Volume 1 · Number 1 · April 1991
GENE EXPRESSION

MINI REVIEW

Tailored tails and transcription initiation: the carboxyl terminal domain of RNA polymerase II

David M. Chao and Richard A. Young

Whitehead Institute for Biomedical Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts

The largest subunit of RNA polymerase II contains a curious structure at its carboxyl terminus: a series of tandem heptapeptide repeats of the consensus sequence, pro-thr-ser-pro-ser-tyr-ser. This conserved, essential, and phosphorylated structure is called the carboxyl terminal domain (CTD, a. k. a. "the tail"). The CTD has been implicated in the response of the transcriptional apparatus to signals from transcriptional activators at certain promoters. We review here the experimental background to current thinking on the function of this unusual tail.

The largest subunit of every RNA polymerase II examined so far contains a CTD (Falkenburg et al., 1987; Allison et al., 1985; Nonet et al., 1987; Bird and Riddle, 1989; Allison et al., 1988; Zehring et al., 1988; Corden et al., 1985; Evers et al., 1989a and 1989b; Smith et al., 1989). In contrast, RNA polymerases I and III and bacterial and viral RNA polymerases do not contain CTDs (Allison et al., 1985; Memet et al., 1988; Puhler et al., 1989; Ovchinnikov et al., 1982; Broyles and Moss, 1986). The consensus sequence of the heptapeptide repeat is conserved among the CTDs of all species studied so far except for the kinetoplastids, whose CTDs are similar in amino acid composition to those of other organisms but lack a discernable repeat (Evers et al., 1989a; Smith et al., 1989; Evers et al., 1989b). The CTDs of various species vary both in their length and in their overall conservation. In general, more complex species have more repeats in their CTDs. For example, the yeast RNA polymerase II tail has twenty-six or twenty-seven repeats, while the murine CTD has fifty-two (Allison et al., 1985; Nonet et al., 1987; Corden et al., 1985).

Such a remarkable degree of conservation implies that the RNA polymerase II CTD performs a function of vital importance. Indeed, approximately half of the CTD is required for the viability of yeast, mouse, and Drosophila cells (Nonet et al., 1987; Allison et al., 1988; Bartolemei et al., 1988; Zehring et al., 1988). In yeast, deletions leaving less than ten of the twenty-seven repeats cause lethality, while deletions leaving between eleven and thirteen repeats result in cold-sensitivity, temperature-sensitivity, and inositol auxotrophy (Nonet et al., 1987; Nonet and Young, 1989; Scafe et al., 1990a). The severity of these growth defects increases with progressive truncation of the CTD, indicating that truncation of the CTD progressively decreases the domain's functional efficiency.

In fact, more detailed analysis of these yeast mutants has revealed that truncation of the CTD progressively decreases the ability of the transcriptional apparatus to respond to signals from certain upstream activating sequences (UASs), the yeast version of enhancers (Scafe et al., 1990b). In a CTD truncation mutant containing eleven repeats, transcription of *INO1*, a gene central to inositol biosynthesis, is induced to only ten percent of the wild-type level. In the same mutant, transcription of another gene,

Correspondence: Richard A. Young, Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142
Tel (617)258-5218 Fax (617)258-9872
 1052-2166/91/0101/001-04$2.00

GAL10, is induced to only forty percent of the wild-type level. Transcription of *HIS4* is induced to nearly wild-type levels in the mutant. The relative transcriptional defects observed with the intact genes precisely parallel the relative defects observed with hybrid promoter constructs fusing the *INO1*, *GAL10* and *HIS4* UASs to the *CYC1* TATA element and start site. Furthermore, as the length of the RNA polymerase II CTD is reduced from twenty-seven to eleven repeats, the induced level of expression from the *INO1* UAS and *GAL10* UAS constructs drops off gradually at first and then precipitously at a threshold number of repeats (Scafe et al., 1990b). In addition, truncation of the CTD magnifies transcriptional defects in a variety of Gal4 activation mutants, while extension of the CTD suppresses these defects (Allison and Ingles, 1989). These data implicate the CTD in the response of the transcriptional apparatus to signals from a subset of UASs.

Transcript elongation does not appear to be affected by truncation of the CTD. Drosophila RNA polymerase II from which the CTD has been proteolytically removed is able to perform non-specific initiation and elongation (Zehring et al., 1988). In addition, DNA downstream of the promoter does not appear to influence UAS-dependent defects observed with truncation of the RNA polymerase II CTD in yeast (Scafe et al., 1990).

The notion that the CTD is involved in the regulation of transcription initiation at a subset of promoters is consistent with the results of in vitro transcription experiments. Drosophila RNA polymerase II lacking the CTD is able to initiate transcription in vitro from the Drosophila actin 5C promoter at levels similar to those obtained with the wild-type enzyme (Zehring et al., 1988). Similarly, in HeLa cell extracts in which the endogenous RNA polymerase II is immunologically inactivated, calf thymus RNA polymerase II lacking the CTD is fully capable of accurately initiating transcription from the adenovirus major late promoter (Thompson et al., 1989). However, in the same system, RNA polymerase II lacking the CTD is unable to accurately initiate transcription from the murine dihydrofolate reductase promoter (Thompson et al., 1989). At promoters that appear to be CTD-independent in vitro, RNA polymerase II lacking a CTD responds normally to stimulation by Sp1 (Zehring and Greenleaf, 1990) and the Major Late Transcription Factor (Buratowski and Sharp, 1990).

Why does efficient transcriptional initiation at only a subset of promoters appear to require an intact CTD? One possibility is that multiple mechanisms exist for transcriptional activation and that only a subset requires CTD function. Another possibility is that the CTD has a role at most or all promoters and that various promoters differ in their ability to compensate for removal of the CTD.

Does the CTD play a central role in a mechanism of transcriptional activation or an accessory role in a mechanism that can function at a suboptimal level without the CTD? For instance, at some promoters the CTD may be an essential intermediate in the transmission of an activation signal from upstream factors to RNA polymerase. Alternatively, the CTD may increase the efficiency of a step in transcriptional activation, such as an interaction between upstream factors and general factors bound at the TATA element. The available evidence permits only speculation on this important issue.

Two types of molecular interactions have been proposed for the CTD during transcription initiation. The first postulates direct or indirect interactions between the CTD and transcription factors. The second posits interactions between the CTD and promoter DNA.

The CTD may influence activation by interacting either directly or indirectly with transcription factors (Allison et al., 1985; Corden et al., 1985; Sigler, 1988). For instance, it has been proposed that the acidic activation domains of transcription factors interact with the abundant hydroxyl groups of the CTD (Sigler, 1988). It is also possible that activation may occur through cofactors or adaptors (Pugh and Tjian, 1990; Berger et al., 1990; Kelleher et al., 1990; Lewin, 1990) which themselves may interact with the CTD.

The CTD might also exert its effect on transcription by binding to DNA and displacing histones or other DNA-binding proteins that influence gene regulation (Corden et al., 1985; Suzuki, 1990). Synthetic peptides containing CTD sequences can bind to supercoiled DNA in vitro, an observation providing some biochemical support for this class of models (Suzuki, 1990). However, the physiological relevance of this observation is not yet clear.

The RNA polymerase II CTD is rich in pro-

lines and hydrophilic amino acids, a characteristic leading some investigators to propose that the CTD projects from the main body of the enzyme as a rod-like structure (Sigler, 1988; Matsushima et al., 1990; Corden, 1990; Suzuki, 1990). Because the CTD appears to be involved in the response to signals from upstream factors, the notion that the CTD is an antenna-like structure is intriguing. Direct evidence that the CTD projects from the body of polymerase and adopts a rod-like conformation is not yet available.

A portion of the RNA polymerase II molecules in cells contain highly phosphorylated CTDs (Cadena and Dahmus, 1987; Kolodziej et al., 1990). Phosphorylation may serve to modulate the CTD's activity. For example, phosphorylation might affect the CTD's interaction with transcription factors or DNA. The observation that the unphosphorylated form of the CTD becomes phosphorylated during a step in initiation suggests that the CTD's phosphorylation switches RNA polymerase II from a resting to an active state (Bartholomew, 1986; Payne et al., 1989; and Laybourn and Dahmus, 1990). The recent purification of kinases that specifically phosphorylate the CTD in vitro should facilitate studies to determine the physiological role of the CTD's phosphorylation (Lee and Greenleaf, 1989; Cisek and Corden, 1989).

Precisely how the CTD is involved in transcriptional initiation is still unclear but is under intense study. Identifying proteins that bind to the CTD and influence its activity is likely to provide significant clues to the molecular mechanisms of CTD-dependent transcriptional activation. Given the complexity of transcriptional regulation and the peculiarity of the CTD, it would not be surprising to find a colorful cast of characters involved in this process.

Acknowledgments

We thank Anthony Koleske and Craig Thompson for helpful discussions. Supported by NIH grant GM34365 and a Burroughs Wellcome Scholarship.

The costs of publishing this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC Section 1734 solely to indicate this fact.

References

L. A. Allison, M. Moyle, M. Shales, and C. J. Ingeles (1985), Cell 42, 599–610.

L. A. Allison, J. K. C. Wong, V. D. Fitzpatrick, M. Moyle, and C. J. Ingles (1988), Mol Cell Biol 8, 321–329.

L. A. Allison and C. J. Ingles (1989), Proc Natl Acad Sci USA 86, 2794–2798.

M. S. Bartolomei, N. F. Halden, C. R. Cullen, and J. F. Corden (1988), Mol Cell Biol 8, 330–339.

B. Bartholomew, M. E. Dahmus, and C. F. Meares (1986), J Biol Chem 261, 14226–14231.

S. L. Berger, W. D. Cress, A. Cress, S. J. Triezenberg, and L. Guarente (1990), Cell 61, 1199–1208.

D. M. Bird and D. L. Riddle (1989), Mol Cell Biol 9, 4119–4130.

S. Buratowski and P. A. Sharp (1990), Mol Cell Biol 10, 5562–5564.

S. S. Broyles and B. Moss (1986), Proc Natl Acad Sci USA 83, 3141–3145.

L. Cisek and J. Corden (1990), Nature 339, 679–684.

D. Cadena and M. Dahmus (1987), J Biol Chem 262, 12468–12474.

J. L. Corden, D. L. Cadena, J. M. Ahaern, and M. E. Dahmus (1985), Proc Natl Acad Sci USA 82, 7934–7938.

J. Corden (1990), Trends Biochem Sci 15, 383–387.

R. Evers, A. Hammer, J. Kock, W. Jess, P. Borst, S. Memet, and A. W. C. A. Cornelissen (1989a), Cell 56, 585–597.

R. Evers, A. Hammer, and A. W. C. A. Cornelissen (1989b), Nucl Acids Res 17, 3404–3413.

D. Falkenburg, B. Dworniczak, D. Faust, and E. K. F. Bautz (1987), J Mol Biol 195, 929–937.

R. J. Kelleher, P. M. Flanagan, and R. D. Kornberg (1990), Cell 61, 1209–1215.

P. A. Koledziej, N. A. Woychik, S. M. Liao, and R. A. Young (1990), Mol Cell Biol 10, 1915–1920.

P. J. Laybourn and M. E. Dahmus (1990), J Biol Chem 265, 13165–13173.

J. M. Lee and A. L. Greenleaf (1989), Proc Natl Acad Sci USA 86, 3624–3628.

B. Lewin (1990), Cell 61, 1161–1164.

N. Matsushima, C. E. Creutz, and R. H. Kretsinger (1990), Proteins: Structure, Function, and Genetics 7, 125–155.

S. Memet, M. Gouy, C. Marck, A. Sentenac, and J. M. Buhler (1988), J Biol Chem 263, 2830–2839.

M. Moyle, J. S. Lee, W. F. Anderson, and C. J. Ingles (1989), Mol Cell Biol 9, 5750–5753.

M. Nonet, D. Sweetser, and R. A. Young (1987), Cell 50, 909–915.

M. Nonet and R. A. Young (1989), Genetics 123, 715–724.

Y. A. Ovchinnikov, G. S. Monastyrskaya, V. V. Gubanov, S. O. Guryev, I. S. Salomatina, T. M. Shuvaea, V. M. Lipkin, and E. D. Sverdlov (1982), Nucl Acids Res 10, 4035–4044.

J. M. Payne, P. J. Laybourn, and M. E. Dahmus (1989), J Biol Chem 263, 19621–19629.

G. Puhler, H. Leffers, F. Gropl, P. Palm, H. P. Klenk, F. Lottspeich, R. Garret, and W. Zillig (1989), Proc

Natl Acad Sci USA 86, 4569–4573.
B. F. Pugh and R. Tjian (1990), Cell 61, 1187–1197.
G. S. Roeder, C. Beard, M. Smith, and S. Keranen (1985), Mol Cell Biol 5, 1543–1553.
C. Scafe, C. Martin, M. Nonet, S. Podos, S. Okamura, and R. A. Young (1990a), Mol Cell Biol 10, 1270–1275.
C. Scafe, D. Chao, J. Lopes, J. Hirsch, S. Henry, and R. A. Young (1990b), Nature 347, 491–494.
P. Sigler (1988), Nature 333, 210–212.
J. L. Smith, J. R. Levin, C. J. Ingles, and N. Agabian (1989), Cell 56, 815–827.
P. W. Sternberg, M. J. Stern, E. Clark, and I. Herskowitz (1987), Cell 48, 567–577.
M. Suzuki (1990), Nature 344, 562–565.
N. E. Thompson, T. H. Steinberg, D. B. Aronson, R. R. Burgess (1989), J Biol Chem 264, 11511–11520.
W. A. Zehring, J. M. Lee, J. R. Weeks, R. S. Jokerst, and A. L. Greenleaf (1988), Proc Natl Acad Sci USA 85, 3698–3702.
W. A. Zhering and A. L. Greenleaf (1990), J Biol Chem 265, 8351–8353.