BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

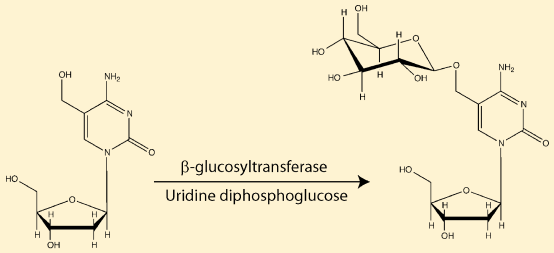
pubs.acs.org/biochemistry

Biochemical Characterization of Recombinant β-Glucosyltransferase and Analysis of Global 5-Hydroxymethylcytosine in Unique Genomes

Jolyon Terragni, Jurate Bitinaite, Yu Zheng, and Sriharsa Pradhan*

New England Biolabs, Inc., 240 County Road, Ipswich, Massachusetts 01938, United States

ABSTRACT: 5-Hydroxymethylcytosine (5-hmC) is an enzymatic oxidative product of 5-methylcytosine (5-mC). The Ten Eleven Translocation (TET) family of enzymes catalyze the conversion of 5-mC to 5-hmC. Phage-encoded glucosyltransferases are known to glucosylate 5-hmC, which can be utilized to detect and analyze the 5-hmC as an epigenetic mark in the mammalian epigenome. Here we have performed a detailed biochemical characterization and steady-state kinetic parameter analysis of T4 phage β-glucosyltransferase (β-GT). Recombinant β-GT glucosylates 5-hmC DNA in a nonprocessive manner, and binding to either 5-hmC DNA or uridine diphosphoglucose (UDP-glucose) substrates is random, with both binary complexes being catalytically competent. Product inhibition studies with β-GT demonstrated that UDP is a competitive inhibitor with respect to UDP-glucose and a mixed inhibitor with respect to 5-hmC DNA. Similarly, the glucosylated-5-hmC (5-ghmC) DNA is a competitive inhibitor with respect to 5-hmC DNA and mixed inhibitor with respect to UDP-glucose. 5-hmC DNA binds ∼10 fold stronger to the β-GT enzyme when compared to its glucosylated product. The numbers of 5-hmC on target sequences influenced the turnover numbers for recombinant β-GT. Furthermore, we have utilized recombinant β-GT to estimate global 5-hmC content in a variety of genomic DNAs. Most of the genomic DNAs derived from vertebrate tissue and cell lines contained 5-hmC. DNA from mouse, human, and bovine brains displayed 0.5−0.9% of the total nucleotides as 5-hmC, which was higher compared to the levels found in other tissues. A comparison between cancer and healthy tissue genomes suggested a lower percentage of 5-hmC in cancer, which may reflect the global hypomethylation of 5-mC observed during oncogenesis.



5-Hydroxymethylcytosine (5-hmC) Glucosyl-5-hydroxymethylcytosine (5-ghmC)

Addition of sugar residues by glycosidic bond formation in living organisms is facilitated by large numbers of glycosyltransferases. The glycosyltransferases utilize unique cosubstrates, which are generally a sugar donor along with an acceptor molecule with specific linkage specificity. One such example of a glycosyltransferase is β-glucosyltransferase of the *Escherichia coli* T4 bacteriophage. The β-GT transfers a glucose residue from the cosubstrate UDP-glucose to the 5-hmC base, as found in the T4 double-stranded DNA genome, converting it to β-glucosyl-5-hydroxymethylcytosine. The role of glucosylation for T4 phage survival upon infection of a host *E. coli* cell is well-documented.[1,2] Furthermore, glucosylation is also implicated in phage specific gene expression by influencing transcription.[3−5]

Although 5-hmC is predominantly found in lower organisms, such as T4 bacteriophage, it is also present in mammalian cells. Earlier documentation of the tissue specific presence of 5-hmC[6] and the recent rediscovery of its presence in embryonic stem cells (ESC)[7] and brain DNAs, especially in Purkinje neurons,[8] have generated interest. Although the biochemical pathways of the oxidation of 5-methylcytosine (5-mC) to 5-hmC are not completely understood, there is clear evidence of a family of enzymes that mediate these reactions. The TET family of enzymes (Tet1−Tet3) specifically bind 5-mC and catalyze its conversion to 5-hmC.[7] Human Tet2 mutations are associated with myeloid malignancies, perhaps because of compromises in their catalytic activity, resulting in impaired oxidation of 5-mC.[9]

The role of 5-hmC in biology is not yet clear, but 5-hmC has been postulated to influence chromatin structure, recruit specific factors, or constitute an intermediate component in cytosine demethylation. Indeed, oxidation of 5-mC by Tet1 promotes active DNA demethylation in the adult mammalian brain. This process is aided by AID/APOBEC deaminase followed by base excision repair processes.[10] Tet1 and Tet2 are Oct4-regulated enzymes that together sustain 5-hmC levels in mouse ESCs and are induced concomitantly with 5-hmC during reprogramming of fibroblasts to induced pluripotent stem cells.[11] In mouse ES cells, Tet1 is preferentially bound to CpG-rich sequences at promoters of both transcriptionally active and polycomb-repressed genes. Thus, Tet1 in mouse ESC may have a dual function in gene activation and repression.[12] In human ESC, 5-hmC is present in gene bodies and enhancer elements, and this presence correlates with gene expression.[13] Furthermore, recent studies have demonstrated that reprogramming of the paternal genome upon fertilization in mammals involves genome-wide oxidation of 5-mC to 5-hmC and that the balance between these two DNA modifications is

Received: September 20, 2011
Revised: December 16, 2011
Published: January 9, 2012

 © 2012 American Chemical Society

inextricably linked with pluripotency and lineage commitment in mouse ESC.[14−16]

Because 5-hmC has been implicated in stem cell biology and cancer, several antibody-based immunoprecipitation methods have been developed for genome-wide mapping and/or quantification. In one of the approaches, termed GLIB (glucosylation, periodate oxidation, biotinylation), a combination of β-GT and chemical steps is employed to isolate DNA fragments containing as few as one 5-hmC.[17] The second approach involves conversion of 5-hmC to cytosine 5-methylenesulfonate (CMS) by treatment of genomic DNA with sodium bisulfite, followed by immunoprecipitation of CMS-containing DNA with a specific antiserum to CMS.[9] All these processes require high-throughput sequencing to determine the 5-hmC mapping on genes. Although antibody-based global 5-hmC level quantification is possible, the linearity of the signal decreases with higher DNA concentrations. Similarly, other global 5-hmC level quantifications can be accomplished by high-performance liquid chromatography (HPLC)−mass spectroscopy methods. Recently, another study has used β-GT to measure global 5-hmC determination utilizing radioactive UDP-[^{3}H]glucose,[18] thereby establishing β-GT as an important enzyme for measuring global 5-hmC levels. Although high-resolution crystal structures and the effects of substrate and metal binding for β-GT were studied a decade ago, there is an apparent lack of information about the basic biochemical parameters of this enzyme. In this work, we have studied the biochemical characterization and reaction mechanism of purified recombinant β-GT (EC 2.4.1.27) from *E. coli*. This enzyme belongs to glycosyltransferase family 63 (http://www.cazy.org/). Using the optimal reaction conditions, we have further utilized a radioactive labeling assay with β-GT to quantitatively measure 5-hmC levels across many different tissue specific DNA samples, including several novel sample types such as *Arabidopsis* and bovine.

■ EXPERIMENTAL PROCEDURES

Recombinant β-GT and MfeI Restriction Digest Protection Assay. Recombinant β-GT was expressed in *E. coli*, and the purified enzyme was obtained from New England Biolabs, Inc. (NEB, catalog no. M0357S). For the MfeI restriction digest protection assay, the glucosylation reactions were conducted in 70 μL of NEB buffer 4 containing 0.188 nM T4-*gt* DNA (the total concentration of 5-hmC residues was ∼11.2 μM), 50 μM UDP-glucose, and 30 nM β-GT. Aliquots (10 μL) were withdrawn from the reaction mixture after incubation for 0, 5, 10, 20, 30, 45, or 60 min at 37 °C. The glucosylation reactions were stopped by heating each sample for 20 min at 70 °C. After heat inactivation of β-GT, 1 μL (10 units) of MfeI restriction endonuclease (NEB) was added to each sample and each mixture incubated at 37 °C for 1 h to cleave nonglucosylated T4-*gt* DNA. The restriction reaction was quenched by adding 0.3 volume of gel loading buffer [60 mM EDTA (pH 8.0), 50% glycerol, 0.2% SDS, and 0.02% bromophenol blue], and the products were separated by electrophoresis in a 1% agarose gel. The ethidium bromide-stained gel was visualized under UV light.

Glucosylation Assay for 5-hmC DNA. UDP-[1-^{3}H]-glucose (UDP-[^{3}H]glucose, American Radiolabeled Chemical, Inc., catalog no. ART 0525) was diluted with cold UDP-glucose (NEB) to form a 0.225 mM stock solution. A standard glucosylation assay for 5-hmC quantification consisted of a fixed concentration of UDP-[^{3}H]glucose and a known quantity of purified mutant T4-*gt* DNA (NEB), where all cytosine residues are modified 5-hmC (mutations in both α and β-GT), and they were mixed with different concentrations of recombinant β-GT in 1× NEB buffer 4 [50 mM potassium acetate, 20 mM Tris acetate, 10 mM magnesium acetate, and 1 mM DTT (pH 7.9)] at 25 °C. Reaction mixtures were incubated at 25 °C for various time intervals, and 25 μL of each reaction mixture was mixed with 5 μL of 400 μM cold UDP-glucose and flash-frozen on ethanol and dry ice for processing. The reaction mixtures were thawed and immediately applied to a 2.5 cm DE81 membrane (GE Healthcare, catalog no. 3658-325) under air pressure using a vacuum manifold (Millipore). The applied reaction mixture was washed in 3 × 1 mL of 0.2 M ammonium bicarbonate, 3 × 1 mL of water, and 3 × 1 mL of ethanol. Membranes were air-dried and placed in scintillation vials. To the dried filter was added 3 mL of scintillation fluid; the solution was mixed, and tritium incorporation was measured for 1 min. All glucosylation reaction values were corrected for nonspecific binding of UDP-[^{3}H]glucose to the processed filters. Background values were determined in the absence of enzyme or substrate but in the presence of UDP-[^{3}H]glucose in the reaction mixture. The counting efficiency of incorporation of [^{3}H]glucose was determined to be ∼50% by using internal standards. Therefore, during all our calculations, a correction value of 2 was included. Data were plotted by either linear or nonlinear regression analysis using GraphPad PRISM 4 (GraphPad Software, Inc.).

DNA Substrate for the 5-hmC Assay. T4-*gt* DNA was obtained from NEB. Duplex DNA substrates containing 2, 6, 12, or 24 residues of 5-hmC were made by polymerase chain reaction (PCR) in the presence of d(5-hm)CTP nucleoside in place of 5-dCTP using Phusion high-fidelity DNA polymerase (NEB, catalog no. M0530S). The sequences of the synthesized duplex DNAs are as follows:[13] 2 × 5-hmC-DNA, 5′ TACTCTATACTCTACTCATCATTACAATATATATATAATTAATTATAATTAA**CG**AAATTATAATTTATAATTAATTAATATGAGATATGAGATGTGTATG 3′; 6 × 5-hmC-DNA, 5′ TACTCTATACTCTACTCATCATTACAATATATATATAATTAATTATAATT**CGCG**AAATTA**CG**ATTTATAATTAATTAATATGAGATATGAGATGTGTATG 3′; 12 × 5-hmC DNA, 5′ TACTCTATACTCTACTCATCATTACAATATATATATAT**CG**TTAA**CG**ATAATT**CGCGCG**ATTA**CG**ATTTATAATTAATTAATATGAGATATGAGATGTGTATG 3′; 24 × 5-hmC DNA, 5′ TACTCTATACTCTACTCATCATTACA**CGCGCG**ATAT**CG**TTAA**CG**ATAATT**CGCGCG**ATTA**CG**AT**CG**ATAA**CGCG**TTAATATGAGATATGAGATGTGTATG 3′. PCR products were biotinylated by virtue of using 5′-biotinylated PCR primers in a standard PCR with Phusion high-fidelity DNA polymerase (NEB). All hydroxymethyl CpG dinucleotides are shown in bold.

The double-stranded hemi-5-hydroxymethylated and symmetrical 5-hydroxymethylated synthetic oligonucleotides, synthesized at NEB, were utilized as 5-hmC substrates. Oligonucleotide pairs were annealed in 10 mM Tris (pH 8.0), 50 mM NaCl, and 1 mM EDTA. The sequences for the double-stranded synthetic oligonucleotides were as follows, with the hydroxymethylated cytosines shown in bold type: hemi-5-hmC (**A**), 5′ CGGCGTTTCCGGGTTCCATAGGCTCCGCC**C**GGACTCTGATGACCAGGGCATCACA 3′ and 3′ GCCGCAAAGGCCCAAGGTATCCGAGGCGGGCCTGAGACTACTGGTCCCGTAGTGT 5′; hemi-5-hmC (**B**), 5′ CGGCGTTTCCGGGTTCCATAGGCTCCGCCCGGACTCTGATGACCAGGGCATCACA 3′ and 3′ GCCGCAAAGGCCCAAGGTATCCGAGGCGGG**C**CTGAGACTACTGGTCCCGTAGTGT 5′; symmetrical 5-hmC (**C**),

5′ CGGCGTTTCCGGGTTCCATAGGCTCCGCCCG-GACTCTGATGACCAGGGCATCACA 3′ and 3′ GCCGCA-AAGGCCCAAGGTATCCGAGGCGGGCCTGAGAC-TACTGGTCCCGTAGTGT 5′.

Processivity Studies with Recombinant β-GT. For processivity studies, a master mix (300 μL) of 1.35 μM biotinylated DNA that contained 24 5-hmC residues (24 × 5-hmC), 50 μM UDP-[^{3}H]glucose, and 0.01 μM recombinant β-GT was incubated on ice for 3 min. At that point, the reaction mixture was brought to room temperature and divided into two samples. A 12-fold excess of nonbiotinylated 24 × 5-hmC was added to one of the samples and the sample mixed. To the other sample was added an equal volume of water. After incubation for 1 min at 25 °C, 25 μL of reaction mixture was removed from both samples, at intervals of 30 s, and mixed with 5 μL of 400 μM cold UDP-glucose and flash-frozen on ethanol and dry ice for processing.

Initial Velocity and Product Inhibition Studies with Recombinant β-GT. All initial velocity studies were conducted under identical conditions as described in Glucosylation Assay for 5-hmC DNA. For initial velocity studies with 5-hmC containing substrate DNA, the concentration of 5-hmC was varied (0.03−1.0 μM) while the UDP-glucose and β-GT concentrations were kept at 2.5 and 0.01 μM, respectively, in the incubation mix. Similarly, for initial velocity studies with the UDP-glucose substrate, the concentration of UDP-glucose was varied (2.5−50.0 μM), while the 5-hmC DNA and β-GT concentrations were kept at 0.025 and 0.01 μM, respectively, in the incubation mix. The amount of product formed was measured in counts per minute. The concentration of product formed ([^{3}H]glucose transfer, N) was calculated using the formula $N = (Y - B)/(EF)$, where Y is the measured counts per minute from the experiment, B is the counts per minute blank, E is the counting efficiency (0.5), and F is the counts per minute per micromolar free UDP-[^{3}H]glucose. The amount of product formed per minute is measured as N/T, where T is the time of the reaction. The initial rate of product formation versus substrate concentration was fit to a one-site binding model (hyperbola). The velocity equation in the absence of product is as follows.

$$v = (V_{max}[S])/(K_m + [S]) \quad (1)$$

where [S] is [UDP-glucose][5-hmC]. This describes binding of a substrate or ligand to an enzyme that follows the law of mass action. V_{max} is the maximal velocity, and K_m is the concentration of substrate needed to reach the half-maximal velocity. All data points were analyzed by nonlinear regression to yield V_{max} and K_m. From the V_{max} value, the turnover number (k_{cat}) was deduced from the known enzyme concentration, where Et is the enzyme concentration.

$$k_{cat} = V_{max}/[Et]$$

Initial velocity studies in the presence of either of the products were performed to deduce the reaction mechanism of recombinant β-GT. The enzyme has two products in the reaction, UDP and 5-ghmC. The UDP (Sigma, catalog no. U4125) solution was made in Milli-Q water at 100 μM. For glucosylated DNA, wild-type T4 DNA (T4-*wt*) (NEB) was used. The difference between T4-*gt* and T4-*wt* was the fact that all the cytosine residues in T4-*gt* are 5-hydroxymethylcytosine. To perform the product inhibition reactions, a series of reactions were performed with fixed amounts of end product. In the experiment with UDP as the end product, the concentrations of reactants were as follows: 2.5 μM 5-hmC, 2−40 μM UDP, and 2.5−50 μM UDP-[^{3}H]glucose or 0.03−1.0 μM 5-hmC, 25 μM UDP, and 2.5−50 μM UDP-[^{3}H]glucose. Similarly, in the experiment with 5-ghmC as the end product, the concentrations of reactants were 0.125−7.2 μM for 5-hmC and 2.5−50 μM for UDP-[^{3}H]glucose. The data obtained from inhibition studies were fit to equations stipulated for competitive or mixed inhibition as described previously.[19]

Global 5-hmC Quantification in Mammalian Genomes by a Glucosylation Assay. Genomic DNA samples from E14, COS-7, control ES (E14), Tet1 KD ES (E14 *Tet1*−/−), U387, NIH-3T3, HEK293, HCT 116 WT, HCT 116 (*DNMT1*−/−), E14 day 1, E14 day 3, E14 day 5, E14 day 7, HeLa, and IMR 90 cells were purified with Qiagen Puregene Core Kit A (catalog no. 1042601). Genomic DNA samples were treated with Qiagen RNase A (catalog no. 19101) to remove RNA contamination. Genomic DNA tissue samples from mouse brain, liver, kidney, heart, and lung were purchased from BioChain. These genomic DNA samples were treated with Agilent Technologies RNace (RNase A and RNase T1, catalog no. 400720) to ensure samples were RNA free. For the β-GT glucosylation assay, 1 μg of purified DNA was incubated with 25 μM UDP-[^{3}H]glucose (125 μM stock) and 5 units of β-GT enzyme in a total volume of 25 μL. The β-GT and reaction mixtures were incubated at 37 °C for 2 h. The reaction mixtures were then applied to DE81 membranes and processed as described above.

The femtomole readings were converted to % 5-hmC and % modified cytosines in the genome with the following equation:

$$\text{total \% 5-hmC in genome} = (\text{moles of 5-hmC measured}) / [(\text{grams of DNA measured})/(\text{MW of DNA base})]$$

RESULTS

Preliminary Characterization of Recombinant β-GT. Commercially available recombinant β-GT was homogeneous with a single band on a Coomassie-stained sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) gel (Figure 1A). The enzyme was tested for glucosylation activity on 5-hmC containing T4 phage DNA (T4-*gt* DNA). Initially, we developed a restriction enzyme-based assay method in which T4-*gt* DNA was resistant to cleavage by MfeI after the glucosylation reaction. Judging by the resistance to cleavage by MfeI as a function of recombinant β-GT-mediated catalysis over time, we determined that the 30 nM enzyme is highly efficient in glucosylating 0.188 nM T4-*gt* DNA in 10 min of the reaction (Figure 1B). This result suggested the linear phase of a β-GT reaction to be less than 10 min.

Linearity of the Glucosylation Reaction by Recombinant β-GT Using UDP-[^{3}H]Glucose. To determine the linear reaction phase of recombinant β-GT, we added a fixed concentration of β-GT and a fixed concentration of DNA with four different concentrations of cosubstrate UDP-glucose. From the reaction mixture, 25 μL was withdrawn at 1 min time intervals for quantitative estimation of glucosylation. At all four different concentrations of UDP-glucose (5, 10, 25, and 50 μM), the reaction essentially remained linear up to 5 min (Figure 2A). We also conducted another set of experiments by

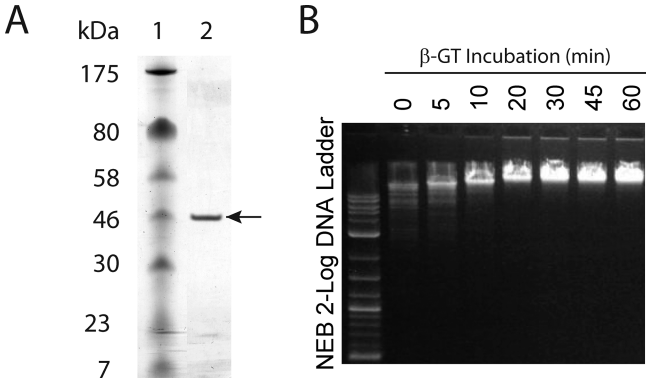


Figure 1. Purity and initial characterization of the β-GT enzyme. (A) Coomassie blue-stained SDS–PAGE gel showing 8 μg (80 units) of recombinant β-GT enzyme. (B) Glucosylation of 5-hmC DNA (T4 phage *gt* −/− DNA) with recombinant β-GT protects it from cleavage by MfeI. A time course of the glucosylation reaction followed by MfeI digestion is shown. The arrow indicates T4 phage *gt* −/− DNA.

varying the recombinant β-GT concentration and keeping the DNA concentrations fixed with two different UDP-glucose concentrations (5 and 25 μM). Indeed, the glucosylation reaction remained virtually linear with enzyme concentration between 0.0025 and 0.02 μM (Figure 2B). From these data, we chose an optimal enzyme concentration to be 0.01 μM in our reactions and the optimal reaction time to be 2 min at 25 °C.

β-GT Glucosylates 5-hmC in a Distributive Manner. 5-Hydroxymethylcytosine is capable of accepting one glucose group from UDP-glucose by β-GT. This reaction could happen in a processive or distributive manner. To determine the reaction process, we incubated the biotin-tagged 5-hmC containing DNA in the presence of a limiting amount of β-GT and excess UDP-[^{3}H]glucose. We preincubated this reaction on ice for 3 min to promote ternary complex formation. This reaction mix was split into two equal portions, and to one half was added 12-fold competitor DNA containing 5-hmC but without a biotin tag. After incubation for 1 min at 25 °C, duplicate aliquots were collected every 30 s and quenched. The biotinylated substrates were captured with streptavidin beads, and radioactive incorporation was measured. In the event of a processive mechanism, the enzyme would remain bound to the biotinylated substrate and would not be released to the competitor until all the available 5-hmCs are glycosylated. In a distributive mechanism, the enzyme is expected to fall off the substrate after one catalytic reaction and then use the available pool of substrates for the next reaction. Because the enzyme-loaded substrate is biotinylated, the amount of products in the presence and absence of competitor DNA could be captured by streptavidin magnetic beads. Should the enzyme be released from the biotinylated DNA to competitor DNA, the biotinylated DNA will not be glucosylated. Indeed, in the presence of competitor DNA, β-GT failed to glucosylate the preloaded substrate DNA (Figure 3),

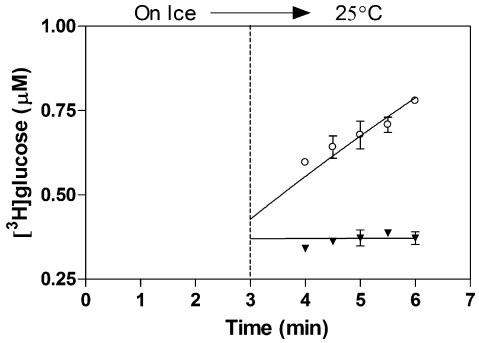


Figure 3. Nonprocessive glucosylation catalyzed by recombinant β-GT. Incorporation of UDP-[^{3}H]glucose after the enzyme was preincubated on ice with 50 μM UDP-[^{3}H]glucose and 1.35 μM 5-hmC DNA. Three minutes after the start of the reaction the mixture was divided into two equal portions, one chased with nonbiotinylated 5-hmC DNA (▼) and the other with an equal volume of water (○) as described in Experimental Procedures. After the chase, both reaction mixtures were incubated at (25 °C) for 1 min and then monitored at 30 s intervals by processing 25 μL of the reaction mixture in duplicate. Measurements were obtained from streptavidin magnetic beads with captured glucosylated [^{3}H]-5-hmC.

as the reaction curve remained parallel to the *x*-axis. The control reaction showed a linear reaction progression. Therefore, β-GT appears to follow a distributive mechanism of reaction.

Initial Velocity Studies with Variable UDP-Glucose and 5-hmC DNA. Initial velocity experiments allow the

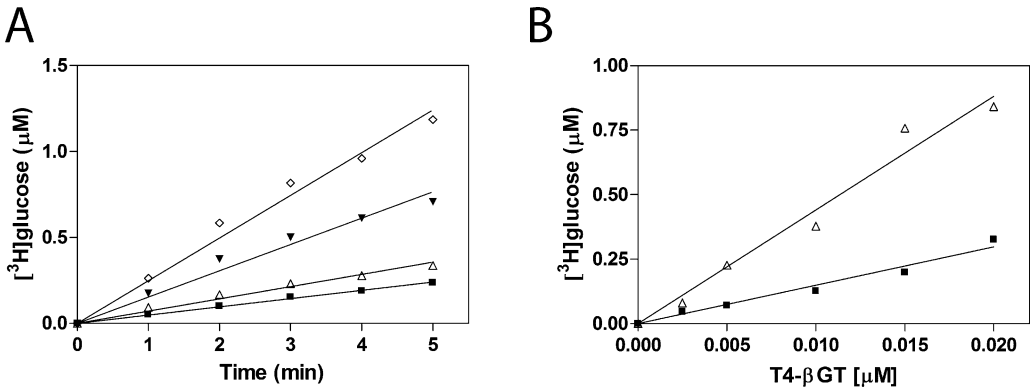


Figure 2. Linearity of glucosylation by recombinant β-GT. (A) Linearity of glucosylase reaction as a function of time with four different concentrations of UDP-[^{3}H]glucose: 50 (◇), 25 (▼), 10 (△), and 5 μM (■). All reaction mixtures contained 0.01 μM β-GT and 2.5 μM 5-hmC and were incubated at 25 °C. Twenty-five microliters of the reaction mixture was spotted on DE81 at 1 min time intervals and processed as described in Experimental Procedures. (B) Linearity of the glucosylase reaction as a function of recombinant β-GT enzyme concentration (0.00025–0.02 μM) with 50 (△) or 25 μM UDP-[^{3}H]glucose (■). Twenty-five microliters of the reaction mixture was spotted on DE81 after incubation with the appropriate concentration of the β-GT enzyme for 2 min.

dx.doi.org/10.1021/bi2014739 | *Biochemistry* 2012, 51, 1009–1019

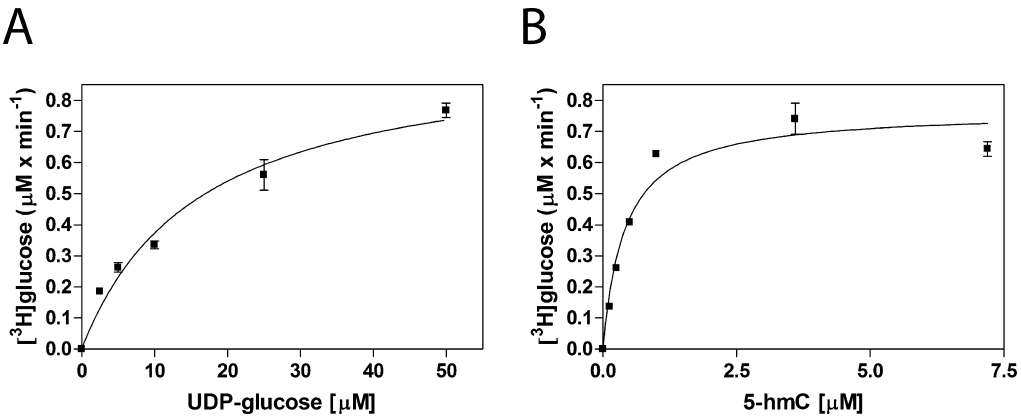


Figure 4. Substrate−velocity curves of recombinant β-GT. (A) Recombinant β-GT activity with UDP-[^{3}H]glucose substrate. Glucosylation reactions were conducted at UDP-[^{3}H]glucose substrate concentrations of 0, 2.5, 5, 10, 25, and 50 μM and fixed enzyme and 5-hmC DNA concentrations of 0.01 and 2.5 μM, respectively. Product formation is plotted vs substrate concentration, and nonlinear regression was performed to determine $K_m^{\text{UDP-glucose}}$ values. (B) Recombinant β-GT activity with a 5-hmC DNA substrate. Glucosylation reactions were conducted with 5-hmC DNA substrate concentrations of 0, 0.125, 0.25, 0.5, 1, 3.6, and 7.2 μM and fixed enzyme and UDP-[^{3}H]glucose concentrations of 0.01 and 50 μM, respectively. Product formation is plotted vs substrate concentration, and nonlinear regression was performed to determine $K_m^{\text{5-hmC}}$ values.

Table 1. Steady-State Kinetic Parameters of Recombinant β-GT[a]

substrate	$K_m^{\text{5-hmC}}$ (μM)	$K_m^{\text{UDP-glucose}}$ (μM)	$k_{cat}/K_m^{\text{5-hmC}}$ (μM min^{-1})	$k_{cat}^{\text{5-hmC}}$ ($\times 10^6$ M^{-1} min^{-1})
T4-*gt* DNA	0.4 ± 0.1	16 ± 3	192	77
2 × 5-hmC DNA	3 ± 0.6	NA	56	168
6 × 5-hmC DNA	0.6 ± 0.1	NA	166	100
12 × 5-hmC DNA	0.23 ± 0.06	NA	269	62
24 × 5-hmC DNA	0.22 ± 0.05	NA	264	58

[a]The T4-*gt* DNA substrate has every cytosine hydroxymethylated. The 2 to 24 × hmC DNA substrates are 100 bp double-stranded PCR products containing a number of hmC residues equivalent to their identifiers. All tested samples were assessed in duplicate (±standard deviation). See Experimental Procedures for sequences of 100 bp fragments.

determination of kinetic constants for a specific enzyme and its substrate, which in turn establish parameters for the kinetic mechanism of a reaction. In the case of recombinant β-GT, there are two different substrates, UDP-glucose and 5-hmC DNA. To determine the kinetic constants, we varied the UDP-glucose concentration from 0 to 50 μM, keeping the 5-hmC DNA in excess (2.5 μM) and recombinant β-GT at 0.01 μM. Similarly, for determination of 5-hmC DNA kinetic parameters, we varied the 5-hmC DNA concentration between 0 and 7.5 μM and kept the UDP-glucose in excess (50 μM) and the enzyme concentration at 0.01 μM. We plotted glucosylation as a function of UDP-glucose (Figure 4A) or 5-hmC DNA (Figure 4B). The apparent $K_m^{\text{5-hmC}}$ and $K_m^{\text{UDP-glucose}}$ values were 0.41 and 16 μM, respectively (Table 1). The turnover number for β-GT on T4-*gt* DNA was 77 min^{-1} (Table 1).

Product Inhibition Studies of Recombinant β-GT with UDP and 5-ghmC DNA. To provide insight into the reaction mechanisms, we conducted product inhibition studies. An equilibrium scheme for enzyme turnover in the presence or absence of the product is shown in Figure 5A. Because of the structural similarities between substrate UDP-glucose and UDP, 5-hmC, and 5-ghmC DNAs, the product molecule of an enzymatic reaction is expected to be competitive with the original substrate molecule. The product molecule may bind directly at the active site, blocking further binding of the substrate molecule, or it can bind at the proximity of the active site and influence structural constraints for substrate binding. In a bisubstrate reaction, as in the case of recombinant β-GT, the calculated inhibition constants are apparent, for the given concentration of the cosubstrate, which is fixed. For glucosyltransferases such as β-GT, transfer of glucose from UDP-glucose to 5-hmC DNA generates two different products, UDP and 5-ghmC. Because β-GT is a bisubstrate enzyme, a competitive inhibitor of one of the substrate-binding sites will display either a competitive or a mixed inhibition pattern, depending on which substrate binds first to the enzyme. Varying the amounts of UDP-glucose with different fixed amounts of UDP yielded a set of double-reciprocal plots that converged on the *y*-axis intercept (Figure 5B), suggesting UDP is a competitive inhibitor with respect to UDP-glucose. Nonlinear regression analysis of the same series of reactions displayed similar V_{max} values but different K_m^{app} values (data not shown). The K_i value of UDP was ~9 μM (Table 1) and was calculated by plotting K_m^{app} versus inhibitor concentration (see the inset of Figure 5B). The K_i value for UDP is 2-fold lower than the K_m value for UDP-glucose (Table 1), suggesting a higher affinity for UDP by the β-GT enzyme. Varying the amounts of 5-hmC DNA with different fixed concentrations of UDP resulted in a set of reciprocal plots that all converged to the left of the *y*-axis (Figure 5C), indicating that UDP acts as a mixed inhibitor with respect to 5-hmC DNA.

Similarly, varying levels of 5-hmC DNA with different fixed amounts of 5-ghmC DNA yielded a set of double-reciprocal plots that did converge on the *y*-axis (Figure 5D), revealing that 5-ghmC acts as a competitive inhibitor with respect to 5-hmC. The K_i value of 5-ghmC was ~4 μM (Table 2). The K_i value for 5-ghmC is 10-fold lower than the K_m value for 5-hmC (Tables 1 and 2), suggesting a much higher affinity of the β-GT enzyme for 5-hmC. Varying amounts of UDP-glucose with different fixed concentrations of 5-ghmC resulted in a set of reciprocal

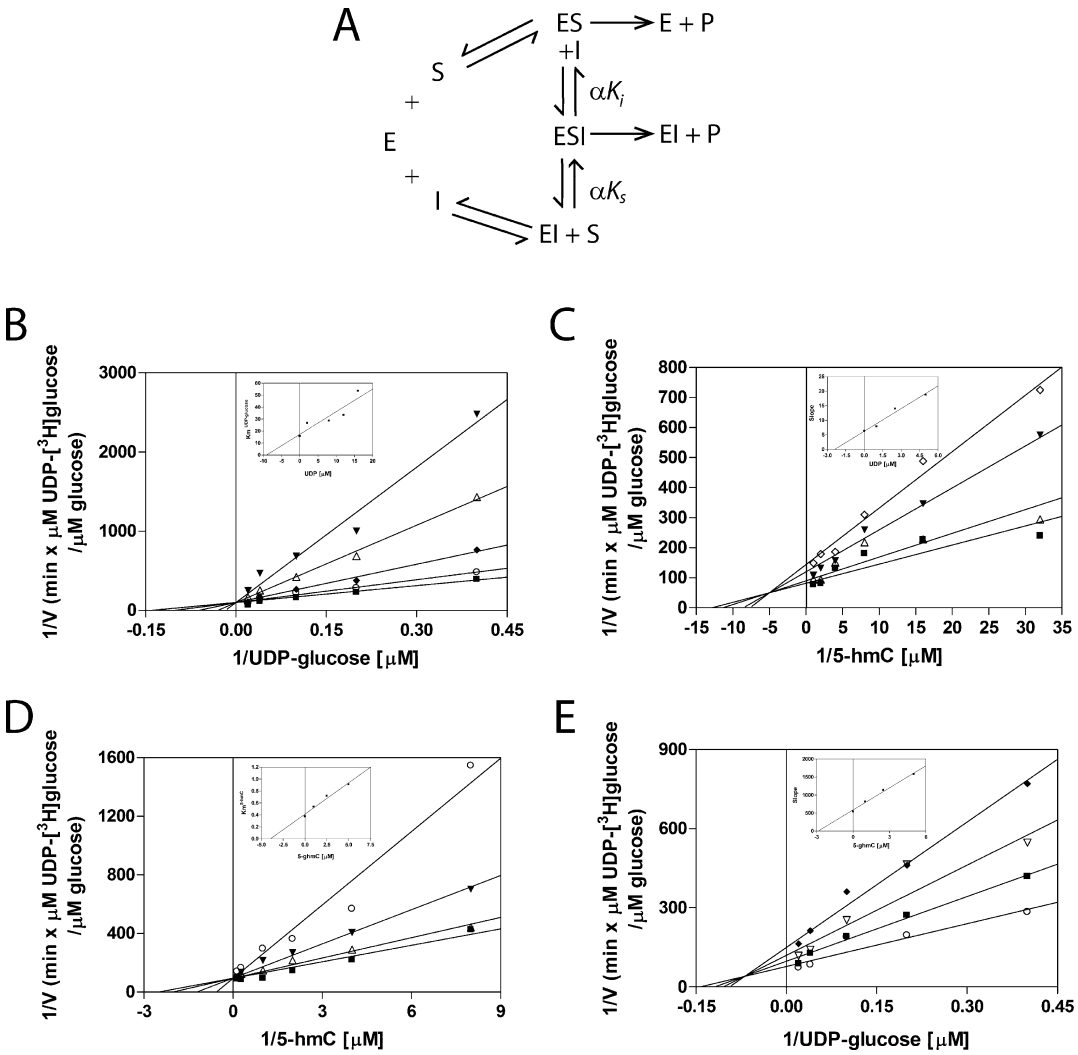


Figure 5. Reaction mechanism of recombinant β-GT and double-reciprocal plots of the initial velocity vs substrate concentration. (A) Scheme detailing the turnover of recombinant β-GT enzyme in the presence or absence of product inhibitors. The substrate 5-hmC or UDP-glucose is S. The product glucosylated 5-hmC is I. The inhibitor constant is K_i. The binary inhibitor constant is αK_i. The degree of modification the first substrate exerts on binding of the second substrate is α. (B) Double-reciprocal plots for fixed 5-hmC substrate and variable UDP concentrations of 0 (■), 2 (○), 8 (◆), 12 (△), and 16 μM (▼). Linear regression was performed, and $1/v$ was plotted vs 1/[UDP-[^{3}H]glucose]. The inset replot of $K_m^{\text{UDP-glucose}}$ vs UDP was used to calculate the K_i value of UDP as a competitive inhibitor with respect to the formation of a β-GT−UDP-glucose complex. (C) Double-reciprocal plots for fixed UDP-[^{3}H]glucose substrate and variable UDP concentrations of 0 (■), 20 (△), 30 (▼), and 40 μM (◊). Linear regression was performed, and $1/v$ was plotted vs 1/[5-hmC]. The inset replot of slope vs UDP was used to calculate the αK_i of UDP as a mixed inhibitor with respect to the formation of a β-GT−5-hmC complex. (D) Double-reciprocal plots for fixed UDP-[^{3}H]glucose substrate and variable glucosylated 5-hmC concentrations of 0 (■), 1 (△), 2.5 (▼), and 5 (○). Linear regression was performed, and $1/v$ was plotted vs 1/[5-hmC DNA]. The inset replot of $K_m^{\text{5-hmC}}$ vs 5-ghmC was utilized to calculate the K_i value of 5-ghmC as a competitive inhibitor with respect to the formation of a β-GT−5-hmC complex. (E) Double-reciprocal plots for 5-hmC concentration and variable glucosylated 5-hmC concentrations of 0 (○), 1 (■), 2.5 (▽), and 5 μM (◆). Linear regression was performed, and $1/v$ was plotted vs 1/[UDP-[^{3}H]glucose]. The inset replot of slope vs 5-ghmC was used to calculate the αK_i of 5-ghmC as a mixed inhibitor with respect to formation of a β-GT−UDP-glucose complex.

Table 2. Product Inhibition Pattern of Recombinant β-GT[a]

product	variable substrate	fixed substrate	type of inhibition	K_i (μM)	αK_i (μM)
UDP	UDP-glucose	5-hmC	competitive	9	
UDP	5-hmC	UDP-glucose	mixed	2	72
5-ghmC	5-hmC	UDP-glucose	competitive	4	
5-ghmC	UDP-glucose	5-hmC	mixed	3	5

[a]The DNAs used for the assay were 5-hmC DNA (T4-*gt* DNA) and 5-ghmC DNA (T4-*wt* DNA) and tested in duplicate (±standard deviation).

plots that converged to the left of the *y*-axis, suggesting a mixed inhibition pattern (Figure 5E). The αK_i values were obtained by plotting $1/V_{max}$ versus inhibitor concentration (data not shown). The K_i and αK_i for UDP and 5-hmC DNA were 2 and 72 μM, respectively. Similarly, the K_i and αK_i for 5-ghmC DNA and UDP-glucose were 3 and 5 μM, respectively (Table 2).

Differential Glucosylation by Recombinant β-GT on 5-hmC-Containing Substrates. Because β-GT is a T4 phage-encoded enzyme in which most of the 5-hmC is converted to glucosylated 5-hmC, we determined if an increased density of 5-hmC on a specific sequence of DNA can affect the steady-state kinetic parameters of the enzyme. We amplified via PCR DNAs with 2, 6, 12, or 24 × 5-hmC sites and used them as substrates. The enzyme bound strongly to DNA containing larger numbers of 5-hmC residues (>12) as

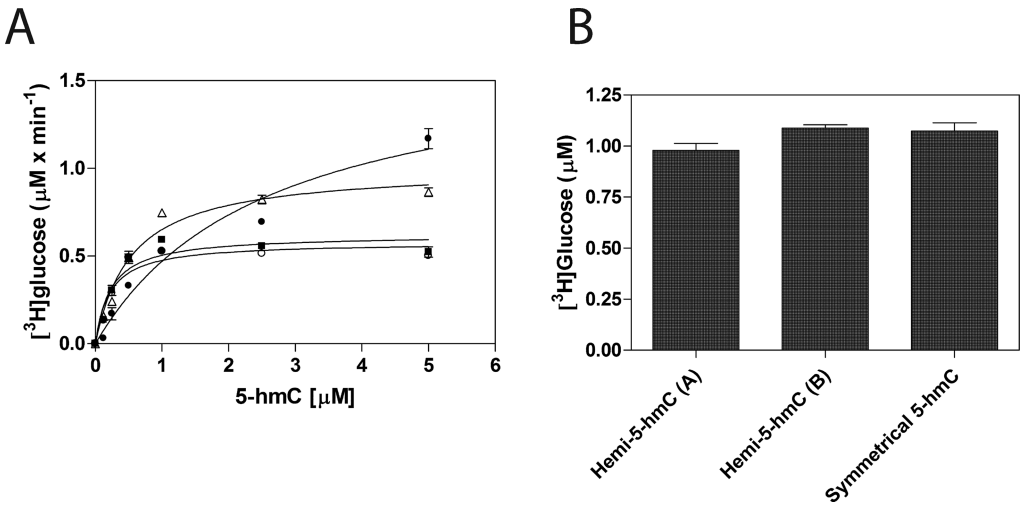


Figure 6. Glucosylation of different 5-hmC-containing substrates by recombinant β-GT. (A) Recombinant β-GT activity with 100 bp double-stranded DNA substrates containing 2 (●), 6 (△), 12 (■), or 24 cytosines (○) that are hydroxymethylated. Glucosylation reactions were performed at 5-hmC concentrations of 0, 0.125, 0.25, 0.5, 1, and 2.5 μM and fixed enzyme and UDP-[^{3}H]glucose concentrations of 0.01 and 50 μM, respectively, in a total volume of 25 μL. (B) End point analysis of recombinant β-GT activity with 55 bp double-stranded substrates containing hemi-5-hmC pair A, hemi-5-hmC pair B, or symmetrical 5-hmC. Glucosylation reactions were performed with 1 μM 5-hmC, 50 μM UDP-[^{3}H]glucose, and 5 units of recombinant β-GT in a 25 μL reaction mixture for 2 h. The reaction mixture was spotted in duplicate on DE81 membranes, and glucosyl group incorporation was measured.

indicated by the lower K_m values (Figure 6A and Table 1). Stronger binding also influenced the reaction velocity, resulting in lower turnover numbers. For example, the DNA samples described above with 2, 6, 12, and 24 × 5-hmC displayed turnover numbers of 168, 100, 62, and 58, respectively (Table 1).

In another assay, we studied the kinetic parameters of recombinant β-GT on hemihydroxymethylated versus fully hydroxymethylated symmetrical CG sequence. Once again, the symmetrical 5-hmC-containing oligonucleotide duplex displayed a higher turnover number, 310, than the hemi-5hmC corresponding sequences, which exhibited numbers of 8 and 5 (Table 3). Previous studies that have utilized β-GT to measure the levels of 5-hmC have not observed β-GT's strong preference for the glucosylation of symmetrical 5-hmC over hemi-5-hmC. When the oligonucleotide duplexes at identical molarities of 5-hmC were again incubated with β-GT, but this time for a longer period of time (2 h), this revealed that β-GT is capable of saturation glucosylation on all conformations of 5-hmC (Figure 6B). From the experiments described above, we conclude that the numbers of 5-hmC residues on DNA can influence β-GT catalysis, especially the rate of catalysis, and hemihydroxymethylated DNA is not the preferred substrate of the enzyme.

Table 3. Hemi versus Symmetrical 5-hmC as a Substrate for Recombinant β-GT[a]

substrate	$K_m^{5\text{-hmC}}$	$k_{cat}/K_m^{5\text{-hmC}}$ (μM min^{-1})	$k_{cat}^{5\text{-hmC}}$ ($\times 10^6$ M^{-1} min^{-1})
hemi-5-hmC (**A**)	0.4 ± 0.1	20	8
hemi-5-hmC (**B**)	0.2 ± 0.1	25	5
symmetrical 5-hmC	2.0 ± 0.5	155	310

[a]Hemi-5-hmC (**A** and **B**) and symmetrical 5-hmC are 55 bp double-stranded synthetic oligonucleotides containing either a single hemi-5-hmC or two symmetrical 5-hmC residues. All tested samples were assessed in duplicate (±standard deviation). See Experimental Procedures for sequences of 55 bp fragments.

Determination of Global 5-hmC Levels in Various Genomes by Glucosylation. Because β-GT-mediated transfer of a tritiated glucose molecule is efficient, we extended our studies to global genomic 5-hmC quantification by [^{3}H]glucose incorporation. First, to determine if the transfer of [^{3}H]glucose to hydroxymethylated cytosine is linear with regard to total 5-hmC content, a 100 bp double-stranded PCR fragment containing 24 × 5-hmC was utilized as a standard substrate for the β-GT enzyme. Unlike our previous glucosylase reactions, which characterized the β-GT enzyme, these assays were conducted for 2 h at 37 °C to ensure complete glucosylation of all 5-hmC residues. The glucosylation reaction was strictly linear (r = 0.9997) within the range of 12−500 fmol of standards when cold UDP-glucose was used in a 4-fold excess over UDP-[^{3}H]glucose (Figure 7A). Furthermore, at least 92% of the total 5-hmC in these standards was labeled with UDP-[^{3}H]glucose, revealing that these conditions allowed for the sensitive and accurate measurement of total 5-hmC content.

The 5-hmC levels ranged from 2 to 30 fmol of 5-hmC/μg of genomic DNA tested, corresponding to the percentage coverage values listed in Table 4. This revealed that the highest percentages of 5-hmC tested were in the human, mouse, and cow brain tissue samples (0.5, 0.9, and 0.7% of total nucleotides, respectively). The 0.9% of global 5-hmC that we recorded in mouse brain tissue was only slightly lower than the ∼1.1−1.5% of 5-hmC recorded by Szwagierczak et al. in different types of mouse brain tissues.[18] Furthermore, another report using thin layer chromatography and high-pressure liquid chromatography coupled to mass spectrometry had very similar results, quantifying global 5-hmC levels in Purkinje tissue to be 0.6%,[8] which was similar to the value of 0.5% that we measured in the human brain tissue sample using our experimental protocol.

Plants utilize a DEMETER class of DNA glucosylases to target and remove 5-mC from their genomes.[20,21] This negates the need for plants to have a Tet1 homologue to convert 5-mC into the 5-hmC intermediate required to possibly promote DNA demethylation. Furthermore, there are no known homologues of

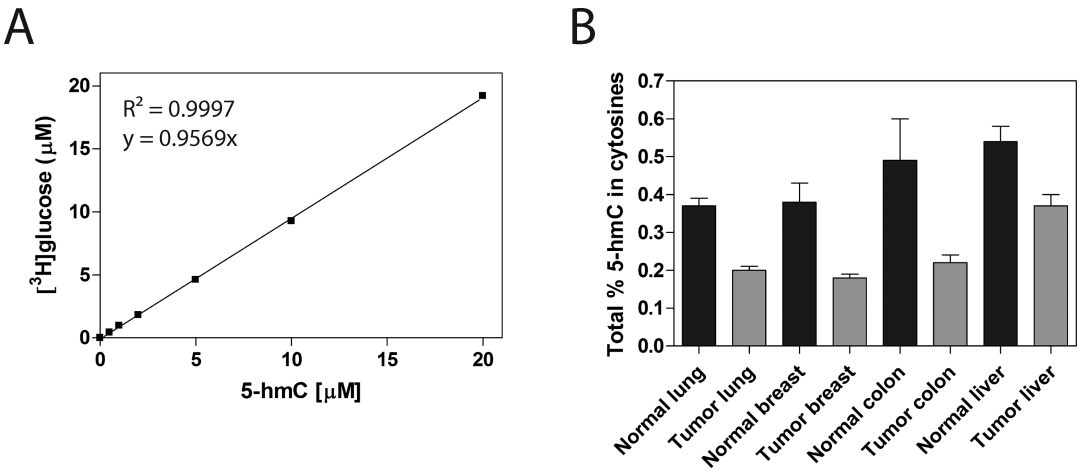


Figure 7. Global 5-hmC levels determined in the genomes of various tissues, and corresponding tumor samples, by recombinant β-GT. (A) Calibration curve using different amounts of 0, 0.5, 1, 2, 5, 10, and 20 μM 5-hmC with 5 units of recombinant β-GT and 25 μM UDP-[^{3}H]glucose in a total reaction volume of 25 μL. The reaction mixture was spotted in duplicate on DE81 membranes, and glucosyl group incorporation was measured. The linear relationship between [^{3}H]glucose incorporation and the molarity of 5-hmC is 0.9997. (B) Glucosylation is utilized to measure global 5-hmC levels in matched normal and tumor genomic DNA samples, as described in Experimental Procedures. Note that all tumor samples have significantly lower levels of 5-hmC when compared to the matched normal sample ($p < 0.001$).

the Tet family of oxidases in any species of plants. In agreement with these findings, the plant genomes tested displayed low levels of 5-hmC (<0.07% of the total nucleotides). Similarly, sea urchin (*Lytechinus variegatus*) embryos had small amounts of genomic 5-hmC at early and mid gastrula (<0.07% of total nucleotides), although there was an increase in the level of 5-hmC (0.11%) 36 hours post fertilization. Analysis of cancer and matched tissue genomes showed reduced amounts of 5-hmC (Figure 7B). For example, normal lung, breast, and colon versus cancer genomes were 0.15% versus 0.08%, 0.16% versus 0.07%, and 0.20% versus 0.09% of total nucleotides, respectively (Table 4). In the colorectal cancer cell line HCT116, the level of 5-hmC was 0.12%, a value reduced to 0.07% in DNMT1 null cells in a HCT116 cell background, which was hypomethylated for 5-mC. Similarly, wild-type mouse ESC displayed 0.27% 5-hmC as compared to 0.22% in Tet1 null ES cells, suggesting a loss of Tet1 function reduces the level of 5-hmC in the genome.

■ DISCUSSION

Modified DNA bases are widespread in living organisms from viruses to mammals. In T4 phages, most of the cytosine residues are present as 5-hydroxymethylcytosine and are often completely glucosylated by virally encoded β-GT.[2,22] Similarly, in the mammalian genome, a percentage of cytosine residues are 5-mC, and a percentage of that is further oxidized to 5-hmC. In 1972, formic acid hydrolysis followed by chromatographic analysis was used to detect 5-hmC in murine brain and liver DNA. Using this method, 5-hmC was estimated to comprise ~15% of the total cytosine residues.[23] Recently, β-GT was used in combination with a radioactive UDP-glucose to determine global 5-hmC levels. In our previous report, we successfully utilized a glucosylation step using a recombinant β-GT in combination with restriction enzyme cleavage to quantify the percentage of C, 5-mC, and 5-hmC in mammalian genomes.[24] To understand the detailed biochemical properties of recombinant β-GT, we determined a series of steady-state kinetic properties and the reaction mechanisms of the enzyme.

Recombinant β-GT glucosylates double-stranded 5-hmC DNA very efficiently, as shown by its ability to completely glucosylate ~100-nucleotide substrates and T4-*gt* DNA when the enzyme is given sufficient time. However, clusters of six or more 5-hmC residues on the 100 bp duplex substrate led to reduced turnover numbers and K_m values, indicating that β-GT binds tightly to clustered 5-hmC on DNA. Such behavior stands in contrast to that of other transferases such as DNA methyltransferase 3A (Dnmt3A), where increases in the density of methylation target sites resulted in higher rates of DNA methylation.[25] Crystallographic studies of β-GT were not able to distinguish between a processive or distributive mode of glucosylation between adjacent 5-hmC residues. A processive mechanism seems unlikely because the active site of the β-GT is completely closed upon substrate DNA binding, requiring unfolding after catalysis to release the UDP product and bind new UDP-glucose substrate.[26,27] This hypothesis was experimentally confirmed in our studies, as the chase nonbiotinylated substrate was able to remove the enzyme and UDP-glucose from the prebound biotinylated substrate DNA. This reaction mechanism was confirmed by product inhibition studies in which β-GT appears to follow a bi-bi random reaction mechanism. Similar reaction mechanisms are common for other histone and DNA methyltransferases, such as lysine methyltransferases (G9a[28] and SUV39H[19]) and DNA cytosine 5-methyltransferases (DNMT1,[29] M.*Hha*I,[30] and M.*Bam*H1[31]). However, in vivo, most of the enzymes may reside in complexes, and the reaction pathways and product formation may be different.

Cell culture-based studies have identified a strong correlation between DNA hypomethylation (reduction of 5-mC) and tumorigenesis.[32,33] In a recent study, 5-mC was purposely converted to 5-hmC in adult mouse brain, and this promoted demethylation through a process requiring base excision repair. Furthermore, the 5-hmC is actively removed by the activation-induced deaminase/APOBEC family of cytidine deaminases in mammalian cells.[10] This evidence would indicate that late stage tumorigenic samples should have lower levels of 5-hmC, when compared to match normal tissue, because 5-hmC is the precursor to the hypomethylation. To test this hypothesis, we measured 5-hmC levels in matched pairs of human genomic DNA from brain, lung, breast, colon, and liver tissues that were

Table 4. 5-hmC Content of Various Genomes and in Tumor and Normal Tissue Samples[a]

genomic DNA sample	genome size (bp)	% GC content	fmol of 5-hmC/μg of DNA	total % 5-hmC in genome	total % 5-hmC in cytosines	DNA
E14 cell line	2.65×10^9	42.0	9.0 ± 0.9	0.30	0.71	mouse
3T3 cell line			3.6 ± 0.2	0.12	0.28	mouse
mouse ESC (Tet1 control)			8.2 ± 0.7	0.27	0.64	mouse
mouse ESC (Tet1 −/−)			6.6 ± 0.0	0.22	0.52	mouse
mouse ESC, day 1			8.1 ± 0.1	0.27	0.63	mouse
mouse ESC, day 3			5.9 ± 0.1	0.19	0.46	mouse
mouse ESC, day 5			8.1 ± 0.4	0.27	0.64	mouse
mouse ESC, day 7			7.4 ± 1.2	0.24	0.58	mouse
adult mouse spleen tissue			4.3 ± 1.1	0.14	0.34	mouse
adult mouse brain tissue			27 ± 0.5	0.90	2.1	mouse
adult mouse liver tissue			6.4 ± 0.3	0.21	0.51	mouse
adult mouse kidney tissue			10 ± 1.2	0.33	0.79	mouse
adult mouse heart tissue			10 ± 2.2	0.33	0.79	mouse
adult mouse lung tissue			8.6 ± 0.4	0.28	0.67	mouse
male rat	2.51×10^9	42.0	9.2 ± 0.9	0.30	0.73	rat
female rat			7.1 ± 1.1	0.23	0.56	rat
U87 cell line	3.08×10^9	41.0	5.8 ± 2.4	0.19	0.47	human
293 cell line			7.3 ± 3.0	0.24	0.59	human
HCT cell line			3.5 ± 1.1	0.12	0.28	human
HCT cell line* (DNMT1 −/−)			2.2 ± 0.0	0.07	0.17	human
HeLa cells			3.4 ± 1.8	0.11	0.27	human
IMR90 cells			4.0 ± 2.6	0.13	0.32	human
human normal brain tissue			15 ± 1.2	0.48	1.2	human
human normal lung tissue			***4.6 ± 0.3***	***0.15***	***0.37***	human
human normal breast tissue			***4.7 ± 0.7***	***0.16***	***0.38***	human
human normal colon tissue			***6.1 ± 1.4***	***0.20***	***0.49***	human
human normal liver tissue			***6.8 ± 0.5***	***0.22***	***0.54***	human
human tumor lung tissue			***2.5 ± 0.2***	***0.08***	***0.20***	human
human tumor breast tissue*			***2.2 ± 0.1***	***0.07***	***0.18***	human
human tumor colon tissue			***2.7 ± 0.3***	***0.09***	***0.22***	human
human tumor liver tissue			***4.7 ± 0.4***	***0.15***	***0.37***	human
human metastatic colon tissue			2.6 ± 0.6	0.09	0.21	human
L. variegatus,* early gastrula	8.14×10^8 [b]		2.2 ± 0.0	0.07	N/A	sea urchin
L. variegatus,* late gastrula			2.2 ± 0.0	0.07	N/A	sea urchin
L. variegatus,* 36 h post fertilization			3.3 ± 0.4	0.11	N/A	sea urchin
*Arabidopsis**	1.25×10^8		2.2 ± 0.0	0.07	N/A	plant
soy bean*	1.15×10^9		2.2 ± 0.0	0.07	N/A	plant
rice*	3.90×10^8		2.2 ± 0.1	0.07	N/A	plant
bovine brain tissue	2.87×10^9	41.7	21 ± 0.8	0.69	1.7	other mammals
male dog	2.38×10^9	41.0	5.8 ± 0.2	0.19	0.47	other mammals
chicken embryo	1.05×10^9	50.0	2.3 ± 0.2	0.08	0.15	other mammals
rabbit liver tissue	2.60×10^9		4.1 ± 0.6	0.13	N/A	other mammals
Rhesus monkey heart	3.09×10^9	40.7	6.4 ± 0.2	0.21	0.52	other mammals
COS cell line (Vervet monkey)	3.00×10^9		4.7 ± 0.7	0.16	N/A	other mammals

[a]Genomic 5-hmC content was determined by measuring the transfer of [^{3}H]glucose to 5-hmC by recombinant β-GT in 1 μg of total genomic DNA. Background measurements, which were obtained from β-GT glucosylation of 1 μg of synthetic poly(dI-dC) DNA containing no 5-hmC, were subtracted from all values for genomic DNA samples. Genomic samples marked with an asterisk had ≤2 fmol of 5-hmC/μg of DNA and ≤0.07 total % 5-hmC in the genome. The total percentage of 5-hmC in cytosines was not calculated for genomes without percent GC determinations. Values of genomic 5-hmC content for human matched normal and tumor tissues are bold and italicized. [b]Genome of *L. variegatus* is estimated.

either normal or tumorgenic. On average, all the tumor samples had significantly ($p < 0.001$) lower levels of genomic 5-hmC than their matched normal tissue (Figure 7B). This evidence lends credibility to the possibility that 5-hmC is acting as a precursor to hypomethylation or 5-hmC reductions are a reflection of reduced precursor 5-mC. Indeed, 5-mC levels in a large series of tumor DNA samples are significantly lower than what was observed in their matched normal tissue.[34]

Comparison between 5-hmC levels in control HCT116 cell and DNMT1 null background also demonstrated reduced levels of 5-hmC due to a genetic knockout of the methyltransferase gene, suggesting that a decrease in the level of the 5-mC precursor may lead to less oxidative product.[18] There is one area of concern when utilizing β-GT to measure global 5-hmC levels, the presence of 5-hydroxymethyluracil (5-hmU) in the tested genomic DNA samples. 5-hmU, which is also a substrate

for β-GT,[35] arises from oxidation of the methyl group on thymine in genomic DNA.[36] When utilizing β-GT for genomic 5-hmC measurements, the presence of 5-hmU can give false measurements of elevated 5-hmC levels. This was not a concern for our tested mammalian genomic DNA samples because in these biological systems 5-hydroxymethyluracil-DNA glycosylase (hmUDG) will glucosylate 5-hmU, leading to its subsequent removal by nucleotide excision repair,[37] yet no homologue for hmUDG has been discovered in either plants or sea urchins; therefore, care needs to be taken when interpreting the genomic 5-hmC levels with β-GT in lower organisms.

Our β-GT-based measurements of global 5-hmC levels during ES cell differentiation were further validated for their accuracy by comparing the data with HPLC-based measurement of the genomic DNA that was hydrolyzed as nucleobases.[8] Indeed, the patterns of 5-hmC changes were very similar in both assay systems. A study by Szwagierczak et al. used a comparable radioactive technique to measure total 5-hmC content, and even though they utilized different DNA standards, they also identified a similar linear relationship between the β-GT transfer of [^{3}H]glucose and total 5-hmC content (r = 0.9991).[18] However, in this earlier study, the 5-hmC levels were measured only in different mouse tissues at different stages of development. We expanded upon this study by establishing a more comprehensive measurement of 5-hmC that was then applied to more than 40 different genomic DNA samples, spanning 13 different prokaryotic and eukaryotic species. In conclusion, β-GT-mediated global determination of 5-hmC levels is consistent with other analytical methods, and because of the simplicity of glucose incorporation measurement, smaller amounts of DNA can be used in the high-throughput format.

■ AUTHOR INFORMATION

Corresponding Author

*New England Biolabs, Inc., 240 County Rd., Ipswich, MA 01938. Telephone: (978) 380-7227. Fax: (978) 350-9211. E-mail: pradhan@neb.com.

Funding

Research was supported by SBIR Grant 4R44GM095209-02 from the National Institutes of Health.

■ ACKNOWLEDGMENTS

We thank Drs. Pierre-Olivier Estéve, Shannon Kinney, Thomas C. Evans, and William Jack for reading the manuscript. We thank Dr. Mala Samaranayake for providing recombinant β-GT and Dr. Shannon Kinney for isolating DNA from the human and mouse cell lines. We kindly thank Dr. Cynthia A. Bradham for growing and collecting the *L. variegatus* and Dr. Devora Cohen-Karni for isolating genomic DNA from the *L. variegatus* and chicken embryo samples; and for her helpful discussion into the measurement of 5-hmC in different genomes. We also thank Dr. Donald G. Comb and Mr. James V. Ellard and New England Biolabs, Inc., for research support.

■ ABBREVIATIONS

TET, Ten Eleven Translocation; β-GT, β-glucosyltransferase; 5-hmC, 5-hydroxymethylcytosine; 5-hmU, 5-hydroxymethyluracil; 5-mC, 5-methylcytosine; 5-ghmC, glucosylated 5-hmC; UDP-glucose, uridine diphosphoglucose; ESC, embryonic stem cells; GLIB, glucosylation, periodate oxidation, biotinylation; CMS, cytosine 5-methylenesulfonate; UDP-[^{3}H]glucose, UDP-[1-^{3}H]glucose; Dnmt3A, DNA methyltransferase 3A.

■ REFERENCES

(1) Josse, J., and Kornberg, A. (1962) Glucosylation of deoxyribonucleic acid. III. α- and β-glucosyl transferases from T4-infected *Escherichia coli*. *J. Biol. Chem. 237*, 1968−1976.

(2) Kornberg, S. R., Zimmerman, S. B., and Kornberg, A. (1961) Glucosylation of deoxyribonucleic acid by enzymes from bacteriophage-infected *Escherichia coli*. *J. Biol. Chem. 236*, 1487−1493.

(3) Dharmalingam, K., and Goldberg, E. B. (1979) Restriction in vivo. III. General effects of glucosylation and restriction on phage T4 gene expression and replication. *Virology 96*, 393−403.

(4) Cox, G. S., and Conway, T. W. (1973) Template properties of glucose-deficient T-even bacteriophage DNA. *J. Virol. 12*, 1279−1287.

(5) Ruger, W. (1978) Transcription of bacteriophage T4 DNA in vitro: Selective initiation with dinucleotides. *Eur. J. Biochem. 88*, 109−117.

(6) Penn, N. W. (1976) Modification of brain deoxyribonucleic acid base content with maturation in normal and malnourished rats. *Biochem. J. 155*, 709−712.

(7) Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L., and Rao, A. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science 324*, 930−935.

(8) Kriaucionis, S., and Heintz, N. (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science 324*, 929−930.

(9) Ko, M., Huang, Y., Jankowska, A. M., Pape, U. J., Tahiliani, M., Bandukwala, H. S., An, J., Lamperti, E. D., Koh, K. P., Ganetzky, R., Liu, X. S., Aravind, L., Agarwal, S., Maciejewski, J. P., and Rao, A. (2010) Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature 468*, 839−843.

(10) Guo, J. U., Su, Y., Zhong, C., Ming, G. L., and Song, H. (2011) Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell 145*, 423−434.

(11) Koh, K. P., Yabuuchi, A., Rao, S., Huang, Y., Cunniff, K., Nardone, J., Laiho, A., Tahiliani, M., Sommer, C. A., Mostoslavsky, G., Lahesmaa, R., Orkin, S. H., Rodig, S. J., Daley, G. Q., and Rao, A. (2011) Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell 8*, 200−213.

(12) Wu, H., D'Alessio, A. C., Ito, S., Xia, K., Wang, Z., Cui, K., Zhao, K., Eve Sun, Y., and Zhang, Y. (2011) Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature 473*, 389−393.

(13) Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S., and Jacobsen, S. E. (2011) 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol. 12*, R54.

(14) Iqbal, K., Jin, S. G., Pfeifer, G. P., and Szabo, P. E. (2011) Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc. Natl. Acad. Sci. U.S.A. 108*, 3642−3647.

(15) Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C. J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J. (2011) 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat. Commun. 2*, 241.

(16) Ficz, G., Branco, M. R., Seisenberger, S., Santos, F., Krueger, F., Hore, T. A., Marques, C. J., Andrews, S., and Reik, W. (2011) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature 473*, 398−402.

(17) Pastor, W. A., Pape, U. J., Huang, Y., Henderson, H. R., Lister, R., Ko, M., McLoughlin, E. M., Brudno, Y., Mahapatra, S., Kapranov, P., Tahiliani, M., Daley, G. Q., Liu, X. S., Ecker, J. R., Milos, P. M., Agarwal, S., and Rao, A. (2011) Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature 473*, 394−397.

(18) Szwagierczak, A., Bultmann, S., Schmidt, C. S., Spada, F., and Leonhardt, H. (2010) Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res. 38*, e181.
(19) Chin, H. G., Patnaik, D., Esteve, P. O., Jacobsen, S. E., and Pradhan, S. (2006) Catalytic properties and kinetic mechanism of human recombinant Lys-9 histone H3 methyltransferase SUV39H1: Participation of the chromodomain in enzymatic catalysis. *Biochemistry 45*, 3272−3284.
(20) Gehring, M., Huh, J. H., Hsieh, T. F., Penterman, J., Choi, Y., Harada, J. J., Goldberg, R. B., and Fischer, R. L. (2006) DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell 124*, 495−506.
(21) Morales-Ruiz, T., Ortega-Galisteo, A. P., Ponferrada-Marin, M. I., Martinez-Macias, M. I., Ariza, R. R., and Roldan-Arjona, T. (2006) DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. *Proc. Natl. Acad. Sci. U.S.A. 103*, 6853−6858.
(22) Georgopoulos, C. P., and Revel, H. R. (1971) Studies with glucosyl transferase mutants of the T-even bacteriophages. *Virology 44*, 271−285.
(23) Penn, N. W., Suwalski, R., O'Riley, C., Bojanowski, K., and Yura, R. (1972) The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. *Biochem. J. 126*, 781−790.
(24) Kinney, S. M., Chin, H. G., Vaisvila, R., Bitinaite, J., Zheng, Y., Esteve, P. O., Feng, S., Stroud, H., Jacobsen, S. E., and Pradhan, S. (2011) Tissue-specific Distribution and Dynamic Changes of 5-Hydroxymethylcytosine in Mammalian Genomes. *J. Biol. Chem. 286*, 24685−24693.
(25) Holz-Schietinger, C., and Reich, N. O. (2010) The inherent processivity of the human de novo methyltransferase 3A (DNMT3A) is enhanced by DNMT3L. *J. Biol. Chem. 285*, 29091−29100.
(26) Morera, S., Imberty, A., Aschke-Sonnenborn, U., Ruger, W., and Freemont, P. S. (1999) T4 phage β-glucosyltransferase: Substrate binding and proposed catalytic mechanism. *J. Mol. Biol. 292*, 717−730.
(27) Morera, S., Lariviere, L., Kurzeck, J., Aschke-Sonnenborn, U., Freemont, P. S., Janin, J., and Ruger, W. (2001) High resolution crystal structures of T4 phage β-glucosyltransferase: Induced fit and effect of substrate and metal binding. *J. Mol. Biol. 311*, 569−577.
(28) Patnaik, D., Chin, H. G., Esteve, P. O., Benner, J., Jacobsen, S. E., and Pradhan, S. (2004) Substrate specificity and kinetic mechanism of mammalian G9a histone H3 methyltransferase. *J. Biol. Chem. 279*, 53248−53258.
(29) Bacolla, A., Pradhan, S., Roberts, R. J., and Wells, R. D. (1999) Recombinant human DNA (cytosine-5) methyltransferase. II. Steady-state kinetics reveal allosteric activation by methylated DNA. *J. Biol. Chem. 274*, 33011−33019.
(30) Wu, J. C., and Santi, D. V. (1987) Kinetic and catalytic mechanism of HhaI methyltransferase. *J. Biol. Chem. 262*, 4778−4786.
(31) Lindstrom, W. M. Jr., Malygin, E. G., Ovechkina, L. G., Zinoviev, V. V., and Reich, N. O. (2003) Functional analysis of BamHI DNA cytosine-N4 methyltransferase. *J. Mol. Biol. 325*, 711−720.
(32) McKenna, E. S., and Roberts, C. W. (2009) Epigenetics and cancer without genomic instability. *Cell Cycle 8*, 23−26.
(33) Kinney, S. R., and Pradhan, S. (2011) Regulation of expression and activity of DNA (cytosine-5) methyltransferases in mammalian cells. *Prog. Mol. Biol. Transl. Sci. 101*, 311−333.
(34) Gama-Sosa, M. A., Slagel, V. A., Trewyn, R. W., Oxenhandler, R., Kuo, K. C., Gehrke, C. W., and Ehrlich, M. (1983) The 5-methylcytosine content of DNA from human tumors. *Nucleic Acids Res. 11*, 6883−6894.
(35) Kurzeck, J. (2002) Functional and Structural Research on the β-Glucosyltransferase of Bacteriophage T4. Ph.D. Dissertation, Ruhr-Universität Bochum, Bochum, Germany.
(36) Cerutti, P. A. (1976) in *DNA base damage induced by ionizing radiation* (Wang, S. Y., Ed.) Vol. *II*, Academic Press, New York.
(37) Hollstein, M. C., Brooks, P., Linn, S., and Ames, B. N. (1984) Hydroxymethyluracil DNA glycosylase in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A. 81*, 4003−4007.