Convergent evolution of chromatin modification by structurally distinct enzymes: comparative enzymology of histone H3 Lys27 methylation by human polycomb repressive complex 2 and vSET

Brooke M. SWALM¹,², Kenneth K. HALLENBECK¹, Christina R. MAJER², Lei JIN, Margaret Porter SCOTT², Mikel P. MOYER², Robert A. COPELAND² and Tim J. WIGLE²,³

Epizyme Inc., 400 Technology Square, Cambridge, MA 02139, U.S.A.

INTRODUCTION

Histone lysine methylation is a dynamic process that exerts control over gene expression by influencing the conformational state of chromatin and by also serving as a recognition locus for the recruitment of transcription factors, enzymes and other proteins to selective locations on the histone. Sequential methylation of lysine results in four distinct states of the lysine side chain (zero, one, two or three methyl groups on the ε-nitrogen), each associated with unique transcriptional consequences. All of these methylation reactions are catalysed by a family of enzymes known as the PKMTs (protein lysine methyltransferases). PKMTs universally use SAM (S-adenosylmethionine) as the methyl group donor and transfer the methyl group to the ε-nitrogen of a lysine side chain through an SN2 reaction mechanism [1,2]. With the exception of the enzyme DOT1L, all known human PKMTs share a similar catalytic domain of approximately 130 amino acids (13.5 kDa) that aligns well with the canonical SET domain [3]. The SET domain contains recognition elements for SAM and lysine binding and confers substrate specificity through interactions between the histone protein surrounding the methyl-accepting lysine and enzyme residues in close proximity to the mouth of the lysine-binding channel [4,5].

Methylation of H3K27 (histone H3 Lys27) is a transcriptionally repressive mark that plays a critical role during development and differentiation, and is implicated in several forms of human cancer [6–14]. In multicellular organisms, this reaction is exclusively catalysed by a multi-protein complex referred to as PRC2 (polycomb repressive complex 2) that contains either of the SET-domain PKMTs EZH (enhancer of zeste homologue) 1 (PRC2EZH1) or EZH2 (PRC2EZH2) [15].

This same reaction is also catalysed by a PKMT of Paramecium bursaria chlorella virus 1, known as vSET [16–19]. vSET represents the minimal protein structural unit for PKMT activity, consisting of 119 amino acids (∼13.5 kDa) that aligns well with the canonical SET domain of larger PKMTs [16]. The small size of this viral enzyme makes it ideal for structural studies aimed at understanding the critical elements of substrate recognition and catalysis by PKMTs. Indeed, high-resolution crystal structures of vSET have previously been reported (PDB codes 3KMT, 3KMJ and 3KMA) [17,19,20]. In contrast, the large multi-protein PRC2EZH2 complex has to date not been amenable to crystallographic methods. Given the keen interest in PRC2EZH2 as a potential target for cancer drug discovery [21,22], we speculated whether it is possible to use vSET as a structurally defined surrogate of PRC2EZH2 for the design of inhibitors and other active site-directed ligands. This approach rests on the reasonable assumption that the common enzymatic activity of PRC2 and vSET is conferred by common structural features of molecular recognition.

In the present study we test this underlying assumption by comparing the enzymatic details of catalysis by these two similar preferences for methylation of H3K27, they diverge in terms of their permissiveness for catalysing methylation of alternative histone lysine sites, their relative preferences for utilization of multimeric macromolecular substrates, their active site primary sequences and, most importantly, their sensitivity to inhibition by drug-like small molecules. The cumulative data led us to suggest that EZH2 and vSET have very distinct active site structures, despite the commonality of the reaction catalysed by the two enzymes. Hence, the EZH2 and vSET pair of enzymes represent an example of convergent evolution in which distinct structural solutions have developed to solve a common catalytic need.

Key words: convergent evolution, epigenetics, enhancer of zeste homologue 2 (EZH2), histone H3 Lys27 (H3K27), methyltransferase, vSET.
enzymes. We find that, although the overall mechanism of catalysis is similar for the two enzymes, active site interactions with ligands nevertheless differ in significant ways. Together, these two enzymes thus represent an example of convergent evolution of enzyme function [23,24].

MATERIALS AND METHODS
Reagents and equipment
The vSET gene was synthesized based on a published sequence (GenBank® accession number AAC96946.1) and cloned into the pET28a vector (Novagen) between NdeI and HindIII sites with a N-terminal His tag and thrombin cleavage site. The protein was expressed in BL21-Gold(DE3) Escherichia coli (Agilent) in LB medium with 50 μg/ml kanamycin and induced by 0.3 mM IPTG at 16 °C for 16 h. The harvested cell pellet was suspended in lysis buffer containing 25 mM Tris/HCl, pH 7.6, 300 mM NaCl, 5 % (v/v) glycerol and 5 mM 2-mercaptoethanol, and cells were lysed by sonication. Cell debris was cleared by centrifugation and the supernatant was added on to a Ni-NTA (Ni²⁺-nitrilotriacetate; Qiagen) column equilibrated with lysis buffer. The Ni-NTA column was first washed with lysis buffer, followed by lysis buffer supplemented with 20 mM imidazole, then with buffer containing 25 mM Tris/HCl, pH 8, 300 mM NaCl, 5 % (v/v) glycerol, 5 mM 2-mercaptoethanol and 50 mM imidazole. The protein was eluted with buffer containing 25 mM Tris/HCl, pH 8, 200 mM NaCl, 5 % (v/v) glycerol, 5 mM 2-mercaptoethanol and 250 mM imidazole. Fractions containing the target protein were pooled, dialysed against buffer to remove imidazole and 250 mM imidazole. Fractions containing vSET were pooled and dialysed against lysis buffer to remove imidazole and then concentrated. The concentrated sample was loaded on to a S-75 column (GE Healthcare) and equilibrated with buffer containing 20 mM Tris/HCl, pH 8, 200 mM NaCl and 5 mM 2-mercaptoethanol. Fractions containing vSET were pooled and concentrated for assay use, yielding a stock that was >90 % pure as judged by capillary electrophoresis.

Four-component PRC2[2Z2H] (EZH2, Suz12, EED and RbAp48) was purified to >95 % purity and 1:1 stoichiometry (judged by capillary electrophoresis) as previously described using a FLAG tag on the EED subunit [11]. Chicken erythrocyte mono- and oligo-nucleosomes were purified as described previously [25]. Flashplates (384-well) and Microscint 0 scintillation fluid were purchased from PerkinElmer. Multiscan HTS glass fibre filter-binding plates (96-well) were obtained from Millipore. [3H]SAM was obtained from American Radiolabeled Chemicals with a specific activity of 80 Ci/mmol. Unlabelled SAM, SAH (S-adenosylhomocysteine) and Sinefungin were obtained from Sigma–Aldrich. Recombinant histone H3 was purchased from New England Biolabs, and recombinant histone H3/H4 tetramer and recombinant histone H2A/H2B/H3/H4 octamer were produced by XTAL Biotechnologies at >95 % purity. All peptides were synthesized and HPLC purified to >95 % purity by 21st Century Biochemicals. EPZ005687 was synthesized by Epizyme as previously described [21] and GSK126 was purchased from Xcessbio.

Flashplates and filter-binding plates were read on a TopCount NXT microplate reader (PerkinElmer), and flashplates were washed in a Bioket Elx-405 with 0.1 % Tween 20 before being read. All enzymatic assays were performed in 384-well and 96-well V-bottom polystyrene microplates (Greiner).

Scanning for activity against a protein substrate panel using filter-binding microplate assays
For histone substrates, 1× assay buffer containing 20 mM Bicine, pH 7.6, 0.002 % Tween 20, 0.005 % bovine skin gelatin and 0.5 mM DTT was used, and, when used with nucleosomes, the same assay buffer was supplemented with 100 mM KCl. Reactions (50 μl) were carried out at 25 °C in 96-well polystyrene microplates and contained 10 nM enzyme, 200 nM [3H]SAM and 200 nM protein substrate. Reactions were terminated by adding an excess of unlabelled SAM to outcompete the incorporation of [3H]SAM. Quenched reactions were added to 96-well filter-binding plates and the membranes were washed three times with 200 μl of 10 % tricarboxylic acid followed by washing once with 200 μl of 95 % ethanol. The membranes were air dried and 30 μl of Microscint 0 was added before reading on a TopCount NXT instrument.

Scanning for enzymatic activity against a histone peptide panel using flashplate format
A library of biotinylated histone peptides was solubilized in either water or DMSO, and 1 μl was spotted into 384-well polystyrene microplates. Reactions (50 μl) were carried out at 25 °C in 384-well polystyrene microplates and contained 8 nM vSET or 8 nM PRC2[2Z2H]. 250 nM [3H]SAM and 1 μM peptide. Reactions were terminated after 1 h by adding an excess of unlabelled SAM to outcompete the incorporation of [3H]SAM. The reaction mixture was transferred to a flashplate, incubated for 1 h at room temperature (25 °C), then washed with 0.1 % Tween 20 and read on a TopCount NXT instrument.

Determination of steady-state mechanism using flashplate format with peptide substrates
Peptide and SAM were titrated, and 8 nM vSET or 8 nM PRC2[2Z2H] were added to initiate the reaction. Reactions (50 μl) were carried out in assay buffer at 25 °C in 384-well polystyrene microplates and samples at various time points were taken by adding an excess of unlabelled SAM to outcompete the incorporation of [3H]SAM. The reaction mixture was transferred to a flashplate, incubated for 1 h at room temperature, then washed with 0.1 % Tween 20 and read on a TopCount NXT instrument. Double substrate titrations were fitted to the following equation for a ternary complex mechanism to determine steady-state Kₐ values:

\[
\frac{E}{[A][B]} = \frac{k_{cat}[E][A][B]}{K_A K_B + K_A [A] + K_B [B] + [A][B]}
\]

where \( E \) is enzyme, \( A \) is substrate \( A \) (SAM) and \( B \) is substrate \( B \) (peptide).

Self-assembled monolayer desorption/ionization MS analysis of peptide methylation
Reactions (50 μl) were carried out in assay buffer at 25 °C and contained 4 nM vSET enzyme, 50 nM peptide substrate and 1 μM SAM. Reactions were terminated by the addition of 100 mM NaCl and a 2 μl sample of each reaction was analysed by SAMDI Tech using self-assembled monolayer desorption/ionization time-of-flight MS [26].

Determination of enzyme inhibition \( K_i \) values
Inhibitors were pre-incubated with vSET or PRC2[2Z2H] in assay buffer for 30 min at 25 °C. Reactions (50 μl) were initiated by the addition of SAM and a peptide representing histone H3 residues 21–44 containing C-terminal biotin (appended to a C-terminal amide-capped lysine). The final concentrations
enzymatic characteristics ([28] and T. Wigle, unpublished work). Hence, for the rest of the present paper, we focus our attention on comparisons between PRC2EZH2 and vSET; unless otherwise stated, the characteristics reported for PRC2EZH2 are similar to those found for PRC2EZH1.

The degree of sequence identity between the human and viral proteins, on the other hand, is quite low (23% identity relative to the SET domain of EZH2). Despite the significant differences between the enzyme active sites at the primary sequence level, the data summarized in Figure 1(A) do not necessarily imply significant differences in protein folding, hence three-dimensional architecture. However, a clear distinction in enzymatic activity between PRC2EZH2 and vSET is realized at the three-dimensional structure level in that vSET folds into a tertiary structure that supports H3K27 methylation as a homodimer [20], without the need for additional protein subunits (Figure 1B), whereas EZH2 is only active in the context of the PRC2 complex (Figure 1C), requiring a minimum of two additional protein subunits, EED and Suz12, for activity and is optimally active in a four- or five-component complex with RbAp48 and AEBP2 (adipocyte enhancer-binding protein 2) [30,31]. In the present study, we explore similarities and differences between vSET and PRC2EZH2 enzymatic activity under identical buffer and temperature conditions.

Despite these structural differences, PRC2EZH2 and vSET display similar preferences for substrates containing the equivalent of the H3K27 site. To investigate this, we tested the ability of each enzyme to catalyse methylation of large physiologically relevant substrates, i.e. nucleosomes and histones, and a peptide library representing all lysine sites on human histones H3 and H4. Figure 2 reveals that oligonucleosome substrates stimulate higher activity from both vSET and PRC2EZH2 than do mononucleosome substrates. These data are consistent with the ability of these enzymes to rapidly propagate the methylation of H3K27 along contiguous stretches of chromatin. Both enzymes are postulated to use H3K27me3 (trimethylated H3K27) recognition on one nucleosome unit to anchor and physically place the enzyme in close proximity to a neighbouring unmethylated H3K27 residue to enhance the efficiency of methylation. Although the overall ability to use H3K27me3, the product of both enzymes’ enzymatic activity, to stimulate additional proximal H3K27 methylation is conceptually similar for PRC2EZH2 and vSET, this process differs at the structural and mechanical level. vSET accomplishes this nucleosome-walking using anti-co-operative homodimers, where one inactive vSET molecule bound to H3K27me3 places an activated vSET molecule in close proximity to an unmethylated H3K27 residue [20]. PRC2EZH2, however, utilizes the WD40 repeat domain present in EED subunit to recognize H3K27me3 and orientate the EZH2 subunit to a proximal unmethylated H3K27 residue [32,33].

Additionally, both enzymes show distinct preferences for recombinant histone H3 as substrate, with little or no activity detectable on recombinant histone H4. A clear difference between the enzymes is apparent when histone H3/H4 tetramers and histone H2A/H2B/H3/H4 octamers are used as substrates. PRC2EZH2 shows a significant enhancement of activity on these substrates, whereas the activity for vSET is actually reduced with these substrates relative to recombinant histone H3. This observation can be rationalized by previous studies indicating that the RbAp48 subunit of PRC2 binds to histone H4 [34,35], which is present in the recombinant histone tetramers and octamers. Combining this observation with the recognition of H3K27me3 by EED, PRC2EZH2 activity appears to be more dependent on recognition elements distal to the direct site of methyl transfer (e.g. H3K27) in the context of protein multimers as substrates.
To investigate further the specificity of each enzyme for particular amino acid sequences, the velocity of methylation catalysed by PRC2EZH2 and vSET was measured against a set of overlapping peptides walking the length of histone H3 and histone H4 in five-amino-acid increments. Inspection of Figure 3 reveals that there is a strong preference of both enzymes for histone H3 peptides. vSET showed modest activity on peptides that contained H3K9 (histone H3 Lys9), and much greater activity on those that contained H3K27. In contrast, PRC2EZH2 was more selective and only showed activity on peptides that contained H3K27. The ARKS amino acid sequence is common between selective and only showed activity on peptides that contained H3K27. The ARKS amino acid sequence is common between the H3K27 residue and the H3K27me3 peptide when SAM is saturating. Conversely, enzyme activation observed when using the H3K27me3 peptide does not allow one to distinguish between a compulsory or random ordered mechanism of catalysis.

Both PRC2EZH2 and vSET catalyse the mono-, di- and tri-methylation of H3K27. Previous data have demonstrated that human PRC2EZH2 displays a clear pattern of substrate use with respect to the methylation state of H3K27, thus PRC2EZH2 is most efficient at catalysing the first methylation reaction (from zero to one methyl group), and progressively less efficient at catalysing di- and tri-methylation of H3K27. As shown in Figure 5(A), vSET shows this same pattern of substrate utilization. The kinetic parameters determined from the steady-state experiments illustrated in Figure 5(A) are summarized in Table 1, where they are compared with the corresponding values for human PRC2EZH2.

The sequential catalysis of mono-, di- and tri-methylation of H3K27 can be accomplished by either a processive or distributive mechanism of catalysis. Processive catalysis presupposes that the enzyme remains bound to a particular substrate molecule until all three rounds of methylation have been completed. In contrast, a distributive mechanism implies dissociation and rebinding of enzyme and substrate after each round of catalysis. Recent studies of the combined activities of the wild-type and Tyr641/Ala677 mutant PRC2EZH2 in subsets of non-Hodgkin’s lymphoma cell lines clearly demonstrate that the enzyme functions through a distributive mechanism of catalysis [11]. To determine whether vSET operates by a processive or distributive mechanism, we followed the accumulation of each reaction product (mono-, di- and tri-methylated peptide) as a function of time after initiating the reaction with unmethylated peptide substrate. The results of these studies are illustrated in Figure 5(B). For a distributive mechanism the various reaction products are expected to accumulate sequentially so that one would expect to see a monotonic diminution of the substrate concentration, transient and sequential accumulation then loss of the monomethyl and dimethyl intermediates, and a lag followed by saturable accumulation of the final trimethyl product. In contrast, for a processive mechanism one would expect to see essentially

Figure 2  The activity of vSET and PRC2EZH2 on a panel of nucleosome and histone substrates

The activity of the enzymes was compared by performing a reaction with 8 nM enzyme, 200 nM [3H]SAM and 200 nM of the indicated substrate. Mono- and oligo-nucleosomes were purified from chicken erythrocytes, whereas recombinant tetramer and octamer respectively refer to histone H3A4 tetramers or histone H2A/H2B/H3/H4 octamers. The activity was measured by capturing the substrates on to a filter binding plate and scintillation counting to determine the incorporation of [3H-labelled methyl groups. Results are means ± S.D. for three experiments.
monotonic diminution of substrate and accumulation of the final product, with no accumulation of any intermediate species beyond the concentration of the enzyme present in the reaction (4 nM in the present case). The data presented in Figure 5(B) unambiguously demonstrate that vSET methylates H3K27 by a distributive mechanism, in concordance with previously reported data for this enzyme and for human PRC2EZH2.

To investigate further the apparent similarities between the catalytic mechanisms of PRC2EZH2 and vSET, the interaction of each enzyme with a series of ligands directed at the SAM- and lysine-binding pockets were compared. Table 1 contrasts the $K_m$ values of the methyl donor SAM and histone H3 substrate peptides containing un-, mono- and di-methylated Lys$^{27}$ for each enzyme. In addition, the $K_i$ of the SAM-mimetic sinefungin and the product-based inhibitors SAH or histone H3 peptide with trimethylated Lys$^{27}$ were directly compared. Inspection of these data reveals that PRC2EZH2 and vSET appear to have a similar affinity for nucleoside-based ligands, but differ significantly in their interactions with histone H3-based ligands. The $K_m$ values for SAH and $K_i$ values for SAH and $K_i$ are within 3-fold of one another, whereas sinefungin does not appreciably inhibit either enzyme up to 100 μM. However, the $K_m$ values for histone H3 peptides with zero, one or two methyl groups on Lys$^{27}$ vary between the two

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Type</th>
<th>$K_m$ (μM)</th>
<th>$K_i$ (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAM</td>
<td>Substrate</td>
<td>1.300 $±$ 0.320</td>
<td>1.244 $±$ 0.195</td>
</tr>
<tr>
<td>SAH</td>
<td>Product</td>
<td>1.299 $±$ 0.260</td>
<td>6.907 $±$ 0.210</td>
</tr>
<tr>
<td>Sinefungin</td>
<td>Inhibitor</td>
<td>$&gt;50$</td>
<td>$&gt;50$</td>
</tr>
<tr>
<td>H3 21–44, K27me0</td>
<td>Substrate</td>
<td>0.021 $±$ 0.003</td>
<td>0.157 $±$ 0.012</td>
</tr>
<tr>
<td>H3 21–44, K27me1</td>
<td>Substrate</td>
<td>0.012 $±$ 0.006</td>
<td>0.337 $±$ 0.026</td>
</tr>
<tr>
<td>H3 21–44, K27me2</td>
<td>Substrate</td>
<td>0.008 $±$ 0.004</td>
<td>0.144 $±$ 0.011</td>
</tr>
<tr>
<td>H3 21–44 K27me3</td>
<td>Product</td>
<td>0.810 $±$ 0.170</td>
<td></td>
</tr>
<tr>
<td>EPZ005687</td>
<td>Inhibitor</td>
<td>$&gt;100$</td>
<td>0.019 $±$ 0.006†</td>
</tr>
<tr>
<td>GSK126</td>
<td>Inhibitor</td>
<td>$&gt;100$</td>
<td>0.002 $±$ 0.001†</td>
</tr>
</tbody>
</table>

†This value is the mean $±$ S.D. for six replicate experiments.

As previously described [11,38,39], the peptides displayed sigmoidal behaviour with PRC2EZH2 and the data were fitted using a $K_{1/2}$ calculation rather than classic Michaelis–Menten fits to determine the concentration of peptide resulting in half-maximal velocity.

Table 1 Comparing the interaction of vSET and PRC2EZH2 with a series of active site ligands
enzymes 8–28-fold. Even more profound is the effect seen when attempting to measure the inhibition of enzymatic activity using an H3 peptide bearing a trimethylated Lys37. vSET is inhibited by this product peptide with a Ki of 0.8 μM, whereas PRC2 activity is enhanced 2–3-fold over the same concentration range of product peptide. The activation of EZH2 by the H3K27me3 product is consistent with a previous study on this effect [32] and is more evidence that PRC2 is more tightly regulated by interactions with distal sites on the nucleosome substrate.

Perhaps the most striking difference between vSET and PRC2 is their relative sensitivity to small organic SAM-interactions with distal sites on the nucleosome substrate.

In contrast with the potent inhibition of PRC2 EZH2 by this product peptide with a Ki of 0.8 μM, whereas PRC2 activity is enhanced 2–3-fold over the same concentration range of product peptide. The activation of EZH2 by the H3K27me3 product is consistent with a previous study on this effect [32] and is more evidence that PRC2 is more tightly regulated by interactions with distal sites on the nucleosome substrate.

In summary, despite the commonality of reaction catalysed, there appear to be meaningful differences in active site structure and interactions with small-molecule inhibitors between PRC2 and vSET that suggest convergent evolution of H3K27 methylation function for these two evolutionarily distant enzymes. Despite similarities in the interaction and utilization of nucleoside-based ligands, there are several key differences that exist between these enzymes with respect to size, overall structure and substrate recognition. Furthermore, the lack of inhibition of vSET by EPZ005687 and GSK126 reflects significant divergence of active site structure between these two enzymes that fundamentally precludes the use of vSET as a meaningful surrogate for PRC2 in studies aimed at understanding the structure–function relationship for the latter enzyme and especially in the design of pharmacological agents based on inhibition of enzyme activity.

AUTHOR CONTRIBUTION
Margaret Scott, Mikel Moyer, Robert Copeland and Tim Wigle designed the experiments. Brooke Swalm, Kenneth Hallenbeck, Christina Majer and Tim Wigle performed the experiments. Lei Jin and Margaret Scott supplied the enzymes. Robert Copeland and Tim Wigle wrote the paper.

ACKNOWLEDGMENTS
We thank Mr Michael Scholle at SAMDI Tech for assistance with self-assembled monolayer desorption/ionization time-of-flight MS and Dr Kevin Kuntz for supplying the inhibitor EPZ005687 and for helpful discussion in the preparation of this paper.

FUNDING
This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

REFERENCES
Convergent evolution of histone H3 Lys$^{27}$ methylation by human PRC2 and vSET


27 Cheng, Y. and Prusoff, W. H. (1973) Relationship between the inhibition constant ($K_i$) and the concentration of inhibitor which causes 50 per cent inhibition ($I_{50}$) of an enzymatic reaction. Biochem. Pharmacol. 22, 3099–3108


29 Reference deleted


Received 25 March 2013/6 May 2013; accepted 17 May 2013
Published as BJ Immediate Publication 17 May 2013, doi:10.1042/BJ20130439


