Journal of Structural Biology 194 (2016) 357–367

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Structural basis for recognition of histone H3K36me3 nucleosome by human *de novo* DNA methyltransferases 3A and 3B

Grégoire Rondelet [a,*], Thomas Dal Maso [a], Luc Willems [b], Johan Wouters [a]

[a] Department of Chemistry, University of Namur, 61 rue de Bruxelles, B-5000 Namur, Belgium
[b] Molecular and Cellular Epigenetics (GIGA) and Molecular Biology (Gembloux Agro-Bio Tech), University of Liège (ULg), 4000 Liège, Belgium

ARTICLE INFO

Article history:
Received 27 November 2015
Received in revised form 14 March 2016
Accepted 15 March 2016
Available online 15 March 2016

Keywords:
DNA methylation
Methyltransferases
DNMT3A
DNMT3B
PWWP
H3K36me3
Structure
Nucleosome

ABSTRACT

DNA methylation is an important epigenetic modification involved in chromatin organization and gene expression. The function of DNA methylation depends on cell context and is correlated with histone modification patterns. In particular, trimethylation of Lys36 on histone H3 tail (H3K36me3) is associated with DNA methylation and elongation phase of transcription. PWWP domains of the *de novo* DNA methyltransferases DNMT3A and DNMT3B read this epigenetic mark to guide DNA methylation. Here we report the first crystal structure of the DNMT3B PWWP domain–H3K36me3 complex. Based on this structure, we propose a model of the DNMT3A PWWP domain–H3K36me3 complex and build a model of DNMT3A (PWWP-ADD-CD) in a nucleosomal context. The trimethylated side chain of Lys36 (H3K36me3) is inserted into an aromatic cage similar to the "Royal" superfamily domains known to bind methylated histones. A key interaction between trimethylated Lys36 and a conserved water molecule stabilized by Ser270 explains the lack of affinity of mutated DNMT3B (S270P) for the H3K36me3 epigenetic mark in the ICF (Immunodeficiency, Centromeric instability and Facial abnormalities) syndrome. The model of the DNMT3A-DNMT3L heterotetramer in complex with a dinucleosome highlights the mechanism for recognition of nucleosome by DNMT3s and explains the periodicity of *de novo* DNA methylation.

1. Introduction

DNA methylation occurs at CpG dinucleotides in mammalian cells and is catalyzed by DNA methyltransferases (DNMTs). DNMT3A and DNMT3B are both involved in *de novo* methylation establishing the methylation pattern in genome during embryogenesis while DNMT1 maintains the pattern during chromosome replication. DNA methylation is essential for cell differentiation and development but is also involved in pathologies like cancer (Bird, 2002). The function of DNA methylation depends on cell context and is correlated with histone modification patterns (Jones, 2012). In particular, trimethylation of Lys36 on histone H3 tail (H3K36me3) is associated with gene body methylation in embryonic stem (ES) cells and elongation phase of transcription (Baubec et al., 2015; Hahn et al., 2011; Lee and Shilatifard, 2007). Both DNMT3A and DNMT3B PWWP domains read H3 tails containing H3K36me3 to guide DNA methylation (Baubec et al., 2015; Dhayalan et al., 2010). A point mutation in the DNMT3B PWWP domain (S270P) leads to loss of recognition with H3K36me3-modified nucleosome (Baubec et al., 2015; Ge et al., 2004) and a decrease in DNA methylation at pericentromeric satellite repeat II as observed for ICF syndrome (Chen et al., 2004; Shirohzu et al., 2002).

DNMT3A and DNMT3B possess a C-terminal catalytic domain (CD) and an N-terminal part with a regulatory function mediated by the ADD (ATRX–DNMT3–DNMT3L) domain and a nucleosome recognition PWWP fold. Similar to the chromodomain, MBT and Tudor domains, the PWWP domain is a member of the "Royal" superfamily domains which recognize methylated histone tails through a conserved aromatic cage (Qin and Min, 2014). The H3K36me3-binding ability was established for the PWWP domains of BRPF1 (Vezzoli et al., 2010), DNMT3A (Dhayalan et al., 2010), DNMT3B (Baubec et al., 2015), PSIP1, MSH6, ZMYND11 + H3.3K36me3 (Qin and Min, 2014), LEDGF (Pradeepa et al., 2012), and Tudor domains of PHF1 and PHF19 (Ballaré et al., 2012; Musselman et al., 2012). The PWWP domain contains an anti-parallel β-barrel-like fold formed by five β-strands (β1–β5) (Fig. 1A), where a short 3_{10} helix is found between β4 and β5 (η2), an insertion motif of different lengths between β2 and β3 (η1), and a C-terminal helix bundle of 1–6 α-helices (Qiu et al., 2002). PWWP domains of DNMT3A and DNMT3B are characterized by a short motif insertion (η1) and five α-helices following the β-barrel. The conserved Pro-Trp-Trp-Pro (PWWP) motif becomes

* Corresponding author.
E-mail address: gregoire.rondelet@unamur.be (G. Rondelet).

http://dx.doi.org/10.1016/j.jsb.2016.03.013
1047-8477/

SWWP and is found in the β2 strand. DNMT3A and DNMT3B PWWP domains would synergistically bind both histone and DNA through their conserved aromatic cage for the recognition of H3K36me3 epigenetic mark, and a positively charged surface that interacts with DNA (Qin and Min, 2014; Qiu et al., 2002). Recently, DNMT3B was shown to be involved in the selective targeting of transcribed genes in mouse stem cells (Baubec et al., 2015). This association occurs through the binding of DNMT3B PWWP domain to trimethylated lysine 36 on histone H3.

Here we report the first crystal structure of the DNMT3B PWWP domain in complex with histone H3K36me3. Based on this structure, we propose a model for the DNMT3A PWWP domain–H3K36me3 complex and predict a structure of DNMT3A in a nucleosomal context to propose a mechanism of nucleosome recognition and *de novo* DNA methylation.

2. Materials and methods

2.1. Cloning, expression and purification of DNMT3B PWWP domain

The plasmid coding for the N-terminal human PWWP domain of DNMT3B was obtained from Addgene (plasmid 32044, C. Arrowsmith) as a bacterial stab culture. The plasmid containing the pET28-MHL vector codes for an N-terminal His_6-tagged fusion protein with integrated TEV protease site. Plasmids were recovered from kanamycin-resistant colonies and purified following the Addgene protocol. Plasmid was sent for sequencing to Beckman Coulter Genomics (Hope end, Takeley CM22 6TA, Essex United Kingdom). Plasmid was transformed into *Escherichia coli* strain Rosetta™ 2 (DE3) competent cells (Novagen®). Cells were grown in TB medium containing 50 μg/ml kanamycin and 34 μg/ml chloramphenicol at 37 °C. When the OD 600 nm was about 1.5, the temperature was reduced to 18 °C and expression of PWWP domain was induced with 1 mM IPTG overnight. The cells were recovered by centrifugation at 7000 rpm for 30 min at 4 °C. The harvested cells were resuspended in ice-cold lysis buffer (50 mM HEPES, pH 7.4, 500 mM NaCl, 2 mM β-mercaptoethanol, 5% glycerol, 0,1% CHAPS) with protease inhibitor cocktail (Complete Mini EDTA-Free, 1 tablet for 10 mL of solution, Roche). Cells were lysed on ice by sonication using a cell disrupter. Crude extract were centrifuged at 9000 rpm for 30 min at 4 °C to recover the lysate. The supernatant was filtered through a 0.2 μm filter (Whatman® FP 30/0.2, GE Healthcare) and loaded onto a 5 ml HisTrap™ FF crude Ni-Chelating column (GE Healthcare) equilibrated with 20 mM HEPES, pH 7.4, 500 mM NaCl, 5% glycerol, 50 mM imidazole. The charged column was washed with the equilibration buffer until a stable baseline was attained and the elution of PWWP domain of DNMT3B was performed with a linear gradient from 50 mM to 1 M imidazole over a total volume of 50 ml. After pooling the appropriate fractions, the solution was dialyzed overnight against 20 mM HEPES, pH 7.4 and 250 mM NaCl in presence of TEV protease (50 μg/mg of protein) to cleave the His-tag. After this, the solution was reloaded onto a 5 ml HiTrap™ IMAC FF Ni-Chelating column (GE Healthcare) and the column was washed with 10 column volumes of 20 mM HEPES, pH 7.4, 500 mM NaCl, 5% glycerol and 50 mM imidazole to elute the cleaved protein. Then, the protein was dialyzed against equilibration buffer (20 mM PIPES, pH 6.5). The sample was filtered through a 0.2 μm filter (Whatman® FP 30/0.2, GE Healthcare) and was loaded into an AKTA FPLC system (GE-Amersham Biosciences) with a cation exchange column (1 ml, Resource™ S, GE Healthcare) to purify to homogeneity. The charged column was washed with the equilibration buffer until a stable baseline was attained and the elution of the protein was performed with a linear gradient of NaCl up to 1 M concentration over a total volume of 20 ml. The solution was concentrated in a stirred cell using a 5 kDa MWCO membrane (Vivaspin 15R; Sartorius Stedim Biotech GmbH, Göttingen, Germany) up to 22 mg/mL prior to crystallization. All the different stages of expression and purification were followed by SDS–PAGE.

2.2. H3K36me3 peptide production and purification

H3K36me3 peptide (SAPATGGV{K(Me3)}KPHRYR) 28–42 was purchased from GenScript (Piscataway, NJ, USA) with a purity up to 95% using reverse-phase HPLC.

2.3. Crystallization and structure determination of DNMT3B PWWP domain in complex with histone H3K36me3

Crystallization trials were performed using the Hampton Research Crystal Screen kit 1–2 and the sitting-drop vapor-diffusion method in 96-well plates at room temperature. Drops consisted of 1 μl of DNMT3B PWWP domain at 22 mg/ml – peptide mixture (1:5 M ratio) plus 1 μl of reservoir solution equilibrated against a reservoir volume of 50 μl. Crystals of complex between DNMT3B PWWP domain and H3K36me3 peptide were grown against a reservoir consisting of 1.6 M Sodium citrate tribasic pH 6.5 (Hampton Research Crystal Screen 2 No. 28). A single-crystal (crystal size: 0.45 mm × 0.4 mm) was cryoprotected in 20% glycerol and flash-cooled in liquid nitrogen. Data set was collected at SOLEIL Synchrotron, Gif-sur-Yvette, France on beamline PROXIMA 2 (PX2-A) using an ADSC Q315 detector at a wavelength of 0.9801 Å. The data were processed using XDS/XSCALE (Kabsch, 2010). Initial phases were calculated by molecular replacement using the program PHASER in PHENIX (Adams et al., 2010; McCoy et al., 2007) with experimental reflection data set (Fo) and the search model (Fc) DNMT3B PWWP domain in complex with bis–tris ligand solved at 2.0 Å resolution (Wu et al., 2011) (PDB code: 3QKJ). The complex was built using the *Coot* program (Emsley et al., 2010) and refined with the program PHENIX (Adams et al., 2010). Single group occupancy of the peptide was set to 0.5 on both monomers and refined. After refinement, occupancy of 0.8 for each peptide was retained and gave good agreement with B values of surrounding residues. A summary of the data-collection and refinement statistics is presented in Table 1. Ramachandran plot for the final model shows 93.92% residues in favored regions and 6.08% in allowed regions. Figures were drawn using both Pymol (DeLano, 2002) and Discovery Studio (Visualizer, 2013).

2.4. Modeling of the DNMT3A PWWP domain in complex with histone H3K36me3

First, we validated the method by docking the H3K36me3 peptide into the DNMT3B PWWP domain (PDB code: 5CIU). For the validation procedure, the assessment of model quality was evaluated by the backbone RMSD (RMSBB) between the crystallographic structure and the peptide model. Flexpepdock retrieved the same binding mode and backbone positioning for eight models among the ten top scoring models, and the RMSBB of the best peptide model with the native peptide was 1.0 Å (<2 Å) (Fig. S2A). To model the H3K36me3 peptide ($H3_{32-38}$K36me3)-DNMT3A PWWP domain complex, we created an initial starting structure of the H3K36me3 peptide within the aromatic cage of DNMT3A PWWP domain by superposition with the DNMT3B PWWP domain–H3K36me3 structure, expecting the peptide to be in the correct binding site. We merged the estimated conformation for the H3K36me3 peptide with the DNMT3A PWWP domain receptor into an input PDB file for FlexPepDock refinement on the server. We also provided a Rosetta atom-pair constraints file to fix the trimethyl-ammonium group of Lys36 inside the aromatic cage during the simulation. The new structure was then refined in 600 independent FlexPepDock simulations. 300 simulations were performed in high-

resolution mode (full-atom mode) and, in order to increase samples of the conformational space, 300 simulations including a low-resolution pre-optimization step prior to the high-resolution refinement were performed. After a prepacking to remove internal clashes, FlexPepDock applied a Monte Carlo minimization to iteratively (ten iterative cycles) optimize the rigid body orientation, the backbone and side chain flexibility of the peptide. During this minimization, side chain flexibility of protein receptor was also considered. The total 600 models were ranked based on the Rosetta full-atom energy score (Rohl et al., 2004). Among the ten top scoring models, nine were found to have the same binding mode and backbone positioning (Fig. S2B).

2.5. Construction of the DNMT3A (PWWP-ADD-CD) structure

We selected as input structures the opened form and catalytically active structure for CD-ADD of DNMT3A with H3K4me0 epigenetic mark (Guo et al., 2014) (PDB code: 4U7T) and the DNMT3A PWWP domain (Wu et al., 2011) (PDB code: 3LLR). The docking servers Cluspro 2.0 (Comeau et al., 2004a,b; Kozakov et al., 2006, 2013) and Zdock 3.0.2 (Pierce et al., 2014) served to identify the interface interaction and the conformation of the complex. These servers perform both an automated rigid-body docking by using the fast Fourier transform correlation method and explore all the different binding modes by combination of translation and rotation of the ligand. Cluspro clusters the low-energy docked conformations based on the RMSD and classify the clusters according to their size. Zdock uses an energy-based scoring function to evaluate each pose. For the calculations, known binding sites of the DNMT3A as the DNMT3L-DNMT3A (catalytic domain) interface and the DNA binding site of the PWWP domain were not considered. After calculations, we get the same conformation for the lowest scoring function value for Zdock and Cluspro and identified a common binding interface between the ADD domain and the PWWP domain (Fig. S3). Moreover, the obtained conformation was retrieved seven times among the ten top scoring models of Zdock. The resulting models from Zdock and Cluspro were then subjected to rigid-body minimization and side-chain conformation optimization using the "docking-local-refine" on the RosettaDock server (Chaudhury et al., 2011; Lyskov and Gray, 2008; Lyskov et al., 2013). The lowest RosettaDock binding score model (ROSETTADOCK binding score was −522.4 for Zdock and −534.8 for Cluspro) shown in orange in Fig. 6A was selected for structural analysis with the nucleosome (Figs. 6C and 7).

2.6. Expression and purification of M.HhaI

LB medium containing 100 μg/ml ampicillin was inoculated with *Escherichia coli* K-12 strain ER1727 containing the pUHE25H-haI plasmid. When the OD600 nm reached 0.6, expression of M.*Hha*I was induced with 1 mM IPTG. After 3 h of induction at 37 °C, the cells were collected by centrifugation at 4000 rpm for 30 min at 4 °C. The cell paste was resuspended in ice-cold lysis buffer (10 mM Hepes (pH 7.0), 5 mM EDTA, 10% (v/v) glycerol, 0.1% (v/v) β-mercaptoethanol). Cells were lysed on ice by sonication using a cell disrupter. After centrifugation at 4000 rpm for 1 h at 4 °C, the supernatant was discarded and the pellet was resuspended in a high-salt buffer (10 mM Tris, 5 mM EDTA (pH 7.5), 10% (v/v) glycerol, 0.1% (v/v) β-mercaptoethanol, 500 mM NaCl) to re-solubilize the protein. The high-salt suspension was centrifuged at 14000 rpm for 1 h at 4 °C, the supernatant was kept and the pellet undergoing the same processing with high-salt buffer. Precipitation of nucleic acids from the supernatant was performed by progressive addition of a half volume of protamine sulfate solution (10 mg/ml, solubilized in high-salt buffer). After incubation at room temperature for 5 min, the solution was centrifuged at 14000 rpm for 40 min at 4 °C and the supernatant dialyzed against 50 mM Tris (pH 6.7), 5 mM EDTA, 10% (v/v) glycerol, 0.1% (v/v) β-mercaptoethanol, 100 mM NaCl. All the following steps of purification were performed at 4 °C. The sample was filtered through a 0.2 μm filter (Whatman® FP 30/0.2, GE Healthcare) and loaded onto a cation exchange column (1 ml, HiTrap™ SP FF, GE Healthcare). The charged column was washed with the equilibration buffer (50 mM Tris (pH 6.7), 5 mM EDTA, 10% (v/v) glycerol, 0.1% (v/v) β-mercaptoethanol, 100 mM NaCl) until a stable baseline was attained and the elution of the M.*Hha*I was realized with a linear gradient from 100 mM to 500 mM NaCl over a total volume of 50 ml. After pooling the appropriate fractions, the solution was concentrated in a stirred cell using a 10 kDa MWCO membrane (Vivaspin 15R; Sartorius Stedim Biotech GmbH, Göttingen, Germany) until ~5 ml and dialyzed against 50 mM Tris (pH 7.4), 5 mM EDTA, 10% (v/v) glycerol, 0.1% (v/v) β-mercaptoethanol, 100 mM NaCl. The sample was filtered through a 0.2 μm filter (Whatman® FP 30/0.2, GE Healthcare) and was injected into an AKTA FPLC system (GE-Amersham Biosciences) with an anion exchange column (1 ml, Mono-Q HR 5/5, GE Healthcare) equilibrated with 50 mM Tris (pH 7.4), 5 mM EDTA, 10% (v/v) glycerol, 0.1% (v/v) β-mercaptoethanol, 100 mM NaCl. The charged column was washed with the equilibration buffer until a stable baseline was attained. The collected flow-through fractions containing the purified M.*Hha*I were pooled and concentrated up to 10 mg/mL concentration prior to crystallization. All the different stages of expression and purification were followed by SDS–PAGE.

2.7. Crystallization and structure determination of M.HhaI in complex with SAH and a short DNA duplex

The crystallization conditions were based and adapted from previously published data (O'Gara et al., 1996, 1998). The oligonucleotides used to form the 12 bp duplex were purchased from Eurogentec (Seraing, Belgium).

5′-T-C-C-A-T-G-C-G-C-T-G-A-C--3′
3′--G-G-T-A-C-G-X-G-A-C-T-G-T-5′

Table 1
Data collection and refinement statistics.

	DNMT3B PWWP–H3K36me3	M.*Hha*I-DNA-SAH
Data collection		
Space group	P 32 2 1 (No. 154)	H 3 2 (No. 155)
Cell dimensions		
a, *b*, *c* (Å)	73.44, 73.44, 158.20	95.33, 95.33, 314.69
α, *β*, *γ* (°)	90, 90, 120	90, 90, 120
Resolution (Å)	50.0–2.24 (2.32–2.24)[a]	47.7–1.59 (1.65–1.59)[a]
R_{sym} or R_{merge}	6.2 (63.3)	4.5 (69.4)
$I/\sigma I$	18.50 (3.19)	24.96 (1.71)
Completeness (%)	99.74 (99.96)	98.50 (90.27)
Redundancy	6.9 (7.0)	6.6 (3.2)
Refinement		
Resolution (Å)	36.72–2.24	36.55–1.59
No. reflections	24436	72699
R_{work}/R_{free}	22.4/25.9	16.5/18.8
No. atoms		
Protein	1997	2606
Peptide/DNA	100 (peptide)	484 (DNA)
Ligand/ion	6 (citrate)	26 (SAH); 40 (sulfate)
Water	163	634
B-factors		
Protein	45.90	16.81
Peptide/DNA	52.60 (peptide)	23.20 (DNA)
Ligand/ion	55.14 (citrate)	27.74 (SO_4)
Water	49.30	30.70
R.m.s. deviations		
Bond lengths (Å)	0.009	0.006
Bond angles (°)	1.20	1.16

[a] Values in parentheses are for highest-resolution shell.

These two synthetic complementary strands of oligonucleotides were designed with a single 5′ thymidine (T)-overhangs at both ends and an abasic residue was incorporated with a dSpacer (X) at the target base pair (GC → GX). The two strands were hybridized in TE buffer at 90 °C for 5 min, followed by room temperature incubation for 1 h. Crystallization trials were performed using the hanging-drop vapor-diffusion method at room temperature with purified M.*Hha*I concentrated at 10 mg/ml, SAH (dissolved in water) and DNA with following molar ratio: 1:3:1.3. Drops consisted of 2 µl of M.*Hha*I-SAH-DNA mixture plus 2 µl of reservoir solution equilibrated against a reservoir volume of 700 µl. Crystals of complex between M.*Hha*I, SAH and DNA were grown against a reservoir consisting of 50 mM Citrate pH 5.6 and 1.8 M ammonium sulfate. A single-crystal (crystal size: 0.3 mm × 0.2 mm) was cryoprotected in 20% glycerol and flash-cooled in liquid nitrogen. Data set was collected at SOLEIL Synchrotron, Gif sur Yvette, France on beamline PROXIMA 2 (PX2-A) using an ADSC Q315r detector at a wavelength of 0.9801 Å. The data were processed using XDS/XSCALE (Kabsch, 2010). A summary of the data-collection and refinement statistics is presented in Table 1. Initial phases were calculated by molecular replacement using the program PHASER in PHENIX (Adams et al., 2010; McCoy et al., 2007) with experimental reflection data set (Fo) and the search model (Fc) solved at 2.4 Å resolution (O'Gara et al., 1998) (PDB code: 9MHT). The complex was built using the *Coot* program (Emsley et al., 2010) and refined with the program PHENIX (Adams et al., 2010). Ramachandran plot for the final model shows 97.85% residues in favored regions, 1.85% in allowed regions and 0.31% in disallowed regions. Figures were drawn using both Pymol (DeLano, 2002) and Discovery Studio (Visualizer, 2013).

3. Results and discussion

3.1. Structure of the DNMT3B PWWP domain in complex with histone H3K36me3

We determined the crystal structure of the DNMT3B PWWP domain in complex with the epigenetic mark H3K36me3 ($H3_{32-38}$K36me3) to gain structural insight into this recognition (Fig. 1A–D). Compared to the open form (Qiu et al., 2002), the loop between the β1 and β2 strand undergoes an induced fit that closes the aromatic cage in order to enhance recognition of this epigenetic mark. The histone peptide occupies a surface groove formed by the β1 strand, the loop between β1 and β2, and the β4 strand (Fig. 1A and B). The trimethylated side chain of Lys36 is well resolved in the density map and is inserted into the aromatic cage formed by the three aromatic residues Phe236, Trp239 (β2 strand) and Trp263 (β3 strand) (Fig. 1C). The trimethyl-ammonium group

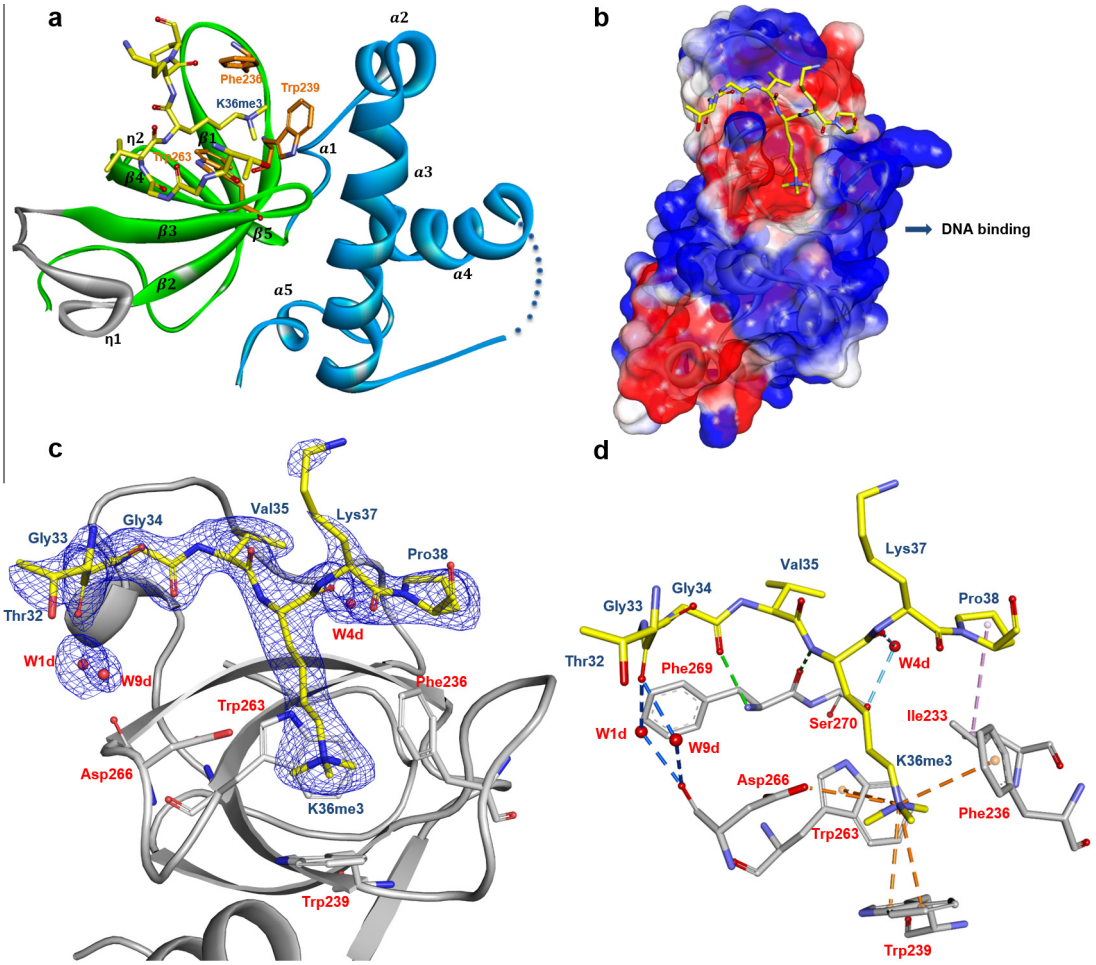

Fig. 1. Structure of the DNMT3B PWWP domain in complex with H3K36me3 peptide. (A) Solid Ribbon representation of the crystal structure of the complex between the DNMT3B PWWP domain and the H3K36me3 peptide. (B) DNMT3B PWWP domain surface colored by electrostatic potential of the residues (red: negatively charged area and blue: positively charged area) with the peptide represented in stick. (C) 2Fo − Fc electron density map contoured at 1 σ for the H3K36me3 peptide (blue mesh) with the stick representation of the residues of the aromatic cage: Phe236, Trp263, Asp266 and Trp239. (D) Selected intermolecular interactions of the peptide with the DNMT3B PWWP domain (Table S2).

forms interaction with this aromatic cage by van der Waals and π-cation interactions with aromatic side chains of Phe236, Trp239 and Trp263, and electrostatic–cation interaction with the carboxylate group of Asp266 (Fig. 1D). These interactions are common to different chromatin-binding proteins in complex with the H3K36me3 peptide (Qin and Min, 2014). Another important common feature of "Royal" superfamily domains for binding preferentially H3K36me3 involves hydrophobic contacts surrounding the trimethylated lysine 36. In our complex, Pro38 is involved in a hydrophobic contact with the side chain of Ile233 as observed for the PHF1 Tudor domain (Cai et al., 2013; Musselman et al., 2012) and Val35 for the PWWP domain of BRPF1 (Vezzoli et al., 2010; Wu et al., 2011). For complex between epigenetic marks H3K4me3, H3K9me3, and H3K27me3, and other chromatin-binding proteins, we observe polar contacts in this surface area (Cheng et al., 2013; Pena et al., 2006; Sanulli et al., 2015). Besides that, an additional key interaction involved the CO main chain group of trimethylated Lys36 with a conserved water molecule stabilized by O atom of Ser270. The loss of this last strong hydrogen bond (D_{A-B} = 2.41 Å in chain C and 2.59 Å in chain D) in the ICF syndrome (S270P) explains the lack of methyltransferase recognition and activity on nucleosomal substrates (Baubec et al., 2015; Dhayalan et al., 2010; Shirohzu et al., 2002). The rest of the modified peptide is involved in intermolecular hydrogen-bonds with backbone amino and carbonyl groups: CO of Thr32 and conserved waters stabilized by CO of Asp266, CO of Gly34 and NH of Phe269, and NH of trimethylated Lys36 and CO of Phe269. The previously discussed intermolecular interactions between DNMT3B PWWP domain and H3K36me3 peptide are presented in Table S2. Some of these observed interactions can be related to previous mutational studies. Mutations among conserved residues of DNMT3A and DNMT3B PWWP domains were performed in the β-barrel (for DNMT3B: $W_{239}P_{240}$-ST; D_{266}-A) and led to the loss of ability of both DNMT3A and DNMT3B to bind to H3K36me3 modified nucleosomes (Baubec et al., 2015; Chen et al., 2004; Dhayalan et al., 2010; Ge et al., 2004). As observed for other H3K36me3 binding domains, van der Waals and electrostatic interactions contribute to a specific association of DNMT3B PWWP domain with H3K36me3 peptide.

One additional observation is that the trimethylated lysine can be engaged differently as observed for the PHF1 and PHF19 Tudor domains (Fig. 2A). We observed a flip of 180 °C as these domains lack the α-helix bundle and interact with another surface of the nucleosome-core. H3K36me3 peptide adopts the same binding mode with both DNMT3B PWWP and BRPF1 (Fig. 2B) (Vezzoli et al., 2010). However, the β-β-α insertion motif of the BRPF1 interacts with peptide residues preceding the trimethylated Lys36 compared to the DNMT3B PWWP exhibiting a short α-helix motif (η1 insertion motif).

Experiments to confirm binding of the H3K36me3 peptide to the DNMT3B PWWP domain *in vitro* were attempted, so far without success. In particular, due to unanticipated technical problems inherent to the system under study (precipitation and non-specific binding), we were not able to confirm the interaction using different biophysical approaches (ITC, DSF, and Bio-layer interferometry).

3.2. Modeling of the DNMT3A PWWP domain in complex with histone H3K36me3

As we were not able to crystallize the complex between the DNMT3A PWWP domain and H3K36me3 peptide, we performed a protein-peptide docking based on our structural study. DNMT3A and DNMT3B PWWP domains are small domains of ~150 residues

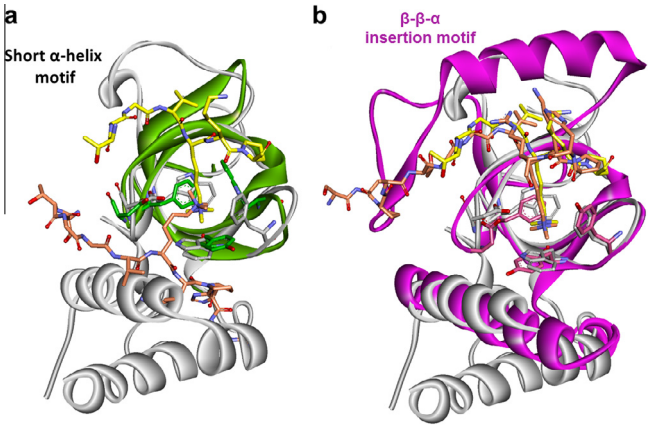

Fig. 2. Different orientations and interactions of H3K36me3 peptide observed among the "Royal" superfamily. (A) Superposition of the DNMT3B PWWP domain (gray) – H3K36me3 (yellow) complex with the PHF1 (green) – H3K36me3 (orange) complex (Musselman et al., 2012) (PDB code: 4HCZ). The absence of a helix bundle for the PHF1 allows the H3K36me3 to be engaged differently than with DNMT3B PWWP. (B) Superposition of the DNMT3B PWWP domain (gray) – H3K36me3 (yellow) complex with the BRPF1 (purple) – H3K36me3 (orange) complex (Vezzoli et al., 2010) (PDB code: 2X4W).

very similar in structural organization and amino acid sequence (Fig. S1). They share 53% of identity in sequence and 67% of similarity using blastp. Based on our crystal structure and the high structural similarity between the two domains, we decided to dock this short peptide ($H3_{32-38}$K36me3) inside the DNMT3A PWWP domain (Wu et al., 2011) (PDB code: 3LLR). We used for this the Rosetta FlexPepDock protocol introduced for the refinement of coarse starting structure of peptide-protein complex into high-resolution models (Raveh et al., 2010) (see Section 2). The proposed model provides a binding mode of the H3K36me3 peptide similar to the one observed in the crystallographic structure with the DNMT3B PWWP domain (Fig. 3A and B). Residues interacting with the H3K36me3 peptide are conserved among DNMT3A and DNMT3B PWWP domains (Fig. 4 and Table S1, S3). The trimethyl-ammonium group is stabilized into the cage formed by Phe303, Trp306, Trp330, and Asp333 (Phe236, Trp239, Trp263 and Asp266 for DNMT3B). The backbone of the peptide is involved in intermolecular hydrogen-bonding interactions with Phe336 ($Phe269_{3B}$) and Ser337 ($Ser270_{3B}$). A hydrophobic contact is observed between Pro38 and the isopropyl-group of Leu300 ($Ile233_{3B}$).

3.3. DNMT3A/3B PWWP domains in complex with the histone H3K36me3 nucleosome core particle

In addition to their histone tail binding abilities, PWWP domains of DNMT3A and DNMT3B interact non-specifically with DNA through a basic surface adjacent to the histone binding site (Fig. 1B) (Purdy et al., 2010; Qiu et al., 2002). Individually, different PWWP domains bind free DNA in the μM range and H3K36me3 peptide in the μM–mM range which is quite low (Lukasik et al., 2006; Qiu et al., 2002; Yu et al., 2012). However, the cooperative binding of these partners can increase the binding affinity up to 10^4-fold and enhance specificity of PWWP domains for H3K36me3-NCP (nucleosome core particle) (Musselman et al., 2012; Wang et al., 2014). This seems to be a general property of PWWP modules and involves conserved interactions at the same

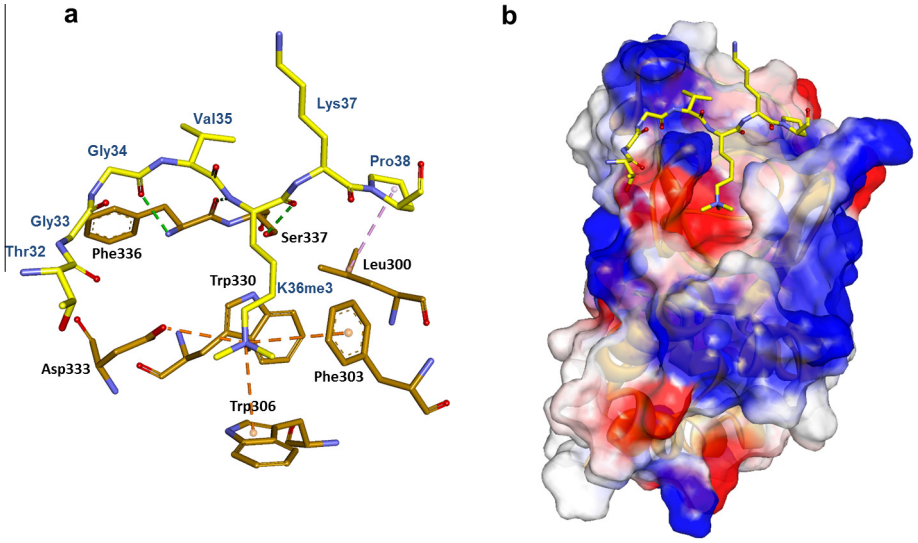

Fig. 3. Molecular modeling of interactions between the DNMT3A PWWP domain and H3K36me3 peptide ($H3_{32-38}$K36me3). (A) Intermolecular interactions of the top scoring model of the DNMT3A PWWP domain in complex with the H3K36me3 peptide. (B) DNMT3A PWWP domain surface colored by electrostatic potential of the residues (red: negatively charged area and blue: positively charged area) with the peptide represented in stick.

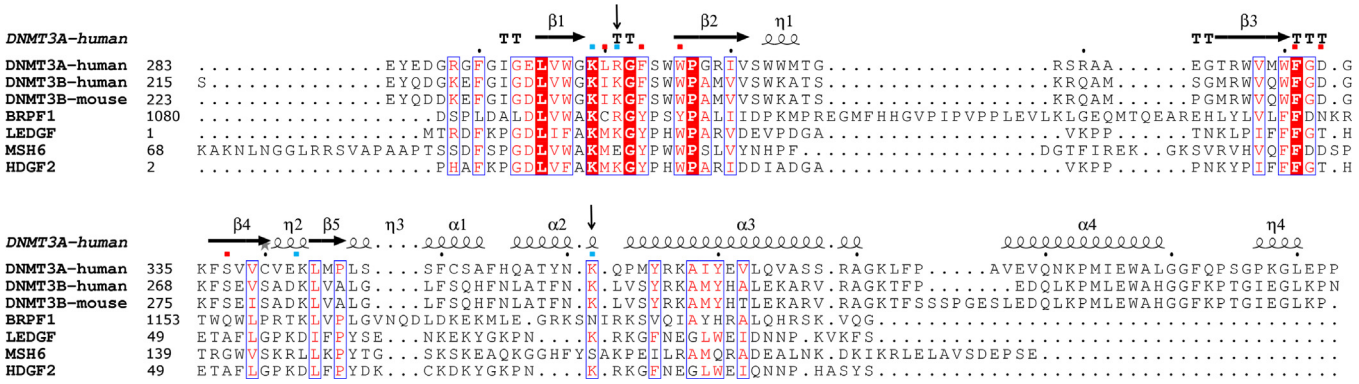

Fig. 4. Structure-based sequence alignment of homologous PWWP domains reveals conserved residues involved in nucleosomal DNA binding. This multiple structure and sequence alignments of homologous PWWP domains was performed with PROMALS3D (Pei et al., 2008) and formatted with Espript (Gouet et al., 1999). In the family of PWWP domains, a hydrophobic interface interacts with the histone tail and an adjacent more basic surface interacts with the negative phosphate backbone of the DNA in order to increase selectivity and affinity with the nucleosome. Residues involved in the H3K36me3 peptide binding: ■ DNA-binding residues: ■. For LEDGF protein, residues Lys16 (↓) and Lys73 (↓) have a large contribution to the intermolecular energy (van Nuland et al., 2013); they correspond to Arg301 and Lys361 in DNMT3A, and Lys234 and Lys294 in DNMT3B.

interface (Eidahl et al., 2013; Qin and Min, 2014; van Nuland et al., 2013). Indeed, multiple structure and sequence alignments of homologous PWWP domains (Fig. 4) show that these domains present the same aromatic cage to bind the histone tail and the basic surface have conserved residues to interact with the nucleosomal DNA. The question remains how the DNMT3A and DNMT3B PWWP domains bind the H3K36me3 epigenetic mark which is near the nucleosome core. To investigate this mechanism of recognition, we superimposed the DNMT3A and DNMT3B PWWP domains with the previously studied LEDGF PWWP domain, sharing high structural similarity, bound to H3K36me3-NCP (Eidahl et al., 2013). We observe that the basic surface is positioned on the DNA wrapped around histone core and is likely to interact with the phosphate backbone of DNA through the same conserved residues as those of LEDGF PWWP domain (Figs. 4 and 5A, B). DNMT3A and DNMT3B PWWP domains are situated between two DNA duplexes from where the epigenetic mark H3K36me3 emerges to interact with them (Figs. 5A and 6C). The selectivity of DNMT3A and DNMT3B for the H3K36me3-NCP is then mediated by interactions of the PWWP domain through the aromatic cage for H3K36me3 recognition and the basic surface which could bind two DNA duplexes of the nucleosome. Experiments showing this direct interaction between the PWWP domain and H3K36me3 nucleosomes were performed in previous studies for both DNMT3A and DNMT3B by GST pull-down assays (Baubec et al., 2015; Dhayalan et al., 2010). These interactions are disrupted by different point mutations (as mentioned earlier in the description of the complex between the DNMT3B PWWP domain and histone H3K36me3).

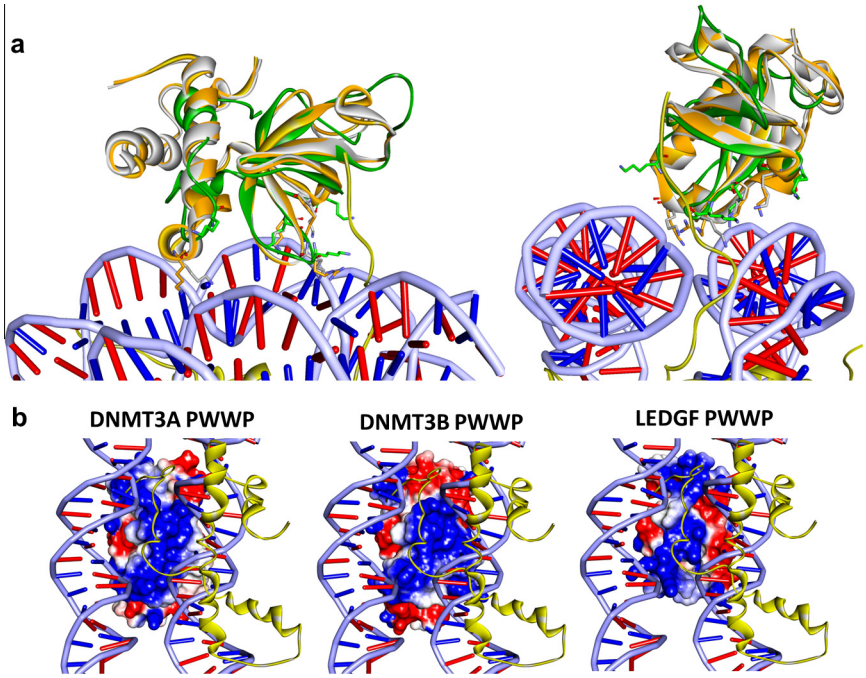

Fig. 5. PWWP domains in complex with H3K36me3-NCP. (A) HADDOCK model of the complex between the LEDGF PWWP (green) domain and H3K36me3-NCP (Eidahl et al., 2013) superimposed with DNMT3A (orange) (root-mean-square deviation of 1.9 Å) and DNMT3B (gray) (root-mean-square deviation of 1.9 Å) PWWP domains. (B) PWWP domains surface colored by electrostatic potential of the residues (red: negatively charged area and blue: positively charged area). The positively charged surface is positioned between two DNA strands and contains the following residues for the DNMT3A PWWP domain: Lys299, Arg301, Lys343 and Lys 361. These residues correspond to Lys232, Lys 234, Lys276, and Lys 294 in the DNMT3B PWWP domain. They are situated on β1, on the loop between β1 and β2, between β4 and β5 (n2), on the loop between α2 and α3, and on helix α3.

3.4. DNMT3A (PWWP-ADD-CD) in a nucleosomal context

DNA methylation functions in specific cellular and genomic contexts. For example, DNA methylation is dependant of a number of post-translational modifications on histones to recruit and activate the *de novo* DNA methylation complex. DNMT3L, which lacks the PWWP domain and the catalytic domain, recognizes specifically the unmethylated histone H3 tail (H3K4me0) through the ADD domain (Table S1) (Eustermann et al., 2011; Ooi et al., 2007; Otani et al., 2009). This protein is expressed especially in oocytes to stabilize the conformation of the active loop and stimulates the activity of DNMT3A and DNMT3B for the establishment of the genomic imprinting (Jia et al., 2007; Smallwood et al., 2011; Suetake et al., 2004). The epigenetic mark H3K4me0 is also bound by DNMT3A *via* the same ADD domain to allosterically activates the DNMT3A (Guo et al., 2014; Li et al., 2011; Otani et al., 2009). Without this recognition, the ADD domain interacts specifically with the catalytic domain and prevents interaction with the DNA substrate inhibiting the activity. Binding of this epigenetic mark H3K4me0 to the ADD domain disrupts this autoinhibitory structure and an important conformational change permits to the catalytic domain to recognize the DNA. The ADD domain of DNMT3B binds the H3K4me0 (1–19) with the same affinity as DNMT3A suggesting a similar mechanism (Table S1) (Zhang et al., 2010). H3K36me3 epigenetic mark, as explained before, is read by both DNMT3A and DNMT3B PWWP domains to recruit these enzymes on the nucleosomal DNA in the pericentromeric heterochromatin regions and gene bodies (Bachman et al., 2001; Baubec et al., 2015; Chen et al., 2004; Dhayalan et al., 2010). The combination of histone modifications permissive unmethylated H3K4me0 and trimethylated H3K36me3 seems therefore necessary for the DNMT3A and DNMT3B to access and methylate DNA (Stewart et al., 2015; Tomizawa et al., 2012).

Different structural and biophysical data are available for separated domains of DNMT3s with individual parts of nucleosome. However, no structural information is available for the recognition of DNMT3s with the integral nucleosome. In this regard, we propose here a model of the DNMT3A (PWWP-ADD-CD) in a nucleosomal context. The crystal structure of the active complex of DNMT3A (ADD-CD)-DNMT3L-H3K4me0 peptide has recently been solved without the PWWP domain (Guo et al., 2014) (PDB code: 4U7T). Based on this structure, we predict the DNMT3A (PWWP-ADD-CD) structure including the missing PWWP domain by using three protein–protein docking servers emerging from the Critical Assessment of PRedicted Interactions (CAPRI) experiment (Janin et al., 2003) (see Section 2). This reconstructed DNMT3A structure (Fig. 6A) shows that the β-strands, β2 and β3, and the insertion motif η1 (α-helix motif) of the PWWP domain are positioned at the interface with the ADD domain (Figs. 6A and S3). The PWWP and the ADD domains are associated by strong electrostatic interactions and diverse hydrogen bonds. In addition, due to the high structural and sequence similarity of the catalytic domain of DNMT3A and the M.*Hha*I, a bacterial DNMT we solved in complex with a short DNA duplex, we superimposed these domains (root-mean-square deviation of 1.2 Å). The short oligonucleotide is in the continuity of the nucleosomal DNA and can be connected easily to the border of the nucleosome to form a contiguous DNA segment (Fig. 6B and C). This approach yields a complete structural model of the H3K36me3-nucleosome-DNMT3A complex which can explain DNMT3A recruitment to a genomic site (Fig. 6C). The PWWP domain specifically binds the H3K36me3-NCP followed by an activation of the catalytic domain through the binding of the H3K4me0 with the ADD domain to methylate the nearby cytosine (Fig. 8A). We also introduced in this study, the DNMT3L from the active complex of DNMT3A (ADD-CD)-DNMT3L-H3K4me0 peptide (Guo et al., 2014) and the dimerization of DNMT3A

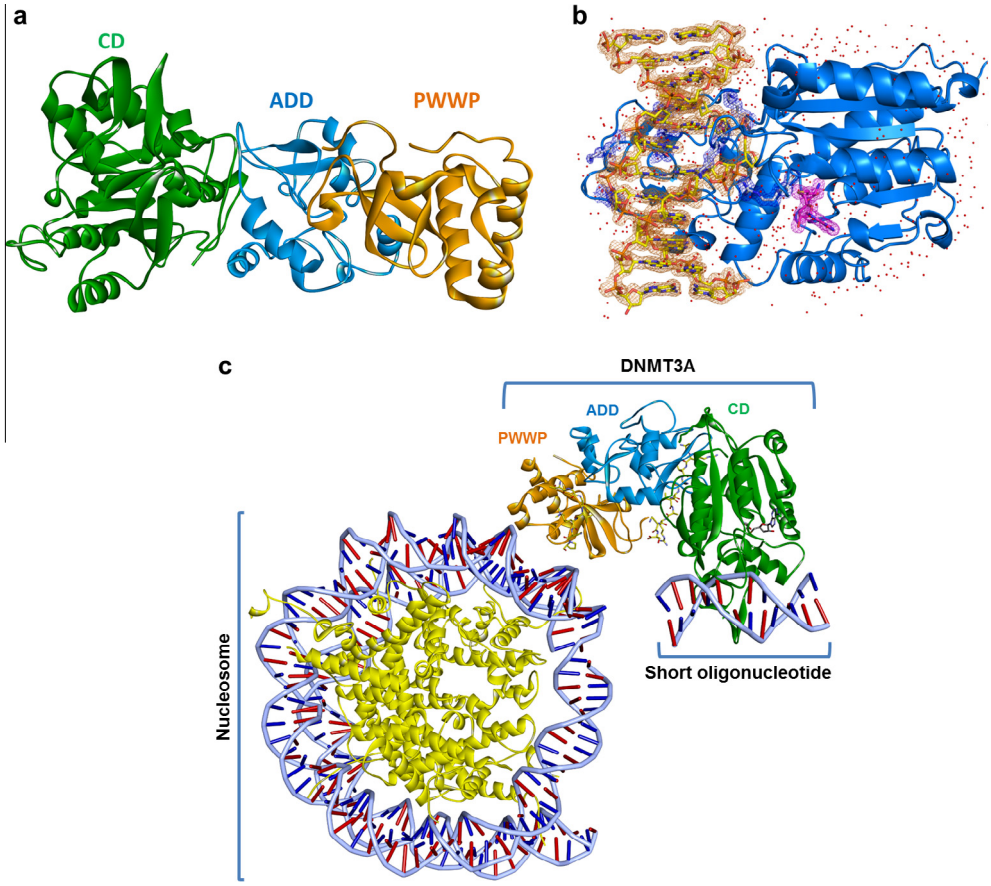

Fig. 6. (A) Cartoon view of the complete reconstructed DNMT3A (PWWP-ADD-CD structure). (B) The bacterial M.*Hha*I (a DNA (cytosine-5) methyltransferase from *Haemophilus haemolyticus*) in complex with DNA and SAH (*S*-adenosyl-L-homocysteine). 2Fo − Fc electron density map contoured at 1.0 σ for short oligonucleotide (orange mesh), SAH (pink mesh) and selected residues interacting with DNA (blue mesh). (C) Structural model of the DNMT3A in complex with a nucleosome.

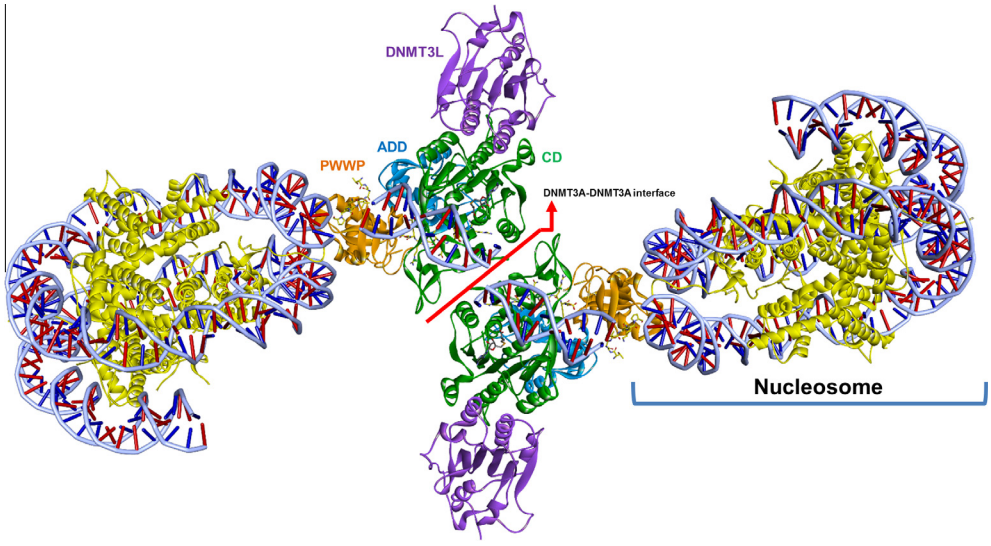

Fig. 7. Structural model of DNMT3A (PWWP-ADD-CD)-DNMT3L heterotetramer in complex with H3K36me3-modified dinucleosome. Two active sites of distinct DNMT3A are located on the linker DNA region between and at the border of the nucleosomes. The length of the linker DNA is ~30-bp and corresponds to the average length in the pericentromeric heterochromatin regions (Chodavarapu et al., 2010). The center-to-center distance between the core particles is 20 nm.

through the catalytic domain (Chodavarapu et al., 2010; Jia et al., 2007; Yazdi et al., 2015). This gives a heterotetramer (DNMT3L-DNMT3A-DNMT3A-DNMT3L) in complex with a dinucleosome exhibiting a bent linker DNA (Figs. 7 and 8B). The DNMT3A-DNMT3A dimer is responsible for methylation of DNA in a periodic pattern of ~10-bp between two CpG sites corresponding to a

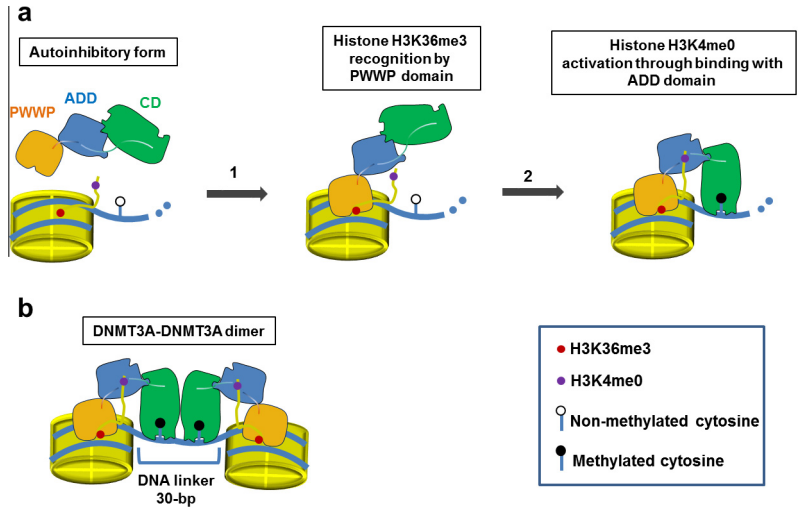

Fig. 8. Proposed mechanism for recognition of nucleosome and DNA methylation by DNMT3A (for clarity, the DNMT3L was removed). (A) Sequential recognition mechanism of DNMT3A for nucleosome and methylation activation with H3 tail: **1**-Nucleosome recognition by PWWP domain through interaction with methylated histone H3 tail (H3K36me3). **2**-Histone H3 tail activation (H3K4me0) through binding with ADD domain. ADD domain of DNMT3A interacts with the catalytic domain and inhibits its activity by preventing it to interact with the DNA. (B) Periodicity in *de novo* methylation patterns of ~10-bp between two CpG sites.

helical turn as previously demonstrated (Jia et al., 2007; Lister et al., 2009; Smallwood et al., 2011). Moreover, this model highlights, as described in the literature, that *de novo* DNA methylation occurs on the linker DNA region between nucleosomes and at the border of the nucleosome (Baubec et al., 2015; Felle et al., 2011a; Morselli et al., 2015; Takeshima et al., 2008). Indeed, DNMT3A and DNMT3B prefer to target the nucleosome-bound DNA through a synergetic effect to methylate the linker DNA region between two nucleosomes which is more accessible (Felle et al., 2011b; Morselli et al., 2015; Zhang et al., 2010). Furthermore, the nucleosomes are more regularly connected in the pericentromeric heterochromatin regions with a linker DNA of ~30-bp which corresponds to the linker DNA size of our model (Chodavarapu et al., 2010). The four ADD domains present in the heterotetrameric complex (DNMT3L-DNMT3A-DNMT3A-DNMT3L) could read the H3K4me0 state for the DNA methylation establishment. The question may be asked whether DNMT3L could interact with the second copy of H3 tail of the histone octamer from adjacent nucleosomes. This could lead to a synergistic effect of binding proteins on nucleosome. This model needs further testing to answer these questions, nevertheless taken together our observations emphasize the functional importance of the combination of histone tail modification status for the formation of a stable *de novo* methylation complex to methylate DNA in a periodic pattern.

It is still difficult to understand the difference of methylation activity between DNMT3A and DNMT3B. Indeed, as for DNMT3A, DNMT3B methylates the linker DNA with a periodicity of 10-bp and a periodicity corresponding to the nucleosome repeat length of ~180 bp (nucleosomal DNA (147-bp) + linker DNA (~30-bp)) (Baubec et al., 2015; Cokus et al., 2008; Morselli et al., 2015). DNMT3B binding to nucleosome is also guided by the presence of two epigenetic marks, the H3K4me0 and the H3K36me3 peptides (Baubec et al., 2015; Morselli et al., 2015). DNMT3A and DNMT3B act in a same nucleosome context in terms of organization of chromatin but methylate different regions of the genome. DNMT3B is however more enriched within gene bodies (Baubec et al., 2015; Jin et al., 2012). DNMT3A and DNMT3B are also distributed differently across the genome and this depends on the cellular context (Jin et al., 2012). The mechanism of DNMT3s targeting to specific site of methylation throughout the genome is actually more complex and involves a series of events including the timing of expression (Lees-Murdock et al., 2005). The differences of recruitment of DNMT3s across the genome may be influenced by other specific histone tail modifications. Their distinct biological functions can be explained by the N-terminal sequence which shares very low sequence identity (10.9% using the EMBOSS needle program) between the two DNMT3s. In our model we do not present this N-terminal part of DNMT3A (amino acids 1–281) as no structural information is available for neither DNMT3A nor DNMT3B. This N-terminal part of both DNMT3A/3B seems important for anchoring to nucleosomes (Baubec et al., 2015; Jeong et al., 2009). Indeed, based on multiple sequence alignments including this sequence, we found relate domains involved in DNA binding including the PDS5 homolog B protein (uniprotkb: Q6TRW4). This protein is a regulator which stabilizes cohesion complex association with chromatin suggesting that the N-terminal part of the DNMT3A could interact with DNA through its 21 arginines and 22 lysines (Fig. S1). Furthermore, this N-terminal part of both DNMT3A/3B could be important to mediate protein–protein interactions to discriminate the recruitment of these two enzymes on the nucleosomal-DNA (Geiman et al., 2004; Li et al., 2006; Rigbolt et al., 2011; Sarraf and Stancheva, 2004; Velasco et al., 2010).

4. Conclusion

Here, we report for the first time the crystal structure of the *de novo* DNA methyltransferase DNMT3B PWWP domain in complex with the epigenetic mark $H3_{32-38}K36me3$. This structure emphasizes that a conserved water molecule mediates strong hydrogen bonding with the epigenetic mark explaining the loss of affinity in the ICF syndrome with the mutated DNMT3B PWWP domain (S270P). Based on our model, the DNMT3A PWWP domain involves nearly the same residues to interact with the H3K36me3 epigenetic mark which has the same conformation as our crystal structure. In a nucleosomal context, DNTM3A and DNMT3B PWWP domains appear to engage the same binding surface and residues to interact with the H3K36me3 nucleosome core particle. The complete structure of the DNMT3A-DNMT3L heterotetramer in complex with a dinucleosome provides structural information about DNA methylation on the linker DNA region of nucleosomes and the *de novo* DNA methylation patterns in agreement with the findings of the literature.

Data deposition

Coordinates and diffraction data for the structures of DNMT3B PWWP–H3K36me3 and M.*Hha*I-DNA-SAH have been deposited in the Protein Data Bank with accession codes 5CIU and 5CIY, respectively.

Acknowledgements

We thank Professor Mamuka Kvaratskhelia for kindly providing us the structural model of the LEDGF PWWP domain in complex with H3K36me3 – nucleosome core particle. This work was supported by the Belgian National Fund for Scientific Research (Grant F.R.S. – FNRS – Télévie 7.4.532.15.F.). Authors also thank A. Matagne, G. Feller and J. Vandenameele (ULg; Belgium) for advice and comments during preliminary biophysical assays.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jsb.2016.03.013.

References

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221.

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

Ballaré, C., Lange, M., Lapinaite, A., Martin, G.M., Morey, L., Pascual, G., Liefke, R., Simon, B., Shi, Y., Gozani, O., 2012. Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. Nat. Struct. Mol. Biol. 19, 1257–1265.

Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A., Schubeler, D., 2015. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature 520, 243–247.

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

Cai, L., Rothbart, S.B., Lu, R., Xu, B., Chen, W.-Y., Tripathy, A., Rockowitz, S., Zheng, D., Patel, D.J., Allis, C.D., 2013. An H3K36 methylation-engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. Mol. Cell 49, 571–582.

Chaudhury, S., Berrondo, M., Weitzner, B.D., Muthu, P., Bergman, H., Gray, J.J., 2011. Benchmarking and analysis of protein docking performance in Rosetta v3. 2. PLoS One 6, e22477.

Chen, T., Tsujimoto, N., Li, E., 2004. The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin. Mol. Cell. Biol. 24, 9048–9058.

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. J. Biol. Chem. 288, 1329–1339.

Chodavarapu, R.K., Feng, S., Bernatavichute, Y.V., Chen, P.-Y., Stroud, H., Yu, Y., Hetzel, J.A., Kuo, F., Kim, J., Cokus, S.J., 2010. Relationship between nucleosome positioning and DNA methylation. Nature 466, 388–392.

Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M., Jacobsen, S.E., 2008. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. Nature 452, 215–219.

Comeau, S.R., Gatchell, D.W., Vajda, S., Camacho, C.J., 2004a. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. Bioinformatics 20, 45–50.

Comeau, S.R., Gatchell, D.W., Vajda, S., Camacho, C.J., 2004b. ClusPro: a fully automated algorithm for protein–protein docking. Nucleic Acids Res. 32, W96–W99.

DeLano, W.L., 2002. The PyMOL molecular graphics system.

Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R.Z., Ragozin, S., Jeltsch, A., 2010. The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. J. Biol. Chem. 285, 26114–26120.

Eidahl, J.O., Crowe, B.L., North, J.A., McKee, C.J., Shkriabai, N., Feng, L., Plumb, M., Graham, R.L., Gorelick, R.J., Hess, S., 2013. Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. Nucleic Acids Res. 41, 3924–3936.

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501.

Eustermann, S., Yang, J.-C., Law, M.J., Amos, R., Chapman, L.M., Jelinska, C., Garrick, D., Clynes, D., Gibbons, R.J., Rhodes, D., 2011. Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. Nat. Struct. Mol. Biol. 18, 777–782.

Felle, M., Hoffmeister, H., Rothammer, J., Fuchs, A., Exler, J.H., Langst, G., 2011a. Nucleosomes protect DNA from DNA methylation in vivo and in vitro. Nucleic Acids Res. 39, 6956–6969.

Felle, M., Hoffmeister, H., Rothammer, J., Fuchs, A., Exler, J.H., Längst, G., 2011b. Nucleosomes protect DNA from DNA methylation in vivo and in vitro. Nucleic Acids Res. 39, 6956–6969.

Ge, Y.-Z., Pu, M.-T., Gowher, H., Wu, H.-P., Ding, J.-P., Jeltsch, A., Xu, G.-L., 2004. Chromatin targeting of de novo DNA methyltransferases by the PWWP domain. J. Biol. Chem. 279, 25447–25454.

Geiman, T.M., Sankpal, U.T., Robertson, A.K., Chen, Y., Mazumdar, M., Heale, J.T., Schmiesing, J.A., Kim, W., Yokomori, K., Zhao, Y., 2004. Isolation and characterization of a novel DNA methyltransferase complex linking DNMT3B with components of the mitotic chromosome condensation machinery. Nucleic Acids Res. 32, 2716–2729.

Gouet, P., Courcelle, E., Stuart, D.I., 1999. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 15, 305–308.

Guo, X., Wang, L., Li, J., Ding, Z., Xiao, J., Yin, X., He, S., Shi, P., Dong, L., Li, G., 2014. Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. Nature 517, 640–644.

Hahn, M.A., Wu, X., Li, A.X., Hahn, T., Pfeifer, G.P., 2011. Relationship between gene body DNA methylation and intragenic H3K9me3 and H3K36me3 chromatin marks. PLoS One 6 e18844.

Janin, J., Henrick, K., Moult, J., Eyck, L.T., Sternberg, M.J., Vajda, S., Vakser, I., Wodak, S.J., 2003. CAPRI: a critical assessment of predicted interactions. Proteins Struct. Funct. Bioinf. 52, 2–9.

Jeong, S., Liang, G., Sharma, S., Lin, J.C., Choi, S.H., Han, H., Yoo, C.B., Egger, G., Yang, A.S., Jones, P.A., 2009. Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. Mol. Cell. Biol. 29, 5366–5376.

Jia, D., Jurkowska, R.Z., 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, B., Ernst, J., Tiedemann, R.L., Xu, H., Sureshchandra, S., Kellis, M., Dalton, S., Liu, C., Choi, J.-H., Robertson, K.D., 2012. Linking DNA methyltransferases to epigenetic marks and nucleosome structure genome-wide in human tumor cells. Cell Rep. 2, 1411–1424.

Jones, P.A., 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. 13, 484–492.

Kabsch, W., 2010. Xds. Acta Crystallogr. D Biol. Crystallogr. 66, 125–132.

Kozakov, D., Brenke, R., Comeau, S.R., Vajda, S., 2006. PIPER: an FFT-based protein docking program with pairwise potentials. Proteins Struct. Funct. Bioinf. 65, 392–406.

Kozakov, D., Beglov, D., Bohnuud, T., Mottarella, S.E., Xia, B., Hall, D.R., Vajda, S., 2013. How good is automated protein docking? Proteins Struct. Funct. Bioinf. 81, 2159–2166.

Lee, J.-S., Shilatifard, A., 2007. A site to remember: H3K36 methylation a mark for histone deacetylation. Mutat. Res. 618, 130–134.

Lees-Murdock, D.J., Shovlin, T.C., Gardiner, T., De Felici, M., Walsh, C.P., 2005. DNA methyltransferase expression in the mouse germ line during periods of de novo methylation. Dev. Dyn. 232, 992–1002.

Li, B.-Z., Huang, Z., Cui, Q.-Y., Song, X.-H., Du, L., Jeltsch, A., Chen, P., Li, G., Li, E., Xu, G.-L., 2011. Histone tails regulate DNA methylation by allosterically activating de novo methyltransferase. Cell Res. 21, 1172–1181.

Li, H., Rauch, T., Chen, Z.-X., Szabó, P.E., Riggs, A.D., Pfeifer, G.P., 2006. The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A interact directly and localize to promoters silenced in cancer cells. J. Biol. Chem. 281, 19489–19500.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M., 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462, 315–322.

Lukasik, S.M., Cierpicki, T., Borloz, M., Grembecka, J., Everett, A., Bushweller, J.H., 2006. High resolution structure of the HDGF PWWP domain: a potential DNA binding domain. Protein Sci. 15, 314–323.

Lyskov, S., Gray, J.J., 2008. The RosettaDock server for local protein–protein docking. Nucleic Acids Res. 36, W233–W238.

Lyskov, S., Chou, F.-C., Conchúir, S.Ó., Der, B.S., Drew, K., Kuroda, D., Xu, J., Weitzner, B.D., Renfrew, P.D., Sripakdeevong, P., 2013. Serverification of molecular modeling applications: the Rosetta Online Server that Includes Everyone (ROSIE). PLoS One 8 e63906.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674.

Morselli, M., Pastor, W.A., Montanini, B., Nee, K., Ferrari, R., Fu, K., Bonora, G., Rubbi, L., Clark, A.T., Ottonello, S., 2015. In vivo targeting of de novo DNA methylation by histone modifications in yeast and mouse. Elife 4 e06205.

Musselman, C.A., Avvakumov, N., Watanabe, R., Abraham, C.G., Lalonde, M.-E., Hong, Z., Allen, C., Roy, S., Nuñez, J.K., Nickoloff, J., 2012. Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. Nat. Struct. Mol. Biol. 19, 1266–1272.

O'Gara, M., Klimasauskas, S., Roberts, R.J., Cheng, X., 1996. Enzymatic C5-cytosine methylation of DNA: mechanistic implications of new crystal structures for HhaL methyltransferase-DNA-AdoHcy complexes. J. Mol. Biol. 261, 634–645.

O'Gara, M., Horton, J.R., Roberts, R.J., Cheng, X., 1998. Structures of HhaI methyltransferase complexed with substrates containing mismatches at the target base. Nat. Struct. Mol. Biol. 5, 872–877.

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

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

Pei, J., Kim, B.-H., Grishin, N.V., 2008. PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic Acids Res. 36, 2295–2300.

Pena, P.V., Davrazou, F., Shi, X., Walter, K.L., Verkhusha, V.V., Gozani, O., Zhao, R., Kutateladze, T.G., 2006. Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. Nature 442, 100–103.

Pierce, B.G., Wiehe, K., Hwang, H., Kim, B.-H., Vreven, T., Weng, Z., 2014. ZDOCK server: interactive docking prediction of protein–protein complexes and symmetric multimers. Bioinformatics 30, 1771–1773.

Pradeepa, M.M., Sutherland, H.G., Ule, J., Grimes, G.R., Bickmore, W.A., 2012. Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. PLoS Genet. 8 e1002717.

Purdy, M.M., Holz-Schietinger, C., Reich, N.O., 2010. Identification of a second DNA binding site in human DNA methyltransferase 3A by substrate inhibition and domain deletion. Arch. Biochem. Biophys. 498, 13–22.

Qin, S., Min, J., 2014. Structure and function of the nucleosome-binding PWWP domain. Trends Biochem. Sci. 39, 536–547.

Qiu, C., Sawada, K., Zhang, X., Cheng, X., 2002. The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. Nat. Struct. Mol. Biol. 9, 217–224.

Raveh, B., London, N., Schueler-Furman, O., 2010. Sub-angstrom modeling of complexes between flexible peptides and globular proteins. Proteins Struct. Funct. Bioinf. 78, 2029–2040.

Rigbolt, K.T., Prokhorova, T.A., Akimov, V., Henningsen, J., Johansen, P.T., Kratchmarova, I., Kassem, M., Mann, M., Olsen, J.V., Blagoev, B., 2011. System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation. Sci. Signal. 4 (rs3-rs3).

Rohl, C.A., Strauss, C.E., Misura, K.M., Baker, D., 2004. Protein structure prediction using Rosetta. Methods Enzymol. 383, 66–93.

Sanulli, S., Justin, N., Teissandier, A., Ancelin, K., Portoso, M., Caron, M., Michaud, A., Lombard, B., Da Rocha, S.T., Offer, J., 2015. Jarid2 methylation via the PRC2 complex regulates H3K27me3 deposition during cell differentiation. Mol. Cell 57, 769–783.

Sarraf, S.A., Stancheva, I., 2004. Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. Mol. Cell 15, 595–605.

Shirohzu, H., Kubota, T., Kumazawa, A., Sado, T., Chijiwa, T., Inagaki, K., Suetake, I., Tajima, S., Wakui, K., Miki, Y., 2002. Three novel DNMT3B mutations in Japanese patients with ICF syndrome. Am. J. Med. Genet. 112, 31–37.

Smallwood, S.A., Tomizawa, S.-I., Krueger, F., Ruf, N., Carli, N., Segonds-Pichon, A., Sato, S., Hata, K., Andrews, S.R., Kelsey, G., 2011. Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. Nat. Genet. 43, 811–814.

Stewart, K.R., Veselovska, L., Kim, J., Huang, J., Saadeh, H., Tomizawa, S.-I., Smallwood, S.A., Chen, T., Kelsey, G., 2015. Dynamic changes in histone modifications precede de novo DNA methylation in oocytes. Genes Dev.

Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H., Tajima, S., 2004. DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. J. Biol. Chem. 279, 27816–27823.

Takeshima, H., Suetake, I., Tajima, S., 2008. Mouse Dnmt3a preferentially methylates linker DNA and is inhibited by histone H1. J. Mol. Biol. 383, 810–821.

Tomizawa, S., Nowacka-Woszuk, J., Kelsey, G., 2012. DNA methylation establishment during oocyte growth: mechanisms and significance. Int. J. Dev. Biol. 56, 867–875.

van Nuland, R., van Schaik, F., Simonis, M., van Heesch, S., Cuppen, E., Boelens, R., Timmers, H., van Ingen, H., 2013. Nucleosomal DNA binding drives the recognition of H3K36-methylated nucleosomes by the PSIP1-PWWP domain. Epigenetics Chromatin 6, 12.

Velasco, G., Hubé, F., Rollin, J., Neuillet, D., Philippe, C., Bouzinba-Segard, H., Galvani, A., Viegas-Péquignot, E., Francastel, C., 2010. Dnmt3b recruitment through E2F6 transcriptional repressor mediates germ-line gene silencing in murine somatic tissues. Proc. Natl. Acad. Sci. 107, 9281–9286.

Vezzoli, A., Bonadies, N., Allen, M.D., Freund, S.M., Santiveri, C.M., Kvinlaug, B.T., Huntly, B.J., Göttgens, B., Bycroft, M., 2010. Molecular basis of histone H3K36me3 recognition by the PWWP domain of Brpf1. Nat. Struct. Mol. Biol. 17, 617–619.

Visualizer, D.S., 2013. Release 4.0. Accelrys Software Inc., San Diego, CA, USA.

Wang, J., Qin, S., Li, F., Li, S., Zhang, W., Peng, J., Zhang, Z., Gong, Q., Wu, J., Shi, Y., 2014. Crystal structure of human BS69 Bromo-ZnF-PWWP reveals its role in H3K36me3 nucleosome binding. Cell Res. 24, 890–893.

Wu, H., Zeng, H., Lam, R., Tempel, W., Amaya, M.F., Xu, C., Dombrovski, L., Qiu, W., Wang, Y., Min, J., 2011. Structural and histone binding ability characterizations of human PWWP domains. PLoS One 6 e18919.

Yazdi, P.G., Pedersen, B.A., Taylor, J.F., Khattab, O.S., Chen, Y.-H., Chen, Y., Jacobsen, S. E., Wang, P.H., 2015. Nucleosome organization in human embryonic stem cells. PLoS One 10, e0136314.

Yu, Q., Wen, Z., Chen, Z., Yan, W., Weiwei, W., Jiahai, Z., Zhiyong, Z., Guohong, L., Yunyu, S., Xiaoming, T., 2012. Solution structure of the Pdp1 PWWP domain reveals its unique binding sites for methylated H4K20 and DNA. Biochem. J. 442, 527–538.

Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., Fischle, W., 2010. Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. Nucleic Acids Res. 38, 4246–4253.