

Vincent Cura, Nathalie Troffer-Charlier, Jean-Marie Wurtz, Luc Bonnefond and Jean Cavarelli*

Département de Biologie Structurale Intégrative, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Université de Strasbourg, CNRS UMR7104, INSERM U596, 1 Rue Laurent Fries, F-67404 Illkirch, France

Correspondence e-mail: cava@igbmc.fr

Structural insight into arginine methylation by the mouse protein arginine methyltransferase 7: a zinc finger freezes the mimic of the dimeric state into a single active site

Protein arginine methyltransferase 7 (PRMT7) is a type III arginine methyltransferase which has been implicated in several biological processes such as transcriptional regulation, DNA damage repair, RNA splicing, cell differentiation and metastasis. PRMT7 is a unique but less characterized member of the family of PRMTs. The crystal structure of full-length PRMT7 from *Mus musculus* refined at 1.7 Å resolution is described. The PRMT7 structure is composed of two catalytic modules in tandem forming a pseudo-dimer and contains only one AdoHcy molecule bound to the N-terminal module. The high-resolution crystal structure presented here revealed several structural features showing that the second active site is frozen in an inactive state by a conserved zinc finger located at the junction between the two PRMT modules and by the collapse of two degenerated AdoMet-binding loops.

Received 25 April 2014

Accepted 18 June 2014

PDB reference: PRMT7, 4c4a

1. Introduction

Methylation of arginine is a prevalent post-translational modification found in eukaryotes, and is mediated by protein arginine methyltransferases (PRMTs). PRMTs are implicated in a wide array of biological processes such as regulation of transcription, RNA metabolism and DNA damage repair (Bedford & Clarke, 2009; Yang & Bedford, 2013). Since deregulation of these processes appears to be implicated in the pathogenesis of various diseases such as human cancers (Cha & Jho, 2012), understanding the mechanism of action of PRMTs at the atomic scale is therefore crucial for both fundamental biology and pharmacological applications. PRMTs catalyse the transfer of a methyl group from S-adenosyl-L-methionine (AdoMet) to the side-chain N atoms of arginine residues to form methylated arginines and S-adenosyl-L-homocysteine (AdoHcy). The target arginine residues, which are mainly located within glycine-rich and arginine-rich patches (GAR motifs), can be monomethylated or dimethylated. The members of mammalian PRMTs are classified into three main types based on their methylation activity: type I [such as PRMT1, PRMT3, PRMT6, PRMT8 and PRMT4, also known as coactivator-associated arginine methyltransferase 1 (CARM1)] dimethylate arginine asymmetrically, while type II (such as PRMT5) dimethylate arginine symmetrically and type III only monomethylate arginine. PRMT7 was first identified in a genetic screen for susceptibility to chemotherapeutic cytotoxicity (Gros *et al.*, 2003). PRMT7 was initially demonstrated to be capable of generating both monomethylated arginine (MMA) and symmetrically dimethylated arginine (sDMA) depending on the substrate that was used in the assays (Miranda *et al.*, 2004; Lee *et al.*, 2005). This activity has recently been revisited and PRMT7 is now defined as the sole known PRMT member that

Table 1

Data statistics for the MmPRMT7 structure.

Values in parentheses are for the outer shell.

Crystallization	
Reservoir	100 mM BTP pH 7.0, 12.6% PEG 8000, 3.2% PEG 20 000
Volume (μl)	1 + 1
Temperature ($^{\circ}\text{C}$)	17
Cryoprotectant	15% PEG 400
Data processing	
Wavelength (\AA)	0.979
Resolution range (\AA)	29.35–1.60 (1.63–1.60)
Space group	$P4_32_12$
Unit-cell parameters (\AA , $^{\circ}$)	$a = b = 97.3$, $c = 168.8$, $\alpha = \beta = \gamma = 90$
Total reflections	933784 (43825)
Unique reflections	107019 (5199)
Average multiplicity	8.7 (8.4)
Completeness (%)	99.9 (99.9)
$\langle I/\sigma(I) \rangle$	13.9 (0.5)
Wilson B factor (\AA^2)	23.1
$R_{\text{meas}}^{\dagger}$ (%)	8.0 (432.6)
$CC_{1/2}^{\ddagger}$	0.999 (0.181)
Refinement	
Resolution range (\AA)	29.35–1.70
$R_{\text{work}}/R_{\text{free}}$	0.160/0.190
No. of atoms	
Total	5882
Protein	5138
Ligands	52
Water	692
No. of protein residues	641
R.m.s. deviations \S	
From ideal bond lengths (\AA)	0.007
From ideal bond angles ($^{\circ}$)	1.13
Ramachandran favoured (%)	97.7
Ramachandran outliers (%)	0.16
Average B factor (\AA^2)	
Overall	35.7
Macromolecules	34.7
Ligands	48.9
Solvent	42.1
$MolProbity$ validation	
Rotamer outliers (%)	1.4
C^{β} outliers	0
Clashscore	1.95
Overall score	1.14

$\dagger R_{\text{meas}}$ is the redundancy-independent merging R factor (Diederichs & Karplus, 1997). \ddagger The mean $I/\sigma(I)$ falls below 2.0 in the outer shell at 1.78 \AA and below 1.0 at 1.7 \AA . \S Using ideal values from Engh & Huber (1991).

only generates MMA residues (Zurita-Lopez *et al.*, 2012). PRMT7, together with PRMT5, is involved in the methylation of histone H3 arginine 2, a marker for the activation of transcription (Migliori *et al.*, 2012), and histone H4 arginine 3, a marker for the repression of transcription of several genes implicated in DNA repair (Karkhanis *et al.*, 2012). PRMT7 also methylates eukaryotic elongation factor 2 (eEF2), a function coordinated by PRMT5 and regulated by the basic fibroblast growth factor (bFGF; Jung *et al.*, 2011). Finally, PRMT7 plays a role in male germline-imprinted gene methylation through its interaction with the epigenetic regulator CTCFL (CCCTC-binding factor-like), also known as BORIS (brother of regulator of imprinted sites; Jelinic *et al.*, 2006).

PRMT7 contains 692 amino acids in mouse (and in human) and is one of the two PRMTs that harbour two putative

AdoMet-binding motifs in tandem, possibly as a result of gene duplication (Miranda *et al.*, 2004). The PRMT catalytic module belongs to the class I type of AdoMet-dependent methyltransferases (Katz *et al.*, 2003) and harbours a set of four conserved sequence motifs. However, the second PRMT7 catalytic module is poorly conserved, including residues within the four motifs, and may not be functional (Miranda *et al.*, 2004; Krause *et al.*, 2007). The PRMTs that catalyse the formation of DMA are active as dimers, an oligomeric state illustrated by the crystal structures that have been solved (Weiss *et al.*, 2000; Zhang *et al.*, 2000; Zhang & Cheng, 2003; Troffer-Charlier *et al.*, 2007; Yue *et al.*, 2007; Cheng *et al.*, 2011; Sun *et al.*, 2011; Antonysamy *et al.*, 2012; Ho *et al.*, 2013; Wang, Zhu, Caceres *et al.*, 2014). However, PRMT7 already contains two catalytic modules in tandem, which raises questions about its structural organization and oligomerization. The correlation between the unusual architecture of PRMT7 and its MMA-formation activity also remains elusive. In the present study, we describe the crystal structure of full-length PRMT7 from *Mus musculus* solved and refined *de novo* at 1.7 \AA resolution. Structure determination has been described elsewhere (Cura *et al.*, 2014).

2. Materials and methods

2.1. Cloning, expression, purification and crystallization

The details, which illustrate the need to screen several homologues of the target protein in parallel and different expression systems to increase the chances of producing diffracting crystals, have been published in Cura *et al.* (2014). Briefly, the *M. musculus prmt7* gene was cloned into a baculovirus expression system as a GST fusion. The recombinant protein was purified to homogeneity by affinity-purification and size-exclusion chromatography. The GST tag was also removed during the process. The protein was stored at 4 $^{\circ}\text{C}$ in 20 mM Tris–HCl pH 8.0, 150 mM NaCl, 5 mM tris(2-carboxyethyl)phosphine (TCEP). The previously defined crystallization conditions (Cura *et al.*, 2014) were refined to improve the diffraction limit of the crystals. The crystals used for X-ray diffraction were obtained from an MmPRMT7 sample at 3.7 mg ml $^{-1}$ incubated for 20 min at room temperature with 1 mM AdoHcy and 100 mM NDSB-256 (the nondetergent sulfobetaine dimethylbenzylammonium propane sulfonate; Sigma). Hanging drops were set up at 290 K by mixing 1 μl protein solution and 1 μl reservoir solution consisting of 100 mM bis-tris propane (BTP) pH 7.0, 3.2% (w/v) PEG 20 000, 12.6% (w/v) PEG 8000. Single diffracting crystals grew to 100 \times 100 \times 200 μm within 10 d. They were cryoprotected by a short soak in 3.2% (w/v) PEG 20 000, 12.6% (w/v) PEG 8000, 100 mM BTP pH 7.0, 15% (v/v) PEG 400 and flash-cooled in liquid nitrogen for data collection.

2.2. Data collection, processing and structure refinement

High-resolution data sets were collected on the PROXIMA 1 beamline at synchrotron SOLEIL using a PILATUS 6M detector (Dectris) at 100 K and processed with XDS (Kabsch,

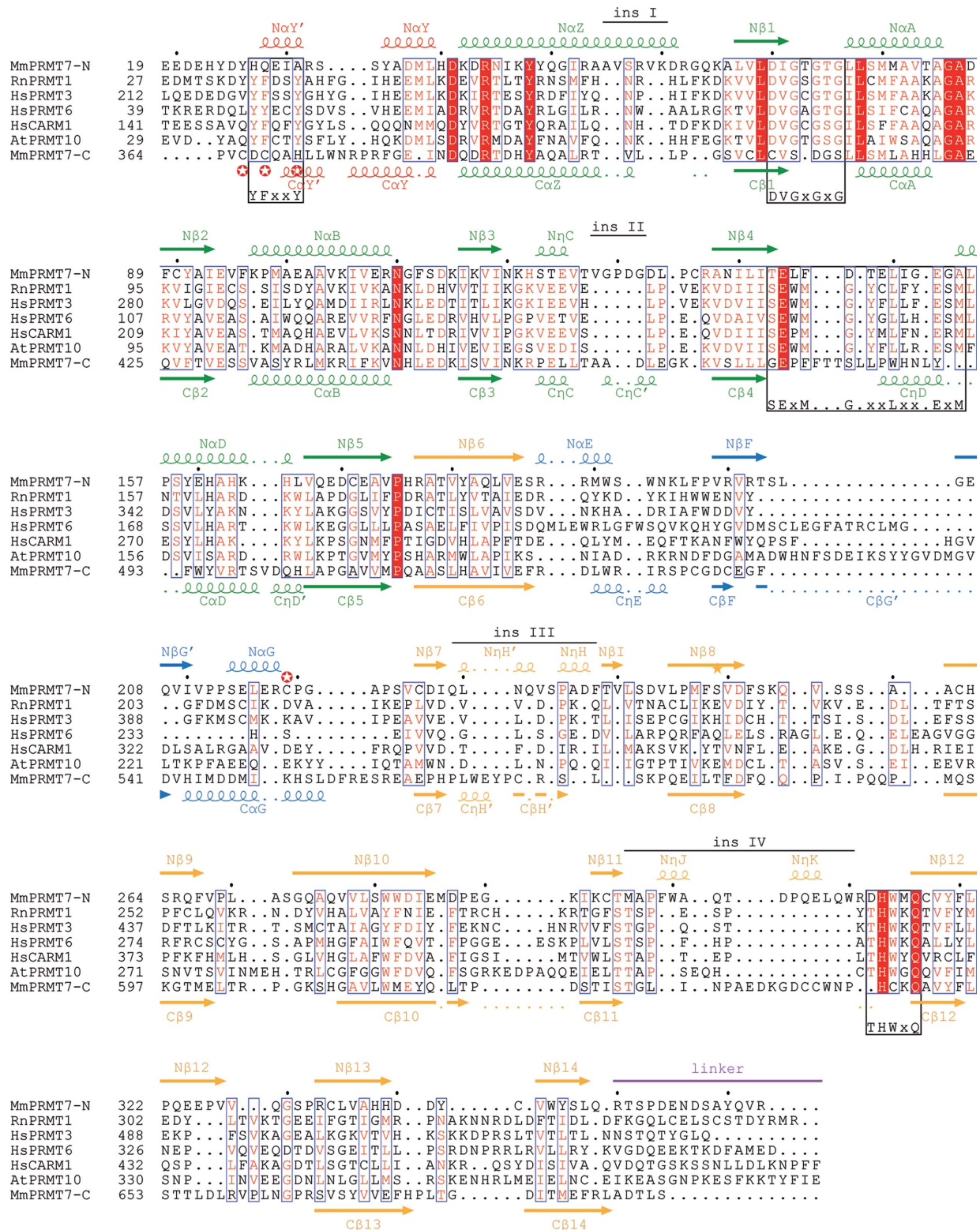


Figure 1

Structure-based sequence alignment of MmPRMT7 with other PRMTs. The structures used for the alignment are RnPRMT1 (PDB entry 1or8; Zhang & Cheng, 2003), HsPRMT3 (PDB entry 4hsg; Structural Genomics Consortium, unpublished work), HsCARM1 (PDB entry 4ikp; Structural Genomics Consortium, unpublished work), HsPRMT6 (PDB entry 4hc4; Structural Genomics Consortium, unpublished work) and AtPRMT10 (PDB entry 3r0q; Cheng *et al.*, 2011). Strictly conserved residues are shown in white on a red background. Partially conserved residues are shown in red and framed in blue. The schematic representation of the secondary-structure elements of MmPRMT7 is drawn above the alignment for the N module and below for the C module. The elements are coloured red for amino-terminal helices, green for the AdoMet-binding domain, yellow for the β -barrel domain and blue for the dimerization arm. The numbering of the elements follows the rules established for other PRMTs in the literature (Troffer-Charlier *et al.*, 2007). The four signature motifs are framed. The four MmPRMT7 residues involved in zinc binding are indicated by red circled stars. The figure was drawn with *ESPrpt* (Gouet *et al.*, 1999) and further edited.

2010) and *HKL-2000* (Otwinowski & Minor, 1997). The structure was solved by molecular replacement using the previous *de novo* model obtained by single anomalous dispersion (SAD) at 2.1 Å resolution (Cura *et al.*, 2014) and refined against a high-resolution native data set at 1.7 Å. The model was built with *Coot* (Emsley *et al.*, 2010) and refined with *PHENIX* (Adams *et al.*, 2010) and *BUSTER* (Bricogne *et al.*, 2011). The statistics are summarized in Table 1. The model geometry was checked and validation was performed using *MolProbity* (Chen *et al.*, 2010). All other crystallographic calculations were carried out with the *CCP4* package (Winn *et al.*, 2011). The model contains two small molecules present in the crystallization condition (one PEG molecule and a sulfobetaine molecule) that show ‘local ligand density fit’ highlighted by the validation server. The $2mF_o - DF_c$ weighted electron-density map contoured at 0.355 e \AA^{-3} (1.3σ) in the area of the AdoHcy molecule is shown in Supplementary Fig. S3¹.

The atomic coordinates and experimental data have been deposited in the Protein Data Bank (PDB entry 4c4a). All structural figures were prepared with *CueMol* (<http://www.cuemol.org>) or *PyMOL* (Schrödinger).

2.3. SAXS experiments and data processing

Small-angle X-ray scattering data were collected using a BioSAXS-1000 camera (Rigaku) mounted on a MicroMax-007 HF rotating-anode X-ray generator (Rigaku) at a sample-to-detector distance of 500 mm, covering the momentum-transfer range $0.008 < q < 0.65 \text{ \AA}^{-1}$ [$q = 4\pi\sin(\theta)/\lambda$, where 2θ is the scattering angle and $\lambda = 0.154 \text{ nm}$ is the X-ray wavelength], in four frames (30 min each) to check for possible radiation damage. The q scale was calibrated by silver behenate powder diffraction and all data were collected to a maximum q of 0.35 \AA^{-1} . All scattering measurements were carried out at 17°C in a thermostated quartz capillary with a diameter of 1.5 mm. MmPRMT7 samples were solubilized in 20 mM Tris pH 8.5, 150 mM NaCl, 2 mM TCEP and measured at several concentrations.

The SAXS data were processed by standard procedures using the *ATSAS* software package (Petoukhov *et al.*, 2012). The scattering intensity $I(0)$ and the radius of gyration R_g were evaluated using the Guinier approximation with *PRIMUS* (Konarev *et al.*, 2003). These parameters were also determined from the entire scattering pattern using *GNOM*, which computes the distance distribution function $P(r)$ and the maximum particle dimension D_{max} (Svergun, 1992). *CRY SOL* was used to fit the theoretical scattering pattern to the experimental data (Svergun *et al.*, 1995).

For *ab initio* model construction, multiple runs were performed using *DAMMIF* (Franke & Svergun, 2009) to verify the stability of the solution. Ten models were averaged using *DAMAVAR* (Volkov & Svergun, 2003) and superimposed on the X-ray structure using *SUPCOMB* (Kozin & Svergun, 2001).

¹ Supporting information has been deposited in the IUCr electronic archive (Reference: KW5095).

3. Results

3.1. Overall structure

The structure of the binary complex of *M. musculus* PRMT7 (MmPRMT7) with AdoHcy was crystallized in space group $P4_32_12$ and the structure was refined to 1.7 Å resolution (Table 1). Residues 1–26, 354–360, 633–636 and 655–664 were not visible in the final electron-density map. The MmPRMT7

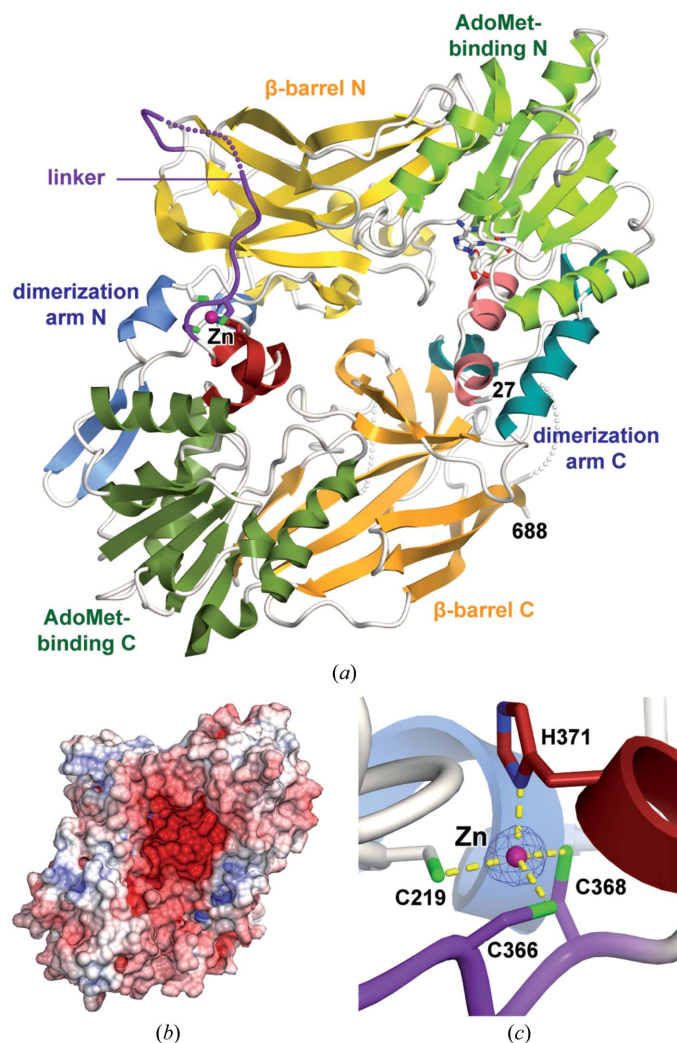


Figure 2

Overall structure and zinc-binding motif of MmPRMT7. (a) Cartoon representation of an MmPRMT7 monomer with the same colour code as in Fig. 1. MmPRMT7 is composed of two PRMT catalytic modules (N and C modules). Each module is composed of three parts: an AdoMet-binding domain, a β -barrel domain and a dimerization arm. The nomenclature used here is AdoMet-binding N, β -barrel N and dimerization arm N for the N module and AdoMet-binding C, β -barrel C and dimerization arm C for the C module. The first and last visible residues are indicated. Disordered regions are drawn as dotted lines. The AdoHcy molecule bound to the the N module is drawn as a stick model. The three cysteines and the histidine coordinating the zinc ion are drawn as ball-and-stick models. The zinc ion is shown in purple. (b) Electrostatic potential mapped on the solvent-accessible surface of MmPRMT7. The model is in the same orientation as in (a). The potential was calculated with *APBS* (Baker *et al.*, 2001). The colour ranges from blue for positively charged areas to red for negatively charged areas. (c) Close-up view of the zinc-binding motif. Residue numbers are indicated. The $2mF_o - DF_c$ weighted electron-density map was contoured at 2.46 e \AA^{-3} (9σ).

monomer is composed of two PRMT catalytic modules in tandem bridged by a 19-residue linker (residues 351–369). The general architecture is very similar to that of known PRMT dimer structures, with the two modules related by a twofold symmetry axis. Each module can be decomposed into an N-terminal AdoMet-binding domain (residues 27–174 and 370–511) and a C-terminal β -barrel domain (residues 176–350 and 513–688) connected by a conserved proline residue (Pro175 and Pro512). The β -barrel domain is interrupted by a helix wing (residues 188–219 and 527–553) called the dimerization arm that mediates most of the interactions between the two modules. In both modules the last strand of the β -barrel (β 15) is missing compared with other PRMTs (Figs. 1 and 2*a*) and the first PRMT YFxxY motif in the α Y' helix is poorly conserved (Supplementary Fig. S1). On the other hand, both MmPRMT7 modules present additional helices in the β -barrel domains that contribute to closing the hole that is present in all other PRMT dimers. The absence of a hole in the

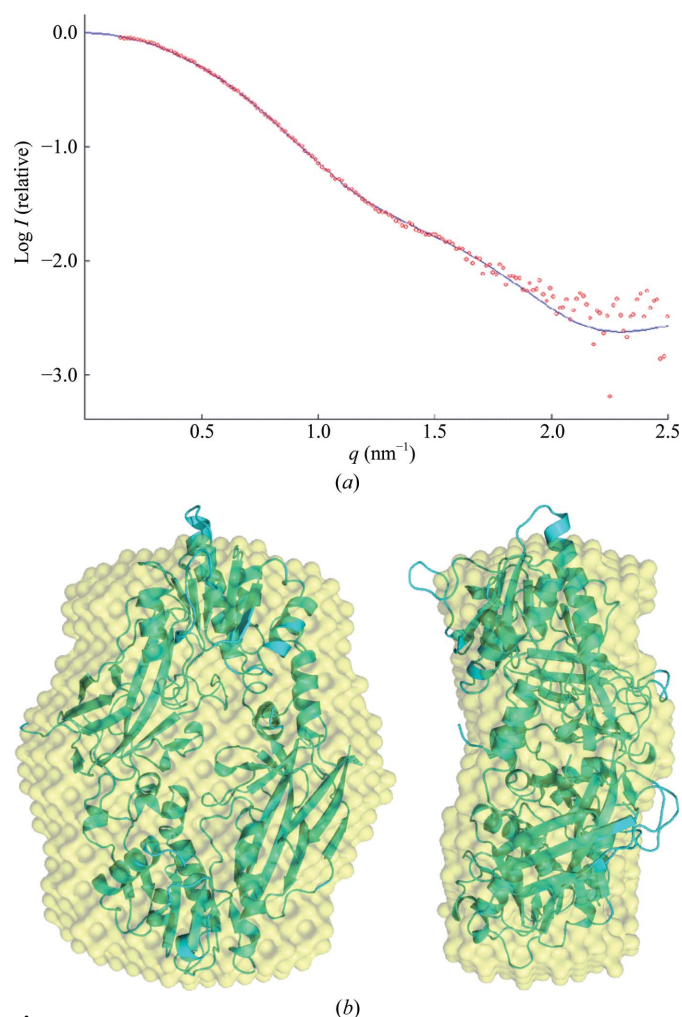


Figure 3
SAXS experiment. (a) SAXS curves of MmPRMT7. Experimental data are denoted by red circles, with the fit from the crystallographic monomeric structure shown as a blue solid line. (b) Two 90° rotated views showing the fit of the crystallographic monomeric structure (shown in ribbon representation) to the SAXS *ab initio* model (represented in yellow).

Table 2
Small-angle X-ray scattering data.

Protein concentration (mg ml ⁻¹)	1.8
$I(0)$ (relative) from $P(r)$	1.264 ± 0.005
R_g from $P(r)$ (nm)	2.80 ± 0.07
$I(0)$ (relative) from Guinier	1.278 ± 0.011
R_g from Guinier (nm)	2.92 ± 0.03
D_{max} (nm)	8.5 ± 0.5

MmPRMT7 structure compared with those of type I and type II PRMTs (Weiss *et al.*, 2000; Zhang *et al.*, 2000; Zhang & Cheng, 2003; Troffer-Charlier *et al.*, 2007; Yue *et al.*, 2007; Cheng *et al.*, 2011; Antonysamy *et al.*, 2012; Wang, Zhu, Chen *et al.*, 2014) was confirmed by a small-angle X-ray scattering (SAXS) model (Fig. 3) and therefore is not an artefact from the crystallization process. Analysis of the electrostatic potential with *APBS* (Baker *et al.*, 2001) reveals the presence of a negatively charged area in the central depression where the target peptide would bind (Fig. 2*b*).

3.2. A monomer which mimics the classical PRMT dimer

Dimer formation is known to be essential for the catalytic activity of PRMTs (Weiss *et al.*, 2000; Zhang *et al.*, 2000; Zhang & Cheng, 2003). In a few instances, higher oligomeric states have been observed (Lin *et al.*, 1996; Tang *et al.*, 1998; Rho *et al.*, 2001; Lim *et al.*, 2005; Troffer-Charlier *et al.*, 2007; Herrmann *et al.*, 2009; Antonysamy *et al.*, 2012). MmPRMT7 consists of two PRMT modules in tandem and is observed to be a monomer in solution. This assignment is based on several observations: (i) in the crystallization paper (Cura *et al.*, 2014), MmPRMT7 was shown to be a monomer by gel-filtration and DLS experiments; (ii) native mass-spectrometric analysis showed the presence of both monomeric and dimeric species, with the monomer as the main species (data not shown); (iii) SAXS analysis gave a radius and a profile in agreement with a monomeric state in solution (Fig. 3 and Table 2); (iv) MmPRMT7 has been crystallized in two different crystal forms ($P4_32_12$ and $I222$), implying different packing interactions; and (v) *PISA* analysis (Krissinel & Henrick, 2007) also predicted no stable higher oligomer states.

3.3. Comparison of the two modules in tandem

When compared with each other, the two PRMT7 modules (named N and C) are somewhat different, with a root-mean-square deviation (r.m.s.d., calculated on equivalent C α atoms) of 3.14 Å (Fig. 4 and Supplementary Table S1). These differences are not clustered in a specific domain (AdoMet-binding, β -barrel or dimerization arm) but are spread all over the module. The N module has the lowest r.m.s.d. with PRMT4/CARM1 (2.42 Å) and the C module has the lowest r.m.s.d. with PRMT5 (2.69 Å). When comparing the individual domains separately, the N and C AdoMet-binding domains have r.m.s.d. values with all type I PRMTs of around 1.6 and 2.3 Å, respectively (Fig. 4*a* and Supplementary Table S2). For the β -barrel domains, the lowest r.m.s.d. values are 2.39 Å between the N domain and CARM1 and 2.49 Å between the C domain and PRMT1 (Supplementary Table S3). Overall, the

MmPRMT7 N module is more similar to the other PRMTs than the C module. The MmPRMT7 dimerization arms display β -hairpins at their tips, with the C arm being shorter than the N arm. In the N arm the second α -helix (N α G) is disrupted by the zinc ion binding to residue Cys219 (Fig. 2c), whereas in the C arm the second helix (C α G) is properly folded. The conformations of the MmPRMT7 dimerization arms are similar to those of PRMT1 and PRMT3, but their orientations are somewhat different owing to the reduced size of their first helices (N α E and C η E; Figs. 4b and 4c). Both dimerization arms display similar orientations relative to the module as those observed for PRMT6 and PRMT3, which is consistent with the apparent twofold axis observed for the MmPRMT7 structure. In summary, the conformation and orientation of the MmPRMT7 dimerization arm result in a specific orientation of the modules relative to each other, so that the space between them is shaped like a cone with the bottom closed by insertion loops.

3.4. MmPRMT7 has no hole between the two modules in tandem

The interface between the N and C modules of MmPRMT7 buries a solvent-accessible surface area of 2760 Å². The interface can be divided into three regions: the two classical PRMT interfaces between the dimerization arm from one module and the AdoMet-binding domain from the other module and the interface that closes the hole observed in the other PRMT dimer structures. The first two interfaces are mainly composed of hydrophobic and hydrogen-bonding interactions. Between the C dimerization arm and the N AdoMet-binding domain, residues Leu527 and Ile530 of helix C η E, residue Phe540 of strand C β G' and residues Val542, Ile544, Met545, Met548, Ile549 and Leu553 of helix C α G make extensive hydrophobic interactions with residues Ile30 and Tyr35 of helix N α Y, residue Tyr49 of helix N α Z, residues Thr75, Leu77, Met80 and Met81 of helix N α A and residues Met99 and Ile106 of helix N α B (Fig. 5a). A set of 14 hydrogen bonds involving main or side chains further strengthens this interface. Between the N dimerization arm and the C AdoMet-binding domain, residues Trp193 and Leu196 in the N α E/N β F loop, residues Val199, Val201 and Thr203 of strand N β F, residue Ile210 of strand N β G' and residue Leu216 of helix N α G make extensive hydrophobic interactions with residue Trp374 in the C α Y'/C α Y loop, residues Phe379 and Ile382 of helix C α Y, residue Thr388 of helix C α Z, residues Leu413, Met416, Met417, His420 and His421 of helix C α A and residues Leu348, Ile442 and Val445 of helix C α B (Fig. 5b). A set of 16 hydrogen bonds involving main or side chains further strengthens this interface. Finally, the extra loops in MmPRMT7 compared with other PRMTs (Fig. 1) create the third interface that fills the gap between the two modules (Fig. 5c). Several hydrophobic interactions between bulky residues and two hydrogen bonds (Asn229–Asp568 and Asp312–Tyr569) form this interface. In summary, the interface between the MmPRMT7 N and C modules is a continuum that spans between the tips of the two dimerization arms. It

involves 66 and 69 residues in each catalytic module and incorporates 34 hydrogen bonds and eight salt bridges.

3.5. AdoHcy recognition in the active N module

One AdoHcy molecule is bound in a deep pocket of the N AdoMet-binding domain which is composed of the N-termini of the N α helices and the C-termini of strands N β 1, N β 2 and N β 4. The AdoHcy molecule adopts an extended conformation (Fig. 6a), as seen in other PRMTs, and can be divided into three parts. For the homocysteine moiety, residue Met38 makes van der Waals contacts with the S atom, Arg44 interacts with the carboxyl group, and the Gly72 carbonyl O atom makes a hydrogen bond to the amine group. For the ribose

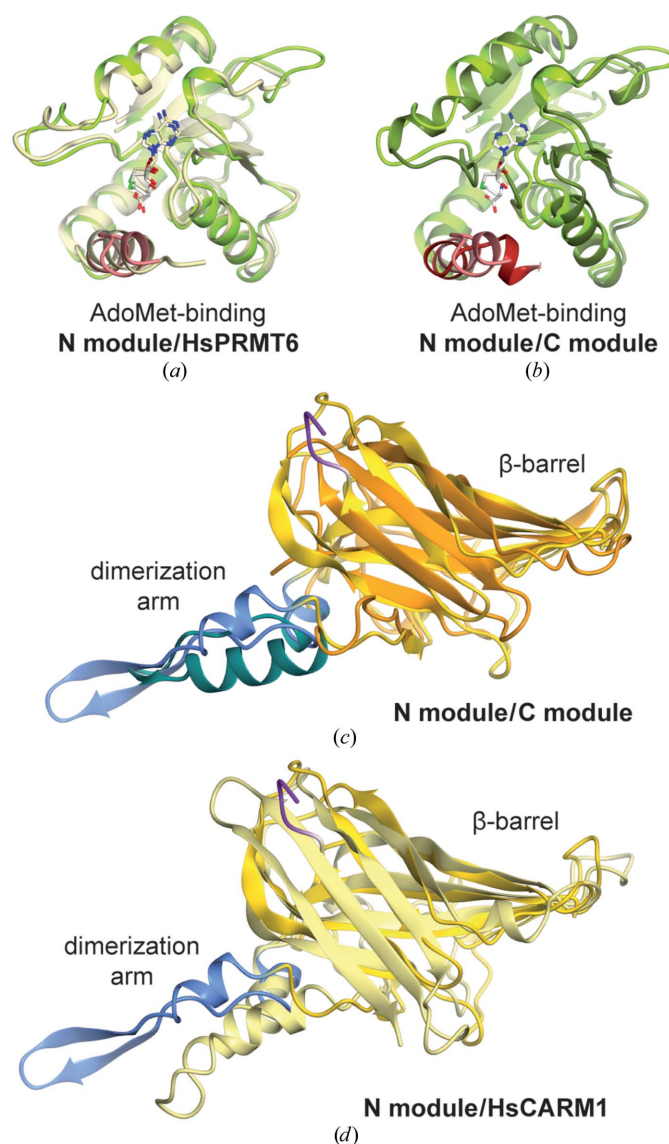


Figure 4 Comparison of individual module structures of MmPRMT7 with other PRMTs. Superposition of the AdoMet-binding domains of MmPRMT7 module N on HsPRMT6 (beige) (a) or on MmPRMT7 module C (b). Superposition of the β -barrel domains of MmPRMT7 module C on module N (c) or on HsCARM1 (d). The colour code for MmPRMT7 is the same as in Fig. 1.

moiety, the Tyr35 hydroxyl group hydrogen-bonds to the ribose 3'-hydroxyl group, and Glu94 makes hydrogen bonds to the two ribose hydroxyl groups. For the adenine moiety, the Val95 backbone amine hydrogen-bonds to N3, the Ser123 hydroxyl group and backbone amine make hydrogen bonds to N1 and N6, and the Ser158 hydroxyl group hydrogen-bonds to N1. The coordinated water molecule interacting with N7 on the adenine ring is conserved in several type I PRMT structures. In addition, the hydrophobic residues Ile71, Val95, His122 and Leu145 make van der Waals interactions with the adenine ring. Overall, the AdoMet-binding mode in MmPRMT7 is very similar to that observed in other PRMTs.

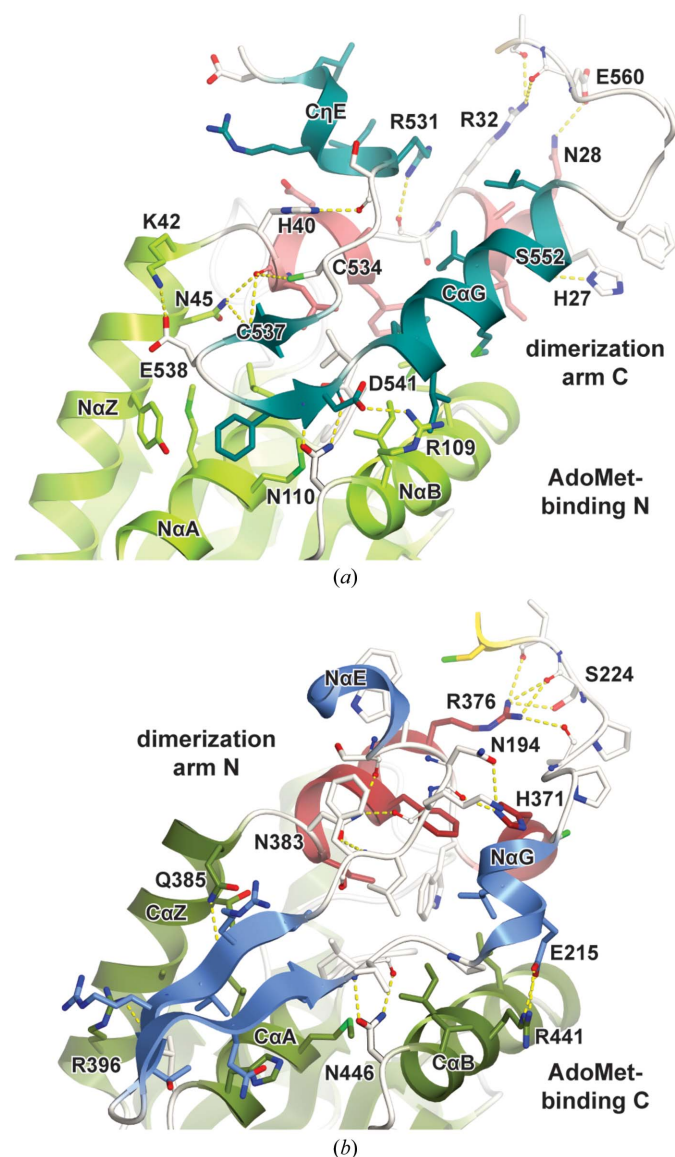


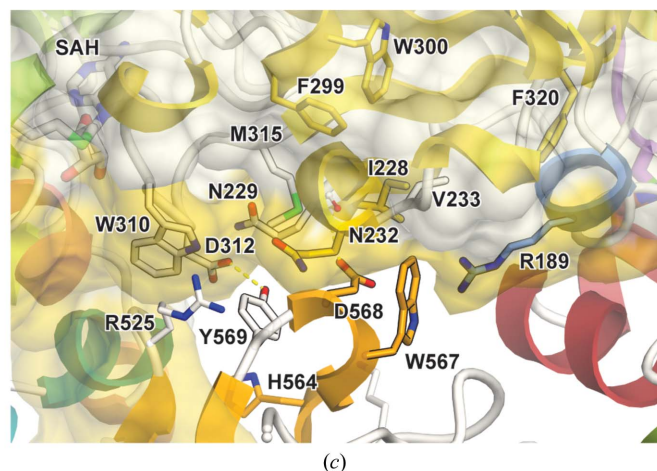
Figure 5
Interfaces between the two MmPRMT7 modules. (a) Interface between dimerization arm C and AdoMet-binding N. (b) Interface between dimerization arm N and AdoMet-binding C. Residue side chains involved in hydrogen bonds are drawn as stick models. Residue main chains involved in hydrogen bonds are drawn as ball-and-stick models. Residues mediating hydrogen bonds through their side chain are numbered. Secondary-structure elements are labelled. The conserved residues Asn110 and Asn446 are outlined in black. (c) Interface between the two β -barrel domains. Residues involved in the interface are outlined in black. Other residues that fill the hole between the two modules are also drawn and labelled. AdoHyc is indicated as SAH. The colour code is the same as in Figs. 1 and 2.

3.6. An unexpected structural zinc ion stabilizes the PRMT7 tandem-repeated modules

The electron-density map reveals that the connection between the N and C modules of MmPRMT7 is stabilized by a metal ion (Fig. 2c). It involves three cysteine residues (Cys219, Cys366 and Cys368) and one histidine residue (His371). One cysteine is located at the end of the N dimerization arm (Cys219), whereas the three other residues are close to each other and are located at the end of the linker between the two modules (366–371). The nature of the ion bound to the mouse PRMT7 monomer was assigned as zinc based on three observations: (i) the strength of the peak in the electron-density map and the observed metal–donor atom distances, (ii) its coordination by three cysteine residues and one histidine residue and (iii) the difference of 63 Da between the theoretical molecular weight and that measured by mass spectrometry in native conditions (data not shown). The residues for zinc coordination are conserved in PRMT7 sequences with tandem-repeated modules, with the exception of the plant PRMT7. The MmPRMT7 zinc finger can be classified as ‘zinc necklace subtype C-terminal’ based on its structural conformation (Krishna *et al.*, 2003; Andreini *et al.*, 2011). This class of zinc finger is found in proteins with diverse functions and has a structural role in most cases.

3.7. Collapsed and nonproductive AdoMet-binding pocket in the C module

The most important difference between the two MmPRMT7 modules is the collapse of the AdoMet-binding pocket in the C module (Fig. 6b) compared with the conserved pocket of the active N module. This collapse leads to the formation of a nonproductive AdoMet-binding pocket. Several correlated structural features are used to lock this



putative AdoMet-binding pocket into a nonproductive state: (i) the conformation and sequence composition of the two loops (motif 2 and double-E loop) involved in cofactor binding or the catalytic mechanism, (ii) the zinc finger which freezes the position of helices $C\alpha Y'/C\alpha Y$ and (iii) an arginine residue (Arg378) which occupies the canonical arginine

substrate-binding pocket. In a classical PRMT active site, motif 2 is located at the C-terminal end of strand $\beta 1$, while the double-E loop is located at the C-terminal end of strand $\beta 4$ (Fig. 1). In both loops of the MmPRMT7 C module, conserved residues involved in AdoMet binding are missing or are replaced by bulkier amino acids (Fig. 1 and 6c). The first

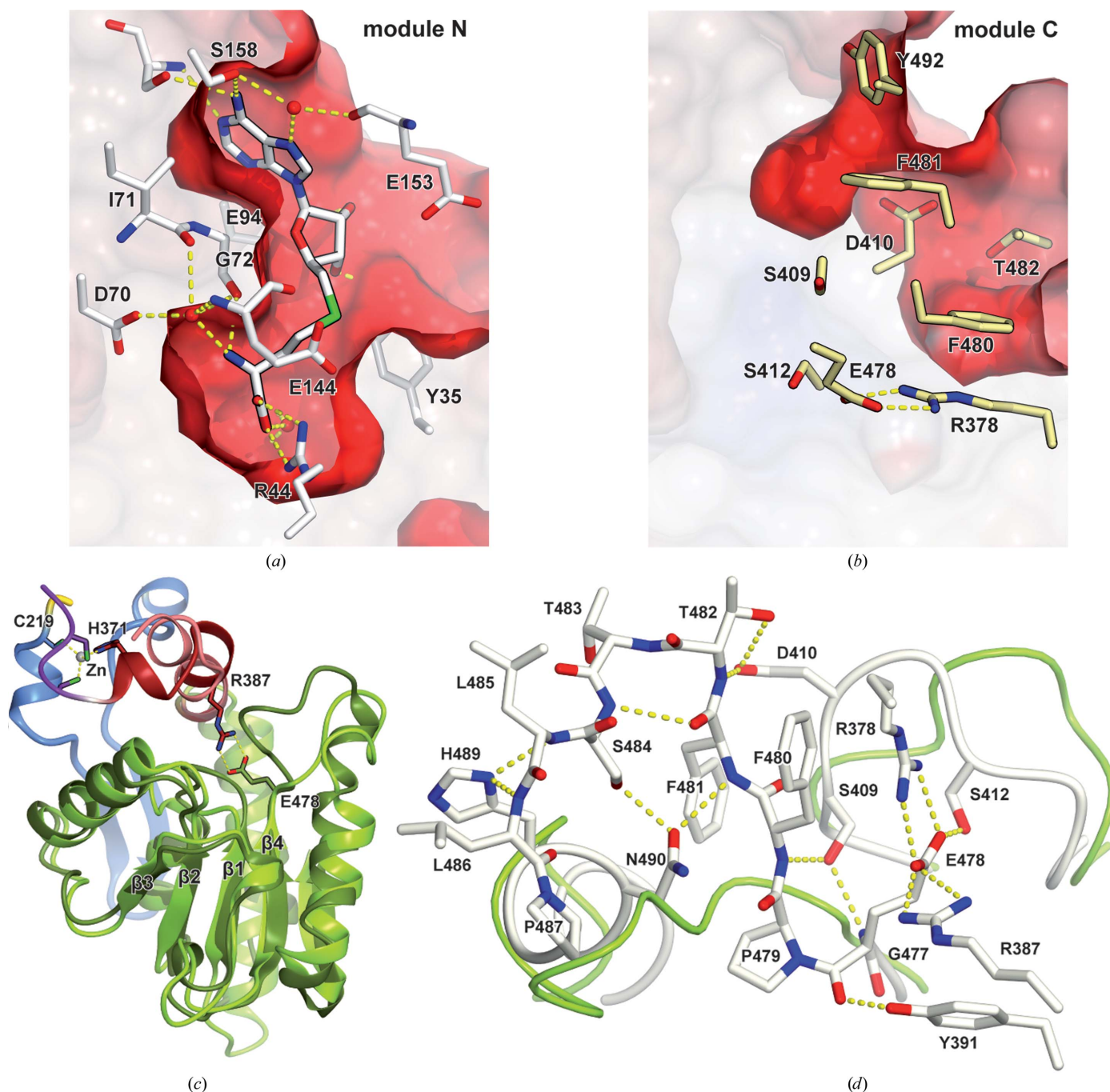


Figure 6
Collapse of the cofactor-binding pocket in MmPRMT7 domain C. (a) AdoMet-binding pocket of module N. The pocket is drawn as a molecular surface cut in half and coloured with the electrostatic potential calculated in Fig. 2(b). The bound AdoHcy molecule and the MmPRMT7 residues interacting with it are drawn as stick models. Water molecules involved in the hydrogen-bond network are drawn as red spheres. Hydrogen bonds are drawn as yellow dotted lines. (b) Collapse of the cofactor-binding pocket in MmPRMT7 module C. The residues that fill the potential pocket are drawn as stick models. (c) Superposition of the N and C AdoMet-binding domains illustrating the conformation of the loops surrounding the pocket. The strands of the β -sheet are numbered. (d) Detailed conformation and interaction network of the residues of MmPRMT7 domain C that are responsible for the collapse of the pocket. Domain N elements are coloured green and domain C elements are coloured white.

residue of the double-E loop, Glu478, is tightly maintained through four hydrogen bonds to Arg378 and Arg387 of helices $C\alpha Y$ and $C\alpha Z$. The *cis* conformation of residue Pro479 is responsible for a sharp turn of the loop towards the domain compared with the double-E loop from the N module. This conformation is maintained by a hydrogen bond to the Tyr391 hydroxyl group (Fig. 6*d*). The hydrogen bonds between the Asn490 hydroxyl group of helix $C\eta D$ and the Phe481 backbone amine and the Ser484 hydroxyl group strongly restrain the conformation of the inactive double-E loop. The loop from motif 2 also moves inwards and several hydrogen bonds between the loops completely lock the active site. Residues Ser409, Asp410 and Ser412 from motif 2 form hydrogen bonds to residues Gly477, Pro479, Thr482 and Glu478 from the double-E loop.

It has been shown that helices $\alpha Y'/\alpha Y$ show a disorder-to-order transition upon substrate binding (Troffer-Charlier *et al.*, 2007). The Zn finger (Cys219–Cys366–Cys368–His371) of MmPRMT7 is located just before helices $C\alpha Y'/C\alpha Y$ and freezes the position of these two short helices. As described above, Arg378, which is located between $C\alpha Y'$ and $C\alpha Y$, is fixed by salt-bridge interactions with the so-called double-E loop. Moreover, this arginine residue of the protein occupies the pocket which is expected to bind the arginine of the substrate to be methylated. Together, these structural features ensure a nonproductive AdoMet-binding site in the C module.

3.8. Target arginine-binding pocket: still more to learn

In an attempt to understand the structural basis of the different activities of the PRMT types, all available PRMT structures were superimposed on the conserved residues in the target arginine-binding pocket, namely the methionine in helix αY , the two glutamates of the double-E loop and the tryptophan of the THW loop. One representative structure was

chosen for each type of PRMT based on the structure resolution and the conformation of the residues.

The formation of asymmetric DMA (aDMA) by type I PRMT4/CARM1 and type I PRMT1 involves two conserved methionines (Gui *et al.*, 2014; Fig. 7*a*) and the THW loop [Met163, Met269 and His145 numbered as in mouse CARM1 (MmCARM1); Cura *et al.*, in preparation]. In type II PRMT5 (Fig. 7*b*), which produces sDMA, those residues are replaced by a phenylalanine and two serines, respectively (Phe327, Ser446 and Ser578 numbered as in HsPRMT5). A Phe-to-Met mutation at this position in PRMT5 (F379M and F327M in CePRMT5 and HsPRMT5, respectively) has also been shown to induce the formation of a mixture of aDMA and sDMA (observed for wild-type PRMT5; Sun *et al.*, 2011). A Met-to-Phe mutation at this position (M48F, corresponding to M163F in MmCARM1) enables type I PRMT1 to generate both aDMA and sDMA (Gui *et al.*, 2014). This methionine is conserved in the type III MmPRMT7 (Met38). However, from molecular modelling it seems that a second methionine (Met269 in MmCARM1) and the histidine of the THWP loop (the DHW loop in MmPRMT7) are structurally positioned to favour the production of aDMA *versus* sDMA by the type I CARM1. This methionine residue is replaced in the sequence by an alanine (Ala155) in the type III MmPRMT7 (Fig. 7*c*). PRMT7 and CARM1 differ significantly in this region in terms of sequence and local structure, therefore producing a rearrangement of the open cavity. The size and shape of Ala155 (MnPRMT7) is smaller compared with Met269 (MnCARM1) (Ala has a surface and a volume of 88.6 Å² and 115 Å³, respectively; Met has a surface and a volume of 162.9 Å² and 185 Å³, respectively). The position of the C α atom of Ala155 of PRMT7 is shifted away from the arginine methylation site by ~2.8 Å compared with the C α atom of Met269 of CARM1. Moreover, the side-chain atoms of those two residues do not point to the same region in space, providing more room in the

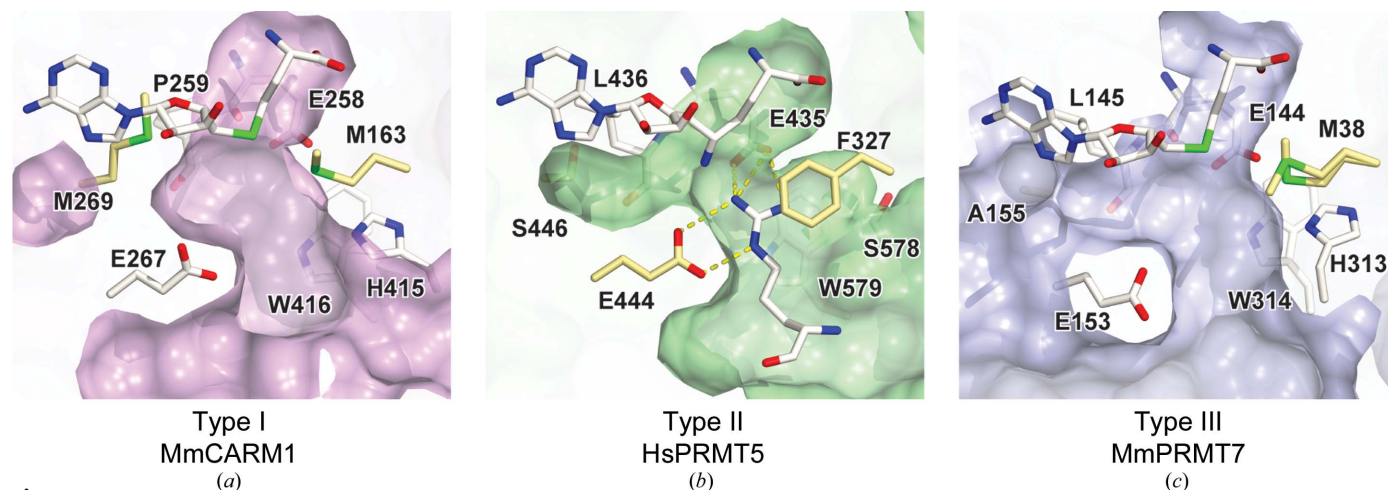


Figure 7

Target arginine-binding pocket in each PRMT type. The active site of MmPRMT7 may accommodate sDMA. Target arginine-binding pockets of (a) MmCARM1 (PDB entry 3b3f; Troffer-Charlier *et al.*, 2007), (b) HsPRMT5 (PDB entry 4gqb; Antonyamy *et al.*, 2012) and (c) MmPRMT7. Conserved residues surrounding the pockets are drawn as stick models and labelled. An M48F mutation in type I PRMT1 (corresponding to M163F in MmCARM1) produces sDMA and aDMA. In type II PRMT5, Phe327 is crucial for producing sDMA. A Phe-to-Met mutation produces both sDMA and aDMA. In type I MmCARM1, Met269 may be also important in producing mainly aDMA. This methionine is replaced by a serine (Ser446) in PRMT5. MmPRMT7 shows features similar to type I PRMT and Ala155 may produce space to accommodate aDMA.

binding site for the arginine guanidinium moiety. Such additional space would be sufficient to accommodate a methyl group, and suggests that the type III MmPRMT7 may be able to accommodate sDMA in its active site. Therefore, the presence of only one active PRMT module in MnPRMT7 cannot by itself rationalize its activity. Further biochemical and structural data are required for a better understanding of the formation of different methylarginines (MMA, aDMA and sDMA) by PRMTs.

3.9. Phylogenetic analysis

To understand the evolutionary history of PRMT7, a phylogenetic tree of PRMT7 modules was built using sequences from various organisms ranging from plants to mammals, including PRMT7 sequences from the Trypanosomatida order, which contain a single PRMT module (Supplementary Fig. S2). On the tree, each PRMT7 module is on a separate branch comprising all species. The trypanosomatid PRMT7s form a separate branch unrelated to the N or C modules. This tree suggests that PRMT7 module duplication occurred early in evolution before speciation. The fact that the trypanosomatid PRMT7s branch out and have only one domain suggests that they diverged early on and that the lack of the second module is not the result of domain loss after duplication. In evolution PRMT7 is closely related to PRMT10, which also contains two PRMT modules in tandem as a result of duplication. It is likely that the duplication event occurred before the separation of PRMT7 from PRMT10.

4. Discussion

PRMTs are usually active as homodimers, with each monomer containing a catalytic core module to which additional domains may be added (Weiss *et al.*, 2000; Zhang *et al.*, 2000; Zhang & Cheng, 2003; Troffer-Charlier *et al.*, 2007; Yue *et al.*, 2007; Cheng *et al.*, 2011; Antonysamy *et al.*, 2012; Wang, Zhu, Chen *et al.*, 2014). The catalytic core module can be divided into three parts: (i) an AdoMet-binding domain conserved in AdoMet-dependent methyltransferases, (ii) a β -barrel domain and (iii) a dimerization arm. Despite extensive research by many teams in a highly competitive field, and X-ray structures of the catalytic domains of several PRMTs, at least three challenges remain to be solved by structural biology: (i) the characterization at the molecular level of functional macromolecular complexes involving full-length PRMT targets, (ii) deciphering rules to explain how a given PRMT recognizes and binds its peptide substrate(s) and (iii) revealing how PRMTs achieve specific arginine methylation on different target sites.

Nevertheless, the classical scheme is that dimers are necessary to produce dimethylated arginine in protein substrates in a two-step process (Lee *et al.*, 1977; Gary & Clarke, 1998), with monomethylation occurring in one monomer followed by a second methylation event occurring in the second active site, possibly without release of the protein substrate. The dimer interface will be involved in the forma-

tion of a binary complex with the protein substrate. As the methylated sites of PRMT substrates are usually located on flexible or unfolded terminal extensions, small conformational changes may be required to translocate the target peptide from one active site to the other. Whether PRMT-catalyzed dimethylation results from a distributive or a processive mechanism remains controversial and requires more structural and biochemical data (Osborne *et al.*, 2007; Obianyano *et al.*, 2008; Lakowski & Frankel, 2008; Kölbl *et al.*, 2009; Rust *et al.*, 2011; Obianyano & Thompson, 2012; Antonysamy *et al.*, 2012; Gui *et al.*, 2013). This dynamic biological process involving specific protein recognition and a catalytic mechanism requiring the hydrolysis of two AdoMet molecules remains to be elucidated at the atomic level.

With few exceptions, PRMT7 contains the two classical PRMT monomers in a single polypeptide chain, with the second module lacking the sequence motifs or residues shown to be required for the methylation reaction. The 1.7 Å resolution crystal structure presented here revealed several structural features showing that the second active site of tandem-duplicated mouse PRMT7 is frozen in an inactive state by a conserved zinc finger at the junction between the two PRMT modules and by the collapse of degenerated AdoMet-binding loops. This zinc finger is conserved in tandem-duplicated PRMT7 with the exception of PRMT7 from plants.

During the last stages of the submission of this paper, two PRMT7 structures were described in papers that were in the press: PRMT7 from *Trypanosoma brucei* (TbPRMT7) and *Caenorhabditis elegans* (CePRMT7) (Wang, Zhu, Cacaes *et al.*, 2014; Hasegawa *et al.*, 2014). In the medium-resolution structure of CePRMT7, the authors observed a ‘huge and unclear’ electron density at the level of the four coordinating residues (Hasegawa *et al.*, 2014) and assigned it as covalent bonds between the cysteines and the histidine. It is more likely that a zinc ion is also tightly bound in this structure. In TbPRMT7 (Wang, Zhu, Cacaes *et al.*, 2014), since the enzyme is composed of a single PRMT module, there is no zinc-binding motif. However, the presence of two additional short helices at the N-terminal end that participate in the interface between the two monomers may rigidify the dimer in an analogous manner to the zinc-binding motif.

In the ‘classical’ PRMT structures with a hole between the two monomers, one face of the dimer usually displays negatively charged patches compatible with the fixation of the target proteins. In the case of histones, the N-terminal tail, where the modification will take place, would pass through the hole to reach the catalytic sites on each PRMT monomer (Fig. 2*b*). In the case of PRMT7, the crystal structure shows that this access is blocked by the relative orientations of the PRMT7 modules and additional loops within the β -barrel. Therefore, the target peptide-binding mode might differ substantially in PRMT7 compared with other PRMTs.

PRMT7 is the only known type III PRMT, although this assignment was the subject of debate. At first glance, the inactivity of the second PRMT7 module may give a simple explanation for its MMA-only forming activity. However, TbPRMT7, which consists of only one PRMT module and

forms homodimers, was also shown to only catalyze MMA formation (Wang, Zhu, Cacares *et al.*, 2014). Therefore, the number of active AdoMet-binding domains in PRMT7 is not correlated to the MMA or DMA formation activity.

In the biochemical studies published on mammalian PRMT7 the only monomethylated peptide tested was H4 (1–21) with a monomethylarginine at position H4R3 (Feng *et al.*, 2013). However, H4R3 does not seem to be an *in vitro* methylation target for PRMT7, which on the contrary targets H4R17 and H4R19 (Feng *et al.*, 2013). For TbPRMT7, H4R3 was shown to be an *in vitro* methylation target, whereas the monomethylated H4R3me1 peptide was not a substrate (Wang, Zhu, Cacares *et al.*, 2014). The authors hypothesize that the residue Gln329, which replaces histidine in the THW motif of type I PRMTs, prevented the binding of MMA in the pocket. However, mutation of this residue to histidine did not allow DMA formation, demonstrating that this residue by itself was insufficient to account for the enzyme activity type. Furthermore, the authors misaligned the PRMT7 residues in the THW loop. Indeed, tandem-duplicated PRMT7s such as MmPRMT7 and CePRMT7 all share a conserved histidine residue in the THW loop. Only trypanosomid PRMT7 has an MQW sequence instead. These single-module PRMT7s differ from tandem-duplicated PRMT7s at the level of their primary sequences. Although both catalyze the formation of MMA, their catalytic mechanisms might be slightly different.

A first simple hypothesis to explain the formation of only MMA by PRMT7 is that monomethylated arginine cannot be further accommodated in the active site. As described above, inspection of the MmPRMT7 target arginine-binding pocket and comparisons of the structures of type III MmPRMT7, type I PRMT4/CARM1 and type II PRMT5 does not clearly rationalize the inability to catalyze a second methylation in the active N module.

A second hypothesis may also be formulated. As dimethylation of an arginine substrate by a single monomer first requires the release of the AdoHcy product and the binding of a second AdoMet cofactor, conformational changes involved in such cofactor release and upload may require or induce release of the MMA protein substrate. If the MMA protein substrate has lower affinity than the nonmodified protein, this may explain why dimethylated substrates are not found. Unfortunately, to our knowledge the affinity of MmPRMT7 for a monomethylated peptide compared with a nonmethylated peptide is not known.

As detailed above, more structures and biological data are necessary in order to fully understand the common and idiosyncratic features controlling MMA/aDMA/sDMA formation by PRMTs and the formation of complexes with full-length protein substrates.

Finally, the biological relevance of MMA remains to be further elucidated (Strahl *et al.*, 2001; Kirmizis *et al.*, 2009). Type I and type II PRMTs are also known to generate and release monomethylated peptides as part of the dimethylation reaction. The identification of MMA residues has recently begun with the advances in mass-spectrometric techniques (Afjehi-Sadat & Garcia, 2013). Monomethylated arginine

residues have been found on several histone and nonhistone PRMT substrates (Bremang *et al.*, 2013). On several residues, both monomethylated and dimethylated arginine modifications have been identified (Gayatri & Bedford, 2014). Presently, tudor domains are the only known specific reader for methylated arginines (Chen *et al.*, 2011). Among them, only the tudor domain from protein TDRD3 (tudor domain-containing protein 3) is known to bind MMA, albeit with less affinity than DMA (Liu *et al.*, 2012).

5. Related literature

The following references are cited in the Supporting Information for this article: Felsenstein (1989), McWilliam *et al.* (2013) and Néron *et al.* (2009).

This work was supported by institutional support from CNRS, Université de Strasbourg, INSERM, Instruct, part of the European Strategy Forum on Research Infrastructures (ESFRI) supported by national member subscriptions as well as the French Infrastructure for Integrated Structural Biology (FRISBI) (ANR-10-INSB-05-01) and grants from Agence Nationale de la Recherche (ANR-08-PCVI-0037-01), Association pour la Recherche contre le Cancer (ARC; Nos. A09/4/5005 and SFI20121205902). The authors are indebted to the baculovirus common resources of the IGBMC for recombinant virus production and thank the members of the IGBMC common services and the members of the structural biology platform of IGBMC, especially Pierre Poussin-Courmontagne and Alastair G. McEwen, for technical assistance. The authors thank the members of SOLEIL PROXIMA 1 beamline, the Swiss Light Source, Paul Scherrer Institut, Villigen, Switzerland and the European Synchrotron Radiation Facility–European Molecular Biology Laboratory Joint Structural Biology groups for use of the beamline facilities and for help during data collection.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Afjehi-Sadat, L. & Garcia, B. A. (2013). *Curr. Opin. Chem. Biol.* **17**, 12–19.
- Andreini, C., Bertini, I. & Cavallaro, G. (2011). *PLoS One*, **6**, e26325.
- Antonyasamy, S. *et al.* (2012). *Proc. Natl Acad. Sci. USA*, **109**, 17960–17965.
- Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.
- Bedford, M. T. & Clarke, S. G. (2009). *Mol. Cell*, **33**, 1–13.
- Bremang, M., Cuomo, A., Agresta, A. M., Stugiewicz, M., Spadotto, V. & Bonaldi, T. (2013). *Mol. Biosyst.* **9**, 2231–2247.
- Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O. S., Vornrhein, C. & Womack, T. O. (2011). *BUSTER v2.10.0*. Cambridge: Global Phasing.
- Cha, B. & Jho, E. H. (2012). *Expert Opin. Ther. Targets*, **16**, 651–664.
- Chen, C., Nott, T. J., Jin, J. & Pawson, T. (2011). *Nature Rev. Mol. Cell Biol.* **12**, 629–642.
- Cheng, Y., Frazier, M., Lu, F., Cao, X. & Redinbo, M. R. (2011). *J. Mol. Biol.* **414**, 106–122.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.

- Cura, V., Troffer-Charlier, N., Lambert, M.-A., Bonnefond, L. & Cavarelli, J. (2014). *Acta Cryst.* **F70**, 80–86.
- Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Engh, R. A. & Huber, R. (1991). *Acta Cryst.* **A47**, 392–400.
- Felsenstein, J. (1989). *Cladistics*, **5**, 164–166.
- Feng, Y., Maity, R., Whitelegge, J. P., Hadjikyriacou, A., Li, Z., Zurita-Lopez, C., Al-Hadid, Q., Clark, A. T., Bedford, M. T., Masson, J. Y. & Clarke, S. G. (2013). *J. Biol. Chem.* **288**, 37010–37025.
- Franke, D. & Svergun, D. I. (2009). *J. Appl. Cryst.* **42**, 342–346.
- Gary, J. D. & Clarke, S. (1998). *Prog. Nucleic Acid Res. Mol. Biol.* **61**, 65–131.
- Gayatri, S. & Bedford, M. T. (2014). *Biochim. Biophys. Acta*, **1839**, 702–710.
- Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.
- Gros, L., Delaporte, C., Frey, S., Decesse, J., de Saint-Vincent, B. R., Cavarec, L., Dubart, A., Gudkov, A. V. & Jacquemin-Sablon, A. (2003). *Cancer Res.* **63**, 164–171.
- Gui, S., Gathiaka, S., Li, J., Qu, J., Acevedo, O. & Hevel, J. M. (2014). *J. Biol. Chem.* **289**, 9320–9327.
- Gui, S., Wooderchak-Donahue, W. L., Zang, T., Chen, D., Daly, M. P., Zhou, Z. S. & Hevel, J. M. (2013). *Biochemistry*, **52**, 199–209.
- Hasegawa, M., Toma-Fukai, S., Kim, J. D., Fukamizu, A. & Shimizu, T. (2014). *FEBS Lett.* **588**, 1942–1948.
- Herrmann, F., Pably, P., Eckerich, C., Bedford, M. T. & Fackelmayer, F. O. (2009). *J. Cell Sci.* **122**, 667–677.
- Ho, M.-C., Wilczek, C., Bonanno, J. B., Xing, L., Seznec, J., Matsui, T., Carter, L. G., Onikubo, T., Kumar, P. R., Chan, M. K., Brenowitz, M., Cheng, R. H., Reimer, U., Almo, S. C. & Shechter, D. (2013). *PLoS One*, **8**, e57008.
- Jelincic, P., Stehle, J. C. & Shaw, P. (2006). *PLoS Biol.* **4**, e355.
- Jung, G. A., Shin, B. S., Jang, Y. S., Sohn, J. B., Woo, S. R., Kim, J. E., Choi, G., Lee, K.-M., Min, B. H., Lee, K.-H. & Park, G. H. (2011). *Exp. Mol. Med.* **43**, 550–560.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Karkhanis, V., Wang, L., Tae, S., Hu, Y.-J., Imbalzano, A. N. & Sif, S. (2012). *J. Biol. Chem.* **287**, 29801–29814.
- Katz, J. E., Dlakić, M. & Clarke, S. (2003). *Mol. Cell. Proteomics*, **2**, 525–540.
- Kirmizis, A., Santos-Rosa, H., Penkett, C. J., Singer, M. A., Green, R. D. & Kouzarides, T. (2009). *Nature Struct. Mol. Biol.* **16**, 449–451.
- Kölbel, K., Ihling, C., Bellmann-Sickert, K., Neundorff, I., Beck-Sickinger, A. G., Sinz, A., Kühn, U. & Wahle, E. (2009). *J. Biol. Chem.* **284**, 8274–8282.
- Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 1277–1282.
- Kozin, M. B. & Svergun, D. I. (2001). *J. Appl. Cryst.* **34**, 33–41.
- Krause, C. D., Yang, Z.-H., Kim, Y.-S., Lee, J.-H., Cook, J. R. & Pestka, S. (2007). *Pharmacol. Ther.* **113**, 50–87.
- Krishna, S. S., Majumdar, I. & Grishin, N. V. (2003). *Nucleic Acids Res.* **31**, 532–550.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- Lakowski, T. M. & Frankel, A. (2008). *J. Biol. Chem.* **283**, 10015–10025.
- Lee, J.-H., Cook, J. R., Yang, Z.-H., Mirochnitchenko, O., Gunderson, S. I., Felix, A. M., Herth, N., Hoffmann, R. & Pestka, S. (2005). *J. Biol. Chem.* **280**, 3656–3664.
- Lee, H. W., Kim, S. & Paik, W. K. (1977). *Biochemistry*, **16**, 78–85.
- Lim, Y., Kwon, Y.-H., Won, N. H., Min, B.-H., Park, I.-S., Paik, W. K. & Kim, S. (2005). *Biochim. Biophys. Acta*, **1723**, 240–247.
- Lin, W.-J., Gary, J. D., Yang, M. C., Clarke, S. & Herschman, H. R. (1996). *J. Biol. Chem.* **271**, 15034–15044.
- Liu, K., Guo, Y., Liu, H., Bian, C., Lam, R., Liu, Y., Mackenzie, F., Rojas, L. A., Reinberg, D., Bedford, M. T., Xu, R.-M. & Min, J. (2012). *PLoS One*, **7**, e30375.
- McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P. & Lopez, R. (2013). *Nucleic Acids Res.* **41**, W597–W600.
- Migliori, V. *et al.* (2012). *Nature Struct. Mol. Biol.* **19**, 136–144.
- Miranda, T. B., Miranda, M., Frankel, A. & Clarke, S. (2004). *J. Biol. Chem.* **279**, 22902–22907.
- Néron, B., Ménager, H., Maufrais, C., Joly, N., Maupetit, J., Letort, S., Carrere, S., Tuffery, P. & Letondal, C. (2009). *Bioinformatics*, **25**, 3005–3011.
- Obianyo, O., Osborne, T. C. & Thompson, P. R. (2008). *Biochemistry*, **47**, 10420–10427.
- Obianyo, O. & Thompson, P. R. (2012). *J. Biol. Chem.* **287**, 6062–6071.
- Osborne, T. C., Obianyo, O., Zhang, X., Cheng, X. & Thompson, P. R. (2007). *Biochemistry*, **46**, 13370–13381.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V. & Svergun, D. I. (2012). *J. Appl. Cryst.* **45**, 342–350.
- Rho, J., Choi, S., Seong, Y. R., Cho, W.-K., Kim, S. H. & Im, D.-S. (2001). *J. Biol. Chem.* **276**, 11393–11401.
- Rust, H. L., Zurita-Lopez, C. I., Clarke, S. & Thompson, P. R. (2011). *Biochemistry*, **50**, 3332–3345.
- Strahl, B. D., Briggs, S. D., Brame, C. J., Caldwell, J. A., Koh, S. S., Ma, H., Cook, R. G., Shabanowitz, J., Hunt, D. F., Stallcup, M. R. & Allis, C. D. (2001). *Curr. Biol.* **11**, 996–1000.
- Sun, L., Wang, M., Lv, Z., Yang, N., Liu, Y., Bao, S., Gong, W. & Xu, R.-M. (2011). *Proc. Natl Acad. Sci. USA*, **108**, 20538–20543.
- Svergun, D. (1992). *J. Appl. Cryst.* **25**, 495–503.
- Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.
- Tang, J., Gary, J. D., Clarke, S. & Herschman, H. R. (1998). *J. Biol. Chem.* **273**, 16935–16945.
- Troffer-Charlier, N., Cura, V., Hassenboehler, P., Moras, D. & Cavarelli, J. (2007). *EMBO J.* **26**, 4391–4401.
- Volkov, V. V. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 860–864.
- Wang, C., Zhu, Y., Caceres, T. B., Liu, L., Peng, J., Wang, J., Chen, J., Chen, X., Zhang, Z., Zuo, X., Gong, Q., Teng, M., Hevel, J. M., Wu, J. & Shi, Y. (2014). *Structure*, **22**, 756–768.
- Wang, C., Zhu, Y., Chen, J., Li, X., Peng, J., Zou, Y., Zhang, Z., Jin, H., Yang, P., Wu, J., Niu, L., Gong, Q., Teng, M. & Shi, Y. (2014). *PLoS One*, **9**, e87267.
- Weiss, V. H., McBride, A. E., Soriano, M. A., Filman, D. J., Silver, P. A. & Hogle, J. M. (2000). *Nature Struct. Biol.* **7**, 1165–1171.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yang, Y. & Bedford, M. T. (2013). *Nature Rev. Cancer*, **13**, 37–50.
- Yue, W. W., Hassler, M., Roe, S. M., Thompson-Vale, V. & Pearl, L. H. (2007). *EMBO J.* **26**, 4402–4412.
- Zhang, X. & Cheng, X. (2003). *Structure*, **11**, 509–520.
- Zhang, X., Zhou, L. & Cheng, X. (2000). *EMBO J.* **19**, 3509–3519.
- Zurita-Lopez, C. I., Sandberg, T., Kelly, R. & Clarke, S. G. (2012). *J. Biol. Chem.* **287**, 7859–7870.