Commentary

A positive addition to a negative tail's tale

Arno L. Greenleaf

Biochemistry Department, Duke University Medical Center, Durham, NC 27710

The C-terminal repeat domain (CTD) of the largest subunit of RNA polymerase II is composed of multiple repeats of the consensus heptamer sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser (YSPTSPS in single-letter code) (for reviews, see refs. 1–3); thus five of the seven residues of each repeat are potential sites of phosphorylation. Indeed, hyperphosphorylation of this domain has been documented in organisms from yeast to human (one diagnostic for hyperphosphorylation of the CTD is a marked mobility change of the largest subunit in SDS gels; unphosphorylated subunit "IIa" migrates with an apparent molecular mass of ≈215 kDa, whereas hyperphosphorylated subunit "IIo" migrates with an apparent molecular mass of ≈240 kDa). With CTD lengths ranging from 26 repeats in yeast to 52 repeats in mammals, one could imagine that nuclear protein kinases encountering this domain might think they had found Paradise. Until now, however, only serine/threonine kinases would have been thought to have this experience. This perception changes with the publication of the article by Baskaran, Dahmus, and Wang (4) in this issue, which demonstrates that the CTD is also subject to tyrosine phosphorylation.

Previous work on *in vivo*-labeled mammalian RNA polymerase II had detected CTD-derived phosphoserine and phosphothreonine; no phosphotyrosine had been observed (5, 6). In the current work, extra precautions were taken to preserve any phosphorylated tyrosine (*P*-Tyr) that might be present on the CTD in HeLa nuclei: extract preparation procedures were modified, and high levels of *P*-Tyr phosphatase inhibitors were included. These steps led to detecting *P*-Tyr in digestions of RNA polymerase II largest subunit purified by immunoprecipitation and SDS/gel electrophoresis. Under the conditions used approximately equal amounts of *P*-Tyr and phosphorylated threonine and about three to five times more phosphorylated serine were detected.

Further experiments reported by Baskaran *et al.* (4) demonstrate that *in vitro*, the CTD can be phosphorylated by a tyrosine kinase known to be found in the nucleus, c-Abl. The c-Abl kinase can add up to ≈30 phosphates to tyrosines in the CTD and, as is the case for serine/threonine hyperphosphorylation, this modification produces a shift in SDS/gel mobility. In contrast to c-Abl, a different tyrosine kinase, c-Src, does not phosphorylate the CTD *in vitro*. This additional information forces a significant reevaluation of our ideas about CTD phosphorylation.

One obvious question raised by these findings concerns the identity of the tyrosine kinase(s) that phosphorylates the CTD *in vivo*. The previously known and now reported properties of c-Abl are consistent with its playing such a role in the nucleus, but additional tests will be required to test critically this possibility. Unpublished data cited by Baskaran *et al.* (4) already suggest the involvement of other as-yet-unidentified tyrosine kinase activities because they observed CTD tyrosine phosphorylation in a c-Abl-negative 3T3 cell line (their *Discussion*). It is a bit unsettling that all previous searches for CTD kinase activities, in extracts of fungal, animal, or plant cells, have yielded serine or serine/threonine kinases. Of course, there might be several explanations for this situation, including low abundance of tyrosine kinase activities or inappropriate assay conditions. The current report will certainly stimulate attempts to detect *P*-Tyr in the CTD of RNA polymerase II subunits from different organisms and to identify the responsible kinases. The hint that c-Abl may be involved is exciting and, if borne out, will provide new ideas about mechanisms of *abl* oncogene-mediated transformation.

A basic question, of course, which is still not fully answered by experimental tests, is "What is the function of CTD phosphorylation?" *In vitro* experiments over the last few years have suggested the following scenario: RNA polymerase II with unphosphorylated CTD (RNA polymerase IIA) enters into preinitiation complexes, the CTD interacting with transcription factor TFIID and possibly other factors; concurrent with initiation and/or beginning productive elongation the CTD becomes hyperphosphorylated, such that the elongating transcriptase is RNA polymerase IIO (ref. 7 and the references therein). This scenario leads to the frequent suggestion that CTD phosphorylation mediates "release" of RNA polymerase II from preinitiation or promoter-proximal paused complexes, thereby allowing productive elongation. The presence of CTD kinase activity in preinitiation complexes and the apparent association of CTD kinase activity with transcription factor TFIIH are consistent with this suggestion (refs. 8–10 and the references therein). More recent *in vivo* experiments are also in general agreement with this overall picture (11). It should be pointed out, however, that despite the identification and characterization of a number of different CTD kinases, the identity of the kinase(s) acting on the CTD *in vivo* has not been rigorously established.

All the previously characterized CTD kinase activities, including those found associated with preinitiation complexes or initiation factors, are specific for serine or serine/threonine. If they represent the activity/activities that in the above scenario generate actively elongating RNA polymerase II0, then a prediction would be that elongating RNA polymerase II0 is phosphorylated on serine and threonine but not on tyrosine. This prediction might be tested by using the approach, frequently used by Dahmus and colleagues, of identifying productively elongating RNA polymerase II by UV crosslinking to nascent transcripts (see ref. 12, for example) and coupling this identification with analyses designed to detect *P*-Tyr. The outcome of such experiments will importantly shape further thinking and experimentation. If elongating RNA polymerase II contains *P*-Tyr, it will then be critical to identify when and by which activity the phosphates were added. Do preinitiation complexes also contain CTD tyrosine kinase activity, for example? If so, what is the identity of the kinase?

The question of *P*-Tyr presence on elongating RNA polymerase II also, of course, has implications for thinking about possible functions of the CTD in *post*initiation phases of transcription and the regulation of those functions. Several roles have been suggested for the hyperphosphorylated CTD during transcript elongation. One suggestion is that the hyperphosphorylated CTD represents an example of an acidic polymer that facilitates passage of polymerase through nucleosomes by catalyzing the displacement of histone H2A/H2B dimers (13).

8928/Dec:23/1939

Another suggestion is that the phosphorylated CTD is a docking or attachment site for certain RNA-processing components (14). Whatever the actual functions of the phospho-CTD during elongation, the inventory of activities able to modulate those functions may be much larger than previously thought if it includes tyrosine kinases; the implications of this eventuality would be far-reaching.

Alternatively, if actively elongating RNA polymerase II does not contain *P*-Tyr, very different models would ensue. Could, for example, tyrosine and serine/threonine phosphorylation be mutually exclusive events with opposite functional consequences? Might CTD hyperphosphorylation on tyrosines render RNA polymerase II incapable of initiation and represent a mechanism for shutting down transcription in part or entirely at certain critical points in development or in the cell cycle?

In addition to the above questions and speculations, this current report also invites a reevaluation of previous experiments. For instance, because hyperphosphorylation on tyrosines can apparently cause the RNA polymerase IIa → IIo mobility shift, previous experiments that monitored the ratio of subunit forms migrating as "IIa" and "IIo" as a means to measure CTD phosphorylation state in crude extracts or *in vivo* may have been influenced in unappreciated ways by tyrosine phosphorylation. For example, attempts to observe differences in the IIa/IIo ratio as a function of position in the cell cycle failed to reveal any cell-cycle-dependent changes (ref. 15; S. Hardin and A.L.G., unpublished work). However, had the relative levels of serine/threonine vs. tyrosine phosphorylation actually changed, the qualitatively different phosphorylation state could have been obscured by the relatively constant total amount of slower migrating subunit ("IIo"), which we now realize can be caused by different kinds of hyperphosphorylation (see also ref. 1).

Another earlier experiment showed that in a *Saccharomyces cerevisiae* mutant strain lacking a functional *CTK1* gene, the gene encoding the catalytic subunit of a well-characterized yeast CTD kinase, very little "IIo" subunit could be detected by antibodies directed against the serine/threonine-phosphorylated CTD; in the same strain both IIa and IIo could be detected by antibodies to a non-CTD portion of the largest subunit (16). These apparently contradictory results might conceivably be explained by invoking tyrosine phosphorylation. Namely, if the subunit migrating at the IIo position and detected by the non-CTD antibodies were present because of hyperphosphorylation on tyrosine, it would not have been detected by the anti-phospho-CTD antibodies used because they were raised against the serine/threonine-phosphorylated CTD. This explanation should be subject to relatively easy experimental assessment.

Finally, a recent report revealed that the CTD in mammalian RNA polymerase II can be modified, not only by addition of phosphate groups but also by the addition of O-linked sugars, specifically *O*-GlcNAc (17). In that work CTD glycosylation and phosphorylation were found to be mutually exclusive events. However, because *P*-Tyr may not have been preserved in the preparations of RNA polymerase II analyzed, the relationship between glycosylation and tyrosine phosphorylation is actually not yet clear.

Overall, the unexpected findings presented in the paper by Baskaran *et al.* (4) emphasize how little we really know about the functions of the CTD and about mechanisms and consequences of its phosphorylation. They suggest that in our attempts to understand this unusual domain and its posttranslational modifications we are only at the beginning of a complex epic rather than near the end of a simple tale.

1. Corden, J. L. (1990) *Trends Biochem. Sci.* **15**, 383–387.
2. Young, R. A. (1991) *Annu. Rev. Biochem.* **60**, 689–715.
3. Dahmus, M. E. & Dynan, W. S. (1992) in *Transcriptional Regulation*, eds. Yamamoto, K. & McKnight, S. (Cold Spring Harbor Lab. Press, Plainview, NY), Vol. 1, pp. 109–129.
4. Baskaran, R., Dahmus, M. E. & Wang, J. Y. J. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 11167–11171.
5. Cadena, D. L. & Dahmus, M. E. (1987) *J. Biol. Chem.* **262**, 12468–12474.
6. Zhang, J. & Corden, J. L. (1991) *J. Biol. Chem.* **266**, 2290–2296.
7. Chesnut, J. D., Stephens, J. H. & Dahmus, M. E. (1992) *J. Biol. Chem.* **267**, 10500–10506.
8. Gileadi, O., Feaver, W. J. & Kornberg, R. D. (1992) *Science* **257**, 1389–1392.
9. Lu, H., Zawel, L., Fisher, L., Egly, J. M. & Reinberg, D. (1992) *Nature (London)* **358**, 641–645.
10. Serizawa, H., Conaway, R. C. & Conaway, J. W. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 7476–7480.
11. Weeks, J. R., Hardin, S. E., Shen, J., Lee, J. M. & Greenleaf, A. L. (1993) *Genes Dev.*, in press.
12. Bartholomew, B., Dahmus, M. E. & Meares, C. F. (1986) *J. Biol. Chem.* **261**, 14226–14231.
13. Hansen, J. C. & Ausio, J. (1992) *Trends Biochem. Sci.* **17**, 187–191.
14. Greenleaf, A. L. (1993) *Trends Biochem. Sci.* **18**, 117–119.
15. Kolodziej, P. A., Woychik, N., Liao, S. M. & Young, R. A. (1990) *Mol. Cell. Biol.* **10**, 1915–1920.
16. Lee, J. M. & Greenleaf, A. L. (1991) *Gene Exp.* **1**, 149–167.
17. Kelly, W. G., Dahmus, M. E. & Hart, G. W. (1993) *J. Biol. Chem.* **268**, 10416–10424.