The Pennsylvania State University
The Graduate School
Graduate Program in Plant Physiology

DISSECTING THE MECHANISM UNDERLYING EPIGENETIC ACTIVATION OF THE MAIZE *SPM* TRANSPOSON BY THE ELEMENT-ENCODED TNPA PROTEIN

A Thesis in
Plant Physiology
by
Hongchang Cui

Submitted in Partial Fulfillment
of the Requirements
for the Degree of
Doctor of Philosophy

August 2003

The thesis of Hongchang Cui has been reviewed and approved by the following:

Nina V. Fedoroff
Willaman Professor of Life Sciences
Evan Pugh Professor
Thesis Co-Advisor
Chair of Committee

Jerry L.Workman
Associate Investigator, Howard Hughes Medical Institute
Paul Berg Professor of Biochemistry and Molecular Biology
Thesis Co-Advisor

Hong Ma
Professor of Biology

Teh-hui Kao
Professor of Biochemistry and Molecular Biology
Head of the Graduate Program in Plant Physiology

* Signatures are on file in the Graduate School.

ABSTRACT

The activity of the maize *Spm* transposon is epigenetically regulated. The promoter and its downstream GC-rich sequence are extensively methylated when the transposon is inactive and unmethylated when the element is active. A methylated, inactive *Spm* transposon can be activated by an active element, and the activation is accompanied by loss of DNA methylation in both the promoter and the GC-rich sequence. Previous studies have identified TnpA, one of the *Spm*-encoded transposase proteins, as the trans-acting factor that mediates the epigenetic activation of an inactive *Spm*. However, how TnpA promotes DNA demethylation has yet to be determined.

To facilitate elucidating the underlying mechanism, I have developed a novel assay system that permits demethylation of the *Spm* sequence to be controlled by inducing expression of TnpA in transgenic tobacco cells. Using this inducible DNA demethylation system, I show that TnpA-mediated DNA demethylation occurs at a rate much faster than that attributable to interference with methylation maintenance during DNA replication. This observation strongly suggests an active process.

In further studies, I show that TnpA is a weak transcriptional activator and that deletions that disrupt its ability to activate transcription also eliminate its ability to promote DNA demethylation. Since the fusion protein between the truncated protein and the viral VP16 activation domain has stronger DNA demethylation activity as well as transcriptional activity, it is thus concluded that TnpA's transcriptional activity plays an important role in promoting DNA demethylation.

Although TnpA-mediated demethylation of the *Spm* sequence is rapid over the whole process, initial DNA demethylation is very slow. Moreover, DNA demethylation has been observed only in callus and suspension cells, but not in leaves that have ceased to divide. Using inhibitors of DNA replication and cell division, I show that DNA replication is required as well for TnpA-mediated *Spm* demethylation. By using a gel mobility shift assay I show that the binding affinity of TnpA to fully methylated DNA fragments derived from *Spm* termini is much lower than its affinity for the same fragments when hemi-methylated or unmethylated. Based on these observations, a two-step DNA demethylation mechanism is suggested where TnpA binds to the post-replicative, hemi-methylated *Spm* sequence and promotes demethylation either by creating an appropriate demethylation substrate or by itself participating in or recruiting a demethylase complex.

In animals cells several enzymatic activities have been identified that appear to be capable of converting 5-methylated cytosine (5mC) into normal cytosine through distinct mechanisms. Although there are cytological observations that suggest the presence of DNA demethylase in plant cells, no biochemical evidence has ever been obtained. In the present study, I have developed an *in vitro* DNA demethylation assay for which a hemimethylated DNA fragment derived from *Spm* promoter and GC-rich sequence is used as the substrate. A DNA demethylase activity is detected in nuclear extract that is prepared from suspension cultured tobacco cells.

To investigate whether TnpA is itself a sequence-specific DNA demethylase or it recruits a DNA demethylase activity to the *Spm* sequence, I have purified native TnpA protein and its truncated derivatives from tobacco suspension cells by

coimmunoprecipitation. No DNA demethylase activity is detected in the TnpA-containing immunoprecipitate, however, indicating the lack of a specific physical interaction between TnpA and DNA demethylase. These results clearly rule out the possibility that TnpA is a sequence-specific DNA demethylase.

To further explore the connection between transcriptional activation and DNA demethylation, I have generated transgenic tobacco lines with the *Spm* promoter and GC-rich sequence under the control of a glucocorticoid-inducible promoter. DNA demethylation is observed in lines that show inducible transcription through the *Spm* sequence. *Spm* demethylation is rapid, again suggesting an active DNA demethylation process. Taken together, these observations strongly suggest that TnpA promotes *Spm* demethylation by facilitating the access of DNA demethylase activities to methylated DNA as a consequence of transcriptional activation.

Since several other trans-acting factors that are known to promote DNA demethylation are also transcription factors, the present study has likely revealed a common mechanism for epigenetic regulation. The theoretical and practical implications of the present findings are discussed. Lastly, several future directions and approaches are briefly described.

TABLE OF CONTENTS

Chapter 8 CONCLUSION

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

This work has been conducted under the supervision of Dr. Nina Fedoroff. DNA methylation and the underlying epigenetic mechanisms have been an area of long-lasting interest in animal studies, cancer research particularly, and such research with plants has become intense in recent years. At the time I embarked on this project, however, understanding in this area, especially about DNA demethylation, was still very rudimentary. Not surprisingly, I have encountered many challenges, both technical and conceptual, in the early stage of my studies. One of the challenges has been to develop a new assay system for dissecting the mechanism by which TnpA promotes DNA demethylation. Although some of my ideas worked well, yielding interesting results, for a long time I had difficulty working out the key assay system. Eventually I was successful with the assay, which has allowed me to address many important questions. However, this would not have been possible without the invaluable advice and much-needed encouragement as well as continuing support from Dr. Fedoroff. As my mentor, Dr. Fedoroff has also helped me profoundly in my developing critical thinking skills as well as understanding of the Western culture. There is no doubt that the training in her lab has set a new stage for me and this experience will continue to benefit my future academic career. I am deeply indebted to her for giving me the opportunity to study and work in her laboratory.

I also would like to express my gratitude to Dr. Teh-hui Kao, Dr. Hong Ma and Dr. Jerry Workman for their service on my committee and invaluable suggestions. I am particularly thankful to Dr. Jerry Workman for his willingness to serve as my co-advisor.

My special thanks go to my wife, Yueling Hao, for her understanding and unconditional support of my studies. In order for me to focus on research, she has given up her job to stay at home taking care of our children and me. In the course of my studies, my second child was born and this has further increased her workload. Since she is not eligible for work in this country, she has to manage to make the both ends meet for an already tight budget. She surely has sacrificed a lot for me and certainly deserves to share all the credits I have earned so far.

Chapter 1

INTRODUCTION

DNA methylation is an almost universal phenomenon in living organisms ranging from bacteria to the human being (Palmer and Marinus, 1994; Yoder and Bestor, 1996). In prokaryotic cells, methylation occurs on both adenosine and cytosine, giving rise to 6'-N-methyl adenosine and 5-methylcytosine (5mC) respectively, and is functionally important in such processes as methylation-directed DNA repair (Modrich, 1989), control of DNA replication initiation (Abeles et al., 1993), as well as protection of the host cell against foreign nucleic acid invasion (Palmer and Marinus, 1994). By contrast, DNA methylation in eukaryotes is normally found as 5mC only. Whether 5mC maintains these prototypic functions is still unclear, but there are indications that it may play a role in controlling initiation of DNA replication (Rein et al., 1997). Extensive studies over the past decades, however, have shown that DNA methylation is an important epigenetic modification for normal growth and development in eukaryotes.

DNA Methylation is Essential for Growth and Development

In mammals, DNA methylation is linked to a variety of phenomena such as parental imprinting (Constancia et al., 1998; Reik and Walter, 1998), X-inactivation (Allaman-Pillet et al., 1998), and allelic exclusion of gene expression (Mostoslavsky et al., 1998).

Since 5mC is normally associated with virus sequences, transposons and retrotransposons as well as other repetitive sequences (Kochanek et al., 1995; ten Lohuis et al., 1995; Bingham, 1997; Yoder et al., 1997; Fedoroff, 2000), DNA methylation has been regarded as a safeguarding mechanism against genome instability. DNA methylation has also been shown to repress homologous recombination (Maloisel and Rossignol, 1998).

The importance of DNA methylation to eukaryotic organisms has been revealed by studies of animal and plant mutants that have a reduced level of DNA methylation. In the mouse, a deficiency in DNA methylation results in embryonic death early in development (Li et al., 1992; Okano et al., 1999). A variety of diseases have been associated with a genome-wide decrease in DNA methylation (Hendrich, 2000), among them Rett syndrome (Van den Veyver and Zoghbi, 2000) and immunodeficiency centromeric instability and facial anomalies (ICF) (Hansen et al., 1999). Although the survival of *Arabidopsis ddm1* mutant is not affected, its growth and development are seriously impaired (Vongs et al., 1993; Kakutani et al., 1996). Deficiency in genome-wide DNA methylation causes abnormal flower development (Jacobsen and Meyerowitz, 1997; Jacobsen et al., 2000; Soppe et al., 2000) and loss of imprinting (Baroux et al., 2002; Grini et al., 2002). Moreover, many tissue- and cell-specific genes have specific patterns of DNA methylation, and maintenance of this methylation pattern seems essential for normal cell growth and differentiation (Rossi et al., 1997; Leegwater et al., 1998; Samac et al., 1998). Programmed changes in the DNA methylation pattern of flower specific genes has been reported during the process of vernalization, a treatment that promotes flowering (Sheldon et al., 1999). In human cancer cells, genes that are essential for cell

proliferation are often found to be associated with aberrant DNA methylation (Schmutte and Jones, 1998).

Biochemistry of DNA Methylation

Maintenance of DNA Methylation

DNA methylation is heritable, that is, it can be passed through DNA replication and to succeeding generations. DNA methylation is maintained by maintenance DNA methyltransferase (Bestor and Verdine, 1994; Adams, 1995). In animals, two maintenance DNA methyltransferase genes have been identified, *DNMT1* and *DNMT2*, but only DNMT1 has been shown to have DNA methyltransferase activity (Hsieh, 1999a; Robert et al., 2003). It has been reported recently that not only the C-terminal catalytic domain, but also the N-terminal regulatory domain, is essential for DNA methylation (Araujo et al., 2001). The lack of the N-terminal domain in DNMT2 is the likely reason why DNMT2 does not have methyltransferase activity (Okano et al., 1998), although it contains all the catalytic domains (Yoder and Bestor, 1998). In contrast, plants have two functional DNA methyltransferase genes. In addition to MET1, which is the counterpart of DNMT1 in animals (Jeddeloh et al., 1998), a new methyltransferase, CHROMOMETHYLASE 3, appears to be a plant specific DNA methyltransferase. Unlike DNMT1 or MET1, which maintains methylation in CpG dinucleotides (Finnegan and Kovac, 2000), CHROMOMETHYLASE 3 (CMT3) is responsible for the

maintenance of CpXpG methylation, which is a major DNA modification in plants but essentially absent in animals (Jeddeloh and Richards, 1996; Lindroth et al., 2001).

Consistent with a role of DNMT1 in maintenance of DNA methylation, *in vitro* studies have shown that DNMT1 recognizes and preferentially methylates hemimethylated duplex DNA. Moreover, maintenance of DNA methylation appears to be coupled with DNA replication, as DNMT1 colocalizes at the replication loci with PCNA, a DNA polymerase accessory factor (Leonhardt et al., 1992; Chuang et al., 1997; Liu et al., 1998). Physical interaction between DNMT1 and PCNA has also been reported (Araujo et al., 2001).

De novo DNA Methylation

Unmethylated DNA sequences, either exogenous or endogenous, can also be methylated, which is defined as *de novo* DNA methylation. Since DNMT1 also shows *de novo* methyltransferase activity by an *in vitro* assay and for a long time has remained the only known DNA methyltransferase, it has been regarded as the *de novo* methylation methyltransferase as well. However, *in vivo* studies have clearly demonstrated that DNMT1 functions only in maintenance of DNA methylation (Hsieh, 1999a). In the meantime, other proteins with *de novo* DNA methyltransferase activity have been identified.

In most organisms, *de novo* DNA methyltransferases are distinct from the maintenance DNA methyltransferases. The fungus *Neurospora* is an exception, however, where the only DNA methyltransferase, DIM-2, is responsible for both the maintenance and *de novo* DNA methylation (Kouzminova and Selker, 2001). In mammals, two *de*

novo DNA methyltransferases, DNMT3a and DNMT3b, are well characterized. Both proteins show *de novo* methyltransferase activity by both *in vitro* (Okano et al., 1999) and *in vivo* assays (Hsieh, 1999a) and use unmethylated DNA as the preferred substrate (Yokochi and Robertson, 2002). DNMT3b has a weaker methyltransferase activity and different cellular localization than DNMT3a (Bachman et al., 2001). DNMT3a is primarily involved in CpG methylation, but also catalyzes cytosine methylation in an asymmetric (CpN) context (Ramsahoye et al., 2000).

Recent studies show that DNMT3a and 3b may also play an important role in maintenance of DNA methylation in cooperation with DNMT1. DNMT3a and DNMT3b could form a complex and colocalize with DNMT1 in the nucleus (Kim et al., 2002). Since hemimethylated DNA of a repetitive sequence can be converted to fully methylated DNA by overexpression of DNMT3a and DNMT3b, but not DNMT1, it is suggested that maintenance of DNA methylation in some sequences could be partly achieved through *de novo* methylation (Liang et al., 2002). This accounts for the observation that substantial level of DNA methylation is still detected in human cells lacking the *DNMT*1 gene (Rhee et al., 2000).

A mammalian *de novo* methyltransferase-like gene, *DNMT3L*, has been identified as well (Aapola et al., 2000). Unlike DNMT3a and DNMT3b, DNMT3L lacks DNA methyltransferase catalytic domains (Bourc'his et al., 2001). Nevertheless, it still stimulates *de novo* methylation in maternal imprinting (Bourc'his et al., 2001). Subsequent studies have revealed that DNA methylation due to DNMT3L is achieved by recruitment of the DNMT3a protein (Chedin et al., 2002; Hata et al., 2002).

Two homologs of the animal *de novo DNA* methyltransferase genes, *DRM1* and *DRM2*, have been identified in *Arabidopsis* (Cao et al., 2000), and there is genetic evidence that they are indeed required for *de novo* DNA methylation (Cao and Jacobsen, 2002a). All the conserved methyltransferase motifs are present in these proteins but in a noncanical order, which gives their names (DRM stands for Domain Rearranged Methyltransferase) (Cao et al., 2000). Unlike *DNMT3a* and *DNMT3b*, *DRM1* and *DRM2* appear to be redundant genes because only double mutants show decrease in DNA methylation (Cao and Jacobsen, 2002a). Moreover, *DRM1* and *DRM2* also show overlapping function with *CMT3* in maintaining CpNpG methylation (Cao and Jacobsen, 2002b). These findings indicate that plants have evolved some unique feature in controlling DNA methylation.

DNA Methylation Represses Transcription

In eukaryotes, a major function of DNA methylation is to represse transcription. In some cases, DNA methylation may inhibit transcription directly by reducing the binding affinity between transcription factors and their target promoter sequences (Bednarik et al., 1991; Inamdar et al., 1991; Kochanek et al., 1995). In others, DNA methylation does not apparently affect DNA binding by transcription factors, but transcription is still efficiently repressed (Holler et al., 1988; Silke et al., 1995; Rossi et al., 1997). Obviously, in the latter situation DNA methylation alone is insufficient to bring about transcription inhibition, suggesting the involvement of additional mechanisms. In agreement with this view, a family of methylated-DNA binding proteins (MBDs) has been identified (Meehan

et al., 1989; Lewis et al., 1992) (Hendrich and Bird, 1998), and most of them can target to methylated DNA for transcription repression (Fujita et al., 1999; Ng et al., 1999; Ng et al., 2000; Hendrich et al., 2001).

Methylated DNA Binding Proteins

The essential role of MBDs for transcription repression by DNA methylation has been elegantly demonstrated in studies of MeCP2, which, due to its abundance in the nucleus, has received the earliest attention. MeCP2 contains two domains that are required for transcription repression: one is a methylated DNA binding domain and the other a repressor domain (Nan et al., 1997). In *Drosophila melanogaster*, which lacks substantial level of DNA methylation (Lyko et al., 2000), transcription activity from the leukosialin promoter does not differ substantially whether the promoter is unmethylated or methylated. Strikingly, however, transcription from the methylated promoter is significantly inhibited when the human *MeCP2* gene is expressed in *Drosophila* (Kudo, 1998).

DNA Methylation Functions through Chromatin Remodeling

Despite a general correlation between DNA methylation and transcription repression, numerous examples exist where heavily methylated DNA is still actively transcribed (Garrick et al., 1996; Walsh and Bestor, 1999). These observations support the view that DNA methylation alone is not sufficient for gene silencing. They further suggest that the presence of MBDs does not necessarily leads to transcriptional repression. More recent studies have revealed that transcription repression by DNA

methylation interfaces with another gene regulatory mechanism, i.e., chromatin remodeling and chromatin protein modification, histones in particular (Geiman and Robertson, 2002; Robertson, 2002).

A functional link between DNA methylation and chromatin remodeling is suggested by an early study on transcription regulation of the HSV tk gene by DNA methylation (Kass et al., 1997). Kass *et al* found that methylated templates are initially assembled into active transcription complexes and transcription inactivation occurred only when chromatin structure was reconstituted (Kass et al., 1997). In agreement with a key role of chromatin remodeling in transcriptional regulation, heavily methylated and transcriptionally inactive genes are generally associated with heterochromatin, a facultatively condensed form of chromatin (Geiman and Robertson, 2002; Robertson, 2002) and transcriptional activation results in chromatin decondensation (Tumbar et al., 1999; Muller et al., 2001; Nye et al., 2002).

Further studies have shown that all known transcription repressor-like methyl-DNA binding proteins, including MBD1, MBD2, MBD3 as well as MeCP2, are components of larger repressor complexes that contain histone deacetylase activity (Yoder and Bestor, 1996; Razin, 1998; Ballestar and Wolffe, 2001). Hyperacetylation of histones, on the 9th lysine of histone H3 particularly, is a hallmark of transcriptionally active genes. These observations suggest that histone modification may determine the function of DNA methylation. Evidence that histone deacetylase is indeed essential for MBD-mediated gene silencing is derived from the observation that repression of transcription by MBD1 and MBD2 is relieved by trichostatin A, a specific histone deacetylase inhibitor (Ng et al., 2000; Magdinier and Wolffe, 2001). Moreover, disruption of the histone deacetylase

activity in the MeCP1 complex abolishes its ability to represses transcription from methylated DNA (Feng and Zhang, 2001). Interestingly, many transcriptional factors such as GCN5 and the viral VP16 protein have been shown to contain histone acetyltransferase activity or interact with other proteins with such an activity (Tumbar et al., 1999; Dyda et al., 2000; Marmorstein, 2001). Notably, yeast and *Drosophila*, which lack substantial levels of DNA methylation, still show a full spectrum of epigenetic phenomena (Paro et al., 1998; Lyko et al., 2000). Although some MBD-like proteins have been identified in *Drosophila*, they appear to repress transcription independently of DNA methylation (Roder et al., 2000; Ballestar et al., 2001). DNA methylation has thus come to be regarded as an additional layer of regulation enhancing gene silencing and maintaining low level of background transcription.

Interplay between DNA Methylation, Chromatin Remodeling, Histone Modifications and Gene Silencing

More recent studies have revealed intimate links between DNA methylation, chromatin remodeling, histone modifications, and gene silencing. On the one hand, chromatin remodeling and histone modifications are essential not only for the function but also for the maintenance or occurrence of DNA methylation. On the other hand, DNA methylation is able to alter the structure of chromatin and the status of histone modifications by recruiting chromatin-remodeling proteins. By such a positive feedback

mechanism, transcriptional repression by DNA methylation is enhanced, ensuring robust control of gene regulation.

Dependence of DNA Methylation on Chromatin Remodeling and Histone Modifications

An important role of chromatin proteins in maintaining DNA methylation is uncovered by the finding in plants that mutations in *DDM1*, a ATP-dependent SWI/SNF-like chromatin remodeling gene, cause a genome-wide decrease in DNA methylation (Jeddeloh et al., 1999). As the *ddm1* mutation does not affect the expression or the activity of DNA methyltransferase (Bartee and Bender, 2001), it is likely that the DDM1 protein itself is required for the maintenance of DNA methylation. A similar phenotype has been observed in mammals that lack the related chromatin remodeling proteins LSH in the mouse (Dennis et al., 2001) and ATRX in the human (Gibbons et al., 2000a; Bourc'his and Bestor, 2002).

How chromatin-remodeling proteins affect the status of DNA methylation is still unclear, but there is convincing data that histone methylation, particularly at the 9th lysine of histone H3 (H3mK9), and heterochromatin formation are key determinants of DNA methylation. In the fungus *Neurospora*, a mutation in *DIM-5*, the sole histone methyltransferase gene, leads to a complete loss of DNA methylation (Tamaru and Selker, 2001; Selker et al., 2002). Mutations in *KRYPTONITE*, a histone methyltransferase gene in *Arabidopsis*, cause loss of DNA methylation as well, albeit primarily in the CpNpG context (Jackson et al., 2002). Moreover, the *ddm1* and *kryp* mutants have a reduced level of histone methylation in regions where the level of DNA methylation decreases (Gendrel et al., 2002). The latter study also showed that

heterochromatin protein HP1, but not methylated H3, interacts with CMT3, the CpNpG DNA methyltransferase (Jackson et al., 2002). The role of HP1 in heterochromatin formation has been demonstrated in *Drosophila* (Li et al., 2003). These studies have therefore revealed an important role of heterochromatin in DNA methylation. Indeed, both the *ddm1* and *atrx* mutants show defects in chromatin condensation (Gibbons et al., 2000; Probst et al., 2003).

In addition to methylation, histones have other modifications such as acetylation, phosphorylation and ubiquitinization. Each seems to play a different role in gene regulation (Richards and Elgin, 2002). Since acetylation and methylation cannot occur simultaneously at the same position (K9), acetylated histones must be deacetylated before being methylated. In the *ddm1* and *kryp* mutants, a decrease in the extent of DNA methylation is associated not only with the loss of histone methylation, but also with histone hyperacetylation (Gendrel et al., 2002; Johnson et al., 2002). Histone deacetylation thus could play a critical role in the occurrence of DNA methylation.

Transcriptional Gene Silencing Leads to DNA Methylation

Heterochromatin, histone hypoacetylation and methylation at K9 are generally associated with transcriptionally inactive sequences such as the centromeres, transposons and other repetitive sequences (Avramova, 2002; Richards and Elgin, 2002). Since this is also true for such organisms as yeast and *Drosophila*, which are either deficient in DNA methylation or lack substantial levels of DNA methylation (Schotta et al., 2002), DNA methylation could be secondary to transcriptional inactivation. Supporting this view, silencers derived from *Drosophila* not only repress transcription but also become

methylated in transgenic mice (Brenton et al., 1999). Conversely, mammalian sequences that remain methylated in their host behave as silencers in *Drosophila* without involvement of DNA methylation (Drewell et al., 2000). Importantly, accumulating evidence indicates that DNA methyltransferases do not bind DNA themselves but are recruited to target DNA sequences through sequence-specific DNA binding transcriptional repressors (Fuks et al., 2001; Kishimoto et al., 2001; Zilberman et al., 2003). This is also true for DNMT3L, a DNA methyltransferase homolog (Aapola et al., 2000) that has no DNA methyltransferase activity (Aapola et al., 2002). Recently DNMT3L has been shown to be a transcriptional repressor that interacts with histone deacetylase and directs DNA methylation through recruitment of *de novo* DNA methyltransferases DNMT3A and DNMT3B (Deplus et al., 2002). Although there is disagreement over whether DNA methylation determines the status of histone methylation and hypoacetylation (Soppe et al., 2002), most studies strongly suggest that transcriptional gene silencing determines or precedes the occurrence of DNA methylation.

Feedback of DNA Methylation on Gene Silencing, Histone Modifications and Chromatin Remoldelling

Once established, DNA methylation not only maintains gene silencing, but also affects transcription and DNA methylation in flanking sequences by altering the status of histone modification and chromatin structure. As discussed previously, there exist a number of MBDs that are associated with histone deacetylase. All DNA methyltransferases are associated with histone deacetylases (Burgers et al., 2002).

Hypoacetylated histones arising from the binding of MBDs and DNA methyltransferases would become potential sites for histone methyltransferase, which in turn modifies the histones and eventually effects DNA methylation. By such a positive feedback mechanism, DNA methylation and heterochromatin can be spread into flanking sequences, and in the case of X-inactivation it even spreads to the whole chromosome (Brockdorff, 2002).

Spreading of DNA methylation is not unlimited, however, due to the counteraction by transcriptional activators. Many transcriptional co-activators such GCN5 and P300/CBP have histone acetyltransferase activity and they can prevent the formation of heterochromatin by maintaining histones in a hyperacetylated state (Frit et al., 2002). Transcriptional activators can also resolve compact heterochromatin by recruiting chromatin-remodeling factors like SWI/SNF proteins (Varga-Weisz, 2001; Li et al., 2002). The boundary of heterochromatin thus seems to be determined by the relative strength of the two opposing forces: histone hyperacetylation resulting from transcriptional activation and histone hypoacetylation and methylation as a consequence of repressor binding (Dillon and Festenstein, 2002). Since different genes need to be transcribed at different stages of development, the boundary is not static, but dynamic.

DNA Methylation is Reversible

DNA methylation is both heritable and reversible. Methylated DNA can become demethylated in one of two ways, designated “passive” and “active” (Hsieh, 2000; Kress et al., 2001; Reik et al., 2001). Passive DNA demethylation occurs by the replication of

methylated DNA without remethylation. This gradually dilutes the concentration of methyl groups and methylated DNA strands in a cell population. Passive DNA demethylation is a predictable function of cell division time, since fully methylated DNA becomes hemimethylated after one cycle of replication and only half of the daughter molecules are unmethylated after two replication cycles. Active DNA demethylation occurs rapidly and can, in some cases, be independent of DNA replication. In mammals, genome-wide DNA demethylation occurs in germ cells, as well as in early embryonic development (Dean et al., 2001; Reik et al., 2001). Loss of DNA methylation in the paternal genome after fertilization is rapid (within 4 hrs) and independent of DNA replication (Mayer et al., 2000; Oswald et al., 2000; Santos et al., 2002). However, the heavily methylated maternal genome is not demethylated as rapidly as the paternal genome (Oswald et al., 2000; Dean et al., 2001). Little is known about methylation and demethylation during plant development, but it appears that DNA is actively demethylated during pollen development (Oakeley et al., 1997).

DNA demethylation also occurs during later stages of plant and animal development, but in a sequence-specific manner. Genome-wide DNA demethylation in early animal embryonic development is followed by *de novo* methylation (Reik et al., 2001). The new methylation pattern is stably propagated during development, except for the selective demethylation of genes in specific tissues or at particular developmental stages (Migeon et al., 1991; Brunk et al., 1996; Kirillov et al., 1996; Grange et al., 2001). Tissue-specific gene demethylation has been observed in response to hormone treatment (Wilks et al., 1984; Grange et al., 2001; Thomassin et al., 2001). Gene-specific DNA demethylation has also been detected during vernalization in certain flower-specific

genes (Sheldon et al., 1999; Soppe et al., 2000). The most extensively studied demethylation mechanisms in plants are those associated with transposable elements and there is evidence for both global demethylation of plant transposons and sequence-specific demethylation (Fedoroff et al., 1995; Jeddeloh et al., 1999; Hirochika et al., 2000; Bartee and Bender, 2001).

Biochemistry of DNA Demethylation**.**

While substantial progress has been made in identifying and analyzing the multiple DNA methyltransferases of both higher plants and animals, relatively little is known about the biochemistry of DNA demethylation. Several DNA demethylation mechanisms have been identified in animal cells. The nucleotide excision repair (NER) pathway entails nicking of the DNA duplex at 5-methylcytosine (5mC) residues and removal of the nucleotide (Jost, 1993; Vairapandi and Duker, 1993). The base excision repair (BER) pathway commences with the cleavage of the glycosidic bond and removal of the 5-methylcytosine by a DNA glycosylase (Jost and Jost, 1995; Jost et al., 1995; Vairapandi and Duker, 1996). The resulting gap from the NER pathway or the abasic site from BER is subsequently replaced by cytosine as a result of DNA repair activities (Jost and Jost, 1995; Jost et al., 1995; Vairapandi and Duker, 1996; Scharer and Jiricny, 2001). Two distinct 5-methylcytosine DNA glycosylase activities have been reported, one is specific for fully methylated DNA (Vairapandi and Duker, 1996) and the other has a preference for hemimethylated DNA as the substrate (Jost and Jost, 1995; Jost et al., 1995). Both 5-methylcytosine DNA glycosylase enzymes appear to require an RNA moiety for activity (Fremont et al., 1997; Jost et al., 1997; Vairapandi et al., 2000).

DNA demethylation can also be achieved by deamination of 5mC to thymine (T), followed by repair of the G/T mismatch by the G/T mismatch DNA glycosylase activity of methyl-binding protein 4 (MBD4) and DNA repair enzymes that remove and repair the abasic residue with cytosine (Bellacosa et al., 1999; Hendrich et al., 1999; Petronzelli et al., 2000).

DNA demethylation by direct removal of the methyl group from 5-mC is controversial. Although it has been claimed that the methyl DNA-binding protein 2 (MBD2) has such a '*bona fide*' DNA demethylase activity (Bhattacharya et al., 1999), efforts in several laboratories have failed to replicate these findings (Ng et al., 1999; Wade et al., 1999). On the contrary, evidence that MBD2 is a transcriptional repressor has steadily accumulated (Ng et al., 1999; Boeke et al., 2000; Tatematsu et al., 2000; Magdinier and Wolffe, 2001; Yu et al., 2001). Moreover, active DNA demethylation was not affected even in cells that are deficient in the MDB2 protein, suggesting that MBD2 does not have detectable DNA demethylase activity (Santos et al., 2002). Direct enzymatic demethylation has long been considered unlikely because it is highly unfavorable energetically, but it is not impossible (Cedar and Verdine, 1999). Hence the question of whether DNA is demethylated by direct removal of methyl groups remains open.

5-methylcytosine could be demethylated as well by an oxidative demethylation mechanism. The bacterial AlkB protein, a 2-oxoglutarate-dependent and iron-dependent oxygenase, recently has been shown to correct DNA with alkylated adenine and cytosine (Falnes et al., 2002; Trewick et al., 2002). AlkB can use both double- and single-

stranded DNA as substrate, but has a preference for single-stranded DNA. Homologs of the bacterial *AlkB* gene have been identified in human and other mammals.

Although the animal enzymes described above have been implicated in DNA demethylation, whether they participate in regulating DNA methylation *in vivo* still remains mysterious. Much less has been learned about the biochemistry of DNA demethylation in plants. Very recently, two *Arabidopsis* DNA glycosylase genes, *ROS1* and *DEMETER*, have been shown to encode DNA glycosylase. Whereas ROS1 was shown to be essential for preventing DNA methylation and maintaining active transcription of a stress-inducible gene in the presence of homologous transgenes, DEMETER does not appear to function in regulating DNA methylation at all (Fremont et al., 1997).

Sequence-Specific DNA Demethylation**.**

Several *trans*-acting factors have been shown to promote sequence-specific DNA demethylation, among them the ubiquitous transcription factor Sp1 (Silke et al., 1995), as well as NF-kappa B (Kirillov et al., 1996), the EBNA-1 protein (Hsieh, 1999b) and the glucocorticoid receptor (Grange et al., 2001; Thomassin et al., 2001). There is evidence that these proteins promote active DNA demethylation, rather than just interfering with maintenance DNA methylation, another mechanism by which DNA can be demethylated (Wilks et al., 1984; Matsuo et al., 1998; Hsieh, 1999b; Grange et al., 2001; Lin and Hsieh, 2001; Thomassin et al., 2001). The target specificity of DNA demethylation by transcription factors is likely to be determined by the DNA binding domain, but transcription activation may also play a role (Matsuo et al., 1998). Consistent with the

view that transcription activation is important for DNA demethylation, a recent study indicates that histone acetylation, which generally accompanies transcription, might mark the site for DNA demethylase action (Cervoni and Szyf, 2001). Studies on the NF-Kappa B factor indicate that additional proteins may be required as well (Matsuo et al., 1998). Interestingly, 5-methylcytosine DNA glycosylase, which can catalyze DNA demethylation through the DNA repair pathway, has been shown to interact with the retinoid or the estradiol receptors, both of which are known to promote sequence-specific DNA demethylation (Zhu et al., 2001; Jost et al., 2002). However, whether this is the case for other trans-acting factors still remains to be determined.

The Maize *Spm* Transposon is Epigenetically Regulated

Some of the earliest evidence for epigenetic regulation of gene expression came from McClintock's elegant genetic studies on the maize *Suppressor-mutator* (*Spm*) transposable element. McClintock reported that an active *Spm* could undergo what she termed a "change of phase", by which she meant a heritable, but reversible, genetic inactivation (McClintock, 1958, 1963). Peterson, who identified the element independently but named it as "Enhancer" (*En)*, made similar observations (Peterson, 1966). In further studies, McClintock identified epigenetic variants of the *Spm* transposon exhibiting different developmental patterns of activation and inactivation (McClintock, 1957-58, 1961, 1968, 1971). Importantly, she deduced that an active *Spm* element could transiently reactivate an inactive one when they were brought together by a genetic cross

(McClintock, 1958, 1959, 1971), suggesting that *Spm* encodes a trans-acting factor that can reactivate a genetically silent transposon.

More recent experiments showed that the genetic activity of *Spm* elements is correlated with the extent of DNA methylation of the transposon's 5' end comprising its promoter and the adjacent GC-rich sequence, termed the downstream control region (DCR) (Banks et al., 1988). Both are unmethylated in an active *Spm* element and become progressively methylated as the heritability of the inactive state increases (Banks and Fedoroff, 1989).

A moderately methylated inactive *Spm* transposon is readily reactivated by the genetic introduction of an active element (Fedoroff, 1989). Transient reactivation of an inactive element requires the simultaneous presence of an active element. A heavily methylated element, termed a "cryptic" *Spm*, is extremely resistant to either spontaneous or *Spm*-mediated reactivation. However, exposure to an active *Spm* element gradually promotes the heritable reactivation of both inactive and cryptic *Spm* transposons. Heritable activation of the latter is a slow process that occurs over several plant generations and requires continuous exposure to the active element (Fedoroff, 1989).

Activation of a genetically inactive transposon is correlated with the loss of DNA methylation in the promoter and the DCR sequences in maize (Banks et al., 1988). *Spm* undergoes genetic inactivation and methylation of the 5' terminus in transgenic tobacco, as it does in maize (Masson and Fedoroff, 1989). Further studies have identified TnpA, one of the two *Spm*-encoded proteins required for transposition, as the protein that promotes demethylation of the transposon's 5' terminal sequences (Schläppi et al., 1993; Schläppi et al., 1994).

TnpA is a sequence-specific DNA-binding protein. It recognizes slight variations of a 12-bp sequence (CCGACACTCTTA) present in 9 copies at the 5' or promoter end of the element and in 15 copies at the 3' end of the element (Masson et al., 1987; Gierl et al., 1988). The DNA binding domain is located between amino acids 120 and 420 at the N-terminus of the protein (Gierl et al., 1988; Trentmann et al., 1993). At its C-terminus is a leucine-zipper domain that is involved in protein dimerization (Trentmann et al., 1993). As homodimers TnpA can thus bind two cognate binding sites on one or two DNA fragments. Accordingly, TnpA could bring together the two ends of the *Spm* sequence, forming the transposition complex. Notably, the TnpA protein easily aggregates and can bind its cognate DNA only under reduced condition (Gierl et al., 1988; Trentmann et al., 1993).

The TnpA-binding motif is present in both direct and inverted orientations at both ends of the *Spm* transposon. The orientation and the number of binding site appear to have an effect on TnpA binding. It has been reported that TnpA forms a more stable complex with DNA fragments with two binding sites in a tail-to-tail orientation than in other orientations (Trentmann et al., 1993). Cooperative TnpA binding between TnpA binding sites is suggested by the observation that DNA fragments with an increasing number of binding site form complexes with TnpA at decreasing protein concentrations (Raina et al., 1998). These observations could explain why truncation in both ends of the *Spm* element delays the timing and frequency of transposition (Masson et al., 1987; Raina et al., 1998).

The *Spm* promoter sequence is co-extensive with the TnpA-binding region at the 5' end of the element and lacks a conventional TATA sequence (Raina et al., 1993).

Truncation from the 5’ end gradually reduces but never eliminates its promoter activity (Raina et al., 1993). Inactivation of the promoter entails methylation of both the promoter and the GC-rich DCR, which encodes the untranslated leader sequence of the single *Spm* transcription unit. TnpA-mediated promoter demethylation encompasses both the promoter and DCR sequence (Banks et al., 1988), although there is no TnpA-binding sites in the DCR (Raina et al., 1998). There is evidence that both the N-terminal DNA binding domain and C-terminal protein dimerization domain are essential for promoting demethylation of the *Spm* sequence (Schläppi et al., 1994; Raina et al., 1998). These observations imply that demethylation does not occur by a simple competition between TnpA and DNA methyltransferases for binding to the TnpA promoter (Schläppi et al., 1994; Raina et al., 1998).

References

Aapola, U., Kawasaki, K., Scott, H.S., Ollila, J., Vihinen, M., Heino, M., Shintani, A., Minoshima, S., Krohn, K., Antonarakis, S.E., Shimizu, N., Kudoh, J., and Peterson, P. (2000). Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. Genomics **65,** 293-298.

Aapola, U., Liiv, I., and Peterson, P. (2002). Imprinting regulator DNMT3L is a transcriptional repressor associated with histone deacetylase activity. Nucleic Acids Res. **30,** 3602-3608.

Aapola, U., Kawasaki, K., Scott, H.S., Ollila, J., Vihinen, M., Heino, M., Shintani, A., Minoshima, S., Krohn, K., Antonarakis, S.E., Shimizu, N., Kudoh, J., and Peterson, P. (2000). Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. Genomics **65,** 293-298.

Abeles, A., Brendler, T., and Austin, S. (1993). Evidence of two levels of control of P1 oriR and host oriC replication origins by DNA adenine methylation. J. Bacteriol. **175,** 7801-7807.

Adams, R.L. (1995). Eukaryotic DNA methyltransferases--structure and function. Bioessays **17,** 139-145.

Allaman-Pillet, N., Djemai, A., Bonny, C., and Schorderet, D.F. (1998). Methylation status of CpG sites and methyl-CpG binding proteins are involved in the promoter regulation of the mouse Xist gene. Gene Expr. **7,** 61-73.

Araujo, F.D., Croteau, S., Slack, A.D., Milutinovic, S., Bigey, P., Price, G.B., Zannis-Hajopoulos, M., and Szyf, M. (2001). The DNMT1 target recognition domain resides in the N terminus. J. Biol. Chem. **276,** 6930-6936.

Avramova, Z.V. (2002). Heterochromatin in animals and plants. Similarities and differences. Plant Physiol. **129,** 40-49.

Bachman, K.E., Rountree, M.R., and Baylin, S.B. (2001). Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. J. Biol. Chem. **276,** 32282-32287.

Ballestar, E., and Wolffe, A.P. (2001). Methyl-CpG-binding proteins. Targeting specific gene repression. Eur. J. Biochem. **268,** 1-6.

Ballestar, E., Pile, L.A., Wassarman, D.A., Wolffe, A.P., and Wade, P.A. (2001). A Drosophila MBD family member is a transcriptional corepressor associated with specific genes. Eur. J. Biochem. **268,** 5397-5406.

Banks, J.A., and Fedoroff, N. (1989). Patterns of developmental and heritable change in methylation of the Suppressor-mutator transposable element. Dev. Genet. **10,** 425-437.

Banks, J.A., Masson, P., and Fedoroff, N. (1988). Molecular mechanisms in the developmental regulation of the maize Suppressor-mutator transposable element. Genes Dev. **2,** 1364-1380.

Baroux, C., Spillane, C., and Grossniklaus, U. (2002). Genomic imprinting during seed development. Adv. Genet. **46,** 165-214.

Bartee, L., and Bender, J. (2001). Two Arabidopsis methylation-deficiency mutations confer only partial effects on a methylated endogenous gene family. Nucleic Acids Res. **29,** 2127-2134.

Bednarik, D.P., Duckett, C., Kim, S.U., Perez, V.L., Griffis, K., Guenthner, P.C., and Folks, T.M. (1991). DNA CpG methylation inhibits binding of NF-kappa B proteins to the HIV- 1 long terminal repeat cognate DNA motifs. New Biol. **3,** 969-976.

Bellacosa, A., Cicchillitti, L., Schepis, F., Riccio, A., Yeung, A.T., Matsumoto, Y., Golemis, E.A., Genuardi, M., and Neri, G. (1999). MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. Proc. Natl. Acad. Sci. USA **96,** 3969-3974.

Bestor, T.H., and Verdine, G.L. (1994). DNA methyltransferases. Curr. Opin. Cell. Biol. **6,** 380-389.

Bhattacharya, S.K., Ramchandani, S., Cervoni, N., and Szyf, M. (1999). A mammalian protein with specific demethylase activity for mCpG DNA. Nature **397,** 579-583.

Bingham, P.M. (1997). Cosuppression comes to the animals. Cell **90,** 385-387.

Boeke, J., Ammerpohl, O., Kegel, S., Moehren, U., and Renkawitz, R. (2000). The minimal repression domain of MBD2b overlaps with the methyl-CpG-binding domain and binds directly to Sin3A. J. Biol. Chem. **275,** 34963-34967.

Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B., and Bestor, T.H. (2001). Dnmt3L and the establishment of maternal genomic imprints. Science **294,** 2536-2539.

Bourc'his, D., and Bestor, T.H. (2002). Helicase homologues maintain cytosine methylation in plants and mammals. Bioessays **24,** 297-299.

Brenton, J.D., Drewell, R.A., Viville, S., Hilton, K.J., Barton, S.C., Ainscough, J.F., and Surani, M.A. (1999). A silencer element identified in Drosophila is required for imprinting of H19 reporter transgenes in mice. Proc. Natl. Acad. Sci. USA **96,** 9242-9247.

Brockdorff, N. (2002). X-chromosome inactivation: closing in on proteins that bind Xist RNA. Trends Genet. **18,** 352-358.

Brunk, B.P., Goldhamer, D.J., and Emerson, C.P., Jr. (1996). Regulated demethylation of the myoD distal enhancer during skeletal myogenesis. Dev. Biol. **177,** 490-503.

Burgers, W.A., Fuks, F., and Kouzarides, T. (2002). DNA methyltransferases get connected to chromatin. Trends Genet. **18,** 275-277.

Cao, X., Springer, N.M., Muszynski, M.G., Phillips, R.L., Kaeppler, S., and Jacobsen, S.E. (2000). Conserved plant genes with similarity to mammalian de novo DNA methyltransferases. Proc. Natl. Acad. Sci. USA **97,** 4979-4984.

Cao, X., and Jacobsen, S.E. (2002a). Role of the arabidopsis DRM methyltransferases in de novo DNA methylation and gene silencing. Curr. Biol. **12,** 1138-1144.

Cao, X., and Jacobsen, S.E. (2002b). Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. Proc. Natl. Acad. Sci. USA **99 Suppl 4,** 16491-16498.

Cedar, H., and Verdine, G.L. (1999). Gene expression. The amazing demethylase. Nature **397,** 568-569.

Cervoni, N., and Szyf, M. (2001). Demethylase activity is directed by histone acetylation. J. Biol. Chem. **276,** 40778-40787.

Chedin, F., Lieber, M.R., and Hsieh, C.L. (2002). The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. Proc. Natl. Acad. Sci. USA **99,** 16916-16921.

Chuang, L.S., Ian, H.I., Koh, T.W., Ng, H.H., Xu, G., and Li, B.F. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science **277,** 1996-2000.

Constancia, M., Pickard, B., Kelsey, G., and Reik, W. (1998). Imprinting mechanisms. Genome Res. **8,** 881-900.

Dean, W., Santos, F., Stojkovic, M., Zakhartchenko, V., Walter, J., Wolf, E., and Reik, W. (2001). Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. Proc. Natl. Acad. Sci. USA **98,** 13734-13738.

Dennis, K., Fan, T., Geiman, T., Yan, Q., and Muegge, K. (2001). Lsh, a member of the SNF2 family, is required for genome-wide methylation. Genes Dev. **15,** 2940-2944.

Deplus, R., Brenner, C., Burgers, W.A., Putmans, P., Kouzarides, T., de Launoit, Y., and Fuks, F. (2002). Dnmt3L is a transcriptional repressor that recruits histone deacetylase. Nucleic Acids Res. **30,** 3831-3838.

Dillon, N., and Festenstein, R. (2002). Unravelling heterochromatin: competition between positive and negative factors regulates accessibility. Trends Genet. **18,** 252-258.

Drewell, R.A., Brenton, J.D., Ainscough, J.F., Barton, S.C., Hilton, K.J., Arney, K.L., Dandolo, L., and Surani, M.A. (2000). Deletion of a silencer element disrupts H19 imprinting independently of a DNA methylation epigenetic switch. Development **127,** 3419-3428.

Dyda, F., Klein, D.C., and Hickman, A.B. (2000). GCN5-related N-acetyltransferases: a structural overview. Annu. Rev. Biophys. Biomol. Struct. **29,** 81-103.

Falnes, P.O., Johansen, R.F., and Seeberg, E. (2002). AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli. Nature **419,** 178-182.

Fedoroff, N. (1989). The heritable activation of cryptic Suppressor-mutator elements by an active element. Genetics **121,** 591-608.

Fedoroff, N., Schläppi, M., and Raina, R. (1995). Epigenetic regulation of the maize Spm transposon. Bioessays **17,** 291-297.

Fedoroff, N. (2000). Transposons and genome evolution in plants. Proc. Natl. Acad. Sci. USA **97,** 7002-7007.

Feng, Q., and Zhang, Y. (2001). The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes. Genes Dev. **15,** 827-832.

Finnegan, E.J., and Kovac, K.A. (2000). Plant DNA methyltransferases. Plant Mol. Biol. **43,** 189-201.

Fremont, M., Siegmann, M., Gaulis, S., Matthies, R., Hess, D., and Jost, J.P. (1997). Demethylation of DNA by purified chick embryo 5-methylcytosine-DNA glycosylase requires both protein and RNA. Nucleic Acids Res. **25,** 2375-2380.

Frit, P., Kwon, K., Coin, F., Auriol, J., Dubaele, S., Salles, B., and Egly, J.M. (2002). Transcriptional activators stimulate DNA repair. Mol. Cell **10,** 1391-1401.

Fujita, N., Takebayashi, S., Okumura, K., Kudo, S., Chiba, T., Saya, H., and Nakao, M. (1999). Methylation-mediated transcriptional silencing in euchromatin by methyl-CpG binding protein MBD1 isoforms [In Process Citation]. Mol. Cell. Biol. **19,** 6415-6426.

Fuks, F., Burgers, W.A., Godin, N., Kasai, M., and Kouzarides, T. (2001). Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. EMBO J. **20,** 2536-2544.

Garrick, D., Sutherland, H., Robertson, G., and Whitelaw, E. (1996). Variegated expression of a globin transgene correlates with chromatin accessibility but not methylation status. Nucleic Acids Res. **24,** 4902-4909.

Geiman, T.M., and Robertson, K.D. (2002). Chromatin remodeling, histone modifications, and DNA methylation-how does it all fit together? J. Cell Biochem. **87,** 117-125.

Gendrel, A.V., Lippman, Z., Yordan, C., Colot, V., and Martienssen, R.A. (2002). Dependence of heterochromatic histone H3 methylation patterns on the Arabidopsis gene DDM1. Science **297,** 1871-1873.

Gibbons, R.J., McDowell, T.L., Raman, S., O'Rourke, D.M., Garrick, D., Ayyub, H., and Higgs, D.R. (2000). Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. Nat. Genet. **24,** 368-371.

Gierl, A., Lutticke, S., and Saedler, H. (1988). TnpA product encoded by the transposable element En-1 of Zea mays is a DNA binding protein. EMBO J. **7,** 4045-4053.

Grange, T., Cappabianca, L., Flavin, M., Sassi, H., and Thomassin, H. (2001). In vivo analysis of the model tyrosine aminotransferase gene reveals multiple sequential steps in glucocorticoid receptor action. Oncogene **20,** 3028-3038.

Grini, P.E., Jurgens, G., and Hulskamp, M. (2002). Embryo and Endosperm Development Is Disrupted in the Female Gametophytic capulet Mutants of Arabidopsis. Genetics **162,** 1911-1925.

Hansen, R.S., Wijmenga, C., Luo, P., Stanek, A.M., Canfield, T.K., Weemaes, C.M., and Gartler, S.M. (1999). The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. Proc. Natl. Acad. Sci. USA **96,** 14412-14417.

Hata, K., Okano, M., Lei, H., and Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development **129,** 1983-1993.

Hendrich, B., and Bird, A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol. Cell. Biol. **18,** 6538-6547.

Hendrich, B., Hardeland, U., Ng, H.H., Jiricny, J., and Bird, A. (1999). The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. Nature **401,** 301-304.

Hendrich, B. (2000). Methylation moves into medicine. Curr. Biol. **10,** R60-63.

Hendrich, B., Guy, J., Ramsahoye, B., Wilson, V.A., and Bird, A. (2001). Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. Genes Dev. **15,** 710-723.

Hirochika, H., Okamoto, H., and Kakutani, T. (2000). Silencing of retrotransposons in arabidopsis and reactivation by the ddm1 mutation. Plant Cell **12,** 357-369.

Holler, M., Westin, G., Jiricny, J., and Schaffner, W. (1988). Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. Genes Dev. **2,** 1127-1135.

Hsieh, C.L. (1999a). In Vivo Activity of Murine De Novo Methyltransferases, Dnmt3a and Dnmt3b. Mol. Cell. Biol. **19,** 8211-8218.

Hsieh, C.L. (1999b). Evidence that protein binding specifies sites of DNA demethylation. Mol. Cell. Biol. **19,** 46-56.

Hsieh, C.L. (2000). Dynamics of DNA methylation pattern. Curr. Opin. Genet. Dev. **10,** 224-228.

Inamdar, N.M., Ehrlich, K.C., and Ehrlich, M. (1991). CpG methylation inhibits binding of several sequence-specific DNA- binding proteins from pea, wheat, soybean and cauliflower. Plant Mol. Biol. **17,** 111-123.

Jackson, J.P., Lindroth, A.M., Cao, X., and Jacobsen, S.E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature **416,** 556-560.

Jacobsen, S.E., and Meyerowitz, E.M. (1997). Hypermethylated SUPERMAN epigenetic alleles in arabidopsis. Science **277,** 1100-1103.

Jacobsen, S.E., Sakai, H., Finnegan, E.J., Cao, X., and Meyerowitz, E.M. (2000). Ectopic hypermethylation of flower-specific genes in Arabidopsis. Curr. Biol. **10,** 179-186.

Jeddeloh, J.A., and Richards, E.J. (1996). mCCG methylation in angiosperms. Plant J. **9,** 579-586.

Jeddeloh, J.A., Bender, J., and Richards, E.J. (1998). The DNA methylation locus DDM1 is required for maintenance of gene silencing in Arabidopsis. Genes Dev. **12,** 1714-1725.

Jeddeloh, J.A., Stokes, T.L., and Richards, E.J. (1999). Maintenance of genomic methylation requires a SWI2/SNF2-like protein. Nat. Genet. **22,** 94-97.

Johnson, L., Cao, X., and Jacobsen, S. (2002). Interplay between two epigenetic marks. DNA methylation and histone H3 lysine 9 methylation. Curr. Biol. **12,** 1360-1367.

Jost, J.P. (1993). Nuclear extracts of chicken embryos promote an active demethylation of DNA by excision repair of 5-methyldeoxycytidine. Proc Natl Acad Sci U S A **90,** 4684-4688.

Jost, J.P., and Jost, Y.C. (1995). Mechanism of active DNA demethylation during embryonic development and cellular differentiation in vertebrates. Gene **157,** 265-266.

Jost, J.P., Siegmann, M., Sun, L., and Leung, R. (1995). Mechanisms of DNA demethylation in chicken embryos. Purification and properties of a 5-methylcytosine-DNA glycosylase. J. Biol. Chem. **270,** 9734-9739.

Jost, J.P., Fremont, M., Siegmann, M., and Hofsteenge, J. (1997). The RNA moiety of chick embryo 5-methylcytosine- DNA glycosylase targets DNA demethylation. Nucleic Acids Res. **25,** 4545-4550.

Jost, J.P., Thiry, S., and Siegmann, M. (2002). Estradiol receptor potentiates, in vitro, the activity of 5-methylcytosine DNA glycosylase. FEBS Lett. **527,** 63-66.

Kakutani, T., Jeddeloh, J.A., Flowers, S.K., Munakata, K., and Richards, E.J. (1996). Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. Proc. Natl. Acad. Sci. USA **93,** 12406-12411.

Kass, S.U., Landsberger, N., and Wolffe, A.P. (1997). DNA methylation directs a time-dependent repression of transcription initiation. Curr. Biol. **7,** 157-165.

Kim, G.D., Ni, J., Kelesoglu, N., Roberts, R.J., and Pradhan, S. (2002). Co-operation and communication between the human maintenance and de novo DNA (cytosine-5) methyltransferases. EMBO J. **21,** 4183-4195.

Kirillov, A., Kistler, B., Mostoslavsky, R., Cedar, H., Wirth, T., and Bergman, Y. (1996). A role for nuclear NF-kappaB in B-cell-specific demethylation of the Igkappa locus. Nat. Genet. **13,** 435-441.

Kishimoto, N., Sakai, H., Jackson, J., Jacobsen, S.E., Meyerowitz, E.M., Dennis, E.S., and Finnegan, E.J. (2001). Site specificity of the Arabidopsis METI DNA methyltransferase demonstrated through hypermethylation of the superman locus. Plant Mol. Biol. **46,** 171-183.

Kochanek, S., Renz, D., and Doerfler, W. (1995). Transcriptional silencing of human Alu sequences and inhibition of protein binding in the box B regulatory elements by 5'-CG-3' methylation. FEBS Lett. **360,** 115-120.

Kouzminova, E., and Selker, E.U. (2001). dim-2 encodes a DNA methyltransferase responsible for all known cytosine methylation in Neurospora. EMBO J. **20,** 4309-4323.

Kress, C., Thomassin, H., and Grange, T. (2001). Local DNA demethylation in vertebrates: how could it be performed and targeted? FEBS Lett. **494,** 135-140.

Kudo, S. (1998). Methyl-CpG-binding protein MeCP2 represses Sp1-activated transcription of the human leukosialin gene when the promoter is methylated. Mol. Cell. Biol. **18,** 5492-5499.

Leegwater, P.A., De Abreu, R.A., and Albertioni, F. (1998). Analysis of DNA methylation of the 5' region of the deoxycytidine kinase gene in CCRF-CEM-sensitive and cladribine (CdA)- and 2-chloro-2'- arabino-fluoro-2'-deoxyadenosine (CAFdA)-resistant cells. Cancer Lett. **130,** 169-173.

Leonhardt, H., Page, A.W., Weier, H.U., and Bestor, T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell **71,** 865-873.

Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F., and Bird, A. (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell **69,** 905-914.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell **69,** 915-926.

Li, G., Hall, T.C., and Holmes-Davis, R. (2002). Plant chromatin: development and gene control. Bioessays **24,** 234-243.

Li, Y., Danzer, J.R., Alvarez, P., Belmont, A.S., and Wallrath, L.L. (2003). Effects of tethering HP1 to euchromatic regions of the Drosophila genome. Development **130,** 1817-1824.

Liang, G., Chan, M.F., Tomigahara, Y., Tsai, Y.C., Gonzales, F.A., Li, E., Laird, P.W., and Jones, P.A. (2002). Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. Mol. Cell. Biol. **22,** 480-491.

Lin, I.G., and Hsieh, C.L. (2001). Chromosomal DNA demethylation specified by protein binding. EMBO Rep. **2,** 108-112.

Lindroth, A.M., Cao, X., Jackson, J.P., Zilberman, D., McCallum, C.M., Henikoff, S., and Jacobsen, S.E. (2001). Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. Science **292,** 2077-2080.

Liu, Y., Oakeley, E.J., Sun, L., and Jost, J.P. (1998). Multiple domains are involved in the targeting of the mouse DNA methyltransferase to the DNA replication foci. Nucleic Acids Res. **26,** 1038-1045.

Lyko, F., Ramsahoye, B.H., and Jaenisch, R. (2000). DNA methylation in Drosophila melanogaster. Nature **408,** 538-540.

Magdinier, F., and Wolffe, A.P. (2001). Selective association of the methyl-CpG binding protein MBD2 with the silent p14/p16 locus in human neoplasia. Proc. Natl. Acad. Sci. USA **98,** 4990-4995.

Maloisel, L., and Rossignol, J.L. (1998). Suppression of crossing-over by DNA methylation in Ascobolus. Genes Dev. **12,** 1381-1389.

Marmorstein, R. (2001). Structure and function of histone acetyltransferases. Cell Mol. Life Sci. **58,** 693-703.

Masson, P., Surosky, R., Kingsbury, J.A., and Fedoroff, N.V. (1987). Genetic and molecular analysis of the *Spm*-dependent *a-m2* alleles of the maize *a* locus. Genetics **117,** 117-137.

Masson, P., and Fedoroff, N.V. (1989). Mobility of the maize suppressor-mutator element in transgenic tobacco cells. Proc. Natl. Acad. Sci. USA **86,** 2219-2223.

Matsuo, K., Silke, J., Georgiev, O., Marti, P., Giovannini, N., and Rungger, D. (1998). An embryonic demethylation mechanism involving binding of transcription factors to replicating DNA. EMBO J. **17,** 1446-1453.

Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haaf, T. (2000). Demethylation of the zygotic paternal genome. Nature **403,** 501-502.

McClintock, B. (1957). Genetic and cytological studies of maize. Carnegie Inst. Wash. Yr. Bk. **56,** 393-401.

McClintock, B. (1958). The suppressor-mutator system of control of gene action in maize. Carnegie Inst. Wash. Yr. Bk. **57,** 415-429.

McClintock, B. (1959). Genetic and cytological studies of maize. Carnegie Inst. Wash. Yr. Bk. **58,** 452-456.

McClintock, B. (1961). Further studies of the suppressor-mutator system of control of gene action in maize. Carnegie Inst. Wash. Yr. Bk. **60,** 469-476.

McClintock, B. (1963). Further studies of gene-control systems in maize. Carnegie Inst. Wash. Yr. Bk. **62,** 486-493.

McClintock, B. (1968). The states of a gene locus in maize. Carnegie Inst. Wash. Yr. Bk. **66,** 20-28.

McClintock, B. (1971). The contribution of one component of a control system to versatility of gene expression. Carnegie Inst. Wash. Yr. Bk. **70,** 5-17.

Meehan, R.R., Lewis, J.D., McKay, S., Kleiner, E.L., and Bird, A.P. (1989). Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. Cell **58,** 499-507.

Migeon, B.R., Holland, M.M., Driscoll, D.J., and Robinson, J.C. (1991). Programmed demethylation in CpG islands during human fetal development. Somat. Cell. Mol. Genet. **17,** 159-168.

Modrich, P. (1989). Methyl-directed DNA mismatch correction. J. Biol. Chem. **264,** 6597-6600.

Mostoslavsky, R., Singh, N., Kirillov, A., Pelanda, R., Cedar, H., Chess, A., and Bergman, Y. (1998). Kappa chain monoallelic demethylation and the establishment of allelic exclusion. Genes Dev. **12,** 1801-1811.

Muller, W.G., Walker, D., Hager, G.L., and McNally, J.G. (2001). Large-scale chromatin decondensation and recondensation regulated by transcription from a natural promoter. J. Cell Biol. **154,** 33-48.

Nan, X., Campoy, F.J., and Bird, A. (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell **88,** 471-481.

Ng, H.H., Zhang, Y., Hendrich, B., Johnson, C.A., Turner, B.M., Erdjument-Bromage, H., Tempst, P., Reinberg, D., and Bird, A. (1999). MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. Nat. Genet. **23,** 58-61.

Ng, H.H., Jeppesen, P., and Bird, A. (2000). Active repression of methylated genes by the chromosomal protein MBD1. Mol. Cell. Biol. **20,** 1394-1406.

Nye, A.C., Rajendran, R.R., Stenoien, D.L., Mancini, M.A., Katzenellenbogen, B.S., and Belmont, A.S. (2002). Alteration of large-scale chromatin structure by estrogen receptor. Mol. Cell. Biol. **22,** 3437-3449.

Oakeley, E.J., Podesta, A., and Jost, J.P. (1997). Developmental changes in DNA methylation of the two tobacco pollen nuclei during maturation. Proc. Natl. Acad. Sci. USA **94,** 11721-11725.

Okano, M., Xie, S., and Li, E. (1998). Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. Nucleic Acids Res. **26,** 2536-2540.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell **99,** 247-257.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. Curr. Biol. **10,** 475-478.

Palmer, B.R., and Marinus, M.G. (1994). The dam and dcm strains of Escherichia coli--a review. Gene **143,** 1-12.

Paro, R., Strutt, H., and Cavalli, G. (1998). Heritable chromatin states induced by the Polycomb and trithorax group genes. Novartis Found. Symp. **214,** 51-61; discussion 61-56, 104-113.

Peterson, P.A. (1966). Phase variation of regulatory elements in maize. Genetics **54,** 249-266.

Petronzelli, F., Riccio, A., Markham, G.D., Seeholzer, S.H., Stoerker, J., Genuardi, M., Yeung, A.T., Matsumoto, Y., and Bellacosa, A. (2000). Biphasic kinetics of the human DNA repair protein MED1 (MBD4), a mismatch-specific DNA N-glycosylase. J. Biol. Chem. **275,** 32422-32429.

Probst, A.V., Fransz, P.F., Paszkowski, J., and Scheid, O.M. (2003). Two means of transcriptional reactivation within heterochromatin. Plant J. **33,** 743-749.

Raina, R., Cook, D., and Fedoroff, N. (1993). Maize Spm transposable element has an enhancer-insensitive promoter. Proc. Natl. Acad. Sci. USA **90,** 6355-6359.

Raina, R., Schläppi, M., Karunanandaa, B., Elhofy, A., and Fedoroff, N. (1998). Concerted formation of macromolecular Suppressor-mutator transposition complexes. Proc. Natl. Acad. Sci. USA **95,** 8526-8531.

Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P., and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc. Natl. Acad. Sci. USA **97,** 5237-5242.

Razin, A. (1998). CpG methylation, chromatin structure and gene silencing-a three-way connection [In Process Citation]. EMBO J. **17,** 4905-4908.

Reik, W., and Walter, J. (1998). Imprinting mechanisms in mammals. Curr. Opin. Genet. Dev. **8,** 154-164.

Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. Science **293,** 1089-1093.

Rein, T., Zorbas, H., and DePamphilis, M.L. (1997). Active mammalian replication origins are associated with a high-density cluster of mCpG dinucleotides. Mol. Cell. Biol. **17,** 416-426.

Rhee, I., Jair, K.W., Yen, R.W., Lengauer, C., Herman, J.G., Kinzler, K.W., Vogelstein, B., Baylin, S.B., and Schuebel, K.E. (2000). CpG methylation is maintained in human cancer cells lacking DNMT1. Nature **404,** 1003-1007.

Richards, E.J., and Elgin, S.C. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. Cell **108,** 489-500.

Robert, M.F., Morin, S., Beaulieu, N., Gauthier, F., Chute, I.C., Barsalou, A., and MacLeod, A.R. (2003). DNMT1 is required to maintain CpG methylation and aberrant gene silencing in human cancer cells. Nat. Genet. **33,** 61-65.

Robertson, K.D. (2002). DNA methylation and chromatin - unraveling the tangled web. Oncogene **21,** 5361-5379.

Roder, K., Hung, M.S., Lee, T.L., Lin, T.Y., Xiao, H., Isobe, K.I., Juang, J.L., and Shen, C.J. (2000). Transcriptional repression by Drosophila methyl-CpG-binding proteins. Mol. Cell. Biol. **20,** 7401-7409.

Rossi, V., Motto, M., and Pellegrini, L. (1997). Analysis of the methylation pattern of the maize opaque-2 (O2) promoter and in vitro binding studies indicate that the O2 B-Zip protein and other endosperm factors can bind to methylated target sequences. J. Biol. Chem. **272,** 13758-13765.

Samac, S., Rice, J.C., and Ehrlich, M. (1998). Analysis of methylation in the 5' region of the human alpha- galactosidase A gene containing a binding site for methylated DNA- binding protein/RFX1-4. Biol. Chem. **379,** 541-544.

Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev. Biol. **241,** 172-182.

Scharer, O.D., and Jiricny, J. (2001). Recent progress in the biology, chemistry and structural biology of DNA glycosylases. Bioessays **23,** 270-281.

Schläppi, M., Smith, D., and Fedoroff, N. (1993). TnpA trans-activates methylated maize Suppressor-mutator transposable elements in transgenic tobacco. Genetics **133,** 1009-1021.

Schläppi, M., Raina, R., and Fedoroff, N. (1994). Epigenetic regulation of the maize Spm transposable element: novel activation of a methylated promoter by TnpA. Cell **77,** 427-437.

Schmutte, C., and Jones, P.A. (1998). Involvement of DNA methylation in human carcinogenesis. Biol. Chem. **379,** 377-388.

Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., and Reuter, G. (2002). Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. EMBO J. **21,** 1121-1131.

Selker, E.U., Freitag, M., Kothe, G.O., Margolin, B.S., Rountree, M.R., Allis, C.D., and Tamaru, H. (2002). Induction and maintenance of nonsymmetrical DNA methylation in Neurospora. Proc. Natl. Acad. Sci. USA **99 Suppl 4,** 16485-16490.

Sheldon, C.C., Burn, J.E., Perez, P.P., Metzger, J., Edwards, J.A., Peacock, W.J., and Dennis, E.S. (1999). The FLF MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. Plant Cell **11,** 445-458.

Silke, J., Rother, K.I., Georgiev, O., Schaffner, W., and Matsuo, K. (1995). Complex demethylation patterns at Sp1 binding sites in F9 embryonal carcinoma cells. FEBS Lett. **370,** 170-174.

Soppe, W.J., Jacobsen, S.E., Alonso-Blanco, C., Jackson, J.P., Kakutani, T., Koornneef, M., and Peeters, A.J. (2000). The late flowering phenotype of fwa mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. Mol. Cell **6,** 791-802.

Soppe, W.J., Jasencakova, Z., Houben, A., Kakutani, T., Meister, A., Huang, M.S., Jacobsen, S.E., Schubert, I., and Fransz, P.F. (2002). DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis. Embo J **21,** 6549-6559.

Tamaru, H., and Selker, E.U. (2001). A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. Nature **414,** 277-283.

Tatematsu, K.I., Yamazaki, T., and Ishikawa, F. (2000). MBD2-MBD3 complex binds to hemi-methylated DNA and forms a complex containing DNMT1 at the replication foci in late S phase. Genes Cells **5,** 677-688.

ten Lohuis, M., Muller, A., Heidmann, I., Niedenhof, I., and Meyer, P. (1995). A repetitive DNA fragment carrying a hot spot for de novo DNA methylation enhances expression variegation in tobacco and petunia. Plant J. **8,** 919-932.

Thomassin, H., Flavin, M., Espinas, M.L., and Grange, T. (2001). Glucocorticoid-induced DNA demethylation and gene memory during development. EMBO J. **20,** 1974-1983.

Trentmann, S.M., Saedler, H., and Gierl, A. (1993). The transposable element En/Spm-encoded TNPA protein contains a DNA binding and a dimerization domain. Mol. Gen. Genet. **238,** 201-208.

Trewick, S.C., Henshaw, T.F., Hausinger, R.P., Lindahl, T., and Sedgwick, B. (2002). Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. Nature **419,** 174-178.

Tumbar, T., Sudlow, G., and Belmont, A.S. (1999). Large-scale chromatin unfolding and remodeling induced by VP16 acidic activation domain. J. Cell Biol. **145,** 1341-1354.

Vairapandi, M., and Duker, N.J. (1993). Enzymic removal of 5-methylcytosine from DNA by a human DNA-glycosylase. Nucleic Acids Res. **21,** 5323-5327.

Vairapandi, M., and Duker, N.J. (1996). Partial purification and characterization of human 5-methylcytosine-DNA glycosylase. Oncogene **13,** 933-938.

Vairapandi, M., Liebermann, D.A., Hoffman, B., and Duker, N.J. (2000). Human DNA-demethylating activity: A glycosylase associated with RNA and PCNA. J. Cell Biochem. **79,** 249-260.

Van den Veyver, I.B., and Zoghbi, H.Y. (2000). Methyl-CpG-binding protein 2 mutations in Rett syndrome. Curr. Opin. Genet. Dev. **10,** 275-279.

Varga-Weisz, P. (2001). ATP-dependent chromatin remodeling factors: nucleosome shufflers with many missions. Oncogene **20,** 3076-3085.

Vongs, A., Kakutani, T., Martienssen, R.A., and Richards, E.J. (1993). Arabidopsis thaliana DNA methylation mutants. Science **260,** 1926-1928.

Wade, P.A., Gegonne, A., Jones, P.L., Ballestar, E., Aubry, F., and Wolffe, A.P. (1999). Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. Nat. Genet. **23,** 62-66.

Walsh, C.P., and Bestor, T.H. (1999). Cytosine methylation and mammalian development. Genes Dev. **13,** 26-34.

Wilks, A., Seldran, M., and Jost, J.P. (1984). An estrogen-dependent demethylation at the 5' end of the chicken vitellogenin gene is independent of DNA synthesis. Nucleic Acids Res. **12,** 1163-1177.

Yoder, J.A., and Bestor, T.H. (1996). Genetic analysis of genomic methylation patterns in plants and mammals. Biol. Chem. **377,** 605-610.

Yoder, J.A., Walsh, C.P., and Bestor, T.H. (1997). Cytosine methylation and the ecology of intragenomic parasites. Trends Genet. **13,** 335-340.

Yoder, J.A., and Bestor, T.H. (1998). A candidate mammalian DNA methyltransferase related to pmt1p of fission yeast. Hum. Mol. Genet. **7,** 279-284.

Yokochi, T., and Robertson, K.D. (2002). Preferential methylation of unmethylated DNA by Mammalian de novo DNA methyltransferase Dnmt3a. J. Biol. Chem. **277,** 11735-11745.

Yu, F., Zingler, N., Schumann, G., and Stratling, W.H. (2001). Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. Nucleic Acids Res. **29,** 4493-4501.

Zhu, B., Benjamin, D., Zheng, Y., Angliker, H., Thiry, S., Siegmann, M., and Jost, J.P. (2001). Overexpression of 5-methylcytosine DNA glycosylase in human embryonic kidney cells EcR293 demethylates the promoter of a hormone-regulated reporter gene. Proc. Natl. Acad. Sci. USA **98,** 5031-5036.

Zilberman, D., Cao, X., and Jacobsen, S.E. (2003). ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science **299,** 716-719.

Chapter 2

TNPA PROMOTES RAPID DEMETHYLATION OF THE *SPM* SEQUENCE

Introduction

Previous studies have identified TnpA, one of the transposase proteins encoded by the *Spm* transposon, as the trans-acting factor that mediates both genetic activation and demethylation of the *Spm* sequence (Banks et al., 1988; Banks and Fedoroff, 1989; Fedoroff, 1989; Schläppi et al., 1993; Schläppi et al., 1994). However, the mechanism by which TnpA promotes DNA demethylation is unknown.

One possible mechanism by which TnpA promotes demethylation of the *Spm* sequence is blockage of DNA remethylation, generally termed passive DNA demethylation. Multiple copies of the consensus TnpA binding site have been identified in the *Spm* promoter sequence (Raina et al., 1993), and there is evidence that the DNA binding domain is indispensable for TnpA's ability to promote *Spm* demethylation (Schläppi et al., 1994). Conceivably, during DNA replication TnpA could interfere with the binding of maintenance DNA methyltransferases to the strand that is newly synthesized and thus unmethylated, causing loss of DNA methylation.

Active DNA demethylation, or enzymatic removal of methylated cytosine, has been reported in plants as well as in animals (Aoyama and Chua, 1997; Oakeley et al., 1997;

Hsieh, 1999; Oswald et al., 2000; Dean et al., 2001; Reik et al., 2001), although it does not seem to occur in some organisms, such as *Xenopus laevis* (Stancheva et al., 2002). Thus, it is possible that an active process could be involved in TnpA-mediated DNA demethylation. Since a passive mechanism is solely dependent on the number of cell cycles, whereas an active process is not, these two possible mechanisms can be distinguished by examining the rate of DNA demethylation relative to the DNA replication rate. In order to study the mechanism underlying TnpA-mediated DNA demethylation, I first sought to develop a novel assay system that would meet the following criteria. First, the assay should be compatible with ongoing DNA replication, as DNA demethylation may be achieved through a passive mechanism. Second, TnpA expression should be controllable so that initiation of DNA demethylation can be accurately timed. Finally, the *Spm* sequence should remain methylated before interaction with TnpA. To this end, I used an inducible promoter to control the expression of the TnpA protein in existing transgenic tobacco plants that carry a methylated *Spm* sequence. I also prepared hemimethylated DNA containing the *Spm* sequence and examined its sensitivity to methylation-sensitive enzymes SalI and Eco109I, which were used for assaying DNA methylation in the *Spm* sequence. The results show that TnpA-mediated DNA demethylation is rapid, suggesting the involvement of an active DNA demethylation mechanism.

Materials and Methods

Plant Materials and Treatments

Nicotiana tabacum L. cv Petite Havana line SR1 plants were used in this study. Transgenic plants were obtained by the *Agrobacterium* leaf-disc transformation method. Where indicated, callus cultures were established from leaves of month-old F1 plants derived from crosses of *Spm-LUC* plants by plants carrying the dexamethasone-inducible construct for FLAG-TnpA or FLAG-$TnpA_{420}$. Callus was maintained on agar medium containing 1X Murashige minimal organics medium (MS) (Life Technologies), 0.8% agar, 3% sucrose, 1 mg/L α-naphthaleneacetic acid (NAA) (Life Technologies) and 0.1 mg/L 6-benzylaminopurine (BA) (Sigma-Aldrich). Expression of TnpA and $TnpA_{420}$ was induced in callus by incubating the callus overnight in liquid growth medium containing 10 μM dexamethasone, then transferring it to an agar plate containing 1 mg/L NAA and 0.1 mg/L BA. For continuous TnpA induction in callus, 1 ml liquid growth medium containing 10 μM dexamethasone was added to the agar plate every 2 days. A similar regime was followed for other chemical treatments.

The cell doubling time was determined by weighing callus, transferring it onto fresh callus maintenance medium, and then reweighing it daily.

Plasmid Constructs

To clone the *FLAG-TnpA* cDNA, a 750-bp fragment containing the FLAG epitope sequence at the 5' end was PCR amplified from pRR483 (Raina et al., 1998) using primer pairs FLAG-1 (5'ATCTTATGGACTACAAGGACGAC3') and FLAG-2

(5'ATACATCATACCCTTTACAGC3'). The Vent® DNA polymerase (New England Biolabs) was used for the PCR reaction so that the amplified product was blunt-ended and ready for subsequent cloning. After restriction with SplI, the PCR products were resolved on a low melting-point agarose gel and the longer fragment (626 bp) was recovered and purified. This DNA fragment and a SplI-SpeI fragment from pRR483 were then cloned by a tripartite ligation into pBluescript II KS (+) (Stratagene) that was cut with EcoRV and SpeI. Clones with the full-length *FLAG-TnpA* cDNA insert (pFLAG-TnpA) were identified by PCR using the primer pair FLAG-1 and FLAG-2 and were further confirmed by sequencing.

pFLAG-$TnpA_{420}$ were cloned by ligating the XhoI to SplI fragment from pFLAG-TnpA and the SplI to SalI fragment from pMS198 (Schläppi et al., 1994) (the SalI site was filled in using the DNA polymerase Klenow fragment) together into pBluescript II KS (+) that was already cut with XhoI and EcoRV. pFLAG-$TnpA_{420}$ encodes the N-terminal 420 amino acids of TnpA, as a result of a nonsense mutation introduced by site-directed mutagenesis at the 3' end of the cDNA for TnpA (Schläppi et al., 1996).

For inducible protein expression, the XhoI to SpeI fragment from pFLAG-TnpA and pFLAG-$TnpA_{420}$ were cloned into pTA7002 (Aoyama and Chua, 1997), giving rise to pTA-FLAG-TnpA and pTA-FLAG-$TnpA_{420}$**,** respectively. For constitutive protein expression, the cDNAs for *FLAG-TnpA* was first cloned as an XhoI to XbaI fragment from pFLAG-TnpA into pAVA120 (von Arnim et al., 1998), yielding pAVA-FLAG-TnpA. The expression cassette was then subcloned as a HindIII fragment into pCGN1549 (McBride and Summerfelt, 1990), giving rise to pCGN-35S-FLAG-TnpA.

Spm Methylation Analysis

Genomic DNA was prepared using a CTAB protocol as described (http://www.arabidopsis.org/info/Protocols_Mundy2.html). Six μg of genomic DNA were digested in a 40 μl reaction for 16 hrs with 20 units each of EcoRV and the methylation-sensitive enzymes EcoO109I or SalI (New England BioLabs). The digest was resolved on a 1.25% Seakem LE agarose gel (BMA) and transferred to Hybond+ (Amersham) nylon membrane. A BamHI-EcoRV fragment from pDC107, part of the luciferase gene, was used as the probe for subsequent Southern analysis. The signal was quantified using a Phosphorimager, and the extent of DNA methylation is represented as the percentage of the total signal in the methylated band.

Preparation and Restriction Analysis of Hemimethylated DNA

A 647-bp DNA fragment containing the *Spm* promoter and DCR, amplified by PCR from pDC105 (Raina et al., 1993) using primers KS and SK, was used for determining the sensitivity of hemimethylated DNA to EcoO109I and SalI. 0.2 μg DNA was digested for 1 hr in a 20 μl reaction with 5 units of either enzyme and the digest was resolved on a 1.8% MetaPhor® agarose gel (BMA). The amount of DNA cleaved was quantified using the AlphaImager™ 2200 gel documentation and analysis system (Alpha Innotech Corp.).

Methylated DNA was prepared either by using SssI methylase (New England BioLabs) or by incorporating 5-methyl-dCTP (Roche Applied Science) during PCR amplification. Hemimethylated DNA was prepared either by annealing complementary unmethylated and methylated single-stranded (ss) DNA or by incorporating 5-methyl-dCTP into the complementary strand of an unmethylated ssDNA template. Less than

fully methylated DNA was prepared by synthesizing DNA using a mixture of 5-methyl-dCTP and dCTP. The composition of DNA was calculated based on the random probability of dCTP and 5-methyl-dCTP incorporation from a substrate mixture. For example, at a 5-methyl-dCTP:dCTP ratio of 2:1, the probability of 5-methyl-dCTP incorporation at each cytosine residue is 2/3; thus, for the EcoO109I cleavage site in the *Spm-Luc* transgene, GGGTCCC, hemimethylated DNA with all cytosines methylated would comprise 2/3 x 2/3 x 2/3 =8/27 of the total DNA, whereas those with 2 or 1 cytosines methylated would be 2/3 x 2/3 x 1/3 x 3 =12/27 and 2/3 x 1/3 x 1/3 x 3 =6/27 respectively. Likewise, unmethylated DNA would be 1/3 x 1/3 x 1/3 =1/27.

A biotin-streptavidin technique was used to prepare ssDNA. Briefly, dsDNA was PCR-amplified using one primer with 5' biotin modification and the other without modification. After the dsDNA was bound to streptavidin-paramagnetic beads and washed thoroughly following the vendor's protocol (CPG Inc.), the unmodified ssDNA was released by incubation in 0.1N NaOH for 8 min. The eluate was neutralized by adding 0.1 volume 1N HCl and 1M Tris-HCl, pH8.0, and the ssDNA was further desalted using the Qiagen PCR purification kit.

To examine enzyme sensitivity under the digestion conditions used for analysis of genomic DNA, approximately 1 ng ^{32}P-labeled hemimethylated DNA was added to 6 μg tobacco genomic DNA and digested with 20 units of SalI in a 40 μl reaction for 16 hours. 1/5 of the digest was resolved on a 2% agarose gel. After drying, the gel was exposed to an X-ray film for visualization. DNA was labeled by preincubation with Klenow and dTTP at 30° C for 15 min, followed by addition of dATP, dGTP and ^{32}P-dCTP (10 μCi in

a 20 µl reaction) and an additional 15-min incubation. All DNA preparations were purified using Qiagen columns.

Results

Inducible TnpA-Mediated Demethylation of the *Spm* Sequence

To study the mechanism of TnpA-mediated *Spm* demethylation, the early events in TnpA-DNA interactions need to be examined. The *TnpA* cDNA was therefore expressed from a glucocorticoid-inducible promoter (Aoyama and Chua, 1997) in transgenic

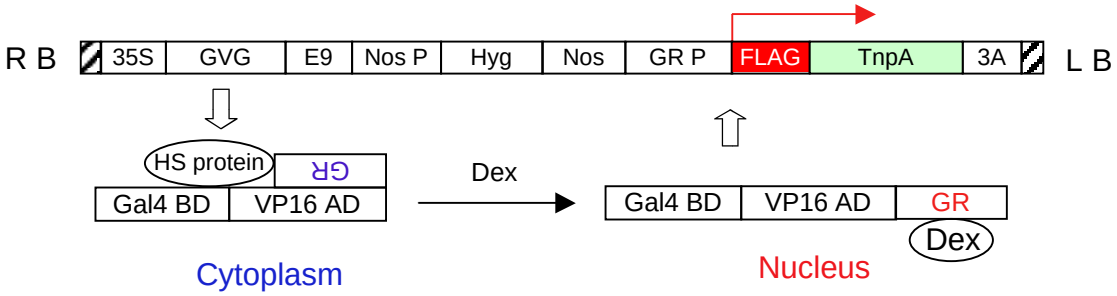


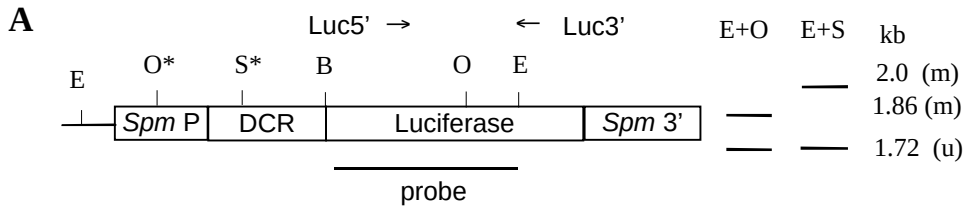
Figure 1. T-DNA region of the binary vector used for inducible expression of FLAG-TnpA, deletion derivatives, and fusion proteins. FLAG, the FLAG polypeptide sequence; Dex, dexamethasone, GVG, a hybrid transcription factor consisting of the Gal4 binding domain (Gal4 BD), the viral VP16 activation domain (VP16 AD) and the glucocorticoid receptor domain (GR). GVG is constitutively expressed under the CaMV 35S promoter (35S), however, in the absence of glucocorticoid ligands it is bound by a heat shock protein (HS protein) and remains inactive in the cytoplasm . Upon binding of dexamethasone (Dex), GVG is activated and transported into the nucleus, where it promotes transcription from the GVG responsive promoter (GR P). Nos P, the nopaline synthase promoter; hyg, the bacterial *HPTII* gene, conferring hygromycin resistance; Nos, the nopaline synthase poly (A) addition sequence; E9 and 3A, the poly (A) addition sequences of pea ribulose bisphosphate carboxylase small subunits rcbS-E9 and rcbS-3A, respectively.

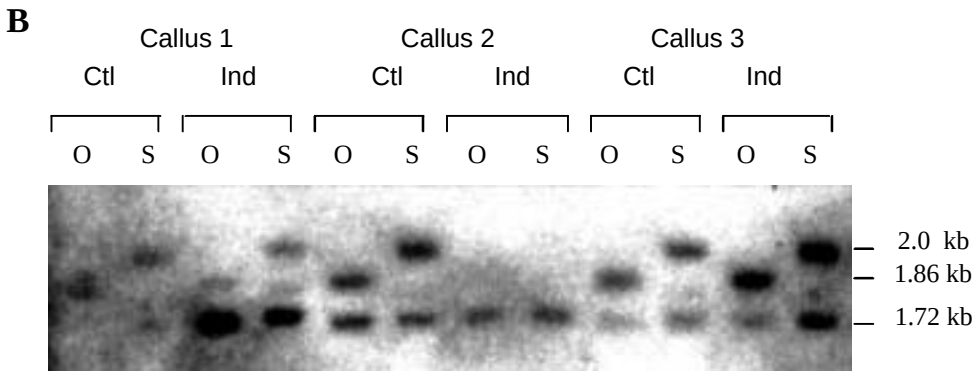
tobacco lines containing a methylated *Spm* sequence in a luciferase reporter construct designated *Spm-LUC* (Schläppi et al., 1994) (Fig. 1 and 2A). To facilitate protein monitoring, a FLAG-epitope tag was fused in-frame to TnpA's N-terminus based on the previous observation that an N-terminal 120-amino acid deletion did not affect its ability to mediate *Spm* demethylation (Schläppi et al., 1994). To determine whether the FLAG-tagged TnpA protein promotes demethylation, a plant expressing the fusion protein from a Cauliflower Mosaic Virus (CaMV) 35S promoter was crossed to a plant containing a heavily methylated *Spm-LUC* transgene and was monitored for the methylation of the *Spm* sequence. The *Spm* sequence was demethylated in all 6 progeny plants examined (not shown), indicating that the FLAG-TnpA retains its demethylation activity.

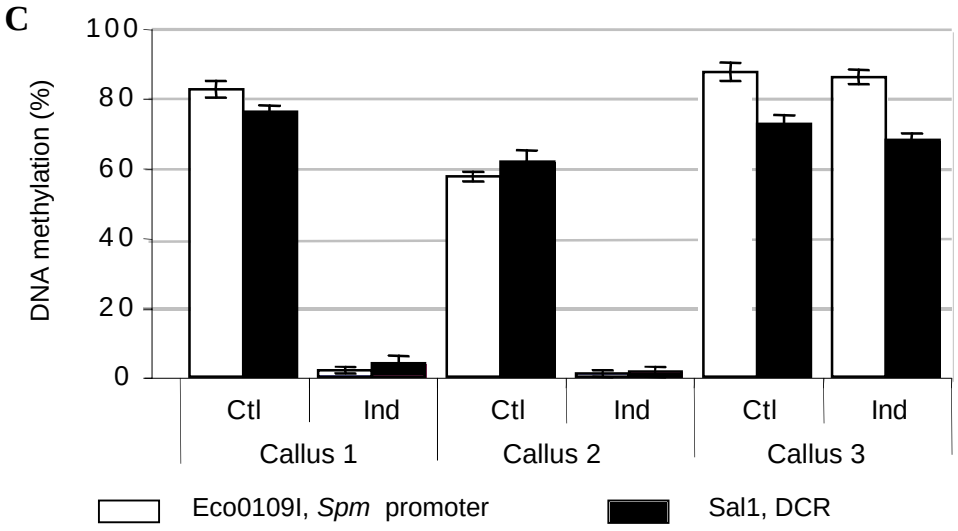
The inducible TnpA construct, designated pTA-FLAG-TnpA, was transformed into an *Spm-LUC* transgenic tobacco line, SR1-83, in which the *Spm* sequence showed about 50% methylation. The *Spm* sequence was demethylated in the resulting transgenic calli and methylation was restored to a very low level even in the second generation of regenerated plants. I therefore screened many independent lines of transgenic plants that contain only the *Spm-LUC* construct and selected a line in which the *Spm* sequence was the most heavily methylated (SR1-1, about 90% and 70% methylation at the sites monitored in the promoter and DCR, respectively), and introduced the dexamethasone-inducible TnpA construct by a genetic cross with a transgenic pTA-FLAG-TnpA plant. The *Spm* promoter was methylated in leaves of the F1 plants, but methylation was unaffected by dexamethasone.

Because cell division has ceased in leaves and continuing cell division may be necessary for demethylation, callus cultures were established from leaves of 1 month-old

F1 progeny plants carrying both pTA-FLAG-TnpA and *Spm-LUC* constructs. These also maintained a high level of *Spm* methylation in the absence of dexamethasone, but exhibited dexamethasone-inducible demethylation. Figure 2A shows a diagram of the *Spm* promoter region in the *Spm-LUC* reporter gene construct that was tested for methylation and the location of the methylation-sensitive restriction endonuclease sites used to monitor its methylation status, as well as a representation of the expected band sizes for the methylated and unmethylated fragments. The two monitored sites, EcoO109I and Sal1, are respectively located in the *Spm* promoter and the DCR. Both sites in the *Spm* sequence became fully sensitive to restriction in some calli but not in others after 3 weeks of induction with dexamethasone. Fig. 2B shows a Southern blot of genomic DNA digested with EcoRV and either EcoO109I or SalI. In Fig. 2C the data obtained from replicates of assays on the same calli are expressed as the percentage of the total signal in the band corresponding to the methylated fragment (% methylation). The extent of *Spm* methylation observed in the absence of dexamethasone shows some variation among independently derived callus lines (Fig. 2C). This variation may be attributable to small differences in the background expression levels of the *TnpA* transgene, since the prolonged maintenance of callus lines containing the gene results in the gradual disappearance of methylation. Sibling lines lacking the *TnpA* gene maintain *Spm* methylation levels. Nonetheless, as evident in Fig. 2C, replicate measurements made on an individual callus line at a similar growth stage, both with and without dexamethasone induction, are highly reproducible.


A
Luc5' →
← Luc3'
E+O
E+S
kb
E
O*
S*
B
O
E
Spm P
DCR
Luciferase
Spm 3'
2.0 (m)
1.86 (m)
1.72 (u)
probe


B
Callus 1
Callus 2
Callus 3
Ctl
Ind
O
S
2.0 kb
1.86 kb
1.72 kb


C
DNA methylation (%)
100
80
60
40
20
0
Ctl
Ind
Callus 1
Callus 2
Callus 3
Eco0109I, Spm promoter
Sal1, DCR

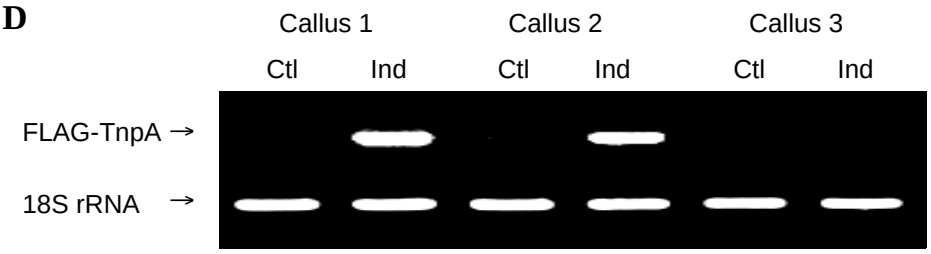

D
Callus 1
Callus 2
Callus 3
Ctl
Ind
FLAG-TnpA →
18S rRNA →

Figure 2. Inducible TnpA-mediated *Spm* demethylation. (A) The *Spm-LUC* reporter gene construct. *Spm* P, the *Spm* promoter; *Spm* 3', the sequence of 3' end of *Spm*. Luc 5' and Luc 3' are primers for RT-PCR analysis. The diagram shows the location in the 5' end of *Spm* of the restriction sites used to monitor its methylation status. The diagram on the right shows the sizes of the genomic DNA fragments expected using the *LUC* gene probe. EcoO109I* and SalI* are methylation-sensitive sites in *Spm* P and DCR; m, methylated; u, unmethylated; E, EcoRV; O, EcoO109I; S, SalI; B, BamHI. (B) A Southern blot of genomic DNA extracted from transgenic tobacco calli containing the constructs shown in Fig. 1 and 2A before (control, ctl) and after induction (induced, ind) with 10 μM dexamethasone for 20 days. (C) Data obtained from replicate experiments on the same callus lines used in B were quantified using a Phosphorimager and expressed as the % of the total signal in the band migrating at the position of the uncleaved fragment (mean and standard error). (D) RT-PCR assay of the *FLAG-TnpA* transcript in the calli used in Fig. 2B and C before and after 48 hrs of Dex induction.

One of the 3 calli assayed in Fig. 2 showed no decrease in DNA methylation upon dexamethasone induction. To determine whether the TnpA gene was expressed in all induced calli, the *TnpA* mRNA level was monitored by RT-PCR. As shown in Fig. 2D, *TnpA* mRNA was detectable in the two calli that showed demethylation after dexamethasone induction, but not in the callus that did not show demethylation of the *Spm* sequence. The data in Fig. 2C represent replicates of induction experiments performed with 3 independent calli derived from a single cross. Similar results have been obtained with 2-3 independent callus lines derived from 4 different crosses between SR1-1 plants and plants containing the inducible *TnpA* construct. Callus lines were checked by PCR to verify that both the *Spm-LUC* and dexamethasone-inducible *TnpA* constructs were present. About half (55%) of the callus lines containing both constructs exhibited dexamethasone inducible demethylation and all of these also showed dexamethasone-inducible *TnpA* mRNA expression, detectable by RT-PCR, as shown in Fig. 2D. In the

remaining lines, dexamethasone was unable to induce either demethylation or *TnpA* mRNA expression. Because *Spm* methylation remained unchanged after inducible expression of the FLAG-$TnpA_{420}$ protein (Figure 8, Chapter 4), the observed DNA demethylation in the TnpA-inducible lines cannot be attributed to the inducible system. It follows that demethylation of the *Spm* sequence can be controlled by inducing expression of TnpA.

TnpA Promotes Rapid *Spm* Demethylation

The complete loss of *Spm* methylation observed after 20 days of dexamethasone induction in the foregoing experiments suggests that TnpA-mediated DNA demethylation is an active process. If demethylation occurs by interference with remethylation of newly replicated DNA, the extent of demethylation should depend on the number of cell divisions and some methylation should still be detectable at 20 days, the precise amount depending on the exact doubling time. The extent of *Spm* DNA demethylation was therefore determined after treatment of calli with 5-aza-cytosine (5-aza-C) or 5-aza-deoxycytosine (5-aza-dC), both of which are incorporated into DNA in place of cytosine, but cannot be methylated (Bender et al., 1999). Little effect on *Spm* methylation was observed after 10 days of treatment with either compound at concentrations of either 10 or 50 μM. Even after 20 days of treatment, methylation was reduced by less than 50% at the higher concentration of 5-aza-dC (Fig. 3A). By contrast, no methylation was detected in calli treated with dexamethasone, either in the presence or absence of 5-aza-dC. Thus TnpA-mediated demethylation is either more rapid or more efficient than the demethylation resulting from the incorporation of 5-aza-dC.

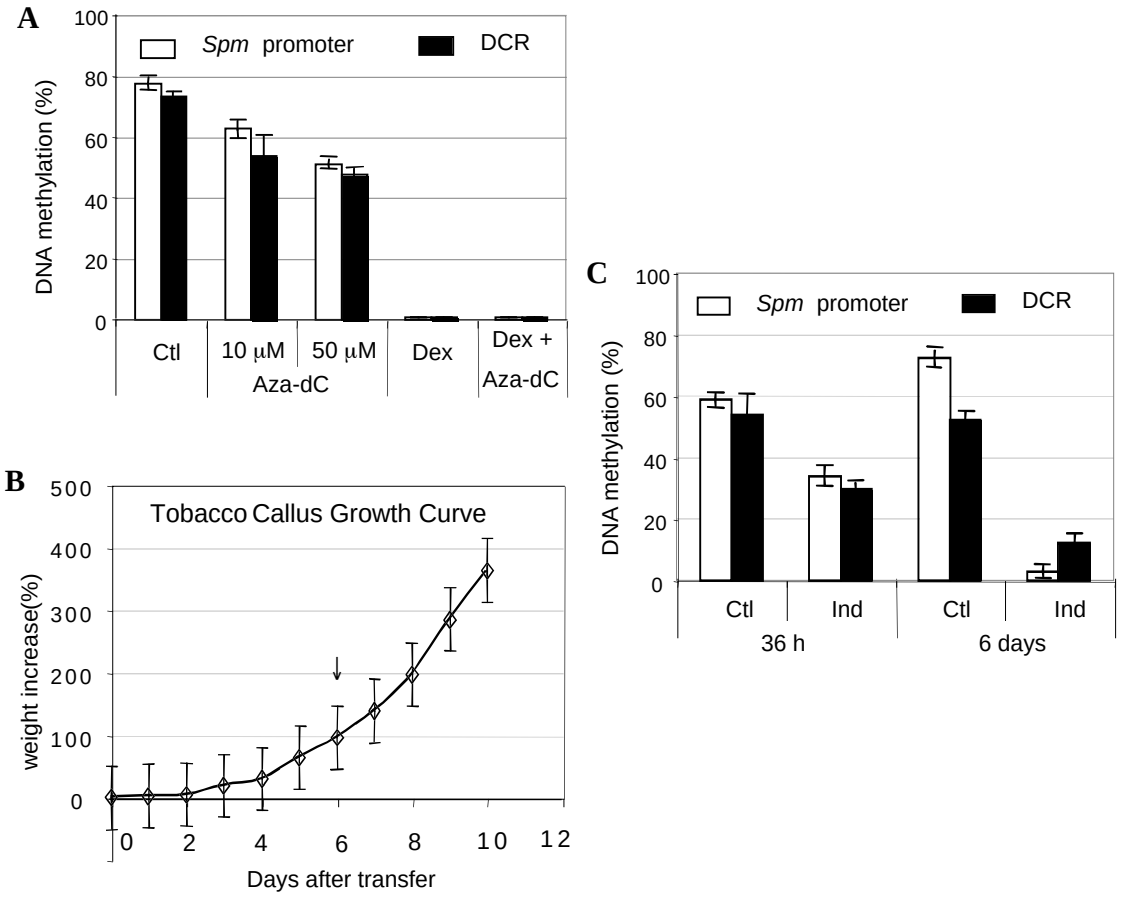


Figure 3. Rapid DNA demethylation after TnpA expression. (A) *Spm* promoter and DCR methylation in genomic DNA from transgenic tobacco calli after treatment for 20 days with 5-aza-deoxycytosine (aza-dC), 10 μM dexamethasone (Dex) or both. (B) The growth of tobacco callus after transfer to fresh medium. Each time point represents the mean (and standard error) of fresh weight measurements of 6 calli. The arrow indicates the initial doubling time. (C) *Spm* promoter and DCR methylation after 36 hrs and after 6 days of TnpA induction on 10 μM Dex. The values are the means and standard errors of measurements from two replicate experiments performed with each callus line at a similar growth stage.

To determine how rapidly the *Spm* sequence was demethylated after induction of TnpA expression relative to the cell doubling time, the callus growth curve was first determined under our experimental conditions (Fig. 3B). There is an initial lag phase,

giving an initial doubling time of 6-7 days, followed by an exponential growth phase with a doubling time of about 2 days. The methylation status of the *Spm* sequence was examined under precisely these growth conditions 1.5 and 6 days after transfer to dexamethasone-containing medium. As shown in Figure 3C, the *Spm* DNA was almost completely demethylated after 6 days of dexamethasone treatment and significant loss of methylation could already be detected by 36 hrs. If DNA demethylation is a consequence of interference with remethylation of newly replicated DNA, the *Spm* sequence should be hemimethylated after one cycle of DNA replication. If the hemimethylated DNA is still resistant to restriction by methylation-sensitive enzymes SalI and EcoO109I, no change in methylation is expected after a single doubling time, unless the DNA is actively demethylated.

Restriction Analysis of Hemimethylated DNA

Although it has been reported that the fully hemimethylated SalI site with both of its C residues methylated is resistant to cleavage, there is no information about either the sensitivity of the fully hemimethylated EcoO109I site or on either site when less than fully methylated (Nelson et al., 1993). A DNA fragment containing the *Spm* promoter and DCR was therefore methylated using the bacterial methyltransferase SssI, which only methylates the cytosine (C) residue in CpG dinucleotides. The sequence of the EcoO109I cleavage site in the *Spm* promoter is GGGTCCC with the last C followed by a G, placing it in the CG dinucleotide context. The SalI site sequence is GTCGAC. Thus, SssI methylates only one of the 3 C residues in the EcoO109I site and one of the two C residues in the SalI site.

As shown in Figure 4A, double-strand DNA methylated by SssI is completely resistant to SalI, but is still sensitive to EcoO109I. Because genomic *Spm* DNA is largely resistant to EcoO109I in the present experiments, it is likely that the EcoO109I site contains more than 1 methylated C residue (5mC), although it is possible that resistance is conferred by methylation of an internal C, but not the site's terminal C. To methylate all of the C residues in one strand, hemimethylated DNA was prepared by incorporating

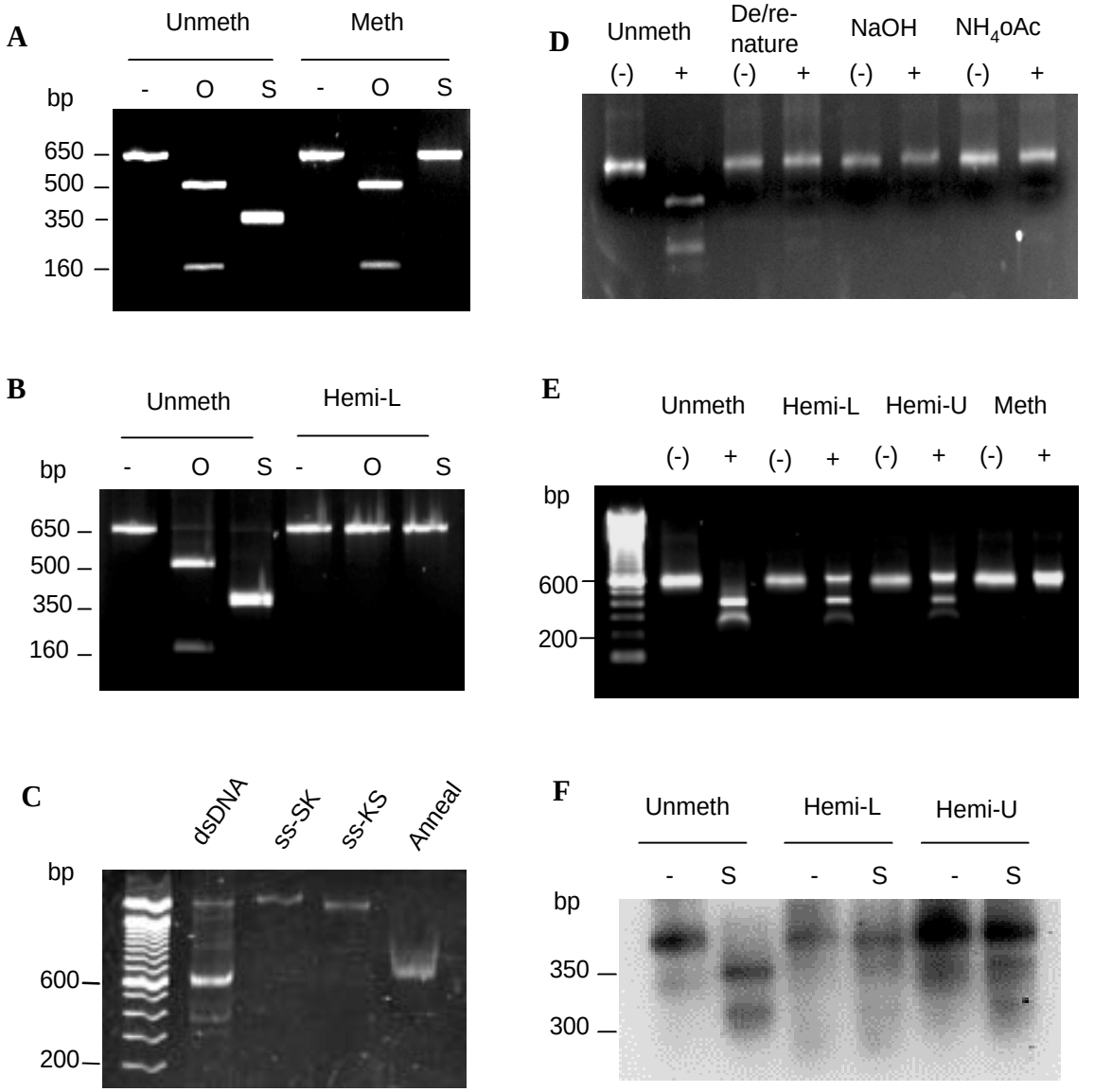


Figure 4. Sensitivity to restriction enzymes of hemimethylated DNA containing the *Spm* promoter and DCR. (A) EcoO109I (O) and SalI (S) digests of a PCR-amplified 674-bp fragment either unmethylated (Unmeth) or methylated (Meth) with SssI. The two fragments generated by SalI differ by only 27 bp and are not resolved on this gel. (B) EcoO109I (O) and SalI (S) digests of the same DNA fragment unmethylated (Unmeth) or hemi-methylated at all C residues in the lower strand (Hemi-L) as described in Methods. (C) Native PAGE gel purification of hemimethylated DNA (Anneal) by annealing of single-stranded (ss) DNA preparations. Ss-SK and ss-KS are the upper and lower strands, respectively. Double-stranded DNA (dsDNA) amplified by PCR was loaded as a control. (D) SalI restriction of unmethylated (unmeth) and SssI methylated dsDNA after treatments with the conditions indicated, as used for ssDNA preparation. (E) Sensitivity to SalI of the *Spm* sequence when it is unmethylated (unmeth), methylated by SssI in the lower (Hemi-L), the upper strand (Hemi-U) or both strands (Meth). 0.5 ug DNA was digested with 5 U SalI for 1hr. (F) An autoradiogram of a gel used to fractionate a 16-hr SalI digest of 6 μg tobacco genomic DNA mixed with 1 ng of the ^{32}P-labeled unmethylated (Unmeth) or hemi-methylated *Spm* promoter fragment with one C methylated in the lower strand (Hemi-L) or the upper strand (Hemi-U).

5-methyl-dCTP into one of the two strands as described in Methods. Hemimethylated DNA prepared in this way is completely resistant to both EcoO109I and SalI (Fig. 4B).

To determine the resistance of a singly methylated hemimethylated SalI site, hemimethylated DNA was prepared from SssI-methylated DNA (Fig. 4C, see Fig. 5 for a digramatic representation). Singly hemimethylated DNA is more resistant to SalI than unmethylated DNA, but could be digested when a large excess of SalI was used (Fig. 4E). Control experiments showed that methylated DNA is very stable under the condition used for hemimethylated DNA preparation, and thus the incomplete resistance of hemimethylated DNA to SalI restriction is not an artifact (Fig. 4D). To ask whether the singly hemi-methylated SalI site is digested under the conditions used to analyze *Spm*

sequence methylation in genomic DNA, ^{32}P-labeled singly hemimethylated *Spm* DNA was prepared using SssI methylation and a trace amount of it was added to a genomic DNA digest (1 ng labeled DNA/6 mg genomic DNA). As shown in Fig. 3F, the singly hemimethylated DNA was almost completely resistant to SalI digestion, irrespective of the strand into which the methyl group was introduced (hemi-L or -U).

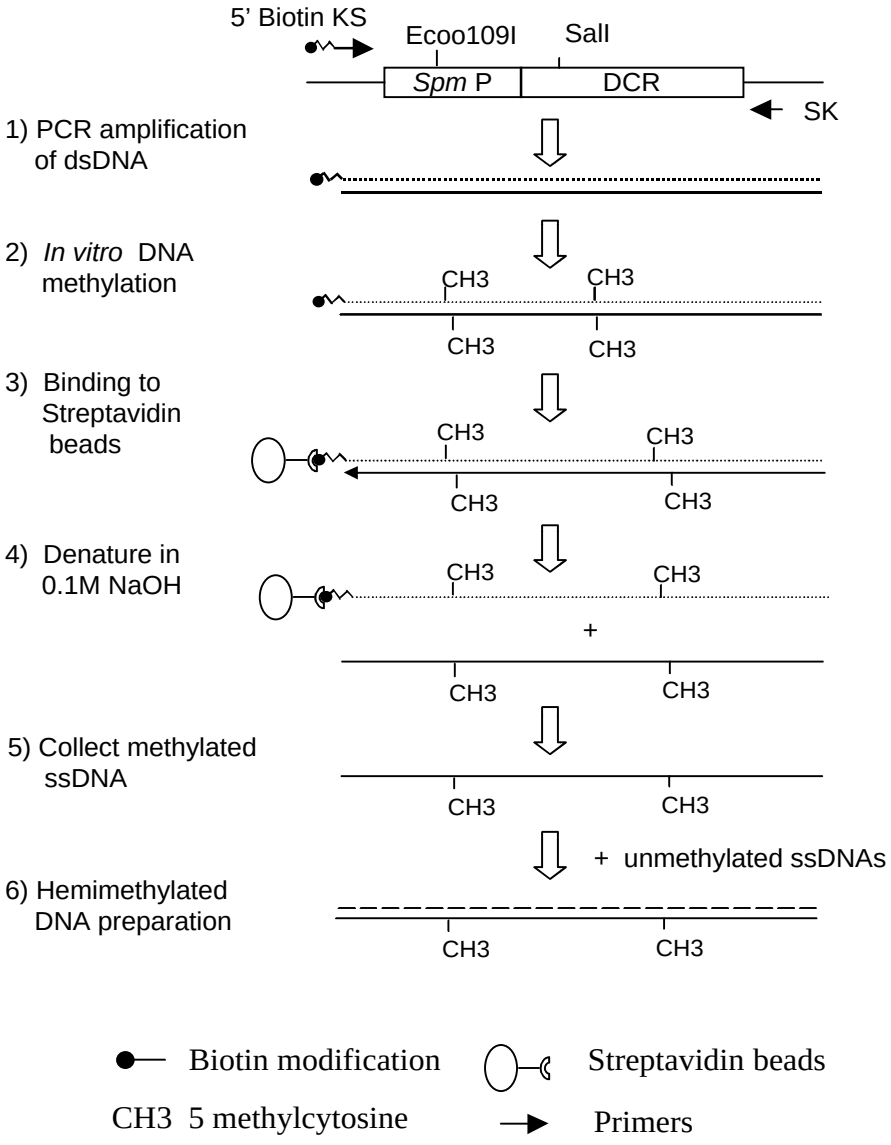


Figure 5. A diagram depicting the method for preparation of hemimethylated DNA using biotin/streptavidin interaction. See methods for details.

To examine the sensitivity of hemimethylated DNA with fewer 5mC residues per site to EcoO109I, hemimethylated DNA was prepared using a mixture of 5-methyl-dCTP and dCTP at ratios of 3:1 and 2:1. From the ratio of methylated to unmethylated dCTP, the expected frequencies of tri-methylated, di-methylated, monomethylated and unmethylated EcoO109I sites can be calculated (see Methods). Based on this calculation, the extent of digestion can in turn be predicted if just unmethylated and monomethylated sites are cut (Table 1, 15.6% at 3:1 and 25.9% at 2:1) and if dimethylated sites are cut, as well (Table 1, 57.8% at 3:1 and 70.4% at 2:1). The observed extent of EcoO109I cleavage, also given in Table 1, is very close to that predicted if only the unmethylated and monomethylated sites can be cleaved in hemimethylated DNA (Table 1, 15.3% and 22.4%). It is therefore concluded that hemimethylated DNA with any 2 of its 3 C's methylated at the EcoO109I site is resistant to digestion. Since genomic *Spm* DNA is highly resistant to EcoO109I and is therefore likely to have at least two methylated C residues in its EcoO109I site, it follows that the newly replicated, hemimethylated site will also be resistant to cleavage by the enzyme.

The foregoing results clearly show that the hemimethylated EcoO109I and SalI sites are resistant to enzymatic digestion when methylated at all internal C residues. They further provide evidence that under the conditions employed in the present experiments, both sites are resistant even when less than fully hemimethylated. The initial doubling time for the callus cultures under the assay conditions was about 6 days, with subsequent doubling times of about 2 days. If *Spm* DNA is demethylated by TnpA's interference with remethylation of hemimethylated DNA, it should be fully resistant to restriction by SalI and EcoO109I after one doubling time (day 6), 50% resistant after 2 (day 8) and

25% resistant after 3 doubling times (day 10). Since methylation was reduced by about 40% after 36 hrs, much less than the initial doubling time, and virtually gone by 6 days, we conclude that TnpA-mediated DNA demethylation is an active process.

Table I. Expected and observed extent of EcoO109I cleavage of hemimethylated DNA with different number of methylated cytosine residues (5mC) at the recognition site. The hemimethylated DNA was generated by copying unmethylated single-stranded DNA using Tag DNA polymerase and a mixture of dCTP and 5-methyl-dCTP.

	EcoO109I site (GGGTCCC[a])			
Ratio[b] 5m-dCTP:dCTP	Expected DNA composition[c]	5mC (1) + 5mC (0)	5mC (2) + 5mC (1) + 5mC (0)	DNA cleaved
3:1	5mC (3): 42.2% 5mC (2): 42.2% 5mC (1): 14.0% 5mC (0): 1.6%	15.6%	57.8%	15.3%
2:1	5mC (3): 29.6% 5mC (2): 44.5% 5mC (1): 22.2% 5mC (0): 3.7%	25.9%	70.4%	22.4%

[a] EcoO109I recognition sequence in the *Spm* promoter.

[b] The ratio of 5-methyl-dCTP to dCTP used in preparing hemimethylated DNA.

[c] 5mC(N), the calculated percentage of DNA with the indicated number (N) of 5mC residues in the EcoO109I site (see Methods for calculation).

Discussion

An Inducible DNA Demethylation Assay

Most studies on DNA demethylation rely on the use of *in vitro* premethylated DNA, which is transfected either transiently or stably into cells. There are some caveats with this approach, however. First of all, naked DNA introduced by this method readily gets demethylated even in the absence of the trans-acting demethylation-promoting factors, presumably by general DNA demethylase activities in the cell (Qu and Ehrlich, 1999; Cervoni and Szyf, 2001). Because known DNA demethylase activities do not show sequence specificity (Jost et al., 1995; Neddermann et al., 1996; Bhattacharya et al., 1999; Vairapandi et al., 2000; Jost et al., 2001), the observed demethylation may not at all reflect the activity of the trans-acting factor under investigation. In the case of transient transfection, a large amount of DNA is generally used but only a small fraction can get into the cells and even less DNA is demethylated, making it difficult to detect the product of DNA demethylation.

Preliminary studies in the laboratory indicate that in tobacco cells there is a DNA demethylase activity that does not have sequence specificity (Chen and Fedoroff, unpublished observation). To overcome the problems associated with transient transfection, I have thus developed an *in vivo* assay system for dissecting the mechanism of TnpA-mediated DNA demethylation.

By controlling DNA demethylation through inducible TnpA expression, we are able to characterize the initial stages of DNA demethylation through several cell division cycles in the native chromatin context. The importance of chromatin structure in

controlling DNA methylation has been revealed in recent studies (see Introduction). This has not been possible previously because methylation was studied either in transformed cells or in regenerated plants many cell generations after the introduction of a *TnpA* gene or *Spm* transposon by transformation. The ability to induce DNA demethylation will facilitate further molecular analysis of TnpA-mediated DNA demethylation. To our knowledge, this is the first inducible DNA demethylation system ever reported. This approach should be useful as well for studies of other trans-acting factors that promote sequence-specific DNA demethylation.

Active DNA demethylation in plants

Although there are both cytological and biochemical evidence that active DNA demethylation occurs in animals cells, very little is known about plants in this regard. The first evidence that DNA is actively demethylated in plants is derived from the observation that DNA methylation is dramatically reduced in the generative but not the vegetative nucleus of germinating pollen in the absence of DNA replication (Oakeley et al., 1997). The present study has provided further support for the existence of an active DNA demethylation mechanism in plants and represents the first example of sequence-specific DNA demethylation mediated by trans-acting factors. More recently, the ROS protein in *Arabidospsis* has been shown to be able to nick methylated DNA, implicating its role in DNA demethylation through a DNA repair pathway (Gong et al., 2002).

Sensitivity of Hemimethylated DNA

DNA methylation-sensitive restriction enzymes are often used for studies of DNA methylation and demethylation. However, artifacts may result if this method is not used properly, because the extent of digestion of some restriction sites is dependent on not only the extent of methylation (the number of 5mC and whether one or both strands are methylated) but also the amount of enzyme and DNA used in the restriction reaction, as manifested by this and other studies (Nelson et al., 1993). In the present study, I have found that although the SalI recognition site is completely resistant to SalI restriction when methylated at both strands, it is only partially resistant to this enzyme when only one strand is methylated with a single 5mC. While hemimethylated DNA with one cytosine residue methylated is still restrictable, hemimethylated DNA with both cytosine residues methylated becomes completely resistant to SalI restriction. This study also showed that DNA hemimethylated at the SalI site with one 5mC is protected from SalI restriction by an excess of unrelated DNA that also contains SalI restriction sites. Similarly, the EcoO109I site is sensitive to restriction when only one cytosine is methylated but becomes completely resistant when more than two cytosine residues methylated. Since not all restriction sites have been characterized regarding sensitivity to restriction when they are variously methylated, caution must be exercised when methylation-sensitive enzymes are used for studies of DNA methylation and demethylation. The methods for preparation of hemimethylated DNA developed in this study would be helpful in characterizing the sensitivity of other restriction sites.

References

Aoyama, T., and Chua, N.H. (1997). A glucocorticoid-mediated transcriptional induction system in transgenic plants. Plant J. **11,** 605-612.

Banks, J.A., Masson, P., and Fedoroff, N. (1988). Molecular mechanisms in the developmental regulation of the maize Suppressor-mutator transposable element. Genes Dev. **2,** 1364-1380.

Banks, J.A., and Fedoroff, N. (1989). Patterns of developmental and heritable change in methylation of the Suppressor-mutator transposable element. Dev. Genet. **10,** 425-437.

Bender, C.M., Gonzalgo, M.L., Gonzales, F.A., Nguyen, C.T., Robertson, K.D., and Jones, P.A. (1999). Roles of cell division and gene transcription in the methylation of CpG islands. Mol. Cell. Biol. **19,** 6690-6698.

Bhattacharya, S.K., Ramchandani, S., Cervoni, N., and Szyf, M. (1999). A mammalian protein with specific demethylase activity for mCpG DNA [see comments]. Nature **397,** 579-583.

Cervoni, N., and Szyf, M. (2001). Demethylase activity is directed by histone acetylation. J. Biol. Chem. **276,** 40778-40787.

Dean, W., Santos, F., Stojkovic, M., Zakhartchenko, V., Walter, J., Wolf, E., and Reik, W. (2001). Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. Proc. Natl. Acad. Sci. USA **98,** 13734-13738.

Fedoroff, N. (1989). The heritable activation of cryptic Suppressor-mutator elements by an active element. Genetics **121,** 591-608.

Gong, Z., Morales-Ruiz, T., Ariza, R.R., Roldan-Arjona, T., David, L., and Zhu, J.K. (2002). ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase. Cell **111,** 803-814.

Hsieh, C.L. (1999). Evidence that protein binding specifies sites of DNA demethylation. Mol. Cell. Biol. **19,** 46-56.

Jost, J.P., Siegmann, M., Sun, L., and Leung, R. (1995). Mechanisms of DNA demethylation in chicken embryos. Purification and properties of a 5-methylcytosine-DNA glycosylase. J. Biol. Chem. **270,** 9734-9739.

Jost, J.P., Oakeley, E.J., Zhu, B., Benjamin, D., Thiry, S., Siegmann, M., and Jost, Y.C. (2001). 5-Methylcytosine DNA glycosylase participates in the genome-wide loss of DNA methylation occurring during mouse myoblast differentiation. Nucleic Acids Res. **29,** 4452-4461.

McBride, K.E., and Summerfelt, K.R. (1990). Improved binary vectors for Agrobacterium-mediated plant transformation. Plant Mol. Biol. **14,** 269-276.

Neddermann, P., Gallinari, P., Lettieri, T., Schmid, D., Truong, O., Hsuan, J.J., Wiebauer, K., and Jiricny, J. (1996). Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase. J. Biol. Chem. **271,** 12767-12774.

Nelson, P.S., Papas, T.S., and Schweinfest, C.W. (1993). Restriction endonuclease cleavage of 5-methyl-deoxycytosine hemimethylated DNA at high enzyme-to-substrate ratios. Nucleic Acids Res. **21,** 681-686.

Oakeley, E.J., Podesta, A., and Jost, J.P. (1997). Developmental changes in DNA methylation of the two tobacco pollen nuclei during maturation. Proc. Natl. Acad. Sci. USA **94,** 11721-11725.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. Curr. Biol. **10,** 475-478.

Qu, G., and Ehrlich, M. (1999). Demethylation and expression of methylated plasmid DNA stably transfected into HeLa cells [In Process Citation]. Nucleic Acids Res. **27,** 2332-2338.

Raina, R., Cook, D., and Fedoroff, N. (1993). Maize Spm transposable element has an enhancer-insensitive promoter. Proc. Natl. Acad. Sci. USA **90,** 6355-6359.

Raina, R., Schläppi, M., Karunanandaa, B., Elhofy, A., and Fedoroff, N. (1998). Concerted formation of macromolecular Suppressor-mutator transposition complexes. Proc. Natl. Acad. Sci. USA **95,** 8526-8531.

Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. Science **293,** 1089-1093.

Schläppi, M., Smith, D., and Fedoroff, N. (1993). TnpA trans-activates methylated maize Suppressor-mutator transposable elements in transgenic tobacco. Genetics **133,** 1009-1021.

Schläppi, M., Raina, R., and Fedoroff, N. (1994). Epigenetic regulation of the maize Spm transposable element: novel activation of a methylated promoter by TnpA. Cell **77,** 427-437.

Schläppi, M., Raina, R., and Fedoroff, N. (1996). A highly sensitive plant hybrid protein assay system based on the Spm promoter and TnpA protein for detection and analysis of transcription activation domains. Plant Mol. Biol. **32,** 717-725.

Stancheva, I., El-Maarri, O., Walter, J., Niveleau, A., and Meehan, R.R. (2002). DNA methylation at promoter regions regulates the timing of gene activation in Xenopus laevis embryos. Dev. Biol. **243,** 155-165.

Vairapandi, M., Liebermann, D.A., Hoffman, B., and Duker, N.J. (2000). Human DNA-demethylating activity: A glycosylase associated with RNA and PCNA. J. Cell. Biochem. **79,** 249-260.

von Arnim, A.G., Deng, X.W., and Stacey, M.G. (1998). Cloning vectors for the expression of green fluorescent protein fusion proteins in transgenic plants. Gene **221,** 35-43.

Chapter 3

TNPA IS A *SPM*-SPECIFIC TRANSCRIPTIONAL ACTIVATOR

Introduction

A prominent feature of TnpA-mediated DNA demethylation is that it is not confined to the *Spm* promoter, which contains multiple TnpA binding sites (Raina et al., 1998), but also occurs in the downstream GC-rich sequence termed the DCR, which encodes the first untranslated exon of the full-length *Spm* transcript (Masson et al., 1991). Since there are no TnpA binding sites in the DCR (Raina et al., 1998), it is unlikely that DNA demethylation of the *Spm* sequence, particularly the DCR, is conferred by TnpA alone. The extended region of DNA demethylation suggests that other trans-acting factors might be involved as well. Interestingly, a leucine-zipper domain has been identified at the C-terminus of the TnpA protein (Trentmann et al., 1993), it is therefore possible that TnpA recruits other proteins that participate in the demethylation of the *Spm* sequence .

To investigate this possibility, I first attempted to screen for plant proteins that interact with TnpA using the yeast two-hybrid interaction trap system (Fields and Song, 1989). Surprisingly, TnpA strongly activated transcription as indicated by two reporter genes in the yeast cells. That TnpA is an *Spm*-specific transcriptional activator is confirmed by further studies using a transgenic line of tobacco that shows inducible TnpA expression and contains an unmethylated *Spm* sequence. Since the *Spm* sequence

remains unmethylated during the assay, transcriptional activation cannot be attributed to DNA demethylation. The present study resolves a long-standing paradoxical observation that TnpA acts as both a transcriptional activator and repressor depending on the methylation status of the *Spm* promoter.

Materials and Methods

Plant materials and treatments

Calli for expression of FLAG-TnpA$_{540}$-VP16 were generated and maintained in similar ways for the FLAG-TnpA expressing lines as described in Chapter 2. Suspension cultures were derived from transgenic calli and were maintained at 24°C with shaking (120 rpm) in MS medium supplemented with 3% sucrose and 0.2 mg/L 2,4-dichlorophenoxyacetic acid (2,4-D) (Life Technologies). Expression of the FLAG-tagged proteins was induced in suspension cultures by adding 10 μM dexamethasone (Sigma-Aldrich) to the growth medium. Induction of protein expression in calli was performed in the same was as described in Chapter 2.

Plasmid Constructs

All FLAG-tagged clones were derivatives of pFLAG-TnpA. pFLAG-TnpA$_{540}$-Vp16 was obtained by ligating the XhoI-SplI fragment of pFLAG-TnpA and the SplI-NcoI fragment (the NcoI site was filled in using the Klenow fragment) from pMS249 (Schläppi et al., 1996) together into the pBluescript II KS (+) vector cut with XhoI and

EcoRV. pTA-FLAG-TnpA$_{540}$-Vp16 was cloned by ligating a ClaI to SpeI fragment from pFLAG-TnpA$_{540}$-VP16 and pTA7002 that was cut with XhoI and SpeI and the XhoI site was partially filled in to the adenine residue.

The construct for expression of the fusion protein between the Gal4 DNA-binding domain and TnpA in yeast cells was cloned as follows. The yeast expression vector pAS2 was first cut with NcoI, filled in with Klenow and then cut with SalI. pRR466 (Raina et al., 1993) was cut with BamHI, filled in with Klenow and then cut with SalI to release the cDNAs for TnpA, which were then cloned into pAS2, giving rise to pAS2-TnpA.

Transcription Assays

Plasmid pAS2-TnpA as well as the vector pAS2 was transformed into yeast strain Y190. Yeast transformation and the subsequent transcription assay were conducted according to the protocol described for the HybriZAP® two-hybrid cDNA Gigapack® cloning kit (Stratagene).

RT-PCR was used to assay transcriptional activation of the *Spm-LUC* gene by TnpA in plant cells. Total RNA was prepared using either the Qiagen RNeasy Plant Mini kit or the GibcoBRL® TRIzol® reagent. To remove residual DNA, RNA was treated with DNase I for 5 min as recommended by the vendor (Life Technologies). For reverse transcription, 1 μl of 100 ng/μl total RNA was mixed with 3 μl of 100 ng/μl random primer and 8 μl RNase-free water, then denatured at 70°C for 10 min. The solution was chilled on ice and collected by centrifugation. A premixed solution of the following reagents was then added: 4 μl of 5x first strand buffer (Life Technologies), 2 μl of 0.1 M DTT, 1 μl 10 μM dNTPs and 2 units RNase block inhibitor. The solution was incubated at 25° C for 10 min, and then at 42° C for 2 min, followed by the addition of 1 μl

Superscript RT (Life Technologies) and incubation at 42° C for 50 min. The reaction was terminated by heating at 70° C for 15 min and 1-2 μl was used in a 20 μl PCR reaction. Simultaneous amplification of the transcripts for the *FLAG-TnpA* gene and the luciferase (*LUC*) reporter gene was achieved by including both primer sets in the same reaction. For semi-quantitative analysis, the PCR reaction was stopped after 15 cycles, fractionated on an agarose gel and then transferred to Hybond^{+} nitrocellulose membrane (Amersham). The amplification products for the *TnpA* and luciferase genes were detected by Southern blot analysis. To ensure that equal amounts of total RNA had been used in the reaction, the 18S rRNA was also amplified by RT-PCR for 10 cycles using the primer pair 18S-5' (5'TACCGTCCTAGTCTCAACCA3') and
18S-3' (5'AACATCTAAGGGCATCACAG3').

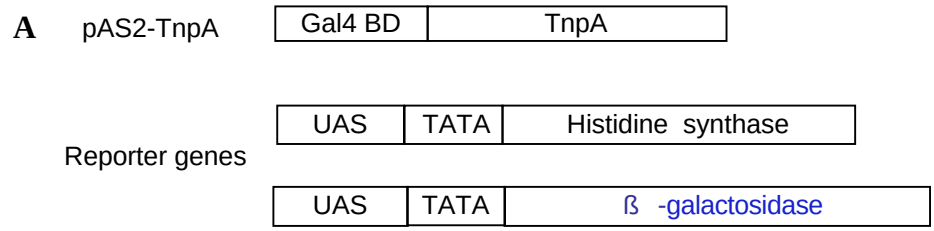
Western Blot Assay

The FLAG-tagged TnpA expressed in transgenic tobacco plants were detected in transgenic plants extracts using the polyclonal anti-FLAG antibody (Sigma-Aldrich). Total plant protein was extracted by grinding plant tissue in 2 volumes of TBS buffer (50 mM Tris-HCl, pH 7.8, 150 mM NaCl) supplemented with 1 mM PMSF and 1x protease inhibitor cocktail for plants (Sigma-Aldrich). The extract was centrifuged at 14,000 rpm at 4° C for 10 min, the supernatant was mixed with 6x SDS loading buffer (Scopes, 1998), boiled for 5 min, and 30 μl was loaded on an 8% polyacrylamide gel for fractionation. Proteins expressed in both yeast and plant cells were visualized using the WesternBreeze^{TM} chemifluorescence immunodetection system (Invitrogen).

Results

TnpA Activates Transcription in Yeast Cells

To identify proteins that could interact with TnpA, I first used the yeast two-hybrid system. However, attempts to use TnpA and TnpA fragments as bait (Fig. 6A) revealed that a Gal4 fusion with the full-length TnpA protein strongly activates transcription of both the histidine synthase and β-galactosidase reporter genes (Fig. 6B).



B

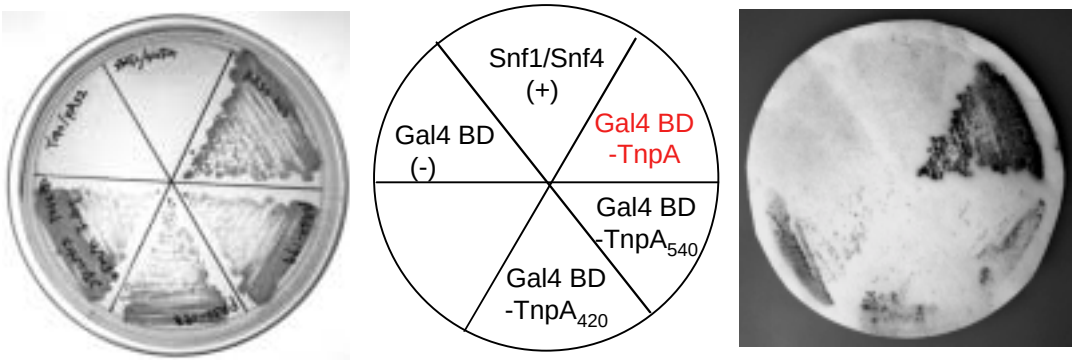
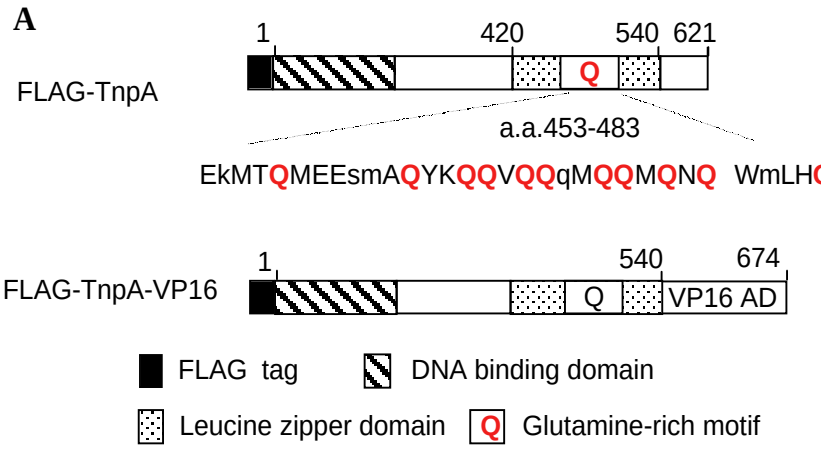


Figure 6. TnpA is a transcriptional activator in yeast. (A) Construct used for expression of the fusion protein between the Gal4 binding domain (BD) and TnpA as well as the reporter genes for transcriptional activity assay in yeast cells. UAS, Gal4 binding sequence.
(B) Growth of transformed yeast cells on histidine (-) medium (left) and filter paper lift assay for β-galactosidase activity (right). Cells expressing the Gal4 BD was used a negative control. Snf1/Snf4 cells express the Gal4 BD-SNF1 and SNF4-Gal4 activation domain fusion proteins and serve as a weak positive control. The middle panel is a diagram showing the position of yeast cells transformed with the constructs as indicated.

TnpA is a Weak Transcriptional Activator in Plant Cells

To assess the transcriptional activity of TnpA on an unmethylated *Spm* promoter, I used suspension cultures derived from tobacco callus lines containing a dexamethasone-inducible *TnpA* gene and an *Spm-LUC* gene in which the *Spm* promoter was not methylated. As a positive control, suspension cultures were also generated from calli that express the FLAG-TnpA$_{540}$-VP16 fusion protein under the control of the dexamethasone-inducible promoter (Fig 7A). After addition of dexamethasone into the cell culture, the *FLAG-TnpA* transcript peaked at 24-48 hrs (Fig. 7B-1). Expression of the tagged proteins peaked at about 72 hrs (Figs. 7C). Induction of TnpA enhanced transcription of the luciferase reporter gene, albeit weakly, and maximal induction was observed when the TnpA protein, rather than its transcript, reached maximum (Fig. 7B-1). Activation of transcription from the *Spm* promoter by TnpA was reproducible and was observed when independent lines were examined (data not shown). Notably, there is some background level of transcription even in the absence of TnpA induction, which is likely due to leaky expression from the inducible promoter or the *Spm* promoter. Nevertheless, repression of transcription was not observed, even at elevated concentrations of dexamethasone (data not shown). Because the strong TnpA induction could artifactually enhance the *LUC* mRNA signal, the blot was stripped and reprobed with just a *LUC* probe. The result confirmed that the *LUC* transcript level indeed increased (Fig. 7B-2). By contrast, although induction of the *FLAG-TnpA$_{540}$-VP16* gene was very slight, it substantially enhanced transcription of the *LUC* gene from the *Spm* promoter (Fig. 7, B-1 and B-2). Taken together, these results showed that TnpA can stimulate transcription from the

unmethylated *Spm* promoter and is therefore a transcriptional activator, albeit a much weaker one than the FLAG-TnpA$_{540}$-VP16 fusion protein.



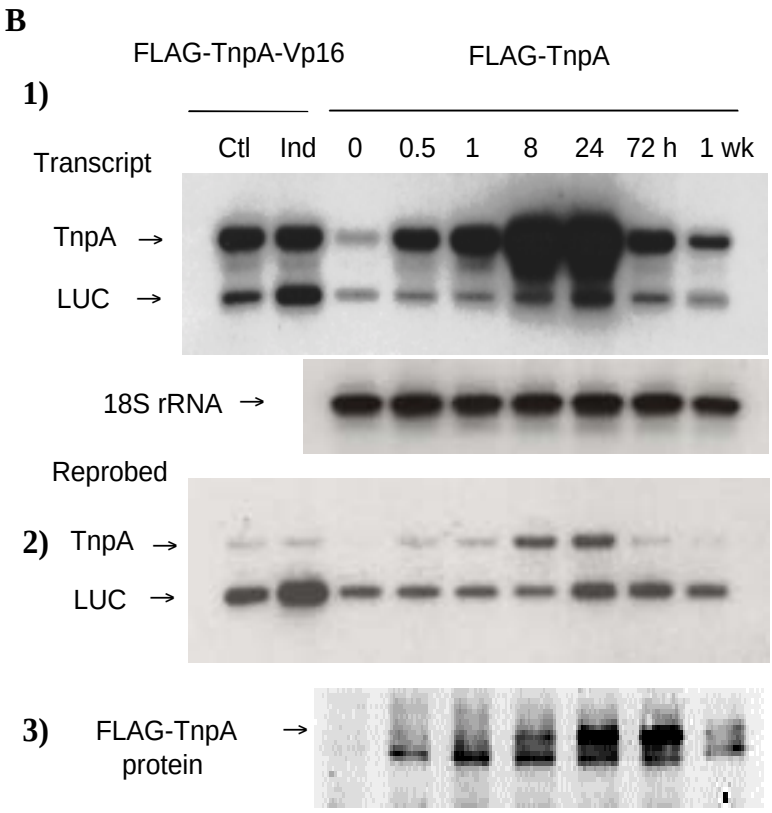


Figure 7. TnpA is a weak transcriptional activator in tobacco cells. (A) The top diagram shows the domain structure of the TnpA protein. The bottom diagram shows the structure of the FLAG-tagged TnpA-VP16 fusion protein. (B-1) Analysis of *TnpA* and luciferase (*LUC*) transcripts following induction of the *FLAG-TnpA-VP16* and *FLAG-TnpA* genes in suspension-cultured transgenic tobacco cells. RNA was amplified by RT-PCR (15 cycles) followed by Southern analysis for signal detection. Equal amounts of *TnpA* and *LUC* probes were used; the 18S rRNA was amplified as a loading control. (B-2) To better reveal the change in *LUC* transcript level in FLAG-TnpA expressing cells, the membrane used in B-1 was stripped and reprobed with only the *LUC* probe; the TnpA bands represent residual label from the initial hybridization. (C) Western analysis of proteins in the corresponding cultures using FLAG antibodies.

Discussion

TnpA has been shown to activate transcription from the *Spm* promoter when methylated, but represses transcription from an unmethylated *Spm* promoter . Paradoxically, however, these studies also showed that the *Spm* promoter has stronger transcription activity when it becomes demethylated due to the presence of TnpA (Schläppi et al., 1994). The present study shows that TnpA is a weak transcriptional activator of the unmethylated *Spm* promoter in plant cells, although its transcription activation domain functions as a strong transcriptional activator when fused to a heterologous DNA binding domain in yeast. Since the *Spm* sequence is unmethylated even before TnpA expression, the observed transcriptional activation is unlikely a consequence of DNA demethylation. DOPA, the TnpA homolog encoded by the *Spm*-like maize *Doppia* transposon, recently has been reported to be a transcriptional activator (Bercury et al., 2001). Interestingly, analysis of the TnpA sequence using a motif-recognition program (http://www.motif.genome.ad.jp/) revealed the presence of a glutamine-rich motif, a common transcription activation domain motif, within the C-terminal leucine-zipper domain (Fig. 7A). However, further studies are needed to determine whether it is responsible for the observed transcriptional activation. In yeast cells glutamine-rich domain activates transcription weakly.

The paradoxical observation about TnpA in the previous studies could arise from the intrinsic drawbacks of the transcription assays used in that experiment. In the case of the transient assay, where TnpA expression construct was introduced into plant cells by the biolistic delivery method, transcription repression could occur as a consequence of

'self-interference' or 'coactivator dilution-out' due to protein over-expression (Kannan and Tainsky, 1999; Nucifora et al., 2001). Similar situation could have occurred in the studies using stable transgenic plants, where TnpA was expressed under strong promoters (Schläppi et al., 1994). Another possible cause for the observed repression by TnpA of transcription from an unmethylated *Spm* promoter is the high background activity of the reporter gene in the line used for the transcription assay. Conceivably, a protein with an intermediate transcription activity can act as an activator on a promoter that is located in a transcriptionally inactive locus or as a repressor when the reporter gene is in an actively transcribing region. In this sense, transcription activation or repression by a trans-acting factor is only relative to the background level of the reporter gene. Accordingly, transcription repression by TnpA in one line does not necessarily mean that TnpA also represses transcription in other lines. To further investigate this possibility, it would be interesting to examine the transcription activity of TnpA in a line that shows high luciferase activity using the TnpA inducible strategy. This experiment is worth doing since other proteins with dual properties have been reported elsewhere (Ohlsson et al., 2001).

In the present study transcription assay has been conducted in the same line of transgenic plants that has an unmethylated *Spm* sequence, the results therefore should better reflect TnpA's intrinsic activity. Although transcription activation from the *Spm* promoter by TnpA is rather weak, it is nevertheless reproducible and consistent when independent transgenic lines are examined. Since the transcription activity before and after induction is measured in the same transgenic line, variability due to sample heterogeneity is minimized, thus permitting reliable detection of even small changes in

transcriptional activity. Variability has been further reduced by using suspension cultures, which has a more homogeneous cell constitution.

The resolution of the apparent paradox that TnpA functions as both an activator and repressor may also lie in the protein concentration-dependent competition between TnpA's functions in transcription and transposition. TnpA is a bi-functional protein, and the *Spm* sequence comprising its promoter likewise has two functions. TnpA is required both for transcription from the transposon's promoter and for transposition of *Spm* (Masson et al., 1991; Schläppi et al., 1993). There are multiple TnpA binding sites at both transposon ends (Masson et al., 1987). Those at the 5' end are required for promoter function and the TnpA binding sites at both ends are involved in transposition (Masson et al., 1987; Raina et al., 1993).

Previous experiments to assess the ability of TnpA to activate transcription from an unmethylated *Spm* promoter were carried out in transient transfection assays at high DNA concentrations (Schläppi et al., 1996). The experiments described here were carried out by inducing expression of a TnpA gene in vivo and measuring transcription from an integrated chromosomal copy of the *Spm-LUC* reporter gene. Hence the concentrations of both the target promoter and the transcription factor are likely to have been much lower in these experiments than in the previous transient expression assays (Schläppi et al., 1996). Because the *Spm* promoter region is co-extensive with the TnpA-binding region at the 5' terminus (Raina et al., 1993), the likeliest explanation is that TnpA-mediated association of *Spm* termini competes with other types of protein-protein interactions involving TnpA, including transcriptional activation. Evidence that it is TnpA and not the inherent weakness of the *Spm* promoter that limits transcriptional

activity is provided by the observation that the $TnpA_{540}$-VP16 fusion protein is a strong transcriptional activator of the *Spm* promoter (Schläppi et al., 1996). The observation that TnpA is a much weaker transcriptional activator of the *Spm* promoter than might be anticipated from the ability of the Gal4-TnpA fusion protein to activate transcription in yeast is consistent with the interpretation that TnpA is a transcriptional activator at low concentrations, but becomes a repressor by homodimerization as its concentration increases. However, the possibility can not be ruled out that TnpA's transcriptional activity is modulated by co-repressors present in plant cells, but not in yeast. In maize, host factors that bind to the subterminal sequences of Ac/Ds and the Mutator element and regulate their transpositional activity have been identified (Becker and Kunze, 1996; Jarvis et al., 1997), and these proteins could bind to the *Spm* sequence as well and affect its promoter activity.

References

Aoyama, T., and Chua, N.H. (1997). A glucocorticoid-mediated transcriptional induction system in transgenic plants. Plant J. **11,** 605-612.

Becker, H.A., and Kunze, R. (1996). Binding sites for maize nuclear proteins in the subterminal regions of the transposable element Activator. Mol. Gen. Genet. **251,** 428-435.

Bercury, S.D., Panavas, T., Irenze, K., and Walker, E.L. (2001). Molecular analysis of the Doppia transposable element of maize. Plant Mol. Biol. **47,** 341-351.

Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. Nature **340,** 245-246.

Jarvis, P., Belzile, F., Page, T., and Dean, C. (1997). Increased Ac excision (iae): Arabidopsis thaliana mutations affecting Ac transposition. Plant J. **11,** 907-919.

Kannan, P., and Tainsky, M.A. (1999). Coactivator PC4 mediates AP-2 transcriptional activity and suppresses ras-induced transformation dependent on AP-2 transcriptional interference. Mol. Cell. Biol. **19,** 899-908.

Masson, P., Surosky, R., Kingsbury, J.A., and Fedoroff, N.V. (1987). Genetic and molecular analysis of the *Spm*-dependent *a-m2* alleles of the maize *a* locus. Genetics **117,** 117-137.

Masson, P., Strem, M., and Fedoroff, N. (1991). The tnpA and tnpD gene products of the Spm element are required for transposition in tobacco. Plant Cell **3,** 73-85.

Nucifora, F.C., Jr., Sasaki, M., Peters, M.F., Huang, H., Cooper, J.K., Yamada, M., Takahashi, H., Tsuji, S., Troncoso, J., Dawson, V.L., Dawson, T.M., and Ross, C.A. (2001). Interference by huntingtin and atrophin-1 with cbp-mediated transcription leading to cellular toxicity. Science **291,** 2423-2428.

Ohlsson, R., Renkawitz, R., and Lobanenkov, V. (2001). CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. Trends Genet. **17,** 520-527.

Raina, R., Cook, D., and Fedoroff, N. (1993). Maize Spm transposable element has an enhancer-insensitive promoter. Proc. Natl. Acad. Sci. USA **90,** 6355-6359.

Raina, R., Schläppi, M., Karunanandaa, B., Elhofy, A., and Fedoroff, N. (1998). Concerted formation of macromolecular Suppressor-mutator transposition complexes. Proc. Natl. Acad. Sci. USA **95,** 8526-8531.

Schläppi, M., Smith, D., and Fedoroff, N. (1993). TnpA trans-activates methylated maize Suppressor-mutator transposable elements in transgenic tobacco. Genetics **133,** 1009-1021.

Schläppi, M., Raina, R., and Fedoroff, N. (1994). Epigenetic regulation of the maize Spm transposable element: novel activation of a methylated promoter by TnpA. Cell **77,** 427-437.

Schläppi, M., Raina, R., and Fedoroff, N. (1996). A highly sensitive plant hybrid protein assay system based on the Spm promoter and TnpA protein for detection and analysis of transcription activation domains. Plant Mol. Biol. **32,** 717-725.

Scopes, R.a.S.J. (1998). Analysis of Proteins. In "Current Protocols in Molecular Biology". **2.,** 10.10.11-10.10.20.

Trentmann, S.M., Saedler, H., and Gierl, A. (1993). The transposable element *En/Spm*-encoded TNPA protein contains a DNA binding and a dimerization domain. Mol. Gen. Genet. **238,** 201-208.

Chapter 4

THE TRANSCRIPTIONAL ACTIVITY OF TNPA IS ESSENTIAL FOR ITS ABILITY TO PROMTE *SPM* DEMETHYLATION

Introduction

Among the trans-acting factors that are known to promote sequence-specific DNA demethylation, most are transcriptional factors, Sp1 (Silke et al., 1995), NF-kappa B (Kirillov et al., 1996; Matsuo et al., 1998) and the glucocorticoid receptor (Grange et al., 2001; Thomassin et al., 2001) for example. This observation suggests that transcription activation plays a positive role in DNA demthylation. Indeed, general transcription machinery has been shown to be essential for DNA demethylation mediated by the NF-kappa B transcriptional activator, although actual transcription does not seem to be required (Matsuo et al., 1998). Chromatin remodeling and transcriptional factor binding have been also shown to precede DNA demethylation (Matsuo et al., 1998; Grange et al., 2001; Thomassin et al., 2001).

DNA demethylation mediated by non-transcriptional factors has also been reported. For instance, the Epstein-Barr virus latent replication origin protein EBNA-1 is shown to promote DNA demethylation of its binding sites, oriP (Hsieh, 1999). Subsequent studies of the LacI repressor protein indicate that DNA binding is both necessary and sufficient for targeted DNA demethylation (Lin et al., 2000; Lin and Hsieh, 2001). In these cases,

however, sequences beyond the binding sites are not demethylated, a sharp contrast to the extended demethylation pattern by TnpA and other transcriptional activators. Moreover, although the LacI protein has been shown to be able to prevent DNA methylation at its binding site, there is no evidence for active DNA demethylation by this protein (Lin et al., 2000; Lin and Hsieh, 2001). This is also true of CTCF, a conserved zinc-finger protein that is essential for maintaining imprinting and X-inactivation (Wolffe, 2000; Chao et al., 2002). Distinct from other trans-acting factors, however, CTCF can only maintain its binding sites unmethylated in the X-chromosome that is to be inactivated (Bell and Felsenfeld, 2000; Szabo et al., 2000; Wolffe, 2000; Chao et al., 2002). It is not clear how such allele-specific activity is achieved, as the CTCF protein can function as both a transcription activator (Yang et al., 1999) and repressor (Awad et al., 1999; Ohlsson et al., 2001), dependent upon posttranslational modification (Delgado et al., 1999; Klenova et al., 2001) and interaction with other trans-acting factors that are activators (Chernukhin et al., 2000) or repressors (Lutz et al., 2000).

Since TnpA turns out to be an *Spm* specific transcriptional activator, it is therefore interesting to determine whether this activity is required for its ability to promote *Spm* demethylation. Using truncated TnpA proteins and fusion proteins with a strong transcription activation domain, the present study show a close correlation between the ability of TnpA to activate transcription and the ability to promote DNA demethylation, suggesting a causal relationship between transcriptional activation and *Spm* demethylation.

Materials and Methods

Plant materials and treatments

Callus and suspension cultures as well as treatments were the same as described in previous chapters.

Plasmid Constructs

The construct for expressing FLAG-TnpA has been described previously. All other FLAG-tagged clones were derivatives of pFLAG-TnpA. pFLAG-$TnpA_{540}$ and pFLAG-$TnpA_{420}$ were cloned by ligating the XhoI to SplI fragment from pFLAG-TnpA and the SplI to SalI fragment from pMS199 or pMS198 (Schläppi et al., 1994) (the SalI site was filled in using the DNA polymerase Klenow fragment) together into pBluescript II KS (+) that was already cut with XhoI and EcoRV. pFLAG-$TnpA_{540}$ and pFLAG-$TnpA_{420}$ encode the N-terminal 540 and 420 amino acids of TnpA, respectively, as a result of a nonsense mutation introduced by site-directed mutagenesis at the 3' end of the cDNA for TnpA (Schläppi et al., 1996). pFLAG-$TnpA_{540}$-Vp16 was obtained by ligating the XhoI-SplI fragment of pFLAG-TnpA and the SplI-NcoI fragment (NcoI site was filled in using the Klenow fragment) from pMS249 (Schläppi et al., 1996) together into the pBluescript II KS (+) vector at the XhoI and EcoRV restriction sites.

For inducible protein expression, the XhoI to SpeI fragment from pFLAG-TnpA, pFLAG-$TnpA_{540}$ and pFLAG-$TnpA_{420}$ were cloned into pTA7002 (Aoyama and Chua, 1997), giving rise to pTA-FLAG-TnpA, pTA-FLAG-$TnpA_{540}$ and pTA-FLAG-$TnpA_{420}$, respectively. pTA-FLAG-$TnpA_{540}$-Vp16 was cloned by ligating the ClaI to SpeI fragment

from pFLAG-TnpA$_{540}$-VP16 and pTA7002 that was cut with XhoI and SpeI and the XhoI site was partially filled in to the adenine residue. For constitutive protein expression, the *FLAG-TnpA* cDNA and its deletion derivatives were first cloned as an XhoI to XbaI fragment from pFLAG-TnpA into pAVA120 (von Arnim et al., 1998), yielding pAVA-FLAG-TnpA, pAVA-FLAG-TnpA$_{540}$ and pAVA-FLAG-TnpA$_{420}$, respectively. The expression cassette was then subcloned as a HindIII fragment into pCGN1549 (McBride and Summerfelt, 1990) or pCam1300 (http://www.cambia.org.au/main/r_et_camvec.htm), giving rise to pCGN-35S-FLAG-TnpA, pCam1300-35S-FLAG-TnpA$_{540}$ and pCam1300-35S-FLAG-TnpA$_{420}$, respectively.

Constructs for expression in yeast cells of fusion proteins comprising the Gal4 DNA-binding domain and TnpA, as well as its deletion derivatives, were cloned as follows. The yeast expression vector pAS2 was first cut with NcoI, filled in with Klenow and then cut with SalI. pRR466 (Raina et al., 1993), pMS199 and pMS198 (Schläppi et al., 1994) were cut with BamHI, filled in with Klenow and then cut with SalI to release the cDNAs for TnpA and its deletion derivatives (TnpA$_{540}$ and TnpA$_{420}$), which were then cloned into pAS2, giving rise to pAS2-TnpA, pAS2-199, pAS198, respectively.

Protein Blot Assays

Protein expression in yeast cells was analyzed by Western blotting using a rapid method for lysate preparation (Langlands and Prochownik, 1997) and a polyclonal antibody against the Gal4 binding domain (Santa Cruz Biotech.). The FLAG-tagged TnpA and its deletion derivatives FLAG-TnpA$_{540}$ and FLAG-TnpA$_{420}$ in transgenic plants cells were detected using the polyclonal anti-FLAG antibody (Sigma-Aldrich). Total

plant protein was extracted by grinding plant tissue in 2 volumes of TBS buffer (50 mM Tris-HCl, pH 7.8, 150 mM NaCl) supplemented with 1 mM PMSF and 1x protease inhibitor cocktail for plants (Sigma-Aldrich). The extract was centrifuged at 14,000 rpm at 4° C for 10 min, the supernatant was mixed with 6x SDS loading buffer (Scopes, 1998), boiled for 5 min, and 30 μl was loaded on an 8% polyacrylamide gel for fractionation. Proteins expressed in both yeast and plant cells were visualized using the WesternBreeze™ chemifluorescence immunodetection system (Invitrogen).

Other Assays

The DNA methylation assay and the transcription assay have been described in the Methods of Chapter 2 and Chapter 3, respectively. Gel mobility shift assays were carried out as described (Raina et al., 1998) using an unmethylated DNA fragment with 2 TnpA binding sites in the tail-to-tail orientation (Raina et al., 1993). Conditions for protein immunoprecipitation were described in detail in Chapter 7.

Results

Truncated TnpA Derivatives Lack the Ability to Promote DNA Demethylation

The finding that TnpA is a transcriptional activator prompted me to ask whether this activity is required for the ability of TnpA to promote *Spm* demethylation. SR1-1 plants carrying the methylated *Spm-LUC* construct were crossed with plants carrying the *FLAG-TnpA*$_{540}$ or the *FLAG-TnpA*$_{420}$ construct under the control of the CaMV 35S

promoter. The extent of *Spm* methylation in the F1 plants expressing either truncated protein was almost the same as that in parental SR1-1 plants, confirming the previous report that the C-terminus is essential for TnpA-promoted demethyation (Schläppi et al., 1994). The truncated proteins were also expressed using the dexamethasone-inducible system. Callus tissue derived from transgenic F1 plants from a cross between SR1-1 plants and transgenic plants with inducible expression of the truncated FLAG-TnpA$_{540}$ or

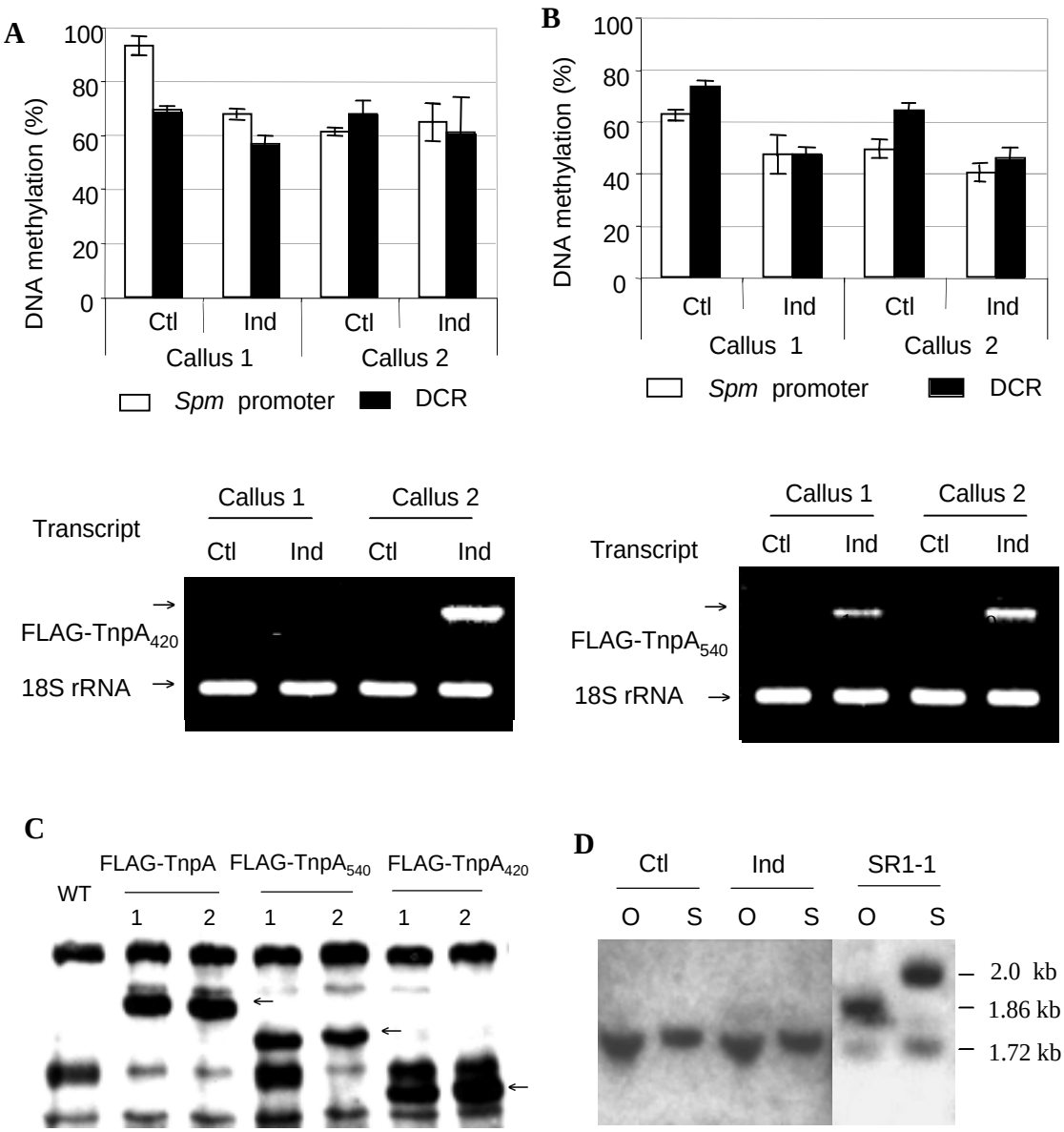


Figure 8. Demethylation of the *Spm* sequence by TnpA deletion derivatives and the TnpA-VP16 fusion protein. (A,B) *Spm* promoter and DCR methylation before (ctl) and after (ind) induction with 10 μM Dex for 20 days in transgenic tobacco calli containing either an inducible (A) FLAG-$TnpA_{420}$ or (B) FLAG-$TnpA_{540}$ construct. The lower panels in A and B show the results of RT-PCR amplification of *FLAG-$TnpA_{420}$* or *FLAG-$TnpA_{540}$* transcripts before and after 48 hrs of induction. The values shown are the means (and standard errors) of measurements from two replicate experiments performed with the same lines of calli at similar growth stages. (C) Western analysis of FLAG-TnpA and its deletion derivatives (indicated by arrows) in extracts of transgenic plants expressing the proteins from a CaMV 35S promoter. (D) *Spm* promoter and DCR methylation, as evidenced by digestion with EcoO109I (O) and SalI (S), respectively, in DNA from transgenic tobacco calli with the FLAG-TnpA-VP16 fusion protein expressed from the Dex-inducible promoter and in the parent SR1 line used for the transformation (the methylated fragments migrate more slowly than the unmethylated fragments, as diagrammed in Fig 2).

FLAG-$TnpA_{420}$ protein was treated with dexamethasone for 20 days. Although the *FLAG-$TnpA_{420}$* gene was expressed in many calli, its expression had no effect on *Spm* methylation (Fig. 8A), whereas expression of the *FLAG-$TnpA_{540}$* gene resulted in a small decrease in *Spm* methylation, particularly in the promoter sequence (Fig. 8B). After a comparable period of induction, calli expressing the full-length FLAG-TnpA showed complete loss of methylation (Fig. 2B, chapter 2). Analysis of proteins by Western blotting using anti-FLAG antibodies showed that TnpA and both of its truncated derivatives are expressed at comparable levels in the transgenic plants that did not show change in DNA methylation (Fig. 7 and Fig. 8C).

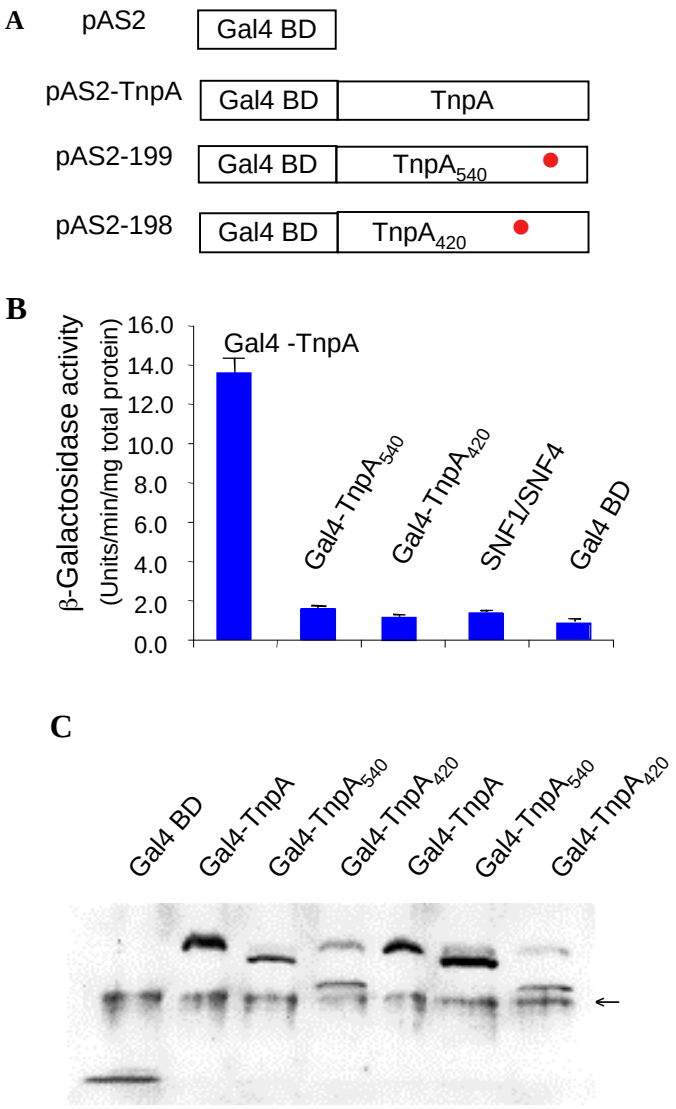
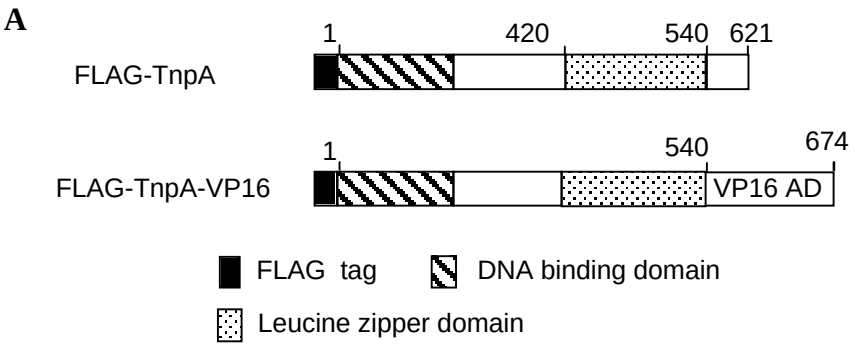


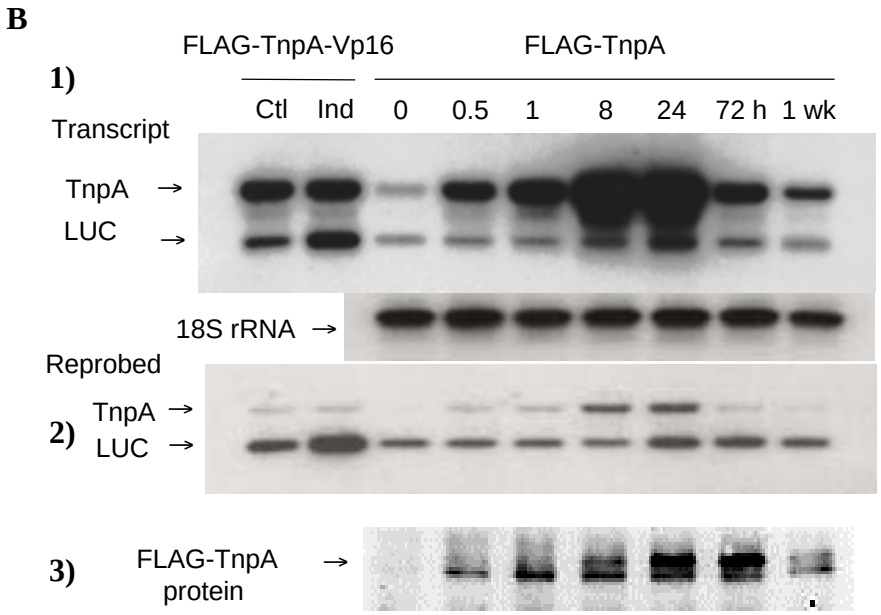
Figure 9. Truncated TnpA proteins do not activate transcription in yeast cells.

(A) Constructs used for expression of Gal4 fusion proteins in yeast cells. ● indicates the position of the stop codon introduced into the *TnpA* cDNA to truncate the proteins at amino acid 420 or 540; the full-length protein is 621 amino acids. (B) Quantitative β-galactosidase assay in transformed yeast cells. Snf1/Snf4 cells express the Gal4 BD-SNF1 and SNF4-Gal4 activation domain fusion proteins and serve as a weak positive control. Represented values are the means (and standard errors) of 3 replicates. (C) Western blot analysis of the Gal4-TnpA fusion protein and its truncated derivatives in yeast using anti-Gal4 BD antibodies. The arrow marks a nonspecific band that reflects total protein loading.

Truncated TnpA Derivatives Lack Transcription Activity

To better compare the transcriptional activity of TnpA and its truncated derivatives, I first measured β-galactosidase activity in yeast cells expressing their GAL4 fusion proteins (Fig. 9A) using a quantitative assay. Neither truncated derivative activated transcription significantly in yeast cells (Fig. 9B). Transcription activity of FLAG-$TnpA_{540}$ and FLAG-$TnpA_{420}$ was then determined using transgenic tobacco lines that showed good inducibility of these truncated proteins and have the same methylated *Spm-LUC* reported construct for studies of TnpA. As shown in Fig. 10C-1, deletion of just the C-terminal 80 amino acids eliminates TnpA's ability to activate the *Spm* promoter in plant cells (Fig. 10 C-1). No transcriptional activity was observed for FLAG-$TnpA_{420}$ either (not shown). The lack of transcriptional activity is not due to protein instability as the truncated proteins were expressed at similar levels as the full-length protein in both yeast and plant cells (Fig. 9C and Fig. 10C-2). Notably, however, there is an accumulation of full-length TnpA protein in the yeast cells, and this becomes more significant when the cell gets old. By contrast, full-length TnpA protein was not detected in plant cells expressing the truncated proteins. Since the truncations were generated by introduction of nonsense codons into the full-length *TnpA* cDNA, the full-length TnpA protein is likely a translational reading-through product. The observed difference in full-length TnpA protein accumulation between yeast and plant cells may reflect their different ability to manage missense mutations. Taken together, the results from this study clearly showed that the C-terminal domain of TnpA is necessary for its ability to both activate transcription and promote DNA demethylation.


A
1
420
540
621
FLAG-TnpA
1
540
674
FLAG-TnpA-VP16
VP16 AD
FLAG tag
DNA binding domain
Leucine zipper domain


B
FLAG-TnpA-Vp16
FLAG-TnpA
1)
Ctl
Ind
0
0.5
1
8
24
72 h
1 wk
Transcript
TnpA
LUC
18S rRNA
Reprobed
2)
TnpA
LUC
3)
FLAG-TnpA protein

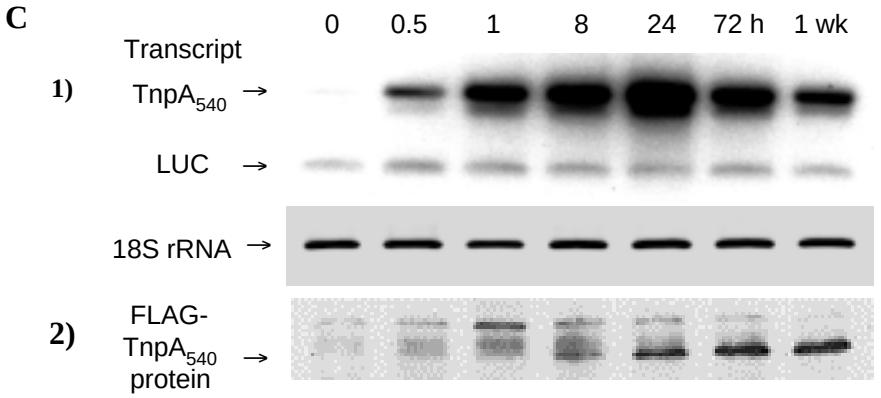

C
0
0.5
1
8
24
72 h
1 wk
Transcript
1)
TnpA540
LUC
18S rRNA
2)
FLAG-TnpA540 protein

Figure 10. Truncated TnpA protein does not activate transcription in tobacco cells.
(A) The top diagram shows the domain structure of the TnpA protein and the positions at which proteins were truncated. The bottom diagram shows the structure of the FLAG-tagged TnpA-VP16 fusion protein. (B-1, C-1) Analysis of *TnpA* and luciferase (*LUC*) transcripts following induction of (B) the *FLAG-TnpA-VP16* and *FLAG-TnpA* genes, and (C) the *FLAG-TnpA*$_{540}$ gene in suspension-cultured transgenic tobacco cells. RNA was amplified by RT-PCR (15 cycles) followed by Southern analysis for signal detection. Equal amounts of *TnpA* and *LUC* probes were used; the 18S rRNA was amplified as a loading control. (B-2) To better reveal the change in *LUC* transcript level in FLAG-TnpA expressing cells, the membrane used in B-1 was stripped and reprobed with only the *LUC* probe; the TnpA bands represent residual label from the initial hybridization. (B-3) and (C-2) Western analysis of proteins in the corresponding cultures using anti-FLAG antibodies. (B) is the same as Fig 7.B and is included here for the convenience of comparison.

Truncated TnpA Derivatives Expressed in Plant Cells Binds to its Cognate DNA

The lack of transcriptional and DNA demethylation activity in the truncated proteins could be due to the loss of DNA binding ability. Although it has been shown that both truncated proteins TnpA$_{540}$ and TnpA$_{420}$ prepared by *in vitro* translation can still bind well to its cognate DNA binding sites derived from the *Spm* promoter sequence (Gierl et al., 1988), whether the *in planta* expressed proteins have the same ability has not yet been determined . To investigate this possibility, I purified the FLAG tagged truncated proteins in their native forms from suspension cultured tobacco cells by immunoprecipitation using agarose bead-coupled anti-FLAG antibodies and by elution with FLAG peptide solution. As shown in Fig. 11, the truncated proteins were still able to bind to DNA containing TnpA binding sites.

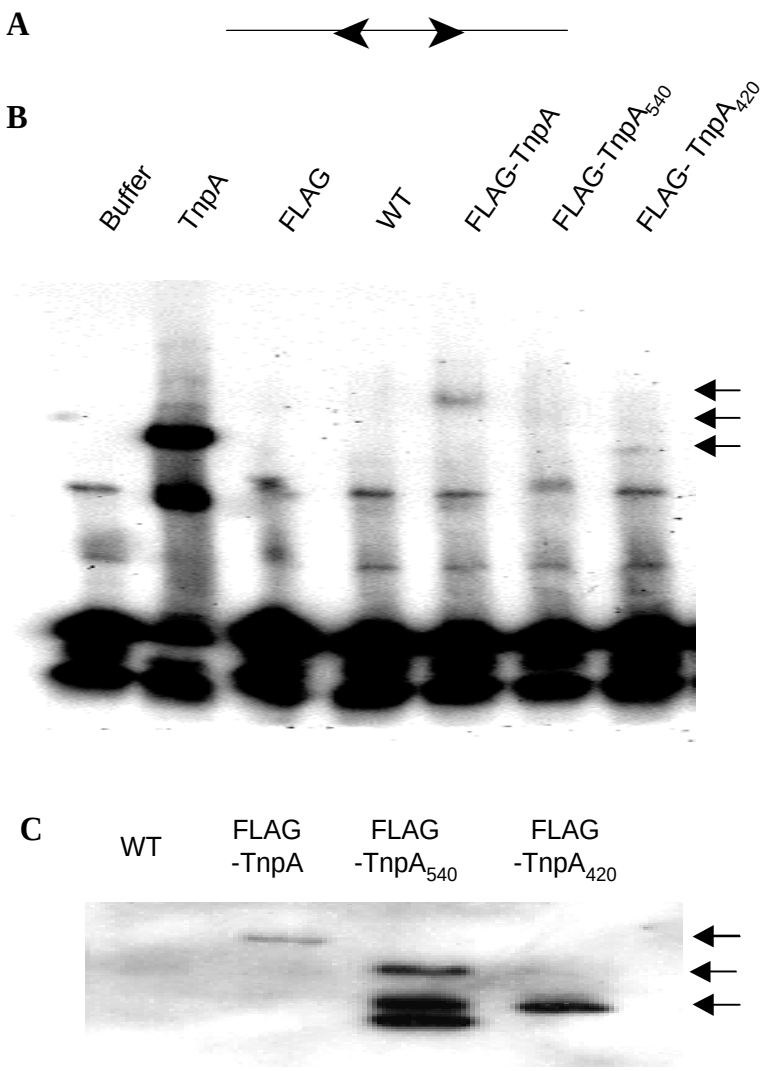


Figure 11. Gel mobility shift assay showing binding of $TnpA_{540}$ and $TnpA_{420}$ to DNA derived from the *Spm* promoter sequence. (A) Diagram of the DNA fragment used for the mobility assay, which contains 2 TnpA binding sites.

(B) Mobility shift assay using immunoprecipitates (IP) prepared from transgenic tobacco cells expressing FLAG-TnpA and its truncated derivatives FLAG-$TnpA_{540}$ and FLAG-$TnpA_{420}$. TnpA, purified TnpA protein expressed in bacteria and used as a positive control. Protein extract buffer (Buffer), the FLAG peptide (FLAG) and an IP prepared from wildtype tobacco extracts (WT) were used as negative controls. The arrows indicate the positions of the DNA-protein complexes. (C) Western blot detection of FLAG tagged proteins in the IP shown in Fig 11.B.

FLAG-TnpA$_{540}$-VP16 fusion protein promotes DNA demethylation

The observed correlation between the loss of TnpA's ability to stimulate transcription and demethylation in the deletion derivatives suggests a causal relationship. I therefore investigated whether the addition of a transcriptional activation domain affects the ability of the truncated proteins to promote *Spm* demethylation. In all calli and regenerated plants expressing the FLAG-TnpA$_{540}$-VP16 fusion protein from the CaMV 35S promoter, the *Spm* sequence is completely unmethylated regardless of the initial methylation level (not shown). The TnpA$_{540}$-VP16 fusion protein was also expressed from the dexamethasone-inducible promoter in transgenic tobacco lines with high levels of *Spm* methylation, as we did for TnpA. None of the lines containing the inducible construct was able to maintain methylation of the *Spm* sequence even in the absence of induction (Figure 8D). Since transgenic tobacco lines containing an inducible *TnpA* gene can maintain *Spm* methylation for a long period of time in the absence of dexamethasone treatment, the consistent absence of *Spm* methylation in lines expressing the FLAG-TnpA$_{540}$-VP16 fusion protein is believed to be attributable to its much greater strength as a transcriptional activator of the *Spm* promoter when compared with intact TnpA. Thus the low level of FLAG-TnpA$_{540}$-VP16 expression in the absence of inducer probably suffices for DNA demethylation. As noted earlier, lines containing the inducible TnpA construct lose *Spm* methylation after several months in culture. These observations suggest that the ability to activate transcription is essential for TnpA's ability to promote *Spm* demethylation.

Discussion

Transcription Activation and DNA Demethylation

Although the ability of DNA methylation to repress transcription is well understood, the connection between transcription activation and DNA demethylation is poorly understood. The present study has revealed a causal connection between transcriptional activation and DNA demethylation. One the one hand, deletion of the C-terminal 80 amino acids of TnpA clearly abolishes both its transcriptional and demethylation activities, although the $TnpA_{540}$ protein is still able to dimerize (Gierl et al., 1988; Trentmann et al., 1993). On the other hand, loss of methylation in the *Spm* sequence was also observed in transgenic tobacco calli and plants containing a *FLAG-$TnpA_{540}$-VP16* gene, even when the fusion protein was expressed from an inducible promoter in the absence of inducer. This shows that the TnpA's transcriptional activation domain can be replaced by a heterologous one and further suggests that the much stronger transcription activation capacity of the fusion protein compared to that of intact TnpA increases its demethylation activity. Because TnpA can activate transcription from an unmethylated *Spm* promoter, as shown in chapter 3, it is believed that DNA demethylation is a consequence, rather a cause, of transcriptional activation by TnpA.

An important aspect of TnpA-mediated DNA demethylation is that it is not confined to the *Spm* promoter, which contains multiple repeats of the TnpA binding site, but also includes the GC-rich sequence encoding the untranslated leader of the *Spm* transcript (Banks et al., 1988; Masson and Fedoroff, 1989). This pattern of DNA demethylation resembles that due to other transcription factors (Matsuo et al., 1998;

Grange et al., 2001; Thomassin et al., 2001) and different from that mediated by the viral DNA replication protein, EBNA-1, and the repressor protein LacI. Neither EBNA-1 nor LacI is able to promote DNA demethylation of sequences beyond their binding sites (Hsieh, 1999; Lin et al., 2000; Lin and Hsieh, 2001). Since the chromatin remodeling that is attendant on transcriptional activation is not limited to the protein binding sites, the extended region of DNA demethylation is consistent with the notion that DNA demethylation is associated with the formation of a transcription initiation complex.

DNA Binding and DNA Demethylation

DNA binding has been shown to be both necessary and sufficient for sequence-specific DNA demethylation mediated by some trans-acting factors such as the EBNA-1 protein (Hsieh, 1999) and the LacI repressor protein (Lin et al., 2000; Lin and Hsieh, 2001). This is apparently not the case for TnpA-mediated demethylation of the *Spm* sequence. Although the DNA binding domain appears essential for determining target sequence specificity, it alone does not promote DNA demethylation (Schläppi et al., 1994). Whether expressed under a constitutive or inducible promoter, neither of the truncated TnpA proteins, $TnpA_{540}$ or $TnpA_{420}$, causes significant reduction in the extent of methylation in the *Spm* sequence. The loss of the ability to promote DNA demethylation in these truncated proteins does not seem to arise from protein instability (Fig. 8C and Fig.10), nor from their inability to bind the *Spm* sequence (Gierl et al., 1988 and Fig. 11). Interestingly, although $TnpA_{540}$ still contains the leucine-zipper domain and the putative glutamine-rich transcription activation motif, it does not show significantly higher transcriptional activity than $TnpA_{420}$. What causes such a dramatic change in this

property of the $TnpA_{540}$ protein is unclear, but improper protein folding is regarded as the most likely explanation. It is worth pointing out, however, that the observed DNA binding ability of the truncated proteins shown by the *in vitro* assay does not necessarily reflect their *in vivo* activity. It is likely that the transcriptional activator domain is required to facilitate the access to DNA in the chromatin context. In any event, the present study supports the conclusion that TnpA's transcription activity is essential for TnpA-mediated *Spm* demethylation.

The Mechanism Underlying TnpA-Mediated DNA Demethylation

The mechanism by which TnpA promotes *Spm* demethylation can be summarized in the model shown in Fig. 12. Binding of TnpA nucleates the assembly of a transcription initiation complex that either contains or attracts additional proteins required for DNA demethylation. Although the possibility has not been definitively ruled out that TnpA itself has a demethylase activity, initial efforts to demonstrate such an activity have not been successful (Chen and Fedoroff, unpublished). Moreover, given the tight binding of TnpA to the *Spm* promoter, but not the DCR, it appears unlikely that the protein itself demethylates the DCR (Raina et al., 1998). *Spm* demethylation is depicted as preceding transcription in Fig. 12, but the temporal relationship between DNA demethylation, chromatin remodeling and transcription *per se* has yet to be examined. It remains to be determined whether TnpA recruits a DNA demethylase (either directly or indirectly) or whether the changes in chromatin structure attendant on formation of the transcription initiation complex simply increase the accessibility of the sequence to such proteins.

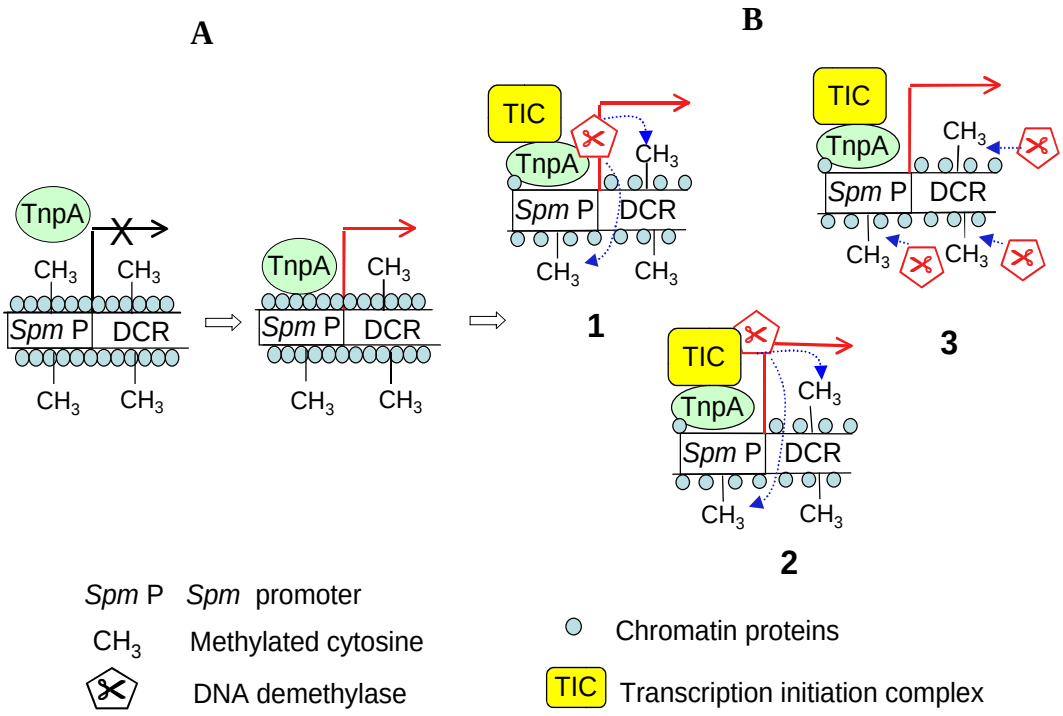


Figure 12. A model depicting the role of TnpA in *Spm* demethylation.

(A) TnpA binds to the *Spm* promoter, assembles a transcription initiation complex.

(B) TnpA either actively recruits additional proteins necessary for demethylation directly (1) or indirectly (2), or restructures chromatin to increase the accessibility of the *Spm* sequence to DNA demethylase (3).

References

Aoyama, T., and Chau, N.H. (1997). A glucocorticoid-mediated transcriptional induction system in transgenic plants. Plant J. **11,** 605-612.

Awad, T.A., Bigler, J., Ulmer, J.E., Hu, Y.J., Moore, J.M., Lutz, M., Neiman, P.E., Collins, S.J., Renkawitz, R., Lobanenkov, V.V., and Filippova, G.N. (1999). Negative transcriptional regulation mediated by thyroid hormone response element 144 requires binding of the multivalent factor CTCF to a novel target DNA sequence. J. Biol. Chem. **274,** 27092-27098.

Banks, J.A., Masson, P., and Fedoroff, N. (1988). Molecular mechanisms in the developmental regulation of the maize Suppressor-mutator transposable element. Genes. Dev. **2,** 1364-1380.

Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature **405,** 482-485.

Chao, W., Huynh, K.D., Spencer, R.J., Davidow, L.S., and Lee, J.T. (2002). CTCF, a candidate trans-acting factor for X-inactivation choice. Science **295,** 345-347.

Chernukhin, I.V., Shamsuddin, S., Robinson, A.F., Carne, A.F., Paul, A., El-Kady, A.I., Lobanenkov, V.V., and Klenova, E.M. (2000). Physical and functional interaction between two pluripotent proteins, the Y-box DNA/RNA-binding factor, YB-1, and the multivalent zinc finger factor, CTCF. J. Biol. Chem. **275,** 29915-29921.

Delgado, M.D., Chernukhin, I.V., Bigas, A., Klenova, E.M., and Leon, J. (1999). Differential expression and phosphorylation of CTCF, a c-myc transcriptional regulator, during differentiation of human myeloid cells. FEBS Lett. **444,** 5-10.

Gierl, A., Lutticke, S., and Saedler, H. (1988). TnpA product encoded by the transposable element En-1 of Zea mays is a DNA binding protein. EMBO J. **7,** 4045-4053.

Grange, T., Cappabianca, L., Flavin, M., Sassi, H., and Thomassin, H. (2001). In vivo analysis of the model tyrosine aminotransferase gene reveals multiple sequential steps in glucocorticoid receptor action. Oncogene **20,** 3028-3038.

Hsieh, C.L. (1999). Evidence that protein binding specifies sites of DNA demethylation. Mol. Cell. Biol. **19,** 46-56.

Kirillov, A., Kistler, B., Mostoslavsky, R., Cedar, H., Wirth, T., and Bergman, Y. (1996). A role for nuclear NF-kappaB in B-cell-specific demethylation of the Igkappa locus. Nat. Genet. **13,** 435-441.

Klenova, E.M., Chernukhin, I.V., El-Kady, A., Lee, R.E., Pugacheva, E.M., Loukinov, D.I., Goodwin, G.H., Delgado, D., Filippova, G.N., Leon, J., Morse, H.C., 3rd, Neiman, P.E., and Lobanenkov, V.V. (2001). Functional phosphorylation sites in the C-terminal region of the multivalent multifunctional transcriptional factor CTCF. Mol. Cell. Biol. **21,** 2221-2234.

Langlands, K., and Prochownik, E.V. (1997). A rapid method for the preparation of yeast lysates that facilitates the immunodetection of proteins generated by the yeast two-hybrid system. Anal. Biochem. **249,** 250-252.

Lin, I.G., Tomzynski, T.J., Ou, Q., and Hsieh, C.L. (2000). Modulation of DNA binding protein affinity directly affects target site demethylation. Mol. Cell. Biol. **20,** 2343-2349.

Lin, I.G., and Hsieh, C.L. (2001). Chromosomal DNA demethylation specified by protein binding. EMBO Rep. **2,** 108-112.

Lutz, M., Burke, L.J., Barreto, G., Goeman, F., Greb, H., Arnold, R., Schultheiss, H., Brehm, A., Kouzarides, T., Lobanenkov, V., and Renkawitz, R. (2000). Transcriptional repression by the insulator protein CTCF involves histone deacetylases. Nucleic Acids Res. **28,** 1707-1713.

Masson, P., and Fedoroff, N.V. (1989). Mobility of the maize suppressor-mutator element in transgenic tobacco cells. Proc. Natl. Acad. Sci. USA **86,** 2219-2223.

Matsuo, K., Silke, J., Georgiev, O., Marti, P., Giovannini, N., and Rungger, D. (1998). An embryonic demethylation mechanism involving binding of transcription factors to replicating DNA. EMBO J. **17,** 1446-1453.

McBride, K.E., and Summerfelt, K.R. (1990). Improved binary vectors for Agrobacterium-mediated plant transformation. Plant Mol. Biol. **14,** 269-276.

Ohlsson, R., Renkawitz, R., and Lobanenkov, V. (2001). CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. Trends Genet. **17,** 520-527.

Raina, R., Cook, D., and Fedoroff, N. (1993). Maize Spm transposable element has an enhancer-insensitive promoter. Proc. Natl. Acad. Sci. USA **90,** 6355-6359.

Raina, R., Schläppi, M., Karunanandaa, B., Elhofy, A., and Fedoroff, N. (1998). Concerted formation of macromolecular Suppressor-mutator transposition complexes. Proc. Natl. Acad. Sci. USA **95,** 8526-8531.

Schläppi, M., Raina, R., and Fedoroff, N. (1994). Epigenetic regulation of the maize Spm transposable element: novel activation of a methylated promoter by TnpA. Cell **77,** 427-437.

Schläppi, M., Raina, R., and Fedoroff, N. (1996). A highly sensitive plant hybrid protein assay system based on the Spm promoter and TnpA protein for detection and analysis of transcription activation domains. Plant Mol. Biol. **32,** 717-725.

Scopes, R.a.S.J. (1998). Analysis of Proteins. Current Protocols in Molecular Biology. **2.,** 10.10.11-10.10.20.

Silke, J., Rother, K.I., Georgiev, O., Schaffner, W., and Matsuo, K. (1995). Complex demethylation patterns at Sp1 binding sites in F9 embryonal carcinoma cells. FEBS Lett. **370,** 170-174.

Szabo, P., Tang, S.H., Rentsendorj, A., Pfeifer, G.P., and Mann, J.R. (2000). Maternal-specific footprints at putative CTCF sites in the H19 imprinting control region give evidence for insulator function. Curr. Biol. **10,** 607-610.

Thomassin, H., Flavin, M., Espinas, M.L., and Grange, T. (2001). Glucocorticoid-induced DNA demethylation and gene memory during development. EMBO J. **20,** 1974-1983.

Trentmann, S.M., Saedler, H., and Gierl, A. (1993). The transposable element *En/Spm*-encoded TNPA protein contains a DNA binding and a dimerization domain. Mol. Gen. Genet. **238,** 201-208.

von Arnim, A.G., Deng, X.W., and Stacey, M.G. (1998). Cloning vectors for the expression of green fluorescent protein fusion proteins in transgenic plants. Gene **221,** 35-43.

Wolffe, A.P. (2000). Transcriptional control: imprinting insulation. Curr. Biol. **10,** R463-465.

Yang, Y., Quitschke, W.W., Vostrov, A.A., and Brewer, G.J. (1999). CTCF is essential for up-regulating expression from the amyloid precursor protein promoter during differentiation of primary hippocampal neurons. J. Neurochem. **73,** 2286-2298.

Chapter 5

DNA REPLICATION IS REQUIRED FOR TNPA-MEDIATED *SPM* DEMETHYLATION

Introduction

Trans-acting factor-mediated DNA demethylation can be achieved through two alternative pathways. Passive DNA demethylation results from blocking DNA binding by DNA methyltransferases to newly synthesized DNA, and thus preventing remethylation after DNA replication. By contrast, active DNA demethylation involves a DNA demethylase, which is recruited to the target sequence by the trans-acting factors. These two conceptually different mechanisms can both be involved in the same DNA demethylation process mediated by a specific trans-acting factor, however, as has been suggested for the NF-kappa B transcriptional factor (Matsuo et al., 1998) and the Epstein-Barr virus latent replication origin protein EBNA-1 (Hsieh, 1999).

The experiments described in Chapter 2 clearly demonstrated that TnpA-mediated demethylation of the *Spm* sequence involves an active mechanism. Although the whole process is very rapid, DNA demethylation is rather slow initially. For example, after 36 hrs of TnpA induction with the inducible system, when the *TnpA* transcript level is almost maximal and a high level of TnpA protein has already accumulated in the cells, the extent of DNA demethylation is only 30-40%. Nearly complete DNA demethylation

is expected if a purely active mechanism is involved. This observation suggests that DNA replication is required as well in TnpA-mediated DNA demethylation. Supporting this view, DNA demethylation was not observed after induction of TnpA in leaves in which cell division has ceased.

To investigate this possibility, I determined the role of DNA replication in TnpA-mediated DNA demethylation using inhibitors of DNA replication and cell division. The results show that *Spm* demethylation is inhibited when DNA replication is arrested. I further show that TnpA has a much-reduced binding affinity for the methylated *Spm* promoter sequence when compared with either the unmethylated or hemimethylated promoter. The requirement for TnpA to initiate DNA demethylation can thus be explained by the low affinity of TnpA for fully methylated DNA.

Materials and Methods

Plant Material and Treatments

Unless specified, all plant materials and treatments were the same as described in previous chapters. Inhibition of DNA replication in callus was achieved by treatment with a combination of aphidicolin and olomoucine (Sigma-Aldrich) at concentrations of 15 μM and 50 μM, respectively. The efficiency of DNA replication inhibition was evaluated by monitoring incorporation of bromodeoxyuridine (BrdU) (Sigma-Aldrich) into DNA using a mouse monoclonal antibody to BrdU (Sigma-Aldrich), alkaline phosphatase-conjugated antibodies to mouse IgG (Sigma), and the WesternBreeze™

chemifluorescence immunodetection system (Invitrogen).

Gel Mobility Shift Assay

Gel mobility shift assays were carried out as described (Raina et al., 1998). DNA sequences with different numbers of TnpA binding sites and differing methylation states were prepared by PCR amplification from plasmids containing either synthetic oligonucleotides or *Spm* promoter fragments (Raina et al., 1993). Primer KS (5'TCGAGGTCGACGGTATC3') and DCR 3' primer (5'CACGACGGCTGTAGA3') or SK (5'CGCTCTAGAACTAGTGGATC3') were used for amplification of DNA fragments with 1 and 6 (derived from promoter deletion) or 2 TnpA binding sites (cloned synthetic oligonucleotide). Unmethylated and fully methylated DNA sequences were prepared by PCR amplification with either dCTP (Life Technologies) or 5m-dCTP (Roche Applied Science) respectively, in the reaction. Hemimethylated DNA fragments were prepared by annealing complementary ssDNA with one unmethylated and the other fully methylated, generated by asymmetric PCR. Optimal ssDNA amplification was achieved when the asymmetric PCR was run for a total of 40 cycles with DNA template at a concentration of 1 ng/50 μl and primer concentrations of 2 μM and 0.2 μM, respectively. 3.5% NuSieve 3:1 agarose (BMA) and 8% native PAGE gel were used for ssDNA analysis and preparation, respectively.

DNA was labeled by preincubation with Klenow and dTTP at 30° C for 15 min, followed by addition of dATP, dGTP and ^{32}P-dCTP (10 μCi in a 20 μl reaction) and an additional 15-min incubation. All DNA preparations were purified using Qiagen columns.

Other Assays

The DNA methylation assay and the transcription assay have been described in the Methods of Chapter 2 and Chapter 3, respectively.

Results

Inhibitors of DNA Replication and Cell Division Interfere with TnpA-Mediated *Spm* Demethylation

Although TnpA expression in tobacco cells is nearly maximal 36 hrs after induction, *Spm* demethylation is far from complete at that time. This could be the result of a slow demethylation process, the lack of TnpA expression in some cells, or a requirement that DNA replicate before demethylation can occur, as suggested by the observation that demethylation can be induced by dexamethasone in callus cultures, but not in the leaves from which they were derived. I therefore asked whether interfering with DNA synthesis and cell division interferes with *Spm* demethylation. Calli containing a methylated *Spm-LUC* transgene and a dexamethasone-inducible *FLAG-TnpA* gene were treated with aphidicolin, a DNA synthesis inhibitor (Sala et al., 1980), and olomoucine, an inhibitor of the Cdc2/Cdk2 kinase that blocks the G1-S cell cycle transition (Glab et al., 1994). The efficiency of DNA replication inhibition was examined by monitoring the incorporation of bromodeoxyuridine (BrdU) into newly replicated DNA using anti-BrdU antibody. At the concentration used in the experiments, DNA replication was completely

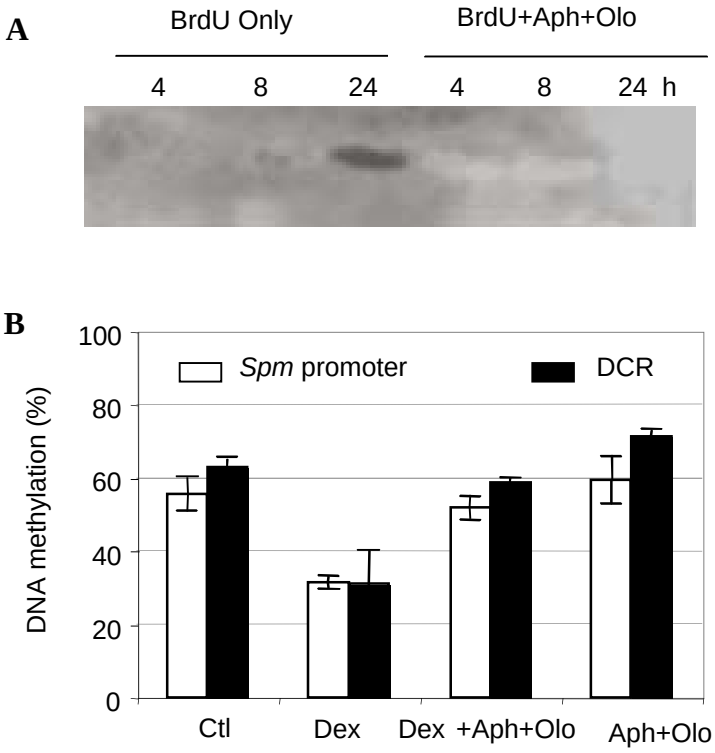


Figure 13. Inhibitors of DNA synthesis and cell division interfere with TnpA-mediated *Spm* demethylation. (A) Western analysis of bromodeoxyuridine (BrdU) incorporation into DNA following addition of BrdU to tobacco callus without (BrdU only) or with aphidicolin (15 μM) and olomoucine (50 μM) (BrdU+Aph+Olo). (B) *Spm* promoter and DCR methylation in transgenic tobacco calli treated with aphidicolin (Aph) and olomoucine (Olo) for 2 days, with or without dexamethasone (Dex) induction of TnpA expression. The values shown are the means (and standard errors) of measurements from two replicate experiments performed on the same callus lines at a similar growth stage.

arrested (Fig. 13A). I then compared the levels of *Spm* methylation after 2 days of treatment with the inhibitors, with dexamethasone alone, and with a combination of all three compounds. Treatment of calli with aphidicolin and olomoucine slightly enhanced *Spm* methylation, while treatment with dexamethasone alone resulted in *Spm*

demethylation (Fig. 13B). Treatment of calli with aphidicolin and olomoucine prevented dexamethasone-induced TnpA-mediated *Spm* demethylation. This suggests that DNA replication is a rate-limiting step in TnpA-mediated demethylation of the *Spm* sequence.

TnpA Binds More Strongly to Unmethylated and Hemimethylated than to Methylated DNA

It has been reported previously that DNA methylation reduces the affinity of TnpA for oligonucleotides containing its consensus 12-bp binding site, CCGACACTCTTA (Gierl et al., 1988). To assess the effect of DNA methylation on the binding affinity of DNA fragments derived from the *Spm* promoter, I compared TnpA binding to unmethylated, hemimethylated and fully methylated synthetic oligonucleotides and DNA fragments derived from the 5' end of the *Spm* transposon (Raina et al., 1998). The electrophoretic mobility shift assay was used to compare the binding of DNA fragments containing 1, 2 and 6 TnpA binding sites. Fully methylated derivatives, as well as both hemimethylated substrates were prepared by the method developed in this study as described (see Methods).

TnpA bound strongly to unmethylated and hemimethylated DNA fragments, irrespective of which strand was methylated. For the fragment containing 1 binding site, the amount of bound protein was slightly greater for the unmethylated fragment than for the hemimethylated fragments, but the difference was not large (Fig. 14A). Similar results were obtained with synthetic oligonucleotides (2 binding sites) and with *Spm*-derived DNA fragments (6 binding sites) (Fig. 14B and 14C). These observations are consistent with the previous report that TnpA's affinity for fully methylated

oligonucleotides containing 1 or 2 TnpA binding sites is less than for hemi- or un-methylated oligonucleotides (Gierl et al., 1988). However, more of the promoter-derived fragment containing 6 TnpA binding sites was in DNA-protein complexes than of the DNA fragments containing 1 or 2 binding sites. *Spm* promoter-derived fragments with multiple binding sites formed large complexes, spreading the labeled DNA over a large fraction of the gel (Fig. 14C).

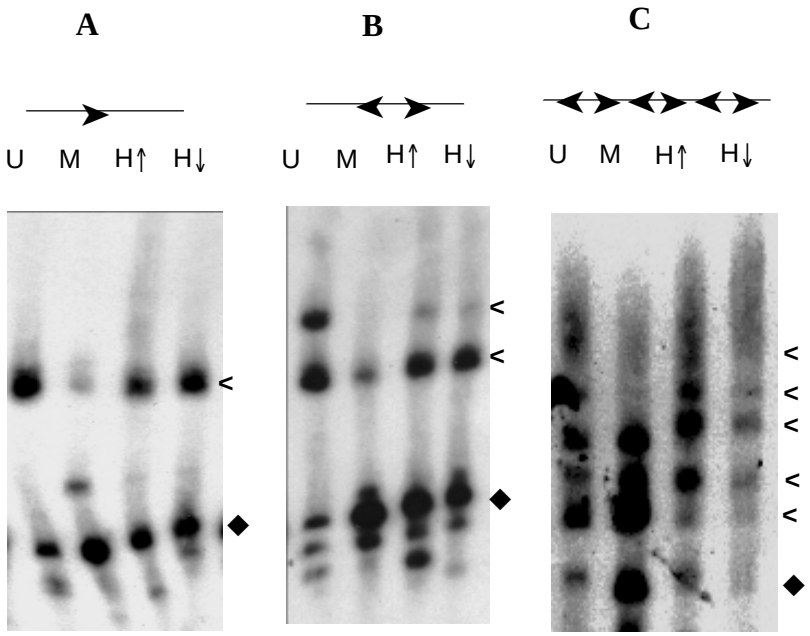


Figure 14. Electrophoretic mobility shift assay of TnpA binding to DNA fragments with (A) one, (B) two and (C) six TnpA binding sites (➤). The DNA fragments were unmethylated (U), fully methylated (M), methylated in the upper strand (H↑), or methylated in the lower strand (H↓).

◆Free DNA; < TnpA-bound DNA**.**

The formation of such large complexes has been reported previously, and it has been noted that the concentration of TnpA required for binding decreases with increasing

number of binding sites per DNA fragment (Raina et al., 1998). In the present experiments, fully methylated DNA fragments bound TnpA less effectively than hemi-methylated DNA fragments in all cases. However, methylated DNA fragments containing 6 TnpA binding sites bound a larger fraction of the input DNA than those with fewer binding sites. Nonetheless, the TnpA-DNA complexes that formed with methylated DNA migrated faster than those formed with unmethylated or hemimethylated fragments, suggesting that they are smaller. This, in turn, indicates that fewer binding sites per methylated DNA molecule were occupied by TnpA. These results are consistent with the observation that DNA replication and cell cycle inhibitors interfere with TnpA-mediated demethylation and support the interpretation that DNA replication is necessary to produce the hemimethylated *Spm* sequence to which TnpA can bind.

Discussion

DNA can be demethylated by interfering with maintenance methylation or by active elimination of methyl groups, methylated bases or methylated nucleotides (Li et al., 1992; Jost et al., 1995; Vairapandi and Duker, 1996; Bhattacharya et al., 1999). DNA demethylation by interference with remethylation of newly replicated DNA is a slow process. It requires two rounds of replication to produce the first fully unmethylated daughter molecules and further reduces the amount of methylated DNA by just twofold at each subsequent replication. The results of the present study reveal that demethylation of

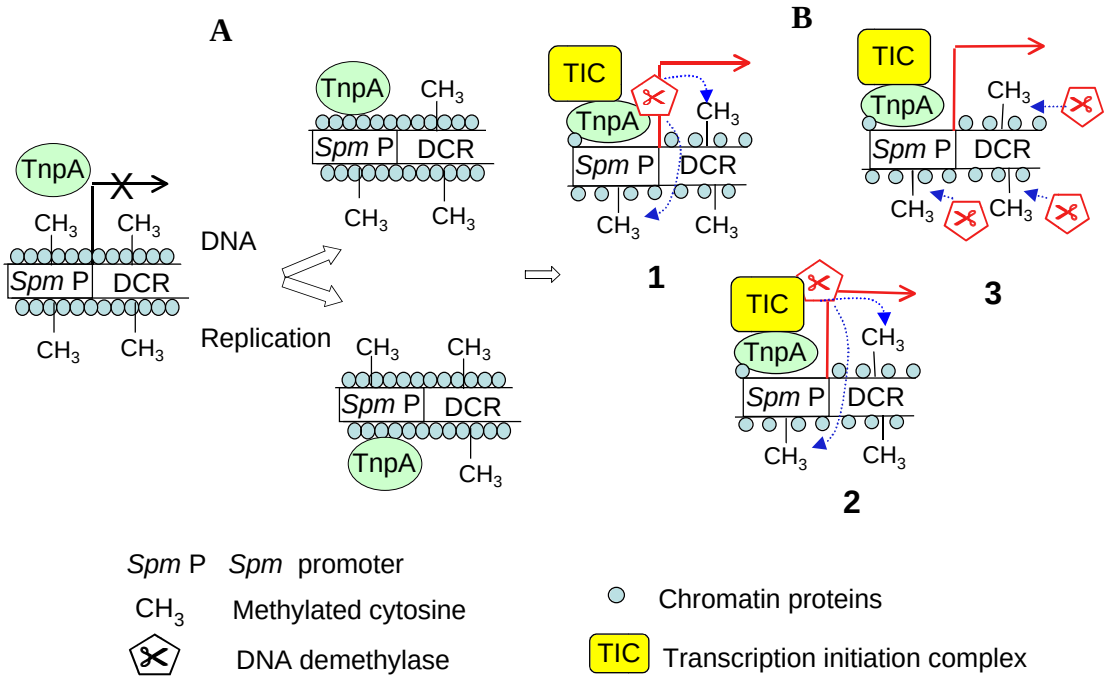


Figure 15. A modified model depicting the role of TnpA in DNA demethylation. (A) TnpA does not bind the fully methylated *Spm* sequence. (B) In dividing cells, TnpA binds to newly replicated, hemimethylated DNA. (C) TnpA assembles a transcription initiation complex and either actively recruits additional proteins necessary for demethylation directly (1) or indirectly (2), or restructures chromatin to increase the accessibility of the *Spm* sequence to demethylation (3).

the *Spm* sequence occurs much more rapidly than can be accounted for by interference with maintenance methylation. However, inhibition of DNA replication and cell cycle progression interferes with the ability of TnpA to demethylate DNA. The likely resolution of this dichotomy lies in the much greater affinity of TnpA for unmethylated and hemimethylated DNA than for fully methylated DNA (Fig. 14). Since demethylation commences soon after TnpA induction in unsynchronized cell cultures, it seems likely that TnpA binding to newly replicated molecules is followed rapidly by demethylation

and that DNA replication is the rate-limiting step for TnpA-mediated demethylation. This interpretation is consistent with our observation that TnpA expression has no effect on *Spm* methylation in leaves, where cell division has ceased. Fig.15 is a revised model, illustrating the two-step mechanism for TnpA-mediated demethylation of the *Spm* sequence.

References

Bhattacharya, S.K., Ramchandani, S., Cervoni, N., and Szyf, M. (1999). A mammalian protein with specific demethylase activity for mCpG DNA [see comments]. Nature **397,** 579-583.

Gierl, A., Lutticke, S., and Saedler, H. (1988). TnpA product encoded by the transposable element En-1 of Zea mays is a DNA binding protein. EMBO J. **7,** 4045-4053.

Glab, N., Labidi, B., Qin, L.X., Trehin, C., Bergounioux, C., and Meijer, L. (1994). Olomoucine, an inhibitor of the cdc2/cdk2 kinases activity, blocks plant cells at the G1 to S and G2 to M cell cycle transitions. FEBS Lett. **353,** 207-211.

Hsieh, C.L. (1999). Evidence that protein binding specifies sites of DNA demethylation. Mol. Cell. Biol. **19,** 46-56.

Jost, J.P., Siegmann, M., Sun, L., and Leung, R. (1995). Mechanisms of DNA demethylation in chicken embryos. Purification and properties of a 5-methylcytosine-DNA glycosylase. J. Biol. Chem. **270,** 9734-9739.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell **69,** 915-926.

Matsuo, K., Silke, J., Georgiev, O., Marti, P., Giovannini, N., and Rungger, D. (1998). An embryonic demethylation mechanism involving binding of transcription factors to replicating DNA. EMBO J. **17,** 1446-1453.

Raina, R., Cook, D., and Fedoroff, N. (1993). Maize Spm transposable element has an enhancer-insensitive promoter. Proc. Natl. Acad. Sci. USA **90,** 6355-6359.

Raina, R., Schläppi, M., Karunanandaa, B., Elhofy, A., and Fedoroff, N. (1998). Concerted formation of macromolecular Suppressor-mutator transposition complexes. Proc. Natl. Acad. Sci. USA **95,** 8526-8531.

Sala, F., Parisi, B., Burroni, D., Amileni, A.R., Pedrali-Noy, G., and Spadari, S. (1980). Specific and reversible inhibition by aphidicolin in the alpha-like DNA polymerase of plant cells. FEBS Lett. **117,** 93-98.

Vairapandi, M., and Duker, N.J. (1996). Partial purification and characterization of human 5-methylcytosine-DNA glycosylase. Oncogene **13,** 933-938.

Chapter 6

DNA DEMETHYLASE ACTIVITY IS PRESENT IN DIVIDING TOBACCO CELLS

Introduction

Programmed change in DNA methylation is essential for normal growth and development of both animals and plants. In animals, active DNA demethylation has been well documented and several proteins seem to be able to convert 5-methylated cytosine (5mC) to cytosine through distinct mechanisms. MBD2, a methyl-DNA binding protein, is shown to be capable of removing the methyl group from 5mC, thus yielding unmethylated DNA (Bhattacharya et al., 1999). Although the existence of such a *bona fide* DNA demethylase has gained further support from subsequent studies from the same lab (Cervoni et al., 1999; Ramchandani et al., 1999; Detich et al., 2002), it has not been replicated by others, which have identified MBD2 as a transcription repressor in a large transcription repression complex (Ng et al., 1999; Zhang et al., 1999). By contrast, an accumulating body of evidence indicates that a major mechanism of active DNA demethylation is through the removal of the methylated base from 5mC by a 5 methylcytosine DNA glycosylase (5MCDG) activity and subsequent replacement of the abasic site by a cytosine via the DNA repair pathway (Jost and Jost, 1995; Vairapandi and Duker, 1996; Zhu et al., 2000). DNA demethylation could also occur as an indirect

consequence of 5mC deamination, which gives rise to a T/G mismatch pair and is the primary cause of mutation in methylated DNA (Wang et al., 1982). There is evidence that MBD1, a protein with a N-terminal methyl-binding domain and a C-terminal DNA glycosylase domain (Wu et al., 2003), is involved in repairing the G/T mismatch pair (Hendrich et al., 1999; Petronzelli et al., 2000). More recently, an oxidation DNA demethylation mechanism has been proposed for the bacterial AlkB protein, which appears to have a preference for methylated DNA in single-stranded DNA (Falnes et al., 2002; Trewick et al., 2002). A similar mechanism might have evolved in higher organisms, as homologs of the *AlkB* gene have been identified in animals.

Although there are cytological observations that suggest the existence of DNA demethylase activities in plant cells (Oakeley and Jost, 1996), no biochemical evidence has ever been obtained. The present study of TnpA-mediated DNA demethylation also indicates that there are DNA demethylases in plants. To investigate this possibility as well as the mechanism by which TnpA promotes *Spm* demethylation, I developed an *in vitro* assay for sensitive detection of DNA demethylase activity. I also optimized the conditions for preparing nuclear extracts from tobacco suspension cells. A DNA demethylase activity was indeed detected in the nuclear extracts. The *in vitro* DNA demethylase assay will be especially useful for further studies of the mechanism underlying TnpA-mediated DNA demethylation.

Materials and Methods

Plant materials

All plant materials used in this study and treatments have been described in previous chapters.

Whole Cell and Nuclear Extract Preparation

Whole cell extracts were prepared from tobacco suspension cells essentially following the protocol of (Cooke and Penon, 1990). A protoplast-based method was used for preparation of tobacco nuclear extracts with some modifications (Roberts and Okita, 1991). Suspension culture was filtered through a layer of miracloth (Calbiochem-Novabiochem Corporation, La Jolla, CA). 50 g cells were washed 2-3 times with protoplast medium (10 mM MES-KOH, pH5.8, 0.4 M D-mannitol) and then resuspended at 4 ml/g cells in protoplast medium containing 1% cellulase "Onozuka R-10" and 0.1% macerozyme (SERVA product, Crescent Chemical Co, Inc, Islandia, NY). Cells were incubated at 28 °C for 90 min on a shaker (100 rpm) and protoplasts were collected by centrifugation at 2500 g for 10 min at 4°C. After washing 3 times with protoplast medium, the pellet was resuspended in 1 vol of ice-cold lysis buffer (25 mM MES-KOH, pH5.8, 18% sucrose, 0.15 mM spermine, 0.15 mM spermidine, 3 mM PMSF, 0.2 μg/ml pepstatin, 2 μg/ml leupeptin, 10 mM β-mercaptoethanol and 20 mM NaF). The mixture was vacuum-filtered through one layer of nylon mesh (75 μM pore size), and the filtrate was centrifuged at 2500 g for 10 min at 2 °C. The pelleted nuclei were gently resuspended in 20 ml of 0.5X protoplast medium and, centrifuged at 2500 g for 5 min at

4 °C. The nuclei were then resuspended immediately in 2 vol of nuclear extract buffer (25 mM HEPES-KOH, pH7.9, 25% glycerol, 4mM $MgSO_4$, 0.2mM EDTA, 1mM EGTA, 10mM NaF, 10 mM β-mercaptoethanol, 3 μg/ml pepstatin, 2 μg/ml leupeptin and 1mM PMSF). 4M ammonium sulfate was added to a final concentration of 0.42 M and the suspension was shaken gently at 4 °C for 30 min, then centrifuged at 200,000 g for 1 hr at 2 °C. The supernatant was dialyzed for 2-3 hrs immediately at 4°C against 500 ml dialysis buffer (20 mM HEPES-KOH, pH7.9, 20% glycerol, 100 mM KOAc and 0.2 mM EGTA, 0.1 mM EDTA, 10 mM β-mercaptoethanol and 0.5 mM PMSF). After two changes of dialysis buffer, the extract was centrifuged at 10,000 g for 5 min at 2 °C. The supernatant was aliquoted and frozen in liquid N2 and stored at –80 °C.

DNA demethylation assay

The substrate for the *in vitro* DNA demethylation assay is a hemimethylated DNA fragment containing the *Spm* promoter and DCR. It was generated by a biotin/streptavidin-based method described in Chapter 2. To determine DNA demethylase activity, 1-5 ng ^{32}P labelled hemimethylated DNA was incubated with protein extracts at 37 °C for 4 hrs in a 50 μl reaction containing 25 mM HEPES, pH7.6, 40 mM KCl, 1 mM $MgCl_2$, 0.01mM $ZnCl_2$, 0.5 mM EDTA, 0.5 mM DTT, 0.5mM ATP, 0.1% BSA and 1 mM dNTPs. 2.5 μg pBluescript plasmid DNA was added into the reaction as carrier either at the start or end of the reaction. DNA was then purified using a Qiagen PCR purification column and eluted in 50 μl ddH_2O. 10 μl of the eluate was digested with 2-4 U SalI in a 20 μl reaction for 2 hrs at 37 °C and the digest was resolved on a 1.8% Metaphor® agarose gel (BMA, Rockland, ME) in 1X TBE buffer. The gel was dried and

expose to X-ray film. Where specified, endonuclease IV (MBI Fermentas) was also used to nick the DNA at abasic sites.

Results

Development of an *in vitro* DNA demethylation assay

My previous experiments suggested that DNA demethylase activities participate in TnpA-mediated DNA demethylation (Chapater 2). To further understand the underlying mechanism, I need a DNA demethylation assay that permits sensitive detection of DNA demethylase activity. As suggested by the observation that rapid DNA demethylation occurs after one cycle of DNA replication, hemimethylated *Spm* DNA might be the preferred substrate for the DNA demethylase activiti(es) involved in TnpA-mediated DNA demethylation (Chapter 5). I therefore used hemimethylated DNA containing the *Spm* promoter and DCR sequence as the substrate for an *in vitro* DNA demethylation assay.

DNA was methylated *in vitro* using the bacterial methylase SssI, which catalyzes the formation of 5-methylcytosine (5mC) in the CpG dinucleotide context. Hemimethylated DNA prepared by this method contains only one 5mC in the SalI recognition site. Since removal of the single 5mC in the SalI site will render the DNA sensitive to SalI restriction, SssI methylated DNA should permit sensitive detection of low DNA demethylase activity.

Surprisingly, however, the hemimethylated DNA is not completely resistant to SalI restriction (Fig. 4E). This is not due to instability of methylated DNA under the conditions used for the preparation of ssDNA, as shown previously (Fig 4D). Notably, however, more SalI enzyme was needed for complete digestion of hemimethylated DNA than unmethylated DNA, indicating that hemimethylated DNA is partially resistant to SalI restriction. I therefore reasoned that for a mixture of hemimethylated and unmethylated DNA, a range of SalI concentrations could be defined within which the unmethylated DNA would be completely restricted, whereas the hemimethylated DNA

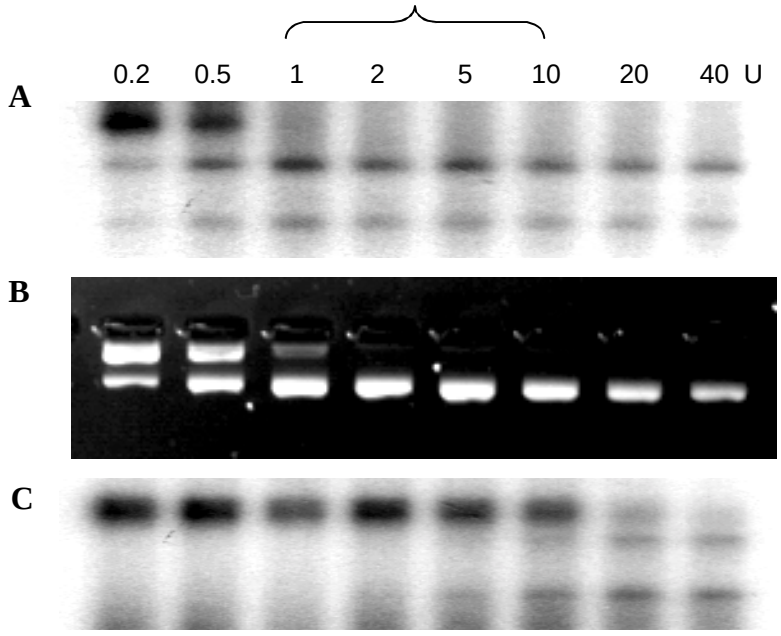


Figure 16. Optimization of conditions for an *in vitro* DNA demethylation assay. Restriction of ^{32}P-labeled (A) unmethylated and (C) hemimethylated substrate DNA by various amounts of SalI enzyme in a 20 μl reaction containing 0.5 μg unmethylated pBluescript plasmid DNA, the digestion pattern of which is shown in (B). The bracket indicated the amount of SalI enzyme (1-5 U) within which DNA demethylation can be performed.

would remain intact.

To test this possibility, I added trace amount of radioactively labeled substrate DNA (about 1 ng), either unmethylated or hemimethylated, into a restriction reaction that contains a larger amount of unlabeled pBluescript plasmid DNA (0.5 μg), which contains a single SalI site. As shown in Fig.16 A and B, the radioactively labeled unmethylated DNA and the pBluescript plasmid were equally digested at the same SalI concentrations and complete restriction was achieved at 1 U of SalI. By contrast, at this concentration of SalI radioactively labeled hemimethylated DNA was intact and its complete restriction was not observed until SalI was increased to 10 U (Fig. 16 C). These results indicate that under the specified assay condition, unmethylated DNA arising from demethylation of hemimethylated DNA could be reliably detected by SalI restriction within the range of 2-5 U per reaction.

DNA Demethylase Activity is Present in Tobacco Cultured Cells

To determine whether there is a DNA demethylase activity in plant cells, I used nuclear extracts prepared from wildtype tobacco suspension cells because of the following considerations. As DNA demethylation is a nuclear event, TnpA and other proteins involved in this process should be concentrated in a nuclear extract, thus facilitating subsequent protein purification. Using a nuclear extract should also help to minimize protein degradation and enzyme inactivation due to contamination by phenolic compounds in the cytoplasm (Lax et al., 1986; Chen et al., 1993).

Initially I used two methods for nuclear extract preparation, one by grinding the whole cells in liquid nitrogen and the other by first preparing protoplast (see Methods).

The quality of these preparations was then evaluated by assaying DNase activity. As shown in Fig 17A, whereas very little DNase activity was detected in the nuclear extract prepared by the first method, strong DNase activity was present in the nuclear extract prepared using the protoplast protocol. For subsequent DNA demethylation assay, I have therefore used the protoplast-based method for nuclear extract preparation.

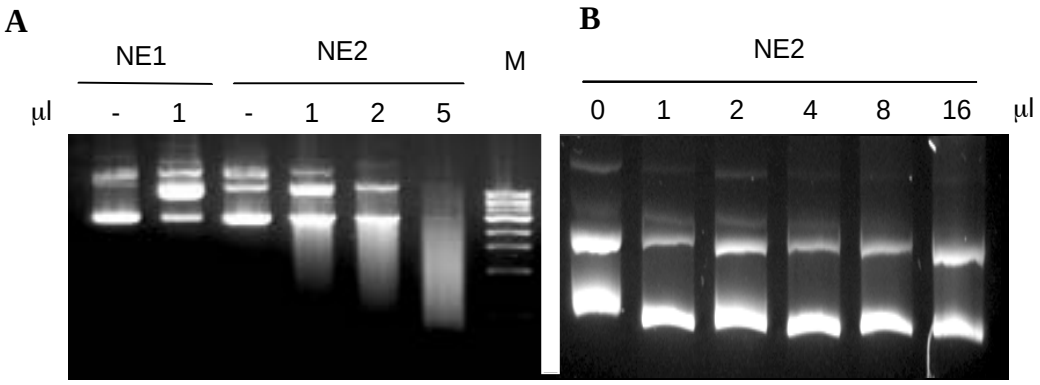


Figure 17. DNase activity in tobacco nuclear extracts and inhibition under the conditions for the *in vitro* DNA demethylation assay. 2.5 µg pBluescript plasmid DNA was incubated at 37 °C for 4 hrs with various amount of tobacco nuclear extract prepared via protoplasting (NE2) in the absence (A) or presence (B) of EDTA. The quality of the nuclear extract prepared by grinding in liquid nitrogen (NE1) was compared by the DNase assay. M is 1kb DNA ladder.

The presence of strong DNase activity in the nuclear extract presents a problem with the *in vitro* DNA demethylation assay, because DNA degradation interferes with detection of demethylated DNA. Since Mg^{2+} alone is sufficient for the activity of most enzymes, while both Mg^{2+} and Ca^{2+} are required for DNases (Latham, G. Ambion, inc. unpublished data), I added the metal chelator EDTA to the DNA demethylation reaction at a concentration significantly lower than that of Mg^{2+} but high enough to chelate

residual Ca^{2+} (see Methods). Under these conditions, DNase activity is completely inhibited (Fig 17 B).

The DNA demethylation assay was then performed in the presence of EDTA. SalI sensitive DNA was observed after incubation with tobacco nuclear extracts, but not after mock treatment (Fig. 18). Moreover, the signal for cleaved DNA intensified when the amount of nuclear extract was increased (Fig. 18). These results clearly demonstrated the presence of an enzymatic activity in tobacco suspension cells that nicks methylated DNA.

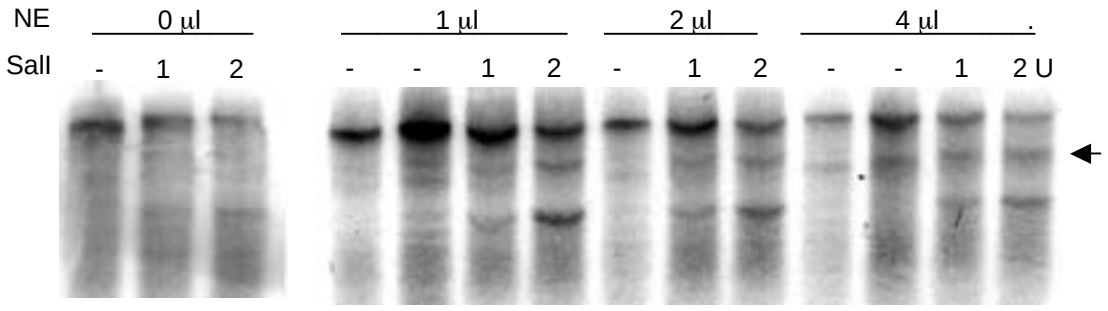


Figure 18. A representative *in vitro* DNA demethylation assay, showing the presence of a DNA demethylase activity in a nuclear extract (NE) prepared from wild type (SR1) tobacco suspension cells. The band indicated by the arrow is DNA fragment resulting from SalI restriction of the substrate DNA.

Discussion

A Sensitive *in vitro* DNA Demethylation Assay

To determine whether there is a DNA demethylase activity in plant cells, I developed an *in vitro* DNA demethylation assay using hemimethylated DNA containing the *Spm* promoter and the DCR as the substrate. This assay will permit the detection of general DNA demethylase activities in plant cells, as well as an activity specific for TnpA. Because TnpA has much higher binding affinity for the *Spm* promoter sequence when hemimethylated than fully methylated (Chapter 5), using hemimethylated DNA enhances the interaction between TnpA and its substrate DNA, thus increasing the sensitivity for detection of TnpA-associated DNA demethylase activity.

Because fully methylated DNA is completely resistant to SalI restriction and hemimethylated DNA is digestable at high enzyme concentrations, the *in vitro* DNA demethylation assay can also be carried out with fully methylated DNA substrate by increasing the amount of SalI in the reaction. This modified DNA demethylation assay will be useful in characterizing the substrate specificity of plant DNA demethylases in further studies.

Distinct DNA Demethylase Activities in Plant Cells

Although there is evidence suggesting the presence of DNA demethylase activities in plant cells, the present study represents the first effort to directly test such a possibility. My results showed that there is indeed a methylated DNA nicking activity in tobacco suspension cells. The nicked DNA cannot be attributed to DNases in the nuclear extract,

because DNase activity was completely inhibited by the metal chelator EDTA. Although it is still unknown whether this DNA nicking activity is specific to methylated DNA, it clearly uses methylated DNA as substrate and is believed to be a DNA demethylase. Since multiple DNA demethylases activities may be present in the crude nuclear extracts, prior purification is necessary in further experiments to characterize the activity.

Simultaneously with this finding it was reported that the ROS1 protein in *Arabidopsis*, a DNA glycosylase, is able to nick methylated DNA specifically (Gong et al., 2002). Unlike the activity detected in this study, however, the ROS1 protein only nicks DNA methylated with MspI methylase, which methylates the external cytosine in a CCGG context, but not DNA methylated with SssI methylase that methylates the cytosine in the CG dinucleotide context. Because the DNA substrate used in the present study was methylated using the SssI methylase, it is very likely that the DNA demethylase activity detected in our assay is distinct from ROS1. Interestingly, two *ROS1* homologs have been identified in *Arabidopsis*. One of them is the *DEMETER*, which encodes an embryo-specific DNA glycosylase but does not seem to have an *in vivo* function in DNA demethylation (Choi et al., 2002). The function of the second homolog has not been characterized, and therefore it would be interesting to see whether it encodes a DNA glycosylase that catalyzes the removal of 5mC from the CG dinucleotide context.

Tissue-Specific and Temporal Expression of DNA Demethylase

As described previously (Chapter 5), TnpA-mediated DNA demethylation was not observed in fully expanded leaves that have ceased to divide. This could be due to the inability of TnpA to get access to methylated *Spm* sequence in the absence of DNA

replication, as DNA replication was shown to be required for TnpA to initiate DNA demethylation (Chapter 5). Another possible cause is the lack of DNA demethylase in the leaf cells. In animals, the 5-methylcytosine DNA glycosylase is exclusively expressed in dividing cells (Jost and Jost, 1995; Schwarz et al., 2000). This observation raises the interesting possibility that DNA demethylase activity in plants may also be expressed in a tissue-specific or temporal manner. This may help to explain the phenotype of the *mom* mutation of *Arabidopsis* that causes release of gene silencing but no change in DNA methylation (Amedeo et al., 2000). The *mom* mutation is a weak allele affecting only a transgene at later stage of development when cells have ceased dividing (Probst et al., 2003).

References

Amedeo, P., Habu, Y., Afsar, K., Scheid, O.M., and Paszkowski, J. (2000). Disruption of the plant gene MOM releases transcriptional silencing of methylated genes. Nature **405,** 203-206.

Bhattacharya, S.K., Ramchandani, S., Cervoni, N., and Szyf, M. (1999). A mammalian protein with specific demethylase activity for mCpG DNA. Nature **397,** 579-583.

Cervoni, N., Bhattacharya, S., and Szyf, M. (1999). DNA demethylase is a processive enzyme. J. Biol. Chem. **274,** 8363-8366.

Chen, Z., Ricigliano, J.W., and Klessig, D.F. (1993). Purification and characterization of a soluble salicylic acid-binding protein from tobacco. Proc. Natl. Acad. Sci. USA **90,** 9533-9537.

Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J.J., Goldberg, R.B., Jacobsen, S.E., and Fischer, R.L. (2002). DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. Cell **110,** 33-42.

Cooke, R., and Penon, P. (1990). In vitro transcription from cauliflower mosaic virus promoters by a cell-free extract from tobacco cells. Plant Mol Biol **14,** 391-405.

Detich, N., Theberge, J., and Szyf, M. (2002). Promoter-specific activation and demethylation by MBD2/demethylase. J. Biol. Chem. **277,** 35791-35794.

Falnes, P.O., Johansen, R.F., and Seeberg, E. (2002). AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli. Nature **419,** 178-182.

Gong, Z., Morales-Ruiz, T., Ariza, R.R., Roldan-Arjona, T., David, L., and Zhu, J.K. (2002). ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase. Cell **111,** 803-814.

Hendrich, B., Hardeland, U., Ng, H.H., Jiricny, J., and Bird, A. (1999). The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. Nature **401,** 301-304.

Jost, J.P., and Jost, Y.C. (1995). Mechanism of active DNA demethylation during embryonic development and cellular differentiation in vertebrates. Gene **157,** 265-266.

Lax, S.R., Lauer, S.J., Browning, K.S., and Ravel, J.M. (1986). Purification and properties of protein synthesis initiation and elongation factors from wheat germ. Methods Enzymol. **118,** 109-128.

Ng, H.H., Zhang, Y., Hendrich, B., Johnson, C.A., Turner, B.M., Erdjument-Bromage, H., Tempst, P., Reinberg, D., and Bird, A. (1999). MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. Nat. Genet. **23,** 58-61.

Oakeley, E.J., and Jost, J.P. (1996). Non-symmetrical cytosine methylation in tobacco pollen DNA. Plant. Mol. Biol. **31,** 927-930.

Petronzelli, F., Riccio, A., Markham, G.D., Seeholzer, S.H., Genuardi, M., Karbowski, M., Yeung, A.T., Matsumoto, Y., and Bellacosa, A. (2000). Investigation of the substrate spectrum of the human mismatch-specific DNA N-

glycosylase MED1 (MBD4): fundamental role of the catalytic domain. J. Cell Physiol. **185,** 473-480.

Probst, A.V., Fransz, P.F., Paszkowski, J., and Scheid, O.M. (2003). Two means of transcriptional reactivation within heterochromatin. Plant J. **33,** 743-749.

Ramchandani, S., Bhattacharya, S.K., Cervoni, N., and Szyf, M. (1999). DNA methylation is a reversible biological signal. Proc. Natl. Acad. Sci. USA **96,** 6107-6112.

Roberts, M.W., and Okita, T.W. (1991). Accurate in vitro transcription of plant promoters with nuclear extracts prepared from cultured plant cells. Plant. Mol. Biol. **16,** 771-786.

Schwarz, S., Bourgeois, C., Soussaline, F., Homsy, C., Podesta, A., and Jost, J.P. (2000). A CpG-rich RNA and an RNA helicase tightly associated with the DNA demethylation complex are present mainly in dividing chick embryo cells. Eur. J. Cell Biol. **79,** 488-494.

Trewick, S.C., Henshaw, T.F., Hausinger, R.P., Lindahl, T., and Sedgwick, B. (2002). Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. Nature **419,** 174-178.

Vairapandi, M., and Duker, N.J. (1996). Partial purification and characterization of human 5-methylcytosine-DNA glycosylase. Oncogene **13,** 933-938.

Wang, R.Y., Kuo, K.C., Gehrke, C.W., Huang, L.H., and Ehrlich, M. (1982). Heat- and alkali-induced deamination of 5-methylcytosine and cytosine residues in DNA. Biochim. Biophys. Acta. **697,** 371-377.

Wu, P., Qiu, C., Sohail, A., Zhang, X., Bhagwat, A.S., and Cheng, X. (2003). Mismatch repair in methylated DNA. Structure and activity of the mismatch-specific thymine glycosylase domain of methyl-cpg-binding protein mbd4. J. Biol. Chem. **278,** 5285-5291.

Zhang, Y., Ng, H.H., Erdjument-Bromage, H., Tempst, P., Bird, A., and Reinberg, D. (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. Genes Dev. **13,** 1924-1935.

Zhu, B., Zheng, Y., Hess, D., Angliker, H., Schwarz, S., Siegmann, M., Thiry, S., and Jost, J.P. (2000). 5-methylcytosine-DNA glycosylase activity is present in a cloned G/T mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex. Proc. Natl. Acad. Sci. USA **97,** 5135-5139.

Chapter 7

DNA DEMETHYLASE ACTIVITY IS RECRUITED TO THE *SPM* SEQUENCE AS A RESULT OF TRANSCRIPTIONAL ACTIVATION

Introduction

Although several proteins in animal cells have been shown to be capable of converting 5mC to cytosine, whether and how they participate in sequence-specific DNA demethylation still remain largely unknown. In some cases, DNA demethylases seem to be targeted to DNA sequences by trans-acting factors through physical interactions. For example, 5-methylcytosine DNA glycosylase, which could catalyze DNA demethylation through a DNA repair pathway, has been shown recently to interact with the retinoid and the estradiol receptors, both are known to promote sequence-specific DNA demethylation (Zhu et al., 2001; Jost et al., 2002).

The previous findings that there is a DNA demethylase activity in plant cells and that TnpA has a leucine-zipper domain in its C-terminus strongly suggest that a DNA demethylase might be recruited to the *Spm* promoter sequence through direct or indirect protein-protein interactions. To test this possibility, I purified TnpA and possible interacting proteins from plant cells expressing the TnpA protein using a co-immunoprecipitation approach. DNA demethylase activity was not detected in TnpA-containing immunoprecipitate, however. The present study clearly ruled out the

possibility that TnpA is a sequence-specific DNA demethylase. Results from this experiment and selective precipitation by ammonium sulfate also failed to show direct interaction between TnpA and DNA demethylase activities. Since demethylation of the *Spm* sequence was observed after induction of transcription from the *Spm* sequence under the control of an inducible promoter, it is suggested that a DNA demethylase gains access to the *Spm* sequence passively as a result of transcriptional activation by the TnpA protein.

Materials and Methods

Plant Materials

Unless specified, all plant materials and treatments used in this study have been described in previous chapters. Transgenic tobacco lines containing the pTA-U+D-Luc-Spm3' construct for inducible transcription activation was generated using the Agribacterium-mediated leaf-disc transformation method as used for generation of other transgenic lines.

Constructs

To clone pTA-*Spm*-LUC-Spm3', the construct for inducible transcription through the *Spm* sequence, the *Spm*-LUC reporter construct was first subcloned from pMS73 as an approximately 4.7 kb fragment (Schläppi, 1994) into pBluescript II KS (+) and then

cloned as an Asp718 (filled-in with Klenow) to SpeI fragment into pTA7001 or pTA7002 (Aoyama and Chua, 1997) that was cut with XhoI and SpeI (XhoI filled-in with Klenow).

Immunoprecipitation

The FLAG-tagged TnpA protein and its truncated derivatives, FLAG-TnpA $_{540}$ and FLAG-TnpA$_{420}$, were pulled down using anti-FLAG® M2-agarose affinity gel (A-2220, Sigma-Aldrich, St. Louis, MO). Tobacco suspension cultures that express each protein were first filtered through a layer of miracloth (Calbiochem-Novabiochem Corporation, La Jolla, CA), weighed and grounded on ice in 3X volume of ice-cold extraction buffer, which is composed of 50 mM Tris-HCl, pH 7.6 at 4 °C, 450 mM NaCl, 10 mM EDTA, 1% Triton X-100, 50 mM β-mercaptoethanol, 3 mM PMSF, 1X plant protease inhibitor cocktail Sigma-Aldrich), 10% glycerol and 1% BSA. The extract was centrifuged at 15000 rpm, 4°C for 15 min, and 1 ml supernatant was transferred to a tube with 40 μl anti-FLAG agarose beads, which had been pretreated following the vendor's recommendations. After incubation at 4 °C for 6 hrs on a shaker (Barnstead/thermolyne, Dubuque, IOWA), the beads were washed 1-3 times with a buffer that contains 50 mM Tris-HCl, pH 7.6 at 4°C, 150 mM NaCl, 10 mM EDTA, 50 mM β-mercaptoethanol, 1 mM PMSF, 1X plant protease inhibitor cocktail and 10% glycerol and 0.1%-1% Triton X-100 depending on the required stringency. The gel was finally washed in washing buffer that contains no Triton X-100.

For protein detection only, the gel was boiled in 40 μl 2X SDS protein sample buffer and the supernatant was analyzed by a western blot assay. To release the FLAG-tagged protein in its native form, the gel was incubated with 40 μl of 150 ng/μl FLAG

peptide solution at 4 °C for 45 min with shaking. The supernatant was collected after a brief centrifugation at 10,000 rpm, 4°C. The elution was repeated once and the supernatants were combined.

Ammonium Sulfate Precipitation

Selective ammonium sulfate precipitation of proteins from tobacco extract was conducted following standard protocols (Lovrien and Matulis, 1997). Briefly, tobacco suspension cultures were first filtered through a layer of miracloth (Calbiochem-Novabiochem Corporation, La Jolla, CA). 50 g cells were weighed and ground in 50 ml ice-cold extraction buffer (Tris-HCl, pH 7.6, 150 mM NaCl, 10% glycerol, 10 mM EDTA, 50 mM 2-mercaptoethanol, 5 mM PMSF and 5mM NaF). The supernatant was then collected after the homogenate was centrifuged at 40,000 rpm, 4 °C for 1 hr using a Ti70 rotor (Beckman Ultracentrifuge). $(NH_4)_2SO_4$ powder was added into the extract with stirring to achieve the desired saturation. The slurry was stirred for additional 20 min, then centrifuged at 17,500 rpm at 4 °C for 20 min. The supernatant was transferred to a beaker and the process was repeated to precipitate proteins at higher $(NH_4)_2SO_4$ concentrations.

The pellets from each fraction were dissolved in 2 ml protein extraction buffer and centrifuged at 15000 rpm, 4 °C for 15 min to remove undissolved material. Each fraction was dialyzed against 100X vol of dialysis buffer (20mM HEPES-KOH, pH7.9, 20% glycerol, 100 mM KOAc, 0.2 mM EGTA, 0.1 mM EDTA, 2 mM DTT and 0.5 mM PMSF) using cellulose ester tubing (Spectra/por® CE, MWCO of 3500, CAT# 131162). The buffer was changed every 4-6 hrs and this was repeated twice. Finally, the solution

was centrifuged at 15,000 rpm, 4 °C for 20 min and the supernatant was aliquoted and stored at –80C.

DNA Demethylation Assay and Other Assays

DNA demethylase assay was performed as described in Chapter 6. In some cases, hemimethylated DNA containing the DCR sequence alone was used as the substrate, which was amplified by PCR as a 493 bp fragment from pDC131 using primer pairs KS and SK (Raina, 1993). Silver staining of protein gels was performed using the Silver Stain Plus kit (BioRad, Hercules, CA).

All other assays such as Western blots, gel mobility shift assays, RT-PCR and DNA methylation assays were performed as described previously.

Results

Optimization of Conditions for TnpA Immunoprecipitation

To investigate the possibility that TnpA interacts physically with a DNA demethylase activity, it is necessary to purify TnpA along with possible interacting proteins from transgenic tobacco cells that express the TnpA protein. This was attempted by co-immunoprecipitation (IP) using anti-FLAG antibody affinity gel (Methods), as TnpA and its truncated derivatives $TnpA_{540}$ and $TnpA_{420}$ have been expressed as FLAG-tagged fusion proteins in transgenic tobacco cells. Another advantage of this approach is that, along with the use of proteomics tools, it would permit quick identification of the proteins that participate in TnpA-mediated DNA demethylation.

Although the fusion protein between the FLAG tag and bacterial alkaline phosphotase (FLAG-BAP) was efficiently pulled down in the reconstruction experiment when purified FLAG-BAP protein was added to a tobacco extract, the initial effort to pull down the FLAG-TnpA protein from tobacco cells was not successful (not shown). The

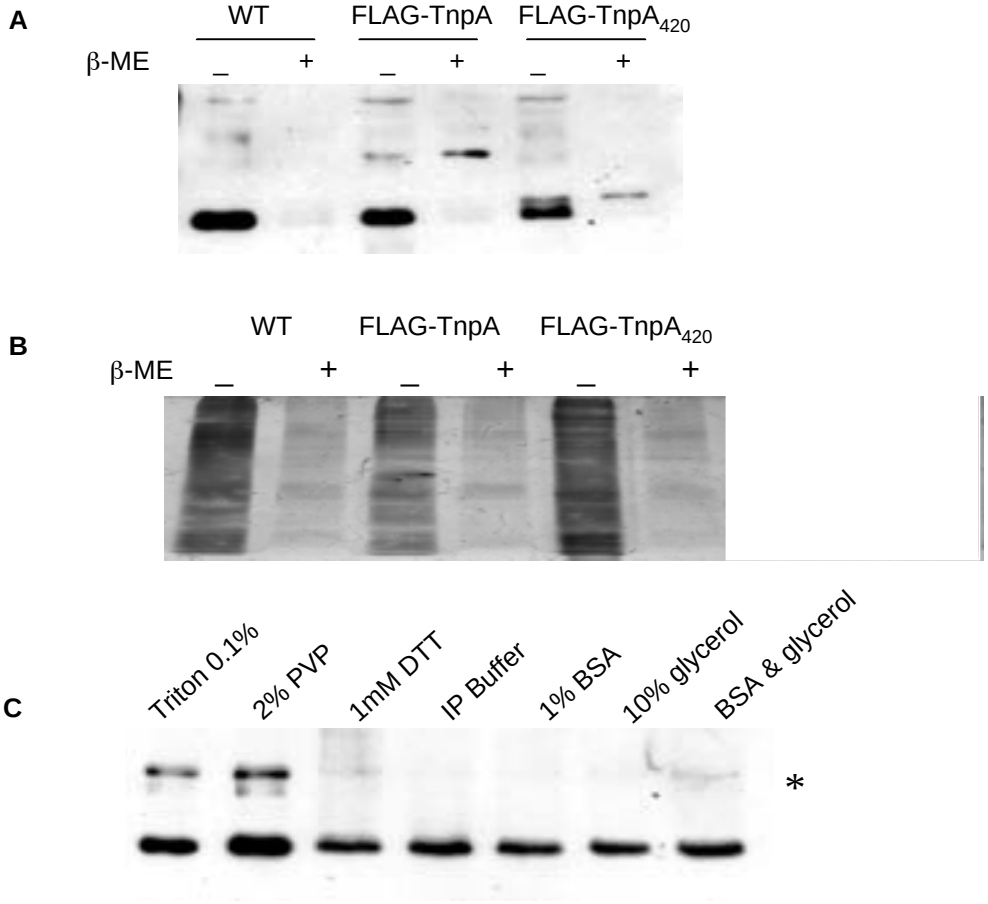


Figure 19. Immunoprecipitation (IP) of FLAG-tagged TnpA protein and derivatives from tobacco suspension cultured cells. (A) Western blot assay of FLAG-TnpA or FLAG-$TnpA_{420}$ in immunoprecipitates prepared in the presence (+) or absence (-) of 50 mM β-mercaptoethanol (β-ME). As a negative control, IP was conducted as well with a wildtype (WT) extract. (B) Silver staining of the immunoprecipitates in (A). (C) IP of FLAG-TnpA in the presence of the chemicals as indicated. The band marked by "*" is a non-TnpA signal arising from cross reactivity.

failure is partly due to some intrinsic property of the protein, because IP efficiency was very low even with purified FLAG-TnpA in extraction buffer (not shown). Since no TnpA protein was recovered at all when it was added to a plant extract, some components in the tobacco extract may exacerbate this problem.

To optimize the conditions for TnpA immunoprecipitation, I therefore compared the efficiency of immunopreciptation under various conditions, including different salt, glycerol and detergent concentrations as well as in the presence of bovine serum albumin (BSA) and polyvinylpolypyrrolidine (PVP). None of the tested conditions significantly and consistently improved the IP efficiency (not shown). However, an inverse relationship was observed between IP efficiency and the intensity of brown color, suggesting that oxidation of polyphenolic compounds might interfere with immunoprecipitation of the TnpA protein (Lax et al., 1986; Chen et al., 1993).

In order to prevent oxidation of phenolic compounds, I therefore increased the concentration of the reducing reagent β-mercaptoethanol in the protein extraction buffer. Remarkably, I was able to consistently pull down the FLAG-TnpA protein and its truncated derivatives when 50 mM β-mercaptoethanol was used (Fig. 19A). Use of other chemicals that supposedly increase specificity or stabilize protein complexes did not further improve the efficiency of immunoprecipitation (Fig 19 C). Interestingly, it has been reported previously that the presence of a reducing reagent, such β-mercaptoethanol or DTT, is essential for maintaining the DNA binding ability of the TnpA protein (Trentman, 1991). A likely explanation for this observation is that under the oxidative condition proteins with protein-protein interaction domains tend to form functionally inactive aggregate.

TnpA is not itself, nor does it Bind to, a DNA Demethylase

To be compatible with the *in vitro* DNA demethylation assay, the FLAG-tagged TnpA protein was released from the affinity gel in its native form by incubation with FLAG peptide solution. As a negative control, native FLAG-TnpA$_{540}$ and FLAG-TnpA$_{420}$

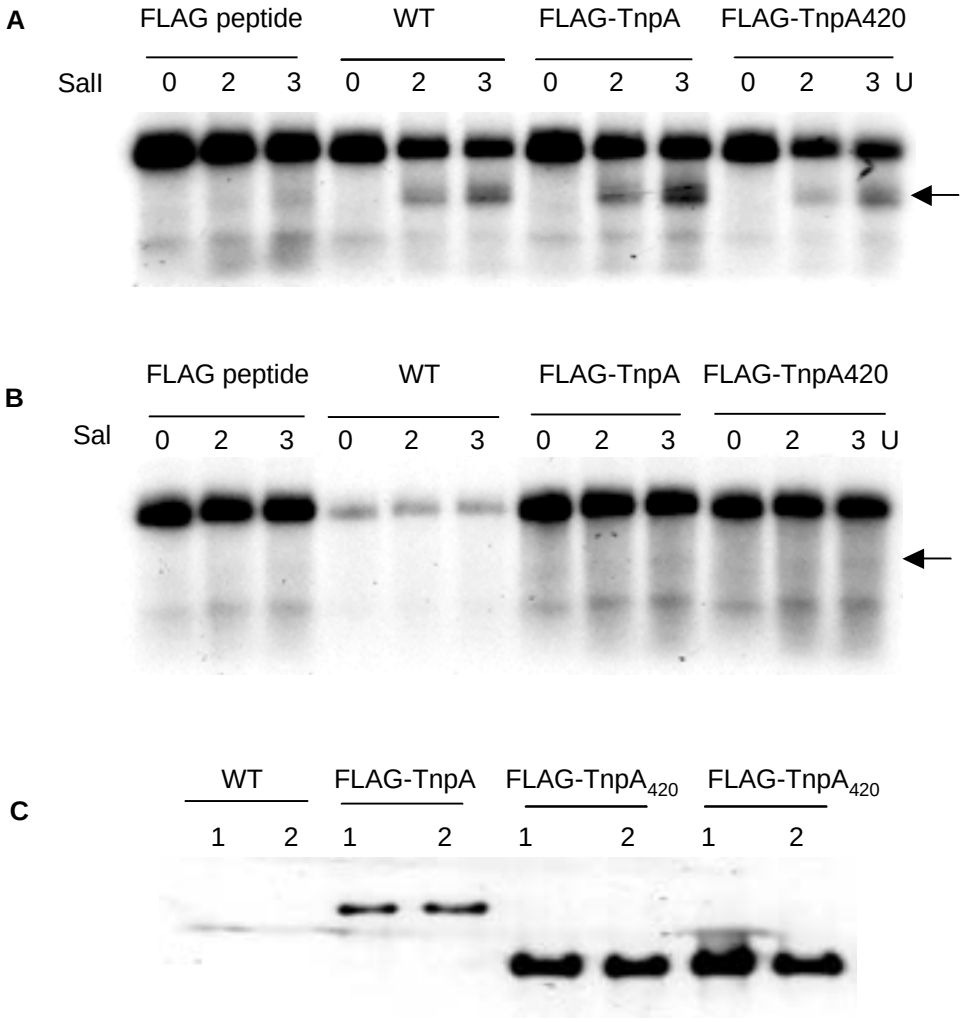


Figure 20. TnpA does not associate specifically with a DNA demethylase activity. DNA demethylase activity in immunoprecipitates (2 μl each) prepared by (A) mild washing or (B) extensive washing. The arrow indicates the position of SalI cleaved DNA as a result of DNA demethylation. (C) Western blot analysis of the samples used in (B).

proteins were also prepared, in addition to a mock immunoprecipitation using wildtype tobacco cell extract. This mild elution condition should also help to maintain the integrity of possible TnpA-containing protein complexes. For the same reason, the IP experiment was initially carried out using mild washing conditions (washing once with 0.1% Triton X100).

As shown in Fig. 20A, DNA demethylase activity was clearly detected in all immunoprecipitated samples. The observed SalI-sensitive DNA was not due to DNase activity, which was inhibited under the assay condition (not shown). This observation confirmed the previous one that there is a DNA demethylase activity in plant cells. However, after increasing the stringency of washing in subsequent IP experiments (multiple washes with buffers containing 1% Triton X-100), DNA demethylase activity was dramatically reduced in all immunoprecipitate samples (Fig. 20B), although the FLAG-G tagged proteins were still present (Fig 20C). DNA demethylase activity detected in the immunoprecipitate without extensive washing is thus believed to have arisen from nonspecific interaction between proteins and the affinity gel. As revealed by silver staining of the gel, many proteins were still present in both the mock immunoprecipitate and the FLAG-TnpA immunoprecipirate (Fig 19B). These results suggested that TnpA does not associate with a demethylase, nor does it contain such an activity.

The failure to detect TnpA-associated DNA demethylase activity in the foregoing experiments could be due to the inability of the FLAG-TnpA protein to bind the DNA substrate. As shown in Fig. 21, the TnpA protein purified by immunoprecipitation binds well to a DNA fragment containing its cognate binding sites. However, because of the

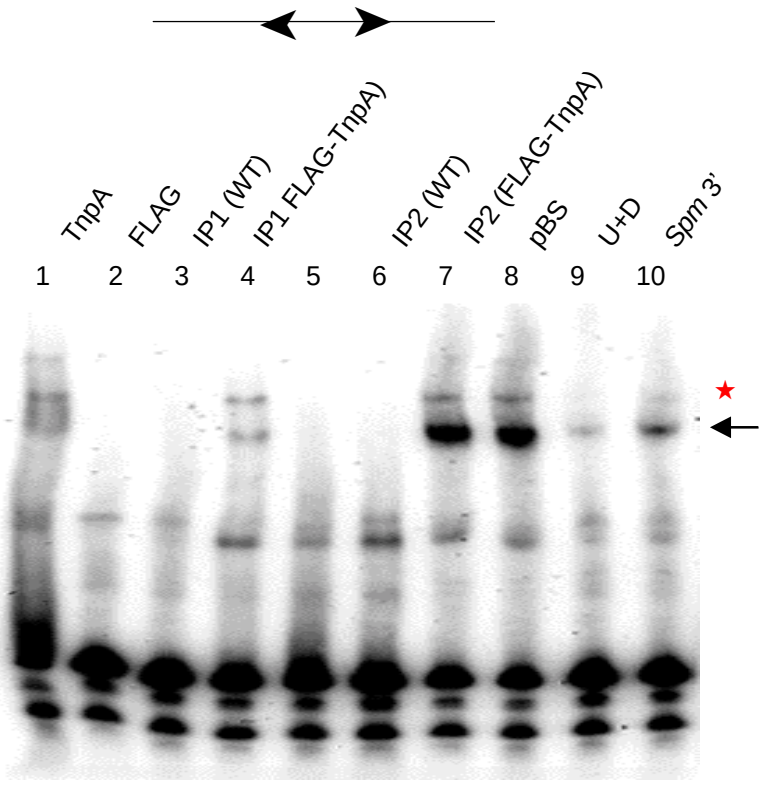


Figure 21. TnpA protein expressed in tobacco cells binds well to the *Spm* promoter sequence. Native FLAG-TnpA protein prepared by immunoprecipitation from tobacco cells was assayed by mobility shift assay for binding to a synthesized DNA fragment containing two TnpA binding sites (shown above). The arrow and asterisk mark the positions of TnpA-DNA complexes. TnpA purified from bacteria (TnpA) was used as a positive control, the FLAG peptide (FLAG) and mock immunoprecipitates prepared from wildtype tobacco extracts (WT) served as negative controls. IP1 and IP2 represent two preparations of immunoprecipitates. In lane 8-10 was added 20 ng competitor DNA as indicated. pBS, pBluescript plasmid; U+D, a 565 bp HindIII-BamHI fragment cut from pDC105 containing the *Spm* promoter and DCR; *Spm*3', a 1.2 kb HindIII-PstI DNA fragment from pMS66 containing the 3' end sequence of the *Spm* transposon (Schläppi et al., 1994).

strong DNA binding, TnpA-associated DNA demethylase activity could be anchored in the *Spm* promoter region and thus unable to reach the downstream DCR to effect DNA

demethylation. To check this possibility, I performed a DNA demethylation assay using a hemimethylated DNA fragment containing the DCR sequence only as substrate (Methods). Again, TnpA-associated DNA demethylase activity was not observed (Fig 22).

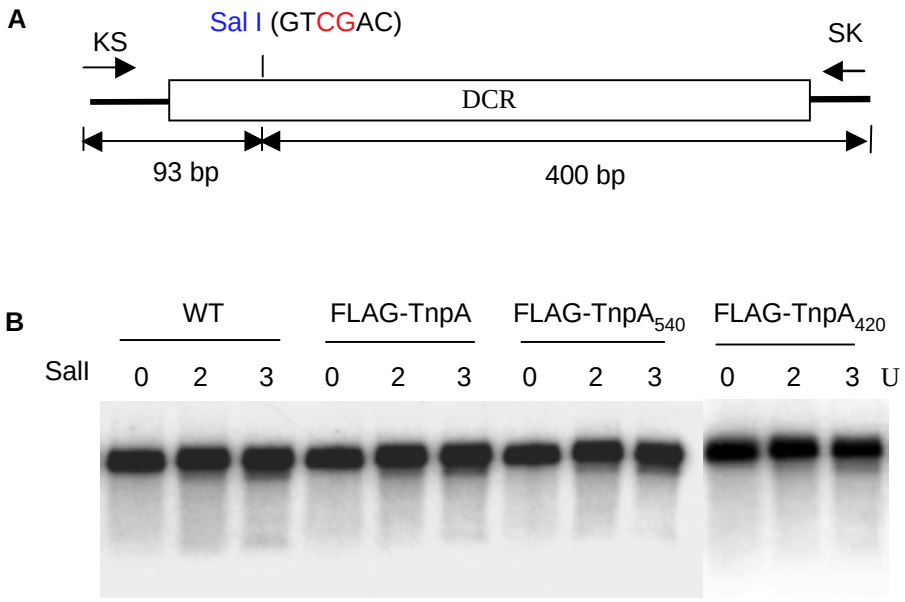


Figure 22. DNA demethylation assay using hemimethylated DNA containing DCR only. (A) A diagram showing the primers KS and SK used for PCR amplification of the DNA substrate and the expected DNA fragments after SalI restriction. (B) DNA demethylase activity in immunoprecipitates as used for the experiment shown in Figure 20B.

The foregoing experiments do not rule out the possibility that TnpA binds a 5-methylcytosine DNA glycosylase (5MCDG). If a 5MCDG activity participates in TnpA-mediated DNA demethylation, the resulting abasic sites at 5mC may become resistant to SalI restriction, as is the case for EcoRII (Petrauskene et al., 1995), and thus evade detection under the current assay conditions. To investigate this possibility, endonuclease

IV, an apurinic/apyrimidinic (AP) endonuclease that nicks DNA specifically at abasic sites (Kaboev et al., 1985; Masters et al., 1991), was used to initiate DNA nicking before SalI restriction. No change in the amount of SalI sensitive DNA was observed, however, indicating the absence of DNA glycosylase in the TnpA-containing immunoprecipitate (Fig 23).

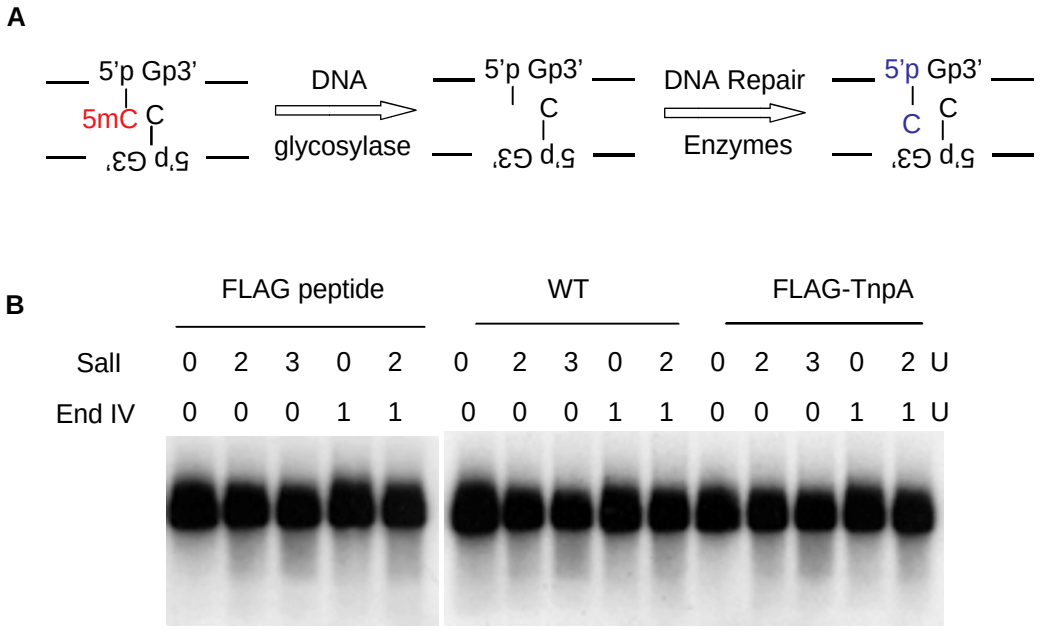


Figure 23. Detection of abasic sites in DNA using endonuclease IV.
(A) Diagram showing DNA demethylation through a DNA repair pathway involving DNA glycosylase. (B) Using bacterial endonuclease IV (End IV) and SalI for detection of DNA demethylase activity in the immunoprecipitates as used for the experiment shown in Fig 20B.

Another possibility for the lack of DNA demethylase activity in the TnpA immunoprecipitate is that proteins that interact with TnpA might be washed off under the more stringent IP condition. To address this question, I precipitated proteins from TnpA

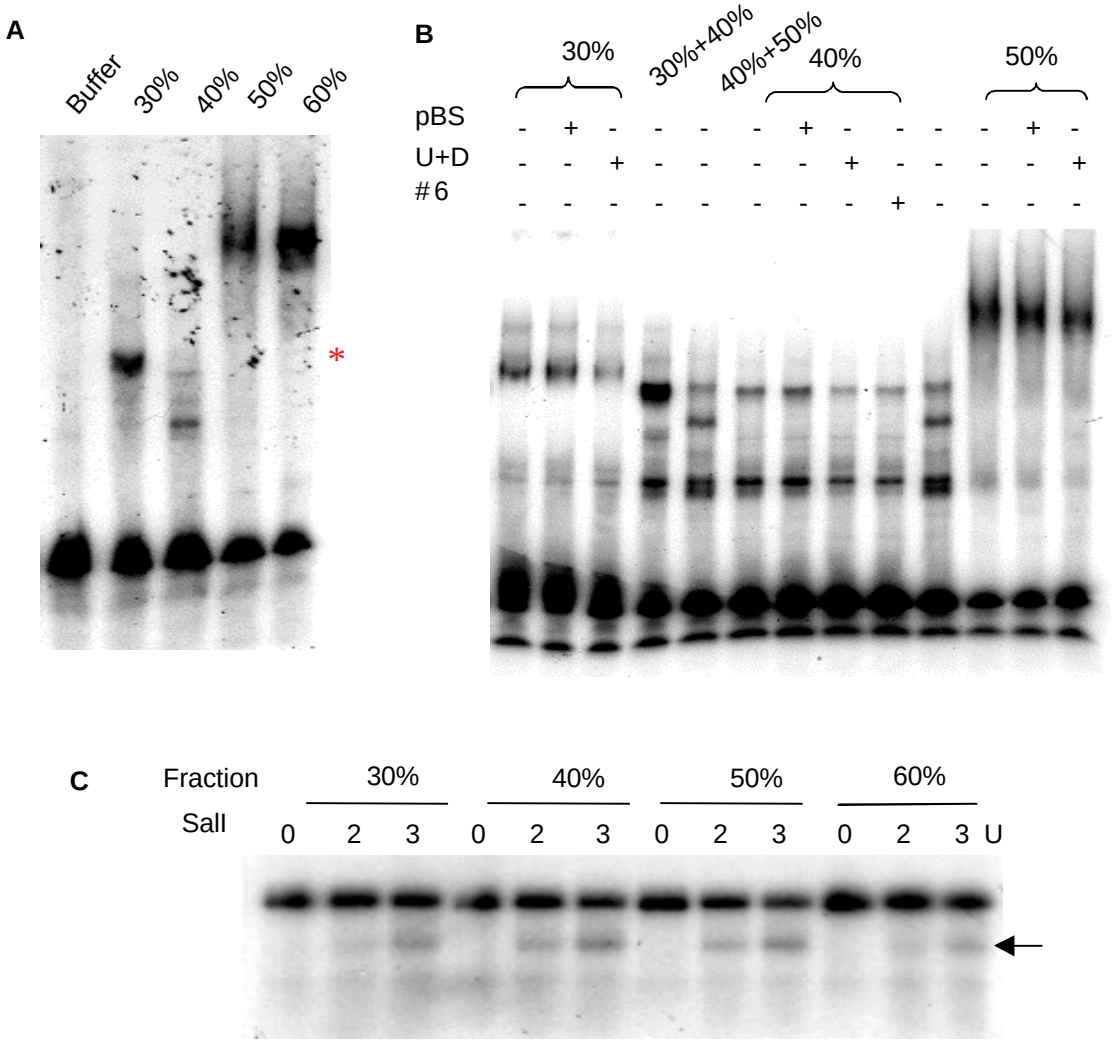


Figure 24. TnpA does not co-fractionate with DNA demethylase activity.

Mobility shift assay of precipitates at the concentrations of $AmSO_4$ indicated (A) in the absence or (B) presence of competitor DNA. The asterisk marks the position of DNA-TnpA complexes. pBS and U+D, same DNA fragment as described in Figure 21. #6, a 134bp PCR-amplified DNA fragment containing 2 TnpA binding site, as described in Chapter 4; The buffer used to dissolve the sample was used a control. 30%+40% and 40%+50%, mixtures of equal volumes of the 30% and 40% or 40% and 50% fractions respectively. (C) DNA demethylase activity in different fractions of $AmSO_4$ precipitation, as indicated by the band marked by the arrow.

expressing tobacco cells at increasing concentrations of ammonium sulfate and

determined their DNA demethylase activity. Due to the 'salting out' effect, proteins tend to remain in their complex at high $(NH_4)_2SO_4$ concentrations. If TnpA interacts with a DNA demethylase activity, TnpA and DNA demethylase activity would be expected to co-fractionate. While almost all FLAG-tagged TnpA protein was precipitated at 30% saturation of $(NH_4)_2SO_4$, as revealed by mobility shift assay (Fig. 24A and B), only a small portion of DNA demethylase activity was present in this fraction (Fig. 24 C). By contrast, most DNA demethylase activity seems to be in the fraction corresponding to 40% $(NH_4)_2SO_4$ saturation, which contains little TnpA protein (Fig 24).

Taken together, all the experiments described above indicate that TnpA is unlikely to be associated physically with a DNA demethylase.

Spm Demethylation after Transcriptional Activation

The finding that TnpA's transcriptional activity is required for it to promote DNA demethylation, as well as the failure to detect DNA demethylase activity in TnpA-containing immunoprecipitates, strongly suggests an indirect mechanism for the recruitment of DNA demethylase activity to the *Spm* sequence. To further investigate the relationship between transcription activation and DNA demethylation, I generated transgenic tobacco plants in which the *Spm* promoter and DCR along with a luciferase reporter gene construct (Schläppi et al., 1994) are placed downstream of the glucocorticoid-inducible (GR) promoter (Aoyama and Chua, 1997), the same vector as used for controlling TnpA expression described in previous experiments (Fig 25A). This will make it possible to assemble the transcription initiation complex on the GR promoter upstream from the *Spm* promoter, then transcribe through the *Spm* sequence.

In most transgenic lines, the *Spm* sequence remained either methylated or unmethylated after dexamethasone treatment to activate transcription (Not shown).

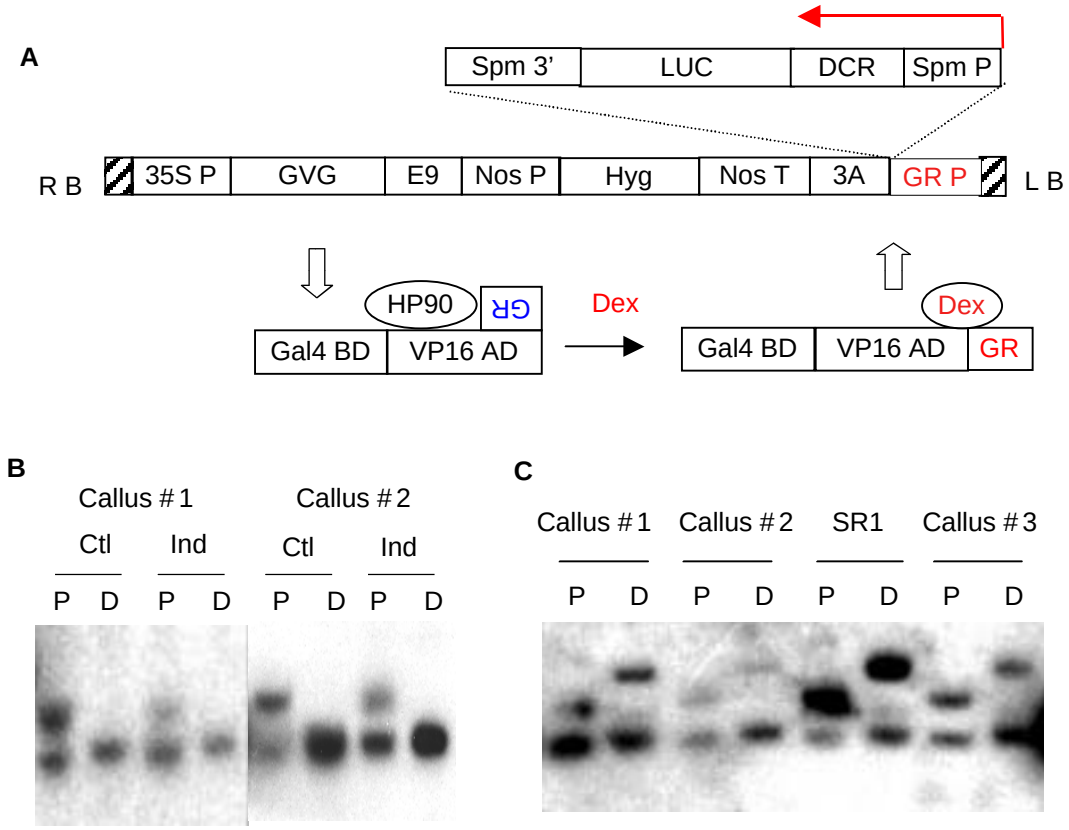


Figure 25. *Spm* demethyaltion after induction of transcription. (A) T-DNA region of the binary vector used for controlling transcription through the *Spm* sequence. GVG, a hybrid transcription factor consisting of the Gal4 binding domain (Gal4 BD), the viral VP16 activation domain (VP16 AD) and the glucocorticoid receptor domain (GR). *Spm* P, the *Spm* promoter sequence*;* DCR, the promoter downstream sequence. The *Spm* P-DCR-Luc-*Spm* 3' sequence is transcribed as a single transcription unit. See the legend in Figure 1 for detail. (B) DNA methylation in the *Spm* promoter (P) and DCR (D) sequence before (Ctl) and after (Ind) induction treatment in two independent transgenic tabacco calli with the construct shown in (A). (C) *Spm* methylation in transgenic tobacco calli, which have lost methylation due to TnpA expression, 2 weeks after transfer to non-induction medium. SR1 is the line that does not have the TnpA-expression construct. In (B) and (C), the upper and lower bands represent methylated and unmethylated DNA, respectively.

However, two of 11 independent transgenic lines examined showed significant loss of DNA methylation after 48 hrs of induction (Fig. 25B). As neither dexamethasone nor the transcription factor GVG for inducible transcription interferes with DNA methylation in the *Spm* sequence (Chapter 2), the present observation strongly supports a causal connection between transcriptional activation and DNA demethylation.

Notably, none of the transgenic lines showed an increase in DNA methylation, although there was clear evidence for the presence of *de novo* DNA methylase activity in the cells under study. In transgenic lines that contain the construct for inducible TnpA expression and have completely lost DNA methylation in the *Spm* sequence after prolonged induction treatment, the *Spm* sequence regained methylation after depletion of the inducing chemical (Fig. 25C),

Discussion

Recruitment by TnpA of DNA Demethylase Activity to the *Spm* Sequence

TnpA could promote active demethylation of the *Spm* sequence through one or a combination of the following possible mechanisms. First, TnpA could itself be a sequence-specific DNA demethylase. Although it has been shown that purified recombinant TnpA protein overexpressed in bacteria does not have DNA demethylase activity, the possibility that TnpA contains DNA demethylase activity has not been formally ruled out, as some eukaryotic modifications might be essential for TnpA's ability to promote DNA demethylation. A more likely possibility is that TnpA recruits a

distinct protein that has DNA demethylase activity to the *Spm* sequence. Recruitment of this DNA demethylase could be achieved through direct interaction with TnpA or indirectly through a protein complex. Since the transcription activation domain of TnpA has been shown to be indispensable for TnpA's ability to promote DNA demethylation, a third possible scenario is that DNA demethylases gains access to the *Spm* sequence as a consequence of transcription activation, chromatin remodeling in particularly.

To distinguish these possibilities, I used co-immunoprecipitation to purify TnpA and possible interacting proteins from tobacco cells that express TnpA as a FLAG tagged protein and determined DNA demethylase activity in the immunoprecipitate using an *in vitro* DNA demethylation assay. If TnpA itself is a DNA demethylase or it recruits such an activity through protein-protein interactions, a DNA demethylase activity should be detectable in the immunoprecipitate that contains TnpA, but not in ones containing its truncated derivatives or in a mock immunoprecipitate from wildtype tobacco extract.

Although a DNA demethylase was detected in all immunoprecipitates under mild washing conditions, no such activity was found to be specifically associated with the TnpA protein regardless of the method used for detecting DNA demethylase activity. This result clearly rules out the possibility that TnpA itself is a sequence-specific DNA demethylase. The co-immunoprecipitation study does not support the hypothesis that TnpA interacts with a DNA demethylase, either. It is not very likely that a TnpA-associating DNA demethylase was washed off the affinity gel during the co-immunoprecipitation experiment because TnpA and DNA demethylase did not co-fractionate when proteins in whole cell extract was selectively precipitated by ammonium sulfate. Results from these experiments and the inducible transcription approach

described below strongly suggest that DNA demethylase activities are recruited to the *Spm* sequence passively as a consequence of transcriptional activation (see below).

DNA Demethylation in Response to Transcription Activation

A causal connection between transcription and DNA demethylation has been identified in the present studies on TnpA (Chapter 2 and 3) and has gained support from the observation that no direct interaction between TnpA and DNA demethylase activity was revealed in the foregoing experiments. To further examine the role of transcription in DNA demethylation, I used a glucocorticoid-inducible promoter to control transcriptional activation from the *Spm* sequence. Demethylation of the *Spm* sequence was observed soon after transcription induction (48 hrs), again suggesting an active process. Since transcription activation was achieved by the binding of GVG, a chimeric transcription factor consisting of the gal4 binding domain, the viral VP16 activation domain and the glucocorticoid receptor domain, DNA demethylation cannot be attributed to the TnpA protein. These results further suggest that TnpA promotes DNA demethylation through a mechanism that is common with other transcriptional factors that are also able to promote sequence-specific DNA demethylation.

One possible scenario is that DNA demethylase activity is recruited to actively transcribing DNA sequences via the transcription machinery. There is evidence that the animal G/T mismatch DNA glycosylase, which has been shown to be capable of converting 5mC to normal cytosine through a DNA repair pathway (Zhu et al., 2000), interacts with the transcriptional coactivator p300/CBP (Tini et al., 2002). As most known DNA demethylation-promoting trans-acting factors are transcriptional activators

(Matsuo et al., 1998; Grange et al., 2001; Thomassin et al., 2001; Zhu et al., 2001; Jost et al., 2002), it is likely that they recruit a DNA demethylase activity via the p300 protein or similar coactivators. Interestingly, p300/CBP homologs and interacting proteins have been identified in plants, which share a highly conserved structure to their animal counterparts (Bordoli et al., 2001; Yuan and Giordano, 2002; Lafarge and Montane, 2003), implying that a similar DNA demethylation mechanism might have evolved in plants.

DNA demethylation could result from transcription *per se*. It has long been known that damaged nucleotides in the actively transcribed strand are preferentially repaired, a process termed transcription-coupled repair (TCD) (Li and Bockrath, 1995), which has been found in a wide range of living organisms including plants (Mellon et al., 1996; Leadon and Avrutskaya, 1998; Sturm and Lienhard, 1998). A direct approach to test this possibility is to activate transcription initiation while inhibiting transcription *per se* using transcription inhibitors such as α-amanitin or cordycepin (Tumbar et al., 1999; Nye et al., 2002). The present inducible transcription system would enable such an experiment, because the transcriptional activator GVG, which activates transcription from the inducible promoter, is constitutively expressed and its activation only requires the binding of glucocorticoid ligands. Consequently, assembly of the transcription initiation complex and chromatin decondensation will not be affected by α-amanitin or cordycepin. Such an experiment is impossible with the inducible TnpA system described previously (Chapter 2), as transcription activation from the *Spm* sequence requires the TnpA protein, which requires *de novo* transcription.

There is accumulating evidence indicating that chromatin remodeling rather than transcription *per se* plays a major role in facilitating the access of DNA demethylase to methylated DNA. Although DNA demethylation mediated by the transcription factor NF-κβ requires the presence of the general transcription machinery, it occurs even when transcription is inhibited (Matsuo et al., 1998). It has also been suggested that hyperacetylated histone may serve as a mark for targeting DNA demethylase (Cervoni and Szyf, 2001). Probably the most telling evidence regarding the role of chromatin structure in DNA demethylation is derived from studies of the *mom1* and *ddm1* mutations in *Arabidopsis*. Both DDM1 and DDM1 proteins are SWI2/SNF2-like chromatin remodeling protein (Amedeo et al., 2000; Brzeski and Jerzmanowski, 2003). Whereas mutation in the *DDM1* gene causes genome-wide reduction in DNA methylation (Jeddeloh et al., 1999) and release of gene silencing (Jeddeloh et al., 1998), the *mom1* mutant only shows derepression of gene silencing without a noticeable change in DNA methylation (Amedeo et al., 2000).

Since transcription is generally associated with chromatin remodeling, these observations seem to argue against a role of chromatin remodeling in DNA demethylation. Remarkably, however, a recent study has revealed that the transgene whose transcription is affected by the *mom* mutation is still in a heterochromatin context (Probst et al., 2003). Since the *ddm1* mutant is also associated with heterochromatin decondensation in addition to histone hyperacetylation and loss of histone methylation (Gendrel et al., 2002; Probst et al., 2003), heterochromatin structure could be the major factor affecting the accessibility of DNA demethylase to methylated DNA.

The viral VP16 activation domain is among the best characterized to activate transcription through increasing the level of histone acetylation and remodeling the chromatin structure (Tumbar et al., 1999; Muller et al., 2001; Nye et al., 2002). The present finding that *Spm* demethylation followed transcriptional activation by the GVG transcription factor thus lends further support to the chromatin remodeling-based mechanism. Moreover, this inducible system would permit a dynamic study of the interrelationship between DNA demethylation and transcription activation, change in histone methylation and acetylation as well as chromatin remodeling.

Transcriptional Gene Silencing and DNA Methylation

The experiment using an inducible promoter to control transcription through the *Spm* sequence has also yielded information regarding the mechanism that controls *de novo* methylation and gene silencing of the *Spm* sequence. DNA methylation can result from gene silencing at the level of transcription (TGS) or posttranscription (PTGS) (Sijen et al., 2001; Vaucheret and Fagard, 2001). In contrast to TGS, where transcription is completely arrested, PTGS is characterized by active transcription from the promoter and an insignificant steady-state level of full-length transcripts as a result of RNA degradation. Since TGS and PTGS often occur to repetitive sequences, and dsRNA has been found to be a potent silencer of gene expression (Matzke et al., 2000), it has been hypothesized that silencing and methylation of transgenes may be due to aberrant RNA derived from inverted repeats in the transgene (Muskens et al., 2000). If *Spm* methylation results from such a PTGS mechanism, transcription activation should enhance DNA methylation. However, this possibility is apparently not supported by our observation,

because none of the lines examined showed an increase in DNA methylation, regardless of the initial methylation status of the *Spm* sequence.

Another interesting observation from this study is that inducible *Spm* demethylation seems to depend on how the inducible *Spm-LUC* transgene is constructed. Only a small

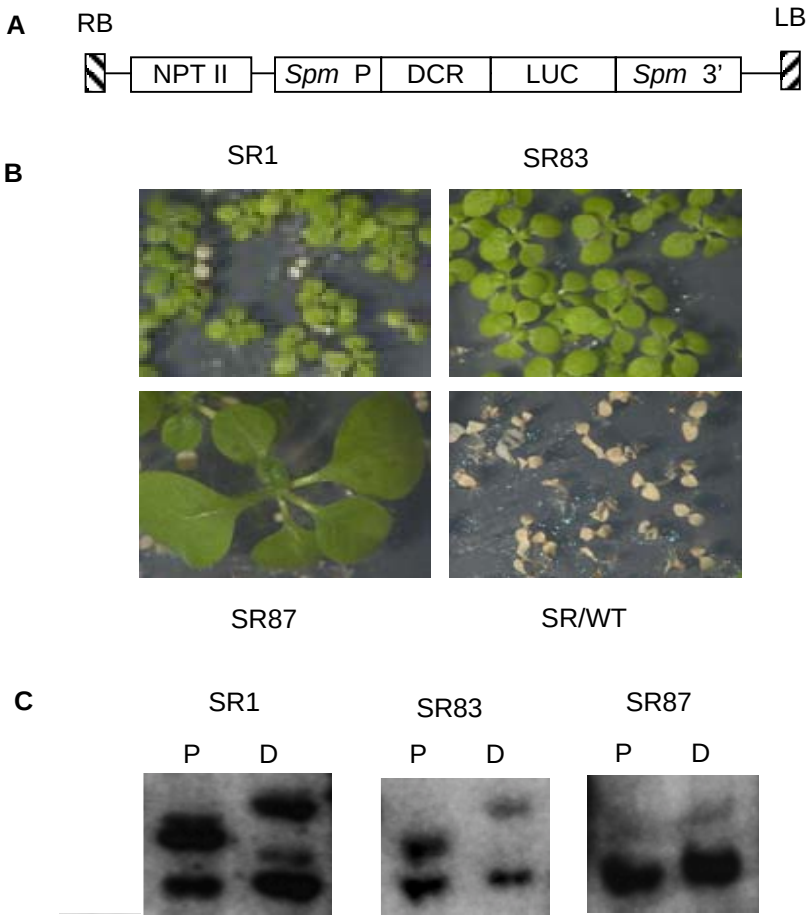


Figure 26. Transgenes in close proximity are co-regulated.

(A) A diagramatic representation of the *Spm* sequence and the *NPT II* gene, which confers resistance to kanamycin, in the vector used for tobacco transformation. (B) Three lines of transgenic tobacco plants grown on MS medium containing 50 mg/L kanamycin, and (C) DNA methylation of the *Spm* promoter (P) and the DCR (D) sequence. The upper and lower bands in each lane represent methylated and unmethylated DNA, respectively.

number of transgenic lines showed inducible *Spm* demethylation when the 5' end of the *Spm-LUC* construct is placed adjacent to the border of the T-vector (facing inward, Fig 25A). By contrast, no such transgenic lines were obtained when the 5' end of the *Spm-LUC* construct is placed adjacent to the hygromycin expression cassette (facing outward).

One likely explanation for this observation is that the *Spm* sequence and the GVG gene are more coregulated when they are in closer proximity. DNA methylation and silencing of flanking sequences has been reported in animal studies (Pikaart et al., 1998) and a similar phenomenon was noted in the present study of transgenic tobacco plants with the *Spm-LUC* construct. As shown in Fig. 26, plants with more extensive DNA methylation in the *Spm* sequence grew much slower than those with slight DNA methylation when grown on medium containing kanamycin, resistance to which is conferred by the *NPT* gene in the same vector. The transgene did not disrupt the normal growth of the transgenic plants, however, when they were grown in normal medium. This observation suggests that gene silencing or methylation in the *Spm* sequence may have spread over to the flanking *NPT* gene.

A similar situation may occur in the transgenic lines with the inducible *Spm*-LUC construct. One prediction from the hypothesis described above is that transcription was efficiently activated in lines with unmethylated DNA, but was not in those with methylated DNA. Such a correlation would lend further support to the foregoing observation that transcriptional activation causes DNA demethylation. In transgenic lines with the *Spm*-LUC transgene facing inwards, however, the GVG gene and the *Spm* sequence are flanked by different sequences and thus may behave more independently.

To overcome the problem of cosilencing, the inducible transcription factor GVG can be expressed using a different vector in lines with heavy DNA methylation.

References

Amedeo, P., Habu, Y., Afsar, K., Scheid, O.M., and Paszkowski, J. (2000). Disruption of the plant gene MOM releases transcriptional silencing of methylated genes. Nature **405,** 203-206.

Aoyama, T., and Chua, N.H. (1997). A glucocorticoid-mediated transcriptional induction system in transgenic plants. Plant J. **11,** 605-612.

Bordoli, L., Netsch, M., Luthi, U., Lutz, W., and Eckner, R. (2001). Plant orthologs of p300/CBP: conservation of a core domain in metazoan p300/CBP acetyltransferase-related proteins. Nucleic Acids Res. **29,** 589-597.

Brzeski, J., and Jerzmanowski, A. (2003). Deficient in DNA methylation 1 (DDM1) defines a novel family of chromatin-remodeling factors. J. Biol. Chem. **278,** 823-828.

Cervoni, N., and Szyf, M. (2001). Demethylase activity is directed by histone acetylation. J. Biol. Chem. **276,** 40778-40787.

Chen, Z., Ricigliano, J.W., and Klessig, D.F. (1993). Purification and characterization of a soluble salicylic acid-binding protein from tobacco. Proc. Natl. Acad. Sci. USA **90,** 9533-9537.

Gendrel, A.V., Lippman, Z., Yordan, C., Colot, V., and Martienssen, R.A. (2002). Dependence of heterochromatic histone H3 methylation patterns on the Arabidopsis gene DDM1. Science **297,** 1871-1873.

Grange, T., Cappabianca, L., Flavin, M., Sassi, H., and Thomassin, H. (2001). In vivo analysis of the model tyrosine aminotransferase gene reveals multiple sequential steps in glucocorticoid receptor action. Oncogene **20,** 3028-3038.

Jeddeloh, J.A., Bender, J., and Richards, E.J. (1998). The DNA methylation locus DDM1 is required for maintenance of gene silencing in Arabidopsis. Genes Dev. **12,** 1714-1725.

Jeddeloh, J.A., Stokes, T.L., and Richards, E.J. (1999). Maintenance of genomic methylation requires a SWI2/SNF2-like protein. Nat. Genet. **22,** 94-97.

Jost, J.P., Thiry, S., and Siegmann, M. (2002). Estradiol receptor potentiates, in vitro, the activity of 5-methylcytosine DNA glycosylase. FEBS Lett. **527,** 63-66.

Kaboev, O.K., Luchkina, L.A., and Kuziakina, T.I. (1985). Apurinic and apyrimidinic DNA endonuclease of extremely thermophilic Thermothrix thiopara. J. Bacteriol. **164,** 878-881.

Lafarge, S., and Montane, M.H. (2003). Characterization of Arabidopsis thaliana ortholog of the human breast cancer susceptibility gene 1: AtBRCA1, strongly induced by gamma rays. Nucleic Acids Res. **31,** 1148-1155.

Lax, S.R., Lauer, S.J., Browning, K.S., and Ravel, J.M. (1986). Purification and properties of protein synthesis initiation and elongation factors from wheat germ. Methods Enzymol. **118,** 109-128.

Leadon, S.A., and Avrutskaya, A.V. (1998). Requirement for DNA mismatch repair proteins in the transcription- coupled repair of thymine glycols in Saccharomyces cerevisiae. Mutat. Res. **407,** 177-187.

Li, B.H., and Bockrath, R. (1995). Benefit of transcription-coupled nucleotide excision repair for gene expression in u.v.-damaged Escherichia coli. Mol. Microbiol. **18,** 615-622.

Lovrien, R.E., and Matulis, D. (1997). Selective Precipitation of Proteins. In "Current Protocols in Protein Science". Edited by Coligan, JE, Dun, BM, Ploegh HL, Speicher, DW and Wingfield, PT **4.5.1-4.5.36**.

Masters, C.I., Moseley, B.E., and Minton, K.W. (1991). AP endonuclease and uracil DNA glycosylase activities in Deinococcus radiodurans. Mutat. Res. **254,** 263-272.

Matsuo, K., Silke, J., Georgiev, O., Marti, P., Giovannini, N., and Rungger, D. (1998). An embryonic demethylation mechanism involving binding of transcription factors to replicating DNA. EMBO J. **17,** 1446-1453.

Matzke, M.A., Mette, M.F., and Matzke, A.J. (2000). Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates. Plant Mol. Biol. **43,** 401-415.

Mellon, I., Rajpal, D.K., Koi, M., Boland, C.R., and Champe, G.N. (1996). Transcription-coupled repair deficiency and mutations in human mismatch repair genes. Science **272,** 557-560.

Muller, W.G., Walker, D., Hager, G.L., and McNally, J.G. (2001). Large-scale chromatin decondensation and recondensation regulated by transcription from a natural promoter. J. Cell Biol. **154,** 33-48.

Muskens, M.W., Vissers, A.P., Mol, J.N., and Kooter, J.M. (2000). Role of inverted DNA repeats in transcriptional and post-transcriptional gene silencing. Plant Mol. Biol. **43,** 243-260.

Nye, A.C., Rajendran, R.R., Stenoien, D.L., Mancini, M.A., Katzenellenbogen, B.S., and Belmont, A.S. (2002). Alteration of large-scale chromatin structure by estrogen receptor. Mol. Cell. Biol. **22,** 3437-3449.

Petrauskene, O.V., Gromova, E.S., Romanova, E.A., Volkov, E.M., Oretskaya, T.S., and Shabarova, Z.A. (1995). DNA duplexes containing methylated bases or non-nucleotide inserts in the recognition site are cleaved by restriction endonuclease R.EcoRII in presence of canonical substrate. Gene **157,** 173-176.

Pikaart, M.J., Recillas-Targa, F., and Felsenfeld, G. (1998). Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. Genes Dev. **12,** 2852-2862.

Probst, A.V., Fransz, P.F., Paszkowski, J., and Scheid, O.M. (2003). Two means of transcriptional reactivation within heterochromatin. Plant J. **33,** 743-749.

Schläppi, M., Raina, R., and Fedoroff, N. (1994). Epigenetic regulation of the maize Spm transposable element: novel activation of a methylated promoter by TnpA. Cell **77,** 427-437.

Sijen, T., Vijn, I., Rebocho, A., van Blokland, R., Roelofs, D., Mol, J.N., and Kooter, J.M. (2001). Transcriptional and posttranscriptional gene silencing are mechanistically related. Curr. Biol. **11,** 436-440.

Sturm, A., and Lienhard, S. (1998). Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast. Plant J. **13,** 815-821.

Thomassin, H., Flavin, M., Espinas, M.L., and Grange, T. (2001). Glucocorticoid-induced DNA demethylation and gene memory during development. EMBO J. **20,** 1974-1983.

Tini, M., Benecke, A., Um, S.J., Torchia, J., Evans, R.M., and Chambon, P. (2002). Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription. Mol. Cell **9,** 265-277.

Tumbar, T., Sudlow, G., and Belmont, A.S. (1999). Large-scale chromatin unfolding and remodeling induced by VP16 acidic activation domain. J. Cell Biol. **145,** 1341-1354.

Vaucheret, H., and Fagard, M. (2001). Transcriptional gene silencing in plants: targets, inducers and regulators. Trends Genet. **17,** 29-35.

Yuan, L.W., and Giordano, A. (2002). Acetyltransferase machinery conserved in p300/CBP-family proteins. Oncogene **21,** 2253-2260.

Zhu, B., Zheng, Y., Hess, D., Angliker, H., Schwarz, S., Siegmann, M., Thiry, S., and Jost, J.P. (2000). 5-methylcytosine-DNA glycosylase activity is present in a cloned G/T mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex. Proc. Natl. Acad. Sci. USA **97,** 5135-5139.

Zhu, B., Benjamin, D., Zheng, Y., Angliker, H., Thiry, S., Siegmann, M., and Jost, J.P. (2001). Overexpression of 5-methylcytosine DNA glycosylase in human embryonic kidney cells EcR293 demethylates the promoter of a hormone-regulated reporter gene. Proc. Natl. Acad. Sci. USA **98,** 5031-5036.

Chapter 8

CONCLUSION

Mechanism of TnpA-Mediated DNA Demethylation

Programmed DNA methylation and demethylation are essential for the normal growth and development of both animals and plants. Although much has been learned about the mechanisms that control *de novo* and maintenance DNA methylation (Bestor, 2000; Finnegan and Kovac, 2000; Kishimoto et al., 2001; Lindroth et al., 2001; Cao and Jacobsen, 2002a, b), our understanding of DNA demethylation is still rudimentary. Genome-wide demethylation, followed by the imposition of a gender- or tissue-specific methylation pattern, is likely to be one mechanism by which differential epigenetic control is achieved. Differential gene expression is also achieved in development by targeted demethylation of different subgroups of genes in different tissues. In animals, some trans-acting factors are known to promote demethylation of DNA sequences to which they bind. Several proteins have also been found to be capable of converting 5-methylcytosine into cytosine by different routes. However, whether and how these enzymes participate in the sequence-specific DNA demethylation mediated by trans-acting factors is largely unknown. Virtually nothing is known in plants, although there is evidence that DNA demethylation occurs in plants as well.

In this study, I carried out experiments aimed at dissecting the mechanism by which the maize *Spm* transposon-encoded TnpA protein promotes demethylation of the *Spm* promoter and its downstream GC-rich sequence. Using an inducible promoter to control TnpA expression in transgenic tobacco cells that harbor a methylated *Spm* transgene, I was able to follow the process of DNA demethylation during DNA replication. I found that TnpA-mediated DNA demethylation is an active process, much more rapid than can be explained by blockage of DNA remethylation. This observation suggests that TnpA is either a sequence-specific DNA demethylase or it recruits such an enzyme to the *Spm* sequence. To further understand the underlying mechanism, I developed an *in vitro* DNA demethylation assay that is designed specifically for sensitive detection of TnpA-associated DNA demethylase activity. Indeed, I detected DNA demethylase activity in nuclear extracts prepared from tobacco suspension cells. By an immunoprecipitation approach, I purified native TnpA and its truncated derivatives from tobacco cells that overexpress these proteins. However, no direct interactions between TnpA and DNA demethylase activity was detected. Since the purified TnpA protein still binds well to the *Spm* promoter sequence, my results clearly rule out the possibility that TnpA itself is a DNA demethylase. They further suggest that a DNA demethylase is recruited to the *Spm* sequence through a mechanism that does not involve direct protein-protein interaction.

A prominent feature of TnpA-mediated DNA demethylation is that not only the *Spm* promoter, which contains multiple TnpA binding sites, but also its downstream sequence, gets demethylated. In further experiments, I found that the TnpA protein is a transcriptional activator specific to the *Spm* promoter and this activity appears to be essential for TnpA to promote DNA demethylation. On the one hand, truncated TnpA

derivatives with a deletion of the C-terminal 80 or 190 amino acids lack both the abilities to activate transcription and to promote *Spm* demethylation. On the other hand, the fusion protein between the truncated $TnpA_{540}$ protein and the viral VP16 activation domain promotes *Spm* demethylation more efficiently than the full-length TnpA protein. Demethylation of the *Spm* sequence was also observed when transcription from the *Spm* sequence was activated from an upstream inducible promoter by a VP16 activation domain-containing transcription factor. These findings strongly suggest that a DNA demethylase activity is recruited to the *Spm* sequence as a consequence of transcriptional activation. Since most other trans-acting factors that promote sequence-specific DNA demethylation are known transcriptional activators, the present finding may have revealed a common mechanism of active DNA demethylation. Conceivably, a DNA demethylase could be targeted to actively transcribing sequences through the transcription initiation complex or the transcription elongation complex. Alternatively, a DNA demethylase may gain access to methylated DNA passively as a result of chromatin remodeling.

Although TnpA-mediated DNA demethylation is very rapid, it occurs at a rather slow rate in the initial stage of TnpA-DNA interaction. In agreement with this observation, the present study demonstrated that DNA replication is required for TnpA to initiate the demethylation event. The requirement for an additional step lies in the fact that TnpA's binding is much weaker to fully methylated DNA compared with hemimethylated DNA. Based on these observations, a two-step mechanism is proposed for TnpA-mediated DNA demethylation: 1) An initial passive process where TnpA binds to post replicative, unmethylated DNA and prevents its remethylation; 2) A subsequent

active mechanism, which follows transcription activation by TnpA and involves a DNA demethylase activity.

Significance of the Present Findings

The present study has revealed two essential elements of sequence-specific active DNA demethylation mediated by trans-acting factors. Whereas the DNA-binding ability defines target specificity, the transcriptional activation domain appears to facilitate the access of DNA demethylases to methylated DNA, which is otherwise embedded in compact heterochromatin. When DNA binding is affected by DNA methylation in the target sequence, DNA replication is required for the trans-acting factors to bind hemimethylated DNA. Of course, another critical prerequisite for active DNA demethylation is that DNA demethylase activity must be available when trans-acting factor is expressed.

This understanding can account for several distinct DNA demethylation patterns mediated by other trans-acting factors. For example, a one-step active mechanism has been reported for DNA demethylation by transcription factors such as the retinoid and estradiol receptors, as well as the glucocorticoid receptor (Grange et al., 2001; Thomassin et al., 2001; Zhu et al., 2001; Jost et al., 2002), whose binding to DNA is through other proteins that are already anchored in the target DNA and thus is not affected by DNA methylation. In the case of TnpA, however, an initial passive process is required for

TnpA to bind to hemimethylated DNA and usher in the subsequent step of active mechanism.

A similar two-step DNA demethylation mechanism has been reported for the transcriptional activator NF-kappa B (Matsuo et al., 1998) and the Epstein-Barr virus latent replication origin protein EBNA-1 (Hsieh, 1999). Although the EBNA-1 protein has not been associated with transcriptional activity, it is well known that DNA replication is generally accompanied by transcription and histone hyperacetylation (Vogelauer et al., 2002). Notably, DNA demethylation due to the EBNA-1 protein is restricted to the region where the protein binds. This is in sharp contrast to that mediated by transcriptional factors such as TnpA, which promote demethylation of sequences well beyond the binding sites. The LacI repressor represents another type of trans-acting factors. Although it has been shown to be able to maintain its binding sites unmethylated, there is no evidence that it can promote active DNA demethylation (Lin et al., 2000; Lin and Hsieh, 2001). In this case, the observed DNA demethylation is probably a result of maintenance of DNA unmethylation rather than loss of DNA methylation. Similar activity has been reported for the CTCF protein, a multifunctional protein (Ohlsson et al., 2001) that can prevent DNA methylation of its binding site in the imprinting center (Hark et al., 2000; Chao et al., 2002). Unlike the LacI protein, however, CTCF binds to its binding sites only in the maternal X chromosome that will undergo inactivation. Importantly, binding of CTCF to its cognate sequence is inhibited by DNA methylation and demethylation in the paternal X-chromosome has not been observed (Hark et al., 2000). Although CTCF can be a transcriptional activator or repressor dependent on the locus as well as protein modifications or interaction partners, it appears to be a

transcriptional repressor in the imprinting locus because it interacts with a histone deacetylase (Chernukhin et al., 2000; Lutz et al., 2000; Klenova et al., 2001). This observation underscores the importance of the transcriptional domain of trans-acting factors in promoting DNA demethylation.

The present findings have practical implications as well. Genetic engineering, which entails the expression of exogenous genes to confer desired traits in plants, has become a powerful tool for crop improvement. However, a major difficulty that has surfaced in this approach is that transgenes are often transcriptionally silenced and in many cases become methylated (Vaucheret et al., 1998; Matzke et al., 2000). Understanding how genes are selectively demethylated is likely to have practical applications in improving our ability to maintain transgenes in an active state.

Another practical application of understanding targeted gene demethylation is to enhance our ability to use homologous recombination to alter gene sequences *in situ*. DNA methylation not only interferes with transcription, but also with recombination, possibly due to the formation of compact chromatin structures around methylated DNA. This has been shown in the fungus *Ascobolus* (Maloisel and Rossignol, 1998) and is very likely to be true in other organisms with extensive DNA methylation, particularly plants (Puchta, 2002). The ability to minimize homologous ectopic recombination is undoubtedly an important mechanism for maintaining stability in genomes with large populations of transposons and retrotransposons (Fedoroff, 1999). But the extremely low frequency of homologous recombination in plants presents a major obstacle to the successful use of homologous recombination to inactive genes, as well as to modify gene sequences for improved agricultural performance or nutritional quality. Targeted DNA

demethylation could therefore be useful to break the barrier that prevents homologous recombination.

The present study on the mechanism by which the *Spm*-encoded TnpA protein promotes *Spm* demethylation has not only filled a significant gap in our understanding of epigenetic regulation of gene expression, but may also allow us to manipulate the expression of agronomically important genes. Because the principle of DNA demethylation appears similar in plants and animals, the knowledge obtained from this study could be useful as well in designing new strategies for gene therapy.

Future Directions

The main question that remains to be answered in future studies is how DNA demethylase activity is recruited to the *Spm* sequence. Specifically, it is necessary to determine whether DNA demethylase activity is present in the transcription initiation complex or the transcription elongation complex or whether it reaches methylated DNA as a result of chromatin remodeling. Moreover, there is also a need to identify and characterize the DNA demethylase activity that participates in TnpA-mediated *Spm* demethylation.

Although a DNA demethylase activity was detected in the nuclear extracts prepared from tobacco suspension cultured cells, no physical interaction between this enzyme and the TnpA protein was revealed in the present study. However, a DNA demethylase that weakly interacts with TnpA might have been lost during the stringent washing step in

preparing the TnpA-immunoprecipitate. To address this question, native blue-gel electrophoresis (BN-PAGE) can be used to compare the apparent molecular size of TnpA in cell extract and the TnpA-containing immunoprecipitate. The BN-PAGE method is a modified 2-D gel electrophoresis technique that was originally developed for characterization of protein complexes in mitochondria and later found wide applications in proteomics (Schagger and von Jagow, 1991; Schagger, 2001; Brookes et al., 2002; Nijtmans et al., 2002). If TnpA turns out to be in a protein complex, this approach can help in optimizing the conditions for immunoprecipitation.

As a transcriptional activator, TnpA could facilitate the binding of a DNA demethylase to the *Spm* sequence through the transcription machinery or by exposing binding sites in a chromatin context. TnpA may interact with other proteins only when bound to DNA. For example, conformational changes and increased protein-protein interaction induced by DNA binding have been reported for the Fos and Jun proteins, which form heterodimers through a leucine-zipper domain (Patel et al., 1990). Thus, it would be interesting to determine whether a DNA demethylase is recruited to the *Spm* promoter sequence in the presence of TnpA.

It has been demonstrated that TnpA binds strongly to the *Spm* promoter sequence when unmethylated and hemimethylated, but not fully methylated. By forming homodimers, TnpA also creates large intermolecular DNA-protein complexes (Raina et al., 1998). To facilitate purification of such complexes, biotinylated hemimethylated *Spm* promoter fragments will be used to pull down TnpA and DNA-binding proteins whose binding to DNA depends on TnpA-mediated transcriptional activation. Use of

biotinylated DNA would allow easy recovery of the DNA-protein complex using paramagnetic streptavidin beads (Kroeger and Abraham, 1997).

Another question that remains to be answered is whether transcription *per se* is required for *Spm* demethylation. This can be investigated by using the inducible transcription system, as has been discussed in Chapter 7.

In the light of the recent revelation about the interplay between histone modifications, chromatin structure and DNA demethylation (Richards and Elgin, 2002) as well as the current finding that transcriptional activation leads to *Spm* demethylation, it would be very interesting to examine changes in these elements during the course of DNA demethylation. Changes in histone modification from methylation to hyperacetylation has recently been reported during transcription activation at the dihydrofolate reductase promoter (Nicolas et al., 2003). In this regard, both the TnpA inducible system and the inducible *Spm* transcription system would be useful because they permit a dynamic study of the DNA demethylation process.

The inducible system for controlling TnpA expression or transcription at the *Spm* sequence has yielded much information that leads to the current understanding of the mechanism of TnpA-mediated DNA demethylation. Still, there are many more questions that can be addressed using the inducible system. For example, in combination with bisulfite sequencing (Clark et al., 1994), a technique for mapping all cytosine sites, it would permit one to examine the dynamic change in DNA methylation in the course of TnpA expression. Such a study would address a range of interesting questions. For example, do the two strands get demethylated differentially? Does demethylation occur distributively or progressively from the TnpA binding sites? Does 5mC in every context

(CpG, CpNpG and non-asymmetric position) get demethylated with the same efficiency? Gain of DNA methylation has been observed in transgenic tobacco calli after withdrawal of the inducing chemical. This will allow a detailed and dynamic study of the process of *de novo* methylation.

Dissecting the mechanism by which the *Spm* sequence gets methylated is also important for us to understand how TnpA reverses this epigenetic mark. One of the efforts that have been made is to reproduce the *Spm* epigenetic phenomenon (DNA methylation and demethylation) in *Arabidopsis*, so as to take advantage of the available rich genetic resources. Among hundreds of transgenic plants examined, however, none show a trace of methylation in the *Spm* sequence. This observation reflects some fundamental difference between tobacco and *Arabidopsis*. One possible reason for the lack of DNA methylation in the *Spm* sequence in transgenic *Arabidopsis* plants is the lack of trans-acting factors that recognize the *Spm* sequence for *de novo* methylation. As histone deacetylation and methylation have been shown to be the determinants of DNA methylation, it would be interesting to test whether fusion of a histone deacetylase or methyltransferase domain to truncated TnpA derivatives would bring about *de novo* methylation of the *Spm* sequence. Alternatively, the *Spm* sequence could remain unmethylated because of factors that bind and maintain an active status of the *Spm* transgene. These possibilities need to be examined as well in future studies.

Lastly, but not least importantly, the DNA demethylase activity that participates in TnpA-mediated DNA demethylation needs to be identified and characterized. If it turns out that this activity is not in the transcription complex, it might be necessary to purify DNA demethylase activities from cell extracts by other separation methods. After the

corresponding genes are cloned, their roles in TnpA-mediated *Spm* demethylation can be determined by examining TnpA's ability to promote *Spm* demethylation in transgenic plants with the respective gene silenced.

References

Bestor, T.H. (2000). The DNA methyltransferases of mammals. Hum. Mol. Genet. **9,** 2395-2402.

Brookes, P.S., Pinner, A., Ramachandran, A., Coward, L., Barnes, S., Kim, H., and Darley-Usmar, V.M. (2002). High throughput two-dimensional blue-native electrophoresis: A tool for functional proteomics of mitochondria and signaling complexes. Proteomics **2,** 969-977.

Cao, X., and Jacobsen, S.E. (2002a). Role of the Arabidopsis DRM Methyltransferases in De Novo DNA Methylation and Gene Silencing. Curr. Biol. **12,** 1138-1144.

Cao, X., and Jacobsen, S.E. (2002b). Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. Proc. Natl. Acad. Sci. USA.

Chao, W., Huynh, K.D., Spencer, R.J., Davidow, L.S., and Lee, J.T. (2002). CTCF, a candidate trans-acting factor for X-inactivation choice. Science **295,** 345-347.

Chernukhin, I.V., Shamsuddin, S., Robinson, A.F., Carne, A.F., Paul, A., El-Kady, A.I., Lobanenkov, V.V., and Klenova, E.M. (2000). Physical and functional interaction between two pluripotent proteins, the Y-box DNA/RNA-binding factor, YB-1, and the multivalent zinc finger factor, CTCF. J. Biol. Chem. **275,** 29915-29921.

Clark, S.J., Harrison, J., Paul, C.L., and Frommer, M. (1994). High sensitivity mapping of methylated cytosines. Nucleic Acids Res. **22,** 2990-2997.

Finnegan, E.J., and Kovac, K.A. (2000). Plant DNA methyltransferases. Plant Mol. Biol. **43,** 189-201.

Grange, T., Cappabianca, L., Flavin, M., Sassi, H., and Thomassin, H. (2001). In vivo analysis of the model tyrosine aminotransferase gene reveals multiple sequential steps in glucocorticoid receptor action. Oncogene **20,** 3028-3038.

Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature **405,** 486-489.

Hsieh, C.L. (1999). Evidence that protein binding specifies sites of DNA demethylation. Mol. Cell. Biol. **19,** 46-56.

Jost, J.P., Thiry, S., and Siegmann, M. (2002). Estradiol receptor potentiates, in vitro, the activity of 5-methylcytosine DNA glycosylase. FEBS Lett. **527,** 63-66.

Kishimoto, N., Sakai, H., Jackson, J., Jacobsen, S.E., Meyerowitz, E.M., Dennis, E.S., and Finnegan, E.J. (2001). Site specificity of the Arabidopsis METI DNA

methyltransferase demonstrated through hypermethylation of the superman locus. Plant Mol. Biol. **46,** 171-183.

Klenova, E.M., Chernukhin, I.V., El-Kady, A., Lee, R.E., Pugacheva, E.M., Loukinov, D.I., Goodwin, G.H., Delgado, D., Filippova, G.N., Leon, J., Morse, H.C., 3rd, Neiman, P.E., and Lobanenkov, V.V. (2001). Functional phosphorylation sites in the C-terminal region of the multivalent multifunctional transcriptional factor CTCF. Mol. Cell. Biol. **21,** 2221-2234.

Kroeger, K.M., and Abraham, L.J. (1997). Magnetic bead purification of specific transcription factors using mutant competitor oligonucleotides. Anal. Biochem. **250,** 127-129.

Lin, I.G., Tomzynski, T.J., Ou, Q., and Hsieh, C.L. (2000). Modulation of DNA binding protein affinity directly affects target site demethylation. Mol. Cell. Biol. **20,** 2343-2349.

Lin, I.G., and Hsieh, C.L. (2001). Chromosomal DNA demethylation specified by protein binding. EMBO Rep. **2,** 108-112.

Lindroth, A.M., Cao, X., Jackson, J.P., Zilberman, D., McCallum, C.M., Henikoff, S., and Jacobsen, S.E. (2001). Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. Science **292,** 2077-2080.

Lutz, M., Burke, L.J., Barreto, G., Goeman, F., Greb, H., Arnold, R., Schultheiss, H., Brehm, A., Kouzarides, T., Lobanenkov, V., and Renkawitz, R. (2000). Transcriptional repression by the insulator protein CTCF involves histone deacetylases. Nucleic Acids Res. **28,** 1707-1713.

Maloisel, L., and Rossignol, J.L. (1998). Suppression of crossing-over by DNA methylation in Ascobolus. Genes Dev. **12,** 1381-1389.

Matsuo, K., Silke, J., Georgiev, O., Marti, P., Giovannini, N., and Rungger, D. (1998). An embryonic demethylation mechanism involving binding of transcription factors to replicating DNA. EMBO J. **17,** 1446-1453.

Matzke, M.A., Mette, M.F., and Matzke, A.J. (2000). Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates. Plant Mol. Biol. **43,** 401-415.

Nicolas, E., Roumillac, C., and Trouche, D. (2003). Balance between Acetylation and Methylation of Histone H3 Lysine 9 on the E2F-Responsive Dihydrofolate Reductase Promoter. Mol. Cell. Biol. **23,** 1614-1622.

Nijtmans, L.G., Henderson, N.S., and Holt, I.J. (2002). Blue Native electrophoresis to study mitochondrial and other protein complexes. Methods **26,** 327-334.

Ohlsson, R., Renkawitz, R., and Lobanenkov, V. (2001). CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. Trends Genet. **17,** 520-527.

Patel, L., Abate, C., and Curran, T. (1990). Altered protein conformation on DNA binding by Fos and Jun. Nature **347,** 572-575.

Puchta, H. (2002). Gene replacement by homologous recombination in plants. Plant Mol. Biol. **48,** 173-182.

Raina, R., Schläppi, M., Karunanandaa, B., Elhofy, A., and Fedoroff, N. (1998). Concerted formation of macromolecular Suppressor-mutator transposition complexes. Proc. Natl. Acad. Sci. USA **95,** 8526-8531.

Richards, E.J., and Elgin, S.C. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. Cell **108,** 489-500.

Schagger, H., and von Jagow, G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. Anal. Biochem. **199,** 223-231.

Schagger, H. (2001). Blue-native gels to isolate protein complexes from mitochondria. Methods Cell Biol. **65,** 231-244.

Thomassin, H., Flavin, M., Espinas, M.L., and Grange, T. (2001). Glucocorticoid-induced DNA demethylation and gene memory during development. EMBO J. **20,** 1974-1983.

Vaucheret, H., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J.B., Mourrain, P., Palauqui, J.C., and Vernhettes, S. (1998). Transgene-induced gene silencing in plants. Plant J. **16,** 651-659.

Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B.J., and Grunstein, M. (2002). Histone acetylation regulates the time of replication origin firing. Mol. Cell **10,** 1223-1233.

Zhu, B., Benjamin, D., Zheng, Y., Angliker, H., Thiry, S., Siegmann, M., and Jost, J.P. (2001). Overexpression of 5-methylcytosine DNA glycosylase in human embryonic kidney cells EcR293 demethylates the promoter of a hormone-regulated reporter gene. Proc. Natl. Acad. Sci. USA **98,** 5031-5036.

Curriculum Vitae

Hongchang Cui

Education

Ph.D., Plant Physiology, Penn State University. 2003.
MS in Botany, Department of Biology, Peking University, Beijing. 1990.
BS in Botany, Department of Biology, Peking University, Beijing. 1987.

Awards and Fellowships

1. The Alumni Association Dissertation Award, Penn State University. 2003.
2. Keystone conference scholarship for the meeting "Epigenetics in Development and Disease." Taos, New Mexico. 2002
3. Nina Fedoroff Teaching Assistant Award, Penn State University. 2000.
4. Life Sciences Consortium Fellowship, Penn State University.1997-99.
5. Third prize for Advancement in Science & Technology, Beijing Municipal government, 1995.
6. First Prize for academic achievement, Beijing Academy of Agriculture & Forestry, 1994.
7. Second prize for Advancement in Science & Technology, Beijing Municipal government, 1993.
8. "Guanghua" Award for Excellence in Graduate Studies, Peking University. 1990.
9. Award for Academic Excellence in Undergraduate Studies, Peking University. 1987.

Major Publications

1. Cui, H. and N.V. Fedoroff, 2002. Inducible DNA demethylation mediated by the maize *Spm* transposon-encoded TnpA protein. Plant Cell, 14 (11): 2883-99.
2. Cui, H., J. Liu, and S. Sun, 1999. Improved conditions for the purification of phosphoglucomutase (PGM) isoenzymes from Chinese cabbage and studies on their molecular properties. Acta Agriculture Borealis-Sinica, 14(1):134-140.
3. Cui, H., L. Li, X. Zheng and S. Sun, 1998. Purification of the heat-tolerance related isozyme of phosphoglucomutase from Chinese cabbage. Acta Agriculture Borealis-Sinica, 13(4): 86-92.
4. Cui, H., 1996. Investigation of the major browning inhibitory factors in honey. Acta Agriculture Borealis-Sinica, 11(4): 125-128.
5. Cui, H. and X. Zheng, 1993. Application of HPLC method to seed variety identification of Chinese cabbage. Seed, 6: 53-58.
6. Cui, H., 1993. Optimization of HPLC conditions for the separation of water-soluble proteins in Chinese cabbage seeds. Seed, 4: 7-12.
7. Cui, H. and Z. Li. Studies on the regeneration of the vascular system in new bark of *Broussonetia paperyferus* Vent. after girdling. Collection of Graduate Theses of Peking University, 1990.