DNA methylation detection based on difference of base content

DNA methylation detection based on difference of base content

Shinobu Sato,[1,2] Keiichi Ohtsuka,[1,2] Satoshi Honda,[1] Yusuke Sato,[1] and Shigeori Takenaka[1,2,*]

[1]Department of Applied Chemistry, Kyushu Institute of Technology, Kitakyushu, Japan

[2]Research Center for Bio-microsensing Technology, Kyushu Institute of Technology, Kitakyushu, Japan

shige@che.kyutech.ac.jp

Abstract. Methylation frequently occurs in cytosines of CpG sites to regulate gene expression. The identification of aberrant methylation of certain genes is important for cancer marker analysis. The aim of this study was to determine the methylation frequency in DNA samples of unknown length and/or concentration. Unmethylated cytosine is known to be converted to thymine following bisulfite treatment and subsequent PCR. For this reason, the AT content in DNA increases with an increasing number of methylation sites. In this study, the fluorescein-carrying bis-acridinyl peptide (FKA) molecule was used for the detection of methylation frequency. FKA contains fluorescein and two acridine moieties, which together allow for the determination of the AT content of double-stranded DNA fragments. Methylated and unmethylated human genomes were subjected to bisulfide treatment and subsequent PCR using primers specific for the CFTR, CDH4, DBC1, and NPY genes. The AT content in the resulting PCR products was estimated by FKA, and AT content estimations were found to be in good agreement with those determined by DNA sequencing. This newly developed method may be useful for determining methylation frequencies of many PCR products by measuring the fluorescence in samples excited at two different wavelengths.

1. Introduction

Cancer-suppressing genes are inactivated by the aberrant methylation of cytosine in the CpG islands of their promoter regions [1]. Such inactivation leads to cancer via loss of physiological function of these genes or due to a loss of transcription factor activity resulting from secondary methylation of related genes [2]. The detection of methylation frequency in cancer-suppressing genes has thus become an important aspect of cancer diagnosis. Cytosine methylation can be detected by (1) bisulfite sequencing [3], (2) methylation-specific PCR [4], or (3) combined bisulfite restriction analysis (COBRA) methods [5]. These methods, however, have limitations such as time-consuming procedures with tedious steps and the formation of unclear internally formed PCR products or DNA sequencing capacity. To overcome these limitations, several improved methods have been developed [6-13]. To overcome the limitations of method 1, direct detection of methylated cytosine was assessed without bisulfite treatment. Okamoto and co-workers succeeded in detecting methylated cytosine based on the specific reaction with osmium complex [6], while Niwa and co-workers were able to discriminate between

methylated and unmethylated cytosine electrochemically based on differences in redox potentials [7]. Another approach was to use surface plasmon resonance (SPR) with a methylated cytosine-specific antibody [8]. Since unmethylated cytosine is converted to thymine after bisulfite treatment and subsequent PCR, the AT content is increased with increasing methylation frequency. Accordingly, methylation frequency can be determined based on T_m values of DNA compared with unmethylated DNA [9]. However, this method has a limitation that the amplification region for PCR analysis is subject to T_m value limitations, where higher T_m values are not measured accurately. The limitations associated with methods 2 and 3 were addressed by combining DNA hybridization after bisulfite treatment with subsequent PCR. We achieved electrochemical methylation detection in PCR products after bisulfite treatment coupled with DNA-immobilized electrode and ferrocenylnaphthalene diimide (FND) as a double-stranded DNA-specific ligand [10-13]. High sensitivity detection of the hybrid species was achieved using a surface-enhanced Raman scattering method [14]. Primer and target regions are discriminated precisely in these methods. Here, a bis-acridinyl peptide carrying fluorescein (FKA, Fig. 1A) [15] was applied to estimate methylation frequency in PCR products, where FKA exhibits almost no fluorescence in aqueous solutions because of intramolecular quenching by fluorescein being sandwiched between two acridine moieties. When FKA is bound to double-stranded DNA, the fluorescein of FKA is eliminated by the intercalation of the acridine portion of the peptide between the base pairs of the DNA duplex and thus the fluorescence is recovered. Interestingly, the fluorescence of fluorescein (λ_{em} = 517 nm) was observed via FRET when excited at 422 nm based on the acridine chromophore and this fluorescence correlated with AT content, since the acridine fluorescence was quenched due to binding to GC base pairs (Fig. 1B). The fluorescence (517 nm) obtained with an excitation wavelength of 492 nm, however, correlated with the amount of DNA present (Fig. 1B). The AT content of DNA was therefore estimated by the fluorescence ratio of ΔFI_{422}:ΔFI_{492} ($\Delta FI_{422}/\Delta FI_{492}$) [15]. Since unmethylated cytosine is converted to thymine after bisulfite treatment and PCR, AT content differs between methylated and unmethylated PCR products. The AT content in DNA can be detected by FKA without knowing the DNA concentration and thus we hypothesized that FKA can be used to estimate the AT content of PCR products of unknown length and sequence. To confirm this hypothesis, the methylation frequencies in PCR products representing portions of the *CDH4* [16], *CFTR* [17], *DBC1* [17], and *NPY* [17] genes—all with different lengths and sequences—were assessed using FKA.

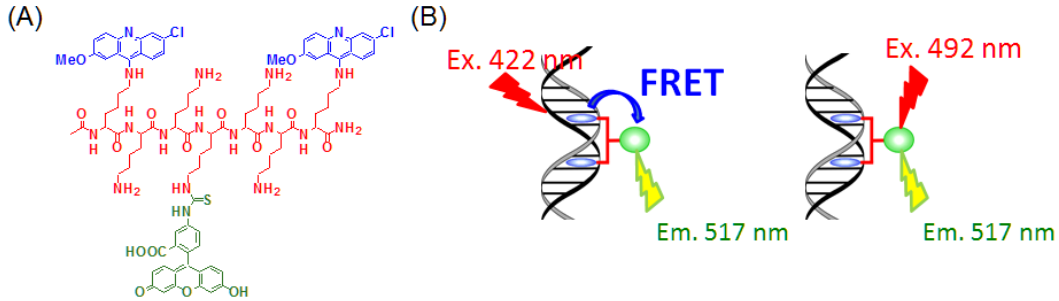

Fig. 1. (A) Chemical structure of the fluorometric peptide-type intercalator FKA and (B) the principle of AT content estimation with FKA.

2. Experimental

2.1. Optimization of experimental conditions

Fluorescence measurements were carried out using 96-well titer plates (Nunc) containing 100 µL/well of either blank solution (0.5 µM FKA, 5 mM Tris-HCl [pH 8.3], 0.5 mM EDTA) or test solution (20

μM (-bp) dsDNA, 0.5 μM FKA, 5 mM Tris-HCl (pH 8.3), and 0.5 mM EDTA) as well as 100 mM, 200 mM, or 400 mM NaCl. A Molecular Devices SPECTRA MAX GEMINI XS instrument was used with excitation wavelengths of either 422 nm or 492 nm and an emission wavelength of 517 nm. Double-stranded DNA species with different AT contents were used: poly[(dA-dT)]$_2$, *Clostridium perfringens* DNA, calf thymus DNA, *Escherichia coli* DNA, *Micrococcus lysodeikticus* DNA, and poly[(dG-dC)]$_2$, which have known AT contents of 100%, 73%, 57%, 28%, and 0%, respectively.
The fluorescence readings obtained were evaluated using equations 1, 2, and 3 below, where $FI_{ex.422, DNA}$ and $FI_{ex.422, blank}$ represent fluorescence at 517 nm when excited at 422 nm in the presence (DNA) or absence (blank) of DNA duplex, respectively; and $FI_{ex.492, DNA}$ and $FI_{ex.492, blank}$ represent fluorescence at 517 nm when excited at 492 nm in the presence (DNA) or absence (blank) of DNA duplex.

$$\Delta FI_{422} = FI_{ex.422, DNA} - FI_{ex.422, blank} \quad (1)$$

$$\Delta FI_{492} = FI_{ex.492, DNA} - FI_{ex.492, blank} \quad (2)$$

From (1) and (2), $\text{F-Ratio} = \Delta FI_{422}/\Delta FI_{492}$ (3)

2.2 Material collection
Oligonucleotide template genes or PCR primers (Table 1) were purchased from Genenet (Fukuoka, Japan). The template DNA sequences of the *CFTR* gene were as follows [16]: methylated *CFTR* gene (CFTR-M), 5′-GTT TTG GGT TTG GCG GAT TTT GAC GCG AAG GAG GGT TTA GGA AGT TTT TCG GGG AGT CGG TTT TTT CGT CGG TGG TTT TTT TTG TTT TTT AGC GTT GTT AAT TGG ATT TAA AGA GAG GTC GGA TTG TCG TTT ATT TGC GGG-3′, unmethylated *CFTR* gene (CFRT-U): 5′-GGT TTT GGG TTT GGT GGA TTT TGA TGT GAA GGA GGG TTT AGG AAG TTT TTT GGG GAG TTG GTT TTT TTG TTG GTG GTT TTT TTT GTT TTT TAG TGT TGT TAA GAT TTA AAG AGA GGT TGT GAT TGT TGT TTA TTT GTG GGA TG-3′. CpGenome™ Universal Methylated DNA (Lot PSO 1566253, Merck Millipore) and CpGenome™ Universal Unmethylated DNA (Lot Number PSO 1579958, Merck Millipore) were used as methylated and unmethylated human genome DNA, respectively. The PrimeStar (TAKARA Bio) DNA polymerase was used to amplify oligonucleotide PCR templates, while ZymoTaq PreMix (Zymo research) was used to amplify genomic DNA after bisulfite treatment.

2.3. Bisulfite treatment
CpGenome™ Universal Methylated DNA (Lot PSO 1566253) and CpGenome™ Universal Unmethylated DNA (Lot Number PSO 1579958) were used as methylated and unmethylated samples, respectively. The above genomic DNA species (1.0 μg) were treated with the CpGenome™ FAST DNA Modification Kit (Merck Millipore) according to the manufacturer's instructions. The concentrations of the bisulfite-treated human genome templates were estimated by Nanodrop (Thermo Scientific) to be 22.9 ng/μl (Methylated, Lot PSO 1566253) and 17.4 ng/μl (Unmethylated, Lot Number PSO 1579958).

2.4. Polymerase Chain Reaction (PCR)
Table 1 shows the sequences of the primers used in PCR. Total reaction volumes of 100 μL were prepared: 4.0 μL CFTR-M (1 nM) or CFTR-U (1 nM) (or 1.0 μL of 22.9 or 17.4 ng/μL template DNA), 4 μL each of F- and R-primers (10 μM), and 50 μL PrimeStar (or ZymoTaq). PCR was carried out under the following conditions: 95°C for 5 min, followed by 30 cycles of 94°C for 15 sec, annealing temperature [different for each template] for 5 sec, and 72°C for 10 sec. The resulting PCR products were purified using a QIAquick PCR Purification Kit (QIAGEN) and were eluted in 30 μl elution buffer. The purified PCR products were assessed by 12.5% native polyacrylamide gel electrophoresis (PAGE).

2.5. PCR product sequencing
Sequencing of PCR products was carried out by Hokkaido System Science (Sapporo, Japan).

2.6. Fluorescence analysis of PCR products

Fluorescence measurements were carried out in microtiter plates containing 100 µL/well of the following solution: 0.6 µM-bp PCR product, 0.5 µM FKA, 5 mM Tris-HCl (pH 8.3), and 0.5 mM EDTA as well as 100 mM, 200 mM, or 400 mM NaCl using a Molecular Devices SPECTRA MAX GEMINI XS instrument with excitation wavelengths of 422 nm or 492 nm and an emission wavelength of 517 nm. The resulting fluorescence values were evaluated using equation (3).

Table 1. Primer sequences used in this experiment

Name	**Sequence 5'→3'**	**Product size (bp)**	**Tm (°C)**	**PCR Cycles**
CFTR_MF	GTTTTGGGTTTGGCGGATTTTGACGC	142	46 56*	40 30*
CFTR_MR	CCCGCAAATAAACGACAATCGCGAC			
CFTR_UF	GGTTTTGGGTTTGGTGGATTTTGATGT	146	46 56*	40 30*
CFTR_UF	CATCCCACAAATAAACAACAATCACAAC			
CDH4_MF1	TTGTAGTTTCGAGCGCGC	117	45	40
CDH4_MR1	CCCGACTCCGAAAACCAAA			
CDH4_MF2	GTTTTCGGTGTCGGGTATC	105	45	40
CDH4_MR2	CGACAACTTACCCGAAACG			
DBC1_MF	ACGCGATCCCTTTAAATACTCGTACG	131	60	40
DBC1_MR	GAGGAGAGACGGGAGGTCGTTTCG			
NPY_MF	TTAAAACCCTCTAACCGAAAACTTCCG	131	45	40
NPY_MR	ACGATTAGCGCGGTATTTTCGTCGG			

*Custom Synthesized Template oligonucleotide

Results and Discussion

3.1 Optimization of detection conditions

Fluorescence at 515 nm was measured for 20–50 µM-bp $[poly(dA\text{-}dT)]_2$, calf Thymus DNA, or $[poly(dG\text{-}dC)]_2$ using 0.50 µM FKA in 5 mM Tris-HCl (pH 8.3), 0.5 mM EDTA, and 200 mM NaCl excited at 422 nm or 492 nm and evaluated in terms of F-Ratio calculated by equation (3). F-Ratios were shown to differ depending on AT content, but were found to be stable at different DNA concentrations (Fig. 2).

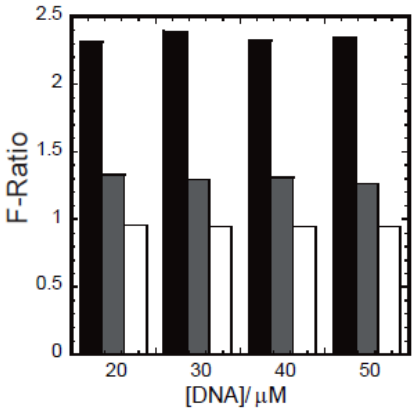

Fig. 2. Estimation of AT content for DNA samples of unknown concentration. Black bar: [poly(A-T)]$_2$, gray bar: calf thymus DNA, white bar: [poly(dG-dC)]$_2$. Experiments were carried out for 20–50 μM-bp DNA in 5 mM Tris-HCl (pH 8.3), 0.5 mM EDTA, and 200 mM NaCl containing 0.5 μM FKA.

Here, six different PCR products were prepared using four different genes (Table 1). These PCR products were fixed at a concentration of 0.6 μM per double-stranded DNA, while the PCR products of 105–146 bp in length corresponded to concentrations of 63–88 μM-bp and were thus subjected to AT content quantification using F-Ratio. In the next step, the effect of salt concentration on AT content estimation by FKA was assessed with NaCl concentrations of 100, 200, or 400 mM using poly[(dA-dT)]$_2$, *C. perfringens* DNA, calf thymus DNA, *E. coli* DNA, *M. lysodeikticus* DNA, and poly[(dG-dC)]$_2$ with AT contents of 100%, 73%, 57%, 28%, and 0%, respectively. Figure 3 shows that the F-Ratios varied with AT content and that the greatest change in F-Ratio was observed with 400 mM NaCl for DNA with >50% AT content. Since these PCR products had AT contents of 37.8–65.1%, subsequent experiments were carried out with 400 mM NaCl.

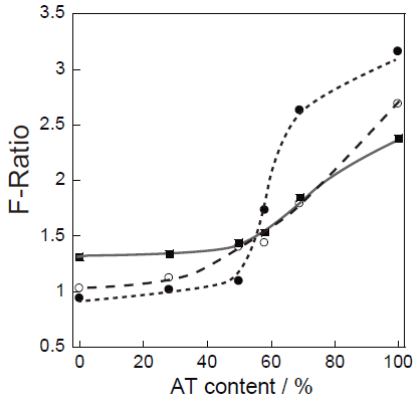

Figure 3. F-Ratio depends on AT content and NaCl concentration. Experiments were carried out with 20 μM DNA, 5 mM Tris-HCl (pH 8.3), 0.5 mM EDTA, 0.5 μM FKA, and NaCl at 100 (■), 200 (○), or 400 (●) mM.

3.2 AT content estimation using the model system

As an initial trial use of the newly developed model system, AT content was estimated by FKA for assessing methylation frequency in the *CFTR* gene, the methylation of which is associated with bowel cancer [17]. The CFTR-M (AT content: 65.1%) and CFTR-U (AT content: 56.8%) DNA species were used as methylated and unmethylated *CFTR* genes following bisulfite treatment, after which PCR products were prepared using these DNA species as templates. A single band of ~140 bp in size was

obtained in native PAGE analysis of the PCR products (data not shown). F-Ratios of these PCR products were determined using FKA and were found to be 1.65 and 1.05 for the methylated and unmethylated samples, respectively, which is in agreement with the data shown in Fig. 3. The CFTR-M and CFTR-U DNA species were combined at ratios of 0:1, 0.25:0.75, 0.5:0.5, 0.75:0.25, and 1:0 ratio to yield AT contents of 65.1, 63.0, 61.0, 58.9, and 56.8%, respectively. The resulting F-Ratios are shown in Fig. 4 and were found to correlate well with the theoretical values (Fig. 3).

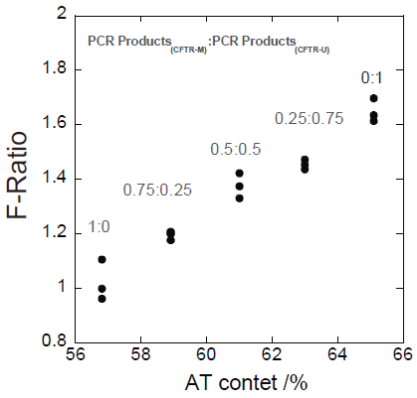

Figure 4. Detection of methylation frequency in the *CFTR* gene by FKA. Experiments were carried out with 0.6 μM PCR product in 5 mM Tris-HCl (pH 8.3), 0.5 mM EDTA, and 400 mM NaCl containing 0.5 μM FKA.

3.3 Methylation detection in genomic samples

The *CFTR*, *CDH4*, *DBC1*, *and NPY* genes were subjected to PCR after bisulfite treatment. CpGenome™ Universal Methylated DNA (Lot PSO 1566253) and CpGenome™ Universal Unmethylated DNA (Lot Number PSO 1579958) were used as template DNA and the oligonucleotides listed in Table 1 were used as primers. The *CFTR*, *CDH4*, *DBC1*, and *NPY* genes correlate with bladder cancer, large intestinal cancer, lung cancer (non-small cell), and neuropeptide Y (a factor indirectly causing cancer), respectively. The resulting PCR products each yielded single bands of the expected sizes in native PAGE (data not shown). Although the methylation frequencies and AT contents of these products are not known, they were determined by DNA sequencing (Fig. A1). The AT contents of the PCR products are listed in Table 2, where the DNA sequencing results are shown together with the theoretical values (assuming completely methylated genomic DNA). The PCR products were then tested using the 0.5 μM FKA method and the resulting F-Ratios were plotted against AT content (Fig. 5). The resulting F-Ratios were in good agreement with the expected AT content, yielding good correlation between the two parameters.

Table 2. AT content determined from DNA sequencing compared with theoretical values.

AT content (%)	CFTR-M	CFTR-U	CDH4 Region 1	CDH4 Region 2	DBC1	NPY
Theoretical	56.8	65.1	37.8	42.9	43.5	59.5
Based on sequencing	56.8	63.0	37.8	51.4	43.5	59.5

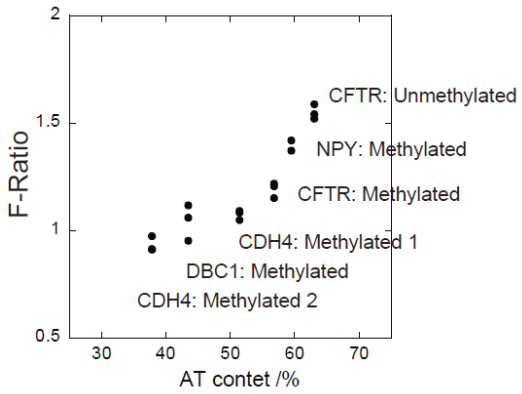

Fig. 5. Relationship between F-Ratio and AT content of methylated and unmethylated PCR products after bisulfite treatment. Experiments were carried out with 0.6 μM PCR product in 5 mM Tris-HCl (pH 8.3), 0.5 mM EDTA, and 400 mM NaCl containing 0.5 μM FKA.

4. Conclusion

Methylation frequency based on AT content was estimated by FKA for the *CFTR*, *CDH4*, *DBC1*, and *NPY* genes amplified from human genome samples subjected to bisulfite treatment. This method does not allow for the location of methylated site of PCR products to be determined; however, methylation frequency was estimated in double-stranded DNA samples of different lengths and/or unknown concentrations using this simple procedure. This newly developed method may be useful in determining methylation frequencies of many PCR products by simply measuring the fluorescence in samples excited at two different wavelengths. Furthermore, it is possible to estimate the methylation frequency in a DNA sample via real-time methods using an RT-PCR instrument and carrying out the reactions in the presence of FKA after bisulfite treatment of sample DNA. For this method to be effective in real-time, one would need to ensure that the polymerase is not inhibited during the PCR stage of the process.

References

[1] Ohtani N, Fujita T, Aoike A, Osifchin N, Robbins P and Sakai T 1993 *Oncogene*, **8** 1063.
[2] Abe M, Ohira M, Kaneda A, Yagi Y, Yamamoto S, Kitano Y, Takato T, Nakagawara A and Ushijima T 2005 *Cancer Res.* **65** 828.
[3] Frommer M, Mcdonald L E, Millar D S, Collis C M, Watt F, Grigg G W, Molloy P L and Paul C L 1992 *Proc. Natl. Acad. Sci. USA* **89** 1827.
[4] Herman J G, Graff J R, Myöhänen S, Nelkin B D and Baylin S B 1996 *Proc. Natl. Acad. Sci. USA* **93** 9821.
[5] Miotto E, Sabbioni S, Veronese A, Calin G A, Gullini S, Liboni A, Gramantieri L, Bolondi L, Ferrazzi E, Gafa R, Lanza G and Negrini M 2004 *Cancer Res.* **64** 8156.
[6] Tanaka K, Tainaka K, Umemoto T, Nomura A and Okamoto A 2007 J. Am. Chem. Soc. **129**14511.
[7] Kurita R and Niwa O 2012 *Anal. Chem.* **84** 7533.
[8] Kato D, Sekioka N, Ueda A, Kurita R, Hirono S, Suzuki K, Niwa O 2008 *J. Am. Chem. Soc.* **130** 3716.
[9] Guldberg P, Worm J, Grønbaek K 2002 *Methods* **27** 121.
[10] Sato S, Kondo H, Hokazono K, Irie T, Ueki T, Waki M, Nojima T and Takenaka S 2006, *Anal. Chim. Acta*, **578** 82.
[11] Sato S, Tsueda M and Takenaka S 2010 *J. Organomet. Chem.* **695** 1858.
[12] Sato S, Tsueda M, Kanezaki Y and Takenaka S 2012 *Anal. Chim. Acta*. **715**, 42.

[13] Sato S, Saeki T, Tanaka T, Kanezaki Y, Hayakawa M, Haraguchi K, Kodama M, Nishihara T, Tominaga K and Takenaka S 2014 *Appl. Biochem. Biotech.* **174** 869.
[14] Wang Y, Wee E J H and Trau M 2015 *Chem. Commun.***51** 10953.
[15] Ueyama U, Mizuki K, Nojima T and Takenaka S 2004 *Analyst* **129** 886.
[16] Bibikova M, Lin Z, Zhou L, Chudin E, Garcia E W, Wu B, Doucet D, Thomas N J, Wang Y, Vollmer E, Goldmann T, Seifart C, Jiang W, Barker DL, Chee M S, Floros J and Fan J B 2006 *Genome Res.* **16** 383.
[17] Miotto E, Sabbioni S, Veronese A, Calin GA, Gullini S, Liboni A, Gramantieri L, Bolondi L, Ferrazzi E, Gafà R, Lanza G and Negrini M 2004 *Cancer Res*. **64** 8156.

Acknowledgments
This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (MEXT).