Cell Reports
Report

Cell PRESS
Open ACCESS

Molecular Basis for the Regulation of the H3K4 Methyltransferase Activity of PRDM9

Hong Wu,[1] Nikolas Mathioudakis,[2,3] Boubou Diagouraga,[4] Aiping Dong,[1] Ludmila Dombrovski,[1] Frédéric Baudat,[4] Stephen Cusack,[2,3] Bernard de Massy,[4,*] and Jan Kadlec[2,3,*]
[1]Structural Genomics Consortium, University of Toronto, MaRS Center, South Tower, 7th Floor, 101 College Street, Toronto, ON M5G 1L7, Canada
[2]European Molecular Biology Laboratory, Grenoble Outstation, 6 rue Jules Horowitz, 38042 Grenoble, France
[3]Unit for Virus Host-Cell Interactions, University Grenoble Alpes-EMBL-CNRS, 6 rue Jules Horowitz, 38042 Grenoble, France
[4]Institute of Human Genetics, UPR 1142, CNRS, 141 rue de la Cardonille, 34396 Montpellier, France
*Correspondence: bernard.de-massy@igh.cnrs.fr (B.d.M.), kadlec@embl.fr (J.K.)
http://dx.doi.org/10.1016/j.celrep.2013.08.035

SUMMARY

PRDM9, a histone lysine methyltransferase, is a key determinant of the localization of meiotic recombination hot spots in humans and mice and the only vertebrate protein known to be involved in hybrid sterility. Here, we report the crystal structure of the PRDM9 methyltransferase domain in complex with a histone H3 peptide dimethylated on lysine 4 (H3K4me2) and S-adenosylhomocysteine (AdoHcy), which provides insights into the methyltransferase activity of PRDM proteins. We show that the genuine substrate of PRDM9 is histone H3 lysine 4 (H3K4) and that the enzyme possesses mono-, di-, and trimethylation activities. We also determined the crystal structure of PRDM9 in its autoinhibited state, which revealed a rearrangement of the substrate and cofactor binding sites by a concerted action of the pre-SET and post-SET domains, providing important insights into the regulatory mechanisms of histone lysine methyltransferase activity.

INTRODUCTION

In humans, the PRDM family consists of 17 proteins that play important roles in a wide range of development processes, including stem cell identity maintenance or cell differentiation and their deregulation results in various cancers (Hohenauer and Moore, 2012). PRDM proteins are characterized by the presence of an N-terminal PR (PRDI-BF1 and RIZ1 homology) domain followed by multiple zinc fingers which confer DNA binding activity. PRDM2, -3, -8, -9, and -16 were shown to possess histone methyltransferase activity, whereas some other members were suggested to recruit histone-modifying enzymes (Eom et al., 2009; Hayashi et al., 2005; Hohenauer and Moore, 2012; Pinheiro et al., 2012). PR domains are only distantly related to the classical SET methyltransferase domains. The key signature motifs of SET domains are poorly conserved in the PRDM family, and the available crystal structures systematically lack bound substrate and cofactor. Thus, it remains unclear how the catalytic mechanism of PRDM proteins relates to that of SET domains (Hohenauer and Moore, 2012).

PRDM9 is the only member of the family (apart from its paralog PRDM7 in human) to contain a domain related to *Kr*üppel-*a*ssociated *b*ox (KRAB) often present in transcription repressors (Birtle and Ponting, 2006) but has been shown to catalyze methylation of H3K4me2 to H3K4me3 (Hayashi et al., 2005), a modification typically enriched at transcription start sites (Barski et al., 2007). PRDM9 may actually not be involved in transcription regulation but is a key factor in specifying the sites of meiotic recombination in mouse and human (Baudat et al., 2010; Myers et al., 2010; Parvanov et al., 2010). This function is mediated by the DNA binding specificity of its zinc finger array, and it has been proposed that PRDM9 binds to specific sites in the genome of oocytes and spermatocytes, where its methyltransferase activity leads to a local enrichment of H3K4me3 and recruits the meiotic recombination machinery (Grey et al., 2011). Whereas the importance of H3K4me3 catalyzed by Set1 in the activity of meiotic recombination sites was recently demonstrated in *S. cerevisiae* (Acquaviva et al., 2013; Sommermeyer et al., 2013), in mammals, the role of H3K4me3 in this process remains unclear. Furthermore, PRDM9 is involved in hybrid sterility, a function potentially linked to its role in recombination (Mihola et al., 2009). Here, we present crystal structures of its catalytic domain in an autoinhibited state and in complex with the H3K4me2 substrate peptide, providing insights into the substrate specificity, catalytic mechanism, and regulation of PRDM proteins.

RESULTS AND DISCUSSION

Structure of the PRDM9 PR/SET Domain in Complex with the H3K4me2 and AdoHcy

The structure of the PR/SET domain of mouse PRDM9 (mPRDM9) methyltransferase (residues 198–368) in complex with histone H3K4me2 peptide and AdoHcy (a methylation

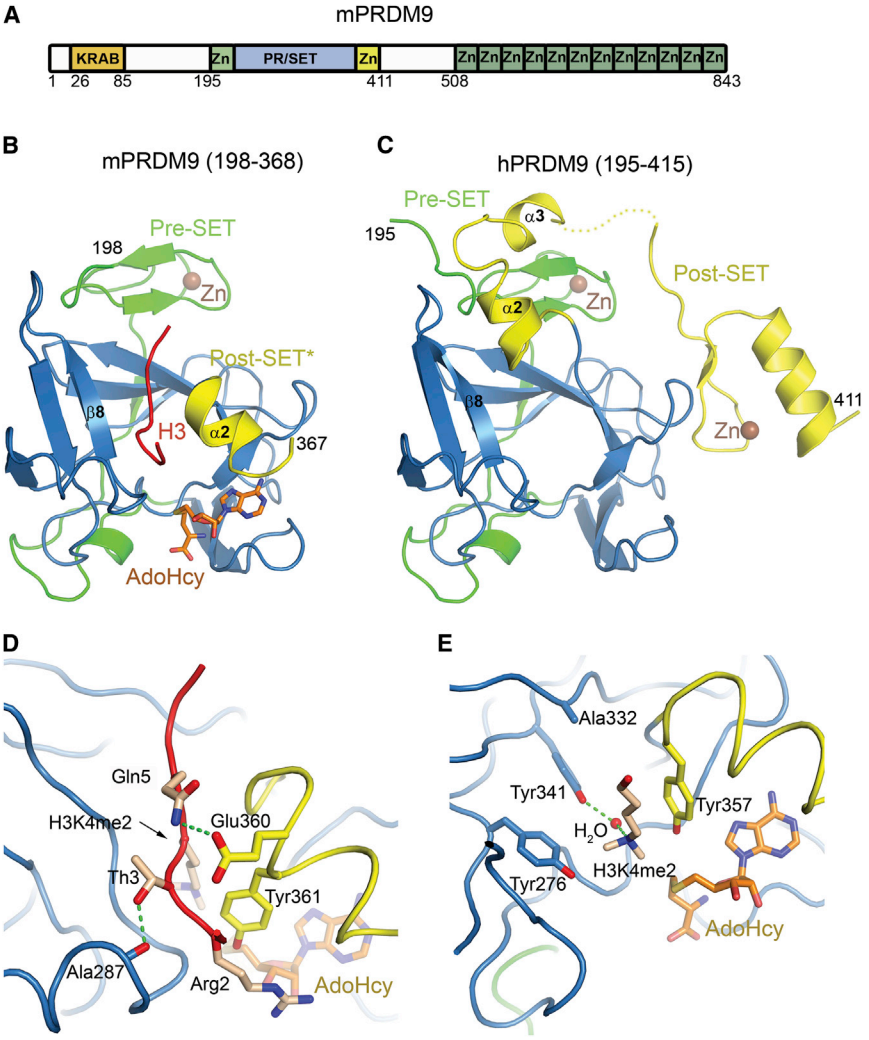

Figure 1. Crystal Structures of the Catalytic Domain of PRDM9

(A) Schematic representation of the domain structure of mouse PRDM9.

(B) Ribbon diagram of the mPRDM9 PR/SET domain in complex with H3K4me2 peptide and AdoHcy. The SET domain (residues 245–358) is shown in blue, pre-SET domain in green, and truncated post-SET (post-SET*) in yellow. The H3K4me2 peptide (in red) binds between strand β8 and helix α2.

(C) Structure of the catalytic domain of hPRDM9 in its unbound form colored as in (B). Helix α2 is oriented in the opposite direction blocking the substrate binding cleft. The structure contains an additional Zn finger within its post-SET domain.

(D) Details of H3K4me2 peptide recognition by mPRDM9. Only hydrogen bonds formed by peptide side chains are shown.

(E) The environment of H3K4me2 ε-amino group in the mPRDM9 catalytic site. Ala332 corresponds to F/Y switch residue Y305 in SETD7 (Del Rizzo et al., 2010).

See also Figure S1.

reaction product) was determined by X-ray crystallography (Figure 1). The asymmetric unit contains two mPRDM9 molecules: one in a ternary complex with the H3 peptide and AdoHcy (Figure 1B) and one in its unbound form (Figures S1A and S1B). Overall, the mPRDM9 PR/SET domain topology corresponds to other SET domains, with the conserved central SET domain fold (Figures 2A, 2B, and S2) flanked by pre-SET and post-SET regions. The pre-SET domain spans residues 198–244 and forms a Zn finger that is connected to the SET domain with a long linker (residues 222–244). The construct used contains only a short post-SET region fragment (residues 359–368; referred to as post-SET*) that is only folded in the presence of AdoHcy and the peptide (Figure S1B).

Although the two SET domain signature motifs (NHS/CxxPN and ELxF/YDY; x being any amino acid; Qian and Zhou, 2006) are not well conserved in the PRDM family, the AdoHcy is bound by mPRDM9 in a conventional way (Figures 2C and 2D), indicating that only Asn320 of the NHS/CxxPN motif (320-NCARDDEEQN in mPRDM9) is necessary for cofactor binding. Surprisingly, in PRDM3, -8, and -16 that were shown to possess methyltransferase activity (Eom et al., 2009; Hohenauer and Moore, 2012; Pinheiro et al., 2012), this invariant asparagine is replaced by an arginine or glutamine residue (Figure 2C). These proteins, thus, do not seem to have the potential of binding AdoHcy in the way that is highly conserved among SET-domain-containing methyltransferases and PRDM9. It is not clear whether arginine or glutamine residues can functionally substitute for Asn320. In PRDM4, for which the methyltransferase activity has not been reported, we could show that the corresponding arginine residue can block the AdoHcy binding site (Figure 2E).

The H3K4me2 peptide binds to PRDM9 in a cleft between strand β8 and helix α2 similar to other SET domains (Figures 1D and 1E). Electron density was interpretable for residues 1–7. In addition to backbone interactions with strand β8 and helix α2, the peptide also makes several side chain contacts with PRDM9 that might determine its specific recognition (Figure 1D). These include a hydrogen bond between Gln5 and the post-SET Glu360 and Thr3 interaction with backbone carbonyl of Ala287. Arg2 stacks against Tyr361 and makes a hydrogen bond with AdoHcy.

Substrate Specificity and the Active Site of PRDM9

PRDM9 was proposed to methylate H3K4me2, but not H3K4 nor H3K4me1 (Hayashi et al., 2005). We analyzed the ability of mPRDM9 (198–368) to bind unmodified H3K4 and H3K4me2 peptides by isothermal titration calorimetry and could show that mPRDM9 binds the two peptides with dissociation constants of 102 μM and 43 μM, respectively (Figures S3A and

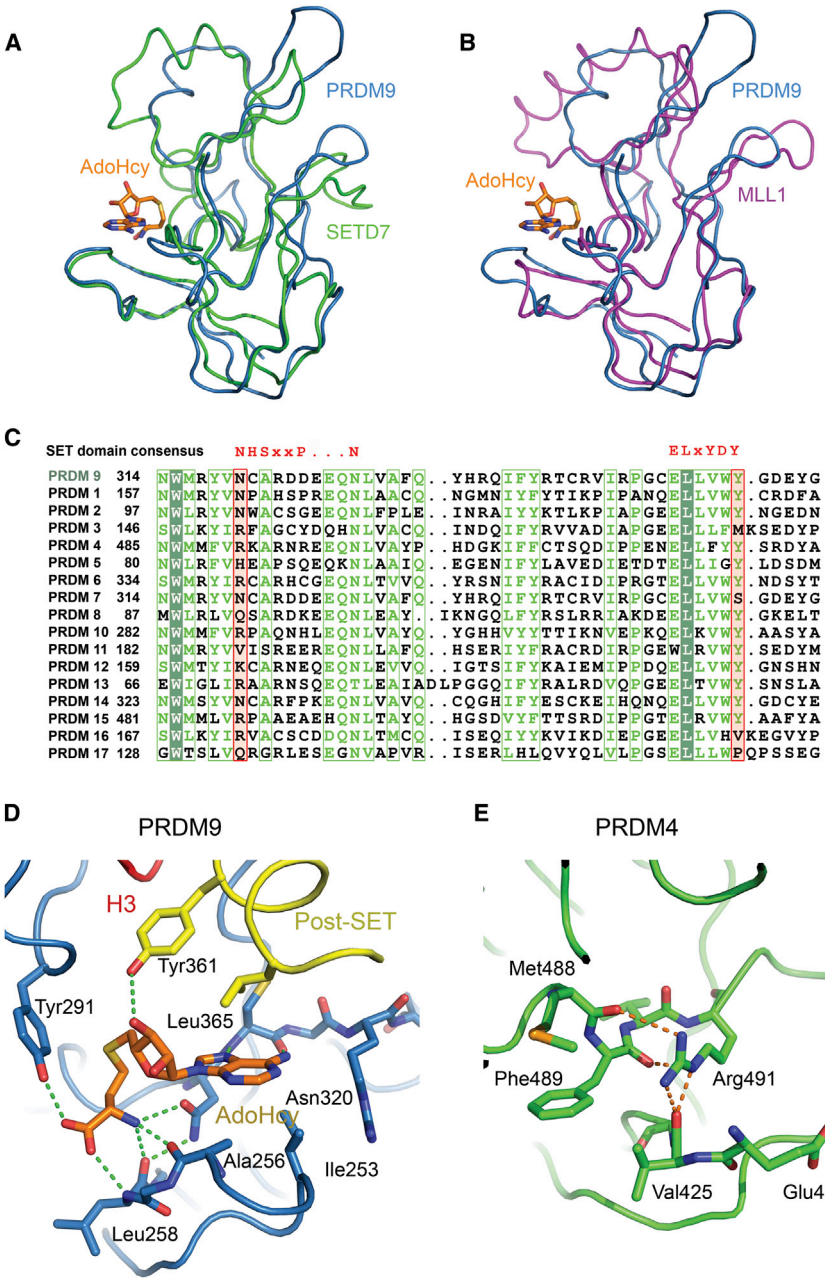

Figure 2. AdoHcy Binding by PRDM Proteins

(A) Comparison of the core SET domain of PRDM9 (residues 245–358) with residues 215–336 of SETD7 (PDB code 1O9S; root-mean-square deviation of 1.83 Å for 98 Cα atoms).

(B) Superposition of the core SET domain of PRDM9 onto MLL1 residues 3830–3945 (PDB code 2W5Z; root-mean-square deviation of 1.6 Å for 96 Cα atoms).

(C) Sequence alignment of PRDM proteins covering the two SET domain consensus regions (in red). Identical residues are in green boxes. Residues corresponding to PRDM9 Asn320 (AdoHcy binding) and Tyr357 (active site), invariant in other SET domains, are highlighted by a red frame.

(D) Details of AdoHcy binding by PRDM9. The AdoHcy interacts with Asn320, main chain atoms of the AGLG motif (Gly257, Gly259), Tyr 291 of the I-SET, and Tyr361 and Leu365 of the post-SET helix α2.

(E) Arg491 in PRDM4 (PDB code 3DB5) corresponding to PRDM9 Asn320 blocks the AdoHcy binding site.

See also Figure S2.

S3B). Additionally, MALDI-TOF mass spectrometry analysis revealed that mPRDM9 can mono-, di-, and trimethylate unmodified H3K4 peptide (Figures 3A and 3B). Finally, in a coupled fluorescent methylation assay, both H3K4 and H3K4me2 peptides were efficiently methylated by mPRDM9 (198–368), whereas H4K20 peptide was not (Figure 3C). The activity on H3K4 was lower but still significant, probably reflecting the lower binding affinity (Figure 3C). Together, these results clearly show that unmodified H3K4 is a genuine substrate of mPRDM9 and that mPRDM9 possesses mono-, di-, and trimethylation activities.

To explore more widely the substrate specificity of PRDM9, we tested the activity of mPRDM9 (198–368) on an array of 379 distinct peptides corresponding to different regions of H2A, H2B, H3, and H4 with various modifications (Figure S3C; Table S1). Significant methyltransferase activity was detected on H3 (1–19) peptides, confirming H3K4 (me0, -1, and -2) to be efficient substrates of PRDM9 (Figure 3D). Arg2 methylation had no significant effect on the H3K4 methylation within H3 (1–19) peptides carrying K9me3 modification (Figure 3D). Surprisingly, PRDM9 can also methylate H3K9 (me0, -1, and, -2) substrates (Figure 3D) and potentially H3K36 based on the signal observed on the H3 (26–45) peptide (Figure S3C). Although no enrichment for H3K9me3 could be detected at a mouse meiotic recombination site bound by PRDM9 (Buard et al., 2009), additional assays for substrate specificities in vitro and in vivo could certainly be interesting. Other SET domain methyltransferases, such as Ash1 from *D. melanogaster*, do display activities on various substrates which may be regulated in vivo (Beisel et al., 2002).

H3K4me2 is located in the channel formed by Trp293, Tyr357, and Tyr361. In the active site, the H3K4me2 ε-amino group is surrounded by three conserved tyrosine residues: Tyr357, Tyr276, and Tyr341 (Figure 1E). Invariant tyrosines corresponding to Tyr357 have been proposed to be directly involved in catalysis (Smith and Denu, 2009). We show that the Y357F mutation completely abolished the in vitro activity of mPRDM9 (Figures 3E and S3C). Similarly, Y276F and Y341F mutants

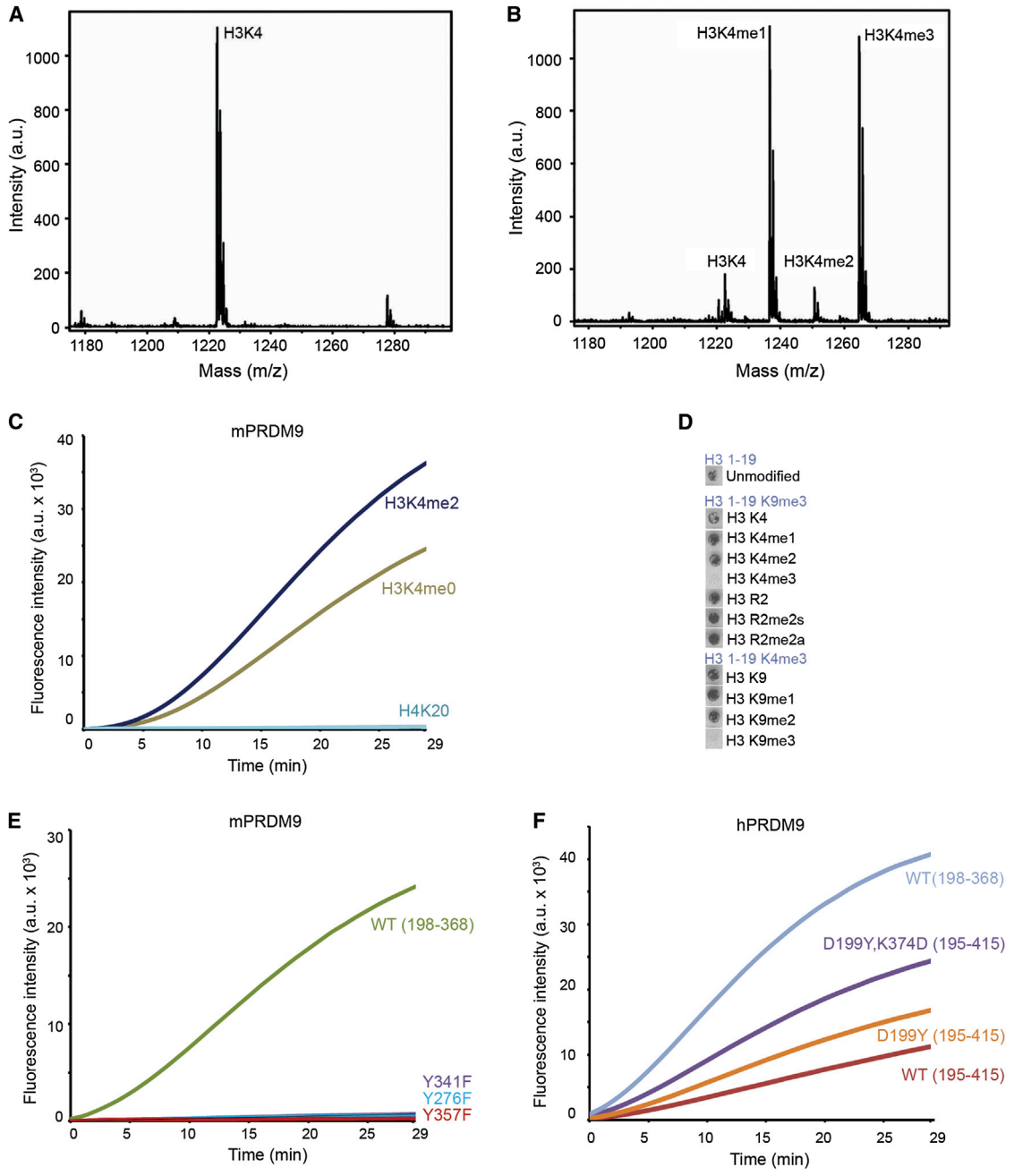

Figure 3. Substrate Specificity of mPRDM9

(A and B) MALDI-TOF mass spectrometry analysis of the H3K4 peptide methylation by PRDM9 (198–368) is shown. Whereas H3K4 peptide incubated in the absence of mPRDM9 is not methylated (A), after incubation with mPRDM9 (198–368), single, double, and triple methylation could be detected (B). a.u., arbitrary units.

(C) Coupled fluorescent methylation assay. H3K4 and H3K4me2 peptides are efficiently methylated by mPRDM9. No methylation is observed for H4K20 peptide.

(D) mPRDM9 methyltransferase activity on histone H3 peptides. Top panel shows methylation of H3 (1–19) peptide. Middle panel shows methylation of H3 (1–19) R8me2s/K9me3 peptides with various levels of K4 methylation and of K4me1/R8me2s/K9me3 peptide with indicated R2 modifications. The bottom panel shows R2me2s/R8me2s/K4me3 peptides with various levels of K9 methylation. No methylation is detected on H3 (1–19) K4me3/K9me3 peptides.

(E) Mutagenesis of key catalytic residues. Y357F, Y276F, and Y341F mutations abolish the methyltransferase activity of mPRDM9 (198–368).

(F) hPRDM9 (195–415) can methylate the H3K4 peptide. The methylation activity is ~4.4 times lower than for hPRDM9 (198–368). Mutations of Asp199 of the pre-SET domain and Lys374 of the post-SET domain, disrupting the pre-SET/post-SET-inhibitory interaction result in a higher methyltransferase activity (1.6 times higher for D199Y and 2.4 times higher for D199Y, K374D). The methyltransferase activity was calculated using linear portion of the curves between minute 7 and 15. Gel filtration and thermal shift measurement of melting temperature confirmed the structural integrity of these mutants (data not shown).

See also Figure S3.

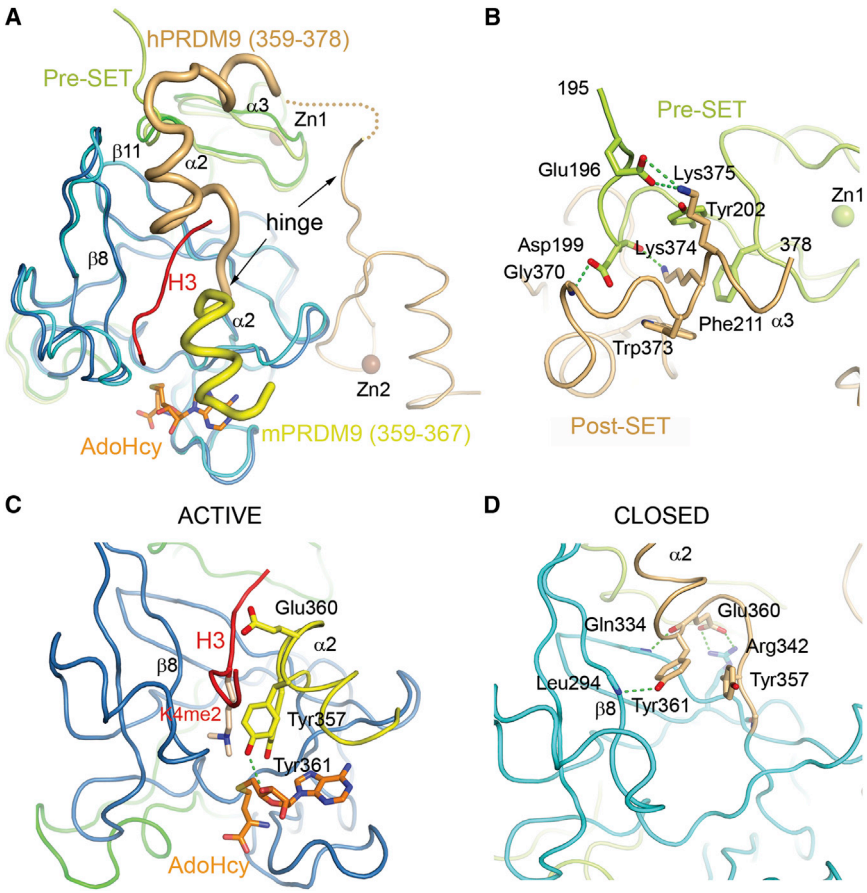

Figure 4. Conformational Changes of the PRDM9 Active Site

(A) Superposition of mPRDM9 (198–368) and hPRDM9 (195–415) structures. The post-SET domains are highlighted in yellow and brown, respectively. Proposed hinge regions are shown.

(B) Details of the interaction between the pre-SET and post-SET domain of hPRDM9. Gly370 forms a hydrogen bond with Asp199. The side chain of Lys374, which is surrounded by Tyr202, Phe211, and Trp373, binds to backbone carbonyl of Asp199. Lys375 makes a salt bridge with Glu196.

(C) Glu360 and Tyr361 participate in the substrate and cofactor binding in mPRDM9. Coloring corresponds to Figure 1B.

(D) Glu360 stabilizes the autoinhibitory conformation of hPRDM9 by a salt bridge with Arg342, and Tyr361 blocks the substrate binding cleft. The catalytic Tyr357 is also displaced.

See also Figure S4.

were inactive in methylation assay using the H3K4 peptide (Figure 3E). Surprisingly, Tyr276 was previously reported to be dispensable for methyltransferase activity on H3 (Hayashi et al., 2005). The integrity of the mutated proteins was verified by gel filtration and thermal shift measurement of melting temperature (Y276F decreased thermal stability of PRDM9 by 5°C; data not shown). Interestingly, Tyr357 is not conserved in PRDM3 and PRDM16 (Figure 2C). Thus, these PRDM proteins are either not truly active methyltransferases or they use different catalytic residues than those conserved in PRDM9 and all other SET domains. Tyr341 helps coordinate a water molecule in the active site (Figure 1E). In our structure, one of the methyl groups of the H3K4me2 would clash with the modeled position of the S-adenosyl methionine (AdoMet) donor methyl. Thus, we interpret our structure as being in a product conformation. Before the final methylation reaction, the lysine would need to be rotated within the catalytic site, possibly to position the clashing methyl group into that occupied by the active site water, as suggested for SETD7 (Del Rizzo et al., 2010). In zebrafish, both Tyr276 and Tyr341 are substituted with phenylalanine (Figure S4A), possibly indicating altered catalytic activity. Gly278 was suggested to be a key residue for the activity of mPRDM9 (Hayashi et al., 2005). The structure reveals that Gly278 is located just upstream of strand β7 forming a β sheet with β8 and β9 and is not in proximity to the substrate or cofactor binding site (11 Å from the H3K4me2 ε-amino group). Thus, this residue is unlikely to have a direct effect on catalysis but appears to be important for structural integrity of the domain because a G278A mutant cannot be expressed in soluble form in bacteria (data not shown).

Structure of the Autoinhibited PRDM9 PR/SET Domain

The post-SET domain is disordered in the unbound mPRDM9 molecule in the crystal (Figure S1B). Similarly, the post-SET domains of DIM-5 or MLL1 become fully folded only in the presence of the AdoHcy/AdoMet cofactor and substrate (Southall et al., 2009; Zhang et al., 2002). To better understand the role of the post-SET domain, we produced crystals of a larger fragment of mouse PRDM9 that, however, only diffracted to a low resolution. We were, however, able to solve the structure of human PRDM9 fragment spanning residues 195–415 (hPRDM9; 90% sequence identity with mPRDM9; Figure 1C). Compared to mPRDM9 (198–368), this fragment includes a slightly longer pre-SET region and an extended post-SET domain including a Zn finger, which the structure shows loosely associates with the SET domain via Glu326 and Arg345. Despite the presence of the substrate peptide and AdoHcy in the crystallization solution, hPRDM9 (195–415) crystallized in an unbound form. Unexpectedly, the hPRDM9 structure reveals that, in the absence of the substrate and cofactor, the post-SET region is ordered but undergoes a conformational change and binds across the SET domain, blocking the peptide binding cleft (Figure 4A). Consequently, also the AdoHcy binding site becomes incomplete suggesting that this is an autoinhibitory conformation. This position of the post-SET domain is stabilized by several interactions with the pre-SET domain Zn finger (Figures 4A and 4B). The last residue in the same position between the two structures is Val355. The following residues that form helix α2 (residues 359–364) in the ligand bound structure are rotated by ~180° around Val355

(Figure 4A). The new helix α2 is then formed by residues 362–367 packing against strands β8 and β11. Glu360 and Trp356 make backbone hydrogen bonds with Leu332 and Gln334 of β11. Glu360, which is involved in the substrate recognition in the complex structure, makes a salt bridge interaction with Arg342, whereas Tyr361 that forms the substrate lysine binding channel in the complex blocks the peptide binding cleft forming a hydrogen bond with Leu294 (Figures 4C and 4D). The following helix α3 packs against the pre-SET domain as shown in Figure 4B. We tested whether hPRDM9 (195–415) maintains methyltransferase activity and find that it is reduced from that of hPRDM9 (198–368) but still significant (Figure 3F). Thus, it seems that the presence of AdoMet and the peptide substrate is sufficient to overcome the inhibition and trigger hPRDM9 methylation activity. Additionally, we could show that mutations predicted to disrupt the inhibitory interaction between the pre-SET and post-SET domains (D199Y and K374D) increased the activity of hPRDM9 (195–415; Figure 3F). We propose that PRDM9 residues 355–356 and the linker between helix α3 and the post-SET Zn finger function as hinge regions that enable translocation of the helix α2 between the active conformation where Glu360 and Tyr361 participate in substrate lysine and cofactor binding and an inhibitory conformation stabilized by Glu360, where Tyr361 blocks the substrate binding site (Figure 4A). It remains to be established whether the transition between the stable rather than disordered conformation of the post-SET domain and its active state is regulated by other factors or posttranslational modifications during meiosis. This is a unique example of a complete rearrangement of the substrate and cofactor binding sites of a histone lysine methyltransferase by a concerted action of the pre-SET and post-SET domains.

EXPERIMENTAL PROCEDURES

Expression, Purification, and Crystallization

A His-tag fusion of mouse PRDM9 (residues 198–368) was expressed in *E. coli* BL21Star (DE3) from pProEXHTb vector. The protein was first purified by affinity chromatography using Ni^{2+} resin. After His-tag cleavage with tobacco etch virus (TEV) protease, the protein was further purified by a second Ni^{2+} column and size-exclusion chromatography. Pure PRDM9 was concentrated with 2 mM AdoHCy and 2 mM H3K4me2 peptide (ARTKme2QTARK-Y; Y added to facilitate quantification) to 9 mg/ml in a buffer containing 20 mM Tris, pH 7.0, 150 mM NaCl, and 5 mM β-mercaptoethanol. The best-diffracting crystals grew within 3 days at 5°C in a solution containing 0.2 M ammonium sulfate; 0.1 M Bis-Tris, pH 5.5; and 25% w/v polyethylene glycol (PEG) 3350. For data collection at 100 K, crystals were snap frozen in liquid nitrogen with a solution containing mother liquor and 30% (v/v) glycerol.

hPRDM9 (residues 195–415) was expressed as a His-tag fusion in *E. coli* BL21 (DE3) V2R-pRARE from pET28-MHL vector. Harvested cells were resuspended in phosphate-buffered saline, supplemented with 250 mM NaCl, 5 mM imidazole, 3 mM β-mercaptoethanol, 5% glycerol, 0.1% 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate, and 1 mM phenylmethanesulfonylfluoride. After affinity purification on 5 ml HiTrap Chelating column (GE Healthcare), the protein was further purified by size-exclusion chromatography. After TEV protease cleavage, hPRDM9 was purified to homogeneity by ion-exchange chromatography on Source 30Q column (10 × 10) (GE Healthcare). Pure hPRDM9 protein (8 mg/ml) was crystallized in a solution containing 23% PEG 3350; 0.2 M ammonium acetate; and 0.1 M BisTris, pH 5.5. The crystal was frozen in liquid nitrogen using 15% ethylene glycol as cryoprotectant.

hPRDM4 (residues 390–540) was expressed as a His-tag fusion in *E. coli* BL21 (DE3) Codon Plus RIL (Stratagene) from pET28a-LIC vector. The overexpressed protein was purified as described above for hPRDM9. Purified PRDM4 (10 mg/ml) was crystallized using hanging drop vapor diffusion method at 20°C by mixing 1.5 μl of the protein solution with 1.5 μl of the reservoir solution containing 23% PEG 3350; 0.2 M ammonium acetate; and 0.1 M BisTris, pH 6.5. The crystal was frozen in liquid nitrogen using Paratone-N as cryoprotectant.

Data Collection and Structure Determination

Crystals of mPRDM9 (198–368) belong to the space group $P2_12_12_1$ with unit cell dimensions a = 55.7 Å, b = 78.2 Å, and c = 107.6 Å. The asymmetric unit contains two PRDM9 molecules and has a solvent content of 60%. A complete native data set was collected to a resolution of 2.3 Å on beamline ID29 at the European Synchrotron Radiation Facility (ESRF). The data were processed using XDS (Kabsch, 2010). The structure was solved by molecular replacement with PHASER (McCoy et al., 2005) using the structure of PRDM11 (Protein Data Bank [PDB] code 3RAY; 44% sequence identity) as a search model. Using prime-and-switch density modification of program RESOLVE (Terwilliger, 2000) enabled to us to obtain a clearly interpretable map. The structure was built in COOT (Emsley and Cowtan, 2004) and refined in REFMAC5 (Murshudov et al., 1997) to final R factor of 20.8% and R_{free} of 24.5% (Table S2) with all residues in allowed (97% in favored) regions of the Ramachandran plot, as analyzed by MOLPROBITY (Davis et al., 2004).

X-ray diffraction data for hPRDM9 was collected at 100 K at beamline 19-ID of Advanced Photon Source at Argonne National Laboratory. Crystals of hPRDM9 (195–415) belong to the space group $P2_1$ with unit cell dimensions of a = 54.7 Å, b = 48.8 Å, c = 78.7 Å, and β = 100°. Data were processed using HKL-3000 suite (Otwinowski and Minor, 1997). The structure of hPRDM9 was determined by single-wavelength anomalous dispersion (SAD) phasing method using a SeMet-substituted crystal. SOLVE/RESOLVE (Terwilliger and Berendzen, 1999) were used for heavy atom site search, phase improvement, and initial model building. Program BUCCANEER (Cowtan, 2006) and ARP/wARP (Perrakis et al., 1999) were used for automatic model building. COOT (Emsley and Cowtan, 2004) was used for model building and visualization. Structure was refined with REFMAC5 (Murshudov et al., 1997) to final R factor of 19.6% and R_{free} of 26.4% with all residues in allowed (96% in favored) regions of the Ramachandran plot, as analyzed by MOLPROBITY (Davis et al., 2004). Crystal diffraction data and refinement statistics for the structure are displayed in Table S2. Representative parts of the $2F_o$–F_c electron density maps calculated using the refined models are shown in Figures S4B and S4C.

X-ray diffraction data for hPRDM4 were collected at 100 K on RIGAKU FR-E DW. Crystals of hPRDM4 belong to the space group *I*422 with unit cell dimensions of a = 107.3 Å, b = 107.3 Å, and c = 133.6 Å. Data were processed using HKL-2000 suite (Otwinowski and Minor, 1997). The structure of hPRDM4 was determined by SAD phasing method using a SeMet-substituted crystal. The structure was solved by using SHELXD (Sheldrick, 2008). COOT (Emsley and Cowtan, 2004) was used for model building and visualization. Structure was refined with REFMAC5 (Murshudov et al., 1997) to final R factor of 22.2% and R_{free} of 29.8% with all residues in allowed (92.1% in favored) regions of the Ramachandran plot, as analyzed by MOLPROBITY (Davis et al., 2004). Crystal diffraction data and refinement statistics for the structure are displayed in Table S2.

Methyltransferase Assays

Methyltransferase activity of PRDM9 was analyzed using a continuous, fluorescent-coupled assay SAMfluoro (G-Bioscience). The AdoHcy product of the methylation reaction is further converted by a supplied mixture of three enzymes to hydrogen peroxide that reacts with 10-acetyl-3,7-dihydroxyphenoxazine to produce a fluorescent compound Resorufin. The production of Resorufin was monitored for 30 min at room temperature using an Infinite 200 PRO plate reader (Tecan). The methylation reactions, containing 0.26 or 0.3 μM enzyme and 0.22 mM peptide substrate, were set up according to manufacturer instructions. The tested substrate includes H3K4 (ARTKQTARK-Y), H3K4me2 (ARTKme2QTARK-Y), and H4K20 (Y-GKGGKGLGKGGAKRHR KVLRD) peptides. All peptides include an extra tyrosine residue for quantification. All the reactions were performed at least in duplicates, and the differences between methylation rates calculated from the linear parts of corresponding curves (above background level) were within 10%.

For mass spectrometry analysis, 1 mM H3K4 peptide was incubated with 5 mM AdoMet in the presence or absence of 10 μM mPRDM9 (198–368) at room temperature for 2 hr in 20 mM Tris 7.0 and 150 mM NaCl. The reaction mixture was analyzed by MALDI-TOF mass spectrometry. The duplicate experiment yielded essentially identical results.

Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) experiments were performed at 25°C, using an ITC200 microcalorimeter (MicroCal). Experiments included 26 injections of 1.5 μl of 2.7 mM peptide solution into the sample cell containing 60 μM of mPRDM9 (198–368) in 20 mM Tris 7.0, 150 mM NaCl, and 5 mM β-mercaptoethanol. The initial data point was deleted, and the last point after saturation was subtracted from the data sets. Binding isotherms were fitted with a one-site binding model by nonlinear regression using Origin Software version 7.0 (MicroCal).

Methyltransferase Activity Assay on Peptide Array

Modified histone peptide arrays (Active Motif, catalog number 13005) were preincubated in methylation buffer (50 mM Tris-HCl, pH 9.0; 100 mM NaCl; 5 mM dithiothreitol; 10 μM $ZnCl_2$) for 20 min, followed by incubation with methylation buffer containing 0.58 μM of (methyl-3H)-S-adenosyl-L-methionine (Perkin Elmer) and 0.2 μM of either wild-type or Y357F mutant mPRDM9 (198–368) for 10, 30, or 60 min at ambient temperature. The arrays were washed four times for 5 min with washing buffer (50 mM NH_4HCO_3 and 0.1% SDS) and then dried and rinsed with Amplify NAMP100 solution (GE Healthcare). The arrays were completely dried and exposed to the Carestream Kodak Biomax MR films in darkness at −80°C for 1–3 days.

ACCESSION NUMBERS

Protein Data Bank coordinates for the crystal structure of the mPRDM9 (198–368) and hPRDM9 (195–415) have been deposited with accession codes 4C1Q and 4IJD, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2013.08.035.

ACKNOWLEDGMENTS

We thank the Grenoble European Molecular Biology Laboratory (EMBL)-European Synchrotron Radiation Facility (ESRF)-L'Institut Laue-Langevin (ILL)-Institut de Biologie Structurale (IBS) Partnership for Structural Biology for access to structural biology instrumentation, notably L. Signor of the IBS for his help with mass spectrometry, the High Throughput Crystallisation Laboratory group of EMBL Grenoble for performing initial screening crystallization trials, A. Palencia for help with the isothermal titration calorimetry, and J. Perard for help with the methylation assays. We thank the ESRF-EMBL Joint Structural Biology Group for access to and assistance on the ESRF synchrotron beamlines. This work was supported by the EU FP7-funded Network of Excellence EpiGeneSys awarded to S.C. and B.d.M. B.d.M. is supported by the Centre National de la Recherche Scientifique, the Agence Nationale de la Recherche (09-BLAN-0269-01), and the Fondation pour la Recherche Médicale. "The SGC" is a registered charity (number 1097737) that receives funds from AbbVie, Boehringer Ingelheim, the Canada Foundation for Innovation, the Canadian Institutes for Health Research, Genome Canada through the Ontario Genomics Institute (OGI-055), GlaxoSmithKline, Janssen, Lilly Canada, the Novartis Research Foundation, the Ontario Ministry of Economic Development and Innovation, Pfizer, Takeda, and the Wellcome Trust (092809/Z/10/Z).

Received: May 21, 2013
Revised: July 13, 2013
Accepted: August 21, 2013
Published: October 3, 2013

REFERENCES

Acquaviva, L., Székvölgyi, L., Dichtl, B., Dichtl, B.S., de La Roche Saint André, C., Nicolas, A., and Géli, V. (2013). The COMPASS subunit Spp1 links histone methylation to initiation of meiotic recombination. Science *339*, 215–218.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823–837.

Baudat, F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., Coop, G., and de Massy, B. (2010). PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. Science *327*, 836–840.

Beisel, C., Imhof, A., Greene, J., Kremmer, E., and Sauer, F. (2002). Histone methylation by the Drosophila epigenetic transcriptional regulator Ash1. Nature *419*, 857–862.

Birtle, Z., and Ponting, C.P. (2006). Meisetz and the birth of the KRAB motif. Bioinformatics *22*, 2841–2845.

Buard, J., Barthès, P., Grey, C., and de Massy, B. (2009). Distinct histone modifications define initiation and repair of meiotic recombination in the mouse. EMBO J. *28*, 2616–2624.

Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr. D Biol. Crystallogr. *62*, 1002–1011.

Davis, I.W., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2004). MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. Nucleic Acids Res. *32*(Web Server issue), W615–W619.

Del Rizzo, P.A., Couture, J.F., Dirk, L.M., Strunk, B.S., Roiko, M.S., Brunzelle, J.S., Houtz, R.L., and Trievel, R.C. (2010). SET7/9 catalytic mutants reveal the role of active site water molecules in lysine multiple methylation. J. Biol. Chem. *285*, 31849–31858.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Eom, G.H., Kim, K., Kim, S.M., Kee, H.J., Kim, J.Y., Jin, H.M., Kim, J.R., Kim, J.H., Choe, N., Kim, K.B., et al. (2009). Histone methyltransferase PRDM8 regulates mouse testis steroidogenesis. Biochem. Biophys. Res. Commun. *388*, 131–136.

Grey, C., Barthès, P., Chauveau-Le Friec, G., Langa, F., Baudat, F., and de Massy, B. (2011). Mouse PRDM9 DNA-binding specificity determines sites of histone H3 lysine 4 trimethylation for initiation of meiotic recombination. PLoS Biol. *9*, e1001176.

Hayashi, K., Yoshida, K., and Matsui, Y. (2005). A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. Nature *438*, 374–378.

Hohenauer, T., and Moore, A.W. (2012). The Prdm family: expanding roles in stem cells and development. Development *139*, 2267–2282.

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

McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., and Read, R.J. (2005). Likelihood-enhanced fast translation functions. Acta Crystallogr. D Biol. Crystallogr. *61*, 458–464.

Mihola, O., Trachtulec, Z., Vlcek, C., Schimenti, J.C., and Forejt, J. (2009). A mouse speciation gene encodes a meiotic histone H3 methyltransferase. Science *323*, 373–375.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

Myers, S., Bowden, R., Tumian, A., Bontrop, R.E., Freeman, C., MacFie, T.S., McVean, G., and Donnelly, P. (2010). Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. Science *327*, 876–879.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Parvanov, E.D., Petkov, P.M., and Paigen, K. (2010). Prdm9 controls activation of mammalian recombination hotspots. Science *327*, 835.

Perrakis, A., Morris, R., and Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. *6*, 458–463.

Pinheiro, I., Margueron, R., Shukeir, N., Eisold, M., Fritzsch, C., Richter, F.M., Mittler, G., Genoud, C., Goyama, S., Kurokawa, M., et al. (2012). Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity. Cell *150*, 948–960.

Qian, C., and Zhou, M.M. (2006). SET domain protein lysine methyltransferases: Structure, specificity and catalysis. Cell. Mol. Life Sci. *63*, 2755–2763.

Sheldrick, G.M. (2008). A short history of SHELX. Acta Crystallogr. A *64*, 112–122.

Smith, B.C., and Denu, J.M. (2009). Chemical mechanisms of histone lysine and arginine modifications. Biochim. Biophys. Acta *1789*, 45–57.

Sommermeyer, V., Béneut, C., Chaplais, E., Serrentino, M.E., and Borde, V. (2013). Spp1, a member of the Set1 Complex, promotes meiotic DSB formation in promoters by tethering histone H3K4 methylation sites to chromosome axes. Mol. Cell *49*, 43–54.

Southall, S.M., Wong, P.S., Odho, Z., Roe, S.M., and Wilson, J.R. (2009). Structural basis for the requirement of additional factors for MLL1 SET domain activity and recognition of epigenetic marks. Mol. Cell *33*, 181–191.

Terwilliger, T.C. (2000). Maximum-likelihood density modification. Acta Crystallogr. D Biol. Crystallogr. *56*, 965–972.

Terwilliger, T.C., and Berendzen, J. (1999). Automated MAD and MIR structure solution. Acta Crystallogr. D Biol. Crystallogr. *55*, 849–861.

Zhang, X., Tamaru, H., Khan, S.I., Horton, J.R., Keefe, L.J., Selker, E.U., and Cheng, X. (2002). Structure of the Neurospora SET domain protein DIM-5, a histone H3 lysine methyltransferase. Cell *111*, 117–127.