Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structures of histone methyltransferase SET7/9 in complexes with adenosylmethionine derivatives

Hideaki Niwa,[a]‡ Noriko Handa,[a]‡ Yuri Tomabechi,[a] Keiko Honda,[a] Mitsutoshi Toyama,[a] Noboru Ohsawa,[a] Mikako Shirouzu,[a] Hiroyuki Kagechika,[b] Tomoya Hirano,[b] Takashi Umehara[a] and Shigeyuki Yokoyama[a,c]*

[a]RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan, [b]Institute of Biomaterials and Bioengineering, Tokyo Medical and Dental University, 2-3-10 Kanda-Surugadai, Chiyoda-ku, Tokyo 101-0062, Japan, and [c]Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

‡ These authors contributed equally to this work.

Correspondence e-mail: yokoyama@riken.jp, yokoyama@biochem.s.u-tokyo.ac.jp

Received 9 August 2012
Accepted 31 December 2012

PDB References: SET7/9, complex with DAAM-3, 3vv0; complex with AAM-1, 3vuz

SET7/9 is a protein lysine methyltransferase that methylates histone H3 and nonhistone proteins such as p53, TAF10 and oestrogen receptor α. In previous work, novel inhibitors of SET7/9 that are amine analogues of the coenzyme *S*-(5′-adenosyl)-L-methionine (AdoMet) have been developed. Here, crystal structures of SET7/9 are reported in complexes with two AdoMet analogues, designated DAAM-3 and AAM-1, in which an *n*-hexylaminoethyl group or an *n*-hexyl group is attached to the N atom that replaces the S atom of AdoMet, respectively. In both structures, the inhibitors bind to the coenzyme-binding site and their additional alkyl chain binds in the lysine-access channel. The N atom in the azaalkyl chain of DAAM-3 is located at almost the same position as the *N*-methyl C atom of the methylated lysine side chain in the substrate–peptide complex structures and stabilizes complex formation by hydrogen bonding to the substrate-binding site residues of SET7/9. On the other hand, the alkyl chain of AAM-1, which is a weaker inhibitor than DAAM-3, binds in the lysine-access channel only through hydrophobic and van der Waals interactions. Unexpectedly, the substrate-binding site of SET7/9 complexed with AAM-1 specifically interacts with the artificial N-terminal sequence of an adjacent symmetry-related molecule, presumably stabilizing the alkyl chain of AAM-1.

1. Introduction

Post-translational modifications (PTMs) of histones, such as methylation and acetylation, play key roles in altering chromatin accessibility and regulating gene expression in eukaryotes (Strahl & Allis, 2000; Jenuwein & Allis, 2001; Taverna *et al.*, 2007). Lysine methylation, the covalent monomethylation, dimethylation or trimethylation of the ε-amino group of a substrate lysine residue, has emerged as one of the crucial PTMs for regulation of the function of not only histones but also many other nonhistone proteins (Lachner & Jenuwein, 2002; Martin & Zhang, 2005). SET7/9 is a protein lysine methyltransferase that methylates histone H3 at Lys4 (H3K4) and several other nonhistone proteins such as the tumour suppressor p53, the TATA box-binding protein (TBP) complex component TAF10 and oestrogen receptor α (ERα) (Wang *et al.*, 2001; Nishioka *et al.*, 2002; Wilson *et al.*, 2002; Chuikov *et al.*, 2004; Kouskouti *et al.*, 2004; Subramanian *et al.*, 2008). Similar to many other protein lysine methyltransferases, SET7/9 contains *S*-(5′-adenosyl)-L-methionine (AdoMet) as the cofactor of the enzyme in its catalytic SET

domain (Fig. 1). SET7/9 transfers the methyl group of the AdoMet to the ε-amino group of a specific lysine residue of the substrate protein within its substrate-binding cavity, as confirmed by the structures of SET7/9 cocrystallized with AdoMet or *S*-(5′-adenosyl)-L-homocysteine (AdoHcy), as well as with their analogue sinefungin, an antifungal analogue of AdoMet that inhibits DNA and histone methyltransferases, without a substrate (Jacobs *et al.*, 2002; Kwon *et al.*, 2003) and with a substrate peptide (Xiao *et al.*, 2003; Chuikov *et al.*, 2004; Couture, Collazo *et al.*, 2006; Subramanian *et al.*, 2008; Estève *et al.*, 2011). Proton dissociation from the targeted lysine residue proximal to the AdoMet of SET7/9 (*i.e.* from Lys-NH_3^+ with AdoMet to Lys-NH_2 with AdoMet + H^+) is assumed to occur prior to methylation by AdoMet. A water channel is considered to be specifically formed in the presence of AdoMet to allow the proton to escape to the solvent (Zhang & Bruice, 2007), which determines whether subsequent lysine methylation can occur. The water molecules in the active site function as placeholders that align the lysine ε-amino group for methyl transfer with AdoMet (Del Rizzo *et al.*, 2010).

Aberrant epigenetic modification by the methylation of histones or nonhistone proteins is often linked with cancerous cells and/or tissues and is now regarded as one of the major markers of carcinogenesis. For example, the EZH2 subunit of the PRC2 complex, which methylates histone H3 at Lys27 (H3K27), is overexpressed or mutated in many different types of cancer cells (Chase & Cross, 2011). In addition, SET7/9 methylates the tumour-suppressor protein p53 at Lys372, influencing the regulation of p53-mediated gene expression (Chuikov *et al.*, 2004). Furthermore, SET7/9 methylates ERα at Lys302, which stabilizes ERα and is necessary for the efficient recruitment of ERα to its target genes (Subramanian *et al.*, 2008). Since ERα is overexpressed in approximately two-thirds of breast cancer cases and the breast cancer-associated mutation K303R causes increased methylation at Lys302 (Subramanian *et al.*, 2008), small-molecule compounds that inhibit histone methyltransferases such as SET7/9 are considered to be potential anticancer agents.

Several histone methyltransferase inhibitors have been developed to date (Jones, 2012): chaetocin and BIX-01294 have been identified as selective inhibitors of SUV39 and G9a, respectively (Kubicek *et al.*, 2007; Greiner *et al.*, 2005), and DZNep, one of the most potent AdoHcy hydrolase inhibitors, effectively depletes the cellular levels of PRC2 components and inhibits histone H3K27 methylation (Tan *et al.*, 2007; Glazer *et al.*, 1986). Additionally, several derivatives of the methylation-reaction coenzymes have also been reported to be inhibitors: AdoHcy, one of the methylation-reaction products, is a nonselective inhibitor of many methyltransferases (Yao *et al.*, 2011; Richon *et al.*, 2011), the nonreactive AdoMet analogues sinefungin (Fig. 1) and AzaAdoMet inhibit AdoMet-dependent methyltransferases such as DNA methyltransferase (Reich & Mashhoon, 1990; Thompson *et al.*, 1999), and AdoMet-analogous selective inhibitors of DOT1L, a histone H3K79 methyltransferase, have also been described (Yao *et al.*, 2011). However, the development of more specific and potent inhibitors for the respective methyltransferases is still needed because the known AdoMet-analogous inhibitors such as AdoHcy and sinefungin are promiscuous methyltransferase inhibitors. In addition, only a few such inhibitors are known and their structural information is quite limited. This is especially true for the methyltransferase SET7/9, for which selective inhibitors have not been reported.

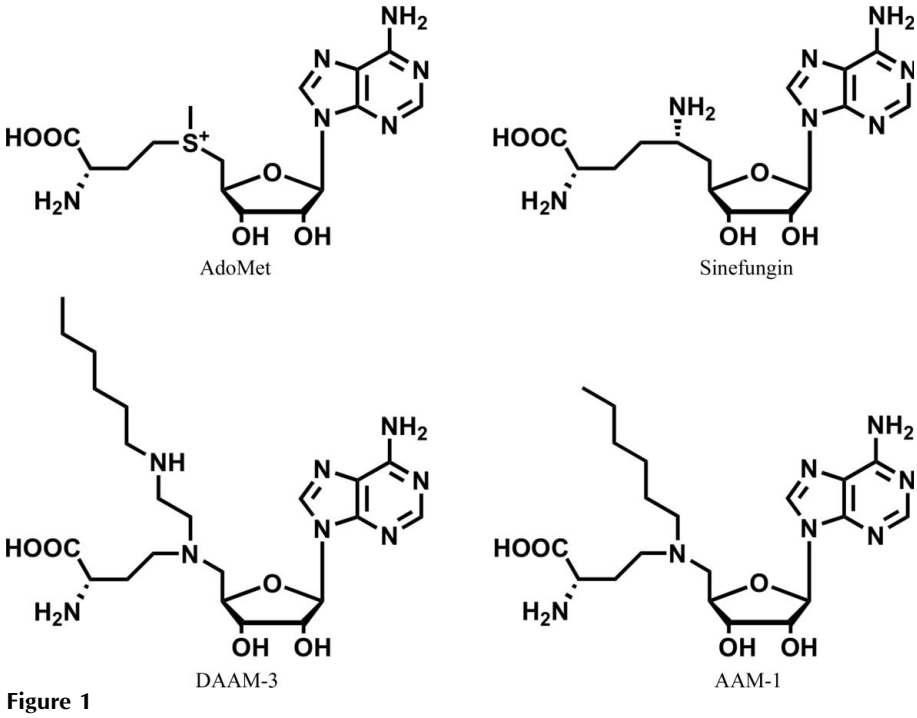


Figure 1
Structures of AdoMet and AdoMet analogues. DAAM-3 and AAM-1 have an *n*-hexylaminoethyl group and an *n*-hexyl group, respectively, attached to the N atom that replaces the S atom of AdoMet.

In a previous study, we reported the design and synthesis of novel amine analogues of AdoMet bearing various alkylamino groups as inhibitors of SET7/9 (Mori *et al.*, 2010). Among the several AdoMet analogues that we developed, two are designated DAAM-3 (DiAzaAdoMet-3; compound 1c in the previous report) and AAM-1 (AzaAdoMet-1; compound 1e in the previous report). DAAM-3 bears an azaalkyl group and AAM-1 bears an alkyl group at the reactive methyl group of AdoMet (see Fig. 1; Mori *et al.*, 2010). As indicated in the original report, DAAM-3 showed approximately 50% inhibition against SET7/9 at 10 μM in an ELISA assay, whereas AAM-1 showed almost no inhibition at the same concentration.

In this study, we report crystal structures of SET7/9 in complexes with the two AdoMet analogues DAAM-3 and AAM-1. The present crystallographic

analyses provide the structural bases for the efficient inhibition of SET7/9 by DAAM-3 and the stabilization of AAM-1, a weaker inhibitor than DAAM-3, within SET7/9 in the crystal.

2. Materials and methods

2.1. Protein expression and purification

The SET domain of human SET7/9 (residues 111–366) with an N-terminal histidine-affinity tag and a tobacco etch virus (TEV) protease-cleavage site was synthesized using the *Escherichia coli* cell-free system by the dialysis method (Kigawa *et al.*, 2004). The protein was purified by chromatography on a HisTrap column (GE Healthcare Biosciences) and subjected to TEV protease digestion. The cleaved histidine-affinity tag was removed by a second passage through the HisTrap column. The SET domain containing an N-terminal artificial Gly-Ser-Ser-Gly-Ser-Ser-Gly sequence was further purified by chromatography on a Mono Q column (GE Healthcare Biosciences) and a Superdex 200 column (GE Healthcare Biosciences) and was concentrated to 19 mg ml^{-1} in 20 m*M* Tris–HCl buffer pH 7.5 containing 0.2 *M* NaCl and 5 m*M* DTT.

2.2. Crystallization, data collection and structure determination

Crystals of SET7/9 complexed with either DAAM-3 or AAM-1 were obtained by cocrystallization under similar conditions to those used for the SET7/9–AdoMet complex (Kwon *et al.*, 2003). Briefly, the protein was incubated with the coenzyme analogues overnight before crystallization. Crystals were grown by the hanging-drop method at 277 K with a reservoir solution consisting of 0.1 *M* Tris–HCl pH 8.3–8.5, 28–30% PEG 6000. Crystals were cryoprotected in well solution containing 20–25% glycerol and were flash-cooled in liquid nitrogen. Data for the DAAM-3 complex and the AAM-1 complex were collected on BL32XU at SPring-8, Harima, Japan and on AR-NW12A at PF-AR, Photon Factory, KEK, Tsukuba, Japan, respectively. All data were processed using the *HKL*-2000 program suite (Otwinowski & Minor, 1997). The data-processing statistics are summarized in Table 1. Molecular replacement was performed with *Phaser* (McCoy *et al.*, 2007) using the coordinates of the SET7/9–AdoMet complex (PDB entry 1n6a; Kwon *et al.*, 2003) as the search model. Model building was accomplished with *Coot* (Emsley & Cowtan, 2004) and refinement was performed in *PHENIX* (Adams *et al.*, 2010). The models in the figures were depicted using *PyMOL* (http://www.pymol.org).

3. Results

3.1. Overall structures

The crystal structures of the SET domain of SET7/9 complexed with the AdoMet analogues DAAM-3 and AAM-1 (Fig. 1) were determined at 2.0 and 2.5 Å resolution, respectively. Both crystals were obtained by cocrystallization and belonged to space group $P2_12_12$. Crystallographic statistics are summarized in Table 1.

Table 1
Data-collection, phasing and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data set	DAAM-3 complex	AAM-1 complex
Data collection and processing		
Temperature (K)	100	100
Space group	$P2_12_12$	$P2_12_12$
Unit-cell parameters		
a (Å)	101.84	101.81
b (Å)	39.13	39.22
c (Å)	67.32	67.05
No. of crystals	1	1
Wavelength (Å)	1.000	1.000
Resolution range (Å)	50.0–2.00 (2.07–2.00)	50.0–2.50 (2.59–2.50)
Unique reflections	18855	9883
Measured reflections	131115	69938
Multiplicity	7.0	7.1
Completeness (%)	100.0 (100.0)	99.9 (100.0)
R_{merge}† (%)	8.1 (71.8)	13.5 (62.5)
$\langle I/\sigma(I)\rangle$	25.7 (2.8)	16.0 (3.2)
Model refinement		
Resolution range (Å)	28–2.00	28–2.50
No. of reflections	18802	9781
No. of protein atoms	1848	1921
No. of water molecules	119	69
No. of compound atoms	35	32
Mean *B* factors (Å)		
Protein	46.79	56.18
Solvent	46.81	47.08
Compound	60.32	63.84
R_{work}/R_{free}‡ (%)	18.6/23.2	18.3/24.7
Stereochemistry		
R.m.s.d. for bond lengths (Å)	0.008	0.007
R.m.s.d. for bond angles (°)	1.084	1.111
Residues in the Ramachandran plot		
Favoured region (%)	94.4	96.3
Allowed regions (%)	5.6	3.7
Outliers (%)	0	0
PDB entry	3vv0	3vuz

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where *hkl* indicates unique reflection indices and *i* indicates symmetry-equivalent indices. ‡ $R_{work} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$ for all reflections; R_{free} was calculated using randomly selected reflections (6%).

The SET-domain structures in the DAAM-3 (Fig. 2*a*) and AAM-1 (Fig. 2*b*) complexes are almost identical, with a C^α root-mean-square deviation of 0.28 Å, and are also similar to that in the SET7/9–H3–AdoHcy ternary complex (Fig. 2*c*; PDB entry 1o9s; Xiao *et al.*, 2003), with C^α root-mean-square deviations of 0.49 and 0.44 Å, respectively. A superimposition of these three backbone structures is shown in Fig. 2(*d*). The binding of DAAM-3, AAM-1 and AdoHcy in the channel for lysine methylation is depicted in Figs. 2(*e*), 2(*f*) and 2(*g*), respectively.

The bound AdoMet analogues DAAM-3 and AAM-1 were unambiguously identified in the electron-density maps (Figs. 3*a* and 3*b*). Both AdoMet analogues bound to the coenzyme-binding site, and the binding modes of the adenosine and the α-amino-acid moieties were essentially identical to those of the reacted coenzyme AdoHcy (Figs. 2*a*–2*d*). The AdoMet analogues have an *n*-hexylaminoethyl group (DAAM-3) or an *n*-hexyl group (AAM-1) attached to the N atom that replaces

the S atom of AdoMet (Mori *et al.*, 2010; see Fig. 1). Both functional groups were inserted into the narrow target lysine-access channel (Figs. 2*e* and 2*f*).

In the AAM-1 complex, the N-terminal tag-cleaved artificial sequence Gly-Ser-Ser-Gly-Ser-Ser-Gly which belongs to the adjacent symmetry-related molecule in the crystal bound to the substrate peptide-binding site on the opposite surface of SET7/9 (Fig. 2). On the other hand, no electron density was observed at the substrate peptide-binding site in the DAAM-3 complex. The same N-terminal artificial sequence seems to be flexible in the DAAM-3 complex, as judged from the lack of electron density.

3.2. DAAM-3 binding mode

DAAM-3 is a much stronger SET7/9 inhibitor than AAM-1 (Mori *et al.*, 2010). Its azaalkyl group is significantly associated with the channel through possible direct hydrogen-bonding interactions between the N atom and the side chain of Asn265 or

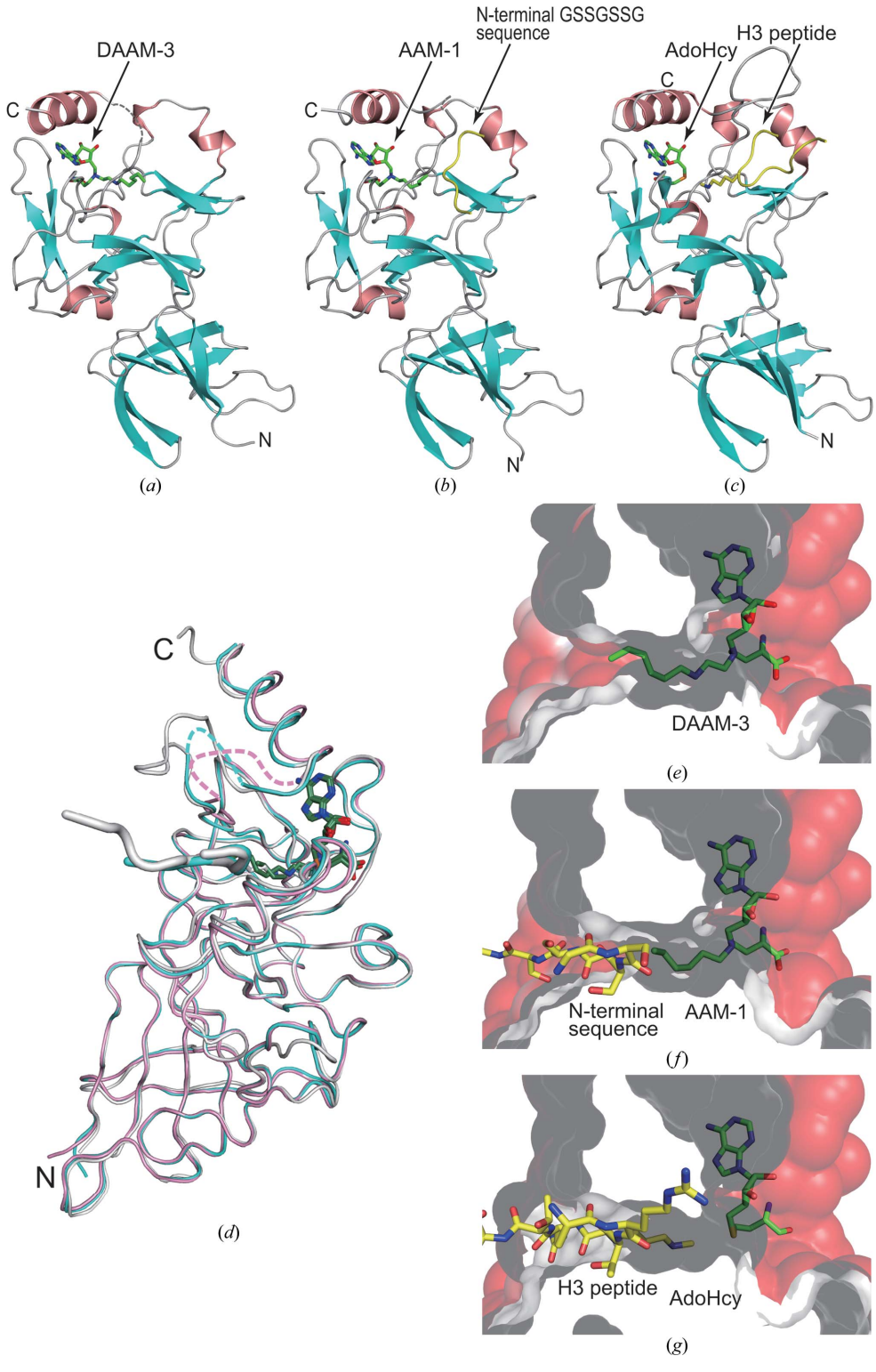
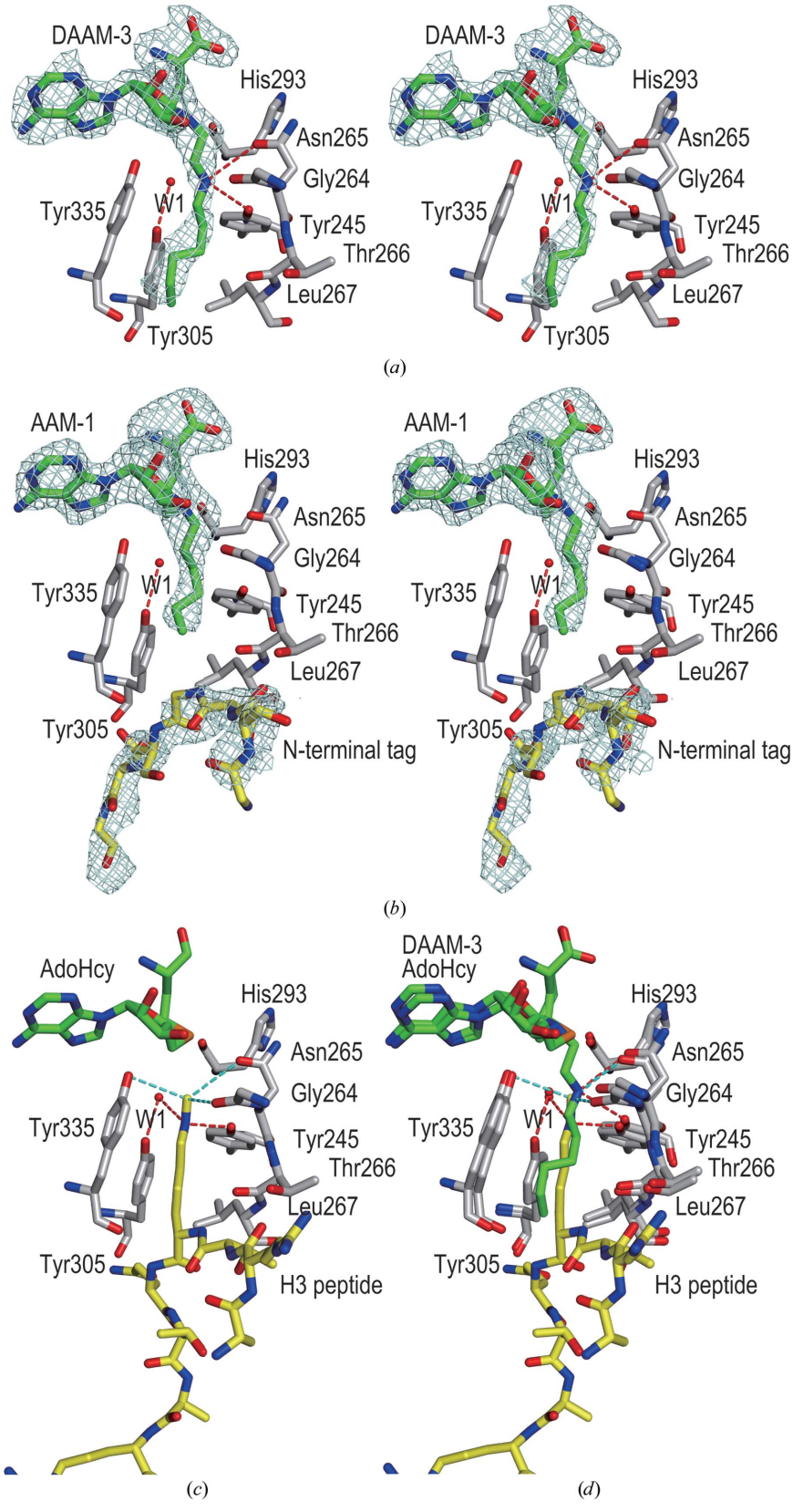


Figure 2
Compound recognition by SET7/9. Ribbon diagrams showing SET7/9 complexed with (*a*) DAAM-3, (*b*) AAM-1 and (*c*) AdoHcy–H3 (grey; PDB entry 1o9s; Xiao *et al.*, 2003). Helices are coloured salmon, β-strands cyan and random coils grey. The bound N-terminal artificial Gly-Ser-Ser-Gly-Ser-Ser-Gly sequence in (*b*) and the H3 peptide in (*c*) are coloured yellow. The disordered regions are represented by dotted lines. The bound AdoMet analogues in (*a*) and (*b*) and AdoHcy in (*c*) are represented by stick models, with C, O, N and S atoms coloured green, red, blue and orange, respectively. (*d*) Diagram showing the superimposition of the SET7/9 complexes with DAAM-3 (pink), AAM-1 (cyan) and AdoHcy–H3 (grey; PDB entry 1o9s; Xiao *et al.*, 2003). Disordered regions are represented by dotted lines. Bound peptides are represented by thick tubes (the N-terminal tag-cleaved artificial sequence Gly-Ser-Ser-Gly-Ser-Ser-Gly is cyan and H3 is grey). The compounds are represented by stick models, with C, O, N and S atoms coloured green, red, blue and orange, respectively. (*e*) A semi-transparent surface representation of SET7/9 complexed with DAAM-3, showing the coenzyme-binding site, the lysine channel and the substrate peptide-binding site. The electronegative regions are red. DAAM-3 is represented by a stick model. (*f*) SET7/9 complexed with AAM-1. AAM-1 and the N-terminal tag-cleaved artificial sequence Gly-Ser-Ser-Gly-Ser-Ser-Gly are represented by stick models with green and yellow C atoms, respectively. (*g*) SET7/9 complexed with AdoHcy–H3 (PDB entry 1o9s; Xiao *et al.*, 2003). AdoHcy and H3 are represented by stick models with green and yellow C atoms, respectively.

Tyr245. Both hydrogen-bond distances are 3.4 Å (Fig. 3*a*). The *n*-hexyl group of DAAM-3 binds within the target lysine-access channel through hydrophobic and van der Waals interactions (Fig. 3*a*).

Superimposition of the structure of the DAAM-3 complex onto that of the histone H3–AdoHcy ternary complex (Fig. 3*c*; PDB entry 1o9s; Xiao *et al.*, 2003) revealed that the N atom in the DAAM-3 azaalkyl group is located at almost the same position as the *N*-methyl C atom of the methylated lysine (Fig. 3*d*). The most distant C atom from the substituted N atom in the DAAM-3 azaalkyl group binds at almost the same position as the C^{β} atom of the target lysine (Fig. 3*d*).



3.3. AAM-1 binding mode

The AAM-1 alkyl group lacking the N atom bound to the target lysine-access channel through hydrophobic and van der Waals interactions (Fig. 3*b*). The main-chain and side-chain conformations of the channel-lining amino acids and the location of the tightly bound water molecule (W1) which functions as a placeholder (Del Rizzo *et al.*, 2010) are almost identical to those in the DAAM-3 complex and the H3–AdoHcy ternary complex (Xiao *et al.*, 2003; see Fig. 3*d*). The most distal C atom in the AAM-1 alkyl group binds at almost the same position as the C^{δ} atom of the target lysine in the channel (Fig. 3*d*).

Figure 3
Binding modes of the AdoMet analogues and the bound sequences to SET7/9. (*a*) DAAM-3 binding by SET7/9 (stereoview). The simulated-annealing OMIT $\alpha_{\rm calc}(|F_{\rm o}| - |F_{\rm c}|)$ map was calculated without DAAM-3 and is contoured at 2.5σ. The bound DAAM-3 and the interacting residues are depicted by stick models with green and grey C atoms, respectively. O, N and S atoms are coloured red, blue and orange, respectively. The tightly bound catalytically important water molecule (W1) is represented by a red sphere. Possible direct hydrogen bonds are indicated by dashed red lines. (*b*) Binding of AAM-1 and the N-terminal artificial Gly-Ser-Ser-Gly-Ser-Ser-Gly sequence by SET7/9 (stereoview). The Gly-Ser-Ser-Gly-Ser-Ser-Gly sequence binds to the substrate peptide-binding cleft of the adjacent symmetry-related SET7/9 molecule. The simulated-annealing OMIT $\alpha_{\rm calc}(|F_{\rm o}| - |F_{\rm c}|)$ map was calculated without AAM-1 and the Gly-Ser-Ser-Gly-Ser-Ser-Gly sequence and is contoured at 2.5σ. The bound AAM-1 and the Gly-Ser-Ser-Gly-Ser-Ser-Gly sequence are depicted by stick models with green and yellow C atoms, respectively. The colouring of the other atoms is the same as in (*a*). (*c*) AdoHcy and H3 peptide binding by SET7/9 (PDB entry 1o9s; Xiao *et al.*, 2003). The bound AdoHcy and H3 peptide are depicted by stick models with green and yellow C atoms, respectively. The *N*-methyl C atom of the methylated lysine may form weak nonclassical hydrogen bonds, indicated by dashed cyan lines. The colouring of the other atoms is the same as in (*b*). (*d*) Superimposition of (*a*) and (*c*).

3.4. Peptide-binding mode in the AAM-1 complex

In the crystal of the AAM-1 complex, the N-terminal artificial Gly-Ser-Ser-Gly-Ser-Ser-Gly sequence interacts with the substrate peptide-binding cleft of the adjacent symmetry-related SET7/9 protein molecule (Figs. 2*b* and 2*f*). The second Gly (Gly4; numbering from the N-terminus) perfectly matches the main-chain binding site of the target lysine (Fig. 4). The Gly4 residue also contacts the AAM-1 alkyl group (Fig. 3*b*); the distance between the C^{α} atom of Gly4 and the most distal C atom in the alkyl group of AAM-1 is 3.5 Å.

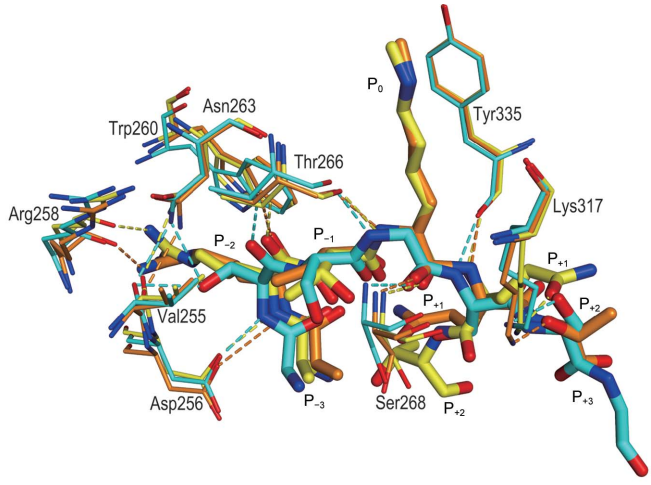
The consensus recognition sequence for SET7/9-mediated lysine methylation is (R/K)-(S/T/A)-K (positions P_{-2} to P_0, where K is the methylation target lysine) and the binding mode of this short motif is structurally conserved (Chuikov *et al.*, 2004; Couture, Collazo *et al.*, 2006). In the AAM-1 complex, the main chain of the Ser2-Ser3-Gly4 sequence interacts with the substrate peptide-binding site in an almost identical manner to the consensus recognition motif in the peptide complexes (Fig. 4). Although the P_{-2} position of the consensus motif is either Arg or Lys, Ser2 is located at the corresponding position in the AAM-1 complex. The side chain of Ser2 forms hydrogen bonds to the side chain of Asn263 and the main-chain carbonyl of Val255 (Fig. 4). In the reported crystal structures of the ternary complexes (Chuikov *et al.*, 2004; Couture, Collazo *et al.*, 2006; Estève *et al.*, 2011; Xiao *et al.*, 2003; Subramanian *et al.*, 2008), the P_{+1} residues adopt two alternative conformations: a β-strand geometry and a sharp left-handed helical turn. In the AAM-1 complex, Ser5, which corresponds to the P_{+1} residue, adopts a β-strand geometry (Fig. 4). The side chain of Ser6 forms a hydrogen bond to the side chain of Lys317, thus stabilizing the β-strand geometry (Fig. 4).

4. Discussion

The crystal structures of SET7/9 in complex with AdoHcy and sinefungin have previously been reported. The crystal structure of SET7/9 complexed with DAAM-3 presented here is the third structure of SET7/9 complexed with an AdoMet analogue inhibitor to be reported. The structure of SET7/9 complexed with DAAM-3 reveals the similarity between the interaction modes of the DAAM-3 azaalkyl group and those of the methylated lysine. The position of the N atom in the azaalkyl group is almost identical to that of the *N*-methyl C atom of the methylated lysine within the lysine-access channel (Fig. 3*d*). In the H3–AdoHcy ternary complex the *N*-methyl C atom forms weak nonclassical hydrogen bonds to the side chains of Asn265 and Tyr335 and the main-chain carbonyl of Gly264, with distances between the *N*-methyl C atom and the O atoms of 3.6, 3.4 and 3.3 Å, respectively, in one of the two complexes in the asymmetric unit. These distances are 3.9, 3.1 and 3.9 Å, respectively, in the other complex (Fig. 3*c*). The C—H···O hydrogen bonding in the SET-domain active site plays an important role in lysine methylation (Couture, Hauk *et al.*, 2006). In the H3–AdoHcy ternary complex the lysine side-chain N atom additionally forms hydrogen bonds to the side chain of Tyr245 and the tightly bound water molecule W1 (Fig. 3*c*).

In our previous study, we examined the inhibitory activities of AdoMet analogues in which the various alkyl moieties attached to the ethylamine N atom were methyl, *n*-propyl, phenethyl and *n*-hexyl groups (Mori *et al.*, 2010). All of these compounds showed inhibitory activity toward SET7/9. Among them, DAAM-3, which contains the longest alkyl group, *n*-hexyl, was slightly more potent than the others (Mori *et al.*, 2010). Therefore, the hydrophobic and van der Waals interactions of the azaalkyl group in DAAM-3 seem to contribute modestly to its SET7/9 inhibitory activity and the hydrogen-bonding interactions of the N atom in the azaalkyl group are considered to be more important. The current structure suggests that the N atom in the azaalkyl group forms possible hydrogen bonds that are analogous to those formed by either the *N*-methylated C atom (with Asn265) or the lysine side-chain N atom (with Tyr245) in the H3–AdoHcy ternary complex. A higher resolution structure will be necessary for more precise evaluation of the hydrogen bonding.

The alkyl chain of AAM-1 is shorter than that of DAAM-3 by one N and two C atoms. Only hydrophobic and van der Waals interactions were observed between SET7/9 and the *n*-hexyl chain of AAM-1. Our previous biochemical results (Mori *et al.*, 2010) were consistent with our structure showing fewer hydrogen bonds and van der Waals interactions with



	P_{-3}	P_{-2}	P_{-1}	P_0	P_{+1}	P_{+2}	P_{+3}
N-terminal tag	Gly1	Ser2	Ser3	Gly4	Ser5	Ser6	Gly7
H3 peptide	Ala1	Arg2	Thr3	Lys4	Gln5	Thr6	(Ala7)

Figure 4
Close-up view of the peptide-binding cleft interactions. The structures of the AAM-1 complex and the two H3 peptide complexes in the asymmetric unit (chains *K* and *L* in PDB entry 1o9s) are superimposed. The N-terminal Gly-Ser-Ser-Gly-Ser-Ser-Gly sequence (β-strand geometry), the H3 peptide adopting the β-strand geometry (chain *L*) and the H3 peptide adopting the sharp left-handed helical turn (chain *K*) are depicted by thick stick models, with the C atoms coloured cyan, orange and yellow, respectively. O, N and S atoms are coloured red, blue and orange, respectively. The interacting residues are depicted by stick models and the colours of the atoms are the same as those of the peptides in each complex. Hydrogen bonds are indicated by dashed lines with the same colours as those of the C atoms in the complex.

AAM-1 compared with DAAM-3. Despite the low inhibitory activity, the AAM-1 complex was crystallized under conditions similar to those of the DAAM-3 and AdoMet complexes. In the crystal of the AAM-1 complex the N-terminal artificial Gly-Ser-Ser-Gly-Ser-Ser-Gly sequence unexpectedly bound to the substrate peptide-binding cleft and contacted the alkyl chain of AAM-1 in the lysine-access channel from the opposite side of the coenzyme-binding site, in which the adenosine and α-amino-acid moieties of AAM-1 are bound. This contact may play an important role in stabilizing the interaction of AAM-1 with SET7/9 in the crystal.

The main-chain binding mode of the Ser2-Ser3-Gly4 sequence in the AAM-1 complex is almost identical to that of the consensus recognition sequence (R/K)-(S/T/A)-K in the peptide complex. Although the P_{-2} residue of the consensus recognition sequence is a long positively charged amino acid, Arg or Lys, the corresponding residue in the AAM-1 complex is a small neutral amino acid, Ser2. In the complex, the side chain of Ser2 forms a hydrogen bond to the N^{δ} atom of Asn263 and the main-chain carbonyl of Val255, and these atoms also form a hydrogen bond between them (Fig. 4). The hydrogen-bonding interactions observed at Ser2 suggest that Ser at the P_{-2} position is accommodated well in the substrate-binding cleft of SET7/9. However, the residue preference at the P_{-2} position of SET7/9 is Lys and, to a lesser extent, Arg, according to a previous biochemical study using a comprehensive H3 mutant peptide array (Dhayalan *et al.*, 2011). Taken together, these results imply that the hydrogen bonding from N^{ζ} of Lys at P_{-2} (*e.g.* p53 and TAF10) or N^{η} of Arg at P_{-2} (*e.g.* histone H3 and ERα) to the main-chain carbonyl group of Arg258 in SET7/9, which is absent in the case of Ser2 at the P_{-2} position (this study), is critically important for SET7/9 to select its substrate peptide.

SET7/9 exhibits plasticity in recognizing various P_{+1} residues, which adopt two alternative conformations: a β-strand geometry and a sharp left-handed helical turn (Couture, Collazo *et al.*, 2006). In the crystal of SET7/9 complexed with histone H3, the two complexes in the asymmetric unit each adopt different conformations at the P_{+1} residue Gln5 (Fig. 4; Xiao *et al.*, 2003). In the complex in which Gln5 adopts the β-strand geometry, the side chain of Thr6 at the P_{+2} position forms a hydrogen bond to the side chain of Lys317 (Fig. 4; Xiao *et al.*, 2003). The interaction mode and the conformations of Ser5 and Ser6 in the AAM-1 complex are similar to those of the P_{+1} and P_{+2} residues in this H3 complex with β-strand geometry (Fig. 4). On the other hand, while the P_{+1} residue in the DNMT1 complex is also Ser (Ser143), it adopts a sharp left-handed helical turn conformation and forms a hydrogen bond to Lys317 (Estève *et al.*, 2011). Since the bound sequence in the AAM-1 complex is connected to the N-terminus of SET7/9, the steric constraints may impose only the β-strand geometry on its Ser5. In some cases of peptide binding to SET7/9, the determinant of the P_{+1} residue conformation may exist in the sequence following the P_{+2} position.

The crystal structures presented here provide information about the binding of both AdoMet-analogue inhibitors and peptides by the SET domain of SET7/9. Although small-molecule compounds that inhibit SET7/9 are potentially useful for the regulation of cancers such as breast cancer, no selective SET7/9 inhibitors have been reported to date. The present crystal structures will facilitate the further design of specific and potent inhibitors for histone methyltransferases, especially SET7/9.

We thank the BL26 and BL32XU beamline staff at SPring-8 and the AR-NW12A beamline staff at the Photon Factory, KEK for their help in data collection, and Ms T. Nakayama and Ms A. Ishii for their clerical assistance. This study was supported in part by the RIKEN Program for Drug Discovery and Medical Technology Platform (DMP), by the Targeted Proteins Research Program (TPRP) of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and by the Japan Society for the Promotion of Science (JSPS) through the 'Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program)' initiated by the Council for Science and Technology Policy (CSTP).

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Chase, A. & Cross, N. C. (2011). *Clin. Cancer Res.* **17**, 2613–2618.
Chuikov, S., Kurash, J. K., Wilson, J. R., Xiao, B., Justin, N., Ivanov, G. S., McKinney, K., Tempst, P., Prives, C., Gamblin, S. J., Barlev, N. A. & Reinberg, D. (2004). *Nature (London)*, **432**, 353–360.
Couture, J.-F., Collazo, E., Hauk, G. & Trievel, R. C. (2006). *Nature Struct. Mol. Biol.* **13**, 140–146.
Couture, J.-F., Hauk, G., Thompson, M. J., Blackburn, G. M. & Trievel, R. C. (2006). *J. Biol. Chem.* **281**, 19280–19287.
Del Rizzo, P. A., Couture, J.-F., Dirk, L. M. A., Strunk, B. S., Roiko, M. S., Brunzelle, J. S., Houtz, R. L. & Trievel, R. C. (2010). *J. Biol. Chem.* **285**, 31849–31858.
Dhayalan, A., Kudithipudi, S., Rathert, P. & Jeltsch, A. (2011). *Chem. Biol.* **18**, 111–120.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Estève, P. O., Chang, Y., Samaranayake, M., Upadhyay, A. K., Horton, J. R., Feehery, G. R., Cheng, X. & Pradhan, S. (2011). *Nature Struct. Mol. Biol.* **18**, 42–48.
Glazer, R. I., Hartman, K. D., Knode, M. C., Richard, M. M., Chiang, P. K., Tseng, C. K. H. & Marquez, V. E. (1986). *Biochem. Biophys. Res. Commun.* **135**, 688–694.
Greiner, D., Bonaldi, T., Eskeland, R., Roemer, E. & Imhof, A. (2005). *Nature Chem. Biol.* **1**, 143–145.
Jacobs, S. A., Harp, J. M., Devarakonda, S., Kim, Y., Rastinejad, F. & Khorasanizadeh, S. (2002). *Nature Struct. Biol.* **9**, 833–838.
Jenuwein, T. & Allis, C. D. (2001). *Science*, **293**, 1074–1080.
Jones, P. (2012). *Med. Chem. Commun.* **3**, 135–161.
Kigawa, T., Yabuki, T., Matsuda, N., Matsuda, T., Nakajima, R., Tanaka, A. & Yokoyama, S. (2004). *J. Struct. Funct. Genomics*, **5**, 63–68.
Kouskouti, A., Scheer, E., Staub, A., Tora, L. & Talianidis, I. (2004). *Mol. Cell*, **14**, 175–182.
Kubicek, S., O'Sullivan, R. J., August, E. M., Hickey, E. R., Zhang, Q., Teodoro, M. L., Rea, S., Mechtler, K., Kowalski, J. A., Homon, C. A., Kelly, T. A. & Jenuwein, T. (2007). *Mol. Cell*, **25**, 473–481.
Kwon, T., Chang, J. H., Kwak, E., Lee, C. W., Joachimiak, A., Kim, Y. C., Lee, J. & Cho, Y. (2003). *EMBO J.* **22**, 292–303.
Lachner, M. & Jenuwein, T. (2002). *Curr. Opin. Cell Biol.* **14**, 286–298.
Martin, C. & Zhang, Y. (2005). *Nature Rev. Mol. Cell Biol.* **6**, 838–849.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Mori, S., Iwase, K., Iwanami, N., Tanaka, Y., Kagechika, H. & Hirano, T. (2010). *Bioorg. Med. Chem.* **18**, 8158–8166.

Nishioka, K., Chuikov, S., Sarma, K., Erdjument-Bromage, H., Allis, C. D., Tempst, P. & Reinberg, D. (2002). *Genes Dev.* **16**, 479–489.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Reich, N. O. & Mashhoon, N. (1990). *J. Biol. Chem.* **265**, 8966–8970.
Richon, V. M., Johnston, D., Sneeringer, C. J., Jin, L., Majer, C. R., Elliston, K., Jerva, L. F., Scott, M. P. & Copeland, R. A. (2011). *Chem. Biol. Drug Des.* **78**, 199–210.
Strahl, B. D. & Allis, C. D. (2000). *Nature (London)*, **403**, 41–45.
Subramanian, K., Jia, D., Kapoor-Vazirani, P., Powell, D. R., Collins, R. E., Sharma, D., Peng, J., Cheng, X. & Vertino, P. M. (2008). *Mol. Cell*, **30**, 336–347.
Tan, J., Yang, X., Zhuang, L., Jiang, X., Chen, W., Lee, P. L., Karuturi, R. K., Tan, P. B., Liu, E. T. & Yu, Q. (2007). *Genes Dev.* **21**, 1050–1063.
Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. & Patel, D. J. (2007). *Nature Struct. Mol. Biol.* **14**, 1025–1040.
Thompson, M. J., Mekhalfia, A., Hornby, D. P. & Blackburn, G. M. (1999). *J. Org. Chem.* **64**, 7467–7473.
Wang, H., Cao, R., Xia, L., Erdjument-Bromage, H., Borchers, C., Tempst, P. & Zhang, Y. (2001). *Mol. Cell*, **8**, 1207–1217.
Wilson, J. R., Jing, C., Walker, P. A., Martin, S. R., Howell, S. A., Blackburn, G. M., Gamblin, S. J. & Xiao, B. (2002). *Cell*, **111**, 105–115.
Xiao, B., Jing, C., Wilson, J. R., Walker, P. A., Vasisht, N., Kelly, G., Howell, S., Taylor, I. A., Blackburn, G. M. & Gamblin, S. J. (2003). *Nature (London)*, **421**, 652–656.
Yao, Y., Chen, P., Diao, J., Cheng, G., Deng, L., Anglin, J. L., Prasad, B. V. & Song, Y. (2011). *J. Am. Chem. Soc.* **133**, 16746–16749.
Zhang, X. & Bruice, T. C. (2007). *Biochemistry*, **46**, 14838–14844.