Reversible Phosphorylation of the C-terminal Domain of RNA Polymerase II*

Michael E. Dahmus

From the Section of Molecular and Cellular Biology, Division of Biological Sciences, University of California, Davis, California 95616

RNA polymerase (RNAP) II is responsible for the synthesis of pre-mRNA in eukaryotic cells. The subunit structure of RNAP II is similar to that of other RNAPs in that it is comprised of two large subunits with a molecular weight in excess of 100,000 and a collection of smaller subunits (1, 2). However, the largest subunit of RNAP II is unique in that it contains an unusual domain at its C terminus comprised of tandem repeats of the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser (3). The consensus repeat has been conserved in evolution although the number of repeats present varies in different species. RNA polymerase II of mammalian cells contains 26 repeat units; the consensus repeat, and yeast contains 26–27 copies, whereas other eukaryotes contain an intermediate number of repeats. Although this C-terminal domain (CTD) plays an essential role in transcription catalyzed by RNAP II, it is absent from RNAPs I and III. The CTD of yeast and mammalian RNAP II was first reported about 10 years ago and is shown in Fig. 1 (4, 5). This domain has provided a focal point for the analysis of RNAP II structure-function relationships. Although our understanding of the CTD has increased considerably in the ensuing 10 years, its precise role in transcription remains to be established.

Temporal Relationship between Phosphorylation of CTD and Progression of RNAP II Through Transcription Cycle

Apart from the extensive repetition of the consensus repeat, the CTD is unusual in that it is heavily phosphorylated at a specific phase of the transcription cycle (3, 6). RNAP II containing an unmodified CTD is referred to as RNAP II A, whereas RNAP II containing a hyperphosphorylated CTD is referred to as RNAP II O. The largest subunit of RNAPs II A and II O is designated IIA and IIO, respectively. Although it has not been possible to map or quantify the number of sites phosphorylated in vivo, the number appears to be in excess of 50 (7). Serine is the predominant site of phosphorylation with a low level of phosphorylation on threonine and tyrosine (6, 8, 9).

RNAPs IIA and IIO have distinct roles in the transcription cycle. It is now generally accepted that RNAP II containing an unphosphorylated CTD, namely RNAP II A, assembles into a preinitiation complex on the promoter (10–13). Presumably, protein-protein interactions mediated by the unphosphorylated CTD play a role in the positioning of RNAP II at the start site of transcription. Phosphorylation of the CTD is catalyzed by a CTD kinase that stably associates with the preinitiation complex. Transcript elongation is catalyzed by RNAP IIO (6, 14, 15). Therefore, phosphorylation of the CTD accompanies the transition of RNAP II from a preinitiation complex to a stable elongation complex. Although CTD phosphorylation is temporally correlated with promoter clearance and thought to be a prerequisite to the formation of a stable elongation complex, the precise role of CTD phosphorylation remains obscure.

The idea that phosphorylation of the CTD at multiple sites serves to disrupt interactions between the unmodified CTD and proteins necessary for the formation of a stable preinitiation complex remains an attractive possibility. Upon completion of the transcript, RNAP IIO must be dephosphorylated by CTD phosphatase to re-generate RNAP IIA and complete the cycle. The transcription cycle of RNAP II is schematically represented in Fig. 2.

Role of CTD in Preinitiation Complex Formation and in Mediating Activity of Transcriptional Regulators

To understand the involvement of the CTD in assembly of the preinitiation complex, it is necessary to consider the complex array of proteins that participate in the early phase of transcription. Assembly of a preinitiation complex and the initiation of transcription are dependent on the presence of multiple general transcription factors (GTFs) (16, 17). These factors, designated TFIIA, -II B, -II D, -II E, -II F, and -II H, in addition to RNAP II are sufficient to support a basal level of transcription from a variety of eukaryotic promoters. Although this complement of factors is sufficient to support transcription in reconstituted systems, additional proteins appear to be involved in vivo. An important clue that additional proteins are required came from the analysis of second site mutations that suppress the conditional phenotype of CTD truncation mutants (18, 19). These SRB genes (suppressors of RNA polymerase II) encode proteins that are involved in transcription and interact with RNAP II (20).

Recently, a holoenzyme form of RNAP II has been described in yeast that is comprised of the core enzyme, the GTFs TFII B, TFII F, TFII H, the products of all nine SRB genes, GAL11, SUG1, and components of the SWI/SNF complex (19–22). The SWI/SNF complex is a general transcriptional regulator involved in chromatin remodeling. A second form of the yeast holoenzyme has been described that also includes the global transcriptional regulators Sin4 and Rgr1 but apparently lacks TFII B, TFII H, and the SWI/SNF complex (23, 24). The difference in holoenzyme composition may arise from different methods of purification that lead to the loss of specific components of the holoenzyme or from differences in either growth conditions or strains of yeast. Alternatively, multiple forms of the holoenzyme may exist. The holoenzyme differs functionally from core RNAP II in that it is responsive to transcriptional regulators in vitro assays. The multiprotein complex containing SRBs and certain GTFs is stable in the absence of RNAP II and has been termed the mediator (19, 24). The mediator appears to associate with core RNAP II via direct interactions with the CTD (24). This observation is consistent with early results, which indicate that the CTD plays an essential role in mediating the response to various transcriptional regulators (25–27). The mammalian holoenzyme, although less well characterized, is reported to contain the GTFs, TFII E, TFII F, and TFII H, in addition to SRB homologues and proteins involved in DNA repair (56). The human RNAP II holoenzyme is comprised of approximately 80 polypeptides, only some of which have been identified.

The discovery of holo-RNAP II has caused a reconsideration of how preinitiation complexes might assemble on the promoter. Analysis of transcription in reconstituted reactions indicates that preinitiation complexes can assemble by the sequential and ordered association of individual GTFs and RNAP II with the promoter (see Fig. 2A). Alternatively, a macromolecular complex containing multiple GTFs, SRBs, and core RNAP II can assemble independent of the promoter and bind directly to DNA (see Fig. 2B). It is not yet possible to distinguish which reaction scheme more closely resembles how transcription complexes form in vivo. However, in either case, the CTD likely plays a critical role by mediating the interaction of core RNAP II with the factors necessary for preinitiation complex assembly and response to transcriptional regulators. Since neither RNAP I nor RNAP III must integrate the input from such a diverse array of regulatory proteins, the involvement of the CTD in mediating the input from multiple regulators could in part account for the fact that only RNAP II contains a CTD.

*This minireview will be reprinted in the 1996 Minireview Compendium, which will be available in December, 1996.

The abbreviations used are: RNAP, RNA polymerase; CTD, C-terminal domain; GTF, general transcription factor; SRB, suppressor of RNAP II.
Minireview: C-terminal Domain of RNA Polymerase II

Identification of Proteins That Interact with the Unphosphorylated CTD

Critical to our understanding of CTD function is an identification of proteins that make direct contact with the CTD and an understanding of how these interactions are influenced by phosphorylation of the CTD. The fact that mutations in SRB genes can restore wild-type phenotype to cells containing CTD truncations indicates that both the CTD and SRBs are involved in the same functional process. Furthermore, the mediator, which is thought to interact directly with CTD-less RNA polymerase II (RNAP II) in a non-CTD-dependent manner, does not assemble into a preinitiation complex on the dihydrofolate reductase promoter. The requirement for the CTD appears to correlate with the absence of a TATA element and may reflect a fundamental difference in the way transcription complexes assemble on different promoters (28).

Identification of Proteins That Interact with the Unphosphorylated CTD

Critical to our understanding of CTD function is an identification of proteins that make direct contact with the CTD and an understanding of how these interactions are influenced by phosphorylation of the CTD. The fact that mutations in SRB genes can restore wild-type phenotype to cells containing CTD truncations indicates that both the CTD and SRBs are involved in the same functional process. Furthermore, the mediator, which is thought to interact directly with CTD-less RNA polymerase II (RNAP II) in a non-CTD-dependent manner, does not assemble into a preinitiation complex on the dihydrofolate reductase promoter. The requirement for the CTD appears to correlate with the absence of a TATA element and may reflect a fundamental difference in the way transcription complexes assemble on different promoters (28).

Role of CTD Phosphorylation in Establishment of Elongation-competent Transcription Complex

The finding that phosphorylation of the CTD prevents RNAP II from assembling into a preinitiation complex and that there is a temporal relationship between promoter clearance and CTD phosphorylation led to the idea that phosphorylation is the trigger that releases RNAP II from the initiated complex. However, it is now clear that transcription from at least some promoters in defined in vitro systems is not dependent on CTD phosphorylation (31, 32). In less defined systems, transcription appears to be dependent on CTD kinase activity suggesting that CTD phosphorylation may be obligatory in vivo (33). Experiments have not been reported that would distinguish between a requirement based on the physical release of RNAP II from the initiated complex and a requirement for a phosphorylated CTD to establish a stable elongation complex. For example, in the latter case the highly phosphorylated CTD may destabilize nucleosomes, thereby facilitating the progression of RNAP II along the DNA template. Therefore, the possibility exists that CTD phosphorylation plays no direct role in the initiation process but is temporally correlated with initiation because it is essential for the formation of a competent elongation complex.

This idea is consistent with the finding that transcription complexes paused near the transcriptional start site on a number of Drosophila genes contain RNAP II A (34). The induction of transcription and the release of RNAP II from the paused complex correlate with phosphorylation of the CTD. One possibility is that RNAP II is still tethered to promoters associated with the promoter by recruited CTD and phosphorylation triggers a CTD-dependent process but is temporally correlated with initiation because it is essential for the formation of a competent elongation complex.

This idea is consistent with the finding that transcription complexes paused near the transcriptional start site on a number of Drosophila genes contain RNAP II A (34). The induction of transcription and the release of RNAP II from the paused complex correlate with phosphorylation of the CTD. One possibility is that RNAP II is still tethered to promoters associated with the promoter by recruited CTD and phosphorylation triggers a CTD-dependent process but is temporally correlated with initiation because it is essential for the formation of a competent elongation complex.

The phosphorylation of the CTD may play a role in transcription-coupled nucleotide excision repair. Of special interest is the association of multiple DNA repair activities with CTD kinase in the general transcription factor TFIIH (38-41). The observation that TFIIE and TFIIH interact with the unphosphorylated CTD and that TFIIH interacts directly with TFIIH provides a mechanism for the recruitment of TFIIH to RNA polymerase II at multiple steps in the transcription cycle. The CTD is phosphorylated at the time of transcript initiation by a CTD kinase that stably interacts with the preinitiation complex, most likely TFIIH. The phosphorylated CTD becomes incorporated into the CTD and may be a CTD kinase. Alternatively, RNAP II may be a CTD kinase that facilitates the formation of an initiation complex.

Role of CTD Phosphorylation in Establishment of Elongation-competent Transcription Complex

The finding that phosphorylation of the CTD prevents RNAP II from assembling into a preinitiation complex and that there is a temporal relationship between promoter clearance and CTD phos-
several observations. CTD kinase is intrinsic to TFIIH, a factor that appears to be involved in promoter clearance and hence functions in the transcription cycle at the time of CTD phosphorylation. The CTD kinase associated with TFIIH has recently been identified as the cyclin-dependent kinase (Cdk) MO15/Cdk-7 in vertebrates (46, 47) and KIN28 in yeast (38). KIN28 is required for RNA synthesis, and RNAP II phosphorylation is dramatically reduced at the restrictive temperature in a kin28-ts mutant (48). TFIIH from yeast has been fractionated into two forms, one involved in transcription and one in nucleotide excision repair (41, 45, 49). The TFIIH core is comprised of five subunits including RAD3, TFB1, and SSL1. The form that functions in transcription, designated holo-TFIIH, consists of the core in association withSSL2 and the two subunits of the TFIIH-associated CTD kinase, designated TFIIK (50).

A second CTD kinase known to assemble into the preinitiation complex is a cyclin-dependent kinase comprised of SRB10 and SRB11 (51). Although the in vitro phosphorylation of holo-RNAP II containing a srb10 mutant enzyme is reduced greater than 10-fold, the in vitro transcriptional activity of the mutant enzyme is unchanged. Nevertheless, the finding that the SRB10/11 kinase is essential for transcriptional activation by galactose in yeast suggests that this kinase plays a role in transcriptional regulation in vivo.

Multiple CTD kinases appear to be involved in the phosphorylation of RNAP II in vivo. This idea is supported by the observation that disruption of the largest subunit of a yeast CTD kinase, designated CTK1 and distinct from KIN28, results in a diminished level of RNAP II phosphorylation (52, 53). Therefore, RNAP II phosphorylation is diminished by a disruption in the activity of either KIN28 or CTK1. The possibility that certain putative CTD kinases function in vivo to regulate the activity of other protein kinases that phosphorylate the CTD cannot be excluded. Finally, the observation that the CTD can be phosphorylated on tyrosine suggests that multiple CTD kinases function in vivo (8). The physiological significance of multiple CTD kinases is not known. One possibility is that phosphorylation of the CTD plays an essential role in transcription, and redundancy has been built into the enzymes that catalyze this reaction. It is also possible that different promoters utilize different protein kinases and/or different protein kinases function at specific times in the transcription cycle. Finally, the observation that a unique form of RNAP II appears to be recruited to discrete nuclear domains when transcription is inhibited suggests that a specific CTD kinase(s) may influence the subnuclear localization of RNAP II (54).

A CTD phosphatase has been purified from a HeLa cell transcription extract and appears to selectively dephosphorylate the CTD of RNAP IIO (43, 44). The regulation of CTD phosphatase activity is complex and appears to involve an interaction of CTD phosphatase with a docking site on RNAP II that is distinct from the CTD (44). Furthermore, the activity of CTD phosphatase is stimulated by TFIIF, and the stimulatory activity of TFIIF is inhibited by TFIIB. These properties of CTD phosphatase are, therefore, consistent with the idea that it functions to dephosphorylate RNAP IIO upon completion of a transcript, thereby regenerating RNAP IIA for preinitiation complex formation (see Fig. 2). TFIIF is known to directly interact with RNAP II and to play a role in the recruitment of RNAP II to the preinitiation complex. Accordingly, the association of TFIIF with RNAP II upon completion of the transcript would stimulate the dephosphorylation reaction (see Fig. 2). Finally, TFIIB may suppress the stimulatory activity of TFIIF in the preinitiation complex, thereby preventing a futile cycle of CTD phosphorylation-dephosphorylation. Regulating access to the docking site on RNAP II by which CTD phosphatase gains access to the CTD may be important in regulating the de-
phosphorylation of RNAP II during the elongation phase of transcription.

Potential Regulatory Significance of CTD Phosphorylation

Since RNAP IIA and IIO have distinct roles in the transcription cycle, CTD kinases and CTD phosphatase can act as positive or negative regulators of transcription depending on the point in the transcription cycle at which they function. For example, phosphorylation of the CTD concomitant with transcript initiation might stimulate transcription whereas phosphorylation of free RNAP II would reduce the amount of RNAP IIA available for recruitment to the promoter and hence inhibit transcription. Conversely, CTD phosphatase that dephosphorylates RNAP II in the initiated or elongation complex may well inhibit transcription whereas dephosphorylation of RNAP IIO upon completion of the transcript would stimulate transcription. A major challenge is to not only enumerate the CTD kinases and phosphatases that modulate the level of RNAP II phosphorylation in vivo but to understand how these enzymes are regulated and the consequences they have on the activity of RNAP II at discrete steps in the transcription cycle.

Summary and Perspectives

The CTD of RNAP II is unusual with respect to both the high level of repetition of the consensus repeat and its high level of phosphorylation. The fact that each round of transcription is associated with the reversible phosphorylation of the CTD is consistent with the idea that RNAP IIA and IIO have distinct roles in the transcription cycle. Indeed, RNAP IIA has been shown to selectively assemble into preinitiation complexes on a number of promoters, whereas transcript elongation has been shown to be catalyzed by RNAP IIO. The precise role played by the unphosphorylated CTD during initiation and the phosphorylated CTD during elongation has not been established. The extended nature of the CTD and the multiplicity of proteins that appear to interact with core RNAP II via the CTD suggest that a primary nature of the CTD and the multiplicity of proteins that appear to interact with core RNAP II via the CTD suggest that a primary

Acknowledgments—I thank Shari Blewins, Delphine Gerber, Brett Kaiser, Alain Lehman, Nick Marshall, and Tricia Marshall for critical reading of the manuscript.

REFERENCES

1. Young, R. A. (1991) Annu. Rev. Biochem. 60, 689–715
2. Sawadogo, M., and Sentenac, A. (1990) Annu. Rev. Biochem. 59, 711–754
3. Dahmus, M. E. (1984) Proc. Nucl. Acids Res. Mol. Biol. 10, 143–179
4. Allison, L. A., Moyle, M., Shakes, M., and Ingle, C. J. (1985) Cell 22, 599–610
5. Corden, J. L., Cadena, D. L., Ahearn, J. M., and Dahmus, M. E. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 7934–7938
6. Cadena, D. L., and Dahmus, M. E. (1987) J. Biol. Chem. 262, 12246–12247
7. Payne, J. M., and Dahmus, M. E. (1993) J. Biol. Chem. 268, 80–87
8. Baskaran, K., Dahmus, M. E., and Wang, Y. Y. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 11617–11621
9. Zhang, J., and Corden, J. L. (1991) J. Biol. Chem. 266, 2290–2296
10. Lu, H., Flores, O., Weinmann, R., and Reinberg, D. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 10004–10008
11. Laybourn, P. J., and Dahmus, M. E. (1990) J. Biol. Chem. 265, 13165–13173
12. Chernev, J. D., Stephens, J. H., and Dahmus, M. E. (1992) J. Biol. Chem. 267, 10500–10506
13. Kang, M. E., and Dahmus, M. E. (1993) J. Biol. Chem. 268, 25033–25040
14. Bartholomew, B., Dahmus, M. E., and Meares, C. F. (1986) J. Biol. Chem. 261, 14226–14231
15. Weeks, J. R., Hardin, S. E., Shen, J., Lee, J. M., and Greenleaf, A. L. (1993) Genes & Dev. 7, 2329–2344
16. Conaway, R. C., and Conaway, J. W. (1993) Annu. Rev. Biochem. 62, 161–190
17. Zawel, L., and Reinbu, J. D. (1993) J. Biol. Chem. 268, 533–543
18. Kolesa, J. S., Buratowski, S., Nonet, M., and Young, R. A. (1992) Cell 69, 883–894
19. Thompson, C. M., Kolesa, A. J., Chao, D. M., and Young, R. A. (1993) Cell 73, 1361–1375
20. Hengartner, C. J., Thompson, C. M., Zhang, J., Chao, D. M., Liao, S. M., Kolesa, A. J., Okamura, S., and Young, R. A. (1995) Genes & Dev. 9, 897–910
21. Kolesa, A. J., and Young, R. A. (1994) Nature 368, 466–469
22. Wilson, C. J., Chao, D. M., Imbalzano, A. N., Schnitzler, G. R., Kingston, R. E., and Young, R. A. (1996) Cell 84, 235–244
23. Li, Y., Bjorklund, S., J. Iang, Y.W., Kim, Y. J., Lane, W. S., Stillman, D. J., and Kornberg, R. D. (1994) Proc. Natl. Acad. Sci. U. S. A. 92, 10864–10868
24. Kim, Y. J., Bjorklund, S., Li, Y., Sayre, M. H., and Kornberg, R. D. (1994) Cell 77, 599–608
25. Allison, L. A., and Ingle, C. J. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 2794–2798
26. Scolf, C., Chao, D., Loepke, J., Hirsch, J. P., Henry, S., and Young, R. A. (1990) Nature 347, 491–494
27. Liao, S. M., Taylor, J. C., Kingston, R. E., and Young, R. A. (1991) Genes & Dev. 5, 2433–2440
28. Bueymer, A. B., Strasheim, L. A., McMahon, S. L., and Farnham, P. J. (1995) J. Biol. Chem. 270, 6798–6807
29. Usheva, A., Maldonado, E., Goldring, A., Lu, H., Houbavi, C., Reinberg, D., and Aloni, Y. (1992) Cell 69, 871–881
30. Kang, M. E., and Dahmus, M. E. (1995) J. Biol. Chem. 270, 23390–23397
31. Serizawa, H., Conaway, J. W., and Conaway, R. C. (1993) Nature 363, 371–374
32. Makela, T. P., Parvin, J. D., Kinn, H., Huo, L., Buschel, D. A., Donahue, T. F., Friedman, E. C., and Kornberg, R. D. (1995) Cell 80, 21–28
33. Maxon, M. E., Goodrich, A. J., and Tjian, R. (1994) Genes & Dev. 8, 515–524
34. Chambers, R. S., and Dahmus, M. E. (1994) J. Biol. Chem. 269, 26243–26248
35. Chambers, R. S., Wang, B. Q., Burdon, Z. F., and Dahmus, M. E. (1995) J. Biol. Chem. 270, 14962–14969
36. Roy, R., Adamczewski, J. P., Seroz, T., Vermeulen, W., Tassan, J. P., Schaeffer, F., Nong, E. A., Hoejmoekers, J. H. J., and Young, R. A. (1994) Cell 79, 1093–1101
37. Schaeffer, L., Roy, R., Humber, S., Monozlal, V., Vermeulen, W., Hoejmoekers, J. H. J., Chamber, P., and Ely, G. J. (1993) Science 260, 58–63
38. Drapkin, R., Sancar, A., and Reinberg, D. (1994) Cell 77, 9–12
39. Svejstrup, J. G., Wang, Z., Feaver, W. J., Wu, X., Buschel, D. A., Donahue, T. F., Friedman, E. C., and Kornberg, R. D. (1995) Cell 80, 21–28
40. Maxon, M. E., Goodrich, A. J., and Tjian, R. (1994) Genes & Dev. 8, 515–524
41. Chambers, R. S., and Dahmus, M. E. (1994) J. Biol. Chem. 269, 26243–26248
42. Chambers, R. S., Wang, B. Q., Burdon, Z. F., and Dahmus, M. E. (1995) J. Biol. Chem. 270, 14962–14969