Organic & Biomolecular Chemistry

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/c4ob01734j

Synthesis and evaluation of protein arginine *N*-methyltransferase inhibitors designed to simultaneously occupy both substrate binding sites†

Matthijs van Haren, Linda Quarles van Ufford, Ed E. Moret and Nathaniel I. Martin*

Received 13th August 2014,
Accepted 23rd October 2014

DOI: 10.1039/c4ob01734j

www.rsc.org/obc

The protein arginine *N*-methyltransferases (PRMTs) are a family of enzymes that function by specifically transferring a methyl group from the cofactor *S*-adenosyl-L-methionine (AdoMet) to the guanidine group of arginine residues in target proteins. The most notable is the PRMT-mediated methylation of arginine residues that are present in histone proteins which can lead to chromatin remodelling and influence gene transcription. A growing body of evidence now implicates dysregulated PRMT activity in a number of diseases including various forms of cancer. The development of PRMT inhibitors may therefore hold potential as a means of developing new therapeutics. We here report the synthesis and evaluation of a series of small molecule PRMT inhibitors designed to simultaneously occupy the binding sites of both the guanidino substrate and AdoMet cofactor. Potent inhibition and surprising selectivity were observed when testing these compounds against a panel of methyltransferases.

Introduction

The methylation of arginine side chains is an abundant post-translational modification encountered across a range of organisms from bacteria and fungi to plants and mammals. Approximately 2% of the arginine residues in total protein extracts from rat liver nuclei are methylated[1,2] and similar levels of arginine methylation appear to be present in the human proteome.[3,4] Arginine methylation is performed by a dedicated family of enzymes known as the protein arginine *N*-methyltransferases (PRMTs). PRMTs employ a bisubstrate mechanism using the methyl-group donor *S*-adenosyl methionine (AdoMet) to methylate substrate proteins/peptides with the concomitant formation of the by-product *S*-adenosyl homocysteine (AdoHcy) (Fig. 1). The first methylation step catalyzed by PRMTs yields monomethyl-arginine (MMA), an intermediate that is typically further methylated to form asymmetric dimethyl-arginine (aDMA) or symmetric dimethyl-arginine (sDMA).[5]

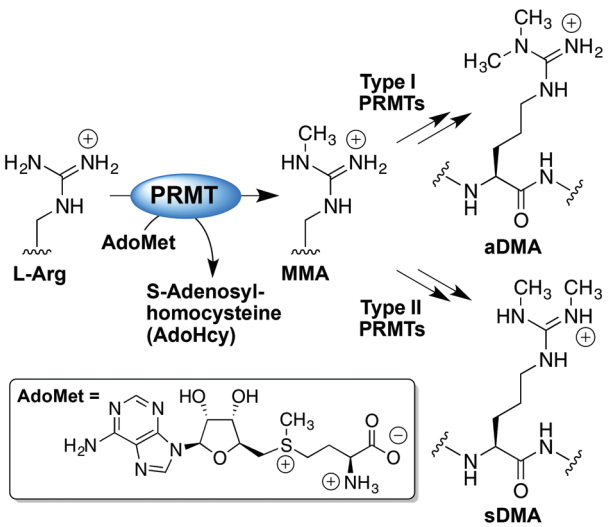


Fig. 1 The PRMT catalyzed methylation of arginine residues to generate monomethyl arginine (MMA), asymmetric dimethyl arginine (aDMA), and symmetric dimethyl arginine (sDMA).

PRMT-mediated arginine methylation is intimately involved in a range of cellular functions, including transcriptional regulation, RNA processing, signal transduction and DNA repair.[5,6] While dysregulated arginine methylation is implicated in a number of pathogenic conditions including cardiovascular disease and inflammation, it is in the areas of epigenetics and

Medicinal Chemistry & Chemical Biology Group, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Universiteitsweg 99, 3584 CG Utrecht, The Netherlands. E-mail: n.i.martin@uu.nl

† Electronic supplementary information (ESI) available: ^{1}H and ^{13}C NMR spectra for all new compounds and analytical RP-HPLC traces for the final compounds **1–6**, and dose–response curves used for IC_{50} determination for **1–6**. See DOI: 10.1039/c4ob01734j

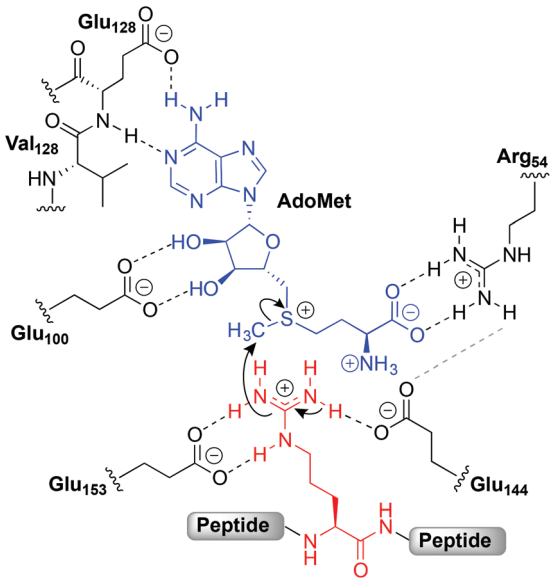


Fig. 2 Proposed PRMT "S_N2-like" bisubstrate mechanism with key active site residues shown (PRMT1 numbering scheme). AdoMet is presented in blue and the arginine residue of a target peptide is presented in red.

oncology that the PRMTs have received most attention. The aberrant modification of histones is now an established contributing factor in cancer development and there is mounting evidence that suggests PRMT activity is important in cancer progression.[7–9] Many PRMTs are overexpressed in breast cancer[10] and PRMT4 (also known as CARM1 for coactivator-associated arginine methyltransferase 1) is associated with prostate and colorectal cancers.[11–13] In addition, PRMT5 levels are enhanced in various human lymphoid cancer cells, including transformed chronic lymphocytic leukemia cell lines.[14] Of particular note are recent studies in which genetic knockdowns of either *PRMT1* or *PRMT6* were shown to suppress the growth of several cancer cell lines.[9] These results suggest that the inhibition of PRMT activity may be a viable approach to developing novel anticancer therapeutics.[7,15]

Among the eleven known human PRMTs, the vast majority of arginine methylation (>85%) is attributable to PRMT1 which exhibits a wide substrate specificity, preferentially methylating arginine residues flanked by one or more glycines.[5,16] By comparison, the other PRMTs have more narrowly defined substrate tolerances.[5,16] While the PRMTs vary significantly in size and sequence, a conserved core of amino acids comprising the catalytic site is strictly maintained.[5,16] The published crystal structures of PRMT1, 3, 4, and 5 exhibit virtually identical folds and also point to a specific set of active site residues as being key for catalysis (Fig. 2).[17–21] Of particular note is the "double-E loop" motif common to all PRMTs, containing the Glu_{144} and Glu_{153} residues (PRMT1 numbering), known to be critical for substrate recognition.[22] As illustrated in Fig. 2, the PRMTs operate *via* a bisubstrate mechanism wherein the guanidine group of an arginine-containing peptide substrate is precisely orientated (*via* a hydrogen-bond network with Glu_{144} and Glu_{153}) in proximity to the electrophilic methylsulfonium group of AdoMet leading to an "S_N2-like" substitution reaction.

With these conserved active site features in mind we set out to design a series of novel inhibitors aimed at specifically exploiting both the adenosine binding domain and "double-E loop" unique to the PRMTs (compounds **1–6**, Fig. 3A). Compounds **1–6** all contain the adenosine group of the AdoMet cofactor which is further connected to a guanidine moiety *via* a variable linker. This design strategy was envisioned to impart selective binding towards PRMTs relative to other AdoMet-dependent methyltransferases. Specifically it was expected that important interactions with key residues conserved in the PRMT active site, such as Glu_{144} and Glu_{153}, would be maintained (Fig. 2). Of note in the structures of the proposed inhibitors is the omission of the amino acid moiety that is present in the AdoMet cofactor. This omission was made based upon the findings of Thompson and coworkers who

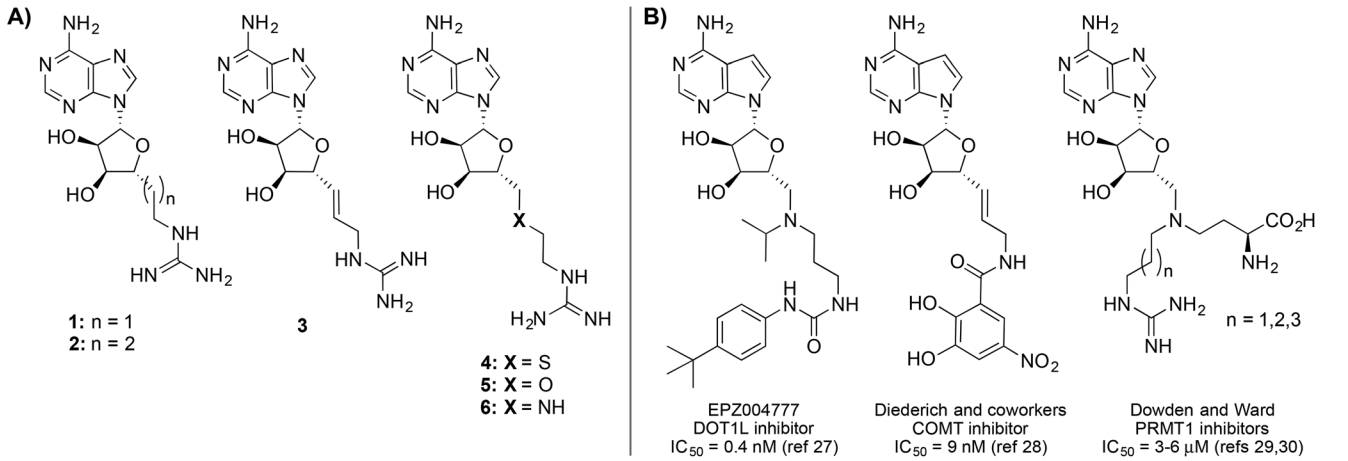


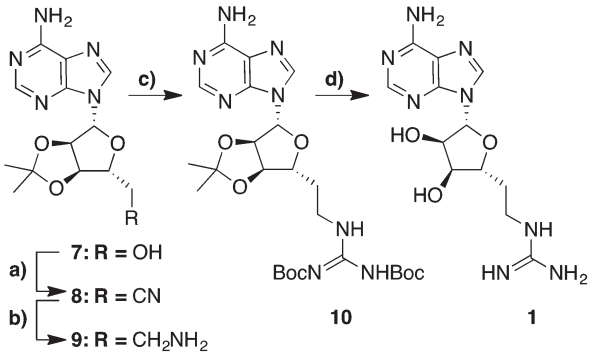
Fig. 3 (A) Bisubstrate PRMT inhibitors **1–6** prepared in the present study and (B) previously described methyltransferase inhibitors containing the adenosine moiety.

showed that mutation of the active site Arg_{54} residue that interacts with the AdoMet carboxyl group has very little impact on PRMT catalysis.[22] Our group has previously reported the preparation of various peptide-based PRMT inhibitors bearing modified arginine residues in which the guanidine moiety is substituted in such a way as to convert a PRMT substrate peptide into an inhibitor.[23–26] In the present study we shift our focus to the preparation of small molecule inhibitors designed to simultaneously occupy the binding sites of both the guanidine unit of the peptide substrate and the adenosine moiety of the AdoMet cofactor.

Compounds that contain the adenosine moiety have also been previously developed as effective inhibitors against other methyltransferases (Fig. 3B). Most notable are the recently disclosed inhibitors developed by Epizyme against DOT1L, a lysine methyltransferase implicated in mixed-lineage leukemia. One of Epizyme's early lead compounds, designated as EPZ004777, is an extremely potent (IC_{50} = 0.4 nM) and selective DOT1L inhibitor capable of killing leukemia cells *in vitro* and shows significant extension of survival in a mouse leukemia model.[27] In another example, the Diederich group successfully developed potent bisubstrate inhibitors of catechol-*O*-methyltransferase (COMT), an enzyme that has been proposed as a target in the development of anti-Parkinson's disease treatments.[28] In addition, within the PRMT field itself, the groups of Dowden and Ward recently described a series of bisubstrate inhibitors with IC_{50} values in the low micromolar range.[29,30] While somewhat similar to the PRMT inhibitors we here describe, the compounds prepared by Dowden and Ward differ notably in the length of the spacer connecting the adenosine and guanidine units. When considering the PRMT transition state model illustrated in Fig. 2, a spacer consisting of three atoms appears to most closely approximate the positioning of the adenosine and guanidine groups relative to each other. In the compound series that Dowden and Ward evaluated, the spacer length used to link the adenosine and guanidine units ranged from five to seven atoms. In contrast, when designing inhibitors **1–6** we were specifically interested in examining the impact of shorter spacers consisting of two to four atoms with the aim of more accurately mimicking the PRMT transition state geometry. As shown in Fig. 3A inhibitors **2** and **3** both contain three-carbon spacers while the spacer in compound **1** is one atom shorter. Also prepared were inhibitors **4–6** which contain four atom spacers each including a different heteroatom. The activities of all 6 inhibitors were then evaluated against a panel of three PRMTs including PMRT1, 4 and 6 as well as G9a, a well-characterized lysine methyltransferase.

Results and discussion

The synthetic routes used in the preparation of PRMT inhibitors **1–6** all started from the commercially available 2′,3′-*O*-isopropylideneadenosine building block **7**. The synthesis of **1** (Scheme 1) began with the one carbon homologation of **7** *via* a

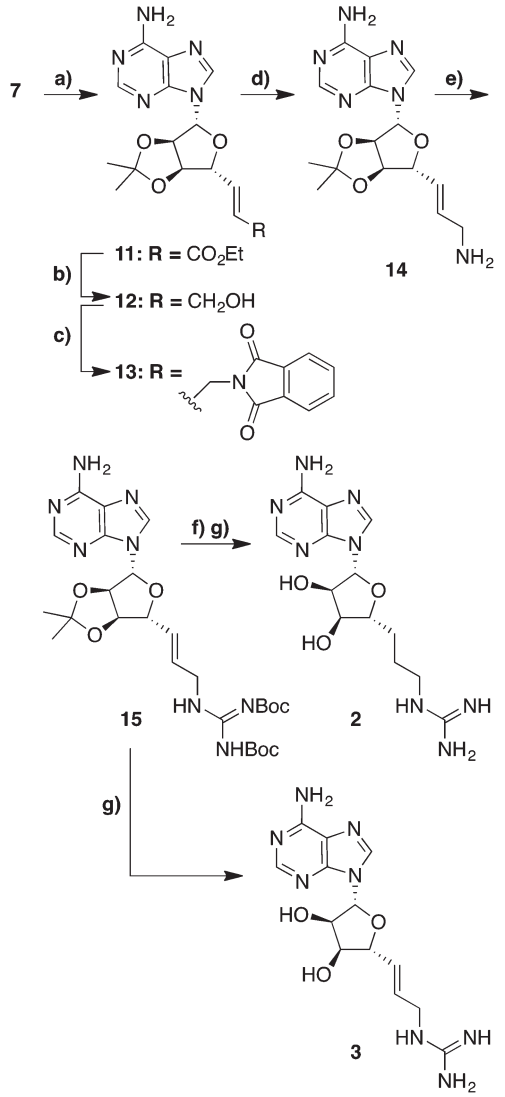


Scheme 1 Synthesis of inhibitor **1**. *Reagents and conditions*: (a) acetone cyanohydrin, PPh_3, DEAD, THF, 90%; (b) PtO_2, AcOH, H_2, quant.; (c) *N*,*N*′-di-Boc-*N*″-triflylguanidine, NEt_3, CH_2Cl_2, 75%; (d) TFA–CH_2Cl_2 (1 : 1) followed by TFA–CH_2Cl_2–H_2O (1 : 1 : 0.3), 89%.

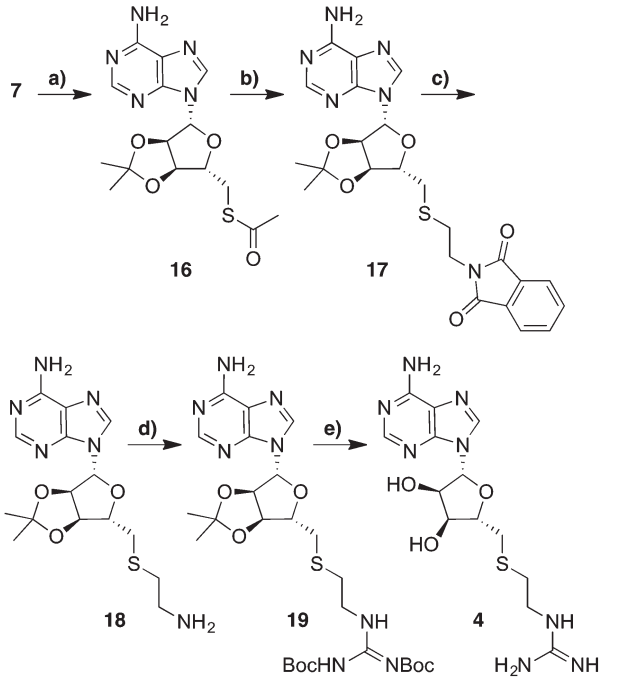
Mitsunobu reaction employing PPh_3, DEAD, and acetone cyanohydrin to yield the known nitrile **8**.[28] Subsequent reduction to the amine followed by treatment with *N*,*N*′-bis(*tert*-butoxycarbonyl)-*N*″-triflylguanidine (Goodman's reagent) provided the protected guanidine species **10**. Global deprotection was achieved by treating **10** with 50% TFA in CH_2Cl_2 to remove the Boc groups after which a small volume of water was added to the reaction mixture leading to the deprotection of the isopropylidene group to yield compound **1**. This two-step procedure was found to be necessary to ensure a clean and complete deprotection, while avoiding degradation of the desired product.

The synthesis of inhibitors **2** and **3** (Scheme 2) began with a one pot IBX oxidation of **7** to yield the intermediate aldehyde which was directly converted to alkene **11** *via* a Wittig reaction with triphenylcarbethoxymethylenephosphorane. Reduction of **11** with DIBAL yielded an alcohol which was transformed into intermediate **13** by a Mitsunobu reaction employing PPh_3, DEAD, and phthalimide.[28] Removal of the phthalimide group with methylamine provided amino compound **14** which was in turn converted into protected guanidine **15** by treatment with Goodman's reagent. Hydrogenation of **15** followed by deprotection using the procedure described above for **1** yielded compound **2** while compound **3** was obtained by the direct deprotection of **15**.

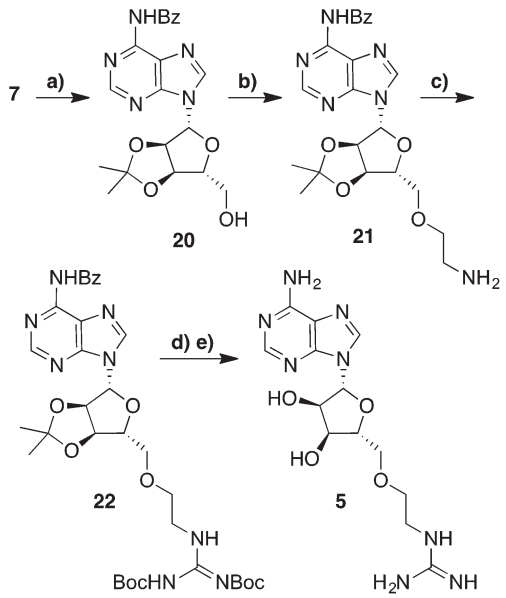
The preparation of thioether-linked inhibitor **4** (Scheme 3) began with the conversion of **7** to thioacetate intermediate **16** *via* the Mitsunobu reaction with PPh_3, DIAD, and thioacetic acid. A one-pot deacetylation and alkylation procedure was then applied to convert **16** into the phthalimide **17**. Phthalimide removal using methylamine followed by treatment with Goodman's reagent yielded protected guanidine **19**. Deprotection as described above yielded inhibitor **4**. The synthesis of oxyether-linked inhibitor **5** (Scheme 4) began with benzoyl protection of the adenosine unit in **7** to give intermediate **20**. Alkylation with chloroethylamine then provided direct access to amine **21** after which treatment with Goodman's reagent generated the protected guanidine **22**. Removal of the Boc and isopropylidene groups as described above followed by



Scheme 2 Synthesis of inhibitors **2** and **3**. *Reagents and conditions*: (a) IBX, $Ph_3P{=}CHCO_2Et$, DMSO, 79%; (b) DIBAL-H, hexane, CH_2Cl_2, 78%; (c) phthalimide, PPh_3, DEAD, THF, 83%; (d) $MeNH_2$, EtOH, 94%; (e) *N*,*N*′-di-Boc-*N*″-triflylguanidine, NEt_3, CH_2Cl_2, 92%; (f) 10% Pd/C, H_2, EtOH, 47%; (g) TFA–CH_2Cl_2 (1 : 1) followed by TFA–CH_2Cl_2–H_2O (1 : 1 : 0.3), 61% for 2 and 70% for 3.



Scheme 3 Synthesis of inhibitor **4**. *Reagents and conditions*: (a) thioacetic acid, PPh_3, DIAD, THF, 85%; (b) 2-bromoethyl-phthalimide, $NaOCH_3$, MeOH, 49%; (c) methylamine, EtOH, 89%; (d) *N*,*N*′-di-Boc-*N*″-triflylguanidine, NEt_3, CH_2Cl_2, 92%; (e) TFA–CH_2Cl_2 (1 : 1) followed by TFA–CH_2Cl_2–H_2O (1 : 1 : 0.3), 74%.

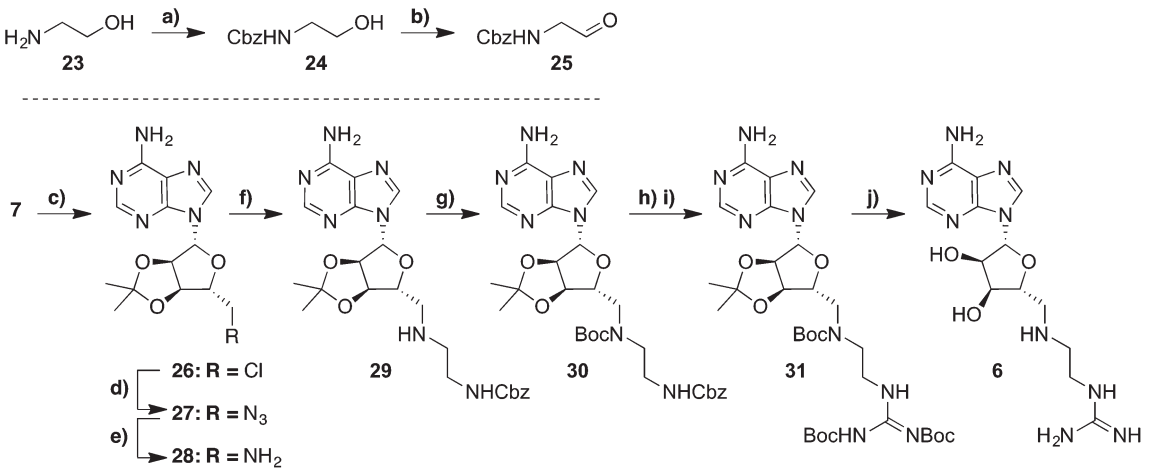


Scheme 4 Synthesis of inhibitor **5**. *Reagents and conditions*: (a) TMS-Cl, BzCl, pyridine, 77%; (b) chloroethylamine, NaH, DMF, 32%; (c) *N*,*N*′-di-Boc-*N*″-triflylguanidine, NEt_3, CH_2Cl_2, 87%; (d) TFA–CH_2Cl_2 (1 : 1) followed by TFA–CH_2Cl_2–H_2O (1 : 1 : 0.3); (e) NH_4OH, MeOH, 79% over 2 steps.

treatment with ammonium hydroxide to remove the benzoyl group yielded inhibitor **5**.

The preparation of inhibitor **6** bearing an amine spacer started with chlorination of **7** followed by conversion to azide **27** (Scheme 5). After hydrogenation to generate amine **28**, reductive amination with the known aldehyde **25** (prepared from ethanolamine) yielded the Cbz protected intermediate **29**. Following Boc protection and Cbz group removal, treatment with Goodman's reagent yielded protected guanidine **31** after which global deprotection led to compound **6**. For compounds **1–6** final purifications were performed using preparative RP-HPLC followed by concentration of pure fractions and lyophilisation to yield each compound as an amorphous solid.

Compounds **1–6** were evaluated as inhibitors of PRMT1, 4, and 6 as these three PRMTs are responsible for the majority of arginine methylation *in vivo* and also display good *in vitro*



Scheme 5 Synthesis of inhibitor **6**. *Reagents and conditions*: (a) Cbz-Cl, NEt_3, CH_2Cl_2, 63%; (b) sulfur trioxide-pyridine complex, DIPEA, DMSO, CH_2Cl_2, 38%; (c) thionyl chloride, HMPA, 77%; (d) NaN_3, DMF, 90%; (e) 10% Pd/C, H_2, MeOH, quant.; (f) 25, $Na(OAc)_3BH$, DCE, 44%; (g) Boc_2O, NEt_3, CH_2Cl_2, 50%; (h) 10% Pd/C, H_2, MeOH; (i) *N*,*N*′-di-Boc-*N*″-triflylguanidine, NEt_3, CH_2Cl_2, 97% over 2 steps; (j) TFA–CH_2Cl_2 (1 : 1) followed by TFA–CH_2Cl_2–H_2O (1 : 1 : 0.3), 74%.

activity. In addition, the lysine methyltransferase G9a was also included to assess the inhibitor selectivity against other methyltransferase families. Enzyme activity was measured using an established chemiluminescence-based assay wherein a substrate peptide derived from either the histone H4 or H3 tail is treated with AdoMet and the methyltransferase of interest. After incubation, a primary antibody is then added to specifically bind the methylated arginine or lysine residue followed by washing and treatment with a horseradish peroxidase (HRP) conjugated secondary antibody. In the final step, an HRP substrate is added to generate a chemiluminescent signal that is measured using a standard microplate reader. Inclusion of potential inhibitors in the first step of this process in turn allows for the convenient detection and quantitation of enzyme inhibition. Compounds **1–6** were initially screened against each enzyme at a threshold concentration of 50 μM. The known methyltransferase inhibitor *S*-adenosyl-L-homocysteine (AdoHcy), a by-product of the AdoMet cofactor resulting from methylation, was also included as a reference compound. In cases where no appreciable inhibition was observed at 50 μM further inhibition studies were not performed. Conversely, in cases where >50% inhibition was observed at the threshold inhibitor concentration of 50 μM, complete IC_{50} curves were generated (Table 1).

As shown in Table 1, compounds **1–3** bearing the shorter two- and three-atom spacers between the adenosine and guanidine groups generally displayed potent inhibition of the PRMTs tested while having no measurable effect on the activity of the lysine methyltransferase G9a. In contrast, compounds **4–6** which contain longer four-atom spacers were found for the most part to be devoid of activity at the threshold concentration tested. These findings would appear to be in line with the transition state model used to illustrate PRMT catalysis (Fig. 2). As described above in the design strategy section, compounds **2** and **3**, containing the 3-carbon linkers, were expected to most closely approximate the spacing between the two substrates in the PRMT transition state. In addition, owing to the presence of the double bond in its alkyl spacer, compound **3** is more rigid and as such could be expected to be a more potent PRMT inhibitor than compound **2**. The results of the inhibition studies however revealed compound **3** to be a slightly weaker inhibitor of both PRMT1 and 6 compared to compound **2** containing a fully saturated spacer. In general, compound **2** was shown to be a more active, albeit less selective, inhibitor of the three PRMTs tested. In contrast, compounds **1** and **3** were particularly potent and selective inhibitors of PRMT4 with IC_{50} values of 120 nM and 150 nM respectively. The near 100-fold selectivity of both **1** and **3** for PRMT4 *vs.* PRMT1 is of particular note in light of previously reported findings with structurally similar PRMT inhibitors. Specifically, Dowden and Ward found that compounds containing longer (five- to seven-atom long) spacers between the adenosine and guanidine groups were generally more active against PRMT1 and displayed little or no inhibition of PRMT4.[29,30] By comparison, the selectivity exhibited by com-

Table 1 Inhibition constants[a] for compounds **1–6** against PRMT1, 4, 6 and PKMT G9a

	PRMT1	PRMT4	PRMT6	G9a
AdoHcy	6.21 ± 0.56	0.67 ± 0.19	0.20 ± 0.25	16.64 ± 6.43
1	11.09 ± 2.77	0.12 ± 0.02	20.23 ± 8.67	*>50*
2	1.30 ± 0.38	0.56 ± 0.25	0.72 ± 0.33	*>50*
3	16.96 ± 3.73	0.15 ± 0.05	5.15 ± 1.27	*>50*
4	*>50*	*>50*	*>50*	3.18 ± 2.67
5	*>50*	*>50*	*>50*	*>50*
6	*>50*	*>50*	3.20 ± 3.93	*>50*

[a] IC_{50} values (μM) based on the best fit of dose–response curves generated from at least seven unique inhibitor concentrations (each concentration measured in duplicate). Curve fitting performed with GraphPad Prism 5.0.

pounds **1** and **3** in the present study suggests that by incorporating shorter spacers preferential inhibition of PRMT4 *vs.* PRMT1 can be achieved. Also worth mentioning among the generally inactive compounds **4–6** (bearing longer four-atom spacers) are two notable and somewhat surprising exceptions. As shown in Table 1, thioether linked analogue **4** showed good inhibition of G9a (IC_{50} 3.18 μM) while the amino-linked species **6** was specific for PRMT6 (IC_{50} 3.20 μM).

The binding modes of compounds **1–3** were next evaluated *via* docking studies. Ideally, a comparison of these compounds docked into the active sites of PRMT1, 4, and 6 might be expected to provide insight into the selectivity observed in the inhibition assays. This, however, is not possible given that the published crystal structure for PRMT1 (from rat, 1OR8.pdb[17]) is not suitable for docking studies as has been previously noted by others.[30–32] Specifically, the PRMT1 structure was obtained at pH 4.7 far from the value of 8.0–8.5 required for the enzyme activity resulting in a structure with a disordered N-terminus. In addition, the recently reported structure of PRMT6 (from the parasite *Trypanosoma brucei*, 4LWP.pdf[33]) is unfortunately not applicable for docking analysis as it also contains a disordered N-terminus. By comparison, the crystal structure of human PRMT4 obtained by Bertrand and co-workers is well suited to docking studies.[20] Compounds **1–3** were thus docked into the PRMT4 active site (PDB code 2Y1W[20]) revealing a number of key interactions. As illustrated in Fig. 4 the hydrogen bond network around the adenine ring is well conserved for each compound. Due to its shorter length, the linker in compound **1** is fully extended so as to allow the guanidine group to hydrogen bond with Glu258 and Met260. This extended conformation however leads to loss of a hydrogen bonding interaction between Ser217 and the furanose ring. Compounds **2** and **3** which contain longer three-atom linkers are able to hydrogen bond with Glu258 and Met260 without the loss of the interaction with Ser217. In addition, both **2** and **3** benefit from H-bonding interactions between their guanidine units and Met163 and Met269 respectively. It is interesting to note that in each of the docking analyses performed, hydrogen bonding interactions between the guanidine moieties of **1–3** and Glu258 were found but none were detected with Glu267. As described above (Fig. 2), the "double-E loop" is conserved across all PRMTs and is critical for PRMT catalysis (in the case of PRMT4 Glu258 and Glu267 correspond to these two residues). The observation that the guanidine groups in compounds **1–3** do not appear to simultaneously interact with both of these glutamate side chains suggests that the compounds may be optimized to further enhance binding.

The potent PRMT4 inhibition displayed by compounds **1–3** places them amongst the most active PRMT4 inhibitors reported to date. Only the pyrazole-based compounds reported by researchers at both MethylGene and Bristol-Myers exhibit more potent PRMT4 inhibition, however they lack activity in cell-based assays.[34,35] The groups of Bedford and Mai also recently generated novel PRMT4 inhibitors which, while generally less active, showed very good PRMT4 selectivity *versus* other methyltransferases and in some cases exhibited activity in cellular models of disease.[36] As a preliminary measure of cellular activity, we evaluated the effects of compounds **1–3** on cell proliferation using a standard MTT assay with Caco-2 colon cancer cells and MCF-7 breast cancer cells. These investigations revealed that compounds **1–3** have no significant effect

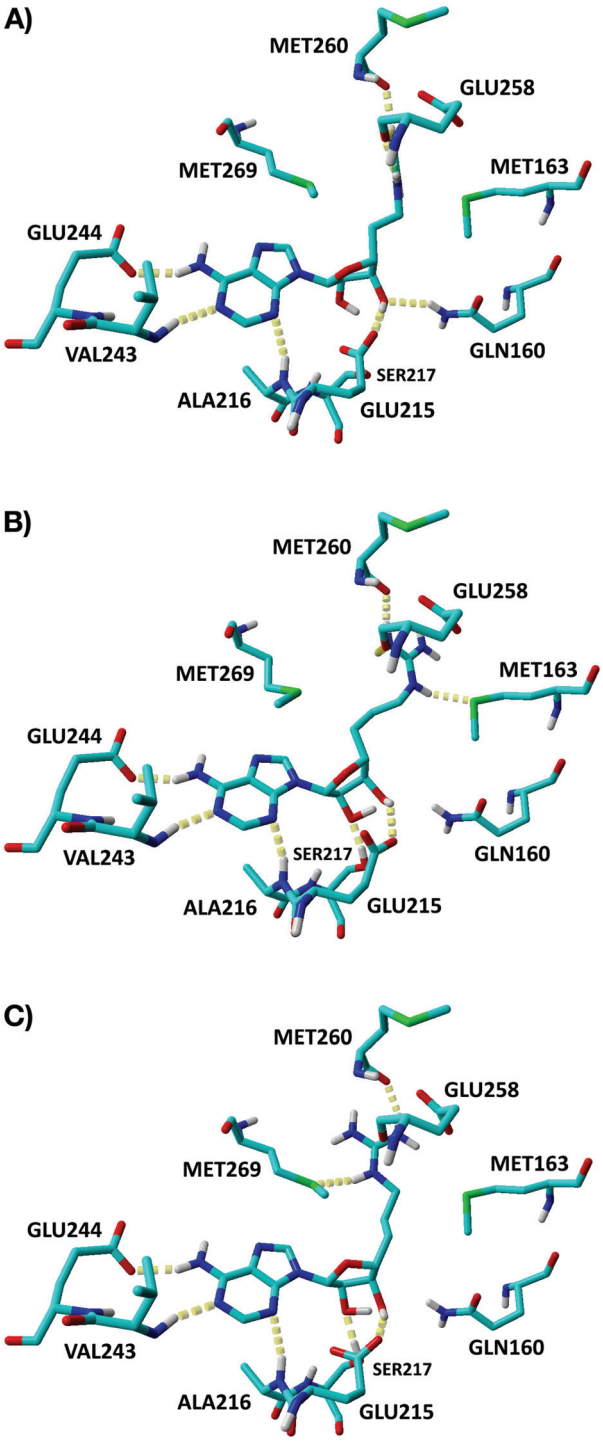


Fig. 4 Proposed hydrogen bond network after docking and minimization of compounds **1–3** in the PRMT4 active site (aliphatic hydrogens omitted for clarity). (A) Compound **1**; (B) compound **2**; (C) compound **3**.

on the viability of these cell lines at concentrations up to 100 μM. This may be due to the inability of the highly charged compounds to enter the cells and suggests that inhibitors containing less polar analogues of the guanidine moiety (*i.e.* amidine or *N*-hydroxy guanidine) may be more effective.

Conclusion

We here describe the design and synthesis of a series of PRMT inhibitors designed to simultaneously occupy both the AdoMet and peptide substrate binding sites. Given that the active site architectures of all human PRMTs are highly conserved, the selectivity observed amongst the inhibitors prepared is rather surprising. Of particular note is the potent PRMT4 inhibition achieved by those inhibitors with shorter two- and three-atom linkers connecting the adenosine and guanidine moieties. By comparison, analogues containing four-atom linkers were generally inactive. Compound **4** bearing a four-atom thioether linkage was, however, found to effectively inhibit the lysine methyltransferase G9a and none of the PRMTs tested. In addition, compound **6** with a four-atom linker comprising of a secondary amine was found to inhibit PRMT6 exclusively. Docking studies were also performed with inhibitors **1–3** and PRMT4 and indicate the possible modes of binding that support the strong inhibition observed with these three compounds. These docking studies also suggest possibilities for further improving the inhibitor affinity by introducing hydrogen-bonding substituents on the guanidine moiety to allow for simultaneous interactions with both Glu258 and Glu267. To this end, future investigations will involve the incorporation of *N*-hydroxy and *N*-amino substituted guanidines[26,37] into the inhibitor structures here described. Aside from the possibility of enhancing inhibition, such modifications may also serve to improve the cell permeability of these inhibitors. The results of these ongoing studies will be reported in due course.

Experimental

Reagents and general methods

All reagents employed were of American Chemical Society (ACS) grade or finer and were used without further purification unless otherwise stated. Compounds **8**,[28] **9**,[28] **11–14**,[28] **16**,[38] **20**,[39] **24**,[40] **25**,[41] **26**,[42] **27**,[43] **28**,[44] and **29–30**[45] were synthesized according to previously described procedures. All known compounds prepared had NMR spectra, mass spectra, and optical rotation values consistent with the assigned structures. All reactions and fractions from column chromatography were monitored by thin layer chromatography (TLC) using plates with a UV fluorescent indicator (normal SiO_2, Merck 60 F254). One or more of the following methods were used for visualization: UV absorption by fluorescence quenching; phosphomolybdic acid : ceric sulfate : sulfuric acid : H_2O (10 g : 1.25 g : 12 mL : 238 mL) staining; $KMnO_4$ staining; PPh_3 staining; ninhydrin staining. Flash chromatography was performed using Merck type 60, 230–400 mesh silica gel. The final compounds **1–6** were purified by preparative scale RP-HPLC using a Reprosil-Pur C18 column (10 μm, 250 × 22 mm) eluted with a water–methanol gradient moving from 0% to 50% MeOH (0.1% TFA) over 60 minutes at a flow-rate of 1.4 mL min^{-1} with UV detection at 214 nm. Purity was confirmed to be ≥95% by analytical RP-HPLC using a Reprosil-Pur C18 column (5 μm, 250 × 4.6 mm) eluted with a water–methanol gradient moving from 0% to 50% MeOH (0.1% TFA) over 40 minutes at a flow rate of 0.7 mL min^{-1} with UV detection at 214 nm and 254 nm.

Instrumentation for compound characterization

^{1}H NMR spectra were recorded at 300 MHz or 400 MHz with chemical shifts reported in parts per million (ppm) downfield relative to tetramethylsilane (TMS) or H_2O (δ 4.79). ^{1}H NMR data are reported in the following order: multiplicity (s, singlet; d, doublet; t, triplet; q, quartet and m, multiplet), coupling constant (J) in hertz (Hz) and the number of protons. When appropriate, the multiplicity is preceded by br, indicating that the signal was broad. ^{13}C NMR spectra were recorded at 75.5 MHz with chemical shifts reported relative to $CDCl_3$ (δ 77.16). The ^{13}C NMR spectra of compounds **1–6** (as their TFA salts) were recorded in D_2O and therefore referenced to the TFA quartet at 116.6. High-resolution mass spectrometry (HRMS) analysis was performed using an ESI instrument.

Synthetic procedures and compound characterization

2-(2-((3a*R*,4*R*,6*R*,6a*R*)-6-(6-Amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)ethyl)-*N*,*N*′-di-Boc-guanidine (10). To a suspension of amine 9[28] (570 mg, 1.5 mmol) in CH_2Cl_2 (2 mL), a solution of *N*,*N*′-bis(*tert*-butoxycarbonyl)-*N*″-triflylguanidine (392 mg, 1.0 mmol) and NEt_3 (416 μL, 3.0 mmol) in CH_2Cl_2 (3 mL) was added. The mixture was stirred overnight at room temperature and diluted with CH_2Cl_2 (15 mL). The organic layer was washed with 1 N $KHSO_4$ (10 mL), saturated $NaHCO_3$ (10 mL) and brine (10 mL), dried over Na_2SO_4 and filtered. Evaporation *in vacuo* and column chromatography (SiO_2: EtOAc–MeOH (98 : 2)) afforded **10** (422 mg, 75%) as a white powder. R_f 0.41 (95 : 5 EtOAc–MeOH). ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.41 (br m, 1H), 8.30 (s, 1H), 7.87 (s, 1H), 6.15 (br s, 2H), 6.02 (s 1H), 5.44 (d, J = 6.3 Hz, 1H), 4.97–4.87 (m, 1H), 4.23 (q, J = 6.6 Hz, 1H), 3.67–3.56 (m, 1H), 3.43–3.33 (m, 1H), 2.01 (q, J = 6.2 Hz, 2H), 1.57 (s, 3H), 1.45 (s, 9H), 1.36 (s, 9H), 1.34 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 163.5, 155.9, 155.7, 153.1, 152.8, 149.3, 120.3, 114.8, 90.1, 85.6, 84.0, 83.9, 82.8, 79.1, 38.1, 32.1, 28.3, 27.9, 27.2, 25.4. HRMS (ESI): calculated for $C_{25}H_{38}N_8O_7$ $[M + Na]^+$ 585.2761, found 585.2773.

2-(2-((2*R*,3*S*,4*R*,5*R*)-5-(6-Amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)ethyl)guanidine (1). Compound **10** (56.5 mg, 0.1 mmol) is dissolved in a 1 : 1 mixture of TFA and CH_2Cl_2 (10 mL) and stirred for 2 hours at room temperature. H_2O (1.5 mL) is added, and the mixture is stirred for 30 minutes at room temperature. Evaporation *in vacuo*, purifi-

cation by preparative HPLC and freeze-drying of the pure fractions afforded the final compound **1** (38.8 mg, 89%) as a white powder. ^{1}H NMR (400 MHz, D_2O): δ 8.46 (s, 1H), 8.45 (s, 1H), 6.12 (d, J = 4.7 Hz, 1H), 4.87 (t, J = 5.0 Hz, 1H), 4.35 (t, J = 5.1 Hz, 1H), 4.27–4.19 (m, 1H), 3.37 (t, J = 6.6 Hz, 2H), 2.17–2.06 (m, 1H). ^{13}C NMR (75 MHz, D_2O): δ 157.0, 150.2, 148.5, 144.7, 143.3, 119.3, 89.2, 82.6, 73.9, 73.5, 38.4, 31.9. HRMS (ESI): calculated for $C_{12}H_{18}N_8O_3$ $[M + H]^+$ 323.1580, found 323.1578.

2-((*E*)-3-((3a*R*,4*R*,6*R*,6a*R*)-6-(6-Amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)allyl)-*N*,*N*′-di-Boc-guanidine (15). To a suspension of amine 14[28] (332 mg, 1.0 mmol) in CH_2Cl_2 (2 mL), a solution of *N*,*N*′-bis(*tert*-butoxycarbonyl)-*N*″-triflylguanidine (392 mg, 1.0 mmol) and NEt_3 (139 μL, 1.0 mmol) in CH_2Cl_2 (2 mL) was added. The mixture was stirred until completion followed by TLC (EtOAc–MeOH (9 : 1)). The mixture was diluted with CH_2Cl_2 (6 mL) and the organic layer was washed with 1 N $KHSO_4$ (10 mL), saturated $NaHCO_3$ (10 mL) and brine (10 mL), dried over Na_2SO_4 and filtered. Evaporation *in vacuo* and column chromatography (SiO_2: EtOAc–MeOH (98 : 2)) afforded compound 15 (526 mg, 92%) as a white powder. R_f 0.43 (9 : 1 EtOAc–MeOH); ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.36 (s, 1H), 8.34 (t, J = 5.1 Hz, 1H), 7.89 (s, 1H), 6.09 (d, J = 2.1 Hz, 1H), 6.01 (br s, 2H), 5.79–5.76 (m, 2H), 5.50 (dd, J = 6.3, 2.1 Hz, 1H), 5.01 (dd, J = 6.3, 3.6 Hz, 1H), 4.68 (t, J = 4.2 Hz, 1H), 4.01–3.97 (m, 2H), 1.63 (s, 3H), 1.49 (s, 18H), 1.40 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 163.4, 155.9, 155.6, 153.2, 153.1, 149.4, 129.7, 129.1, 120.2, 114.6, 90.3, 87.3, 84.6, 84.2, 83.2, 79.4, 41.9, 28.3, 28.0, 27.1, 25.4. HRMS (ESI): calculated for $C_{26}H_{38}N_8O_7$ $[M + H]^+$ 575.2942, found 575.2921.

2-(2-((2*R*,3*S*,4*R*,5*R*)-5-(6-Amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)propyl)guanidine (2). To a solution of compound **15** (250 mg, 0.44 mmol) in EtOH (10 mL), a suspension of Pd/C (10%, 200 mg, excess) in EtOH (5 mL) was added. The mixture was hydrogenated (5 bar H_2) for 72 hours. The catalyst was removed by filtration over celite and the solvent was evaporated *in vacuo*. Purification by column chromatography (SiO_2: EtOAc–MeOH (98 : 2)) afforded the saturated intermediate (118 mg, 47%) as a white powder, which was subjected to the deprotection and purification procedure as described for compound **1** to afford the final compound **2** (39 mg, 61%) as a white powder. ^{1}H NMR (400 MHz, D_2O): δ 8.46 (s, 1H), 8.45 (s, 1H), 6.12 (d, J = 4.9 Hz, 1H), 4.88–4.81 (m, 1H), 4.31 (t, J = 5.1 Hz, 1H), 4.22–4.14 (m, 1H), 3.24 (t, J = 6.7 Hz, 2H), 1.99–1.53 (m, 4H). ^{13}C NMR (75 MHz, D_2O): δ 157.0, 150.6, 148.6, 145.3, 142.8, 119.2, 88.6, 84.7, 74.0, 73.4, 41.0, 29.9, 24.6. HRMS (ESI): calculated for $C_{13}H_{20}N_8O_3$ $[M + H]^+$ 337.1737, found 337.1733.

2-((*E*)-3-((2*R*,3*S*,4*R*,5*R*)-5-(6-Amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)allyl)guanidine (3). Starting from compound **15** (100 mg, 0.17 mmol), the deprotection and purification procedure as described for compound **1** afforded the final compound **3** (41 mg, 70%) as a white powder. ^{1}H NMR (400 MHz, D_2O): δ 8.45 (m, 2H), 6.17 (d, J = 4.0 Hz, 1H), 5.97–5.91 (m, 2H), 4.86–4.81 (m, 1H), 4.65 (br s, 1H), 4.39 (t, J = 5.2 Hz, 1H), 3.91 (s, 2H). ^{13}C NMR (75 MHz, D_2O): δ 157.1, 150.3, 148.4, 144.9, 142.9, 130.1, 128.1, 119.2, 89.0, 84.5, 73.9, 73.8, 42.0. HRMS (ESI): calculated for $C_{13}H_{18}N_8O_3$ $[M + H]^+$ 335.1580, found 335.1576.

2-(2-((((3a*S*,4*S*,6*R*,6a*R*)-6-(6-Amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methyl)thio)-ethyl)isoindoline-1,3-dione (17). To a solution of thioacetate 16[38] (500 mg, 1.37 mmol) in MeOH (30 mL), 2-bromoethylphthalimide (536 mg, 2.10 mmol) was added and the mixture was cooled to −20 °C under a N_2 atmosphere. Sodium methoxide (165 mg, 3.06 mmol) was added and the mixture was stirred for 90 minutes at −20 °C and for 72 hours at room temperature. The solvent was evaporated *in vacuo* and extracted from water–chloroform (3 × 30 mL). The organic layers were combined, dried over Na_2SO_4 and filtered. Evaporation *in vacuo* and column chromatography (SiO_2: CH_2Cl_2–MeOH (9 : 1)) afforded compound **17** (332 mg, 49%) as a white foam. R_f 0.61 (9 : 1 CH_2Cl_2–MeOH). ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.37 (s, 1H), 7.95 (s, 1H), 7.88–7.79 (m, 2H), 7.66–7.75 (m, 2H), 6.09 (d, J = 2.1 Hz, 1H), 5.98 (br s, 2H), 5.49 (dd, J = 6.3, 2.1 Hz, 1H), 5.08 (dd, J = 6.3, 3.3 Hz, 1H), 4.41 (td, J = 6.7, 3.3 Hz, 1H), 3.82 (t, J = 6.9 Hz, 2H), 2.92 (dd, J = 6.8, 2.7 Hz, 2H), 2.83 (t, J = 6.9 Hz, 2H), 1.61 (s, 3H), 1.39 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 168.0, 155.9, 153.2, 149.1, 134.0, 131.9, 123.3, 120.2, 114.5, 90.7, 86.7, 84.1, 83.7, 36.8, 33.7, 30.4, 27.1, 25.3. HRMS (ESI): calculated for $C_{23}H_{24}N_6O_5S$ $[M + H]^+$ 497.1607, found 497.1662.

9-((3a*R*,4*R*,6*S*,6a*S*)-6-(((2-Aminoethyl)thio)methyl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)-9*H*-purin-6-amine (18). For deprotection of the phthalimide, compound **17** (271 mg, 0.55 mmol) was dissolved in a solution of methylamine in EtOH (33%, 9 mL) and stirred at room temperature for 16 hours. The solvent was evaporated *in vacuo* and the residue was redissolved in $CHCl_3$ (20 mL). The organic layer was extracted with 10% aqueous acetic acid (25 mL). The aqueous layer was washed with $CHCl_3$ (3 × 20 mL), the pH was adjusted to >12 using 2 N NaOH and then was extracted with $CHCl_3$ (4 × 20 mL). The final organic layers were dried over Na_2SO_4 and filtered and the solvents were evaporated *in vacuo* to afford free amine **18** (177 mg, 89%). R_f 0.18 (9 : 1 CH_2Cl_2–MeOH). ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.26 (s, 1H), 7.88 (s, 1H), 6.67 (s, 2H), 6.03 (d, J = 2.1 Hz, 1H), 5.47 (dd, J = 6.4, 2.1 Hz, 1H), 5.01 (dd, J = 6.4, 3.1 Hz, 1H), 4.32 (td, J = 6.9, 3.1 Hz, 1H), 2.85–2.61 (m, 4H), 2.53 (t, J = 6.3 Hz, 2H), 1.54 (s, 3H), 1.52 (br s, 2H), 1.33 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 156.1, 153.1, 149.1, 120.2, 114.4, 90.8, 86.9, 84.0, 83.8, 41.0, 36.7, 34.0, 27.1, 25.3. HRMS (ESI): calculated for $C_{15}H_{22}N_6O_3S$ $[M + H]^+$ 367.1552, found 367.1535.

2-(2-((((3a*S*,4*S*,6*R*,6a*R*)-6-(6-Amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methyl)-thio)ethyl)-*N*,*N*′-di-Boc-guanidine (19). Following the procedure as described for compound **15**, amine **18** (160 mg, 0.44 mmol) was guanidinylated. After purification by column chromatography (SiO_2: CH_2Cl_2–MeOH (9 : 1)), compound **19** was obtained as a clear oil (245 mg, 92%). R_f 0.63 (9 : 1 CH_2Cl_2–MeOH). ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.50 (t, J = 5.3 Hz, 1H), 8.25 (s, 1H), 7.90 (s, 1H), 6.54 (s, 2H), 6.03 (d, J = 1.8 Hz, 1H), 5.44 (dd, J = 6.4,

1.8 Hz, 1H), 5.00 (dd, J = 6.3, 3.1 Hz, 1H), 4.32 (td, J = 6.7, 3.1 Hz, 1H), 3.50 (q, J = 6.2 Hz, 2H), 2.79 (m, 2H), 2.64 (t, J = 6.4 Hz, 2H), 1.54 (s, 3H), 1.40 (s, 18H), 1.32 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 163.4, 156.0, 155.9, 153.0, 149.1, 120.1, 114.4, 90.6, 86.5, 84.0, 83.6, 83.1, 79.2, 39.8, 33.9, 31.6, 28.2, 28.0, 27.0, 25.3. HRMS (ESI): calculated for $C_{26}H_{40}N_8O_7S$ $[M + Na]^+$ 631.2638, found 631.2685.

1-(2-((((2*S*,3*S*,4*R*,5*R*)-5-(6-Amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)thio)ethyl)guanidine (4). Starting from protected guanidine **19** (140 mg, 0.23 mmol), the deprotection and purification procedure as described for compound **1** afforded the final compound **4** (63 mg, 74%). ^{1}H NMR (400 MHz, D_2O): δ 8.49 (s, 1H), 8.41 (s, 1H), 6.12 (d, J = 4.7 Hz, 1H), 4.85 (t, J = 5.0 Hz, 1H), 4.43 (t, J = 5.1 Hz, 1H), 4.36–4.28 (m, 1H), 3.37 (t, J = 6.5 Hz, 2H), 3.12–2.95 (m, 2H), 2.78 (t, J = 6.5 Hz, 2H). ^{13}C NMR (75 MHz, D_2O): δ 157.4, 150.5, 148.8, 145.1, 143.5, 119.5, 89.2, 84.5, 74.2, 72.9, 41.2, 34.3, 31.6. HRMS (ESI): calculated for $C_{13}H_{20}N_8O_3S$ $[M + H]^+$ 369.1457, found 369.1465.

***N*-(9-((3a*R*,4*R*,6*R*,6a*R*)-6-((2-Aminoethoxy)methyl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)-9*H*-purin-6-yl)benzamide (21).** To a solution of benzoyl-protected adenosine **20**[39] (617 mg, 1.5 mmol) in dry DMF (10 mL) under Ar at 0 °C, NaH (dispersion in mineral oil (50%), 720 mg, 15 mmol) was added over a period of 1 hour. The mixture was stirred for 1 hour allowing it to warm to room temperature. The mixture was then cooled to −5 °C and chloroethylamine hydrochloride (1.05 g, 9.0 mmol) was added. The mixture was stirred for 5.5 hours, allowing it to warm to room temperature, and quenched with MeOH (6 mL). Evaporation *in vacuo* and column chromatography with the product absorbed on silica (SiO_2: first CH_2Cl_2–MeOH (9 : 1), then CH_2Cl_2–MeOH–NEt_3 (9 : 1 : 0.5)) afforded amine **21** (220 mg, 32%) as an off-white powder. R_f 0.32 (9 : 1 : 0.5 CH_2Cl_2–MeOH–NEt_3); ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.72 (s, 1H), 8.54 (s, 1H), 7.97–7.88 (m, 2H), 7.55–7.33 (m, 3H), 6.39 (br s, 2H), 6.19 (d, J = 2.1 Hz, 1H), 5.38 (dd, J = 6.2, 2.1 Hz, 1H), 5.12 (dd, J = 6.2, 2.6 Hz, 1H), 4.44–4.35 (m, 1H), 3.69–3.48 (m, 4H), 3.02–2.86 (m, 2H), 1.55 (s, 3H), 1.34 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 165.6, 152.4, 151.7, 149.2, 142.9, 133.3, 132.7, 128.6, 128.0, 123.4, 114.1, 90.8, 86.1, 84.6, 81.2, 71.0, 69.0, 39.8, 27.0, 25.3. HRMS (ESI): calculated for $C_{22}H_{26}N_6O_5$ $[M + H]^+$ 455.2043, found 455.2002.

***N*-(9-((3a*R*,4*R*,6*R*,6a*R*)-6-((2-(*N*,*N*′-Di-Boc-guanidino)-ethoxy)methyl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]-dioxol-4-yl)-9*H*-purin-6-yl)benzamide (22).** Following the procedure as described for compound **15**, amine **21** (220 mg, 0.48 mmol) was guanidinylated. After purification by column chromatography (SiO_2: CH_2Cl_2–MeOH (98 : 2)), compound **22** was obtained as a white powder (291 mg, 87%). R_f 0.24 (98 : 2 CH_2Cl_2–MeOH). ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.30 (br s, 1H), 8.73 (s, 1H), 8.44 (t, J = 4.8 Hz, 1H), 8.27 (s, 1H), 8.03–7.92 (m, 2H), 7.57–7.36 (m, 3H), 6.22 (d, J = 2.4 Hz, 1H), 5.33 (dd, J = 6.1, 2.4 Hz, 1H), 4.99 (dd, J = 6.1, 2.0 Hz, 1H), 4.56–4.48 (m, 1H), 3.75–3.19 (m, 6H), 1.58 (s, 3H), 1.40 (s, 9H), 1.38 (s, 9H), 1.35 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 164.8, 163.3, 156.1, 153.2, 151.4, 149.4, 133.5, 132.7, 128.7, 127.9, 123.4, 114.0, 92.0, 86.1, 85.0, 83.2, 81.7, 79.4, 71.1, 69.4, 40.0, 28.2, 28.0, 27.1, 25.3. HRMS (ESI): calculated for $C_{33}H_{44}N_8O_9$ $[M + Na]^+$ 719.3129, found 719.3180.

1-(2-(((2*R*,3*S*,4*R*,5*R*)-5-(6-Amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methoxy)ethyl)guanidine (5). Starting from protected guanidine **22** (56 mg, 0.08 mmol), the deprotection procedure as described for compound **1** afforded the benzoyl-protected intermediate which was subsequently deprotected overnight in a mixture of NH_4OH (3 mL) and MeOH (5 mL). Purification by preparative HPLC and freeze-drying as described for compound **1** afforded the final compound **5** (22.4 mg, 79%). ^{1}H NMR (400 MHz, D_2O): δ 8.47 (s, 1H), 8.41 (s, 1H), 6.17 (d, J = 4.6 Hz, 1H), 4.84–4.80 (m, 1H), 4.48 (t, J = 5.0 Hz, 1H), 4.40–4.33 (m, 1H), 3.95–3.79 (m, 2H), 3.79–3.65 (m, 2H), 3.44–3.32 (m, 2H). ^{13}C NMR (75 MHz, D_2O): δ 157.4, 150.3, 148.5, 144.9, 142.9, 119.1, 89.0, 84.0, 74.5, 70.6, 70.4, 69.4, 41.5. HRMS (ESI): calculated for $C_{13}H_{20}N_8O_4$ $[M + H]^+$ 353.1686, found 353.1653.

2-(2-((((3a*R*,4*R*,6*R*,6a*R*)-6-(6-Amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methyl)-(Boc)amino)ethyl)-*N*,*N*′-di-Boc-guanidine (31). The Cbz-protected intermediate **30**[45] (148 mg, 0.25 mmol) was deprotected overnight using 10% palladium on activated charcoal (105 mg, excess) in methanol (6 mL) at room temperature under a H_2 atmosphere. After removal of the catalyst over celite and concentration the deprotected amine was guanidinylated following the procedure as described for compound **15**. After purification by column chromatography (SiO_2: CH_2Cl_2–MeOH (925 : 75)), compound **31** was obtained as a clear oil (168 mg, 97% over two steps). R_f 0.54 (9 : 1 CH_2Cl_2–MeOH). ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.40–8.20 (m, 2H), 7.89 (s, 1H), 6.44 (s, 2H), 6.01 (m, 1H), 5.40 (dd, J = 6.3, 1.9 Hz, 1H), 4.94 (m, 1H), 4.51–4.21 (m, 1H), 3.80–3.12 (m, 6H), 1.53 (s, 3H), 1.40 (s, 18H), 1.37 (s, 9H), 1.31 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 163.4, 156.3, 155.9, 155.6, 153.1, 149.2, 120.2, 114.6, 90.5, 85.7, 84.7, 83.8, 83.0, 82.4, 80.3, 79.1, 49.0, 47.0, 39.0, 28.3, 28.0, 27.8, 27.2, 25.4. HRMS (ESI): calculated for $C_{31}H_{49}N_9O_9$ $[M + Na]^+$ 714.3551, found 714.3565.

1-(2-((((2*R*,3*S*,4*R*,5*R*)-5-(6-Amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)amino)ethyl)-guanidine (6). Starting from protected guanidine **31** (160 mg, 0.23 mmol), the deprotection and purification procedure as described for compound **1** afforded the final compound **6** (61 mg, 74%). ^{1}H NMR (400 MHz, D_2O): δ 8.46–8.41 (m, 2H), 6.16 (d, J = 4.0 Hz, 1H), 4.88 (t, J = 4.0 Hz, 1H), 4.54–4.43 (m, 2H), 3.72–3.54 (m, 4H), 3.39 (t, J = 5.9 Hz, 2H). ^{13}C NMR (75 MHz, D_2O): δ 157.4, 150.3, 148.4, 144.7, 143.9, 119.7, 90.2, 79.8, 73.6, 71.8, 49.8, 46.5, 37.8. HRMS (ESI): calculated for $C_{13}H_{21}N_9O_3$ $[M + H]^+$ 352.1846, found 352.1834.

Methyltransferase inhibition assays

Methyltransferase inhibition assays were performed as previously described[46] using commercially available chemiluminescent assay kits for PRMT1, PRMT4 and PRMT6 and G9a (purchased from BPS Bioscience, San Diego, CA, USA). The enzymatic reactions were conducted in duplicate at room

temperature for 1 hour (20 minutes for PRMT1) in substrate-coated well plates at a final reaction volume of 50 μL containing the manufacturer's proprietary assay buffer, AdoMet (at a concentration of 5-times the respective K_m value for each enzyme), the methyltransferase enzyme: PRMT1 (80 ng per reaction), PRMT4 (200 ng per reaction), PRMT6 (180 ng per reaction), G9a (160 ng per reaction), and inhibitors **1–6** (in the range of concentrations: 0.001–100 μM in water). Before addition of the AdoMet, the enzyme was first incubated with the inhibitor for 15 minutes at 37 °C. Positive controls were performed in the absence of inhibitors using water to keep the final volume consistent. Blanks and substrate controls were performed in the absence of the enzyme and AdoMet respectively. Following the enzymatic reactions, 100 μL of primary antibody (recognizing the respective immobilized methylated arginine product) was added to each well and the plate was incubated at room temperature for an additional 1 hour. Then 100 μL of secondary horseradish peroxidase (HRP)-conjugated antibody was added to each well and the plate was incubated at room temperature for additional 30 minutes. Finally, 100 μL of an HRP substrate mixture was added to the wells and the luminescence was measured directly using a standard microplate reader. In all cases enzyme activity measurements were performed in duplicate at each of the inhibitor concentrations evaluated.

The luminescence data were analysed using GraphPad Prism (version 5.0). The luminescence of the positive control (L_p) in each dataset was defined as 100% activity. This value was included in the IC_{50} graphs at a concentration of two log values below the lowest concentration tested. The luminescence data of the negative controls (L_n) in each dataset were subtracted from the obtained luminescence data. The percent activity in the presence of each inhibitor was calculated according to the following equation: % activity = $(L - L_n)/(L_p - L_n)$, where L = the luminescence in the presence of the compound, L_n = the luminescence in the absence of the enzyme, and L_p = the luminescence in the absence of the inhibitor. The percent activity values were plotted as a function of inhibitor concentrations and fitted using non-linear regression analysis of the Sigmoidal dose–response curve generated using the equation $Y = B + (T - B)/1 + 10^{((\text{Logic50}-X)\times \text{Hill Slope})}$, where Y = percent activity, B = minimum percent activity, T = maximum percent activity, X = the logarithmic concentration of the compound and Hill Slope = slope factor or Hill coefficient. The IC_{50} value was determined by the concentration resulting in a half-maximal percent activity. The standard errors were reported using the symmetrical CI function. The IC_{50} values measured for AdoHcy which served as a reference compound are similar to those reported by the supplier of the methyltransferase kits.

Cell proliferation assays

The MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assay was based on a known procedure.[47] Caco-2 colon cancer cells (HTB-22) and MCF-7 breast cancer cells (HTB-37) were obtained from the American Type Culture Collection. Cells were grown in DMEM (Dulbecco's modified eagle's medium) supplemented with 10% FBS and 10 000 U per mL penicillin and 10 mg per mL streptomycin at 37 °C under a 5% CO_2 atmosphere in air. Monolayers were passaged on reaching ~90% confluency with trypsin–EDTA. The MCF-7 and Caco-2 cells were plated in duplicate into 96-well plates in media at a concentration of 5000 cells per well. At 40–50% confluency (48 hours post-seeding) compounds **1–3** were added at concentrations of 1, 10 and 100 μM respectively to each well plate (each compound concentration applied to four separate wells per plate). The plates were then incubated for 24 hours in the presence of the test compounds after which the media was removed and the cells incubated for 2 hours in fresh media. The cells were next treated with a MTT solution at a final concentration of 0.1 mg ml^{-1} in a serum-free medium for 4 hours at 37 °C/5% CO_2. After 4 hours, the plates were centrifuged and the MTT solution was removed. DMSO was added to dissolve the blue formazan product and UV absorption was measured at 550 nm.

Docking studies

Compounds **1–3** were built in YASARA (Version 13.9.8[48]), based on the *S*-adenosyl homocysteine structure in complex with PRMT3[49] (PDB code 1F3L). After aligning 1F3L to the structure of PRMT4[20] (PDB code 2Y1W) using MUSTANG[50] compounds **1–3** were in a good starting position for docking to PRMT4, using the default run_docklocal macro. According to this macro the ligand and all protein sidechains within 7 Å of the ligand were energy minimized using the steepest descent and simulated annealing with the NOVA forcefield,[51] followed by docking with AutoDock Vina.[52] Of the best docking pose, all ligand and protein atoms were again energy minimized with the steepest descent and simulated annealing with the NOVA forcefield, to yield the images used in generating Fig. 4.

Acknowledgements

We thank Javier Sastre Toraño for HRMS analysis. Financial support provided by Utrecht University and the Netherlands Organization for Scientific Research (VIDI grant to NIM) is acknowledged.

Notes and references

1 L. C. Boffa, J. Karn, G. Vidali and V. G. Allfrey, *Biochem. Biophys. Res. Commun.*, 1977, **74**, 969–976.
2 W. K. Paik and S. Kim, in *Advances in Enzymology and Related Areas Molecular Biology*, John Wiley & Sons, Inc., 2006, pp. 227–286.
3 M. Bremang, A. Cuomo, A. M. Agresta, M. Stugiewicz, V. Spadotto and T. Bonaldi, *Mol. BioSyst.*, 2013, **9**, 2231–2247.
4 K. B. Sylvestersen, H. Horn, S. Jungmichel, L. J. Jensen and M. L. Nielsen, *Mol. Cell. Proteomics*, 2014, **13**, 2072–2088.
5 M. T. Bedford and S. G. Clarke, *Mol. Cell*, 2009, **33**, 1–13.

6 M. T. Bedford and S. Richard, *Mol. Cell*, 2005, **18**, 263–272.
7 Y. Yang and M. T. Bedford, *Nat. Rev. Cancer*, 2013, **13**, 37–50.
8 R. A. Copeland, M. P. Moyer and V. M. Richon, *Oncogene*, 2013, **32**, 939–946.
9 M. Yoshimatsu, G. Toyokawa, S. Hayami, M. Unoki, T. Tsunoda, H. I. Field, J. D. Kelly, D. E. Neal, Y. Maehara, B. A. Ponder, Y. Nakamura and R. Hamamoto, *Int. J. Cancer*, 2011, **128**, 562–573.
10 I. Goulet, G. Gauvin, S. Boisvenue and J. Cote, *J. Biol. Chem.*, 2007, **282**, 33009–33021.
11 C. Y. Ou, M. J. LaBonte, P. C. Manegold, A. Y. So, I. Ianculescu, D. S. Gerke, K. R. Yamamoto, R. D. Ladner, M. Kahn, J. H. Kim and M. R. Stallcup, *Mol. Cancer Res.*, 2011, **9**, 660–670.
12 H. Hong, C. Kao, M. H. Jeng, J. N. Eble, M. O. Koch, T. A. Gardner, S. Zhang, L. Li, C. X. Pan, Z. Hu, G. T. MacLennan and L. Cheng, *Cancer*, 2004, **101**, 83–89.
13 Y. R. Kim, B. K. Lee, R. Y. Park, N. T. Nguyen, J. A. Bae, D. D. Kwon and C. Jung, *BMC Cancer*, 2010, **10**, 197.
14 L. Wang, S. Pal and S. Sif, *Mol. Cell. Biol.*, 2008, **28**, 6262–6277.
15 R. A. Copeland, M. E. Solomon and V. M. Richon, *Nat. Rev. Drug Discovery*, 2009, **8**, 724–732.
16 S. S. Wolf, *Cell. Mol. Life Sci.*, 2009, **66**, 2109–2121.
17 X. Zhang and X. Cheng, *Structure*, 2003, **11**, 509–520.
18 X. Zhang, L. Zhou and X. Cheng, *EMBO J.*, 2000, **19**, 3509–3519.
19 W. W. Yue, M. Hassler, S. M. Roe, V. Thompson-Vale and L. H. Pearl, *EMBO J.*, 2007, **26**, 4402–4412.
20 J. S. Sack, S. Thieffine, T. Bandiera, M. Fasolini, G. J. Duke, L. Jayaraman, K. F. Kish, H. E. Klei, A. V. Purandare, P. Rosettani, S. Troiani, D. Xie and J. A. Bertrand, *Biochem. J.*, 2011, **436**, 331–339.
21 S. Antonysamy, Z. Bonday, R. M. Campbell, B. Doyle, Z. Druzina, T. Gheyi, B. Han, L. N. Jungheim, Y. Qian, C. Rauch, M. Russell, J. M. Sauder, S. R. Wasserman, K. Weichert, F. S. Willard, A. Zhang and S. Emtage, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 17960–17965.
22 H. L. Rust, C. I. Zurita-Lopez, S. Clarke and P. R. Thompson, *Biochemistry*, 2011, **50**, 3332–3345.
23 T. M. Lakowski, P. 't Hart, C. A. Ahem, N. I. Martin and A. Frankel, *ACS Chem. Biol.*, 2010, **5**, 1053–1063.
24 P. 't Hart, T. M. Lakowski, D. Thomas, A. Frankel and N. I. Martin, *ChemBioChem*, 2011, **12**, 1427–1432.
25 P. 't Hart, D. Thomas, R. van Ommeren, T. M. Lakowski, A. Frankel and N. I. Martin, *Med. Chem. Commun.*, 2012, **3**, 1235–1244.
26 D. Thomas, T. Koopmans, T. M. Lakowski, H. Kreinin, M. I. Vhuiyan, S. A. Sedlock, J. M. Bui, N. I. Martin and A. Frankel, *ChemBioChem*, 2014, **15**, 1607–1613.
27 S. R. Daigle, E. J. Olhava, C. A. Therkelsen, C. R. Majer, C. J. Sneeringer, J. Song, L. D. Johnston, M. P. Scott, J. J. Smith, Y. Xiao, L. Jin, K. W. Kuntz, R. Chesworth, M. P. Moyer, K. M. Bernt, J. C. Tseng, A. L. Kung, S. A. Armstrong, R. A. Copeland, V. M. Richon and R. M. Pollock, *Cancer Cell*, 2011, **20**, 53–65.
28 C. Lerner, B. Masjost, A. Ruf, V. Gramlich, R. Jakob-Roetne, G. Zurcher, E. Borroni and F. Diederich, *Org. Biomol. Chem.*, 2003, **1**, 42–49.
29 J. Dowden, W. Hong, R. V. Parry, R. A. Pike and S. G. Ward, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 2103–2105.
30 J. Dowden, R. A. Pike, R. V. Parry, W. Hong, U. A. Muhsen and S. G. Ward, *Org. Biomol. Chem.*, 2011, **9**, 7814–7821.
31 R. Ragno, S. Simeoni, S. Castellano, C. Vicidomini, A. Mai, A. Caroli, A. Tramontano, C. Bonaccini, P. Trojer, I. Bauer, G. Brosch and G. Sbardella, *J. Med. Chem.*, 2007, **50**, 1241–1253.
32 Y. Feng, M. Li, B. Wang and Y. G. Zheng, *J. Med. Chem.*, 2010, **53**, 6028–6039.
33 C. Wang, Y. Zhu, J. Chen, X. Li, J. Peng, Y. Zou, Z. Zhang, H. Jin, P. Yang, J. Wu, L. Niu, Q. Gong, M. Teng and Y. Shi, *PLoS One*, 2014, **9**, e87267.
34 M. Allan, S. Manku, E. Therrien, N. Nguyen, S. Styhler, M. F. Robert, A. C. Goulet, A. J. Petschner, G. Rahil, A. Robert Macleod, R. Deziel, J. M. Besterman, H. Nguyen and A. Wahhab, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 1218–1223.
35 T. Huynh, Z. Chen, S. Pang, J. Geng, T. Bandiera, S. Bindi, P. Vianello, F. Roletto, S. Thieffine, A. Galvani, W. Vaccaro, M. A. Poss, G. L. Trainor, M. V. Lorenzi, M. Gottardis, L. Jayaraman and A. V. Purandare, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 2924–2927.
36 D. H. Cheng, S. Valente, S. Castellano, G. Sbardella, R. Di Santo, R. Costi, M. T. Bedford and A. Mai, *J. Med. Chem.*, 2011, **54**, 4928–4932.
37 C. A. Mooney, S. A. Johnson, P. t Hart, L. Quarles van Ufford, C. A. de Haan, E. E. Moret and N. I. Martin, *J. Med. Chem.*, 2014, **57**, 3154–3160.
38 M. Pignot, G. Pljevaljcic and E. Weinhold, *Eur. J. Org. Chem.*, 2000, 549–555.
39 A. M. Jawalekar, N. Meeuwenoord, J. G. O. Cremers, H. S. Overkleeft, G. A. van der Marel, F. P. J. T. Rutjes and F. L. van Delft, *J. Org. Chem.*, 2008, **73**, 287–290.
40 Y. Kobayashi, T. Shinozuka and O. Kanno, 1-Methylcarbapenem Derivatives, *Patent US* 2004014962 (A1), 2004, Sankyo Company Ltd.
41 D. D. Long, T. J. Church, J. R. Jacobsen, L. Jiang, D. R. Saito, I. Stergiades, P. M. Van Dyke, S. Dalziel and L. M. Preza, 8-Azabicyclo[3.2.1]octane compounds as mu opioid receptor antagonists, *Patent US* 2007219278 (A1), 2007, Theravance Inc.
42 K. C. Tang, R. Mariuzza and J. K. Coward, *J. Med. Chem.*, 1981, **24**, 1277–1284.
43 J. Fuentes, J. M. Illangua, F. J. Sayago, M. Angulo, C. Gasch and M. A. Pradera, *Tetrahedron: Asymmetry*, 2004, **15**, 3783–3789.
44 P. Ciuffreda, A. Loseto and E. Santaniello, *Tetrahedron*, 2002, **58**, 5767–5771.
45 R. Chesworth, K. W. Kuntz, E. J. Olhava and M. A. Patane, Modulators of histone methyltransferase, and methods of

use thereof, *Patent WO* 2012082436 (A2), 2012, Epizyme Inc.

46 M. Vedadi, D. Barsyte-Lovejoy, F. Liu, S. Rival-Gervier, A. Allali-Hassani, V. Labrie, T. J. Wigle, P. A. DiMaggio, G. A. Wasney, A. Siarheyeva, A. P. Dong, W. Tempel, S. C. Wang, X. Chen, I. Chau, T. J. Mangano, X. P. Huang, C. D. Simpson, S. G. Pattenden, J. L. Norris, D. B. Kireev, A. Tripathy, A. Edwards, B. L. Roth, W. P. Janzen, B. A. Garcia, A. Petronis, J. Ellis, P. J. Brown, S. V. Frye, C. H. Arrowsmith and J. Jin, *Nat. Chem. Biol.*, 2011, **7**, 648–648.

47 T. Mosmann, *J. Immunol. Methods*, 1983, **65**, 55–63.

48 E. Krieger, T. Darden, S. B. Nabuurs, A. Finkelstein and G. Vriend, *Proteins*, 2004, **57**, 678–683.

49 X. Zhang, L. Zhou and X. D. Cheng, *EMBO J.*, 2000, **19**, 3509–3519.

50 A. S. Konagurthu, J. C. Whisstock, P. J. Stuckey and A. M. Lesk, *Proteins*, 2006, **64**, 559–574.

51 E. Krieger, G. Koraimann and G. Vriend, *Proteins: Struct., Funct., Genet.*, 2002, **47**, 393–402.

52 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455–461.

