Cell, Vol. 122, 13–16, July 15, 2005, Copyright ©2005 by Elsevier Inc. DOI 10.1016/j.cell.2005.06.034

The Role of the RNAi Machinery in Heterochromatin Formation

Minireview

Michael Wassenegger*
RLP-AgroScience GmbH
AlPlanta-Institute for Plant Research
Breitenweg 71
67435 Neustadt
Germany

Considerable effort has been devoted to the characterization of RNA interference (RNAi), a posttranscriptional gene silencing mechanism involving small RNA-containing effector complexes. Recent studies have revealed that components of the RNAi machinery are associated not only with target RNA cleavage and impairment of target RNA translation but also with the formation of heterochromatin. There is increasing evidence that RNA-mediated chromatin modifications play an important role in epigenetic transcriptional gene silencing.

RNAi acts mainly posttranscriptionally (see minireview by Filipowicz [2005] in this issue of *Cell*), but components of the RNAi machinery can also be involved in nuclear processes leading to heterochromatin formation and transcriptional gene silencing (TGS). RNA-mediated heterochromatin formation, here referred to as nuclear RNAi, appears to be a natural epigenetic gene regulation mechanism. This mechanism is active in most eukaryotes and controls heritable changes in gene expression that are not caused by mutations. Nuclear RNAi functions as a surveillance mechanism against foreign nucleic acids, e.g., retroelements and transposons; is involved in the regulation of developmental genes; and contributes to accurate chromosome segregation during cell division. Depending on the organism, nuclear RNAi can engage specific processes, e.g., DNA methylation and/or RNA amplification; however, targeting of a homologous chromosomal region for chromatin modifications by an RNA inducer is a common theme. In this minireview, each step in the nuclear RNAi process is examined by considering examples from *Schizosaccharomyces pombe*, plants, and mammals. Unresolved issues are discussed, and a model of nuclear RNAi is presented.

Nuclear RNA Interference in Schizosaccharomyces pombe

The mechanism of RNA-mediated heterochromatin formation in *S. pombe* is relatively well understood (Grewal and Rice, 2004). Only a brief overview of the *S. pombe* system is presented here. The initial step requires bidirectional transcription of target loci to produce primary dsRNAs. Small interfering RNAs (siRNAs) are generated from these dsRNAs by the RNase III-type endonuclease Dicer. siRNAs are incorporated into the *R*NA-induced *i*nitiator of *t*ranscriptional gene *s*ilencing (RITS) complex via an argonaute family protein (Ago1), guiding RITS to complementary sites of the genome. It is unclear whether RITS binds only to nascent target transcripts or pairs to corresponding DNA. Either way, binding of the complex enables recruitment of chromatin-modifying proteins, including a histone H3 lysine 9 (H3K9) methyltransferase (Clr4). Methylation of H3K9 likely stabilizes the binding of RITS to chromatin. This dependence on the H3K9 methyltransferase may argue for RITS/DNA interactions in addition to RNA binding. Alternatively, nascent transcripts may be chromatin bound, allowing RITS to guide histone methyltransferases to proximal DNA regions. Once stabilized, RITS interacts with the RNA-directed RNA polymerase complex (RDRC), required for the production of secondary dsRNA and thus for amplification of the silencing signal (Motamedi et al., 2004). RDRC consists of the *S. pombe* RNA-directed RNA polymerase (RdRP) Rdp1, a putative helicase (Hrr), and a protein (Cid12) associated with RNA polyadenylation. The role of RNA polyadenylation in RdRP-mediated dsRNA amplification is unknown, but RdRP activity is critical for generation of secondary siRNAs and heterochromatin assembly. A self-perpetuating loop thus exists in which binding of RITS to chromatin allows RDRC to operate in *cis* to produce secondary dsRNA using nascent transcripts as templates. Importantly, current models predict that, even if a high concentration of homologous dsRNA is available from an alternative source, RdRP-mediated secondary dsRNA synthesis is mandatory. This requirement indicates that silencing and heterochromatin maintenance depend on primary transcription of target loci.

In *S. pombe*, it is unclear what determines whether siRNAs initiate RNAi or nuclear RNAi. dsRNA hairpin constructs can induce either the RNAi pathway alone (Sigova et al., 2004) or both pathways (Schramke and Allshire, 2003). Since *S. pombe* has only single copies of Dicer and Ago, the putative RNA-cleavage complex and RITS must contain the same AGO protein, and the dsRNA is processed by the same Dicer, ruling out the possibility that these proteins directly determine the subcellular localization of either complex. Thus, the factor(s) that determines whether the siRNA/AGO complex becomes a RITS or RNA-cleavage complex remain to be elucidated. The localization of siRNA precursors may be important, as suggested by Sigova et al. (2004). In this study, the dsRNA hairpin contained an intron separating two inverted repeats (IRs) that were complementary to the target RNA. Efficient splicing promoted accumulation of the dsRNA in the cytoplasm, leading to dsRNA processing and complex assembly in this cell compartment and initiating target RNA cleavage. A hairpin RNA that initiated both RNAi and nuclear RNAi did not contain an intron. It may have been present in both nucleus and cytoplasm, resulting in both RITS and RNA cleavage activities. Notably, a dependence on the presence of a RNA inducer in the nucleus was also observed for RNA-directed DNA methylation (RdDM) in plants (Mette et al., 2000) (see below).

*Correspondence: michael.wassenegger@agroscience.rlp.de

Table 1. Proteins Involved in RNA-Mediated DNA Methylation and Heterochromatin Formation

S. pombe	*A. thaliana*	Mammals	Protein Class
Dcl1	DCL3	Dicer	RNase III-type endonuclease
–	HEN1	(?)	RNA methyltransferase
Ago1	AGO4, [?]	AGO?	Small RNA binding protein[a]
–	DRM2	Dnmt3b	De novo DNA methyltransferase
–	MET1	Dnmt1	Maintenance DNA methyltransferase
–	CMT3	–	Maintenance DNA methyltransferase[b]
–	AtMBD5, AtMBD6, [?]	MBD?	Methyl binding protein
–	HDA6, [?]	HDAC?	Histone deacetylase
–	DDM1, DRD1, [?]	Lsh, [?]	Chromatin remodeling factor
Clr4	SUVH2, SUVH4, [?]	Suv39h1, Suv39h2, [?]	Histone H3 lysine 9 methyltransferase
Rdp1	RDR2	–	RNA-directed RNA polymerase
–	Pol IV	–	DNA-dependent RNA polymerase
Chp1	(?)	(?)	Chromodomain protein
Tas3	(?)	(?)	Unknown
Hrr1	(?)	(?)	Putative helicase
Cid12	(?)	(?)	RNA polyadenylating protein

?, not clear which member of the gene family is involved; [?], likely involvement of further members of the gene family; (?), involvement of orthologous proteins in nuclear RNAi not known; –, no known homologues.
[a] Component of functional si- and mi-RNPs.
[b] Plant-specific methyltransferase.

Nuclear RNAi in Plants and Mammals

Plants provide excellent systems for the study of RNAi and nuclear RNAi. Indeed, many seminal discoveries were made in plants, including posttranscriptional gene silencing (PTGS, the plant equivalent of RNAi), RdDM, RdRPs, siRNAs, and AGO proteins. However, nuclear RNAi in plants has several unique features (Table 1), such as a fourth DNA-dependent RNA polymerase (Pol IV) (Herr et al., 2005; Onodera et al., 2005) and an RNA-methylating enzyme (HEN1) (Yu et al., 2005). The relevance of these proteins to other systems is not yet clear. On the other hand, biochemical analysis of components involved in heterochromatin formation has been more substantial in mammalian systems. Since plants combine RNAi components from both *S. pombe* and mammals, plants can serve as an overall guide to RNA-mediated heterochromatin formation (Figure 1).

Initiation of Nuclear RNAi: The Origin of Primary dsRNA

Initiation of heterochromatin formation requires production of primary dsRNAs. In plants and yeast, three different scenarios can be envisioned. First, dsRNA may be produced through bidirectional transcription or transcription of IRs. Second, a target sequence may not be transcribed at all, but dsRNA would originate from an unlinked homologous transcribed locus or an exogenous source. Third, dsRNA may be produced by an RdRP on a single-stranded RNA template. Although RdRP may produce dsRNAs in either the cytoplasm or the nucleus, it appears that in both yeast and *Arabidopsis thaliana*, RdRP activity associated with heterochromatin formation is nuclear. In plants, it is likely that all three processes for dsRNA production contribute equally to natural epigenetic gene regulation. Since mammals lack RdRPs, dsRNA delivery in these organisms is based on the two former processes.

DNA Methylation as Part of Nuclear RNAi in Plants and Mammals

One important difference between RNA-mediated heterochromatin formation in *S. pombe* and in plants and mammals is the association of nuclear RNAi with de novo DNA methylation in the latter two systems. Methylation appears to act upstream of de novo heterochromatin formation: an RNA inducer targets a DNA sequence for methylation, followed by recruitment of the heterochromatin forming machinery. What is the advantage of a DNA methylation system? In *S. pombe*, heterochromatin maintenance through the cell cycle requires the permanent presence of an inducer. In contrast, DNA methylation in symmetrical CpG nucleotide groups can be maintained after genome replication even in the absence of inducer. Thus, maintenance of methylation results in inheritance of repressive epigenetic marks, helping to preserve gene repression.

RNA-Directed De Novo DNA Methylation in Plants

Establishment of de novo methylation patterns in plants requires a dsRNA inducer in the nucleus (Mette et al., 2000). The process also requires the SWI2/SNF2-like protein DRD1, at least at some loci. How the RNA guides the major de novo DMTase DRM2 to homologous DNA is unknown (Cao et al., 2003). A RdDM complex containing DRD1 may directly bind one or both strands of the primary dsRNA. The complex may then scan the genome for homologous regions, opening the helix to allow pairing of the RNA with the complementary DNA (Figure 1). The complex may then recruit DRM2, resulting in the methylation of one or both DNA strands.

It is difficult to determine whether one or both strands are initially methylated because different strands may be methylated in different cells. In addition, the maintenance methylation machinery may act on the opposite, unmethylated strand at symmetric CpG sites, the preferential substrate for the methyltransferase MET1. Either way, the result will be the (incomplete) methylation of cytosine residues at symmetrical CpG and CpHpGs (H = A, C, or T), and asymmetrical CpHpH sites.

Alternatively, siRNAs, rather than long dsRNAs (>24 nt), may bind the RdDM complex (Figure 1). Plants pro-

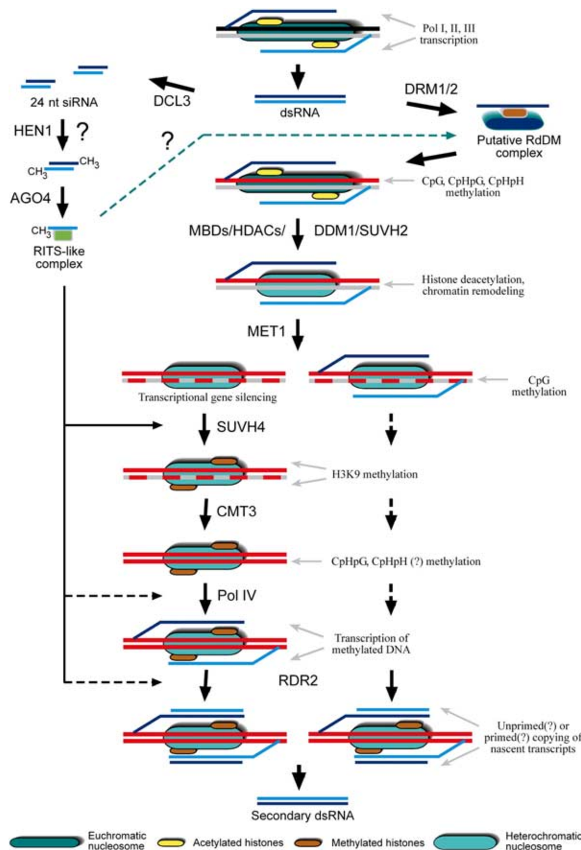

Figure 1. Mechanism for RNA-Mediated Initiation of DNA Methylation and Heterochromatin Formation in Plants

Unmethylated DNA strands are indicated by black and gray lines, respectively. Fully methylated DNA strands are in red, and CpG-methylated DNA strands are indicated by a gray/red line. RNA sense and antisense strands are dark blue and light blue lines, respectively.

duce both 21 and 24 nt short RNAs. The latter, originating from processing for the initiating dsRNA by nuclear Dicer-like (DCL3), seem to be specialized for nuclear RNAi and are thought to guide the RdDM complex. However, in *A. thaliana*, 24 nt siRNAs homologous to unmethylated regions are observed, indicating that 24 nt siRNAs per se can not induce RdDM (Xie et al., 2004). Moreover, in one case, RdDM occurred in the absence of detectable corresponding 24 nt siRNAs. Thus, further studies are needed to conclusively determine which type of RNA binds the RdDM complex in plants. The 24 nt siRNAs may also be involved in processes leading to production of secondary dsRNA (see below). If so, 24 nt siRNAs indirectly contribute to RdDM by increasing the concentration of dsRNA.

RNA-Directed De Novo DNA Methylation in Mammals

Dicer-deficient mouse embryonic stem cells are defective in DNA methylation and histone modifications (Kanellopoulou et al., 2005), suggesting that 21 nt siRNAs are required for RdDM in mammals. Kawasaki and Taira (2004) reported on siRNA-directed de novo methylation in mammalian cells supporting the idea that an RdDM-like mechanism exists in mammals. They found that the mammalian de novo DMTase DNMT3b, most closely related to the plant de novo DMTase (DRM2), is essential for siRNA-mediated DNA methylation. In addition, significant CpHpG methylation within regions of siRNA complementarity was found, indicating that non-CpG methylation may be based on a RNA-mediated mechanism. Compared to plants, the overall frequency of methylated non-CpG sites is low in mammals, perhaps due to spatial or temporal regulation of RdDM. Like asymmetric methylation in plants (see below), non-CpG methylation cannot be maintained in mammals by the maintenance DMTase Dnmt1 (the plant MET1 homolog). Examination of differentiated cells therefore reveals primarily methylated CpG sites. Thus, mammalian RdDM may occur predominantly during early embryogenesis. Accordingly, Dnmt3a and -3b are highly expressed in embryos and downregulated in differentiated cells, and transgenic mice lacking one or both activities die at embryonic stages or shortly after birth (Okano et al., 1999).

Completion of De Novo Methylation Requires Chromatin Remodeling

Opposite strand methylation and/or establishment of the full methylation pattern appears to require chromatin modifications. The former resembles maintenance methylation after DNA replication, since replication results in hemimethylated DNA. Methyl binding proteins (MBDs) may be recruited to the methylated DNA. At least one of the *A. thaliana* MBDs, AtMBD6, forms a complex with a histone deacetylase (HDAC) (Zemach and Grafi, 2003), and a plant HDAC (HDA6) is required for the maintenance of RdDM-induced de novo CpG methylation (Aufsatz et al., 2002). Thus, the plant maintenance DMTase MET1 may act on deacetylated, hemimethylated loci to methylate the opposite strand (Figure 1). In addition to histone deacetylation, the SWI2/SNF2 nucleosomal remodeling factor DDM1 appears to be essential for maintenance of CpG methylation and is known to bind MBDs in *A. thaliana*. Thus, DDM1 is targeted to methylated DNA and serves to further maintain CpG methylation (Zemach et al., 2005). The DDM1-mediated chromatin structure likely enables direct access of the maintenance methylation machinery to (hemi)methylated DNA. A connection between DDM1 and the plant H3K9 methyltransferase, SUVH2, has been recently reported (Naumann et al., 2005), potentially coupling histone methylation with MBDs and chromatin remodeling by DDM1. In mammals, MBDs also recruit HDACs, and the maintenance DMTase Dnmt1 interacts directly with HDACs, supporting the dependence of maintenance methylation on histone deacetylation. In addition, Lsh, the mouse DDM1 homolog, appears to regulate the accessibility of chromatin to DMTases that maintain CpG methylation.

Methylation of opposite strand CpHpG sites likely depends on H3K9 methylation in plants and is directed by the 24 nt siRNAs. Guiding the 24 nt siRNAs to the correct sites may involve formation of a RITS-like complex containing an Ago protein (AGO4) (Zilberman et al., 2004). Formation of this complex requires the RNA-methylating enzyme HEN1 (Figure 1). HEN1 stabilizes microRNAs (miRNAs), another class of small noncoding RNAs (see minireview by Sontheimer and Carthew [2005] in this issue of *Cell*), by methylating the miRNA 3′ overhangs (Yu et al., 2005). However, plants carrying

a mutation that abolished miRNA biosynthesis showed normal methylation and silencing of transposons, arguing against absolute requirement for miRNAs in nuclear RNAi (Lippman et al., 2003). Although Yu and coworkers failed to methylate 21 nt siRNAs in vitro, it remains possible that HEN1 associates with cofactors to methylate 24 nt siRNA in vivo. These stabilized 24 nt RNAs may bind AGO4, and the putative siRNA/AGO4 complex may then target the plant H3K9 methyltransferase, SUVH4, to complementary DNA. Alternatively, the complex may bind to nascent transcripts instead of DNA. H3K9 methylation-induced chromatin modifications would make the DNA readily accessible to the maintenance DMTase CMT3 that is unique to plants. Thus, de novo CpHpG methylation by DRM2 would be copied to the opposite strand, resulting in the final methylation pattern.

Secondary dsRNA Production

As in *S. pombe*, the primary dsRNA can trigger its own amplification in plants, a process essential for reestablishment of CpHpH- and SUVH4-mediated H3K9 methylation. Secondary dsRNA synthesis appears to depend on the coordinated action of Pol IV and RDR2. What are the substrates of these enzymes? Pol IV is proposed to transcribe methylated DNA (Herr et al., 2005; Onodera et al., 2005), though it is unlikely to transcribe all methylated sequences. If it did, dsRNAs or siRNAs corresponding to the bulk of the genome would be detectable. More likely, Pol IV is guided to (hemi)-methylated DNA by the 24 nt siRNAs (Figure 1). Alternatively, it can not be excluded that Pol IV uses nascent transcripts as substrates. The existence of two functionally diversified Pol IV complexes (Kanno et al., 2005) suggests multiple roles for these enzymes in RdDM. Since not all sequences targeted for de novo methylation are transcribed, the simplest model is that Pol IV acts on methylated DNA rather than nascent transcripts, and that transcription proceeds only along the methylated DNA, preventing RNA synthesis from adjacent unmethylated regions.

Transcripts synthesized by Pol IV may serve as templates for RDR2, generating the second RNA stand. If RDR2 transcription is primed by siRNAs, only sequences corresponding to the original dsRNA trigger would be transcribed. Such a mechanism would allow secondary dsRNA production to become independent from primary transcription by RNA polymerases I, II, and III and from the presence of exogenous dsRNA, creating a self-perpetuating RNAi loop (Figure 1). After DNA replication, the secondary dsRNA would ensure maintenance of CpHpG methylation and RdDM. Both processes would help to establish the dense methylation patterns required for RNA-mediated heterochromatin formation. The absence of RdRP-mediated dsRNA production in mammals has two major implications. Maintenance of H3K9 methylation must involve an RdRP-independent mechanism, and, unlike plants, the capacity for preserving RdDM activity may be localized or spatially and temporally restricted to cells that produce the primary dsRNA.

Conclusions

S. pombe, plants, and mammals utilize small RNAs to establish de novo DNA methylation patterns and/or to maintain epigenetic marks. Recent studies indicate that RNAi/RNAi components are also involved in DNA elimination in *Tetrahymena thermophila* and in heterochromatin silencing in *Drosophila melanogaster*. We therefore appear to have only touched the tip of the iceberg. At present, it is difficult to develop models that account for all of the observations. For example, in plants, the initiation and/or maintenance of de novo methylation and chromatin modification can differ for individual endogenous and transgenic loci (Xie et al., 2004). The same may apply for RNA-mediated heterochromatin formation at different loci in mammals. siRNA-mediated mechanisms probably require unique components not involved in genomic imprinting and X chromosome inactivation. However, RNA-mediated epigenetic gene regulation is an exciting area that is sure to keep us busy and interested in the future.

Selected Reading

Aufsatz, W., Mette, M.F., van der Winden, J., Matzke, M., and Matzke, A.J.M. (2002). EMBO J. *21*, 6832–6841.

Cao, X., Aufsatz, W., Zilberman, D., Mette, F.M., Huang, M.S., Matzke, M., and Jacobsen, S.E. (2003). Curr. Biol. *13*, 2212–2217.

Filipowicz, W. (2005). Cell *122*, this issue, 17–20.

Grewal, S.I.S., and Rice, J.C. (2004). Curr. Opin. Cell Biol. *16*, 230–238.

Herr, A.J., Jensen, M.B., Dalmay, T., and Baulcombe, D.C. (2005). Science *308*, 118–120.

Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Genes Dev. *19*, 489–501.

Kanno, T., Huettel, B., Mette, M.F., Aufsatz, W., Jaligot, E., Daxinger, L., Kreil, D.P., Matzke, M., and Matzke, A.J.M. (2005). Nat. Genet. Published online May 29, 2005. 10.1038/ng1580

Kawasaki, H., and Taira, K. (2004). Nature *431*, 211–217.

Lippman, Z., May, B., Yordan, C., Singer, T., and Martienssen, R. (2003). PLoS Biol. *1*, e7. 10.1371/journal.pbio.0000067

Mette, M.F., Aufsatz, W., van der Winden, J., Matzke, M.A., and Matzke, A.J.M. (2000). EMBO J. *19*, 5194–5201.

Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Cell *119*, 789–802.

Naumann, K., Fischer, A., Hofmann, I., Krauss, V., Phalke, S., Irmler, K., Hause, G., Aurich, A.C., Dorn, R., Jenuwein, T., and Reuter, G. (2005). EMBO J. *24*, 1418–1429.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). Cell *99*, 247–257.

Onodera, Y., Haag, J.R., Ream, T., Costa Nunes, P., Pontes, O., and Pikaard, C.S. (2005). Cell *120*, 613–622.

Schramke, V., and Allshire, R. (2003). Science *301*, 1069–1074.

Sigova, A., Rhind, N., and Zamore, P.D. (2004). Genes Dev. *18*, 2359–2367.

Sontheimer, E.J., and Carthew, R.W. (2005). Cell *122*, 9–12.

Xie, Z., Johansen, L.K., Gustafson, A.M., Kasschau, K.D., Lellis, A.D., Zilberman, D., Jacobsen, S.E., and Carrington, J.C. (2004). PLoS Biol. *2*, 642–652.

Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R.W., Steward, R., and Chen, X. (2005). Science *307*, 932–935.

Zemach, A., and Grafi, G. (2003). Plant J. *34*, 565–572.

Zemach, A., Li, Y., Wayburn, B., Ben-Meir, H., Kiss, V., Avivi, Y., Kalchenko, V., Jacobsen, S.E., and Grafi, G. (2005). Plant Cell *17*, 1549–1558.

Zilberman, D., Cao, X., Johansen, L.K., Xie, Z., Carrington, J.C., and Jacobsen, S.E. (2004). Curr. Biol. *14*, 1214–1220.