SOSS-INTAC: a new gatekeeper of genomic integrity at the interface of transcription and R-loops

Natalia Gromak

Sir William Dunn School of Pathology, University of Oxford, Oxford, OX1 3RE, UK

Correspondence: Natalia Gromak; email: natalia.gromak@path.ox.ac.uk

Summary

A recent Nature paper by Xu et al [1] describes an important link between RNA polymerase II promoter-proximal pausing and genome stability maintenance orchestrated by liquid droplet formation to reduce unwanted R-loop accumulation.

TEXT

Precise regulation of gene expression and maintenance of genome integrity are two fundamental processes in any living organism. However, we still do not fully understand how communication between these is orchestrated. Transcription by RNA polymerase II (Pol II) in eukaryotic organisms requires negotiation of promoter-proximal pausing before Pol II can enter productive elongation. One of the major players in this process is the 1.59 Megadalton Integrator-PP2A complex (INTAC), consisting of the 14-subunit RNA cleavage Integrator complex (INTS1-14) and protein phosphatase 2A (PP2A) [2] (Figure 1A). INTAC is critical for termination of promoter-proximal paused Pol II, as the Integrator complex cleaves the nascent RNA close to Pol II exit channel and PP2A dephosphorylates promoter-proximal paused Pol II to inhibit elongation [3]. However, how INTAC is recruited during transcription to execute these functions represents one of the key unanswered questions in the field. The paper by Xu et al sheds light on the underlying molecular mechanism of this process.

The maintenance of genome integrity in human cells is assisted by a heterotrimeric sensor of ssDNA (SOSS) complex, which includes INTS3/SOSS-A, the ssDNA-binding protein SSB1/SOSS-B and SOSS-C [4], acting through the ATM-signaling pathway (Figure 1A). INTS3 was shown to be a central adaptor stabilizing SOSS complex assembly on DNA. Moreover, INTS3 is a component of both INTAC and SOSS complexes. Previous studies have hinted at the existence of a functional connection between INTAC and SOSS complexes, since SSB1 can interact with multiple subunits of the Integrator complex, including INTS3. Xu et al have now used comprehensive biochemical experiments to uncover an interaction between INTAC and SOSS, which results in the formation of a stable SOSS-INTAC complex, providing cross-talk between transcription and genome stability machineries. This is supported by chromatin immuno-precipitation and sequencing (ChIP-seq) experiments for SSB1, INTS3 and INTS5, which demonstrate that SOSS-INTAC binds to promoters and enhancers. Interestingly, SSB1 is required for INTAC recruitment to genes, and this is mediated by the ssDNA-binding capacity of SSB1 as shown by EMSA *in vitro* and by correlation between SOSS-INTAC occupancy and ssDNA presence *in vivo* by kethoxal-assisted single-stranded DNA sequencing (KAS-seq).

R-loops are three-stranded structures composed of an RNA/DNA hybrid and unpaired ssDNA, which have emerged as important regulators of multiple biological processes [5]. R-loops occupy ~10% of human genome, where they are particularly enriched over promoters and terminators [6]. R-loops can promote Pol II termination at the 3'ends of human genes, a process which also requires an RNA/DNA helicase Senataxin [7]. Xu et al now demonstrate that promoter-proximal R-loops promote early Pol II termination by recruitment of INTAC-SOSS complex to chromatin through the N-terminal domain of SSB1 which recognizes ssDNA, released by promoter-proximal R-loop. This in turn brings the endonuclease domain of INTS11 into play to cleave TSS-proximal transcripts that then restrict local R-loops by recruiting exonucleases such as exosome complex (DIS3 and EXOSC10) and Xrn2.

Consequently, two molecular processes are facilitated: termination of TSS-paused Pol II and removal of DNA damage prone R-loops (Figure 1B). In essence, this study proposes a novel mechanism of R-loop-dependent promoter-proximal Pol II termination acting through SOSS-INTAC complex.

Unexpectedly, the function of SSB1 to recruit SOSS-INTAC relies on its ability to form liquid droplets, mediating SOSS-INTAC condensate formation, as demonstrated *in vitro* and in cells. The authors identified critical residues within the conserved C-terminal intrinsically disordered region (IDR) of SSB1 which are important for condensate formation, and these correspond to residues mutated in cancer. The condensation capacity of SSB1 acts to generally restrict R-loops *in vivo*, a process important to prevent transcription-replication conflicts and undesired genomic instability. It has been observed that R-loop-interacting proteins are enriched in IDRs, highlighting their potential ability to create multiple interactions, where R-loops may act as a functional scaffold for regulatory processes and also contribute to condensates formation [8].

The demonstrated ability of SSB1, INTS2 and INTS11 to restrict R-loop formation that otherwise causes increased DNA damage and delayed replication fork progression, suggests that the entire SOSS-INTAC complex is likely to function to maintain genome stability. In line with this, dysregulation of SSB1 and Integrator is linked to cancer and developmental defects. Previous studies in bacteria have demonstrated that SSB can recruit RNase H to resolve R-loops at replication sites [9]. It is possible that SSB1 may recruit additional co-factors in human cells which help to counteract promoter-proximal R-loops. Remarkably, endonucleolytic activity of the INTAC is required for R-loop removal. However, future studies using rapid depletion approaches for RNA exonucleases and mapping Integrator cleavage site in relation to R-loop position will provide insight into kinetics of this process. It is likely that helicases which can unwind either RNA (i.e. MTR4, as shown in this paper) or R-loops may also be involved in these promoter-proximal Pol II termination events. Given the prevalence of R-loops at other genomic locations, it is not clear if this termination mechanism is unique to promoter-proximal R-loops. Understanding how SOSS-INTAC recruitment to R-loops is regulated would help to decipher functional differences between R-loops, a critical unresolved question in the field.

R-loops accumulate in pathological conditions, including cancer and neurodegeneration. Can SOSS-INTAC complex promote unscheduled R-loop-mediated Pol II termination and therefore help to resolve R-loops in disease? Previously it was demonstrated that structure-specific ssDNA endonucleases XPG/XPF, components of the transcription-coupled nucleotide excision repair machinery, can process R-loops if they exceed physiological levels [10]. These two pathways may co-exist in pathological conditions. Thus, an understanding of the interplay between SOSS-INTAC complex and these R-loop nucleases warrants further study.

In summary, this study reveals an important link between transcription and genome stability regulation and showcases the functional significance of protein/RNA phase separation. Understanding how SOSS-INTAC regulates promoter-proximal Pol II pausing will be important for future investigations of pausing regulation in other contexts beyond transcription, such as replication and DNA damage response. Notably, these findings uncover a new regulatory function for promoter-proximal R-loops, further expanding the repertoire of these structures in health and disease.

Acknowledgements
NG is supported by the MRC Senior Fellowship (MR/X020487/1) and EPA Fund (Dunn School of Pathology, University of Oxford).

Declaration of Interests

The author declares no competing interests.

Figure 1. Role of SOSS-INTAC complex in co-ordinating transcription and genome stability.

A. Schematic diagram depicting formation of SOSS-INTAC complex. Based on currently available structural and biochemical information SOSS-INTAC complex can be divided into six modules, including the backbone (INTS1, INTS2 and INTS7), shoulder (INTS5 and INTS8), endonuclease (INTS4, INTS9 and INTS11), phosphatase (INTS6, PP2A-A and PP2A-C), auxiliary (INTS10/13/14/15) and SOSS (INTS3, SSB1/2, INIP) modules. The structural placement of INTS12 is currently unclear. SOSS-INTAC ensures efficient transcription and maintenance of genome stability.

B. (I) In physiological conditions, transcription by Pol II results in formation of promoter-proximal R-loop. (II) The SSB1 subunit of SOSS interacts with single stranded DNA (ssDNA), released by R-loop, to recruit SOSS-INTAC to promoters and drives condensate formation. (III) RNA cleavage by SOSS-INTAC complex promotes RNA degradation by a combination of Xrn2 and exosome activities, leading to premature promoter-proximal Pol II termination and R-loop attenuation. Cancer-associated mutations of SSB1 impairing condensation or disrupting SOSS-INTAC complex lead to loss of promoter-proximal Pol II termination and aberrant accumulation of R-loops, resulting in DNA damage and genome instability.

References

1. Xu, C., Li, C., Chen, J., Xiong, Y., Qiao, Z., Fan, P., Li, C., Ma, S., Liu, J., Song, A., et al. (2023). R-loop-dependent promoter-proximal termination ensures genome stability. Nature. 10.1038/s41586-023-06515-5.
2. Zheng, H., Qi, Y., Hu, S., Cao, X., Xu, C., Yin, Z., Chen, X., Li, Y., Liu, W., Li, J., et al. (2020). Identification of Integrator-PP2A complex (INTAC), an RNA polymerase II phosphatase. Science *370*. 10.1126/science.abb5872.
3. Fianu, I., Chen, Y., Dienemann, C., Dybkov, O., Linden, A., Urlaub, H., and Cramer, P. (2021). Structural basis of Integrator-mediated transcription regulation. Science *374*, 883-887. 10.1126/science.abk0154.
4. Huang, J., Gong, Z., Ghosal, G., and Chen, J. (2009). SOSS complexes participate in the maintenance of genomic stability. Mol Cell *35*, 384-393. 10.1016/j.molcel.2009.06.011.
5. Garcia-Muse, T., and Aguilera, A. (2019). R Loops: From Physiological to Pathological Roles. Cell *179*, 604-618. 10.1016/j.cell.2019.08.055.
6. Sanz, L.A., Hartono, S.R., Lim, Y.W., Steyaert, S., Rajpurkar, A., Ginno, P.A., Xu, X., and Chedin, F. (2016). Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. Mol Cell *63*, 167-178. 10.1016/j.molcel.2016.05.032.
7. Skourti-Stathaki, K., Proudfoot, N.J., and Gromak, N. (2011). Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol Cell *42*, 794-805. 10.1016/j.molcel.2011.04.026.
8. Dettori, L.G., Torrejon, D., Chakraborty, A., Dutta, A., Mohamed, M., Papp, C., Kuznetsov, V.A., Sung, P., Feng, W., and Bah, A. (2021). A Tale of Loops and Tails: The Role of Intrinsically Disordered Protein Regions in R-Loop Recognition and Phase Separation. Front Mol Biosci *8*, 691694. 10.3389/fmolb.2021.691694.
9. Wolak, C., Ma, H.J., Soubry, N., Sandler, S.J., Reyes-Lamothe, R., and Keck, J.L. (2020). Interaction with single-stranded DNA-binding protein localizes ribonuclease HI to DNA

replication forks and facilitates R-loop removal. Mol Microbiol *114*, 495-509. 10.1111/mmi.14529.

10. Sollier, J., Stork, C.T., Garcia-Rubio, M.L., Paulsen, R.D., Aguilera, A., and Cimprich, K.A. (2014). Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. Mol Cell *56*, 777-785. 10.1016/j.molcel.2014.10.020.