

Carbamoyltransferase Enzyme Assay: *In vitro* Modification of 5-hydroxymethylcytosine (5hmC) to 5-carbamoyloxymethylcytosine (5cmC)

Weiwei Yang¹, Nan Dai¹, Yu-Cheng Lin[§], William Johnson¹, Romualdas Vaisvila¹, Peter Weigle¹, Yan-Jiun Lee¹, Ira Schildkraut¹, Ivan R. Corrêa Jr.¹ and Laurence Ettwiller^{1,*}

¹New England Biolabs, Ipswich, United States

[§]Present address: School of Dentistry, National Yang Ming Chiao Tung University, Taiwan

*For correspondence: ettwiller@neb.com

Abstract

Nucleic acids in living organisms are more complex than the simple combinations of the four canonical nucleotides. Recent advances in biomedical research have led to the discovery of numerous naturally occurring nucleotide modifications and enzymes responsible for the synthesis of such modifications. In turn, these enzymes can be leveraged towards toolkits for DNA and RNA manipulation for epigenetic sequencing or other biotechnological applications. Here, we present the protocol to obtain purified 5-hydroxymethylcytosine carbamoyltransferase enzymes and the associated assays to convert 5-hydroxymethylcytosine to 5-carbamoyloxymethylcytosine *in vitro*. We include detailed assays using DNA, RNA, and single nucleotide/deoxynucleotide as substrates. These assays can be combined with downstream applications for genetic/epigenetic regulatory mechanism studies and next-generation sequencing purposes.

Keywords: DNA modification, RNA modification, 5-hydroxymethylcytosine, 5-carbamoyloxymethylcytosine, Carbamoyltransferase

This protocol was validated in: eLife (2021), DOI: 10.7554/eLife.70021

Background

Modification and hyper modification of canonical nucleotides can be found in all domains of life. In bacteriophages, nucleotide modifications are often used as a strategy to evade bacterial immunity, most notably the restriction/modification system. Thus, bacteriophages tend to have a broad array of modifications present throughout their genomes. For example, the genome of T4 bacteriophages uses hydroxymethylated deoxycytidine triphosphate (dCTP) as substrate during DNA synthesis, effectively replacing dC with 5dhmC. Incorporated 5dhmC is further post-synthetically glycosylated, resulting in DNA containing glycosylated 5-hydroxymethylcytosine (Lehman and Pratt, 1960). Recent studies have aimed at expanding the *repertoire* of known modifications by mining the microbiomes for DNA-modifying enzymes (Ferrer et al., 2005).

In a previous study, we developed the Metagenomics Genome–Phenome Association (MetaGPA) framework to associate genetic data with phenotypic traits at the level of an entire microbiome (Yang et al., 2021). As a case study, we applied MetaGPA to sewage and ocean microbiome samples and successfully identified enzymes associated with cytosine modifications in the DNA. Experimental follow-up on these associations allowed us to identify and validate a novel DNA/RNA modifying enzyme, the 5-hydroxymethylcytosine carbamoyltransferase. This enzyme catalyzes the formation of a previously unknown form of cytosine modification, the 5-carbamoyloxymethylcytosine, using 5-hydroxymethylcytosine in DNA, RNA, or single nucleotides as substrates. Here, we describe the protocol for cloning, expressing, and purifying the enzyme from *E. coli*. We also describe the protocol for assaying this enzyme *in vitro* on both nucleic acid polymers and single nucleotides, which greatly expands its potential applications. As shown previously (Yang et al., 2021), the carbamoyltransferase reaction can be included in a next-generation sequencing-based method for genomic mapping of modified cytosines. Other possible applications include genetic engineering, DNA/RNA polymerase kinetics, and post-synthetic modification regulations.

Materials and Reagents

A. Consumables

1. Low retention 10 µL tips (VWR, catalog number: 89135-572)
2. Low retention 200 µL tips (VWR, catalog number: 76175-406)
3. Low retention 1,000 µL tips (VWR, catalog number: 76175-412)
4. 1.5 mL microcentrifuge tube (Eppendorf, catalog number: 022363204)
5. 8 strip 0.2 mL PCR tubes (VWR, catalog number: 20170-010)
6. 15 mL centrifuge tubes (VWR, catalog number: 21008-103)
7. 50 mL centrifuge tubes (VWR, catalog number: 21008-242)
8. 5 mL serological pipets (VWR, catalog number: 89130-908)
9. 10 mL serological pipets (VWR, catalog number: 89130-888)
10. 25 mL serological pipets (VWR, catalog number: 89130-900)
11. Weigh boats (any source)
12. Gloves (any source)
13. Oligo clean & concentrator kit (Zymo Research, catalog number: D4060)
14. Oligo clean-up and concentrator kit (Norgen Biotek, catalog number: 34100)

B. For carbamoyltransferase enzyme expression and purification

1. Bacteria culture tube (any source)
2. Carbamoyltransferase plasmid pET28-Kan-his-CT (available upon request)
3. Competent DH5- α *E. coli* cells (New England Biolabs, catalog number: C2987H)
4. T7 Express competent *E. coli* cells (New England Biolabs, catalog number: C2566)
5. 2 L flask (VWR)

6. Kanamycin stock solution, 40 mg/mL (self-prepared, powder from Americanbio, catalog number: CU01100-00200)
7. IPTG, 500 mM (self-prepared, powder from TEKnova, catalog number: I3305)
8. Phenylmethanesulfonyl fluoride (PMSF), 1 M (self-prepared, powder from MilliporeSigma, catalog number: 78830)
9. HisTrap HP 5 mL column (Cytiva, catalog number: 17524801)
10. 10%–20% Tris-Glycine SDS-PAGE gel (Thermo, Invitrogen, catalog number: XP10205BOX)
11. Tris-glycine SDS running buffer, 10× (Thermo, Invitrogen, catalog number: LC2675)
12. Protein ladder (New England Biolabs, catalog number: P7717S)
13. Blue protein loading dye with reducing reagent (New England Biolabs, catalog number: B7703S)
14. DTT (MilliporeSigma, catalog number: D9779)
15. SimplyBlue SafeStain reagent (Thermo, Life Technologies, catalog number: LC6065)
16. HiTrap Q 5 mL column (Cytiva, catalog number: 17115401)
17. Bradford protein assay kit (Thermo, catalog number: PI23200)
18. Dialysis tubing, 6–8 kD (VWR, catalog number: 28170-138)
19. LB agar plate containing 40 µg/mL kanamycin (self-prepared)
20. LB media (see Recipes)
21. Iron II solution (see Recipes)
22. Carbamoylphosphate solution (see Recipes)
23. TE buffer (see Recipes)
24. Cell lysis buffer (see Recipes)
25. HisTrap column buffer A (see Recipes)
26. HisTrap column buffer B (see Recipes)
27. HiTrap Q column buffer A (see Recipes)
28. HiTrap Q column buffer B (see Recipes)
29. Dilution buffer (see Recipes)
30. Dialysis buffer (see Recipes)
31. Ultrapure EDTA, 0.5 M, pH 8.0 (ThermoFisher, catalog number: 15575020), store at RT
32. Tris-HCl, 1 M, pH 8.0 (Amresco, catalog number: E298-NEB-2L), store at RT
33. Tris-HCl, 1 M, pH 7.5 (Amresco, catalog number: E199-NEB-2L), store at RT
34. NaCl (MilliporeSigma, catalog number: 203505), store at RT
35. Glycerol (self-prepared), store at RT
36. Imidazole (MilliporeSigma, catalog number: I2399), store at RT
37. Tryptone (MilliporeSigma, catalog number: T7293), store at RT
38. Yeast extract (MilliporeSigma, catalog number: Y1625), store at RT
39. Nuclease-free water (New England Biolabs, catalog number: B1500), store at RT
40. Ammonium iron (ii) sulfate hexahydrate, 500 µM (MilliporeSigma, catalog number: 203505), store at RT

C. For carbamoyltransferase enzyme assay

1. Phage T4gt genomic DNA (self-prepared), store in TE buffer at -20 °C
Note: E. coli bacteriophage T4 is known to have cytosine glucosylation and hydroxymethylation in its genome (Flodman et al., 2020). T4gt is a mutant deficient in DNA glucosyltransferase, and thus its genomic DNA contains 5-hydroxymethylcytosine. Up to 95% of cytosines in the T4gt genomic DNA are replaced by 5-hydroxymethylcytosine. The propagation of T4gt was done using T7 express (NEB, catalog number: C2566), and the bacterial host was grown in LB medium at 37 °C. The host was infected with phage at the log phase until the culture reached lysis. The phages in supernatant were pelleted by 10% PEG8000 and centrifugated. The genomic DNA of T4gt phage was extracted using phenol/chloroform.
2. 5hmC DNA oligo: 5'-TGTCCGATAGACT/5dhmC/TACGCA (IDT, 5'-TGTCCGATAGACT/i5HydMe-dC/TACGCA), store at -20 °C
3. 5hmC RNA oligo: 5'-GGAGTGAGAAGATGGT/5hmC/TAGGTGTTTATTGGTGATGAA (synthesized)

- with *in vitro* transcription), store at -20 °C
4. 5-hydroxymethyl dCTP, 100 mM (Zymo Research, catalog number: D1045), store at -20 °C
 5. 5-hydroxymethyl CTP, 100 mM (Trilink Biotechnologies, catalog number: N-1087), store at -20 °C
 6. 5'-hydroxymethylcytosine Carbamoyltransferase, store at -20 °C
Note: Please see Procedure Section A for expression and purification of the enzyme.
 7. Proteinase K (New England Biolabs, catalog number: P8107), store at -20 °C
 8. Thermolabile Proteinase K (New England Biolabs, catalog number: P8111), store at -20 °C
 9. NEBuffer 2.1, 10× (New England Biolabs, catalog number: B7202), store at -20 °C
 10. Ammonium iron (ii) sulfate hexahydrate, 500 µM (MilliporeSigma, catalog number: 203505, see recipe section), store at RT
Note: Iron (ii) solution should be freshly prepared for the experiment
 11. Carbamoylphosphate, 100 mM (MilliporeSigma, catalog number: C4135, see Recipes), store at -20 °C
Note: Carbamoylphosphate solution should be freshly prepared for the experiment.
 12. ATP, 100 mM (New England Biolabs, catalog number: N0450), store at -20 °C
 13. Nuclease-free water (New England Biolabs, catalog number: B1500), store at RT
 14. 1× TE buffer (self-prepared, see Recipes), store at RT
 15. Tris-HCl, 1 M, pH 8.0 (Amresco, catalog number: E298-NEB-2L), store at RT
 16. Tris-HCl, 1 M, pH 7.5 (Amresco, catalog number: E199-NEB-2L), store at RT
 17. NaCl, 5 M (self-prepared), store at RT
 18. Glycerol (self-prepared), store at RT
 19. Ultrapure EDTA, 0.5 M, pH 8.0 (ThermoFisher, catalog number: 15575020), store at RT
 20. Zymo oligo clean & concentrator kit (Zymo Research, catalog number: D4060)
 21. Norgenbiotek oligo clean-up and concentrator kit (Norgenbiotek, catalog number: 34100)
 22. Nucleoside Digestion Mix (New England Biolabs, catalog number: M0649)

Equipment

1. Thermal cycler (Bio-Rad, T100, catalog number: 1861096)
2. Fast protein liquid chromatography (FPLC) system (GE Healthcare, model: AKTA GO)
3. Gel electrophoresis power supply (Thermo, model: EC105)
4. SDS-PAGE gel electrophoresis system (Invitrogen, model: Xcell surelock mini-cell, catalog number: EI0001)
5. Gel photography system (AlphaImager HP)
6. Sonicator for *E. coli* cell disruption (Misonix, model: S-4000)
7. Bacteria incubation shaker (New Brunswick Scientific, model: Innova 44)
8. Agar plate incubator at 37 °C (Napco, model: 310)
9. UV/Visible spectrophotometer (Amersham Biosciences, model: Ultrospec 2100 pro)
10. Centrifuge (Beckman Coulter, model: Avanti J-E)
11. Table-top centrifuge (Eppendorf, model: 5415D)
12. Vortex shaker
13. Water bath or heat block
14. Ice bath
15. Pipette

Procedure

A. Expression and purification of His-carbamoyltransferase protein

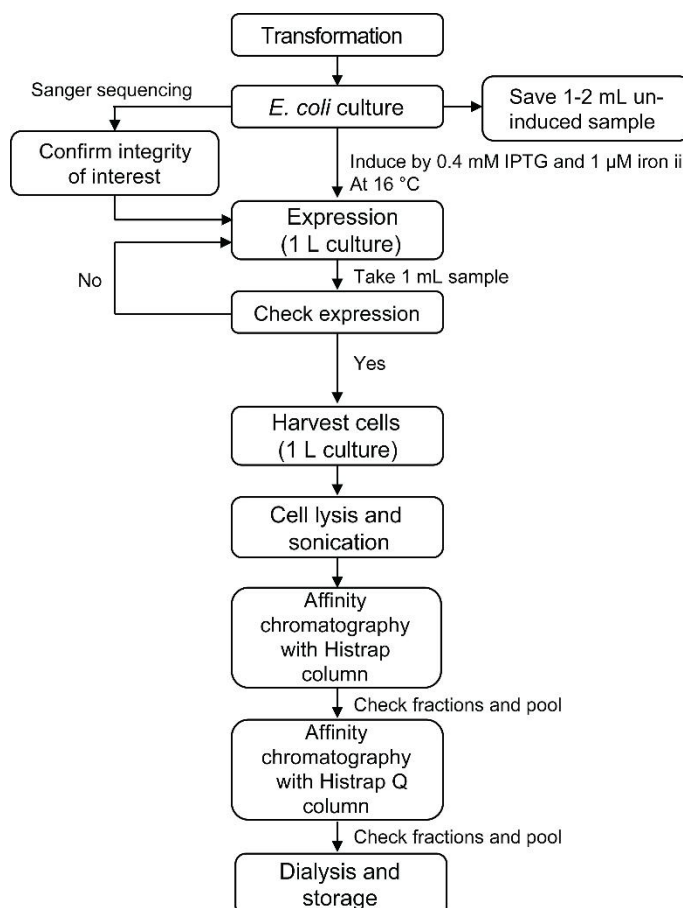


Figure 1. Major steps for expression and purification of protein.

1. Transform 50 μL of competent DH5-alpha cells with 10–50 ng of pET28-Kan-his-CT plasmid (follow the manufacturer protocol) and spread 20–30 μL on an agar LB plate containing 40 $\mu\text{g}/\text{mL}$ kanamycin. Incubate the plate overnight at 37 $^{\circ}\text{C}$.
2. Inoculate 10–50 mL of LB medium containing 40 $\mu\text{g}/\text{mL}$ kanamycin in a bacteria culture tube or shaking flask with a single colony from the plate. Grow overnight at 37 $^{\circ}\text{C}$ while shaking at 250 rpm.
Optional: Verify the integrity of the insert by sanger sequencing.
3. In the next morning, inoculate 1 L of LB medium containing 40 $\mu\text{g}/\text{mL}$ kanamycin in a shaking flask with 10–25 mL overnight culture (roughly 1:40 to 1:100 ratio). Grow the cells at 37 $^{\circ}\text{C}$ while shaking at 250 rpm until $\text{OD}_{600\text{nm}}$ reaches 0.6. Save 1–2 mL cell culture as un-induced control. Pellet the 1–2 mL cells by centrifuging for 5 min at $10,000 \times g$ and freeze in -20°C for short-term storage.
Note: This step may take several hours. Check $\text{OD}_{600\text{nm}}$ every 1–1.5 h in the first few hours and then more frequently. Be aware that the cells grow fast in log phase and the $\text{OD}_{600\text{nm}}$ can increase rapidly when near 0.6.
4. Transfer the cells to a shaker pre-chilled to 16 $^{\circ}\text{C}$ and let it sit for 20 min to cool down the cell culture.
Note: This step is performed to guarantee that carbamoyltransferase protein expression is induced at its optimal temperature ($\sim 16^{\circ}\text{C}$).
5. Add IPTG at 0.4 mM (stock concentration is 0.4 M, 1,000 \times) and 1 μM iron II (stock concentration is 500

μM , 500 \times) to the culture, and continue growth at 16 °C while shaking at 250 rpm for ~24 h. Carbamoyltransferase protein expression is confirmed in step A6, before the entire 1 L volume of culture is processed in step A7.

Note: Iron (ii) is added to facilitate protein folding.

6. Check protein expression with SDS-PAGE gel.
 - a. Harvest 1 mL culture, and pellet cells by centrifuging for 5 min at 10,000 \times g.

Note: The purpose of taking 1 mL from the 1 L culture is for the confirmation of protein expression before the entire 1 L culture is harvested and processed. Note that this 1 mL culture is induced, and the 1 mL control culture taken in step A3 is un-induced.
 - b. Resuspend the pellet and un-induced cell pellets (saved in step A3) in 1 mL cell lysis buffer.
 - c. Lyse the cells in a 1.5 mL microcentrifuge tube on ice with 3–5 cycles of 20 s on/20 s off pulses using sonicator with a microtip. 10 μL is saved as crude lysate. The rest of the volume (1 mL – 10 μL = 990 μL) is used in step A6d.
 - d. Pellet the inclusion bodies by centrifuging the sonicated cell suspension at full speed for 10 min at 4°C, and transfer supernatant to a new 1.5 mL microcentrifuge tube.
 - e. Mix 10 μL crude lysate taken in step A6c, 10 μL nuclease-free water, and 10 μL blue protein loading dye (3 \times with reducing reagent added) to get a final volume of 30 μL .
 - f. Mix 10 μL of soluble samples (supernatant from step A6d), 10 μL of nuclease-free water, and 10 μL of blue protein loading dye (3 \times with reducing reagent added) to get a final volume of 30 μL .
 - g. Heat samples from steps 6Ae and 6Af at 95 °C for 5 min.
 - h. Assemble the SDS-PAGE gel and gel running apparatus, fill with 1 \times running buffer, load the samples and protein ladder, and run gel at 150 V for approximately 1 h.
 - i. Stain the gel with SimplyBlue staining reagent, take picture of the gel, and confirm expression of carbamoyltransferase protein in crude lysate and soluble portion.

Note: The estimated size of carbamoyltransferase enzyme is 63kDa. Presence of protein in soluble lysate should be confirmed before large-scale purification in the next step.
7. Harvest cells from 1 L cell culture from step A5 by centrifuging for 15 min at 5,000 \times g at 4 °C.
8. Cell resuspension and sonication:
 - a. Directly resuspend cell pellet in 30 mL of HisTrap buffer A (or 4 mL buffer per gram of pellet).
 - b. Sonicate the cells on ice with 5 s on/5 s off pulses for 30–40 cycles. Set power to max level.

Note: The number of cycles needed for complete disruption of cells depends on the model of sonicator. It is suggested to do a pre-test to determine the best sonication condition before performing the actual experiment.
 - c. Centrifuge the sonicated cell suspension for 30 min at 23,000 \times g at 4 °C.
 - d. Transfer supernatant (~30 mL) and mix with equal volume of HisTrap buffer A for loading to the AKTA purifier. Add the appropriate volume of 1M PMSF to a final concentration of 0.4 mM to prevent protein degradation.

Note: Save 0.1 mL sample as before purification control. Here, the total loading volume is about 60 mL, and the suggested loading volume range is between 50 and 100 mL.
9. First purification using HisTrap HP 5 mL column on AKTA.
 - a. Equilibrate the column with 10 column volume (CV) of HisTrap buffer A.
 - b. Load sample at 2 mL/min.
 - c. Wash out unbound proteins of the column using 10 \times CV HisTrap buffer A.
 - d. Elute protein using a linear gradient from 0%–100% HisTrap buffer B for a total of 10 \times CV. Set fraction volume to collect 2 mL per fraction tube.
 - e. Check purified proteins in each fraction by adding 5 μL of sample with 15 μL of nuclease-free water and 10 μL of blue protein loading dye (3 \times , reducing agent added) to get a final volume of 30 μL .
 - f. Heat samples at 95 °C for 5 min.
 - g. Load the samples and run SDS-PAGE gel (see **Figure 2** for example of protein purification with HisTrap column).
 - h. Stain the gel with SimplyBlue staining reagent, and take picture of the gel.
 - i. Pool fractions with concentrated proteins.

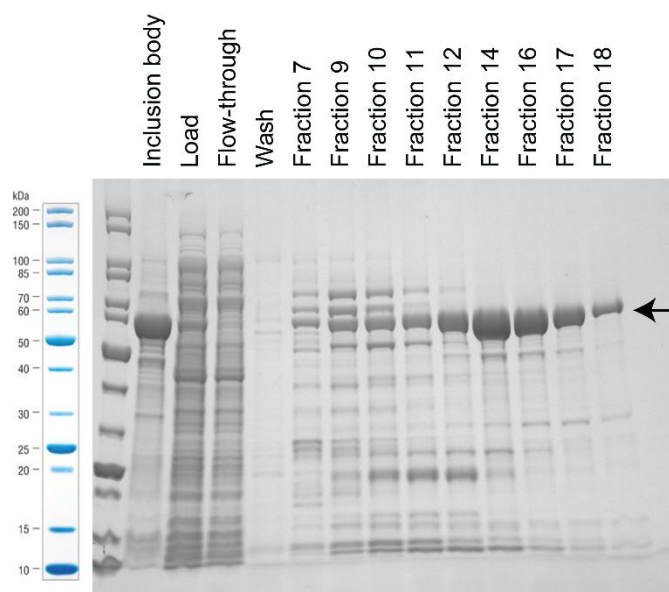


Figure 2. Stained SDS-PAGE gel with samples from the purification with HisTrap column.

Arrow marks band of carbamoyltransferase expression at 63 kDa. For this particular example, we pooled fraction 12–18.

10. Second purification using HiTrap Q column on AKTA.

- a. Mix pooled fractions with 7× volume of dilution buffer so that the final salt concentration is approximately 50 mM.
- b. Equilibrate the column with 10 column volume (CV) of Q buffer A.
- c. Load sample at 3 mL/min.
- d. Wash out unbound proteins of the column using 10× CV Q buffer A.
- e. Elute protein using a linear gradient from 0%–100% Q buffer B for a total of 10× CV. Set fraction volume to collect 2 mL per fraction tube.
- f. Check purified proteins in each fraction by adding 5 µL of sample with 15 µL of nuclease-free water and 10 µL of blue protein loading dye (3×, reducing agent added) to get final volume of 30 µL.
- g. Heat samples at 95 °C for 5 min.
- h. Load the samples and run SDS-PAGE gel (see **Figure 3** for example of protein purification with HiTrap Q column).
- i. Stain the gel with SimplyBlue staining reagent (following manufacturer instructions), and take picture of the gel.
- j. Pool fractions with concentrated proteins.

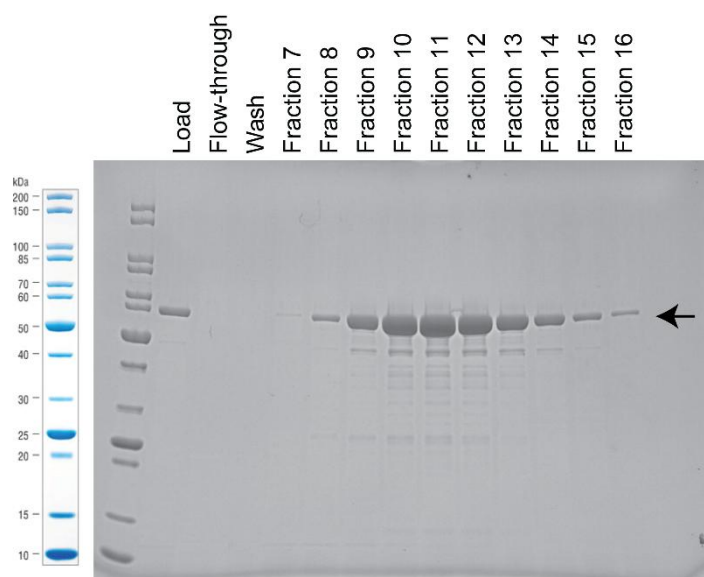


Figure 3. Stained SDS-PAGE gel with samples from the purification with Hitrap Q column.

Arrow marks band of carbamoyltransferase expression at 63 kDa. For this particular example, we pooled fraction 9–13.

11. Transfer purified protein to a dialysis tubing bag (MWCO: 6–8 kD). Dialyze overnight against dialysis buffer at 4 °C.
12. Carefully transfer samples from dialysis tubing bag and determine concentration of purified carbamoyltransferase using the Bradford protein assay kit (following manufacturer instructions).
13. Aliquot purified carbamoyltransferase into 50 µL per microcentrifuge tube and store at -20 °C.

B. dsDNA enzymatic assay

1. DNA denaturation: 5-hydroxymethylcytosine carbamoyltransferase has a preference for ssDNA; therefore, it is critical to denature dsDNA to ssDNA for optimal conversion.
 - a. Prepare 1–2 µg genomic DNA in 5 µL of nuclease-free water or TE buffer in PCR tube
Note: The reaction with T4gt genomic DNA (containing 90–99% 5hmC) is shown as example.
 - b. Heat to denature dsDNA in a thermal cycler set at 95 °C for 10 min.
 - c. Immediately transfer the tube into an ice bath to avoid re-annealing.
Critical: DNA denaturation is a reversible process. When the temperature is cooled down slowly, denatured ssDNA will anneal again. Thus, it is critical to transfer the heated sample and immediately embed the tube completely in the ice bath.
2. Add reaction mix (see **Table 1**) to denatured DNA sample, keep the tube on ice, and mix thoroughly by pipetting up and down for 10 times.

Table 1. Reaction mix for dsDNA modification

| Reagent | Stock Conc. | Final Conc. | 50 μ L Rxn. |
|---------------------------------|-------------|-------------|---------------------------|
| 5hmC denatured DNA | - | 1–2 μ g | 5 μ L |
| NEBuffer 2.1 | 10 \times | 1 \times | 5 μ L |
| Iron II solution | 500 μ M | 10 μ M | 1 μ L |
| Carbamoylphosphate ² | 100 mM | 10 mM | 5 μ L |
| ATP | 100 mM | 5 mM | 2.5 μ L |
| Enzyme | - | 1–7 μ M | See footnote ¹ |
| Nuclease-free water | - | - | Bring up to 50 μ L |
| Total | - | - | 50 μ L |

¹ The volume of enzyme added to the reaction should not exceed 10% of the total volume.

² Carbamoylphosphate has a short half-life at room temperature. Therefore, it is recommended to always keep the solution on ice and to add the enzyme right before the incubation step.

- Incubate the reaction in a thermal cycler set to 30 °C for 3 h.
- Inactivate the reaction by adding 2 μ L of Proteinase K, mix thoroughly, and continue incubating at 37 °C for 30 min.
- Follow the Zymo oligo clean & concentrator kit manual to purify DNA product, resuspend, and recover DNA in 20 μ L TE buffer or Elution buffer provided in the kit.
- After purification, the conversion rate can be measured by quantitative liquid chromatography (see Data analysis).
- Use the purified reactions for downstream analysis, or otherwise store at -20 °C.

C. ssDNA oligo and RNA oligo enzymatic assay

- Prepare 5hmC ssDNA or RNA oligo: no denaturation is needed for single-stranded oligos; just prepare 1 μ g of oligos in suitable volume between 1–5 μ L with nuclease-free water or TE buffer.
Note: Examples of ssDNA or RNA oligo sequences are provided in the Materials and Methods section.
- Add reaction mix (see **Table 2**) to oligo sample, keep the tube on ice, and mix thoroughly by pipetting up and down 10 times.

Table 2. Reaction mix for ssDNA or RNA modification

| Reagent | Stock Conc. | Final Conc. | 50 μ L Rxn. |
|---------------------------------|-------------|-------------|---------------------------|
| ssDNA/RNA oligo | - | 1 μ g | 1–5 μ L |
| NEBuffer 2.1 | 10 \times | 1 \times | 5 μ L |
| Iron II solution | 500 μ M | 10 μ M | 1 μ L |
| Carbamoylphosphate ² | 100 mM | 10 mM | 5 μ L |
| ATP | 100 mM | 5 mM | 2.5 μ L |
| Enzyme | - | 1–7 μ M | See footnote ¹ |
| Nuclease-free water | - | - | Bring up to 50 μ L |
| Total | - | - | 50 μ L |

¹ The volume of enzyme added to the reaction should not exceed 10% of the total reaction volume.

² Carbamoylphosphate has a short half-life at room temperature. Therefore, it is recommended to always keep the solution on ice and to add the enzyme right before the incubation step.

3. Incubate the reaction in a thermal cycler set to 30 °C for 3 h.
4. Inactivate the reaction by adding 2 µL of Proteinase K, mix thoroughly, and continue incubating at 37 °C for 30 min.
5. For purification of short oligos ≤21 nt, we generally see better recovery with Norgenbiotek Oligo clean-up and concentrator kit. Otherwise, follow the Zymo oligo clean & concentrator kit manual. Resuspend and recover converted DNA/RNA oligo in 20 µL of TE buffer or Elution buffer provided in the kit.
6. After purification, the conversion rate can be measured by quantitative liquid chromatography (see Data analysis).
7. Use the purified reactions for downstream analysis, or otherwise store at -20 °C.

D. Single nucleotide/deoxynucleotide enzymatic assay

1. Prepare 1:10 dilution of dhmCTP or hmCTP (dilute to 10 mM) with nuclease-free water.
2. Add reaction mix (see **Table 3**) to sample, keep the tube on ice, and mix thoroughly by pipetting up and down 10 times.

Table 3. Reaction mix for dhmCTP or hmCTP modification

| Reagent | Stock Conc. | Final Conc. | 50 µL Rxn. |
|---------------------------------|-------------|-------------|---------------------------|
| dhmCTP or hmCTP | 10 mM | 0.2 mM | 1 µL |
| NEBuffer 2.1 | 10× | 1× | 5 µL |
| Iron II solution | 500 µM | 10 µM | 1 µL |
| Carbamoylphosphate ² | 100 mM | 10 mM | 5 µL |
| ATP | 100 mM | 5 mM | 2.5 µL |
| Enzyme | - | 1–7 µM | See footnote ¹ |
| Nuclease-free water | - | - | Bring up to 50 µL |
| Total | - | - | 50 µL |

¹ The volume of enzyme added to the reaction should not exceed 10% of total reaction volume.

² Carbamoylphosphate has a short half-life at room temperature. Therefore, it is recommended to always keep the solution on ice and to add the enzyme right before the incubation step.

3. Incubate the reaction in a thermal cycler set to 30 °C for 3 h.
4. Inactivate the reaction by adding 2 µL of Proteinase K, mix thoroughly, and continue incubating at 37 °C for 30 min. Because no purification will be performed to clean up Proteinase K, if it remains a concern for downstream analysis, use thermolabile proteinase K instead and perform additional heat inactivation at 55 °C for 10 min.
5. After the reaction, the conversion rate can be measured by quantitative liquid chromatography (see Data analysis).
6. Use the reactions for downstream analysis, or otherwise store at -20 °C.

Data analysis

Quantification of the enzymatic reaction is done by measuring the conversion rate of 5dhmC/5hmC to 5dcmC/5cmC using liquid chromatography. To prepare nucleosides for quantitative analysis, we use Nucleoside Digestion Mix (New England Biolabs, catalog number: M0649) and follow the manual for incubation. No additional purification is required for LC-MS analysis. LC-MS analysis was performed on an Agilent 1200 Series LC/MS system equipped with a G1315D diode array detector and a 6120 Single Quadrupole Mass Detector in both positive (+ESI) and negative (-ESI) electrospray ionization modes. LC was performed on a Waters Atlantis T3 column (4.6×150 mm, $3 \mu\text{m}$) with a gradient mobile phase consisting of aqueous ammonium acetate (10 mM, pH 4.5) and methanol. The identity of each peak was confirmed by MS. The relative abundance of each nucleoside was determined by the integration of each peak at 260 nm or its respective UV absorption maxima. Since the extinction coefficient constant for 5dcmC or 5cmC was not available, the relative quantification was performed by measuring the decrease of 5dhmC or 5hmC. The total abundance of 5dhmC/5hmC and converted 5dcmC/5cmC was normalized to a no enzyme treatment control.

Figure 4 shows an example of nucleoside peaks with a reaction performed on 5dhmC ssDNA (marked in red, sequence provided in Materials and Methods). A no enzyme reaction control is also included (marked in blue). Note the complete disappearance of 5dhmC in the red sample, which suggested almost complete conversion of 5dhmC to 5dcmC. Three replicate reactions were performed for statistical analysis.

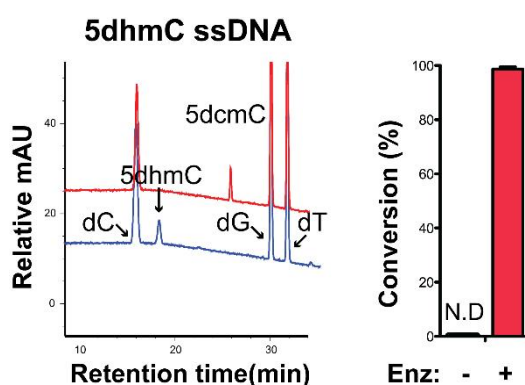


Figure 4. Quantification of carbamoyltransferase conversion rate in ssDNA oligo containing an internal 5dhmC.

Notes

1. The optimal concentration of enzyme for the assay depends on the active units of purified enzyme, and thus can vary from batch to batch. We recommend conducting enzyme serial dilutions to first determine the optimal molarity concentration for the reaction. For calculation of enzyme molarity, use the formula: $\text{Molarity} = \text{Moles of enzyme} / \text{volume of reaction} = \text{Mass} / (\text{Molecular weight} * \text{volume of reaction})$. The molecular weight of carbamoyltransferase is 63,000 g/mol.
2. For dsDNA enzymatic assay, we normally see up to 70% conversion using the T4gt genomic DNA; the conversion of 5dhmC to 5dcmC on ssDNA (sequence provided in Materials and Reagents) can reach nearly 100% (example data shown in Data analysis; we suggest using purification batch with ssDNA conversion rate no less than 95%); the conversion for single nucleotide/deoxynucleotide is expected to be between 60%–90%.

Recipes

1. LB media

- Weigh 10 g of tryptone, 5 g of yeast extract, and 10 g of sodium chloride.
- Dissolve in 1 L H₂O.
- Sterilize by autoclaving.

2. Iron II solution (prepare fresh, this recipe makes 50 mL)

- Weigh 10 mg of ammonium iron (ii) sulfate hexahydrate powder.
- Dissolve in 5 mL of nuclease-free water to get 5 mM.
- Dilute 1:10 with nuclease-free water to get stock concentration at 500 µM.

3. Carbamoylphosphate solution (prepare fresh)

- Weigh 18 mg of carbamoyl phosphate powder.
- Dissolve in 1 mL of nuclease-free water to get stock concentration at 100 mM.

4. TE buffer

| Reagent | Final concentration | Amount |
|------------------------|---------------------|---------|
| EDTA (0.5 M, pH 8.0) | 1 mM | 0.2 mL |
| Tris-HCl (1 M, pH 8.0) | 10 mM | 1 mL |
| Nuclease-free water | n/a | 98.8 mL |
| Total | n/a | 100 mL |

5. Cell lysis buffer

| Reagent | Final concentration | Amount |
|------------------------|---------------------|--------|
| NaCl (5 M) | 250 mM | 5 mL |
| Tris-HCl (1 M, pH 7.5) | 20 mM | 2 mL |
| Nuclease-free water | n/a | 93 mL |
| Total | n/a | 100 mL |

6. HisTrap Buffer A

| Reagent | Final concentration | Amount |
|---------------------------|---------------------|--------|
| NaCl (5 M) | 250 mM | 25 mL |
| Tris-HCl (1 M, pH 7.5) | 20 mM | 10 mL |
| Imidazole (2.5 M, pH 8.0) | 20 mM | 4 mL |
| Nuclease-free water | n/a | 461 mL |
| Total | n/a | 500 mL |

7. HisTrap Buffer B

| Reagent | Final concentration | Amount |
|---------------------------|---------------------|--------|
| NaCl (5 M) | 500 mM | 30 mL |
| Tris-HCl (1 M, pH 7.5) | 20 mM | 6 mL |
| Imidazole (2.5 M, pH 8.0) | 750 mM | 90 mL |
| Nuclease-free water | n/a | 174 mL |
| Total | n/a | 300 mL |

8. Dilution buffer

| Reagent | Final concentration | Amount |
|------------------------|---------------------|--------|
| Tris-HCl (1 M, pH 7.5) | 20 mM | 6 mL |
| Glycerol | 5% | 15 mL |
| Nuclease-free water | n/a | 279 mL |
| Total | n/a | 300 mL |

9. HiTrap Q Buffer A

| Reagent | Final concentration | Amount |
|------------------------|---------------------|--------|
| NaCl (5 M) | 50 mM | 5 mL |
| Tris-HCl (1 M, pH 7.5) | 20 mM | 10 mL |
| Glycerol | 5% | 25 mL |
| Nuclease-free water | n/a | 460 mL |
| Total | n/a | 500 mL |

10. HiTrap Q Buffer B

| Reagent | Final concentration | Amount |
|------------------------|---------------------|--------|
| NaCl (5 M) | 1.5 M | 90 mL |
| Tris-HCl (1 M, pH 7.5) | 20 mM | 6 mL |
| Glycerol | 5% | 15 mL |
| Nuclease-free water | n/a | 189 mL |
| Total | n/a | 300 mL |

11. Dialysis buffer (same as protein storage buffer)

| Reagent | Final concentration | Amount |
|------------------------|---------------------|----------|
| Glycerol | 50% | 500 mL |
| Tris-HCl (1 M, pH 7.5) | 20 mM | 20 mL |
| NaCl (5 M) | 100 mM | 20 mL |
| DTT (1 M) | 1 mM | 1 mL |
| Nuclease-free water | n/a | 459 mL |
| Total | n/a | 1,000 mL |

Acknowledgments

This work is funded by New England Biolabs Inc.

Competing interests

All authors are employees of New England Biolabs Inc., a manufacturer and vendor of molecular biology reagents, including several enzymes and buffers used in this work.

References

- Ferrer, M., Martinez-Abarca, F. and Golyshin, P. N. (2005). [Mining genomes and 'metagenomes' for novel catalysts](#). *Curr Opin Biotechnol* 16, 588-593.
- Flodman, K., Correa Jr, I. R., Dai, N., Weigele, P. and Xu, S. Y. (2020). [In vitro Type II restriction of bacteriophage DNA with modified pyrimidines](#). *Front Microbiol* 11: 604618.
- Lehman, I. R. and Pratt, E. A. (1960). [On the structure of the glucosylated hydroxymethylcytosine nucleotides of coliphages T2, T4, and T6](#). *J Biol Chem* 235: 3254-3259.
- Yang, W., Lin, Y. C., Johnson, W., Dai, N., Vaisvila, R., Weigele, P., Lee, Y. J., Correa Jr, I. R., Schildkraut, I. and Ettwiller, L. (2021). [A genome-phenome association study in native microbiomes identifies a mechanism for cytosine modification in DNA and RNA](#). *Elife* 10: 70021.