Published in final edited form as:

Adv Enzyme Regul. 2010 ; 50(1): 104–110. doi:10.1016/j.advenzreg.2009.12.005.

Structural basis for H3K4 trimethylation by yeast Set1/ COMPASS

Yoh-Hei Takahashi and Ali Shilatifard

Stowers Institute for Medical Research, 1000 East 50th Street, Kansas City, MO 64110

Abstract

Histone methylation on lysine 4 of histone H3 (H3K4) is a hallmark of activity of the transcribed regions on eukaryotic chromatin. H3K4 can be mono-, di- and trimethylated by Set1/COMPASS. In this review, we will discuss recent findings regarding the role of the Y/F switch by the catalytic domain of Set1 in the regulation of H3K4 methylation by Set1/COMPASS.

Introduction

The basic genetic information that orchestrates development and differentiation of living organisms is encoded within the DNA. How such information is interpreted during the proper stage of development depends mainly on the interaction of DNA with proteins called histones. Histones package DNA into nucleosomes, which is major part of the packaging and organization of DNA within the nucleus (Kornberg, 1974, 2007; Luger et al., 1997). The N-terminal tails of histones protrude outward from the nucleosomes and are subject to various posttranslational modifications regulating the accessibility of DNA to factors involved in many DNA-mediated phenomena such as DNA replication, repair/recombination and transcription (Luger et al., 1997). To date, acetylation, phosphorylation, methylation, ubiquitination, sumoylation, ADP-ribosylation, deimination and proline isomerization are characterized as histone posttranslational modifications (Nowak and Corces, 2004; Shilatifard, 2006; Zhang, 2003). Whereas histone lysine acetylation and serine/threonine phosphorylation generally appear to regulate gene expression in a positive way (Berger, 2002, 2007), the impact of histone lysine methylation on transcription often is dependent on the site of the modification and its biological context (Shilatifard, 2006; Zhang, 2003). Furthermore, the extent of the methylation of lysine (mono-, di-, and tri-) adds further to the complexity of the relationships between the lysine methylations and those biological outputs (Shilatifard, 2008).

In particular, histone H3 lysine 4 (H3K4) methylation is a prominent example of the complex process involved in transcriptional regulation (Shilatifard, 2006). The first H3K4 methylase was identified in yeast and contains SET domain protein Set1 as its catalytic subunit (Miller et al., 2001). Set1 in yeast is associated within a macromolecular complex named COMPASS (complex proteins associated with Set1) and is capable of mono-, di- and

© 2009 Elsevier Ltd. All rights reserved.

*To whom correspondence and proofs should be sent: Ali Shilatifard, Stowers Institute for Medical Research, 1000 East 50th Street, Kansas City, MO 64110, Office: (816) 926-4465, ASH@Stowers.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

trimethylating H3K4 (Krogan et al., 2002; Miller et al., 2001; Roguev et al., 2001; Shilatifard, 2006).

The subunits COMPASS, which are named Cps60 through Cps15 according to their apparent molecular weight, are all required for proper H3K4 methylation by the complex (Krogan et al., 2002; Schneider et al., 2005; Shilatifard, 2006). For example, several subunits such as Set1, Cps50 and Cps30 are necessary for the mono-, di- and trimethylation of histone H3K4, whereas Cps40 and Cps25 are necessary just for the trimethylation of histone H3K4 (Schneider et al., 2005). Following the purification of COMPASS from yeast, it was demonstrated that the mammalian MLLs also exist in COMPASS-like complexes capable of methylating the lysine 4 of histone H3 (Hughes et al., 2004). We now know that there are at least seven COMPASS and COMPASS-like complexes in mammalian cells all capable of methylating H3K4 (Shilatifard, 2008). Indeed many of the subunits shared from yeast to human, also function similarly in both organisms (Shilatifard, 2008).

Our Global Proteomic Analysis in Yeast (GPS) studies in yeast also demonstrated that histone H2B monoubiquitination at lysine 123 residue (K123) is required for proper methylation by COMPASS (Dover et al., 2002). We have shown that histone H2B monoubiquitination is detected at both the promoters and coding regions of the gene. The enzymatic machinery implementing H2B monoubiquitination was originally identified in yeast (Dover et al., 2002; Wood et al., 2003). Yeast Rad6, is an E2 ubiquitin-conjugating enzyme (Robzyk et al., 2000), and yeast Bre1, its E3 ubiquitin ligase, directly interacts with Rad6 to form a stable complex and monoubiquitinate H2BK123 (Wood et al., 2003). Similar to COMPASS, Rad6 and Bre1 are found in both yeast and mammals (Kim et al., 2009; Kim et al., 2005; Pavri et al., 2006; Zhu et al., 2005). Following the studies in yeast demonstrating a role for the Rad6/Bre1 complex in the regulation of COMPASS's function, it was demonstrated that human hRad6A and hRad6B and hBre1A and hBre1B can monoubiquitinate histone H2B and function in H3K4 methylation in vivo (Shilatifard, 2008). In this manuscript, we will review our structural and molecular understanding of factors and machineries involved in the regulation of COMPASS's H3K4 trimethylase activity.

Components of COMPASS involved in the regulation of H2B monoubiquitination H3K4 methylation crosstalk

Since composition and catalytic activity of COMPASS is conserved from yeast to human, yeast has provided an outstanding model for a molecular understanding of H3K4 methylation by COMPASS and COMPASS-like complexes in mammals. Our initial studies in yeast demonstrated that COMPASS purified from strains lacking H2B monoubiquitination to have several fold less of the Cps35 subunit associated within COMPASS (Lee et al., 2007). Furthermore, COMPASS purified from an H2B monoubiquitination deficient background lacks the enzymatic ability to trimethylate histone H3 (Lee et al., 2007). These observations pointed towards Cps35 as the subunits of COMPASS involved in regulating the crosstalk between H2B monoubiquitination and H3K4 trimethylation. In support of this observation, Cps35 interacts with chromatin in a monoubiquitination-dependent manner (Lee et al., 2007; Takahashi et al., 2009). Furthermore, the Cps35 homologue in humans, the Wdr82 protein, is a component of the human Set1 complex (hCOMPASS) (Lee and Skalnik, 2008; Wu et al., 2008) and requires H2B monoubiquitination for its interaction with chromatin (Wu et al., 2008).

Molecular mechanism installed within COMPASS for the regulation of H3K4 di- and trimethylation

As discussed above, the highly intricate network is found to regulate H3K4 methylation outside COMPASS. Recent studies start to clarify some novel aspects of differential regulation functioning within COMPASS to achieve the proper H3K4 methylation states. Ahead of the latest studies, Schneider et al., set out a comparative GPS by using all of the three specific antibodies towards the mono-, di-, and trimethylation of H3K4, and in combination with the biochemical analysis, determined each function of Cps40 and Cps60 (presumably through the Cps60-Cps25 subcomplex) in transitioning from di- to trimethylation and from mono- through trimethylation, respectively (Schneider et al., 2005). The roles of the WD repeat proteins, Cps50 and Cps30, in the regulation of H3K4 methylation have remained elusive because of a lack of instability of Set1 within Cps50 or Cps30-deficient COMPASS (Schneider et al., 2005; Dehe et al., 2006). Following studies (Dou et al., 2006; Patel et al., 2008a; Steward et al., 2006) using a defined *in vitro* system of recombinant proteins of MLL1 C-terminal SET domain-containing fragments and other core components (WDR5 (hCps30), RbBP5 (hCps50), and ASH2L (hCps60)), set out to investigate these roles in reconstitution assays. We confirmed the specific Cps60-mediated regulation for H3K4 di- and trimethylation (Steward et al., 2006). However, H3K4 trimethylation was not observed for the core MLL1 complex reconstituted with recombinant proteins prepared by using *E. coli* as a host strain (Patel et al., 2008a).

In addition to the requirement of the subunits of COMPASS in regulating the pattern of mono-, di-, and trimethylation, several key residues within Set1's catalytic pocket are also capable of regulating the trimethylation activity of COMPASS (Takahashi et al., 2009). One such residue is the Tyrosine 1052 (Y1052) of Set1. Our studies have demonstrated that Y1052 plays an essential role in the regulation of the available space within the catalytic pocket of Set1, and thereby regulating the transitioning from mono- to di- to trimethylation (Figure 1). This residue Y1052 is of great interest to us due to the fact that it corresponds to a site known as the F/Y switch governing product specificities for several of the known histone methyltransferases (Figure 1-2). SET domain monomethyltransferases such as SET8 and SET7/9 bear a Y in this position, and the SET domain di- and trimethyltransferases, such as G9A and DIM-5, harbor phenylalanine or other hydrophobic residues in this site. COMPASS is an exception to this model as the SET domain of Set1 possesses a tyrosine in the F/Y switch position, but the enzyme is capable of catalyzing the mono-, di-, and trimethylation of H3K4.

We have proposed that a subunit of COMPASS, Cps40, which is required for proper H3K4 trimethylation by the complex, is the subunit responsible for this exception to the F/Y switch model for COMPASS. We put into action a combinatorial analysis of COMPASS in both the presence and absence of several of its subunits including *cps40* and *cps60*, and site-directed mutagenesis focusing on the putative F/Y switch of the active site, which governs the product specificities (Takahashi et al., 2009). We proved that the F/Y switch model is certainly operative in Set1 of COMPASS by substituting the Tyr1052 with Phe (Y1052F), with the specific enhancement of the H3K4 trimethylation activity exhibited *in vivo* and *in vitro* (Takahashi et al., 2009). More importantly, the enhancement effects suppressed the H3K4 trimethylation defect of Cps40-deficient COMPASS and even reached both the H3K4 di- and trimethylation defects of COMPASS lacking Cps60, as well. Those results suggested not only that the F/Y switch rule dominates all of the SET-domain proteins including the SET1 family, but also that the Tyr1052 residue of Set1 functions together with Cps40 and Cps60 to implement the H3K4 di- and trimethylations properly (Takahashi et al., 2009) (Figure 2). The SET1 family of proteins may have evolved to obtain the intrinsically down-regulated catalytic domain to adapt to the necessity of the differential and diverse levels of

H3K4 methylation regulation through the integration of the signals of their various unique and common subunits.

Apart from the SET domain, several regulatory domains or motifs within the Set1 polypeptide were identified through genetic and biochemical analyses. Schlichter and Cairns characterized and suggested that the RRM (RRM1) and autoinhibitory domains N-terminal to the SET domain to be positive and negative modulators toward H3K4 di- and trimethylation, respectively (Schlichter and Cairns, 2005). The following study found another RRM domain (RRM2) adjacently to the C-terminal to RRM1 and the destabilizing effect on Set1 of the RRM1 mutation (Tresaugues et al., 2006), implying that the inactivation of H3K4 methylation of the RRM1 mutant Set1 may be due to the loss of COMPASS's integrity, not the catalytic effects. It is of note that the amino acid sequences from the N-terminal to the SET domain are relatively divergent than the above mentioned RRM1, RRM2, and that the auto-inhibitory domains are only conserved among *S. cerevisiae* Set1, though not among human Set1A and Set1B, and the MLL1-4 proteins. Prominently, in recent studies, the WDR5-interaction motif, Win motif, which is strictly conserved among human Set1A, Set1B, and the MLL1-4 proteins, was identified immediately upstream of the SET domain of MLL1 together with the co-crystals of WDR5 and a Win peptide from MLL1 (Patel et al., 2008a; Patel et al., 2008b; Song and Kingston, 2008). Their strong affinities with the dissociation constant of sub-micromolars between the recombinant core MLL1 fragment and WDR5 in a stoichiometric ratio can be attributed solely to the Win motif, suggesting that the Win motif is the critical determinant for the binding to WDR5 of the human Set1 and MLL proteins within hCOMPASS and hCOMPASS-like complexes. For unknown reasons, the Win motif is not obviously conserved in *S. cerevisiae* Set1 (data not shown). However, the C-terminally truncated Set1 that includes the N-SET region could be stably purified with Cps30 (WDR5 in humans) and Cps50. Cps30 may still bind to Set1 in a similar way to WDR5 and MLL1, but through different residues within the N-SET region.

Recently, Vitaliano-Prunier et al. have proposed a Cps35-mediated activation of COMPASS, however, by a different mechanism than that which we have proposed (Vitaliano-Prunier et al., 2008). Vitaliano-Prunier have suggested that Cps40 of COMPASS accomplishes H3K4 trimethylation crosstalk based on the observations that Cps35 was subject to ubiquitination by the Rad6/Bre1 complex when both Cps35 and ubiquitin are overexpressed in yeast (Vitaliano-Prunier et al., 2008). Vitaliano-Prunier and colleagues also have proposed that such ubiquitination of Cps35 is involved in the recruitment and association of Cps40 with COMPASS and chromatin. The Cps35 ubiquitination remains quite controversial because the observed Cps35 ubiquitination was under overexpression conditions for both Cps35 and ubiquitin, and the *in vitro* reconstituted analysis using the purified components were not performed (Vitaliano-Prunier et al., 2008). Indeed, Vitaliano-Prunier and colleagues are not able to detect even a 1% ubiquitination of Cps35 when they overexpress both Cps35 and ubiquitin. Furthermore, our analysis of the H2B monoubiquitination deficient cells such as *rad6Δ* and the H2BK123R strains clearly demonstrates a stable Cps40 association on chromatin in an H2B monoubiquitination-independent manner as well as its stable assembly within COMPASS (Takahashi et al., 2009). Furthermore, when COMPASS is purified from strains lacking Cps35, the same levels of Cps40 are purified with COMPASS suggesting that Cps35 has very little to do in recruiting Cps40 to COMPASS as suggested by Vitaliano-Prunier and colleagues (Takahashi et al., 2009). Overall, our studies suggest that H2B monoubiquitination inducing the Cps35 assembly into and activating COMPASS is a distinct mechanism from the Cps40-mediated adjustments of the H3K4 trimethylation activity in conjunction with the Y1052 residue in the Set1 catalytic site.

H2B monoubiquitination is not the only crosstalk for H3K4 methylation by COMPASS

In addition to H2BK123 monoubiquitination, the surfaces of the nucleosome can play important roles for the implementation of H3K4 methylation. The scanning histone mutagenesis with alanine (SHIMA) library analysis elegantly showed and elucidated a novel patch on the histone H2A/H2B dimer and a lysine residue of the histone H3 N-terminal tail (H3K14) as trans- and cis-regulatory regions for H3K4 methylation, working possibly through the effects on the H2B ubiquitination machineries and/or the direct interactions with COMPASS (Nakanishi et al., 2008). Interestingly, it was recently revealed that other histone lysine methylations such as H3K36 and H3K79 also require the respective histone trans-regulatory regions and corresponding interactions with their methyltransferase enzymes such as H4K44 and the N-terminal acidic patch of Set2, and the H4R17-H18-R19 basic patch and the C-terminal acidic patch of Dot1. Compared with these two trans-regulatory pathways, which are obviously based on the direct charge-based electrostatic interactions between each enzyme and histone protein in the nucleosome, the COMPASS-associated histone regulatory regions are relatively complicated, because the H2A/H2B patch consists of a mixture of both acidic (H2AE65-N69-D73) and basic (H2BH112-R119) residues, the latter of which is proximal to the monoubiquitination site (K123) of H2B, and furthermore, some of those residues affect H2B monoubiquitination positively and others do so negatively. Nevertheless with similar H3K4 di-and/or trimethylation defects observed when measuring bulk histones from the different mutants (Nakanishi et al., 2008), further detailed genetic and biochemical analyses are required for an exact understanding of those these regulation mechanisms. However, histone regulatory pathways requiring these novel crosstalks among histone methyltransferases and nucleosome histone surfaces are certainly one of the emerging concepts in the studies of the epigenetic histone posttranslational modification network.

Summary

A series of systematic and candidate approaches centered on H3K4 methylation using yeast as a model system, has given rise to numerous insights into a trans-histone tail crosstalk between H2B monoubiquitination and H3K4 methylation, as well as structures and functions of the proteins mediating these effects. The resulting molecular framework is being applied to many other multi-cellular eukaryotes, leading to the establishment of the widely conserved molecular axis of H2B monoubiquitination-H3K4 methylation as a regulator of developmental gene expression programs. However, current knowledge of H2B monoubiquitination-H3K4 methylation machineries is quite far from the precise understanding of how such histone modifications are introduced, maintained, and most importantly involved in facilitating and/or repressing gene expression. Indeed, relationships between H2B monoubiquitination-H3K4 methylation and transcription are obviously highly tangled, in which H2B monoubiquitination, apart from H3K4 methylation, independently activates gene expression while both modifications not only affect transcriptional efficiency, but also necessitate the existing transcription for the implementation through the interaction with the PAF complex. In addition to this interdependency, inaccessibility of the whole transcription system, including chromatin, activators and repressors, GTF, Pol II, and innumerable upstream and downstream chromatin modifying complexes associated with H2B monoubiquitination and H3K4 methylation, to defined *in vitro* analyses, have hindered quantitative evaluations of H2B monoubiquitination and H3K4 methylation, and impeding chromatin researchers from defining their direct biological outputs. These difficulties are common for many other histone modifications as well; to overcome them, it is inevitable to search for and complete the full complement of histone posttranslational modification and histone modifying enzymes. Recent high throughput analyses targeting histone proteins through mutagenesis (Shima et al., 2008; Dai et al., 2008) and future studies based on those histone mutant libraries definitely accelerate completing the list of the network of histone

posttranslational modifications and their associations with the biological outputs including transcription.

In relation to H3K4 methylation, the emerging and exciting question is in regards to the enzymatic regulations of multi-protein COMPASS and hCOMPASS-like complexes. Although, since the identification of the SET domain as a histone methyltransferase, characterizations of monomeric SET domain enzymes have thoroughly advanced biophysically and biochemically as described above, those of multi-subunit complexes such as COMPASS considerably lag behind. Recent determination of a crystal structure of the SET domain of MLL1 as well as the identification of the catalytic key residue of the Phe/Tyr switch will help further the understanding of the molecular regulation unique to the multi-protein COMPASS complex. However, most eagerly awaited is the structural information of the holo-COMPASS complex exhibiting the full H3K4 mono-, di-, and trimethyltransferase activities. The configuration(s) of the SET domain with other non-catalytic subunits, as well as any alteration(s) observed on the SET domain structure, must provide valuable insight for not only a deeper understanding of the molecular mechanisms of H3K4 methylation, but also for the development of novel therapeutics and chemical drugs targeting the catalytic SET domain and/or regulating non-catalytic subunits for the treatment of MLL1-mediated leukemia and cancer diseases.

Acknowledgments

We thank Dr. Edwin Smith for critical reading of this manuscript and Laura Shilatifard for editorial assistance. The work in Shilatifard's laboratory is supported by the grants R01GM069905 and R01CA89455 from the National Institute of Health to ASH.

References

- Berger SL. Histone modifications in transcriptional regulation. *Curr Opin Genet Dev* 2002;12(2):142–8. [PubMed: 11893486]
- Berger SL. The complex language of chromatin regulation during transcription. *Nature* 2007;447(7143):407–12. [PubMed: 17522673]
- Dou Y, Milne TA, Ruthenburg AJ, Lee S, Lee JW, Verdine GL, Allis CD, Roeder RG. Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol* 2006;13(8):713–9. [PubMed: 16878130]
- Dover J, Schneider J, Tawiah-Boateng MA, Wood A, Dean K, Johnston M, Shilatifard A. Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *J Biol Chem* 2002;277(32):28368–71. [PubMed: 12070136]
- Hughes CM, Rozenblatt-Rosen O, Milne TA, Copeland TD, Levine SS, Lee JC, Hayes DN, Shanmugam KS, Bhattacharjee A, Biondi CA, Kay GF, Hayward NK, Hess JL, Meyerson M. Menin associates with a trithorax family histone methyltransferase complex and with the *hoxc8* locus. *Mol Cell* 2004;13(4):587–97. [PubMed: 14992727]
- Kim J, Guermah M, McGinty RK, Lee JS, Tang Z, Milne TA, Shilatifard A, Muir TW, Roeder RG. RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. *Cell* 2009;137(3):459–71. [PubMed: 19410543]
- Kim J, Hake SB, Roeder RG. The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interactions. *Mol Cell* 2005;20(5):759–70. [PubMed: 16337599]
- Kornberg RD. Chromatin structure: a repeating unit of histones and DNA. *Science* 1974;184(139):868–71. [PubMed: 4825889]
- Kornberg RD. The molecular basis of eukaryotic transcription. *Proc Natl Acad Sci U S A* 2007;104(32):12955–61. [PubMed: 17670940]
- Krogan NJ, Dover J, Khorrami S, Greenblatt JF, Schneider J, Johnston M, Shilatifard A. COMPASS, a histone H3 (Lysine 4) methyltransferase required for telomeric silencing of gene expression. *J Biol Chem* 2002;277(13):10753–5. [PubMed: 11805083]

- Lee JH, Skalnik DG. Wdr82 is a C-terminal domain-binding protein that recruits the Setd1A Histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. *Mol Cell Biol* 2008;28(2):609–18. [PubMed: 17998332]
- Lee JS, Shukla A, Schneider J, Swanson SK, Washburn MP, Florens L, Bhaumik SR, Shilatifard A. Histone Crosstalk between H2B Monoubiquitination and H3 Methylation Mediated by COMPASS. *Cell* 2007;131(6):1084–96. [PubMed: 18083099]
- Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 1997;389(6648):251–60. [PubMed: 9305837]
- Miller T, Krogan NJ, Dover J, Erdjument-Bromage H, Tempst P, Johnston M, Greenblatt JF, Shilatifard A. COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc Natl Acad Sci U S A* 2001;98(23):12902–7. [PubMed: 11687631]
- Nakanishi S, Sanderson BW, Delventhal KM, Bradford WD, Staehling-Hampton K, Shilatifard A. A comprehensive library of histone mutants identifies nucleosomal residues required for H3K4 methylation. *Nat Struct Mol Biol* 2008;15(8):881–8. [PubMed: 18622391]
- Nowak SJ, Corces VG. Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet* 2004;20(4):214–20. [PubMed: 15041176]
- Patel A, Dharmarajan V, Cosgrove MS. Structure of WDR5 bound to mixed lineage leukemia protein-1 peptide. *J Biol Chem* 2008a;283(47):32158–61. [PubMed: 18829459]
- Patel A, Vought VE, Dharmarajan V, Cosgrove MS. A conserved arginine-containing motif crucial for the assembly and enzymatic activity of the mixed lineage leukemia protein-1 core complex. *J Biol Chem* 2008b;283(47):32162–75. [PubMed: 18829457]
- Pavri R, Zhu B, Li G, Trojer P, Mandal S, Shilatifard A, Reinberg D. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* 2006;125(4):703–17. [PubMed: 16713563]
- Robzyk K, Recht J, Osley MA. Rad6-dependent ubiquitination of histone H2B in yeast. *Science* 2000;287(5452):501–4. [PubMed: 10642555]
- Roguev A, Schaft D, Shevchenko A, Pijnappel WW, Wilm M, Aasland R, Stewart AF. The *Saccharomyces cerevisiae* Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *Embo J* 2001;20(24):7137–48. [PubMed: 11742990]
- Schlichter A, Cairns BR. Histone trimethylation by Set1 is coordinated by the RRM, autoinhibitory, and catalytic domains. *Embo J* 2005;24(6):1222–31. [PubMed: 15775977]
- Schneider J, Wood A, Lee JS, Schuster R, Dueker J, Maguire C, Swanson SK, Florens L, Washburn MP, Shilatifard A. Molecular regulation of histone H3 trimethylation by COMPASS and the regulation of gene expression. *Mol Cell* 2005;19(6):849–56. [PubMed: 16168379]
- Shilatifard A. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu Rev Biochem* 2006;75:243–69. [PubMed: 16756492]
- Shilatifard A. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Curr Opin Cell Biol* 2008;20(3):341–8. [PubMed: 18508253]
- Song JJ, Kingston RE. WDR5 interacts with mixed lineage leukemia (MLL) protein via the histone H3-binding pocket. *J Biol Chem* 2008;283(50):35258–64. [PubMed: 18840606]
- Steward MM, Lee JS, O'Donovan A, Wyatt M, Bernstein BE, Shilatifard A. Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. *Nat Struct Mol Biol* 2006;13(9):852–4. [PubMed: 16892064]
- Takahashi YH, Lee JS, Swanson SK, Saraf A, Florens L, Washburn MP, Trievel RC, Shilatifard A. Regulation of H3K4 trimethylation via Cps40 (Spp1) of COMPASS is monoubiquitination independent: implication for a Phe/Tyr switch by the catalytic domain of Set1. *Mol Cell Biol* 2009;29(13):3478–86. [PubMed: 19398585]
- Vitaliano-Prunier A, Menant A, Hobeika M, Geli V, Gwizdek C, Dargemont C. Ubiquitylation of the COMPASS component Swd2 links H2B ubiquitylation to H3K4 trimethylation. *Nat Cell Biol* 2008;10(11):1365–71. [PubMed: 18849979]
- Wood A, Krogan NJ, Dover J, Schneider J, Heidt J, Boateng MA, Dean K, Golshani A, Zhang Y, Greenblatt JF, Johnston M, Shilatifard A. Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. *Mol Cell* 2003;11(1):267–74. [PubMed: 12535539]

- Wu M, Wang PF, Lee JS, Martin-Brown S, Florens L, Washburn M, Shilatifard A. Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. *Mol Cell Biol* 2008;28(24):7337–44. [PubMed: 18838538]
- Zhang Y. Transcriptional regulation by histone ubiquitination and deubiquitination. *Genes Dev* 2003;17(22):2733–40. [PubMed: 14630937]
- Zhu B, Zheng Y, Pham AD, Mandal SS, Erdjument-Bromage H, Tempst P, Reinberg D. Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. *Mol Cell* 2005;20(4):601–11. [PubMed: 16307923]

Figure 1. Homology model of the catalytic domain of Set1

a.) Ribbon representation of the secondary structure of the Set1 catalytic domain illustrating the SET domain (blue) and PostSET motif (magenta). A histone H3 substrate peptide spanning residues 1-6 (orange carbon atoms) and the product S-adenosyl-L-homocysteine (AdoHcy; green carbons) are docked in the active site. The side chains of the four cysteines that coordinate the Zn ion in the PostSET domain are also illustrated. **b.)** Active site of Set1. Residues comprising the enzyme's active sites are shown with gray carbon atoms, while the residues mutated in this study are highlighted with magenta carbons. AdoHcy and H3K4 are depicted as in (a). Hydrogen bonds are illustrated by cyan dashed lines.

Figure 2. Model illustrating the Cps40 and 60 function in the regulation of HMTase activity of Set1/COMPASS

Cps40 and Cps60 (depicted in yellow diamond) induces the conformational change of Set1/SET domain such that pulling back Tyr1052 (light purple trapezoid) so that its HMTase active site encompassing the substrate of di-methylated lysine gets to be more open making it easier to accommodate tri-methylated lysine. In the Y1052F mutant Set1, mono- and di-methylation proceed similarly to wild-type. However, F1052 (dark purple) lacking the hydroxyl group (blue small trapezoid) from the phenol ring side chain, allows enough space for methyl transfer to the trimethylated H3 peptide without the aid of Cps40. This model is consistent with our previous observation that the loss of Cps60 and Cps40 results in the loss H3K4 trimethylation (Schinder et al., 2005). Our molecular modeling and experimental results demonstrating complementation of H3K4 trimethylation in a SetY1052F in the absence of Cps40 strongly supports this model. Arg1013 (blue rectangle), whose substitution results in the complete loss of H3K4 methylation, should be involved in holding and/or the arrangement of the cofactor SAM on its own or via an interaction with Y993 (Figure 1b).