Supplemental Figures:

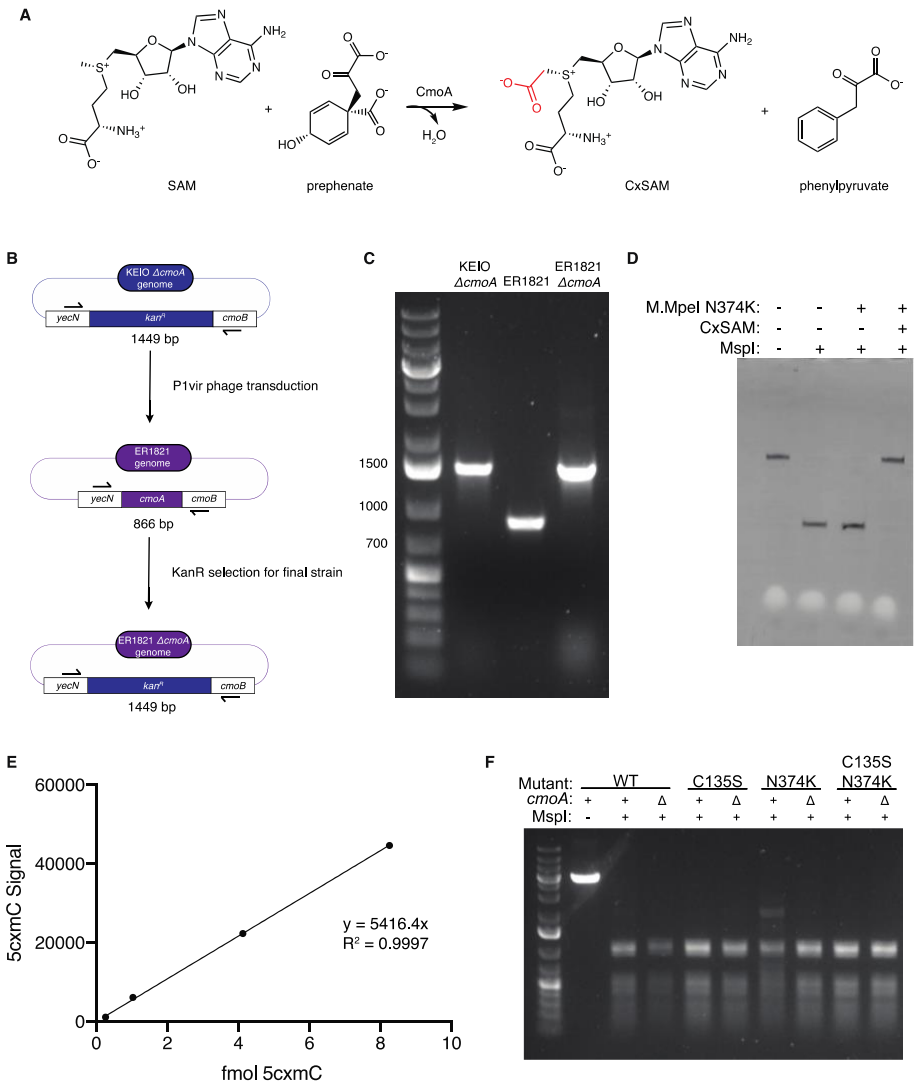

Figure S1, related to Figure 2: Validation of CmoA knockout and LC-MS/MS standard. (A) CmoA catalyzes the reaction between SAM and prephenate to yield water, CxSAM, and phenylpyruvate (Kim et al., 2013). (B) Transduction with bacteriophage P1vir was used to move a kanamycin resistance cassette from KEIO *ΔcmoA* to ER1821 cells (Baba et al., 2006) (C) Colony PCR was used to validate ER1821 *ΔcmoA* cells with primers flanking the endogenous *cmoA* locus. (D) MspI assay (see methods) confirming >98% carboxymethylation of 5cxmC oligonucleotide standard. When interpreted with Figure 1C, this experiment additionally shows that 5cxmC completely blocks MspI digestion, suggesting that this modification has a new gain-of-function ability to protect genomic DNA from restriction digestion in a way that 5mC cannot. (E) Representative LC-MS/MS standard curve for 5cxmC standard. (F) *In vivo* modification assay as described in Figure 1B. Upon overexpression, M.MpeI modifies its encoding plasmid using endogenous SAM or CxSAM. The modification can be assessed after plasmid isolation and restriction digestion. An MspI-resistant modification appears only in the presence of both M.MpeI N374K and an intact *cmoA* gene.

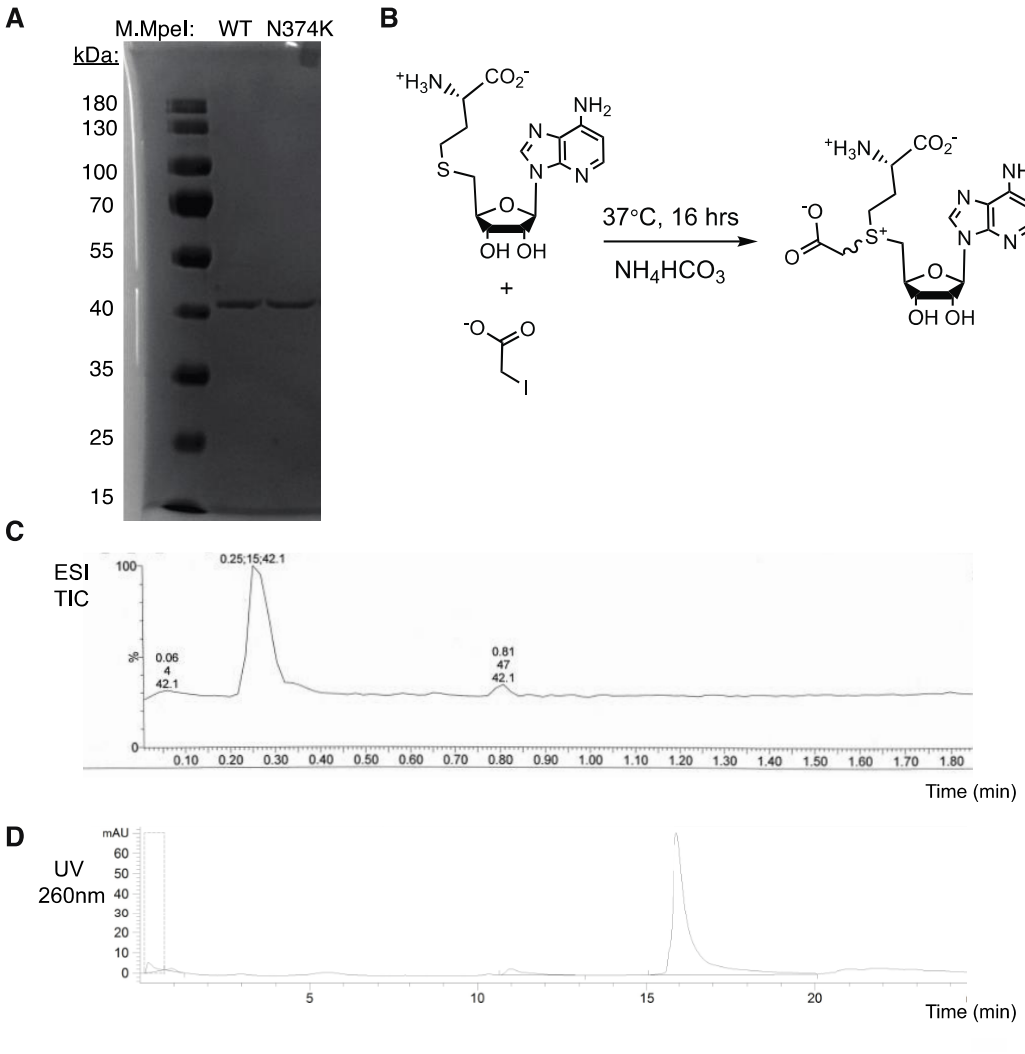

Figure S2, related to Figure 3: Validation of *in vitro* reagents. (A) SDS-PAGE of M.MpeI enzymatic preps used in this study. (B) Chemical synthesis of CxSAM. (C) LC-MS ESI$^+$ Total Ion Current (TIC) signal with observed mass of 443.5. In addition to the trace shown, HRMS was also obtained, identifying a mass of 443.1360 (mDa = -0.2, PPM = -0.5, Theoretical Mass: 443.1343). (D) HPLC purification of CxSAM showing single UV 260 nm peak.

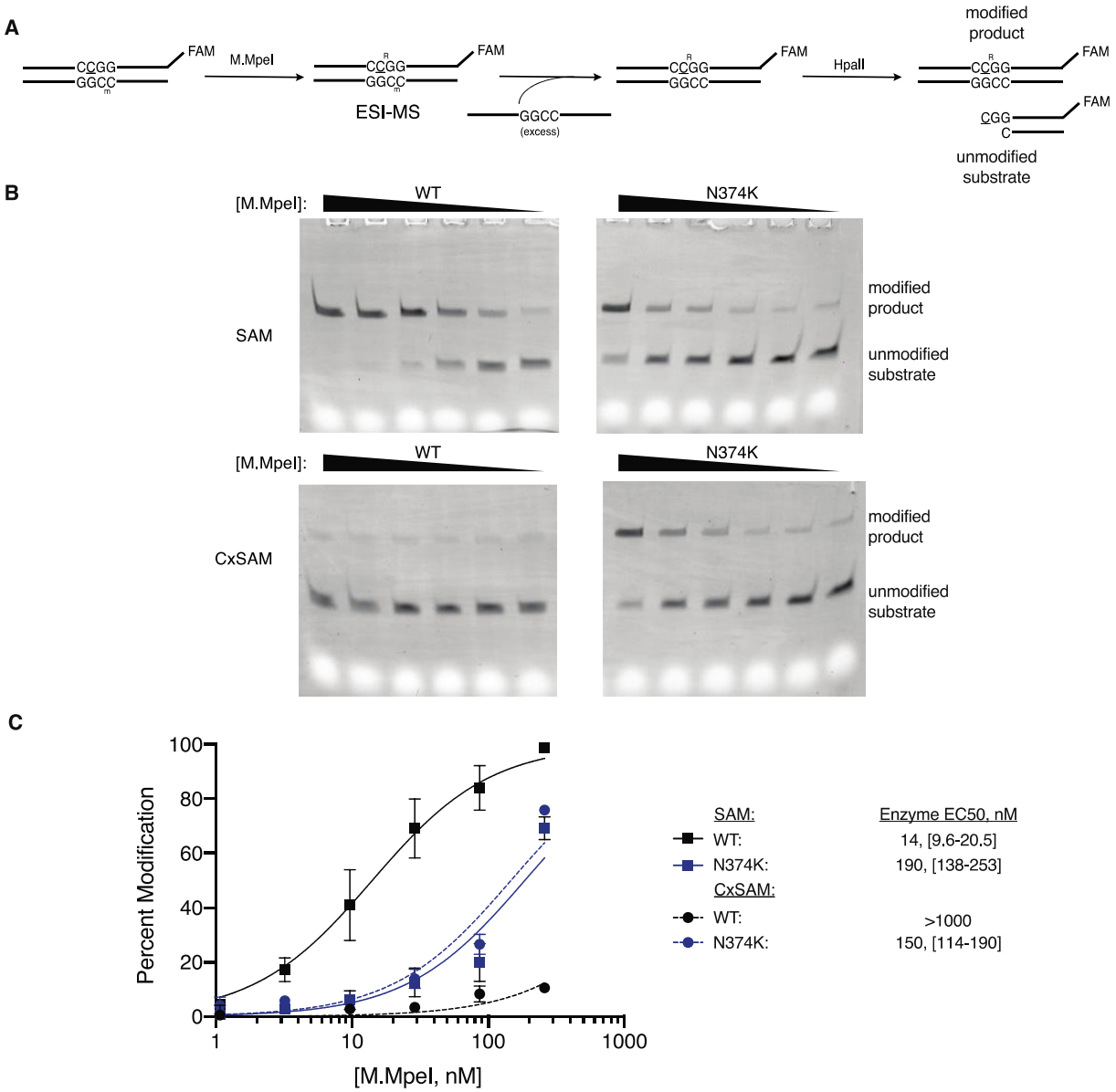

Figure S3, related to Figure 3: Quantitative oligonucleotide assay. Assay design was previously validated with homologous methyltransferases (Nabel et al., 2017). a) M.MpeI N374K was incubated with excess SAM or CxSAM and a hemimethylated CpG substrate containing a fluorophore label as shown. ESI-MS was obtained to confirm carboxymethylation of the hemimethylated substrate (expected top strand: 8877.9, observed top strand: 8876.7). HpaII digest was used to visualize total modification of the top strand after bottom strand exchange. b) Representative oligonucleotide assay gels. c) Enzyme dilution curve showing quantitative, relative activities of M.MpeI WT and N374K towards SAM and CxSAM. Points represent mean ± s.e. (n = 3 independent replicates). EC_{50} values were calculated, and 95% Confidence Intervals are reported in brackets.

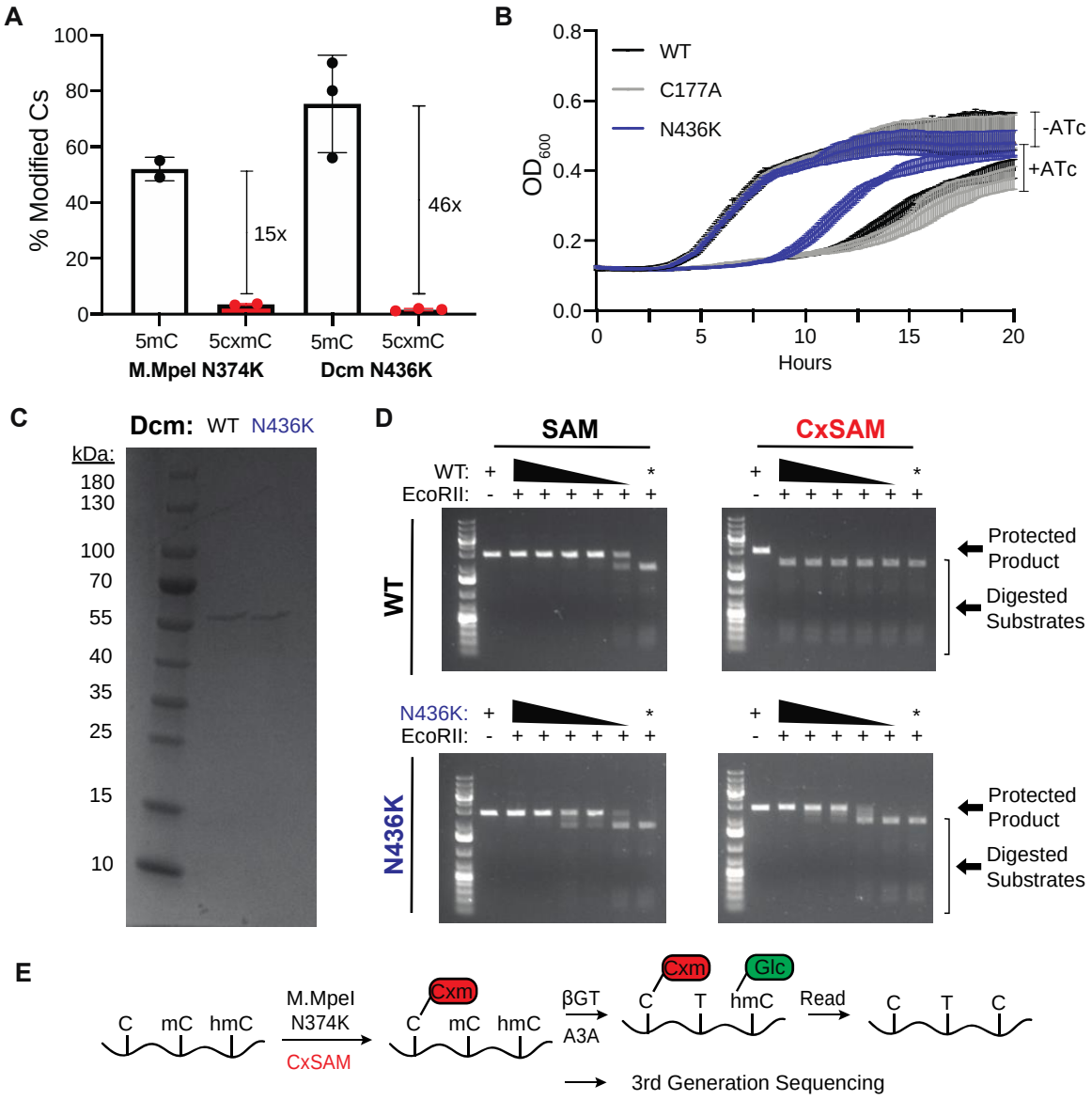

Figure S4, related to Figure 4: Generalizability and applications of DNA carboxymethyltransferases. (A) LC-MS/MS detection of 5mC and 5cxmC nucleosides catalyzed by M.MpeI N374K and Dcm N436K *in vivo* overexpression. M.MpeI experiments show percent of CpG modification on plasmid DNA. Dcm experiments are shown as a percent of CCWGG modification on gDNA. The relative 5mC vs 5cxmC modification levels are also labelled. These are the same experimental data as Figure 2E and Figure 4D, respectively. (B) Growth curves for *dam*⁻/*dcm*⁻ *E. coli* upon transformation with Dcm expression plasmids. Overnight cultures were diluted into a plate reader and monitored for 20 hrs using OD_{600}. Protein overexpression was induced with or without anhydrotetracycline (ATc). Points represent mean ± s.d. (n = 3 biological replicates). (C) SDS-PAGE of Dcm enzymatic preps used in this study. (D) Unmodified pUC19 plasmid DNA was incubated with an excess of SAM or CxSAM and 3-fold serial dilutions of Dcm WT or the N436K mutant to yield methylated or carboxymethylated DNA.

The negative control lane (*) contains the highest concentration of enzyme with no SAM or CxSAM substrate. Digestion with EcoRII fragments only unmodified DNA, allowing for visualization of digested substrate vs protected product. Dcm N436K transfers both CxSAM and SAM at comparable efficiency *in vitro* while Dcm WT only transfers SAM. (E) CpG carboxymethyltransferase can create 5-carboxymethylcytosine, which could permit differentiation of C, mC, and hmC by either enzymatic epigenetic sequencing methods or third generation sequencing (Schutsky et al., 2018).