Switching Condensates: The CTD Code Goes Liquid

Bede Portz1,  and James Shorter1, 2, *

Condensates containing RNA polymerase II (Pol II) materialize at sites of active transcription. Young and coworkers now establish that C-terminal domain phosphorylation regulates Pol II partitioning into distinct condensates connected with transcription initiation or splicing. This advance hints that distinct condensates with specialized functional compositions might choreograph distinct stages of transcription.

Liquid–liquid phase separation (LLPS) of RNA polymerase II (Pol II) mediated by its repetitive C-terminal domain (CTD) may explain clustering of Pol II at sites of active transcription [1–4]. Recent work by Young and colleagues now connects CTD-mediated LLPS with the “CTD code” [5,6]. The CTD code specifies how largely unphosphorylated Pol II is initially recruited to gene promoters and how, as transcription progresses, the CTD becomes dynamically phosphorylated and dephosphorylated in specific patterns [7]. Thus, specific patterns of CTD phosphorylation predominate at 5’ and 3’ ends of genes, with each Pol II phosphosioform recruiting specific interactomes. In this way, initiation factors give way to elongation and mRNA splicing factors, and eventually termination factors, as Pol II traverses a gene [8]. Guo et al. now fuse the concepts of CTD transcriptional condensates with decades of biology delineating how the CTD code might orchestrate transcription and cotranscriptional events [5]. The stage appears to now be set for elucidating nuanced transcriptional condensates of different functional compositions that might drive distinct stages of the transcription cycle.

The Pol II CTD can form liquid droplets and partition into hydrogels comprised of protein low-complexity domains with transcription activation functions [1,9]. In both cases, decreased CTD length or CTD phosphorylation reduced partitioning into these dense phases [1,9]. These observations suggested that Pol II is recruited to promoters via LLPS of the CTD and initiation machinery and subsequently released from promoter-associated condensates as a function of CTD phosphorylation, which accrues after transcription initiation [7]. Now, Guo et al. confirm that CTD phosphorylation leads to dissociation of the CTD from mediator condensates and disambiguate that CTD phosphorylation also promotes CTD recruitment to splicing condensates [5]. Thus, the CTD code may govern the CTD interactome as well as the CTD partition coefficient for different functional condensates comprised of distinct components.

Guo et al. reconstituted separate condensates in vitro containing either the CTD and mediator components, or the CTD and splicing factors, and combined these studies with sophisticated imaging of proteins and nascent RNA in cells [5]. In vitro, CTD:mediator condensates exhibited liquid-like behavior [5]. In cells, mediator and Pol II colocalized with nascent RNA from Nanog and Trim28 [5]. Nascent RNA from these genes overlapped with or were immediately adjacent to nuclear speckles (another liquid-like condensate), which contain a host of splicing factors [5]. Notably, unphosphorylated Pol II colocalized more strongly with mediator, while Serine 2 phosphorylated Pol II, a mark for transcription elongation, colocalized more strongly with splicing factors [5]. Proximity of splicing factors with mediator was reduced by drugs that inhibit splicing, whereas colocalization of phosphorylated Pol II with splicing factors was reduced by drugs that inhibit CTD kinases [5].

The differential colocalization of Pol II depending on CTD phosphorylation, coupled with the proximity of actively transcribed loci to clusters of splicing factors and speckle components, suggested that CTD phosphorylation might enable Pol II to migrate from initiation condensates to splicing condensates [5]. In support of this hypothesis, Guo et al. reconstituted two distinct CTD condensates. The first contained purified mediator, into which the unphosphorylated CTD partitioned more strongly than CTDs phosphorylated by CDK7 or CDK9 [5]. Conversely, phosphorylated CTD partitioned more favorably than unphosphorylated CTD into condensates containing nuclear-speckle components, SRSF1 and SRSF2, two proteins involved in splicing [5]. The CTD concentration relative to mediator or splicing factors in the two condensates remains a key question, the answer to which may reveal which component scaffolds each condensate versus merely partitions into it. Indeed, is the CTD acting as a scaffold or a client for these condensates [10]? Moreover, does CTD status as scaffold or client shift as a function of phosphorylation? Another open question is whether or not phospho-specific CTD condensates are a general feature of active genes, or are they unique to the super-enhancer regulated genes studied [5]? Regardless, these findings connect CTD condensates to the CTD code, with initiation condensates giving way to splicing condensates as a function of CTD phosphorylation.

What exactly do the separate phospho-specific CTD condensates generated in vitro tell us about transcription...
in vivo? Does transcription require eviction of Pol II from one condensate and migration into another? Or does one condensate simply evolve into another during the course of transcription in a process driven by CTD phosphorylation? We consider these two models below (Figure 1).

In the first model, there are distinct initiation and splicing condensates, with Pol II leaving the former and entering the latter as it traverses a gene and CTD phosphorylation increases (Figure 1A). Here, these droplets could be distinct and separate immiscible entities, or adjacent drops could touch and transiently exchange components, perhaps akin to kiss-and-run fusion events. Droplet fusion can occur much faster than internal mixing, which may help explain the proximity, but incomplete overlap, of mediator and splicing factors. This model requires elongating Pol II to exit one droplet and enter another, raising questions about how cells would sort which polymerases to extract and which to leave behind?

One possibility is that genic DNA radiates outward from mediator droplets in a 5' to 3' orientation, with CTD phosphorylation and the energy of NTP hydrolysis driving Pol II across the droplet energy barrier, leaving behind the non-transcribing and hypophosphorylated Pol II to prime re-initiation. Sorting might also arise from phosphorylation by one or more kinases applied progressively and primed by an initiation-associated mark, ensuring that elongating Pol II is multiply phosphorylated more rapidly than Pol II that has yet to initiate.

In the second model, there is but a single droplet about a transcribed gene or cluster of genes, with the partition coefficient of specific components evolving over time as a function of CTD phosphorylation (Figure 1B). In this model, CTD phosphorylation could tune the partition function of CTD-associated factors, including mediator and splicing factors, for each condensate. Here, gene output would be self-limiting, with initiation ceasing as the partition function for mediator into the droplet decreases and the partition function for splicing factors into the condensate increases. This condensate could conceivably mature into a termination condensate, temporarily shutting down transcription.
machinery. We note that this model has the advantage of being self-limiting, as an initiation droplet gives way to an elongation droplet, and perhaps so on to a termination droplet. In this model, the CTD and perhaps other components may be expected to have partition coefficients associated with a threshold level of phosphorylation, whereby the composition of a droplet can regulate itself via the net phosphorylation levels of one or more condensate scaffolds or clients.

More complex in vitro reconstitution and visualization of catalytically ‘active droplets’, including regulatory kinases, could help distinguish between these possibilities. Ideally, such experiments would also include specific RNA transcripts, such as Trim28 and Nanog, given the established roles for RNA in regulating condensates. However, this type of sophisticated reconstitution might be challenging. Perhaps more feasible is tracking whether a single Pol II can be captured traversing from initiation to splicing condensates in vivo via super-resolution microscopy. Likewise, can a single initiation condensate be observed to evolve into a splicing condensate in vivo? Finally, are these condensates a cause or effect of transcription or splicing? Do these condensates truly create specialized microenvironments optimized for the biochemistry of these reactions? The answers to these and other questions will help reveal the mechanisms by which transcriptional and splicing condensates, and potentially transcription and splicing, are regulated.

4. Lu, F. et al. (2019) The C-terminal domain of RNA polymerase II is a multivalent targeting sequence that supports Drosophila development with only consensus heptads. Mol. Cell 73, 1232–1242

References