Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation

Tatsuo Kanno¹, M. Florian Mette², David P. Kreil³, Werner Aufsatz¹, Marjori Matzke¹ and Antonius J.M. Matzke¹

¹Gregor Mendel Institute of Molecular Plant Biology
Austrian Academy of Sciences
UZAII, Althanstrasse 14
A-1090 Vienna, Austria

²Institute of Plant Genetics and Crop Plant Research
Corrensstrasse 3
D-06466 Gatersleben, Germany

³Dept. of Genetics / Inference Group
University of Cambridge
Cambridge, CB2 3EH
UK

Correspondence should be addressed to M.M.
Tel: +43-662-63961-26
Fax: +43-662-63961-53
Email: marjori.matzke@gmi.oeaw.ac.at
Summary (201 words)

In plants, RNA can induce de novo cytosine methylation of homologous DNA by a mechanism that is poorly understood. Cytosines in all sequence contexts become modified in response to RNA signals \cite{1, 2}. Recent work has implicated the de novo DNA methyltransferases (DMTases), DRM1 DRM2, in establishing RNA-directed methylation of the constitutive nopaline synthase promoter \cite{3} and the DMTase MET1 \cite{4} and the putative histone deacetylase HDA6 \cite{5} in maintaining or enhancing CpG methylation induced by RNA. Despite the identification of enzymes that catalyze epigenetic modifications in response to RNA signals, it is unclear how RNA targets DNA for methylation. A screen for mutants defective in RNA-directed DNA methylation identified a novel putative chromatin remodeling protein, DRD1. This protein belongs to a previously undefined, plant-specific subfamily of SWI2/SNF2-like proteins most similar to the RAD54/ATRX-subfamily. In \textit{drd1} mutants, RNA-induced non-CpG methylation is almost eliminated at a target promoter, resulting in reactivation, while methylation of centromeric and rDNA repeats is unaffected. Thus, unlike the SNF2-like proteins DDM1/Lsh1 \cite{6, 7} and ATRX \cite{8, 9}, which regulate methylation of repetitive sequences, DRD1 is not a global regulator of cytosine methylation. DRD1 is the first SNF2-like protein implicated in an RNA-guided epigenetic modification of the genome.

RNA-directed DNA methylation requires a double stranded RNA that is processed by a Dicer-like activity into short RNAs 21-26 nucleotides in length \cite{10-12}. Double stranded RNAs containing promoter sequences can induce methylation of homologous promoter regions and transcriptional gene silencing \cite{4, 5, 10-14}. To study RNA-directed DNA methylation of a tissue-specific promoter, we established a two-component transgene system based on the seed-specific α’ promoter in \textit{Arabidopsis} (Fig. 1A, Supplementary information). In this system, silencing (Fig. 1B,C) and methylation (Fig. 2A) of an α’GFP reporter gene are triggered by an α’ promoter hairpin RNA that is transcribed from an inverted DNA repeat by an unrelated constitutive promoter. To identify proteins of the silencing machinery, seeds of a homozygous silenced α’GFP line (termed DT7-3) were mutagenized using EMS. The
treated seeds were germinated and the resulting M1 plants were allowed to self fertilize, producing M2 seeds. Silencing-defective mutants were detected by screening M2 seeds for green fluorescence (Fig. 1B). From pools of M2 seeds harvested from ~ 50,000 M1 plants, 59 putative \textit{drd} mutants (defective in RNA-directed DNA methylation) were recovered and placed into three complementation groups. These mutants are recessive as indicated by resiliencing of the \textit{\textalpha'GFP} target gene in seeds obtained by backcrossing to wild type DT7-3 plants (Fig. 1C). We report here on \textit{drd1}.

Northern blot analysis demonstrated that the synthesis of \textit{\textalpha'}promoter double stranded RNA (not shown) and processing to \textit{\textalpha'} promoter short RNAs occurred normally in the \textit{drd1} mutant (Fig. 2B). Hence, release of silencing was not due to impaired production of RNA signals. DNA methylation of the target \textit{\textalpha'} promoter was examined in the \textit{drd1} mutant by using methylation-sensitive restriction enzymes and bisulfite sequencing. Non-CpG methylation (analyzed by the restriction enzymes abbreviated F, S, Ps, Pa, E and B) was absent in leaf DNA isolated from the \textit{drd1} mutant and only trace amounts were detected in seed DNA. By contrast, CpG methylation (analyzed by the enzyme abbreviated H) was largely unaffected in leaves and seeds of \textit{drd1} plants, as demonstrated by the distinctive double band that is also observed in wild type DT7-3 plants (Fig. 2A). The results obtained from the Southern blot analysis were confirmed by bisulfite sequencing, which revealed a dramatic decrease in CpNpG and CpNpN methylation in the \textit{drd1} mutant accompanied by approximately wild type DT7-3 levels of CpG methylation (Fig. 3 and Suppl. Fig. 1). Thus non-CpG methylation induced by RNA in the \textit{\textalpha'} promoter silencing system requires DRD1. By contrast, neither CpG nor non-CpG methylation is reduced to a detectable extent in centromeric or rDNA repeats in the \textit{drd1} mutant (Fig. 4). Although it is not yet known whether RNA triggers cytosine methylation of these repetitive sequences in plants, these results suggest that DRD1 acts locally, not globally, to regulate levels of non-CpG methylation.

The \textit{drd1} mutation was mapped (Supplementary information) and found to correspond to a putative chromatin remodeling protein CHR35 (At2g16390) [15], which is a member of a previously uncharacterized subfamily of SNF2-like proteins that is unique to plants. The
The DRD1 subfamily can be defined by four ProDom [16] domains (Fig. 5). These overlap with matches to the functional signatures SNF2_N and HELICc, which together constitute the SWI/SNF ATPase domain essential for chromatin remodeling activity [17]. The \textit{drd1-1} mutation consists of a G to R change in the putative Mg\textsuperscript{2+}-binding site of SNF2_N. Five additional \textit{drd1} alleles (\textit{drd1-2}, \textit{drd1-3}, \textit{drd1-4}, \textit{drd1-5}, \textit{drd1-6}) were identified and sequenced. They all contain a mutation in strongly conserved or functionally implicated regions of the SWI/SNF ATPase domain (Fig. 5).

The DRD1 subfamily comprises six additional members, including a clear DRD1 homolog in rice (BAC84084) (Suppl. Fig. 2). CHR34 (At2g21450) is the \textit{Arabidopsis} protein most similar to DRD1, still sharing all six ProDom domains. Another rice protein (AAM15781) is highly similar to DRD1 and also contains all six domains. The remaining three members [At1g05480, T25N20.14 (Q9ZVY9, similar to CHR31), and CHR40 (At3g24340)] have only four of the six ProDom domains in common with DRD1. There are some \textit{Arabidopsis} proteins outside this well defined group that have strong local similarities to DRD1 but are likely to function differently, such as At2g25020, which is shorter and lacks the ATP-binding site, or the pair CHR38 (At3g42670) and CHR42 (At5g20420) from the proposed Clade A [15], which have only three of the four required domains and these are spread far apart.

The most closely related non-plant proteins are the \textit{Drosophila} RAD54 ortholog and an uncharacterized protein from fission yeast (P87114), which share with DRD1 only the putative Mg\textsuperscript{2+}-binding site of SNF2_N and the ATP binding site of HELICc. Although much more distant, these are also shared by the next closest SNF2 subfamily, the ATRX-like proteins. RAD54 is a chromatin remodeling factor that is specialized for homologous DNA repair [18]. ATRX localizes to pericentromeric heterochromatin and is essential for the expression of human alpha-globin genes by an unknown mechanism [8]. Mutations in human \textit{ATRX} lead to ATR-X syndrome (alpha-thalassemia/mental retardation, X-linked) and changes in CpG methylation of some repetitive sequences [9]. The \textit{Arabidopsis} genome encodes two proteins in these subfamilies [15]: a RAD54 homolog (CHR25; At3g19210) and
an ATRX-like protein (CHR20; At1g08600). Proteins of the *Arabidopsis* DRD1 subfamily are thus related to, but distinct from, presumably authentic RAD54 and ATRX-like proteins.

The recovery of the *drd1* mutant in our screen probably reflects the apparent sensitivity of the α’ promoter to non-CpG methylation. In contrast to the constitutive nopaline synthase promoter [5, 13], the α’ promoter is not appreciably silenced by CpG methylation, since nearly wild type levels of CpG methylation at this promoter remain when silencing is released in *drd1* mutants. These results support the idea that individual promoters vary in their sensitivity to cytosine methylation in different sequence contexts [2]. Whether the inverted repeat associated, RNA-directed DNA methylation pathway studied here is related to the RNAi-dependent silencing and *de novo* methylation of direct repeats in the *FWA* gene [19] remains to be determined.

Identifying DRD1 as a regulator of RNA-induced cytosine methylation establishes a function for a previously undefined subfamily of plant-specific SNF2-like proteins. DRD1 joins DDM1/Lsh1 [6, 7] and ATRX [8, 9], which maintain CpG methylation, on the list of SNF2-like proteins important for DNA methylation. The apparent preference of DRD1 for non-CpG methylation and the lack of close homologs in animals distinguish it from these other proteins. Whether DRD1 is involved in RNA-directed *de novo* methylation or acts to maintain non-CpG methylation induced by RNA remains to be determined. However, the heavy loss in *drd1* plants of CpNpN methylation, which is not efficiently maintained in the absence of the RNA trigger [10, 13], suggests a direct relationship between DRD1 activity and RNA signals. Given the relatedness of DRD1 to RAD54, it is intriguing to consider possible mechanistic similarities between RNA-directed DNA methylation and homologous DNA repair. In each case, the respective chromatin remodeling factor could facilitate a homology search on duplex DNA, nucleosome displacement and DNA unpairing, leading to heteroduplex formation and recruitment of enzyme complexes. In the RNA-directed DNA methylation pathway this could create an RNA-DNA hybrid that attracts DMTases, thus accounting for the extraordinary specificity of cytosine methylation, which is largely restricted to the region of RNA-DNA sequence similarity [1, 5, 13]. The lack of DRD1 homologs outside of the plant
kingdom may mean that RNA-directed DNA methylation occurs only in plants. Alternatively, RAD54 or ATRX-like proteins may serve this function in other organisms.

**Experimental procedures**

**Plant transformation**

*Arabidopsis thaliana* plants ecotype Columbia were transformed with *Agrobacterium tumefaciens* harboring either the target or silencer constructs (Supplementary information) using the floral dip method [20]. Transformants were selected on Murashige and Skoog medium containing either 50 µg/ml kanamycin (target construct) or 20 µg/ml hygromycin (silencer construct). The plant lines used in this study were the homozygous α’GFP target line and the corresponding ‘Double Transformant’ line DT7-3, which contains both the target and the silencer constructs.

**GFP visualization and Western blot analysis**

GFP fluorescence of mature seeds was visualized under the fluorescence microscope MZ FLIII (Leica, Wetzlar, Germany) with a GFP2 filter. The image was captured using a Colour CoolView Camera (Photonic Science, Kent, UK) and optimized using Adobe Photoshop. GFP protein was detected on Western blots using ECL Advance Western Blotting Detection kit (Amersham, Vienna, Austria) as recommended by the manufacturer using Living Colors A.v. antibody (Clontech, BD Biosciences, Belgium) for the first antibody and Anti-rabbit Ig, HRP-Linked whole AB (Amersham, Vienna, Austria) for the second antibody. The result was visualized on Kodak X-OMAT film. The film was scanned and the signals were quantified using ImageMaster 1D (Amersham, Vienna, Austria).

**RNA analysis**

α’ promoter double stranded (ds) RNA and short RNA were isolated from fresh rosette leaves of plants just starting to flower and analyzed by Northern blotting as described previously [5, 10, 13] with slight modifications. A ³²P-labeled α’ promoter DNA probe was used for dsRNA detection and ³²P-labeled sense or antisense RNA probe were used for small RNA detection on Northern blots, respectively.
DNA analysis

**Arabidopsis** total DNA for methylation analysis was isolated from dry mature seeds and fresh rosette leaves of BC1F3 *drd1-1* plants, fourth generation wild type DT7-3 plants, and fourth generation α’GFP plants using a DNeasy Plant Maxi Kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. Target α’ promoter methylation was analyzed using methylation sensitive restriction enzymes and Southern blot hybridization according to previously published protocols [5, 10, 13]. To probe the Southern blots, a 32P-labeled sense RNA corresponding to the first 220 bp of the GFP protein coding region was used. Methylation in the centromeric and rDNA repeats was analyzed as described previously [22].

For bisulfite sequencing, DNA preparations were treated with sodium bisulfite and subsequent steps were carried out according to standard protocols [5, 12]. Degenerate primers for amplification after bisulfite treatment were: 5’-AYGYGATAGAAAAAYAAAATATAG-3’ and 5’-CTTTACTCATTRTTATCTCC-3’ for the α’ promoter top strand. The number of PCR amplifications products sequenced: *drd1* leaf, 18; *drd1* seed, 16; DT7-3 leaf, 16; DT7-3 seeds, 21. Methylation appears slightly higher with bisulfite sequencing, which samples a subset of PCR amplification products, than with methylation-sensitive restriction enzymes, which assess methylation in the entire DNA sample. The general changes in methylation detected with both approaches are the same. Negligible methylation was detected by bisulfite sequencing in DNA isolated from seeds and leaves of the α’GFP target line (data not shown).

Acknowledgements

We thank Johannes van der Winden for DNA sequencing, Adriana Danielopol for assistance with plant work, and Ortrun Mittelsten Scheid for *som8/ddm1*-5 seeds. We are grateful to Cereon Genomics for marker information. This work was supported by grants to M.M. and A.J.M.M. from the Austrian Fonds zur Förderung der wissenschaftlichen Forschung (Grant nrs. Z21-MED and P-15611-B07) and the European Union (Contract QLK3-2000-00078).
D.P.K. acknowledges support by a Research Fellowship of the U.K. Medical Research Council (grant nr. G81/555).

References


15. Plant Chromatin Database (http://www.chromdb.org): see the proposed Clade A.


Legends to Figures

**Fig. 1.** (A) The target construct contains an α′ promoter- driven green fluorescent protein reporter gene (α′p-GFP), which is flanked by genes encoding resistance to kanamycin (Np-NPTII) and nopaline synthase (NOS). The silencer construct contains a hygromycin resistance marker (19Sp-HPT) and an inverted repeat of α′ promoter sequences separated by Np sequences (black bar), which is transcribed by the 35S promoter of cauliflower mosaic virus. pUC18 or pUC119 sequences containing an *E. coli* origin of replication enable plasmid rescue of transgene inserts. Abbreviations: LB and RB: left and right T-DNA borders, respectively; Np: NOS promoter; T: NOS transcription terminator. (B) Selfed seeds of: the homozygous α′GFP target line (no silencing; 100% GFP-positive), the DT7-3 line containing the silencer construct in the hemizygous state (segregating silencer; 75% GFP-negative) or the homozygous state (all GFP-negative), and the heterozygous drd1, 2 and 3 mutants (segregating recessive drd mutation; 25% GFP-positive). (C) Western blot analysis of GFP protein in seeds of the α′GFP target line (lane 1), wild type DT7-3 (lane 2), homozygous drd1,2 and 3 mutants (lanes 3, 5, 7), and their backcross progeny, where silencing is re-established (lanes 4, 6, 8).

**Fig. 2.** (A) DNA methylation analysis using enzymes sensitive to non-CpG methylation (F, S, Ps, Pa, E, B) and CpG methylation (H). Each enzyme was added to a standard double digest of *Kpn*I and *Nde*I, which are methylation-insensitive (‘-’ lanes). The dots at the left indicate the position of the methylated fragment. DNA was isolated from leaves and seeds of a control α′GFP plant (top; no methylation), a wild type DT7-3 plant (middle; increased CpG and non-CpG methylation induced by α′ promoter double stranded RNA), and the drd1-1 mutant (bottom; loss of predominantly non-CpG methylation). Positions of restriction enzymes in the 270 bp target α′ promoter and GFP probe (hatched) are shown at the bottom. The black portion indicates the region of RNA-DNA sequence identity. Abbreviations of enzymes and their recognition sequences (C sensitive to methylation indicated by the
superscript ‘m’): *Hpa*CH4IV: A*m*CGT; *Fnu*4HI: G*m*C*m*NG*m*C, if N is C; *Scr*Fl: C*m*CNGG; *Pst*I: m*CTG*m*CAG; *Pst*I: T*m*CATGA; *Eco*T22I: ATG*m*CAT; *Bam*HI: GGAT*m*C*m*C.

(B) α’ promoter short RNAs (~ 21-24 nt) of both polarities in a wild type DT7-3 plant (lanes 1) and a *drd1-1* mutant plant (lanes 3). Lanes 2 (left) and (right) contain RNA isolated from the control α’GFP plant. Ethidium bromide staining of the major RNA size class in the samples is shown as a loading control [21]. The position of a 23 nt DNA oligonucleotide size marker is indicated by the arrow.

**Fig. 3.** Bisulfite sequence analysis. The percent methylation in the top DNA strand of the target α’ promoter in DNA isolated from leaves and seeds of the *drd1-1* mutant (top) and a wild type DT7-3 plant (bottom). The map in the middle indicates the positions of methylation-sensitive restriction enzyme sites in the α’ promoter (Fig. 2) and the region of RNA-DNA sequence identity (black portion). CpG, black; CpNpG, blue; CpNpN, red. The individual CpG and CpNpG profiles are shown in Supplementary Fig. 2.

**Fig. 4.** Methylation analysis of repetitive sequences. The enzymes used are sensitive to methylation in CpG and CpNpG (*Hpa*I; m*C*m*CGG); only CpNpG (*Msp*I: m*CCGG); or any type depending on the sequence context of the recognition sequence (*Sau*3A: GAT*m*C; the methylation-insensitive isoschizomer is *Nde*I). In untransformed *Arabidopsis* (Col-0) and wild type DT7-3 plants, the digestion patterns of the 180 bp centromeric repeat (A) and rDNA repeats (B) reflect substantial methylation of CpGs and non-CpGs as indicated by increased cleavage with the methylation-insensitive isoschizomer (M or N) compared to the methylation–sensitive isoschizomer (H or S, respectively). Identical patterns are observed in *drd1-1* plants, indicating no change in the levels of CpG or non-CpG methylation. In contrast, mutants defective in DDM1, a global regulator of cytosine methylation [6, 20] exhibit substantial reductions of CpG and non-CpG methylation, as revealed by similar cleavage with the methylation-sensitive and methylation–insensitive isoschizomers.
Fig. 5. Domain structure of DRD1 and positions of \textit{drd1} mutations and known functional signatures. The four ProDom domains defining the DRD1 family are shown in bold (I, III, IV, VI). The six domains present in DRD1 and its closest homologs (I-VI) are PD423058, PD217428, PD000441, PD210292, PD690098, PD039514. A putative binding site for Mg$^{++}$ is at 492-495 (mutated in \textit{drd1-1}) and for ATP at Q803, R807 (mutated in \textit{drd1-2}) and R810 (mutated in \textit{drd1-5}). NLS, predicted bipartite nuclear localization signal. The ‘#’ sign indicates a premature stop codon.
Supplementary information

α’ promoter silencing system and plant transformation

The α’ promoter fragment used in this study is a 270 bp fragment from the gene encoding the α’ subunit of β-conglycinin, a soybean seed storage protein. The fragment corresponds to positions –316 to –46 relative to the translational start codon as given in DDBJ/EMBL/GenBank accession No. M13759 and –257 to +13 relative to the transcription start site [1]. For the α’GFP target construct, a plasmid containing 35Spro-smRS-GFP-NOSter (soluble modified red shifted GFP) [2] was obtained as clone CD3-327 from the Arabidopsis Biological Resource Center of Ohio State University. In this clone, the 35S promoter was replaced with a fragment containing the α’ promoter fragment. The plasmid was linearized with HindIII and ligated into the HindIII of BV4 [3] in the orientation shown in Fig. 1. The 35S-α’proIR silencer construct contains the α’ promoter fragment in sense orientation with respect to the 35S promoter and an α’ promoter fragment in the antisense orientation. The two α’ promoter fragments are separated by a 298 bp spacer containing the NOSpro (-264 to +63 with respect to the transcriptional start site as described in DDBJ/EMBL/Gene Bank accession No.V00087) in sense orientation relative to the 35S promoter assembled on pUC18. The plasmid was linearized with HindIII and inserted into BV:19S-HPT [4] in the orientation shown in Fig. 1.

By using plasmid rescue, the transgene insertion sites were found to be in intergenic regions on chromosome 2 in BAC clone T6A23, position 88894 (left T-DNA border) and 88914 (right T-DNA border) (target construct); and chromosome 5 in P1 clone MCL19, position 37705 (left T-DNA border; right border not sequenced) (silencer construct).

Genetic mapping of DRD1

For genetic mapping and cloning of the DRD1 gene, a mapping population was obtained by crossing BC1F3 drd 1-1 plants (genotype: homozygous target locus, homozygous silencer locus, homozygous drd1-1) to ecotype Landsberg erecta. F2 seeds of this cross were pre-screened for hygromycin resistance (indicating presence of the silencer
construct) and green fluorescence. In plants containing the silencer construct, only plants that are homozygous for the \textit{drd1-1} mutation show green fluorescence. Using an initial mapping population of 27 F2 plants and a set of cleaved amplified polymorphic sequence (CAPS) markers that detect polymorphisms between the Columbia and Landsberg ecotypes \cite{5}, the \textit{drd1-1} mutation was found to be linked to the THY1 marker (www.arabidopsis.org) on chromosome 2 (7.1Mb: 0 recombinants per 27 F2 plants). With an increased mapping population of \~{} 350 F2 plants, the \textit{drd1-1} mutation was found to reside within an approximately 264 kb physical interval between INDEL (insertion/deletion) marker CER453031 (Cereon Genomics; one recombinant) and CER459637 (one recombinant). This interval includes 75 annotated ORFs (www.arabidopsis.org). By annotation \cite{6}, one ORF could be considered a \textit{bona fide} candidate for encoding the DRD1 protein: At2g16390, which encodes a putative chromatin remodeling protein of the SNF2 family. The \textit{DRD1} gene was PCR-amplified from mutant (\textit{drd1-1}, \textit{drd1-2}, \textit{drd1-3}, \textit{drd1-4}, \textit{drd1-5} and \textit{drd1-6}) and wild type DT7-3 plants. The PCR fragments were purified with QIAquick PCR purification kit (Qiagen, Hilden, Germany) and sequenced using Thermo Sequenase (Amersham, Vienna, Austria) and dye-labeled primers (Biolegio, Malden, The Netherlands). The sequencing reactions were resolved on a LICOR sequencer (LICOR, Omaha, NE). Primers for PCR amplification were as follows: CHR35-1F, 5’-ACATTATATGCATGTCGTTCC-3’; CHR35-1R, 5’-AGTTTCTGCTCGTCTTACC-3’ these primers were also used for sequencing. Additional sequencing primers were as follows: CHR35-2, 5’-TGAGAAGACGATGGATAATC-3’, CHR35-3, 5’-GTAATTTTGTAGGAGGAC-3’, CHR35-4, 5’-CATTTTCGTTGGAGGAC-3’, CHR35-5, 5’-CTACGCTAATTGTGTCGTTCC-3’, CHR35-7, 5’-ACTCCAAGGATGAGGAC-3’, CHR35-8, 5’-GTGTTCTTAATCTGTCC-3’.

Supplementary Figure 1: Individual profiles of CpG and CpNpG methylation in the target \textit{\alpha’} promoter as determined by bisulfite sequencing. The complete profile, including CpNpN methylation, is shown in Fig. 3.
Supplementary Figure 2: Amino acid alignments of DRD1, *A. thaliana* protein CHR34, and two rice (*Oryza sativa*) homologs, BAC84084 and AAM15781. In addition, homologs of
DRD1 were identified by Smith-Waterman alignment of the CAST masked sequence [7] to proteins from a comprehensive non-redundant database [8, 9]. Functional motifs and domains were studied using integrated databases of conserved sequence patterns and their respective search tools [10].

References for Supplementary information and figures

8. http://www.ebi.ac.uk/MPsrch/
A

Target construct

Silencer construct

B

\[
\text{\(\alpha'\text{GFP}\)} \\
\text{DT7-3 (Henni)} \\
\text{DT7-3 (Homo)}
\]

drd1  

C

1 2 3 4 5 6 7 8

Fig. 1  Kanno et al.
Fig. 2 Kanno et al.
Fig. 3 Kanno et al.
Fig. 4  Kanno et al.
Supplementary figure 1,  Kanno et al.