UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

Exploration of Genetically Encoded Sensors for DNA Methylation in Live Cells

Permalink

https://escholarship.org/uc/item/6pk9q745

Author

Fang, Chih-Yu

Publication Date

2018

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Exploration of Genetically Encoded Sensors for DNA Methylation in Live Cells

A Thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Science

in

Cell, Molecular and Developmental Biology

by

Chih Yu Fang

September 2018

Thesis Committee:
Dr. Prudence Talbot, Chairperson
Dr. Huiwang Ai
Dr. Yinsheng Wang

Copyright by
Chih Yu Fang
2018

The Thesis of Chih Yu Fang is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Dr. Huiwang Ai, for his guidance and support. I would also like to thank my thesis committee members, Dr. Yinsheng Wang and Dr. Prudence Talbot, for their valuable input and suggestions.

ABSTRACT OF THE THESIS

Exploration of Genetically Encoded Sensors for DNA Methylation in Live Cells

by

Chih Yu Fang

Master of Science, Graduate Program in Cell, Molecular, and Developmental Biology
University of California, Riverside, September 2018
Dr. Prudence Talbot, Chairperson

Epigenetic modification refers to heritable changes that influence gene function without changing the DNA sequence. Besides histone modification, DNA methylation and hydroxymethylation have been reported to be major epigenetic factors involved in cellular events. Histone modifications often refer to the various post-translational modifications applicable to histone proteins, including methylation, acetylation, phosphorylation, ubiquitination, and sumoylation, all of which typically influence gene expression by changing the chromatin structure or recruiting histone modifiers.[1] DNA methylation is a process through which methyl groups are added to DNA, particularly cytosine bases, typically leading to gene repression.[2] DNA methylation has been identified to be an important participant in embryogenesis, cell division, and cell development. Unusual DNA methylation leads to various diseases, such as cancers and neurological disorders.[3-5] Therefore, technologies that can detect DNA methylation have

become in high demand in the biomedical and pharmaceutical science fields. Several recent studies have reported fluorescent reagents for live-cell labeling of DNA methylation, but these reagents suffer from high background noise, low specificity, and implementation difficulties. Genetically encoded fluorescent protein (FP)-based probes have recently been demonstrated as powerful tools for imaging targets and processes in living cells and tissues. In this research project, we explore dimerization-dependent red FPs to enable the development of genetically encoded fluorogenic reagents to label global DNA methylation in live cells. This new ability to image live-cell genome-wide DNA methylation is expected to catalyze a large array of basic and translational studies related to epigenetics and genetics at the single-cell level.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv

ABSTRACT OF THE THESIS v

Chapter 1 1

Introduction 1

1.1. Theory of Epigenetics 1

1.2. DNA Methylation and DNA Methyltransferases 2

1.3. DNA Demethylation 4

1.4. DNA Methylation and Histone Modifications 7

1.5. Methyl-CpG-binding Domain Proteins and MeCP2 8

1.6. DNA Methylation and Cancer Development 9

2. Fluorescent Proteins 11

2.1. Discovery of Green Fluorescent Protein 11

2.2. Derivatives of Green Fluorescent Protein 11

2.3. Discovery of Red Fluorescent Protein 13

2.4. Derivatives of Red Fluorescent Protein 14

2.5. Fluorescent Protein-based biosensors 16

2.6. Dimerization-Dependent Fluorescent Protein-based Biosensors 17

3. Fluorescent Biosensors for DNA Methylation 20

Chapter 2 21

2.1. Introduction 21

2.2. Experimental Section 22

2.2.1. Materials, Reagents, and General Methodology 22

2.2.2. Plasmid Construction 22

2.2.3. Bacterial Transformation 29

2.2.4. Electro-competent Cell Preparation 29

2.2.5. Plasmid Large-Prep 30

2.2.6. Cell Culture 31
2.2.7. Live-Cell Fluorescence Imaging 31
2.3. Results and Discussion 32
Chapter 3 52
Conclusion and Outlook 52
References 54

TABLE OF FIGURES

Figure 1. Cytosine (C) is methylated to 5-methylcytosine (5mC) 3

Figure 2. Proposed mechanisms of DNA methylation and demethylation cycle 6

Figure 3. Conceptual schematic of ddFP technology 19

Figure 4. Conceptual scheme of ddRFP-based probe for DNA methylation 34

Figure 5. Construct of pAB-M 35

Figure 6. Construct of pATB-M 36

Figure 7. Live-cell imaging of methylated DNA 38

Figure 8. Live-cell imaging of methylated DNA 40

Figure 9. Live-cell imaging of methylated DNA 41

Figure 10. Construct of pE-M 43

Figure 11. Live-cell imaging of methylated DNA 44

Figure 12. Live-cell imaging of methylated DNA 45

Figure 13. Live-cell imaging of methylated DNA 46

Figure 14. Live-cell imaging of methylated DNA 47

Figure 15. Construct of pAB-mM 49

Figure 16. Construct of pATB-mM 50

Figure 17. Live-cell imaging of methylated DNA 51

Chapter 1

Introduction

1. DNA Methylation

1.1. Epigenetics

Epigenetic changes are genetic controls by factors that influence gene function and expression without altering the DNA sequence; these changes typically switch genes on or off. The major epigenetic factors involving changes in cellular events are DNA methylation and histone modification.[5] Histone modifications are various post-translational modifications, such as methylation, acetylation, phosphorylation, ubiquitination, and sumoylation, that typically influence transcriptional expression by modulating the chromatin structure. DNA methylation is a chemical process in which methyl groups are typically added to cytosines; it generally occurs during cell development and transcriptional repression[1] at the C-5 cytosine position to form 5-methylcytosine (5mC) via DNA methyltransferases (DNMTs) acting at CpG dinucleotides in mammals.[6]

1.2. DNA Methylation and DNA Methyltransferases

DNA methylation is a chemical process in which methyl groups are added to cytosine at the C-5 position to form 5mC via the DNMT family (Figure 1). The DNMT family includes DNMT1, DNMT2, DNMT3A, DNMT3B, and DNMT3L.[7] DNMT1 mainly maintains methyltransferases and preferentially targets hemimethylated DNA during DNA replication and repair.[8] DNMT2 methylates tRNA instead of DNA.[9] DNMT3A and DNMT3B specifically bind to non-methylated CpG dinucleotides (CpG islands) and perform methylation during cell development, particularly during DNA synthesis and replication.[7,8,10] DNMT3L lacks a catalytic domain, and, instead of methylating DNA, it scaffolds other methyltransferases to methyl group donors to increase their enzymatic efficiency.[10] Deficiency in any DNMT member could lead to DNA methylation failure,[7] which, in turn, promotes various diseases and tumor development. These findings reflect the significant and essential role of DNMTs in mammalian cell and embryonic development.[3]

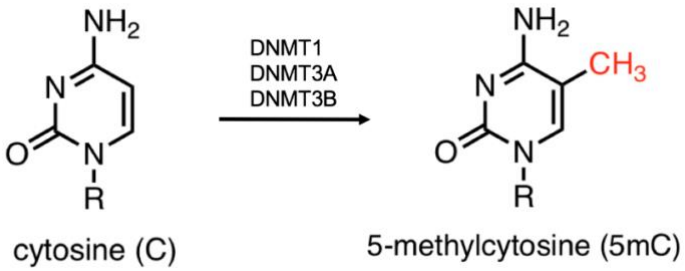


Figure 1. Cytosine (C) is methylated to 5-methylcytosine (5mC). DNA methylation is mainly performed by three DNA methyltransferases, DNMT1, DNMT3A, and DNMT3B (Song 2016. *Bioessays*).

1.3. DNA Demethylation

DNA methylation was initially regarded as an irreversible epigenetic modification until the discovery of 5-hydroxymethyl-2′-deoxycytidine (5hmC), an intermediate that facilitates DNA methylation.[11,12] DNA demethylation is indispensable in early embryonic and normal development,[13] and recent studies have proposed several mechanisms of DNA demethylation (**Fig. 2**). Oxidation of 5mC has been demonstrated to participate in DNA demethylation by the Ten-Eleven Translocation (TET) family of enzymes.[11,12] TET1, TET2, and TET3 hydroxylases sequentially oxidize 5mC to 5-hydroxymethyl-2′-deoxycytidine (5hmC), then to 5-formylcytosine (5fC), and finally to 5-carboxylcytosine (5caC).[12] Activation-induced cytosine deaminases (AIDs) have also been linked to DNA demethylase functions wherein thymine is subsequently replaced by cytosine during 5mC conversion, resulting in a T – G base pair mismatch. AID deficiency significantly affects DNA demethylation in primordial germ cells.[13] In all proposed DNA demethylation mechanisms described above, thymine DNA glycosylase (TDG) plays a very important role, converting intermediates, such as thymine, 5fC, and 5caC, back to native cytosine via the base excision repair (BER) pathway without undergoing methylation.[14-16] $TDG^{-/-}$ mice are non-viable, which suggests that TDG is indispensable in mammals.[17] The discovery of TDG revealed that reversible DNA methylation and demethylation is strictly regulated to maintain normal development and function in mammalian organisms. Knockout of *mbd4* in mice influenced parathyroid hormone-induced DNA demethylation, thereby suggesting that MBD4 is essential during DNA demethylation in mammalian cells.[18,19] A previous study also showed that DNMT1 is not involved in the conversion of

5mC to 5hmC in the DNA demethylation cycle.[20] 5hmC, which is considered an intermediate produced during the DNA demethylation cycle, is particularly abundant in nerve cells and ESCs and highly expressed in brain tissue.[21] Another recent study indicated that MBD from the *mecp2* gene not only binds to 5mC but also targets 5hmC.[22] These findings suggest that 5hmC may play a vital role in regulating DNA methylation in nerve cells; however, the role of 5hmC in DNA demethylation and the mechanism of DNA demethylation remain poorly understood and require further elucidation.[11,12]

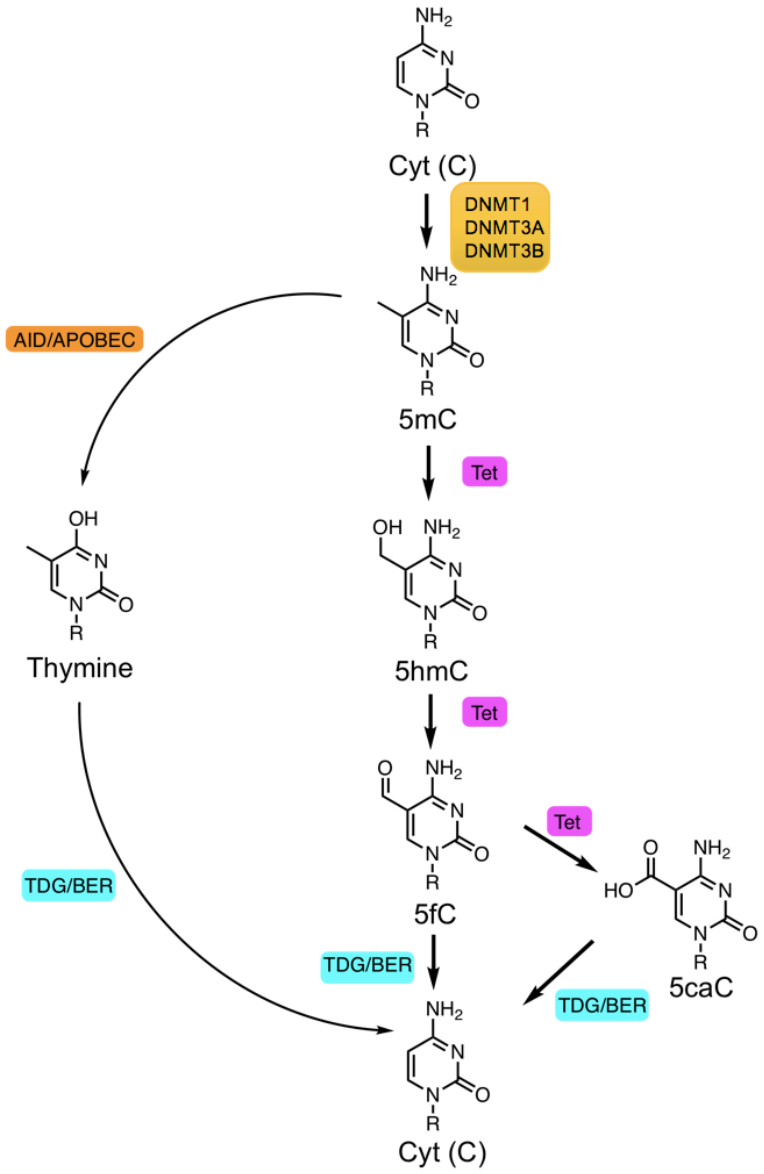


Figure 2. Proposed mechanisms of the DNA methylation and demethylation cycle. Cytosine is converted to thymine by activation-induced cytidine deaminase/apolipoprotein B mRNA-editing enzyme complex (AID/APOBEC). TET enzymes sequentially catalyze oxidation of 5mC to 5hmC, and then to 5fC, and then finally to 5caC. Thymine, 5fC, and 5caC are recognized by the TDG/BER pathway and converted back to unmodified cytosine (Moore et al. 2012. *Neuropsychopharmacology*).

1.4. DNA Methylation and Histone Modification

Some specific interactions have been reported between histone modification and DNA methylation. Histone H3 trimethylated at lysine 9 (H3K9me3) and lysine 27 (H3K27me3), as well as SUV39H1/2 and EZH2, have been identified to cooperate with some DNMT members in the histone and DNA methylation network.[23,24] The histone-lysine N-methyltransferase, SUV39H1/2, primarily trimethylates H3K9. Mutant *Suv39h1/2* shows a decrease in methylation DNA level and satellite repeats during embyogenesis.[23] Another histone-lysine methyltransferase, EZH2, which is a Polycomb Group (PcG) protein responsible for H3K9 and H3K27 methylation, recruits DMNTs to target some of its target gene promoters.[24] Upon the discovery of EZH2, downstream EZH2 effectors show a significant decrease in H3K9me3.[24] A transcriptional repressor, heterochromatin protein 1 (HP1), specifically binds methylated H3K9 and has been identified to be an intermediate promoting cooperation between DNA and histone methylation.[25] HP1 preferentially binds to H3K9me and targets DNMT3A through the methyltransferase G9A.[26] Collaboration of methylated H3K9 and HP1 has been identified as the connection between DMMT and H3K9me, thus further indicating that the HP1-H3K9 methylation system directly influences DNA methylation patterns.[27] G9A and G9A-like protein (GLP) are euchromatin-associated methyltransferases known to be SUV39H1/2 family members. G9A, which mainly performs monomethylation and dimethylation of H3K9 (H3K9me and H3K9me2), inhibits its target genes through collaboration between GLP, H3K9, and DNA methylation at the transcriptional level.[26,28,29] Knockout of both *G9a* and *GLP* show significantly reduced DNA

methylation levels among G9A and GLP downstream genes in embryonic stem cells.[26,28,29] Other evidence has also determined the crosstalk between DNA methylation and histone modification, in which both DMNT1 and G9A target H3K9me2 during DNA replication. G9A and H3K9 methylation is significantly delayed when DMNT1 is eliminated.[30,31] Other studies show that some genes are suppressed and H3K9 methylation is significantly decreased in some cancer cell lines after treatment of cells with 5-aza-2-deoxycytidine, which has been identified to specifically eliminate DNA methylation.[32] These previous studies clearly contribute to the topic at hand; however, many unknown links between DNA methylation and histone modifications require further investigation.

1.5. Methyl-CpG-binding domain proteins and MeCP2

Methyl-CpG-binding domain proteins (MBDs) provide a platform for networks between DNA methylation and histone modifications. MBDs comprise the human proteins MECP2, MBD1, MBD2, MBD3, and MBD4.[33] Each MBD member precisely binds to and interacts with methylated DNA and typically suppresses gene transcription.[33] Deletion of *mbd1* revealed that H3K9 methylation is regulated by the histone-lysine methyltransferase, SETDB1, through cooperation with MBD1.[8,34] MBD1 and MBD2 are capable of suppressing transcription downstream of methylated gene promoters.[35] Other SRA domains in plants can also recognize methylated DNA. Only two proteins in the SRA family can bind to methylated DNA in mammals: UHRF1, which is responsible for patterning DNA methylation during DNA synthesis, and UHRF2, which targets both 5mC and 5hmC.[36] MeCP2 is subject to X-inactivation and essential for normal function

in nerve cells.[37] MeCP2 has been identified to pattern histone deacetylation by targeting chromosomes during DNA methylation.[38] Most known MeCP2 gene mutations lead to Rett Syndrome, which is a genetic disease affecting intellectual disability in girls and women (men usually do not survive this mutation).[37] Rett Syndrome typically causes decreases in speech and motor function capabilities, seizures, and growth retardation.[37,39] Patients often suffer from skeletal diseases, including severe scoliosis and osteoporosis, in their childhood.[40-42] MeCP2 also plays dual roles in recruiting corepressor complexes to inhibit transcriptional regulation and promote transcriptional activation.[38,43-46]

1.6. DNA Methylation and Cancer Development

DNA methylation and demethylation are highly associated with cancer development. In general, genetic alterations are common in tumor cells, further promoting tumor progression and cancer development. Cooperation between DNA methylation and histone modification, especially bivalent chromatin patterning, significantly interferes with tumor progression by suppressing pro-differentiated genes; by contrast, disruption of epigenetic markers promotes malignant tumor growth in progenitor cells.[47,48] In cancer cells, low levels of methylation promote oncogene activation to affect tumor invasion and metastasis. During tumor progression, tumor suppressor genes are always hypermethylated and silenced when DNA demethylation loses function.[49] It is possible to lead to cancer development when genes undergo either hypermethylation or hypomethylation.[49] Such evidence indicates that the methylation level among specific genes, for example, tumor suppressor genes, must be strictly

regulated by several mechanisms of crosstalk between DNA methylation and histone modification to maintain normal development.

Epigenetic modifications are significant for mammalian development. Interference of epigenetic modifications at the transcriptional level may trigger malignant cell growth and lead to pathological phenotypes, such as Rett Syndrome, neuronal disorders, or carcinomas.[27-29] DNA methylation, as a significant regulator of transcriptional expression, is capable of repressing chromatin and affecting histones by collaborating with DNMTs, histone modification systems, and other cofactors to further influence the entire epigenome.[11] Analysis of DNA methylation promotes the understanding of epigenetic networks. However, to date, the roles of epigenetic markers, including those of DNA methylation and histone modification, in normal development and diseases, remain poorly understood despite many proposed mechanisms. Cooperating networks among DNA methylation and other epigenetic modifications involve complicated processes, which may occur individually or simultaneously, and rely on the collaboration of various factors. Many unknown components and mechanisms involved in this network have yet to be elucidated.

2. Fluorescent Proteins

2.1. Discovery of Green Fluorescent Protein

In the early 1960s, Osamu Shimomura and Frank Johnson, who worked at the Friday Harbor Laboratories of the University of Washington, were the first scientists to extract and purify a calcium-dependent bioluminescent protein from the jellyfish *Aequorea victoria*; this protein was named aequorin and produced a blue light emission upon catalyzing the oxidization of a coelenterazine substrate with a calcium cofactor.[50] During aequorin extraction, a protein that lacked blue light emission but could exhibit bright green fluorescence under ultraviolet (UV) excitation was found.[50] This autofluorescent protein was later identified to cooperate with aequorin by absorbing the excited light from aequorin and emitting green light during the conversion of calcium-induced luminescence signals into green luminescence.[51] Due to this characteristic, the autofluorescent protein was eventually called green fluorescent protein (GFP).[50,51] While the GFP initially isolated from *Aequorea* was discovered and identified several decades ago, its significance as a potent and robust molecular probe has yet to be fully realized. In fact, this protein is still commonly used in GFP-fused probes to monitor transcriptional expression in bacteria.

2.2. Derivatives of Green Fluorescent Protein

The stability and brightness of GFP is generally affected by mutations to its structure, further leading to a decrease in its quantum yield and maturation. Native GFP is

quite stable and produces significant fluorescence. However, its maximum excitation is very close to the excitation scope of UV light, which typically causes damage to mammalian cells; thus, the protein is unsuitable for live-cell imaging. When exposed to UV light, the serine of GFP at position 65 is replaced by threonine (S65T), resulting in maximum excitation at 488 nm.[52] This mutation produces the very popular GFP variant, enhanced GFP (EGFP), which is among the most popular fluorescent probes used in cell labeling experiments today.[52] Another mutation to the GFP structure, where threonine at position 203 is replaced with tyrosine (T203Y), generates another GFP variant called yellow fluorescent protein (YFP), which features excitation and emission wavelengths of approximately 20 nm longer than those of the original protein.[53-55] Unfortunately, enhanced YFP (EYFP) shows poor stability and loses fluorescence on account of its sensitivity to chloride ions and acidic environments.[53] Improved versions of the protein for better environmental adaption have been investigated. The blue variant of GFP (BFP) is produced by replacement of tyrosine at position 66 with histidine (Y66H) in the GFP structure.[56-58] BFP features an excitation peak of around 380 nm and an emission peak of 448 nm. However, the enhanced version of BFP (EBFP) presents only around 25% of the brightness of EGFP, and its excitation nearly overlaps with that of UV light; thus, the practical applications of BFP are limited.[56-58] When the tyrosine at position 66 is replaced with tryptophan (Y66W) in the GFP structure, a variant of GFP that possesses cyan fluorescence is obtained. Cyan fluorescent protein (CFP) features two excitation peaks at 433 and 445 nm and two maximum emissions at 475 and 503 nm.[59] Improved versions of CFP (ECFP) possess higher brightness and photostability, but, because ECFP has only

around 40% of the brightness of EGFP, it is still not a useful tool in cell labeling experiments.[59-61] Moreover, as the wavelengths of EBFP and ECFP are close to those of UV light, optically improved and specialized instruments are required for imaging; as such, these proteins are not widely used in general microscope systems.[61] Despite these disadvantages, however, EBFP and, especially, ECFP are popular in fluorescence resonance energy transfer (FRET) experiments. ECFP shows a minor peak at around 440 nm, which enables it to better resist photobleaching compared with other BFP derivatives.[62,63] An improved version of ECFP has a higher quantum yield and simpler excitation mode than previous versions. Coupling improves ECFP and EYFP, efficiently reduces undesired signals, and presents higher contrast. Thus, ECFP and EYFP are the most popular combination of fluorescence proteins in FRET-based experiments, and many researchers have focused developing better cyan variant probes.

2.3. Discovery of Red Fluorescent Protein

Although *A. victoria* jellyfish GFP and its derivatives possess many advantages in multi-color cell labeling experiments, the construction and characteristics of these proteins cannot be used to successfully monitor specific organs or the entire organism, which are significantly more transparent to red light.[64] Thus, the demand for fluorescent proteins that feature the same quantum yield and photostability under red-shifted excitation and emission wavelengths has steadily increased. Unfortunately, the development of red fluorescent protein (RFP) extracted from *A. victoria* jellyfish has not been successful, and, thus, some researchers have turned to explore and investigate other

species, such as the tropical reef coral Anthozoa, which can also produce autofluorescence.[64,65] The first coral-derived fluorescent protein (FP) was extracted from *Discosoma striata* and called DsRed.[64] After full maturation, DsRed exhibits a maximum excitation peak at 558 nm and a maximum emission peak at 583 nm.[66,67] However, maturation of DsRed is slow and produces green fluorescence emission during chromophore maturation. DsRed features an obligate tetramer that easily forms large aggregates in live cells. Moreover, the FP is toxic when expressed in live cells.[67] Owing to these disadvantages, the applications of DsRed are relatively restricted compared with those of GFP and other GFP variants. Mutagenesis removed several amino acids from the protein and formed the second generation of DsRed called DsRed2. DsRed2 presents faster chromophore maturation and reduces protein aggregation and toxicity, making it more compatible than DsRed for applications in live cells. However, DsRed2 still easily forms oligomeric tetramers.[68,69] Another RFP, HcRed, was isolated from *Heteractis crispa*.[70] HcRed presents oligomeric characteristics and tends to form dimers; it also easily aggregates with DsRed and produces much less fluorescence than other RFPs.[70] Thus, unless their oligomeric features are improved, DsRed and HcRed are not considered suitable options for protein labeling experiments in live cells.

2.4. Derivatives of Red Fluorescent Protein

In their natural state, oligomeric forms are not commonly been observed in *A. victoria* GFP. GFP is largely free from aggregations and typically considered a monomeric FP when expressing in mammalian cells. Most FPs tend to aggregate and act

as oligomeric structures, especially RFP and its variants. When most RFPs are expressed in cellular compartments, the protein concentration may be high enough to trigger oligomerization.[71] Oligomeric aggregation is a potential issue when performing FP-labeling experiments, particularly FRET, because such aggregation can introduce artifacts. Therefore, improved monomeric or dimeric RFPs are important directions for further development. The first generation of monomeric DsRed proteins (mRFP1), which was created by around 30 amino acid replacements to the DsRed protein structure, is more compatible with cell labeling experiments because of its monomeric characteristic and fast chromophore maturation.[71] Unfortunately, this new derivative loses fluorescence and photobleaches quite rapidly compared with native RFPs.[71] In a direct evolutionary approach, monomeric RFPs were successfully derived through direct mutations targeting the chromophore Q66 and Y67 residues, which are regarded as critical residues possibly determining the color of mRFP1.[72] These new monomeric derivatives were named mFruits and generally featured maximum emissions between 540 and 610 nm.[72] For example, one mFruit family member, mCherry, has an emission peak at 610 nm and features around 50% of the brightness of EGFP; this protein high photostability and is commonly used in cell imaging.[72,73] Another mFruit member, mApple, features high brightness, photostability, and fast maturation, hence its popularity in protein labeling experiments.[73] The appearance of new RFP derivatives significantly and successfully fills the gap between the mostly red-shifted jellyfish GFPs and coral reef RFPs.[73] The discovery of these new FP derivatives further promotes the development of improved variants with higher extinction coefficients, quantum yields, and photostability and faster

maturation. Interestingly, the tendency to develop RFP derivatives has turned toward understanding and dissecting far-RFPs (far-red FPs). RFPs and far-red FPs are necessary for application in mammalian systems because the longer emission wavelengths of RFP derivatives can more easily penetrate tissues or organs than other FPs and their derivatives. For instance, mKate and the improved versions mKate and mNeptune can achieve emission wavelengths of around 650 nm.[74,75] The applications of RFPs and far-red FPs are being given greater consideration in studies in biomedical science. Researchers have focused on the creation and discovery of improved RFPs and far-red FP derivatives to afford genetically encoded FPs that can be formally applied to detect desired tissues or organs in the human body.

2.5. Fluorescent Protein-based Biosensors

FP-containing biosensors have grown in popularity quite rapidly. These FPs are engineered to produce huge numbers of colored and improved variants. With the rapid evolution of FP technology, genetically encoded FPs, such as CFP, BFP, YFP, and RFP, and their derived forms have become very powerful analytical tools. These optical biosensors are capable of specifically and clearly not only labeling proteins but also detecting signal transductions in live cells without chemical, synthetic, or antibody-labeled tags. The first genetically encoded fluorescent probe was developed to detect the activity of calcium ions and calmodulin in HeLa cells from Dr. Atsushi Miyawaki's laboratory.[76] Thereafter, various FP-based biosensors have been rapidly developed for investigating different cellular events. For example, FP-based biosensors have been

designed to monitor enzymes acting as a kinase or second-messenger in signaling pathways or enzymes participating in oxidation – reduction reactions or sensing environmental changes.[77] These optical probes have been applied in ratiometric FRET, fluorescence lifetime imaging microscopy (FLIM), or two-photon imaging for deeper studies into molecular or cellular events.[77] As a reliable and powerful tool, genetically encoded FP-based biosensors are important for dissecting cellular functions in organs or live animals.

2.6. Dimerization-dependent Fluorescent Protein-based Biosensors

In Alford *et al.*'s 2012 paper, the idea of dimerization-dependent RFP (ddRFPs) biosensors was first proposed (Figure 3). One domain of ddRFPs has very dim fluorescence while the other domain has no fluorescence. Instead, this domain increases fluorescence of the first domain when the two dimerize.[78] Two ddGFPs and ddYFPs have been engineered for dimeric optimization to monitor mitochondria and the endoplasmic reticulum.[79] However, owing to the high affinity of ddGFPs and ddYFPs, these probes cannot clearly differentiate the associated and disassociated states of proteins and, therefore, produce undesired signals. In Ding *et al.*'s 2015 paper, the ddFP approach was extended for use in ratiometric dimerization-dependent GFP and RFP. Similarly, a domain was engineered to have no fluorescence but increase the fluorescence of other partner domains.[80] For example, a nuclear localization sequence was fused to the domain that had no fluorescence (NLS-dB), a nuclear export sequence was fused to RFP (NLS-RA), and a nuclear localization sequence was fused to GFP (NLS-GA). Red fluorescence

was observed when dimeric NLS-dB and NES-RA were formed and linked by a caspase-3 substrate sequence outside of the nucleus. After treatment with caspase, a caspase-3 substrate sequence was cleaved and green fluorescence was observed when dimeric NLS-dB and NLS-GA were formed upon NLS-dB entering the nucleus. Besides monitoring protease activity, this approach has also been used to study calcium-dependent interactions and PKA activity. Previous research shows the applications of ddFP-based probes to further the investigation of protein – protein interactions and provide a novel strategy for developing new FP-based biosensors.

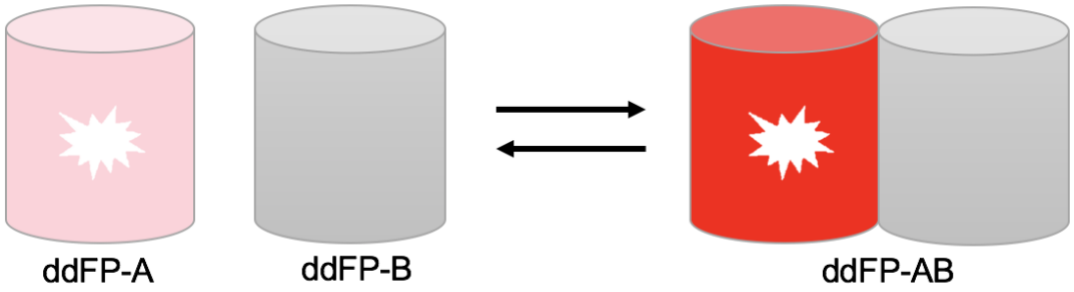


Figure 3**.** Conceptual schematic of dimerization-dependent fluorescent protein (ddFP) technology. Dimeric state of FP (ddFP-AB) is reversibly formed by weak or non-fluorescent monomers (ddFP-A or ddFP-B) (Alford et al. 2012. *Chem Biol*).

3. Fluorescent Biosensors for DNA Methylation

DNA methylation participates in cell division and development and is essential in embryogenesis. Abnormal DNA methylation leads to various diseases, such as cancers and neurological disorder.[3-5] Techniques to detect DNA methylation are in high demand in the biomedical and pharmaceutical science fields. Interest in the mechanism of DNA methylation during cell division has steadily grown, and developing a probe that can reduce the interference of undesirable signals and detect methylated DNA is important. Several probes fusing specific binding domains have been developed to understand DNA methylation patterns. For example, MBD could be fused to EGFP to monitor DNA methylation patterns in live cells.[81] In addition, DNA methylation could be specifically detected by a chemical fluorogenic probe using a fluorogenic switching mechanism in live cells.[82] To date, no FP-based biosensor that can clearly illuminate methylated DNA in high contrast is yet available, and most probes require further optimization. In this research project, to clearly detect DNA methylation patterns by using FP-based biosensors, a ddRFP probe was applied instead of a GFP-based probe, which may produce undesired background nice. The design of a ddRFP probe described in this work is expected to specifically and clearly target the methylated DNA to enable further examination the differential DNA methylation levels in human cells.

Chapter 2

DNA methylation Indicators Based on Dimerization-Dependent Red Fluorescent Proteins

2.1. Introduction

DNA methylation is essential in cell division and development. Abnormal DNA methylation leads to various diseases, such as cancers and neurological disorder.[3-5] Interest in the mechanism of DNA methylation during cell division has steadily increased. Thus, developing a probe that can specifically and clearly illuminate methylated DNA is important. Unfortunately, no FP-based biosensor that can clearly detect methylated DNA with high contrast is yet available, and most of the available probes require further optimization to expand their applications. In this research project, ddRFP probes were developed to clearly detect DNA methylation patterns via an FP-based biosensor. The dimerization characteristic of ddRFP biosensors produces a fluorescence switching mechanism that can clearly differentiate associated and dissociated states, which means methylated and non-methylated DNA may be clearly examined and differentiated. An appropriate design of a ddRFP probe is expected to enable the successful patterning of methylated DNA and examination of differential DNA methylation levels in human cells.

2.2. Experimental Section

2.2.1. *Materials, Reagents, and General Methodology*

Synthetic DNA oligonucleotides for cloning and construction were purchased from Eurofin Genomics (Louisville, KY, USA), and restriction endonucleases were purchased from New England Biolabs (Ipswich MA, USA) or Thermo Scientific Fermentas (Vilnius, Lithuania). Polymerase chain reaction (PCR) products and digestion were purified by gel electrophoresis and extracted by Syd Laboratories gel extraction columns (Malden MA, USA). DNA plasmid was purified by Syd Laboratories mini-prep columns. DNA sequencing analysis was carried out by Eurofin Genomics (Louisville, KY, USA), and live-cell microscope imaging was performed through confocal microscopy (Leica) in Dr. Huiwang Ai's laboratory at the Department of Molecular Physiology and Biological Physics, University of Virginia, School of Medicine.

2.2.2. *Plasmid Construction*

pcDNA3-NLS-dA-MBD-NLS-dB-MBD and pcDNA3-NLS-dA-MBD-T2A-NLS-dB-MBD

PCR was conducted to amplify the DNA fragments of the designed RFP sensor constructs, designed EGFP sensor construct, and MBDs derived from the human MeCP2 gene. PCR was performed to amplify the domain A fragment using oligos dRA-F2 and dRA-R and single polypeptide FPX biosensor for calcium ion plasmid (Addgene #60887) as the template. PCR to amplify domain A with a nuclear localization sequence fragment was performed by using oligos dRA-F1 and dRA-R and single polypeptide FPX

biosensor for calcium ion plasmid (Addgene #60887) as the template. PCR was performed to amplify domain B by using oligos NLS-dB-F2 and NLS-dB-R and single polypeptide FPX biosensor for calcium ion plasmid (Addgene #60887) as the template. PCR to amplify domain B with a nuclear localization sequence fragment was performed by using oligos F-NLS-dB-F1 and NLS-dB-R and the amplified domain B as the template. PCR to amplify domain B with the nuclear localization sequence and T2A peptide fragment were first carried out by using oligos T2A-NLS-dB-F2 and NLS-dB-R and domain B as the template and then next with oligos T2A-NLS-dB-F1 and NLS-dB-R and the first PCR product as the template. PCR was performed to amplify MBDs for domain A fusion with the nuclear localization sequence (NLS-dA) by using oligos hMBD-F and hMBD-R and pDEST-hMeCP2-GFP plasmid (Addgene #48078) as the template. PCR to amplify MBDs for domain B fusion with the nuclear localization sequence (NLS-dB) was performed by using oligos dB-hMBD-F and F-dB-hMBD-xbaI-R and pDEST-hMeCP2-GFP plasmid (Addgene #48078) as the template. PCR to amplify MBDs for domain B fusion with T2A peptide and the nuclear localization sequence (T2A-NLS-dB) by using oligos dB-hMBD-F and F-dB-hMBD-xhoI-R and pDEST-hMeCP2-GFP plasmid (Addgene #48078) as the template. Overlap PCR to amplify NLS-dA fusing with MBDs (NLS-dA-MBD) was performed by using oligos dRA-F1 and hMBD-R and the PCR product NLS-dA and MBDs for NLS-dA as templates. Overlap PCR to amplify NLS-dB fused with MBDs (NLS-dB-MBD) was performed by using oligos F-NLS-dB-F1 and F-dB-hMBD-xbaI-R and the PCR product NLS-dB and MBDs for NLS-dB as templates. Overlap PCR to amplify T2A-NLS-dB fused with MBDs

(T2A-NLS-dB-MBD) was performed by using oligos T2A-NLS-dB-F1 and F-dB-hMBD-xhoI-R and PCR product NLS-dB and MBDs for T2A-NLS-dB as templates. The PCR product NLS-dA-MBD was digested with KpnI and EcoRI and then ligated into a predigested pcDNA3 expression vector to afford pcDNA3-NLS-dA-MBD. The ligation product was used to transform *Escherichia coli* DH10B competent cells, which were then plated on Luria–Bertani (LB) broth agar plates supplemented with ampicillin (50 ug/mL) and arabinose (0.2 g/mL). Colonies were picked and cultured with LB broth, and plasmids were extracted and preliminarily confirmed by digestion with KpnI and EcoRI; further validation was performed by Eurofin Genomics. The PCR product NLS-dB-MBD was digested with XhoI and XbaI and then ligated into a predigested pcDNA3-NLS-dA-MBD expression vector to afford pcDNA3-NLS-dA-MBD-NLS-dB-MBD. To generate pcDNA3-NLS-dA-MBD-NLS-dB-MBD, the ligation product was used to transform *E. coli* DH10B competent cells, which were then plated on LB broth agar plates supplemented with ampicillin (50 ug/mL) and arabinose (0.2 g/mL). Colonies were picked and cultured with LB broth, and plasmids were then extracted and preliminarily confirmed by digestion with XhoI and XbaI; further validation was performed by Eurofin Genomics. The PCR product T2A-NLS-dB-MBD was digested with EcoRI and XhoI and then ligated into a predigested pcDNA3-NLS-dA-MBD expression vector to afford pcDNA3-NLS-dA-MBD-T2A-NLS-dB-MBD. To generate pcDNA3-NLS-dA-MBD-T2A-NLS-dB-MBD, the ligation product was used to transform *E. coli* DH10B competent cells, which were then plated on LB broth agar plates supplemented with ampicillin (50 ug/mL) and arabinose (0.2 g/mL). Colonies were picked and cultured with

LB broth, and plasmids were extracted and preliminarily confirmed by digestion with EcoRI and XhoI. Further validation was performed by Eurofin Genomics.

pcDNA3-NLS-dAt-MBD

PCR was carried out to amplify domain A with a stop codon fragment by using oligo dRA-F2 and dRA-R and single polypeptide FPX biosensor for calcium ion plasmid (Addgene #60887) as the template. PCR was carried out to amplify domain A with a nuclear localization sequence and stop codon fragment by using dRA-F1 and dRA-R and single polypeptide FPX biosensor for calcium ion plasmid (Addgene #60887) as the template. PCR to amplify MBDs for domain A with a nuclear localization sequence and stop codon fragment was performed by using oligos hMBD-F and F-EGFP-hMBD-R and pDEST-hMeCP2-GFP plasmid (Addgene #48078) as the template. Overlap PCR was performed to amplify NLS-dAt fused with MBDs (NLS-dAt-MBD) by using oligos dRA-F1 and F-EGFP-hMBD-R and the PCR products NLS-dAt and MBDs for NLS-dAt as templates. The PCR product NLS-dAt-MBD was digested with KpnI and EcoRI and then ligated into a predigested pcDNA3 expression vector to afford pcDNA3-NLS-dAt-MBD. The ligation product was used to transform *E. coli* DH10B competent cells, which were then plated on LB broth agar plates supplemented with ampicillin (50 ug/mL) and arabinose (0.2 g/mL). Colonies were picked and cultured with LB broth, and plasmids were extracted and preliminarily confirmed by digestion with KpnI and EcoRI. Further validation was performed by Eurofin Genomics.

pcDNA3-NLS-EGFP-MBD

PCR reaction was carried out to amplify the EGFP fragment by using oligos NLS-EGF-F2 and NLS-EGF-R and pBAD-EGFP as the template. PCR was performed to amplify EGFP fused with the nuclear localization sequence fragment (NLS-EGFP) by using oligos dRA-F1 and NLS-EGF-R and the PCR product EGFP as the template. PCR was performed to amplify MBDs for EGFP by using oligos EGFP-hMBD-F and F-EGFP-hMBD-R as the template. Overlap PCR was performed to amplify EGFP fused with MBDs for EGFP (NLS-EGFP-MBD) by using oligos dRA-F1 and F-EGFP-hMBD-R and the PCR products NLS-EGFP and MBDs for EGFP as templates. The PCR product NLS-EGFP-MBD was digested with KpnI and EcoRI and then ligated into a predigested pcDNA3 expression vector to afford pcDNA3-EGFP-MBD. The ligation product was used to transform *E. coli* DH10B competent cells, which were then plated on LB broth agar plates supplemented with ampicillin (50 ug/mL) and arabinose (0.2 g/mL). Colonies were picked and cultured with LB broth, and plasmids were extracted and preliminarily confirmed by digestion with KpnI and EcoRI. Further validation was performed by Eurofin Genomics.

pcDNA3-NLS-dA-mMBD-NLS-dB-mMBD and pcDNA3-NLS-dA-mMBD-T2A-NLS-dB-mMBD

PCR was performed to amplify the domain A fragment by using oligos dRA-F2 and dRA-R and single polypeptide FPX biosensor for calcium ion plasmid (Addgene #60887) as the template. PCR was performed to amplify domain A with a nuclear

localization sequence fragment by using oligos dRA-F1 and dRA-R and single polypeptide FPX biosensor for calcium ion plasmid (Addgene #60887) as the template. PCR was performed to amplify domain B by using oligos NLS-dB-F2 and NLS-dB-R and single polypeptide FPX biosensor for calcium ion plasmid (Addgene #60887) as the template. PCR was performed to amplify domain B with a nuclear localization sequence fragment by using oligos F-NLS-dB-F1 and NLS-dB-R and amplification of domain B as the template. PCR was performed to amplify domain B with a nuclear localization sequence and T2A peptide fragment were first carried out by using oligos T2A-NLS-dB-F2 and NLS-dB-R and domain B as the template and then by using oligos T2A-NLS-dB-F1 and NLS-dB-R and the first PCR product as the template. PCR was performed to amplify R111G mutant MBDs (mMBDs) by using oligos Temp-R111G-mMBD-F and Temp-R111G-mMBD-R and pDEST-hMeCP2-GFP plasmid (Addgene #48078) as the template. PCR was performed to amplify mMBDs for domain A fusion with the nuclear localization sequence (NLS-dA) by using oligos hMBD-F and hMBD-R and mMBDs as the template. PCR was performed to amplify mMBDs for domain B fusion with the nuclear localization sequence (NLS-dB) by using oligos dB-hMBD-F and F-dB-hMBD-xbaI-R and mMBDs as the template. PCR was performed to amplify mMBDs for domain B fusion with T2A peptide and the nuclear localization sequence (T2A-NLS-dB) by using oligos dB-hMBD-F and F-dB-hMBD-xhoI-R and mMBDs as the template. Overlap PCR was performed to amplify NLS-dA fused with mMBDs (NLS-dA-mMBD) by using oligos dRA-F1 and hMBD-R and PCR product NLS-dA and mMBDs for NLS-dA as templates. Overlap PCR was performed to amplify NLS-dB fused with mMBDs (NLS-

dB-MBD) by using oligos F-NLS-dB-F1 and F-dB-hMBD-xbaI-R and PCR product NLS-dB and mMBDs for NLS-dB as templates. Overlap PCR was performed to amplify T2A-NLS-dB fusing with mMBDs (T2A-NLS-dB-mMBD) by using oligos T2A-NLS-dB-F1 and F-dB-hMBD-xhoI-R and PCR product NLS-dB and mMBDs for T2A-NLS-dB as templates.

The PCR product NLS-dA-mMBD was digested with KpnI and EcoRI and then ligated into a predigested pcDNA3 expression vector to afford pcDNA3-NLS-dA-mMBD. The ligation product was used to transform *E. coli* DH10B competent cells, which were then plated on LB broth agar plates supplemented with ampicillin (50 ug/mL) and arabinose (0.2 g/mL). Colonies were picked and cultured with LB broth, and plasmid were extracted and preliminarily confirmed by digestion with KpnI and EcoRI. Further validation was performed by Eurofin Genomics. The PCR product NLS-dB-mMBD was digested with XhoI and XbaI and then ligated into a predigested pcDNA3-NLS-dA-mMBD expression vector to afford pcDNA3-NLS-dA-mMBD-NLS-dB-MBD. To generate pcDNA3-NLS-dA-mMBD-NLS-dB-mMBD, the ligation product was used to transform *E. coli* DH10B competent cells, which were then plated on LB broth agar plates supplemented with ampicillin (50 ug/mL) and arabinose (0.2 g/mL). Colonies were picked and cultured with LB broth, and plasmids were extracted and preliminarily confirmed by digestion with XhoI and XbaI. Further validation was performed by Eurofin Genomics. The PCR product T2A-NLS-dB-mMBD was digested with EcoRI and XhoI and then ligated into a predigested pcDNA3-NLS-dA-mMBD expression vector to afford pcDNA3-NLS-dA-mMBD-T2A-NLS-dB-mMBD. To generate pcDNA3-NLS-dA-

mMBD-T2A-NLS-dB-mMBD, the ligation product was used to transform *E. coli* DH10B competent cells, which were then plated on LB broth agar plates supplemented with ampicillin (50 ug/mL) and arabinose (0.2 g/mL). Colonies were picked and cultured with LB broth, and plasmids were extracted and preliminarily confirmed by digestion with EcoRI and XhoI. Further validation was performed by Eurofin Genomics.

2.2.3. *Bacterial Transformation*

Overnight liquid cultures was used for obtain electro-competent cells, which were diluted to 1:100 or 1:50 in 500 or 1000 mL of 2xYT media with ampicillin (100 mg/mL), shaken at 350 rpm, and incubated in 37 °C until OD_{600} = 0.8. The cells were washed three times by centrifugation at 8000 rpm for 5 min and resuspension in 15% glycerol, after which they were finally aliquoted, separated into tubes, and rapidly frozen in either a CO_2–acetone or liquid nitrogen bath. Tubes containing electro-competent cells were stored at –80 °C for until use.

2.2.4. *Electro-competent Cell Preparation*

2xYT was used in all solid and liquid culture media containing the appropriate antibiotics and grown for 12 h. All transformed cells were electroporated at 1.5 kV and outgrown in SOC media at 37 °C. The transformed cells were shaken at 350 rpm, incubated at 37 °C, and grown on solid agar media at 37 °C for 12 – 48 h.

2.2.5. *Plasmid Large-Prep*

Transformed plasmid *E. coli* DH10B cells were harvested from 100–200 mL cultures by centrifuging for 15 min at 4700 rpm. The supernatant was discarded and the pellet was resuspended in 5 mL of P1 solution in a 50 mL conical tube. Then, 5 mL of P2 solution was added to the tube, which was subsequently inverted for mixing at room temperature for 2 min. Exactly 5 mL of PX solution was added to the tube, which was then inverted for mixing and centrifuged 4700 rpm for 15 min at 4 °C. The supernatant was transferred into a new 50 mL conical tube. An equal volume of isopropanol was added, and the 50 mL tube was centrifuged at 4700 rpm for 30 min at 4 °C. The supernatant was discarded without disrupting the pellet, and the tube was inverted to completely remove the solvent. Thereafter, the pellet was dissolved in 0.9 mL of P1 by pipetting, and the mixture was transferred into a 2 mL Eppendorf tube. Another 0.9 mL of 4 °C PY solution was added to the Eppendorf tube, which was then vortexed and centrifuged at 15000 rpm for 10 min at 4 °C. The supernatant was discarded without disrupting the pellet, and the tube was inverted to completely remove the solvent. Thereafter, the pellet was dissolved in 0.1 mL of P1 solution and 0.8 mL of ddH_2O by gentle vortexing. Exactly 0.9 mL of PhOH:Chl was added to the Eppendorf tube, and the tube was vortexed vigorously and then centrifuged 15000 rpm for 3 min at 4°C. The clear aqueous top layer was collected and transferred into a new 2 mL Eppendorf tube. An equal volume of chloroform was added to the tube, which was then vortexed and centrifuged at 15000 rpm for 3 min at 4 °C. The clear aqueous top layer was collected and transferred into a new 2 mL Eppendorf tube. Exactly 180 µL of 3M NaOAc (pH 5.3)

and 900 μL of cold isopropanol were added to the tube and mixed. The tube was centrifuged at 15000 rpm for 3 min at 4 ℃. The supernatant was discarded without disrupting the pellet, and 1 mL of cold 70% EtOH was added to the tube to rinse the pellet. The supernatant was removed once more, and the Eppendorf tube was inverted and dried. The pellet was dissolved in 200–400 μL of ddH_2O. The P1 and P2 solutions used for large-prep are identical to those used for mini-prep plasmid extraction. All listed regents were prepared before the experiment.

2.2.6. *Cell Culture*

Human embryonic kidney 293T (HEK293T) cells were cultured in Dulbecco's Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) at 37 ℃. MDA-MB-468 breast cancer cells were cultured in Leibovitz's L-15 medium supplemented with 20% FBS at 37 ℃ without CO_2.

2.2.7. *Live-Cell Fluorescence Imaging*

HEK293T cells were transfected with *pcDNA3-NLS-dA-MBD-NLS-dB-MBD* or *pcDNA3-NLS-dA-MBD-T2A-NLS-dB-MBD* or *pcDNA3-NLS-dAt-MBD* or *pcDNA3-NLS-EGFP-MBD* or *pcDNA3-NLS-dA-mMBD-NLS-dB-mMBD* or *pcDNA3-NLS-dA-mMBD-T2A-NLS-dB-mMBD* by using PEI or Lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's protocol. After incubating transfected cells at 37 ℃ for 24 h, the cells were washed with Dulbecco's phosphate-buffered saline (DPBS) and then incubated in fresh DMEM containing 10% FBS for further imaging. Microscopic images

of the cells were recorded using a confocal laser-scanning microscope (Leica) with excitation at 488 nm for *pcDNA3-NLS-EGFP-MBD* and at 532 nm for *pcDNA3-NLS-dA-MBD-NLS-dB-MBD* or *pcDNA3-NLS-dA-MBD-T2A-NLS-dB-MBD* or *pcDNA3-NLS-dAt-MBD* or *pcDNA3-NLS-dA-mMBD-NLS-dB-mMBD* or *pcDNA3-NLS-dA-mMBD-T2A-NLS-dB-mMBD*. To inhibit DNA methylation, HEK293T cells were incubated with 5-AzadC in DMEM (2 mL) after transfection for 5 h and then further incubated at 37 °C for 24 h and washed with DPBS. The cells were subsequently incubated in fresh DMEM containing 10% FBS for further imaging. Microscopic images of the cells were obtained following the protocol described above. To inhibit alpha-ketoglutarate-dependent enzymes, HEK293T cells were incubated with N-oxalylglycine (NOG) in DMEM (2 mL) after transfection for 5 h and then further incubated at 37 °C for 24 h and washed with DPBS. The cells were subsequently incubated in fresh DMEM containing 10% FBS for further imaging. Microscopic images of the cells were obtained following the protocol described above. All listed reagents were prepared before the experiments.

2.3. Results and Discussion

2.3.1. *Design of dimerization-dependent red fluorescent proteins probes for detecting methylated DNA*

As interest in the mechanism of DNA methylation during cell division has steadily grown, developing a probe that can decrease the interference of undesired signals is necessary to accurately detect methylated DNA. To achieve this goal, GFP probes were

not considered into this research because the proteins involved produce undesirable background noise. Instead, ddRFP probes developed from the Alford SC *et al.* 2012 paper were applied in this study.[78] Here, two domains, A and B, were used in the ddRFP probes to detect methylated DNA. Domain A was derived from dTomato and shows fluorescence; domain B was also derived from dTomato but does not possess fluorescence. Instead, domain B acts forms a heterodimer with domain A and enhances its fluorescence intensity. The Alford SC *et al* 2012 paper revealed a 10-fold increase in fluorescence intensity upon dimerization of domains A and B.[78] We thus intended to develop a fluorescent probe that can specifically detect methylated DNA based on this ddRFP system (Figure 4). MBDs derived from the human MeCP2 gene and reported to be able to specifically bind to methylated DNA were fused to domains A and B. As we hoped that our probe could specifically target methylated DNA in the nucleus, a nuclear localization sequence was also fused to domains A and B. The first construct of the methylated DNA sensor aimed to link NLS-dA-MBD and NLS-dB-MBD together for expression in a pcDNA3 vector (Figure 5). However, owing to the oligomeric characteristic of ddRFPs, domains A and B may not show 1:1 expression in the first construct and form aggregates in the nucleus. Thus, the T2A peptide, which can achieve self-cleaving by ribosome skipping, was fused between domains A and B to afford 1:1 expression in the second construct (Figure 6). We envisioned these two constructs to show different localization and fluorescence intensities.

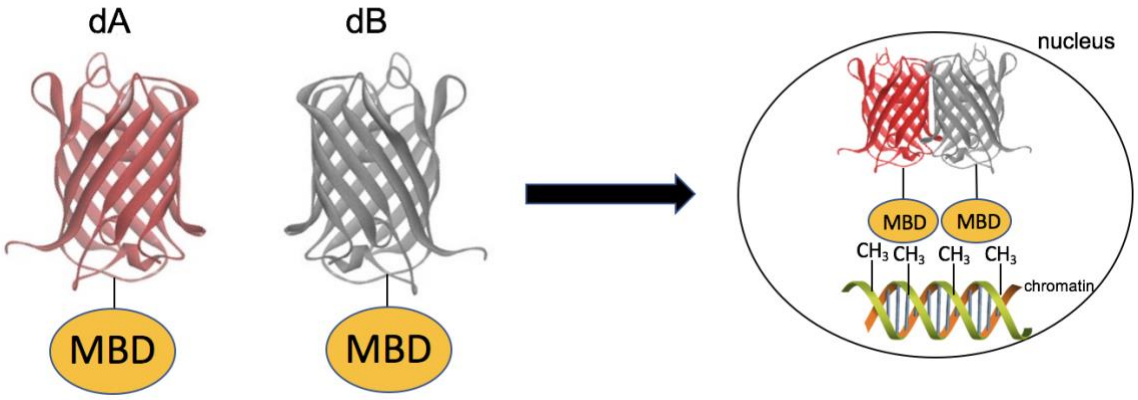


Figure 4. Conceptual scheme of ddRFP-based probe for DNA methylation. The monomeric state of domain A shows very dim fluorescence while the monomeric state of domain B has no fluorescence. However, domain B increases the fluorescence of domain A upon dimerization. This figure illustrates our strategy wherein methyl-CpG-binding domain proteins (MBDs) are fused with domains A and B, after which the domains are dimerized and used to target methylated DNA within the nucleus of a cell.

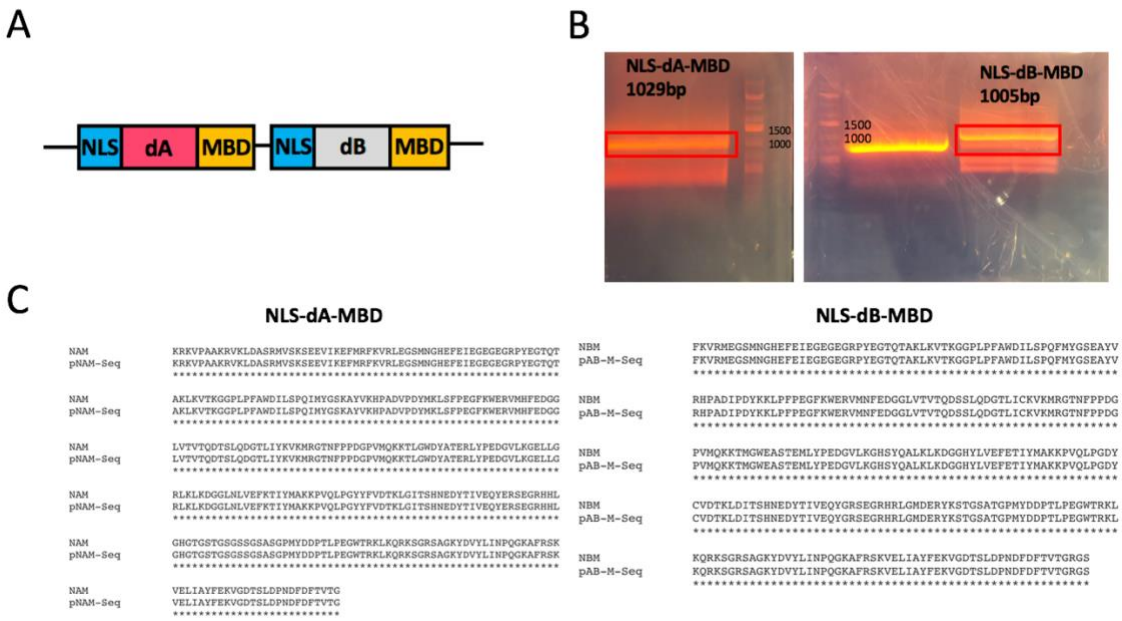


Figure 5. Construct of pAB-M. (A) Design of the pAB-M construct. (B) PCR results were developed by gel electrophoresis. (C) The pAB-M construct was aligned and validated by sequencing.

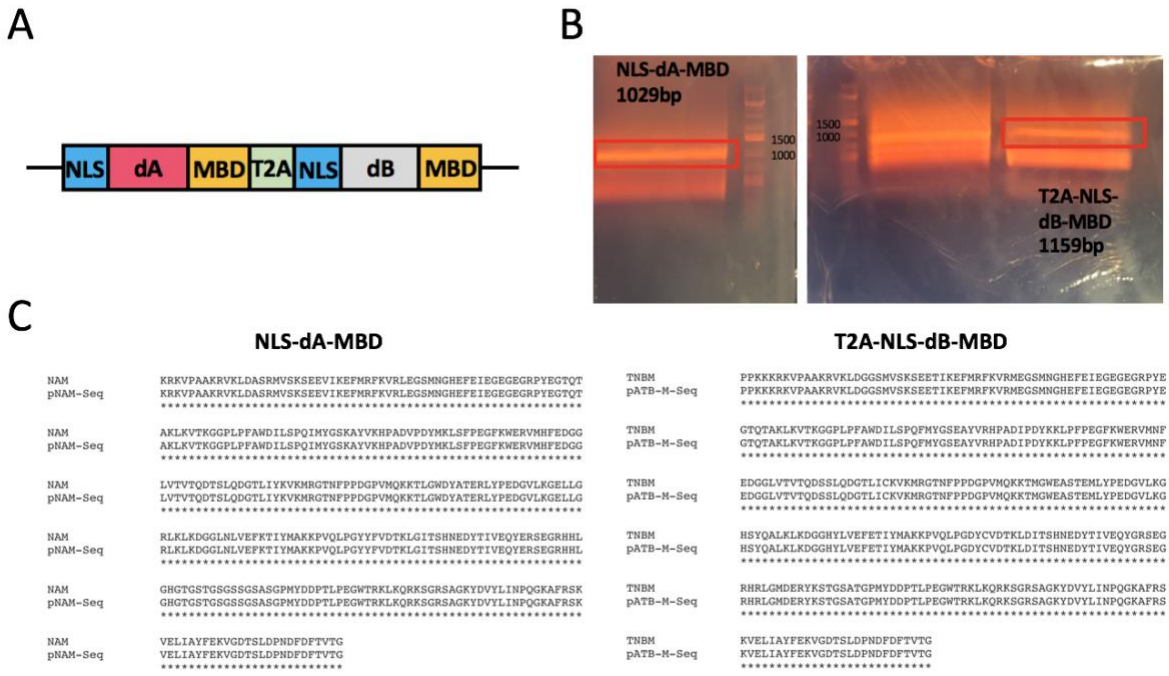


Figure 6. Construct of pATB-M. (A). Design of the pATB-M construct. (B) PCR results were developed by gel electrophoresis. (C) The pATB-M construct was aligned and validated by sequencing.

2.3.2. *Live-cell imaging of methylated DNA*

For live-cell imaging of methylated DNA, the two constructs, NLS-dA-MBD-NLS-dB-MBD (pAB-M) and NLS-dA-MBD-T2A-NLS-dB-MBD (pATB-M), including the MBD derived from the human MeCP2 gene and nuclear localization sequences, were transfected into HEK293T cells for imaging. Small fluorescent puncta that could overlap with nuclei in cells observed under bright-field microscopy were found. These results reveal that the ddRFP probes pAB-M and pATB-M clearly pattern methylated DNA in living cells (Figure 7).

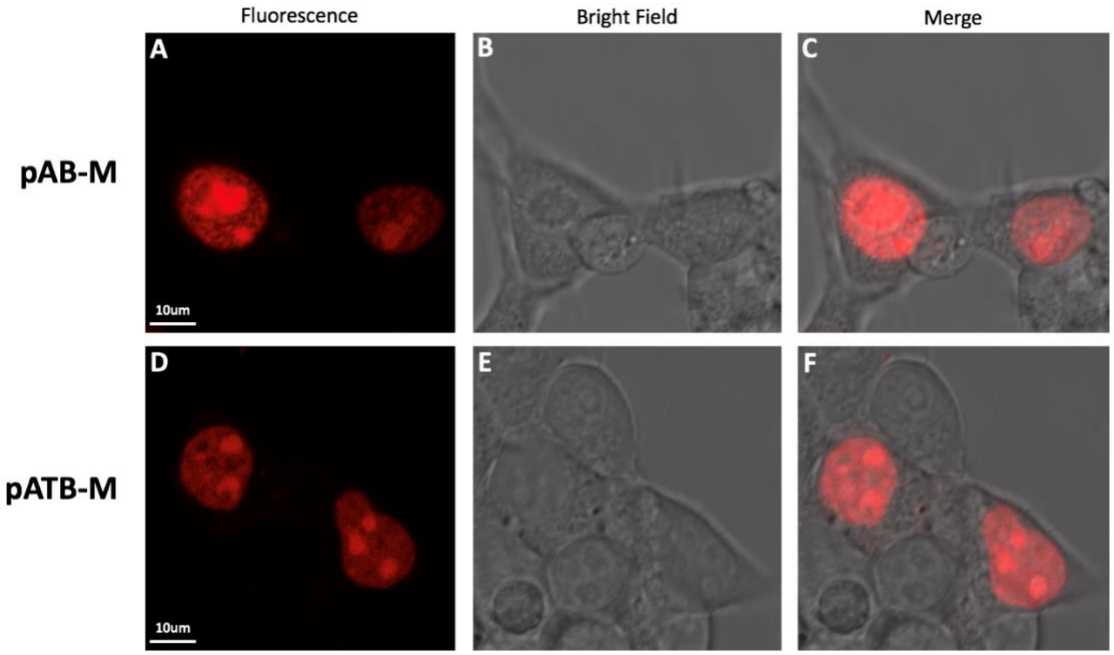


Figure 7. Live-cell imaging of methylated DNA. (A–C) pAB-M and (D–F) pATB-M were expressed in HEK293T cells.

Further analyses using a DNA methylation inhibitor, 5-azacytidine (5AZA)[83], and an alpha-ketoglutarate-dependent enzyme inhibitor, NOG[84], were used to validate whether the puncta are generated from the binding of the ddRFP probes to methylated DNA. Theoretically, cells lose fluorescence when treated with 5-AZA and display fluorescence when treated with NOG. In the first construct, pAB-M, fluorescence intensity was not expected to differ because domains A and B are strictly linked. The second construct, pATB-M, however, was expected to show a difference in fluorescence intensity owing to the separation of domains A and B by the T2A peptide (Figure 8). Quantitative analyses were performed to measure differences in fluorescence between these two constructs after treatment of cells with 5AZA or NOG. The results show a loss of fluorescence after treatment with 5AZA and no significant change in fluorescence after treatment NOG (Figure 9). This finding is reasonable because NOG only inhibits the alpha-ketoglutarate-dependent enzymes involved in DNA demethylation. However, other 5mC deoxygenase enzymes and cofactors are still available in complex with DNA demethylation. Thus, we can infer that no significant fluorescence changes should be expected because the efficiency by which the inhibition of DNA demethylation process might not apparent. More data are required to examine this phenomenon.

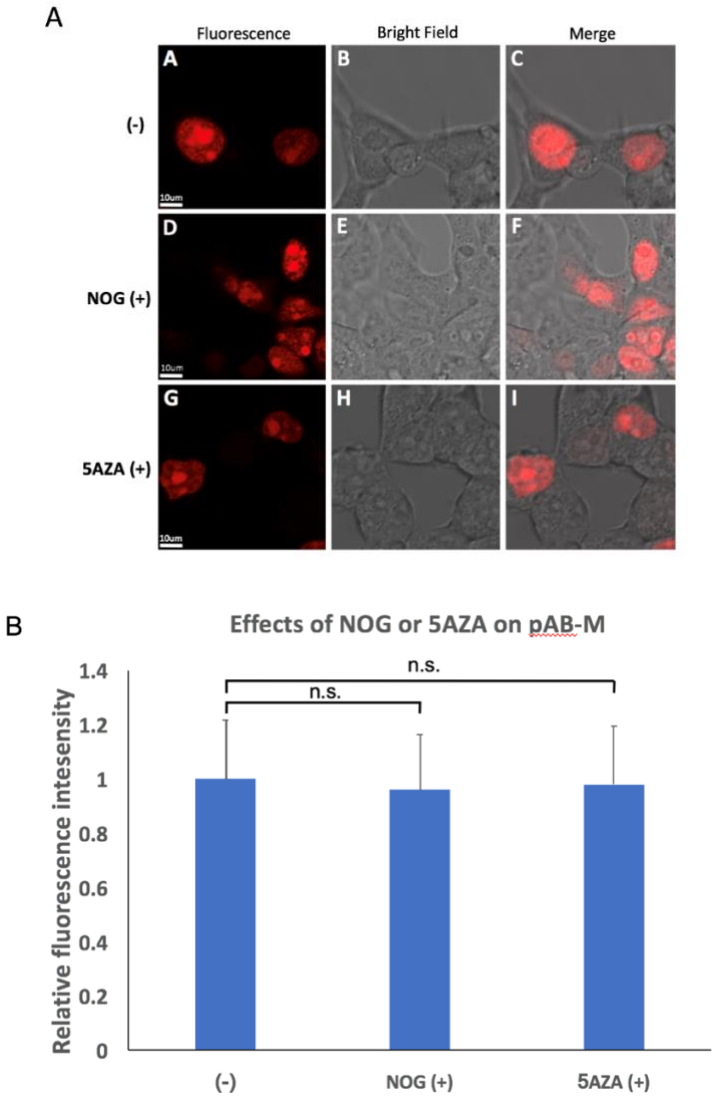


Figure 8. Live-cell imaging of methylated DNA. (A) Effect of NOG or 5AZA on fluorescence in the nuclei of cells. HEK293T cells were transfected with plasmid pAB-M in the absence or presence of NOG (1 mM) or 5AZA (2 μM) for 24 h. Scale bar, 10 μm. (B) Imaging-based quantification of methylated DNA. Fluorescence in nuclei was quantified by treatment with or without NOG (1 mM) or 5AZA (2 μM) for 24 h. Data are shown as the mean ± SD of at least six cells from different experiments. n.s.= non-significant difference.

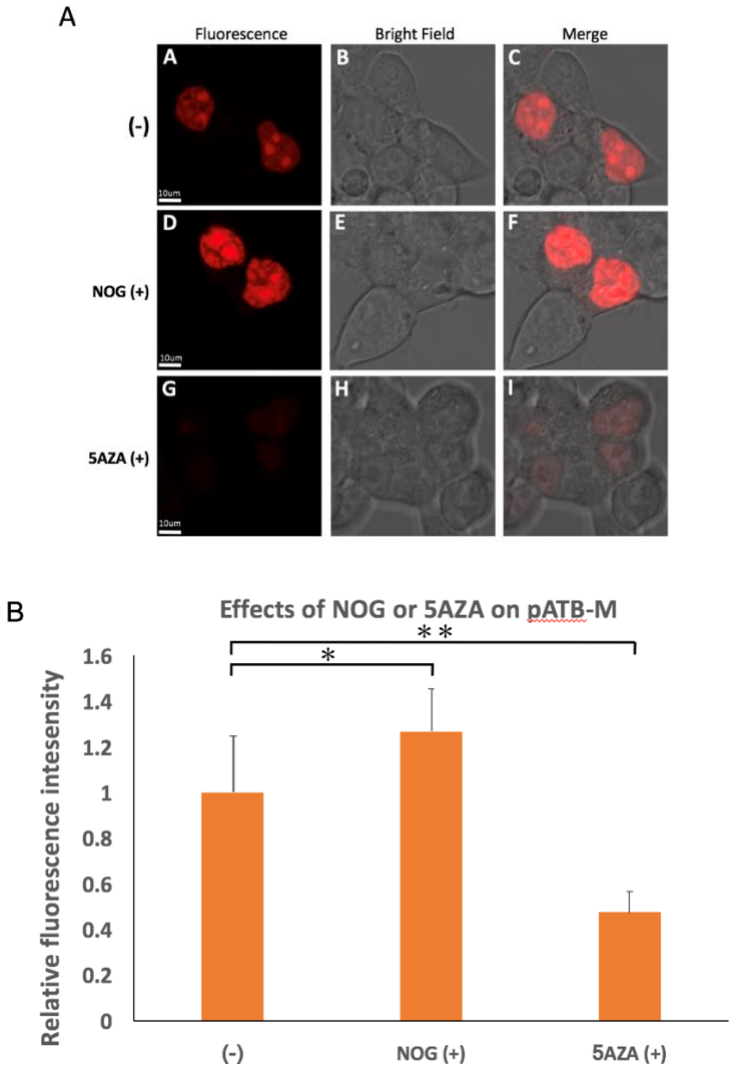


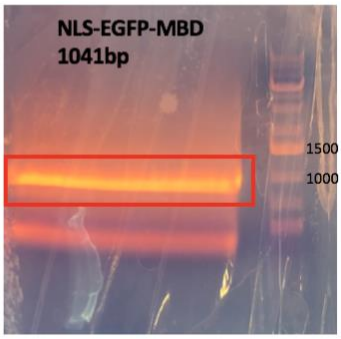
Figure 9. Live-cell imaging of methylated DNA. (A) Effect of NOG or 5AZA on fluorescence in the nuclei of cells. HEK293T cells were transfected with plasmid pATB-M in the absence or presence of NOG (1 mM) or 5AZA (2 μM) for 24 h. Scale bar, 10 μm. (B) Imaging-based quantification of methylated DNA. Fluorescence in nuclei was quantified after treatment with or without NOG (1 mM) or 5AZA (2 μM) for 24 h. Data are shown as the mean ± SD of at least six cells from different experiments. ∗p<0.05, ∗∗p<0.01.

For compound treatment experiments, we compared the performance of EGFP probes with that of our ddRFP probes. EGFP was fused with the MBD derived from the human MeCP2 gene and a nuclear localization sequence (pE-M) and assayed in HEK293T cells (Figure 10). The fluorescent puncta detected from the pE-M probe was similar to those of the pAB-M and pATB-M probes. However, pE-M clearly showed high brightness, which led to significant background noise and ambiguous localization (Figures 11 – 12). Owing to the strong autofluoresence generated by the EGFP-based probes, the fluorescence intensity remained at the same level regardless of the amount of methylated DNA available. In the present research project, we addressed this issue by using ddRFP-based probes, which feature very dim fluorescence when only domain A is present but show increased fluorescence when both domains A and B dimerize around methylated DNA regions. These results suggest that the ddRFP-based probes are able to clearly examine methylated DNA levels in cellular experiments with lower fluorescence intensity but clearer localization compared with EGFP-based probes because the former possess fluorescence intensity-switching characteristics. We also compared differences in the fluorescence intensities of domain A only and the ddRFP probes. Domain A was fused with MeCP2 MBD and a nuclear localization sequence (pA-M) and subsequently assayed in HEK293T cells. The results reveal the significant fluorescence intensity difference between pA-M and the ddRFP probes and demonstrate that the dimeric state was successfully formed and that fluorescence was enhanced when domains A and B are close to each other around methylated DNA regions (Figures 13 – 14).

A



B



C

NLS-EGFP-MBD

```
NEM             RKVPAAKRVKLDASRMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTL
pE-M-Seq        RKVPAAKRVKLDASRMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTL
                ************************************************************

NEM             KFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDG
pE-M-Seq        KFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDG
                ************************************************************

NEM             NYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVN
pE-M-Seq        NYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVN
                ************************************************************

NEM             FKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFV
pE-M-Seq        FKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFV
                ************************************************************

NEM             TAAGITLGMDELYKGSASGPMYDDPTLPEGWTRKLKQRKSGRSAGKYDVYLINPQGKAFR
pE-M-Seq        TAAGITLGMDELYKGSASGPMYDDPTLPEGWTRKLKQRKSGRSAGKYDVYLINPQGKAFR
                ************************************************************

NEM             SKVELIAYFEKVG
pE-M-Seq        SKVELIAYFEKVG
                *************
```

Figure 10. Construct of pE-M. (A) Design of the pE-M construct. (B) PCR results were developed by gel electrophoresis. (C) The pE-M construct was aligned and validated by sequencing.

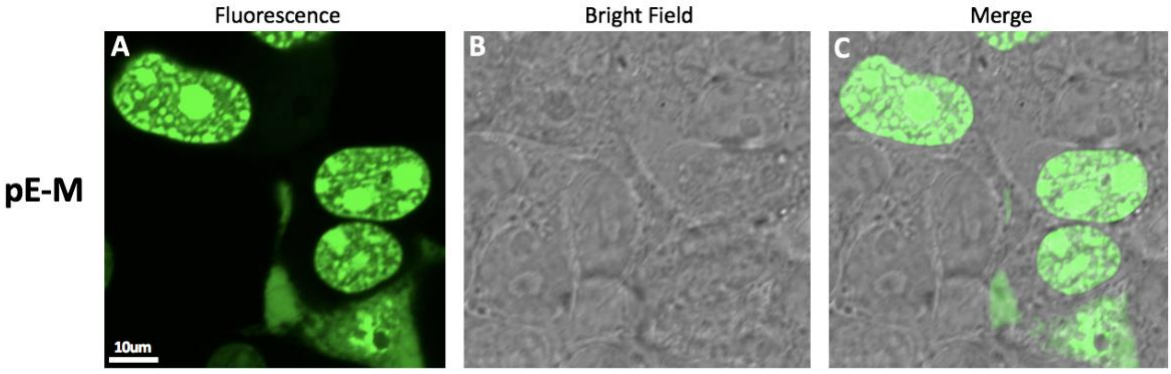


Figure 11. Live-cell imaging of methylated DNA. (A–C) pE-M was expressed in HEK293T cells.

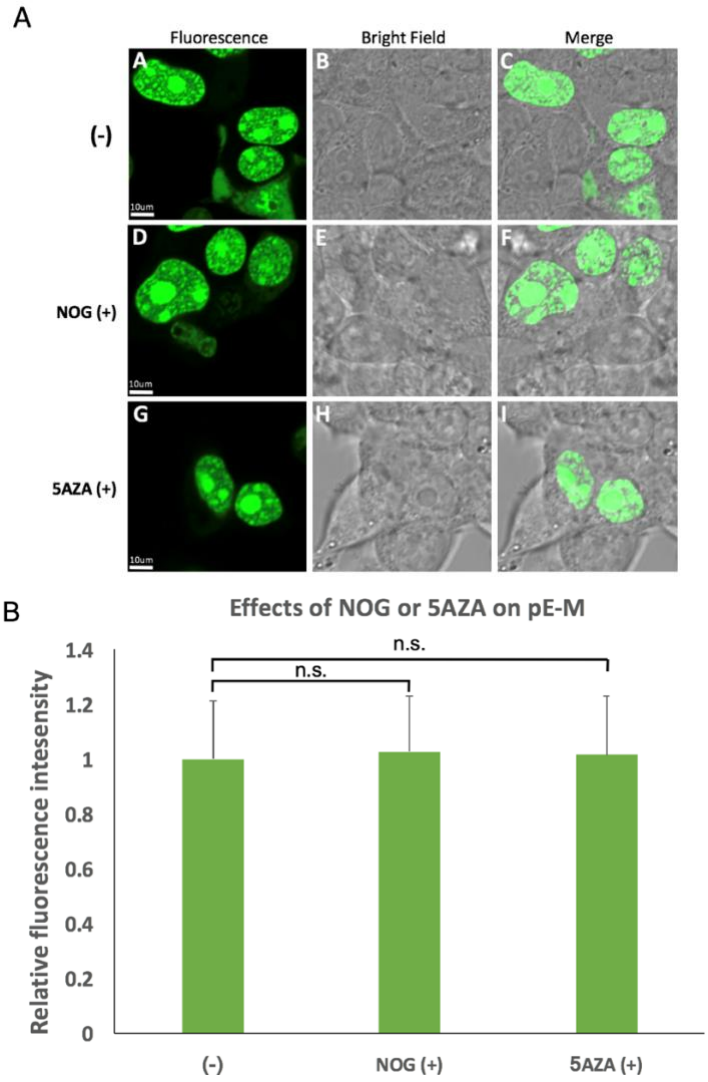


Figure 12. Live-cell imaging of methylated DNA. (A) Effect of NOG or 5AZA on fluorescence in the nuclei of cells. HEK293T cells were transfected with plasmid pAB-M in the absence or presence of NOG (1 mM) or 5AZA (2 μM) for 24 h. Scale bar, 10 μm. (B) Imaging-based quantification of methylated DNA. Fluorescence in nuclei was quantified after treatment with or without NOG (1 mM) or 5AZA (2 μM) for 24 h. Data are shown as the mean ± SD of at least six cells from different experiments. n.s.= non-significant difference.

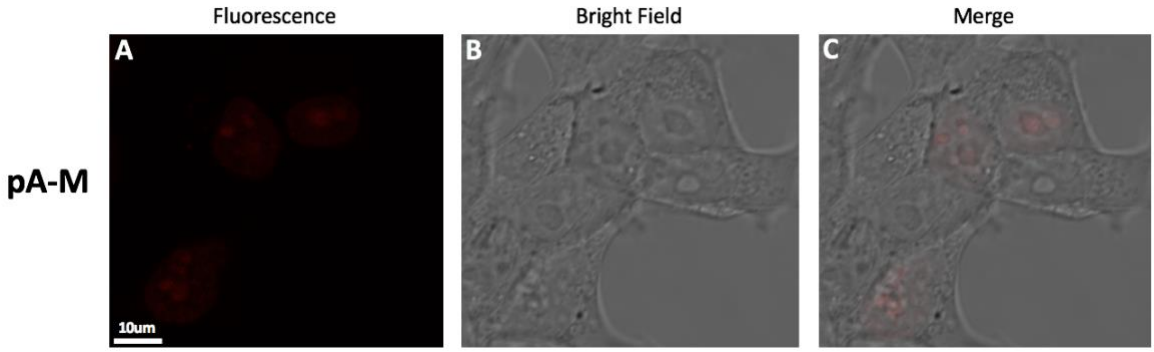


Figure 13. Live-cell imaging of methylated DNA. (A–C) pA-M was expressed in HEK293T cells.

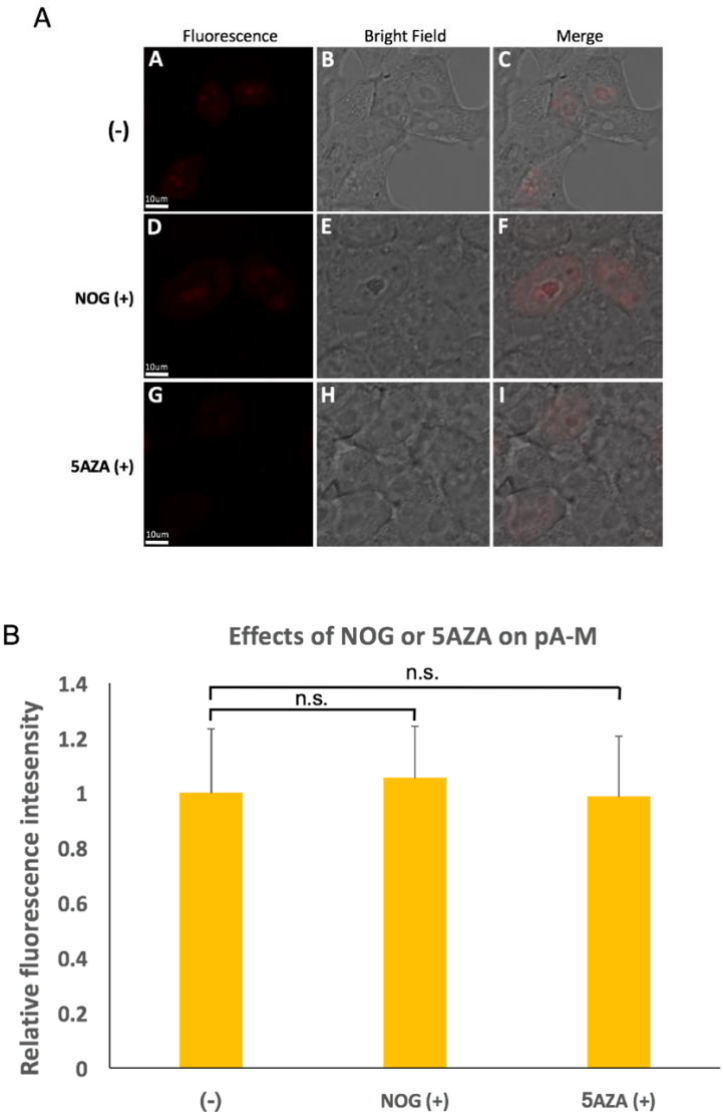


Figure 14. Live-cell imaging of methylated DNA. (A) Effect of NOG or 5AZA on fluorescence in the nuclei of cells. HEK293T cells were transfected with plasmid pAB-M in the absence or presence of NOG (1 mM) or 5AZA (2 μM) for 24 h. Scale bar, 10 μm. (B) Imaging-based quantification of methylated DNA. Fluorescence in nuclei was quantified after treatment with or without NOG (1 mM) or 5AZA (2 μM) for 24 h. Data are shown as the mean ± SD of at least six cells from different experiments. n.s.= non-significant difference.

Next, we investigated differences among localization phenotypes based on the original and mutant MBDs in ddRFP probes when detecting methylated DNA. Previous studies have shown a single-point mutation at R111G that abolishes MeCP2 MBD binding to methyl substrates.[85] Based on this study, an R111G MBD mutant was fused to the ddRFPs and a nuclear localization sequence (pAB-mM and pATB-mM) and subsequently assayed in HEK293T cells. The fluorescent puncta generated from the ddRFP-mMBD probes differed from those produced by the ddRFP-MBD probes, which indicates that the methylated DNA regions in the nucleus were targeted by the latter probes (Figures 15 – 17).

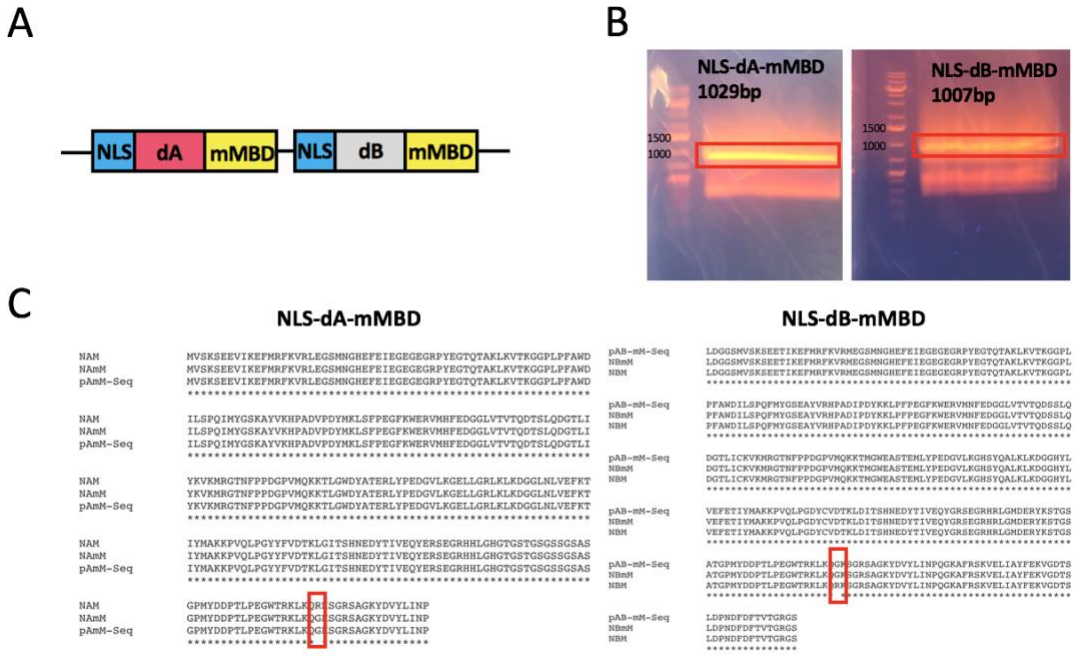


Figure 15. Construct of pAB-mM. (A) Design of the pAB-mM construct. (B) PCR results were developed by gel electrophoresis. (C) The pAB-M construct was aligned and validated by sequencing.

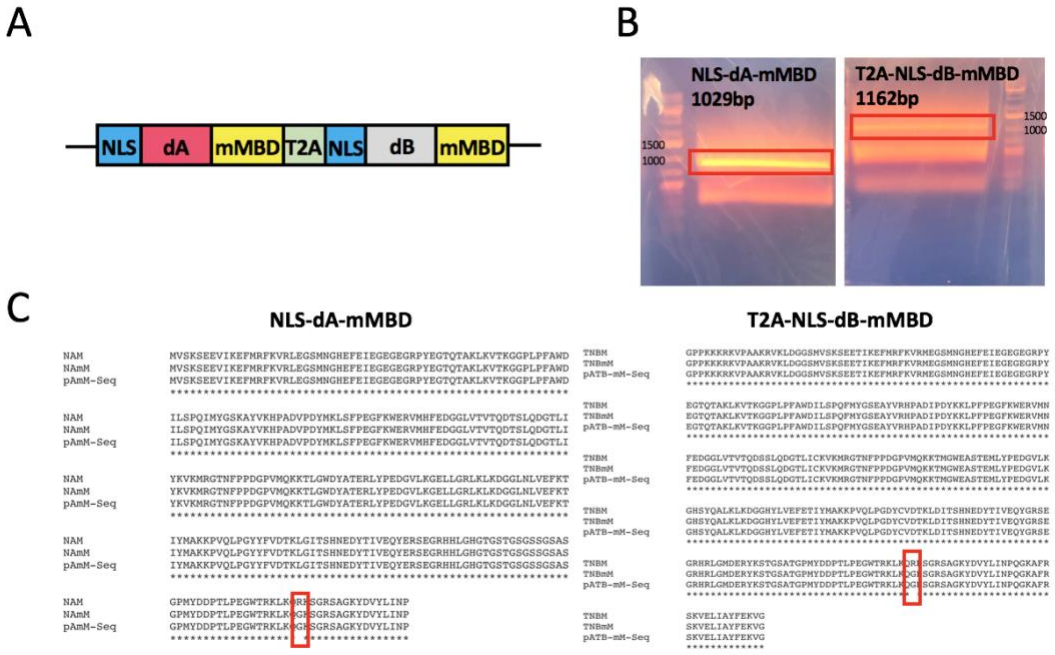


Figure 16. Construct of pATB-mM. (A) Design of the pATB-mM construct. (B) PCR results were developed by gel electrophoresis. (C) The pATB-mM construct was aligned and validated by sequencing.

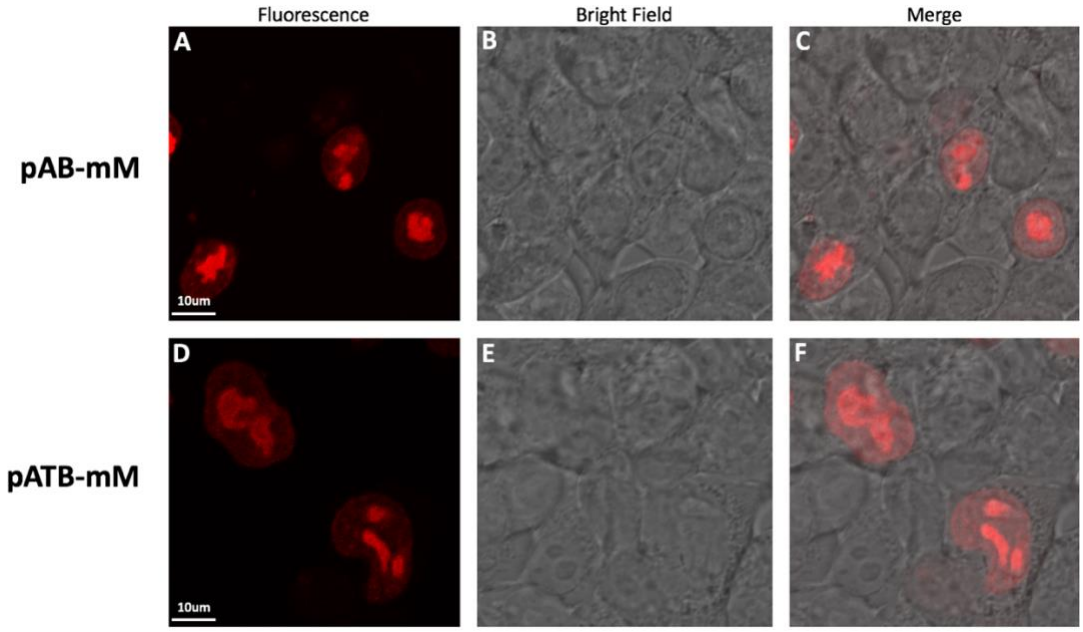


Figure 17. Live-cell imaging of methylated DNA. (A – C) pAB-mM and (D – F) pATB-mM were expressed in HEK293T cells.

Chapter 3

Conclusion and Outlook

In this research project, we developed a dimerization-dependent RFP probe to image endogenous methylated DNA. Preliminary data indicate that methylated DNA was patterned and monitored by the developed ddRFP probes in mammalian cells. The dimerization characteristic and T2A-peptide version of the ddRFP biosensor produce a fluorescence switching mechanism that can clearly differentiate associated and dissociated states, which means methylated and non-methylated DNA can be clearly examined; this development solves the issue of EGFP-based probes that may produce undesired background fluorescence. However, this project is not yet complete and much work remains to be done to determine whether ddRFPs are appropriate sensors for detecting DNA methylation. Recent studies have indicated that MBD from the MeCP2 gene not only binds to 5mC but also targets 5hmC.[21,22] Thus, a gel shift assay is required to further investigate whether MBD binds to 5mC or 5hmC. Sensors for detecting 5hmC have not yet been developed for mammalian cells. Through the design of ddRFP-based probes for methylated DNA described in this project, we may be able to monitor 5hmC levels in various diseases, particularly brain diseases, in the future. 5hmC is abundant in the nerve cells and brain tissue of mammals, which means it may play a significant role in neuronal and brain development.[21] Nevertheless, much in this area remains unknown. Given further development, we may be able to apply the developed ddRFP probe to

detect 5mC or 5hmC levels in different cancers and further dissect how DNA methylation patterns develop in various cancers. More importantly, our strategies and preliminary data from this project clearly show the potential and novelty of building up new epigenetic FP-based biosensors for the next generation.

References

1 Bird, A. DNA methylation patterns and epigenetic memory. *Genes & development* **16**, 6-21, doi:10.1101/gad.947102 (2002).

2 Weinhold, B. Epigenetics: The Science of Change. *Environmental health perspectives* **114**, A160-A167 (2006).

3 Kulis, M. & Esteller, M. DNA methylation and cancer. *Advances in genetics* **70**, 27-56, doi:10.1016/b978-0-12-380866-0.60002-2 (2010).

4 Lu, H., Liu, X., Deng, Y. & Qing, H. DNA methylation, a hand behind neurodegenerative diseases. *Frontiers in aging neuroscience* **5**, doi:10.3389/fnagi.2013.00085 (2013).

5 Wen, K. *et al.* The Role of DNA Methylation and Histone Modifications in Neurodegenerative Diseases: A Systematic Review. *Plos one* **11**, doi:10.1371/journal.pone.0167201 (2016).

6 Robertson, K. D. DNA methylation and human disease. *Nature reviews. Genetics* **6**, 597-610, doi:10.1038/nrg1655 (2005).

7 Bestor, T. H. The DNA methyltransferases of mammals. *Human molecular genetics* **9**, 2395-2402 (2000).

8 Probst, A. V., Dunleavy, E. & Almouzni, G. Epigenetic inheritance during the cell cycle. *Nature reviews. Molecular cell biology* **10**, 192-206, doi:10.1038/nrm2640 (2009).

9 Goll, M. G. *et al.* Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science (New York, N.Y.)* **311**, 395-398, doi:10.1126/science.1120976 (2006).

10 Kareta, M. S., Botello, Z. M., Ennis, J. J., Chou, C. & Chedin, F. Reconstitution and mechanism of the stimulation of de novo methylation by human DNMT3L. *The journal of biological chemistry* **281**, 25893-25902, doi:10.1074/jbc.M603140200 (2006).

11 Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science (New York, N.Y.)* **324**, 929-930, doi:10.1126/science.1169786 (2009).

12 Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science (New York, N.Y.)* **324**, 930-935, doi:10.1126/science.1170116 (2009).

13 Popp, C. *et al.* Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463**, 1101-1105, doi:10.1038/nature08829 (2010).

14 Raiber, E. A. *et al.* Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase. *Genome biology* **13**, R69, doi:10.1186/gb-2012-13-8-r69 (2012).

15 Shen, L. *et al.* Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell* **153**, 692-706, doi:10.1016/j.cell.2013.04.002 (2013).

16 Song, C. X. *et al.* Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell* **153**, 678-691, doi:10.1016/j.cell.2013.04.001 (2013).

17 Cortazar, D. *et al.* Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature* **470**, 419-423, doi:10.1038/nature09672 (2011).

18 Jiricny, J. & Menigatti, M. DNA Cytosine demethylation: are we getting close? *Cell* **135**, 1167-1169, doi:10.1016/j.cell.2008.12.008 (2008).

19 Kim, M. S. *et al.* DNA demethylation in hormone-induced transcriptional derepression. *Nature* **461**, 1007-1012, doi:10.1038/nature08456 (2009).

20 Valinluck, V. & Sowers, L. C. Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer research* **67**, 946-950, doi:10.1158/0008-5472.Can-06-3123 (2007).

21 Song, J. & Pfeifer, G. P. Are there specific readers of oxidized 5-methylcytosine bases? *BioEssays: news and reviews in molecular, cellular and developmental biology* **38**, 1038-1047, doi:10.1002/bies.201600126 (2016).

22 Mellen, M., Ayata, P., Dewell, S., Kriaucionis, S. & Heintz, N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* **151**, 1417-1430, doi:10.1016/j.cell.2012.11.022 (2012).

23 Lehnertz, B. *et al.* Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Current biology: CB* **13**, 1192-1200 (2003).

24 Vire, E. *et al.* The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**, 871-874, doi:10.1038/nature04431 (2006).

25 Fuks, F., Hurd, P. J., Deplus, R. & Kouzarides, T. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic acids research* **31**, 2305-2312 (2003).

26 Tachibana, M. *et al.* G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes & development* **16**, 1779-1791, doi:10.1101/gad.989402 (2002).

27 Cheutin, T. *et al.* Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science (New York, N.Y.)* **299**, 721-725, doi:10.1126/science.1078572 (2003).

28 Tachibana, M., Matsumura, Y., Fukuda, M., Kimura, H. & Shinkai, Y. G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *The EMBO journal* **27**, 2681-2690, doi:10.1038/emboj.2008.192 (2008).

29 Tachibana, M., Sugimoto, K., Fukushima, T. & Shinkai, Y. Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. *The journal of biological chemistry* **276**, 25309-25317, doi:10.1074/jbc.M101914200 (2001).

30 Egger, G. *et al.* Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. *Proceedings of the national academy of sciences of the United States of America* **103**, 14080-14085, doi:10.1073/pnas.0604602103 (2006).

31 Esteve, P. O. *et al.* Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes & development* **20**, 3089-3103, doi:10.1101/gad.1463706 (2006).

32 Wozniak, R. J., Klimecki, W. T., Lau, S. S., Feinstein, Y. & Futscher, B. W. 5-Aza-2'-deoxycytidine-mediated reductions in G9A histone methyltransferase and histone H3 K9 di-methylation levels are linked to tumor suppressor gene reactivation. *Oncogene* **26**, 77-90, doi:10.1038/sj.onc.1209763 (2007).

33 Du, Q., Luu, P. L., Stirzaker, C. & Clark, S. J. Methyl-CpG-binding domain proteins: readers of the epigenome. *Epigenomics* **7**, 1051-1073, doi:10.2217/epi.15.39 (2015).

34 Sarraf, S. A. & Stancheva, I. Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. *Molecular cell* **15**, 595-605, doi:10.1016/j.molcel.2004.06.043 (2004).

35 Fujita, N. *et al.* Mechanism of transcriptional regulation by methyl-CpG binding protein MBD1. *Molecular and cellular biology* **20**, 5107-5118 (2000).

36 Frauer, C. *et al.* Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. *Plos one* **6**, e21306, doi:10.1371/journal.pone.0021306 (2011).

37 Neul, J. L. *et al.* Rett Syndrome: Revised Diagnostic Criteria and Nomenclature. *Annals of neurology* **68**, 944-950, doi:10.1002/ana.22124 (2010).

38 Nan, X. *et al.* Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**, 386-389, doi:10.1038/30764 (1998).

39 Hagberg, B., Aicardi, J., Dias, K. & Ramos, O. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Annals of neurology* **14**, 471-479, doi:10.1002/ana.410140412 (1983).

40 Guidera, K. J., Borrelli, J., Jr., Raney, E., Thompson-Rangel, T. & Ogden, J. A. Orthopaedic manifestations of Rett syndrome. *Journal of pediatric orthopedics* **11**, 204-208 (1991).

41 Percy, A. K. *et al.* Profiling Scoliosis in Rett Syndrome. *Pediatric research* **67**, 435-439, doi:10.1203/PDR.0b013e3181d0187f (2010).

42 Zysman, L., Lotan, M. & Ben-Zeev, B. Osteoporosis in Rett syndrome: A study on normal values. *The scientific world journal* **6**, 1619-1630, doi:10.1100/tsw.2006.266 (2006).

43 Gabel, H. W. *et al.* Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. *Nature* **522**, 89-93, doi:10.1038/nature14319 (2015).

44 Guo, J. U. *et al.* Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nature neuroscience* **17**, 215-222, doi:10.1038/nn.3607 (2014).

45 Lewis, J. D. *et al.* Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* **69**, 905-914 (1992).

46 Li, Y. *et al.* Global transcriptional and translational repression in human-embryonic-stem-cell-derived Rett syndrome neurons. *Cell stem cell* **13**, 446-458, doi:10.1016/j.stem.2013.09.001 (2013).

47 Feinberg, A. P., Ohlsson, R. & Henikoff, S. The epigenetic progenitor origin of human cancer. *Nature reviews. Genetics* **7**, 21-33, doi:10.1038/nrg1748 (2006).

48 Ohm, J. E. & Baylin, S. B. Stem cell chromatin patterns: an instructive mechanism for DNA hypermethylation? *Cell cycle (Georgetown, Tex.)* **6**, 1040-1043, doi:10.4161/cc.6.9.4210 (2007).

49 Kisseljova, N. P. & Kisseljov, F. L. DNA demethylation and carcinogenesis. *Biochemistry. Biokhimiia* **70**, 743-752 (2005).

50 Shimomura, O., Johnson, F. H. & Saiga, Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *Journal of cellular and comparative physiology* **59**, 223-239 (1962).

51 Morise, H., Shimomura, O., Johnson, F. H. & Winant, J. Intermolecular energy transfer in the bioluminescent system of Aequorea. *Biochemistry* **13**, 2656-2662 (1974).

52 Heim, R. & Tsien, R. Y. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current biology: CB* **6**, 178-182 (1996).

53 Miyawaki, A., Griesbeck, O., Heim, R. & Tsien, R. Y. Dynamic and quantitative Ca2+ measurements using improved cameleons. *Proceedings of the national academy of sciences of the United States of America* **96**, 2135-2140 (1999).

54 Miyawaki, A., Nagai, T. & Mizuno, H. Engineering fluorescent proteins. *Advances in biochemical engineering/biotechnology* **95**, 1-15 (2005).

55 Wachter, R. M., Elsliger, M. A., Kallio, K., Hanson, G. T. & Remington, S. J. Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein. *Structure (London, England : 1993)* **6**, 1267-1277 (1998).

56 Cubitt, A. B. *et al.* Understanding, improving and using green fluorescent proteins. *Trends in biochemical sciences* **20**, 448-455 (1995).

57 Cubitt, A. B., Woollenweber, L. A. & Heim, R. Understanding structure-function relationships in the Aequorea victoria green fluorescent protein. *Methods in cell biology* **58**, 19-30 (1999).

58 Heim, R., Prasher, D. C. & Tsien, R. Y. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the national academy of sciences of the United States of America* **91**, 12501-12504 (1994).

59 Patterson, G., Day, R. N. & Piston, D. Fluorescent protein spectra. *Journal of cell science* **114**, 837-838 (2001).

60 Patterson, G. H., Knobel, S. M., Sharif, W. D., Kain, S. R. & Piston, D. W. Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophysical journal* **73**, 2782-2790, doi:10.1016/s0006-3495(97)78307-3 (1997).

61 Yang, T. T. *et al.* Improved fluorescence and dual color detection with enhanced blue and green variants of the green fluorescent protein. *The Journal of biological chemistry* **273**, 8212-8216 (1998).

62 Patterson, G. H., Piston, D. W. & Barisas, B. G. Forster distances between green fluorescent protein pairs. *Analytical biochemistry* **284**, 438-440, doi:10.1006/abio.2000.4708 (2000).

63 Shaner, N. C., Patterson, G. H. & Davidson, M. W. Advances in fluorescent protein technology. *Journal of cell science* **120**, 4247-4260, doi:10.1242/jcs.005801 (2007).

64 Gross, L. A., Baird, G. S., Hoffman, R. C., Baldridge, K. K. & Tsien, R. Y. The structure of the chromophore within DsRed, a red fluorescent protein from coral. *Proceedings of the national academy of sciences of the United States of America* **97**, 11990-11995, doi:10.1073/pnas.97.22.11990 (2000).

65 Baird, G. S., Zacharias, D. A. & Tsien, R. Y. Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proceedings of the national academy of sciences of the United States of America* **97**, 11984-11989, doi:10.1073/pnas.97.22.11984 (2000).

66 Wiehler, J., von Hummel, J. & Steipe, B. Mutants of Discosoma red fluorescent protein with a GFP-like chromophore. *FEBS letters* **487**, 384-389 (2001).

67 Yarbrough, D., Wachter, R. M., Kallio, K., Matz, M. V. & Remington, S. J. Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-A

resolution. *Proceedings of the national academy of sciences of the United States of America* **98**, 462-467, doi:10.1073/pnas.98.2.462 (2001).

68 Bevis, B. J. & Glick, B. S. Rapidly maturing variants of the Discosoma red fluorescent protein (DsRed). *Nature biotechnology* **20**, 83-87, doi:10.1038/nbt0102-83 (2002).

69 Yanushevich, Y. G. *et al.* A strategy for the generation of non-aggregating mutants of Anthozoa fluorescent proteins. *FEBS letters* **511**, 11-14 (2002).

70 Gurskaya, N. G. *et al.* GFP-like chromoproteins as a source of far-red fluorescent proteins. *FEBS letters* **507**, 16-20 (2001).

71 Campbell, R. E. *et al.* A monomeric red fluorescent protein. *Proceedings of the national academy of sciences of the United States of America* **99**, 7877-7882, doi:10.1073/pnas.082243699 (2002).

72 Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nature biotechnology* **22**, 1567-1572, doi:10.1038/nbt1037 (2004).

73 Shaner, N. C. *et al.* Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nature methods* **5**, 545-551, doi:10.1038/nmeth.1209 (2008).

74 Lin, J. Y., Lin, M. Z., Steinbach, P. & Tsien, R. Y. Characterization of engineered channelrhodopsin variants with improved properties and kinetics. *Biophysical journal* **96**, 1803-1814, doi:10.1016/j.bpj.2008.11.034 (2009).

75 Shcherbo, D. *et al.* Bright far-red fluorescent protein for whole-body imaging. *Nature methods* **4**, 741-746, doi:10.1038/nmeth1083 (2007).

76 Miyawaki, A. *et al.* Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin. *Nature* **388**, 882-887, doi:10.1038/42264 (1997).

77 Oldach, L. & Zhang, J. Genetically encoded fluorescent biosensors for live-cell visualization of protein phosphorylation. *Chemistry & biology* **21**, 186-197, doi:10.1016/j.chembiol.2013.12.012 (2014).

78 Alford, S. C. A Fluorogenic Red Fluorescent Protein Heterodimer. *Chemistry & biology***19**, 353-360, doi:10.1016/j.chembiol.2012.01.006 (2012).

79 Alford, S. C. Dimerization-Dependent Green and Yellow Fluorescent Proteins. *ACS synthetic biology* **1**, 569-575, doi:10.1021/sb300050j (2012).

80 Ding, Y. *et al.* Ratiometric biosensors based on dimerization-dependent fluorescent protein exchange. *Nature methods* **12**, 195-198, doi:10.1038/nmeth.3261 (2015).

81 Yamagata, K. DNA methylation profiling using live-cell imaging. *Methods (San Diego, Calif.)* **52**, 259-266, doi:10.1016/j.ymeth.2010.04.008 (2010).

82 Hori, Y. *et al.* Synthetic-Molecule/Protein Hybrid Probe with Fluorogenic Switch for Live-Cell Imaging of DNA Methylation. *Journal of the American chemical society* **140**, 1686-1690, doi:10.1021/jacs.7b09713 (2018).

83 Christman, J. K. 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene* **21**, 5483-5495, doi:10.1038/sj.onc.1205699 (2002).

84 Hausinger, R. P. FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes. *Critical reviews in biochemistry and molecular biology* **39**, 21-68, doi:10.1080/10409230490440541 (2004).

85 Galvão, T. C. & Thomas, J. O. Structure-specific binding of MeCP2 to four-way junction DNA through its methyl CpG-binding domain. *Nucleic acids research* **33**, 6603-6609, doi:10.1093/nar/gki971 (2005).