Minireview

Transcriptional Elongation by RNA Polymerase II and Histone Methylation*

Mark Gerber† and Ali Shilatifard‡§¶¶

From the †Department of Biochemistry and the ‡St. Louis University Cancer Center, St. Louis University School of Medicine, St. Louis, Missouri 63104

mRNA synthesis in eukaryotic organisms is a key biological process that is regulated at multiple levels. From the covalent modifications of chromatin by a number of chromatin remodeling complexes during the initiation and activation steps of transcription to the processing of mRNA transcripts, a very large consortium of proteins and multiprotein complexes is critical for gene expression by RNA polymerase II. The list of proteins essential for the successful synthesis of mRNA continues to grow at a rapid pace. Recent advances in this area of research have been focused on transcription through chromatin. In this article, we will review the recent literature linking the key biochemical process of transcriptional elongation by RNA polymerase II to histone methylation by COMPASS, Dot1p, and Set2 methyltransferases.

The RNA Polymerase II Elongation Factors

RNA polymerase II transcription proceeds through multiple stages, designated preinitiation, initiation, and elongation. Historically, efforts to understand the elongation phase of the transcription cycle lagged behind efforts geared toward unraveling the processes of preinitiation and initiation. However, during the past several years an immense amount of research has been performed to identify and characterize transcription factors that regulate the elongation stage of mRNA synthesis by RNA polymerase II. Although there has been significant progress in the biochemical characterization of these factors, little is known about their physiological role in development in multicellular organisms or at what stages of transcription elongation they function. However, recent studies have begun to address the roles for some of the RNA polymerase II elongation factors using in vivo model systems.

The known elongation factors fall into at least three functional classes. The first class is composed of factors involved in drug-induced or sequence-dependent arrest and includes the SII (also known as TFIIS) family of elongation factors and PTEF-b (positive transcription elongation factor b) (1, 2). SII was the first RNA polymerase II elongation factor to be purified and is known to promote efficient RNA polymerase II elongation by preventing DNA sequence-dependent premature arrest (3, 4). Furthermore, SII and its bacterial counterparts are the only known factors that can reactivate arrested polymerase (5). PTEF-b is a heterodimeric factor typically composed of cyclin T and Cdk9 and has been shown to phosphorylate the C-terminal domain (CTD) of the largest subunit of RNA polymerase II, a critical event in the transition from initiation to elongation (6). In vivo cross-linking studies in yeast have demonstrated that phosphorylation of serine 5 of the pol II CTD is localized to promoter and promoter distal regions of genes and is therefore a marker for early elongation complexes (7). However, phosphorylation of serine 2 of the CTD exhibits a nearly complementary pattern of localization and is found throughout the coding regions and nearer the 3′-end genes, suggesting that serine 2 phosphorylation is a mark for RNA polymerase engaged in the processive phases of transcription elongation (7).

The second class of factors functions to regulate the rate of elongation through chromatin and includes FACT (facilitates chromatin transcription) (8). A third class operates by increasing the catalytic rate of elongation by altering the Kcat and/or the Vmax of pol II and includes TFIIF, the ELL family of proteins (3, 4, 9–14). ELL, which is found in translocation with the MLL gene in patients with acute myeloid leukemia, was first biochemically isolated based on its ability to increase the catalytic rate of transcription elongation by RNA polymerase II (13, 15). Three mammalian and one Drosophila homologue of the ELL family have been identified and characterized, and biochemical and genetic evidence now suggests that (a) ELL is an essential gene, (b) ELL associates with actively elongating RNA pol II in vivo, and (c) ELL is required for the efficient expression of specific genes (13, 14, 16–18). Additionally, recent observations have refuted the previously conceived notion that elongation factors that have similar biochemical activities have overlapping or redundant functions in cells (18).

In Saccharomyces cerevisiae, the Paf1 complex was both genetically and biochemically isolated and demonstrated to function at both the initiation and elongation stages of transcription (19–24). Most importantly, recent studies have demonstrated that the Paf1 complex is physically associated with elongating RNA polymerase II (22–24). However, although the Paf1 complex has been shown to associate with elongating RNA polymerase, its exact roles in transcription elongation remained unclear until recently.

Chromosomal Modifications by Histone Methylation

A critical step in the activation of a gene is to make the DNA near the promoter regions more accessible to the basal transcription machinery and RNA polymerase II. In its natural state, the two-meter long DNA is packed around a core histone octamer (containing two copies of histones H2A, H2B, H3, and H4) called a nucleosome (25, 26). This structure can be further compacted with the association of histone H1, making the DNA template inaccessible to the transcriptional machinery. Compacted chromatin must be “loosened” for transcription and gene expression to occur. This process is accomplished by a set of enzymes known as chromatin remodeling or modifying complexes. These complexes function by one of two mechanisms. The first mechanism is dependent upon ATP hydrolysis and catalyzes a shifting of nucleosomes along DNA (27–29). The second mechanism involves the covalent modification of the histone proteins that make up the nucleosomes (30, 31). These modifications include acetylation, phosphorylation, methylation, and ubiquitination. Although modifications of histones were identified many years ago (32, 33), it is only recently that concentrated research from many laboratories has begun to elucidate how they function in regulation of gene expression. Most of these modifications such as ubiquitination are also involved in posttranslational modification of other components of the transcription machinery in regulation of transcription (34).

Although histone acetylation has been actively investigated for a number of years, histone methylation (and the enzymes that catalyze this process) is not as thoroughly understood. Recent efforts have begun to shed light on how methylation of histones is important in the regulation of gene expression, and these efforts have centered around the identification and characterization of a number of histone methyltransferase (HMT) complexes, some of which

* This minireview will be reprinted in the 2003 Minireview Compendium, which will be available in January, 2004.

† To whom correspondence should be addressed: Dept. of Biochemistry, St. Louis University School of Medicine, 1402 South Grand Blvd., St. Louis, MO 63104. Tel.: 314-577-8137; Fax: 314-268-5737; E-mail: shilatifard@slu.edu.

‡ The abbreviations used are: CTD, C-terminal domain; pol II, polymerase II; HMT, histone methyltransferase.

Published, JBC Papers in Press, May 22, 2003, DOI 10.1074/jbc.R300014200

© 2003 by The American Society for Biochemistry and Molecular Biology, Inc.
will be described here. Histone methylation occurs at a number of residues within all four histones in the nucleosomal core, suggesting that this type of modification could have a significant effect on the regulation of chromatin structure. Recently, the identification of a number of novel HMTs has implicated histone methylation as an important event in biological processes such as telomeric silencing, transcriptional activation, and transcriptional repression (35–49). SUV39 was identified in *Drosophila* as a suppressor of position-effect variegation and was the first histone methyltransferase to be identified (50). Both trithorax (trx) and polycomb (Pc) group proteins contain a 130–140-amino acid motif called the SET domain, which is found in a variety of chromatin-associated proteins. This domain takes its name from the *Drosophila* proteins *Su(var)3–9*, Enhancer of zeste (E(z)), and Trithorax. Searching the data base for other SET domain-containing proteins has revealed many other potential histone methyltransferase candidates in a phylogenetically diverse group of organisms. These SET domain-containing proteins are involved in methylation of lysine residues of different histones. Fig. 1 illustrates the diverse pattern of histone methylation on lysines 4, 9, and 79 of histone H3 and the localization pattern of serine 2 phospho-RNA polymerase II on *Drosophila* polytene chromosomes. For example, methylation of histone H3 on lysine 9 occurs at the fused mass of heterochromatin known as the chromocenter and is associated with position-effect variegation and transcriptional silencing (51). Because the transcriptionally active (i.e., elongating) form of RNA polymerase II is phosphorylated on serine 2 of its CTD, a comparison of the pattern of distribution for lysine 4 and 79 methylated histone H3 with this form of RNA polymerase II indicates that these modifications appear to colocalize. This observation therefore implies that lysine 4 and 79 methylation of histone H3 is associated with the transcriptionally active euchromatin.

In *S. cerevisiae*, there are six SET domain-containing proteins. We will focus on Set1 and Set2 in this article. Yeast Set1 was originally purified as part of a multiprotein complex, termed COMPASS (COMplex of Proteins ASSociated with Set1) (42, 43, 52). Related members in this family of proteins include those encoded by the polycym and trithorax groups of genes, such as the mammalian MLL gene and *Drosophila* TRX. In early studies, COMPASS was implicated in telomeric silencing, as a number of deletion strains for the various components of this complex showed defects in silencing genes located in telomeric regions (42, 43). COMPASS can also methylate lysine 4 of histone H3, and this modification is linked to telomeric silencing (42, 43, 52).

To identify proteins involved in the methylation of histone H3 by COMPASS, we developed Global Proteome analysis of *S. cerevisiae* (GPS) to test by Western blotting extracts of each of the nonessential yeast gene deletion mutants for defects in methylation of lysine 4 of histone H3 (47). Employing this novel method, it was demonstrated that methylation of histone H3 by COMPASS requires the ubiquitination of lysine 123 of histone H2B in a process involving the ubiquitin-conjugating enzyme, Rad6 (47). Allis and colleagues (49) also reported the same observation independently employing a different method. Because Rad6 is an E2 ubiquitin-conjugating enzyme involved in diverse biological pathways such as the N-end rule pathway, DNA repair, recombination, and transcription, many laboratories have been searching to identify the specific E3 ligase that brings Rad6 to the transcription pathway. Employing GPS, we identified the C3HC4 ring finger protein Bre1 as the E3 ligase that is required for ubiquitination of histone H2B (48, 54). These observations also help to solidify the premise that there is indeed cross-talk between pathways used to covalently modify histone octamers.

The Set2 protein was purified and characterized based on its nucleosome-specific HMT activity (39). Set2 is responsible for the methylation of the lysine 36 residue of histone H3 in *S. cerevisiae* (39). Set2 has been demonstrated to be a potent repressor of transcription, suggesting a role for lysine 36 methylation in down-regulating gene expression (39). However, lysine 36 methylation is observed in transcriptionally active macronuclei in *Tetrahymena*, supporting an additional role for this modification in the potentiation of transcription (55). Recent evidence from several laboratories has provided a clearer picture of the role of Set2 in transcription, as will be detailed below.

The Dot1 protein was originally identified in a genetic screen for high copy suppression of telomeric silencing (36, 56). This effect was later shown to require Dot1-mediated methylation of the lysine 79 residue of histone H3, which is found within the histone globular domain rather than the tail of the protein, where a large portion of modifications occur (36, 40). Unlike the Set family of proteins, Dot1 does not contain a SET domain but instead methylates proteins via a methylase fold. Early observations of the role of lysine 79 methylation by Dot1 demonstrated that this modification of histone H3 inhibits binding of the Sir silencing proteins at the telomerases and therefore regulates telomeric silencing (36, 40). By analogy, it is likely that methylation of histone H3 on lysine 4 also inhibits binding of the Sir silencing proteins at telomerases and hence is associated with active loci as in other organisms. Most importantly, both histone H3 lysine 79 and lysine 4 methylation seem to occur throughout the genome, but it appears slightly higher at the coding regions of active genes (57, 58).

**A Role for Histone Methylation of Lysine 4 of Histone H3 in Transcription Elongation**

As mentioned previously, COMPASS, the yeast Set1-containing complex, is the HMT that catalyzes the methylation of lysine 4 of histone H3 (42, 43, 46). Employing GPS, we recently identified several other proteins required for methylation of lysine 4 by COMPASS. Among this group of proteins are some components of the Paf1 complex that have been found to associate with the elongating form of RNA polymerase II (59). The Paf1 complex localizes to promoters and to the body of genes and is required for the expression of a number of genes, mainly those responding to the Pkc1/mitogen-activated protein kinase-signaling cascade (60). Further investigation has revealed that the physical interaction between COMPASS and RNA polymerase II requires the presence of some of the components of the Paf1 complex. Although the interaction between COMPASS and RNA polymerase II requires the Paf1 complex, the association of Paf1 with RNA polymerase II is not dependent upon COMPASS, indicating a role for the Paf1 complex as a "platform" for COMPASS interaction with RNA polymerase II (59).

Biochemical studies have also demonstrated that COMPASS interacts with the form of RNA polymerase II whose CTD is phos-
phorylated by TFIIH.

of the preinitiation complexes, the CTD of RNA polymerase II is phoshorylated by TFIIH and ubiquitinated histone H2B (48). 

ubiquitin-conjugating enzyme is recruited to the promoter by its E3 ligase Bre1 and ubiquitates histone H2B (48).

A Role for Histone Methylation of Lysine 79 of Histone H3 in Transcription Elongation

As mentioned above, Dot1 methylase is the sole HMT responsible for the methylation of lysine 79 in S. cerevisiae (36, 40, 62). This modification has been hypothesized exclusively to regulate repression of genes in telomeric regions by facilitating Sir protein association with chromatin (36, 40). However, recent reports indicate that this modification is also found in euchromatic regions of the genome (Fig. 1). As with lysine 4 methylation, we have also demonstrated that the components of the Paf1 complex are also required for lysine 79 methylation in S. cerevisiae (59).

Although biochemical evidence to support an interaction between Dot1 and RNA polymerase II via the Paf1 complex is not yet available, the striking similarities between the regulation of Dot1-mediated lysine 79 methylation and COMPASS-mediated lysine 4 methylation suggest a similar mechanism by which Dot1 is recruited to chromatin. These observations are also consistent with a model in which the Paf1 complex functions as a “platform” for interaction of histone methyltransferases with elongating RNA polymerase II.

A Role for Histone Methylation of Lysine 36 of Histone H3 in Transcription Elongation

As previously described, the Set2 protein is the HMT required for the methylation of lysine 36 of histone H3 (39). In efforts to understand the role of Set2 as a transcriptional regulatory factor and to investigate its potential targeting mechanism, we and others have recently demonstrated an interaction between Set2 and RNA polymerase II (63–66). Chromatin immunoprecipitation experiments have demonstrated that Set2 and methylation of lysine 36 of histone H3 are associated with the actively transcribed coding regions of several genes. Most interestingly, interaction of Set2 with RNA polymerase II requires the CTD of the largest subunit of pol II. Specifically, Set2 preferentially associates with the serine 2 phosphorylated form of pol II, indicating that it associates with both early elongating and processively elongating RNA polymerase II. Phosphorylation of the CTD of the large subunit of pol II is essential for Set2 to bind RNA polymerase II and histone H3 on lysine 36 (64–66).

Work performed in several laboratories has indicated that interaction of Set2 with the elongating pol II is lost in deletion strains lacking the CTD serine 2 kinase, Ctk1 (65, 66). Interestingly, Ctk1 deletion strains are defective for lysine 36 methylation, thus revealing the mechanism targeting Set2 to chromatin. Unlike Dot1 and COMPASS, methylation of lysine 36 by Set2 does not require ubiquitination of lysine 123 of histone H2B. Rather, it appears that phosphorylation is the key event leading to lysine 36 methylation, implicating a selective pathway for this modification. However, similar to COMPASS and Dot1, interaction of Set2 with the elongating RNA polymerase II requires the presence of the Paf1 complex, further supporting a role for the Paf1 complex as a “platform” for the interaction of histone methyltransferases with the elongating RNA polymerase II. Alternatively, it is also possible that the Paf1 complex may function as a “platform” for other enzymatic processes during the initiation and the elongation phase of transcription.

Fig. 2. The process of histone modification by ubiquitination and methylation at the promoter and during the early and processive stages of transcription elongation. A and B, the Rad6 ubiquitin-conjugating enzyme is recruited to the promoter by its E3 ligase Bre1 and ubiquitates histone H2B (48). C, after the formation of the preinitiation complexes, the CTD of RNA polymerase II is phosphorylated by TFIIH. D and E, the Paf1 complex functions as a “platform” for the recruitment of COMPASS resulting in the methylation of the fourth lysine of histone H3 at the promoter and early elongation complexes (59). E and F, following the phosphorylation of the CTD by Ctk1 complex, COMPASS departs and again the Paf1 complex functions as a “platform” for recruitment of Set2 methyltransferase resulting in the methylation of the lysine 36 of histone H3. GTF, general transcription factor; SAM, S-adenosylmethionine.
A Role for Histone Methylation through the Elongating RNA Polymerase II as a Mark of “Transcriptional Memory” for Recently Transcribed Genes

One possible role of histone methylation on the body of protein coding genes may be to function as a molecular memory for recently transcribed genes. Specifically, methylation of histone H3 on lysines 4, 36, and 79 and the requirement of the Paf1 complex for this histone modification are consistent with this idea. Several recent studies have demonstrated that lysine 4 of histone H3 can be methylated in coding regions during the transcription of a gene and that this histone methylation persists for a considerable time after transcriptional inactivation and the dissociation of COMPASS (53, 57, 59, 61). Therefore, histone methylation on the fourth lysine of histone H3 may inform the cell that transcription of a given gene has occurred in the recent past but is not necessarily happening at the present time. Although it has been demonstrated that histone methylation clearly lasts for a significant portion of an individual cell cycle, this modification is not faithfully transmitted to all daughter cells (61). Thus, histone marking by lysine 4 methylation appears to provide memory for recently transcribed genes that is mechanistically distinct from the epigenetic memory that occurs in position-effect variegation and transcriptional silencing via methylation on lysine 9 of histone H3.

If histone H3 lysine 4 methylation functions as a mark for recently transcribed genes, why then do cells also use methylation of histone H3 lysine 36 and lysine 79 for the same process? One possible explanation of this observation is that different regions of the body of a transcribed gene may be marked with different types of methylation. For example, histone H3 lysine 4 methylation is cleared by a factor that has described methylation may be marked with different types of methylation. Several recent studies have demonstrated that lysine 4 of histone H3 can be methylated in coding regions during the transcription of a gene and that this histone methylation persists for a considerable time after transcriptional inactivation and the dissociation of COMPASS (53, 57, 59, 61). Therefore, histone methylation on the fourth lysine of histone H3 may inform the cell that transcription of a given gene has occurred in the recent past but is not necessarily happening at the present time. Although it has been demonstrated that histone methylation clearly lasts for a significant portion of an individual cell cycle, this modification is not faithfully transmitted to all daughter cells (61). Thus, histone marking by lysine 4 methylation appears to provide memory for recently transcribed genes that is mechanistically distinct from the epigenetic memory that occurs in position-effect variegation and transcriptional silencing via methylation on lysine 9 of histone H3.

REFERENCES

1. Reines, D., and Mote, J., Jr. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 1917–1921.

2. Marshall, N. F., and Price, D. H. (1995) J. Biol. Chem. 270, 12335–12338.

3. Reines, D., Conaway, J. W., and Conaway, R. C. (1996) Trends Biochem. Sci. 21, 351–355.

4. Conaway, R. C. and Conaway, J. W. (1996) Annu. Rev. Biochem. 68, 301–319.

5. Wind, M. and Reines, D. (2000) Bioessays 22, 327–336.

6. Marshall, N. F., Peng, J., Xie, Z., and Price, D. H. (1996) J. Biol. Chem. 271, 27176–27183.

7. Komarnitsky, P., Cho, E. J., and Buratowski, S. (2000) Genes Dev. 14, 2452–2460.

8. Orphanides, G., LeRoy, G., Chang, C. H., Luse, D. S., and Reinberg, D. (1998) Cell 92, 105–116.

9. Price, D. H., Sluder, A. E., and Greenleaf, A. L. (1989) Mol. Biol. Cell. 9, 1446–1475.

10. Bruderer, J. N., Jackson, K. W., Conaway, R. C., and Conaway, J. W. (1993) J. Biol. Chem. 268, 25587–25593.

11. Garrett, K. P., Tan, S., Bradsher, J. N., Lane, W. S., Conaway, J. W., and Price, D. H. (1996) Bioessays 22, 327–336.

12. Marshall, N. F., Peng, J., Xie, Z., and Price, D. H. (1996) J. Biol. Chem. 271, 27176–27183.

13. Komarnitsky, P., Cho, E. J., and Buratowski, S. (2000) Genes Dev. 14, 2452–2460.

14. Orphanides, G., LeRoy, G., Chang, C. H., Luse, D. S., and Reinberg, D. (1998) Cell 92, 105–116.

15. Price, D. H., Sluder, A. E., and Greenleaf, A. L. (1989) Mol. Biol. Cell. 9, 1446–1475.

16. Bruderer, J. N., Jackson, K. W., Conaway, R. C., and Conaway, J. W. (1993) J. Biol. Chem. 268, 25587–25593.

17. Garrett, K. P., Tan, S., Bradsher, J. N., Lane, W. S., Conaway, R. C., and Conaway, J. W. (1993) J. Biol. Chem. 268, 25587–25593.

18. Duan, D. R., Pause, A., Burgess, W. H., Aso, T., Chen, D. Y., Garrett, K. P., Conaway, R. C., Conaway, J. W., Linehan, W. M., and Klausner, R. D. (1995) Science 269, 1402–1406.

19. Shilatifard, A., Lane, W. S., Jackson, K. W., Conaway, R. C., and Conaway, J. W. (1996) Science 271, 1873–1876.

20. Miller, T., Williams, K., Johnstone, R. W., and Shilatifard, A. (2000) J. Biol. Chem. 275, 30628–30635.

21. Thirman, M. J., Levin, D. A., Kobayashi, H., Simon, M. C., and Rowley, J. D. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 12110–12114.

22. Shilatifard, A., Duan, D. R., Hargus, D., Florence, C., Schubach, W. H., Conaway, J. W., and Conaway, R. C. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 3639–3643.

23. Gerber, M., Ma, J., Dean, K., Eisenberg, J. C., and Shilatifard, A. (2001) EMBO J. 20, 6104–6114.

24. Eisenberg, J. C., Ma, J., Gerber, M. A., Christiansen, A., Kennison, J. A., and Shilatifard, A. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 8984–8989.

25. Stolzmann, L. C., Eisenmann, D. M., and Arndt, K. M. (1997) Mol. Cell. Biol. 17, 4490–4500.
