JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Selective Inhibitors of Histone Methyltransferase DOT1L: Design, Synthesis, and Crystallographic Studies

Yuan Yao,[†,§] Pinhong Chen,[†,§] Jiasheng Diao,[†] Gang Cheng,[†] Lisheng Deng,[†] Justin L. Anglin,[†] B. V. Venkataram Prasad,[‡] and Yongcheng Song*[,†]

[†]Department of Pharmacology and [‡]Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, 1 Baylor Plaza, Houston, Texas 77030, United States

Supporting Information

ABSTRACT: Histone H3-lysine79 (H3K79) methyltransferase DOT1L plays critical roles in normal cell differentiation as well as initiation of acute leukemia. We used structure- and mechanism-based design to discover several potent inhibitors of DOT1L with IC_{50} values as low as 38 nM. These inhibitors exhibit only weak or no activities against four other representative histone lysine and arginine methyltransferases, G9a, SUV39H1, PRMT1 and CARM1. The X-ray crystal structure of a DOT1L−inhibitor complex reveals that the N6-methyl group of the inhibitor, located favorably in a predominantly hydrophobic cavity of DOT1L, provides the observed high selectivity. Structural analysis shows that it will disrupt at least one H-bond and/or have steric repulsion for other histone methyltransferases. These compounds represent novel chemical probes for biological function studies of DOT1L in health and disease.

The human genome is packed into chromatins, which are composed of millions of repetitive units known as nucleosomes. A single nucleosome includes a fragment of DNA (∼147 base pairs) wound around a disklike histone octamer consisting of two histone H2A, H2B, H3, and H4 proteins. Post-translational epigenetic modifications on several lysine and arginine residues of histones, such as methylation and acetylation, control the accessibility of the DNA, thereby regulating the expression or silencing of a gene.[1] It has been widely recognized that in addition to gene mutations, aberrant epigenetic modifications play an important role in the initiation of many diseases, such as cancer.[2−4] Great interest has therefore been generated in studying histone-modifying enzymes (e.g., histone methyltransferases) as well as their functions in pathogenesis. Histone methyltransferases include a large family of dozens of histone lysine methyltransferases (HKMTs) and histone/protein arginine methyltransferases (PRMTs),[5,6] many of which have recently been found to play critical roles in cell differentiation, gene regulation, DNA recombination, and damage repair.[7] Therefore, small-molecule inhibitors of histone methyltransferases represent useful chemical probes for these biological studies as well as potential therapeutics.[8] However, very few inhibitors of HKMTs and PRMTs have been discovered and developed.[8,9]

We are particularly interested in human histone lysine methyltransferase DOT1L,[10,11] which is highly conserved from yeasts to mammals. DOT1L is a unique HKMT in that unlike all other

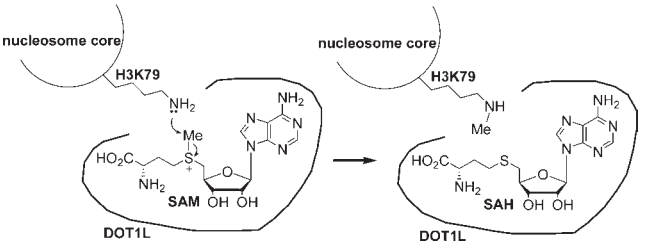


Figure 1. Mechanism of catalysis of DOT1L.

HKMTs containing a SET domain (which are class-V methyltransferases), it belongs to the class-I methyltransferase family. In addition, DOT1L is the only known enzyme that specifically catalyzes methylation of the histone H3-lysine79 (H3K79) residue located in the nucleosome core structure, while other methylation sites are in the unordered N-terminal tail of the histone. Moreover, the clinical importance of DOT1L as well as the H3K79 methylation is that DOT1L has been found to be necessary and sufficient for the initiation and maintenance of leukemia with mixed-lineage leukemia (MLL) gene translocations.[12−14] This type of leukemia accounts for ∼75% of infant and ∼10% of adult acute leukemia and has a particularly poor prognosis.[15] DOT1L therefore represents a novel target for intervention. It is of interest that while this manuscript was being revised for publication, a DOT1L inhibitor that possesses selective activity against MLL leukemia was disclosed.[16]

DOT1L catalyzes an S_N2 reaction of the H3K79 ε-NH_2 of the substrate nucleosome with the methyl group of *S*-(5′-adenosyl)-L-methionine (SAM), which is the cofactor of the enzyme, as schematically illustrated in Figure 1. One of the reaction products, *S*-(5′-adenosyl)-L-homocysteine (SAH), has been known to be a nonselective inhibitor of many methyltransferases, including DOT1L.[17] We also found that it inhibits recombinant human DOT1L (catalytic domain 1−472)[10] with a K_i value of 160 nM (Table 1). However, SAH cannot be used as a probe in cell biology or in vivo, since it is quickly degraded to become adenosine and homocysteine by SAH hydrolase,[18] keeping a cellular SAM/SAH molar ratio of ∼40:1.[19] In addition, selectivity is of importance for a DOT1L inhibitor to be a useful probe, since other histone lysine and arginine methyltransferases also use SAM and histone/nucleosome as their cofactor and substrate, respectively.[5,6]

Received: July 11, 2011
Published: September 21, 2011

dx.doi.org/10.1021/ja206312b | *J. Am. Chem. Soc.* 2011, 133, 16746–16749

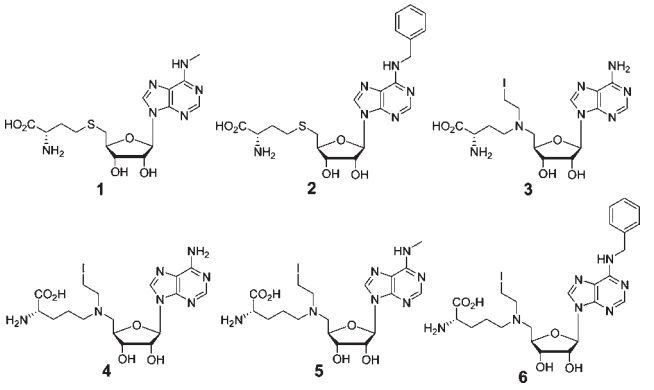
We analyzed the crystal structure of the DOT1L−SAM complex[11] as well as those of all other histone methyltransferases available in the Protein Data Bank and found one structural feature that is unique to the binding of SAM to DOT1L, which can be exploited to design selective DOT1L inhibitors. As shown in Figure S1A in the Supporting Information (SI), the 6-NH_2 group of SAM forms only one H-bond with DOT1L with a large hydrophobic cavity nearby. However, the 6-NH_2 group of bound SAM or SAH has two H-bonds with PRMTs (which also are class-I methyltransferases), such as CARM1 (also known as PRMT4), as shown in Figure S1B. All other HKMTs, such as G9a, are SET-domain methyltransferases having a completely different structure. The binding conformation of SAM/SAH to these enzymes is distinct from that of DOT1L, with the 6-NH_2 group facing toward the protein and forming two H-bonds (Figure S1C). We thus hypothesized that N6-substituted SAH analogues such as **1** and **2** (Chart 1)would be potent and selective DOT1L inhibitors. This turned out to be the case. Compounds **1** and **2** were synthesized from N6-substituted adenosine (see the Experimental Section in the SI). Compound **1**, having only one extra $-CH_3$ group in comparison with SAH, was still found to be a potent DOT1L inhibitor, with a K_i value of 290 nM (Table 1 and Figure S2), but it possesses only weak or no inhibitory activities against two PRMTs (CARM1 and PRMT1) and two HKMTs (G9a and SUV39H1), with K_i values of 22.7 to >100 μM (Table 1). In contrast, SAH remains an inhibitor of all these enzymes, with K_i values of 0.4−4.9 μM. Similarly, compound **2**, N6-benzyl-SAH, has good activity toward DOT1L (K_i = 1.1 μM) but is very weak against CARM1 and PRMT1 (K_i = 18 and 21.2 μM, respectively) and inactive toward G9a and SUV39H1 (Table 1).

Table 1. K_i or IC_{50} values (μM) of Methyltransferase Inhibitors

	DOT1L	CARM1	PRMT1	G9a	SUV39H1
SAH[a]	0.16	0.40	0.86	0.57	4.9
1[a]	0.29	>100	22.7	>100	>100
2[a]	1.1	18	21.2	>100	>100
3[b]	15.7	46.4	22.0	>100	>100
4[b]	0.038	1.1	2.7	1.8	>100
5[b]	0.12	>100	>100	>100	>100
6[b]	0.11	>100	>100	>100	>100

[a] K_i values. [b] IC_{50} values.

Chart 1. Structures of DOT1L Inhibitors



Next, X-ray crystallography was used to investigate how compound **1** binds to DOT1L, with a particular interest in the binding site of the N6-methyl group that provides excellent selectivity. We determined the crystal structure of the DOT1L−**1** complex at 2.5 Å. Details of the data processing and refinement are shown in Table S1 in the SI, and the overall structure and protein−ligand interactions in the DOT1L−**1** complex are illustrated in Figure S3. As shown in Figure 2A, the protein and the SAH moiety of the inhibitor are superimposed with those of the previously reported DOT1L−SAM structure,[11] with a root-mean-square deviation (rmsd) of 0.2 Å. As a result, all of the 10 H-bonds as well as the other interactions between the ligand and the protein remain essentially intact (Figure 2B), in agreement with the potent inhibitory activity of **1**. The N6-methyl group is nicely inserted into a hydrophobic cavity, surrounded by Phe223, Leu224, Val249, Lys187, and Pro133 (Figure 2B,C). In addition, its orientation allows the 6-NH group to form a H-bond with Asp222 that is important to the binding of the adenine ring.

It is therefore clear that introducing a substituent at N6 does not significantly affect the binding of SAH to DOT1L. However, our experiments show the N6-substituted SAH analogues **1** and **2** cannot bind to other HKMTs and PRMTs strongly (Table 1), suggesting that any substitution at this position disrupts at least

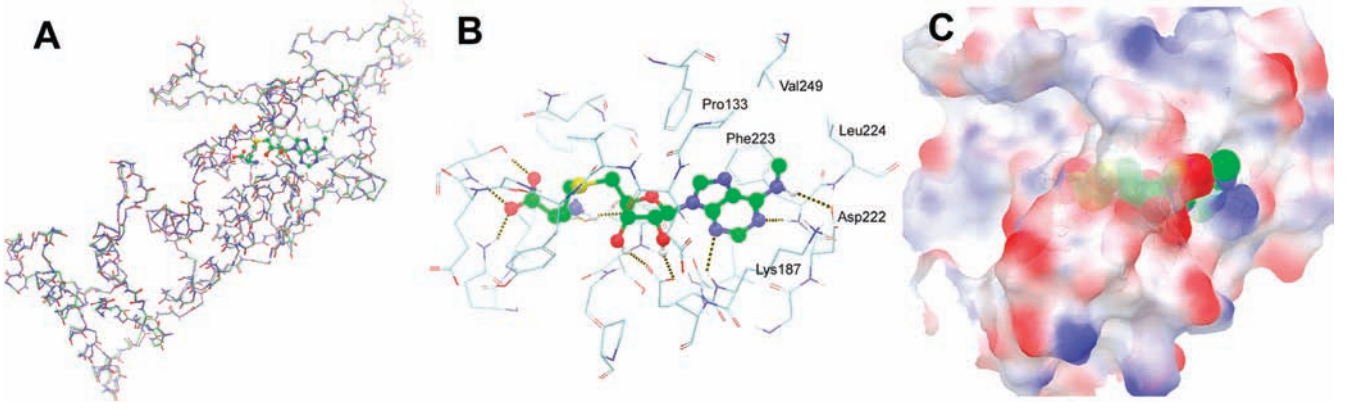
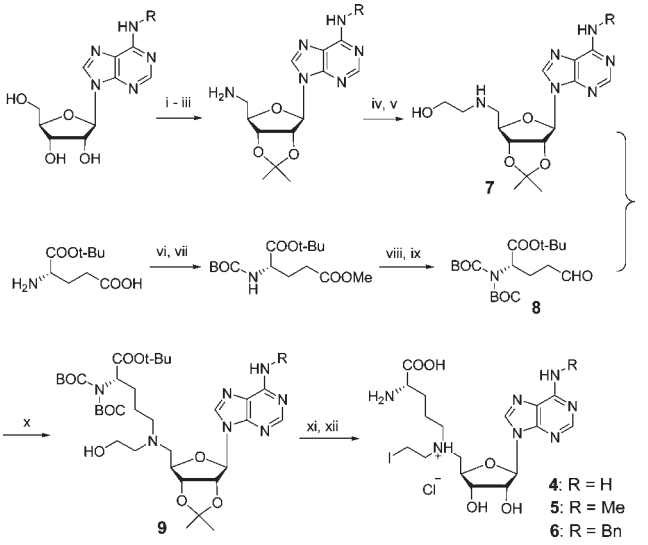


Figure 2. X-ray crystal structure of the human DOT1L−**1** complex. (A) Superposition of the structures of DOT1L−**1** (with C atoms in green) and DOT1L−SAM (in purple), with an rmsd of 0.2 Å. For clarity, only protein backbones are shown. (B) Close-up view of the active site of the DOT1L−**1** structure, with 10 H-bonds shown as dotted lines. (C) Electrostatic potential surface (with 25% transparency) of the DOT1L−**1** complex, showing that the N6-methyl group of **1** is located in a hydrophobic cavity. **1** is shown as a space-filling model.

dx.doi.org/10.1021/ja206312b |*J. Am. Chem. Soc.* 2011, 133, 16746–16749

Scheme 1. General Synthesis of Compounds 4−6[a]



[a] Reagents and conditions: (i) acetone, $SOCl_2$; (ii) phthalimide, PPh_3, diisopropyl azodicarboxylate; (iii) NH_2NH_2, 80 °C; (iv) ethyl bromoacetate, NEt_3; (v) $LiAlH_4$; (vi) BOC_2O; (vii) ClCOOMe, DMAP, NEt_3; (viii) BOC_2O, DMAP; (ix) DIBAL, −78 °C; (x) $NaCNBH_3$, HCl, MeOH; (xi) PPh_3, I_2, imidazole, 0 °C; (xii) HCl-dioxane.

one H-bond and/or changes the binding conformation of the adenine ring, thereby causing a considerable affinity loss. In addition, for SET-domain HKMTs, any N6-substituent leads to intolerable steric repulsion with the protein, preventing these compounds from binding. These results show that N6-substituted SAH analogues are selective inhibitors of DOT1L and provide a structural basis for further inhibitor design and development.

A mechanism-based inhibitor design was exploited to find selective DOT1L inhibitors with improved potency. Compound **3** (Chart 1) was initially synthesized. The rationale is that it can undergo intramolecular cyclization at neutral pH to form a reactive aziridinium intermediate[20,21] that may be covalently bound to the ε-NH_2 group of H3K79 (Figure S4). Compound **3** was found to exhibit only weak enzyme inhibition against DOT1L, with an IC_{50} value of 15.7 μM. We reasoned that compound **4** with one more $-CH_2-$ group could be a better inhibitor, since the two C−N bonds (∼1.47 Å each) in **3** are considerably shorter than the C−S bonds (∼1.82 Å) in SAM/SAH. The crystal structures of DOT1L show that SAM as well as **1** bind to the protein in a fully extended conformation, suggesting that the amino acid moiety of **3** might not be able to reach its optimal binding site in DOT1L. Compound **4** has not been made before, and our synthetic route is shown in Scheme 1. The 2′,3′-dihydroxyls of adenosine were selectively protected with an acetonide, after which the 5′-hydroxyl was converted to an $-NH_2$ group, using a Mitsunobu reaction followed by treatment with hydrazine. The product was alkylated with ethyl bromoacetate and reduced with $LiAlH_4$ to afford compound **7**. The *tert*-butyl ester of L-glutamic acid was first protected with one *tert*-butoxycarbonyl (BOC) group and its δ-carboxyl converted to a methyl ester. It was necessary to protect the amino group with a second BOC before reduction to give aldehyde **8**. Compounds **7** and **8** were subjected to a reductive amination to produce compound **9**, whose free hydroxyl group was converted to an iodide with PPh_3/I_2, affording compound **4** after acidic deprotection.

Compound **4** was found to be an extremely potent inhibitor of DOT1L (IC_{50} = 38 nM; Table 1), almost quantitatively inactivating DOT1L. Interestingly, it possesses relatively weak or no inhibitory activity toward other methyltransferases, with IC_{50} values of 1.1 to >100 μM, respectively, showing a high selectivity (>29-fold). It is remarkable that because of the complicated enzyme kinetics of histone methyltransferases involving covalent binding of inhibitor **4** (or **3**) to the substrate, we measured IC_{50} values for each enzyme using a minimal enzyme concentration (50−100 nM), K_m of SAM, and saturated concentration of the substrate. Under these assay conditions, the IC_{50} values may be used to compare the relative inhibitory abilities of each compound across these enzymes. Although **4** does not have an N6-substituent, the locally more hydrophobic environment at the binding site of the putative aziridinium intermediate of **4** in DOT1L might account for the selectivity, since it could protect the highly reactive aziridinium cation from nonspecific hydrolysis. The corresponding sites in other histone methyltransferases are either exposed to the solvent (for SET-domain HKMTs) or polar (for PRMTs). We synthesized compounds **5** and **6**, which are N6-substituted analogues of **4**, using the general approach in Scheme 1. These two compounds also exhibited potent activity against DOT1L, with IC_{50} values of 120 and 110 nM, respectively (Table 1). As expected, their N6-methyl and benzyl groups provide excellent selectivity: **5** and **6** are essentially inactive against other methyltransferases, showing that these compounds could have wide applications in probing the biological functions of DOT1L.

In summary, this work is of interest for a number of reasons. First, DOT1L, a specific histone H3K79 methyltransferase, plays a critical role in normal cell differentiation as well as the initiation and maintenance of acute leukemia with MLL gene translocations. DOT1L inhibitors therefore represent novel chemical probes for functional studies of DOT1L as well as potential therapeutics for leukemia. Second, we used structure- and mechanism-based design to synthesize and identify several potent DOT1L inhibitors with IC_{50} values as low as 38 nM. These compounds exhibit only weak or no inhibitory activities on four other representative histone lysine and arginine methyltransferases. Third, we determined the crystal structure of the DOT1L−**1** complex, which revealed the structural basis for the excellent selectivity. The methyl group of the inhibitor is located favorably in a hydrophobic cavity of DOT1L, while it disrupts at least one H-bond and/or has steric repulsions for all other histone methyltransferases. This finding should have implications for the future design and development of DOT1L inhibitors.

■ ASSOCIATED CONTENT

Supporting Information. Figures S1−S4, Table S1, experimental section, and complete ref 16. This material is available free of charge via the Internet at http://pubs.acs.org. Coordinates and structure factors of the DOT1L−**1** complex have been deposited in the Protein Data Bank as entry 3SR4.

■ AUTHOR INFORMATION

Corresponding Author
ysong@bcm.edu

Author Contributions
[§]These authors contributed equally.

■ ACKNOWLEDGMENT

This work was supported by a grant (RP110050) from the Cancer Prevention and Research Institute of Texas (CPRIT) to Y.S. and a grant (Q1279) from the Robert A. Welch Foundation to B.V.V.P. The recombinant human DOT1L (1-472) expression plasmid was kindly provided by Dr. Yi Zhang (University of North Carolina). We thank the staff of the X-ray Crystallography Facility at Baylor College of Medicine for assistance with data collection. We also thank Dr. Timothy Palzkill (Baylor College of Medicine) for useful discussion on enzyme kinetics.

■ REFERENCES

(1) Kouzarides, T. *Cell* **2007**, *128*, 693.
(2) Jones, P. A.; Baylin, S. B. *Cell* **2007**, *128*, 683.
(3) Wilson, C. B.; Rowell, E.; Sekimata, M. *Nat. Rev. Immunol.* **2009**, *9*, 91.
(4) Tsankova, N.; Renthal, W.; Kumar, A.; Nestler, E. J. *Nat. Rev. Neurosci.* **2007**, *8*, 355.
(5) Cheng, X.; Collins, R. E.; Zhang, X. *Biophys. Biomol. Struct.* **2005**, *34*, 267.
(6) Schubert, H. L.; Blumenthal, R. M.; Cheng, X. *Trends Biochem. Sci.* **2003**, *28*, 329.
(7) Bhaumik, S. R.; Smith, E.; Shilatifard, A. *Nat. Struct. Mol. Biol.* **2007**, *14*, 1008.
(8) Copeland, R. A.; Solomon, M. E.; Richon, V. M. *Nat. Rev. Drug Discovery* **2009**, *8*, 724.
(9) Cole, P. A. *Nat. Chem. Biol.* **2008**, *4*, 590.
(10) Feng, Q.; Wang, H.; Ng, H. H.; Erdjument-Bromage, H.; Tempst, P.; Struhl, K.; Zhang, Y. *Curr. Biol.* **2002**, *12*, 1052.
(11) Min, J.; Feng, Q.; Li, Z.; Zhang, Y.; Xu, R. M. *Cell* **2003**, *112*, 711.
(12) Okada, Y.; Feng, Q.; Lin, Y.; Jiang, Q.; Li, Y.; Coffield, V. M.; Su, L.; Xu, G.; Zhang, Y. *Cell* **2005**, *121*, 167.
(13) Krivtsov, A. V.; Armstrong, S. A. *Nat. Rev. Cancer* **2007**, *7*, 823.
(14) Krivtsov, A. V.; Feng, Z.; Lemieux, M. E.; Faber, J.; Vempati, S.; Sinha, A. U.; Xia, X.; Jesneck, J.; Bracken, A. P.; Silverman, L. B.; Kutok, J. L.; Kung, A. L.; Armstrong, S. A. *Cancer Cell* **2008**, *14*, 355.
(15) Hilden, J. M.; Dinndorf, P. A.; Meerbaum, S. O.; Sather, H.; Villaluna, D.; Heerema, N. A.; McGlennen, R.; Smith, F. O.; Woods, W. G.; Salzer, W. L.; Johnstone, H. S.; Dreyer, Z.; Reaman, G. H. *Blood* **2006**, *108*, 441.
(16) Daigle, S. R.; et al. *Cancer Cell* **2011**, *20*, 53.
(17) Richon, V. M.; Johnston, D.; Sneeringer, C. J.; Jin, L.; Majer, C. R.; Elliston, K.; Jerva, L. F.; Scott, M. P.; Copeland, R. A. *Chem. Biol. Drug Des.* **2011**, *78*, 199.
(18) Borchardt, R. T.; Keller, B. T.; Patel-Thombre, U. *J. Biol. Chem.* **1984**, *259*, 4353.
(19) Chiba, P.; Wallner, C.; Kaiser, E. *Biochem. Biophys. Acta* **1988**, *971*, 38.
(20) Osborne, T.; Weller Roska, R. L.; Rajski, S. R.; Thompson, P. R. *J. Am. Chem. Soc.* **2008**, *130*, 4574.
(21) Weller, R. L.; Rajski, S. R. *ChemBioChem* **2006**, *7*, 243.