Biochemical reconstitution and phylogenetic comparison of human SET1 family core complexes involved in histone methylation

Stephen A. Shinsky, Kelsey E. Monteith, Susan Viggiano and Michael S. Cosgrove

Department of Biochemistry and Molecular Biology, SUNY Upstate Medical University
750 East Adams Street, Syracuse, New York 13210

Running title: SET1 family complexes have different product specificities

To whom correspondence should be addressed: Michael S. Cosgrove, Ph.D., Department of Biochemistry and Molecular Biology, SUNY Upstate Medical University, 750 East Adams Street, Syracuse, NY 13210, Email: cosgrovm@upstate.edu

Keywords: SET1, Mixed Lineage Leukemia, MLL, product specificity, phylogeny, histone methylation, WDR5, RbBP5, Ash2L, DPY-30

Background: The six human SET1 family core complexes catalyze methylation of histone H3 lysine 4 (H3K4).

Results: Different SET1 family core complexes catalyze different levels of H3K4 methylation.

Conclusion: Product specificity of the SET1 family is correlated with evolutionary lineage.

Significance: Core complex subunits differentially regulate the product specificity of different SET1 family members.

ABSTRACT

Mixed Lineage Leukemia protein-1 (MLL1) is a member of the SET1 family of histone H3 lysine 4 (H3K4) methyltransferases that are required for metazoan development. MLL1 is the best-characterized human SET1 family member, which includes MLL1-4 and Setd1A/B. MLL1 assembles with WRAD (WDR5, RbBP5, Ash2L, DPY-30) to form the MLL1 core complex, which is required for H3K4 dimethylation and transcriptional activation. Since all SET1 family proteins interact with WRAD in vivo, it is hypothesized they are regulated by similar mechanisms. However, recent evidence suggests differences among family members that may reflect unique regulatory inputs in the cell. Missing is an understanding of the intrinsic enzymatic activities of different SET1 family complexes under standard conditions. In this investigation, we reconstituted each human SET1 family core complex and compared subunit assembly and enzymatic activities. We found that in the absence of WRAD, all but one SET domain catalyzes at least weak H3K4 monomethylation. In the presence of WRAD, all SET1 family members showed stimulated monomethyltransferase activity, but differed in their di- and trimethylation activities. We found that these differences are correlated with evolutionary lineage, suggesting these enzyme complexes have evolved to accomplish unique tasks within metazoan genomes. To understand the structural basis for these differences, we employed a “phylogenetic scanning mutagenesis” assay and identified a cluster of amino acid substitutions that confer a WRAD-dependent gain-of-function dimethylation activity on complexes assembled with the MLL3 or Drosophila Trithorax proteins. These results form the basis for understanding how WRAD differentially regulates SET1 family complexes in vivo.

Methylation of histone H3 at lysine 4 (H3K4) is required for the epigenetic maintenance of transcriptionally active forms of chromatin in eukaryotes (1). H3K4 can be mono-, di- or trimethylated, with each state associated with distinct genomic locations and functional outcomes (2). For example, H3K4 trimethylation (H3K4me3) is enriched in the promoters of active genes, whereas H3K4me2 is localized throughout open reading frames (3,4). H3K4me1 is enriched in active enhancer elements (5) and is associated
SET1 family complexes have different product specificities

with gene silencing in lower eukaryotes (6-9). How cells achieve and maintain different H3K4 methylation states at distinct genomic regions is an important unsolved question in biology. This is particularly important given that genetic alteration of the genes that encode the enzymes that regulate H3K4 methylation are associated with a number of human developmental disorders and cancers (10-17).

In humans, Mixed Lineage Leukemia protein-1 (MLL1, also known as HRX, ALL1, KMT2A) catalyzes methyl group transfer from s-adenosylmethionine (AdoMet) to the epsilon amino group of H3K4 (18), and is frequently altered in poor prognosis acute leukemias (19-21). MLL1 is a member of the SET1 family of H3K4 methyltransferases, which is conserved from yeast to humans. It has been suggested that in budding yeast, SET1p is the sole H3K4 methyltransferase, as deletion of the SET1 gene results in global reduction of H3K4 methylation (6).

Drosophila melanogaster has three SET1 family members: Trithorax (Trx), trithorax-related (Trr) and dSET1 (22), whereas humans have six related SET1 family members: MLL1-4 (23-26), and Setd1A/B (27,28). Phylogenetic analysis shows that the human SET1 family reduces to three distinct clades with MLL1 and MLL4 (also known as MLL2, Wbp7, KMT2B) being most closely related to Trx, MLL2 (MLL4, ALR, KMT2D) and MLL3 (KMT2C) being most closely related to Trr, and Setd1A and Setd1B being most closely related to Drosophila dSET1 and yeast SET1p (Figure 1a,b) (29,30).

SET1 family members share the properties that they all catalyze H3K4 methylation using the evolutionarily conserved Suppressor of Variegation, Enhancer of Zeste, Trithorax (SET) domain (31), and all interact with an evolutionarily conserved sub-complex called WRAD (WD-40 repeat protein 5 (WDR5), Retinoblastoma binding protein 5 (RbBP5), Absent small homeotic-2-like (Ash2L), and Dumpy-30 (DPY-30)) (32-35). MLL1 interacts with WRAD to form the MLL1 core complex, which is required for multiple H3K4 methylation in vitro and in vivo (32,36). In vitro biochemical studies have shown that the isolated MLL1 SET domain catalyzes predominantly weak H3K4 monomethyltransferase activity (36). However, when in complex with WRAD, the rates of H3K4 mono- and dimethylation are markedly increased (36). The molecular mechanisms for how WRAD increases these activities are not well understood.

While MLL1 has served as a paradigm for the mechanism of action of human SET1 family enzymes, several recent reports suggest different family members are controlled by distinct regulatory mechanisms. For example, while the MLL1 core complex has predominantly mono- and dimethyltransferase activity in vitro (36), a similar complex assembled with MLL3 shows only monomethylation activity (37,38). Indeed, MLL2/3 complexes are linked to H3K4 monomethylation at active enhancers (39). In contrast, it has been suggested that Setd1A/B complexes catalyze the bulk of H3K4 trimethylation in cells (40,41), but is dependent on other factors that are unique to Setd1A/B complexes, such as WDR82 (SWD2 in yeast) (40,41) and CFP1 (SPP1 in yeast) (8,42). A further complication comes from the observation that histone H2B monoubiquitination is required for H3K4 trimethylation in a potential cross-talk mechanism (43-46). In yeast, this effect appears to be mediated by SPP1 and the N-SET domain of SET1p (42,45), but may also be a context dependent phenomenon (47). It is unclear if these enzymatic activities are intrinsic to SET1 family complexes or if they are the result of direct or indirect regulation of SET1 family SET domains within cells. Lacking is a rigorous biochemical comparison of all human SET1 family core complexes under standard conditions.

In this investigation, we reconstituted each human SET1 family core complex from individual subunits and compared complex assembly and enzymatic activity using well-defined standard in vitro assays. We found that in the absence of WRAD, all but one SET domain catalyzes at least weak H3K4 monomethylation. The Setd1A SET domain is inactive in the absence of WRAD, likely due to a disordered AdoMet binding pocket. We also found that all SET1 family SET domains assemble with WRAD, but that the requirement for WDR5 for complex assembly and enzymatic activity differs among family members. In the presence of WRAD, all SET1 family SET domains showed stimulated monomethyltransferase activity, but differ in their abilities to catalyze H3K4 di- and trimethylation. We found that these differences in product specificity are correlated.
with evolutionary lineage, suggesting these enzyme complexes have evolved to accomplish unique tasks within metazoan genomes. To begin to understand the structural basis for differences in product specificity, we employed a phylogenetic scanning mutagenesis assay and identified a non-active site SET domain surface that is required for WRAD-dependent H3K4 dimethylation by SET1 family core complexes. These results form the basis for our understanding of how the incorporation of H3K4 methylation in mammalian genomes may be regulated.

MATERIALS & METHODS

Materials

WDR5 antibody was obtained from Abcam (ab22512). RbBP5 and Ash2L antibodies were obtained from Bethyl (A300-109A and A300-489A respectively). An HRP-conjugated donkey anti-rabbit antibody was obtained from GE Healthcare. A polyclonal anti-GST antibody was obtained from GE Healthcare (2725701). Histone H3 peptides were synthesized by GeneScript and contained residues 1-20 followed by GGK-biotin and were either unmodified, mono-methylated, or di-methylated at H3K4. All peptides were purified to greater than 95% purity. Furthermore, all peptides were blocked by amidation of the C-terminus. MCF-7 cell extracts were obtained from Santa Cruz (sc-24793).

Protein Expression/Purification

Human SET1 family constructs consisting of residues MLL1 (3745-3969) (UniProtKB ID Q03164), MLL2 (5319-5537) (UniProtKB ID O14686), MLL3 (4689-4911) (UniProtKB ID Q8NEZ4), MLL4 (2490-2715) (UniProtKB ID Q9UMN6), Setd1A (1474-1708) (UniProtKB ID O15047), and Setd1B (1684-1924) (UniProtKB ID Q9UPS6) were sub-cloned into pGST parallel expression vectors (48) and individually expressed in Escherichia coli (Rosetta II (DE3) pLysS; Novagen) and purified as previously described (49). WRAD components were further purified and buffer exchanged by gel filtration chromatography (Superdex 200, GE) pre-equilibrated with 20mM Tris (pH 7.5), 300mM NaCl, 1mM Tris(2-carboxyethyl)phosphine and 1 μM ZnCl2(Buffer 2). SET1 family mutants were prepared bysubjecting DNA constructs to site-directed mutagenesis (QuickChange II XL, Agilent) and expressed and purified as described above.

Methyltransferase Activity Assays

Histone H3 methyltransferase assays were performed by incubating GST-tagged SET1 family members with a stoichiometric amount of WRAD (3 μM), 1 μCi of [3H]-AdoMet (Perkin Elmer Inc.), and 100 μM of histone H3 peptides that were unmodified or previously mono-, or di-methylated at H3K4. Reactions were incubated at 15°C for 6 hours. 15°C was chosen as the incubation temperature due to SET domain instability at higher temperatures. Isolated SET1 family SET domains (5 μM) were assayed by incubating with 1 μCi of [3H]-AdoMet (Perkin Elmer Inc.), and 100 μM of histone H3 peptides that were unmodified or previously mono-, or di-methylated at H3K4. These reactions were incubated at 15°C for 8 hours. All reactions were quenched with SDS-loading buffer and separated by SDS-PAGE using a 4-12% Bis-Tris gel (Life Technologies) run at 200V for 30 minutes. The gels were enhanced for 30 minutes (Enlightning, Perkin Elmer Inc.) then dried for 2.5 hours at 72°C under constant vacuum. The dried gels were exposed to film (Kodak Biomax MS Film) for 4-24 hours.

Liquid scintillation counting (LSC) was performed by excising bands corresponding to histone H3 peptides, which were dissolved in 750 μL of Solvable (Perkin Elmer Inc.), incubated at room temperature for 30 minutes followed by incubation at 50 °C for 3 hours. The solubilized volume of each sample was transferred to liquid scintillation vials containing 10mL of Ultima Gold XL liquid scintillation cocktail (Perkin Elmer Inc.). Samples were dark adapted for 1 hour then counted for 5 minutes each with a two-sigma error cut off using an all purpose scintillation counter (Beckman Coulter).
MALDI TOF mass spectrometry methyltransferase assays were performed by incubating 7 μM SET1 family SET domain with 7 μM WRAD, 250 μM AdoMet (Cayman Chemicals), and 10 μM of H3_1-20 peptide (unmodified) at 15°C for 24 hours. The reactions were quenched with 0.5% TFA and then mixed 1:5 with α-cyano-4-hydroxycinnamic acid and shot on a Bruker Autoflex III mass spectrometer (SUNY College of Environmental Sciences & Forestry; Syracuse, NY) in reflectron mode. Final shots were averaged from 200 shots per spot at 5 different positions. Relative methylation levels were quantitated using mMass (50). Reaction progress curves were global fitted to irreversible consecutive reaction models using Dynafit (51).

**3H-AdoMet Crosslinking Assays**

Purified GST-tagged SET1 family proteins (wild type and mutant) alone or assembled with WRAD (3 μM) were incubated with 1 μCi of 3H-AdoMet (Perkin Elmer Inc.) at 15°C for 3 hours then incubated on ice for 1 hour. Half the volume of each sample was quenched with SDS loading buffer and the other half was exposed to UV light (254nm) in a Stratalinker oven at a distance of ~15 cm for 30 minutes on ice. The UV exposed samples were quenched with SDS-PAGE using a 4-12% Bis-Tris gel (Invitorgen) run at 200V for 30 minutes. The gels were enhanced for 30 minutes (Enlightning, Perkin Elmer Inc.) then dried for 2.5 hours at 72°C under constant vacuum. The dried gels were exposed to film (Kodak Biomax MS Film) for 3 days. Bands corresponding to SET domains were excised and counted by LSC as described above.

**GST Pull-downs/Immunoblots**

GST tagged SET1 proteins were pre-incubated with a stoichiometric amount of purified WRAD components (3 μM) for 1 hour at 4°C before being added to pre-washed agarose beads coated with glutathione (Thermo Fisher) and incubated for an additional 2 hours at 4°C. The beads were washed 3x with Buffer 2 supplemented with 0.05% Triton X-100 and 0.05% sodium deoxycholate. The complexes were eluted from beads by boiling the samples at 95°C in SDS-loading buffer for 10 minutes. Samples of the supernatant were run on a 4-12% Bis-Tris gel (Life Technologies) and either stained with Coomassie Brilliant Blue or transferred to a PVDF membrane (Life Technologies) at 30V for 1 hour. PVDF membranes were blocked for 1 hour with a 5% non-fat milk solution then incubated with primary antibody (1:12,000) for 1 hour at room temperature. Blots were washed 4 times then incubated with an HRP-conjugated anti-rabbit secondary antibody (1:10,000) for 1 hour at room temperature. Blots were washed an additional 4 times then visualized by chemiluminescence (Clarity Western, BioRad) on a BioRad ChemiDoc MP Imager using the chemiluminescence setting.

For pull-down assays using cell extracts, 3 μM of GST-tagged SET domains was incubated with 100 μg of MCF-7 cell extracts for 16 hours at 4°C. Following the initial incubation, 20 μL of a 50:50 slurry of glutathione agarose beads was added to each sample and incubated for an additional 2 hours at 4°C. Samples were washed 3x in radioimmunoprecipitation (RIPA) buffer. Samples were eluted from beads by boiling the samples at 95°C in SDS-loading buffer for 10 minutes. Samples of the supernatant were separated on a 4-12% Bis-Tris gel (Life Technologies) and Western blotted as described above. Primary antibodies were used at a dilution of 1:2000, and the secondary antibody was used at a dilution of 1:5000.

**Phylogenetic Scanning Mutagenesis (PSM) assays**

Mutant and wild type constructs (GST-tagged) were expressed in Escherichia coli (Rosetta II (DE3) pLysS; Novagen) in 5 mL cultures by inducing with 1 mM IPTG (MLL1/MLL3) or 750 μM IPTG (Trx) and growing for 24 hours at 16°C. Cells were harvested by centrifugation and lysed by resuspending pellets in Buffer 1 supplemented with a complete protease inhibitor tablet (Roche), 1x BugBuster (Novagen), and 0.25mg/mL DNAse A at 4°C for 3 hours with rotation. Lysates were harvested by centrifugation. The expression of each mutant was assessed by running samples of the lysate on a 4-12% Bis-Tris gel (Invitorgen) and then western blotted as described above. The primary α-GST antibody (GE Healthcare) was used at a dilution of 1:10,000 (MLL3 mutants) or 1:7500 (Trx mutants), and the secondary antibody...
SET1 family complexes have different product specificities

(Jackson Immuno Research) was used at a dilution of 1:10,000.

Methyltransferase assays were performed by incubating each lysate with 3 μM WRAD, 250 μM H3 peptide (unmodified or monomethylated), and 1 μCi (Trx) or 2 μCi (MLL3) 3H-AdoMet (Perkin Elmer Inc). Samples were incubated at 15°C for 6 hours (Trx) or 15 hours (MLL3). The reactions were quenched with SDS-loading buffer and separated by SDS-PAGE using a 4-12% Bis-Tris gel (Life Technologies) run at 200V for 30 minutes. The gels were enhanced for 30 minutes (Enlightning, Perkin Elmer Inc.) then dried for 2.5 hours at 72°C under constant vacuum. The dried gels were exposed to film (Kodak Biomax MS Film) for 24 hours (Trx) or 2 days (MLL3). Bands corresponding to H3 peptides were excised and counted by LSC as described above.

RESULTS

Human SET1 family SET domains are predominantly monomethyltransferases

Previous structural studies have shown that the product specificity of SET domain enzymes depends on the presence of Phe or Tyr residue at a specific position in the SET domain active site, called the Phe/Tyr switch position (52-56). SET1 family SET domains all possess tyrosine at the switch position, which limits active site volume and predicts monomethyltransferase activity. We previously demonstrated that the isolated MLL1 SET domain is predominantly an H3K4 monomethyltransferase, and that replacement of Y3942 with Phe converts it into a processive trimethyltransferase (36). To determine if the other members of the human SET1 family catalyze a similar degree of methylation, we purified each recombinant human SET1 family SET domain as a GST-fusion protein from E. coli and compared histone methyltransferase activity and complex assembly in the presence and absence of WRAD. GST-SET domain constructs contained the SET and post-SET domains as well as the conserved WDR5 interaction (Win) motif (Figure 1b), which we and others previously showed is crucial for MLL1 core complex assembly in vitro and in vivo (49,57-59).

To determine the product specificity of the isolated human SET1 family constructs, we compared the enzymatic activity of each SET domain on histone H3 peptides that were unmodified or previously mono-, or di-methylated at H3K4. When 3H-AdoMet and unmodified histone H3 peptide (H3K4me0) were incubated with each GST-SET domain protein, only MLL1 and Setd1B showed activity in the fluorogram after a 4 hour exposure to film (Figure 1c). After a 24-hour exposure, weak activity was detected for MLL2, MLL3 and MLL4, but not the Setd1A protein (Figure 1c). With the H3K4me0 peptide substrate, MLL1 showed the greatest amount of activity, followed by Setd1B, MLL2, MLL3 and MLL4 (Figure 1d). A similar pattern of activity was observed upon liquid scintillation counting of excised peptide bands (Figure 1d). When H3K4me1 was used as a substrate, weak activity was observed with the MLL1 and Setd1B enzymes (Figure 1c). No activity was observed when H3K4me2 was the substrate. These results are consistent with previous experiments suggesting that SET1 family SET domains preferentially catalyze H3K4 monomethylation.

Human SET1 family core complexes display different product specificities

We previously showed that when MLL1 is assembled with WRAD under single turnover conditions, significant H3K4 mono- and dimethylation activities are observed with trace amounts of trimethylation after 24 hours (36). To determine if all human SET1 family core complexes show a similar product specificity, we compared complex formation and enzymatic activity among all six human SET1 family core complexes. GST-pull-down experiments were used to compare the ability of each GST-SET domain to pull-down recombinant WRAD subunits in vitro. As shown in Figure 2a, all GST-SET1 family fusion proteins enriched all WRAD components (Lanes 1-6), compared to the GST-only control (Lane 7). Similarly, each GST-SET1 family fusion protein enriched endogenous WRAD subunits after incubation with MCF-7 cell extracts (Figure 2b). These results demonstrate that all SET1 family SET domains form complexes with WRAD in vitro.

We next compared enzymatic activity of each complex using H3 peptides that were either
SET1 family complexes have different product specificities

unmodified or previously mono- or dimethylated at Lys4. As shown in previous studies (36), the MLL1 core complex showed significant activity with H3K4me0 and H3K4me1 peptides, with trace amounts of activity with the H3K4me2 peptide (Figure 2c, lane 1 and Figure 2d). A similar activity profile was observed when GST-MLL4 was mixed with WRAD (Figure 2c, lane 2). These results suggest that MLL1 and MLL4 core complexes are predominantly H3K4 mono- and dimethyltransferases. In contrast, the complexes assembled with the other human SET1 family SET domains showed striking differences. For example, while MLL2 and MLL3 complexes showed WRAD stimulated activity with H3K4me0 peptides, they showed little or no activity with H3K4me1 and H3K4me2 peptides (Figure 2c, lanes 3 and 4). These results suggest that MLL2 and MLL3 core complexes are predominantly monomethyltransferases. In contrast, while we were not able to detect activity with the Setd1A without WRAD (Figure 1c), with WRAD, it displayed activity with all three peptides (Figure 2c, lane 5). A similar pattern was observed with the Setd1B enzyme (Figure 2c, lane 6). These results suggest that Setd1A/B complexes catalyze mono-, di-, and trimethylation of H3K4.

To confirm the observed product specificities using a different assay, we incubated each enzyme complex with the H3K4me0 peptide under single turnover conditions and monitored methylation over time by quantitative MALDI TOF mass spectrometry. As shown in Figure 3a, (top panel), the MLL1 core complex converts most of the H3 peptide into the me2 form, with a small amount converted into the me3 form at 24 hours. The complex assembled with MLL4 showed similar product specificity, but with a slower overall rate (Figure 3a,b,c). MLL2 and MLL3 complexes, in contrast, showed predominantly monomethyltransferase activity after 24 hours, whereas Setd1A/B complexes showed di- and trimethyltransferase activities (Figure 3a,b). For enzyme complexes catalyzing multiple methylation (MLL1, MLL4, Setd1A/B), fitting of reaction progress curves to irreversible consecutive reactions models showed significant accumulation of the H3K4me1 and H3K4me2 intermediates (Figure 3b), with rate constants for monomethylation (k1) that were ~4-8-fold greater than that for dimethylation (k2) (Figure 3b, c). In addition, the rates constants for dimethylation by the Setd1A/B complexes were ~2-6-fold greater than that for trimethylation (k3) (Figure 3b, c). These results are consistent with non-processive methylation.

Taken together, these results are in agreement with that of the radiometric assays, and suggest that different SET1 family core complexes catalyze different degrees of H3K4 methylation.

WRAD increases AdoMet binding in several SET1 family members

To begin to understand why WRAD increases the H3K4 monomethylation activity of all SET1 family complexes, we performed UV-dependent AdoMet cross-linking studies in the presence and absence of WRAD. A complicating feature of these studies is that we previously demonstrated that the isolated MLL1 SET domain undergoes a robust intramolecular automethylation activity in the absence of UV (60). Indeed, in this investigation, we observed that the isolated MLL4 and MLL2 SET domains also undergo weaker automethylation reactions compared to that of MLL1 (Figure 4a). In addition, when MLL1 is assembled with WRAD, we previously found that MLL1 methylates Ash2L in an intramolecular (intra-complex) manner (60). In this investigation, we observed a similar Ash2L methylation reaction in complexes assembled with MLL1, MLL4, MLL2, Setd1A and Setd1B proteins in the absence of UV light (Figure 4b, lower panels, lanes 2,6,10,18, & 22, respectively). No Ash2L methylation was detected in the complex assembled with the MLL3 SET domain (Figure 4b, lower panel, lane 14). Despite these activities, we found that UV-dependent cross-linking could be used to interrogate AdoMet binding in the presence and absence of WRAD. For example, all isolated GST-SET proteins except Setd1A showed an increase in radioactivity in the presence of UV-light compared to that observed in the absence of UV light (Figure 4b, upper panels). AdoMet cross-linking was abolished in each SET domain variant in which the AdoMet binding asparagine (equivalent to N3906 in MLL1) was replaced with alanine (Figure 4b, MLL1 N3906A, MLL4 N2652A, MLL2 N5474A, MLL3 N4848A, Setd1A N1646A, Setd1B N1862A). These results suggest that all SET domain proteins can bind AdoMet, with the exception of the Setd1A protein.
SET1 family complexes have different product specificities

When similar assays were performed in the presence of WRAD, all SET domain proteins showed UV-dependent AdoMet crosslinking (Figure 4b, lower panels). The Setd1A protein, which did not show evidence of crosslinking in the absence of WRAD, showed robust UV-dependent cross-linking in the presence of WRAD (lane 17). These results suggest that WRAD rescues the AdoMet binding defect on the GST-Setd1A protein. Indeed, WRAD significantly increased SET domain AdoMet crosslinking in MLL4, MLL2, Setd1A and Setd1B complexes when excised bands were quantitated by liquid scintillation counting (Figure 4c). In contrast, Ash2L methylation was unaffected in the presence or absence of UV light, indicating that it is the result of an enzymatic reaction that is catalyzed by the SET domain subunit, as previously suggested for MLL1 (60). The only complex that did not display Ash2L methylation was the MLL3 core complex (Figure 4b, lane 15).

These results suggest that WRAD induces a conformational change in SET1 family SET domains that results in increased binding of AdoMet for the majority of SET1 family members.

Functions of WRAD subunits in human SET1 family core complexes

From the results presented above, WRAD stimulates H3K4 monomethylation in all SET1 family core complexes and is required for the di- and trimethylation activity of a subset of complexes. To further dissect the roles of WRAD subunits, we compared enzymatic activities in SET1 family complexes in which each WRAD subunit was systematically deleted. When the WDR5 subunit was omitted, we noticed that all SET1 family core complexes retained significant activity with the K3K4me0 substrate (Figure 5a, lanes 8–13 and Figure 5b). Interestingly, in the absence of WDR5, the MLL3 core complex displayed a significant stimulation in enzymatic activity with the H3K4me0 substrate (compare lanes 4 and 11). These results suggest that WDR5 partially inhibits the H3K4 monomethyltransferase activity of the MLL3 core complex. To characterize WDR5 inhibition further, we titrated WDR5 into the MLL3 core complex and found that inhibition occurs only when the concentration of WDR5 approaches the concentration of the other subunits (Figure 5c). Therefore, a near- or super-stoichiometric concentration of WDR5 partially inhibits the enzymatic activity of the MLL3 core complex.

When the H3K4me1 or H3K4me2 peptides were used as substrates, significant difference among SET1 family members were observed upon omission of WDR5. H3K4 dimethylation activity is significantly reduced with MLL1 (Figure 5a, compare lanes 1 and 8 and Figure 5b) and Setd1A complexes (Figure 5a, lanes 5 and 12). In contrast, little or no differences were observed with MLL4, MLL2 and Setd1B complexes (Figure 5a and 5b). Despite stimulation of the monomethylation activity of the MLL3 complex upon omission of WDR5, little activity was observed when H3K4me1 or H3K4me2 peptides were used as substrates (Figure 5a, lane 11). Taken together, these results suggest that MLL1 and Setd1A complexes depend on WDR5 for their complete di- and trimethylation activities. WDR5 appears to be somewhat dispensable for the enzymatic activity of the MLL2, MLL3, MLL4 and Setd1B complexes.

In contrast to that of WDR5, omission of RbBP5 or Ash2L subunits completely abolished WRAD stimulated effects on all human SET1 family core complexes (Figures 5d and e). The activity of each complex resembles that of the isolated SET domain in the absence of WRAD on all three substrates (Figure 1c). These results suggest that RbBP5 and Ash2L are critical for the enzymatic activities of human SET1 family core complexes. Omission of DPY-30, conversely, did not significantly affect the activities of SET1 family core complexes on all three substrates (Figure 5f).

Phylogenetic Scanning Mutagenesis reveals SET domain surface involved in WRAD dependent product specificity regulation

To begin to identify amino acids responsible for differences in product specificity among SET1 family complexes, we developed a high-throughput “Phylogenetic-scanning mutagenesis” assay, where systematic amino acid substitutions are made between SET domain paralogs or orthologs and assayed for gain-of-function activities. The assay works by expressing wild type or variant GST-SET domain proteins in 5 mL cultures of E. coli, followed by western
SET1 family complexes have different product specificities

blotting with anti-GST antibodies to compare and normalize expression levels (Figure 6a). The extracts are then incubated with \(^{\text{3}}\text{H}-\text{AdoMet} \) and the H3K4me0/1 peptide substrates and assayed directly for H3K4 dimethyltransferase activity in the presence and absence of purified WRAD. Gain-of-function substitutions are defined as those variants showing significantly increased H3K4 dimethylation activity (p \(\leq 0.01 \)) compared to that of the wild type enzyme.

We used phylogenetic scanning mutagenesis to identify residues that confer a gain-of-function dimethyltransferase activity on the MLL3 core complex, which is predominantly an H3K4 monomethyltransferase (Figure 2c). We performed a multiple sequence alignment with MLL1 and MLL3 orthologs and identified at least 17 positions that were conserved or semi-conserved within each phylogenetic clade, but different between clades (Figure 6b). We then systematically replaced each position in MLL3 with the corresponding amino acid in MLL1 and assayed for enzymatic activity using the H3K4me0 and H3K4me1 peptides as substrates. The resulting heat map shows that when H3K4me0 was the substrate there was very little variation in activity among MLL3 variants with the exception of Y4774G, which was significantly reduced compared to that of the complex assembled with wild type MLL3 (Figure 6c). Indeed western blotting revealed that the Y4774G GST-fusion protein did not express in \( \text{E. coli} \), explaining why no activity was observed (Figure 6d, right panel, lane 8). In contrast, when H3K4me1 was used as the substrate, four of the 17 variants (R4779P, Y4786F, Q4877Y, and H4900N) showed a modest WRAD-dependent increase in dimethylation activity compared to that of wild type MLL3 (Figure 6c and e). These results cannot be explained by differences in expression levels as western blotting revealed that each gain of function variant expressed at a level similar to that of wild type MLL3 in \( \text{E. coli} \) (Figure 6c and e). These results cannot be explained by differences in expression levels as western blotting revealed that each gain of function variant expressed at a level similar to that of wild type MLL3 in \( \text{E. coli} \) (Figure 6c). In addition, there was little variation in activity when the H3K4me0 peptide was used as substrate, suggesting that similar amounts of folded protein were assayed (Figure 6c).

To determine if differences in AdoMet binding account for the gain-of-function activities, we purified the gain-of-function variants and compared AdoMet UV-crosslinking in the presence of WRAD with that of the wild type MLL3 core complex. The results showed little changes in AdoMet crosslinking (0.9-2-fold) suggesting that differences in AdoMet binding alone do not account for increased dimethylation activity (Figure 6f).

We then mapped the gain-of-function mutations on the three-dimensional structure of the MLL1 SET domain, and noticed that all gain-of-function amino acid positions cluster around a solvent exposed concave surface that is distinct from the SET domain active site cleft (Figure 7). Modeling reveals that each gain-of-function position tolerates multiple rotomers of each mutant side chain without steric clash, making it likely that the mutations do not significantly alter the overall structure of the SET domain. These results suggest that this surface is involved in WRAD-dependent dimethylation activity of the MLL1 core complex.

We next asked if phylogenetic scanning mutagenesis could be used to determine if variation at the same amino acid positions can explain why the \( \text{Drosophila} \) Trx ortholog of MLL1 does not catalyze H3K4 dimethylation when mixed with human WRAD. We previously found that like the human MLL1 SET domain, the Trx SET domain forms a complex with human WRAD and displays stimulated monomethyltransferase activity (61). However, unlike that of MLL1, Trx does not catalyze appreciable H3K4 dimethylation when assembled with human WRAD (61). Interestingly, several of the gain-of-function amino acid positions identified in the MLL3 screen showed an evolutionary split between vertebrate and invertebrate MLL1 orthologs (which includes Trx) (Figure 8a). We replaced the four gain-of-function and one non-gain-of-function amino acid positions in the Trx SET domain with the corresponding amino acids in MLL1 and assayed for H3K4 dimethyltransferase activity in the presence and absence of WRAD. The results revealed that three of the four amino acid substitutions showed a WRAD-dependent gain-of-function dimethylation activity in the Trx SET domain (Figure 8b-d). These results suggest that this surface is involved in the WRAD-dependent product specificity differences among SET1 family core complexes.

DISCUSSION
In this investigation, we have reconstituted and characterized the biochemical properties of all six human SET1 family core complexes under standard conditions in vitro. We found that all but one SET1 family SET domains catalyze predominantly H3K4 monomethylation, which is stimulated by interaction with WRAD. In addition, we found that the SET domains interact with WRAD with similar stoichiometries in vitro, but that the requirement for individual WRAD subunits for enzymatic activity differed. While all SET1 family complexes require RbBP5 and Ash2L for full activity, the MLL3 core complex does not require WDR5 for enzymatic activity, and appears to be inhibited by WDR5 when in stoichiometric excess. We found that WDR5 is required for the assembly of MLL1 and Setd1A core complexes, but not for the assembly of MLL2-4 and Setd1B. Interestingly, these differences are inversely correlated with the differences in affinity observed between WDR5 and Win motif peptides derived from each SET1 family member. Paradoxically, MLL1 and Setd1A Win motif peptide sequences bind to WDR5 with the weakest affinity (0.5-2 μM compared to 30-100 nM for MLL2-4,Set1d1b) (57), yet require WDR5 for complex formation and multiple H3K4 methylation. Understanding the structural and functional basis for these differences will be important for translational efforts to target individual SET1 family members in diseased cells. Indeed, it has been shown that the MLL1-WDR5 interaction can be targeted for inhibition by peptides and small molecules in vitro and in mammalian cells (38,49,57). Even though WDR5 is not required for the enzymatic activity of all SET1 family complexes, it has been retained in each complex for some additional reason, possibly for gene targeting. For example, it has been shown that WDR5 binds to a long non-coding RNA called HOTTIP and is involved in recruitment of MLL family complexes to specific genomic loci (62,63).

The major differences we observed among SET1 family complexes is that product specificities varied in a manner correlated with evolutionary lineage, which is surprising given their high degree of amino acid conservation. Since WRAD subunits were identical, these results indicate that SET domain amino acid sequence variation accounts for the observed product specificities. However, SET domain active site residues are strictly conserved and are predicted to adopt similar positions in three-dimensional models (Figure 7b), implying that product specificity differences are due to amino acid sequence variation outside of the active site cleft. Indeed, we recently described the characterization of a non-active site SET domain surface that is mutated in MLL2 in human Kabuki Syndrome and in non-Hodgkin’s lymphomas (64). We found that this surface, called the Kabuki Interaction surface (KIS), is required for the interaction of MLL1 with the RbBP5-Ash2L heterodimer and for H3K4 dimethylation by the MLL1 core complex (64). However, these residues are strictly conserved among SET1 family members, and cannot therefore be responsible for the product specificity differences observed in this investigation.

We therefore developed a high-throughput “phylogenetic scanning mutagenesis” assay to systematically switch amino acid residues between orthologs and paralogs, and found a group of four amino acid positions that cluster on a common surface and appear to be involved in the H3K4 dimethylation reaction of the MLL1 core complexes in a WRAD-dependent manner. Interestingly, these amino positions are highly conserved among vertebrate MLL1 orthologs, but are different among invertebrate orthologs. These residues map to a solvent exposed surface that is opposite the KIS cluster, and distinct from the SET domain active site cleft. We propose that amino acid differences in this SET domain surface results in subtly different interactions with WRAD and “fine-tunes” the number of methyl groups that can be transferred to the histone substrate. The molecular mechanisms for how this works are unknown. One possibility is that different interactions between each SET domain and WRAD causes different allosteric alterations of SET domain active site residues, resulting in different product specificities. This mechanism is consistent with the “one-active site” model for multiple methylation, where mono-, di- and trimethylation all occur within the SET domain active site. However, it is also possible that the amino acid sequence variation alters a surface that is involved in formation of a second active site, one that is created at the interface between WRAD and a non-active site SET domain surface (36). This so called “two-active site” model is supported
by the demonstration that WRAD possesses an intrinsic H3K4 methyltransferase activity that is independent of the enzymatic activity of the MLL1 SET domain (36,65). Further experiments will be required to distinguish between these mechanisms.

In summary, we have established that different SET1 family complexes catalyze varying degrees of H3K4 methylation when compared under identical conditions in vitro. These differences appear to be due in part to amino acid variation outside of the SET domain active site cleft. We suggest that variation in this SET domain surface regulates product specificity either through allosteric changes in SET domain active site residues, or through modulation of the activity of a second active site. This information will facilitate future experiments that will distinguish these possibilities. In addition, this work forms the basis for understanding how additional regulatory inputs alter the enzymatic activities of human SET1 family core complexes in cells.

ACKNOWLEDGEMENTS:
We would like to thank Dr. David Skalnik for the Setd1A and Setd1B plasmids and Dr. Peter Harte for the Trx plasmid. We are also grateful to Vicky Lyle at the DNA sequencing core facility at SUNY Upstate Medical University. We also thank Dr. Dave Amberg and Dr. Brian Haarer for the anti-GST and HRP-conjugated anti-goat antibody. In addition, we thank Dr. Nilda Alicea-Velazquez and Kevin Namitz for helpful comments on this manuscript. Lastly, we would also like to thank Dr. Nilda Alicea-Velazquez for technical assistance with the UV cross-linking experiments. This work is supported in part by National Institutes of Health grant 1R01CA140522 (to M.S.C.).

REFERENCES:


SET1 family complexes have different product specificities
SET1 family complexes have different product specificities

are tethered together selectively by the cell-proliferation factor HCF-1. **Genes Dev** **17**, 896-911


SET1 family complexes have different product specificities


SET1 family complexes have different product specificities

Modeller. *Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis ... (et al.) Chapter 5*, Unit 5 6
FIGURE LEGENDS

Figure 1. Human SET1 family members predominantly catalyze monomethylation of H3K4. (a) Phylogenetic cluster analysis (Clustal Omega (66)) of SET1 family members using full-length protein sequences from S. cerevisiae (ySet1), Drosophila (dSet1, Trx, Trr), and humans (MLL1-4, Setd1A/B). (b) Schematic representation of full-length human SET1 family proteins. The catalytic SET domain is shown in light pink, the post-SET region is shown in green, and the WDR5 interaction (Win) motif is shown in blue. Dotted lines represent the N-terminus of the recombinant constructs used in this study, beginning with the residues noted above each construct. (c) Comparison of histone methyltransferase activity among human SET1 family SET domains. Upper panels show Coomassie blue stained SDS-PAGE gels, and the lower panels show 3H-methyl incorporation by fluorography after a four-hour exposure (middle panels), and after a 24-hour exposure (lowest panels). The control lane shows the activity of the MLL1 SET domain on 100 µM unmodified H3 peptide, which is included on each gel. The control lanes are from the same gel at the same exposure but were cropped for clarity. (d) Quantification of radioactivity from excised histone H3 bands by LSC. Data are normalized to the activity level of the control lane on each gel. Error bars represent the standard error of measurement between three independent experiments.

Figure 2. SET1 family core complexes catalyze different levels of H3K4 methylation. (a) Comparison of core complex assembly with each human SET1 family member by in vitro GST pull-down assays from purified components. Individual GST-tagged SET domains were incubated with purified WRAD components and glutathione coated agarose beads. The upper panel shows a Coomassie blue stained gel, and the lower panels show the Western blot. Purified GST is used as a negative control (lane 7). Purified individual WRAD subunits were run on the gel (lanes 8-11) to compare to the migration of WRAD components from the pull-down lanes (lanes 1-8). (b) Comparison of core complex assembly with each human SET1 family member by in vitro pull-down experiments from MCF-7 breast cancer cell extracts. Individual GST-tagged SET domains were incubated with cell extracts and pulled down with glutathione agarose beads. WRA components were detected by Western blotting. The upper panel shows a Ponceau S stained PVDF membrane, and the lower panels show the Western blot. (c) Comparison of core complex methyltransferase activities among SET1 family members in complex with WRAD. The upper panels show Coomassie blue stained SDS-PAGE gels, and the lower panels show 3H-methyl incorporation after 4 hours as shown by fluorography. The control lane shows the activity of the MLL1 SET domain on 100 µM unmodified H3 peptide, which is included on each gel. (d) Quantification of radioactivity from excised histone H3 bands by LSC. Data are normalized to the activity level of the control lane on each gel. Error bars represent the standard error of measurement between five independent experiments.

Figure 3. SET1 family core complex single turnover kinetics. (a) MALDI TOF mass spectrometry showing histone methylation of an unmodified H3 peptide after 24 hours for each human SET1 family core complex. (b) Reaction progress curves global fitted to irreversible consecutive reaction models using DynaFit. Each time point represents the mean percentage of total integrated area for each species in MALDI TOF reactions. Error bars represent ± standard deviation from duplicate measurements. (c) Summary of rate constants derived from global fitting of reaction progress curves as described in Methods.

Figure 4. AdoMet crosslinking studies of the human SET1 family members. (a) A subset of SET1 family members undergo auto-methylation. Isolated SET domains were incubated with 3H-AdoMet in the presence (+) or absence (-) of 250 µM unmodified H3 peptide. The upper panel shows a Coomassie blue stained SDS-PAGE gel and the lower panel shows 3H-methyl incorporation by fluorography after a 24 hour exposure to film. (b) UV exposure was used to cross-link 3H-AdoMet to isolated SET1 family SET domains (upper panels) or SET1 family core complexes (lower panels). All panels represent 3H-methyl incorporation by fluorography after a three-day exposure to film. A single asparagine to alanine mutation in the AdoMet binding pocket of each SET domain abolishes 3H-AdoMet cross-linking. (c) Quantification
of \(^3\)H-AdoMet cross-linking in the presence and absence of WRAD. Bands corresponding to the SET domains of each wild-type SET1 family member were excised and quantified by liquid scintillation counting as described in Methods. Error bars represent the standard error of measurement from three independent experiments.

**Figure 5. Functions of WRAD components in the SET1 family of core complexes.** (a) Comparison of H3K4 methylation activity by SET1 family core complexes assembled with and without WDR5. The upper panels show Coomassie blue stained SDS-PAGE gels, and the lower panels show \(^3\)H-methyl incorporation after a 4 hour exposure to film. (b) Quantification of methyltransferase activity among SET1 family core complexes assembled with and without WDR5 by liquid scintillation counting. Error bars represent the standard error of measurement from three-five independent experiments. (c) Titration of WDR5 into the MLL3 RAD complex. A 3 \(\mu\)M MLL3-RAD complex was assembled with increasing amounts of WRD5 and tested for methyltransferase activity when an unmodified H3 peptide was the substrate. WDR5 was titrated in 0.5 \(\mu\)M increments in lanes 1-9 and in 1 \(\mu\)M increments for lanes 10-11 (range is from 0-6 \(\mu\)M). (d-f) Comparison of H3K4 methylation activity by SET1 family core complexes assembled with and without RbBP5 (d), Ash2L (e), and DPY-30 (f). The upper panels show Coomassie blue stained SDS-PAGE gels, and the lower panels show \(^3\)H-methyl incorporation after 4 hours as shown by fluorography. All gels contain the activity of the isolated MLL1 SET domain on 100 \(\mu\)M unmodified H3 peptide as a control.

**Figure 6. Phylogenetic scanning mutagenesis reveals a cluster of mutations that enhance the dimethylation activity of the MLL3 core complex.** (a) Schematic of the phylogenetic scanning mutagenesis assay. Mutant constructs were expressed in 5 mL cultures and lysates were analyzed for protein expression. Purified WRAD components along with \(^3\)H-AdoMet and H3 peptides (unmodified or previously mono-methylated at H3K4) were incubated with the lysates. Fluorography and liquid scintillation counting were used to analyze activity. (b) Representative sequence alignment of SET1 family SET domains (Clustal Omega). The MLL1 homologs are highlighted in blue and the MLL3 homologs are highlighted in green. The two boxed positions represent gain-of-function hits from the screen. (c) Heat map of methyltransferase activities of wild type and mutant MLL3 constructs. The data represent the Log2-fold change (mutant/wild type) in activity of MLL3 constructs with the indicated peptide. * represents significant increase in dimethylation activity (p< 0.01). (d) Assessment of the expression level of mutant MLL3 constructs. Lysate samples from mutant and wild type constructs were separated by SDS-PAGE, transferred to PVDF membranes and blotted with an \(\alpha\)-GST antibody. The upper panels show Ponceau S stained PVDF membranes and the lower panels show the Western blots. The control represents an untransformed E. coli lysate that was induced with 1mM IPTG. (e) Representative gel of gain-of-function MLL3 mutants. The upper panel depicts a Coomassie blue stained SDS-PAGE gel, and the lower panel shows \(^3\)H-methyl incorporation after 2 days as shown by fluorography. (f) UV-exposure was used to cross-link \(^3\)H-AdoMet to wild type or MLL3 variant core complexes. \(^3\)H-AdoMet crosslinking was quantitated by LSC of excised SET domain bands. Error bars represent standard error of measurement from three independent experiments.

**Figure 7.** (a) A surface representation of the crystal structure of the MLL1 SET domain bound to histone H3 peptide (yellow) and S-adenosyl-L-homocysteine (SAH) (green) (PDB:2W5Z) (67). The position of the 4 gain-of-function mutants are highlighted in blue and noted with their MLL3 numbering. The location of the previously identified Kabuki interaction surface (KIS) is noted. (b) Homology modeling of the conserved active-site residues in the human SET1 family predicts that they adopt similar three-dimensional positions. Models were generated in Modeler (68) using the MLL1 structure (PDB:2W5Z) as a template.

**Figure 8.** Gain-of-function positions identified in MLL3 enhance WRAD-dependent dimethylation by the complex assembled with Drosophila Trx. (a) Representative alignment of MLL1 orthologs
reveals amino acid positions that show a split between invertebrates (boxed) and vertebrates. (b) Heat map of methyltransferase activities of wild type and mutant Trx constructs. The data represent the Log2-fold change (mutant/wild type) in activity of Trx constructs with the indicated peptide. (c) Representative gel of gain-of-function Trx mutants. The upper panel shows a Coomassie blue stained SDS-PAGE gel, and the lower panel shows 3H-methyl incorporation after 24 hours of fluorography. (d) Assessment of the expression level of mutant Trx constructs. Lysate samples from mutant and wild type constructs were separated by SDS-PAGE, transferred to PVDF membranes and blotted with an α-GST antibody. The upper panel shows a Ponceau S stained PVDF membrane and the lower panel shows the Western blot. The control represents an untransformed E. coli lysate that was induced with 0.75 mM IPTG.
Figure 2

(a) Western blot analysis showing expression levels of various proteins. 

(b) Close-up view of selected bands from the Western blot.

(c) Another set of Western blot images highlighting different proteins.

(d) Quantitative analysis of protein expression levels using a bar graph.
Figure 3

(a) Intensity vs. m/z for MLL1, MLL2, MLL3, and Setd1A and Setd1B. The peaks at m/z values indicate the intensity of different methylated forms.

(b) Methylation (%) vs. Time (min.) for MLL1, MLL2, MLL3, and Setd1A and Setd1B. The graphs illustrate the time course of methylation changes.

(c) Table of Rate Constants (hr$^{-1}$) for MLL complexes and Setd1A and Setd1B WRAD.

<table>
<thead>
<tr>
<th>Complex</th>
<th>Rate Constants (hr$^{-1}$)</th>
<th>$k_2$</th>
<th>$k_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLL1 WRAD</td>
<td>1.50 ± 0.10</td>
<td>0.36 ± 0.04</td>
<td>N/A $^a$</td>
</tr>
<tr>
<td>MLL4 WRAD</td>
<td>0.15 ± 0.009</td>
<td>0.04 ± 0.005</td>
<td>N/A</td>
</tr>
<tr>
<td>MLL2 WRAD</td>
<td>0.29 ± 0.002</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>MLL3 WRAD</td>
<td>0.08 ± 0.004</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Setd1A WRAD</td>
<td>0.67 ± 0.05</td>
<td>0.08 ± 0.01</td>
<td>0.04 ± 0.01</td>
</tr>
<tr>
<td>Setd1B WRAD</td>
<td>0.85 ± 0.08</td>
<td>0.19 ± 0.03</td>
<td>0.03 ± 0.01</td>
</tr>
</tbody>
</table>

$^a$ Rate constants derived from global fitting of the data as described in methods ± standard error of the fit.

$^b$ N/A = non applicable, activities were either not present or too low to obtain reliable estimates of rate constants from global fitting.
Figure 8

<table>
<thead>
<tr>
<th>Species</th>
<th>Amino Acid Sequence</th>
<th>KPCXKK</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>H. sapiens</td>
<td>VGVYRSPIHKRCGLCPRND</td>
<td>KPCXKK</td>
<td>3963</td>
</tr>
<tr>
<td>M. musculus</td>
<td>VGVYRSPIHKRCGLCPRND</td>
<td>KPCXKK</td>
<td>3960</td>
</tr>
<tr>
<td>M. domestica</td>
<td>VGVYRSPIHKRCGLCPRND</td>
<td>KPCXKK</td>
<td>4006</td>
</tr>
<tr>
<td>D. rerio</td>
<td>VGVYRSAPIHKRCGLCPRND</td>
<td>KPCXKK</td>
<td>4213</td>
</tr>
<tr>
<td>G. gallus</td>
<td>VGVYRSPIHKRCGLCPRND</td>
<td>KPCXKK</td>
<td>4092</td>
</tr>
<tr>
<td>A. mellifera</td>
<td>VGVYRSPIHKRCGLDIER</td>
<td>KPCXKK</td>
<td>3486</td>
</tr>
<tr>
<td>C. floridanus</td>
<td>VGVYRSPIHKRCGLDIER</td>
<td>KPCXKK</td>
<td>3212</td>
</tr>
<tr>
<td>T. castaneum</td>
<td>VGVYRSPIHKRCGLDIER</td>
<td>KPCXKK</td>
<td>2337</td>
</tr>
<tr>
<td>A. gambiae</td>
<td>VGVYRSPIHKRCGLDIER</td>
<td>KPCXKK</td>
<td>4919</td>
</tr>
<tr>
<td>D. melanogaster</td>
<td>VGVYRSAPIHKRCGLCTDKR</td>
<td>KPCXKK</td>
<td>3720</td>
</tr>
</tbody>
</table>

**Figure 8a**

**Figure 8b**

**Figure 8c**

**Figure 8d**
Biochemical Reconstitution and Phylogenetic Comparison of Human SET1 Family
Core Complexes Involved in Histone Methylation
Stephen A. Shinsky, Kelsey E. Monteith, Susan Viggiano and Michael S. Cosgrove

J. Biol. Chem. published online January 5, 2015

Access the most updated version of this article at doi: 10.1074/jbc.M114.627646

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts