Structural Insights into Methylated DNA Recognition by the Methyl-CpG Binding Domain of MBD6 from *Arabidopsis thaliana*

Yutaka Mahana, Izuru Ohki, Erik Walinda, Daichi Morimoto, Kenji Sugase, and Masahiro Shirakawa*

**ABSTRACT:** Cytosine methylation is an epigenetic modification essential for formation of mature heterochromatin, gene silencing, and genomic stability. In plants, methylation occurs not only at cytosine bases in CpG but also in CpHpG and CpHpH contexts, where H denotes A, T, or C. Methyl-CpG binding domain (MBD) proteins, which recognize symmetrical methyl-CpG dinucleotides and act as gene repressors in mammalian cells, are also present in plant cells, although their structural and functional properties still remain poorly understood. To fill this gap, in this study, we determined the solution structure of the MBD domain of the MBD6 protein from *Arabidopsis thaliana* and investigated its binding properties to methylated DNA by binding assays and an in-depth NMR spectroscopic analysis. The AtMBD6 MBD domain folds into a canonical MBD structure in line with its binding specificity toward methyl-CpG and possesses a DNA binding interface similar to mammalian MBD domains. Intriguingly, however, the binding affinity of the AtMBD6 MBD domain toward methyl-CpG-containing DNA was found to be much lower than that of known mammalian MBD domains. The main difference arises from the absence of positively charged residues in AtMBD6 that supposedly interact with the DNA backbone as seen in mammalian MBD/methyl-CpG-containing DNA complexes. Taken together, we have established a structural basis for methyl-CpG recognition by AtMBD6 to develop a deeper understanding how MBD proteins work as mediators of epigenetic signals in plant cells.

**INTRODUCTION**

A certain percentage of cytosine bases in eukaryotic DNA exist in an epigenetically modified form, as 5-methylcytosine. Cytosine methylation plays essential roles in numerous vital functions such as repression of gene expression, organization of the chromatin structure, and inactivation of transposons. In animal cells, cytosine methylation occurs mostly at symmetrical CpG sequences and is achieved by the action of DNA methyltransferases. 5-Methylcytosine may be further oxidized to 5-hydroxymethylcytosine. These two modifications are examples of distinct epigenetic marks, which can then be specifically recognized (“read-out”) with specific reader domains in many proteins.

Methyl-CpG binding domain (MBD) proteins were first identified in mammals that have very high CpG methylation levels as high as 70–80% as chromatin regulators that recognize the cytosine methylation at CpG dinucleotides (hereafter: methyl-CpG sites). Within the MBD family, several “canonical” MBD proteins from humans (*Homo sapiens*, Hs) have been intensively studied. For example, HsMeCP2 and HsMBD1 act as transcriptional repressors by binding to the methyl-CpG site and then recruiting histone methyltransferases and histone deacetylases, which promote chromatin condensation (i.e., formation of mature heterochromatin). As another example of a canonical MBD protein, the HsMBD4 MBD domain preferably recognizes (mismatch) TpG/methyl-CpG and hydroxymethyl-CpG/methyl-CpG sites, which are intermediate products in the course of active DNA demethylation. In addition to the MBD domain, these MBD proteins contain one or more functional domains that define the specific function of these proteins such as the transcription repression domain in HsMBD1 and the glycosylase domain in HsMBD4.

In plants, cytosine bases in not only CpG but also symmetrical CpHpG and asymmetrical CpHpH contexts (H represents A, T, or C) can be methylated. All these methylation patterns are thought to be established by the RNA-directed DNA methylation mechanism and maintained through different regulatory pathways. Previous studies demonstrated that genome-wide levels of cytosine methylation in plants largely deviate among species; 24% CpG, 6.7% CpHpG, and 1.7% CpHpH are methylated in thale cress (*Arabidopsis thaliana*, At), whereas high methylation rates of 87–88% CpG, 67–68% CpHpG, and 41–43% CpHpH are...
observed in rice cultivars. Epigenetic mark reader domains, such as MBD proteins, are also found in plants. Interestingly, however, domain organization of plant MBD proteins is different from that of mammalian MBD proteins. The differences in the sequence contexts and the recognition machineries of cytosine methylation between mammals and plants indicate that epigenetic regulation via DNA methylation in plant cells contains multiple pathways that do not exist in mammalian cells. However, the molecular mechanism by which each methylation pattern is read out, interpreted, and translated into downstream signals is still largely unclear.

In A. thaliana, 13 proteins, AtMBD1—AtMBD12 and AtIDM1, were identified as putative MBD proteins on the basis of their sequence homology with mammalian MBD proteins. Previous studies have shown by fluorescence microscopy that at least three of them, AtMBD5, AtMBD6, and AtMBD7, are localized to chromocenters abundant in cytosine methylation. Although the subnuclear distribution indicated that these MBD proteins recognize methylated DNA, both their binding specificity and affinity toward methylated DNA in various sequence contexts remain elusive. Furthermore, no structural information on the MBD domains from plants has been published.

Here, we focused on one particular MBD protein from A. thaliana, AtMBD6. AtMBD6 does not contain any known functional domains except the MBD domain. On the basis of the subnuclear colocalization observed by fluorescence microscopy and in vitro binding examined by pull-down assays, it has been speculated that AtMBD6 is involved in maintenance of DNA methylation mediated by an ATP-dependent DNA helicase, AtDDM1. Another previous study suggested by yeast two-hybrid assays and Förster resonance energy transfer experiments that AtMBD6 interacts with proteins involved in the RNA-directed DNA methylation pathway, including an RNA binding protein, AtAGO4, and a histone deacetylase AtHDA6. However, many fundamental properties of AtMBD6 such as the structure and the binding specificity toward methylated DNA remain to be elucidated. In the present study, we characterized the AtMBD6 MBD domain using NMR spectroscopy to establish a structural basis toward understanding the roles of MBD domains in the complex of epigenomic regulation in plant cells.

**Binding Specificity of the AtMBD6 MBD Domain toward Methylated DNA.** First, we examined the binding preference of the MBD domain of AtMBD6 (residues 78—140; MBD<sub>AtMBD6</sub>) toward DNA harboring single CpG, CpApG, or CpApA sites in methylated or nonmethylated states (Table 1) using a gel shift assay. MBD<sub>AtMBD6</sub> showed specific binding to the DNA harboring the methyl-CpG site and roughly the same degree of nonspecific binding to all other DNA samples examined (Figures 1 and S1). Since a previous study had indicated that AtMBD6 could form both a homodimer and a heterodimer with AtMBD5 in vitro, we considered that formation of a homodimer might alter the observed binding specificity of MBD<sub>AtMBD6</sub> toward methylated DNA. However, the two-fold serial dilutions of a highly concentrated solution of MBD<sub>AtMBD6</sub> showed no detectable changes in the backbone amide chemical shifts, indicating that MBD<sub>AtMBD6</sub> exists as a monomer in solution even at concentrations as high as 200 μM (Figure S2). Taken together, these results indicated that MBD<sub>AtMBD6</sub> is a canonical monomeric MBD domain that specifically recognizes the methyl-CpG sites in DNA.

### Table 1. Sequences of the DNA Constructs Used in This Study

| name | sequence<sup>a</sup> |
|------|----------------------|
| CG   | CCTACTCCGTCCTACG 3′ |
| MG   | CAGTACCGCTCAGACC 3′ |
| CAG  | GGTGAGCAAGAGGC 3′ |
| MAG  | GGTGAGMAAGAGGCC 3′ |
| CAA  | CCTCTCGTCCTCCG 3′ |
| MAA  | CCACTCGGTCTCCG 3′ |

<sup>a</sup>M represents 5-methylcytosine.

**Figure 1.** Binding specificity of the AtMBD6 MBD domain toward methylated DNA. Gel shift assay of DNA binding by MBD<sub>AtMBD6</sub>. The names of the DNA samples, indicated at the top of the panel, correspond to the sequences shown in Table 1.

**Solution Structure of the AtMBD6 MBD Domain.** To gain structural insights into the DNA binding of MBD<sub>AtMBD6</sub>, we determined the solution structure of MBD<sub>AtMBD6</sub> in the free form by NMR spectroscopy. In the initial structure calculation, interproton distance restraints from three-dimensional NOESY spectra and dihedral angle restraints derived from backbone and β-carbon chemical shifts were used. The lowest-energy structure of the NOE-based structure calculation was then used as the initial structure for further structural refinement applying residual dipolar coupling (RDC) restraints in addition. The obtained 20 minimum energy structures were well defined, as indicated by the root-mean-square deviation (RMSD) for the backbone of 0.6 ± 0.2 Å except for a loop region (residues 93—98) and a disordered C-terminal region (residues 128—140) (Figure 2A and Table 2).

The structure is composed of three β-strands, an α-helix, and a flexible loop L1 that connects the β-strands β1 and β2 (Figure 2B). This fold is indeed the canonical fold of mammalian MBD domains. Accordingly, the overall structure is quite similar to the solution and crystal structures of mammalian MBD domains in complex with methyl-CpG-containing DNA reported previously (Figure 2C). The structural flexibility of the L1 loop is consistent with relatively small backbone heteronuclear NOE values previously observed for residues in this loop and very similar observations have also been reported for the free forms of the MBD domains of HsMBD1 and HsMeCP2. In addition, four important
residues including two arginine residues termed as arginine fingers (RFs), which are conserved among mammalian MBD domains, are also present at the corresponding position in the structure of MBDAtMBD6 (Figure 2C, shown as sticks). In the reported complexes of human MBD domains and DNA, these residues were shown to form unique structural motifs to recognize the methyl-CpG sites via hydrophobic interactions with the methyl groups of 5-methylcytosine bases and direct or water-mediated hydrogen bonds with 5-methylcytosine and guanine bases.18,19 Taken together, our data showed that MBDAtMBD6 adopts the canonical MBD fold in solution, which is in fine agreement with the result of the gel shift assay showing the binding specificity of MBDAtMBD6 toward methyl-CpG-containing DNA.

**Table 2. Statistics of the AtMBD6 MBD Domain NMR Structure Determination**

|                | distance restraints | short range (|i − j| ≤ 1) | 445 |
|----------------|---------------------|-----------------|-----|
|                | medium range (1 < |i − j| < 5)      | 125 |
|                | long range (|i − j| ≥ 5)    | 164 |
|                | total               | 734 |
|                | 1H-15N RDC restraints | 37 |
| RMSD statistics (residues 78–92, 99–127) | backbone | 0.6 ± 0.2 Å |
|                | heavy atoms         | 1.1 ± 0.1 Å |
| Ramachandran plot statistics (residues 78–127) | residues in the most favored regions | 89.2% |
|                | residues in additionally allowed regions | 9.0% |
|                | residues in generously allowed regions | 1.8% |
|                | residues in disallowed regions | 0.0% |

Table 2. Statistics of the AtMBD6 MBD Domain NMR Structure Determination

Interestingly, several peaks exhibited significant line broadening when the molar ratio of DNA to MBD was in the range of 0.1–0.5 (Figure 3A, enlarged view). In addition, the magnitude of line broadening appeared to be roughly proportional to the degree of change in chemical shift upon binding. This observation indicated that MBDAtMBD6 binds to methyl-CpG-containing DNA in the fast to intermediate exchange regime on the NMR timescale; in general, this regime is associated with dissociation constants $K_d$ ranging from several micromolar to several tens of micromolar.25–27 In order to determine the $K_d$ value for this interaction, we fitted the chemical shift changes during the course of the titration to a two-state binding model (Figure 3D). The $K_d$ value was obtained as 40.2 ± 0.5 μM, which is reasonable for the observed fast to intermediate exchange regime. Previous studies have revealed that several mammalian MBD domains show high affinity toward methyl-CpG-containing DNA with $K_d$ values from several tens of nanomolar to a few micromolar.24,26 Therefore, our results clearly show that the binding affinity of MBDAtMBD6 toward methyl-CpG-containing DNA.
DNA is much lower in comparison to that of the canonical MBD domains, even though MBD AtMBD6 shares the DNA binding interface with its mammalian counterparts.

Consequence of Variation of Important Amino Acids in the AtMBD6 MBD Domain.

To understand the comparably weak binding of MBD AtMBD6 to methyl-CpG-containing DNA, we performed sequence and structural alignment of MBDAtMBD6 with human and other A. thaliana MBD domains (Figures 4A, S4, and S5). MBDAtMBD6 possesses all the key residues that have been shown to be responsible for specific recognition of the methyl-CpG sites in canonical MBD domains,8,19 in line with the aforementioned results. However, two critical differences were found in the amino acid sequences. First, the C-terminal region of MBD AtMBD6 (residues 127−149) shows poor homology with its human counterparts. Previous studies on human MBD domains demonstrated that this region forms a “hairpin loop” structure that is also associated with DNA binding.23 Therefore, we hypothesize that the affinity of MBD AtMBD6 toward DNA is reduced as a consequence of the absence of the hairpin loop (Figure 4B).19 The absence of this positively charged residue therefore supports our idea that the hairpin loop is missing in MBDAtMBD6 that is, not only in the free but also even in the DNA-bound state. The remaining two residues are located near the two ends of the L1 loop and contact the phosphate backbone of DNA in the reported structures of canonical MBD-DNA complexes (Figure 4C).18,19 Previous studies demonstrated that the positive charge of these residues plays a crucial role in the DNA binding by human and chicken MBD domains.18,23,31 Therefore, we hypothesized that the absence of these important electrostatic contributions to the binding energy significantly reduces the binding affinity of MBDAtMBD6 toward methyl-CpG-containing DNA.

To explore this hypothesis, we constructed a single-point mutant S100R MBD AtMBD6 and conducted NMR titration experiments under the same conditions as performed with wildtype (WT) MBDAtMBD6. As the molar ratio of DNA to MBD was raised, cross-peaks of S100R MBD AtMBD6 corresponding to the free state gradually disappeared with very slight changes in the chemical shift, while the peak intensity of the bound state increased (Figures 4D and S6A). Several cross-peaks of the free state did not move straight toward the bound state, suggesting that an intermediate state detectable on the chemical shift timescale might be involved in the DNA binding of S100R MBDAtMBD6. These observations indicated that the off-rate of the binding was lower than that of WT MBDAtMBD6.

DNA is much lower in comparison to that of the canonical MBD domains, even though MBD AtMBD6 shares the DNA binding interface with its mammalian counterparts.

Consequence of Variation of Important Amino Acids in the AtMBD6 MBD Domain. To understand the comparably weak binding of MBD AtMBD6 to methyl-CpG-containing DNA, we performed sequence and structural alignment of MBD AtMBD6 with human and other A. thaliana MBD domains (Figures 4A, S4, and S5). MBD AtMBD6 possesses all the key residues that have been shown to be responsible for specific recognition of the methyl-CpG sites in canonical MBD domains,8,19 in line with the aforementioned results. However, two critical differences were found in the amino acid sequences. First, the C-terminal region of MBD AtMBD6 (residues 127−149) shows poor homology with its human counterparts. Previous studies on human MBD domains demonstrated that this region forms a “hairpin loop” structure that is also associated with DNA binding.23 Therefore, we hypothesize that the affinity of MBD AtMBD6 toward DNA is reduced as a consequence of the absence of the hairpin loop. Second, three positively charged residues conserved among human MBD proteins (R17, K23, and R30 in HsMBD1) correspond to uncharged residues (V87, T93, and S100, respectively) in AtMBD6. In canonical human MBD proteins, the first of these three residues stabilizes the hairpin loop by electrostatic interactions with backbone carbonyl groups in the hairpin loop (Figure 4B).19 The absence of this positively charged residue therefore supports our idea that the hairpin loop is missing in MBD AtMBD6 that is, not only in the free but also even in the DNA-bound state. The remaining two residues are located near the two ends of the L1 loop and contact the phosphate backbone of DNA in the reported structures of canonical MBD-DNA complexes (Figure 4C).18,19 Previous studies demonstrated that the positive charge of these residues plays a crucial role in the DNA binding by human and chicken MBD domains.18,23,31 Therefore, we hypothesized that the absence of these important electrostatic contributions to the binding energy significantly reduces the binding affinity of MBD AtMBD6 toward methyl-CpG-containing DNA.
and the exchange between the intermediate and bound states was in the slow exchange regime on the chemical shift timescale. Taken together, our results suggest that the S100R mutant has an increased affinity toward methyl-CpG-containing DNA as compared with WT MBDAtMBD6, in line with the hypothesis based on the sequence alignment. To obtain the $K_d$ value of this interaction, we tried several methods including this NMR titration, isothermal titration calorimetry, and fluorescence polarization. Unfortunately, however, we could not obtain a reliable $K_d$ value, likely because this interaction is not a simple two-state exchange and several properties of this interaction and MBDAtMBD6 itself such as binding enthalpy and molecular weight are not suitable for quantification of the binding affinity by these methods. Therefore, hereafter, we focus on the differences in the chemical shift change upon DNA binding between WT and S100R MBDAtMBD6. For most of the residues, the CSD values between the free and bound states were similar to the corresponding CSD values of WT MBDAtMBD6; however, the four residues (V101, D102, R117, and E119) showed large differences in the chemical shift change upon DNA binding between WT and S100R MBDAtMBD6 (Figure 4E, left panel; see also Figures 3B and S6B). Among these four residues, only R117 experienced a “downward” change, suggesting that this residue is less involved in the DNA binding in the S100R mutant compared to WT MBDAtMBD6. By contrast, V101, D102, and E119 showed “upward” changes, suggesting that these residues became more involved in DNA binding due to the introduction of the S100R mutation. It should also be noted that R117 and E119 are not close to the mutation site (more than 10 Å apart from S/R100). Intriguingly, we found that these four residues are located near either of the RFs (Figure 4E, right). Overall, the main distinction between the CSD values of WT and S100R MBDAtMBD6 results from the difference in the chemical shifts of the bound state, not the free state (Figures S6C,D). Furthermore, as the molar ratio of DNA over protein increased, an aliased side-chain signal emerged in the $^1$H−$^{15}$N HSQC spectra at a markedly field-shifted $^1$H chemical shift (Figure S6A). Based on a $^{15}$N-edited NOESY-HSQC spectrum, this resonance could be identified as the $\epsilon$-NH resonance of R92 of the structural element RF1. By contrast, this signal was not observed in the case of WT MBDAtMBD6, further underlining the differential involvement of RF1 in WT and S100R MBDAtMBD6. Thus, taken together, these results suggest that “exchanging” positively charged residues responsible for electrostatic interactions with the DNA backbone (in human MBD domains) to uncharged residues (in MBDAtMBD6) specifically affects local conformational states around the two distinct RFs in the bound state. In other words, the reduced binding affinity of MBDAtMBD6 toward methyl-CpG-containing DNA can be considered a result of the lack of important nonspecific interactions and the consequent alterations in the DNA

Figure 4. Consequence of variation of important amino acids in the AtMBD6 MBD domain. (A) Comparison of the amino acid sequence of MBDAtMBD6 with the sequences of human MBD domains. Identical and homologous residues are colored in black and gray, respectively. Asterisks indicate the key residues responsible for the specific binding of the canonical MBD domains to methyl-CpG-containing DNA. Daggers refer to the positively charged residues conserved among human MBD proteins but absent in AtMBD6. (B) Comparison of MBDAtMBD6 and the MBD domain of HsMeCP2 on the basis of presence and stabilization of the hairpin loop. Polar contacts are indicated by yellow dashed lines. In the description of the structure of HsMeCP2, the DNA moiety is omitted for clarity. (C) Comparison of MBDAtMBD6 and the MBD domain of HsMeCP2 on the coordinates of two important residues at the L1 loop that contact the DNA backbone. (D) Spectral changes of S100R MBDAtMBD6 upon binding to methyl-CpG-containing DNA (MG). For comparison to WT MBDAtMBD6 see also the enlarged inset in Figure 3A. (E) Correlation plot of the normalized CSD values upon binding between WT MBDAtMBD6 and the S100R mutant (left). Mapping of the residues displaying significant differences in the correlation on the structure of WT MBDAtMBD6 (right).
conformational CpG site stabilizes the dynamic L1 loop, thereby reducing its flexibility. Previous structural and chemical shift analyses indicated that these sequences are especially recognized methyl-CpG-containing DNA, suggesting that AtMBD6 is not directly involved in epigenetic regulation via methylation in CpHpG and CpHpH contexts. While we found that the L1 loop of MBDAtMBD6 shows high structural flexibility prior to DNA binding, interestingly, previous structural and chemical shift analyses indicated that the specific binding of human MBD domains to the methyl-CpG site stabilizes the dynamic L1 loop, thereby reducing its conformational flexibility. Since the amino acid composition in this region of AtMBD6 is similar to that of human MBD proteins, the L1 loop of MBDAtMBD6 would also recognize the major groove of DNA by hydrophobic interactions and hydrogen bonds to become rigid upon binding to methyl-CpG-containing DNA.

The C-terminal region of MBDAtMBD6 shows no noticeable sequence homology with human MBD domains and was found to be highly disordered in the solution structure, in line with the negative backbone heteronuclear NOE values reported in our initial study on MBDAtMBD5. This implies that the requirements for formation of a stabilized hairpin loop in plant MBD domains are the same as in human MBD domains: both a tripeptide motif FBF (B represents D or N) at the respective position in the C-terminal region (see Figure 4A) and a positively charged residue that forms polar contacts with the hairpin loop at the beginning of β1. According to amino acid sequence similarity (Figure S4), MBD proteins from A. thaliana can be classified into two groups by the presence and absence of the hairpin loop. However, our results indicate that the presence or absence of the hairpin loop alone does not correlate with the binding ability of the MBD domains to methyl-CpG-containing DNA in A. thaliana.

**DISCUSSION**

**Structural Properties of the AtMBD6 MBD Domain.** Among the 13 MBD proteins from A. thaliana, AtMBD6, AtMBD5, and AtMBD7 show the highest sequence homology in the MBD domain with mammalian MBD domains that specifically recognize the methyl-CpG sites. Correspondingly, the overall structure of MBDAtMBD6 in the free form is almost identical to the canonical structures of known mammalian MBD domains, except for the C-terminal region. The insertion of a single residue, N109, in the short loop between β2 and β3 of MBDAtMBD6 (Figure 4A) has almost no effect on the overall structure of the MBD core (Figure 2). On the basis of sequence homology, other MBD proteins from A. thaliana also seem to adopt the canonical MBD fold, although one or more residues in the L1 loop are missing except for AtMBD5 and AtMBD8 that do possess these residues (i.e., they have a “full” L1 loop).

While we found that the L1 loop of MBDAtMBD6 shows high structural flexibility prior to DNA binding, interestingly, previous structural and chemical shift analyses indicated that the specific binding of human MBD domains to the methyl-CpG site stabilizes the dynamic L1 loop, thereby reducing its conformational flexibility. Since the amino acid composition in this region of AtMBD6 is similar to that of human MBD proteins, the L1 loop of MBDAtMBD6 would also recognize the major groove of DNA by hydrophobic interactions and hydrogen bonds to become rigid upon binding to methyl-CpG-containing DNA.

The C-terminal region of MBDAtMBD6 shows no noticeable sequence homology with human MBD domains and was found to be highly disordered in the solution structure, in line with the negative backbone heteronuclear NOE values reported in our initial study on MBDAtMBD5. This implies that the requirements for formation of a stabilized hairpin loop in plant MBD domains are the same as in human MBD domains: both a tripeptide motif FBF (B represents D or N) at the respective position in the C-terminal region (see Figure 4A) and a positively charged residue that forms polar contacts with the hairpin loop at the beginning of β1. According to amino acid sequence similarity (Figure S4), MBD proteins from A. thaliana can be classified into two groups by the presence and absence of the hairpin loop. However, our results indicate that the presence or absence of the hairpin loop alone does not correlate with the binding ability of the MBD domains to methyl-CpG-containing DNA in A. thaliana.

**DNA Binding Specificity of the AtMBD6 MBD Domain.** The hitherto published reports of DNA binding specificity of AtMBD6, which were determined on the basis of gel shift assay experiments, are not in mutual agreement, for example, one group claims that AtMBD6 binds all types of methylated DNA in plants, whereas other groups claim that AtMBD6 specifically recognizes the methyl-CpG sites. Possible reasons for this discrepancy include artifacts derived from the unstructured regions of AtMBD6, the presence of an N-terminal affinity tag, and differences in the experimental buffer composition. To resolve this contradiction, we had prepared high-purity tag-free MBDAtMBD6 samples and qualitatively evaluated its DNA binding specificity using all types of methylated DNA at the same time, that is, on the same gel. Under the conditions employed here, MBDAtMBD6 specifically recognized methyl-CpG-containing DNA, suggesting that AtMBD6 is not directly involved in epigenetic regulation via methylation in CpHpG and CpHpH contexts.

**Mechanism of Weak Binding of AtMBD6 to Methyl-CpG-Containing DNA.** We revealed by solution NMR spectroscopy that MBDAtMBD6 possesses a similar binding interface to methyl-CpG-containing DNA as canonical mammalian MBD domains, albeit with one to three orders of magnitude lower binding affinity. One of the reasons for the comparably low affinity of MBDAtMBD6 is the absence of two positively charged residues that have been shown to electrostatically interact with the DNA backbone in other MBD domains (T93 and S100 in AtMBD6). NMR titration experiments using a single-point mutant to “correct” one of these atypical residues, S100R, indicated that the difference between MBDAtMBD6 and canonical mammalian MBD domains lies in the chemical environment in the vicinity of the two RFs (V101, D102, R117, and E119 in AtMBD6) in the DNA-bound state. However, the backbone amide CSD values of the RFs themselves were not significantly changed by the S100R mutation (see Figure S6D), indicating that only their side-
chain guanidino groups might be involved in the mechanistic difference in DNA binding affinity between AtMBD6 and canonical MBD proteins. A simple effect of the S100R mutation is an attractive electrostatic interaction between the side chain of the introduced arginine R100 and the DNA backbone, as observed in the reported structures of canonical MBD domains in complex with methyl-CpG-containing DNA. Another possible aspect of this mutation may be an electrostatic repulsion between the positively charged guanidino groups of R100 and the RF1, that is, R92. These effects seem to explain the observed changes in the vicinity of RF1 but not in RF2. Importantly, D121 and E137 in HsMeCP2 corresponding to D102 and E119 in AtMBD6 form salt bridges with the guanidino groups of the two RFs to fix them in appropriate positions for specific recognition of the methyl-CpG site in the complex (Figure 5A, left panel).19 Therefore, the observed differences between WT MBDAtMBD6 and the S100R mutant indicate that the absence of important local interactions with the DNA backbone in MBDAtMBD6 makes the retention of RF1 and RF2 (cooperative binding) in a proper conformation insufficient (Figure 5A, right panel, Figure 5B). Indeed, the downfield-shifted RF1 side-chain resonance is reminiscent of the methyl-CpG-specific binding mode exhibited by the RF1 motif of several MBD domains from human and even the invertebrate Ephidatia muelleri.24,25 The facts that in the complex, the side chain of D121 in HsMeCP2 also forms a polar contact with its backbone amide group (3.2 Å distance between one of the two carboxy oxygen atoms and the amide nitrogen atom) and that D102 in WT MBDAtMBD6 showed only a small CSD value upon DNA binding support our conclusion that the RF1 motif is not properly formed in WT MBDAtMBD6. Collectively, the conformational incompleteness of the RF motifs in AtMBD6 in the complex appears to increase the off-rate, resulting in a significant reduction in the binding affinity toward methyl-CpG-containing DNA, although we stress again that MBDAtMBD6 does still preserve the binding specificity.

Our analysis further indicates that the DNA binding affinity of AtMBD5 is likely to be even lower than that of AtMBD6, as the important tyrosine (Y104 in AtMBD6) is substituted to phenylalanine (Figures 2C and S4) and, thus, AtMBD5 could not form the water-mediated hydrogen bond network that has been implicated in DNA base recognition by HsMeCP2.13 The sequence homology of the MBD domains of AtMBD7 with mammalian MBD domains is slightly lower than that of AtMBD6. Therefore, the DNA binding affinity of each individual MBD unit is likely to be even lower than that of AtMBD6. However, AtMBD7 may still retain a moderate binding affinity toward methyl-CpG-containing DNA by cooperative binding of the three tandem MBD domains to multiple methyl-CpG sites.

Possible Mechanism of Subnuclear Localization of AtMBD6 Haboring an MBD Domain with Low DNA Binding Affinity. Although the binding affinity of the MBD domain toward methyl-CpG-containing DNA is significantly reduced, AtMBD6 localizes to highly methylated heterochromatin.15,34 Given that 24% of the CpG sites in genomic DNA are methylated in A. thaliana,35 the density of the methyl-CpG sites may aid proper subnuclear localization of AtMBD6 even with a comparably low DNA binding affinity. In addition, regions other than the MBD domain itself might also contribute to binding of AtMBD6 toward methyl-CpG-containing DNA, although at present, no significant sequence homology with any known DNA-binding domain is detected for the N- or C-terminal regions of AtMBD6.

Another possible explanation for the localization of AtMBD6 to heterochromatin would be that binding partners compensate for the low DNA binding affinity of AtMBD6. One of the putative AtMBD6-binding proteins, AtDDM1, possesses a DNA helicase domain and was shown to bind to both free DNA and nucleosomal DNA.39 Another potential binding partner of AtMBD6, AtAGO4, plays a crucial role in the RNA-directed DNA methylation and thus might also bind to DNA with sequence preference. Indeed, a previous study showed by fluorescence recovery analysis that AtMBD6 is highly mobile at perinucleolar chromocenters, while a small fraction of the protein is relatively immobile.17 We assume that the immobile fraction that tightly binds to chromatin is involved in formation of protein–protein complexes and the mobile fraction corresponds to the free molecules.

F CONCLUSIONS

In this study, we revealed that MBDAtMBD6 specifically recognizes methyl-CpG sites and characterized its structural properties at the atomic level. MBDAtMBD6 binds methyl-CpG-containing DNA with a significantly reduced affinity compared to mammalian MBD domains, while its DNA binding interface is conserved. Future structural studies of AtMBD6 in complex with methyl-CpG-containing DNA and analysis of its interactions with previously identified binding partners ought to pave the way of understanding the mechanism by which AtMBD6 serves as a repressor of gene expression in plant cells.

M MATERIALS AND METHODS

Sample Preparation. The expression vector pGEX-6P-1 encoding the MBD domain of AtMBD6 (UniProtKB entry Q9LTJ1, residues 78–140) was used for protein overexpression in bacterial cells. Site-directed mutagenesis to generate S100R MBDAtMBD6 was achieved by inverse PCR. Proteins were prepared as previously described21 with a few modifications. Briefly, Escherichia coli BL21 (DE3) cells carrying the expression plasmids were grown in Lennox’s LB media or M9 minimal media containing 2 g/L uniformly 13C-labeled 0-glucose (Cambridge Isotope Laboratories) or 4 g/L unlabeled 0-glucose (Nacalai Tesque), 1 g/L 15N-labeled ammonium chloride (Cambridge Isotope Laboratories), and as antibiotic selection pressure 100 mg/L ampicillin. Proteins were expressed for 20 h at 16 °C and purified using glutathione affinity chromatography. After on-column digestion of the N-terminal glutathione S-transferase-tag by HRV3C protease overnight at 4 °C, proteins were eluted from the column and further purified by size-exclusion chromatography. Purity and integrity of the samples were verified by SDS-PAGE and MALDI-TOF mass spectrometry. The yields of proteins per 1 L culture were 4–5 mg for LB and 2–4 mg for M9.

All DNA samples used in this study were purchased from Hokkaido System Science. These DNA sequences are summarized in Table 1. Single-stranded DNA samples were dissolved in 10 mM Tris buffer (pH 8.0 at 4 °C) containing 50 mM sodium chloride for the gel shift assay or in the titration buffer containing 20 mM sodium phosphate (pH 6.5 at 25 °C) and 150 mM sodium chloride for NMR titration experiments. Annealing was performed by incubating the solutions for 3 min at 98 °C and then gradually lowering the temperature to 4 °C over a total duration of 140 min.
DNA Binding Assay. The mixture of 80 pmol annealed DNA (final 4 μM) and 400 pmol protein (final 20 μM) was incubated in 20 μL of the binding buffer containing 25 mM Tris, 25 mM boric acid, 150 mM sodium chloride, 5% glycerol, and 1 mM dithiothreitol for 30 min at 4 °C. Then, 5 μL of the reaction sample was loaded onto an 8% polyacrylamide gel containing 25 mM Tris and 25 mM boric acid and subjected to electrophoresis at 150 V for 60 min at 4 °C (i.e., in the cold room). Double-stranded DNA was stained with ethidium bromide, and the corresponding fluorescent bands were detected by UV irradiation using a Chemidoc XRS Plus system (Bio-Rad).

General NMR Spectroscopy. All NMR experiments were performed at 20 °C using an AVANCE II 700 MHz NMR spectrometer (Bruker) equipped with a TCI cryogenic probe with the exception of in-phase/anti-phase (IPAP) HSQC experiments, which were performed on an AVANCE 600 MHz NMR spectrometer (Bruker) equipped with a TXI cryogenic probe. NMR samples contained 5% deuterium oxide (Cambridge Isotope Laboratories) and were measured in 5 mm-diameter Shigemi tubes (Shigemi). Sodium 3-(trimethylsilyl)-1-propanesulfonate (Tokyo Chemical Industry) was used as an external standard of the 1H chemical shift; 13C and 15N chemical shifts were calibrated indirectly.40 All acquired NMR data (i.e., free induction decays) were processed using NMRPipe,41 and the resulting spectra were further analyzed using various software packages (see below).

NMR Analysis of the Homo-oligomerization Status of the AtMBD6 MBD Domain. Two-dimensional 1H−15N HSQC spectra of 13C, 15N-labeled MBDAtMBD6 in the titration buffer were acquired with sequentially decreasing the concentration of the protein by twofold dilution. The concentration of MBDAtMBD6 in these experiments ranged from 200 to 3 μM. Analysis of the obtained spectra was performed using CcpNmr Analysis.42

NOESY Experiments. Three-dimensional 15N-edited NOESY-HSQC spectra of 13C, 15N-labeled MBDAtMBD6 in the titration buffer were acquired with sequentially decreasing the concentration of the protein by twofold dilution. The concentration of MBDAtMBD6 in these experiments ranged from 200 to 3 μM. Analysis of the obtained spectra was performed using CcpNmr Analysis.42

NOESY Experiments. Three-dimensional 15N-edited NOESY-HSQC spectra of 13C, 15N-labeled MBDAtMBD6 in the titration buffer were acquired with sequentially decreasing the concentration of the protein by twofold dilution. The concentration of MBDAtMBD6 in these experiments ranged from 200 to 3 μM. Analysis of the obtained spectra was performed using CcpNmr Analysis.42

IPAP HSQC Experiments. IPAP 1H−15N HSQC experiments were performed using 0.5 mM 13C, 15N-labeled MBDAtMBD6 samples in the presence and absence of pentaethylene glycol monododecyl ether (C12E5) biceps as alignment medium.45 The bicelle solution was prepared by gradually adding a total of 15 μL 1-hexanol (Tokyo Chemical Industry) to a solution composed of 50 μL of C12E5 (Sigma-Aldrich), 200 μL of the NMR buffer, and 50 μL of deuterium oxide. The final concentration of C12E5 in the NMR sample was 4%. IPAP HSQC (F1: 7246.377 Hz/S12 complex points, F2: 1520.447 Hz/S350 complex points) experiments were conducted consecutively twice for each sample. Each acquired IPAP spectrum was split into two distinct spectra (IP + AP and IP − AP) in TopSpin (Bruker) and analyzed with CcpNmr Analysis. The averaged backbone amide RDC constants were used to generate RDC restraints in the structure refinement.

Structure Determination of the AtMBD6 MBD Domain in the Free Form. The initial structure calculation of MBDAtMBD6 in the free form was performed using CYANA40 with interproton distance restraints derived from NOESY spectra (NOE peaks were automatically assigned based on the chemical shifts of backbone and side-chain resonances reported previously21 and backbone dihedral angle restraints estimated by TALOS+.22 The lowest-energy structure of the 20 structures calculated using CYANA was used as an initial structure in the subsequent refinement procedure. Structure refinement was performed using 500 cycles of a combination of simulated annealing and energy minimization using XPLOR-NIH19,64 while additionally applying RDC-based restraints. The top 20 structures of the 41 structures that meet the acceptance criteria are reported here with the statistics shown in Table 2. The coordinates of the solution structure of MBDAtMBD6 in the free form were deposited in the Protein Data Bank with accession ID 7D8K.

NMR Titration and Resonance Assignment of the AtMBD6 MBD Domain in the Bound State. A solution of 4.5 mM double-stranded methyl-CpG-containing DNA termed as MG (see Table 1) was added in a stepwise manner to 0.6 mM 13C, 15N-labeled WT MBDAtMBD6 or 15N-labeled S100R MBDAtMBD6 in the titration buffer. After each addition of the ligand solution, a 1H−15N HSQC spectrum was acquired. To obtain backbone resonance assignments of WT MBDAtMBD6 in the bound state, HNCO, HN(CA)CO, CBCA(CO)NH, and HNCACB experiments40 were carried out after the final titration experiment. The triple resonance spectra were analyzed using MagRO-NMRView, and initial assignments were performed using FLYA51 followed by manual verification and correction of the automated assignments. For each residue, the SSP score was evaluated from its chemical shifts of amide proton, amide nitrogen, α-carbon, β-carbon, and carbonyl carbon with the SSP program.52 In the case of the S100R mutant, resonances of backbone amide groups were assigned using three-dimensional 15N-edited NOESY-HSQC spectra with CcpNmr Analysis. The normalized CSD value was calculated as follows53

\[
\Delta \delta = \frac{1}{2} \left( \Delta \delta_{1H}^2 + \left( \frac{\Delta \delta_{15N}}{7} \right)^2 \right)
\]

in which \(\Delta \delta_{1H}\) and \(\Delta \delta_{15N}\) are the CSD values in the 1H and 15N dimensions of the 1H−15N HSQC spectra, respectively. The titration curves of WT MBDAtMBD6 were fitted to a classical two-state binding model using GLOVE34 to derive the dissociation constant \(K_d\) as a global parameter. In this model, the observed chemical shift change \(\Delta \delta_{obs}\) (where \(\Delta \delta_{obs}\) is quantified separately for the 1H and 15N dimensions) is represented as follows28

\[
\Delta \delta_{obs} = \frac{\Delta \delta_{fre}}{2|P|_0} \left[ |P|_0 + |I|_0 + K_d \right] - \sqrt{(|P|_0 + |I|_0 + K_d)^2 - 4|P|_0|I|_0}.
\]

in which \(\Delta \delta_{obs}\) is the CSD between the free state and bound state. \(|P|_0\) and \(|L|_0\) denote the total concentrations of the protein and ligand, respectively. Fitting errors were estimated.
by 100 steps of Monte Carlo simulation implemented in GLOVE.

**ASSOCIATED CONTENT**

Supporting Information
The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.1c04917.

Gel shift assay of DNA binding by the AtMBD6 MBD domain at different protein concentrations; overlaid $^1$H-$^15$N HSQC spectrum of the AtMBD6 MBD domain at different concentrations; comparison of the SSP scores of the AtMBD6 MBD domain in the free and DNA-bound states; sequence alignment of the MBD domains in *A. thaliana*; structural superposition of the AtMBD6 MBD domain and human MBD domains; and details on the binding of the S100R AtMBD6 MBD domain to methyl-CpG-containing DNA (PDF)

**Accession Codes**
The UniProtKB entry ID of MBD6 from *A. thaliana* is Q9L7J1. The atomic coordinates of the solution structure of the AtMBD6 MBD domain in the free form were deposited in the Protein Data Bank as an entry 7D8K.

**AUTHOR INFORMATION**

Corresponding Author
Masahiro Shirakawa — Department of Molecular Engineering, Kyoto University, Kyoto 615-8510, Japan; Phone: (81) 75-383-2535; Email: shirakawa@moleng.kyoto-u.ac.jp; Fax: (81) 75-383-2541

Authors
Yutaka Mahana — Department of Molecular Engineering, Kyoto University, Kyoto 615-8510, Japan; orcid.org/0000-0001-6218-6726

Izuru Ohki — Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan

Erik Walinda — Graduate School of Medicine, Kyoto University, Kyoto 606-8501, Japan; orcid.org/0000-0003-1882-6401

Daichi Morimoto — Department of Molecular Engineering, Kyoto University, Kyoto 615-8510, Japan; orcid.org/0000-0002-7672-2136

Kenji Sugase — Department of Molecular Engineering, Kyoto University, Kyoto 615-8510, Japan; orcid.org/0000-0001-8623-7743

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.1c04917

Author Contributions
Y.M. conducted all experiments. I.O. established the concept of this study. Y.M. and E.W. wrote the manuscript. E.W., D.M., and K.S. participated in experimental design and NMR data analysis. M.S. obtained funding. All authors reviewed and commented on the manuscript.

Notes
The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

This work was supported by the Japan Agency for Medical Research and Development Grant JP16gm0510004 and Japan Society for the Promotion of Science KAKENHI Grants JP20H00453 and JP18K19297.

**REFERENCES**

(1) Goll, M. G.; Bestor, T. H. Eukaryotic Cytosine Methyltransferases. *Annu. Rev. Biochem.* 2005, 74, 481–514.

(2) Law, J. A.; Jacobsen, S. E. Establishing, Maintaining and Modifying DNA Methylation Patterns in Plants and Animals. *Nat. Rev. Genet.* 2010, 11, 204–220.

(3) Li, E.; Zhang, Y. DNA Methylation in Mammals. *Cold Spring Harbor Perspect. Biol.* 2014, 6, a019133.

(4) Meehan, R. R.; Lewis, J. D.; McKay, S.; Kleiner, E. L.; Bird, A. P. Identification of a Mammalian Protein That Binds Specifically to DNA Containing Methylated CpGs. *Cell* 1989, 58, 499–507.

(5) Hendrich, B.; Bird, A. Identification and Characterization of a Family of Mammalian Methyl-CpG Binding Proteins. *Mol. Cell. Biol.* 1998, 18, 6538–6547.

(6) Du, Q.; Luu, P.-L.; Stirzaker, C.; Clark, S. J. Methyl-CpG-Binding Domain Proteins: Readers of the Epigenome. *Epigenomics* 2015, 7, 1051–1073.

(7) Hendrich, B.; Hardeeland, U.; Ng, H.-H.; Jiricny, J.; Bird, A. The Thymine Glycosylase MBD4 Can Bind to the Product of Deamination at Methylated CpG Sites. *Nature* 1999, 401, 301–304.

(8) Otani, J.; Arita, K.; Kato, T.; Kinoshita, M.; Kimura, H.; Suetake, I.; Tajima, S.; Ariyoshi, M.; Shirakawa, M. Structural Basis of the Versatile DNA Recognition Ability of the Methyl-CpG Binding Domain of Methyl-CpG Binding Domain Protein 4. *J. Biol. Chem.* 2013, 288, 6351–6362.

(9) Henderson, I. R.; Jacobsen, S. E. Epigenetic Inheritance in Plants. *Nature* 2007, 447, 418–424.

(10) Cokus, S. J.; Feng, S.; Zhang, X.; Chen, Z.; Merriman, B.; Haudenschild, C. D.; Pradhan, S.; Nelson, S. F.; Pellegrini, M.; Jacobsen, S. E. Shotgun Bisulphite Sequencing of the Arabidopsis Genome Reveals DNA Methylation Patterning. *Nature* 2008, 452, 215–219.

(11) Matzke, M. A.; Mosher, R. A. RNA-Directed DNA Methylation: An Epigenetic Pathway of Increasing Complexity. *Nat. Rev. Genet.* 2014, 15, 394–408.

(12) Zhang, H.; Lang, Z.; Zhu, J.-K. Dynamics and Function of DNA Methylation in Plants. *Nat. Rev. Mol. Cell Biol.* 2018, 19, 489–506.

(13) Garg, R.; Narayana Chevala, V.; Shankar, R.; Jain, M. Divergent DNA Methylation Patterns Associated with Gene Expression in Rice Cultivars with Contrasting Drought and Salinity Stress Response. *Sci. Rep.* 2015, 5, 14922.

(14) Springer, N. M.; Kaeppler, S. M. Evolutionary Divergence of Monocot and Dicot Methyl-CpG-Binding Domain Proteins. *Plant Physiol.* 2005, 138, 92–104.

(15) Zemach, A.; Li, Y.; Wayburn, B.; Ben-Meir, H.; Kiss, V.; Avivi, Y.; Kalchenko, V.; Jacobsen, S. E.; Grafi, G. DDM1 Binds Arabidopsis Methyl-CpG Binding Domain Proteins and Affects Their Subnuclear Localization. *Plant Cell* 2005, 17, 1549–1558.

(16) Parida, A. P.; Sharma, A.; Sharma, A. K. AtMBD6, a Methyl CpG Binding Domain Protein, Maintains Gene Silencing in Arabidopsis by Interacting with RNA Binding Proteins. *J. Biosci.* 2017, 42, 57–68.

(17) Zemach, A.; Gaspan, O.; Grafi, G. The Three Methyl-CpG-Binding Domains of AtMBD7 Control Its Subnuclear Localization and Mobility. *J. Biol. Chem.* 2008, 283, 8406–8411.

(18) Ohki, I.; Shimotake, N.; Fujita, N.; Jee, J.-G.; Ikegami, T.; Nakao, M.; Shirakawa, M. Solution Structure of the Methyl-CpG Binding Domain of Human MBD1 in Complex with Methylated DNA. *Cell* 2001, 105, 487–497.

(19) Ho, K. L.; McNae, I. W.; Schmiedeberg, L.; Klose, R. J.; Bird, A. P.; Walkinshaw, M. D. McCP2 Binding to DNA Depends upon Hydration at Methyl-CpG. *Mol. Cell* 2008, 29, 525–531.

(20) Walavalkar, N. M.; Cramer, J. M.; Buchwald, W. A.; Scarsdale, J. N.; Williams, D. C. Solution Structure and Intramolecular Exchange of Methyl-Cytosine Binding Domain Protein 4 (MBD4) on DNA Suggests a Mechanism to Scan for M CpG/TpG Mismatches. *Nucleic Acids Res.* 2014, 42, 11218–11232.
ally Active and at Least One, AtMBD11, Is Crucial for Normal Encoding. Methyl-CpG-Binding Domain Proteins Are Transcription-K.; Aalen, R. B. Ten Members of the Arabidopsis Gene Family were characterized. Hall, C.; Siraliev-Perez, E.; Walavalkar, N. M.; Sperlazza, M. J.; Williams, D. C. Probing the Dynamic Solution Structure of the Methyl-CpG-Binding Domain of the MBD6 from Arabidopsis Thaliana. Dimethylation in Vivo. Mechanism for 5mC Recognition by the Arabidopsis SUVH5 SRA Domain. L. M.; Reinberg, D.; Patel, D. J.; Jacobsen, S. E. A Dual Flip-out Mechanism for 5mC Recognition by the Arabidopsis SUVH5 SRA Domain and Its Impact on DNA Methylation and H3K9 Dimethylation in Vivo. Jacobsen, S. E. Mechanism of DNA Methylation-Directed Histone Methylation by KRYPTONITE. Mol. Cell 2014, 55, 495–504.

(23) Ohki, I.; Shimotake, N.; Fujita, N.; Nakao, M.; Shirakawa, M. Solution Structure of the Methyl-CpG-Binding Domain of the Methyl-CpG-Dependent Transcriptional Repressor MBDD1. EMBO J. 1999, 18, 6655–6661.

(24) Liu, K.; Xu, C.; Lei, M.; Yang, A.; Loppnau, P.; Hughes, T. R.; Min, J. Structural Basis for the Ability of MBD Domains to Bind Methyl-CG and CG T Sites in DNA. J. Biol. Chem. 2018, 293, 7344–7354.

(25) Cramer, J. M.; Scarsdale, J. N.; Walavalkar, N. M.; Buchwald, W. A.; Ginder, G. D.; Williams, D. C. Probing the Dynamic Distribution of Bound States for Methylcytosine-Binding Domains on DNA. J. Biol. Chem. 2014, 289, 1294–1302.

(26) Cramer, J. M.; Pohlmann, D.; Gomez, F.; Mark, L.; Kornegay, B.; Hall, C.; Siralese-Perez, E.; Walavalkar, N. M.; Sperlazza, M. J.; Bilinovich, S.; Prokop, J. W.; Hill, A. L.; Williams, D. C. Methylation Specific Targeting of a Chromatin Remodeling Complex from Sponges to Humans. Sci. Rep. 2017, 7, 40674.

(27) Sperlazza, M. J.; Bilinovich, S. M.; Sinanan, L. M.; Javier, F. R.; Williams, D. C. Structural Basis of MeCP2 Distribution on Non-CpG Methylated and Hydroxymethylated DNA. J. Mol. Biol. 2017, 429, 1581–1594.

(28) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G., III; Rance, M.; Skelton, N. J. Protein NMR Spectroscopy, 2nd Ed.; Elsevier, 2007.

(29) Klekner, I. R.; Foster, M. P. An Introduction to NMR-Based Approaches for Measuring Protein Dynamics. Biochim. Biophys. Acta 2011, 1814, 942–968.

(30) Hashimoto, H.; Liu, Y.; Upadhyay, A. K.; Chang, Y.; Howerton, S. B.; Vertino, P. M.; Zhang, X.; Cheng, X. Recognition and Potential Mechanisms for Replication and Erasure of Cytosine Hydroxymethylation. Nucleic Acids Res. 2012, 40, 4841–4849.

(31) Scarsdale, J. N.; Webb, H. G.; Ginder, G. D.; Williams, D. C. Solution Structure and Dynamic Analysis of Chicken MBDD2 Methyl Binding Domain Bound to a Target-Methylated DNA Sequence. Nucleic Acids Res. 2011, 39, 6741–6752.

(32) Zernach, A.; Grafi, G. Characterization of Arabidopsis Thaliana Methyl-CpG-Binding Domain (MBD) Proteins. Plant J. 2003, 34, 565–572.

(33) Berg, A.; Meza, T. J.; Mahić, M.; Thorstensen, T.; Kristiansen, A.; Aalen, R. B. Ten Members of the Arabidopsis Gene Family Encoding Methyl-CpG-Binding Domain Proteins Are Transcriptionally Active and at Least One, AtMBD11, Is Crucial for Normal Encoding. Nucleic Acids Res. 2003, 31, S291–S304.

(34) Scbebba, F.; Bernacchia, G.; De Bastiani, M.; Evangelista, M.; Cantonii, R. M.; Cellia, R.; Locci, M. T.; Pittio, L. Arabidopsis MBD Proteins Show Different Binding Specificities and Nuclear Localization. Plant Mol. Biol. 2003, 53, 755–771.

(35) Ito, M.; Koike, A.; Koizumi, N.; Sano, H. Methylated DNA-Binding Proteins from Arabidopsis. Plant Physiol. 2003, 133, 1747–1754.

(36) Johnson, L. M.; Bostick, M.; Zhang, X.; Kraft, E.; Henderson, I.; Callis, J.; Jacobsen, S. E. The SRA Methyl-Cytosine-Binding Domain Links DNA and Histone Methylation. Curr. Biol. 2007, 17, 379–384.

(37) Rajakumara, E.; Law, J. A.; Simanshu, D. K.; Voigt, P.; Johnson, L. M.; Reinberg, D.; Patel, D. J.; Jacobsen, S. E. A Dual Flip-out Mechanism for 5mC Recognition by the Arabidopsis SUVH5 SRA Domain and Its Impact on DNA Methylation and H3K9 Dimethylation in Vivo. Genes Dev. 2011, 25, 137–152.

(38) Du, J.; Johnson, L. M.; Groth, M.; Feng, S.; Hale, C. J.; Li, S.; Vashihsit, A. A.; Gallego-Bartolome, J.; Wohlschlegel, J. A.; Patel, D. J.; Jacobson, S. E. Mechanism of DNA Methylation-Directed Histone Methylation by KRYPTONITE. Mol. Cell 2014, 55, 495–504.

(39) Brzeski, J.; Jerzmanowski, A. Deficient in DNA Methylation 1 (DDM1) Defines a Novel Family of Chromatin-Remodeling Factors. J. Biol. Chem. 2003, 278, 823–828.

(40) Markley, J. L.; Bax, A.; Arata, Y.; Hilbers, C. W.; Kaptein, R.; Sykes, B. D.; Wright, P. E.; Wüthrich, K. Recommendations for the Presentation of NMR Structures of Proteins and Nucleic Acids. Pure Appl. Chem. 1998, 70, 117–142.

(41) Delaglio, F.; Grzesiek, S.; Vuister, G.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. J. Biomol. NMR 1995, 6, 277–293.

(42) Franken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. Proteins 2005, 59, 687–696.

(43) Kobayashi, N.; Iwahara, J.; Koshiba, S.; Tomizawa, T.; Tochio, N.; Güntert, P.; Kigawa, T.; Yokoyama, S. KUJIRA, A Package of Integrated Modules for Systematic and Interactive Analysis of NMR Data Directed to High-Throughput NMR Structure Studies. J. Biomol. NMR 2007, 39, 31–52.

(44) Johnson, B. A.; Blevins, R. A. NMR View: A Computer Program for the Visualization and Analysis of NMR Data. J. Biomol. NMR 1994, 4, 603–614.

(45) Rücker, M.; Otting, G. Alignment of Biological Macromolecules in Novel Nonionic Liquid Crystalline Media for NMR Experiments. J. Am. Chem. Soc. 2000, 122, 7793–7797.

(46) Downing, A. K. Protein NMR Techniques, 2nd ed.; Humana Press, 2004; Vol. 278.

(47) Shen, Y.; Delaglio, F.; Cornils, Gu.; Bax, A. TALOS+: A Hybrid Method for Predicting Protein Backbone Torsion Angles from NMR Chemical Shifts. J. Biomol. NMR 2009, 44, 213–223.

(48) Schwieters, C. D.; Kuszniewski, J. J.; Tjandra, N.; Marius Clore, G. The Xplor–NIH NMR Molecular Structure Determination Package. J. Magn. Reson. 2003, 160, 65–73.

(49) Schwieters, C.; Kuszniewski, J.; Marius Clore, G. Using Xplor–NIH for NMR Molecular Structure Determination. Proc. Nal. Mag. Reson. Spectrosc. 2006, 48, 47–62.

(50) Sattler, M.; Schleucher, J.; Griesinger, C. Heteronuclear Multidimensional NMR Experiments for the Structure Determination of Proteins in Solution Employing Pulsed Field Gradients. Proc. Nal. Mag. Reson. Spectrosc. 1999, 34, 93–158.

(51) Schmidt, E.; Güntert, P. A New Algorithm for Reliable and General NMR Resonance Assignment. J. Am. Chem. Soc. 2012, 134, 12817–12829.

(52) Marsh, J. A.; Singh, V. K.; Jia, Z.; Forman-Kay, J. D. Sensitivity of Secondary Structure Propensities to Sequence Differences between α- and γ-Synuclein: Implications for Fibrillation. Protein Sci. 2006, 15, 2795–2804.

(53) Williamson, M. P. Using Chemical Shift Perturbation to Characterise Ligand Binding. Proc. Nal. Mag. Reson. Spectrosc. 2013, 73, 1–16.

(54) Sugase, K.; Konuma, T.; Lansing, J. C.; Wright, P. E. Fast and Accurate Fitting of Relaxation Dispersion Data Using the Flexible Software Package GLOVE. J. Biomol. NMR 2013, 56, 275–283.