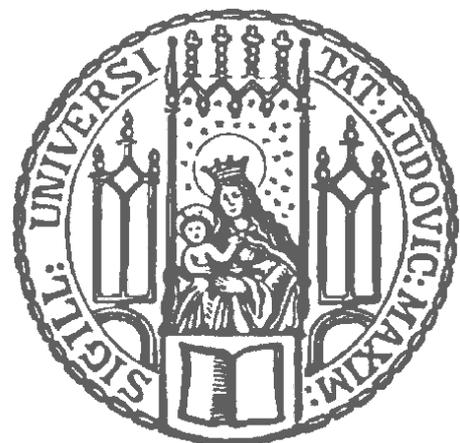# Role and regulation of DNA methylation in mouse embryonic stem cells

Daniela Meilinger

München 2011

Role and regulation of DNA methylation in mouse embryonic stem cells

Daniela Meilinger

Dissertation an der Fakultät für Biologie

der Ludwig-Maximilians-Universität München

vorgelegt von

Daniela Meilinger aus Koblenz

München, den 20. September 2011

Erstgutachter: Prof. Dr. Heinrich Leonhardt

Zweitgutachter: PD Dr. Anna Friedl

Tag der mündlichen Prüfung: 07.11.2011

Table of Contents

Summary

DNA methylation belongs to a complex epigenetic regulatory network of processes and has been shown to be essential for normal embryonic development and plays a crucial role in gene silencing. Dnmt1 is the major DNA methyltransferase responsible for the inheritance of DNA methylation during replication and its catalytic activity is tightly regulated by its N-terminal domain, as well as accessory interacting factors. Our primary goal was to gain further insight on the role and regulation of crucial subunits within the N-terminal domain of Dnmt1 and decipher their involvement in the enzyme's catalytic activity in living cells. Moving towards this end, we investigated the importance of the PCNA binding domain (PBD) and the CXXC zinc finger motif in mouse embryonic stem (ES) cells. For this purpose, we performed complementation and rescue assays in *dnmt1*[-/-] mouse ES cells, which allowed the analysis of the catalytic activity of Dnmt1 *in vivo*. In addition to classical methylation detection methods, such as bisulfite sequencing and COBRA, we established a quantitative method to analyze DNA methylation profiles by pyrosequencing. Using these approaches we showed that both N-terminal subunits of Dnmt1, PBD and CXXC zinc finger motif seem to be dispensable for the enzyme's catalytic activity in living cells, although the PBD interaction with PCNA, enhances the activity of Dnmt1 by two-fold.

Next we attempted to address a rather controversial function of Dnmt1, which is its possible involvement in *de novo* methylation of unmethylated DNA sequences *in vivo, a* function, which is first and foremost established by Dnmt3a and Dnmt3b. To this aim, we developed a new FACS based approach for the generation of stable ES cells lines and succeeded in establishing triple knockout (*dnmt1*[-/-], *dnmt3a*[-/-] *dnmt3b*[-/-]) ES cells that stably express a Dnmt1 expression construct. Pyrosequencing was employed as a highly sensitive method, in order to monitor even subtle changes on the methylation level changes of those cell lines, aiming to clarify the possible *de novo* methylation activity of Dnmt1. Using this approach, we succeeded in clearly demonstrating that Dnmt1 has no detectable *de novo* methylation activity in this system, favoring its unique function in maintaining DNA methylation.

Finally, due to the important known interaction between Dnmt1 and Np95, which has been shown to be crucial for DNA methylation maintenance, we were interested in looking into a possible interaction between Np95 and the *de novo* methyltransferases, Dnmt3a and Dnmt3b and its functional significance. By co-immunoprecipitation experiments, we found that there is an even stronger interaction between Np95 and the *de novo* methyltransferases compared to its interaction with Dnmt1. To resolve a possible function and therefore, the physiological role of this interaction, we developed a fluorescent CMV promoter-silencing assay *in vivo*. By the use of this assay we were able to detect an involvement of several epigenetic factors in this silencing mechanism. Interestingly, we found that apart from Np95, Dnmt3a and Dnmt3b, additionally histone-modifying

Summary

enzymes, such as G9a, are involved in silencing of the CMV promoter. Moreover, we clearly showed that DNA methylation is not required for the initiation of this silencing process. These findings point towards a silencing mechanism where histone modifications mediate the first step, which is then followed by DNA methylation. Still, to which extend this mechanism exists *in vivo* had to be further investigated. Considering the above, we proceeded by studying the involvement of Np95 and the active methyltransferases on silencing of endogenous promoter regions *in vivo*. For this purpose we differentiated wild type, *dnmt* and *np95* knockout cell lines into embryoid bodies and monitored methylation level changes of pluripotency promoter regions during differentiation. In this system and in contrast to our previous findings, we observed that silencing of the pluripotency factors *oct4* and *nanog* is initially independent of Np95 and the DNA methyltransferases Dnm1, Dnmt3a and Dnmt3b. Taken together, our data favor a silencing process that is initially independent of DNA methylation and suggest that DNA methylation is required for long term and stable gene silencing.

1. [Introduction]

C.H. Waddington first introduced the Epigenetic Landscape as a metaphor for biological development in 1945. He coined the term epigenetics from epi (Greek for over, above) and genetics to describe the study of "casual interactions between genes and their products, which bring their phenotypes into being." (Waddington,1954). In mammals and other multicellular organisms, all cells carry the same genetic information encoded in the deoxyribonucleic acid (DNA) sequence. Yet different cell lineages show a broad range of functional and morphologic diversity, which results from different patterns of transcriptionally active and silent genes. Epigenetic research studies the mechanisms that generate and maintain different patterns of gene expression without changing the DNA sequence. On one hand, epigenetic processes stabilize expression patterns to preserve cell type identity, while on the other hand, they must allow cells to respond to environmental and developmental cues and therefore must be highly dynamic. The molecular basis of epigenetic processes is complex and includes DNA and histone modifications. A large number of histone modifying enzymes can alter histone proteins through methylation, acetylation, phosphorylation and ubiquitination.

The best-characterized DNA modification is the covalent addition of a methyl group to the 5-carbon atom of cytosine (C), preferentially in a context followed by a guanine (G). These dinucleotides are called CpG sites and are surprisingly underrepresented in the genome, most likely because they are mutation hot spots (Coulondre et al, 1978) leading to CpG depletion during evolution. Apart from single CpG sites there are regions dense of CpG sites called CpG islands. In the human genome, around 30.000 CpG islands were identified. They are often found in promoter regions and first exons of many housekeeping genes where they are associated with regulatory functions (Bird, 1986; Antequera, 2003).

In contrast to sparsely distributed CpG sites, which are highly methylated, CpG islands and sites within regulatory elements are often unmethylated (Antequera & Bird, 1993). Methylation in regulatory regions is inversely correlated with transcriptional activity of the associated gene. The correlation of methylation and gene activity makes the methylation analysis of regulatory regions an attractive tool to study gene activity. DNA methylation is involved in transcriptional repression by at least two described mechanisms. On one hand,

DNA methylation may directly preclude binding of transcription factors by preventing proteins that specifically bind methylated CpG sites to bind to their recognition site. On the other hand, methylcytosine-binding proteins recruit repressive chromatin modifiers and remodeling factors (Bird, 1986). DNA methylation is involved in gene silencing and has been shown to be crucial for embryonic development (Okano et al, 1999), genomic imprinting, (Li et al, 1993) X-inactivation in mammals (Beard et al, 1995) and silencing of retroviral DNA elements (Dong et al, 2008).

5-methylcytosine (5-mC) is often considered the 5th base in the genome. Recently, a family of proteins (TET) has been described that catalyzes the conversion from methylcytosine to hydroxymethylcytosine, which might be considered as the 6th base in the genome (Tahiliani et al, 2009). This modification has initially been found in genomic DNA of neurons (Kriaucionis & Heintz, 2009) and mouse embryonic stem cells (Tahiliani et al, 2009). The function of this modification is still unclear. Hydroxymethylcytosine might be an intermediate product in the pathway of active demethylation and an epigenetic signal.

1.1   Role of DNA Methylation in Development and Disease

During development and differentiation distinct sets of genes are either activated or silenced by different epigenetic processes. These processes involve DNA methylation, histone modification, polycomb repressive complexes, as well as non-coding RNAs and lead to functional and morphological specialization of cell types and lineages. Once established, these patterns are maintained through cell division cycles and thus contribute to the identity of the differentiated state.

Mammalian development is accompanied by dramatic changes in the epigenetic state, especially methylation of tissue specific genes, as well as global DNA methylation levels. Genome wide epigenetic reprogramming occurs in two waves of demethylation, first in early embryos and second in primordial germ cells (PGC's) (Reik et al, 2001). During preimplantation, the paternal genome is rapidly and most likely actively demethylated while the maternal genome becomes passively demethylated through cell divisions until blastocyst stage (Reik et al, 2001). These passive and active demethylation processes remove most of the parental methylation marks. Subsequently, a wave of *de novo*

methylation establishes new cell type specific methylation profiles after implantation. Genomic imprinting is a specialized developmental mechanism where genes are allele specifically expressed depending on the parental origin. Only a small number of all genes are special imprinted genes and they are often organized in clusters.

To investigate the role of methylation, mice with reduced methyltransferase activity were generated (Li et al, 1992; Li et al, 1993). The analysis of imprinted genes in these mice, such as H19, insulin-like growth factor 2 (Igf-2), and Igf-2 receptor (Igf-2r) has clearly demonstrated that DNA methylation is involved in the control of differential expression (Li et al, 1993). In PCGs methylation of imprints are erased and according to the parental origin newly established. However, imprinted genes are protected from this wave of erasure and *de novo* methylation that takes place after fertilization (Morgan et al, 2005).

Specific patterns of gene expression are important to control cell identity and proliferation. Misregulation of DNA methylation can lead to tumorigenesis. Global, as well as gene-specific changes of DNA methylation have been observed in the analysis of tumor cells. While cancer cells are locally hypermethylated they show global hypomethylation, which is often accompanied by chromosomal instability and nuclear disorganization (Dodge et al, 2005). Additionally, hypomethylation can lead to oncogene expression, which directly affects growth control of cancer cells. Gaudet and colleagues explored the involvement of DNA hypomethylation in tumorigenesis by generating hypomorphic *dnmt1* mice (Gaudet, 2003). Those mice showed a 90% reduced Dnmt1 expression level, were growth-restricted and developed aggressive T-cell tumors several months after birth. Local hypermethylation often occurs in regulatory regions and CpG islands of tumorsuppressor genes and thus leads to their silencing. During tumorigenesis, epigenetic silencing often affects key signaling pathways by either disruption or activation of involved genes and is proposed to be an initial event (Feinberg & Tycko, 2004; Yamada & Jackson-Grusby, 2005). The effect of DNA methylation was also extensively studied in the human colon cancer cell line HCT116, where hypermethylation of the *sfrp* genes was found to cause activation of the *Wnt* signaling pathway, and thereby aberrantly increased proliferation and expansion of stem-cell populations (Taketo, 2004). Moreover, in colon, lung and breast cancer cells the CDKN/p16 regulatory region was *de novo* methylated and thereby inactivated (Gonzalez-Quevedo et al, 2004).

Introduction

Epigenetic mutations are in contrast to genetic alterations reversible, which makes them an interesting target for cancer therapy. For instance, DNA methyltransferase inhibitors like the nucleoside analogue 5`-aza-deoxycytidine (Decitabine) are already used in clinical studies for the treatment of breast and colon cancer (Yoo & Jones, 2006).

Apart from involvement in tumor formation, epigenetic disregulations have also been shown to be involved in several other diseases, such as ICF (Immune deficiency, centromeric instability and facial abnormalities), Fragile X and RETT syndrome as well as the imprinting disorders Prader-Willi and Angelmann syndrome (Amir et al, 1999; Basehore & Friez, 2009; Zeschnigk et al, 1997; Hansen et al, 1999; Cassidy & Schwartz, 1998). The ICF syndrome is caused by a missense mutation in the *dnmt3b* gene affecting the catalytic C-terminal domain. Dnmt3b is one of the *de novo* methyltransferases and this mutation results in a loss of function. ICF patients carry mutations in both Dnmt3b alleles resulting in hypomethylation of pericentromeric heterochromatin and genomic instability (Hansen et al, 1999; Xu et al, 1999). Interestingly, ICF is the only known epigenetic syndrome that is based on a genetic mutation in one of the DNA methyltransferase genes.

Methyl cytosine binding protein 2 (MeCP2) is a protein that selectively binds methylation marks and induces chromatin compacting by recruiting remodeling factors. Mutations in the X-chromosome located *mecp2* gene results in the neurodevelopmental disorder RETT and leads to a defective readout of DNA methylation and chromosome instability (Nan & Bird, 2001). Highest expression levels of MeCP2 have been found in brain tissues, which might explain the neuronal symptoms. Another X-chromosome linked epigenetic disorder is the Fragile-X syndrome. It is caused by an expansion of a trinucleotide repeat in the regulatory region of the FMR1 (fragile X mental retardation-1) gene (Basehore & Friez, 2009). A trinucleotide expansion of over 200 repeats is subsequently methylated, which then results in silencing of the FMR1 gene. This gene is required for neural development and its silencing leads to mental retardation in patients (Basehore & Friez, 2009; Stöger et al, 1997). The imprinting diseases Prader-Willi (PWS) and Angelmann (AS) syndrome are special cases of epigenetic disorders (Cassidy & Schwartz, 1998). Many imprinted genes are organized in clusters. One of those clusters is located on chromosome 15 and a uniparental disomy 15, or a deletion in this region results in either PWS or AS. The clinical

14

symptoms differ drastically and depend on whether and where the deletion occurs in the maternally (AS) or paternally (PWS) inherited chromosome.

1.2 DNA methyltransferases

DNA methylation is established and maintained by a highly conserved family of DNA methyltransferases (Dnmts). So far, five members have been described (Dnmt1, 2, 3a, 3b and 3L) yet only three of them have been shown to be active DNA methyltransferases (Dnmt1, Dnmt3a and Dnmt3b).
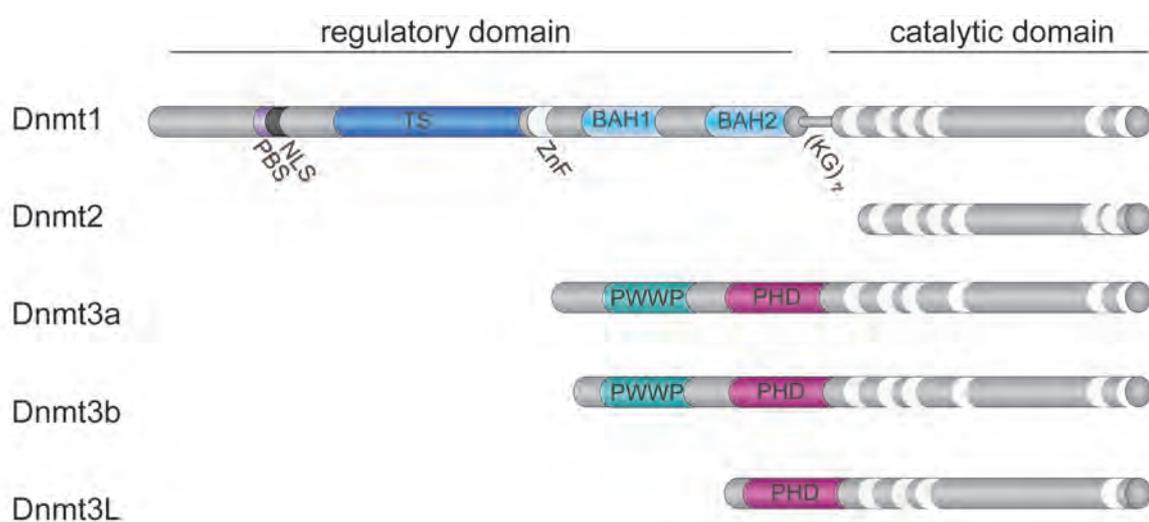
Figure 1-1 Schematic representation of the mammalian DNA methyltransferase family. All Dnmts exhibit a similar catalytic domain that features highly conserved catalytic motifs (I-X) The Dnmts differ in their regulatory region. Dnmt1 contains the PCNA binding domain (PBD), the pericentric heterochromatin targeting sequence (TS), a CXXC type zinc finger motif (ZnF), and two (BAH). The regulatory domains of Dnmt3a and 3b comprise a PWWP domain and a ATRX-like PHD domain. Dnmt2 has no regulatory domain and consists sonly of a catalytic part. The Dnmt3 like protein Dnmt3L has a small regulatory part including a PHD domain but lacks the essential motifs for catalytic activity.

All active Dnmts consist of a unique variable N-terminal regulatory domain and a C-terminal catalytic part that is highly conserved, even throughout different species (Goll & Bestor, 2005). Active Dnmts catalyze the transfer of a methylgroup from S-adenosyl L-methionine (SAM) to the 5`-position of a cytosine base. Dnmt1, Dnmt3a and Dnmt3b are the best-characterized members of the DNA methyltransferase family comprising different functions and target sites. Dnmt3a and 3b establish methylation marks *de novo* during

development whereas Dnmt1 maintains DNA methylation during DNA replication. Even though Dnmt2 contains all conserved catalytic motifs, it shows only residual DNA methyltransferase activity (Hermann et al, 2003). Instead, Dnmt2 has been shown to methylate cytoplasmic tRNA (Goll et al, 2006). The Dnmt3-like protein Dnmt3L is a cofactor of the *de novo* methyltransferases. Interaction stimulates the catalytic activity of Dnmt3a and 3b and the interaction is essential for methylation of retrotransposable elements and maternal imprints (Bourc'his et al, 2001; Hata et al, 2002; Hu et al, 2008). Dnmt3L is the only member that lacks crucial catalytic motifs and is catalytically inactive (Aapola et al, 2001).

1.2.1 Establishment of new methylation patterns - Dnmt3a and Dnm3b

The establishment of new methylation marks predominately happens in a highly controlled way during embryogenesis and development of germ cells, and is catalyzed by members of the Dnmt3 family. Dnmt3a, Dnmt3b and their isoforms were first described in 1999 and are the only known active *de novo* methyltransferases (Okano et al, 1999). They are highly expressed in embryonic stem cells, down-regulated during differentiation and only little expression can be detected in somatic lineages, thus underlining a predominant role of *de novo* methylation during embryogenesis (Okano et al, 1999; Okano et al, 1998). Inactivation of both *de novo* methyltransferase genes results in early embryonic lethality, although embryonic stem (ES) cells (ESCs) show no measurable proliferation defect and undifferentiated cells grow normally under tissue culture conditions. Interestingly, knockout mice lacking either Dnmt3a or Dnmt3b die at different stages during embryonic development. *Dnmt3a* null mice showed a significant growth defect and died roughly 4 weeks after birth. In contrast, *dnmt3b* null mice were not viable and several developmental defects were observed in these embryos. This significant difference suggests a major role of Dnmt3b during early development while Dnmt3a has a predominant role later in development (Okano et al, 1999). Methylation of repetitive sequences, such as major and minor satellite sequences, and intracisternal A-particle (IAP) elements were analyzed and compared in single and double null (*dnmt3a*$^{-/-}$, *dnmt3b*$^{-/-}$; DKO) ES cells. Interestingly, only cells lacking Dnmt3b or both Dnmt3's show an extensive lack of methylation at minor satellite repeats, suggesting that minor satellites are exclusively methylated by Dnmt3b.

Like all active eukaryotic DNA methyltransferases, both *de novo* methyltransferases are comprised of a large regulatory N-terminal domain and a highly conserved C-terminal catalytic domain. The regulatory part is essential for protein-protein interactions, the correct targeting to specific sites and harbors several conserved domains like the PWWP and ATRX homology domain. Although Dnmt3a and Dnmt3b are highly similar in their protein structure they have distinct functions and different target sites.

The ATRX-like PHD domain of Dnmt3a and Dnmt3b is similar to a domain of the chromatin associated X-linked ATRX protein, suggesting a role in protein-protein or protein-DNA interaction (Argentaro et al, 2007). The PHD domain of Dnmt3a interacts with histone deacethylase 1 (HDAC1) and is involved in transcriptional repression of tissue specific gene (Aoki et al, 1998; Fuks et al, 2001; Bachman et al, 2001).

While structure and function of ATRX domain is similar in Dnmt3a and Dnmt3b, the PWWP domain has distinct functions in both enzymes and is thought to determine the functional differences of both enzymes. Dnmt3a and Dnmt3b are enriched at pericentric heterochromatin (PH) in ES cells. Mutations in the PWWP domain disrupt the association and correct targeting to PH and abolishes the capability of Dnmt3a and Dnmt3b to methylate satellite repeats (Chen et al, 2004). However, localization studies with GFP fusion proteins in mouse embryonic fibroblasts (NIH cells) suggest that localization of Dnmt3a and Dnmt3b is similar but not strictly identical. In those cells, Dnm3a was stronger and more obvious enriched at PH than Dnmt3b, which might be due to the preference of Dnmt3a to methylate major satellites while Dnmt3b preferentially methylates minor satellites (Chen et al, 2004). Mutations in the PWWP domain of Dnmt3b result, among others, in a lack of methylation of minor satellites (Xu et al, 1999). This might explain the chromosomal instability that accompanies the ICF syndrome. Taken together, Dnmt3a and Dnm3b are both required for *de novo* methylation but have distinct functions and target sites.

1.2.2   Maintenance of DNA methylation patterns - Dnmt1

Once established, DNA methylation marks have to be maintained after and during each replication cycle. Among the active methyltransferases, Dnmt1 was described first and later shown to be essential for maintaining DNA methylation during DNA replication in

proliferating cells (Bestor & Ingram, 1983; Leonhardt et al, 1992). Null mutation of the Dnmt1 gene results in drastically reduced, but stable levels of methylation. These *dnmt1*[-/-] ES cells carry only about 20% of the normal methylation levels, are impaired in their differentiation ability and the homozygous embryos cannot develop further than the 8-somite stage (Li et al, 1992; Lei et al, 1996). However, *dnmt1*[-/-] ES cells grow surprisingly normal, while a knockout in differentiated somatic cells leads to a p53 induced apoptosis (Robertson et al, 2000).

Among the methyltransferase family, Dnmt1 has the largest N-terminal regulatory region, which comprises several unique domains that define function and sequence specificity of the respective protein. In contrast to the *de novo* methyltransferases, Dnmt1 is regulated in a cell cycle dependent manner and protein expression levels peak during S-Phase (Robertson et al, 2000). The association of Dnmt1 with the replication machinery mediated by interaction of the PCNA binding domain (PBD) of Dnmt1 and the proliferation cell nuclear antigen (PCNA) is proposed to be an efficient mechanism to couple replication and maintenance of methylation (Chuang et al). Moreover, Dnmt1 has been found at sites of DNA damage and it has been described that PCNA recruits Dnmt1 via its PBD domain to those sites (Mortusewicz et al, 2005). In this study we further investigated the dynamics and role of the interaction of Dnmt1 with the replication machinery (Schermelleh et al, 2007), as it will be discusses later on (see chapter 3.2)

In addition to the association with the replication machinery, Dnmt1 localizes at heterochromatic regions especially in late S-Phase until early G1. This association with constitutive heterochromatin is mediated by the highly conserved and unique TS (targeting sequence) domain of Dnmt1 (Leonhardt et al, 1992). This targeting is independent of the replication machinery, histone modifications like H3K9 trimethylation and interacting proteins such as the histone modifying enzyme SUV39H1 and histone binding protein HP1 (Easwaran et al., 2004). More importantly, it has been shown that the TS domain directly interacts with the nuclear protein Np95 (Uhrf-1) and thereby recruits Dnmt1 to hemimethylated DNA sites (Sharif et al, 2007; Bostick et al, 2007; Achour et al, 2008). Recent structural studies indicate that the TS domain may have an autoinhibitory effect on the catalytic activity of Dnmt1 (Syeda et al, 2011). The Np95 protein could act in this

context as an activator of Dnmt1, which will be discussed later in more detail (see chapter 3.2).

Apart from PBD and TS, the regulatory N-terminal part comprises a CXXC Zinc finger, two Bromo-Adjacent Homology (BAH) domains and several nuclear localization signals (NLS). The zinc finger is composed of eight conserved cysteine residues that can bind two zinc ions. This motif is thought to be essential for allosteric activation, catalytic activity of the enzyme, and was also shown to interact with the catalytic domain (Fatemi et al, 2001; Margot et al, 2000). DNA binding properties of the CXXC zinc finger of Dnmt1 have been discussed controversially. While the mouse Dnmt1 motif was described to bind preferentially to hemimethylated sites, the human DNMT1 zinc finger has been shown to bind to unmethlytated CpG sites (Pradhan et al, 2008; Fatemi et al, 2001). Given that these studies used not only different species but also different start and endpoints of the constructs, binding properties of this motif still remains to be extensively studied. This specific motif and its function was subject of this work and our results will be discussed in light of recent publications later on.

The regulatory part of Dnmt1 is linked via a flexible glycin-lysin linker to the highly conserved catalytic domain. Even though the catalytic domain contains all motives essential for catalytic activity described for bacterial methyltransferases, expression of the isolated catalytic domain of Dnmt1 did not show any catalytic activity (Margot et al, 2000). In fact, most of the N-terminal region including the TS domain is essential for protein activity and deletion of larger parts results in complete abolishment of enzymatic activity *in vivo*. Structure and function of Dnmt1 is extremely sensitive to deletions by disturbing the complex folding of the protein, arguing for a highly complex intra- and Intermolecular interplay between several domains and other proteins (Fellinger et al, 2009; Zimmermann et al, 1997).

1.2.3 DNA methyltransferase Crosstalk

A well established model divides the family of methyltransferases based on their functions. This model also supports the hypothesis that the methyltransferases have unique, non-overlapping functions; either in maintenance of DNA methylation during replication (Dnmt1) or in establishing new methylation marks (Dnmt3a and Dnmt3b). However, a closer look at

available data suggests that crosstalk between the members of the family is fundamentally important and Dnmt1, Dnmt3 and Dnmt3b have overlapping roles in maintenance and *de novo* methylation.
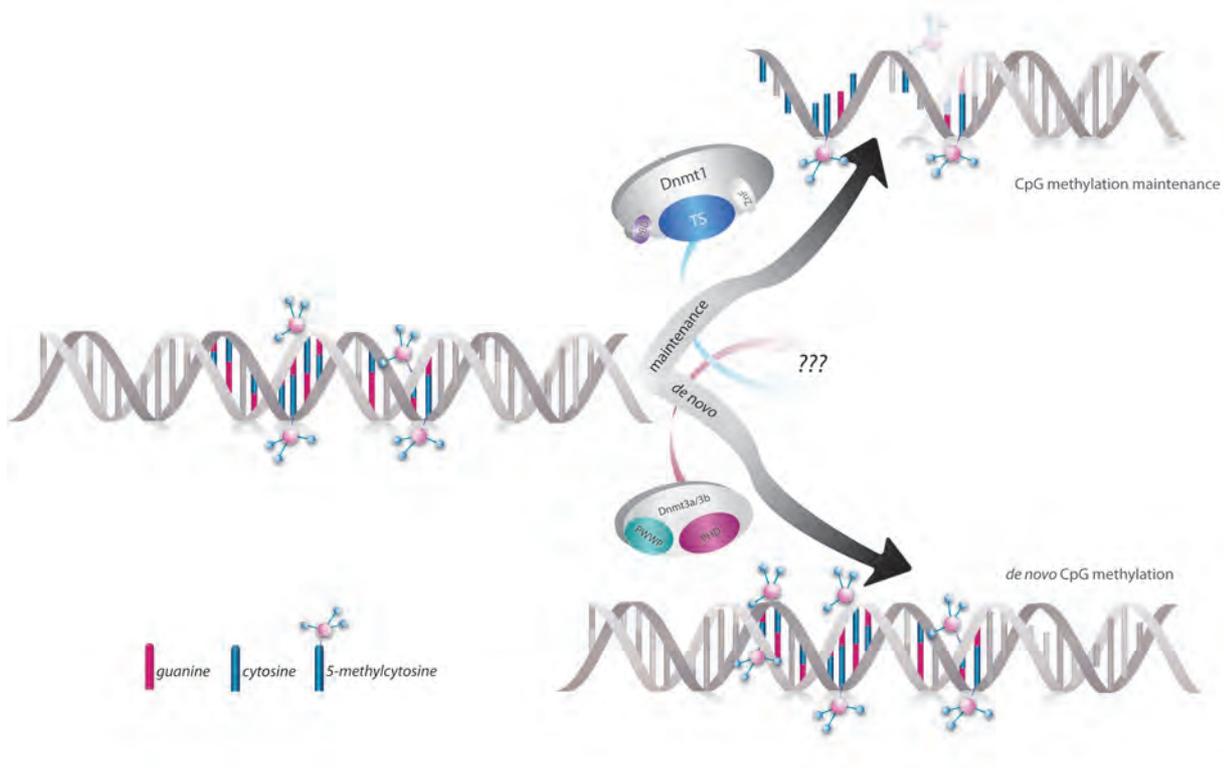

Figure 1-2 Maintenance and *de novo* methylation. Dnmt1 is the major enzyme that is maintains DNA methylation during replication. Dnmt3a and Dnmt3b are *de novo* methyltransferases that establish new methylation marks predominately during embryogenesis and differentiation. However, Dnmt3a and Dnmt3b have been shown to be involved in maintenance of special sequences and Dnmt1 is discussed to have de novo methylation activity, yet to which extent has still to be investigated.

As mentioned above, Dnmt1 has a strong preference for hemi-methylated DNA *in vitro*, underlining its predominant role in maintaining DNA methylation during replication (Hermann et al, 2004b; Pradhan et al, 1999; Fatemi et al, 2001). However, it remains to be discussed whether Dnmt1 might also play a role in *de novo* methylation of certain sites within the genome. In contrast to all previous studies and clear classification of Dnmt's in maintenance and *de novo* methyltransferases, studies indicated an additional role of Dnmt1 in *de novo* DNA methylation processes. By using a recombinant human DNMT1 protein in an *in vitro* methylation assay, binding preference for hemi-methylated over unmethylated DNA were compared (Pradhan et al, 1999). Although DNMT1 showed a strong preference for hemimethylated DNA, this study suggests an additional *de novo* methylation activity on unmethylated DNA substrate. In addition, hypermethylation of

imprinted promoter regions have been found in tumor cells overexpressing Dnmt1, which also could hint to an involvement of Dnmt1 in *de novo* methylation (Biniszkiewicz et al, 2002). However, clear data supporting a *de novo* activity of Dnmt1 in living cells is still missing.

The division of *de novo* and maintenance methyltransferases is not strict and can be influenced in both directions. Dnmt1 and the *de novo* methyltransferase Dnmt3a cooperate during *de novo* methylation, which was shown by a fivefold stimulation of methylation activity if both enzymes are present *in vitro* (Fatemi et al, 2002). In addition, Dnmt1 and Dnmt3a efficiently methylated nucleosomal DNA *in vitro* and their N-terminal domains are essential for the binding ability to nucleosoms. Nucelosomes, compared to naked DNA reduce the accessibility of DNA binding proteins and enzymes to their substrate but reflects the native target form in living cells. Isolated catalytic domains of Dnmt1 and Dnmt3b, as well as the bacterial SssI methyltransferase are not able to bind DNA that is organized in higher order chromatin structures including nucleosomes. Only full-length Dnmt enzymes are able to bind nucleosomal DNA, indicating the significance of intra- and intermolecular interactions in regulating the methylatransferase activity including targeting, unwrapping the chromatin substrate and finishing the process by its enzymatic activity (Gowher et al, 2005; Okuwaki & Verreault, 2004)).

Interestingly, Dnmt3a and Dnmt3b are also involved in maintenance of methylation patterns of specific sequences as well as global methylation patterns (Chen et al, 2003). Embryonic stem cells that lack both enzymes (*dnmt3a$^{-/-}$, dnmt3b$^{-/-}$* ; DKO) clearly show progressive loss of several analyzed sequences, including imprinted genes, repetitive elements, and non-imprinted genes (Liang et al, 2002). Moreover, 30% of the CpG sites in DKO cells were found to be hemimethylated. Given the fact that Dnmt1 is still present in those cells, this result underlines that Dnmt1 alone is not sufficient to fully maintain the methylation at those sites.

In addition to the active methyltransferases, a Dnmt-like protein (Dnmt3L) is enzymatically inactive but essential for activity of the *de novo* methyltransferases (Aapola et al, 2004; Hata et al, 2002). Dnmt3L is highly expressed in ES and germ cells. In germ cells, Dnmt3L interacts with Dnmt3a leading to a methylation dependent silencing of retrotransposons

(Bourc'his et al, 2001). Furthermore, in ES cells maternal methylation-specific imprinting is dependent on the involvement of Dnmt3L (Bourc'his et al, 2001; Hata et al, 2002).

Knockout ES cells that lack all three active methyltransferases (TKO) have no residual methylation, but grow surprisingly normal (Tsumura et al, 2006). Due to the complete abolishment of methylation as well as the lack of active methyltransferases, TKO cells are an attractive tool to study interaction, interplay and distinct functions of Dnmt1, Dnmt3a and Dnmt3b, which was a main subject of this work and will be discussed in detail later on (see chapter 3.3).

1.3   Epigenetic Crosstalk

The epigenetic code is composed of several factors including DNA methylation, chromatin modifications, non-coding RNAs and polycomb repressive complexes and influence cooperatively chromatin structures. In addition, a complex network is involved to determine and preserve a certain epigenetic state. Interactions of methyltransferases with chromatin modifying enzymes as well as methylcytosine binding proteins and other regulating proteins are essential key factors in this network. Together, DNA methylation and histone modifications regulate gene expression and chromatin structure in a cooperative way.

Several groups have shown that Dnmt1 interacts with Np95, also known as Uhrf1 and that this interaction is essential for maintaining DNA methylation (Bostick et al, 2007; Sharif et al, 2007). Mouse ES cells with a genetic ablation of Np95 have a surprisingly similar loss of global as well as local, gene specific methylation comparable to *dnmt1* null ES cells and both knockout mice show a similar developmental arrest (Sharif et al, 2007). These similarities underline the dependence of faithful DNA methylation maintenance on the interaction and presence of both enzymes. Np95 is a multi-domain protein that is able to bind hemi-methylated CpG sites via its SET and RING-finger associated (SRA) domain. It also mediates loading of Dnmt1 to heterochromatin replication sites. Correct localization and targeting of Np95 is dependent on methylated sites and the presence of Dnmt1. Immunofluorescence studies compared localization of both proteins and showed that they co-localize throughout S-Phase. Localization of Np95 to replication foci was completely

abolished in unmethylated TKO ES cells, while a higher enrichment of Np95 to replication foci was observed in *dnmt1* null ES cells, which have a higher amount of hemimethylated sites (Sharif et al, 2007). Taken together, these findings underline the crucial role of Np95 in maintenance of DNA methylation and a key role in targeting Dnmt1 to hemimethylated sites during S-Phase. Moreover, Np95 plays a additional role in linking DNA methylation with histone modification and binds specifically to histone 3 lysine 9 trimethylated tails (H3K9me3) via a tandem tudor domain (Rottach et al., 2009a). Recently, a new member of the Np95 family, Np97 (Uhrf 2), has been found that harbors similar domains like Np95 (Bronner et al., 2007). However, little is know about the functional relevance of Np97 apart from its involvement in degradation of polyglutamine aggregates (Iwata et al., 2009).

Histone 3 and 4 (H3K4) tails are highly modified and modifications at those sites contribute and determine chromatin structure and compaction. In particular, histone 3 lysine 9 (H3K9) methylation has been shown to correlate with regions of silenced genes within eu- and heterochromatic regions, while H3K4 methylation is often linked to active genes in euchromatin regions (Cowell et al, 2002; Litt et al, 2001). It was shown that the histone-modifying enzyme G9a is specific for mono- and di- methylation, while the Suv39 class is responsible for trimethylation of H3K9 (Rea et al, 2000; Tachibana et al, 2008). Consequently, genetic loss of G9a decreases H3K9 methylation. Interestingly, Dnmt1 and G9a directly interact with each other and they have a surprisingly similar cellular distribution during replication and show similar embryonic lethality (Dong et al, 2008). In addition, knockdown of Dnmt1 results in an impaired G9a loading onto chromatin. Those findings indicate a coordinate link of histone methylation and the maintenance of DNA methylation.

G9a not only interacts with Dnmt1 and but also with Np95 and the lymphoid specific helicase (Lsh) that is involved in chromatin remodeling and associated with pericentric heterochromatin (Jarvis et al, 1996; Myant et al, 2011). Moreover, Lsh has been found to interact with Dnmt's and histone deacetylases (HDAC1 and HDAC2) suggesting a crucial role in linking histone remodeling complexes and DNA methylation for a cooperative transcriptional repression (Myant & Stancheva, 2008; Zhu et al, 2006). Knockout of LSH was shown to be lethal and crucial for embryonic development. Global DNA methylation, as well as methylation of repetitive and gene specific sequences is drastically reduced in

lsh-/- ES cells (Myant & Stancheva, 2008). These findings suggest an important role of Lsh in transcriptional repression by DNA methylation as well as interaction with histone deacetylases and thereby contributing to heterochromatic formation.

Apart from histone deacethylases, histone de*methylases* are also linked to stable DNA methylation. A lysine-specific demethylase 1 (LSD1) has a dual role and specifically demethylates H3K4 and H3K9, but more astonishing also demethylates Dnmt1 and thereby stabilizes the protein. *Lsd1* knock out ES cells have reduced levels of Dnmt1 and subsequently also reduced global DNA methylation levels (Wang et al, 2009). These results propose a link of DNA and histone methylation by a LSD1 that on one hand removes histone methylation marks, while it stabilizes Dnmt1 on the other hand and thereby assures faithful DNA methylation maintenance.

Methylcytosine binding proteins (MBDs) are also key players in the complicated epigenetic network that links DNA methylation and histone modifications. The mammalian family of MBD's consists of five members (MeCP2 and MBD1-4) whereas MeCP2, MBD1, MBD2 and MBD4 specifically bind methylcytosine. In particular, a heterodimer of MBD2 and MBD3 recognizes and binds to hemimethylated CpG sites and directly interacts with Dnmt1 at late replication foci (Tatematsu et al, 2000). MeCP2 was the first described family member, binds to methylcytosine and subsequently recruits HDACs for further gene silencing (Nan et al, 1996).
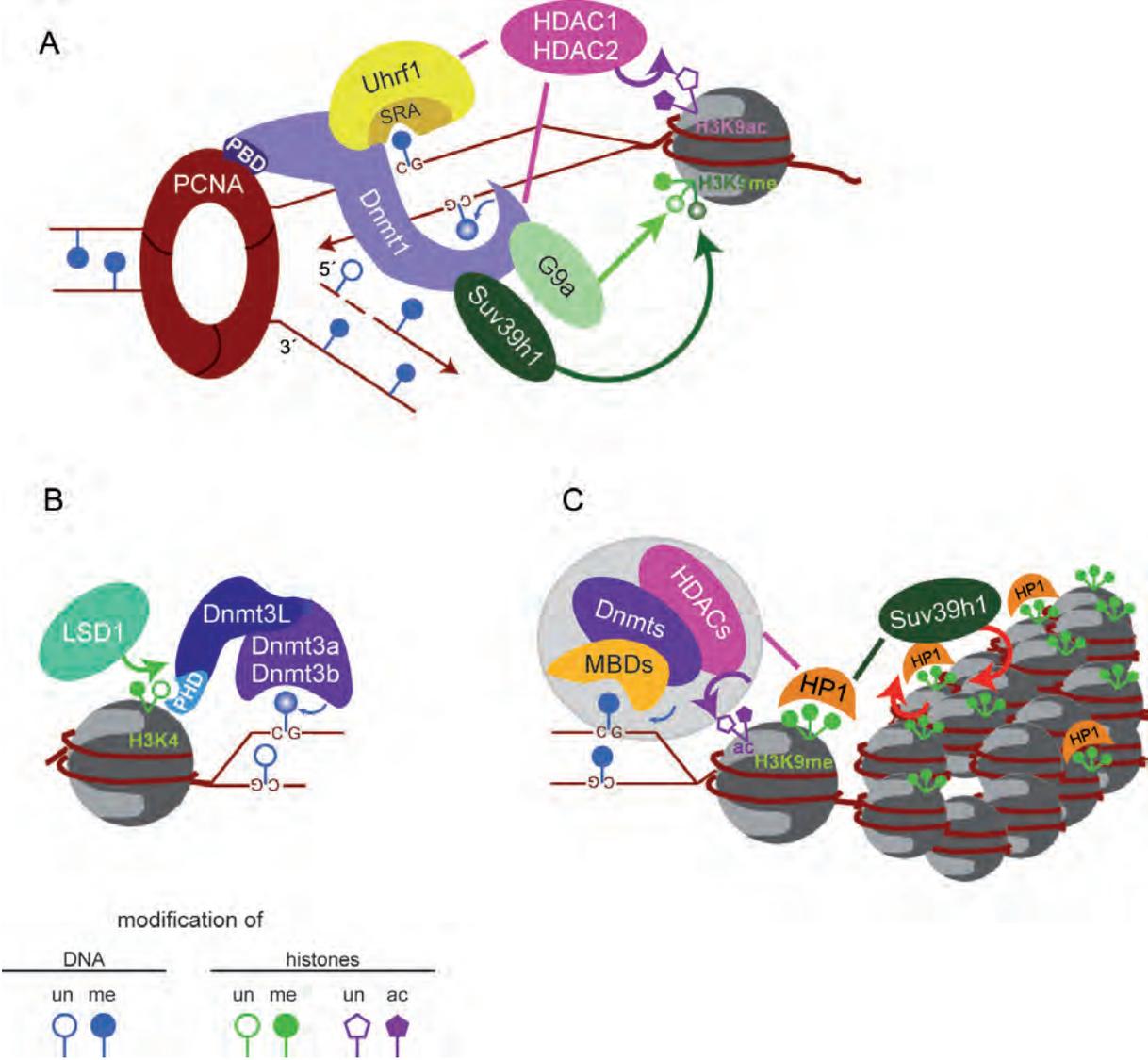
Figure 1-3 Epigenetic crosstalk between DNA methylation and histone modifications. (A) Dnmt1 binds to PCNA via the PBD domain to the replication fork. Np95 (Uhrf1) binds to hemimethylated DNA, recruits Dnmt1 and thereby allows DNA methylation maintenance. Histone modifying enzymes such as Suv39, G9a and HDAC1/2 interact with Dnmt1 and Np95 and enforces further gene silencing by removing acetyl groups from histone tails (HADAC1/2) and adding methyl groups to H3K9 (G9a, Suv39). (B) De novo methyltransferases Dnmt3a and Dnmt3b interact with the Dnmt3-like protein Dnmt3L, which serves as an activating and regulating factor. LSD1 removes methylation from H3K9 where Dnmt3L can bind. (C) Long term gene silencing is mediated by HP1, which recruits Suv39 to generate more binding sites for HP1. (Rottach et.al. 2009)

In summary, Dnmt1 and Np95 are loaded onto the replication machinery and interaction of Dnmt's and Np95 with chromatin modeling factors such as G9a, Suv39h, Lsh and LSD1 enforces further gene silencing (Myant & Stancheva, 2008; Viré et al, 2006). Methylcytosin reading proteins such as the MBD family bind methylated CpG sites and contribute to the compaction of the chromatin.

1.4 Mouse embryonic stem cells

Pluripotency describes the capability of cells to differentiate into all three germ layers: ectoderm, endoderm and mesoderm. Pluripotent cells can only be found in early embryos and exist in a very short time frame during development. Mouse embryonic stem (ES) cells are obtained by isolating cells from the inner cell mass (ICM) of blastocysts and they maintain their pluripotent state and retain an undifferentiated state under special tissue culture conditions throughout many divisions (self-renewal). ES cells can differentiate *in vitro* and *in vivo* into all three germ layers, which makes them a valuable tool to study differentiation and early developmental processes. Moreover, ES cells can genetically be manipulated and knockouts of several important genes of the epigenetic network are viable although somatic cells lacking those proteins undergo apoptosis.

Among several other transcription factors Oct4, Sox2 and Nanog are essential to maintain the pluripotent state and cell identify of the ICM as well as ES cell lineages (Avilion et al, 2003; Babaie et al, 2007; Chambers et al, 2003). They are central players in a transcriptional network that on one hand activates genes necessary for survival while on the other hand they help repressing other genes that are involved in cellular differentiation processes. Upon differentiation, Oct4, Sox2 and Nanog are silenced and down regulated while cell type specific genes become activated (Boyer et al, 2005). Routinely, ES cells are grown on a layer of feeder cells or in the presence of the cytokine Leukemia Inhibitory Factor (LIF) to maintain their pluripotent state by inhibiting differentiation processes. LIF stimulates self-renewal of ESC's through activation of the STAT-signaling pathway and upon removal of LIF spontaneous differentiation is initiated (Hall et al, 2009; Niwa et al, 2009).

To investigate the role of epigenetic factors during differentiation, ES cells can be differentiated *in vitro*. Several methods are used and established to differentiate ES cells. Differentiation can either be induced spontaneously by removal of LIF or by using various molecular factors that force differentiation into certain cellular lineages. A commonly used method is based on the three dimensional aggregation of ES cells into embryoid bodies (EBs) in absence of LIF (Keller, 1995). To generate EBs several methods can be used, either by culturing ES cells on bacterial plates that do not allow cell adhesion or by using

the "hanging drop" method. Under these conditions ES cells are able to aggregate and form a colony of differentiated cells, which resembles early embryonic development. The hanging drop method has been shown to be more efficient in forming EBs and provides a more homogenous differentiation process.
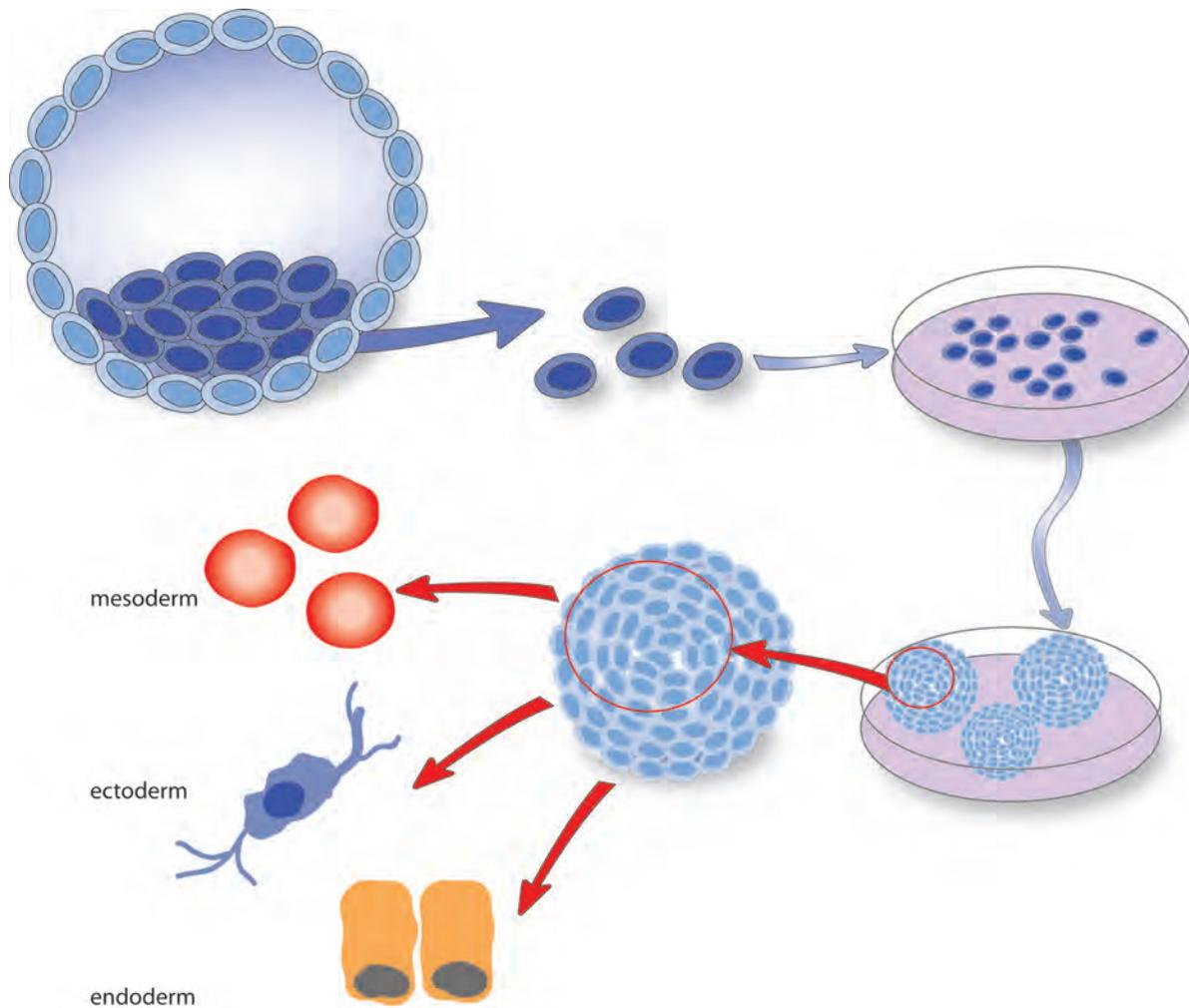

Figure 1-4: Differentiation of ES cells into embryoid bodies (EB). Isolating cells from the ICM of blastocysts generates ES cells. They can be kept in culture over several passages without losing their pluripotent stage and self-renewal ability. Embryoid body formation is based on the three dimensional aggregation of ES cells under LIF removal and resembles early embryonic development. ES cells can differentiate into cell types from all three germ layers: mesoderm, ectoderm and endoderm.

Upon stimulation of ESC's with chemical defined reagents it is possible to direct the differentiation process into specific cell lineages. Among many other factors retinoic acid (RA) is widely used. It regulates growth control and plays a critical role in neuronal differentiation. RA treatment of ESCs induces growth arrest and differentiation into different neuronal cells such as radial glia cells (Bibel et al, 2004).

The capacity of ES cells to differentiate and their pluripotency can also be tested *in vivo* by injecting ES cells into immune-deficient mice (Doetschman et al, 1985). The injected cells proliferate and differentiate into teratomas. Cells of all three germ layers can be found in those tumor cells including also complex organ-like structures (Martin, 1981). Routinely, this is a common mode to test new ES cell lines for their pluripotency and differentiation ability.

As mentioned before, ES cells lacking epigenetic key players such as DNA methyltransferases or histone modifying enzymes are able to tolerate genome-wide hypomethylation without being impaired in their ability to keep their undifferentiated characteristics and proliferation. However, they are often impaired and deficient in their differentiation capability. In this study the role of epigenetic factor during differentiation was addressed by methylation analysis and comparison of different ES lines that lack epigenetic factors (see chapter 2.7).

1.5   Mapping DNA Methylation

DNA methylation is altered in different tissues, various diseases and methylation levels change drastically during differentiation. Therefore, analysis of DNA methylation is a powerful early diagnosis tool in cancer diagnostics. The sensitivity of the selected method is crucial; especially when early methylation changes in cancer cells are investigated. Moreover, analysis and comparison of methylation in different cell types as well as in knockout cell lines can help understanding the complex regulation in the epigenetic network.

1.5.1   Global Methylation

Several methods are established to determine global methylation, including Immunostaining, methylation-specific restriction, and high performance separation techniques such as high-performance capillary electrophoresis (HPCE) and high-performance liquid electrophoresis (HPLC) (Kuo et al, 1980; Fraga et al, 2000). Immunostaining with a 5-mC antibody is commonly used in tissue samples, as well as in cell culture. However, Immunostaining lacks high sensitivity and is not able to quantitatively measure the overall methylcytosine content. To obtain more accurate quantitative data,

methylation-specific restriction is used where genomic DNA is treated and digested with a methylation-specific enzyme.

The most accurate methods to determine global methylation are HPLC and a HPEC based method, which has been shown to be more sensitive. For a HPLC analysis, genomic DNA is treated with ribonuclease enzymes to obtain single nucleosides, which are subsequently quantified. These methods provide highly accurate quantitative and specific data of the overall amount of methylcytosine. However, a large amount of genomic DNA for each sample is required. Recently, genome wide methylation analysis protocols that involve next generation high-throughput deep sequencing become available. The biggest challenge here is to overcome the huge amount of bioinformatics data sets that are generated, which also prevents these methods to be used in routine analysis. However, deep sequencing methods are very useful in clinical samples that involve analysis of global methylation levels as well as genome wide CpG promoter analysis in comparison to the transcriptome.

1.5.2 Gene-Specific Methylation

While global methylation measures the overall 5-methylcytosine content and genome wide analysis is often accompanied with a high amount of data sets, methylation analysis of specific promoter regions is widely used. To study sequence specific methylation levels and patterns, many methods are based on bisulfite treatment of the DNA followed by PCR amplification (Clark et al, 1994; Frommer et al, 1992). During bisulfite conversion cytosine is selectively deaminated to uracil, while 5-methylcytosine remain unchanged. In the subsequent PCR reaction methylated cytosine are amplified as cytosine and unmethylated cytosine results in a thymine, subsequently methylation levels can be determined by different methods. An effective method is the combined bisulfite restriction analysis (COBRA) where a specific sequence is amplified in a PCR reaction after bisulfite treatment. The PCR product is digested with enzymes that distinguish methylated from unmethylated sites. The digested fragments are separated on an agarose gel, thus resulting in a semi-quantitative analysis (Xiong & Laird, 1997). However, COBRA is limited to the restrictions sites created in the analyzed sequences and only few CpG sites within the amplified sequence can be analyzed with this assay.

Classical bisulfite sequencing is often referred to as the gold standard in analyzing sequence specific DNA methylation levels, due to the fact that it provides detailed information of individual molecules at single CpG sites resolution. For this method, the bisulfite PCR is cloned into a bacterial expression vector and single clones are sequenced. This approach allows a detailed analysis of all CpG sites within each analyzed molecule. Despite the fact that this method allows to determine the status of each CpG site within a single molecule, generating quantitative data requires sequencing of a large number of clones. A method based on real-time sequencing technology overcomes these limitations (Ronaghi, 2001).
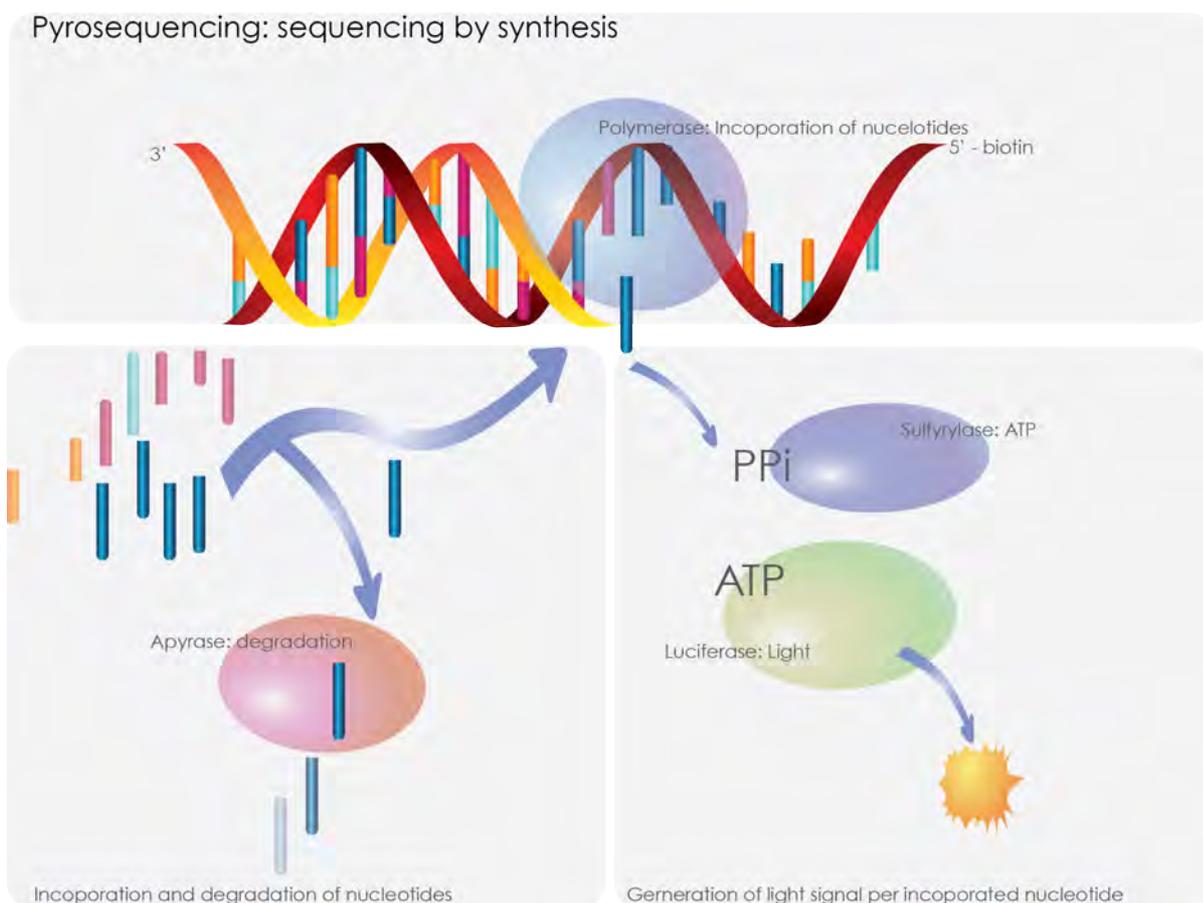
Figure 1-5 Principle of pyrosequencing. A quantitative sequencing by synthesis method in which incorporated nucleotides are monitored bioluminometrically. Each nucleotide that is incorporated by a polymerase releases a pyrophosphate (PPi), which is used by sulfyrylase to generate ATP. Subsequently, ATP is used by luciferase to generate a light signal that is proportional to the amount of incorporated nucleotides. Apyrase degrades not incorporated nucleotides.

Pyrosequencing is a sequencing-by-synthesis method in which the incorporation of nucleotides is monitored bioluminometrically. The intensity of response is directly

proportional to the amount of incorporated nucleotides. After bisulfite conversion of the genomic DNA, a region is amplified using a biotinylated primer and the PCR pool is subsequently pyrosequenced. By measuring the amount of cytosine (unmethylated cytosine) over the amount of thymine (methylated cytosine) at each CpG site a quantitative analysis is obtained (Tost et al, 2003; Tost & Gut, 2007). This study focused on the analysis of selected single copy gene promoters that are normally methylated in ES cells such as the *skeletal α-actin* promoter (Warnecke & Clark, 1999). This gene is expressed and unmethylated only in skeletal muscle cells and therefore serves as a good target to study re-methylation events in hypomethylated knock out ES cells. Repetitive elements, such as major and minor satellite sequences and IAP sequences were also studied to obtain a general overview of the methylation status within the studied genome.

1.6 Aim of this work

The aim of this study was to get further inside into the regulatory mechanism of epigenetic maintenance throughout cell cycle and differentiation, focusing on the maintenance DNA methyltransferase 1 and its interacting partner Np95.

The first part of this study focused on the regulatory function and role of the PBD domain and the CXXC zinc finger motif of Dnmt1 in catalytic activity (see chapters 0 and 2.1). Dnmt1 interacts with the replication machinery by a PBD mediated interaction with the replication processivity factor PCNA and the CXXC zinc finger motif has been shown to be involved allosteric activation of Dnmt1. To address these questions, we performed rescue assays in hypomethylated *dnmt1*[-/-] ESC's with a GFP-Dnmt1 construct that is either deficient in its binding ability to PCNA or a CXXC zinc finger deletion construct. We analyzed the mutant's ability to restore DNA Methylation in promoter regions and repetitive elements.

In the following the hypothesis was addressed whether Dnmt1 has an additional role in *de novo* methylation besides it's predominate role in maintaining DNA methylation during replication (see chapter 2.3 and 2.4) Overlapping roles for Dnmt1 and for the *de novo* methyltransferases Dnmt3a and Dnmt3b has been controversially discussed. To address this question, we used mouse *dnmt1*[-/-]*dnmt3a*[-/-]*dnmt3b*[-/-] (TKO) ESC's that are devoid of

any methylation. These cell lines are deficient in all three active DNA methyltransferases: Dnmt1, Dnmt3a and Dnmt3b which results in a complete lack of methylation. This makes them a powerful tool to study activity and regulation of methyltransferases *in vivo*. Due to the slow kinetics of the methyltransferase reaction, it was imperative to generate stable cell lines expressing the respective DNA methyltransferase. In this study a new FACS based approach in generating stable GFP fusion protein expressing cell lines was established. We analyzed promoter regions and repetitive elements for *de novo* methylation using a pyrosequencing-based assay to obtain quantitative data.

During the generation of stable cell lines, we observed that the commonly used cytomegalovirus (CMV) promoter for ectopic expression constructs was silenced in the respective ESCs. To obtain further inside into the mechanism of CMV promoter silencing in mouse ESCs we established a promoter-silencing assay using fluorescent reporters (see chapter 2.6). We analyzed several ESCs deficient for key players in the epigenetic network to address the question whether DNA methylation or histone modification changes are required in this process. Interestingly, we showed that in the absence of the Dnmt1 interacting partner Np95, mouse ES cells are not able to silence the CMV promoter, although all three DNA methyltransferases are present in those cells. Consequently, we determined the role of Np95 and the DNA methyltransferases in epigenetic silencing of endogenous promoters by differentiating mouse ESCs into embryoid bodies (see chapter 2.7 and 2.8). By analysis of the methylation profile of pluripotency gene promoter regions we addressed the role of DNA methyltransferases as well as Np95 in silencing of pluripotency genes during differentiation.

2. [Results]

Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation

2.1 Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation

Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation

Lothar Schermelleh[1], Andrea Haemmer[1], Fabio Spada[1], Nicole Rösing[1], Daniela Meilinger[1], Ulrich Rothbauer[1], M. Cristina Cardoso[2] and Heinrich Leonhardt[1,*]

[1]Ludwig Maximilians University Munich (LMU), Department of Biology II, 82152 Martinsried, Germany and [2]Max Delbrück Center for Molecular Medicine (MDC), 13125 Berlin, Germany

ABSTRACT

Postreplicative maintenance of genomic methylation patterns was proposed to depend largely on the binding of DNA methyltransferase 1 (Dnmt1) to PCNA, a core component of the replication machinery. We investigated how the slow and discontinuous DNA methylation could be mechanistically linked with fast and processive DNA replication. Using photobleaching and quantitative live cell imaging we show that Dnmt1 binding to PCNA is highly dynamic. Activity measurements of a PCNA-binding-deficient mutant with an enzyme-trapping assay in living cells showed that this interaction accounts for a 2-fold increase in methylation efficiency. Expression of this mutant in mouse *dnmt1*$^{-/-}$ embryonic stem (ES) cells restored CpG island methylation. Thus association of Dnmt1 with the replication machinery enhances methylation efficiency, but is not strictly required for maintaining global methylation. The transient nature of this interaction accommodates the different kinetics of DNA replication and methylation while contributing to faithful propagation of epigenetic information.

INTRODUCTION

Genomic DNA in mammalian cells is commonly methylated at position 5 of cytosine residues in CpG sequences. This epigenetic modification plays an important role in the regulation of gene expression and chromatin structure and is essential for normal development, cell differentiation, X chromosome inactivation and genomic imprinting (1,2). Cell-type-specific methylation patterns are established *de novo* in early developmental stages by the action of Dnmt3a and 3b and are then maintained through subsequent cell generations primarily by the action of

Dnmt1 (3,4). In somatic cells, Dnmt1 is the predominant DNA methyltransferase in terms of abundance, contribution to global methyltransferase activity and to genomic methylation levels (2,5). Reduction of Dnmt1 levels leads to hypomethylation, genomic instability and cancer (6). Aberrant genomic methylation is often associated with human disease and tumorigenesis (7).

Due to its strong preference for hemimethylated substrate DNA *in vitro* (8–10) and its accumulation at replication sites during S-phase (11,12) Dnmt1 is thought to act on hemimethylated CpG sites generated during DNA replication. The association of Dnmt1 with replication factories was proposed as an efficient mechanism for coupling maintenance of genomic methylation patterns to DNA replication (11). Later, direct interaction with the replication processivity factor PCNA was shown and a small PCNA-binding domain (PBD) was mapped to the N-terminal region of Dnmt1 (12,13). This domain was also shown to be necessary for recruitment of Dnmt1 to DNA damage sites, suggesting that it is responsible for coupling DNA repair and restoration of methylation patterns (14). Thus, direct interaction with PCNA would ensure that methylation patterns are faithfully preserved in different situations involving DNA synthesis. However, DNA replication is highly processive taking about 0.035 s per nucleotide (15), while *in vitro* steady-state kinetic analysis of purified recombinant Dnmt1 revealed rather low turn-over rates of about 70–450 s per methyl group transfer (16). Although DNA methylation by Dnmt1 may be faster *in vivo*, it is not likely to come close to the 3–4 orders of magnitude faster DNA replication. In fact, cytosine methylation is a highly complex reaction involving recognition of hemimethylated CpG sites, binding of S-adenosylmethionine, flipping of the cytidine base out of the double helix, formation of a covalent bond between the enzyme and the cytidine, transfer of the methyl group and release of the covalent bond by β-elimination (17–19). These considerations leave open

---

*To whom correspondence should be addressed. Tel: +49-89-2180-74232; Fax: +49-89-2180-74236; Email: h.leonhardt@lmu.de

the question of how DNA replication and methylation are kinetically and mechanistically coordinated.

In addition to Dnmt1 several other factors directly and indirectly involved in DNA replication, such as DNA Ligase I, Fen1, CAF-1 and Cyclin A have been shown to redistribute to replication foci during S-phase (20–24). Many of these factors have been found to interact directly with PCNA, which forms a homotrimeric ring around the DNA helix and serves as a platform for tethering them to the replication machinery (25). Even taking into account the trivalent nature of the PCNA ring, the sheer number of its potential binding partners during replication makes it clear that they cannot all possibly bind at the same time in a constitutive manner.

The functional relevance of the interaction between Dnmt1 and PCNA and its contribution to the maintenance of epigenetic information after DNA replication, however, remains unclear. We have addressed this question by comparing the kinetics and activity of GFP-tagged wild-type Dnmt1 and PCNA-binding-deficient mutants in live cell assays and Dnmt1 deficient embryonic stem (ES) cells. Our data show that the interaction of Dnmt1 with PCNA is highly transient, increases the efficiency of postreplicative methylation by 2-fold, but is not required for restoring CpG methylation in Dnmt1 deficient ES cells.

MATERIALS AND METHODS

Expression constructs

The expression constructs RFP-PCNA, GFP-Ligase, GFP-Dnmt1^{wt}, GFP-Dnmt1$^{\Delta 1-171}$, GFP-Dnmt1^{C1229W} as well as the PBD-GFP construct were described earlier (13,21,26,27). The GFP-Dnmt1^{Q162E} and GFP-Dnmt1^{F169S} expression constructs were derived from the GFP-Dnmt1^{wt} construct by overlap extension PCR mutagenesis (28,29) using the outer forward primer 5′-CAG ATC TCG AGC TCA AGC TTC-3′, the inner reverse primer 5′-GTG TCA AGC TCT GAT AGG ACC AGC-3′, the inner forward primers 5′-GAACCACCAG GGAGACCACCATC-3′ for Q162E and 5′-CACGGCTC ACTCCACGAAGG-3′ for F169S and the outer reverse primer 5′-CTGGAATGACCGAGACGCAGTCG-3′. The final PCR fragments containing the mutations were digested with BglII and HindIII and exchanged with the corresponding fragment in the GFP-DNMT1^{wt} construct. Mutations were confirmed by DNA sequencing and molecular size of fusion proteins was tested by expression in HEK 293T cells and western blot analysis. For stable transfections we inserted the cassette containing the wt Dnmt1 cDNA fused to GFP from the GFP-Dnmt1^{wt} construct into the pCAG-IRESblast vector (30).

Cell culture, transfection and FACS-sorting

Human embryonic kidney (HEK) 293T cells and mouse C2C12 myoblasts were cultured in DMEM supplemented with 10% and 20% fetal calf serum, respectively, and 50 μg/ml gentamycin. HEK 293T cells were transfected with polyethylenimine (Sigma) (31). For live cell observations C2C12 myoblasts were grown to 30-40% confluence on Lab-Tek chamber slides (Nunc) or μ-slides (Ibidi) and transfected with TransFectin transfection reagent (Bio-Rad) according to the manufacturer's instructions. Cells were then incubated overnight before performing live cell analysis. Nuclear localization of GFP-Dnmt1^{wt} was identical to endogenous Dnmt1 as determined by immunostaining with an affinity purified polyclonal antiserum against the N-terminal domain of mouse Dnmt1. GFP-Dnmt1 localization was not affected by additional co-expression of RFP-PCNA (controls not shown). Dnmt1 immunostaining showed that typical expression levels of transfected GFP-Dnmt1 constructs in cells selected for live cell imaging were comparable to those of endogenous Dnmt1 protein (Supplementary Figure 1). For stable expression of GFP-Dnmt1^{wt}, C2C12 cells were grown in a p100 tissue culture dish and transfected as described earlier. Cells were then cultured with 10 μg/ml blasticidin for at least 20 days before homogeneity and levels of expression were determined by fluorescence microscopy and western blotting (Supplementary Figure 2). Mouse wild type and *dnmt1$^{-/-}$* J1 ES cells (s allele) (5) were cultured without feeder cells in gelatinized flasks as described (27). J1 cells were transfected with Transfectin (BioRad) 3–4 h after seeding and GFP-positive cells were sorted with a FACS Vantage SE cell sorter (Becton–Dickinson).

Co-immunoprecipitation

HEK 293T cells were transiently transfected with expression plasmids as described above. After 48 h about 70–90% of the cells expressed the GFP constructs as determined by fluorescence microscopy. Extracts from $\sim 1 \times 10^7$ cells were prepared in 200 μl of lysis buffer (20 mM Tris/HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 2 mM PMSF, 0.5% NP40). After centrifugation supernatants were diluted to 500 μl with lysis buffer without NP40. Extracts were incubated with 1 μg of a GFP-binding protein coupled to sepharose (manuscript in preparation) for 1 h at 4°C with constant mixing. Immunocomplexes were pulled down by centrifugation. The supernatant was removed and 50 μl were collected (referred to as non-bound). The beads were washed twice with 1 ml of dilution buffer containing 300 mM NaCl and resuspended in SDS-PAGE sample buffer. Proteins were eluted by boiling at 95°C and subjected to SDS-PAGE followed by immunoblotting. Antigens were detected with a mouse monoclonal anti-GFP antibody (Roche) and a rat monoclonal anti-PCNA antibody (32).

In vitro methyltransferase assay

Extracts from HEK 293T cells expressing the indicated GFP constructs were prepared and immunoprecipitations were performed as described above. After washing with dilution buffer containing 300 mM NaCl the beads were washed twice with assay buffer (100 mM KCl, 10 mM Tris pH 7.6, 1 mM EDTA, 1 mM DTT) and resuspended in 500 μl of assay buffer. After adding 30 μl of methylation mix {[^{3}H]-SAM (S-adenosyl-methionine); 0.1 μCi (Amersham Biosciences), 1.67 pmol/μl hemimethylated ds 35 bp DNA (50 pmol/μl), 160 ng/μl BSA} incubation was

carried out for 2.5 h at 37°C. The reactions were spotted onto DE81 cellulose paper filters (Whatman) and the filters were washed 3 times with 0.2 M (NH$_4$)HCO$_3$, once with ddH$_2$O and once with 100% ethanol. After drying, radioactivity was measured by liquid scintillation. Samples without enzyme and with 2 µg of purified human recombinant DNMT1 were used as negative and positive controls, respectively.

Combined bisulfite restriction analysis (COBRA)

Genomic DNA was isolated by the phenol–chloroform method (33) and bisulfite treatment was as described (34) except that deamination was carried out for 4 h at 55°C. Primer sets and PCR conditions for CpG islands of *skeletal α-actin*, *H19* (region A) and *dnmt1o* promoters, *Xist* exon 1 and intracisternal type A particle long terminal repeats (IAP LTRs) were as described (35–39). PCR products were digested with the following enzymes: *skeletal α-actin* and *dnmt1o* promoters and IAP LTRs with HpyCH4IV (New England BioLabs); *H19* with Bsh136I and *Xist* with TaqI (both from Fermentas). Digests were separated by agarose electrophoresis except for IAP LTR fragments, which were separated in 10% acrylamide gels. Digestion fragments were quantified from digital images using ImageJ software (http://rsb.info.nih.gov/ij/). The results were corrected for PCR bias, which was calculated as described (40). Briefly, COBRA assays were performed on genomic DNA from untransfected *Dnmt1*$^{-/-}$ J1 cells methylated *in vitro* with recombinant SssI methyltransferase (New England BioLabs) and mixed in different proportions with unmethylated DNA from the same cells (Supplementary Figure 6). Bias curves and corrections were calculated using WinCurveFit (Kevin Raner Software). For each amplified sequence digestion with restriction enzymes whose recognition sequence includes cytosine residues in a non-CpG context was used to control for complete bisulfite conversion except for the *skeletal α-actin* promoter, where bisulfite sequencing revealed about 99% conversion efficiency in all samples (Supplementary Figure 5).

Live cell microscopy, FRAP analysis and trapping assay

Live cell imaging and FRAP experiments were performed on a TCS SP2 AOBS confocal laser scanning microscope (Leica) using a 63 × /1.4 NA Plan-Apochromat oil immersion objective. The microscope was equipped with a heated environmental chamber set to 37°C. Fluorophores were excited with the 488 nm line of an argon laser and a 561 nm solid state diode laser. Confocal image series were typically recorded with a frame size of 256 × 256 pixels, a pixel size of 100 nm, and with 150 ms time intervals. The laser power was typically set to 2–4% transmission with the pinhole opened to 3 Airy units. For FRAP analysis, half of the nucleus was photobleached for 300 ms with all laser lines of the argon laser set to maximum power at 100% transmission. Typically 20 prebleach and 400 postbleach frames were recorded for each series. Quantitative evaluation was performed using ImageJ. The mean fluorescence intensities of the bleached and unbleached region for each time point were background subtracted and normalized to the mean of the last

10 prebleach values (single normalized). These values were divided by the respective total nuclear fluorescence in order to correct for total loss of nuclear fluorescence as well as for the gain of nuclear fluorescence due to import from the cytoplasm over the time course (double normalized). Since the fluorescence in the bleached region differed from cell to cell and typically did not reach background level the values were also normalized to zero. Here, only the distal two-thirds of the unbleached part in the first postbleach frame were considered as reference to take into account fast diffusing molecules invading into the bleached half (triple normalized). For each construct and cell cycle stage 6–10 nuclei were averaged and the mean curve as well as the standard error of the mean (SEM) was calculated. Half times of recovery were calculated from the mean curves.

The trapping assay to measure postreplicative methylation efficiency in living cells was previously described (27). 5-Aza-2′-deoxycytidine (Sigma) was added at a final concentrations of 30 µM and cells were incubated for the indicated periods before performing FRAP experiments. Microscope settings were as described above except that a smaller ROI (3 µm × 3 µm) was selected and the time interval was set to 208 ms. FRAP data were double normalized as described above.

For presentation, we used linear contrast enhancement on entire images. For color figures we chose magenta as false color for red fluorescence to accommodate for colorblindness.

RESULTS

Point mutations within the PBD abolish the interaction of Dnmt1 with PCNA

We recently demonstrated a loss of association with replication foci in early and mid S-phase of a GFP-Dnmt1 fusion construct with a deletion of the first 171 amino acids which includes most of the PBD (GFP-Dnmt1$^{\Delta 1\text{-}171}$) (13). To address the function of the PBD in early and mid S-phase more specifically and to exclude potential misfolding of the protein due to the large deletion we have generated GFP-Dnmt1 constructs bearing single point mutations of highly conserved residues within the PIP (PCNA-interacting peptide)-Box (41) of the PBD (GFP-Dnmt1^{Q162E} and GFP-Dnmt1^{F169S}) (Figure 1). Either of these mutant constructs were expressed in C2C12 mouse myoblasts together with RFP-PCNA, which served as S-phase marker (42). Both Dnmt1 constructs showed a diffuse nuclear distribution in early and mid S-phase cells (Figure 2B and Supplementary Figure 3), in contrast to the wild-type Dnmt1 construct (GFP-Dnmt1^{wt}) which was concentrated at replication foci during early and mid S-phase (Figure 2A). In late S-phase and G2, however, the wild type as well as all the PBD mutant constructs, including GFP-Dnmt1$^{\Delta 1-171}$, were similarly concentrated at chromocenters (Figure 2A and B and Supplementary Figure 1), confirming the additional binding to heterochromatin in late S-phase mediated by the TS domain (13). Co-immunoprecipitation experiments confirmed that the Q162E point mutation
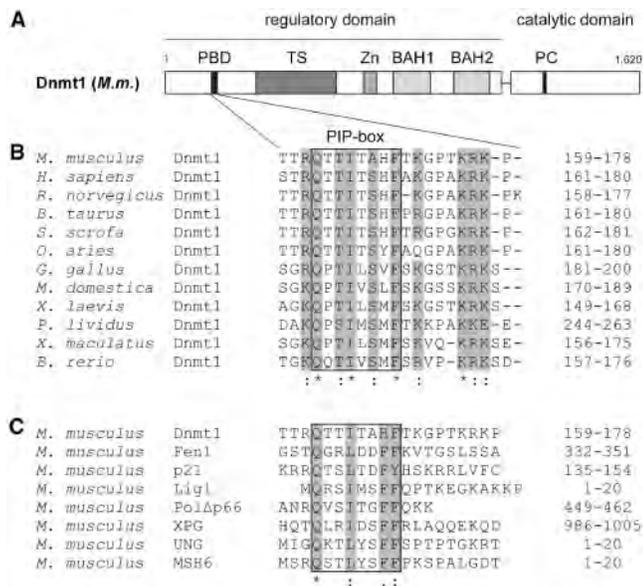
Figure 1. Structure of Dnmt1 and alignment of PCNA-binding domains (PBD). (**A**) Schematic representation of the mouse Dnmt1 consisting of an N-terminal regulatory and a C-terminal catalytic domain. The PBD, the TS domain, the Zn-binding domain (Zn), the two bromo-adjacent homology domains (BAH) and the catalytic center (PC motif) are highlighted. (**B**) Alignment of Dnmt1 PBDs from different species. Highly conserved residues are gray shaded. Accession numbers: *Mus musculus* P13864; *Homo sapiens* P26358; *Rattus norvegicus* Q9Z330; *Bos taurus* Q6Y856; *Sus scrofa* Q4TTV6; *Ovis aries* Q865V5; *Gallus gallus* Q92072; *Monodelphis domestica* Q8MJ28; *Xenopus laevis* Q6GQH0; *Paracentrotus lividus* Q27746; *Xiphophorus maculatus* Q9I8X6; *Brachydanio rerio* Q8QGB8. (**C**) Alignment of PBDs in different proteins from *Mus musculus*. Gray shaded amino acids are highly conserved within the PIP-Box. Accession numbers: Dnmt1 P13864; Fen1 P39749; p21 P39689; DNA Ligase1 P37913; Polymerase δ/subunit Δp66 Q9EQ28; XPG P39749; UNG:P97931 MSH6 P54276.

abolishes the interaction between Dnmt1 and PCNA (Figure 2C). Similar results were obtained with GFP-Dnmt1^{F169S} and GFP-Dnmt1$^{\Delta 1-171}$ (Supplementary Figure 3). Thus, both Q162E and F169S point mutations prevent accumulation of Dnmt1 at replication foci during early to mid S-phase, while localization at constitutive heterochromatin in late S-phase and G2 is not affected. These results clearly confirm the role of the PBD in mediating the interaction between Dnmt1 and PCNA *in vivo*.

The PBD-mediated interaction of Dnmt1 at replication sites is highly transient

To investigate the dynamics of the PBD-mediated interaction at replication sites we measured fluorescence recovery after photobleaching (FRAP) of GFP-Dnmt1^{wt} throughout S-phase. GFP-Dnmt1^{wt} and RFP-PCNA were co-expressed in C2C12 cells and a small square region of interest (ROI) was bleached (Figure 3). Consistent with earlier observations (26,43) RFP-PCNA showed hardly any recovery within the observation period of 100 s, whereas in the same period GFP-Dnmt1 recovered fully, with very similar kinetics in early and mid S-phase, but notably slower in late S-phase. This result shows that the binding of Dnmt1 at replication sites is more dynamic

during early and mid S-phase than in late S-phase, when Dnmt1 is likely slowed down by the additional interaction with chromatin mediated by the TS domain (13).

To address the kinetic properties of the PBD-mediated binding to replication sites more specifically, we performed quantitative FRAP analysis of GFP-Dnmt1^{wt} and GFP-Dnmt1^{Q162E} in G1 and early/mid S-phase. As replication sites are not homogeneously distributed in the nucleus, we chose to bleach half nuclei (half-FRAP) to ensure that the bleached region contains a representative number of potential binding sites (Figure 4A). In early/mid S-phase nuclei GFP-Dnmt1^{wt} recovered with a halftime ($t_{1/2}$) of 4.7 ± 0.2 s and reached complete equilibration (t_{∞}) in about 56 s. These values place Dnmt1 among the more dynamic factors involved in chromatin transactions previously determined with half-FRAP analyses (44). In comparison to GFP-Dnmt1^{wt}, GFP-Dnmt1^{Q162E} showed a slightly increased mobility ($t_{1/2} = 4.4 \pm 0.2$ s; $t_{\infty} \sim 45$ s), which is likely due to the lack of binding to PCNA rings at replication sites. In the absence of active replication sites in G1, GFP-Dnmt1^{wt} and Dnmt1^{Q162E} showed nearly identical kinetics, which were remarkably similar to the kinetics of Dnmt1^{Q162E} in early/mid S-phase. These data indicate that PCNA binding has only a minor contribution to Dnmt1 kinetics in S phase.

The recovery rates measured for the full-length constructs were considerably slower than the rate of GFP alone ($t_{1/2} = 0.8 \pm 0.1$ s; $t_{\infty} \sim 10$ s), which was used to control for unspecific binding events. As $t_{1/2}$ is roughly proportional to the cubic root of the molecular mass (45,46), the ~8-fold size difference of GFP alone to the full-length construct (27 and 210 kDa, respectively) would only account for about a 2-fold slower recovery. Instead, GFP-Dnmt1 full-length constructs recover more slowly, pointing to one (or more) additional and yet uncharacterized cell cycle independent interaction(s).

In order to dissect the PBD-mediated interaction from superimposing effects caused by other potential interactions we assayed the FRAP kinetics of a GFP fusion with the isolated PBD of Dnmt1, i.e. amino acids 159–178 (PBD-GFP). We found that in S–phase the recovery was only about 2 times slower than GFP alone ($t_{1/2} = 1.5 \pm 0.1$ s; $t_{\infty} \sim 16$ s), thus confirming the highly transient nature of the PBD interaction with PCNA. In non-S-phase cells, PBD-GFP showed an increase in mobility ($t_{1/2} = 1.1 \pm 0.1$ s; $t_{\infty} \sim 11$ s) similar to that observed with the full-length wild-type construct.

For direct comparison we further analyzed another PCNA-interacting enzyme, DNA Ligase I fused to GFP (GFP-Ligase), and found a similar mobility shift in S phase ($t_{1/2} = 2.4 \pm 0.3$ s; $t_{\infty} \sim 27$ s) compared to non-S-phase cells ($t_{1/2} = 2.1 \pm 0.3$ s; $t_{\infty} \sim 17$ s). Thus, with all three PCNA-interacting GFP fusion proteins (GFP-Dnmt1^{wt}, PBD-GFP and GFP-Ligase), but not with the PCNA-binding mutant, the transient association with the replication machinery (PCNA) caused a slower recovery as compared to G1/non-S-phase (Figure 4B, inset).

Together the limited but significant contribution of the PBD-mediated interaction to the kinetic properties of GFP-Dnmt1 and the high mobility of PBD-GFP in
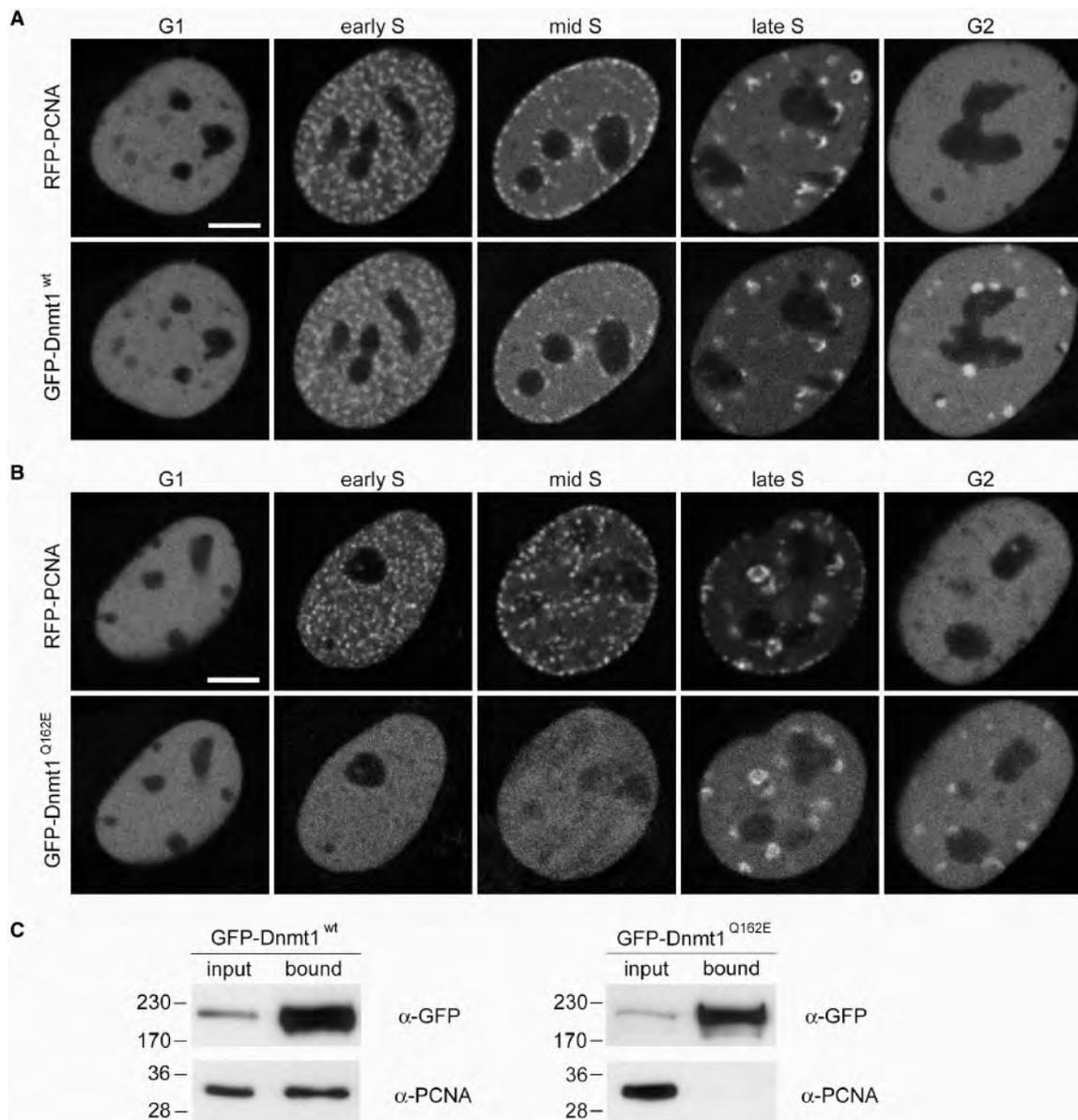
Figure 2. Mutation of the PBD abolishes Dnmt1 interaction with PCNA and prevents accumulation at replication sites in early and mid S-phase. (**A**, **B**) Confocal mid sections of living mouse C2C12 myoblasts expressing either GFP-Dnmt1^{wt} (**A**) or GFP-Dnmt1^{Q162E} (**B**). Cells were co-transfected with RFP-PCNA to identify replication foci and to distinguish S-phase stages. Scale bars: 5 μm. GFP-Dnmt1^{wt} accumulates at replication sites throughout S phase where it co-localizes with RFP-PCNA. In G2 cells a fraction of Dnmt1 remains associated with the late replicating pericentric heterochromatin. In contrast, GFP-Dnmt1^{Q162E} shows a fully dispersed nuclear distribution in early and mid S-phase stages, whereas in late S-phase and G2 association with centromeric heterochromatin is still observed. (**C**) Endogenous PCNA efficiently co-immunoprecipitates with GFP-Dnmt1^{wt} but not with GFP-Dnmt1^{Q162E}. Cell extracts were prepared from HEK 293T cells expressing either GFP-Dnmt1^{wt} or GFP-Dnmt1^{Q162E}. Precipitated proteins were detected by immunostaining with antibodies against GFP and PCNA, respectively.

S–phase clearly indicate that binding of Dnmt1 to PCNA at replication foci is highly transient.

Interaction with PCNA enhances methylation efficiency *in vivo* only moderately

Next we investigated the contribution of the highly transient interaction with PCNA to the postreplicative methylation activity of Dnmt1. Earlier it was shown that N-terminal deletions of mouse Dnmt1 comprising the PBD did not alter catalytic activity *in vitro* (47,48). To test

the catalytic activity of the GFP-Dnmt1 constructs used in this study they were expressed in HEK 293T cells, immunopurified and directly assayed for methyltransferase activity *in vitro*. While the catalytically inactive GFP-Dnmt1^{C1229W} mutant (27) displayed only background activity, GFP-Dnmt1^{Q162E} and GFP-Dnmt1$^{\Delta1-171}$ exhibited enzymatic activity comparable to GFP-Dnmt1^{wt} (Figure 5D), indicating that neither the Q162E point mutation nor deletion of the first 171 amino acids affect the enzymatic activity of Dnmt1.
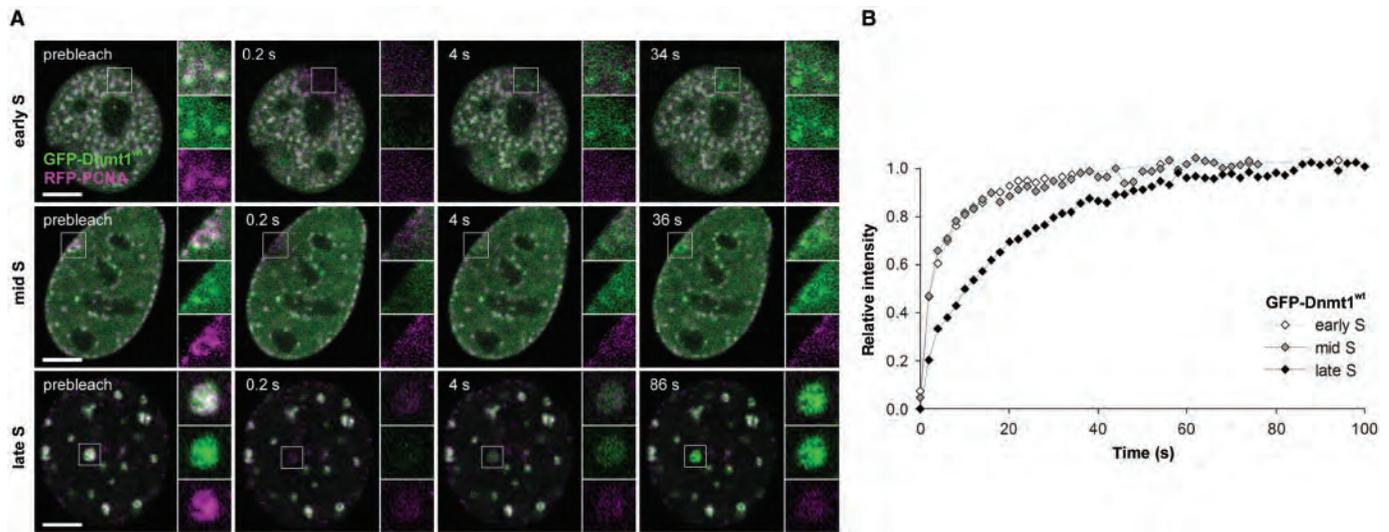
Figure 3. Transient binding of Dnmt1 at replication sites in different S-phase stages. (**A**) GFP-Dnmt1[wt] (green) and RFP-PCNA (magenta) co-expressed in mouse C2C12 cells co-localize at replication sites in early, mid and late S-phase. Fluorescence bleaching of a small region of interest reveals fast recovery of GFP-Dnmt1[wt], whereas RFP-PCNA shows almost no recovery within the observation period. (**B**) Quantitative evaluation of the FRAP experiments shown in (**A**) reveal a significantly faster recovery of GFP-Dnmt1[wt] in early and mid S-phase compared to late S-phase. Scale bars: 5 μm.

To establish whether the binding to PCNA is required for postreplicative methylation *in vivo*, we tested the activity of GFP-Dnmt1[wt] and GFP-Dnmt1[Q162E] in living cells using a recently developed trapping assay (27). This assay takes advantage of the catalytic mechanism of DNA (cytosine-5) methyltransferases which involves transient formation of a covalent complex with the C6 position of the cytosine residue. When the cytosine analogue 5-aza-2′-deoxycytidine (5-aza-dC) is incorporated into the DNA during replication the covalent complex of Dnmt1 and 5-aza-dC cannot be resolved and Dnmt1 is trapped at the site of action. Time-dependent immobilization, i.e. trapping of GFP-tagged Dnmt1 at replication foci can be visualized and measured by FRAP and reflects enzymatic activity of the fusion protein.

C2C12 cells co-transfected with RFP-PCNA and either GFP-Dnmt1[wt] or GFP-Dnmt1[Q162E] as well as C2C12 cells stably expressing GFP-Dnmt1[wt] were incubated in the presence of 30 μM 5-aza-dC. In early S-phase the focal enrichment of GFP-Dnmt1[wt] at replication sites increased over time reflecting the accumulation of immobilized enzyme and after 40 min GFP-Dnmt1[wt] was completely immobilized (Figure 5A). Similar kinetics were observed in mid S-phase (data not shown). As shown above GFP-Dnmt1[Q162E] displayed a diffuse nuclear distribution in early S-phase cells (Figure 5B). However, with prolonged incubation in the presence of 5-aza-dC an increasing focal accumulation at replication sites was observed. Quantitative FRAP analysis revealed that the immobilization rate of GFP-Dnmt1[Q162E], which is a direct measure of its enzymatic activity, was only about 2-fold slower than GFP-Dnmt1[wt], resulting in complete trapping after ~90 min. These results indicate that the PCNA-binding-deficient mutant binds to DNA and is catalytically engaged at hemimethylated sites generated during replication.

Immobilization of GFP-Dnmt1[wt] at replication sites does not prevent progression of the replication machinery

We then probed the stability of the interaction between Dnmt1 and the replication machinery by trapping Dnmt1 with 5-aza-dC and long-term live imaging. In the case of stable interaction, covalent immobilization of Dnmt1 would be expected to stall the progression of the replication machinery. C2C12 cells co-transfected with the GFP-Dnmt1[wt] and RFP-PCNA constructs were incubated in the presence of 10 μM 5-aza-dC and individual S-phase cells were imaged at consecutive time points for an extended period of time (Figure 6). Progressive separation of GFP-Dnmt1[wt] and RFP-PCNA foci could be clearly observed over a time period of ~2 h, indicating that trapping of Dnmt1 did not prevent the progression of replication factories. This result is consistent with the FRAP kinetics of GFP-Dnmt1[wt] demonstrating the transient nature of the interaction between Dnmt1 and PCNA.

GFP-Dnmt1[Q162E] rescues CpG methylation in Dnmt1[−/−] ES cells

To investigate the contribution to maintenance of methylation patterns by the PBD-mediated interaction with PCNA *in vivo* we transiently expressed either GFP-Dnmt1[wt] or GFP-Dnmt1[Q162E] in Dnmt1 deficient mouse ES cells, which are severely hypomethylated in all genomic compartments (5,49). GFP-positive cells were isolated by FACS sorting 24h and 48h after transfection and methylation of single-copy sequences and intracisternal type A particle (IAP) interspersed repetitive elements was analyzed by COBRA. An increase in methylation at all tested sites was observed in cells expressing either the wild type or the mutant Dnmt1 constructs already 24h after transfection (Figure 7 and Supplementary Figure 4A).
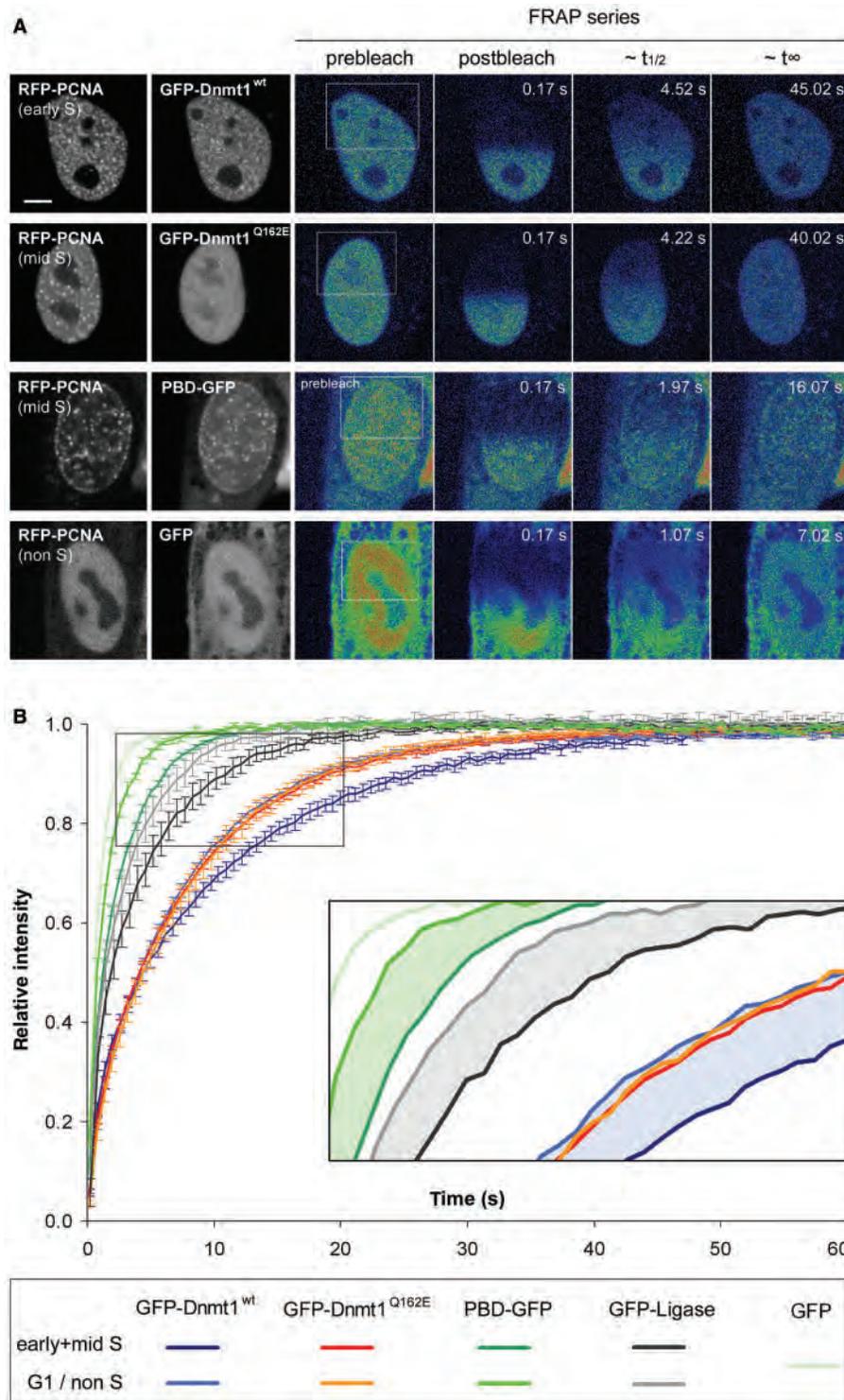
Figure 4. Effect of the PBD on protein mobility in S-phase cells. (**A**) Half nucleus FRAP series of C2C12 cells expressing GFP-Dnmt1[wt] and GFP-Dnmt1[Q162E], PBD-GFP and GFP alone as indicated in the second column. Cell-cycle stages were identified by the subnuclear pattern of co-expressed RFP-PCNA (first column). Selected pre- and postbleach frames are shown in false color. Lower signal-to-noise ratio of FRAP series images is due to the higher imaging frame rate. Rectangles indicate the bleached ROI. The column marked with $\sim t_{1/2}$ displays frames corresponding to the half time of fluorescence recovery. The column marked with $\sim t_{\infty}$ displays the frames corresponding to approximately the time points when fluorescence recovery reached the plateau. Bar: 5 μm. (**B**) Quantitative evaluation of FRAP experiments. Wild-type Dnmt1 shows a small but significant decrease in mobility in early/mid S-phase (dark blue curve) as compared to G1 phase (light-blue curve). A similar mobility shift between early/mid S and non-S-phase cells is also seen for PBD-GFP (dark green and green curve, respectively) and for GFP-Ligase (dark gray and gray curve, respectively). No such shift is observed for the PCNA-binding-deficient Dnmt1 mutant (red and orange curve, respectively). GFP alone (light-green curve) is shown as reference. For clarity only every fourth time point is displayed; data are represented as mean ± SEM. Kinetic shifts are indicated by shadings in the enlarged inset.

Further substantial increase of methylation 48 h after transfection was observed only for the *skeletal α-actin* promoter where the methylation level approached that observed in wild-type cells. The result for the *skeletal α-actin* promoter were confirmed and extended by bisulfite sequencing (Supplementary Figure 5). It was previously reported that re-expression of wild-type Dnmt1 in *dnmt1*[−/−] ES cells does not lead to restoration of methylation at imprinted genes since passage through the germ line is needed for re-establishment of methylation patterns in these sequences (49). We analyzed the promoter of the imprinted *H19* gene to control for the
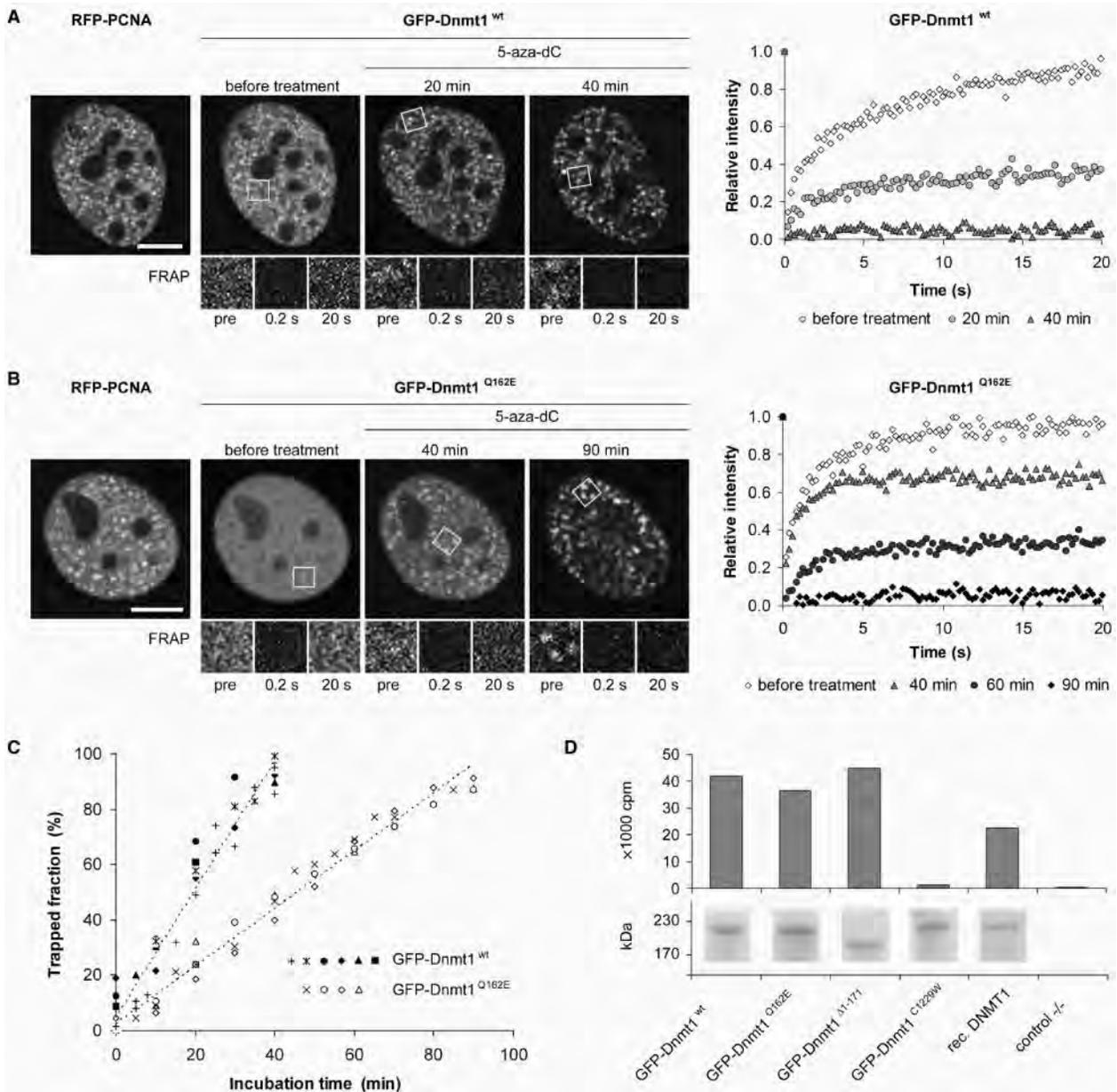

Figure 5. PCNA-binding-deficient mutant Dnmt1 is catalytically active and shows moderately reduced postreplicative methylation activity *in vivo*. (**A–C**) Trapping assay for the analysis of postreplicative methylation activity *in vivo*. (**A, B**) Confocal mid sections (upper panel) and corresponding FRAP series (lower panel) of a C2C12 mouse myoblast cell co-transfected with RFP-PCNA and either GFP-Dnmt1[wt] (**A**) or GFP-Dnmt1[Q162E] (**B**) before and at selected time points after incubation with 5-aza-dC. Before drug treatment the observed cells were in early S-phase as indicated by the RFP-PCNA pattern (left). Boxes indicate bleached ROI's, which are shown magnified in the lower panels at indicated time points of the FRAP series. Quantitative FRAP analysis is shown on the right. Scale bars: 5 μm. (**C**) Time-dependent increase of immobile fractions of GFP-Dnmt1[wt] and GFP-Dnmt1[Q162E]. Individual cells assayed at consecutive time points are indicated by closed (GFP-Dnmt1[wt]) and open (GFP-Dnmt1[Q162E]) symbols. Character symbols represent different cells each assayed at a single time point. + symbol indicates data points obtained with C2C12 cells stably expressing GFP-Dnmt1[wt] (C2C12-GMT1) without co-expression of RFP-PCNA. (**D**) *In vitro* methyltransferase assay for GFP-Dnmt1[wt], GFP-Dnmt1[Δ1-171], GFP-Dnmt1[Q162E] and the catalytically inactive mutant GFP-Dnmt1[C1229W]. GFP fusions were expressed in HEK 293T cells, immunopurified and the amount of [³H]CH₃ transferred to a hemimethylated oligonucleotide substrate was measured. To normalize for the amount of precipitated protein aliquots were analyzed by SDS-PAGE and Coomassie staining (lower panel).

specificity of our assay and found that indeed expression of neither GFP-Dnmt1[wt] nor GFP-Dnmt1[Q162E] resulted in increased methylation of this sequence (Figure 7). To further evaluate methylation of repetitive sequences we
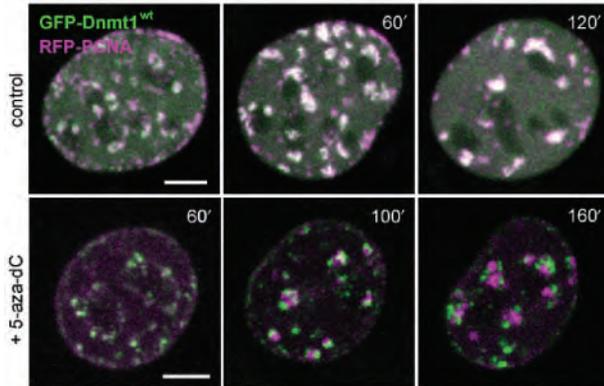

Figure 6. Immobilization of Dnmt1 does not prevent progression of DNA replication. S-phase C2C12 myoblasts expressing GFP-Dnmt1[wt] (green) and RFP-PCNA (magenta) were imaged without drug treatment (upper row) and at the indicated time points after addition of 10 μM 5-aza-dC to the medium (lower row). In the control cell the two constructs largely co-localize during transition from mid to late S-phase, whereas in the presence of 5-aza-dC progressive separation of green and red foci indicate immobilization of GFP-Dnmt1[wt] at postreplicative hemimethylated sites and progression of replication foci containing RFP-PCNA. Projections of confocal mid sections are shown. Scale bar: 5 μm.

stained transiently transfected *dnmt1*[−/−] cells with an antibody against 5-methylcytidine that detects highly methylated satellite repeats present at mouse chromocenters (pericentromeric heterochromatin). Restoration of high DNA methylation levels at chromocenters was observed in both cells expressing GFP-Dnmt1[wt] and GFP-Dnmt1[Q162E] (Supplementary Figure 4B).

The experimental procedures employed did not allow detection of significant differences in remethylation kinetics between the two constructs. Nevertheless, these results show that the PCNA-binding-deficient mutant is, like wild-type Dnmt1, able to rescue methylation of both single copy and repetitive sequences *in vivo*.

DISCUSSION

Faithful replication of genetic and epigenetic information is crucial to ensure the integrity and identity of proliferating cells. Earlier work has demonstrated that the maintenance methyltransferase Dnmt1 binds to the replication processivity factor PCNA and is thus recruited to replication sites (12). The interaction between Dnmt1 and the replication machinery was proposed as a mechanism for coupling maintenance of genomic methylation to DNA replication (11). By traveling along with the replication machinery Dnmt1 would be able to restore symmetrical methylation as soon as hemimethylated CpG sites are generated. The estimated kinetics of DNA replication *in vivo* and DNA methylation by Dnmt1
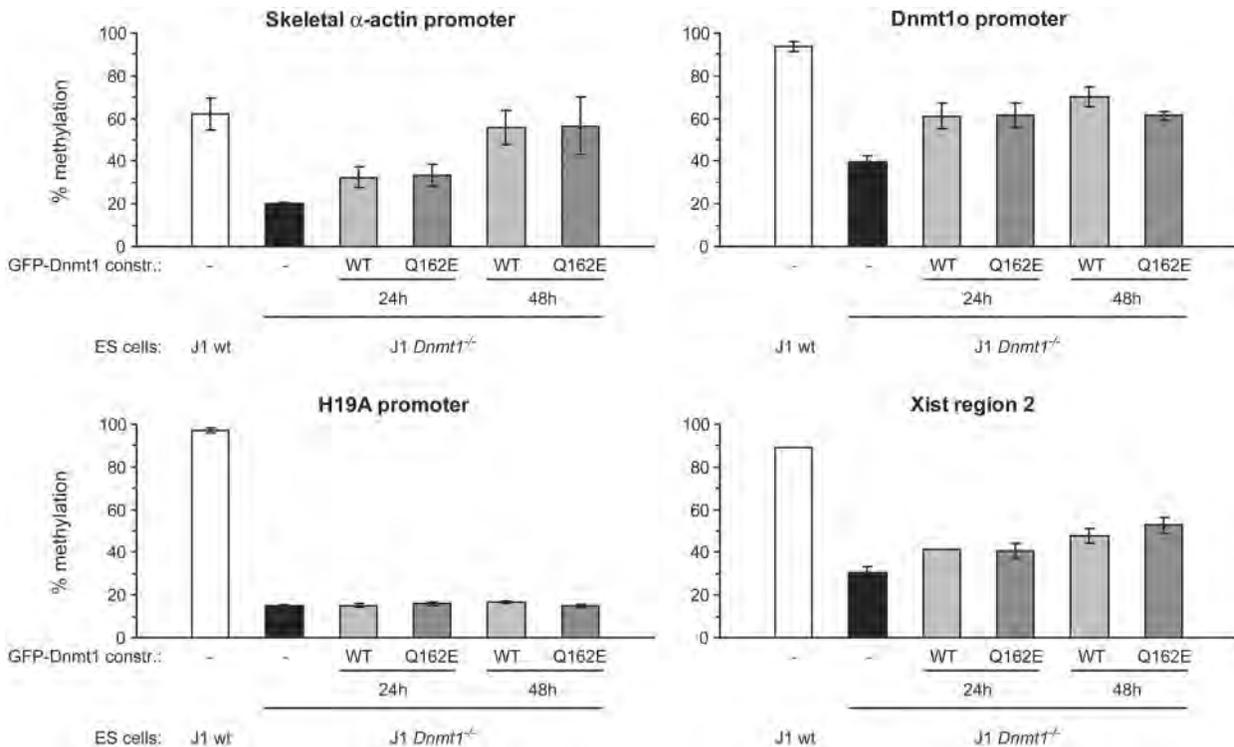

Figure 7. PCNA-binding-deficient mutant Dnmt1 restores methylation of single-copy sequences in *dnmt1*[−/−] ES cells. Mouse *dnmt1*[−/−] ES cells were transiently transfected with either the GFP-Dnmt1[wt] or GFP-Dnmt1[Q162E] constructs and FACS-sorted 24h and 48 h after transfection. Methylation of single-copy sequences was assayed by COBRA. Shown are methylation percentages at assayed CpG sites with respect to genomic DNA from *dnmt1*[−/−] ES cells methylated *in vitro* with the recombinant methyltransferase *Sss*I (100% methylation; see also Supplementary Figure 6). Mean values and standard errors of duplicate amplifications from each of two independent experiments are indicated.

in vitro, however, differ by 3–4 orders of magnitude arguing against a stable coupling. Although DNA methylation may be faster *in vivo* it is not likely to come much closer to the DNA replication rate. In other words, it is reasonable to assume that methylating a cytosine takes far longer than to incorporate it. Also, DNA replication is essentially a continuous process, while methylated CpG sites are encountered discontinuously in vertebrate genomes. Stable binding of Dnmt1 to the replication machinery would stall replication at each hemimethylated CpG site. Here we show that the interaction of Dnmt1 with the replication machinery *in vivo* is highly dynamic and that immobilization of Dnmt1 at postreplicative hemimethylated sites does not prevent the progression of DNA replication. The transient nature of the interaction between Dnmt1 and the replication machinery allows rapid release and transfer of Dnmt1 to hemimethylated substrate sites, reconciling the paradox of the relative rates of DNA replication and methylation. According to basic principles of enzyme kinetics the local enrichment of Dnmt1 at replication foci would increase the postreplicative methylation rate. At the same time, transient binding of Dnmt1 enables also other replication factors to interact with PCNA. Similar binding dynamics have been shown for the interaction of PCNA with DNA Ligase I and Fen1 (26) and may thus represent a common feature of PBD-containing factors.

Interestingly, the interaction between Dnmt1 and PCNA is believed to be a major mechanism for the methylation maintenance activity of Dnmt1, but its functional relevance had never been tested experimentally. Here we show two lines of evidence that this interaction is not crucial for the maintenance of methylation patterns in mammalian cells. First, postreplicative methylation rate of wild-type Dnmt1 measured *in vivo* is only 2-fold faster than that of a PCNA-binding-deficient mutant. Second, methylation of both single copy and repetitive sequences in Dnmt1 deficient ES cells was restored by this PCNA-binding-deficient Dnmt1 mutant with efficiency comparable to wild-type Dnmt1. The maintenance of DNA methylation without direct coupling to the replication machinery could in part be explained by the preference of Dnmt1 for hemimethylated sites (8–10). Also, genetic manipulation in the mouse indicate that Dnmt1 is at least 5-fold more abundant than necessary for maintaining normal methylation levels (6,50). Thus, the combined effect of the affinity for hemimethylated sites, relative abundance and simple diffusion could explain the relatively fast immobilization of PCNA-binding-deficient mutants in the presence of 5-aza-dC. In addition, the ability of Dnmt1 to methylate nucleosomal DNA *in vitro* (51–53), suggests that maintenance of DNA methylation is not necessarily restricted to the short time window before nucleosome assembly.

Recent structural data on Ligase I:PCNA and FEN-1:PCNA complexes indicate that for these enzymes PCNA does not simply serves as a loading platform for the replication machinery, but also causes allosteric activation (56,57). The data presented here cannot rule out a similar mechanism for Dnmt1, but clearly show that interaction with PCNA is not a prerequisite for enzyme activity

in vivo. The transient nature of this interaction also argues against PCNA as a classic processivity factor for postreplicative DNA methylation. The major role of the PCNA interaction most likely is to increase the local Dnmt1 concentration and thereby enhance methylation efficiency at replication sites.

Notably, it is still unclear whether the role and regulation of Dnmt1 is similar in different cell types and species. While Dnmt1 is clearly essential for maintenance of DNA methylation in mouse cells (5,54), it seemed dispensable in human tumor HCT116 cells (32). However, two reports recently showed that DNMT1 is essential for maintenance of DNA methylation also in these human tumor cells (32,55). It turned out that genomic methylation was maintained by a residual, truncated DNMT1 form lacking the PBD, arguing that PCNA binding is not strictly required in these cells (32).

Also, the requirement of Dnmt1 for cell viability remains unsettled. In mouse fibroblasts inactivation of the *dnmt1* gene caused a continuous loss of genomic methylation leading to apoptotic cell death after several cell division cycles (54). Similar results were obtained after depletion of DNMT1 activity by RNA interference in human cells (32). Surprisingly, conditional deletion of the *DNMT1* locus in the same cells caused immediate apoptotic cell death long before substantial passive loss of genomic methylation could occur (58) arguing for additional roles of DNMT1. In this regard, the association of Dnmt1 with heterochromatin in G2 phase and occasionally in mitosis in mouse cells (13) would fit well with the mitotic catastrophe observed upon deletion of the *DNMT1* gene in HCT116 cells. In addition, the PBD-mediated association of Dnmt1 with repair sites (14) may indicate a direct role in the maintenance of genome integrity (59). Clearly, more experiments are required to resolve the species and cell-type-specific role and regulation of Dnmt1.

In summary, we demonstrate that the association of Dnmt1 with the replication machinery is not strictly required for maintaining global methylation but still enhances *in vivo* methylation efficiency by 2-fold. Whereas the benefit of Dnmt1 to be directly recruited to replication foci seems subtle in short-term cell culture experiments, it may be more relevant in long-lived organisms and in situations where the nuclear concentration of Dnmt1 is limiting. Indeed, Dnmt1 levels vary considerably in different tissues and developmental stages (60,61). Based on sequence features of the *dnmt1* gene a modular structure was proposed to originate from an ancestral DNA methyltransferase that evolved by stepwise acquisition of new domains (47). Thus, the improved efficiency of postreplicative methylation achieved by the PBD-mediated transient binding to PCNA likely represents an additional safety mechanism, which was acquired in the course of evolution and contributes to the faithful maintenance of epigenetic information over the entire lifespan of complex organisms.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Conflict of interest statement. None declared.

REFERENCES

1. Bird,A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev.*, **16**, 6–21.
2. Li,E., Bestor,T.H. and Jaenisch,R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, **69**, 915–926.
3. Goll,M.G. and Bestor,T.H. (2005) Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.*, **74**, 481–514.
4. Hermann,A., Gowher,H. and Jeltsch,A. (2004) Biochemistry and biology of mammalian DNA methyltransferases. *Cell. Mol. Life Sci.*, **61**, 2571–2587.
5. Lei,H., Oh,S., Okano,M., Juttermann,R., Goss,K., Jaenisch,R. and Li,E. (1996) De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development*, **122**, 3195–3205.
6. Gaudet,F., Hodgson,J.G., Eden,A., Jackson-Grusby,L., Dausman,J., Gray,J.W., Leonhardt,H. and Jaenisch,R. (2003) Induction of tumors in mice by genomic hypomethylation. *Science*, **300**, 489–492.
7. Jones,P.A. and Baylin,S.B. (2002) The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.*, **3**, 415–428.
8. Bestor,T.H. and Ingram,V.M. (1983) Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proc. Natl Acad. Sci. USA*, **80**, 5559–5563.
9. Pradhan,S., Talbot,D., Sha,M., Benner,J., Hornstra,L., Li,E., Jaenisch,R. and Roberts,R.J. (1997) Baculovirus-mediated expression and characterization of the full-length murine DNA methyltransferase. *Nucleic Acids Res.*, **25**, 4666–4673.
10. Hermann,A., Goyal,R. and Jeltsch,A. (2004) The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J. Biol. Chem.*, **279**, 48350–48359.
11. 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.
12. Chuang,L.S.-H., Ian,H.-I., Koh,T.-W., Ng,H.-H., Xu,G. and Li,B.F.L. (1997) Human DNA-(Cytosine-5) Methyltransferase-PCNA Complex as a Target for p21WAF1. *Science*, **277**, 1996–2000.
13. Easwaran,H.P., Schermelleh,L., Leonhardt,H. and Cardoso,M.C. (2004) Replication-independent chromatin loading of Dnmt1 during G2 and M phases. *EMBO Rep.*, **5**, 1181–1186.
14. Mortusewicz,O., Schermelleh,L., Walter,J., Cardoso,M.C. and Leonhardt,H. (2005) Recruitment of DNA methyltransferase I to DNA repair sites. *Proc. Natl Acad. Sci. USA*, **102**, 8905–8909.
15. Jackson,D.A. and Pombo,A. (1998) Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. *J. Cell Biol.*, **140**, 1285–1295.
16. Pradhan,S., Bacolla,A., Wells,R.D. and Roberts,R.J. (1999) Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. *J. Biol. Chem.*, **274**, 33002–33010.
17. Klimasauskas,S., Kumar,S., Roberts,R.J. and Cheng,X. (1994) HhaI methyltransferase flips its target base out of the DNA helix. *Cell*, **76**, 357–369.
18. Chen,L., MacMillan,A.M., Chang,W., Ezaz-Nikpay,K., Lane,W.S. and Verdine,G.L. (1991) Direct identification of the active-site nucleophile in a DNA (cytosine-5)-methyltransferase. *Biochemistry*, **30**, 11018–11025.
19. Santi,D.V., Garrett,C.E. and Barr,P.J. (1983) On the mechanism of inhibition of DNA-cytosine methyltransferases by cytosine analogs. *Cell*, **33**, 9–10.
20. Montecucco,A., Savini,E., Weighardt,F., Rossi,R., Ciarrocchi,G., Villa,A. and Biamonti,G. (1995) The N-terminal domain of human DNA ligase I contains the nuclear localization signal and directs the enzyme to sites of DNA replication. *Embo J.*, **14**, 5379–5386.
21. Cardoso,M.C., Joseph,C., Rahn,H.P., Reusch,R., Nadal-Ginard,B. and Leonhardt,H. (1997) Mapping and use of a sequence that targets DNA ligase I to sites of DNA replication in vivo. *J. Cell Biol.*, **139**, 579–587.
22. Cardoso,M.C., Leonhardt,H. and Nadal-Ginard,B. (1993) Reversal of terminal differentiation and control of DNA replication: cyclin A and Cdk2 specifically localize at subnuclear sites of DNA replication. *Cell*, **74**, 979–992.
23. Sobczak-Thepot,J., Harper,F., Florentin,Y., Zindy,F., Brechot,C. and Puvion,E. (1993) Localization of cyclin A at the sites of cellular DNA replication. *Exp. Cell Res.*, **206**, 43–48.
24. Krude,T. (1995) Chromatin assembly factor 1 (CAF-1) colocalizes with replication foci in HeLa cell nuclei. *Exp. Cell Res.*, **220**, 304–311.
25. Maga,G. and Hubscher,U. (2003) Proliferating cell nuclear antigen (PCNA): a dancer with many partners. *J. Cell Sci.*, **116**, 3051–3060.
26. Sporbert,A., Domaing,P., Leonhardt,H. and Cardoso,M.C. (2005) PCNA acts as a stationary loading platform for transiently interacting Okazaki fragment maturation proteins. *Nucleic Acids Res.*, **33**, 3521–3528.
27. Schermelleh,L., Spada,F., Easwaran,H.P., Zolghadr,K., Margot,J.B., Cardoso,M.C. and Leonhardt,H. (2005) Trapped in action: direct visualization of DNA methyltransferase activity in living cells. *Nat. Methods*, **2**, 751–756.
28. Ho,S.N., Hunt,H.D., Horton,R.M., Pullen,J.K. and Pease,L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, **77**, 51–59.
29. Ito,W., Ishiguro,H. and Kurosawa,Y. (1991) A general method for introducing a series of mutations into cloned DNA using the polymerase chain reaction. *Gene*, **102**, 67–70.
30. Chen,T., Ueda,Y., Dodge,J.E., Wang,Z. and Li,E. (2003) Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol. Cell. Biol.*, **23**, 5594–5605.
31. Boussif,O., Lezoualc'h,F., Zanta,M.A., Mergny,M.D., Scherman,D., Demeneix,B. and Behr,J.P. (1995) A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc. Natl Acad. Sci. USA*, **92**, 7297–7301.
32. Spada,F., Haemmer,A., Kuch,D., Rothbauer,U., Schermelleh,L., Kremmer,E., Carell,T., Langst,G. and Leonhardt,H. (2007) DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells. *J. Cell Biol.*, **176**, 565–571.
33. Sambrook,J. and Russell,D.W. (2001) *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
34. Warnecke,P.M., Stirzaker,C., Song,J., Grunau,C., Melki,J.R. and Clark,S.J. (2002) Identification and resolution of artifacts in bisulfite sequencing. *Methods*, **27**, 101–107.
35. Warnecke,P.M. and Clark,S.J. (1999) DNA methylation profile of the mouse skeletal alpha-actin promoter during development and differentiation. *Mol. Cell. Biol.*, **19**, 164–172.
36. Warnecke,P.M., Biniszkiewicz,D., Jaenisch,R., Frommer,M. and Clark,S.J. (1998) Sequence-specific methylation of the mouse H19

gene in embryonic cells deficient in the Dnmt-1 gene. *Dev. Genet.*, **22**, 111–121.

37. Ko,Y.-G., Nishino,K., Hattori,N., Arai,Y., Tanaka,S. and Shiota,K. (2005) Stage-by-stage change in DNA methylation status of Dnmt1 locus during mouse early development. *J. Biol. Chem.*, **280**, 9627–9634.

38. McDonald,L.E., Paterson,C.A. and Kay,G.F. (1998) Bisulfite genomic sequencing-derived methylation profile of the xist gene throughout early mouse development. *Genomics*, **54**, 379–386.

39. Hajkova,P., Erhardt,S., Lane,N., Haaf,T., El-Maarri,O., Reik,W., Walter,J. and Surani,M.A. (2002) Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.*, **117**, 15–23.

40. Warnecke,P., Stirzaker,C., Melki,J., Millar,D., Paul,C. and Clark,S. (1997) Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA. *Nucleic Acids Res.*, **25**, 4422–4426.

41. Warbrick,E. (1998) PCNA binding through a conserved motif. *Bioessays*, **20**, 195–199.

42. Easwaran,H.P., Leonhardt,H. and Cardoso,M.C. (2005) Cell cycle markers for live cell analyses. *Cell Cycle*, **4**.

43. Sporbert,A., Gahl,A., Ankerhold,R., Leonhardt,H. and Cardoso,M.C. (2002) DNA polymerase clamp shows little turnover at established replication sites but sequential de novo assembly at adjacent origin clusters. *Mol. Cell*, **10**, 1355–1365.

44. Phair,R.D., Scaffidi,P., Elbi,C., Vecerova,J., Dey,A., Ozato,K., Brown,D.T., Hager,G., Bustin,M. *et al.* (2004) Global nature of dynamic protein-chromatin interactions in vivo: three-dimensional genome scanning and dynamic interaction networks of chromatin proteins. *Mol. Cell Biol.*, **24**, 6393–6402.

45. Sprague,B.L. and McNally,J.G. (2005) FRAP analysis of binding: proper and fitting. *Trends Cell Biol.*, **15**, 84–91.

46. Reits,E.A. and Neefjes,J.J. (2001) From fixed to FRAP: measuring protein mobility and activity in living cells. *Nat. Cell Biol.*, **3**, E145–E147.

47. Margot,J.B., Aguirre-Arteta,A.M., Di Giacco,B.V., Pradhan,S., Roberts,R.J., Cardoso,M.C. and Leonhardt,H. (2000) Structure and function of the mouse DNA methyltransferase gene: Dnmt1 shows a tripartite structure. *J. Mol. Biol.*, **297**, 293–300.

48. Vilkaitis,G., Suetake,I., Klimasauskas,S. and Tajima,S. (2005) Processive methylation of hemimethylated CpG sites by mouse Dnmt1 DNA methyltransferase. *J. Biol. Chem.*, **280**, 64–72.

49. Tucker,K.L., Beard,C., Dausmann,J., Jackson-Grusby,L., Laird,P.W., Lei,H., Li,E. and Jaenisch,R. (1996) Germ-line passage is required for establishment of methylation and expression patterns of imprinted but not of nonimprinted genes. *Genes Dev.*, **10**, 1008–1020.

50. Gaudet,F., Rideout,W.M.3rd, Meissner,A., Dausman,J., Leonhardt,H. and Jaenisch,R. (2004) Dnmt1 expression in pre- and postimplantation embryogenesis and the maintenance of IAP silencing. *Mol. Cell. Biol.*, **24**, 1640–1648.

51. Okuwaki,M. and Verreault,A. (2004) Maintenance DNA methylation of nucleosome core particles. *J. Biol. Chem.*, **279**, 2904–2912.

52. Robertson,A.K., Geiman,T.M., Sankpal,U.T., Hager,G.L. and Robertson,K.D. (2004) Effects of chromatin structure on the enzymatic and DNA binding functions of DNA methyltransferases DNMT1 and Dnmt3a in vitro. *Biochem. Biophys. Res. Commun.*, **322**, 110–118.

53. Gowher,H., Stockdale,C.J., Goyal,R., Ferreira,H., Owen-Hughes,T. and Jeltsch,A. (2005) De novo methylation of nucleosomal DNA by the mammalian Dnmt1 and Dnmt3A DNA methyltransferases. *Biochemistry*, **44**, 9899–9904.

54. Jackson-Grusby,L., Beard,C., Possemato,R., Tudor,M., Fambrough,D., Csankovszki,G., Dausman,J., Lee,P., Wilson,C. *et al.* (2001) Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nat. Genet.*, **27**, 31–39.

55. Egger,G., Jeong,S., Escobar,S.G., Cortez,C.C., Li,T.W., Saito,Y., Yoo,C.B., Jones,P.A. and Liang,G. (2006) Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. *Proc. Natl Acad. Sci. USA*, **103**, 14080–14085.

56. Chapados,B.R., Hosfield,D.J., Han,S., Qiu,J., Yelent,B., Shen,B. and Tainer,J.A. (2004) Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. *Cell*, **116**, 39–50.

57. Pascal,J.M., Tsodikov,O.V., Hura,G.L., Song,W., Cotner,E.A., Classen,S., Tomkinson,A.E., Tainer,J.A. and Ellenberger,T. (2006) A flexible interface between DNA ligase and PCNA supports conformational switching and efficient ligation of DNA. *Mol. Cell*, **24**, 279–291.

58. Chen,T., Hevi,S., Gay,F., Tsujimoto,N., He,T., Zhang,B., Ueda,Y. and Li,E. (2007) Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells. *Nat. Genet.*, **39**, 391.

59. Brown,K.D. and Robertson,K.D. (2007) DNMT1 knockout delivers a strong blow to genome stability and cell viability. *Nat. Genet.*, **39**, 289–290.

60. Ratnam,S., Mertineit,C., Ding,F., Howell,C.Y., Clarke,H.J., Bestor,T.H., Chaillet,J.R. and Trasler,J.M. (2002) Dynamics of Dnmt1 methyltransferase expression and intracellular localization during oogenesis and preimplantation development. *Dev. Biol.*, **245**, 304–314.

61. Robertson,K.D., Uzvolgyi,E., Liang,G., Talmadge,C., Sumegi,J., Gonzales,F.A. and Jones,P.A. (1999) The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. *Nucleic Acids Res.*, **27**, 2291–2298.
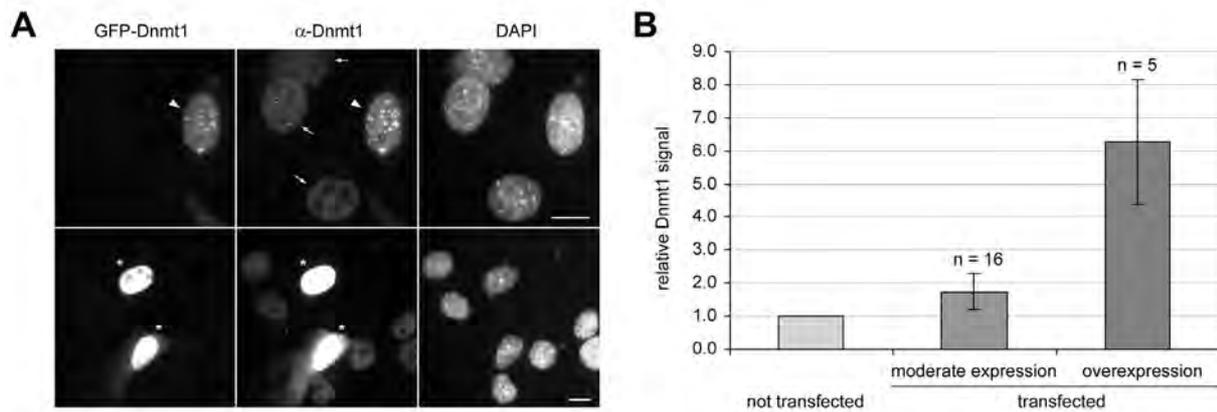
Supplementary Figure 1. Comparison of endogenous and GFP-tagged Dnmt1 levels in transiently transfected C2C12 myoblasts. Cells were fixed with 3.7% formaldehyde 24 h after transfection with the GFP-Dnmt1[wt] expression construct. Endogenous and GFP-tagged Dnmt1 were detected with polyclonal rabbit antibodies against the N-terminal domain of Dnmt1 and Alexa555 coupled secondary antibodies. DNA was counterstained with 4',6-diamidino-2-phenylindole (DAPI). Widefield epifluorescence images were acquired with an Axiophot 2 microscope (Zeiss) equipped with a 63x/1.4 NA oil objective and a 12-bit CCD camera. (**A**) The upper row shows exemplary images with a late S phase cell moderately expressing GFP-Dnmt1 typically chosen for live cell microscopy (arrowhead). Dnmt1 immunostaining shows slightly higher fluorescence compared to untransfected cells (arrows) due to the ectopically expressed GFP-Dnmt1 in addition to endogenous Dnmt1. Overexpressing cells with much brighter nuclear fluorescence (lower row, asterisks) were not considered for live cell analyses. Scale bars: 10 µm. (**B**) Quantification of total Dnmt1 levels in transfected and untransfected cells. Average nuclear antibody signal (endogenous Dnmt1 and GFP-Dnmt1) from transfected and untransfected cells was determined for each image. Signal ratios of moderately and overexpressing transfected cells to untransfected cells were calculated for each image and then averaged. Error bars indicate SD.
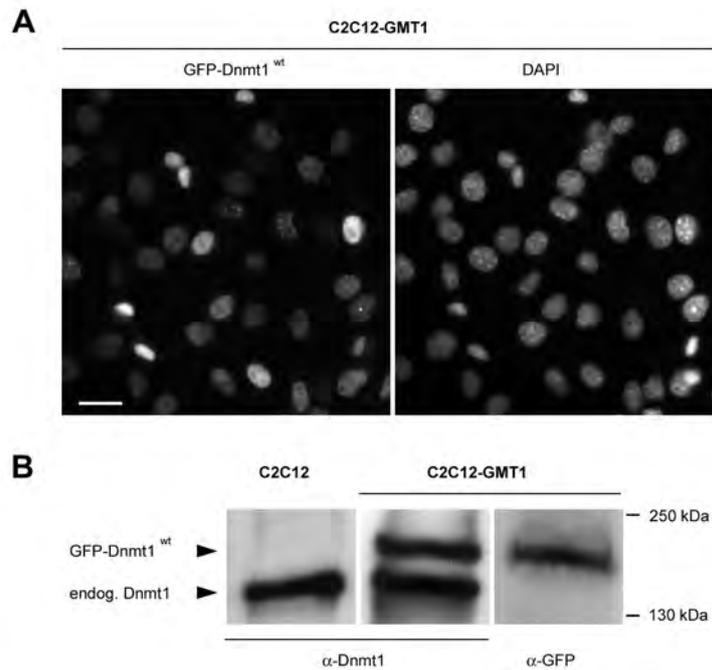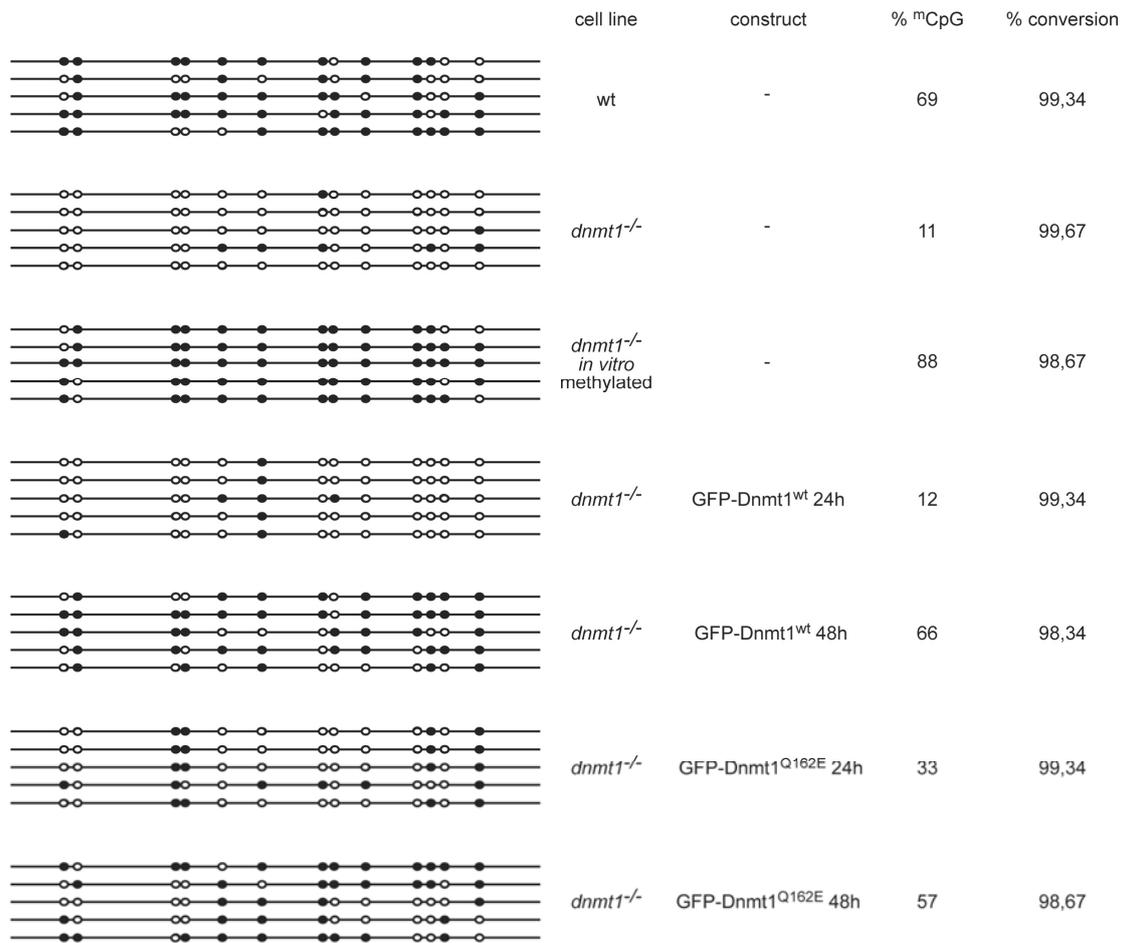
Supplementary Figure 2. Comparison of endogenous and GFP-tagged Dnmt1 levels in stably transfected C2C12 myoblasts (C2C12-GMT1) after 20 days of selection with 10 µg/ml blasticidin. (**A**) Cells were fixed with 3.7% formaldehyde and DNA was counterstained with 4',6-diamidino-2-phenylindole (DAPI). Widefield epifluorescence images were acquired with an Axiophot 2 microscope (Zeiss) equipped with a 63x/1.4 NA oil objective and a 16-bit CCD camera. About 75% of the cells show stable expression of GFP-Dnmt1. Scale bar: 10 µm. (**B**) Western blot analysis to compare the levels of GFP-Dnmt1[wt] and endogenous Dnmt1 in stably transfected C2C12 cells. The detection was performed with an antibody against the N-terminal domain of Dnmt1 (Grohmann et al., 2005, *BMC Dev Biol* **5**, 18) and a monoclonal anti-GFP antibody (Roche). Untransfected C2C12 cells were used as a control. After correction for the fraction of expressing cells, the expression level of GFP-Dnmt1[wt] amounts to about 90% of the endogenous Dnmt1.
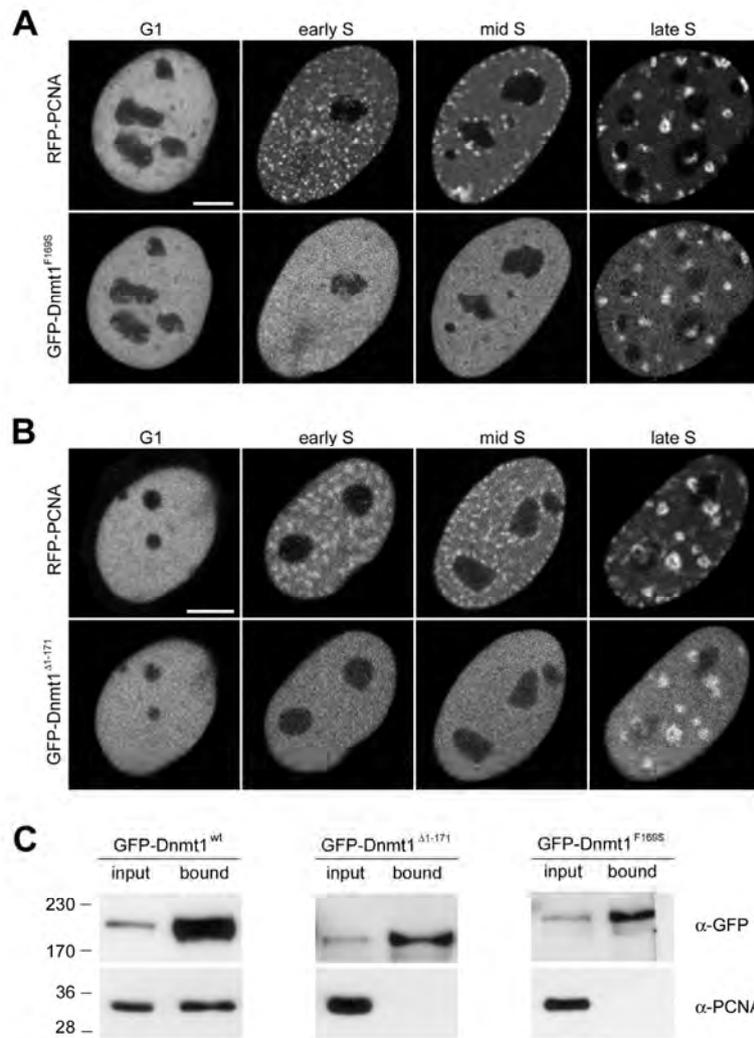
Supplementary Figure 3. Mutations of the PBD abolish Dnmt1 interaction with PCNA and prevent accumulation at replication sites in early and mid S phase. (**A, B**) Confocal mid sections of living mouse C2C12 myoblasts expressing either GFP-Dnmt1^{F169S} (**A**) or GFP-Dnmt1$^{\Delta1-171}$ (**B**). Both constructs show a dispersed nuclear distribution in early and mid S phase stages, whereas in late S phase association with pericentric heterochromatin is observed. Cells were co-transfected with RFP-PCNA to identify replication foci and to distinguish S phase stages. Scale bars: 5 μm. (**C**) Endogenous PCNA efficiently co-immunoprecipitates with GFP-Dnmt1^{wt} but not with GFP-Dnmt1^{F169S} and GFP- Dnmt1$^{\Delta1-171}$. GFP fusion constructs were expressed in HEK 293T cells. Precipitated proteins were detected by immunostaining with antibodies against GFP and PCNA, respectively.
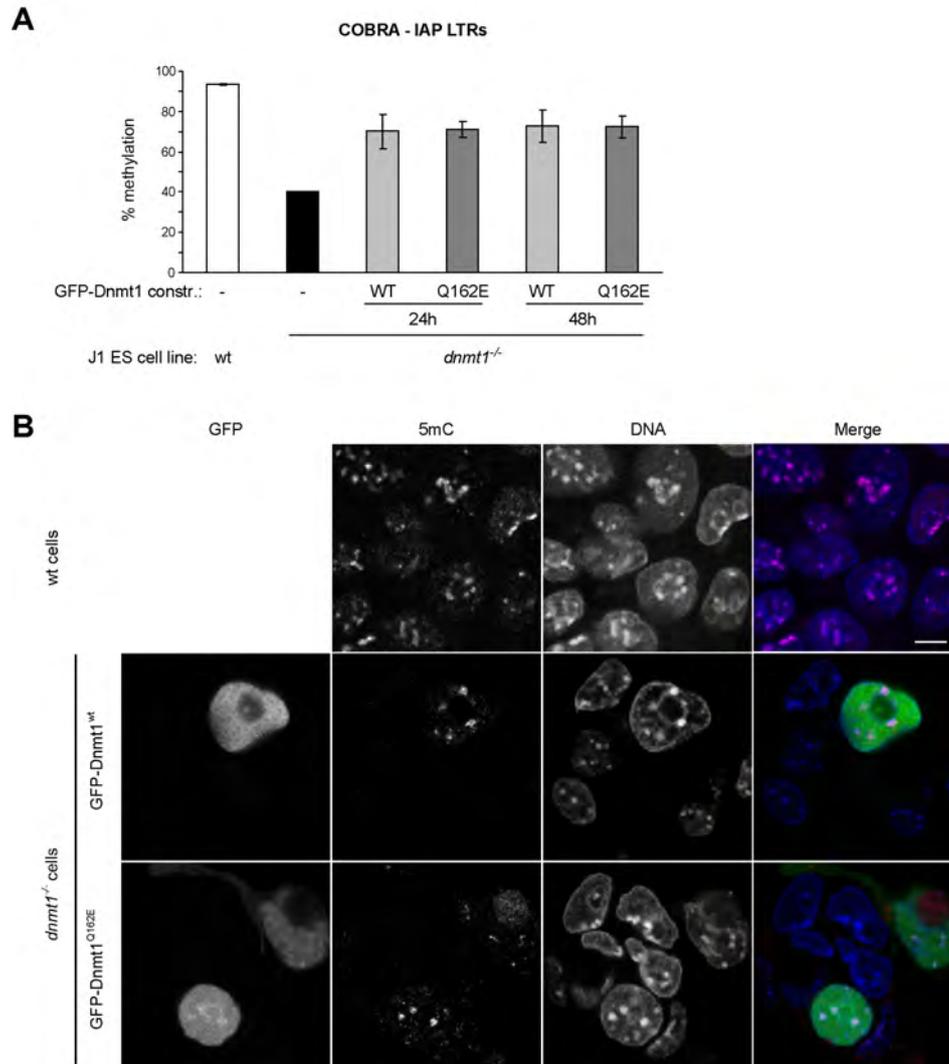
Supplementary Figure 4. PCNA binding-deficient Dnmt1 mutants restore methylation of repetitive DNA sequences in Dnmt1 null ES cells. (**A**) Mouse *dnmt1*[-/-] J1 ES cells were transiently transfected with either the GFP-Dnmt1[wt] or GFP-Dnmt1[Q162E] constructs and FACS-sorted 24 and 48 h after transfection as in Figure 7. Methylation of IAP LTRs was assayed by COBRA. Shown are cumulative methylation percentages at 2 HpyCH4IV sites (ACGT) with respect to genomic DNA from *dnmt1*[-/-] ES cells methylated *in vitro* with the recombinant methyltransferase *Sss*I (100% methylation). Mean values and standard errors of duplicate amplifications from each of two independent experiments are indicated. (**B**) 48 h after transfection with either GFP-Dnmt1[wt] or GFP-Dnmt1[Q162E] Dnmt1 null ES cells were stained with an antibody against 5-methylcytidine (5mC) and their DNA counterstained with TO-PRO-3 as described (Schermelleh et al., 2005, *Nat Methods* **2**, 751). Staining of wild type ES cells is shown as a control in the first row. Individual channels are indicated at the top and in the merge channel GFP, 5mC and TO-PRO-3 stainings are shown in green, red and blue, respectively. Cells expressing either GFP-Dnmt1[wt] or GFP-Dnmt1[Q162E] showed remethylation of repetitive DNA at chromocenters. Similar results were obtained with the GFP-Dnmt1[F169S] and GFP-Dnmt1[Δ1-171] mutant constructs (data not shown). Scale bar: 5 μm.

cell line	construct	% ^{m}CpG	% conversion
wt	-	69	99,34
dnmt1$^{-/-}$	-	11	99,67
dnmt1$^{-/-}$ *in vitro* methylated	-	88	98,67
dnmt1$^{-/-}$	GFP-Dnmt1^{wt} 24h	12	99,34
dnmt1$^{-/-}$	GFP-Dnmt1^{wt} 48h	66	98,34
dnmt1$^{-/-}$	GFP-Dnmt1^{Q162E} 24h	33	99,34
dnmt1$^{-/-}$	GFP-Dnmt1^{Q162E} 48h	57	98,67

Supplementary Figure 5. Bisulfite sequencing of the skeletal α-actin promoter in wt, Dnmt1 null and Dnmt1 null J1 ES cells transiently transfected with either the GFP-Dnmt1^{wt} or GFP-Dnmt1^{Q162E} constructs as in Figure 7. Transfected cells were FACS-sorted and genomic DNA isolated 24 and 48 h after transfection as indicated. Genomic DNA from *dnmt1$^{-/-}$* ES cells methylated *in vitro* with recombinant *Sss*I methyltransferase was used as a control. PCR products were cloned with the StrataClone PCR Cloning Kit (Stratagene) and individual clones sequenced by automated sequencing at MWG Biotech. Black lines represent sequences from individual clones. Open and closed circles represent unmethylated and methylated cytosines, respectively. Global percentages of converted cytosines and methylation at CpG sites within the analyzed sequence are indicated on the right.
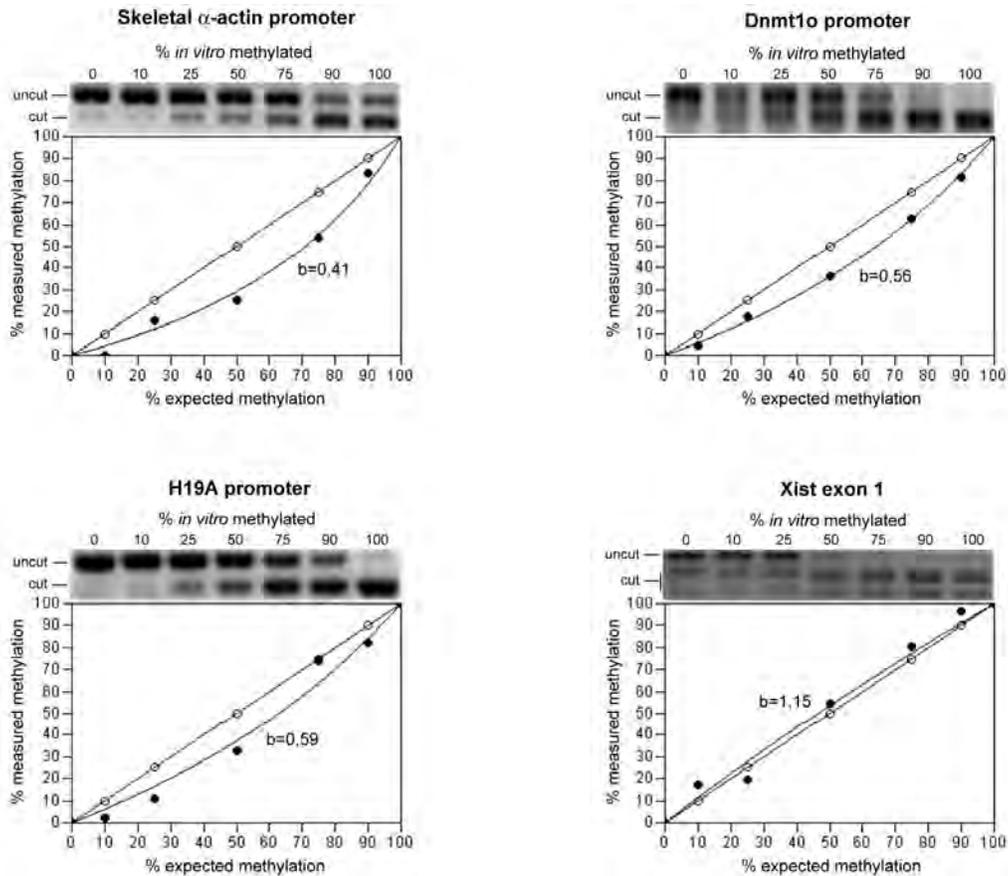
Supplementary Figure 6. Determination of PCR bias in COBRA assays. Genomic DNA from untransfected *Dnmt1*$^{-/-}$ J1 cells was methylated *in vitro* with recombinant SssI methyltransferase (New England BioLabs) and mixed with unmethylated DNA from the same cells in the proportions reported at the top of each lane. COBRA assays for the indicated sequences were performed on each sample as described in Materials and Methods. At the top of each panel the digests of bisulfite PCR fragments from these genomic DNA mixtures are shown. Uncut fragments, corresponding to unmethylated molecules, and restriction fragments (cut; only one out of two for *skeletal α-actin*, *Dnmt1o* and *H19A* promoters and two out of three for *Xist* exon 1), corresponding to methylated molecules, are shown. Measured methylation percentages were plotted against expected ones. Closed and open circles represent measured values and values expected in the absence of bias. Best fit curves and bias coefficients (b) were generated and calculated, respectively, using WinCurveFit (Kevin Raner Software) and according to the equation:

$$\frac{x}{100-x} \times b = \frac{y}{100-y}$$

where *x* is the expected value and *y* the measured one.

Different binding properties and functional relevance of the CXXC zinc finger domains in Dnmt1 and Tet1

2.2   Different binding properties and functional relevance of CXXC zinc finger domains in Dnmt1 and Tet1

PLoS one

Different Binding Properties and Function of CXXC Zinc Finger Domains in Dnmt1 and Tet1

Carina Frauer[1][9], Andrea Rottach[1][9], Daniela Meilinger[1], Sebastian Bultmann[1], Karin Fellinger[1¤a], Stefan Hasenöder[1¤b], Mengxi Wang[1], Weihua Qin[1], Johannes Söding[2], Fabio Spada[1]*, Heinrich Leonhardt[1]*

1 Department of Biology II and Center for Integrated Protein Science Munich (CIPS^M), Ludwig Maximilians University Munich, Planegg, Germany, 2 Gene Center Munich, Ludwig Maximilians University Munich, Munich, Germany.

Abstract

Several mammalian proteins involved in chromatin and DNA modification contain CXXC zinc finger domains. We compared the structure and function of the CXXC domains in the DNA methyltransferase Dnmt1 and the methylcytosine dioxygenase Tet1. Sequence alignment showed that both CXXC domains have a very similar framework but differ in the central tip region. Based on the known structure of a similar MLL1 domain we developed homology models and designed expression constructs for the isolated CXXC domains of Dnmt1 and Tet1 accordingly. We show that the CXXC domain of Tet1 has no DNA binding activity and is dispensable for catalytic activity in vivo. In contrast, the CXXC domain of Dnmt1 selectively binds DNA substrates containing unmethylated CpG sites. Surprisingly, a Dnmt1 mutant construct lacking the CXXC domain formed covalent complexes with cytosine bases both in vitro and in vivo and rescued DNA methylation patterns in dnmt1^−/− embryonic stem cells (ESCs) just as efficiently as wild type Dnmt1. Interestingly, neither wild type nor ΔCXXC Dnmt1 re-methylated imprinted CpG sites of the H19a promoter in dnmt1^−/− ESCs, arguing against a role of the CXXC domain in restraining Dnmt1 methyltransferase activity on unmethylated CpG sites.
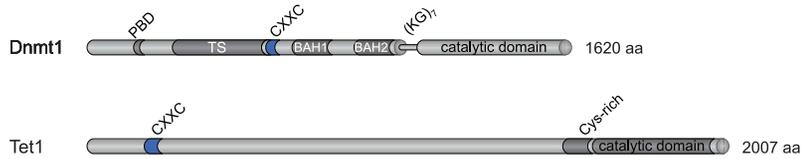
Introduction

In mammals DNA methylation is restricted to cytosine residues and mainly involves CpG dinucleotides. CpG methylation is widespread across mammalian genomes, including gene bodies regardless of their transcriptional activity [1–4]. However, highly CpG-rich regions (CpG islands) are refractory to methylation and mostly coincide with promoters of constitutively active genes. The methylation state of other regulatory sequences with moderate to low CpG density, including promoters and enhancers, shows developmental and/or tissue-specific variations and positively correlates with a transcriptionally silent state [1,3–8]. Dense methylation of repetitive sequences is also thought to maintain these elements in a silent state and thus contribute to genome stability [9–11]. In mammals cytosine methylation is catalyzed by a family of DNA methyltransferases (Dnmts) [12]. Dnmt3a and Dnmt3b establish methylation patterns during embryonic development of somatic as well as germ cell lineages and, consistently, show developmental stage and tissue specific expression patterns. In contrast, Dnmt1 is ubiquitous and generally the most abundant DNA methyltransferase in mammalian tissues, where it associates with the replication machinery and restores symmetrical methylation at hemimethylated CpG sites generated by the semi-

conservative DNA replication process [13]. Thus, Dnmt1 maintains methylation patterns with high fidelity and is essential for embryonic development and genome integrity [9,14,15].

Dnmt1 is a large enzyme with a complex domain structure that likely evolved by fusion of at least three genes [16]. It comprises a regulatory N-terminal region and a C-terminal catalytic domain connected by a linker of seven glycine-lysine repeats (Figure 1A)[17]. The N-terminal part contains a PCNA binding domain (PBD), a heterochromatin targeting sequence (TS), a CXXC-type zinc finger domain and two Bromo-Adjacent Homology domains (BAH1 and BAH2). The C-terminal domains of mammalian Dnmts contain all ten catalytic motifs identified in bacterial DNA (cytosine-5) methyltransferases [12]. Thus, prokaryotic and mammalian cytosine methyltransferases are thought to adopt the same catalytic mechanism. However, the C-terminal domain of Dnmt1 is the only DNA methyltransferase domain in Dnmts that is not catalytically active when expressed separately. Indeed, interaction with the N-terminal part is required for allosteric activation of the enzyme [18]. Remarkably, the first 580 amino acids (aa) of human DNMT1 are dispensable for both enzymatic activity and substrate recognition, whereas deletion of the first 672 aa results in an inactive enzyme [19]. Interestingly, this truncation eliminates part of the CXXC domain, suggesting an involvement of this domain in allosteric activation. However, addition
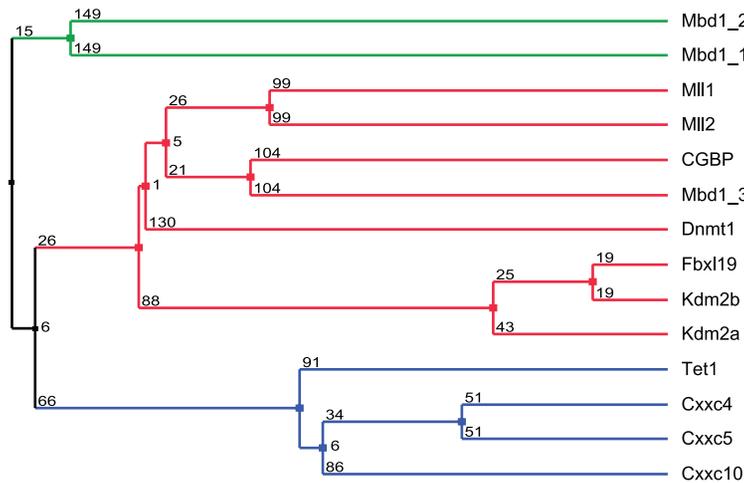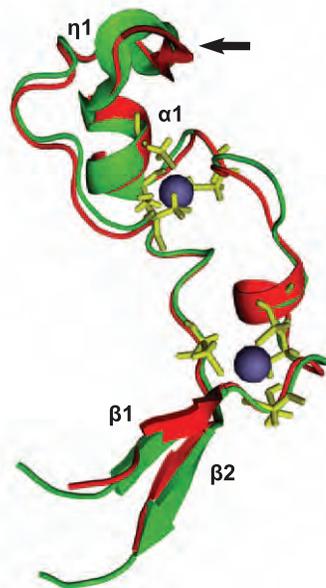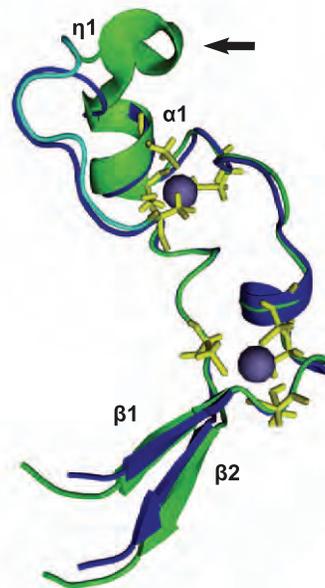
Figure 1. Sequence and predicted structural homology of CXXC domains. (A) Schematic representation of the domain structure in Dnmt1 and Tet1. The catalytic domain and the N-terminal region of Dnmt1 are connected by seven lysine-glycine repeats [(KG)$_7$]. PBD: PCNA binding domain; TS: targeting sequence; CXXC: CXXC-type zinc finger domain; BAH1 and 2: bromo-adjacent homology domain; NLS: nuclear localization signal; Cys-rich: cysteine rich region. **(B)** Alignment of mammalian CXXC domains. Numbers on the right side indicate the position of the last amino acid in the corresponding protein. The Mbd1a isoform contains three CXXC motifs (Mbd1_1-3). Absolutely conserved residues, including the eight cysteines involved in zinc ion coordination are highlighted in red and the conserved KFGG motif is in red bold face. Positions with residues in red face share 70% similarity as calculated with the Risler algorithm [66]. At the top residues of MLL1 involved in β sheets β1 and β2 (black arrows), α helices α1 and α2 and strict α turns (TTT) are indicated. All sequences are from *M. musculus*. Accession numbers (for GenBank unless otherwise stated): Dnmt1, NP_034196; Mll1, NP_001074518; Mll4, O08550 (SwissProt); CGBP, NP_083144; Kdm2a, NP_001001984; Kdm2b, NP_001003953; Fbxl19, NP_766336; Mbd1, NP_038622; CXXC4/Idax, NP_001004367; CXXC5, NP_598448; CXXC10 (see Materials and Methods). **(C)** A homology tree was generated from the alignment in (B). The three subgroups of CXXC domains identified are in different colors. Average distances between the sequences are indicated. **(D–E)** Homology models of the mouse Dnmt1 (D; red) and Tet1 (E; blue) CXXC domains superimposed to the CXXC domain of MLL1 (green; [35]). MLL1 residues that were described to contact DNA according to chemical shift measurements [35] are cyan in (E), while cysteines involved in coordination of the two zinc ions are yellow. Arrows point to the KFGG motif in MLL1 and Dnmt1. The locations of α helices and β sheets are indicated as in (B).
doi:10.1371/journal.pone.0016627.g001

of an N-terminal fragment containing the isolated CXXC domain to the catalytic domain was not sufficient for catalytic activation [20].

CXXC-type zinc finger domains are found in several other proteins with functions related to DNA or chromatin modification, including the histone H3 lysine 4 (H3K4) methyltransferases mixed-lineage leukaemia (MLL) proteins 1 and 4, the CpG-binding protein (CGBP, also known as Cfp1 or CXXC1), the methyl-CpG binding domain protein 1 (MBD1), the H3 lysine 36 (H3K36) demethylases KDM2A and B (also known as JHD1A/FBXL11 and JHD1B/FBXL10) and the MLL1 fusion partner TET1 (Figure 1A) [21–28]. The CXXC domains of some of these proteins were shown to mediate specific binding to double stranded DNA templates containing unmethylated CpG sites [21,22,29,30]. A region of Dnmt1 which mainly includes the CXXC domain (aa 628–753) was also shown to bind Zn ions and DNA [20,31,32]. However, available data on the selectivity of this DNA binding activity are conflicting. Whereas a fragment including aa 613–748 of mouse Dnmt1 was shown to bind DNA with a slight preference for hemimethylated CpG sites [20], aa 645–737 of human DNMT1 were shown to selectively bind unmethylated DNA [32]. As these studies used different constructs and species, the selectivity of DNA binding by the CXXC domain of Dnmt1 with regard to CpG methylation state and the role of the CXXC domain in allosteric activation and substrate discrimination remain to be firmly established.

Notably, not all CXXC domains show DNA binding activity, as exemplified by the fact that only one out of three CXXC domains in MBD1 binds DNA [29]. Interestingly, TET1 was recently shown to be a 2 oxoglutarate- and Fe(II)-dependent dioxygenase responsible for converting genomic 5-methylcytosine (mC) to 5-hydroxymethylcytosine (hmC) [33,34]. However, it is not known whether the CXXC domain of TET1 is involved in recognition of methylated DNA substrates.

Here we report a functional study and characterization of the DNA binding activity for the CXXC domains of mouse Dnmt1 and Tet1 proteins. We generated isolated CXXC domain and deletion constructs based on structural homology models to minimize structural alterations. We show that the CXXC domain of Dnmt1 preferentially binds DNA substrates containing unmethylated CpG sites, but does not contribute significantly to the DNA binding properties of the full length enzyme and is dispensable for its catalytic activity *in vitro* and *in vivo*. In addition, we found that the CXXC domain of Tet1 does not bind DNA *in vitro* and is also dispensable for catalytic activity of Tet1 *in vivo*.

Results

Sequence homology and structural modeling identify distinct CXXC domain subtypes

Dnmt1 contains a zinc finger domain of the CXXC type, which is present in several mammalian proteins including MLL1

(Figure 1A–C) and is highly conserved among Dnmt1 sequences from various animal species (Figure S1 in File S1). The primary structure of CXXC domains spans two clusters of 6 and 2 cysteine residues separated by a stretch of variable sequence and length. Sequence alignment and homology tree construction identified three distinct groups of CXXC domains (Figure 1B and C). The sequence between the two cysteine clusters in the CXXC domains of Dnmt1, CGBP/Cfp1, Fbxl19, Mll1, Mll2 and Kdm2 proteins and CXXC domain 3 of Mbd1 is highly conserved and contains a KFGG motif. The two other homology groups, including the CXXC domains 1 and 2 of Mbd1 on one side and those of Tet1, Cxxc4/Idax, Cxxc5/RINF and Cxxc10 on the other side, lack the KFGG motif and diverge from the first group and from each other in the sequence between the cysteine clusters. We generated structural homology models for the CXXC domains of mouse Dnmt1 and Tet1 using the NMR structure of the MLL1 CXXC domain as a template (Figure 1D and E)[35]. The CXXC domains of these proteins adopt an extended crescent-like structure that incorporates two Zn^{2+} ions each coordinated by four cysteine residues. The peptide of the MLL1 CXXC domain predicted to insert into the major groove of the DNA double helix (cyan in Fig. 1E) is located on one face of the structure and is contiguous to the KFGG motif [35]. The predicted structure of the Tet1 CXXC domain lacks the short 3$_{10}$ helix (η1 in Figure 1E) formed by residues PKF and partially overlapping the KFGG motif, but is similar to the MLL1 CXXC domain in the region of the DNA-contacting peptide. However, each of the two predicted β-strands in Tet1 carries three positive charges, whereas there is only one or no charged residue in the C-terminal strands of the CXXC domains in MLL1 and Dnmt1. Depending on the orientation of the positively charged side chains, it cannot be excluded that the charge density prevents strand pairing in the Tet1 CXXC domain.

The Dnmt1 CXXC domain binds unmethylated DNA

To investigate the binding properties of the Dnmt1 CXXC domain, we generated a GFP fusion construct including aa 652–699 (GFP-CXXC^{Dnmt1}). According to our homology model the ends of this fragment form an antiparallel β-sheet that structurally delimits the domain as in MLL1. We first compared the localization and mobility of GFP-CXXC^{Dnmt1} and GFP in mouse C2C12 myoblasts. While GFP was diffusely distributed in both nucleus and cytoplasm, GFP-CXXC^{Dnmt1} was exclusively nuclear with a punctuated pattern throughout the nucleoplasm and was enriched in nucleoli, a pattern independent of cell cycle stage (Figure 2A and Figure S2 in File S1). Enrichment in the nucleus and nucleoli is frequently observed with constructs containing stretches with high density of basic residues. After photobleaching half of the nuclear volume we observed a slower fluorescence recovery rate for GFP-CXXC^{Dnmt1} than for GFP (Figure 2B). To rule out a contribution of nucleolar interactions to the slower kinetics of GFP-CXXC^{Dnmt1},

we separately bleached nucleoplasmic and nucleolar regions and found that GFP-CXXC^{Dnmt1} has even faster kinetics within the nucleolus (Figure S3 in File S1). These results are consistent with a binding activity of GFP-CXXC^{Dnmt1} in the nucleus and very transient, unspecific binding in the nucleolus. To investigate whether the CXXC domain of Dnmt1 binds DNA and its possible selectivity with respect to CpG methylation we used a recently developed fluorescent DNA binding assay [36,37]. GFP-CXXC^{Dnmt1} was transiently expressed in HEK293T cells, immunopurified with the GFP-trap (Figure S4 in File S1) and incubated with fluorescent DNA substrates containing either no CpG site or one central un-, hemi- or fully methylated CpG site in direct competition. As shown in Figure 2C, GFP-CXXC^{Dnmt1} displayed a significant preference for the substrate containing one unmethylated CpG site, which increased substantially with a five-fold higher concentration of the DNA substrates (Figure S5 in File S1). These results are consistent with the reported binding preference of the CXXC domains in human DNMT1 and other factors belonging to the same CXXC homology group [21,22,29,32]. Notably, the CXXC domains 1 and 2 of Mbd1 lack the KFGG motif and do not bind DNA, while mutation of this motif prevented DNA binding by the CXXC domain of MLL1 [29,38]. Therefore,

we generated a GFP-CXXC^{Dnmt1} construct where the KFGG motif was mutated to AAGG (GFP-CXXC$^{Dnmt1KF/AA}$, Figure S4 in File S1) to test the requirement of the KFGG motif for binding by the CXXC domain of Dnmt1. The mutant domain showed significantly decreased binding to all DNA substrates and complete loss of preferential binding to the unmethylated substrate in vitro (Figure 2B). In addition, GFP-CXXC$^{Dnmt1KF/AA}$ showed faster recovery after photobleaching (FRAP) in vivo compared to the corresponding wild type construct (Figure 2C). These results further support the importance of the KFGG motif for DNA binding by CXXC domains.

The CXXC domain of Tet1 shows no specific DNA binding activity and is dispensable for enzymatic activity in vivo

It was recently shown that Tet1 oxidizes genomic mC to hmC. However, the mechanism by which Tet1 is targeted to genomic mC is not known. Our model for the structure of the Tet1 CXXC domain diverged from the structure of the MLL1 CXXC domain with respect to the KFGG motif but not to the DNA-contacting peptide, suggesting that the Tet1 CXXC domain may still bind DNA. To test this we generated a GFP-tagged Tet1 CXXC
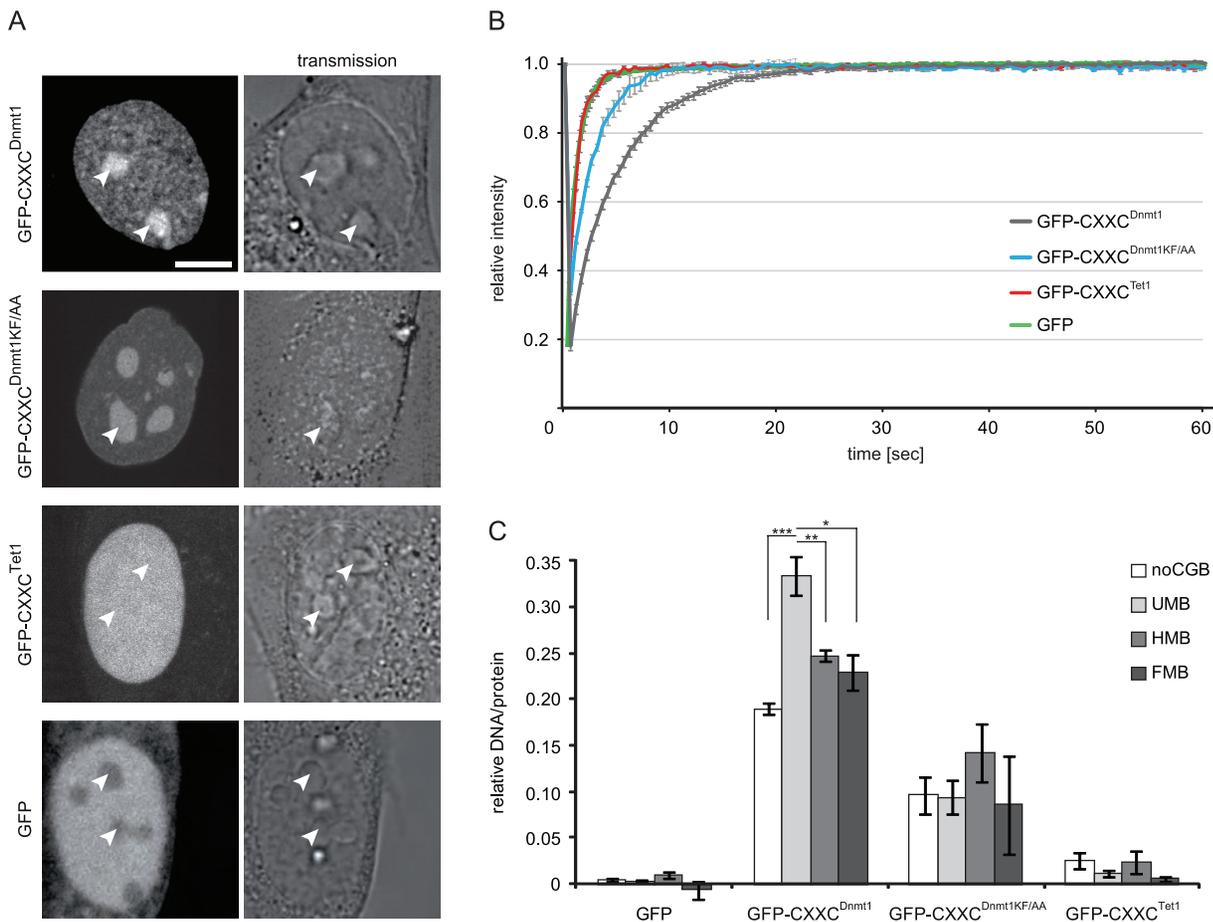
Figure 2. Properties of isolated Dnmt1 and Tet1 CXXC domains. (A–B) Subcellular localization (A) and binding kinetics (B) of GFP-CXXC^{Dnmt1}, GFP-CXXC$^{Dnmt1KF/AA}$, GFP-CXXC^{Tet1} and GFP in mouse C2C12 myoblasts. Localization and binding kinetics were independent from the cell cycle stage (Figures S2 and S5 in File S1). Arrowheads in (A) point to nucleoli. Scale bar: 5 μm. Binding kinetics were analyzed by FRAP. **(C)** DNA binding specificity of the Dnmt1 and Tet1 CXXC domains. GFP, GFP-CXXC^{Dnmt1}, GFP-CXXC$^{Dnmt1KF/AA}$ and GFP-CXXC^{Tet1} were pulled down from extracts of transiently transfected HEK293T cells and incubated with fluorescent DNA substrates containing no CpG site or one central un-, hemi- or fully methylated CpG site in direct competition (noCGB, UMB, HMB, FMB, respectively). Shown are the mean DNA/protein ratios and corresponding standard errors from 5 (GFP), 4 (GFP-CXXC^{Dnmt1} and GFP-CXXC$^{Dnmt1KF/AA}$) and 2 (GFP-CXXC^{Tet1}) independent experiments. * $P = 0.01$; ** $P = 0.007$; ***$P = 0.001$.
doi:10.1371/journal.pone.0016627.g002

construct (GFP-CXXC^{Tet1}) following the same criteria as for GFP-CXXC^{Dnmt1} and investigated its cellular localization, *in vivo* binding kinetics and *in vitro* DNA binding activity. GFP-CXXC^{Tet1} was prevalently nuclear with a homogeneous distribution including nucleoli that was independent of cell cycle stage (Figure 2A and Figure S6 in File S1). After photobleaching GFP-CXXC^{Tet1} showed very fast recovery kinetics similar to GFP (Figure 2B) and its DNA binding activity *in vitro* was also similar to the background levels of the GFP control (Figure 2C). We conclude that the isolated CXXC domain of Tet1 has no specific DNA binding activity. Together with the observation that the CXXC domains 1 and 2 of Mbd1 also lack the KFGG motif and do not bind DNA [29] and that mutation of this motif reduced DNA binding by the CXXC domains of both Dnmt1 (Figure 2C) and MLL1 [38], this result indicates that the KFGG motif is a major determinant for DNA binding by CXXC domains.

To assess whether the CXXC domain is required for catalytic activity of Tet1 we generated a GFP-Tet1 fusion construct and a corresponding mutant lacking the CXXC domain (GFP-Tet1$^{\Delta CXXC}$). In C2C12 myoblasts GFP-Tet1 and GFP-Tet1$^{\Delta CXXC}$ showed punctuated nuclear patterns that did not depend on the cell cycle stage (Figure 3A and data not shown). The same constructs were transfected in HEK293T cells and global levels of genomic hmC were measured using a recently described hmC glucosylation assay [39]. Overexpression of GFP-Tet1 and GFP-Tet1$^{\Delta CXXC}$ determined a similar 5-fold increase of genomic hmC levels relative to control samples overexpressing GFP (Figure 3B), indicating that the CXXC domain is not required for enzymatic activity of Tet1 *in vivo*.

Deletion of the CXXC domain does not affect the activity of Dnmt1 *in vitro*

To explore the role of the CXXC domain in Dnmt1 function we generated GFP-Dnmt1 fusion constructs where the CXXC domain, as defined by our homology model, was deleted. We reasoned that precise deletion of the entire structure delimited by the antiparallel β-sheet (Figure 1D) would have the highest chances to preserve native folding of the rest of the protein. We introduced this deletion in GFP fusion constructs encoding either the full length Dnmt1 or the isolated N-terminal region (GFP-Dnmt1$^{\Delta CXXC}$ and GFP-NTR$^{\Delta CXXC}$, respectively; Figure 4A and

Figure S4 in File S1). We then compared DNA binding properties, catalytic activity and interaction between N-terminal region and C-terminal catalytic domain of ΔCXXC and corresponding wild type constructs. Competitive DNA binding assays with the same set of substrates as used for the experiments with GFP-CXXC^{Dnmt1} and GFP-CXXC^{Tet1} reported above (Figure 2C) showed that both GFP-Dnmt1 and GFP-Dnmt1$^{\Delta CXXC}$ bind DNA independently of the presence and methylation state of a CpG site (Figure 4B). As the isolated CXXC domain preferentially bound the substrate containing an unmethylated CpG site, the result with GFP-Dnmt1 and GFP-Dnmt1$^{\Delta CXXC}$ indicates that the CXXC domain contributes negligibly to the DNA binding specificity of the full-length enzyme.

Several groups reported that interaction between the N-terminal region and the C-terminal catalytic domain of Dnmt1 leads to allosteric activation of Dnmt1 [16,18–20,40]. To test whether the CXXC domain is involved in this intramolecular interaction, we co-expressed either GFP-tagged N-terminal region (GFP-NTR) or GFP-NTR$^{\Delta CXXC}$ constructs with a Cherry- and His-tagged C-terminal domain (Ch-CTD-His) in HEK293T cells and performed co-immunoprecipitation experiments. Ch-CTD-His co-precipitated both GFP-NTR and GFP-NTR$^{\Delta CXXC}$, indicating that the CXXC domain is dispensable for the interaction between the N-terminal region and the C-terminal domain of Dnmt1 (Figure 4C).

To investigate whether the CXXC domain is needed for enzymatic activity or substrate recognition, we tested formation of the covalent complex with cytosine and transfer of the methyl group for GFP-Dnmt1 and GFP-Dnmt1$^{\Delta CXXC}$. We first employed an assay to monitor covalent complex formation that exploits the formation of an irreversible covalent bond between the enzyme and the mechanism-based inhibitor 5-aza-2-deoxycytosine (5-aza-dC). This results in permanent trapping of the enzyme by DNA substrates containing 5-aza-dC, as opposed to the reversible complex formed with substrates containing the natural substrate 2-deoxycytosine (dC) [36]. GFP-Dnmt1 and GFP-Dnmt1$^{\Delta CXXC}$ were incubated with fluorescent DNA substrates containing either dC (binding) or 5-aza-dC (trapping) at a single CpG site in direct competition. DNA-protein complexes were then isolated by GFP pulldown and molar DNA/protein ratios were calculated from fluorescence measurements (Figure 4D). Covalent complex
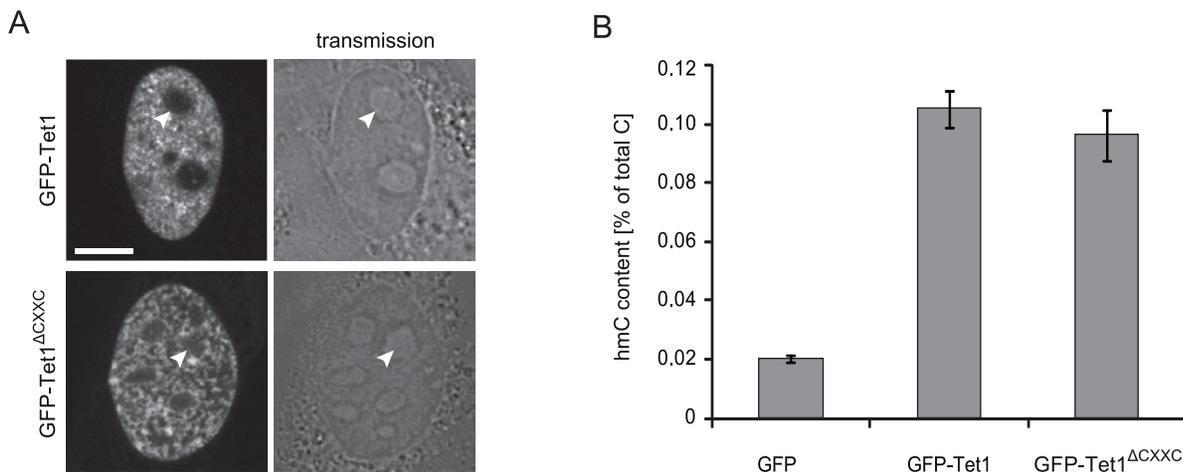
Figure 3. Cellular localization and *in vivo* catalytic activity of GFP-Tet1 and GFP-Tet1$^{\Delta CXXC}$. (A) Live images of C2C12 myoblasts expressing GFP-Tet1. Scale bar: 5 µm. (B) Genomic hmC content in HEK293T cells overexpressing GFP, GFP-Tet1 and GFP-Tet1$^{\Delta CXXC}$. Shown are mean values and standard deviation of hmC percentage over total cytosine for three measurements from one transfection.
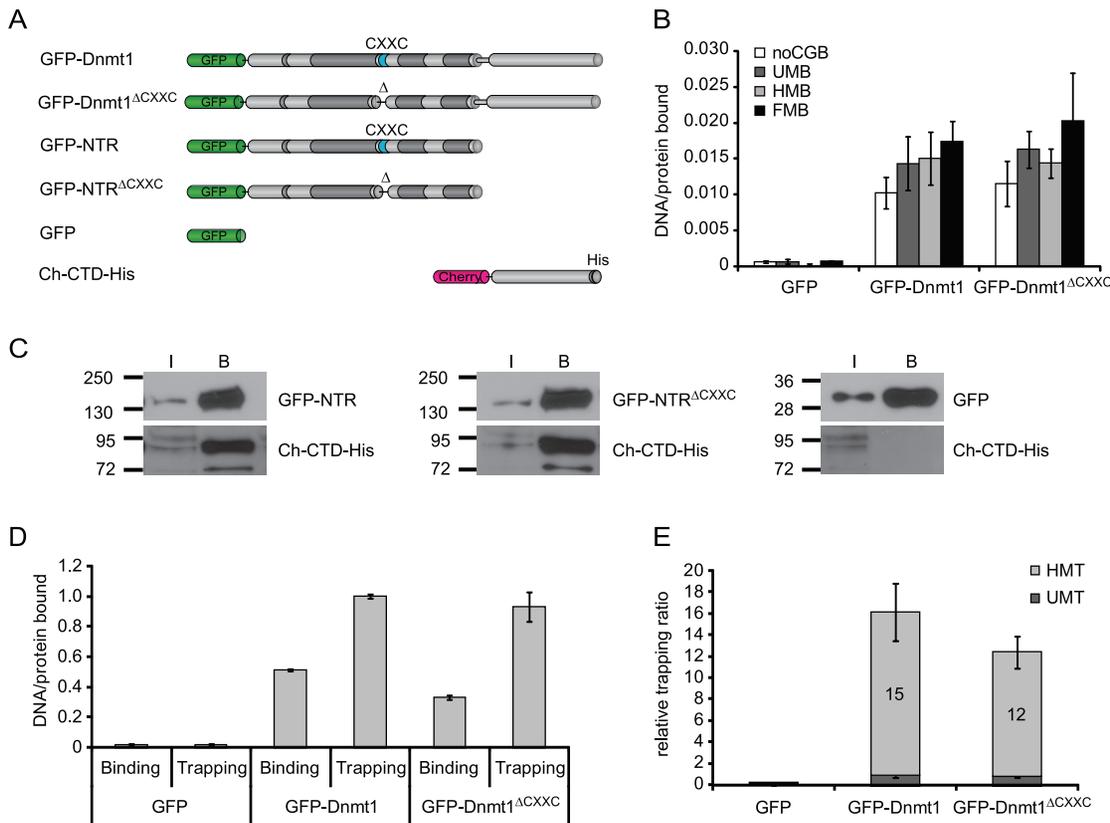doi:10.1371/journal.pone.0016627.g003

Figure 4. DNA binding specificity, intramolecular interaction and trapping of wild-type Dnmt1 and CXXC deletion constructs *in vitro*. (**A**) Schematic representation of Dnmt1 expression constructs. (**B**) DNA binding specificity of GFP-Dnmt1 and GFP-Dnmt1$^{\Delta CXXC}$ were assayed as described in Figure 2C. (**C**) Co-immunoprecipitation of the C-terminal domain of Dnmt1 (Ch-CTD-His) and the N-terminal region with and without deletion of the CXXC domain (GFP-NTR and GFP-NTR$^{\Delta CXXC}$, respectively). GFP fusions were detected using an anti-GFP antibody, while the C-terminal domain construct was detected using an anti-His antibody. GFP was used as negative control. I = input, B = bound. (**D**) Comparison of binding and trapping activities for GFP-Dnmt1 and GFP-Dnmt1$^{\Delta CXXC}$ to monitor irreversible covalent complex formation with hemimethylated substrates. (**E**) Relative covalent complex formation rate of GFP-Dnmt1 and GFP-Dnmt1$^{\Delta CXXC}$ on substrates containing one un- (UMT) or hemi-methylated CpG site (HMT) in direct competition. The trapping ratio for GFP-Dnmt1 on unmethylated substrate was set to 1. In (D) and (E) the means and corresponding standard deviations of triplicate samples from three independent experiments are shown. GFP was used as negative control.
doi:10.1371/journal.pone.0016627.g004

formation was then estimated by comparing trapping and binding activities. GFP-Dnmt1 and GFP-Dnmt1$^{\Delta CXXC}$ showed comparable covalent complex formation rates (trapping/binding ratios), which were about 15- and 12-fold higher for hemi- than unmethylated substrates, respectively (Figure 4E). Together with the data from binding experiments (Fig. 4B), this result indicates that the preference of Dnmt1 for hemimethylated substrates is determined at the covalent complex formation step rather than upon DNA binding. Furthermore, the CXXC domain clearly does not play a major role in determining either the efficiency or the methylation state-specificity of covalent complex formation.

Next, we tested whether deletion of the CXXC domain affects the ability of Dnmt1 to transfer [³H]methyl groups from the donor S-adenosylmethionine (SAM) to a poly(dI·dC)-poly(dI·dC) substrate, a standard DNA methyltransferase activity assay. This showed that *in vitro* GFP-Dnmt1 and GFP-Dnmt1$^{\Delta CXXC}$ are equally active methyltransferases (Figure S7 in File S1). This result is in contrast with a previous report showing that deletion of aa 647–690 in human DNMT1 encompassing the CXXC domain resulted in a drastic loss of catalytic activity [32]. However, according to our homology model the deletion by Pradhan *et al.* would eliminate the predicted N-terminal β-strand (β1 in Figure 1) preventing the formation of the antiparallel β-sheet and potentially distort the

folding of the rest of the protein. This is in contrast with our GFP-Dnmt1$^{\Delta CXXC}$ mutant that was designed to retain the β-sheet structure. To test whether this may account for the observed discrepancy, we generated GFP fusion constructs of wild type human DNMT1 and the same deletion as reported by Pradhan *et al.* and tested covalent complex formation with 5-aza-dC containing DNA substrates as described above. While the human wild type construct showed the same preference for hemimethylated over unmethylated trapping substrates as the mouse constructs, this preference was clearly reduced for the human CXXC deletion mutant (Figure S8 in File S1). This result is consistent with the loss of enzymatic activity shown by Pradhan *et al.* for this mutant and together with the retention of trapping and catalytic activity by the different deletion in GFP-Dnmt1$^{\Delta CXXC}$ suggests that disruption of the antiparallel β-sheet delimiting the CXXC domain results in further distortion and loss of activity of the enzyme.

In conclusion, we showed that, *in vitro*, deletion of the CXXC domain does not affect the interaction between N-terminal region and C-terminal domain, DNA binding, the preference for hemimethylated substrates upon covalent complex formation and the methyltransferase activity of Dnmt1. Together, these data strongly argue against an involvement of the CXXC domain in allosteric activation of Dnmt1.
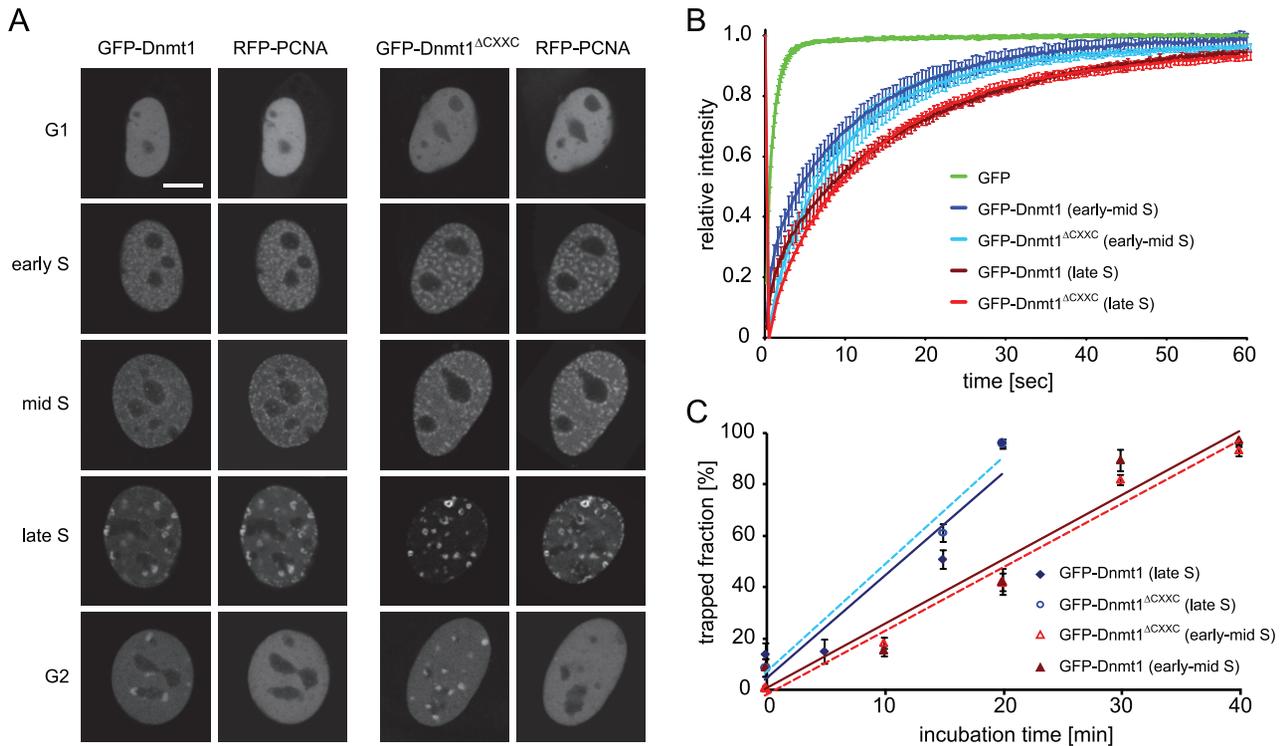
Figure 5. Cell cycle dependant cellular localization, protein mobility and trapping of wild-type Dnmt1 and CXXC deletion constructs in mouse C2C12 myoblasts. (**A**) Cell cycle dependent localization of GFP-Dnmt1 and GFP-Dnmt1$^{\Delta CXXC}$ constructs. Scale bar: 5 µm. (**B**) Analysis of binding kinetics of GFP-Dnmt1 and GFP-Dnmt1$^{\Delta CXXC}$ in early and late S-phase cells by FRAP. The recovery curve for GFP is shown for comparison. (**C**) *In vivo* trapping by FRAP analysis in cells treated with 5-aza-dC. The trapped enzyme fraction is plotted over time for early and late S-phase cells. For each construct three to six cells in early-mid and late S phase were analysed per time point. Shown are mean values ± SEM. In (A–C) RFP-PCNA was cotransfected to identify cell cycle stages in living cells.
doi:10.1371/journal.pone.0016627.g005

Deletion of the CXXC domain does not affect Dnmt1 activity *in vivo*

We then undertook a functional characterization of the GFP-Dnmt1$^{\Delta CXXC}$ construct *in vivo*. We first compared localization and binding kinetics of GFP-Dnmt1 or GFP-Dnmt1$^{\Delta CXXC}$ in mouse C2C12 myoblasts co-transfected with RFP-PCNA, which served as S-phase marker [41]. GFP-Dnmt1$^{\Delta CXXC}$ showed the same cell-cycle dependent nuclear localization pattern as previously shown for GFP-Dnmt1 and endogenous Dnmt1 (Figure 5A)[42,43]. Interaction with PCNA via the PBD directs Dnmt1 to replication foci throughout S-phase. In addition, in late S-phase and G2 Dnmt1 is enriched at chromocenters, clusters of pericentric heterochromatin (PH) that are observed as discrete domains densely stained by DNA dyes in mouse interphase cells. Association of Dnmt1 with PH at these stages is mediated by the TS domain [42]. Thus, the CXXC domain clearly does not contribute to the subnuclear localization of Dnmt1 at this level of resolution.

We also compared the mobility of GFP-Dnmt1 and GFP-Dnmt1$^{\Delta CXXC}$ in living C2C12 myoblasts by FRAP analysis (Figure 5B). These experiments revealed that the kinetics of Dnmt1 is not significantly affected by deletion of the CXXC domain in early-mid as well as late S-phase.

To test covalent complex formation in living cells, we used a previously established trapping assay [44]. Mouse C2C12 myoblasts were co-transfected with RFP-PCNA and either GFP-Dnmt1 or GFP-Dnmt1$^{\Delta CXXC}$ and treated with 5-aza-dC. Immobilization of the Dnmt1 constructs at the site of action was then measured by FRAP analysis (Figure 5C). GFP-Dnmt1 and

GFP-Dnmt1$^{\Delta CXXC}$ showed very similar trapping kinetics, the immobile enzyme fraction reaching nearly 100% after 20 and 40 minutes in early-mid and late S-phase, respectively. This result clearly shows that the CXXC domain is dispensable for covalent complex formation also *in vivo*.

Finally, we compared the ability of GFP-Dnmt1 and GFP-Dnmt1$^{\Delta CXXC}$ to restore DNA methylation patterns in mouse *dnmt1$^{-/-}$* ESCs. Cells transiently expressing either GFP-Dnmt1 or GFP-Dnmt1$^{\Delta CXXC}$ were FACS sorted 48 h after transfection. Isolated genomic DNA was then bisulfite treated and fragments corresponding to major satellite repeats, intracisternal type A particle (IAP) interspersed repeats, *skeletal α-actin* and *H19a* promoters were amplified and subjected to pyrosequencing (Figure 6). As shown previously [43], under these conditions GFP-Dnmt1 partially restored methylation of major satellite and IAP repeats and the *skeletal α-actin* promoter, but not of the imprinted *H19a* promoter, as establishment of the methylation imprint requires passage through the germ line [45]. Methylation patterns of all these sequences in cells expressing GFP-Dnmt1$^{\Delta CXXC}$ were very similar to those in GFP-Dnmt1 expressing cells, including the lack of (re-) methylation at the *H19a* promoter. These results suggest that the CXXC domain is not required for maintenance of DNA methylation patterns by Dnmt1 and does not restrain the DNA methyltransferase activity of Dnmt1 on unmethylated CpG sites. Thus, the CXXC domain does not play a major role in subcellular localization, it does not contribute to the global binding kinetics of Dnmt1 and, consistent with the *in vitro* data reported above, is dispensable for maintaining DNA methylation patterns in living cells.
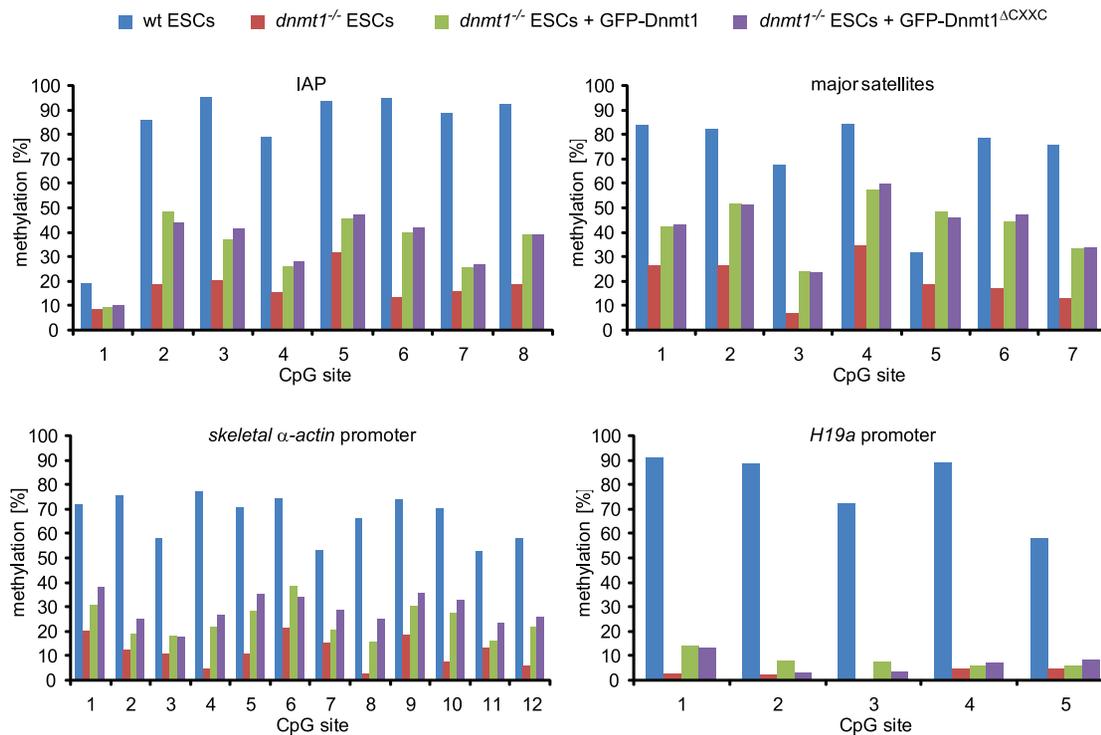
Figure 6. The CXXC deletion construct of Dnmt1 restores methylation in *dnmt1* null cells. Mouse *dnmt1*$^{-/-}$ ESCs transiently expressing GFP-Dnmt1 or GFP-Dnmt1$^{\Delta CXXC}$ were isolated by FACS-sorting 48 h after transfection and CpG methylation levels within the indicated sequences were analyzed by bisulfite treatment, PCR amplification and direct pyrosequencing. Methylation levels of untransfected wild type and *dnmt1*$^{-/-}$ ESCs are shown for comparison.
doi:10.1371/journal.pone.0016627.g006

Discussion

We generated homology models based on the reported structure of the MLL1 CXXC domain to design isolated CXXC domain constructs and CXXC domain deletion mutants for Dnmt1 and Tet1 with minimal probability of structural alteration. According to these models CXXC domains are delimited by an antiparallel β-sheet, a discrete structural element. Our data show that the CXXC domain of mouse Dnmt1 preferentially binds DNA substrates containing unmethylated CpG sites as previously shown for CXXC domains of human DNMT1 and other mammalian proteins. We note that sequences C-terminal to the corresponding peptide in CGBP/Cfp1 were reported to be required for DNA binding *in vitro* [22] and that only a significantly larger peptide spanning the CXXC-3 domain of Mbd1a was tested for DNA binding. However, sequences C-terminal to CXXC domains are not conserved (Figure 1B) and our data show that they are not required for DNA binding by the CXXC domain of Dnmt1. Nevertheless, all the CXXC domains reported to selectively bind unmethylated CpG sites cluster in a distinct homology group and contain the KFGG motif. The latter was shown to be crucial for DNA binding by the CXXC domain of MLL1 [38] and here we extend this observation to the CXXC domain of Dnmt1. Sequence alignment reveals two distinct CXXC domain homology groups that lack the KFGG motif (Figure 1A). Consistent with a role of this motif in DNA binding, members of these groups such as CXXC-1/2 of Mbd1 [29] and the CXXC domain of Tet1 (this study) show no DNA binding activity. While no specific function is known for CXXC-1/2 of Mbd1, the CXXC domain of Tet1 is closely related to those in CXXC4/Idax and CXXC5/RINF that were shown to mediate protein-protein interactions [46–48]. This suggests that the CXXC domain of Tet1, rather than mediating DNA binding, may function as a protein-protein interaction domain. However, our data do not rule out the possibility that the DNA binding properties of the CXXC domain within the context of full length Tet1 may be different from those of the isolated domain. Nevertheless, we show that the CXXC domain is not required for enzymatic activity of Tet1 *in vivo*.

Although we observed a clear DNA binding activity by the isolated CXXC domain of Dnmt1, we found that, within the context of the full length enzyme, this domain is dispensable for overall DNA binding properties, preference for hemimethylated substrates upon covalent complex formation, methyltransferase activity and allosteric activation as well as for the ability to restore methylation of representative sequences in *dnmt1* null ESCs. Consistent with our data, a recent report showed a preference of the CXXC domain of human DNMT1 for substrates containing unmethylated CpG sites [32]. However, the same report showed that deletion of the CXXC domain from the human enzyme results in a significant decrease in methyltransferase activity on hemimethylated substrates *in vitro* and 25% lower methylation at rDNA repeats upon overexpression in HEK293 cells, suggesting a dominant negative effect of the deletion construct. These discrepancies may be due to the fact that the fragment deleted by Pradhan *et al.* includes the N-terminal strand of the predicted antiparallel β-sheet, potentially leading to disruption of native folding, to species-specific differences and/or to the analysis of non-physiological expression levels in HEK293 cells. In our trapping assay the same human deletion mutant showed reduced covalent complex formation, consistent with loss of enzymatic activity. The report from Pradhan *et al.* also showed that mutation of cysteine 667 to glycine within the CXXC domain of human

DNMT1 disrupts DNA binding and enzymatic activity. However, as this point mutation involves one of the zinc coordinating residues it is not unlikely to alter peptide folding with negative consequences potentially extending beyond the CXXC domain and including reduced enzymatic activity. In this respect the dominant negative effect observed upon overexpression of this mutant may be explained by the prevalent occurrence of Dnmt1 as a dimer [49]. These observations, together with preserved ability for covalent complex formation and catalytic activity of our CXXC domain deletion, support the validity of our homology model-driven approach for functional characterization of the CXXC domain. In addition, our genetic complementation approach constitutes a rather physiologic functional assay. However, due to the transient approach and the analysis of genomic methylation at only a few representative sequences, subtle or highly sequence specific effects of deletion of the CXXC domain cannot be excluded.

It was recently shown that binding of Cfp1/CGBP and KDM2A to CpG islands through their CXXC domains leads to local enrichment and depletion of H3K4 and H3K36 methylation, respectively [26,30]. Analogously, Dnmt1 may bind CpG islands through its CXXC domain. However, this interaction would not lead to a straightforward functional interpretation as CpG islands with high CpG density are generally refractive to DNA methylation and a function of Dnmt1 as a *de novo* DNA methyltransferase is not well established. It could be envisaged that binding to unmethylated CpG sites/islands by the CXXC domain may have a negative effect on the enzymatic activity of Dnmt1 and restrain its function as a *de novo* DNA methyltransferase. However, we show that in *dnmt1* null ESCs methylation of the imprinted *H19a* promoter is not restored upon expression of either wild type or ΔCXXC Dnmt1 constructs, arguing against a negative regulatory function of the CXXC domain.

Notably, binding of unmethylated CpG sites by KFGG motif-containing CXXC domains does not exclude a role in protein-protein interaction as the CXXC domain of MLL1 was reported to interact with both DNA and Polycomb Repressive Complex 1 components HPC2/CBX4 and BMI-1 [21,50]. Therefore, it is possible that the CXXC domain of Dnmt1 has regulatory functions in specific cell types or developmental stages that may involve DNA binding and/or interaction with other proteins. The generation of dedicated animal models may be instrumental for testing these possibilities.

Materials and Methods

Bioinformatic methods

Alignments were performed using the ClustalW2 software [51]. The CXXC domain homology tree (Figure 1C) was generated from the alignment in Figure 1B with Jalview 2.4 by unweighted pair group method with arithmetic mean (UPGMA). The neighbor-joining method gave the same result. Average distances between the sequences were calculated using the BLOSSUM62 matrix. The human CXXC10 coding sequence [52] was determined by assembling ESTs AI438961, BX114363, BX492895, BU633058.1, AW207644.1 and the genomic sequence AC073046.7. The putative translational start site is located 16308 bp upstream of the annotated transcriptional start site of *TET3*. A partial coding sequence of murine Cxxc10 containing the CXXC domain was identified by aligning the human CXXC10 protein sequence to the ORFs present in NT_039353.7 upstream of the *tet3* gene from position 35663306 to 35808487. A very high match was found 13266 nt upstream of *tet3* at positions 35676374-35676572 of NT_039353.7. To build

homology models for the CXXC domains of Dnmt1 (aa 645–696) and Tet1 (aa 561–614), we submitted the respective sequences to the HHpred server [53]. The best template was the CXXC domain of MLL1 (PDB-ID: 2J2S). The 49 residues of the CXXC domain in Dnmt1 can be aligned to this domain with 45% sequence identity and only a single amino acid gap after residue 661 (Figure 1B). 3D models were calculated with the homology modeling software MODELLER [54] (version 9.5) using this alignment. Distance restraints were given to MODELLER to enforce a distance of 2.3 ± 0.1 Å between the eight sulphurs in the Zn-coordinating cysteines and the Zn^{2+} ions. TM-align [55] was used to superpose the model structure with the template domain. Images were generated using the PyMol Molecular Graphics System (Version 1.3, Schrödinger, LLC). The quality of the models and the underlying alignments were checked with DOPE [56] and Verify3D [57] and results for both models were found to be comparable to the MLL1 template structure (2J2S).

Expression constructs

Fusion constructs were generated using enhanced green fluorescent protein, monomeric red fluorescent protein or monomeric cherry and are here referred to as GFP, RFP and Cherry fusions, respectively. Mammalian expression constructs for GFP, mouse GFP-Dnmt1, GFP-NTR and human RFP-PCNA were described previously [42,44,49,58]. The deletion construct GFP-Dnmt1^{ΔCXXC} was obtained by replacing the sequence coding for aa 655–696 with three alanine codons in the GFP-Dnmt1 construct as described [59]. The GFP-DNMT1^{ΔCXXC} construct was generated by subcloning the sequence coding for human DNMT1^{ΔCXXC} from the homonymous construct by Pradhan *et al.* [32] in the pEGFP-C2 vector (Clonetech). To generate GFP-Tet1 three partially overlapping fragments spanning the Tet1 coding sequence were amplified using E14 ESCs cDNA as template. The fragments were then joined by overlap extension PCR and inserted into the pCAG-GFP-IB vector [43]. To generate GFP-Tet1^{ΔCXXC} aa 569-621 of murine Tet1 were deleted from GFP-Tet1 using a type IIs restriction endonuclease approach as described [60]. To generate GFP-CXXC^{Dnmt1} and GFP-CXXC^{Tet1} sequences coding for the respective CXXC domains (aa 643-700 for Dnmt1 and 561-614 for Tet1) were amplified by PCR using the GFP-Dnmt1 expression construct and cDNA from E14 ESCs as templates, respectively. PCR fragments were then inserted into the pCAG-GFP-IB vector. GFP-NTR^{ΔCXXC} was obtained by replacing the BglII-XhoI fragment of GFP-NTR with the same fragment of GFP-Dnmt1^{ΔCXXC}. Ch-CTD-His was generated by replacing the GFP coding sequence in a GFP-CTD construct [49] with the Cherry coding sequence. All constructs were confirmed by sequencing.

Cell culture, transfection and cell sorting

HEK293T cells [61] and mouse C2C12 myoblasts [62] were cultured in DMEM supplemented with 50 µg/ml gentamicin and 10% and 20% fetal calf serum, respectively. For expression of fusion proteins HEK293T cells were transfected with polyethylenimine (Sigma). For live cell imaging, C2C12 cells were grown to 40% confluence on Lab-Tek chambers (Nunc) or µ-slides (Ibidi) and transfected with TransFectin transfection reagent (BioRad) according to the manufacturer's instructions. Mouse ESCs were cultured as described [63] and transfected with FuGENE HD (Roche) according to the manufacturer's instructions. ESCs were sorted with a FACS Aria II instrument (Becton Dickinson). The *dnmt1*$^{-/-}$ J1 ESCs used in this study are homozygous for the c allele [14].

In vitro DNA binding and trapping assays

In vitro DNA binding and trapping assays were performed as described previously [36,37] with the following modifications. DNA substrates labeled with four different ATTO fluorophores (Tables S1 and S2 in File S1) were used at a final concentration of 125 nM each in the pull-down assay with immobilized GFP fusions. After removal of unbound substrate, the amounts of protein and DNA were determined by fluorescence intensity measurements with a Tecan Infinite M1000 plate reader using calibration curves from purified GFP or DNA coupled ATTO fluorophores, respectively. The following excitation/emission ± detection bandwidth settings were used: 490/511±10 nm for GFP, 550/580±15 nm for ATTO550, 600/630±15 nm for ATTO590, 650/670±10 nm for ATTO647N and 700/720±10 nm for ATTO700. Cross detection of GFP and different ATTO dyes was negligible with these settings. Binding and trapping ratios were calculated dividing the concentration of bound DNA substrate by the concentration of GFP fusion on the beads.

In vivo mC hydroxylation assay

Genomic DNA was isolated from HEK293T cells 24 h after transfection with the GFP-Tet1 and GFP-Tet1$^{\Delta CXXC}$ constructs and global hmC levels were measured using the *in vitro* glucosylation assay as previously described [63], except that 100 nM β-glucosyltransferase and only UDP-[^{3}H]glucose donor (0.43 μM) were used.

Co-immunoprecipitation

Co-immunoprecipitation was performed as described previously [49,64]. Shortly, HEK293T cells were transiently co-transfected with expression plasmids for GFP fusions and the Ch-CTD-His construct, harvested and lysed. GFP fusions were pulled down using the GFP-Trap [65] (Chromotek) and subjected to western blotting using anti-GFP (Roche or Chromotek) and anti-His (Invitrogen) monoclonal antibodies.

Live cell microscopy, FRAP analysis and live cell trapping assay

Live cell imaging and FRAP experiments were performed as described previously [43]. For each construct 6-15 nuclei were averaged and the mean values as well as the standard errors were calculated. For presentation, we used linear contrast enhancement on entire images. The DNA methyltransferase trapping assay was described previously [44]. Briefly, transfected cells were incubated with 30 μM 5-aza-dC (Sigma) for the indicated periods of time before photobleaching experiments. FRAP analysis was performed with a confocal laser scanning microscope (TCS SP5, Leica)

equipped with a 63×/1.4 NA Plan-Apochromat oil immersion objective. Microscope settings were as described except that a smaller region of interest (3 μm×3 μm) was selected for photobleaching. Mean fluorescence intensities of the bleached region were corrected for background and for total loss of nuclear fluorescence over the time course, and normalized by the mean of the last 10 prebleach values.

DNA Methylation Analysis

Genomic DNA was isolated with the QIAmp DNA Mini Kit (Qiagen) and 1.5 μg were bisulfite converted using the EZ DNA Methylation-Gold Kit (Zymo research) according to the manufacturer's instructions. Primer sets and PCR conditions for IAP-LTR, *skeletal α-actin* and *H19* promoters were as described [43]. Primer sequences for major satellites were AAAATGAGAAA-CATCCACTTG (forward primer) and CCATGATTTT-CAGTTTTCTT (reverse primer). For amplification we used Qiagen Hot Start Polymerase in 1x Qiagen Hot Start Polymerase buffer supplemented with 0.2 mM dNTPs, 0.2 μM forward primer, 0.2 μM reverse primer, 1.3 mM betaine (Sigma) and 60 mM tetramethylammonium-chloride (TMAC, Sigma). Promoter regions and IAP-LTR were amplified with two subsequent (nested) PCR reactions and major satellite repeats were amplified with a single amplification reaction. Pyrosequencing reactions were carried out by Varionostic GmbH (Ulm, Germany). Pyrosequencing primers are listed in Table S3 in File S1.

Supporting Information

File S1   Tables S1–S3, Figures S1–S8 and Supplemental methods.
(PDF)

References

1. Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, et al. (2009) Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol 27: 361–368.
2. Suzuki MM, Bird A (2008) DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 9: 465–476.
3. Laurent L, Wong E, Li G, Huynh T, Tsirigos A, et al. (2010) Dynamic changes in the human methylome during differentiation. Genome Res 320: 320–331.
4. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, et al. (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462: 315–322.
5. Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, et al. (2008) Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. Mol Cell 30: 755–766.
6. Deng J, Shoemaker R, Xie B, Gore A, LeProust EM, et al. (2009) Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. Nat Biotechnol 27: 353–360.
7. Schmidl C, Klug M, Boeld TJ, Andreesen R, Hoffmann P, et al. (2009) Lineage-specific DNA methylation in T cells correlates with histone methylation and enhancer activity. Genome Res 19: 1165–1174.
8. Edwards JR, O'Donnell AH, Rollins RA, Peckham HE, Lee C, et al. (2010) Chromatin and sequence features that define the fine and gross structure of genomic methylation patterns. Genome Res 20: 972–980.
9. Gaudet F, Hodgson JG, Eden A, Jackson-Grusby L, Dausman J, et al. (2003) Induction of tumors in mice by genomic hypomethylation. Science 300: 489–492.
10. Walsh CP, Chaillet JR, Bestor TH (1998) Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet 20: 116–117.
11. Xu G-L, Bestor TH, Bourc'his D, Hsieh C-L, Tommerup N, et al. (1999) Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. Nature 402: 187–191.
12. Goll MG, Bestor TH (2005) Eukaryotic cytosine methyltransferases. Annu Rev Biochem 74: 481–514.

13. Leonhardt H, Page AW, Weier HU, Bestor TH (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell 71: 865–873.

14. Lei H, Oh SP, Okano M, Juttermann R, Goss KA, et al. (1996) De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development 122: 3195–3205.

15. Li E, Bestor TH, Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 69: 915–926.

16. Margot JB, Aguirre-Arteta AM, Di Giacco BV, Pradhan S, Roberts RJ, et al. (2000) Structure and function of the mouse DNA methyltransferase gene: Dnmt1 shows a tripartite structure. J Mol Biol 297: 293–300.

17. Spada F, Rothbauer U, Zolghadr K, Schermelleh L, Leonhardt H (2006) Regulation of DNA methyltransferase 1. Adv Enzyme Regul 46: 224–234.

18. Zimmermann C, Guhl E, Graessmann A (1997) Mouse DNA methyltransferase (MTase) deletion mutants that retain the catalytic domain display neither de novo nor maintenance methylation activity in vivo. Biol Chem 378: 393–405.

19. Pradhan S, Esteve PO (2003) Allosteric activator domain of maintenance human DNA (cytosine-5) methyltransferase and its role in methylation spreading. Biochemistry 42: 5321–5332.

20. Fatemi M, Hermann A, Pradhan S, Jeltsch A (2001) The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA. J Mol Biol 309: 1189–1199.

21. Birke M, Schreiner S, Garcia-Cuellar MP, Mahr K, Titgemeyer F, et al. (2002) The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. Nucl Acids Res 30: 958–965.

22. Lee JH, Voo KS, Skalnik DG (2001) Identification and characterization of the DNA binding domain of CpG-binding protein. J Biol Chem 276: 44669–44676.

23. Cross SH, Meehan RR, Nan X, Bird A (1997) A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. Nat Genet 16: 256–259.

24. Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, et al. (2006) Histone demethylation by a family of JmjC domain-containing proteins. Nature 439: 811–816.

25. Frescas D, Guardavaccaro D, Bassermann F, Koyama-Nasu R, Pagano M (2007) JHDM1B/FBXL10 is a nucleolar protein that represses transcription of ribosomal RNA genes. Nature 450: 309–313.

26. Blackledge NP, Zhou JC, Tolstorukov MY, Farcas AM, Park PJ, et al. (2010) CpG Islands Recruit a Histone H3 Lysine 36 Demethylase. Mol Cell 38: 179–190.

27. Lorsbach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, et al. (2003) TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). Leukemia 17: 637–641.

28. Ono R, Taki T, Taketani T, Taniwaki M, Kobayashi H, et al. (2002) LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). Cancer Res 62: 4075–4080.

29. Jorgensen HF, Ben-Porath I, Bird AP (2004) Mbd1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains. Mol Cell Biol 24: 3387–3395.

30. Thomson JP, Skene PJ, Selfridge J, Clouaire T, Guy J, et al. (2010) CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature 464: 1082–1086.

31. Bestor TH (1992) Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. EMBO J 11: 2611–2617.

32. Pradhan M, Esteve PO, Chin HG, Samaranayke M, Kim GD, et al. (2008) CXXC domain of human DNMT1 is essential for enzymatic activity. Biochemistry 47: 10000–10009.

33. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, et al. (2009) Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. Science 324: 930–935.

34. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, et al. (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature 466: 1129–1133.

35. Allen MD, Grummitt CG, Hilcenko C, Min SY, Tonkin LM, et al. (2006) Solution structure of the nonmethyl-CpG-binding CXXC domain of the leukaemia-associated MLL histone methyltransferase. EMBO J 25: 4503–4512.

36. Frauer C, Leonhardt H (2009) A versatile non-radioactive assay for DNA methyltransferase activity and DNA binding. Nucl Acids Res 37: e22.

37. Rottach A, Frauer C, Pichler G, Bonapace IM, Spada F, et al. (2010) The multi-domain protein Np95 connects DNA methylation and histone modification. Nucl Acids Res 38: 1796–1805.

38. Ayton PM, Chen EH, Cleary ML (2004) Binding to nonmethylated CpG DNA is essential for target recognition, transactivation, and myeloid transformation by an MLL oncoprotein. Mol Cell Biol 24: 10470–10478.

39. Szwagierczak A, Bultmann S, Schmidt CS, Spada F, Leonhardt H (2010) Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. Nucleic Acids Research 38: e181.

40. Bacolla A, Pradhan S, Larson JE, Roberts RJ, Wells RD (2001) Recombinant human DNA (cytosine-5) methyltransferase. III. Allosteric control, reaction order, and influence of plasmid topology and triplet repeat length on methylation of the fragile X CGG.CCG sequence. J Biol Chem 276: 18605–18613.

41. Easwaran HP, Leonhardt H, Cardoso MC (2005) Cell Cycle Markers for Live Cell Analyses. Cell Cycle 4: 453–455.

42. Easwaran HP, Schermelleh L, Leonhardt H, Cardoso MC (2004) Replication-independent chromatin loading of Dnmt1 during G2 and M phases. EMBO Rep 5: 1181–1186.

43. Schermelleh L, Haemmer A, Spada F, Rosing N, Meilinger D, et al. (2007) Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. Nucl Acids Res 35: 4301–4312.

44. Schermelleh L, Spada F, Easwaran HP, Zolghadr K, Margot JB, et al. (2005) Trapped in action: direct visualization of DNA methyltransferase activity in living cells. Nat Methods 2: 751–756.

45. Tucker KL, Beard C, Dausmann J, Jackson-Grusby L, Laird PW, et al. (1996) Germ-line passage is required for establishment of methylation and expression patterns of imprinted but not of nonimprinted genes. Genes Dev 10: 1008–1020.

46. Andersson T, Södersten E, Duckworth JK, Cascante A, Fritz N, et al. (2009) CXXC5 Is a Novel BMP4-regulated Modulator of Wnt Signaling in Neural Stem Cells. J Biol Chem 284: 3672–3681.

47. Hino S-i, Kishida S, Michiue T, Fukui A, Sakamoto I, et al. (2001) Inhibition of the Wnt Signaling Pathway by Idax, a Novel Dvl-Binding Protein. Mol Cell Biol 21: 330–342.

48. London TBC, Lee H-J, Shao Y, Zheng J (2004) Interaction between the internal motif KTXXXI of Idax and mDvl PDZ domain. Biochem Biophys Res Commun 322: 326–332.

49. Fellinger K, Rothbauer U, Felle M, Langst G, Leonhardt H (2009) Dimerization of DNA methyltransferase 1 is mediated by its regulatory domain. J Cell Biochem 106: 521–528.

50. Xia ZB, Anderson M, Diaz MO, Zeleznik-Le NJ (2003) MLL repression domain interacts with histone deacetylases, the polycomb group proteins HPC2 and BMI-1, and the corepressor C-terminal-binding protein. Proc Natl Acad Sci U S A 100: 8342–8347.

51. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res 22: 4673–4680.

52. Katoh M (2004) Identification and characterization of human CXXC10 gene in silico. Int J Oncol 25: 1193–1199.

53. Söding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. Nucl Acids Res 33: W244–248.

54. Sali A, Potterton L, Yuan F, van Vlijmen H, Karplus M (1995) Evaluation of comparative protein modeling by MODELLER. Proteins 23: 318–326.

55. Zhang Y, Skolnick J (2005) TM-align: a protein structure alignment algorithm based on the TM-score. Nucl Acids Res 33: 2302–2309.

56. Shen M-y, Sali A (2006) Statistical potential for assessment and prediction of protein structures. Protein Sci 15: 2507–2524.

57. Eisenberg D, Lüthy R, Bowie JU (1997) VERIFY3D: Assessment of protein models with three-dimensional profiles. In: Carter CWJ, Sweet RM, eds. Methods Enzymol: Academic Press. pp 396–404.

58. Sporbert A, Domaing P, Leonhardt H, Cardoso MC (2005) PCNA acts as a stationary loading platform for transiently interacting Okazaki fragment maturation proteins. Nucl Acids Res 33: 3521–3528.

59. Fellinger K, Leonhardt H, Spada F (2008) A mutagenesis strategy combining systematic alanine scanning with larger mutations to study protein interactions. Anal Biochem 373: 176–178.

60. Ko J-K, Ma J (2005) A rapid and efficient PCR-based mutagenesis method applicable to cell physiology study. Am J Physiol Cell Physiol 288: C1273–1278.

61. DuBridge RB, Tang P, Hsia HC, Leong PM, Miller JH, et al. (1987) Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. Mol Cell Biol 7: 379–387.

62. Blau HM, Pavlath GK, Hardeman EC, Chiu CP, Silberstein L, et al. (1985) Plasticity of the differentiated state. Science 230: 758–766.

63. Szwagierczak A, Bultmann S, Schmidt CS, Spada F, Leonhardt H (2010) Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. Nucl Acids Res 38: e181.

64. Meilinger D, Fellinger K, Bultmann S, Rothbauer U, Bonapace IM, et al. (2009) Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells. EMBO Rep 10: 1259–1264.

65. Rothbauer U, Zolghadr K, Muyldermans S, Schepers A, Cardoso MC, et al. (2007) A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. Mol Cell Proteomics 7: 282–289.

66. Mohseni-Zadeh S, Brezellec P, Risler JL (2004) Cluster-C, an algorithm for the large-scale clustering of protein sequences based on the extraction of maximal cliques. Comput Biol Chem 28: 211–218.

Different binding properties and function of CXXC zinc finger domains in Dnmt1 and Tet1

Carina Frauer[#], Andrea Rottach[#], Daniela Meilinger, Sebastian Bultmann, Karin Fellinger, Stefan Hasenöder, Mengxi Wang, Weihua Qin, Johannes Söding[§], Fabio Spada* and Heinrich Leonhardt*

Department of Biology and Center for Integrated Protein Science Munich (CIPS[M]), Ludwig Maximilians University Munich, Planegg-Martinsried and [§]Gene Center Munich, Ludwig Maximilians University Munich, Munich, Germany.

*Corresponding authors:
FS: f.spada@lmu.de; Fax +49 89 2180-74236; Tel. +49 89 2180-74230
HL: h.leonhardt@lmu.de; Fax +49 89 2180-74236; Tel. +49 89 2180-74232

[#]These authors contributed equally to this work

SUPPLEMENTAL INFORMATION FILE

Tables S1-3

Figures S1-8

Supplemental Methods

Table S1. Sequences of DNA oligonucleotides used for preparation of double stranded DNA substrates. M: 5-methylcytosine.

Name	Sequence
CG-up	5'- CTCAACAACTAACTACCATCCGGACCAGAAGAGTCATCATGG -3'
MG-up	5'- CTCAACAACTAACTACCATCMGGACCAGAAGAGTCATCATGG -3'
noCG-up	5'- CTCAACAACTAACTACCATCTGGACCAGAAGAGTCATCATGG -3'
Fill-In-550	5'- ATTO550-CCATGATGACTCTTCTGGTC -3'
Fill-In-590	5'- ATTO590-CCATGATGACTCTTCTGGTC -3'
Fill-In-647N	5'- ATTO647N-CCATGATGACTCTTCTGGTC -3'
Fill-In-700	5'- ATTO700-CCATGATGACTCTTCTGGTC -3'

Table S2. DNA substrates used for the *in vitro* DNA binding and trapping assays.

Name	CpG site	Label	Oligo I	Oligo II	dCTP reaction	Purpose
noCGB 700	no CpG site	700	noCG-up	Fill-In-700	dCTP	Binding
UMB 550	unmethylated	550	CG-up	Fill-In-550	dCTP	Binding
UMB 590		590		Fill-In-590		
UMB 647N		647N		Fill-In-647N		
UMB 700		700		Fill-In-700		
UMT 550		550		Fill-In-550	5-aza-dCTP	Trapping
HMB 590	hemimethylated	590	MG-up	Fill-In-590	dCTP	Binding
HMB 647N		647N		Fill-In-647N		
HMT 550		550		Fill-In-550	5-aza-dCTP	Trapping
HMT 647N		647N		Fill-In-647N		
FMB 647N	fully methylated	647N	MG-up	Fill-In-647N	5methyl dCTP	Binding

Table S3. Primers used for pyrosequencing. Each primer is biotinylated at the 5' end.

Name	Sequence
skeletal α-actin-1	5'- AGTTGGGGATATTTTTTATA -3'
skeletal α-actin-1b	5'- TTTTGGTTAGTGTAGGAGAT -3'
skeletal α-actin-2	5'- TGGGAAGGGTAGTAATATTT -3'
H19-1	5'- ATAGTTATTGTTTATAGTTT -3'
H19-2	5'- AGGAATATGTTATATTTAT -3'
IAP LTR-1	5'- CCCTAATTAACTACAACCCA -3'
IAP LTR-2	5'- TGTAGTTAATTAGGGAGTGA -3'
Major Satellite-1	5'- AAAATGAGAAATATTTATTTG -3'
Major Satellite-2	5'- GAGAAATATATACTTTAGGA -3'
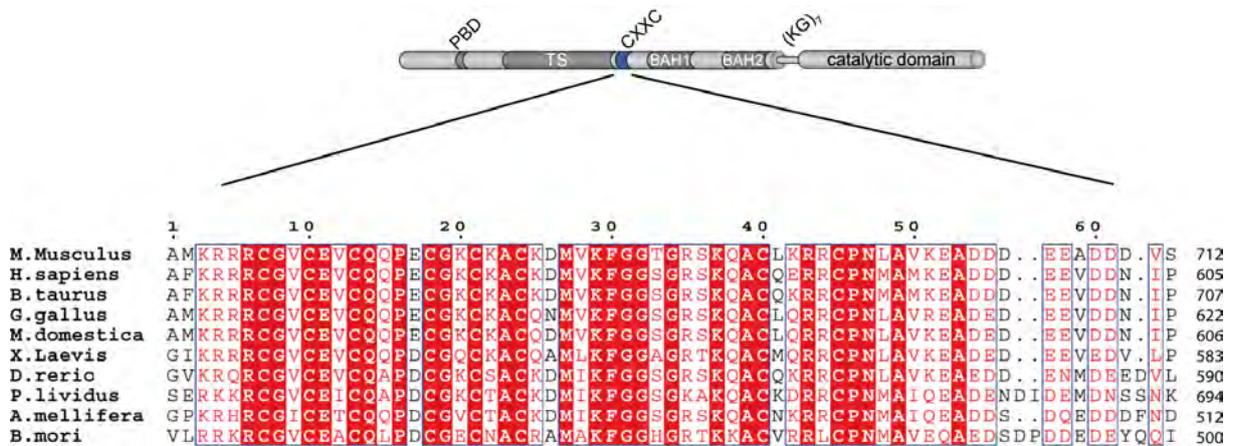
Figure S1. Dnmt1 domain structure and alignment of Dnmt1 CXXC domains from different species. Numbers on the right side indicate the position of the last amino acid in each sequence. PBD: PCNA binding domain; TS: targeting sequence; CXXC: CXXC-type zinc finger domain; BAH1 and 2: bromo-adjacent homology domain; (KG)7: seven lysine-glycine repeats. Absolutely conserved residues are highlighted in red. Positions with residues in red face share 70% similarity as calculated with the Risler algorithm {Mohseni-Zadeh, 2004 #133}. The alignment was generated with ClustalW2 and displayed with ESPript 2.2. GenBank accession numbers are: *Mus musculus*: NP_034196; *Homo sapiens*: NP_001124295; *Bos taurus*: NP_872592; *Monodelphis domestica*: NP_001028141; *Gallus gallus*: NP_996835*; Xenopus laevis*: NP_001084021; *Danio rerio*: NP_571264; *Paracentrotus lividus*: Q27746 (Swiss Prot); *Apis mellifera*: NP_001164522 (Dnmt1a); *Bombyx mori*: NP_001036980.

Figure S2. The cellular localization of GFP-CXXC^{Dnmt1} is independent of cell cycle stage. Live images of C2C12 mouse myoblasts cotransfected with expression constructs for GFP-CXXC^{Dnmt1} and RFP PCNA. The latter served for identification of the cell cycle stage. Scale bar: 5 µm.
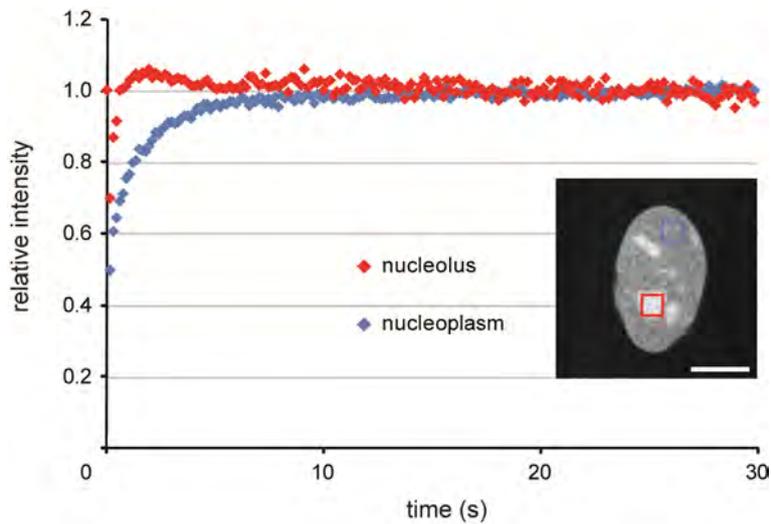
Figure S3. Differential mobility of GFP-CXXC[Dnmt1] in nucleoli and nucleoplasm of mouse C2C12 myoblasts measured by FRAP analysis. Identical regions of interest over the nucleoplasm or nucleoli (as exemplified in the inset) were bleached and recovery curves were recorded over 30 seconds. GFP-CXXC[Dnmt1] kinetics is faster in nucleoli than in the nucleus, which indicates more transient (possibly unspecific) binding in the former than in the latter. Scale bar: 5 µm.
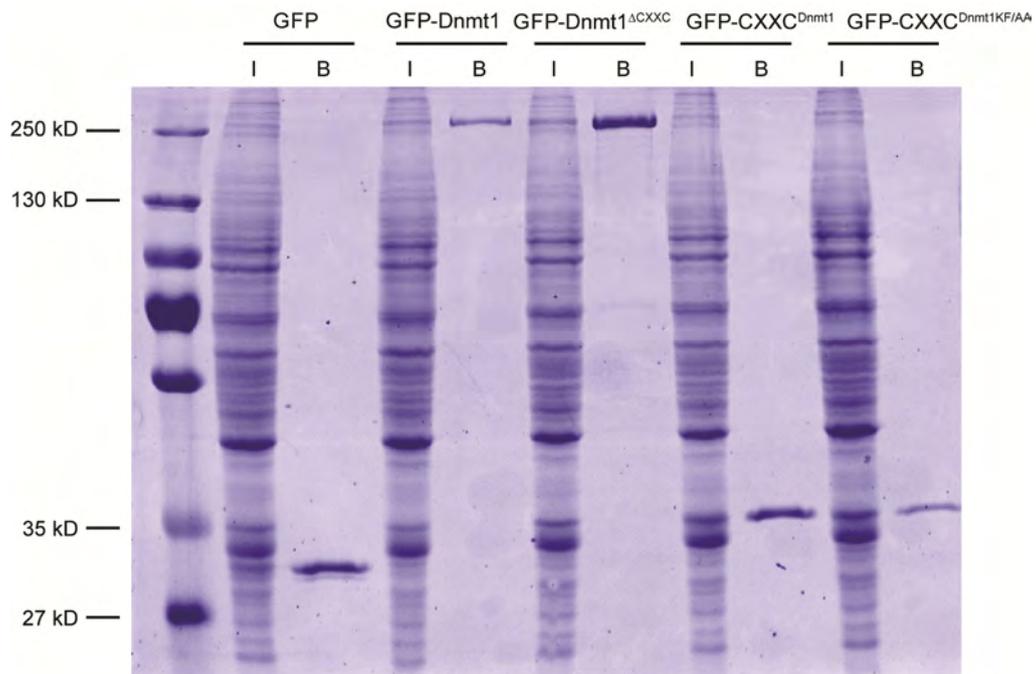
Figure S4. GFP fusion pulldowns from transiently transfected HEK293T cells using the GFP-trap. Shown is a SDS polyacrylamide gel stained with coomassie blue. I = input (1%); B = bound (10%).
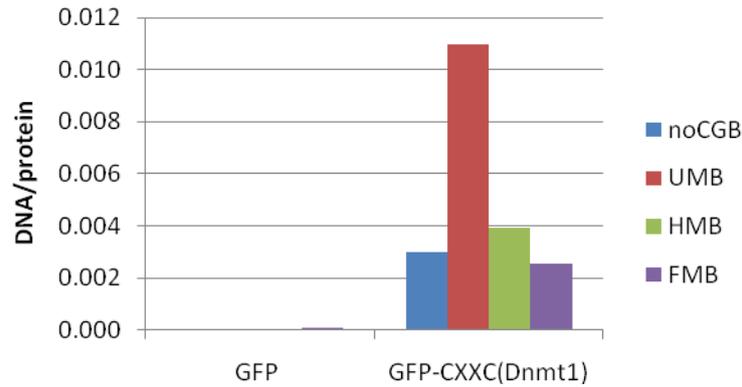
Figure S5. The CXXC domain of Dnmt1 preferentially binds unmethylated CpG sites. GFP and GFP-CXXC^{Dnmt1} purified from transiently transfected HEK293T cells with the GFP trap were challenged with fluorescent DNA substrates containing no CpG site or one central un-, hemi- or fully methylated CpG site in direct competition (noCGB, UMB, HMB and FMB, respectively) as in Figure 2C, except that a five-fold higher concentration (625 nM) of each substrate was used.
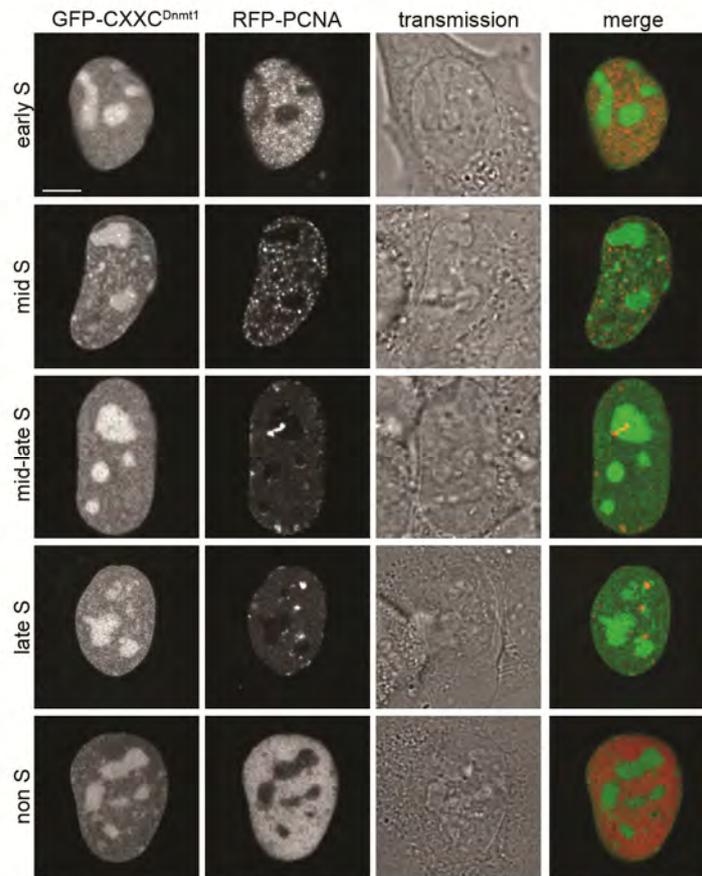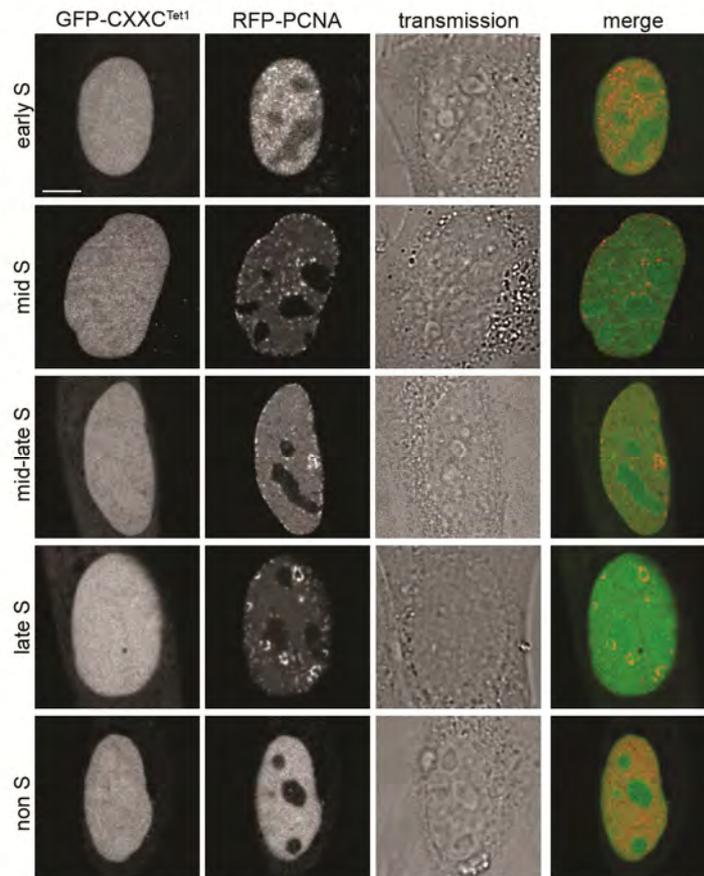
Figure S6. The cellular localization of GFP-CXXC^{Tet1} is independent of cell cycle stage. Live images of C2C12 mouse myoblasts cotransfected with expression constructs for GFP-CXXC^{Tet1} and RFP PCNA. The latter served for identification of the cell cycle stage. Scale bar: 5 µm.
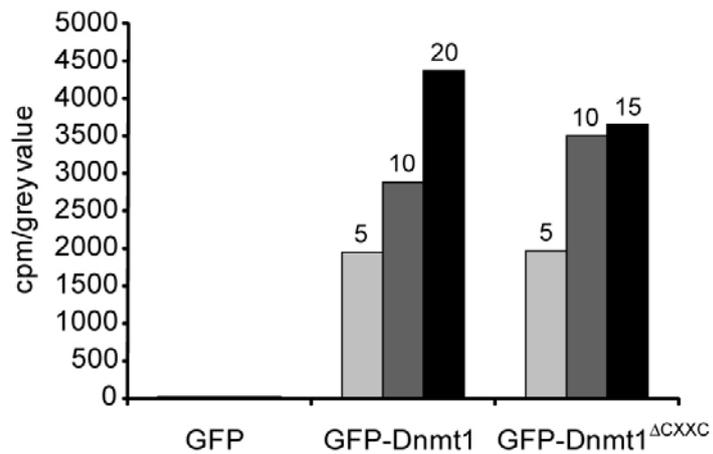
Figure S7. Radioactive methyltransferase activity assay for GFP Dnmt1 and GFP-Dnmt1$^{\Delta CXXC}$. The transfer of [3H]-methyl groups to poly(dI•dC)-poly(dI•dC) substrate was measured for increasing volumes of GFP fusion proteins immunopurified from transiently transfected HEK293T cells. Counts per minute (cpm) were normalized to the relative protein concentration as determined by SDS-PAGE analysis. GFP was used as negative control. Numbers above the bars indicate the volume (µl) of protein solution added.
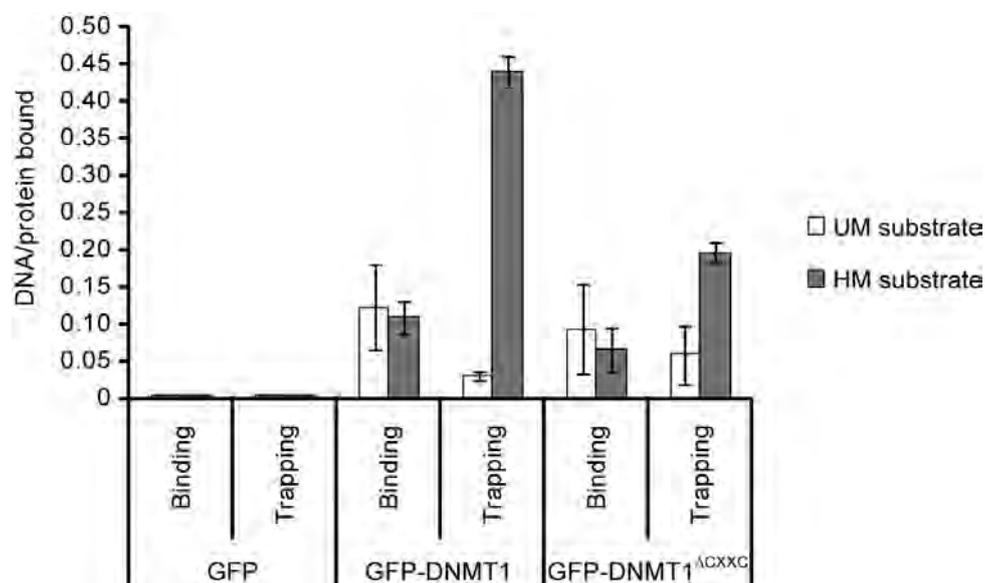
Figure S8. Competitive DNA binding and trapping assays for human GFP-DNMT1 and GFP-DNMT1$^{\Delta CXXC}$. GFP, GFP-DNMT1 and GFP-DNMT1$^{\Delta CXXC}$ were purified from transfected HEK293T cells using the GFP-trap and incubated with fluorescent DNA substrates containing one central unmethylated (UM) or hemimethylated (HM) CpG site in direct competition. Both substrates contained either dC (binding) or 5 aza dC (trapping) on the strand opposite to the differentially methylated one. The comparison of binding and trapping ratios reflects irreversible covalent complex formation. Note the reduction in trapping of GFP-DNMT1$^{\Delta CXXC}$ relative to GFP-DNMT1 by the hemimethylated substrate. Shown are mean values and standard deviation of DNA/protein ratios from two independent experiments.

SUPPLEMENTARY METHODS

In vitro methyltransferase activity assay

Eight milligrams of His-tagged GFP-binding protein (GBP; Chromotek) were coupled to 1ml Ni-NTA agarose beads (Qiagen) by incubating for 2 h at 4°C in PBS and unbound protein was washed out twice with PBS. Extracts of HEK293T cells expressing GFP or a GFP fusions were prepared in 200 µl lysis buffer II (50 mM NaH_2PO_4 pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.5 % Tween-20, 2 mM $MgCl_2$, 1 mg/ml DNaseI, 2 mM PMSF, 1X mammalian protease inhibitor mix). After centrifugation, supernatants were diluted to 500 µl with immunoprecipitation buffer II (50 mM NaH_2PO_4 pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.05 % Tween-20) and precleared by incubation with 25 µl of equilibrated Ni-NTA agarose beads for 30 min at 4°C followed by centrifugation. Precleared extracts were then incubated with 40 µg of His-tagged GFP-trap coupled to Ni-NTA beads for 2 h at 4°C with constant mixing. GFP or GFP fusions were pulled down by centrifugation at 540 *g*. After washing twice with wash buffer II (50 mM NaH_2PO_4 pH 8.0, 300 mM NaCl, 20 mM imidazole, 0.05 % Tween-20), complexes were eluted with 60 µl of elution buffer (10 mM Tris pH 7.5, 100 mM KCl, 1 mM EDTA, 1 mM DTT, 250 mM imidazole) for 10 min at 25°C with constant mixing. 10 µl aliquots of all eluates were subjected to western blot analysis using mouse or rat monoclonal antibodies to GFP (Roche and Chromotek, respectively) and quantified by densitometry. Indicated volumes of eluate were incubated with 1 µg of poly(dI·dC)-poly(dI·dC) substrate (Sigma), 0.5 µg/µl of BSA and 1 µCi of S-adenosyl-[³H-methyl]-methionine in 50 µl of trapping buffer (10 mM Tris pH 7.5, 100 mM KCl, 1 mM EDTA, 1 mM DTT) for 60 min at 37°C. 15 µl of each sample were spotted onto blotting paper and the DNA was precipitated with ice cold 5 % TCA. After washing twice with 5% TCA and once with cold 70 % ethanol, paper filters were air dried and analyzed by scintillation in 4 ml scintillation cocktail (Rotiszint® eco plus, Roth) for 5 min.

De novo activity of DNA methyltransferase 1 in triple knock-out mouse ES cells *(manuscript in preparation)*

2.3   *De novo* activity of DNA methyltransferase 1 in triple knock-out mouse ES cells

Dnmt1 was the first DNA methyltransferase to be described in mammalian cells and is the major enzyme involved in maintenance of DNA methylation patterns of post-replicative DNA sequences (Bestor & Ingram, 1983; Hermann et al, 2004a; Leonhardt et al, 1992). Dnmt1 has a strong preference for hemi-methylated sites *in vitro*, although *in vitro* studies revealed residual activity on unmethylated DNA sequences (Frauer & Leonhardt, 2009; Hermann et al, 2004a; Pradhan et al, 1999). Moreover, in cancer cells, overexpression of Dnmt1 was associated with hypermethylation at specific promoter regions which argues for a *de novo* methylation activity of Dnmt1 (Biniszkiewicz et al, 2002). The question that remains unanswered is whether Dnmt1 has *de novo* methylation activity on unmethylated DNA sequences in living cells. To this aim, we analyzed the ability of Dnmt1 to restore methylation patterns in methylation devoid triple knockout ES cell line (TKO) (Tsumura et al, 2006).

Material and Methods

Expression constructs and cell lines: The GFP-Dnmt1 expression construct was described previously (Frauer et al, 2011). The GFP-Dnmt3a and GFP-Dnmt3b constructs were kindly provided by (Okano et al, 1998), triple knock out cell lines were generated and kindly provided by (Tsumura et al, 2006).

Cell culture and transfection: Mouse ES cells were cultured without feeder cells in gelatinized flasks and in DMEM supplemented with 16% fetal calf serum, 1000 U/ml LIF, ß-mercaptoethanol and L-Glutamin as described previously (Frauer et al., 2011). Cells were transfected with a GFP-Dnmt expression construct immediately after splitting with FuGENE HD (Roche, Mannheim) according to the manufactures protocol.

Rescue assay and stable cell line generation: GFP-positive cells were sorted 48 hours after transfection with a FACS Vantage SE cell sorter or a FACS Aria II (Becton –Dickinson), respectively. Sorted cells were either directly lysed to isolate genomic DNA (48 hours rescue assay) or plated under low density for stable cell line generation. The remaining GFP-positive cells were subsequently sorted until a stable pool of GFP-expressing cells was obtained. After a stable pool of cells was generated, single cells were sorted into a

gelatinized 96 well plate to generate single clones. Clones were cultured, expanded and proteins were extracted to analyze the expression of the GFP fusion protein. Expression levels were compared to the endogenous protein levels in wild type ES cells and clones expressing similar levels were chosen.

Genomic DNA Isolation and Bisulfite Treatment: Genomic DNA was isolated using QIAmp DNA Mini kit (Qiagen) kit and 500 ng-1.5 μg DNA was bisulfite treated using either EpiTect (Qiagen) or EZ DNA Methylation-Gold Kit (Zymo Research) following the manufactures conditions.

Methylation Analysis based on Pyrosequencing: Bisulfite treated DNA was amplified in a PCR reaction with specific Primer sets and PCR conditions for CpG Islands of *skeletal α-actin*, *H19a*, *dnmt1o* promoter, *Xist* exon 1, intracisternal type A particle long terminal repeat (IAP) and major satellite repeats (see table Table 1 and Table 2).

Table 1: Bisulfite primer sequences for single copy gene promoters.

promoter	forward primer sequence	reverse primer sequence	length	CpG
skeletal- α actin				
	GGGGTAGATAGTTGGGGATATTTTT	CCTACTACTCTAACTCTACCCTAAAT A-bio	309bp	13
dnmt1o	GTTGTTTTTTGGTTTTTGTGGGTA	CAACCTTAACAACACAACTAAAATA		
	GTTGTTTTTTGGTTTTTGTGGGTA	CAACTATACACTATCAAATAACCT-bio	272bp	4
H19a	GATTAGATAGTATTGAGTTTGTTTGGAGT	CCTAAAATACTAAACTTAAATAACCCA CAA		
	GAG AAA ATA GTT ATT GTT TAT AGT TTT	ACCATTTATAAATTCCAATACCAAAAA TAA-bio	317bp	10
xist	GTTAATTAATGTAGAAGAATTTTTAGTGTT TA	TTATTTAAGGAGTTTTGGGGGAATATT T		
	GTTAATTAATGTAGAAGAATTTTTAGTGTT TA	TTTAATAAGATGTTAGAATTGTAATTTT TGTG -bio	458bp	19

Major satellite repeats were amplified in a single PCR reaction and two PCR reactions were pooled using a Millipore purification kit, while all other sequences were amplified in a nested PCR reaction. The amplified PCR product was subjected to pyrosequencing carried out by varionostic GmbH, Ulm.

Table 2: Bisulfite primer sequences for repetitive sequences

sequence	forward primer sequence	reverse primer sequence	length	CpG
MajSat	AAATGAGAAAtATttAtTTG	CCATaATTTTCAaTTTTCTT -bio	234bp	8
IAP	GTTGTTTTTTGGTTTTTGTGGGTA	CAACCTTAACAACACAACTAAAATA		
	GTTGTTTTTTGGTTTTTGTGGGTA	CAACTATACACTATCAAATAACCT-bio	259bp	4

Methylation analysis in TKO cells transiently rescued with GFP-Dnmt1

We investigated the activity of Dnmt1 on unmethylated DNA in TKO ESC's by re-expression of Dnmt1 in transient as well as stable rescue experiments. For this aim, we used triple knockout (TKO) ESC's that lack all the three major methyltransferases Dnmt1, Dnmt3a and Dnmt3b and therefore are devoid of methylation. We performed a transient rescue assay where TKO cells were transfected with GFP-Dnmt1, GFP-Dnmt3a and GFP-Dnmt3b and analyzed selected sequences to determine the ability to methylate unmethylated DNA sequences *in vivo*. In a transient rescue assay TKO cells were transfected with the respective plasmids and FACS sorted after 48 hours. Genomic DNA was isolated, bisulfite treated and subjected to pyrosequencing for a quantitative methylation analysis. TKO cells rescued with GFP-Dnmt1 remained unmethylated in all of the analyzed sequences after 48 hours. In comparison we performed the same approach with GFP-Dnmt3a and GFP-Dnmt3b, respectively, which showed that both *de novo* methyltransferases are active and able to re-methylate unmethylated DNA sequences in TKO cells.

Establishment of stable TKO cell lines and methylation analysis by pyrosequencing

To overcome limitations of a transient rescue assay as well as to analyze long-term effects of Dnmt1 expression in TKO cells, we additionally established TKO cell lines, which stably express GFP-Dnmt1. For a comparison of different methyltransferase activities we also generated stable pools of GFP-Dnmt3a and GFP-Dnmt3b TKO cell lines, respectively. For this aim we established a new FACS based approach to generate stable TKO cell lines

80

since TKO cells are resistant to the most commonly used selective markers (Tsumura et al, 2006). After several rounds of FACS sorting for GFP positive cells and their subsequent propagation all cell lines were stable. We generated pools of TKO cells stably expressing GFP-Dnmt1, GFP-Dnmt3a and GFP-Dnmt3b. Furthermore the pool of TKO cells stably expressing GFP-Dnmt1 was further subcloned and single clones were selected for further methylation analysis. Genomic DNA was isolated, bisulfite treated and subjected to pyrosequencing as described for the transient rescue experiment. Moreover, we analyzed three different pools of GFP-Dnmt1 stable cell lines over several passages to address dynamics and kinetics of a potential *de novo* methylation activity of Dnmt1 and to detect a possible accumulation of methylation over time. Genomic DNA samples were obtained three and six weeks after transfection and subjected to pyrosequencing.

A *skeletal α-actin* promoter -
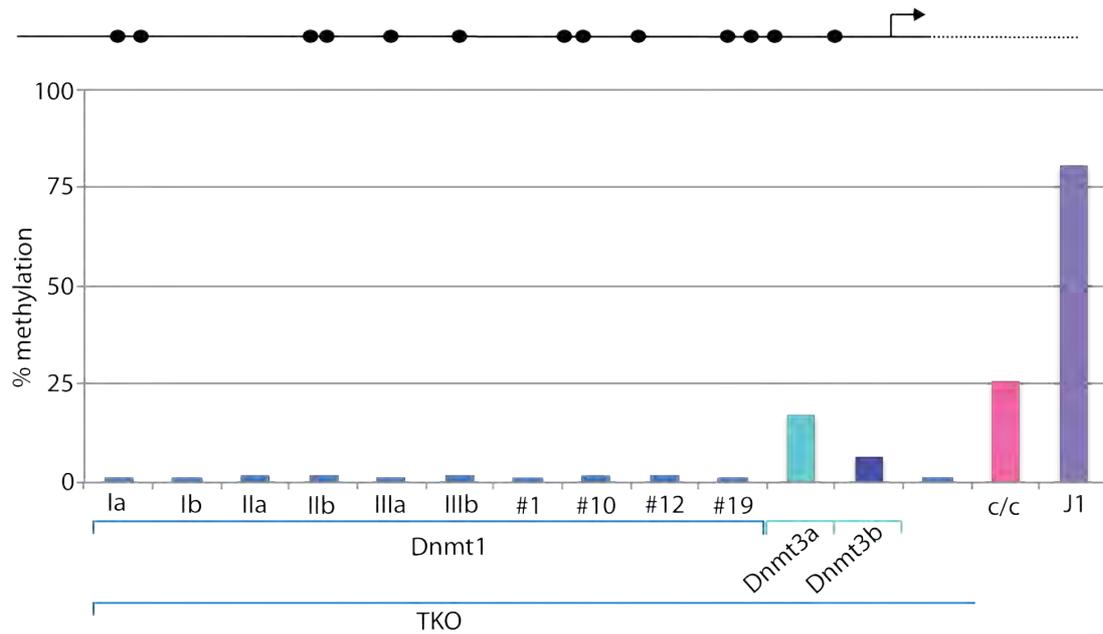
B *dnmt1o* promoter -
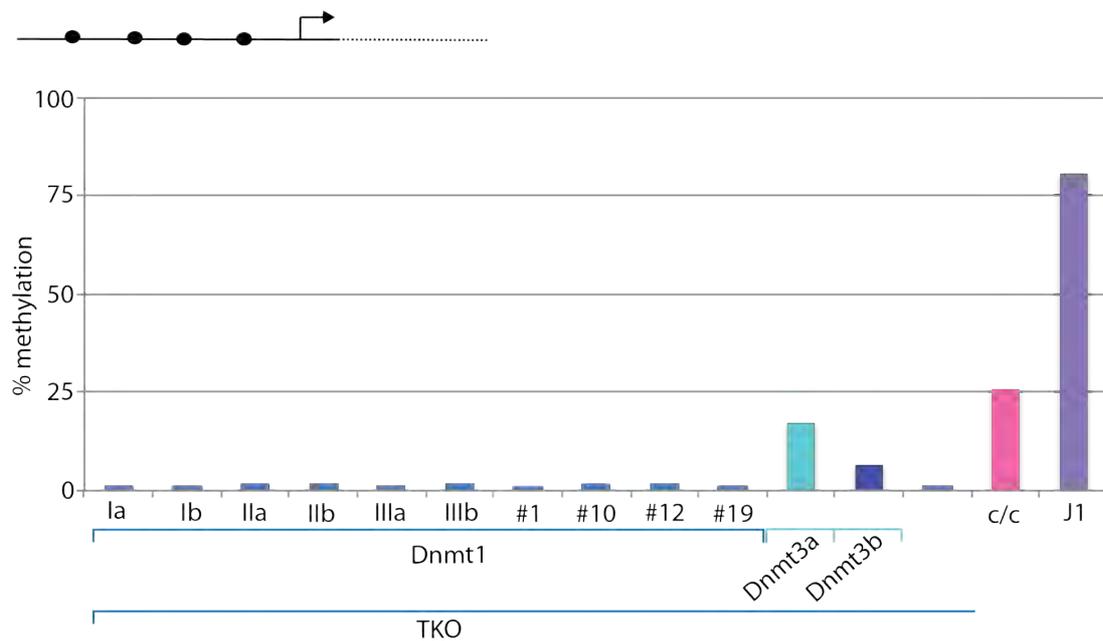
Figure 2-1: Quantitative methylation analysis of single copy gene promoters in TKO cells rescued with GP-Dnmt1, GFP-Dnmt3a and GFP-Dnmt3b. Quantitative methylation levels of (A) 13 CpG sites were averaged for *skeletal α-actin* promoter and (B) 4 CpG sites for *dnmt1o* promoter for each analyzed sample. Methylation of wild type and *dnm1⁻/⁻* are displayed for comparison. TKO cells were rescued with GFP-Dnmt1 and three different pools of stable cell lines (I-III) were analyzed 3 weeks (a) and 6 weeks (b) after stable transfection. Four different stable clones (#1,#10,#12,#19) were selected whose *dnmt1* expression levels are similar compared to the endogenously expressed Dnmt1 in wild type cells (data not shown). CpG Methylation was not detected at any time point in any of the analyzed TKO GFP-Dnmt1 cell lines (A,B). TKO cells rescued with GFP-Dnmt3a and GFP-Dnmt3b showed detectable methylation in the *skeletal α-actin* promoter (A) and in the *dnmt1o* promoter (B) indicating *de novo* methylation activity.

A *xist* promoter
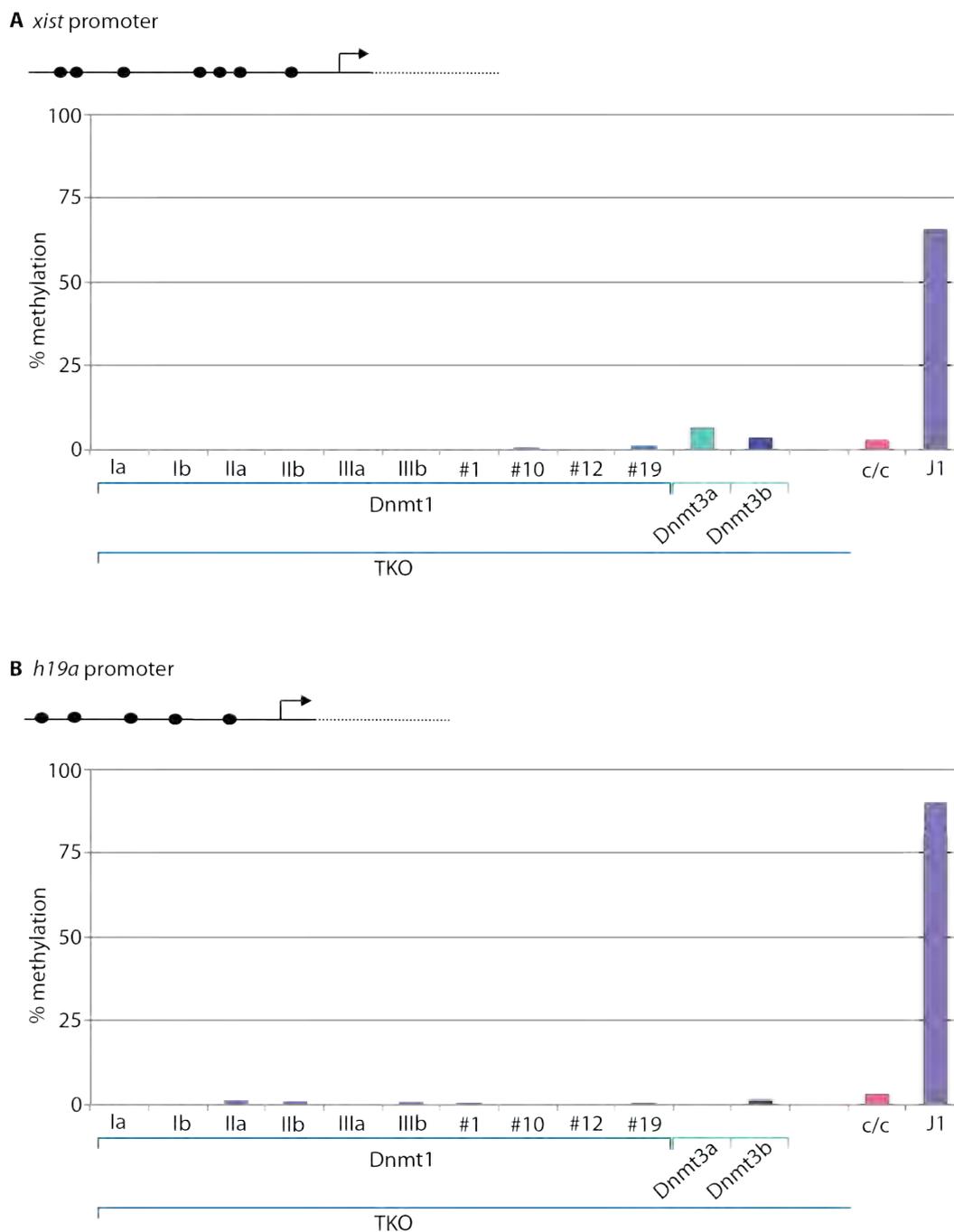
B *h19a* promoter

Figure 2-2: Quantitative methylation analysis of imprinted promoter H19a and *xist* promoter in TKO cells rescued with GP-Dnmt1, GFP-Dnmt3a and GFP-Dnmt3b. Quantitative methylation levels of (A) 6 CpG sites were averaged for *xist* promoter and (B) 5 CpG sites for *h19a* promoter for each analyzed sample. Methylation of wild type and *dnm1*[-/-] are displayed as comparison. TKO cells were rescued with GFP-Dnmt1 and three different pools of stable cell lines (I-III) were analyzed 3 weeks (a) and 6 weeks (b) after stable transfection. Four different stable clones (#1,#10,#12,#19) were selected whose *dnmt1* expression levels are similar compared to the endogenously expressed Dnmt1 in wild type cells (data not shown). CpG Methylation was not detected at any time point in any of the analyzed TKO GFP-Dnmt1 cell lines (A,B). TKO cells rescued with GFP-Dnmt3a and GFP-Dnmt3b showed detectable methylation in the *xist* promoter (A) indicating *de novo* methylation activity, while no methylation was detectable in the *h19a* promoter (B).

A major satellite repeats
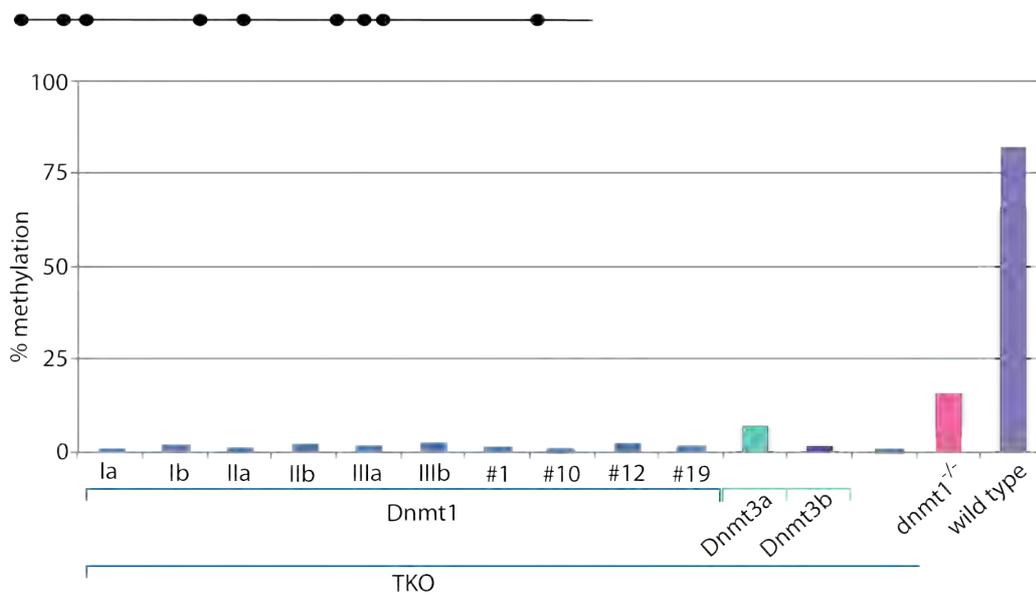
B IAP repeat
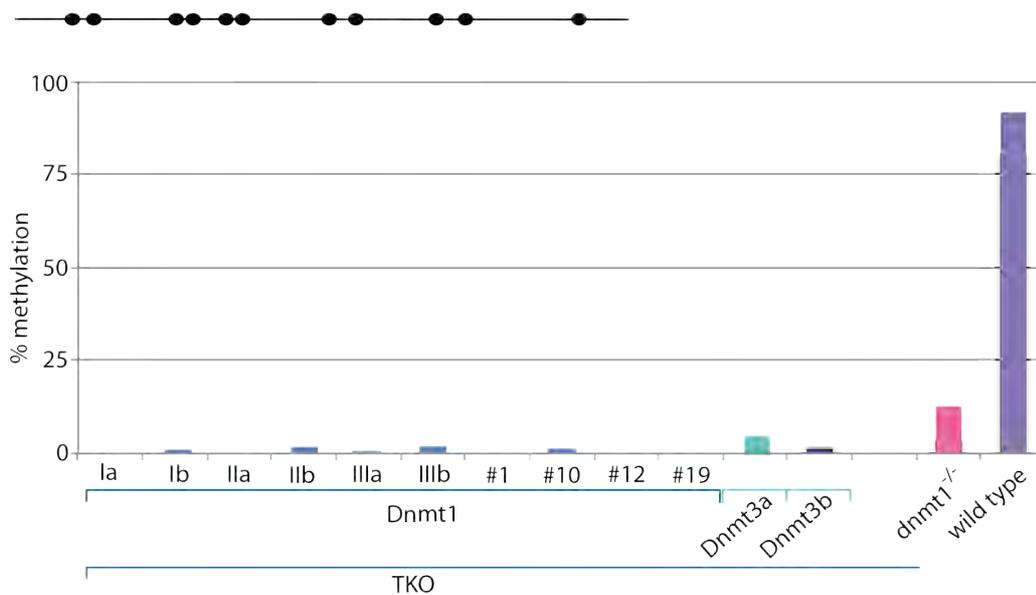
Figure 2-3 Quantitative methylation analysis of repetitive sequences major satellites and IAP elements TKO cells rescued with GP-Dnmt1, GFP-Dnmt3a and GFP-Dnmt3b. Quantitative methylation levels of (A) 8 CpG sites were averaged for major satellites and (B) 11 CpG sites for IAP elements for each analyzed sample. Methylation of wild type and *dnm1*[−/−] are displayed as comparison. TKO cells were rescued with GFP-Dnmt1 and three different pools of stable cell lines (I-III) were analyzed 3 weeks (a) and 6 weeks (b) after stable transfection. Four different stable clones (#1,#10,#12,#19) were selected whose *dnmt1* expression levels are similar compared to the endogenously expressed Dnmt1 in wild type cells (data not shown). CpG Methylation was not detected at any time point in any of the analyzed TKO GFP-Dnmt1 cell lines (A,B). TKO cells rescued with GFP-Dnmt3a and GFP-Dnmt3b showed detectable methylation in major satellites (A) and in the IAP elements (B) indicating *de novo* methylation activity.

Methylation analysis in TKO cells stably expressing GFP-Dnmt1, GFP-Dnmt3a and GFP-Dnmt3b

The analysis of GFP-Dnmt3a and GFP-Dnmt3b stable cell lines show low but measureable *de novo* methylation in all analyzed sequences, except for the case of *H19a*, which requires passage through the germ line to become methylated (Tucker et al, 1996). We were able to detect a 10-15-fold increase in methylation levels of GFP-Dnmt3a and GFP-Dnmt3b stable cell lines compared to untreated TKO cells. However, these methylation levels do not by far reach those found in *dnmt1$^{-/-}$* ES cells where both *de novo* methyltransferases are expressed. Therefore, our data points towards the important role of a synergetic function between Dnmt3a and Dnmt3b in *de novo* methylation activity.

Surprisingly, in none of the analyzed GFP-Dnmt1 clones and stable pools was DNA methylation detectable above background, even after several passages of cultivation (see Figure 2-1, Figure 2-2, Figure 2-3). These findings clearly show that Dnmt1 alone is not able to *de novo* methylate unmethylated DNA sequences *in vivo* and argues against several studies claiming that Dnmt1 has *de novo* methylation activity apart from its clear involvement in maintenance. To clarify whether proximal pre-existing CpG sites are able to initiate *de novo* methylation of Dnmt1 further experiments have to be performed in living cells. To this aim introduction of methylation at specific sites followed by a rescue assay with Dnmt1 would be necessary. Moreover, to investigate whether interaction between the three active methyltransferases might activate de novo activity of Dnmt1, double-rescue experiments of the active methyltransferases in combination with their respective catalytically inactive "mutants" are necessary to be performed.

DNA-methylation of CpG dyads and markov modeling of DNA-methylation control in mammals

(manuscript in preparation)

2.4 DNA-methylation of CpG dyads and markov modeling of DNA-methylation control in mammals

DNA-methylation of CpG dyads and markov modeling of DNA-methylation control in mammals

(proposed author list)

Julia Arand[a], David Spieler[b], Tommy Karius[a,1], Daniela Meilinger[c], Alexander Meissner[d], Masaki Okano[e], Thomas Jenuwein[f], Guoliang Xu[g], Heinrich Leonhardt[c], Verena Wolf[b] and Jörn Walter[a,2]

a University of Saarland, FR 8.3, Biological Sciences, Genetics / Epigenetics, Campus A2.4, 66123 Saarbrücken, Germany,

b University of Saarland, FR 6.2, Computer Science Department, Campus E1.3, 66123 Saarbrücken, Germany,

c LMU München, Biozentrum Martinsried, Dept. Bio II., Grosshaderner Str. 2, 82152 Martinsried, Germany,

d Broad Institut, 7 Divinity Avenue, Rm 207, Cambridge, MA, 02138, USA,

e RIKEN Center for Developmental Biology, Lab for Mammalian Epigenetic Studies, 2-2-3 Minatojima-Minamimachi, Chuo-ku, Kobe 650-0047, Japan,

f MPI for Immunology and Epigenetics, Stübeweg 51, 79108 Freiburg, Germany,

g Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, 320 Yueyang Road, Shanghai 200031, China,

1 current address: Laboratoire de Biologie Moléculaire et Cellulaire du Cancer (LBMCC), Hôpital Kirchberg 9, rue Steichen 2540, Luxemburg

2 corresponding author: j.walter@mx.uni-saarland.de, University of Saarland, FR 8.3, Biological Sciences, Genetics / Epigenetics, Campus A2.4, 66123 Saarbrücken, Germany, Phone: 0681/302-4367, Fax: 0681/302-2703

Abstract

Repetitive elements such as LINE1, B1, IAP-LTR-retrotransposons and major Satellites, comprise a large portion of the mammalian genome and are major targets of DNA-methylation mediated silencing (1) (2). Using 454 bisulfite sequencing combined with hairpin linker technology we mapped the CpG dyad methylation on both strands of the four major classes of repetitive elements at high resolution. A comparative analysis of sequence reads generated from somatic cells, wt and mutant ESCs with deficiencies for Dnmt1, Dnmt3a, Dnmt3b, Dnmt3L, NP95 and double deficiency for Dnmt3a/3b and Suv39h1/2 allowed us to assign cell type differences and enzyme dependent methylation changes. The data set was used to develop a comparative Hidden Markov Model (HMM) to estimate the relative enzyme rates for *de novo* and maintenance methylation. The HMM shows that i) Dnmts contribute with varying degrees to CpG dyad methylation and ii) Dnmts have distinct preferences for the methylation of individual types of repetitive elements. Besides a predominant methylation maintenance role of Dnmt1 (and NP95) the model also predicts a substantial *de novo* methylation activity for Dnmt1 at IAPs and major satellites. Conversely, Dnmt3a and Dnmt3b have a strong preference for hemimethylated CpG dyads contributing to maintain full methylation following replication at certain elements. Finally, we identified CpA methylation at specific locations in major satellite sequences of ESCs catalysed by distinct position dependent activities of Dnmt3a or Dnmt3b, respectively. Dnmt3L enhances this CpA methylation activity along with CpG methylation for both enzymes.

Introduction

DNA methylation at the C-5 positions of cytosine (5mC) is a key epigenetic modification in mammals essential for normal development of mammals (3). Cytosine methylation is predominantly found in CpG dinucleotide context and about 70 to 80% of all CpGs are methylated (4). These methylated CpGs are usually located in CpG poor regions and often in repetitive sequences (4, 5). About 40% of the genome consists of repetitive elements. Four main groups of repetitive elements can be discriminated: long interspersed nuclear elements (LINEs), short interspersed nuclear elements (SINEs), long-terminal repeat (LTR) retrotransposons and (peri-) centromeric satellites. For these elements, the maintaining of DNA methylation during development and aging is important for transcriptional silencing and genome stability (1, 6).

The establishment and maintenance of methylation patterns at palindromic CpG sequences (CpG dyads) is performed by three catalytically active DNA methyltransferases (Dnmts). *In vitro* experiments suggest that Dnmt1 prefers hemimethylated CpG (hemi-mCpG) dyads and maintains the methylation pattern on the newly synthesised strand after replication (maintenance methylation). In *vitro* Dnmt1 shows a low activity on unmethylated CpG dyads. Dnmt3a and Dnmt3b methylate DNA *de novo*, independent of the methylation status of the complementary CpG position (7). I*n vitro* analyses furthermore suggest that methylation by Dnmt3b and the maintenance function of Dnmt1 mostly occur in a processive manner, whereas Dnmt3a and the "*de novo* function" of Dnmt1 are distributive (8-10). However, other groups observe a processive methylation activity for DNMT3a (11). Although there is a lot of *in vitro* data available on Dnmts and their possible functions, little is known about their concerted action *in vivo* at dispersed repetitive elements. Data of ESCs with individual and combined Dnmt knockouts indicated preferences of Dnmts for specific repetitive elements (2). In addition Dnmt activities are modulated by Nuclear protein of 95 kDa (Np95; also known as Uhrf1) and Dnmt3L. Dnmt3L, a cofactor for the *de novo* methyltransferases, is reported to stimulate Dnmt3a/3b activity, to be needed for *de novo* establishment for imprint methylation and furthermore to enhance the processive methylation activity of human Dnmt3a (11-13). Np95 is recruiting Dnmt1 to hemimethylated DNA (14, 15) and is interacting with Dnmt3a/3b for gene silencing (16).

Different mathematical models were developed to simulate the kinetics of DNA methylation (17-20). Initial DNA methylation analysis using hairpin bisulfite sanger sequencing allowed to gain a first insight on symmetrical methylation of DNA strands and such data were used for initial static mathematical modelling (21, 22).

In this paper, we present a first comprehensive high resolution methylation analysis for both DNA strands using hairpin linker technology. This allowed us to determine the methylation symmetry along individual chromosomal fragments of four distinct classes of repetitive elements in wt ESCs and ESCs depleted for Dnmt1, Dnmt3a, Dnmt3b, Dnmt3a/b, Dnmt3L, Np95 and Suv39h. The complex data set was used to further develop hidden Markov models (HMM) which describe the relative contribution of Dnmts for maintenance and *de novo* methylation (17, 18, 22). This provides a number of novel insights in the complexity and dynamic of DNA methylation control.

Results

Outline and quality monitoring of the hairpin bisulfite sequencing strategy

We designed specific hairpin linker protocols to amplify representative fragments of the four major classes of repetitive elements from bisulfite treated mouse DNA. The sequence elements selected were i) major Satellites, ii) IAPLTR1, a class of LTR-retrotransposons, iii) the 5' untranslated region of L1MdTf, a long interspersed element (LINE) and iv) a conserved portion of B1 elements, representing a class of short interspersed elements (SINEs) (see also Supplementary Table S1 for detailed references). In this paper, we conveniently refer to the repeat elements as mSat, IAP, L1 and B1, respectively. Following amplification the PCR products were sequenced on a 454 GS-FLX sequencer with an average read length of 200-400bp covering 3 to 7 CpG dyads of the respective amplicons.

The addition of an unmethylated hairpin linker containing several unmodified cytosines allowed us to directly monitor the bisulfite conversion rates per sequenced

molecule. In the linker sequences the conversion rates ranged from 99,1- 99,4% for mSat, IAP and L1 sequences, respectively. Sequences of B1 elements showed a slightly lower conversion rate of 97.8%, probably due to the more degenerate sequence composition and occasional back-folding (Supplementary Fig. S1a, Supplementary Table S2).

In contrast to conventional single strand bisulphite sequencing, hairpin bisulfite sequencing allows to unambiguously distinguish between unmethylated and mutated CpG sites. We identify mutated CpGs in all repetitive element to various extend (see white positions in Fig.1). A particular abundance of mutated CpGs was found in B1 elements with 50% of CpGs being mutated to TpG (see also Supplementary Fig. S1b). Previous single strand bisulfite sequencing accounted such TpGs as unmethylated positions estimating the total methylation of B1 elements to be only 10%. (23). When correcting for mutated sites we find B1 elements to be up to 80% methylated in wt cells.

Analysis of the methylation symmetry at CpG dyads in various repetitive elements

First, we compared the DNA methylation patterns of repetitive elements between mouse wt ESC lines, mouse embryonic liver and cultured mouse embryonic fibroblasts (MEFs) (Fig.1). The comparison included separate analysis of the three ES cell lines J1 (24), E14 (15) and WT26 (25) which are the "parental" lines for the Dnmt knockouts (J1) (3, 26, 27), the Np95 knockout (E14) (15) and the Suv39h knockout (WT26) analysed (25).

MSat, IAPs and B1 are highly methylated in all wt ESC and somatic cells. Their overall methylation ranges from 60-95% at CpG dyads. The same holds for L1 elements in somatic cells (> 74%). In ESCs, however, L1 methylation levels only reach a max. of 30-52%. Moreover ESCs appear to have a significant general variation in their overall methylation levels. The ESC line J1 is significantly higher methylated at all repetitive elements compared to E14 and WT26. Also between MEFs and embryonic liver we detected a substantial difference in the overall methylation status (Fig. 2a).

Following a precise alignment to reference sequences using BiQAnalyzerHT (28) we calculated the amount of hemimethylated CpG dyads (hemi-mCpGs) for all sequences.

Laird *et al*. and Burden *et al.* reported values between 4.8% (sperm) and 20.8% (leukocytes) of hemi-mCpGs for human LINE1 elements (21, 29). Shao *et al.* observed 7% hemi-mCpG in human Satellite 2 sequences in different tissues (30). In mouse we also observe 8 to 25% of all CpGs in a hemimethylated status (Fig. 2b). MEFs show the lowest and least variable rate of hemi-mCpG sites across all analysed elements (8 to 10.5%) while in embryonic liver the highest amount of hemi-mCpGs is detected across all repetitive elements (20 to 25%). Strikingly, unlike the almost equal occurrence of hemi-mCpGs at various repetitive elements in differentiated cells, the degree of hemimethylation is much more variable in repetitive elements of wt ESCs. While hemimethylation levels of IAPs are rather low (9 to 12%) more than 20% of CpGs in L1 are hemimethylated in ESCs (Fig.2b). Moreover the extend of hemi-mCpGs at mSat greatly varied between the three wt ESC lines.

Loss of single Dnmts has distinct effects on the methylation status of repetitive elements

Next, we analysed the contributions of Dnmts and cofactors for the maintenance of methylation by comparing hairpin bisulfite sequence data of ESCs mutated for Dnmt1, Dnmt3a, Dnmt3b, Dnmt3a and 3b (DKO), Dnmt3L and Np95 (UHRF1), respectively. An overview of the methylation patterns is given in Fig. 3.

Deletion of Dnmt1 caused a high reduction of DNA methylation in all analysed elements when compared to the corresponding wt ESC. MSat methylation was reduced by 66%, IAP by 72%, L1 by 76% and B1 by 74%, respectively. This tendency was also observed in previous low resolution data obtained for a subset of repetitive sequence elements (2, 26). The effects of Np95 KO were very similar to the Dnmt1 KO (see also Bostick *et al.* (14)).

In contrast to a general increased hypomethylation at all elements in Dnmt1/Np95 KOs, the loss of Dnmt3 activities led to element and enzyme specific differences. While methylation at IAP and mSat did not change in Dnmt3a or Dnmt3b single KO`s, the double KO, analysed at passage 8, lead to a clear decrease of CpG methylation at mSat (24%)

and IAPs (17%). Methylation at L1 and B1 did not change in the Dnmt3b single KO but was strongly decreased in the Dnmt3a single KO by 64% for L1 and 37% for B1. Hence while either the loss of Dnmt3a or Dnmt3b can be compensated by the other enzyme for IAPs and mSat the situation is more complex at B1 and L1, respectively. Finally also Dnmt3L contributes to the maintenance of methylation. In the Dnmt3L KO methylation decreases at all elements but much less pronounced as in the Dnmt3a/3b DKO arguing for a stimulatory effect of Dnmt3L on both *de novo* Dnmts. Note that Dnmt3L KO cells were at passage 15 and underwent already almost twice the amount of replications than the Dnmt3a/b DKO.

Loss of Dnmt1 and Np95 KO leads to an increase of hemi-mCpG sites

For all elements we observe a strong increase in the relative amount of hemi-mCpGs (in regard to total methylation) in Dnmt1 KO and Np95 KO (Fig. 4a), along with a huge loss of overall methylation. This observation highlights the important role of Dnmt1 in maintaining symmetrical CpG methylation. However, in both Dnmt1 and Np95 null backgrounds we still find a considerable amount of sequences with fully methylated CpG dyads which in addition are not sporadically distributed (Fig. 3, Fig. 4b).

Chromosomal sequences with hemi-mCpG sites on only the upper or lower strand, respectively, were much more frequently compared to sequences with (dispersed) hemi-mCpG sites on both upper and lower strands. Such mosaical hemimethylation was found in <2% of mSat, <4,4% of IAP, <7% of L1 and <3,4% of B1 reads (see Fig. 4b). In Dnmt1 KO cells, mosaical hemi-mCpGs were enriched in IAP, B1 and mSat but not in L1 sequences. In Np95KO ESCs this mosaical hemimethylation increases further in all four repetitive elements indicating that the contribution of Np95 and Dnmt1 to maintain methylation pattern are partially distinct. Along this line, the Dnmt1 KO shows a 3,8 (for IAPs) and 4,8 (for mSat) fold increase of fully methylated sequences as compared to NP95 KO (Fig. 4b). In contrast to the Dnmt1 KO and Np95 KO, respectively, the relative abundance and distribution of hemimethylated sites does not differ between Dnmt3a and Dnmt3b single KOs and DKO.

CpA methylation is present at major Satellites in ESCs

Along with the CpG methylation, we identified sites of non-CpG methylation in mSat sequences of all WT ESC lines (J1, E14 and WT26). In none of the other elements we found evidence for non-CpG methylation. In mSat sequences cytosines at positions 4, 6, 11 (top strand) and 22 and 28 (botton strand) showed significant methylation levels above the conversion background (1.1%) as defined by linker sequence conversion (see above). All five unconverted (methylated) positions are found in the CpA sequence context (Fig. 5a+5b). The highest CpA levels were present at position 6 where up to 10% of all sequences of the ESC line J1 contained an unconverted cytosine at this position. It is interesting to note in most of the sequence reads only a single CpA methylated position was detected (e.g. in 80% of J1 reads with CpAs) (Supplementary Figure S4).   In differentiated cells (embryonic liver and MEFs), no methylated CpA positions were detected above background. This is in line with previous findings by Lister *et al.* in human ESCs (31) (Fig. 5c).

Dnmt3 family is responsible for non-CpG methylation in major satellites

To address the question, which Dnmts are responsible for CpA methylation we compared the relative occurrence of methylated CpAs between wt and the different KO ESC lines (Fig. 5d). In Dnmt1 KO the methylation at all five CpA positions remained unchanged despite of a greatly reduced CpG methylation level. Notably, sequences of Dnmt1KO ESCs with high overall CpG methylation showed increased CpA methylation other than the residual sequences with relatively low CpG methylation. Moreover in Dnmt1KO sequences, where we observe CpA methylation, the proportion of hemi-mCpGs is enriched on the upper strand, i.e. in the strand where we also observe increased CpA methylation in wt ESCs (Fig. 6).

Together these findings clearly point towards an exclusive Dnmt3a and 3b mediated CpA methylation. Surprisingly, the relative contribution of Dnmt3a and Dnmt3b, respectively to CpA methylation differed at the five position. Whereas CpA methylation at position 6 and 11 is greatly reduced in a Dnmt3a KO, the loss of Dnmt3b diminishes CpA methylation at

position 4, 22 and 28. In Dnmt3a/3b DKO CpA methylation above background is completely absent. The loss of Dnmt3L greatly reduces the methylation at most positions except position 28.

Suv39h KO decreases CpG methylation level at mSat but not L1, B1 or IAP and has no influence on non CpG methylation

The Suv39h1/2 mediated modification of histone H3 at position 9 was reported to influence the targeting of DNA methylation. We therefore included ESCs and MEFs with KO for Suv39h1 and Suv39h2 (Suv39dn) in our analysis. Suv39dn ESCs were reported to lack H3K9 trimethylation and Dnmt3b localisation at pericentric heterochromatin. Lehnertz *et al.* reported reduced DNA methylation (by southern blot) at mSat in Suv39h KO but not at minor Satellites or a C-type retrovirus (25). Our hairpin bisulfite analysis confirmed this finding on a sequencing basis. DNA-methylation at major satellites is reduced by 20% in Suv39dn ESCs but does not at B1, IAP and Line1 elements. Surprisingly the effect on mSat methylation is almost absent in dnMEFs which retain 95% of wt methylation (Fig. 7a). Finally despite of the proposed interaction of Suv39h with Dnmt3b at mSat we do not observe any influence of the Suv39h absence on CpA methylation particularly not at the Dnmt3b specific positions 4, 22 and 28 (Fig. 7b).

Hidden Markov model predicts methylation efficiencies of Dnmts

The precise determination of fully methylated, hemimethylated and unmethylated CpG dyads allowed us to calculate the methylation characteristics for the different Dnmts in a modified version of the basic HMM proposed by Sontag *et al*. (Sontag *et al.* 2006). The comparative data set of some 20.000 sequences (around 84.000 CpG positions) including Dnmt KO`s, allowed us to compute maximum likelihood estimates for both methylation efficiencies on unmethylated and hemimethylated CpG dyads separately (Supplementary Table 3). In our model we also accounted for measurement errors based on exact bisulfite conversion rates and we approximated the variances of estimated methylation efficiencies in order to determine the quality of the estimation. As opposed to previous models, we

also did not assume that steady-states are reached in the KO ESC lines. Instead, we estimate the amount of cell divisions and infer parameters during the transient phase of the system, since at least Dnmt3a/3b DKO shows a processive loss of DNA methylation with increasing passage number (Chen *et al.* 2003).

In our calculations we assumed that *de novo* methylation has the same probability to methylate unmethylated and hemi-mCpG dyads. We list the estimated efficiency values and standard deviations in Table 1, where the entries are the probabilities of Dnmt1, Dnmt3a and Dnmt3b mediated *de novo* or maintenance methylation in one DNA replication. The approximated standard deviations showed that for most elements and enzymes efficiencies were accurately estimated. For L1, a high standard deviation of the maintenance function for Dnmt3a showed that here the parameter is not identifiable. This is probably due to the small number of hemi- and fully methylated dyads in the KO's of L1, which could be consulted for the model development.

To substantiate the appropriateness of our model and the accuracy of our estimated methylation efficiencies, we then predicted the DNA methylation level for the parental wt ESC line (Table 2). Indeed we found excellent predictions for mSat, IAP and L1 (maximum error rate of 1.8%, 2.1% and 3.3%, respectively) while the prediction for B1 had a maximal error rate of 12.1%.

Based on the HMM calculations the *de novo* methylation efficiency of Dnmt1 at IAP and mSat is about four to five times lower than the maintenance methylation efficiency (IAP: 4.1, Sat: 4.7). At L1, Dnmt1 does not show a remarkable *de novo* methylation activity (for L1 <76 fold less compared to maintenance methylation). Interestingly, for the *de novo* methyltransferase Dnmt3a we observe a 1.8 (mSat) to 3 (IAP) fold higher contribution to methylation maintenance, which contrasts *in vitro* derived data (Okano *et al.* 1998). The contribution of Dnmt3b to methylation maintenance is even higher for IAP and mSat.

Discussion

In our work we present the first high resolution DNA bisulfite methylation map of repetitive elements in double stranded DNA. Comparative analysis of wt and KO mouse ES lines revealed detailed insights in the relative contributions of Dnmts to CpG dyad and non-CpG methylation. In a HMM, we predicted the relative contribution of Dnmts to maintenance and *de novo* methylation.

Distribution and specificity of hemimethylation

Our experimental approach allowed us to unambiguously assign DNA methylation at about 140.000 CpG dyads on both DNA strands. Besides a general prevalence for symmetrical methylation we found a substantial portion of hemimethylated CpG dyads in all cell types analysed. The presence of such hemi-mCpGs can be explained by three different mechanisms: the improper recognition and maintenance activity of Dnmt1/NP95, the selective *de novo* methylation or active DNA demethylation. The different levels of hemimethylated CpGs in MEFs or embryonic liver cannot simply be explained by a variable expression of Dnmt1. At least on the transcriptional level Dnmt1, Dnmt3a and 3b appear to be equally expressed in both cell types (Supplementary Figure S3). This suggests that in liver the Dnmt1/NP95 maintenance fidelity is less pronounced or alternatively DNA demethylation is more pronounced at repetitive elements. The recently discovered 5hmC may play a major role for both processes. Since bisulfite modification does not discriminate between 5mC and 5hmC our sequencing data represent an unknown mixture of methylated and hydroxymethylated cytosines (32, 33). 5hmC was reported to be abundant in ESCs and some tissues, but less so in cultured cells ((34, 35). The relative abundance of 5hmC could therefore explain the observed effects on methylation maintenance, however the effect of 5hmC on 5mC maintenance is not yet clear. While 5hmC was shown not to be bound by Dnmt1 in vitro (36), Np95 apparently recognizes and binds to 5hmC containing DNA (37). In ESCs, which posses a relative abundance of 5hmC we observe variable frequencies of hemi-mCpGs at the different repetitive element classes. In L1 sequences an intermediate level of overall methylation is linked to a high level of hemimethylated sites (20%). Apparently the maintenance methylation is less

accurate at L1 in ESCs. From our data we see that a strong cooperativity of Dnmt1, Dnmt3a and Dnmt3b is needed to maintain methylation at L1 (a model of the contribution and interaction of enzymes is shown in Fig. 8). The increased presence of 5hmCs at L1 sequences might contribute to an increased dynamic variation at L1 sequences in particular. Ficz *et al.* indeed observe an enrichment of 5hmC in L1 5'UTRs (38) while 5hmC is not enriched at IAPLTR1 (Supplementary Fig. S3). Assuming that there is active DNA demethylation in ESCs, as shown e.g. for Nanog (39), it could be that certain L1 are prone to 5hmC induced demethylation resulting in hemimethylated sites and a lower methylation level. This scenario would be comparable the partial active DNA demethylation found at Line1 elements in mouse zygotes, while IAPs escape demethylation (40, 41). In addition, the increased *de novo* methylation activity of Dnmt1 at IAP, calculated with the HMM, and the increased "maintenance" functions of Dnmt3a and Dnmt3b might partially compensate for these "maintenance errors".

Control of CpA methylation

DNA methylation outside of the CpG context was initially detected in mESCs using nearest neighbour analysis. This analysis revealed a strong prevalence for the CpA context (42). Recent genome wide single stranded bisulfite sequencing using Illumina short reads identified non-CpG methylation at various sequence contexts in human ESCs at rather high rates of 24.5% (human) of all methylated Cs (31) In the genome wide analysis no clear assignment of non-CpG methylation at repetitive elements has been performed. In our data set covering about 170.000 CpG positions and 10^6 bases we exclusively detect CpA methylation confined to only mSat sequences. This CpA methylation amounts to 5% of all methylated cytosines in mSat. It is possible that the primer based amplification of our analysis caused some selection against non-CpG methylation and we therefore underestimate the amount of non-CpG methylation. Still the CpA position dependent methylation remains outstanding. The CpA methylation is exclusively dependent on Dnmts 3a and 3b in combination with Dnmt3L. Neither the absence of Dnmt1 nor Np95 altered the CpA methylation (see the model in Fig.9). Moreover the unchanged CpA methylation in Suv39hdn cells reveals that the proposed protective function of H3K9 trimethylation for

non-CpG methylation may not be true for mSat (43). The sequence analysis of Dnmt1 KOs unambiguously shows that CpA methylation is linked to Dnmt3a and Dnmt3b mediated CpG methylation. Along this line CpA positions are highly co-methylated with some neighbouring CpG positions (Supplementary S4 and S5). This indicates that the CpA methylation result as a position specific "side reaction" of Dnmt3a and Dnmt3b stimulated by Dnmt3L. In line with this Holz-Schietinger *et al.* show that Dnmt3L increases the processivity of Dnmt3a (11). Finally Dnmt3L is highly expressed in embryonic stem cells compared to somatic cells, where we do not find any evidence for CpA methylation (44).

Modelling of Dnmts maintenance and de novo methylation activities

The HMM applied to estimate methylation efficiencies extends previous modelling approaches: i) by separating the effects of all three Dnmts based on KO data we could estimate *de novo* and maintenance methylation probabilities for each enzyme independently. ii) our experimental strategy allowed to precisely assign CpG dyads and account for experimental measurement errors (bisulfite conversion and mutation errors) as well as the number of cell divisions (passages). iii) we employed numerical techniques to infer optimal methylation efficiencies since analytic solutions of our more complex model are infeasible. By integrating all these parameters we could functionally extend the previous models developed by Genereux *et al*. and Sontag et al. (17, 22). Morever byond prediction our validations (see Table 2) demonstrate the appropriateness of the model at least for mSat and IAP, where experimental data is abundantly available. For L1, the prediction is very accurate even though maintenance activity of Dnmt3a is difficult to estimate. For B1, the prediction is rather weak, which indicates that further extensions of the model may be necessary to predict B1 methylation level. Since we presume based on the influence of Dnmt3a and 3b single and double KO that Dnmt3a can widely replace Dnmt3b activity, the model extension first should go in this direction. In the future the model should also account for maintenance and *de novo* methylation effect caused by hydroxymethylcytosine - which at least is enriched in Line 1 elements of ESCs. However given the fact that evidence for direct effects of 5hmC on local demethylation, maintenance and *de novo* methylation, respectively, and a clear distinction between 5mC

and 5hmC in sequences is still lacking the inclusion of 5hmC in the HMM is not yet feasible.

For Dnmt1, HMM indicates a significant *de novo* methylation probability in ESCs (up to 15%), depending on the repetitive element/sequence. *In vitro* experiments analysing the methylation activity of Dnmt1 show 2 to 50 fold higher preference for hemimethylated CpG dyads, dependent on the substrate or conditions (45). Pre-existing methylated sites *in vitro* are shown to enhance the *de novo* methylation efficiency of Dnmt1 (46, 47). Our data corroborate these observations *in vivo*, linking an increased methylation to a higher *de novo* methylation activity of Dnmt1 (IAP, mSat > L1). Differential regulation of the CXXC domain binding capacities at the repetitive elements could osculate the *de novo* methylation activity of Dnmt1 (48). The fidelity for Dnmt1 to methylate hemi-mCpG dyads was shown to be 95% to 96% *in vitro* (47). Our HMM predicts fidelities of methylating hemi-mCpGs for IAP and mSat of 0.90 and 0.93, respectively, which fits quite well to the *in vitro* data. For a comparison with the *in vitro* derived data, we did not discriminate between *de novo* and maintenance methylation at hemimethylated CpG dyads (see values in Supplementary Table 3). At L1 elements, the fidelity decreases to 75%, according to our HMM mainly due to missing *de novo* methylation. Lower fidelity in comparison to *in vitro* derived data could arise from the presence of 5hmC at L1 elements (as discussed above).

DNA methylation patterns of Np95 and Dnmt1 KO are very similar underpinning their strong functional link in maintenance methylation. Nevertheless a few differences in the KO patterns of both enzymes point towards functional differences. E.g. hemi-mCpG sites on the upper and lower strand differ between both KOs. Np95 KO shows a 2 fold enriched proportion of mosaical hemimethylation compared to Dnmt1. This enrichment is seen at IAP and mSat sequences, where Dnmt1 show relatively high *de novo* methylation activity in HMM but is not so pronounced at B1 and L1 elements, where HMM only deduces a rather low Dnmt1 *de novo* methylation activity. This suggests that the distributive *de novo* methylation function of Dnmt1 is not influenced by Np95. Accordingly, compared to Dnmt1 KO, Np95 KO still allows the distributive *de novo* methylation by Dnmt1, leading to more hemimethylated sites distributed on both sites of the strand.

For Dnmt3a and Dnmt3b *de novo* and maintenance methylation contributions differ extensively across sequence elements. Such context dependent effects were not addressed in former *in situ* and *in vitro* modelling studies (7, 17) and may become only evident in the native chromatin context. HMM revealed a significant "maintenance" contribution for both enzymes at repetitive elements. The contribution of Dnmt3a and 3b to maintain full methylation at CpG dyads following replication might be attributed to targeted and enhanced *de novo* activity stimulated by the presence of CpG methylation density. Some studies show that Dnmt3a/3b can strongly bind to nucleosomes containing (hemi)methylated DNA (49, 50). By this Dnmt3a/3b could be triggered to "*de novo*" methylate hemimethylated sites following replication to maintain full methylation in the absence of Dnmt1. Indeed in Dnmt3a/3b fractions bound to nucleosomes SINEs and LINEs were found to be enriched and for these elements we observe high need for Dnmt3a/3b to maintain methylation. The activity of Dnmt3a and 3b is enhanced by Dnmt3L (11) apparently also for single stranded modification as apparent from the distribution of CpA sites which we almost exclusively detect only one of the two DNA strands in single reads. Finally, it is conceivable to speculate from our data that NP95 may also influence the methylation activity of Dnmt3a/b. Interactions of NP95 with both enzymes have been observed (16).

Conclusion

In summary our comparative deep sequencing analysis including Dnmt KO ESCs revealed a complex scenario of sequence, element and cell specific control of DNA methylation pattern control at CpG dyads. The data allowed to construct greatly improved HMMs to infer DNA methylation dynamics and efficiencies for the *de novo* and maintenance methylation function for all three Dnmts. Most strikingly we were able to allocate CpA methylation of major satellites in ESCs to distinct activities of Dnmt3a and 3b strongly stimulated by Dnmt3L.

1. Yoder, J. A., Walsh, C. P. & Bestor, T. H. (1997) *Trends Genet* **13,** 335-40.

2. Liang, G., Chan, M. F., Tomigahara, Y., Tsai, Y. C., Gonzales, F. A., Li, E., Laird, P. W. & Jones, P. A. (2002) *Mol Cell Biol* **22,** 480-91.

3. Okano, M., Bell, D. W., Haber, D. A. & Li, E. (1999) *Cell* **99,** 247-57.

4. Ehrlich, M., Gama-Sosa, M. A., Huang, L. H., Midgett, R. M., Kuo, K. C., McCune, R. A. & Gehrke, C. (1982) *Nucleic Acids Res* **10,** 2709-21.

5. Gama-Sosa, M. A., Midgett, R. M., Slagel, V. A., Githens, S., Kuo, K. C., Gehrke, C. W. & Ehrlich, M. (1983) *Biochim Biophys Acta* **740,** 212-9.

6. Bourc'his, D. & Bestor, T. H. (2004) *Nature* **431,** 96-9.

7. Okano, M., Xie, S. & Li, E. (1998) *Nat Genet* **19,** 219-20.

8. Hermann, A., Goyal, R. & Jeltsch, A. (2004) *J Biol Chem* **279,** 48350-9.

9. Gowher, H. & Jeltsch, A. (2001) *J Mol Biol* **309,** 1201-8.

10. Gowher, H. & Jeltsch, A. (2002) *J Biol Chem* **277,** 20409-14.

11. Holz-Schietinger, C. & Reich, N. O. (2010) *J Biol Chem* **285,** 29091-100.

12. Bourc'his, D., Xu, G. L., Lin, C. S., Bollman, B. & Bestor, T. H. (2001) *Science* **294,** 2536-9.

13. Gowher, H., Liebert, K., Hermann, A., Xu, G. & Jeltsch, A. (2005) *J Biol Chem* **280,** 13341-8.

14. Bostick, M., Kim, J. K., Esteve, P. O., Clark, A., Pradhan, S. & Jacobsen, S. E. (2007) *Science* **317,** 1760-4.

15. Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T. A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., Tajima, S., Mitsuya, K., Okano, M. & Koseki, H. (2007) *Nature* **450,** 908-12.

16. Meilinger, D., Fellinger, K., Bultmann, S., Rothbauer, U., Bonapace, I. M., Klinkert, W. E., Spada, F. & Leonhardt, H. (2009) *EMBO Rep*.

17. Sontag, L. B., Lorincz, M. C. & Georg Luebeck, E. (2006) *J Theor Biol* **242,** 890-9.

18. Lacey, M. R. & Ehrlich, M. (2009) *Stat Appl Genet Mol Biol* **8,** Article 40.

19. Otto, S. P. & Walbot, V. (1990) *Genetics* **124,** 429-37.

20. Pfeifer, G. P., Steigerwald, S. D., Hansen, R. S., Gartler, S. M. & Riggs, A. D. (1990) *Proc Natl Acad Sci U S A* **87,** 8252-6.

21. Laird, C. D., Pleasant, N. D., Clark, A. D., Sneeden, J. L., Hassan, K. M., Manley, N. C., Vary, J. C., Jr., Morgan, T., Hansen, R. S. & Stoger, R. (2004) *Proc Natl Acad Sci U S A* **101,** 204-9.

22. Genereux, D. P., Miner, B. E., Bergstrom, C. T. & Laird, C. D. (2005) *Proc Natl Acad Sci U S A* **102,** 5802-7.

23. Jeong, K. S. & Lee, S. (2005) *Biochem Biophys Res Commun* **335,** 1211-6.

24. Li, E., Bestor, T. H. & Jaenisch, R. (1992) *Cell* **69,** 915-26.

25. Lehnertz, B., Ueda, Y., Derijck, A. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T. & Peters, A. H. (2003) *Curr Biol* **13,** 1192-200.

26. Lei, H., Oh, S. P., Okano, M., Juttermann, R., Goss, K. A., Jaenisch, R. & Li, E. (1996) *Development* **122,** 3195-205.

27. Hata, K., Okano, M., Lei, H. & Li, E. (2002) *Development* **129,** 1983-93.

28. Lutsik, P., Feuerbach, L., Arand, J., Lengauer, T., Walter, J. & Bock, C. (2011) *Nucleic Acids Res* **39 Suppl 2,** W551-6.

29. Burden, A. F., Manley, N. C., Clark, A. D., Gartler, S. M., Laird, C. D. & Hansen, R. S. (2005) *J Biol Chem* **280,** 14413-9.

30. Shao, C., Lacey, M., Dubeau, L. & Ehrlich, M. (2009) *Epigenetics* **4,** 165-75.

31. Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., Nery, J. R., Lee, L., Ye, Z., Ngo, Q. M., Edsall, L., Antosiewicz-Bourget, J., Stewart, R., Ruotti, V., Millar, A. H., Thomson, J. A., Ren, B. & Ecker, J. R. (2009) *Nature* **462,** 315-22.

32. Hayatsu, H. & Shiragami, M. (1979) *Biochemistry* **18,** 632-7.

33. Huang, Y., Pastor, W. A., Shen, Y., Tahiliani, M., Liu, D. R. & Rao, A. (2010) *PLoS One* **5,** e8888.

34. Kriaucionis, S. & Heintz, N. (2009) *Science* **324,** 929-30.

35. Kinney, S. M., Chin, H. G., Vaisvila, R., Bitinaite, J., Zheng, Y., Esteve, P. O., Feng, S., Stroud, H., Jacobsen, S. E. & Pradhan, S. (2011) *J Biol Chem*.

36. Valinluck, V. & Sowers, L. C. (2007) *Cancer Res* **67,** 946-50.

37. Frauer, C., Hoffmann, T., Bultmann, S., Casa, V., Cardoso, M. C., Antes, I. & Leonhardt, H. (2011) *PLoS One* **6,** e21306.

38. Ficz, G., Branco, M. R., Seisenberger, S., Santos, F., Krueger, F., Hore, T. A., Marques, C. J., Andrews, S. & Reik, W. (2011) *Nature* **473,** 398-402.

39. Ito, S., D'Alessio, A. C., Taranova, O. V., Hong, K., Sowers, L. C. & Zhang, Y. (2010) *Nature* **466,** 1129-33.

40. Lane, N., Dean, W., Erhardt, S., Hajkova, P., Surani, A., Walter, J. & Reik, W. (2003) *Genesis* **35,** 88-93.

41. Wossidlo, M., Arand, J., Sebastiano, V., Lepikhov, K., Boiani, M., Reinhardt, R., Scholer, H. & Walter, J. (2010) *Embo J* **29,** 1877-88.

42. Ramsahoye, B. H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A. P. & Jaenisch, R. (2000) *Proc Natl Acad Sci U S A* **97,** 5237-42.

43. Lister, R., Pelizzola, M., Kida, Y. S., Hawkins, R. D., Nery, J. R., Hon, G., Antosiewicz-Bourget, J., O'Malley, R., Castanon, R., Klugman, S., Downes, M., Yu, R., Stewart, R., Ren, B., Thomson, J. A., Evans, R. M. & Ecker, J. R. (2011) *Nature* **471,** 68-73.

44. Su, A. I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K. A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., Cooke, M. P., Walker, J. R. & Hogenesch, J. B. (2004) *Proc Natl Acad Sci U S A* **101,** 6062-7.

45. Goyal, R., Reinhardt, R. & Jeltsch, A. (2006) *Nucleic Acids Res* **34,** 1182-8.

46. Tollefsbol, T. O. & Hutchison, C. A., 3rd (1997) *J Mol Biol* **269,** 494-504.

47. Vilkaitis, G., Suetake, I., Klimasauskas, S. & Tajima, S. (2005) *J Biol Chem* **280,** 64-72.

48. Song, J., Rechkoblit, O., Bestor, T. H. & Patel, D. J. (2011) *Science* **331,** 1036-40.

49. Jeong, S., Liang, G., Sharma, S., Lin, J. C., Choi, S. H., Han, H., Yoo, C. B., Egger, G., Yang, A. S. & Jones, P. A. (2009) *Mol Cell Biol* **29,** 5366-76.

50. Sharma, S., De Carvalho, D. D., Jeong, S., Jones, P. A. & Liang, G. (2011) *PLoS Genet* **7,** e1001286.
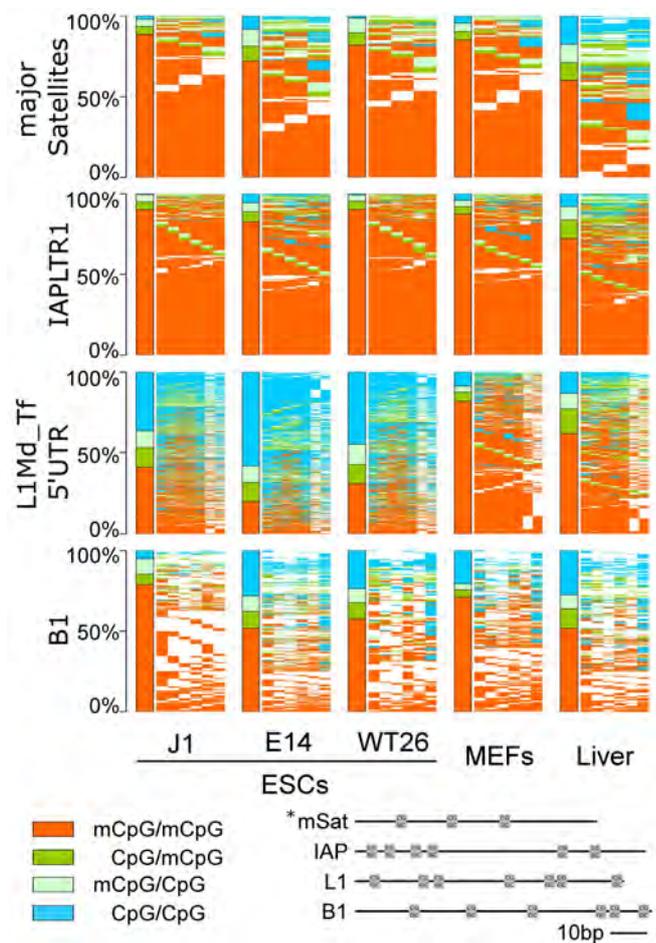
Figures:

Fig. 1 DNA methylation pattern of CpG dyads in WT ESCs, MEFs and embryonic liver

The bars sum up the DNA methylation status of all CpG dyads. The map next to the bar represents the distribution of methylated sites. Each column shows neighboured CpG dyads, and each line represents one sequence read. Red visualises fully methylated CpG dyads, light green and dark green hemi-mCpG dyads on the upper and lower strand, respectively, blue shows unmethylated CpG dyads and white strand for positions which are mutated or due sequence quality not analysable. The reads in the map are sorted first by fully methylated sites and then by hemi-m CpG dyads. *This picture shows the distance of the CpG dyads to each other.
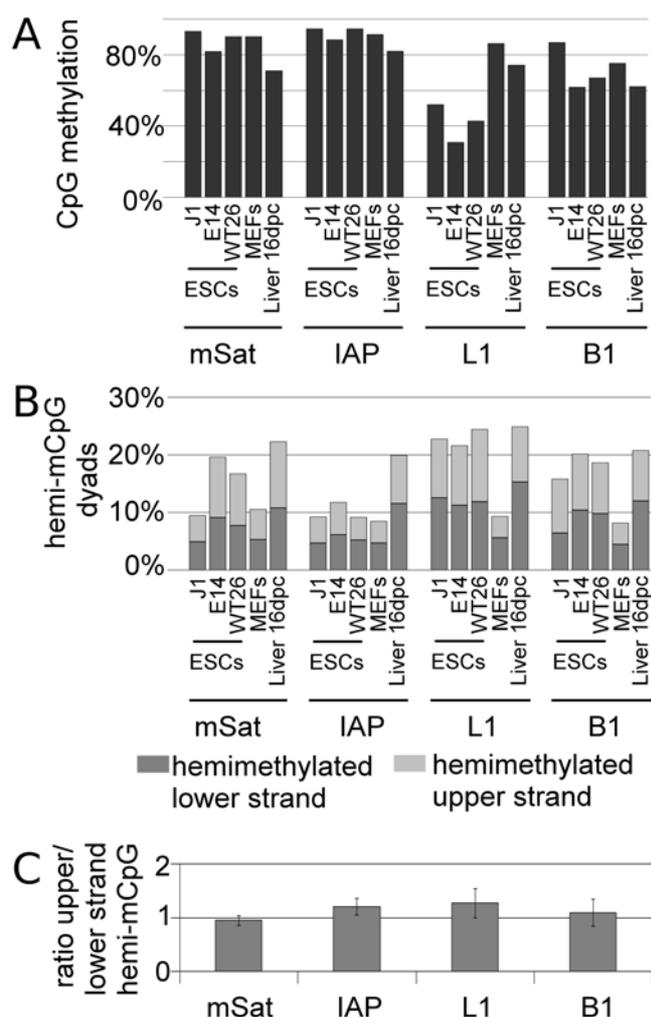
Fig. 2: DNA Methylation of WT cells at major Satellites (mSat), IAPLTR (IAP), L1Md_Tf (L1) and B1

A: overall methylation level of WT cells Overall methylation level was calculated from all CpG dyads, hemimethylated sites assessed as 1 methylated site and fully methylated sites as 2 methylated sites. B: amount of hemimethylated sites splitted by occurrence of the methylation on the upper or lower strand. C: relation of hemimethylated sites methylated on the upper or lower strand
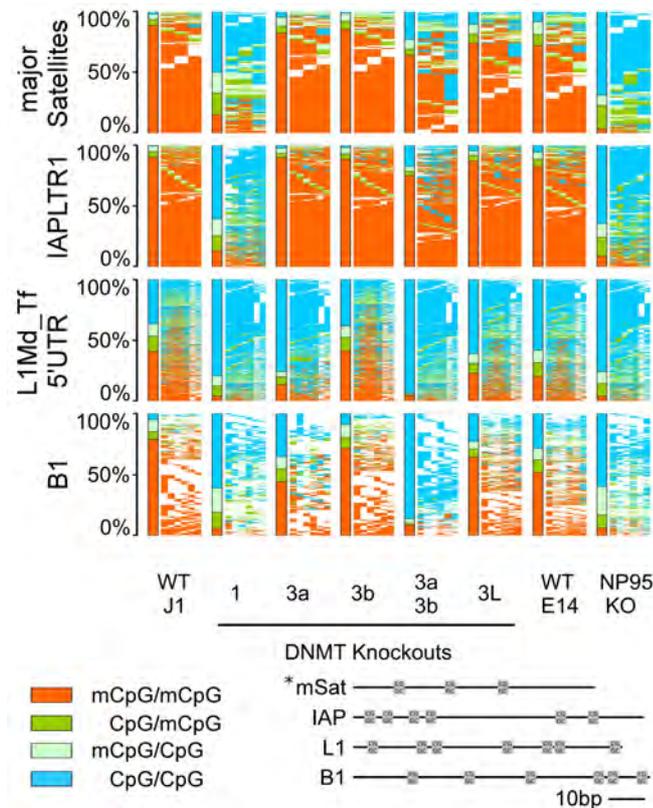
Fig. 3 Methylation pattern map of ESCs depleted in Dnmts or factors of the methylation machinery

The bars sum up the methylation status of all CpG dyads. The map next to the bar represents the distribution of methylated sites. Each column shows neighboured CpG dyads, and each line represents one sequence read. Red visualises fully methylated CpG dyads, light green and dark green hemi-mCpG dyads on the upper and lower strand, respectively, blue shows unmethylated CpG dyads and white strand for positions which are mutated or due sequence quality not analysable. The reads in the map are sorted first by fully methylated sites and then by hemi-m CpG dyads. *This picture shows the distance of the CpG dyads to each other.
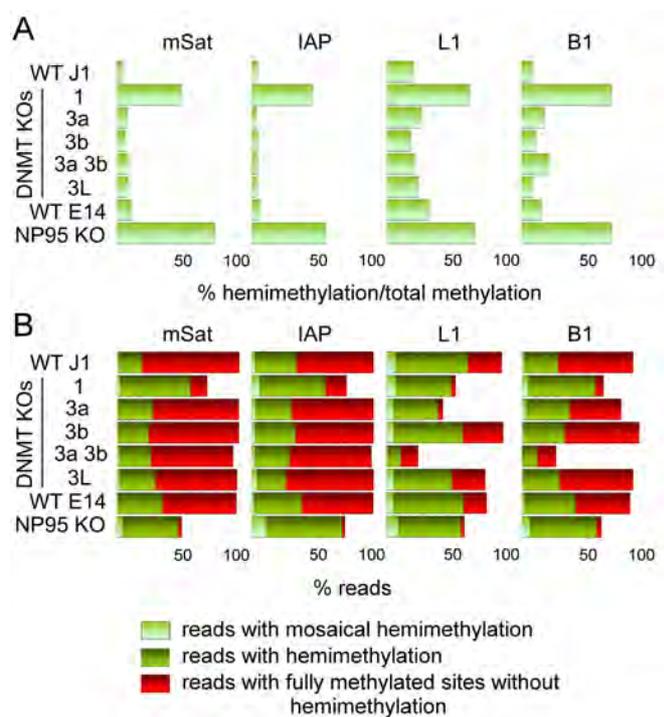
Fig. 4 Hemimethylated sites in ESCs depleted in DNMTs or factors of the methylation machinery

A: Relative amount of hemimethylated sites The relative amount of hemimethylated sites was calculated by dividing the amount of hemimethylated sites with the overall methylation.

B: Distribution of methylated/hemimethylated sites along reads Reads showing any methylation contributed to the fraction "reads with methylated sites". The fraction of these sequences with any hemimethylated site is classified as fraction of "reads with hemimethylated sites". Out of these reads which show hemimethylated sites on the upper and lower strand at the same time contribute to the fraction of "reads with hemimethylated sites on both strands".
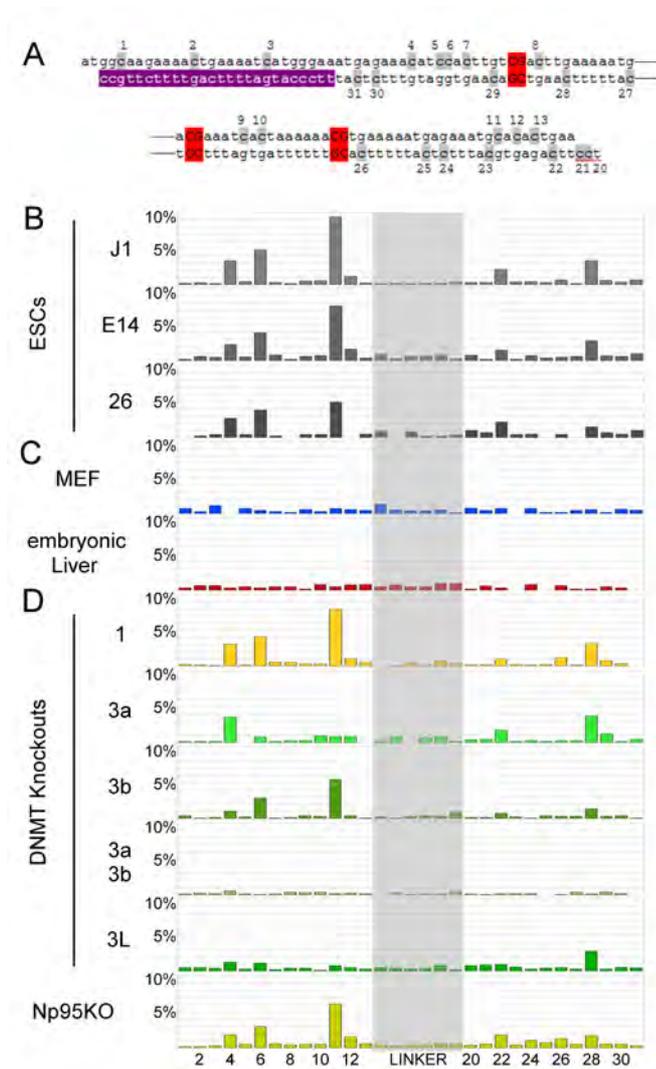
Fig. 5 non-CpG methylation of major Satellites

A major satellite genomic sequence of the hairpin bisulfite PCR product. Non CpGs are marked grey with the corresponding number attached. CpGs are marked red. Purple shows the location of the lower primer.

B-D non-CpG methylation of mESCs (B), differentiated cells (C) and Dnmt and Np95 KOs (D) at major satellites. Only CpA positions show methylation up to ten percent (position 4, 6, 11, 22 and 28). Knockouts of Dnmt3 family shows decrease of CpA methylation on different sites.

Fig. 6: Correlation of CpA and CpG methylation. A: The graph represents the relative amount of reads per CpG methylation level; in grey for reads showing no CpA methylation and in black for reads showing CpA methylation. Re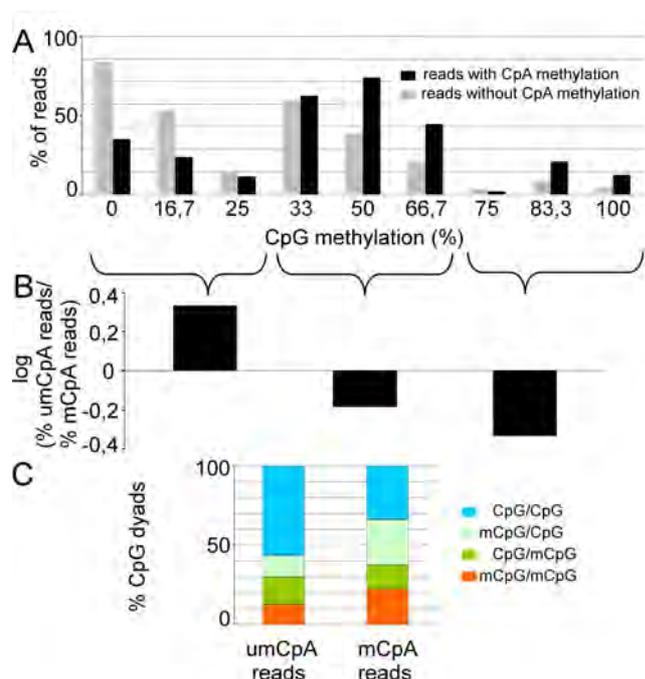ads showing CpA methylation are depleted in the fraction of reads with low CpG methylation level and enriched in reads showing 50% or more CpG methylation. **B:** ratio between non CpA methylated reads and CpA methylated reads divided into three fraction by CpG methylation level (0-25%, 33-66,7%, 75-100%). **C:** Distribution of un-, hemi- and fully methylated CpG dyads in the reads showing CpA methylation or no CpA methylation. The fraction of reads showing CpA methylation is enriched in fully and hemimethylated CpG dyads. Hemimethylated CpG dyads, methylated on the upper strand, are mainly enriched. Interestingly, on the upper strand, we also observe the main part of CpA methylation.
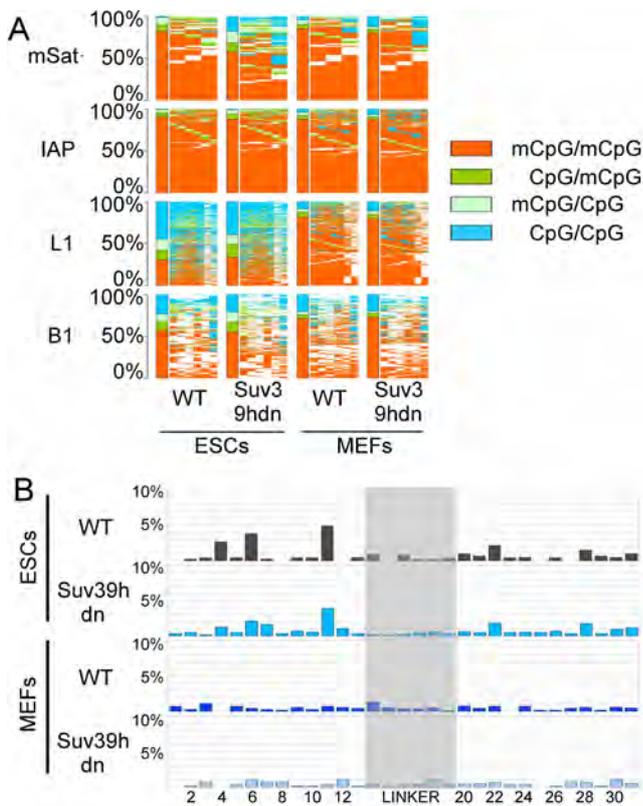
Fig. 7 Influence of Suv39h double Knockout on methylation pattern

A Methylation pattern map of CpG dyads The bars sum up the methylation status of all CpG dyads. The map next to the bar represents the distribution of methylated sites. Each column shows neighboured CpG dyads, and each line represents one sequence read.

B non CpG methylation. We could not detect a changed non CpG methylation pattern in the Suv39KO cells.

112
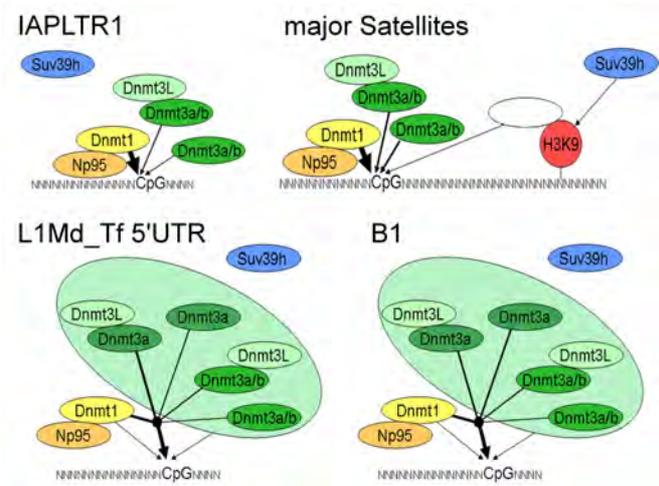
Fig. 8 Model for CpG methylation at different repetitive elements

The arrows' thickness shows the methylation efficiencies of the enzymes and enzyme complexes.

Dnmt3a/b in one circle represents compensable function of both. The black circle represents cooperativity of the enzymes and enzyme complexes. The influence of enzymes without any linkage is analysed and not found.

Fig. 9 Model for CpA methylation at major satellites

The thickness of the arrows shows the methylation efficiencies of the enzymes and enzyme complexes. The influence of enzymes without any linkage is analysed and not found. Two arrows based on one enzyme represent different methylation activities depending on the CpA position.

Cooperative DNA and histone binding by Uhrf2 links the two major repressive epigenetic pathways

2.5 Cooperative DNA and histone binding by Uhrf2 links the two major repressive epigenetic pathways

Cooperative DNA and Histone Binding by Uhrf2 Links the Two Major Repressive Epigenetic Pathways

Garwin Pichler, Patricia Wolf, Christine S. Schmidt, Daniela Meilinger, Katrin Schneider, Carina Frauer, Karin Fellinger, Andrea Rottach, and Heinrich Leonhardt*

Ludwig Maximilians University Munich, Department of Biology II and Center for Integrated Protein Science Munich (CIPS^M), Großhaderner Str. 2, 82152 Planegg-Martinsried, Germany

ABSTRACT

Gene expression is regulated by DNA as well as histone modifications but the crosstalk and mechanistic link between these epigenetic signals are still poorly understood. Here we investigate the multi-domain protein Uhrf2 that is similar to Uhrf1, an essential cofactor of maintenance DNA methylation. Binding assays demonstrate a cooperative interplay of Uhrf2 domains that induces preference for hemimethylated DNA, the substrate of maintenance methylation, and enhances binding to H3K9me3 heterochromatin marks. FRAP analyses revealed that localization and binding dynamics of Uhrf2 in vivo require an intact tandem Tudor domain and depend on H3K9 trimethylation but not on DNA methylation. Besides the cooperative DNA and histone binding that is characteristic for Uhrf2, we also found an opposite expression pattern of *uhrf1* and *uhrf2* during differentiation. While *uhrf1* is mainly expressed in pluripotent stem cells, *uhrf2* is upregulated during differentiation and highly expressed in differentiated mouse tissues. Ectopic expression of Uhrf2 in *uhrf1*$^{-/-}$ embryonic stem cells did not restore DNA methylation at major satellites indicating functional differences. We propose that the cooperative interplay of Uhrf2 domains may contribute to a tighter epigenetic control of gene expression in differentiated cells. J. Cell. Biochem. 112: 2585–2593, 2011. © 2011 Wiley-Liss, Inc.

KEY WORDS: UHRF1; UHRF2; DNA METHYLATION; HISTONE MODIFICATIONS; EPIGENETICS

DNA methylation and histone modifications are major epigenetic marks involved in the regulation of gene expression, inheritance of chromatin states, genome stability, and cellular differentiation [Bird, 2002; Kouzarides, 2007; Reik, 2007]. Misregulation of epigenetic pathways, like erroneous DNA methylation, may lead to cancer and other diseases [Jones and Baylin, 2007]. Open questions concern the crosstalk and mechanistic link between different epigenetic signals.

Genome-scale DNA methylation studies revealed a connection between DNA methylation and histone modifications. Specifically, DNA methylation correlates with the absence of H3K4 methylation and presence of H3K9 methylation [Meissner et al., 2008]. This correlation may in part be caused by DNA methyltransferases specifically recognizing histone modifications. For instance, the de novo DNA methyltransferase Dnmt3a and its cofactor Dnmt3L specifically recognize unmethylated H3K4 mediated by the ATRX-Dnmt3-Dnmt3L (ADD) domain [Ooi et al., 2007; Otani et al., 2009]. Dnmt1, which is involved in maintenance methylation during DNA

replication and DNA repair [Leonhardt et al., 1992; Mortusewicz et al., 2005], specifically methylates hemimethylated DNA [Bestor and Ingram, 1983; Pradhan et al., 1997] and associates with constitutive heterochromatin via its targeting sequence (TS) domain [Easwaran et al., 2004].

Recently, Uhrf1 (also known as Np95 or ICBP90) has been shown to link DNA and histone modifications and has emerged as an essential cofactor for the maintenance of genomic DNA methylation. Genetic ablation of *uhrf1* leads to remarkable genomic hypomethylation, a phenotype similar to *dnmt1*$^{-/-}$ embryonic stem cells (ESCs) [Bostick et al., 2007; Sharif et al., 2007]. Uhrf1 binds hemimethylated DNA via a SET and RING associated domain (SRA) domain and targets Dnmt1 to its substrate of maintenance DNA methylation [Bostick et al., 2007; Sharif et al., 2007; Arita et al., 2008; Avvakumov et al., 2008; Hashimoto et al., 2008; Qian et al., 2008; Rottach et al., 2010]. This targeting activity of Uhrf1 is based on specific binding to the heterochromatin mark H3K9me3 via a tandem Tudor domain (TTD) [Karagianni et al., 2008; Rottach et al.,

2010]. In addition, Uhrf1 interacts with Dnmt3a and Dnmt3b and with histone modifying enzymes like HDAC1, G9a, and Tip60 [Unoki et al., 2004; Achour et al., 2009; Kim et al., 2009; Meilinger et al., 2009]. Finally, Uhrf1 displays E3 ubiquitin ligase activity for histone H3 [Citterio et al., 2004] and is involved in large scale reorganization of chromocenters [Papait et al., 2008].

Interestingly, a second member of the Uhrf family, Uhrf2, harbors similar domains [Bronner et al., 2007]. Until now, the only known function of Uhrf2 is a role in intranuclear degradation of polyglutamine aggregates [Iwata et al., 2009]. In this study, we systematically investigated the function and interplay of distinct Uhrf2 domains in DNA and histone tail substrate recognition and report first hints on cell-type specific functions of Uhrf1 and Uhrf2.

MATERIALS AND METHODS

EXPRESSION CONSTRUCTS

Expression constructs for GFP, RFP-PCNA, Uhrf1-GFP, and GFP constructs of Dnmt1 were described previously [Sporbert et al., 2005; Fellinger et al., 2009; Meilinger et al., 2009]. All Uhrf2 expression constructs were derived by PCR from mouse *uhrf2*-myc cDNA (MR210744, ORIGENE). To obtain GFP fusion constructs, the *uhrf1* cDNA [Rottach et al., 2010] was replaced by *uhrf2* encoding PCR fragments in the pCAG-*uhrf1*-GFP vector. The deletion and point mutant expression constructs were derived from the corresponding wild-type constructs by overlap extension PCR [Ho et al., 1989] and PCR-based mutagenesis. The following start and end amino acids were chosen: Uhrf2 tandem Tudor domain, amino acids 118–312; Uhrf2 PHD domain, amino acids 325–395; Uhrf2 tandem Tudor–PHD domain, amino acids 118–395; Uhrf1 tandem Tudor–PHD domain, amino acids 121–370. The linker exchange constructs were derived by PCR using overlapping primers that contained the partial linker sequence. Amino acid sequences of the linkers: Uhrf1: KERRPLIASPSQPPA; Uhrf2: GAHPISFADGKF. All constructs were verified by DNA sequencing. Throughout this study enhanced GFP constructs were used and for simplicity referred to as GFP fusions.

CELL CULTURE, TRANSFECTION, CELL SORTING, AND DIFFERENTIATION

HEK293T cells, MEFs, and ESCs were cultured and transfected as described [Schermelleh et al., 2007; Rottach et al., 2010] with the exception that Lipofectamin (Invitrogen) was used for transfection of MEFs. E14 *uhrf1*$^{-/-}$ ESCs were transfected with Uhrf1-GFP and Uhrf2-GFP expression constructs using FuGENE HD (Roche) according to the manufacturer's instructions. ESCs were sorted for GFP positive cells 48 h after transfection with a FACS Aria II instrument (Becton Dikinson). ESC strains wt E14, wt J1, and E14 *uhrf1*$^{-/-}$ were cultured and differentiated to embryoid bodies as described [Szwagierczak et al., 2010]. The ESC strain wt JM8A3.N1 (EUCOMM, Germany) was cultured in Knockout D-MEM (Gibco-BRL, Grand-Island, NY) medium containing 10% fetal bovine serum (PAA Laboratories GmbH, Austria), 0.1 mM β-mercaptoethanol (Gibco-BRL), 2 mM L-glutamine, 100 U/ml penicillin, 100 μg/ml streptomycin (PAA Laboratories GmbH). The medium was supple-mented with 1,000 U/ml recombinant mouse LIF (Millipore, Temecula, CA).

RNA ISOLATION, CDNA SYNTHESIS, AND QUANTITATIVE REAL–TIME PCR

RNA isolation and cDNA synthesis were performed as described [Szwagierczak et al., 2010]. Equal amounts of cDNA were used for Real-time PCR with TaqMan Gene Expression Master Mix (Applied Biosystems) on the 7500 Fast Real-time PCR System (Applied Biosystems) according to the manufacturer's instructions. The following TaqMan Gene expression assays were used: Gapdh (Assay ID: Mm99999915_g1), uhrf1 (Assay ID: Mm00477865_m1) and uhrf2 (Assay ID: Mm00520043_m1). Gene expression levels were normalized to Gapdh and calculated using the comparative C_T Method (ΔΔC_T Method).

IN VITRO DNA BINDING AND HISTONE–TAIL PEPTIDE BINDING ASSAY

The in vitro binding assays were performed as described previously [Frauer and Leonhardt, 2009; Rottach et al., 2010]. NoCpG DNA substrates were produced in a primer extension reaction [Frauer and Leonhardt, 2009] others by hybridization of two DNA oligos (Supplementary Fig. S7B–D). Histone-tail peptides were purchased as TAMRA conjugates (PSL, Germany; Supplementary Fig. S7A). Peptides were added in a molar ratio 1.5:1 (peptide/GFP fusion) and the binding reaction was performed at RT for 15 min with constant mixing. For combined assays, samples were additionally incubated with either H3K9me3 or H3K9ac histone-tail peptides in a molar ratio 1.5:1 (peptide/GFP fusion) or increasing amount of DNA substrate as indicated. The binding reaction was performed at RT for 60 min with constant mixing.

IMMUNOFLOURESCENCE STAINING AND ANTIBODIES

For immunostaining, MEF cells and ESCs were grown on cover slips and transiently transfected with Uhrf2-GFP (MEF cells), or co-transfected with Uhrf2-GFP and RFP-PCNA (ESCs). Cells were fixed with 2.0% or 3.7% formaldehyde in PBS and permeabilized in PBS containing 0.2% Triton X-100. The post-translational histone modification H3K9me3 was detected via a rabbit primary antibody (Active Motif) and a secondary anti-rabbit antibody conjugated to Alexa Fluor 594 (Molecular Probes, Eugene, OR). The antibodies were diluted 1:1,000 or 1:500, respectively, in PBS containing 0.02% Tween-20 and 2% BSA. GFP-Binder (ChromoTek, Germany) was used to boost GFP signals and was labeled with Alexa Fluor 488. Cells were counterstained with DAPI and mounted in Vectashield (Vector Laboratories, Burlingame, CA). Images of the cells were obtained using a TCS SP5 AOBS confocal laser scanning microscope (Leica, Wetzlar, Germany) with a 63x/1.4 NA Plan-Apochromat oil immersion objective. GFP, Alexa Fluor 488, RFP, and Alexa Fluor 594 were excited with a 488-nm argon laser and a 561-nm diode laser, respectively. Image series were recorded with a frame size of 512×512 pixels, a pixel size of 100 nm and with a detection pinhole size of 1 Airy Unit.

LIVE CELL MICROSCOPY AND FLUORESCENCE RECOVERY AFTER PHOTOBLEACHING (FRAP) ANALYSIS

Live cell imaging and FRAP analyses were performed as described [Schermelleh et al., 2007] with the exception that imported images were intensity normalized, converted to 8-bit and Gauss-filtered (2 pixel radius). Data sets showing lateral movement were corrected by image registration using the StackReg plug-in of ImageJ [Abramoff et al., 2004] starting with a frame when approximately half recovery was reached. Within the first 30 s after bleaching, images were taken every 150 ms and then in intervals of 1 s.

DNA METHYLATION ANALYSIS

Genomic DNA was isolated with the QIAmp DNA Mini Kit (Qiagen) and 1.5 μg were bisulfite converted using the EZ DNA Methylation-Gold Kit (Zymo research) according to the manufacturer's instructions. Primer sequences for major satellites were AAAAT-GAGAAACATCCACTTG (forward primer) and CCATGATTTT-CAGTTTTCTT (reverse primer). For amplification we used Qiagen Hot Start Polymerase in 1× Qiagen Hot Start Polymerase buffer supplemented with 0.2 mM dNTPs, 0.2 μM forward primer, 0.2 μM reverse primer, 1.3 mM betaine (Sigma) and 60 mM tetramethy-lammonium-chloride (TMAC, Sigma). Major satellites were amplified in a single amplification and pyrosequencing reactions were carried out by Varionostic GmbH (Ulm, Germany).

STATISTICAL ANALYSIS

Results were expressed as means ± SD or means ± SEM. The difference between two mean values was analyzed by Student's t-test and was considered as statistically significant in case of $P < 0.05$ (*) and highly significant for $P < 0.001$ (**).

RESULTS

OPPOSITE EXPRESSION PATTERN OF *UHRF1* AND *UHRF2* DURING DIFFERENTIATION

Recently, Uhrf1 has emerged as an essential factor for the maintenance of DNA methylation. Sequence analyses revealed that Uhrf2 harbors five recognizable domains similar to Uhrf1 (Fig. 1A), but its role in the regulation of DNA methylation is still unclear. We compared the expression pattern of *uhrf1* and *uhrf2* in ESCs and somatic cells, during differentiation and in differentiated mouse tissues (Fig. 1B–D and Supplementary Fig. S1). Interestingly, both genes show opposite expression patterns; while *uhrf1* is expressed in ESCs and down regulated during differentiation, which is consistent with previous reports [Muto et al., 1995; Fujimori et al., 1998; Hopfner et al., 2000], *uhrf2* is upregulated and highly expressed in differentiated mouse tissues. The switch in the expression pattern argues against a functional redundancy of both genes and is consistent with the drastic loss of DNA methylation in *uhrf1*$^{-/-}$ ESCs despite the presence of intact *uhrf2* alleles. Therefore, the opposite expression pattern of both genes suggests different functional roles of *uhrf1* and *uhrf2* in development.
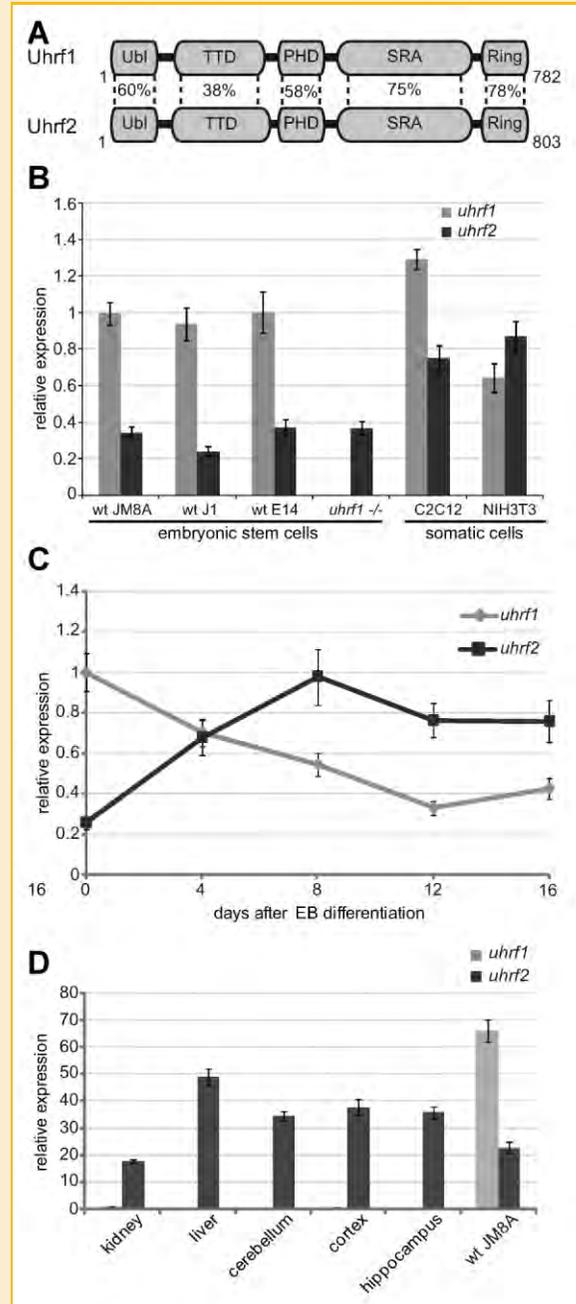

Fig. 1. Opposite expression pattern of *uhrf1* and *uhrf2* during differentiation. A: Schematic outline of the multi-domain architecture of Uhrf1 in comparison to Uhrf2. An N-terminal ubiquitin-like domain (Ubl) is followed by a tandem Tudor domain (TTD), a plant homeodomain (PHD), a SET and RING associated (SRA) domain and a C-terminal really interesting new gene (RING) domain. Numbers indicate primary sequence similarities of single domains determined by BlastP search [Altschul, 1991]. Expression analysis of *uhrf1* and *uhrf2* by Real-time PCR in ESCs and somatic cells (B), during differentiation of wt J1 ESCs (C) and in various adult mouse tissues in comparison to the expression data in ESCs (D). Expression levels are relative to *uhrf1* in wtJM8A (B), day 0 of differentiation (C) and to kidney (D) (*uhrf1* set to 1). Shown are means ± SD of at least two independent experiments.

COOPERATIVE BINDING OF REPRESSIVE EPIGENETIC MARKS BY UHRF2

To investigate DNA and histone-tail binding preferences of Uhrf2 in vitro, we used a versatile binding assay developed for GFP fusion proteins [Rothbauer et al., 2008; Frauer and Leonhardt, 2009;

Rottach et al., 2010]. Similar to Uhrf1, histone-tail peptide binding assays revealed that Uhrf2 preferentially binds to H3(1–20) and H3K9me3 peptides (Fig. 2A). This binding activity of Uhrf2 is mediated by the TTD but not the PHD domain (Fig. 2B). Consistently, acetylation of H3K9, underrepresented in heterochromatin,
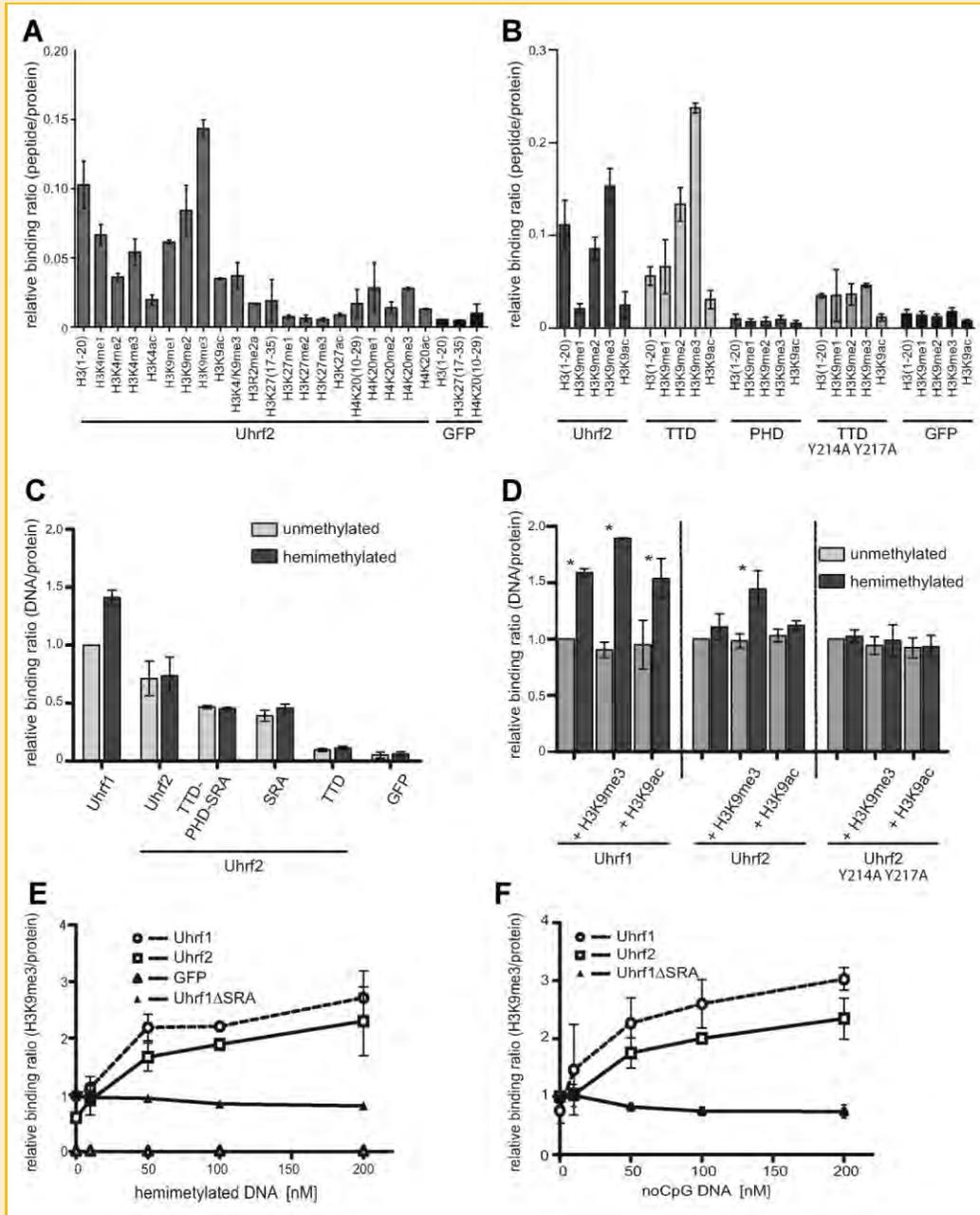

Fig. 2. Cooperative binding of repressive epigenetic marks by Uhrf2. In vitro binding ratios of fluorescently labeled substrate over bound GFP fusion proteins were determined. A: Histone H3- and H4-tail binding specificities of Uhrf2. Shown are means ± SD of biological duplicates. B: Histone H3 tail binding specificity of Uhrf2, its tandem Tudor domain (TTD), its PHD domain and its TTD mutant (Y214A Y217A). Shown are means ± SEM of at least three independent experiments. C: DNA binding properties of Uhrf1, Uhrf2 and of single (SRA, TTD) and combined Uhrf2 domains (TTD–PHD–SRA). Shown are means ± SEM of three independent experiments. D: DNA binding properties of Uhrf1, Uhrf2 and Uhrf2 Y214A Y217A in combination with histone–tail peptide binding. Shown are means ± SD of three independent experiments (Uhrf1, Uhrf2) and of two independent experiments (Uhrf2 Y214A Y217A). Values were normalized to the binding ratio of each GFP fusion for unmethylated DNA without histone–tail peptide. Statistical significance of differences between the binding ratios with un- and hemimethylated DNA is indicated; *$P < 0.05$. E + F: H3K9me3 peptide binding by Uhrf1, Uhrf2, and Uhrf1ΔSRA with increasing concentrations of DNA substrate containing either one central hemimethylated (E) or noCpG site (F). Shown are means ± SD of biological duplicates. Values were normalized to the binding ratio of Uhrf1ΔSRA without DNA.

prevented the binding of Uhrf2 and its TTD. The binding of Uhrf1 to H3K9me3 is mediated by an aromatic cage in the TTD [Rottach et al., 2010]. Site-directed mutagenesis of Uhrf2 changing the two conserved tyrosine residues to alanine (Y214A Y217A) (Supplementary Fig. S2) abolished specific peptide binding (Fig. 2B) and supports a function of the aromatic cage in H3K9me3 recognition.

Whereas Uhrf1 preferentially binds to hemimethylated DNA, Uhrf2 failed to show a preference for hemi-over unmethylated DNA (Fig. 2C). These differences in DNA binding preferences between Uhrf1 and Uhrf2 were confirmed by electrophoretic mobility shifts (Supplementary Fig. S3). To further investigate the functional interplay between DNA and histone binding we performed combined binding assays (Fig. 2D). Interestingly, binding to heterochromatin-specific H3K9me3 peptides induced a significant preference of Uhrf2 for hemi-over unmethylated DNA. Uhrf1 already on its own showed preference for hemimethylated DNA that was further enhanced by binding to H3K9me3 peptides. To test the specificity of this cooperativity we mutated the aromatic cage in Uhrf2 that is necessary for H3K9me3 histone-tail peptide binding. The mutated Uhrf2 (Y214A Y217A) showed comparable DNA binding activity as the wild-type Uhrf2 but addition of heterochromatin-specific H3K9me3 peptides did not induce preference for hemi-over unmethylated DNA (Fig. 2D).

In the reverse experiment, addition of DNA enhanced binding of Uhrf1 and Uhrf2 to the H3K9me3 peptide (Fig. 2E,F). This was not observed for the DNA binding mutant of Uhrf1 (Uhrf1ΔSRA) which showed constant peptide binding with increasing DNA concentrations. These findings suggest that single binding events of distinct Uhrf2 domains lead to multivalent engagement of different repressive epigenetic marks. In fact, multivalent engagement of DNA and histone tail peptides via the SRA domain and the TTD, respectively, results in affinity enhancement and additional specificity for hemimethylated DNA, the substrate of maintenance methylation.

CELLULAR LOCALIZATION AND DYNAMICS OF UHRF2 DEPEND ON HISTONE H3K9 METHYLATION

To monitor the subcellular localization of Uhrf2, we expressed Uhrf2-GFP constructs in cells with different genetic backgrounds. In wild type (wt) ESCs, Uhrf2 is localized in the nucleus and is enriched at pericentric heterochromatin (PH) (Fig. 3A,B and Supplementary Fig. S4A–C). To investigate which epigenetic marks at PH are recognized by Uhrf2 we determined the localization of Uhrf2 in genetically modified ESCs either lacking all three major DNA methyltransferases Dnmt1, Dnmt3a, and Dnmt3b (TKO) [Tsumura et al., 2006] or ESCs lacking the two major H3K9 methyltransferases Suv39H1/H2 (*Suv39h dn*) [Lehnertz et al., 2003]. TKO cells are practically devoid of genomic DNA methylation and *Suv39h dn* ESCs show substantially reduced H3K9me3 levels. We found Uhrf2 localized at PH in TKO but not in *Suv39h dn* ESCs, indicating that localization of Uhrf2 is dependent on H3K9 but not on DNA methylation (Fig. 3A). Consistently, immunostaining of wt mouse embryonic fibroblasts (MEFs) showed co-localization of Uhrf2 and H3K9me3 marks at PH, which was not observed in *Suv39h dn* MEFs [Peters et al., 2001] (Fig. 3B). Also, mutations in the TTD (Uhrf2 Y214A Y217A) that abolished binding to H3K9me3 peptides in vitro
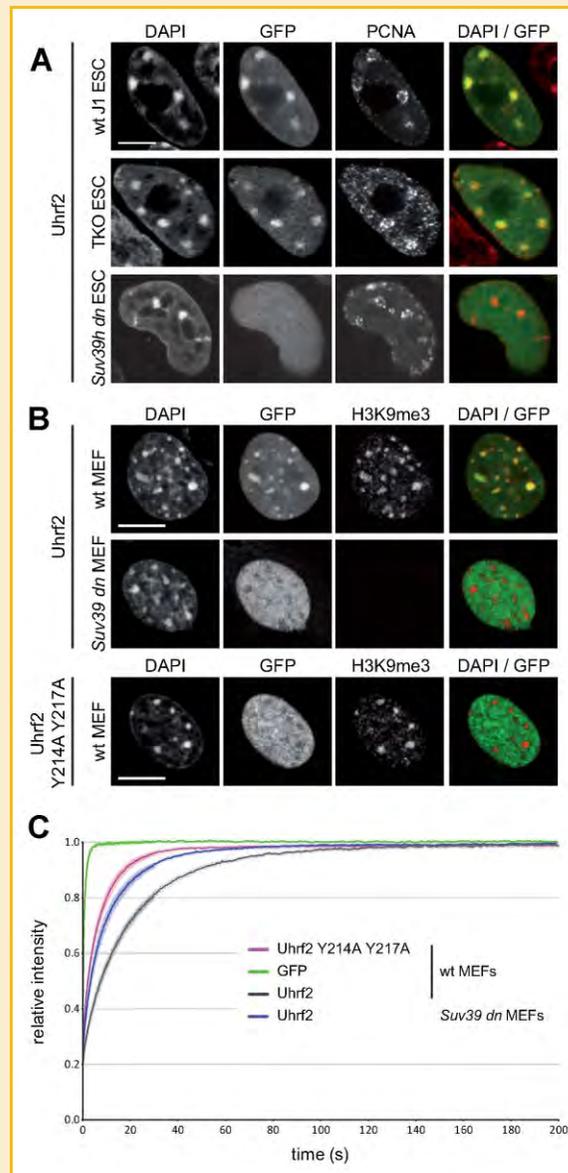

Fig. 3. Cellular localization and dynamics of Uhrf2 depend on histone H3K9 methylation. A: Confocal mid sections of fixed wt J1, TKO and *Suv39h dn* ESCs transiently expressing Uhrf2–GFP and RFP–PCNA and counterstained with DAPI, which preferentially highlights PH. Merged images are displayed on the right side (GFP: green; DAPI: red). Scale bar 5 μm. B: Confocal mid sections of fixed wt MEFs and *Suv39h dn* MEFs transiently expressing Uhrf2–GFP or Uhrf2 Y214A Y217A–GFP were immunostained for H3K9me3 and counterstained with DAPI. Merged images are displayed on the right side (GFP: green; DAPI: red). Scale bar 5 μm. C: Dynamics of Uhrf2–GFP and Uhrf2 Y214A Y217A–GFP in living MEFs determined by half nucleus FRAP analysis. GFP is shown as reference. Curves represent means ± SEM from at least 8 nuclei.

disrupted enrichment at PH in wt MEFs (Fig. 3B). The dependence of Uhrf2 localization on H3K9me3 was also confirmed by quantitative correlation analysis (Supplementary Fig. S4D,E).

To investigate the effect of H3K9me3 on the dynamics of Uhrf2 in living cells we performed quantitative fluorescence recovery after photobleaching (FRAP) analyses in wt and *Suv39h dn* MEFs. We chose to bleach half nuclei to include a representative number of

interactions from different nuclear domains and structures in the bleached area [Rottach et al., 2010]. Recovery of Uhrf2-GFP fluorescence in *Suv39h dn* MEFs (half-time $t_{1/2} = 5.9 \pm 0.6\,s$) and of the TTD mutant in wt MEFs ($t_{1/2} = 3.2 \pm 0.4\,s$) was considerably faster than the recovery of Uhrf2-GFP in wt MEFs ($t_{1/2} = 11.8 \pm 0.6\,s$) pointing to a crucial role of H3K9me3 in Uhrf2 dynamics in living cells (Fig. 3C). Taken together, these results clearly demonstrate that the interaction of Uhrf2 with the heterochromatin mark H3K9me3 is required for the localization at PH and affects binding dynamics in living cells.

COOPERATIVE BINDING OF THE COMBINED UHRF2 TTD–PHD DOMAIN

Recently, several studies showed multivalent binding to histone-tail peptides [Ruthenburg et al., 2007]. In case of Uhrf1 and Uhrf2, the TTD is followed by a second histone-tail binding domain, a PHD domain (Fig. 1A). As the isolated PHD domains of Uhrf1 and Uhrf2 did not show binding to H3 histone-tail peptides (Fig. 2B) [Rottach et al., 2010], we tested whether the combination of the PHD and the TTD results in cooperative histone-tail binding. Surprisingly, the combined TTD–PHD domain of Uhrf2 displayed a fourfold increased binding to H3K9me2/me3 in comparison to the single TTD, which was not observed for the corresponding construct of Uhrf1 (Figs. 2B and 4A).

Sequence alignments of the combined domains revealed two striking differences between Uhrf1 and Uhrf2. Firstly, Uhrf2 harbors an additional stretch of 33 highly conserved amino acids present in the TTD (Supplementary Fig. S5A). Secondly, the linker region between the TTD and PHD domain of Uhrf2 is highly conserved, whereas this region is highly diverse in Uhrf1 (Supplementary Fig. S5A). To test which sequence is responsible for the observed cooperative interplay between PHD and TTD, we generated and tested different hybrid and deletion constructs (Supplementary Fig. S5B). Notably, replacement of the native linker in the Uhrf2 TTD–PHD construct by the Uhrf1 linker caused decreased relative binding ratios to H3K9me2/3 comparable to the single Uhrf2 TTD (Fig. 4B). Transferring the Uhrf2 linker to the Uhrf1 TTD–PHD construct as well as deletion of the Uhrf2 stretch region did not affect the binding to H3K9me3 peptides (Fig. 4B).

These results suggest that the cooperative interplay of different Uhrf2 domains, which is responsible for the increased binding to heterochromatin marks, is dependent on the highly conserved linker region connecting the TTD and PHD domains. A similar functional importance of linker sequences has been described for BPTF and histone lysine demethylases [Li et al., 2006; Horton et al., 2010].

UHRF1 AND UHRF2 ARE NOT FUNCTIONALLY REDUNDANT IN ESCS

To investigate whether Uhrf1 and Uhrf2 are functionally redundant we performed interaction and rescue assays. Like Uhrf1, also Uhrf2 interacts with Dnmts (Supplementary Fig. S6) suggesting a similar function in DNA methylation. To test for such a functional role, we ectopically expressed Uhrf2-GFP or Uhrf1-GFP in *uhrf1*$^{-/-}$ ESCs and determined DNA methylation levels at major satellites by pyrosequencing. While ectopic expression of Uhrf1-GFP led to significant increase of DNA methylation levels at CpG sites of major satellite DNA in *uhrf1*$^{-/-}$ ESCs, Uhrf2-GFP did not restore DNA
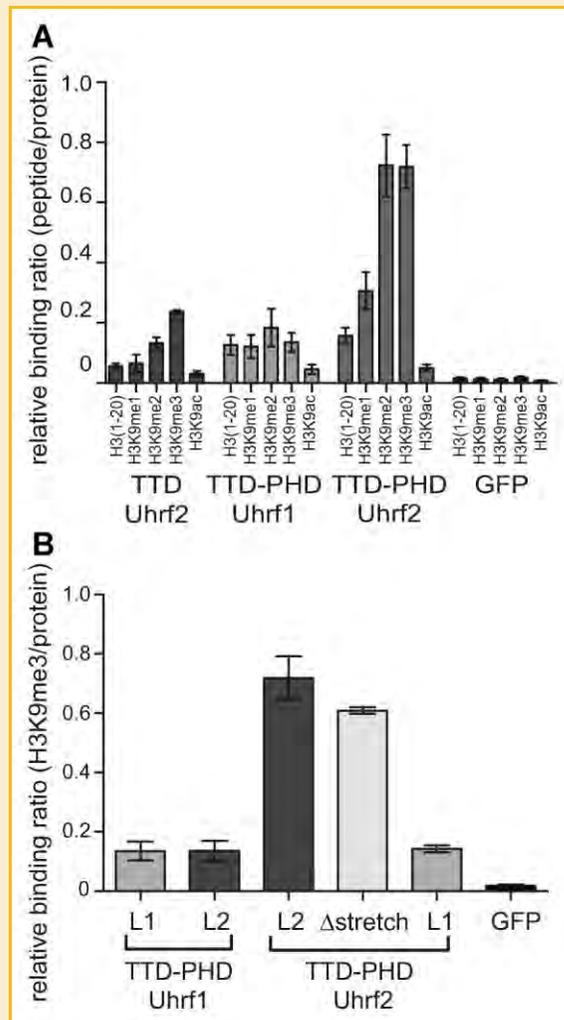

Fig. 4. Cooperative binding of the combined tandem Tudor–PHD domain of Uhrf2. A: Histone H3 N-terminal tail binding specificity of the TTD of Uhrf2 and of the combined TTD and PHD domain (TTD–PHD) of Uhrf1 and Uhrf2. Shown are means ± SEM from at least six independent experiments. B: Histone H3K9me3 binding of the combined TTD–PHD domains of Uhrf1 and Uhrf2, hybrid proteins (L1 and L2 specify inserted linker sequences derived from Uhrf1 and Uhrf2, respectively) and a stretch deletion Uhrf2 construct. Shown are means ± SEM from at least three independent experiments.

methylation at these sites (Fig. 5). These results point to functional differences between Uhrf1 and Uhrf2 in vivo.

DISCUSSION

Over the past decades many different histone modifications were discovered that are involved in epigenetic gene regulation. A key question is how these histone marks are linked to DNA methylation pattern and how this complex epigenetic information is integrated and translated into defined chromatin structures and gene expression levels. Epigenetic regulators that bind DNA and histone marks are ideally suited to link these pathways and intramolecular interactions between different binding domains may contribute to
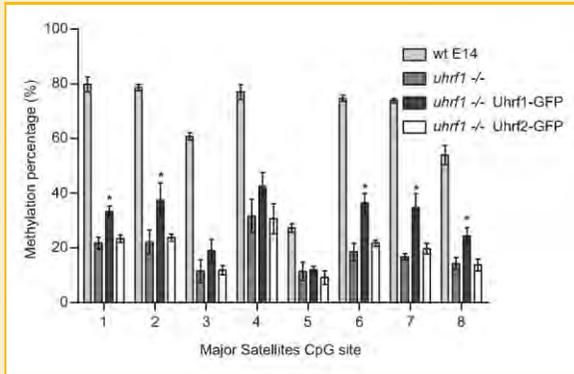
Fig. 5. Uhrf1 and Uhrf2 are not functionally redundant in ESCs. DNA methylation analysis of wt E14 ESCs, uhrf1−/− ESCs and of uhrf1−/− ESCs ectopically expressing Uhrf1–GFP or Uhrf2–GFP. ESCs transiently expressing Uhrf1–GFP and Uhrf2–GFP were isolated by FACS sorting 48 h after transfection and CpG methylation levels of major satellites repeats were analysed by bisulfite treatment, PCR amplification and direct pyrosequencing. Statistical significance of differences in DNA methylation levels between uhrf1−/− ESCs and uhrf1−/− ESCs with ectopically expressed Uhrf1–GFP or Uhrf2–GFP are indicated; $*P < 0.05$. Shown are means ± SD from three independent experiments.

substrate specificity and epigenetic regulation [Hashimoto et al., 2009].

Recently, Uhrf1, an essential factor for the maintenance of DNA methylation, has been shown to bind to repressive DNA and histone modifications via an SRA and a tandem Tudor domain, respectively. Here we provide the first systematic characterization of the second member of the Uhrf family, Uhrf2, and demonstrate that Uhrf2 binds to the H3K9me3 heterochromatin mark via an aromatic cage of a tandem Tudor domain (TTD). Mutations in the aromatic cage abolished binding to H3K9me3 histone-tail peptides in vitro and prevented enrichment of Uhrf2 at pericentric heterochromatin in vivo. Interestingly, similar mutations in the aromatic cage of Uhrf1 prevented repression of $p16^{INK4A}$ [Nady et al., 2011] suggesting a link between H3K9me3 binding and a function of Uhrf proteins in gene repression.

Our results point to a complex regulation of substrate recognition by Uhrf2 involving cooperative binding domains and critical linker sequences. In contrast to Uhrf1, preferential binding of Uhrf2 to hemimethylated DNA, the substrate of DNA maintenance methylation, was only induced upon simultaneous binding to H3K9me3 histone-tail peptides. Binding of Uhrf1 and Uhrf2 to DNA in turn enhanced binding to H3K9me3 histone-tail peptides. Consistently, SILAC-based proteomic analysis identified enrichment of UHRF1 at nucleosomes containing repressive DNA and H3K9 methylation marks [Bartke et al., 2010]. Together, these data demonstrate a cooperative interplay between DNA and histone tail binding domains of Uhrf1 and Uhrf2. A similar effect was reported for MSL3 that specifically binds to H4K20me1 via a chromodomain only in the presence of DNA [Kim et al., 2010].

An additional level of complexity was added by recent studies showing multivalent binding of histone-tail peptides by mixed two-effector modules [Ruthenburg et al., 2007]. Notably, the combined TTD–PHD domain of Uhrf2, but not of Uhrf1, showed enhanced

binding to H3K9me3 histone-tail peptides. This cooperativity was dependent on the highly conserved linker region connecting the TTD and PHD domains. Similarly, an important role was attributed to the linker sequence between the histone binding domain (PHD) and the histone modifying domain of jumanji histone lysine demethylases [Horton et al., 2010].

The dramatic loss of DNA methylation in uhrf1−/− ESCs [Bostick et al., 2007; Sharif et al., 2007] is remarkable, especially considering the presence of the uhrf2 gene, which encodes a highly similar protein as demonstrated in this study. As one possible explanation for this lack of functional redundancy we found, in contrast to uhrf1, relatively low uhrf2 mRNA levels in ESCs, which were not affected by genetic uhrf1 ablation. Moreover, both genes also show opposite expression patterns during differentiation. The failure of ectopically expressed Uhrf2 to restore DNA methylation in uhrf1 deficient cells clearly points to functional differences between both proteins in vivo. However, more definitive insights into the specific function(s) of Uhrf2 will require targeted mutations and subsequent analyses of pluripotent as well as differentiated cells. Based on the cooperative binding of Uhrf2 domains to repressive DNA and histone marks we propose that Uhrf2 might contribute to a tighter control of gene repression in differentiated cells as compared to a less stringent control by Uhrf1 in pluripotent ESCs.

REFERENCES

Abramoff MD, Magelhaes PJ, Ram SJ. 2004. Image processing with ImageJ. Biophotonics Int 11:36–42.

Achour M, Fuhrmann G, Alhosin M, Ronde P, Chataigneau T, Mousli M, Schini-Kerth VB, Bronner C. 2009. UHRF1 recruits the histone acetyltransferase Tip60 and controls its expression and activity. Biochem Biophys Res Commun 390:523–528.

Altschul SF. 1991. Amino acid substitution matrices from an information theoretic perspective. J Mol Biol 219:555–565.

Arita K, Ariyoshi M, Tochio H, Nakamura Y, Shirakawa M. 2008. Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. Nature 455:818–821.

Avvakumov GV, Walker JR, Xue S, Li Y, Duan S, Bronner C, Arrowsmith CH, Dhe-Paganon S. 2008. Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. Nature 455:822–825.

Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T. 2010. Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell 143:470–484.

Bestor TH, Ingram VM. 1983. Two DNA methyltransferases from murine erythroleukemia cells: Purification, sequence specificity, and mode of interaction with DNA. Proc Natl Acad Sci USA 80:5559–5563.

Bird A. 2002. DNA methylation patterns and epigenetic memory. Genes Dev 16:6–21.

Bostick M, Kim JK, Esteve PO, Clark A, Pradhan S, Jacobsen SE. 2007. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science 317:1760–1764.

Bronner C, Achour M, Arima Y, Chataigneau T, Saya H, Schini-Kerth VB. 2007. The UHRF family: Oncogenes that are drugable targets for cancer therapy in the near future? Pharmacol Ther 115:419–434.

Citterio E, Papait R, Nicassio F, Vecchi M, Gomiero P, Mantovani R, Di Fiore PP, Bonapace IM. 2004. Np95 is a histone-binding protein endowed with ubiquitin ligase activity. Mol Cell Biol 24:2526–2535.

Easwaran HP, Schermelleh L, Leonhardt H, Cardoso MC. 2004. Replication-independent chromatin loading of Dnmt1 during G2 and M phases. EMBO Rep 5:1181–1186.

Fellinger K, Rothbauer U, Felle M, Langst G, Leonhardt H. 2009. Dimerization of DNA methyltransferase 1 is mediated by its regulatory domain. J Cell Biochem 106:521–528.

Frauer C, Leonhardt H. 2009. A versatile non-radioactive assay for DNA methyltransferase activity and DNA binding. Nucleic Acids Res 37:e22.

Fujimori A, Matsuda Y, Takemoto Y, Hashimoto Y, Kubo E, Araki R, Fukumura R, Mita K, Tatsumi K, Muto M. 1998. Cloning and mapping of Np95 gene which encodes a novel nuclear protein associated with cell proliferation. Mamm Genome 9:1032–1035.

Hashimoto H, Horton JR, Zhang X, Bostick M, Jacobsen SE, Cheng X. 2008. The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. Nature 455:826–829.

Hashimoto H, Horton JR, Zhang X, Cheng X. 2009. UHRF1, a modular multi-domain protein, regulates replication-coupled crosstalk between DNA methylation and histone modifications. Epigenetics 4:8–14.

Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77:51–59.

Hopfner R, Mousli M, Jeltsch JM, Voulgaris A, Lutz Y, Marin C, Bellocq JP, Oudet P, Bronner C. 2000. ICBP90, a novel human CCAAT binding protein, involved in the regulation of topoisomerase IIalpha expression. Cancer Res 60:121–128.

Horton JR, Upadhyay AK, Qi HH, Zhang X, Shi Y, Cheng X. 2010. Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases. Nat Struct Mol Biol 17:38–43.

Iwata A, Nagashima Y, Matsumoto L, Suzuki T, Yamanaka T, Date H, Deoka K, Nukina N, Tsuji S. 2009. Intranuclear degradation of polyglutamine aggregates by the ubiquitin-proteasome system. J Biol Chem 284:9796–9803.

Jones PA, Baylin SB. 2007. The epigenomics of cancer. Cell 128:683–692.

Karagianni P, Amazit L, Qin J, Wong J. 2008. ICBP90, a novel methyl K9 H3 binding protein linking protein ubiquitination with heterochromatin formation. Mol Cell Biol 28:705–717.

Kim JK, Esteve PO, Jacobsen SE, Pradhan S. 2009. UHRF1 binds G9a and participates in p21 transcriptional regulation in mammalian cells. Nucleic Acids Res 37:493–505.

Kim D, Blus BJ, Chandra V, Huang P, Rastinejad F, Khorasanizadeh S. 2010. Corecognition of DNA and a methylated histone tail by the MSL3 chromodomain. Nat Struct Mol Biol 17:1027–1029.

Kouzarides T. 2007. Chromatin modifications and their function. Cell 128:693–705.

Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH. 2003. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol 13:1192–1200.

Leonhardt H, Page AW, Weier HU, Bestor TH. 1992. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell 71:865–873.

Li H, Ilin S, Wang W, Duncan EM, Wysocka J, Allis CD, Patel DJ. 2006. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature 442:91–95.

Meilinger D, Fellinger K, Bultmann S, Rothbauer U, Bonapace IM, Klinkert WE, Spada F, Leonhardt H. 2009. Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells. EMBO Rep 10:1259–1264.

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R, Lander ES. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454:766–770.

Mortusewicz O, Schermelleh L, Walter J, Cardoso MC, Leonhardt H. 2005. Recruitment of DNA methyltransferase I to DNA repair sites. Proc Natl Acad Sci USA 102:8905–8909.

Muto M, Utsuyama M, Horiguchi T, Kubo E, Sado T, Hirokawa K. 1995. The characterization of the monoclonal antibody Th-10a, specific for a nuclear protein appearing in the S phase of the cell cycle in normal thymocytes and its unregulated expression in lymphoma cell lines. Cell Prolif 28:645–657.

Nady N, Lemak A, Walker JR, Avvakumov GV, Kareta MS, Achour M, Xue S, Duan S, Allali-Hassani A, Zuo X, Wang YX, Bronner C, Chedin F, Arrowsmith CH, Dhe-Paganon S. 2011. Recognition of multivalent histone states associated with heterochromatin by UHRF1. J Biol Chem. [Epub ahead of print].

Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, Cheng X, Bestor TH. 2007. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 448:714–717.

Otani J, Nankumo T, Arita K, Inamoto S, Ariyoshi M, Shirakawa M. 2009. Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. EMBO Rep 10:1235–1241.

Papait R, Pistore C, Grazini U, Babbio F, Cogliati S, Pecoraro D, Brino L, Morand AL, Dechampesme AM, Spada F, Leonhardt H, McBlane F, Oudet P, Bonapace IM. 2008. The PHD domain of Np95 (mUHRF1) is involved in large-scale reorganization of pericentromeric heterochromatin. Mol Biol Cell 19:3554–3563.

Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schofer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A, Opravil S, Doyle M, Sibilia M, Jenuwein T. 2001. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell 107:323–337.

Pradhan S, Talbot D, Sha M, Benner J, Hornstra L, Li E, Jaenisch R, Roberts RJ. 1997. Baculovirus-mediated expression and characterization of the full-length murine DNA methyltransferase. Nucleic Acids Res 25:4666–4673.

Qian C, Li S, Jakoncic J, Zeng L, Walsh MJ, Zhou MM. 2008. Structure and hemimethylated CpG binding of the SRA domain from human UHRF1. J Biol Chem 283(50):34490–34494.

Reik W. 2007. Stability and flexibility of epigenetic gene regulation in mammalian development. Nature 447:425–432.

Rothbauer U, Zolghadr K, Muyldermans S, Schepers A, Cardoso MC, Leonhardt H. 2008. A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. Mol Cell Proteomics 7:282–289.

Rottach A, Frauer C, Pichler G, Bonapace IM, Spada F, Leonhardt H. 2010. The multi-domain protein Np95 connects DNA methylation and histone modification. Nucleic Acids Res 38:1796–1804.

Ruthenburg AJ, Li H, Patel DJ, Allis CD. 2007. Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol 8:983–994.

Schermelleh L, Haemmer A, Spada F, Rosing N, Meilinger D, Rothbauer U, Cardoso MC, Leonhardt H. 2007. Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. Nucleic Acids Res 35:4301–4312.

Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-Koseki Y, Toyoda T, Okamura K, Tajima S, Mitsuya K, Okano M, Koseki H. 2007. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature 450:908–912.

Sporbert A, Domaing P, Leonhardt H, Cardoso MC. 2005. PCNA acts as a stationary loading platform for transiently interacting Okazaki fragment maturation proteins. Nucleic Acids Res 33:3521–3528.

Szwagierczak A, Bultmann S, Schmidt CS, Spada F, Leonhardt H. 2010. Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. Nucleic Acids Res 38:e181.

Tsumura A, Hayakawa T, Kumaki Y, Takebayashi S, Sakaue M, Matsuoka C, Shimotohno K, Ishikawa F, Li E, Ueda HR, Nakayama J, Okano M. 2006. Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes Cells 11:805–814.

Unoki M, Nishidate T, Nakamura Y. 2004. ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain. Oncogene 23: 7601–7610.

Supplementary Information
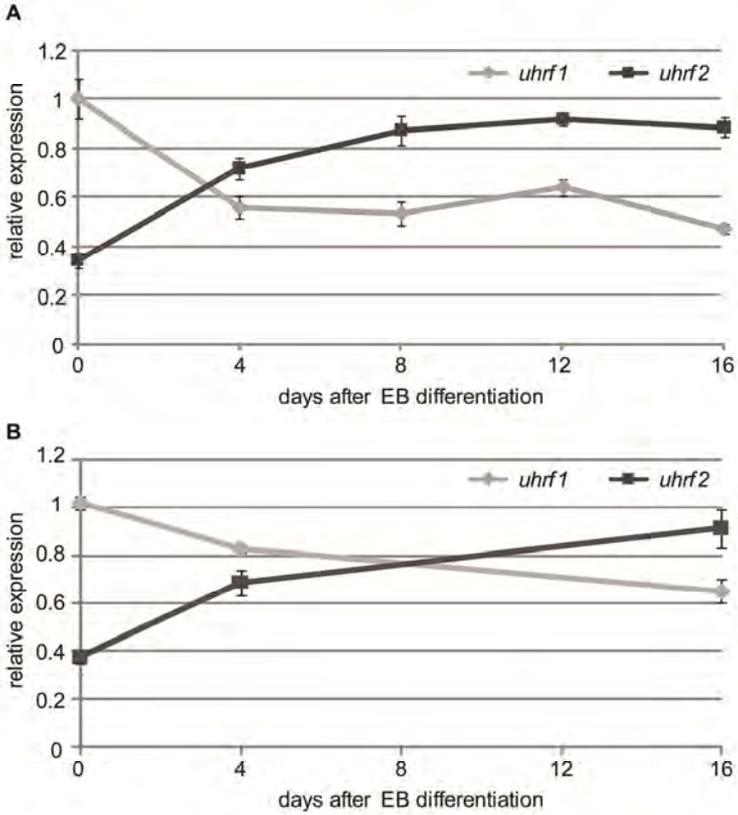
Supplementary Figure S1. Opposite expression pattern of *uhrf1* and *uhrf2*. Expression analyses of *uhrf1* and *uhrf2* by Real-time PCR during differentiation of ESCs with two different genetic backgrounds (wt E14 (**A**) and wt JM8A (**B**)). Transcript levels of *uhrf1* at day 0 of EB formation are used as reference point (set to 1). Shown are means ± SD from three technical replicates of one biological experiment.
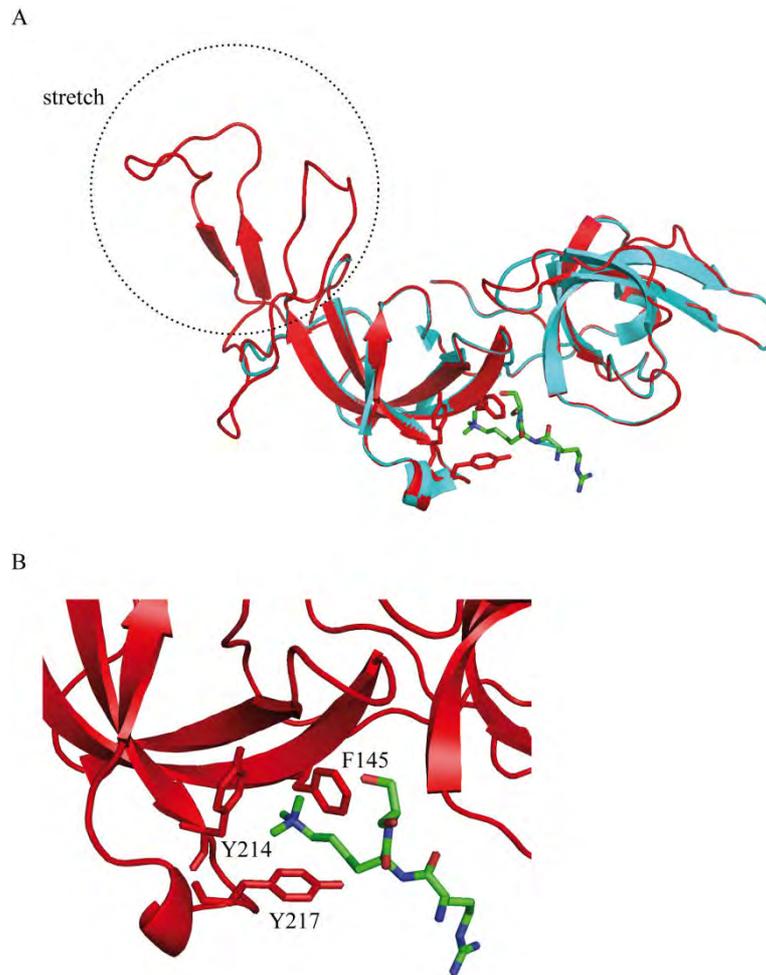
Supplementary Figure S2. Model of the tandem Tudor domain (TTD) of Uhrf2. (**A**) A model of the TTD of Uhrf2 was generated using SWISS Model [Arnold et al., 2006; Guex and Peitsch, 1997] with the solved structure of the TTD of Uhrf1 (PDB: 3DB3) as template. Both structures, the Uhrf2 model in red and the Uhrf1 template in cyan, are superimposed in PyMOL [Schrodinger, 2010]. (**B**) H3K9me3 is embedded in an aromatic cage formed by three aromatic residues of Uhrf2.

A

GFP ATTO550 unmethylated ATTO647N hemimethylated

B
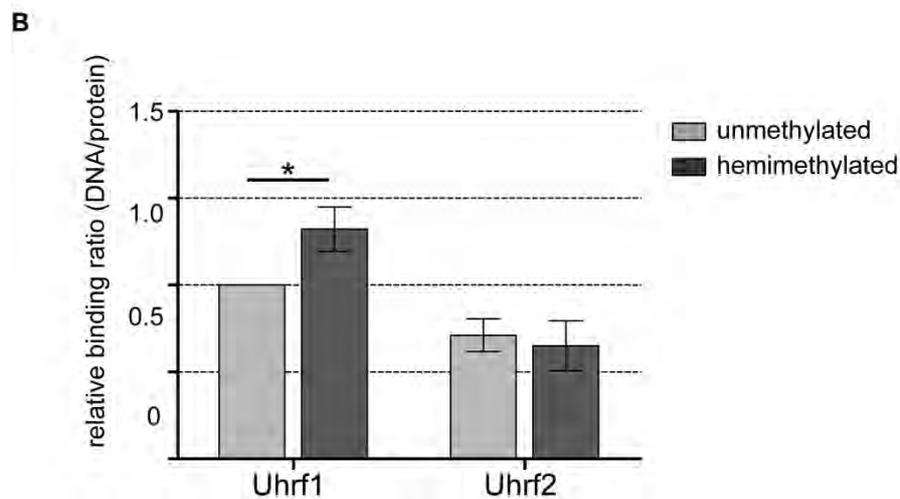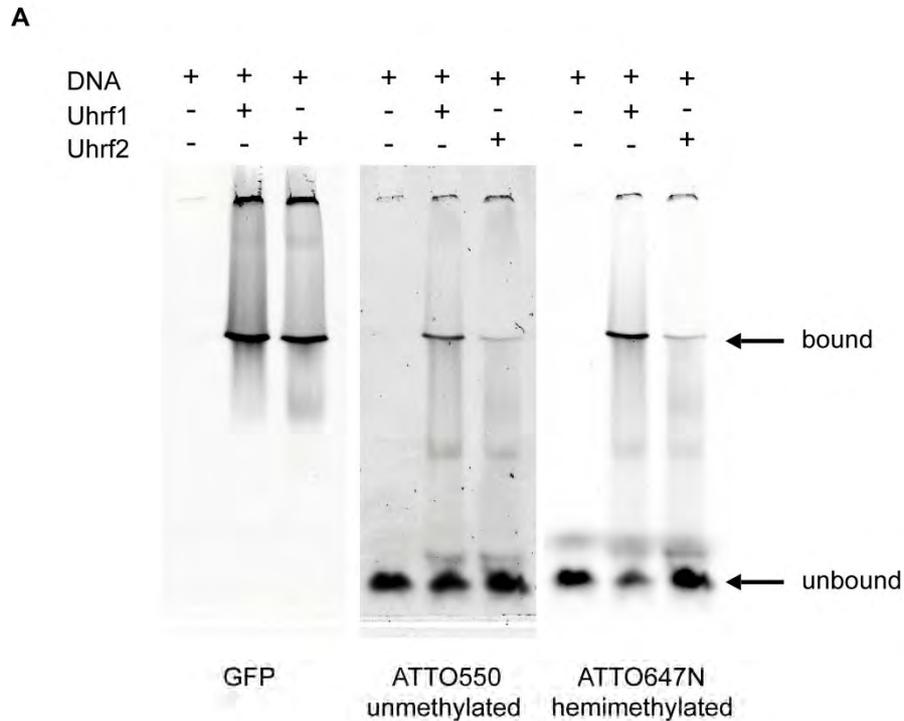
Supplementary Figure S3. Electrophoretic mobility shift of Uhrf1 and Uhrf2. (**A**) Un- and hemimethylated DNA substrates (1 pmol each in direct competition) were incubated with 0.63 pmol purified Uhrf1-GFP or Uhrf2-GFP. Samples were subjected to 3.5% non-denaturing PAGE and analyzed with a fluorescence scanner (Typhoon TRIO scanner, GE Healthcare) to detect ATTO550 (unmethylated substrate), ATTO647N (hemimethylated substrate) and GFP. (**B**) Band intensities were quantified with ImageJ [Abramoff, 2004]. To quantify bound DNA/protein ratios, grey values of unbound DNA bands were subtracted from the corresponding DNA input bands and subsequently normalized by the grey values of the GFP bands. All values were normalized to the relative binding ratio of Uhrf1 to unmethylated substrate. Shown are means ± SD from three independent experiments. Statistical significance between the binding ratios of un- and hemimethylated DNA is indicated; *P < 0.05.
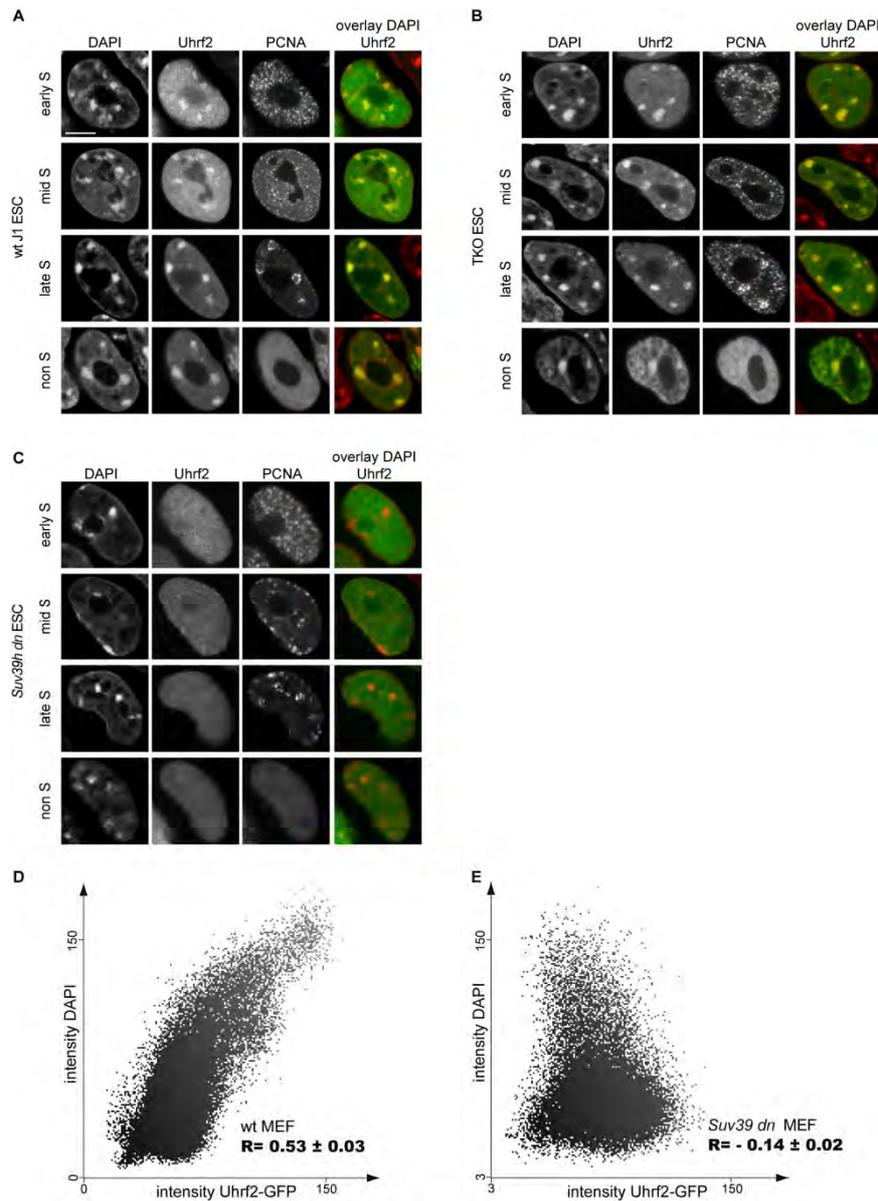
Supplementary Figure S4. Cell-cycle dependent localization of Uhrf2 in cells with different genetic backgrounds. Confocal mid sections of fixed wt J1 (**A**), TKO (**B**) and *Suv39h dn* ESCs (**C**), transiently expressing Uhrf2-GFP. Cells were co-transfected with a RFP-PCNA expression vector to distinguish S phase stages [Sporbert et al., 2005] and counterstained with DAPI. Merged images are displayed on the right. Scale bar 5 µm. In wt J1 and TKO ESCs the Uhrf2 fusion protein accumulates at pericentric heterochromatin independent of the cell-cycle stage and methylation levels (**A**) (**B**). In contrast, Uhrf2-GFP shows a fully dispersed nuclear distribution in *Suv39h dn* cells indicating the dependency on H3K9me3 methylation for localization at PH *in vivo* (**C**). (**D**) and (**E**) Scatter blot of GFP-Uhrf2 and DAPI signals in wt MEFs and *Suv39h dn* MEFs. The corresponding Pearson correlation coefficients R ± SEM are calculated from ten analysed cells. The software Volocity (Perkin Elmer) was used for analysis, selecting the cell nucleus as region of interest. Note that Pearson correlation coefficients range from +1 to -1 for perfect to no co-localization.
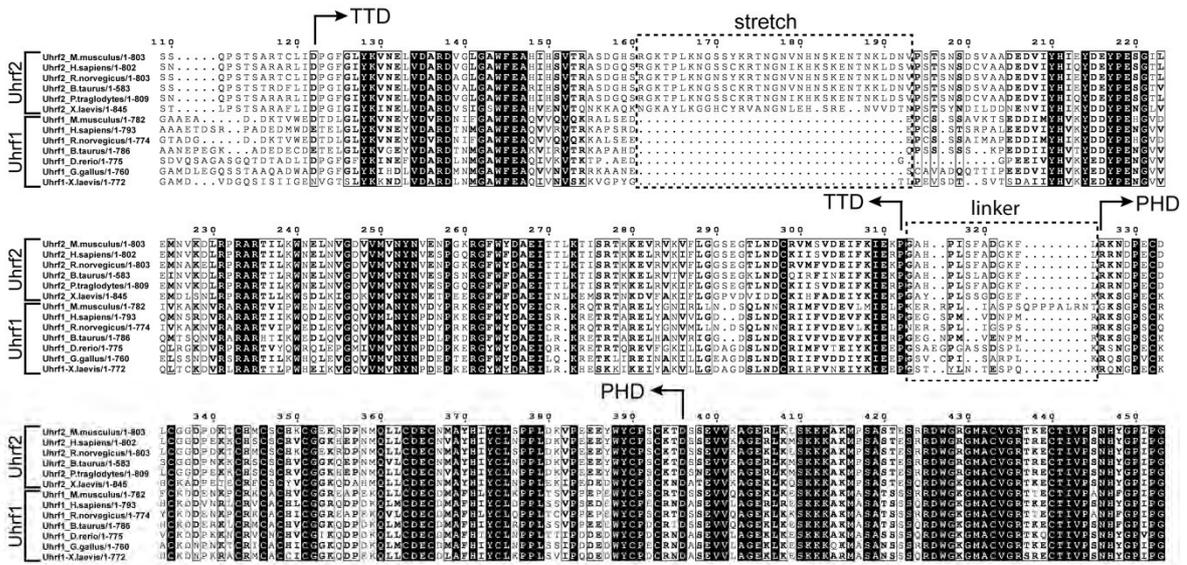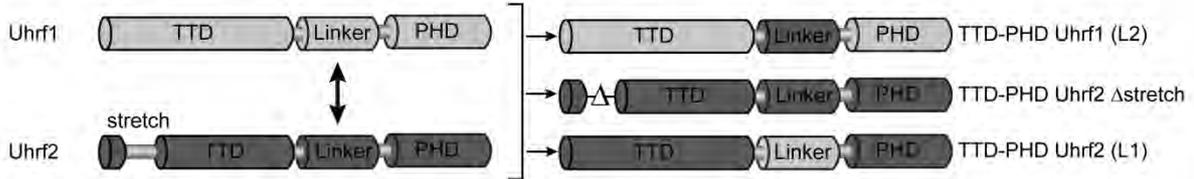
Supplementary Figure S5. Alignment and recombination of Uhrf1 and Uhrf2 domains. (**A**) Alignment of the tandem Tudor domain (TTD) and PHD domains from vertebrate Uhrf2 and Uhrf1 orthologs. Accession numbers for Uhrf2: *Homo sapiens* CAH74119.1; *Bos taurus* AAI48950.1; *Mus musculus* Q7TMI3; *Rattus norvegicus* NP_001101055.1; *Pan troglodytes* XP_528534.2; *Xenopus laevis* AAI28674.1. Accession numbers for Uhrf1: *Homo sapiens* Q96T88.1; *Bos taurus* AAI51672.1; *Mus musculus* Q8VDF2.2; *Rattus norvegicus* Q7TPK1.2; *Dario rerio* NP_998242.1; *Xenopus laevis* AAI28674.1, *Gallus gallus* XP_418269.2. Arrows show the start and end positions of the TTD and PHD domains. Absolutely conserved residues are black shaded, while positions showing conservative substitutions are boxed with residues in bold face. The additional stretch region found in the TTD of Uhrf2 and the linker region between TTD and PHD finger are boxed with dotted black lines. (**B**) Schematic outline of engineered constructs including the deletion of the stretch region and the swapping of linker sequences.
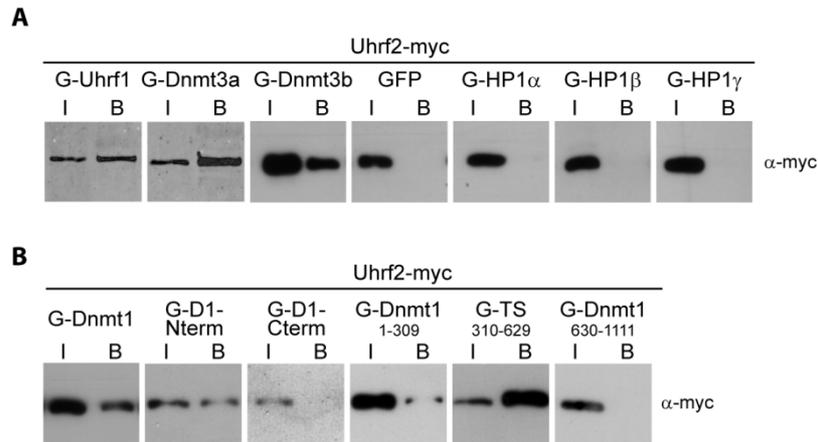
Supplementary Figure S6. Uhrf2 interacts with Uhrf1, Dnmt1 and Dnmt3a/b. (**A**) Co-immunoprecipation of Uhrf2-myc and GFP-Uhrf1, GFP-Dnmt3a, GFP-Dnmt3b, GFP-HP1α, GFP-HP1β, GFP-HP1γ or GFP transiently co-expressed in HEK293T cells. Note that Uhrf2 interacts with Uhrf1, Dnmt3a and Dnmt3b. (**B**) Co-immunoprecipation of Uhrf2-myc and GFP-Dnmt1 constructs transiently co-expressed in HEK293T cells: GFP-Dnmt1 (G-Dnmt1), GFP-fusions of the N-terminal and C-terminal part of Dnmt1 (G-D1-Nterm, G-D1-Cterm) and truncated Dnmt1 constructs (G-Dnmt1 1-309, G-TS 310-629, G-Dnmt1 630-1111). Note that Uhrf2 interacts with full-length Dnmt1, the N-terminal part and the targeting sequence (G-TS 310-629). One percent of input (I) relative to bound fractions (B) was loaded. Co-immunoprecipitation was performed using the GFP trap [Rothbauer et al., 2008]. Co-precipitated myc-tagged proteins were detected using a mouse monoclonal primary anti-myc antibody (Invitrogen, Germany) and an HRP- or Cy5-conjugated secondary anti-mouse antibody (Sigma, Germany, or Jackson ImmunoResearch Laboratories, USA, respectively).

A

Peptide name	Peptide sequence	Peptide labelling
H3K4me1	ART X1 QTARKSTGGKAPRKQLK	
H3K4me2	ART X2 QTARKSTGGKAPRKQLK	
H3K4me3	ART X3 QTARKSTGGKAPRKQLK	
H3K4ac	ART Z QTARKSTGGKAPRKQLK	
H3K4/9un	ARTKQTARKSTGGKAPRKQLK	TAMRA at C-terminus
H3K9me1	ARTKQTAR X1 STGGKAPRKQLK	
H3K9me2	ARTKQTAR X2 STGGKAPRKQLK	
H3K9me3	ARTKQTAR X3 STGGKAPRKQLK	
H3K9ac	ARTKQTAR Z STGGKAPRKQLK	
H3R2me2a	A X4 TKQTARSTGGKAPRKQLK	
H3K4me3K9me3	ART X3 QTAR X3 STGGKAPRKQLK	
H3K27un	RKQLATKAARKSAPATGGVK	
H3K27me1	RKQLATKAAR X1 SAPATGGVK	
H3K27me2	RKQLATKAAR X2 SAPATGGVK	
H3K27me3	RKQLATKAAR X3 SAPATGGVK	
H3K27ac	RKQLATKAAR Z SAPATGGVK	TAMRA at N-terminus
H4K20un	LGKGGAKRHRKVLRDNIQGI	
H4K20me1	LGKGGAKRHR X1 VLRDNIQGI	
H4K20me2	LGKGGAKRHR X2 VLRDNIQGI	
H4K20me3	LGKGGAKRHR X3 VLRDNIQGI	
H4K20ac	LGKGGAKRHR Z VLRDNIQGI	

X1: Lysine(me1); X2: Lysine(me2); X3: Lysine(me3); X4: Arginine(me2 asymmetric) Z: Lysine(ac)

B

DNA substrate	DNA sequence	DNA labelling
CGup	CTCAACAACTAACTACCATCCGGACCAGAAGAGTCATCATGG	no
MGup	CTCAACAACTAACTACCATCMGGACCAGAAGAGTCATCATGG	no
noCpG	CTCAACAACTAACTACCATCCTGACCAGAAGAGTCATCATGG	no
um647N	CCATGATGACTCTTCTGGTCCGGATGGTAGTTAGTTGTTGAG	ATTO647N at 5′end
um700	CCATGATGACTCTTCTGGTCCGGATGGTAGTTAGTTGTTGAG	ATTO700 at 5′end
Fill-In-550	CCATGATGACTCTTCTGGTC	ATTO550 at 5′end
Fill-In-590	CCATGATGACTCTTCTGGTC	ATTO590 at 5′end
Fill-In-647N	CCATGATGACTCTTCTGGTC	ATTO647N at 5′end
Fill-In-700	CCATGATGACTCTTCTGGTC	ATTO700 at 5′end

C
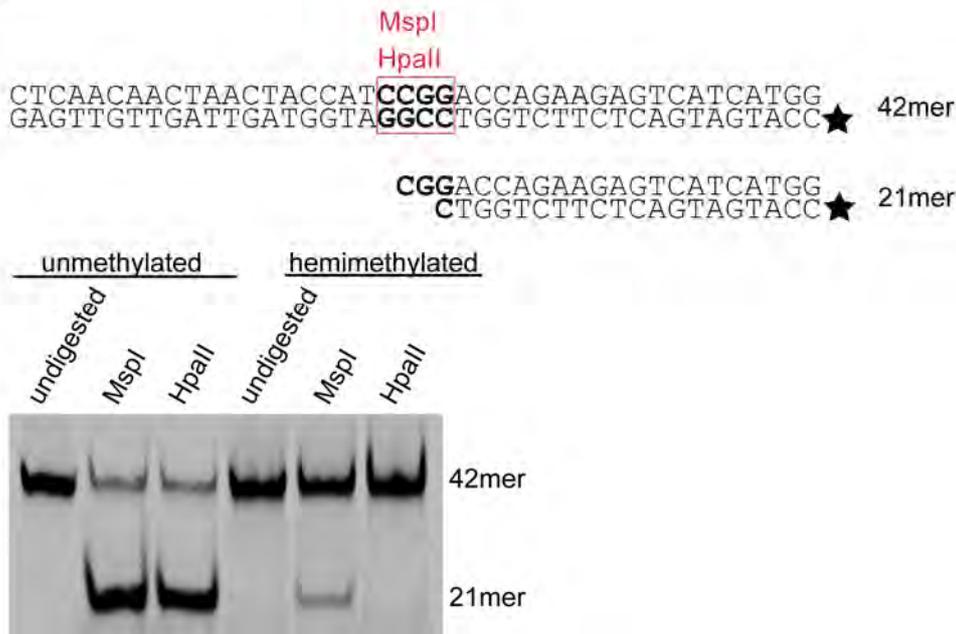
Supplementary Figure S7. Histone-tail peptide and DNA sequences and quality control of DNA substrates. (**A**) Amino acid sequence of TAMRA-labelled peptides for *in vitro* histone-tail peptide binding assays. Histone-tail peptides were purchased as TAMRA conjugates (PSL, Germany). (**B**) DNA oligos used for preparation of double-stranded probes for *in vitro*

DNA binding assays. M: 5-methyl-cytosine. For hybridization, DNA oligos were mixed in equimolar amounts, heated to 92°C and cooled down to room temperature. DNA substrates for Figure 2F were completed in a primer extension reaction. By using a control set of DNA probes with identical sequence but different fluorescent labels we observed effects due to probe preparation and/or unspecific binding of ATTO dyes (data not shown). The values obtained from the control set were used to normalize every probe/protein pair. (**C**) Quality control of DNA substrates. Un- and hemimethylated DNA substrates (2 pmol; Atto647N labelled) were digested with 1 unit MspI or HpaII and analyzed by 15% non-denaturing PAGE for CpG methylation. Note that unmethylated DNA substrate is digested by both enzymes, whereas hemimethylated substrate is only cut by MspI. Enzyme recognition motifs are boxed and asterisks represent ATTO labels.

Supplementary References

Abramoff MD, Magelhaes, P.J., Ram, S.J. 2004. Image Processing with ImageJ. Biophotonics International 11:36-42.

Arnold K, Bordoli L, Kopp J, Schwede T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22:195-201.

Guex N, Peitsch MC. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18:2714-23.

Rothbauer U, Zolghadr K, Muyldermans S, Schepers A, Cardoso MC, Leonhardt H. 2008. A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. Mol Cell Proteomics 7:282-9.

Schrodinger, LLC. 2010. The PyMOL Molecular Graphics System, Version 1.3editor^editors.

Sporbert A, Domaing P, Leonhardt H, Cardoso MC. 2005. PCNA acts as a stationary loading platform for transiently interacting Okazaki fragment maturation proteins. Nucleic Acids Res 33:3521-8.

Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells

2.6 Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells

scientific report

Np95 interacts with *de novo* DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells

Daniela Meilinger[1]*, Karin Fellinger[1]*, Sebastian Bultmann[1], Ulrich Rothbauer[1], Ian Marc Bonapace[2], Wolfgang E.F. Klinkert[3], Fabio Spada[1]+ & Heinrich Leonhardt[1]++

[1]Department of Biology II, Center for Integrated Protein Science Munich, Ludwig Maximilians University Munich, Planegg-Martinsried, Germany, [2]Department of Structural and Functional Biology, University of Insubria, Busto Arsizio, Italy, and

[3]Department of Neuroimmunology, Max Planck Institute of Neurobiology, Am Klopferspitz, Martinsried, Germany

Recent studies have indicated that nuclear protein of 95 kDa (Np95) is essential for maintaining genomic methylation by recruiting DNA methyltransferase (Dnmt) 1 to hemi-methylated sites. Here, we show that Np95 interacts more strongly with regulatory domains of the *de novo* methyltransferases Dnmt3a and Dnmt3b. To investigate possible functions, we developed an epigenetic silencing assay using fluorescent reporters in embryonic stem cells (ESCs). Interestingly, silencing of the cytomegalovirus promoter in ESCs preceded DNA methylation and was strictly dependent on the presence of either Np95, histone H3 methyltransferase G9a or Dnmt3a and Dnmt3b. Our results indicate a regulatory role for Np95, Dnmt3a and Dnmt3b in mediating epigenetic silencing through histone modification followed by DNA methylation.

[1]Department of Biology II, Center for Integrated Protein Science Munich (CIPSM), Ludwig Maximilians University Munich, Großhaderner Street 2, 82152 Planegg-Martinsried, Germany

[2]Department of Structural and Functional Biology, University of Insubria, Via da Giussano 12, 21052 Busto Arsizio (VA), Italy

[3]Department of Neuroimmunology, Max Planck Institute of Neurobiology, Am Klopferspitz 18, 82152 Martinsried, Germany

*These authors contributed equally to this work

+Corresponding author. Tel: + 49 89 218074230; Fax: + 49 89 218074236; E-mail: f.spada@lmu.de

++Corresponding author. Tel: + 49 89 218074232; Fax: + 49 89 218074236; E-mail: h.leonhardt@lmu.de

INTRODUCTION

In mammals, DNA methylation contributes to the establishment and maintenance of cell-type-specific gene expression programmes, imprinting, X-chromosome inactivation and genome stability (Bird, 2002). The majority of genomic methylation occurs at cytosine residues within CpG dinucleotides and is catalysed by the DNA methyltransferases (Dnmt) 1, 3a and 3b. Dnmt1 is responsible for maintaining genomic methylation, whereas Dnmt3a and Dnmt3b are mainly involved in *de novo* establishment of methylation patterns during cellular differentiation (Leonhardt *et al*, 1992; Li *et al*, 1992; Lei *et al*, 1996; Okano *et al*, 1999; Spada *et al*, 2007). Nuclear protein of 95 kDa (Np95; also known as Uhrf1) has recently been identified as an essential co-factor for maintaining genomic methylation (Bostick *et al*, 2007; Sharif *et al*, 2007; Achour *et al*, 2008). *dnmt1*−/− and *np95*−/− embryonic stem cells (ESCs) and embryos have similar reduced levels of DNA methylation. In addition, Np95 interacts with Dnmt1, binds hemi-methylated CpG sites through its Set and Ring associated (SRA) domain and both Np95 and Dnmt1 accumulate at replication sites (Uemura *et al*, 2000; Bostick *et al*, 2007; Papait *et al*, 2007; Arita *et al*, 2008; Avvakumov *et al*, 2008; Hashimoto *et al*, 2008). Thus, it has been proposed that Np95 mediates maintenance of genomic methylation by recruiting Dnmt1 to hemi-methylated CpG sites generated during replication.

Here, we investigated a possible involvement of Np95 in epigenetic regulation beyond its role in Dnmt1-mediated maintenance of DNA methylation. We found that Np95 interacts with the *de novo* methyltransferases, Dnmt3a and Dnmt3b, and mediates promoter silencing before DNA methylation is detected.

RESULTS AND DISCUSSION
Np95 interacts with Dnmt3a and Dnmt3b

Immunoprecipitation experiments showed that different isoforms of both *de novo* methyltransferases Dnmt3a and Dnmt3b interact
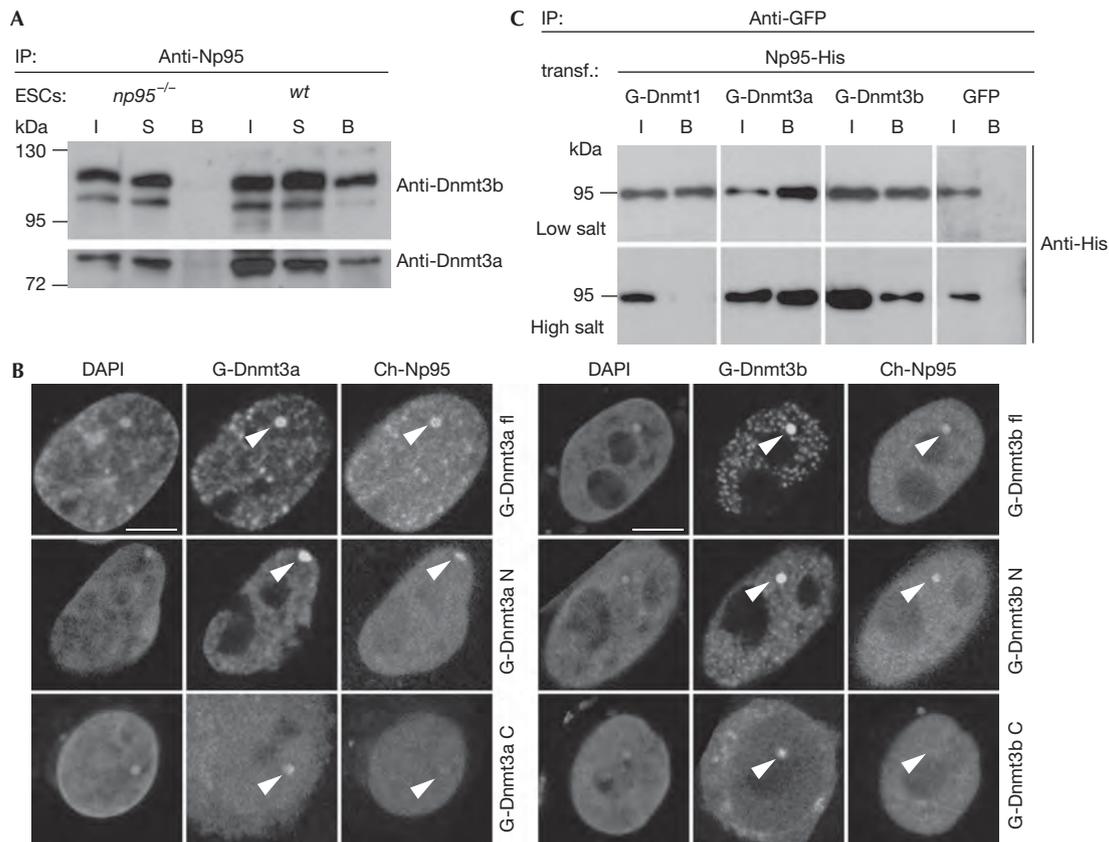
Fig 1 | Np95 interacts with *de novo* methyltransferases Dnmt3a and Dnmt3b. (A) Co-immunoprecipitation of Dnmt3a and Dnmt3b with Np95 in wt and *np95⁻/⁻* E14 ESCs. The Dnmt3a2 isoform is shown in the lower panel. **(B)** F2H shows recruitment of Cherry–Np95 (prey) at the *lac* operator array (indicated by arrowheads) when GFP fusions of full-length Dnmt3a and Dnmt3b1 (G-Dnmt3a/b fl) or their amino-terminal regions (G-Dnmt3a/b N) are used as bait and not with their isolated C-terminal catalytic domains (G-Dnmt3a/b C). Scale bars, 5 μm. **(C)** Co-immunoprecipitation of Np95-His with GFP-tagged Dnmt1, Dnmt3a and Dnmt3b1 (G-Dnmt) transiently co-expressed in HEK293T cells. Co-expression of GFP was used as the control. In the upper row, immunoprecipitations carried out in the presence of 150 mM NaCl throughout the procedure are shown, whereas in the lower row, immunoprecipitation and wash buffers were carried out using 300 and 500 mM NaCl, respectively. Two per cent of input and supernatant relative to bound fractions were loaded in (**A**) and (**C**). B, bound; Dnmt, DNA methyltransferase; ESCs, embryonic stem cells; F2H, fluorescent two-hybrid assay; GFP, green fluorescent protein; HEK293T, human embryonic kidney 293T; I, input; Np95, nuclear protein of 95 kDa; S, supernatant; wt, wild type.

with Np95 in wild-type (wt) ESCs, including the more abundant Dnmt3a2 and Dnmt3b1 (Fig 1A). Furthermore, using a green fluorescent protein (GFP) trap (Rothbauer *et al*, 2008), we co-immunoprecipitated endogenous Dnmt1 and isoforms of Dnmt3a and Dnmt3b with a GFP–Np95 fusion construct transiently expressed in *np95⁻/⁻* ESCs and, vice versa, endogenous Np95 co-immunoprecipitated with GFP–Dnmt3a or GFP–Dnmt3b1 fusions in *dnmt3a* and *3b* double knockout (DKO) ESCs (supplementary Fig S1A,B online). In addition, we observed co-immunoprecipitation of endogenous DNMT3b and inverted CCAAT box binding protein of 90 kDa—the human homologue of Np95—from human embryonic kidney 293T (HEK293T) cell extracts (supplementary Fig S1C online). We confirmed the interaction of Np95 with Dnmt3a/b by using a recently developed fluorescent two hybrid assay (F2H; Zolghadr *et al*, 2008). GFP–Dnmt3 fusion constructs were used as bait by tethering them to a *lac* operator array present in baby hamster kidney (BHK) cells, so that the array was visible as a distinct

nuclear spot of enriched GFP fluorescence (Fig 1B). A Cherry–Np95 fusion (prey) accumulated at this spot only when GFP fusions of full-length Dnmt3a and Dnmt3b1 or their amino-terminal regions were used as bait and not when their isolated Carboxy-terminal catalytic domains were used. We further mapped the interaction of Np95 with Dnmt3a/b through co-immunoprecipitation of deletion constructs and isolated domains transiently expressed in HEK293T cells (supplementary Fig S2 online). The results were consistent with those produced by F2H: the N-terminal regions of Dnmt3a and Dnmt3b1, but not their C-terminal catalytic domains, interacted with Np95. Deletion of the PHD or PWWP domains of Dnmt3a and Dnmt3b did not eliminate the interaction with Np95. We then determined the domains of Np95 involved in this interaction. We found that the SRA domain and the N-terminal 298 amino acids of Np95, which include the ubiquitin-like domain, interacted with Dnmt3a and Dnmt3b1, whereas the PHD domain and the C-terminal 132 amino acids, including the Ring domain, did not. Furthermore,
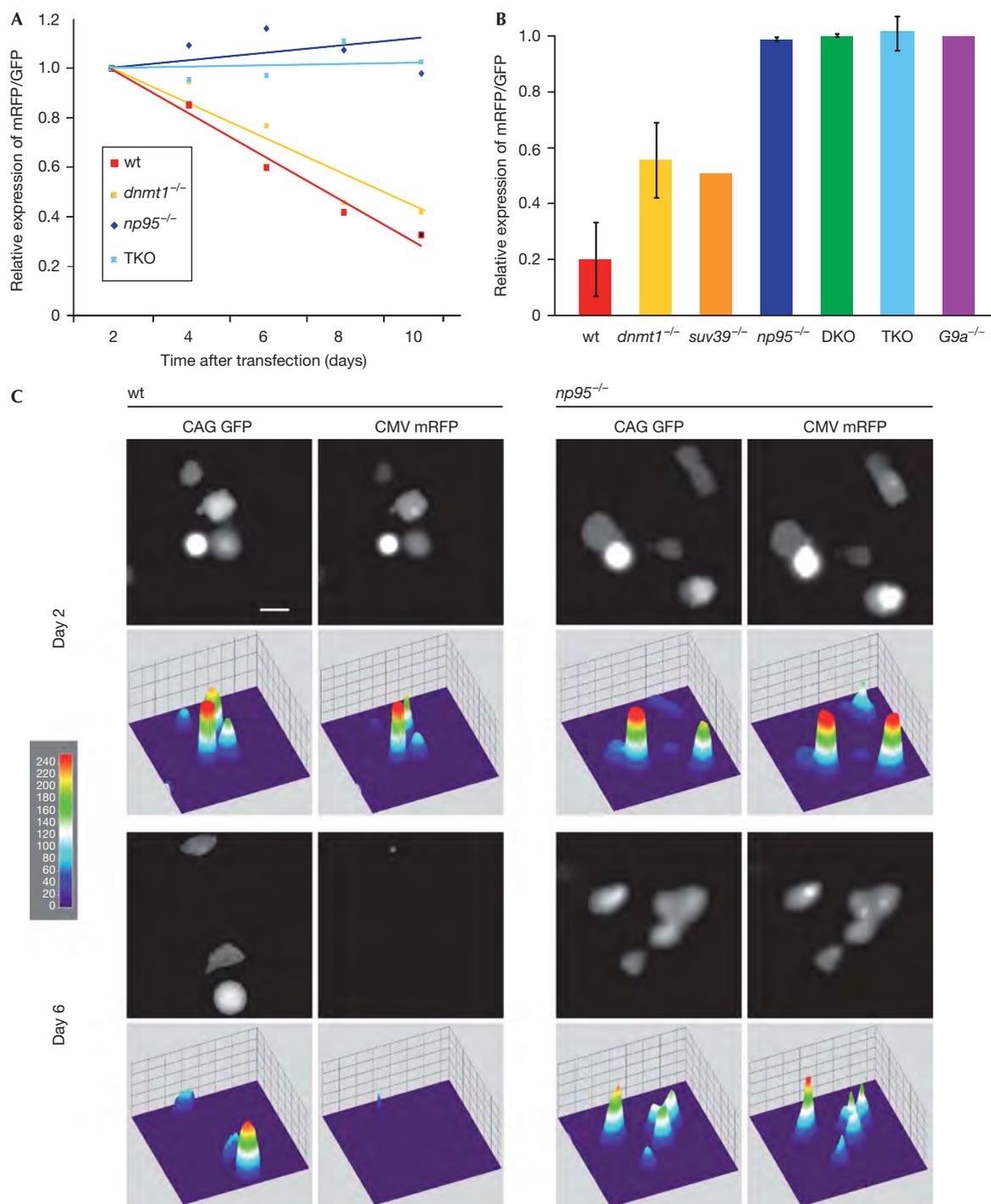
Fig 2 | Promoter silencing activity in wild-type and mutant ESCs. ESCs with the indicated phenotypes were transiently co-transfected with CMV promoter-driven mRFP and CAG promoter-driven GFP reporter constructs. Between 90 and 150 images per sample were acquired either (**A**) every second day after transfection from a single experiment or (**B**) only at days 2 and 7–10 after transfection from 3–5 independent experiments. Relative levels of red over green fluorescence are shown with values for day 2 (first day of imaging) set to 1 (**A**). Wild-type J1 and E14 cells gave similar results (data not shown). $suv39^{-/-}$ stands for Suv39h1/2 double null ESCs. (**C**) Representative images of wt and $np95^{-/-}$ E14 ESCs co-transfected as in (**A**) and (**B**) (upper panels) and respective heat map intensity plots (lower panels). Scale bar, 15 μm. CAG, CMV early enhancer/chicken β actin promoter; CMV, cytomegalovirus promoter; DKO, double knockout; Dnmt, DNA methyltransferase; ESCs, embryonic stem cells; GFP, green fluorescent protein; mRFP, monomeric red fluorescent protein; Np95, nuclear protein of 95 kDa; TKO, triple knockout; wt, wild type.

we observed co-immunoprecipitation of endogenous Np95 with GFP–Dnmt3a and GFP–Dnmt3b transiently expressed in *dnmt1*⁻/⁻ ESCs, indicating that Dnmt3a and Dnmt3b interact with Np95 independently of Dnmt1 (supplementary Fig S1D online). To compare the relative association between endogenous Np95 and Dnmts, we re-probed the blot in Fig 1A with a Dnmt1 antibody and observed a substantially weaker signal for the co-immunoprecipitated Dnmt1 relative to the input than in the case of Dnmt3a2 and Dnmt3b1 (supplementary Fig S3A online). To compare further the stability of Np95 interactions with the Dnmts, we transiently co-expressed Np95-His with GFP–Dnmt1, GFP–Dnmt3a or GFP–Dnmt3b1 in HEK293T cells and immunoprecipitated with the GFP trap in the presence of different salt concentrations (Fig 1C; supplementary Fig S3B online). Interestingly, under high salt conditions, the interaction between Np95-His and GFP–Dnmt1 was lost, whereas co-immunoprecipitation of GFP–Dnmt3a and GFP–Dnmt3b1 remained relatively unaffected. These data clearly indicate that Np95 interacts more strongly with the *de novo* methyltransferases, Dnmt3a and Dnmt3b, than with Dnmt1.

Np95, Dnmt3a/3b and G9a mediate epigenetic silencing

As DNA methylation has a central role in epigenetic silencing, we investigated the requirement of DNA methyltransferases and Np95 for promoter silencing in ESCs. We found that, on transient transfection of wt ESCs, constructs driven by the cytomegalovirus (CMV) promoter were rapidly silenced, as opposed to longer-lasting expression of constructs driven by the chimeric CMV early enhancer/chicken β actin (CAG) promoter (Fig 2), which is consistent with the popularity of the CAG promoter for stable transgene expression in ESCs and mice. We then established an epigenetic silencing assay based on this observation. ESCs were co-transfected with two distinct plasmids, one expressing mono-meric red fluorescent protein (mRFP) under the CMV promoter, the other expressing GFP driven by the CAG promoter. mRFP and GFP expression was monitored after transfection for up to ten days by using automated image acquisition and quantification of fluorescent signals (supplementary Fig S4A online). The ratio between mRFP and GFP expression declined steadily in wt ESCs, reflecting preferential silencing of the CMV promoter (Fig 2). By contrast, DKO ESCs and ESCs lacking all three major DNA methyltransferases (*dnmt1*, *3a* and *3b* triple knockout) showed no preferential silencing of the CMV promoter. Surprisingly, *np95*⁻/⁻ ESCs were also unable to silence the CMV promoter, whereas *dnmt1*⁻/⁻ ESCs showed only partly reduced silencing under these conditions. Similar results were obtained on swapping GFP and mRFP reporter sequences, ruling out potential artefacts due to differences in their coding sequences or stability of the reporter proteins (supplementary Fig S4B online). Thus, despite expressing a full complement of DNA methyltransferases, ESCs lacking Np95 are as deficient in promoter silencing activity as ESCs lacking all three major Dnmts. We next investigated whether silencing of the CMV promoter correlates with CpG methylation. Interestingly, promoter methylation was detected only ten days after transfec-tion and was lower in *np95*⁻/⁻ than in wt ESCs, whereas none of the *dnmt* mutant ESCs showed appreciable DNA methylation (Fig 3; supplementary Fig S5A online). At the same time no obvious methylation was detected in any of the ESC lines within the CpG island of the CAG promoter construct (supplementary Fig S5B online). Thus, CMV promoter silencing depends on the
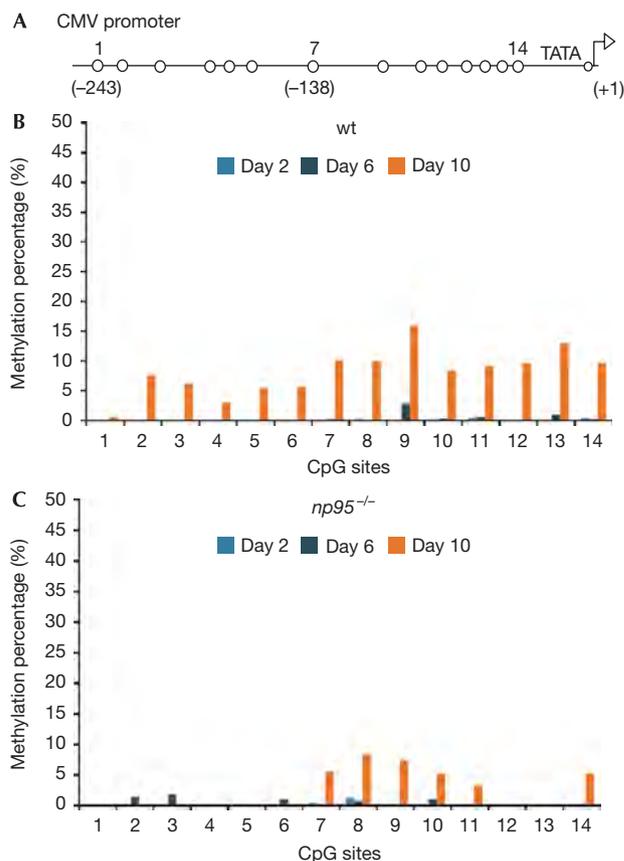
Fig 3 | Methylation of the CMV promoter 2, 6 and 10 days after transfection. Wild-type and *np95*⁻/⁻ ESCs were transfected as in Fig 2 and GFP-positive cells were sequentially sorted at the indicated days after transfection. Total DNA was isolated from sorted cells and bisulphite-treated. A proximal part of the CMV promoter was amplified and subjected to pyrosequencing. (**A**) Schematic drawing of the 14 proximal CpG sites analysed (shown as open circles). The numbers above correspond to the CpG sites shown in (**B**) and (**C**) and numbers in brackets refer to the position of CpG sites with respect to the transcription start site. (**B,C**) Methylation percentage at individual CpG sites for (**B**) wt and (**C**) *np95*⁻/⁻ ESCs as measured by pyrosequencing. CMV, cytomegalovirus; ESCs, embryonic stem cells; GFP, green fluorescent protein; wt, wild type.

presence of both Np95 and *de novo* Dnmts, but ensues well before *de novo* methylation of the promoter is detected. This prompted us to investigate the involvement of repressive histone methylation as a possible mechanism for the observed silencing. We found that in the absence of histone H3 lysine 9 methyl-transferases (H3K9MTs), G9a or Suv39h1/2, silencing of the CMV promoter was completely abolished or reduced, respectively, indicating that G9a and, in part, Suv39h1/2 are also required for silencing (Fig 2B).

The results shown here indicate that Np95 interacts with Dnmt3a and Dnmt3b and mediates silencing of the CMV promoter by mechanisms that are, at least initially, independent of *de novo* DNA methylation. Importantly, our data also show the involvement of H3K9MTs, G9a and Suv39h1/2 in CMV promoter silencing. H3K9MTs were reported to associate with *de novo*

Dnmts, and major satellite repeats were found to be hypomethylated in ESCs lacking either Suv39h1/2 or Dnmt3 enzymes. However, major satellite transcript levels were altered only in Suv39h1/2-deficient cells and not in Dnmt3-deficient cells (Fuks et al, 2003; Lehnertz et al, 2003; Martens et al, 2005). A recent study showed that G9a, Dnmt1, Dnmt3a and Dnmt3b are required for normal methylation at long terminal repeats of endogenous retrotransposable elements, although transcription of these elements was increased in Dnmt-deficient ESCs, but not G9a-deficient ESCs (Dong et al, 2008). Furthermore, recent studies have shown that G9a interacts with Dnmt3a and Dnmt3b and mediates *de novo* methylation of the *oct4*, *nanog* and *dnmt3l* promoters on retinoic-acid-induced differentiation of ESCs (Feldman et al, 2006; Li et al, 2007; Epsztejn-Litman et al, 2008). However, two of these studies showed that neither G9a nor *de novo* Dnmts are required to silence the *oct4* promoter, and G9a was also found to be dispensable for silencing the *nanog* and *dnmt3l* promoters (Feldman et al, 2006; Epsztejn-Litman et al, 2008). In the third study, *nanog*, but not *oct4*, was shown to be silenced in differentiating ESCs lacking both Dnmt3a and Dnmt3b (Li et al, 2007). We found that silencing of *oct4* during embryoid body differentiation is largely independent from the presence of Np95 as well as all three major Dnmts, and occurs in the absence of DNA methylation (D.M., F.S., S.B., and H.L., unpublished data). These data, together with our findings on silencing of the CMV promoter in ESCs, indicate that Dnmts, Np95 and H3K9MTs mediate silencing through many mechanisms that do not necessarily involve DNA methylation and might depend on the presence of different *cis* elements and an intricate interplay with other epigenetic and transcription factors. Interestingly, Np95 was recently shown to interact with G9a (Kim et al, 2009) and here we show that silencing of the CMV promoter in ESCs strictly depends on Np95 and on *de novo* Dnmts as well as G9a. Taken together these observations suggest that Np95, *de novo* Dnmts and G9a might be involved in a common silencing pathway.

In summary, our data clearly support a crucial role of Np95 in epigenetic silencing mediated by *de novo* DNA and histone methyltransferases, and make Np95 an attractive target for epigenetic reprogramming strategies.

METHODS

Cell culture and transfection. HEK293T cells, BHK cells and ESCs were cultured and transfected as described by Schermelleh et al (2007), except FuGENE HD (Roche, Mannheim, Germany) was used for transfection of ESCs. The *dnmt1*[−/−] J1 ESCs used in this study were homozygous for the c allele (Lei et al, 1996). BHK cells were co-transfected on glass coverslips with GFP–Dnmt3 and Cherry–Np95 constructs using Transfectin (Bio-Rad, Munich, Germany) according to the manufacturer's instructions. Cell fixation and microscopy were carried out as described by Zolghadr et al (2008).

Co-immunoprecipitation. ESCs and HEK293T cell extracts were prepared in lysis buffer (20 mM Tris–HCl (pH 7.5), 0.5 mM EDTA, 2 mM phenylmethyl sulphonyl fluoride and 0.5% NP40) containing 150 or 300 mM NaCl (high-salt condition) and diluted with lysis buffer without NP40. GFP trap (Rothbauer et al, 2008) and a specific rabbit antiserum (Citterio et al, 2004) were used for immunoprecipitation of GFP fusions and endogenous Np95, respectively. GFP trap and protein G beads (Sigma, Taufkirchen,

Germany) were washed with dilution buffer containing increasing salt concentrations (150 and 300 mM, or 300 and 500 mM NaCl for the high-salt condition) and re-suspended in SDS–PAGE sample buffer. The following mouse monoclonal antibodies were used for immunoblotting: anti-His (C-terminal, Invitrogen, Karlsruhe, Germany), anti-Dnmt3a (clone 64B1446, Imgenex, San Diego, CA, USA) and anti-Dnmt3b (clone 52A1018, Abcam, Cambridge, UK). Np95 was detected with the same antiserum used for immunoprecipitation and a rabbit antiserum was used for detection of Dnmt1 (Grohmann et al, 2005). Horseradish peroxidase-conjugated rabbit anti-mouse or goat anti-rabbit secondary antibodies (Sigma) and ECL Plus reagent (GE Healthcare, Munich, Germany) were used for detection.

Silencing assay. ESCs were co-transfected with pCAG-eGFP-IB and pCMV-mRFP as described above and images from live cells were acquired at the indicated time points with an InCell Analyser 1000 (GE Healthcare) using a ×20 air objective (NA = 0.45) and standard filter settings for GFP and RFP. A total of 90–150 images were acquired for each channel, using the same exposure time throughout the time course. Cells were passaged every second day and images were taken 4–5 h after seeding. Images were analysed using ImageJ v1.42a software. To calculate fluorescent reporter expression, pictures were processed using a Gaussian blur algorithm (radius (sigma) = 2), and a threshold for maximal signal and minimal background coverage was adjusted and applied to each channel (supplementary Fig S4A online). The threshold was converted into area selection and the total size of the selected area was measured.

DNA methylation analysis. ESCs were transfected as in the silencing assay with pCAG-eGFP-IB and pCMV-mRFP, and GFP-positive cells were sequentially sorted with a FACSVantage or FACSAria II (Becton Dickinson, Heidelberg, Germany) at days 2, 6 and 10 after transfection. After each sorting, total DNA was isolated using the QIAmp DNA Mini kit (Qiagen, Hilden, Germany) and bisulphite treated with the EZ DNA Methylation-Gold kit (Zymo research, Orange, CA, USA). The following primers were used for PCR amplification: CMV-forward TGGGAT TTTTTTATTTGGTAGT; CMV-reverse ATGGGAGTTTGTTTTGG TATTA; CAG-forward GGAGAGGTGAGGAGGTAGTTAATTAGA and CAG-reverse CCCCAAACCCCTCAAAACTT. Pyrosequencing was carried out by Varionostic GmbH (Ulm, Germany).

Supplementary information is available at *EMBO reports* online (http://www.emboreports.org).

CONFLICT OF INTEREST
The authors declare that they have no conflict of interest.

REFERENCES
Achour M et al (2008) The interaction of the SRA domain of ICBP90 with a novel domain of DNMT1 is involved in the regulation of VEGF gene expression. *Oncogene* **27:** 2187–2197

Arita K, Ariyoshi M, Tochio H, Nakamura Y, Shirakawa M (2008) Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature* **455:** 818–821

Avvakumov GV, Walker JR, Xue S, Li Y, Duan S, Bronner C, Arrowsmith CH, Dhe-Paganon S (2008) Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. *Nature* **455:** 822–825

Bird A (2002) DNA methylation patterns and epigenetic memory. *Genes Dev* **16:** 6–21

Bostick M, Kim JK, Esteve P-O, Clark A, Pradhan S, Jacobsen SE (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317:** 1760–1764

Citterio E, Papait R, Nicassio F, Vecchi M, Gomiero P, Mantovani R, Di Fiore PP, Bonapace IM (2004) Np95 is a histone-binding protein endowed with ubiquitin ligase activity. *Mol Cell Biol* **24:** 2526–2535

Dong KB *et al* (2008) DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. *EMBO J* **27:** 2691–2701

Epsztejn-Litman S *et al* (2008) *De novo* DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat Struct Mol Biol* **15:** 1176–1183

Feldman N, Gerson A, Fang J, Li E, Zhang Y, Shinkai Y, Cedar H, Bergman Y (2006) G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol* **8:** 188–194

Fuks F, Hurd PJ, Deplus R, Kouzarides T (2003) The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res* **31:** 2305–2312

Grohmann M, Spada F, Schermelleh L, Alenina N, Bader M, Cardoso MC, Leonhardt H (2005) Restricted mobility of Dnmt1 in preimplantation embryos: implications for epigenetic reprogramming. *BMC Dev Biol* **5:** 18

Hashimoto H, Horton JR, Zhang X, Bostick M, Jacobsen SE, Cheng X (2008) The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature* **455:** 826–829

Kim JK, Esteve PO, Jacobsen SE, Pradhan S (2009) UHRF1 binds G9a and participates in p21 transcriptional regulation in mammalian cells. *Nucleic Acids Res* **37:** 493–505

Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH (2003) Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* **13:** 1192–1200

Lei H, Oh S, Okano M, Juttermann R, Goss K, Jaenisch R, Li E (1996) *De novo* DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122:** 3195–3205

Leonhardt H, Page AW, Weier HU, Bestor TH (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* **71:** 865–873

Li E, Bestor TH, Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69:** 915–926

Li JY *et al* (2007) Synergistic function of DNA methyltransferases Dnmt3a and Dnmt3b in the methylation of Oct4 and Nanog. *Mol Cell Biol* **27:** 8748–8759

Martens JH, O'Sullivan RJ, Braunschweig U, Opravil S, Radolf M, Steinlein P, Jenuwein T (2005) The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J* **24:** 800–812

Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell* **99:** 247–257

Papait R, Pistore C, Negri D, Pecoraro D, Cantarini L, Bonapace IM (2007) Np95 is implicated in pericentromeric heterochromatin replication and in major satellite silencing. *Mol Biol Cell* **18:** 1098–1106

Rothbauer U, Zolghadr K, Muyldermans S, Schepers A, Cardoso MC, Leonhardt H (2008) A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. *Mol Cell Proteomics* **7:** 282–289

Schermelleh L, Haemmer A, Spada F, Rosing N, Meilinger D, Rothbauer U, Cristina Cardoso M, Leonhardt H (2007) Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. *Nucleic Acids Res* **35:** 4301–4312

Sharif J *et al* (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450:** 908–912

Spada F, Haemmer A, Kuch D, Rothbauer U, Schermelleh L, Kremmer E, Carell T, Langst G, Leonhardt H (2007) DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells. *J Cell Biol* **176:** 565–571

Uemura T, Kubo E, Kanari Y, Ikemura T, Tatsumi K, Muto M (2000) Temporal and spatial localization of novel nuclear protein NP95 in mitotic and meiotic cells. *Cell Struct Funct* **25:** 149–159

Zolghadr K, Mortusewicz O, Rothbauer U, Kleinhans R, Goehler H, Wanker EE, Cardoso MC, Leonhardt H (2008) A fluorescent two-hybrid (F2H) assay for direct visualization of protein interactions in living cells. *Mol Cell Proteomics* **7:** 2279–2287
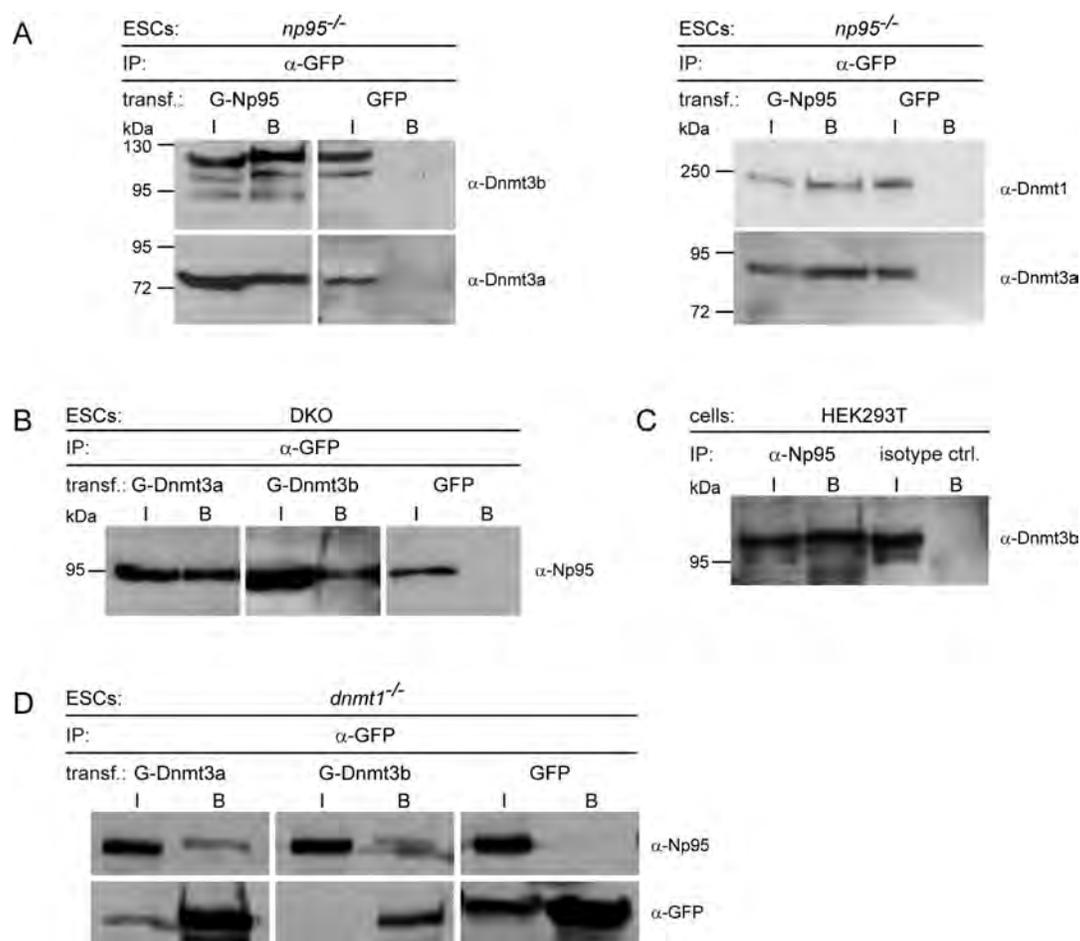
Supplementary Figure S1. Np95 interacts with *de novo* methyltransferases Dnmt3a and 3b. (**A**) Co-immunoprecipitation of endogenous Dnmt3a2 (left and right), Dnmt3b isoforms (left) and Dnmt1 (right) with GFP-Np95 transiently expressed in *np95*[-/-] ESCs. Left and right panels are from independent experiments where 2 and 4% of input (I) relative to bound (B) fractions was loaded, respectively. (**B**) Co-immunoprecipitation of endogenous Np95 with either GFP-Dnmt3a (left panel) or GFP-Dnmt3b1 (central panel) transiently expressed in DKO ESCs. Transient expression of GFP was used as control (right panel). GFP and GFP fusions were immunoprecipitated with GFP-trap as in experiments shown in Fig. 1B. 2% of input (I) relative to bound (B) fractions was loaded. (**C**) Co-immunoprecipitation of endogenous ICBP90/UHRF1 and DNMT3b in HEK293T cells. Antibodies to mouse proteins cross-react with the respective human homologues. 4% of input (I) relative to bound (B) fractions was loaded. (**D**) Co-immunoprecipitation of endogenous Np95 with either GFP-Dnmt3a or GFP-Dnmt3b1 (lower panel) transiently expressed in *dnmt1*[-/-] ESCs.
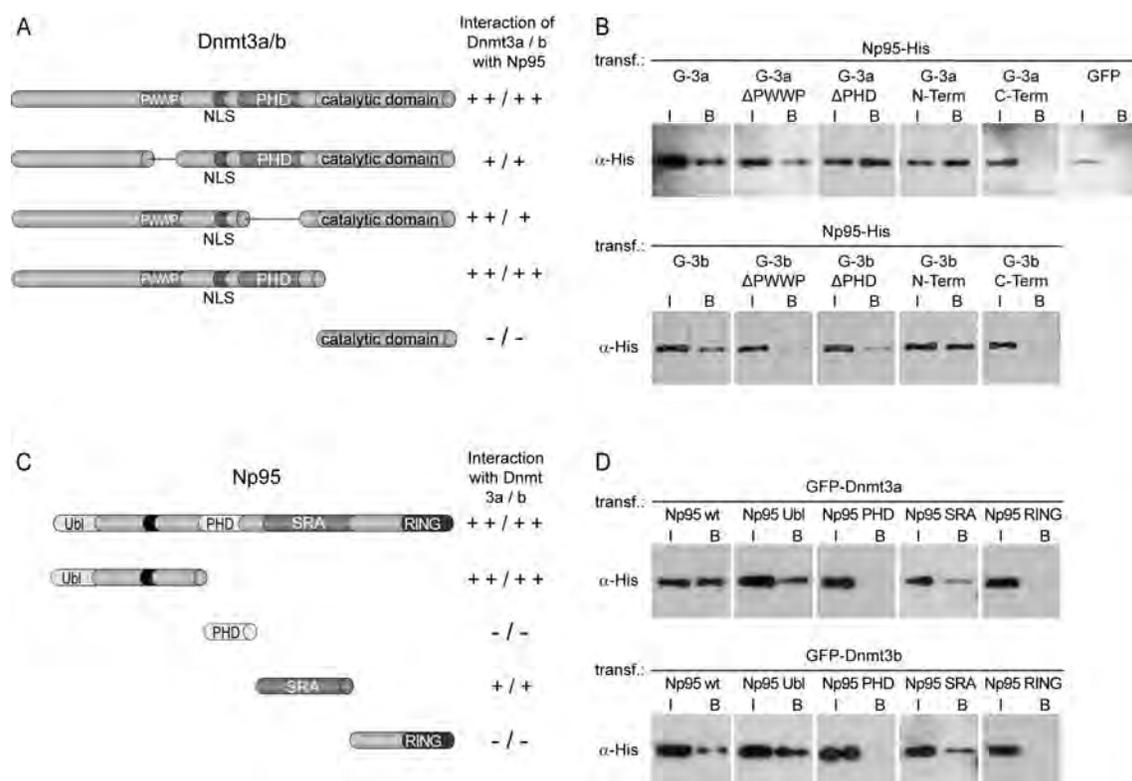
Supplementary Figure S2. Mapping the interaction domains of Dnmt3a/b and Np95. (**A**) Schematic representation of GFP-Dnmt3a/b fusion constructs used for mapping the interaction site with Np95 (N-terminal GFP tag is not shown). (**B**) Co-immunoprecipitation of Np95-His with GFP-Dnmt3 constructs (G-3a/b) from extracts of transiently transfected HEK293T cells. (**C**) Schematic representation of Np95-His constructs used for mapping the interaction site with Dnmt3a/b. (**D**) Co-immunoprecipitation of Np95-His domains shown in c with GFP-Dnmt3a/b constructs from extracts of transiently transfected HEK293T cells. G indicates the GFP fusion. The GFP-trap was used for all the immunoprecipitations in panels B and D. 0.5% of input (I) and 40% of bound (B) fractions were loaded. PWWP, domain with conserved pro-trp-trp-pro motif; NLS, nuclear localization signal; PHD, plant homeodomain; Ubl, ubiquitin-like domain; SRA, set and ring associated domain; RING, really interesting new gene domain. Results of mapping are scored by + or -.
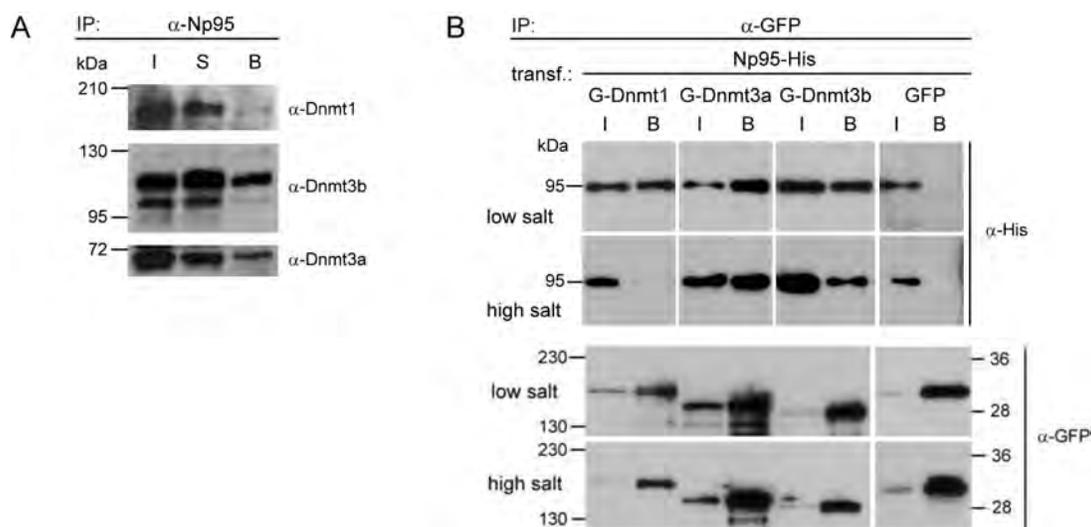
Supplementary Figure S3. Relative stability of Np95 interactions with Dnmt1, 3a and 3b. (**A**) The Dnmt3a/3b blot in Fig. 1A (shown here as middle and lower panels) was reprobed with an anti-Dnmt1 antibody (upper panel) to compare the relative amounts of endogenous Dnmts associated with Np95. The lower panel shows the Dnmt3a2 isoform. (**B**) The blots in Fig. 1C (shown here in the two upper panels) were reprobed with an anti-GFP antibody (two lower panels) to reveal that similar amounts of each GFP construct were immunoprecipitated in low and high salt conditions. Both results point to a tighter association of Np95 with Dnmt3a and 3b as compared to Dnmt1.
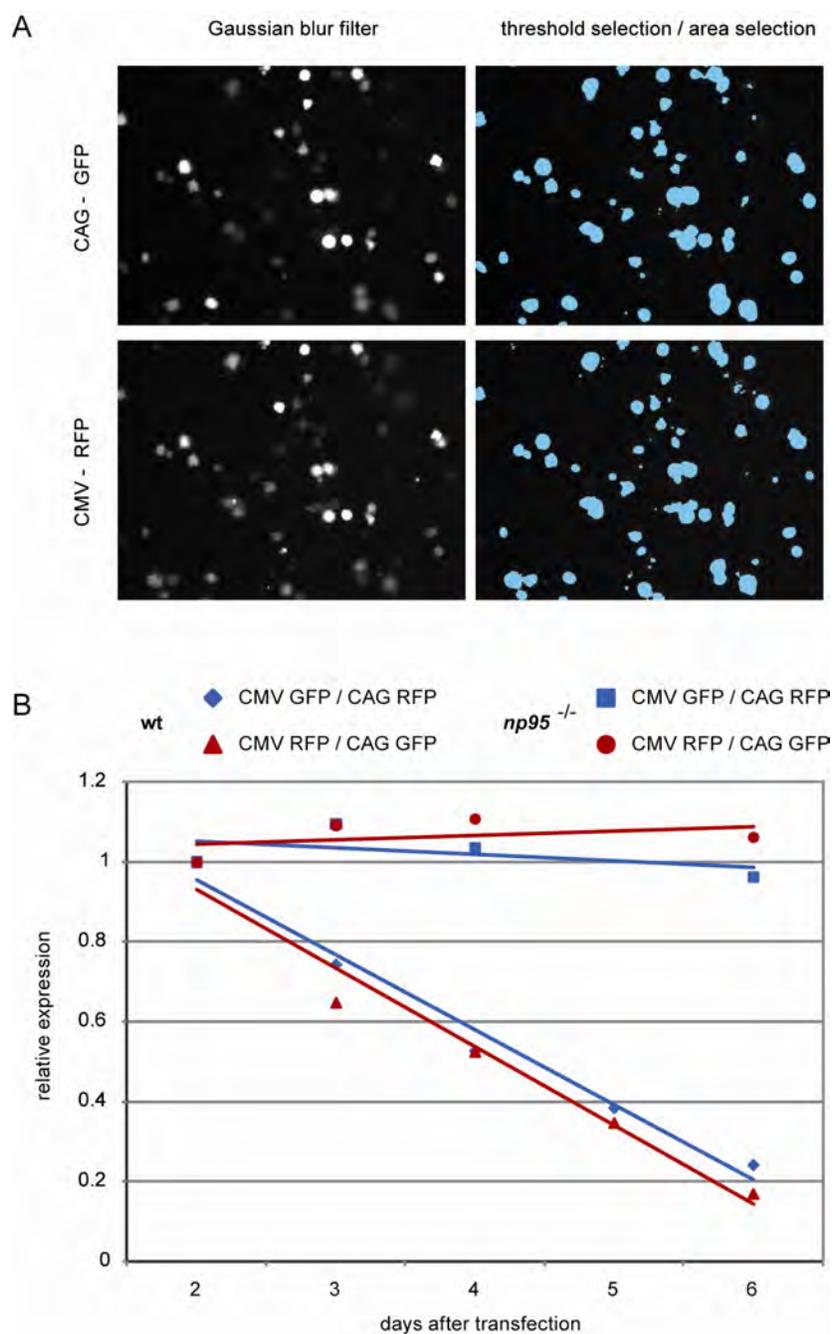
Supplementary Figure S4. (**A**) Automated procedure for quantification of fluorescent signals from digital micrographs for the promoter silencing assay. A macro was written for the ImageJ software that applies a Gaussian blur filter (left panel) and signal thresholding (right panel) to raw images (data not shown) and then calculates the total signal area. (**B**) Silencing assay results are not affected by the choice of fluorescent reporter. wt and $np95^{-/-}$ ESC were cotransfected with either CMV-driven mRFP and CAG-driven GFP (red) or CAG-driven mRFP and CMV-driven GFP (blue) expression constructs and the ratio of CMV- over CAG-driven fluorescence was quantified at the indicated time points after transfection as for Figure 2A.
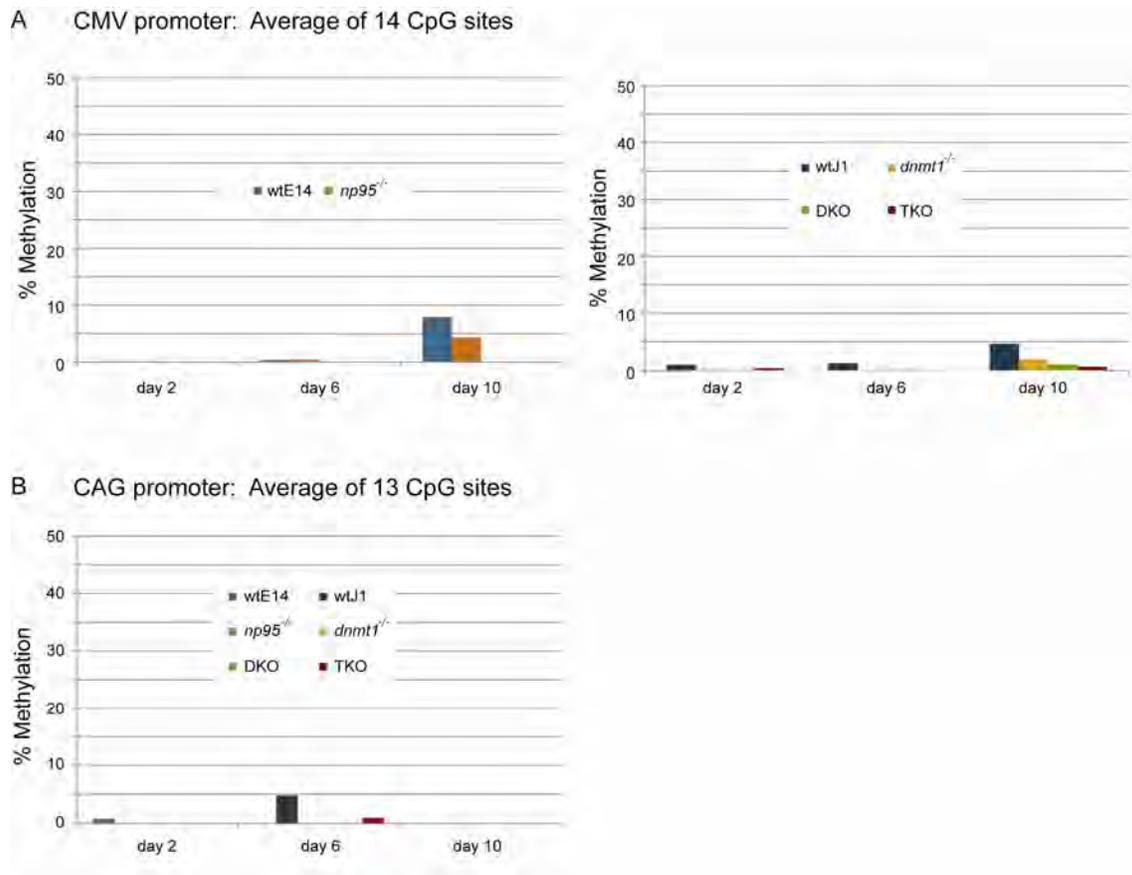
Supplementary Figure S5. Methylation analysis of CMV and CAG promoters 2, 6 and 10 days after transfection. ESCs with the indicated genotypes were transfected, sorted, total DNA was isolated and bisulfite treated as for Fig. 3. **(A)** The same proximal region of the CMV promoter was amplified and pyrosequenced as in Fig. 3. **(B)** A fragment of the CAG construct containing 13 CpG sites was amplified and subjected to pyrosequencing. The analyzed fragment spans across the 3' part of the promoter, the first exon and the 5' part of the first intron of the chicken ß-actin gene and is part of a CpG island. Methylation percentages at individual CpG sites within the respective promoter sequences are averaged. The plot on the left of A was derived from the same data presented in Fig. 3C.

Supplementary methods

Plasmid construction. The CMV-driven enhanced GFP construct was from Clontech (pEGFP-C1). To generate the CMV-driven mRFP construct (pCMV-mRFP) the coding sequence for eGFP in pEGFP-C1 was replaced with that for mRFP from pRSETB-mRFP (Campbell et al, 2002; provided by Roger Tsien). To create CAG-driven eGFP, mRFP and mCherry expression constructs (pCAG-eGFP-IB, pCAG-mRFP-IB and pCAG-mCherry-IB, respectively) sequences coding for the respective fluorescent proteins from pEGFP, pRSETB-mRFP and pRSETB-mCherry (Shaner et al, 2004; also provided by R. Tsien) were inserted downstream to the CAG promoter in the pCAG-IRESblast vector(Chen et al, 2003). The expression construct for Np95-His was described previously (Citterio et al, 2004). To generate expression constructs for GFP-Np95, Ch-Np95, GFP-Dnmt3a and GFP-Dnmt3b1 the sequences coding for Np95, Dnmt3a or Dnmt3b1 were then transferred from the respective CMV promoter-driven constructs (Chen et al, 2003; Citterio et al, 2004) to either pCAG-eGFP-IB or pCAG-mCherry-IB downstream to sequences coding for the fluorescent protein. GFP-Dnmt3a and GFP-Dnmt3b1 deletion constructs were generated by overlap extension mutagenesis(Ho et al, 1989) to remove the following amino acids from Dnmt3a and 3b1, respectively: 278-343 and 223-287 (ΔPWWP); 485-582 and 435-532 (ΔPHD). GFP fusion constructs of N-terminal regions (aa 1-629 and 1-580) and C-terminal domains (aa 630-908 and 581-859) of Dnmt3a and 3b, respectively, were generated by PCR cloning using full length constructs as templates. All constructs were characterised by sequencing and immunoblotting.

Supplementary references

Campbell RE, Tour O, Palmer AE, Steinbach PA, Baird GS, Zacharias DA, Tsien RY (2002) A monomeric red fluorescent protein. Proc Natl Acad Sci U S A 99(12): 7877-7882

Chen T, Ueda Y, Dodge JE, Wang Z, Li E (2003) Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Mol Cell Biol 23(16): 5594-5605

Citterio E, Papait R, Nicassio F, Vecchi M, Gomiero P, Mantovani R, Di Fiore PP, Bonapace IM (2004) Np95 is a histone-binding protein endowed with ubiquitin ligase activity. Mol Cell Biol 24(6): 2526-2535

Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77(1): 51-59

Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE, Tsien RY (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat Biotechnol 22(12): 1567-1572

Role of DNA methylation in silencing of pluripotency genes during embryonic development

(*manuscript in preparation*)

2.7 Role of DNA methylation in silencing of pluripotency genes during embryonic stem cell differentiation

We have recently shown that Np95 is involved in a common silencing process of ectopically expressed promoter constructs with Dnmt3a, Dnm3b and the histone methyltransferase G9a ((Meilinger et al, 2009) see chapter 2.6). Moreover, we showed that Np95 interacts with the *de novo* methyltransferases Dnmt3a and Dnmt3b. These findings pointed us towards the investigation of a possible Np95 role in silencing of endogenous promoter regions. We addressed this question by differentiating various knockout cell lines and compared expression and methylation level changes of the pluripotency factors *oct4* and *nanog* during differentiation*.*

Material and Methods

Cell lines: Wild type (J1 and E14)*, dnmt1$^{-/-,}$ DKO (dnmt3a$^{-/-}$dnmt3b$^{-/-}$), TKO (dnmt1$^{-/-}$, dnmt3a$^{-/-}$dnmt3b$^{-/-}$), and np95$^{-/-}$* cell lines described previously (Meilinger et al, 2009).

Cell culture, transfection and stable cell line generation: Mouse ES cells were cultured without feeder cells in gelatinized flasks and in DMEM supplemented with 16% fetal calf serum, 1000U/ml LIF, ß-mercaptoethanol and L-Glutamin as described previously (Frauer et al., 2011). Cell were transfected immediately after splitting with FuGENE HD (Roche, Mannheim) according to manufactures protocol.

GFP-positive cells were sorted 48 hours after transfection with a FACS Aria II (Becton – Dickinson). Sorted cells were plated under low density for stable cell line generation. The remaining GFP-positive cells were subsequently sorted and cultured until a stable pool of GFP-expressing cells was obtained.

Differentiation into embryoid bodies: a single cell suspension was obtained after resuspending in ESC medium without LIF. Cells were counted and adjusted to 3×10^4 cells/ml and 20 µl drops were spotted on the lid of a bacterial dish. The lid was turned and put back on the dish filled with PBS to prevent desiccation. After 2 or 4 days of incubation, EBs were collected and propagated in ESC medium without LIF. To prevent adhesion, EBs were transferred into a bacterial dish for further cultivation. Samples were collected every second or fourth day and genomic DNA as well as RNA was extracted and purified.

Genomic DNA and RNA Isolation from embryoid bodies: Genomic DNA was isolated using QIAmp DNA Maxi kit following the protocol for tissue lysis and isolation. Lysis was carried out 1-4 hours in lysis buffer to guarantee complete lysis. For simultaneous purification of DNA and RNA samples were collected in 1ml TRIzol reagent (Invitrogen) and either lysed by repetitive pipetting or by using a Dounce Homogenisator at later time points. Genomic DNA was further purified using a QIAmp Mini column and bisulfite treated using DNA Methylation Gold kit (Zymo Research).

Methylation Analysis based on Pyrosequencing: Bisulfite treated DNA was amplified in a PCR reaction with specific primer sets and nested PCR conditions for *oct4* and *nanog* promoter regions (see Table 3). The amplified PCR product was subjected to pyrosequencing carried out by varionostic GmbH, Ulm.

Table 3: Bisulfite primer sequences for *oct4* and *nanog* promoter

promoter	forward primer sequence	reverse primer sequence	length	CpG
oct4	GTTGTTTTGTTTTGGTTTTG GATAT	GTTAGAGGTTAAGGTTAGAGGGTGG		
	ATGGGTTGAAATATTGGGTT TATTTA	GTTAGAGGTTAAGGTTAGAGGGTGG-bio	454bp	12
nanog	GAGGATGTTTTTTAAGTTTTT TTT	TTATTATATTGATATGAGTGTGGG		
	AATGTTTATGGTGGATTTTG TAGGT	TTATTATATTGATATGAGTGTGGG-bio	369bp	6

DNA methylation and expression of oct4 and nanog during EB differentiation

We investigate a possible function of Np95 in silencing of endogenous promoter regions by differentiating *np95⁻ᐟ⁻, dnmt1⁻ᐟ⁻* ES cells and their respective wild types into EBs and compared expression and methylation level changes of the pluripotency factors *oct4* and *nanog*. Various differentiation protocols are available, but EB formation by the hanging-drop method provides the most homogenous population. The differentiation process in EBs is similar to the one observed in early embryonal development. During EB formation cells are not forced into specific cell lineages and thus this method allows differentiation into cells from all three germ layers. We controlled the size of the EBs by using a defined number of cells. EBs were harvested after 2-4 days and transferred into bacterial dishes to

avoid adhesion and further cultivation. To compare methylation of *oct4* and *nanog* promoter with transcript levels, we harvested EBs every 2-4 days and analyzed methylation level changes of *oct4* and *nanog* promoter regions as well as transcript levels. For methylation analysis, we isolated and bisulfite treated genomic DNA before subjecting it to pyrosequencing.

The pyrosequencing analysis of the *oct4* promoter in wild type ES cells shows that methylation steadily increases after 4 days of initiation of differentiation and reaches a plateau after 16 days (see Figure 2-5). The results are strikingly similar in both analyzed wild type cell lines: E14 and J1, which are the respective wild type backgrounds for $np95^{-/-}$ ES cells and the *dnmt* knockout ES cells (TKO and $dnmt1^{-/-}$). Notably, expression of *oct4* declines fast after initiation of differentiation and already at day 4 lowest transcript levels were detected indicating a complete down regulation (provided by Christine Schmidt). Similar effects can be observed in the analysis of the *nanog* promoter, yet methylation levels are significantly lower and reach only 40% methylation compared to approximately 70% that were found in the *oct4* promoter sequence (see Figure 2-6). Interestingly, we found down regulation of *oct4* and *nanog* in $np95^{-/-}$ and $dnmt1^{-/-}$ knock out cells to be surprisingly similar compared to the respective wild type cell lines and with a similar kinetics. Moreover, methylation levels were established at day 4 of differentiation, which is comparable to data obtained in wild type cell lines. However, methylation levels by far do not reach those observed in the respective wild type cells and stagnate at day 8 with around 20% and 10% in the *oct4* promoter and the *nanog* promoter, respectively. Furthermore, methylation levels decrease slightly after day 8 in $dnmt1^{-/-}$ ES cells but not in $np95^{-/-}$ ES cells. This observed difference was unexpected considering that $np95^{-/-}$ and $dnmt1^{-/-}$ ES cells have such strikingly similar methylation levels on both single copy promoter sequences and repetitive elements (see figure Figure 2-4) and given the fact that enzymatic activity of Dnmt1 is directly depended on the interaction with Np95 (Bostick et al, 2007; Sharif et al, 2007).
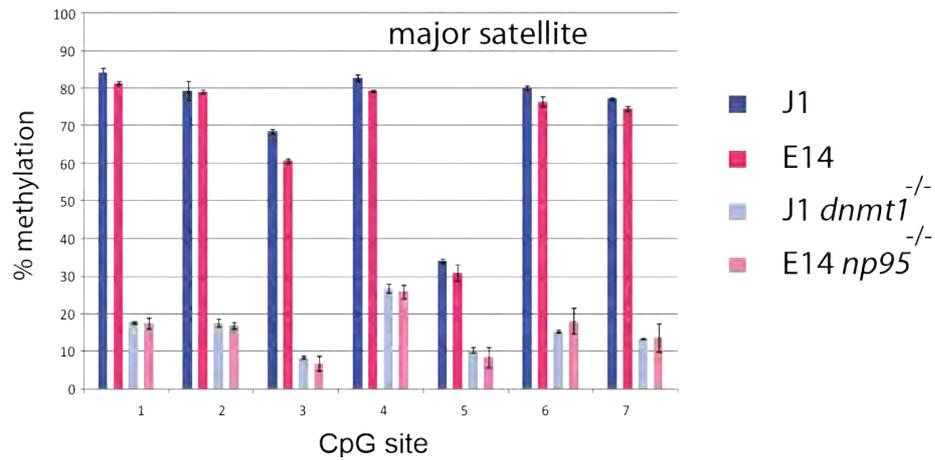
Figure 2-4: Comparison of methylation in wild type, *dnmt1*[-/-] and *np95*[-/-] ES cells. CpG methylation of major satellite repeats was analyzed by bisulfite treatment, PCR amplification and pyrosequencing for quantitative analysis. Both wild type cell lines J1 and E14 have methylation levels of around 80% at each CpG site. Strikingly, similar recused methylation levels were found in *dnmt1*[-/-] and *np95*[-/-] ES cells.

A methylation - *oct4* promoter
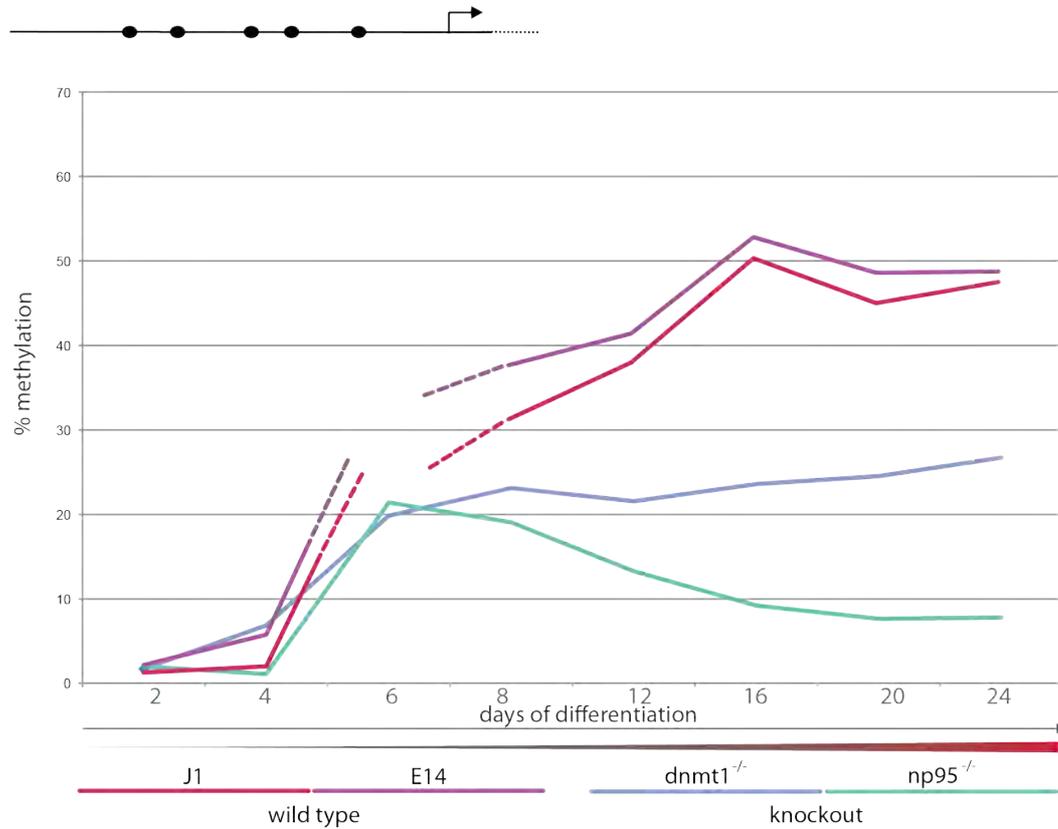
B expression - *oct4*
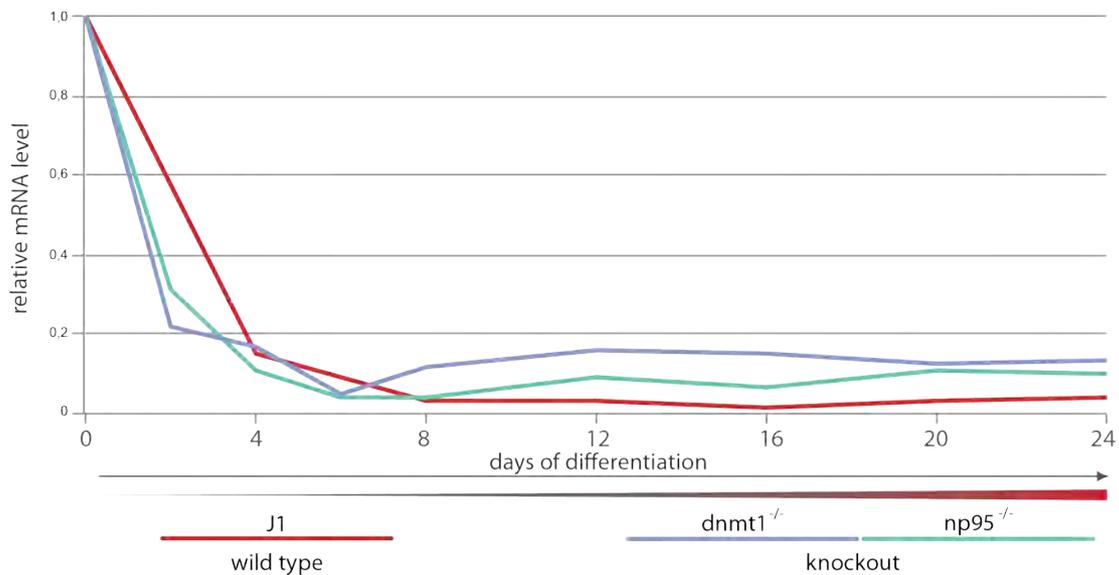
Figure 2-5: Methylation and expression of pluripotency factor oct4 during EB differentiation in wild type and knockout ES cells. Two wild type cell lines (J1, E14) were differentiated into EB up to 24 days and oct4 promoter methylation was analyzed every 2-4 days. Methylation levels from 5 CpG sites were averaged at each time point for each cell lines. (B) Simultaneously to genomic DNA extraction, mRNA was prepared for relative *oct4* mRNA expression at each time point (provided Christine Schmidt).

A methylation - *nanog* promoter
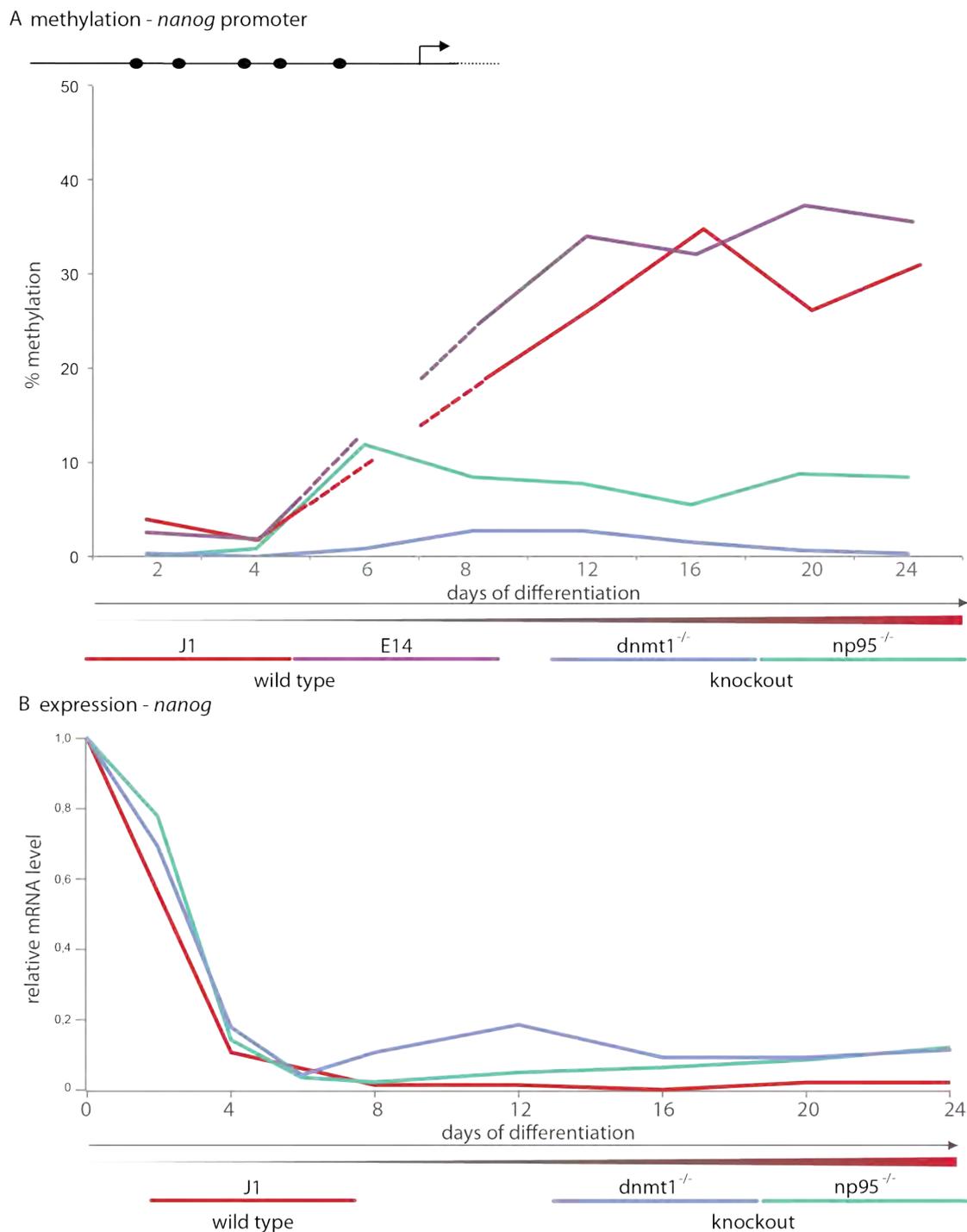
B expression - *nanog*

Figure 2-6 Methylation and expression of pluripotency factor *nanog* during EB differentiation in wild type and knockout ES cells: Two wild type cell lines (J1, E14) were differentiated into EB up to 24 days and nanog promoter methylation was analyzed every 2-4 days. Methylation levels from 5 CpG sites were averaged at each time point for each cell lines. (B) Simultaneously to genomic DNA extraction, mRNA was prepared for relative *nanog* mRNA expression at each time point (provided by Christine Schmidt).

Taken together, the comparison of the methylation level profiles in *np95⁻ᐟ⁻* and *dnmt1⁻ᐟ⁻* knockout cell lines revealed that *de novo* methylation carried out by Dnmt3a and Dnmt3b is not impaired in *np95⁻ᐟ⁻*ES cells, as we expected from the CMV promoter silencing data. Moreover, our data suggest that interaction of Np95 with the *de novo* methyltransferases Dnmt3a and Dnmt3b ((Meilinger et al, 2009) see chapter 2.7) is not involved in silencing of endogenous promoters. However, consistent with our previous findings that silencing of the CMV promoter precedes DNA methylation, our data on silencing of the pluripotency factors *oct4* and *nanog* clearly shows that initial down regulation is independent of DNA methylation and DNA methyltransferases.

Down regulation of oct4 in TKO cells

In addition to the analysis of wild type and *np95⁻ᐟ⁻* and *dnmt1⁻ᐟ⁻* ES cells we differentiated TKO cells and TKO cells stably expressing GFP-Dnmt1 (see chapter 2.3) into EBs and determined methylation of *oct4* and *nanog* promoter as well as their transcript levels during differentiation. We found that both; *oct4* and *nanog* mRNA levels decrease steadily during the first 4 days and stagnate at about 20%, indicating an initial down regulation but incomplete silencing of *oct4*. These findings strongly support further that the initial down regulation process of pluripotency genes is independent of DNA methylation and DNA methyltransferases.
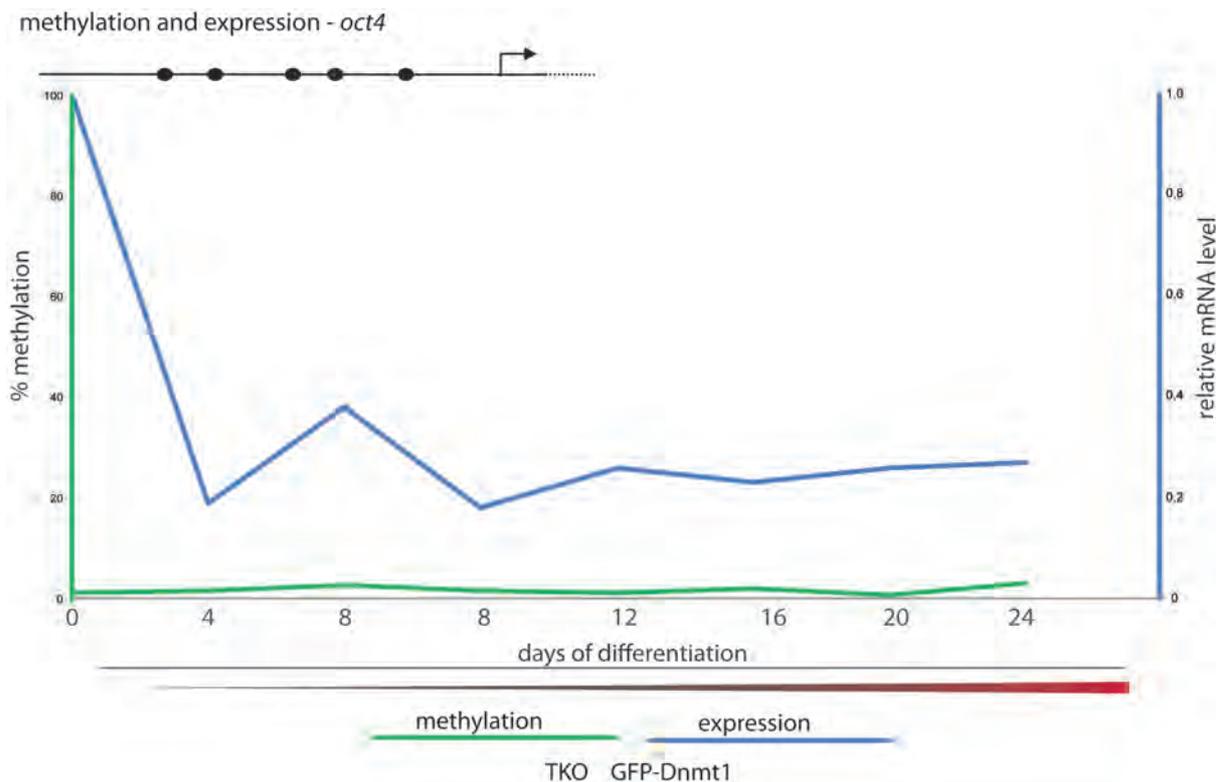
Figure 2-7 Methylation and expression of *oct4* in TKO cells rescued with Dnmt1. TKO cells stably expressing GFP-Dnmt1 were differentiated over 24 days into EBs. Methylation of *oct4* promoter was analyzed in those cells and no methylation was detectable at any time point. In comparison, mRNA levels of *oct4* were determined at the same time points during differentiation and down regulation was detectable reaching a stable relative level of around 20% (provided by Christine Schmidt).

Identification of differentiation area

(manuscript in preparation)

2.8 Identification of differentiating area
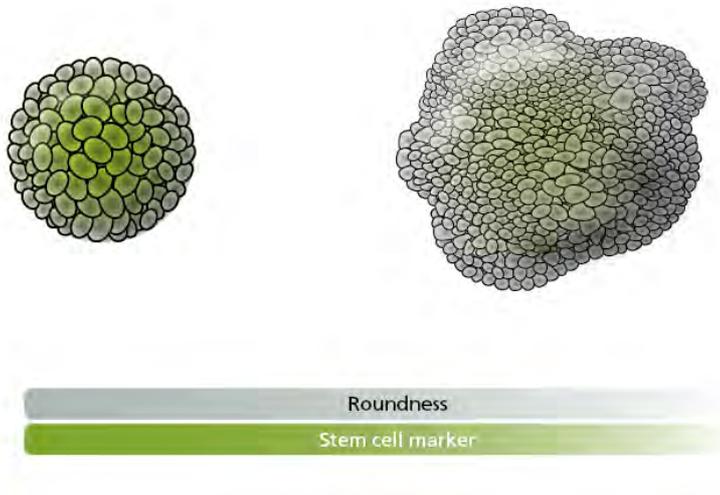
APPLICATION NOTE

Quantitative Analysis of Embryoid Bodies using Operetta

Key Features

- 3D confocal imaging
- Cell identification in maximum intensity projection image
- Quantification of marker positive cells

Identification of differentiating area

Background

Embryoid bodies (EB) are **spherical three-dimensional cell aggregates** derived from embryonic stem cells which serve as model system for embryonic development. The formation of EBs allows studying cellular and molecular interactions in a three-dimensional manner in the early stages of embryogenesis. The differentiation process finally leads to a large variety of differentiated cell types that can be studied and characterized *in vitro* [Desbaillets *et al.,* 2000; Itskovitz-Eldor *et al.,* 2000; Son *et al.,* 2011].

High content imaging enables quantitative analysis of fluorescent marker intensities combined with cellular morphology and textures, and has proven to be a precise and rapid tool for characterizing cells. Here, we present an imaging approach for cell identification and stem cell marker quantification in a three-dimensional object using the Operetta® High Content Imaging System and the Harmony® High Content Imaging and Analysis Software. Classification of stem cell marker positive and negative cells allowed the discrimination of differentiated and undifferentiated areas of the EBs.

Application

ES cells (ESCs) were used that express GFP under control of the Oct4 promotor a well known stem cell marker [Rodriguez *et al*., 2007]. To study differentiation process ESCs were differentiated into EB using the hanging drop method. ESCs were counted and adjusted to 3×10^4 cells/ml in ESC medium without LIF and 20µl drops were spotted on the lid of a bacterial dish. The lid was turned and put back on the dish filled with PBS to prevent desiccating. After 4 days of incubation in hanging drops, EBs were collected and propagated in ESC medium without LIF. EBs were collected at day 4 and day 8 washed with PBS and fixed with 3,7% formaldehyde for 30 minutes.

Fixed 4-day and 8-day old EBs were permeabilized with 0.5% Triton X-100 (15 min). The nuclei and actin cytoskeleton were stained by incubating them with Hoechst 33342 (20 µM) and Rhodamine Phalloidin (Invitrogen®, final conc. 1 unit/ml), respectively. Finally the EBs were transferred to a 384-well CellCarrier™ microtiter plate (PerkinElmer, 6007550) for imaging.

The images were acquired on the Operetta using the confocal mode and a 20X high NA objective. A z-stack of 80 planes with a total distance of 160 µm was measured. For image analysis the stack processing method "maximum intensity projection" was used. With this procedure all planes of the stack measurement are reduced to a single image. The maximum intensity value was applied to each pixel of the generated projection image gets across the stack planes. The projection images (Figure 1) were used to further calculate the morphology and intensity properties of the 4-day and 8-day old EBs.
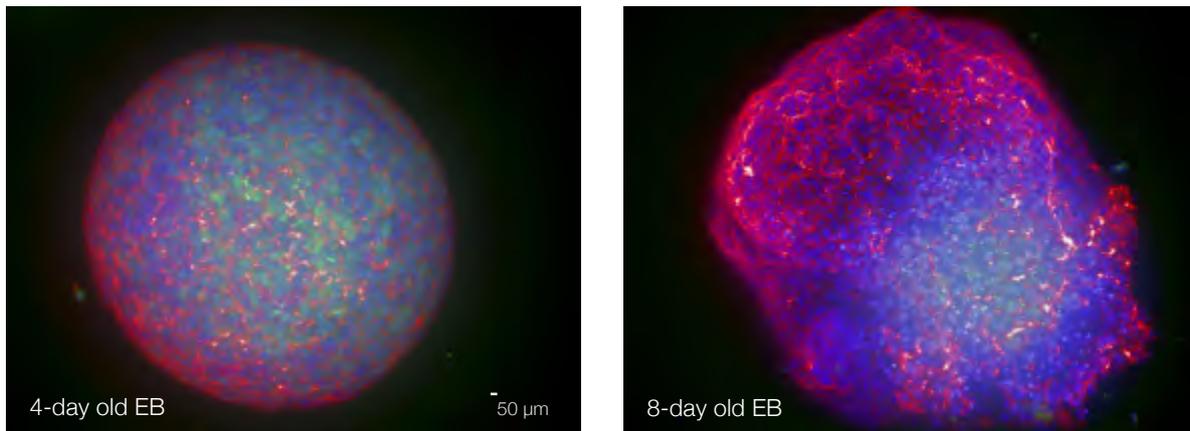
Figure 1

Figure 1: Maximum intensity projection images of 4-day old EB (left) and 8-day old EB (right). The maximum intensity projection was generated using all 80 planes of the acquired z-stack. Images show a false color overlay of nuclei (blue), actin (red) and Oct4 (green). Compared to the 4-day old EB, the 8-day old EB is larger, more complex and expresses less stem cell marker Oct4, indicating advanced differentiation.

For image analysis, the EB region could be clearly discriminated from background by texture-based segmentation using the actin image. By calculating the morphology properties of the EB region, there was already a significant readout comparing area and roundness of the bodies. While the total area of the 8-day old EB is clearly increased, the roundness decreases due to an uneven and complex structured surface of the body (Figure 3). The Hoechst channel enabled further segmentation of the EB into single nuclei. By calculating the Oct4 intensity properties of each nucleus, cells were subdivided into Oct4 intensity classes allowing the generation of a histogram (Figure 3). The histogram clearly shows that the main proportion of cells from the 4-day old EB show a high Oct4 intensity, while the 8-day old body contains mainly cells with lower intensities. This provides further evidence of the advanced differentiation of the growing EB. Setting an Oct4 intensity threshold allows the separation of the cells into marker positive and marker negative cell populations, enabling the identification of differentiated and undifferentiated regions inside the body (Figure 2).
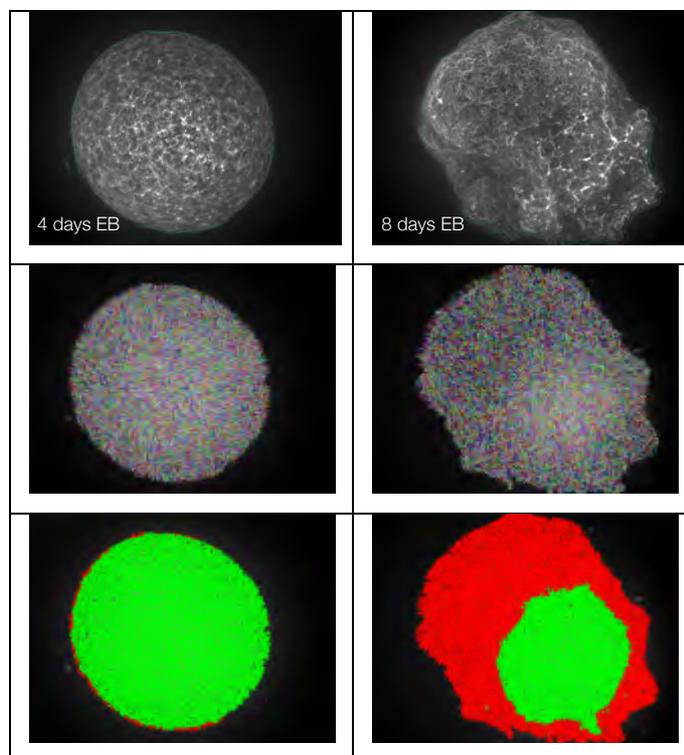
Figure2   Figure3
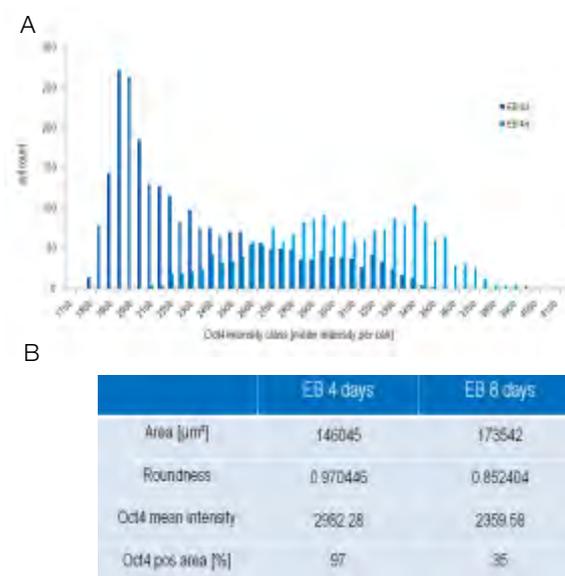
Figure 2: Image segmentation and identification of Oct4 positive/negative area in the maximum intensity projection image. Texture-based segmentation enables the EB region (upper panel, green border) to be found. The generated nuclei mask in the EB region is used to calculate the intensity properties of each nucleus in the Oct4-channel (middle panel). Setting an intensity threshold allows a marker positive and a marker negative cell population (lower panel) to be selected, and an undifferentiated area (shown in green) and differentiating area (shown in red) can be located.

Figure 3: Quantification of several readouts indicating differentiation of EB. The Oct4 mean intensities of all detected nuclei were subdivided into intensity classes. The histogram (A) shows the number of nuclei in each of the intensity classes for both EBs. The 4-day old EB (light blue bars) mainly contains cells with high Oct4 intensity whereas the 8-day old body (dark blue bars) essentially contains cells with low Oct4 intensity. The table (B) summarizes a selection of the most important quantitative readouts evaluating the progress of EB differentiation. As the increased area of the 8-day old body indicates growth, the decreased values for roundness and Oct4 positive area simultaneously imply differentiation of the body.

Conclusion

In this study, we analyzed three-dimensional embryoid bodies (EBs) using the Operetta High Content Imaging System. The applied method of confocal stack measurement combined with analysis of the maximum intensity projection image is a robust approach for characterizing 3D bodies. The maximum intensity projection comprises the extracted information from all planes rather than each plane individually so that image analysis can be achieved much more quickly.

Alongside the quantification of area, roundness and Oct4 intensity the Harmony High Content Imaging and Analysis Software enables the discrimination of differentiated and undifferentiated regions of the body by classifying cells according to their Oct4 expression level. For the EBs studied here we observed a differentiation in the rim and outer regions while the center remained undifferentiated after 8 days of incubation. It would be quite straightforward to further characterize the differentiation process by including additional differentiation markers into the high content imaging approach.

Desbaillets I, Ziegler U, Groscurth P, Gassmann M (2000): Embryoid bodies: an in vitro model of mouse embryogenesis. Exp Physiol., 85(6), 645-651.

Itskovitz-Eldor J, Schuldiner M, Karsenti D, Eden A, Yanuka O, Amit M, Soreq H, Benvenisty N (2000): Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. Mol Med., 6(2), 88-95.

Rodriguez RT, Velkey JM, Lutzko C, Seerke R, Kohn DB, O'Shea KS, Firpo MT (2007): Manipulation of OCT4 Levels in Human Embryonic Stem Cells Results in Induction of Differential Cell Types. Exp Biol Med., 232(10), 1368 - 1380.

Son M, Kim H, Kim M, Cho YS (2011): Physical Passaging of Embryoid Bodies Generated from Human Pluripotent Stem Cells. PLoS One, 6(5), e19134.

Authors

Stefan Letzsch

PerkinElmer
Cellular Technologies Germany GmbH
Cellular Imaging & Analysis
Hamburg, DE

Daniela Meilinger
LMU Biocenter
Munich, DE

3. [Discussion]

3.1 Analyzing DNA methylation in living cells

The main focus of this study was the analysis of DNA methylation changes as a result of the regulation and activity of DNA methyltransferases in living cells. DNA methyltransferase activity is commonly measured *in vitro* using DNA substrates in either radioactive or fluorescent assays. Within this context, it is important to stress that results of methyltransferase assays might differ drastically, dependent on whether they are performed *in vitro* or *in vivo.* DNA substrates used in *in vitro* assays are usually naked DNA oligonucleotides, where physiological compaction into nucleosomes is missing. Hence, analysis of DNA methyltransferase activity in its physiological environment in living cells is ultimately the most important experimental assay to determine methyltransferase activity. To this aim, complementation or rescue assays in the respective knockout cell lines are used to determine the activity of DNA methyltransferases and their "mutants" *in vivo.* Knock out cells are transfected with the respective DNA methyltransferase expression vector followed by a methylation analysis to determine the ability of the protein to re-methylate *in vivo*. A rescue assay can either be performed transiently or by generating stably expressing cell lines, which allows studying long-term effects concerning status and levels of DNA methylation in addition to various other applications like biochemical approaches and/or microscopy. We aimed to investigate DNA methyltransferase activity in living cells and therefore improved the generation of stable ES cell lines by a new FACS based approach (see chapter 2.3). This new protocol enabled us to establish stable cell lines faster than conventional methods. Furthermore, we improved classic DNA methylation protocols, to obtain reliable, reproducible and quantitative data.

3.1.1 Improved Stable ES Cell Line Generation

To generate a cell line stably expressing a fluorescent fusion protein, cells are transfected with respective expression constructs of a protein-of-interest and subsequent rounds of FACS sorting for positive fluorescent cells and expansion are followed including intervals of propagation in tissue culture. This results in the creation of a cell population that stably expresses the fluorescent fusion protein. This new FACS based approach can also be

adapted for other stable cell line production like somatic cell lines and has several crucial advantages.

(i) This approach allows the generation of a pool of stably expressing cells in a shorter amount of time compared to classic methods based on solely antibiotic selection. Only 2-3 weeks of subsequent sorting is necessary to obtain a pool of stably expressing cells and adjacent single cell sorting can be used to generate single clones. In addition, it is possible to generate several different cell lines in parallel when using this approach. Furthermore it is possible to combine this approach with antibiotic selection, which makes it even easier to generate multiple stable cell lines in parallel. (ii) By using fluorescent fusion proteins, a high throughput microscopy system can be used to screen single clones in a 96 well format to compare expression levels and choose clones with different expression levels for further analysis. (iii) Those stable cell lines can be used for protein-protein interaction studies using the GFP-trap technique without the limitations of low transient transfection efficiencies (Rothbauer et al, 2008). (iv) Our approach allows generating double stable cell lines by using two different fluorescent markers in only one transfection round.

Having in hand these cell lines with fluorescent marker proteins opens the whole field of combining *in vitro* and *in vivo* studies with only one cell line, without the need to repeatedly perform transient transfections. Moreover, it offers the possibility to study protein-protein interactions *in vitro* and *in vivo*, cell cycle depended localization, protein kinetics and methyltransferase activity in living cells (see figure Figure 3-1).
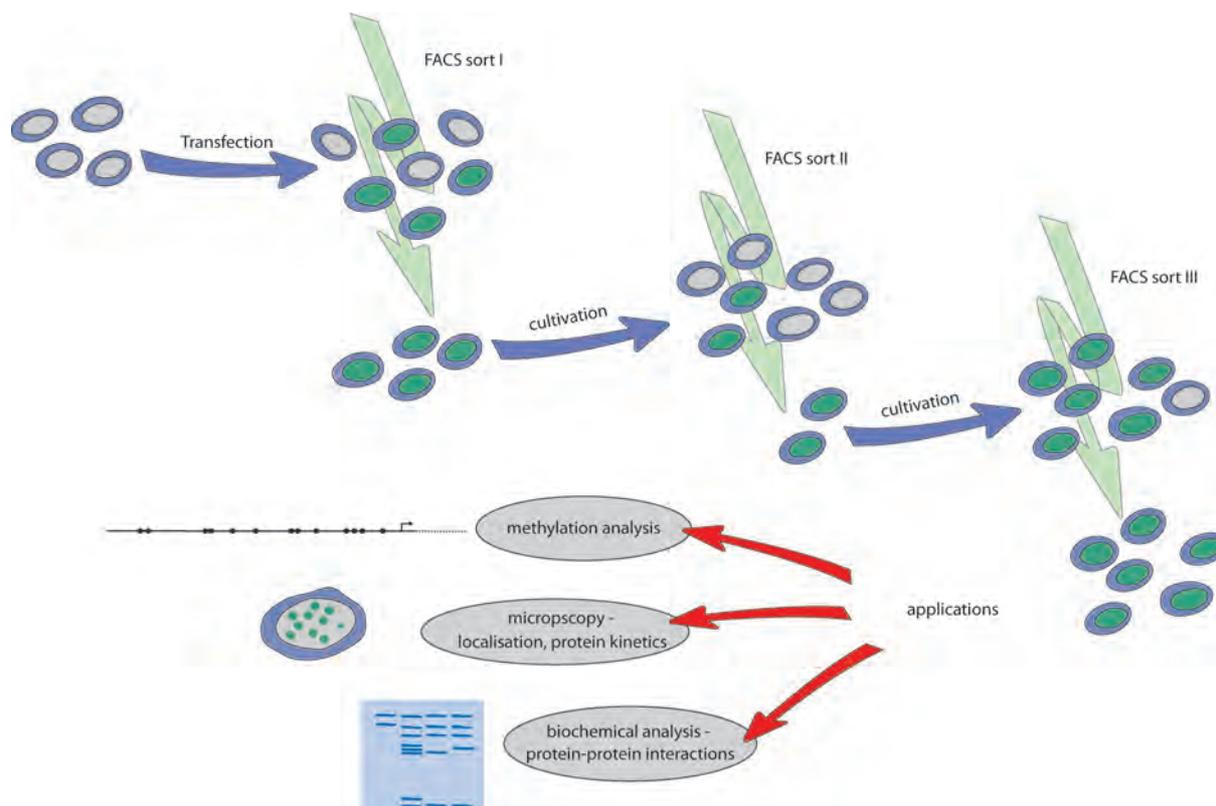
Figure 3-1 Overview FACS bases approach to generate stable cell lines for methylation analysis. ES cells are sorted repeatedly for GFP positive cells, indicated by the green nucleus in the cartoon. After the first round of sorting, cells are cultivated and sorted again for the remaining GFP positive cells. After a third round of sorting, a stably expressing pool of GFP positive cells is generated. Possible following applications are methylation analysis, microscopy including localization studies as well as protein kinetics analysis with FRAP and biochemical analyses including the study of protein-protein interactions.

3.1.2   Improved DNA methylation analysis protocol

Due to the importance of detailed methylation analysis in studying DNA methyltransferase activity, gene specific regulation, as well as cancer diagnostics, various protocols to determine DNA methylation levels have been established. Bisulfite treatment of DNA is often considered as the gold standard in DNA methylation analysis and many methods are based on the selective deamination of cytosine but not methylated cytosine bases and subsequent PCR amplification of a defined region (Frommer et al, 1992). Based on this reaction principle, two protocols, Combined Bisulfite Restriction Analysis (COBRA) and classical bisulfite sequencing are established and commonly used. However, both protocols (COBRA and bisulfite sequencing) have clear disadvantages and limitations. More specifically, bisulfite treatment and PCR amplification can result in a methylation dependent creation of new restriction sites, which is used to determine the methylation

level at the respective site in an COBRA assay (Xiong & Laird, 1997). However, methylation analysis in COBRA is limited to the created restriction site. In classical bisulfite sequencing protocols, the PCR product is cloned into a target vector and individual clones are selected and sequenced, resulting in detailed information of each CpG sites within the amplified sequence. Despite the fact that this method grants an analysis of the methylation status of each CpG site within single molecules, obtaining quantitative data requires sequencing of a large number of clones, which is very laborious and cost intensive.

To overcome these limitations and to detect even subtle changes in the methylation profile, we established a pyrosequencing methylation analysis protocol. Pyrosequencing is based on bisulfite treatment, followed by amplification of a specific DNA sequence with one biotinylated primer to allow a sequencing by synthesis application (Colella et al, 2003; Ronaghi, 2001). Pyrosequencing is on the one hand a mid- to high throughput approach but enables on the other hand the simultaneous quantification and analysis at single CpG site resolution. Moreover, the method has a detection limit of only 1-5% methylation (Tost et al, 2003). Although the generation of quantitative methylation data with only one-step PCR amplification is beneficial, pyrosequencing can only measure amplicons up to a length of 120 base pairs (bp), which makes it necessary to use several sequencing reactions when analyzing whole CpG islands and large promoters.

Since the method is based on PCR amplification a possible bias must be considered. Differences in secondary structure formation between methylated versus unmethylated DNA after bisulfite treatment may result in different amplification efficiencies during PCR. Indeed, such observations have already been documented during the methylation analysis of the human p16 promoter, where unmethylated DNA was amplified with a greater efficiency than the methylated sequences (Warnecke et al, 1997). To minimize such bias we systematically tested PCR conditions for the sequences subjected to pyrosequencing. The addition of TMAC (tetramethylammonium chloride) to a PCR reaction has been reported to reduce bias in normal PCR amplifications, while the addition of betaine (*N,N,N,-trimethylglycine*) has been shown to drastically reduce a PCR induced bias in bisulfite treated PCR amplifications (Chevet et al, 1995; Voss et al, 1998). Methylation analysis of control DNA titrations (0%, 10%, 25%, 50%, 75%, 90% and 100% methylated) by pyrosequencing showed expected methylation values at each site and in each analyzed

sequence. These results confirm that our bisulfite PCR protocols have nearly no detectable bias and that both, methylated and unmethylated sequences are amplified with a similar efficiency.

Recently, a new modified base, 5-hydroxymethylcytosine (5-hmC) was discovered, that shows chemically similar properties to 5mC and thus might be indistinguishable in standard bisulfite based methods (Kriaucionis & Heintz, 2009; Tahiliani et al, 2009). Indeed, a study showed that 5-hmC, like 5-mC is resistant to deamination, indicating that standard bisulfite methods, including pyrosequencing are not able to distinguish 5-hmC from 5-mC (Huang et al, 2010). Thus, it is important to take this finding into consideration when analyzing methylation data that is obtained from bisulfite-based methods, in particular in cells with a high amount of 5-hmC such as ES cells.

3.2 Role and dynamics of the regulatory subunits of Dnmt1

Faithful maintenance of DNA methylation is coupled with replication of DNA during S-Phase (Chuang et al, 1997; Easwaran et al, 2004; Leonhardt et al, 1992). Dnmt1 is the main maintenance methyltransferase in mammalian cells and interacts with the replication machinery via its PCNA binding domain (PBD). It has been previously proposed that this interaction is crucial for proper maintenance of DNA methylation especially in early S-Phase (Easwaran et al, 2004). In this study we further investigated the role of the Dnmt1 interaction with the replication machinery by a detailed analysis of the PBD-PCNA interaction using different experimental approaches ((Schermelleh et al, 2007) see chapter 2.1). By comparing PBD domains of different species, a highly conserved region termed PCNA-interacting peptide (PIP)-Box was detected and specific point mutations were introduced. These PBD point mutants abolish interaction with PCNA *in vitro* and *in vivo* and one point mutant Dnmt1^{Q162E} was further used to determine the ability to restore methylation in mouse *dnmt1*$^{-/-}$ ES cells. Interestingly, we found that Dnmt1^{Q162E} was able to restore methylation of single copy genes as well as of repetitive sequences in a similar fashion compared to wild type Dnmt1 suggesting that the catalytic activity is not impaired in Dnmt1^{Q162E}.

Discussion

Our data clearly imply that the PBD-PCNA interaction is highly transient, but not strictly required for the restoration of global methylation levels *in vivo* ((Schermelleh et al) see chapter 2.1). However, this interaction enhances the methylation efficiency by a factor of two and may thus function as an additional safety mechanism. This enhanced efficiency might be important in cell types that contain low Dnmt1 expression levels (Ratnam et al, 2002; Robertson et al, 1999). In addition, Dnmt1 has also been found at sites of DNA repair mediated by a similar targeting mechanism via the PBD-PCNA interaction (Mortusewicz et al, 2005), thus contributing to genomic integrity by coupling epigenetic and genetic repair pathways. Furthermore, also other domains of the regulatory N-terminus of Dnmt1 e.g. the TS domain are known to be responsible for targeting Dnmt1 to heterochromatic sites (Leonhardt et al, 1992; Easwaran et al, 2004).

The TS domain is the largest regulatory subdomain within the N-terminal domain of Dnmt1 and is a crucial structure involved in targeting Dnmt1 to heterochromatin, N-C terminal interaction, protein-protein interaction and catalytic activity (Easwaran et al, 2004; Fellinger et al, 2009; Leonhardt et al, 1992; Margot et al, 2000). Furthermore, the TS domain interacts with Np95 (Uhrf-1), a multi domain protein that is crucial and essential for DNA methylation (Bostick et al, 2007; Sharif et al, 2007; Fellinger, 2009). Strikingly, Np95 knockout ES cells have a similar loss of genomic methylation compared to Dnmt1 knockout ES cells underlining the crucial role of the TS-Np95 interaction in maintenance of DNA methylation (Feldman et al, 2006; Bostick et al, 2007; Sharif et al, 2007)see also figure Figure 2-4). Interestingly, the catalytic activity of a Dnmt1 mutant protein, lacking the complete TS domain, is not impaired as measured *in vitro* assays, while the catalytic activity is abolished in living cells (unpublished data Weihua Qin). Given the fact that *in vitro* assays were performed on naked DNA, which is not the physiological substrate of Dnmt1, a discrepancy between *in vivo* and *in vitro* studies might occur and underlines the importance to analyze the regulation and function of Dnmt1 in living cells.

Several recent studies, resolving crystal structures of Dnmt1 fragments provided new hints to understand function and regulation of Dnmt1 (Song et al, 2011; Syeda et al, 2011; Takeshita et al, 2011). According to the crystal structure, the TS domain binds to the catalytic domain of Dnmt1 and thereby inhibits the catalytic activity of Dnmt1, the TS domain has to be released from the catalytic pocket in order to allow substrate binding by

the catalytic domain (Takeshita et al, 2011). Thus, the crystal structure of the TS domain suggests an inhibitory role of the TS for the catalytic activity of Dnmt1. Furthermore it was proposed that interaction of the TS domain with Np95 would then allow Dnmt1 to bind to its DNA target site (Syeda et al, 2011). This autoinhibitory role for the TS domain together with our own data showing that deletion of the TS leads to an abolishment of the catalytic activity of Dnmt1 sheds new light onto the functions of the TS domain and underlines its crucial role in Dnmt1. Further assays performed in living cells would be necessary to confirm the autoinhibitory role of the TS domain. For this purpose, mutagenesis based on the crystal structure followed by a rescue assay in ES cells would be necessary to determine the effect of Dnmt1 subdomains on catalytic activity in living cells.

Given the critical role of the TS domain in catalytic activity and targeting of Dnmt1 to heterochromatin, this domain has been intensively studied in our own lab and is subject of ongoing work. We aim to determine potential interaction partners and characterize the respective binding properties in different cellular states. FRAP experiments using the isolated TS domain as well as several N-terminal deletion constructs demonstrated a surprisingly strong association to chromatin in the nucleus (unpublished data, Daniela Meilinger diploma thesis). This interaction seems to be independent of DNA methylation levels and Np95 interaction (unpublished data, Andrea Rottach). Preliminary data show that the isolated TS domain binds to H3K9me3 peptides *in vitro* (unpublished data, Patricia Wolf), a finding that is consistent with the clear enrichment of the TS domain within heterochromatin *in vivo.*

The CXXC zinc finger motif of Dnmt1 has been proposed to be essential for substrate binding, allosteric activation and catalytic activity of Dnmt1, and was suggested to be involved in N-and C-terminal interaction (Margot et al, 2000; Fatemi et al, 2001; Pradhan et al, 2008). The available data regarding the binding specifies of the CXXC zinc finger of Dnmt1 are controversial, since they have used different construct length, places and size of deletion of Dnmt1 CXXC zinc finger and also used constructs from different species. To address the binding specificity and the function of the CXXC zinc finger of Dnmt1 we generated several deletion constructs and performed a detailed *in vitro* and *in vivo* analysis of those constructs (Frauer et al, 2011) see chapter2.1). In this study we used a precise deletion of the mouse zinc finger motif (aa 652-697), trying to prevent disruption of the

surrounding protein structure, as well as the isolated domain for *in vitro* analysis. To resolve the function of the CXXC finger domain and to clarify the role in DNA binding we performed an *in vitro* binding assay developed in our laboratory (Frauer & Leonhardt, 2009). Our data showed that the CXXC zinc finger preferentially binds to unmethylated DNA *in vitro*. However, deletion of the zinc finger within the context of the full-length protein (Dnmt1ΔZn) did not impair the capability to restore DNA methylation in living *cells.*

Taken together, our results suggest that the CXXC zinc finger domain of Dnmt1 is negligible for faithful DNA methylation maintenance in ES cells since deletion of the CXXC Zn finger within the context of the full-length protein (Dnmt1ΔZn) is able to restore methylation patterns similar to the wild type Dnmt1 enzyme. However, since only few sequences were analyzed, possible sequence specific effects cannot be excluded. Genome wide mapping of methylation would be a possible approach to study effects of Dnmt1ΔZn on specific sequences within the genome. In addition, long-term effects on a possible impaired DNA methylation maintenance cannot be excluded due to the fact that these results were done in a transient rescue experiment. Therefore, a cell line stably expressing Dnmt1ΔZn or even a mouse model would be necessary to address eventual long-term effects on DNA methylation maintenance. Furthermore, a mouse model would allow studying potential effects during differentiation, development and different cell types by following methylation level changes during differentiation and detailed analysis of different cell types.

One possible function for the CXXC zinc finger domain of Dnmt1 could be an auto-inhibitory effect due to the binding specifies of the zinc finger to unmethylated DNA sequences and therefore prevent aberrant *de novo* methylation of Dnmt1. Indeed, recent structural studies further support this hypothesis by showing co-crystal structures of Dnmt1 with unmethylated DNA where the CXXC zinc finger is proposed to bind to unmethylated DNA (Song et al, 2011). By binding to unmethylated DNA the CXXC zinc finger prevents the catalytic pocket of Dnmt1 to bind to its substrate. In addition, an *in vitro* DNA methylation activity assays could furthermore show an increased methylation activity of Dnmt1 on unmethylated substrates when the zinc finger is missing thus proposing a *de novo* methylation activity (Song et al, 2011). Notably, rescue experiments in ES cells using

wild type Dnmt1 and Dnmt1ΔZn showed that methylation at the imprinted promoter region of H19a remains unmethylated in both rescue experiments arguing against a possible *de novo* activity of Dnmt1 lacking the CXXC zinc finger. To fully investigate whether deletion of the CXXC zinc finger in Dnmt1 would lead to an increased *de novo* activity of Dnmt1 in living cells, rescue assays in unmethylated TKO cells should be performed and genome wide methylation levels analyzed.
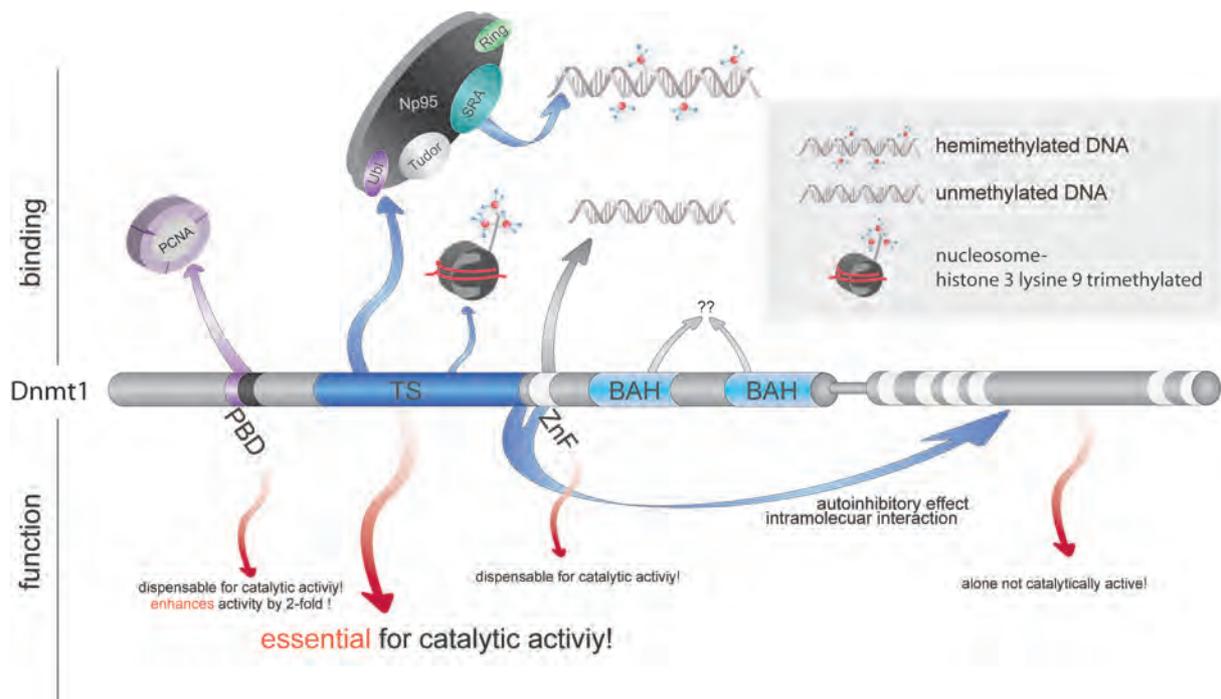
Figure 3-2 Domain overview of Dnmt1 with essential interaction partners and their influence on catalytic activity. The PBD domain of Dnmt1 interacts with PCNA enhances catalytic activity by two fold but deletion of this domain does not affect catalytic activity. The TS domain interacts with the UBL domain of Np95, H3K9me3 and is essential for catalytic activity. Np95 binds via the SRA domain to methylated DNA and recruits Dnmt1. The Zn finger binds unmethylated DNA (*in vitro*). Both, TS and Zn are involved in intramolecular N-C terminal interaction and thereby influence catalytic activity of Dnmt1 (autoinhibitory effect). The Zn finger seems dispensable for catalytic active *in vivo*.

3.3 Dnmt1 and its role beyond DNA methylation maintenance

Dnmt1 maintains DNA methylation patterns after replication and shows a clear preference for hemimethylated CpG sites, which occurs in post-replicative DNA sequences (Bestor & Ingram, 1983; Hermann et al, 2004b; Pradhan et al, 1999; Leonhardt et al, 1992). Whether Dnmt1 also features an additional *de novo* activity *in vivo* is yet unclear and still remains to be addressed. Dnmt1 has low, but detectable activity on unmethylated substrates in *vitro* (Frauer & Leonhardt, 2009). Moreover, overexpression of an Dnmt1

transgene in cancer cells lead to hypermethylation of imprinted promoter regions which might also hint to a possible *de novo* activity of Dnmt1 (Biniszkiewicz et al, 2002). The *de novo* activity of Dnmt1 was shown to be enhanced by proximal methylated CpG sites (Goyal et al, 2006; Tollefsbol & Hutchison, 1997). Notably, those studies were performed using *in vitro* assays and clear evidence for a *de novo* activity of Dnmt1 has yet to be shown in living cells. Up until now, determination of Dnmt1 methylation activity *in vivo* has been performed by rescuing *dnmt1*$^{-/-}$ ES cells and by a subsequent analysis of methylation level changes after re-introducing Dnmt1. However, since the *de novo* methyltransferases Dnmt3a and Dnmt3b are still expressed in those cells, a clear distinction between maintenance and *de novo* activity of Dnmt1 has not been possible so far. Therefore, the generation of a methylation devoid triple knockout ES cell line (TKO) that lacks all three major methyltransferases, Dnmt1, Dnmt3a and Dnmt3b offered a new possibility to study methyltransferase activity *in vivo* (Tsumura et al, 2006). These TKO made it possible to distinguish *de novo* methylation activity from maintenance methylation activity by re-expressing Dnmts and a careful subsequent methylation analysis.

First, we transiently transfected TKO cells with a GFP-Dnmt1 expression construct and sorted for GFP-positive cells after 48 hours (see chapter 2.3). Notably, in none of the analyzed sequences was methylation above background detectable, which lead us to the establishment of stable expressing cell lines to investigate long-term effects and to be able to detect even rare *de novo* methylation events. With our new FACS based approach to generate stable cell lines (see chapter 3.1.1), we were able to generate TKO cell lines that stably express GFP-Dnmt1 and analyze the ability to restore methylation of unmethylated DNA. We established three different stable pools of TKO cells expressing GFP-Dnmt1 and generated single clones that express GFP-Dnmt1 comparable to endogenous *dnmt1* expression in wild type cells. We analyzed samples from different time points, 3 and 6 weeks after initial transfection, to monitor a possible accumulation of methylation over time. The methylation analysis of different stably expressing GFP-Dnmt1 cell lines and the analysis at different time points clearly shows that methylation levels in those rescued TKO cell lines are below the detection limit and indicate no measurable *de novo* activity of Dnmt1 in those cells. Clearly, our data provide strong evidence that Dnmt1 alone is not able to *de novo* methylate unmethylated substrate *in vivo*. However, we only analyzed

selected sequences, thus cannot exclude that methylation occurs at sites, which have been not accessed in our experimental approach. Within this context, it would be important to perform global as well as genome wide methylation analysis.

As mentioned above, it has been shown *in vitro* that Dnmt1 does have a *de novo* activity particularly if proximal CpG sites were already methylated (Goyal et al, 2006; Tollefsbol & Hutchison, 1997). Recent work supports the hypothesis that Dnmt1 is involved in spreading of CpG methylation ((Arand et al) see chapter 2.4). A detailed bisulfite sequencing analysis based on a hairpin linker approach was used to determine methylation levels in *dnmt1⁻ᐟ⁻*, *dnmt3a⁻ᐟ⁻ dnmt3b⁻ᐟ⁻* (DKO), and *np95⁻ᐟ⁻* ES cell lines. Using this approach it was possible to discriminate hemimethylated, fully methylated and unmethylated sites. Interestingly, based on this data set *de novo* activity of Dnmt1 at major satellite and IAP sequences was predicted via a mathematical model (see chapter 2.4). It should be noted that these data were obtained by comparing cell lines that are hypomethylated but not completely devoid of methylation as it is the case in TKO cells, thus are not contradicting with the inability of Dnmt1 to *de novo* methylate unmethylated substrate as we have observed in our rescue experiments outline above. Taken together, our data imply that Dnmt1 *de novo* activity *in vivo* seems to require the presence of pre-existing methylated CpG sites. To address whether pre-existing proximal methylated sites in the TKO cell system initiates *de novo* methylation of Dnmt1, introduction of methylation in specific and clearly determined regions would be necessary to investigate this question *in vivo*.

Finally, it should be noted that Dnmt1 is part of part of a complex epigenetic network where interacting factors exert an important accessory function and therefore in most cases not assessed in *in vitro* assays. However, those interacting factors are essential when determine and analyzing DNA methylation activity. To further investigate whether the interaction of Dnmt1 with Dnmt3a and Dnmt3b might influence or even activate a potential *de novo* activity of Dnmt1, it would be interesting to perform double rescue experiments in TKO cells with catalytic inactive Dnmt3a and Dnmt3b, respectively.

Clearly, our data contradict *in vitro* data predicting *de novo* methylation activity of Dnmt1 on unmethylated DNA. It should be noted that *in vitro* assays normally use naked DNA

oligonucleotides, where nucleosomes and thus chromatin compacting is missing. Additionally, the lack of interacting partner might influence Dnmt1 activity, leading to differences between *in vitro* and *in vivo* data. In TKO cells we were able to analyze the activity of Dnmt1 on DNA substrates within its physiological environment.

3.4   Role of Np95 in methylation-dependent promoter silencing

Np95 is an essential cofactor of DNA methylation maintenance. It is postulated that Np95 binds to hemimethylated sites and recruits Dnmt1 by its interaction with the TS domain of Dnmt1 (Achour et al, 2008; Bostick et al, 2007; Sharif et al, 2007). Dnmt1 and Np95 knockout cell lines show a strikingly similar hypomethylation and developmental arrest (Bostick et al, 2007; Muto et al, 2002) indicating that the lack of Np95 results in a loss of DNA methylation maintenance comparable to a Dnmt1 loss of function. Given the fundamental importance of Np95 in DNA methylation we raised the questions whether Np95 is involved in *de novo* methylation processes. To this aim, we investigated the role of Np95 as well as the methyltransferases Dnmt1, Dnmt3a and Dnmt3b in exogenous and endogenous promoters silencing processes ((Meilinger et al, 2009) see chapters 2.6 and 2.7).

3.4.1   Role of Np95 and DNA methyltransferases in exogenous promoter silencing processes

By co-immunoprecitation experiments we found strong evidence that Np95 interacts with both *de novo* methyltransferases Dnmt3a and Dnmt3b *in vitro* and *in vivo* (Meilinger et al, 2009) see chapter 2.6). Dnmt3a and Dnmt3b play crucial roles in gene silencing by setting new methylation marks in regulatory gene regions and we raised the question whether the interaction with Np95 plays a role in this process.
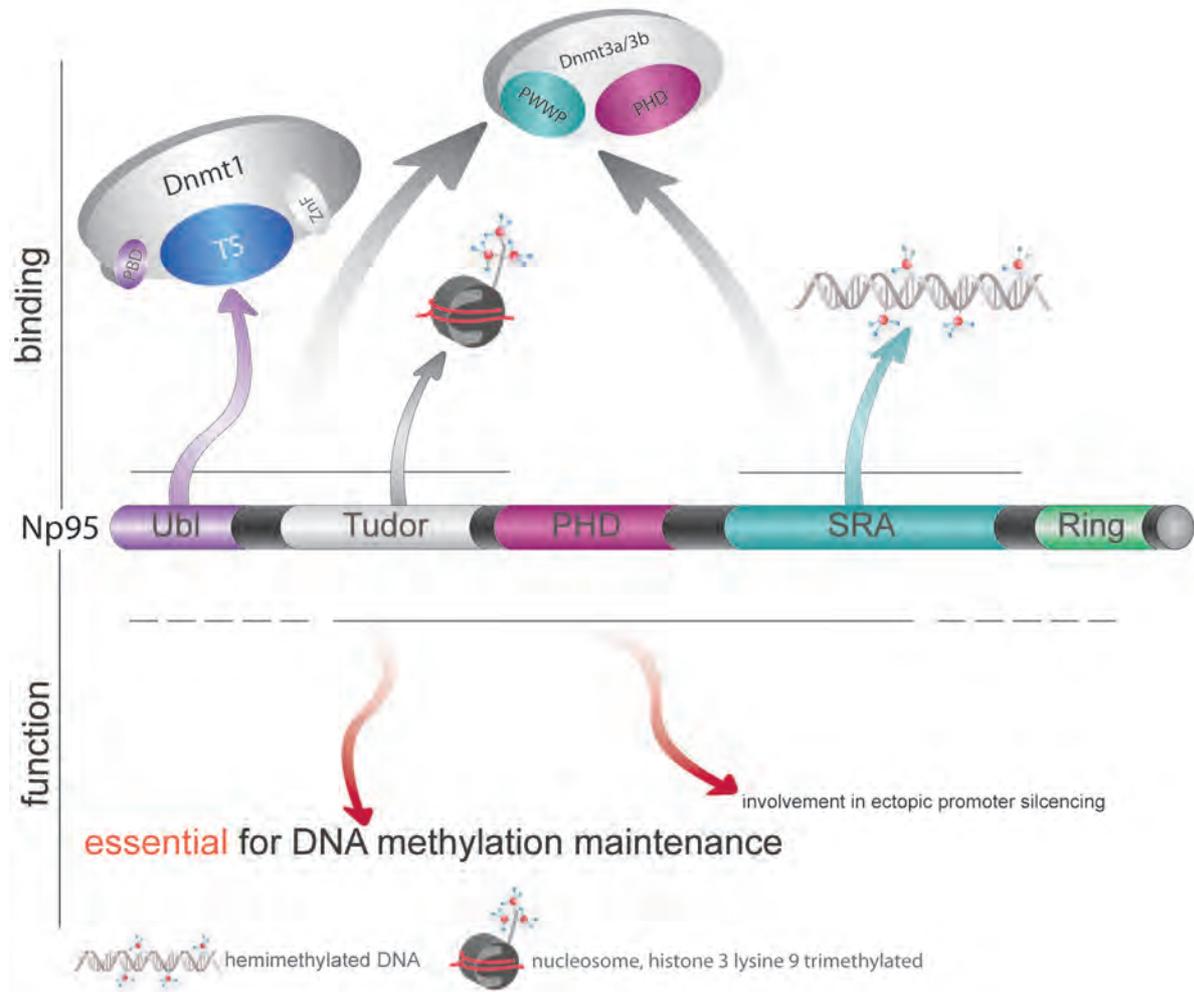
Figure 3-3: Domain overview of Np95 and interacting partners. The Ubl domain of Np95 interacts with the TS domain of Dnmt1, which is an essential interaction for DNA methylation maintenance. The Tudor domain binds H3K9me3 and SRA domain binds to hemimethylated DNA. Ubl, Tudor and SRA domain are involved in the interaction with the *de novo* methyltransferases Dnmt3a/3b. Np95 is involved in ectopic promoter silencing in vivo.

In order to further investigate the functional significance of this interaction we developed a fluorescent reporter-silencing assay in ES cells ((Meilinger et al, 2009); see chapter 2.6). For this approach we utilized the observation that the commonly used cytomegalovirus (CMV) promoter for expression in mammalian cells is rapidly silenced in transient transfections of mouse ES cells. In contrast, the chimeric CMV early enhancer/chicken b-actin (CAG) promoter remains stably active in wild type ES cells and thus is often used to generate stable cell lines. Accordingly, ES cells were simultaneously transfected with a red fluorescent protein (RFP) construct driven by the CMV promoter and a GFP construct driven by the CAG promoter. The fluorescent signal of both fluorescent markers was monitored up to 10 days and acquisition was performed using an automated imaging

Discussion

system allowing a medium-throughput approach. With this assay we analyzed the decrease of the CMV driven fluorescence in comparison to the CAG driven GFP decrease in various cell lines lacking proteins that are involved in DNA methylation, as well as histone modifications. The fluorescence of the CMV promoter steadily decreased in wild type cells whereas *dnmt1*[-/-] ES cells showed a partly decreased silencing effect. Surprisingly, there was no decrease detectable in *np95*[-/-], *dnmt3a*[-/-]/*dnmt3b*[-/-] (DKO) and TKO cells, arguing for an impaired silencing ability in those cell lines. Interesting was the impaired silencing effect observed in *np95*[-/-] ES cells, especially considering that those cells express all three active methyltransferases. We argue that the selective decrease of the CMV promoter driven fluorescence reflects a promoter silencing effect and we further explored the role of DNA methylation in this process by pyrosequencing both CAG and CMV promoter regions. Interestingly, we found that methylation was only detectable at the end point of our analysis in the CMV promoter region in wild type cells, suggesting that methylation plays no distinct role in the initiation of the observed silencing effect. Our data imply that the *de novo* methyltransferases Dnmt3a/3b and Np95 are involved in silencing of the CMV promoter.

The dispensability/independence of DNA methylation in the initial silencing mechanism leads us to investigate the involvement of repressive histone marks and histone modifying proteins. For this purpose we analyzed knockout cell lines that are devoid of the H3K9 methyltransferases G9a or Suv39h1/2, respectively. Notably, these histone methyltransferases are known to interact with Np95 as well as Dnmt1, Dnmt3a and Dnmt3b (Epsztejn-Litman et al, 2008; Kim et al, 2009). Furthermore, interaction of G9a with the *de novo* methyltransferase mediates *de novo* methylation of *oct4* and *nanog* promoters during differentiation (Epsztejn-Litman et al, 2008; Feldman et al, 2006). Data derived from our silencing assay indicated that G9a is required for silencing of the CMV promoter similar to a similar extent as Np95, Dnmt3a and Dnmt3b. Suv39h1/2 seems to be only partially involved. With regard to previous findings that plasmids are chromatinized upon transfection in mammalian cells (Mehta et al, 2009; Riu et al, 2007), it was expected to find histone modifying enzymes involved in silencing of the CMV promoter Whether a silenced CMV promoter exhibits heterochromatin marks such as H3K9me similar to native

marks observed in mammalian cells should be addressed by performing a Chromatin Immuno-precipitation (ChIP) assay.

Our data clearly show that Np95 not only interacts with Dnmt1, but also with Dnm3a and Dnmt3b. Furthermore, we were able to demonstrate an involvement of Np95, Dnmt3a Dnmt3b and G9a in a silencing process that is initially independent of DNA methylation. Surprisingly, Dnmt1 is only partially involved in this process. With these observations we provide an additional co-regulatory role of Np95 with the *de novo* methyltransferases Dnmt3a and Dnmt3b and therefore provide deeper insight in the complex network of epigenetic silencing processes.

3.4.2   Role of Np95 and DNA methyltransferases in endogenous promoter silencing

Although our silencing assay provides a powerful tool to screen factors involved in silencing of ectopically expressed promoters, it is imperative to investigate the role of those factors on endogenous promoter silencing. It is of profound interest whether exogenous and endogenous promoter silencing share a common pathway. To this aim, we wanted to investigate the role of DNA methyltransferases as well as Np95 in silencing of pluripotency genes o*ct4* and *nanog* and the involvement of DNA methylation during differentiation (see chapter 2.7). *Oct4* and *nanog* are pluripotency genes that undergo silencing and methylation upon differentiation (Li et al, 2007). We differentiated *np95$^{-/-}$,dnmt1$^{-/-}$* as well as their respective wild type cells into embryoid bodies (EBs) and analyzed changes of the methylation levels of *oct4* and *nanog* promoters. These results were further compared to the transcript levels of *oct4* and *nanog* during differentiation.

Consistent with our previous findings on CMV promoter silencing, down regulation of *oct4* and *nanog* transcripts precedes promoter methylation. Interestingly, methylation levels of *oct4* and *nanog* promoters were similar in *np95$^{-/-}$* and *dnmt1$^{-/-}$* knockout cell lines and reached a plateau around 6 days after initiation of differentiation. This methylation is most likely carried out by the *de novo* methyltransferases Dnmt3a and Dnmt3b since they are both present in the analyzed cells and a synergistic function of Dnmt3a and Dnmt3b in methylation of *oct4* and *nanog* promoters has been reported previously (Li et al., 2007). However, the similar observed initial methylation increase in *np95$^{-/-}$* and *dnmt1$^{-/-}$* knockout cell lines argues that *de novo* methylation catalyzed by Dnmt3a and Dnmt3b is not

impaired in *np95*$^{-/-}$ES cells. These findings do not support our hypothesis that interaction of Np95 with the *de novo* methyltransferases Dnmt3a/3b is involved in silencing of endogenous promoters. To test whether *de novo* activity of Dnmt3a and Dnmt3b is not or only slightly impaired in the absence of Np95, one could establish a stable cell line expressing a Np95 construct where interaction with Dnmt3a and Dnmt3b is disrupted, followed by an analysis of methylation levels during differentiation.

Furthermore, methylation levels by far do not reach those observed in wild type cells and more importantly a slight but notable decrease in methylation levels was observed in *dnmt1*$^{-/-}$ ES cells most likely due to the lack of DNA methylation maintenance. Surprisingly, no such decrease was observed in *np95*$^{-/-}$ ES cells, even though the catalytic activity of Dnmt1 is highly controlled by its interaction with Np95. These findings could argue for a residual methylation activity of Dnmt1 independent of Np95, or for the concerted action of the *de novo* methyltransferases Dnmt3a and Dnmt3b in maintaining DNA methylation maintenance. Another possibility explaining this residual methylation would be a functional redundancy of Np97 to Np95 in maintaining DNA methylation. Np97 is a recently discovered protein that comprises similar domain structures as Np95. In addition we have shown recently that Np97 is upregulated during differentiation (Pichler et.al. 2011; see chapter 2.5), which could hint to a compensating role Np97.

Taken together, our data show that initial silencing of the pluripotency genes *oct4* and *nanog* does not depend on the involvement of Np95, Dnmt1 and Dnmt3a/b and argue for a different silencing pathway than we have observed in CMV promoter silencing. Similar to the involvement of histone methyltransferases in silencing of the CMV promoter, histone modification and histone modifying enzymes are also involved in silencing of endogenous promoters (Myant et al, 2011; Feldman et al, 2006). It was shown that G9a mediated methylation of H3K9 and silencing of the *oct4* promoter during retinoic acid treatment precedes methylation of the promoter region (Feldman et al, 2006). These data imply a model where histone modifications like H3K9me3 are involved in an initial silencing processes and histone modifications as well as interaction between histone modifying enzymes and DNA methyltransferases are essential for long term silencing processes.

3.5 Differentiation of ES cells lacking all three major methyltransferases and silencing of pluripotency genes

With the CMV promoter silencing assay we showed that silencing of the retroviral CMV promoter was impaired in TKO cells (see chapter 2.6). When differentiating TKO cells into embryoid bodies (EBs), we observed that the EBs were smaller when compared to wild type EBs. Interestingly, we did not observe a change in apoptosis rate, which has been suggested to increase during differentiation of hypomethylated cells (Lei et al, 1996; Panning & Jaenisch, 1996; Jackson et al, 2004). To test whether the differentiation potential of TKO is impaired in comparison to wild type ES cells, we analyzed the expression of pluripotency genes *oct4* and *nanog*, which should decrease with the level of differentiation. During differentiation of TKO cells into EBs *oct4* and *nanog* mRNA levels decline during the first 4 days of differentiation in a similar timely manner as during wild type differentiation. This was a remarkable observation considering that TKO cells have no active methyltransferases, but are still able to down regulate pluripotency genes *in vivo*, indicating that no active DNA methyltransferases are involved in this initial silencing process. Our findings suggest that TKO cells have differentiation potential, yet it has to be determined to which extent. In accordance to our findings, recent studies showed that TKO cells are able to develop into embryonic and extraembryonic cell types (Sakaue et al, 2010). It would be of profound interest to characterize TKO EB's and determine their differentiation potential by analyzing genome wide expression profiles as well as cell type and lineage specific marker proteins, which is in work progress.

However, mRNA levels were not completely down regulated as observed in wild type cells and stagnate at a residual level of around 20%, arguing for an important role of methylation at a later stage in the silencing process and suggest the requirement of DNA methylation for complete silencing.
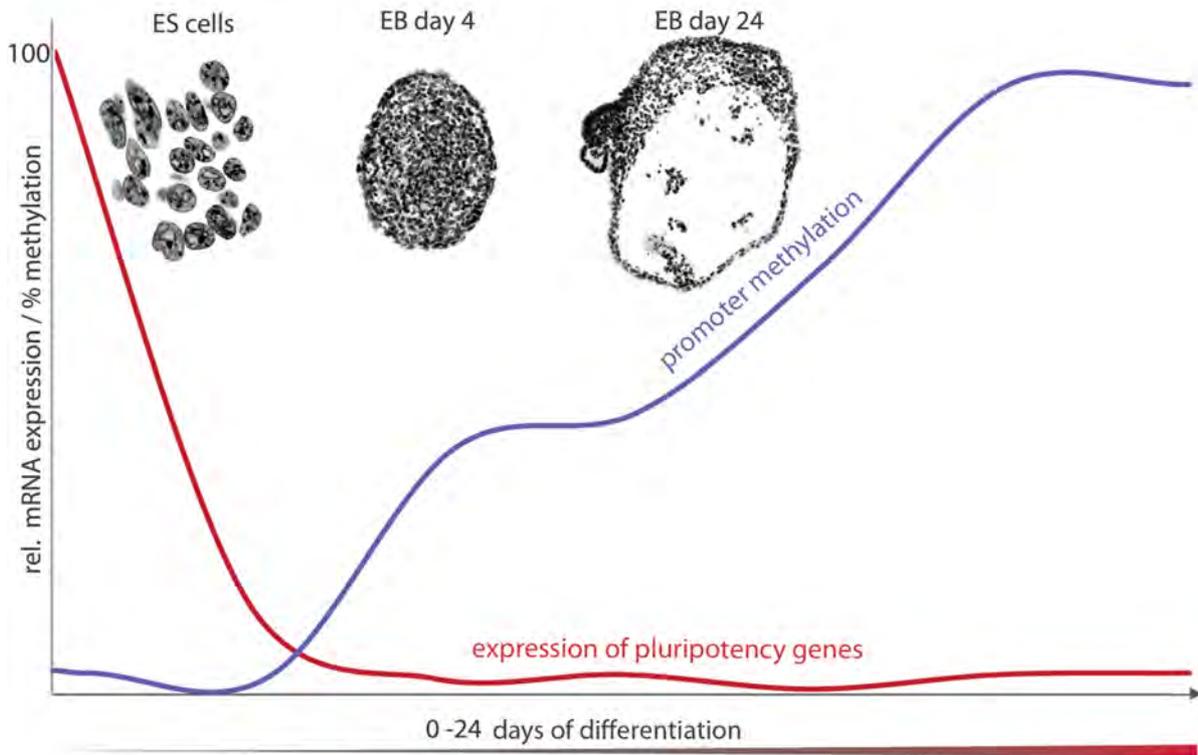
Figure 3-4: Schematic representation of relative expression level changes and quantitative methylation level changes during EB differentiation. ES cells and subsequent EB formation is shown with representative images in the upper panel at day 0, day 4 and day 20 of differentiation. ES cells and EBs are stained with DAPI and kindly provided by Irina Solovei. The expression levels of pluripotency genes decline fast after onset of differentiation (blue line). In contrast, methylation levels begin to increase with a delay at day 4 of differentiation (red line).

Our data provide strong evidence that methylation and DNA methyltransferases are not involved in the initial process of differentiation and that down regulation of the pluripotency genes *oct4* and *nanog* do not primarily depend on the contribution of any known active methyltransferases. In fact, DNA methylation is a process that occurs later in the differentiation process and might be involved in long term silencing. We are currently working on the question whether cells that are impaired in DNA methyltransferase activity are able to reverse the differentiation process. To address this question, we generated stable wild type and TKO cell lines that express a fluorescent marker (GFP) under the endogenous *oct4* promoter (see chapter 2.8). Having in hand a fluorescent read out, we are able to follow the differentiation and de-differentiation process by FACS as well as by fluorescence microscopy. Preliminary data show that a recovery of the GFP signal by the addition of LIF and therefore subsequent initiation of a de-differentiation process is evident

only in TKO but not wild type cells. These results would support the hypothesis that methylation is involved in a long-term silencing pathway.

In conclusion we investigated the role of the Dnmt1 subdomains PBD and the CXXC zinc finger motif in faithful methylation maintenance and our data suggest that both domains seem dispensable for enzymatic activity of Dnmt1 and maintenance of DNA methylation in living cells. However, the PBD domain enhances DNA methylation efficiency by two-fold. Moreover, we provide data regarding the involvement of Dnmt1 in *de novo* methylation, which was controversially discussed previously and we clearly showed that Dnmt1 is not able to *de novo* methylate unmethylated DNA sequences in living cells. Given the fact the Np95 is such a crucial factor in DNA methylation maintenance and links DNA methylation with histone modifications, we investigated its role in *de novo* methylation and silencing of promoter sequences. We found strong evidence that Np95, together with the *de novo* methyltransferases Dnm3a/Dnmt3b and the histone methyltransferase G9a is involved in this silencing pathway. The above findings initiated a detailed methylation and expression analysis of pluripotency factors during differentiation. We obtained clear data supporting the hypothesis that methylation is dispensable for initiation of differentiation but suggests that methylation is required for stable and long term gene silencing.

Discussion

4.   [Annex]

4.1   References

Aapola U, Lyle R, Krohn K, Antonarakis SE & Peterson P (2001) Isolation and initial characterization of the mouse Dnmt3l gene. *Cytogenet Cell Genet* **92:** 122–126

Aapola U, Maenpaa K, Kaipia A & Peterson P (2004) Epigenetic modifications affect Dnmt3L expression. *Biochem J* **380:** 705–713

Achour M, Jacq X, Rondé P, Alhosin M, Charlot C, Chataigneau T, Jeanblanc M, Macaluso M, Giordano A, Hughes AD, Schini-Kerth VB & Bronner C (2008) The interaction of the SRA domain of ICBP90 with a novel domain of DNMT1 is involved in the regulation of VEGF gene expression. *Oncogene* **27:** 2187–2197

Amir RE, den Veyver Van IB, Wan M, Tran CQ, Francke U & Zoghbi HY (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* **23:** 185–188

Antequera F (2003) Structure, function and evolution of CpG island promoters. *Cell. Mol. Life Sci.* **60:** 1647–1658

Antequera F & Bird A (1993) CpG islands. *EXS* **64:** 169–185

Aoki K, Meng G, Suzuki K, Takashi T, Kameoka Y, Nakahara K, Ishida R & Kasai M (1998) RP58 associates with condensed chromatin and mediates a sequence-specific transcriptional repression. *J Biol Chem* **273:** 26698–26704

Arand J, Spieler D, Karius T, Meilinger D, Meissner A, Okano M, Jenuwein T, Xu G, Leonhardt H, Wolf V & Walter J *DNA-methylation of CpG dyads and markov modeling of DNA-methylation control in mammals*

Argentaro A, Yang J-C, Chapman L, Kowalczyk MS, Gibbons RJ, Higgs DR, Neuhaus D & Rhodes D (2007) Structural consequences of disease-causing mutations in the ATRX-DNMT3-DNMT3L (ADD) domain of the chromatin-associated protein ATRX. *Proc Natl Acad Sci USA* **104:** 11939–11944

Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N & Lovell-Badge R (2003) Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* **17:** 126–140

Babaie Y, Herwig R, Greber B, Brink TC, Wruck W, Groth D, Lehrach H, Burdon T & Adjaye J (2007) Analysis of Oct4-dependent transcriptional networks regulating self-renewal and pluripotency in human embryonic stem cells. *Stem Cells* **25:** 500–510

Bachman KE, Rountree MR & Baylin SB (2001) Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. *J Biol Chem* **276:** 32282–32287

Basehore MJ & Friez MJ (2009) Molecular analysis of Fragile X syndrome. *Curr Protoc Hum Genet* **Chapter 9:** Unit 9.5

Beard C, Li E & Jaenisch R (1995) Loss of methylation activates Xist in somatic but not in embryonic cells. *Genes Dev* **9:** 2325–2334

Bestor TH & Ingram VM (1983) Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proc Natl Acad Sci USA* **80:** 5559–5563

Bibel M, Richter J, Schrenk K, Tucker KL, Staiger V, Korte M, Goetz M & Barde Y-A (2004) Differentiation of mouse embryonic stem cells into a defined neuronal lineage. *Nat. Neurosci.* **7:** 1003–1009

Biniszkiewicz D, Gribnau J, Ramsahoye B, Gaudet F, Eggan K, Humpherys D, Mastrangelo M-A, Jun Z, Walter J & Jaenisch R (2002) Dnmt1 overexpression causes genomic hypermethylation, loss of imprinting, and embryonic lethality. *Molecular and Cellular Biology* **22:** 2124–2135

Bird AP (1986) CpG-rich islands and the function of DNA methylation. *Nature* **321:** 209–213

Bostick M, Kim JK, Estève P-O, Clark A, Pradhan S & Jacobsen SE (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317:** 1760–1764

Bourc'his D, Xu GL, Lin CS, Bollman B & Bestor TH (2001) Dnmt3L and the establishment of maternal genomic imprints. *Science* **294:** 2536–2539

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R & Young RA (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122:** 947–956

Cassidy SB & Schwartz S (1998) Prader-Willi and Angelman Syndromes: Disorders of Genomic Imprinting. *Medicine (Baltimore)* **77 N2 -:** 140–151

Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S & Smith A (2003) Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113:** 643–655

Chen T, Ueda Y, Dodge JE, Wang Z & Li E (2003) Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Molecular and Cellular Biology* **23:** 5594–5605

Chen T, Tsujimoto N & Li E (2004) The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin. *Molecular and Cellular Biology* **24:** 9048–9058

Chevet E, Lemaître G & Katinka MD (1995) Low concentrations of tetramethylammonium

chloride increase yield and specificity of PCR. *Nucleic Acids Res* **23:** 3343–3344

Chuang LS, Ian HI, Koh TW, Ng HH, Xu G & Li BF (1997) Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science* **277:** 1996–2000

Clark SJ, Harrison J, Paul CL & Frommer M (1994) High sensitivity mapping of methylated cytosines. *Nucleic Acids Res* **22:** 2990–2997

Colella S, Shen L, Baggerly KA, Issa JP & Krahe R (2003) Sensitive and quantitative universal Pyrosequencing methylation analysis of CpG sites. *Biotech.* **35:** 146–150

Coulondre C, Miller JH, Farabaugh PJ & Gilbert W (1978) Molecular basis of base substitution hotspots in Escherichia coli. *Nature* **274:** 775–780

Cowell IG, Aucott R, Mahadevaiah SK, Burgoyne PS, Huskisson N, Bongiorni S, Prantera G, Fanti L, Pimpinelli S, Wu R, Gilbert DM, Shi W, Fundele R, Morrison H, Jeppesen P & Singh PB (2002) Heterochromatin, HP1 and methylation at lysine 9 of histone H3 in animals. *Chromosoma* **111:** 22–36

Dodge JE, Okano M, Dick F, Tsujimoto N, Chen T, Wang S, Ueda Y, Dyson N & Li E (2005) Inactivation of Dnmt3b in mouse embryonic fibroblasts results in DNA hypomethylation, chromosomal instability, and spontaneous immortalization. *J Biol Chem* **280:** 17986–17991

Doetschman TC, Eistetter H, Katz M, Schmidt W & Kemler R (1985) The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J Embryol Exp Morphol* **87:** 27–45

Dong KB, Maksakova IA, Mohn F, Leung D, Appanah R, Lee S, Yang HW, Lam LL, Mager DL, Schuebeler D, Tachibana M, Shinkai Y & Lorincz MC (2008) DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. *EMBO J* **27:** 2691–2701

Easwaran HP, Schermelleh L, Leonhardt H & Cardoso MC (2004) Replication-independent chromatin loading of Dnmt1 during G2 and M phases. *EMBO reports* **5:** 1181–1186

Epsztejn-Litman S, Feldman N, Abu-Remaileh M, Shufaro Y, Gerson A, Ueda J, Deplus R, Fuks F, Shinkai Y, Cedar H & Bergman Y (2008) De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat Struct Mol Biol* **15:** 1176–1183

Fatemi M, Hermann A, Pradhan S & Jeltsch A (2001) The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA. *Journal of Molecular Biology* **309:** 1189–1199

Fatemi M, Hermann A, Gowher H & Jeltsch A (2002) Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. *Eur J Biochem* **269:** 4981–4984

Feinberg AP & Tycko B (2004) The history of cancer epigenetics. *Nat. Rev. Cancer* **4:** 143–153

Feldman N, Gerson A, Fang J, Li E, Zhang Y, Shinkai Y, Cedar H & Bergman Y (2006) G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol* **8:** 188–194

Fellinger K (2009) Analysis of Protein Interactions Controlling DNA Methyltransferases. *Dissertation* 1–150

Fellinger K, Rothbauer U, Felle M, Längst G & Leonhardt H (2009) Dimerization of DNA methyltransferase 1 is mediated by its regulatory domain. *J Cell Biochem* **106:** 521–528

Fraga MF, Rodríguez R & Cañal MJ (2000) Rapid quantification of DNA methylation by high performance capillary electrophoresis. *Electrophoresis* **21:** 2990–2994

Frauer C & Leonhardt H (2009) A versatile non-radioactive assay for DNA methyltransferase activity and DNA binding. *Nucleic Acids Res* **37:** e22

Frauer C, Rottach A, Meilinger D, Bultmann S, Fellinger K, Hasenöder S, Wang M, Qin W, Söding J, Spada F & Leonhardt H (2011) Different binding properties and function of CXXC zinc finger domains in Dnmt1 and Tet1. *PLoS ONE* **6:** e16627

Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL & Paul CL (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci USA* **89:** 1827–1831

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

Gaudet F (2003) Induction of Tumors in Mice by Genomic Hypomethylation. *Science* **300:** 489–492

Goll MG & Bestor TH (2005) Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* **74:** 481–514

Goll MG, Kirpekar F, Maggert KA, Yoder JA, Hsieh C-L, Zhang X, Golic KG, Jacobsen SE & Bestor TH (2006) Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science* **311:** 395–398

Gonzalez-Quevedo R, Garcia-Aranda C, Moran A, De Juan C, Sanchez-Pernaute A, Torres A, Diaz-Rubio E, Balibrea J, Benito M & Iniesta P (2004) Differential impact of p16 inactivation by promoter methylation in non-small cell lung and colorectal cancer: Clinical implications. *Int J Oncol* **24:** 349–355

Gowher H, Stockdale CJ, Goyal R, Ferreira H, Owen-Hughes T & Jeltsch A (2005) De

novo methylation of nucleosomal DNA by the mammalian Dnmt1 and Dnmt3A DNA methyltransferases. *Biochemistry* **44:** 9899–9904

Goyal R, Reinhardt R & Jeltsch A (2006) Accuracy of DNA methylation pattern preservation by the Dnmt1 methyltransferase. *Nucleic Acids Res* **34:** 1182–1188

Hall J, Guo G, Wray J, Eyres I, Nichols J, Grotewold L, Morfopoulou S, Humphreys P, Mansfield W, Walker R, Tomlinson S & Smith A (2009) Oct4 and LIF/Stat3 Additively Induce Kruppel Factors to Sustain Embryonic Stem Cell Self-Renewal. *Cell Stem Cell* **5:** 597–609

Hansen RS, Wijmenga C, Luo P, Stanek AM, Canfield TK, Weemaes CM & Gartler SM (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 & Li E (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129:** 1983–1993

Hermann A, Schmitt S & Jeltsch A (2003) The human Dnmt2 has residual DNA-(cytosine-C5) methyltransferase activity. *J Biol Chem* **278:** 31717–31721

Hermann A, Gowher H & Jeltsch A (2004a) Biochemistry and biology of mammalian DNA methyltransferases. *Cell. Mol. Life Sci.* **61:** 2571–2587

Hermann A, Goyal R & Jeltsch A (2004b) The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J Biol Chem* **279:** 48350–48359

Hu Y-G, Hirasawa R, Hu J-L, Hata K, Li C-L, Jin Y, Chen T, Li E, Rigolet M, Viegas-Péquignot E, Sasaki H & Xu G-L (2008) Regulation of DNA methylation activity through Dnmt3L promoter methylation by Dnmt3 enzymes in embryonic development. *Hum Mol Genet* **17:** 2654–2664

Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR & Rao A (2010) The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS ONE* **5:** e8888–

Jackson M, Krassowska A, Gilbert N, Chevassut T, Forrester L, Ansell J & Ramsahoye B (2004) Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. *Molecular and Cellular Biology* **24:** 8862–8871

Jarvis CD, Geiman T, Vila-Storm MP, Osipovich O, Akella U, Candeias S, Nathan I, Durum SK & Muegge K (1996) A novel putative helicase produced in early murine lymphocytes. *Gene* **169:** 203–207

Keller GM (1995) In vitro differentiation of embryonic stem cells. *Curr. Opin. Cell Biol.* **7:** 862–869

Kim JK, Estève P-O, Jacobsen SE & Pradhan S (2009) UHRF1 binds G9a and participates in p21 transcriptional regulation in mammalian cells. *Nucleic Acids Res* **37:** 493–505

Kriaucionis S & Heintz N (2009) The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. *Science* **324:** 929–930

Kuo KC, McCune RA, Gehrke CW, Midgett R & Ehrlich M (1980) Quantitative reversed-phase high performance liquid chromatographic determination of major and modified deoxyribonucleosides in DNA. *Nucleic Acids Res* **8:** 4763–4776

Lei H, Oh SP, Okano M, Jüttermann R, Goss KA, Jaenisch R & Li E (1996) De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122:** 3195–3205

Leonhardt H, Page AW, Weier HU & Bestor TH (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* **71:** 865–873

Li E, Bestor TH & Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69:** 915–926

Li E, Beard C & Jaenisch R (1993) Role for DNA methylation in genomic imprinting. *Nature* **366:** 362–365

Li J-Y, Pu M-T, Hirasawa R, Li B-Z, Huang Y-N, Zeng R, Jing N-H, Chen T, Li E, Sasaki H & Xu G-L (2007) Synergistic function of DNA methyltransferases Dnmt3a and Dnmt3b in the methylation of Oct4 and Nanog. *Molecular and Cellular Biology* **27:** 8748–8759

Liang G, Chan MF, Tomigahara Y, Tsai YC, Gonzales FA, Li E, Laird PW & Jones PA (2002) Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Molecular and Cellular Biology* **22:** 480–491

Litt M, Simpson M, Gaszner M, Allis C & Felsenfeld G (2001) Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* **293:** 2453–2455

Margot JB, Aguirre-Arteta AM, Di Giacco BV, Pradhan S, Roberts RJ, Cardoso MC & Leonhardt H (2000) Structure and function of the mouse DNA methyltransferase gene: Dnmt1 shows a tripartite structure. *Journal of Molecular Biology* **297:** 293–300

Martin GR (1981) Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci USA* **78:** 7634–7638

Mehta AK, Majumdar SS, Alam P, Gulati N & Brahmachari V (2009) Epigenetic regulation of cytomegalovirus major immediate-early promoter activity in transgenic mice. *Gene* **428:** 20–24

Meilinger D, Fellinger K, Bultmann S, Rothbauer U, Bonapace IM, Klinkert WEF, Spada F & Leonhardt H (2009) Np95 interacts with de novo DNA methyltransferases, Dnmt3a

and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells. *EMBO reports* **10:** 1259–1264

Morgan HD, Santos F, Green K, Dean W & Reik W (2005) Epigenetic reprogramming in mammals. *Hum Mol Genet* **14 Spec No 1:** R47–58

Mortusewicz O, Schermelleh L, Walter J, Cardoso MC & Leonhardt H (2005) Recruitment of DNA methyltransferase I to DNA repair sites. *Proc Natl Acad Sci USA* **102:** 8905–8909

Muto M, Kanari Y, Kubo E, Takabe T, Kurihara T, Fujimori A & Tatsumi K (2002) Targeted disruption of Np95 gene renders murine embryonic stem cells hypersensitive to DNA damaging agents and DNA replication blocks. *J Biol Chem* **277:** 34549–34555

Myant K & Stancheva I (2008) LSH cooperates with DNA Methyltransferases to repress transcription. *Molecular and Cellular Biology* **28:** 215–226

Myant K, Termanis A, Sundaram AYM, Boe T, Li C, Merusi C, Burrage J, Las Heras de JI & Stancheva I (2011) LSH and G9a/GLP complex are required for developmentally programmed DNA methylation. *Genome Res* **21:** 83–94

Nan X & Bird A (2001) The biological functions of the methyl-CpG-binding protein MeCP2 and its implication in Rett syndrome. *Brain Dev.* **23 Suppl 1:** S32–7

Nan X, Tate P, Li E & Bird A (1996) DNA methylation specifies chromosomal localization of MeCP2. *Molecular and Cellular Biology* **16:** 414–421

Niwa H, Ogawa K, Shimosato D & Adachi K (2009) A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature* **460:** 118–122

Okano M, Xie S & Li E (1998) Cloning and characterization of a family of novel mammalian DNA (cytosine-5)

Okano M, Bell DW, Haber DA & Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99:** 247–257

Okuwaki M & Verreault A (2004) Maintenance DNA methylation of nucleosome core particles. *J Biol Chem* **279:** 2904–2912

Panning B & Jaenisch R (1996) DNA hypomethylation can activate Xist expression and silence X-linked genes. *Genes Dev* **10:** 1991–2002

Pradhan M, Estève P-O, Chin HG, Samaranayke M, Kim G-D & Pradhan S (2008) CXXC domain of human DNMT1 is essential for enzymatic activity. *Biochemistry* **47:** 10000–10009

Pradhan S, Bacolla A, Wells RD & Roberts RJ (1999) Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. *J Biol Chem* **274:** 33002–33010

Ratnam S, Mertineit C, Ding F, Howell CY, Clarke HJ, Bestor TH, Chaillet JR & Trasler JM (2002) Dynamics of Dnmt1 methyltransferase expression and intracellular localization during oogenesis and preimplantation development. *Developmental Biology* **245:** 304–314

Rea S, Eisenhaber F, O'Carroll N, Strahl B, Sun Z, Schmid M, Opravil S, Mechtler K, Ponting C, Allis C & Jenuwein T (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406:** 593–599

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

Riu E, Chen Z-Y, Xu H, He C-Y & Kay MA (2007) Histone modifications are associated with the persistence or silencing of vector-mediated transgene expression in vivo. *Mol. Ther.* **15:** 1348–1355

Robertson K, Uzvolgyi E, Liang G, Talmadge C, Sumegi J, Gonzales F & Jones P (1999) The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. *Nucleic Acids Res* **27:** 2291–2298

Robertson KD, Ait-Si-Ali S, Yokochi T, Wade PA, Jones PL & Wolffe AP (2000) DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nat Genet* **25:** 338–342

Ronaghi M (2001) Pyrosequencing sheds light on DNA sequencing. *Genome Res* **11:** 3–11

Rothbauer U, Zolghadr K, Muyldermans S, Schepers A, Cardoso MC & Leonhardt H (2008) A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. *Mol. Cell Proteomics* **7:** 282–289

Sakaue M, Ohta H, Kumaki Y, Oda M, Sakaide Y, Matsuoka C, Yamagiwa A, Niwa H, Wakayama T & Okano M (2010) DNA Methylation Is Dispensable for the Growth and Survival of the Extraembryonic Lineages. *Curr Biol* **20:** 1452–1457

Schermelleh L, Haemmer A, Spada F, Rösing N, Meilinger D, Rothbauer U, Cardoso MC & Leonhardt H (2007) Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. *Nucleic Acids Res* **35:** 4301–4312

Sharif J, Muto M, Takebayashi S-I, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-Koseki Y, Toyoda T, Okamura K, Tajima S, Mitsuya K, Okano M & Koseki H (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450:** 908–912

Song J, Rechkoblit O, Bestor TH & Patel DJ (2011) Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation. *Science* **331:** 1036–1040

Stöger R, Kajimura TM, Brown WT & Laird CD (1997) Epigenetic variation illustrated by DNA methylation patterns of the fragile-X gene FMR1. *Hum Mol Genet* **6:** 1791–1801

Syeda F, Fagan RL, Wean M, Avvakumov GV, Walker JR, Xue S, Dhe-Paganon S & Brenner C (2011) The RFTS domain is a DNA-competitive inhibitor of Dnmt1. *J Biol Chem*

Tachibana M, Matsumura Y, Fukuda M, Kimura H & Shinkai Y (2008) G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *EMBO J* **27:** 2681–2690

Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L & Rao A (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324:** 930–935

Takeshita K, Suetake I, Yamashita E, Suga M, Narita H, Nakagawa A & Tajima S (2011) Structural insight into maintenance methylation by mouse DNA methyltransferase 1 (Dnmt1). *Proc Natl Acad Sci USA* **108:** 9055–9059

Taketo MM (2004) Shutting down Wnt signal-activated cancer. *Nat Genet* **36:** 320–322

Tatematsu KI, Yamazaki T & 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

Tollefsbol TO & Hutchison CA (1997) Control of methylation spreading in synthetic DNA sequences by the murine DNA methyltransferase. *Journal of Molecular Biology* **269:** 494–504

Tost J & Gut IG (2007) DNA methylation analysis by pyrosequencing. *UNKNOWN* **2:** 2265–2275

Tost J, Dunker J & Gut IG (2003) Analysis and quantification of multiple methylation variable positions in CpG islands by Pyrosequencing. *Biotech.* **35:** 152–156

Tsumura A, Hayakawa T, Kumaki Y, Takebayashi S-I, Sakaue M, Matsuoka C, Shimotohno K, Ishikawa F, Li E, Ueda HR, Nakayama J-I & Okano M (2006) Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* **11:** 805–814

Tucker KL, Beard C, Dausmann J, Jackson-Grusby L, Laird PW, Lei H, Li E & Jaenisch R (1996) Germ-line passage is required for establishment of methylation and expression patterns of imprinted but not of nonimprinted genes. *Genes Dev* **10:** 1008–1020

Viré E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, van Eynde A, Bernard D, Vanderwinden J-M, Bollen M, Esteller M, Di Croce L, de Launoit Y & Fuks F (2006) The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439:** 871–874

Voss K, Roos K, Nonay R & Dovichi N (1998) Combating PCR bias in bisulfite-based cytosine methylation analysis. Betaine-modified cytosine deamination PCR. *Anal Chem* **70:** 3818–3823

Wang J, Hevi S, Kurash JK, Lei H, Gay F, Bajko J, Su H, Sun W, Chang H, Xu G, Gaudet F, Li E & Chen T (2009) The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nat Genet* **41:** 125–129

Warnecke PM & Clark SJ (1999) DNA methylation profile of the mouse skeletal alpha-actin promoter during development and differentiation. *Molecular and Cellular Biology* **19:** 164–172

Warnecke PM, Stirzaker C, Melki JR, Millar DS, Paul CL & Clark SJ (1997) Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA. *Nucleic Acids Res* **25:** 4422–4426

Xiong Z & Laird PW (1997) COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res* **25:** 2532–2534

Xu GL, Bestor TH, Bourc'his D, Hsieh CL, Tommerup N, Bugge M, Hulten M, Qu X, Russo JJ & Viegas-Péquignot E (1999) Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402:** 187–191

Yamada Y & Jackson-Grusby L (2005) Opposing effects of DNA hypomethylation on intestinal and liver carcinogenesis. In *Proceedings of the …*

Yoo CB & Jones PA (2006) Epigenetic therapy of cancer: past, present and future. *Nat Rev Drug Discov* **5:** 37–50

Zeschnigk M, Schmitz B, Dittrich B, Buiting K, Horsthemke B & Doerfler W (1997) Imprinted segments in the human genome: different DNA methylation patterns in the Prader-Willi/Angelman syndrome region as determined by the genomic sequencing method. *Hum Mol Genet* **6:** 387–395

Zhu H, Geiman TM, Xi S, Jiang Q, Schmidtmann A, Chen T, Li E & Muegge K (2006) Lsh is involved in de novo methylation of DNA. *EMBO J* **25:** 335–345

Zimmermann C, Guhl E & Graessmann A (1997) Mouse DNA methyltransferase (MTase) deletion mutants that retain the catalytic domain display neither de novo nor maintenance methylation activity in vivo. *Biol Chem* **378:** 393–405

4.2   Abbreviations

5-aza-dC	5-aza-2'-deoxycytidine
5hmC	5 hydroxymethyl-cytosine
5mC	5 methyl-cytosine
aa	amino acids
amp	ampicilline
BAH	bromo adjacent homology domain
CAG	chicken beta actin
CMV	cytomegalie virus
COBRA	combined bisulfite restriction analysis
CpG	cytosine-phosphatidyl-guanine
DAPI	4',6-diamidino-2'-phenylindole dihydrochloride
DKO	dnmt3a-/-3b-/- double knockout ESCs
DNA	desoxyribonucleid acid
Dnmt	DNA methyltransferase
EB	embryoid body
ES	embryonic stemcells
ESCs	ES cells
F2H	fluorescence two-hybrid
FACS	fluorescence activated cell sorting
FRAP	fluorescence recovery after photo bleaching
GFP	green fluorescent protein
H3K4me	histone 3 lysine 4 methylation
H3K9me3	histone 3 lysine 8 trimethylation
HDACs	histone deacetylases
HP1	heterochromatin protein 1
IAP	intracisternal A-type particle
ICF	immunodeficiency, centromere instability and facial anomalies
ICM	inner cell mass
IPS	induced pluripotent stem cell
LIF	leukemia inhibitory factor
LSD1	lysine specific demethylase
Lsh	lymphoid specific helicase
MBD	methyl CpG binding domain
MeCP2	Methyl CpG binding protein 2
mRFP	monomeric red fluorescent protein
NLS	nuclear localization signal
Np	nuclear protein
Np95	nuclear protein 95kDa
Np97	nuclear protein 97kDa
PBD	PCNA binding domain

PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PGC	primordial germ cell
PH	pericentic heterochromatin
PHD	plant homeo domain
PWWP	proline-tryptophane-tryptophane-proline motif
RF	replication foci
RING	really interesting new gene
S-phase	synthesis phase
SRA	SET and RING associated domain
SUV39H1	suppressor of variegation 3-9 homolog 1
TKO	dnmt1$^{-/-}$3a$^{-/-}$3b$^{-/-}$ triple knockout ESCs
TMAC	tetramethylammonium chloride
TS	targeting sequence
UBL	ubiquitin-like domain
UHRF	ubiquitin-like, containing PHD and RING finger domains
wt	wildtype
ZnF	zinc finger

4.3   Contributions

Declaration of contributions to "Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation" (see chapter 2.1)

This study was conceived by Heinrich Leonhardt and Lothar Schermelleh. I performed the bisulfite sequencing methylation analysis of the *skeletal α-actin* promoter and prepared supplementary figure 5. I established and performed a COBRA assay for IAP-LTRs and prepared figure 4 (A). The manuscript was written with Lothar Shermelleh and Heinrich Leonhardt.

Declaration of contributions to "Different binding properties and functional relevance of CXXC zinc finger domains in Dnmt1 and Tet1" (see chapter 2.2)

This study was conceived by Heinrich Leonhardt, Carina Frauer, Andrea Rottach and Fabio Spada. I performed the rescue assay in ES cells and analyzed methylation by pyrosequencing. I contributed and prepared figure 5, wrote the figure legend and the corresponding material and methods section. Heinrich Leonhardt, Carina Frauer and Andrea Rottach wrote the manuscript.

Declaration of contributions to "De novo Activity of DNA Methyltransferase 1 in Triple Knock-out mouse ES cells"(see chapter 2.3)

I conceived and laid out the project aims together with Heinrich Leonhardt and Fabio Spada. I generated the stably expressing Dnmt1 TKO cell lines and performed the transient rescue assays in TKO cells. I designed and established the protocols for the bisulfite pyrosequencing methylation analysis for all analyzed sequences. I developed a FACS based assay to establish stable TKO cells and performed the FACS sorting myself. I wrote the preliminary manuscript with the help of Heinrich Leonhardt.

Declaration of contributions to "DNA-methylation of CpG dyads and markov modeling of DNA-methylation control in mammals"(see chapter 2.4)

This study was conceived by Jörn Walter and Julia Arand. I isolated the genomic DNA of the knock out ES cells. Jörn Walter and Julia Arand wrote the manuscript.

Declaration of contributions to "Cooperative DNA and histone binding by Uhrf2 links the two major repressive epigenetic pathways" (see chapter 2.5)

This study was conceived by Garwin Pichler and Heinrich Leonhardt. I performed the rescue assay in ES cells and analyzed the methylation analysis by pyrosequencing. I contributed and prepared figure 5 and wrote the corresponding material and methods section. Garwin Pichler wrote the manuscript with the help of Heinrich Leonhardt.

Declaration of contributions to "Np95 interacts with de novo DNA methyltransferases Dnmt3a and 3b and mediates epigenetic silencing."(see chapter 2.6)

This study was conceived by me, Heinrich Leonhardt and Fabio Spada. I developed the CMV promoter silencing assay in ES cells and performed the silencing assay experiments with the help of Sebastian Bultmann. I performed the cell sorting myself on a FACS ARIA II. I designed and performed the methylation analysis of CAG and CMV promoter. I prepared figures 2, 3 and supplementary figures 4, 5 wrote the corresponding figure legends and materials and methods sections. Karin Fellinger designed and performed all biochemical experiments and identified and mapped the interactions described. I helped writing the manuscript together with Karin Fellinger, Fabio Spada and Heinrich Leonhardt.

Declaration of contributions to "Role of DNA methylation in silencing of pluripotency genes during embryonic stem cell differentiation" (see chapter 2.7)

I conceived and laid out the project aims together with Sebastian Bultmann, Heinrich Leonhardt and Fabio Spada. I differentiate ES cells into EBs with the help of Sebastian Bultmann. I designed and established the protocols for the bisulfite pyrosequencing methylation analysis for the pluripotency genes *oct* and *nanog.* I performed the methylation analysis of the promoter regions during differentiation. I established the stable cell lines expressing GFP under the *oct4* promoter together with Sebastian Bultmann. Christine Schmidt established and performed the expression analysis of the pluripotency genes *oct* and *nanog.* I wrote the preliminary manuscript with the help of Heinrich Leonhardt.

Declaration of contributions to "Identification of differentiating area" (see chapter 2.8)

This application note was conceived by Stefan Letzsch (PerkinElmer). I established the stable cell lines expressing GFP under the *oct4* promoter, differentiated them into EBs and fixed samples at day 4 and 8 with the help of Sebastian Bultmann. I wrote the material and methods section. The application note was written Stefan Letzsch.

4.4 Declaration According to the "Promotionsordnung der LMU München für die Fakultät Biologie"

Declaration According to the "Promotionsordnung der LMU München für die Fakultät Biologie"

Betreuung: Hiermit erkläre ich, dass die vorgelegte Arbeit an der LMU von Herrn Prof. Dr. Leonhardt betreut wurde.

Anfertigung: Hiermit versichere ich ehrenwörtlich, dass die Dissertation selbstständig und ohne unerlaubte Hilfsmittel angefertigt wurde. Über Beiträge, die im Rahmen der kumulativen Dissertation in Form von Manuskripten in der Dissertation enthalten sind wurde im Kapitel 4.3 Rechenschaft abgelegt und die eigenen Leistungen wurden aufgelistet.

Prüfung: Hiermit erkläre ich, dass die Dissertation weder als ganzes noch in Teilen an einem anderen Ort einer Prüfungskommission vorgelegt wurde. Weiterhin habe ich weder an einem anderen Ort eine Promotion angestrebt oder angemeldet oder versucht eine Doktorprüfung abzulegen.

München, den 20. September 2011

_____ (Daniela Meilinger)

4.5 Acknowledgment

First and foremost I would like to thank Prof. Dr. Heinrich Leonhardt for giving me the opportunity to conduct my PhD thesis in his group. I am very grateful that you gave me the chance to stay in your group for my PhD after finishing my diploma. Thank you for your tremendous support during the last years and for hours of inspirational scientific discussions. Last but not least, I would like to thank you for giving me the opportunity to stay even further – I'm looking forward to it!

A big thank you belongs to Fabio Spada with whom I had the pleasure to work during the last years. Clearly, I learned a lot from you! Thanks for your support and the endless scientific discussion.

A super special thank you to Anja Gahl, without whom I wouldn't have been able to survive on a daily basis. Anja, thank you so very much for being a wonderful friend and for your endless support both scientifically and personally. And thanks for kicking my a** from time to time when I needed it!

I would also like to thank Andrea Rottach, who is not only a colleague but also a friend. Thank you, Andrea, for being just you and supporting me, especially during these horrible writing times! I'm looking forward to working with you in the future.

Many thanks to Sebastian Bultmann. Bulti, it was such a pleasure working with you. Clearly, it was both super fun and very successful! Thanks!

Thanks to Susanne! I have no idea what we all did before we had you! Thanks for doing all the work without even having to ask!

A special thank you to Uli Rothbauer and Lothar Schermelleh. You guys are awesome and it was a great pleasure working with you! Uli, thanks for all your support.

I would like to thank Christine Schmidt and Carina Frauer. Thanks girls, it was great working with you. And thanks, Carina, for sharing your scientific mind with me!

A warm thank you goes to Katharina, who brought so much wonderful energy into lab. Thank you, Katharina, for being who you are!

Thanks to all the members of the group for the open and friendly atmosphere. It was a pleasure working with all of you!! Weihua: I so very much enjoyed working with you! Daniel: thanks for always making sure there is enough milk, I very much appreciate it! Andreas, Garwin, Patricia, Ola, Katrin, Jürgen, Wen, Mengxi, Kourosh, Jaqueline, Udo, Jonas, Irina and Boris, thanks for being such great colleagues.

Thanks to Yolanda. Girl, I have no idea what I would have done these last weeks without you! You are awesome!

Thanks to Heather and Heather. Thanks for cheering for me during the last years. I'm grateful to have found such wonderful friends. Heather, thanks for making my English sound better. And Pushy: Just plain and simple thanks for everything! <3

The most special thank you to my friend Katrin. I have no idea how I would have survived these past years without you. Having the chance to work with my best friend was wonderful. Now, that our ways part - I will miss you horribly. But, I know you will have a successful career – I'm so very proud of you!

Biggest thanks goes to my parents. Mama und Papa: Danke für einfach alles. Kein Kind kann sich bessere Eltern wünschen als Euch. Ich bin stolz eure Tochter zu sein und ich liebe Euch!

The most heartfelt thanks goes to my husband. Without you I would not be in Munich and would not have had the chance to meet all those wonderful people above. Thanks for being you, for your love, patience, endless support and for simply letting me be who I am! I love you and I know how happy it makes you that this part of our life is finally over ;)

List of publications

Schermelleh L, Haemmer A, Spada F, Rösing N, **Meilinger D**, Rothbauer U, Cardoso MC & Leonhardt H (2007) Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. *Nucleic Acids Res* **35:** 4301–4312

Meilinger D*, Fellinger K,* Bultmann S, Rothbauer U, Bonapace IM, Klinkert WEF, Spada F & Leonhardt H (2009) Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells. *EMBO reports* **10:** 1259–1264

Frauer C, Rottach A, **Meilinger D**, Bultmann S, Fellinger K, Hasenöder S, Wang M, Qin W, Söding J, Spada F & Leonhardt H (2011) Different binding properties and function of CXXC zinc finger domains in Dnmt1 and Tet1. *PLoS ONE* **6:** e16627

Pichler G, Wolf P, Schmidt CS, **Meilinger D**, Schneider K, Frauer C, Fellinger K, Rottach A & Leonhardt H (2011) Cooperative DNA and histone binding by Uhrf2 links the two major repressive epigenetic pathways. *J Cell Biochem* **112:** 2585–2593

Casas-Delucchi CS, van Bemmel JG, Haase S, Herce HD, Nowak D, **Meilinger D**, Stear JH, Leonhardt H & Cardoso MC (2011) Histone hypoacetylation is required to maintain late replication timing of constitutive heterochromatin. *Nucleic Acids Res*

Arand J, Spieler D, Karius T, **Meilinger D**, Meissner A, Okano M, Jenuwein T, Xu G, Leonhardt H, Wolf V & Walter J (2011) DNA-methylation of CpG dyads and markov modeling of DNA-methylation control in mammals (manuscript in preparation)

Letzsch S & **Meilinger D** Identification of differentiating area. (2011) Application Note, PerkinElmer