REVIEW

Structural biology and chemistry of protein arginine methyltransferases

Matthieu Schapira*[ab] and Renato Ferreira de Freitas[a]

Protein arginine methyltransferases (PRMTs), an emerging target class in drug discovery, can methylate histones and other substrates, and can be divided into three subgroups, based on the methylation pattern of the reaction product (monomethylation, symmetrical or asymmetrical dimethylation). Here, we review the growing body of structural information characterizing this protein family, including structures in complex with substrate-competitive and allosteric inhibitors. We outline structural differences between type I, II and III enzymes and propose a model underlying class-specificity. We analyze the structural plasticity and diversity of the substrate, cofactor and allosteric binding sites, and propose that the conformational dynamics of PRMTs can be exploited towards the discovery of allosteric inhibitors that would antagonize conformationally active states.

Introduction

PRMTs catalyze the transfer of a methyl group from the cofactor *S*-adenosylmethionine (SAM) to the guanidinium group of arginine side-chains on nuclear and cytoplasmic proteins including histones. There are nine known PRMTs in humans, but the large number of putative methyltransferases suggests that this number may increase.[1,2] PRMTs can be divided into three classes: Type I (PRMT1, the major Type I PRMT, PRMT2, PRMT3, CARM1/PRMT4, PRMT6, and PRMT8) can mono- and asymmetrically di-methylate arginines; Type II (PRMT5) mono- and symmetrically di-methylate their substrates, and Type III (PRMT7) only mono-methylate the guanidinium group.[3] PRMT substrates are involved in multiple and diverse cellular mechanisms such as chromatin mediated signaling (PRMT1, PRMT2, CARM1, PRMT5, PRMT6 and PRMT7 methylate histone tails[4,5]), the ribosomal machinery (PRMT3 methylates the 40S ribosomal protein S2 (ref. 6)), and DNA repair or mRNA splicing[7]; PRMTs affect cell growth, proliferation and differentiation,[5] can be deregulated in cancer and other disease conditions, and are emerging as potential therapeutic targets.[8–10]

Domain architecture

The nine human PRMTs are organized around a conserved catalytic core composed of a Rossman fold, where the cofactor binds, and a β-barrel that contributes to substrate binding[11] (Fig. 1A). Protein–protein interaction modules are sometimes observed at the N-terminus of the PRMT core. These include an SH3 domain in PRMT2, a zinc finger in PRMT3, a pleckstrin homology (PH) domain in CARM1,[12] and a TIM barrel in PRMT5.[13] It is expected that these structural domains participate in substrate binding, recruit other proteins that are necessary for the formation of a catalytically active complex, or contribute to the homo-oligomerization of PRMT subunits. For instance, the TIM barrel of PRMT5 interacts with MEP50, a binding partner that is necessary for full PRMT5 catalytic activity, and largely contributes to the homo-tetramerization of the enzyme (discussed below).[13]

Structural overview of Type I and III PRMTs

Class I PRMT structures solved to date adopt a head-to-tail dimeric arrangement, where a dimerization arm projecting off of the β-barrel interacts with the Rossman fold of another subunit (Fig. 1B).[12,14–18] A dynamic α-helix at the N-terminus of the Rossman fold also contributes to inter-subunit contacts in the active state of rCARM1.[12] This important and conserved structural element is composed of 2 segments, α-X and α-Y that fold onto the bound cofactor, and participates in recruitment and proper positioning of the substrate peptide. As a result, the cofactor becomes buried in the core of the Rossman fold, and a substrate-binding site at the boundary of the Rossman fold, the β-barrel and the α-helix becomes clearly delineated (Fig. 1C). This suggests a mechanism where cofactor binding stabilizes a conformationally active state of the α-helix (Fig. 1D). Cofactor-binding alone may however not be sufficient as the α-helix was captured in a great diversity of conformational states in holo-structures of PRMT1, PRMT3, CARM1, and PRMT6.[12,18,19]

While no structure of mammalian PRMT7 was reported, the CePRMT7 structure reveals a variation on the head-to-tail dimerization theme observed in Type I PRMTs:[14] the N-terminal

[a]Structural Genomics Consortium, University of Toronto, Toronto, ON, M5G 1L7, Canada. E-mail: matthieu.schapira@utoronto.ca; Tel: +1-416-978-3092

[b]The Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON, M5S 1A8, Canada

This journal is © The Royal Society of Chemistry 2014
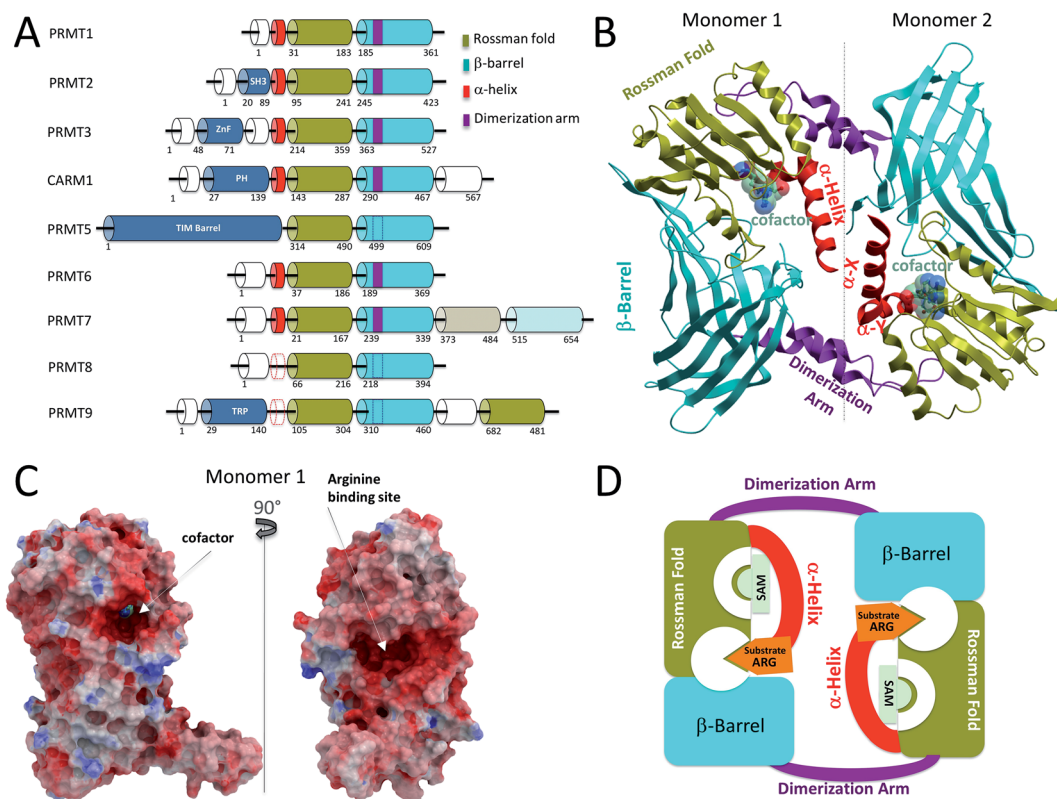
Med. Chem. Commun., 2014, **5**, 1779–1788 | **1779**

Fig. 1 Structural overview of Type I PRMTs. (A) Domain architecture of human PRMTs. The C-terminal Rossman fold and β-barrel of PRMT7, shown in lighter colors, have low sequence similarity with template sequences, and are catalytically inactive. Dashed lines indicate putative structural elements for which no structural data is available. (B) Canonical dimeric arrangement of Type I PRMTs, illustrated by the structure of CARM1 (PDB code 3B3F). In CePRMT7, the C-terminal catalytic core replaces the second subunit. (C) Connolly surface representation of CARM1 colored by electrostatic potential highlighting the cofactor and substrate binding sites. (D) Schematic representation illustrating the structural mechanism of Type I PRMTs: folding of the α-helix on the cofactor completes the formation of the substrate binding site.

catalytic core composed of a Rossman fold and a β-barrel adopts a canonical PRMT structure, but a quaternary arrangement places a C-terminal pseudo-catalytic core that is unique to PRMT7 in a position that mimics the second subunit in Type I PRMTs and in TbPRMT7.[14,16] An important pair of glutamic acids that are conserved at the catalytic sites of active PRMTs (detailed later) is missing from the C-terminal CePRMT7 core, and a loop partially occludes the cofactor pocket.

In depth structural analysis of rCARM1 by Troffer-Charlier demonstrated a profound structural rearrangement of the conserved Type I α-helix, which flips by 180° and partially converts into a β-strand in the apo conformation[12] (Fig. 2A and B). This dramatic motion should re-position the PH domain located immediately at the N-terminus of the α-helix to the other side of the CARM1 dimer, which probably necessitates dissociation of the dimer, and suggests a regulatory mechanism whereby cofactor binding and enzymatic activity correlate with positioning of the PH domain and its binding partners relative to the CARM1 catalytic core (Fig. 2B).

A catalytic Rossman fold is also observed in non-SET domain lysine methyltransferases,[20] DNA, RNA and small-molecule methyltransferases[21–23] (Fig. 2C). In all these structures, a secondary element located at the N- or C-terminus of the Rossman fold and positioned in the 3D space at the site of PRMTs' α-helix contributes to cofactor binding. This element is an α-helix in Type I and III PRMTs, as well as in DNMTs, and a loop in small-molecule, lysine, RNA methyltransferases, and PRMT5 (Fig. 2C). The structural plasticity of the corresponding loop was demonstrated in the case of lysine methyltransferase DOT1L: the size and hydrophobicity of the cofactor binding pocket are significantly increased when the loop adopts a conformationally inactive state, which is captured in complex with pico-molar small molecule inhibitors that compete with the cofactor.[24,25] The fact that the structure of this dynamic element varies with the nature of the substrate (Fig. 2C), and its proximity to the site of methyl transfer, supports the idea that it plays an important role in substrate recognition.

Structural features unique to PRMT5

A recent crystal structure of PRMT5 revealed that the core Rossman fold and β-barrel were conserved in this Type II methyltransferase (Fig. 3A and B) but, in contrast with the α-X and α-Y helices observed in Type I enzymes, PRMT5 harbors a short α-Y helix preceded by a large loop that links the catalytic Rossman fold to an N-terminal, PRMT5-specific TIM-barrel.[13] The TIM-barrel in turn binds MEP50, a WD-repeat protein that is a core component of PRMT5-containing chromatin
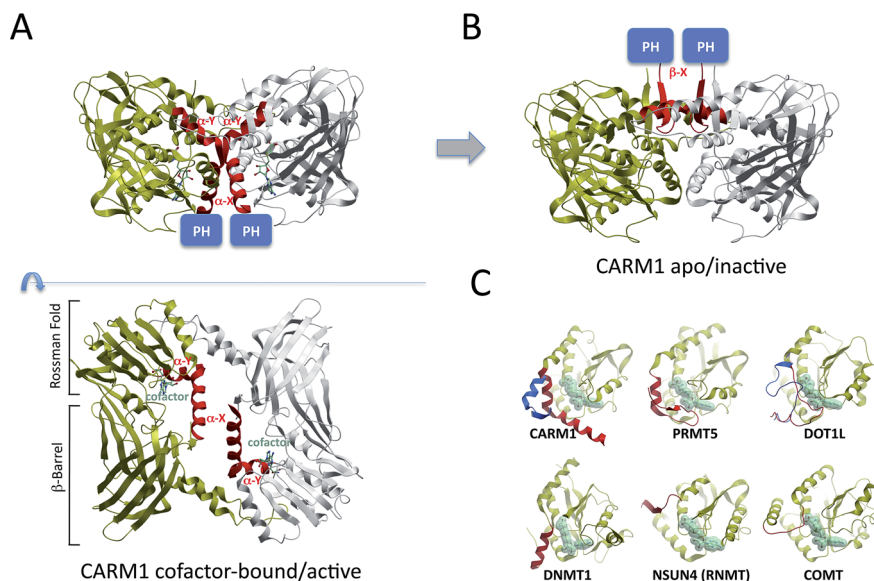
Fig. 2 Structural dynamics of the Type I α-helix. The α-helix observed in Type I PRMTs is structurally dynamic and can adopt strikingly divergent conformations in the active (A − PDB: 3B3F) and inactive (B − PDB: 3B3J) states, correlating with distinct positioning of distant structural modules, such as the PH domain of CARM1 (shown), the SH3 domain of PRMT2, or the ZnF domain of PRMT3, relative to the catalytic domain. (C) Substrate specificity of Rossman-fold methyltransferases may rely on the nature of a regulatory element (red; inactive state: blue) N-terminal to the conserved Rossman fold (yellow). PDB codes: active CARM1: 3B3F; inactive CARM1: 3B3J; PRMT5: 4GQB; active DOT1L: 1NW3; inactive DOT1L: 4ER5; DNMT1: 3PTA; NSUN4: 4FP9; COMT: 1VID. (C is inspired from ref. 40).
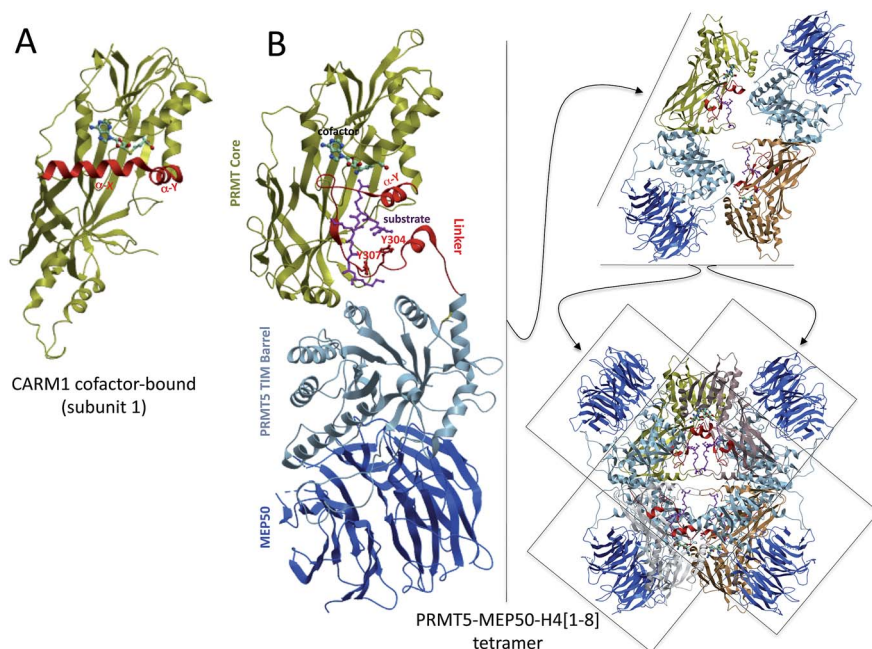

Fig. 3 Structural overview of PRMT5. (A) The structure core of CARM1 is shown as a reference [PDB: 3B3F]. (B) PRMT5 monomers have a conserved catalytic core composed of a Rossman fold and a β-barrel; specific structural features include a TIM barrel that recruits MEP50, and a linker region that contributes to cofactor and substrate binding. Interactions between the TIM barrel and Rossman fold + linker mediate dimerization; orthogonal stacking of dimers produces the PRMT5 tetramer [4GQB].

complexes and acts as a structural scaffold for the recruitment of substrates[13] (Fig. 3B). The loose secondary structural arrangement of the linker domain suggests that it corresponds to a structurally variable region, as observed in the corresponding α-helix of Type I PRMTs.

It is unclear whether the α-helix of Type I PRMTs interacts with residues of the target peptide other than the substrate arginine,[16] but the corresponding linker segment of PRMT5 makes extensive interactions with the backbone and side-chains of at least 8 histone residues centered on the substrate

This journal is © The Royal Society of Chemistry 2014

Med. Chem. Commun., 2014, **5**, 1779–1788 | **1781**

arginine (Fig. 3B).[13] Three tyrosine side-chains from the PRMT5 linker (Y297, Y304 and Y307) are phosphorylated by an oncogenic mutant of the kinase JAK2, resulting in impaired methyltransferase activity.[26] Two of these (Y304 and Y307) are making direct interactions with the substrate histone, which would be disrupted by phosphorylation (Fig. 3B), and phosphorylation of the third would also dramatically affect the electrostatics and geometry of the substrate binding site. In a related mechanism, phosphorylation of Ser217 in CARM1 abrogates its interaction with Tyr154 of the α-X helix, and prevents locking of the α-helix in a conformationally active state.[27] Both Tyr154 and Ser217 are conserved in Type I PRMTs, suggesting that phosphorylation affecting structural elements at the catalytic site may be a common mechanism to regulate the activity of PRMTs.

In addition to the presence of a TIM barrel and a linker region, another PRMT5-specific feature is its tetrameric form. Dimerization is not mediated by a dimerization arm, as observed in Type I PRMTs, and in CePRMT5,[28] but is stabilized by extensive interactions between the TIM barrel of one subunit and the Rossman fold and linker of another. Orthogonal stacking of two dimers generates a tetrameric arrangement at the heart of which lies the substrate binding site, while MEP50 is recruited at the periphery (Fig. 3B). Electron microscopy analysis conducted on XlPRMT5 indicates that the bulk of the interaction with nucleoplasmin, a PRMT5 substrate, is taking place with MEP50 rather than PRMT5 itself, and it was proposed that a channel linking MEP50 to the substrate arginine binding pocket may guide the peptide substrate to the site of methyl transfer.[29] The PRMT5–MEP50 complex has about fifty times stronger affinity than PRMT5 for a substrate H4 peptide,[13] supporting the idea that MEP50 contributes to the recruitment of diverse PRMT5 substrates.

The fact that MEP50 – a core component of PRMT5 complexes necessary for full PRMT5 activity – interacts mainly with the TIM barrel of PRMT5, the position of which is tightly correlated with the conformation of the linker segment, suggests a model where interactions at the TIM barrel would dictate the activation state of the linker domain (α-helix in CARM1). This would echo a structural mechanism proposed for CARM1 where protein–protein interactions at the PH domain would dictate the conformational state of the α-helix and control enzymatic activity. Further structural analysis will be necessary to support or disprove this model.

Structural determinants of product specificity

Structural determinants underlying asymmetrical dimethylation by Type I PRMTs, symmetrical dimethylation by Type II PRMTs and monomethylation by Type III enzymes have not been clearly established, but recent structures of PRMT5 and CePRMT7 in complex with a substrate arginine as well as CARM1 in complex with potent inhibitors occupying the substrate arginine site are providing important insights.[12,13,16] A first observation is that mutating F327 to methionine turns PRMT5 into a dual Type I and Type II PRMT,[28] suggesting that the catalytic mechanism is conserved between the two PRMT

classes, and relies on the only two catalytic site residues that are absolutely conserved across all PRMTs: a pair of glutamates from the Rossman fold "double-E loop" that is probably involved in deprotonating and activating the methyl-accepting nitrogen.[13,18] If Type I, II and III enzymes share the same catalytic mechanism, steric factors are probably responsible for product specificity. While F327 stacks against the substrate guanidinium group in the PRMT5 structure and holds the substrate arginine tightly in place, a more flexible methionine, conserved at this position in Type I enzymes (M163 in CARM1), probably allows sufficient spatial adjustment of the substrate arginine to accommodate two methyl groups on the same nitrogen, though the efficiency of asymmetrical di-methylation seems limited[28] (Fig. 4).

In Type I PRMTs, a conserved YFxxY motif (Y150, F151 and Y154 in CARM1) adopts a particular conformation where the 3 aromatic rings stack in an orthogonal fashion with each other and with the adenine ring of the cofactor; this stabilizes a conformational state where the two hydroxyl groups of the tyrosine side-chains form hydrogen-bonds with one of the conserved glutamates of the double-E loop (E267 in CARM1). The other carboxylic oxygen of the glutamate forms a hydrogen bond with the substrate guanidinium. This structural arrangement has two consequences: (1) it locks in place the substrate arginine, and (2) it generates a large Type I-specific volume surrounding the methyl-accepting nitrogen of the arginine substrate, compatible with binding of unmethylated arginine for monomethylation, or monomethylated arginine for asymmetrical dimethylation. This is in contrast with the tight environment of the other ω-nitrogen of the arginine substrate, which is buttressed against the side-chain of M163 and forms hydrogen-bonds with the side-chains of H415 and the second "double E-loop" glutamate (E258), leaving no room for deposition of a methyl group. Importantly, all these residues are conserved in Type I PRMTs. In summary, the substrate arginine of CARM1 is locked in place by three hydrogen-bonds with the double-E loop glutamates (E258, E267), as well as H415, and by steric constraints from M163[17]. This arrangement is further stabilized by the YFxxY motif (Y150, F151 and Y154) that stacks against the cofactor and denies any rotational freedom at E267. Importantly, these residues are conserved exclusively in Type I PRMTs, and four of them are part of the dynamic α-helix that adopts distinct conformations in the active and inactive states (Fig. 4).

The situation is very different in PRMT5: while a large volume is observed at the methyl-accepting nitrogen in CARM1, this volume is closed in PRMT5 due to a difference in the relative orientation of the substrate guanidinium group and E444 (E267 in CARM1).[28] Methylation of this nitrogen would preclude its repositioning at the same site for asymmetrical dimethylation. Instead, extra volume is present on the other side of the guanidinium group, due to the replacement of a histidine group in Type I enzymes (H415 in CARM1) with a serine group (S578 in PRMT5). This volume can accommodate a monomethylated N-ω-nitrogen, accompanied by placement of the unmethylated N′-ω-nitrogen at the site of methyl transfer,

1782 | Med. Chem. Commun., 2014, 5, 1779–1788

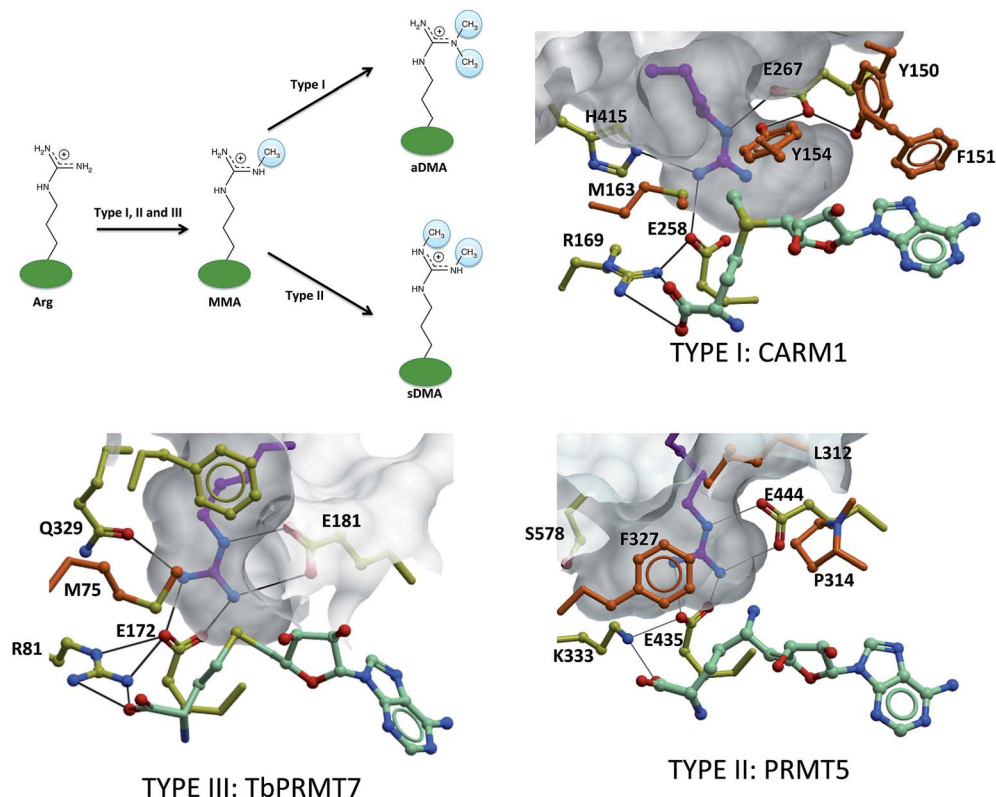This journal is © The Royal Society of Chemistry 2014

Fig. 4 Structural determinants for product specificity of Type I, II and III PRMTs. Type I and II PRMTs asymmetrically and symmetrically dimethylate arginine side-chains respectively, while Type III PRMTs mono-methylate only. The conserved "double-E loop" glutamates are shown, along with residues dictating class specificity in the crystal structures of CARM1 [PDB: 2Y1X], PRMT5 [4GQB] and TbPRMT7 [4M38]. The substrate arginine (positioned manually in the case of CARM1 by superimposing the guanidinium group on the alanine moiety of a co-crystallized inhibitor [2Y1X]) is shown in magenta. Orange: residues from the α-helix of CARM1 and TbPRMT7, and from the linker region of PRMT5.

an arrangement that is in agreement with symmetrical dimethylation (Fig. 4).

Finally, in TbPRMT7, the orientation of the guanidinium group relative to E181 is similar to that observed in PRMT5 (relative to E444), which antagonizes asymmetrical dimethylation. Hydrogen bonding of the N'-ω-nitrogen with Q329 recapitulates the effect of H415 in CARM1 and prevents symmetrical dimethylation[16] (Fig. 4). Further supporting this model, the Q329H TbPRMT7 mutant retains Type III specificity.[16] The fact that the Q329S mutant largely loses catalytic activity suggests that the substrate arginine is no longer stabilized in a conformation where it aligns with the departing methyl, maybe due to the replacement of F327, which is tightly stacked against the substrate guanidinium group in PRMT5, with a more permissive methionine (M75) in PRMT7.

In summary, we propose a model where variations between Type I, Type II and Type III PRMTs in spatial constraints and in hydrogen-bonds at the substrate N- and N'-ω-nitrogens are controlled by class-specific combinations of residues (M163, H415, Y150, F151, and Y154 in CARM1, S578 and F327 in PRMT5, Q329 and M75 in TbPRMT7) that dictate product specificity. An open question is whether these class-specific features could be exploited to design class-specific PRMT inhibitors.

Structural diversity of cofactor and substrate binding sites

The α-helix of Type I PRMTs and the linker domain of PRMT5 fold on the cofactor and participate in the formation of the substrate binding site (Fig. 5A and B). The cofactor site is entirely buried and the substrate site is more open, but sufficiently enclosed to be targeted by potent inhibitors. For instance, the substrate-binding pocket of CARM1 has been co-crystallized with 30 nM inhibitors (Fig. 5A).[30] Bisubstrate inhibitors composed of cofactor analogs that extend into the methyl-arginine binding channel have also been described,[31] but these extremely polar compounds are not drug-like and reflect the challenge of designing cofactor competitors that are sufficiently polar to have high affinity for the hydrophilic cofactor binding site, but sufficiently hydrophobic to cross cell membranes. An important question is whether the structural diversity of these sites is sufficient to develop highly selective chemical probes that could be used as tools for target validation and may serve as the starting point for drug discovery.[32,33]

When clustering methyltransferases based on the structural diversity of their cofactor site using the atomic property fields method,[34] SET domain protein methyltransferases form a cluster of their own, while PRMTs are grouped with small molecule-, DNA- and RNA-methyltransferases (Fig. 6A). This reflects a risk of off-target activity when targeting the cofactor site of PRMTs. This is in agreement with the fact that, unlike SET domain
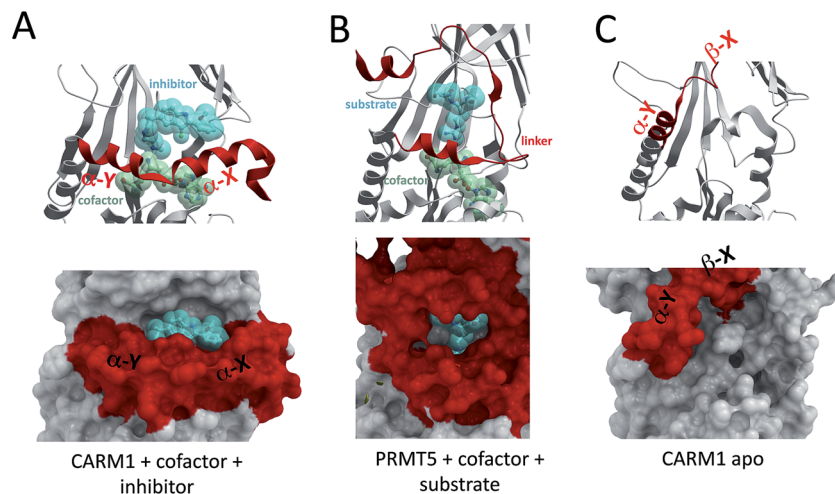
This journal is © The Royal Society of Chemistry 2014

Med. Chem. Commun., 2014, 5, 1779–1788 | 1783

Fig. 5 Structural plasticity of the cofactor and substrate binding sites. The conformationally dynamic α-helix of CARM1 [PDB code 2Y1W] (A) and the corresponding linker domain of PRMT5 [4GQB] (B) contribute to the formation of both cofactor and substrate binding pockets. The cofactor binding pocket is absent in the inactive state of CARM1 [3B3J] (C), and a novel pocket is formed at the interface of the α-Y helix and the structure core.

protein methyltransferases, PRMTs share a cofactor-binding Rossman fold with non-protein methyltransferases (Fig. 2C). Previous work indicated that the structural diversity of the cofactor binding pocket of protein methyltransferases is similar to that of the ATP binding pocket of kinases.[35] Since it is generally possible to develop selective kinase inhibitors, the same should apply to protein methyltransferases in general, and to PRMTs in particular. Alignment of residues lining the cofactor site of Type I PRMTs shows that the pocket is generally conserved, but highlights a few variable positions, such as F138, M164 or A216 (CARM1 nomenclature) that may be exploited to develop isoform-selective inhibitors (Fig. 6B). Nevertheless, the relatively high sequence conservation at the cofactor site should represent a challenge for drug design.

The situation is a little more favorable at the substrate-binding site. A multiple sequence alignment of residues lining potent CARM1 inhibitors that occupy the substrate binding pocket[30] shows high sequence conservation at the substrate arginine channel, but significant diversity at the other end of the pocket, at positions occupied by S146, Q149, N266, R446 or S448 in CARM1 (Fig. 6C). The fact that the two 30 nM CARM1 inhibitors are inactive against PRMT1 and PRMT3 (ref. 36 and 37) supports the idea that selective inhibition can be achieved at the substrate binding site, though profiling these inhibitors against other PRMTs would be informative. Interestingly, different warheads in each inhibitor are anchored in the methyl-arginine binding channel of CARM1, but in both cases, hydrogen bonds are formed with the side-chains of H415 and the double E-loop glutamate E258, two residues that are absolutely conserved in Type I PRMTs, as well as the backbone carbonyl of M260 (Fig. 6C). This suggests that potent Type I PRMT inhibitors binding at the substrate-binding pocket should efficiently exploit these positions. Considering the high sequence conservation at the methyl-arginine channel, combinatorial libraries focused on these two warheads should be enriched in Type I PRMT inhibitors.

Allosteric inhibition of PRMT3

Compounds from a third chemical series were co-crystallized at an allosteric pocket of PRMT3, located at the base of the dimerization arm (Fig. 7A).[15,38] In spite of their binding mode, the compounds do not prevent homodimerization of PRMT3. This unexpected binding site is 15 Å away from the site of methyl transfer, and it is surprising that binding at such a distant pocket could have an inhibitory effect, but the mechanistic relevance of the crystal structure was confirmed by mutagenesis studies showing that mutating residues at the allosteric site (ex K392A) preserved the catalytic activity of the enzyme, but reduced the effect of the inhibitor (Fig. 7B).[15] The structure clearly shows that the benzothiadiazole moiety of the compound is inserted in an enclosed cavity of the β-barrel of a PRMT3 subunit, while the 2-oxo-phenyl end of the inhibitor extends out of the cavity and is buttressed against the α-helix of the other PRMT3 subunit (Fig. 7B). Interestingly, the α-helix is in an inactive state, suggesting that the inhibitor induces some conformational strain that antagonizes proper folding of the α-X helix on the cofactor. While the reported inhibitors only have IC$_{50}$s in the low micromolar range, it is expected that potent compounds could be targeting this site.

The entrance of the pocket is occluded in the structure of PRMT3 in complex with the cofactor, due to an alternate rotameric state of R396 (PDB code 2FYT) (Fig. 7C). The fact that the pocket is not seen in other PRMT structures does therefore not necessarily mean that it is unique to PRMT3. A multiple sequence alignment of Type I PRMTs based on PRMT3 residues lining the co-crystallized inhibitor reveals significant sequence variability, indicating that this pocket can be exploited towards the development of selective inhibitors (Fig. 7D).

Unexplored allosteric binding pockets

The dynamic nature of the α-helix of Type I PRMTs translates into significant structural plasticity at both cofactor and substrate binding sites. In the structure of apo CARM1, the

1784 | *Med. Chem. Commun.*, 2014, **5**, 1779–1788

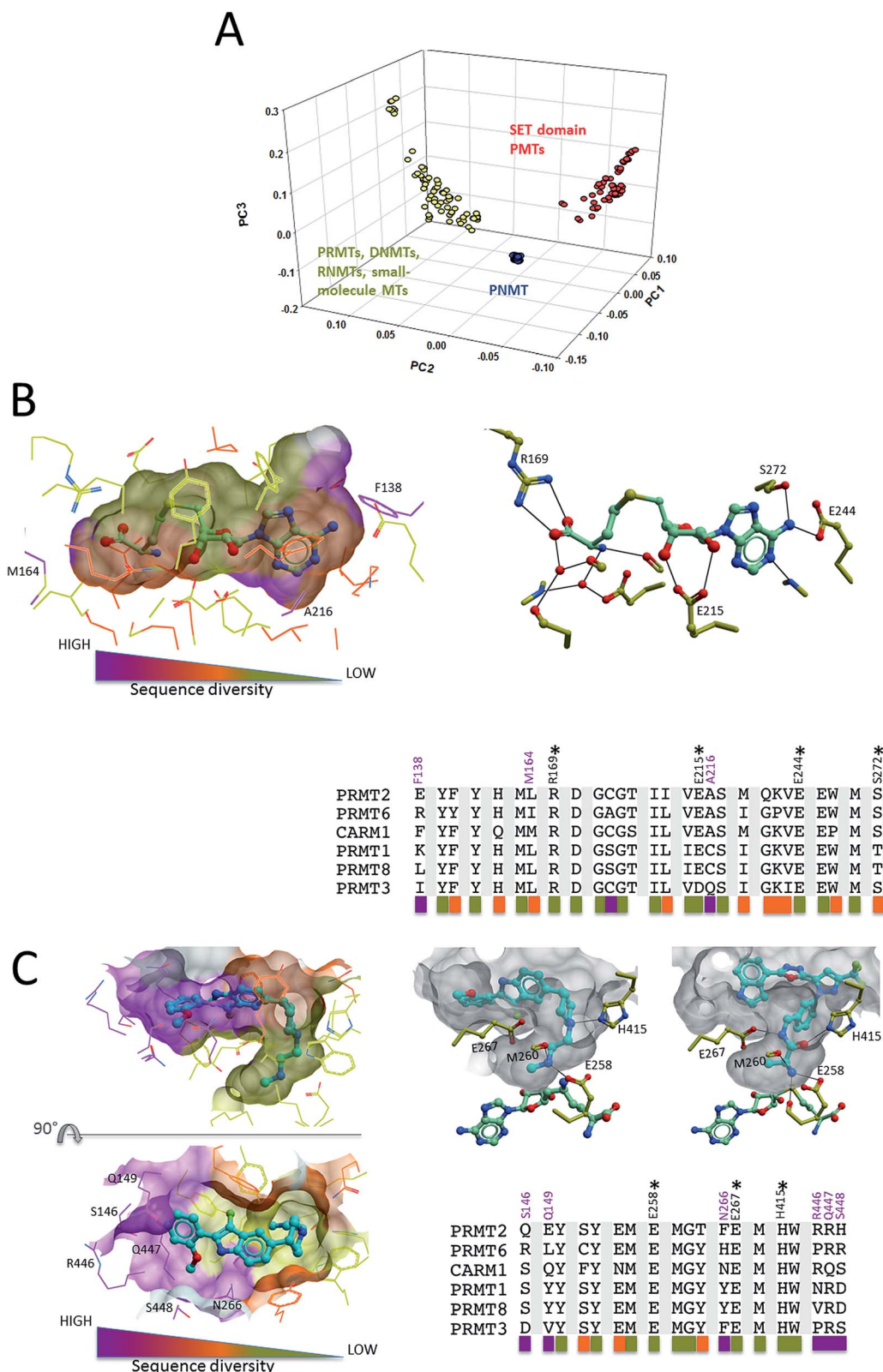This journal is © The Royal Society of Chemistry 2014

Fig. 6 Structural diversity of the cofactor and substrate binding pockets. (A) Clustering of human methyltransferases in the protein databank based on the structural similarity of their cofactor site (Molsoft's APF method was used[34]). (B) Side-chains within 5 Å of the cofactor in the CARM1 structure (PDB code 3B3F) were extracted from a multiple sequence alignment of Type I PRMTs. Sequence conservation at these positions was used for color-coding. Conserved residues forming direct hydrogen-bonds with the cofactor are marked with an "*". Positions with significant sequence variability are highlighted in magenta (CARM1 numbering). (C) The same procedure as in (B) was applied to the substrate binding pocket, defined by side-chains within 5 Å of 2 substrate competitors co-crystallized to CARM1 (PDB codes 2Y1W and 2Y1X). (PNMT: phenyl-ethanolamine N-methyltransferase).
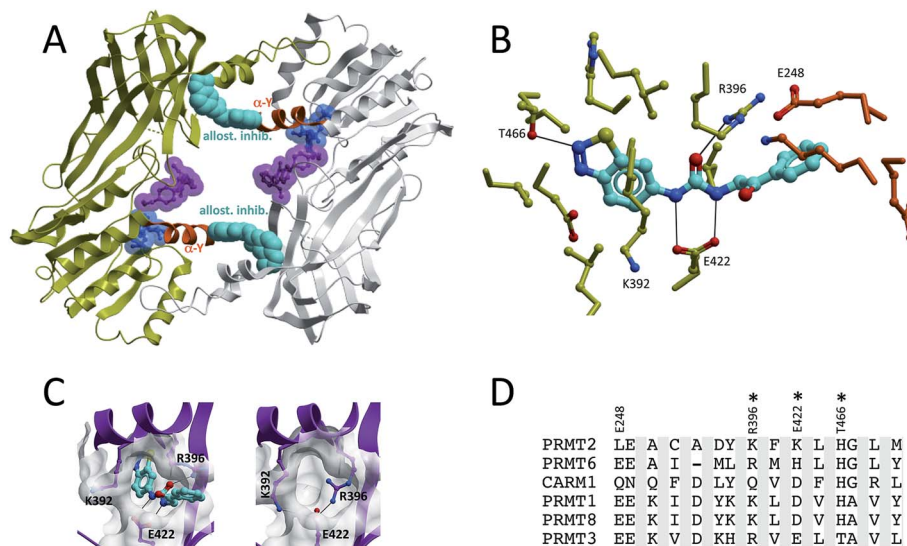
This journal is © The Royal Society of Chemistry 2014

Med. Chem. Commun., 2014, 5, 1779–1788 | 1785

Fig. 7 Allosteric inhibition of PRMT3. (A and B) An allosteric PRMT3 inhibitor binds in a cavity of the β-barrel, at the base of the dimerization arm, and is buttressed against the α-helix (orange) of the other PRMT3 subunit. The cofactor-binding pocket is highlighted by a cofactor molecule (blue), and the substrate binding pocket by a substrate competitor (magenta) co-crystallized to CARM1 (superimposed CARM1 structure not shown). Both the cofactor and CARM1 inhibitor are shown as references, but are absent from the PRMT3 structure. (C) Entrance to the allosteric pocket is occluded by R396 in the absence of the inhibitor. (D) Sequence diversity of residues lining the allosteric pocket (constructed as in Fig. 6).
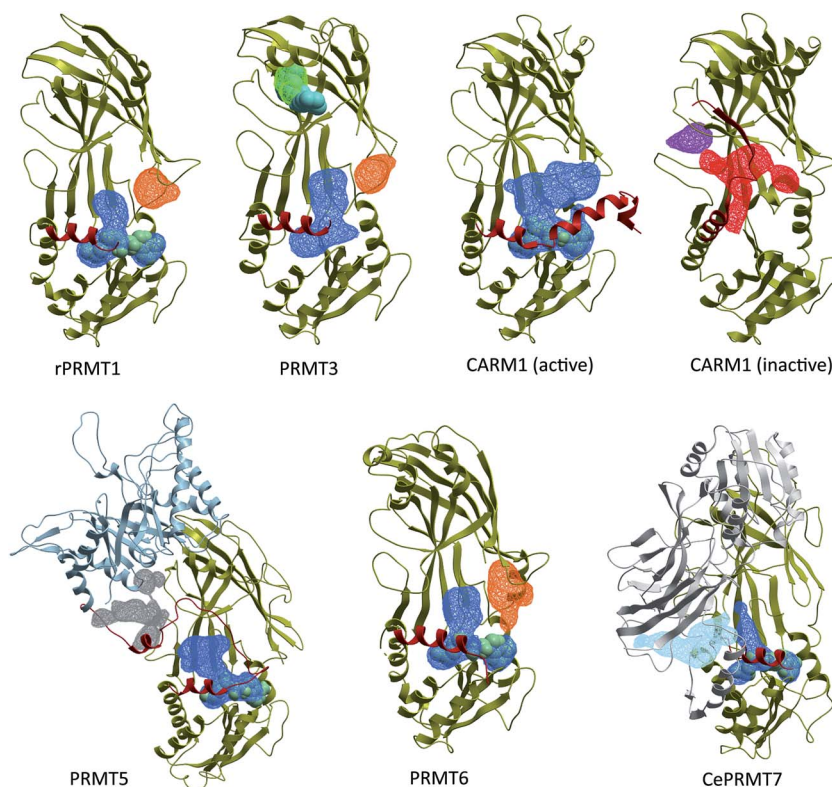

Fig. 8 The PRMT pocketome. ICMPocketFinder (Molsoft, San Diego) was used to map pockets present in PRMT structures. Blue: merged cofactor and substrate binding pockets. Green: experimentally validated PRMT3 allosteric pocket. Orange: unexplored allosteric pocket observed in PRMT1, PRMT3 and PRMT6. Red: novel CARM1 pocket formed by the inactive α-helix. Light blue: pocket formed at the interface of the C-terminal β-barrel and pseudo-dimerization arm of CePRMT7 (C-terminal pseudo-PRMT core shown in white). PDB codes: PRMT1: 1OR8; PRMT3: 3SMQ; CARM1 (active): 2Y1W; CARM1 (inactive): 3B3J; PRMT5: 4GQB; PRMT6: 4HC4; CePRMT7: 3WSS. Red: α-helix of Type I/III PRMTs and linker domain of PRMT5.

1786 | *Med. Chem. Commun.*, 2014, **5**, 1779–1788

This journal is © The Royal Society of Chemistry 2014

cofactor binding pocket is gone, and a profound remodeling takes place at what used to be the substrate binding site, where an entirely new cavity emerges from the relocation of the α-helix (Fig. 5C and 8). Compounds that occupy this site would presumably lock the enzyme in an inactive conformation.

In an attempt to identify other potential allosteric binding sites, we used Molsoft's ICMPocketFinder[39] to systematically search for binding pockets in PRMT structures available in the PDB (Fig. 8). In addition to the substrate and cofactor binding sites, and the PRMT3 allosteric site, this exercise reveals the existence of unexplored binding pockets. First, a cavity present in PRMT1, PRMT3, PRMT6 and apo-CARM1 is located at the interface of the Rossman fold and the β-barrel, immediately next to the second double-E loop glutamate (E355 in PRMT3). The entrance to this cavity is occupied by the α-helix in the active conformation of CARM1, suggesting a model where inhibitors occupying this site would destabilize the catalytically competent conformation of the α-helix. Another pocket is identified at the interface of the linker domain and the TIM-barrel of PRMT5. It is located in the vicinity of the substrate-binding site, and compounds binding at this site may affect recruitment of PRMT5 substrates. A large cavity is also present at the junction of the C-terminal β-barrel and the pseudo-dimerization arm of CePRMT7, immediately next to the N-terminal catalytic site (Fig. 8, light blue). However, CePRMT7 shares only 32% sequence identity with hPRMT7, and the corresponding pocket may not be present in the latter. Finally, a pocket is found at the back of the β-barrel of inactive CARM1 (Fig. 8, purple). However, unlike the pocket targeted by PRMT3 allosteric inhibitors, this cavity is facing away from the other subunits, and the putative effect of ligands binding at this site is unclear.

Conclusion

PRMTs share a catalytic Rossman fold with DNA-, RNA-, and small molecule-methyltransferases, as well as some lysine protein methyltransferases, and novel enzymes with PRMT activity may emerge from the human methyltransferasome.[1] We propose a model where steric constraints dictate class (i.e. product) specificity of PRMTs. Structural determinants for substrate specificity, another dimension to the signaling versatility of PRMTs, remain largely unexplored, and more structures such as the recent PRMT5- or TbPRMT7-histone 4 complexes will be important to address this question. Another area for future research is the influence of protein–protein interaction modules such as the PH domain of CARM1, or the Zn finger of PRMT3 on the activation state of PRMTs. This may necessitate the introduction of additional components of chromatin complexes, as was done in the case of PRMT5.[13] The discovery and co-crystal structure of potent CARM1 inhibitors demonstrated that the substrate-binding pocket of PRMTs is chemically tractable. The cofactor-binding site is also very enclosed, but low diversity and high hydrophilicity make it a challenging target. More promising is the opportunity to develop allosteric inhibitors that exploit the structural plasticity of PRMTs, as was demonstrated in the case of PRMT3.[15,38]

References

1 T. C. Petrossian and S. G. Clarke, *Mol. Cell. Proteomics*, 2011, **10**, M110000976.

2 V. M. Richon, D. Johnston, C. J. Sneeringer, L. Jin, C. R. Majer, K. Elliston, L. F. Jerva, M. P. Scott and R. A. Copeland, *Chem. Biol. Drug Des.*, 2011, **78**, 199–210.

3 C. I. Zurita-Lopez, T. Sandberg, R. Kelly and S. G. Clarke, *J. Biol. Chem.*, 2012, **287**, 7859–7870.

4 A. Di Lorenzo and M. T. Bedford, *FEBS Lett.*, 2011, **585**, 2024–2031.

5 Y. Yang and M. T. Bedford, *Nat. Rev. Cancer*, 2013, **13**, 37–50.

6 R. Swiercz, D. Cheng, D. Kim and M. T. Bedford, *J. Biol. Chem.*, 2007, **282**, 16917–16923.

7 M. T. Bedford and S. G. Clarke, *Mol. Cell.*, 2009, **33**, 1–13.

8 B. Cha and E. H. Jho, *Expert Opin. Ther. Targets*, 2012, **16**, 651–664.

9 R. A. Copeland, M. E. Solomon and V. M. Richon, *Nat. Rev. Drug Discovery*, 2009, **8**, 724–732.

10 C. D. Krause, Z. H. Yang, Y. S. Kim, J. H. Lee, J. R. Cook and S. Pestka, *Pharmacol. Ther.*, 2007, **113**, 50–87.

11 X. Cheng, R. E. Collins and X. Zhang, *Annu. Rev. Biophys. Biomol. Struct.*, 2005, **34**, 267–294.

12 N. Troffer-Charlier, V. Cura, P. Hassenboehler, D. Moras and J. Cavarelli, *EMBO J.*, 2007, **26**, 4391–4401.

13 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.

14 M. Hasegawa, S. Toma-Fukai, J. D. Kim, A. Fukamizu and T. Shimizu, *FEBS Lett.*, 2014, **588**, 1942–1948.

15 A. Siarheyeva, G. Senisterra, A. Allali-Hassani, A. Dong, E. Dobrovetsky, G. A. Wasney, I. Chau, R. Marcellus, T. Hajian, F. Liu, I. Korboukh, D. Smil, Y. Bolshan, J. Min, H. Wu, H. Zeng, P. Loppnau, G. Poda, C. Griffin, A. Aman, P. J. Brown, J. Jin, R. Al-Awar, C. H. Arrowsmith, M. Schapira and M. Vedadi, *Structure*, 2012, **20**, 1425–1435.

16 C. Wang, Y. Zhu, T. B. Caceres, L. Liu, J. Peng, J. Wang, J. Chen, X. Chen, Z. Zhang, X. Zuo, Q. Gong, M. Teng, J. M. Hevel, J. Wu and Y. Shi, *Structure*, 2014, **22**, 756–768.

17 W. W. Yue, M. Hassler, S. M. Roe, V. Thompson-Vale and L. H. Pearl, *EMBO J.*, 2007, **26**, 4402–4412.

18 X. Zhang and X. Cheng, *Structure*, 2003, **11**, 509–520.

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

This journal is © The Royal Society of Chemistry 2014

Med. Chem. Commun., 2014, **5**, 1779–1788 | 1787

20 J. Min, Q. Feng, Z. Li, Y. Zhang and R. M. Xu, *Cell*, 2003, **112**, 711–723.

21 J. Song, O. Rechkoblit, T. H. Bestor and D. J. Patel, *Science*, 2011, **331**, 1036–1040.

22 H. Spahr, B. Habermann, C. M. Gustafsson, N. G. Larsson and B. M. Hallberg, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 15253–15258.

23 J. Vidgren, L. A. Svensson and A. Liljas, *Nature*, 1994, **368**, 354–358.

24 A. Basavapathruni, L. Jin, S. R. Daigle, C. R. Majer, C. A. Therkelsen, T. J. Wigle, K. W. Kuntz, R. Chesworth, R. M. Pollock, M. P. Scott, M. P. Moyer, V. M. Richon, R. A. Copeland and E. J. Olhava, *Chem. Biol. Drug Des.*, 2012, **80**, 971–980.

25 W. Yu, E. J. Chory, A. K. Wernimont, W. Tempel, A. Scopton, A. Federation, J. J. Marineau, J. Qi, D. Barsyte-Lovejoy, J. Yi, R. Marcellus, R. E. Iacob, J. R. Engen, C. Griffin, A. Aman, E. Wienholds, F. Li, J. Pineda, G. Estiu, T. Shatseva, T. Hajian, R. Al-Awar, J. E. Dick, M. Vedadi, P. J. Brown, C. H. Arrowsmith, J. E. Bradner and M. Schapira, *Nat. Commun.*, 2012, **3**, 1288.

26 F. Liu, X. Zhao, F. Perna, L. Wang, P. Koppikar, O. Abdel-Wahab, M. W. Harr, R. L. Levine, H. Xu, A. Tefferi, A. Deblasio, M. Hatlen, S. Menendez and S. D. Nimer, *Cancer Cell*, 2011, **19**, 283–294.

27 Q. Feng, B. He, S. Y. Jung, Y. Song, J. Qin, S. Y. Tsai, M. J. Tsai and B. W. O'Malley, *J. Biol. Chem.*, 2009, **284**, 36167–36174.

28 L. Sun, M. Wang, Z. Lv, N. Yang, Y. Liu, S. Bao, W. Gong and R. M. Xu, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 20538–20543.

29 M. C. Ho, C. Wilczek, J. B. Bonanno, L. Xing, J. Seznec, T. Matsui, L. G. Carter, T. Onikubo, P. R. Kumar, M. K. Chan, M. Brenowitz, R. H. Cheng, U. Reimer, S. C. Almo and D. Shechter, *PLoS One*, 2013, **8**, e57008.

30 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.

31 J. Dowden, W. Hong, R. V. Parry, R. A. Pike and S. G. Ward, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 2103–2105.

32 M. E. Bunnage, E. L. Chekler and L. H. Jones, *Nat. Chem. Biol.*, 2013, **9**, 195–199.

33 S. V. Frye, *Nat. Chem. Biol.*, 2010, **6**, 159–161.

34 M. Totrov, *BMC Bioinf.*, 2011, **12**(1), S35.

35 V. Campagna-Slater, M. W. Mok, K. T. Nguyen, M. Feher, R. Najmanovich and M. Schapira, *J. Chem. Inf. Model.*, 2011, **51**, 612–623.

36 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.

37 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.

38 F. Liu, F. Li, A. Ma, E. Dobrovetsky, A. Dong, C. Gao, I. Korboukh, J. Liu, D. Smil, P. J. Brown, S. V. Frye, C. H. Arrowsmith, M. Schapira, M. Vedadi and J. Jin, *J. Med. Chem.*, 2013, **56**, 2110–2124.

39 J. An, M. Totrov and R. Abagyan, *Mol. Cell. Proteomics*, 2005, **4**, 752–761.

40 W. Yu, D. Smil, F. Li, W. Tempel, O. Fedorov, K. T. Nguyen, Y. Bolshan, R. Al-Awar, S. Knapp, C. H. Arrowsmith, M. Vedadi, P. J. Brown and M. Schapira, *Bioorg. Med. Chem.*, 2013, **21**, 1787–1794.

1788 | *Med. Chem. Commun.*, 2014, **5**, 1779–1788

This journal is © The Royal Society of Chemistry 2014