

Supporting Information for
Leveraging Intracellular ALDH1A1 Activity for Selective Cancer Stem-like Cells Labeling and Targeted Treatment via in vivo Click Reaction

Yang Bo^{1‡}, Jingyi Zhou^{1‡}, Kaimin Cai^{1,2‡}, Ying Wang¹, Yujun Feng¹, Wenming Li¹, Yunjiang Jiang¹, Shanny Hsuan Kuo³, Jarron Roy⁴, Chelsea Anorma⁵, Sarah H. Gardner⁶, Long M. Luu², Gee W. Lau³, Yan Bao⁷, Jefferson Chan^{5,8}, Hua Wang^{1,4,8,9,10,11*}, Jianjun Cheng^{1,2,4,5,8,9,10,11*}

*Correspondence should be addressed to chengjianjun@westlake.edu.cn; huawang3@illinois.edu.

This PDF file includes:

Supporting text
Figures S1 to S29
Scheme S1 to S4
Chemical synthesis
SI References

Supporting text

Materials. All chemicals purchased from Sigma-Aldrich (St. Louis, MO, USA), Thermo Fisher Scientific (Waltham, MA, USA), Chem-Impex (Wood Dale, IL, USA) or Oakwood Chemical (Estill, SC, USA) were reagent grade and used without further purification unless otherwise stated. DBCO-Cy5 was purchased from Click Chemistry Tools (Scottsdale, AZ, USA). Dulbecco's Modified Eagle Medium (DMEM) was obtained from Invitrogen (Carlsbad, CA, USA). Iscove's Modified Dulbecco's Medium were obtained from American Type Culture Collection (Manassas, VA, USA). Fetal Bovine Serum (FBS) was purchased from Lonza Walkersville Inc. (Walkersville, MD, USA). BD Falcon culture plates were purchased from Fisher Scientific (Hampton, NH, USA). ProLong Gold antifade reagent was purchased from Life Technologies (Carlsbad, CA, USA). Anhydrous dichloromethane (DCM), tetrahydrofuran (THF) and dimethylformamide (DMF) were prepared by passing the solvent through alumina columns. DBCO-PEG₄-biotin, streptavidin–horseradish peroxidase (HRP), and Pierce ECL western blotting substrate were purchased from Thermo Fisher Scientific (Waltham, MA, USA).

Instrumentation. High performance liquid chromatography (HPLC) analysis was performed on a Shimadzu CBM-20A system (Shimadzu, Kyoto, Japan) equipped with an SPD-20A PDA detector (190–800 nm), a RF-20A fluorescence detector, and an analytical C18 column (Shimadzu, 3 μ m, 50 $\times$ 4.6 mm, Kyoto, Japan). Preparative HPLC was performed on a CombiFlash®*R*_f system (Teledyne ISCO, Lincoln, NE, USA) equipped with a RediSep®*R*_f HP C18 column (Teledyne ISCO, 30 g, Lincoln, NE, USA). Lyophilization was conducted in a Labconco FreeZone lyophilizer (Kansas City, MO, USA). Nuclear Magnetic Resonance (NMR) spectra were recorded on a Varian U500 (500 MHz) or VXR500 (500 MHz) or a Bruker Carver B500 (500 MHz) spectrometer. Electrospray ionization (ESI) mass spectra were obtained from a Waters ZMD Quadrupole Instrument (Waters, Milford, MA, USA). Flow cytometry analysis was conducted on a BD FACS Canto 6-color flow cytometry analyzer (BD, Franklin Lakes, NJ, USA) or a BD Accuri™ C6 Flow Cytometer (BD, Franklin Lakes, NJ, USA). Confocal laser scanning microscopy (CLSM) images were taken by using a Zeiss LSM 700 Confocal Microscope (Carl Zeiss, Thornwood, NY, USA). Western blotting protein bands were imaged on an ImageQuant LAS 4000 gel imaging system (GE, Pittsburgh, PA, USA). *In vivo* and *ex vivo* images of animals and tissues were taken on a Bruker *In Vivo* Imaging System (Bruker, Billerica, MA, USA). Frozen tumor tissues were embedded with optimum cutting temperature (O.C.T.) compound (Sakura Finetek USA, Torrance, CA, USA), sectioned by a Leica CM3050S Cryostat, and mounted onto glass slides.

Chemical synthesis and characterization. Detailed synthetic route and characterization are listed in the supporting information file.

Cell culture. MDA-MB-231, LS174T, A549, and HEK293T cells were purchased from American Type Culture Collection (Manassas, VA, USA). Cells were cultured in DMEM containing 10% FBS, 100 units/mL Penicillin G and 100 μ g/mL streptomycin (Invitrogen, Carlsbad, CA, USA) at 37 °C in 5% CO₂ humidified incubator. K562 cells were obtained from Prof. Paul Hergenrother (UIUC, Chemistry). K562 cells were cultured as a suspension in Iscove's Modified Dulbecco's Medium (IMDM, ATCC) supplemented with 10% FBS, 100 units/mL Penicillin G and 100 μ g/mL streptomycin (Invitrogen, Carlsbad, CA, USA) at 37 °C in 5% CO₂ humidified incubator.

Animals. Female athymic nude mice and BALB/c mice were purchased from Charles River (Wilmington, MA, USA). Feed and water were available ad libitum. Artificial light was provided in a 12 h/12 h cycle. The animal protocol was reviewed and approved by the Illinois Institutional Animal Care and Use Committee (IACUC) of the University of Illinois at Urbana–Champaign. For all the animal studies, mice were randomly allocated to each group. The investigator was aware of the group allocation during the animal studies, as demanded by the experimental designs.

Expression and Purification of ALDH1A1 and Related Isoforms. The expression of ALDH1A1 was conducted with the established method. (1, 2) Expression constructs for recombinant human ALDH1A1, ALDH1A3, ALDH2, ALDH4A1, and ALDH5A1 were generously provided by Prof. Daria Mochly-Rosen (Stanford, Chemical and Systems Biology). *E. coli* BL21(DE3) cells were transformed with each of the above constructs to produce the corresponding ALDH isoforms. Target proteins were then eluted by wash buffer (20 mM sodium phosphate pH 7.4, 0.5 M NaCl, 500 mM imidazole). The purity of the eluted protein

was determined to be $\geq 95\%$ by SDS-PAGE, and the concentration of protein was determined by standard BCA assay. Protein was stored with 50% glycerol at -80°C .

AAMCHO Isoform Selectivity Assay. Activation of AAMCHO was assessed using 20 units of each ALDH isoform. All enzymatic reactions were performed in 50 mM triethanolamine buffer (pH 7.4) with 2.5 mM NAD^{+} and 5% v/v DMSO in a 1 mL quartz cuvette at room temperature. Before measurement, AAMCHO (1 μM) was added to an HPLC vial. At the designated time point, 100 μL aliquot was taken out and quenched by 100 μL methanol and 300 μL acetonitrile. Followed by incubation with 1 μL 10 mM DBCO-Cy5 at 37°C for 1 h and detected by HPLC with a fluorescence detector (excitation: 650 nm, emission: 688 nm). All scans were normalized to the signal from AAMCHO in 50 mM TEA and 2.5 mM NAD^{+} (without enzyme). Synthesized AAMOH also went through the procedure above as a standard reference.

siRNA Knockdown of ALDH1A1. K562 cells were on a poly-L-lysine (Trevigen) coated Tissue Culture 24-well-plate (Costar) 1 day before treatment with siRNA. Both the negative control scrambled siRNA (Sigma-Aldrich, MISSION siRNA Universal Negative Control #1) as well as the ALDH1A1 siRNA (Sigma Aldrich, SASI_Hs01_00244056) was applied at either 1 nM or 10 nM final concentration and pre-mixed with the Lipofectamine 2000 (Thermo Fisher). After treatment, cells were incubated in Opti-MEM (Gibco) at 37°C , 5% CO_2 for 8 h. The Opti-MEM was then removed and replaced with full IMDM growth media. Cells were incubated with 50 μM AAMCHO (while PBS and Ac_4ManAz incubation was used as control) at 37°C , 5% CO_2 for 48 h. Then the cells were probed with 20 μM DBCO-Cy5 for 1 h and subjected to flow cytometry analysis.

Western blotting analysis of azide-sugar labeled cells. AAMCHO or Ac_4ManAz (50 μM) was added to the cell culture, and cells were incubated at 37°C for 48 h. For the ALDH1A1 inhibition group, the corresponding sugar was spiked with 50 μM disulfiram. After washing with PBS, cells were lysed with lysis buffer (50 mM Tris-HCl, 1% SDS, pH 7.4) and the lysate was incubated at 4°C for 30 min. Following the removal of the lysate debris by centrifuge at 10000 rpm, the protein concentration of cell lysate was quantified by using standard BCA assay. After adjusting the amount of protein, iodoacetamide solution (60 mM stock in DMSO, 1 μL in 20 μL lysate) was added and incubated at 37°C for 1 h to fully block the free thiol groups in the proteome. DBCO-(PEG)₄-biotin (10 mM, 1 μL in 20 μL lysate) was added to the mixture and incubated overnight at 37°C . The protein mixture was mixed with 2 $\times$ sample buffer, heated for 10 min at 95°C , subjected to electrophoresis on a 4-20% gel (MiniPROTEAN TGX precast gel, Bio-Rad) and then transferred to PVDF membrane (Mini format 0.2 μm PVDF Trans-Blot Turbo™ Transfer pack, Bio-Rad) using a Trans-Blot Turbo™ Transfer System (Bio-Rad). The protein concentration was confirmed by reversible Ponceau S staining (Sigma-Aldrich, P7170). The membrane was incubated with streptavidin-HRP (Invitrogen) at room temperature for 1 h and washed with TBST at room temperature while shaking for 10 min. The blots were developed with ECL Western Blotting substrate and imaged with LAS 4010 Luminescent image analyzer (GE Healthcare).

Trypan blue cell viability assay. K562 cells were plated in a 24-well plate at a density of 500,000 cells/mL in full IMDM media supplemented with either 10 μM or 50 μM AAMCHO or $\text{Ac}_4\text{N-AAM}$. At 12, 24, and 48 h, the cells were suspended with trypsin, and 10 μL sample was taken from each of the samples and mixed 1:1 with a 0.4% (w/v) trypan blue solution in PBS. Samples were incubated for 1 min at room temperature before being loaded onto a hemocytometer, where live and dead cells were counted.

K562 labeling with AAMCHO in hypoxia conditions. Using a hypoxic chamber, K562 cells were grown in full IMDM media with 1% O_2 , 5% CO_2 , and 94% N_2 for 72 hours. As a control, cells were grown in a standard normoxia incubator supplied with 5% CO_2 , while co-incubated with 50 μM AAMCHO, or $\text{Ac}_4\text{N-AAM}$, or Ac_4ManAz as control. After 48 h, the cells were collected by centrifugation and stained with 20 μM DBCO-Cy5 for 1 h. The cells were then spun down and resuspended in PBS. Anti-CD34-FITC and anti-CD38-VioBlue were added and incubated for 20 minutes on ice before subjecting to flow cytometry analysis.

Tumorsphere Culture and Imaging. (1) *Confocal microscopy imaging:* Stem cell tumorsphere formation from MDA-MB-231 breast cancer cells was performed as described previously.^(1, 2) Cells were resuspended and diluted to a density of 2000 cells/ mL in DMEM/F12 (Sigma-Aldrich) supplemented with 2% B27 supplement (Thermo Fisher), 40 ng/mL rhFGF-2 (Miltenyi Biotec), and 20 ng/mL rhEGF (Gibco). Cells were

plated in ultralow attachment six-well plates (Corning) and incubated at 37 °C with 5% CO₂. The media were exchanged with full DMEM media to allow differentiation over 2, 6, 12, and 24 h, respectively, and then incubated with AAMCHO for 48 h. After rinsing three times with phosphate buffer, the labeled tumorsphere was incubated with 20 μM DBCO-Cy5 for 1 h and fixed with 4% PFA solution. The imaging was performed on a Zeiss LSM700 Confocal Microscope. (2) *Flow cytometry analysis*: The mammosphere was labeled by AAMCHO or Ac₄ManAz for 48 h after being treated with FBS for 0 h or 36 h. Single cell suspensions were collected via trypsin digesting and pipetting, and incubated with 50 μM DBCO-Cy5 for 30 min prior to flow cytometry analysis. For cell sorting, the isolated cells from either the sphere culture or xenograft tumor were counterstained with 1 μg/mL FITC-conjugated anti-CD44 (Santa Cruz, sc-7297 FITC) for 30 min. Sorting was performed using FACSaria II (BD Biosciences).

Sphere-Forming Assay. Sorted single cell suspensions were plated (4000 cells/well) in 6 well ultralow attachment culture plates (Corning) in DMEM/F12 medium supplemented with 2% B27 supplement (Thermo Fisher), 40 ng/mL rhFGF-2 (Miltenyi Biotec), and 20 ng/mL rhEGF (Gibco). The suspension was cultured for 7 days, and then the sphere number was quantified under the microscope. The effective sphere-forming threshold is set above 50 μm in diameter.

In vivo tumorigenesis assay. 10⁴ cells of each group collected from FACS sorting were suspended in HBSS/Matrigel (1/1, v/v, 30 μL) and inoculated into the mammary fat pad of 6-week-old female athymic nude mice. Tumor volume and body weight of mice were measured at designated days. The tumor volume was calculated using the formula (length) × (width)² / 2, where the long axis diameter was regarded as the length, and the short axis diameter was regarded as the width. At Day 28, all mice were euthanized, and tumor tissues were collected and weighed.

Quantification of azide-sialic acid in cell or tissue lysate. The quantification of conjugated azide-monosaccharide was based on an established method using Cu(I) catalyzed click reaction with coumarin-alkyne. The tissue or cell lysate was homogenized with a mechanical disruptor and then disrupted by an ultrasonicator with lysis buffer (50 mM Tris-HCl, 1% SDS, pH 7.4). The protein concentration of each sample was quantified with standard BCA protocol. 90 μL of lysate was mixed with 10 μL pure acetic acid while heated at 80° C for 3 h. Then, the mixture was centrifuged at 10000 rpm for 5 min. The supernatant was mixed with a newly prepared reaction kit containing 7-Ethynylcoumarin (50 μM, E1092, TCI Chemicals) in DMSO, THPTA (5 mM, Tris(3-hydroxypropyltriazolylmethyl) amine) (762342, Sigma-Aldrich) in DMSO, CuSO₄ (5 mM) in ddH₂O, and ascorbic acid (30 mM, 1043003, Sigma-Aldrich) in ddH₂O. The reaction mixture was shaken at 37° C overnight and detected by HPLC equipped with a fluorescence detector (λ_{Ext}: 328 nm, λ_{Emi}: 415 nm). Neu5NAz was synthesized following a previously reported procedure.(3) The Neu5NAz amount is normalized by cell counting for cell lysate or protein concentration for tissue lysate.

In vivo AAMCHO mediated CSC labeling. Mammosphere from MDA-MB-231 cells was performed as described previously. For non-CSC cells, the mammospheres were treated with 10 % FBS and full DMEM media to allow differentiation over 36 h. Before *in vivo* inoculation, the CSC or non-CSC single cell suspension was prepared by treating the mammosphere with trypsin and pipetting gently. CSCs and non-CSCs (10⁵ cells in 50 μL HBSS) were subcutaneously inoculated into the left and right flank, respectively. After a week, AAMCHO (60 mg/kg) was i.v. injected once daily for three consecutive days. Mice i.v. administered with Ac₄ManAz (40 mg/kg) or PBS were used as controls. (1) *In vivo* fluorescence imaging: At 24 h post the last injection of AAMCHO or Ac₄ManAz or PBS, DBCO-Cy5 (5 mg/kg) was i.v. injected and its biodistribution was monitored using the Bruker In Vivo Xtreme Imaging System. The excitation/emission filter of 630 nm/700 nm was used. Tumors and organs were harvested from mice at 48 h p.i. of DBCO-Cy5 and imaged *ex vivo*. Fluorescence intensity at selected ROIs was quantified using the Bruker imaging software. After *ex vivo* imaging, tumors were bisected. Half of the tumors were directly frozen in O.C.T. compound, sectioned into 8 μm thickness, stained with Hoechst 33342 (1 μg/ml, Invitrogen) and FITC conjugated CD44 antibody at 1:500 dilution (Santa Cruz, sc-7297 FITC), imaged under a confocal laser scanning microscope. The other half of the tumors were minced and digested with 1mg/mL Collagenase (Type IV, Gibco) digestion buffer. Suspensions were filtered through 70 μm cell strainers and were counterstained with 1 μg/ mL FITC conjugated CD44 antibody (Santa Cruz, sc-7297 FITC) at room temperature for 30 min. The cell suspension was washed by centrifugation and subjected to flow cytometry analysis. (2) Azide-sialic acid quantification: In a separate study, at 24 h post the last injection of AAMCHO

or Ac₄ManAz or PBS, tumors and major organs were harvested for azide-sialic acid quantification following the above-mentioned procedures.

***In vivo* tissue penetration and biodistribution.** The orthotopic tumor model was established by inoculation of single cell suspension prepared from MDA-MB-231 mammosphere (10⁶ cells per mice) into the mammary fat pad of athymic nude mice. After 10 days, AAMCHO (60 mg/kg) was i.v. injected once daily for three consecutive days. At 24 h post the last injection of AAMCHO, DBCO–Cy5 (5 mg/kg) and FITC-conjugated anti-CD44 (5 mg/kg) were i.v. injected. After 24 h, tumors were bisected, and flash-frozen in optimum cutting temperature (OCT) compound on dry ice and sectioned for histology on a cryostat (10- μ m slices). The tumor slices were stained with mouse monoclonal IgG₁ κ anti-mouse CD31 and m-IgG kappa BP-CFL 488 (sc-376764 and sc-516176, Santa Cruz), counterstained with Hoechst 33342 (1 μ g/ml, Invitrogen), and imaged under a confocal laser scanning microscope.

β -glucuronidase induced degradation of DBCO-MMAE. Recombinant Human β -Glucuronidase (6144-GH, R&D systems) was activated by suspending (40 mg/mL) in assay buffer (100 mM NaOAc in ddH₂O, pH = 3.5). DBCO-MMAE (1.4 mg/mL) was added to the activated β -Glucuronidase solution or blank assay buffer or 10% FBS and incubated at 37 °C. Aliquots of the mixture were taken out for HPLC measurements at selected time points.

Development of multidrug-resistant cell line. The multidrug-resistant (MDR) cell lines A549-*R* and LS174T-*R* were developed according to the drug-resistant cell line protocol. A549 or LS174T cells were incubated with 5 μ M cisplatin or 400 nM oxaliplatin in each passage, respectively. Treatment time: 1, 2, 4, 12, 24, 48, 72 hours (repeat 3 times, except for 72 h treatment which was repeated 10 times). The medium was changed to discard the apoptotic cells and allow the cell growing to 70% confluence for each time treatment. After passage, the drug-resistant inducing efficiency was tested by MTT assay.

Long-term efficacy study of AAMCHO + DBCO-MMAE against orthotopic MDA-MB-231 tumors. MDA-MB-231 tumors were established in 6-week-old female athymic nude mice by orthotopic injection of MDA-MB-231 cells (1.5 $\times$ 10⁶ cells in HBSS/Matrigel (50 μ L, 1/1, v/v)) into the fat pad under the mammary gland. When the tumors reached $\sim$ 100 mm³, mice were randomly divided into four groups (group 1: AAMCHO + DBCO-MMAE; group 2: DBCO-MMAE; group 3: MMAE; group 4: PBS; n = 7 or 8). AAMCHO (60 mg/kg) was i.v. injected to group 1 mice once daily for three consecutive days (Day 0, 1, 2). DBCO-MMAE (8 mg/kg) was i.v. injected on Days 3. For group 3, i.p. injection of MMAE (0.25 mg/kg, maximum tolerance dose) was conducted at Days 3. Each mouse's tumor volume and body weight were measured twice a week. In a separate study for anti-cancer efficacy against the late-stage tumor, when the tumors reached $\sim$ 500 mm³, mice were randomly divided into four groups (group 1: AAMCHO + DBCO-MMAE; group 2: DBCO-MMAE; group 3: MMAE; group 4: PBS; n = 7 or 8). AAMCHO (60 mg/kg) was i.v. injected to group 1 mice once daily for three consecutive days (Day 0, 1, 2). DBCO-MMAE (20 mg/kg) was i.v. injected on Days 3.

Anti-cancer efficacy study against 4T1 lung metastases in BALB/c mice. 4T1 metastatic cancer was established by tail vein injection of luciferase-engineered 4T1 cells (1 $\times$ 10⁵ cells in 200 μ L HBSS) into 6-week-old BALB/c mice on Day 0. Mice were then randomly divided into five groups (group 1: AAMCHO + DBCO-MMAE; group 2: DBCO-MMAE; group 3: MMAE; group 4: AAMCHO; group 5: PBS; n = 7 for each group). AAMCHO (60 mg/kg) was i.v. injected once daily for 3 d (Day 2, 3, and 4). DBCO-MMAE (8 mg/kg, i.v.) or MMAE (0.25 mg/kg, i.p., maximum tolerated dose) was injected on day 5. Each mouse's body weight and food intake were measured every other day. Lung metastases were monitored via bioluminescence imaging of BALB/c mice using the Bruker In Vivo Xtreme imaging system every 4 days starting from Day 4. For bioluminescence imaging, D-luciferin potassium salt (150 mg/kg) was intraperitoneally injected 3 min before imaging. Bioluminescence imaging data were processed using the Bruker MI SE imaging software. After the last imaging on Day 13, all mice were sacrificed under anesthesia. The lung and heart of each animal were resected as a whole, weighed, and injected with 10% formalin into trachea until the lungs inflated. Tumor nodules on the lungs were counted under a dissecting microscope. All lung tissues were paraffin embedded, sectioned with a thickness of 5 μ m, and stained with hematoxylin and eosin (H&E). All the lung sections were then scanned and analyzed. The surface area of tumors and lungs were measured to calculate the percentage of tumor surface area over the total lung surface area ($A_{\text{tumor}}/A_{\text{total}}$).

Long-term efficacy study of AAMCHO + DBCO-MMAE against LS174T-R tumors. Multidrug-resistant LS174TR tumors were established in 6-week-old female athymic nude mice by subcutaneous injection of LS174TR cells (1.5×10^6 cells in HBSS/Matrigel (1/1, v/v, 50 μ l)) into both flanks. When the tumors reached ~ 200 mm³, mice were randomly divided into four groups (group 1: AAMCHO + DBCO-MMAE; group 2: DBCO-MMAE; group 3: Oxaliplatin; group 4: PBS; n = 7 or 8). For group 1, AAMCHO (60 mg/kg) was i.v. injected on Days 0, 1, 2, followed by i.v. injection of DBCO-MMAE (8 mg/kg) at Day 3. For group 2, i.v. injection of DBCO-MMAE (8 mg/kg) was conducted at Day 3. For group 3, i.p. injection of Oxaliplatin (5 mg/kg, maximum tolerance dose) was conducted at Day 3. For Group 4, PBS was i.v. injected as control. Tumor volume and body weight of mice were measured every other day. The tumor volume was calculated using the formula (length) $\times$ (width)² / 2, where the long axis diameter was regarded as the length, and the short axis diameter was regarded as the width. An animal exited the study due to tumor volume or treatment-related death, the final tumor volume recorded for the animal was used to calculate the mean tumor volume at a subsequent time. The tumors were bisected after sacrifice, minced, and digested with 1 mg/mL Collagenase (Type IV, Gibco) digestion buffer. Suspensions were filtered through 70 μ m cell strainers, permeabilized with Triton-X100 (0.1%), and were counterstained with Oct-4 or Sox-2 antibodies separately, followed by staining with FITC-conjugated IgG antibody. The cell suspension was washed by centrifugation and subjected to flow cytometry analysis.

Histopathology evaluation. After the above-mentioned anti-cancer efficacy study against 4T1 lung metastasis, the liver, heart, kidneys, spleen, and spinal cord (cervical, thoracic, and lumbar) were fixed in formalin, paraffin-embedded, sectioned with a thickness of 5 μ m, and stained with H&E. Tissues were analyzed by a board-certified pathologist to investigate treatment-mediated toxicity.

Statistical analysis. Statistical analysis was performed using GraphPad Prism v6 and v8. Sample variance was tested using the F test. For samples with equal variance, the significance between the groups was analyzed by a two-tailed student's t-test. For samples with unequal variance, a two-tailed Welch's t-test was performed. For multiple comparisons, a one-way analysis of variance (ANOVA) with a post hoc Fisher's LSD test was used. The results were deemed significant at $0.01 < *P \leq 0.05$, highly significant at $0.001 < **P \leq 0.01$, and extremely significant at $0.01 < ***P \leq 0.001$, $****P < 0.0001$.

Supplementary Figures

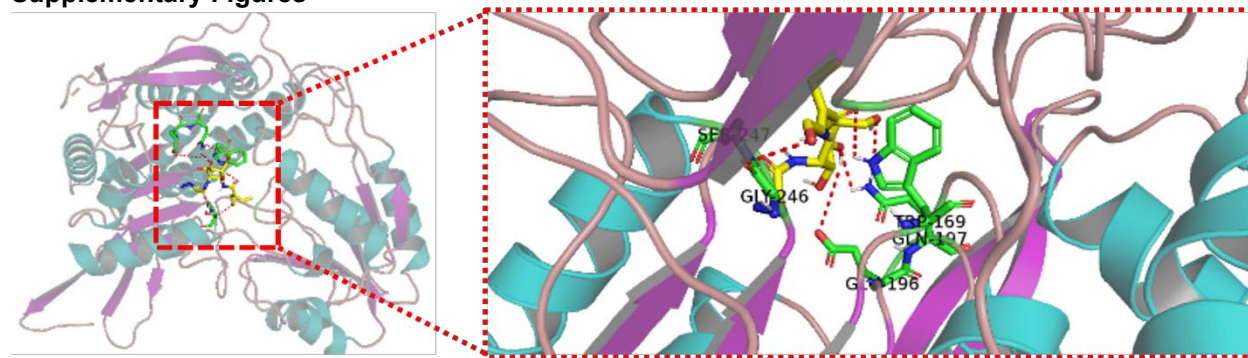


Fig. S1. Unbiased docking of ALDH1A1 and AAMCHO was performed using the AutoDock Vina. All non-peptidic residues in the ALDH1A1 crystal structure (PDB ID: 4WB9) were removed prior to docking. The hydrogen bonds between ALDH1A1 and deacetylated AAMCHO were shown in red lines.

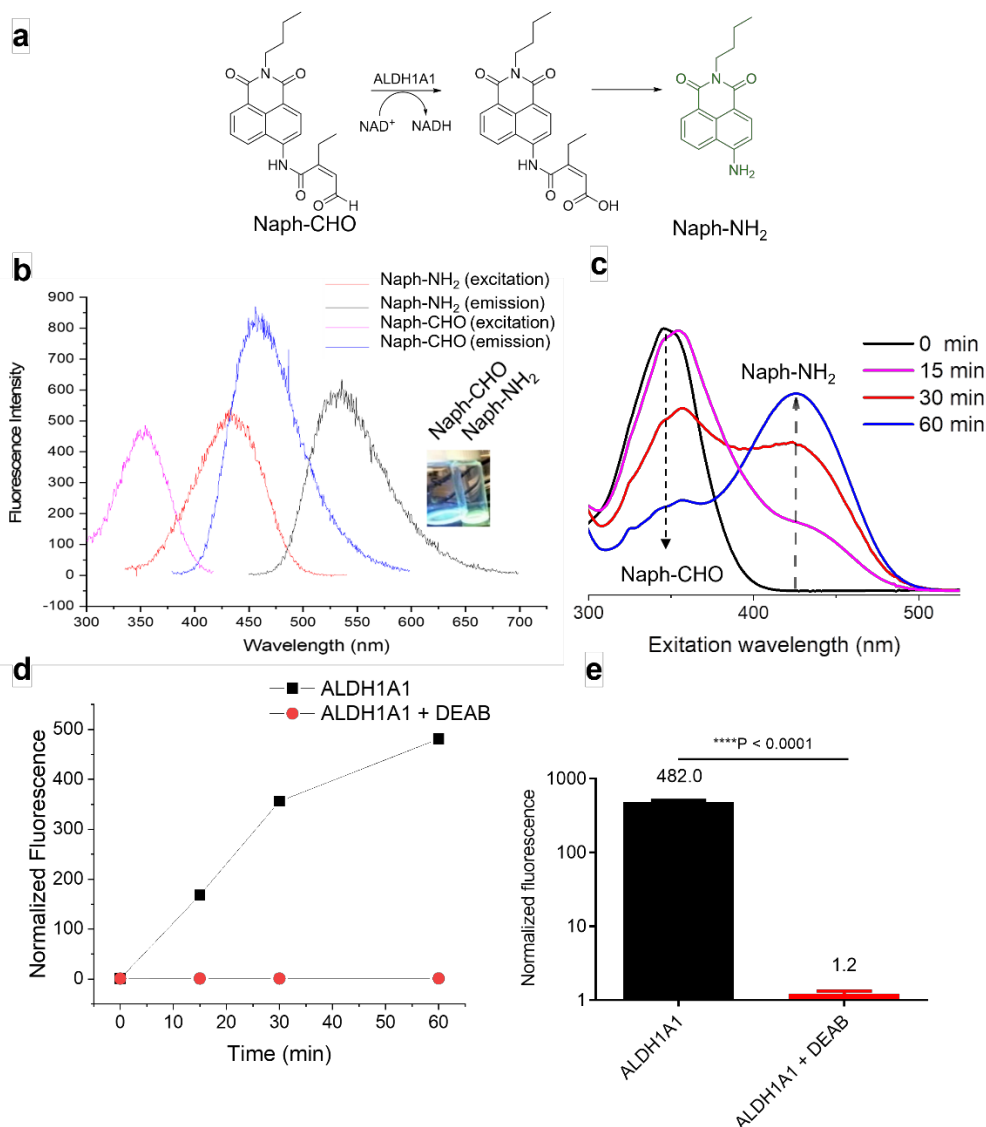


Fig. S2. Naph-CHO is developed as a fluorescence reporter for ALDH1A1 and verifies the ALDH1A1-responsive moiety. (a) Scheme of ALDH1A1-mediated degradation of Naph-CHO into Naph-NH₂. (b) Excitation and emission spectra of Naph-CHO (5 µg/mL in methanol) and Naph-NH₂ (200 ng/mL in methanol). (c-e) 1 µM Naph-CHO was incubated with 100 nM ALDH1A1. Reactions were conducted with 2.5 mM NAD⁺ and 50 mM TEA (pH 7.4) at room temperature. (c) Excitation spectra of the reaction solution at 0, 15, 30, and 60 min, respectively. (d) Normalized fluorescence intensity of Naph-CHO (1 µM) incubated with ALDH1A1 (100 nM) in the presence or absence of ALDH inhibitor 4-diethylaminobenzaldehyde (DEAB, 100 nM). λ_{Ex}: 425 nm, λ_{Em}: 550 nm. (e) Fluorescence intensity of Naph-CHO after treatment with ALDH1A1 in the presence or absence of DEAB (100 nM) for 60 min (n=3). λ_{Ex}: 425 nm, λ_{Em}: 550 nm. All the numerical data are presented as mean ± SD (0.01 < *P ≤ 0.05; 0.001 < **P ≤ 0.01; 0.0001 < ***P ≤ 0.001, ****P ≤ 0.0001).

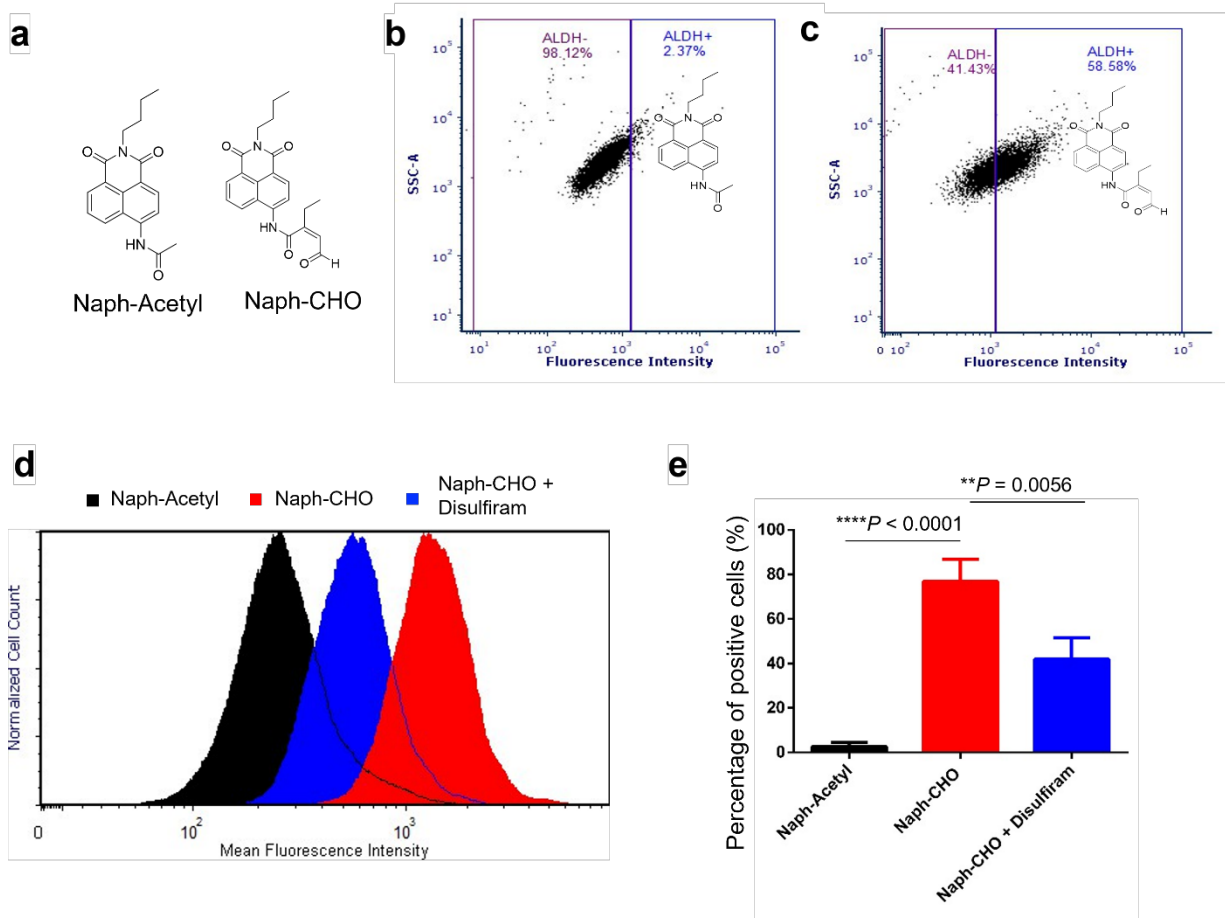


Fig. S3. Naph-CHO mediates the detection of ALDH1A1⁺ cells. (a) Structure of ALDH1A1-responsive Naph-CHO and non-responsive Naph-Acetyl. (b-c) Representative flow cytometry plot showing the staining of K562 cells by (b) Naph-Acetyl or (c) Naph-CHO. (d) Flow cytometry Histogram of K562 cells stained with Naph-CHO for 24 h. Naph-Acetyl was used as a control. Disulfiram is an inhibitor of ALDH1A1. (e) Percentage of fluorescence-positive K562 cells after 24-h treatment with Naph-Acetyl, Naph-CHO, or Naph-CHO + disulfiram (n=3). All the numerical data are presented as mean $\pm$ SD ($0.01 < *P \leq 0.05$; $0.001 < **P \leq 0.01$; $0.0001 < ***P \leq 0.001$, $****P \leq 0.0001$).

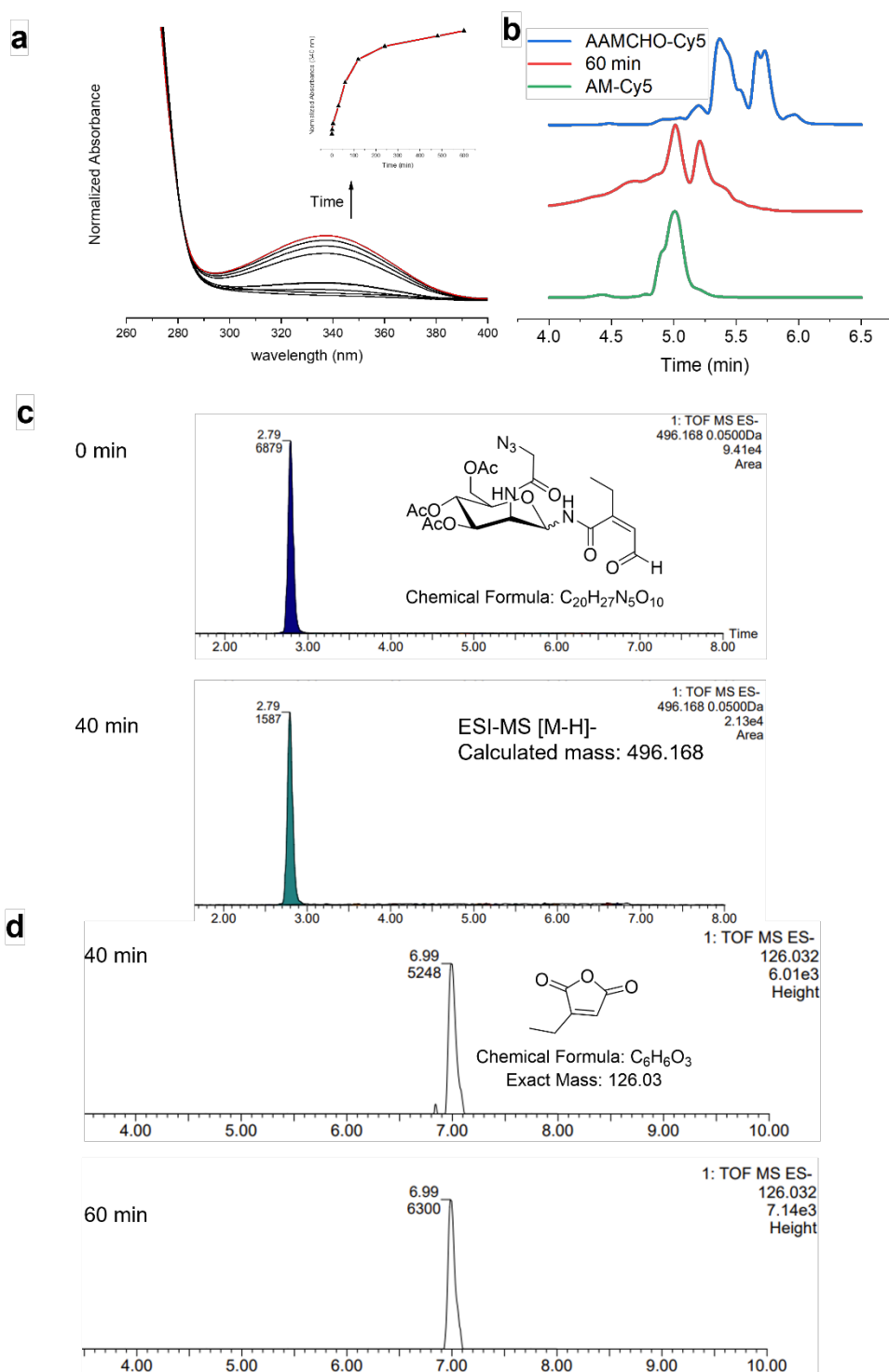


Fig. S4. LC-MS validates the degradation of AAMCHO in the presence of ALDH1A1. (a) ALDH1A1 enzymatic activity validation. Propionaldehyde was incubated with recombinant ALDH1A1 and NAD^+ at room temperature. UV absorbance at 340 nm over time indicates the formulation of NADH. (b) HPLC traces show degradation of AAMCHO in the presence of ALDH1A1. The HPLC is coupled with a fluorescence detector (excitation: 650 nm emission: 688 nm). At selected time points, 100 μ L aliquot was collected, quenched by acetonitrile/methanol, and conjugated with DBCO-Cy5, prior to HPLC measurement. The

conjugate between as-synthesized AM and DBCO-Cy5 was used as a reference. (c) LC-MS traces of AAMCHO were detected at a mass value of 496.168 Da, after treatment with ALDH1A1 for 0 and 40 min, respectively. (d) LC-MS traces of degradation byproduct were detected at a mass value of 126.032 Da, after treatment with ALDH1A1 for 40 and 60 min, respectively.

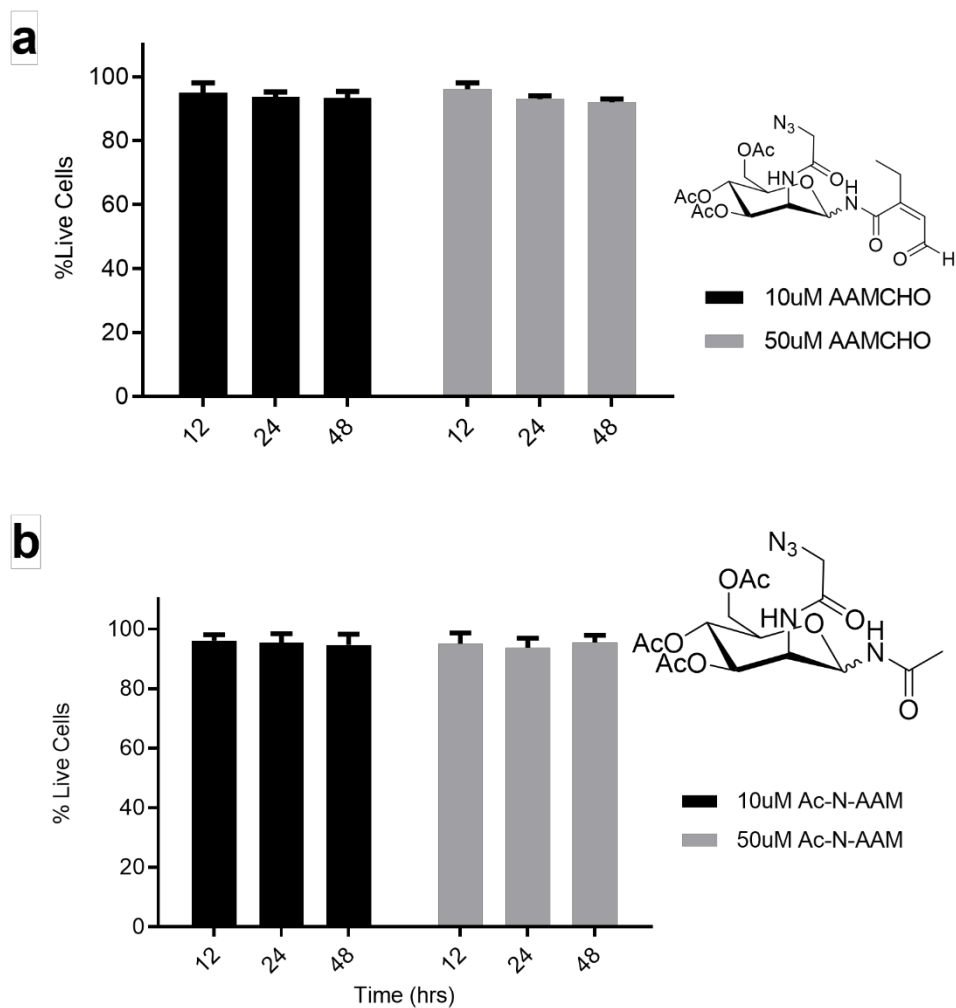


Fig. S5. AAMCHO and Ac-N-AAM show minimal cytotoxicity. Shown are the viability of K562 cells treated with (a) AAMCHO or (b) Ac-N-AAM for 12, 24, and 48 h, respectively (n=3). All the numerical data are presented as mean $\pm$ SD.

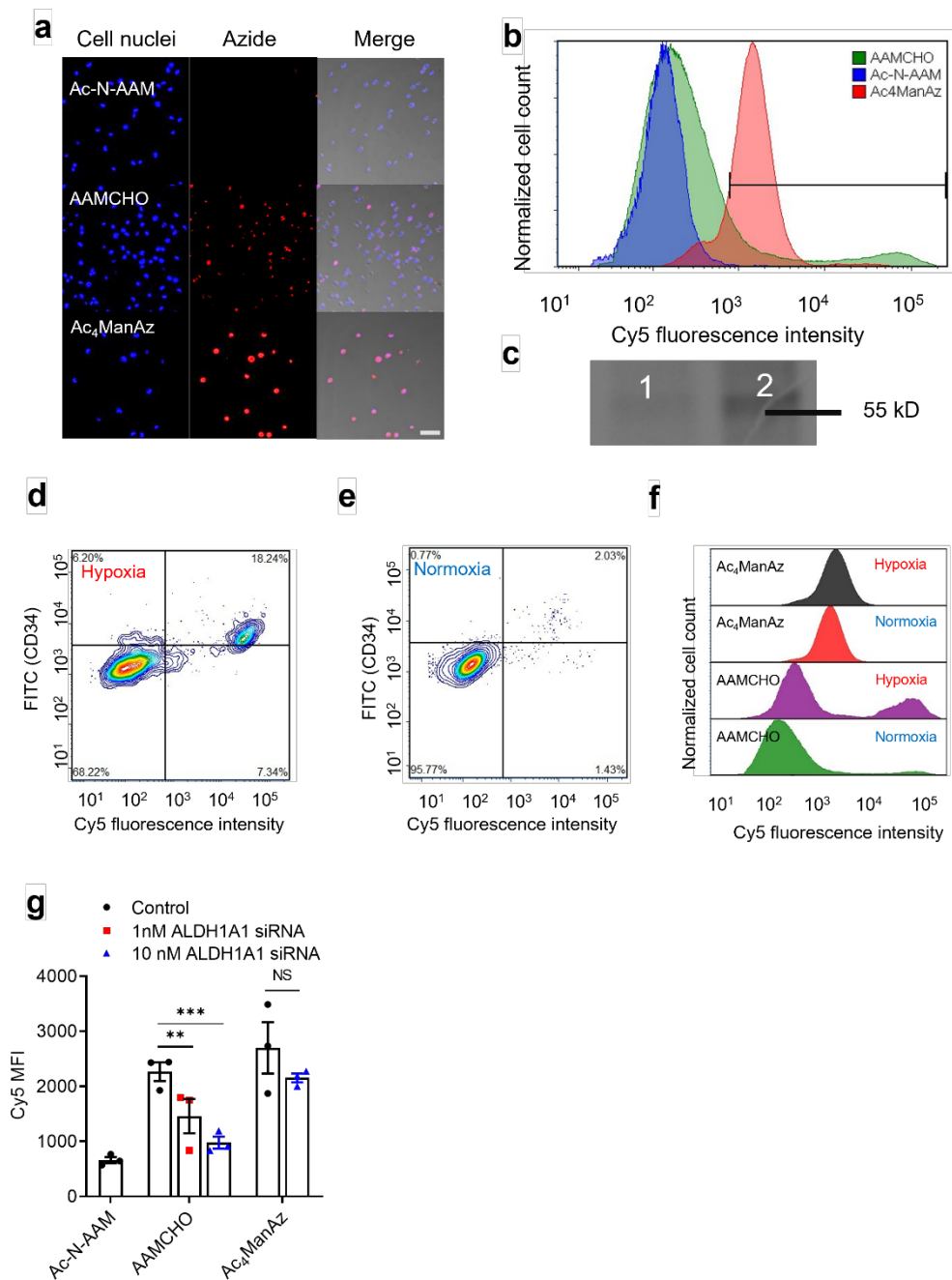


Fig. S6. AAMCHO metabolically label cancer cells with azide groups in an ALDH1A1-dependent manner. (a) Representative CLSM image of K562 cells treated with different azide-sugars for three days and incubated with DBCO-Cy5 for 1 h. Scale bar: 100 μ m. (b) Representative flow cytometry histogram of K562 cells labeled with azide-sugars and detected with DBCO-Cy5. (c) Western blot analysis of ALDH1A1 expression in K562 cells treated with (1) 10nM siRNA specific for ALDH1A1 or (2) scrambled siRNA. (d-f) K562 cells were cultured with AAMCHO under hypoxia (1% O₂) or normoxia (23% O₂) conditions for three days and stained with DBCO-Cy5 and FITC-conjugated anti-CD34. CD34⁺ cells were pre-gated for analysis. (d, e) Representative flow cytometry plots of K562 cells labeled with AAMCHO cultured under (d) hypoxia or (e) normoxia conditions. (f) Representative Cy5 fluorescence histograms of K562 cells after 3-day hypoxia or normoxia. (g) Bar graph of Cy5 MFI for Control, 1nM ALDH1A1 siRNA, and 10 nM ALDH1A1 siRNA across Ac-N-AAM, AAMCHO, and Ac₄ManAz treatments.

treatment with azide-sugars and probed with DBCO-Cy5. **(g)** Cy5 mean fluorescence intensity of K562 cells after 72-h incubation with different azide-sugars in the presence of ALDH1A1 siRNA under hypoxia conditions. Cells were stained with DBCO-Cy5 for 1 h for azide detection. All the numerical data are presented as mean $\pm$ SD ($0.01 < *P \leq 0.05$; $0.001 < **P \leq 0.01$; $0.0001 < ***P \leq 0.001$, $****P \leq 0.0001$).

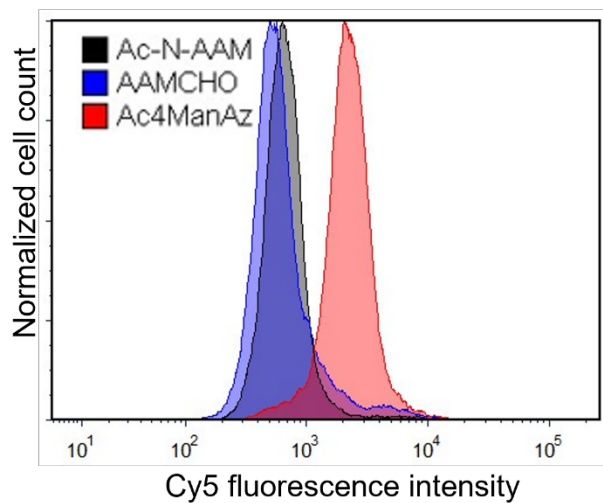
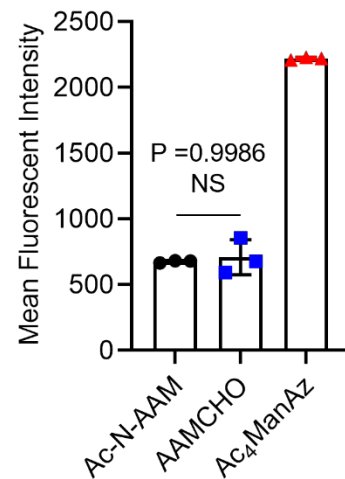
a**b**

Fig. S7. ALDH1A1-negative HEK293T cells exhibit minimal labeling by AAMCHO. (a) Representative flow cytometry histogram of ALDH1A1-negative HEK293T cells treated with different azide-sugars and probed with DBCO-Cy5. (b) Mean Cy5 fluorescence intensity of ALDH1A1-negative HEK293T cells after pretreated with azide-sugars for three days and incubated with DBCO-Cy5 for 1 h. All the numerical data are presented as mean $\pm$ SD ($0.01 < *P \leq 0.05$; $0.001 < **P \leq 0.01$; $0.0001 < ***P \leq 0.001$, $****P \leq 0.0001$).

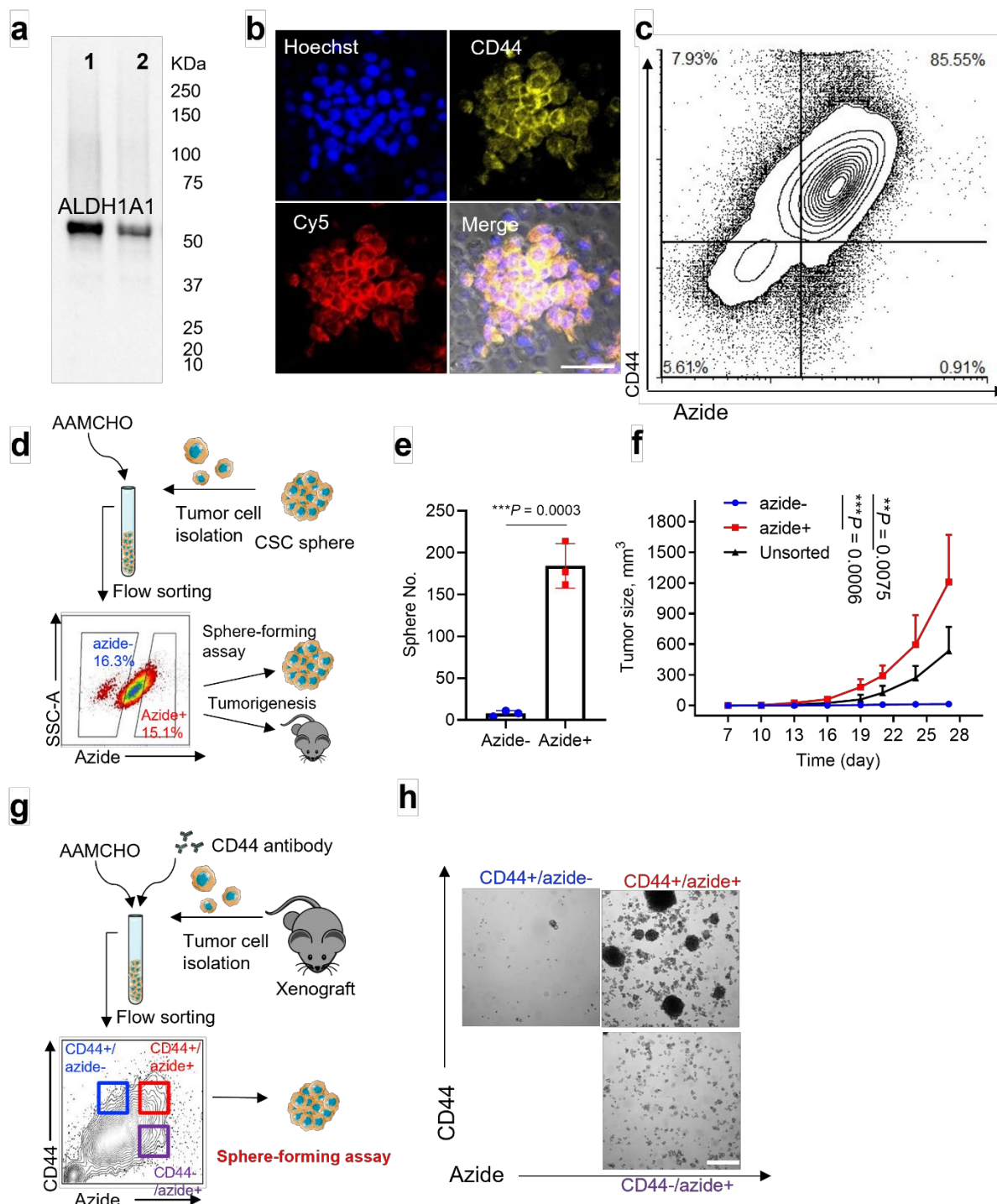


Fig. S8. AAMCHO specifically labels CSCs, and the labeled population exhibits self-renewal properties *in vitro* and *in vivo*. (a) Western blot analysis of ALDH1A1 expressed by MDA-MB-231 cells. 1: MDA-MB-231 cells isolated from CSC spheres cultured under serum-free conditions. 2: Regular MDA-MB-231 cells. (b) CLSM images of MDA-MB-231 spheres labeled with AAMCHO and detected with DBCO-Cy5 and FITC conjugated anti-CD44. (c) Representative plot and quadrants quantification of MDA-MB-231 spheres labeled with AAMCHO and detected with DBCO-Cy5 and FITC conjugated anti-CD44. (d-f) Single-cell suspension was isolated from the MDA-MB-231 CSC sphere, labeled with AAMCHO + DBCO-Cy5.

Azide- and azide+ populations were sorted and characterized by *in vitro* sphere-forming assay or inoculated subcutaneously into mice for tumorigenesis test. **(d)** Sorting strategy for azide+ and azide- subpopulations. **(e)** Sphere forming assays of different sub-populations. Spheres were counted on Day 7 (n = 3). **(f)** Tumorigenesis test: Tumor volume of MDA-MB-231. 10k cells were injected into the mammary fat pad of nude mice (n = 6) xenografts derived from different sub-populations (unsorted, azide-, and azide+) over time. Statistical comparison on Day 27 is shown. **(g-h)** Single-cell suspension was isolated from MDA-MB-231 xenograft, labeled with AAMCHO + DBCO-Cy5 and FITC conjugated CD44 antibody. Three populations were sorted and characterized by *in vitro* sphere-forming assay. **(g)** Schematics of the experimental design. **(h)** Representative image of the spheres formed from different subpopulations. Scale bar, 100 μ m. All the numerical data are presented as mean $\pm$ SD (0.01 < *P $\leq$ 0.05; 0.001 < **P $\leq$ 0.01; 0.0001 < ***P $\leq$ 0.001, ****P $\leq$ 0.0001).

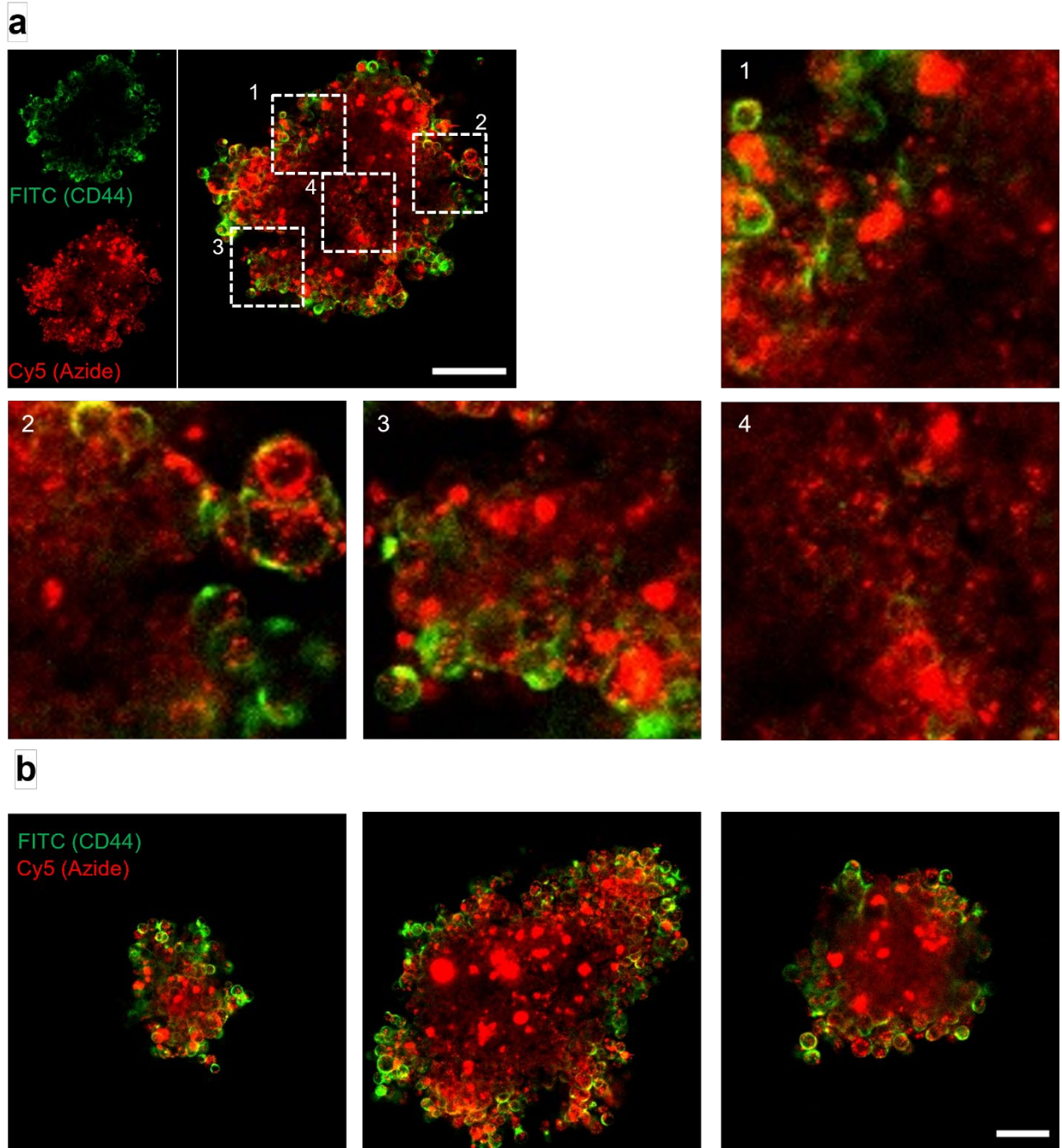


Fig. S9. AAMCHO + DBCO-Cy5 exhibits deeper penetration than anti-CD44 antibodies in tumorspheres. (a-b) CLSM image of MDA-MB-231 tumorsphere after a 3-day labeling with AAMCHO and 1-h incubation with DBCO-Cy5 and FITC conjugated anti-CD44. Scale bar, 50 μ m. (a) Zoomed-in images of the barriers (1,2,3) and the inner areas (4) of the tumorspheres are shown. (b) Three representative images MDA-MB-231 tumorsphere after a 3-day labeling with AAMCHO and 1-h incubation with DBCO-Cy5 and FITC conjugated anti-CD44. Scale bar, 50 μ m.

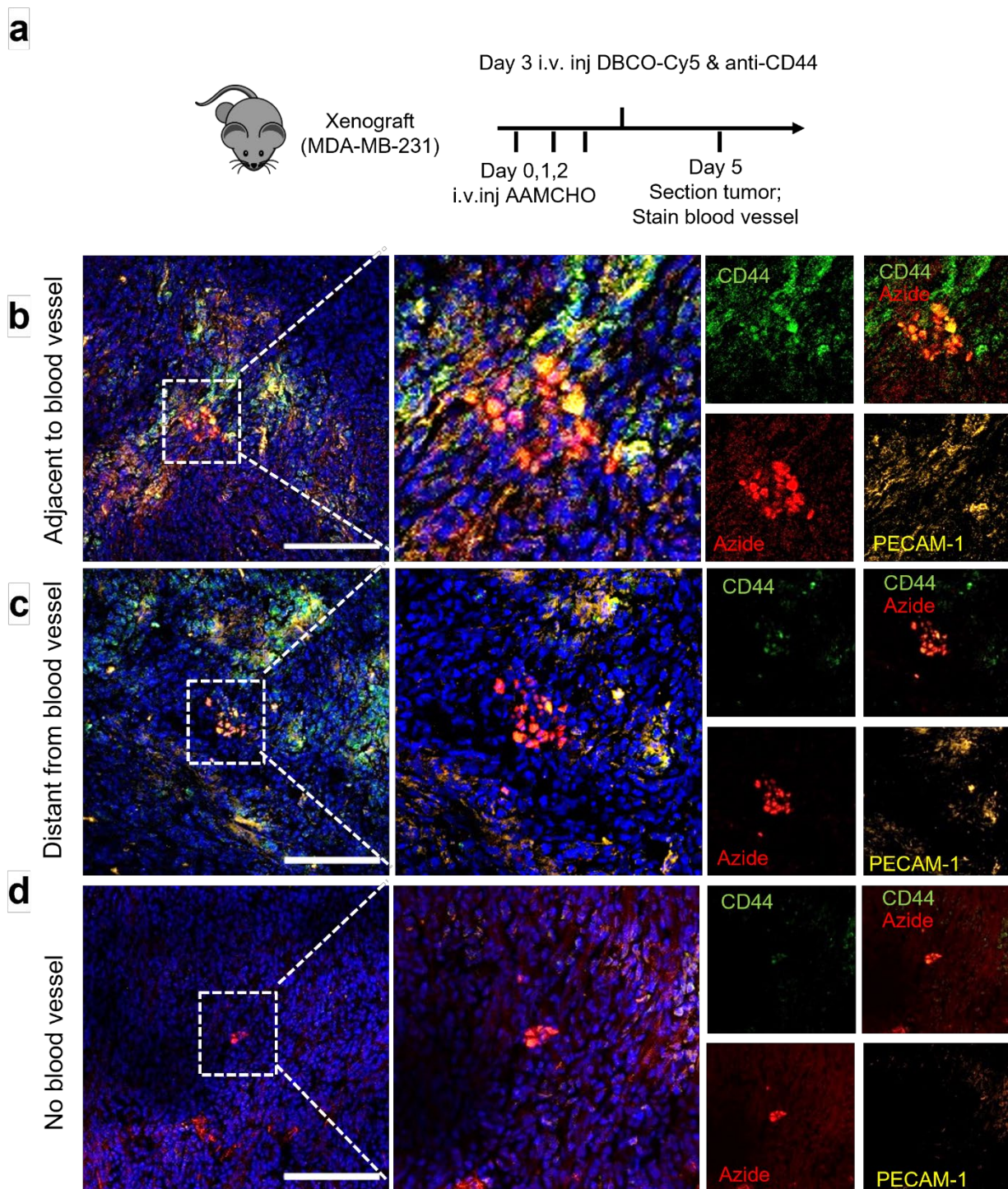


Fig. S10. Micro-distribution of DBCO-Cy5 and FITC conjugated anti-CD44 in azide-labeled MDA-MB-231 tumors. (a-d) After MDA-MB-231 Xenograft reached $\sim 50 \text{ mm}^3$, AAMCHO was intravenously injected on days 0, 1, and 2, followed by intravenous injection of DBCO-Cy5 and FITC conjugated anti-CD44 on day 3. Tumors were sectioned and stained on day 5. (a) Schematics of the experimental design. (b-d) Shown are CLSM images of labeled CSCs (b) adjacent to, (c) distant from, or (d) without blood vessels. The blood vessels were stained with platelet endothelial cell adhesion molecule 1 (PECAM-1) antibody and

CFL-555 secondary antibody (yellow). The cell nuclei were stained with Hoechst 33342. Regions of interest are zoomed in for analyzing the distribution of Cy5 and FITC signals. Scale bar, 200 μm .

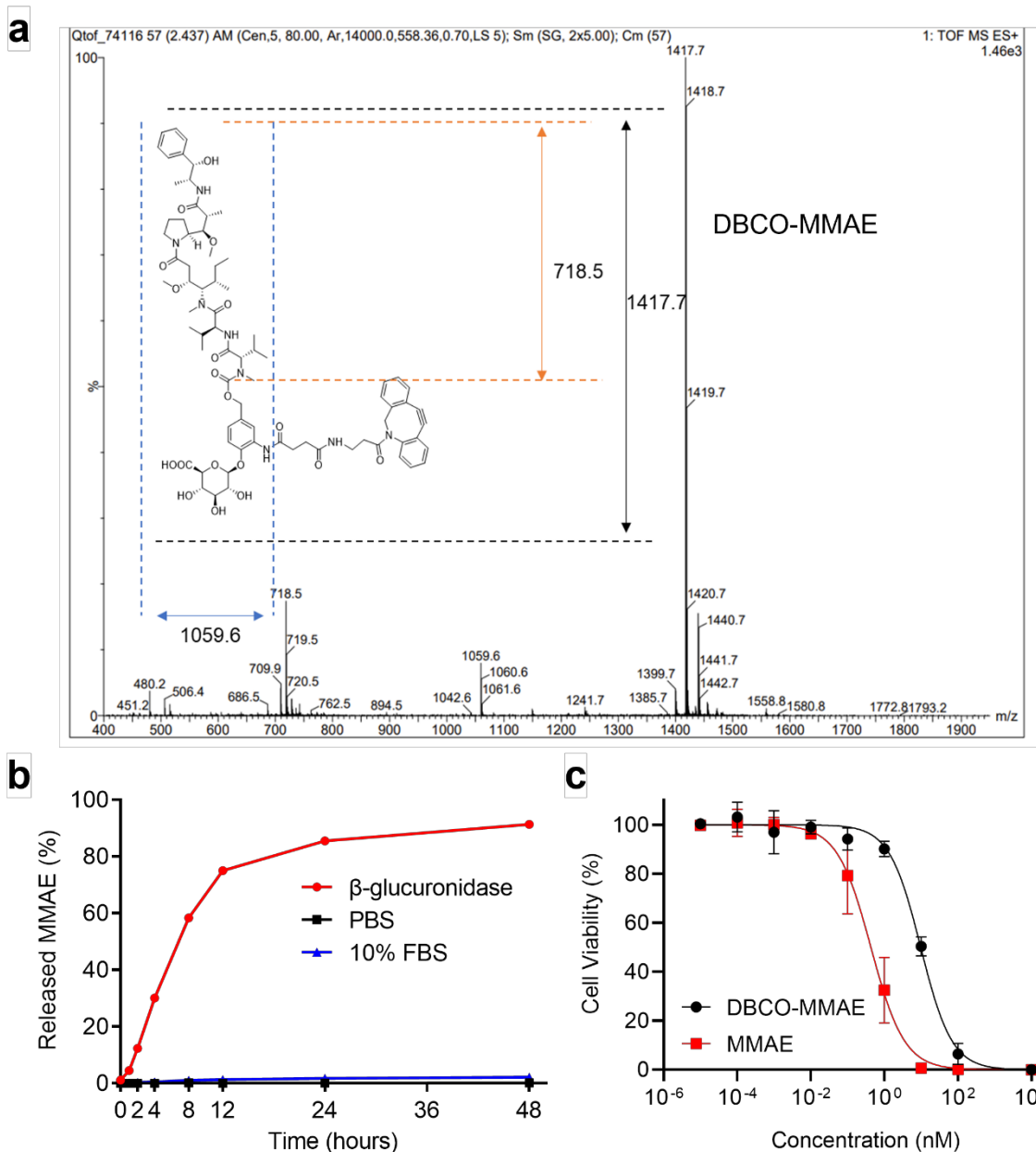
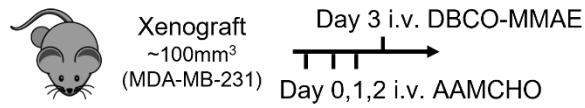
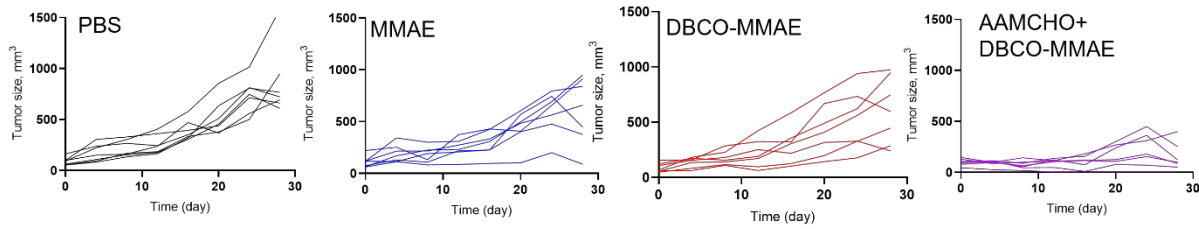


Fig. S11. Synthesis and Characterization of DBCO-MMAE. (a) The ESI-Mass spectrum of DBCO-MMAE. (b) The release profile of DBCO-MMAE in the presence of activated β -glucuronidase, PBS, and 10% FBS, respectively. (c) Cytotoxicity of DBCO-MMAE and MMAE against MDA-MB-231 cells. All the numerical data are presented as mean $\pm$ SD.

a**b****c**

Group A vs Group B		Day						
		4	8	12	16	20	24	28
AAMCHO+DBCO-MMAE	PBS	n.s.	*	*	**	**	***	***
DBCO-MMAE	PBS	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	*
MMAE	PBS	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	*
DBCO-MMAE	MMAE	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
AAMCHO+DBCO-MMAE	DBCO-MMAE	n.s.	n.s.	*	*	*	**	***
AAMCHO+DBCO-MMAE	MMAE	n.s.	n.s.	*	*	**	***	***

Fig. S12. AAMCHO mediates targeted delivery of DBCO-MMAE and results in significantly improved anti-tumor efficacy against MDA-MB-231 TNBC. (a) Timeframe of the study. (b) Tumor volume of individual mice in each group. (c) Statistical comparisons of the tumor volume over time. ($0.01 < *P \leq 0.05$; $0.001 < **P \leq 0.01$; $0.0001 < ***P \leq 0.001$, $****P \leq 0.0001$).

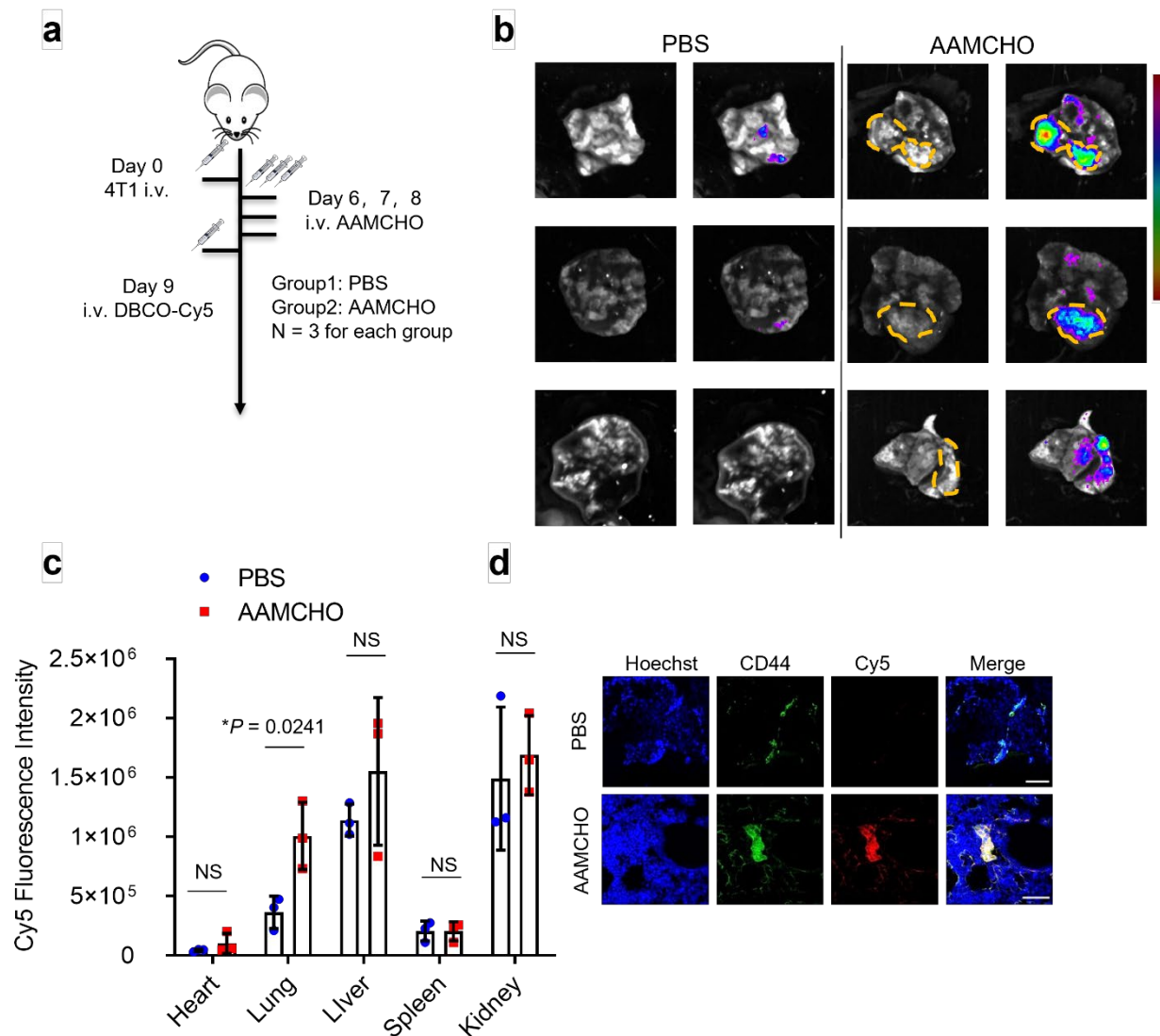


Fig. S13. AAMCHO can label metastatic 4T1 TNBC colonies *in vivo*. (a) Timeframe of *in vivo* labeling study. 4T1 metastatic cancer was established by tail vein injection of 4T1 cells (1×10^5 cells in 200 μ l HBSS) into BALB/c mice on day 0. AAMCHO was i.v. injected once daily for 3 days (days 6, 7, and 8). DBCO-Cy5 was i.v. injected on day 9. (b) *Ex vivo* cy5 fluorescence images of harvested lungs. Metastatic tumor nodules are marked in orange. Fluorescence images were overlaid with brightfield images. (c) Average Cy5 fluorescence intensity of different tissues ($n = 3$). (d) CLSM images of lung tissue sections. Tissue sections were stained with FITC-conjugate anti-CD44 (Green) and Hoechst 33342 (blue). Scale bar, 50 μ m. All the numerical data are presented as mean $\pm$ SD ($0.01 < *P \leq 0.05$; $0.001 < **P \leq 0.01$; $0.0001 < ***P \leq 0.001$, $****P \leq 0.0001$).

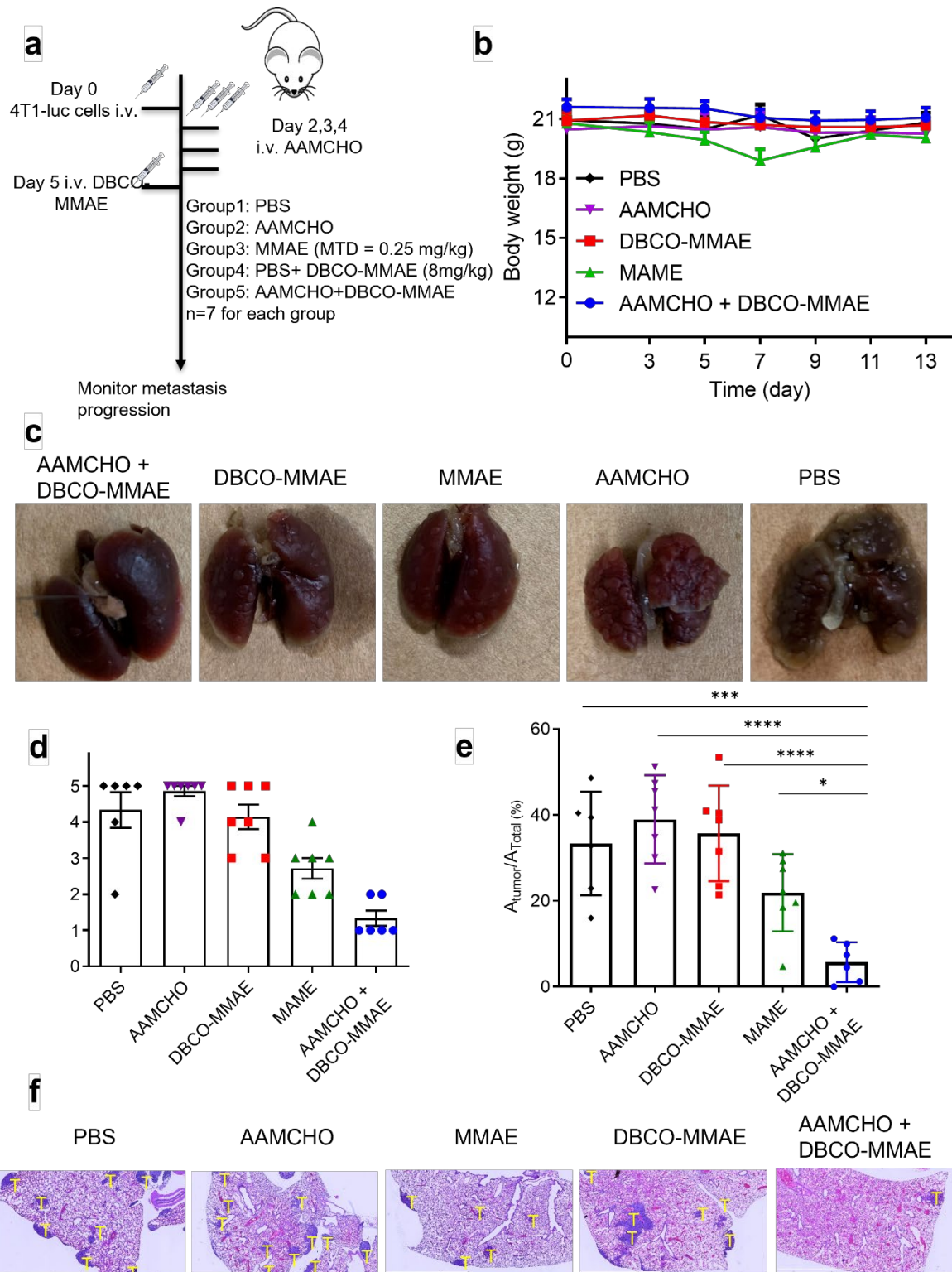


Fig. S14. AAMCHO + DBCO-MMAE inhibits the growth of metastatic 4T1 TNBC. (a) Timeframe of the efficacy study. (b) Body weight of mice over the course of the efficacy study. (c) Lung pictures at necropsy are shown for each treatment group on day 13. (d) Scaling of the lung tissues with metastatic tumors. Lungs were excised and scaled based on the severity of metastases. Lungs with over 90% surface occupied by tumors were scored 5; lungs with 70-90% surface occupied by tumors scored 4; lungs with 50-70% surface

occupied by tumors scored 3; lungs with 30-50% surface occupied by tumors scored 2; lungs with less than 30% occupied by tumors scored 1. **(e)** Percentage of tumor surface area over total lung tissue area for each group (n=7). **(f)** Representative images of H&E-stained lung tissues from each group. Fixed lung tissues were embedded in paraffin, sectioned into 5- μ m thickness, and stained with H&E. T: metastatic tumor nodule. Scale bar, 200 μ m. All the numerical data are presented as mean $\pm$ SD (0.01 < *P $\leq$ 0.05; 0.001 < **P $\leq$ 0.01; 0.0001 < ***P $\leq$ 0.001, ****P $\leq$ 0.0001).

a

Organs	Histopathological lesions	PBS	AAMCHO	MMAE	DBCO-MMAE	AAMCHO + DBCO-MMAE
Spleen	Lymphoid Depletion	-	-	+ / +++ (5/2)	- / + (4/3)	-
	Extramedullary Hematopoiesis (Increased)	- / + (6/1)	- / + (5/2)	+ / +++ (5/2)	- / + (5/2)	- / + (4/3)
Liver	Single Cell Necrosis	-	-	- / + (1/6)	- / + (6/1)	- / + (5/2)
	Biliary Hyperplasia	-	- / + (6/1)	- / + (2/5)	- / + (6/1)	- / + (6/1)
	Portal Triaditis	-	- / + (5/2)	-	- / + (6/1)	- / + (6/1)

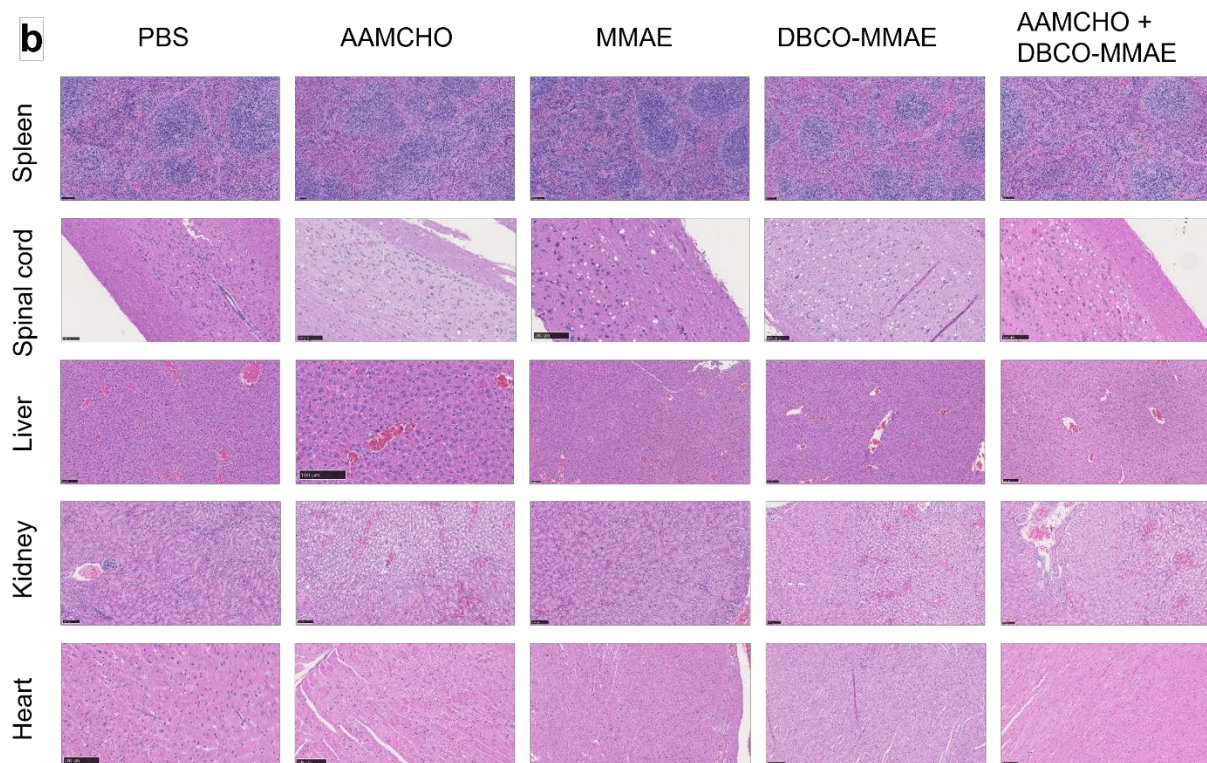
b

Fig. S15. DBCO-MMAE shows significantly reduced acute toxicity compared to MMAE. (a) A summary of the pathological diagnostics observed in different groups. A dose of 0.25 mg/kg was used for MMAE, while 8 mg/kg was used for DBCO-MMAE. Grade criteria: + (mild); +++ (marked); - (negative). (b) Representative images of H&E-stained tissues from each group. Fixed tissues were embedded in paraffin, sectioned into 5- μ m thickness, and stained with H&E.

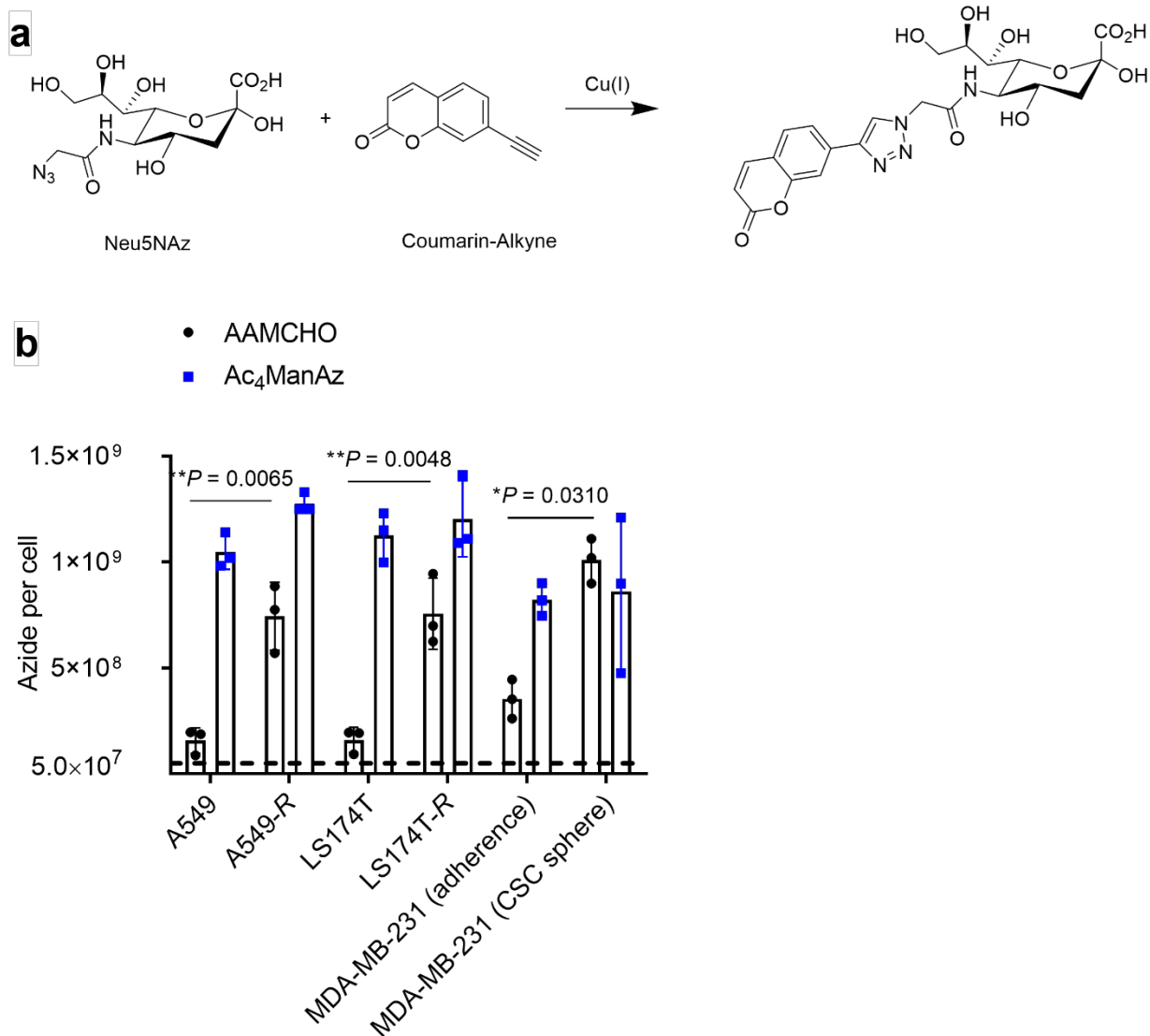


Fig. S16. Quantitative analysis of azide-sialic acid expression of cells labeled by AAMCHO. (a) Quantification of azide-sialic acid (Neu5Az) in cell or tissue lysate by reacting them with coumarin-alkyne via Cu(I) catalyzed Click reaction and quantifying coumarin-bearing compounds using HPLC equipped with a fluorescence detector (λ_{Ex} : 328 nm, λ_{Em} : 415 nm). The Neu5NAz amount is normalized by cell counting. (b) Number of azide groups per cell after treatment with AAMCHO for 72 h (n=3). All the numerical data are presented as mean $\pm$ SD (0.01 < *P $\leq$ 0.05; 0.001 < **P $\leq$ 0.01; 0.0001 < ***P $\leq$ 0.001, ****P $\leq$ 0.0001).

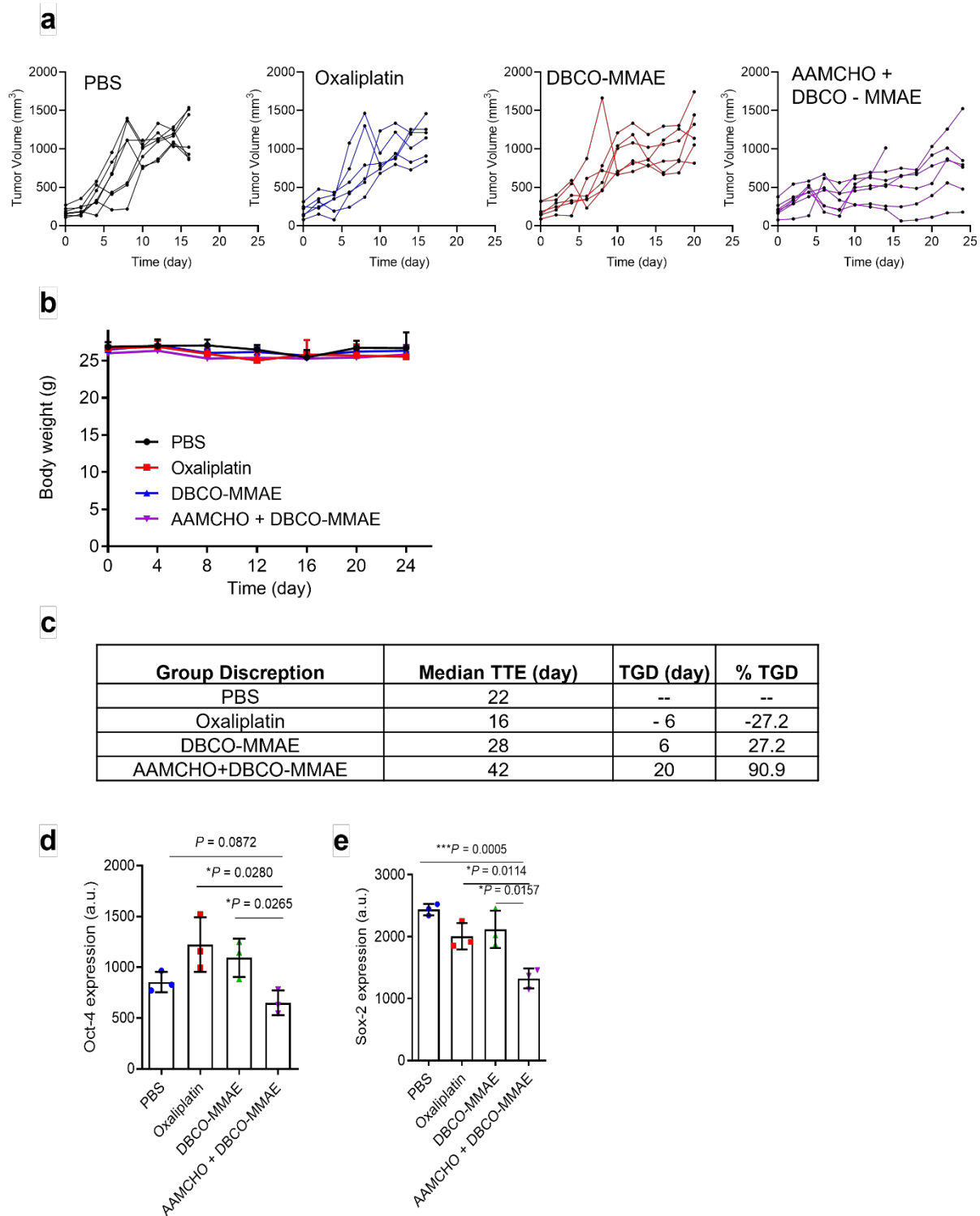


Fig. S17. AAMCHO-mediated targeted delivery of DBCO-MMAE improves the anticancer efficacy against LS174T-R tumors while downregulating the markers of CSC. (a) Tumor volume of individual mice in each group. (b) Body weight of mice from different groups over the course of the efficacy study. (c) Survival analysis of mice from different groups. TTE: time to the endpoint. TGD: tumor growth delay; TGD = TTE (treated group) – TTE (PBS group). %TGD = 100% × TGD/TTE (PBS group). (d-e) Relative

expressions of the (d) Oct-4 and (e) Sox-2 in the tumor collected from the mice. All the numerical data are presented as mean $\pm$ SD (0.01 < *P $\leq$ 0.05; 0.001 < **P $\leq$ 0.01; 0.0001 < ***P $\leq$ 0.001, ****P $\leq$ 0.0001).

General procedure

1) NaOMe/MeOH
ClCH₂COOCH₂COCl
TEA/MeOH

2) NaN₃
H₂O/MeOH, 65 °C

3) Ac₂O/DMAP, Py

1) PMe₃/THF, THF
2) H₂O, rt

ClCH₂COOCH₂COCl
Pyridine, THF, 0 °C

Ph₃PCH₂CHO
0 °C to r.t.

ManNH₂.HCl

1

2

3

4

AAMCHO

Scheme S1. Synthesis of **AAMCHO**.

To a solution of D-mannosamine hydrochloride (2.00 g, 9.28 mmol) in methanol (10 mL) was dropwise added 0.5 N NaOMe in MeOH (18.6 mL, 9.28 mmol) at 0 °C. The mixture was stirred at room temperature for 30 min. Then triethylamine (0.94 g, 1.3 mL, 9.28 mmol) and chloroacetic anhydride (1.90 g, 11.14 mmol, 95%) were added at 0 °C. The reaction mixture was stirred at room temperature overnight and the was monitored with TLC (ethyl acetate/MeOH/H₂O = 2/2/1, one drop ammonium hydroxide solution). Then H₂O (3 mL) and sodium azide (2.41 g, 37.12 mmol) were added at room temperature. After stirring overnight at 65 °C, the reaction was monitored by TLC (CH₂Cl₂/MeOH = 5/1, one drop water). The solid was filtered out and the filtrate was concentrated. The residue was suspended in pyridine (15 mL), then DMAP (113.6 mg, 0.93 mmol) and acetic anhydride (7.58 g, 7 mL, 74.24 mmol) were added at 0 °C. The reaction mixture was stirred at room temperature overnight and the progress was monitored with TLC (CH₂Cl₂/MeOH = 5/1, one drop water). The reaction was quenched with methanol and stirred at room temperature for 30 min. After concentration, the residue was dissolved in EtOAc (50 mL) and washed successively with HCl (1 M), brine and sat. NaHCO₃(aq). The organic layer was dried over Na₂SO₄ and filtered. The filtrate was concentrated, and the crude product was purified by column chromatography (hexane/ethyl acetate 1:1) to afford the desired product 1 (2.60 g, 65% yield, mixture of α/β anomers) as white foam.

$R_f = 0.40$ (hexane/ethyl acetate = 1/2).

¹H NMR (500 MHz, CDCl₃) δ 6.64 (d, *J* = 9.5 Hz, 0.75H), 6.56 (d, *J* = 8.0 Hz, 1H), 6.04 (d, *J* = 1.5 Hz, 1H), 5.89 (d, *J* = 1.5 Hz, 0.75H), 5.34 (dd, *J* = 10.5, 4.5 Hz, 1H), 5.22 (t, *J* = 10.0 Hz, 1H), 5.16 (t, *J* = 9.5 Hz, 0.75H), 5.06 (dd, *J* = 9.5, 3.5 Hz, 0.75H), 4.73 (ddd, *J* = 9.5, 3.5, 1.5 Hz, 0.75H), 4.62 (ddd, *J* = 9.0, 4.0, 2.0 Hz, 1H), 4.25 (dd, *J* = 7.5, 4.5 Hz, 0.75H), 4.23 (dd, *J* = 7.5, 4.0 Hz, 1H), 4.15 (dd, *J* = 12.0, 2.0 Hz, 0.75H), 4.12 – 4.02 (m, 5.5H), 3.82 (ddd, *J* = 9.5, 4.5, 2.0 Hz, 0.75H), 2.18 (s, 3H), 2.11 (s, 7.5H), 2.06 (s, 5.25H), 2.01 (s, 2.25H), 2.00 (s, 3H).

3,4,6-tri-O-acetyl-2-azideacetamido-2-deoxy-D-mannopyranosyl azide (2):

To a solution of 1 (1.60 g, 3.72 mmol) in DCM (16 mL) were added trimethylsilyl azide (642.3 mg, 740 μ L, 5.58 mmol), SnCl₄ (193.7 mg, 87 μ L, 0.744 mmol) and AgClO₄ (154.2 mg, 0.744 mmol) at 0 °C. After being stirred overnight at room temperature, the reaction was monitored by TLC (hexane/ethyl acetate = 2/3). Then the reaction was quenched with sat. NaHCO₃(aq) and the precipitate was filtered out through a pad of Celite. The filtrate was separated, and the aqueous phase was further extracted with DCM. The combined organic layer was washed with brine, dried over Na₂SO₄, filtered, and concentrated. The crude product was purified by column chromatography on silica gel (hexane/ethyl acetate = 3:2) to afford 2 (1.37 g, 89% yield).

R_f = 0.35 (hexane/ethyl acetate = 2:3);

¹H NMR (500 MHz, CDCl₃) δ 6.61 (d, *J* = 9.0 Hz, 1H), 5.36 (d, *J* = 1.5 Hz, 1H), 5.24 (dd, *J* = 10.0, 4.0 Hz, 1H), 5.12 (t, *J* = 9.5 Hz, 1H), 4.50 (ddd, *J* = 9.0, 4.5, 2.0 Hz, 1H), 4.26 (dd, *J* = 13.0, 5.5 Hz, 1H), 4.17 – 4.16 (m, 1H), 4.15 – 4.14 (m, 1H), 4.06 (d, *J* = 16.5 Hz, 1H), 4.02 (d, *J* = 16.5 Hz, 1H), 2.11 (s, 3H), 2.05 (s, 3H), 1.98 (s, 3H).

¹³C NMR (125 MHz, CDCl₃) δ 170.60, 169.86, 169.78, 166.93, 88.36, 70.37, 68.41, 65.32, 61.96, 52.47, 50.10, 20.77, 20.73, 20.70.

HRMS (ESI) Calculated for C₁₄H₂₀N₇O₈ [M + H]⁺: 414.1373; Found: 414.1382.

(2R,3S,4R,5S)-2-(acetoxymethyl)-5-(2-azideacetamido)-6-(2-oxobutanamido) tetrahydro-2H-pyran-3,4-diyl diacetate (4):

To a solution of 2 (200.0 mg, 0.484 mmol) in THF (2 mL) was dropwise added PMe₃ (484 μ L, 1 M in THF, 0.484 mmol) at 0 °C. After being stirred for 30 mins, the ice bath was removed, and the reaction was stirred for an additional 5 minutes at room temperature. H₂O (87.0 mg, 87 μ L, 4.84 mmol) was added and the reaction was stirred overnight at room temperature. Then the mixture was concentrated to afford the crude glycosyl amine 3. In another flask, to a solution of 2-ketobutyric acid (247.1 mg, 2.420 mmol) in DCM (3 mL) were added triethylamine (244.4 mg, 336.6 μ L, 2.420 mmol) and oxalyl chloride (246.0 mg, 164 μ L, 1.936 mmol) at room temperature under N₂ atmosphere. After being stirred at for 30 mins, the reaction was concentrated to afford the crude acid chloride. The crude glycosyl amine 3 was dissolved in DCM (1 mL) and pyridine (153.5 mg, 157 μ L, 1.936 mmol) was added at room temperature under N₂ atmosphere. Then the acid chloride was dissolved in DCM (2 mL) and was dropwise added to the above mixture at 0 °C. After being stirred overnight at the same temperature, the reaction was monitored by TLC (hexane/ethyl acetate = 1/1). Then the reaction was quenched with sat. NaHCO₃(aq) and the mixture was separated. The aqueous phase was further extracted with DCM. The combined organic layer was washed with brine, dried over Na₂SO₄, filtered, and concentrated. The crude product was purified by column chromatography on silica gel (hexane/ethyl acetate = 1:1) to afford product (57.0 mg, 25% yield).

R_f = 0.2 (hexane/ethyl acetate = 1:1);

¹H NMR (500 MHz, CDCl₃): δ 7.61 (d, *J* = 9.2 Hz, 1H), 6.66 (d, *J* = 9.8 Hz, 1H), 5.42 (d, *J* = 9.2 Hz, 1H), 5.22 – 5.04 (m, 2H), 4.69 (dd, *J* = 9.7, 3.4 Hz, 1H), 4.24 (dd, *J* = 12.5, 4.3 Hz, 1H), 4.17 (s, 2H), 4.14 (s, 1H), 3.82 (ddd, *J* = 9.7, 4.4, 2.2 Hz, 1H), 2.93 (q, *J* = 7.2 Hz, 2H), 2.11 (s, 3H), 2.07 (s, 3H), 1.99 (s, 3H), 1.10 (t, *J* = 7.2 Hz, 3H).

¹³C NMR (125 MHz, CDCl₃): δ 197.51, 170.49, 169.98, 169.60, 167.88, 158.66, 74.53, 71.85, 64.79, 61.79, 52.50, 50.43, 31.21, 20.63, 7.53.

HRMS (ESI) *m/z*: calculated for C₁₈H₂₆N₅O₁₀ [M + H]⁺ 472.1680, found 472.1674.

(2R,3S,4R,5S)-2-(acetoxymethyl)-5-(2-azideacetamido)-6-((Z)-2-ethyl-4-oxobut-2-enamido)tetrahydro-2H-pyran-3,4-diyl diacetate (AAMCHO)

To a solution of 4 (57.0 mg, 0.121 mmol) in THF (200 μ L) was dropwise added a solution of (Triphenylphosphoranylidene)acetaldehyde (147.2mg, 0.484 mmol) in THF (500 μ L) at 0 °C. After being stirred 30 mins, the ice bath was removed, and the reaction was warmed up to 45 °C and stirred overnight. The reaction was monitored by TLC (hexane/ethyl acetate = 1/2). Then the solvent was removed under

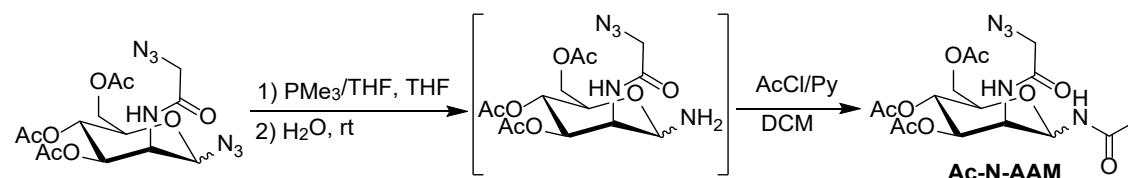
vacuum and resuspended in ethanol (2 mL) followed by addition of anhydrous ZnCl_2 (131.9 mg, 0.968 mmol), stirred overnight. The mixture was filtered, and the solution phase was washed with brine, dried over Na_2SO_4 , filtered, and concentrated. The crude product was purified by column chromatography on silica gel (hexane/ethyl acetate = 8:1) to afford product (24.0 mg, 39% yield).

R_f = 0.3 (hexane/ethyl acetate = 1:2);

^1H NMR (500 MHz, CDCl_3 , mixture of anomers 1:3 α : β) δ 9.70 (d, J = 7.7 Hz, 1H), 9.63 (d, J = 7.9 Hz, 1H), 7.03 (d, J = 8.9 Hz, 1H), 6.97 (d, J = 8.0 Hz, 2H), 6.94 – 6.80 (m, 2H), 6.58 (dt, J = 5.9, 1.9 Hz, 2H), 6.38 (dd, J = 15.2, 7.7 Hz, 1H), 6.27 (dd, J = 15.2, 7.8 Hz, 1H), 5.62 – 5.51 (m, 3H), 5.30 – 5.10 (m, 4H), 5.05 (q, J = 6.7 Hz, 8H), 4.93 – 4.77 (m, 1H), 4.67 (s, 1H), 4.38 – 4.20 (m, 3H), 4.22 – 4.12 (m, 1H), 4.06 (d, J = 7.8 Hz, 2H), 3.86 (d, J = 9.1 Hz, 3H), 2.58 (dd, J = 12.9, 7.4 Hz, 1H), 2.30 (t, J = 6.8 Hz, 2H), 2.14 (s, 3H), 2.13 (s, 3H), 2.09 (s, 3H), 2.04 (s, 3H), 2.02 (s, 3H), 2.00 (s, 2H), 1.16 – 1.13 (m, 4H).

^{13}C NMR (125 MHz, CDCl_3) δ 141.14, 139.24, 132.13, 132.05, 131.89, 128.61, 128.52, 80.49, 79.32, 74.30, 73.02, 65.38, 61.99, 60.42, 53.37, 49.91, 20.65, 18.04, 14.20, 11.42.

HRMS (ESI) m/z : calculated for $\text{C}_{20}\text{H}_{28}\text{N}_5\text{O}_{10}$ $[\text{M} + \text{H}]^+$ 498.1836, found 498.1832.



Scheme S2. Synthesis of **Ac-N-AAM**.

2-azideacetamido-1-N-acetyl-3,4,6-tri-O-acetyl-2-deoxy-D-glucopyranosylamine (Ac-N-AAM):

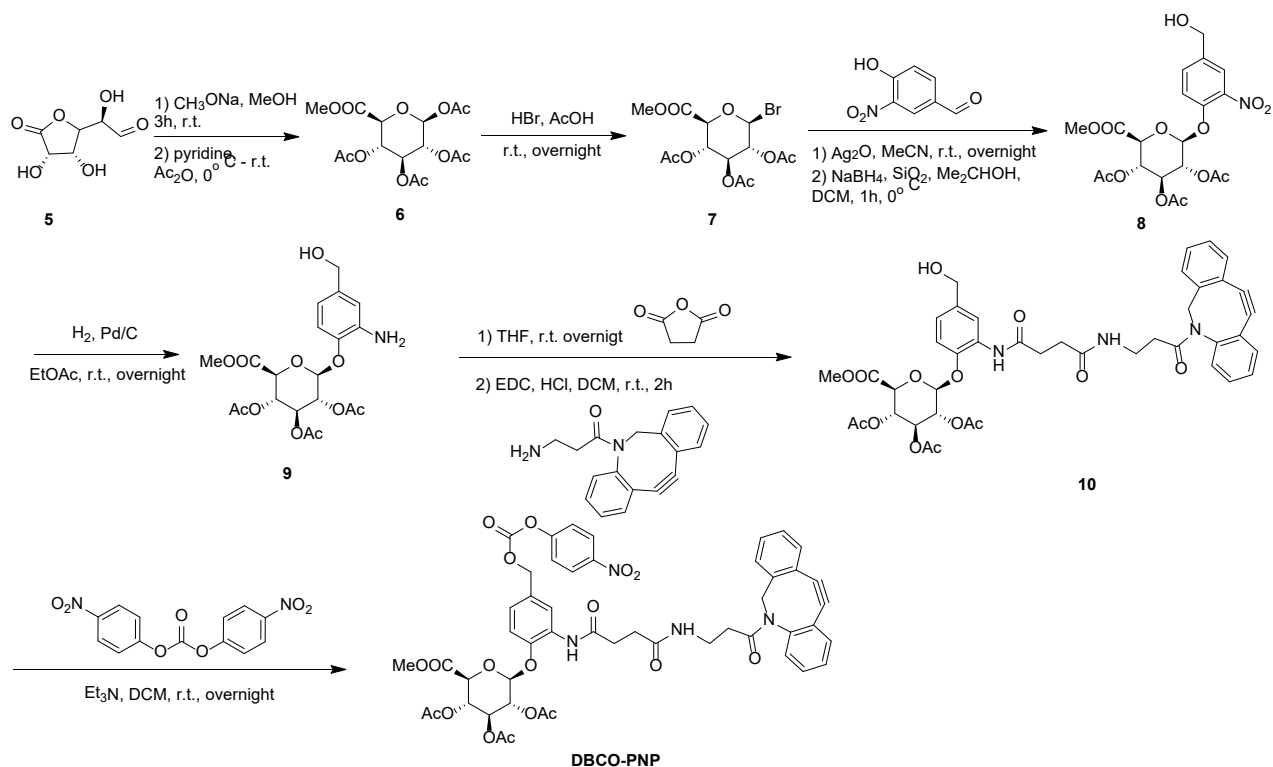
To a solution of **2** (30.0 mg, 0.073 mmol) in THF (1 mL) was dropwise added PMe_3 (73 μL , 1 M in THF, 0.073 mmol) at 0 °C. After being stirred for 30 mins, the ice bath was removed and the reaction was stirred for an additional 5 minutes at room temperature. H_2O (13.0 mg, 13 μL , 0.73 mmol) was added and the reaction was stirred overnight at room temperature. Then the mixture was concentrated, and the residue was dissolved in DCM (1 mL). Pyridine (46.0 mg, 47 μL , 0.584 mmol) and acetyl chloride (23.2 mg, 21 μL , 0.292 mmol) were added at 0 °C. After being stirred overnight at the same temperature, the reaction was monitored by TLC (hexane/acetone = 1/1). Then the mixture was concentrated, and the crude was purified by column chromatography on silica gel (hexane/acetone = 2:1) to afford **Ac-N-AAM** (12.1 mg, 39% yield).

R_f = 0.3 (hexane/acetone = 1:1);

^1H NMR (500 MHz, CDCl_3) δ 6.67 (d, J = 9.5 Hz, 1H), 6.40 (d, J = 9.0 Hz, 1H), 5.48 (d, J = 9.0 Hz, 1H), 5.15 – 5.07 (m, 2H), 4.64 (dd, J = 9.5, 3.5 Hz, 1H), 4.25 (dd, J = 12.5, 4.5 Hz, 1H), 4.15 (d, J = 16.5 Hz, 1H), 4.11 (d, J = 17.0 Hz, 1H), 4.10 (dd, J = 12.0, 2.0 Hz, 1H), 3.82 (ddd, J = 9.5, 4.5, 2.5 Hz, 1H), 2.10 (s, 3H), 2.06 (s, 3H), 2.01 (s, 3H), 1.98 (s, 3H).

^{13}C NMR (125 MHz, CDCl_3) δ 170.66, 170.04, 169.88, 169.83, 168.17, 76.78, 74.29, 72.18, 65.06, 62.04, 52.70, 50.84, 23.45, 20.80, 20.78, 20.78.

HRMS (ESI) Calculated for $\text{C}_{16}\text{H}_{24}\text{N}_5\text{O}_9$ $[\text{M} + \text{H}]^+$: 430.1574; Found: 430.1567.



Scheme S3 Synthesis of DBCO-DNP

Synthesis of 6 To a round bottom flask was added compound 5 (20 g, 0.114 mol) and methanol (150 mL). Then the stirring mixture was added 25% CH₃ONa in MeOH (0.8 mL, 0.0035 mol). The mixture was stirred at room temperature for 3 h. The mixture was concentrated using a rotary evaporator. The crude was dissolved in pyridine (100 mL). The flask was placed in an ice bath. To the stirring mixture was added acetic anhydride (70 mL, 0.74 mol). The mixture turned to a dark solution. The mixture was warmed up to room temperature slowly, stirred at room temperature overnight. The mixture was poured into water (400 mL), stirred at room temperature for 30 min. The mixture was extracted with EtOAc (250 mL) 5 times. The combined organic layer was divided into 2 equal portions. Each portion was extracted with saturated NH₄Cl (250 mL), and saturated NaHCO₃ (250 mL) 2 times. The organic layer was dried over MgSO₄, concentrated using a rotary evaporator. The crude product was suspended in a mixture of Hexane/EtOAc (150 mL, 2/1 v/v). The solid was collected by vacuum filtration, washed with EtOAc to give compound 6 as a white solid (12 g, 28%).

Synthesis of 7 To a 500 mL round bottom flask was added compound 6 (12 g, 0.032 mol), and AcOH (20 mL). To the stirring mixture was added HBr in AcOH (150 mL). The mixture was stirred at room temperature overnight. The mixture was poured into water (500 mL), extracted with DCM (400 mL) 3 times. The organic layer was dried over MgSO₄, concentrated using a rotary evaporator to dryness to give compound 7 as a yellow gel (11.6 g, 92%).

Synthesis of 8 To a 500 mL round bottom flask was added compound 7 (11.6 g, 0.029 mol), 3-nitro-4-hydroxyl phenylaldehyde (4.9 g, 0.029 mol), and CH₃CN (200 mL). To the stirring mixture was added Ag₂O (7.4 g, 0.032 mol). The mixture was stirred at room temperature overnight. The mixture was filtered through a pad of Celite. The filtrate was concentrated using a rotary evaporator to give a brown solid. The solid was transferred to a 1 L round bottom flask. The solid was dissolved in DCM (500 mL), and *i*-PrOH (100 mL). To the stirring mixture was added SiO₂ (40 g). The flask was placed in an ice bath. To the stirring mixture was added NaBH₄ (2.3 g, 0.061 mol). The mixture was stirred at 0 °C for 1 h. The reaction was monitored by TLC using 100% EtOAc. The mixture was poured into water (500 mL), filtered through a pad of Celite. The organic layer was isolated. The aqueous layer extracted with DCM (500 mL). The combined organic layer was dried over MgSO₄, concentrated to give compound 8 as a yellow solid (10.1 g, 72%).

LRMS (ESI) Calcd for $C_{20}H_{23}NO_{13}$ $[M + Na]^+$: 508.3 Found: 508.1.

Synthesis of 9 A 250 mL round bottom flask was flushed with N_2 . To the flask was added 10% Pd/C (0.1 g), and EtOAc (100 mL). To the mixture was added compound 8 (1.0 g, 2.1 mmol). The flask was closed with a septum, flushed with H_2 using balloon. The mixture was stirred under H_2 atmosphere overnight. The reaction was monitored by TLC using Hexane/EtOAc = 1/2. The mixture was filtered through a pad of Celite. The colorless filtrate was concentrated using a rotary evaporator, dried under high vacuum to give compound 14 as a white foamy solid (0.92 g, 98%).

1H NMR (500 MHz, DMSO- d_6) δ 6.81 (d, J = 8.2 Hz, 1H), 6.67 (d, J = 2.0 Hz, 1H), 6.46 (dd, J = 8.2, 2.0 Hz, 1H), 5.53 – 5.37 (m, 2H), 5.15 – 4.93 (m, 3H), 4.71 – 4.56 (m, 3H), 4.32 (d, J = 5.7 Hz, 2H), 3.65 (s, 3H), 2.03 (d, J = 11.7 Hz, 9H).

LRMS (ESI) Calcd for $C_{20}H_{26}NO_{11}$ $[M + H]^+$: 456.1 Found: 456.1.

Synthesis of 10 To a 500 mL 1 neck round bottom flask was added compound 9 (3.93 g, 8.6 mmol), and THF (100 mL). To the stirring mixture was added succinic anhydride (0.9 g, 8.96 mmol). The mixture was stirred at room temperature for overnight. The mixture was concentrated using a rotary evaporator to give a pale brown foamy solid. The solid was mixed with DBCO-amine (2.4 g, 8.7 mmol), and DCM (125 mL). To the stirring mixture was added EDC·HCl (1.7 g, 8.9 mmol). The mixture was stirred at room temperature for 2 h. The reaction was monitored by TLC using EtOAc/MeOH = 9/1. The mixture was concentrated using a rotary evaporator, purified by Silica column using EtOAc/MeOH from 99/1 to 19/1. Fractions containing product were combined, concentrated to give compound 10 as a foamy off white solid (3.91 g, 56%).

1H NMR (500 MHz, DMSO- d_6) δ 8.59 (s, 1H), 7.86 (s, 1H), 7.76 – 7.55 (m, 4H), 7.53 – 7.27 (m, 8H), 7.02 (s, 2H), 5.51 (dd, J = 20.1, 8.7 Hz, 2H), 5.22 – 5.01 (m, 4H), 4.71 (d, J = 9.9 Hz, 1H), 4.40 (d, J = 5.7 Hz, 2H), 3.64 (s, 3H), 3.13 (dd, J = 13.8, 7.5 Hz, 1H), 3.01 – 2.88 (m, 1H), 2.28 (t, J = 7.1 Hz, 3H), 2.01 (d, J = 1.7 Hz, 9H), 1.84 (dd, J = 16.5, 7.4 Hz, 1H).

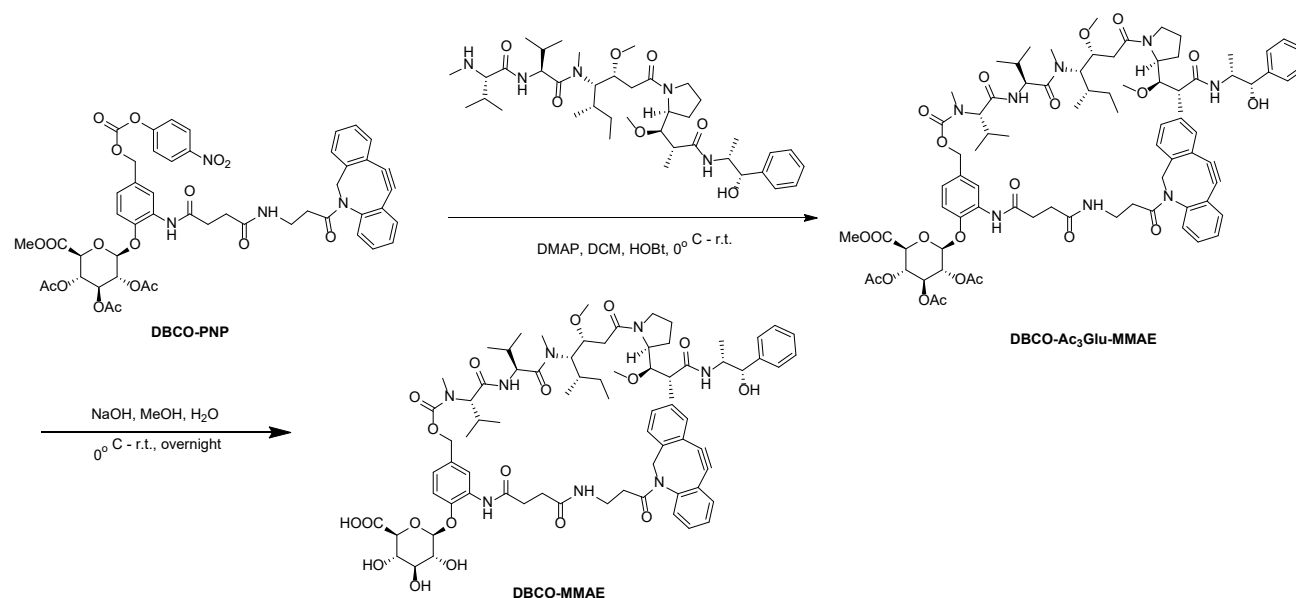
LRMS (ESI) Calcd for $C_{42}H_{44}N_3O_{14}$ $[M + H]^+$: 814.3 Found: 814.2.

Synthesis of DBCO-PNP To a 100 mL round bottom flask was added compound 10 (1.0 g, 1.23 mmol), DCM (30 mL), and Et_3N (0.4 mL, 2.86 mmol). To the stirring mixture was added bis (4-nitrophenyl) carbonate (0.8 g, 2.63 mmol). The mixture was stirred at room temperature overnight. The reaction was monitored by TLC using EtOAc/MeOH = 9/1. The mixture was concentrated using a rotary evaporator, purified by Silica gel column using EtOAc/MeOH from 100/0 to 19/1. Fractions containing product were combined, concentrated using a rotary evaporator to give compound DBCO-PNP as a white foamy solid (0.52 g, 43%)

1H NMR (500 MHz, $CDCl_3$) δ 8.40 (s, 1H), 8.10 (d, J = 16.0 Hz, 1H), 7.68 (d, J = 7.4 Hz, 1H), 7.45 – 7.27 (m, 4H), 7.06 (d, J = 7.5 Hz, 1H), 6.95 (d, J = 11.3 Hz, 1H), 6.28 (t, J = 6.1 Hz, 1H), 5.48 – 5.27 (m, 4H), 5.20 – 5.04 (m, 1H), 4.62 (s, 1H), 4.21 (t, J = 9.1 Hz, 1H), 4.14 (q, J = 7.1 Hz, 2H), 3.76 (s, 3H), 3.45 – 3.36 (m, 1H), 3.24 (t, J = 8.9 Hz, 1H), 2.84 – 2.60 (m, 3H), 2.59 – 2.35 (m, 4H), 2.19 – 2.02 (m, 9H), 2.01 – 1.89 (m, 2H), 1.27 (t, J = 7.1 Hz, 2H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 172.97, 171.64, 171.12, 170.43, 170.28, 169.86, 169.39, 166.55, 163.41, 155.55, 153.03, 151.06, 148.36, 145.38, 132.10, 130.04, 129.57, 129.16, 128.62, 128.37, 127.92, 127.24, 126.09, 125.66, 125.30, 124.55, 122.94, 122.54, 121.82, 120.99, 115.63, 114.86, 101.69, 72.61, 71.15, 71.02, 70.66, 69.13, 60.43, 55.54, 53.21, 35.89, 34.59, 31.99, 30.49, 20.50, 12.42.

LRMS (ESI) Calcd for $C_{49}H_{47}N_4O_{18}$ $[M + H]^+$: 979.3 Found: 979.3.



Scheme S4 Synthesis of DBCO-MMAE

Synthesis of DBCO-Ac₃Glu-MMAE To a 100 mL round bottom flask was added compound DBCO-PNP (190 mg, 0.19 mmol), and DCM (20 mL). The flask was placed in an iced water bath. To the stirring mixture was added DMAP (30 mg, 0.25 mmol), then MMAE (140 mg, 0.19 mmol). The reaction was warmed up slowly to room temperature and stirred overnight. The reaction was monitored by TLC using EtOAc/MeOH = 19/1. The mixture was purified by Si column using EtOAc/MeOH from 100/0 to 19/1. Fractions containing product were combined, concentrated using a rotary evaporator to give product as a white solid (165 mg, 55%).

LRMS (ESI) Calcd for C₈₂H₁₀₉N₈O₂₂ [M + H]⁺: 1557.8; Found: 1557.7.

Synthesis of DBCO-MMAE To a 25 mL 1 neck round bottom flask was added DBCO-Ac₃Glu-MMAE (165 mg, 0.11 mmol), MeOH (9 mL), and aqueous NaOH 0.1 M (1 mL, 0.1 mmol). The mixture was stirred at room temperature overnight. The mixture was concentrated using a rotary evaporator. The crude was purified by C18 column chromatography using H₂O/CH₃CN from 100/0 to 60/40. Fractions containing product were combined, lyophilized to get product.

LRMS (ESI) Calcd for C₇₅H₁₀₁N₈O₁₉ [M + H]⁺: 1417.7; Found: 1417.7.

NMR and Mass Spectra

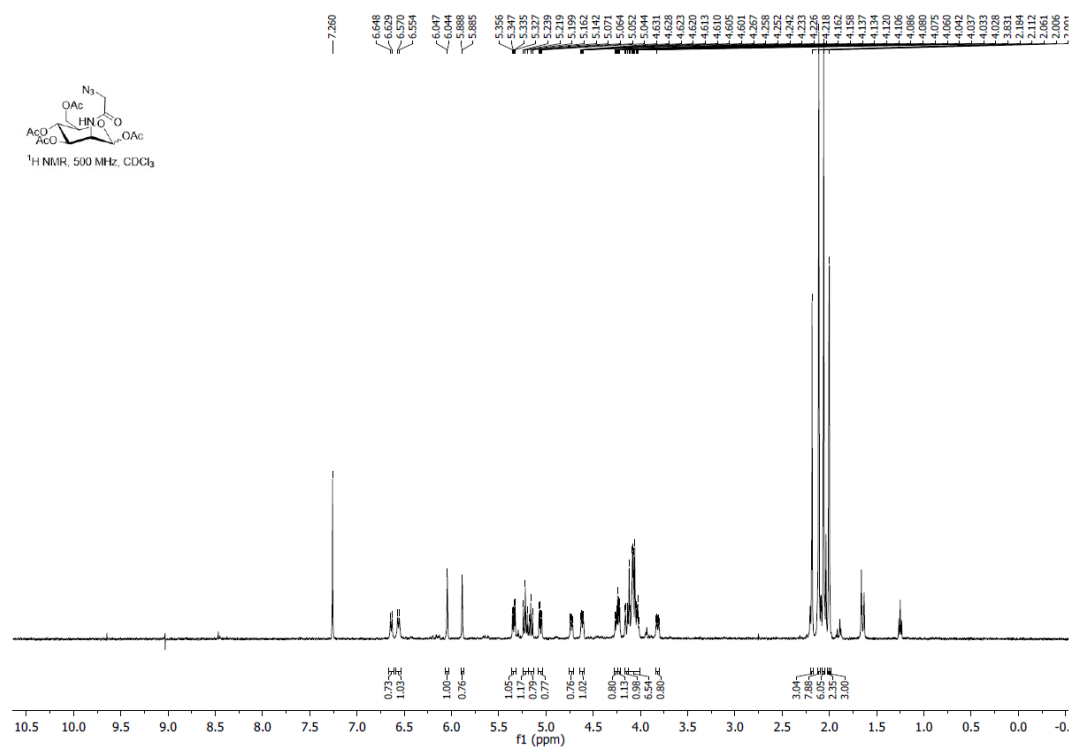


Figure S18 ¹H NMR of **1 Ac₄ManAz**

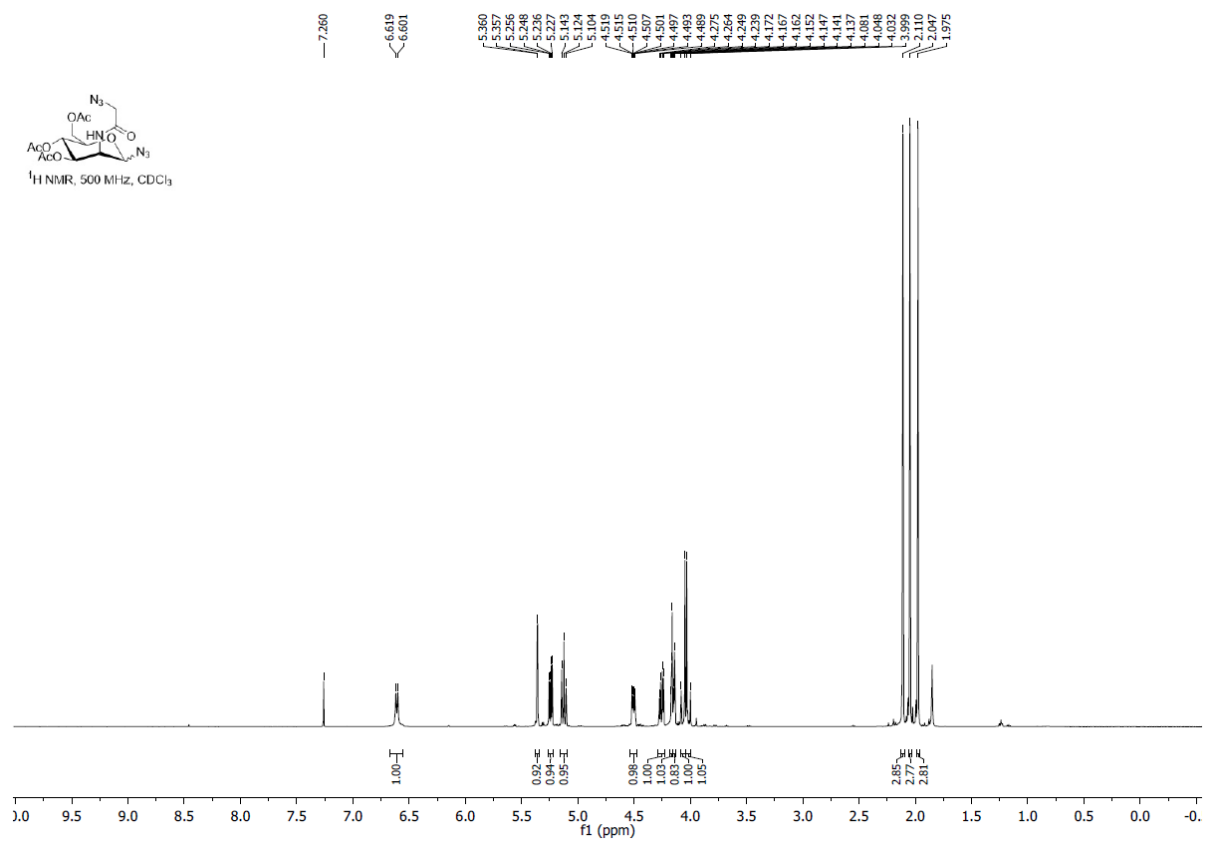


Figure S19 ^1H NMR of **2**

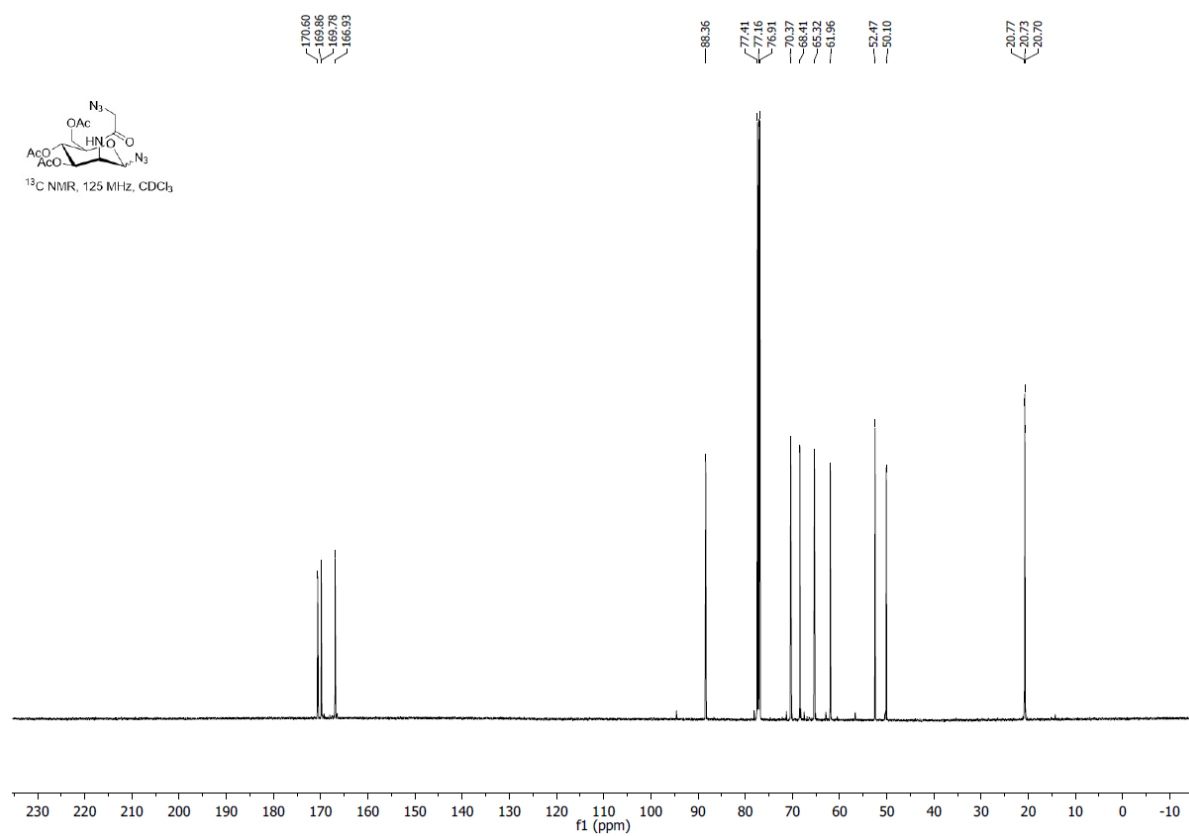


Figure S20 ^{13}C NMR of **2**

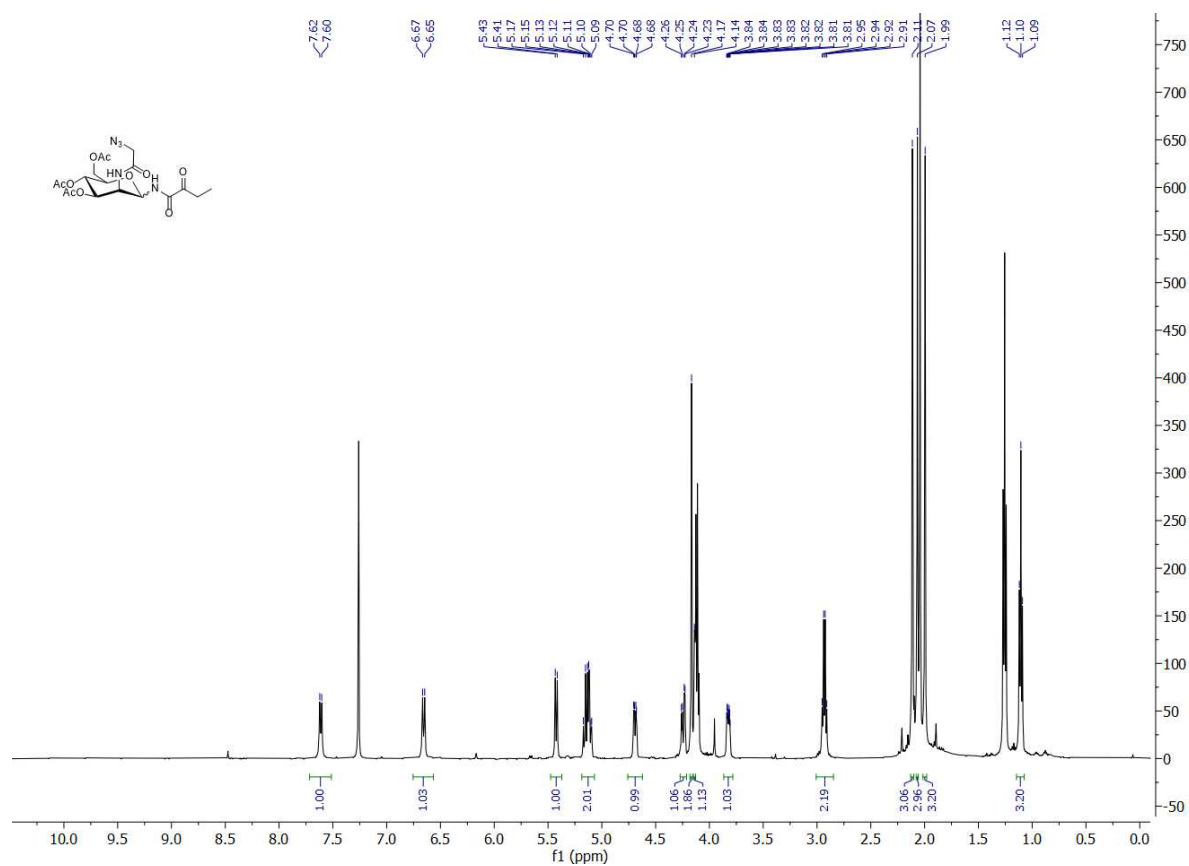


Figure S21 ¹H NMR of **4**

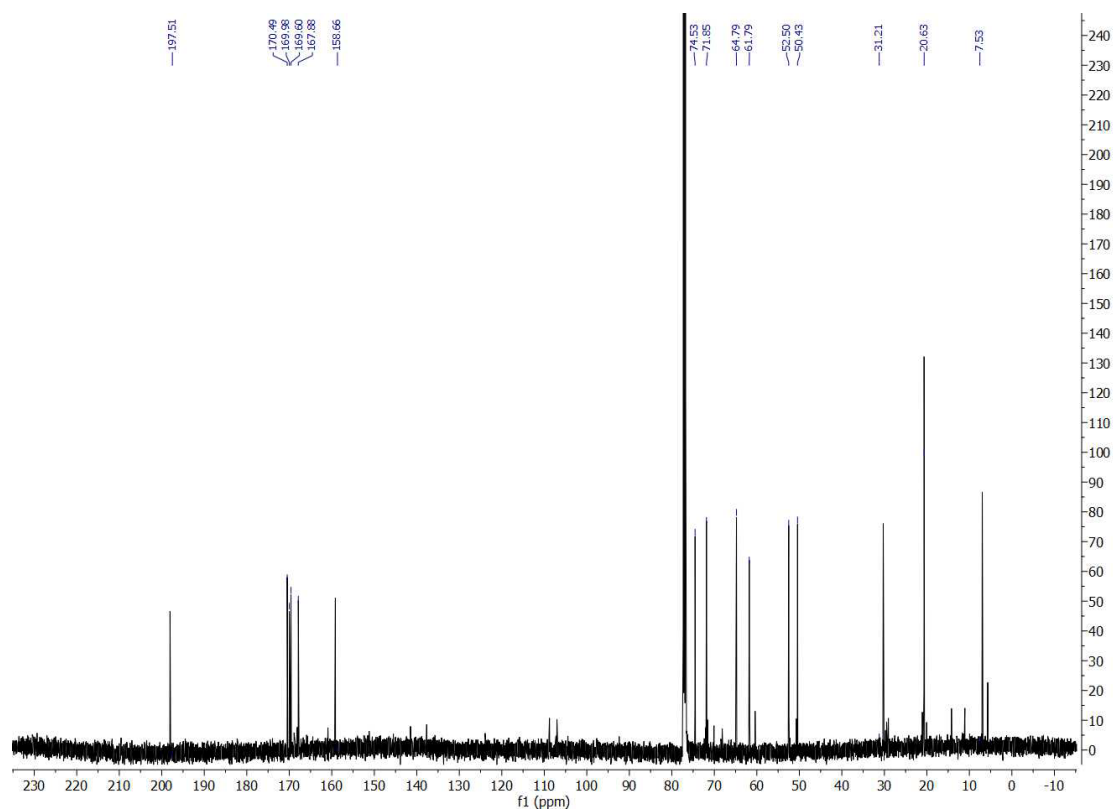


Figure S22 ¹³C NMR of 4

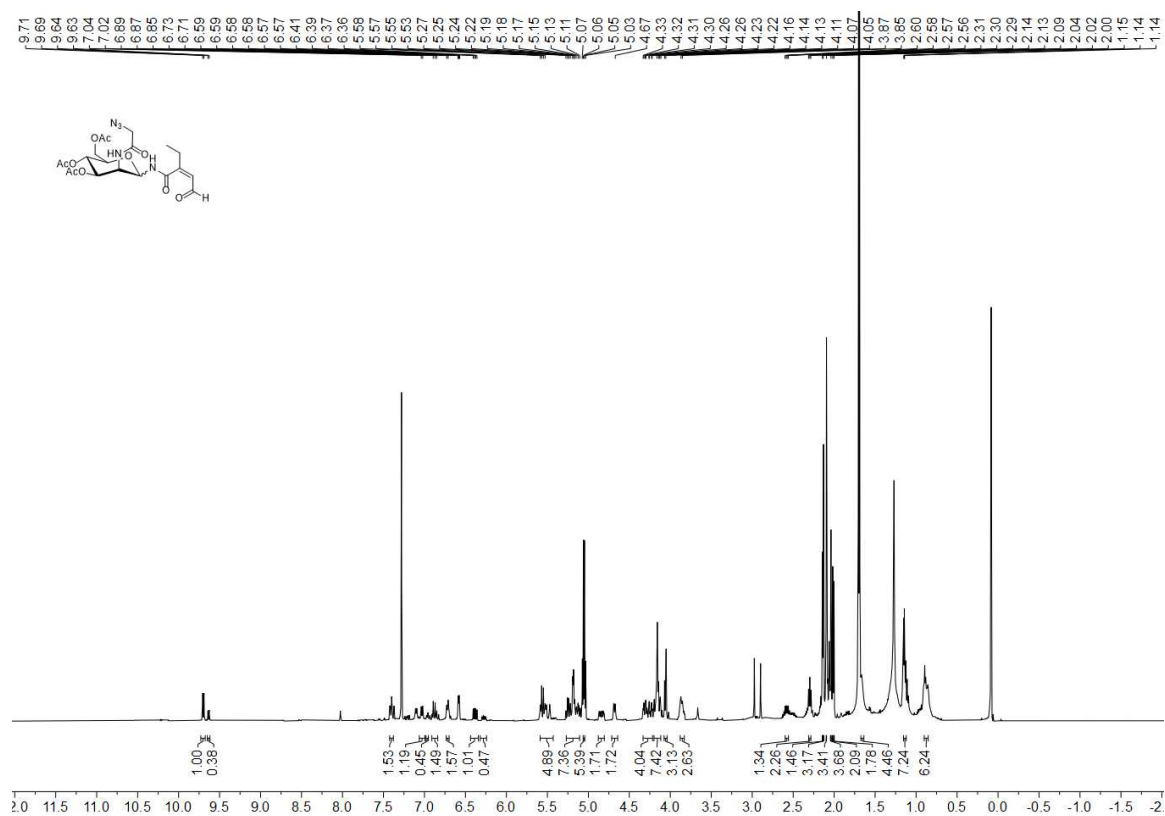


Figure S23 ¹H NMR of AAMCHO

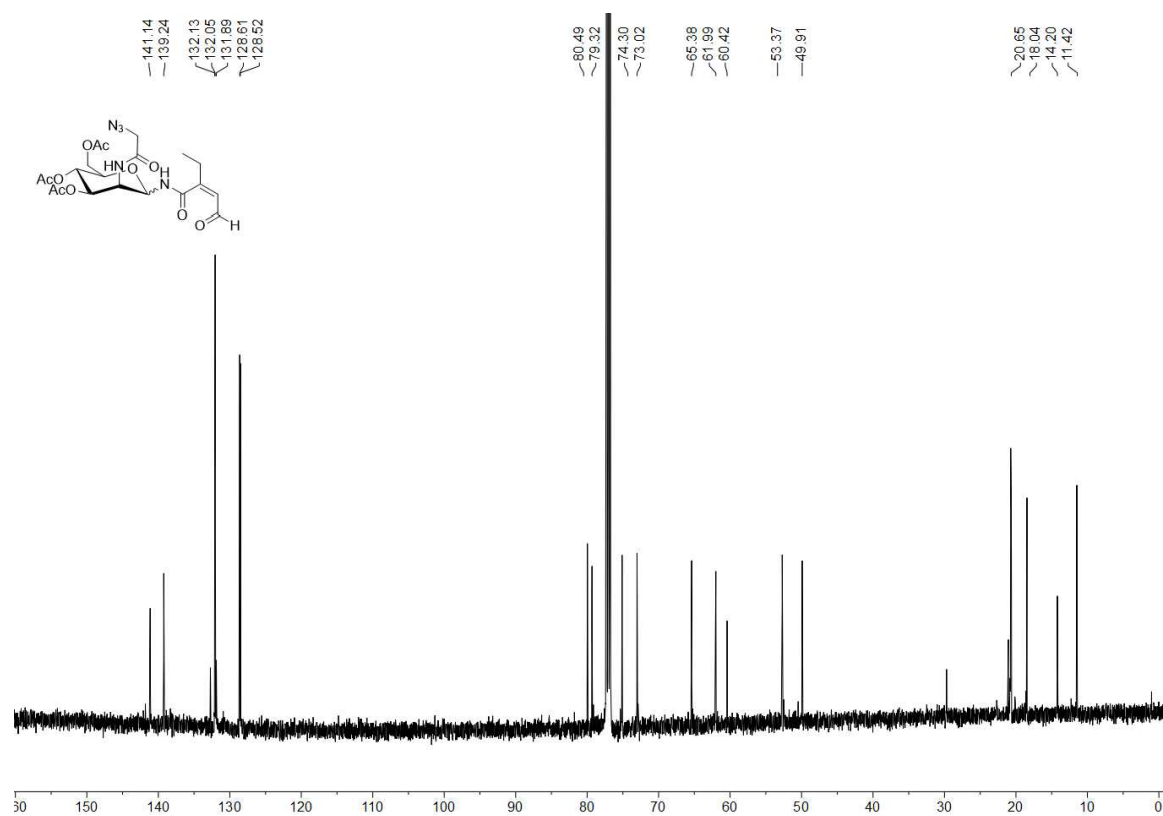


Figure S24 ^{13}C NMR of AAMCHO

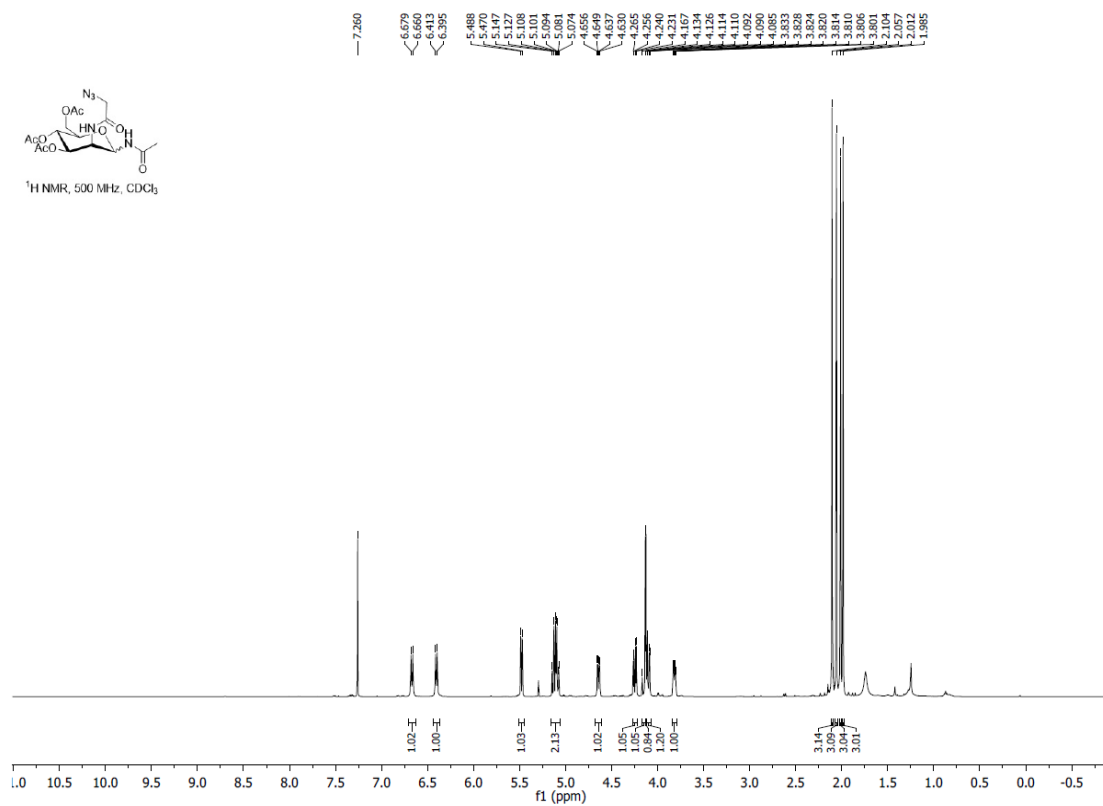


Figure S25 ¹H NMR of Ac-N-AAM

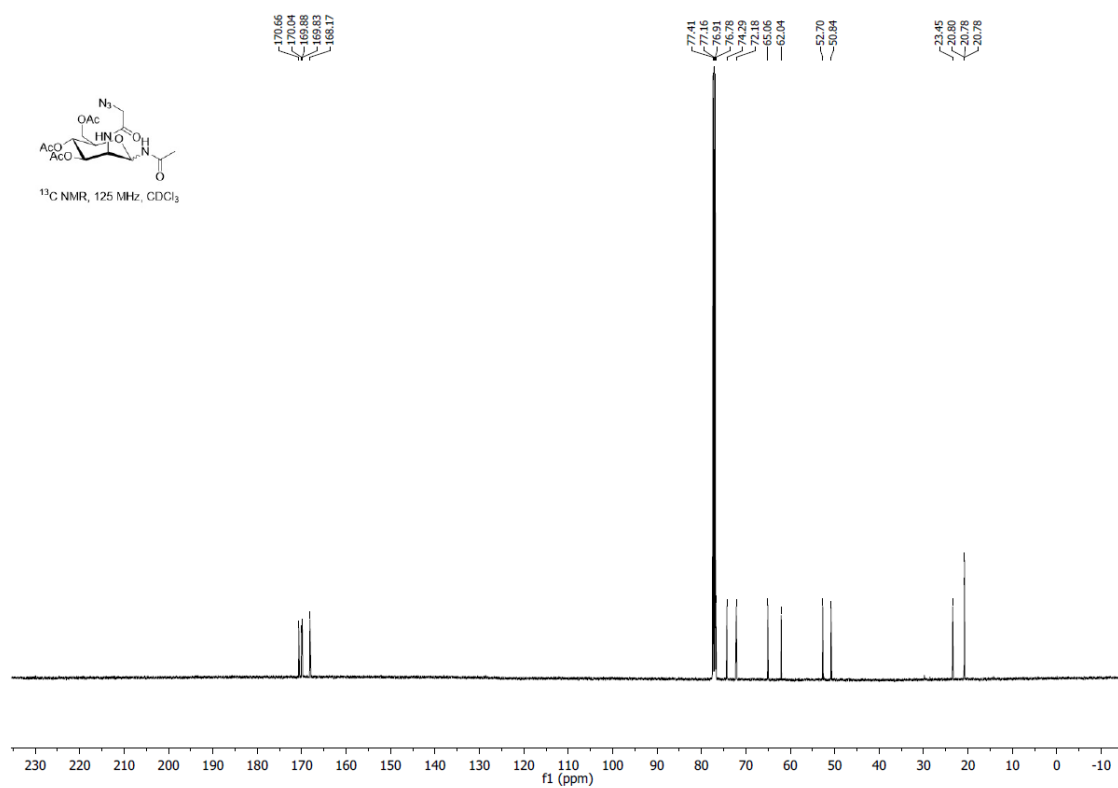


Figure S26 ¹³C NMR of Ac-N-AAM

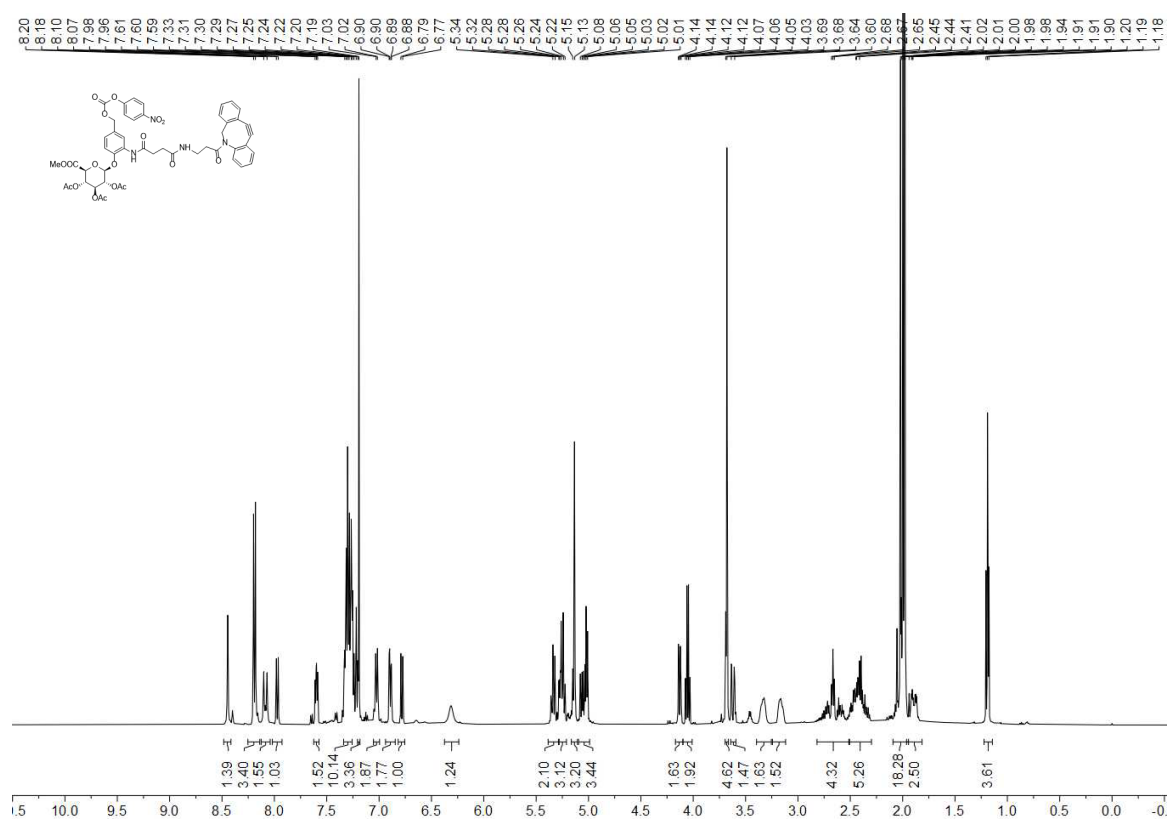


Figure S27 ¹H NMR of DBCO-PNP

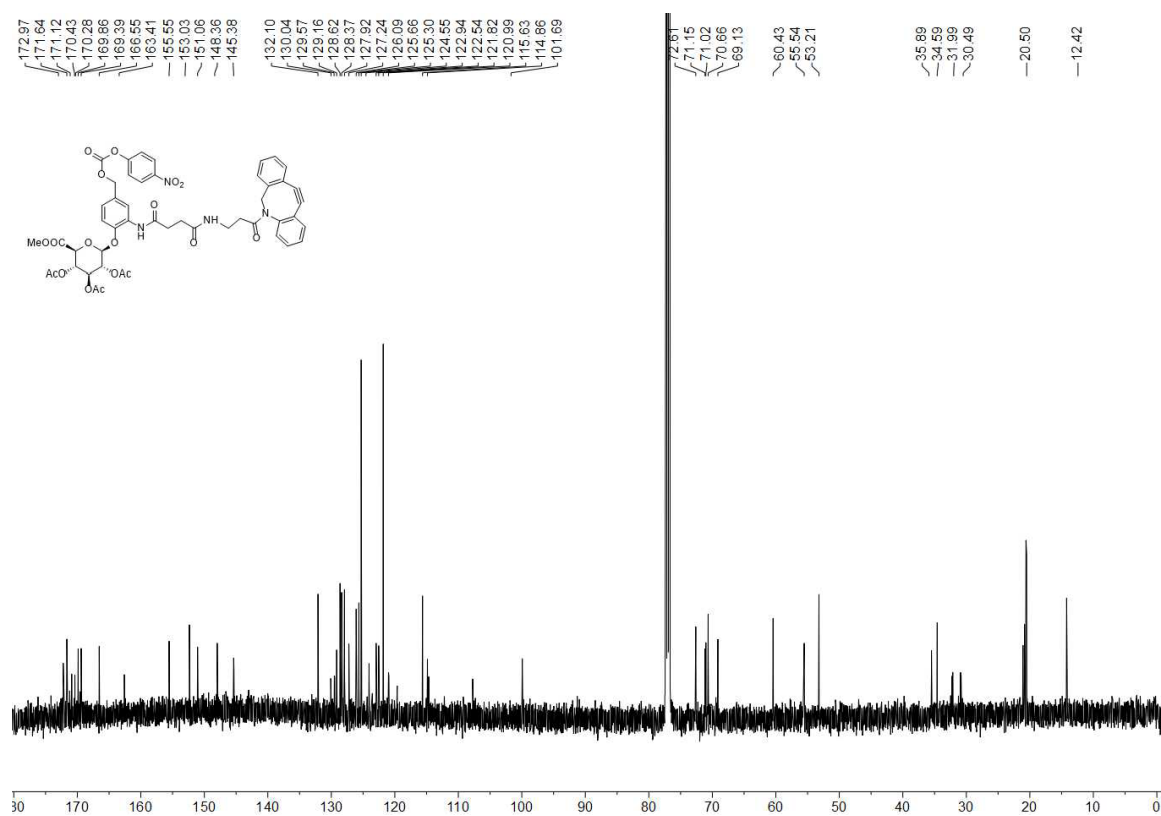


Figure S28 ^{13}C NMR of DBCO-PNP

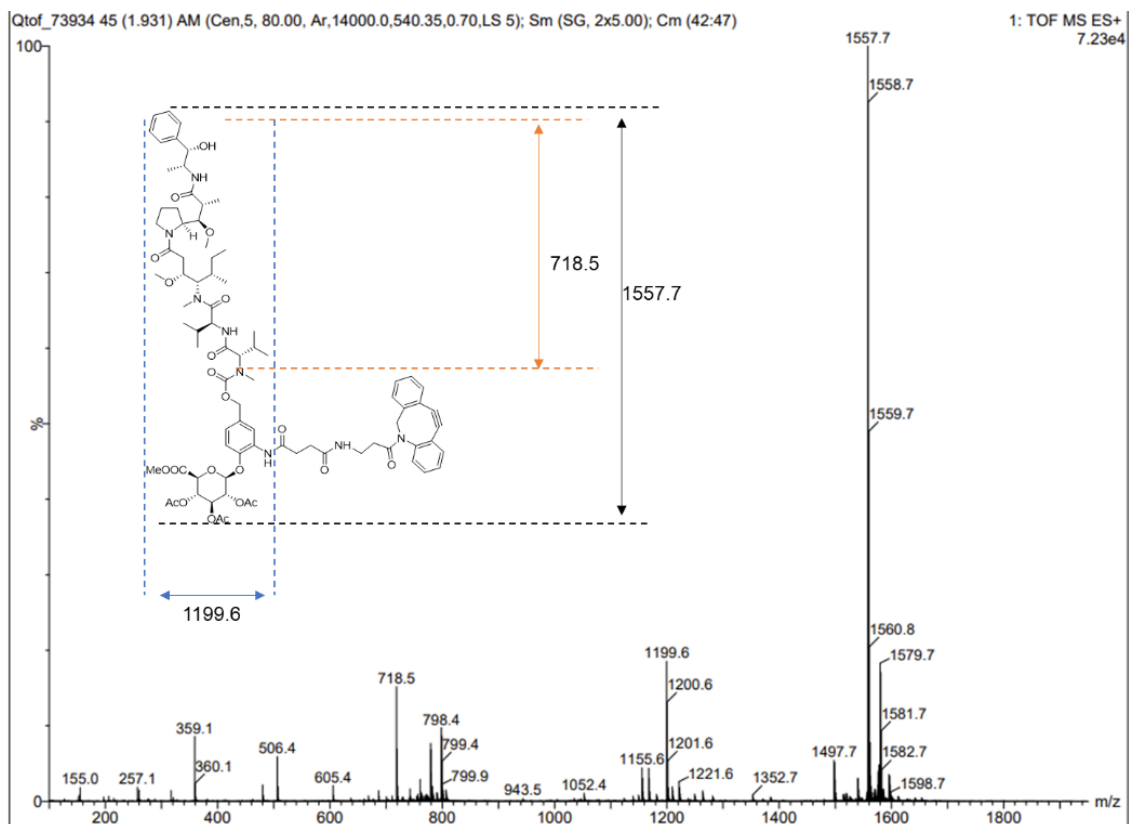


Figure S29 ESI-MS of DBCO-Ac₃Glu-MMAE

References

1. C. Anorma *et al.*, Surveillance of Cancer Stem Cell Plasticity Using an Isoform-Selective Fluorescent Probe for Aldehyde Dehydrogenase 1A1. *ACS central science* **4**, 1045-1055 (2018).
2. T. E. Bearrood, C. D. Anorma, J. Chan (2022) Selective fluorescent probe for aldehyde dehydrogenase. (US Patent 11,407,725).
3. A. S. Kooner, H. Yu, X. Chen, Synthesis of N-glycolylneuraminic acid (Neu5Gc) and its glycosides. *Frontiers in immunology* **10**, 2004 (2019).