Regulation of the catalytic activity and specificity of DNA nucleotide methyltransferase 1

Von der Fakultät Energie-, Verfahrens- und Biotechnik
der Universität Stuttgart zur Erlangung der Würde
eines Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte Abhandlung

Vorgelegt von

Pavel Bashtrykov

aus Moskau, Russland

Hauptberichter: Prof. Dr. Albert Jeltsch

Mitberichter: Prof. Dr. Dieter H. Wolf

Tag der mündlichen Prüfung: 12 Feb 2014

Institut für Biochemie der Universität Stuttgart

2014

Эта работа посвящается моему папе, Баштрыкову Павлу Михайловичу,
с благодарностью за его любовь, терпение и педагогический талант.

This work is dedicated to my dad, Pavel Mikhailovich Bashtrykov,
with gratitude for his love, patience and pedagogical talent.

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich diese Arbeit selbst verfasst und dabei keine anderen als die angegeben Quellen und Hilfsmittel verwendet habe.

Stuttgart, 7 November 2013

Pavel Bashtrykov

Acknowledgements

I want to thank my supervisor, Prof. Dr. Albert Jeltsch, for providing me the opportunity to work in his lab, for teaching me and for sharing enthusiasm and love to the science.

I am very grateful to my direct supervisor and best friend, Dr. Sergey Ragozin, for his invaluable assistance and the transfer of knowledge, for his optimism and sense of humor that helped me during my work on the dissertation.

I am thankful to Prof. Dr. Dieter H. Wolf for being the co-referee of my PhD thesis. I appreciate very much the work, which was done by Prof. Dr. Stephan Nußberger, Prof. Dr. Arnd G. Heyer, Prof. Dr. (apl.) Christina Wege, Prof. Dr. Peter Scheurich being in the PhD committee.

I want to thank all my colleagues for a friendly and supportive environment in the lab and lots of useful advices and comments.

I am grateful to my wife, Inna Bashtrykova, for her love, patience, understanding and support.

I would like to thank all my relatives who invested their energy and passion in my intellectual growth and always believe in me.

TABLE OF CONTENTS

List of publications

Bashtrykov P, Jankevicius G, Smarandache A, Jurkowska RZ, Ragozin S, Jeltsch A. *Specificity of Dnmt1 for methylation of hemimethylated CpG sites resides in its catalytic domain.* **Chemistry & Biology.** 2012;19(5):572-8.

Bashtrykov P, Ragozin S, Jeltsch A. *Mechanistic details of the DNA recognition by the Dnmt1 DNA methyltransferase.* **FEBS letters.** 2012;586(13):1821-3.

Ceccaldi A, Rajavelu A, Ragozin S, Senamaud-Beaufort C, Bashtrykov P, Testa N, et al. *Identification of novel inhibitors of DNA methylation by screening of a chemical library.* **ACS Chemical Biology.** 2013;8(3):543-8.

Bashtrykov P, Jankevicius G, Jurkowska RZ, Ragozin S, Jeltsch A.
Uhrf1 stimulates the activity and specificity of the maintenance DNA methyltransferase Dnmt1 by an allosteric mechanism. **J. Biol. Chem.** 2014;289(7):4106-15.

Bashtrykov P, Rajavelu A, Hackner B, Ragozin S, Carell T, Jeltsch A. *Targeted mutagenesis results in an activation of DNA methyltransferase 1 and confirms an autoinhibitory role of its RFTS domain.* **ChemBioChem.** 2014 in press.

List of abbreviations

5caCyt	5-carboxylcytosine
5fCyt	5-fluorocytosine
5hmCyt	5-hydroxymethylcytosine
5mCyt	5-methylcytosine
AdoHcy	S-adenosyl-*L*-homocystein
AdoMet	S-adenosyl-*L*-methionine
BAH1/2	Bromo-adjacent homology 1 and 2 domains
BER	Base excision repair
CGBP	CpG binding protein
CXXC	Cys-X-X-Cys domain
DMAP1	DNA methyltransferase associated protein 1
Dnmt1	DNA methyltransferase 1
Dnmt3a	DNA methyltransferase 3a
Dnmt3b	DNA methyltransferase 3b
Dnmt3L	DNA methyltransferase 3L
Dnmts	DNA methyltransferases
*Eco*Dam	*Escherichia coli* DNA adenine methyltransferase
M.*Eco*RV	*Escherichia coli* methyltransferase *Eco*RV
EHMT1	Euchromatic histone-lysine N-methyltransferase 1
ESs	Embryonic stem cells
Fig	Figure
GK linker	Glycine-lysine linker
GST	Glutathione S-transferase
HDAC1	Histone deacetylase 1
H2A	Histone 2A
H2B	Histone 2B
H3K9me2/3	Histone 3 lysine 9 di-/trimethylation
H3K9me3	Histone 3 lysine 9 trimethylation
H3R2	Histone 3 arginine 2
H4K20me2	Histone 4 lysine 20 dimethylation
HP1	Heterochromatin protein 1
M.*Hae*III	*Haemophilus aegypticus* methyltransferase III
M.*Hha*I	*Haemophilus haemolyticus* methyltransferase I
M.*Hpa*II	*Haemophilus parainfluenzae* methyltransferase II
M.*Sss*I	*Spiroplasma* sp. strain MQ1 methyltransferase I
M.*Taq*I	*Thermus aquaticus* methyltransferase I
MBD1	Methyl-CpG binding domain protein 1
MLL	Mixed Lineage Leukemia
NFAT1	Nuclear Factor of Activated T-cells 1
NF-kB	Nuclear Factor kappa B
Ni-NTA	Ni^{2+}-Nitrilotriacetic acid
Np95	Nuclear protein 95
ORC1	Origin of Replication Complex 1
PBD	PCNA binding domain
PCNA	Proliferating cell nuclear antigen
PGCs	Primordial germ cells

PHD	Plant homeodomain
PML	Promyeloic Leukemia
PMTs	Post-translational modifications
RFTS	Replication Foci Targeting Sequence domain
RING	Really Interesting New Gene domain
SETD7	SET domain containing (lysine methyltransferase) 7
Sir3	Silent information regulator 3
SP1	Specificity protein 1
SP3	Specificity protein 3
SRA	SET and RING-associated domain
STAT3	Signal transducer and activator of transcription 3
SUV39H1	Suppressor of variegation 3-9 homolog 1
T4Dam	Bacteriophage T4 DNA adenine methyltransferase
TDG	Thymine DNA glycosylase
Tet	Ten-eleven translocation
TIP60	Tat-interacting protein of 60 kDa
TRD	Target recognition domain
TTD	Tandem Tudor domain
Ubl	Ubiquitin-like domain
USP7	Ubiquitin specific peptidase 7
Uhrf1	Ubiquitin-like, PHD and RING finger domain-containing 1
ZMET2	*Zea mays* methyltransferase 2

Abstract

DNA nucleotide methyltransferase 1 (Dnmt1) is mainly responsible for the maintenance of DNA methylation in mammals and plays a crucial role in the epigenetic control of gene expression. Dnmt1 recognizes and methylates hemimethylated CpG sites formed during DNA replication. In the present work, the mechanistic details of the substrate recognition by the catalytic domain of Dnmt1, the possible role of the CXXC and RFTS domains of Dnmt1 in the regulation of specificity and activity of Dnmt1, and the influence of the Ubiquitin-like PHD and RING finger domain-containing 1 (Uhrf1) protein on the enzymatic properties of Dnmt1 was investigated.

Using modified substrates, the functional roles of individual contacts of the Dnmt1 catalytic domain with the CpG site of the DNA substrate were analysed. The data show that the interaction with the 5-methylcytosine:guanine pair is required for the catalytic activity of Dnmt1, whereas the contacts to the non-target strand guanine are not important, since its replacement with adenine increased the activity of Dnmt1.

It was proposed that the CXXC domain binding to unmethylated CpG sites increases the specificity of Dnmt1 for hemimethylated DNA. Our data showed that the CXXC domain does not influence the enzyme's specificity in the full-length Dnmt1. In contrast, mutagenesis in the catalytic domain introducing an M1235S exchange resulted in a significant reduction in specificity. Therefore, the readout for the hemimethylated DNA occurs within its catalytic domain.

It was observed in a crystal structure that the RFTS domain of Dnmt1 inhibits the activity of the enzyme by binding to the catalytic domain and blocking the entry of the DNA. By amino acid substitution in the RFTS domain its positioning within the catalytic domain was destabilized and a corresponding increase in the catalytic rate was observed, which supports this concept and suggests a possible mechanism to allosterically regulate the activity of Dnmt1 in cells.

Uhrf1 has been shown to target Dnmt1 to replicated DNA, which is essential for DNA methylation. Here it is demonstrated that Uhrf1 as well as its isolated SRA domain increase the activity and specificity of Dnmt1 in an allosteric mechanism. The stimulatory effect was independent of the SRA domain's ability to bind hemimethylated DNA. The RFTS domain of Dnmt1 is required for the stimulation, since its deletion or blocking of its interaction with the SRA domain, significantly reduced the ability of

Uhrf1 to increase the activity and specificity of Dnmt1. Uhrf1, therefore, plays multiple roles that support DNA methylation including targeting of Dnmt1, its stimulation and an increase of its specificity.

Zusammenfassung

Die DNA-Methyltransferase 1 (Dnmt1) ist hauptverantwortlich für die Konservierung der DNA-Methylierung bei Säugetieren und spielt eine entscheidende Rolle in der epigenetischen Kontrolle der Genexpression. Dnmt1 erkennt und methyliert hemimethylierte CpG-Stellen, die während der DNA-Replikation gebildet werden. In der vorliegenden Arbeit wurden mechanistische Details der Substraterkennung durch die katalytische Domäne von Dnmt1, die mögliche Rolle der CXXC- und RFTS-Domänen in der Regulation der Spezifität und Aktivität von Dnmt1 sowie der Einfluss des Ubiquitin-like PHD- und RING-Finger-Domänen enthaltenden 1 (Uhrf1) Proteins auf die enzymatischen Eigenschaften von Dnmt1 untersucht.

Mit verschiedenen modifizierten Substraten wurde die funktionelle Rolle einzelner Kontakte der katalytischen Domäne von Dnmt1 mit der CpG-Stelle der Substrat-DNA untersucht. Unsere Daten zeigen, dass die Interaktion mit dem 5-Methylcytosin:Guanin-Paar für die katalytische Aktivität von Dnmt1 notwendig ist, während die Kontakte zum im Gegenstrang liegenden Guanin offenbar nicht von Bedeutung sind, da der Austausch dieses Guanins gegen Adenin zu einer erhöhten Aktivität von Dnmt1 führte.

In der Literatur wurde vorgeschlagen, dass die CXXC-Domäne durch die Bindung an unmethylierte DNA die Spezifität von Dnmt1 für hemimethylierte DNA erhöhen kann. Wir konnten allerdings zeigen, dass die CXXC-Domäne von Dnmt1 die Spezifität des Enzyms nicht beeinflusst. Im Gegensatz dazu führte der Austausch M1235S in der katalytischen Domäne von Dnmt1 zu einer signifikanten Reduktion der Spezifität. Daher muss die Erkennung hemimethylierter DNA innerhalb der katalytischen Domäne von Dnmt1 stattfinden.

Die Untersuchung einer Kristallstruktur ergab, dass die RFTS-Domäne die Aktivität von Dnmt1 durch Bindung an die katalytische Domäne und die Blockierung der Eintrittsstelle der DNA hemmt. Durch Aminosäuresubstitutionen in der RFTS-Domäne konnte deren Positionierung innerhalb der katalytischen Domäne destabilisiert werden, was zu einer entsprechenden Erhöhung der katalytischen Rate führte. Unsere Beobachtung unterstützt dieses Konzept und zeigt einen möglichen Mechanismus auf, mit dem die Aktivität von Dnmt1 in Zellen allosterisch reguliert werden kann.

Uhrf1 rekrutiert Dnmt1 an kürzlich replizierte DNA. Außerdem konnten wir zeigen, dass Uhrf1 sowie seine isolierte SRA-Domäne die Aktivität und Spezifität von Dnmt1 nach einem allosterischen Mechanismus erhöht. Diese stimulierende Wirkung war unabhängig von der Fähigkeit der SRA-Domäne, hemimethylierte DNA zu binden. Die RFTS-Domäne war für die Stimulation erforderlich, da ihre Entfernung oder die Blockade der Wechselwirkung mit der SRA-Domäne die Fähigkeit von Uhrf1, die Aktivität und Spezifität von Dnmt1 zu steigern, deutlich reduziert. Unsere Daten zeigen, dass Uhrf1 bei der Unterstützung der DNA-Methylierung mehrere Aufgaben erfüllt, welche die Rekrutierung, Stimulation und Steigerung der Spezifität von Dnmt1 umfassen.

1. Introduction

1.1 Epigenetic phenomena in eukaryotes

Regulation of the retrieval of genomic information is one of the most important tasks for living organisms. The maximal complexity of genome regulation is reached in multicellular organisms, especially in mammals, because they contain more than one hundred different cell types, which have stable and distinct phenotypes. Still, all the cells originate from one original precursor cell, the zygote, and basically all have the same genome. Cellular specialization happens due to differential gene expression, which is in turn epigenetically controlled. One popular definition of epigenetics is a "Change in phenotype that is heritable but does not involve DNA mutation" (Gottschling, 2007, Epigenetics, Cold Spring Harbor Press). There are several epigenetic signals including DNA methylation, histone modifications, histone variants, and non-coding RNAs, which all together regulate chromatin structure, control gene expression, and genome stability (Bonasio et al, 2010). We will focus on DNA methylation as a main topic of the present work.

1.2 DNA methylation as a paradigm of epigenetic signalling

1.2.1 DNA methylation is a durable, reversible, and heritable mark

DNA methylation is a universal biochemical phenomenon found in bacteria, plants, fungi, and animals (Jeltsch, 2010). DNA methylation in mammals occurs mainly at position 5 of cytosine in the context of CpG dinucleotides (Bird, 2002; Jeltsch, 2002). In addition, embryonic stem cells also have some methylation of cytosine in a non-CG context (Lister et al, 2009). Approximately 70-80% of CpG dinucleotides are methylated, and the distribution of the methylation through the genome, the so called DNA methylation pattern, is not random, but has cell-type specific characteristics (Law & Jacobsen, 2010). Normally, palindromic CpG sites are either methylated on both strands or completely unmethylated, which explains the fundamental basis for the inheritance of DNA methylation patterns. Semiconservative DNA replication leads to the synthesis of the new daughter strand without methylation marks, which results in the

generation of hemimethylated CpG sites from the parental fully methylated CpG sites, while unmethylated parental sites remain unmethylated after DNA replication. Therefore, DNA methylation patterns present on the parental strand can be transferred to the daughter strand by an enzyme capable to recognize and methylate hemimethylated CpG sites (Holliday & Pugh, 1975; Riggs, 1975). Indeed, it was shown that DNA methylation is a very stable mark, which can be preserved for many cell divisions (Schubeler et al, 2000).

CpG sites are non-homogeneously distributed across the genome. Interestingly, CpG-poor regions are usually methylated, whereas CpG-rich regions are hypomethylated. CpG-rich regions, also known as CpG islands, are found within the promoters of many genes (Bird, 1986). DNA methylation is a repressive epigenetic signal, since high methylation levels of CpG islands in promoter regions usually correlate with a low transcriptional activity of the corresponding genes (Tate & Bird, 1993). DNA methylation represses gene transcription via two mechanisms. 1) Methylation of CpG islands recruits methyl-CpG-binding proteins, which interact with other factors and form repressive complexes regulating chromatin structure and inhibiting transcription (Jones et al, 1998; Nan et al, 1998). 2) Methylation of cytosines blocks sequence-specific DNA binding of transcription factors and directly inhibits transcription initiation (Bell & Felsenfeld, 2000; Prendergast et al, 1991). DNA methylation-mediated gene repression is involved in several biological processes, such as cellular differentiation (Hemberger et al, 2009), embryonic development (Latham et al, 2008; Reik, 2007), parental control of imprinted genes (Kelsey & Feil, 2013; Li et al, 1993), X chromosome inactivation (Barakat & Gribnau, 2012; Panning & Jaenisch, 1998), and silencing of repetitive elements and transposons (Bestor & Bourc'his, 2004; Walsh et al, 1998). Dysregulation of DNA methylation underlies various human diseases, including cancer, psychiatric disorders, and developmental abnormalities (Bergman & Cedar, 2013; Grayson & Guidotti, 2013; Jones & Baylin, 2007; Portela & Esteller, 2010; Yin et al, 2012).

DNA methylation is a durable, but reversible epigenetic signal (Auclair & Weber, 2012). Methylation of cytosines is catalysed by enzymes called DNA methyltransferases (Dnmts). DNA methylation is set by *de novo* Dnmts (**Fig. 1**). The established DNA methylation patterns are normally copied after every DNA replication

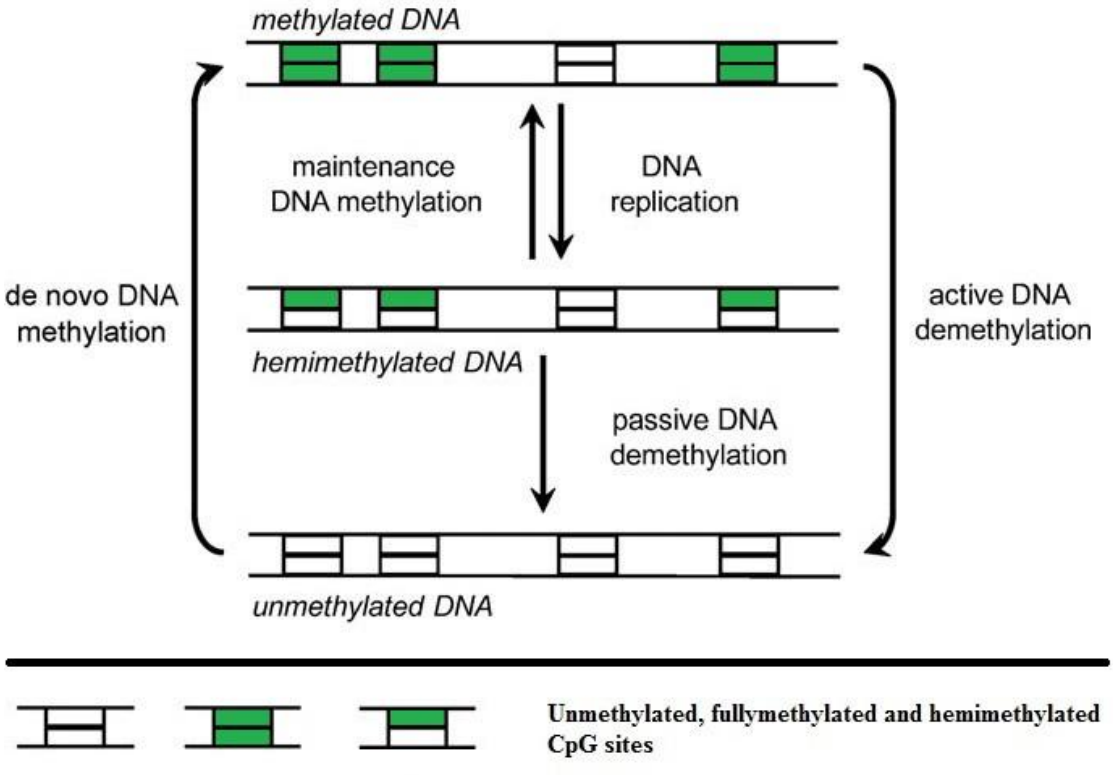

Figure 1. DNA methylation in eukaryotes (from Jurkowska et al. (2011) with modifications).
DNA methylation pattern is set by *de novo* DNA methylation and preserved by the maintenance DNA methylation. The DNA methylation mark can be erased via passive or active DNA demethylation processes.

and formation of the hemimethylated DNA. This task is taken over by the maintenance DNA methyltransferase, which specifically methylates hemimethylated CpG sites. The removal of methylation marks can utilize either passive or active DNA demethylation mechanisms. The passive DNA demethylation may happen as a result of DNA replication if the maintenance DNA methylation is inhibited and formed hemimethylated CpG sites are not remethylated. This process is relatively slow and requires several cell divisions before a significant dilution of the DNA methylation pattern is achieved. Active DNA demethylation is based on a specific enzymatic activity resulting in an erasure of methylation marks and is independent of DNA replication (Dalton & Bellacosa, 2012; Kinney & Pradhan, 2013).

The removal of the chemically inert methyl group is nontrivial, and the mechanisms underlying this process are not completely understood. Several demethylation pathways have been proposed (Wu & Zhang, 2010): The first one starts with the deamination of 5-methylcytosine (5mCyt) resulting in the formation of thymine and a TG mismatch. Then, a base excision repair (BER) process is initiated with the

removal of the thymine base by the thymine DNA glycosylase (TDG) followed by other steps resulting in the introduction of cytosine. The second pathway utilizes a direct oxidation of methyl groups (Jurkowski & Jeltsch, 2011). Recently, a new Ten-eleven translocation (Tet) family of proteins was identified (Tan & Shi, 2012). Tet1, Tet2, and Tet3 enzymes catalyse the oxidation of 5mCyt to 5-hydroxymethylcytosine (5hmCyt). For their dioxygenase activity, Tet proteins require α-ketoglutarate as a co-substrate and molecular oxygen. 5hmCyt can be further oxidized by Tet enzymes to 5-formylcytosine and 5-carboxylcytosine (5caCyt) or deaminated by other factors to 5-hydroxymethyluracil. The final step of 5mCyt oxidation to 5caCyt may be followed by decarboxylation of the 5caCyt leading to the formation of cytosine, but an enzyme responsible for this activity has not yet been identified. All modified bases can be excised by DNA glycosylases and a following BER process that finally leads to demethylation. It was proposed that the occurrence of oxygen-based demethylation in nature developed after the increase of atmospheric oxygen and gave a burst to the formation of multicellular organisms (Jeltsch, 2013).

The investigation of the genome-wide distribution of DNA methylation revealed a massive reprogramming of the DNA methylation patterns during the embryonic development of mammals (Saitou et al, 2012; Seisenberger et al, 2013). A first global DNA demethylation event takes place immediately after fertilization. The paternal genome loses its methylation mark before the DNA replication in the zygote probably via active demethylation mechanisms employing Tet enzymes. In contrast, the maternal genome undergoes a slower demethylation through several cell divisions, which ends only in the early blastocyst. It is assumed that demethylation in that case utilizes a passive mechanism. Then, shortly before and after implantation, massive and genome wide *de novo* DNA methylation takes place accompanying the differentiation of pluripotent stem cells and the formation of specific cell lines. DNA methylation patterns of differentiated somatic cells are mainly preserved during through the following cell divisions in the embryo and the adult organism, although some specific changes in DNA methylation occur during the development of some cell types. In contrast, primordial germ cells (PGCs) undergo a second demethylation event between embryonic day 7.5 and 13.5 (in mouse), and the further development of PGCs into specialized gametes results in an additional wave of *de novo* DNA methylation.

1.2.2 DNA methyltransferases

Initially, two functional types of DNA methyltransferases were proposed by Riggs, Holliday, and Pugh in 1975 (Holliday & Pugh, 1975; Riggs, 1975). They postulated that *de novo* enzymes, setting the methylation mark, and maintenance enzymes, responsible for the preservation of the DNA methylation, should exist. The maintenance enzyme(s) must exhibit strong preference towards hemimethylated CpG sites. Within the next 20 years of research, enzymes corresponding to the predicted activities were discovered and characterized (Jeltsch, 2002).

All Dnmts transfer the methyl group from the cofactor S-adenosyl-*L*-methionine (AdoMet) to the position 5 of the cytosine base. They all contain a C-terminal catalytic domain and an N-terminal regulatory part. The catalytic domain contains ten characteristic amino acid motifs, which are conserved between eukaryotic and prokaryotic C5-cytosine methyltransferases (Cheng, 1995).

The family of mammalian DNA methyltransferases contains three enzymes – Dnmt1, Dnmt3a, and Dnmt3b as well as one catalytically inactive protein – Dnmt3L (**Fig. 2**). Dnmt3a and Dnmt3b are *de novo* methyltransferases required for setting DNA methylation patterns during embryogenesis and gametogenesis (Kaneda et al, 2004; Okano et al, 1999), and the Dnmt3L protein is a potent activator of Dnmt3s (Gowher et al, 2005; Jia et al, 2007; Jurkowska et al, 2011b). In addition to their main role in the establishment of methylation patterns, Dnmt3 enzymes were shown to be involved also in the maintenance DNA methylation of repetitive elements (Jones & Liang, 2009;

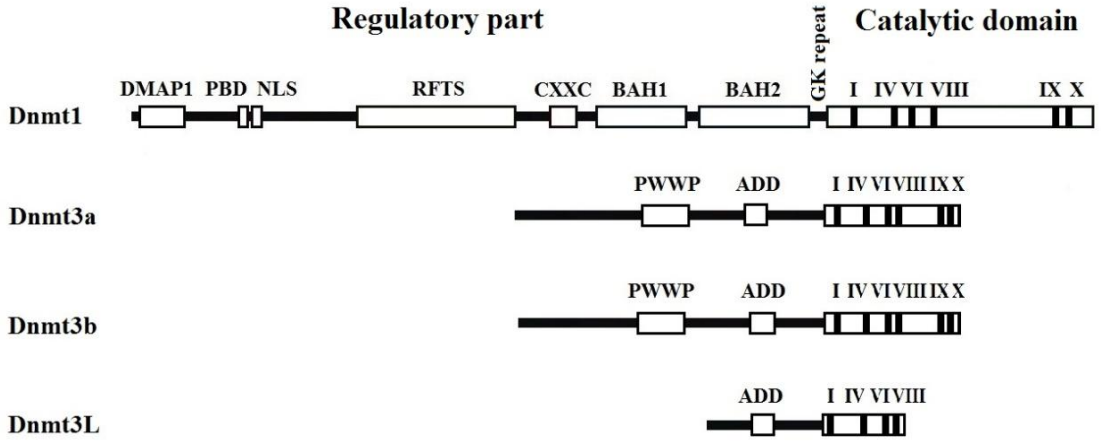

Figure 2. Domain structures of mammalian DNA methyltransferases.
Mammalian Dnmts consist of the N-terminal variable regulatory part and the C-terminal catalytic domain harbouring conserved amino acid motifs required for catalysis.

Liang et al, 2002). Dnmt1 is a maintenance DNA methyltransferase with preference for methylation of hemimethylated CpG sites, which is the subject of this work and will be described in details in the following chapters.

1.3 Dnmt1 function, structure, and regulation

1.3.1 History, catalytic demands, and determinants

Dnmt1 was the first mammalian DNA methyltransferase which was cloned and biochemically characterized (Bestor et al, 1988). Dnmt1 is the main enzyme responsible for the maintenance DNA methylation. The targeted disruption of the *Dnmt1* gene leads to a threefold reduction of the DNA methylation level in embryonic stem (ES) cells. Experiments with *Dnmt1* disruption in mice showed that embryos are delayed in development and die shortly after gastrulation, indicating that the Dnmt1 function is essential for embryonic development (Lei et al, 1996; Li et al, 1992). Also, Dnmt1 is involved in the silencing of imprinted genes, X chromosome inactivation, and maintenance of the pluripotency of mesenchymal stem cells (Li et al, 1993; Sado et al, 2000 Tsai, 2012). Recently, this finding was confirmed and validated in an animal model, harbouring a Dnmt1 catalytically inactive mutant (Takebayashi et al, 2007). Mice containing this mutation in both alleles showed a severe phenotype similar to the Dnmt1 knockout animals. This indicates that the loss of catalytic activity of Dnmt1 is responsible for the embryonic lethality.

Dnmt1 is highly abundant in proliferating cells, and conversely its expression is low in quiescent cells (Robertson et al, 2000). Dnmt1 is a cell cycle-regulated protein, and its expression and cellular localization changes during the cell cycle. Dnmt1 expression can be detected during all phases of the cell cycle, but maximum abundance is reached in the S phase during DNA replication (Kimura et al, 2003; Lee et al, 1996). Dnmt1 is a nuclear protein and has a characteristic subnuclear localization. During the S phase, Dnmt1 associates with replication foci, which are regions of active DNA replication and formation of hemimethylated CpG sites. In the early S phase, it forms a punctuate pattern, corresponding to replication forks in the euchromatin. In the middle and late S phase, Dnmt1-containing structures become lager, less numerous, and adopt a

toroidal conformation. These structures were shown to be replicating centromeric heterochromatin. In other cell cycle phases, the enzyme does not have a specific localization and is diffusely distributed in the nucleus (Easwaran et al, 2004; Leonhardt et al, 1992; Liu et al, 1998).

As a maintenance DNA methyltransferase, Dnmt1 specifically methylates hemimethylated CpG sites. The preference of Dnmt1 for the hemimethylated DNA over unmethylated DNA determined *in vitro* is in the range of ten- to 40-fold (Bashtrykov et al, 2012a; Fatemi et al, 2001; Goyal et al, 2006; Hermann et al, 2004; Jeltsch, 2006; Pradhan et al, 1999). This variability is mainly due to the variation of the methylation rate of unmethylated CpG sites and also related to different assay conditions. Dnmt1 is a highly processive enzyme, methylating up to 30 CpG sites without dissociation from the substrate (Bestor & Ingram, 1983; Goyal et al, 2006; Hermann et al, 2004; Vilkaitis et al, 2005). The processive methylation takes place only on one DNA strand, indicating that Dnmt1 slides along the newly synthesized strand and methylates hemimethylated sites after DNA replication (Hermann et al, 2004).

1.3.2 Domain organization of Dnmt1

Dnmt1 is a 180 kDa single-chain polypeptide, containing 1620 amino acids in the mouse and 1616 amino acids in the human protein (Jurkowska et al, 2011a). Structurally, Dnmt1 consists of two parts (**Fig. 2**, page 16), a C-terminal catalytic domain (amino acids 1140-1620 of the mouse protein) and an N-terminal multidomain regulatory part, which are connected by a linker of six glycine-lysine dipeptides (GK linker). The catalytic domain of Dnmt1 contains ten amino acid motifs, which are characteristic for DNA methyltransferases. Despite the high similarity to other DNA methyltransferases, the isolated catalytic domain is enzymatically inactive, which indicates the requirement of the N-terminal part for catalysis (Fatemi et al, 2001; Margot et al, 2000; Zimmermann et al, 1997).

Several domains can be distinguished within the N-terminal part of Dnmt1: a DNA methyltransferase associated protein 1 (DMAP1) interaction domain, a Proliferating cell nuclear antigen (PCNA) binding domain (PBD), a Nuclear Localization Signal (NLS), a Replication Foci Targeting Sequence (RFTS) domain, a zinc finger domain, and two Bromo-adjacent homology 1 and 2 (BAH1/2) domains.

The DMAP1 interaction domain binds the DNA methyltransferase associated protein 1and recruits it to the replication forks during S phase (Lee et al, 2010; Rountree et al, 2000). DMAP1 is involved in transcription repression and DNA repair, and it interacts with many other proteins. The PBD domain is responsible for the interaction of Dnmt1 with the proliferative cell nuclear antigen (PCNA) at the replication foci and contributes to the maintenance DNA methylation (Chuang et al, 1997; Egger et al, 2006; Schermelleh et al, 2007). Targeting of Dnmt1 to the nucleus is mediated by three NLS sequences, mapped within the 650 N-terminal amino acids (Cardoso & Leonhardt, 1999). The RFTS domain was initially discovered as a domain targeting Dnmt1 to the replication forks during S phase (Leonhardt et al, 1992). Later it was found to mediate the interaction of Dnmt1 with heterochromatin during the G2 and M phases (Easwaran et al, 2004). The targeting of Dnmt1 to the replicating chromatin is mediated by its direct interaction with the SET and RING-associated (SRA) domain of the Ubiquitin-like, PHD and RING finger domain-containing 1 (Uhrf1) protein (Achour et al, 2008). Mutations in the RFTS domain of Dnmt1 cause neurological diseases, including autosomal dominant cerebellar ataxia, narcolepsy, and deafness (Winkelmann et al, 2012). The zinc finger domain of Dnmt1, known as the Cys-X-X-Cys (CXXC) domain, binds to the DNA containing unmethylated CpG dinucleotides (Frauer et al, 2011; Pradhan et al, 2008). It has eight conserved cysteine residues coordinating two zinc cations and forms a crescent-like fold (Song et al, 2011). It requires zinc for DNA binding (Lee et al, 2001). Several related CXXC domains were found in other proteins involved in the modification of DNA and histones, such as the CpG binding protein (CGBP), the methyl-CpG binding domain protein 1 (MBD1), Ten-eleven translocation 1 (Tet1), and Mixed Lineage Leukemia (MLL) (Ayton et al, 2004; Frauer et al, 2011 2002; Jorgensen et al, 2004; Lee et al, 2001). It was shown that the CXXC domain is crucial for the enzymatic activity of Dnmt1 (Pradhan et al, 2008), and it was proposed to be responsible for the specificity of Dnmt1 towards hemimethylated DNA (Song et al, 2011) (see below). Two BAH1/2 domains have a so far unknown function in Dnmt1. Other proteins containing similar BAH motifs are involved in the regulation of transcription and replication. The BAH domain of the Origin of Replication Complex 1 (ORC1) protein improves the association of ORC with chromatin and is involved in the activation of origins of replication (Noguchi et al, 2006). Recently, it was found that the

ORC1 BAH domain recognizes histone H4 lysine 20 dimethylation (H4K20me2), a chromatin mark enriched at origins of replication (Kuo et al, 2012). The BAH domain from *Zea mays* methyltransferase 2 (ZMET2), a plant methyltransferase responsible for CHG methylation, binds histone H3 lysine 9 dimethylation (H3K9me2) (Du et al, 2012). The BAH domain of the yeast silent information regulator 3 (Sir3) protein interacts with histones H3 and H4 tails and plays a role in gene silencing (Sampath et al, 2009).

DNA replication is a very processive reaction and takes approximately 0.035 sec per incorporated nucleotide (Jackson & Pombo, 1998). In contrast, *in vitro* studies showed that Dnmt1 methylates with a turnover rate in the range of 0.1-1 turnover/min (Hermann et al, 2004; Pradhan et al, 1999; Pradhan et al, 1997; Song et al, 2011). This is not enough to follow the replication fork and to copy a methylation pattern of 56 million CpG sites of the human genome in a reasonable time during the S phase (Egger et al, 2006). More to that, as mentioned above, the preference of Dnmt1 for the methylation of hemimethylated CpG sites is only ten- to 40-fold, which is not sufficient to guarantee an accurate copying of methylation patterns. In fact, Dnmt1 shows significant activity towards unmethylated CpG sites *in vitro*. The *de novo* methylation of CpG islands in promoter regions of tumor suppressor genes by Dnmt1 found in cancer cells makes it a promising candidate for anti-cancer therapy and a target for synthesis of specific inhibitors (Ceccaldi et al, 2013; Feltus et al, 2003; Jair et al, 2006). Hence, mechanisms increasing the efficiency and specificity of Dnmt1 should exist *in vivo*. This hypothesis got a lot of experimental support during the last years. In the following chapters, we discuss intrinsic properties of Dnmt1 making it suitable for the maintenance work as well as additional factors, such as interacting partners, which facilitate maintenance DNA methylation modulating properties of Dnmt1.

1.3.3 Crystallographic studies / intrinsic properties of Dnmt1

For a long time, detailed structural information of Dnmt1 was not available. However, several crystal structures of Dnmt1 were published within the last three years and provided novel and very important mechanistic insights into substrate recognition by Dnmt1. They gave rise to several models attempting to explain the contribution of

Dnmt1's N-terminal part to the maintenance of DNA methylation. Two domains, the CXXC and the RFTS, apart from the catalytic domain, were proposed to influence the enzymatic properties of Dnmt1 – but all structures available so far are truncated proteins, missing considerable parts of the N-terminus.

First, in 2011, Song and colleagues crystallized a murine Dnmt1 fragment comprising the residues 650-1602 including the CXXC, BAH1/2 and catalytic domains. It was co-crystallized with S-adenosyl-*L*-homocystein (AdoHcy) and a 19 base pair double-stranded DNA oligonucleotide (Song et al, 2011). The DNA contained two unmethylated CpG sites. The solved crystal structure revealed all domains as well as DNA and AdoHcy at 3.0 Å resolution (**Fig. 3**). The core of the structure was formed by the catalytic domain with the AdoHcy bound in the catalytic centre. The catalytic domain can be subdivided into two functional subdomains, a methyltransferase and a Target recognition domain (TRD).

An analysis of the structure of the methyltransferase domain revealed that is has a fold common for all class I methyltransferases (Cheng, 1995). Comparison of the catalytic domain structure with the structure of the well-studied bacterial methyltransferase I from *Haemophilus haemolyticus* (M.*Hha*I) (Klimasauskas et al, 1994) demonstrated a high similarity of the methyltransferase subdomains of both

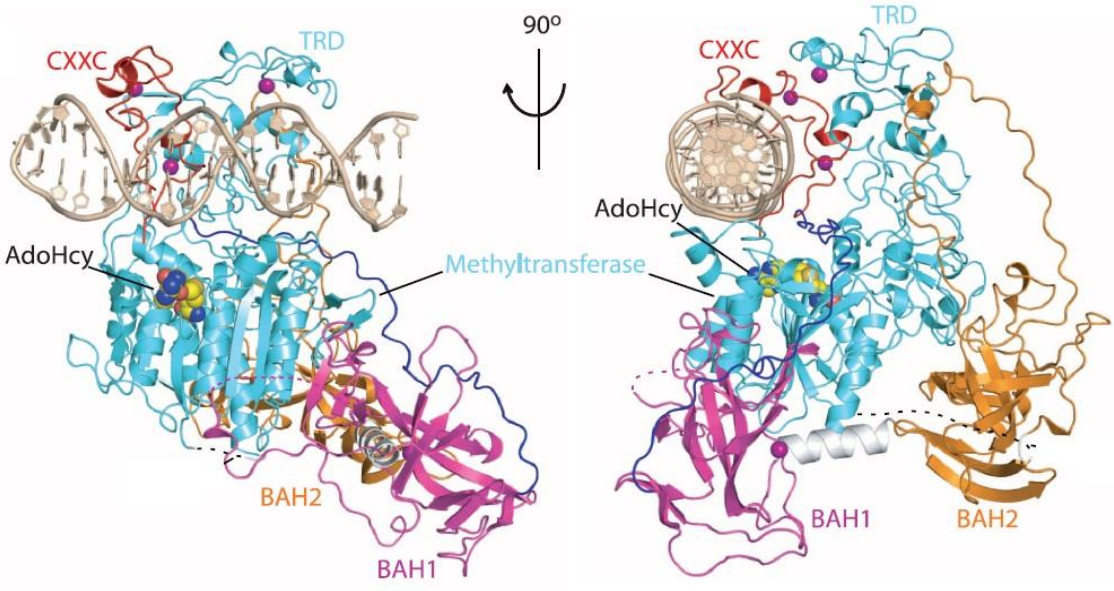

Figure 3. Crystal structure of the Dnmt1 (650-1602)-DNA complex
The CXXC, BAH1, BAH2, and catalytic domains are shown in red, magenta, orange, and cyan, respectively. The double helix DNA is bound to the CXXC domain. The catalytic cleft is blocked by the CXXC domain and the CXXC-BAH1 linker. Adopted from Song et al. (2011) with modifications.

enzymes (Song et al, 2011). M.*Hha*I was crystallized in a complex with DNA, which was bound to the catalytic cleft formed by the methyltransferase and TRD subdomains. This was different in the Dnmt1 structure – here the DNA was bound to the CXXC domain and kept aside from the DNA binding site of the catalytic domain. Surprisingly, the DNA binding cleft of Dnmt1 was occupied by the CXXC domain and an acidic peptide linker connecting the CXXC and BAH1 domains. The structure of the CXXC domain was similar to the crescent-like structure initially described for the CXXC domain of the MLL1 protein (Allen et al, 2006). The CXXC domain formed specific contacts with one CpG site by penetrating the major grove of the DNA with a loop containing four residues, namely Arg684-Ser685-Lys686-Gln687. The side chains of the residues Lys686 and Gln687 were involved in a specific recognition of guanine bases of the CpG site. The backbone of the residues Ser685 and Lys686 formed contacts with the cytosine bases of the CpG dinucleotides. These specific interactions enable the CXXC domain to discriminate the methylation state of CpG sites, since methylation of any cytosine would cause steric clashes and prevent the formation of the specific protein-DNA contacts. Based on these structural observations, Allen et al. suggested the following mechanistic model: The CXXC domain specifically binds unmethylated CpG sites inducing a conformational change in Dnmt1, such as the acidic CXXC-BAH1 linker repels DNA from the catalytic domain, and, thereby, prevents the *de novo* methylation activity of Dnmt1. This model is in a good agreement with the maintenance function of Dnmt1. Dnmt1 is a highly specific enzyme, it recognizes and preferentially methylates hemimethylated CpG sites. In contrast, Dnmt3a does not discriminate the methylation state of the CpG sites and was shown to methylate also in a non-CpG context (Fatemi et al, 2001; Gowher & Jeltsch, 2001). In order to confirm the model, several Dnmt1 variants were purified and their specificity towards substrate DNA was analysed (Song et al, 2011). The authors were able to demonstrate that the Dnmt1 fragment (650-1602) methylates hemimethylated DNA 75 times faster than unmethylated DNA. A shorter Dnmt1 variant without CXXC domain (amino acid residues 717-1602) had only an elevenfold preference towards the hemimethylated DNA. A similar drop in the specificity was achieved by the mutagenesis of the residues involved in the recognition of unmethylated CpG sites by the CXXC domain. The Dnmt1 (650-1602) K686A/Q687A mutant demonstrated only a tenfold preference for

the hemimethylated substrate. Subsequently, Song and co-workers showed that the CXXC domain increases the specificity of the truncated Dnmt1 variants by disfavouring the unmethylated substrate.

Before the information discussed above became available, we were investigating the influence of the CXXC domain on the specificity of the full length Dnmt1. Since the crystal structure of the Dnmt1 fragment (650-1602) had not yet been published, amino acid residues of the CXXC domain involved in the recognition of unmethylated CpG dinucleotides were not known. We compared sequences of CXXC domains of Dnmt1 proteins of different species with residues involved in the specific DNA recognition by the MLL CXXC domain. Based on this analysis, we proposed several residues in the Dnmt1 CXXC domain to be involved in the DNA binding. We investigated the importance of these residues for the DNA recognition and the contribution of the CXXC-DNA interactions to the substrate specificity of Dnmt1. The results of the study were published (Bashtrykov et al, 2012a) (Appendix 1). They will be discussed in chapter 3.2.1.

The RFTS domain of Dnmt1, responsible for the targeting of the enzyme to the replication foci and interaction with Uhrf1 protein, was proposed as an internal regulatory factor of Dnmt1 activity. In 2011, Takeshita and co-workers published the structure of a Dnmt1 fragment containing residues 291-1620 (Takeshita et al, 2011). This Dnmt1 fragment contained the RFTS, CXXC, BAH1/2, and catalytic domains (**Fig. 4**). The structure of the catalytic domain was identical with the structure obtained by Song et al. (2011). The folding of the BAH1/2 domains and their location relative to the catalytic domain were highly similar in both structures as well. The most striking difference was found at the N-terminus of the crystallized Dnmt1 fragments. In Takeshita's crystal structure, the DNA binding pocket of Dnmt1 is occupied by the RFTS domain, and the CXXC domain is located further away from the catalytic domain. The surface of the RFTS domain involved in the interaction with the catalytic domain has a negative electrostatic potential, mimicking DNA. Additionally, the position of the RFTS domain within the DNA binding pocket is stabilized by several hydrogen bonds and ion pairs with the residues from the catalytic domain, namely E531-K1537, D532-R1576, D554-S1495, A594-H1504, and L593-T1505. Since the RFTS domain blocks the catalytic pocket of Dnmt1, the authors investigated a possible

influence of the RFTS domain on the DNA methylation activity. It was found that the substrate specificity of both Dnmt1 (291-1620) and Dnmt1 (602-1620) lacking the RFTS domain was similar. Also, they found that the activation energy of Dnmt1 (291-1620) with the RFTS domain was three times higher in comparison to Dnmt1 (602-1620). The authors concluded that this extra energy is needed to remove the RFTS domain from the catalytic pocket and to adopt an active conformation of Dnmt1.

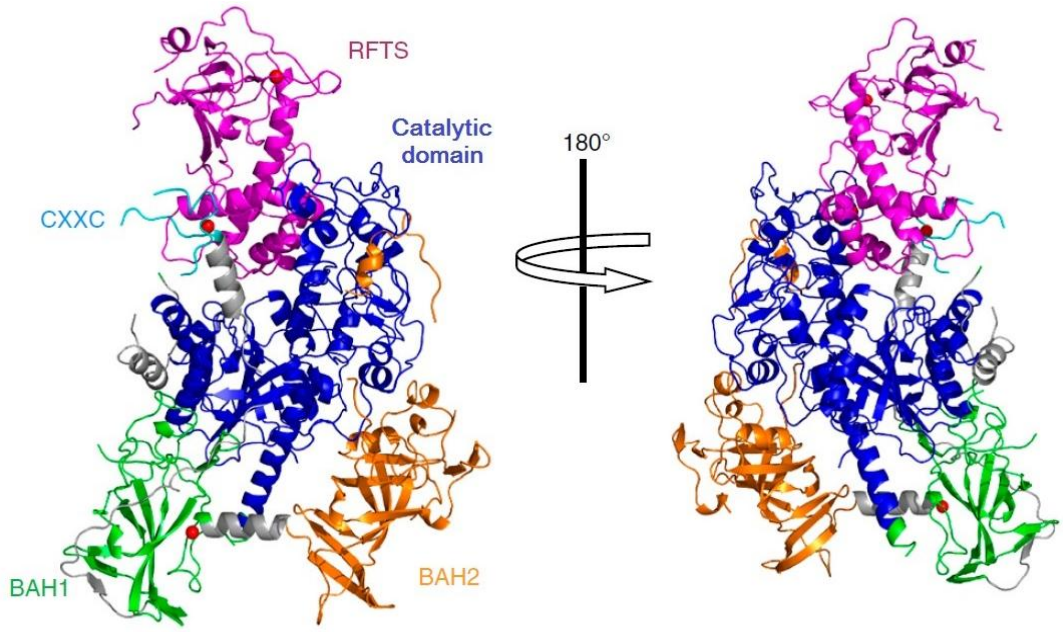

Figure 4. Crystal structure of Dnmt1 (291-1620).
The RFTS domain (magenta) occupies the catalytic pocket (dark blue) of Dnmt1 and prevents its binding with DNA. Adopted from Takeshita et al. (2011).

Independently, another paper investigating the contribution of the RFTS domain to the Dnmt1 DNA methylation was published (Syeda et al, 2011). Two Dnmt1 fragments, one containing the RFTS domain (residues 351-1616) and another without the RFTS domain (residues 651-1616), were purified, and their DNA binding and DNA methylation activities were examined. It was found that the Dnmt1 (651-1616) variant binds a hemimethylated 12 base pairs DNA oligonucleotide much stronger than Dnmt1 (351-1616). Addition of the isolated RFTS domain (351-600) inhibited DNA binding by the Dnmt1 (651-1616) fragment. These results allowed the authors to conclude that the RFTS domain inhibits DNA binding by the catalytic domain of Dnmt1. Furthermore, using the DNA methylation assay with hemimethylated DNA, it was found that the

addition of the RFTS domain to the reaction mixture strongly inhibits the DNA methylation activity of the Dnmt1 (651-1616) variant. This inhibition is competitive for the DNA substrate with a Ki ~100 nM. To identify the structural properties of the RFTS domain underlying such an inhibitory effect, the RFTS domain (residues 351-600) was crystallized, and the obtained structure was solved with 2.3 Å resolution. The authors proposed three acidic loops in the RFTS domain to be involved in the interaction with three basic patches of the Dnmt1 catalytic domain. They proposed an autoinhibitory role of the RFTS domain and speculated that the Dnmt1-binding protein, such as Uhrf1, is required for the removal of the RFTS domain from the DNA binding cleft and the consequent activation of Dnmt1.

We performed a detailed study in order to analyse the autoinhibitory model proposed by both Syeda et al. (2011) and Takeshita et al. (2011). We designed Dnmt1 mutants to have a reduced interaction between the RFTS domain and the catalytic domain. The manuscript describing the results of the comparative analysis of the catalytic activities of those mutants is in preparation (Appendix 5) and it will be discussed in chapter 3.2.2.

As mentioned above, it was shown that Dnmt1 has a high specificity for the methylation of CpG sites (Fatemi et al, 2001) and a preference towards hemimethylated sites (Fatemi et al, 2001; Pradhan et al, 1999). The structural features responsible for the recognition of the CpG substrate and the discrimination of the methyl-cytosine by the catalytic domain of Dnmt1 were identified only in 2012 by Song and colleagues (Song et al, 2012). This group crystallized a second truncated Dnmt1 (731-1602) variant (containing only the BAH1, BAH2, and catalytic domains) in a complex with a 12 base pairs DNA oligonucleotide (**Fig. 5**). The DNA contained one hemimethylated CpG site in the centre with a 5-fluorocytosine (5fCyt) within the CpG dinucleotide of the target strand. This cytosine analogue is known to form a covalent complex with the cysteine of the DNA methyltransferase's active centre, since the final step of the methyltransferase reaction is the deprotonation at position 5 leading to the β-elimination of the cysteine SH group, which is not possible in the case of 5fCyt, in which hydrogen is replaced by fluor (Cheng, 1995). Thus, Dnmt1 (731-1602) formed a productive complex with the DNA in the presence of AdoMet.

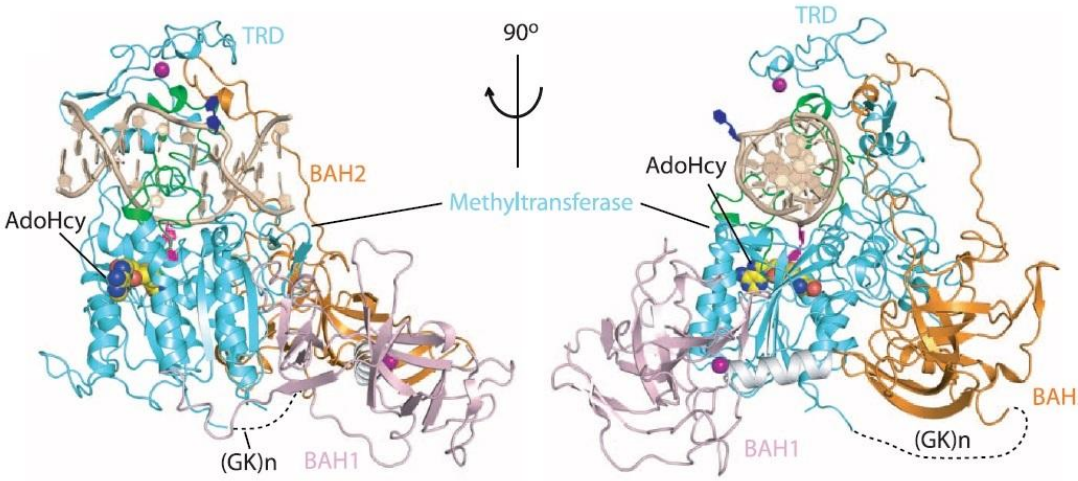

Figure 5. Crystal structure of the Dnmt1 (731-1602) complex with hemimethylated DNA.
Adopted from Song et al. (2012).

The crystal structure revealed that the folding of both BAH domains and the methyltransferase subdomain is similar to the structures discussed earlier (Song et al, 2011; Takeshita et al, 2011). This was the first reported structure of Dnmt1 with the DNA bound in the catalytic cleft. The methyltransferase and TRD subdomains formed many contacts with the DNA. The target 5fCyt was flipped-out of the DNA helix and inserted into the catalytic pocket. The 5fCyt was methylated and covalently bound to the Cys1229 of the catalytic loop. Dnmt1 formed contacts with the 5fCyt via conservative residues found previously in M.*Hha*I.

The basis for the specific recognition of the hemimethylated CpG site was found in the structure as well (**Fig. 6A**). The TRD subdomain formed a hydrophobic surface (residues Cys1501, Leu1502, Trp1512, Leu1515, and Met1535) in the major grove around the methyl group of the 5mCyt. Two loops of the TRD subdomain, penetrating into the major grove, and the catalytic loop, invading the minor grove of the DNA, formed contacts with the DNA. These contacts are involved in the recognition of the CpG dinucleotide and the stabilization of the distorted DNA structure. Recognition of the 5mCyt was taken over by the Arg1237, which contacts the O2 atom of the 5mCyt, and Met1535 forming a hydrogen bond with the N4 atom of the 5mCyt (**Fig. 6B**). Recognition of the Gua of the 5mCyt:Gua base pair was mediated by a hydrogen bond between the backbone amide of Lys1537 and the O6 atom of Gua and a water-mediated hydrogen bond of Gln1538 with the N7 atom of Gua (**Fig. 6B**). The Gua of the non-

target strand formed three hydrogen bonds with Dnmt1: Gly1234 contacts N1, Asn1236 contacts N2, and the side chain of Lys1537 contacts O6 atom of the Gua (**Fig. 6C**). The space in the DNA helix left after the flipping of the target 5fCyt was occupied by Met1235 and Leu1537 residues (**Fig. 6D**). More to that, an unexpected reorganization in the DNA structure was observed at the 5´ flank of the 5fCyt-Gua dinucleotide (**Fig. 6C, D**). The orphan Gua of the non-target strand formed a non-canonical Gua:Gua base pair

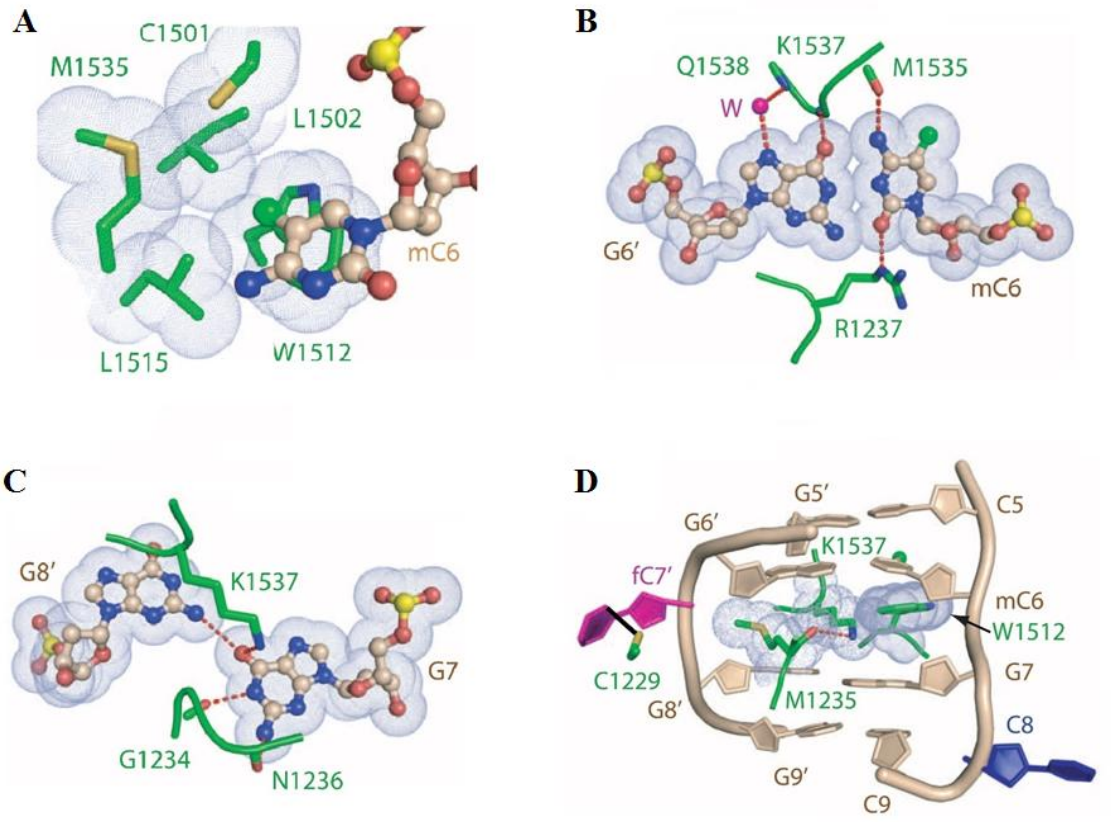

Figure 6. Recognition of the hemimethylated CpG site and structural rearrangement of the DNA in the Dnmt1-DNA complex.
A. The residues of the TRD subdomain form a hydrophobic surface around the methyl group of the 5mCyt. **B.** Recognition of the 5mCyt:Gua base pair by direct and water-mediated (purple W) hydrogen bonds. **C.** Formation of the non-canonical Gua:Gua base pair and recognition of the non-target strand Gua. **D.** Distortion of the DNA structure over the CpG site. The target Cyt (purple, fC7´) and the non-target strand Cyt (blue, C8) are rotated out of the DNA helix. Adopted from Song et al. (2012).

with Gua at the 5´ flank of the target 5fCyt. The now orphaned Cyt of this flanking Gua:Cyt base pair flipped out of the DNA helix away from the target 5fCyt. Furthermore, Song et al. investigated the importance of the individual residues involved in interaction with the DNA for the Dnmt1 specificity. Selected residues, namely Cys1501, Leu1502, Trp1512, Leu1515, and Met1535, were mutagenized, Dnmt1

mutants were purified and assayed *in vitro*. The obtained results showed that all residues forming the hydrophobic surface for the methyl group of the 5mCyt are important for the catalysis as replacement by Ala (Met in the case of Cys1501) leads to a significant decrease of the methylation activity towards both hemimethylated and unmethylated substrates, but they did not dramatically influence the specificity of Dnmt1.

We studied the specific protein-DNA contacts identified in the crystal structure and investigated their importance for the recognition of the CpG dinucleotide and the catalytic activity of Dnmt1. Additionally, the structural rearrangements at the 3´ flank of the 5mCyt-Gua dinucleotide and the corresponding contacts with the protein must be sequence specific, suggesting the possibility of Dnmt1 flanking preferences, which we also analysed. We introduced an approach allowing us to probe individual DNA-protein contacts by using modified DNA substrates without mutagenizing Dnmt1. The obtained results were published (Bashtrykov et al, 2012b) (Appendix 2) and will be discussed in chapter 3.1.

1.3.4 Interaction partners of Dnmt1

More than forty proteins interacting with Dnmt1 have been discovered so far (reviewed in Qin et al., 2011). The effects of interacting partners on Dnmt1 can be divided into two groups. First, local targeting of Dnmt1: A direct interaction of Dnmt1 with transcription factors, such as Specificity protein 1 (SP1), Specificity protein 3 (SP3), and Signal transducer and activator of transcription 3 (STAT3) in complexes with transcriptional regulators was shown in several studies, this interaction contributes to the targeted suppression of certain genes (Esteve et al, 2007; Robertson et al, 2000; Zhang et al, 2005). Interaction of Dnmt1 with chromatin binding proteins like Suppressor of variegation 3-9 homolog 1 (SUV39H1), Euchromatic histone-lysine N-methyltransferase 1 (EHMT1), and Heterochromatin protein 1 (HP1) may recruit Dnmt1 to heterochromatin, to facilitate DNA methylation and ensure a stable repression at the corresponding region (Fuks et al, 2003; Kim et al, 2009; Xu et al, 2011). Second, apart from the discussed local targeting of Dnmt1 to specific genomic regions, the

PCNA and Uhrf1 proteins were shown to facilitate maintenance DNA methylation after DNA replication.

Interaction of Dnmt1 with PCNA

It was found that Dnmt1 is co-localized with replication foci by directly interacting with PCNA (Chuang et al, 1997; Easwaran et al, 2004; Leonhardt et al, 1992). PCNA, also known as a replication clamp, is a component of eukaryotic replication forks responsible for the processivity of DNA replication. This observation led to a model employing the idea that the interaction of Dnmt1 with PCNA connects DNA replication with the remethylation of newly synthesized DNA. However, later experiments showed that the interaction of Dnmt1 with PCNA is transient and not essential for maintenance methylation (Egger et al, 2006; Schermelleh et al, 2007). Another model was suggested assuming that Dnmt1 is recruited by PCNA to a newly replicated DNA. There, Dnmt1 binds to the DNA, and by the linear diffusion it can move along the DNA and find hemimethylated CpG sites (Jeltsch, 2008). Experiments with Dnmt1 mutants lacking interaction with PCNA showed that efficiency of Dnmt1 DNA methylation dropped by twofold, but this interaction was not indispensable (Egger et al, 2006; Spada et al, 2007).

Interaction of Dnmt1 with Uhrf1

Uhrf1, also known as a nuclear protein 95 (Np95), is an important epigenetic player. Genetic studies on mice demonstrated that a knockout of Uhrf1 leads to developmental disorders and early embryonic lethality (Bostick et al, 2007; Sharif et al, 2007). An investigation of mouse *Uhrf1-/-* embryonic stem cells elicited a dramatic decrease of the global DNA methylation level, disregulation of imprinted genes, and retrotransposons. This phenotype was similar to the *Dnmt1-/-* phenotype, though the expression of Dnmt1 was not affected in Uhrf1 knockout cells. Hence, it was speculated that Uhrf1 regulates the maintenance DNA methylation through another mechanism. Indeed, it was found that Uhrf1 interacts with Dnmt1 and recruits it to the replication foci and heterochromatin during DNA replication, thus guiding Dnmt1 to the sites where maintenance DNA methylation is required (Bostick et al, 2007; Sharif et al, 2007).

Similarly as in the case of Dnmt1, the expression of Uhrf1 correlates with cell proliferation. It is not expressed in quiescent cells, but highly abundant in both proliferating normal and cancer cells (Fujimori et al, 1998; Hopfner et al, 2000; Unoki et al, 2004). Its presence is important for the S phase entry of the cells (Bonapace et al, 2002). Uhrf1 itself is a cell cycle-regulated protein. Its expression reaches its maximum during the S phase and declines in the G1 phase (Miura et al, 2001; Uemura et al, 2000), suggesting that the function of Uhrf1 is mainly connected to the S phase. Uhrf1 is a nuclear protein and has a characteristic subnuclear localization. In the early and partially in the mid S phase, Uhrf1 is co-localized with PCNA at the replication foci, the sites of

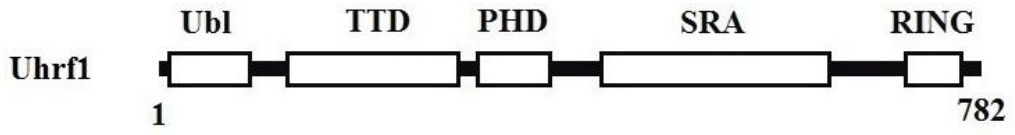

Figure 7. Domain structure of Uhrf1.
Uhrf1 harbours an Ubiquitin-like domain (Ubl), a tandem Tudor domain (TTD), a Plant homeodomain (PHD), a SET and RING-associated (SRA) domain, and a Really Interesting New Gene (RING) domain.

active DNA replication. However, this co-localization with PCNA is transient and in the mid and late S phase Uhrf1 is located at the PCNA-negative heterochromatin (Bostick et al, 2007; Miura et al, 2001). It was shown that Uhrf1 is involved in cancerogenesis by promoting hypermethylation of tumor suppressor genes in primary human cancer cells (Daskalos et al, 2011; Jin et al, 2010). Uhrf1 harbours five defined domains (**Fig. 7**), and at least four of them are involved in the regulation of epigenetic processes, their functions are described in the text below.

An Uhrf1 SET and RING associated (SRA) domain is a DNA binding unit. It was shown that the SRA domain recognizes hemimethylated CpG sites formed during DNA replication (Arita et al, 2008; Avvakumov et al, 2008; Bostick et al, 2007; Hashimoto et al, 2008; Qian et al, 2008). Different structures of the SRA domain co-crystallized with hemimethylated DNA reveal that the SRA domain contacts DNA from both minor and major grooves. Strikingly, recognition of the 5-methylcytosine is accompanied by its flipping out of the DNA helix. The flipped base is anchored within a binding pocket of the SRA, and the position of the orphan Guanine was stabilized by a protein loop filling up the space in the DNA helix left by the 5mCyt. Although base

flipping was known for many enzymes involved in DNA repair and for DNA methyltransferases (Klimasauskas et al, 1994; Tubbs et al, 2007; Yang et al, 2008), it was quite unexpected for the SRA as a reading domain. Based on this observation, it was proposed that Uhrf1 binds hemimethylated CpG sites, recruits Dnmt1, and hands over CpG sites to Dnmt1, thus facilitating the maintenance DNA methylation (Arita et al, 2008). Finally, Achour et al. (2007) found that the Uhrf1 SRA domain is responsible for interaction with Dnmt1. It binds to Dnmt1 amino acid residues 401-615, which correspond to the Dnmt1 RFTS domain. In this thesis, Dnmt1 interaction with the Uhrf1 protein and its SRA domain was further investigated with a focus on their effect on the activity of Dnmt1. A manuscript describing the results has been submitted for publication (Appendix 4), and the results will be discussed in chapter 3.3.

A tandem Tudor domain (TTD) of Uhrf1 recognizes a heterochromatin mark histone 3 lysine 9 dimethylation (H3K9me3) in a combination with unmethylated lysine 4 of histone H3 (H3K4me0). The recognition of H3K9me3 is performed by the first Tudor subdomain. Binding to H3K9me3 is important for heterochromatic localization of Uhrf1 and for the regulation of gene expression (Nady et al, 2011). Interaction of Uhrf1 with H3K9me is required for the maintenance DNA methylation, since the Uhrf1 H3K9me3-binding defective mutant cannot restore the DNA methylation level in Uhrf1 knockdown cells (Rothbart et al, 2012). In addition to that, the interaction of Uhrf1 with H3K9me3 is not sensitive to phosphorylation of serine 10 of histone H3 (H3S10), a mark which appears during the M phase and leads to the dissociation of chromatin bound factors from chromatin during mitosis (Fischle et al, 2005; Hirota et al, 2005), suggesting that Uhrf1 may regulate some processes during the M phase.

An UHRF1 plant homeodomain (PHD) was known to be involved in the reorganization of pericentromeric heterochromatin during the replication of the DNA (Papait et al, 2008). The PHD finger binds the histone H3 tail and recognizes the unmodified arginine residue 2 of histone H3 (H3R2) (Hu et al, 2011; Rajakumara et al, 2011; Wang et al, 2011). The recognition of H3R2 is not required for the localization of Uhrf1 at heterochromatin, but it is important for the regulation of genes located in euchromatin. Recent crystallographic studies revealed that Uhrf1 can bind both H3R2 and H3K9me3 of one histone H3 tail simultaneously by interaction via the PHD and TTD domains (Arita et al, 2012; Cheng et al, 2013; Xie et al, 2012), and the coordinated

recognition of both histone marks is required for maintenance DNA methylation (Rothbart et al, 2013). In addition, Liu et al. (2013) demonstrated that the co-localization of Uhrf1 with heterochromatin and its ability to recruit Dnmt1 and regulate the maintenance DNA methylation required the binding to at least one of two epigenetic signals – to hemimethylated DNA and/or histone H3K9me2/3. This finding is supported by the observation that Uhrf1 is targeted to heterochromatin even at low levels of global DNA methylation (Rottach et al, 2010).

A Really Interesting New Gene (RING) domain is responsible for the E3 ubiquitin ligase activity of Uhrf1 (Citterio et al, 2004). Uhrf1 ubiquitinates Dnmt1, and this results in a proteasome-mediated degradation of the protein (Agoston et al, 2005; Qin et al, 2011b). It was also shown that the RING finger ubiquitinates histone H3 (Citterio et al, 2004; Karagianni et al) (Citterio et al, 2004; Karagianni et al, 2008). Ubiquitination of histones regulates the activity of chromatin, for example histone H2A and histone H2B ubiquitination is a mark of active genes (Muratani & Tansey, 2003). Uhrf1 regulates the expression of the Promyeloic Leukemia (PML) protein, a known tumor suppressor, through ubiquitination-mediated degradation. Uhrf1 is overexpressed in many types of cancer cell and the downregulation of PML by Uhrf1's E3 ligase activity may contribute to tumorogenesis (Guan et al, 2013). Uhrf1 also ubiquitinates the Tat-interacting protein of 60 kDa (TIP60), which leads to the inhibition of TIP60-dependent p53 activation and may contribute to tumorigenesis (Dai et al, 2013). Summing it up, Uhrf1 reads different histone marks, recognizes hemimethylated CpG sites, and it is involved in several epigenetic processes.

2. Aim of the study and summary of results

The maintenance of DNA methylation patterns is an important task mainly fulfilled by the Dnmt1 methyltransferase. A lot of publications within the last several years provided insight into the properties of Dnmt1 that lead to its preferential methylation of hemimethylated CpG sites, which is due to the specific recognition of a single methyl group in the active site of the enzyme. It was one aim of this study to investigate the molecular details of the substrate recognition to shed light on this fascinating and very important process. However, the direct readout of the hemimethylation in the active site of Dnmt1 is not sufficient for the proper replication of DNA methylation patterns, and there might be additional mechanisms regulating the activity, specificity, stability, and targeting of Dnmt1 during the maintenance of DNA methylation. In the present PhD thesis, we investigated the allosteric regulation of Dnmt1 and its interaction with the Uhrf1 protein, which might facilitate Dnmt1 to maintain DNA methylation patterns. Following questions were addressed and investigated.

2.1 Specific DNA recognition by the Dnmt1 catalytic domain

Dnmt1 recognizes and preferentially methylates hemimethylated CpG sites. Recently, the amino acid residues forming base-specific contacts with the CpG site and the methyl group of the 5mCyt, which are responsible for the substrate recognition of Dnmt1, were identified (Song et al, 2012). We tested the importance of the isolated DNA-protein contacts for DNA recognition and catalytic activity of Dnmt1. For this, we used a DNA substrate ("parental" substrate) containing a single CpG site and several near cognate substrates, which differ from it by the replacement of one or two nucleotides within the CpG site or its 5´ flank base pair. By *in vitro* DNA methylation assays we determined the rate of methylation of the "parental" substrate by the purified wild-type Dnmt1 and compared it with the rate of methylation of the near cognate substrates. We observed that the contacts with the 5mCyt:Gua base pair are important for the catalysis, since the replacement of any of these nucleotides resulted in a dramatic reduction of the methylation activity of Dnmt1. Recognition of the methyl group of the

5mCyt increased the activity of Dnmt1 tenfold compared with an unmethylated Cyt. The contacts with the non-target strand Gua are not required for the substrate recognition, since its exchange to Ade increased the activity of Dnmt1 2.5-fold. Finally, we showed that Dnmt1 has no preference for the 5´ flanking sequence of the CpG site, suggesting that some conformational changes of the DNA observed in this region in a crystal structure of a Dnmt1-DNA complex (Kuo et al, 2012) do not occur in solution. The results are described in the manuscript (Bashtrykov et al, 2012b) which is attached to this thesis as Appendix 2.

2.2 Influence of the CXXC domain on the specificity of Dnmt1

The CXXC domain of Dnmt1 binds DNA containing unmethylated CpG sites. It was proposed that the DNA binding of the CXXC domain increases the specificity of Dnmt1 towards the hemimethylated DNA by abolishing the access of unmethylated sites to the catalytic site of the enzyme. Since so far this model was only supported by biochemical data obtained from truncated Dnmt1 variants, we tested it again using the full-length Dnmt1 protein. Using a site-directed mutagenesis, we generated several variants of the isolated CXXC domain and investigated their DNA binding by an electrophoretic mobility shift assay. The amino acid residues involved in the DNA binding of the CXXC domain were identified. Based on these data, we generated full-length Dnmt1 variants lacking the DNA binding of the CXXC domain and compared their substrate specificity with the full-length wild-type Dnmt1. We found that the loss of the DNA binding of the CXXC domain did not decrease the specificity of the enzyme, in fact some Dnmt1 variants showed an even higher preference for the hemimethylated substrate over the unmethylated substrate as compared with the wild-type Dnmt1. In contrast, the exchange of the M1235 residue, which is involved in the recognition of the CpG site by the catalytic domain, led to a significant reduction of the preference of Dnmt1 for the hemimethylated DNA. Thus, the specificity of Dnmt1 towards the hemimethylated CpG sites resides within its catalytic domain. The results are described in the manuscript (Bashtrykov et al, 2012a) which is attached to this thesis as Appendix 1.

2.3 Regulation of the activity of Dnmt1 by conformational changes

Recently, the RFTS domain was proposed to function as an autoinhibitor of Dnmt1's activity. It was shown that in the truncated Dnmt1 (291-1620) variant the RFTS domain locates within the DNA binding cleft preventing the access of a substrate to the catalytic site. This specific position of the RFTS domain is stabilized by several hydrogen bonds to the catalytic domain. To investigate the validity of this model in the full-length Dnmt1, we generated Dnmt1 variants D554R and E572R/D575R, which contain exchanges of the amino acids involved in the formation of these hydrogen bonds, which might therefore destabilize the binding of the RFTS domain to the catalytic domain. The Dnmt1 variants showed a 2.5- and fourfold increase in the activity, respectively, in comparison with the wild-type Dnmt1. Our results strongly support the assumed autoinhibitory role of the RFTS domain in the regulation of Dnmt1's activity. The results are described in the manuscript (Bashtrykov et al., in preparation) which is attached to this thesis as Appendix 5.

2.4 Influence of the Uhrf1 protein on the enzymatic properties of Dnmt1

Uhrf1 is a crucial component of the DNA methylation machinery. It was shown that Uhrf1 recognizes hemimethylated CpG sites via its SRA domain and binds Dnmt1. The direct interaction of Uhrf1 with Dnmt1 results in the targeting of the enzyme to freshly replicated DNA regions containing hemimethylated CpG sites. Uhrf1 is required for the maintenance of the DNA methylation patterns. In this thesis, we show that Uhrf1 and its isolated SRA domain allosterically stimulate the activity of Dnmt1 and increase its specificity to the hemimethylated substrate *in vitro*. The stimulatory effect requires a direct interaction of Uhrf1 with the RFTS domain of Dnmt1 and can be abolished by deletion of the RFTS domain or by the exchange of the residues (E406R/D407R) of the RFTS domain, which are involved in the interaction with Uhrf1. Also, we found that the DNA binding of the SRA domain is not essential for the stimulation of Dnmt1's activity, since the SRA domain variant lacking the DNA binding increases the activity of the enzyme as efficient as the wild-type SRA domain. Finally, we proposed a model

explaining the stimulatory effect of Uhrf1 on Dnmt1's activity by Uhrf1 facilitating the removal of the RFTS domain from the DNA binding cleft of the catalytic domain leading to the transition of Dnmt1 into a catalytically active conformation. The results are described in the manuscript (Bashtrykov et al., in reviewing) which is attached to this thesis as Appendix 4.

2.5 Screening for inhibitors of Dnmt1

Aberrant DNA methylation is one of the key mechanisms underlying the pathogenesis of cancer. Dnmt1 was found to be responsible for the hypermethylation of promoters of tumor suppressor genes in different types of cancer. In this regard, the inhibition of the activity of Dnmt1 might be one of the possible approaches to anti-cancer therapy. We participated in a collaboration for screening for effective and selective inhibitors of DNA methyltransferases and tested compounds using recombinant purified Dnmt1 and *in vitro* DNA methylation assays. The results are described in the manuscript (Ceccaldi et al, 2013) which is attached to this thesis as Appendix 3.

3. Discussion

Multicellular organisms, like mammals, originate from a single cell, a zygote, which during embryogenesis divides and differentiates into many cell types. All these cells contain identical genetic information, but show different phenotypes, which are controlled by epigenetic mechanisms. Phenotypes can be inherited and do not associate with changes in the DNA sequence. Epigenetic signals, namely DNA methylation, modifications of histone tails, histone variants, and non-coding RNAs are heritable and reversible marks, which regulate gene expression and chromatin state. DNA methylation in mammals takes place at position 5 of cytosine, primarily within CpG dinucleotides in a cell-type specific manner. DNA methylation is a paradigm of epigenetic signalling on the basis of our at least partial understanding of the mechanistic details of the inheritance of DNA methylation patterns at a molecular level. DNA methylation patterns are established by *de novo* DNA methyltransferases, which methylate palindromic CpG sites on both DNA strands. DNA replication results in the formation of hemimethylated CpG sites, since the newly synthesized DNA strand contains no methylation marks. In order to preserve and transmit DNA methylation patterns to the next cell generations, hemimethylated CpG sites have to be remethylated. Dnmt1 is the key enzyme in the inheritance of DNA methylation patterns. It is a so called maintenance methyltransferase, which recognizes hemimethylated CpG sites and methylates the unmodified strand. Therefore, the concept of the heritability of DNA methylation patterns is based on the ability of Dnmt1 to discriminate hemimethylated CpG sites from unmethylated ones and preferentially methylate them. Hence, the understanding of the molecular basis of the specificity of Dnmt1 has been a major scientific goal for many years in the field of epigenetics.

3.1 Specific DNA recognition by the Dnmt1 catalytic domain

DNA methyltransferases are highly abundant enzymes. More than one thousand methyltransferases have been sequenced, which recognize and methylate several hundreds of unique DNA sequences (http://www.neb.com/rebase). The majority of these enzymes are prokaryotic DNA methyltransferases appearing as part of a

restriction-modification system responsible for the defence of the host genome against invading DNA, like bacteriophages. Several prokaryotic methyltransferases have been crystallized, and their substrate recognition mechanisms are well understood. Interestingly, among such a variety of prokaryotic enzymes, there is not one methyltransferase that were able to recognize hemimethylated DNA (Jeltsch, 2002). This is of course related to the fact that bacteria typically do not contain and do not propagate methylation patterns, so that an enzyme with a specificity for hemimethylated sites is not needed. From this point of view, the mammalian Dnmt1 (and its homologues in other species) is a unique methyltransferase, which is capable of sequence recognition on two different levels. First, Dnmt1 recognizes the CpG dinucleotide sequence, and second, the methylation state of the CpG site. The ability of Dnmt1 to bind specifically CpG sites and also preferentially methylate hemimethylated CpG sites is the fundamental basis for its function as a maintenance DNA methyltransferase (Holliday & Pugh, 1975; Pradhan et al, 1999; Riggs, 1975). It was shown that the recognition of the DNA sequence by the prokaryotic methyltransferases is performed by domains which correspond to the methyltransferase and TRD subdomains of the catalytic domain of Dnmt1 (Klimasauskas et al, 1994).

The molecular details of the substrate recognition by Dnmt1 became available only in 2012, when Song and co-workers solved the first crystal structure of the Dnmt1 (731-1602) fragment in a complex with a hemimethylated DNA (Song et al, 2012). In the complex, the DNA is bound within the catalytic cleft formed by the methyltransferase and TRD subdomains of the catalytic domain. The crystal structure allowed the determination of the specific protein-DNA contacts involved in the recognition of the nucleotides of the CpG site, of detailed interactions with the methyl group of the 5mCyt, and revealed structural rearrangements of the DNA (**Fig. 8**). The distortion of the DNA structure was identified around the CpG site and at the 5´ flank of the target Cyt-Gua dinucleotide. The target Cyt is rotated out of the DNA helix and embedded within the catalytic pocket of Dnmt1. The flipping of the target cytosine is a conserved mechanism found in all DNA methyltransferases (Cheng, 1995; Cheng & Roberts, 2001; Jeltsch, 2002; Roberts & Cheng, 1998). The Met1235 side chain penetrates into the DNA and occupies the space in the DNA helix which is left after the flipping of the target Cyt. The adjoining space on the parental strand is occupied by the

side chain of Lys1537. The orphan Gua shifts along the DNA helix toward the 3´ end, and the 3´ Cyt rotates out of the DNA helix in a direction opposite to the target Cyt. Hence, double flipping was observed. The translocated orphan Gua of the non-target strand formed a non-canonical Gua:Gua base pair with the Gua at the 5´ flank of the target Cyt.

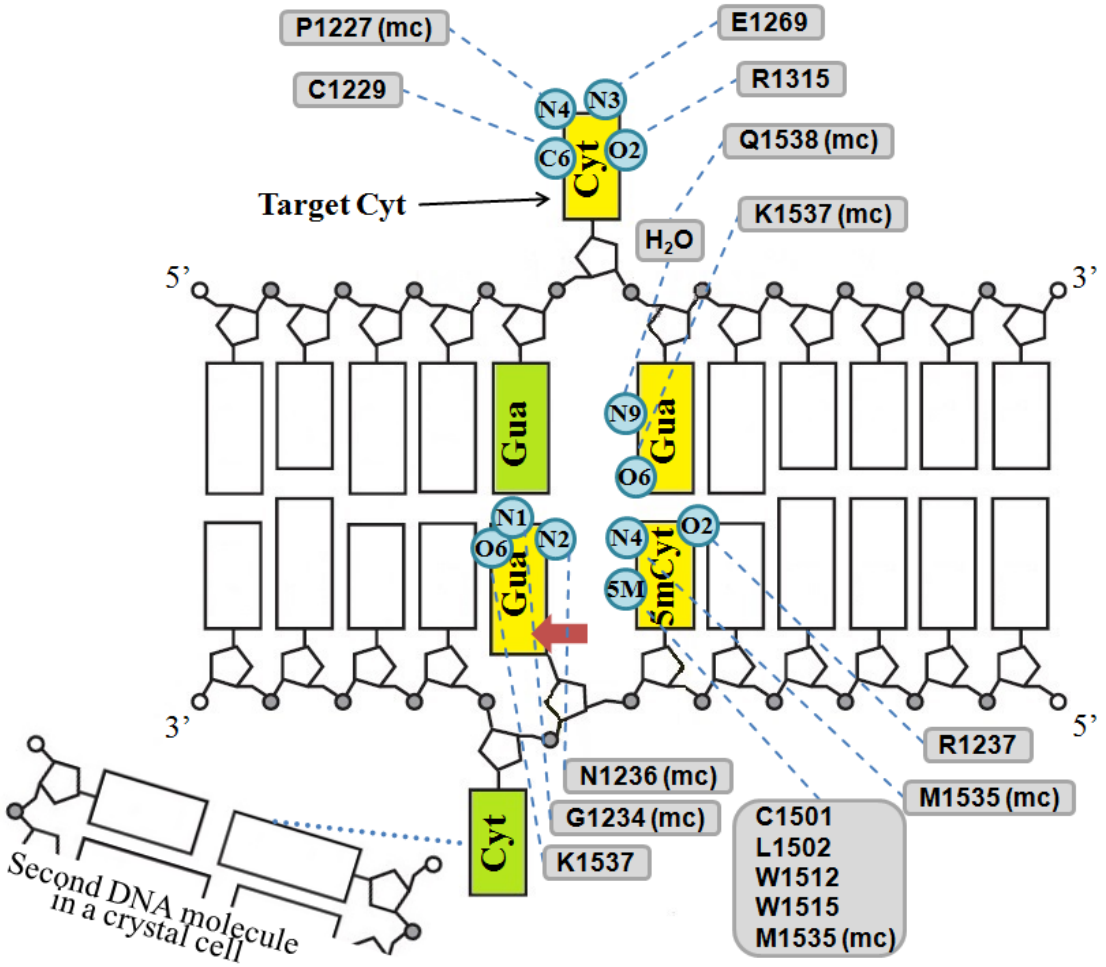

Figure 8. Schematic representation of Dnmt1-DNA contacts and structural rearrangements of the DNA found in the Dnmt1-DNA complex.
Nucleotides of the CpG site are shown in yellow, and the 5´ flank base pair is shown in green. Amino acid residues, of which side chains or main chains (mc) form contacts (blue lines) with nucleotides, are indicated. The red arrow shows the direction of the translocation of the non-target strand Gua. Adopted from Song et al. (2012) with modifications.

Similar, but not identical DNA rearrangements were observed previously in the complexes of prokaryotic methyltransferases with DNA. The smallest distortion of the DNA was found in the M.*Hha*I-DNA complex, where only the flipping of the target Cyt was observed, but the other parts of the DNA remained in the classical B-form. The

interhelical position of the orphan Gua is stabilized by the interaction with the Gln237 side chain occupying the empty space in the DNA left by the flipped Cyt (Klimasauskas et al, 1994). In the bacteriophage T4's DNA adenine methyltransferase (T4Dam)-DNA structure, an analogous interaction was observed for a Thy-Ser112 "base-amino acid" pair. In addition, the Phe111 side chain intercalates the DNA between the orphan Thy and the adjacent Ade:Thy base pair (Horton et al, 2005). The rearrangements in the *Haemophilus aegypticus* methyltransferase III (M.*Hae*III)-DNA complex were more significant and similar to those found in the Dnmt1-DNA complex, but in this case the orphan Gua translocated toward the 5´ end and formed a base pair with the Cyt at the 3´ end of the target Cyt while its partner stayed inside the DNA helix (Reinisch et al, 1995). A double base flipping was found in the *Escherichia coli* DNA adenine methyltransferase (*Eco*Dam)-DNA complex (Horton et al, 2006). In this case, the target Ade and its base pair partner Thy are flipped out, the empty space in the DNA is reduced by a compression. The orphan Thy is flexible and may adopt an intrahelical or extrahelical position. *Thermus aquaticus* methyltransferase I (M.*Taq*I), an N6-adenine DNA methyltransferase, also rotates the target Ade out of the DNA helix. Interestingly, the extrahelical position of the Ade is stabilized by the compression of the DNA at the target base pair resulting in the relocation of the orphan Thy into a central position, in which the Thy would sterically repulse an interhelical Ade (Goedecke et al, 2001). Hence, the distortion of the DNA substrate bound to the enzyme can be found quite often, but the importance of the rearrangements observed in the Dnmt1-DNA complex still needed to be clarified.

We investigated the importance of the individual protein-DNA contacts recognizing CpG dinucleotides and the methyl group of the 5mCyt for the DNA methylation activity of Dnmt1. A well-known way to analyse protein-DNA interactions is the mutagenesis of protein residues followed by the functional analysis of the mutants. This approach allows the testing for individual DNA-residue interactions, but has a strong disadvantage: A decrease of the enzymatic activity may be a consequence of the disruption of an important contact or a change of the folding of the mutant. In order to avoid the possible influence of the mutagenesis on the activity of Dnmt1, we mutagenized the sequence of the DNA substrate used in the methylation reaction. We synthesized a set of 20-mer double-stranded DNA substrates. The parental substrate

contains one central CpG site, the other substrates derived from it harbour specific nucleotide exchanges within the CpG site or its 5´ flank (**Table 1**). Then, methylation rates of these substrates by the full-length wild-type Dnmt1 were determined *in vitro*.

Table 1. List of the substrates used in the DNA methylation assay and methylation activities of Dnmt1.
Sequences of both strands of the double-stranded DNA substrates are given. M – 5-methylcytosine, I – inosine, the CpG sites are shaded grey, mutagenized nucleotides are in red, the target cytosine is in green. Activities are expressed as a ratio of the rate of methylation of the corresponding substrate to the rate of methylation of substrate 1. Values are given as an average of three independent experiments ± standard error of the mean. Taken from Bashtrykov et al. (2012b) with modifications.

Substrate	Sequence	Relative methylation activity
1	GGTAGCAGG**C**GGCCTCCAAG CCATCGTCCGMCGGAGGTTC	1.0
2	GGTAGCAGG**C**GCCCTCCAAG CCATCGTCCG**C**GGGAGGTTC	0.110±0.025
3	GGTAGCAGG**C**GGCCTCCAAG CCATCGTCCG**T**CGGAGGTTC	0.015±0.002
4	GGTAGCAGG**C**IGCCTCCAAG CCATCGTCCGMCGGAGGTTC	0.377±0.019
5	GGTAGCAGG**C**AGCCTCCAAG CCATCGTCCGMCGGAGGTTC	0.014±0.001
6	GGTAGCAGG**C**AGCCTCCAAG CCATCGTCCG**T**CGGAGGTTC	0.002±0.001
7	GGTAGCAGG**C**GGCCTCCAAG CCATCGTCC**A**MCGGAGGTTC	2.37±0.13
8	GGTAGCAG**C**CGGCCTCCAAG CCATCGTC**G**GMCGGAGGTTC	1.22±0.11
9	GGTAGCAG**A**CGGCCTCCAAG CCATCGTC**T**GMCGGAGGTTC	1.21±0.03
10	GGTAGCAG**T**CGGCCTCCAAG CCATCGTC**A**GMCGGAGGTTC	1.34±0.10
11	GGTAGCAG**C**CGGCCTCCAAG CCATCGTC**G**G**C**CGGAGGTTC	0.103±0.016
12	GGTAGCAG**A**CGTCCTCCAAG CCATCGTC**T**G**C**AGGAGGTTC	0.112±0.020
13	GGTAGCAG**T**CGACCTCCAAG CCATCGTC**A**G**C**TGGAGGTTC	0.098±0.015

First, we investigated the importance of the methyl group of the 5mCyt and nucleotides within the hemimethylated CpG site for the Dnmt1 activity. The replacement of the 5mCyt by Cyt (**Fig. 9A, B**), leading to the formation of an unmethylated CpG site, resulted in a tenfold reduction of the Dnmt1 activity (substrates 1 and 2). This result is consistent with published data and demonstrates that the

hydrophobic surface around the methyl group of the 5mCyt, formed by residues Cys1501, Leu1502, Trp1512, Leu1515, and Met1535, contributes to the specificity of the enzyme. The exchange of 5mCyt by Thy (substrate 3), which contains a methyl group at position 5 as the 5mCyt (**Fig. 9C**), led to an approximately 60-fold drop in DNA methylation. Such a strong reduction in the activity despite the presence of the methyl group reflects the importance of the hydrogen bond between the atom N4 of 5mCyt and the main-chain carbonyl oxygen of Met1535, which discriminates 5mCyt from Thy. The disruption of this contact may also break the hydrophobic surface recognizing the methyl group, since Met1535 is involved in the formation of it. In the future, it might prove interesting to test another substrate, the replacement of 5mCyt by Ura, which differs from Thy by the absence of the 5-methyl group. It would be interesting to check if the substitution of 5mCyt to Ura leads to a stronger decrease in the methylation activity (in comparison to 5mCyt to Thy replacement) due to inability to form a hydrophobic surface around the methyl group.

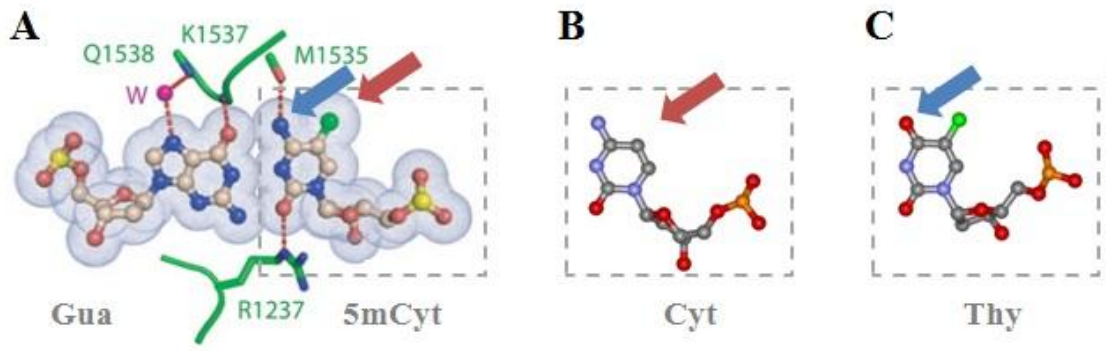

Figure 9. Recognition of the 5mCyt by Dnmt1.
A. The 5mCyt:Gua base pair and base-specific contacts (direct and water-mediated [purple W]) are shown (taken from Song et al. [2012] with modifications). The importance of the contacts with 5mCyt was tested on substrates containing either Cyt (**B**) or Thy (**C**) at this position. The methyl group at position 5 of the 5mCyt (green ball), indicated by the red arrow, is absent in Cyt. The amino group at position 4 of the 5mCyt, indicated by the blue arrow, is replaced by a keto group in Thy.

The substrates with the replacement of the target strand Gua by Ino or Ade (substrates 4 and 5) were methylated about 2.5 and 70 times slower, respectively, than the original substrate. This difference can be explained by the specific contact of Lys1537 to the atom O6 of Gua. This oxygen is present in Ino and absent in Ade (**Fig. 10**). Finally, changing both nucleotides of the 5mCyt:Gua pair to Thy:Ade (substrate 6)

completely blocked the activity of Dnmt1. Our results demonstrated that the methylation of a cytosine within a non-CpG context, as found in ES cells, cannot be introduced and propagated by Dnmt1, and support the studies emphasizing the importance of the Dnmt3s for the non-CpG methylation (Lister et al, 2009; Ramsahoye et al, 2000; Ziller et al, 2011).

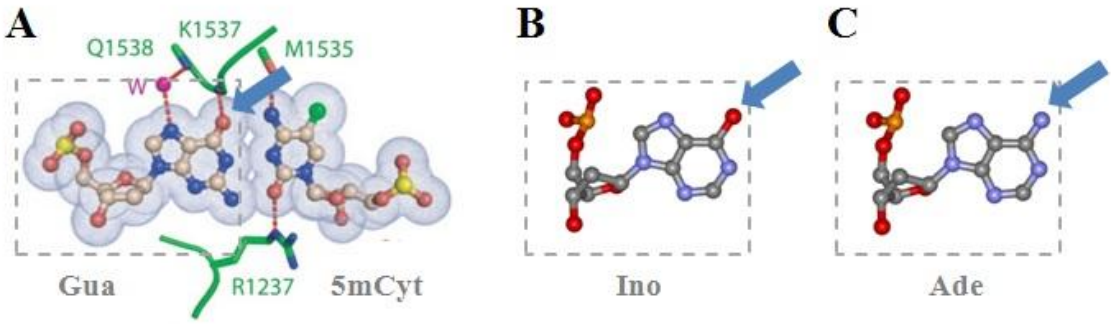

Figure 10. Recognition of the target strand Gua by Dnmt1.
A. The 5mCyt:Gua base pair and base-specific contacts (direct and water-mediated [purple W]) are shown (taken from Song et al. [2012] with modifications). The importance of the recognition of the Gua was tested on substrates containing either Ino (**B**) or Ade (**C**) at this position. The O6 atom of the Gua, indicated by the blue arrow, is present in Ino and replaced by an amino group in Ade.

In order to test the significance of the contacts with the non-target strand Gua, it was exchanged with Ade (substrate 7), which resulted in 2.5-fold stimulation of Dnmt1 activity. These surprising data indicate that the contacts with the non-target strand Gua found in the crystal structure are not important for the catalysis (**Fig. 11**).

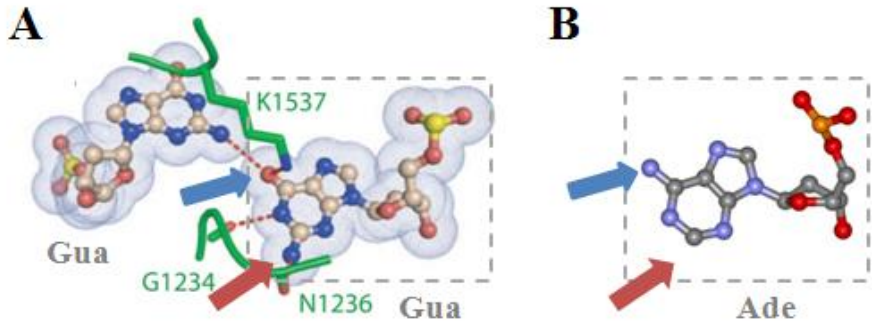

Figure 11. Recognition of the non-target strand Gua by Dnmt1.
A. The non-canonical Gua:Gua base pair is shown (taken from Song et al. [2012] with modifications). The recognition of the non-target strand Gua is performed by specific contacts with the O6 atom (blue arrow) and the N2 atom (red arrow). The importance of these contacts was tested on the substrate containing Ade (**B**) at this position, in which the O6 atom and the N2 atom are absent.

The increase of the methylation activity can be explained by the fact that this Ade forms a non-canonical base pair with the target Cyt, which has to be flipped out of the DNA helix, which is easier in the case of the weaker Cyt:Ade pair. A similar increase of the methylation activity of substrates containing the target nucleotide in a mismatch base pair has already been observed in other cases. For example, a substrate with a Cyt:Thy mismatch in the recognition sequence (G<u>C</u>TATC/GATA<u>T</u>C) was methylated by M.*Eco*RV (DNA-(adenine-N^6)-methyltransferase from *Escherichia coli)* much faster than the substrate containing Cyt:Gua pair (G<u>C</u>TATC/GATA<u>G</u>C) (Roth & Jeltsch, 2001).

It is known that apart from their strict specificity for recognition sequences, enzymes may also have a preference for certain flanking sequences outside of their recognition motif. For example, the Dnmt3a and Dnmt3b *de novo* methyltransferases methylate cytosine within the CG recognition sequence, but prefer purine bases at the 5´-end and pyrimidine bases at the 3´-end of the CpG site (Handa & Jeltsch, 2005; Lin et al, 2002). Recently, the flanking preference of Dnmt3a was refined: Thy is preferred over purine bases at position -2, Ade over Gua at -1, a pyrimidine base at +1, and Ade and Thy over Gua at +3 (Jurkowska et al, 2011c). The flanking sequence preference of Dnmt1 has not been investigated in detail. According to the crystal structure (Song et al, 2012), no base-specific contacts were found within the 3´ flank of the CpG site. In contrast, the structural rearrangements at the 5´ flank of the CpG dinucleotide must be sensitive to the sequence of the 5´ flank nucleotide, which forms the non-canonical base pair with the non-target strand Gua (**Fig. 8**, page 39). To test this hypothesis, we synthesized substrates with different 5´ flank pair (**Table 1**, substrates 1, 2, 8-13) and determined their methylation rate. However, the DNA methylation assays showed no significant difference in the Dnmt1 activity between substrates with different 5´ flanking base pairs (neither in hemimethylated nor unmethylated state). Hence, biochemical data do not support the rearrangements of the DNA found in the crystal structure. Additionally, a detailed investigation of the crystal structure revealed that the flipped out Cyt of the 5´ flank pair does not form contacts with the protein. Instead, the Cyt is stacking and forms a hydrogen bond with the DNA from the nearest crystal cell (**Fig. 12**). The formation of this contact is not possible in the solution, thus these structural rearrangements are more likely crystallographic artefacts. The absence of a

flanking preference of Dnmt1 perfectly coincides with the function of the enzyme. Dnmt1 as a maintenance methyltransferase should remethylate all hemimethylated CpG sites generated during DNA replication, regardless of their flanking sequences. Interestingly, it was found earlier that Dnmt1 has a preference for CCGG sequences during its *de novo* activity, this site was methylated ten times more often than other sites analysed in that study (Goyal et al, 2006). The discrepancy between those findings and our results can be explained by the fact that we have investigated the flanking preference only at the 5´ flank of the CpG site, and the 3´ flank has not been tested. The selection of the 5´ flank for our studies was made upon observations in the crystal structure of the Dnmt1-DNA complex, where structural rearrangements in the DNA were observed at the 5´ flank of the CpG site (Song et al, 2012). Hence, more detailed follow-up studies including both flanks are required to answer the question of whether Dnmt1 has flanking preferences or not.

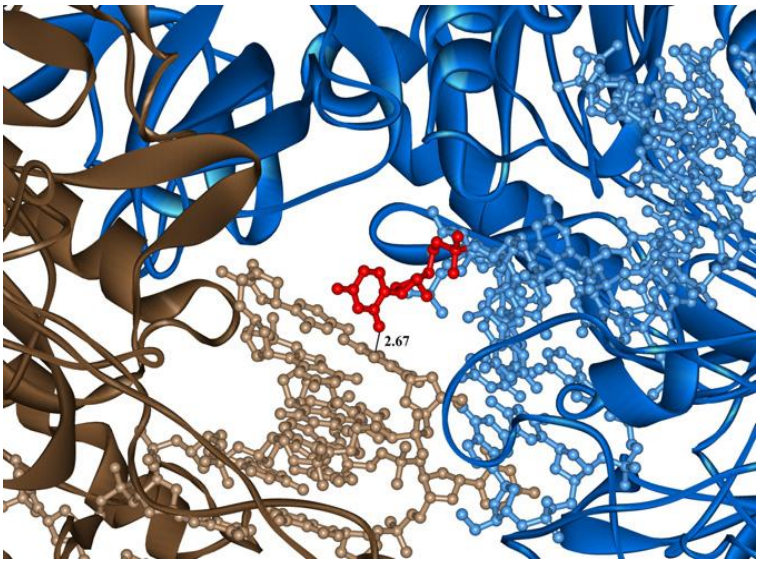

Figure 12. Enlargement of the interface between two Dnmt1 molecules in the structure of the Dnmt1-DNA complex.
The non-target strand Cyt (red) of one Dnmt1-DNA complex (blue) is flipped out of the DNA helix and forms a hydrogen bond and stacking interaction to the end of the DNA of a second Dnmt1-DNA complex (brown) present in the crystal cell. Adopted from Bashtrykov et al. (2012b).

In summary, all these data demonstrate that Dnmt1 has a strong specificity for the CpG dinucleotide, and it is unlikely that Dnmt1 is involved in the non-CpG methylation. The specific protein-DNA contacts discriminating the bases of the 5mCyt:Gua base pair are required for the DNA methylation activity. In contrast, the

contacts to the non-target strand Gua are not important for the activity of Dnmt1. The formation of the hydrophobic surface around the methyl group of the 5mCyt, responsible for the distinction between the 5mCyt and Cyt, increases the activity of Dnmt1 and underlines the preference of the enzyme for the hemimethylated DNA. Finally, we showed that Dnmt1 has no flanking sequence preference at least at the 5´ flank of the CpG site. This result suggests that the rearrangements of the DNA structure observed in the Dnmt1-DNA complex do not have a functional significance and most probably do not occur in solution.

3.2 Allosteric regulation of Dnmt1 by CXXC and RFTS domains

Prokaryotic C5-cytosine methyltransferases are small enzymes (for example, M.*Hha*I consists of 327 amino acid residues) and have homology with the catalytic domain of eukaryotic methyltransferases (Cheng, 1995). As described in the previous chapter, the catalytic domain of Dnmt1 contains conserved amino acid motifs required for catalysis and forms sequence-specific contacts with the DNA, essential for the recognition of the CpG site and its hemimethylated state. Surprisingly, despite the presence of all elements needed for catalysis and specificity, the isolated catalytic domain of Dnmt1 is enzymatically inactive (Bacolla et al, 2001; Fatemi et al, 2001; Margot et al, 2000; Zimmermann et al, 1997). Dnmt1 has an additional N-terminal part harbouring approximately 1100 amino acid residues, which is not present in prokaryotic methyltransferases. Several experiments demonstrated that the N-terminal part of Dnmt1 participates in the regulation of the Dnmt1 activity and serves as a platform for the interaction with other proteins (Jurkowska et al, 2011b). It was shown that the N-terminus has at least two DNA binding regions and directly interacts with the catalytic domain (Araujo et al, 2001; Fatemi et al, 2001; Margot et al, 2003). However we did not have a clear understanding of how the N-terminus regulates the activity of Dnmt1. Two models suggesting regulatory mechanism of its CXXC and RFTS domains have been proposed and experimentally investigated in this study.

3.2.1 Influence of the CXXC domain on the specificity of Dnmt1

The CXXC domain (amino acid residues 650 – 699) is a DNA binding module localized approximately in the middle of the N-terminal part of Dnmt1 (Fatemi et al, 2001). Homologues of the CXXC domain were found in many proteins interacting with DNA and histones (Ayton et al, 2004; Birke et al, 2002; Jorgensen et al, 2004; Lee et al, 2001). It was found that the CXXC domain specifically binds unmethylated CpG sites, though the primary function of Dnmt1 requires recognition of hemimethylated DNA (Frauer et al, 2011; Pradhan et al, 2008). This surprising observation formed a basis for investigations of the possible role of the CXXC domain in the regulation of the activity and specificity of Dnmt1. Before the beginning of our study, it was reported that the CXXC domain would be indispensable for the activity of Dnmt1, since the deletion of

the CXXC domain or the point mutation of one of the conserved cysteines results in a dramatic decrease of the Dnmt1 enzymatic activity (Pradhan et al, 2008). In contrast, in another study no contribution of the CXXC domain for the Dnmt1 methylation activity, DNA binding, or preference for hemimethylated substrate was observed (Frauer et al, 2011). Therefore, apart from the specific recognition of unmethylated CpG sites, the function(s) of the CXXC domain of Dnmt1 was unknown.

We investigated the possible influence of DNA binding to the CXXC domain on the methylation activity and specificity of Dnmt1. In order to test this, we planned to generate Dnmt1 variants with mutations in the CXXC to disrupt its DNA binding and then compare the specificity and activity of the Dnmt1 CXXC mutants with the Dnmt1 wild-type protein. At the time of the beginning of the project, no crystal structure of the Dnmt1 CXXC domain was available and, thus, the residues involved in the formation of protein-DNA contacts were not yet identified. We made an alignment of Dnmt1 CXXC domains from several animal species and compared it with the CXXC domain of the Mixed-Lineage Leukemia (MLL) protein, for which the residues forming contacts with DNA had already been determined (Allen et al, 2006). Based on this analysis, we selected two conserved regions of basic residues, namely R652, K653, K654, K655, and K691, R692, R693, as candidates for mutagenesis, since patches of basic residues are involved in electrostatic interaction with DNA. Additionally, in the MLL CXXC domain, a loop containing residues 1182 – 1188, corresponding to R648 and K686 of the Dnmt1 CXXC, mediates specific contacts with DNA. In addition, to study the importance of the DNA recognition by the catalytic domain, we performed a modelling and selected M1235 as a residue, since it might be involved in the formation of a hydrophobic pocket for the methyl group of a 5mCyt.

At that time, the first crystal structure of the truncated Dnmt1 containing the CXXC domain in a complex with unmethylated DNA was solved (Song et al, 2011). It was observed that the DNA is bound to the CXXC domain, which together with the CXXC-BAH1 linker occupied the catalytic cleft preventing the access of the substrate to the catalytic domain. Based on these structural findings, an autoinhibitory model was proposed, in which the CXXC domain was postulated to increase the specificity of Dnmt1 for hemimethylated CpG sites by sequestering unmethylated CpGs from the catalytic domain. The model was supported by biochemical data demonstrating that the

deletion of the CXXC domain as well as a mutagenesis of the key residues (K686 and Q687) required for the CXXC domain DNA binding results in a decrease of the specificity towards hemimethylated CpG sites. However, these results were obtained in the context of the truncated Dnmt1 variants lacking 650 residues from the N-terminus, consequently, it remained unclear whether the model is valid for the full-length protein.

To test the involvement of the selected residues for the DNA binding of the CXXC domain, five mutants of the isolated CXXC domain were prepared: 4S (R652S, K653S, K654S, K655S), 3S (K691S, R692S, R693S), R684S, K686S, and Q687A (the last variant was added based on the crystal structure data). The DNA binding of the CXXC wild-type and five obtained mutants were tested by two methods: an electrophoretic mobility shift assay and the nitrocellulose filter binding with the DNA substrate containing one unmethylated CpG site. All mutants except Q687A showed a strong reduction of the DNA binding demonstrated by both methods (**Fig. 13**). The finding of residues responsible for the DNA binding to the CXXC domain allowed us to study the influence of the DNA binding of the CXXC domain on the activity and specificity of Dnmt1.

In order to investigate possible influences of the DNA binding of the CXXC domain, we generated the 4S, 3S, R684S, and K686S mutants designed by us and the K686A/Q687A mutant used by Song et al. (2011) in the context of the full-length

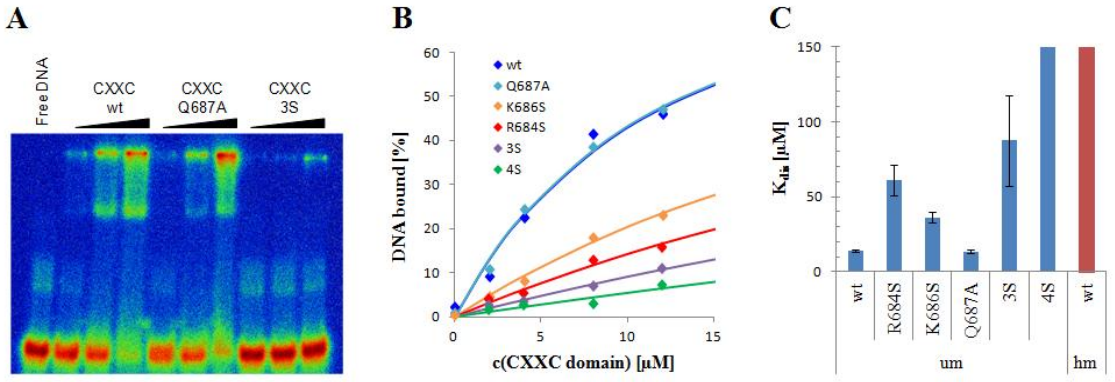

Figure 13. DNA binding of the CXXC domain wild-type and mutants.
A. Example of the gel obtained in the gel shift experiments with the CXXC domain wild-type (wt) and mutants. **B.** Example of the quantification obtained from the nitrocellulose filter binding experiments. **C.** Binding constants of the CXXC domain variants derived from the quantitative analysis of the nitrocellulose filter binding experiments with unmethylated (um) and hemimethylated (hm) DNA. Error bars indicate the standard error of the mean of the averages. Adopted from Bashtrykov et al. (2012a) with modifications.

Dnmt1. Only the 4S, R684S, and K686A/Q687A variants were purified, and their substrate specificity was determined. Surprisingly, the 4S and R684S mutants demonstrated a similar preference for the hemimethylated substrate over the unmethylated one as does the wild-type Dnmt1 (**Fig. 14A**). The specificity of the K686A/Q687A variant was even higher, approximately 17-fold. These results were confirmed by an additional method, in which the specificity of the Dnmt1 variants was determined with a substrate containing two CpG sites, one of which is hemimethylated and one of which is unmethylated. The specificity of the wild-type Dnmt1 was approximately 60-fold, and none of the mutants demonstrated a significant reduction of the specificity. These results obtained by us did not agree with the findings made by Song and co-workers (2011), who demonstrated a strong reduction of the specificity of Dnmt1 variants lacking either the CXXC domain or its DNA binding. Thus, the autoinhibition model proposed and promoted by Song et al. (2011) for the truncated Dnmt1 variants does not appear to be valid for the full-length protein, at least under our experimental conditions. Also, we observed that the 4S and K686A/Q687A mutants (in agreement with the Song et al. [2011] data) are more active than the Dnmt1 wild-type (**Fig. 14B**). This can be explained by the fact that the CXXC domain is an additional DNA binding site in the Dnmt1 protein, and its removal increases the rate of the DNA binding by the catalytic domain. Finally, we purified the Dnmt1 M1235S variant and

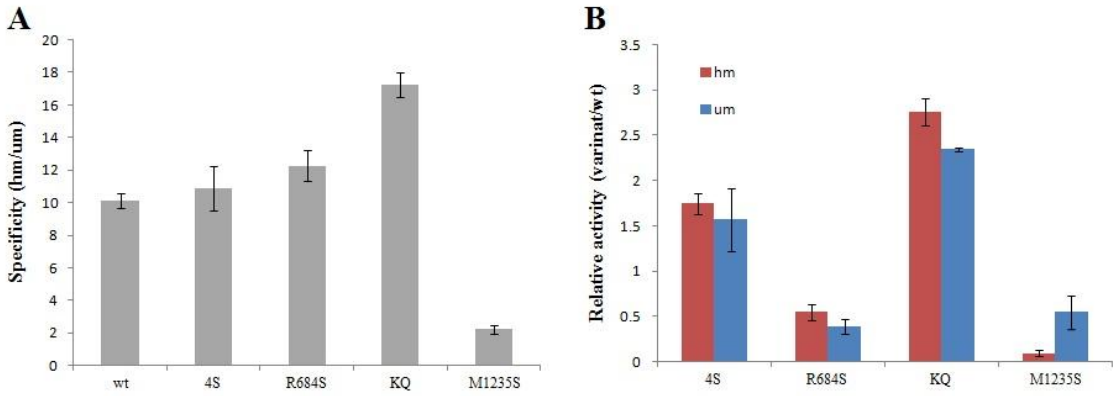

Figure 14. Substrate specificities and catalytic activities of Dnmt1 variants.
A. The substrate specificities of Dnmt1 variants were determined as a ratio of the rate of methylation of the hemimethylated substrate to the rate of methylation of the unmethylated substrate. **B.** Catalytic activities of the Dnmt1 mutants for methylation of the hemimethylated and unmethylated substrates are shown in relation to the activity of the wild-type Dnmt1 (variant/wt). Specificities and activities were averaged over three to six independent experiments. Error bars indicate the standard error of the mean. Adopted from Bashtrykov et al. (2012a) with modifications.

determined its activity and specificity. It was observed that the mutagenesis within the catalytic domain resulted in a tenfold reduction of the methylation activity of the hemimethylated substrate and an only twofold reduction of the methylation of the unmethylated substrate by M1235S mutant (**Fig. 14B**). This led to a 2.2-fold decrease of the specificity (**Fig. 14A**). Later, the crystal structure of the Dnmt1 (731-1602) variant was solved and published. It was found that the M1235 variant penetrates the DNA helix and contacts the hemimethylated CpG site (Song et al, 2012). Therefore, we demonstrated that the specificity of Dnmt1 for the hemimethylated DNA mainly resides in the catalytic domain.

It is an outstanding finding that the maintenance methyltransferase, having a preference for hemimethylated DNA, harbours a domain able to bind unmethylated CpG sites. A possible explanation for this finding could be that the CXXC domain contributes to the specificity of Dnmt1 by playing the role of a filter and preventing methylation of unmethylated CpG sites by binding to the undesired substrate. We could not confirm this hypothesis *in vitro* comparing the preferences of the Dnmt1 wild-type and its variants lacking the CXXC domain DNA binding for methylation of hemimethylated and unmethylated substrates. In another experiment, in which the substrate contained both type of sites (unmethylated and hemimethylated CpGs) and the enzyme was therefore able to "choose" between them, the Dnmt1 wild-type type demonstrated an even higher preference for the hemimethylated CpG site than in the previous experiment. These results may better reflect the situation in a cell. Again, in the Dnmt1 variants lacking the CXXC domain, the DNA binding showed a specificity similar to that obtained for the wild-type enzyme. *In vivo* studies demonstrated that a Dnmt1 mutant with a deletion of the CXXC domain rescued the DNA methylation pattern in D*nmt1-/-* embryonic stem cells as efficiently as the wild-type Dnmt1 (Frauer et al, 2011). Thus, these data also contradict the role of the CXXC domain in restraining Dnmt1 methylation activity on unmethylated CpG sites. Though the results obtained by us and the other study mentioned above disagree with the model suggested by Song et al. (2011), it is difficult to believe that all the intricate structural details observed in the Dnmt1-DNA complex, in which the DNA was bound to the CXXC domain, are merely artefacts. Therefore, we propose a modified model, which can explain this difference.

This model will be discussed in chapter 3.2.2 and investigated in future follow-up studies.

Another possible role of the CXXC domain in the regulation of Dnmt1 specificity should be discussed with regard to the *de novo* activity of Dnmt1. It was shown that Dnmt1 is responsible for the hypermethylation of promoter CpG islands in cancer cells (Feltus et al, 2003; Jair et al, 2006). Though Dnmt1 has a strong preference (ten- to 40-fold) for the hemimethylated DNA, it is more active at the unmethylated CpG sites then Dnmt3a/3b and can be used as a *de novo* methyltransferase under certain conditions. Dnmt1 can be recruited to unmethylated promoters by interacting partners leading to methylation of CpG islands and gene silencing (Hervouet et al, 2010; Jin et al, 2010). In this case, the binding of unmethylated CpG sites by the CXXC domain might increase the residence time of Dnmt1 at unmethylated CpG islands and enforce their methylation. Evidence for such a dual role of Dnmt1 is accumulating, which suggests that a tight control of the Dnmt1 activity is required to provide a precise DNA methylation only at desired genomic regions.

Taken together, our results demonstrate that the preference of Dnmt1 for hemimethylated DNA is controlled mainly by its catalytic domain. An autoinhibitory model proposing an important role of the CXXC domain in the regulation of Dnmt1's specificity and found in truncated Dnmt1 variants (Song et al, 2012) was not confirmed by us in the context of the full-length Dnmt1, which is in agreement with another study that could not show any influence of the CXXC domain on the specificity of Dnmt1, too (Frauer et al, 2011). However, it is possible that the effect observed by Song et al. (2011) is due to the flexibility of the CXXC domain in the truncated Dnmt1, which can also be realized *in vivo* upon an interaction with other proteins resulting in a stronger preference for the hemimethylated CpG sites than determined *in vitro*.

3.2.2 Regulation of the activity of Dnmt1 by conformational changes

The RFTS, a domain responsible for the localization of Dnmt1 at the replication forks (Leonhardt et al, 1992), was proposed as an autoinhibitor of Dnmt1 activity. This idea was based on the crystal structure of the Dnmt1 (291-1620) fragment, which showed that the DNA binding cleft of the catalytic domain is occupied by the RFTS domain (Takeshita et al, 2011). The position of the RFTS domain in this structure is

stabilized by the electrostatic interaction between the negatively charged surface of the RFTS and the positively charged surface of the catalytic cleft as well as by several hydrogen bonds. In this conformation, the catalytic domain cannot bind DNA. Based on this it was postulated that the displacement of the RFTS domain from its position, which is necessary to change Dnmt1 to a methylation potent conformation, would require energy, thus should be a mechanism regulating this process (Takeshita et al, 2011). A direct biochemical confirmation of the RFTS domain inhibition of the Dnmt1 activity was provided by Syeda et al. (2011) who showed that a Dnmt1 (621-1616) fragment lacking the RFTS domain binds DNA and polynucleosomes stronger than the Dnmt1 (351-1616) fragment containing the RFTS domain. The DNA methylation activity of the Dnmt1 (351-1616) variant was weaker compared to the Dnmt1 (621-1616) variant. Furthermore, the DNA binding and DNA methylation of the Dnmt1 (621-1616) variant was inhibited by the addition of the isolated RFTS domain. The authors concluded that the RFTS domain is a concurrent inhibitor of Dnmt1 enzymatic activity and proposed that an interaction with an additional factor, like Uhrf1 protein, might release the RFTS domain from the DNA binding cleft and activate Dnmt1.

We were interested in investigating the inhibitory effect of the RFTS domain. The inhibition of the Dnmt1 methylation activity is based on the specific position of the RFTS domain within the catalytic cleft of the enzyme. The removal of the RFTS domain opens the access for the DNA to the catalytic site. Weakening of the RFTS-catalytic domain interaction should decrease the energy required for the transition of Dnmt1 to the open state and stimulate DNA methylation. To test this hypothesis, we designed Dnmt1 variants in which the hydrogen bonds involved in the RFTS-catalytic domain interaction were destroyed by mutagenesis. Again, we based our experiments on the full-length protein, since the experience with the CXXC domain studies and different results obtained by Song et al. (2011) and by us using a truncated and full-length Dnmt1 argued against a domain deletion approach. Three regions in the RFTS domain were subjected to mutagenesis leading to the generation of Dnmt1 (E531R/D532R), Dnmt1 (D554R), and Dnmt1 (E572R/D575R) variants (**Fig. 15**). In all cases, the acidic residues of the RFTS domain were replaced by arginines, which should not only break the hydrogen bonds, but also weaken the electrostatic interaction with the catalytic domain. Finally, we purified only the Dnmt1 (D554R) and Dnmt1

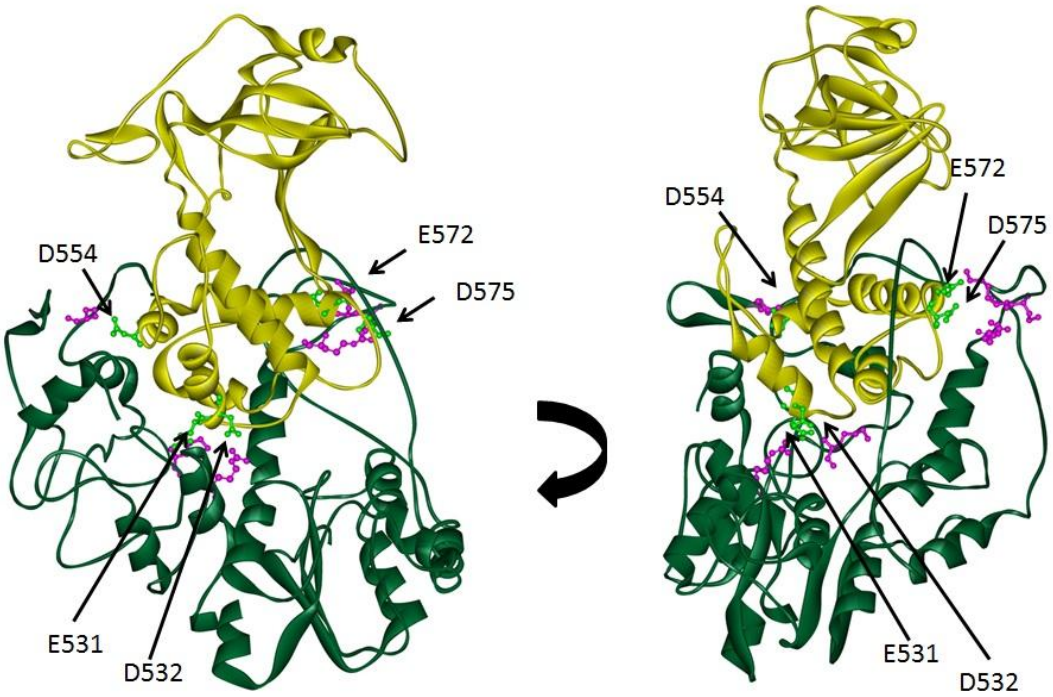

Figure 15. The RFTS domain binds to the catalytic domain of Dnmt1 via hydrogen bonds.
The RFTS domain (yellow ribbon) and parts of the CXXC and catalytic domains (green ribbon) are shown. The position of the RFTS domain within the DNA-binding pocket is stabilized by several hydrogen bonds, formed by residues of the RFTS domain (light green ball-and-stick representation) and residues of the CXXC and catalytic domains (purple ball-and-stick representation). Residues of the RFTS domain selected for mutagenesis are indicated by arrows. The crystal structure was taken from Takeshita et al. (2011). Adopted from Bashtrykov et al. (manuscript is in preparation, Appendix 5).

(E572R/D575R) mutants, which were investigated *in vitro* for their DNA methylation activity. We found that the D554R and E572R/D575R mutants methylated a hemimethylated substrate 2.5- and fourfold faster, respectively, than the Dnmt1 wild-type. The comparison of the specificity of the proteins revealed that both mutants have a 16-fold preference for the hemimethylated over unmethylated DNA, which equals the specificity of the Dnmt1 wild-type obtained under these conditions. These results demonstrate that the destabilization of the RFTS domain binding to the catalytic cleft stimulates the activity of Dnmt1 and strongly support the model proposed by both Syeda et al. (2011) and Takeshita et al. (2011).

Autoinhibiton is a widespread phenomenon that negatively regulates protein functions via intramolecular interactions and conformational changes. It may inhibit ligand binding, subcellular localization, or enzymatic activity of proteins. The function may be implemented by post-translation modifications (PTMs), cleavage of the

inhibitory domain, or by interaction with other proteins or molecules (Pufall & Graves, 2002). For example, phosphorylation of Nuclear Factor kappa B (NF-kB) on serine 276 prevents an interaction between N- and C-terminal regions and stimulates its transcriptional activity (Zhong et al, 1998). The Protein Kinase A forms an inactive heterotetrameric complex consisting of regulatory and catalytic subunits. The binding of cAMP to the regulatory subunits results in its conformational changes, dissociation from the catalytic subunit, and activation of the enzyme (Boettcher et al, 2011; Kim et al, 2006). The transcription factor Nuclear Factor of Activated T-cells 1 (NFAT1) has 13 conserved serine residues phosphorylation of which is required to expose a nuclear export signal and to mask a nuclear localization signal. In a phosphorylated state, NFAT1 localizes in the cytoplasm. Dephosphorylation of all 13 residues is necessary to expose the nuclear localization signal resulting in the activation of NFAT1 and its translocation of into the nucleus (Okamura et al, 2000).

The autoinhibitory effect of the RFTS domain of Dnmt1 may explain the discrepancy between the fast process of the remethylation of CpG sites *in vivo* after DNA replication and the relatively slow rate of DNA methylation by Dnmt1 determined *in vitro*. One may propose that there is an equilibrium of the two Dnmt1 states. The first is a "closed" conformation, in which the catalytic site of Dnmt1 is blocked by the RFTS domain. The second state is an "open" conformation, which Dnmt1 adopts upon the removal of the RFTS domain from the catalytic cleft (**Fig. 16**, transition I). The equilibrium between the two states of Dnmt1 is shifted towards the "closed" conformation, thus only a small proportion of the Dnm1 molecules methylate DNA, which reflects the low turnover rate in *in vitro* DNA methylation. So, the real turnover rate can be much higher than measured *in vitro*. The equilibrium can be shifted towards the "open" conformation of Dnmt1, which can be achieved *in vivo* by post-translational modifications or via interaction with other proteins. This regulatory mechanism could help to prevent uncontrolled DNA methylation activity of Dnmt1 and suggests the existence of pathways activating the enzyme.

Based on these results, we propose a modified model of the regulation of Dnmt1 including the possible role of the CXXC domain (**Fig. 16**, transition II). In the "closed" conformation, the regulation of Dnmt1's activity is ruled by the RFTS domain. Its translocation induces the transition of Dnm1 to the "open", catalytically active

conformation. In this state, the flexible CXXC domain may adopt another localization, which enables it to influence the specificity of the enzyme by binding to undesired unmethylated CpG sites.

In order to test this appealing model, Dnmt1 variants containing mutations in both the CXXC and RFTS domains should be generated. The specificity of such combined mutants should be compared with the specificity of Dnmt1 variants with mutations only in the RFTS domain. If the model is correct, then the mutagenesis in the CXXC domain abolishing its DNA binding will decrease the specificity of Dnmt1 RFTS domain mutants.

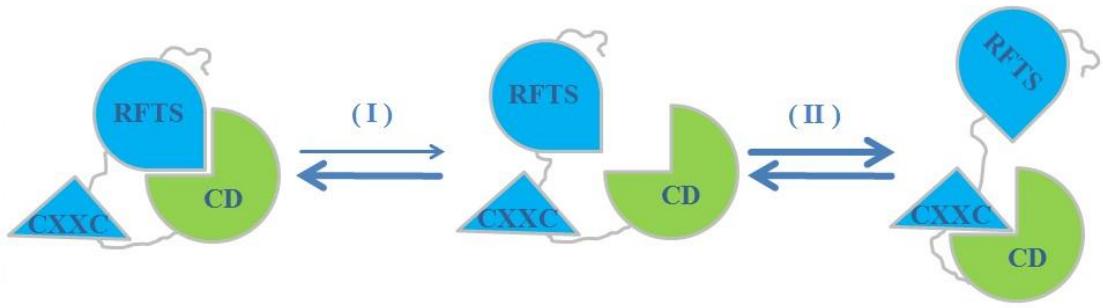

Figure 16. Model of the regulation of Dnmt1's activity and specificity by the RFTS and CXXC domains.
The RFTS domain inhibits the activity of the catalytic domain (CD) by occupying its DNA binding cleft, which corresponds to the "closed" inactive state of Dnmt1. The dissociation of the RFTS domain from the catalytic domain results in a transition of Dnmt1 to the "open" active state (I). Equilibrium between the "closed" and "open" states is shifted towards the inactive "closed" conformation. In the "open" state, the CXXC domain might occupy a position in which it can influence the specificity of Dnmt1.

3.3 Influence of the Uhrf1 protein on the enzymatic properties of Dnmt1

During the last years of investigations, many proteins interacting with Dnmt1 were identified (Qin et al, 2011a). Among them, PCNA and Uhrf1 occupy a special position. Both proteins play an important role in the maintenance DNA methylation. The direct interaction of Dnmt1 with PCNA increases the amount of Dnmt1 at the replication forks, which are the sites of DNA synthesis (Chuang et al, 1997; Leonhardt et al, 1992). This transient interaction facilitates the binding of Dnmt1 to the newly replicated DNA, where the enzyme moves by linear diffusion to find hemimethylated CpG sites (Egger et al, 2006; Jeltsch, 2008; Schermelleh et al, 2007). Though the binding of Dnmt1 to PCNA stimulates maintenance DNA methylation, it is not essential, and the loss of this interaction leads only to a twofold prolongation of the remethylation, but does not change the final level of DNA methylation (Egger et al, 2006; Spada et al, 2007). It was shown that Uhrf1 binds to hemimethylated CpG sites formed during DNA replication and interacts with Dnmt1. This interaction promotes maintenance DNA methylation by increasing the occupancy of Dnmt1 at the hemimethylated DNA. Interestingly, a knockout of the *Uhrf1* gene in mouse ES cells results in a phenotype similar to *Dnmt1-/-*. ES cells showed a massive loss of DNA methylation and impaired association of Dnmt1 with the replication forks (Bostick et al, 2007; Sharif et al, 2007). This indicates that the interaction of Dnmt1 with Uhrf1 is crucial for maintaining the DNA methylation patterns and might include not only targeting, but possibly even other effects. The mechanisms underlying such a strong phenotype are still unknown.

We investigated the possible influence of Uhrf1 on the DNA methylation activity of Dnmt1. For this we measured *in vitro* the activity of Dnmt1 in the presence of purified full-length Uhrf1 in the methylation reaction. We found that pre-incubation with Uhrf1 stimulates the methylation of the hemimethylated substrate by Dnmt1 about four times (**Fig. 17A**). This result suggests that the interaction with Uhrf1 directly stimulates the activity of Dnmt1. Surprisingly, Uhrf1 also increased the activity of Dnmt1 on an unmethylated substrate. In this case, the stimulatory effect was weaker, only about twofold, which results in a twofold increase of the preference of Dnmt1 for

the hemimethylated DNA in the presence of Uhrf1 (**Fig. 17B**). Therefore, we have shown for the first time that Uhrf1 stimulates the activity and specificity of Dnmt1 *in vitro*, which might facilitate the maintenance DNA methylation *in vivo*. Since there was previous evidence that the Uhrf1 SRA domain interacts with Dnmt1 (Achour et al, 2008), we investigated if this domain is involved in the stimulation of the Dnmt1 activity by Uhrf1. We demonstrated that the pre-incubation of Dnmt1 with the isolated SRA domain also leads to an increase of its activity, though this effect was weaker (twofold for the hemimethylated substrate) compared to the full-length Uhrf1 (**Fig. 17**).

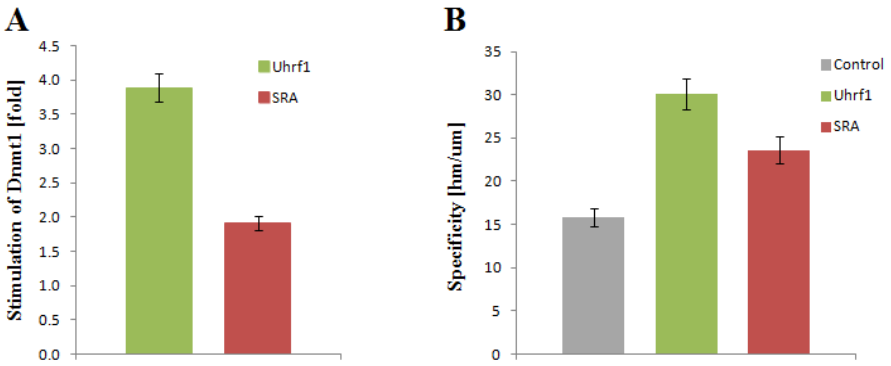

Figure 17. Uhrf1 and its SRA domain increase the activity and specificity of Dnmt1.
A. Stimulation of Dnmt1 activity is expressed as the ratio of the rates of methylation after pre-incubation with Uhrf1 or SRA to the rate of methylation after pre-incubation with the dialysis buffer. **B.** Uhrf1 and the SRA domain increase the specificity of Dnmt1. Specificity of Dnmt1 is determined as the ratio of the rate of methylation of the hemimethylated substrate (hm) divided by the rate of methylation of unmethylated substrate (um). Pre-incubation with Uhrf1 or SRA domain increased specificity of Dnmt1 from 16-fold to 30- and 24-fold, respectively. The stimulation of the activity and specificities were averaged over three to six independent experiments. Error bars indicate the standard error of the mean of the averages. Adopted from Bashtrykov et al. (manuscript has been submitted for a review, Appendix 4) with modifications.

As we know, Uhrf1 binds hemimethylated CpG sites formed after DNA replication via its SRA domain (Bostick et al, 2007). A special point of interest was to verify whether the DNA binding by the SRA domain is a prerequisite for stimulation of Dnmt1 activity. To investigate this, we generated several SRA domain mutants, namely D474R, R436E, R438E, and R496E, and measured their DNA binding by an electrophoretic mobility shift assay. Our results indicate that all SRA variants show a reduced DNA binding in comparison to the SRA wild-type (**Fig. 18A**). For the following experiments, we used the SRA R438E mutant (it showed an almost complete loss of interaction with the DNA) in DNA methylation assays with Dnmt1 and observed

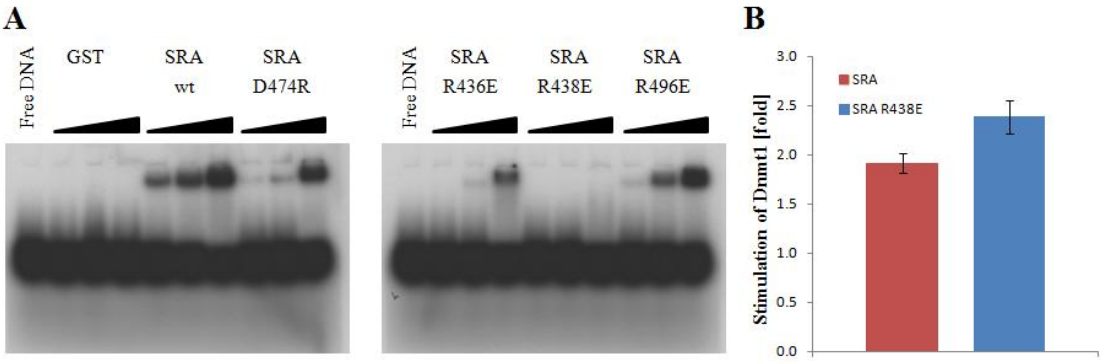

Figure 18. The SRA domain stimulates the activity of Dnmt1 in a DNA binding independent manner.
A. Example of results obtained in the electrophoretic mobility shift assays. The GST-tagged SRA domain wild type, the mutants (concentrations: 0.25, 0.5 and 1 μM), and the GST control (concentrations: 1.25, 2.5 and 5 μM) were pre-incubated with radioactively labeled hemimethylated DNA and resolved on a gel. The strongest reduction of the DNA binding was shown by the SRA R438E mutant. **B.** Stimulation of the activity of Dnmt1 by the SRA R438E mutant. The stimulation was averaged over three to six independent experiments. Error bars indicate the standard error of the mean of averages. Adopted from Bashtrykov et al. (manuscript has been submitted for a review, Appendix 4).

that pre-incubation with the SRA R438E mutant stimulated the activity of Dnmt1 even slightly stronger than pre-incubation with the SRA wild-type (**Fig. 18B**).

Additionally, pre-incubation of Uhrf1 or the SRA domain with the hemimethylated substrate before the start of the methylation reaction inhibited the activity of Dnmt1, which suggests that there exists a competition between Dnmt1 and Uhrf1 for the DNA. These data are in agreement with the results published by Felle et al. (2011), who used a similar experimental set-up and observed a decrease of Dnmt1 activity if the substrate was pre-incubated with Uhrf1. Therefore, the interaction of Uhrf1 with the DNA is necessary for the recruitment of Dnmt1 to the replicated DNA containing hemimethylated CpG sites, but the stimulation of Dnmt1 by Uhrf1 is a process independent of the DNA-binding. Recently it was found that the recognition of the H3R2 and H3K9me3 histone marks by the PHD and TTD domains, respectively, apart from the hemimethylated CpG binding by the SRA domain, are required for the proper localization of Uhrf1 at the heterochromatin and for the targeting of Dnmt1 to the DNA (Liu et al, 2013; Rothbart et al, 2013). So, Uhrf1 is a multivalent epigenetic player able to recognize hemimethylated DNA and repressive chromatin marks to guide Dnmt1 during maintenance DNA methylation.

Apart from this, our results allowed us to clarify the mode of Dnmt1 recruitment to the hemimethylated DNA by Uhrf1. Two models of recruitment were proposed:

Uhrf1 binds a hemimethylated CpG site via its SRA domain and targets Dnmt1 1) to the region with hemimethylated DNA or 2) to the same CpG site. According to the first model, Uhrf1 binds Dnmt1 leading to the recruitment of Dnmt1 to the regions with hemimethylated CpG sites, since Uhrf1 is enriched on replicated hemimethylated DNA. The second model proposed by Arita et al. (2008) implies a direct "hand-over" of the CpG site bound by the SRA to Dnmt1. The authors solved the structure of the SRA domain co-crystallized with hemimethylated DNA and found that the 5-methylcytosine is flipped out of the DNA helix. Also, they constructed a model of the Dnmt1 catalytic domain bound to hemimethylated DNA using the structure of the M.*Hha*I methyltransferase as template (Cheng & Blumenthal, 1996), since the crystal structure of the Dnmt1 was not solved at that time. Further, this model was overlaid with the SRA-DNA crystal structure. It was observed that Dnmt1 and the SRA approach the CpG site from opposite sides of the DNA helix and that both proteins may coexist on the DNA. The only steric clash in that model appeared in the major groove between the DNA recognition loops of the proteins. Based on this simulation, a model, in which Uhrf1 transfers the hemimethylated CpG site directly to Dnmt1, was proposed. If the "hand-over" model is correct, then a pre-incubation of the DNA with the SRA domain leading to a complex formation would increase the Dnmt1 methylation activity. As we discussed earlier, pre-incubation of DNA neither with the SRA domain nor with the full-length Uhrf1, stimulated DNA methylation by Dnmt1, but inhibited its activity. Therefore, our results argue against the "hand-over" model, and the first model is the most plausible. Also, the crystal structure of the truncated Dnmt1 with the hemimethylated DNA showed that the recognition of the substrate requires the formation of many protein-DNA contacts at both minor and major DNA groves over the CpG site (Song et al, 2012), which may cause more steric clashes than predicted by Arita et al. (2008). Thus, the first model, suggesting that the interaction with Uhrf1 targets Dnmt1 and increases its abundance at the genomic regions containing hemimethylated CpG sites, is more likely.

Finally, we aimed at mapping the domain(s) of Dnmt1 which are involved in the interaction with Uhrf1 and mediate the stimulatory effect. Early studies showed that the Uhrf1 SRA domain interacts with two regions of Dnmt1: the RFTS domain and residues 1081-1408 corresponding partially to the adjacent parts of BAH2 and the catalytic

domains (Achour et al, 2008; Bostick et al, 2007). We found that the SRA domain strongly interacts with the isolated RFTS domain, whereas the binding to the truncated Dnmt1 (731-1602) variant lacking the RFTS domain was significantly reduced (**Fig. 19A**). As was mentioned previously, the RFTS domain is responsible for the localization of Dnmt1 at the replication foci (Leonhardt et al, 1992), where Dnmt1 co-localizes with Uhrf1 (Bostick et al, 2007; Sharif et al, 2007). Thus, one may speculate that the localization of Dnmt1 at sites of DNA replication is (at least partially) based on its interaction with the Uhrf1 protein via the RFTS domain of Dnmt1. To investigate the importance of the interaction between the RFTS domain and Uhrf1 for the stimulation of the Dnmt1 activity, we obtained a Dnmt1 E406R/D407R variant, in which a mutagenesis of the RFTS domain abolished binding to the SRA domain (**Fig. 19A**). *In*

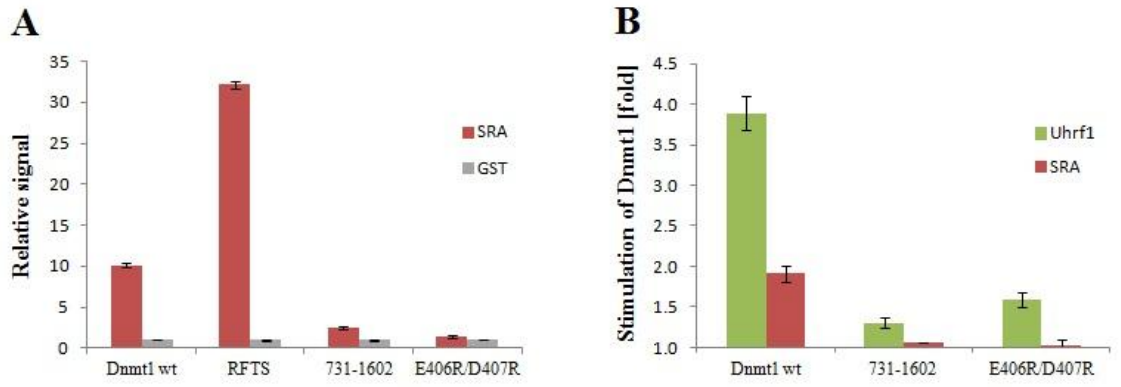

Figure 19. Stimulation of Dnmt1's activity by Uhrf1 is mediated via its interaction with the RFTS domain of Dnmt1.
A. Interaction between GST (Glutathione S-transferase) tagged SRA domain and Dnmt1 variants was tested by AlphaScreen technology. The results are presented as the ratio between the Alpha signal obtained after incubation of Dnmt1 variants with GST-SRA and the Alpha signal obtained after incubation of the same Dnmt1 variant with GST. The Alpha signal was averaged over three independent experiments. Error bars indicate the standard error of the mean of averages. **B.** Stimulation of the activity of Dnmt1 variants by Uhrf1 and its isolated SRA domain. The stimulation was averaged over three to six independent experiments. Error bars indicate the standard error of the mean of averages. Adopted from Bashtrykov et al. (manuscript has been submitted for a review, Appendix 4) with modifications.

vitro experiments showed that Uhrf1 and the SRA domain stimulate the methylation activity of the Dnmt1 E406R/D407R mutant significantly less than of the wild-type Dnmt1. Additionally, the truncated Dnmt1 (731-1602) variant lacking the RFTS domain completely lost the stimulation of Uhrf1 or its SRA domain (**Fig. 19B**). Therefore, we conclude that the stimulation of Dnmt1 DNA methylation activity by Uhrf1 is mediated mainly via the interaction between the RFTS domain and the SRA domain of the proteins. As discussed above, the RFTS domain binds to the DNA binding cleft of

Dnmt1 and inhibits its activity by blocking the access of the DNA to the catalytic site (Syeda et al, 2011; Takeshita et al, 2011), and the destabilization of the RFTS domain position results in the activation of the enzyme (Bashtrykov et al., manuscript in preparation). Thus, one of the possible explanations for the stimulatory effect of Uhrf1 on Dnmt1 activity observed by us is a model of Syeda et al. (2011), proposing that Uhrf1 facilitates the translocation of the RFTS domain from the DNA binding cleft resulting in the transition of Dnmt1 into the "open" catalytically potent conformation. Our results show that the targeting and stimulatory effects of Uhrf1 explain why the knockout of the *Uhrf1* gene results in a drastic decrease of the global DNA methylation (Bostick et al, 2007; Sharif et al, 2007), whereas the disruption of the PCNA-Dnmt1 interaction leads only to the delay of the remethylation, but does not influence the level of DNA methylation (Egger et al, 2006; Schermelleh et al, 2007; Spada et al, 2007). Hence, Uhrf1 targets Dnmt1 to the genomic regions containing hemimethylated DNA and stimulates the activity and specificity of the enzyme by facilitating its transition to the "open" state.

PCNA and Uhrf1 facilitate remethylation of the genome recruiting Dnmt1 during DNA replication. Apart from this, it was shown that transcription factors and other proteins may target Dnmt1 for local methylation at certain genes and genomic regions (Esteve et al, 2007; Kim et al, 2009; Xu et al, 2011; Zhang et al, 2005). The targeting of DNA methyltransferases and other DNA and chromatin modifying enzymes is a common approach. For example, the setting of DNA methylation patterns by Dnmt3a and Dnmt3b and the establishment of the epigenetic marks at imprinted genes are targeted processes (Kelsey & Feil, 2013; Law & Jacobsen, 2010).

Additionally, certain features of Dnmt1 can be regulated by post-translational modifications (PMTs). As we know, PMTs may alter the activity, specificity, and interaction with other proteins as well as the abundance of the enzymes by modulating their degradation and stability. Several post-translational modifications of Dnmt1 have been found already, including phosphorylation, methylation, acetylation, ubiquitination, and sumoylation. Phosphorylation of serine 515 regulates the interaction of the N-terminal part of Dnmt1 with the catalytic domain and increases the catalytic activity of the enzyme (Glickman et al, 1997; Goyal et al, 2007). Casein kinase 1 delta/epsilon phosphorylates serine 146, which decreases the DNA binding affinity of Dnmt1

(Sugiyama et al, 2010). Dnmt1 is phosphorylated at serine 154 by cyclin-dependent kinases 1, 2, and 5 *in vitro* and *in vivo,* but no clear biochemical characterization of the function of this modification has been conducted (Lavoie & St-Pierre, 2011). Hervouet et al. (2010) showed that the interaction of Dnmt1 with PCNA and Uhrf1 is regulated by phosphorylation at serines 127 and 143. Phosphorylation of Dnmt1 at serine 127 inhibits binding of Uhrf1 and double phosphorylation of serines 127 and 143 prevents its interaction with PCNA and Uhrf1. Dnmt1 is a cell cycle-regulated protein and its abundance is increased in the S phase, when its activity is required, and declines shortly after finishing of the DNA replication. Several post-translational modifications are involved in the regulation of Dnmt1's stability. It was found that Dnmt1 acetylation by Tip60 followed by ubiquitination by Uhrf1 leads to proteosomal degradation of Dnmt1. In contrast, the stability of Dnmt1 is increased by the opposite action of Histone deacetylase 1 (HDAC1) and deubiquitinase Ubiquitin specific peptidase 7 (USP7) (Du et al, 2010; Qin et al, 2011b). In addition, phosphorylation of serine 143 stabilizes Dnmt1 and prevents its methylation at lysine 142 by SET domain containing (lysine methyltransferase) 7 (SETD7), a mark found to induce degradation of Dnmt1 (Esteve et al, 2011). Further studies on PTMs of Dnmt1 are needed, including a careful and detailed analysis of the functional consequences of modification.

Finally, based on our results and data obtained by others, we propose three levels of the regulation of Dnmt1 activity and specificity (**Fig. 20**). The Dnmt1 catalytic domain and its intrinsic properties are at the first level of the regulation of Dnmt1. The catalytic domain has available all conserved motifs essential for the methyltransferase activity. It forms specific contacts with the DNA required for the recognition of the CpG site and its hemimethylated state. These features underlie the specificity of Dnmt1 for the hemimethylated CpG sites. The second level of control consists of the RFTS and CXXC domains, which by intramolecular interactions regulate the functions of Dnmt1. The RFTS domain positioned within the DNA binding cleft keeps the enzyme in the "closed" inactive conformation, which prevents the uncontrolled activity of Dnmt1 including undesired *de novo* DNA methylation. The transition of Dnmt1 into the "open" conformation upon the translocation of the RFTS domain from the DNA binding cleft makes it catalytically potent. In this state, the CXXC domain, having flexible linkers with other domains, might adopt a position allowing it to improve the specificity of

Dnmt1 sequestrating unmethylated CpG sites from the catalytic domain (which is still to be shown). The third level of regulation includes external factors controlling Dnmt1.

Interacting partners (PCNA, Uhrf1, transcription factors, etc.) of Dnmt1 regulate its activity recruiting the enzyme to the genomic regions, where the DNA methylation activity is required. Interaction with Uhrf1 also causes the transition of Dnmt1 to the "open" state and stimulates its activity and specificity. The partitioning between "open" and "closed" conformations can be modulated as well by the post-translational modifications of Dnmt1. Additionally, the activity of Dnmt1 is controlled at the level of the abundance of the enzyme in the nucleus. All these mechanisms discussed work together and orchestrate the maintenance of the DNA methylation patterns.

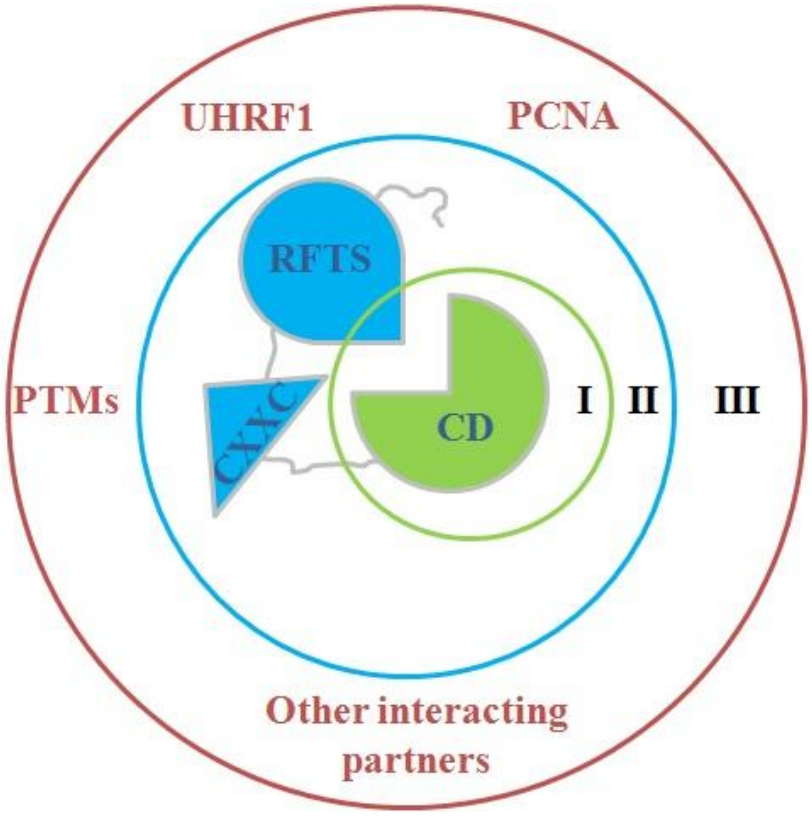

Figure 20. Three levels of regulation of Dnmt1 activity and specificity.
I – the catalytic domain (CD) of Dnmt1 harbours intrinsic specificity towards hemimethylated CpG sites. II – the CXXC and RFTS domains of Dnmt1 are involved in the intramolecular regulation of specificity and activity of the enzyme. III – PCNA, Uhrf1, and other interacting partners recruit and activate Dnmt1 at the regions in which DNA methylation is required. PTMs are also involved in the regulation of the activity and abundance of Dnmt1.

4. Prospective directions / Outlook

The maintenance of the DNA methylation patterns is a critical epigenetic phenomenon, based on the ability of Dnmt1 to recognize and specifically methylate hemimethylated CpG sites. In the present work, several mechanisms facilitating the specificity of this process have been investigated and discussed. Apart from this, several follow-up experiments could be proposed, which might contribute to a deeper understanding of the maintenance DNA methylation by Dnmt1.

1. One of the open questions is a possible flanking sequence preference of Dnmt1. Here, we showed that Dnmt1 does not have flanking preference for the base pair at the 5´ flank of the CpG site. These experiments were inspired by the findings in the crystal structure of Dnmt1 (Song et al, 2012), consequently, the 3´ flank base pair was not investigated. However, in an earlier study from our group strong evidence in favour of flanking sequence preferences in the methylation of unmethylated substrates was provided (Goyal et al, 2006). In order to clarify this question, a more general and systematic analysis should be conducted using hemimethylated and unmethylated substrates. These experiments should include the 3´ flank base pair of the CpG site as well as more distant positions at both flanks (-2, +2, -3, +3, etc.).
2. Our experiments with the Dnmt1 mutants lacking the DNA binding by the CXXC domain showed that the DNA binding of the CXXC domain is irrelevant for the substrate specificity of the full-length Dnmt1. We were surprised by this finding, since the crystal structure of Dnmt1 with DNA showing the DNA binding to the CXXC domain, is well resolved and provides many structural details, which make it is hard to believe that this structure is completely wrong or irrelevant. However, it is possible that the RFTS domain inhibits the activity of the CXXC domain in the closed conformation of Dnmt1 by occupying the DNA binding pocket so that the CXXC domain cannot influence the specificity of Dnmt1. In the “open” active state, the CXXC domain might adopt another conformation and contribute to the preference of Dnmt1 for the hemimethylated CpG sites. To investigate this hypothesis, we need to induce the transition of Dnmt1 variants lacking the DNA binding by the CXXC domain into the “open” state. If the model is correct, than in the “open” state these variants will demonstrate a reduced preference for the hemimethylated substrate in comparison to the Dnmt1 variant containing the wild-

type CXXC domain. For this purpose, the 4S mutant generated in the present study could be used. The transition to the "open" state can be stimulated by the additional mutations in the RFTS domain (E572R/D575R) or by pre-incubation of the Dnmt1 variants with Uhrf1.

3. The influence of the CXXC domain and the Uhrf1 protein on the specificity of Dnmt1 can also be tested on a long DNA substrate containing multiple CpG sites. It is possible to generate three types of substrates containing 1) only unmethylated CpG sites, 2) only hemimethylated CpG sites, and 3) a mixed pattern of both hemimethylated and unmethylated CpGs. The specificity of the Dnmt1 variants discussed above can be determined on these substrates in the presence of Uhrf1 or without it. Dnmt1 is a processive enzyme able to slide along the DNA and to methylate several CpG site without dissociation from the DNA. Hence, a long DNA containing multiple CpG sites is a more natural substrate for Dnmt1.
4. We have shown that Uhrf1 increases the activity of Dnmt1 by inducing its transition to an "open" conformation. It would be most interesting to investigate whether Uhrf1 stimulates the DNA binding of Dnmt1 and/or its catalytic rate. Using a long hemimethylated substrate, this can be tested by determining the fraction of methylated DNA molecules and the number of methyl groups introduced into each DNA molecule under conditions of (slight) excess of the substrate. If Uhrf1 increases the rate of DNA methylation, we should see more methylation events during the same reaction time resulting in the formation of longer stretches of methylated CpG sites in comparison to the sample without Uhrf1. If Uhrf1 increases the rate of DNA binding, more Dnmt1-DNA complexes should be formed in the same time of pre-incubation of both partners in the presence of Uhrf1. This will lead to a higher fraction of DNA molecules containing methylation.
5. Special attention should be paid to *in vivo* experiments. In the present study, we generated hyperactive Dnmt1 variants (D554R and E572R/D575R) and the Dnmt1 (E406R/D407R) mutant, which lost binding to Uhrf1. In order to investigate the relevance of the *in vitro* effects of these Dnmt1 variants in cells, one might conduct the following experiments. A cell line with *Dnmt1* knockout should be transiently transfected with the Dnmt1 wild-type or the obtained mutants. Several days later,

genomic DNA can be isolated and the global or site-specific DNA methylation by Dnmt1 can be compared to that of the variants.

6. Finally, an investigation of the post-translational modifications of Dnmt1, namely phosphorylation, is one of our ongoing projects. Several dozens of phosphorylation sites have been found in Dnmt1 (http://www.phosphosite.org), but biochemical analyses have been carried out for only some of them. Since many of these Dnmt1 modifications were discovered in cancer cells, an investigation of the functional consequences of them on catalysis might elucidate the role played by Dnmt1 in cancerogenesis. A mutational study of the phosphorylation is feasible in particular for serine residues (the most abundant type of phosphorylation), because amino acid exchanges can be introduced into Dnmt1 (and other proteins) mimicking the unmodified and the modified state. The glutamic acid mimics the phosphorylation of serine residues, since it has a similar size and charge. The unmodified state can be achieved by mutagenesis of serine to alanine.

We believe that these experiments will help to further improve our understanding of this complicated enzyme, which plays a central role in human development and disease processes.

5. References

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

Agoston AT, Argani P, Yegnasubramanian S, De Marzo AM, Ansari-Lari MA, Hicks JL, Davidson NE, Nelson WG (2005) Increased protein stability causes DNA methyltransferase 1 dysregulation in breast cancer. *The Journal of biological chemistry* **280:** 18302-18310

Allen MD, Grummitt CG, Hilcenko C, Min SY, Tonkin LM, Johnson CM, Freund SM, Bycroft M, Warren AJ (2006) Solution structure of the nonmethyl-CpG-binding CXXC domain of the leukaemia-associated MLL histone methyltransferase. *The EMBO journal* **25:** 4503-4512

Araujo FD, Croteau S, Slack AD, Milutinovic S, Bigey P, Price GB, Zannis-Hadjopoulos M, Szyf M (2001) The DNMT1 target recognition domain resides in the N terminus. *The Journal of biological chemistry* **276:** 6930-6936

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

Arita K, Isogai S, Oda T, Unoki M, Sugita K, Sekiyama N, Kuwata K, Hamamoto R, Tochio H, Sato M, Ariyoshi M, Shirakawa M (2012) Recognition of modification status on a histone H3 tail by linked histone reader modules of the epigenetic regulator UHRF1. *Proceedings of the National Academy of Sciences of the United States of America* **109:** 12950-12955

Auclair G, Weber M (2012) Mechanisms of DNA methylation and demethylation in mammals. *Biochimie* **94:** 2202-2211

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

Ayton PM, Chen EH, Cleary ML (2004) Binding to nonmethylated CpG DNA is essential for target recognition, transactivation, and myeloid transformation by an MLL oncoprotein. *Molecular and cellular biology* **24:** 10470-10478

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

Barakat TS, Gribnau J (2012) X chromosome inactivation in the cycle of life. *Development* **139:** 2085-2089

Bashtrykov P, Jankevicius G, Smarandache A, Jurkowska RZ, Ragozin S, Jeltsch A (2012a) Specificity of Dnmt1 for methylation of hemimethylated CpG sites resides in its catalytic domain. *Chemistry & biology* **19:** 572-578

Bashtrykov P, Ragozin S, Jeltsch A (2012b) Mechanistic details of the DNA recognition by the Dnmt1 DNA methyltransferase. *FEBS letters* **586:** 1821-1823

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

Bergman Y, Cedar H (2013) DNA methylation dynamics in health and disease. *Nature structural & molecular biology* **20:** 274-281

Bestor T, Laudano A, Mattaliano R, Ingram V (1988) Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *Journal of molecular biology* **203:** 971-983

Bestor TH, Bourc'his D (2004) Transposon silencing and imprint establishment in mammalian germ cells. *Cold Spring Harbor symposia on quantitative biology* **69:** 381-387

Bestor TH, Ingram VM (1983) Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proceedings of the National Academy of Sciences of the United States of America* **80:** 5559-5563

Bird A (2002) DNA methylation patterns and epigenetic memory. *Genes & development* **16:** 6-21

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

Birke M, Schreiner S, Garcia-Cuellar MP, Mahr K, Titgemeyer F, Slany RK (2002) The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. *Nucleic acids research* **30:** 958-965

Boettcher AJ, Wu J, Kim C, Yang J, Bruystens J, Cheung N, Pennypacker JK, Blumenthal DA, Kornev AP, Taylor SS (2011) Realizing the allosteric potential of the tetrameric protein kinase A RIalpha holoenzyme. *Structure* **19:** 265-276

Bonapace IM, Latella L, Papait R, Nicassio F, Sacco A, Muto M, Crescenzi M, Di Fiore PP (2002) Np95 is regulated by E1A during mitotic reactivation of terminally differentiated cells and is essential for S phase entry. *The Journal of cell biology* **157:** 909-914

Bonasio R, Tu S, Reinberg D (2010) Molecular signals of epigenetic states. *Science* **330:** 612-616

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

Cardoso MC, Leonhardt H (1999) DNA methyltransferase is actively retained in the cytoplasm during early development. *The Journal of cell biology* **147:** 25-32

Ceccaldi A, Rajavelu A, Ragozin S, Senamaud-Beaufort C, Bashtrykov P, Testa N, Dali-Ali H, Maulay-Bailly C, Amand S, Guianvarc'h D, Jeltsch A, Arimondo PB (2013) Identification of novel inhibitors of DNA methylation by screening of a chemical library. *ACS chemical biology* **8:** 543-548

Cheng J, Yang Y, Fang J, Xiao J, Zhu T, Chen F, Wang P, Li Z, Yang H, Xu Y (2013) Structural insight into coordinated recognition of trimethylated histone H3 lysine 9 (H3K9me3) by the plant homeodomain (PHD) and tandem tudor domain (TTD) of UHRF1 (ubiquitin-like, containing PHD and RING finger domains, 1) protein. *The Journal of biological chemistry* **288:** 1329-1339

Cheng X (1995) Structure and function of DNA methyltransferases. *Annual review of biophysics and biomolecular structure* **24:** 293-318

Cheng X, Blumenthal RM (1996) Finding a basis for flipping bases. *Structure* **4:** 639-645

Cheng X, Roberts RJ (2001) AdoMet-dependent methylation, DNA methyltransferases and base flipping. *Nucleic acids research* **29:** 3784-3795

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

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

Dai C, Shi D, Gu W (2013) Negative Regulation of the Acetyltransferase TIP60-p53 Interplay by UHRF1 (Ubiquitin-like with PHD and RING Finger Domains 1). *The Journal of biological chemistry* **288:** 19581-19592

Dalton SR, Bellacosa A (2012) DNA demethylation by TDG. *Epigenomics* **4:** 459-467

Daskalos A, Oleksiewicz U, Filia A, Nikolaidis G, Xinarianos G, Gosney JR, Malliri A, Field JK, Liloglou T (2011) UHRF1-mediated tumor suppressor gene inactivation in nonsmall cell lung cancer. *Cancer* **117:** 1027-1037

Du J, Zhong X, Bernatavichute YV, Stroud H, Feng S, Caro E, Vashisht AA, Terragni J, Chin HG, Tu A, Hetzel J, Wohlschlegel JA, Pradhan S, Patel DJ, Jacobsen SE (2012) Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. *Cell* **151:** 167-180

Du Z, Song J, Wang Y, Zhao Y, Guda K, Yang S, Kao HY, Xu Y, Willis J, Markowitz SD, Sedwick D, Ewing RM, Wang Z (2010) DNMT1 stability is regulated by proteins coordinating deubiquitination and acetylation-driven ubiquitination. *Science signaling* **3:** ra80

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

Egger G, Jeong S, Escobar SG, Cortez CC, Li TW, Saito Y, Yoo CB, Jones PA, Liang G (2006) Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. *Proceedings of the National Academy of Sciences of the United States of America* **103:** 14080-14085

Esteve PO, Chang Y, Samaranayake M, Upadhyay AK, Horton JR, Feehery GR, Cheng X, Pradhan S (2011) A methylation and phosphorylation switch between an adjacent lysine and serine determines human DNMT1 stability. *Nature structural & molecular biology* **18:** 42-48

Esteve PO, Chin HG, Pradhan S (2007) Molecular mechanisms of transactivation and doxorubicin-mediated repression of survivin gene in cancer cells. *The Journal of biological chemistry* **282:** 2615-2625

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

Feltus FA, Lee EK, Costello JF, Plass C, Vertino PM (2003) Predicting aberrant CpG island methylation. *Proceedings of the National Academy of Sciences of the United States of America* **100:** 12253-12258

Fischle W, Tseng BS, Dormann HL, Ueberheide BM, Garcia BA, Shabanowitz J, Hunt DF, Funabiki H, Allis CD (2005) Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438:** 1116-1122

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

Fujimori A, Matsuda Y, Takemoto Y, Hashimoto Y, Kubo E, Araki R, Fukumura R, Mita K, Tatsumi K, Muto M (1998) Cloning and mapping of Np95 gene which encodes

a novel nuclear protein associated with cell proliferation. *Mammalian genome : official journal of the International Mammalian Genome Society* **9:** 1032-1035

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

Glickman JF, Pavlovich JG, Reich NO (1997) Peptide mapping of the murine DNA methyltransferase reveals a major phosphorylation site and the start of translation. *The Journal of biological chemistry* **272:** 17851-17857

Goedecke K, Pignot M, Goody RS, Scheidig AJ, Weinhold E (2001) Structure of the N6-adenine DNA methyltransferase M.TaqI in complex with DNA and a cofactor analog. *Nature structural biology* **8:** 121-125

Gowher H, Jeltsch A (2001) Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites. *Journal of molecular biology* **309:** 1201-1208

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. *The Journal of biological chemistry* **280:** 13341-13348

Goyal R, Rathert P, Laser H, Gowher H, Jeltsch A (2007) Phosphorylation of serine-515 activates the Mammalian maintenance methyltransferase Dnmt1. *Epigenetics : official journal of the DNA Methylation Society* **2:** 155-160

Goyal R, Reinhardt R, Jeltsch A (2006) Accuracy of DNA methylation pattern preservation by the Dnmt1 methyltransferase. *Nucleic acids research* **34:** 1182-1188

Grayson DR, Guidotti A (2013) The dynamics of DNA methylation in schizophrenia and related psychiatric disorders. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* **38:** 138-166

Guan D, Factor D, Liu Y, Wang Z, Kao HY (2013) The epigenetic regulator UHRF1 promotes ubiquitination-mediated degradation of the tumor-suppressor protein promyelocytic leukemia protein. *Oncogene* **32:** 3819-3828

Handa V, Jeltsch A (2005) Profound flanking sequence preference of Dnmt3a and Dnmt3b mammalian DNA methyltransferases shape the human epigenome. *Journal of molecular biology* **348:** 1103-1112

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

Hemberger M, Dean W, Reik W (2009) Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nature reviews Molecular cell biology* **10:** 526-537

Hermann A, Goyal R, Jeltsch A (2004) The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *The Journal of biological chemistry* **279:** 48350-48359

Hervouet E, Lalier L, Debien E, Cheray M, Geairon A, Rogniaux H, Loussouarn D, Martin SA, Vallette FM, Cartron PF (2010) Disruption of Dnmt1/PCNA/UHRF1 interactions promotes tumorigenesis from human and mice glial cells. *PloS one* **5:** e11333

Hirota T, Lipp JJ, Toh BH, Peters JM (2005) Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* **438:** 1176-1180

Holliday R, Pugh JE (1975) DNA modification mechanisms and gene activity during development. *Science* **187:** 226-232

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

Horton JR, Liebert K, Bekes M, Jeltsch A, Cheng X (2006) Structure and substrate recognition of the Escherichia coli DNA adenine methyltransferase. *Journal of molecular biology* **358:** 559-570

Horton JR, Liebert K, Hattman S, Jeltsch A, Cheng X (2005) Transition from nonspecific to specific DNA interactions along the substrate-recognition pathway of dam methyltransferase. *Cell* **121:** 349-361

Hu L, Li Z, Wang P, Lin Y, Xu Y (2011) Crystal structure of PHD domain of UHRF1 and insights into recognition of unmodified histone H3 arginine residue 2. *Cell research* **21:** 1374-1378

Jackson DA, Pombo A (1998) Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. *The Journal of cell biology* **140:** 1285-1295

Jair KW, Bachman KE, Suzuki H, Ting AH, Rhee I, Yen RW, Baylin SB, Schuebel KE (2006) De novo CpG island methylation in human cancer cells. *Cancer research* **66:** 682-692

Jeltsch A (2002) Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. *Chembiochem : a European journal of chemical biology* **3:** 274-293

Jeltsch A (2006) On the enzymatic properties of Dnmt1: specificity, processivity, mechanism of linear diffusion and allosteric regulation of the enzyme. *Epigenetics : official journal of the DNA Methylation Society* **1:** 63-66

Jeltsch A (2008) Reading and writing DNA methylation. *Nature structural & molecular biology* **15:** 1003-1004

Jeltsch A (2010) Molecular biology. Phylogeny of methylomes. *Science* **328:** 837-838

Jeltsch A (2013) Oxygen, epigenetic signaling, and the evolution of early life. *Trends in biochemical sciences* **38:** 172-176

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

Jin W, Chen L, Chen Y, Xu SG, Di GH, Yin WJ, Wu J, Shao ZM (2010) UHRF1 is associated with epigenetic silencing of BRCA1 in sporadic breast cancer. *Breast cancer research and treatment* **123:** 359-373

Jones PA, Baylin SB (2007) The epigenomics of cancer. *Cell* **128:** 683-692

Jones PA, Liang G (2009) Rethinking how DNA methylation patterns are maintained. *Nature reviews Genetics* **10:** 805-811

Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, Wolffe AP (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature genetics* **19:** 187-191

Jorgensen HF, Ben-Porath I, Bird AP (2004) Mbd1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains. *Molecular and cellular biology* **24:** 3387-3395

Jurkowska RZ, Jurkowski TP, Jeltsch A (2011a) Structure and function of mammalian DNA methyltransferases. *Chembiochem : a European journal of chemical biology* **12:** 206-222

Jurkowska RZ, Rajavelu A, Anspach N, Urbanke C, Jankevicius G, Ragozin S, Nellen W, Jeltsch A (2011b) Oligomerization and binding of the Dnmt3a DNA methyltransferase to parallel DNA molecules: heterochromatic localization and role of Dnmt3L. *The Journal of biological chemistry* **286:** 24200-24207

Jurkowska RZ, Siddique AN, Jurkowski TP, Jeltsch A (2011c) Approaches to enzyme and substrate design of the murine Dnmt3a DNA methyltransferase. *Chembiochem : a European journal of chemical biology* **12:** 1589-1594

Jurkowski TP, Jeltsch A (2011) Burning off DNA methylation: new evidence for oxygen-dependent DNA demethylation. *Chembiochem : a European journal of chemical biology* **12:** 2543-2545

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:** 900-903

Karagianni P, Amazit L, Qin J, Wong J (2008) ICBP90, a novel methyl K9 H3 binding protein linking protein ubiquitination with heterochromatin formation. *Molecular and cellular biology* **28:** 705-717

Kelsey G, Feil R (2013) New insights into establishment and maintenance of DNA methylation imprints in mammals. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* **368:** 20110336

Kim C, Vigil D, Anand G, Taylor SS (2006) Structure and dynamics of PKA signaling proteins. *European journal of cell biology* **85:** 651-654

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

Kimura F, Seifert HH, Florl AR, Santourlidis S, Steinhoff C, Swiatkowski S, Mahotka C, Gerharz CD, Schulz WA (2003) Decrease of DNA methyltransferase 1 expression relative to cell proliferation in transitional cell carcinoma. *International journal of cancer Journal international du cancer* **104:** 568-578

Kinney SR, Pradhan S (2013) Ten eleven translocation enzymes and 5-hydroxymethylation in mammalian development and cancer. *Advances in experimental medicine and biology* **754:** 57-79

Klimasauskas S, Kumar S, Roberts RJ, Cheng X (1994) HhaI methyltransferase flips its target base out of the DNA helix. *Cell* **76:** 357-369

Kuo AJ, Song J, Cheung P, Ishibe-Murakami S, Yamazoe S, Chen JK, Patel DJ, Gozani O (2012) The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. *Nature* **484:** 115-119

Latham T, Gilbert N, Ramsahoye B (2008) DNA methylation in mouse embryonic stem cells and development. *Cell and tissue research* **331:** 31-55

Lavoie G, St-Pierre Y (2011) Phosphorylation of human DNMT1: implication of cyclin-dependent kinases. *Biochemical and biophysical research communications* **409:** 187-192

Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature reviews Genetics* **11:** 204-220

Lee GE, Kim JH, Taylor M, Muller MT (2010) DNA methyltransferase 1-associated protein (DMAP1) is a co-repressor that stimulates DNA methylation globally and

locally at sites of double strand break repair. *The Journal of biological chemistry* **285:** 37630-37640

Lee JH, Voo KS, Skalnik DG (2001) Identification and characterization of the DNA binding domain of CpG-binding protein. *The Journal of biological chemistry* **276:** 44669-44676

Lee PJ, Washer LL, Law DJ, Boland CR, Horon IL, Feinberg AP (1996) Limited up-regulation of DNA methyltransferase in human colon cancer reflecting increased cell proliferation. *Proceedings of the National Academy of Sciences of the United States of America* **93:** 10366-10370

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

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

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

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

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

Lin IG, Han L, Taghva A, O'Brien LE, Hsieh CL (2002) Murine de novo methyltransferase Dnmt3a demonstrates strand asymmetry and site preference in the methylation of DNA in vitro. *Molecular and cellular biology* **22:** 704-723

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462:** 315-322

Liu X, Gao Q, Li P, Zhao Q, Zhang J, Li J, Koseki H, Wong J (2013) UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9. *Nature communications* **4:** 1563

Liu Y, Oakeley EJ, Sun L, Jost JP (1998) Multiple domains are involved in the targeting of the mouse DNA methyltransferase to the DNA replication foci. *Nucleic acids research* **26:** 1038-1045

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

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

Miura M, Watanabe H, Sasaki T, Tatsumi K, Muto M (2001) Dynamic changes in subnuclear NP95 location during the cell cycle and its spatial relationship with DNA replication foci. *Experimental cell research* **263:** 202-208

Muratani M, Tansey WP (2003) How the ubiquitin-proteasome system controls transcription. *Nature reviews Molecular cell biology* **4:** 192-201

Nady N, Lemak A, Walker JR, Avvakumov GV, Kareta MS, Achour M, Xue S, Duan S, Allali-Hassani A, Zuo X, Wang YX, Bronner C, Chedin F, Arrowsmith CH, Dhe-Paganon S (2011) Recognition of multivalent histone states associated with heterochromatin by UHRF1 protein. *The Journal of biological chemistry* **286:** 24300-24311

Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393:** 386-389

Noguchi K, Vassilev A, Ghosh S, Yates JL, DePamphilis ML (2006) The BAH domain facilitates the ability of human Orc1 protein to activate replication origins in vivo. *The EMBO journal* **25:** 5372-5382

Okamura H, Aramburu J, Garcia-Rodriguez C, Viola JP, Raghavan A, Tahiliani M, Zhang X, Qin J, Hogan PG, Rao A (2000) Concerted dephosphorylation of the transcription factor NFAT1 induces a conformational switch that regulates transcriptional activity. *Molecular cell* **6:** 539-550

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

Panning B, Jaenisch R (1998) RNA and the epigenetic regulation of X chromosome inactivation. *Cell* **93:** 305-308

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

Portela A, Esteller M (2010) Epigenetic modifications and human disease. *Nature biotechnology* **28:** 1057-1068

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

Pradhan S, Bacolla A, Wells RD, Roberts RJ (1999) Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. *The Journal of biological chemistry* **274:** 33002-33010

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

Prendergast GC, Lawe D, Ziff EB (1991) Association of Myn, the murine homolog of max, with c-Myc stimulates methylation-sensitive DNA binding and ras cotransformation. *Cell* **65:** 395-407

Pufall MA, Graves BJ (2002) Autoinhibitory domains: modular effectors of cellular regulation. *Annual review of cell and developmental biology* **18:** 421-462

Qian C, Li S, Jakoncic J, Zeng L, Walsh MJ, Zhou MM (2008) Structure and hemimethylated CpG binding of the SRA domain from human UHRF1. *The Journal of biological chemistry* **283:** 34490-34494

Qin W, Leonhardt H, Pichler G (2011a) Regulation of DNA methyltransferase 1 by interactions and modifications. *Nucleus* **2:** 392-402

Qin W, Leonhardt H, Spada F (2011b) Usp7 and Uhrf1 control ubiquitination and stability of the maintenance DNA methyltransferase Dnmt1. *Journal of cellular biochemistry* **112:** 439-444

Rajakumara E, Wang Z, Ma H, Hu L, Chen H, Lin Y, Guo R, Wu F, Li H, Lan F, Shi YG, Xu Y, Patel DJ, Shi Y (2011) PHD finger recognition of unmodified histone H3R2 links UHRF1 to regulation of euchromatic gene expression. *Molecular cell* **43:** 275-284

Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proceedings of the National Academy of Sciences of the United States of America* **97:** 5237-5242

Reik W (2007) Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* **447:** 425-432

Reinisch KM, Chen L, Verdine GL, Lipscomb WN (1995) The crystal structure of HaeIII methyltransferase convalently complexed to DNA: an extrahelical cytosine and rearranged base pairing. *Cell* **82:** 143-153

Riggs AD (1975) X inactivation, differentiation, and DNA methylation. *Cytogenetics and cell genetics* **14:** 9-25

Roberts RJ, Cheng X (1998) Base flipping. *Annual review of biochemistry* **67:** 181-198

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

Roth M, Jeltsch A (2001) Changing the target base specificity of the EcoRV DNA methyltransferase by rational de novo protein-design. *Nucleic acids research* **29:** 3137-3144

Rothbart SB, Dickson BM, Ong MS, Krajewski K, Houliston S, Kireev DB, Arrowsmith CH, Strahl BD (2013) Multivalent histone engagement by the linked tandem Tudor and PHD domains of UHRF1 is required for the epigenetic inheritance of DNA methylation. *Genes & development* **27:** 1288-1298

Rothbart SB, Krajewski K, Nady N, Tempel W, Xue S, Badeaux AI, Barsyte-Lovejoy D, Martinez JY, Bedford MT, Fuchs SM, Arrowsmith CH, Strahl BD (2012) Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nature structural & molecular biology* **19:** 1155-1160

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

Rountree MR, Bachman KE, Baylin SB (2000) DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nature genetics* **25:** 269-277

Sado T, Fenner MH, Tan SS, Tam P, Shioda T, Li E (2000) X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation. *Developmental biology* **225:** 294-303

Saitou M, Kagiwada S, Kurimoto K (2012) Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. *Development* **139:** 15-31

Sampath V, Yuan P, Wang IX, Prugar E, van Leeuwen F, Sternglanz R (2009) Mutational analysis of the Sir3 BAH domain reveals multiple points of interaction with nucleosomes. *Molecular and cellular biology* **29:** 2532-2545

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

Schubeler D, Lorincz MC, Cimbora DM, Telling A, Feng YQ, Bouhassira EE, Groudine M (2000) Genomic targeting of methylated DNA: influence of methylation on transcription, replication, chromatin structure, and histone acetylation. *Molecular and cellular biology* **20:** 9103-9112

Seisenberger S, Peat JR, Hore TA, Santos F, Dean W, Reik W (2013) Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* **368:** 20110330

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

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

Song J, Teplova M, Ishibe-Murakami S, Patel DJ (2012) Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation. *Science* **335:** 709-712

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

Sugiyama Y, Hatano N, Sueyoshi N, Suetake I, Tajima S, Kinoshita E, Kinoshita-Kikuta E, Koike T, Kameshita I (2010) The DNA-binding activity of mouse DNA methyltransferase 1 is regulated by phosphorylation with casein kinase 1delta/epsilon. *The Biochemical journal* **427:** 489-497

Syeda F, Fagan RL, Wean M, Avvakumov GV, Walker JR, Xue S, Dhe-Paganon S, Brenner C (2011) The replication focus targeting sequence (RFTS) domain is a DNA-competitive inhibitor of Dnmt1. *The Journal of biological chemistry* **286:** 15344-15351

Takebayashi S, Tamura T, Matsuoka C, Okano M (2007) Major and essential role for the DNA methylation mark in mouse embryogenesis and stable association of DNMT1 with newly replicated regions. *Molecular and cellular biology* **27:** 8243-8258

Takeshita K, Suetake I, Yamashita E, Suga M, Narita H, Nakagawa A, Tajima S (2011) Structural insight into maintenance methylation by mouse DNA methyltransferase 1 (Dnmt1). *Proceedings of the National Academy of Sciences of the United States of America* **108:** 9055-9059

Tan L, Shi YG (2012) Tet family proteins and 5-hydroxymethylcytosine in development and disease. *Development* **139:** 1895-1902

Tate PH, Bird AP (1993) Effects of DNA methylation on DNA-binding proteins and gene expression. *Current opinion in genetics & development* **3:** 226-231

Tubbs JL, Pegg AE, Tainer JA (2007) DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by O6-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. *DNA repair* **6:** 1100-1115

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

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

Vilkaitis G, Suetake I, Klimasauskas S, Tajima S (2005) Processive methylation of hemimethylated CpG sites by mouse Dnmt1 DNA methyltransferase. *The Journal of biological chemistry* **280:** 64-72

Walsh CP, Chaillet JR, Bestor TH (1998) Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nature genetics* **20:** 116-117

Wang C, Shen J, Yang Z, Chen P, Zhao B, Hu W, Lan W, Tong X, Wu H, Li G, Cao C (2011) Structural basis for site-specific reading of unmodified R2 of histone H3 tail by UHRF1 PHD finger. *Cell research* **21:** 1379-1382

Winkelmann J, Lin L, Schormair B, Kornum BR, Faraco J, Plazzi G, Melberg A, Cornelio F, Urban AE, Pizza F, Poli F, Grubert F, Wieland T, Graf E, Hallmayer J, Strom TM, Mignot E (2012) Mutations in DNMT1 cause autosomal dominant cerebellar ataxia, deafness and narcolepsy. *Human molecular genetics* **21:** 2205-2210

Wu SC, Zhang Y (2010) Active DNA demethylation: many roads lead to Rome. *Nature reviews Molecular cell biology* **11:** 607-620

Xie S, Jakoncic J, Qian C (2012) UHRF1 double tudor domain and the adjacent PHD finger act together to recognize K9me3-containing histone H3 tail. *Journal of molecular biology* **415:** 318-328

Xu B, Zeng DQ, Wu Y, Zheng R, Gu L, Lin X, Hua X, Jin GH (2011) Tumor suppressor menin represses paired box gene 2 expression via Wilms tumor suppressor protein-polycomb group complex. *The Journal of biological chemistry* **286:** 13937-13944

Yang CG, Yi C, Duguid EM, Sullivan CT, Jian X, Rice PA, He C (2008) Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA. *Nature* **452:** 961-965

Yin LJ, Zhang Y, Lv PP, He WH, Wu YT, Liu AX, Ding GL, Dong MY, Qu F, Xu CM, Zhu XM, Huang HF (2012) Insufficient maintenance DNA methylation is associated with abnormal embryonic development. *BMC medicine* **10:** 26

Zhang Q, Wang HY, Marzec M, Raghunath PN, Nagasawa T, Wasik MA (2005) STAT3- and DNA methyltransferase 1-mediated epigenetic silencing of SHP-1 tyrosine phosphatase tumor suppressor gene in malignant T lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 6948-6953

Zhong H, Voll RE, Ghosh S (1998) Phosphorylation of NF-kappa B p65 by PKA stimulates transcriptional activity by promoting a novel bivalent interaction with the coactivator CBP/p300. *Molecular cell* **1:** 661-671

Ziller MJ, Muller F, Liao J, Zhang Y, Gu H, Bock C, Boyle P, Epstein CB, Bernstein BE, Lengauer T, Gnirke A, Meissner A (2011) Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS genetics* **7:** e1002389

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

Author's contribution

1. Bashtrykov P, Jankevicius G, Smarandache A, Jurkowska RZ, Ragozin S, Jeltsch A. *Specificity of Dnmt1 for methylation of hemimethylated CpG sites resides in its catalytic domain.* **Chemistry & Biology.** 2012;19(5):572-8.

   Pavel Bashtrykov (P.B.) performed site-directed mutagenesis, expression and purification of the full-length Dnmt1 mutants and conducted DNA methylation assay to determine the specificity of these Dnmt1 variants (Figure 3), as well as participated in data analysis and data interpretation. Supervised A.S.

2. Bashtrykov P, Ragozin S, Jeltsch A. *Mechanistic details of the DNA recognition by the Dnmt1 DNA methyltransferase.* **FEBS letters.** 2012;586(13):1821-3.

   P.B. performed all experiments, participated in data analysis and interpretation and prepared Figure 2 of the manuscript.

3. Ceccaldi A, Rajavelu A, Ragozin S, Senamaud-Beaufort C, Bashtrykov P, Testa N, et al. *Identification of novel inhibitors of DNA methylation by screening of a chemical library.* **ACS Chemical Biology.** 2013;8(3):543-8.

   P.B. performed the Dnmt1 inhibition assay and contributed to its analysis.

4. Bashtrykov P, Jankevicius G, Jurkowska RZ, Ragozin S, Jeltsch A.*Uhrf1 stimulates the activity and specificity of the maintenance DNA methyltransferase Dnmt1 by an allosteric mechanism.* **J. Biol. Chem.** 2014;289(7):4106-15.

   P.B. contributed to developing the concept and design of the study, performed all experiments (except the stimulation of Dnmt1 by Uhrf1 and the SRA domain (Figure 2), which was performed with the contribution of G.J.), participated in data analysis, data interpretation and in writing of the manuscript.

5. Bashtrykov P, Rajavelu A, Hackner B, Ragozin S, Carell T, Jeltsch A. *Targeted mutagenesis results in an activation of DNA methyltransferase 1 and confirms an autoinhibitory role of its RFTS domain.* **ChemBioChem.** 2014 in press.

   P.B. contributed to developing the concept and design of the study, performed all experiments (except the experiment presented in Figure 5), as well as participated in data analysis, data interpretation and in writing of the manuscript.

Appendix

Publications

Bashtrykov P, Jankevicius G, Smarandache A, Jurkowska RZ, Ragozin S, Jeltsch A. *Specificity of Dnmt1 for methylation of hemimethylated CpG sites resides in its catalytic domain.* **Chemistry & Biology.** 2012;19(5):572-8.

Bashtrykov P, Ragozin S, Jeltsch A. *Mechanistic details of the DNA recognition by the Dnmt1 DNA methyltransferase.* **FEBS letters.** 2012;586(13):1821-3.

Ceccaldi A, Rajavelu A, Ragozin S, Senamaud-Beaufort C, Bashtrykov P, Testa N, et al. *Identification of novel inhibitors of DNA methylation by screening of a chemical library.* **ACS Chemical Biology.** 2013;8(3):543-8.

Bashtrykov P, Jankevicius G, Jurkowska RZ, Ragozin S, Jeltsch A.
Uhrf1 stimulates the activity and specificity of the maintenance DNA methyltransferase Dnmt1 by an allosteric mechanism. **J. Biol. Chem.** 2014;289(7):4106-15.

Bashtrykov P, Rajavelu A, Hackner B, Ragozin S, Carell T, Jeltsch A. *Targeted mutagenesis results in an activation of DNA methyltransferase 1 and confirms an autoinhibitory role of its RFTS domain.* **ChemBioChem.** 2014 in press.