Analyst

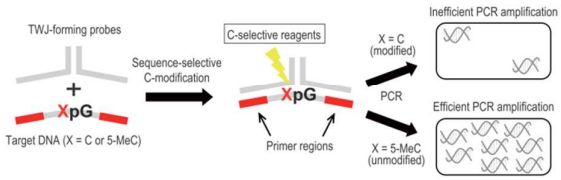
Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/analyst

A table of contents entry

Sequence-selective modification of a single cytosine using junction-forming DNA probes was applied for the quantitative-PCR-based detection of methylated cytosine.



Detection of single methylated cytosine using junction-forming DNA probes

Kenta Takanashi and Teru Kato†

Graduate School of Bionics, Computer and Media Sciences, Tokyo University of Technology, 1404-1 Katakura, Hachioji, Tokyo 192-0982, Japan

† To whom correspondence should be addressed.

E-mail: kato@stf.teu.ac.jp

Abstract

DNA methylation is an epigenetic mechanism for transcriptional regulation. The methylation process controls cellular differentiation and is defective in many diseases including cancer. Therefore, the development of a simple method for analysing cytosine methylation in a target gene is required. Here we report a conceptually new method for sequence-selective chemical modification of a single cytosine in single-stranded DNA (ssDNA) using two DNA probes to form a DNA three-way junction with the ssDNA. The method was successfully used in a simple quantitative polymerase-chain-reaction-based assay for discrimination of a single methylated cytosine.

5-Methylcytosine (M) is generated by enzymatic methylation of cytosine (C) and plays an important role in epigenetic gene regulation.[1] In mammals, M frequently appears in CpG dinucleotides of genomic DNA and is a promising biomarker for the diagnosis of cancer and other diseases.[2,3] Currently, the most commonly used method for the evaluation of the CpG methylation level is bisulfite sequencing.[4,5] However, this method requires cumbersome, tedious experimental procedures such as cloning the sequences of bisulfite-treated DNA and sequencing numerous clones that contain numerous non-CpG cytosines.[6,7] Therefore, a simple method for methylation analysis that targets only 'methylatable' cytosine (*e.g.* cytosine of CpG) in a long DNA molecule is highly desirable. A C/M-selective chemical reaction[8-11] and a technique to guide the reaction reagents to the methylatable cytosine[8,9] might constitute such a method. To guide the reaction reagents, Okamoto *et al.* developed an artificial DNA probe covalently tethered to a bipyridine ligand.[9] When the probe was hybridized with a long target DNA containing M, the ligand positioned opposite M induced sequence-selective and M-selective osmium complexation with M. However, preparation of DNA probes tethered to ligands or reactive species often requires post-synthetic modification of the probes or custom synthesis of phosphoramidite derivatives of the nucleotide that bears those ligands or reactive species, which complicates the method.[9,12,13] To guide reaction reagents using only DNA probes prepared by a standard synthetic method[8], we used two DNA probes that were partially complementary to each other and to the target DNA; the probes thus hybridized with the target and folded into a fully matched three-way-junction (TWJ) structure (Fig. 1). A TWJ is a secondary structure in which three stems are connected at the branching point; TWJs deserve attention as building blocks for applications in nanotechnology and molecular recognition.[14-17] Coll *et al.* and we have reported small molecules selectively bound to a nanoscale cavity at the branching point of a fully matched TWJ. [16,17] We have also previously confirmed that nucleotide bases at the

branching point of the fully matched TWJ were hyper-reactive against reagents such as osmium tetroxide, probably because of the exposure of these bases surrounding the cavity to such reagents; however, the bases in the three stems were fairly protected from the reagents by base-pair stacking.[14,17] This unique property of TWJ encouraged us to demonstrate modification of a single C in target DNA using C-selective reagents, *i.e.* bisulfite and aminooxy derivatives, which have been reported to add to C in ssDNA under mild conditions, whereas M was resistant to the addition and remained unmodified.[11] When a TWJ is formed by the hybridization of the target with two DNA probes whose sequences have been designed to place C at the branching point of TWJ (Fig. 1), C should be sequence-selectively modified by the bisulfite and aminooxy reagents. To detect the modified C, we monitored the obstruction of *Taq* DNA polymerase-catalysed primer extension and PCR using the target containing the modified C as a template because the adduct of a cytosine residue with a bisulfite and an aminooxy reagent contained in ssDNA has been reported to obstruct polymerase-catalysed extension of primer DNA annealed to the ssDNA.[18]

To prove the concept, we prepared a 56mer target DNA1 that contains an XpG dinucleotide (X = C or M) and 11 other Cs in its sequence (Supplementary Table 1). Two DNA probes (FM-TWJ probes 1 and 2) to place XpG at the branching point of the TWJ were also prepared (Fig. 2A). DNA1 (0.5 pmol) was incubated in the absence or presence of FM-TWJ probes 1 and 2 (2.0 pmol) with 1.0 M sodium bisulfite (pH 5.0) and 1.0 M *O*-carboxymethylhydroxylamine (CMH, pH 5.0), which is a highly water-soluble aminooxy reagent, for 5 h at 25°C. After DNA1 was recovered by ethanol precipitation, an excess amount of 5′-FAM-labeled 18mer primer complementary to the 3′-region of DNA1 was annealed to DNA1 and extended by *Taq* DNA polymerase under thermal cycling conditions to suppress any interference with primer annealing by residual DNA probes (Fig. 2B). The products were

analysed by polyacrylamide gel electrophoresis under denaturing conditions (Fig. 2C). When non-incubated DNA1 (X = C or M) was used as a template for primer extension, most of the 5′-FAM-labeled primer extended fully (upper band of lanes 1 and 2, Fig. 2C). The extension of the primer annealed to DNA1 (X = C or M) that had been incubated in the absence of FM-TWJ probes did not yield the fully extended product (lanes 3 and 4), probably because of failure of primer annealing and blockage of primer extension caused by modification of several Cs contained in DNA1. When DNA1 (X = C) incubated in the presence of FM-TWJ probes was used as a template (lane 5), a distinct band caused by blockage of the primer extension was observed near the G29 band of the Maxam–Gilbert A+G marker, opposite X that had been placed at the branching point of the TWJ (Figs. 2A, B). In addition, a small amount of the fully extended product was observed in lane 5. In contrast, extension of the primer annealed to DNA1 (X = M) incubated in the presence of FM-TWJ probes was not blocked near the G29 band of the marker (lane 6) and yielded a substantial amount of the fully extended product. These results suggest that the two FM-TWJ probes that formed the TWJ with 56mer DNA1 successfully guided addition of bisulfite and CMH to C placed at the branching point and suppressed addition to other Cs in DNA1 by placing them in the stems of the TWJ. Thus, one of the 12 Cs in the 56mer target DNA1 was chemically modified in a sequence-selective and C-selective manner by using the FM-TWJ probes, and this modification could be monitored by primer-extension analysis catalysed by *Taq* DNA polymerase.

To construct a prototype for simple methylation analysis based on this sequence-selective C-modification, we performed a model experiment of single C/M discrimination using quantitative PCR. A 56mer target DNA1 (1.5 fmol, X = C or M) was incubated with sodium bisulfite and CMH in the absence or presence of FM-TWJ probes 1 and 2 (6 pmol each) for 3 h, recovered by ethanol precipitation and analysed by quantitative PCR. The Ct values of the

incubated and non-incubated samples were acquired using primers 1 and 2 (Fig. 3A). The amount of unmodified DNA1 that can serve as a template for PCR amplification was determined from its Ct values using a standard curve of non-incubated DNA1 (Supplementary Fig. 1). The unmodified DNA1 rate (%) in the samples was calculated by dividing the amount of unmodified DNA1 by that in the non-incubated samples and was normalized to the rate in the non-incubated samples, which was defined as 100%. After incubation in the presence of FM-TWJ probes, approximately 38% of DNA1 (X = M) remained unmodified, whereas only 1.5% of DNA1 (X = C) remained unmodified (Fig. 3B). In the presence of the FM-TWJ probes, an approximately 25-fold increase in the unmodified DNA1 (X = M) rate relative to that of DNA1 (X = C) was observed. In the absence of the FM-TWJ probes, the Ct values of the DNA1 (X = C or M) were approximately 24 (Fig. 3A), indicating that more than 99.9% of DNA1 was modified. These results indicate that DNA1 (X = M) was fairly protected from modification at the branching point of TWJ, whereas DNA1 (X = C) was extensively modified at the branching point of TWJ, resulting in a very low unmodified DNA rate. Thus, the substantial increase in the unmodified target DNA (X = M) rate relative to that of the target DNA (X = C), which is easily determined by quantitative PCR, can serve as an index for the discrimination of a single C/M placed at the branching point of TWJ.

Probes used to form a single-base bulge and mismatch with ssDNA have been applied for the sequence-selective cleavage and modification of ssDNA.[8,19] A 56mer target DNA1′ with the same sequence as DNA1, except for a CXpG trinucleotide (X = C or M) containing two consecutive cytosine residues at the centre of the sequence, was subsequently designed. Target DNA1′ was designed to compare the C/M-discrimination ability of the probes forming the fully matched TWJ (FM-TWJ probes 1′ and 2) with that of various probes that hybridize with the target DNA1′ and form secondary structures, including a single-base bulge, a mismatch and a

mismatched TWJ (Supplementary Fig. 2). After incubation with sodium bisulfite and CMH, the unmodified DNA1′ rate in the presence of various probes was evaluated by quantitative PCR using primers 1 and 2, as previously described (Fig. 4). In the presence of FM-TWJ probes 1′ and 2, 28% of DNA1′ (X = M) and 3.4% of DNA1′ (X = C), respectively, remained unmodified, resulting in an approximately 8-fold increase in the unmodified DNA1′ (X = M) rate compared with that of DNA1′ (X = C). In the presence of two probes forming the TWJ with a mismatch at the branching point (mismatched TWJ probes 1 and 2), only 11% of DNA1′ (X = M) was unmodified, indicating that an M–M mismatch at the branching point changed the local structure of the TWJ and caused modification of C adjacent to MpG. Probes that formed a bulge and a mismatch with DNA1′ resulted in a high unmodified DNA rate, irrespective of whether X was C or M, which in turn, resulted in a small increase or decrease in the unmodified DNA1′ (X = M) rate compared with that of DNA1′ (X = C). The reason for the observed high unmodified DNA1′ rates may be that these probes contain two primer regions at both termini and can thus serve as a template for PCR amplification when they remain unmodified after incubation; in contrast, FM-TWJ probes, which are split into two parts, cannot serve as a PCR template. Consequently, of all the probes tested, FM-TWJ probes 1′ and 2 exhibited the greatest increase in the unmodified DNA1′ (X = M) rate relative to that of DNA1′ (X = C) and also exhibited the most efficient C/M-discrimination ability when hybridized with DNA1′.

To further demonstrate the potential applicability of the FM-TWJ probes as a tool for simple methylation analysis, we prepared a 54mer target DNA2 whose sequence is a part of the tumour suppressor *p16* gene promoter (1878-1931, GenBank AF022809.1) and FM-TWJ probes 3 and 4 to form a TWJ with DNA2 (Supplementary Fig. 3). A 54mer DNA complementary to DNA2 (cDNA2) was also prepared to perform discrimination of a single C/M in double-stranded DNA2. After incubation of single-stranded and double-stranded 54mer target

DNA2 (15 fmol, X = C or M) with sodium bisulfite and CMH in the presence of FM-TWJ probes 3 and 4 (6 pmol each), the unmodified DNA2 rate was determined by quantitative PCR using primers 3 and 4, as previously described. When a single-stranded target DNA2 was incubated in the presence of FM-TWJ probes, 56% of DNA2 (X = M) and 6.7% of DNA2 (X = C), respectively, remained unmodified. Similarly, when a double-stranded target DNA2 was used, 34% of DNA2 (X = M) and 4.4% of DNA2 (X = C), respectively, remained unmodified. In both cases, an approximate 8-fold increase in the unmodified DNA2 (X = M) rate relative to that of DNA2 (X = C) was confirmed. In addition, a single-stranded DNA2 in the absence of FM-TWJ probes was tested, and less than 0.1% of DNA2 was unmodified irrespective of whether X was C or M. We also confirmed that the methylation level of a single cytosine was quantifiable in a similar manner. A linear relationship between Ct and the logarithm of the methylation ratio of X was observed using mixtures of a double-stranded DNA2 (X = M) and DNA2 (X = C) with a fixed amount (150 amol) and varying ratios (Fig. 6).

In summary, we developed a new method for sequence-selective modification of a single cytosine using C-selective reagents bisulfite and CMH and probes that form fully matched TWJs with target DNAs, which were prepared by a standard synthetic method. Furthermore, we detected the existence of only one methyl group in a 54mer double-stranded DNA using the combination of sequence-selective C-modification and quantitative PCR. Two methods for the sequence-selective chemical discrimination of a single C/M using either bulge-forming probes[8] or bipyridine-tethered probes[9] have been previously reported; a single methylation site in a 60mer target DNA[8] and even in genomic DNA[9] was successfully modified using M-selective osmium complexation and was discriminated using quantitative PCR after recovery of the DNA sample. However, the need for custom synthesis of the bipyridine-tethered nucleotide derivative[9] and the risk of forming hazardous osmium tetroxide during the modification[8, 9] limit

the applicability of these methods. Although improvement of the detection limit, which can be achieved by optimization of the probe/primer design and reaction conditions, is an issue for the analysis of genomic DNA, we anticipate that this study will be the first step toward a simple, sequence-selective methylation assay that will surpass conventional methods, including bisulfite sequencing.

Acknowledgements

This study was supported in part by A-Step (AS251Z03023P) from the Japan Science and Technology Agency and by a Grant-in-Aid for Scientific Research (No. 22501037) from the Japan Society for the Promotion of Science.

References

1. M. M. Suzuki and A. Bird, *Nat. Rev. Genet.*, 2008, **9**, 465-476.

2. Y. Gruenbaum, R. Stein, H. Cedar and A. Razin, *FEBS Lett.*, 1981,**124**, 67-71.

3. M. Esteller, *N. Engl. J. Med.*, 2008, **358**, 1148-1159.

4. H. Hayatsu, Y. Wataya, K. Kai and S. Iida, *Biochemistry*, 1970, **9**, 2858-2865.

5. R. Shapiro, R. E. Servis and M. Welcher, *J. Am. Chem. Soc.*, 1970, **92**, 422-424.

6. H. Hayatsu, *Mutat. Res.*, 2008, **659**, 77-82.

7. C. Grunau, S. J. Clark and A. Rosenthal, *Nucleic Acids Res.*, 2001, **29**, e65.

8. A. Okamoto, K. Tainaka and T. Kamei, *Org. Biomol. Chem.*, 2006, **4**, 1638-1640.

9. K. Tanaka, K. Tainaka, T. Umemoto, A. Nomura and A. Okamoto, *J. Am. Chem. Soc.*, 2007, **129**, 14511-14517.

10. S. Bareyt and T. Carell, *Angew. Chem. Int. Ed.*, 2008, **47**, 181-184.

11. Y. Oka, T. Peng, F. Takei and K. Nakatani, *Org. Lett.*, 2009, **11**, 1377-1379.

12. M. Komiyama, *J. Biochem.*, 1995, **118**, 665-670.

13. M. Chatterjee and S. E. Rokita, *J. Am. Chem. Soc.*, 1991, **113**, 5116-5117.

14. D. Duckett and D. M. J. Lilley, *EMBO J.* 1990, **9**, 1659-1664.

15. N. C. Seeman, *Angew. Chem. Int. Ed*., 1998, **37**, 3220-3238.

16. A. Oleksy, A. G. Blanco, R. Boer, I. Usón, J. Aymamí, A. Rodger, M. J. Hannon and M. Coll, *Angew. Chem. Int. Ed*., 2006, **45**, 1227-31.

17. T. Kato, K. Yano, K. Ikebukuro and I. Karube, *Nucleic Acids Res*., 2000, **28**, 1963-1968.

18. T. Bessho, N. Nitta, K. Negishi and H. Hayatsu, *Nucleic Acids Res.*, 1992, **20**, 4213-4220.

19. J. Brunner and J. K. Barton, *J. Am. Chem. Soc.*, 2006, **128**, 6772-6773.

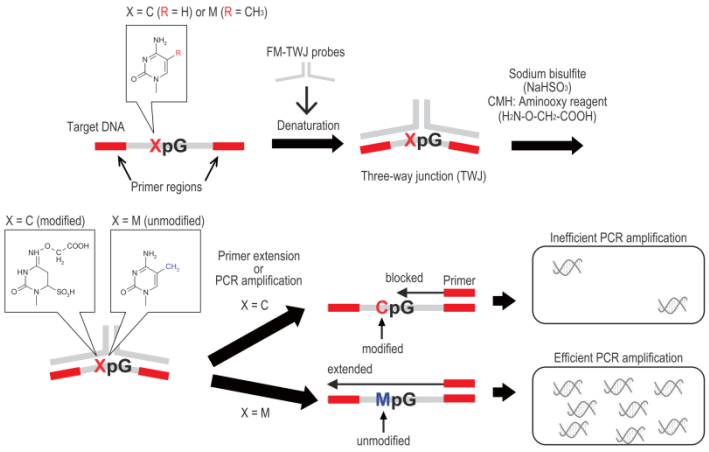


Fig. 1 Schematic representation of sequence-selective C-modification using two DNA probes to form a fully matched TWJ with the target DNA and subsequent C/M discrimination using primer extension and a quantitative PCR assay.

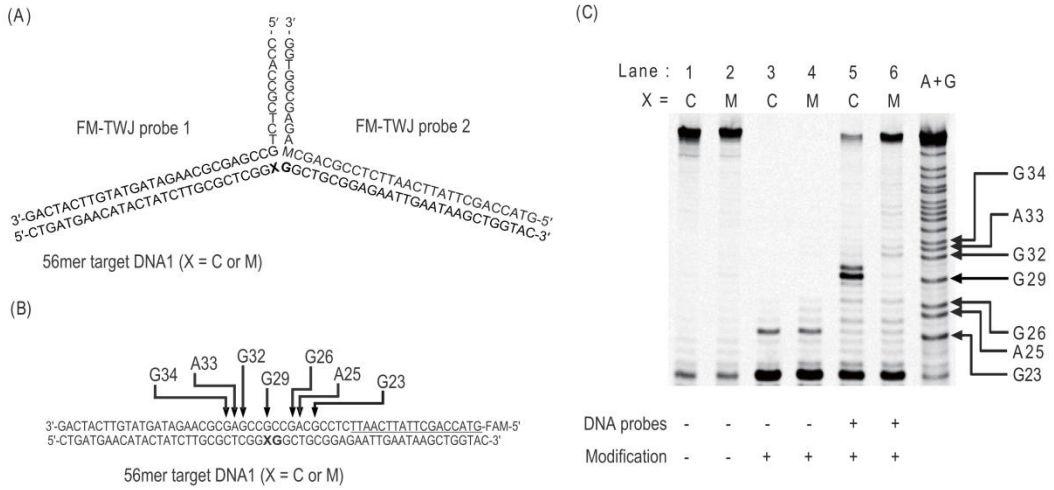


Fig. 2 Discrimination of a single C/M in 56mer target DNA1. (A) Fully matched TWJ formed by 56mer DNA1 and the FM-TWJ probes 1 and 2. Bold letters denote the XpG of interest. (B) Fully extended 5′-FAM-labeled 18mer primer annealing to 56mer DNA1. The underlined letters indicate the 18mer primer sequence. Bold letters denote the XpG of interest. (C) PAGE analysis of the extended 5′-FAM-labeled 18mer primers. The primers extended after annealing to non-incubated DNA1 (lanes 1 and 2) and to DNA1 incubated in the absence (lanes 3 and 4) and presence (lanes 5 and 6) of the FM-TWJ probes 1 and 2. A+G, Maxam–Gilbert A+G ladder of 5′-FAM-labeled 56mer DNA complementary to DNA1 (5'-FAM-labeled-cDNA1).

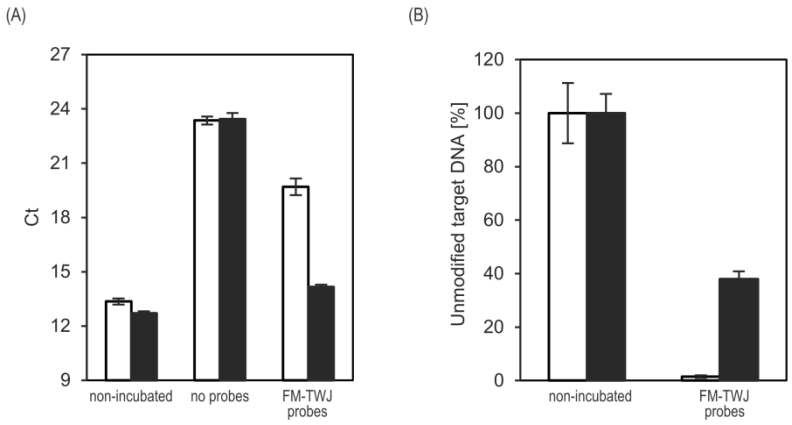


Fig. 3 Quantitative PCR assay for discrimination of a single C/M in 56mer target DNA1. (A) Ct values of non-incubated DNA1 (non-incubated) and incubated DNA1 in the absence (no probes) and presence (FM-TWJ probes) of FM-TWJ probes 1 and 2; white bars: DNA1 (X = C), black bars: DNA1 (X = M). Average values from three independent experiments are shown with error bars indicating the standard deviations. (B) Unmodified DNA1 rate (%) of non-incubated DNA1 (non-incubated) and incubated DNA1 in the presence of FM-TWJ probes 1 and 2 (FM-TWJ probes); white bars: DNA1 (X = C), black bars: DNA1 (X = M). Average values from three independent experiments are shown with error bars indicating the standard deviations.

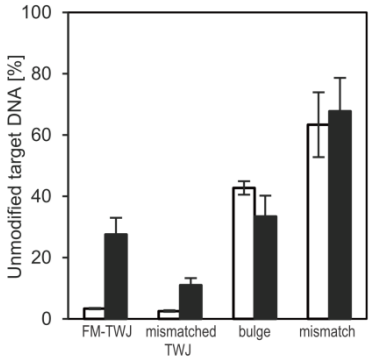


Fig. 4 Comparison of the C/M discrimination ability of various probes evaluated by quantitative PCR. Unmodified DNA1′ rate (%) of incubated DNA1′ in the presence of FM-TWJ probes 1′ and 2 (FM-TWJ), mismatched TWJ probes 1 and 2 (mismatched TWJ), a bulge probe (bulge) and a mismatch probe (mismatch); white bars: DNA1′ (X = C), black bars: DNA1′ (X = M). The amount of unmodified DNA1′ was determined from its Ct values using a standard curve of non-incubated DNA1′ (data not shown). Average values from three independent experiments are shown with error bars indicating the standard deviations.

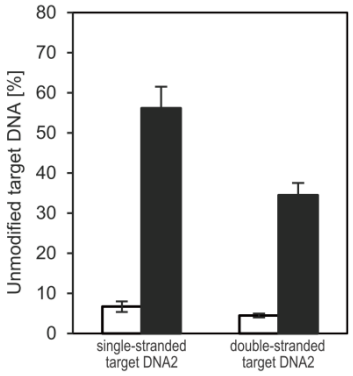


Fig. 5 Quantitative PCR assay for discrimination of a single C/M in single-stranded and double-stranded 54mer target DNA2. Unmodified DNA2 rate (%) of incubated singe-stranded and double-stranded DNA2 in the presence of FM-TWJ probes 3 and 4; white bars: DNA2 (X = C), black bars: DNA2 (X = M). The amounts of unmodified singe-stranded and double-stranded DNA2 were determined from their Ct values using standard curves of non-incubated singe-stranded and double-stranded DNA2 (data not shown). Average values from three independent experiments are shown with error bars indicating the standard deviations.

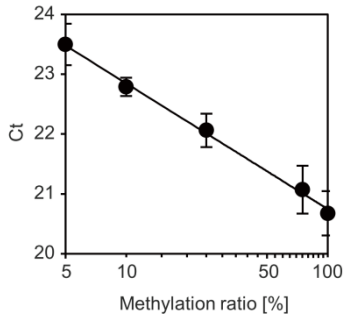


Fig. 6 Linear relationship between Ct and the methylation ratio in the experiment using 150 amol mixtures of a double-stranded target DNA2 (X = C) and DNA2 (X = M). Average values from three independent experiments are shown with error bars indicating the standard deviations.