Supporting Information

Chemical Modification-Assisted Bisulfite Sequencing (CAB-Seq) for 5-Carboxylcytosine Detection in DNA

Xingyu Lu[1], Chun-Xiao Song[1], Keith Szulwach[2], Zhipeng Wang[1], Payton Weidenbacher[1], Peng Jin[2],Chuan He[1]*

[1]Department of Chemistry and Institute for Biophysical Dynamics, The University of Chicago, 929 East 57th Street, Chicago, Illinois, 60637, USA.

[2]Department of Human Genetics, Emory University School of Medicine, Atlanta, Georgia, USA

Table of Contents	Page

A. General Materials

All oligonucleotides containing modified cytosines were synthesized using Applied Biosystems 392 DNA synthesizer with modified dC-CE phosphoramidite (Glen Research). All synthetic oligonucleotides were then purified by denaturing PAGE, the molecular weight of short oligo is confirmed by MALDI-TOF (Figiure S2-S4). The oligonucleotides without modified bases were purchased from Operon.

Synthetic modified oligonucleotides sequence:
9 mer DNA: 5'-GAC XGG AGT-3'; X = 5hmC, 5fC, 5caC.
11 mer DNA: 5'-CGA GTX AAG GC-3'; X = 5caC
12 mer DNA: 5'-GCTGXGTGCGC-3'; X = 5caC
76 mer DNA:
5'-CCTCACCATCTCAACCAATATTATATTATGTGTATATXGXGTATTTTGTGTT
ATAATATTGAGGGAGAAGTGGTGA-3'; X = caC

B. Methods

EDC-catalyzed coupling reaction

a. 1-step coupling: Double stranded 5caC DNA (10 μmol), amine (10 mM), EDC (2 mM), Mes (pH = 6.0, 75 mM) were incubated in aqueous solution at 37 °C for 1.0 h.

b. 2-step coupling: Double stranded 5caC DNA (10 μmol), NHS (20 mM), EDC (2 mM), Mes (pH = 5.0, 75 mM) were incubated in aqueous solution at 37 °C for 0.5 h. Buffer exchange to sodium phosphate (pH = 7.5, 100 mM), NaCl (150 mM), amine (10 mM), incubated at 37 °C for 1.0 h

Synthesis of (4-aminomethyl)benzylazide

0.325 g NaN_3 (Sigma-Aldrich) was dissolved in DMSO for 2h. Then 4-(bromomethyl)benzylamine HBr (Astatech. Inc.) in DMSO was added into the mixture. The reaction was stirred at room temperature for 2h. Then 20 mL H2O was slowly added. The mixture was extracted with CH_2Cl_2 and washed by 1 M NaOH and H_2O. The pure product was gained by flash column on silica as a pale yellow oil as previous reported.[1] ^{1}H NMR (500 MHz) (CD3Cl) δ: 7.34 (d, *J* = 8.0 Hz, 2H; Ar*H*), 7.29 (d, *J* = 8.0 Hz, 2H; Ar*H*), 4.32 (s, 2H; CH_2N_3), 3.88 (s, 2H; CH_2NH_2), 1.49 (s, 2H, CH_2NH_2) (See Figure S9). HRMS Calcd. for C8H11N4: 163.0984 ($[M+H]^+$), Found: 163.0909 ($[M+H]^+$).

Copper free Click reaction on N_3-5caC and DTT cleavage

The copper free click chemistry was performed with addition of 150 μM dibenzocyclooctyne modified biotin into the 10 μM N_3-5caC DNA solution, and the reaction mixture was incubated for 2 h at 37 °C. The reaction was purified with QIAquick Nucleotide Removal Kit (QIAGEN). The disulfide bond in purified product can be cleaved by 100 mM DTT for 2 h at 37 °C.

HPLC analysis of EDC-mediated coupling reaction

0.1 nmol 12/16 mer double stranded 5caC-containing DNA was labeled under standard condition. The reaction yield was calculated by HPLC with ion-exchange column equilibrated with buffer A (10 mM sodium phosphate, pH = 12.0) and buffer B (10 mM sodium phosphate, pH = 12.0, 1 M NaCl).

Cu(I)-catalyzed click reaction on N_3-5caC DNA

The Cu(I)-catalyzed click reaction was performed with 10 nmol 11/15 mer double stranded 5caC-containing DNA, 0.5 mM ascorbic acid, 0.5 mM Cu(II)-TBTA (Lumiprobe), 100 mM TEAA buffer, 10 mM biotin-PEG4- biotin (Click Chemistry Tool) in 50% DMSO aqueous solution at room temperature overnight. The product was purified by acetone precipitation. The purified DNA was digested by nuclease P1 and alkaline phosphatase (Sigma-Aldrich). The digested single nucleosides were separated by HPLC with C18 reverse-phase column equilibrated with buffer A (5 mM ammonium acetate, pH 7.5) and buffer B (5mM ammonium, 0.01% TFA, 60% CH_3CN). The purified peaks were collected and analyzed by HRMS.

Bisulfite treatment of the 76 mer double stranded 5caC-containing DNA and colony picking

100 ng ds 76 mer double stranded 5caC-containing DNA was EDC labeled under standard condition. After EDC treatment, the DNA was purified with QIAquick Nucleotide Removal Kit (QIAGEN) and then applied to EpiTect Bisulfite Kit (QIAGEN) following the supplier's instruction. After PCR amplification with Hotstar Taq polymerase (QIAGEN) (forward primer: 5′-CCCTTT TATTATTTTAATTAATATTATATT-3′; reverse primer: 5′-CTCCGACATTATCACTACCATCAACCACCCATCCTACCTGGACTACA TTCTTATTCAGTATTCACCACTTCTCCCTCAAT-3′), the PCR product was purified using PCR purification kits (QIAGEN) and sent for Sanger sequencing.

For the colony picking, the PCR product was inserted into the vector following the instruction of the TOPO TA Cloning kit (Invitrogen) and sent for Sanger sequencing (Figure S10).

Bisulfite reaction on single nucleoside

50 μM single nucleosides were treated with saturated sodium bisulfite (Sigma-Aldrich) solution at 60 °C. The reaction was monitored by HPLC at each time point with C18 reverse-phase column equilibrated with buffer A (5 mM ammonium acetate, pH 7.5) and buffer B (5 mM ammonium, 0.01% TFA, 60% CH_3CN).

qPCR of 76 mer DNA

qPCR is performed with 10 pg of 76 mer dC, 5caC, N3-5caC containing DNA using SYBR® Green Real-Time PCR Master Mixes (Invitrogen). The relative amplification efficiency was calculated by $Rx=2^{(Ct_x/Ct_{dC})}$. (forward primer: 5'-CCTCACCATCTCAACCAATA-3'; reverse primer 5'-TCACCACTTCTCCCTCAAT-3'.)

C. Supplemental Tables and Figures

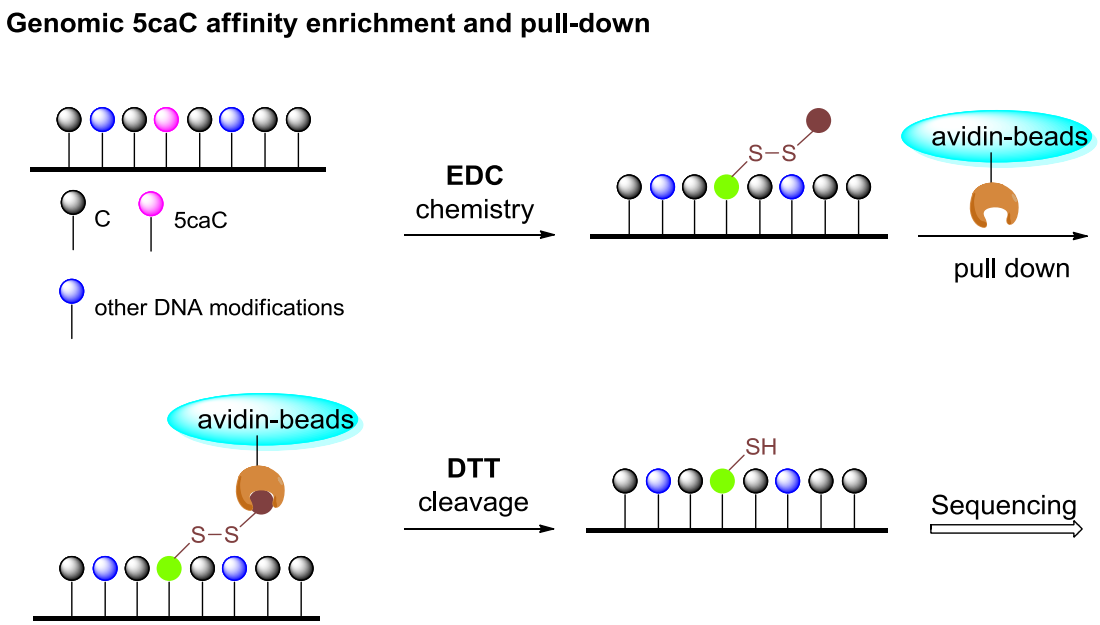

Figure S1. 5caC pull-down strategy using the optimized EDC chemistry. The 5caC-containing DNA fragments may be selectively labeled with a biotin tag. The avidin-beads can capture the biotin group and enrich the 5caC-containing fragments. After DTT cleavage of the disulfide bond, the 5caC-containing fragments can be eluted and used for sequencing.

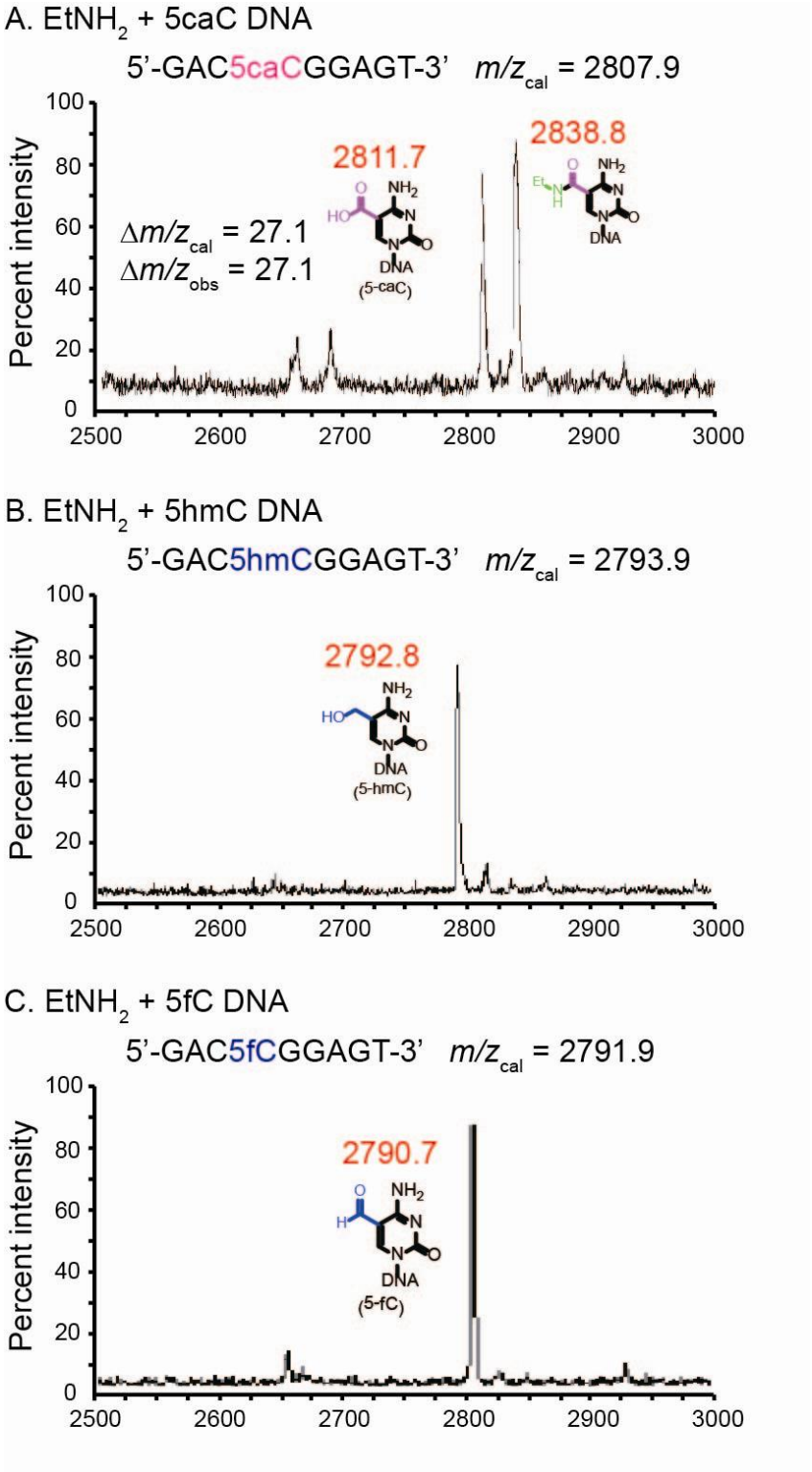

Figure S2. EDC-mediated amide bond formation on different DNA modifications. MALDI-TOF of the EDC-mediated coupling between ethylamine and 5caC, 5hmC, and 5fC, respectively, with the calculated molecular weight and observed molecular weight indicated. Reactions were performed in duplex DNA with the complementary strand; however, MS monitored the single-stranded DNA containing the modification.

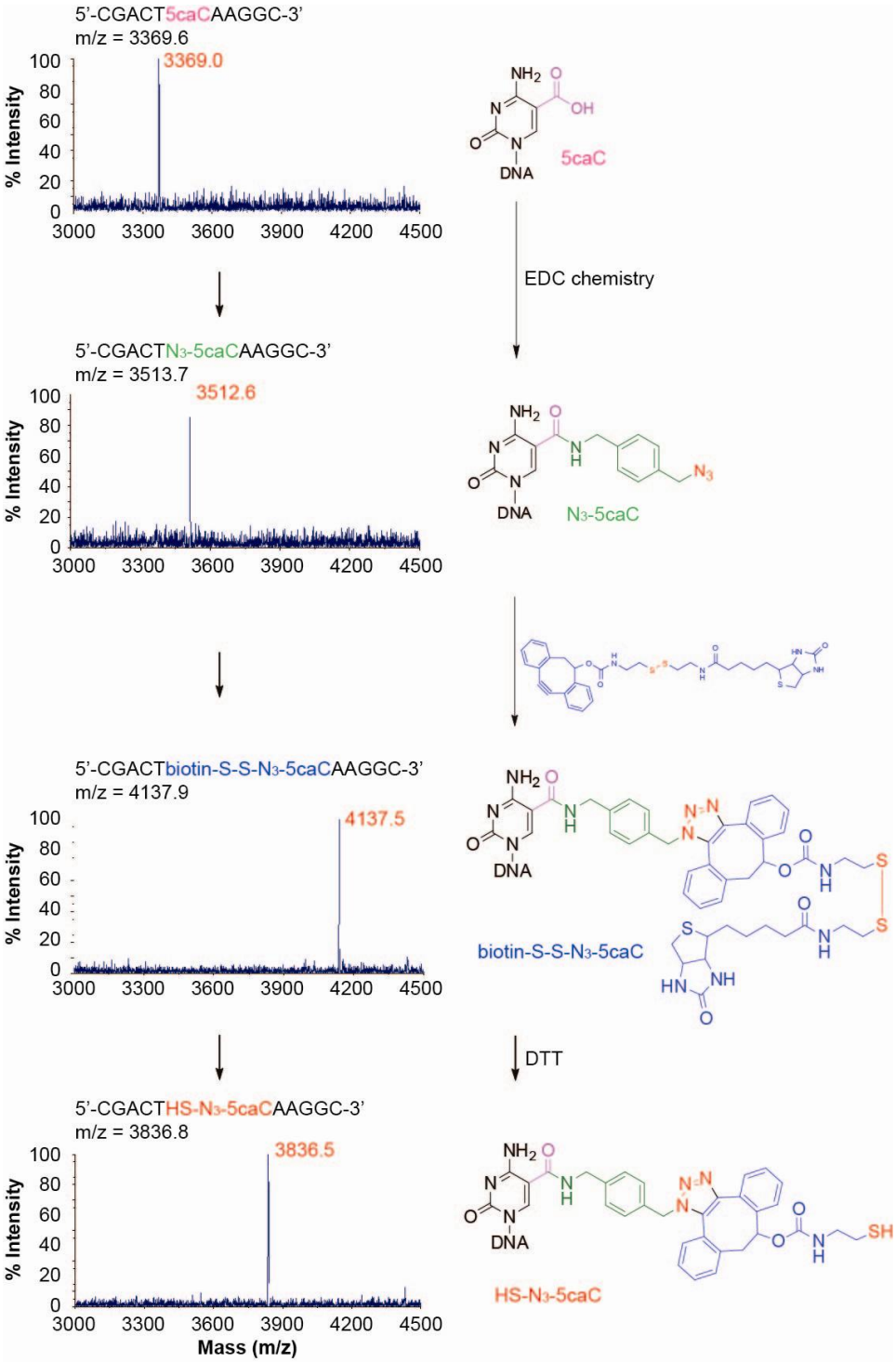

Figure S3. Mass spectrum of EDC-based 5caC labeling and biotinylation. MALDI-TOF of 5caC-, N_3-5caC-, and biotin-S-S-N_3-5caC-containing 11 mer DNA, respectively, with the calculated molecular weight and observed molecular weight indicated. Corresponding reactions of the EDC-based coupling and the subsequent copper-free click chemistry to yield biotin-S-S-N_3-5caC in duplex DNA are shown. Reactions were performed in duplex DNA with the complementary strand; however, MS monitored the single-stranded DNA containing the modification.

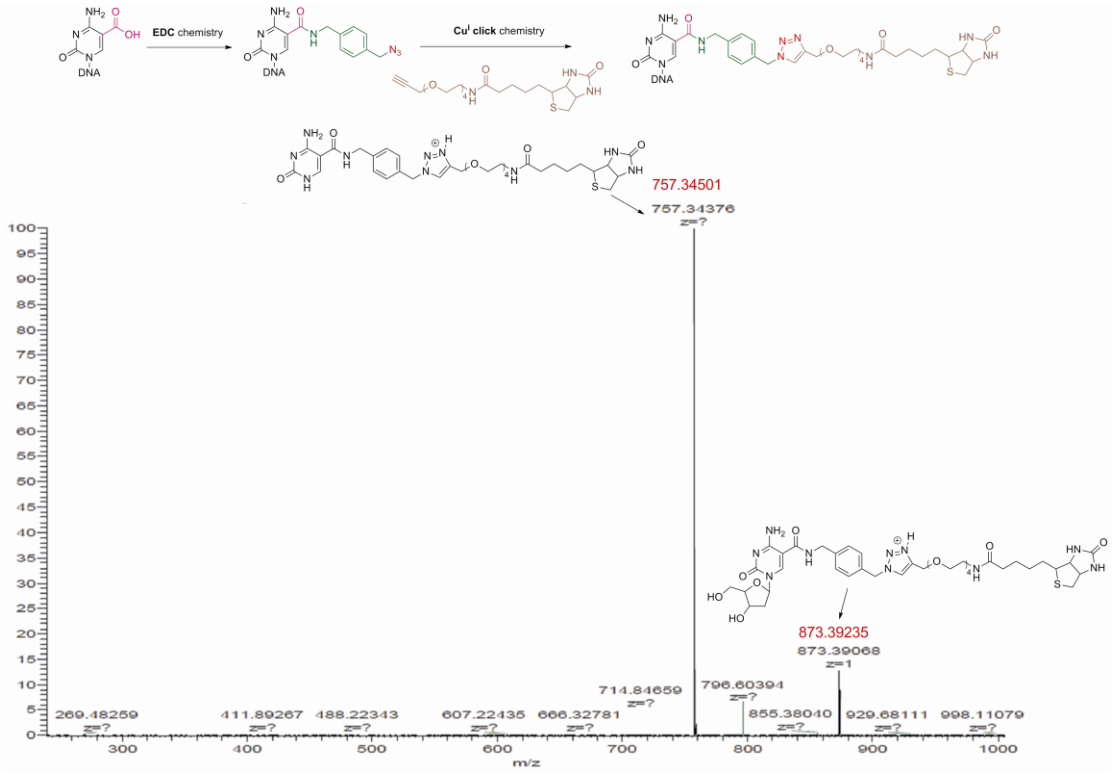

Figure S4. HRMS of labeled 5caC-containing nucleoside. 11/15 mer dsDNA with 5caC was labeled by the EDC-mediated reaction and subsequent Cu(I)-catalyzed click chemistry. After digestion and HPLC separation, the labeled 5caC nucleoside was subject to high resolution mass spectra.

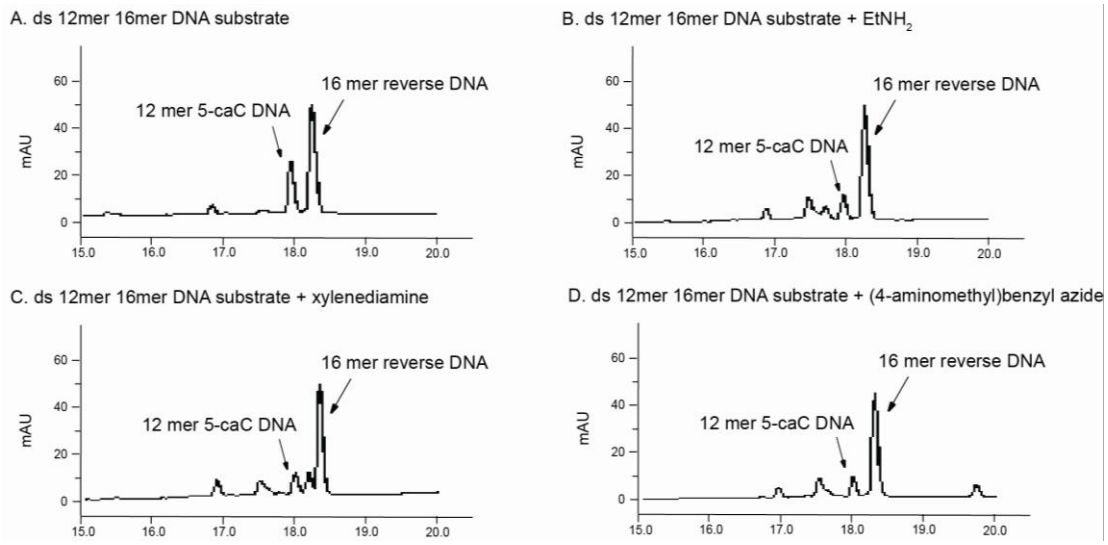

Figure S5. HPLC of EDC-mediated reactions between active primary amines and 5caC-containing DNA. The 12 mer 5caC-containing DNA and 16 mer complementary DNA were indicated. The reaction was monitored by HPLC with ion-exchange column equilibrated with buffer A (10 mM sodium phosphate, pH=12.0) and buffer B (10 mM sodium phosphate, pH = 12.0, 1 M NaCl).

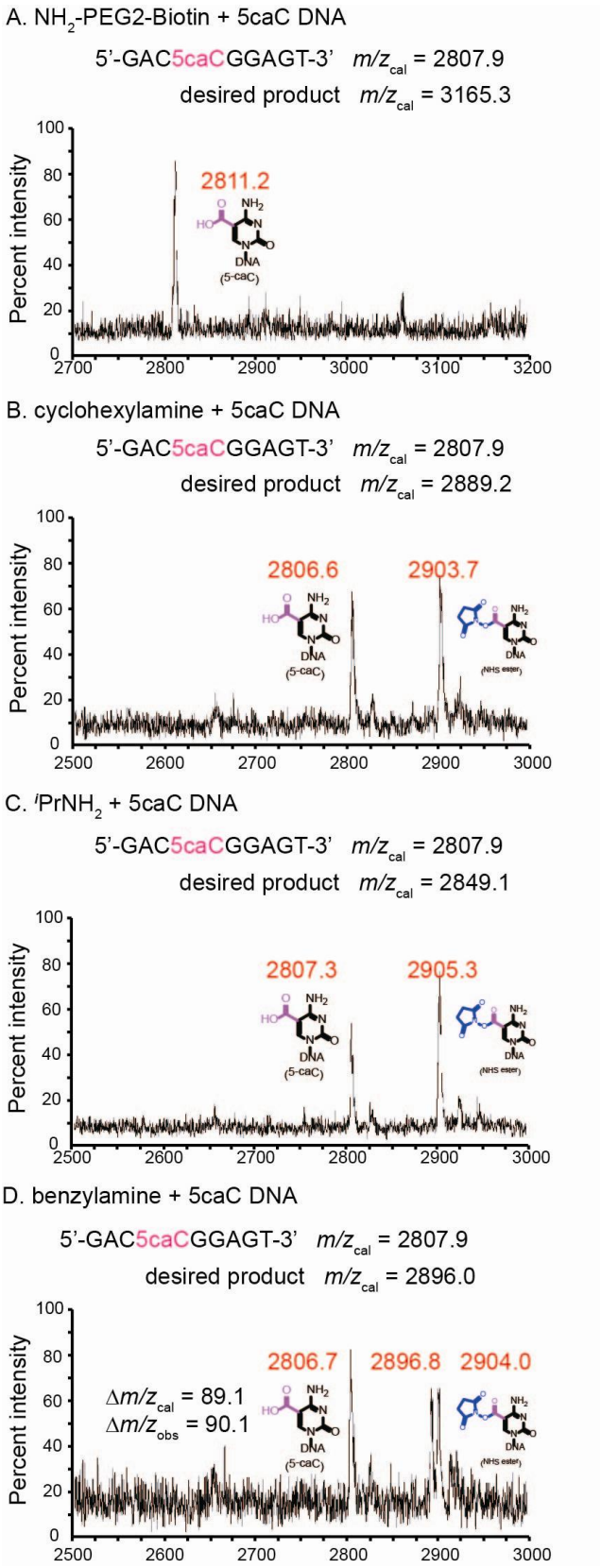

Figure S6. Mass spectrum of bulky primary amines in the EDC-mediated reaction. MALDI-TOF of EDC-mediated coupling between bulky primary amines and 5caC-containing DNA with the calculated molecular weight and observed molecular weight indicated.

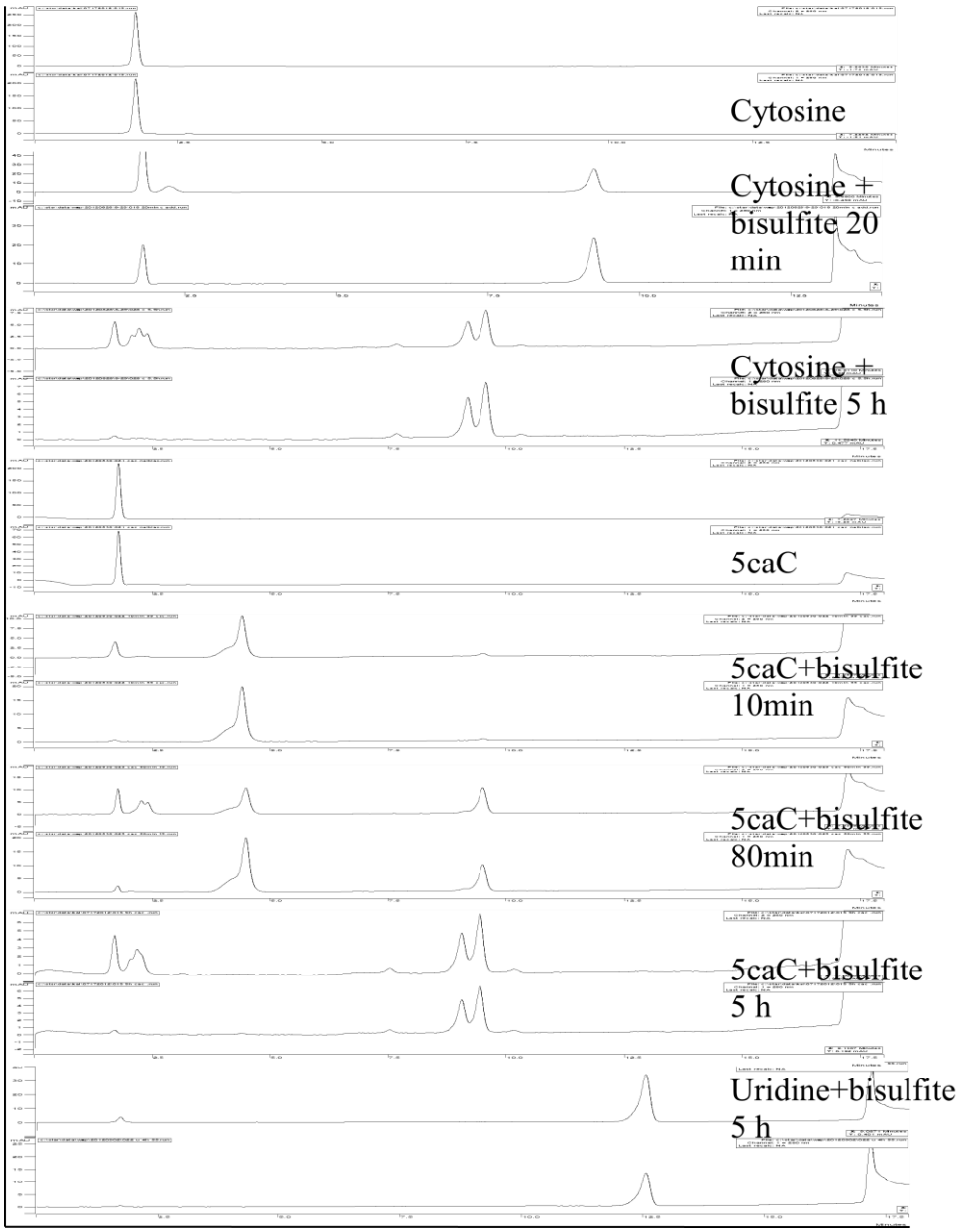

Figure S7. HPLC analysis of bisulfite treatment of normal cytosine and 5caC. 5caC bisulfite intermidiate appears to convert to the same bisulfite intermidiate from cytosine, and then to the deamination final product.

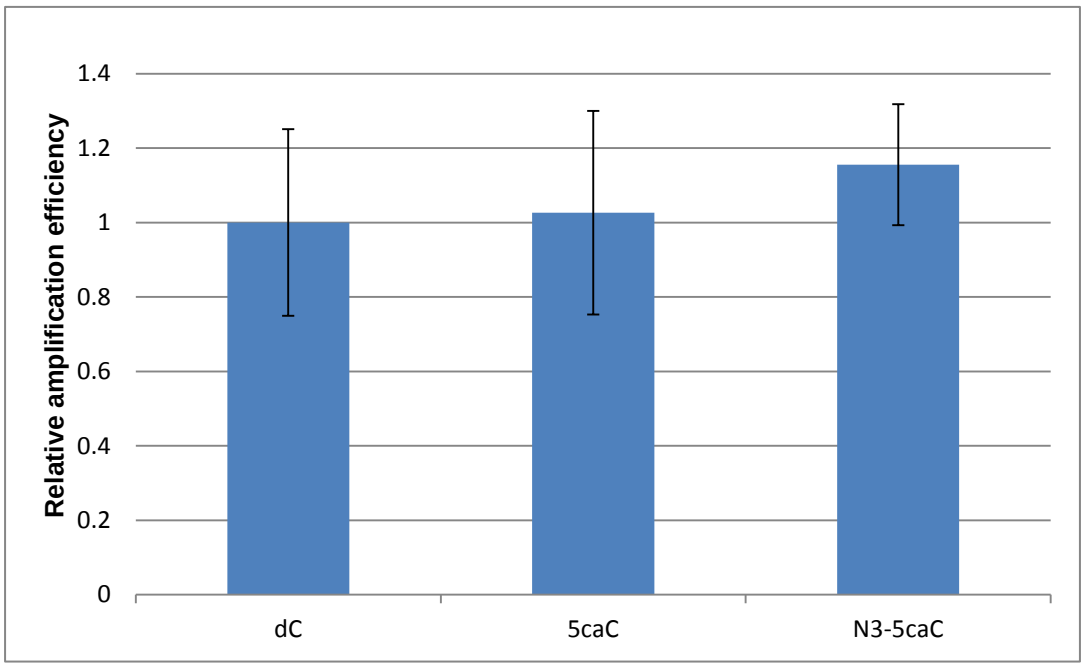

Figure S8. qPCR of 76 mer dsDNA (same sequence shown in the General Matierals section) containing a normal cytosine, 5caC, and EDC-labeled 5caC, respectively. The amplification efficiency is normalized to normal cytosine-containing DNA.

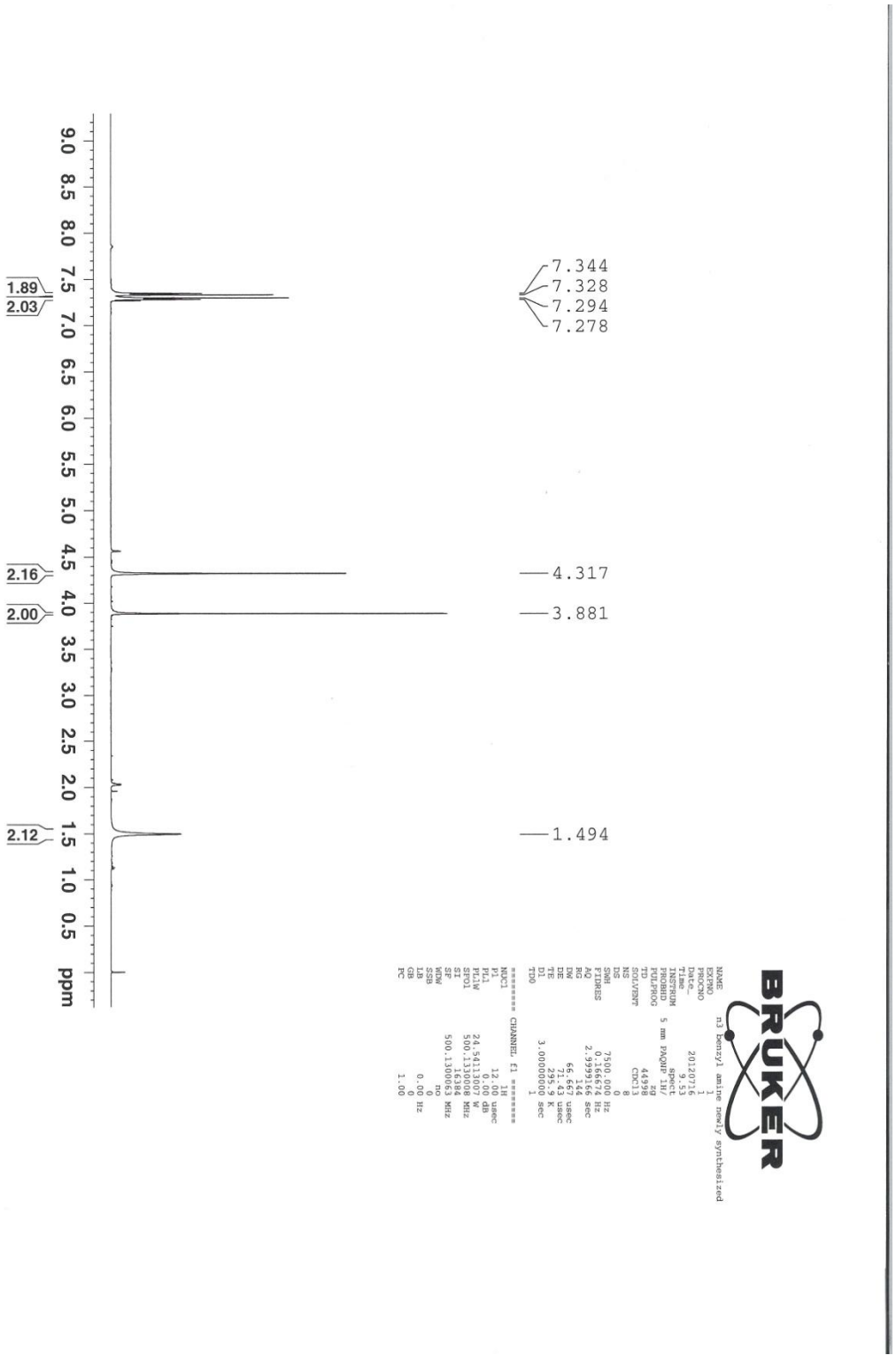

Figure S9. ^{1}HNMR spectra of (4-aminomethyl)benzylazide.

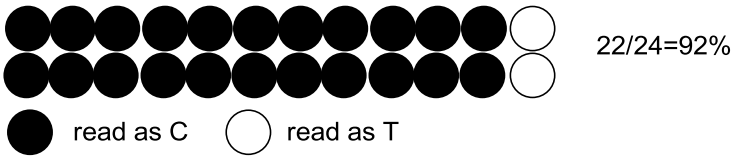

Figure S10. Colony picking and Sanger sequencing of protected 76 mer 5caC-containing oligo after bisulfite treatment.

Reference:

(1) Lau, K. N.; Chow, H. F.; Chan, M. C.; Wong, K. W. *Angew. Chem. Int. Ed.* **2008**, *47*, 6912.