Organic & Biomolecular Chemistry

Dynamic Article Links

Cite this: *Org. Biomol. Chem.*, 2011, **9**, 7814

www.rsc.org/obc

PAPER

Small molecule inhibitors that discriminate between protein arginine *N*-methyltransferases PRMT1 and CARM1†

James Dowden,*[a] Richard A. Pike,[a] Richard V. Parry,[b] Wei Hong,[a] Usama A. Muhsen[a] and Stephen G. Ward[b]

Received 6th July 2011, Accepted 23rd August 2011
DOI: 10.1039/c1ob06100c

Protein arginine *N*-methyltransferases (PRMTs) selectively replace N–H for N–CH_3 at substrate protein guanidines, a post-translational modification important for a range of biological processes, such as epigenetic regulation, signal transduction and cancer progression. Selective chemical probes are required to establish the dynamic function of individual PRMTs. Herein, model inhibitors designed to occupy PRMT binding sites for an arginine substrate and *S*-adenosylmethionine (AdoMet) co-factor are described. Expedient access to such compounds by modular synthesis is detailed. Remarkably, biological evaluation revealed some compounds to be potent inhibitors of PRMT1, but inactive against CARM1. Docking studies show how prototype compounds may occupy the binding sites for a co-factor and arginine substrate. Overlay of PRMT1 and CARM1 binding sites suggest a difference in a single amino acid that may be responsible for the observed selectivity.

Introduction

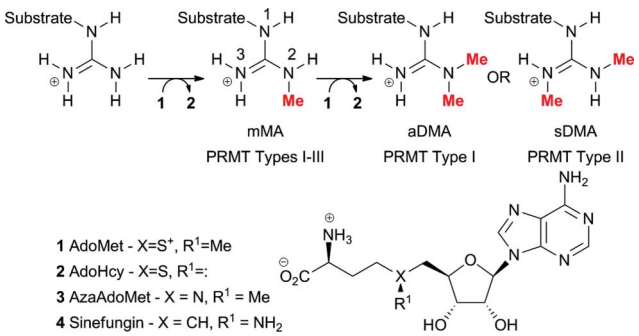
Chemical modulators that address individual protein arginine methyltransferases (PRMTs, EC 2.1.1.125) have potential applications for elucidating their dynamic contribution to cell biology, or informing the development of novel therapeutics.[1,2]

PRMTs co-localise the versatile cellular co-factor *S*-adenosyl methionine (AdoMet, **1**, Scheme 1) and guanidine functionality of the arginine-bearing target substrate to achieve site-specific guanidine *N*-methylation,[3] leading to monomethylarginine (mMA). Depending on the PRMT type, further methylation occurs at either the same (N^2), or adjacent (N^3) nitrogen, to furnish asymmetric or symmetrical dimethylarginine substitution (aDMA or sDMA) respectively (Scheme 1).[4]

Irregular PRMT activity and arginine methylation patterns have emerged as biomarkers for cardiovascular conditions,[5,6] viruses,[7] immune disease[8] and cancer.[9] For example, aggressive prostate and breast cancers cells both show over-expression of PRMT 4 (also known as co-factor associated arginine methyltransferase, CARM1).[10–12]

Chemical interference has great potential for learning about PRMT functions and correcting aberrant processes mediated by them.[1,2,13] PRMT-mediated peptide alkylation by an AdoMet derived nitrogen mustard gave potent inhibitors within the adduct mixture.[14] Recently, peptides featuring a reactive chloroacetamidine were demonstrated to give selective covalent inhibition of PRMT1, but not PRMT3 or CARM1.[15,16] Peptides incorporating N^{η}-alkylated arginines have been shown to inhibit PRMTs 1 and 6 about 10-fold more strongly than CARM1.[17] The same group recently described further peptide inhibitors featuring N^{η}-arginine substituents resembling the homocysteine part of **1**.[18]

In principle, small molecules may be readily adapted for *in vivo* studies. The first reported PRMT inhibitor 'AMI-1',[19] inspired reports of several related suramin-like analogues;[20,21] most recently, the sulfonate functionality of AMI-1 has been replaced by carboxylates while maintaining PRMT inhibition.[22] Virtual screening[23] and hit-to-lead studies have revealed a number of potential inhibitor scaffolds from which pyrazole[24–27] and benzo[d]imidazole[28] derivatives represent the most potent and selective inhibitors so far. Interestingly, AdoHcy induced organisation of the binding site was recently confirmed to be essential for pyrazole compounds to achieve potency.[29]



Scheme 1 Regiochemical classification of PRMTs and the structure of co-substrate AdoMet **1**, co-product AdoHcy **2** and inhibitors **3** and **4**.

[a]*School of Chemistry, University of Nottingham, University Park, Nottingham, UK, NG7 2RD. E-mail: james.dowden@nottingham.ac.uk; Fax: +44 115 9513565; Tel: +44 115 9513566*
[b]*Department of Pharmacy and Pharmacology, University of Bath, Claverton Down, Bath, UK, BA2 7AY*

† Electronic supplementary information (ESI) available: characterisation data for novel precursors, copies of NMR spectra and LC-MS data for **16–24**; table of measured distances for docked structures and images for docked structures of **19** and **20**. See DOI: 10.1039/c1ob06100c

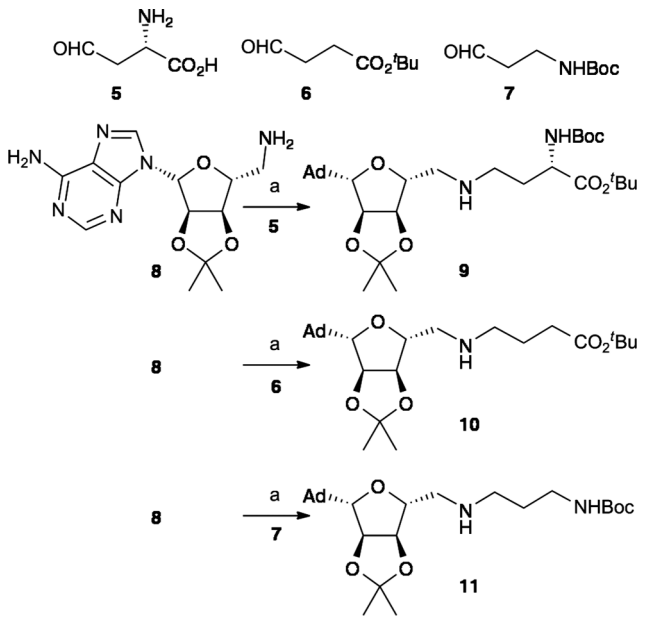
Small molecules that target both AdoMet and arginine binding sites offer an excellent platform to design selective PRMT inhibitors. We recently reported protoype PRMT1 inhibitors based on the latter principle.[30] Herein, we report comparison of a wider panel of inhibitiors against PRMT1 and CARM1. Remarkably, some of these compounds discriminate between PRMT1 and CARM1. Docking studies and an overlay of crystal structures that form a model that may explain the observed selectivity are also reported.

Results and discussion

Chemistry

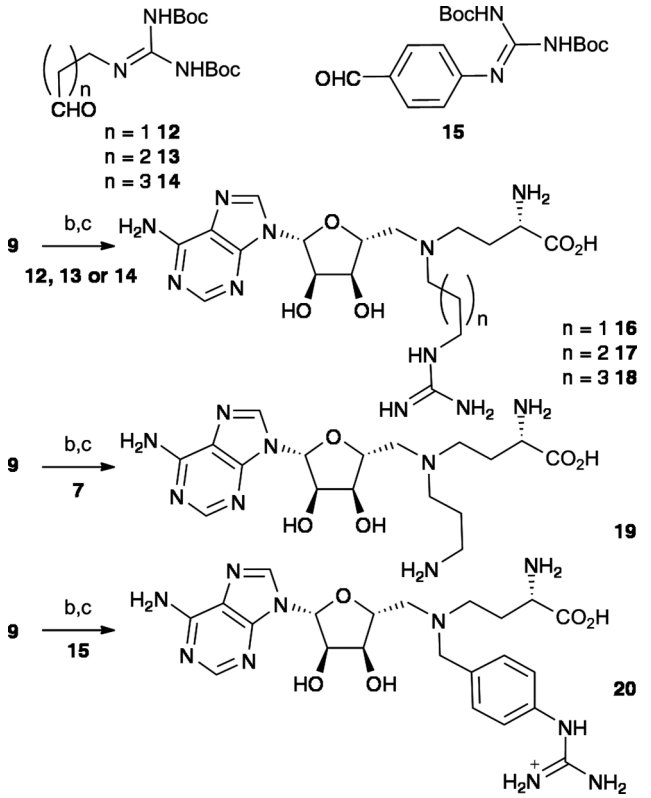
The inhibitor design extends non-reactive AdoMet analogues, such as AzoAdoMet **3**, to display functionality that can occupy the arginine binding site. In principle inhibitor specificity may be introduced by accessing non-covalent interactions with residues unique to the individual PRMTs.[29,31] A modular, three-step synthetic strategy was developed to give expedient access to AzoAdoMet analogues from readily available aldehydes.

Reductive alkylation of acetonide-protected 5′-amino-5′-deoxyadenosine **8**[32] with aldehyde **5**, derived from aspartic acid,[33] gave the AzoAdoHcy **9** in good yield (Scheme 2).[30] Compounds bearing *tert*-butyl butyrate **10**, or *N*-boc propylamine **11**, side chains were made from aldehydes **6**[34] and **7**[35] respectively.



Scheme 2 *Reagents and conditions*: a, $NaBH(OAc)_3$, $ClCH_2CH_2Cl$, rt, 73% (**9**),[30] 23% (**10**), 45% (**11**).

Elaboration of intermediate secondary amine **9** by further reductive alkylation with aldehydes **12**,[36] **13**,† **14**† and global deprotection gave guanidines **16–18** as previously described (Scheme 3).[30] Equivalent reactions with carbamate **7** gave analogue **19**, featuring a primary amine rather than a guanidine group after deprotection. Aldehyde **15**† was subject to the same sequence to give aryl guanidine **20**. Amines **10** and **11** were similarly reacted with guanidine **12** to give carboxylic acid **21** or amine **22**, respectively (Scheme 4). The symmetrical diamine **23** was obtained



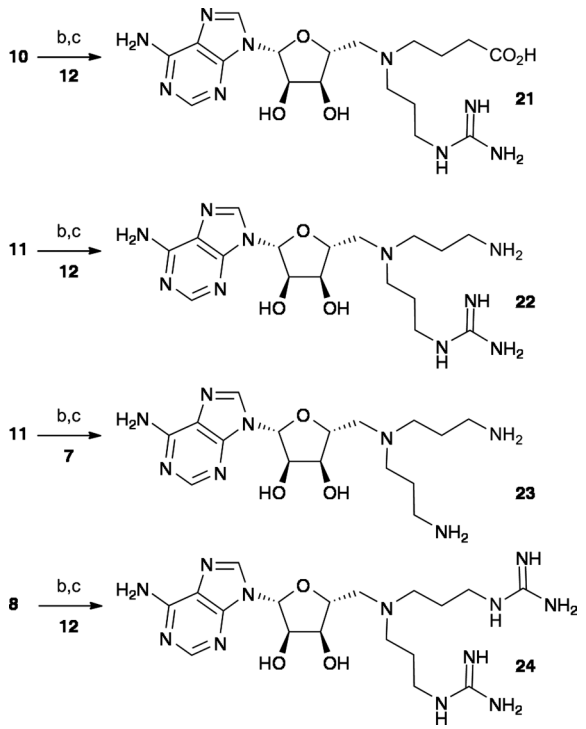
Scheme 3 *Reagents and conditions*: b, $NaBH(OAc)_3$, $ClCH_2CH_2Cl$, rt; c, TFA/H_2O, rt, Amberlyst IRA-400(Cl^-) (**16–18**),[30] 71% (**19**), 71% (**20**).

by further incorporation of amine **11** with carbamate **7**; the related symmetrical guanidine analogue **24** was similarly prepared using aldehyde **12**.[36]

In all cases, the global protecting group ablation proceeded smoothly in aqueous trifluoroacetic acid, and the concentrated products were subjected to ion exchange using Amberlyst IRA-400(Cl^-) and freeze drying to deliver the inhibitors **16–25** as hygroscopic hydrochloride salts.

Biochemical evaluation

Previously, the natural product sinefungin **4** was confirmed to be a potent but promiscuous inhibitor that inhibits PRMT1, CARM1 and SET7 with similar potency (IC_{50} ~ 1–3 μM),[30] thus providing a positive control for inhibition. We also previously reported a propargyl analogue of AzoAdoMet (not shown, X = N, R = $CH_2C{\equiv}CH$, Fig. 1 for general structure) as a negative control that does not inhibit either methyltransferase, indicating that additional functionality is required for inhibition.[30] PRMT1 was inhibited by analogues **16–18** bearing guanidine functionality regardless of the length of the alkyl linker, but the lysine methyltransferase SET7 was not inhibited, with at least 50% of enzyme activity remaining at 100 μM inhibitor concentrations. Indeed, none of the compounds reported herein inhibit the lysine methyltransferase.[37] Evaluation of the same inhibitors against CARM1 provided interesting new results. The analogue bearing the guanidine *via* a three carbon linker, **16**, inhibited CARM1 to approximately the same degree as PRMT1 (IC_{50} ~ 13.3 ± 8.7 μM). Remarkably, the compounds that display guanidine *via* a four or five carbon linker, **17** and **18** respectively, were very poor



Scheme 4 *Reagents and conditions*: b, $NaBH(OAc)_3$, $ClCH_2CH_2Cl$, rt; c, TFA/H_2O, rt, then Amberlyst IRA-400(Cl^-) 79% (**21**), 35% (**22**), 75% (**23**), 54% (**24**).

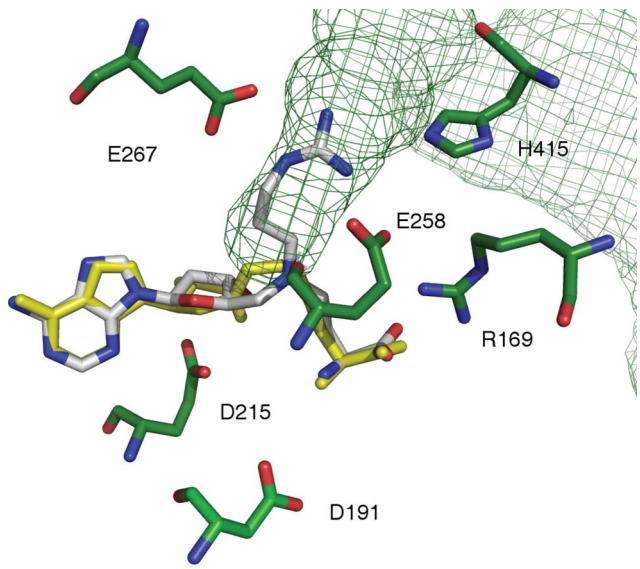


Fig. 1 Result of docking inhibitor **16** (white backbone) with CARM1 binding site (2Y1X.pdb, selected residues green backbone), also AdoHcy **2** (yellow), substrate arginine channel (green mesh).

inhibitors of CARM1, thus displaying remarkable discrimination between these enzymes. Interestingly, analogue **19**, that replaces the guanidine with an amine, expected to protonate under the assay conditions, was also a potent inhibitor of PRMT1, but not CARM1.

We next set out to investigate the effect of structural modifications to the prototype inhibitor design, specifically with a view to reduce the overall polarity of the compound, with a longer term goal of achieving greater cell permeability. Analogue **21**, that retained the carboxyl group but not the primary amine, did not inhibit any methyltransferases. On the other hand, the complementary analogue **22**, that retained the primary amine but not the carboxylic acid, inhibited PRMT1 and CARM1 with similar IC_{50} values (PRMT1 2.2 ± 1.8 μM, CARM1 3.3 ± 3.0 μM) to the prototype inhibitors. The latter analogue presents amine and guanidine functional groups *via* propyl linkers, both of which will be positively charged under the assay conditions. In principle, symmetrical bis-amine **23** and bis-guanidine **24** derived compounds might be expected to behave in a similar manner and indeed both compounds inhibit PRMT1 and CARM1 with similar IC_{50} values (1.1 ± 1.3 and 2.1 ± 2.3 μM for PRMT1 and 1.5 ± 1.5 and 4.6 ± 6.3 μM for CARM1, respectively, see Table 1).

Table 1 Inhibition constants for PRMT1, CARM1, and SET7

Compound Number	PRMT1[a,b]	CARM1[a,c]	SET7 [a,c]
Sinefungin **4**	0.7 ± 0.3	2.4 ± 3.6	[2.5][30]
16	6.2 ± 3.8[30]	13.3 ± 8.7	>100[d]
17	2.9 ± 0.8[30]	>100[d]	>100[d]
18	5.6 ± 4.5[30]	—[e]	—[e]
19	3.9 ± 1.8	—[e]	—[e]
20	14.5 ± 4.1	—[e]	—[e]
21	>100[d]	—[e]	—[e]
22	2.2 ± 1.8	3.3 ± 3.0	—[e]
23	1.5 ± 1.5	1.1 ± 1.3	—[e]
24	2.1 ± 2.3	4.6 ± 6.3	—[e]

[a] Mean IC_{50} ± standard deviation (μM) based on >triplicate experiments. [b] Substrate = P3, the arginine glycine rich region of Src-associated protein of 68 kDa (Sam68). [c] Substrate = Histone H3. [d] >50% enzyme activity at 100 μM. [e] No activity.

At present, the prototype compounds are somewhat flexible, which allows them to adopt a range of conformations that might fit a range of target enzymes in principle, thus, increasing rigidity might deliver greater specificity in future. As a tentative step toward exploring more rigid inhibitors, the AzoAdoMet analogue featuring a benzylic linker **20** was synthesised; subsequent evaluation revealed that this compound appeared to inhibit PRMT1 (IC_{50} 5.6 ± 4.7 μM) and not CARM1 in a similar way to the related analogues featuring four and five carbon flexible linkers.

Docking studies

Docking studies were performed to examine how the inhibitors (*i.e.* **16**) might be accommodated within a protein arginine methyltransferase active site. Unfortunately, PRMT1 crystal structures (*e.g.* 1or8.pdb)[38] are inappropriate for docking studies,[20,23] since they were obtained at a pH at which the enzyme is not active and are also disordered at the *N*-terminus. A recent crystal structure of CARM1 featuring AdoHcy **2** and a potent indole inhibitor simultaneously bound at the active site has been described (2Y1X.pdb).[29] Unlike previously reported CARM1 structures, an additional helix (αW) that interacts with the adenine group of AdoHcy **2** was observed.

Flexible docking of **16** to a rigid CARM1 binding site from this structure was carried out using AutoDock Vina[39] within the PyRX virtual screening tool.[40]

The most stable predicted binding mode revealed excellent overlap between the docked position of prototype inhibitor **16** and that of AdoHcy **2** in the native crystal structure (Fig. 1).

For example, interactions between the aminoacid component with arginine R169 and aspartic acid D191 are maintained. These models do not provide an obvious explanation for the difference in inhibition observed for compounds featuring only carboxylate **21** or amine **22** functionality as yet. Recent studies of PRMT1 mutants suggest that the arginine (R54) equivalent to R169 is not essential for binding of AdoMet, however.[41]

Most notable is that the remaining guanidine functionality of **16** extends into the channel believed to accommodate the substrate arginine during catalysis. Furthermore, this guanidine is well placed to gain optimal interactions with glutamic acid residues E258 and E267 reported to be essential for binding and catalysis respectively.[38,41] In the docked structure, the distance between the guanidinium and these residues is ~3.2 Å.† This analysis supports the original design principle and will be a valuable tool for planned virtual screening.

The most stable binding orientation predicted for the four and five methylene linked compounds **17** (not shown) and **18** (Fig. 2), are very similar to that for **16**, but the longer linkers push the guanidinium functional group beyond the glutamic acid residue (E258), to ~5.9 Å and ~6.9 Å respectively,† which may explain the reduced CARM1 inhibition observed for these compounds (Fig. 2). Overlay of the structure for PRMT1 with the docked structure of **18** with CARM1 reveals different residues near the substrate binding channel. While this does not represent a docked structure with PRMT1, a hypothesis can be formed assuming that the binding sites are similar. Specifically, PRMT1 features a glutamic acid residue (Glu47) close to the putative substrate channel, whereas CARM1 features an asparagine residue (Asn162). Thus, electrostatic interactions between the guanidine of **17** and **18** and the carboxylate of Glu47, within approximately 4.4 Å and 3.2 Å in the overlay structure,† may facilitate inhibition of PRMT1 by **17** and **18**.

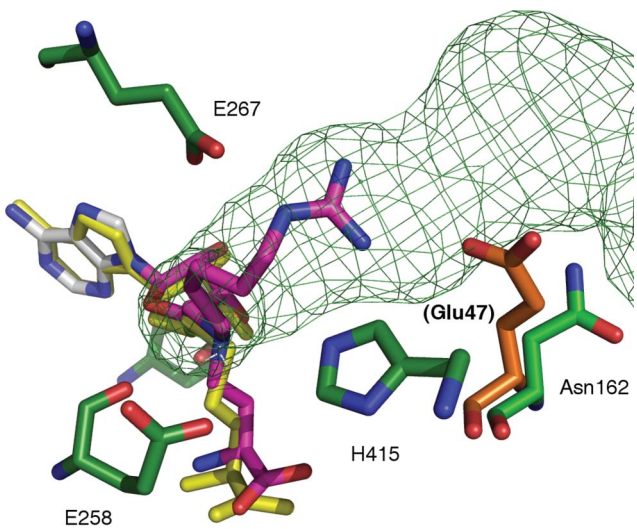


Fig. 2 Result of docking inhibitor **18** (magenta backbone) with CARM1 binding site (2Y1X.pdb, selected residues green backbone), also AdoHcy **2** (yellow), substrate arginine channel (green mesh). PRMT1 residue superimposed, Glu47 (orange, bold in parentheses). Note: different angle with respect to Fig. 1 to emphasise proximity.

The selective inhibition of PRMT1, but not CARM1 by the analogue displaying the benzylic guanidine **20** may also be explained by this model since the glutamic acid (E258) is ~8.6 Å away in CARM1, but the glutamic acid of PRMT1 (Glu47) is ~4.5 Å away in the overlay structure.† PRMT3 also features a carboxylate at this position (Glu229),[42] thus it will be interesting to learn whether these compounds inhibit this enzyme in future.

Analogue **19**, with an amine in place of guanidine compared to compound **16**, docked with a best fit binding mode of similar orientation to the latter compound in the CARM1 binding site.† It is less obvious from this model of binding as to how selective inhibition of PRMT1 but not CARM1 occurs for this compound.

As mentioned above, significant changes to the structure of CARM1 occur upon binding AdoHcy **2**.[43] Indeed this reorganisation is required for some inhibitors to achieve potency.[29] Presumably, compounds based on an extended AdoHcy template, especially **16** and **22–24**, are able induce fit to the CARM1 binding site. Molecular dynamics calculations that account for the motion of individual residues and domains are desirable, but will also require much more processor time. Future crystallographic and NMR binding studies will be valuable for revealing information about the structure and dynamics of non-covalent interactions between protein and these inhibitors.

Conclusions

These results demonstrate a common scaffold that can be readily adapted to inhibit individual PRMTs. Small molecules designed to occupy PRMT binding sites for both arginine substrate and AdoMet co-factor have been demonstrated as inhibitors with potency comparable to sinefungin. Remarkably, selective discrimination between individual PRMTs was observed, despite relatively simple variations between the prototype inhibitor structures. Confirmation of the bisubstrate mode of inhibition will require more detailed binding and kinetic studies which are the subject of ongoing studies. In the meantime, *in silico* docking studies identify substantial overlap with the AdoMet for **16** with the linker-supported guanidinium functional group extending into the arginine substrate binding site as designed. The overlay of PRMT1 (1or8.pdb)[38] and CARM1 (2Y1X.pdb)[29] crystal structures highlighted differences in a key residue corresponding to Glu47 within the PRMT1 binding site, Asn162 within CARM1, that may be reached by the longer linkers **17** and **18**, but not the three carbon linkers **16** or **19**. The modular synthesis demonstrated herein is well suited to the efficient production of further analogues. Studies to achieve selective inhibition by exploiting further sequence variations between various PRMT binding sites are under way.

Experimental

Chemistry: general

All chemicals were purchased from commercial sources and used as supplied unless stated. THF was freshly distilled over sodium benzophenone ketyl, Et_3N and CH_2Cl_2 were distilled from CaH_2. Thin layer chromatography was performed using silica gel 60 precoated aluminium plates (0.20 mm thickness) from Macherey-Nagel, with visualisation by UV light (254 nm) or exposure to potassium permanganate solution. Column chromatography was performed on silica gel (particle size 40–63 μm) from Fischer Chemicals. Melting points were determined using a Stuart SMP3

melting point apparatus and are uncorrected. Optical rotations were obtained at ambient temperature using a Jasco DIP-370 digital polarimeter. IR spectra were recorded in the solid phase using a Thermo AVATAR 320 FT-IR or from $CHCl_3$ using a Bruker Tensor 27 spectrometer. ^{1}H and ^{13}C NMR spectra (δ [ppm], J [Hz]) were recorded on Bruker AV400, DPX400 and AV500 spectrometers. Broadening of NMR signals was observed, presumably due to slow conformational exchange. This effect was most obvious in ^{13}C spectra and could be recovered by recording spectra at 70 °C on a JEOL EX270 spectrometer.† Mass spectra were recorded using a service LC-TOF (Bruker micrOTOF), running in an open-access mode. LC-MS data† were obtained using a Varian "Prostar" System comprising: Varian 410 autosampler, 2 × Varian 210 pump (binary system), Varian Polaris 5 μm C18-A 250 mm × 4.6 mm reverse phase column, Varian 325 UV detector monitoring at 254 nm and a Varian 310-MS TQ mass spectrometer operating in ESI mode. All runs were performed with a flow rate of 1 ml min^{-1} and eluted with the following linear gradient solvent system: t = 0 mins, 10% CH_3CN: 90% H_2O; t = 2 mins, 10% CH_3CN: 90% H_2O; t = 20 min, 100% CH_3CN: 0% H_2O; t = 28 min, 100% CH_3CN: 0% H_2O; t = 30 min, 10% CH_3CN: 90% H_2O.

General procedure for first reductive amination: method 1

5′-Amino-5′-deoxy-2′,3′-*O*,*O*-(1-methylethylidene)adenosine **8**[32] (1.2 mol eq.) and aldehyde **5**,[33] **6**[34] or **7**[35] (1.0 mol eq.) were suspended in $ClCH_2CH_2Cl$ (5 mL) at room temperature under Ar and the mixture gently heated with vigorous stirring to obtain a solution. $NaBH(OAc)_3$ (1.4 mol eq.) was slowly added portionwise and the reaction left for 2–4 h at room temperature under Ar. The reaction was then quenched by the addition of saturated aqueous Na_2CO_3 solution (5 mL). The mixture was extracted with CH_2Cl_2 (3 × 50 mL) and the combined organic layers were dried over $MgSO_4$ and concentrated *in vacuo* to yield the crude product. The product was purified by column chromatography on silica gel eluting with MeOH : CH_2Cl_2, 5 : 95 unless specified otherwise.

General procedure for second reductive amination: method 2

Secondary amines **9**,[30] **10** or **11** (1.0 mol eq.) and aldehydes **12**,[36] **13–15**† (1.1 mol eq.) with $MgSO_4$ (10 mol eq.) were used, then as method 1 above.

General procedure for deprotection: method 3

Intermediates† from method 2 above were dissolved in TFA (4 mL) and water (0.10 mL) and stirred overnight. The mixture was concentrated in vacuo, dissolved in water (5 mL) and washed with EtOAc (2 × 5 mL). The aqueous layer was concentrated *in vacuo* to ~1 mL then applied to a column of Amberlite IRA-400(Cl^-) ion exchange resin, eluted with water, and fractions containing product were freeze dried to give the hydrochloride salts **16–18**,[30] **19–24**.

5′-deoxy-5′-[[(4-(1,1-dimethylethoxy)-4-oxobutyl]amino]-2′,3′-*O*-(1-methylethylidene)-adenosine (10)

Method 1. White solid (0.59 g, 23%), m_p 49–51 °C (CH_2Cl_2); $[\alpha]_D^{19}$ −32.5 (c 1.05 in $CHCl_3$); ν_{max}/cm^{-1} 3011 (m), 2984 (m), 2938 (m), 1720 (s), 1633 (s), 1589 (m), 1475 (m), 1457 (m), 1424 (m), 1370 (s), 1330 (m); ^{1}H NMR δ_H (400 MHz; $CDCl_3$) 8.35 (1H, s, Ar*H*), 7.93 (1H, s, Ar*H*), 6.00 (1H, d, J = 2.8, 1′-*H*), 5.95 (2H, br s, adenosine-NH_2), 5.49-5.44 (1H, m, 2′-*H*), 5.08-5.03 (1H, m, 3′-*H*), 4.42–4.36 (1H, m, 4′-*H*), 3.00 (1H, br s, N*H*), 2.87–3.00 (2H, m, 5′-CH_2), 2.72–2.56 (2H, m, CH_2NH), 2.27 (2H, t, J = 7.3, COCH_2), 1.82–1.72 (2H, m, CH_2CH$_2$NH), 1.63 (3H, s, CH_3), 1.44 (9H, s, C(CH_3)$_3$), 1.40 (3H, s, CH_3); ^{13}C NMR δ_C (100 MHz; $CDCl_3$) 172.8 (C), 155.6 (C), 153.0 (CH), 149.4 (C), 139.9 (CH), 120.4 (C), 114.6 (C), 91.0 (CH), 85.3 (CH), 83.4 (CH), 82.2 (CH), 80.2 (C), 51.1 (CH_2), 49.0 (CH_2), 33.2 (CH_2), 28.1 (CH_3), 27.3 (CH_3), 25.4 (CH_3), 25.1 (CH_2); m/z (ESI) 449.3 (M + H$^+$, 100%) Found 449.2512 $C_{21}H_{33}N_6O_5$ requires 449.2507.

5′-deoxy-5′[(3-*tert*-butoxycarbonylaminopropyl)amino]-2′,3′-*O*-(1-methylethylidene)-adenosine (11)

Method 1. White solid (0.40 g, 45%), m_p 55–57 °C (CH_2Cl_2); $[\alpha]_D^{19}$ −10.5 (c 0.35 in $CHCl_3$); ν_{max}/cm^{-1} 3414 (w), 3305 (w), 2982 (m), 2931 (m), 2855 (w), 1707 (s), 1632 (s), 1588 (w), 1500 (m), 1473 (m), 1423 (w), 1369 (s); ^{1}H NMR δ_H (400 MHz; $CDCl_3$) 8.30 (1H, s, Ar*H*), 7.95 (1H, s, Ar*H*), 6.29 (2H, br s, adenosine-NH_2), 6.00 (1H, d, J = 3.1, 1′-*H*), 5.46 (1H, dd, J = 6.3, 3.1 2′-*H*), 5.33 (1H, br s, N*H*), 5.03 (1H, dd, J = 6.3, 3.3, 3′-*H*), 4.40-4.35 (1H, m, 4′-*H*), 3.25–3.08 (1H, br m, CONHCH_2), 2.96–2.84 (2H, m, 5′-CH_2), 2.78 (1H, s, N*H*), 2.74–2.66 (1H, m, CH$_2$CH_aH$_b$NH), 2.66–2.57 (1H, m, CH$_2$CH$_a$$H_b$NH), 1.68–1.61 (2H, m, C$H_2CH_2$NH), 1.62 (3H, s, C$H_3$), 1.40 (9H, s, C(C$H_3$)$_3$), 1.39 (3H, s, C$H_3$); ^{13}C NMR δ_C (100 MHz; $CDCl_3$) 156.2 (C), 155.8 (C), 153.1 (CH), 149.3 (C), 139.9 (CH), 120.1 (C), 114.7 (C), 90.9 (CH), 85.1 (CH), 83.4 (CH), 82.2 (CH), 79.1 (C), 51.2 (CH_2), 47.6 (CH_2), 39.0 (CH_2), 29.5 (CH_2), 28.4 (CH_3), 27.3 (CH_3), 25.4 (CH_3); m/z (ESI) 464.3 (M + H$^+$, 100%) Found 464.2611 $C_{21}H_{34}N_7O_5$ requires 464.2616; Calc. for $C_{21}H_{33}N_7O_5$.

5′-[(3(*S*)-3-amino-3-carboxypropyl)][(3-aminopropyl)amino]-5′-deoxy-adenosine (19)

Method 2, then 3. Pale yellow solid (0.19 g, 71% two steps), m_p 137–138 °C (dec.) (H_2O); $[\alpha]_D^{26}$ +24.2 (c 1.00 in H_2O); ν_{max}/cm^{-1} 2970 (m), 2920 (m), 1686 (s), 1611 (w), 1505 (m), 1412 (m), 1322 (w), 1229 (m), 1128 (m), 1049 (s); ^{1}H NMR δ_H (400 MHz; D_2O) 8.40 (2H, s, Ar*H*), 6.10 (1H, d, J = 3.7, 1′-*H*), 4.57–4.88 (1H, m, 2′-*H*), 4.41–4.49 (2H, m, 3′-*H* and 4′-*H*), 3.88–3.96 (1H, m, C*H*CH$_2$), 3.63–3.80 (2H, m, 5′-CH_2), 3.41–3.58 (2H, m, CHCH$_2$CH_2), 3.29–3.39 (2H, m, H$_2$NCH$_2$CH$_2$CH_2N), 2.95–3.05 (2H, m, H$_2$NCH_2CH$_2$CH$_2$N), 2.26–2.39 (1H, m, CHCH_aH$_b$), 2.12–2.23 (1H, m, CHCH$_a$$H_b$), 2.01–2.12 (1H, m, H$_2NCH_2CH_2CH_2$N); ^{13}C NMR δ_C (100 MHz; D_2O) 171.6 (C), 149.9 (C), 148.1 (C), 144.4 (CH), 143.5 (CH), 119.3 (C), 90.0 (CH), 78.0 (CH), 73.0 (CH), 71.5 (CH), 55.3 (CH_2), 51.5 (CH and CH_2), 50.4 (CH_2), 36.4 (CH_2), 24.5 (CH_2), 21.5 (CH_2); ^{13}C NMR δ_C (68 MHz; D_2O, 70 °C) 172.4, 150.9, 148.9, 145.4, 144.1, 120.0, 90.8, 78.9, 73.8, 72.3, 56.3, 52.9, 51.6, 37.3, 25.4, 22.4; m/z

(ESI) 425.2 (M + H^+, 100%), Found 425.2256 $C_{17}H_{29}N_8O_5$ requires 425.2255.

5′-[[(3(*S*)-3-amino-3-carboxypropyl]][(4-guanidinophenyl)-methyl]amino]-5′-deoxy-adenosine (20)

Method 2, then 3. Pale yellow solid (0.24 g, 71% two steps), m_p 138-139 °C (dec.) (H_2O); $[\alpha]_D^{28}$ +34.9 (*c* 1.00 in H_2O); ν_{max}/cm^{-1} 2973 (m), 2902 (m), 1682 (s), 1605 (w), 1577 (w), 1507 (w), 1408 (m), 1223 (m), 1055 (s), 925 (w), 853 (w), 828 (m), 744 (s); ^1H NMR δ_H (400 MHz; D_2O) 8.35 (1H, s, Ar*H*), 8.32 (1H, s, Ar*H*), 7.47 (2H, d, *J* = 8.4, Ar*H*), 7.24 (2H, d, *J* = 8.4, Ar*H*), 6.10 (1H, d, *J* = 3.2, 1′-*H*), 4.57–4.82 (1H, m, 2′-*H*), 4.37–4.52 (4H, m, 3′-*H*, 4′-*H* and ArCH_2N), 3.90 (1H, dd, *J* = 8.8, 4.4, C*H*CH_2), 3.61–3.71 (2H, m, 5′-CH_2), 3.45–3.61 (2H, m, $CHCH_2CH_2$), 2.31–2.44 (1H, m, $CHCH_aH_b$), 2.16–2.30 (1H, m, $CHCH_aH_b$); ^{13}C NMR δ_C (100 MHz; D_2O) 171.5 (C), 155.9 (C), 149.9 (C), 147.9 (C), 144.3 (CH), 143.4 (CH), 136.2 (C), 132.7 (CH), 127.6 (C), 125.4 (CH), 119.1 (C), 90.2 (CH), 78.0 (CH), 73.2 (CH), 71.6 (CH), 57.6 (CH_2), 54.7 (CH_2), 51.4 (CH), 51.3 (CH_2), 24.6 (CH_2); *m*/*z* (ESI) 559.2 (100%, [M + HCO_2]$^+$), 537.2 (38%, M + Na^+), 515.2 (31%, M + H^+) Found 515.2462 $C_{22}H_{31}N_{10}O_5$ requires 515.2473.

5′-[(3-carboxypropyl)][(3-guanidinopropyl)amino]-5′-deoxy-adenosine (21)

Method 2, then 3. Pale yellow solid (0.24 g, 79% two steps), m_p 115–6 °C (dec.) (H_2O); $[\alpha]_D^{28}$ +13.7 (*c* 1.00 in H_2O); ν_{max}/cm^{-1} 2986 (m), 2971 (m), 2935 (w), 1685 (s), 1626 (m), 1509 (w), 1409 (m), 1396 (m), 1321 (w), 1224 (m), 1126 (m), 1048 (s); ^1H NMR δ_H (400 MHz, D_2O) 8.41 (1H, s, Ar*H*), 8.40 (1H, s, Ar*H*), 6.11 (1H, d, *J* = 4.1, 1′-*H*), 4.56–4.87 (1H, m, 2′-*H*), 4.43–4.50 (2H, m, 3′-*H* and 4′-*H*), 3.71–3.81 (1H, m, 5′-CH_aH_b), 3.59–3.68 (1H, m, 5′-CH_aH_b), 3.13–3.33 (6H, m, $COCH_2CH_2CH_2$ and $NHCH_2CH_2CH_2$), 2.31–2.46 (2H, m, $COCH_2$), 1.81–2.02 (4H, m, $COCH_2CH_2CH_2$ and $NHCH_2CH_2CH_2$); ^{13}C NMR δ_C (100 MHz, D_2O) 176.5 (C), 156.7 (C), 150.1 (C), 148.0 (C), 144.7 (CH), 143.5 (CH), 119.3 (C), 90.1 (CH), 78.2 (CH), 72.9 (CH), 71.7 (CH), 54.7 (CH_2), 50.8 (CH_2 × 2), 38.0 (CH_2), 30.2 (CH_2), 22.6 (CH_2), 18.4 (CH_2); *m*/*z* (ESI) 452.2 (M + H^+, 100%) Found 452.2341 $C_{18}H_{30}N_9O_5$ requires 452.2364.

5′-deoxy-5′-[(3-guanidinopropyl)][(3-aminopropyl)amino]adenosine (22)

Method 2, then 3. Pale yellow solid (0.15 g, 35% two steps), m_p 115–116 °C (dec.) (H_2O); $[\alpha]_D^{26}$ +12.2 (*c* 1.10 in H_2O); ν_{max}/cm^{-1} (neat) 3149 (w), 3054 (w), 2989 (w), 1680 (s), 1624 (m), 1507 (w), 1419 (w), 1322 (w), 1226 (w), 1127 (m), 1053 (m); ^1H NMR δ_H (400 MHz; D_2O) 8.40 (1H, s, Ar*H*), 8.39 (1H, s, Ar*H*), 6.11 (1H, d, *J* = 3.5, 1′-*H*), 4.75–4.84 (1H, m, 2′-*H*), 4.39–4.51 (2H, m, 3′-*H* and 4′-*H*), 3.73–3.82 (1H, m, 5′-CH_aH_b), 3.62–3.70 (1H, m, 5′-CH_aH_b), 3.26–3.37 (4H, m, $H_2NCH_2CH_2CH_2N$ and $H_2NCH_2CH_2CH_2N$), 3.08–3.21 (2H, m, $C(N)NHCH_2CH_2CH_2N$), 2.96–3.04 (2H, m, $H_2NCH_2CH_2CH_2N$), 1.98–2.14 (2H, m, $C(N)NHCH_2CH_2CH_2N$), 1.81–1.98 (2H, m, $C(N)NHCH_2CH_2CH_2N$); ^{13}C NMR δ_C (100 MHz; D_2O) 156.7 (C), 150.1 (C), 148.0 (C), 144.7 (CH), 143.3 (CH), 119.3 (C), 90.2 (CH), 79.4 (CH), 73.1 (CH), 71.5 (CH), 54.8 (CH_2), 51.1 (CH_2), 38.0 (CH_2), 36.4 (CH_2), 33.3 (CH_2), 22.6 (CH_2), 21.7 (CH_2); *m*/*z* (ESI) 423.3 (M + H^+, 100%) Found 423.2576 $C_{17}H_{31}N_{10}O_3$ requires 423.2575.

5′-[di(3-aminopropyl)amino]-5′-deoxy-adenosine (23)

Method 2, then 3. Pale yellow solid (0.28 g, 75% two steps), m_p 110–111 °C (dec.) (H_2O); $[\alpha]_D^{23}$ +16.5 (*c* 1.00 in H_2O); ν_{max}/cm^{-1} 2987 (m), 2971 (m), 2901 (m), 1684 (w), 1406 (m), 1394 (m), 1231 (w), 1077 (s), 1066 (s), 1047 (s), 949 (w); ^1H NMR δ_H (400 MHz; D_2O) 8.42 (1H, s, Ar*H*), 8.41 (1H, s, Ar*H*), 6.14 (1H, d, *J* = 3.5, 1′-*H*), 4.71–4.90 (1H, m, 2′-*H*), 4.50 (1H, dd, *J* = 6.8, 5.4, 3′-*H*), 4.45 (1H, ddd, *J* = 9.2, 6.8, 2.4, 4′-*H*), 3.76 (1H, dd, *J* = 14.2, 9.2, 5′-CH_aH_b), 3.70 (1H, dd, *J* = 14.2, 2.4, 5′-CH_aH_b), 3.37 (4H, t, *J* = 8.4, $(H_2NCH_2CH_2CH_2)_2N$), 3.02 (4H, t, *J* = 7.7, $(H_2NCH_2CH_2CH_2)_2N$), 2.00–2.18 (4H, m, $(H_2NCH_2CH_2CH_2)_2N$); ^{13}C NMR δ_C (100 MHz; D_2O) 150.1 (C), 148.1 (C), 144.6 (CH), 143.2 (CH), 119.3 (C), 90.1 (CH), 77.5 (CH), 73.1 (CH), 71.4 (CH), 55.1 (CH_2), 50.7 (CH_2), 36.4 (CH_2), 21.5 (CH_2); *m*/*z* (ESI) 381.2 (M + H^+, 100%) Found 381.2362 $C_{16}H_{29}N_8O_3$ requires 381.2357.

5′-deoxy-5′-[di(3-guanidinopropyl)amino]adenosine (24)

Method 1, 2, then 3. Pale yellow solid (0.11 g, 54%), m_p 80–82 °C (H_2O); $[\alpha]_D^{23}$ +11.9 (*c* 1.00 in H_2O); ν_{max}/cm^{-1} 3671 (w), 2987 (s), 2970 (s), 2901 (s), 1665 (m), 1406 (m), 1394 (m), 1381 (m), 1249 (m), 1230 (m), 1074 (s), 1066 (s), 1052 (s), 1028 (s), 898 (w); ^1H NMR δ_H (400 MHz; D_2O) 8.44 (1H, s, Ar*H*), 8.43 (1H, s, Ar*H*), 6.14 (1H, d, *J* = 3.7, 1′-*H*), 4.59–4.75 (1H, m, 2′-*H*), 4.41–4.55 (1H, m, 3′-*H* and 4′-*H*), 3.81 (1H, dd, *J* = 14.2, 10.0, 5′-CH_aH_b), 3.61–3.74 (1H, m, 5′-CH_aH_b), 3.12–3.37 (8H, m, $NH_2CH_2CH_2CH_2N$ × 2), 1.84–2.05 (4H, m, $NH_2CH_2CH_2CH_2N$ × 2); ^{13}C NMR δ_C (100 MHz; D_2O) 156.7 (C), 150.0 (C), 148.0 (C), 144.5 (CH), 143.5 (CH), 119.3 (C), 90.3 (CH), 78.2 (CH), 73.1 (CH), 71.6 (CH), 54.7 (CH_2), 51.2 (CH_2), 38.1 (CH_2), 22.9 (CH_2); *m*/*z* (ESI) 465.3 (M + H^+, 100%) Found 465.2793 $C_{18}H_{33}N_{12}O_3$ requires 465.2793.

Enzymatic studies

Preparation of GST-linked fusion proteins. The coding sequences of required proteins were cloned into a pGEX-5X-1 expression vector (Amersham Pharmacia Biotech). GST fused protein was expressed in DH5α cells grown in LB-medium. 200 mL cultures with an A_{600} of 0.6 were induced with 0.5 mM IPTG and grown for 3 h at 30 °C. After centrifugation of cells at 4000*g*, the pellet was resuspended in 5 mL of GST-binding buffer (PBS, pH 7.3) containing one protease inhibitor tablet (Complete, Roche, Mannheim, Germany) for 50 mL of buffer. Cells were lysed using a high-pressure cell disruptor (Constant Cell Disruption Systems). The resulting lysate was centrifuged at 20000*g* for 20 min at 4 °C. GST fusion protein was purified from soluble extracts by binding to glutathione agarose beads for 2 h at 4 °C.

Preparation of enzyme substrates. P3 is a protein sequence encoding the arginine glycine rich region of Src-associated protein of 68 kDa (Sam68). For PRMT-1 methylation assays, GST-P3 substrate protein was released from the glutathione agarose beads by three incubations with 200 µL of 20 mM glutathione in PBS (pH7.4). After each hour rotating at 4 °C beads were sedimented by centrifugation in a microfuge at 1500 rpm for 5 min and

supernatant removed and aliquots retained. CARM1 and SET-7 enzyme assays used 1 mg mL^{-1} Histone H3 (Sigma Aldrich).

Inhibition assays. *In vitro* methylation reactions were performed according to published procedures.[19] Briefly, methylation reactions were carried out in the presence of [^{3}H]AdoMet (79 Ci $mmol^{-1}$ from a 12.6 μM stock solution in dilute HCl/ethanol 9 : 1, pH 2.0–2.5, Amersham Biosciences) and reaction buffer (200 mM NaCl, 20 mM Tris pH 8.0). Various concentrations of each compound were incubated with either GST-CARM1 or GST-SET-7 with histone H3 substrate, or GST-PRMT1 with GST-P3 substrate. Substrates (5 μg) were incubated with recombinant enzymes (5 μg) in the presence of 0.5 μM [^{3}H]AdoMet for 90 min at 30 °C in a final volume of 50 μL. Reactions were terminated by addition of 20 μL SDS-PAGE sample buffer and boiling for 5 min, and reaction product was resolved by 10% SDS-PAGE. Proteins were transferred to nitrocellulose membrane, and visualised with Ponceau S stain. Substrate proteins were excised and incorporated tritium quantified by scintillation counting (Perkin Elmer Tri-carb 2800TR). Results were plotted as % enzyme activity of control samples, and IC_{50} values were determined from at least three separate assays and are presented as mean values ± S.E.M.

Docking studies

AutoDock Vina (version 1.1.2)[39] was used within the PyRX virtual screening tool (version 0.8)[40] using default settings. Docking was performed at the binding site of CARM1 chain B (2Y1X.pdb)[29] with ligands removed. The receptor was treated as rigid within a 25 Å^3 box centred upon the location of the sulfur atom of SAH in the original structure. Default settings were used, which treated all bonds as rotatable except for the furan conformation which was fixed. Ligand files (*.pdbqt) were prepared for docking within PyRX from minimised 3D ligand co-ordinates (*.mol2) generated in OpenBabel (version 2.3.0)[44,45] from 2D ligands co-ordinates (*.mol) with chirality determined from parity flags. Graphical outputs, including overlay, were generated using PyMol (version 0.99);[46] pathways from the binding site were calculated using the CAVER plugin (version 2.1)[47] starting from the same co-ordinates as the sulfur atom of AdoHcy **2**.

Acknowledgements

Financial support from the Medical Research Council Grant ref: G0700840, the University of Nottingham (Studentship to W.H.) and The Higher Committee for Education Development Iraq (Studentship to UAM) is gratefully acknowledged. We thank Oreste Acuto (University of Oxford, UK) for the gift of GST-P3 plasmid DNA and helpful discussions. We thank Mark Bedford (University of Texas, MD Anderson Cancer Center, USA) for the gift of GST-PRMT1, GST-CARM1 and GST-SET-7 plasmid DNA.

Notes and references

1 R. A. Copeland, M. E. Solomon and V. M. Richon, *Nat. Rev. Drug Discovery*, 2009, **8**, 724–732.
2 P. A. Cole, *Nat. Chem. Biol.*, 2008, **4**, 590–597.
3 B. C. Smith and J. M. Denu, *Biochim. Biophys. Acta.*, 2009, **1789**, 45–57.
4 M. T. Bedford, *J. Cell Sci.*, 2007, **120**, 4243–4246.
5 A. O. Yildirim, P. Bulau, D. Zakrzewicz, K. E. Kitowska, N. Weissmann, F. Grimminger, R. E. Morty and O. Eickelberg, *Am. J. Respir. Cell Mol. Biol.*, 2006, **35**, 436–443.
6 X. Chen, F. Niroomand, Z. F Liu, A. Zankl, H. A. Katus, L. Jahn and C. P. Tiefenbacher, *Basic Res. Cardiol.*, 2006, **101**, 346–353.
7 M. C. Boulanger, C. Liang, R. S. Russell, R. T. Lin, M. T. Bedford, M. A. Wainberg and S. Richard, *J. Virol.*, 2005, **79**, 124–131.
8 F. Blanchet, A. Cardona, F. A. Letimier, M. S. Hershfield and O. Acuto, *J. Exp. Med.*, 2005, **202**, 371–377.
9 M. Yoshimatsu, G. Toyokawa, S. Hayami, M. Unoki, T. Tsunoda, H. I. Field, J. D. Kelly, D. E. Neal, Y. Maehara, B. A. J. Ponder, Y. Nakamura and R. Hamamoto, *Int. J. Cancer*, 2011, **128**, 562–573.
10 D. B. Seligson, S. Horvath, T. Shi, H. Yu, S. Tze, M. Grunstein and S. K. Kurdistani, *Nature*, 2005, **435**, 1262–1266.
11 (*a*) S. Majumder, Y. Liu, O. H. Ford, J. L. Mohler and Y. E. Whang, *Prostate*, 2006, **66**, 1292–1301; (*b*) H. Hong, C. H. Kao, M. H. Jeng, J. N. Eble, M. O. Koch, T. A. Gardner, S. B. Zhang, L. Li, C. X. Pan, Z. Q. Hu, G. T. MacLennan and L. Cheng, *Cancer*, 2004, **101**, 83–89.
12 S. El Messaoudi, E. Fabbrizio, C. Rodriguez, P. Chuchana, L. Fauquier, D. H. Cheng, C. Theillet, L. Vandel, M. T. Bedford and C. Sardet, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 13351–56.
13 E. M. Bissinger, R. Heinke, W. Sippl and M. Jung, *Med. Chem. Commun.*, 2010, **1**, 114–124.
14 T. C. Osborne, R. L. Roska, S. R. Rajski and P. R. Thompson, *J. Am. Chem. Soc.*, 2008, **130**, 4574–4575.
15 K. L. Bicker, O. Obianyo, H. L. Rust and P. R. Thompson, *Mol. BioSyst.*, 2011, **7**, 48–51.
16 O. Obianyo, C. P. Causey, T. C. Osborne, J. E. Jones, Y. H. Lee, M. R. Stallcup and P. R. Thompson, *ChemBioChem*, 2010, **11**, 1219–1223.
17 T. M. Lakowski, P. 't Hart, C. A. Ahern, N. I. Martin and A. Frankel, *ACS Chem. Biol.*, 2010, **5**, 1053–1063.
18 P. 't Hart, T. M. Lakowski, D. Thomas, A. Frankel and N. I. Martin, *ChemBioChem*, 2011, **12**, 1427–1432.
19 D. Cheng, N. Yadav, R. W. King, M. S. Swanson, E. J. Weinstein and M. T. Bedford, *J. Biol. Chem.*, 2004, **279**, 23892–23899.
20 Y. Feng, M. Y. Li, B. H. Wang and Y. G. Zheng, *J. Med. Chem.*, 2010, **53**, 6028–6039.
21 K. Bonham, S. Hemmers, Y. H. Lim, D. M. Hill, M. G. Finn and K. A. Mowen, *FEBS J.*, 2010, **277**, 2096–2108.
22 S. Castellano, C. Milite, R. Ragno, S. Simeoni, A. Mai, V. Limongelli, E. Novellino, I. Bauer, G. Brosch, A. Spannhoff, D. Cheng, M. T. Bedford and G. Sbardella, *ChemMedChem*, 2010, **5**, 398–414.
23 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.
24 A. V. Purandare, Z. Chen, T. Huynh, S. Pang, J. Geng, W. Vaccaro, M. A. Poss, J. O'Connell, K. Nowak and L. Jayaraman, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 4438–4441.
25 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. Déziel, J. M. Besterman, H. Nguyen and A. Wahhab, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 1218–1223.
26 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.
27 H. Wan, T. Huynh, S. Pang, J. Geng, 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**, 5063–5066.
28 E. Therrien, G. Larouche, S. Manku, M. Allan, N. Nguyen, S. Styhler, M. F. Robert, A. C. Goulet, J. M. Besterman, H. Nguyen and A. Wahhab, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 6725–6732.
29 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.
30 J. Dowden, W. Hong, R. V. Parry, R. A. Pike and S. G. Ward, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 2103–2105.
31 V. Campagna-Slater, M. W. Mok, K. T. Nguyen, M. Feher, R. Najmanovich and M. Schapira, *J. Chem. Inf. Model.*, 2011, **51**, 612–623.
32 A. M. Reeve and C. A. Townsend, *Tetrahedron*, 1998, **54**, 15959–15974.
33 R. M. Werner, O. Shokek and J. T. Davis, *J. Org. Chem.*, 1997, **62**, 8243–8246.
34 M. Kolb, J. Barth, J.-G. Heydt and M. J. Jung, *J. Med. Chem.*, 1987, **30**, 267–272.

35 S. Hourcade, A. Ferdenzi, P. Retailleau, S. Mons and C. Marazano, *Eur. J. Org. Chem.*, 2005, 1302–1310.
36 N. S. Freeman,Y. Tal-Gan, S. Klein, A. Levitzki and C. Gilon, *J. Org. Chem.*, 2011, **76**, 3078–3085.
37 S. Mori, K. Iwase, N. Iwanami, Y. Tanaka, H. Kagechika and T. Hirano, *Bioorg. Med. Chem.*, 2010, **18**, 8158–8166.
38 X. Zhang and X. D. Cheng, *Structure*, 2003, **11**, 509–520.
39 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455–461.
40 L. K. Wolf, *Chem. Eng. News*, 2009, **87**, 31.
41 H. L. Rust, C. I. Zurita-Lopez, S. Clarke and P. R. Thompson, *Biochemistry*, 2011, **50**, 3332–3345.
42 X. Zhang, L. Zhou and X. D. Cheng, *EMBO J.*, 2000, **19**, 3509–3519.
43 N. Troffer-Charlier, V. Cura, P Hassenboehler, D. Moras and J. Cavarelli, *EMBO J.*, 2007, **26**, 4391–4401.
44 R. Guha, M. T. Howard, G. R. Hutchison, P. Murray-Rust, H. Rzepa, C. Steinbeck, J. K. Wegner and E. Willighagen, *J. Chem. Inf. Model.*, 2006, **46**, 991–998.
45 The Open Babel Package, version 2.3.0 http://openbabel.sourceforge.net/ (accessed Jan 2011).
46 W. L. DeLano *The PyMOL Molecular Graphics System*. (2008) DeLano Scientific LLC, Palo Alto, CA, USA.
47 M. Petrek, M. Otyepka, P. Banás, P. Kosinová, J. Koca and J. Damborský, *BMC Bioinformatics*, 2006, **7**, 316.