

Supplemental information to Goyal et al.

Detailed Experimental Procedures

Enzymes and oligonucleotides

Wild type murine *dnmt1* gene was cloned in pFastBacHTc vector in frame with a His₆-tag. The Dnmt1 was purified from Sf21 cells basically as described (1,2). The Dnmt1 preparations were >90% pure as determined from Coomassie stained SDS gels (Suppl. Fig. 1). Oligonucleotides were purchased from Thermo Hybaid in HPLC purified form.

Generation of hemimethylated substrates

Two different PCR products were amplified from lambda DNA using two pairs of primers (sub1f: 5' P-AAC TGA TGC AGG ATA TCC AGC AGG AAA CAC-3', sub1r: 5'- CTC AAA CAT TTA CTA CAT TTA TCC TTA ATA -3', sub2f: 5' P-GAA GGA CAA CCT GAA GTC CAG GCA GTTG -3'. sub2r: 5'- GTG TAT GAC CAC CAG AGC CTT TTG TTC TGC- 3' where P denotes for the presence of a phosphate group). Substrate 1 comprises 634 bps containing 54 CG sites. The length of substrate 2 was 566 bps. It contains 44 CG sites and 8 CCGG sites. These PCR products were purified and used either directly as unmethylated substrates or converted into hemimethylated substrates as described (2,3). Briefly, for the generation of hemimethylated substrates, the PCR products were methylated either at CCGG sites by M.HpaII (NEB) or at all CG sites by M.SssI (kindly provided by Dr. M. Roth). Methylation was carried out using 10-20 U of enzyme, 5 µg of DNA and 2 mM AdoMet in 50 µl of buffer recommended by NEB for the respective enzyme for 1 h at 37°C. Completion of the methylation reaction was checked by restriction protection assay using several restriction enzymes that are sensitive towards CG methylation. Then the DNA was purified PCR purification columns (Macherey & Nagel) and 5 µg of DNA were treated with 30 units of λ exonuclease which specifically digests the 5'-phosphorylated strand. Digestion with λ exonuclease was judged for completion on a native polyacrylamide gel, in which no double stranded PCR

product was detectable after digestion. Subsequently, the methylated single strand was used as template for a primer extension reaction where the ssDNA was incubated with forward primer, dNTPs and Taq polymerase for 3 min at 95 °C, 30 s at 65 °C, 5 min at 72 °C, 1 min at 90 °C, 30 s at 62 °C, 5 min at 72 °C, 2 min at 88 °C, 30 s at 60 °C and 15 min at 72 °C. Hemimethylated DNA was purified using PCR purification columns from (Macherey & Nagel) and quantified by using an extinction coefficient of $\epsilon_{260 \text{ nm}} = 20 \text{ cm}^{-1} \text{ Lg}^{-1}$.

Sequences of the hemimethylated substrates

Substrate 1:

CTCAGGCATTTGCTGCATTTATCCTTCGTGATATCGGACGTTGGCTGGTCATATTCATCCCGCGA
CAGCCCGGACCGCTATAACCGCACTCGTACCGCGATAGGTCCAGGTGCAGGTGTTGGCCAGC
ATGATACCGTCCCGGAAAAACAGCGCCATCCGTTTCCGTCCGGCGTGGACAGTACAAAGGAGGC
ACTCACCGCGCTCAGTTCGCTGCACTGCTCAATGCGCCAGCGGCTGATCACCTCCTGCTCCGG
ATCGGCGTAACTGTTTCCCGTTGACGAAGTTCACCGCATCCAGAAAACGGGCGTAAACCTTACCG
CCGGACCACCGTTCTGCGACCAGACTCTGCATATCTTCCCGCCATCCCGGTGACCATACCGTA
CAGGTTAGAAACCGTCAGCGTGGGGCGCGTACTGGTGCCTTTGCCATTCAGTTCAAACCGCT
CCCCTGAATGGGATACCGGCTGATACTGTCGCCCCTGCCAGGTGACCGGCTCACCTTTTTCGTTC
TGCTCATTACAGAAAAATAACCGTTCTCCACCGACCTCTGTCAGGTCGATTCCCAGAGCACC
ACCGCTGGCCGACTGCTCCGCACGGGTGCATTTCATTCAGTGTTCCTGCGGGATATCCTGCATCA
GTT

Methylation data could not be obtained for the first CG site of the substrate.

Substrate 2:

TGTATGACCGACCAGAGCGTTTTGTTCTGCAGCTGGTCTGTGGTGCTGTCGGCGTTCATC
CTGCGCATCCGGATATTAAACCGGGCGCGGCGGCAGGTTACCCATCACCACCGAGGCCAG
ATACTGCGAGGTGGTTTTGCCCTTAATGGTGTGATGTCTTTTTCGTCACCCAGCCACCGT
TACCGTTGTATCTGAACCAGCAGGCGGACTTCCGACCGGATTCCTGTACCCCTTTGAGGTG
GTTTCCACCAGTGCCTGTACACCGAAGGTAAAGCGCAGACCGGTCGATGTTTGCAGACCGT
AATGGTCGGGTGATCGGCGTGTTCATATTTCACTTCCGTTACCAGCACCGTCTCGGAGC
CGGAGGATTCAAATCCCTCCGGCGGAGTCTGCTCCTGCTCACCAGCCCGGAACACCACC
GTGACACCGGATATGTTGGTATTCCCCTCAGTGTCCAGCACCGGCGTACTGTTTCAGCAG

CACGCTTTTTAAGCCATCCACCGGACCTTCAATCGGCCCTTCGCTGATGGCATCGATCA
CACTCAGCAACTGCGTGGACTTCAGGTTGTCCTTC

Methylation of oligonucleotide substrates

DNA methylation activity of the purified proteins was measured as described (4) by the incorporation of tritiated methyl groups from labeled [methyl-3H]-AdoMet into the biotinylated oligonucleotide substrate which contains a single CG site in unmethylated (CG30um: 5'- Bt-TTG CAC TCT CCT CCC GGA AGT CCC AGC TTC-3'/5'- GAA GCT GGG ACT TCC GGG AGG AGA GTG CAA-3', where Bt denotes a Biotin group) or hemimethylated form (CG30hm: 5'- Bt-TTG CAC TCT CCT CCC GGA AGT CCC AGC TTC-3'/5'- GAA GCT GGG ACT TC^mC GGG AGG AGA GTG CAA-3'). The methylation reactions were carried out at concentrations of 1 μ M DNA and 0.76 μ M of labeled AdoMet (3048 GBq/mmol, Perkin Elmer Life Sciences) in methylation buffer (20 mM HEPES, pH 7.0, 1 mM EDTA) at 37 °C using enzyme concentrations of 100 nM. At each time point 2 μ l of the reaction mixture were removed and analyzed as described (4). In some experiments fully methylated CG30fm was added as an allosteric activator (CG30fm: 5'- Bt-TTG CAC TCT CCT CC^mC GGA AGT CCC AGC TTC-3'/5'- GAA GCT GGG ACT TC^mC GGG AGG AGA GTG CAA-3').

Methylation of long hemimethylated DNA

For methylation 3 nM DNA, 5 nM Dnmt1, 1 mM AdoMet were incubated in methylation buffer (20 mM HEPES, pH 7.0, 1 mM EDTA, 30 mM NaCl, 7% glycerol) and 25 μ g/ml BSA at 37°C for 10 min. The methylation reaction was stopped after 10 min by freezing the reaction mixture in liquid nitrogen. 5 μ l buffer (10 mM Tris-HCl, 10 mM MgCl₂, 50 mM NaCl and 1 mM DTT) and 7.5 μ l (20 mg/ml) of Proteinase K (MBI Fermentas) was added. After incubation for 30 min, the reaction was purified over PCR purification columns (Macherey & Nagel) and eluted in 100 μ l water. Subsequently, the DNA was

subjected to bisulfite conversion performed as described (5-9) using the following primers.

Substrate 1 upper strand: 5'-AAT TGA TGT AGG ATA TTT AGT AGG AAA TAT-3' and 5'-CTC AAA CAT TTA CTA CAT TTA TCC TTA ATA-3' (annealing temperature: 58°C)

Substrate 1 lower strand: 5'-TTT AGG TAT TTG TTG TAT TTA TTT-3' and 5'-AAC TAA TAC AAA ATA TCC AAC AAA-3' (annealing temperature: 55°C)

Substrate 2 upper strand: 5'-GAA GGA TAA TTT GAA GTT TAG GTA GTT GT-3' and 5'-TAT ATA ACC ACC AAA CCC TTT TAT TCT A-3' (annealing temperature at 62°C)

Substrate 2 lower strand: 5'-GAT TAG AGT TTT TTG TTT TGT AGT-3' and 5'-CCT AAA ATC CAA ACA ATT ACT AAA TAT A-3' (annealing temperature 58°C)

PCR was carried out with the following program: 1x 4 min 95 °C, 35 x (20 s 95 °C, 30 s at annealing temperature and 1 min at 72 °C) 1x 10 min 72 °C. The PCR products were purified using PCR purification columns from Macherey & Nagel and cloned using the TOPO-TA kit (Invitrogen) as recommended by the supplier.

Plasmid DNA was purified from positive clones and sequences using T7, M13 forward and/or M13reverse primers. At least 25 clones were sequenced for each experiment; regions with incomplete conversion were disregarded. The data were aligned and analyzed with respect to the methylation frequencies at all CG sites. The level of methylation by M.HpaII and M.SssI enzymes was checked by the bisulfite protocol and found to be 87 % and 99 %. With M.HpaII there was 0.8% erroneous methylation at non CCGG sites.

Linear diffusion analysis of Dnmt1 movement

In order to describe the distribution of numbers of methyl groups incorporated in a processive reaction, we apply the theory of linear diffusion. Since there is no energy input for the movement of Dnmt1 along the DNA, it follows a random walk. For the reason of simplicity we will consider the movement of the enzyme from one CG site to the next as one diffusional step. In a random walk the number of steps required to reach the next target site equals the square of the distance to be traveled. In our experiment this distance is increasing with the increase of the length of the methylated stretch. We will consider the movement of the enzyme from one CG site to the next as one diffusional step, ignoring that the sites are at difference distances to each other. If the enzyme is located at the edge of a stretch of n fully methylated sites, the average distance d to the next hemimethylated CG site is given by $d_{GC}=(n+1)/2$ CG to CG steps. If we consider the average number of base pairs between each site on our substrate ($d_{BP,avg}=10.6$), the distance to the next available hemimethylated CG site can be also expressed in base pairs steps $d_{BP}=(n+1) \times d_{BP,avg}/2$. In a random walk, d^2 steps are required to travel the distance of d . If the enzyme has the probability P_{dif} to diffuse one step instead of dissociating from the DNA, the probability P of the enzyme to reach the next hemimethylated site is given by (10):

$$P = \exp(-d^2/P_{dif})$$

Therefore the probability to extend an existing stretch of methylated sites decreases with n .

Simulation of the propagation of the methylation information

We assume at generation 0 two populations of DNA molecules are present having 0 and 100% methylation at one particular CG site. We assume an average level of keeping the methylation mark of 95% and a preference for methylation at hemimethylated sites of 20-fold. Then, the probability of an unmethylated site of becoming methylated is 4.95% and the level of methylation after the next generation ($i+1$) can be calculated by:

$$P_{i+1} = P_i \times 0.9405 + 0.0495$$

where P_i is the methylation probability at generation i .

The discrimination between both populations after i generations (D_i) is given by $D_i = P_i(\text{met})/P_i(\text{unmet})$, where $P_i(\text{met})$ is the probability of the methylated pool being methylated in generation i , and $P_i(\text{unmet})$ is the probability of the unmethylated pool to be methylated at that time. If k independent methylation sites are considered, the probability of k sites being methylated is given by $P_i(\text{met})^k$ and the discrimination between the both pools $D(k)$ is given by $D(k) = d^k$.

References cited in Supplemental part

1. Fatemi, M., Hermann, A., Pradhan, S. and Jeltsch, A. (2001) The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA. *J Mol Biol*, **309**, 1189-1199.
2. Hermann, A., Goyal, R. and Jeltsch, A. (2004) The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J Biol Chem*, **279**, 48350-48359.
3. Thomas, E., Pingoud, A. and Friedhoff, P. (2002) An efficient method for the preparation of long heteroduplex DNA as substrate for mismatch repair by the Escherichia coli MthLS system. *Biol Chem*, **383**, 1459-1462.
4. Roth, M. and Jeltsch, A. (2000) Biotin-avidin microplate assay for the quantitative analysis of enzymatic methylation of DNA by DNA methyltransferases. *Biol Chem*, **381**, 269-272.
5. Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L. and Paul, C.L. (1992) A genomic sequencing protocol that yields a positive display of 5- methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. USA*, **89**, 1827-1831.
6. Hermann, A., Schmitt, S. and Jeltsch, A. (2003) The human Dnmt2 has residual DNA-(cytosine-C5) methyltransferase activity. *J Biol Chem*, **278**, 31717-31721.
7. Warnecke, P.M., Stirzaker, C., Song, J., Grunau, C., Melki, J.R. and Clark, S.J. (2002) Identification and resolution of artifacts in bisulfite sequencing. *Methods*, **27**, 101-107.
8. Grunau, C., Clark, S.J. and Rosenthal, A. (2001) Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res*, **29**, E65-65.
9. Clark, S.J., Harrison, J., Paul, C.L. and Frommer, M. (1994) High sensitivity mapping of methylated cytosines. *Nucleic Acids Res*, **22**, 2990-2997.
10. Ehbrecht, H.J., Pingoud, A., Urbanke, C., Maass, G. and Gualerzi, C. (1985) Linear diffusion of restriction endonucleases on DNA. *J Biol Chem*, **260**, 6160-6166.

Suppl. Table 1

experiment	substrate analyzed for methylation	# of hemimethylated CG sites		#of unmethylated CG sites	
		available (on all sequenced clones)	methylated	available (on all sequenced clones)	methylated
substrate 1 unmethylated (alone)	substrate 1	-	-	1953	85
substrate 1 hemimethylated (alone)	substrate 1	932	545	-	-
substrate 1 (hemimethylated and substrate 2 (unmethylated) in competition	substrate 2	-	-	871	44
	substrate 1	937	531	-	-
substrate 2 hemimethylated at CCGG sites (alone)	substrate 2	146	78	607	28

Suppl. Figures

Fig. 1. Purification of Dnmt1. The figure shows a Coomassie stained denaturing polyacrylamide gel loaded with the Dnmt1 preparation and a molecular weight maker.

