = 8.0 Hz, 1 H), 7.87 (d, J = 8.8 Hz 1 H), 7.61 (d, J = 8.8 Hz, 1 H), 7.17 (d, J = 8.8 Hz, 1 H), 7.14 (s, 1 H), 3.62-3.76 (m, 4 H), 3.18-3.14 (m, 4 H), 2.89 (s, 3 H), 2.16 (s, 3 H); LCMS calc’d for C₂₀H₂₂N₆O₃ [M+H]⁺: 395.4 measured 395.4.

Compound 2. A suspension of N-(4-(6-(4-methylpiperazin-1-yl)-1H-benzo[d]imidazol-2-yl)-2-nitrophenyl) acetamide (450 mg, 1.14 mmol) and 2.28 mL of NaOH (1 M, 2.28 mmol) in a mixture of MeOH (6 mL) and H₂O (1.5 mL) was stirred at 70°C for 4 h. The reaction mixture was acidified with 1 N
HCl to adjust pH 3, precipitate was collected by filtration, and washed with MeOH to afford the brown solid product (273 mg, 68% yield). 1H NMR (400 MHz, DMSO-d6) δ 8.76 (s, 1H), 8.15 (d, J = 8.0 Hz, 1H), 7.77 (s, 1H), 7.38 (d, J = 8.8 Hz, 1H), 7.14 (d, J = 8.8 Hz, 1H), 6.92 (d, J = 8.4 Hz, 1H), 3.17-3.10 (m, 4H), 2.42-2.49 (m, 4H), 2.24 (s, 3H); LCMS calc’d for C18H20N6O2 [M+H]+: 353.4 measured 353.4.

**Compound 3.** A mixture of 4-(6-(4-methylpiperazin-1-yl)-1H-benzo[d]imidazol-2-yl)-2-nitroaniline (215 mg, 0.69 mmol) and 10% Pt/C (catalytic amount) in MeOH/ EtOAc (3/3 mL) was hydrogenated at 1 atm for 14 h. The crude was filtered through Celite pad, washed with EtOAc, filtrate was concentrated in vacuo to afford brown solid product (146 mg, 66%). 1H NMR (400 MHz, DMSO-d6) δ 7.38-7.36 (m, 2H), 7.20 (d, J = 8.0 Hz, 1H), 6.91 (s, 1H), 6.83 (d, J = 8.0 Hz, 1H), 6.60 (d, J = 8.4 Hz, 1H), 4.92 (s, 2H), 4.66 (s, 2H), 3.11-3.07 (m, 4H), 2.49-2.45 (m, 4H), 2.21 (s, 3H); LCMS calc’d for C36H36N6O3 [M+H]+: 601.7 measured 601.7.

**Compound 4.** A mixture of 4-(6-(4-methylpiperazin-1-yl)-1H-benzo[d]imidazol-2-yl)benzene-1,2-diamine (50 mg, 0.16 mmol) and benzyl 4-(4-formylphenoxy)butanoate (46 mg, 0.16 mmol) in nitrobenzene (1.6 mL) was stirred at 130 °C for 24 h. The solution was cooled to room temperature, and an excess of hexanes was added to precipitate the product. The precipitated crude product was collected by filtration and then purified by ISCO column chromatography (9:1 ethyl acetate/methanol containing 0.1% triethylamine) to afford brown solid (70 mg, 75%). 1H NMR (400 MHz, DMSO-d6) δ 8.36 (s, 1H), 8.15 (dd, J = 9.2, 3.2 Hz, 2H), 8.02-8.00 (m, 2H), 7.62-7.58 (m, 2H), 7.39-7.35 (m, 5H), 7.13 (d, J = 7.2 Hz, 1H), 6.95-6.93 (m, 2H), 5.13 (s, 2H), 4.11 (t, J = 6.0 Hz, 2H), 3.14 (s, 4H), 2.87-2.83 (m, 2H), 2.08-2.03 (m, 2H); LCMS calc’d for C36H36N6O3 [M+H]+: 601.7 measured 601.7.

**Compound 5.** To a solution of benzyl 4-(4-(6-(4-methylpiperazin-1-yl)-1H,3'H-[2,5'-bibenzo[d]imidazol]-2'-yl)phenoxy)butanoate (70 mg, 0.12 mmol) in anhydrous DMSO (0.69 mL) was added DIPEA (0.037 mL, 0.212 mmol) and HBTU (24 mg, 0.064 mmol) and stirred for 5 min. To this was added a solution of tert-butyl (2-(2-aminoethoxy)ethoxy)ethyl carbamate (16 mg, 0.064 mmol) in anhydrous DMF (1 mL). The mixture was stirred for 2 h at room temperature, concentrated, and purified by ISCO column chromatography (50-70% methanol in ethyl acetate) to afford a pale yellow solid (28 mg, 71%). 1H NMR (CD3OD) δ 8.24 (s, 1H), 8.06 (d, J = 9.2 Hz, 2H), 7.97 (brd, 1H), 7.71 (brd, 1H), 7.51 (d, J = 8.8 Hz, 1H), 7.15 (s, 1H), 7.11 (d, J = 8.8 Hz, 2H), 7.07 (dd, J = 2.4, 8.8 Hz, 1H), 4.10 (t, J = 6.4 Hz, 2H), 3.58 (s, 4H), 3.55 (t, J = 5.6 Hz, 2H), 3.51-3.48 (m, 2H), 3.39 (t, J = 5.6 Hz, 2H), 3.27-3.22 (m, 4H), 3.20 (t, J = 5.6 Hz, 2H), 2.71-2.67 (m, 4H), 2.44 (t, J = 7.2 Hz, 2H), 2.39 (s, 3H), 2.16-2.11 (m, 2H), 1.42 (s, 9H); LCMS calc’d for C40H52N8O6 [M+H]+: 741.4 measured 741.4.

To a solution of Boc-amine (19 mg, 0.026 mmol) in dichloromethane (1.5 mL) was added TFA (0.6 mL) at room temperature. The reaction mixture was stirred for 1 h and concentrated in vacuo. The crude
product was used next step without further purification; LCMS calc’d for C35H44N8O4 [M+H]+: 640.3 measured 641.3.

**Hoechst-Cy5 probe.** To a solution of Cy5-NHS-ester (15 mg, 0.026 mmol) in anhydrous DMSO (2 mL) was added DIPEA (0.07 mL, 0.039 mmol) at room temperature. The reaction mixture was stirred under Ar at room temperature for 2 h, the solvent was removed in vacuo, purified by reverse phase Gilson HPLC using semi-preparative C18 column (20-95% linear gradient of MeCN and 0.1 % aqueous TFA) to give respective fluorescent probe as a blue solid (18 mg, 63 %); LCMS calc’d for C67H81N10O5 [M+H]+: 1107.4 measured 1106.4.
Figure S1. Synthetic route of Hoechst-Cy5.
Figure S2. Representative background-corrected raw images and reconstructed super-resolution images.

(A1-D1) Background-corrected single-frame raw images of DNA labeled with Hoechst-Cy5, TOTO-3, Hoechst-JF646 and Live-650 in Figure 2. (A2-D2, A3-D3, A4-D4) The corresponding super-resolution images reconstructed with PathSTORM. Scale bars: 10 µm in A1-A2, 1 µm in A3 and 200 nm in A4.
Figure S3. Conventional wide-field fluorescence images of genomic DNA from normal and pathological tissue.

These wide-field fluorescence images correspond to the same set of super-resolution images shown in Figure 5. The zoom-in regions are the same as those shown in Figure 5.
Figure S4. Additional two-color STORM images of genomic DNA and lamin A/C from tissue with colorectal adenocarcinoma.

The two-color STORM images are from different nuclei (labeled 1, 2 and 3) from patients with colorectal adenocarcinoma besides the image shown in Figure 7. These images showed a certain level of heterogeneous structure in chromatin and nuclear lamina.
Figure S5. Two-color conventional wide-field and super-resolution images of H3K4me3 and genomic DNA.

(A) Two-color conventional wide-field fluorescence images of nucleus in cells stained with H3K4me3 conjugated with CF568 (green) and DNA labelled by EdU-Alexa647 (red). Some cells with H3K4me3 signals were not labeled with EdU-Alexa647, suggesting those without newly synthesized DNA. This is an example with mis-labeled DNA due to lack of proliferation. (B) Various labeling efficiency of DNA in cells with different proliferation. For example, cell 1 and cell 2 showed different labeling density of DNA due to different proliferation rate.
Figure S6. Representative raw images and reconstructed super-resolution images of genomic DNA in cultured cells.

Single-frame raw images of genomic DNA without (A1-E1) and with background correction (A4-E4), labeled with Hoechst-Cy5, Alexa647, TOTO-3, Hoechst-JF646 and Live-650, respectively, in cultured cells. The super-resolution images were reconstructed with ThunderSTORM (A2-E2, A3-E3) and PathSTORM (A5-E5, A6-E6). The image quality from Cy5 and Alexa647 remains the best. PathSTORM significantly improves the overall quality of STORM images via background correction and algorithms accounting for overlapping emitters in single molecule localization.
Figure S7. Representative two-color STORM images of genomic DNA and replication foci.

Two-color STORM images of genomic DNA labeled with Hoechst-Cy5 (red), and replication foci labeled with pulse-labeling of EdU-CF568 for a duration of (A) 30 minutes, (B) one hour and (C) three hours (green). These images showing the capability of Hoechst-Cy5 to image genomic DNA combined with replication foci pause labeled with EdU.
Figure S8. Illustration of watershed-based segmentation of DNA nanodomains within the cell nucleus.

(A) The reconstructed STORM images of DNA from a single nucleus labeled by Hoechst-Cy5. (B) The magnified region from the white box in (A). (C) The segmented DNA nanodomains by watershed from the same region in (B).