



# Quaternary structure and regulation of mammalian DNA methyltransferase 3a (Dnmt3a)

by

Arumugam Rajavelu

A Thesis submitted in partial fulfilment  
of the requirements for the degree of

Doctor of Philosophy in Biochemistry

Approved, Thesis Committee

---

Prof. Dr. Albert Jeltsch

---

Prof. Dr. Neil Brockdorff

---

Prof. Dr. Klaudia Brix

Date of Defense: 04.10.2011

---

School of Engineering and Science  
Jacobs University Bremen  
Bremen, Germany

# Declaration

Here by, I declare that the research work has been carried and given in the thesis are done by Arumugam Rajavelu and not copied from anywhere else.

Thanking you

Sincerely,

Arumugam Rajavelu

Date : 02.11.11

Place : Bremen

## List of publications

1. **Rajavelu A**, Jurkowska R, Fritz J, Jeltsch A (2011). Function and disruption of DNA Methyltransferase 3a cooperative DNA binding and nucleoprotein filament formation. *Nucleic Acids Res* (2011), in press.
2. Jurkowska RZ<sup>#</sup>, **Rajavelu A**<sup>#</sup> Anspach N, Urbanke C, Jankevicius G, Ragozin S, Nellen W, Jeltsch A (2011). Oligomerization and Binding of the Dnmt3a DNA Methyltransferase to Parallel DNA Molecules: Heterochromatic localization and role of Dnmt3L. *J Biol Chem.* 286 (27); 24200-7.

<sup>#</sup> These authors contributed equally to this study

3. **Rajavelu A**, Tulyasheva Z, Jaiswal R, Jeltsch A<sup>\*</sup>, Kuhnert N<sup>\*</sup> (2011). The inhibition of the mammalian DNA methyltransferase 3a (Dnmt3a) by dietary black tea and coffee polyphenols. *BMC Biochem.* 12; 16. **Highly accessed.**

<sup>\*</sup> Corresponding authors

4. Ceccaldi A, **Rajavelu A**, Champion C, Rampon C, Jurkowska R, Jankevicius G, Sénamaud-Beaufort C, Ponger L, Gagey N, Dali Ali H, Tost J, Vríz S, Ros S, Dauzonne D, Jeltsch A, Guianvarc'h D, Arimondo PB (2011). C5-DNA Methyltransferase Inhibitors: From Screening to Effects on Zebrafish Embryo Development. *Chembiochem.* 14; 12 (9); 1337-45.
5. Dhayalan A, **Rajavelu A**, Rathert P, Tamas R, Jurkowska RZ, Ragozin S, Jeltsch A (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J Biol Chem.* 285(34); 26114-20.

6. Zhang Y, Jurkowska R, Soeroes S, **Rajavelu A**, Dhayalan A, Bock I, Rathert P, Brandt O, Reinhardt R, Fischle W, Jeltsch A (2010). Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res.* 38(13); 4246-53.

## Manuscripts submitted

1. Deplus R<sup>#</sup>, Blanchon L<sup>#</sup>, **Rajavelu A**<sup>#</sup>, Boukaba H<sup>#</sup>, *et al* (Submitted for publication)  
<sup>#</sup> These authors contributed equally to this study
2. Halby L, Sénamaud-Beaufort C, Ajjan S, Ceccaldi A, Drujon T, **Rajavelu A**, *et al* (Submitted for publication)
3. Ceccaldi A, **Rajavelu A**, *et al* (Submitted for publication)

## Acknowledgements

I would like to thank my supervisor and mentor Prof. Dr. Albert Jeltsch for his excellent supervision, valuable discussions and constant encouragement towards science. I would like to express my sincere gratitude to him for providing me the opportunity to work in his lab and introducing me to the field of Epigenetics.

I am thankful to Prof. Neil Brockdorff, Oxford University, London, UK & Prof. Klaudia Brix, Jacobs University Bremen, Germany for being co-referees of my PhD thesis.

Many thanks to Prof. Francois Fuks, Free University of Brussels, Belgium, Prof. Paola B. Arimondo and Alexandre Ceccaldi, MNHN, CNRS INSERM, France and Prof. Nikolai Kuhnert, Jacobs University Bremen, Germany, for their collaboration and excellent support in my projects.

I would like thank Prof. Jürgen Fritz, Jacobs University Bremen, Germany & Prof. Wolfgang Nellen for collaboration on AFM experiments. Many thanks to Prof. Mathias Winterhalter, Jacobs University Bremen, Germany for extending his lab facility for fluorescence depolarization and FRET experiments.

I am thankful to Dr. Renata Jurkowska, Dr. Arunkumar Dhayalan, Dr. Sergey Ragozine, Dr. Yingying Zhang, Dr. Tomasz Jurkowski and all other senior members in our group, for their contribution to my work and providing me the friendly atmosphere in lab. I would like to thank especially Arun for many useful scientific and many other discussions.

Many thanks to my lab colleagues Abu, Ina, Qazi, Martin, Razvan, Pavel, Raluca, Hany and others for making friendly atmosphere in the lab. I am grateful for the help provided by Sandra Becker.

Many thanks to my Indian friends especially Srikanth, Mahendran, Shaik, Sanjay, Abhishek, Mahesh, Suneetha, Raghu, Binit, Rajesh, Sunil and many others for their friendly support during my stay at Jacobs.

I am always grateful to my family members, especially to my sisters for their constant support. Finally, I dedicate this whole work to my parents.

# Contents

List of publications.....	II
Acknowledgements.....	IV
Contents.....	V
List of figures.....	VII
List of abbreviations.....	VIII
Abstract.....	IX
1. Introduction.....	1
1.1 Gene regulation and development.....	1
1.2 Epigenetic players.....	3
1.3 DNA methylation.....	4
1.3.1 CpG Islands.....	6
1.4 DNA methyltransferases.....	7
1.4.1 Discovery of the Dnmt3 family.....	8
1.4.2 Biological role of the Dnmt3 enzymes.....	9
1.4.3 Isoforms of Dnmt3a and Dnmt3b.....	11
1.5 DNA Methyltransferase-Like regulatory protein (Dnmt3L).....	12
1.6 Structure of Dnmt3a-C/Dnmt3L-C complex.....	12
1.7 Multimerization of Dnmt3a-C/Dnmt3L-C on DNA.....	15
1.8 Post translational modification of proteins.....	16
1.9 Regulation of DNA methylation and Dnmt3a activity.....	17
1.9.1 Dynamic regulation of DNA methylation in mammals.....	17
1.9.2 Targeting of Dnmt3a to genomic regions.....	19
1.10 Control of DNA methylation.....	20
1.10.1 DNA methyltransferase inhibitors.....	21
1.10.2 Mechanisms of action of DNA methyltransferase inhibitors.....	21
2. Aims of the study and summary of key results.....	24

3. Discussion.....	26
3.1 Regulation of chromatin interaction of Dnmt3a.....	26
3.1.1 Dnmt3a ADD domain interacts with H3K4 unmethylated tail.....	26
3.1.2 Dnmt3a PWWP domain reads H3K36 trimethylation mark.....	27
3.1.3 Dnmt3L reorganizes quaternary structure of Dnmt3a and regulates Its localization in cells.....	29
3.2 Cooperative DNA binding of Dnmt3a and its function.....	32
3.2.1 Role of cooperative multimerization of Dnmt3a.....	35
3.3 Regulation of Dnmt3a activity and localization by post translational modification.....	36
3.3.1 Modulation of Dnmt3a activity by post translational modification.....	36
3.3.2 Compartmentalization of Dnmt3a depends on its regulator and PTM.....	37
3.4 Searching new inhibitors for Dnmt3a.....	39
3.4.1 Developing a high-throughput screening assay.....	39
3.4.2 Inhibition of Dnmt3a by black tea and coffee polyphenols.....	40
3.4.3 Synthesis of new DNMT inhibitors derivatives of procainamide.....	42
4. Conclusion.....	44
5. References.....	45
6. List of publications and author's contribution.....	57
7. Annex – 1.....	60
8. Annex – 2 confidential part available only to the reviewers.....	62

## List of figures

Fig 1. Depiction of Epigenetics function.....	2
Fig 2. Schematic picture of metabolism of the methyl group in mammalian cells.....	6
Fig 3. Schematic picture of the structure of eukaryotic DNA methyltransferases.....	8
Fig 4. Structure of the Dnmt3a-C and Dnmt3L-C heterotetrameric complex.....	14
Fig 5. Schematic representation of cooperative multimerization of Dnmt3a/Dnmt3L Complexes.....	15
Fig 6. Dynamics of DNA methylation in mammals.....	19
Fig 7. Structure of the DNMT inhibitors.....	21
Fig 8. Modeling of Dnmt3a hexameric structure with DNA.....	31
Fig 9. Modeling of Dnmt3a complex with DNA .....	34
Fig 10. Compartmentalization of Dnmt3a.....	38
Fig 11. Structures of chlorogenic compounds.....	41
Fig 12. The conjugated derivatives of procainamide.....	43

## List of abbreviations

ADD	– ATRX-Dnmt3a-Dnmt3L domain
AFM	– Atomic Force Microscopy
ATRX	– ATP dependent helicase domain
AZA	– 5-Azacytidine
CK2	– Casein kinase 2
CpG Island	– Cytosine-phosphate-Guanine rich DNA sequence
CXXC	– Cysteine-rich CXXC Zinc binding domain
DMR	– Differentially Methylated Regions
Dnm3a-C	– DNA methyltransferase 3a catalytic domain
DNMT	– DNA methyltransferase
Dnmt1	– DNA methyltransferase 1
Dnmt3a	– DNA methyltransferase 3a
Dnmt3b	– DNA methyltransferase 3b
Dnmt3L	– DNA methyltransferase 3L
Dnmt3L-C	– DNA methyltransferase 3L C-terminal domain
EGCG	– Epi-gallocatechin gallate
EMSA	– Electrophoretic Mobility Shift Assay
ES cells	– Embryonic stem cells
FRET	– Fluorescence Resonance Energy Transfer
H3K36me3	– Histone 3 lysine 4 trimethylated
H3K4me3	– Histone 3 lysine 4 trimethylated
H3K9me3	– Histone 3 lysine 4 trimethylated
IC <sub>50</sub>	– Inhibitory concentration 50
ICF	– Immunodeficiency, centromere instability and facial abnormalities
K3K4	– Histone 3 lysine 4
M.Hha1	– Methyltransferase from <i>Haemophilus haemolyticus</i>
Mtase	– Methyltransferase
NIH 3T3 cells	– Mouse embryonic fibroblast cell line
PHD	– Plant homeo domain
PWWP	– Proline Tryptophan Tryptophan Proline rich domain
SAH	– S-adenosyl-L-homocysteine
SAM	– S-adenosyl-L-methionine

# Abstract

The Dnmt3 family includes two catalytically active *de novo* methyltransferases (Dnmt3a and Dnmt3b) and one catalytically inactive regulatory factor called Dnmt3L. The N-terminal part of Dnmt3a consist of an ADD and a PWWP domain, we have shown that the ADD domain interacts with H3K4 unmethylated tails and the PWWP domain interacts with H3K36me3 marks. These interactions are responsible for the regulation of Dnmt3a's activity as well as for targeting to heterochromatin. We have identified that Dnmt3a and Dnmt3L form a heterotetramer that methylates two CpG sites on opposite DNA strands in a distance of 8-10 bps. We have shown that Dnmt3a forms a linear multimer and it binds to many DNA molecules oriented in parallel. This process is required for the heterochromatic localization of Dnmt3a in cells. The regulator protein Dnmt3L reorganizes the quaternary structures of Dnmt3a, thereby prevents multimerization. In addition to formation of protein multimers, Dnmt3a complexes bind cooperatively to DNA forming protein-DNA filaments. This might be required for generation of 8-10 bp periodicity patterns at imprinted genes and tight packing of Dnmt3a at the heterochromatic region. In addition, we have identified that the enzymatic activity Dnmt3a is regulated by phosphorylation. Casein kinase 2 mediated phosphorylation of the Dnmt3a negatively regulates its enzymatic activity in cells. Adding to existing evidence, our data uncovered novel mechanism of regulation of Dnmt3a activity, which are required for the proper DNA methylation patterns in cells. Also, using a newly developed high throughput assay, we have identified novel inhibitors to Dnmt3a from chemical libraries and identified new compounds purified from the dietary black tea and coffee polyphenols, which inhibit the Dnmt3a activity moderately.

# 1. Introduction

Epigenetics is a branch of biology, which deals with the study of the changes in gene expression, which occur in organisms with differentiated cells. The most recent definition is “changes in the phenotype of the cells that are heritable but do not involve DNA mutation” [Gottschling, 2004]. The term epigenetics was introduced by Waddington, “Epi” is a latin prefix meaning above, over, outside or beyond [Waddington, 1942]. Epigenetic mechanisms are used for the formation and storage of cellular information in response to transient environmental signals. The initial population of embryonic stem (ES) cells is identical in both, genotype and phenotype, but during development, the ES cells follow distinct differentiation pathways. The mature cells become phenotypically different although they remain genotypically identical [Levenson and Sweatt, 2005]. There are many studies and examples which show the importance of epigenetic mechanisms in the information storage at the cellular level. Once an embryonic precursor cell is triggered by an environmental signal to differentiate into a particular cell type (for instance, a liver cell), that cell and its subsequent daughter cells are required to undergo thousands of cell divisions without losing their identity. To maintain the cell’s identity and cellular differentiation, the genetic information must be expressed at right time and right place (Fig. 1). The information responsible for this cannot be contained on the DNA sequence itself but is encoded by heritable epigenetic mechanisms [Levenson and Sweatt, 2005].

## 1.1. Gene regulation and development

In eukaryotes, DNA is not naked, but packed into a highly organised nucleoprotein structure called chromatin, which maintains genes in an inactive state by restricting access of the transcription machinery i.e. RNA polymerase and its

accessory factors. The fundamental unit of chromatin is the nucleosome, which is composed of 147 base pairs of DNA wrapped around the histone octamer [Li and Reinberg, 2011]. The core histone proteins (H2A, H2B, H3 & H4) are globular and highly basic proteins. These histone proteins have unstructured N terminal tails and undergo several modifications which influence the local gene transcription [Li and Reinberg, 2011]. Consecutive nucleosomes are assembled into 11-nm fibers (resembling beads on a string), which are then compacted in association with the linker histones (H1) into a poorly defined, higher-order structure and eventually into individual chromosomes [Felsenfeld and Groudine, 2003].

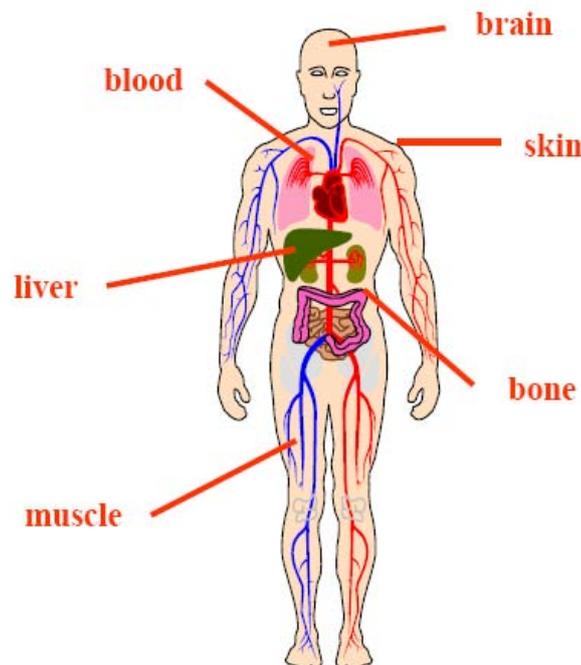


Fig.1. Depiction of epigenetic function: Cellular development and gene expression in mammals is mediated by epigenetic mechanisms. (Adopted from Albert Jeltsch, MolLife).

According to the current models chromatin functions as a structural device for packaging DNA as well as in regulation of gene expression. Compaction of the DNA with histones in chromatin influences the accessibility for transcription factors and cofactors and thus regulates gene expression. It is known that the establishment of

patterns of gene expression determines the cellular phenotype and this can be influenced by how DNA is packed into chromatin. In eukaryotes not all the chromatin is equal, rather it is divided into two functional units. The untranscribed chromatin is packed in highly condensed heterochromatin, while transcribing regions are more accessible to transcription factors and form euchromatin [Richards and Elgin, 2002; Margueron and Reinberg, 2010].

DNA methylation has been implicated in playing an important role in the control mechanisms that govern gene expression in mammals [Jaenisch and Bird, 2003]. An inverse correlation between DNA methylation and gene expression was demonstrated for a large number of genes, indicating that DNA methylation plays a vital role to suppress gene expression [Weber and Schübeler, 2007; Zhang *et al.*, 2009; Boon *et al.*, 2002; Kim *et al.*, 2005]. The transcriptional suppression of genes by CpG methylation could be either by direct interference with the binding of transcription factors or by binding of multiprotein repressor and co-repressor complexes, which results in the formation of an inactive chromatin structure [Fuks *et al.*, 2001; Fuks *et al.*, 2003]. Association of repressors and co-repressors with the methylated DNA mediates a permanent inactivation of genes leading to a stable repressed state.

## 1.2. Epigenetic players

There are three molecular epigenetic players mainly involved in gene regulation *i.e.* DNA methylation, histone modifications and non-coding RNA [Cheng and Blumenthal, 2008; Hermann *et al.*, 2004; Gangaraju and Lin, 2009]. Regulation of gene expression is tightly controlled by these epigenetic players [Jeltsch, 2002; Li and Reinberg, 2011; Mochizuki, 2010]. Misregulation of epigenetic systems has

potential deleterious effect on control of gene expression and chromatin accessibility, which leads to diseased state of cells.

### 1.3. DNA Methylation

One of the important epigenetic modifications in unicellular and multicellular organisms is DNA methylation, which regulates and mediates the function of the genes and differentiation of cells. In mammals and other vertebrates, DNA methylation occurs at the C5 position of cytosine bases, mostly at the cytosine base followed by guanine (CpG dinucleotide). The complementary strand of a CpG site is also methylated and these two methyl groups are positioned in the major groove of double stranded DNA [Ohki *et al.*, 2001]. The 5-methylcytosine constitutes 3-8% of the total cytosine bases of the DNA, therefore it could be considered as a fifth base of mammalian DNA [Hermann *et al.*, 2004]. It has many functions *viz*, the control of cellular differentiation and development [Cheng and Blumenthal, 2010; Jurkowska *et al.*, 2011; Kim *et al.*, 2009], the maintenance of chromosomal integrity [Lengauer *et al.*, 1999], parental imprinting [Holmes and Soloway, 2006], the inhibition of gene expression [Jaenisch and Bird, 2003], X-chromosome inactivation [Chow and Brown, 2003] and it also plays a role in the development of the immune system [Fitzpatrick and Wilson, 2003].

DNA methylation is important for normal cellular differentiation and viability in mammals, deletion of any of the active DNA methyltransferase enzyme in mice is lethal [Li *et al.* 1992; Okano *et al.*, 1999]. Abnormal methylation patterns of genomes are associated with several human diseases, such as Rett syndrome, ICF (immunodeficiency, centromere instability and facial abnormalities) and Fragile X syndrome. Rett syndrome is a neurological disorder resulting in mental retardation, which is caused by mutations in one of the methyl-CG binding proteins (MeCP2)

[Kriaucionis and Bird, 2003], ICF syndrome is caused by point mutations occurring in the gene encoding for Dnmt3b methyltransferase [Xu *et al.*, 1999; Hansen *et al.*, 1999]. Aberrant methylation of an amplified repeat in the promoter region of FMR1 gene leads to the Fragile X syndrome, which is associated with hypermethylation of genes and changes in the chromatin structure [Robertson and Wolffe, 2000]. Hypermethylation of tumor suppressor genes and hypomethylation of the genome of somatic cells contribute to genomic instability and the development of cancer [Clark and Melki, 2002].

Though methylation of CpG sites in the mammalian genome has variety of important functions for normal cellular development and differentiation, methylated cytosines represent mutational hot spots. The deamination of cytosine bases leads to the formation of uracil, which can be efficiently recognized and repaired by the uracil deglycosylase pathway. However, deamination of a methylated cytosine base leads to the formation of thymine, which can not be recognized by this repair pathway, hence presence of methylated cytosine bases increases the mutational rate of DNA [Hermann *et al.*, 2004].

The methylation of DNA is carried out in a reaction catalyzed by DNA methyltransferases. These enzymes use the cofactor SAM as a donor of an activated methyl group, as a result 5-methylcytosine is created and the SAH is released from the enzyme. In living eukaryotic cells folate is the source of the C1 units needed to regenerate SAM from SAH. The activated methyl groups from SAM are transferred into DNA, RNA and proteins by methyltransferases (Fig 2).

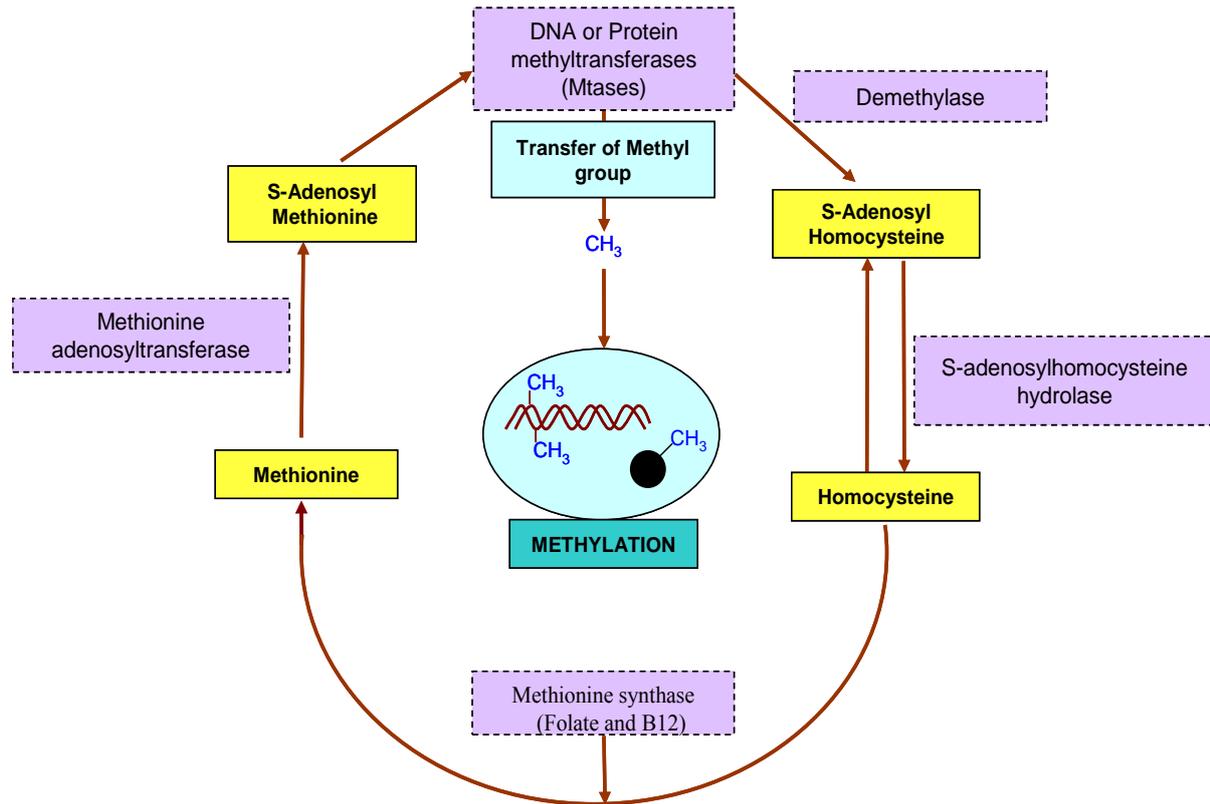


Fig 2. Schematic representation of metabolism of the methyl group in mammalian cells. Transfer of  $\text{CH}_3$  group from S-adenosyl-L-methionine to DNA, RNA and proteins by methyltransferases (Mtases).

### 1.3.1. CpG Islands

The most important targets of the DNA methylation in mammalian genomes are cytosine bases within the dinucleotide CpG. A high density of the CpG dinucleotides is present at the transcription start sites of approximately 50% of all mammalian genes, which generally are unmethylated [Ishikhes and Zhang, 2000]. The co-localization of CpG islands with promoters suggest that regulation of CpG island methylation may have role in transcription and indeed CpG islands can become methylated during mammalian development. Having numerous methylation sites, the immunity of the CpG islands for the methylation is one of the unresolved paradoxes of the biology of the CpG island and probably holds a key functional significance. The possible hypothesis for the absence of the DNA methylation at the

CpG island are that the presence of transcription factors on the DNA may prevent the accessibility of the DNA methyltransferases to the DNA. In addition, it has been suggested that CpG island could be possible target for the active DNA demethylation in early embryogenesis [Hsieh, 1999].

#### 1.4. DNA methyltransferases

The mammalian DNA methylation machinery consists of DNA methyltransferase 1 (Dnmt1) and the DNA methyltransferase 3 family (Dnmt3) [Gowhar, 2002; Hermann, 2004]. Dnmt1 is generally referred to as maintenance methyltransferase, because it maintains the methylation pattern of CpG sites during each round of DNA replication [Hermann *et al.*, 2004; Goyal *et al.*, 2006; Chen and Li, 2006; Mortusewicz *et al.*, 2005]. Dnmt1 shows preference towards hemimethylated substrates *in vitro* and this property of the enzyme makes it suitable for its *in vivo* function i.e copying the methylation pattern of the parental strand to the daughter strand during DNA replication [Jeltsch *et al.*, 2006]. Dnmt1 is the largest protein among the Dnmts, it was the first DNA methyltransferase to be discovered and well characterized. Studies have shown that this enzyme localizes to the DNA replication fork [Leonhardt *et al.*, 1992]. Also, it has been shown that association of the Dnmt1 with replication machinery increases the efficiency of DNA methylation at replication foci [Schermelleh *et al.*, 2007; Spada *et al.*, 2006]. The Dnmt3a and Dnmt3b enzymes are designated *de novo* methyltransferases, as they do not have a preference for hemimethylated or unmethylated DNA substrates [Okano *et al.*, 1998]. They are responsible for the establishment of initial DNA methylation patterns during embryonic development which gets inherited through cell division [Lei *et al.*, 1996]. The Dnmt3 family includes two active enzymes (Dnmt3a and Dnmt3b) and one

inactive regulatory factor called DNA methyltransferase-like protein (Dnmt3L) [Okano *et al.*, 1999; Chedin *et al.*, 2002].

The distinct feature of all Dnmts is that they have a common structure of a C-terminal catalytic domain, which contains up to 10 amino acid motifs which are conserved among prokaryotes and eukaryotes [Gowher *et al.*, 2002; Hermann *et al.*, 2004]. Motifs I and X have a role in cofactor binding, motifs IV and VI are involved in catalysis. The conserved region between motifs VIII and IX represents the so called Target Recognition Domain (TDR), which is involved in DNA recognition (Fig 3).

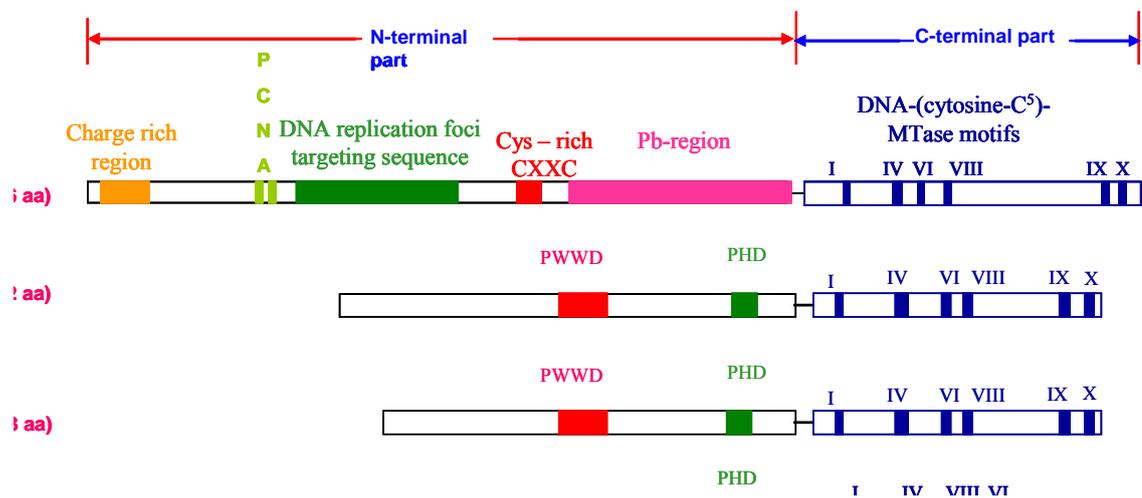


Fig 3. Schematic picture of the structure of eukaryotic DNA methyltransferases, roman letters at the C-terminal part of the proteins indicate the conserved motifs in the catalytic domains of DNA methyltransferases [Adopted from Hermann A, *et al.*, 2004].

### 1.4.1 Discovery of the Dnmt3 family

Dnmt1 null mutant embryonic stem cells still contained a small percentage of methylated DNA and showed detectable DNA methyltransferase activity [Li *et al.*, 1992]. Additionally, provirus DNA integrated in Dnmt1 null mutant embryonic stem cells became methylated, suggesting the presence of additional methyltransferase

activity [Lei *et al.*, 1996]. In 1998 Okano *et al.*, first cloned and characterized the Dnmt3 family of enzymes (Dnmt3a and Dnmt3b) that methylate both hemimethylated and unmethylated DNA at the same rate [Okano *et al.*, 1998]. This family of enzymes is present in Mouse, Human, Arabidopsis, Maize and other organisms, it is heavily expressed in embryonic tissues but only low expression is observed in differentiated cells [Okano *et al.*, 1998; Jeltsch, 2002]. These methyltransferases methylate CpG dinucleotides during embryogenesis, which explains the importance of these enzymes for the *de novo* methylation of DNA [Okano *et al.*, 1998; Fatemi *et al.*, 2002; Gowher *et al.*, 2005; Gowher *et al.*, 2006]. Dnmt3a is expressed ubiquitously and the expression of Dnmt3b occurs at very low levels in most tissues except testis, thyroid gland and bone marrow [Xie *et al.*, 1999]. The expression of Dnmt3b increases profoundly in various tumor cell lines, suggesting that it might play a role in carcinogenesis [Robertson *et al.*, 1999]. The third member of the Dnmt3 family, called Dnmt3L, shares clear homology with Dnmt3a and Dnmt3b enzymes, but lacks the DNA methyltransferase activity. Dnmt3L has been identified later by sequence homology to Dnmt3a and Dnmt3b and it functions as regulator of Dnmt3a and Dnmt3b [Bourc'his *et al.*, 2001; Chedin *et al.*, 2002; Hata *et al.*, 2002; Gowher *et al.*, 2005].

#### 1.4.2 Biological role of the Dnmt3 enzymes

The *de novo* methyltransferases Dnmt3a and Dnmt3b are important for the normal embryonic development of mice [Okano *et al.*, 1999]. Mouse Dnmt3b *-/-* knockout embryos die *in utero* with multiple developmental defects, whereas Dnmt3a knockout mice develop but die shortly after birth. In contrast, the Dnmt3L *-/-* mice are viable and do not show any distinguishable morphological defects. The female Dnmt3L *-/-* mice are fertile but fail to deliver viable pups, because these embryos

show defects in neuronal development. In contrast, the male Dnmt3L  $-/-$  are sterile and fail to produce mature spermatozoa. Detailed studies showed that Dnmt3L is essential for the establishment of genomic imprints in oocytes and male germ cells. To study the role of Dnmt3a and Dnmt3b in germ line development, conditional knock out mice were generated. The Dnmt3b conditional germ line knockout shows no apparent phenotype, whereas the phenotype of corresponding Dnmt3a knockout mice is characterized by an altered methylation of imprinted sequences in both male and female germ cells [Bourc'his *et al.*, 2001; Bourc'his and Bestor, 2004]. These results indicate that Dnmt3a and Dnmt3L are required for the methylation of most imprinted loci in germ cells. Although Dnmt3a and Dnmt3b have similar sequence and biochemical properties, these enzymes exhibit distinct biological functions. Dnmt3a methylates single copy genes, retrotransposons and it is involved in the establishment of the genomic imprints during germ cell development [Hata *et al.*, 2002; Bourc'his *et al.*, 2001]. Dnmt3b is responsible for methylation of pericentromeric satellite regions containing repetitive DNA [Bourc'his and Bestor, 2004]. It has been shown that Dnmt3b methylates CpG sites within the pericentromeric regions of the chromosomes. These regions contain a type of satellite DNA, called satellite 2 and 3, which is normally heavily methylated. Mutations within the Dnmt3b gene (mostly in the C-terminal domain of the enzyme) which reduces its activity and is associated with a human genetic disorder called immunodeficiency, centromere instability and facial anomalies syndrome (ICF syndrome). The affected patients show facial dimorphism, mental retardation and prolonged respiratory, cutaneous and gastrointestinal infections. The ICF syndrome is associated with a specific loss of DNA methylation at the satellites 2 and 3 of the pericentromeric region of the chromosomes 1, 9 and 16 [Xu *et al.*, 1999; Hansen *et al.*, 1999].

### 1.4.3 Isoforms of Dnmt3a and Dnmt3b

The DNA methylation pattern varies from tissue to tissue and it is still largely unknown how the methyltransferases generate this dynamic methylation in mammals. It has been shown that the *de novo* DNA methyltransferase 3a or 3b gene produces different protein isoforms at different developmental stages using alternative internal promoters or alternative splicing. Two isoforms of Dnmt3a were identified in the cells, the shorter isoform (Dnmt3a2) is transcribed from an internal start site under the control of an intronic promoter and lacks the first 223 (human), 219 (mouse) amino acid residues of full length protein [Chen *et al.*, 2002]. The full length protein is the major form in adult tissues and it co-localizes with heterochromatin, while Dnmt3a2 is predominant in embryonic stem cells and embryonic carcinoma cells and it localizes to euchromatin, suggesting that the N-terminal part of the full length Dnmt3a might have a targeting function [Weisenberger *et al.*, 2002]. Six isoforms of Dnmt3b have been reported, all the transcripts results from alternative splicing of different exons and these isoforms are expressed at different levels during embryonic development. Among the six Dnmt3b isoforms only Dnmt3b1 and Dnmt3b2 showed enzymatic activity in DNA Mtase assay, whereas Dnmt3b3, Dnmt3b4 and Dnmt3b5 appeared to be inactive, which was expected because these isoforms lack important active site motifs in the catalytic domain [Weisenberger *et al.*, 2002; Hermann *et al.*, 2004; Gopalakrishnan *et al.*, 2009].

The murine gene products of the Dnmt3a and Dnmt3b comprise 908, 853 amino acid residues, respectively and share 37% of sequence identity [Jeltsch *et al.*, 2002]. Dnmt3a and Dnmt3b have similar domain arrangements, both contain a variable N-terminal region followed by a PWWP domain, ADD domain (cys-rich CXXC Zinc binding domain), which is similar to the ATRX or PHD domain, followed by the C-terminal catalytic domain [Cheng *et al.*, 2008] (Fig 3). Studies have shown

that PWWP domain of Dnmt3a and Dnmt3b targets these enzymes to the satellite DNA at the pericentric heterochromatin [Chen *et al.*, 2004; Ge *et al.*, 2004].

### 1.5. DNA Methyltransferase-Like Regulatory protein (Dnmt3L)

Dnmt3L is a small protein, it contains an N-terminal ADD domain and C-terminal domain with homology to the catalytic domain of Dnmt3a and Dnmt3b, which contains many mutations at important residues, which explains why the protein lacks DNA methyltransferase activity. Dnmt3L is expressed specifically in germ cells during gametogenesis and embryonic stages and was shown to be essential for the establishment of genomic imprints in oocytes [Hata *et al.*, 2002]. It has been demonstrated that Dnmt3L interacts with the catalytic domains of the Dnmt3a and Dnmt3b and it acts as a positive regulator of mammalian DNA methyltransferases by stimulating their activity up to 15 fold [Gowher *et al.*, 2005; Chedin *et al.*, 2002]. It has been reported that Dnmt3L reads the unmethylated H3K4 via its ADD domain and the catalytic domain of Dnmt3L interacts with Dnmt3a-C and brings the Dnmt3a to establish a heritable DNA methylation pattern [Ooi *et al.*, 2007].

### 1.6. Structure of the Dnmt3a-C/Dnmt3L-C complex

Dnmt3a comprises of an N-terminal regulatory domain and a C-terminal catalytic domain. Unlike the Dnmt1 C-terminal domain, which lost its Mtase activity in the isolated form, the catalytic domain of the Dnmt3a (Dnmt3a-C) alone retains substantial methyltransferase activity [Gowher *et al.*, 2002]. It has been shown that Dnmt3L interacts with the unmethylated H3K4 tail via its ADD domain and through its catalytic domain interacts with the C-terminal domain of the Dnmt3a [Margot *et al.*, 2003; Gowher *et al.*, 2005; Ooi *et al.*, 2007]. Jia *et al.* (2007) solved the structure of Dnmt3a-C in complex with Dnmt3L-C, this complex contains two monomers of

Dnmt3a-C and two monomers of Dnmt3L-C (3L-3a-3a-3L) forming a tetramer. The overall complex is elongated (about 160x60x50 Å) with a butterfly shape (Jia *et al.*, 2007) (Fig 4A). Dnmt3a-C and Dnmt3L-C both have the classical fold characteristic for the SAM dependent DNA methyltransferases. The crystallographic studies show the presence of the cofactor product SAH only with Dnmt3a-C not with Dnmt3L-C (Fig 4A). This result strongly suggests that Dnmt3a-C is the catalytic component of this complex, which is supported by lack of Mtase activity by Dnmt3L [Jia *et al.*, 2007, Gowher *et al.*, 2005].

The tetrameric complex has two 3a-3L interfaces and one 3a-3a interface. The 3a-3L interaction is mediated by hydrophobic interactions represented by two pairs of phenylalanines F728 and F768 of mouse Dnmt3a and F297 and F337 of mouse Dnmt3L, whereas the 3a-3a interaction is mediated by two charged amino acids R881 and D872 (Fig 4A) [Jia *et al.*, 2007; Jurkowska *et al.*, 2008]. The introduction of a R881A mutation in Dnmt3a-C at the 3a-3a interface eliminates the polar interactions and abolished the catalytic activity of Dnmt3a-C. A F728A mutation in Dnmt3a-C at the 3a-3L interface eliminated the hydrophobic interaction, as well as the catalytic activity of the tetrameric complex. These results illustrate the importance of these interfaces for the catalytic activity of the tetrameric complex [Jurkowska *et al.*, 2008]. Modeling of the double stranded DNA into the 3L-3a-3a-3L tetrameric complex suggests that the active sites of the two Dnmt3a are located in the major groove of DNA about 40 Å apart from each other such that the two central Dnmt3a-C subunits could methylate two CG sites separated by 10 base pairs on opposite DNA strands in one binding event (Fig 4B).

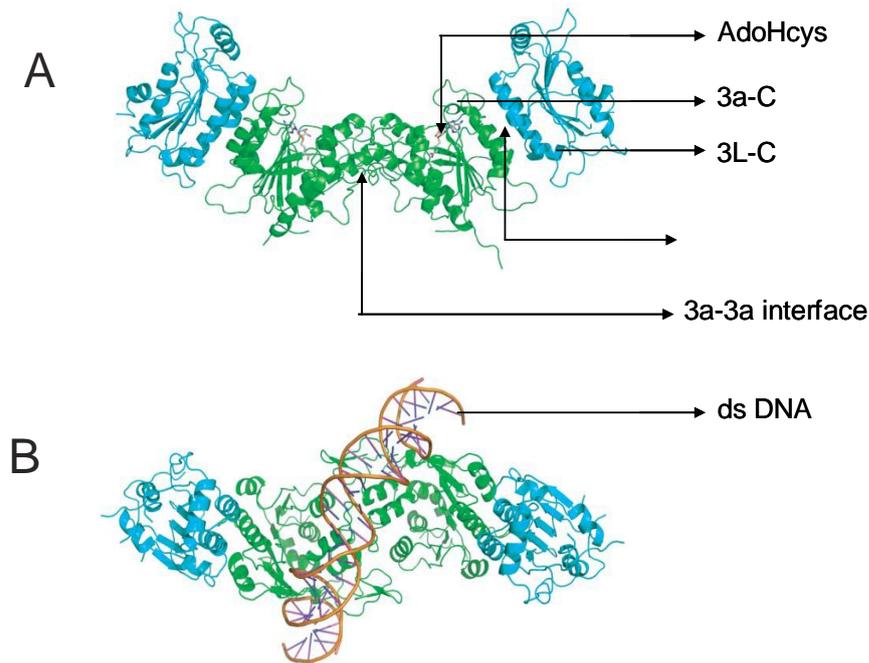


Fig 4. Structure of the Dnmt3a-C and Dnmt3L-C heterotetrameric complex (3L-3a-3a-3L): A) Location of the hydrophobic (3a-3L) and hydrophilic (3a-3a) interfaces in the tetrameric complex and presence of AdoCys within Dnmt3a-C. B) Modelling of DNA with 3a/3L tetramer indicates location of the two active sites in the major groove of the DNA double helix. [Picture adopted from Jia *et al.*, (2007)].

The interesting question is why the catalytic domains of Dnmt3a should dimerize? The possible explanation for the 3a-3a dimerization could be that the DNA binding surface area of Dnmt3a-C is relatively small and Dnmt3a-C dimerization effectively doubles this area, which may be needed for good DNA binding. In addition, the domain has two active sites and it can methylate two CpG sites by single binding event [Jia *et al.*, 2007]. In the Dnmt3a/Dnmt3L tetramer, RD interface is required for the DNA binding, where the two active sites of the Dnmt3a are located at the distance of 8-10 bp in opposite DNA strands. The geometry of two active sites in 3a/3L tetramer suggested that it could methylate efficiently two CpG sites separated at 8-10 bp distance in opposite DNA strands [Jurkowska *et al.*, 2008; Jia *et al.*, 2007].

## 1.7. Multimerization of Dnmt3a-C/Dnmt3L-C on DNA

Jurkowska *et al* (2008), showed a cooperative multimerization of Dnm3a-C/Dnm3L-C complexes on the DNA using Electrophoretic Mobility Shift Assay (EMSA) that was further confirmed by using scanning force microscopy [Jurkowska *et al.*, 2008]. The DNA methylation patterns on the longer DNA substrates showed that the Dnmt3a alone or Dnmt3a/Dnmt3L complexes generated a 8-10 bp methylation pattern on the same DNA strand, indicating a regular arrangement of Dnmt3a complexes on the DNA. Jurkowska *et al.*, (2008) used hairpin bisulfite approach to study geometry of two active sites in the 3a/3L tetramer, showing that one tetramer methylates two CpG sites in the opposite DNA strands. In addition, the regular arrangement of the Dnmt3a multimeric complex on DNA generated a 8-10 bp methylation pattern on the same DNA strand. The biochemical studies, hairpin bisulphite experiments and modeling of Dnmt3a suggested that two Dnmt3a-C molecules from adjacent tetramers can approach each other, the left complex would interact with the CpG site in the upper strand of the DNA whereas the right complex would interact with the lower strand (Fig.5).

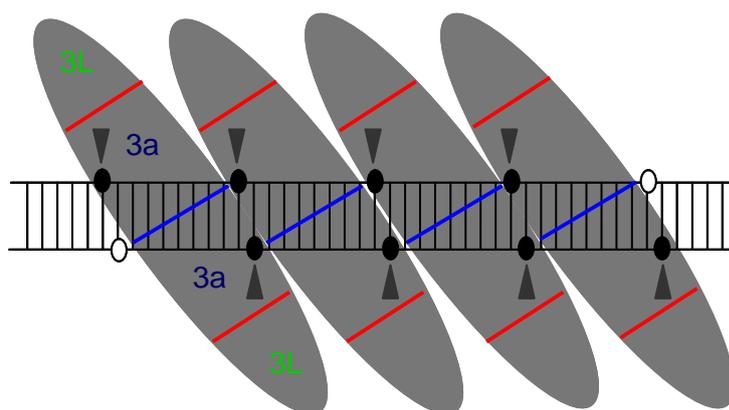


Fig 5. Schematic picture of Dnmt3a/Dnmt3L multimerization: Representation of cooperative multimerization of Dnmt3a/Dnmt3L complexes on the DNA. Generation of 8-10 bp periodic methylation pattern on DNA by Dnmt3a/Dnmt3L complex.

The results of EMSA and scanning force microscopy validated that 3a/3L tetramers cooperatively multimerizes on the DNA next to each other cooperatively and form a nucleoprotein filament [Jurkowska *et al.*, 2008]. To form a cooperative complex on the DNA, the Dnmt3a complex requires a functional interface to interact with the adjacent complex. Formation of cooperative nucleoprotein filament by Dnmt3a on the DNA could efficiently methylate the target DNA substrates containing CpG sites present in a periodic pattern of 8-10 bp. Such kind of periodicity has been observed in differentially methylated DNA sequences associated with maternally imprinted genes [Jia *et al.*, 2007; Jurkowska *et al.*, 2008].

## 1.8. Post translational modification of proteins

In mammalian cells, the majority of proteins are subjected to modification at the translational level as well as post translationally. The modification could be either the addition of chemical group to the side chain of an amino acid or the cleavage of the protein at a specific site to become an activated form of the protein [Bradshaw *et al.*, 1998]. The simplest modification of proteins is the addition of small chemical groups like acetyl, methyl or phosphate groups to amino acid side chains or the termini of the polypeptide. So far 150 different modifications have been found in many proteins, the most common modifications are acetylation, methylation, phosphorylation, glycosylation and ubiquitination. In the context of epigenetic regulation, the modification of histone tails influences the local gene transcription and translation. The most important modifications, which influence gene transcription is the methylation of histone 3 lysine 4 which leads to activation of gene transcription, whereas the methylation of lysine 9 and lysine 27 of histone 3 tails cause repression of gene transcription [Black and Whetstine, 2011; Klose and Zhang, 2007; Kouzarides, 2007]. The unstructured N termini of histone proteins are also subjected

to other modifications, including acetylation and phosphorylation [Voigt and Reinberg, 2011]. Phosphorylation and dephosphorylation of serine, threonine and tyrosine residues are the best known modifications involved in reversible, activation and inactivation of enzyme activity, changing the localization of proteins, regulation of protein-protein interaction and also plays an important role in signalling cascade pathway [Pawson, 2002; Seo *et al.*, 2003]. Post-translational phosphorylation of proteins is one of the most common reversible protein modifications that occur in animal cells. Addition of a phosphate group is carried out by kinases, whereas the dephosphorylation of proteins is done by phosphatases.

## 1.9. Regulation of DNA methylation and Dnmt3a activity

Why the regulation of Dnm3a or Dnmt3b enzyme in cells is required? Writing of a precise DNA methylation at different developmental stages, needs regulation of Mtases. Its well known that the post translational modification of proteins are important for the regulation of their activity [Seger *et al.*, 1994; Olsen *et al.*, 2006] and Dnmt3a or Dnmt3b might undergo the posttranslational modifications that might plan a role in their regulation. The regulation of Dnm3a is required during cell cycle, where after the S-phase of cell cycle, the unmethylated DNA needs to be methylated. Also the regulation of Mtases activity could be very important to prevent the aberrant DNA methylation.

### 1.9.1 Dynamic regulation of DNA methylation in mammals

Coordinated and tightly controlled gene expression in mammalian cells is mediated by DNA methylation patterns, which are set during embryogenesis by the

Dnmt3a and Dnmt3b DNA methyltransferases [Bourc'his *et al.*, 2001; Hata *et al.*, 2002; Hermann *et al.*, 2004]. During early development of mammals, shortly after fertilization, a genome wide active demethylation takes place. The fertilized cells remain unmethylated after multiple rounds of cell division until morula stage when *de novo* methylation is established by Dnmt3a and Dnmt3b during blastocyst development [Reik *et al.*, 2001; Wu and Morris, 2001]. To reach the unmethylated state, there are two possible mechanisms; one could be existence of an active demethylase, the other passive demethylation. In active demethylation, methyl groups from DNA removed by demethylases, whereas in passive demethylation, exclusion of Dnmt1 from the nucleus, dilutes the methylation during each round of replication to generate unmethylated DNA [Howell *et al.*, 2001] (Fig. 6).

The differentiation of cells is tightly controlled by DNA methylation, during differentiation, the not needed genes are specifically suppressed and Dnmt3a and/or Dnmt3b probably contributed to this gene silencing. Thus, the DNA methylation pattern is dynamically regulated in mammals, after setting the initial methylation pattern during development by the *de novo* DNA methyltransferases, while it is maintained by the Dnmt1 during each round of cell division [Chen *et al.*, 2006; Mortusewicz *et al.*, 2005]. In addition to the Dnmt1, recent studies show that Dnmt3a and Dnmt3b play an important role in maintaining the DNA methylation of repetitive DNA and heterochromatin in adult cells [Dodge *et al.*, 2005; Jones and Liang, 2009; Feng *et al.*, 2010].

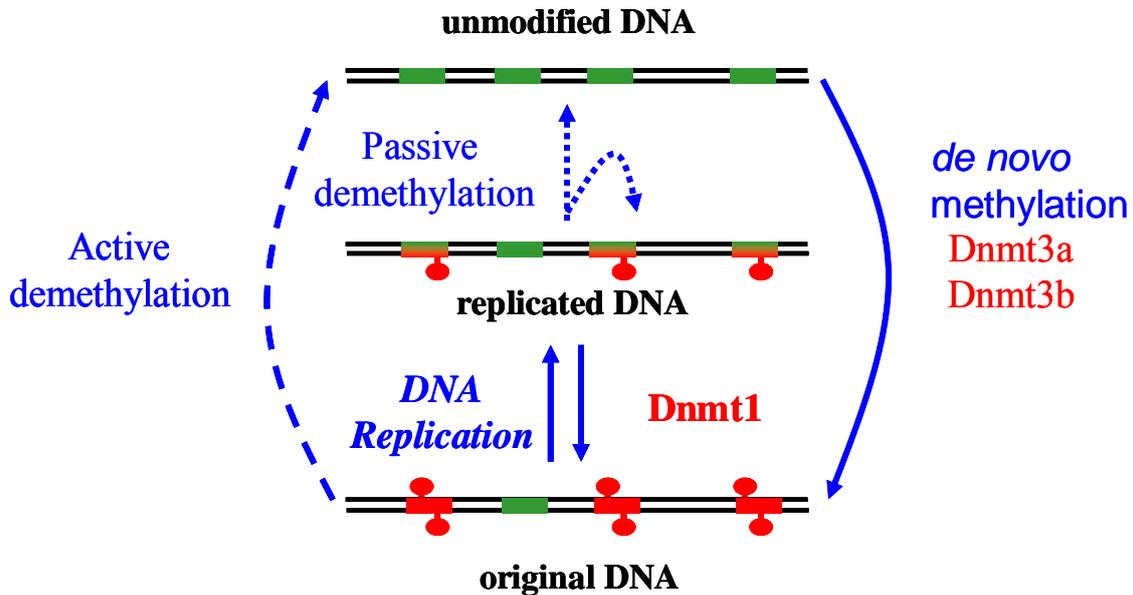


Fig 6. Dynamics of DNA methylation in mammals: *de novo* methylation is carried out by Dnmt3a & Dnmt3b. During each round of DNA replication, DNA methylation is preserved by Dnmt1. The solid line represent that the evidence proved by experiment. Demethylation might takes place either by active demethylase or passive demethylation. Dotted line represent hypothesis yet to be proven [Adopted from Albert Jeltsch, Mol Life].

### 1.9.2 Targeting of Dnmt3a to genomic regions

The mechanisms of targeting Dnmt3a or Dnmt3b to specific genomic loci are not well understood. As mentioned these enzymes consist of a C-terminal catalytic domain and an N-terminal part consisting of a PWWP domain and an ADD domain [Hermann *et al.*, 2002; Cheng *et al.*, 2008]. Recently, it has been shown that interaction of the N-terminal parts of Dnmt3a with chromatin might have a role in targeting to specific genomic region [Jeong *et al.*, 2009]. Interaction of Dnmt3a through its N-terminal part to the ankyrin domain of G9a, recruited it to the Oct3/4 pluripotency determining gene in embryonic stem cells (ES cells) and permanently

suppressed its expression [Litman *et al.*, 2008]. The PWWP domain of Dnmt3a and Dnmt3b specifically targets the enzyme to the heterochromatin [Chen *et al.*, 2004; Ge *et al.*, 2004]. The targeting of Dnmt3a to specific genomic loci could be mediated by post translational modification of Dnmt3a, in which the Dnmt3a is regulated by selective compartmentalization.

## 1.10 Control of DNA methylation

DNA methylation is vital for normal development and maintenance of tissue-specific gene expression patterns in mammals and it should be precisely delivered to specific genomic loci. Misregulation of DNA methylation patterns leads to diseases like ICF (Immunodeficiency, centromeric instability and facial anomalies), X-fragile syndrome and various cancers [Robertson and Wolffe, 2000; Sharma *et al.*, 2010]. Recent evidences show that many cancer cells exhibit a disturbed epigenetic landscape, which often includes a global hypomethylation of the genome that causes abnormal expression of genes and the hypermethylation of gene promoters that silences tumor suppressor genes [Laird, 2005; Ehrlich, 2009; Feinberg and Vogelstein, 1983]. The site-specific hypermethylation leads to tumorigenesis by silencing of tumor suppressor genes. Initially it has been shown that Rb, a tumor suppressor gene associated with retinoblastoma [Greger, 1989] and many other tumor suppressor genes, including p16, MLH1 and BRCA1 undergo gene silencing by hypermethylation [Jones and Baylin, 2002; Jones and Bylin, 2007; Baylin, 2005]. Hypermethylation could be caused by aberrant expression of DNA methyltransferase or abnormal activity of DNA methyltransferases. To reduce hypermethylation one could apply Dnmts inhibitors, which could be used as epigenetic therapy.

### 1.10.1 DNA methyltransferase inhibitors

Recent evidence shows that interference with DNA methylation at the level of methyltransferases is a valid strategy to reactivate suppressed genes. Controlling of DNA methylation by DNMT inhibitors leads to the inhibition of cancer initiation and progression [Christman, 2002]. Two types of DNMT inhibitors are available, which are nucleoside analogues and non-nucleoside analogues. Cytidine and its analogues like azacitidine (5-azacytidine) and decitabine (5-aza-2'-deoxycytidine) are the nucleoside DNMT inhibitors (Fig 7) [Fenaux, 2005], whereas hydralazine, EGCG and RG108 belong to the non-nucleoside inhibitors [Cornacchia *et al.*, 1988; Chavez-Blanco *et al.*, 2006]. Azacitidine and decitabine were the first molecules shown to inhibit DNMTs [Sorm *et al.*, 1964]. These pyrimidine analogues of cytidine are incorporated into RNA or DNA and form covalent complexes with DNA methyltransferases, leading to a depletion of active enzymes

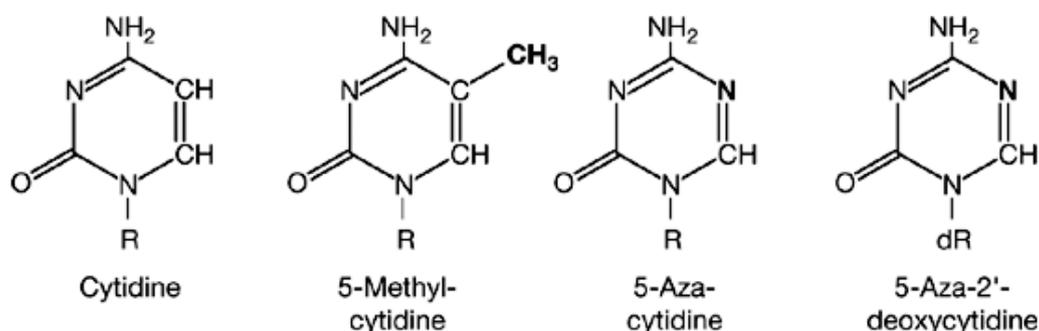


Fig 7. Structure of cytidine and the DNMT inhibitors 5-methylcytidine, azacitidine (5-azacytidine) and decitabine (5-aza-2'-deoxycytidine) (Adopted from Fenaux, 2005).

### 1.9.2. Mechanisms of action of DNA methyltransferase inhibitors

**Azacitidine and decitabine:** The nucleoside analogues azacitidine and decitabine are incorporated into the DNA during DNA replication, where they trap the DNMTs leading to the depletion of DNA methyltransferases. Thereby, these drugs

act as inhibitors for the DNA methylation by preventing DNMTs activity [Jones and Taylor, 1980; Bender *et al.*, 1998; Chuang *et al.*, 2010]. Recently, the FDA (Food and Drug Administration) had been approved 5-Azacytidine (Vidaza) and 5-aza-2'-deoxycytidine (Dacogen<sup>TM</sup>) for the treatment of myelodysplastic syndromes and various solid tumours [Steensma *et al.*, 2009; Gore *et al.*, 2006; Müller *et al.*, 2006; Ulrich-Pur *et al.*, 2001]. They were first synthesized as cytotoxic agents, in the 1980s, these compounds were found to have hypomethylating activity after incorporation into the DNA of actively replicating tumor cells [Jones and Taylor, 1980; Taylor and Jones, 1979]. The cytidine analogues also been used for treatment of many solid tumors [Ulrich-Pur *et al.*, 2001].

**Zebularine:** A nucleoside analogue that inhibits DNMTs. It does not incorporate into RNA but only into the DNA. Zebularine is incorporated into the DNA and traps the DNMTs forming enzyme-DNA adducts, which leads to a reduced methylation in cells [Gravina *et al.*, 2010].

**Epigallocatechin gallate (EGCG), RG108 and Procainamide:** All these inhibitors belong to the non-nucleoside analogue inhibitors family. EGCG is the main compound found in green tea, however during tea fermentation EGCG is completely transformed into higher more complex polyphenols of unknown structure. A recent study by Lee *et al* showed that the EGCG inhibits DNMT1 activity [Lee *et al.*, 2005]. Some of the EGCG is transformed into formally dimeric structures like theaflavins of theasinensins, which are present in larger quantities in black tea. Little is known about the biological activities of any black tea compound, although black tea is by far the most consumed beverage worldwide.

Procainamide was used as antiarrhythmic drug since several years and it has been shown that it induces a significant demethylation in the genomic DNA of cancerous cells [Cornacchia *et al.*, 1998, Viller-Garea *et al.*, 2003]. Procainamide

inhibits Dnmt1 by reducing the affinity of the DNMT to the DNA and cofactor significantly, thereby it causes demethylation [[Gravina et al., 2010](#)].

RG108 is a small molecule, non-nucleoside DNA methyltransferase inhibitor that blocks the active site of Dnmt1 and inhibits the enzyme activity. It has been used as epigenetic reactivation of tumour suppressor genes, it reduces the DNA methylation in cancerous cells [[Stresemann, 2006](#); [Schirmacher, 2006](#)].

## 2. Aims of the study and summary of key results

### 1. Chromatin targeting and regulation of Dnmt3a localization

One important aim of this study was to improve our understanding of the regulation of Dnmt3a and its targeting. Two studies from our laboratory showed that ADD and PWWP domains of Dnmt3a interact with H3K4 unmethylated tail and H3K36me3 respectively. These interactions are responsible for the regulation of enzyme activity as well as targeting of Dnmt3a to heterochromatin. In addition to the N-terminal part, the catalytic domain of the Dnmt3a interacts with Dnmt3L and forms a stable tetramer. We have shown that in the absence of Dnmt3L, Dnmt3a forms dimers via the FF interface, which further oligomerize via their RD interfaces. The oligomerization of Dnmt3a-C creates multiple DNA binding sites, which allow for binding of separate DNA molecules oriented in parallel. Since Dnmt3L does not have an RD interface, it prevents Dnmt3a oligomerization and binding of more than one DNA molecules. We have shown that oligomerization of Dnmt3a is necessary for its proper heterochromatic localization in cells. Over expression of Dnmt3L in cells leads to the release of Dnmt3a from heterochromatic regions and makes it available in euchromatic region [Zhang *et al.*, 2010; Dhayalan *et al.*, 2010; Jurkowska *et al.*, 2011].

### 2. Cooperative multimerization of DNA methyltransferase 3a (Dnmt3a)

The Dnmt3a is cooperatively multimerizes on the DNA. We wanted to investigate the function and mechanisms of this process. We have shown the functional interface in Dnmt3a, which is required for its cooperative multimerization. We have validated the importance of the interface by mutating residues critical for multimerization and DNA binding studies using the gel shift assay and atomic force

microscopy. In addition we have shown that the cooperative multimerization of Dnmt3a may contribute to proper heterochromatic localization of Dnmt3a [[Rajavelu et al., 2011](#)].

### 3. Regulation of Dnmt3a activity post translational modification

During the development of mammals, DNA methylation patterns are dynamically regulated. The generation of such dynamic methylation pattern is likely associated with regulation of enzyme activity. Here, we showed that post translational modification of Dnmt3a controls the DNA methylation pattern. Also, we have observed that post translational modification of Dnmt3a is necessary for its proper heterochromatin localization in cells [[Deplus R et al., in revision](#)].

### 4. Identification of potential inhibitors for Dnmt3a

Aberrant DNA methylation of gene promoters of tumor suppressor genes leads to carcinogenesis. To prevent or revert hypermethylation in the genome, specific DNA methyltransferase inhibitors would be required. Here, we have screened and identified new compounds purified from the black tea, which inhibits the Dnmt3a activity moderately. Also using a newly developed high throughput assay, we have identified novel specific inhibitors for Dnmt3a [[Rajavelu et al., 2011](#); [Ceccaldi et al., 2011](#); [Halby et al., Submitted](#); [Ceccaldi et al., Submitted](#)].

## 3. Discussion

### 3.1 Regulation of chromatin interaction of Dnmt3a

Dnmt3a is localized to the heterochromatin in cells and it has been shown that the Dnmt3a is strongly associated with methylated nucleosomes [Chen *et al.*, 2004; Jeong *et al.*, 2009; Sharma *et al.*, 2011]. A chromatin based interaction of Dnmt3a with histone tails has been proposed for the regulation of DNA methylation in cells. In our laboratory, we have studied the specific interaction of N-terminal domains of Dnmt3a with modified histone tail, which regulates the activity of Dnmt3a enzyme. In addition to N-terminal domains, we have shown that the regulator protein Dnmt3L regulates the localization of Dnmt3a enzyme in cells.

#### 3.1.1 Dnmt3a ADD domain interacts with H3K4 unmethylated tail

Dnmt3a contains N-terminal parts including a PWWP domain and a PHD-like ADD domain [Hermann *et al.*, 2004]. It has been shown that the ADD domain of Dnmt3a interacts with many other proteins and serves as platform for protein-protein interactions. The ADD domain of Dnmt3a shares homology with PHD domain of ATRX, Dnmt3b, Dnmt3L (ADD-ATRX-Dnmt3a-Dnmt3L). A recent study showed that Dnmt3L, a member of the Dnmt3 family, interacts specifically with H3K4 unmethylated histone tails through its ADD domain and this interaction was abolished by H3K4 trimethylation [Ooi *et al.*, 2007]. Recently, Otani *et al.* (2009), solved the structure of Dnmt3a-ADD with unmodified histone tail. Also, they have shown that trimethylation at H3K4 eliminates the interaction of ADD domain H3 tail [Otani *et al.*, 2009]. To understand more about the function of the modification based interaction of Dnmt3a with chromatin, we have used modified histone tail peptide array for interaction studies. We showed that the ADD domains of Dnmt3a and Dnmt3b bind to

unmodified histone 3 tails and do not interact with H3K4me3 peptides. We have observed that the H3–ADD domain interaction stimulates the catalytic activity of Dnmt3a2 or Dnmt3a2-Dnmt3L for methylation of linker DNA of oligonucleosomes carrying specific modifications [Zhang *et al.*, 2010]. The methylation of linker DNA was not affected by H3K9 trimethylation, whereas H3K4 trimethylation decreased the activity of the Dnmt3a2 or Dnmt3a-Dnmt3L complex on nucleosomes.

The Dnmt3L protein expresses only in germ cells where it is essential for *de novo* methylation and setting maternal imprints during development [Bourc'his D *et al.*, 2001; Ooi *et al.*, 2007] but it shows little or no expression in somatic cells [Bourc'his D *et al.*, 2001; Hata *et al.*, 2002] Our data explored a new mode of interaction of Dnmt3a with histone tails, this is independent of Dnmt3L. The Dnmt3a-ADD domain – histone 3 interaction might be necessary for DNA methylation in somatic cells where the Dnmt3L is absent. In conclusion, we provided the first evidence that preferential binding of a Dnmt3a to histone tails carrying specific post-translational modification pattern directly leads to the methylation of DNA bound to the modified chromatin. We demonstrate that this effect can be mediated by ADD domain in Dnmt3a, which binds to chromatin as like the Dnmt3L ADD domain [Ooi *et al.*, 2007].

### 3.1.2 Dnmt3a PWWP domain reads H3K36 trimethylation mark

The PWWP domain belongs to the Royal domain superfamily, members of which were identified to interact with histone tails in various modification states [Maurer-Stroh *et al.*, 2003; Taverna *et al.*, 2007]. The PWWP domains of Dnmt3a and Dnmt3b are essential for heterochromatic targeting and localization in cells [Chen *et al.*, 2004]. Using peptide arrays containing modified histone tails, we have identified that the PWWP domain of Dnmt3a specifically interacts with H3K36 trimethylation

mark and mutation of critical residues within the aromatic binding pocket abolished this interaction [Dhayalan *et al.*, 2010]. The interaction of PWWP with H3K36me3 is very specific but it has only a weak affinity. In cells, the interaction could be strengthened by the interaction of the ADD domain with the unmodified H3K4. In addition, it has been shown that the PWWP domain weakly binds with DNA [Chen *et al.*, 2004; Purdy *et al.*, 2010] and the lysine 36 on the H3 tail is positioned close to the DNA. Hence, the PWWP domain could interact with both the histone tail and the DNA simultaneously. We showed that the H3K36me3-PWWP interaction mediates chromatin binding of Dnmt3a, because the heterochromatic localization of Dnmt3a was disturbed in the D329A mutant, which no longer binds to H3K36me3. Targeting of Dnmt3a by an H3K36me3-PWWP interaction is in agreement with previous results showing heterochromatic localization of Dnmt3a in NIH 3T3 cells [Chen *et al.*, 2004; Ge *et al.*, 2004].

We have observed that the Dnmt3a PWWP-H3K36me3 interaction stimulates the methylation activity of Dnmt3a on chromatin-bound DNA isolated from human cells, whereas the catalytic activity of the D329A mutant (which lost the H3K36me3 interaction) was similar to that of the isolated catalytic domain. Our result suggests an important role for the PWWP domain of Dnmt3a in guiding DNA methylation to chromatin carrying the H3K36me3 mark, which is in agreement with the observation that the PWWP domain is required for the activity of Dnmt3a in the cell [Chen *et al.*, 2004; Shikauchi *et al.*, 2009]. It has been found that H3K36me3 is present in euchromatin in the body of active genes, and its presence is anti-correlated with H3K4me3. The genome-wide distribution of DNA methylation is very similar to that of H3K36me3 methylation on histones which strengthens our model. [Vakoc *et al.*, 2006; Barski *et al.*, 2007; Larschan *et al.*, 2007]. A recent study on genome wide DNA methylation analysis showed that the correlation of H3K36me3 and DNA

methylation was observed at exon-intron boundaries, where the exons were shown to have increased levels of both H3K36me3 and DNA methylation [Hodges *et al.*, 2009]. The direct correlation of H3K36me3 and DNA methylation in gene bodies, suggests that the PWWP – K36me3 interaction has an important role in the regulation of enzyme activity in the gene body of cells.

### 3.1.3 Dnmt3L reorganizes the quaternary structure of Dnmt3a and regulates its localization in cells

#### Dnmt3a homo-oligomerization creates multiple DNA binding sites

In the Dnmt3a-Dnmt3L tetramer, the Dnmt3a-Dnmt3L contact is mediated by the FF interface. We showed that the FF interface of Dnmt3a not only supports the Dnmt3a-Dnmt3L interaction but also a Dnmt3a-Dnmt3a interaction. Therefore, Dnmt3a can form long proteins filaments by alternative usage of its FF and RD interfaces (Fig 8). Since, the RD interface contains the DNA binding surface, such protein filament could bind two or more DNA molecules. Modeling of such complexes suggested that these DNA molecules would be oriented roughly in parallel to each other, such that one linear DNA could not bind to separate RD sites on one Dnmt3a multimer. To study the parallel DNA binding of Dnmt3a oligomer in solution, we have developed a Fluorescence Resonance Energy Transfer (FRET) assay. We have prepared a DNA substrate that consists of two 20 base pair double-stranded regions separated by a 10 nucleotide single-stranded region, which is flexible. Both ends of the DNA were labeled with fluorophores that constitute a FRET donor/ acceptor pair (Cy3/Cy5). In solution, the average distance of both ends was large enough not to permit FRET. However, upon binding of the Dnmt3a oligomer to the DNA substrate, the ends approached each other and FRET was established [Jurkowska *et al.*, 2011].

The disruption of the FF interface or preincubation of Dnmt3L with Dnmt3a abolished the binding of Dnmt3a to parallel DNA [Jurkowska *et al.*, 2011]. Binding of two or more DNA molecules oriented in parallel is an unusual mode of DNA binding that has been never observed with any of the methyltransferase, so far. However it has been observed with few proteins, one well known example is the RecA protein [Bell, 2005].

### Structural rearrangement of Dnmt3a by its regulator Dnmt3L

We have observed that Dnmt3a-C tends to aggregate during protein purification at the dialysis step when the salt concentrations are changed from 500 mM to 200 mM KCl. This aggregation of Dnmt3a is a reversible process, presence of Dnmt3L at the equimolar concentration prevents the oligomerization of Dnmt3a [Jurkowska *et al.*, 2011]. In addition, we have shown that the Dnmt3a-Dnmt3L interface (FF interface) also supports self-interaction of Dnmt3a-C and that Dnmt3a-C dimers can form protein oligomers using their RD interfaces. In analytical ultracentrifugation, Dnmt3a sediments as dimer, tetramer and higher aggregates depending on the concentration of the protein, whereas the 3a/3L complex sediments as defined heterotetramer [Jurkowska *et al.*, 2011]. We therefore suggested that the Dnmt3a multimerization via the FF and RD interface is responsible for its reversible aggregation and Dnmt3L prevents this process. In agreement with our results, previous studies showed that Dnmt3a2 elutes as large multimeric structures and the mass equivalent to the multimeric complex (roughly 8 monomers) in size exclusion chromatography [Jurkowska *et al.*, 2008; Cheng and Blumenthal, 2008; Jia *et al.*, 2007]. The pre-incubation of Dnmt3a with the Dnmt3L protein, reorganized the large multimeric complexes of Dnmt3a into defined tetrameric structures [Kareta *et al.*, 2006; Jurkowska *et al.*, 2011]. In conclusion, we show that the Dnmt3a tend to self

interact and forms a large multimeric complex and this interaction is prevented by Dnmt3L, because it does not contain an RD interface (Fig 8).

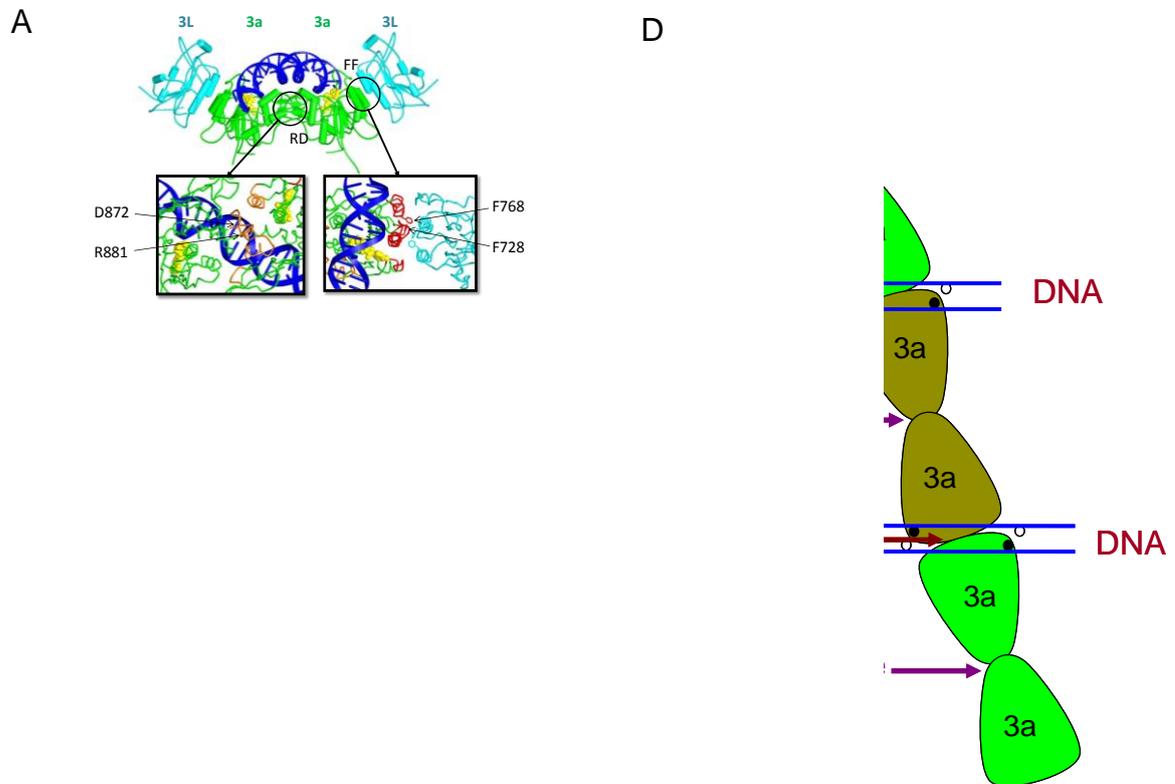


Fig 8. Modeling of Dnmt3a yielded hexameric structure containing two DNA binding sites, which bind two DNA molecules oriented in parallel. A). 3a/3L heterotetramer containing two FF interface and one RD interface. B) and C). Modeling of a Dnmt3a showing a hexamer containing two DNA binding RD interfaces and three FF interfaces. D) Schematic representation of a Dnmt3a hexamer which binds with two DNA molecules oriented in parallel.

### Regulation of the sub-nuclear localization of Dnmt3a by its physiological regulator Dnmt3L

Dnmt3a is localized to the heterochromatin in NIH3T3 cells. Targeting of this enzyme to the heterochromatin is mediated by its PWWP domain and the interaction of the PWWP domain with H3K36me3 [Chen *et al.*, 2004; Dhayalan *et al.*, 2010].

Dnmt3L disrupts the oligomerization of Dnmt3a and prevents the binding of multiple DNA molecules oriented in parallel [Jurkowska *et al.*, 2011]. To explore the functional significance of this type of oligomerization, we have studied its role in heterochromatin localization. Initial results suggested that the Dnmt3a oligomerization plays a role in heterochromatin localization, because disruption of the oligomeric interfaces, led to the loss of the specific heterochromatic localization in NIH3T3 cells [Jurkowska *et al.*, 2011]. Loss of heterochromatin localization by interface mutants could be explained because oligomerized Dnmt3a binds to multiple DNA strands and loss of multiple DNA binding leads to loss heterochromatic localization. Interestingly, the co-expression of Dnmt3a with Dnmt3L in NIH3T3 cells also led in more than 60% of the cells to a loss of the Dnmt3a heterochromatic spots. Co-expression of Dnmt3L F261D which does not interact with Dnmt3a [Jurkowska *et al.*, 2008], had no effect on the localization of the Dnmt3a enzyme in NIH3T3 cells [Jurkowska *et al.*, 2011]. The redistribution of Dnmt3a by its regulator represents a novel mode of action of Dnmt3L in setting DNA methylation patterns. It is well known that Dnmt3L stimulates the activity of Dnmt3a enzyme. In addition, to the stimulatory effect, we showed that Dnmt3L might be involved in the release of Dnmt3a proteins from dense heterochromatin to make it available to act at imprinted differentially methylated regions, which are generally located in euchromatin (Fig 9).

### 3.2 Cooperative DNA binding by Dnmt3a and its function

Previously, from our laboratory it has been shown that Dnmt3a forms a cooperative complex on the DNA at increasing concentration of proteins [Jurkowska *et al.*, 2008]. We confirmed cooperative multimerization of Dnmt3a/3L and Dnmt3a complexes on DNA using by various methods, *viz* EMSA, fluorescence depolarization, AFM and hairpin bisulfite DNA methylation analysis, which revealed

the fluctuations of methylation activities along the DNA molecule. All these experimental evidences strongly suggested that there must be a regular arrangement of Dnmt3a forming a cooperative complex on the DNA [Jia *et al.*, 2007; Jurkowska *et al.*, 2008; Rajavelu *et al.*, 2011]. We have modeled a complex in which two Dnmt3a-C molecules would interact simultaneously with the upper and lower strand of symmetrical CpG sites. For modeling we have used the structure of the bacterial M.HhaI methyltransferase bound to DNA as template [Klimasauskas *et al.*, 1994], which was the first methyltransferase for which base flipping could be shown. In this model, two loops of the two Dnmt3a-C complexes approach each other symmetrically. These contacting loops are enriched with charged amino acids and form a hydrophilic interface, which explains the cooperative multimerization of Dnmt3a-C and Dnmt3a-C/3L-C complexes [Rajavelu *et al.*, 2011].

Modeling of two 3a/3L heterotetramers with DNA suggested that two loops (residues 820-855) of the two Dnmt3a-C complexes interacting with one central CpG site are facing each other (Fig 9). As mentioned earlier these loops created hydrophilic environment to form an interface for interaction of adjacent Dnmt3a complexes leading to a cooperative multimerization of Dnmt3a-C and Dnmt3a-C/3L-C complexes on DNA. To disrupt this cooperative multimerization, we have introduced opposite charged amino acids at this putative interface. Afterwards we expected Dnmt3a complexes to repel each other and lose their ability of cooperative multimerization on the DNA [Rajavelu *et al.*, 2011]. Our results could verify this prediction and we have validated the loss of cooperative multimerization of Dnmt3a by using fluorescence depolarization, EMSA and AFM, in which Dnmt3a mutants binds as single molecule on the DNA [Rajavelu *et al.*, 2011] (Fig 9C).

Fig 9. Modeling of Dnmt3a complex with DNA: A). Model of the central Dnmt3a-C dimers (colored orange and red for one complex and dark green and green for the other) bound to DNA (blue). The flipped target cytosines are shown in pink, AdoMet is in yellow. The expanded version showing the interacting loops from the adjacent complexes. The region comprising amino acids 827-844 is not ordered in the structure (labeled yellow). B). Schematic model of Dnmt3a-C/3L-C heterotetramers binding next to each other to form a cooperative complex on DNA, the central CpG site could be methylated in both DNA strands. The interaction of adjacent 3a/3l tetramers mediated by hydrophilic interface (Marked in black circle), the interacting Dnmt3a-C subunits of the left tetramer are colored red, those of the right tetramer dark green. The RD interface is marked by a green arrow and FF interface is marked by a blue arrow. C). DNA binding and cooperativity of the interface loop mutants studied by EMSA. Dnmt3a-C wild type protein forms a large complex with very low electrophoretic mobility (region 3). Free DNA runs in region 1. The R832E, K837E and R832E/K837E mutants bind to DNA, but form smaller complexes with an electrophoretic mobility that corresponds to one Dnmt3a-C complex being bound to the DNA (region 2).

### 3.2.1 Role of cooperative multimerization of Dnmt3a

In our laboratory, we observed a novel property of the Dnmt3a enzyme, which is that at increasing concentration Dnmt3a forms a cooperative complex on DNA *in vitro*. Initially, it was difficult to understand and interpret this unexpected property of Dnmt3a. To understand the multimerization better, we have set up an *in vitro* assay. We have generated non-polymerizing variants by site directed mutagenesis, which lost their cooperative complex formation. Consistent with previous observation, wild type Dnmt3a-C methylates DNA with 8-10 bp periodicities, whereas non-polymerizing variants lost their cooperative multimerization (Dnmt3a-C R832E and Dnmt3a-C K837E) as well as defined methylation peaks on the hairpin DNA substrates [Jurkowska *et al.*, 2008; Rajavelu *et al.*, 2011].

It is known that Dnmt3a and Dnmt3L are required for methylation at imprinted loci during development [Bourc'his *et al.*, 2001; Kaneda *et al.*, 2004; Bourc'his and Bestor, 2004] and the representation of CpG sites in an 8-10 bp CpG periodicity was shown the DNA sequence of 12 known maternally imprinted DMRs [Jia *et al.*, 2007]. Recent genome wide methylation analyses also showed a methylation at 8-10 bp periodicity in the genome [Zhang *et al.*, 2006; Zhang *et al.*, 2009; Chodavarapu *et al.*, 2010]. The presence of such methylation patterns in the genome suggests that cooperative multimerization of Dnmt3a could generate this periodicity pattern. Dnmt3a and Dnm3L are expressed at high level in germ cells [Watanabe *et al.*, 2002; La Salle *et al.*, 2006] and required for the *de novo* methylation as well as setting the imprints in embryo. It is logical to propose that a high level expression of Dnmt3a in germ cells could lead to multimerization of Dnmt3a and formation of a cooperative complex at the specific loci. The question remains what makes the initiation of multimerization at a specific locus to generate the 8-10 bp periodic methylation patterns in genome? There are many factors which might contribute to the initiation of

multimerization of Dnmt3a, it could be the interaction of Dnmt3a with chromatin or specific genomic sequences itself might initiate the cooperative multimerization.

We have studied the importance of cooperative multimerization of Dnmt3a in heterochromatin localization. Dnmt3a is localized to heterochromatin in NIH3T3 cells, whereas non-polymerizing mutants lost their specific localization pattern in NIH3T3 cells. Tight packing of Dnmt3a to these regions requires the close approximation of adjacent Dnmt3a complexes bound to the same DNA molecule, which requires some interface for adjacent Dnmt3a. *In vitro* this interface could mediate a cooperative multimerization. Dnmt3a R832E/K837E double mutation interferes with such approximation, which leads to weakening of the heterochromatic localization NIH 3T3 cells and loss of cooperative DNA binding *in vitro*.

### 3.3 Regulation of Dnmt3a activity and localization by post translational modification

#### 3.3.1 Modulation of Dnmt3a activity by post translational modification

Though the structure and function of Dnmt3a has been well studied, the generation of DNA methylation patterns and targeting of DNA methyltransferases to specific genomic targets is not well understood. Emerging evidence suggests that an interaction of the N-terminal domains of the enzyme with modified histone tails could mediate the specific targeting [Ooi *et al.*, 2007, Zhang *et al.*, 2010; Otani *et al.*, 2009; Dhayalan *et al.*, 2010]. In addition to the existing evidence, our recent work shows that regulation of global DNA methylation is mediated by post translational modification of Dnmt3a [Deplus *et al.*, 2011 in revision]. The targeting function of post translationally modified Dnmt3a could be mediated by ability of Dnmt3a to interact with one or more components of the histone modification and chromatin-remodelling systems, like the H3K36me3 mark in histone 3 tail, the SUV39H H3K9 Mase

enzymes or the protein LSH, that are involved in the establishment of DNA methylation at heterochromatin [[Lehnertz et al., 2003](#); [Yan et al., 2003](#); [Dhayalan et al., 2010](#)]. Still there remain many questions how the Dnmts are targeted to sub-genomic regions and generate the methylation patterns? ([Fig 10](#)).

### 3.3.2 Compartmentalization of Dnmt3a depends on its regulator and PTM

Targeting of Dnmt3a methyltransferase to sub-genomic regions is very important to control the writing of a precise methylation patterns. In addition to the existing evidence to targeting function of Dnmt3a, here we have provided that Dnmt3a localization depends on two factors ([Fig 10](#)).

The first is the regulator protein Dnmt3L; the interaction of Dnmt3L with Dnmt3a releases the protein from heterochromatin and makes it available to methylate euchromatic DNA. We have provided evidence that the Dnmt3L regulates the compartmentalization of the Dnmt3a proteins within the nucleus [[Jurkowska et al., 2011](#)]. Both the Dnmt3a and Dnmt3L proteins are required for the methylation of differentially methylated regions (DMR) in imprinted genes [[Bourc'his et al., 2001](#); [Kaneda et al., 2004](#); [Bourc'his et al., 2004](#)]. The interaction of Dnmt3L with Dnmt3a might target the enzyme to euchromatin to methylate imprinted genes.

The second factor is post translational modification of Dnmt3a, which modulates the enzyme's activity and localization of the enzyme in the cell. The post translational modification of Dnmt3a could be a dynamic process and might be required at different stages of the cell cycle to regulate the DNA methylation pattern. The two independent studies from our laboratory showed that the localization of Dnmt3a is regulated by two different mechanisms. Interaction of Dnmt3L with Dnmt3a stimulates the activity of enzyme [[Gowher et al., 2005](#)] as well as releases

the Dnmt3a from heterochromatin and targeted to euchromatic regions [Jurkowska *et al.*, 2011]. In this work, we have identified that the Dnmt3a is modified and this modification reduces the activity of Dnmt3a and targets the Dnmt3a to heterochromatin [Deplus *et al.*, in revision]. It is interesting to note that in these fundamentally different process reduction of catalytic activity is associated with heterochromatin localization.

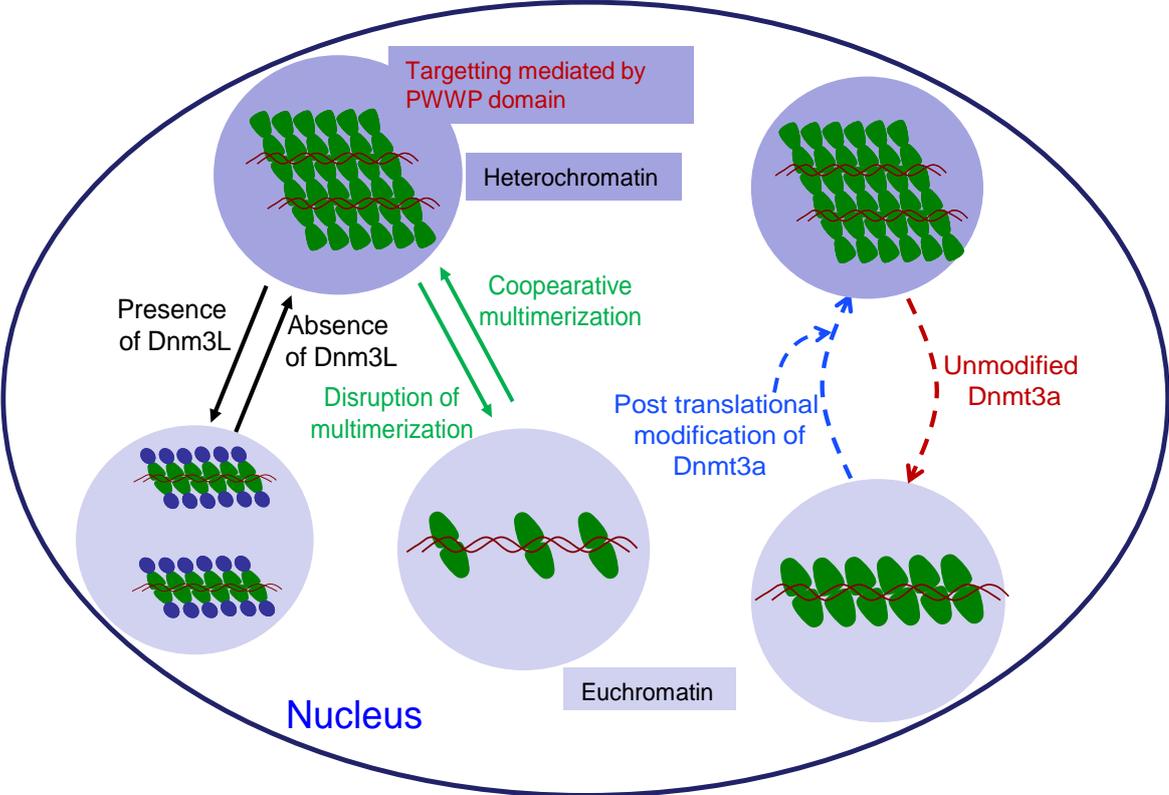


Fig 10. Compartmentalization of Dnmt3a: Dynamics of Dnmt3a localization is regulated by Dnmt3L. Presence of Dnmt3L prevents the multiple DNA binding property of Dnmt3a and releases it from the heterochromatin. A cooperative multimerization of Dnmt3a is required to localize in dense heterochromatic region. Post translational modification of Dnmt3a targets the enzyme to the heterochromatin, whereas unmodified form of Dnmt3a localizes in euchromatin.

### 3.4 Searching for new inhibitors of Dnmt3a

Recent studies indicated that aberrant DNA methylation and misregulation of DNA methyltransferases in mammals contributes to carcinogenesis [Sharma *et al.*, 2010; Robertson and wolffe, 2000]. Hypermethylation could be caused by aberrant expression of DNA methyltransferase or abnormal activity of DNA methyltransferases. One approach to revert the hypermethylation in cells, could be to inhibit the activity of Dnmts by using specific Dnmts inhibitors, which could be further used in epigenetic therapies.

#### 3.4.1 Developing a high-throughput screening assay

To identify potential inhibitors of Dnmt3a, we were involved in a collaboration to develop a novel high-throughput Dnmts screening assay, which is a quick and reliable *in vitro* test that uses fluorescence. This assay allows to study many methyltransferases and to screen a large number chemical compounds. Initially, we have used the Dnmt3a-C/Dnmt3L-C tetrameric complex with 114 flavones and flavanone derivatives to screen for inhibition activity of these compounds on Dnmt3a-C/Dnmt3L-C activity [Ceccaldi *et al.*, 2011]. From our assay, we have identified 2 molecules of 3-nitroflavones and 20 molecules of 3-chloro-3-nitroflavanones, which showed an inhibition on the Dnmt3a-C/Dnmt3L-C activity. Most of the compounds showed IC<sub>50</sub> values in sub-micromolar (1-9  $\mu$ M) range, whereas reference compounds showed 315  $\mu$ M for RG108 and 500  $\mu$ M for genistein [Ceccaldi *et al.*, 2011]. In addition to *in vitro* activity of these identified compounds, the 3-chloro-3-nitroflavanone derivatives showed a developmental effect on zebrafish embryo development similar as the reference compound 5-azacytidine [Ceccaldi *et al.*, 2011].

### 3.4.2 Inhibition of Dnmt3a by black tea and coffee polyphenols

Epi-gallocatechin gallate (EGCG) is the main polyphenolic constituent of green tea and 5-caffeoyl quinic acid is the main phenolic constituent of the green coffee, both has been shown to inhibit the activity of Dnmt1 [Lee *et al.*, 2005; Lee *et al.*, 2006]. Addition of EGCG to cancerous cells leads to decreases the cellular DNA methylation level and reactivated silenced tumor suppressor genes [Nandakumar *et al.*, 2011]. We have screened the inhibitory activity of these and other chlorogenic compounds purified from the black tea compounds on Dnmt3a-C activity. The inhibitory activity N1, N3, N4, N5, N6, N7, N8 and other derivatives were screened for inhibition of Dnmt3a-C (Fig. 11). The compounds N1, N4, N3 and N5 showed a weak inhibition on Dnmt3a-C activity. The compound N6 showed good inhibition on Dnmt3a activity with an IC<sub>50</sub> value at 44 μM, this was the best inhibitor we identified in this series of nutritional compounds. The thearubigin derivatives (N7 & N8) which are structurally similar to theaflavin (N6) also showed a good inhibition with an IC<sub>50</sub> value of 40 μM for N7 and 28 μM for N8. It had been shown previously that EGCG (N1) inhibits the Dnmt1 and decreases the cellular DNA methylation [Lee, 2005; Lee, 2006], here first time we are reporting that the derivatives of EGCG (N1) theaflavin also inhibit Dnmt3a activity.

Fig 11. Structures of polyphenolic compounds: Structures of epigallocatechin EGCG (N1), theaflavin-3-gallate (N3), epigallocatechin (N4), theaflavin 3'-gallate (N5), Theaflavin 3, 3'-digallate (N6). The N6 compound shows significant inhibition on activity of *de novo* DNA methyltransferase 3a (Dnmt3a).

To evaluate the possible biological significance of the IC<sub>50</sub> values of Dnmt3a inhibition, we have inspected available pharmacokinetic data of theaflavins. Recently, Mulder and co-workers reported theaflavin concentrations of 4.2 µg l<sup>-1</sup> in urine 2 h after consumption of 1 cup of black tea containing 8.8 mg total theaflavins [Vermeer *et al.*, 2008]. Another group reported a concentration of theaflavin in colon, small intestine, prostate and liver tissue of 2 µM, whereas the theaflavins N3, N5 and N6 investigated here showed roughly 1 µM, half this value after consumption of one cup of black tea [Henning *et al.*, 2006]. All these physiological concentration

are in the range of the observed IC<sub>50</sub> values suggesting a possible effect *in vivo*. In conclusion, we have shown that the black tea polyphenols, in particular theaflavin 3, 3'-digallate N6 and thearubigin inhibit Dnmt3a activity with physiologically and nutritionally relevant IC<sub>50</sub> values. Therefore, we identified a novel biological target for poly phenols present in black tea that is able to rationalize both anti-carcinogenic activity and mental health performance related with the activity of black tea.

### 3.4.3 Synthesis of new DNMT inhibitor derivatives of procainamide

Procainamide was used as antiarrhythmic drug and it had been shown to induce significant demethylation of genomic DNA in human cancer cells [Lee *et al.*, 2005; Cornacchia *et al.*, 1988]. These compounds are known to bind to CpG sequences in DNA and thereby disturb the DNA methylation. Our collaborators used procainamide as DNA binder to guide a competitive inhibitor of DNMTs to CpG rich regions on genome, such that one could increase the local concentration of the inhibitor at CpG sites. This strategy had been successfully used to target the action of DNA topoisomerases I and II inhibitors to specific DNA sites [Arimondo PB *et al.*, 2001, Duca *et al.*, 2006; Stierle *et al.*, 2008]. Our collaborator had generated a conjugated chemical molecule in which procainamide was attached to the Dnmt3a inhibitor RG108 [Halby *et al.*, 2011, Submitted]. This conjugated molecule was functional without affecting the activity of inhibitors. A series of procainamide–RG108 (N-phthalimidoyl protected tryptophan) conjugated derivatives was generated, among these molecules, derivatives 12 and 13 showed a strong inhibition of Dnmt3a/Dnmt3L activity with IC<sub>50</sub> values of 9  $\mu$ M and 11  $\mu$ M, respectively (Fig 12).

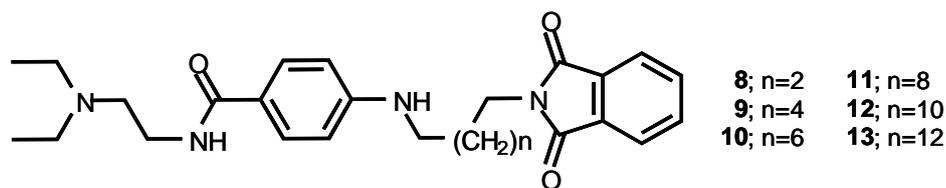


Fig 12. The conjugated derivatives of procainamide: Procainamide–RG108 (N-phthalimidoyl protected tryptophan) conjugated derivatives were generated with modification at position n. The promising compounds are 12 and 13.

The inhibition by these molecules was very selective for the Dnmt3a/Dnmt3L complex. It did not show any inhibition on bacterial methyltransferase (EcoDam) and other mammalian methyltransferase (G9a) [Halby *et al.*, 2011, Submitted]. In conclusion, procainamide-RG108 conjugated derivatives showed a strong inhibition of Dnmt3a/Dnmt3L with high selectivity, which illustrates that this novel approach of DNA methyltransferase inhibition is promising.

## 4. Conclusion

In conclusion, our data discover four novel mechanisms of Dnmt3a regulation. First, multimerization of the Dnmt3a using its alternative interfaces, such that it binds with many DNA molecules oriented in parallel. Disruption of the interfaces led to the loss of heterochromatic localization in cells, which suggests that the multimerization has a role in localization of Dnmt3a. The second mechanism is that interaction of the Dnmt3L with Dnmt3a prevents this multimerization and multiple DNA binding. The regulator protein Dnmt3L also releases the Dnmt3a from heterochromatin and makes it available to methylate euchromatic DNA. This is the first evidence showing that Dnmt3L regulates the compartmentalization of the Dnmt3a proteins within the nucleus [Jurkowska *et al.*, 2011]. The third mechanism is the cooperative multimerization of the Dnmt3a-C on DNA that leads to the formation of nucleoprotein filaments *in vitro* and generates 8-10 bp periodic methylation patterns on DNA substrates. Such cooperative multimerization of Dnmt3a could be important for the generation of methylation pattern on imprinted genes. The fourth mechanism is that Dnmt3a is regulated by post translational modifications, which modulate its activity and the localization of the enzyme in the cell. The post translational modification of Dnmt3a targets the enzyme to the heterochromatin. Why is the regulation of Dnmt3a activity very important? The answers would be that the DNA methylation and other epigenetic signals determine the cell's identity as well as normal development of cell. Regulation of Dnmt3a activity is necessary to avoid over methylation of genome. Misregulation of Dnmt3a activity might lead to aberrant DNA methylation which eventually causes cancer. In conclusion, regulation of Dnmt3a activity helps to write a precise DNA methylation pattern on the genome of mammals.

## 5. References

- Appanah R, Dickerson DR, Goyal P, Groudine M, Lorincz MC (2007). An unmethylated 3' promoter-proximal region is required for efficient transcription initiation. *PLoS Genet.* 3(2), e27.
- Arimondo PB, Bailly C, Bourtouline AS, Moreau P, Prudhomme M, Sun JS, Garestier T, Hélène C (2001). Triple helix-forming oligonucleotides conjugated to indolocarbazole poisons direct topoisomerase I-mediated DNA cleavage to a specific site. *Bioconjug Chem.* 12(4), 501-9.
- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007). High-resolution profiling of histone methylations in the human genome. *Cell.* 129(4), 823-37.
- Baylin SB (2005). DNA methylation and gene silencing in cancer. *Nat Clin Pract Oncol.* 2 Suppl 1, S4-11.
- Bell CE (2005). Structure and mechanism of Escherichia coli RecA ATPase. *Mol Microbiol.* 58(2), 358-66.
- Bender CM, Pao MM, Jones PA (1998). Inhibition of DNA methylation by 5-aza-2'-deoxycytidine suppresses the growth of human tumor cell lines. *Cancer Res.* 58(1), 95-101.
- Black JC, Whetstone JR (2011). Chromatin landscape: methylation beyond transcription. *Epigenetics.* 6(1), 9-15.
- Boon K, Osorio EC, Greenhut SF, Schaefer CF, Shoemaker J, Polyak K, Morin PJ, Buetow KH, Strausberg RL, De Souza SJ, Riggins GJ (2002). An anatomy of normal and malignant gene expression. *Proc Natl Acad Sci U S A.* 99(17), 11287-92.
- Bourc'his D, Bestor TH (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature.* 431(7004), 96-9.
- Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH (2001). Dnmt3L and the establishment of maternal genomic imprints. *Science.* 294(5551), 2536-9.
- Bradshaw RA, Brickey WW, Walker KW (1998). N-terminal processing: the methionine aminopeptidase and N $\alpha$ -acetyl transferase families. *Trends Biochem. Sci.* 23, 263-267.
- Ceccaldi A, Rajavelu A, Champion C, Rampon C, Jurkowska R, Jankevicius G, Sénamaud-Beaufort C, Ponger L, Gagey N, Ali HD, Tost J, Vríz S, Ros S, Dauzonne D, Jeltsch A, Guianvarc'h D, Arimondo PB (2011). C5-DNA methyltransferase inhibitors: from screening to effects on zebrafish embryo development. *Chembiochem.* 12(9), 1337-45.

Ceccaldi A, Rajavelu A, Sénamaud-Beauforta C, Testaa N, Ali-Dalia H, Jurkowska R, Maulay-Baillye C, Amande C, Jeltsch A, Guianvarc'h D and Arimondo PB. Identification of novel inhibitors of the Dnmt3A/3L Catalytic Complex by Screening of a Chemical Library [Submitted for publication].

Champion C, Guianvarc'h D, Sénamaud-Beaufort C, Jurkowska RZ, Jeltsch A, Ponger L, Arimondo PB, Guieysse-Peugeot AL (2010). Mechanistic insights on the inhibition of c5 DNA methyltransferases by Zebularine. *PLoS One*. 5(8), e12388.

Chavez-Blanco A, Perez-Plasencia C, Perez-Cardenas E, Carrasco-Legleu C, Rangel-Lopez E, Segura-Pacheco B, Taja-Chayeb L, Trejo-Becerril C, Gonzalez-Fierro A, Candelaria M, Cabrera G, Duenas-Gonzalez A (2006). Antineoplastic effects of the DNA methylation inhibitor hydralazine and the histone deacetylase inhibitor valproic acid in cancer cell lines. *Cancer Cell Int*. 6, 2.

Chedin F, Lieber MR, Hsieh CL (2002). The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proc Natl Acad Sci U S A*. 99(26), 16916-21.

Chen T, Li E (2006). Establishment and maintenance of DNA methylation patterns in mammals. *Curr Top Microbiol Immunol*. 301, 179-201.

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. *Mol Cell Biol*. 24(20), 9048-58.

Chen T, Ueda Y, Xie S and Li E (2002). A novel Dnmt3a isoforms produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. *J Biol Chem*, 277, 38746-54

Cheng X, Blumenthal RM (2008). Mammalian DNA methyltransferases: a structural perspective. *Structure*. 6(3), 341-50.

Cheng X, Blumenthal RM (2010). Coordinated chromatin control: structural and functional linkage of DNA and histone methylation. *Biochemistry*. 49(14), 2999-3008.

Chodavarapu RK, Feng S, Bernatavichute YV, Chen PY, Stroud H, Yu Y, Hetzel JA, Kuo F, Kim J, Cokus SJ, Casero D, Bernal M, Huijser P, Clark AT, Krämer U, Merchant SS, Zhang X, Jacobsen SE, Pellegrini M (2010). Relationship between nucleosome positioning and DNA methylation. *Nature*. 466(7304), 388-92.

Chow JC, Brown CJ (2003). Forming facultative heterochromatin: silencing of an X chromosome in mammalian females. *Cell Mol Life Sci*. 60(12), 2586-603.

Christman JK (2002). 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene* 21, 5483–5495.

Chuang JC, Warner SL, Vollmer D, Vankayalapati H, Redkar S, Bearss DJ, Qiu X, Yoo CB, Jones PA (2010). S110, a 5-Aza-2'-deoxycytidine-containing dinucleotide, is an effective DNA methylation inhibitor in vivo and can reduce tumor growth. *Mol Cancer Ther.* 9(5), 1443-50.

Clark SJ, Melki J (2002). DNA methylation and gene silencing in cancer: which is the guilty party? *Oncogene.* 21(35), 5380-7.

Cornacchia E, Golbus J, Maybaum J, Strahler J, Hanash S, Richardson B (1988). Hydralazine and procainamide inhibit T cell DNA methylation and induce autoreactivity. *J Immunol.* 140(7), 2197-200.

Dhayalan A, Rajavelu A, Rathert P, Tamas R, Jurkowska RZ, Ragozin S, Jeltsch A (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J Biol Chem.* 285(34), 26114-20.

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(18), 17986-91.

Duca M, Guianvarc'h D, Oussedik K, Halby L, Garbesi A, Dauzonne D, Monneret C, Osheroff N, Giovannangeli C, Arimondo PB (2006). Molecular basis of the targeting of topoisomerase II-mediated DNA cleavage by VP16 derivatives conjugated to triplex-forming oligonucleotides. *Nucleic Acids Res.* 34(6), 1900-11.

Deplus R, Blanchon L, Rajavelu A, Boukaba H, Defrance M, Luciani J, Dedeurwaerder S, Brinkman AB, Simmer F, Müller F, Berdasco M, Putmans P, Calonne E, Litchfield DW, Launoit Y, Jurkowski TP, Stunnenberg HG, Bock C, Fraga MF, Esteller M, Jeltsch A and Fuks F. (2011) [Revision].

Ehrlich M (2009). DNA hypomethylation in cancer cells. *Epigenomics.* 1(2), 239-259.

Fatemi M, Hermann A, Gowher H, Jeltsch A (2002). Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. *Eur J Biochem.* 269(20), 4981-4.

Feinberg AP, Vogelstein B (1983). Hypomethylation of ras oncogenes in primary human cancers. *Biochem Biophys Res Commun.* 111(1), 47-54.

Felsenfeld G, Groudine M (2003). Controlling the double helix. *Nature.* 421(6921), 448-53.

Fenaux P (2005). Inhibitors of DNA methylation: beyond myelodysplastic syndromes. *Nat Clin Pract Oncol.* Suppl 1, S36-44.

Feng J, Zhou Y, Campbell SL, Le T, Li E, Sweatt JD, Silva AJ, Fan G (2010). Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nat Neurosci.* 13(4), 423-30.

Fitzpatrick DR, Wilson CB (2003). Methylation and demethylation in the regulation of genes, cells, and responses in the immune system. *Clin Immunol.* 109(1), 37-45.

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(10), 2536-44.

Fuks F, Hurd PJ, Wolf D, Nan X, Bird AP, Kouzarides T (2003). The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J Biol Chem.* 278(6), 4035-40.

Gangaraju VK, Lin H (2009). MicroRNAs: key regulators of stem cells. *Nat Rev Mol Cell Biol.* 10(2), 116-25.

Ge YZ, Pu MT, Gowher H, Wu HP, Ding JP, Jeltsch A, Xu GL (2004). Chromatin targeting of de novo DNA methyltransferases by the PWWP domain. *J Biol Chem.* 279(24), 25447-54.

Goll MG, Kirpekar F, Maggert KA, Yoder JA, Hsieh CL, Zhang X, Golic KG, Jacobsen SE, Bestor TH (2006). Methylation of tRNA<sup>Asp</sup> by the DNA methyltransferase homolog Dnmt2. *Science.* 311, 395-398.

Gopalakrishnan S, Van Emburgh BO, Shan J, Su Z, Fields CR, Vieweg J, Hamazaki T, Schwartz PH, Terada N, Robertson KD (2009). A novel DNMT3B splice variant expressed in tumor and pluripotent cells modulates genomic DNA methylation patterns and displays altered DNA binding. *Mol Cancer Res.* 7(10), 1622-34.

Gore SD, Baylin S, Sugar E, Carraway H, Miller CB, Carducci M, Grever M, Galm O, Dausers T, Karp JE, Rudek MA, Zhao M, Smith BD, Manning J, Jiemjit A, Dover G, Mays A, Zwiebel J, Murgo A, Weng LJ, Herman JG (2006). Combined DNA methyltransferase and histone deacetylase inhibition in the treatment of myeloid neoplasms. *Cancer Res.* 66(12), 6361-9.

Gottschling DE (2004). Epigenetics--from phenomenon to field. *Cold Spring Harb Symp Quant Biol.* 69, 507-19.

Gowher H, Liebert K, Hermann A, Xu G, Jeltsch A (2005). Mechanism of stimulation of catalytic activity of Dnmt3a and Dnmt3b DNA- (cytosine-C5)-methyltransferases by Dnmt3L. *J Biol Chem.* 280, 13341-13348.

Gowher H, Loutchanwoot P, Vorobjeva O, Handa V, Jurkowska RZ, Jurkowski TP, Jeltsch A (2006). Mutational analysis of the catalytic domain of the murine Dnmt3a DNA-(cytosine C5)-methyltransferase. *J Mol Biol.* 357(3):928-41.

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(29), 9899-904.

Gowher H, Jeltsch A (2002). Molecular enzymology of the catalytic domains of the Dnmt3a and Dnmt3b DNA methyltransferases. *J Biol Chem.* 277(23), 20409-14.

Goyal R, Reinhardt R, Jeltsch A (2006). Accuracy of DNA methylation pattern preservation by the Dnmt1 methyltransferase. *Nucleic Acids Res.* 34(4), 1182-8.

Gravina GL, Festuccia C, Marampon F, Popov VM, Pestell RG, Zani BM, Tombolini V (2010). Biological rationale for the use of DNA methyltransferase inhibitors as new strategy for modulation of tumor response to chemotherapy and radiation. *Mol Cancer.* 9, 305.

Greger V, Passarge E, Höpping W, Messmer E, Horsthemke B (1989). Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. *Hum Genet.* 83(2), 155-8.

Halby L, Sénamaud-Beaufort C, Ajjan S, Ceccaldi A, Drujon T, Rajavelu A, Champion C, Jurkowska R, Lequin O, Nelson WG, Jeltsch A, Guy A, Guianvarc'h D, Ferroud C and Arimondo PB. Rapid synthesis of new DNMT inhibitors derivatives of Procainamide [Submitted for publication].

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 U S A.* 96(25), 14412-7.

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

Henning SM, Aronson W, Niu Y, Conde F, Lee NH, Seeram NP, Lee RP, Lu J, Harris DM, Moro A, Hong J, Pak-Shan L, Barnard RJ, Ziaee HG, Csathy G, Go VL, Wang H, Heber D (2006). Tea polyphenols and theaflavins are present in prostate tissue of humans and mice after green and black tea consumption. *J Nutr.* 136(7), 1839-43.

Hermann A, Gowher H, Jeltsch A (2004). Biochemistry and biology of mammalian DNA methyltransferases. *Cell Mol Life Sci.* 61(19-20), 2571-87.

Hodges E, Smith AD, Kendall J, Xuan Z, Ravi K, Rooks M, Zhang MQ, Ye K, Bhattacharjee A, Brizuela L, McCombie WR, Wigler M, Hannon GJ, Hicks JB (2009). High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. *Genome Res.* 19(9), 1593-605.

Holmes R, Soloway PD (2006). Regulation of imprinted DNA methylation. *Cytogenet Genome Res.* 113(1-4), 122-9.

Howell CY, Bestor TH, Ding F, Latham KE, Mertineit C, Trasler JM, Chaillet JR (2001). Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. *Cell.* 104(6), 829-38.

Hsieh CL (1999). Evidence that protein binding specifies sites of DNA demethylation. *Mol Cell Biol.* 19(1), 46-56.

- Hu JL, Zhou BO, Zhang RR, Zhang KL, Zhou JQ, Xu GL (2009). The N-terminus of histone H3 is required for de novo DNA methylation in chromatin. *Proc Natl Acad Sci U S A*.106(52), 22187-92.
- Ioshikhes IP, Zhang MQ (2000). Large-scale human promoter mapping using CpG islands. *Nat Genet*. 26(1), 61-3.
- Jaenisch R, Bird A (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet*. 245-54.
- Jeltsch A (2002) Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. *Chembiochem*. Apr 2;3(4):274-93.
- Jeltsch A, Nellen W, Lyko F (2006). Two substrates are better than one: dual specificities for Dnmt2 methyltransferases. *Trends Biochem Sci*. 31:306–308.
- Jeong S, Liang G, Sharma S, Lin JC, Choi SH, Han H, Yoo CB, Egger G, Yang AS, Jones PA (2009). Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. *Mol Cell Biol*. 29(19), 5366-76
- Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X (2007). Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature*. 449, 248–251.
- Jones PA, Baylin SB (2002). The fundamental role of epigenetic events in cancer. *Nat Rev Genet*. 3(6), 415-28.
- Jones PA, Baylin SB (2007) The epigenomics of cancer. *Cell*. 128(4), 683-92.
- Jones PA, Liang G (2009). Rethinking how DNA methylation patterns are maintained. *Nat Rev Genet*. 10(11), 805-11.
- Jones PA, Taylor SM (1980). Cellular differentiation, cytidine analogs and DNA methylation. *Cell*. 20(1), 85-93.
- Jurkowska RZ, Anspach N, Urbanke C, Jia D, Reinhardt R, Nellen W, Cheng X, Jeltsch A (2008). Formation of nucleoprotein filaments by mammalian DNA methyltransferase Dnmt3a in complex with regulator Dnmt3L. *Nucleic Acids Res*. 36(21), 6656-63.
- Jurkowska RZ, Jurkowski TP, Jeltsch A (2011). Structure and function of mammalian DNA methyltransferases. *Chembiochem*. 12(2), 206-22.
- Jurkowska RZ, Rajavelu A, Anspach N, Urbanke C, Jankevicius G, Ragozin S, Nellen W, Jeltsch A (2011). Oligomerization and Binding of the Dnmt3a DNA Methyltransferase to Parallel DNA Molecules: HETEROCHROMATIC LOCALIZATION AND ROLE OF Dnmt3L. *J Biol Chem*. 286(27), 24200-24207.
- Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, Sasaki H (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature*. 429(6994), 900-3.

- Kareta MS, Botello ZM, Ennis JJ, Chou C, Chédin F (2006). Reconstitution and mechanism of the stimulation of de novo methylation by human DNMT3L. *J Biol Chem.* 281(36), 25893-902.
- Kim TH, Barrera LO, Zheng M, Qu C, Singer MA, Richmond TA, Wu Y, Green RD, Ren B (2005). A high-resolution map of active promoters in the human genome. *Nature.* 436(7052), 876-80.
- Kim JK, Samaranyake M, Pradhan S (2009). Epigenetic mechanisms in mammals. *Cell Mol Life Sci.* 66(4), 596-612.
- Klimasauskas S, Kumar S, Roberts RJ, Cheng X (1994). HhaI methyltransferase flips its target base out of the DNA helix. *Cell.* 76(2), 357-69.
- Klose RJ, Zhang Y (2007). Regulation of histone methylation by demethylination and demethylation. *Nat Rev Mol Cell Biol.* 8(4), 307-18.
- Kobayashi H, Suda C, Abe T, Kohara Y, Ikemura T, Sasaki H (2006). Bisulfite sequencing and dinucleotide content analysis of 15 imprinted mouse differentially methylated regions (DMRs): paternally methylated DMRs contain less CpGs than maternally methylated DMRs. *Cytogenet Genome Res.* 113(1-4), 130-7.
- Kouzarides T (2007). Chromatin modifications and their function. *Cell.* 23, 128(4), 693-705.
- Kriaucionis S, Bird A (2003). DNA methylation and Rett syndrome. *Hum Mol Genet.* 12, R221-7.
- La Salle S, Trasler JM (2006). Dynamic expression of DNMT3a and DNMT3b isoforms during male germ cell development in the mouse. *Dev Biol.* 1;296(1), 71-82.
- Laird PW (2005). Cancer epigenetics. *Hum Mol Genet.* 14, R65-76.
- Larschan E, Alekseyenko AA, Gortchakov AA, Peng S, Li B, Yang P, Workman JL, Park PJ, Kuroda MI (2007). MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. *Mol Cell.* 28(1), 121-33.
- Lee WJ, Shim JY, Zhu BT (2005). Mechanisms for the inhibition of DNA methyltransferases by tea catechins and bioflavonoids. *Mol Pharmacol.* 68(4), 1018-30.
- Lee WJ, Shim JY, Zhu BT (2005). Mechanisms for the inhibition of DNA methyltransferases by tea catechins and bioflavonoids. *Mol Pharmacol.* 68(4), 1018-30.
- Lee WJ, Zhu BT (2006). Inhibition of DNA methylation by caffeic acid and chlorogenic acid, two common catechol-containing coffee polyphenols.. *Carcinogenesis.* 27(2), 269-77.

Lee BH, Yegnasubramanian S, Lin X, Nelson WG (2005). Procainamide is a specific inhibitor of DNA methyltransferase 1. *J Biol Chem.* 280(49), 40749-56.

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(14), 1192-200.

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(10):3195-205.

Lengauer C, Kinzler KW, Vogelstein B (1999). DNA methylation and genetic instability in colorectal cancer cells. *Proc Natl Acad Sci U S A.* 94(6), 2545-50.

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

Levenson JM, Sweatt JD (2005). Epigenetic mechanisms in memory formation *Nat Rev Neurosci.* 6, 108-18.

Li BZ, Huang Z, Cui QY, Song XH, Du L, Jeltsch A, Chen P, Li G, Li E, Xu GL (2011). Histone tails regulate DNA methylation by allosterically activating *de novo* methyltransferase. *Cell Res.* 1-10.

Li E, Bestor TH, Jaenisch R (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell.* 69(6), 915-26.

Li G, Reinberg D (2011). Chromatin higher-order structures and gene regulation. *Curr Opin Genet Dev.* 21, 175-86.

Litman SE, 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(11), 1176-83.

Loshikhes IP and Zhang MO (2000). Large-scale human promoter mapping using CpG islands. *Nature Genet.* 26, 61-63.

Margot JB, Ehrenhofer-Murray AE, Leonhardt H. (2003) Interactions within the mammalian DNA methyltransferase family. *BMC Molecular Biology.* 4, 7.

Margueron R, Reinberg D (2010). Chromatin structure and the inheritance of epigenetic information.. *Nat Rev Genet.* 11(4), 285-96.

Maurer-Stroh S, Dickens NJ, Hughes-Davies L, Kouzarides T, Eisenhaber F, Ponting CP (2003). The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends Biochem Sci.* 28(2), 69-74.

Mochizuki K (2010). RNA-directed epigenetic regulation of DNA rearrangements. *Essays Biochem.* 48(1), 89-100.

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.

Müller CI, Rüter B, Koeffler HP, Lübbert M (2006). DNA hypermethylation of myeloid cells, a novel therapeutic target in MDS and AML. *Curr Pharm Biotechnol.* 7(5), 315-21.

Nandakumar V, Vaid M, Tollefsbol TO, Katiyar SK (2011). Aberrant DNA hypermethylation patterns lead to transcriptional silencing of tumor suppressor genes in UVB-exposed skin and UVB-induced skin tumors of mice. *Carcinogenesis.* 32(4), 597-604.

Ohki I, Shimotake N, Fujita N, Jee J, Ikegami T, Nakao M, Shirakawa M (2001). Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. *Cell.* 105(4), 487-97.

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.

Okano M, Xie S, Li E (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat Genet.* 19(3), 219-20.

Olsen JV, Blagoev B, Gnäd F, Macek B, Kumar C, Mortensen P, Mann M (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell.*127(3), 635-48.

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(7154), 714-7.

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(11), 1235-1241.

Pawson T (2002). Regulation and targets of receptor tyrosine kinases. *Eur J Cancer.* 38 Suppl 5, S3-10.

Purdy MM, Holz-Schietinger C, Reich NO (2010). Identification of a second DNA binding site in human DNA methyltransferase 3A by substrate inhibition and domain deletion. *Arch Biochem Biophys.* 498(1), 13-22.

Rajavelu A, Tulyasheva Z, Jaiswal R, Jeltsch A, Kuhnert N (2011). The inhibition of the mammalian DNA methyltransferase 3a (Dnmt3a) by dietary black tea and coffee polyphenols. *BMC Biochem.* 12, 16.

Rajavelu A, Jurkowska R, Fritz J, Jeltsch A (2011). Function and disruption of DNA Methyltransferase 3a cooperative DNA binding and nucleoprotein filament formation. *Nucleic Acid Res*, 2011, in press

Reik W, Dean W, Walter J (2011). Epigenetic reprogramming in mammalian development. *Science*. 293(5532), 1089-93.

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

Robertson KD and Wolffe AP (2000). DNA methylation in health and disease. *Nat. Rev. Genet.* 1, 11-19.

Robertson KD, Uzvolgyi E, Liang G, Talmadge C, Sumegi J, Gonzales FA, Jones PA (1999). The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors.. *Nucleic Acids Res*. 27(11), 2291-8.

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(13), 4301-12.

Schirmacher E, Beck C, Brueckner B, Schmitges F, Siedlecki P, Bartenstein P, Lyko F, Schirmacher R (2006). Synthesis and in vitro evaluation of biotinylated RG108: a high affinity compound for studying binding interactions with human DNA methyltransferases. *Bioconjug Chem*. 17(2), 261-6.

Seger R, Seger D, Reszka AA, Munar ES, Eldar-Finkelman H, Dobrowolska G, Jensen AM, Campbell JS, Fischer EH, Krebs EG (1994). Overexpression of mitogen-activated protein kinase kinase (MAPKK) and its mutants in NIH 3T3 cells. Evidence that MAPKK involvement in cellular proliferation is regulated by phosphorylation of serine residues in its kinase subdomains VII and VIII. *J Biol Chem*. 269(41), 25699-709.

Seo GJ, Kim SE, Lee YM, Lee JW, Lee JR, Hahn MJ, Kim ST (2003). Determination of substrate specificity and putative substrates of Chk2 kinase. *Biochem Biophys Res Commun*. 304(2), 339-43.

Sharma S, De Carvalho DD, Jeong S, Jones PA, Liang G (2011). Nucleosomes containing methylated DNA stabilize DNA methyltransferases 3A/3B and ensure faithful epigenetic inheritance. *PLoS Genet*. 7(2), e1001286.

Sharma S, Kelly TK, Jones PA (2010). Epigenetics in cancer. *Carcinogenesis*. 31(1), 27-36.

Shikauchi Y, Saiura A, Kubo T, Niwa Y, Yamamoto J, Murase Y, Yoshikawa H (2009). SALL3 interacts with DNMT3A and shows the ability to inhibit CpG island methylation in hepatocellular carcinoma. *Mol Cell Biol*. 29(7), 1944-58.

Sorm F, Pískala A, Cihák A, Veselý J (1964). 5-Azacytidine, a new, highly effective cancerostatic. *Experientia*. 20(4), 202-3.

Spada F, Rothbauer U, Zolghadr K, Schermelleh L, Leonhardt H (2006). Regulation of DNA methyltransferase 1. *Adv Enzyme Regul*. 46, 224-34.

Steensma DP, Baer MR, Slack JL, Buckstein R, Godley LA, Garcia-Manero G, Albitar M, Larsen JS, Arora S, Cullen MT, Kantarjian H (2009). Multicenter study of decitabine administered daily for 5 days every 4 weeks to adults with myelodysplastic syndromes: the alternative dosing for outpatient treatment (ADOPT) trial. *J Clin Oncol*. 27(23), 3842-8.

Stierlé V, Duca M, Halby L, Senamaud-Beaufort C, Capobianco ML, Laigle A, Jollès B, Arimondo PB (2008). Targeting MDR1 gene: synthesis and cellular study of modified daunomycin-triplex-forming oligonucleotide conjugates able to inhibit gene expression in resistant cell lines. *Mol Pharmacol*. 73(5), 1568-77.

Stresemann C, Brueckner B, Musch T, Stopper H, Lyko F (2006). Functional diversity of DNA methyltransferase inhibitors in human cancer cell lines. *Cancer Res*. 66(5), 2794-800.

Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol*. 14(11), 1025-40.

Taylor SM, Jones PA (1979). Multiple new phenotypes induced in 10T1/2 and 3T3 cells treated with 5-azacytidine. *Cell*. 17(4), 771-9.

Ulrich-Pur H, Kornek GV, Fiebiger W, Schüll B, Raderer M, Scheithauer W (2001). Treatment of advanced hepatocellular carcinoma with biweekly high-dose gemcitabine. *Oncology*. 60(4), 313-5.

Vakoc CR, Sachdeva MM, Wang H, Blobel GA (2006). Profile of histone lysine methylation across transcribed mammalian chromatin. *Mol Cell Biol*. 26(24), 9185-95.

Vermeer MA, Mulder TP, Molhuizen HO (2008). Theaflavins from black tea, especially theaflavin-3-gallate, reduce the incorporation of cholesterol into mixed micelles. *J Agric Food Chem*. 56(24), 12031-6.

Villar-Garea A, Fraga MF, Espada J, Esteller M (2003). Procaine is a DNA-demethylating agent with growth-inhibitory effects in human cancer cells. *Cancer Res*. 63(16), 4984-9.

Voigt P, Reinberg D (2011). Histone tails: ideal motifs for probing epigenetics through chemical biology approaches. *Chembiochem*. 12(2), 236-52.

Waddington CH (1942). "The epigenotype". *Endeavour* 1, 18–20.

- Watanabe D, Suetake I, Tada T, Tajima S (2002). Stage- and cell-specific expression of Dnmt3a and Dnmt3b during embryogenesis. *Mech Dev.* 118(1-2), 187-90.
- Weber M, Schübeler D (2007). Genomic patterns of DNA methylation: targets and function of an epigenetic mark. *Curr Opin Cell Biol.* 19(3), 273-80.
- Weisenberger DJ, Velicescu M, Preciado-Lopez MA, Gonzales FA, Tsai YC, Liang G, Jones PA (2002). Identification and characterization of alternatively spliced variants of DNA methyltransferase 3a in mammalian cells. *Gene.* 298, 91–99.
- Wu Ct, Morris JR (2001) Genes, genetics, and epigenetics: a correspondence. *Science.* 293(5532), 1103-5.
- Xie S, Wang Z, Okano M, Li Y, He WW, Okumura K, Li E (1999). Cloning, expression and chromosome locations of the human DNMT3 gene family. *Gene* 236, 87-95.
- 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(6758), 187-91.
- Yan Q, Cho E, Lockett S, Muegge K (2003). Association of Lsh, a regulator of DNA methylation, with pericentromeric heterochromatin is dependent on intact heterochromatin. *Mol Cell Biol.* 23(23), 8416-28.
- Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, Ecker JR (2006). Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell.* 126(6), 1189-201.
- Zhang Y, Rohde C, Tierling S, Jurkowski TP, Bock C, Santacruz D, Ragozin S, Reinhardt R, Groth M, Walter J, Jeltsch A (2009). DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution. *PLoS Genet.* 5(3), e1000438.
- Zhang Y, Jurkowska R, Soeroes S, Rajavelu A, Dhayalan A, Bock I, Rathert P, Brandt O, Reinhardt R, Fischle W, Jeltsch A (2010). Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res.* 38(13), 4246-53.

## 6. List of publications and author's contribution

1. **Rajavelu A**, Jurkowska R, Fritz J, Jeltsch A. Function and disruption of DNA Methyltransferase 3a cooperative DNA binding and nucleoprotein filament formation. *Nucleic Acids Res* (2011) in press.

**Rajavelu A**, contributed in designing the study and performed all the experiments. **A.R** was involved in data interpretation and manuscript preparation.

2. Jurkowska RZ<sup>#</sup>, **Rajavelu A**<sup>#</sup>, Anspach N, Urbanke C, Jankevicius G, Ragozin S, Nellen W, Jeltsch A (2011). Oligomerization and Binding of the Dnmt3a DNA Methyltransferase to Parallel DNA Molecules: Heterochromatic localization and role of Dnmt3L. *J Biol Chem.* 286 (27); 24200-7.

<sup>#</sup> These authors contributed equally to this study.

**Rajavelu A** has contributed in designing the study and performed the following experiments, mutant generation, protein purification, protein precipitation assay, FRET assay, localization and immunofluorescence microscopy. **A.R** also participated in interpretation of data and manuscript preparation.

3. **Rajavelu A**, Tulyasheva Z, Jaiswal R, Jeltsch A<sup>\*</sup>, Kuhnert N<sup>\*</sup> (2011). The inhibition of the mammalian DNA methyltransferase 3a (Dnmt3a) by dietary black tea and coffee polyphenols. *BMC Biochem.* 12; 16. **Highly accessed**

**Rajavelu A**, contributed in designing the study and performed following experiments, DNMTs protein purification, inhibitors screening, enzyme kinetics. **A.R** also contributed to manuscript preparation. **A.R** Supervised ZT, who was involved in this project.

<sup>\*</sup> Corresponding authors

4. Ceccaldi A, **Rajavelu A**, Champion C, Rampon C, Jurkowska R, Jankevicius G, Sénamaud-Beaufort C, Ponger L, Gagey N, Dali Ali H, Tost J, Vríz S, Ros S, Dauzonne D, Jeltsch A, Guianvarc'h D, Arimondo PB (2011). C5-DNA Methyltransferase Inhibitors: From Screening to Effects on Zebrafish Embryo Development. *Chembiochem*. 14; 12 (9); 1337-45.  
**Rajavelu A**, purified the DNMTs and contributed to inhibitor screening and enzyme kinetics.
5. Dhayalan A, **Rajavelu A**, Rathert P, Tamas R, Jurkowska RZ, Ragozin S, Jeltsch A (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J Biol Chem*. 285(34); 26114-20.  
**Rajavelu A**, contributed in Dnmt3a2 protein purification, DNMT assay, Dnmt3a localization and immunofluorescence microscopy. **A.R** was involved in data analysis and data interpretation.
6. Zhang Y, Jurkowska R, Soeroes S, **Rajavelu A**, Dhayalan A, Bock I, Rathert P, Brandt O, Reinhardt R, Fischle W, Jeltsch A (2010). Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res*. 38(13); 4246-53.  
**Rajavelu A**, purified all DNMTs and involved in data analysis and data interpretation.

## Manuscripts submitted

1. Deplus R<sup>#</sup>, Blanchon L<sup>#</sup>, **Rajavelu A<sup>#</sup>**, Boukaba H<sup>#</sup>, Defrance M, Luciani J, Dedeurwaerder S, Brinkman AB, Simmer F, Müller F, Berdasco M, Putmans P, Calonne E, Litchfield DW, Launoit Y, Jurkowski TP, Stunnenberg HG, Bock C, Fraga MF, Esteller M, Jeltsch A<sup>\*</sup> and Fuks F<sup>\*</sup>.  
**Rajavelu A**, contributed to experiments involving kinase-related assays, performed DNMT enzymatic assays, immunofluorescence experiments, sample preparation for Mass spectrometry. **A.R** participated in designing the experiments and interpretation of the data.

# These authors equally contributed to this study

\* Corresponding authors

2. Halby L, Sénamaud-Beaufort C, Ajjan S, Ceccaldi A, Drujon T, **Rajavelu A**, Champion C, Jurkowska R, Lequin O, Nelson WG, Jeltsch A, Guy A, Guianvarc'h D, Ferroud C and Arimondo PB. Rapid synthesis of new DNMT inhibitors derivatives of Procainamide [[Submitted for publication](#)].  
**Rajavelu A**, performed DNMTs protein purification and contributed to inhibitor screening and enzyme kinetics.
  
3. Ceccaldi A, **Rajavelu A**, Sénamaud-Beauforta C, Testaa N, Ali-Dalia H, Jurkowska R, Maulay-Baillye C, Amande C, Jeltsch A, Guianvarc'h D and Arimondo PB. Identification of novel inhibitors of the Dnmt3A/3L Catalytic Complex by Screening of a Chemical Library [[Submitted for publication](#)].  
**Rajavelu A**, purified the DNMTs and contributed to inhibitor screening and enzyme kinetics.

## Annex-1

1. **Rajavelu A**, Jurkowska R, Fritz J, Jeltsch A (2011). Function and disruption of DNA Methyltransferase 3a cooperative DNA binding and nucleoprotein filament formation. *Nucleic Acids Res* (2011), in press.
2. Jurkowska RZ<sup>#</sup>, **Rajavelu A**<sup>#</sup> Anspach N, Urbanke C, Jankevicius G, Ragozin S, Nellen W, Jeltsch A (2011). Oligomerization and Binding of the Dnmt3a DNA Methyltransferase to Parallel DNA Molecules: Heterochromatic localization and role of Dnmt3L. *J Biol Chem.* 286 (27); 24200-7.
3. **Rajavelu A**, Tulyasheva Z, Jaiswal R, Jeltsch A<sup>\*</sup>, Kuhnert N<sup>\*</sup> (2011). The inhibition of the mammalian DNA methyltransferase 3a (Dnmt3a) by dietary black tea and coffee polyphenols. *BMC Biochem.* 12; 16. Highly accessed.
4. Ceccaldi A, **Rajavelu A**, Champion C, Rampon C, Jurkowska R, Jankevicius G, Sénamaud-Beaufort C, Ponger L, Gagey N, Dali Ali H, Tost J, Vríz S, Ros S, Dauzonne D, Jeltsch A, Guianvarc'h D, Arimondo PB (2011). C5-DNA Methyltransferase Inhibitors: From Screening to Effects on Zebrafish Embryo Development. *Chembiochem.* 14; 12 (9); 1337-45.
5. Dhayalan A, **Rajavelu A**, Rathert P, Tamas R, Jurkowska RZ, Ragozin S, Jeltsch A (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J Biol Chem.* 285(34); 26114-20.
6. Zhang Y, Jurkowska R, Soeroes S, **Rajavelu A**, Dhayalan A, Bock I, Rathert P, Brandt O, Reinhardt R, Fischle W, Jeltsch A (2010). Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res.* 38(13); 4246-53.

## Annex-2

Confidential, available only to examiners

1. Regulation of DNA methylation patterns mediated by post translational modification of DNMT3A.
2. Rapid synthesis of new DNMT inhibitors derivatives of Procainamide
3. Identification of novel inhibitors of the Dnmt3A/3L Catalytic Complex by Screening of a Chemical Library

# Annex-1

# Function and disruption of DNA Methyltransferase 3a cooperative DNA binding and nucleoprotein filament formation

Arumugam Rajavelu<sup>1</sup>, Renata Z. Jurkowska<sup>1</sup>, Jürgen Fritz<sup>2</sup> and Albert Jeltsch<sup>1,\*</sup>

<sup>1</sup>Biochemistry Laboratory and <sup>2</sup>Biophysics Laboratory, School of Engineering and Science, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany

Received June 24, 2011; Revised August 8, 2011; Accepted August 29, 2011

## ABSTRACT

The catalytic domain of Dnmt3a cooperatively multimerizes on DNA forming nucleoprotein filaments. Based on modeling, we identified the interface of Dnmt3a complexes binding next to each other on the DNA and disrupted it by charge reversal of critical residues. This prevented cooperative DNA binding and multimerization of Dnmt3a on the DNA, as shown by the loss of cooperative complex formation in electrophoretic mobility shift assay, the loss of cooperativity in DNA binding in solution, the loss of a characteristic 8- to 10-bp periodicity in DNA methylation and direct imaging of protein–DNA complexes by scanning force microscopy. Non-cooperative Dnmt3a-C variants bound DNA well and retained methylation activity, indicating that cooperative DNA binding and multimerization of Dnmt3a on the DNA are not required for activity. However, one non-cooperative variant showed reduced heterochromatic localization in mammalian cells. We propose two roles of Dnmt3a cooperative DNA binding in the cell: (i) either nucleofilament formation could be required for periodic DNA methylation or (ii) favorable interactions between Dnmt3a complexes may be needed for the tight packing of Dnmt3a at heterochromatic regions. The complex interface optimized for tight packing would then promote the cooperative binding of Dnmt3a to naked DNA *in vitro*.

## INTRODUCTION

One of the important epigenetic modifications of unicellular and multicellular organisms is DNA methylation. In mammals and other vertebrates, DNA methylation occurs

at the C5 position of cytosine bases, mostly at the cytosine residues followed by a guanine (CpG dinucleotides) (1–3). In mammals, DNA methylation has many functions including the control of cellular differentiation and development, the maintenance of chromosomal integrity, parental imprinting, the regulation of gene expression and X-chromosome inactivation. Hypermethylation of tumor-suppressor genes and genome-wide hypomethylation contribute to genomic instability and the development of cancer (4–6). DNA methylation is important for normal cellular differentiation and viability of mammals (3,7), as indicated by the finding that the deletion of any active DNA methyltransferase enzyme in mice is lethal (8,9). Abnormal methylation patterns of genomes are associated with several human diseases, such as Rett syndrome, ICF and Fragile X syndrome. Additionally, DNA methylation plays an important role in brain function (10,11) and changing of DNA methylation patterns is needed for cellular reprogramming (12).

The mammalian DNA methylation machinery consists of the DNA methyltransferase 1 (Dnmt1), which has specificity for hemimethylated DNA and the members of the DNA methyltransferases 3 family, Dnmt3a and Dnmt3b. The Dnmt3 enzymes set up the initial methylation pattern in the genome during embryonic development; however, they also have a role in the preservation of methylation levels at heterochromatin (2,13). The Dnmt3 family also includes one catalytically inactive regulatory factor called DNA methyltransferase 3 like protein (Dnmt3L), which stimulates the activity of *de novo* DNA methyltransferases and assists them at the setting of the initial methylation imprints (14–16). The distinct feature of all Dnmts is that they have a common structure of their C-terminal catalytic domain, which contains 10 amino acids motifs that are conserved among prokaryotic and eukaryotic C5 DNA methyltransferases. Motifs I and X have a role in cofactor binding, motifs IV and VI are involved in catalysis. The conserved region between motifs VIII and IX represents the so called Target Recognition Domain (TDR),

\*To whom correspondence should be addressed. Tel: +49 421 200 3247; Fax: +49 421 200 3249; Email: a.jeltsch@jacobs-university.de

which is involved in DNA recognition. The C-terminal domain adopts a fold characteristic for S-adenosyl-L-methionine (AdoMet) dependent methyltransferases (17). Dnmt3a and Dnmt3b have a similar domain arrangement; both contain a variable N-terminal region followed by a PWWP domain, a Cys-rich Zinc-binding ADD domain and the C-terminal catalytic domain (18). The PWWP domain of Dnmt3a reads the H3K36 trimethylation mark (19) and it targets the enzyme to pericentromeric heterochromatin (19–21). The ADD domain of Dnmt3a binds H3 tails unmethylated at K4 (22,23) and stimulates the activity of Dnmt3a after peptide binding (23,24).

The isolated catalytic domain of Dnmt3a (Dnmt3a-C) is catalytically active (25) and it interacts with the C-terminal domain of Dnmt3L (17), forming a butterfly-shaped elongated tetrameric complex (26). This complex contains two monomers of Dnmt3a-C and two monomers of Dnmt3L-C forming a linear 3L–3a–3a–3L heterotetramer. It has two 3a–3L interfaces and one 3a–3a interface. The 3a–3L interaction is mediated by hydrophobic interactions represented by two pairs of phenylalanines F728 and F768 of mouse Dnmt3a and F297 and F337 of mouse Dnmt3L, whereas the 3a–3a interaction is mediated by two charged amino acids R881 and D872 (26). The active sites of the two Dnmt3a are located in the major groove of the DNA in a distance of ~40 Å, such that they could methylate two CG sites separated by ~8–10 bp in one binding event (26,27). In the absence of Dnmt3L, Dnmt3a-C also self-interacts via the FF interface forming reversible oligomers, which bind to more than one DNA molecule oriented in parallel (28). The multimerization of Dnmt3a via the FF and RD interfaces and binding to parallel DNA molecules contribute to the heterochromatic localization of the enzyme. Since Dnmt3L does not possess an RD interface, its binding to Dnmt3a disrupts Dnmt3a oligomerization and leads to the release of Dnmt3a from heterochromatic regions, which could help methylate euchromatic targets like the differential methylated regions in imprinting centers (28).

We have shown previously that the heterotetrameric Dnmt3a/3L complex, as well as oligomeric Dnmt3a-C complexes, bind non-specifically to DNA and cooperatively multimerize on DNA forming large nucleoprotein filaments (26–28). Cooperative multimerization of Dnmt3a/3L and Dnmt3a complexes on DNA was shown by DNA-binding analyses and scanning force microscopy. DNA methylation studies revealed fluctuations of methylation activities along the DNA molecule, which also indicated that there must be a regular arrangement of Dnmt3a complexes on the DNA (26–28). Cooperative DNA binding implies an interaction of Dnmt3a complexes which bind next to each other on the DNA. In this work, based on modeling we have identified one loop that contributes to the interface of such neighboring complexes. By mutating critical residues in the interface and introducing opposite charges, we disrupted the interaction and prevent multimerization of Dnmt3a complexes on DNA, which confirms the predication from modeling. Non cooperative Dnmt3a variants did not lose DNA binding and retained methylation activity, indicating

that cooperative DNA binding and multimerization of Dnmt3a on the DNA is not required for enzyme activity. Loss of the cooperative multimerization of Dnmt3a on DNA reduced the heterochromatic localization of the enzyme in NIH3T3 cells, suggesting that it has a role in binding of Dnmt3a to heterochromatic regions.

## EXPERIMENTAL PROCEDURES

### Site-directed mutagenesis, protein expression and purification

The sequence encoding the C-terminal domain (residues 608–908) of mouse Dnmt3a (Dnmt3a-C) was cloned into pET-28a (Novagen) with an N-terminal His<sub>6</sub>-tag (17). Based on modeling, we selected 10 hydrophilic and charged amino acids for the interface study and using the megaprimer site-directed mutagenesis method we mutated these residues to opposite charges to break the interaction (29). Mutagenesis was confirmed by restriction marker analysis and DNA sequencing (Supplementary Figure S1). Protein expression was carried out as described (17). The proteins were purified at high micromolar concentrations using Ni-NTA agarose. Each protein was purified at least twice and the purity of the preparations was estimated to be >95% from Coomassie stained SDS gels (Supplementary Figure S2). The concentrations of the proteins were determined by UV spectrophotometry and confirmed by densitometric analysis of Coomassie stained SDS-polyacrylamide gels. Wild-type like folding of the most important mutants was confirmed by circular dichroism spectroscopy (Supplementary Figure S3).

### Methyltransferase activity assay

The methyltransferase activity of Dnmt3a-C wild-type and its interface mutants was measured using biotinylated 30-mer oligonucleotides (Bt-GAG AAG CTG GGA CTT CCG GGA GGA GAG TGC/GCA CTC TCC TCC CGG AAG TCC CAG CTT CTC) containing a single CpG site as described (27). Briefly, DNA methylation was measured by the incorporation of tritiated methyl groups from radioactively labeled AdoMet (Perkin Elmer) into the biotinylated oligonucleotide using the avidin-biotin methylation assay (30). The methylation reactions were carried out in methylation buffer (20 mM HEPES pH 7.2, 1 mM EDTA, 50 mM KCl, 25 µg/ml bovine serum albumin) at 37°C, using 1 µM substrate DNA, 0.76 µM AdoMet and 2 µM Dnmt3a-C wild-type and mutant enzyme. The initial slope of the enzymatic reaction was determined by linear regression. All the kinetic reactions were carried out at least three times. DNA methylation was also measured using a 520-bp DNA fragment containing 40 CpG sites was amplified from phage λ DNA using a biotinylated PCR primer as described (31).

### Methylation pattern analysis by bisulfite conversion

DNA methylation patterns of Dnmt3a-C wild-type and mutants were analyzed by hairpin bisulfite analysis (27,32) using an oligonucleotide substrate containing

nine CG sites in an identical sequence context d(AAT TGA CGA CGA CGA CGA CGA CGA CGA CGA CGA CGAC)/d(GAT CGT CGTC). The annealed oligonucleotide (100 nM) was incubated with 2.5  $\mu$ M enzyme for 2 h in methylation buffer supplemented with 0.32 mM AdoMet (Sigma) at 37°C. The methylation reaction was stopped by freezing the sample in liquid nitrogen, followed by proteinase K digestion (New England Biolabs). Afterwards, the hairpin loop and the adaptors were ligated to the methylated substrates as described (27) and DNA was subjected to bisulfite conversion, cloned into Topo-TA vector (Invitrogen) and individual clones were sequenced essentially as described (33).

In addition, the DNA methylation patterns of Dnmt3a wild-type and mutants was analyzed using a 146-mer DNA substrate amplified from the mammary tumor virus 3' long-terminal repeat nucleosome A-binding site (34) which contains 10 CpG sites in various distances. Methylation reactions and bisulfite conversion were conducted as described above.

#### Electrophoretic mobility shift assay

The DNA-binding affinity and cooperativity of Dnmt3a-C and its mutants was analyzed by electrophoretic mobility shift assay (EMSA) carried out essentially as described (27). Briefly, increasing concentrations of Dnmt3a-C wild-type and its interface mutants were incubated with the Cy-5-labeled 146 bp DNA described above (30 nM) in reaction buffer (20 mM HEPES pH 7.5, 1 mM EDTA, 100 mM KCl, 0.5 mg/ml of bovine serum albumin) containing 0.2 mM sinefungin at room temperature for 20 min to form the protein–DNA complex. The protein–DNA complexes were separated on an 8% non-denaturing acrylamide gel, which was scanned with a phosphor imager system (Fuji). Free DNA and wild-type Dnmt3a-C protein were included on the gels as controls.

#### Fluorescence depolarization

To study equilibrium binding of Dnmt3a to Cy5-labeled oligonucleotide substrates, the change of fluorescence anisotropy caused by protein binding to the DNA was determined using a Cary Eclipse fluorescence spectrophotometer (Varian) equipped with excitation and emission polarizers. Three substrates were used, two 29-mer oligonucleotides (2CG-Cy5 and non-CG-Cy5) and one 60-bp PCR product containing two CpG sites that was amplified from the 146-mer substrate described above (34). All binding substrates were fluorescently (Cy5) labeled at one end.

2CG-Cy5: ACT TGC AAC GGT CCT AAC CGT CAC CTC TT

Non-CG-Cy5: ACT TGC AAC AGT CCT AAC ATT CAC CTC TT

The anisotropy was measured with the excitation wavelength at 633 nm (band-width 10 nm) and the emission wavelength at 665 nm (band-width 10 nm) for 5 s. The reactions were carried out in binding buffer containing 20 mM HEPES pH 7.5, 100 mM KCl, 1 mM EDTA,

2 nM of Cy5-labeled DNA and increasing concentration of protein after incubation at room temperature for 5 min. Each protein concentration was measured in triplicate and the average values were taken for the analysis. Data were least squares fitted using a 1:1 binding model for the 29-mer, where no cooperativity could be detected. For the 60-mer, a cooperative binding model was used to determine the dissociation constant ( $K_D$ ) and Hill coefficient ( $n$ ) of DNA binding.

#### Scanning force microscopy

Scanning force microscopy (SFM) experiments were performed with Dnmt3a-C wild-type and the R832E and K837E mutants using a 509-bp DNA fragment, which can be visualized easily in SFM. The DNA was derived from the CG island upstream of the human SUHW1 gene and contains 58 CG sites roughly equally distributed over the entire DNA length. The sample preparation was carried out essentially as described (27). Briefly, the DNA–protein filaments were formed in a reaction volume of 30  $\mu$ l by incubating 12 nM DNA with 200 nM Dnmt3a-C wild-type or mutants in 50 mM HEPES (pH 7.5), 250 mM NaCl, 1 mM EDTA and 100  $\mu$ M of sinefungin (Sigma). After addition of DNA to the reaction mixture containing the protein, samples were incubated for 20 min at room temperature to allow for DNA binding. Then, the complex solution (1  $\mu$ l) was mixed with 9  $\mu$ l of 5 mM NiCl<sub>2</sub> solution and deposited on freshly cleaved mica (Plano, GmbH), allowed to adhere for 2 min and then washed with sterile water. The sample was then dried using compressed air. Protein–DNA filaments were observed by tapping mode in air using a Multimode SFM with a Nanoscope IIIa controller (Veeco Instruments GmbH, Germany) using RTESPW silicon cantilevers (Veeco) with a nominal spring constant of 50 N/m and a resonance frequency of 150 kHz. All images were obtained with a scanning speed of 1 Hz and at a resolution of 512  $\times$  512 pixels. To remove background slope, raw images were flattened using the Nanoscope software. DNA–protein complexes were considered present if the height of the filament exceeded 150% of the height observed for free DNA molecules and if filaments had an apparent width of at least 20 nm. Filaments were evaluated using the section tool of the Nanoscope V6r12 software and ImageJ software.

#### Cell culture and laser scanning microscopy

The cellular and sub-nuclear localization of the YFP-fused Dnmt3a and its variants in NIH 3T3 cells was investigated essentially as described (19,28). Briefly, cells were grown in DMEM with 10% (v/v) fetal calf serum and 2 mM L-glutamine at 37°C in 5% (v/v) CO<sub>2</sub>. Cells ( $1\text{--}2 \times 10^5$ ) were transfected in six-well plates using FuGENE 6 (Roche, Basel, Switzerland; 1  $\mu$ g total plasmid DNA per well). Transfected NIH3T3 cells were fixed in 4% (w/v) paraformaldehyde. Confocal images were taken using a Carl Zeiss LSM510 (Jena, Germany; software version 3.0). The sub-nuclear localization pattern of Dnmt3a was inspected in detail for 100 cells for the wild-type, 25 cells for the R832E and K837E mutants (which did not

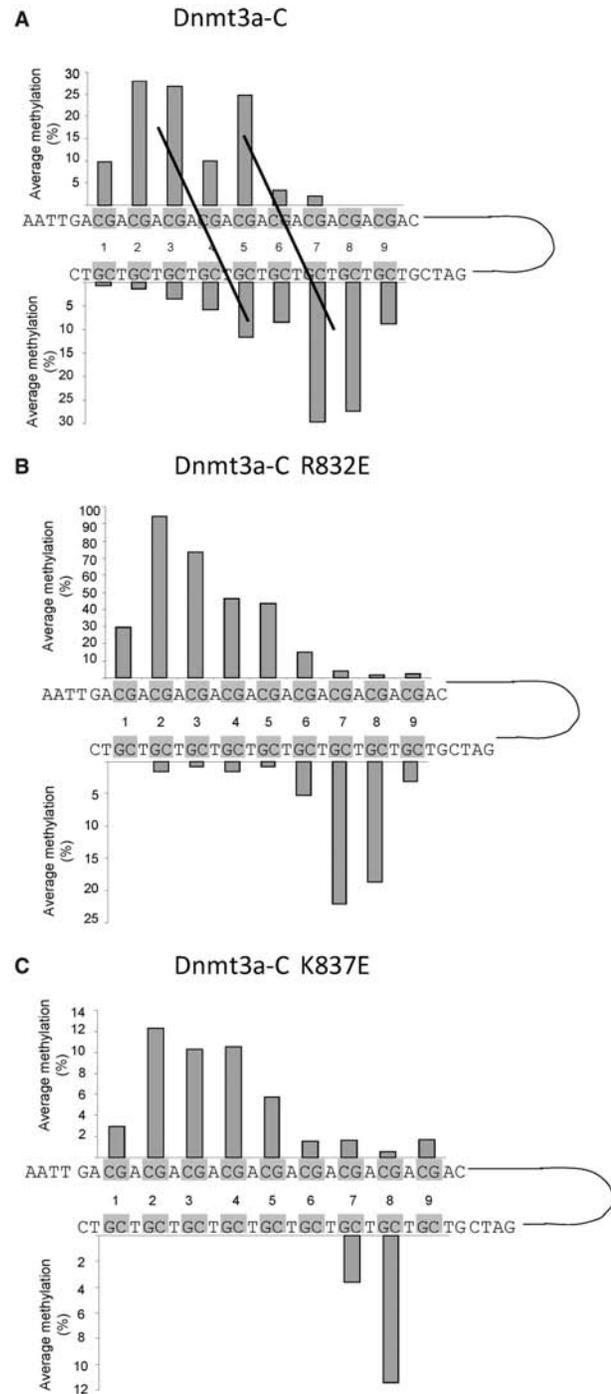
differ from wild-type) and 50 cells for the R832E/K837E double mutant and the observed patterns were divided into three categories: spotty, spotty plus diffused and diffused.

## RESULTS

We have already shown that Dnmt3a cooperatively multimerizes on DNA. To identify the positions of the individual active sites on the DNA, we employed an ‘activity footprint’ approach using a substrate which contains nine CG sites next to each other. By applying the hairpin bisulfite technology, the methylation of both strands of individual molecules could be directly studied. After methylation of this substrate with wild-type Dnmt3a-C (Figure 1A) four peaks of preferential methylation were observed, two of them in the lower strand (first peak at sites 7 and 8 and second peak at site 5) and two in the upper strand (first peak at site 5 and second peak at sites 2 and 3). These peaks are all separated by 8- to 10-bp distances both in the same strand and across the strands. This result is in perfect agreement with results of a previous experiment with the Dnmt3a-C/3L-C complex (27). According to the model of the Dnmt3a/3L heterotetramer in complex with DNA (26,27), the methylation of two sites in a distance of 8–10 bp in opposite DNA strands could be attributed to the activity of the active sites from the two central Dnmt3a subunits of the Dnmt3a or Dnmt3a/3L complex, which interact via the RD interface. In contrast, the preferential methylation of sites in distances of 8–10 bp located in the same DNA strand must be due to the presence of two Dnmt3a complexes binding next to each other on the DNA.

The observation of four peaks of methylation suggests that two Dnmt3a-C complexes were bound to the substrate, each of them containing two active centers connected via an RD interface. The right complex methylates the lower strand sites 7 and 8 with one of its active sites and the upper strand site 5 with the other. The left complex approaches the other one and it methylates the lower strand site 5 with one of its active sites and the upper strand sites 2 and 3 with the other (Supplementary Figure S6). The observation that CpG site 5 was preferentially methylated in both DNA strands indicates that the Dnmt3a-C subunits in adjacent Dnmt3a-C complexes can approach each other and interact with both DNA strands of this CpG site.

Several observations support the significance of the methylation peaks at CpG site 5 in both strands: (i) *P*-values are highly significant (except for the comparison of site 5 and 6 in the lower strand where the *P*-value is 0.078) (Supplementary Figure S4), (ii) the geometry of the peaks is palindromic in the upper and lower strand with methylation peaks at sites 2, 3 and 5 in the upper strand corresponding to methylation peaks at sites 8, 7 and 5 in the lower strand, (iii) methylation preferences in opposite strands occur in a distance of 8–10 bp as expected for the two active sites of one Dnmt3a complex (sites 2 and 3 in the upper strand combined with site 5 in the lower strand, site 5 in the upper strand combined with sites 7 and 8 in



**Figure 1.** Methylation pattern studied by hairpin bisulfite experiments. Double-stranded hairpin oligonucleotide substrates containing nine equally spaced CG sites were methylated by Dnmt3a-C wild-type or R832E and K837E mutants. 185, 170 and 190 clones were analyzed for Dnmt3a-C wild-type (A), R832E (B) and K837E (C), respectively. The lines in (A) connect peaks of methylation separated by ~8 bp in the opposite DNA strand.

the lower strand), (iv) a very similar pattern has been observed before in a completely independent experiment (27). The preferred methylation of both DNA strands at site 5 suggests that two Dnmt3a-C subunits from adjacent complexes can simultaneously approach the upper and the

lower strand of one CpG site (site 5 in this case). We took this observation as the starting point for modeling of the multimerization of Dnmt3a-C complexes on DNA.

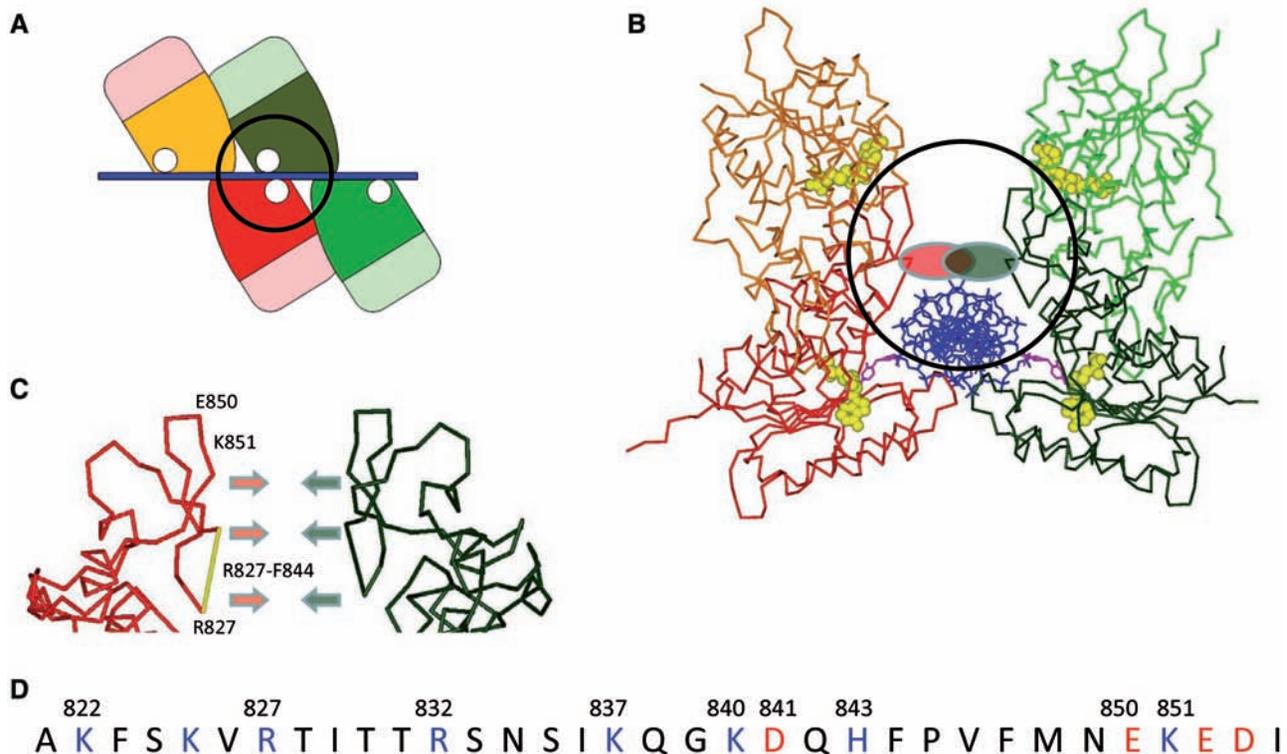
### Modeling of the multimerization of Dnmt3a-C complexes on DNA

To model a complex in which two Dnmt3a-C molecules would interact simultaneously with the upper and lower strand of one CpG site, we have used as a template the structure of the bacterial M.HhaI methyltransferase bound to DNA (35), which was the first methyltransferase for which base flipping could be shown. The complex was used twice and the two DNA strands with the flipped cytosine were annealed with each other to obtain a CpG site with both target cytosines flipped out. Then, Dnmt3a-C molecules were superimposed with the two M.HhaI molecules using the conserved methyltransferase motifs. In the obtained structural model it was also possible to replace each Dnmt3a-C by a Dnmt3a-C/3L-C heterotetramer without larger steric overlaps to obtain a model of two Dnmt3a-C or Dnmt3a-C/3L-C complexes bound to DNA. In this model, two loops (residues 820–855) of the two Dnmt3a-C complexes approach each other symmetrically (Figure 2). This loop is enriched with charged amino acids and may form a

hydrophilic interface, which could explain the co-operative multimerization of Dnmt3a-C and Dnmt3a-C/3L-C complexes on DNA. A large part of this loop (residues 827–844) is not ordered in the structure, such that the detailed interactions could not be deduced. The modeling also suggested that the active sites of the two Dnmt3a-C subunits from adjacent complexes that are interacting with the same DNA strand are located ~40-Å apart from each other, indicating that they could methylate two CG sites separated by 10 bp in one binding event. However, given that this modeling did not take into account any conformational changes of the DNA or the Dnmt3a-C complexes, its results need to be interpreted with caution. It was the aim of the following study to investigate if the putative interface suggested by modeling is involved in multimerization of Dnmt3a-C on DNA, if the multimerization can be influenced by mutations and if non-multimerizing Dnmt3a-C variants would still be able to bind and methylate DNA.

### DNA binding of Dnmt3a-C and its interface variants studied by EMSAs

To study the role of the selected residues in the cooperative multimerization of Dnmt3a-C, we prepared the putative interface mutants K822E, R827E, R832E,



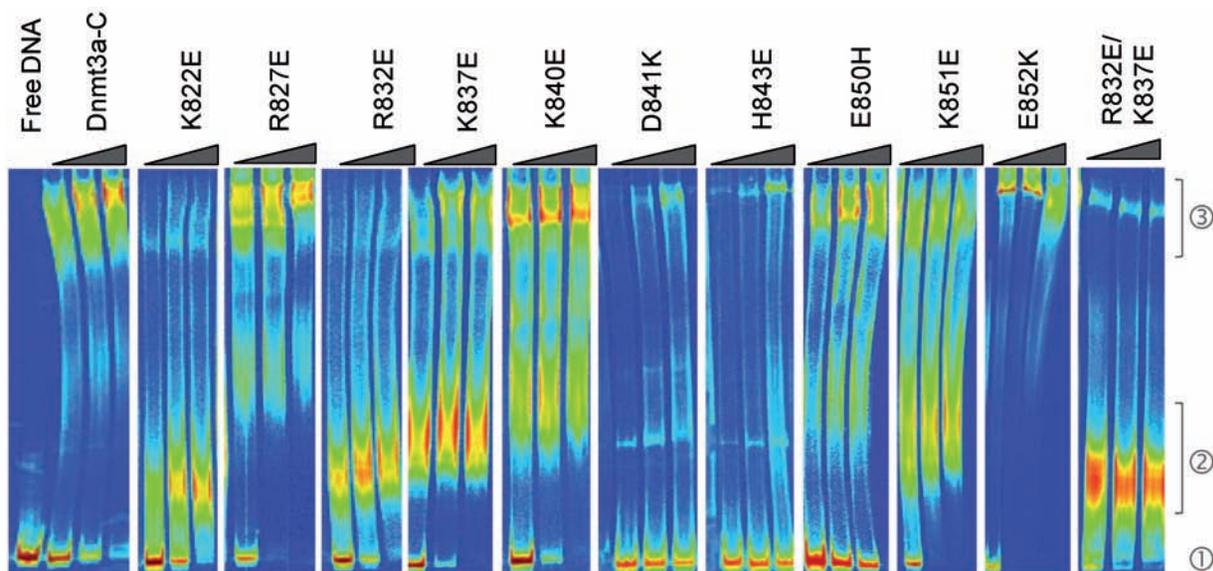
**Figure 2.** Modeling of two Dnmt3a-C complexes bound to DNA. (A) Schematic model of two Dnmt3a-C/3L-C heterotetramers binding next to each other on the DNA, such that a central CpG site could be methylated in both DNA strands. The Dnmt3a-C subunits of the left tetramer are colored orange and red, those of the right tetramer dark green and green. Dnmt3L subunits are colored light red and light green, respectively. (B) Model of the central Dnmt3a-C dimers (colored orange and red for one complex and dark green and light green for the other) bound to DNA (blue). The flipped target cytosines are shown in pink, AdoMet is colored yellow. The red and green ovals indicate the proposed interactions of both loops. Note that the amino acids 827–844 are not ordered in the structure and hence not visible in the model. (C) Detail of the model showing the interacting loops from the adjacent complexes. The disordered region (amino acids 827–844) is colored yellow. (D) Amino acid sequence of the putative interaction loop.

K837E, K840E, D841K, H843E, E850H, K851E and E852K. In each of them, a charged residue was exchanged by a residue carrying an opposite charge. Such exchange should result in a repulsive interaction, if the residue was involved in a charge/charge interaction at the interface and, therefore, disrupt the cooperativity of DNA binding by preventing the binding of two Dnmt3a complexes next to each other. The Dnmt3a-C wild-type and its mutants showed comparable expression in *Escherichia coli* cells. The proteins were purified to similar quality as the wild-type Dnmt3a-C and their folding confirmed by circular dichroism spectroscopy for the R832E and K837E mutants (Supplementary Figure S2 and S3). EMSA were carried out to study the DNA binding of the wild-type Dnmt3a-C and its interface mutants (Figure 3). It had been shown previously that the wild-type Dnmt3a-C multimerizes on the DNA at increasing protein concentration, as indicated by the formation of a protein–DNA complex with very low electrophoretic mobility corresponding to several Dnmt3a-C complexes being bound to the DNA (region 3 in Figure 3) (27). The high degree of cooperativity of the multimerization reaction is indicated by the formation of the multimeric complex without generation of detectable amounts of intermediates. Using identical reaction conditions, the binding of the Dnmt3a-C interface mutants on DNA was studied. The K822E, R832E and K837E mutants shifted the DNA into complexes with an electrophoretic mobility that corresponds to one Dnmt3a-C complex being bound to the DNA (region 2 in Figure 3), which indicates that these variants have lost the ability to multimerize on DNA. The slight differences in the mobilities of these complexes may be explained by the exchanges of charged residues which can affect the

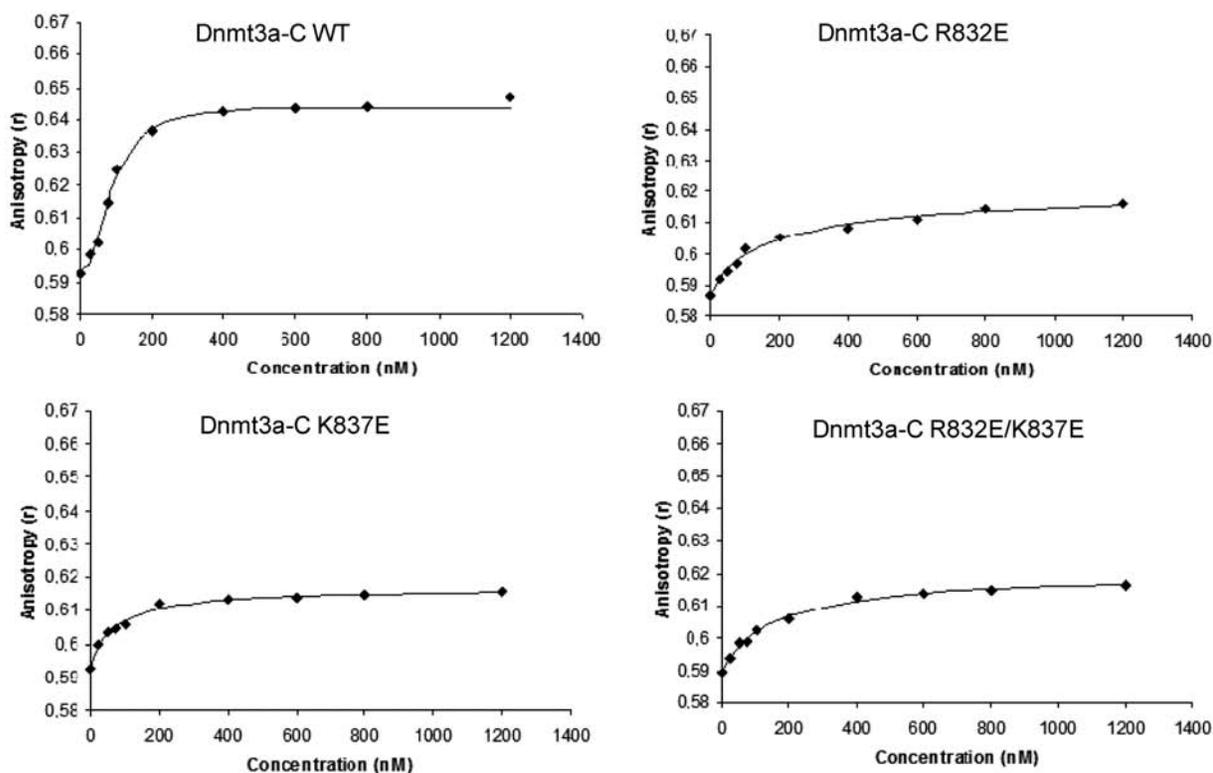
behavior of the mutants in native gel systems. DNA binding of these variants is efficient as illustrated by the disappearance of unbound DNA, which indicates that the loss of multimerization is not a trivial effect caused by a reduction of DNA binding. The K840E and K851E mutants bound strongly to DNA and formed large multimeric complexes, but intermediates were also detectable, indicating lower cooperativity of multimerization than observed with the wild-type Dnmt3a-C. E852K showed better DNA binding than the wild-type enzyme and cooperativity could not be evaluated due to the strong binding. The D841K and H843E mutants bound very weakly to the DNA. In summary, we conclude that mutation of most of the candidate residues led to a loss or reduction of the cooperativity of multimerization of Dnmt3a-C complexes on DNA. The most interesting non-multimerizing variants were R832E and K837E, which both completely lost the ability to form cooperative complexes, but showed good DNA binding. To obtain even stronger phenotypic effects, we have also generated and purified the R832E/K837E double mutant and confirmed its folding by circular dichroism spectroscopy (Supplementary Figure S2 and S3). As expected, this variant showed strong DNA binding and loss of multimerization (Figure 3).

#### DNA binding of Dnmt3a-C and its interface variants studied in solution

To study DNA binding in a homophasic assay, we used Cy5-labeled DNA substrates and determined the increase in fluorescence anisotropy of the Cy5 probe, which accompanied protein binding to the DNA. We first used two 29-mer substrates which either contain two or no



**Figure 3.** DNA binding and cooperativity studied by EMSA. Dnmt3a-C wild-type protein forms a larger complex with very low electrophoretic mobility (region 3). Free DNA runs in region 1. The K822E, R832E, K837E and R832E/K837E mutants bind to DNA, but form smaller complexes with an electrophoretic mobility that corresponds to one Dnmt3a-C complex being bound to the DNA (region 2). The R827E, K840E, E850H and K851E mutants formed large multimeric complexes, but also some intermediates. The D841K and H843E mutants bound very weakly to the DNA. The E852K mutant showed even better DNA binding than the wild-type enzyme and no change in cooperativity.



**Figure 4.** DNA binding and cooperativity studied by fluorescence anisotropy. DNA binding and cooperativity was studied using a 60-mer Cy-5-labeled DNA for the Dnmt3a-C wild-type and R832E, K837E and R832E/K837E proteins at increasing concentration. Dnmt3a-C binds to DNA very strongly with positive cooperativity ( $n = 2.2$ ), suggesting that at least three molecules would bind on the DNA. The R832E, K837E and R832E/K837E mutant proteins showed no cooperativity ( $n = 1.0$ ), but still bound the DNA.

**Table 1.** Summary for the catalytic activities, DNA-binding constants and DNA-binding cooperativity of wild-type Dnmt3a-C and the interface mutants

Protein	Catalytic activity (% of wild-type)	DNA-binding constant ( $K_d$ in nM)	Cooperativity of DNA binding ( $n$ )
Dnmt3a-C WT	100	86	2.2
Dnmt3a-C K822E	5	310	1.0
Dnmt3a-C R827E	12	61	2.2
Dnmt3a-C R832E	48	190	1.0
Dnmt3a-C K837E	23	72	1.0
Dnmt3a-C K840E	28	77	1.3
Dnmt3a-C D841K	67	400	1.0
Dnmt3a-C H843E	38	130	1.0
Dnmt3a-C E850H	110	140	1.1
Dnmt3a-C K851E	42	130	1.0
Dnmt3a-C E852K	110	64	1.3
Dnmt3a-C R832E/K837E	5	151	1.0

All values were averages of two or more experiments. Highest standard deviations were  $\pm 10\%$  for the catalytic activity,  $\pm 25\%$  for the  $K_d$  and  $\pm 15\%$  for the cooperativity values.

CpG sites. Fluorescence anisotropy was measured in solution containing 2 nM of labeled DNA with increasing protein concentration. Fitting to an equilibrium binding model yielded a  $K_d$  of  $190 \pm 10$  nM for the 2-CG substrate and a  $K_d$  of  $130 \pm 12$  nM for the non-CG substrate. This result indicated that Dnmt3a-C bound to the DNA

without specificity for CpG sites. The slightly better binding to the non-CG substrate might be due to the flanking sequence, which has a strong influence on the DNA interaction and DNA methylation of Dnmt3a that is well documented [see (31) and references therein]. It is also possible that the energy needed for flipping of the target cytosine may reduce the binding to CpG site substrates [see (36) and references therein for examples of prokaryotic DNA methyltransferase studied in this respect].

To investigate cooperative DNA binding of Dnmt3a-C, we used a longer Cy5 labeled 60-mer DNA substrate for subsequent binding studies. The cooperativity of the wild-type Dnmt3a-C and mutant DNA binding was determined by fitting the anisotropy values to the Hill equation. The Dnmt3a-C bound the 60-mer DNA strongly (the calculated  $K_D$  value was 86 nM) and the calculated cooperativity value for the wild-type protein was  $n = 2.2$ , indicating that at least three Dnmt3a-C complexes bound to the DNA in a cooperative manner (Figure 4 and Table 1). Although DNA substrates and assay methods were different, there was a generally nice correspondence between the results of the EMSA and fluorescence polarization DNA-binding studies. In contrast to the wild-type enzyme, the variants that showed loss or reduction of cooperativity in DNA binding in the EMSA, like K822E, R832E, K837E, K851E, or the R832E/K837E double mutant had

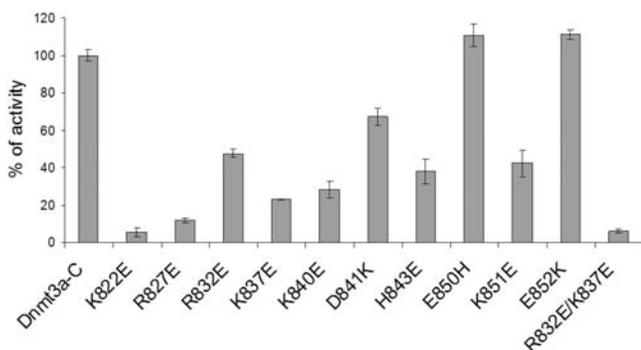
cooperativity values close to  $n = 1$  (Figure 4 and Table 1). As seen in EMSA, R827E bound DNA with similar affinity and cooperativity as the wild-type and E852K bound DNA better. Also weaker DNA binding by D841K, H843E and E850H was observed. We conclude that R832E, K837E, K851E and the R832E/K837E double mutant are most interesting. All of them show complete loss of cooperativity in DNA binding and their binding affinity is not reduced >50%. K837E is particularly striking, because its binding constant is identical to the wild-type Dnmt3a-C (within the limits of error) but cooperativity is completely lost.

### Catalytic activity of Dnmt3a-C and its interface variants

The catalytic activity of the variants was determined using an oligonucleotide substrate containing one CpG target site. The K822E, R827E, K837E, K840E and H843E mutants showed a strong reduction of methyltransferase activity, whereas R832E, D841K and K851E showed only a weak reduction of activity. The activity of the E850H and E852K mutants was not affected (Figure 5 and Table 1). We tested the activity of the R832E, K837E and R832E/K837E double mutant also using a 520-bp substrate which contains 40 CpG sites and did not observe major differences in the relative activities when compared with wild-type Dnmt3a-C (Supplementary Figure S5). In a previous study, we already investigated the R827A and R832A mutants, which showed 45% and 75% of wild-type activity, respectively (37). The stronger loss of activity observed here can be understood, because the charge reversal done here was expected to cause a stronger effect than the exchange by alanine in the previous study. In summary, the R832E and K851E mutants are most interesting, because both of them show ~50% of activity despite a complete loss of cooperativity in DNA binding, indicating that multimerization is not required for catalytic activity.

### Methylation pattern of Dnmt3a-C and its interface variants

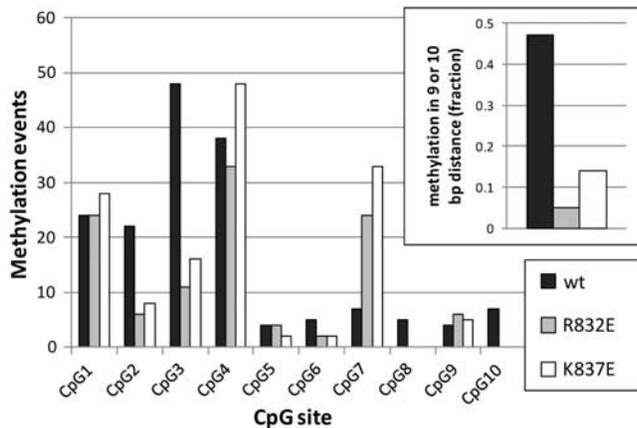
We also performed hairpin bisulfite ‘activity footprint’ experiments with the active non-polymerizing Dnmt3a-C



**Figure 5.** DNA methyltransferase activity of the interface loop mutants. The methyltransferase activity of Dnmt3a-C and its interface mutants was determined using an oligonucleotide substrate containing a single CpG site. All the kinetics were done at least in duplicate. The error bar indicates the standard error of initial slopes.

mutants as described above for the wild-type enzyme (Figure 1). The activity footprint of the active non-polymerizing R832E and K837E mutants showed preservation of the preferential methylation of sites 2 and 3 in the upper strand as well as 7 and 8 in the lower strand (Figure 1B and C). However, both mutants showed clear changes in the methylation pattern including a loss of the methylation at CpG site 5 in the lower strand and broadening of the upper strand methylation peak. The methylation of site 5 in the upper strand was still detected (although somehow less prominent). When taken together, these observations suggest that the right Dnmt3a complex is bound to the substrate in the same manner as with the wild-type Dnmt3a-C (Supplementary Figure S6). Upper strand methylation at sites 2, 3 and 4 indicates that the second (left) Dnmt3a-C complex is present as well. The broadening of the methylation peak in the upper strand observed with the mutants suggests that the left complex is more mobile and can reach more sites. This could be due to the loss of the interaction with the right tetramer caused by the mutations (Supplementary Figure S6). Loss of methylation of the lower strand of site 5 suggests that the repulsion caused by the close approximation of two mutations from adjacent complexes led to a conformational change in the lower subunit of the left complex which caused the loss of methylation of the corresponding target site (Supplementary Figure S6). This result indicates that methylation is not possible in both DNA strands at CpG sites in the interface of two Dnmt3a complexes, which is exactly what was expected from our modeling. In addition, the R832E mutant showed a reduced methylation of the lower strand when compared to the upper one. We speculate that the reason for this change is that the mutations affect the flanking sequence preferences of Dnmt3a (upper strand flanking context is ACGA, lower strand it is TCGT). Altered flanking preferences of the mutants were also observed on the larger substrate (see the next paragraph).

The DNA methylation patterns of wild-type Dnmt3a-C and the R832E and K837E mutants were also investigated with the 146-bp substrate that was also used for the gel shift experiments. Methylation was analyzed by bisulfite analysis of individual clones to derive the overall methylation profiles, which revealed strong differences between the wild-type and mutant enzymes (Figure 6). However, these overall methylation profiles display the combined effects of the mutations on flanking sequence preferences and preferences for co-methylation in certain distances. Since we were mainly interested in the latter, we focused on those clones showing more than one methylation event and collected the distances of all comethylation events for wild-type and mutants. As shown in the insert of Figure 6, with the wild-type enzyme we observed that 47% of all comethylation events occurred in a distance of 9 or 10 bp (the substrate does not contain CpG sites in a distance of 8 bp). This indicates a strong preference for comethylation in such distances, similarly as shown before with two other DNA substrates (26). This preference was completely lost with the mutants, where only 2% and 8%, respectively, of all comethylation events occurred in this distance range. Since only comethylation events were used for this

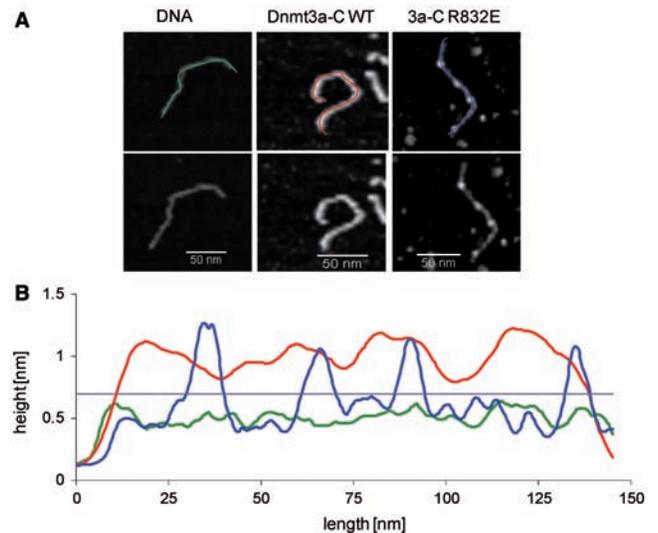


**Figure 6.** Methylation pattern analysis of wild-type Dnmt3a and its R832E and K837E mutants on the 146-mer DNA substrate which has 10 CpG sites. The insert displays the fraction of co-methylation events in a distance of 9 or 10 bp observed with the wild-type enzyme and the mutants.

analysis, this difference cannot be attributed to the lower activity of the mutants. This result confirms our conclusion from the hairpin bisulfite experiment that the preference for co-methylation in a distance of 8–10 bp in the same DNA strand has been lost with the mutants.

#### Imaging of Dnmt3a-C protein–DNA filaments

We used scanning force microscopy to study the structure of Dnmt3a-C wild-type and R832E and K837E mutant DNA complexes. It had been shown previously (27,28) and was confirmed here that Dnmt3a-C/3L-C heterotetramer and Dnmt3a-C complexes form continuous filamentous structures on DNA, while no binding of individual complexes to the DNA was observed. This result confirms the cooperative DNA binding of Dnmt3a-C. Using the section tool of the SFM software and the ImageJ software, we analyzed the height of the filaments and compared with the height of free DNA molecules for about 100 molecules of free DNA and 60 protein–DNA complexes each for wild-type and the two mutants. With wild-type Dnmt3a-C, we identified many stretches, where the protein continuously occupied the DNA (Figure 7 and Supplementary Figure S7). In contrast, the non-cooperative mutants only showed binding of individual protein complexes to the DNA, as indicated by the sharp and narrow peaks in the height profile (Figure 7, Supplementary Figure S7A and B). Such peaks were never observed with the wild-type enzyme. In contrast, long filamentous structures occupied with protein as seen for the wild-type enzyme were never detected with the R832E and K837E mutants, indicating that the DNA-binding properties of the mutants were radically changed. In conclusion, disruption of the Dnmt3a-C interface of adjacent Dnmt3a-C complexes by introducing opposite charges at key residues prevented the binding of Dnmt3a-C complexes next to each other, leading to the binding of individual complexes on the DNA. This result is in agreement with the EMSA data reported above, which indicated that the

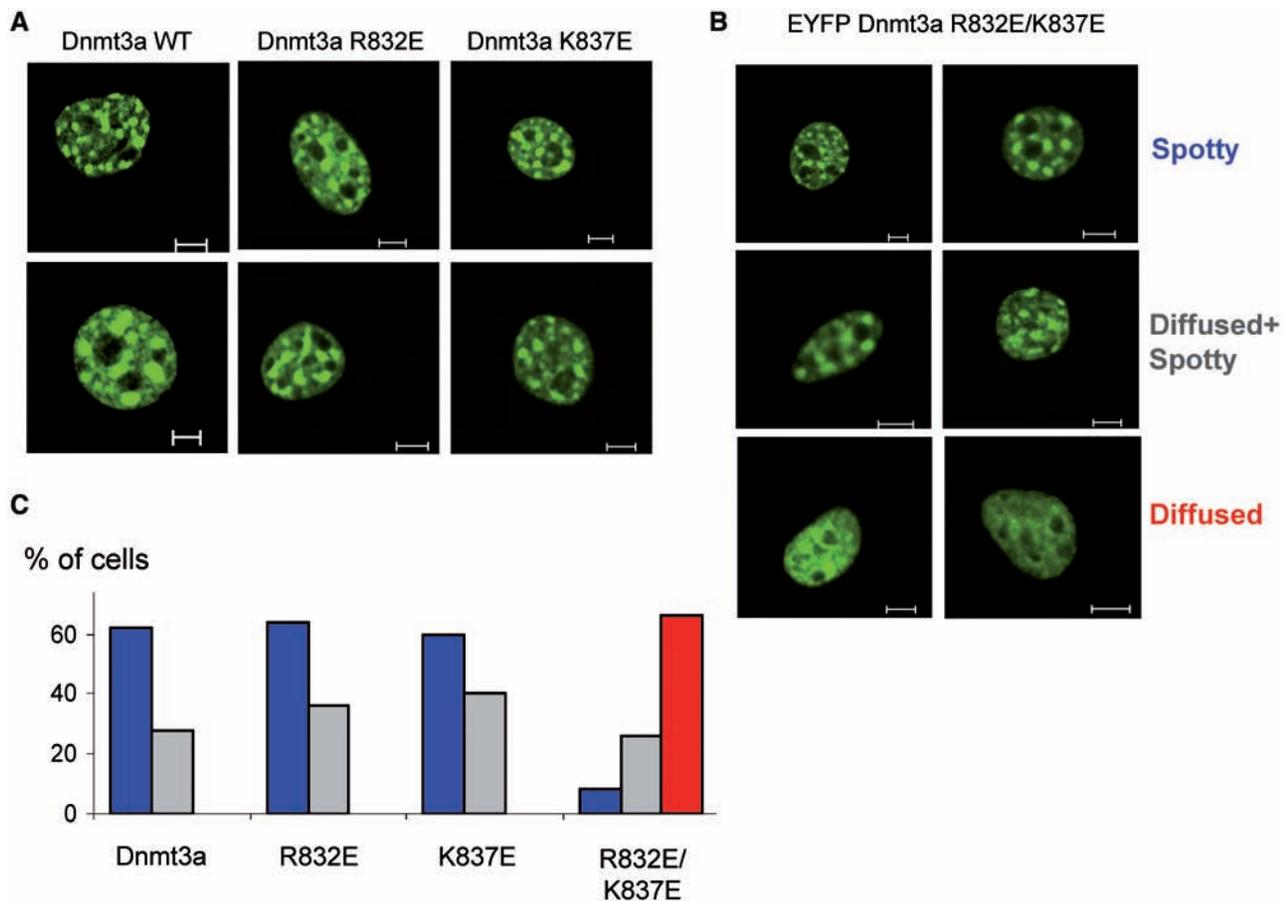


**Figure 7.** Imaging of Dnmt3a-C/DNA complexes by scanning force microscopy. (A) Examples of free DNA, DNA covered with wild-type Dnmt3a-C and DNA with R832E mutant. Wild-type Dnmt3a-C showed complete occupancy of the DNA, whereas the R832E variant showed individual binding events on the DNA. (B) Height analysis of protein–DNA complexes. The height for the free DNA is  $\sim 0.5$  nm, for wild-type Dnmt3a-C a continuous height profile at 1.2 nm is observed, indicative of continuous coverage. In contrast, for the Dnmt3a-C R832E individual binding events with heights of 1.2 nm are separated by free DNA. Additional examples of structures for the R832E and K837E mutants are shown in Supplementary Figure S7A and B.

R832E and K837E mutants still bind to DNA, but they have lost their cooperativity.

#### Heterochromatic localization of Dnmt3a variants in NIH 3T3 cells

The N-terminal PWWP domain of Dnmt3a and the multimerization of its catalytic domain are required for the targeting of this enzyme to the heterochromatin (19–21,28). We were interested to determine if the cooperative DNA binding of Dnmt3a also has an influence on its heterochromatic localization. Hence, we have carried out localization studies with variants that lost cooperative DNA binding. Consistent with previous studies, wild-type Dnmt3a localized exclusively to multiple heterochromatic spots in NIH3T3 cells (19–21,28). Despite their loss of the cooperative DNA binding, the R832E and K837E variants showed a similar localization pattern as the wild-type protein (Figure 8A and C). However, the R832E/K837E double mutant showed an altered sub-nuclear localization pattern with many cells showing a diffused nuclear localization that was never observed with wild-type Dnmt3a (Figure 8B and C). This result indicates that the two mutations in the R832E/K837E double mutant protein prevent the approximation of adjacent Dnmt3a complexes, which interferes with heterochromatic localization of the enzyme. This observation suggests that Dnmt3a binds to heterochromatin by forming close contacts between adjacent Dnmt3a complexes.



**Figure 8.** Localization of Dnmt3a in NIH 3T3 cells. (A) Localization of the eYFP-tagged full-length Dnmt3a wild-type and R832E and K837E mutants in NIH 3T3 cells. The wild-type protein and single mutants show localization to heterochromatic spots (19,28). (B) Localization of the YFP-tagged full-length Dnmt3a R832E/K837E double mutant in NIH 3T3 cells, showing three different patterns of nuclear localization: diffuse, spotty plus diffuse and spotty. (C) Quantification of the localizations patterns observed with wild-type Dnmt3a and the R832E/K837E mutant protein. The scale bars represent 5  $\mu$ m.

## DISCUSSION

Crystallographic studies revealed that the C-terminal domain of Dnmt3L interacts with Dnmt3a-C and forms a butterfly-shaped elongated tetrameric complex (26,27). The active sites of the two Dnmt3a are located in the major groove of DNA  $\sim$ 40-Å apart from each other and methylate two CG sites separated by 8–10 bp in one binding event (26,27). In the absence of Dnmt3L, Dnmt3a-C also interacts via the FF interface forming reversible oligomers, which bind to more than one DNA molecule oriented in parallel (28). In the cell, Dnmt3a is tightly bound to heterochromatin (2,13), as demonstrated by the observation that it can only be eluted from the chromatin by treatment with high salt (38,39). This stable binding requires the interaction of the PWWP domain with H3K36me3 (19) as well as oligomerization of Dnmt3a and binding to parallel DNA molecules (28). The oligomerization of Dnmt3a is disrupted by Dnmt3L, which redistributes the enzyme for methylation of euchromatic targets (28).

In our previous studies, we showed cooperative DNA binding of Dnmt3a-C and Dnmt3a-C/3L-C complexes

(26,27). The cooperative multimerization of proteins on DNA is based on two features, a relatively non-specific DNA binding and a favorable interaction between adjacent protein molecules. Dnmt3a-C indeed binds non-specifically to DNA, such that the first of the pre-conditions for cooperative DNA binding is fulfilled. Here, we have identified the putative interface between Dnmt3a-C complexes bound next to each other on the DNA by modeling and disrupted the interaction by introducing opposite charges at key residues. The mutants retained DNA binding and residual catalytic activity, but lost cooperativity in DNA binding, indicating that we have successfully mapped the oligomerization interface. The interacting loops are partially disordered in the Dnmt3a-C/Dnmt3L-C structure (26), which might be due to the absence of a second enzyme tetramer bound to the DNA. The successful modeling of the interface and its targeted disruption illustrates the good level of understanding we have reached for this important enzyme and its complicated quaternary structure.

Ever since its first observation, the cooperative multimerization of Dnmt3a-C on the DNA was

considered puzzling, because it was difficult to attribute a biological function to this 'in vitro' property. Is there enough naked DNA available in the cell nucleus to allow for the formation of long protein nucleofilaments? What could be the role of the cooperativity in DNA binding? In this study, we reconfirmed the cooperative DNA binding of Dnmt3a-C by employing an EMSA, showing the cooperativity of DNA binding in solution, confirming a characteristic 8- to 10-bp periodicity in DNA methylation in the same DNA strand and direct imaging of protein-DNA complexes by scanning force microscopy. We prepared non-cooperative variants which lost all these properties giving further credence to the experimental approaches. At the same time the non-cooperative mutants did not lose DNA binding and retained methylation activity, which indicates that cooperative DNA binding and multimerization of Dnmt3a on the DNA are not required for enzyme activity—a conclusion that further emphasizes the question of the physiological role of this entire process.

We could imagine two models for the role of cooperative multimerization of Dnmt3a complexes on DNA in cells. One possible interpretation is based on the observation that the formation of a defined structured nucleoprotein complex of Dnmt3a on DNA leads to a characteristic methylation pattern with peaks in a distance of 8–10 bp on the DNA (Figures 5A and 6) (27). Such patterns have been observed in genome-wide DNA methylation analyses (33,40,41) and a similar periodicity of presentation of CpG sites has been found in imprinted Dnmt3a/3L target sites (26), suggesting that they may be of importance. We show here that Dnmt3a-C variants which lost cooperative complex formation did not generate such defined patterns. Hence, multimerization of Dnmt3a and stable filament formation is needed for periodic methylation of DNA. Such binding and periodic DNA methylation could be possible in linker DNA regions or after chromatin remodeling. However, since the 8–10 bp periodicity coincides with the helical repeat length of DNA, it is still unclear if it contains epigenetic information or just reflects the accessibility of CpG sites when the DNA is bound to nucleosomes.

An alternative model is based on the finding that Dnmt3a-C is known to bind very tightly to heterochromatic sites. We have provided evidence that the interaction of the PWWP domain with H3K36me3 is essential for this localization (19), but also the oligomerization of the Dnmt3a protein together with its ability of binding to several DNA molecules oriented in parallel is required (28). We show here that the disruption of the interface of Dnmt3a complexes that is needed for the side by side binding on the DNA, also leads to a disruption of heterochromatic targeting of the enzyme. What emerges is a picture of Dnmt3a binding to very DNA dense regions, which have a condensed and regular structure. We propose that the tight packing of Dnmt3a at these regions requires the close approximation of adjacent Dnmt3a complexes bound to the same DNA molecule. The mutations in the R832E/K837E double mutant protein interfere with such approximation, which could explain the weakening of the heterochromatic

localization. In this model, the tight packing of Dnmt3a at heterochromatic DNA regions is important for the physiology of this enzyme. Tight packing of the enzyme directly requires some favorable interactions between adjacent complexes, which will automatically result in a cooperative binding of the complexes to DNA in an *in vitro* assay, if the general DNA binding is non-specific. Hence, the cooperative binding to naked DNA *in vitro* could be a consequence of the optimization of the interface of Dnmt3a complexes for tight packing of Dnmt3a at heterochromatic regions *in vivo* and not a property directly selected for during molecular evolution of Dnmt3a.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

Thanks are due to Dr Nils Anspach for help with the SFM analyses.

## FUNDING

Funding for open access charge: Deutsche Forschungsgemeinschaft (JE 252/10-1).

*Conflict of interest statement.* None declared.

## REFERENCES

- Cheng, X. and Blumenthal, R.M. (2010) Coordinated chromatin control: structural and functional linkage of DNA and histone methylation. *Biochemistry*, **49**, 2999–3008.
- Jurkowska, R.Z., Jurkowski, T.P. and Jeltsch, A. (2011) Structure and function of mammalian DNA methyltransferases. *ChemBiochem*, **12**, 206–222.
- Kim, J.K., Samaranyake, M. and Pradhan, S. (2009) Epigenetic mechanisms in mammals. *Cell Mol. Life Sci.*, **66**, 596–612.
- Feinberg, A.P. and Tycko, B. (2004) The history of cancer epigenetics. *Nat. Rev. Cancer*, **4**, 143–153.
- Jones, P.A. and Baylin, S.B. (2007) The epigenomics of cancer. *Cell*, **128**, 683–692.
- Berdasco, M. and Esteller, M. (2010) Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Dev. Cell*, **19**, 698–711.
- Meissner, A. (2010) Epigenetic modifications in pluripotent and differentiated cells. *Nat. Biotechnol.*, **28**, 1079–1088.
- Li, E., Bestor, T.H. and Jaenisch, R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, **69**, 915–926.
- Okano, M., Bell, D.W., Haber, D.A. and Li, E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, **99**, 247–257.
- Feinberg, A.P. (2007) Phenotypic plasticity and the epigenetics of human disease. *Nature*, **447**, 433–440.
- Meaney, M.J. and Ferguson-Smith, A.C. (2010) Epigenetic regulation of the neural transcriptome: the meaning of the marks. *Nat. Neurosci.*, **13**, 1313–1318.
- De Carvalho, D.D., You, J.S. and Jones, P.A. (2010) DNA methylation and cellular reprogramming. *Trends Cell. Biol.*, **20**, 609–617.
- Jones, P.A. and Liang, G. (2009) Rethinking how DNA methylation patterns are maintained. *Nat. Rev. Genet.*, **10**, 805–811.
- Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B. and Bestor, T.H. (2001) Dnmt3L and the establishment of maternal genomic imprints. *Science*, **294**, 2536–2539.

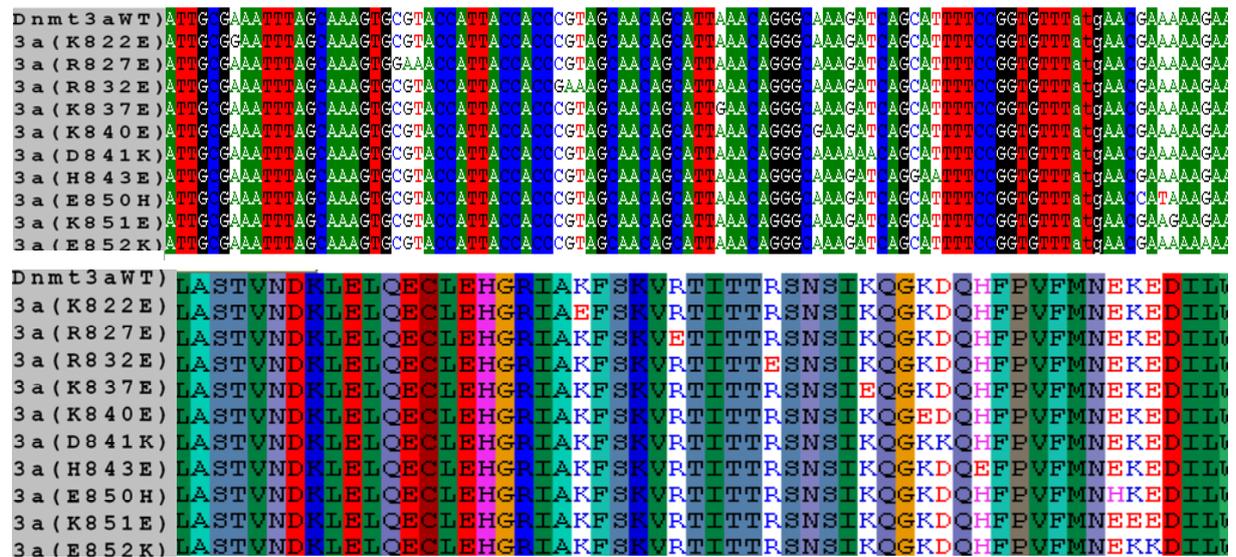
15. Hata, K., Okano, M., Lei, H. and Li, E. (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development*, **129**, 1983–1993.
16. Bourc'his, D. and Bestor, T.H. (2004) Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature*, **431**, 96–99.
17. Gowher, H., Liebert, K., Hermann, A., Xu, G. and Jeltsch, A. (2005) Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. *J. Biol. Chem.*, **280**, 13341–13348.
18. Cheng, X. and Blumenthal, R.M. (2008) Mammalian DNA methyltransferases: a structural perspective. *Structure*, **16**, 341–350.
19. Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R.Z., Ragozin, S. and Jeltsch, A. (2010) The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J. Biol. Chem.*, **285**, 26114–26120.
20. Chen, T., Tsujimoto, N. and Li, E. (2004) The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin. *Mol. Cell Biol.*, **24**, 9048–9058.
21. Ge, Y.Z., Pu, M.T., Gowher, H., Wu, H.P., Ding, J.P., Jeltsch, A. and Xu, G.L. (2004) Chromatin targeting of de novo DNA methyltransferases by the PWWP domain. *J. Biol. Chem.*, **279**, 25447–25454.
22. Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M. and 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.
23. Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., Fischle, W. *et al.* (2010) Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res.*, **38**, 4246–4253.
24. Li, B.Z., Huang, Z., Cui, Q.Y., Song, X.H., Du, L., Jeltsch, A., Chen, P., Li, G., Li, E. and Xu, G.L. (2011) Histone tails regulate DNA methylation by allosterically activating de novo methyltransferase. *Cell Res.*, **21**, 1172–1181.
25. Gowher, H. and Jeltsch, A. (2002) Molecular enzymology of the catalytic domains of the Dnmt3a and Dnmt3b DNA methyltransferases. *J. Biol. Chem.*, **277**, 20409–20414.
26. Jia, D., Jurkowska, R.Z., Zhang, X., Jeltsch, A. and Cheng, X. (2007) Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature*, **449**, 248–251.
27. Jurkowska, R.Z., Anspach, N., Urbanke, C., Jia, D., Reinhardt, R., Nellen, W., Cheng, X. and Jeltsch, A. (2008) Formation of nucleoprotein filaments by mammalian DNA methyltransferase Dnmt3a in complex with regulator Dnmt3L. *Nucleic Acids Res.*, **36**, 6656–6663.
28. Jurkowska, R.Z., Rajavelu, A., Anspach, N., Urbanke, C., Jankevicius, G., Ragozin, S., Nellen, W. and Jeltsch, A. (2011) Oligomerization and binding of the DNMT3A DNA methyltransferase to parallel DNA molecules, heterochromatic localization and role of DNMT3L. *J. Biol. Chem.*, **286**, 24200–24207.
29. Jeltsch, A. and Lanio, T. (2002) Site-directed mutagenesis by polymerase chain reaction. *Methods Mol. Biol.*, **182**, 85–94.
30. Roth, M. and Jeltsch, A. (2000) Biotin-avidin microplate assay for the quantitative analysis of enzymatic methylation of DNA by DNA methyltransferases. *Biol. Chem.*, **381**, 269–272.
31. Jurkowska, R.Z., Siddique, A.N., Jurkowski, T.P. and Jeltsch, A. (2011) Approaches to enzyme and substrate design of the murine Dnmt3a DNA methyltransferase. *Chembiochem*, **12**, 1589–1594.
32. Laird, C.D., Pleasant, N.D., Clark, A.D., Sneed, J.L., Hassan, K.M., Manley, N.C., Vary, J.C. Jr, Morgan, T., Hansen, R.S. and Stoger, R. (2004) Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. *Proc. Natl Acad. Sci. USA*, **101**, 204–209.
33. Zhang, Y., Rohde, C., Tierling, S., Jurkowski, T.P., Bock, C., Santacruz, D., Ragozin, S., Reinhardt, R., Groth, M., Walter, J. *et al.* (2009) DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution. *PLoS Genet.*, **5**, e1000438.
34. Flaus, A. and Richmond, T.J. (1998) Positioning and stability of nucleosomes on MMTV 3'LTR sequences. *J. Mol. Biol.*, **275**, 427–441.
35. 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.
36. Jeltsch, A., Roth, M. and Friedrich, T. (1999) Mutational analysis of target base flipping by the EcoRV adenine-N6 DNA methyltransferase. *J. Mol. Biol.*, **285**, 1121–1130.
37. Gowher, H., Loutchanwoot, P., Vorobjeva, O., Handa, V., Jurkowska, R.Z., Jurkowski, T.P. and Jeltsch, A. (2006) Mutational analysis of the catalytic domain of the murine Dnmt3a DNA-(cytosine C5)-methyltransferase. *J. Mol. Biol.*, **357**, 928–941.
38. Jeong, S., Liang, G., Sharma, S., Lin, J.C., Choi, S.H., Han, H., Yoo, C.B., Egger, G., Yang, A.S. and Jones, P.A. (2009) Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. *Mol. Cell Biol.*, **29**, 5366–5376.
39. Sharma, S., De Carvalho, D.D., Jeong, S., Jones, P.A. and Liang, G. (2011) Nucleosomes containing methylated DNA stabilize DNA methyltransferases 3A/3B and ensure faithful epigenetic inheritance. *PLoS Genet.*, **7**, e1001286.
40. Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S.W., Chen, H., Henderson, I.R., Shinn, P., Pellegrini, M., Jacobsen, S.E. *et al.* (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell*, **126**, 1189–1201.
41. Chodavarapu, R.K., Feng, S., Bernatavichute, Y.V., Chen, P.Y., Stroud, H., Yu, Y., Hetzel, J.A., Kuo, F., Kim, J., Cokus, S.J. *et al.* (2010) Relationship between nucleosome positioning and DNA methylation. *Nature*, **466**, 388–392.

# Function and disruption of DNA Methyltransferase 3a cooperative DNA binding and nucleoprotein filament formation

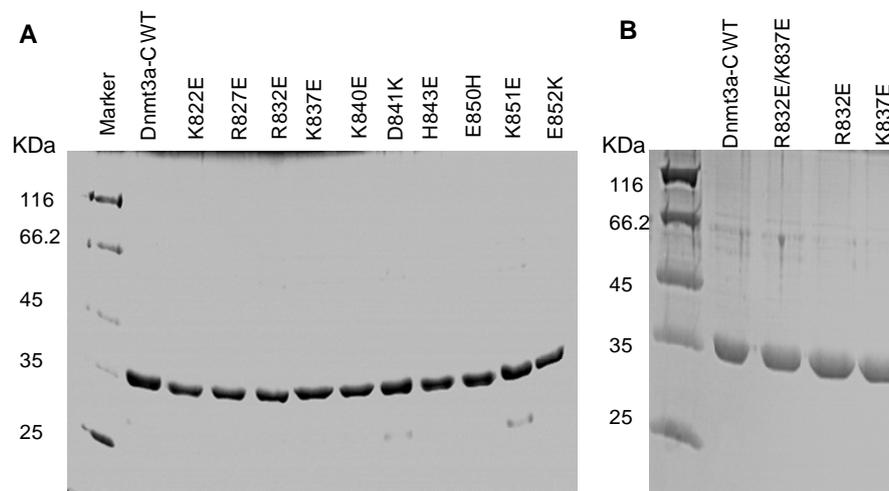
Arumugam Rajavelu, Renata Z. Jurkowska, Jürgen Fritz & Albert Jeltsch

## Supplemental figures and legends

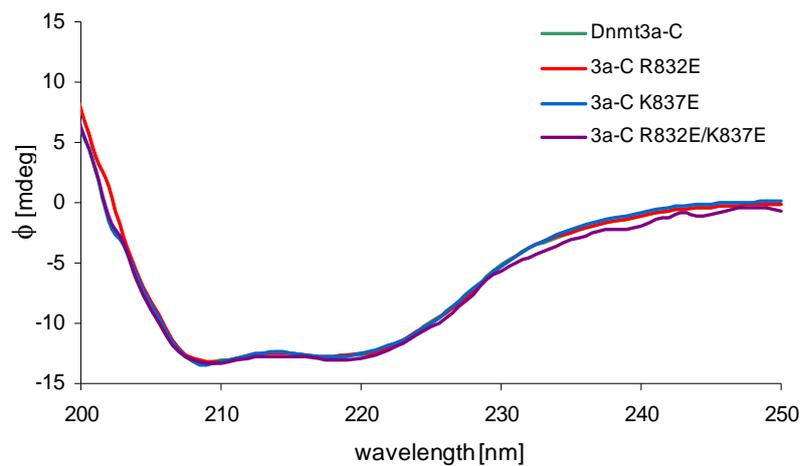
Suppl. Fig. 1. DNA sequencing results of all ten mutants confirming the introduced mutations and absence of secondary mutations.



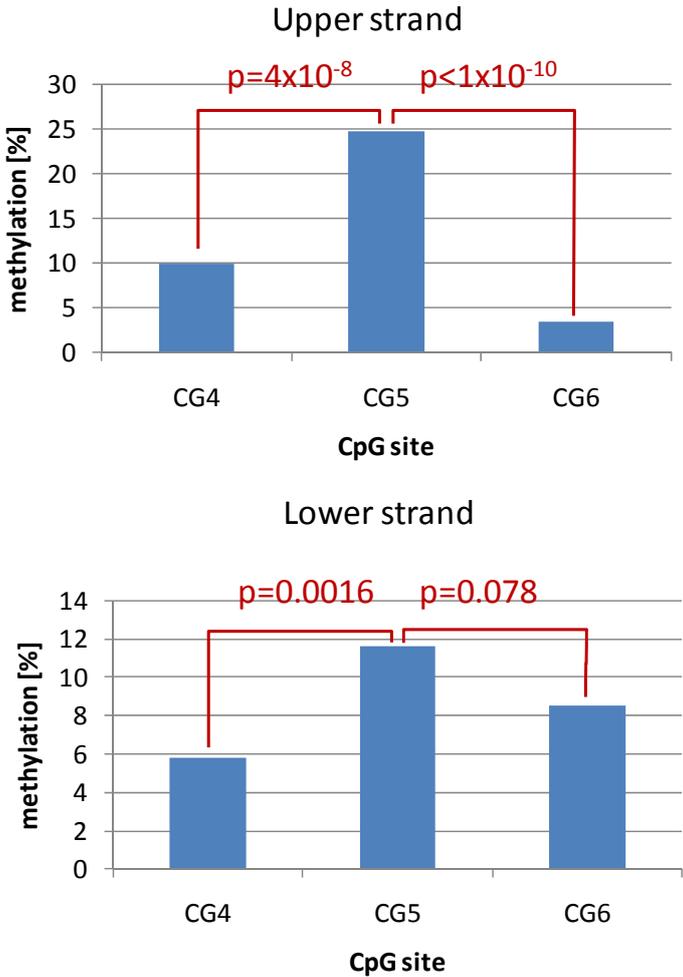
Suppl. Fig. 2A and B. Purification of Dnmt3a-C wild type and the interface mutants. The images show 12% SDS-PAGE gels stained with colloidal Coomassie.



Suppl. Fig. 3. Circular dichroism (CD) spectra of purified Dnmt3a-C wild type and the R832E, K837E and R832E/K837E mutants. CD experiments were carried out using 10  $\mu$ M of the proteins in a Jasco J-810 spectropolarimeter with a 0.1 mm cuvette in buffer containing 20 mM Hepes, 200 mM KCl, 6 % glycerol, 0.2 mM DTT and 1 mM EDTA. The spectra of the variants were scaled to the wild type spectrum to correct for concentration differences and allow better comparison. The matching shape of all curves indicates that the secondary structure composition and folding of all proteins is very similar.

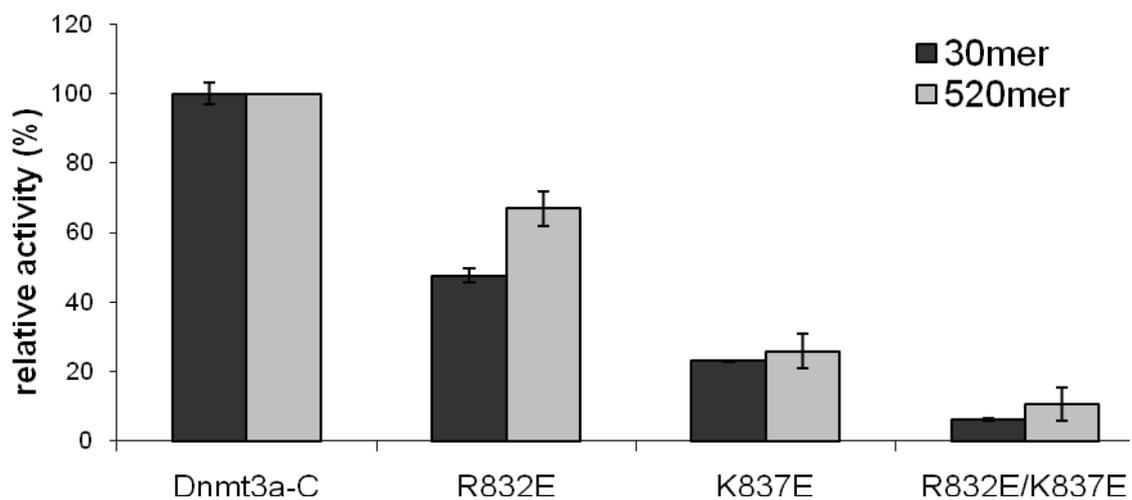


Suppl. Fig. 4. P-values for the methylation peaks at CpG site 5 in Fig. 1A

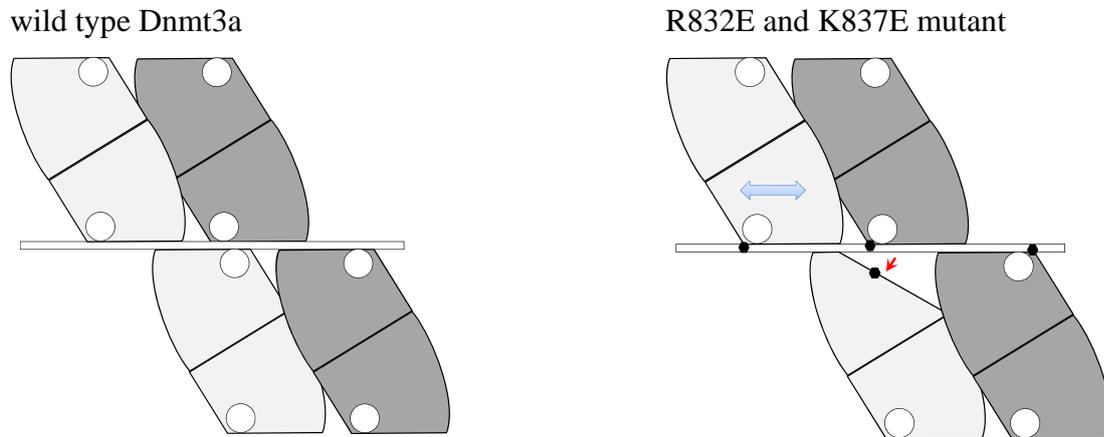


P-values were calculated using Poisson distribution. They assess if the methylation at site 5 is significantly higher than the methylation observed at the sites 4 and 6, by calculating the probability to obtain methylation levels as observed at CpG site 5, assuming the methylation of CpG sites 4 or 6 is expected. The total number of analyzed clones was 185.

Suppl. Fig. 5: Comparison of the catalytic activities of the R832E, K837E and R832E/K837E double mutant on the 520mer DNA substrate and on the 30mer oligonucleotide substrate. To allow for better comparison, the activity of the wild type enzyme with each substrate was set as 100%. The 30mer data were taken from Fig. 5.



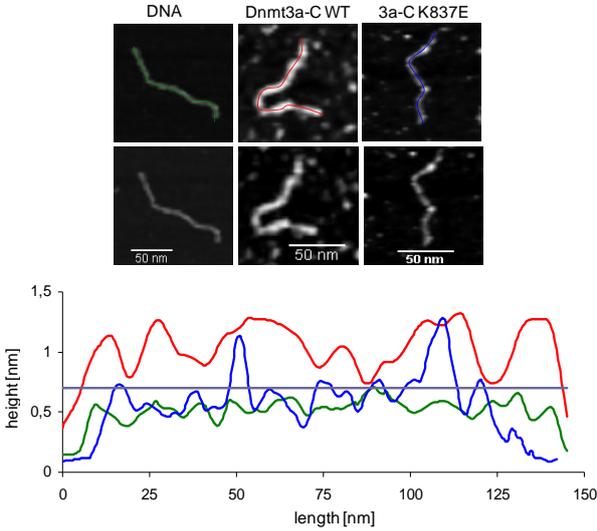
Suppl. Fig. 6: Interpretation of the hairpin bisulfite methylation pattern shown in Fig. 1.



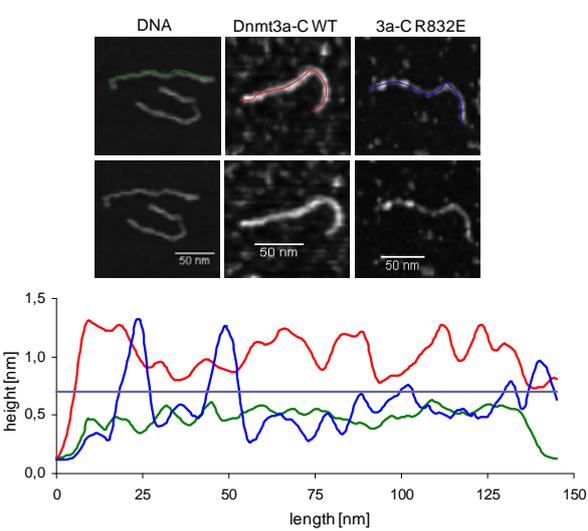
The picture shows a schematic drawing of the possible arrangement of two Dnmt3a tetramers on the hairpin bisulfite substrate. White spheres indicate the regions of the active sites. Mutations are represented by black hexagons. The repulsion caused by the close approximation of two mutations from adjacent tetramers leads to a conformational change in one subunit of the left complex (red arrow) and loss of methylation of the corresponding target site (CpG site 5 in the lower strand in Fig. 1). Having lost the interaction with the right tetramer, the left complex is more mobile and can reach more sites in the upper DNA strand (indicated by the blue double arrow). For details cf. the text of the main manuscript.

Suppl. Fig. 7A and B. Additional examples of SFM images of free DNA, DNA covered with Dnmt3a-C protein and DNA bound with R832E and K837E. For details cf. the legend to Fig. 7.

**A**



**B**



# Oligomerization and Binding of the Dnmt3a DNA Methyltransferase to Parallel DNA Molecules

## HETEROCHROMATIC LOCALIZATION AND ROLE OF Dnmt3L<sup>\*§</sup>

Received for publication, April 26, 2011, and in revised form, May 10, 2011. Published, JBC Papers in Press, May 12, 2011, DOI 10.1074/jbc.M111.254987

Renata Z. Jurkowska<sup>†1</sup>, Arumugam Rajavelu<sup>†1</sup>, Nils Anspach<sup>§2</sup>, Claus Urbanke<sup>¶</sup>, Gytis Jankevicius<sup>||</sup>, Sergey Ragozin<sup>‡</sup>, Wolfgang Nellen<sup>§</sup>, and Albert Jeltsch<sup>†3</sup>

From the <sup>†</sup>Biochemistry Laboratory, School of Engineering and Science, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany, the <sup>§</sup>Abt. Genetik, CINSaT, Universität Kassel, Heinrich-Plett-Str. 40, 34132 Kassel, Germany, the <sup>¶</sup>Medizinische Hochschule, Abteilung Strukturanalyse OE 8830, Carl Neuberg Str. 1, 30625 Hannover, Germany, and the <sup>||</sup>MoLife Graduate Program, School of Engineering and Science, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany

Structural studies showed that Dnmt3a has two interfaces for protein-protein interaction in the heterotetrameric Dnmt3a/3L C-terminal domain complex: the RD interface (mediating the Dnmt3a-3a contact) and the FF interface (mediating the Dnmt3a-3L contact). Here, we demonstrate that Dnmt3a-C forms dimers via the FF interface as well, which further oligomerize via their RD interfaces. Each RD interface of the Dnmt3a-C oligomer creates an independent DNA binding site, which allows for binding of separate DNA molecules oriented in parallel. Because Dnmt3L does not have an RD interface, it prevents Dnmt3a oligomerization and binding of more than one DNA molecule. Both interfaces of Dnmt3a are necessary for the heterochromatic localization of the enzyme in cells. Overexpression of Dnmt3L in cells leads to the release of Dnmt3a from heterochromatic regions, which may increase its activity for methylation of euchromatic targets like the differentially methylated regions involved in imprinting.

The coordinated gene expression during development and cell differentiation is regulated by epigenetic signals comprising DNA methylation, post-translational modification of histone tails, and non-coding RNAs (1). Among them, DNA methylation is a key epigenetic process involved in the control of gene activity (2–3), parental imprinting (4–6), X-chromosome inactivation (7–10), and maintenance of the genome integrity and stability through protection against endogenous retroviruses and transposons (11). Aberrant DNA methylation patterns are associated with several human diseases, including cancer (12–17). In mammals, DNA methylation patterns are established during embryogenesis by the action of the *de novo* DNA methyltransferases Dnmt3a and Dnmt3b, together with their regulatory factor, Dnmt3L (18–19). In addition, the Dnmt3a and Dnmt3b enzymes have a role in the preservation of DNA meth-

ylation at heterochromatin regions (20–21). Disruption of the Dnmt3a or Dnmt3b gene in mice is lethal (22), whereas Dnmt3L knock-out mice are viable and do not show discernable morphological abnormalities (23–25), indicating that Dnmt3a and Dnmt3b are functional in the absence of Dnmt3L.

Dnmt3a and Dnmt3b are nuclear proteins, which stably associate with chromatin containing methylated DNA (21), including mitotic chromosomes, and localize to pericentromeric heterochromatin (26–28). Dnmt3 proteins comprise two parts: a large multidomain N-terminal part of variable size, which has regulatory and targeting functions and a C-terminal catalytic part. The N-terminal part of Dnmt3a and Dnmt3b contains two defined domains: a cysteine-rich ADD (ATR<sub>X</sub>-DNMT3-DNMT3L) domain, which binds to H3 tails unmodified at K4 (29–30) and a PWWP domain, which binds to H3 trimethylated at K36 (31) and is essential for the heterochromatin localization of the enzymes (27–28, 31). The C-terminal parts of Dnmt3 enzymes resemble prokaryotic DNA-(cytosine C5)-methyltransferases and harbor the active centers of the enzymes and are active in an isolated form (32). Dnmt3L is catalytically inactive, directly interacts with the catalytic domains of Dnmt3a and Dnmt3b and stimulates their activity both *in vivo* and *in vitro* (25, 33–37).

The structure of a complex consisting of the C-terminal parts of Dnmt3a and Dnmt3L (Dnmt3a-C/Dnmt3L-C) revealed a linear heterotetramer, in which the two Dnmt3L subunits are positioned at the edges and two Dnmt3a molecules in the center (Fig. 1A) (38–39). The Dnmt3a C-terminal domain provides two interfaces for protein-protein contacts: one hydrophobic FF interface at the Dnmt3a-C/3L-C contact formed by the stacking interaction of four phenylalanine residues (two from each subunit) and one polar RD interface at the Dnmt3a-C/3a-C contact mediated by a hydrogen bonding network between arginine and aspartate residues. The RD interface forms the DNA binding site, while both interfaces are essential for cofactor S-adenosyl-L-methionine binding and catalytic activity of the enzyme (38, 40). We demonstrate here that Dnmt3a-C oligomerizes in a reversible manner. Initial dimer formation occurs through the FF interface, further oligomerization via the RD interface of the FF dimers. Because each RD interface constitutes an independent DNA binding site, Dnmt3a-C oligomers can bind to two or more DNA molecules, oriented next to

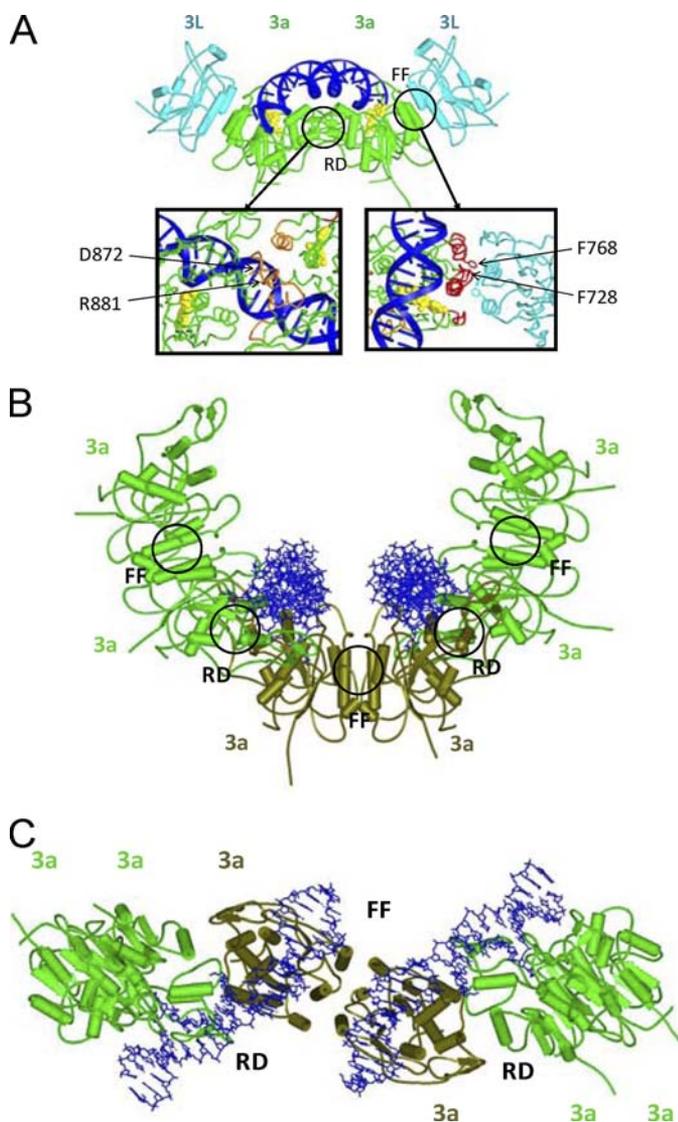
\* This work was supported by the DFG (JE 252/6).

§ The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Figs. S1–S7.

<sup>†</sup> Both authors contributed equally to this work.

<sup>‡</sup> Present address: DME Nanotechnologie GmbH, Am Listholze 82, Hannover, Germany.

<sup>3</sup> To whom correspondence should be addressed: Albert Jeltsch School of Engineering and Science, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany. Tel.: 49-421-200-3247; Fax: 49-421-200-3249; E-mail: a.jeltsch@jacobs-university.de.



**FIGURE 1. Models of Dnmt3a-C/3L-C heterotetramer and Dnmt3a-C oligomers in complex with DNA.** *A*, model of the Dnmt3a/Dnmt3L C-terminal domain tetramer, colored in green for Dnmt3a-C and in cyan for Dnmt3L-C. The cofactor *S*-adenosyl-L-methionine is shown in yellow and the modeled DNA in blue. The close up of the FF and RD interfaces is shown with the relevant amino acid residues labeled. *B* and *C*, modeling of Dnmt3a-C oligomers formed by the interaction of FF-dimers via their RD interface. The figure shows a hexameric complex as an example; the three FF-dimeric building blocks are colored light green and dark olive green. The two DNA molecules bound at the RD interfaces in the Dnmt3a-C hexamer are oriented roughly in parallel. *Panel B* shows a view along the axis of the bound DNA molecules, *panel C* shows a view perpendicular to the axis of the Dnmt3a oligomer.

each other, roughly in parallel. Dnmt3L prevents this oligomerization, because it lacks the RD interface, and it therefore precludes binding of Dnmt3a to more than one DNA molecule. The formation of Dnmt3a oligomers is required for the localization of the enzyme to heterochromatin, and binding of Dnmt3L changes the subnuclear localization of Dnmt3a, which could be one mechanism describing how Dnmt3L stimulates the methylation of imprinted loci.

## EXPERIMENTAL PROCEDURES

**Mutagenesis, Protein Expression, and Protein Purification**—The sequences encoding C-terminal domains of the murine Dnmt3a (residues 608–908) and Dnmt3L (residues 208–421)

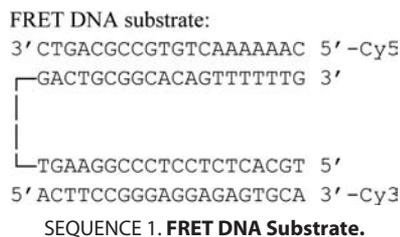
were subcloned as N-terminal His<sub>6</sub> fusion proteins into pET-28a vector (Novagen) (32, 36). Selected Dnmt3a-C and Dnmt3L-C variants were generated using the megaprimer site-directed mutagenesis method as described previously (41). Introduction of the mutations was confirmed by restriction marker site analysis and DNA sequencing. Protein expression was carried out in *Escherichia coli* BL21 (ΔDE3) pLysS (Novagen). Cells were grown at 32 °C, induced with 1 mM IPTG at  $A_{600\text{ nm}}$  of approximately 0.5 and harvested after 3 h. The proteins were purified at high  $\mu\text{M}$  concentrations using Ni-NTA-agarose. Each protein was purified at least twice, and the purity of the preparations was estimated to be greater than 98% (supplemental Fig. S4A). The concentrations of the proteins were determined by UV spectrophotometry using extinction coefficients of  $39290\text{ M}^{-1}\text{ cm}^{-1}$  and  $40330\text{ M}^{-1}\text{ cm}^{-1}$  for Dnmt3a-C and Dnmt3L-C, respectively. Concentration determinations were confirmed by densitometric analysis of Coomassie-stained SDS-polyacrylamide gels. The folding of the mutant proteins was verified by circular dichroism spectroscopy as described (42), indicating that the secondary structure of the R881A variant was indistinguishable from wild-type Dnmt3a-C. The F728 variants were both folded as demonstrated by the strong negative ellipticity at 220 nm, which is indicative of  $\alpha$ -helix content (supplemental Fig. S4B). Both F728 variants displayed a small but significant shift in the ascending limb of the spectrum around 200 nm, suggesting some deviations from the wt folding. The aggregation of Dnmt3a-C shown in Fig. 2A and supplemental Fig. S1A and recovery of enzyme activity from aggregated Dnmt3a-C was only done for illustration. All other experiments reported in this study were conducted with Dnmt3a-C proteins that never went through aggregation, because they were purified at lower protein concentration.

**Analytical Ultracentrifugation**—Analytical ultracentrifugation was done with an An50-Ti 8-place rotor in a Beckman-Coulter model XL-A centrifuge equipped with UV absorption optics. Sedimentation velocity experiments were run at 4 °C and 50 krpm in a buffer containing 10 mM HEPES pH 7.2, 0.2 M KCl, 1.34 M glycerol (10%), 1 mM EDTA, and 0.2 mM DTT. They were evaluated using the program SEDFIT, which transforms the measured data into a diffusion corrected differential sedimentation coefficient distribution (43–44). For every sedimenting species the sedimentation constant was read off the corresponding maximum of this distribution. All distributions were corrected to water at 20 °C using density and viscosity data and partial specific volumes incorporated into SEDNTERP.<sup>4</sup>

**Fluorescence Spectroscopy**—Fluorescence spectroscopy was carried out in a Cary Eclipse fluorimeter (Varian) using a double-labeled two flank oligonucleotide substrate obtained by annealing of three individual oligonucleotides, one 50mer that could hybridize with two 20mers (one of them labeled by Cy3, one by Cy5): 50mer: 5'-TGC ACT CTC CTC CCG GAA GTG TTC TTC GTA GAC TGC GGC ACA GTT TTTTG-3'; Cy3-20-L: 5'-ACT TCC GGG AGG AGA GTG CA-3'-Cy3; Cy5-20-R: Cy5-5'-CAA AAA ACT GTG CCG CAG TC-3'. Control

<sup>4</sup> D. B. Hayes, T. Laue, and J. Philo, unpublished data.

## Oligomerization of Dnmt3a



substrates only containing the Cy3 or Cy5 probe were generated using unlabeled versions of the 20mers. 20-L: 5'-ACT TCC GGG AGG AGA GTG CA-3'; 20-R: 5'-CAA AAA ACT GTG CCG CAG TC-3'.

The DNA substrate used for Fluorescence Resonance Energy Transfer (FRET)<sup>5</sup> experiments was obtained by annealing of the 50mer with one 20-L and one 20-R oligonucleotide. It provides two double-stranded 20 bps stretches, which are separated by a 10 nucleotide single-stranded part, and contain either two (Cy3 and Cy5) or one fluorophore (only Cy3 or only Cy5) (Sequence 1).

The FRET DNA was annealed by mixing all three individual strands at equal concentration (20  $\mu$ M), heating to 86 °C for 5 min, and slowly cooling to ambient temperature. The annealed substrate was purified using Qiagen nucleotide purification kit. FRET experiments were carried out in the presence and absence of Dnmt3a-C proteins in buffer (20 mM HEPES pH 7.5, 1 mM EDTA, 100 mM KCl) containing 100 nM FRET DNA. For the protein-bound sample, 200 nM Dnmt3a-C proteins complex were added and incubated with the DNA at ambient temperature for 20 min to allow for complex formation. The 3a/3L complex was prepared by incubation of Dnmt3a-C and Dnmt3L-C (final concentration of 200 nM each) on ice for 20 min prior to the incubation with DNA. FRET was studied by using Cy3 excitation at 540 nm. Cy5 excitation in control reaction was at 640 nm. Emission and excitation slits were always set at 5 nm. In all experiments, 5 separate scans were taken and averaged. All experiments were carried out at least in duplicate; usually 3–4 repeats were performed.

**Scanning-force Microscopy**—Scanning force microscopy (SFM) experiments were carried out using a 509mer DNA fragment derived from the human CG island upstream of the human *SUHW1* gene and purified catalytic domains of Dnmt3a. DNA-protein filaments were generated in a total volume of 30  $\mu$ l by incubating 12 nM DNA with 200 nM Dnmt3a protein in 50 mM HEPES (pH 7.5), 250 mM NaCl, 1 mM EDTA, and 100  $\mu$ M of sinefungin (Sigma). Samples were incubated for 1 h at room temperature to allow for DNA binding. 1  $\mu$ l of the complex solution was mixed with 9  $\mu$ l of 10 mM MgCl<sub>2</sub> solution and deposited on freshly cleaved mica (Plano, Wetzlar), allowed to adhere for 40 s and then washed with 1 ml of bi-distilled water. The sample was then dried using compressed air. Protein-DNA filaments were observed by tapping mode in air using a Multimode AFM with a Nanoscope III controller (Digital Instruments, Santa Barbara, CA). We used NSTNCHF silicon cantilevers (Nascatec, Stuttgart) with a nominal spring constant of 50 N/m and a resonance frequency of  $\sim$ 150 kHz. All

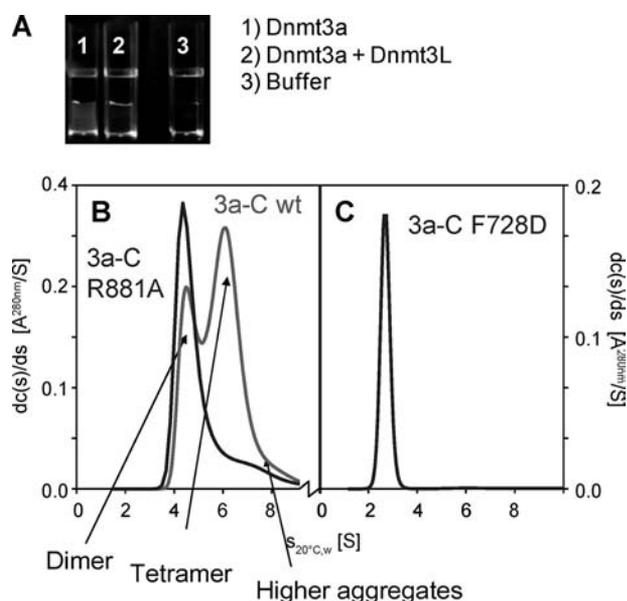
images were obtained with a scanning speed of 0.75–1 Hz and a resolution of 512  $\times$  512 pixels. Experiments were carried out blinded, *i.e.* the operator of the microscope did not know the data obtained in prior experiments. To remove background slope, raw images were flattened using the Nanoscope software. DNA protein complexes were regarded as a filament if the height exceeded 150% of the height observed for DNA molecule alone and were at least 20 nm wide. Filaments were evaluated using the section tool of the Nanoscope V6r12 software.

**Cell Culture and Laser Scanning Microscopy**—The expression construct for murine full length Dnmt3a fused to yellow fluorescent protein (YFP) was obtained from Dr. G. Xu (Shanghai) (45). To prepare a cyan fluorescent protein (CFP) fusion of Dnmt3L, human full-length Dnmt3L (AP001753) was cloned into the pECFP-C1 vector (Invitrogen), using XhoI and EcoRI sites. The interface variants of full length Dnmt3a and Dnmt3L were created in a similar way as the interface variants of the C-terminal domains (see above). NIH3T3 cells were grown in DMEM with 10% (v/v) fetal calf serum and 2 mM L-glutamine at 37 °C in 5% (v/v) CO<sub>2</sub>. Cells (1–2  $\times$  10<sup>5</sup>) were transfected in 6-well plates using FuGENE 6 (Roche, Basel, Switzerland; 1  $\mu$ g of total plasmid DNA per well). In co-transfection experiments equal amounts of Dnmt3a-YFP and Dnmt3L-CFP encoding plasmids were used. Transfected NIH3T3 cells were fixed in 4% (w/v) paraformaldehyde. Confocal images were taken using a Carl Zeiss LSM510 (Jena, Germany; software version 3.0). For the analysis of the co-transfection experiments, the observed localization pattern of Dnmt3a was divided into three categories: spotty, spotty plus diffused and diffused; 100, 54, or 73 cells were analyzed for the Dnmt3a wt, Dnmt3a-Dnmt3L wt, and Dnmt3a-Dnmt3L F261D mutant, respectively.

## RESULTS

**Oligomerization of Dnmt3a-C Depends on Both Interfaces and Is Disrupted by Dnmt3L-C**—We have noticed for several years, that Dnmt3a-C tends to aggregate during purification at the dialysis step when the salt concentration is reduced from 500 to 200 mM KCl if the protein concentration is too high (Fig. 2A and supplemental Fig. S1A). Recently, we observed that the aggregated enzyme preparations could be re-dissolved after dilution and then display almost 80% of the catalytic activity of an enzyme preparation that never went through aggregation (supplemental Fig. S1B). The observation that recovery of enzyme activity did not require refolding indicates that the aggregation process is reversible, and it probably occurs through native protein/protein interfaces. The structure of the Dnmt3a-C/3L-C tetramer shows that Dnmt3a-C has two such interfaces (Fig. 1A), the FF and RD interface. The conservation of key residues of the FF interface between all members of the Dnmt3 family (supplemental Fig. S2), as well as modeling, suggested that Dnmt3a could replace Dnmt3L at the Dnmt3a-3L interface (the FF interface) to form another Dnmt3a/3a contact. Hence, the C-terminal domain of Dnmt3a possesses two putative interfaces for self-interaction, leading to the possibility of forming linear Dnmt3a-C oligomers (Fig. 1, B and C and supplemental Fig. S2B). Dnmt3a-C precipitation experiments in the presence of Dnmt3L-C showed that Dnmt3L-C considerably reduced

<sup>5</sup> The abbreviations used are: FRET, fluorescence resonance energy transfer; YFP, yellow fluorescent protein; CFP, cyan fluorescent protein; SFM, scanning force microscopy.



**FIGURE 2. Oligomerization of Dnmt3a-C.** *A*, reversible aggregation of Dnmt3a-C during protein purification is prevented in the presence of Dnmt3L-C. The figures show samples of Dnmt3a-C (40  $\mu$ M) or Dnmt3a-C/3L-C (40  $\mu$ M each) after reduction of salt concentration from 500 mM KCl to 200 mM KCl by 2 h of dialysis. *B* and *C*, differential sedimentation coefficient ( $S_{20^{\circ}\text{C},\text{w}}$ ) distributions obtained from sedimentation velocity experiments in the analytical ultracentrifuge with Dnmt3a-C wt (2B, red), R881A (2B, blue) and F728D (2C) mutants.

aggregation of Dnmt3a-C (Fig. 2A and supplemental Fig. S1A). The amino acid sequence alignment of Dnmt3a and Dnmt3L shows that Dnmt3L-C lacks an RD interface (supplemental Fig. S2A), suggesting that the RD interface plays an important role in the reversible aggregation of Dnmt3a-C. This conclusion is also in agreement with the salt dependence of the aggregation, because the RD interface is dominated by ionic contacts which strengthen during reduction of salt concentration.

Using analytical ultracentrifugation, we investigated in detail the oligomeric state of the purified wild type Dnmt3a-C that never went through aggregation. The experiment revealed one fraction sedimenting with an  $S_{20^{\circ}\text{C},\text{water}}$ -value of 4.5–4.6 S (Fig. 2B), which corresponds to a dimer with a frictional ratio of 1.35–1.4. The frictional ratio of a folded protein lies between  $\sim 1.2$  for a globular structure and 1.5–1.7 for extended or elongated structures (46). A second, major fraction of Dnmt3a-C sedimented at higher S values ( $\sim 6$ –6.2 S) corresponding to a Dnmt3a-C tetramer with elongated structure (frictional ratio  $1.6 \pm 0.3$ ). Higher aggregates were also observed, depending on the initial concentration of the preparation. In contrast, in ultracentrifugation experiments with purified Dnmt3a-C/3L-C complex at comparable concentrations only the 2:2 heterotetramer of  $\sim 5.4$  S was detected (40), indicating that Dnmt3L-C prevents oligomerization of Dnmt3a-C. These observations are in agreement with published data obtained with Dnmt3a2, an isoform of Dnmt3a also containing most of the N-terminal part, showing that Dnmt3a2 forms large structures of heterogeneous sizes, although binding of Dnmt3L to Dnmt3a2 leads to the formation of specific heterotetrameric complexes (37). To analyze the potential self-interaction of Dnmt3L-C via its FF interface, purified Dnmt3L-C was also studied by analytical ultra-

centrifugation and showed a sedimentation constant of 2.35 S, which corresponds to a monomeric state with a frictional ratio of 1.35 (supplemental Fig. S3).

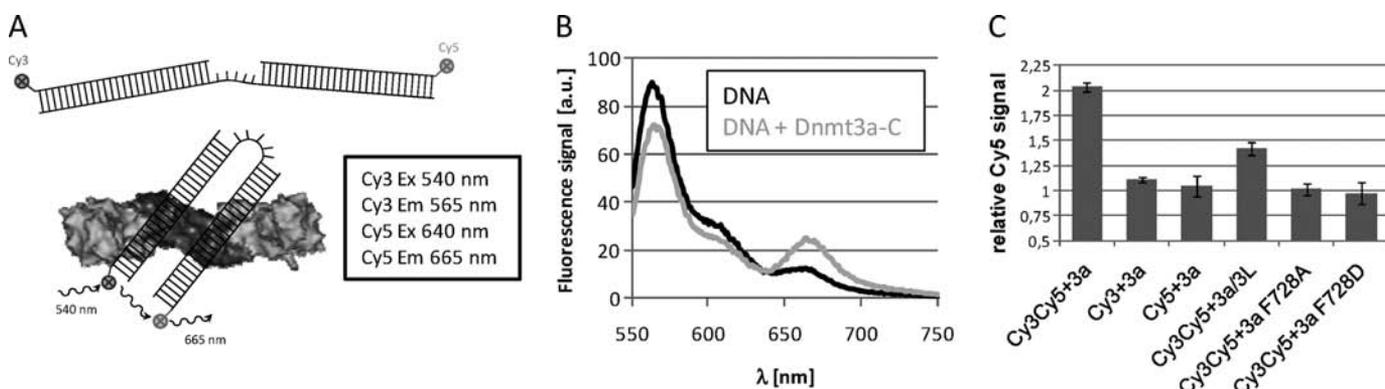
To study the role of the interfaces in Dnmt3a function, we mutated the Arg-881 at the RD interface to alanine (R881A) and the Phe-728 at the FF interface to alanine (F728A) and to aspartic acid (F728D), assuming that the F to D exchange, which introduces a charged residue instead of a hydrophobic one, would result in a stronger phenotype than the F to A mutation. The protein variants were expressed, purified and their folding confirmed by circular dichroism spectroscopy (supplemental Fig. S4). As expected, DNA binding experiments showed that the RD interface constitutes the DNA binding site of Dnmt3a-C homooligomers, because the R881A exchange abrogated DNA binding, while the F728A and F728D variants bound DNA with even slightly better affinity than the wild type enzyme (supplemental Fig. S5). These results replicate previous observations made with the corresponding heterotetrameric Dnmt3a-C/3L-C complexes (40). We have also shown before that all three Dnmt3a-C variants are catalytically inactive (40).

In the ultracentrifugation experiments, the Dnmt3a-C R881A mutant retained its ability to dimerize ( $S_{20^{\circ}\text{C},\text{water}} = 4.3$ –4.4 S) (Fig. 2B). However, tetramers as observed for the wild type protein were not detected. The dimeric state of the R881A mutant demonstrates that the FF interface is fully functional and mediates dimerization of Dnmt3a. In contrast, the F728D mutation converted the protein into a monomer ( $S_{20^{\circ}\text{C},\text{water}} = 2.6$ –2.7 S, frictional ratio 1.5, Fig. 2C), indicating that the intact RD interface does not support self-interaction in absence of DNA. The efficient DNA binding of the FF variants suggests that dimerization across the RD interface can occur in presence of DNA, because DNA binding requires the cooperation of both subunits at the RD interface (38). The Dnmt3a-C F728A variant behaved as a dimer (data not shown), indicating that the F to A exchange did not fully disrupt the FF interface, such that dimerization was still possible under the conditions of the analytical ultracentrifugation, which was performed at relatively high protein concentrations.

Taken together, we show that the FF interface is the primary interface for Dnmt3a-C homodimerization, because the R881A mutant retained its ability to dimerize (via the FF interface), whereas the F728D mutation converted the dimer into a monomer. This finding is in agreement with the structure, because the FF interface is hydrophobic while the RD interface is hydrophilic in nature, and hydrophobic protein/protein interfaces are known to be more stable (47–48). Further reversible oligomerization of Dnmt3a-C is mediated by the salt-dependent interaction of Dnmt3a-C FF dimers via their RD interfaces. This process is prevented by Dnmt3L-C, because it does not contain an RD interface. The requirement for the Dnmt3a-C homodimer to first dissociate before Dnmt3L-C/3a-C complexes can form also explains the long preincubation times of separately purified Dnmt3a-C and Dnmt3L-C needed to reach a maximum stimulation of Dnmt3a-C activity (37). The formation of Dnmt3a-C/3L-C complexes is supported by the monomeric state of Dnmt3L-C.

*Binding of Parallel DNA Molecules by Dnmt3a-C Oligomers—* We modeled the structure of a Dnmt3a-C/3a-C contact via the

## Oligomerization of Dnmt3a



**FIGURE 3. Analysis of binding of parallel DNA molecules by Dnmt3a-C oligomers.** *A*, schematic setup of the FRET assay. In free form, the distance of both ends of the FRET substrate is too large for FRET to occur. After binding of the two double-stranded regions to adjacent DNA binding sites in a Dnmt3a-C oligomer, the ends approach each other and FRET can be established. *B*, fluorescence emission spectra of the free FRET DNA (black line) and of FRET DNA-Dnmt3a-C complexes (grey line) after Cy3 excitation. FRET is indicated by the strong increase in Cy5 emission at 665 nm in the presence of Dnmt3a-C. *C*, Cy5 emission of the FRET DNA after addition of Dnmt3a-C, Dnmt3a-C interface variants or Dnmt3a-C/3L-C complex divided by the corresponding Cy5 emission of the FRET obtained for free DNA. In addition, as a control, experiments with single labeled FRET DNA (either only carrying Cy3 or Cy5 label) were carried out, but no signal change was observed after addition of Dnmt3a-C. Excitation was at 540 nm. Error bars represent the error of the mean derived from the variation of repeated experiments.

FF interface by superimposing a Dnmt3a-C subunit onto one of the Dnmt3L-C molecules of the Dnmt3a-C/3L-C heterotrimer structure. This superposition can be done with good confidence, because of the strong structural and amino acid sequence similarity between Dnmt3a-C and Dnmt3L-C. By combining these Dnmt3a-C/3a-C FF dimers obtained by modeling with Dnmt3a-C/3a-C dimers formed via the RD interface taken from the Dnmt3a-C/3L-C structure, a model of a Dnmt3a-C hexamer containing three FF and two RD interfaces could be generated. In principle, this hexamer could be further extended. Because each RD interface constitutes an independent DNA binding site, such a hexamer could bind two DNA molecules (Fig. 1, *B* and *C*). Interestingly, the modeling suggests that the DNA binding sites are oriented roughly in parallel to each other in a distance of  $\sim 3$ – $4$  nm, which would not allow for one DNA molecule to bind side by side to both sites simultaneously. Instead, binding of two DNA molecules oriented roughly in parallel to each other could occur (Fig. 1, *B* and *C* and [supplemental Fig. S2B](#)).

To study the potential binding of two parallel DNA molecules by the Dnmt3a-C oligomer in solution, we developed a FRET assay, using a DNA substrate that consists of two 20 base pair double-stranded regions separated by a 10 nucleotide single-stranded region, which are flexible. The ends of the DNA were labeled with fluorophores that constitute a FRET donor/acceptor pair (Cy3/Cy5 with a Förster radius of about 56 Å). In the annealed oligonucleotide, the average distance of both ends is large and does not permit FRET. However, after binding of the FRET substrate to adjacent RD interfaces of a Dnmt3a-C oligomer, the ends should approach each other, and FRET could be established (Fig. 3*A*). Indeed, with the free DNA only a small Cy5 signal was observed after excitation of Cy3 at 540 nm (Fig. 3*B*). This signal does not correspond to FRET, but could be attributed to direct excitation of Cy5 (data not shown). However, after addition of Dnmt3a-C to the FRET substrate, we observed a strong increase in Cy5 fluorescence after Cy3 excitation, which is indicative of FRET (Fig. 3*B*). Control experiments confirmed that an increase in Cy5 fluorescence after

addition of Dnmt3a-C was not observed after direct excitation of Cy5 (data not shown) and with identical substrates that only contained the Cy3 or Cy5 probe (Fig. 3*C*). These controls confirm that FRET is detected and parallel DNA binding occurs. Next we investigated the effect of Dnmt3L-C on parallel DNA binding of Dnmt3a-C. A strongly reduced FRET signal was seen after addition of a Dnmt3a-C/3L-C complex when compared with Dnmt3a-C at same concentration. The Dnmt3a-C/3L-C complex was generated by preincubation of Dnmt3a-C and Dnmt3L-C. Under these conditions, it is unlikely that all Dnmt3a-C oligomers are resolved, which explains the residual FRET effect that was still observed. The FRET signal was also lost with the Dnmt3a-C mutants that disrupt the FF interface, such that DNA binding is possible, but no oligomerization of Dnmt3a-C occurs (Fig. 3*C*). These results confirm that Dnmt3a-C oligomers can bind to two DNA molecules oriented in parallel. Dnmt3L-C or mutations at the FF interface of Dnmt3a-C disrupt the oligomers and, thereby, prevent parallel binding of more than one DNA molecule.

Next, we studied the binding of long DNA molecules to Dnmt3a-C oligomers by SFM. We have shown previously that the Dnmt3a-C/3L-C complex polymerizes on DNA, using purified Dnmt3a-C and a 509 bp PCR product, which can be visualized accurately by SFM (40). We now observed that Dnmt3a-C forms DNA-nucleoprotein filaments in the absence of Dnmt3L as well (Fig. 4). However, additional structural features like sharp curvatures and lariats were also seen with Dnmt3a-C, indicating that protein bound stretches from different parts of a long single DNA molecule can interact after forming DNA loops (Fig. 4*C*). Interestingly, about one third of the observed complexes contained two DNA molecules connected by more or less extended protein-DNA filaments (Fig. 4*D*). These connected synaptonemal-like complexes were probably formed by a direct interaction between Dnmt3a-C molecules bound to different DNA molecules oriented in parallel (Fig. 4*E*). Such complexes were not observed in the previous Dnmt3a-C/3L-C experiments. They document the ability of Dnmt3a-C oligomers to bind simultaneously to two DNA molecules oriented in parallel.

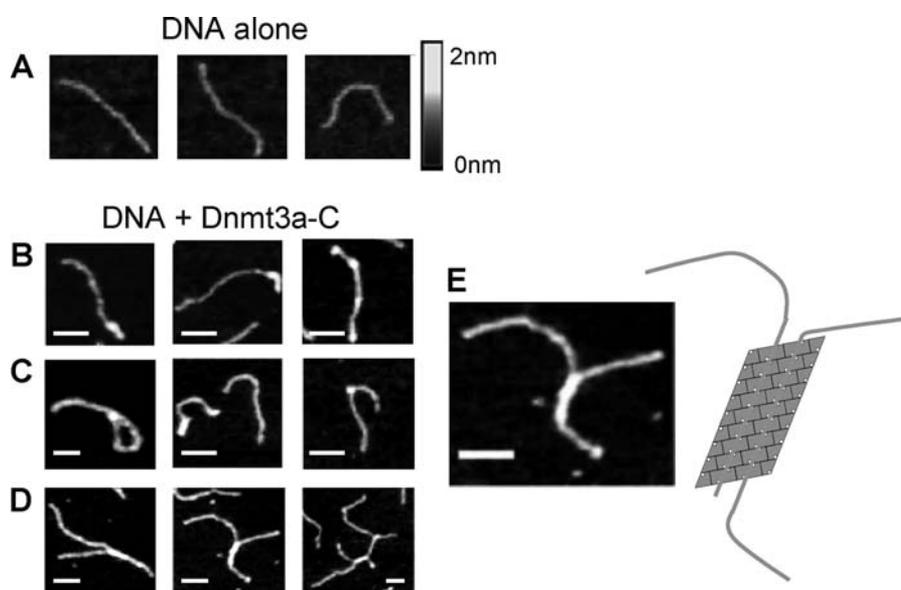


FIGURE 4. **Scanning force microscopy of Dnmt3a-C-DNA filaments.** *A*, images showing free DNA. *B–D*, images showing examples of Dnmt3a-DNA filaments. Regular filaments (*B*), structures like lariats and sharp curvatures (*C*), as well as branched, synaptonemal-like complexes (*D*) can be observed. *E*, magnification of an exemplary picture from *D* and a schematic representation of a DNA oligomer bound to two parallel DNA molecules. The scale bars correspond to 50 nm.

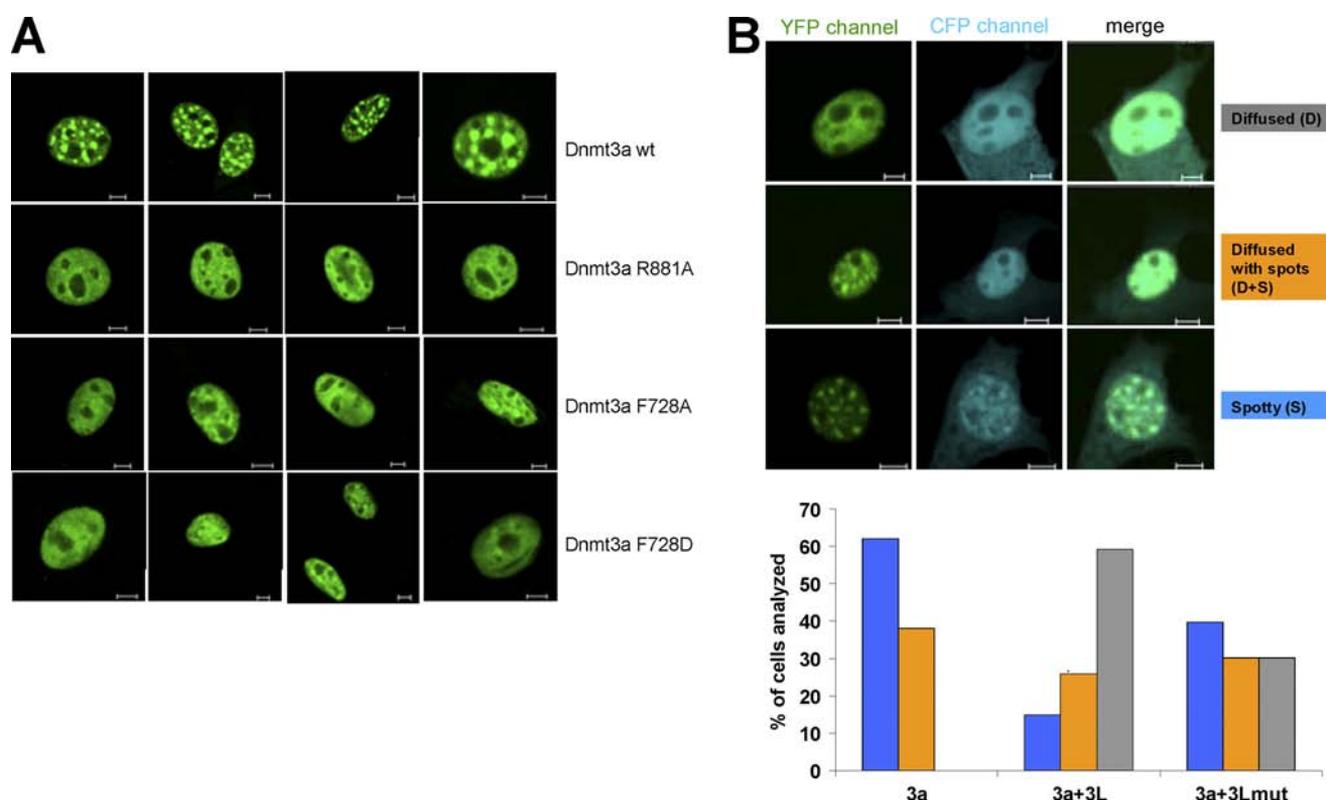


FIGURE 5. **Subnuclear localization of Dnmt3a and Dnmt3L.** *A*, localization of the YFP-tagged full-length Dnmt3a wt and its interface variants R881A, F728A and F728D in NIH3T3 cells. *B*, localization of the YFP-tagged full-length Dnmt3a wt in NIH3T3 cells in the presence and absence of Dnmt3L. Three patterns of nuclear localization were observed: diffused, spotty plus diffused and spotty. The bar diagram shows a quantification of the Dnmt3a localization in the presence and absence of Dnmt3L wt. Additional experiments were conducted with the Dnmt3L F261D interface mutant that partially lost its ability to interact with Dnmt3a. The *blue bars* represent cells with spotty Dnmt3a localization, the *orange bars* spotty plus diffused, and the *gray bars* correspond to diffused localization pattern. The scale bars represent 5  $\mu$ m.

*Subnuclear Localization of Dnmt3a Depends on Both Interfaces and Is Altered by Dnmt3L*—It had been shown before that Dnmt3a localizes to heterochromatic foci in mouse cell lines (26–28, 31) and that this effect depends on both the PWWP and the catalytic domain of Dnmt3a (27–28, 31). To investigate

if oligomerization of Dnmt3a has an effect on its cellular and subnuclear localization, we studied the cellular distribution of YFP-tagged full-length Dnmt3a and its interface mutants R881A, F728A, and F728D in NIH 3T3 cells. As expected, wild type Dnmt3a was located exclusively in subnuclear spots (Fig. 5

## Oligomerization of Dnmt3a

and supplemental Figs. S6 and S7), which were previously shown to co-localize with pericentromeric heterochromatin and reflect the localization of the endogenous enzyme (26–28, 31). The F728A, F728D, and R881A variant fusion proteins showed similar expression levels as the wild type protein (as judged by the intensity of the fluorescence signal) and exclusive nuclear localization as well. However, all variants displayed a great change in the sub-nuclear localization pattern. Instead of the enrichment in distinct heterochromatic spots characteristic of the wild-type protein, a diffuse staining was observed, with the fluorescence signal almost uniformly distributed within the nucleus (Fig. 5A). These results indicate that the ability of Dnmt3a to bind to heterochromatic regions was lost or greatly reduced with the interface mutants indicating that both interfaces are required for the stable localization of Dnmt3a to the heterochromatic foci. A similar change in localization was observed after complete removal of the catalytic domain of Dnmt3a (27).

We next investigated the subnuclear localization of CFP-fused Dnmt3L, which showed a homogenous nuclear and some weak cytoplasmic localization (supplemental Fig. S6). We confirmed that the CFP and YFP channels showed no cross-talk in these experiments (supplemental Fig. S7) and studied nuclear localization of Dnmt3a and Dnmt3L after co-transfection of Dnmt3L-CFP with Dnmt3a-YFP. Dnmt3L showed weak localization to heterochromatic spots in some cells, indicating some targeting influence of Dnmt3a (Fig. 5B). This finding reproduces results of an earlier study (49). Interestingly, we observed a major change in the distribution of Dnmt3a in the nucleus in the absence and presence of Dnmt3L (Fig. 5B). When Dnmt3L was not present, Dnmt3a was localized to the heterochromatic foci in the great majority of cells (62%). Only a minority of cells showed spotty Dnmt3a localization overlaid with some diffuse background (38%), but we never observed cells with uniformly diffused nuclear Dnmt3a localization without any spots. In contrast, in the presence of Dnmt3L, the majority of cells showed a diffused nuclear Dnmt3a localization pattern (60%) and only a minor fraction of cells displayed the spotty localization characteristic for the wild-type Dnmt3a (15 and 25% for spotty and spotty plus diffused, respectively). We performed similar experiments with the Dnmt3L F261D interface mutant, which showed strongly reduced interaction with Dnmt3a *in vitro* (40) and observed a much reduced effect of the Dnmt3L variant on the localization of Dnmt3a. In the presence of Dnmt3L F261D the majority of cells retained the spotty or spotty and diffused Dnmt3a localization pattern and only in a minor fraction of cells Dnmt3a was found diffusely distributed in the nucleus. These results indicate that Dnmt3L interaction reduces the heterochromatic localization of Dnmt3a and redistributes it into euchromatin.

## DISCUSSION

The Dnmt3a C-terminal domain employs two interfaces for protein/protein contacts in the structure of the Dnmt3a/Dnmt3L heterotetrameric complex: one FF interface and one RD interface. We show here that the FF interface also supports self-interaction of Dnmt3a-C and that Dnmt3a-C dimers can form protein oligomers using their RD interfaces. Since each

RD interface constitutes a potential DNA binding site, oligomerization of Dnmt3a creates several DNA binding sites. Our modeling of a Dnmt3a oligomer predicted binding to several DNA molecules oriented next to each other roughly in parallel (Fig. 1 and supplemental Fig. S2B). We document this unusual mode of DNA binding to Dnmt3a-C oligomers by binding experiments performed in solution (FRET) and imaging (SFM). Binding of two DNA molecules in parallel orientation so far has been observed only with few proteins (one noticeable example being RecA (50)) and such DNA binding mode was not anticipated for Dnmt3a.

Dnmt3a has been shown to bind tightly to chromatin (21). It is engaged in two interactions with the H3 tail, mediated by its ADD domain binding to the end of the H3 tail unmodified at K4 (29–30), and by its PWWP domain binding to H3K36me3, which was found to be essential for heterochromatic localization (31). In addition, Dnmt3a polymerizes on the DNA (38, 40) and we show here that Dnmt3a oligomers bind to several DNA molecules oriented in parallel. Since heterochromatin is defined as a region of very high DNA density, the ability of Dnmt3a oligomers to bind to parallel DNA close to each other may contribute to the targeting of the enzyme to such DNA dense regions. This hypothesis is supported by our finding that heterochromatic localization of Dnmt3a was lost with non-oligomerizing Dnmt3a mutants affected at the FF interface and in the presence of Dnmt3L, which disrupts oligomerization and multiple DNA binding because it lacks an RD interface. The structure of heterochromatic DNA is not known. However, at least one model of the 30 nm chromatin fiber, the helical ribbon model (51), proposes an almost parallel arrangement of the linker DNA regions in a spacing that would fit to the model of the Dnmt3a oligomers binding to parallel DNA molecules. Hence, the unusual DNA binding mode of Dnmt3a could represent a novel mechanism of heterochromatic targeting.

We have shown that Dnmt3L disrupts the reversible oligomerization of Dnmt3a-C and it interferes with the simultaneous binding to several DNA molecules to Dnmt3a. We investigated the influence of Dnmt3L on the subnuclear localization of Dnmt3a. Strikingly, more than 60% of the cells lost the heterochromatic localization of Dnmt3a in the presence of Dnmt3L, but not when Dnmt3a was co-expressed with the Dnmt3L F261D mutant, which partially lost its ability to interact with Dnmt3a. This observation suggests a novel mode of action of Dnmt3L in setting DNA methylation imprints. In addition to stimulating the activity of Dnmt3a, it might be involved in the release of Dnmt3a from dense heterochromatin to make it available to act at imprinted differentially methylated regions, which are generally located in euchromatin. The expression of Dnmt3L in germ line cells may redistribute Dnmt3a from heterochromatin to euchromatin and, thereby, increase Dnmt3a availability and DNA methylation activity for the generation of DNA methylation imprints.

*Acknowledgment*—We thank Dr. Xiadong Cheng for insightful discussions and remarks.

## REFERENCES

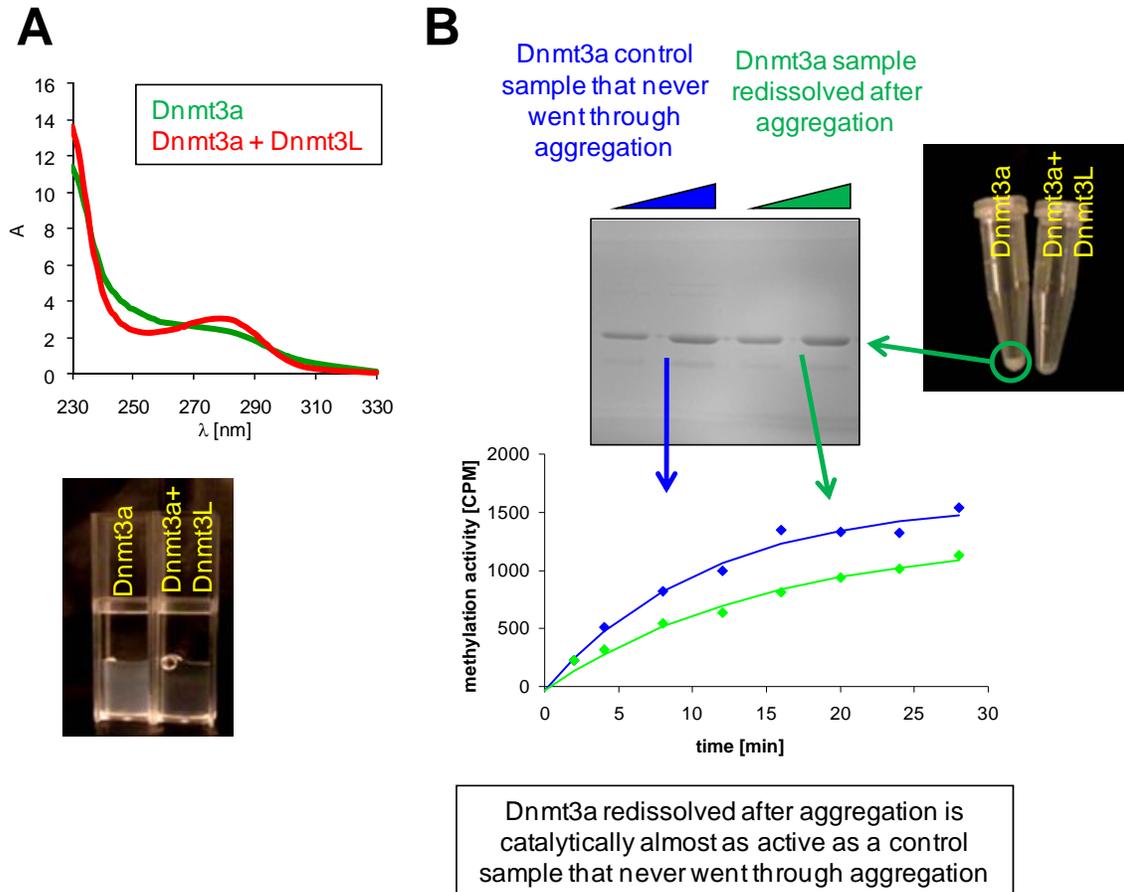
- Allis, C. D., Jenuwein, T., and Reinberg, D. (2007) *Epigenetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Miranda, T. B., and Jones, P. A. (2007) *J. Cell. Physiol.* **213**, 384–390
- Lande-Diner, L., Zhang, J., Ben-Porath, I., Amariglio, N., Keshet, I., Hecht, M., Azuara, V., Fisher, A. G., Rechavi, G., and Cedar, H. (2007) *J. Biol. Chem.* **282**, 12194–12200
- Delaval, K., and Feil, R. (2004) *Curr. Opin. Genet. Dev.* **14**, 188–195
- Hore, T. A., Rapkins, R. W., and Graves, J. A. (2007) *Trends Genet.* **23**, 440–448
- Sha, K. (2008) *Annu Rev. Genomics Hum. Genet.* **9**, 197–216
- Heard, E. (2004) *Curr. Opin. Cell Biol.* **16**, 247–255
- Chang, S. C., Tucker, T., Thorogood, N. P., and Brown, C. J. (2006) *Front Biosci.* **11**, 852–866
- Yen, Z. C., Meyer, I. M., Karalic, S., and Brown, C. J. (2007) *Genomics* **90**, 453–463
- Straub, T., and Becker, P. B. (2007) *Nat. Rev. Genet.* **8**, 47–57
- Howard, G., Eiges, R., Gaudet, F., Jaenisch, R., and Eden, A. (2008) *Oncogene* **27**, 404–408
- Jones, P. A., and Baylin, S. B. (2007) *Cell* **128**, 683–692
- Egger, G., Liang, G., Aparicio, A., and Jones, P. A. (2004) *Nature* **429**, 457–463
- Feinberg, A. P., and Tycko, B. (2004) *Nat. Rev. Cancer* **4**, 143–153
- Rodenhiser, D., and Mann, M. (2006) *Cmaj* **174**, 341–348
- Robertson, K. D. (2005) *Nat. Rev. Genet.* **6**, 597–610
- Esteller, M. (2007) *Nat. Rev. Genet.* **8**, 286–298
- Jurkowska, R., Jurkowski, T. P., and Jeltsch, A. (2011) *Chembiochem* **12**, 206–222
- Klose, R. J., and Bird, A. P. (2006) *Trends Biochem. Sci.* **31**, 89–97
- Dodge, J. E., Okano, M., Dick, F., Tsujimoto, N., Chen, T., Wang, S., Ueda, Y., Dyson, N., and Li, E. (2005) *J. Biol. Chem.* **280**, 17986–17991
- Jeong, S., Liang, G., Sharma, S., Lin, J. C., Choi, S. H., Han, H., Yoo, C. B., Egger, G., Yang, A. S., and Jones, P. A. (2009) *Mol. Cell. Biol.* **29**, 5366–5376
- Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999) *Cell* **99**, 247–257
- Bourc'his, D., Xu, G. L., Lin, C. S., Bollman, B., and Bestor, T. H. (2001) *Science* **294**, 2536–2539
- Bourc'his, D., and Bestor, T. H. (2004) *Nature* **431**, 96–99
- Hata, K., Okano, M., Lei, H., and Li, E. (2002) *Development* **129**, 1983–1993
- Bachman, K. E., Rountree, M. R., and Baylin, S. B. (2001) *J. Biol. Chem.* **276**, 32282–32287
- Chen, T., Tsujimoto, N., and Li, E. (2004) *Mol. Cell. Biol.* **24**, 9048–9058
- Ge, Y. Z., Pu, M. T., Gowher, H., Wu, H. P., Ding, J. P., Jeltsch, A., and Xu, G. L. (2004) *J. Biol. Chem.* **279**, 25447–25454
- Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., and Shirakawa, M. (2009) *EMBO Rep* **10**, 1235–1241
- Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., Fischle, W., and Jeltsch, A. (2010) *Nucleic Acids Res.* **38**, 4246–4253
- Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R. Z., Ragozin, S., and Jeltsch, A. (2010) *J. Biol. Chem.* **285**, 26114–26120
- Gowher, H., and Jeltsch, A. (2002) *J. Biol. Chem.* **277**, 20409–20414
- Chedin, F., Lieber, M. R., and Hsieh, C. L. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 16916–16921
- Chen, Z. X., Mann, J. R., Hsieh, C. L., Riggs, A. D., and Chédin, F. (2005) *J. Cell. Biochem.* **95**, 902–917
- Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H., and Tajima, S. (2004) *J. Biol. Chem.* **279**, 27816–27823
- Gowher, H., Liebert, K., Hermann, A., Xu, G., and Jeltsch, A. (2005) *J. Biol. Chem.* **280**, 13341–13348
- Kareta, M. S., Botello, Z. M., Ennis, J. J., Chou, C., and Chédin, F. (2006) *J. Biol. Chem.* **281**, 25893–25902
- Jia, D., Jurkowska, R. Z., Zhang, X., Jeltsch, A., and Cheng, X. (2007) *Nature* **449**, 248–251
- Cheng, X., and Blumenthal, R. M. (2008) *Structure* **16**, 341–350
- Jurkowska, R. Z., Anspach, N., Urbanke, C., Jia, D., Reinhardt, R., Nellen, W., Cheng, X., and Jeltsch, A. (2008) *Nucleic Acids Res.* **36**, 6656–6663
- Jeltsch, A., and Lanio, T. (2002) *Methods Mol. Biol.* **182**, 85–94
- Gowher, H., Loutchanwoot, P., Vorobjeva, O., Handa, V., Jurkowska, R. Z., Jurkowski, T. P., and Jeltsch, A. (2006) *J. Mol. Biol.* **357**, 928–941
- Schuck, P. (2000) *Biophys J.* **78**, 1606–1619
- Schuck, P., and Rossmanith, P. (2000) *Biopolymers* **54**, 328–341
- 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., and Xu, G. L. (2007) *Mol. Cell. Biol.* **27**, 8748–8759
- Lebowitz, J., Lewis, M. S., and Schuck, P. (2002) *Protein Sci.* **11**, 2067–2079
- Chothia, C., and Janin, J. (1975) *Nature* **256**, 705–708
- Bahadur, R. P., Chakrabarti, P., Rodier, F., and Janin, J. (2003) *Proteins* **53**, 708–719
- Nimura, K., Ishida, C., Koriyama, H., Hata, K., Yamanaka, S., Li, E., Ura, K., and Kaneda, Y. (2006) *Genes Cells* **11**, 1225–1237
- Bell, C. E. (2005) *Mol. Microbiol.* **58**, 358–366
- Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R. R., and Richmond, T. J. (2004) *Science* **306**, 1571–1573

**OLIGOMERIZATION AND BINDING OF THE DNMT3A DNA  
METHYLTRANSFERASE TO PARALLEL DNA MOLECULES,  
HETEROCHROMATIC LOCALIZATION AND ROLE OF  
DNMT3L\***

Renata Z. Jurkowska, Arumugam Rajavelu, Nils Anspach, Claus Urbanke, Gytis Jankevicius, Sergey Ragozin, Wolfgang Nellen & Albert Jeltsch

**Contents:**

Supplemental Figures S1-S7



### Supplemental Figure 1

A) Aggregation of Dnmt3a-C (40  $\mu$ M) occurs during reduction of salt concentration from 500 mM KCl to 200 mM KCl by 2 hours of dialysis. Presence of equal amounts of Dnmt3L-C prevented the aggregation. Aggregation is visible in the UV spectra of the samples and in the turbulent appearance of the Dnmt3a-C solution.

B) Active Dnmt3a can be recovered from aggregated material. Aggregated Dnmt3a-C was dissolved in Dnmt3a-C storage buffer (20 mM Hepes pH 7.2, 200 mM KCl, 1 mM EDTA, 0.2 mM DTT and 10% glycerol). Afterwards the amount of redissolved enzyme was determined and its catalytic activity assessed using an enzyme preparation that never went through precipitation as reference. Kinetics were carried out using 2  $\mu$ M of enzyme, 1  $\mu$ M 30mer oligonucleotide substrate in methylation buffer (20 mM Hepes pH 7.2, 50 mM KCl, 1 mM EDTA, 0.025 mg/ml BSA) at 37°C as described (40). Note that highly active Dnmt3a could be recovered from the precipitated material (with about 80% of specific activity) although the procedure did not include any refolding step.

**A**

→ start of C-terminal domain

```

3a_mouse: CVDLLVGPAAQAAIKEDPWNCYMCGHKGTYGLLRRRKDWPSRLQMFANNHD...QEFDPKVPYPPVPAEKRRKPIRVLSLFDGI : 639
3a_human: CVDLLVGPAAQAAIKEDPWNCYMCGHKGTYGLLRRRKDWPSRLQMFANNHD...QEFDPKVPYPPVPAEKRRKPIRVLSLFDGI : 643
3b_mouse: CLEVLVCAATAEDAKLOEPWSCYMLPQRCHGVLRRRKDWNMRLQDFFTDDELEEFEPKLYPAIPAAKRRPIRVLSLFDGI : 590
3b_human: CLEVLVCAATAEAKLOEPWSCYMLPQRCHGVLRRRKDWNMRLQDFFTSDTGLE...YEAPKLYPAIPAAKRRPIRVLSLFDGI : 584
3L_human: CVDLLVGPPTSCKVHAMSNNVCMCLCLPSSRSGLLQRRKRWRSQLKAFYDRSE...NPLEMFTVPVWRQPPVRLSLEFDI : 199
3L_mouse: CVDLLVGPPTSERINAMACWVCLCLPFSRSGLLQRRKRWREHQLKAFHDEGA...GPMETIKTVSAWKRPVRLSLEFNI : 233

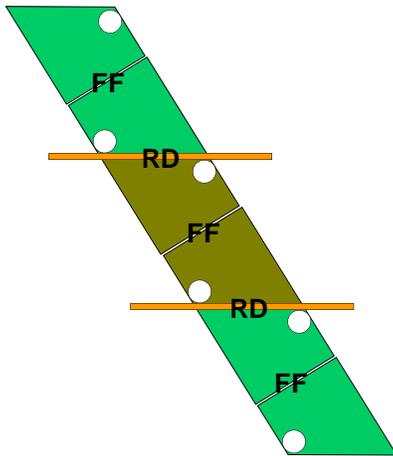
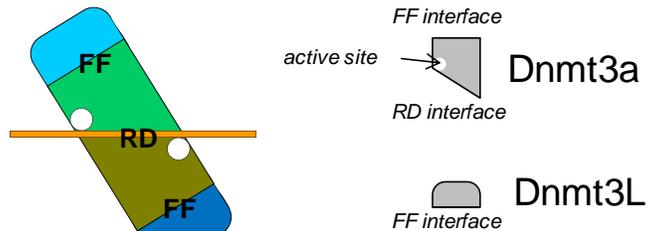
3a_mouse: ATGLLVLDLGIQVDRIYASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIEQWGPFDLVIGGSPCNDSLIVNPAKGLYEG : 722
3a_human: ATGLLVLDLGIQVDRIYASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIEQWGPFDLVIGGSPCNDSLIVNPAKGLYEG : 726
3b_mouse: ATGYLVLKLKELGKVKYVASEVCEESIAVGTVKHEGQIKYVNDVRNITKKNIEEWGPFDLVIGGSPCNDSLIVNPAKGLYEG : 673
3b_human: ATGYLVLKLKELGKVKYVASEVCEESIAVGTVKHEGQIKYVNDVRNITKKNIEEWGPFDLVIGGSPCNDSLIVNPAKGLYEG : 667
3L_human: KK...ELTSLGFLES.....GSDFG...CLKHVVVDVTDIVRQDVEEWGPFDLVYG.....ATFPLGHTCDRP : 255
3L_mouse: DK...VLKSLGFLES.....GSGSGCGTILKYVEDVTNVVRRDVEKMGPFDLVYG.....STCPLGSSCDRC : 291

3a_mouse: TGRLLFFEFYRLLHDARPKEGDDRPFVWLFENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARVFWGNLPGMNRPLASTV : 805
3a_human: TGRLLFFEFYRLLHDARPKEGDDRPFVWLFENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARVFWGNLPGMNRPLASTV : 809
3b_mouse: TGRLLFFEFYHLLNYSRPKEGDRPFVWLFENVVAMKVNDKRDISRFLSNPVMIDAKEVSAHRARVFWGNLPGMNRPLASTV : 756
3b_human: TGRLLFFEFYHLLNYSRPKEGDRPFVWLFENVVAMKVNDKRDISRFLSNPVMIDAKEVSAHRARVFWGNLPGMNRPLASTV : 750
3L_human: PSWYLFQFHRLLQYARPKGSERPFVWLFENVVNLVLNKEDLDVASRFLSEVVTIPDVHGGSTQNAVVRVWSNIPAIRSSRFWAL : 338
3L_mouse: PGWMLFQFHRLLQYALPROESCRPFVWLFENVVNLVLTEDDQETTRRFLQTEAVTLQDVRCDRYQNAVVRVWSNIPGLKS...KEAPL : 373

3a_mouse: NDKLELQECLEHGRIAKFSKVRTITTRSNSIKQKGDQHFVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLAKRQLLGRS : 888
3a_human: NDKLELQECLEHGRIAKFSKVRTITTRSNSIKQKGDQHFVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLAKRQLLGRS : 892
3b_mouse: NDKLELQDCLEFSRTAKLKKVQTIITTKSNSIRQGNQLFPVVMNGKDDVLWCTELERIFGFPAHYTDVSNMGRGARQKLLGRS : 839
3b_human: NDKLELQDCLEYNRIAKLKKVQTIITTKSNSIKQGNQLFPVVMNGKDDVLWCTELERIFGFPAHYTDVSNMGRGARQKLLGRS : 833
3L_human: VSEELSLAQNQSSKLAKWPTKLVKNCPLPLREY...FKYFSTELTSLL... : 387
3L_mouse: TPKEEYLQAGVRSRSKLDAPKVDLLVKNCLPLREY...FKYFSQNSLPL... : 421

3a_mouse: WSVPVIRHLFAPLKEYFACV : 908
3a_human: WSVPVIRHLFAPLKEYFACV : 912
3b_mouse: WSVPVIRHLFAPLKDYFACE : 859
3b_human: WSVPVIRHLFAPLKDYFACE : 853

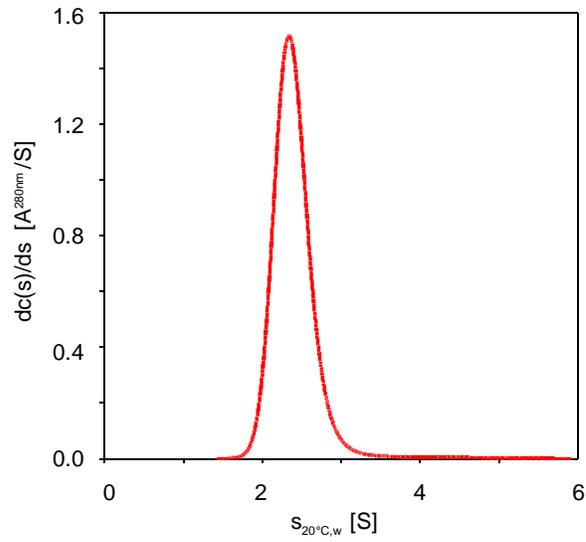
```

**B****C**

## **Supplemental Figure 2**

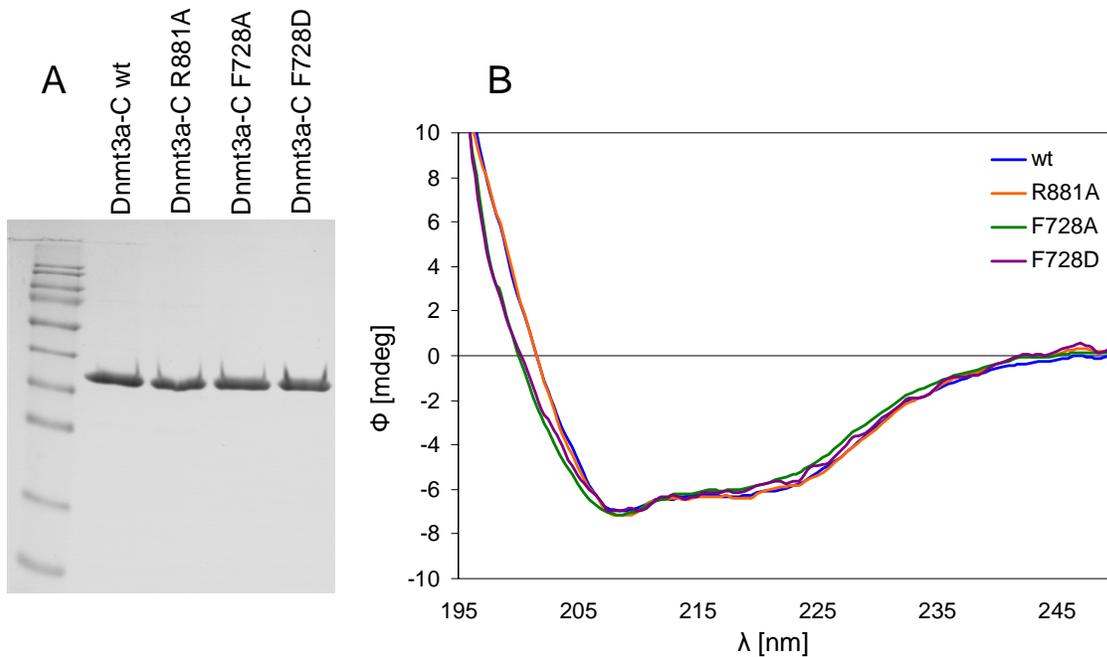
A) Sequence alignment of the C-terminal parts of the human and mouse Dnmt3 enzymes showing the similarity of the FF and RD interfaces. Residues from the RD and FF interfaces are labeled in orange and red, respectively, the main residues involved in the interaction at both interfaces (F728, F768 at the FF interface and R881 and D872 at the RD interface) are shaded in blue.

B and C) Schematic pictures of the multimeric structure of Dnmt3a (B) and the Dnmt3a/3L heterotetramer (C). In either case oligomers are formed by RD interface interaction of dimers formed via the FF interface. Dnmt3a is colored light green and dark olive green, Dnmt3L cyan and blue and bound DNA is shown as orange line. Dnmt3a oligomers can bind to more than one DNA molecule oriented in parallel.



### Supplemental Figure 3

Monomeric state of Dnmt3L-C. Analytical centrifugation analysis of 9  $\mu$ M Dnmt3L-C at 8°C and 45 krpm in 10 mM HEPES pH 7.2, 0.15 M KCl, 1.34 M glycerol (10%), 1 mM EDTA, and 0.2 mM DTT. The figure shows the differential sedimentation coefficient distribution. Dnmt3L-C sedimented with a sedimentation constant of 2.35 corresponding to a monomeric state (frictional ratio 1.35).

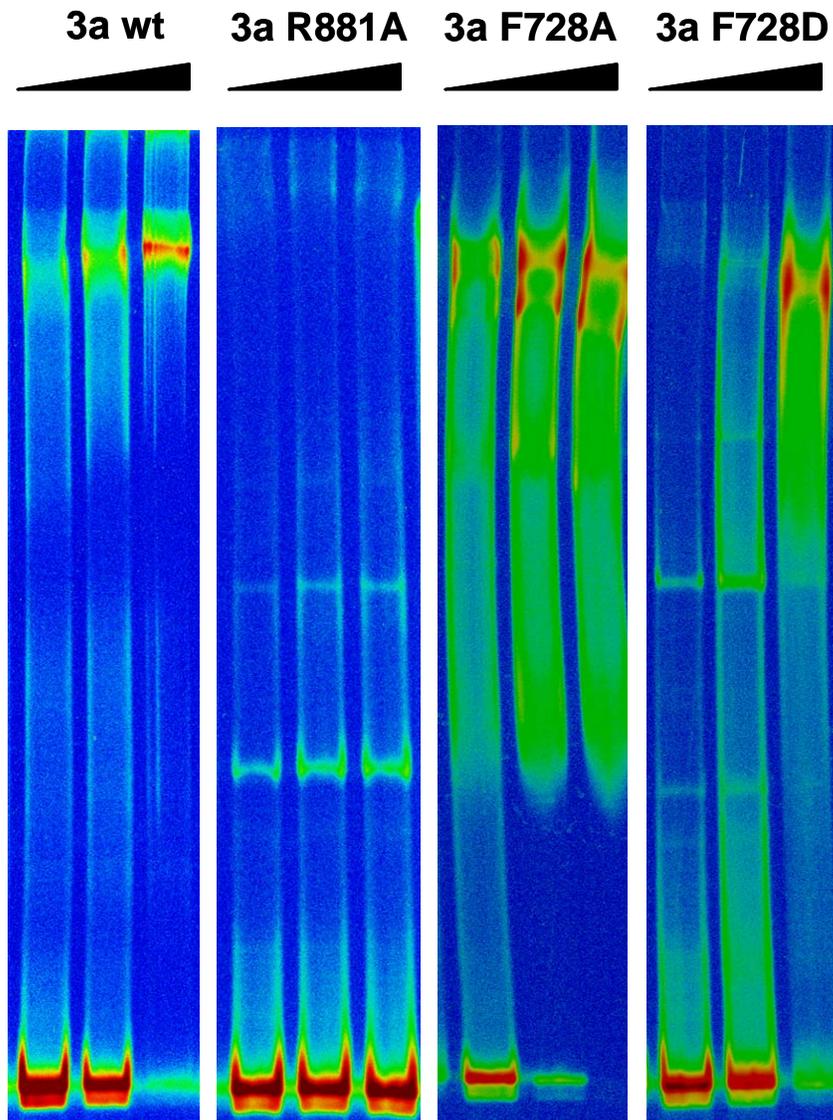


#### Supplemental Figure 4

Purification and folding of Dnmt3a-C wt and its interface variants.

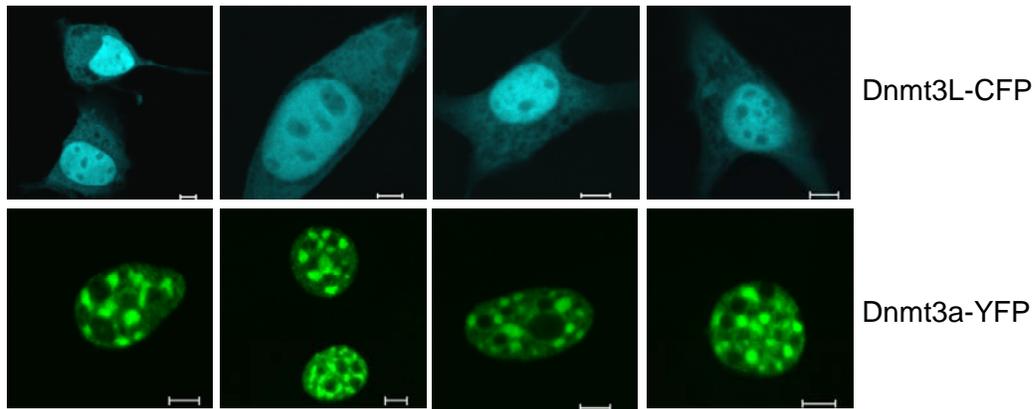
A) SDS-gel stained with colloidal Coomassie showing similar quality and amount of the Dnmt3a-C wt and its interface variants R881A, F728A and F728D. The purity of the proteins was estimated to be 98%.

B) Far UV circular dichroism spectra of the Dnmt3a-C wt and its interface variants. The figure shows superposition of the experimental curves obtained for the wt and mutant proteins. It demonstrates that the folding of the R881A mutant is identical to that of the wt protein. The F728A and F728D mutants show a change in the ascending limb of the spectrum around 200 nm, suggesting some deviation from the wt folding but the CD spectrum clearly indicates that both variants are folded.



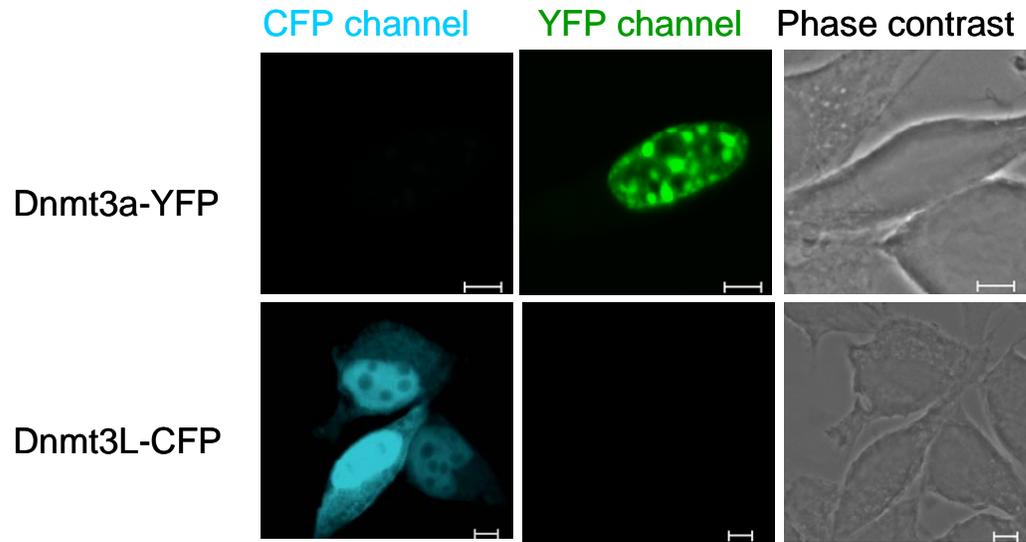
### Supplemental Figure 5

DNA binding and filament formation by Dnmt3a-C wt and its interface mutants. DNA binding of wt Dnmt3a-C, and its R881A, F728A and F728D variants was determined by electromobility shift assays using fluorescently labeled 146-bp long DNA (30 nM) and increasing amounts of proteins (1–10  $\mu$ M) in reaction buffer (20 nM HEPES pH 7.5, 1 mM EDTA, 100 mM KCl, 0.2 mM sinefungin, 0.5 mg/ml BSA) basically as described (38, 40). Sinefungin is an analog of AdoMet that does not allow DNA methylation but supports DNA binding of Dnmt3a-C. The gel was then scanned with a phosphoimager system (Fuji).



### Supplemental Figure 6

Localization of the ectopically expressed, YFP-tagged Dnmt3a and CFP-tagged Dnmt3L in NIH3T3 cells. Dnmt3a is localized uniquely to the nucleus and concentrates in the heterochromatic spots, while Dnmt3L in the absence of Dnmt3a is present both in the nucleus and in the cytoplasm. The scale bars represent 5  $\mu$ m.



**Supplemental Figure 7**

**Localization of Dnmt3a in NIH3T3 cells in the presence and absence of Dnmt3L.** Control experiment showing that under experimental settings used here there is no cross-talk between the CFP and the YFP channels. The scale bars represent 5  $\mu\text{m}$ .

RESEARCH ARTICLE

Open Access

# The inhibition of the mammalian DNA methyltransferase 3a (Dnmt3a) by dietary black tea and coffee polyphenols

Arumugam Rajavelu<sup>2</sup>, Zumrad Tulyasheva<sup>3</sup>, Rakesh Jaiswal<sup>1</sup>, Albert Jeltsch<sup>2\*</sup> and Nikolai Kuhnert<sup>1\*</sup>

## Abstract

**Background:** Black tea is, second only to water, the most consumed beverage globally. Previously, the inhibition of DNA methyltransferase 1 was shown by dietary polyphenols and epi-gallocatechin gallate (EGCG), the main polyphenolic constituent of green tea, and 5-caffeoyl quinic acid, the main phenolic constituent of the green coffee bean.

**Results:** We studied the inhibition of DNA methyltransferase 3a by a series of dietary polyphenols from black tea such as theaflavins and thearubigins and chlorogenic acid derivatives from coffee. For theaflavin 3,3 digallate and thearubigins IC<sub>50</sub> values in the lower micro molar range were observed, which when compared to pharmacokinetic data available, suggest an effect of physiological relevance.

**Conclusions:** Since Dnmt3a has been associated with development, cancer and brain function, these data suggest a biochemical mechanism for the beneficial health effect of black tea and coffee and a possible molecular mechanism for the improvement of brain performance and mental health by dietary polyphenols.

## Background

Black tea is, second only to water, the most consumed beverage globally with an average per capita consumption of around 550 ml per day. The annual production of tea leaves reached a record high in 2008 with a global harvest of 3.75 Mt [1]. Production of dried tea comprises 20% green, 2% oolong and the remainder black. Following black tea, coffee is the third most consumed beverage globally with an annual production of 9.7 Mt and a daily consumption of around 300 ml (data from <http://www.fas.usda.gov/>, obtained 1st March 2011). Strong epidemiological evidence has repeatedly linked the consumption both black tea [2] and coffee [3,4] to a variety of beneficial health effects, among them is the prevention of multifactorial diseases including cancer, cardiovascular disease and neurological disorders as well as a series of psychoactive responses improving

alertness, mood and general mental performance [5-8]. Recently, Unilever made an application for a health claim, in which the black tea beverage should supposedly improve mental alertness and focus, based on studies by Nurk et al. with the activities of the two compounds caffeine and L-theanine as the proposed rationale [9]. While epidemiological studies link two causally unrelated events, e. g. a beneficial health effect with the consumption of a certain diet, with a certain statistical probability, the molecular causes of these epidemiological observations are rarely known. In order to rationalize epidemiological observations, a biological target must be identified that is mechanistically linked to the beneficial health effect reported, as well as the specific molecules contained in the diet that interact with the biological target in question at dietary and physiologically relevant concentrations. The search for such matching pairs of biological targets and dietary compound must be considered an exercise of fishing in the dark, however, where enzymes known to be intimately involved in the area in question need to be systematically screened against secondary metabolites known to be produced by the dietary plant in question.

\* Correspondence: a.jeltsch@jacobs-university.de; n.kuhnert@jacobs-university.de

<sup>1</sup>Chemistry, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany

<sup>2</sup>Biochemistry, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany

Full list of author information is available at the end of the article

Prompted by reports of Fang and co-workers, who have recently reported the inhibition of DNA methyltransferase 1 (Dnmt1) by a series of dietary polyphenols [10] and work by Lee and co-workers on the inhibition of the same enzyme investigating most notably epigallocatechin gallate (EGCG) [11] (the main polyphenolic constituent of green tea) and 5-caffeoyl quinic acid [12] (the main phenolic constituent of the green coffee bean), and Nandakumar, showing the reduction of cellular DNA methylation after admission of (-)-epigallocatechin-3-gallate [13], we decided to screen the interaction of a series of black tea and coffee polyphenols against DNA methyltransferase 3a, another important member of this family of enzymes.

DNA methyltransferases catalyzes methylation of DNA at cytosine residues and play an important role in epigenetic regulation of gene expression, X-chromosome inactivation, genomic imprinting, and development cellular aging and cell differentiation [14,15]. In mammals, DNA methylation is catalyzed mainly by three DNA methyltransferases [15,16]: Dnmt1, Dnmt3a, and Dnmt3b. Dnmt1 has a high preference for hemimethylated DNA and is essential for maintaining the methylation patterns during each round of DNA replication. On the other hand, Dnmt3a and Dnmt3b modify both unmethylated and hemimethylated DNA and are responsible for *de novo* methylation during early development. Errors in DNA methylation contribute to both the initiation and the progression of various cancers [17,18]. In addition, aberrant or missing DNA methylation causes many kinds of diseases which include defects in embryonic development or brain development and neurological defects which are also associated with behavioral changes [19]. Hypermethylation of genes is one of important process in cancer development, typically resulting in the repression of tumor suppressor genes. Preventing the hypermethylation of promoter genes by selective inhibition of methyltransferases could pave a way for cancer treatment [20-22]. Importantly it has been shown that upon use of methyltransferase inhibitors it was possible to reactivate gene silenced by promoter methylation in cancers and thus modulate gene expression. Several efforts are directed at developing small molecules that target DNA methyltransferases and other elements of the machinery, as the proteins that bind to methylated CpG; some are in clinical trials [20-22].

Another important issue of DNA methylation is its function in brain development. Levenson and coworkers showed that Dnmt1 is involved in the formation of hippocampus-dependent long term memory [23]. They found that the promoters for reelin and brain-derived neurotrophic factor (genes implicated in the induction of synaptic plasticity in the adult hippocampus) exhibit

rapid and dramatic changes in cytosine methylation when Dnmt1 activity was inhibited. Moreover, DNA methyltransferase inhibitors like 5-aza-2-deoxycytidine blocked the induction of long term potentiation at Schaffer collateral synapses. Furthermore, Dnmt3a-dependent DNA methylation has been reported to influence transcription of neurogenic genes [24]. Additional studies showed that Dnmt1 and Dnmt3a regulate synaptic function in adult forebrain neurons [25] and Dnmt3a affects plasticity of neurons [26].

Changes in the DNA methylation pattern of regions in the hippocampus are associated with behavioral changes in rat [27]. In addition, Dnmt3a has been recently shown to affect the emotional behaviour [26]. Thus, DNA methylation which is already known to be involved in setting up cellular memory is also involved in brain function. The combination of studies in cell lines and in animal models, coupled with data obtained from post-mortem human material provides compelling evidence that aberrant methylation may contribute to psychiatric diseases like schizophrenia and psychosis [28]. Strong epidemiological evidence suggests that particularly for black tea and green tea there is an inverse relation between intake and significant beneficial effects on patients suffering from psychological disorders [2,5-8]. Currently, no accepted rationale on the molecular level exists that can account for these epidemiological findings. Dnmts are a possible biological target for tea dietary polyphenols suggesting a molecular based rationale for the observed biological activities.

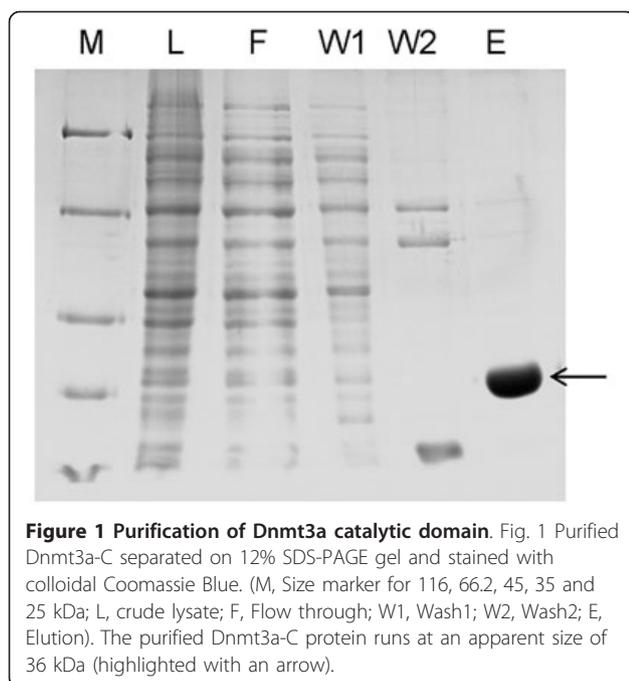
## Results

### Expression and purification of Dnmt3a-C

The catalytic domain of Dnmt3a was expressed and purified following an established protocol [29,30]. The purified protein by Ni-NTA affinity chromatography was >90% homogenous as judged from SDS-PAGE stained with colloidal Coomassie Blue (Figure 1).

### Selection and purification of black tea and coffee polyphenols

Black tea is produced from the young green shoots of the tea plant (*Camellia sinensis*), which are converted to black tea by fermentation [31]. There are two major processes, the 'orthodox' and the 'cut-tear-curl'. In both, the objective is to achieve efficient disruption of the cellular substructure bringing phenolic compounds present in the green tea leaf, mainly flavan-3-ols otherwise known as catechins, into contact with polyphenol oxidases and activating many other enzymes. The catechin substrates are oxidized and extensively transformed into novel dimeric, oligomeric and polymeric compounds. The chemical composition of black tea brew can be divided into (i) a series of well characterized small



**Figure 1 Purification of Dnmt3a catalytic domain.** Fig. 1 Purified Dnmt3a-C separated on 12% SDS-PAGE gel and stained with colloidal Coomassie Blue. (M, Size marker for 116, 66.2, 45, 35 and 25 kDa; L, crude lysate; F, Flow through; W1, Wash1; W2, Wash2; E, Elution). The purified Dnmt3a-C protein runs at an apparent size of 36 kDa (highlighted with an arrow).

molecules including alkaloids (e.g. theobromine and caffeine), carbohydrates and amino acids (including theanine), and a series of glycosylated flavonoids and dimers of catechins including most notably theaflavins together accounting for 30-40% of the dry mass of a typical black tea infusion, and (ii) the heterogeneous and poorly characterized polyphenolic fermentation products accounting for the remaining 60-70% [31]. This material was originally referred to as oxytheotannin and later renamed by Roberts as thearubigins [32].

For this study, we first selected EGCG **N1** and (-)-epigallocatechin **N4** (from green tea, also on occasions found in black tea at low concentrations) as reference compounds. Next we selected the four most common theaflavin derivatives: theaflavin **N2**, theaflavin-3-gallate **N5**, theaflavin 3'-gallate **N3** and theaflavin 3, 3'-digallate **N6** (see Figure 2) [33,34]. All four compounds are found in black tea infusions at concentrations of around 100 mM, making up 2-3% of the total content of dry mass in typical black tea infusion. Theaflavins are structurally closely related the catechins being formal dimers of EGCG obtained through a two electron oxidation followed by C-C bond formation and a benzylic acid type rearrangement leading to the benzotropolone core structure. Next to theaflavins we decided to screen as well two crude thearubigin fractions. We recently proposed that thearubigins contain several thousands polyhydroxylated theaflavin derivatives in equilibrium with their ortho-quinones. Theaflavins were obtained by extraction from black tea infusion followed by purification by preparative HPLC. The purity was assessed by

LC-tandem MS. Thearubigins were obtained from black tea infusions using a protocol developed by Roberts [35].

For coffee polyphenols, we selected a range of naturally occurring and synthetic derivatives of chlorogenic acids. Chlorogenic acids (CGAs) are formally hydroxycinnamate esters of quinic acid with a dietary intake of an estimated 2 g per human per day [36]. We recently reported a total of 70 different chlorogenic acids found in green coffee beans and selected some representative examples containing both caffeic acid and ferulic acid substituents [37-40]. Furthermore, we selected a range of epimers of CGAs produced from the original secondary plant metabolites by roasting of the coffee beans. Representative structures are shown in Figure 2. All CGA derivatives were obtained through chemical synthesis unless stated otherwise.

#### Dnmt3a-C activity and inhibitors screening

The purified Dnmt3a-C was catalytically highly active (Figure 3). For an initial screening of the twenty four inhibitor candidates, Dnmt3a-C DNA methylation kinetics were carried out in the presence of 100  $\mu$ M of compound. Rates of DNA methylation were derived by linear regression of the initial phase of the reaction progress curves. The reaction rates were compared with control reactions carried out after addition of a corresponding volume of DMSO to ensure identical reaction conditions, because DMSO had been shown before to influence the activity of Dnmt3a [41]. As shown in Figure 4, four of the compounds had a substantial inhibitory effect for the *in vitro* Dnmt3a-C activity (N6-N8 and N12). To determine  $IC_{50}$  values, DNA methylation kinetics were carried out in the presence of variable concentrations of the inhibitors, initial slopes derived and the activity profile analysed by fitting of the experimental data to the equation:

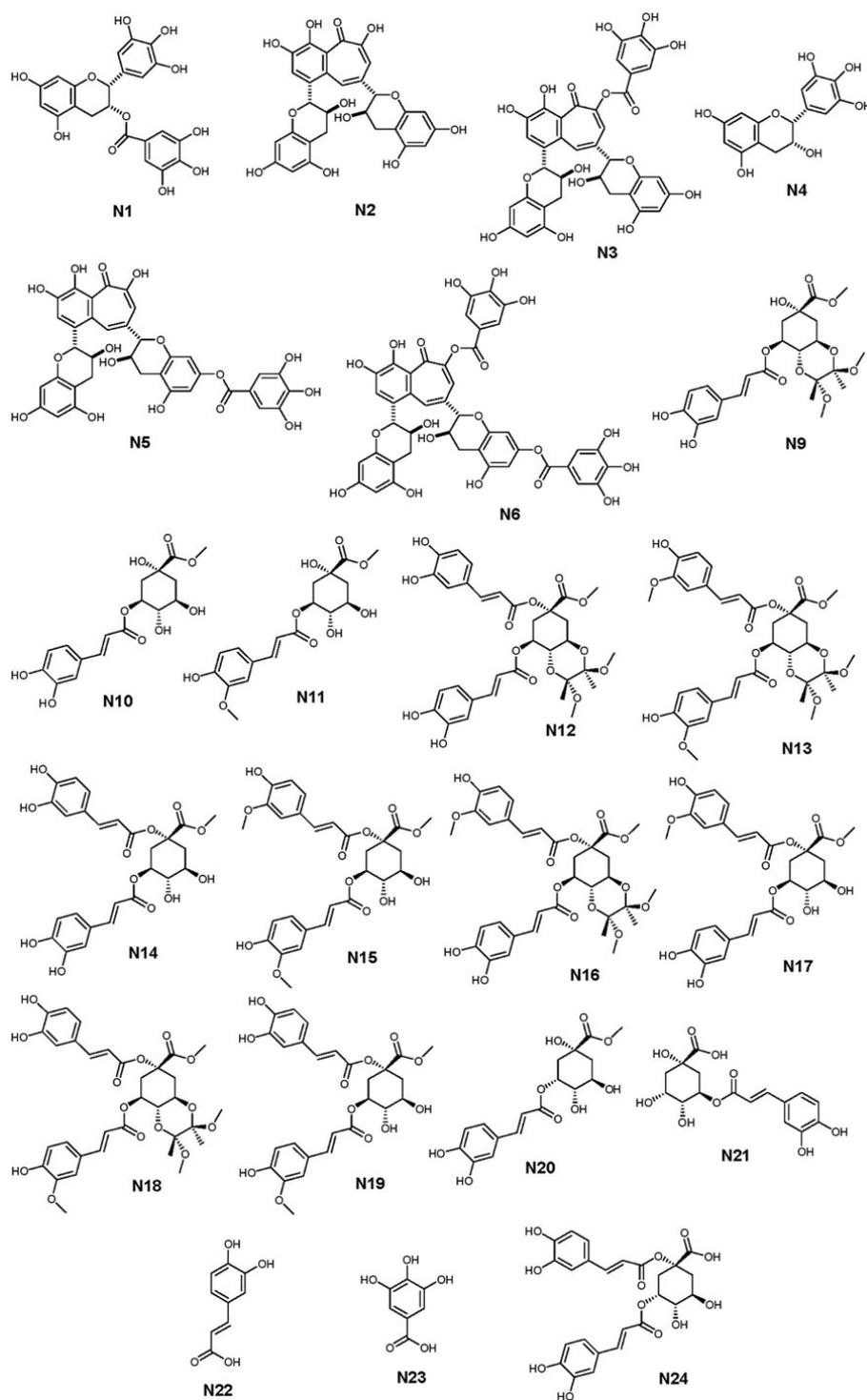
$$A(c_1) = A_0 \times c_1 / (IC_{50} + c_1) + BL$$

with:  $c_1$ , concentration of the inhibitor;  $A(c_1)$ , activity in presence of inhibitor at concentration  $c$ ;  $A_0$ , activity in absence of inhibitor; BL, baseline.

As shown in Figure 5, the  $IC_{50}$  values for the compounds N6-N8 and N12 were all in the lower  $\mu$ M range.

#### Discussion

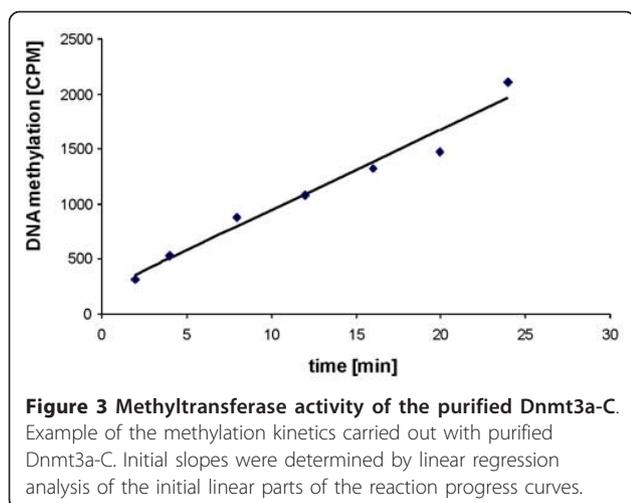
Lee *et al* had showed that caffeic acid and chlorogenic acid inhibit the activity of M.SssI and Dnmt1 and decrease the methylation level at the RAR beta promoter gene in the breast cancer cell lines [12]. Furthermore, they have recently described the inhibition of human Dnmt1 by tea flavanoids such as EGCG, catechin and other flavanoids such as quercetin and myristin, observing  $K_i$  values in the low micromolar range [11]. While Dnmt1 is considered a biological target involved in



**Figure 2** Structures of the compounds tested for Dnmt3a-C inhibition.

cancer development its close relative Dnmt3a, investigated in this study, has been linked to both cancer development and mental performance and health. Therefore, any inhibitory interaction between any of the

screened dietary polyphenols and Dnmt3a might allow identification of compounds that have a positive effect on cancer prevention and improved mental performance.



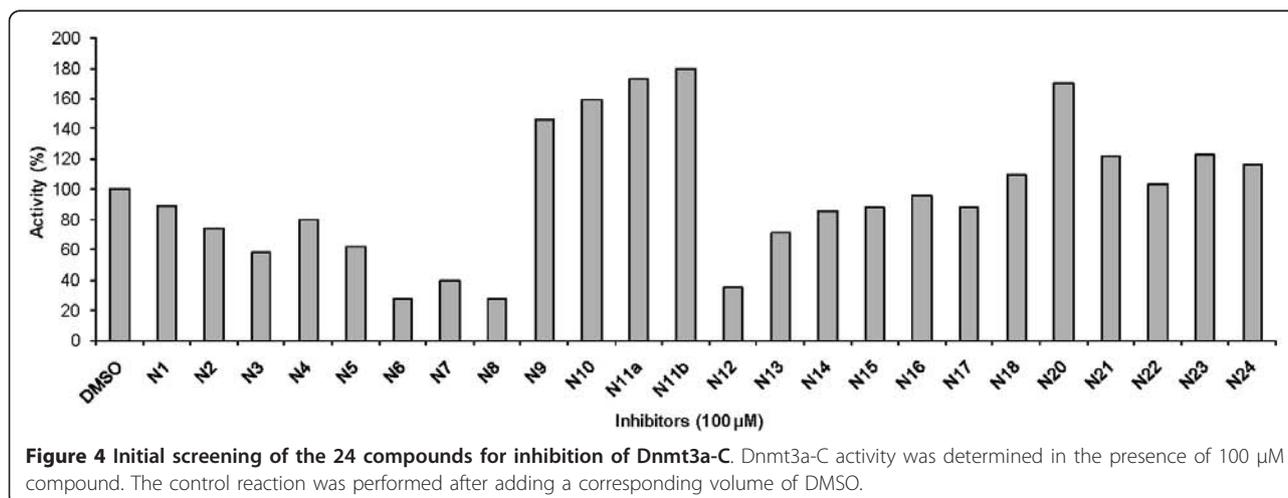
### Black tea polyphenols

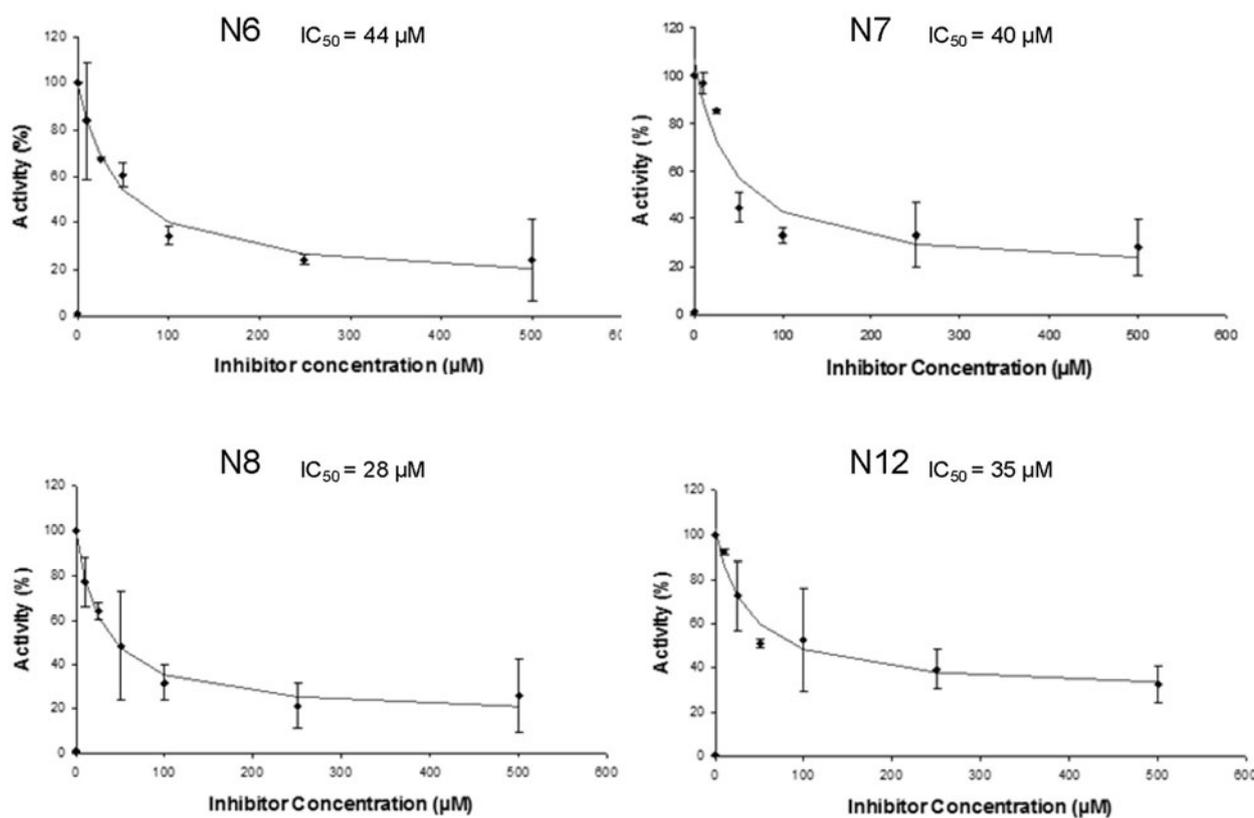
EGCG (N1) with a reported  $IC_{50}$  on Dnmt1 of  $0.21 \mu M$  and epigallocatechin (N4) showed only weak inhibition of Dnmt3a. A slightly increased activity was observed for theaflavin, theaflavin-3-gallate (N3) and theaflavin 3'-gallate (N5) with the gallated derivatives showing a larger inhibitory effect. Theaflavin 3, 3'-digallate (N6) performed best in this series with a measured  $IC_{50}$  value of  $44 \mu M$ . Similarly, the thearubigin fractions performed well in this test with  $IC_{50}$  values of  $40 \mu M$  and  $28 \mu M$ , respectively (molarity calculated by assuming an average molecular weight of  $800 \text{ g/mol}$ ). It has to be noted that according to our knowledge this is the first time that a thearubigin fraction (consumed at a level of 1 Mt per annum) has been investigated in an enzyme assay and found to exhibit inhibitory activity. Previous work on thearubigins biological activity had focused on interference with signalling cascades in the anti-inflammatory response [42-45]. Due to the structural similarity of theaflavins and thearubigins (poly-hydroxy theaflavins),

the inhibition of Dnmt3a does not come as a complete surprise.

To evaluate any possible biological significance of the  $IC_{50}$  values of Dnmt3a inhibition observed here, human pharmacokinetic data need to be consulted. Two published reports address the pharmacokinetic behaviours of theaflavins. Mulder and co-workers report theaflavin concentrations of  $4.2 \mu g \text{ l}^{-1}$  in urine 2h after consumption of 1 cup of black tea containing 8.8 mg total theaflavins [46]. Henning reported a concentration of  $2 \text{ nmol g}^{-1}$  tissue (if converted around  $2 \mu M$ ) of theaflavin in colon, small intestine, prostate and liver target tissue, with all further three theaflavins N3, N5 and N6 investigated here showing roughly  $1 \mu M$ , half this value after consumption of one cup of black tea [47]. Although no plasma concentration values are available for theaflavin derivatives, it can be assumed that the plasma concentration is the same order of magnitude if not even higher when compared with concentrations in target tissues. As the average per capita consumption of black tea is around 550 ml or three cups per day, again a higher physiological concentration must be assumed.

From these data it becomes obvious that out of the compounds investigated theaflavin 3, 3'-digallate N6 is a compound showing reasonable bioavailability. These concentration estimate of  $2 \mu M$  is only roughly by one order of magnitude smaller than the measured  $IC_{50}$  values. Assuming consumption of a black tea beverage rich in theaflavins (a maximum of  $50 \text{ mg l}^{-1}$  has been determined) or repeated consumption of larger quantities of black tea the measured  $IC_{50}$  values for Dnmt3a inhibition, therefore, may have biological significance and inhibition of this enzyme can be expected under physiological conditions after black tea consumption. No data are available on thearubigin pharmacokinetics but since a typical cup of tea contains 60-70% of its dry mass of this mixture of compounds biological significance can as well be assumed.





**Figure 5** Measurement of IC<sub>50</sub> values for compounds N6, N7, N8 and N12. For compounds N6, N7, N8 and N11 DNA methylation kinetics were carried out at different concentration of the compounds to determine the IC<sub>50</sub> value. The error bars show the maximal deviations in repeated experiments. IC<sub>50</sub> values are valid by  $\pm 30\%$ .

Two pieces of further work published recently touch on the problem discussed here and are worth highlighting. Firstly, work by Vauzour et al. showed that dietary polyphenols from berries of similar polarity and structure compared to the polyphenols studied here, are able to cross the blood brain barrier [48], therefore suggesting that brain target tissue could be reached by the compounds under investigation. Secondly, recent work by Müller-Harvey et al. reports an accumulation of tea polyphenols in cell nuclei [49], suggesting that not only target tissue but target cell organelles, in which Dnmt3a methylates DNA can indeed be reached by the compounds under investigation.

#### Coffee polyphenols

Out of the twelve chlorogenic acid derivatives screened, seven showed a minor inhibitory effect on Dnmt3a with one compound 1,3-dicaffeoyl-muco-quinic acid diacetal (N12) showing a good IC<sub>50</sub> value of 35 µM. Since compound N12 is a synthetic derivative, not present in the human diet, this finding has no direct dietary significance. However, the activity of compound N12 clearly indicates that chlorogenic acid derivatives have the

potential to inhibit Dnmt3a and this derivative might serve as a lead compound to screen and identify further dietary compound possessing this interesting biological activity.

Interestingly, all compounds showing inhibitory effects are diacyl quinic acids, whereas monoacyl quinic acids showed no effect at all. As a general trend caffeoyl derivatives seem to be more active if compared to feruloyl derivatives and a 1,3-diacyl regiochemistry appears to be favourable. Similarly gallic acid and caffeic acid had no inhibitory effect at all in contrast to the values reported by Lee & Zhu for Dnmt1 inhibition [12]. Despite the structural similarity of these two enzymes a predictive design of inhibitors targeting both classes of enzymes does not seem possible, which can be turned into an advantage considering that the compounds investigated by us and by Lee show remarkable selectivity for either Dnmt1 or Dnmt3a.

#### Conclusions

We have shown that the black tea polyphenols, in particular theaflavin 3, 3'-digallate N6 and thearubigin fraction inhibit Dnmt3a with a physiologically and

nutritionally relevant IC<sub>50</sub> value and therefore identified a novel biological target that is able to rationalize both anti-carcinogenic activity and mental health and performance related activity of black tea.

## Methods

### Expression and purification of Dnmt3a-C

The mouse Dnmt3a C-terminal domain was expressed and purified as described [29,30]. The purity of protein was determined on 12% SDS-PAGE gel stained with colloidal Coomassie Blue (Figure 1). Protein concentration was determined from the absorbance at 280 nm using an extinction coefficient of 39290 M<sup>-1</sup> cm<sup>-1</sup>.

### DNA methyltransferase activity assay

Kinetics of Dnmt3a-C was analyzed by using a Biotin-Avidin methylation kinetics assay basically as described [50] using a biotinylated oligonucleotide substrate and [methyl-3H]AdoMet.

FP3 5'-TTGCACTCTCCTCCCGGAAGTCC-CAGCTTC-3' FP3-Bt 5'-Bt-GAAGCTGGGACTTCCGG-GAGGAGAGTGCAA-3'; The oligonucleotides were annealed by heating to 86°C for some minutes and slowly cooling down to ambient temperature. The methylation reactions were carried out in methylation buffer [20 mM HEPES pH 7.2, 1 mM EDTA, 50 mM KCl, 25 mg/ml bovine serum albumin (BSA)] at 37°C, using 1 μM substrate DNA, 0.76 μM AdoMet and 2.5 μM Dnmt3a-C. After the methylation reaction, the oligonucleotides were immobilized at various time points on an avidin-coated microplate. The incorporation of [3H] into the DNA was quenched by addition of an excess of unlabeled AdoMet to the binding buffer. Subsequently, unreacted AdoMet was removed by washing five times with PBST containing 0.5 M NaCl. The immobilized DNA was digested with a non-specific endonuclease to release the radioactivity from the microplate. After digestion, 120 μl of the reaction mixture were transferred to a fresh microplate and 160 μl of Microscint-PS scintillation fluid (Perkin Elmer) was added to each well. Finally, the amount of methyl groups transferred to the DNA and the solution obtained after nucleolytic digestion was quantified by using the TopCount NXT liquid scintillation counter. To determine the initial slope, the data were fitted by linear regression of the initial part of the reaction progress curves.

All the inhibitors were prepared in the DMSO at 5 mM stock. For the screening purpose 100 μM concentrations of the inhibitors were used in the reaction mixture. To determine the apparent IC<sub>50</sub> value for the potential inhibitors, different concentration of the inhibitors were used in the reaction mixture (10 μM, 25 μM, 50 μM, 100 μM, 250 μM, 500 μM). The different concentrations of the inhibitors were incubated with Dnmt3a protein for 10 min at room temp. The reaction

was started by adding substrate and cofactor and further incubated at 37°C for another 10 min then the reaction was stopped by adding excess unlabelled AdoMet. The DMSO was used as control in each experimental setup to exclude the possible inhibition effect from the DMSO itself. All the inhibitor kinetics was done at duplicate and standard error was calculated for the two experimental values.

### Isolation and synthesis of inhibitors

EGCG N1 and (-)-epigallocatechin (N4), theaflavin (N2), theaflavin-3-gallate (N3), theaflavin 3'-gallate (N5) and theaflavin 3, 3'-digallate (N6) were from black tea obtained using published procedures (Figure 2) [33,51]. Thearubigin fractions (N7 and N8) were obtained from black tea and characterised using published procedures [33,51]. All chlorogenic acid derivatives were obtained by synthesis using published procedures (Figure 2) [52].

### Acknowledgements

The authors thank Jacobs University Bremen for a scholarship to Rakesh Jaiswal. Technical assistance from Ms. Anja Müller is gratefully acknowledged. This work was supported by DFG grant JE 252/6.

### Author details

<sup>1</sup>Chemistry, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany. <sup>2</sup>Biochemistry, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany. <sup>3</sup>MoLife program, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany.

### Authors' contributions

AR, ZT and RJ conducted and analyzed the experiments. NK and AJ participated in the design of the study and in data analysis and interpretation and drafted the manuscript. All authors read and approved the final manuscript.

Received: 17 March 2011 Accepted: 21 April 2011

Published: 21 April 2011

### References

1. Poulter S: *Daily Mail Online* 2008 [http://dailymail.co.uk].
2. Gardner EJ, Ruxton CHS, Leeds AR: **Black tea - helpful or harmful? A review of the evidence.** *European Journal of Clinical Nutrition* 2007, **61**(1):3-18.
3. Higdon JV, Frei B: **Coffee and health: A review of recent human research.** *Critical Reviews in Food Science and Nutrition* 2006, **46**(2):101-123.
4. Dorea JG, da Costa THM: **Is coffee a functional food?** *British Journal of Nutrition* 2005, **93**(6):773-782.
5. Hamer M, Williams ED, Vuononvirta R, Gibson EL, Steptoe A: **Association between coffee consumption and markers of inflammation and cardiovascular function during mental stress.** *Journal of Hypertension* 2006, **24**(11):2191-2197.
6. Franco R: **Coffee and mental health.** *Atencion Primaria* 2009, **41**(10):578-581.
7. Hozawa A, Kuriyama S, Nakaya N, Ohmori-Matsuda K, Kakizaki M, Sone T, Nagai M, Sugawara Y, Nitta A, Tomata Y, et al: **Green tea consumption is associated with lower psychological distress in a general population: the Ohsaki Cohort 2006 Study.** *American Journal of Clinical Nutrition* 2009, **90**(5):1390-1396.
8. de Mejia EG, Ramirez-Mares MV, Puangphaphant S: **Bioactive components of tea: Cancer, inflammation and behavior.** *Brain Behavior and Immunity* 2009, **23**(6):721-731.
9. Nurk E, Refsum H, Drevon CA, Tell GS, Nygaard HA, Engedal K, Smith AD: **Intake of Flavonoid-Rich Wine, Tea, and Chocolate by Elderly Men and**

- Women Is Associated with Better Cognitive Test Performance. *Journal of Nutrition* 2009, **139**(1):120-127.
10. Fang MZ, Chen DP, Yang CS: Dietary polyphenols may affect DNA methylation. *Journal of Nutrition* 2007, **137**(1):223S-228S.
11. Lee WJ, Shim JY, Zhu BT: Mechanisms for the inhibition of DNA methyltransferases by tea catechins and bioflavonoids. *Molecular Pharmacology* 2005, **68**(4):1018-1030.
12. Lee WJ, Zhu BT: Inhibition of DNA methylation by caffeic acid and chlorogenic acid, two common catechol-containing coffee polyphenols. *Carcinogenesis* 2006, **27**(2):269-277.
13. Nandakumar V, Vaid M, Katiyar SK: (-)-Epigallocatechin-3-gallate reactivates silenced tumor suppressor genes, Cip1/p21 and p16INK4a, by reducing DNA methylation and increasing histones acetylation in human skin cancer cells. *Carcinogenesis* 2011.
14. Klose RJ, Bird AP: Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* 2006, **31**(2):89-97.
15. Jurkowska RZ, Jurkowski TP, Jeltsch A: Structure and function of mammalian DNA methyltransferases. *Chembiochem* 2011, **12**(2):206-222.
16. Jeltsch A: Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. *Chembiochem* 2002, **3**(4):274-293.
17. Jones PA, Baylín SB: The epigenomics of cancer. *Cell* 2007, **128**(4):683-692.
18. Feinberg AP, Tycko B: The history of cancer epigenetics. *Nat Rev Cancer* 2004, **4**(2):143-153.
19. Egger G, Liang GN, Aparicio A, Jones PA: Epigenetics in human disease and prospects for epigenetic therapy. *Nature* 2004, **429**(6990):457-463.
20. Yoo CB, Jones PA: Epigenetic therapy of cancer: past, present and future. *Nature Reviews Drug Discovery* 2006, **5**(1):37-50.
21. Gal-Yam EN, Saito Y, Egger G, Jones PA: Cancer epigenetics: modifications, screening, and therapy. *Annu Rev Med* 2008, **59**:267-280.
22. Kelly TK, De Carvalho DD, Jones PA: Epigenetic modifications as therapeutic targets. *Nat Biotechnol* 2010, **28**(10):1069-1078.
23. Levenson JM, Sweatt JD: Epigenetic mechanisms: a common theme in vertebrate and invertebrate memory formation. *Cellular and Molecular Life Sciences* 2006, **63**(9):1009-1016.
24. Wu H, Coskun V, Tao J, Xie W, Ge W, Yoshikawa K, Li E, Zhang Y, Sun YE: Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science* 2010, **329**(5990):444-448.
25. Feng J, Zhou Y, Campbell SL, Le T, Li E, Sweatt JD, Silva AJ, Fan G: Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nat Neurosci* 2010, **13**(4):423-430.
26. LaPlant Q, Vialou V, Covington HE, Dumitriu D, Feng J, Warren BL, Maze I, Dietz DM, Watts EL, Iniguez SD, et al: Dnmt3a regulates emotional behavior and spine plasticity in the nucleus accumbens. *Nat Neurosci* 2010, **13**(9):1137-1143.
27. Weaver ICG, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR, Dymov S, Szyf M, Meaney MJ: Epigenetic programming by maternal behavior. *Nature Neuroscience* 2004, **7**(8):847-854.
28. Feng J, Fan G: The role of DNA methylation in the central nervous system and neuropsychiatric disorders. *Int Rev Neurobiol* 2009, **89**:67-84.
29. Gowher H, Jeltsch A: Molecular enzymology of the catalytic domains of the Dnmt3a and Dnmt3b DNA methyltransferases. *J Biol Chem* 2002, **277**(23):20409-20414.
30. Jurkowska RZ, Anspach N, Urbanke C, Jia D, Reinhardt R, Nellen W, Cheng X, Jeltsch A: Formation of nucleoprotein filaments by mammalian DNA methyltransferase Dnmt3a in complex with regulator Dnmt3L. *Nucleic Acids Res* 2008, **36**(21):6656-6663.
31. Drynan JW, Clifford MN, Obuchowicz J, Kuhnert N: The chemistry of low molecular weight black tea polyphenols. *Natural Product Reports* 2010, **27**(3):417-462.
32. Roberts EAH, Catwright RA, Oldschool M: *J Sci Food Agric* 1959, **8**:72-80.
33. Kuhnert N, Drynan JW, Obuchowicz J, Clifford MN, Witt M: Mass spectrometric characterization of black tea thearubigins leading to an oxidative cascade hypothesis for thearubigin formation. *Rapid Communications in Mass Spectrometry* 2010, **24**(23):3387-3404.
34. Kuhnert N: Unraveling the structure of the black tea thearubigins. *Archives of Biochemistry and Biophysics* 2010, **501**(1):37-51.
35. Roberts EAH, Myers M: *J Sci Food Agric* 1959, **10**:167-179.
36. Clifford MN: Chlorogenic acids and other cinnamates - nature, occurrence and dietary burden. *Journal of the Science of Food and Agriculture* 1999, **79**(3):362-372.
37. Clifford MN, Johnston KL, Knight S, Kuhnert N: Hierarchical scheme for LC-MSn identification of chlorogenic acids. *Journal of Agricultural and Food Chemistry* 2003, **51**(10):2900-2911.
38. Clifford MN, Knight S, Surucu B, Kuhnert N: Characterization by LC-MSn of four new classes of chlorogenic acids in green coffee beans: Dimethoxycinnamoylquinic acids, diferuloylquinic acids, caffeoyl-dimethoxycinnamoylquinic acids, and feruloyl-dimethoxycinnamoylquinic acids. *Journal of Agricultural and Food Chemistry* 2006, **54**(6):1957-1969.
39. Jaiswal R, Patras MA, Eravuchira PJ, Kuhnert N: Profile and Characterization of the Chlorogenic Acids in Green Robusta Coffee Beans by LC-MSn: Identification of Seven New Classes of Compounds. *Journal of Agricultural and Food Chemistry* 2010, **58**(15):8722-8737.
40. Jaiswal R, Sovdat T, Vivan F, Kuhnert N: Profiling and Characterization by LC-MSn of the Chlorogenic Acids and Hydroxycinnamoylshikimate Esters in Mate (*Ilex paraguariensis*). *Journal of Agricultural and Food Chemistry* 2010, **58**(9):5471-5484.
41. Yokochi T, Robertson KD: Dimethyl sulfoxide stimulates the catalytic activity of de novo DNA methyltransferase 3a (Dnmt3a) in vitro. *Bioorg Chem* 2004, **32**(4):234-243.
42. Lin YL, Tsai SH, Lin-Shiau SY, Ho CT, Lin JK: Theaflavin-3,3'-digallate from black tea blocks the nitric oxide synthase by down-regulating the activation of NF-kappa B in macrophages. *European Journal of Pharmacology* 1999, **367**(2-3):379-388.
43. Bhattacharya U, Halder B, Mukhopadhyay S, Giri AK: Role of oxidation-triggered activation of JNK and p38 MAPK in black tea polyphenols induced apoptotic death of A375 cells. *Cancer Science* 2009, **100**(10):1971-1978.
44. Jochmann N, Lorenz M, von Krosigk A, Martus P, Bohm V, Baumann G, Stangl K, Stangl V: The efficacy of black tea in ameliorating endothelial function is equivalent to that of green tea. *British Journal of Nutrition* 2008, **99**(4):863-868.
45. Lorenz M, Urban J, Engelhardt U, Baumann G, Stangl K, Stangl V: Green and black tea are equally potent stimuli of NO production and vasodilation: new insights into tea ingredients involved. *Basic Research in Cardiology* 2009, **104**(1):100-110.
46. Vermeer MA, Mulder TP, Molhuizen HO: Theaflavins from black tea, especially theaflavin-3-gallate, reduce the incorporation of cholesterol into mixed micelles. *J Agric Food Chem* 2008, **56**(24):12031-12036.
47. Henning SM, Aronson W, Niu YT, Conde F, Lee NH, Seeram NP, Lee RP, Lu JX, Harris DM, Moro A, et al: Tea polyphenols and theaflavins are present in prostate tissue of humans and mice after green and black tea consumption. *Journal of Nutrition* 2006, **136**(7):1839-1843.
48. Vauzour D, Rendeiro C, Corona G, Williams C, Spencer JEP: *Polyphenol commun* 2010, **1**:70-71.
49. Müller-Harvey I, Botchway S, Feucht W, Polster J, Burgos P, Parker A: *Planta Medica* 2010, **76**:1167.
50. Roth M, Jeltsch A: Biotin-avidin microplate assay for the quantitative analysis of enzymatic methylation of DNA by DNA methyltransferases. *Biological Chemistry* 2000, **381**(3):269-272.
51. Kuhnert N, Clifford MN, Muller A: Oxidative cascade reactions yielding polyhydroxy-theaflavins and theacitrins in the formation of black tea thearubigins: Evidence by tandem LC-MS. *Food & Function* 2010, **1**(2):180-199.
52. Kuhnert N, Jaiswal R, Eravuchira P, El-Abassy RM, von der Kammer B, Materny A: Scope and limitations of principal component analysis of high resolution LC-TOF-MS data: the analysis of the chlorogenic acid fraction in green coffee beans as a case study. *Analytical Methods* 2011, **3**(1):144-155.

doi:10.1186/1471-2091-12-16

Cite this article as: Rajavelu et al.: The inhibition of the mammalian DNA methyltransferase 3a (Dnmt3a) by dietary black tea and coffee polyphenols. *BMC Biochemistry* 2011 **12**:16.

# C5-DNA Methyltransferase Inhibitors: From Screening to Effects on Zebrafish Embryo Development

Alexandre Ceccaldi,<sup>[a]</sup> Arumugam Rajavelu,<sup>[b]</sup> Christine Champion,<sup>[a]</sup> Christine Rampon,<sup>[c]</sup> Renata Jurkowska,<sup>[b]</sup> Gytis Jankevicius,<sup>[b]</sup> Catherine Sénamaud-Beaufort,<sup>[a]</sup> Loïc Ponger,<sup>[a]</sup> Nathalie Gagey,<sup>[a]</sup> Hana Dali Ali,<sup>[a]</sup> Jörg Tost,<sup>[d]</sup> Sophie Vrız,<sup>[c]</sup> Sindu Ros,<sup>[e]</sup> Daniel Dauzonne,<sup>[e]</sup> Albert Jeltsch,<sup>[b]</sup> Dominique Guianvarc'h,<sup>[f]</sup> and Paola B. Arimondo<sup>\*,[a]</sup>

DNA methylation is involved in the regulation of gene expression and plays an important role in normal developmental processes and diseases, such as cancer. DNA methyltransferases are the enzymes responsible for DNA methylation on the position 5 of cytidine in a CpG context. In order to identify and characterize novel inhibitors of these enzymes, we developed a fluorescence-based throughput screening by using a short DNA duplex immobilized on 96-well plates. We have screened 114 flavones and flavanones for the inhibition of the murine

catalytic Dnmt3a/3L complex and found 36 hits with IC<sub>50</sub> values in the lower micromolar and high nanomolar ranges. The assay, together with inhibition tests on two other methyltransferases, structure–activity relationships and docking studies, gave insights on the mechanism of inhibition. Finally, two derivatives effected zebrafish embryo development, and induced a global demethylation of the genome, at doses lower than the control drug, 5-azacytidine.

## Introduction

Epigenetics describes the heritable and reversible regulation of gene expression that is not encoded within the DNA sequence. DNA methylation, histones modifications (e.g., methylation and acetylation) and nucleosome dynamics are the main epigenetic modifications.<sup>[1–3]</sup> Epigenetic regulation is implicated in many biological phenomena, including evolution, embryonic development, cell differentiation, genomic imprinting, X chromosome inactivation, memory formation and consolidation, and animal behavior.<sup>[4–6]</sup> They are also involved in various diseases, like cancer.<sup>[7]</sup> A general disruption of the epigenetic landscape is associated with almost every human cancer.<sup>[8]</sup> The genome of malignant cells is globally hypomethylated, which is responsible for aberrant expression of repeated sequences, such as transposons, and leads to global genomic instability. In parallel, most of the tumors show specific hypermethylation of the promoters of certain genes, such as tumor suppressor genes, associated with their silencing. In mammals, DNA methyltransferases (DNMT) catalyze the transfer of a methyl group from the cofactor S-adenosyl-L-methionine (SAM) to the position 5 of cytidines in the context of CpG dinucleotides.<sup>[9]</sup> Interestingly, for the development of antitumor strategies, this observed gain of methylation is reversible, and epigenetic modulators, therefore, represent an attractive new class of chemotherapeutic agents.<sup>[10]</sup> Indeed, DNA methyltransferase inhibitors (DNMTi) have been shown to be able to reactivate tumor suppressor genes and revert malignant cell status.<sup>[11]</sup>

Two types of DNMT inhibitors have been described.<sup>[11,12]</sup> Cytidine analogues have been known for more than thirty years. 5-Azacytidine (Vidaza) and 5-aza-2'-deoxycytidine (Dacogen<sup>TM</sup>) were approved by the FDA in 2004 and 2005, respectively, for the treatment of myelodysplastic syndromes, and cytidine ana-

logues are currently tested against many types of solid tumors. These inhibitors act as suicide substrates for the enzyme and need to be incorporated into the DNA to be active. However, they integrate anywhere in the genome, are highly toxic and some are chemically instable.<sup>[13,14]</sup> More recently, non-nucleoside DNMT inhibitors, showing a much larger chemical diversity, has been discovered.<sup>[11]</sup> Hydralazine, a commonly used vasodilator,<sup>[15]</sup> the green tea polyphenol EGCG<sup>[16,17]</sup> or RG108 (a molecule discovered through a virtual screening)<sup>[18]</sup> have been

[a] A. Ceccaldi, C. Champion, C. Sénamaud-Beaufort, Dr. L. Ponger, Dr. N. Gagey, H. Dali Ali, Dr. P. B. Arimondo<sup>+</sup>  
UMR 7196 CNRS, MNHN INSERM U565  
43 rue Cuvier, 75005 Paris (France)  
Fax: (+33) 5-34-50-34-92  
E-mail: paola.arimondo@dr14.cnrs.fr

[b] A. Rajavelu, Dr. R. Jurkowska, G. Jankevicius, Prof. A. Jeltsch  
Jacobs University Bremen, Campus Ring 1, 28759 Bremen (Germany)

[c] Dr. C. Rampon, Prof. S. Vrız  
UMR 7233 CNRS-Chaire des Processus  
Morphogénétiques Collège de France  
11, Place Marcelin Berthelot 75005 Paris (France)

[d] Dr. J. Tost  
Laboratory for Epigenetics, Centre National de Génotypage  
CEA-Institut de Génomique, 91000 Evry (France)

[e] Dr. S. Ros, Dr. D. Dauzonne  
CNRS UMR176 Institut Curie Centre de Recherche  
26, rue d'Ulm 75005 Paris (France)

[f] Dr. D. Guianvarc'h<sup>+</sup>  
UPMC Paris 06, ENS, CNRS, UMR 7203, FR2769  
4, place Jussieu, 75005 Paris (France)

[<sup>+</sup>] These authors contributed equally to this work.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cbic.201100130>.

shown to inhibit DNMTs. However, their detailed mechanism of action remains largely unknown, their specificity towards DNMTs has yet to be demonstrated and most of them have a very low efficiency. Thus, there is an urgent need to design more efficient and less toxic inhibitors. Here, we describe the discovery of new and potent inhibitors of the murine catalytic Dnmt3a/3L complex with the use of a novel high-throughput screening (HTS) system for DNMT inhibitors.

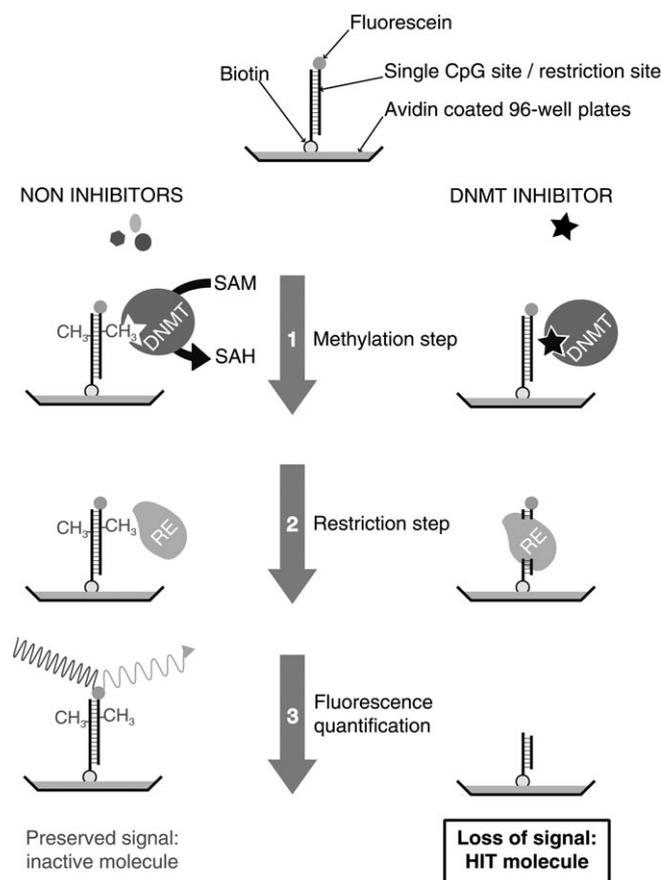
## Results

### The screening test

In order to discover new inhibitors of DNMTs, we developed a medium-/high-throughput screening assay with the following specifications: quick and reliable *in vitro* test that uses fluorescence, compatible with automation and affordable enough to allow large-scale screenings. The versatility of the assay was important to ensure its compatibility with multiple types of methyltransferases. Accordingly, all inhibition mechanisms can be analyzed with the same assay: substrate competition (DNA and/or SAM competition), noncompetitive inhibition, conformational changes and allosteric effects as well as inhibition of protein–protein interactions. Here, we focused on the discovery of new drugs that, in contrast to the above-mentioned cytosine analogues, do not need to be integrated in DNA to be active.<sup>[11]</sup> This functional design called for a simple technical principle to measure DNA methylation inhibition; we chose to detect the loss of fluorescent signal after cleavage of a fluorescently labeled DNA duplex by a methylation-sensitive restriction endonuclease (RE) that is only active if the substrate is not methylated at the CpG site contained in its specific recognition sequence (Figure 1). The 96-well plates were coated for biotin–avidin interactions with a double-modified DNA duplex containing the fluorophore, 6-FAM, at one end and biotin at the other of the complementary strand. When CpG is methylated, the RE can no longer cleave the DNA substrate and fluorescence is retained after washing. In the presence of an inhibitor, the site is not methylated, the DNA is cleaved and the fluorescence is lost after washing. A restriction step with an isoschizomer, which cleaves the same target site but is not sensitive to methylation, serves as a control that ensures the loss of restriction cleavage in the presence of DNMTs is indeed due to DNA methylation.

First, we applied the screen to the murine catalytic complex of Dnmt3a/3L and added the tested compounds of the chemical library during the formation of the catalytic complex. In the case of the Dnmt3a/3L complex, the crucial influence of this protein–protein interaction on Dnmt3a catalytic activity has been shown.<sup>[19]</sup>

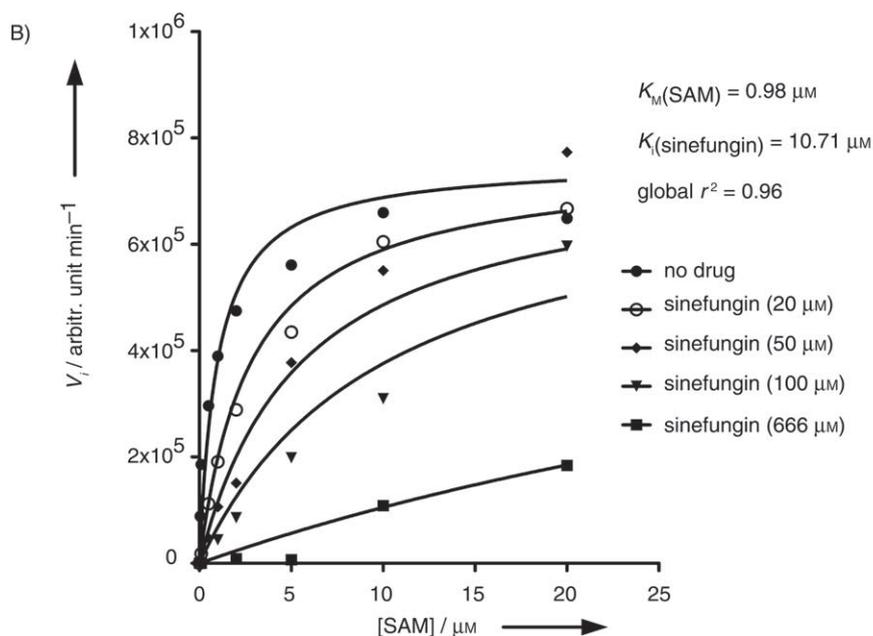
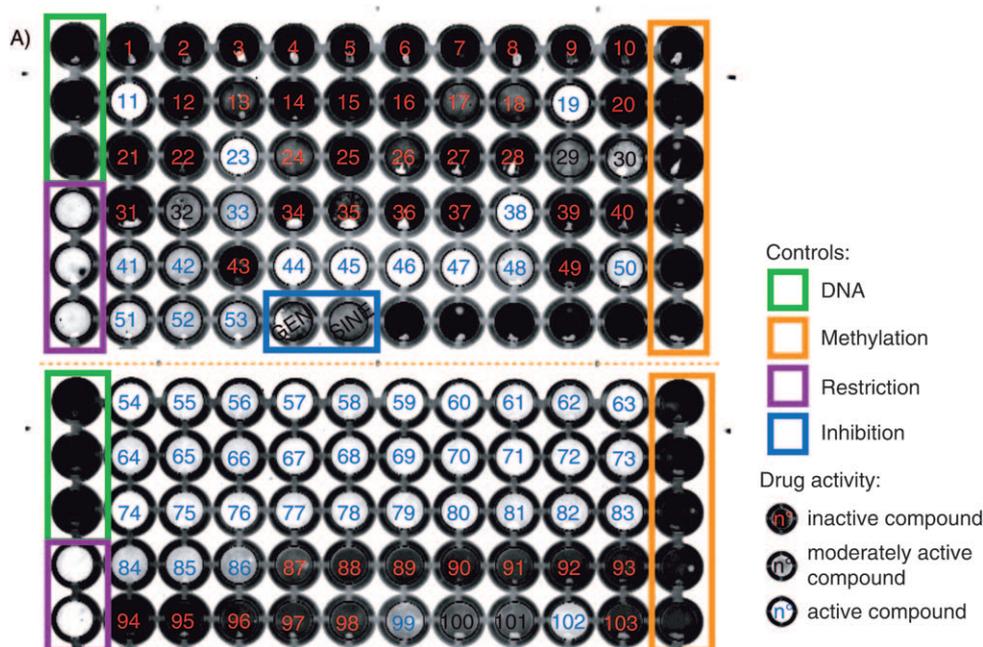
114 flavones and flavanone derivatives were screened for Dnmt3a/3L activity at 500  $\mu\text{M}$  (Figure 2 and Table S1 in the Supporting Information). These families of molecules were chosen due to their chemical proximity with genistein,<sup>[20]</sup> a soybean isoflavone that was shown to inhibit several enzymatic activities including DNMTs. Four types of controls were used to ensure the reliability of each experiment: DNA, methylation,



**Figure 1.** Schematic overview of the screening assay: 96-well microplates were coated with a double modified fluorescent DNA duplex containing a single CpG site through biotin–avidin interactions. Step 1: methylation in the presence of tested compounds, followed by washing. Step 2: methylation sensitive restriction reaction followed by washing. Step 3: measurement of fluorescence and quantification. A hit corresponds to a loss of fluorescent signal since the unmethylated duplex is cut during step 2.

restriction and inhibition (Figure 2). Inhibition controls (blue boxes) offer a comparison with known active inhibitors, such as genistein,<sup>[20]</sup> sinefungin (a SAM analogue)<sup>[21]</sup> and RG108.<sup>[18]</sup> Active compounds (Figure 2, in blue) induced a loss of fluorescence and the wells appear white. Inactive compounds (Figure 2, in red) maintained the fluorescent signal and the wells appear black. Intermediate compounds are written in black. Interestingly, we found a group of more than 50 new molecules that are more active than sinefungin and genistein (Figure 2A and Table S1 in the Supporting Information). To validate the assay, the kinetics of the Dnmt3a/3L complex was measured in the presence of sinefungin, a reference inhibitor of DNMTs that acts as a cofactor competitor (Figure 2B). After global nonlinear regression the data fitted correctly with the competitive model ( $r^2 = 0.96$ ) with apparent  $K_{M(\text{SAM})} = 0.98 \mu\text{M}$ , which is consistent with previous data<sup>[22]</sup> and  $K_{i(\text{sinefungin})} = 10.71 \mu\text{M}$ .

After a secondary screening at 5  $\mu\text{M}$  (Figure S1A in the Supporting Information) we evaluated the  $\text{IC}_{50}$  of the best molecules using the 96-well fluorescence-based enzymatic assay (Figure S2 in the Supporting Information). In these further



**Figure 2.** A) Primary screening at 500  $\mu\text{M}$  of 109 flavone and flavanone derivatives tested as inhibitors of the Dnmt3a/3L C-terminal domains complex in the presence of SAM. Fluorescence detection in each well was measured on a Typhoon scanner and quantified with the ImageQuant software. The dashed line separates two different experiments. The maximal fluorescent signal is given by DNA controls (green box) that did not undergo enzymatic treatment. Restriction controls (purple boxes), corresponding to the minimal fluorescent signal that only underwent the restriction step, were included to avoid false negative results. Methylation controls (orange boxes) underwent all the enzymatic steps and were used to verify the activity of the DNMTs during the assay. Blue box: controls of methylation in the presence of genistein and sinefungin. Inactive compounds are printed in red, active compounds in blue and medium compounds in black. B) SAM competition by sinefungin on the Dnmt3a/3L complex. The  $K_M$  for SAM and  $K_i$  for sinefungin are given with the global  $r^2$  of the competitive inhibition model fitting.

studies, beside the best hits of the flavanone family, we also included the two most active molecules, 3-nitroflavones, **11** and **23**. Table 1 shows the structures and  $\text{IC}_{50}$  values of a subset of the most potent compounds. The entire library of molecules is

reported in Table S1 in the Supporting Information. The best molecule, compound **69**, displayed an in vitro  $\text{IC}_{50}$  of  $371(\pm 53)$  nM, 1000-fold lower than reference compound RG108 in the same test.

The  $\text{IC}_{50}$  values of the most representing compounds measured by this HTS method were compared to the ones obtained in a classical methylation assay based on the incorporation of radioactively labeled  $^3\text{H}$ -SAM in a DNA duplex in solution.<sup>[23]</sup> Despite the different reaction conditions and setup, the two methods gave comparable apparent  $\text{IC}_{50}$  values (Table 1).

### Structure–activity relationships

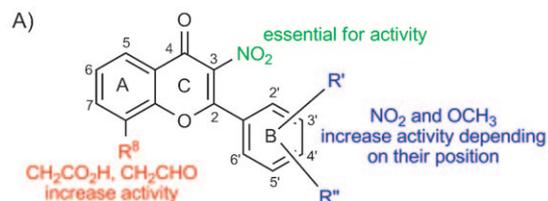
We further investigated the relationship between the structure and activity of the screened library of molecules (Scheme 1 and Table S1 in the Supporting Information). The active compounds belong to two well-identified families:

- 1) The 3-nitroflavones, in which the presence of a nitro substituent on the B ring was critical for activity (compare **19** to **20**, for example). Substituents in position 8, such as  $\text{CH}_2\text{-CHO}$  or  $\text{CH}_2\text{-COOH}$ , increased the activity (compounds **11**, **19** vs. **41**, **16**). Finally, a methoxy group on cycle C enhanced the activity depending on its position (cf. compounds **19**, **23** compared to **24**, **21**).
- 2) The 3-chloro-3-nitroflavones, for which the most potent compound was **69**. The presence of  $\text{NO}_2$  on the B ring was stabilizing (compared to **52**) with the following order: *ortho* > *meta*, *para* (**69** > **47**, **70**). Electron-withdrawing groups, such as chloro and fluoro substituents, increased the activity (**61**, **62**, **64**, **65**, **66** and **113** vs. **52**) in particular in *ortho* for Cl (**64**) and *meta* for F (**113**). The order of the increased activity observed for these substitu-

**Table 1.** Comparison of IC<sub>50</sub> of some interesting compounds measured on Dnmt3a/3L complex by HTS FluorMet and <sup>3</sup>H-SAM assay.

Compound	Chemical structure	HTS IC <sub>50</sub> [μM]	<sup>3</sup> H-SAM IC <sub>50</sub> [μM]
11		9.5 ± 1.1	15 ± 5
23		1.26 ± 0.17	2.37 ± 0.90
47		1.31 ± 0.38	0.92 ± 0.06
62		1.68 ± 0.18	1.76 ± 0.62
63		0.48 ± 0.12	n.d.
64		0.69 ± 0.16	n.d.
68		0.56 ± 0.10	n.d.
69		0.37 ± 0.05	0.49 ± 0.09
70		1.32 ± 0.08	1.03 ± 0.32
74		0.67 ± 0.16	n.d.
L-RG108		315	> 500
genistein		> 500	> 100

n.d.: not determined.

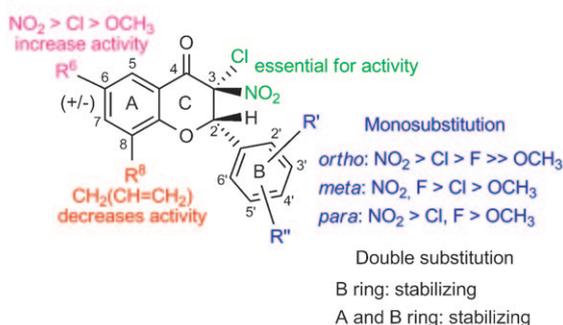


SAR for selected compounds

IC<sub>50</sub>

11: R <sup>a</sup> = CH <sub>2</sub> CHO	R <sup>c</sup> = NO <sub>2</sub>		9.5 μM
45: R <sup>a</sup> = CH <sub>2</sub> CHO			27 μM
41: R <sup>a</sup> = H	R <sup>c</sup> = NO <sub>2</sub>		14 μM
16: R <sup>a</sup> = H	R <sup>c</sup> = NO <sub>2</sub>	R <sup>d</sup> = OCH <sub>3</sub>	inactive
19: R <sup>a</sup> = CH <sub>2</sub> CO <sub>2</sub> H	R <sup>c</sup> = NO <sub>2</sub>	R <sup>d</sup> = OCH <sub>3</sub>	19 μM
20: R <sup>a</sup> = CH <sub>2</sub> CO <sub>2</sub> H	R <sup>c</sup> = H	R <sup>d</sup> = OCH <sub>3</sub>	inactive
21: R <sup>a</sup> = CH <sub>2</sub> CO <sub>2</sub> H	R <sup>c</sup> = H	R <sup>e</sup> = OCH <sub>3</sub>	inactive
23: R <sup>a</sup> = CH <sub>2</sub> CO <sub>2</sub> H	R <sup>c</sup> = NO <sub>2</sub>	R <sup>e</sup> = OCH <sub>3</sub>	1.25 μM
24: R <sup>a</sup> = CH <sub>2</sub> CO <sub>2</sub> H	R <sup>c</sup> = NO <sub>2</sub>	R <sup>e</sup> = H	inactive

B)

**Scheme 1.** Structure–activity relationships of: A) the 3-nitroflavone family, and B) the 3-chloro-3-nitroflavone compounds.

ent was: in *ortho*: NO<sub>2</sub> > Cl > F; in *meta*: F, NO<sub>2</sub> ≥ Cl; in *para*: NO<sub>2</sub> > Cl, F.

The methoxy substitution was always penalizing (**78**, **79** and **80**) but the least in position *meta* (**78**). Double substitutions with one electron-withdrawing group on the B ring increased the activity (compound **63** vs. **62**, and **73** vs. **79**) while methoxy groups alone were greatly destabilizing (**81** and **82**). A ring substitutions associated with B ring substitutions increased the activity when in position 6: the NO<sub>2</sub> group increased activity by threefold (**71** vs. **70**, **68** vs. **66**, and **72** vs. **52**) and the chloro substituent led to a twofold increase (**67** vs. **66**). Again the methoxy substituent was less efficient (**76** vs. **66**). Allyl substitution in position 8 on the A ring clearly decreased the activity (e.g., **53** vs. **52**, **54** vs. **69**). Finally, multiple substitutions with electron-withdrawing groups on the A ring conferred a very good activity (**74** vs. **66**).

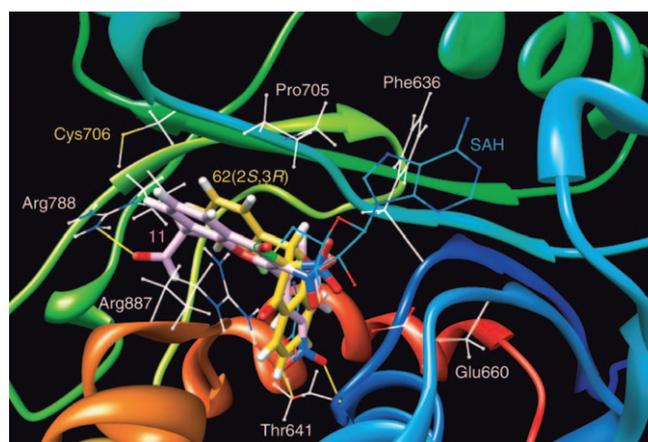
### Inhibition mechanism

Next, we addressed the selectivity and mechanism of action of these new potential Dnmt3a/3L inhibitors by complementary experimental approaches. First, the HTS was applied to the

human DNMT1 (Figure S1 B in the Supporting Information) and the bacterial methyltransferase, M.SssI (Figure S1 C in the Supporting Information). The chemical library was screened at 5  $\mu\text{M}$ . Globally, a similar profile for DNMT1 and M.SssI compared to the catalytic Dnmt3a/3L complex was observed (Figure S1 A in the Supporting Information).

We then investigated whether the hit molecules could affect the activity of other methyltransferases in order to get insights on the mechanism of inhibition. A subset of the most representing compounds were tested against the histone methyltransferase (HMTK) G9a, which methylates histone tails on their lysine residues,<sup>[24]</sup> and the bacterial DNA (adenine N-6)-methyltransferase, EcoDam, which methylates DNA at GATC sites.<sup>[25]</sup> A biotin–avidin microplate kinetic assay based on <sup>3</sup>H-SAM incorporation was used for the quantitative analysis of methyltransferase activities, *in vitro*.<sup>[23]</sup> Interestingly, the flavone **11** showed a very good specificity for the Dnmt3a/3L complex compared to the other two methyltransferases, with an  $\text{IC}_{50}$  of 15  $\mu\text{M}$  against Dnmt3a/3L and 352  $\mu\text{M}$  against G9a HMTK and 290  $\mu\text{M}$  against EcoDam. The chloronitroflavanone **62** showed specificity for the DNA methyltransferases ( $\text{IC}_{50}$  = 14.8, 131.4 and 6.5  $\mu\text{M}$  for Dnmt3a/3L, G9a HMTK and EcoDam, respectively). A similar profile was obtained with compound **69**, the most active molecule on Dnmt3a/3L, and was found to show an even clearer preference for the DNMT EcoDam ( $\text{IC}_{50}$  = 11  $\mu\text{M}$ ) versus G9a HMTK ( $\text{IC}_{50}$  = 450  $\mu\text{M}$ ). On the other hand, the nitro analogues **70** and **47** were less selective against Dnmt3a/3L ( $\text{IC}_{50}$  = 9.5 and 4.5  $\mu\text{M}$ , respectively), more efficient against EcoDam ( $\text{IC}_{50}$  = 2.2 and 0.9  $\mu\text{M}$ , respectively), and showed an excellent inhibition of HMTK G9a ( $\text{IC}_{50}$  = 0.5 and 1  $\mu\text{M}$ , respectively).

Molecular docking studies were carried out on the three enzymes. Compounds **11**, **47**, **62**, **69**, **70** and **71** were docked with Dock 6.4 on the crystal structure of the murine catalytic Dnmt3a/3L complex (PDB ID: 2QRV).<sup>[19]</sup> Since the flavanones exist in the racemic form, both enantiomers *R,S* and *S,R* were docked. All compounds were found in the catalytic pocket, either in the SAM pocket or closer to the DNA pocket (Figure 3 and Table S2 in the Supporting Information). Interestingly, flavone derivative **11** made a potential H bond with an amino acid close to the catalytic cysteine. We observed differences between the enantiomers of the flavanones; the *2S,3R* molecules stretched in the DNA pocket whereas the *2R,3S* compounds tended to be docked in the SAM pocket (Scheme S1 in the Supporting Information) except for **47** and **71**, for which both enantiomers were found closer to the DNA pocket. Strikingly, the  $\text{NO}_2$  group at the C3 position shared by all the compounds was always located in the same pocket formed by Pro705, Phe636, Gly638, Glu660 and Arg887. This was also true for both enantiomers. Interestingly, the potent inhibitors **69** and **71** made a H bond with the catalytic cysteine in position 706 (Table S2 in the Supporting Information). Analogue docking studies were done in the crystal structure of EcoDam (PDB ID: 2G1P; Table S2 in the Supporting Information) and G9a (PDB ID: 3K5K). In this case, the minimal conformation for all compounds corresponded to the one in the SAM pocket of the enzymes.

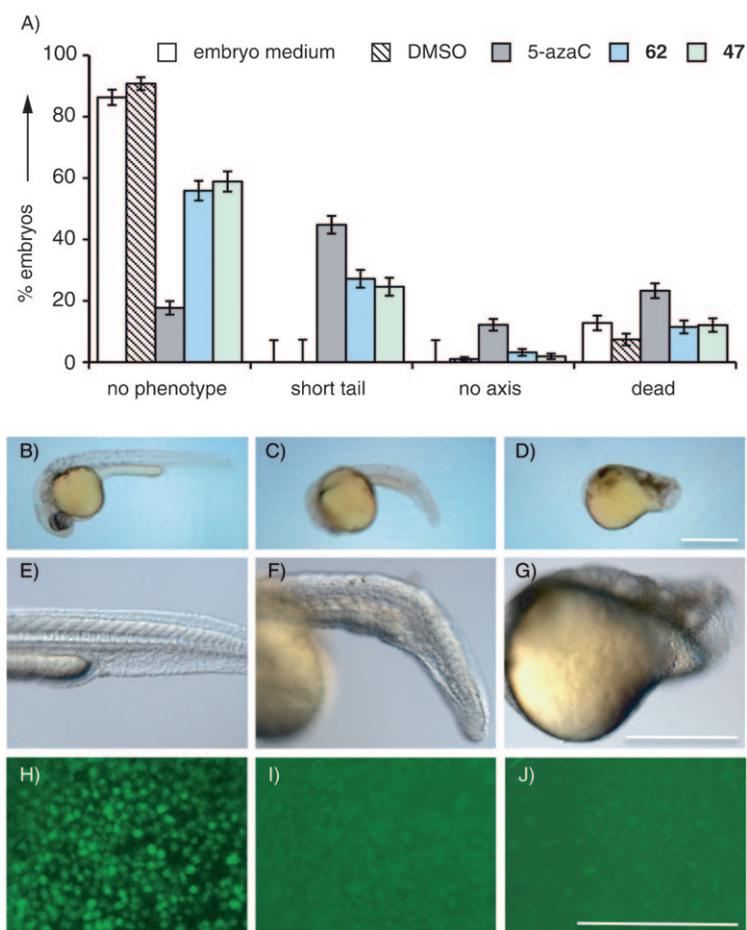


**Figure 3.** Molecular docking of compounds **11** and **62** on the crystal structure of the catalytic Dnmt3a (PDB ID: 2QRV). Inhibitors are shown in stick representation in pink for compound **11** and yellow for **62** (in configuration *2S,3R*). The cofactor SAH (*S*-adenosyl-homocysteine, in blue), the catalytic cysteine and the amino acids close to the 3- $\text{NO}_2$  on the C ring of the inhibitors (< 5 Å) are shown. Heteroatoms are colored: oxygen in red, nitrogen in blue, sulfur in yellow. H bonds are shown in yellow. The SAH molecule was docked as control (in stick representation).

Since the docking studies suggested that the compounds should have a mixed inhibition mechanism, only partially competing with the SAM cofactor, we determined the apparent  $K_M$  and  $V_{max}$  of the Dnmt3a/3L complex in the presence of the inhibitors as a function of the SAM concentration. Indeed, compounds **11**, **47** and **62** gave typical plots of a mixed inhibition model; with increasing concentration of the inhibitor the apparent  $K_M$  increased and the  $V_{max}$  decreased. This demonstrates that the molecule had a different affinity for the Dnmt3a/3L complex depending whether or not a SAM molecule was bound. This result has two major implications. First, our new DNMT inhibitors are not pure SAM competitors. Since SAM is the common cofactor that provides the methyl group of all methyltransferase activities in the cell, SAM competition would be unfavorable for selectivity. Second, the mixed inhibition profile is characteristic for a binding inside the active site of the enzyme, as suggested by docking.

### Zebrafish embryo development model

It is known that inhibition of DNA methylation at the blastula stage induces the loss of tail and abnormal patterning of somites in zebrafish.<sup>[26]</sup> We, therefore, decided to investigate the biological effect of molecules **11**, **23**, **47**, **62** and **69** and a negative control, compound **81**, on the development of zebrafish. Embryos were treated at 16-cells stage with different concentration of the drugs directly added in embryo medium in 24-well plates. The phenotype was analyzed 6, 24 and 48 h after treatment. 5-Azacytidine Cu (5-azaC) was used as a positive control and embryos were incubated with vehicle (DMSO) as a negative control. As depicted in Figure 4, 25  $\mu\text{M}$  5-azaC ( $n = 240$ ) induced the expected aberrant morphology: short or absent tail (59%) as well as abnormal somites in comparison to DMSO treated embryos, which induced no remarkable phe-



**Figure 4.** Biological effect of the flavone and flavanone derivatives on embryo development of zebrafish. As with 5-azaC, zebrafish embryos treated with **62** and **47** drugs lacked tails and showed abnormal somites. A) Quantitative analysis after 5-azaC (25  $\mu\text{M}$ ) or **62** and **47** treatments (2.5  $\mu\text{M}$ ). Error bars are standard errors. Mean comparisons were made by using Wilcoxon's test.  $P$  values less than 0.05 were considered significant (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ). B–G) Representative images of one control embryo (B and E) and two **62** treated embryos displaying short tail (C and F) or no axis phenotype (D and G); scale bar: 500  $\mu\text{m}$ . H–J) Global methylation was assessed by immunolabeling for 5-methylcytosine. Detail of representative embryos after 6 h treatment: control (H), 5-azaC (100  $\mu\text{M}$ ; I) and **62** (2.5  $\mu\text{M}$ ; J); scale bar: 100  $\mu\text{m}$ .

notype in 90% of the embryos ( $n = 185$ ). At 25  $\mu\text{M}$ , compounds **47** and **62** killed all embryos ( $n = 40$  and 58). Strikingly at 2.5  $\mu\text{M}$  these molecules ( $n = 223$  and 237) induced a phenotypic profile comparable to 5-azaC treatment; a third of the embryos had no or short tails. Interestingly, genistein and compound **81** induced only high cytotoxicity (13.5%) and a growth delay (36.5%) (Table S3 in the Supporting Information). Compound **69**, the most active on the purified catalytic Dnmt3a/3L complex, did not give any phenotype at 2.5  $\mu\text{M}$ . Global DNA methylation was assessed by fluorescent immunolabeling with an antibody directed against 5-MedC. A decrease in signal was observed upon treatment with 5-azaC (Figure 4I), compounds **62** (Figure 4J) and **47** as compared to untreated samples. Compounds **81** and **69** did not effect global methylation (Table S3 in the Supporting Information).

## Discussion

In this study we implemented a medium-/high-throughput screening assay for DNMT inhibitors. Our first screening campaign identified a family of new potential lead compounds: two 3-nitroflavones and twenty 3-chloro-3-nitroflavanones. Most of them showed submicromolar activity against Dnmt3a/3L compared to 315 and  $>500 \mu\text{M}$  for the reference compounds RG108 and genistein, respectively. The HTS developed here is robust, versatile and can be applied to different types of methyltransferases, such as murine Dnmt3a/3L, human DNMT1 and bacterial M.SssI. The main technical improvement of the test compared to the HTS found in the literature<sup>[23,27–29]</sup> resides in the fact that the DNA is immobilized on the plate. Thus, first, washing steps can be carried out to limit false positive and negative results due to fluorescent molecules that interfere with the final quantification of signal or compounds that might inhibit the restriction enzyme step. Second, the methylation and DNA cleavage reactions do not need to be carried out in the same buffer, which it is not always possible.

SAR studies showed the importance of the nitro on position 3 of the flavones and the chloronitro motif on position 3 of the flavanones, although the interactions remain unclear between this electron-withdrawing group ( $\text{NO}_2$ ) and the mainly hydrophobic cavity of Dnmt3a formed by amino acids Pro705, Phe636, Gly638, Glu660 and Arg88 in which it was docked. SAR investigations revealed also the positive influence of electron-withdrawing substituents on the ring B; the largest effect was observed with an  $\text{NO}_2$  in *ortho* (**69**), which according to docking studies presents a potential H bond with the catalytic Cys706. We also observed a synergy when electron-withdrawing groups were on ring B and A (**71**).

SAM competition experiments supported docking studies and indicated that these DNMT inhibitors have a mixed inhibition profile, since they interact both with the DNA and SAM pocket. The similarity of the catalytic pocket of methyltransferases makes it difficult to design specific inhibitors. Compounds **47** and **70** are equally active on histone methyltransferase G9A and bacterial adenine methyltransferase, EcoDam. Interestingly, compound **62** has a clear preference for the DNA methyltransferases and compound **11** is specific for the DNA C-5 methyltransferase. Nitroflavone **11** is the most specific compound for Dnmt3a/3L and its minimal energy conformation is found mainly in the DNA pocket, but gave a  $\text{IC}_{50}$  of only 9.5  $\mu\text{M}$ . It will be thus interesting to improve its inhibitory activity while maintaining its selectivity. Finally, it will also be interesting to develop selective inhibitors of the different DNMTs to better understand their functions.<sup>[9]</sup> In our HTS tests, little differences in activity were observed for human DNMT1 and bacterial M.SssI compared to Dnmt3a/3L; this demonstrates the need for new selective screening campaigns.

Flavones and isoflavones have several targets in cells. Derivatives **18** to **25** have been tested as inhibitors of the cell surface aminopeptidase N/CD13 (APN), which is over-expressed in tumors and plays a critical role in angiogenesis.<sup>[30]</sup> Only compound **23** is efficient against the DNMTs and shows only a modest activity against APN. Compound **15**, inactive on DNMTs, has been shown to have a chemopreventive effect on the formation of aberrant crypt foci (ACF) in rats exposed to subcutaneous injection of azoxyethane.<sup>[31]</sup> Finally, compounds **6**, **8**, **9** and **10** have been tested for their inhibition of the NO pathway during apoptosis of leukemic B-cells.<sup>[32]</sup> None of these molecules was active on DNMTs. Therefore, the 3-nitro derivatives that inhibit the DNMTs seem to be specific for these enzymes, and present eventually an inhibition activity on other methyltransferases, as observed above.

The most active inhibitors were tested for their effect on zebrafish development. Zebrafish is becoming increasingly used as a cancer model.<sup>[33]</sup> Large epigenetic changes occur during the first steps of development of the zebrafish embryo.<sup>[34]</sup> For example, it has been recently demonstrated that the *no tail* gene, which is necessary for notochord and tail formation, undergoes a dynamic regulation of its expression by gene specific de novo methylation of its CpG island.<sup>[35]</sup> Interestingly, DNMT inhibitors **47** and **62** gave the same short axis phenotype as 5-azaC on zebrafish embryo development at a lower concentration. Furthermore, fluorescent immunolabeling detection of 5-methyl-deoxycytidine (5-MedC) confirmed these results: treatment with 5-azaC or compounds **47** and **62** induced a decrease in the fluorescent signal associated with 5-MedC. No phenotype was observed with the in vitro inactive parent compound **81**; this suggests a specific effect. Nevertheless, the most potent inhibitor on the purified enzyme, compound **69**, did not give any phenotype. This could reflect a permeability or stability problem of the molecule in vivo. Since it has been shown that Dnmt3a and G9a cooperate for tissue specific development in zebrafish<sup>[36]</sup> and that the knockdowns of Dnmt1 and Dnmt3 give different phenotypic profiles, this model could be used to further investigate the mechanism of action and the in vivo selectivity of the DNMT inhibitors.

## Conclusions

In conclusion, the screening test presented here for DNMT inhibitors identified new compounds with IC<sub>50</sub> values in the sub-micromolar range. Among them, two 3-chloro-3-nitroflavone derivatives effected zebrafish embryo development similar to the reference inhibitor 5-azaC, and induced a short axis phenotype and a global demethylation of the genome. Further developments are underway to implement these compounds as DNA methylation inhibitors, in vivo.

## Experimental Section

**Chemicals and drugs:** 5-Azacytidine was obtained from Sigma and stored as a stock (25 mM) in acetic acid (50%) at  $-20^{\circ}\text{C}$ . RG-108 was synthesized as described in ref. [18]. All the drugs were dissolved in DMSO (100%) and stored at  $-20^{\circ}\text{C}$ . S-Adenosyl-L-me-

thionine (SAM) was from Sigma–Aldrich and preweighted aliquots were stored at  $-80^{\circ}\text{C}$ .

**Flavones and flavanones chemical library:** Compounds **26–30** and **98** are commercially available. Synthesis of compound **25** is described in ref. [37]; compounds **1**, **2**, **31** and **32** are described in ref. [38]; **46** and **51** are described in ref. [39]; **9**, **40**, **41**, **47**, **48**, **52**, **64–70**, **74–80** and **83** are described in ref. [40]; **35** is described in ref. [41]; **88–94** and **97** are described in ref. [42]; **3**, **14**, **85–87**, **95**, **96** and **110** are described in ref. [43]; **13** is described in ref. [44]; **5–8**, **10**, **15**, **16**, **39**, **43**, **45**, **53**, **61–63**, **71–73** and **81** are described in ref. [42]; **17–24**, **44**, **50**, **54–57**, **106** and **107** are described in ref. [30]; **58** is described in ref. [45]; **108** is described in ref. [46]; and **109** is described in ref. [47]. Novel compounds **4**, **11**, **33**, **34**, **36**, **38**, **59**, **60**, **82**, **84**, **99**, **100–105**, and **111–114** were synthesized as described in the Supporting Information.

### Screening test

**Substrate design:** DNA duplexes were made by hybridization of complementary oligonucleotides (Eurogentec, Belgium) bearing a 6-carboxyfluorescein (6-FAM) and a biotin (biot), at the 5'-end of each strand. They contained a single CpG site that overlapped with a restriction site of a methylation sensitive restriction enzyme. Substrates designed for the assay with the Dnmt3a/3L complex, M.SssI and DNMT1 have the following sequences (the CpG site is in bold):

- Dnmt3a/3L and M.SssI: FAM-5'-GCTATATATA **CGTACTGTGA** ACCTACCAG ACATGCACTG-3'; 3'-CGATATATAT **GCATGACACTTGG**-GATGGTC TGACGTGAC-5'-biot
- DNMT1: FAM-5'-GCATATATAT **GAmCGATCCT** GTAGGTCACCT AC-CAGACAT GCACTG-3'; 3'-CGTATATATA **CTGCTAGGAC** ATCCAGT-GATGGTCTGTACG TGAC-5'-biot

where *mC* is 5-methyl-deoxycytidine.

**Coating of the test plates:** Costar 96-well high-binding EIA/RIA plates (ref. 9018) were coated with avidin (1  $\mu\text{g}$  per well; Sigma–Aldrich) in NaHCO<sub>3</sub> (100  $\mu\text{L}$  of 100 mM solution; pH 9.6) at  $4^{\circ}\text{C}$ , overnight. The plate was then washed five times with PBST (1  $\times$  PBS, 0.05% Tween-20) and NaCl (500 mM). The plate could be stored for two weeks at  $4^{\circ}\text{C}$ . To further coat the substrate on the plate, DNA duplex (25 pmol per well) was incubated in PBST (100  $\mu\text{L}$ ) at room temperature for at least 30 min. The plate was finally ready to use after being washed three times with PBST + NaCl (500 mM) and three times with PBST.

**Dnmt3a/3L methylation reaction:** The C-terminal catalytic domain of the murine Dnmt3a (623–908) and the C-terminal domain of DNMT3L (208–421), obtained as described in ref. [19], were preincubated together at 200 nM for 20 min at room temperature in the reaction buffer (20 mM HEPES, pH 7.2, 50 mM KCl, 1 mM EDTA) in a total volume of 55  $\mu\text{L}$  per well of a Greiner 96-well V-form transfer microplate. Tested compounds were added either after the preincubation of the enzymatic complex or during the complex formation step. Aliquots (50  $\mu\text{L}$ ) from the total volume (57  $\mu\text{L}$ ) in each well of the preincubation plate were transferred into the corresponding well of the testing plate coated with the DNA substrate. Freshly dissolved SAM was added to initiate the reaction (20  $\mu\text{M}$  final concentration). Methylation was achieved over 1 h at  $37^{\circ}\text{C}$ . Each well was then washed three times with PBST and NaCl (500 mM) and three times with PBST.

**Methylation reaction with DNMT1:** The human GST-tagged DNMT1 was obtained from BPS Bioscience. The enzyme (350 nM) was incubated in reaction buffer (20 mM HEPES, pH 7.2, 50 mM KCl, 1 mM

EDTA) and SAM (20  $\mu\text{M}$ ) in the presence of the tested compound in a total volume of 50  $\mu\text{L}$  per well of a testing plate. The methylation reaction was achieved at 37  $^{\circ}\text{C}$  during 120 min. Each well was washed as described above.

**Methylation reaction with M.SssI:** DNA methylation by the bacterial CpG methyltransferase M.SssI (New England Biolabs) in the presence of the tested compound in a total volume of 50  $\mu\text{L}$  per well during 1 h at 37  $^{\circ}\text{C}$  was realized under the following conditions: M.SssI (50 nM), SAM (20  $\mu\text{M}$ ), Tris-HCl (10 mM; pH 7.9), NaCl (50 mM),  $\text{MgCl}_2$  (10 mM), DTT (1 mM). Each well was washed as described above.

**Restriction reactions:** For the Dnmt3a/3L and M.SssI assays the methylation-sensitive restriction enzyme, HpyCH4IV (2 units per well; New England Biolabs) was incubated in 50  $\mu\text{L}$  total volume of restriction buffer (10 mM Bis-Tris-propane-HCl, 10 mM  $\text{MgCl}_2$ , 1 mM DTT, pH 7.0) for 1 h at 37  $^{\circ}\text{C}$ . Washing conditions were the same as described above. For the DNMT1 assay the methylation-sensitive restriction enzyme, BfuCI (2 units per well; New England Biolabs) was incubated in 50  $\mu\text{L}$  total volume of the same restriction buffer. Washing conditions were as described above.

**Fluorescence detection:** The fluorescent signal of the DNA substrate was measured on a Typhoon 9410 scanner (GE Healthcare). Fluorescence (excitation at 488 nm; emission at 520 nm) was measured after: 1) coating of the microplate with DNA to verify the capture of sufficient amount of substrate, 2) methylation, and 3) completion of the reaction. Up to ten plates could be scanned in a single run. Quantification was carried out automatically by measuring the sum of pixels in each well.

**Quantification and quality controls:** We used two different quality indicators to calibrate each test. The first one was the global percentage of methylation defined as  $100 \times ((\mu_{\text{meth}} - \mu_{\text{rest}}) / (\mu_{\text{DNA}} - \mu_{\text{rest}}))$ ; where  $\mu_{\text{meth}}$ ,  $\mu_{\text{rest}}$  and  $\mu_{\text{DNA}}$  are the average signal of the methylation, restriction and DNA controls, respectively. The second one was the Z factor described in ref. [48]. We define the Z factor of the assay as the parameter in which  $\text{SD}_{\text{drugs}}$  and  $\text{SD}_{\text{rest}}$  are the observed standard deviations of the signal of the tested drugs and restriction controls, respectively, and  $\mu_{\text{drugs}}$  and  $\mu_{\text{rest}}$  are the means of the corresponding fluorescence signals. The Z' factor is based only on the control wells and serves as an intrinsic quality parameter of the assay, without any assumptions, such as the number of expected hits. It is defined as  $1 - 3 \times ((\text{SD}_{\text{meth}} + \text{SD}_{\text{rest}}) / (\mu_{\text{meth}} - \mu_{\text{rest}}))$ , where  $\mu_{\text{meth}}$  and  $\text{SD}_{\text{meth}}$  are the mean signals corresponding to the methylation controls and the associated standard deviations. Experiments showing a Z' factor  $< 0.20$  were omitted for analysis.

A molecule with a fluorescent signal inferior or equal to the threshold of  $(\mu_{\text{rest}} + 6 \times \text{SD}_{\text{rest}})$  is considered as a fully active compound at the given concentration. For the determination of the  $\text{IC}_{50}$  (concentration of drug needed to obtain 50% of inhibition) each experimental set was performed in triplicate and the results were plotted as relative methylation activity against the log of inhibitor concentration.  $\text{IC}_{50}$  values were evaluated after being fitted the dose-response plots by nonlinear regression through the GraphPad Prism software with constrained variable slope equations. Mean  $\text{IC}_{50}$  values are given  $\pm$  standard error defined as  $\text{SD}_{\text{sample}} / \sqrt{(n-1)}$ , where  $\text{SD}_{\text{sample}}$  is the standard deviation of the observed  $\text{IC}_{50}$  values and  $n$  the number of experiments.

To determine the type of inhibition for selected compounds, the assay was run by using several SAM concentrations (from 20 to 0.05  $\mu\text{M}$ ) with the same amount of DNA to get a maximal coating of oligonucleotides on the plate, and several inhibitor concentra-

tions. The data (initial rates vs. concentrations) were analyzed by using nonlinear regression analysis with the GraphPad Prism software to evaluate the best fitting model of inhibition and the corresponding kinetic parameters.

**Radioactive enzymatic tests:** DNA methylation activity of the Dnmt3a/3L complex was measured by incorporation of tritiated methyl groups from labeled S-[methyl- $^3\text{H}$ ] SAM (specific activity 2.9 TBq  $\text{mmol}^{-1}$ , Perkin-Elmer) into a DNA duplex substrate containing 8-CpG sites, with the following sequence: 5'-AGGGGACGAA GGAGGGAAGG AAGGGCAAGG CGGGGGGGGC TCTGCGAGAG CGCGCCAGC CCCGCCTTCG GGCCCCACAG.

The methylation reactions were carried out in the corresponding buffer (20 mM HEPES pH 7.2, 50 mM KCl, 1 mM EDTA) by using DNA (200 nM), radiolabeled SAM (280 nM) and DNMTs (0.5  $\mu\text{M}$ ). Similar experiments were carried out by using the purified bacterial EcoDam N-6 DNA methyltransferase and the catalytic domain of the human G9a histone H3K9 methyltransferase, as described in refs. [49,24], respectively. Each experimental set was performed in triplicate and the results were plotted as relative methylation activity (expressed in %) against the inhibitor concentration. To obtain the apparent inhibition constant ( $\text{IC}_{50}$ ) data were fitted to the following Equation (1):

$$\text{signal } (C_i) = \text{BL} + 100 \times \text{IC}_{50} / (\text{IC}_{50} + C_i) \quad (1)$$

where BL is baseline,  $C_i$  is the concentration of inhibitor and  $\text{IC}_{50}$  is the concentration to obtain 50% of apparent inhibition.

**Docking studies:** Compounds were docked in the crystal structure of catalytic Dnmt3a (PDB ID: 2QRV, chain A) EcoDam (PDB ID: 2G1P, chain A) and G9a (PDB ID: 3K5K, chain A) with Dock 6.4. Ligand structures were prepared for the docking process with Marvin (Chemaxon). Default parameters were applied. Chimera was used for the graphical visualization.

**Zebrafish experiments:** Fertilized eggs were obtained from natural mating of adult zebrafish maintained under standard conditions.<sup>[50]</sup> Embryos were incubated in an aqueous solution of the various substrates (from stage 16 cells to analysis). Embryos were manually dechorionated before malformations were scored or 5-mC was detected with a specific antibody (anti-5-methylcytosine antibody, Calbiochem). Error bars are statistical errors estimated as  $(\sqrt{p(1-p)/n})$ , where  $p$  is the percentage of embryos exhibiting a phenotype, and  $n$  the total number of embryos investigated (or  $\sqrt{1/n}$  when  $p=0$  or 1). For immunohistochemistry, embryos were treated, as previously described,<sup>[51]</sup> with minor modifications. After incubation with the anti-5-methylcytosine antibody, the embryos were washed again and incubated first with biotinylated anti-mouse secondary antibody (GE Healthcare) and then with streptavidin-fluorescein (GE Healthcare).

## Acknowledgements

The authors thank Lionel Dubost and Dr. Arul Marie of the Mass Spectroscopy Service of the MNHN for mass spectra analysis. This work was supported by ANR JCJC06 JCJC06 0080-01 (to P.B.A.) and by DFG grant JE 252/6 (to A.J.). A.C. was recipient of an ARC fellowship and C.C. of an INSERM-CNRS BDI.

**Keywords:** DNA methylation · DNMT · flavones · high-throughput screening · inhibitors

- [1] A. Bird, *Nature* **2007**, *447*, 396.
- [2] S. L. Berger, T. Kouzarides, R. Shiekhattar, A. Shilatifard, *Genes Dev.* **2009**, *23*, 781.
- [3] S. Sharma, T. K. Kelly, P. A. Jones, *Carcinogenesis* **2010**, *31*, 27.
- [4] J. J. Day, J. D. Sweatt, *Nat. Neurosci.* **2010**, *13*, 1319.
- [5] G. Miller, *Science* **2010**, *329*, 24.
- [6] A. P. Feinberg, *Nature* **2007**, *447*, 433.
- [7] A. Portela, M. Esteller, *Nat. Biotechnol.* **2010**, *28*, 1057.
- [8] M. Esteller, *N. Engl. J. Med.* **2008**, *358*, 1148.
- [9] X. Cheng, R. M. Blumenthal, *Structure* **2008**, *16*, 341.
- [10] T. K. Kelly, D. D. De Carvalho, P. A. Jones, *Nat. Biotechnol.* **2010**, *28*, 1069.
- [11] C. B. Yoo, P. A. Jones, *Nat. Rev. Drug Discovery* **2006**, *5*, 37.
- [12] A. Mai, L. Altucci, *Int. J. Biochem. Cell Biol.* **2009**, *41*, 199.
- [13] J. S. Garcia, N. Jain, L. A. Godley, *Oncotargets Ther.* **2010**, *3*, 1.
- [14] C. Stresemann, F. Lyko, *Int. J. Cancer* **2008**, *123*, 8.
- [15] M. Candelaria, D. Gallardo-Rincon, C. Arce, L. Cetina, J. L. Aguilar-Ponce, O. Arrieta, A. Gonzalez-Fierro, A. Chavez-Blanco, E. de La Cruz-Hernandez, M. F. Camargo, C. Trejo-Becerril, E. Perez-Cardenas, C. Perez-Plasencia, L. Taja-Chayeb, T. Wegman-Ostrosky, A. Revilla-Vazquez, A. Duenas-Gonzalez, *Ann. Oncol.* **2007**, *18*, 1529.
- [16] M. Z. Fang, Y. Wang, N. Ai, Z. Hou, Y. Sun, H. Lu, W. Welsh, C. S. Yang, *Cancer Res.* **2003**, *63*, 7563.
- [17] M. Pandey, S. Shukla, S. Gupta, *Int. J. Cancer* **2010**, *126*, 2520.
- [18] B. Brueckner, R. G. Boy, P. Siedlecki, T. Musch, H. C. Kliem, P. Zielenkiewicz, S. Suhai, M. Wiessler, F. Lyko, *Cancer Res.* **2005**, *65*, 6305.
- [19] D. Jia, R. Z. Jurkowska, X. Zhang, A. Jeltsch, X. Cheng, *Nature* **2007**, *449*, 248.
- [20] W. Qin, W. Zhu, H. Shi, J. E. Hewett, R. L. Ruhlen, R. S. MacDonald, G. E. Rottinghaus, Y. C. Chen, E. R. Sauter, *Nutr. Cancer* **2009**, *61*, 238.
- [21] G. Schluckebier, M. Kozak, N. Bleimling, E. Weinhold, W. Saenger, *J. Mol. Biol.* **1997**, *265*, 56.
- [22] H. Gowher, K. Liebert, A. Hermann, G. Xu, A. Jeltsch, *J. Biol. Chem.* **2005**, *280*, 13341.
- [23] M. Roth, A. Jeltsch, *Biol. Chem.* **2000**, *381*, 269.
- [24] P. Rathert, X. Cheng, A. Jeltsch, *Biotechniques* **2007**, *43*, 602.
- [25] S. Urig, H. Gowher, A. Hermann, C. Beck, M. Fatemi, A. Humeny, A. Jeltsch, *J. Mol. Biol.* **2002**, *319*, 1085.
- [26] C. C. Martin, L. Laforest, M. A. Akimenko, M. Ekker, *Dev. Biol.* **1999**, *206*, 189.
- [27] Y. H. Woo, P. T. Rajagopalan, S. J. Benkovic, *Anal. Biochem.* **2005**, *340*, 336.
- [28] F. Li, M. Papworth, M. Minczuk, C. Rohde, Y. Zhang, S. Ragozin, A. Jeltsch, *Nucleic Acids Res.* **2007**, *35*, 100.
- [29] Y. Ye, J. T. Stivers, *Anal. Biochem.* **2010**, *401*, 168.
- [30] B. Bauvois, M. L. Puiffe, J. B. Bongui, S. Paillat, C. Monneret, D. Dauzonne, *J. Med. Chem.* **2003**, *46*, 3900.
- [31] V. E. Steele, C. W. Boone, D. Dauzonne, C. V. Rao, R. V. Bensasson, *Cancer Res.* **2002**, *62*, 6506.
- [32] C. Billard, C. Quiney, R. Tang, C. Kern, F. Ajchenbaum-Cymbalista, D. Dauzonne, J. P. Kolb, *Ann. N. Y. Acad. Sci.* **2003**, *1010*, 381.
- [33] H. Feitsma, E. Cuppen, *Mol. Cancer Res.* **2008**, *6*, 685.
- [34] A. A. Mhanni, R. A. McGowan, *Dev. Genes Evol.* **2004**, *214*, 412.
- [35] K. Yamakoshi, N. Shimoda, *Genesis* **2003**, *37*, 195.
- [36] A. C. Smith, A. R. Raimondi, C. D. Salthouse, M. S. Ignatius, J. S. Blackburn, I. V. Mizgirev, N. Y. Storer, J. L. de Jong, A. T. Chen, Y. Zhou, S. Revskoy, L. I. Zon, D. M. Langenau, *Blood* **2010**, *115*, 3296.
- [37] G. Atassi, P. Briet, J.-J. Berthelon, F. Collonges, *Eur. J. Med. Chem. Chim. Ther.* **1985**, *20*, 393.
- [38] D. Dauzonne, R. Royer, *Chem. Pharm. Bull.* **1986**, *34*, 1628.
- [39] D. Dauzonne, P. Demerseman, *Synthesis* **1990**, 66.
- [40] D. Dauzonne, C. Grandjean, *Synthesis* **1992**, 677.
- [41] D. Dauzonne, C. Grandjean, *J. Heterocycl. Chem.* **1994**, *31*, 1021.
- [42] D. Dauzonne, C. Monneret, *Synthesis* **1997**, 1305.
- [43] A. Gonzalez De Peredo, S. Leonce, C. Monneret, D. Dauzonne, *Chem. Pharm. Bull.* **1998**, *46*, 79.
- [44] C. Beudot, M. P. De Meo, D. Dauzonne, R. Elias, M. Laget, H. Guiraud, G. Balansard, G. Dumenil, *Mutat. Res.* **1998**, *417*, 141.
- [45] M. H. Pham, N. Auzeil, A. Regazzetti, D. Dauzonne, A. Dugay, M. C. Menet, D. Scherman, G. G. Chabot, *Drug Metab. Dispos.* **2007**, *35*, 2023.
- [46] B. M. Choudary, K. V. S. Ranganath, J. Yadav, M. L. Kantam, *Tetrahedron Lett.* **2005**, *46*, 1369.
- [47] S. Gester, P. Metz, O. Zierau, G. Vollmer, *Tetrahedron* **2001**, *57*, 1015.
- [48] J. H. Zhang, T. D. Chung, K. R. Oldenburg, *J. Biomol. Screening* **1999**, *4*, 67.
- [49] J. R. Horton, K. Liebert, M. Bekes, A. Jeltsch, X. Cheng, *J. Mol. Biol.* **2006**, *358*, 559.
- [50] M. Westerfield, *The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio rerio)*, **1995**, 3rd ed., University of Oregon Press, Eugene, p. 385.
- [51] A. B. MacKay, A. A. Mhanni, R. A. McGowan, P. H. Krone, *Genome* **2007**, *50*, 778.

Received: February 24, 2011

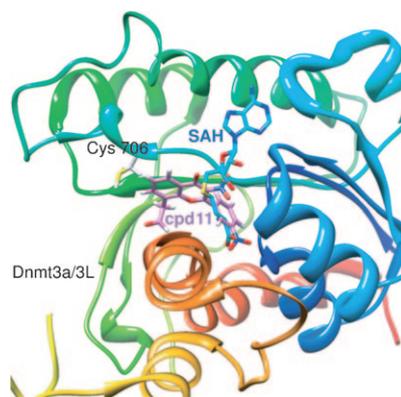
Published online on ■■■■■, 0000

## FULL PAPERS

A. Ceccaldi, A. Rajavelu, C. Champion,  
C. Rampon, R. Jurkowska, G. Jankevicius,  
C. Sénamaud-Beaufort, L. Ponger,  
N. Gagey, H. Dali Ali, J. Tost, S. Vríz,  
S. Ros, D. Dauzonne, A. Jeltsch,  
D. Guianvarc'h, P. B. Arimondo\*



**C5-DNA Methyltransferase Inhibitors:  
From Screening to Effects on  
Zebrafish Embryo Development**



**Screen shots:** A set of flavones and flavanones is described as a new family of nanomolar DNA methyltransferase inhibitors. Structure–activity relationships, enzymatic competition and docking studies (see figure) have provided insights on their mode of action. Two inhibitors effected zebrafish development in a similar fashion to the reference drug, 5-azacytidine.

# The Dnmt3a PWWP Domain Reads Histone 3 Lysine 36 Trimethylation and Guides DNA Methylation<sup>\*S</sup>

Received for publication, November 30, 2009, and in revised form, June 10, 2010. Published, JBC Papers in Press, June 11, 2010, DOI 10.1074/jbc.M109.089433

Arunkumar Dhayalan<sup>‡</sup>, Arumugam Rajavelu<sup>‡</sup>, Philipp Rathert<sup>§</sup>, Raluca Tamas<sup>¶</sup>, Renata Z. Jurkowska<sup>‡</sup>, Sergey Ragozin<sup>‡</sup>, and Albert Jeltsch<sup>‡1</sup>

From the <sup>‡</sup>Biochemistry Laboratory and the <sup>¶</sup>MoLife Graduate Program, School of Engineering and Science, Jacobs University Bremen, 28759 Bremen, Germany and <sup>§</sup>Active Motif Europe, BE-1330 Rixensart, Belgium

The Dnmt3a DNA methyltransferase contains in its N-terminal part a PWWP domain that is involved in chromatin targeting. Here, we have investigated the interaction of the PWWP domain with modified histone tails using peptide arrays and show that it specifically recognizes the histone 3 lysine 36 trimethylation mark. H3K36me3 is known to be a repressive modification correlated with DNA methylation in mammals and heterochromatin in *Schizosaccharomyces pombe*. These results were confirmed by equilibrium peptide binding studies and pulldown experiments with native histones and purified native nucleosomes. The PWWP-H3K36me3 interaction is important for the subnuclear localization of enhanced yellow fluorescent protein-fused Dnmt3a. Furthermore, the PWWP-H3K36me3 interaction increases the activity of Dnmt3a for methylation of nucleosomal DNA as observed using native nucleosomes isolated from human cells after demethylation of the DNA with 5-aza-2'-deoxycytidine as substrate for methylation with Dnmt3a. These data suggest that the interaction of the PWWP domain with H3K36me3 is involved in targeting of Dnmt3a to chromatin carrying that mark, a model that is in agreement with several studies on the genome-wide distribution of DNA methylation and H3K36me3.

In mammals, DNA methylation plays important roles in differentiation, gene regulation, genomic imprinting, X chromosome inactivation, and disease-related processes (1–3). DNA methylation patterns are set during embryogenesis by the Dnmt3a and Dnmt3b DNA methyltransferases and their regulatory factor Dnmt3L (Dnmt3-like) (4–6). However, the mechanisms guiding these enzymes to their target regions are not well understood. Dnmt3a and 3b consist of a C-terminal catalytic domain and an N-terminal part containing a PWWP domain and an ADD<sup>2</sup> domain (1, 7, 8). Biochemical studies provide evidence for a direct interaction of Dnmt3a and 3b with native nucleosomes (9), which could be mediated by the ADD

domain or the PWWP domain. The ADD domains of Dnmt3L and Dnmt3a have been shown to interact with the histone 3 tail unmethylated at Lys<sup>4</sup> (10–12), which can explain the anticorrelation of DNA methylation and the activating H3K4me3 mark as observed in many genome-wide DNA methylation studies (13–16). However, the ADD domain is not directly involved in heterochromatic targeting of Dnmt3a (17, 18).

PWWP domains belong to the Royal domain superfamily, members of which were identified to interact with histone tails in various modification states (19). The PWWP domains of Dnmt3a and 3b are essential for heterochromatic targeting (17, 18). An S333P missense mutation in the Dnmt3a PWWP domain (numbering refers to murine Dnmt3a) led to the loss of chromatin targeting of Dnmt3a (18). This mutation corresponds to the S282P mutation in the PWWP domain of human Dnmt3b, which has been identified in immunodeficiency, centromeric heterochromatin instability, and facial anomalies syndrome patients (20).

Here, we explore the possibility of the interaction of the Dnmt3a PWWP domain with histone peptides using modified histone tail peptide arrays. We discovered its specific interaction with H3K36me3 and show that this interaction is important for the subnuclear localization of Dnmt3a and for its catalytic activity on native chromatin.

## EXPERIMENTAL PROCEDURES

**Cloning, Site-directed Mutagenesis, Expression, and Purification**—The sequence encoding the PWWP domain of murine Dnmt3a (residues 279–420) was subcloned as a GST fusion protein into the pGEX-6P2 vector (GE Healthcare) using BamHI/XhoI sites. The GST-tagged PWWP domain and its variants were expressed and purified as described (21). The full-length murine Dnmt3a2 cloned into pET28a (Novagen) vector was used for the expression and purification of full-length Dnmt3a2 (22), which is the predominant isoform of Dnmt3a in embryonic stem cells (23). The catalytic domain of Dnmt3a (Dnmt3a-CD) was expressed and purified as described (24). To prepare the EYFP-Dnmt3a expression construct used for subnuclear localization studies, Dnmt3a was cloned from a pET expression vector (25) into pcDNA3.1, yielding pcDNA-Dnmt3a. Then EYFP was amplified from pEYFP-C1 (Clontech) and cloned into pcDNA-Dnmt3a in front of the Dnmt3a gene to encode for an EGFP-Dnmt3a fusion protein.

The E290A and D329A mutations in the PWWP domain of Dnmt3a and the D329A mutation in Dnmt3a2 and in EGFP-Dnmt3a were introduced by using a PCR megaprimer muta-

\* This work was supported, in whole or in part, by National Institutes of Health Grants GM068680 and DK082678. This work was also supported by Deutsche Forschungsgemeinschaft Grant JE 252/6-1.

<sup>S</sup> The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Figs. S1–S6.

<sup>1</sup> To whom correspondence should be addressed: School of Engineering and Science, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany. Tel.: 49-421-200-3247; Fax: 49-421-200-3249; E-mail: a.jeltsch@jacobs-university.de.

<sup>2</sup> The abbreviations used are: ADD, ATRX-DNMT3-DNMT3L; GST, glutathione S-transferase; EYFP, enhanced yellow fluorescent protein.

genesis method as described previously (26). Successful mutagenesis and cloning was confirmed by restriction marker site analysis and DNA sequencing.

**Synthesis of Peptide SPOT Arrays**—Peptide arrays were synthesized on cellulose membranes using the SPOT synthesis method (27). Each spot had diameters of 2 mm and contained ~9 nmol of peptide (Autospot Reference Handbook, Intavis AG). Successful synthesis of each peptide was confirmed by bromphenol blue staining of the membranes. H3 26–44 K36me3 (with N-terminal fluorescein label) and unmodified H3 26–44 peptides were purchased from Intavis AG in purified form.

**Binding of Protein Domains to Peptides Arrays**—The cellulose membrane was blocked by incubation in TTBS buffer (10 mM Tris/HCl, pH 8.3, 0.05% Tween 20, and 150 mM NaCl) containing 5% nonfat dried milk at 4 °C overnight. The membrane was then washed once with TTBS buffer and incubated with purified GST-tagged PWWP domain of Dnmt3a or mutant domains (10 nM) at room temperature for 1 h in interaction buffer (100 mM KCl, 20 mM Hepes, pH 7.5, 1 mM EDTA, 0.1 mM dithiothreitol, and 10% glycerol). After washing in TTBS buffer, the membrane was incubated with goat anti-GST antibody (GE Healthcare, catalog number 27-4577-01 at 1:5000 dilution) in TTBS buffer for 1 h at room temperature. Then the membrane was washed three times with TTBS and incubated with horseradish peroxidase-conjugated anti-goat antibody (Invitrogen catalog number 81-1620 1:12000) in TTBS for 1 h at room temperature. Finally, the membrane was submerged in ECL developing solution (GE Healthcare), and the image was captured in x-ray film.

**Fluorescence Depolarization Peptide Binding Experiments**—Fluorescence depolarization experiments were carried out at 25 °C using a Varian Carry Eclipse fluorescence spectrophotometer. For the determination of the binding constant of the PWWP domain to the fluorophore-coupled H3 K36me3 peptide, 0.1 μM of peptide was incubated with increasing concentrations of PWWP domain (0.8–60 μM) in interaction buffer (100 mM NaCl, 50 mM Tris/HCl, pH 8.0, 2 mM EDTA, 0.1% Triton X-100), and fluorescence depolarization was determined (excitation at 494 nm, emission at 524 nm, and excitation and emission slits at 5 nm). The data were fitted to a binary binding equilibrium to determine the equilibrium binding constant. For competition experiments, 30 μM of fluorophore-coupled H3K36me3 was incubated with 30 μM of PWWP domain in interaction buffer, and their interaction was competed by 2- and 5-fold molar excesses of unlabeled unmodified H3 26–44 peptide.

**Preparation of Native Histones and Nucleosomes**—Native histones were isolated from HEK293 cells by acid extraction as described (28). Native nucleosomes were prepared from HEK293 cells by partial micrococcal nuclease (New England Biolabs) digestion of nuclei as described (29), except that after micrococcal nuclease digestion the nuclei were removed by centrifugation, and the supernatant containing nucleosomes (predominantly mononucleosomes) were used (supplemental Fig. S3). For the preparation of demethylated nucleosomes, HEK293 cells were treated with 2 μM 5-aza-2'-deoxycytidine (Sigma-Aldrich) for 72 h (corresponding to three cell

divisions under these conditions), and the cells were allowed to recover for another 48 h in medium without 5-aza-2'-deoxycytidine. From these cells, nucleosomes were prepared as described above.

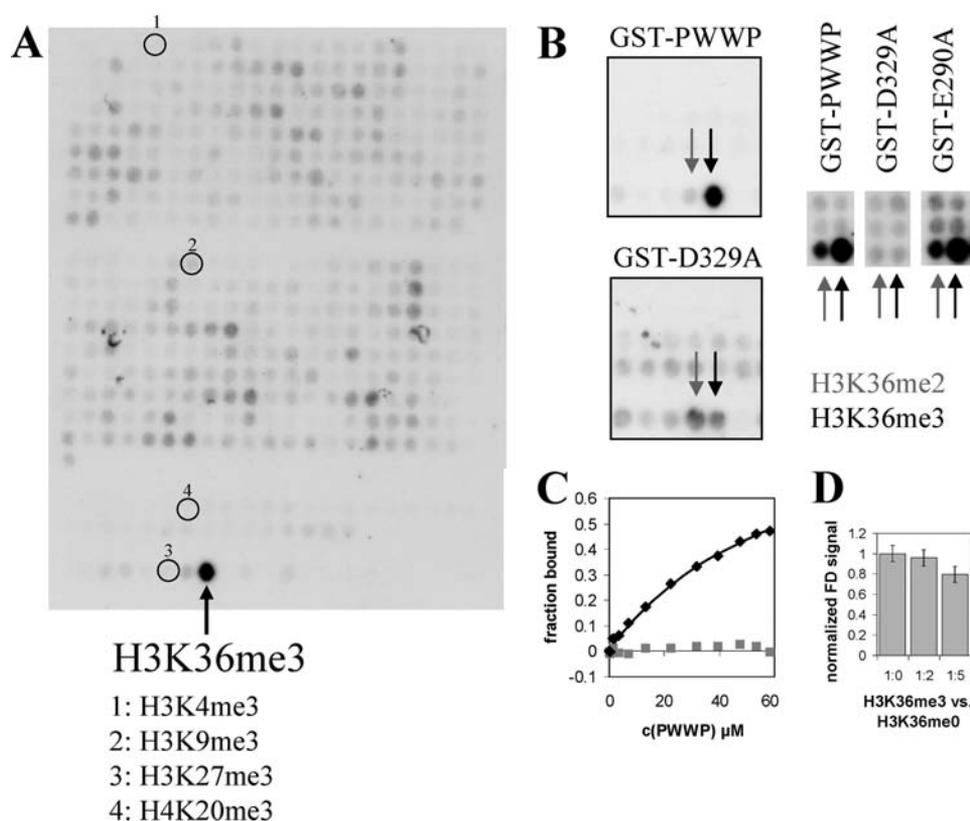
**GST Pulldown Experiments**—For GST pulldown experiments, 20 μl of glutathione-Sepharose 4B beads (Amersham Biosciences) were incubated with 20 μg of GST-tagged PWWP domain of Dnmt3a or mutant domain for 1 h at 4 °C in the interaction buffer (50 mM Tris/HCl, pH 8.0, 100 mM NaCl, 2 mM EDTA, and 0.1% Triton X-100). Then the beads were washed once with interaction buffer and blocked with interaction buffer containing 5% bovine serum albumin for 1 h at 4 °C. 35 μg of native histones or 100 μg of native oligonucleosomes were added to the beads in interaction buffer and incubated for 3 h. Finally, the beads were washed three times with interaction buffer, resuspended in 2× SDS-PAGE loading dye, and boiled for 10 min, and the supernatant containing the bound fractions was separated on 16% SDS-PAGE and transferred to nitrocellulose membrane. The membrane was cut in two parts. The lower part (<20 kDa) was probed with anti-H3K36me3 antibody (Abcam ab9050), and the upper part was probed with anti-GST antibody (GE Healthcare) to provide a gel loading control (supplemental Fig. S4, A and B).

For the specificity analysis, the beads were loaded with GST-tagged PWWP domain and incubated with native oligonucleosomes in interaction buffer as described above, washed three times with washing buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 2 mM EDTA, and 0.1% Triton X-100), resuspended in 2× SDS-PAGE loading buffer and boiled for 10 min. The supernatant was split equally and loaded on 16% SDS-PAGE gels. The separated proteins were transferred to nitrocellulose membrane and probed with different modification specific histone antibodies: anti-H3K36me3 (Abcam ab9050), anti-H3K9me3 (Active Motif 39161), anti-H3K4me3 (Active Motif 39159), and anti-H3K27me3 (Active Motif 39535). Equal loading of bound fractions was confirmed by Ponceau S staining of the blot prior to the antibody staining (supplemental Fig. S4C).

**Cell Culture, Transfection, and Microscopy**—HEK293 and NIH3T3 cells were grown in Dulbecco's modified Eagle's medium with 5% fetal calf serum at 37 °C at 5% CO<sub>2</sub>. The NIH3T3 cells were seeded in coverslips and transfected with EYFP-Dnmt3a or EYFP-Dnmt3a-D329A mutant constructs using FuGENE 6 (Roche Applied Science) according to the manufacturer's instructions. Two days after transfection, the cells were fixed with 4% paraformaldehyde and embedded with Mowiol (Carl Roth). Confocal images were taken using a Zeiss LSM 510 Meta instrument (software version 3.0) and 63× oil immersion objective. An argon laser line at 514 nm was used to excite EYFP fluorescence, and a BP530–550 filter was used for image recording. Fluorescence microscopy pictures were taken using an AXIOPLAN2 microscope equipped with AxioCam HRC camera and Achromplan 100× oil immersion objective with filter sets 4',6'-diamino-2-phenylindole FT395/LP420 and green fluorescent protein BP450–490/FT510 (all from Carl Zeiss).

**In Vitro Methylation of Native Chromatin**—The methylation reactions were performed with either Dnmt3a2 enzyme, Dnmt3a2-D329A mutant, or Dnmt3a-CD in methylation

## Dnmt3a PWWP Interacts with H3K36me3



**FIGURE 1. Peptide binding by the Dnmt3a wild type PWWP domain and its variants.** *A*, interaction of the wild type PWWP domain with histone tail arrays containing different combinations of 107 known and hypothetical modifications at the H3, H4, H2A, H2B, and H1b histone tails. The isolated GST domain does not interact with peptide arrays (data not shown). The spot corresponding to the H3 29–46 K36me3 peptide is indicated with an arrow; some other spots that do not interact with the PWWP domain are labeled with numbers and annotated below the image. *B*, interaction of wild type PWWP domain, D329A, and E290A variants with modified histone tails. The H3K36me3 and H3K36me2 spots on the different arrays used for this study are highlighted with black and gray arrows, respectively. *C*, binding of the H3K36me3 peptide to PWWP domain studied by fluorescence depolarization (black diamonds). The line represents a fit of the data to a binary binding equilibrium, which revealed a  $K_d$  of 64  $\mu\text{M}$ . The gray squares display changes in fluorescence depolarization observed after adding same volumes of buffer. *D*, weak competition of H3K36me3 peptide binding to the PWWP domain by the addition of unmethylated H3K36 peptide confirms specific binding.

buffer (20 mM Hepes, pH 7.5, 2 mM EDTA, and 50 mM KCl) containing 0.75  $\mu\text{M}$  radioactively labeled *S*-adenosyl-*L*-methionine (PerkinElmer Life Sciences) and 1  $\mu\text{g}$  of demethylated native chromatin in 25  $\mu\text{l}$  of total reaction volume. Additional methylation experiments were conducted in the presence of anti-H3K36me3 antibody (Abcam ab9050), anti-H3K9me3 antibody (Active Motif 39161), anti-H3K4me3 antibody (Active Motif 39159), or anti-H3K27me3 antibody (Active Motif 39535) all in 1:5000 final dilution. The methylation reactions were incubated at 37  $^{\circ}\text{C}$  for 2 h, and the reactions were stopped by the addition of excess of unlabeled *S*-adenosyl-*L*-methionine. Then the reaction mixtures were subjected to proteinase K (New England Biolabs) digestion at 42  $^{\circ}\text{C}$  for 2 h, and the nucleic acids in the reaction mixture were purified using Nucleospin PCR purification (Machery Nagel). The amount of radioactivity incorporated into the DNA was measured by liquid scintillation counting and normalized with respect to the total amount of DNA as determined by UV spectroscopy. Methylation of biotinylated oligonucleotide substrates by Dnmt3a2 and its variants was studied by measuring the transfer of titrated methyl groups from labeled *S*-adenosyl-*L*-methionine into biotinylated oligonucleotides (Bt-GAG AAG CTG

GGA CTT CCG GGA GGA GAG  
TGC/GCA CTC TCC TCC CGG  
AAG TCC CAG CTT CTC) using the avidin-biotin methylation assay as described (22, 30).

## RESULTS

**Peptide Array Binding and Structural Modeling**—To investigate whether the Dnmt3a PWWP domain interacts with modified histone peptides, we employed two different modified histone tail peptide arrays containing in total more than 1000 peptides in various combinations of known and hypothetical modification states of the H3, H4, H2A, and H2B tails. We repeatedly observed that the GST-tagged PWWP domain of Dnmt3a interacted with the histone H3 K36 trimethylation mark with high specificity (Fig. 1).

Binding to trimethylated lysine usually involves hydrophobic cages, which often have an acidic residue nearby (19). Because the PWWP domains from Dnmt3a and 3b share high amino acid homology, we used the structure of the Dnmt3b PWWP domain (31) to model the structure of the Dnmt3a PWWP domain and identify such potential binding pockets next to Glu<sup>290</sup> and Asp<sup>329</sup> (supplemental Fig. S1). Trying to disrupt the histone

interaction of PWWP, we mutated Glu<sup>290</sup> and Asp<sup>329</sup> to alanine. These residues are both exposed on the surface of the Dnmt3a PWWP structural model, suggesting that their exchange will not affect the overall domain structure. We did not mutate the buried hydrophobic residues of the predicted pockets to avoid effects on protein folding and stability. The mutant domain proteins could be expressed and purified with similar yields as the wild type domain (supplemental Fig. S2A), and CD spectroscopy confirmed the wild type-like folding of the D329A mutant domain (supplemental Fig. S2B). Binding experiments on peptide arrays showed that there was no significant change in the peptide interaction of the E290A variant, but the D329A variant had completely lost its ability to bind specifically to the H3K36me3 peptide (Fig. 1B).

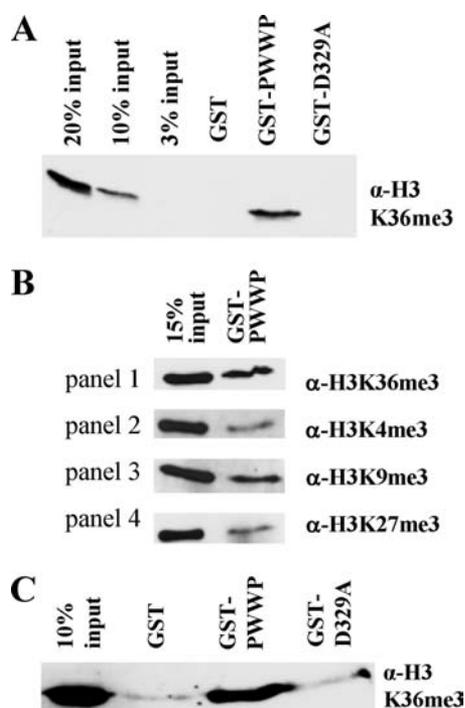
Recently, the structure of the Eaf3 chromo barrel domain was solved in complex with an H3K36me2 analog (32). Over 144 backbone atoms comprising the conserved  $\beta$ -barrel and the adjacent  $\alpha$ -helix fold, the Eaf3 chromo barrel domain structure could be superimposed with the mouse Dnmt3b PWWP domain structure with a root mean square deviation of 1.37  $\text{Å}$  (supplemental Fig. S1). The overlay positioned the methylated Lys<sup>36</sup> residue immediately adjacent to Asp<sup>329</sup> in the Dnmt3a

PWWP domain model. This observation indicates that the biochemical data obtained here are compatible with the structural results obtained with related domains. Furthermore, the Asp<sup>329</sup> residue is close to the Ser<sup>333</sup> residue (supplemental Fig. S1), mutation of which leads to loss of heterochromatic localization (18).

**Equilibrium Peptide Binding Experiments**—The interaction of the PWWP domain with H3K36me3 was confirmed by fluorescence depolarization using a fluorescently labeled H3 26–44 K36me3 peptide (Fig. 1C). The binding constant was determined as 64  $\mu\text{M}$ , which is weak but in the range of results observed with other individual chromatin-binding domains, like BHC80 Plant Homeodomain like domain binding to H3K4me0 peptide: 33  $\mu\text{M}$  (33), L3MBTL1 to the H3K9me1 peptide: 26  $\mu\text{M}$  (34), BRD2 bromodomain 1 to the H4K5ac/K12ac peptide: 360  $\mu\text{M}$  (35), and Brd4 bromodomain 1 binding to H3K14ac: 118  $\mu\text{M}$  (36). Peptide binding of the PWWP-D329A variant was at least five times weaker (data not shown). To assess the binding specificity, binding was competed with unmodified H3 26–44 peptide (Fig. 1D). We observed that at a 5-fold excess of competitor only 20% of binding was lost, corresponding to a 25-fold preference for binding to the methylated peptide.

**Nucleosome and Histone Pulldown Experiments**—The PWWP-H3K36me3 interaction was also confirmed by pull-down experiments. The GST-PWWP domain of Dnmt3a was able to pull down native nucleosomes isolated from human cells, which contained the H3K36me3 mark (Fig. 2A and supplemental Figs. S3A and 4A). In contrast, the PWWP-D329A mutant completely failed to interact with nucleosomes (Fig. 2A and supplemental Fig. S4A). In addition to H3K36me3, the nucleosome fraction bound by the PWWP domain was also enriched with H3K9me3, but the amounts of H3K4me3 and H3K27me3 were reduced in comparison with H3K36me3 (Fig. 2B). The relative reduction in the amounts of H3K4me3 and K27me3 indicates that the pull-down is specific. The presence of the H3K9me3 mark in the pull down was expected on the basis of the overlapping distribution profiles of the H3K36me3 and H3K9me3 marks across transcribed chromatin (37). Using native histones isolated from human cells (supplemental Fig. S3B), we observed that the GST-PWWP domain of Dnmt3a, but not its D329A mutant, interacted with histone H3K36me3 mark (Fig. 2C and supplemental Fig. S4B), indicating that the interaction does not require DNA.

**Subnuclear Localization of Dnmt3a**—We investigated the subnuclear distribution of the EYFP-fused Dnmt3a and its D329A mutant in NIH3T3 cells after transient expression of the proteins. Both proteins were located in the nucleus, which was expected because the nuclear localization signals of Dnmt3a are outside of the PWWP domain (17). Within the nucleus, Dnmt3a was found in large spots corresponding to 4',6'-diamino-2-phenylindole-stained heterochromatin as reported earlier (17, 18) (Fig. 3 and supplemental Fig. S5). In contrast, the mutant protein was found more homogeneously distributed in the nucleus, indicating a partial loss of heterochromatic localization. This result suggests that the interaction of PWWP with K36me3 is important for the localization of Dnmt3a to heterochromatin.



**FIGURE 2. Native chromatin and histone binding of the Dnmt3a PWWP domain and its variants.** A, GST pull-down assay interaction analysis of the PWWP domain and the D329A variant with mononucleosomes purified from human cells. The bound proteins were separated and immunoblotted with anti-H3K36me3 antibody. B, for specificity analysis, the bound fractions were separated and immunoblotted with anti-H3K36me3 antibody, anti-H3K4me3 antibody, anti-H3K9me3 antibody, and H3K27me3 antibody. C, GST pull-down interaction analysis of PWWP domain and D329A mutant domain with native histone proteins, purified from human cells. The bound fractions were separated and immunoblotted with anti-H3K36me3 antibody. In A–C, all of the bands correspond to histone H3.

	YFP-Dnmt3a		YFP-Dnmt3a D329A	
	Examples	Fraction of cells	Examples	Fraction of cells
Homogenous	Not observed	0 %		82 %
Homogenous with some large spots		38 %		18 %
Mainly large spots		62 %	Not observed	0 %

**FIGURE 3. Subnuclear distribution of wild type EYFP-Dnmt3a and EYFP-Dnmt3a-D329A mutant in NIH3T3 cells.** The subnuclear localization patterns were analyzed by laser scanning microscopy and assigned into one of the following categories: localization mainly homogenous, homogenous localization with some large spots, and localization almost exclusively in large spots. The figure shows examples of each type of localization patterns and gives the distribution of occurrence of the different patterns in both experiments in 100 arbitrarily chosen cells.

**DNA Methylation Experiments**—The influence of the PWWP-H3K36me3 interaction on the methylation of nucleosomal DNA by Dnmt3a was investigated using native nucleosomes from human cells prepared after treatment of the cells with 5-aza-2'-deoxycytidine. This treatment reduces DNA methylation at unique sequences and repeats by more than 50% (15, 38). The native nucleosomes carry their endogenous histone modification pattern and partially demethylated DNA. This experimental design allowed us to study the methylation activity of Dnmt3a using a substrate that mimics the *in vivo*

## Dnmt3a PWWP Interacts with H3K36me3

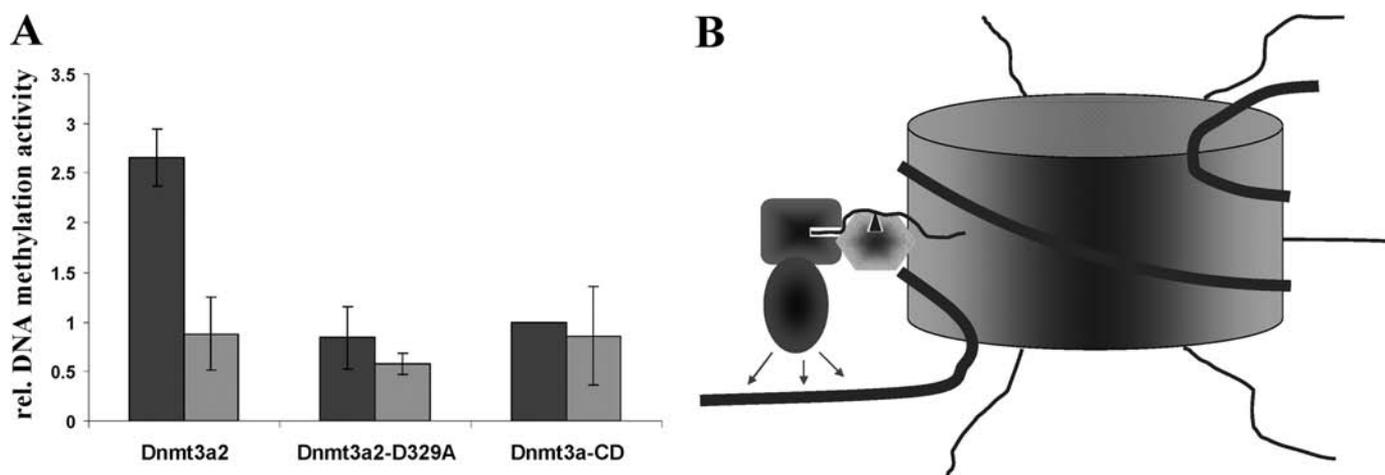


FIGURE 4. *A*, methylation of demethylated native nucleosomes by Dnmt3a enzymes. Relative methylation activity of Dnmt3a2, Dnmt3a2-D329A, and Dnmt3a-CD on demethylated native nucleosomes in the absence (dark gray bars) and the presence (light gray bars) of anti-H3K36me3 antibody. The error bars indicate standard deviations derived from at least three independent experiments. *B*, schematic picture showing how the H3 tail interactions of the Dnmt3a ADD domain (dark grey rectangle) binding to the end of the H3 tail (12) and the PWWP domain (light grey hexagon) binding to Lys<sup>36</sup> (black triangle) may anchor the catalytic domain of the enzyme (dark grey oval) to methylate DNA preferentially in linker region (12).

situation. Demethylated nucleosomes were methylated with Dnmt3a2, a shorter isoform of Dnmt3a lacking the N-terminal variable region (23), the Dnmt3a2 D329A mutant (Dnmt3a2-D329A), and the catalytic domain of Dnmt3a (Dnmt3a-CD) using radioactively labeled *S*-adenosyl-*L*-methionine as cofactor. The amounts of Dnmt3a2, Dnmt3a2-D329A, and Dnmt3a-CD enzymes used in the nucleosome methylation experiments were normalized to have similar activity on oligonucleotide substrates (supplemental Fig. S6). Using the native chromatin as substrate, we observed that Dnmt3a2 is ~2.5-fold more active than Dnmt3a-CD (Fig. 4 and supplemental Fig. S6). The Dnmt3a2 enzyme containing the mutated D329A-PWWP domain, which fails to interact with the H3K36me3 mark, only showed a basal level of activity similar to Dnmt3a-CD (Fig. 4).

To confirm that the increase in activity of the wild type Dnmt3a2 is due to the PWWP-H3K36me3 interaction, we performed the chromatin methylation experiments in the presence of an anti-H3K36me3 antibody. Although the addition of the anti-H3K36me3 antibody had no effect on the activity of Dnmt3a2-D329A or Dnmt3a-CD, it decreased the activity of Dnmt3a2 to a basal level, similar to the level of activity observed with the Dnmt3a2-D329A or Dnmt3a-CD (Fig. 4). Additional unrelated antibodies used as controls had no effect or a very weak effect on the catalytic activity of Dnmt3a2 (supplemental Fig. S6). Taken together, these results demonstrate that the PWWP domain-H3K36me3 interaction increases the activity of Dnmt3a2 for methylation of native nucleosomal DNA.

## DISCUSSION

We show here that the Dnmt3a PWWP domain specifically interacts with H3 tails containing the K36me3 mark. Peptide binding is weak but in the range of binding constants seen with other individual domains interacting with histone peptides. In the cell binding may be further supported by an interaction of the PWWP domain with DNA, because the Lys<sup>36</sup> position of the H3 tail is very close to the DNA, such that the PWWP

domain could interact with both epitopes simultaneously. Interestingly, the DNA binding and peptide binding properties of the PWWP domains from Dnmt3a and 3b seem to be inverse; we detected peptide binding of the Dnmt3a-PWWP domain but not for Dnmt3b (data not shown), whereas DNA binding of the Dnmt3b-PWWP domain has been reported to be stronger than that of Dnmt3a (17).

We show that the H3K36me3-PWWP interaction mediates chromatin binding, because the heterochromatic localization of Dnmt3a was disturbed in the D329A mutant, which no longer binds to H3K36me3. Targeting of Dnmt3a by an H3K36me3-PWWP interaction is in agreement with previous results showing that the Dnmt3a PWWP domain is necessary for heterochromatic localization of Dnmt3a (17, 18). In contrast, the ADD domain of Dnmt3a, which binds to H3 tails unmethylated at Lys<sup>4</sup> (11, 12), does not have a role in heterochromatic localization (17, 18).

We show that the PWWP-H3K36me3 interaction stimulates the methylation activity of Dnmt3a on chromatin-bound DNA isolated from human cells, because the catalytic activity of the D329A mutant dropped to levels similar to the isolated catalytic domain. This result suggests an important role for the PWWP domain in guiding DNA methylation to chromatin carrying the H3K36me3 mark, which is in agreement with the observation that the PWWP domain is required for the activity of Dnmt3a in the cell (17, 39). This model is further supported by the observation that the genome-wide distribution of DNA methylation is very similar to that of K36me3 methylation. Many studies have found that H3K36me3 accumulates in euchromatin in the body of active genes, and its distribution is anticorrelated with H3K4me3 (37, 40–43). Similarly, active genes showed high DNA methylation in the gene bodies, whereas inactive genes did not (44, 45). The strong correlation of DNA methylation, absence of H3K4me3, and presence of H3K36me3 were also observed in two recent genome-wide studies (14, 16). Additionally, a correlation of H3K36me3 and DNA methylation was observed at exon-intron boundaries where exons were

shown to have increased levels of both H3K36me3 (46) and DNA methylation (16).

Like DNA methylation, H3K36me3 is a silencing mark that recruits histone deacetylation and represses intragenic transcriptional initiation (47). Consequently, the Smyd2 histone methyltransferase (which generates K36me2) has a repressive effect on reporter gene expression (48). In *Schizosaccharomyces pombe* the Set2 histone methyltransferase methylates histone H3 lysine 36 at repeat loci during the S phase (49). Consequently, heterochromatic H3K36me3 transiently peaks in the S phase and has an important role in the re-establishment of heterochromatin after DNA replication that acts in a pathway parallel to Clr4. It has been proposed that in mammals Dnmt3a contributes to the maintenance of the DNA methylation at highly methylated regions (like repeats) (50). H3K36 methylation is an evolutionarily conserved mark, suggesting that a transient methylation of H3K36 might occur in mammalian genomes as well and that transient H3K36 methylation might be one of the modifications that recruit Dnmt3a to heterochromatin.

Dnmt3a has been shown to bind tightly to chromatin (9). When combined with the previous findings, our data indicate that there are several mechanisms that contribute to the interaction of Dnmt3a with chromatin; Dnmt3a is engaged in two interactions with the H3 tail, mediated by its ADD domain, which binds to the end of the H3 tail unmodified at Lys<sup>4</sup> (11, 12), and by its PWWP domain, which interacts with H3K36me3 located at the more basal region of the H3 tail (Fig. 4B). The anticorrelation of the H3K4me3 and H3K36me3 marks suggests that in the cell most H3 tails will either bind to both ADD and PWWP or neither of them, such that the targeting effects of both domains are synergistic. In addition, the catalytic domain of Dnmt3a binds to DNA and polymerizes on the DNA (22, 51), and it interacts with many other proteins including Dnmt3L and Dnmt3b (52), which provide additional contact points to chromatin (10, 53), suggesting that a complex network of interactions is involved in targeting DNA methylation.

*Acknowledgment*—Many thanks are due to Dr. X. Cheng for insightful comments and valuable discussions.

## REFERENCES

- Goll, M. G., and Bestor, T. H. (2005) *Annu. Rev. Biochem.* **74**, 481–514
- Klose, R. J., and Bird, A. P. (2006) *Trends Biochem. Sci.* **31**, 89–97
- Reik, W. (2007) *Nature* **447**, 425–432
- Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999) *Cell* **99**, 247–257
- Bourc'his, D., Xu, G. L., Lin, C. S., Bollman, B., and Bestor, T. H. (2001) *Science* **294**, 2536–2539
- Hata, K., Okano, M., Lei, H., and Li, E. (2002) *Development* **129**, 1983–1993
- Hermann, A., Gowher, H., and Jeltsch, A. (2004) *Cell Mol. Life Sci.* **61**, 2571–2587
- Cheng, X., and Blumenthal, R. M. (2008) *Structure* **16**, 341–350
- Jeong, S., Liang, G., Sharma, S., Lin, J. C., Choi, S. H., Han, H., Yoo, C. B., Egger, G., Yang, A. S., and Jones, P. A. (2009) *Mol. Cell Biol.* **29**, 5366–5376
- Ooi, S. K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S. P., Allis, C. D., Cheng, X., and Bestor, T. H. (2007) *Nature* **448**, 714–717
- Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., and Shirakawa, M. (2009) *EMBO Rep.* **10**, 1235–1241
- Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., Fischle, W., and Jeltsch, A. (2010) *Nucleic Acids Res.*, in press
- Weber, M., Hellmann, I., Stadler, M. B., Ramos, L., Pääbo, S., Rebhan, M., and Schübeler, D. (2007) *Nat. Genet.* **39**, 457–466
- Meissner, A., Mikkelsen, T. S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B. E., Nusbaum, C., Jaffe, D. B., Gnirke, A., Jaenisch, R., and Lander, E. S. (2008) *Nature* **454**, 766–770
- Zhang, Y., Rohde, C., Tierling, S., Jurkowski, T. P., Bock, C., Santacruz, D., Ragozin, S., Reinhardt, R., Groth, M., Walter, J., and Jeltsch, A. (2009) *PLoS Genet* **5**, e1000438
- Hodges, E., Smith, A. D., Kendall, J., Xuan, Z., Ravi, K., Rooks, M., Zhang, M. Q., Ye, K., Bhattacharjee, A., Brizuela, L., McCombie, W. R., Wigler, M., Hannon, G. J., and Hicks, J. B. (2009) *Genome Res.* **19**, 1593–1605
- Chen, T., Tsujimoto, N., and Li, E. (2004) *Mol. Cell Biol.* **24**, 9048–9058
- Ge, Y. Z., Pu, M. T., Gowher, H., Wu, H. P., Ding, J. P., Jeltsch, A., and Xu, G. L. (2004) *J. Biol. Chem.* **279**, 25447–25454
- Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D., and Patel, D. J. (2007) *Nat. Struct. Mol. Biol.* **14**, 1025–1040
- Shirohzu, H., Kubota, T., Kumazawa, A., Sado, T., Chijiwa, T., Inagaki, K., Suetake, I., Tajima, S., Wakui, K., Miki, Y., Hayashi, M., Fukushima, Y., and Sasaki, H. (2002) *Am. J. Med. Genet.* **112**, 31–37
- Rathert, P., Dhayalan, A., Murakami, M., Zhang, X., Tamas, R., Jurkowska, R., Komatsu, Y., Shinkai, Y., Cheng, X., and Jeltsch, A. (2008) *Nat. Chem. Biol.* **4**, 344–346
- Jurkowska, R. Z., Anspach, N., Urbanke, C., Jia, D., Reinhardt, R., Nellen, W., Cheng, X., and Jeltsch, A. (2008) *Nucleic Acids Res.* **36**, 6656–6663
- Chen, T., Ueda, Y., Xie, S., and Li, E. (2002) *J. Biol. Chem.* **277**, 38746–38754
- Gowher, H., and Jeltsch, A. (2002) *J. Biol. Chem.* **277**, 20409–20414
- Gowher, H., and Jeltsch, A. (2001) *J. Mol. Biol.* **309**, 1201–1208
- Jeltsch, A., and Lanio, T. (2002) *Methods Mol. Biol.* **182**, 85–94
- Frank, R. (2002) *J. Immunol. Methods* **267**, 13–26
- Shechter, D., Dormann, H. L., Allis, C. D., and Hake, S. B. (2007) *Nat. Protoc.* **2**, 1445–1457
- Brand, M., Rampalli, S., Chaturvedi, C. P., and Dilworth, F. J. (2008) *Nat. Protoc.* **3**, 398–409
- Roth, M., and Jeltsch, A. (2000) *Biol. Chem.* **381**, 269–272
- Qiu, C., Sawada, K., Zhang, X., and Cheng, X. (2002) *Nat. Struct. Biol.* **9**, 217–224
- Xu, C., Cui, G., Botuyan, M. V., and Mer, G. (2008) *Structure* **16**, 1740–1750
- Lan, F., Collins, R. E., De Cegli, R., Alpatov, R., Horton, J. R., Shi, X., Gozani, O., Cheng, X., and Shi, Y. (2007) *Nature* **448**, 718–722
- Kalakonda, N., Fischle, W., Bocconi, P., Gurvich, N., Hoya-Arias, R., Zhao, X., Miyata, Y., Macgrogan, D., Zhang, J., Sims, J. K., Rice, J. C., and Nimer, S. D. (2008) *Oncogene* **27**, 4293–4304
- Umehara, T., Nakamura, Y., Jang, M. K., Nakano, K., Tanaka, A., Ozato, K., Padmanabhan, B., and Yokoyama, S. (2010) *J. Biol. Chem.* **285**, 7610–7618
- Vollmuth, F., Blankenfeldt, W., and Geyer, M. (2009) *J. Biol. Chem.* **284**, 36547–36556
- Vakoc, C. R., Sachdeva, M. M., Wang, H., and Blobel, G. A. (2006) *Mol. Cell Biol.* **26**, 9185–9195
- Rohde, C., Zhang, Y., Stamerjohanns, H., Hecher, K., Reinhardt, R., and Jeltsch, A. (2009) *BioTechniques* **47**, 781–783
- Shikauchi, Y., Saura, A., Kubo, T., Niwa, Y., Yamamoto, J., Murase, Y., and Yoshikawa, H. (2009) *Mol. Cell Biol.* **29**, 1944–1958
- Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007) *Cell* **129**, 823–837
- Larschan, E., Alekseyenko, A. A., Gortchakov, A. A., Peng, S., Li, B., Yang, P., Workman, J. L., Park, P. J., and Kuroda, M. I. (2007) *Mol. Cell* **28**, 121–133
- Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R., and Young, R. A. (2007) *Cell* **130**, 77–88
- Edmunds, J. W., Mahadevan, L. C., and Clayton, A. L. (2008) *EMBO J.* **27**, 406–420

## ***Dnmt3a PWWP Interacts with H3K36me3***

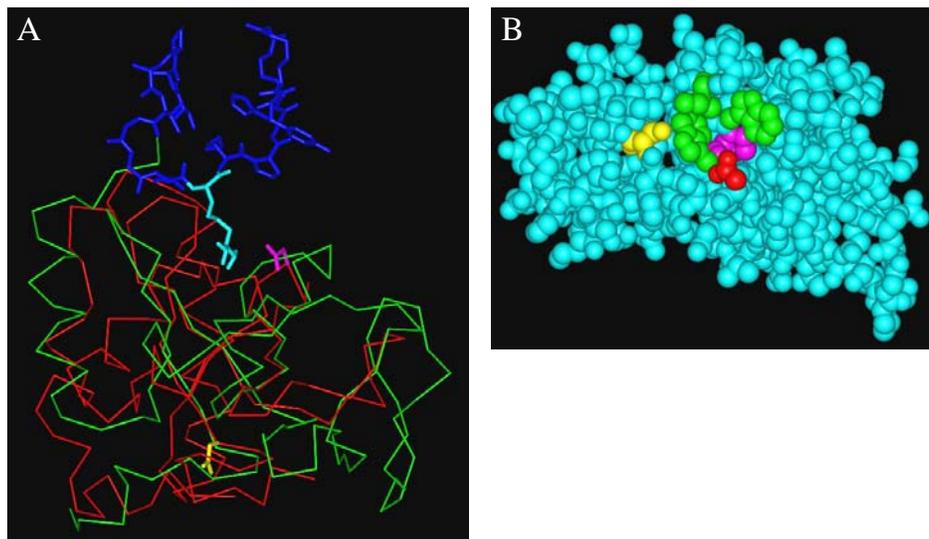
44. Hellman, A., and Chess, A. (2007) *Science* **315**, 1141–1143
45. Ball, M. P., Li, J. B., Gao, Y., Lee, J. H., LeProust, E. M., Park, I. H., Xie, B., Daley, G. Q., and Church, G. M. (2009) *Nat. Biotechnol.* **27**, 361–368
46. Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X. S., and Ahinger, J. (2009) *Nat. Genet.* **41**, 376–381
47. Lee, J. S., and Shilatifard, A. (2007) *Mutat Res.* **618**, 130–134
48. Brown, M. A., Sims, R. J., 3rd, Gottlieb, P. D., and Tucker, P. W. (2006) *Mol. Cancer* **5**, 26
49. Chen, E. S., Zhang, K., Nicolas, E., Cam, H. P., Zofall, M., and Grewal, S. I. (2008) *Nature* **451**, 734–737
50. Jones, P. A., and Liang, G. (2009) *Nat. Rev. Genet.* **10**, 805–811
51. Jia, D., Jurkowska, R. Z., Zhang, X., Jeltsch, A., and Cheng, X. (2007) *Nature* **449**, 248–251
52. 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., and Xu, G. L. (2007) *Mol. Cell. Biol.* **27**, 8748–8759
53. Gopalakrishnan, S., Sullivan, B. A., Trazzi, S., Della Valle, G., and Robertson, K. D. (2009) *Hum. Mol. Genet.* **18**, 3178–3193

# THE DNMT3A PWWP DOMAIN READS HISTONE 3 LYSINE 36 TRIMETHYLATION AND GUIDES DNA METHYLATION

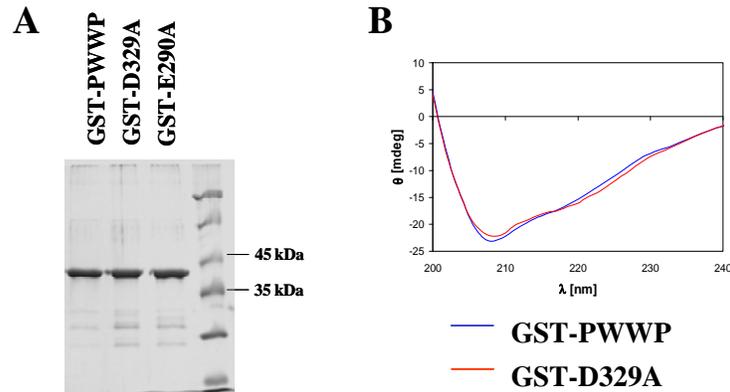
Arunkumar Dhayalan, Arumugam Rajavelu, Philipp Rathert, Raluca Tamas, Renata Z.  
Jurkowska, Sergey Ragozin & Albert Jeltsch

## Supplemental Figures

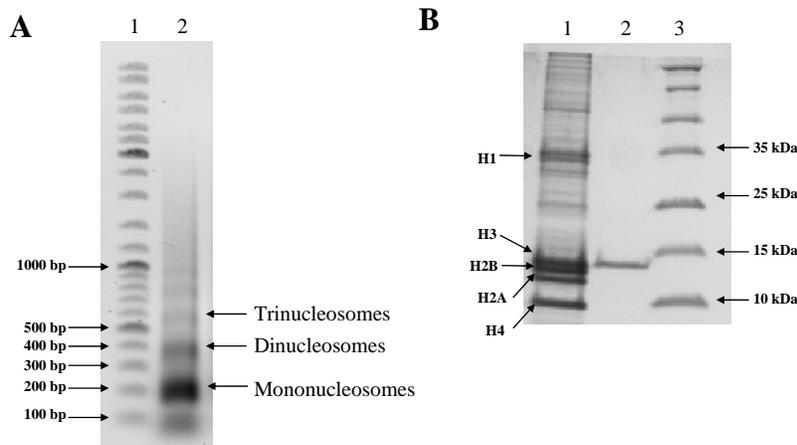
**Supplemental Figure 1:** A) Structure of murine Dnmt3b PWWP domain (red) (ref. 29) superimposed with *Saccharomyces cerevisiae* Eaf3 (green) in complex with H3 peptide (blue) containing an H3K36me2 analog (cyan) (ref. 30). Dnmt3b D273 (corresponding to D329 in the murine Dnmt3a PWWP domain) is colored purple, Dnmt3b D234 (corresponding to E290 in the murine Dnmt3a PWWP domain) is colored yellow. The superposition shows that D329, which is essential for the specific interaction of PWWP with H3K36me3, closely approaches the methylated lysine residue. In contrast E290, exchange of which did not have an effect on H3K36me3 interaction, is located on the opposite side of the protein. B) D273 surface pocket in the structure of the murine Dnmt3b PWWP domain. D273 (colored red) is next to a hydrophobic pocket composed by two aromatic residues (W270 and W246) and one isoleucine (I240) all colored green and P247 (colored pink), which is part of the conserved PWWP motif. Ser277, which corresponds to the S282P ICF mutation in human Dnmt3b, is colored yellow. The side chains of K214 and K275 which partially hide but do not occupy this pocket were omitted for clarity.



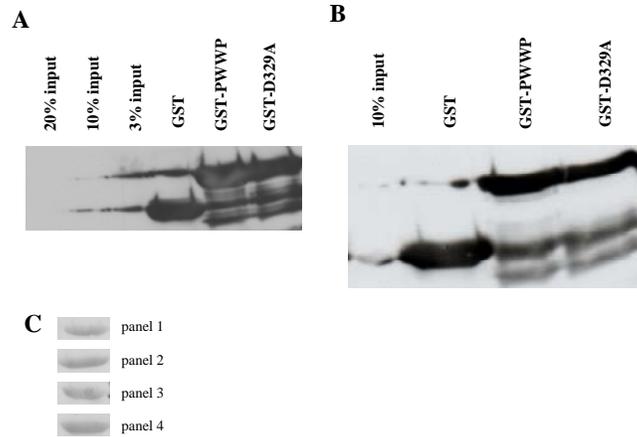
**Supplemental Figure 2:** Purification of GST-tagged PWWP domain variants. A) Coomassie stained gel of purified GST-tagged PWWP domain and its E290A and D329A variants. B) Circular dichroism spectrum of wildtype GST-PWWP and its D329A variant showing identical secondary structure composition.



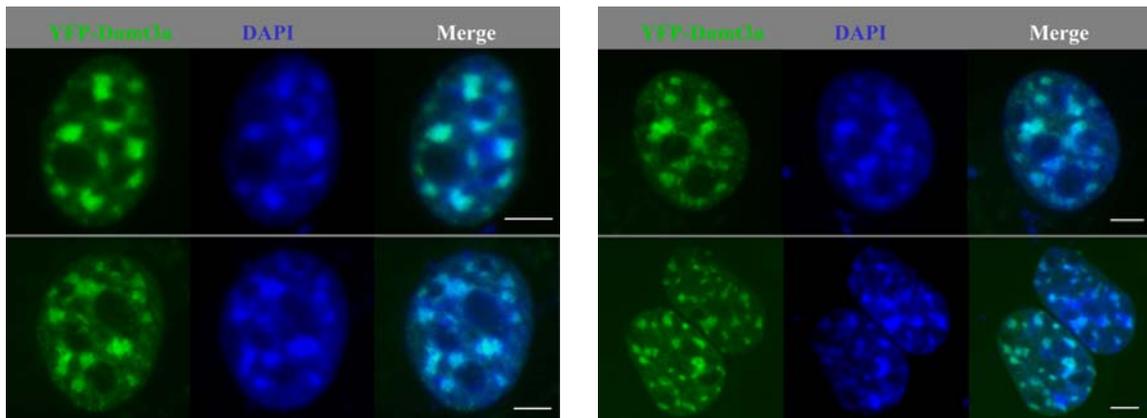
**Supplemental Figure 3:** Native nucleosome and histone preparation from HEK293 cells. A) Native nucleosomes were run on 1 % agarose gel and stained with ethidium bromide. The gel shows the presence of mononucleosomes, dinucleosomes, trinucleosomes and higher molecular weight chromatin with mononucleosomes representing the major part of the preparation (about 90%). Lane 1: DNA molecular weight standard, lane 2: native nucleosomes. B) Native histones run on SDS polyacrylamide gel stained with Coomassie blue. Lane 1: native histones, lane 2: recombinant histone H3 purchased from NEB, lane 3: protein molecular weight standard.



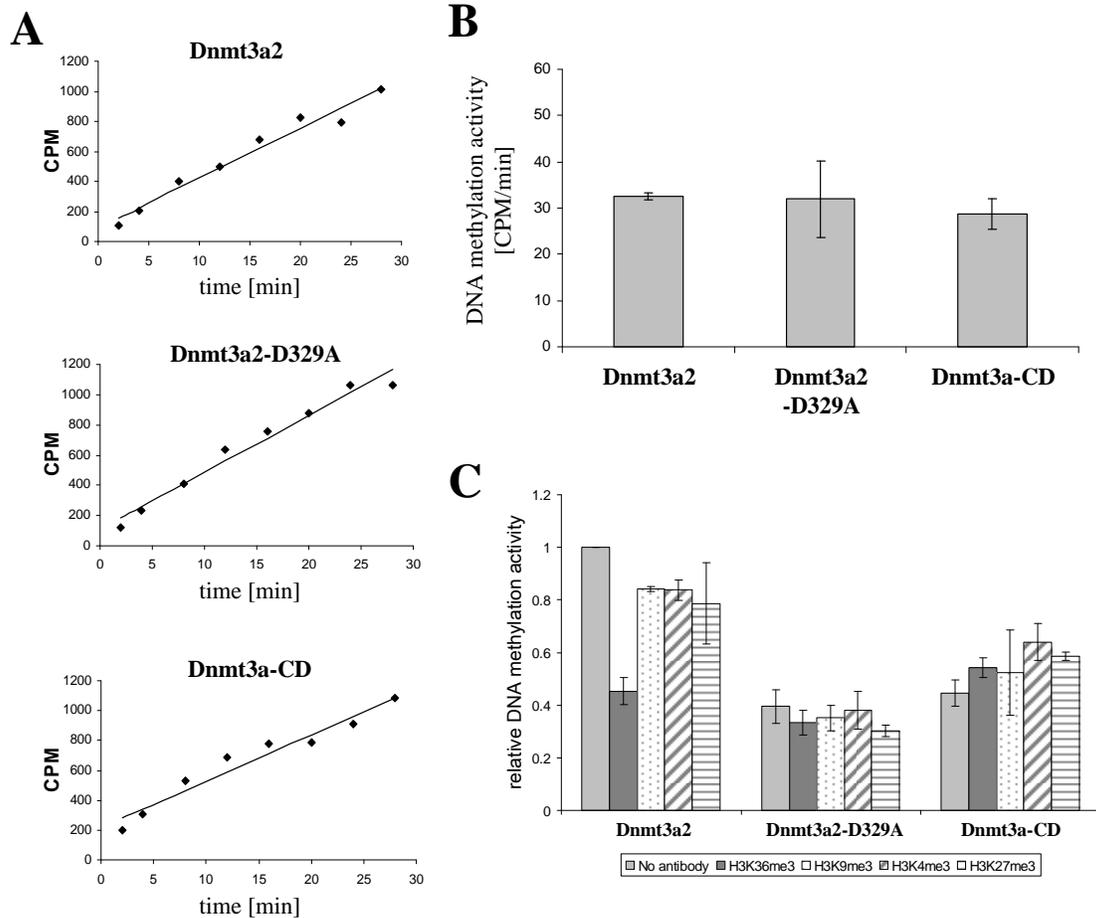
**Supplemental Figure 4:** Loading controls for Figure 2. A) and B) show a different parts of the same blots as shown in Fig. 2A and 2C, respectively, after probing with anti-GST antibody. The bands correspond to GST and the GST-PWWP domain proteins. The staining shows that similar amounts of GST and GST-PWWP domains were loaded for the pull-down. C) Ponceau S staining of the blot shown in Fig. 2B prior to the antibody staining. The bands correspond to the GST-PWWP domain ensuring equal loading of the lanes.



**Supplemental Figure 5:** Exemplary cell images documenting that Dnmt3a foci coincide with DAPI dense heterochromatic foci in NIH3T3 cells as already observed previously (refs. 16 and 17). The bars correspond to 5  $\mu$ M.



**Supplemental Figure 6:** Methylation activity normalization of Dnmt3a2, Dnmt3a2-D329A and Dnmt3a-CD with oligonucleotide substrates. A) Time courses of methylation. B) Averages of methylation activities and standard deviations derived from at least three independent experiments. C) Repetition of the experiments shown in Fig. 4, including additional unrelated antibodies (anti-H3K4me3, anti-H3K9me3 and anti-H3-K27me3) that were used as controls. Error bars indicate standard deviations derived from at least three independent experiments.



# Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail

Yingying Zhang<sup>1</sup>, Renata Jurkowska<sup>1</sup>, Szabolcs Soeroes<sup>2</sup>, Arumugam Rajavelu<sup>1</sup>, Arunkumar Dhayalan<sup>1</sup>, Ina Bock<sup>1</sup>, Philipp Rathert<sup>3</sup>, Ole Brandt<sup>4</sup>, Richard Reinhardt<sup>5</sup>, Wolfgang Fischle<sup>2</sup> and Albert Jeltsch<sup>1,\*</sup>

<sup>1</sup>Biochemistry Laboratory, School of Engineering and Science, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, <sup>2</sup>Laboratory of Chromatin Biochemistry, Max-Planck Institute for Biophysical Chemistry, D-37077 Göttingen, Germany, <sup>3</sup>Active Motif Europe, Av F.D. Roosevelt, 104, BE-1330 Rixensart, Belgium, <sup>4</sup>INTAVIS Bioanalytical Instruments AG, Nattermannallee 1, 50829 Köln and <sup>5</sup>Max Planck Institute for Molecular Genetics, Ihnestrasse 63-73, D-14195 Berlin-Dahlem, Germany

Received December 4, 2009; Revised February 11, 2010; Accepted February 18, 2010

## ABSTRACT

Using peptide arrays and binding to native histone proteins, we show that the ADD domain of Dnmt3a specifically interacts with the H3 histone 1–19 tail. Binding is disrupted by di- and trimethylation of K4, phosphorylation of T3, S10 or T11 and acetylation of K4. We did not observe binding to the H4 1–19 tail. The ADD domain of Dnmt3b shows the same binding specificity, suggesting that the distinct biological functions of both enzymes are not related to their ADD domains. To establish a functional role of the ADD domain binding to unmodified H3 tails, we analyzed the DNA methylation of *in vitro* reconstituted chromatin with Dnmt3a2, the Dnmt3a2/Dnmt3L complex, and the catalytic domain of Dnmt3a. All Dnmt3a complexes preferentially methylated linker DNA regions. Chromatin substrates with unmodified H3 tail or with H3K9me3 modification were methylated more efficiently by full-length Dnmt3a and full-length Dnmt3a/3L complexes than chromatin trimethylated at H3K4. In contrast, the catalytic domain of Dnmt3a was not affected by the H3K4me3 modification. These results demonstrate that the binding of the ADD domain to H3 tails unmethylated at K4 leads to the preferential methylation of DNA bound to chromatin with this modification state. Our *in vitro* results recapitulate DNA methylation patterns observed in genome-wide DNA methylation studies.

## INTRODUCTION

DNA methylation is a major form of epigenetic modification and plays essential roles in gene expression regulation and chromatin structure remodeling (1–3). The methylation state of DNA is closely connected to other epigenetic signals including histone modifications, such as methylation or acetylation, which are known to activate or silence gene expression (4). The methylation of CpG dinucleotides (CpG) in mammalian cells is catalyzed by DNA methyltransferases (Dnmts), comprising Dnmt3a and 3b, which establish DNA methylation patterns during embryonic development and Dnmt1, which maintains the methylation pattern after DNA replication (1,2,5). Dnmt3a and 3b contain large N-terminal parts including a PWWP domain and a PHD-like ADD domain, which interact with other proteins, and a C-terminal domain harboring the catalytic center. The isolated catalytic domains of Dnmt3a and 3b are enzymatically active (6). Another member of the Dnmt3 family, Dnmt3L (Dnmt3-like), is homologous to the Dnmt3 enzymes, but lacks catalytic activity. It acts as a regulatory factor and can stimulate the catalytic activity of Dnmt3a and 3b (7–10).

The ADD domain of Dnmt3L was shown to interact specifically with histone H3 tails that are unmethylated at lysine 4 (11). The Dnmt3a/3L complex forms a heterotetramer (12), suggesting that the interaction of the Dnmt3L ADD domain with the H3 tail could direct DNA methylation by Dnmt3a (11). The ADD domains of Dnmt3a and 3b share considerable homology with Dnmt3L and recently binding of the Dnmt3a ADD domain to H3 tails unmethylated at K4 has been shown and the structure of this complex solved (13). In addition, an

\*To whom correspondence should be addressed. Tel: +49 421 200 3247; Fax: +49 421 200 3249; Email: a.jeltsch@jacobs-university.de

interaction of the Dnmt3a ADD with H4R3me2s peptides has been reported as well (14). Here, we studied the interaction of the Dnmt3a and 3b ADD domains with modified histone tails by applying a hypothesis free peptide array binding approach.

Independent experimental evidence suggesting an influence of histone tail modification on DNA methylation has been provided through several epigenomic studies. Genome-wide DNA methylation and histone modification studies revealed a strong anti-correlation of DNA methylation and histone H3 lysine 4 trimethylation (H3K4me3) (15–18), and a correlation of H3K9me3 with DNA methylation (16). A functional connection of these two silencing marks (DNA methylation and H3K9 methylation) has been observed before in *Neurospora crassa* (19,20), plants (21) and mammalian cells (22), where disruption of the H3K9me3 signal led to a loss of DNA methylation. As described above, biochemical and structural data suggest a direct role of the ADD domain of Dnmt3L in the targeting of Dnmt3a to chromatin unmethylated at H3K4 (11). However, so far experimental evidence for preferential methylation of DNA bound to chromatin, which carries a particular modification pattern, has not been provided. In order to establish the molecular mechanism of DNA methylation guidance by histone modification states, we set up a complete *in vitro* system that allowed us to study the influence of chromatin modifications on the activity of purified Dnmt3a or Dnmt3a/3L. To this end, histones were generated by native peptide ligation to contain specifically H3K4 and H3K9 methylation. The modified histones were assembled into octamers, bound to DNA and the reconstituted oligonucleosomes were used as substrates for methylation with Dnmt3a and Dnmt3a/3L.

## MATERIALS AND METHODS

For details of ‘Materials and Methods’ section; see Supplementary Data.

### Recombinant chromatin preparation

Expression and purification of *Xenopus laevis* histones were performed as described (23). H3K4me3 and H3K9me3 were generated by native protein ligation. Ligation of the activated H3 peptide to the truncated H3 histone and purification of the ligation product was performed as described (24). Assembly of histone octamers containing H3unmod, H3K4me3 and H3K9me3, as well as reconstitution of recombinant oligonucleosomes was performed by salt dialysis as described, using the 12× 200 × 601 template (23,25).

### Purification of Dnmt enzymes

The Dnmt3a and Dnmt3b ADD domains (amino acids 472–610 of murine Dnmt3a and 422–560 of murine Dnmt3b1) were cloned as GST fusion proteins and purified by standard procedures. The expression and purification of Dnmt3a2, Dnmt3L and Dnmt3a-C were carried out as described (10,12). The purity of the proteins was determined on 12% sodium dodecyl sulfate

polyacrylamide gel electrophoresis (SDS-PAGE) gel to be better than 90% (Supplementary Figure S4).

### Binding of protein domains to peptides arrays

CelluSpots arrays were provided by Intavis AG (Köln, Germany). The array was blocked, incubated with purified GST-tagged ADD domains of Dnmt3a or 3b (1 μM) and the binding detected by anti-GST antibody. We confirmed the correct synthesis of the spots containing T3P, K4me2, K4me3, S10P, T11P and H4R3me2s by incubation with modification specific antibodies (Supplementary Figure S5), methylation of the array with different histone lysine methyltransferases and by high-performance liquid chromatography (HPLC) and mass spectrometric analysis of peptides synthesized in parallel and cleaved off from the matrix for characterization (data not shown).

### Bisulfite conversion, subcloning and sequencing

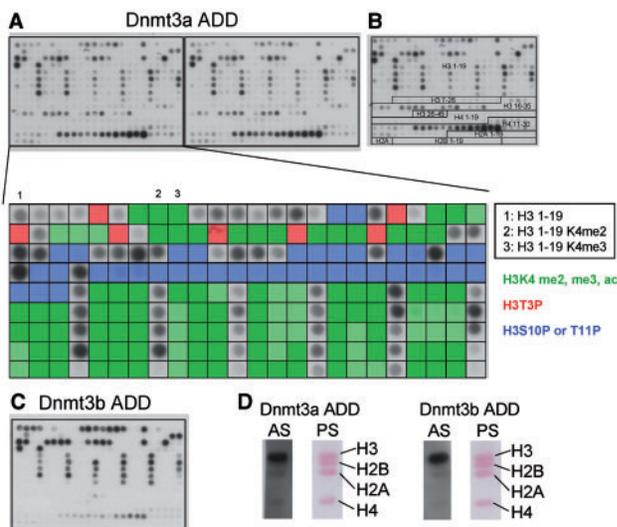
The modified chromatin was methylated with Dnmt3a complexes and the DNA was purified and treated with bisulfite as described (26).

## RESULTS

### Binding of the Dnmt3a ADD domain to modified histone tail peptides

We used an array comprising 384 peptide spots prepared by the CelluSpots method (27) to study the interaction of the purified Dnmt3a and 3b ADD domains with modified histone tails (Figure 1). The array contains peptides from eight different regions of the N-terminal tails of histones, namely H3 1–19, 7–26, 16–35 and 26–45, H4 1–19 and 11–30, H2A 1–19 and H2B 1–19 (Figure 1 and Supplementary Data S1), featuring 59 post-translational modifications in many different combinations. Each array was presenting the same peptides in duplicate for quality control. Binding to the internal duplicates was highly reproducible (Supplementary Figure S1). In addition, the results were confirmed by binding of the ADD domain to peptide arrays originating from an independent peptide synthesis (data not shown).

Peptide binding of the ADD domains from Dnmt3a and 3b was very similar (Figure 1). In general, we observed highly specific binding of the Dnmt3a and 3b ADD domains to differentially modified H3 tails (Figure 1A). Binding to H2A and H2B tails was not observed. Binding at H4 tails was only observed at H4K16ac, if they contained additional modifications in different combinations. Binding of the ADD domain to H4R3me2s peptides that has been reported recently (14) could not be detected (Supplementary Figure S2). We tested binding of the ADD domain to native histones isolated from human cells by western blotting and detected strong binding to the H3 tail, but no binding to H4, H2A or H2B tails (Figure 1D). This result suggests that the special combinations of modifications on H4 leading to ADD binding on the peptide array are not frequent in chromatin isolated from human cells. Based on these results, we



**Figure 1.** Binding of the Dnmt3a and Dnmt3b ADD domains to peptide arrays and native histones. (A) Binding of Dnmt3a ADD domain to peptide arrays comprising 384 different peptides. The enlargement shows the binding to the H3 1–19 peptides. Peptides containing H3K4me2 or me3, H3T3P, H3S10P or H3T11P are shaded green, red and blue, respectively. The positions of the unmodified H3 1–19 as well as the peptides di- and trimethylated at K4 are annotated. (B) Design of the CelluSpots histone tail peptide arrays. For a detailed annotation of all spots cf. Supplementary Data S1. (C) Binding of Dnmt3b ADD domain to peptide arrays comprising 384 different peptides. (D) Binding of ADD domains to native histones isolated from human cells. Histones were separated by polyacrylamide gel electrophoresis and blotted to Nitrocellulose membrane. The membrane was stained with Ponceau S (PS). Then, membranes were incubated with GST tagged ADD domains, washed and ADD binding detected with anti-GST antibody staining (AS).

focused on the H3 tail interaction for the further functional studies.

Binding of H3 tails was disrupted by modification of H3K4 other than monomethylation. In addition, the phosphorylation of T3, S10 or T11 inhibited peptide binding. As shown clearly in Figure 1A, binding of the ADD domain to H3 1–19 peptides is only possible if none of these inhibiting modifications is present. Monomethylation of K4 reduced binding, but did not completely prevent the interaction. In addition, non-natural acetylation of the N-terminus completely disrupted binding, indicating that the free amino terminus is an important contact point as well (Supplementary Figure S3). These results agree with equilibrium peptide binding results for the Dnmt3a ADD domain reported by Otani *et al.* (13), who showed that H3K4me2, H3K4me3 and N-acetylation inhibit H3 tail binding, while trimethylation at K9 has no influence on binding. Also in their experiments, binding to H4 1–19 tails and H4 1–19 modified by R3me2s was very weak or undetectable. Other modifications were not tested in that study.

We conclude that similar to the ADD domain of Dnmt3L, the Dnmt3a and Dnmt3b ADD domains bind to the H3 tail. With Dnmt3L, mono-, di- and trimethylation of K4 was shown to inhibit binding, but the effect

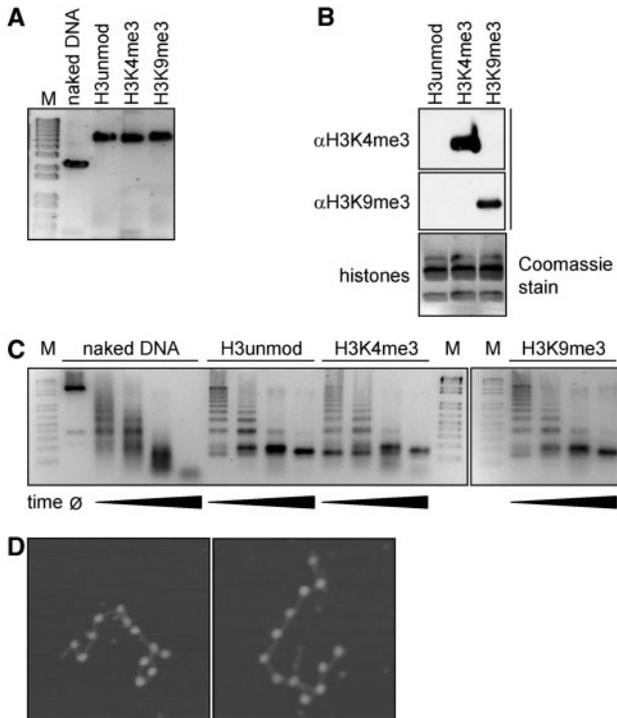
of monomethylation was weaker than that of di- or trimethylation (11). Since we have available active full-length Dnmt3a2 [which is an isoform of Dnmt3a (28) containing the ADD domain] and its catalytic domain (not containing the ADD) (6), we focused our functional studies on the targeting of DNA methylation activity of Dnmt3a2 by its ADD domain.

### Reconstitution of specifically modified oligonucleosomes

With the aim to study the interaction of DNA methylation and histone modifications in nucleosomes, we generated recombinant unmodified chromatin and chromatin containing H3K4me3 and H3K9me3 modifications. Wild-type *Xenopus laevis* histones were expressed and purified. Histone H3 proteins modified at H3K4me3 and H3K9me3 were generated by native protein ligation (24). The assembly of histone octamers containing unmodified H3, H3K4me3 and H3K9me3 and the reconstitution of recombinant oligonucleosomes was performed using the  $12 \times 200 \times 601$  template (12 tandem repeats of a 200 bp 601 DNA sequence). The nucleosome positioning sequence locates at the center of each monomeric 200 bp of the 601 repeat, and the 12 repeats were arranged on the array in a head to tail manner (25,29). The successful assembly of chromatin was verified by native gel electrophoresis (Figure 2A) and analytical ultracentrifugation (data not shown). Identity and purity of the histone proteins was confirmed by SDS-PAGE (Figure 2B) and by mass spectrometry. To further control the saturation level of the oligonucleosomal reconstitution, assembled material was analyzed after digestion with micrococcal nuclease (Figure 2C). In addition, samples were analyzed by scanning force microscopy after mild fixation with glutaraldehyde and deposition on mica support (Figure 2D) (30). At a 1.1 to 1 octamer:DNA ratio, all samples showed the MNase digestion pattern characteristic for nucleosome bound DNA;  $11 \pm 1$  nucleosomes on arrays were observed in SFM, confirming the successful reconstitution of recombinant chromatin.

### Methylation of DNA bound to oligonucleosomes

In the next step, the recombinant chromatin was methylated by full-length Dnmt3a2, the C-terminal catalytic domain of Dnmt3a, and the complex of full-length Dnmt3a2 and Dnmt3L. For methylation analysis, the DNA was isolated and treated with bisulfite. By using primers specific for the DNA sequence of the monomeric 601 repeat (Figure 3), we amplified the PCR products, which represent a mixture of the DNA repeat in all 12 nucleosomes, and analyzed the DNA methylation pattern by sequencing of individual clones. The analyzed region spanned the nucleosome and included 16 out of the 18 CpG sites of each 601 repeat. DNA methylation was investigated in both DNA strands. Control methylation reactions were performed with naked DNA from the  $12 \times 200 \times 601$  template, which was used for the chromatin reconstitution. The Dnmt3a2/3L complex methylated DNA with about 1.5-fold enhanced rate as compared to Dnmt3a2 alone, which is a similar level of stimulation of



**Figure 2.** Analysis of reconstituted oligonucleosomes. (A) Native agarose gel (0.5%, 0.2 $\times$  TBE) of free DNA or oligonucleosomes after reconstitution on 12  $\times$  200  $\times$  601 sequence containing uniformly the indicated H3 species. Gel was stained post-running with ethidium bromide (M, molecular size markers). (B) Reconstituted oligonucleosomes containing the indicated H3 species were run on SDS PAGE gel and analyzed by western blotting using the indicated antibodies. A separately run gel was stained with Coomassie blue. (C) DNA used for reconstitution or assembled oligonucleosomes containing different H3 species were digested with MNase for varying times. Samples were run on agarose gels (1.5%, 1 $\times$  TBE) and stained with ethidium bromide (M, molecular size markers). (D) Scanning force microscopy images of reconstituted oligonucleosomes. At a 1:1:1 octamer:DNA ratio the number of visible nucleosomes on one DNA molecule typically ranged from 10 to 12.

Dnmt3a by Dnmt3L on long DNA substrates as reported in the literature (31). We observed in all reactions that nucleosome bound DNA was methylated with a strong preference at the first three CpG sites, which are located in the linker DNA region immediately upstream of the nucleosomal bound DNA (Figure 3). The activities of all enzymes were reduced on chromatin when compared to naked DNA, with Dnmt3a-C showing stronger reduction than full-length enzymes (Figure 3). In addition, the naked DNA showed a totally different methylation pattern, indicating that the preference for methylation of the first three sites was not related to the DNA sequence.

#### Methylation of DNA bound to specifically modified oligonucleosomes

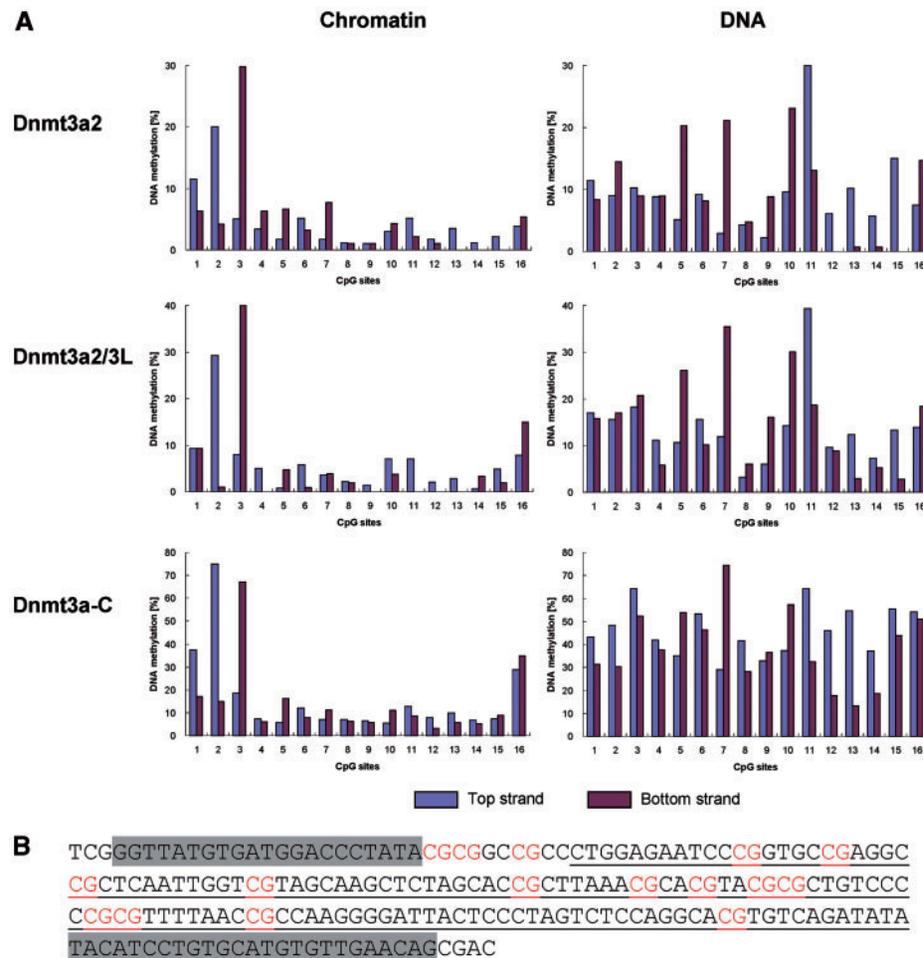
Next, we investigated the methylation of the oligonucleosomes carrying H3K4me3 or H3K9me3 modifications following the same approach (Figure 4). Since we observed the strongest methylation in the first three CpG sites, we restricted the analysis to these sites (Figure 5). Introduction of the H3K9me3 modification

did not significantly change the methylation levels—an observation that is in agreement with previous results obtained using mononucleosomes (32) and with the lack of any effect of H3K9me3 on ADD binding to H3 tails. In contrast, the DNA methylation level introduced into the H3K4me3 modified chromatin by full-length Dnmt3a2 or the complex of full-length Dnmt3a2 with Dnmt3L was strongly decreased when compared to the unmodified chromatin. To analyze whether this reduction of methylation of H3K4me3 modified chromatin was related to the ADD domain of Dnmt3a, we methylated the recombinant chromatin using only the catalytic domain of Dnmt3a (Dnmt3a-C). The results showed that without the ADD domain there was no significant methylation difference between the unmodified chromatin and the chromatin containing H3K4me3, suggesting that it is the K4 methylation specific H3 tail binding of the ADD domain that guides methylation of full-length Dnmt3a in this experiment. This observation also confirms that the different modified chromatin substrates used here are of comparable quality.

#### DISCUSSION

It is an open question how DNA methylation patterns are generated by DNA methyltransferases. For the catalytic domain of Dnmt3a it has been observed that it shows some flanking sequence preferences (33,34) and it favors the methylation of CpG sites in a distance equal to one helical turn of DNA (12,35). Both these properties were shown to influence the DNA methylation in cells, but they cannot explain the generation of a specific DNA methylation pattern during embryogenesis and germ cell development. In addition, being an epigenetic signal, the DNA methylation pattern cannot be determined only by the primary DNA sequence alone.

We have shown that the ADD domains of Dnmt3a and Dnmt3b interact with the H3-tail unmethylated at K4 similar as the ADD domain of Dnmt3L (11). Thereby, these domains assist the binding of full-length Dnmt3a (or the Dnmt3a/Dnmt3L complex) to chromatin. Using designer nucleosomes containing defined modifications, we investigated whether H3K4me3 or H3K9me3 modifications influence the methylation of the chromatin bound DNA by Dnmt3a2 and its catalytic domain (which lacks the N-terminal part including the ADD domain). Nucleosome bound DNA was methylated with a strong preference for the linker DNA region. This finding confirms a previous results obtained with full-length Dnmt3a and Dnmt3a-C using mononucleosomes reconstituted with a different DNA template and a different method to study methylation of chromatin bound DNA (32). We conclude that the core nucleosome is protected against DNA methylation by Dnmt3a, but the linker regions of the DNA are accessible. The finding that histone bound DNA is protected from methylation is consistent with results obtained with chromatin isolated from mammalian cells (36) also showing that addition of H1 further inhibits DNA methylation. We observed that the catalytic domain of Dnmt3a showed a higher activity than



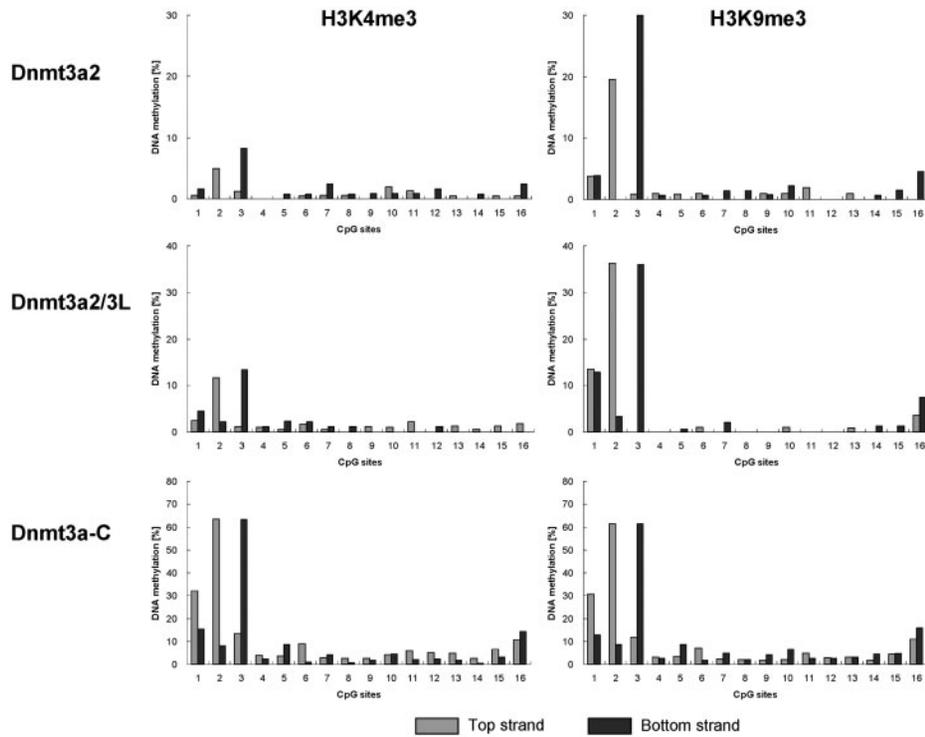
**Figure 3.** DNA methylation of unmodified chromatin and naked DNA by full-length Dnmt3a2, full-length Dnmt3a2/Dnmt3L complex and Dnmt3a-C. DNA methylation analysis was performed for both strands of DNA in chromatin substrates and DNA. Results shown are the average of three independent experiments. Altogether about 250 clones were sequenced for each strand and each data point. (A) Methylation levels of single CpG sites in the upper and lower strand of the DNA. (B) DNA sequence of the 601 monomeric unit (200 bp). The grey color shades the sequences of the forward and reverse primers. The red-color-labeled CpG sites were analyzed in this study. The nucleosome positioning sequence (147 bp) is underlined.

Dnmt3a2. This result was reproducible over many purifications and many years. It may suggest that the N-terminal part of Dnmt3a2 has a repressive regulatory function on the catalytic domain, as it is often seen in other regulated enzymes like proteases or kinases.

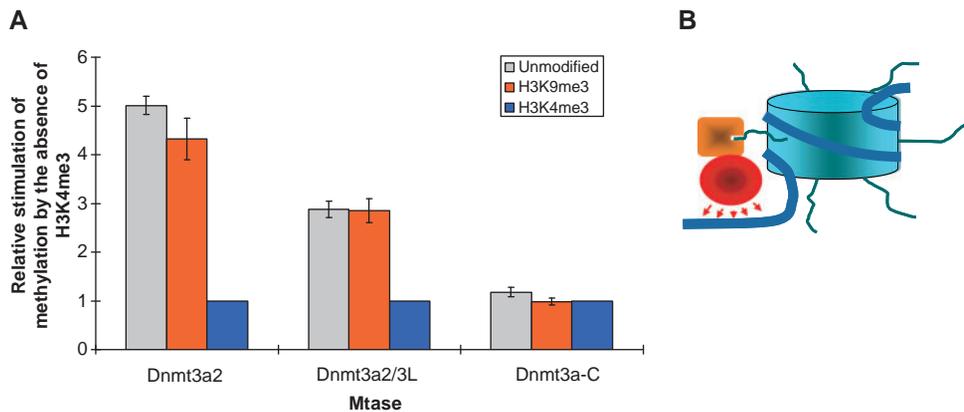
The preferential methylation of linker DNA is observed for Dnmt3a2 and its isolated catalytic domain, indicating that it is not mediated by the ADD domain. However, with Dnmt3a2 the DNA methylation levels of chromatin with unmodified H3 tail or with the H3K9me3 modification is four to five times higher than methylation of H3K4me3 containing chromatin. This difference was not seen with the isolated catalytic domain of Dnmt3a. This result indicates that the disruption of the ADD-H3 tail interaction by di- or tri-methylation of H3K4 interferes with the chromatin recruitment of Dnmt3a2 and methylation of DNA. With Dnmt3a2/3L, methylation of unmodified and H3K9me3 methylated chromatin was about three times higher than with H3K4me3 chromatin.

The reduced stimulation might be explained by the fact that in the Dnmt3a/3L tetramer the Dnmt3L ADD domains also interact with the H3 tails, but Dnmt3L does not contribute a catalytic domain. In contrast, in case of Dnmt3a2, each ADD-H3 tail interaction leads to the recruitment of one catalytic domain to the chromatin bound DNA.

Using peptide array containing several hundred peptides in different modification states, we identified additional modifications of the H3 tail that interfere with binding, namely acetylation of K4 and phosphorylation of T3, S10 or T11. The effects at T3 and K4 can be readily explained in the light of the Dnmt3a ADD structure (13): T3 is in 4.4 Å distance to Glu545, an interaction that would be disrupted by phosphorylation of the threonine. K4 is inserted into a very tight acidic pocket formed by Asp529 and Asp531 that does not allow for its acetylation. The effects in the C-terminal part of the H3 peptide cannot be easily



**Figure 4.** DNA methylation levels of single CpG sites in chromatin with H3K4me3 or H3K9me3 modifications by full-length Dnmt3a2, full-length Dnmt3a2/Dnmt3L and Dnmt3a-C. Results shown are the averages of three independent experiments. Altogether about 250 clones were sequenced for each strand and each data point.



**Figure 5.** Methylation of DNA on modified oligonucleosomes. (A) Methylation of the first three CpG sites by full-length Dnmt3a2, full-length Dnmt3a2/Dnmt3L, but not Dnmt3a-C is stimulated by the absence of H3K4me3 (data taken from Figures 3 and 4). For each comparison, the methylation levels were normalized to H3K4me3, which showed the lowest methylation. Results shown are the average of three independent experiments; error bars indicate the standard deviations. (B) Schematic picture showing how the H3 tail interaction of Dnmt3a's ADD domain (orange) anchors the catalytic domain (red) to methylate the linker DNA regions of nucleosomes.

interpreted, but in the structure the peptide was covalently bound to the N-terminus of the ADD domain, such that the positioning of the C-terminal part of the H3 peptide might not be fully reliable.

Our data provide the first evidence that preferential binding of a DNA methyltransferase to histone tails carrying specific post-translational modification patterns directly leads to the favored methylation of DNA bound to the modified chromatin (Figure 5B). Here, we

demonstrate that this effect can be achieved by the ADD domain in Dnmt3a, which binds to chromatin with the same specificity as the Dnmt3L ADD domain. However, other domains and factors may further contribute to the targeting of Dnmt3a. For example, the PWWP domain of Dnmt3a has been shown to be involved in heterochromatin targeting of the enzyme (37,38). Our unpublished data suggest that this domain is also directly engaged in an interaction with histone tails

(manuscript in preparation). Together with the DNA binding of the catalytic domains, these interactions may be responsible for the strong and direct interaction of Dnmt3a with chromatin (39). Another silencing mark that is associated with DNA methylation is H3K9me3 (19–22). Our results make a direct interaction of Dnmt3a with H3K9me3 (via the ADD domain or any other part of the enzyme) unlikely, suggesting that this correlation might be mediated by additional proteins, like HP1, which reads H3K9me3 and has been shown to directly interact with Dnmt3a (40).

It is interesting that the ADD domains of Dnmt3L (11), Dnmt3a (this work and Ref. 13) and Dnmt3b (this work) all basically have the same preference for binding to H3 tails unmodified at K4. Dnmt3a and 3b show clearly distinct biological functions with Dnmt3a being involved in the setting of parental imprints (41,42) and Dnmt3b in the methylation of pericentromeric repeats (41,43). Our results suggest that the ADD domains of both proteins are not responsible for this difference in function. The finding that the ADD domains from Dnmt3a and Dnmt3L bind to H3 tails with the same specificity suggests that the recruitment of Dnmt3a to chromatin with unmethylated H3 tail may not be the primary function of Dnmt3L.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## FUNDING

The DFG (JE 252/6 and JE 252/7); NIH grant DK08267; and by the Max Planck Society (to W.F. and R.R.). S. S. was recipient of a Boehringer Ingelheim predoctoral fellowship. Funding for open access charge: DFG.

*Conflict of interest statement.* None declared.

## REFERENCES

- Goll, M.G. and Bestor, T.H. (2005) Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.*, **74**, 481–514.
- Klose, R.J. and Bird, A.P. (2006) Genomic DNA methylation: the mark and its mediators. *Trends Biochem. Sci.*, **31**, 89–97.
- Reik, W. (2007) Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature*, **447**, 425–432.
- Kouzarides, T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.
- Hermann, A., Gowher, H. and Jeltsch, A. (2004) Biochemistry and biology of mammalian DNA methyltransferases. *Cell. Mol. Life Sci.*, **61**, 2571–2587.
- Gowher, H. and Jeltsch, A. (2002) Molecular enzymology of the catalytic domains of the Dnmt3a and Dnmt3b DNA methyltransferases. *J. Biol. Chem.*, **277**, 20409–20414.
- Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B. and Bestor, T.H. (2001) Dnmt3L and the establishment of maternal genomic imprints. *Science*, **294**, 2536–2539.
- Hata, K., Okano, M., Lei, H. and Li, E. (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development*, **129**, 1983–1993.
- Chedin, F., Lieber, M.R. and Hsieh, C.L. (2002) The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proc. Natl Acad. Sci. USA*, **99**, 16916–16921.
- Gowher, H., Liebert, K., Hermann, A., Xu, G. and Jeltsch, A. (2005) Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. *J. Biol. Chem.*, **280**, 13341–13348.
- Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.P., Allis, C.D. *et al.* (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature*, **448**, 714–717.
- Jia, D., Jurkowska, R.Z., Zhang, X., Jeltsch, A. and Cheng, X. (2007) Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature*, **449**, 248–251.
- Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M. and 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.
- Zhao, Q., Rank, G., Tan, Y.T., Li, H., Moritz, R.L., Simpson, R.J., Cerruti, L., Curtis, D.J., Patel, D.J., Allis, C.D. *et al.* (2009) PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing. *Nat. Struct. Mol. Biol.*, **16**, 304–311.
- Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Paabo, S., Rebhan, M. and Schubeler, D. (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.*, **39**, 457–466.
- Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B. *et al.* (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*, **454**, 766–770.
- Zhang, Y., Rohde, C., Tierling, S., Jurkowski, T.P., Bock, C., Santacruz, D., Ragozin, S., Reinhardt, R., Groth, M., Walter, J. *et al.* (2009) DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution. *PLoS Genet.*, **5**, e1000438.
- Hodges, E., Smith, A.D., Kendall, J., Xuan, Z., Ravi, K., Rooks, M., Zhang, M.Q., Ye, K., Bhattacharjee, A., Brizuela, L. *et al.* (2009) High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. *Genome Res.*, **19**, 1593–1605.
- Tamaru, H. and Selker, E.U. (2001) A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature*, **414**, 277–283.
- Tamaru, H., Zhang, X., McMillen, D., Singh, P.B., Nakayama, J., Grewal, S.I., Allis, C.D., Cheng, X. and Selker, E.U. (2003) Trimethylated lysine 9 of histone H3 is a mark for DNA methylation in *Neurospora crassa*. *Nature Genet.*, **34**, 75–79.
- Jackson, J.P., Lindroth, A.M., Cao, X. and Jacobsen, S.E. (2002) Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature*, **416**, 556–560.
- Lehnertz, B., Ueda, Y., Derijck, A.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T. and Peters, A.H. (2003) Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.*, **13**, 1192–1200.
- Luger, K., Rechsteiner, T.J., Flaus, A.J., Wayne, M.M. and Richmond, T.J. (1997) Characterization of nucleosome core particles containing histone proteins made in bacteria. *J. Mol. Biol.*, **272**, 301–311.
- Shogren-Knaak, M.A. and Peterson, C.L. (2004) Creating designer histones by native chemical ligation. *Methods Enzymol.*, **375**, 62–76.
- Huynh, V.A., Robinson, P.J. and Rhodes, D. (2005) A method for the in vitro reconstitution of a defined “30 nm” chromatin fibre containing stoichiometric amounts of the linker histone. *J. Mol. Biol.*, **345**, 957–968.
- Zhang, Y., Rohde, C., Tierling, S., Stamerjohanns, H., Reinhardt, R., Walter, J. and Jeltsch, A. (2009) DNA methylation analysis by bisulfite conversion, cloning, and sequencing of individual clones. *Methods Mol. Biol.*, **507**, 177–187.
- Winkler, D.F., Hilpert, K., Brandt, O. and Hancock, R.E. (2009) Synthesis of peptide arrays using SPOT-technology and the CelluSpots method. *Methods Mol. Biol.*, **570**, 157–174.
- Chen, T., Ueda, Y., Xie, S. and Li, E. (2002) A novel Dnmt3a isoform produced from an alternative promoter localizes to

- euchromatin and its expression correlates with active de novo methylation. *J. Biol. Chem.*, **277**, 38746–38754.
29. Lowary,P.T. and Widom,J. (1998) New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.*, **276**, 19–42.
  30. Leuba,S.H. and Bustamante,C. (1999) Analysis of chromatin by scanning force microscopy. *Methods Mol. Biol.*, **119**, 143–160.
  31. Suetake,I., Shinozaki,F., Miyagawa,J., Takeshima,H. and Tajima,S. (2004) DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. *J. Biol. Chem.*, **279**, 27816–27823.
  32. 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.
  33. Handa,V. and Jeltsch,A. (2005) Profound flanking sequence preference of Dnmt3a and Dnmt3b mammalian DNA methyltransferases shape the human epigenome. *J. Mol. Biol.*, **348**, 1103–1112.
  34. Oka,M., Rodic,N., Graddy,J., Chang,L.J. and Terada,N. (2006) CpG sites preferentially methylated by Dnmt3a in vivo. *J. Biol. Chem.*, **281**, 9901–9908.
  35. Jurkowska,R.Z., Anspach,N., Urbanke,C., Jia,D., Reinhardt,R., Nellen,W., Cheng,X. and Jeltsch,A. (2008) Formation of nucleoprotein filaments by mammalian DNA methyltransferase Dnmt3a in complex with regulator Dnmt3L. *Nucleic Acids Res.*, **36**, 6656–6663.
  36. Takeshima,H., Suetake,I. and Tajima,S. (2008) Mouse Dnmt3a preferentially methylates linker DNA and is inhibited by histone H1. *J. Mol. Biol.*, **383**, 810–821.
  37. Ge,Y.Z., Pu,M.T., Gowher,H., Wu,H.P., Ding,J.P., Jeltsch,A. and Xu,G.L. (2004) Chromatin targeting of de novo DNA methyltransferases by the PWWP domain. *J. Biol. Chem.*, **279**, 25447–25454.
  38. Chen,T., Tsujimoto,N. and Li,E. (2004) The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin. *Mol. Cell. Biol.*, **24**, 9048–9058.
  39. Jeong,S., Liang,G., Sharma,S., Lin,J.C., Choi,S.H., Han,H., Yoo,C.B., Egger,G., Yang,A.S. and Jones,P.A. (2009) Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. *Mol. Cell. Biol.*, **29**, 5366–5376.
  40. Fuks,F., Hurd,P.J., Deplus,R. and Kouzarides,T. (2003) The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res.*, **31**, 2305–2312.
  41. Okano,M., Bell,D.W., Haber,D.A. and Li,E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, **99**, 247–257.
  42. Kaneda,M., Okano,M., Hata,K., Sado,T., Tsujimoto,N., Li,E. and Sasaki,H. (2004) Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature*, **429**, 900–903.
  43. Xu,G.L., Bestor,T.H., Bourc'his,D., Hsieh,C.L., Tommerup,N., Bugge,M., Hulten,M., Qu,X., Russo,J.J. and Viegas-Pequignot,E. (1999) Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature*, **402**, 187–191.

# **Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail**

Yingying Zhang, Renata Jurkowska, Szabolcs Soeroes, Arumugam Rajavelu, Arunkumar Dhayalan, Ina Bock, Philipp Rathert, Ole Brandt, Richard Reinhardt, Wolfgang Fischle & Albert Jeltsch

## **Supplemental Information**

### **Materials and methods**

#### **Recombinant chromatin preparation**

Expression and purification of wt *Xenopus laevis* histones was performed as described (1). H3K4me3 and H3K9me3 were generated by native protein ligation (2). The coding sequence for *X.laevis* H3 $\Delta$ 1-20 C21A was amplified by PCR and cloned into the pET3d expression vector. The truncated H3 protein was expressed and purified like the wt histones. The H3 N-terminal peptide containing residues 1-20 and tri-methylated lysines 4 or 9 were synthesized using Fmoc-based solid-phase synthesis and activated at the C-terminus by thioesterification. Ligation of the activated H3 peptide to the truncated H3 histone and purification of the ligation product was performed as described (2). Identity and purity of histones was verified by SDS-PAGE as well as by mass spectrometry.

Assembly of histone octamers containing H3unmod, H3K4me3, and H3K9me3 as well as reconstitution of recombinant oligonucleosomes was performed by salt dialysis as described using the 12 $\times$ 200 $\times$ 601 template (1,3). Briefly, octamers were reconstituted using H3unmod, H3K4me3, or H3K9me3 and purified by gel filtration on Superdex 200 (GE Healthcare). Scavenger DNA corresponding to a fragment of 148 bp length PCR amplified from pUC18 was used in all reconstitutions. Assembly reactions were titrated at different octamer:DNA ratios. Reproducibly, a octamer:DNA

ratio of 1.1:1 resulted in saturated nucleosomal arrays. Reconstituted material was used for all assays without further purification after extensive dialysis against 10 mM TEA, 0.1 mM EDTA, pH 7.5.

### **Binding of ADD domains to peptides arrays**

CelluSpots arrays were provided by Intavis AG (Köln, Germany). The array was blocked by incubation in TTBS buffer (10 mM Tris/HCl pH 8.3, 0.05% Tween-20 and 150 mM NaCl) containing 5% non-fat dried milk at 4 °C overnight, then washed three times with TTBS buffer, and incubated with purified GST-tagged ADD domain of Dnmt3a or 3b (1 µM) at room temperature for 2 hours in interaction buffer (100 mM KCl, 20 mM HEPES, pH 7.5, 1 mM EDTA, 0.1 mM DTT and 10% glycerol). After washing with TTBS buffer three times, the array was incubated with goat anti-GST antibody (GE Healthcare #27- 4577-01) 1:5000 dilution in TTBS buffer for 1 h at room temperature. Then, the array was washed three times with TTBS and incubated with horseradish peroxidase conjugated anti-Goat antibody (Invitrogen #81-1620) 1:12000 in TTBS for 1 h at room temperature. Finally, the array was washed three times with TTBS and submerged in ECL developing solution (Thermo Fisher Scientific) and image was captured in X-ray film.

### **Binding of ADD domains to native histones**

Native histones were isolated from HEK293 cell line by acid extraction method as described (4). For the binding specificity analysis of ADD domains, 7.5 µg of native histones were separated in 16% SDS-PAGE and the histones transferred to the Nitrocellulose membrane by western blotting. After staining with Ponceau S, the Nitrocellulose membrane was blocked by incubation in TTBS buffer (10 mM Tris/HCl pH 8.3, 0.05% Tween-20 and 150 mM NaCl) containing 5% non-fat dried milk at 4 °C overnight. The membrane was then washed once with TTBS buffer, and incubated with purified GST-tagged ADD domain of Dnmt3a or Dnmt3b (1 µM) at room temperature for 90 min in interaction buffer (100 mM KCl, 20 mM HEPES pH

7.5, 1 mM EDTA, 0.1 mM DTT and 10% glycerol). After washing in TTBS buffer, the membrane was incubated with goat anti-GST antibody (GE Healthcare #27-4577-01 at 1:3000 dilution) in TTBS buffer for 1 h at room temperature. Then, the membrane was washed three times with TTBS and incubated with horseradish peroxidase conjugated anti-Goat antibody (Invitrogen #81-1620 1:3000) in TTBS for 1 h at room temperature. Finally, the membrane was submerged in ECL developing solution (Thermo Scientific) and image was captured in X-ray film.

### **Methylation of the DNA in recombinant chromatin**

16 nM chromatin (unmodified and modified with H3K4me3, and H3K9me3) were methylated using 2  $\mu$ M Dnmt3a2 or Dnmt3a-C, in reaction buffer (20 mM Hepes pH 7.2, 1 mM EDTA, 100 mM KCl) in the presence of S-adenosyl-L-methionine (0.32 mM) for 2 h at 37°C. Naked DNA was methylated as a control. Dnmt3a2/3L complex was formed by pre-incubation of equimolar amounts of both proteins at room temperature for 30 min, then used to methylate the chromatin. Duplicate reactions were set up in parallel for each sample type. The reaction was stopped by freezing in liquid nitrogen. Afterwards, proteinase K digestion was performed, followed by clean-up of DNA using DNA purification Kit (Macherey Nagel, Düren, Germany).

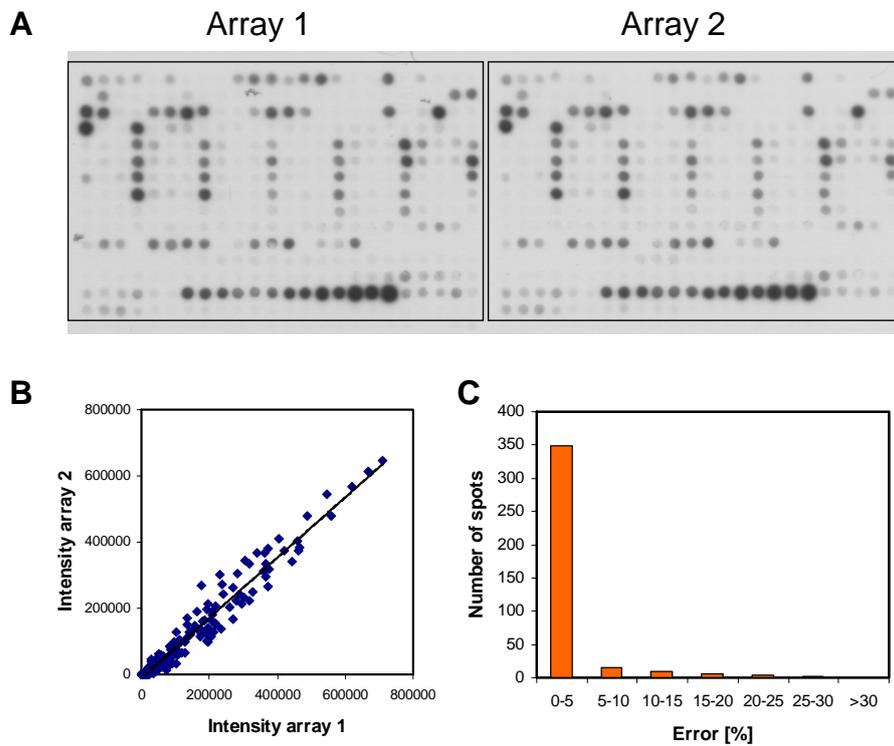
### **Bisulfite conversion, subcloning and sequencing**

The purified DNA was treated with bisulfite as described (5). The bisulfite converted DNA was used as template for PCR using primers designed for the single nucleosome (Top strand: forward GGT TAT GTG ATG GAT TTT ATA and reverse CTA TTC AAY ACA TAC ACA AAA TATA; bottom strand: forward TAA TAT ATG TAT AGG ATG TAT ATA TTT GA and reverse TCR AAT TAT ATA ATA AAC CCT ATAC; with Y representing C or T and R representing A or G). The PCR program was: 95°C, 3 min; (94°C 45 s , 43°C 35 s, 72 °C 20 s) 35 cycles, 72°C 5 min. The PCR products were purified by ChargeSwitch® PCR Clean-Up Kit (Invitrogen) and subcloned using the StrataClone™ Kit (Stratagene). Around 100 clones for each PCR product were picked for each individual experiment and sequenced using ABI BigDye Terminator

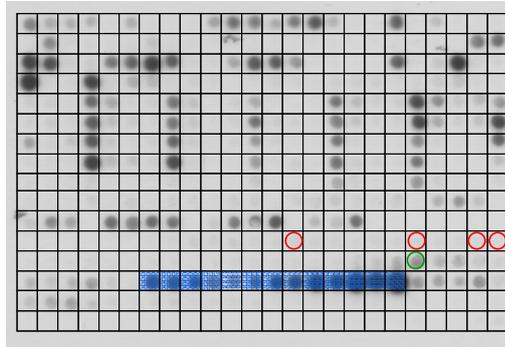
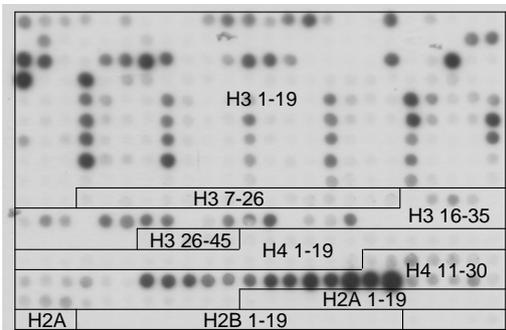
chemistry (BigDye Terminator v3.1 K) using 3730xl ABI 96-capillary sequencer systems equipped with capillaries of 50 cm separation length. BISMA and BDPC was used to perform quality control, derive DNA methylation patterns from the sequencing results, calculate the average methylation level of single CpG sites and present the DNA methylation pattern (6-8).

## Supplemental Figures

Suppl. Fig. 1: Reproducibility of ADD binding to peptide arrays. A) The arrays contain two identical sets of peptides. B) Plots of the binding intensities to the two corresponding spots in the two replicates of the array. C) The relative deviation of intensities to the two corresponding spots after normalization of activity to the strongest spot on the array.



Suppl. Fig. 2: Binding of the Dnmt3a ADD domain to non-H3 1-19 spots. The left part shows the binding pattern of the ADD domain on the histone tail array. On the right side some H4 tail spots are annotated.



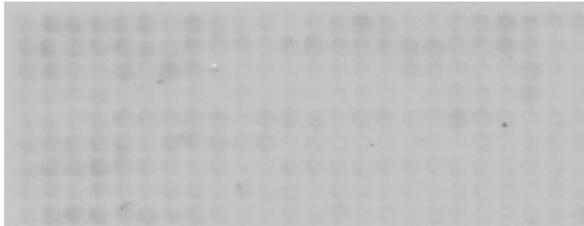
H4R3me2s

H4 11-30 K16ac

H4 11-30 K16ac + additional modifications

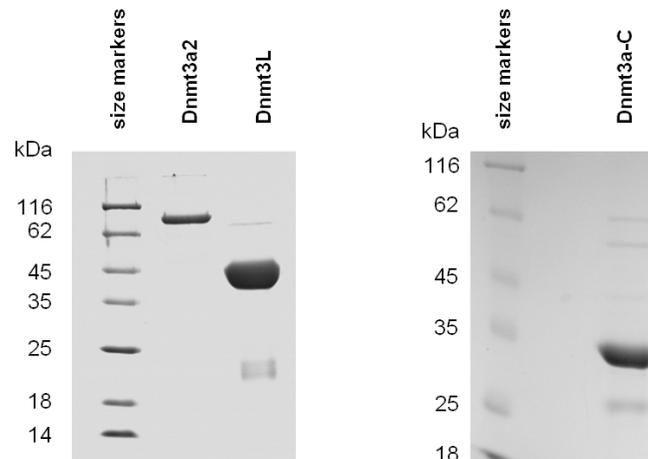
- K12ac K16ac
- K16ac R17me2s
- K16ac R17me2a
- K16ac R19me2s
- K16ac R19me2a
- K16ac K20me1
- K16ac K20me2
- K16ac K20me3
- K16ac K20ac
- K12ac K16ac K20me1
- K12ac K16ac K20me2
- K12ac K16ac K20me3
- K12ac K16ac K20ac

Suppl. Fig. 3: The Dnmt3a ADD domain did not bind to H3 1-19 peptides, if they were methylated at the N-terminus. The picture shows the same array design as in all other figures, with the only difference that the N-termini of all peptides were acetylated. We confirmed by antibody binding, that the peptides spotted in this array are fully functional (data not shown).

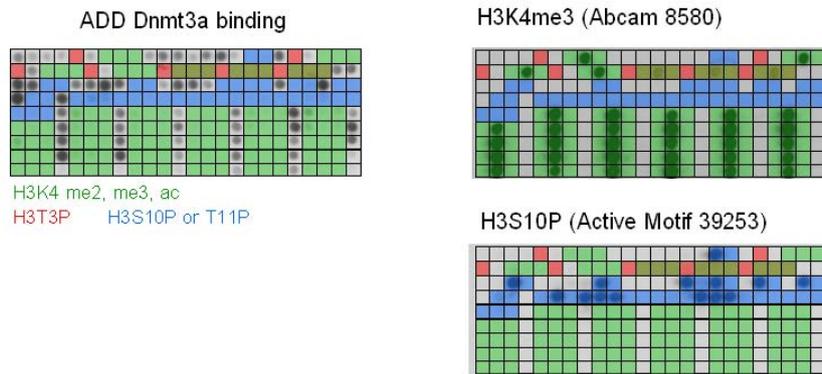


N-terminus acetylated

Suppl. Fig 4: Examples of Coomassie stained SDS polyacrylamide gel pictures of the protein preparations used in this study.



Suppl. Fig. 5: Binding of control antibodies to the peptide arrays used in this study. In the left panel, binding of the Dnmt3a ADD domain is shown. Peptides which carry a modification that prevents binding are shaded in green red and blue. On the right side, the binding of two different antibodies to the arrays is shown, with shading same as for the ADD domain. These results clearly illustrate that the spots containing the H3K4me3 or H3S10P modifications are present on the array.



## Supplemental References

1. Luger, K., Rechsteiner, T.J., Flaus, A.J., Wayne, M.M. and Richmond, T.J. (1997) Characterization of nucleosome core particles containing histone proteins made in bacteria. *Journal of molecular biology*, **272**, 301-311.
2. Shogren-Knaak, M.A. and Peterson, C.L. (2004) Creating designer histones by native chemical ligation. *Methods in enzymology*, **375**, 62-76.
3. Huynh, V.A., Robinson, P.J. and Rhodes, D. (2005) A method for the in vitro reconstitution of a defined "30 nm" chromatin fibre containing stoichiometric amounts of the linker histone. *Journal of molecular biology*, **345**, 957-968.
4. Shechter, D., Dormann, H.L., Allis, C.D. and Hake, S.B. (2007) Extraction, purification and analysis of histones. *Nature protocols*, **2**, 1445-1457.
5. Zhang, Y., Rohde, C., Tierling, S., Stamerjohanns, H., Reinhardt, R., Walter, J. and Jeltsch, A. (2009) DNA methylation analysis by bisulfite conversion, cloning, and sequencing of individual clones. *Methods in molecular biology (Clifton, N.J.)*, **507**, 177-187.
6. Rohde, C., Zhang, Y., Jurkowski, T.P., Stamerjohanns, H., Reinhardt, R. and Jeltsch, A. (2008) Bisulfite sequencing Data Presentation and Compilation (BDPC) web server--a useful tool for DNA methylation analysis. *Nucleic acids research*, **36**, e34.
7. Rohde, C., Zhang, Y., Stamerjohanns, H., Hecher, K., Reinhardt, R. and Jeltsch, A. (2009) New clustering module in BDPC bisulfite sequencing data presentation and compilation Web application for DNA methylation analyses. *BioTechniques*, in press.
8. Rohde, C., Zhang, Y., Reinhardt, R. and Jeltsch, A. (2009) BISMAs - Fast and accurate primary bisulfite sequencing data analysis of individual clones from unique and repetitive sequences. submitted for publication.