The SRA Methyl-Cytosine-Binding Domain Links DNA and Histone Methylation

Lianna M. Johnson,1,2,6 Magnolia Bostick,2,6 Xiaoyu Zhang,2 Edward Kraft,1,6 Ian Henderson,2 Judy Callis,4,5 and Steven E. Jacobsen2,3,*

1 Life Science Core Curriculum
2 Department of Molecular, Cell, and Developmental Biology
3 Howard Hughes Medical Institute
University of California, Los Angeles
Los Angeles, California 90095
4 Plant Biology Graduate Group
5 Department of Molecular and Cellular Biology
University of California, Davis
Davis, California 95616

Summary
Epigenetic gene silencing suppresses transposon activity and is critical for normal development [1, 2]. Two common epigenetic gene-silencing marks are DNA methylation and histone H3 lysine 9 dimethylation (H3K9me2) [1]. In Arabidopsis thaliana, H3K9me2, catalyzed by the methyltransferase KRYPTONITE (KYP/SUVH4), is required for maintenance of DNA methylation outside of the standard CG sequence context [3, 4]. Additionally, loss of DNA methylation in the met1 mutant correlates with a loss of H3K9me2 [5, 6]. Here we show that KYP-dependent H3K9me2 is found at non-CG methylation sites in addition to those rich in CG methylation. Furthermore, we show that the SRA domain of KYP binds directly to methylated DNA, and SRA domains with missense mutations found in loss-of-function kyp mutants have reduced binding to methylated DNA in vitro. These data suggest that DNA methylation is required for the recruitment or activity of KYP and suggest a self-reinforcing loop between histone and DNA methylation. Lastly, we show that SRA domains from two Arabidopsis SRA-RING proteins also bind methylated DNA and that the SRA domains from KYP and SRA-RING proteins prefer methylcytosines in different sequence contexts. Hence, unlike the methyl-binding domain (MBD), which binds only methylated-CpG sequences, the SRA domain is a versatile new methyl-DNA-binding motif.

Results and Discussion
Histone H3K9 Dimethylation Is Correlated with Levels of DNA Methylation
We utilized chromatin immunoprecipitation (ChIP) and genomic bisulfite sequencing to further investigate the relationship between histone methylation and DNA methylation. In Arabidopsis, DNA methylation (5-methyl cytosine) occurs not only at CG residues, but also at CNG and asymmetric cytosines (CNN) [2]. METHYLTRANSFERASE 1 (MET1), a homolog of mammalian Dnmt1, maintains methylation at CG residues [7, 8], whereas CHROMOMETHYLASE3 (CMT3) maintains the majority of CNG methylation and a small amount of the CNN methylation [9, 10]. A third enzyme, DOMAINS REARRANGED METHYLASE 2 (DRM2), which is mainly responsible for maintaining CNN methylation, plays a minor role in the maintenance of CNG methylation and is also required for de novo methylation of cytosines in all sequence contexts [2, 11]. Genomic regions that are enriched in methylated DNA are also found associated with H3K9me2 [12–14].

It has previously been shown that mutations in KYP lead to reduction of non-CG DNA methylation, but not CG methylation, suggesting that non-CG methylation is dependent on histone H3K9me2 [3, 4]. CMT3 is responsible for this non-CG methylation and has been shown in vitro to bind directly to H3 tails trimethylated at lysine 9 and lysine 27 [9, 10, 15]. Conversely, the loss of DNA methylation in met1 mutants is also associated with decreased H3K9me2 at many loci [5, 6, 12, 16]. One possible explanation for this observation is that KYP is in a complex with MET1, such that loss of MET1 causes a loss of KYP activity. A second possibility is that DNA methylation itself directly affects KYP activity or targeting.

To test these ideas, we examined H3K9me2 levels in different DNA methyltransferase mutants and at two loci with very different patterns of CG or non-CG methylation (see Figure S1 in the Supplemental Data available online). AtSN1 is a 170 bp SINE retrotransposon heavily methylated at non-CG and to a lesser extent at CG sites (~13 versus 3 per element; Figure 1A, see also [17]). AtSN1 is also marked by H3K9me2, which is dependent on KYP, illustrated by its loss in a kyp null line, kyp-6 (Figure 1B). Elimination of most of the DNA methylation in the drm1 drm2 cmt3 triple mutant also results in significant loss of H3K9me2 methylation, showing a strong correlation between levels of DNA and histone methylation (Figure 1B). In contrast, the met1-3 mutation showed complete elimination of CG methylation from AtSN1, but only a minor effect on non-CG methylation (Figure 1A), and H3K9me2 was observed at wild-type levels (Figure 1B). This suggests that KYP activity is not affected by loss of MET1 per se and that histone H3 methylation can be maintained at genomic regions containing only non-CG methylation.

Unlike AtSN1, AtCOPIA4 is a CG-rich retrotransposon [18]. Bisulfite sequencing of a 500 nucleotide segment revealed an average of 22 methylated CG residues, 6 methylated CNGs, and 3 methylated asymmetric cytosines in wild-type plants (Figure 1C, see also [19]). This element is also associated with KYP-dependent H3K9me2 (Figure 1D; see also [18]). In drm1 drm2 cmt3, the CNG methylation was greatly reduced, while the CG and histone methylation levels were unchanged (Figures 1C and 1D). However, in met1-3, all DNA
methyl-DNA binding in the KYP amino terminus. Further analysis with various double-stranded oligonucleotides containing only CG, CNN, or CNN methylation revealed that the SRA domain in KYP binds to DNA with methylated cytosines in all contexts with a preference for CNG and CNN sequences (Figure 3B). This preference was supported by competition experiments (Figures 3C and 3H).

To provide additional evidence that the SRA domain was indeed responsible for methyl-DNA binding, we engineered two different SRA mutations into the GST-KYP SRA fusion protein (amino acids 335–601) did not bind to methylated or unmethylated DNA. The addition of an anti-GST antibody to the all-5mC binding reaction blocked overall binding and a reduced ability of the mutant protein (amino acids 1–334, Figure 2C) to discriminate between methylated and unmethylated oligonucleotides with the same sequence (all-C) compared to unmethylated oligonucleotides (Figure 3A). In contrast, a GST-tagged KYP-SET fusion protein (amino acids 335–601) did not bind to methylated or unmethylated DNA. The addition of an anti-GST antibody to the all-5mC binding reaction blocked DNA binding (last two lanes in Figure 3A). These results indicate that the amino terminus of KYP containing the SRA domain binds methylated but not unmethylated DNA.

Further analysis with various double-stranded oligonucleotides containing only CG, CNN, or CNN methylation revealed that the SRA domain in KYP binds to DNA with methylated cytosines in all contexts with a preference for CNG and CNN sequences (Figure 3B). This preference was supported by competition experiments (Figures 3C and 3H).

To provide additional evidence that the SRA domain was indeed responsible for methyl-DNA binding, we engineered two different SRA mutations into the GST-KYP SRA construct (E208K and S200F). We observed weaker overall binding and a reduced ability of the mutant proteins to discriminate between methylated and unmethylated oligonucleotides (Figures 3D and 3E). These results strongly suggest that the SRA domain is responsible for methyl-DNA binding in the KYP amino terminus.

**SUVH6 SRA Domain Has a Binding Specificity Similar to KYP**

In addition to KYP, SRA domains are found in the eight DNA methyltransferase genes (SUVH class) in Arabidopsis [20]. We tested one of these, SUVH6, to determine whether it too could discriminate between methylated and unmethylated DNA. GST-SUVH6 full-length protein (amino acids 1–781) bound methylated, but not unmethylated, oligonucleotides (Figure S2A).
Furthermore, competition experiments showed that GST-SUVH6 had a higher affinity for methylated CNG and CNN substrates than methylated CG, a preference very similar to that of KYP (Figure S2B).

Figure 2. The SRA Domain
(A) ClustalW alignment of the SRA domains from the nine SUVH genes, two Arabidopsis SRA- and RING-domain-containing proteins, mouse np95, and human ICBP90. The site of the KYP S200F [4] mutation is indicated with #, the E208K mutation is indicated with a *, and the R260H [4] mutation is indicated with a +. Arrows show the SRA domain identified by SMART, and the boxed amino acids are those included in the GST-SUVH6-SRA fusion protein. The association of SRA domains with SET domain proteins is unique to the plant kingdom.
(B) Immunofluorescence of H3K9me2 compared to DAPI in nuclei isolated from wild-type, kyp-2 (null allele), and kyp-5 (SRA mutation E208K).
(C) Regions of the SUVH genes contained in the various constructs analyzed in Figure 3.

Figure 2. The SRA Domain
(A) ClustalW alignment of the SRA domains from the nine SUVH genes, two Arabidopsis SRA- and RING-domain-containing proteins, mouse np95, and human ICBP90. The site of the KYP S200F [4] mutation is indicated with #, the E208K mutation is indicated with a *, and the R260H [4] mutation is indicated with a +. Arrows show the SRA domain identified by SMART, and the boxed amino acids are those included in the GST-SUVH6-SRA fusion protein. The association of SRA domains with SET domain proteins is unique to the plant kingdom.

Furthermore, competition experiments showed that GST-SUVH6 had a higher affinity for methylated CNG and CNN substrates than methylated CG, a preference very similar to that of KYP (Figure S2B).

In order to engineer a fusion protein with just the SRA domain (Figure 2C), we first tried cloning the SRA domain predicted by SMART [24] (delineated by arrows in Figure 2A) into vectors containing a His-tag and found...
that the tagged proteins were insoluble. GST or MBP (maltose binding protein) fusions were soluble, but inactive in binding DNA. We reanalyzed the SRA domain by aligning the nine different SUVH SRAs and two SRA-RINGs (see below) to determine whether any additional conserved regions outside the annotated SRA domain existed that could be important for its function. As shown in Figure 2A, a short stretch of amino acids upstream of the annotated SRA domain was found to be conserved. When this extended SRA domain from SUVH6 was fused to GST (SUVH6-SRA), it was completely soluble and highly active in binding methylated DNA. Importantly, like full-length SUVH6, SUVH6-SRA preferred methylated CNG and CNN over CG substrates (Figure 3F). Such a preference was confirmed by competition experiments with methylated CNG substrates and methylated CG or CNG oligonucleotides as competitors (Figure 3G).

SRA Domains Associated with Arabidopsis Ring Proteins Preferentially Bind Methylated CG

In addition to the histone methyltransferases, eight other Arabidopsis proteins contain the SRA domain. Two proteins consist of a single SRA domain, whereas a class of six proteins homologous to Np95 contains the SRA as well as PHD and RING domains. To determine whether the SRA domain found in the SRA-RING proteins also binds methylated DNA, we examined two Arabidopsis SRA-RING genes, which we named ORTH1 (At5g39550) and ORTH2 (At1g57820) (Figure 2A). Three GST fusion constructs were made from ORTH1: PNS contains the PHD domain, the N-terminal RING, and the SRA domain; PN contains the PHD and N-terminal RING domains; and SC contains the SRA and C-terminal RING domains (Figure 4A). In addition, a GST-tagged, full-length ORTH2 was made. Each of the proteins containing an SRA domain (PNS, SC, ORTH2)
was able to bind methylated CG, CNG, or CNN substrates, but significantly less binding was observed to the unmethylated substrates (Figure 4B and Figures S3A and S3B). Interestingly, unlike the KYP and SUVH6 SRA domains that prefer CNG and CNN methylation, ORTH1 and ORTH2 show preferential binding to methylated CG substrates (Figures 4C and 4D, Figures S3C and S3D). Two different mutations introduced into the SRA domain in the PNS construct, S292F and R362H (corresponding to amino acids 200 and 260 based on the KYP sequence [Figure 2A]), resulted in loss of DNA binding (Figures S4A and S4B). These results implicate SRA-RING proteins in functions related to DNA methylation and further support the generality of the function of the SRA as a methyl-cytosine-binding domain. The SRA domain differs from the methyl-CpG-binding domain (MBD) found in several mammalian proteins such as MeCP2, because the MBD binds only methylated CG [25], but the SRA domain can recognize methylated DNA in any sequence context. Further, our data suggest that different SRA domains display preference for methylated cytosines in different sequence contexts.

Conclusions
Our findings extend our understanding of DNA and histone methylation pathways in Arabidopsis thaliana. The observation that the SRA domain can bind directly to methylated DNA suggests that regions rich in either CG or non-CG DNA methylation may attract KYP, leading to H3K9 methylation. Alternatively, KYP could be targeted to heterochromatin by other mechanisms, and the binding of methylated DNA to the SRA domain may be somehow required for KYP activity. KYP-dependent histone methylation would then provide a binding site for CMT3 leading to CNG methylation, which would promote further action by KYP. This model suggests a self-reinforcing feedback loop for the maintenance of DNA and histone methylation and may help to explain the remarkable stability of epigenetic silent states.

The potential targeting of KYP to methylated DNA is similar to the interaction of the histone methyltransferase SETDB1 with MBD1, which links DNA methylation to histone methylation in mammalian chromatin [26]. Thus, although the mechanisms are clearly different, the general phenomenon of DNA methylation targeting histone methylation appears to be conserved between plants and animals.

Supplemental Data
Supplemental Data include four figures, one table, and Experimental Procedures and can be found with this article online at http://www.current-biology.com/cgi/content/full/17/4/379/DC1/.

Acknowledgments
We thank S. Pradhan for oligonucleotides, D. Zilberman for sequencing kyp-5, and A.M. Palanca and W. Turki-Judeh for technical assistance. M.B. was funded by American Heart Association Postdoctoral Fellowship 0625014Y. E.K. was supported by NIH training grant GM0007377-27. Funded by NIH Grant GM60398 to S.E.J and NSF-2010 MCB-0519970 to J.C. S.E.J. is an investigator of the Howard Hughes Medical Institute.

Received: November 20, 2006
Revised: December 28, 2006
Accepted: January 3, 2007
Published online: January 18, 2007

References


