Chromatin-based silencing mechanisms
Judith Bender

Eukaryotic genomes are organized into regions of transcriptionally active euchromatin and transcriptionally inactive heterochromatin. In plant genomes, heterochromatin is marked by methylation of cytosine and methylation of histone H3 at lysine 9. Heterochromatin formation is targeted to transposons as a means of defending the host genome against the deleterious effects of these sequences. Heterochromatin is directed to transposon sequences by transposon-derived aberrant RNA species and functions to prevent unwanted transcription and movement. Formation of heterochromatin at rRNA-encoding genes and centromere-associated repeats might also involve an RNA-based mechanism that is designed to stabilize these potentially labile structures.

Addresses
Department of Biochemistry and Molecular Biology, John Hopkins University, Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, Maryland 21205, USA
e-mail: jbender@mail.jhmi.edu

Abbreviations
CMT chromomethylase
DMTase DNA methyltransferase
dsRNA double-stranded RNA
HMTase histone methyltransferase
HP1 heterochromatin protein 1
RdDM RNA-directed DNA methylation
RdRP RNA-dependent RNA polymerase
RNAi RNA interference
siRNA small interfering RNA

Introduction
In eukaryotes, DNA is packaged into the nucleus of the cell by assembly into chromatin. The fundamental unit of chromatin is the nucleosome, consisting of an octamer with two subunits each of histones H2A, H2B, H3 and H4, which provides a globular core that is wrapped by roughly 150 bp of DNA. Nucleosomal DNA is further packaged by higher-order folding and association with other proteins. Along each chromosome, chromatin is organized into transcriptionally active, less-condensed euchromatin, and transcriptionally inactive, highly-condensed heterochromatin.

In some eukaryotes, including plants and mammals, DNA in regions of heterochromatin is marked by methylation at the 5-position of cytosines. This covalent DNA modification can be inherited on the template strand during DNA replication to guide the maintenance of the mark. In both plants and animals, heterochromatin is also marked by histones carrying particular posttranslational modifications on their amino-terminal tails, which extend outwards from the core of the nucleosome. These modifications include methylation of histone H3 at lysine 9 (H3 mK9), and a lack of acetylation on histones H3 and H4. By contrast, euchromatin is marked by methylation of H3 at lysine 4 (H3 mK4) and acetylation of histones H3 and H4. These combinations of histone modifications have been proposed to constitute a ‘histone code’ that guides chromatin interactions with appropriate chromatin remodeling factors (reviewed in [1,2]). In this review, I discuss recent results that elucidate how heterochromatin is targeted to appropriate regions of the plant genome, and how interplay between DNA methylation and histone modification acts to maintain the heterochromatin state.

RNA-directed DNA methylation
Some, if not all, heterochromatin formation in plant genomes is directed by aberrant RNA species. This RNA-based heterochromatin formation is interrelated with RNA interference (RNAi), another RNA-based silencing mechanism (reviewed in [3]). In RNAi, double-stranded RNA (dsRNA) provides a trigger for the degradation of transcripts that have shared sequence identity. The dsRNA trigger is first diced into small interfering RNAs (siRNAs) of 21–26 nt by Dicer ribonucleases. At least some species of these siRNAs are subsequently taken up by an RNA-induced silencing complex and used to guide the degradation of target transcripts by sequence complementarity. The observation from plant systems that produce high levels of dsRNA, such as RNA virus infections [4,5] or highly transcribed inverted repeat transgenes [6,7,8], is that dsRNA and/or siRNAs can also trigger the methylation of DNA with identical sequences. This RNA-directed DNA methylation (RdDM) is highly sequence-specific, methylation does not significantly spread beyond the boundary of the RNA trigger [9,10].

In plant genomes, DNA methylation is maintained preferentially at cytosines in the symmetric dinucleotide context 5′-CG-3′ (hereafter referred to as CG) but
methylation can also occur in CNG (where N is any nucleotide) and asymmetric contexts. By contrast, in mammalian genomes, methylation is maintained almost exclusively in CG contexts. RdDM typically involves dense methylation of both CG and non-CG cytosines in the target regions. Forward and reverse genetic approaches in Arabidopsis thaliana have implicated two plant-specific types of DNA methyltransferase (DMTase) in non-CG methylation: DRM DMTases are primarily involved in initiating new RdDM imprints and can methylate asymmetric cytosines \[11,12,13^*,14\], whereas the CMT3 DMTase maintains the RdDM imprint preferentially at CNG cytosines \([12,14–16]; \text{Figure 1}\). A third type of DMTase, MET1, is related to the mammalian DMTase Dnmt1 and maintains methylation in CG contexts\[17,18^*\]. Thus, the relative contribution of each type of DMTase to the overall methylation status of a locus depends on the sequence composition of the locus.

The CMT ‘chromomethylase’ class carries a ‘chromo-domain’ interaction motif embedded in its methyltransferase catalytic domain. As the chromo-domains found in the animal heterochromatin protein 1 (HP1) and Polycomb protein bind to H3 mK9 and H3 mK27, respectively \([19–21]\), it is attractive to speculate that the CMT chromo-domain is similarly involved in interactions with methylated histones. In fact, the genetic screens in Arabidopsis that identified cmt3 mutations also identified mutations in the KYP/SUHV4 H3 K9 histone methyltransferase (HMTase) that confer deficiencies in non-CG methylation, indicating a direct link between the H3 mK9 mark and CMT3-mediated DNA methylation \([22^*,23^*]\).

In the fungus Neurospora crassa, DNA methylation is maintained by the DMTase DIM-2, which lacks a chromo-domain \[24\], but also requires a separate

---

**Figure 1**

Establishment and maintenance of heterochromatin. The box represents a duplicated transposon sequence, which is either transcribed (green arrows) or silenced (red crosses), depending on whether DNA methylation occurs on promoter sequences. The histone H3 methylation status corresponding to transcriptional activity is indicated by K4 for methylation on Lys4 (active) or by K9 for methylation on Lys9 (silenced). Question marks indicate unknown histone-modification state for densely methylated but transcribed sequence. DNA methylation is indicated by filled circles, with methylated CG cytosines in black and methylated non-CG cytosines in red. The unbroken and broken arrows for RdDM indicate, respectively, strong and weak interactions of the RNA signal. CMT, DRM and MET are different classes of DNA methyltransferases.
chromo-domain-containing HP1-related protein [25] and H3 mK9 [26]. These observations suggest that the DMTase/chromo-domain/H3 mK9 relationship is conserved among eukaryotes, with variations in whether the functional protein domains are located in separate polypeptide chains. At present, however, there is no evidence that the CG methylation maintained by MET1 in plants is directed by patterns of histone methylation. In addition, the Arabidopsis HP1 homolog LHP1 is not needed to maintain DNA methylation [27*]. Instead, the histone deacetylase HDA6 is needed to reinforce CG methylation patterning in Arabidopsis [27,28,29*].

Transposon defense via RdDM

Transposon sequences present a threat to the integrity of the host plant genome because of their movement to new insertion sites. Even transposon-derived duplications that are no longer capable of movement can potentially promote unwanted rearrangements by recombination-based mechanisms. A primary line of host defense against these dangerous sequences is to target them for heterochromatin formation, which acts to suppress their transcription, movement and recombinational activity. Heterochromatin targeting involves RdDM, as indicated by patterns of CG and non-CG methylation, and transposon-derived siRNAs [28,30,31*,32*].

Analysis of the mobile CACTA transposons in Arabidopsis shows that CG and non-CG methylation patterning function as redundant reinforcements in transposon suppression. When CACTA elements are partially demethylated by mutation of the MET1 CG DMTase or mutation of the CMT3 non-CG DMTase, they show partial transcriptional reactivation but no movement [31**]. Only when the elements are completely demethylated by a double met1 cmt3 mutation can they transpose. Thus, the different pathways of methylation patterning in plants provide different layers of transposon defense.

A puzzling aspect of transposon methylation is how transposons, but not normal host genes, generate the dsRNA triggers for RdDM. Part of the answer to this puzzle comes from genetic studies in Arabidopsis showing that an RNA-dependent RNA polymerase (RdRP) and other putative RNA-processing factors are needed for both RdDM and RNAi triggered by aberrant RNA species that are not themselves double-stranded [7*,33–37]. These results suggest that some single-stranded RNAs act as templates for RdRP-mediated synthesis of a second strand to form dsRNA. Furthermore, when a particular template RNA is converted into dsRNA and diced siRNAs, the siRNAs then have the potential to amplify the pathway either by directly priming RdRP-mediated synthesis or by annealing to and causing structural changes in the template RNA that indirectly facilitate RdRP-mediated synthesis (5**,38; Figure 1).

What features, then, would distinguish transposon-derived transcripts as preferred templates for RdRP? A key aspect of transposons (or transgenes) is that they integrate at random sites in the host genome and, thus, can be potentially transcribed from fortuitous nearby promoters. Such ‘read-through’ transcripts are likely to have disrupted splice sites and translational open reading frames across the junction of host and transposon sequences, which might create cues that shunt them into the RdRP pathway. Another possibility is that independent sense and antisense strand read-through transcripts have the potential to pair with each other to form transposon dsRNA. Read-through RNA triggers of RdDM that encompass transposon sequences from end to end can impair movement at two levels: by silencing internal transposon promoters and by blocking interactions of transposon-encoded proteins with transposon sequences. RNAi provides an additional line of defense against externally driven, transposon-derived transcripts that might otherwise yield protein products that promote transposon movement.

What level of dsRNA would be required to maintain transposon methylation, especially if the transposon sequence were present around the genome in multiple copies, of which only one or a few were the source of dsRNA? Studies on the methylated endogenous PAI genes in Arabidopsis, which are arranged as an inverted repeat and transcribed from a fortuitous external promoter, suggest that extremely low levels of dsRNA can be potent RdDM signals. Only a few PAI transcripts extend beyond a polyadenylation site at the center of the PAI inverted repeat to make dsRNA, and diced siRNAs are below the limit of detection by RNA gel blot [8]. Nonetheless, the inverted repeat promotes dense CG plus non-CG methylation of itself plus two unlinked duplicated PAI sequences through RdDM. Moreover, a rearrangement in the center of the inverted repeat that further suppresses dsRNA levels causes reduced but still very stable methylation of PAI sequences [39]. The PAI genes thus illustrate two key points relevant to transposon defense: even very low levels of aberrant RNAs can efficiently promote RdDM, and a single ‘bad’ member of a family can potentially generate enough RNA to suppress the whole family in trans.

Transposon duplications might consist of two classes: one in which transposon sequences are transcribed from external promoters to make RdDM triggers, and another in which transposon sequences are transcriptionally silent targets (Figure 1). This idea fits in with another observation regarding non-CG methylation, that is, that the KYP/ SUVH4 H3 K9 HMTase, which makes the H3 mK9 mark diagnostic of transcriptional silencing, is needed only to maintain non-CG methylation at a subset of methylation target sequences. For example, the transcribed PAI inverted repeat is not demethylated in a kyp/suvh4 mutant
background [23]. Only the unlinked transcriptionally silent PAI genes are affected in this background. In contrast, all three PAI loci are demethylated in a cmt3 mutant background. These results suggest that, at transcriptionally active loci that produce RNA triggers of RdDM, CMT3 can respond to KYP/SUVH4-independent cues, such as H3 mK9 catalyzed by another HMTase, a different histone modification, and/or direct interactions with accumulated aberrant RNAs. Conversely, at transcriptionally silenced target loci where aberrant RNA interactions might occur more rarely, CMT3 is strongly dependent on the KYP/SUHV4-catalyzed H3 mK9 mark.

**rRNA-encoding and centromere-associated repeats**

Like transposon repeats, the rRNA-encoding gene (rRNA gene) repeats in plant genomes carry both cytosine methylation and the H3 mK9 mark, and yet these sequences produce high levels of RNA transcripts. An explanation for this apparent paradox comes from recent Arabidopsis studies [40,41] of the rRNA genes in nucleolus-organizer regions, which are transcribed by RNA polymerase I, and the 5S rRNA genes, which are transcribed by RNA polymerase III. These studies show that rRNA repeats consist of two populations: one that is methylated and enriched for the H3 mK9 mark, and one that is unmethylated and enriched for the H3 mK4 mark.

For the nucleolus-organizer-region repeats, which lie in subtelomeric regions, chromatin immunoprecipitation was used to show that the fraction enriched in H3 mK9 contains cytosine-methylated rRNA genes, whereas the fraction enriched in H3 mK4 contains unmethylated rRNA genes [40]. For the 5S rRNA repeats, which lie in pericentromeric regions, immunocytology was used to show that a presumably transcriptionally active population of repeats lacking H3 mK9 emerges as a loop away from the heterochromatic population [41]. Thus, although methylation of rRNA genes might be an inevitable consequence of producing high levels of RNA, it could be that these sequences can assemble active domains of repeats that are resistant to methylation by virtue of a higher-order organization within the nucleus. In addition, those repeats that are assembled into heterochromatin might represent a storage form of rRNA function [46,47]. Because of the complexity of centromere repeats, it remains to be determined whether they are organized into subdomains, in which functional regions are interspersed with inactive regions. siRNAs corresponding to centromere sequences have been detected in Arabidopsis [32], arguing that at least a subset of the centromere repeats are transcribed to make signals for RdDM. In addition, centromere repeat sequences in a kyp/sucb4 mutant background are only partially demethylated at CNG residues in comparison with those in a cmt3 mutant background [22,23], suggesting that only a subset of the repeats are transcriptionally silent and carry the H3 mK9 mark of transcriptional silencing.

Condensed centromeric heterochromatin could be required for correct chromosome segregation and could also stabilize centromere repeats against recombination; however, methylation-deficient mutant plants with strong demethylation and decondensation of centromere sequences are viable [43,45]. Thus, either the tight compaction of centromere repeats is not a requirement for function or deleterious effects of decondensed centromeres are not evident for plants grown under laboratory conditions for only a few generations.

**Conclusions**

Although it is clear that dsRNA provides a sequence-specific signal for heterochromatin formation in plants, we are only beginning to understand the underlying mechanisms. Key issues include defining the features that make some transcripts preferred substrates for processing into dsRNA, elucidating how dsRNA-derived species align with cognate DNA sequences, and identifying the components of the RNA–DNA complex that recruit DNA- and histone-modifying enzymes. Heterochromatin directed at transposon sequences prevents the deleterious effects of unchecked transcription and movement; however, the functions of heterochromatin formation at rRNA gene repeats and centromere repeats remain unclear. Mutations that disrupt heterochromatin will provide key tools with which to dissect these functions.

**Acknowledgements**

J Bender’s research on gene silencing is supported by National Institute of Health (NIH) grant GM-61148.

**References and recommended reading**

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

GFP experiments show that spreading is dependent on both transcription of the target gene and a putative RNA-dependent RNA methylation. Plant Cell 1999, 11:2291-2301.


The authors use RNA viruses to deliver dsRNA that corresponded to specific sequences into the plant. Notably, when the viral dsRNA trigger corresponds only to 5’-GF’ sequences of the green fluorescent protein gene (GFP), which is expressed in full length from a transgene in the tobacco or Arabidopsis genomes, GFP siRNAs and GFP DNA methylation spread from the trigger into the ‘P’ segment of the target. Additional experiments show that spreading is dependent on both transcription of the resident GFP transgene and the sde1/sgs2 RdRP. These findings argue that siRNAs located anywhere along a target transcript can potentially activate the transcript as a template for RdRP-catalyzed synthesis of dsRNA. Two endogenous tobacco transcripts are shown to be resistant to RNAi spreading, however, indicating that not all transcripts can be similarly activated.


Several mutations, including the sde1/sgs2 RdRP mutation, have been isolated as suppressors of RNAi triggered by a direct repeat sense strand transgene. These mutations are tested for their effects on RNAi triggered by an inverted repeat transgene for the same target sequence. The mutations do not impair either RNAi or RdDM triggered by the inverted repeat. The results argue that the tested mutations affect a pathway that converts single-stranded RNA templates into dsRNA; this pathway is bypassed when dsRNA is produced directly by transcription of an inverted repeat. These findings agree with previous results [33] showing that block/sde1/sgs2 mutation does not affect RNAi triggered by viral infection, which provides a direct source of dsRNA through viral replication.


The DRM DMTases were originally identified on the basis of their sequence homology to mammalian Dnmt3 DMTases, which have been implicated in de novo methylation. In this study, a double insertion mutant in two potentially redundant Arabidopsis DRM genes is explicitly generated and shown to be defective for de novo methylation of a newly integrated transgene that is efficiently methylated in a wild-type background. In [12], the same authors show that the double drm mutant is defective for maintenance of non-CG methylation at specific methylated endogenous sequences.


The authors describe two insertional mutations in the Arabidopsis CG DMTase gene MET1. Strikingly, heterozygous met1/MET1 progeny derived by self-pollination of the original heterozygous met1/MET1 insertional mutant isolates show loss of CG methylation. Because the recessive met1/MET1 mutations do not affect RNAi or RdDM triggered by the inverted repeat, the results indicate that met1-induced demethylation occurs during haploid gametogenesis. Additional experiments with a methylated transgene reporter show that when the transgene is inherited through met1 female gametogenesis, a higher proportion of progeny shows loss of CG DNA methylation than when the transgene is inherited through met1 male gametogenesis. This finding is consistent with an extra round of postmeiotic DNA replication in female gametogenesis. This work also raises the possibility that the met1 alleles might have significant strong or weak phenotypes than a missense allele that was used in several previous studies [17].


Transgene-induced methylation and silencing of the endogenous Arabidopsis SUPERMAN gene provide a reporter system for mutations that are defective in silencing. This screen yields mutations in the KYP/SUVH4 H3 K9 HMTase [22], the CMT3 DMTase [15], and the AGO4 RNA-processing factor [37] that affects the control of non-CG methylation, particularly in the CNG context that is enriched in SUPERMAN 5’ sequences. This paper provides the first link in a plant system between the H3K9 modification and DNA methylation. In vitro evidence suggests that the LHP1 chromodomain protein functions as a bridge between H3 methylases and DNA methylation. This paper also raises the possibility that the met1 alleles might have significant strong or weak phenotypes than a missense allele that was used in several previous studies [17].


A propagated and silenced Arabidopsis PAl gene is used as a reporter to isolate silencing-defective mutations. Like the SUPERMAN screen [22], this screen yields mutations in the KYP/SUVH4 H3 K9 HMTase [23] and the CMT3 DMTase [16]. Although mutations in both genes reduce non-CG methylation on the silenced PAl gene reporter, they have different effects on the transcribed PAl inverted repeat that triggers methylation on PAl: the inverted repeat is demethylated by cmt3 but not by kup/suvh4 mutations. Furthermore, a new PAl methylation system in plant cannot be established in a cmt3 mutant background but can be established in a kup/suvh4 mutant background. Together, these results argue that only a subset of CMT3-mediated methylation relies on the KYP/SUVH4-mediated H3 K9 mark. Notably, in contrast to the SUPERMAN screen, the PAl screen does not yield apo3 mutations; this difference might reflect the need for processing a dsRNA signal in the SUPERMAN system that differs from the direct production of dsRNA in the PAl system (see [8]).


27. Aufsatz W, Mette MF, Van Der Winden J, Matzke M, Matzke AJ: Two classes of Arabidopsis histone deacetylase HDA6 are required for maintenance of transcriptional gene silencing and determines nuclear organization of DNA repeats. Plant Cell 2004, 16:1021-1034. Mutations in the Arabidopsis histone deacetylase gene HDA6 have been previously isolated in screens for the reactivation of silenced methylated transgenes (e.g. [27]). In this paper, hda6 alleles isolated from two other screens that were based on transgene silencing are tested for their effects on endogenous repetitive sequences, including the nuclear rRNA genes. This analysis shows that the hda6 mutations cause loss of CG methylation, enrichment for acetylated histone H4, and enrichment for H3 methylated on H3K4 on rRNA sequences. These mutations have little or no effect on DNA methylation at other sequences, such as the centromere repeats, or on bulk levels of acetylated H4 and H3, suggesting that HDA6 might have a specific function in rRNA chromatin organization; however, the hda6-induced chromatin changes are not sufficient to increase rRNA expression. The authors speculate that HDA6 might work together with HDT1, another histone deacetylase implicated in rRNA regulation (see [40]).


29. Probst AV, Fagard M, Proux F, Mournant P, Boulcours D, Earlley K, Lawrence RJ, Pikaard CS, Murfett J, Earley K et al.: Arabidopsis histone deacetylase HDA6 is required for maintenance of transcriptional gene silencing and determines nuclear organization of DNA repeats. Plant Cell 2004, 16:1021-1034. Mutations in the Arabidopsis histone deacetylase gene HDA6 have been previously isolated in screens for the reactivation of silenced methylated transgenes (e.g. [27]). In this paper, hda6 alleles isolated from two other screens that were based on transgene silencing are tested for their effects on endogenous repetitive sequences, including the nuclear rRNA genes. This analysis shows that the hda6 mutations cause loss of CG methylation, enrichment for acetylated histone H4, and enrichment for H3 methylated on H3K4 on rRNA sequences. These mutations have little or no effect on DNA methylation at other sequences, such as the centromere repeats, or on bulk levels of acetylated H4 and H3, suggesting that HDA6 might have a specific function in rRNA chromatin organization; however, the hda6-induced chromatin changes are not sufficient to increase rRNA expression. The authors speculate that HDA6 might work together with HDT1, another histone deacetylase implicated in rRNA regulation (see [40]).


31. Kato M, Miura A, Bender J, Jacobsen SE, Kakutani T: Role of CG and non-CG methylation in immobilization of transposons in Arabidopsis. Curr Biol 2003, 13:421-426. This study uses mutations in the Arabidopsis MET1 CG DMTrase and the CMT3 non-CG DMTrase to show that single mutations are insufficient to activate the movement of CACTA transposons, whereas the double met1 cmt3 mutant background shows marked CACTA transposition. These results thus demonstrate that the CG and non-CG methylation patterning pathways in plants provide redundant layers of transposon defense.


40. Lawrence RJ, Earlley K, Pontes O, Silva M, Chen ZJ, Neves N, Viegas W, Pikaard CS: A conserved DNA methylation/histone deacetylation switch regulates rRNA gene dosage control and nucleolar dominance. Mol Cell 2004, 13:599-609. This study shows that the hda6 mutations cause loss of CG methylation, enrichment for acetylated histone H4, and enrichment for H3 methylated on H3K4 on rRNA sequences. These mutations have little or no effect on DNA methylation at other sequences, such as the centromere repeats, or on bulk levels of acetylated H4 and H3 methylated on H3K4. A probable null hda6 mutation also confers decondensation of rRNA but not of centromeric heterochromatin. Together, these results suggest that HDA6 might have a specific function in rRNA chromatin organization; however, the hda6-induced chromatin changes are not sufficient to increase rRNA expression. The authors speculate that HDA6 might work together with HDT1, another histone deacetylase implicated in rRNA regulation (see [40]).


43. Soppe WJ, Jasencakova Z, Houben A, Kakutani T, Meister A, Huang MS, Jacobsen SE, Schubert I, Fransz PF: DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis. EMBO J 2002, 21:6549-6559. Cytological methods are used to understand the effects of methylation deficiency mutations on ‘chromocenters’ corresponding to the heterochromatic rRNA gene and centromere-associated repeat sequences. This study reveals that mutations in the MET1 CG DMTrase or in the DDM1 chromatin-remodeling protein (which affects both CG and non-CG methylation) cause loss of DNA methylation, methylation of H3 lysine 9, and condensation at the chromocenters. Another important finding is that single-copy DNA-methylated sequences are not recruited to the chromocenters and instead can be silenced in a euchromatic context.


