Published in final edited form as:
J Am Chem Soc. 2012 October 31; 134(43): 18004–18014. doi:10.1021/ja307060p.

Sinefungin Derivatives as Inhibitors and Structure Probes of Protein Lysine Methyltransferase SETD2

Weihong Zheng[1,#], **Glorymar Ibáñez**[1,#], **Hong Wu**[2,#], **Gil Blum**[1,3], **Hong Zeng**[2], **Aiping Dong**[2], **Fengling Li**[2], **Taraneh Hajian**[2], **Abdellah Allali-Hassani**[2], **Maria F. Amaya**[2], **Alena Siarheyeva**[2], **Wenyu Yu**[2], **Peter J. Brown**[2], **Matthieu Schapira**[2], **Masoud Vedadi**[2], **Jinrong Min**[2,4,*], and **Minkui Luo**[1,*]

[1]Molecular Pharmacology and Chemistry Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA

[2]Structural Genomics Consortium, University of Toronto, 101 College St., Toronto, ON, M5G 1L7, Canada

[3]Tri-Institutional Training Program in Chemical Biology, Memorial Sloan-Kettering Cancer Center, New York 10065, USA

[4]Department of Physiology, University of Toronto, Toronto, Ontario, M5S 1A8, Canada

Abstract

Epigenetic regulations are involved in numerous physiological and pathogenic processes. Among the key regulators that orchestrate epigenetic signaling are over 50 human protein lysine methyltransferases (PKMTs). Interrogating the functions of individual PKMTs can be facilitated by target-specific PKMT inhibitors. Given the emerging need of such small molecules, we envision an approach to identify target-specific methyltransferase inhibitors by screening privileged small-molecule scaffolds against diverse methyltransferases. Here we demonstrate such feasibility by identifying the inhibitors of SETD2. N-propyl sinefungin (Pr-SNF) was shown to preferentially interact with SETD2 by matching the distinct transition-state features of SETD2's catalytically-active conformer. With Pr-SNF as a structure probe, we further revealed the dual roles of SETD2's post-SET loop on regulating substrate access through a distinct topological reconfiguration. Privileged sinefungin scaffolds are expected to have broad use as structure and chemical probes of methyltransferases.

INTRODUCTION

Epigenetics is defined as heritable phenotypic changes without altering the genotype.[1–2] Among the essential epigenetic regulators are diverse posttranslational modifiers such as protein lysine methyltransferases (PKMTs).[1–3] The human genome encodes more than 50 PKMTs, which use *S*-adenosylmethionine (SAM) as the cofactor and transfer its sulfonium methyl group to the ε-amino group of lysine side chains of specific protein substrates.[3] PKMT-mediated methylation regulates numerous biological functions such as signal transduction, gene transcription, and protein stabilization.[1–2,4] The dysregulation of these

*Corresponding Author: Minkui Luo: luom@mskcc.org, Jinrong Min: jr.min@utoronto.ca.
#These authors made equal contribution.

Supplementary Information. Reagents, assays, crystallography, molecular docking, synthesis, the full author list of Refs 20, 21, 26, Fig. S1–S3, and Table S1–S3). This material is available free of charge via the Internet at http://pubs.acs.org.

Data Deposition. The atomic coordinates and structure factors reported in this paper have been deposited in the Protein Data Bank, www.pdb.org (PDB codes 4H12,4FMU).

events has been linked to various diseases including cancer.[5–6] Given the physiological and pathological relevance of the emerging epigenetic targets, there is an urgent need of small-molecule probes to investigate biochemical properties of individual PKMTs as well as manipulate them pharmacologically.[3,5–6]

With the exception of DOT1L, whose catalytic domain is distinct from other protein methyltransferases,[7] PKMTs harbor a canonical 130-amino-acid SET domain for SAM binding and enzymatic catalysis.[8–10] By exploiting the distinct SAM-binding motif of DOT1L, Daigle *et. al.* developed the SAM analogue EPZ004777 as a potent inhibitor of DOT1L with *in vitro* IC_{50} of 0.3 nM.[11] The difference between DOT1L and other PKMTs also allowed Yao *et. al.* to design 5′-aziridine-based SAM analogues as DOT1L-selective inhibitors.[12] Apart from the rational design approach, medium- or high-throughput screening led to the identification of the inhibitors of SET-domain PKMTs such as chaetocin for *Drosophila melanogaster* SU(VAR)3–9 (likely human SUV39H1), BIX-01294 for G9a (likely its homologue GLP) and AZ505 for SMYD2.[3,13–15] However, the handful of inhibitors cannot satisfy the increased need of PKMT chemical probes toward understanding epigenetic roles of more than 50 human PKMTs.[3] Since most PKMTs rely on the highly-conserved SET domain for SAM binding and less-structured regions for substrate recognition, it seems challenging to develop PKMT inhibitors with both selectivity and potency in a rational manner.[3] Here we envision an approach to screen diverse methyltransferases against privileged small-molecule scaffolds to identify target-specific PKMT inhibitors.

Enzymatic transition state theory argues that even closely-related enzymes may adopt distinct transient structures in the reaction path of enzymatic catalysis and thus be selectively inhibited by the structurally-matched small molecules.[16–17] Molecular dynamic modeling and static structures of PKMTs suggest that transition state stabilization at substrate-cofactor interfaces of PKMTs involves both classic hydrophobic interaction/hydrogen bonds with the lysine side chain of substrates and nonclassic carbon-oxygen hydrogen bonds with SAM's sulfonium methyl moiety (Fig. 1).[18–19] Here we envision developing N-alkyl sinefungins as PKMT inhibitors by capturing certain transition-state characters (Fig. 1). To achieve high affinity to specific PKMTs, such sinefungin analogues are expected to position their secondary amines at the substrate-cofactor interface and the N-alkyl chains at the lysine-binding pocket for optimal interactions (Fig. 1).

Human SETD2 is a tumor-suppressing PKMT implicated in p53-dependent gene regulation, transcription elongation, and intron-exon splicing.[20–26] Aberrant activities of SETD2 and its homologues NSD1/2/3 are also implicated in various developmental syndromes and cancers.[20,27–32] Here we report N-alkyl sinefungin analogues as SETD2-specific inhibitors by matching SETD2's distinct transition-state characters. With the aid of N-propyl sinefungin (Pr-SNF, Fig. 1), we further revealed that SETD2's post-SET loop goes through a remarkable reconfiguration for inhibitor binding, substrate recognition and enzymatic catalysis. Since structurally-diverse sinefungin variants can be examined in a similar manner, the sinefungin-based scaffolds are expected to display broad utility as structure or pharmacological probes of protein methyltransferases.

RESULTS

Synthesis of sinefungin analogues 3a–d (Scheme 1)

As a privileged small-molecule scaffold to screen PKMT inhibitors, sinefungin analogues **3a–d** were prepared from the reported D-ribose derivative **4**.[33] To access N–alkyl (methyl, ethyl or benzyl) **3a**, **3b** and **3d**, the common precursor **4** was alkylated with the corresponding haloalkanes (methyl/ethyl iodide or benzyl bromide) to afford **5a**, **5b** and **5d**,

respectively. In contrast to the ready alkylation of **4** to **5a**, **5b** and **5d**, the reaction of **4** with iodopropane is sluggish (< 15% yield), likely because of the haloalkane's low reactivity. To circumvent this toward the preparation of N-propyl sinefungin (Pr-SNF, **3c**), we adopted an alternative strategy through ozonolysis of the terminal alkene of **4** and then Me_2S workup in the presence of methanol to yield **6**,[34] followed by alkylation with more reactive allyl iodide to furnish **7** (this allyl group serves as the precursor of Pr-SNF's N-propyl moiety). The primary alcohols **8a**–**d** were obtained via either ozonolysis of terminal-alkene-containing **5a**, **5b** and **5d** or I_2-facilated acetal deprotection of **7**,[35] followed by hydride reduction. The mesylation of **8a**–**d** and subsequent iodination afforded **9a**–**d**. The chiral amino acid moiety of **3a**–**d** was introduced by Schöllkopf bislactim ether chiral auxiliary upon the synthesis of **10a**–**d**.[36–37] These intermediates were processed into **11a–d** and then **12a–d** after hydrolytic removal of the pyrazine and isopropylidene, protection of the amino moiety with benzyloxycarbonyl (Cbz) group, acetylation of the ribosyl hydroxyl moieties, and then incorporation of N^6-benzoyladenine under Vorbrüggen conditions.[33] Further conversion of **12a**–**d** by sequential treatments of K_2CO_3 (remove acetates), hydrazine (remove the methyl ester and N^6-benzoyl group) and Pd-catalyzed hydrogenolysis (selectively remove Cbz and reduce allyl group) yielded the desired final products **3a**–**d** (overall yields of 7~11% from **4** to **3a–d**, Scheme 1 and Supplementary Information).

Pr-SNF 3c and N-benzyl sinefungin 3d as potent, selective inhibitors of SETD2

With sinefungin and its derivatives **3a**–**d**, we examined their inhibition profile against a panel of human methyltransferases, including 10 PKMTs (SET7/9, SET8, EZH2, MLL, GLP, G9a, SUV39H2, SETD2, SUV420H1 and SUV420H2) as well as 5 non-PKMTs (PRMT1, PRMT3, CARM1, DOT1L and DNMT1) (Fig. 2). Among the 5×15 panel of small molecules and enzymes, **3c** (Pr-SNF) and **3d** displayed 2~200-fold and 10~100-fold preferences, respectively, to SETD2 over other examined methyltransferases (Fig. 2a and Table S1). The potency of **3c** and 3**d** against SETD2 (apparent IC_{50} of 0.80 ± 0.2μM and 0.48 ± 0.06μM, respectively) is more than 10-fold higher than that of sinefungin **2**, **3a** and **3b** (Fig. 2b and Table S1). This structure-activity-relationship (SAR) indicated that the N-propyl/benzyl moieties of **3c** and **3d**, in contrast to the free amine of sinefungin **2** and N-methyl/ethyl groups of **3a** and **3b**, contribute to the tight interaction with SETD2. For other examined methyltransferases in the 5×15 assay panel against **3a–d**, only SET7/9 and CARM1 show modest μM-range IC_{50} of 1.4~3.0 μM, which are not significantly different from IC_{50} of the parent compound sinefungin **2**, a pan-methyltransferase inhibitor (Fig. 2c and Table S1).[3] With the exception of SETD2 among the examined 15 methyltransferases, appending the small alkyl moieties (*e.g.* methyl/ethyl/propyl groups) to sinefungin deteriorates or has no effect on its IC_{50} (Table S1). In contrast, the strong preference of **3c** and **3d** against SETD2, together with their well-correlated SAR, presents the two sinefungin derivatives as SETD2-specific inhibitors with decent potency and selectivity.

Pr-SNF as a structure probe specific for SETD2's active conformer

To elucidate the molecular mechanism of the potency and selectivity of the sinefungin derivatives as SETD2 inhibitors, we solved the structures of the binary complexes of SETD2's catalytic domain (AWS, SET and post-SET motifs) with *S*-adenosyl-*L*-homocysteine (SAH, which was degraded from SAM under the crystallization condition, Experimental Method) and Pr-SNF **3c** (Fig. 3a). In the presence of SAH, SETD2 adopts a compact configuration with its AWS motif and post-SET motif interwoven with the central SET domain. The SAH molecule is located within a deep pocket formed between SETD2's SET domain and post-SET domain, as conserved among all known SET-domain PKMTs (*e.g.* SET7/9, SET8, G9a, GLP, MLL, SUV39H2, SUV420H1 and SUV420H2).[9,38–39] However, the SETD2-SAH binary complex is distinct by its auto-inhibitory post-SET loop,

which is positioned to prevent substrate binding, and its characteristic Arg1670 residue located in the pocket that is otherwise occupied by substrate lysine (Fig. 3b). Such auto-inhibitory topology has also been reported for NSD1 and ASH1L, two closely-related homologues of SETD2, and proposed to regulate the access of substrates to the PKMTs (Fig. 3c, S2).[10,40–42]

Although the overall structure of the binary complex of SETD2 with Pr-SNF **3c** is similar to that with SAH, a remarkable difference was revealed at the post-SET motif of SETD2 in complex with Pr-SNF (Fig. 3b,d). The overlaid structures of SETD2 with Pr-SNF and SAH showed that the Pr-SNF's propyl group partially extends into the lysine-binding pocket (Fig. 3b,d). To accommodate this N-propyl moiety, which would otherwise clash with the Arg1670 residue in the SETD2-SAH complex, SETD2 orients this arginine 15 Å away from the lysine-binding pocket by flipping the otherwise auto-inhibitory post-SET loop (Fig. 3d). This reconfiguration also vacates SETD2's catalytic site for the entry of substrates (this site is occupied by the post-SET loop in the SETD2-SAH complex). To stabilize this substrate-accessible configuration in the SETD2-Pr-SNF binary complex, SETD2's post-SET loop is glued at a newly-formed hydrophobic core through Tyr1604, Tyr1666, Phe1668, Leu1689 and the hydrocarbon side chain of Arg1670 (Fig. 3d). The structure overlay with the substrate-bound GLP further revealed that Gln1669/Tyr1671 in the SETD2's post-SET loop and the helix between SETD2's Glu1588–Asn1599 region are topologically comparable to GLP's Arg1214/Ile1218 and GLP's Ser1132–Glu1138 helix region (Fig. 3b).[43–44] Given that these residues of GLP play the key role on substrate recognition,[43–44] the comparable regions of SETD2 (Gln1669/Tyr1671 in the post-SET loop; the helix of aa 1588–1599) are expected to participate in substrate binding in a similar manner. Comparing the overlaid structures of Pr-SNF-SETD2 and SAH-SETD2 also revealed a 45-degree rotation of Tyr1666 residue, although its biological relevance remains unknown. Collectively, these results showed that SETD2's methyltransferase domain can adopt at least two alternative conformations through flipping its post-SET loop: an auto-inhibitory close conformation and a substrate-accessible open conformation. Here Pr-SNF **3c** plays a key role as a structure probe through its preferential interaction with the latter.

Effects of mutations of SETD2's post-SET loop on enzyme catalysis

Structural analysis of SETD2 underscored the importance of SETD2's post-SET loop to receive substrate. To confirm the role of SETD2's post-SET loop, we mutated SETD2's F1668, Q1669, R1670 and Y1671 residues to alanine (Fig. 4). The barely-detectable methylation activities of SETD2's F1668A and Q1669A/Y1671A mutants are consistent with their proposed roles in stabilizing SETD2's active open conformer and interacting with substrate, respectively. In contrast, SETD2's R1670A mutant partially retains the methylation activity. To further examine the role of this residue in enzyme catalysis, Arg1670 was systematically replaced with nonpolar/hydrophobic amino acids G/V/I/L/P/F/W as well as polar amino acids K/Q. The overall activity profile indicated that SETD2's Arg1670 site can tolerate modest (*e.g.* A/V/I/L/F) but not extreme hydrophobic substitutions (*e.g.* G/P/W) (Fig. 4). This observation is consistent with the role of the hydrocarbon portion of Arg1670's side chain to stabilize SETD2's active open conformer by interacting with the nearby hydrophobic core. Such interaction is expected to be partially maintained by replacing the arginine with certain hydrophobic residues (Fig. 3d). The complete loss of activity of SETD2's R1670Q mutant versus partial loss of the activity of the R1670K/A/V/I/L/F mutant, also suggests that both the charge and hydrophobic side chain of SETD2's Arg1670 play certain roles on the enzyme catalysis (Fig. 4). These mutagenesis results therefore confirm the roles of SETD2's post-SET loop and its Arg1670 on substrate interaction, as revealed above by SETD2's structures.

Characterization of Pr-SNF as SETD2 inhibitor via enzyme kinetics

The structure of the Pr-SNF-SETD2 binary complex suggests that, to accommodate Pr-SNF, SETD2 needs to adopt the catalytically-active open configuration with its post-SET loop aligned for substrate entry. Given the present difficulty to obtain the structure of the SETD2-substrate complex, such topological switch was validated through analyzing the kinetics of SETD2-catalyzed methylation in the absence or presence of Pr-SNF (Fig. 5). The initial velocities of SETD2-catalyzed methylation were monitored with the H3 peptide (20–50 amino acid) as H3K36 substrate and [^{3}H-Me]-SAM as cofactor. The plots of the initial velocities versus concentrations of peptide substrate or SAM cofactor were first generated in the absence of Pr-SNF. The linear regression did not converge on the Y axis (Fig. 5a) and thus ruled out an ordered sequential mechanism: the ordered SAM/substrate or substrate/SAM binding along the reaction path of SETD2-catalyzed methylation.[45] The converged initial velocities on the X axis versus the concentrations of peptide substrate or SAM further argue that SETD2-catalyzed methylation goes through a random sequential mechanism and the binding of either SAM or substrate does not affect the sequential binding of the other ($\alpha = 1$, Fig. 5c).[45] Plotting the slopes of the double reciprocal kinetics against the concentrations of SAM or substrate gives SETD2's $K_{m,SAM}$ of 1.21 ± 0.05 μM, $K_{m,Substrate}$ of 0.42 ± 0.02 μM and k_{cat} = 0.14 ± 0.01min^{-1} (Fig. 5c, S1a). The similar results were obtained via non-linear regression of the same set of kinetic data (Fig. S1b).

To accurately determine the dissociation constant K_d of **3c** on SETD2, we measured initial velocities of SETD2-catalyzed methylation versus concentrations of peptide substrate and SAM cofactor as the function of the concentration of Pr-SNF **3c**. In the presence of the varied amount of **3c** and the fixed amount of SAM, the double-reciprocal kinetics of SETD2-catalyzed methylation versus peptide substrate converged in the negative region of X axis (Fig. 5b left).[45] This result implicated a non-competitive character between Pr-SNF and peptide substrate, featured as co-existence of the inhibitor-enzyme binary complex and inhibitor-enzyme-substrate ternary complex (Fig. 5c).[45] Given that Pr-SNF's N-propyl moiety and the H3K36 lysine chain could clash if simultaneously occupying SETD2's lysine-binding pocket, the formation of the inhibitor-enzymesubstrate complex is intriguing and can be better elucidated upon solving the structure of the ternary complex.

The converged linear regression in the negative region of Y axis further revealed that the initial binding of either Pr-SNF **3c** or substrate facilitates the following binding of the other ($\beta < 1$) (Fig. 5b left, 5c). In contrast, in the presence of the varied amount of Pr-SNF and the fixed amount of peptide substrate, the double-reciprocal kinetics of SETD2-catalyzed methylation versus SAM cofactor converged on Y axis, consistent with the competitive character between Pr-SNF **3c** and SAM cofactor (Fig. 5b right, 5c). Upon fitting these kinetics with $K_{m,SAM}$ =1.21 ± 0.05 μM, $K_{m,Substrate}$ = 0.42 ± 0.02 μM and $\alpha = 1$ (results above), K_d = 360 ± 15 nM and βK_d = 43 ± 4 nM (β = 0.12 ± 0.01) were obtained for Pr-SNF to bind SETD2 in the absence and presence of substrate, respectively (Fig. 5c, S1b). These values are well consistent with the IC_{50} of **3c** against SETD2 as measured with sub-K_m concentrations of SAM cofactor and peptide substrate (calculated IC_{50} of 0.60 μM versus experimental IC_{50} of 0.80 μM shown in Fig. 2). Our kinetic data together with SETD2's structures thus concluded that SETD2-catalyzed methylation goes through a random sequential mechanism and Pr-SNF **3c** inhibits this process through the formation of either Pr-SNF-SETD2 binary or Pr-SNF-SETD2-substrate ternary complex (Fig. 5c). The small β value of 0.12 further argues that the binding of Pr-SNF to SETD2 significantly facilitates the subsequent recruitment of substrate to form the inhibitor-SETD2-substrate ternary dead complex ($K_{m,Substrate}$ = 0.42 ± 0.02 μM versus $\beta K_{m,Substrate}$ = 0.050 ± 0.005 μM, Fig. 5c). The high affinity of the Pr-SNF-SETD2 binary complex to substrate peptide, as reflected by small $\beta K_{m,Substrate}$ of 0.05 ± 0.005μM, also recapitulates the prior finding

that **3c** preferentially interacts with SETD2's active open conformer and thus position SETD2's post-SET loop in a ready configuration to bind substrate.

Selective inhibition of Pr-SNF 3c on SETD2 versus other SET-domain PKMTs

To further explore the origin of the selectivity of Pr-SNF **3c** on SETD2 versus other SET-domain PKMTs, we superimposed SETD2-bound Pr-SNF to SAM or SAH in complex with SETD2's auto-inhibitory close conformer and other PKMTs. The secondary amine of Pr-SNF **3c** is expected to be protonated at physiological pH and thus carries two hydrogen atoms. In SETD2's active open conformation, Pr-SNF's secondary amine moiety is ideally positioned to form the two hydrogen bonds with the backbone carbonyl of SETD2's R1625 and Y1604 (N-O distances of 2.8 Å and 2.9 Å, respectively) (Fig. 6). The comparable carbonyl groups of SET7/9 have also been proposed to participate in transition-state stabilization and enzymatic catalysis at the substrate-cofactor interface by forming nonclassic oxygen-carbon hydrogen bonds with SAM's sulfonium methyl moiety.[18] In contrast, such carbonyl residues in SETD2's inactive conformer (SETD2 in complex with SAH) and the prior SET domain structures (*e.g.* SET8, MLL, GLP, G9a, SUV39H2, SUV420H1 and SUV420H2) are not well positioned with the less optimally docked N-O distances of > 3.1 Å for the carbonyl groups (Fig. 3,6). In these cases, the formation of the optimal hydrogen bonds would require the PMTs to alter ground-state conformations and thus pay additional energy penalty. Despite the near-optimal N-O distances of 3.1 Å in the structures of SETD8, SUV420H1 and SUV420H2 (Fig. 6), Pr-SNF shows even lower affinity to MLL and SUV39H2 (Table S1). This observation indicates that additional factors may further contribute to the altered affinity of Pr-SNF on PKMTs. In the case of SET7/9, the amide oxygen of the side-chain of N265 can form an alternative hydrogen bond, which may compensate the less optimal carbonyl hydrogen bond and thus cause the modest cross-inhibition observed (Fig. 2,6). This analysis therefore suggests that SETD2's active open conformer contains the two well-positioned backbone carbonyls to form the hydrogen bonds with Pr-SNF and these interactions in part account for its selective inhibition on SETD2 over other PKMTs.

Here we further argue that the N-propyl moiety of Pr-SNF also plays a positive role on its preferential interaction with SETD2's active open conformer, because sinefungin and its N-methyl derivative **3a** do not show such selectivity (Table S1). In the structure of the SETD2-Pr-SNF binary complex, Pr-SNF's N-propyl moiety was located in the hydrophobic binding pocket of substrate lysine formed by SETD2's Tyr1579, Tyr1605, Met1627, Phe1650, Phe1664 and Tyr1666 (Fig. 3d). The readiness of the lysine-binding pocket to accommodate N-alkyl groups is thus essential for Pr-SNF's selectivity on SETD2. Here SETD2's lysine-binding pocket can be explored with the N-alkyl sinefungin analogues **3a**–**d** as structure probes (Fig. 2 and Table S1). The gradually reduced IC_{50} of 100 μM, 8.2 ± 1.2 μM, 0.80 ± 0.20μM and 0.48 ± 0.06μM from **3a**–**d** on SETD2 indicate that the lysine-binding pocket of SETD2 is flexible enough to accommodate bulky propyl or benzyl moiety. In contrast, 10~50-fold drop in potency from **3c** to **3d** suggest that the lysine-binding pockets of SET7/9 and SUV39H2 are not sufficient for **3d**'s benzyl group. These findings therefore substantiate the importance of the matched lysine-binding pocket, besides the two well-positioned backbone carbonyls (Fig. 6), for Pr-SNF's selectivity on SETD2 over other PKMTs.

DISCUSSION

Synthesizing the privileged N-alkyl sinefungin derivatives and screening them against a panel of methyltransferases allowed us to identify Pr-SNF **3c** and N-benzyl sinefungin **3d** as inhibitors of SETD2. Among the examined 10 SET domain PKMTs (SET7/9, SET8, EZH2, MLL, GLP, G9a, SUV39H2, SETD2, SUV420H1 and SUV420H2) and 5 non-SET-domain

methyltransferases (PRMT1, PRMT3, CARM1, DOT1L and DNMT1), Pr-SNF **3c** and N-benzyl sinefungin **3d** showed the best inhibition on SETD2 with apparent IC_{50} of sub-μM (K_d = 0.36 ± 0.02μM for Pr-SNF). Further characterization of Pr-SNF revealed that the presence of SETD2's substrate enhances its K_d by additional 8-fold (βK_d = 0.043 ± 0.004 μM). SETD2 is solely responsible for H3K36 trimethylation and has been characterized as a tumor suppressor of multiple cancers.[22,24,46] Here we present Pr-SNF **3c** (likely N-benzyl sinefungin **3d** as well) as a SETD2-specific inhibitor.

Pr-SNF was further demonstrated to be a valuable structure probe of human SETD2. Pr-SNF **3c** preferentially interacts with SETD2's catalytically-active open conformer. Compared with SETD2's auto-inhibitory close conformer, the active open conformer is featured by a distinct reconfiguration of a post-SET substrate-recognizing loop. In particular, SETD2's Arg1670 residue plays dual roles by orienting the loop to block substrate entry in the inactive close conformer but readily receive the substrate in the active open conformer. The configuration switch also positions SETD2's backbone carbonyl groups to form two optimal hydrogen bonds with Pr-SNF's secondary amine. The presence of Pr-SNF **3c** therefore favors the equilibrium to SETD2's catalytically-active open conformer. Remarkably, NSD1 and ASH1L were also reported to contain similar auto-inhibitory loops to prevent free access of substrates.[40–42,47] SETD2, NSD1 and ASH1L are more closely-related among the SET-domain-containing PKMTs on the basis of their primary sequence (Fig. S2) and their ability to recognize H3K36 as a substrate.[40–42,47] The auto-inhibitory topology may be adapted by the subfamily of PKMTs as a general mechanism to regulate substrate recognition. The switch from the auto-inhibitory to catalytically-active configuration of the subfamily of PKMTs could be modulated by their binding partners in cellular contexts and thus account for their context-specific substrate recognition.[48]

Pr-SNF's preference for SETD2 over 14 other methyltransferases is striking given the overall structural similarity between the 10 examined SET-domain-containing PKMTs.[38–39] Structural analysis and enzyme kinetics of SETD2, together with the results of mutagenesis and molecular docking, indicated that such selectivity relies on the existence of the matched lysine-binding pocket and unique catalytically-active conformer of SETD2. To rationally design inhibitors of protein methyltransferases, several prior efforts focused on conjugating a portion of substrates to a 5′-azo-SAM analogue (bisubstrate-type inhibitors).[3,12,49–50] With exception of the 5′-aziridine-based inhibitor of DOT1L whose structural topology is different from other protein methyltransferases,[12] most bisubstrate-type inhibitors only showed modest IC_{50}.[49–50] More mechanistic studies appear to be necessary to understand their low potency. It has been noticed that protein methyltransferases, though structurally similar in terms of conversed SAM-binding motifs, can display a broad range of affinity to the same ligand (*e.g.* SAM and sinefungin) and this variation, even between closely-related methyltransferases, could not be readily justified according to their static structures.[3,8–10] This observation therefore argues that individual protein methyltransferases may achieve tight interaction with specific ligands by adopting alternative but better matched conformations. Our current success in identifying the N-alkyl sinefungin analogues as SETD2 inhibitors presents the utility and power of using privileged scaffolds to probe these distinct conformations in the course of developing PKMT inhibitors. Given that only a limited number of sinefungin analogues and PKMTs are examined here, we envision a promising use of structural variants of sinefungin as structure and chemical probes to elucidate functions of protein methyltransferases.

CONCLUSION

In this work, we outlined an approach, apart from the conventional high throughput screening, to identify target-specific methyltransferase inhibitors by screening privileged

small-molecule scaffolds against diverse methyltransferases. Among the small set of sinefungin derivatives synthesized here, Pr-SNF and N-benzyl sinefungin were identified as SETD2-specific inhibitors with decent potency and selectivity. The preferential interaction between the N-alkyl sinefungin analogues and SETD2 attributes to the distinct transition-state features of SETD2's catalytically-active conformer. With Pr-SNF as a structure probe, we further revealed a dual role of SETD2's post-SET loop on regulating substrate access through a distinct topological reconfiguration. The current work further argues that even closely-related SET-domain-containing PKMTs, which contain almost identical SAM-binding motifs, can adopt distinct configurations and thus be selectively inhibited by well-designed small molecules. Although sinefungin was regarded as a pan-inhibitor of methyltransferases, we demonstrated that well-designed sinefungin variants can go beyond the pan-inhibitor category and thus stand as lead compounds for further optimization. Given sinefungin contains rich structural motifs including primary amine, carboxylic acid, adenine and ribosyl moieties and thus can be subject to further derivatization, privileged sinefungin scaffolds are expected to show broad use in the course of developing inhibitors and interrogating functions of methyltransferases.

EXPERIMENTAL SECTION

Synthesis and characterization of Pr-SNF 3c

To a stirred solution of **12c** (0.02 mmol) in methanol (10 mL) was added potassium carbonate (14 mg, 0.1 mmol). The resultant mixture was stirred at ambient temperature for 8h, concentrated to dryness and then redissolved in 10 mL water. To the mixture was added hydrazine monohydrate (5 μL, 0.1 mmol). The reaction was stirred for 8 h at ambient temperature, neutralized with 1M aqueous HCl and then concentrated under reduced pressure. This mixture was then dissolved in 6 mL ethanol: water (5:1). To this solution was added 20 μL acetic acid and palladium on activated carbon (15 mg, 10 wt%, wet Degussa type). The subsequent hydrogenation reaction was carried out with hydrogen balloon for 12 h. The reaction mixture was filtered through a short pad of Celite and washed out with 20 mL MeOH and then 20 mL 0.1% TFA/water. The combined filtrates were concentrated under reduced pressure. The resultant crude products were purified by preparative reversed-phase HPLC (XBridge™ Prep C 18 5μm OBD™ 19×150mm) with 0~15 % gradient of acetonitrile in aqueous trifluoroacetic acid (0.1%) in 10 min and a flow rate of 10 mL/min. The fractions containing Pr-SNF were collected. The volatile solvents were removed by SpeedVac. The resultant solution was lyophilized to give the desirable product **3c** with overall 7% from Compound **4**. Pr-SNF was dissolved in water and stored at −20 °C before use.

3a R = Me, 52% yield. ^{1}H-NMR (600 MHz, MeOD): δ 1.96–2.03(m, 2H), 2.05–2.08(m, 2H), 2.25–2.29(m, 2H), 2.64(s, 3H), 3.43–3.45(m, 1H), 3.99–4.03(m, 1H), 4.19–4.22(m, 1H), 4.36(t, 1H, *J* = 5.9 Hz), 4.65(dd, 1H, *J* = 5.4Hz, 3.7 Hz), 6.01(d, 1H, *J* = 3.7 Hz), 8.35(s, 1H), 8.36(s, 1H); ^{13}C-NMR (150 MHz, MeOD): δ 26.54, 27.60, 31.43, 33.55, 53.59, 58.23, 74.86, 75.01, 80.92, 91.85, 117.99(q, *J* = 289.7Hz), 121.15, 143.44, 149.45, 150.10, 154.64, 162.57(q, *J* = 35.5 Hz),171.52; MS(ESI) m/z: 396 $[M+H]^+$; HRMS: calculated for $C_{17}H_{28}N_7O_5$ ($[M+H]^+$) 396.1995, found: 396.1982.

3b R = Et, 56% yield. ^{1}H-NMR (600 MHz, MeOD): δ 1.11(t, 3H, *J* = 7.2 Hz), 1.93–1.97(m, 2H), 1.99–2.07(m, 2H), 2.23–2.27(m, 1H), 2.28–2.32(m, 1H), 3.05(q, 2H, *J* = 7.2 Hz), 3.46–3.48(m, 1H), 3.97(t, 1H, *J* = 6.0 Hz), 4.19–4.22(m, 1H), 4.37(t, 1H, *J* = 6.0 Hz), 4.70(dd, 1H, *J* = 5.4Hz, 3.8 Hz), 5.99(d, 1H, *J* = 3.8 Hz), 8.30(s, 2H); ^{13}C-NMR (150 MHz, MeOD): δ 11.53, 26.99, 27.71, 33.51, 41.92, 53.75, 56.89, 74.57, 75.17, 80.91, 91.78, 118.09(q, *J* = 289.2 Hz), 121.12, 142.79, 150.26, 151.64, 156.02, 162.70(q, *J* = 35.4 Hz),171.77; MS(ESI)

m/z: 410 $[M+H]^+$; HRMS: calculated for $C_{17}H_{28}N_7O_5$ ($[M+H]^+$) 410.2152, found 410.2142.

3c R = Pr, 57% yield. ^{1}H-NMR (500 MHz, MeOD): δ 0.83(t, 3H, *J* = 7.4 Hz), 1.42–1.49(m, 1H), 1.52–1.59(m, 1H), 1.94–2.09(m, 4H), 2.21–2.26(m, 1H), 2.29–2.35(m, 1H), 2.92(t, 2H, *J* = 8.0 Hz) 3.44–3.48(m, 1H), 4.01(t, 1H, *J* =6.0 Hz), 4.19–4.22(m, 1H), 4.40(t, 1H, *J* =6.0 Hz), 4.67(dd, 1H, *J* = 5.4Hz, 3.4 Hz), 6.02(d, 1H, *J* = 3.4 Hz), 8.35(s, 1H), 8.36(s, 1H); ^{13}C-NMR (150 MHz, MeOD): δ 11.20, 20.77, 27.05, 27.64, 33.32, 48.23, 53.54, 57.29, 74.83, 75.16, 80.91, 91.90, 117.92 (q, *J* = 289.4 Hz),121.11, 143.47, 149.35, 150.09, 154.55, 162.44(q, *J* = 35.8 Hz),171.50; MS(ESI) m/z: 424 $[M+H]^+$; HRMS: calculated for $C_{18}H_{30}N_7O_5$ ($[M+H]^+$) 424.2308, found 424.2296.

3d R = Bn, 30% yield. ^{1}H-NMR (600 MHz, MeOD): δ 1.97–2.10(m, 4H), 2.31 (ddd, 1H, *J* = 15.8Hz, 5.8Hz, 3.2 Hz), 2.40–2.45(m, 1H), 3.57–3.59(m, 1H), 3.99(t, 1H, *J* =6.0 Hz), 4.12(d, 1H, *J* = 13.0 Hz), 4.20(d, 1H, *J* = 13.0 Hz), 4.41(t, 1H, *J* =6.0 Hz), 4.70(dd, 1H, *J* = 5.8Hz, 4.0 Hz), 5.49(s, 2H), 5.99(d, 1H, *J* = 3.8 Hz), 7.10(d, 2H, *J* = 7.2 Hz), 7.23(t, 2H, *J* = 7.2 Hz), 7.31(t, 1H, *J* = 7.2 Hz), 8.20(s, 1H), 8.33(s, 1H); ^{13}C-NMR (150 MHz, MeOD): δ 27.09, 27.87, 32.45, 53.71, 54.96, 57.08, 74.33, 74.84, 80.98, 91.88, 121.22, 130.22, 130.55, 130.68, 132.26, 142.88, 150.14, 151.49, 155.86, 162.55(q, *J* = 35.4 Hz), 171.75; MS(ESI) m/z: 472 $[M+H]^+$; HRMS: calculated for $C_{22}H_{30}N_7O_5$ ($[M+H]^+$) 472.2308, found 472.2299.

Protein expression and purification for the assays of enzymatic activities

Full-length SET7/9, SET8 (residues 191–395), SETD2 (residues 1347-1711, native and mutants), GLP (residues 951–1235), G9a (residues 913–1193), SUV39H2 (residues 112–410), PRMT1 (residues 10–352), PRMT3 (residues 211–531) and CARM1 (residues 19–608) were expressed and purified as previously reportedSee supporting information DOT1L (residues 1–420), SUV420H1 (residues 69–335), and SUV420H2 (residues 2–248) containing an N-terminal His tag were overexpressed in *E. coli* BL21 (DE3) V2R-pRARE (SGC) and purified by Ni-NTA column (Qiagen). The EZH2 complex containing EZH2 (residues 1–751), EED (residues 1–441) and SUZ12 (residues 1–739), and the MLL complex containing MLL (residues 3745–3969), WDR5 (residues 1–334) and RBBP5 (residues 1–538) were cloned in a pFastBac Dual vector (Invitrogen) with an N-terminal His_6-tag on MLL or EZH2. Both complexes were expressed in SF9 cells and purified by a Ni-NTA column. Additional purification steps were used if needed.

Biochemical assays of methylation activities

Three assays (the filter-paper assay, Scintillation Proximity Assay and the fiber filterplate assay) were used to determine the activities of methyltransferases according to the readiness of assay reagents and the characters of enzymes. The filter-paper assay, in which ^{3}H-Me of [^{3}H-Me]-SAM is transferred to peptide substrates, followed by filter binding and then quantification with a scintillation counter, was used to examine the activities of G9a, GLP, SUV39H2, SET7/9, SET8, SETD2 (wild type and mutants), PRMT1, PRMT3 and CARM1. The Scintillation Proximity Assay (SPA), in which ^{3}H-Me of [^{3}H-Me]-SAM is transferred to biotinylated substrates, followed by immobilization onto SPA-plate and Topcount plate reading, was applied to determine the activities of SUV420H1, SUV420H2, EZH2 complex, MLL complex, and DNMT1. The fiber filterplate assay, in which ^{3}H-Me of [^{3}H-Me]-SAM is transferred to a protein substrate, followed by acid precipitation, immobilization to fiber filterplate and Topcount plate reading, was used for DOT1L. The IC_{50} values were obtained by fitting inhibition percentage versus inhibitor concentration using GraphPad Prism5 or SigmaPlot software.

Crystallization, data collection and structure determination (see Supplementary Information for details)

Human SETD2 (residues 932–1208) with an N-terminal 6×histidine tag was overexpressed in *E. coli* BL21 (DE3) codon plus RIL strain (Stratagene) and purified by HiTrap Chelating column (GE Healthcare), Superdex200 column (GE Healthcare) and then ion-exchange chromatography. Purified SETD2 protein (10 mg/mL) was complexed with S-adenosyl-*L*-methionine (SAM, which was degraded into SAH during crystallization) (Sigma) at a 1:10 protein to ligand molar ratio, and crystallized using a sitting drop vapor diffusion method. 1 μl of the protein solution was mixed with 1 μl of the reservoir solution containing 30% PEG 2K MME and 0.1 M KSCN at 20°C. For the complex of SETD2 with Pr-SNF, the purified SETD2 was incubated with the inhibitor at a 1:5 protein to inhibitor molar ratio and crystallized using a sitting drop vapor diffusion method. Protein solution 1 μl was mixed with 1 μl of reservoir solution containing 20 % PEG4000, 10 % isopropanol, 0.1 M HEPES (pH 7.5) at 20 °C. X-ray diffraction data for the SETD2-SAH complex was collected at 100K at beamline19ID of Advanced Photon Source (APS). The methyltransferase domain structure of SETD2 in complex with SAH was solved by molecular replacement with human SUV39H2 (PDB code 2R3A) as the search model. The SEDT2-SAH structure was subsequently used as model to solve the structure of the Pr-SNF-SETD2 complex. Contoured omission electronic density maps for SAH, Pr-SNF, the auto-inhibitory post-SET loops were simulated to confirm the ligand binding or loop reconfiguration.

Molecular docking

To compare the SETD2-Pr-SNF complex with SAH-bound SETD2 and other PKMTs, the Pr-SNF ligand was overlaid with the SAM or SAH in complex with the structures of SETD2 (PDB code 4H12), SET8 (PDB code 1ZKK), MLL (PDB code 2W5Z), GLP (PDB code 2RFI), G9a (PDB code 3RJW), SUV39H2 (PDB code 2R3A), SUV420H1 (PDB code 3S8P) and SUV420H2 (PDB code 3RQ4) with ICM version 3.7-2b (Molsoft, San Diego).

Kinetics analysis

The filter-paper assay was used to determine the initial velocities of SETD2 in the presence of varied concentrations of H3 peptide substrate (residues 20–50) and [^{3}H-Me]-SAM cofactor. $K_{m,SAM}$, $K_{m,Substrate}$ and k_{cat} were obtained from the secondary double-reciprocal plots of the initial velocities versus the concentrations of substrate or cofactor according to Eq. S1–S2.[45] The kinetic parameters ($K_{m,SAM}$, $K_{m,Substrate}$ and k_{cat}) were further confirmed upon fitting the same set of initial velocities versus the concentrations of SAM and peptide substrate via non-linear regression (Fig. S1b). To obtain the inhibition constant K_d (K_i) of Pr-SNF on SETD2, the methylation kinetics were measured with the filter-paper assay in the presence of varied amounts of the inhibitor. The secondary double-reciprocal plots of the initial velocities versus the concentration of the inhibitor were the generated and further processed according to Eq. S3–S6 to give K_d.[45]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work is supported by Mr. William H. Goodwin and Mrs. Alice Goodwin Commonwealth Foundation for Cancer Research (M.L.), The Experimental Therapeutics Center of Memorial Sloan-Kettering Cancer Center (M.L.), NIGMS (1R01GM096056, M.L.), NINDS (R21NS071520, M.L.) and the Structural Genomics Consortium (J.M.), which is a registered charity (1097737) that receives funds from the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, Genome Canada through the Ontario Genomics Institute, GlaxoSmithKline, Karolinska Institute, The Knut and Alice Wallenberg Foundation, the Ontario Innovation Trust,

the Ontario Ministry for Research and Innovation, Merck & Co., the Novartis Research Foundation, the Swedish Agency for Innovation Systems, the Swedish Foundation for Strategic Research, and the Wellcome Trust.

References

1. Kouzarides T. Cell. 2007; 128:693. [PubMed: 17320507]
2. Kouzarides T. Cell. 2007; 128:802. [PubMed: 17320515]
3. Luo M. ACS Chem Biol. 2012; 7:443. [PubMed: 22220966]
4. Lee YH, Stallcup MR. Mol Endocrinol. 2009; 23:425. [PubMed: 19164444]
5. Spannhoff A, Hauser AT, Heinke R, Sippl W, Jung M. Chemmedchem. 2009; 4:1568. [PubMed: 19739196]
6. Copeland RA, Solomon ME, Richon VM. Nat Rev Drug Discov. 2009; 8:724. [PubMed: 19721445]
7. Min J, Feng Q, Li Z, Zhang Y, Xu RM. Cell. 2003; 112:711. [PubMed: 12628190]
8. Richon VM, Johnston D, Sneeringer CJ, Jin L, Majer CR, Elliston K, Jerva LF, Scott MP, Copeland RA. Chem Biol Drug Des. 2011; 78:199. [PubMed: 21564555]
9. Campagna-Slater V, Mok MW, Nguyen KT, Feher M, Najmanovich R, Schapira M. J Chem Inf Model. 2011; 51:612. [PubMed: 21366357]
10. Schapira M. Curr Chem Genomics. 2011; 5:85. [PubMed: 21966348]
11. Daigle SR, Olhava EJ, Therkelsen CA, Majer CR, Sneeringer CJ, Song J, Johnston LD, Scott MP, Smith JJ, Xiao Y, Jin L, Kuntz KW, Chesworth R, Moyer MP, Bernt KM, Tseng JC, Kung AL, Armstrong SA, Copeland RA, Richon VM, Pollock RM. Cancer Cell. 2011; 20:53. [PubMed: 21741596]
12. Yao Y, Chen P, Diao J, Cheng G, Deng L, Anglin JL, Prasad BVV, Song Y. J Am Chem Soc. 2011; 133:16746. [PubMed: 21936531]
13. Greiner D, Bonaldi T, Eskeland R, Roemer E, Imhof A. Nat Chem Biol. 2005; 1:143. [PubMed: 16408017]
14. Kubicek S, O'Sullivan RJ, August EM, Hickey ER, Zhan Q, Teodoro ML, Rea S, Mechtler K, Kowalski JA, Homon CA, Kelly TA, Jenuwein T. Mol Cell. 2007; 25:473. [PubMed: 17289593]
15. Ferguson AD, Larsen NA, Howard T, Pollard H, Green I, Grande C, Cheung T, Garcia-Arenas R, Cowen S, Wu J, Godin R, Chen H, Keen N. Structure. 2011; 19:1262. [PubMed: 21782458]
16. Gutierrez JA, Luo M, Singh V, Li L, Brown RL, Norris GE, Evans GB, Furneaux RH, Tyler PC, Painter GF, Lenz DH, Schramm VL. ACS Chem Biol. 2007; 2:725. [PubMed: 18030989]
17. Schramm VL. Annu Rev Biochem. 2011; 80:703. [PubMed: 21675920]
18. Horowitz S, Yesselman JD, Al-Hashimi HM, Trievel RC. J Biol Chem. 2011; 286:18658. [PubMed: 21454678]
19. Zhang XD, Bruice TC. Biochemistry. 2008; 47:2743. [PubMed: 18260647]
20. Zhang J, Ding L, Holmfeldt L, et al. Nature. 2012; 481:157. [PubMed: 22237106]
21. Varela I, Tarpey P, Raine K, et al. Nature. 2011; 469:539. [PubMed: 21248752]
22. Newbold RF, Mokbel K. Anticancer Res. 2010; 30:3309. [PubMed: 20944102]
23. Hu M, Sun XJ, Zhang YL, Kuang Y, Hu CQ, Wu WL, Shen SH, Du TT, Li H, He F, Xiao HS, Wang ZG, Liu TX, Lu H, Huang QH, Chen SJ, Chen Z. Proc Natl Acad Sci USA. 2010; 107:2956. [PubMed: 20133625]
24. Duns G, van den Berg E, van Duivenbode I, Osinga J, Hollema H, Hofstra RM, Kok K. Cancer Res. 2010; 70:4287. [PubMed: 20501857]
25. de Almeida SF, Grosso AR, Koch F, Fenouil R, Carvalho S, Andrade J, Levezinho H, Gut M, Eick D, Gut I, Andrau JC, Ferrier P, Carmo-Fonseca M. Nat Struct Mol Biol. 2011; 18:977. [PubMed: 21792193]
26. Dalgliesh GL, Furge K, Greenman C, et al. Nature. 2010; 463:360. [PubMed: 20054297]
27. Nimura K, Ura K, Shiratori H, Ikawa M, Okabe M, Schwartz RJ, Kaneda Y. Nature. 2009; 460:287. [PubMed: 19483677]
28. Kuo AJ, Cheung P, Chen K, Zee BM, Kioi M, Lauring J, Xi Y, Park BH, Shi X, Garcia BA, Li W, Gozani O. Mol Cell. 2011; 44:609. [PubMed: 22099308]

29. Morishita M, di Luccio E. Biochim Biophys Acta. 2011; 1816:158. [PubMed: 21664949]
30. Pasillas MP, Shah M, Kamps MP. Hum Mutat. 2011; 32:292. [PubMed: 21972110]
31. Cross NC. Oncotarget. 2012; 3:3. [PubMed: 22287508]
32. Yang P, Guo L, Duan ZJ, Tepper CG, Xue L, Chen X, Kung HJ, Gao A, Zou JX, Chen HW. Mol Cell Biol. 2012; 32:3121. [PubMed: 22645312]
33. Ghosh AK, Liu W. J Org Chem. 1996; 61:6175. [PubMed: 11667452]
34. Dubber M, Lindhorst TK. Synthesis. 2001:327.
35. Sun J, Dong Y, Cao L, Wang X, Wang S, Hu Y. J Org Chem. 2004; 69:8932. [PubMed: 15575776]
36. Yin X, Zhao G, Schneller SW. Tetrahedron Lett. 2007; 48:4809. [PubMed: 18612331]
37. Baldwin JE, Adlington RM, Mitchell MB. Tetrahedron. 1995; 51:5193.
38. Qian C, Zhou MM. Cell Mol Life Sci. 2006; 63:2755. [PubMed: 17013555]
39. Cheng XD, Collins RE, Zhang X. Annu Rev Biophys Biomol Struct. 2005; 34:267. [PubMed: 15869391]
40. An S, Yeo KJ, Jeon YH, Song JJ. J Biol Chem. 2011; 286:8369. [PubMed: 21239497]
41. Morishita M, di Luccio E. Biochem Biophys Res Commun. 2011; 412:214. [PubMed: 21806967]
42. Qiao Q, Li Y, Chen Z, Wang M, Reinberg D, Xu RM. J Biol Chem. 2011; 286:8361. [PubMed: 21196496]
43. Wu H, Min JR, Lunin VV, Antoshenko T, Dombrovski L, Zeng H, Allali-Hassani A, Campagna-Slater V, Vedadi M, Arrowsmith CH, Plotnikov AN, Schapira M. Plos One. 2010; 5:e8570. [PubMed: 20084102]
44. Krishnan S, Horowitz S, Trievel RC. Chembiochem. 2011; 12:254. [PubMed: 21243713]
45. Segel, IH. Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems. Wiley Classics Library; 1993.
46. Edmunds JW, Mahadevan LC, Clayton AL. Embo J. 2008; 27:406. [PubMed: 18157086]
47. Wagner EJ, Carpenter PB. Nat Rev Mol Cell Biol. 2010; 13:115. [PubMed: 22266761]
48. Li Y, Trojer P, Xu CF, Cheung P, Kuo A, Drury WJ 3rd, Qiao Q, Neubert TA, Xu RM, Gozani O, Reinberg D. J Biol Chem. 2009; 284:34283. [PubMed: 19808676]
49. t Hart P, Lakowski TM, Thomas D, Frankel A, Martin NI. Chembiochem. 2011; 12:1427. [PubMed: 21560220]
50. Mori S, Iwase K, Iwanami N, Tanaka Y, Kagechika H, Hirano T. Bioorg Med Chem Lett. 2010; 18:8158.

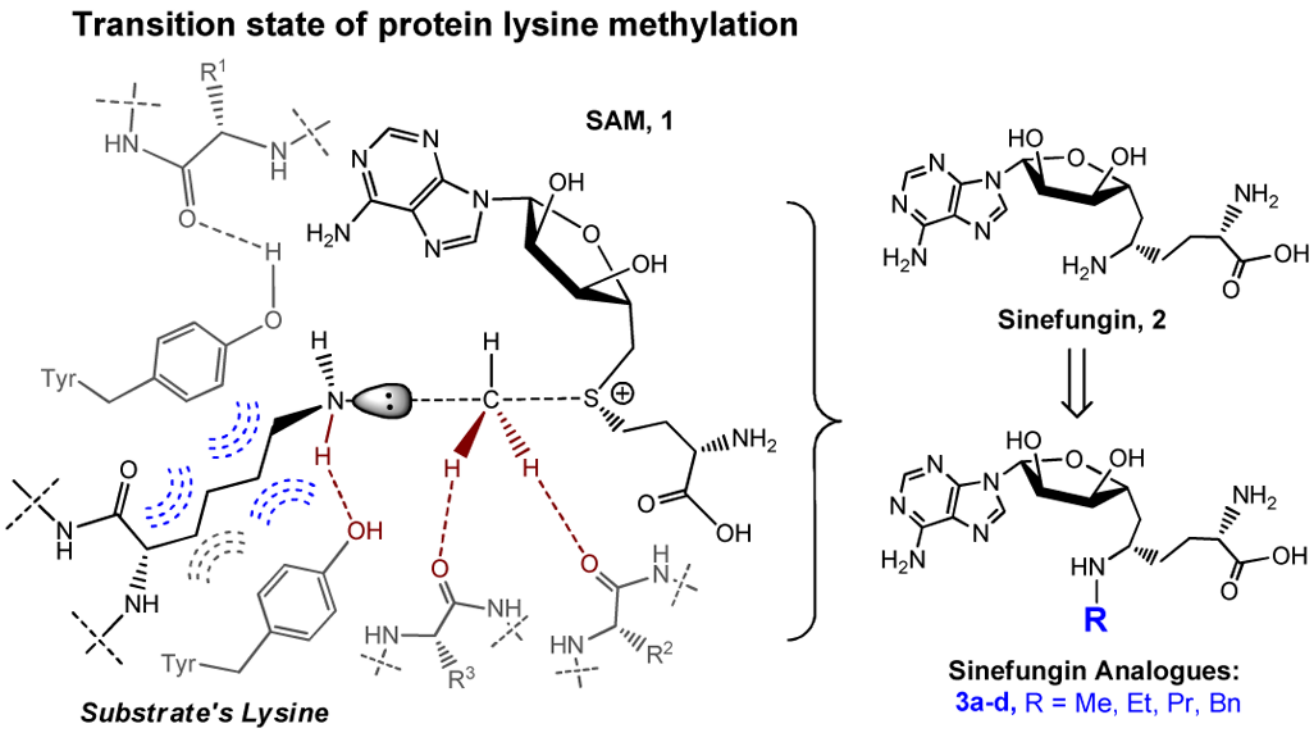


Figure 1. Structures of a proposed transition state of protein lysine methylation and sinefungin analogues as the transition-state mimics

The transition state is featured by classic hydrophobic interaction/hydrogen bonds with the lysine side chain of substrates and nonclassic carbon-oxygen hydrogen bonds with SAM's sulfonium methyl moiety. Sinefungin and its analogues 3a–d are expected to capture certain transition-state characters of specific PKMTs.

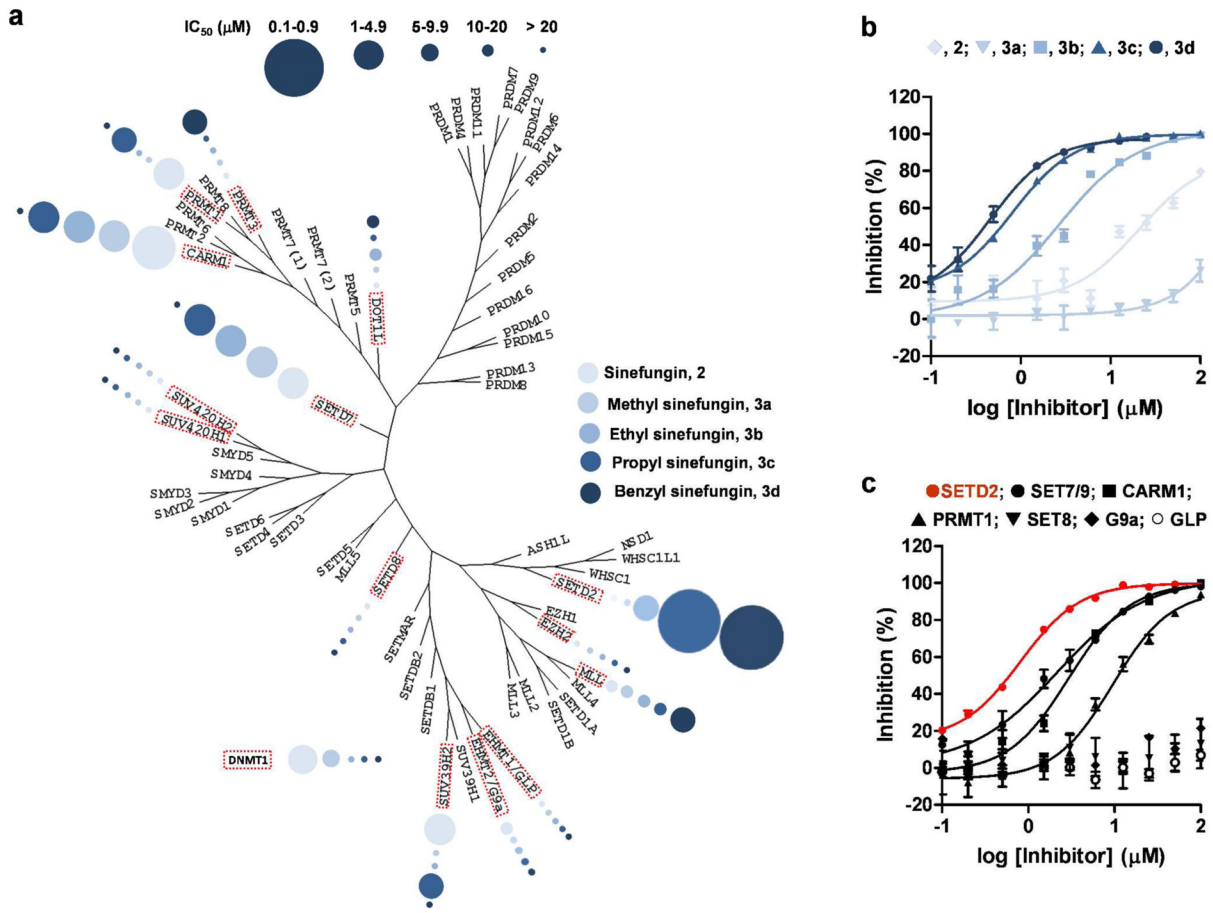


Figure 2. Inhibition profile of sinefungin and its analogues 3a–d against a panel of methyltransferases

(**a**) The magnitude of IC_{50} of sinefungin and its analogues **3a–d** is presented against 15 phylogenetically-arrayed methyltransferases (their IC_{50} values are listed in Table S1). The increased diameters and darkness of the circles reflect higher potency (lower IC_{50}) and larger sizes of the inhibitors, respectively. (**b**) Representative IC_{50} curves of sinefungin and its analogues **3a–d** against SETD2 with the values of 28.4 ± 1.5μM, >100 μM, 8.2 ± 1.2μM, 0.80 ± 0.02 μM and 0.48 ± 0.06 μM for sinefungin and **3a–d**, respectively. (**c**) IC_{50} curves of Pr-SNF **3c** against representative methyltransferases with the values of 0.80 ± 0.02 μM for SETD2, 2.2 ± 0.4μM for SET7/9, 2.96 ± 0.3 μM for CARM1, 9.5 ± 0.4 μM for PRMT1, >100 μM for SET8, G9a and GLP.

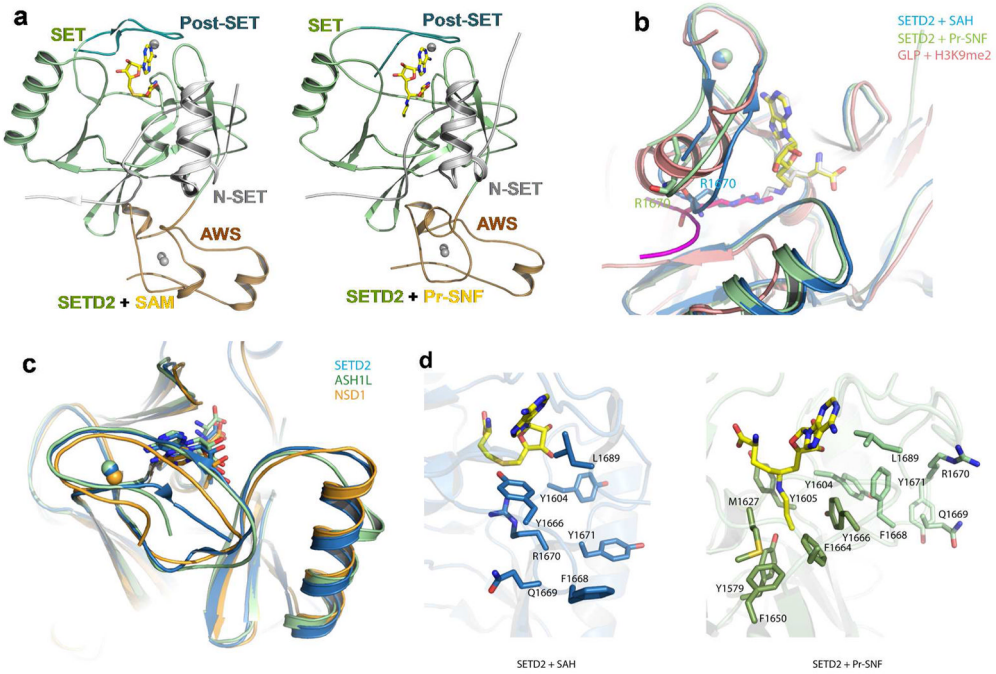


Figure 3. Structures of SETD2 in complex with SAH and Pr-SNF

(a) Overall crystal structures of SETD2 in complex with SAH (left) and Pr-SNF **3c** (right). Here the catalytic domain of human SETD2 (residues 1435–1711) contain the N-terminal extension motif, AWS domain (orange), SET domain (light green) and post-SET motif (cyan) with the ligands (SAH or Pr-SNF) highlighted. (**b**) Superposition of the binary complexes of SAH-SETD2 (blue-yellow, PDB code 4H12), Pr-SNF-SETD2 (green-white, PDB coddle 4FMU) and H3K9me2-GLP (mauve-magenta, PDB code 2RFI). Overlaid structures between SAH-SETD2 and H3K9me2-GLP show that the autoinhibitory loop in the SAH-SETD2 complex and its R1670 (blue) bind the site that would be otherwise occupied by the substrate (magenta) of GLP. In contrast, the loop and its characteristic R1670 are repositioned in the structure of Pr-SNF-SETD2 (green) and overlaid with the post-SET helix (mauve) for binding substrate. **(c)** Superposed structures of the inactive binary complex SAH-SETD2 (blue, PDB code 4H12) and its homologues, ASH1L (green, PDB code 3OPE) and NSD1 (orange, PDB code 3OOI). The three PKMTs have the similar autoinhibitory topology with their post-SET loops. **(d)** The key residues to stabilize the alternative configuration of the post-SET loop and interact with Pr-SNF's N-propyl chain. In the SETD2-Pr-SNF (right) but not SETD2-SAH (left) binary complex, SETD2's post-SET loop is glued at a hydrophobic core consisting of Tyr1604, Tyr1666, Phe1668, Leu1689 and the hydrocarbon side chain of Arg1670. In the SETD2-Pr-SNF, the N-propyl moiety was buried in the hydrophobic binding pocket formed by SETD2's Tyr1579, Tyr1605, Met1627, Phe1650, Phe1664 and Try1666.

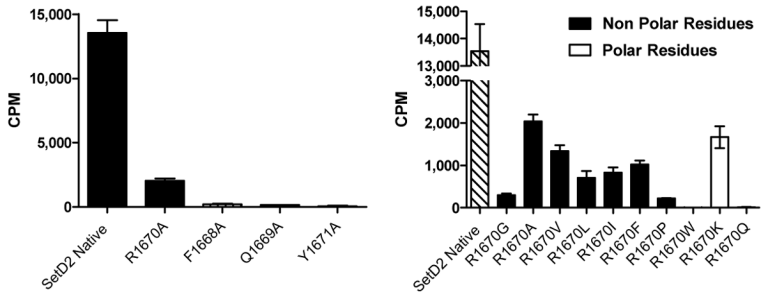


Figure 4. Effects of SETD2's enzyme catalysis

Methylation activities of SETD2 and its mutants were measured through the filter-paper assay with [^{3}H-Me]-SAM as cofactor. The F1668A, Q1669A, R1670G/A/V/L/I/F/P/W/K/Q and Y1671A mutations are within SETD2's post-SET loop, which regulates substrate access. The barely-detectable methylation activities of SETD2's F1668A. Q1669A and Y1671A mutants are consistent with their roles on either stabilizing SETD2's active open conformer or interacting with substrate. SETD2's 1670 site can tolerate modest hydrophobic substitutions (*e.g.* A/V/I/L/F) but not extreme ones (*e.g.* G/P/W).

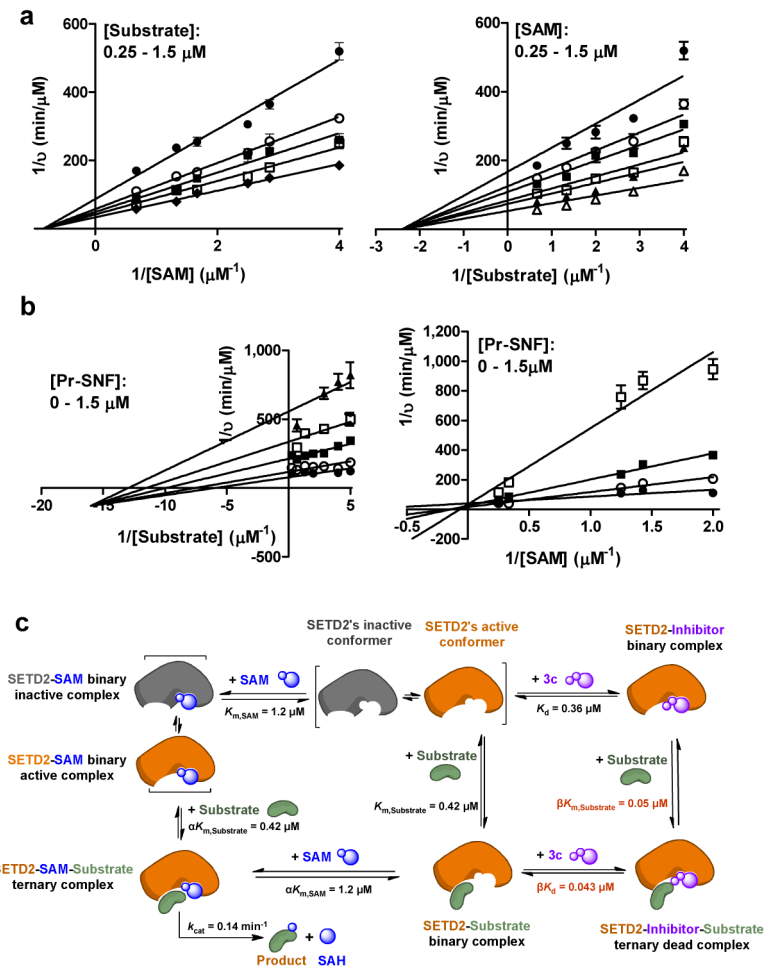


Figure 5. Methylation kinetics of SETD2

The kinetic analysis was carried out with varied concentrations of substrate and cofactor in the absence **(a)** or presence **(b)** of Pr-SNF. **(a)** Initial velocities for SETD2-mediated methylation were measured to generate the Lineweaver–Burk curve versus the concentration of the SAM cofactor (0.25 – 1.5 μM) (left), or the concentration of the H3K36 peptide substrate (0.25 – 1.5 μM) (right). The linear regressions, which converge on the X axis in both cases, suggest a random sequential mechanism. $K_{m,SAM}$ = 1.21 ± 0.05 μM, $K_{m,Substrate}$ = 0.42 ± 0.02 μM, α = 1 and k_{cat} = 0.14 ± 0.01min^{-1} were obtained by plotting the slopes of the Lineweaver–Burk double reciprocal curves against the concentrations of SAM and substrate (Fig. S1a and Supplementary Information). **(b)** Double reciprocal plots of initial velocities of SETD2-mediated methylation versus the concentration of substrate or SAM in the presence of Pr-SNF (0 – 1.5 μM) were generated by varying the concentration of substrate (0.25 – 5 μM) (left) in the presence of the fixed concentration of SAM (0.8 μM) or by varying the concentration of cofactor (0.5 – 4 μM) in the presence of the fixed concentration of the substrate (4μM) (right). The linear regressions, converging on the negative X axis against substrate and on the Y axis against SAM, are consistent with a noncompetitive mechanism between Pr-SNF and substrate, and a competitive mechanism between Pr-SAN and SAM. K_d = 360 ± 15nM and βK_d = 43 ± 4nM (β = 0.12 ± 0.01) were obtained by plotting the slopes of the double reciprocal curves against the concentration of Pr-SNF (Fig. S1b and Supplementary Information). **(c)** Overall mechanisms and kinetic parameters of SETD2-mediated methylation in the absence or presence of Pr-SNF. SETD2-

mediated methylation goes through a random sequential mechanism and can be inhibited by Pr-SNF via forming either the Pr-SNF-SETD2 binary complex or the Pr-SNF-SETD2-substrate ternary complex.

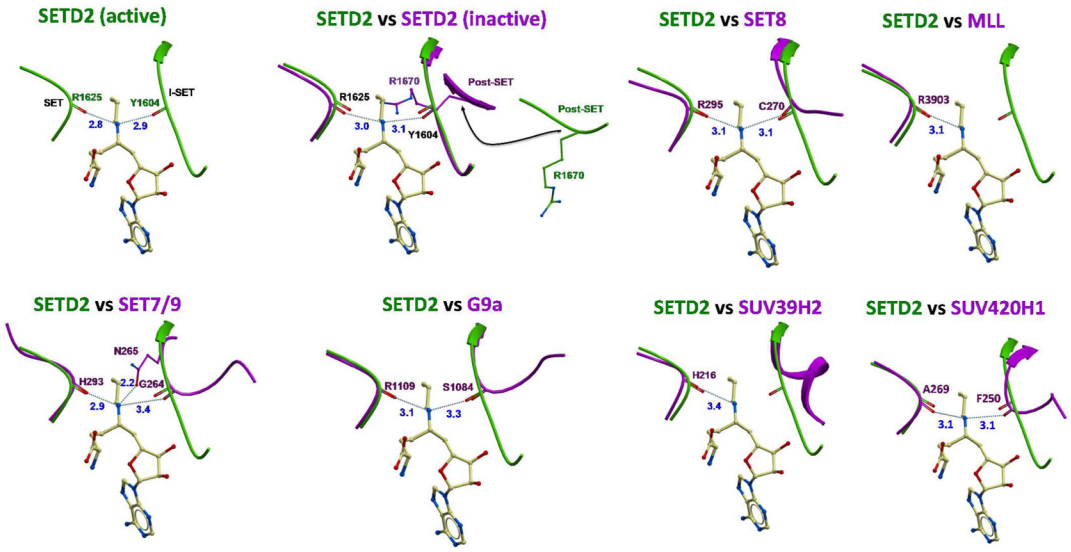
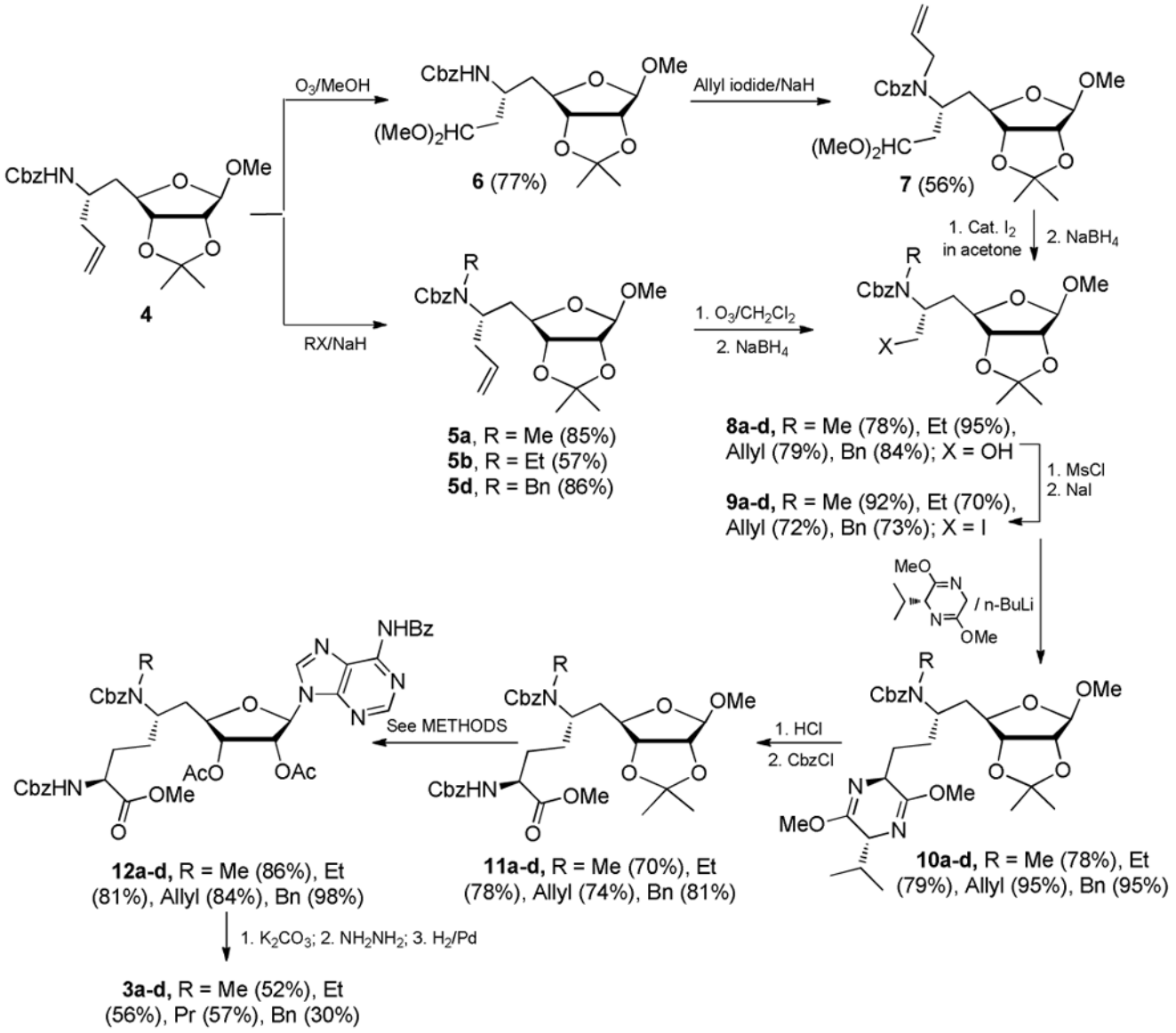


Figure 6. Structural insights of Pr-SNF's selectivity on SETD2 over other PKMTs

Pr-SNF ligand in the Pr-SNF-SETD2 binary complex (PDB code4FMU) was overlaid with SAM or SAH in complex with SET7/9 (PDB code 1O9S), SETD2's autoinhibitory conformer (PDB code 4H12), SET8 (PDB code 1ZKK), MLL (PDB code 2W5Z), G9a (PDB code 3RJW), GLP (PDB code 2RFI), SUV39H2 (PDB code 2R3A), SUV420H1 (PDB code 3S8P) and SUV420H2 (PDB code 3RQ4) with ICM version 3.7-2b (Molsoft, San Diego). The overlaid structures of GLP and SUV420H2 were omitted given their similarity to those of G9a and SUV420H1, respectively. The optimal hydrogen-bonding distances between Pr-SNF's secondary amine and main-chain carbonyls observed in the SETD2 structure may contribute to the high selectivity of the inhibitor.



Scheme 1.
Synthesis of sinefungin analogues 3a–d.