

Supplementary Information

DNA Sequence and Substrate Concentration Shape TET2-Mediated Oxidation of CpG and Non-CpG Methylcytosines

Aninda Sundar Dey, Chayan Bhattacharya, Vidit Minda, Gauresh Shivji and Mridul Mukherji*

Division of Pharmacology & Pharmaceutical Sciences, University of Missouri-Kansas City,
USA

*Correspondence: Division of Pharmacology & Pharmaceutical Sciences, University of
Missouri-Kansas City, MO 64108, USA; e-mail: mukherjim@umkc.edu

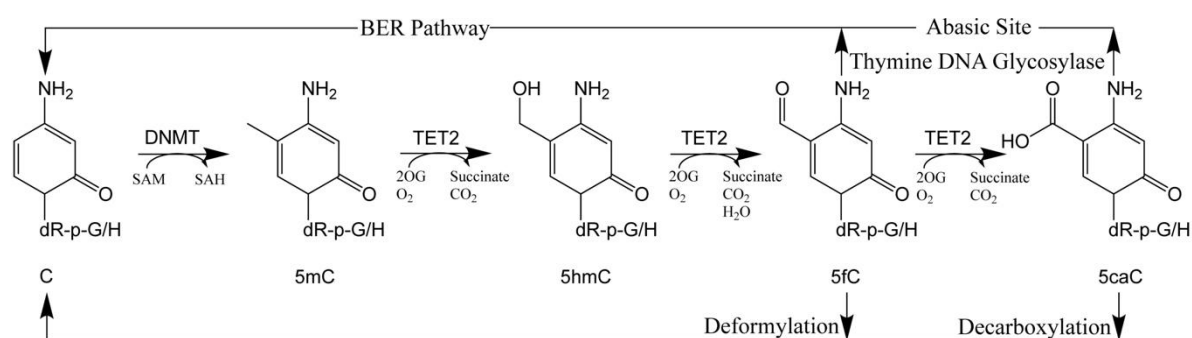


Figure S1. TET2-mediated sequential oxidation of 5mC and pathways for active DNA demethylation. TET2 sequentially oxidizes 5mC to 5hmC, 5fC, and 5caC. The higher oxidation products 5fC and 5caC can be excised by thymine DNA glycosylase (TDG), generating abasic sites that are subsequently processed through the base-excision repair (BER) pathway to restore unmodified cytosine. Alternative deformylation and decarboxylation pathways leading from 5fC and 5caC, respectively, to unmodified cytosine are also shown.

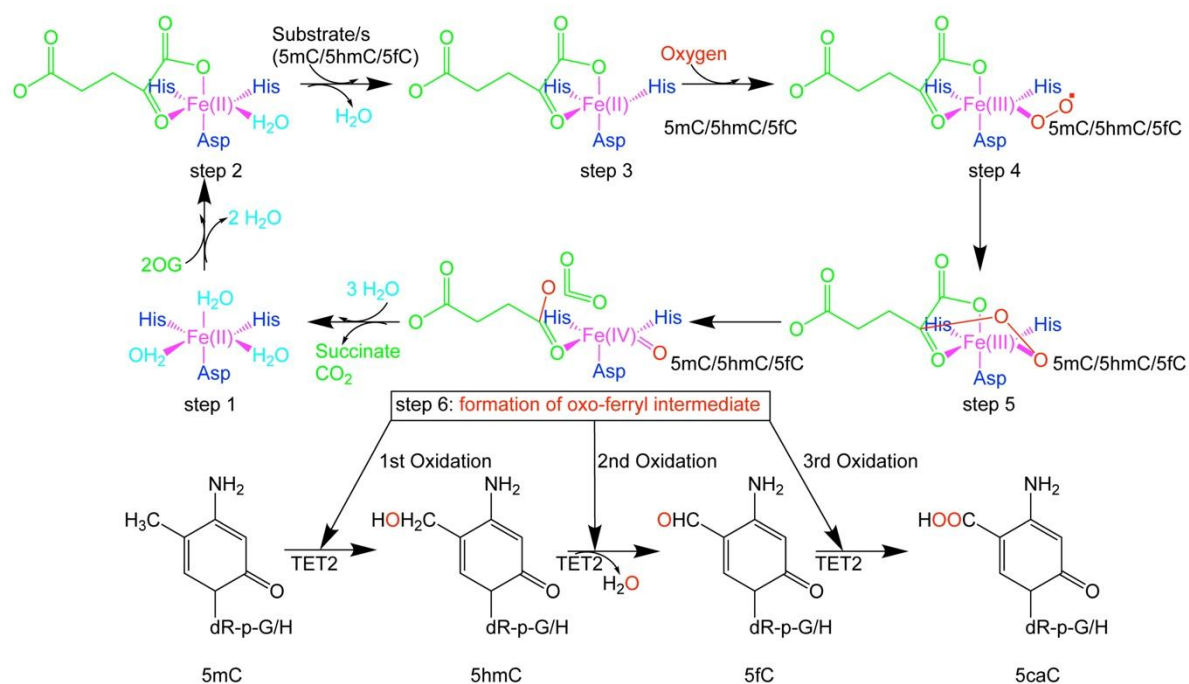


Figure S2. Catalytic cycle of TET2-mediated iterative oxidation of 5mC. TET2 catalyses successive Fe(II)- and 2-oxoglutarate-dependent oxidation cycles, generating 5hmC after the first catalytic cycle, 5fC after the second, and 5caC after the third. Accordingly, Total Oxidation Events (TOE) were calculated by weighting each product according to the cumulative number of oxidation events required for its formation: $\text{TOE} = 1 \times 5\text{hmC} + 2 \times 5\text{fC} + 3 \times 5\text{caC}$.

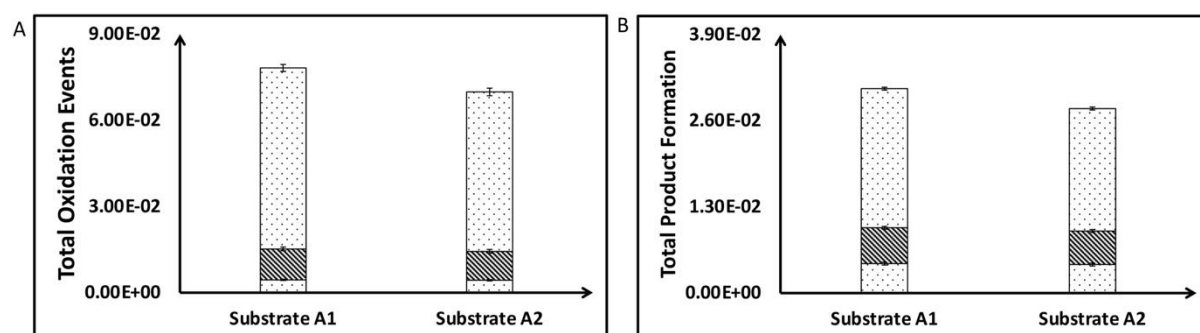


Figure S3. Comparison of TET2-mediated oxidation of the initial mCG substrates A1

and A2. (A) Total Oxidation Events (TOE) and (B) Total Product Formation (TPF) for

substrates A1 and A2. TOE and TPF were calculated from background-corrected LC-MS/MS

peak areas normalized to the deoxycytidine peak area per unmodified cytosine ($A_C / \#C$), as

described in Materials and Methods. Data represent mean $\pm$ SD from three independent

experiments.

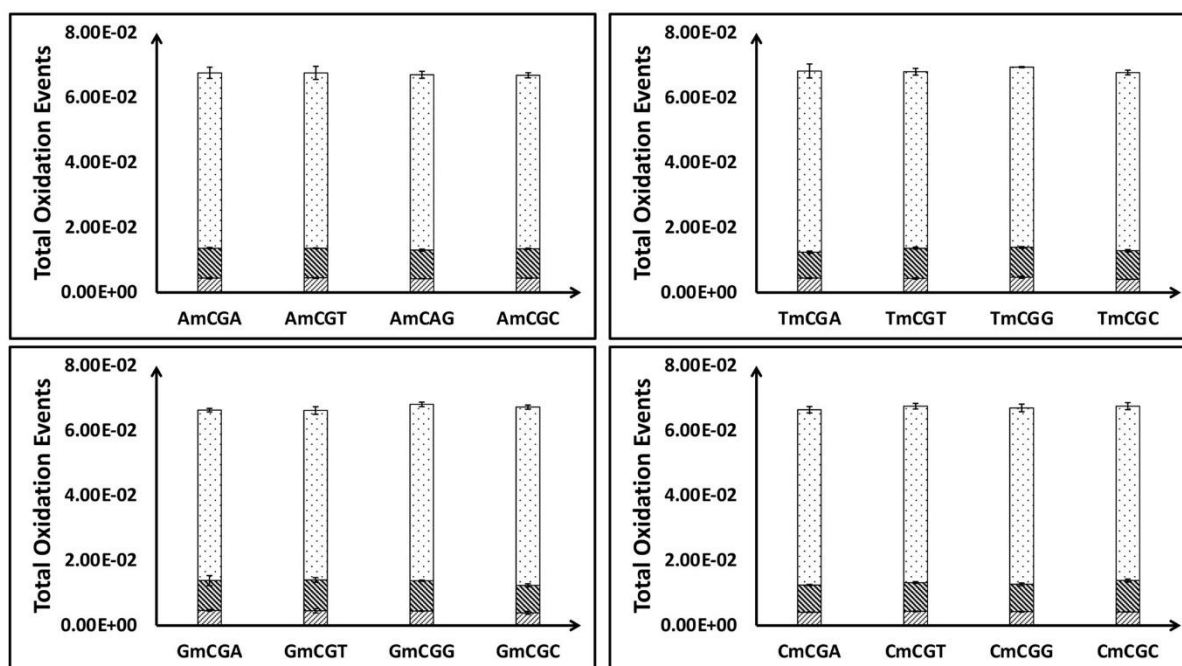


Figure S4. TET2-mediated oxidation of mCG substrates with varied –1 and +2 flanking nucleotides. Comparison of Total Oxidation Events (TOE) for all 16 25-mer dsDNA substrates containing the N–1mCGN+2 motif. Data represent mean $\pm$ SD from three independent experiments.

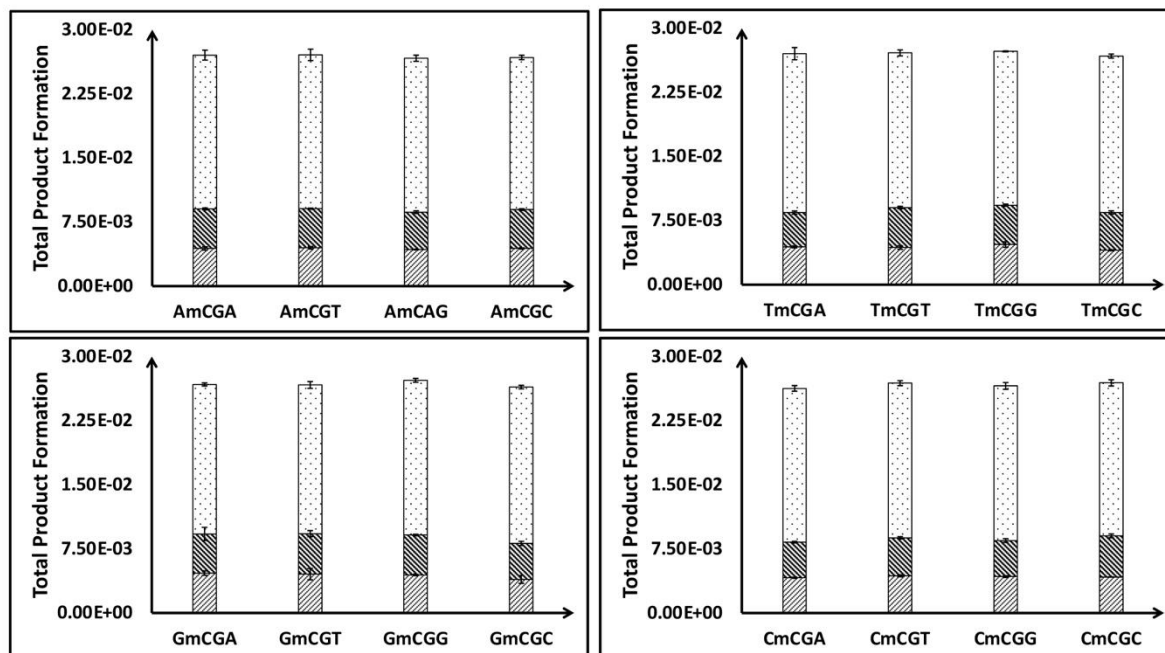


Figure S5. TET2-mediated product formation for mCG substrates with varied -1 and +2 flanking nucleotides. Comparison of Total Product Formation (TPF) for all 16 25-mer dsDNA substrates containing the N-1mCGN+2 motif. Data represent mean $\pm$ SD from three independent experiments.

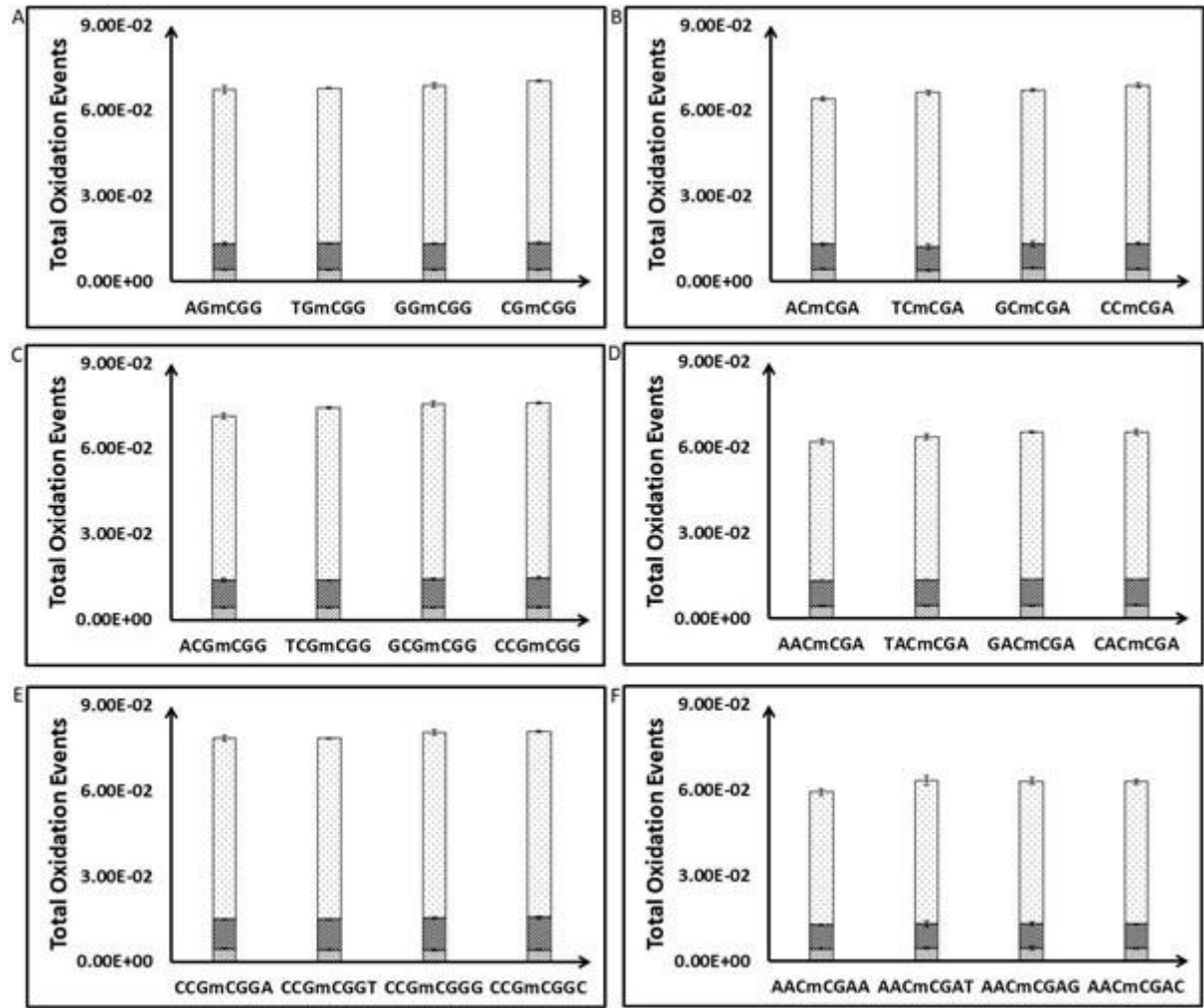


Figure S6. Stepwise analysis of distal flanking-sequence effects on TET2-mediated mCG oxidation. Total Oxidation Events (TOE) were compared following sequential variation of the -2 (A–B), -3 (C–D), and $+3$ (E–F) positions within the preferred (A, C, E) and less-preferred (B, D, F) mCG sequence contexts. Data represent mean $\pm$ SD from three independent experiments.

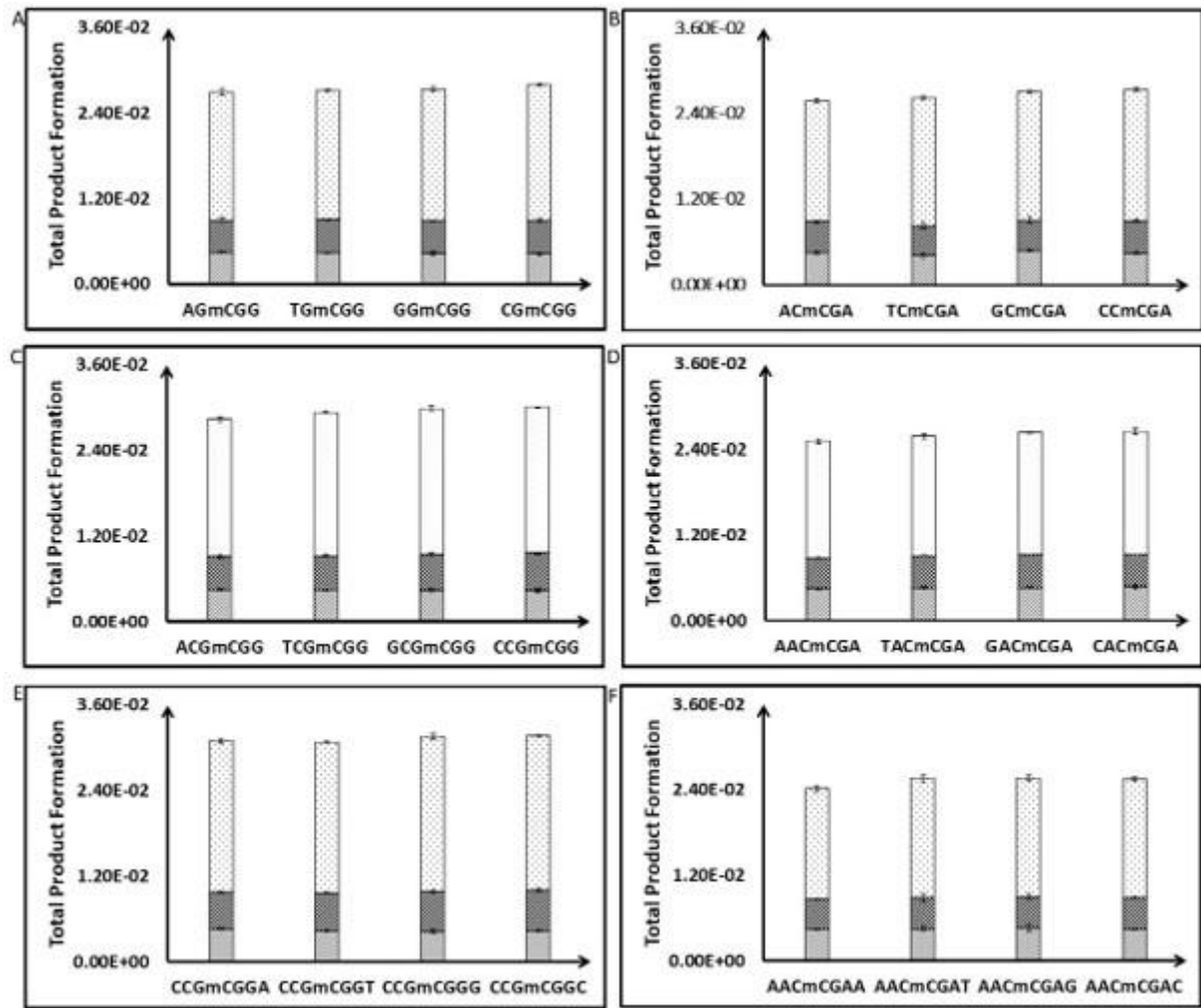


Figure S7. Stepwise analysis of distal flanking-sequence effects on TET2-mediated mCG product formation. Total Product Formation (TPF) was compared following sequential variation of the -2 (A–B), -3 (C–D), and $+3$ (E–F) positions within the preferred (A, C, E) and less-preferred (B, D, F) mCG sequence contexts. Data represent mean $\pm$ SD from three independent experiments.

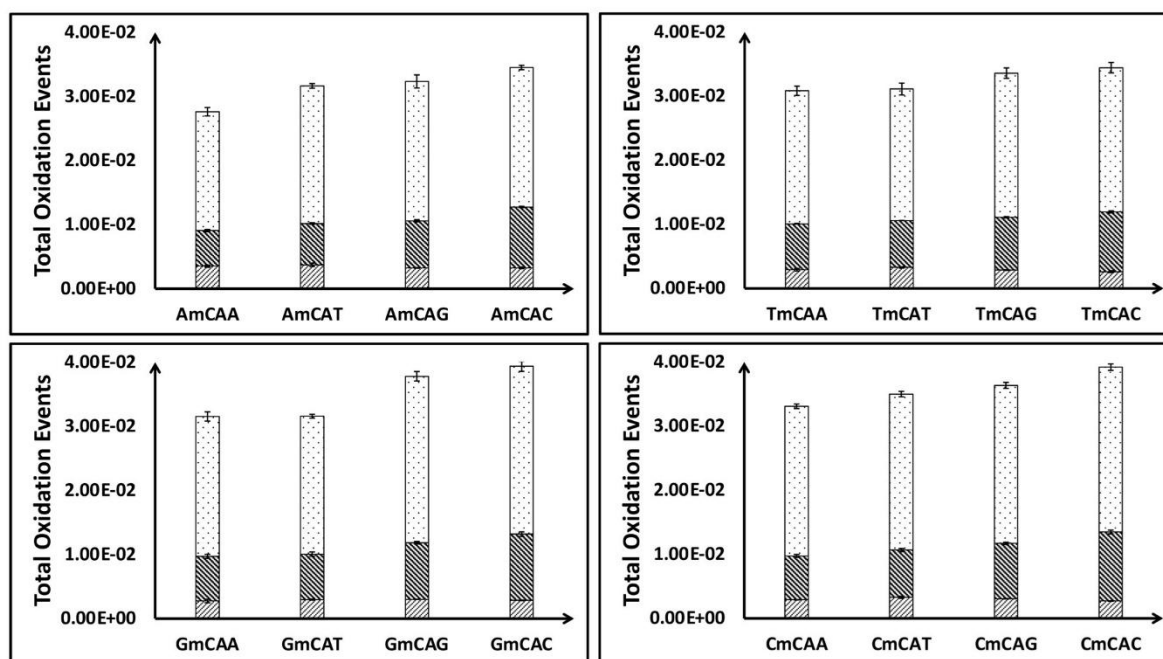


Figure S8. TET2-mediated oxidation of mCA substrates with varied -1 and +2 flanking nucleotides. Comparison of Total Oxidation Events (TOE) for all 16 25-mer dsDNA substrates containing the N-1mCAN+2 motif. Data represent mean $\pm$ SD from three independent experiments.

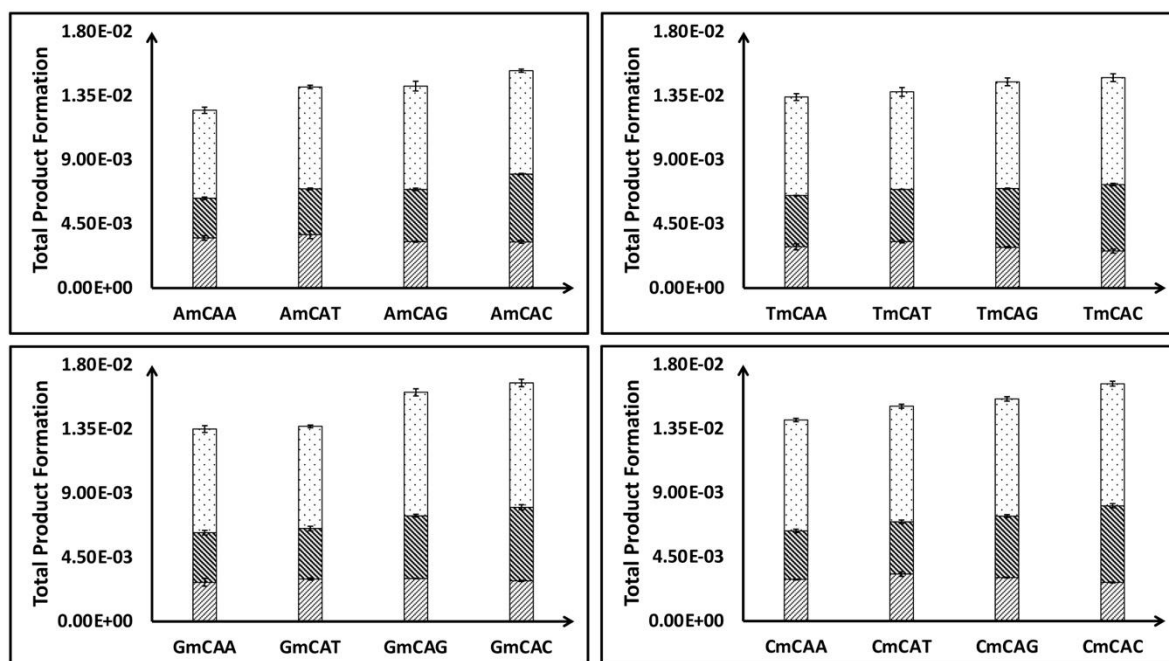


Figure S9. TET2-mediated product formation from mCA substrates with varied –1 and +2 flanking nucleotides. Comparison of Total Product Formation (TPF) for all 16 25-mer dsDNA substrates containing the N–1mCAN+2 motif. TPF represents the summed normalized formation of 5hmC, 5fC, and 5caC. Data represent mean $\pm$ SD from three independent experiments.

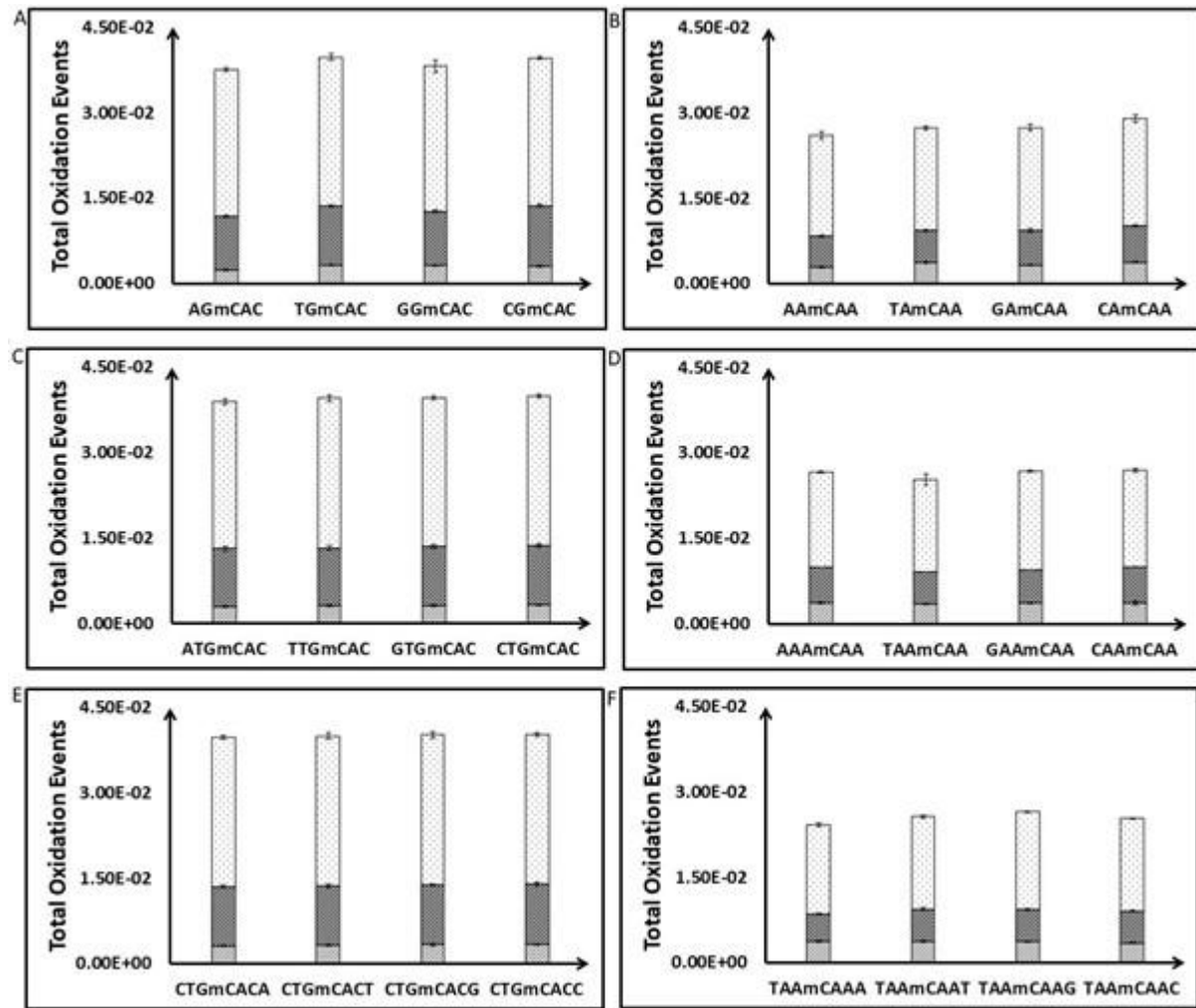


Figure S10. Stepwise analysis of distal flanking-sequence effects on TET2-mediated mCA oxidation. Total Oxidation Events (TOE) were compared following sequential variation of the -2 (A–B), -3 (C–D), and $+3$ (E–F) positions within the preferred (A, C, E) and less-preferred (B, D, F) mCA sequence contexts. Data represent mean $\pm$ SD from three independent experiments.

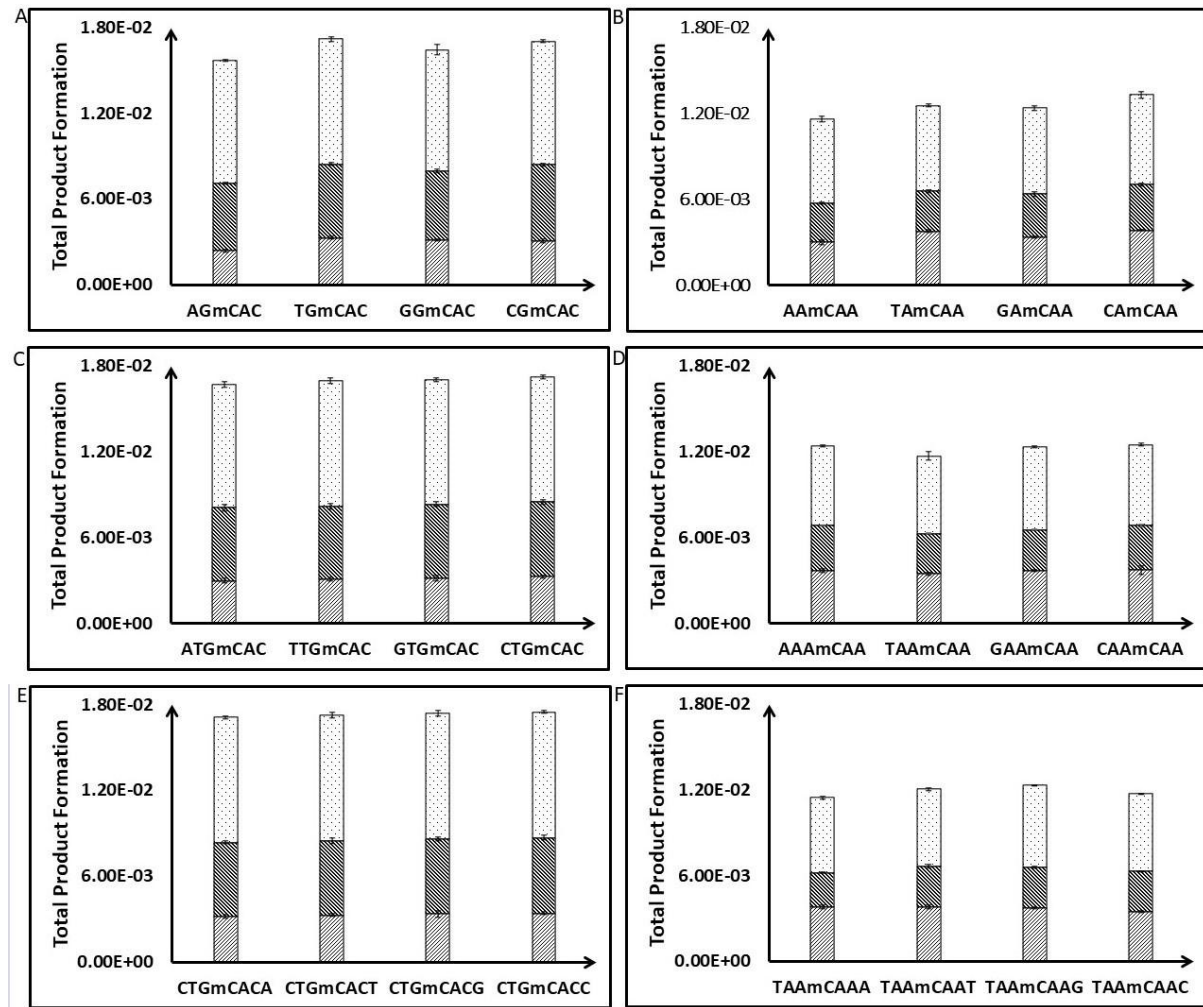


Figure S11. Stepwise analysis of distal flanking-sequence effects on TET2-mediated mCA product formation. Total Product Formation (TPF) was compared following sequential variation of the -2 (A–B), -3 (C–D), and +3 (E–F) positions within the preferred (A, C, E) and less-preferred (B, D, F) mCA sequence contexts. Data represent mean $\pm$ SD from three independent experiments.

DNA Substrate	DNA Sequence
A	
1	5'-AGCCCGCGCCGmCGCCGGTCGAGCGG-3'
2	5'-GCGCCGGTCGAGmCGGCCGCTCCCGC-3'
B	
3	5'-GCGCCGGTCGTAmCGACCGCTCCCGC-3'
4	5'-GCGCCGGTCGTAmCGTCCGCTCCCGC-3'
5	5'-GCGCCGGTCGTAmCGGCCGCTCCCGC-3'
6	5'-GCGCCGGTCGTAmCGCCCGCTCCCGC-3'
7	5'-GCGCCGGTCGTTmCGACCGCTCCCGC-3'
8	5'-GCGCCGGTCGTTmCGTCCGCTCCCGC-3'
9	5'-GCGCCGGTCGTTmCGGCCGCTCCCGC-3'
10	5'-GCGCCGGTCGTTmCGCCCGCTCCCGC-3'
11	5'-GCGCCGGTCGTGmCGACCGCTCCCGC-3'
12	5'-GCGCCGGTCGTGmCGTCCGCTCCCGC-3'
13	5'-GCGCCGGTCGTGmCGGCCGCTCCCGC-3'
14	5'-GCGCCGGTCGTGmCGCCCGCTCCCGC-3'
15	5'-GCGCCGGTCGTcmCGACCGCTCCCGC-3'
16	5'-GCGCCGGTCGTcmCGTCCGCTCCCGC-3'
17	5'-GCGCCGGTCGTcmCGGCCGCTCCCGC-3'
18	5'-GCGCCGGTCGTcmCGCCCGCTCCCGC-3'
C	
19	5'-GCGCCGGTCGAGmCGGCCGCTCCCGC-3'
20	5'-GCGCCGGTCGTGmCGGCCGCTCCCGC-3'
21	5'-GCGCCGGTCGGGmCGGCCGCTCCCGC-3'
22	5'-GCGCCGGTCGCGmCGGCCGCTCCCGC-3'
23	5'-GCGCCGGTCGACmCGACCGCTCCCGC-3'
24	5'-GCGCCGGTCGTcmCGACCGCTCCCGC-3'
25	5'-GCGCCGGTCGGcmCGACCGCTCCCGC-3'
26	5'-GCGCCGGTCGCCmCGACCGCTCCCGC-3'
D	
27	5'-GCGCCGGTCACGmCGGCCGCTCCCGC-3'
28	5'-GCGCCGGTCTCGmCGGCCGCTCCCGC-3'
29	5'-GCGCCGGTCGCGmCGGCCGCTCCCGC-3'
30	5'-GCGCCGGTCCCGmCGGCCGCTCCCGC-3'
31	5'-GCGCCGGTCAACmCGACCGCTCCCGC-3'
32	5'-GCGCCGGTCTACmCGACCGCTCCCGC-3'
33	5'-GCGCCGGTCGACmCGACCGCTCCCGC-3'
34	5'-GCGCCGGTCCACmCGACCGCTCCCGC-3'

DNA Substrate	DNA Sequence
E	
35	5'-GCGCCGGTCCCGmCGGACGCTCCCGC-3'
36	5'-GCGCCGGTCCCGmCGGTCGCTCCCGC-3'
37	5'-GCGCCGGTCCCGmCGGGCGCTCCCGC-3'
38	5'-GCGCCGGTCCCGmCGGCCGCTCCCGC-3'
39	5'-GCGCCGGTCAACmCGAACGCTCCCGC-3'
40	5'-GCGCCGGTCAACmCGATCGCTCCCGC-3'
41	5'-GCGCCGGTCAACmCGAGCGCTCCCGC-3'
42	5'-GCGCCGGTCAACmCGACCGCTCCCGC-3'
F	
43	5'-ACCCGmCGGGGGT-3'
44	5'-ACAACmCGAAGGT-3'
G	
45	5'-GCGCCGGTCGTAmCAACCGCTCCCGC-3'
46	5'-GCGCCGGTCGTAmCATCCGCTCCCGC-3'
47	5'-GCGCCGGTCGTAmCAGCCGCTCCCGC-3'
48	5'-GCGCCGGTCGTAmCACCCGCTCCCGC-3'
49	5'-GCGCCGGTCGTTmCAACCGCTCCCGC-3'
50	5'-GCGCCGGTCGTTmCATCCGCTCCCGC-3'
51	5'-GCGCCGGTCGTTmCAGCCGCTCCCGC-3'
52	5'-GCGCCGGTCGTTmCACCCGCTCCCGC-3'
53	5'-GCGCCGGTCGTGmCAACCGCTCCCGC-3'
54	5'-GCGCCGGTCGTGmCATCCGCTCCCGC-3'
55	5'-GCGCCGGTCGTGmCAGCCGCTCCCGC-3'
56	5'-GCGCCGGTCGTGmCACCCGCTCCCGC-3'
57	5'-GCGCCGGTCGTCmCAACCGCTCCCGC-3'
58	5'-GCGCCGGTCGTCmCATCCGCTCCCGC-3'
59	5'-GCGCCGGTCGTCmCAGCCGCTCCCGC-3'
60	5'-GCGCCGGTCGTCmCACCCGCTCCCGC-3'
H	
61	5'-GCGCCGGTCGAGmCACCCGCTCCCGC-3'
62	5'-GCGCCGGTCGTGmCACCCGCTCCCGC-3'
63	5'-GCGCCGGTCGGGmCACCCGCTCCCGC-3'
64	5'-GCGCCGGTCGCGmCACCCGCTCCCGC-3'
65	5'-GCGCCGGTCGAAmCAACCGCTCCCGC-3'
66	5'-GCGCCGGTCGTAmCAACCGCTCCCGC-3'
67	5'-GCGCCGGTCGGAmCAACCGCTCCCGC-3'
68	5'-GCGCCGGTCGCAmCAACCGCTCCCGC-3'

DNA Substrate	DNA Sequence
I	
69	5' -GCGCCGGTCATGmCACCCGCTCCCGC-3'
70	5' -GCGCCGGTCTTGmCACCCGCTCCCGC-3'
71	5' -GCGCCGGTCGTGmCACCCGCTCCCGC-3'
72	5' -GCGCCGGTCCTGmCACCCGCTCCCGC-3'
73	5' -GCGCCGGTCAAAmCAACCGCTCCCGC-3'
74	5' -GCGCCGGTCTAAmCAACCGCTCCCGC-3'
75	5' -GCGCCGGTCGAAmCAACCGCTCCCGC-3'
76	5' -GCGCCGGTCCAAmCAACCGCTCCCGC-3'
J	
77	5' -GCGCCGGTCCTGmCACACGCTCCCGC-3'
78	5' -GCGCCGGTCCTGmCACTCGCTCCCGC-3'
79	5' -GCGCCGGTCCTGmCACGCGCTCCCGC-3'
80	5' -GCGCCGGTCCTGmCACCCGCTCCCGC-3'
81	5' -GCGCCGGTCTAAmCAAACGCTCCCGC-3'
82	5' -GCGCCGGTCTAAmCAATCGCTCCCGC-3'
83	5' -GCGCCGGTCTAAmCAAGCGCTCCCGC-3'
84	5' -GCGCCGGTCTAAmCAACCGCTCCCGC-3'
K	
85	5' -ACCTGmCACCGGT-3'
86	5' -ACTAAmCAAAGGT-3'
L	
87	5' -GCGCCGGTCCCGmCGGGCGCTCCCGC-FAM-3'
88	5' -GCGCCGGTCAACmCGAACGCTCCCGC-FAM-3'
89	5' -GCGCCGGTCCTGmCACCCGCTCCCGC-FAM-3'
90	5' -GCGCCGGTCTAAmCAAACGCTCCCGC-FAM-3'

Table S1A-L. List of substrates used in this study.