H3K36me2/3 Binding and DNA Binding of the DNA Methyltransferase DNMT3A PWWP Domain Both Contribute to its Chromatin Interaction

Michael Dukatz, Katharina Holzer, Michel Choudalakis, Max Emperle, Cristiana Lungu†, Pavel Bashtrykov and Albert Jeltsch

Department of Biochemistry, *Institute of Biochemistry and Technical Biochemistry, Stuttgart University, Allmandring 31, 70569 Stuttgart, Germany*

Correspondence to Albert Jeltsch: *Department of Biochemistry, Institute of Biochemistry and Technical Biochemistry, Allmandring 31, D-70569 Stuttgart, Germany. Fax: +49 711 685 64392.*
albert.jeltsch@ibtb.uni-stuttgart.de
https://doi.org/10.1016/j.jmb.2019.09.006
Edited by Sepideh Khorasanizadeh

Abstract

The PWWP domain of DNMT3 DNA methyltransferases binds to histone H3 tails containing methylated K36, and this activity is important for heterochromatic targeting. Here, we show that the PWWP domain of mouse DNMT3A binds to H3K36me2 and H3K36me3 with a slight preference for H3K36me2. PWWP domains have also been reported to bind to DNA, and the close proximity of H3K36 and nucleosomal DNA suggests a combined binding to H3K36me2/3 and DNA. We show here that the DNMT3A PWWP domain binds to DNA with a weak preference for AT-rich sequences and that the designed charge reversal R362E mutation disrupts DNA binding. The K295E mutation, as well as K295I recently identified in paraganglioma, a rare neuroendocrine neoplasm, disrupts both DNA and H3K36me2/3 binding, which is in agreement with the proximity of K295 to residues involved in K36me2/3 methyllysine binding. Nucleosome pulldown experiments show that DNA binding and H3K36me2/3 binding are important for the interaction of the DNMT3A PWWP domain with nucleosomes. Localization studies of transiently transfected fluorescently-tagged wild-type and PWWP-mutated full-length DNMT3A indicate that both interactions contribute to the subnuclear localization of DNMT3A in mouse cells. In summary, our data demonstrate that the combined binding of the DNMT3A PWWP domain to the H3 tail containing K36me2/3 and to the nucleosomal or linker DNA is important for its chromatin interaction and subnuclear targeting of DNMT3A in living cells.

Introduction

The N-terminal parts of the DNMT3A and DNMT3B DNA methyltransferases are necessary for the targeting of these enzymes to chromatin and for the regulation of their activity [1,2]. They contain an unstructured part, followed by a PWWP domain and an ADD domain. PWWP domains comprise about 100–130 amino acid residues and have been identified in 22 human proteins listed in the Human Protein Reference Database (HPRD) [3]. The domain name was coined based on a PWWP amino acid motif which is conserved in many PWWP domain sequences. Initially, PWWP domains have been described as potential protein/protein interaction domains [4]. Structural analysis with the PWWP domain of DNMT3B showed that it folds into an N-terminal 5-stranded antiparallel β-barrel with a C-terminal bundle of 5 α-helices docked onto it [5]. Different publications demonstrated nonspecific DNA binding of the DNMT3B [5,6] and DNMT3A [7] PWWP domains.

The PWWP domains of DNMT3A and DNMT3B were shown to be required for chromatin targeting and heterochromatic localization of both enzymes [6,8]. In 2010, it was discovered that the PWWP domain of DNMT3A binds to the N-terminal tail of histone H3 containing H3K36me3 [9]. Trimethyllysine binding was mapped to an aromatic cage similar as observed in other trimethyllysine-binding proteins [9]. Exchange of the D329 residue next to this aromatic pocket by alanine in the mouse DNMT3A

PWWP domain led to the loss of H3K36me3 binding. Follow-up studies with modified histone peptide arrays revealed additional binding activity for H3K36me2 [10–12], but the relative affinities for both methylation states of K36 have not been evaluated so far in quantitative terms. The structure of the DNMT3B PWWP domain bound to an H3 peptide containing H3K36me3 was solved in 2016 [13] and confirmed the binding pocket previously identified in the biochemical studies [9]. The aromatic cage residues in human DNMT3A are F303, W306, and W330; the mouse D329 residue corresponds to D333 in human DNMT3A. The importance of the H3K36me2/3 binding for DNMT3A targeting was demonstrated by showing that the D329A exchange led to the loss of the heterochromatic localization of DNMT3A in full-length DNMT3A fused to EYFP [9]. Cellular DNA methylation experiments revealed that K36 methylation has a targeting effect on DNA methylation and DNMT3 enzymes, which confirms the biological significance of the K36me2/3 binding by the DNMT3 PWWP domains. In DNMT3B knockout mouse ES cell lines, it was reported that intragenic DNA methylation was deposited by DNMT3B [14,15]. This activity was dependent on H3K36me3 introduced by the SETD2 protein lysine methyltransferase and on an intact PWWP domain of DNMT3B. Another important role of H3K36me3 binding to the PWWP domain was described in a study showing DNMT3A and DNMT3B associate in an H3K36me3-dependent manner to most active enhancers in epidermal stem cells, indicating that this binding is mediated by the PWWP domain as well [16]. Recently, it was shown that a knock-in of the D329A mutation into the DNMT3A gene in mice leads to dominant postnatal growth retardation and global changes in DNA methylation [17].

Chromatin binding of the PWWP domain has well-documented disease connections. Mutations in the PWWP domain of DNMT3B were observed in immunodeficiency-centromeric instability-facial anomalies syndrome (ICF) patients and were found to lead to reduced DNA methylation at pericentromeric satellite II repeats [6,8,18]. Moreover, very recently, mutations in the PWWP domain of DNMT3A that disrupted K36me3 binding were observed in microcephalic dwarfism including the W330R mutations of the conserved aromatic cage residue and the D333N mutation of the conserved aspartic acid residue [12]. Interestingly, the loss of K36me3 anchorage of DNMT3A in these patients leads to hypermethylation of DNA methylation canyons and loss of H3K4me3 and H3K27me3 in these regions, similarly as observed in mouse cells with DNMT3A D329A knock-in [17]. In addition, a K299I mutation in the PWWP domain of DNMT3A was recently discovered in the germline of patients with paraganglioma, a rare neuroendocrine neoplasm [19]. This mutation also led to global DNA hypermethylation, but its molecular effects have not yet been studied. In this work, we investigated the role and specificity of methylated H3K36 and DNA binding of the DNMT3A PWWP domain and identified mutations including K299I, which have lost one of these properties, or both. We then show that the combined H3K36me2/3 and DNA binding activities are important for *in vitro* chromatin interaction of the PWWP domain and subnuclear localization of DNMT3A in mouse cells.

Results

The location of the K36 residue on H3 in close vicinity of the DNA on the nucleosome surface (Suppl. Fig. 1A) suggests that a reading domain interacting with H3K36me3 would also come in contact with the nucleosomal or linker DNA. Hence, both functions of PWWP domains described earlier, H3K36me3 binding and DNA binding, can be functionally connected. However, so far the relative role of both interactions in chromatin binding has not been studied in detail. Based on electrostatic potential maps, putative DNA interfaces of the DNMT3A and DNMT3B PWWP domains have been predicted [5,13], but these interfaces have not been mapped experimentally for any of the enzymes. There are different patches of basic residues on the surface of the DNMT3 PWWP domains, which represent potential DNA contact points (Suppl. Fig. 1B). We selected two residues in these regions for further investigation (K295 and R362), both being conserved among human and mouse DNMT3A and DNMT3B PWWP domains (Suppl. Fig. 1C).

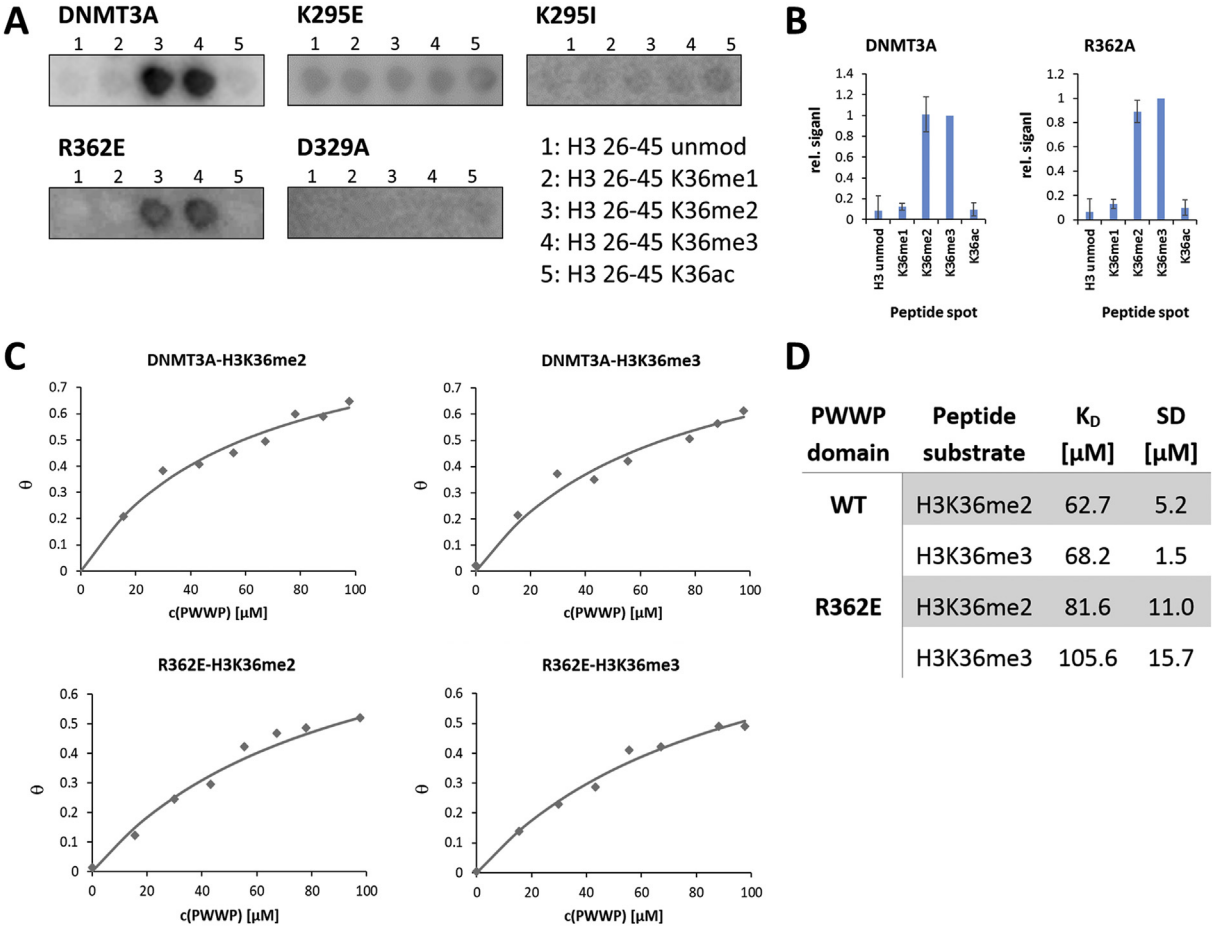
Mutagenesis, protein expression, and purification

To study the effect of the putative DNA binding mutants in the murine DNMT3A PWWP domain, we introduced charge reversal mutations at the selected positions by site-directed mutagenesis and generated K295E and R362E variants in the context of the GST-tagged PWWP domain of mouse DNMT3A. In addition, the disease associated mutation K295I (corresponding to human K299I [19]) was prepared and investigated as well. The D329A mutation affecting the methyllysine binding pocket of the DNMT3A PWWP domain was already available [9] and included in our study as well. The wild-type and mutant PWWP domains were overexpressed in *Escherichia coli* and purified by affinity chromatography (Suppl. Fig. 2A). The purified proteins were investigated by circular dichroism (CD) spectroscopy showing that all CD spectra were superimposable within the noise range (Suppl. Fig. 2B). Thermal unfolding experiments followed by the CD signal at

220 nm revealed a stable fold of all proteins with denaturation temperatures (Tm) of 59–61 °C (Suppl. Fig. 2C). Interestingly, the mutations removing a positive charge from the basic surface patches even led to a small stabilization of the folding. While these findings do not exclude local structural differences between the wild-type and mutant domains, they confirm that all mutant proteins are stably folded.

H3K36me2/3 binding of PWWP mutants

While the DNMT3A PWWP domain was initially reported to bind H3K36me3 [9], following studies with peptide arrays consistently revealed specific binding of H3K36me2 and H3K36me3 with roughly equal intensities [10–12,20]. To further investigate this point and study H3K36me2/3 binding of the PWWP mutants, binding reactions were conducted using CelluSpots peptide arrays, which contain H3 26–45 peptide spots in different modification states (Fig. 1A and B). In agreement with most recent studies, incubation of the arrays with the wild-type PWWP domain reproduced the roughly equal binding to H3K36me2 and H3K36me3. Similar preferences for binding H3K36me2 and H3K36me3 but slightly reduced binding intensities were observed with R362E indicating that this mutation does not drastically affect the H3K36me2/3 interaction. Peptide binding was lost with D329A as described previously [9–11,20]. Additionally, the K295E and K295I mutations disrupted peptide binding and recognition of K36me2/3.



PWWP domain	Peptide substrate	K_D [µM]	SD [µM]
WT	H3K36me2	62.7	5.2
	H3K36me3	68.2	1.5
R362E	H3K36me2	81.6	11.0
	H3K36me3	105.6	15.7

Fig. 1. H3K36me2/3 binding of the wild-type and mutant PWWP domains. A) Example images of wild-type and mutant PWWP domains binding to MODified™ Histone Peptide Arrays [10,11,20]. The picture shows the binding to the H3 26–45 peptide spots present on the array. **B)** Binding intensities of these spots were quantified for wild-type PWWP and R362E mutant on 4 independent arrays. The images show average binding intensities after normalization to the H3K36me3 spot intensity. Error bars display standard deviations. **C)** Examples of the titration experiments of wild-type and R362E PWWP binding to H3 27–43 peptides containing K36me2 or K36me3 studied by fluorescence polarization. The lines represent a fit of the data to a 1:1 binding equilibrium. Theta (θ) indicates the binding saturation of the peptide ($=c_{peptide,\ bound}/c_{peptide,\ total}$). **D)** Average binding constants of 2 independent experiments together with standard deviations.

To determine the peptide binding properties of the wild-type and R362E PWWP domain in a quantitative manner, equilibrium binding experiments to modified H3 27−43 peptides containing K36me2 or K36me3 were conducted by fluorescence polarization (Fig. 1C and D). Our data revealed binding of wild-type PWWP to both peptides with a slight preference for H3K36me2 (binding constants 63 μM for H3K36me2 and 68 μM for H3K36me3). The K_D determined for H3K36me3 is in good agreement with our previous data [9]. Control experiments with an unmodified H3 27−43 peptide showed no binding ($K_d > 400$ μM) (Suppl. Fig. 3). Binding of R362E was slightly weaker than wild-type (1.3 fold), but the relative preference was similar with a slightly stronger binding of K36me2.

DNA binding of PWWP wild type and mutants

DNA binding of the PWWP mutants was studied using a double-stranded 30 bp DNA oligonucleotide which was incubated with different PWWP domain variants followed by gel retardation analysis. Our data showed a strong DNA binding of the wild-type and D329A proteins, which was completely lost in the K295 and R362 mutants (Fig. 2A). Using the 30mer DNA-binding substrate after methylation of its central CpG site with M.SssI revealed no difference in the binding (Fig. 2B). We validated quantitative methylation of the target site (Suppl. Fig. 4) and concluded that DNA binding of the PWWP domain is not influenced by DNA methylation at CpG sites, which is in agreement with a similar observation made previously [7].

To study potential sequence preferences of the PWWP DNA binding in more details, a substrate containing a central CN site (N = A, T, G, or C) flanked by 10 randomized bases was prepared (Fig. 2C), incubated with the GST-tagged PWWP domain and precipitated using GST-beads. Binding reactions with the PWWP R362E GST DNA-binding mutant were conducted as control. The amount of precipitated DNA was analyzed by semiquantitative polymerase chain reaction (PCR) revealing that a PCR product was detected about 5 cycles earlier with GST-PWWP than with GST-PWWP R362E, indicating a roughly 32-fold enrichment of the DNA in the pulldown with the wild-type PWWP (Fig. 2D). The DNA was used for library preparation and subjected to Illumina sequencing. Input DNA was used as control, because the amount of DNA in the GST-PWWP R362E pulldown was insufficient for library generation. NGS provided about 612000 reads for the pulldown DNA and 866000 for the input. There was no enrichment of any base at the central CN site, indicating that the PWWP domain does not have specificity for CpG or any dinucleotide starting with C in agreement with an earlier finding based on gel shift binding studies with two substrates [7]. However, we detected a weak enrichment of A or T bases in the pulldown material when compared with input, which was even elevated if the abundance of WWW (W = A or T) or SSS (S = G or C) trinucleotides was compared (Fig. 2E). To confirm this result, a second pulldown reaction was conducted with a CH (H = A, T, or G) substrate embedded in randomized N_{10} flanks. Input and pulldown were sequenced with 370000 and 710000 reads, and the same analysis for enrichment of A and T bases and trinucleotides in the pulldown was conducted. The results confirmed the initial analysis showing a weak but statistically highly significant enrichment of AT single nucleotides (p-value $= 6.7 \times 10^{-3}$ based on the average and standard deviation [SD], assuming a normal distribution) and WWW trinucleotides (p-value $= 1.6 \times 10^{-11}$ based on the average and SD, assuming a normal distribution). The relatively weak absolute level of enrichment can be explained by the fact that in this experiment only one favored binding site must be present somewhere within the 22 base pairs with variable sequence. We conclude that the PWWP domain binds DNA with a weak preference for AT-rich sequences.

Chromatin interaction of PWWP mutants

To study the chromatin interaction of the PWWP mutants, we isolated oligonucleosomes from HepG2 cells. The preparation consisted of a mixture of mainly mono- and dinucleosomes, with some additional larger oligonucleosomes (Fig. 3A). We conducted a GST pulldown with the PWWP mutants using this oligonucleosome preparation and analyzed the pulldown with antibodies directed against unmodified H3, H3K36me2 and H3K36me3 (Fig. 3B−D). In all experiments we observed a clear pulldown with the wild-type protein, which was specific as indicated by the absence of a pulldown with the isolated GST protein and slightly stronger for H3K36me2. In contrast, the pulldown of unmodified H3 was much weaker (around 5-fold reduced) indicating a specific enrichment of H3K36me2- and H3K36me3-containing nucleosomes. Residual pulldowns (5−10 times reduced) were observed with D329A and R362A indicating that loss of H3K36me2/3 binding (in D329A) or DNA binding (in R362A) caused a reduction in the interaction of PWWP with oligonucleosomes. The observation of a reduced chromatin pulldown of D329A is in agreement with our previous data [9]. Interestingly, the K295E and K299I mutants were no longer able to precipitate oligonucleosomes indicating that the combined loss of DNA and H3K36me2/3 binding causes an aggravated phenotype. Additionally, pulldown experiments were conducted with mononucleosomes prepared by a more extensive MNase digestion (Suppl. Fig. 5). Interestingly, the PWWP

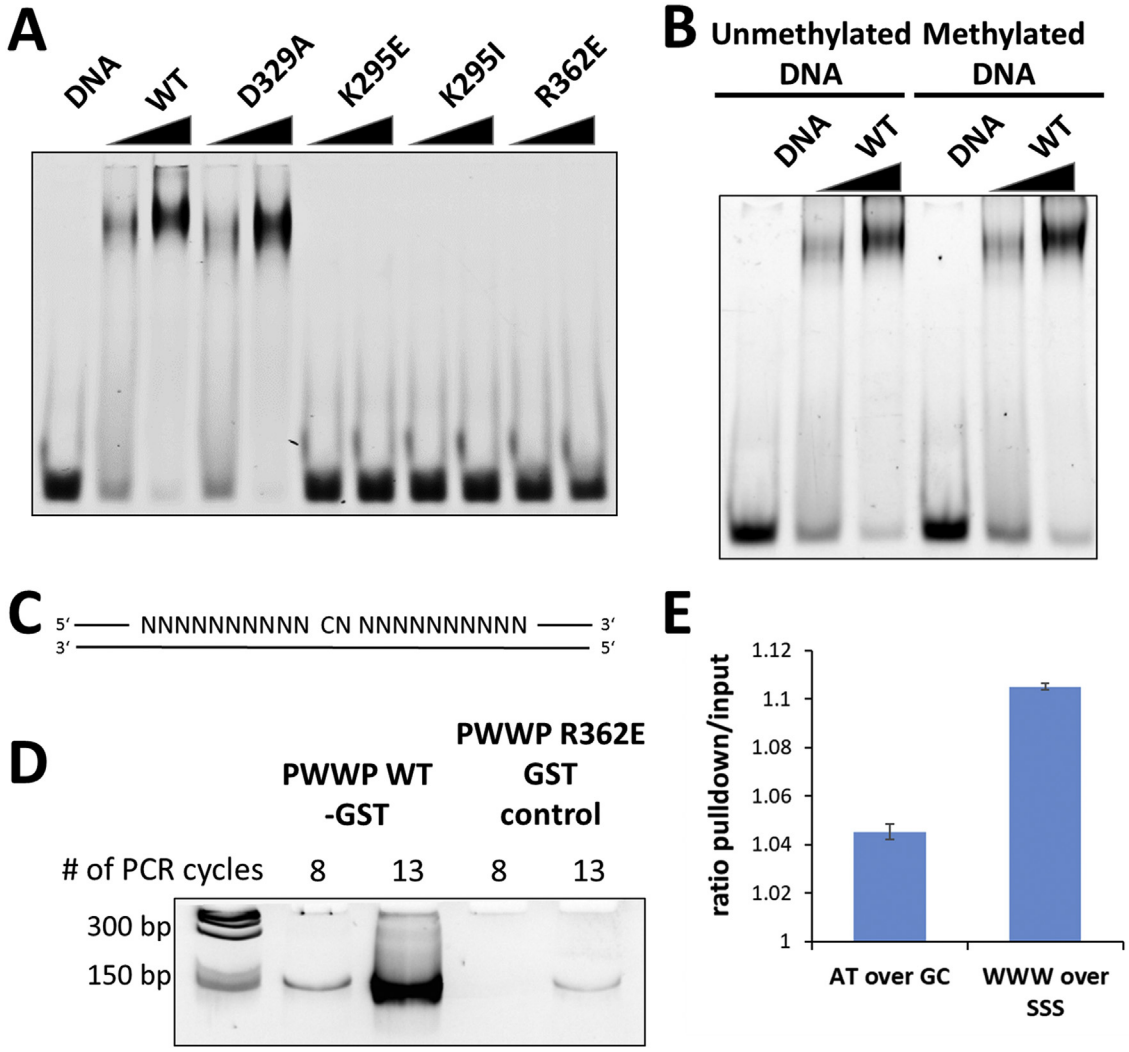


Fig. 2. DNA binding and sequence preferences of the wild-type and mutant PWWP domains. A) DNA binding of the DNMT3A PWWP domain proteins analyzed by a gel shift assay. The Cy-5-labeled 30mer double-stranded oligonucleotide was incubated with PWWP proteins in two concentrations (0.5 μM and 1.5 μM), and the sample separated on a native polyacrylamide gel. The DNA was detected by Cy5 fluorescence using a FUSION advance solo 4 (PeqLab). **B)** DNA binding assay using unmethylated and methylated DNA showing equal efficiency of DNA binding by the PWWP domain. Lanes 1–3 same as in panel A. Lanes 4–6 DNA binding analysis conducted with methylated 30mer DNA (see Suppl. Fig. 4 for validation of the methylation state of the DNA substrate). **C)** Scheme of the DNA substrate with a central CN site embedded in a flanking sequence of 10 randomized bases used for the DNA pulldown. **D)** DNA pulldown with GST-PWWP and GST control using the DNA substrate with randomized flanking sequences. The amount of precipitated DNA was analyzed by semiquantitative polymerase chain reaction (PCR). PCR was conducted for different number of cycles and the PCR product analyzed on an acrylamide gel stained with GelRed®. **E)** Ratio of A or T versus G or C nucleotides in the pulldown and input (left bar). Ratio of WWW (W = A or T) trinucleotides versus SSS (S = G or C) trinucleotides in the pulldown and input (right bar). The figure shows averages of two independent pulldown reactions and the corresponding standard deviation (SD).

domain was not able to precipitate these mononucleosomes, suggesting that the shorter linker DNA regions in the mononucleosomes were not sufficient to support PWWP binding. Altogether, our data show that both H3K36me2/3 and DNA binding are important for the chromatin interaction of the DNMT3A PWWP domain.

Influence of PWWP mutations on the subnuclear localization of DNMT3A

It is well established that DNMT3A localizes to heterochromatic regions, which are forming defined microscopically visible spots in mouse NIH3T3 cells when stained with minor groove binders with DNA

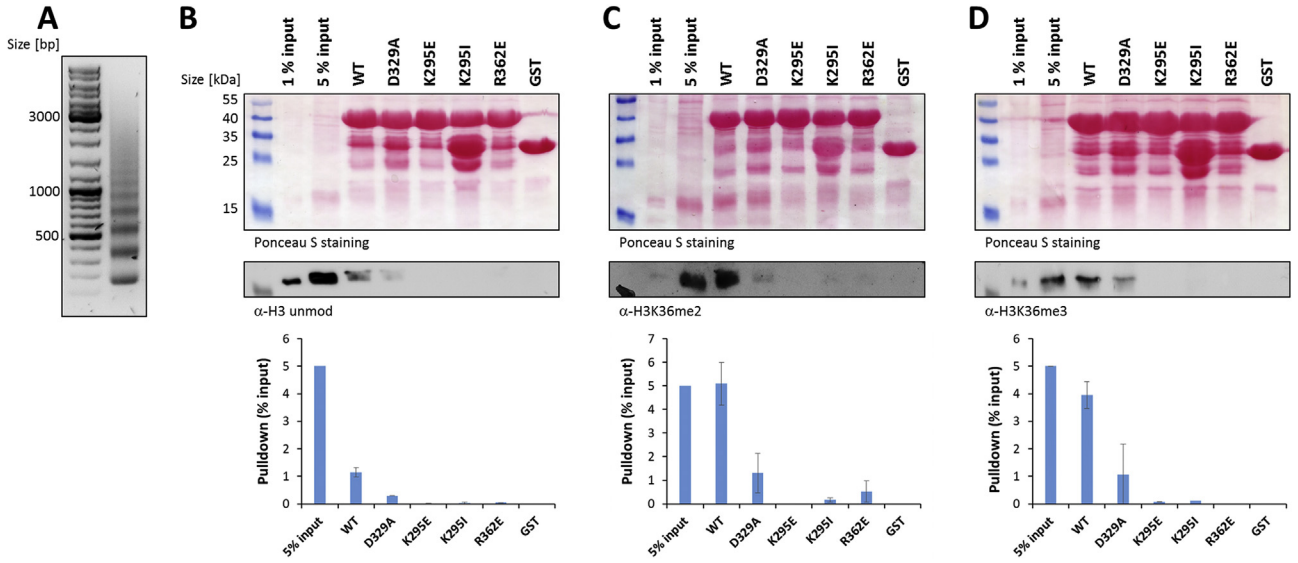


Fig. 3. Oligonucleosome binding of the wild-type and mutant PWWP domains. A) Size distribution of the DNA in oligonucleosomes prepared from HepG2 cells analyzed by agarose gel electrophoresis and stained by GelRed®. **B**–**D)** Oligonucleosomes were incubated with the wild-type and mutant PWWP domains, subjected to GST-pulldown, analyzed by sodium lauryl sulfate (SDS) gel electrophoresis. Afterward, a Western blot was conducted and the overall protein content analysed by Ponceau S staining. Next, the H3 protein in its different modification states was detected with α-H3 unmodified antibody **(B)**, α-H3K36me2 antibody **(C)**, or α-H3K36me3 antibody **(D)**. The bar diagrams show average pulldown efficiencies and standard deviations of 2 (unmodified) or 3 (K36me2 and K36me3) independent repetitions of each experiment.

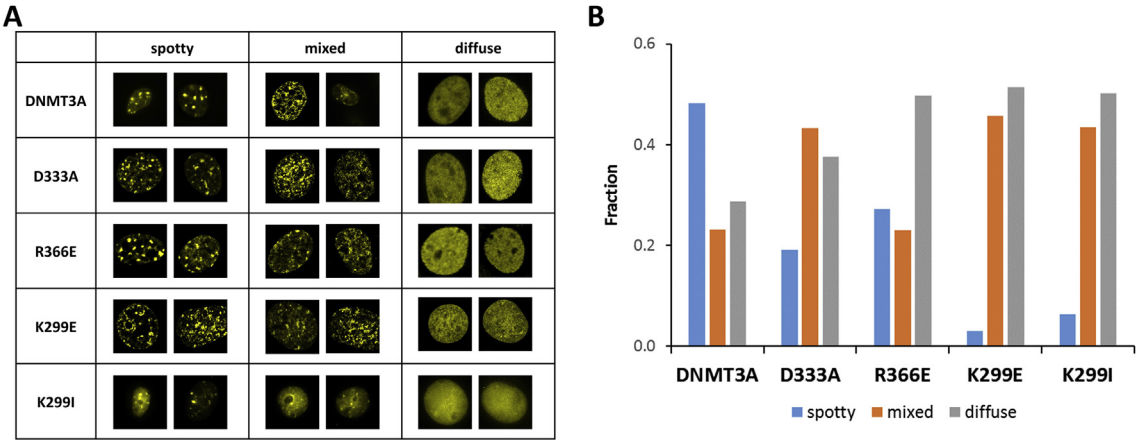


Fig. 4. Heterochromatic localization of Venus-fused human wild-type and mutant DNMT3A1 analyzed by fluorescence microscopy in NIH3T3 cells. A) Example images of cell nuclei sorted by three phenotypes of subnuclear distribution (spotty, mixed, and diffuse). The width of the images corresponds to 24 μm. **B)** Fractions of cells showing a spotty, mixed, or diffuse nuclear distribution. More than 240 cells were analyzed for each DNMT3A1 variant taken from at least three independent transfections. The p-value for the reduction in the fraction of spotty cells is 1.7×10^{-11} for R366E and $<10^{-32}$ for all other mutants (as determined by binomial testing). An analysis of the data observed in the individual experiments is shown in Suppl. Fig. 6.

binding preference for AT-rich regions like DAPI or Hoechst [6,8,9,21,22]. We showed previously that the PWWP domain is essential for this process and the D333A exchange in human DNMT3A1, which disrupts H3K36me2/3 binding, also causes a massive reduction in the heterochromatic localization of fluorophore-tagged human full-length DNMT3A1 [9]. To analyze if the disruption of DNA binding affects the subnuclear localization of DNMT3A1 as well, we introduced the K299E, K299I, and R366E mutations in the Venus-tagged DNMT3A1. Note that the numbering of human DNMT3A is applied in this section, reflecting the usage of the Venus-tagged human full-length DNMT3A1 protein in these experiments. The mouse K295, D329, and R362 residues correspond to human K299, D333, and R366. The expression vectors were transfected into NIH3T3 cells and the subnuclear localization observed by fluorescence microscopy for >240 cells in at least 3 independent experiments. As shown in Fig. 4 and Suppl. Fig. 6, the spotty subnuclear localization of wild-type DNMT3A1 observed in several previous studies [6,8,9,21,22] was reproduced. Costaining with Hoechst confirmed that the spots represent condensed heterochromatic regions (Suppl. Fig. 7) as shown previously [6,8,9,21,22]. The spotty localization of DNMT3A1 is not due to its ectopic overexpression, because it is seen in cells with very low DNMT3A1 expression as well (Suppl. Fig. 8).

Next, the subnuclear localization of the DNMT3A1 PWWP mutants was investigated (Fig. 4). As shown previously [9], the D333A mutation led to a strong reduction of the heterochromatic localization of DNMT3A1 with an increase in the number of cells with diffuse or spotty/diffuse mixed phenotype. The corresponding experiments with the R366E mutant revealed a similar redistribution of DNMT3A1. These data indicate that H3K36me2/3 and DNA binding of the PWWP domain are both important for the heterochromatic localization of DNMT3A in mouse cells, because disruption of H3K36me2/3 binding (in D333A) or DNA binding (in R366E) reduced DNMT3A1 targeting to these loci. The K295E and K299I mutants, which disrupt both DNA and H3K36me2/3 binding led to an even more pronounced almost complete loss of spotty heterochromatic localization of DNMT3A1 indicating a synergistic effect of both binding activities. Importantly, costaining with Hoechst confirmed that cells with homogenous distribution of DNMT3A1 mutants still have intact heterochromatic spots (Suppl. Fig. 9) indicating that the change in localization is indeed due to an altered targeting of the mutant DNMT3A1.

Discussion

The location of the side chain of K36 in the N-terminal tail of histone H3 is in close vicinity to the linker DNA emerging from the histone octamer and the nucleosomal DNA. This suggests that proteins contacting this residue to introduce a modification or read its methylation state will also interact with DNA. This presumption was supported by early findings indicating that, for example, the SETD2 H3K36 protein lysine methyltransferase showed much higher enzymatic activity on nucleosomal substrates than on isolated H3 protein [23]. Moreover, it had

been demonstrated that the PSIP1-PWWP domain binds to the K36 methylated histone H3 tail and nucleosomal DNA in a concerted binding reaction that is needed for high-affinity nucleosome interaction [24]. Indeed, DNA binding of the DNMT3 PWWP domains had been reported already in 2002 [5,7], but so far the connection of H3K36me2/3 and DNA binding has not been investigated for the PWWP domains of DNMT3A or DNMT3B.

Here, we determined by peptide binding and chromatin pulldown that the DNMT3A PWWP domain binds to H3K36me2 and H3K36me3 with a slight preference for H3K36me2, which is in line with a paper published during the revision phase of this manuscript showing binding of the DNMT3A PWWP domain to H3K36me2 and H3K36me3 with about 2-fold preference for K36me2 [25]. In agreement with these observations, the PWWP domain contains an incomplete aromatic cage not providing a structural basis for strong discrimination of trimethyllysine or dimethyllysine. We confirmed DNA binding by the DNMT3A PWWP domain and show that it has a weak preference for AT-rich sequences, but no specificity for CpG sites or their methylation state. Unfortunately, our data did not allow to determine further details of the sequence-dependent DNA interaction of DNMT3A PWWP including the size of the binding site and eventual preference for A and T at particular positions of the binding site. By charge reversal mutations of conserved basic residues on the surface of the DNMT3A PWWP domain, we identified PWWP mutants which lost DNA binding. The R362E mutation disrupted DNA binding, but it only very mildly affects H3K36me2/3 interaction, in agreement to the position of this mutation distant to the aromatic cage binding K36me2/3. Importantly, the R362E mutation led to a loss of chromatin interaction *in vitro* and reduction of the characteristic subnuclear localization of transiently transfected full-length DNMT3A1 in NIH3T3 cells indicating that DNA binding is important in these processes. The K295E mutation led to a loss of DNA binding and H3K36me2/3 binding, in agreement with the position of K295 next to the aromatic cage residues F299 and W302 on the same antiparallel β-sheet (Suppl. Fig. 10). By this, K295 mutations could influence the aromatic cage conformation of the PWWP domain, although K295 itself does not contact the H3 peptide or the methyllysine directly in the available structures. Consequently, the subnuclear localization of DNMT3A was even more severely disrupted by the K295E mutation than by the other two ones. Taken together, our data indicate that the combined binding of the DNMT3A PWWP domain to H3K36me2/3 and DNA contribute to its chromatin interaction.

The DNMT3A K299I mutation (corresponding to mouse K295I), which recently has been reported to occur in paraganglioma [19], behaves similarly as K295E also disrupting DNA binding and H3K36me2/3 interaction. Hence, this disease-associated mutation disrupts both critical functional properties of the PWWP domain strongly suggesting that the pathogenic mechanism of this mutation depends on the loss or weakening of DNMT3A-chromatin interaction similarly as recently shown for the W330R and D333N mutations in the K36me2/3 binding pocket of the human DNMT3A PWWP domain [12]. These results indicate that the combined DNA and H3K36me2/3 binding by the DNMT3A PWWP domain has important implications in human diseases.

Methods

Site-directed mutagenesis, protein expression, and purification

The GST-tagged PWWP domain of mouse DNMT3A (279–420 of NP_001258682) [9] and its mutated protein variants were expressed using a pGEX-6P-2 vector. Mutagenesis was performed using the megaprimer method [26] and confirmed by restriction marker analysis and DNA sequencing. Protein overexpression was conducted at 20 °C in LB medium using BL21 (DE3) Codon + RIL *E.coli* cells (Stratagene), induced with 1 mM IPTG at 0.6–0.8 OD_{600} for 14 h. The cells were collected by centrifugation and resuspended in 20 mM HEPES pH 7.5, 500 mM KCl, 0.2 mM DTT, 1 mM EDTA, and 10% glycerol and disrupted by sonication. Cell debris was removed by centrifugation at 38,000 rcf,. Then, the supernatant was passed over a glutathione Sepharose 4B resin (Amersham Biosciences) and washed with 150 bead volumes washing buffer (20 mM HEPES pH 7.5, 500 mM KCl, 0.2 mM DTT, 1 mM EDTA, and 10% glycerol). The bound proteins were eluted with washing buffer containing 40 mM glutathione and dialyzed against 20 mM HEPES pH 7.5, 200 mM KCl, 0.2 mM DTT, 1 mM EDTA, and 10% glycerol. Proteins were aliquoted and stored at −80 °C. The purity of the preparations was estimated from Coomassie stained sodium lauryl sulfate (SDS) gels to be >95% with the exception of K295I, which always contained a higher level of contaminants. The concentrations of the proteins were determined by UV spectrophotometry and confirmed by densitometric analysis of Coomassie BB stained SDS–polyacrylamide gels.

Circular dichroism spectroscopy

Folding of the purified proteins (10 μM) was analyzed by CD spectroscopy using a J-815 CD spectrophotometer (JASCO Corporation, Tokyo, Japan) in buffer containing 200 mM KCl, 10 mM HEPES pH 7.5, 5% glycerol. The spectra were collected at 20 °C using a 0.1 mm cuvette in a wavelength range between 190 and 250 nm using a scanning speed of 100 nm/min, bandwidth of 1 nm, and data integration time of 1 s. For each sample, 30 scans were collected and averaged. Thermal denaturation was determined by heating the samples from 20 °C to 80 °C with a rate of 2 °C per minute. Data were measured at 220 nm using a

bandwidth of 5 nm and data integration time of 16 s. Melting curves were analyzed as described [27].

Nucleosome isolation

Native nucleosomes were isolated from HepG2 cells by micrococcal nuclease digestion of intact nuclei obtained as described [28]. Briefly, around 20 million cells were resuspended in 5 mL TM 2 buffer (10 mM Tris−HCl pH 7.4, 2 mM $MgCl_2$, 0.5 mM PMSF) supplemented with protease inhibitors (cOmplete ULTRA Tablets, Mini, EDTA-free, Easy pack from Roche, 1 tablet for 15 mL of buffer) and 150 μL of 20% IGEPAL®CA-630 and incubated on ice for 5 min. The lysed cells were centrifuged at 1000 rcf for 10 min at 4 °C, and the pellet was washed with TM 2 buffer. After a second centrifugation step, the pellet containing the cell nuclei was resuspended in TM 2 buffer (600−800 μL for 20×10^6 cells), transferred into an Eppendorf tube, prewarmed at 37 °C, and supplemented with 1 mM $CaCl_2$. To obtain oligonucleosomes, the nuclei were treated with 5 units MNase (New England Biolabs M0247S) per microgram of DNA for 10 min at 37 °C with gentle shaking. The reaction was stopped by adding EGTA to a final concentration of 2 mM. NaCl was added to a final concentration of 300 mM, and the buffer was supplemented with 0.1% Triton X-100. Afterward, the nucleosomes were centrifuged at 13,000 rcf for 10 min at 4 °C. The supernatant containing the soluble oligonucleosomes was collected and immediately used for further experiments. For preparation of mononucleosomes, 10 units of MNase per microgram DNA were used.

Nucleosomal pulldown

For preclearing, 20 μL of glutathione Sepharose 4B beads (GE Healthcare) were washed 3 times with 200 μL DP buffer (16.7 mM Tris/HCl pH8, 100 mM NaCl, 1.1% Triton X-100, 1.2 mM EDTA) supplemented with protease inhibitors (cOmplete ULTRA Tablets, Roche) and incubated with 30 μg of the nucleosomes (based on DNA absorbance) in a final volume of 500 μL DP buffer for 1 h at 8 °C with rotation. Afterward, the beads were spun down for 2 min with 2000 rcf and the precleared supernatant was collected. In parallel, the GST-tagged DNMT3A PWWP domains were bound to equilibrated beads in 200 μL DP-buffer for 15 min at 8 °C. The beads were spun down for 2 min with 2000 rcf and washed 3 times with 200 μL DP-buffer. The precleared nucleosomes were added to the beads with the bound protein and incubated overnight at 8 °C with constant rotation. The beads were washed 3 times with 1 mL PB-buffer (50 mM Tris/HCl pH 8, 100 mM NaCl, 2 mM DTT, 1 mM EDTA, 0.5% Nonidet-P40) and once with 1 mL TE buffer (10 mM Tris/HCl pH 8, 1 mM EDTA). Afterward the beads were incubated at 95 °C with 20 μL 5× LAP (160 mM Tris/HCl pH 6.8, 2% SDS, 5% β-mercaptoethanol, 40% glycerol, and 0.1% bromphenol blue) and loaded on a 15% SDS-polyacrylamide gel. After the gel run, the proteins were blotted onto a nitrocellulose membrane and stained with Ponceau-S. The nitrocellulose membrane was blocked with 5% milk TBS-T (100 mM Tris/HCl pH 7.5, 154 mM NaCl, 0.1% Tween®20) at room temperature for 60 min and washed 3 times with TBS-T for 5 min. For the binding of the primary antibodies (H3 unmodified: abcam ab 1791, 1:5000; H3K36me2: Active Motif 39255, 1:2500; H3K36me3: Abcam ab9050, 1:1000), the membrane was incubated with the antibody in 1% milk TBS-T overnight at 8 °C with gentle shaking. On the next day, the membrane was washed 4 times for 5 min with TBS-T at room temperature followed by incubation with the secondary antibody (ECLTM Anti-Rabbit IgG, Horseradish, GE Healthcare NA934, 1:5000) in 1% milk TBS-T for 60 min at room temperature. The membrane was washed 3 times for 5 min with TBS-T and doused with ECL solution (SuperSignalTM West Femto, Thermo Scientific). Images were captured with FUSION advance solo 4 (PeqLab) under dynamic conditions. Data were quantified using ImageJ.

Gel shifts

Different concentrations of the DNMT3A PWWP domain variants were incubated for 20 min at room temperature with 0.05 μM Cy5-labeled DNA 30mer in buffer (20 mM HEPES, 1 mM EDTA) supplemented with bovine serum albumin (BSA) (1 mg/mL). The 30mer sequence contains one centrally positioned CpG site.

Sequence: Cy5-GAAGCTGGGACTTC**CG**GGAGG AGAGTGCAA

Samples were separated on a 5% acrylamide gels run in 0.5 × TB-buffer (45 mM Tris, 45 mM boric acid, pH 7.5) for 35 min at constant voltage of 110 V. Afterward, the gel was exposed with UV light in a FUSION advance solo 4 (PeqLab) and pictures were captured after filtering with the F-695Y interference filter.

To methylate the DNA binding substrate, 3 μM of the Cy5-labeled 30mer DNA was incubated with 7 μM M.SssI (NEB) in 1 × NEBufferI supplemented with 100 ng/μL BSA (NEB) and 1 mM S-Adenosyl-L-methionine for 3 h at 37 °C. The purification of the methylated DNA was conducted following the protocol #4.1 from Ref. [29] except that the binding to the beads was conducted in a 5 times larger buffer volume than proposed in the protocol.

Live-cell imaging

To determine the cellular localization of full-length human DNMT3A1, the localization of Venus-tagged DNMT3A1 in mouse NIH3T3 cells was investigated basically as described [9,21,22]. NIH3T3 cells were seeded in 35-mm FluoroDish cell culture dishes (World Precision Instruments) and transfected with the corresponding DNMT3A1 expression plasmid using Fugene HD. Imaging was performed 24 h after transfection either using a confocal microscope (LSM 710 Zeiss) or a Cell observer (Zeiss), both equipped with a Plan-Apochromat 63 ×/1.40 Oil DIC M27 objective and an XL-LSM 710 S1 incubation chamber for temperature and CO_2 control. Cells were manually subdivided into phenotypes and counted. *P*-values were calculated with Excel assuming a normal distribution. Live cell staining of heterochromatic spots was conducted using Hoechst 33258.

Peptide array binding

For the binding of the DNMT3A PWWP domain variants to MODified™ Histone Peptide Arrays (Active Motif), the

arrays were washed 2 times 5 min with TBS-T (100 mM Tris/HCl pH 7.5, 154 mM NaCl, 0.1% Tween®20) and blocked in 5% milk TBS-T at 8 °C overnight. Washing of the array was done 3 times 5 min with TBS-T and once with interaction buffer (100 mM KCl, 20 mM HEPES pH 7.5, 1 mM EDTA, 0.1 mM DTT, 10% glycerol). Then, 1 μM of the GST-tagged DNMT3A PWWP variants were incubated on the peptide arrays in interaction buffer for 2 h at room temperature. After the incubation, the membrane was washed 3 times 5 min with TBS-T. The primary anti-GST antibody (1:5000 in TBS-T containing 1% skim milk, GE Healthcare Life Science 27-4577-01, lot 9541184) was incubated for 1 h at room temperature; excess of the antibody was removed by washing 3 times 5 min with TBS-T. The peptide array then was incubated with the horseradish peroxidase conjugated anti-goat antibody (1% milk TBS-T, 1:5000, Sigma-Aldrich A4174, lot 071M4767). Finally, the peptide arrays were washed 4 times with TBS-T and doused with ECL solution (Thermo Fisher Scientific). Images were captured with FUSION advance solo 4 (PeqLab) under dynamic conditions and analyzed using ImageJ.

Equilibrium peptide binding

Equilibrium peptide binding experiments were conducted using H3 (27-43) peptides that contained K36 unmodified, di- or trimethylated and were labeled with FITC or Fluorescein. Binding was studied using a Jasco FP-8300 spectrofluorometer with automatic polarizer FDP-837. Acquisitions were performed at 23 °C, with excitation at 493.0 nm and emission measured at 516.6 nm. Slit width was set to 5 nm for both. 1 μM peptide was dissolved in 125 μL of anisotropy buffer (20 mM HEPES pH 7.5, 100 mM KCl, 0.1 mM DTT, 10% v/v glycerol, passed through a 0.2 μm filter), and the PWWP domain dissolved in dialysis buffer (20 mM HEPES pH 7.5, 200 mM KCl, 1 mM Na_2-EDTA, 60% v/v glycerol, 0.2 mM DTT) was added stepwise. Each acquisition was conducted in 3 technical repeats. Control experiments were conducted with dialysis buffer, and the fluorescence anisotropy values were corrected accordingly. For determination of the K_d-values, the data were fitted to a simple binding equilibrium:

$$\text{Signal} = \text{BL} + \text{F} \times (c_{PWWP}/(c_{PWWP} + K_d))$$

with K_d, equilibrium binding constant; F, signal factor; BL, baseline.

DNA pulldown with the GST-PWWP domain

The DNA substrate for the pulldown was a 67-mer dsDNA with a CpN-site positioned in the center and flanked by 10 randomized bases in both directions. This substrate was generated by a primer-extension reaction using a 67mer single-stranded oligodeoxynucleotide (ordered from IDT) and a primer complementary to its 3′-end and purified with Nucleospin® Gel and PCR Clean-up Kit (Macherey-Nagel). For the GST-pulldown, 0.3 μM DNMT3A PWWP WT and the DNMT3A PWWP R362E were incubated with 0.05 μM DNA substrate in interaction buffer (20 mM HEPES pH 7.5, 1 mM EDTA, 50‰ v/v Tween®20) in a total volume of 100 μL for 20 min at room temperature. The reaction mixtures were added to 20 μL glutathione Sepharose 4B beads (GE Healthcare), which were equilibrated before with wash buffer (20 mM HEPES pH 7.5, 1 mM EDTA, 50‰ v/v Tween®20, 0.36 mM Tris, 6 mM KCl) and incubated with rotation for 30 min at room temperature. Afterward, the beads were washed 3 times for 10 min with 100 μL wash buffer. After washing, 50 μL wash buffer containing 1 μL Proteinase K (NEB, P8107S) was added to the beads and incubated at 50 °C for 30 min. Afterward, the beads were centrifuged for 2 min at 16,000 rcf. The released DNA in the supernatant was purified with NucleoSpin® Gel and PCR Clean-up Kit and eluted with 15 μL elution buffer NE. To test the pulldown efficiency, the eluted fractions were amplified with 2 primers in a PCR reaction with Q5® High-Fidelity DNA Polymerase for 8 and 13 cycles. Samples were loaded on 10% acrylamide TPE gel and separated for 50 min at 110 V constant.

Library generation and next-generation sequencing

Libraries for next-generation sequencing (NGS) of the DNA pulled down by DNMT3A PWWP and the input DNA were generated using a 2-step PCR method (Illumina). For this, input DNA was diluted to similar concentrations as pulldown. Then, the first PCR was conducted with pulldown and diluted input under same conditions as described above for 15 PCR cycles. The 146 bp−long barcoded DNA fragments were purified with NucleoSpin® Gel and PCR Clean-up Kit and eluted with 15 μL elution buffer NE. The eluted DNA fragments were diluted to 1 ng/μL and used as template for a second PCR with the primers that introduce indices with S7 Fusion High-Fidelity DNA Polymerase for 8 cycles, finally generating 215 bp−long PCR products. The samples were loaded on 6% acrylamide TBE gel and separated for 70 min at 90 V constant. The bands at the target size were cut out of the gel and the gel smashed into small pieces. Then, 700 μL ddH_2O was added and incubated at 8 °C overnight with constant rotation. The mixture was spun down with 16,000 rcf for 3 min and the supernatant was transferred into a new low-binding 1.5 mL microcentrifugation tube. Purification of the DNA samples was performed with AMPure XP for PCR Purification (Beckman Coulter) according to manufacturers' protocol. Produced libraries were sequenced in the Max Planck-Genome-Centre Cologne. After Illumina sequencing, the variable parts of the sequences were extracted from the sequence reads using the BWA-MEM and FASTQ Trimmer tools in Galaxy [30]. Next, the occurrences of bases, dinucleotide and trinucleotide sequences at all positions in the variable parts were analyzed in the sequences detected in the pulldown and input using an in-house Visual Basic program.

Acknowledgments

This work has been supported by the DFG (JE 252/10).

Author contributions

AJ devised the project. MD conducted the biochemical studies with contributions of KH. MC did the equilibrium peptide binding experiments. PB contributed to bioinformatics. AJ, ME, CL, and PB provided technical advice and supervised the research. AJ and MD wrote the manuscript draft and prepared the figures. All authors contributed to data interpretation and discussion and read and approved the final manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2019.09.006.

Received 5 March 2019;
Received in revised form 8 September 2019;
Accepted 9 September 2019
Available online 18 October 2019

Keywords:
DNA methylation;
DNA methyltransferase;
Chromatin binding;
PWWP domain;
DNMT3A

†Present address: Institute of Cell Biology and Immunology, Stuttgart University, Allmandring 31, 70569 Stuttgart, Germany.

References

[1] A. Jeltsch, R.Z. Jurkowska, Allosteric control of mammalian DNA methyltransferases – a new regulatory paradigm, Nucleic Acids Res. 44 (2016) 8556–8575.
[2] H. Gowher, A. Jeltsch, Mammalian DNA methyltransferases: new discoveries and open questions, Biochem. Soc. Trans. 46 (2018) 1191–1202.
[3] T.S. Prasad, K. Kandasamy, A. Pandey, Human Protein Reference Database and Human Proteinpedia as discovery tools for systems biology, Methods Mol. Biol. 577 (2009) 67–79.
[4] I. Stec, S.B. Nagl, G.J. van Ommen, J.T. den Dunnen, The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? FEBS Lett. 473 (2000) 1–5.
[5] C. Qiu, K. Sawada, X. Zhang, X. Cheng, The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds, Nat. Struct. Biol. 9 (2002) 217–224.
[6] T. Chen, N. Tsujimoto, E. Li, The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin, Mol. Cell. Biol. 24 (2004) 9048–9058.
[7] M.M. Purdy, C. Holz-Schietinger, N.O. Reich, Identification of a second DNA binding site in human DNA methyltransferase 3A by substrate inhibition and domain deletion, Arch. Biochem. Biophys. 498 (2010) 13–22.
[8] Y.Z. Ge, M.T. Pu, H. Gowher, H.P. Wu, J.P. Ding, A. Jeltsch, et al., Chromatin targeting of de novo DNA methyltransferases by the PWWP domain, J. Biol. Chem. 279 (2004) 25447–25454.
[9] A. Dhayalan, A. Rajavelu, P. Rathert, R. Tamas, R.Z. Jurkowska, S. Ragozin, et al., The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation, J. Biol. Chem. 285 (2010) 26114–26120.
[10] I. Bock, S. Kudithipudi, R. Tamas, G. Kungulovski, A. Dhayalan, A. Jeltsch, Application of Celluspots peptide arrays for the analysis of the binding specificity of epigenetic reading domains to modified histone tails, BMC Biochem. 12 (2011) 48.
[11] G. Kungulovski, I. Kycia, R. Tamas, R.Z. Jurkowska, S. Kudithipudi, C. Henry, et al., Application of histone modification-specific interaction domains as an alternative to antibodies, Genome Res. 24 (2014) 1842–1853.
[12] P. Heyn, C.V. Logan, A. Fluteau, R.C. Challis, T. Auchynnikava, C.A. Martin, et al., Gain-of-function DNMT3A mutations cause microcephalic dwarfism and hypermethylation of Polycomb-regulated regions, Nat. Genet. 51 (2019) 96–105.
[13] G. Rondelet, T. Dal Maso, L. Willems, J. Wouters, Structural basis for recognition of histone H3K36me3 nucleosome by human de novo DNA methyltransferases 3A and 3B, J. Struct. Biol. 194 (2016) 357–367.
[14] T. Baubec, D.F. Colombo, C. Wirbelauer, J. Schmidt, L. Burger, A.R. Krebs, et al., Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation, Nature 520 (2015) 243–247.
[15] F. Neri, S. Rapelli, A. Krepelova, D. Incarnato, C. Parlato, G. Basile, et al., Intragenic DNA methylation prevents spurious transcription initiation, Nature 543 (2017) 72–77.
[16] L. Rinaldi, D. Datta, J. Serrat, L. Morey, G. Solanas, A. Avgustinova, et al., Dnmt3a and Dnmt3b associate with enhancers to regulate human epidermal stem cell homeostasis, Cell Stem Cell 19 (2016) 491–501.
[17] G. Sendzikaite, C.W. Hanna, K.R. Stewart-Morgan, E. Ivanova, G. Kelsey, A DNMT3A PWWP mutation leads to methylation of bivalent chromatin and growth retardation in mice, Nat. Commun. 10 (2019) 1884.
[18] H. Shirohzu, T. Kubota, A. Kumazawa, T. Sado, T. Chijiwa, K. Inagaki, et al., Three novel DNMT3B mutations in Japanese patients with ICF syndrome, Am. J. Med. Genet. 112 (2002) 31–37.
[19] L. Remacha, M. Curras-Freixes, R. Torres-Ruiz, F. Schiavi, R. Torres-Perez, B. Calsina, et al., Gain-of-function mutations in DNMT3A in patients with paraganglioma, Genet. Med. 20 (12) (2018) 1644–1651.
[20] R. Mauser, G. Kungulovski, D. Meral, D. Maisch, A. Jeltsch, Application of mixed peptide arrays to study combinatorial readout of chromatin modifications, Biochimie 146 (2018) 14–19.
[21] R.Z. Jurkowska, A. Rajavelu, N. Anspach, C. Urbanke, G. Jankevicius, S. Ragozin, et al., Oligomerization and binding of the Dnmt3a DNA methyltransferase to parallel DNA molecules: heterochromatic localization and role of Dnmt3L, J. Biol. Chem. 286 (2011) 24200–24207.
[22] A. Rajavelu, R.Z. Jurkowska, J. Fritz, A. Jeltsch, Function and disruption of DNA methyltransferase 3a cooperative

DNA binding and nucleoprotein filament formation, Nucleic Acids Res. 40 (2012) 569–580.
[23] X.J. Sun, J. Wei, X.Y. Wu, M. Hu, L. Wang, H.H. Wang, et al., Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase, J. Biol. Chem. 280 (2005) 35261–35271.
[24] R. van Nuland, F.M. van Schaik, M. Simonis, S. van Heesch, E. Cuppen, R. Boelens, et al., Nucleosomal DNA binding drives the recognition of H3K36-methylated nucleosomes by the PSIP1-PWWP domain, Epigenetics Chromatin 6 (2013) 12.
[25] D.N. Weinberg, S. Papillon-Cavanagh, H. Chen, Y. Yue, X. Chen, K.N. Rajagopalan, et al., The histone mark H3K36me2 recruits DNMT3A and shapes the intergenic DNA methylation landscape, Nature 573 (7773) (2019) 281–286.
[26] A. Jeltsch, T. Lanio, Site-directed mutagenesis by polymerase chain reaction, Methods Mol. Biol. 182 (2002) 85–94.
[27] N.J. Greenfield, Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions, Nat. Protoc. 1 (2006) 2527–2535.
[28] S. Kasinathan, G.A. Orsi, G.E. Zentner, K. Ahmad, S. Henikoff, High-resolution mapping of transcription factor binding sites on native chromatin, Nat. Methods 11 (2014) 203–209.
[29] P. Oberacker, P. Stepper, D.M. Bond, S. Hohn, J. Focken, V. Meyer, et al., Bio-On-Magnetic-Beads (BOMB): open platform for high-throughput nucleic acid extraction and manipulation, PLoS Biol. 17 (2019), e3000107.
[30] E. Afgan, D. Baker, M. van den Beek, D. Blankenberg, D. Bouvier, M. Cech, et al., The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update, Nucleic Acids Res. 44 (2016) W3–W10.