

Journal Pre-proof

Dissect the DNMT3A- and DNMT3B-mediated DNA co-methylation through a covalent complex approach

Linfeng Gao, Hiwot Anteneh, Jikui Song



PII: S0022-2836(19)30660-6

DOI: <https://doi.org/10.1016/j.jmb.2019.11.004>

Reference: YJMBI 66322

To appear in: *Journal of Molecular Biology*

Received Date: 13 August 2019

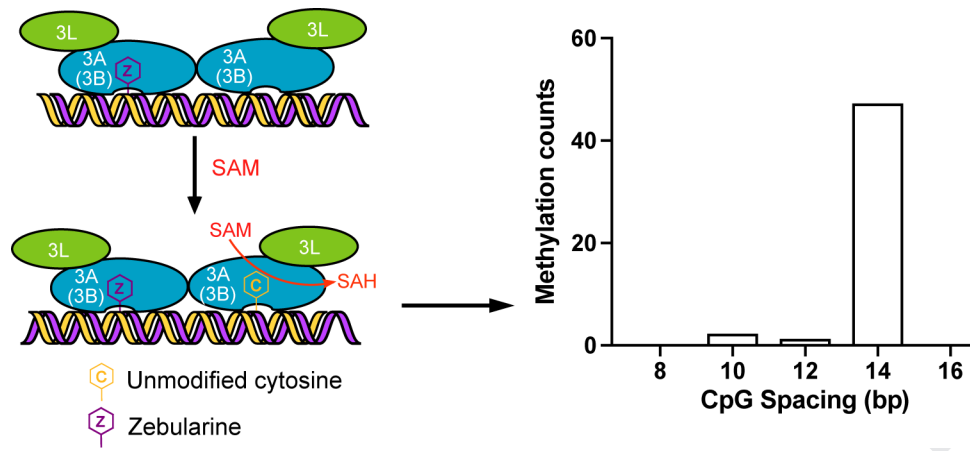
Revised Date: 1 November 2019

Accepted Date: 4 November 2019

Please cite this article as: L. Gao, H. Anteneh, J. Song, Dissect the DNMT3A- and DNMT3B-mediated DNA co-methylation through a covalent complex approach, *Journal of Molecular Biology*, <https://doi.org/10.1016/j.jmb.2019.11.004>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier Ltd.



Dissect the DNMT3A- and DNMT3B-mediated DNA co-methylation through a covalent complex approach

Linfeng Gao¹, Hiwot Anteneh², Jikui Song^{1,2,*}

¹Environmental Toxicology Graduate Program, University of California, Riverside, USA

²Department of Biochemistry, University of California, Riverside, USA

*Correspondence: jikui.song@ucr.edu

Abstract

DNA methylation plays a critical role in regulating gene expression, genomic stability and cell fate commitment. Mammalian DNA methylation, which mostly occurs in the context of CpG dinucleotide, is installed by two *do novo* DNA methyltransferases, DNMT3A and DNMT3B. Oligomerization of DNMT3A and DNMT3B permits both enzymes to co-methylate two CpG sites located on the same DNA substrates. However, how DNMT3A- and DNMT3B-mediated co-methylation contributes to the DNA methylation patterns remain unclear. Here we generated covalent enzyme-substrate complexes of DNMT3A and DNMT3B, and performed bisulfite sequencing-based single turnover methylation analysis on both complexes. Our results showed that both DNMT3A- and DNMT3B-mediated co-methylation preferentially gives rises to a methylation spacing of 14 base pairs, consistent with the previous structural observation for DNMT3A in complex with regulatory protein DNMT3L and CpG DNA. This study provides a novel method for mechanistic investigation of DNMT3A- and DNMT3B-mediated DNA co-methylation.

Introduction

DNA methylation is an important epigenetic mechanism essential for regulating diverse cellular processes, including transcriptional silencing of retrotransposons¹⁻³, genomic imprinting⁴ and X-chromosome inactivation^{5, 6}. In mammals, DNA methylation mainly occurs at the C-5 position of cytosine in the CpG dinucleotide context, accounting for ~70-80% of the CpG sites across the genome⁷. Establishment of DNA methylation is mainly achieved by two *de novo* DNA methyltransferases, DNMT3A and DNMT3B, during gametogenesis and postimplantation development⁸, which are further regulated by DNMT3-like protein (DNMT3L)⁹⁻¹¹. Specific methylation patterns across genome mediate distinct DNA-templated activities, serving as a hallmark in governing cell fate commitment. Dysregulation of DNMT3A- and DNMT3B-mediated *de novo* methylation has been associated with a variety of human diseases, such as cancer and neurological disorders¹². Interestingly, genome-wide methylation analysis has identified a modest enrichment of 8-10 base pair (bp) periodicity among the methylation sites of both mammalian and plant genomes¹³⁻¹⁶. However, the underlying mechanisms remain unclear.

To date, mechanistic understanding of the DNMT3-mediated DNA methylation has mainly come from DNMT3A^{17, 18}. The crystal structure of DNMT3A methyltransferase (MTase) domain in complex with the C-terminal domain of DNMT3L reveals a heterotetrameric fold in the order of DNMT3L-DNMT3A-DNMT3A-DNMT3L¹⁶. The presence of two active sites, separated by a distance of ~40 Å, has lent support to a DNA co-methylation model, in which DNMT3A can co-methylate two CpG sites across the DNA double strands separated by 8-10 base pairs (bps)¹⁶. Along the line, a

subsequent study has developed a hairpin bisulfite analysis-based approach to evaluate the DNMT3A-mediated CpG co-methylation¹⁹. In essence, the double-stranded DNA substrates were first subjected to methylation by mouse DNMT3A-DNMT3L complex. Subsequently, the methylated products were ligated with a hairpin loop at one end, followed by bisulfite conversion and cloning analysis. This approach, through conversion of the two opposite DNA strands into the same strand, permits the measurement of the correlated methylation between the two CpG sites located across the opposite strands of the same DNA substrate. Consistent with the structural modeling analysis, the hairpin bisulfite analysis indicated that the most correlated methylation events occur at the CpG sites on the two opposite strands, with spacing of 8-10 bp¹⁹. These observations therefore led to the proposition that the DNMT3A-mediated CpG co-methylation may contribute to the periodicity of CpG methylation in mammalian genomes^{16, 19}. On the other hand, this approach is challenged by the fact that it could not distinguish whether the two correlated methylation events arise from co-methylation or sequential methylation events (Fig. 1A). In fact, it has been argued that the 10-bp periodic distribution of the WW (W= A, T) and SS (SS= G, C) dinucleotides in nucleosomal DNA, which serves to accommodate the DNA wrapping around histone octamers²⁰, may dictate the periodicity of DNA methylation¹³.

To gain further understanding of the DNMT3A- and DNMT3B-mediated *de novo* methylation, we developed a “single-turnover” methylation assay, in which DNMT3A or DNMT3B is covalently linked to one target site of the substrates, thereby restricting the dissociation of DNMT3A or DNMT3B from the substrate before the next enzymatic attack. Our result unequivocally clarified 14 bp as the co-methylation spacing of

DNMT3A and DNMT3B, providing a basis for mechanistic understanding of DNMT3A- and DNMT3B-mediated CpG co-methylation.

Results

Through a mechanism-based cross-linking method, we have recently reported the crystal structures of the DNMT3A-DNMT3L tetramer covalently bound to DNA substrates in which the target cytosine in the CpG site is replaced by Zebularine (Z), a cytosine analogue²¹, providing mechanistic insight into the productive state of DNMT3A-substrate complex²². To elaborate how the oligomerization of DNMT3A and closely-related DNMT3B affects their co-methylation of CpG sites, we have compared the two crystal structures of the DNMT3A-DNMT3L-CpG DNA complexes from our previous study: One involves the DNMT3A-DNMT3L tetramer bound to two 10 bp-long DNA duplexes each containing one single central CpG/ZpG site, while the other involves the DNMT3A-DNMT3L tetramer bound to a 25-mer self-complimentary DNA substrate containing two CpG/ZpG sites, with the target Zebularines separated by 14 bps. Structural comparison of the two DNMT3A-DNMT3L-DNA complexes indicates a similar protein-DNA interface (Fig. 1B), suggesting that the substrate recognition modes of these complexes were not affected by their different crystal packing conditions. Notably, for the complex of the DNMT3A-DNMT3L tetramer bound to short DNA duplexes, the two DNA molecules do not align with each other to mimic the standard B-form conformation; instead, they are oriented with a slightly titled angle, mimicking the bending conformation of the longer substrate (Fig. 1B). These observations suggest that the 14-bp spacing between the two target sites of the longer substrate may reflect the co-methylation spacing of DNMT3A under the physiological condition.

To clarify how DNMT3B and DNMT3A may mediate DNA co-methylation in solution, we designed a single turnover methylation assay (Fig. 2A,B), based on the stable DNMT3A-DNMT3L-DNA or DNMT3B-DNMT3L-DNA complex formed prior to the methylation reaction (Fig. 2A). In essence, we first generated a covalent complex between the DNMT3A-DNMT3L or DNMT3B-DNMT3L tetramer and an unmethylated DNA substrate containing one single ZpG site as well as multiple CpG sites, with the Zebularine and cytosines separated by 8-, 10-, 12-, 14- and 16-bp, respectively. Next, the DNMT3A-DNMT3L-DNA or DNMT3B-DNMT3L-DNA complex was incubated with cofactor S-adenosyl-L-methionine (SAM), which triggers the methylation reaction with DNMT3A or DNMT3B staying bound to the DNA. Subsequently, DNMT3A or DNMT3B was removed from the DNA substrate through protease digestion, the target strand of the DNA product was amplified by PCR reaction, and the distances between the methylcytosine and the ZpG site was measured by bisulfite sequencing analysis (Fig. 2B and Tables S1-S4). In this study, we particularly focus on the DNA strand that is complementary to the ZpG-containing strand, given the fact that structural analysis of the DNMT3A-DNMT3L-DNA complex indicates an opposite directionality between the co-methylated CpG sites (Fig. 1B). In comparison with previous hairpin bisulfite approach, this newly developed approach offers the advantage to eliminate the effect of DNMT3A- and DNMT3B-mediated sequential methylation, if any.

Analysis of the resulting sequencing data reveals a >96% conversion rate for the CpG site complementary to the ZpG site (Fig. S1), which not only suggests that this site is protected from being methylated during the reaction, but also serves as an internal control to ensure high bisulfite conversion rate. In addition, the sequencing results

indicate a ~94% conversion rate for an external control, the unmethylated CpG DNA that was not subjected to methylation reaction (Fig. S1), which further validates the efficiency of bisulfite conversion.

Out of 157 and 150 sequences analyzed for the methylation products of DNMT3A-DNMT3L and DNMT3B-DNMT3L, respectively, we were able to identify 50 and 59 single-methylated sequences, respectively (Tables S1-S3). Among these, 47 of DNMT3A-DNMT3L-treated and 55 of DNMT3B-DNMT3L-treated sequences yield a methylation spacing of 14 bps (Fig. 2C,D and Tables S2-S3). In addition, we performed the single-turnover methylation assay for DNMT3A-DNMT3L on the DNA substrate in which the Zebularine within the sole ZpG site and the cytosines within the CpG sites are separated by 9-, 11-, 13-, 15- and 17-bp, respectively. Under the same reaction condition, 12 out 36 analyzed sequences were identified as single-methylated sequences (Tables S1 and S4). However, analysis of these sequences failed to reveal any of the CpG spacing from this substrate as dominant methylation event (Fig. 2E and Table S4). Together, these data suggest that two target sites with 14 bp spacing most likely represent the most favorable co-methylation substrate of DNMT3. One caveat of this study is that the DNMT3A- and DNMT3B-mediated DNA co-methylation was examined in the context of the catalytic domains, instead of full-length enzymes, which may lead to different co-methylation behaviors. It is also worth noting that, the minor population of methylation events, with methylation spacing other than 14 bp (Fig. 2C-E), likely reflect the fact that this covalent complex of the DNMT3A-DNMT3L or DNMT3B-DNMT3L tetramer with ZpG DNA remains reversible²³, leading to partial complex

dissociation and association during the enzymatic reaction, and therefore a possible enzymatic action that was not restricted by the original covalent complex.

Discussion

The previously proposed DNA co-methylation spacing model of the DNMT3A-DNMT3L tetramer highlights the importance of DNMT3A oligomerization in defining DNA methylation patterns^{16, 19}. This model fits with genome-wide methylation analysis, which has indicated a 10-bp of methylation periodicity of mammalian genomes^{13, 24, 25}. However, how such oligomerization of DNMT3A or DNMT3B contributes to the correlation of two neighboring methylation events remains controversial²⁰. This study, through single-turnover methylation assay, demonstrates 14 bp as the most favorable co-methylation spacing of DNMT3A and DNMT3B, consistent with the structural observation for the DNMT3A-DNMT3L-DNA complexes.

The apparent discrepancy between the 14-bp co-methylation spacing determined here and the previous observation of 8-10 bp spacing likely arises from the distinct enzymatic behaviors captured by these studies. In particular, 14-bp co-methylation spacing arises from a scenario in which the two active sites of the DNMT3A or DNMT3B dimer undergo the productive catalysis simultaneously (Fig. 3A). Structurally, the co-methylation event observed in this study is accompanied by significant bending of DNA substrates, which might not be energetically favorable in solution when compared with the sequential methylation events along the substrate with the standard B-form conformation. On the other hand, the 8-10 bp methylation spacing may arise from an energetically distinct scenario (Fig. 3B): It is likely that the DNMT3A or DNMT3B dimer

presents one of the monomers to methylate one CpG site first, with the other monomer binds to the DNA in a non-productive manner; upon the completion of the first methylation event, the DNMT3A or DNMT3B dimer presents the second monomer to attack another CpG site that is in close proximity. During this transition, DNMT3A/3B and DNA substrate may undergo a conformational adjustment without being fully dissociated. This latter scenario also permits the conformation of the DNA substrate to stay in a B-form conformation, which is energetically more favorable. A detailed characterization of these enzymatic behaviors awaits further investigation.

How the 14-bp spacing mechanism described here may contribute to the DNA methylation patterns in cells, in particular in the context of the chromatin environment, remains to be elucidated. Nevertheless, it is apparent that the unique architecture of DNMT3A oligomer provides a basis for correlative methylation of two cytosines within a relatively short distance in general, through either co-methylation or a sequential enzymatic action.

Acknowledgments

This work was supported by NIH (1R35GM119721 to J.S.). We would like to thank Dr. Albert Jeltsch for insightful discussion.

Author Contributions

L.G. and H.A. performed experiments. J.S. conceived and designed the study, and prepared the manuscript.

Author Information

The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.S. (jikui.song@ucr.edu)

References

1. Bourc'his, D. & Bestor, T.H. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* **431**, 96-99 (2004).
2. Holliday, R. & Pugh, J.E. DNA modification mechanisms and gene activity during development. *Science* **187**, 226-232 (1975).
3. Walsh, C.P., Chaillet, J.R. & Bestor, T.H. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat Genet* **20**, 116-117 (1998).
4. Li, E., Beard, C. & Jaenisch, R. Role for DNA methylation in genomic imprinting. *Nature* **366**, 362-365 (1993).
5. Panning, B. & Jaenisch, R. RNA and the epigenetic regulation of X chromosome inactivation. *Cell* **93**, 305-308 (1998).
6. Riggs, A.D. X inactivation, differentiation, and DNA methylation. *Cytogenetics and cell genetics* **14**, 9-25 (1975).
7. Ehrlich, M. *et al.* Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res* **10**, 2709-2721 (1982).
8. Okano, M., Bell, D.W., Haber, D.A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-257 (1999).
9. Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B. & Bestor, T.H. Dnmt3L and the establishment of maternal genomic imprints. *Science* **294**, 2536-2539 (2001).
10. Chedin, F., Lieber, M.R. & Hsieh, C.L. The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proc Natl Acad Sci U S A* **99**, 16916-16921 (2002).
11. Hata, K., Okano, M., Lei, H. & Li, E. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129**, 1983-1993 (2002).
12. Robertson, K.D. DNA methylation and human disease. *Nat Rev Genet* **6**, 597-610 (2005).
13. Chodavarapu, R.K. *et al.* Relationship between nucleosome positioning and DNA methylation. *Nature* **466**, 388-392 (2010).
14. Cokus, S.J. *et al.* Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* **452**, 215-219 (2008).
15. Glass, J.L., Fazzari, M.J., Ferguson-Smith, A.C. & Grealley, J.M. CG dinucleotide periodicities recognized by the Dnmt3a-Dnmt3L complex are distinctive at retroelements and imprinted domains. *Mammalian genome : official journal of the International Mammalian Genome Society* **20**, 633-643 (2009).
16. Jia, D., Jurkowska, R.Z., Zhang, X., Jeltsch, A. & Cheng, X. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* **449**, 248-251 (2007).
17. Cheng, X. & Blumenthal, R.M. Mammalian DNA methyltransferases: a structural perspective. *Structure* **16**, 341-350 (2008).
18. Ren, W., Gao, L. & Song, J. Structural Basis of DNMT1 and DNMT3A-Mediated DNA Methylation. *Genes* **9** (2018).
19. Jurkowska, R.Z. *et al.* Formation of nucleoprotein filaments by mammalian DNA methyltransferase Dnmt3a in complex with regulator Dnmt3L. *Nucleic Acids Res* **36**, 6656-6663 (2008).
20. Widom, J. Role of DNA sequence in nucleosome stability and dynamics. *Quarterly reviews of biophysics* **34**, 269-324 (2001).
21. Zhou, L. *et al.* Zebularine: a novel DNA methylation inhibitor that forms a covalent complex with DNA methyltransferases. *J Mol Biol* **321**, 591-599 (2002).
22. Zhang, Z.M. *et al.* Structural basis for DNMT3A-mediated de novo DNA methylation. *Nature* **554**, 387-391 (2018).

23. Champion, C. *et al.* Mechanistic insights on the inhibition of c5 DNA methyltransferases by zebularine. *PLoS one* **5**, e12388 (2010).
24. Zhang, X. *et al.* Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell* **126**, 1189-1201 (2006).
25. Zhang, Y. *et al.* DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution. *PLoS genetics* **5**, e1000438 (2009).

Figure Legends

Figure 1. Co-methylation model of DNMT3A- and DNMT3B. (A) Schematic model of co-methylation vs sequential methylation. (B) Structural overlay of the DNMT3A-DNMT3L tetramer bound to two short DNA duplexes containing single CpG site (PDB 6F57) and the DNMT3A-DNMT3L tetramer bound to one long DNA substrate containing two CpG sites (PDB 5YX2). Both structures suggest co-methylation spacing of 14-bp.

Figure 2. Schematic view of single-turnover methylation assay for measurement of co-methylation spacing. (A) Schematics for generation of the covalent complex of DNMT3A and DNMT3B with Zebularine-containing DNA. (B) Workflow for the single-turnover methylation assay. The Zebularine and unmodified cytosine are denoted as “Z” and “C”, respectively. The target strand contains a 5’ overhang for the purpose of strand-specific PCR amplification. (C) The relative abundance of indicated even-numbered CpG co-methylation spacing mediated by the DNMT3A-DNMT3L complex. The average and standard deviation were derived from analysis of three independent batches of sequencing data (See Table S2). (D) The relative abundance of indicated even-numbered CpG co-methylation spacing mediated by the DNMT3B-DNMT3L complex. The average and standard deviation were derived from analysis of three independent batches of sequencing data (See Table S3). (E) The relative abundance of

indicated odd-numbered CpG co-methylation spacing mediated by the DNMT3A-DNMT3L complex. The average and standard deviation were derived from analysis of two independent batches of sequencing data (See Table S4).

Figure 3. Model for DNMT3A- and DNMT3B-mediated DNA co-methylation. (A) The 14-bp co-methylation spacing arises from the two concurrent methylation events of DNMT3A and DNMT3B. The CpG sites are denoted by letters 'C' and 'G'. The DNMT3A or DNMT3B monomers are represented by oval spheres. The flipped 'C' represents the insertion of target cytosine into the active site of each monomer, corresponding to a productive methylation state. (B) DNMT3A/3B may catalyze the methylation of one CpG site first, and then transit into the catalytic state for methylating the next CpG site. At each catalytic stage, only one monomer undergoes productive methylation, while the other monomer adopts a non-productive state, which does not involve base flipping of target cytosine. During the transition of the methylation events, the DNMT3A/3B dimer may remain associated with the DNA substrate.

Methods

Protein expression and purification. The MTase domains of human DNMT3A or DNMT3B (DNMT3A MTase: residues 562–853; DNMT3B MTase: residues 628–912) or was co-expressed with residues 178–386 of human DNMT3L on a modified pRSFDuet-1 vector (Novagen), which contains a hexahistidine (His₆) and SUMO tag preceding the DNMT3B or DNMT3A sequence. The expression and purification of the DNMT3A-

Preparation of covalent complexes

(DNA^{spacing1}) (Upper strand: 5'-GCATGZGTTCTAACGCGCGCGCGTGAAGGAAGGAAGG-3'; Lower strand: 5'-GGTGGTGGTGGTGATCCTTCCTTCCTTCACGCGCGCGCGTTAGAACGCATGCGTG
GAGATGGAGGAGG-3'; Z: Zebularine) or a 38-mer/69-mer hybrid DNA duplex
(DNA^{spacing2}) (Upper strand: 5'-GCATGXGTTCTAA

GCGCGCGCGCGTGAAGGAAGGAAGG-3'; Lower strand: 5'-

GGTGGTGGTGGTGATCCTTCCTTCCTTCACGCGCGCGCGCTTAGAACGCATGCGT

GGAGATGGAGGAGG-3'; Z: Zebularine) in a 2:1 molar ration and incubated under the

condition of 20 mM Tris-HCl (pH 8), 20% glycerol, and 40 mM DTT at room

temperature. The 5'-overhang sequences in both DNA substrates serve as templates for

strand-specific PCR amplification as well as internal control for bisulfite conversion.

Subsequently, the DNMT3A/3B-DNMT3L-DNA^{spacing1} or DNMT3A/3B-DNMT3L-

DNA^{spacing2} covalent complexes were purified sequentially through ion exchange

chromatography on a HiTrap Q XL column (GE Healthcare) and size-exclusion

chromatography on a HiLoad 16/600 Superdex 200 pg column.

Single-turnover methylation assay. For the enzymatic assays, the DNMT3A/3B-

DNMT3L-DNA^{spacing1} or DNMT3A/3B-DNMT3L-DNA^{spacing2} covalent complexes were

adjusted to 1 μ M and reacted in a buffer containing 32 μ M AdoMet, 50 mM Tris-HCl, pH

8.0, 0.05% β -mercaptoethanol, 5% glycerol and 200 μ g/mL BSA. The reaction was

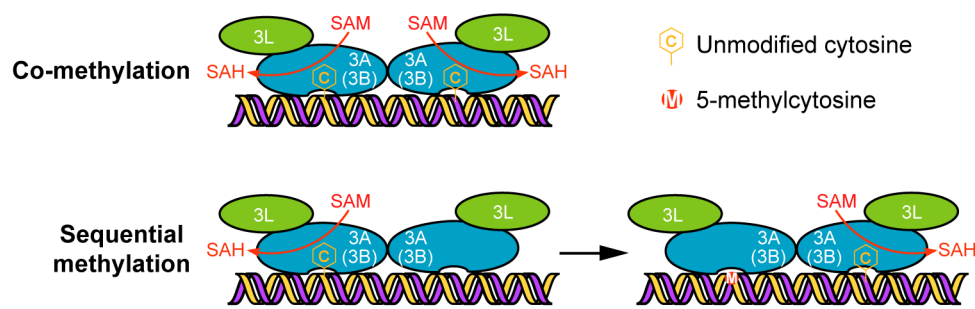
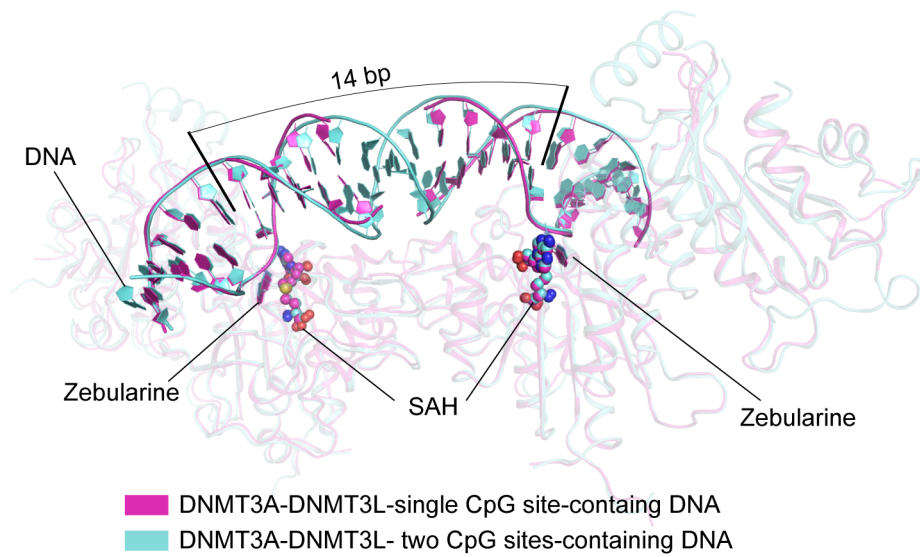
performed at 37°C for 1h. The reaction products were digested with protease K at 42°C

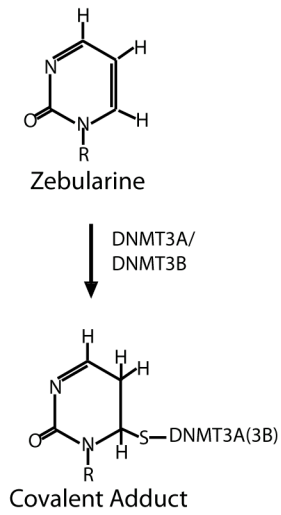
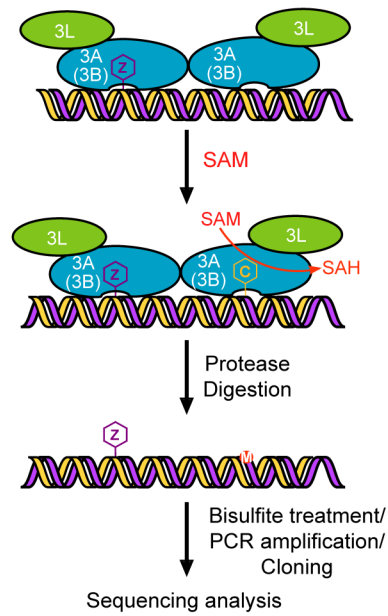
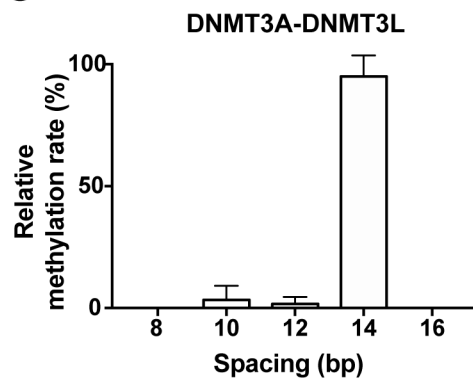
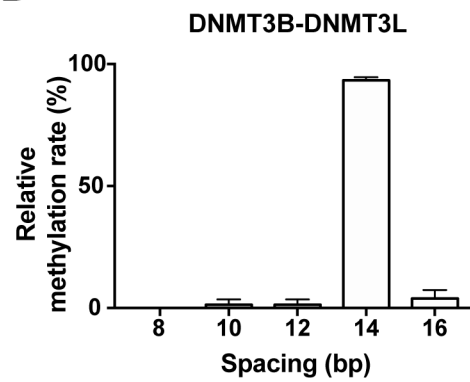
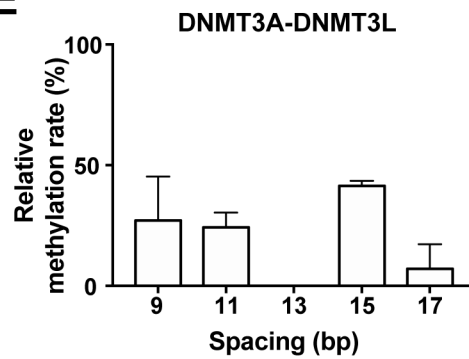
for 2h, purified by agarose gel electrophoresis using GeneJET gel extraction kit

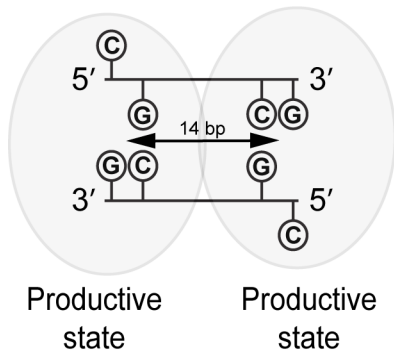
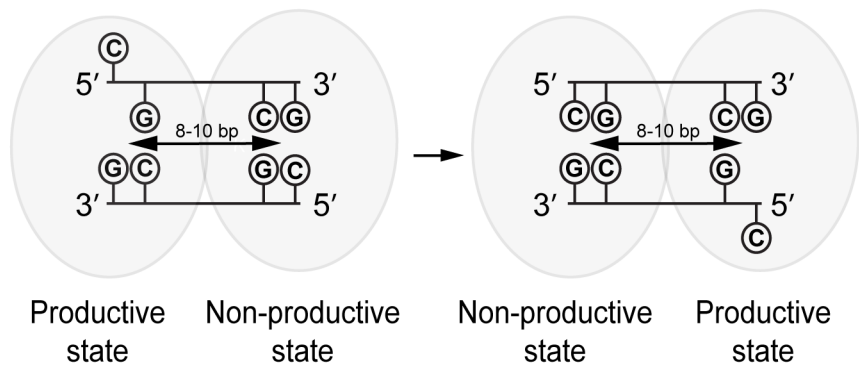
(Thermo Scientific). 100 ng of purified DNA was subjected to bisulfite conversion using

an EZ DNA Methylation Gold Kit (Zymo Research), followed by PCR amplification using

primers: 5'-CCACCACCACCACTA-3' and 5'-GGAGGAGGTAGAGGTG-3'. The PCR products were cloned into pCR4-TOPO vector (Invitrogen) for sequencing analysis.

A**B**

A**B****C****D****E**

A**B**

Highlights

- Formation of covalent complexes between DNMT3A/3B and DNA
- The method performs enzymatic assays with covalent complexes of DNMT3A/3B
- Accurate measurement of single turnover methylation event was developed
- DNMT3A/3B co-methylates CpG DNA with preferred 14-bp spacing