

Small Molecule Inhibitors of Histone Arginine Methyltransferases: Homology Modeling, Molecular Docking, Binding Mode Analysis, and Biological Evaluations

Rino Ragno,^{*,†} Silvia Simeoni,[†] Sabrina Castellano,[‡] Caterina Vicidomini,[‡] Antonello Mai,[†] Antonella Caroli,[†] Anna Tramontano,[§] Claudia Bonaccini,^{§,+} Patrick Trojer,^{||} Ingo Bauer,[⊥] Gerald Brosch,^{*,⊥} and Gianluca Sbardella^{*,‡}

Istituto Pasteur - Fondazione Cenci Bolognetti, Dipartimento di Studi Farmaceutici, Università degli Studi di Roma "La Sapienza", P.le Aldo Moro 5, I-00185 Roma, Dipartimento di Scienze Farmaceutiche, Università degli Studi di Salerno, Via Ponte Don Melillo, I-84084 Fisciano (SA), Dipartimento di Scienze Biochimiche, A. Rossi Fanelli, Università degli Studi di Roma "La Sapienza", P.le Aldo Moro 5, 00185 Roma, Howard Hughes Medical Institute, Department of Biochemistry, Division of Nucleic Acids Enzymology, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, 683 Hoes Lane, Piscataway, New Jersey 08854, and Division of Molecular Biology, Biocenter-Innsbruck Medical University, Fritz-Preglstrasse 3, 6020 Innsbruck, Austria

Received October 17, 2006

The screening of the inhibition capabilities of dye-like small molecules from a focused library against both human PRMT1 and *Aspergillus nidulans* RmtA is reported as well as molecular modeling studies (homology modeling, molecular docking, and 3-D QSAR) of the catalytic domain of the PRMT1 fungal homologue RmtA. The good correlation between computational and biological results makes RmtA a reliable tool for screening arginine methyltransferase inhibitors. In addition, the binding mode analyses of tested derivatives reveal the crucial role of two regions, the pocket formed by Ile12, His13, Met16, and Thr49 and the SAM cysteinic binding site subsite. These regions should be taken into account in the design of novel PRMT inhibitors.

Introduction

Similar to other posttranslational covalent modifications, such as phosphorylation, acetylation, and ubiquitination, histone methylation regulates a broad range of DNA and chromatin-templated nuclear events, including transcription.^{1,2} Histones can be methylated on lysine as well as arginine residues, preferentially on the amino-terminal tails of histones H3 and H4, and this methylation is a stable epigenetic mark that does not alter the overall charge of the histone tails. Nevertheless, the recent identification and structural characterization of "histone demethylases"^{3–12} led us to consider that histone methylation might also play a dynamic role in the epigenetic code.^{3–5} However, with increasing methylation¹³ comes an increase in basicity, hydrophobicity, and an influence on the affinity for anionic molecules such as DNA.^{14,15} As a matter of fact, similarly to histone acetylation, histone methylation can modulate histone interaction with DNA and chromatin associated proteins, resulting in an alteration of nucleosomal structures and functions and ultimately contributing to different biological processes.¹⁶ Histone methyltransferases display remarkable specificity in the level of methylation they catalyze, and the latest findings suggest that this could have functional significance in transcription.^{17,18} The nine mammalian protein arginine methyltransferases (PRMTs) thus far identified^{19,20} share a

highly conserved catalytic domain. Seven of them (including the recently discovered FBXO11/PRMT9) catalyze the transfer of a methyl group from *S*-adenosylmethionine (SAM or AdoMet, **1**, Chart 1) to guanidino nitrogen atoms of arginine residues,²¹ resulting in *S*-adenosylhomocysteine (**2**, Chart 1) and mono- or dimethylated (symmetric or asymmetric) arginine. The fact that PRMTs are known coactivators for nuclear receptors makes them likely candidates to be overexpressed in prostate and breast cancers. The inhibition of both PRMT1 and CARM1 (coactivator-associated arginine methyltransferase-1) can suppress estrogen and androgen receptor-mediated transcriptional activation.²² In addition, the ability of PRMT5, when overexpressed, to promote anchorage-independent cell growth also points to this enzyme as a candidate for deregulation in transformed cellular states. Actually, two types of compounds are used for the inhibition of PRMTs: inhibitors of *S*-adenosylhomocysteine catabolism (like 2',3'-acycloadenosine-2',3'-dialdehyde **3**, Chart 1)²³ and SAM analogues (like methylthioadenosine, **4**, and sinefungin, **5**, Chart 1),^{24,25} but, though effective, they both have displayed limited specificity, indiscriminately targeting all SAM-dependent enzymes. For this reason the development of novel small molecule selective inhibitors of PRMTs is highly desirable.

Recently a few compounds selected through a random screening were described as reversible arginine methyltransferase inhibitors (AMIs) by Bedford and co-workers.²² Being interested in small molecule modulators of epigenetic targets and, particularly, of histone-modifying enzymes,^{26–28} we focused our attention on AMI structures and noticed that all of them were dyes or dye-like derivatives. Particularly, two scaffolds (**A** and **B**, Chart 2) emerged as privileged ones.

This finding, together with the aim to develop a rational approach toward arginine methyltransferase inhibitors, prompted us to explore the affinity for the target binding site of selected AMIs as well as of a focused library (Chart 3) of dye-like small molecule compounds (mostly containing privileged structures scaffolds **A** and **B**).

* To whom correspondence should be addressed. R.R., Tel.: +39-06-4991-3937; Fax: +39-06-491491; e-mail: rino.ragno@uniroma1.it; G.B., Tel.: +43-512-9003-70211; Fax: +43-512-9003-73100; e-mail: gerald.brosch@i-med.ac.at; G.S., Tel.: +39-089-96-9770; Fax: +39-089-96-9602; e-mail: gsardella@unisa.it.

[†] Istituto Pasteur - Fondazione Cenci Bolognetti, Dipartimento di Studi Farmaceutici, Università degli Studi di Roma "La Sapienza".

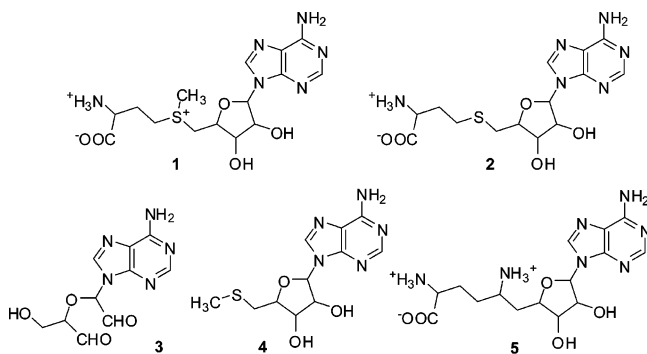
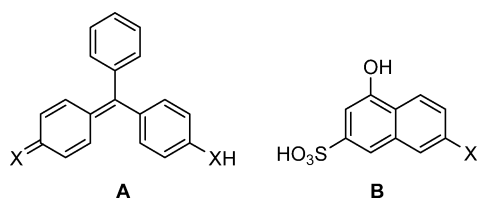
[‡] Università degli Studi di Salerno.

[§] Dipartimento di Scienze Biochimiche, Università degli Studi di Roma "La Sapienza".

^{||} University of Medicine and Dentistry of New Jersey.

[⊥] Biocenter-Innsbruck Medical University.

⁺ Current address: Dipartimento di Scienze Farmaceutiche, Polo Scientifico di Sesto Fiorentino, Università degli Studi di Firenze, Via U. Schiff 6, I-50019 Sesto Fiorentino (FI).

Chart 1. Nonselective SAM-Dependent Methyltransferases Inhibitors**Chart 2.** Privileged Scaffolds

In this paper we report structure-based (SB) and ligand-based (LB) molecular modeling studies (homology modeling, molecular docking, and 3-D QSAR studies) to investigate the binding mode of several small molecule AMIs (**6–20**) in the catalytic domain of the PRMT1 fungal homologue RmtA.²⁹ Along with computational studies, enzymatic assays on both RmtA and PRMT1 using SAM and histones as substrates are also reported.

Chemistry. Derivatives **6e**, **6f**,³⁰ **9**,³¹ **10b**,³² **10c**, and **17**³³ were synthesized as described in the Experimental Section.

All other derivatives were purchased from Lancaster Synthesis, Milan (Italy), or Sigma-Aldrich, Milan (Italy).

Results and Discussion

Biological Results. Compounds **6–20** were tested for their inhibition capabilities against human PRMT1 as well as *Aspergillus nidulans* RmtA (Table 1).

Interestingly and as expected, as a general rule, derivatives **6–20** exhibited similar inhibitory potency against hPRMT1 and RmtA. In fact, IC₅₀ data for the two enzymes correlate quite well, displaying a correlation coefficient of 0.91 (Supporting Information, Figure A).

Derivatives **6c**, **6f**, and **9** confirmed their activity against both tested enzymes. Among the molecules of the A set (Chart 3) we found that derivatives **6a–f** exhibited the same activity rank (**6d** < **6a** < **6b** < **6f** < **6e** < **6c**) against both hPRMT1 and RmtA. In particular, we observed that bromine substituents in the positions 2, 4, 5, and 7 of the xanthenic moiety greatly enhanced inhibitory potency (compare IC₅₀ of compounds **6b** and **6c**). On the contrary, the introduction of an amino group in the position 5 of the phenyl ring was detrimental for the activity (compare derivative **6d** with **6b**) unless counteracted by the effects of bromine substituents (compound **6e**) or of the *N*-dichloro-triazino moiety (derivative **6f**). The formal opening of the xanthenic moiety resulting in the triarylmethane scaffold (compounds **7a–d** and **8**) significantly decreased inhibition capability even when the carboxylic group was replaced with a sulfonic group or with the introduction of bulky alkyl substituents in the positions 2, 4, 5, and 7 (compare IC₅₀ value of compound **6b** with those of compounds **7a**, **7c**, and **7b**, respectively). Similarly, other simplifications or modifications

(compounds **16–19**) of the core structure of compounds **6a–f** yielded inactive or scarcely active derivatives (Table 1).

As far as the molecules of the B set (Chart 3) are concerned, we observed a slightly different activity ranking against RmtA (**10c** < **10b** < **10a** < **11** < **9**) and PRMT1 (**10c** < **10a** < **11** < **10b** < **9**). Yet, in both cases the symmetrical ureidic derivative **9** was the most active. The replacement of one of the two hydroxynaphthalene sulfonic substituents with the shorter and nonpolar phenyl group resulted in a significantly reduced inhibitory potency (compare IC₅₀ values of **9** and **10b**), this decrease being even more marked if the oxygen of the urea moiety was replaced with a sulfur atom (compound **10c**). Therefore, it is not surprising that suramin (derivative **15**), a polar polysulfonated symmetrical urea was significantly more active than **9** in the inhibition of both tested enzymes (compare IC₅₀ values of compounds **9** and **15**). On the contrary, it was unexpected that derivative **12**, with a diazo group replacing the ureidic moiety and only one of the two naphthyl substituents functionalized with polar groups, was comparable to **9** in inhibiting both RmtA and PRMT1 (compare IC₅₀ values of compounds **9** and **12**). Finally, derivatives **13**, **14**, and **20** showed a very low (if any) activity.

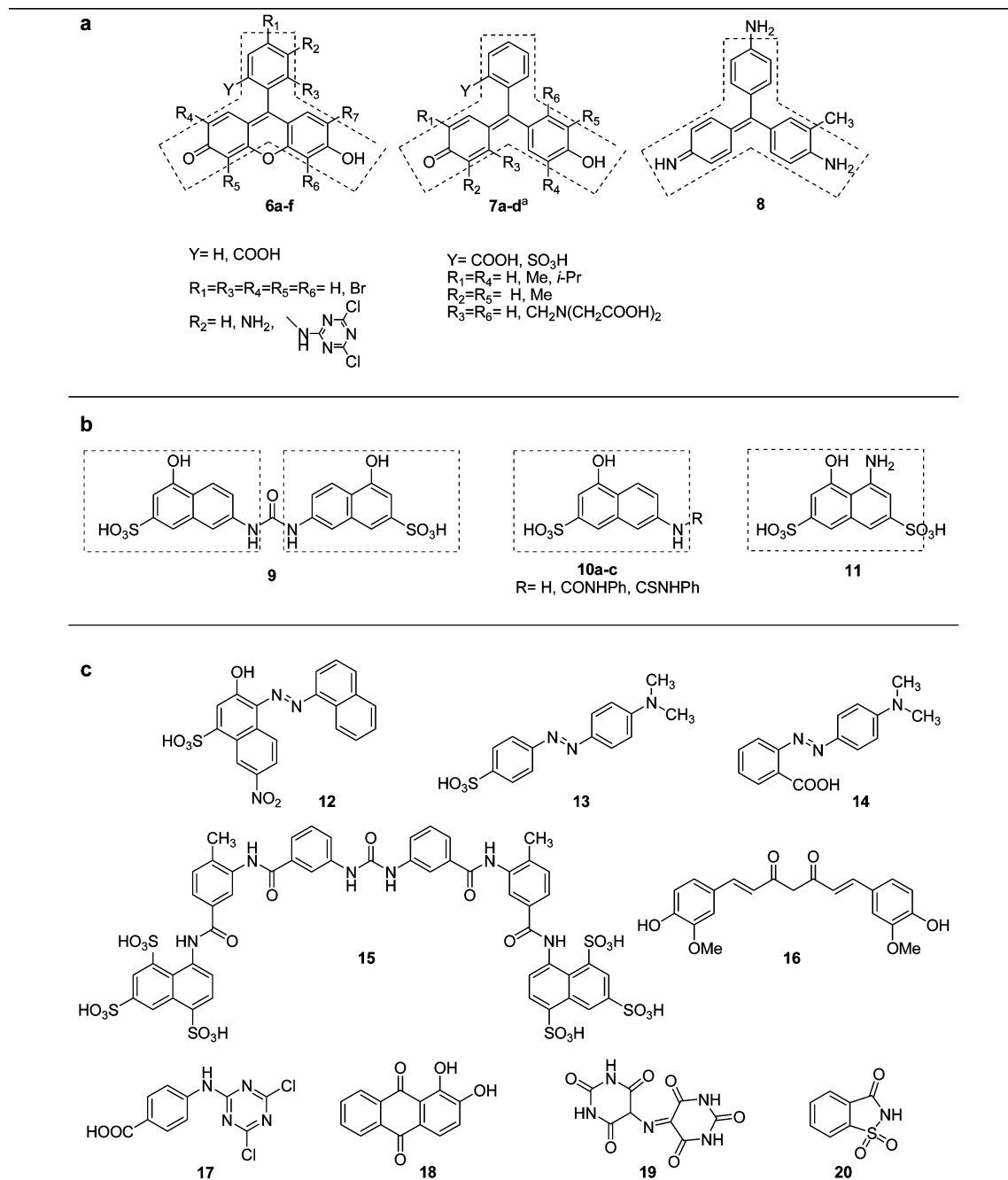
Homology Modeling. Computational approaches to the prediction of protein structures are valuable tools for structure-based research.³⁴ Indeed, in many cases homology models have been shown to be valuable research tools for identification, validation, and optimization of compounds,³⁵ even though the success in using such models is very dependent on their quality, which is in turn correlated to the sequence similarity between the target protein and the template structures used for modeling.³⁶

Our previous studies on *Aspergillus nidulans* genome²⁹ revealed the existence of three distinct PRMTs exhibiting in vitro histone methyltransferase (HMT) activity as recombinant proteins. One of these proteins, termed RmtA, showed significant sequence similarity (Figure 1) and similar biochemical properties compared to human PRMT1.²⁹ As we had access to a large amount of the PRMT1 orthologue RmtA, similarly to our approach to design new HDAC inhibitors,^{26,27,37–44} we decided to verify if RmtA could serve as workhorse for our enzymatic assays and if the RmtA three-dimensional model could be used for computer-aided ligand design studies.

The PDB database^{45,46} contains the crystal structures of three representative PRMTs from two different families: the rat PRMT1 (rPRMT1)⁴⁷ (residues 41–353, pdb entry code 1OR8, *R* = 2.35 Å) and its yeast homologue RMT1/Hmt1⁴⁸ (residues 30–348, pdb entry code 1G6Q, *R* = 2.90 Å), and the rat PRMT3 (rPRMT3) catalytic core⁴⁹ (residues 208–528, pdb entry code 1F3L, *R* = 2.03 Å). All these structures present a conserved core of about 310 residues formed by two domains: the N-terminal catalytic domain, which contains the SAM-binding site, and the C-terminal barrel-like domain. The all-beta structure of the latter is interrupted by an arm-like motif consisting of two alpha helices connected by a short loop, which seems to be implicated in PRMT dimerization and protein structure stabilization by hydrophobic interaction with the outer surface of the catalytic domain of a neighboring molecule. Among the available structures, all with the exception of the yeast RMT1/Hmt1 have been determined in complex with the demethylated cofactor *S*-adenosyl-homocysteine (AdoHcy). Rat PRMT1 has also been complexed with peptides mimicking the protein substrate.

Optimal superposition of the three structures yielded a root-mean-square deviation (rmsd) over the main chain atoms of less

Chart 3. Structures of Derivatives 6–20



a) derivatives containing privileged scaffold A; b) derivatives containing privileged scaffold B; c) derivatives partially containing either scaffold A or scaffold B.

than 1 Å. The sequence identity ranges between 41% and 47% in the conserved core and increases to more than 60% in the N-terminal catalytic domain (Figure 1).⁵⁰ Interestingly, almost any residue within the target sequence is represented in at least one structure, with few exceptions principally located in exposed regions. In particular all the residues observed as forming interactions with both the AdoHcy cofactor and Arg substrate are very well conserved, with only minor differences in hydrophobic side chain dimensions.

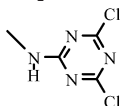
The crystal structures of rat PRMT1 (pdb code 1OR8, 1ORI, 1ORH) were obtained at a nonphysiological pH value (pH 4.7; maximum enzymatic activity at pH 8.5) and therefore are not suitable as target structures for virtual screening. Additionally, an important helical section near to the binding pocket was not resolved in the PRMT1 X-ray structures.⁴⁷ For better results we therefore decided to apply a multiple template approach⁵¹

and to make use of all the structures (from yeast RMT1, rat PRMT1, and rat PRMT3) to build our RmtA homology model by the SWISS-MODEL server.⁵² The obtained model was then subjected to energy minimization in the presence of both SAM and Arg substrates in order to optimize side-chain positions (Figure 2).

Similarly as reported,⁴⁷ SAM and Arg pockets lie close to each other to allow the monomethylation or the dimethylation of the guanidine nitrogen atoms of arginine side chain (Figure 2).

Docking Studies and Binding Mode Analysis of Compound 6c, 6f, and 9. The binding modes of tested compounds 6a–20 into the modeled RmtA were analyzed by the means of the program Autodock,⁵³ and the X-Score⁵⁴ external scoring function was applied to select the binding conformations. To assess the reliability of the Autodock method either the SAM

Table 1. Inhibition Activities (IC₅₀) of **6–20** against hPRMT1 and RmtA

compd	Y	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	R ₇	IC ₅₀ (μM)	
									RmtA	PRMT1
6a	H	H	H	H	OH	H	H	OH	160	126
6b	COOH	H	H	H	H	H	H	H	123	75
6c	COOH	H	H	H	Br	Br	Br	Br	0.18	1.4
6d	COOH	H	NH ₂	H	H	H	H	H	710	519
6e	COOH	Br	NH ₂	Br	Br	Br	Br	Br	22.3	4.8
6f	COOH	H		H	H	H	H	H	73	18.9
7a	COOH	H	H	H	H	H	H	H	296	885
7b	COOH	<i>i</i> -Pr	H	Me	<i>i</i> -Pr	H	Me	-	>217	>217
7c	SO ₃ H	H	H	H	H	H	H	-	>263	1050
7d	SO ₃ H	<i>i</i> -Pr	CH ₂ N(CH ₂ COOH) ₂	Me	<i>i</i> -Pr	CH ₂ N(CH ₂ COOH) ₂	Me	-	>34	>34
8									515	280
9									88	92
10a	H								590	928
10b	CONHPh								1080	529
10c	CSNHPH								1360	990
11									220	540
12									103	96
13									490	1220
14									2100	440
15									5.9	18.1
16									1100	1070
17									406	476
18									1220	1160
19									>164	>164
20									>510	>510

or the Arg substrates rPRMT1 experimental bound conformation were effectively reproduced and successfully selected by the X-Score scoring function (see Supporting Information, Figure B). The DOCK⁵⁵ program was also tried as alternative and faster docking routine. Although DOCK performed rather well in reproducing the SAM binding mode (not shown), it was only partially able to correctly pose the Arg substrate (not shown).

In their report, Cheng et al.²² identified three compounds, AMI-1, AMI-5, and AMI-6 (Cheng's paper numbers), as methyltransferase inhibitors, AMI-1 and AMI-6 being selective for PRMT1 while AMI-5 showed to be active against both lysine and arginine methyltransferases. Therefore, initial binding mode analyses were focused on these three compounds, herein reported as **6c** (AMI-5), **6f** (AMI-6), and **9** (AMI-1).

Interestingly, the Autodock conformations selected by X-Score for these three compounds were found docked in different binding sites (Figure 3). In particular, **6c** binds exclusively to the SAM binding pocket whereas **9** and especially **6f** lie between SAM and Arg binding sites. These results are consistent with the outcomes of kinetics experiments performed by Cheng et al.²²

A deeper inspection revealed that **6c** binds mainly to the SAM-adenosine binding pocket formed by Val45, Gly46, Cys47, Gly48, Val67, Asp68, Met69, Ser70, Gly94, Lys95, Met96, Glu97, Glu112, Met123, and Thr126 residues and makes several hydrophobic (Gly46, Met69, Ser70, Met123, and Thr126) and electrostatic (Gly46, Val67, Asp68, Ser70, and Thr126) interactions and a hydrogen bond between the **6c** carboxylate group and the amidic NH of Met69 (2.6 Å, Figure 3c).

As regards to **9**, about half of its structure is buried exclusively in the Arg site (Glu15, Met16, Glu112, Met114, Tyr116, Glu121, His261, and Trp262, Figures 3a and 3b) with a strong hydrogen bond interaction occurring between one of the sulfonic groups and the amidic NH of His261 (Figure 3b, SO₃⁻⋯NH_{His261} = 2.95 Å) and a second one between the ureidic NH and the arginine anchoring Glu112 side-chain (HN₉⋯OOC_{Glu112}

= 2.80 Å, Figure 3b). The other half of the structure of **9** stretches out in the SAM binding site, slightly overlapping with the cofactor methyl donor group and making at least three hydrogen bonds involving the second sulfonic group and Arg22 guanidine side chain (3.08 Å), Gly48 amidic NH (2.75 Å), and Thr49 hydroxyl group (3.22 Å), respectively. Several hydrophobic (Glu15 and Met16), electrostatic (Glu15, Met16, Gly48, Thr49, Glu112, Met114, and His261), and π - π (Try116, His261, and Trp262) interactions are also detectable between **9** and RmtA.

Similarly and to a higher extent than that of **9**, the **6f** binding conformation partially occupies both SAM and Arg sites (Figure 3d). In particular, part of the fused tricyclic moiety is superimposed with the SAM methionine residue, making at least two hydrogen bonds (phenolic-OH_{6f}⋯OOC_{Asp44} = 3.20 Å; endocyclic-O⋯N-guanidine_{Arg22} = 2.74 Å), while the dichlorotriazine portion fills the guanidine binding site of the Arg substrate, making positive contacts with Trp113, Met114, Gly115, Glu121, Ser122, Met123, and Trp262. A further hydrogen bond interaction is also made between carboxylic group of **6f** and the Gly48 amidic NH (2.89 Å). Several hydrophobic (His14, Met16, Thr49, Glu112, Trp113, Try116, Ser122, and Met123) and electrostatic (Asp44, Gly46, Gly48, Thr49, Glu112, Met114, and Glu121) interactions are also observed between **6f** and RmtA.

Docking Studies and Binding Mode Analysis of Other Compounds. On the basis of biological results, we analyzed the binding mode of compounds showing a measurable inhibitory activity (20 molecules).⁵⁶ We divided the selected docked conformations into three groups: (a) molecules docked in the Arg pocket (DAP: **6a**, **12**, **13**, **16b**,⁵⁷ and **17**, Figure 4a); (b) molecules docked in the SAM pocket (DSP: **6b**, **6c**, **6d**, **10a**, and **18**, Figure 4b), and (c) molecules partially overlapping with both sites (docked in both pockets, DBP: **6e**, **6f**, **7a**, **8**, **9**, **10b**, **10c**, **14**, and **16a**, Figure 4c). Derivative **11** was found to dock only outside either the SAM or Arg binding sites.

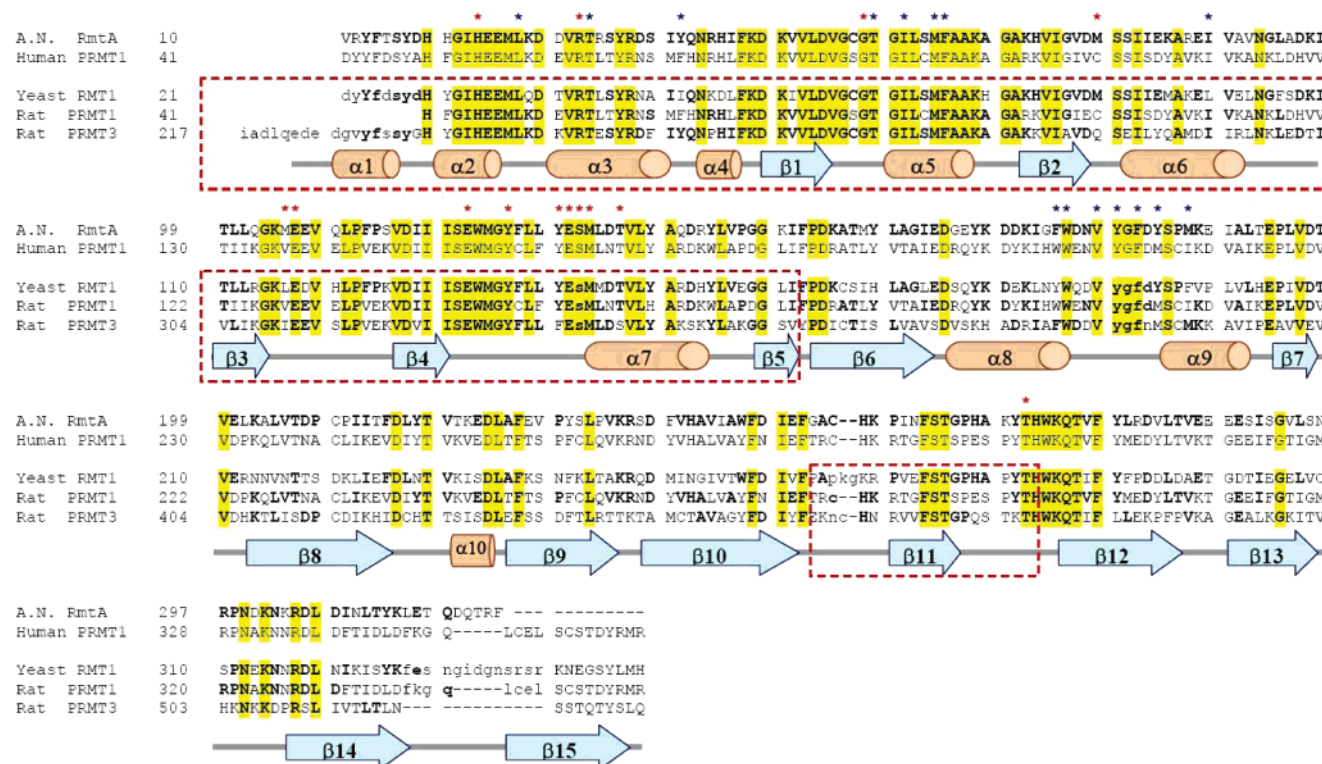


Figure 1. Sequence alignment of the target *A. nidulans* (A.N.) RmtA and human PRMT1 to the structural alignment of yeast RMT1, rat PRMT1, and rat PRMT3. Cylinders and arrows indicate the common secondary structural elements (helices and strands, respectively) in the template structures. Red boxes point out the catalytic domain regions. Residue numbering is indicated on the left of each sequence row. Small letters indicate residues which are not structurally aligned. Identical residues between target and templates are bold. Residues highlighted in yellow are identical in all the sequences. Red asterisks indicate residues having a role in cofactor binding and/or catalytic mechanism; blue asterisks indicate residues important for the dimerization process.

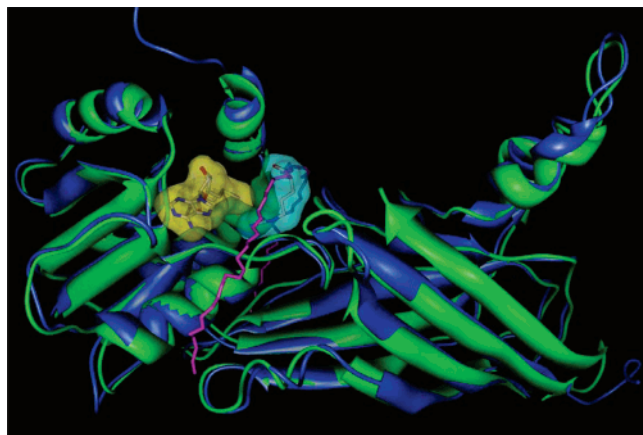


Figure 2. Superimposition of RmtA homology model (blue) and rPRMT1 X-rays structure (green). The SAM binding site, the Arg binding site, and the backbone of histone H4 are highlighted in yellow, cyan, and purple, respectively.

As illustrated in Figure 4a, the molecules of the DAP group occupy the Arg pocket plus a portion of the area surrounding the entry site of the substrate side chain (not shown). The formation of a hydrogen bond with the Glu15 side-chain carboxyl is the most common interaction in this group, noticeable for three compounds (**6a**, **16b**, and **17**) out of five. Among them, **6a**, the second most active molecule of the set, also makes positive interactions with the hydroxyl of Tyr116 and amidic nitrogen of His261. Moreover, the phenyl group seems to make a π - π interaction with the indole ring of Trp262, while the

tricyclic portion is lying between Met16 and His261 side chains (Supporting Information, Figure C).

On the other hand, derivative **12**, the most active of the DAP group, makes important hydrogen bonding interactions between its sulfonic group and the guanidine moiety of Arg295. In addition the unsubstituted naphthyl ring stacks between the Trp262 indole and the Tyr116 phenol rings making π - π interactions. (Supporting Information, Figure D). As regards to the less active derivatives **13**, **16b**, and **17**, although a few relevant interactions (i.e., hydrogen bond with Glu15) are present, as in the case of **6a** and **12**, major hydrophobic and electrostatic interactions are lacking (Supporting Information, Figure E). Additional docking studies performed in the presence of the SAM cofactor revealed similar results with differences only for compounds **12** and **17**. (See Supporting Information, Figure F): derivative **12** displays a different binding pose, while **17** is shifted toward SAM and the negatively charged carboxylate seems to make electrostatic interaction with the positively charged SAM methyl group. No relevant interactions are made between **6a**, **13**, **12**, and **16b** with SAM.

As regards to compounds of DSP group, it is noteworthy that the two similar derivatives **6b** and **6d** display a different binding scenario from the one above-described for **6c**, as they are shifted toward the Arg site and while losing the interactions with amidic nitrogen Asp68, they gain some interactions with key residues at the interface between the SAM and Arg sites (Glu112 or Asp68 carboxylate for **6b** or **6d**, respectively; see Supporting Information, Figure G). Compounds **10a** and **18** occupy the same space but, though making a few important interactions such as a hydrogen bond with the Thr49 hydroxyl, most

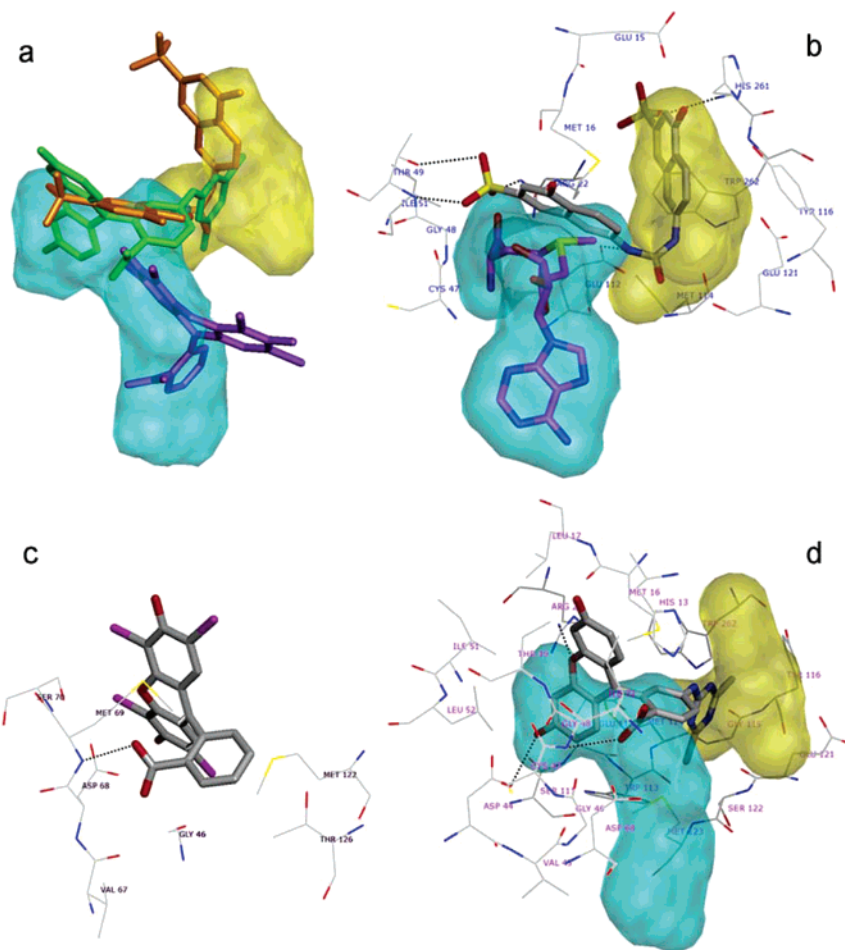


Figure 3. Autodock/X-Score selected binding conformations of compounds **6c**, **6f**, and **9** into RmtA catalytic site. (a) Superimposed conformations of **6c** (purple), **6f** (green), and **9** (orange); (b) binding mode of **9** (in stick, carbon atoms in gray); (c) binding mode of **6c** (in stick); (d) binding mode of **6f** (in stick). The volumes occupied by Arg and SAM are represented in yellow and cyan, respectively. The RmtA residues within 4.0 Å from the docked inhibitor are shown in wire. For the sake of clarity, hydrogen atoms are not displayed.

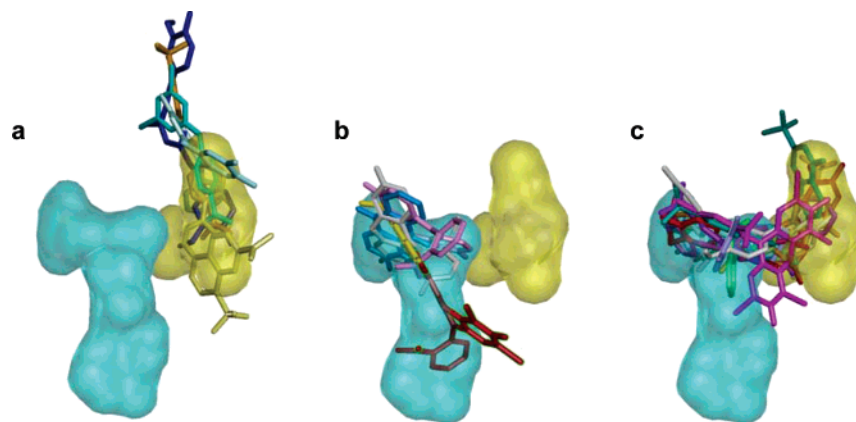


Figure 4. Binding conformations of tested molecules, divided according to the binding modes of **6c**, **6f**, and **9** (see text). In yellow and cyan are represented the Arg and SAM sites, respectively. (a) **6a** (aquamarine), **12** (yellow), **13** (orange), **16b** (red) and **17** (cyan); (b) **6b** (orchid), **6c** (red), **6d** (light gray), **10a** (yellow), and **18** (sky blue); (c) **6e** (magenta), **6f** (light gray), **7a** (green), **8** (cornflower blue), **9** (sky blue), **10b** (purple), **10c** (orange), and **16a** (red).

hydrophobic interactions are lacking due to their smaller size (Supporting Information, Figure H).

The docking area of the third group of molecules (DBP), partially overlapping both Arg and SAM sites, indicates a different interaction profile. Taken together, the molecules of this group make contacts within 3 Å that define a binding site comprising at least 18 residues (Ile12, His13, Glu15, Leu17, Arg22, Gly46, Cys47, Gly48, Thr49, Leu52, Asp68, Glu112, Met114, Gly115, Tyr116, Glu121, His261, and Trp262, Sup-

porting Information, Figure I). Among these residues, Arg22 (**6f**, **10b** and **14**), Gly46 (**6f**, **7a**, **10**, and **14**), Glu112 (**6f**, **7a**, **8**, **9**, **10b**, **10c**, **14**, and **16a**), Met114 (**6f**, **7a**, **10b**, and **10c**), and His261 (**9**, **10c**, and **16a**) are the most involved in ligand binding. The most active compound **6e** shows peculiar hydrogen bonding interactions: a strong one with the hydroxyl group of Tyr116 (3.07 Å) and a weak one with His13 (3.73 Å) (not shown). However, the strong point of **6e** seems to be the number of positive contacts that the six bromine atoms make in mostly

Table 2. Statistical Parameters of the Three 3-D QSAR Models

model	group	probe	no. ^a	variables ^b	PCs ^c	r ²	q ² _{LOO}	SDEP _{LOO}
1	DAP	OH2-C3	5	1240	2	0.99	0.88	0.12
2	DSP	O	5	2641	2	0.99	0.91	0.40
3	DBP	OH2-C3	9	1063	2	0.99	0.93	0.17

^a Number of molecules in the group. ^b Number of GRID variable after the FFD selections. ^c Number of principal components.

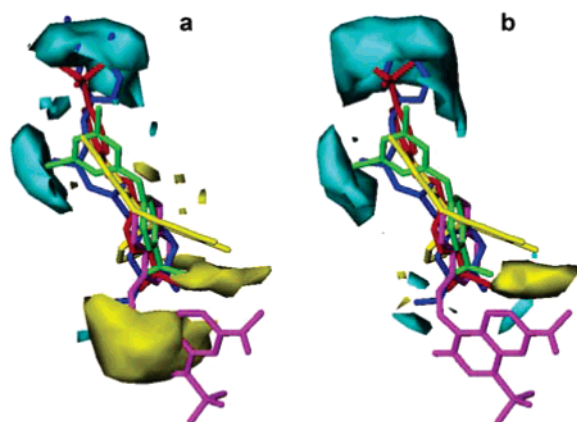


Figure 5. DAP group grid plots of the PLS coefficients for the OH2 (a) and C3 (b) probes. Compounds **6a** (yellow), **12** (magenta), **13** (red), **16b** (blue), and **17** (green) are also displayed. The cyan contours represent negative coefficients under -0.0003 energy value while the yellow contours represent the positive coefficients over 0.0003 .

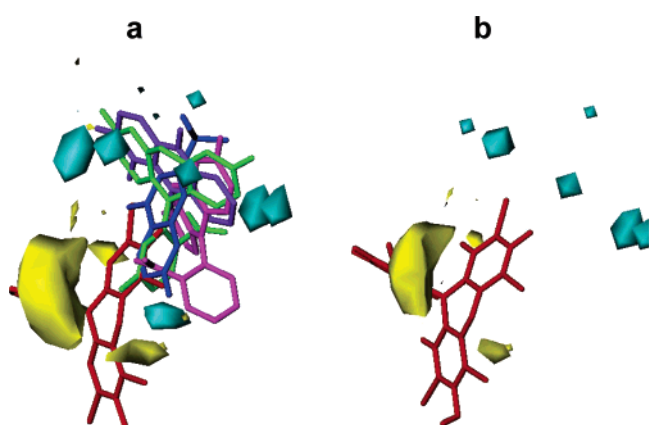


Figure 6. (a) DSP group grid plot of the PLS coefficients. Compounds **6b** (magenta), **6c** (red), **6d** (green), **10a** (blue) and **18** (purple) are also displayed. The cyan contours represent negative coefficients under -0.0005 energy value while the yellow contours represent the positive coefficients over 0.0015 . (b) Activity contribution plot relative to compound **6c**.

filling the Arg binding area and part of SAM binding area (Supporting Information, Figure J). The less active derivatives **7a**, **8**, **10b**, **10c**, **14**, and **16a** though showing docking positions similar to that of **6e** are smaller in volume (data not shown) and make less positive contacts.

In addition to the above-mentioned docking studies, we also generated the human PRMT1 (*hPRMT1*) using the same protocol described for RmtA homology modeling. Subsequently, we performed a 3-D comparison of the RmtA and *hPRMT1* models and we found that only five (13.5%) of the 37 residues that are in contact with the SAM and arginine substrates had different side-chains, and more interestingly, four out of five showed some similarities (Cys47_{RmtA} → Ser39_{hPRMT1}; Val67_{RmtA} → Ile59_{hPRMT1}; Asp68_{RmtA} → Val60_{hPRMT1}; Met69_{RmtA} → Cys61_{hPRMT1}; Met96_{RmtA} → Val88_{hPRMT1}; see Supporting Information, Figure K). Once the *hPRMT1*/SAM/H4 ternary

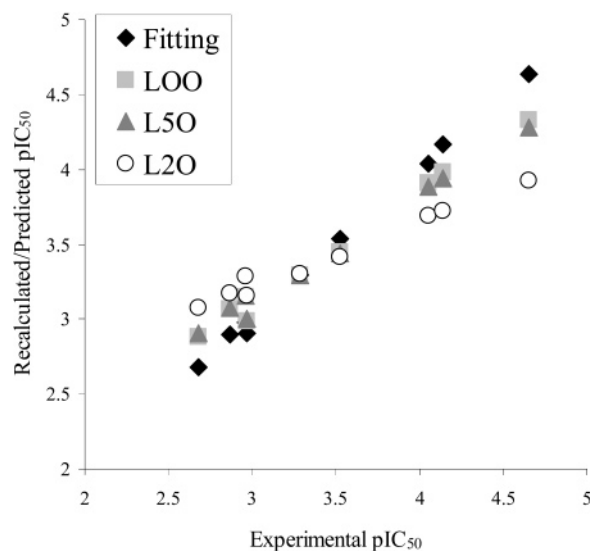


Figure 7. Recalculated and predicted pIC_{50} from model 3. LOO: cross-validation with the leave one out method; L5O: cross-validation with the leave some out method using five groups; L2O: cross-validation with the leave some out method using two groups.

complex was geometry optimized we applied the Autodock/X-SCORE protocol to inspect the binding site of derivatives **6–20** (data not shown). As expected the binding modes of the compounds in the *h-PRMT1* were similar to those described for RmtA with only marginal differences in the structure conformations, consistent with the aforementioned good correlation between pIC_{50} values.

3-D QSAR Studies. As learned from the foregoing docking studies, the three different binding modes adopted by molecules **6–20** could represent structure-based aligned training sets to build 3-D QSAR models useful as validation tools for the docking studies themselves as well as tools to design new PRMT inhibitors. With this aim, three 3-D QSAR models were derived using the GRID/GOLPE procedure.⁵⁸ Obviously, it has to be stressed that due to the limited number of molecules in each group the 3-D QSAR models are of limited validity and were only built to gain qualitative support to the above-described binding modes and to get some insights to design new PRMT inhibitors.

In building the 3-D QSAR models, different GRID probes alone (C3, N1, O, OH, OH2, and DRY) and combinations of them (C3+N1, DRY+N1, DRY+O, O+C3, O+N1, OH2+C3, OH2+DRY, OH+C3, OH+DRY, and OH+N1) were tried, and only the preliminary GOLPE/GRID models showing the higher q^2 values were refined by sequential fractional factorial design (FFD) selections. The combination of the water (OH2) and the methyl (C3) probes gave the best preliminary models for the DAP and DBP groups while only the carbonyl oxygen (O) probe supplied the best model for the DSP group alignment. After having imported the molecular interaction fields (MIF) into GOLPE, a series of fractional factorial design (FFD) selections allowed the definition of the final 3-D QSAR models whose statistical parameters are listed in Table 2. As can be seen, in

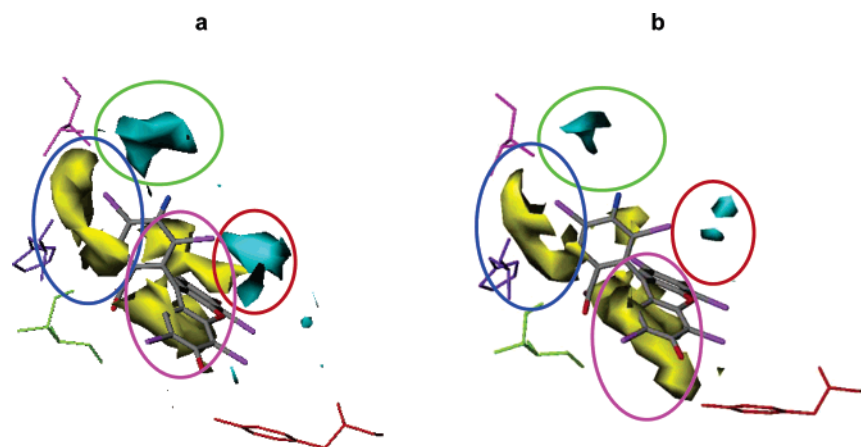


Figure 8. DBP group grid plots of the PLS coefficients for the OH2 (a) and C3 (b) probes. The cyan contours represent negative coefficients under -0.0006 energy value while the yellow contours represent the positive coefficients over 0.0006 . Aminoacids Ile12 (green), His13 (purple), Thr49 (magenta), and Tyr116 (red) are also displayed. Hydrogen atom are not displayed.

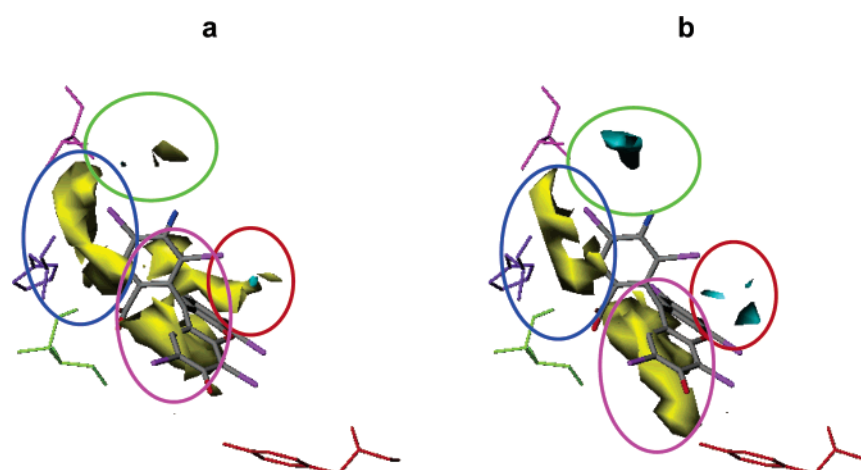


Figure 9. Derivative **6e** activity contribution plots for the OH2 (a) and C3 (b) probes. The **6e** docked conformation is displayed as atom type colors. The cyan contours represent negative activity contribution areas while the yellow contours represent the positive activity contribution areas. Amino acids Ile12 (green), His13 (purple), Thr49 (magenta), and Tyr116 (red) are also displayed. Hydrogen atoms are not displayed.

term of cross-validation data (q^2 and SDEP values), all three models, although developed with a limited number of compounds, were statistically robust with q^2 and SDEP values ranging from 0.88 to 0.93 and from 0.12 to 0.40, respectively. Although in models 1 and 2 the molecules seem not to be optimally aligned, the 3-D QSAR models seem to support the above-described binding modes obtained by docking studies.

As already outlined in the binding mode analysis section, only derivative **12**, the most active out of the DAP group, is fully buried in the arginine substrate site, which is correctly recognized by the PLS analysis of the OH2 and C3 GRID fields. In particular in the OH2 PLS coefficient plot (Figure 5a) a large positive yellow polyhedron runs over the **12** diaza group which is close to the guanidine group of the arginine substrate. Although in the above binding mode analysis the **12** diaza group did not seem to make any significant interaction, the PLS analysis of both the actual field and activity contribution plots (Supporting Information, Figure L) indicates that a π -rich group should be retained in designing new compounds competitive for the arginine site. This consideration is also supported by the observation that less active molecules such as **13** and **16b** possessed hydrogen bonding acceptor groups in the same **12** diaza moiety area. To prevent the model from being guided by the misalignment of **12**, we also built a 3-D QSAR model with only four compounds, and, taking into account the statistical

limits of such a model, the final 3-D QSAR model was characterized with r^2 and q^2 values of 0.99 and 0.84, respectively.

Regarding the 3D-QSAR model 2, it was built using only five molecules which were docked exclusively in the SAM binding site; derivative **6c**, the most active among all tested molecules, belongs to this group. As in the case of compound **12** in model 1, **6c** seems misaligned (Figures 4b and 6), yet a 3-D QSAR model endowed with high r^2 (0.99) and q^2 (0.68) values without compound **6c** was also obtained, confirming that model 2 reflects a structure–activity correlation belonging to all five molecules in the set.

In Figure 6a is reported the PLS coefficients plot of the 3-D QSAR model for the DSP group in which one large yellow polyhedron (positive values) and a few small cyan polyhedrons (negative values) are present. To properly understand the role of the yellow polyhedron, the activity contribution plot for compound **6c** is also reported (Figure 6b). As shown, the most important contribution to the activity seems due to the benzoate group while no polyhedron can be associated to the bromine atoms, the importance of which still emerged from biological assays. This disagreement could be attributed to the fact that the SAM binding site is not fully represented by the aligned molecules, and thus the 3-D QSAR model here only supports the hypothesis that there is some correlation between the inhibitory activity with both the bound conformations and their spatial disposition (alignment).

The 3-D QSAR model 3 was developed on molecules partially occupying both the SAM and Arg binding sites (DBP, Figure 4c). Differently from the above analyzed models, all nine molecules of this group were found superimposed to each other, leading to a robust model (Figure 4c and Supporting Information, Figure M), as also confirmed by high q^2 values obtained by cross-validations using the group method (leave some out, LSO): with five (L5O, $q^2 = 0.90$) and two (L2O, $q^2 = 0.61$) groups the DBP molecules were highly self-predictive (Figure 7).

Analysis of the PLS coefficient plots (Figure 8) revealed that some areas around the structure-based aligned molecules should be further explored by the inclusion of additional ligands. In particular, in both the OH2 and the C3 PLS coefficients plot (Figure 8a and 8b, respectively) at 0.0006 energy level a large yellow polyhedron is noticeable close to the tricyclic moiety of **6e** (magenta circle). This polyhedron is not correlated with any residue of RmtA but lies in an open space, maybe the entrance of SAM. A second yellow polyhedron (blue circle) could be associated to a space close to the dibromoaminobenzoate moiety of **6e** corresponding to a pocket formed by Ile12, His13, Met16, and Thr49 residues. It is noteworthy that this area is also filled by **6f** and **9**, that together with **6e** are the most potent compounds of the DBP group. Two other cyan-colored polyhedrons (negative PLS coefficients) are visible in Figure 8a (OH2 probe) and almost absent in Figure 8b (C3 probe): the polyhedrons highlighted by green circles, mainly associated either with electrostatic (**7a**, **8**, **10b**, and **16a**) or hydrogen bonding (**6e**, **6f**, and **9**) interactions with Arg22 (not shown) and the polyhedrons circled in red, mainly associated either with electrostatic (**6f**, **10c**, and **14**, not shown) or hydrogen bonding (**7a**, **8**, **9**, **10b**, and **10c**, not shown) interactions with either Glu112 side-chain carboxyl (**6f**, **9**, **10b**, **10c**, and **14**, not shown) or Met114 carbonilic oxygen (**6f**, **7a**, **8**, **9**, **10b**, and **10c**, not shown). The negative PLS coefficients generated by a methyl probe (small cyan polyhedrons in Figure 8b) indicate areas that have to be filled to maintain or increase the activity. One of these areas coincides with the binding site of the cysteinic portion of SAM, a site delimited by Arg22, Asp44, Gly46, Cys47, Ile51, Leu52, and Glu112 (not shown).

The aforementioned important role of the interactions of compounds of the DBP group with the pocket composed of Ile12, His13, Met16, and Thr49 residues is more clearly explained in Figure 9, reporting the activity contribution plots relative to derivative **6e**, the most active in this group. In this figure, yellow polyhedrons, associated with positive activity contribution areas, are closely related to this pocket. Again the small cyan polyhedrons, associated with negative activity contribution areas, highlight regions lacking in positive sterical interactions (green and red circles in Figure 9b).

Conclusions

In this report we describe the binding mode analysis of a focused library of small molecule compounds into the catalytic domain of the PRMT1 fungal homologue RmtA as well as enzymatic assays performed on recombinant RmtA and PRMT1 proteins using SAM and histones as substrates. Computational and biological results were in good agreement. Moreover, the good correlation between pIC_{50} values against the two enzymes ($R = 0.91$) supported our hypothesis of using RmtA as a model to screen for arginine methyltransferase inhibitors.

Moreover, the agreement between molecular modeling studies (molecular docking and 3-D QSAR) and biological results allowed us to shed more light on the molecular mechanism of

inhibition of small molecule inhibitors of histone arginine methyltransferases. In particular the structure-based 3-D QSAR high statistical coefficients seem to support the docking experiments. If considered together, the docking and the 3-D QSAR studies allowed a deep binding mode analysis of tested molecules. In addition, these analyses hinted that two regions, the pocket formed by Ile12, His13, Met16, and Thr49 and the SAM cysteinic portion binding site, should be taken into account to design novel active inhibitors.

Further molecular modeling, synthetic and biological efforts are ongoing to develop more robust 3-D QSAR models.

Materials and Methods

Melting points were determined on a Gallenkamp melting point apparatus and are uncorrected. 1H NMR spectra were recorded at 300 MHz on a Bruker Avance 300 spectrometer; chemical shifts are reported in δ (ppm) units relative to the internal reference tetramethylsilane (Me_4Si). Electronic impact mass spectrometry (EI-MS) was performed on a Finnigan LCQ DECA TermoQuest (San José, CA) mass spectrometer. All compounds were routinely checked by TLC and 1H NMR. TLC was performed on aluminum-backed silica gel plates (Merck DC-Alufolien Kieselgel 60 F₂₅₄) with spots visualized by UV light or using a $KMnO_4$ alkaline solution. All solvents were reagent grade and, when necessary, were purified and dried by standard methods. Concentration of solutions after reactions and extractions involved the use of a rotary evaporator operating at a reduced pressure of ca. 20 Torr. Organic solutions were dried over anhydrous sodium sulfate. Analytical results are within $\pm 0.40\%$ of the theoretical values. Reagents were purchased from Lancaster Synthesis, Milan (Italy), or Sigma-Aldrich, Milan (Italy). As a rule, samples prepared for physical and biological studies were dried in high vacuum over P_2O_5 for 20 h at temperatures ranging from 25 to 110 °C, depending on the sample melting point.

Preparation of 4-Amino-3,5-dibromo-2-(2,4,5,7-tetrabromo-6-hydroxy-3-oxo-3H-xanthen-9-yl)benzoic Acid (6e). A stirred mixture of 6-aminofluorescein (0.576 mmol, 0.2 g) in glacial acetic acid (7 mL) was cooled at 0 °C. A solution of bromine (4.606 mmol, 0.74 g) in acetic acid (3 mL) was then added dropwise, and stirring was continued at room temperature for additional 15 min. After completion (TLC monitoring: $CHCl_3$:AcOEt 1:1, on silica gel plate), the reaction mixture was diluted with water (10 mL) and filtered by vacuum filtration. The precipitate was washed with water (3×5 mL) and air-dried in a vacuum desiccator to give TLC pure **6e** in 91% yield (0.82 g; mp: 291–292 °C), MS (EI, 70 eV) m/z 821; 1H NMR (DMSO- d_6): δ 6.53 (s, 1H), 7.10 (s, 1H), 8.07 (s, 1H), 10.79 (s, 1H).

Preparation of 4-(4,6-Dichloro-1,3,5-triazin-2-ylamino)-2-(6-hydroxy-3-oxo-3H-xanthen-9-yl)benzoic Acid (6f).³⁰ A solution of 6-aminofluorescein (0.288 mmol, 0.1 g) in acetone (2 mL) was added dropwise to a solution of cyanuric chloride (0.2851 mmol, 0.053 g) in acetone (3 mL) magnetically stirred at 5 °C. The resulting mixture was kept stirring for 3 h at 3–5 °C (TLC monitoring: AcOEt on silica gel plate) and then filtered to give **6f** as a TLC pure yellow precipitate (0.13 g, yield, 92%; mp: > 350 °C); MS (EI, 70 eV) m/z 494; 1H NMR (DMSO- d_6): δ 6.79–6.88 (m, 4H), 6.95 (s, 2H), 7.73 (s, 1H), 8.18–8.28 (m, 2H), 11.72 (s, 1H).

Preparation of 7,7'-Carbonylbis(azanediy)bis(4-hydroxynaphthalene-2-sulfonic acid) (9).³¹ J-acid (2.00 g, 3.964 mmol) was added to 50 mL of water, with dropwise addition of aqueous NaOH 10% to facilitate dissolution; the pH of the final solution was about 7. The solution was then stirred and heated to 60 °C and triphosgene powder (0.40 g, 1.348 mmol) rapidly added. The resulting mixture was magnetically stirred for 6 h while keeping temperature (60 °C) and pH (7–8) constant. When the reaction was complete, J-acid urea **9** was obtained by salting-out. (1.80 g, Yield 90%). MS (EI, 70 eV) m/z : 504; 1H NMR (DMSO- d_6): δ 7.04 (m, 1H), 7.51 (m, 1H), 7.55 (m, 1H), 8.00 (m, 1H), 8.02 (m, 1H), 9.27 (s, 1H), 9.61 (s, 1H).

Preparation of 4-Hydroxy-7-(3-phenylureido)naphthalene-2-sulfonic Acid (10b).³² A stirred solution of J acid (1.943 mmol, 0.5 g) in 8 mL of alkaline water solution containing NaOH (1.943 mmol, 0.077 g), at room temperature, was treated with a solution of phenyl isocyanate (1.943 mmol, 0.23 g) in tetrahydrofuran (5 mL). The resulting mixture was magnetically stirred at room temperature for 24 h (TLC monitoring: CHCl₃ on silica gel plate), acidified with a water solution of HCl 2 N (3 mL), evaporated under reduced pressure, and purified by column chromatography (silica gel/ethyl acetate:isopropanol: water 12:3:1). (0.49 g, Yield, 70%; mp: > 350 °C); MS (EI, 70 eV) *m/z* 358; ¹H NMR (DMSO-*d*₆): δ 6.92–6.98 (m, 2H), 7.24–7.29 (t, 2H), 7.42–7.48 (m, 4H), 7.85 (s, 1H), 7.95–7.98 (d, 1H), 8.72 (s, 1H), 8.93 (s, 1H), 10.02 (s, 1H).

Preparation of 4-Hydroxy-7-(3-phenylthioureido)naphthalene-2-sulfonic Acid (10c). A stirred solution of J-acid (1.943 mmol, 0.5 g) in 8 mL of alkaline water solution containing NaOH (1.943 mmol, 0.077 g), at room temperature, was treated with a solution of phenyl isothiocyanate (1.943 mmol, 0.278 g) in tetrahydrofuran (5 mL). The resulting mixture was magnetically stirred at room temperature for 12 h (TLC monitoring: CHCl₃ on silica gel plate), acidified with a water solution of HCl 2 N (3 mL), evaporated under reduced pressure, and purified by crystallization (CH₃CN) to yield TLC pure **10c** (0.55 g, Yield: 75%; mp: > 350 °C); MS (EI, 70 eV) *m/z* 374; ¹H NMR (DMSO-*d*₆): δ 7.05–7.10 (m, 2H), 7.31–7.34 (t, 2H), 7.47–7.57 (m, 4H), 7.83–7.84 (s, 1H), 7.97–8.00 (d, 1H), 9.82 (s, 1H), 9.95 (s, 1H), 10.08 (s, 1H).

Preparation of 4-(4,6-Dichloro-1,3,5-triazin-2-ylamino)benzoic Acid (17).³³ A solution of 4-aminobenzoic acid (3.656 mmol, 0.5 g) in acetone (10 mL) was added to a solution of cyanuric chloride (3.61 mmol, 0.67 g) in acetone (12 mL) magnetically stirred at 0 °C. Aqueous sodium hydrogenocarbonate 0.3 M (12 mL) was then added to the mixture over 30 min while maintaining a temperature of 0–5 °C. The solution was stirred for 2 h (TLC monitoring: CHCl₃:ethyl acetate 1:1, on silica gel plate), and then aqueous HCl 12% was added until a precipitate formed which was filtered and washed with acetone. (0.99 g, Yield: 95%; mp: > 350 °C); MS (EI, 70 eV) *m/z* 284; ¹H NMR (DMSO-*d*₆): δ 7.67–7.70 (d, 2H), 7.78–7.95 (d, 2H), 11.37 (s, 1H).

Cloning of Human PRMT1 and *Aspergillus nidulans* RmtA. Human PRMT1 (isoform 2) was amplified from a HeLa cDNA library (Invitrogen) using the following primer set: PRMT1fwd 5'-GGATCCGAGAATTTTGTAGCCACCTTGGCTAA-3' and PRMT1rev 5'-CTCGAGTCAGCGCATCCGGTAGTCGGTGGA-3'. The PCR fragment was subcloned into a pGEM-Teasy vector (Promega), released by digestion with the restriction endonucleases BamHI and XhoI (New England Biolabs), and inserted into the BamHI/XhoI sites of the bacterial expression plasmid pGEX6P1 (Amersham) using T4 Ligase (Roche).

Total *Aspergillus* RNA was used to prepare cDNA using Superscript reverse transcriptase (Life Technologies, Inc.) according to the manufacturer's instructions employing an oligo-dT₁₇ standard primer. For amplification of fragments of *rmtA* from the cDNA of *A. nidulans*, degenerate oligodeoxynucleotide primers were used. Forward primer was 5'-GGNATHCAYGARGARATG-3' based on the amino acid sequence GIHEEM, and reverse primer was 5'-ACNCAYTGGAARCARAC-3' based on THWKQT. A product of 750 bp was isolated and cloned into pGEM-T vector (Promega). A full-length genomic copy of *rmtA* was obtained using cDNA PCR products to screen a genomic library of *Aspergillus nidulans* DNA. For amplification of the 3' end of *rmtA*, the 3' rapid amplification of cDNA ends (RACE) protocol⁵⁹ was performed. The first PCR was performed with forward primer 5'-ATACCGTTC-GAGCTCAAG-3' and an adapter primer of sequence 5'-GACTC-GAGTCGACATCGA-3'. For "nested PCR", primer 5'-GACTA-GAGTACACGATGGA-3' (forward) and dT4 adapter primer 5'-GACTCGAGTCGACATCGATTTT-3' (reverse) were used. Amplified products were cloned into pGEM-T vector (Promega).

Preparation of GST-RmtA and GST-PRMT1 Fusion Proteins. Expression of the GST-PRMT1 fusion protein was performed in BL21 cells. At OD₆₀₀ nm = 0.6 expression was induced with 1

mM isopropyl β-D-thiogalactopyranoside (IPTG), and the cultures were grown for 3 h upon induction. Bacterial cells were harvested by centrifugation (3000g for 20 min), resuspended in sonication buffer (PBS, 0.01% TritonX100, 0.2 mM PMSF, 1 μg/mL aprotinin, 1 μg/mL leupeptin, 1 μg/mL pepstatin), and sonicated three times for 30 s with a sonic dismembrator 550 (Fisher Scientific; amplitude 40%). The sonicated suspension was centrifuged for 20 min at 10000g, and the supernatant was incubated with preequilibrated glutathione-sepharose-4-fast-flow overnight at 4 °C on a rotating wheel. The resin was transferred to a gravity-flow column and washed with five column volumes (CV) of PBS + 0.1% TritonX100. Bound GST-PRMT1 was released using five times one CV of elution buffer (50 mM Tris pH 8.0, 10 mM glutathione). Purified GST-PRMT1 was dialyzed against BC100 buffer (20 mM Tris pH 7.9, 10% glycerol, 100 mM KCl, 0.2 mM EDTA, 10 mM 2-ME, 0.2 mM PMSF), and protein concentration was estimated by SDS-PAGE and Bradford assay.

The coding sequence of RmtA was cloned into a pGEX-5X-1 expression vector (Amersham Pharmacia Biotech). RmtA-Protein was expressed in BL21 cells in LB-medium. Cultures (250 mL) with an A₆₀₀ of 0.4 were induced with a final concentration of 1 mM IPTG and grown for 4 h at 37 °C. After centrifugation of cells at 4000g, the pellet was resuspended in 6 mL of GST-binding buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.3) containing one protease inhibitor tablet (Complete, Roche, Mannheim, Germany) per 50 mL of buffer. For cell lysis, lysozyme was added at a final concentration of 5 mg/mL binding buffer, and cells were passed through a french press with a pressure setting of 1000 psi. The resulting lysate was centrifuged at 20000g for 10 min at 4 °C. GST fusion protein was purified from soluble extracts by binding to a GST-HiTrap column (Amersham Pharmacia Biotech). Proteins were eluted with 50 mM Tris-HCl, 10 mM reduced glutathione, pH 8.0, and assayed for histone methyltransferase activity.

Histone Methyltransferase Assay (HMT). For inhibition assays, affinity purified GST-RmtA and GST-PRMT1 fusion proteins were used as the enzyme source. HMT activities were assayed as described²⁹ using chicken erythrocyte core histones as substrate. A 500 ng amount of GST-RmtA and GST-PRMT1 fusion proteins was incubated with different concentrations of compounds for 15 min at room temperature, and 20 μg of chicken core histones and 0.55 μCi of [³H]-S-adenosyl-L-methionine (SAM) were added. This mixture was incubated for 30 min at 30 °C. Reaction was stopped by TCA precipitation (25% final concentration), and samples were kept on ice for 20 min. Whole sample volumes were collected onto a glass fiber filter (Whatman GF/F) preincubated with 25% TCA. Filters were washed three times with 3 mL of 25% TCA and then three times with 1 mL of ethanol. After the filters were dried for 10 min at 70 °C, radioactivity was measured by liquid scintillation spectrophotometry (3 mL of scintillation cocktail).

Preparation of the RmtA-SAM-H4 Ternary Complex. The RmtA model from the homology model experiment was merged with the S-adenosyl-L-homocysteine (AdoHcy) and the 19mer substrate peptide taken from the experimental PRMT1 structure (pdb entry code 1or8).⁴⁷ The AdoHcy was modified to SAM using an experimental SAM structure extracted from the FTSJ RNA methyltransferase (pdb entry code 1ej0)⁶⁰ which showed a SAM conformation similar to the AdoHcy found in the 1or8 complex. The experimental 19mer peptide lacked of most of the side-chain atoms, which were added using the xLeap module of the AMBER suite.⁶¹ The RmtA-SAM-H4 complex was solvated (SOLVATEOCT command) in a box extending 10 Å with 16658 water molecules (TIP3 model) and neutralized with 10 Na⁺ ions. The solvated complex was then refined by minimization using the SANDER module of AMBER. The parameters for the SAM were calculated using the antechamber module of AMBER, and the atomic charges were calculated using the AM1 Hamiltonian.

Molecular Docking. All the tested molecules were built, starting from ASCII text, using the stand alone version of PRODRG,⁶² in conjunction with the GROMACS suite.⁶³ Docking studies were performed by means of the Autodock 3.0.5 program using a grid

spacing of 0.375 Å and $39 \times 50 \times 56$ number of points that embraced both the SAM and Arg binding sites. The grid was centered on the mass center of the experimental bound AdoMet and Arg substrates. The GA-LS method was adopted using the default setting except for the maximum number of energy evaluations increased from 250000 to 2500000. Autodock generated 100 possible binding conformations for each molecule that were clustered using a tolerance of 2.0 Å. The AutoDockTool (ADT) graphical interface⁶⁴ was used to prepare the enzyme PDBQS file. The protein atom charges as calculated during the complex minimization were retained for the docking calculations.

The SAM and *N*-acetyl,*O*-methyl-capped Arg PRODRG-generated conformations were docked into the RmtA to assess the docking protocol. The SAM and Arg were docked back into their binding sites, and the Xscore⁵⁴-selected conformations showed root-mean-square deviations (rmsd) of 0.60 and 1.47, respectively. The Autodock scoring function did not select conformations with lower rmsd values. Trials with the DOCK program did not successfully dock back the substrates into the RmtA; therefore, we continued to use Autodock. To analyze the docking results, the ADT was used and the Chimera 1.2176 program⁶⁵ was used to produce the images.

Docking studies were not performed on nonactive compounds (**7b**, **7c**, **7d**, and **20**) and on compound **15** due to its large dimensions.

3-D QSAR Studies. GRID Calculations. The interaction energies were calculated by using the program GRID⁶⁶ (version 22) with a grid spacing of 1 Å and the grid dimensions (Å): X_{\min}/X_{\max} , 58.25/20.25; Y_{\min}/Y_{\max} , 69.25/31.25; Z_{\min}/Z_{\max} , 75.25/37.25.

GOLPE Analyses. PLS models were calculated with GOLPE 4.5.12⁶⁷ running on a SGI O2 R10000 equipped with the IRIX operating system 6.5.11. To measure the goodness of the model the statistical indices r^2 , q^2 , and SDEP were employed.

Probe Selection. In this study we started with five probes (C_{sp^3} , DRY, OH2, NH_{amide} , O_{sp^2}) and combinations of them, according to the nature of the RmtA active site. Preliminary PLS analyses were run on the DSP, DAP, and DBP groups, and the probe combination displaying higher values of r^2 and q^2 was selected to develop the final model for each group.

Variable Preselection. The resulting probe-target interaction energies for each compound were unfolded to produce the corresponding one-dimensional vector variables, which were assembled in the so-called X matrix. This matrix was pretreated first by using a cutoff of 5 kcal/mol to produce a more symmetrical distribution of energy values and then zeroing small variable values and removing variables with small standard deviation, using appropriate cutoffs. In addition, variables taking only two and three distribution were also removed.

Smart Region Definition (SRD). A number of seeds (3000) were selected using a D-optimal design criterion in the weight space. Structural differences between different molecules in the series will be reflected in groups of variables, and therefore groups were generated around each seed in the 3-D space. Variables with a distance of no more than 1 Å to the seeds were included in the groups. If two neighboring groups (with a distance smaller than 2 Å) contained the same information the groups were collapsed. The groups were used in the variable selection procedure replacing the original variables. The effect of the groups on the predictivity was evaluated, and groups instead of individual variables were removed from the data file.

Region Selection. The effect of the grouped variables on the predictivity was evaluated using a fractional factorial design (FFD) procedure. A number of reduced models (twice the number of variables) were built removing some of the variables according to the FFD design. The effect of dummy variables (20%) on the predictivity was calculated, and only if a variable had a positive effect on the predictivity larger than the effect of the average dummy variable was included in the final model. The FFD selection was repeated until the r^2 and q^2 value did not increase significantly. In

the FFD selection the cross-validation was conducted using five random groups for 20 times and a maximum of two principal components.

Cross-Validation. The models were validated using the leave one out (LOO) method. Molecules were assigned in a random way to five groups of equal size. Reduced models were built keeping out one group at a time.

Acknowledgment. Many thanks are due to Prof. Gabriele Cruciani and Prof. Sergio Clementi (Molecular Discovery and MIA srl) for the use the GOLPE program in their chemometric laboratory (University of Perugia, Italy) and for having provided the GRID program. This work was partially supported by PRIN 2004 (A.M.), AIRC 2005 (A.M.), and Regione Campania 2003, LR 5/02 (G.S.) grants.

Supporting Information Available: Characterization data for compounds **6e**, **6f**, **9**, **10b**, **10c**, and **17**, sequence homology, and alignment and additional binding mode informations. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Strahl, B. D.; Allis, C. D. The language of covalent histone modifications. *Nature* **2000**, *403*, 41–45.
- (2) Jenuwein, T.; Allis, C. D. Translating the histone code. *Science* **2001**, *293*, 1074–1080.
- (3) Shi, Y.; Lan, F.; Matson, C.; Mulligan, P.; Whetstone, J. R.; Cole, P. A.; Casero, R. A.; Shi, Y. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **2004**, *119*, 941–953.
- (4) Forneris, F.; Binda, C.; Vanoni, M. A.; Battaglioli, E.; Mattevi, A. Human histone demethylase LSD1 reads the histone code. *J. Biol. Chem.* **2005**, *280*, 41360–41365.
- (5) Forneris, F.; Binda, C.; Vanoni, M. A.; Mattevi, A.; Battaglioli, E. Histone demethylation catalysed by LSD1 is a flavin-dependent oxidative process. *FEBS Lett.* **2005**, *579*, 2203–2207.
- (6) Chen, Y.; Yang, Y.; Wang, F.; Wan, K.; Yamane, K.; Zhang, Y.; Lei, M. Crystal structure of human histone lysine-specific demethylase 1 (LSD1). *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13956–13961.
- (7) Stavropoulos, P.; Blobel, G.; Hoelz, A. Crystal structure and mechanism of human lysine-specific demethylase-1. *Nat. Struct. Mol. Biol.* **2006**, *13*, 626–632.
- (8) Yamane, K.; Toumazou, C.; Tsukada, Y.; Erdjument-Bromage, H.; Tempst, P.; Wong, J.; Zhang, Y. JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. *Cell* **2006**, *125*, 483–495.
- (9) Whetstone, J. R.; Nottke, A.; Lan, F.; Huarte, M.; Smolnikov, S.; Chen, Z.; Spooner, E.; Li, E.; Zhang, G.; Colaiacovo, M.; Shi, Y. Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* **2006**, *125*, 467–481.
- (10) Tsukada, Y.; Fang, J.; Erdjument-Bromage, H.; Warren, M. E.; Borchers, C. H.; Tempst, P.; Zhang, Y. Histone demethylation by a family of JmjC domain-containing proteins. *Nature* **2006**, *439*, 811–816.
- (11) Klose, R. J.; Yamane, K.; Bae, Y.; Zhang, D.; Erdjument-Bromage, H.; Tempst, P.; Wong, J.; Zhang, Y. The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature* **2006**, *442*, 312–316.
- (12) Cloos, P. A.; Christensen, J.; Agger, K.; Maiolica, A.; Rappsilber, J.; Antal, T.; Hansen, K. H.; Helin, K. The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. *Nature* **2006**, *442*, 307–311.
- (13) Zhang, Y.; Reinberg, D. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev.* **2001**, *15*, 2343–2360.
- (14) Byvoet, P.; Shepherd, G. R.; Hardin, J. M.; Noland, B. J. The distribution and turnover of labeled methyl groups in histone fractions of cultured mammalian cells. *Arch. Biochem. Biophys.* **1972**, *148*, 558–567.
- (15) Baxter, C. S.; Byvoet, P. Intercalating agents as probes of the spatial relationship between chromatin components. *Biochem. Biophys. Res. Commun.* **1975**, *63*, 286–291.
- (16) Rice, J. C.; Allis, C. D. Histone methylation versus histone acetylation: new insights into epigenetic regulation. *Curr. Opin. Cell Biol.* **2001**, *13*, 263–273.
- (17) Turner, B. M. Memorable transcription. *Nat. Cell Biol.* **2003**, *5*, 390–393.
- (18) Lee, D. Y.; Teyssier, C.; Strahl, B. D.; Stallcup, M. R. Role of protein methylation in regulation of transcription. *Endocr. Rev.* **2005**, *26*, 147–170.

- (19) Bedford, M. T.; Richard, S. Arginine methylation an emerging regulator of protein function. *Mol. Cell* **2005**, *18*, 263–272.
- (20) Cook, J. R.; Lee, J. H.; Yang, Z. H.; Krause, C. D.; Herth, N.; Hoffmann, R.; Pestka, S. FBXO11/PRMT9, a new protein arginine methyltransferase, symmetrically dimethylates arginine residues. *Biochem. Biophys. Res. Commun.* **2006**, *342*, 472–481.
- (21) Gary, J. D.; Clarke, S. RNA and protein interactions modulated by protein arginine methylation. *Prog. Nucleic Acid Res. Mol. Biol.* **1998**, *61*, 65–131.
- (22) Cheng, D.; Yadav, N.; King, R. W.; Swanson, M. S.; Weinstein, E. J.; Bedford, M. T. Small molecule regulators of protein arginine methyltransferases. *J. Biol. Chem.* **2004**, *279*, 23892–23899.
- (23) a) Bartel, R. L.; Borchardt, R. T. Effects of adenosine dialdehyde on S-adenosylhomocysteine hydrolase and S-adenosylmethionine-dependent transmethylation in mouse L929 cells. *Mol. Pharmacol.* **1984**, *25*, 418–424.; b) Schwerk, C.; Schulze-Osthoff, K. Methyltransferase inhibition induces p53-dependent apoptosis and a novel form of cell death. *Oncogene* **2005**, *24*, 7002–7011.
- (24) Casellas, P.; Jeanteur, P. Protein methylation in animal cells. II. Inhibition of S-adenosyl-L-methionine:protein(arginine) N- methyltransferase by analogs of S-adenosyl-L-homocysteine. *Biochim. Biophys. Acta, Nucleic Acids Protein Synth.* **1978**, *519*, 255–268.
- (25) Barbes, C.; Sanchez, J.; Yebra, M. J.; Robert-Gero, M.; Hardisson, C. Effects of sinefungin and S-adenosylhomocysteine on DNA and protein methyltransferases from *Streptomyces* and other bacteria. *FEMS Microbiol. Lett.* **1990**, *69*, 239–243.
- (26) Massa, S.; Mai, A.; Sbardella, G.; Esposito, M.; Ragno, R.; Loidl, P.; Brosch, G. 3-(4-aryloyl-1H-pyrrol-2-yl)-N-hydroxy-2-propenamides, a new class of synthetic histone deacetylase inhibitors. *J. Med. Chem.* **2001**, *44*, 2069–2072.
- (27) Mai, A.; Massa, S.; Ragno, R.; Esposito, M.; Sbardella, G.; Nocca, G.; Scatena, R.; Jesacher, F.; Loidl, P.; Brosch, G. Binding mode analysis of 3-(4-benzoyl-1-methyl-1H-2-pyrrolyl)-N-hydroxy-2-propenamide: a new synthetic histone deacetylase inhibitor inducing histone hyperacetylation, growth inhibition, and terminal cell differentiation. *J. Med. Chem.* **2002**, *45*, 1778–1784.
- (28) Ornaghi, P.; Rotili, D.; Sbardella, G.; Mai, A.; Filetici, P. A novel Gcn5p inhibitor represses cell growth, gene transcription and histone acetylation in budding yeast. *Biochem. Pharmacol.* **2005**, *70*, 911–917.
- (29) Trojer, P.; Dangel, M.; Bauer, I.; Graessle, S.; Loidl, P.; Brosch, G. Histone methyltransferases in *Aspergillus nidulans*: evidence for a novel enzyme with a unique substrate specificity. *Biochemistry* **2004**, *43*, 10834–10843.
- (30) Barskii, V. E.; Ivanov, V. B.; Skliar Iu, E.; Mikhailov, G. I. [Dichlorotriazinylaminofluorescein—a new fluorochrome for cytochemical and histochemical detection of proteins]. *Izv. Akad. Nauk. SSSR Biol.* **1968**, *5*, 744–747.
- (31) Peng, X.; Yu, H.; Hang, Y.; Wang, J. N,N'-phosgenation with triphosgene in the synthesis of direct dyes containing the ureylene group. *Dyes Pigm.* **1996**, *32*, 193–198.
- (32) Adam, J. M.; Kaser, A. Anionic disazo dyes. EP0290384, 1988.
- (33) Birkett, H. E.; Cherryman, J. C.; Chippendale, A. M.; Evans, J. S. O.; Harris, R. K.; James, M.; King, I. J.; McPherson, G. J. Structural investigations of three triazines: Solution-state NMR studies of internal rotation and structural information from solid-state NMR, plus a full structure determination from powder X-ray diffraction in one case. *Magn. Reson. Chem.* **2003**, *41*, 324–336.
- (34) Kuntz, I. D. Structure-based strategies for drug design and discovery. *Science* **1992**, *257*, 1078–1082.
- (35) Hillisch, A.; Pineda, L. F.; Hilgenfeld, R. Utility of homology models in the drug discovery process. *Drug Discovery Today* **2004**, *9*, 659–669.
- (36) Tramontano, A.; Morea, V. Assessment of homology-based predictions in CASP5. *Proteins* **2003**, *53 Suppl. 6*, 352–368.
- (37) Mai, A.; Massa, S.; Ragno, R.; Cerbara, I.; Jesacher, F.; Loidl, P.; Brosch, G. 3-(4-Aroyl-1-methyl-1H-2-pyrrolyl)-N-hydroxy-2-alkylamides as a new class of synthetic histone deacetylase inhibitors. 1. Design, synthesis, biological evaluation, and binding mode studies performed through three different docking procedures. *J. Med. Chem.* **2003**, *46*, 512–524.
- (38) Rosato, R. R.; Grant, S. Histone deacetylase inhibitors in cancer therapy. *Cancer Biol. Ther.* **2003**, *2*, 30–37.
- (39) Mai, A.; Massa, S.; Cerbara, I.; Valente, S.; Ragno, R.; Bottoni, P.; Scatena, R.; Loidl, P.; Brosch, G. 3-(4-Aroyl-1-methyl-1H-2-pyrrolyl)-N-hydroxy-2-propenamides as a new class of synthetic histone deacetylase inhibitors. 2. Effect of pyrrole-C2 and/or -C4 substitutions on biological activity. *J. Med. Chem.* **2004**, *47*, 1098–1109.
- (40) Ragno, R.; Mai, A.; Massa, S.; Cerbara, I.; Valente, S.; Bottoni, P.; Scatena, R.; Jesacher, F.; Loidl, P.; Brosch, G. 3-(4-Aroyl-1-methyl-1H-pyrrol-2-yl)-N-hydroxy-2-propenamides as a new class of synthetic histone deacetylase inhibitors. 3. Discovery of novel lead compounds through structure-based drug design and docking studies. *J. Med. Chem.* **2004**, *47*, 1351–1359.
- (41) Mai, A.; Massa, S.; Rotili, D.; Cerbara, I.; Valente, S.; Pezzi, R.; Simeoni, S.; Ragno, R. Histone deacetylation in epigenetics: an attractive target for anticancer therapy. *Med. Res. Rev.* **2005**, *25*, 261–309.
- (42) Mai, A.; Massa, S.; Rotili, D.; Simeoni, S.; Ragno, R.; Botta, G.; Nebbioso, A.; Miceli, M.; Altucci, L.; Brosch, G. Synthesis and Biological Properties of Novel, Uracil-Containing Histone Deacetylase Inhibitors. *J. Med. Chem.* **2006**, *49*, 6046–6056.
- (43) Mai, A.; Massa, S.; Valente, S.; Simeoni, S.; Ragno, R.; Bottoni, P.; Scatena, R.; Brosch, G. Aroyl-Pyrrolyl Hydroxyamides: Influence of Pyrrole C4-Phenylacetyl Substitution on Histone Deacetylase Inhibition. *ChemMedChem* **2006**, *1*, 225–237.
- (44) Ragno, R.; Simeoni, S.; Valente, S.; Massa, S.; Mai, A. 3-D QSAR studies on histone deacetylase inhibitors. A GOLPE/GRID approach on different series of compounds. *J. Chem. Inf. Model.* **2006**, *46*, 1420–1430.
- (45) Kouranov, A.; Xie, L.; de la Cruz, J.; Chen, L.; Westbrook, J.; Bourne, P. E.; Berman, H. M. The RCSB PDB information portal for structural genomics. *Nucleic Acids Res.* **2006**, *34*, 302–305.
- (46) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (47) Zhang, X.; Cheng, X. Structure of the predominant protein arginine methyltransferase PRMT1 and analysis of its binding to substrate peptides. *Structure* **2003**, *11*, 509–520.
- (48) Weiss, V. H.; McBride, A. E.; Soriano, M. A.; Filman, D. J.; Silver, P. A.; Hogle, J. M. The structure and oligomerization of the yeast arginine methyltransferase, Hmt1. *Nat. Struct. Biol.* **2000**, *7*, 1165–1171.
- (49) Zhang, X.; Zhou, L.; Cheng, X. Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *EMBO J.* **2000**, *19*, 3509–3519.
- (50) Multiple sequence alignment was accomplished by the use of BioEdit software v. 7.0.1 and then manually adjusted. For reference see: Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.* **1999**, *41*, 95–98.
- (51) Venclovas, C. Comparative modeling in CASP5: progress is evident, but alignment errors remain a significant hindrance. *Proteins* **2003**, *53 Suppl. 6*, 380–388.
- (52) Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.* **2003**, *31*, 3381–3385.
- (53) Goodsell, D. S.; Morris, G. M.; Olson, A. J. Automated docking of flexible ligands: applications of AutoDock. *J. Mol. Recognit.* **1996**, *9*, 1–5.
- (54) Wang, R.; Lai, L.; Wang, S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 11–26.
- (55) Shoichet, B. K.; Stroud, R. M.; Santi, D. V.; Kuntz, I. D.; Perry, K. M. Structure-based discovery of inhibitors of thymidylate synthase. *Science* **1993**, *259*, 1445–1450.
- (56) Docking studies were not performed on non-active compounds (**7b**, **7c**, **7d**, and **20**) and on compound **15** due to its large dimensions.
- (57) The two tautomeric forms of compound **16** (keto and enol, indicated as **16a** and **16b**, respectively) docked in different sites.
- (58) Cruciani, G.; Watson, K. A. Comparative molecular field analysis using GRID force-field and GOLPE variable selection methods in a study of inhibitors of glycogen phosphorylase b. *J. Med. Chem.* **1994**, *37*, 2589–2601.
- (59) Frohman, M. A.; Dush, M. K.; Martin, G. R. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 8998–9002.
- (60) Bügl, H.; Fauman, E.; Staker, B.; Zheng, F.; Kushner, S.; Saper, M.; Bardwell, J.; Jakob, U. RNA Methylation under Heat Shock Control. *Mol. Cell* **2000**, *6*, 349–360.
- (61) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.
- (62) Schüttelkopf, A. W.; van Aalten, D. M. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 1355–1363.
- (63) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718.

- (64) Gillet, A.; Sanner, M.; Stoffler, D.; Olson, A. Tangible interfaces for structural molecular biology. *Structure* **2005**, *13*, 483–491.
- (65) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612.
- (66) Goodford, P. J., A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985**, *28*, 849–857.
- (67) *GOLPE*, Multivariate Infometric Analysis Srl., Viale dei Castagni 16, Perugia, Italy, 1999.

JM061213N