

Supplementary Figures for  
A Chemical Probe Targets DNA 5-Formylcytosine Sites and Inhibits TDG Excision, Polymerases Bypass, and  
Gene Expression

Liang Xu, Ying-Chu Chen, Satoshi Nakajima, Jenny Chong, Lanfeng Wang, Li Lan,  
Chao Zhang\*, and Dong Wang\*

Figure S1. Chemical characterization of HMA

Figure S2. TLC analysis of nucleoside reactivity.

Figure S3. Reactivity of HMA with other forms of cytosine.

Figure S4. MALDI analysis of HMA modification.

Figure S5. Structure of TDG binding duplex DNA.

Figure S6. Elongation of DNA pol I through the single-strand DNA template.

Figure S7. Fluorescent behaviours of HMA-linked DNA in the presence of NTS

Figure S8. DNA pol I replication in the presence of free HMA ligand.

Figure S9. RNA pol II transcription in a short time period.

Figure S10. Construction of 5fC-contained plasmid.

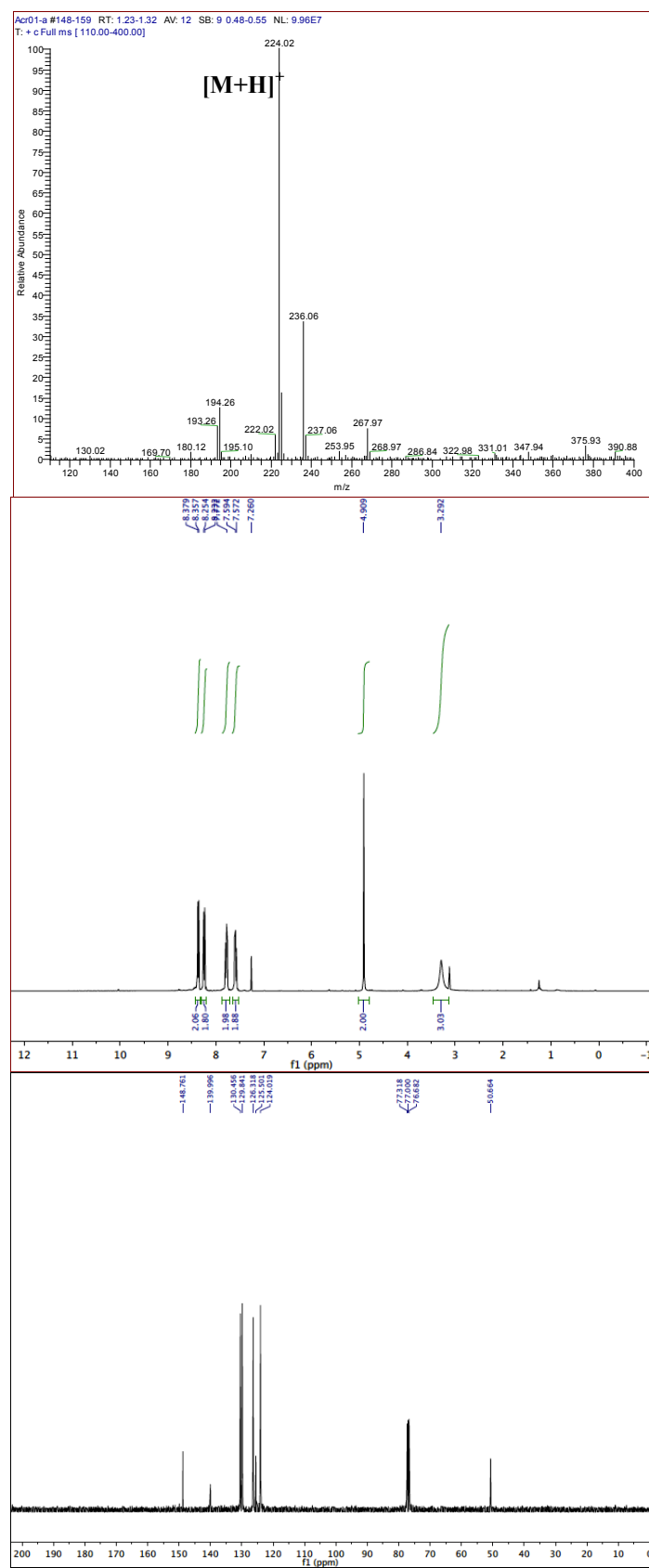
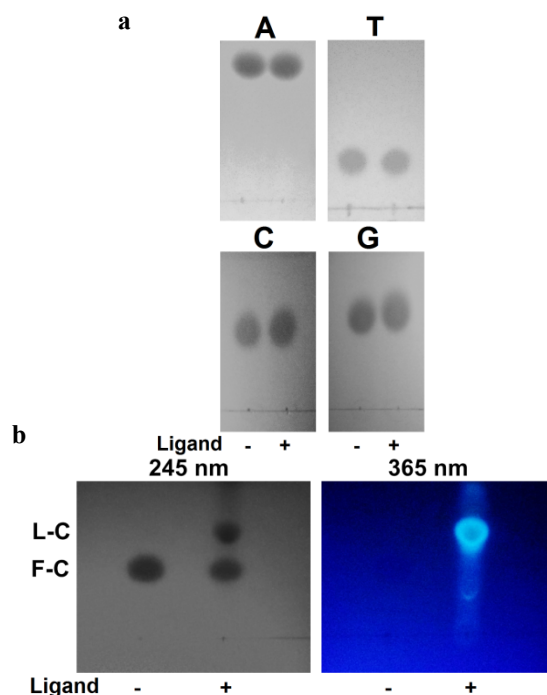
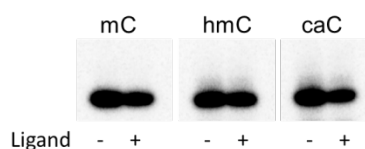


Figure S1. ESI MS,  $^1\text{H}$  and  $^{13}\text{C}$  NMR of HMA.

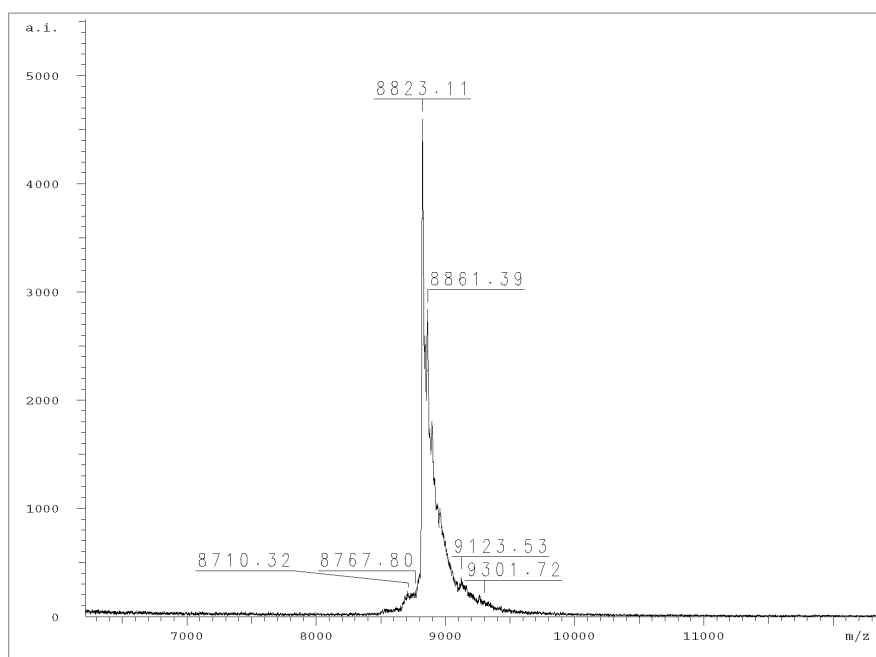


**Figure S2.** TLC images of nucleosides. (a) Images of A,U,G and C after reaction with HMA. No difference between the absence and the presence of ligand was observed. The TLC plates were run for different time periods to get obvious migrations. These images were taken under irradiation of 245 nm UV light. No spot was observed under irradiation of 365 nm UV light. (b) Images of 5-formylcytosine after reaction with HMA. We can clearly see that a new spot appeared after 1 hr reaction with HMA. This spot was also visible under irradiation of 365 nm UV light, which indicated that the polyaromatic ring was linked to this nucleoside. The free ligand migrated much faster than nucleosides, so this fluorescent spot could only be the ligand-modified cytosine.



**Figure S3.** HMA does not target other forms of cytosine. DNA containing 5mC, 5hmC and 5caC were gifts from Dr. Chuan He from University of Chicago.

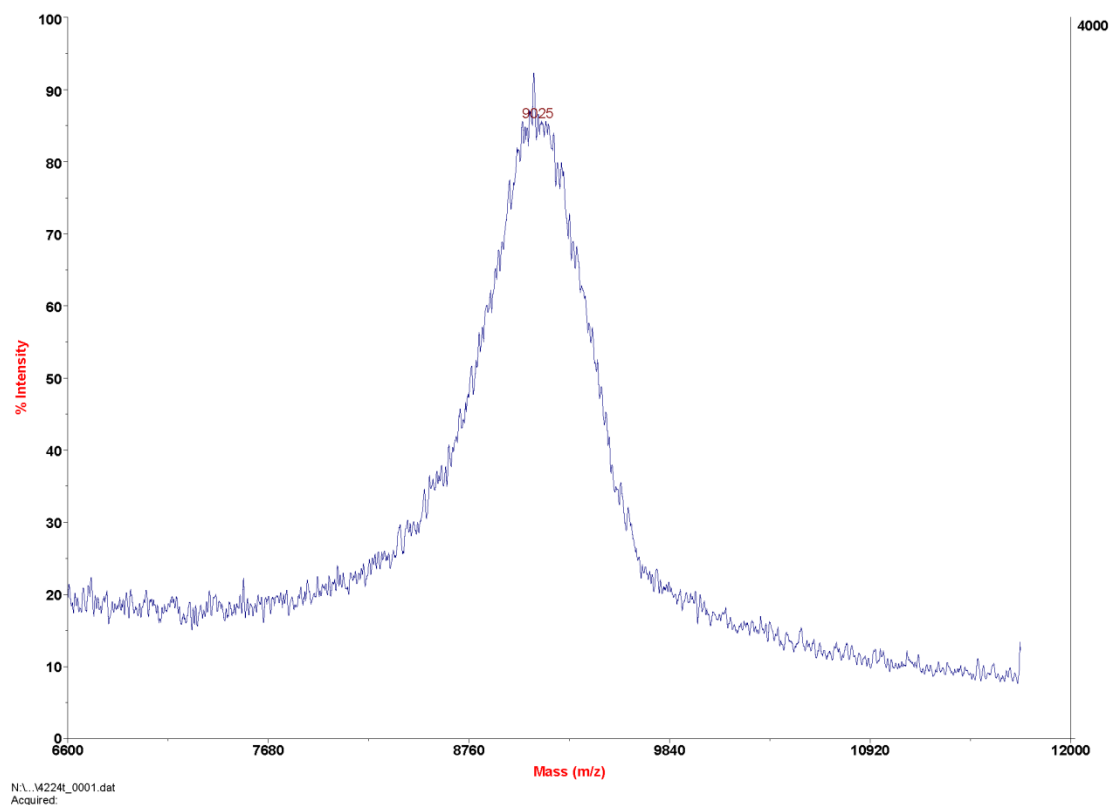
**a**



/D:/data/2012/Wanq/08202012/29FCO-HPA-Laser-75-b/ILin/pdata/1 Administrator Thu Aug 23 15:45:01 2012

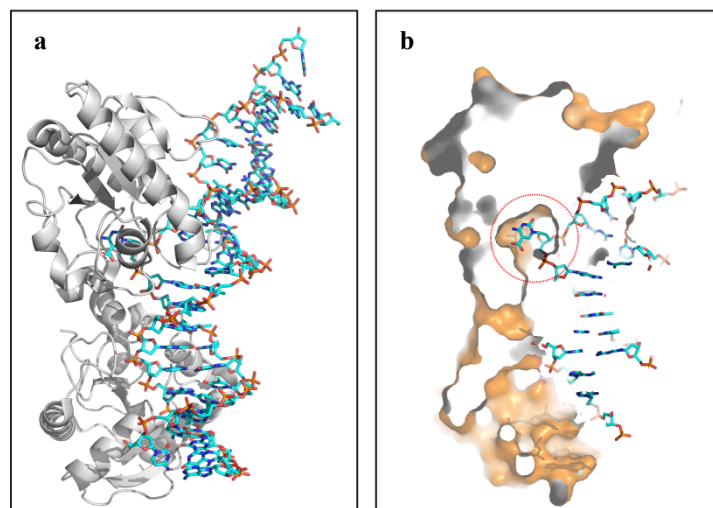
Voyager Spec #1=>SM5=>MC[BP = 1303.9, 7626]

**b**

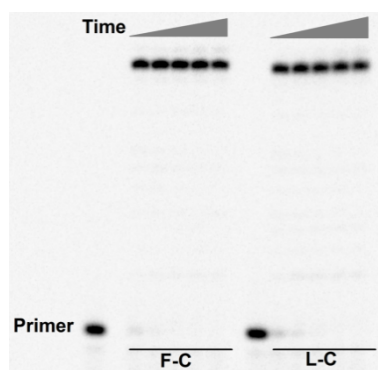


N1\_M224t\_0001.dat  
Acquired:

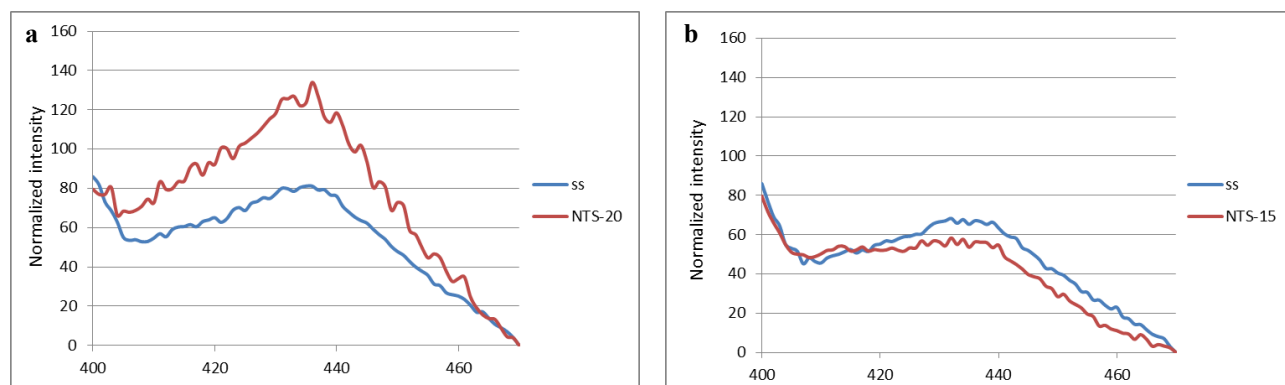
**Figure S4.** MALDI analysis of F-C DNA (a) and L-C DNA (b).



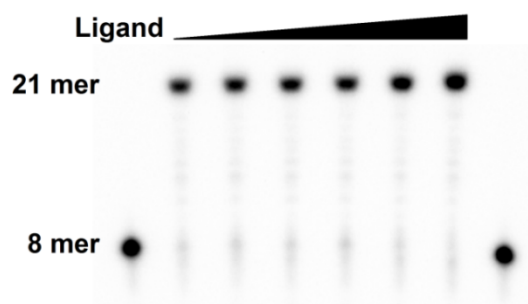
**Figure S5.** Structure of TDG binding DNA duplex with a caC site (PDB: 3UO7).<sup>[3]</sup> The flipped nucleobase binding site is marked by a red circle.



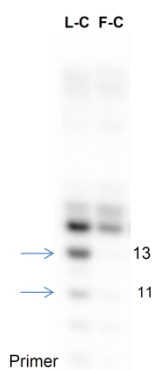
**Figure S6.** Elongation of DNA Pol I through F-C and L-C DNA templates in the absence of non-template strands. Time points here were 0.5, 1, 2, 5, and 10 min. No obvious pausing behaviors were found in this gel.



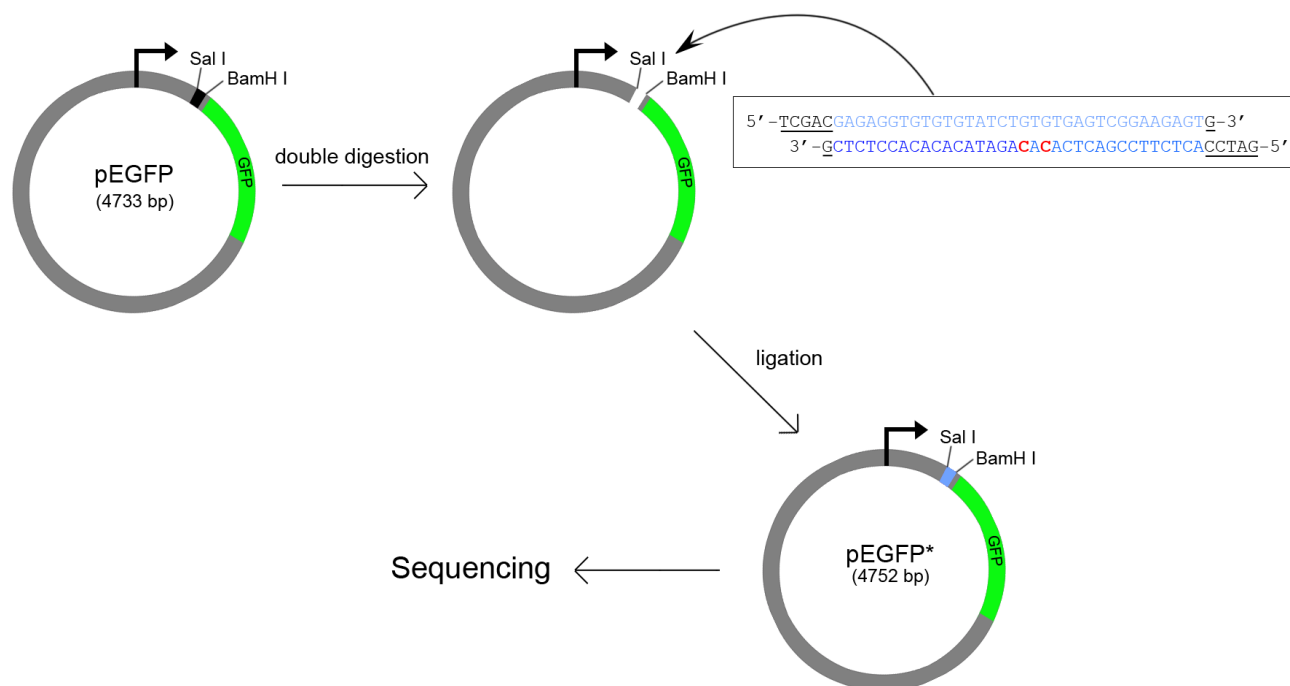
**Figure S7.** Fluorescent spectra of HMA-linked DNA in the presence of long NTS (a) and short NTS (b). Blue curves refer to single-strand L-C DNA only; red curves refer to DNA duplex structures formed by annealing with NTS-20 or NTS-15 respectively. In the presence of long NTS, as the acridine compound was involved into DNA duplex structure, we thus observed fluorescence enhancement in the presence of NTS-20. However, the short NTS in which the acridine ligand did not interact with the duplex could hardly change the fluorescence behavior of acridine. <sup>[1,2]</sup>



**Figure S8.** Elongation of DNA polymerase I through unmodified DNA template with NTS-20 as non-template strand in the presence of free ligand molecule. The concentrations of ligand are 0, 1  $\mu$ M, 10  $\mu$ M, 100  $\mu$ M, 1 mM and 10 mM from the lowest to the highest. It is clearly seen that no different patterns occur with variation of free ligand concentration.



**Figure S9.** RNA primer elongation after 1 min reaction. The pol II transient pausing at 5fC site (the 11<sup>th</sup> position) is strengthened by the HMA ligand modification. The major pausing sites on the L-C template are located at a couple of base pairs downstream from the modified site (the 13<sup>th</sup> -15<sup>th</sup> positions from 3' end).



**Figure S10.** Procedures for construction of 5fC-included plasmid. The red C refers to the 5-formylcytosine site. The modified plasmid was confirmed by sequencing.

References:

1. J. P. Schreiber, and M. P. Daune, *J. Mol. Biol.* 1974, **83**, 487.
2. K. Fukui, and K. Tanaka, *Nucleic Acids Research*, 1996, **24**, 3962.
3. L. Zhang, X. Lu, J. Lu, H. Liang, Q. Dai, G. L. Xu, C. Luo, H. Jiang, and C. He, *Nat. Chem. Biol.* 2012, **8**, 328.