THE LESSER KNOWN WORLD OF RNA POLYMERASES

EDITED BY: Francisco Navarro, Athar Ansari and Olga Calvo

PUBLISHED IN: Frontiers in Molecular Biosciences

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-8897-4114-4
DOI 10.3389/978-2-8897-4114-4

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

THE LESSER KNOWN WORLD OF RNA POLYMERASES

Topic Editors:
Francisco Navarro, University of Jaén, Spain
Athar Ansari, Wayne State University, United States
Olga Calvo, Universidad de Salamanca, Spain

Citation: Navarro, F., Ansari, A., Calvo, O., eds. (2023). The Lesser Known World of RNA Polymerases. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8897-4114-4

Table of Contents

04 ***Editorial: The Lesser Known World of RNA Polymerases***
Olga Calvo, Athar Ansari and Francisco Navarro

08 ***Gene Architecture Facilitates Intron-Mediated Enhancement of Transcription***
Katherine Dwyer, Neha Agarwal, Lori Pile and Athar Ansari

16 ***Eukaryotic RNA Polymerases: The Many Ways to Transcribe a Gene***
Marina Barba-Aliaga, Paula Alepuz and José E. Pérez-Ortín

24 ***Biogenesis of RNA Polymerases in Yeast***
Ana I. Garrido-Godino, Francisco Gutiérrez-Santiago and Francisco Navarro

34 ***Critical Involvement of TFIIB in Viral Pathogenesis***
Michael J. O'Brien and Athar Ansari

42 ***Early Evolution of Transcription Systems and Divergence of Archaea and Bacteria***
Lei Lei and Zachary F. Burton

57 ***Transcribing Genes the Hard Way:* In Vitro *Reconstitution of Nanoarchaeal RNA Polymerase Reveals Unusual Active Site Properties***
Sven Nottebaum and Robert O. J. Weinzierl

66 ***Regulation of RNA Polymerase II Transcription Initiation and Elongation by Transcription Factor TFII-I***
Niko Linzer, Alexis Trumbull, Rukiye Nar, Matthew D. Gibbons, David T. Yu, John Strouboulis and Jörg Bungert

74 ***Specific Features of RNA Polymerases I and III: Structure and Assembly***
Tomasz W. Turowski and Magdalena Boguta

82 ***The Transcription-Repair Coupling Factor Mfd Prevents and Promotes Mutagenesis in a Context-Dependent Manner***
Laura A. Lindsey-Boltz and Aziz Sancar

88 ***Regulation of Eukaryotic RNAPs Activities by Phosphorylation***
Araceli González-Jiménez, Adrián Campos, Francisco Navarro, Andrés Clemente-Blanco and Olga Calvo

99 ***RNA Polymerase III Subunit Mutations in Genetic Diseases***
Elisabeth Lata, Karine Choquet, Francis Sagliocco, Bernard Brais, Geneviève Bernard and Martin Teichmann

129 ***Coupling Between Cell Cycle Progression and the Nuclear RNA Polymerases System***
Irene Delgado-Román and Mari Cruz Muñoz-Centeno

137 ***Coupling Between Production of Ribosomal RNA and Maturation: Just at the Beginning***
Chaima Azouzi, Mariam Jaafar, Christophe Dez, Raghida Abou Merhi, Annick Lesne, Anthony K. Henras and Olivier Gadal

146 ***Several Isoforms for Each Subunit Shared by RNA Polymerases are Differentially Expressed in the Cultivated Olive Tree (*Olea europaea *L.)***
Isabel Fernández-Parras, Jorge Antolín Ramírez-Tejero, Francisco Luque and Francisco Navarro

frontiers in Molecular Biosciences

EDITORIAL
published: 02 December 2021
doi: 10.3389/fmolb.2021.811413

Editorial: The Lesser Known World of RNA Polymerases

Olga Calvo[1]*, *Athar Ansari*[2]* and *Francisco Navarro*[3,4]*

[1]Instituto de Biología Funcional y Genómica, Consejo Superior de Investigaciones Científicas, Universidad de Salamanca, Salamanca, Spain, [2]Department of Biological Science, Wayne State University, Detroit, MI, United States, [3]Departamento de Biología Experimental-Genética, Universidad de Jaén, Jaén, Spain, [4]Centro de Estudios Avanzados en Aceite de Oliva y Olivar, Universidad de Jaén, Jaén, Spain

Keywords: RNA polymerase, transcription, evolution, biogenesis, posttranslational modifications, diseases, splicing, transcription factors

Editorial on the Research Topic

The Lesser Known World of RNA Polymerases

According to the central dogma of molecular biology, genetic information is passed from DNA to RNA to protein. The transmission of information from DNA to RNA is called transcription and is carried out by the RNA polymerases (RNA pols). The RNA pols from bacteria to eukaryotes are multimeric enzymes which show a high degree of conservation in terms of their structure and functionality (Cramer, 2019; Werner and Grohmann 2011) (**Figure 1**). Notably, while bacteria and archaea contain only one RNA pol, most eukaryotes contain three different enzymes, RNA pol I, II and III, with the exception of plants that also have two additional RNA pols, IV and V, which evolved from RNA pol II (Ream et al., 2014). RNA pol I synthesizes the precursor of the three largest rRNAs, RNA pol III produces mostly tRNAs, the 5S rRNA and several short non-translated RNAs and RNA pol II give rise to all mRNAs and many non-coding RNAs, including miRNA. Finally, RNA pol IV and V participate in transcriptional silencing and also in production of non-coding RNAs involved in the development and response to environmental changes (Werner and Grohmann 2011; Ream et al., 2014; Cramer, 2019). The enzyme exhibits a broad evolutionary diversity, functional dynamism and pleiotropic role in biological systems. However, many aspects of RNA pols, including their biogenesis, function, and even their impact in different cellular processes or health, have not been deeply investigated. This special issue is an attempt to cover some of the lesser-known aspects of RNA pol diversity, dynamism, function and evolutionary conservation. In addition, this issue also considers transcription factors as part of the transcriptional machinery.

Despite variation in structure and subunit composition, RNA pols from different organisms harbors conserved features (Lane and Darst, 2010a; Lane and Darst, 2010b). The article by Lei and Burton compares the three-dimensional structure of archaeal, bacterial and eukaryotic RNA pols. Their analyses revealed that the enzyme, in all three domains of life, are of two-double-ψ-β-barrel (2-DPBB) type. In addition to 2-DPBB, the catalytic core of multisubunit RNA pols is comprised of a conserved bridge helix and trigger loop. Lei and Burton propose that the 2-DPBB family of multisubunit RNA pols might have evolved prior to the last universal common cellular ancestor during evolution.

Nanoarchaea is a highly diverged archaeal phylum with many unusual biological features. Nottebaum and Weinzierl describe that several of the key motifs in the active center of *Nanoarchaeum equitans* RNA pol contain unusual and radical substitutions expected to be harmful to the catalytic activity. However, the authors reconstituted a RNA pol complex *in vitro* with transcription activity, concluding that sequence changes do not adversely affect catalytic activity, even if they are unusual and localized in key motifs. Moreover, they identified a stringent

OPEN ACCESS

Edited and reviewed by:
André P. Gerber,
University of Surrey, United Kingdom

***Correspondence:**
Olga Calvo
ocalvo@usal.es
Athar Ansari
bb2749@wayne.edu
Francisco Navarro
fngomez@ujaen.es

Specialty section:
This article was submitted to RNA Networks and Biology, a section of the journal Frontiers in Molecular Biosciences

Received: *08 November 2021*
Accepted: *17 November 2021*
Published: *02 December 2021*

Citation:
Calvo O, Ansari A and Navarro F (2021) Editorial: The Lesser Known World of RNA Polymerases. Front. Mol. Biosci. 8:811413. doi: 10.3389/fmolb.2021.811413

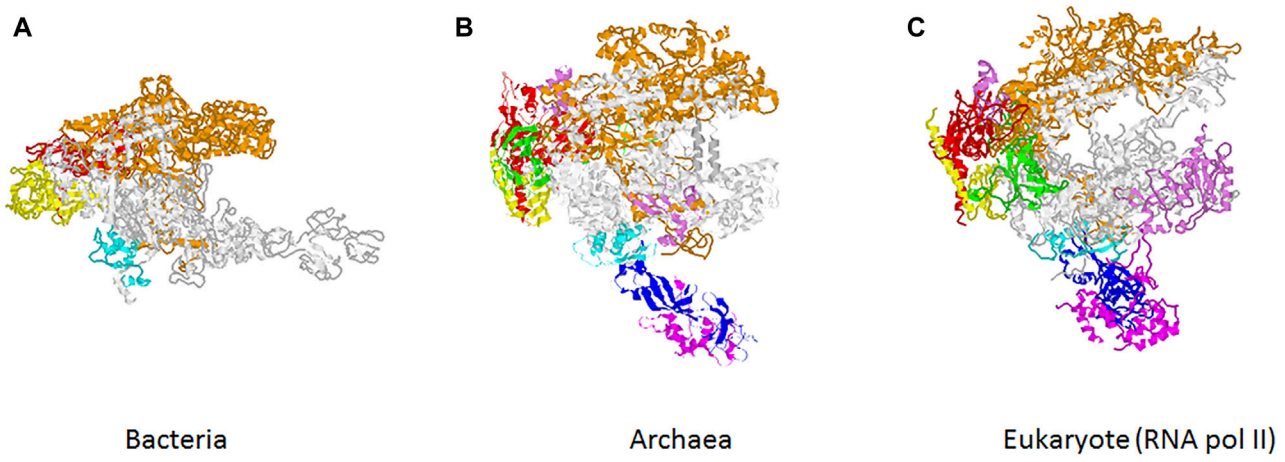


FIGURE 1 | Evolution of RNA polymerases. Structure of RNA polymerase from bacteria **(A)**, archaea **(B)** and eukaryotic RNA pol II **(C)**. Color code corresponds to the previously used for *Saccharomyces cerevisiae* RNA pol II subunits (Armache et al., 2003). Similar colors are used for homologous subunits in bacterial and archaeal RNA pols.

atypical requirement for fluoride ions for maximal RNA pol activity, proposing a model where more "conventional" archea will not use it.

The work by Barba-Aliaga et al. focuses on nuclear RNA pols in eukaryotes and summarizes their evolutionary origin and functional reasons that could have led to their multiplicity. Furthermore, authors discuss the regulation and the homeostasis of the different RNA products that they synthesise. The authors present several studies to show how the coordination between RNA pols activities is necessary for cellular processes, such as the influence of RNA pols on the translation machinery synthesis (ribosomes and tRNAs) or how eukaryotic RNA pols transcription regulation occurs with respect to the changes in cellular volume.

Plants are the only known eukaryotic organisms containing two addition RNA pols (IV and V), in addition to RNA pol I-III, which evolved from RNA pol II (Ream et al., 2013). The work by Fernández-Parras et al., investigates the transcriptional regulation of the RNA pols common subunits genes in olive tree cultivar (*Olea europaea* L. cv. Picual) and shows that they are spatio-temporally regulated, as well as regulated by biotic and abiotic stresses. This work opens questions about the existence of multiple RNA pols variants in polyploid organisms.

A model for the biogenesis of eukaryotic RNA pol II has been proposed based on the bacterial RNA pol formation, a sequential process involving participation of several subassembly complexes that leads to the complete enzyme assembly in the cytoplasm before its nuclear import (Wild and Cramer, 2012). However, despite recent progress, the assembly of RNA pols remains poorly described. The work by Garrido-Godino et al., focuses on the knowledge of biogenesis of RNA pols in yeast. The authors review the mechanisms and proteins (assembly and transport factors) involved in these processes and make the comparison with human factors described previously. In addition, the manuscript by Turowski and Boguta summarizes the current knowledge on the biogenesis of the RNA pol I and III and focuses on the model of their co-translational assembly, based on recent publications, showing the importance of Rpb10, Rpc19 and Rpc40, and of the Rbs1 protein in the assembly of RNA pol I and III complexes.

Notably, mutations of RNA pols lead to diseases and disorders. Some of these are suggested to be associated with RNA pol III assembly. Although the causal relationship between RNA pols mutations and disease development is widely accepted, the associated molecular mechanisms are poorly understood. The work by Lata et al. reviews the current knowledge regarding the functional impact of specific mutations, possible Pol III-related disease-causing mechanisms, and animal models that may help to better understand the links between Pol III mutations and disease. Similarly, a large number of genetic diseases associated with RNA pol I mutations exists. They are collectively called ribosomopathies. The understanding of the precise mechanistic of Pol I transcription opens broad perspectives in health-related research areas.

Azouzi et al. nicely review recent advances in the field of RNA pol I transcription elongation, revealed using nucleotide resolution techniques. These advances showed the connection between the production of rRNA and nascent rRNA folding. Indeed, rRNA folding during transcription seems to be an anti-pausing mechanism favoring transcription elongation because rRNA secondary structures prevent backtracking. Furthermore, they also discuss mechanisms involved in RNA pol I termination. Based on recent discoveries by Darrière et al. (2019), using a super-active RNA polI mutant, the authors propose that premature transcription termination at defined positions can control rRNA production *in vivo*.

There is an enormous interest in deciphering how RNA pols integrate the information that cells receive and how RNA pols are coordinated and communicated to regulate gene expression in response to physiological and pathological conditions. In this regard, Delgado-Román and Muñóz-Centeno proposed that RNA pols I and III activities should be connected to regulate

cell cycle progression. How cell cycle regulation is affected by the balance between the three RNA pols products and RNA pols assembly is discussed. The authors focus on ribogenesis, a process that requires the activity of all three RNA pols (de la Cruz et al., 2018), and discuss how the balanced production of ribosomal components prevents G1 arrest in budding yeast and mammalian cells, which show strong analogies (Bursac et al., 2012; Gómez-Herreros et al., 2013).

González-Jiménez et al. propose that phosphorylation may have a role in the coordination of the three transcription machineries. Various studies have reported that several subunits of RNA pol I, II and III are susceptible to phosphorylation (for instance, Šoštarić, et al., 2018; Lanz et al., 2021). Some of these phosphorylation sites are distributed within subunits common to all three RNA pols. This suggests that phosphorylation events might finely modulate the activities of all RNA pols and give rise to the speculation that they can play a crucial regulatory role in the coordination between the three RNA pols, which, so far, has not been investigated enough. In this review the authors compile all the known phosphorylation sites identified for the three RNA pols, localized most of them within the respective complexes and discussed their possible roles. This is a valuable information for researchers interested in this exciting and promising field of study.

Since their discovery, the biological significance of introns in the eukaryotic genome has remained an enigma. A number of studies in a diversity of eukaryotes have revealed that the process of splicing, which removes an intron from a primary transcript, is often a positive regulator of transcription (Gallegos and Rose 2015; Shaul 2017). The article by Dwyer et al. proposes a novel mechanism of splicing-mediated regulation of transcription by RNA pol II through modulating the gene architecture. The transition in topology of a gene from linear to a loop during cotranscriptional splicing and the mechanism of enhancement of transcription by the looped structure is being discussed.

Transcription by RNA pol II in eukaryotes is facilitated by a number of transcription factors. TFIIB is one such essential general transcription factors (Deng and Roberts 2007). The article by O' Brien and Ansari focuses on a rather unexpected role of TFIIB during viral pathogenesis. The article describes in detail the targeting of TFIIB by viral transcriptional regulators during pathogenesis. Likely reasons for preferred targeting of TFIIB over other general transcription factors by viruses are discussed. This makes TFIIB a potential target of antiviral therapies.

TFII-I is another transcription factor of RNA pol II (Roy 2012). It was originally discovered as an initiator-binding protein that helps in initiation of transcription from TATA-less promoters. Further research revealed that TFII-I is involved in post-initiation steps as well. Linzer et al. discuss multiple aspects of TFII-I participation in the transcription cycle. In addition to affecting initiation from a subset of promoters, TFII-I is involved in transcription elongation by regulating pausing of RNA pol. The involvement of this factor in cancer, neurological and immunological disorders in humans, development in mice, and induction of pluripotency is discussed.

It was known since a long time that the transcriptionally active UV-damaged regions of genome are repaired more efficiently than the non-transcribed regions. The factor responsible for the transcription-coupled repair of damaged DNA in prokaryotes is *mfd* (Selby and Sancar, 1993). The article by Lindsey-Boltz and Sancar discusses the recent advances in three-dimensional structure and single molecule studies pertaining to *mfd*. These studies have revealed that *Mfd* binds stalled RNA pols even in the absence of UV damage and helps the pol operate in hard-to-transcribe regions. The possibility of *mfd*-RNA polymerase interaction contributing to both, promotion and prevention of mutagenesis in a context-dependent manner, is discussed.

The articles in this special issue cover some, but not all, lesser-known aspects of RNA pols and highlight that many mechanistic, structural and/or evolutionary aspects of RNA pols, among others, remain unexplored or are still not well investigated. Futures investigations on RNA pols will greatly help to understand gene expression regulation, where transcription is the bottle neck.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

REFERENCES

Armache, K.-J., Kettenberger, H., and Cramer, P. (2003). Architecture of Initiation-Competent 12-subunit RNA Polymerase II. *Proc. Natl. Acad. Sci.* 100, 6964–6968. doi:10.1073/pnas.1030608100

Bursac, S., Brdovcak, M. C., Pfannkuchen, M., Orsolic, I., Golomb, L., Zhu, Y., et al. (2012). Mutual Protection of Ribosomal Proteins L5 and L11 From Degradation Is Essential for P53 Activation Upon Ribosomal Biogenesis Stress. *Proc. Natl. Acad. Sci.* 109 (50), 20467–20472. doi:10.1073/pnas.1218535109

Cramer, P. (2019). Eukaryotic Transcription Turns 50. *Cell.* 179 (4), 808–812. doi:10.1016/j.cell.2019.09.018

Darrière, T., Pilsl, M., Sarthou, M.-K., Chauvier, A., GentyAudibert, T. S., Audibert, S., et al. (2019). Genetic Analyses Led to the Discovery of a Super-Active Mutant of the RNA Polymerase I. *Plos Genet.* 15, e1008157. doi:10.1371/journal.pgen.1008157

de la Cruz, J., Gómez-Herreros, F., Rodríguez-Galán, O., Begley, V., de la Cruz Muñoz-Centeno, M., and Chávez, S. (2018). Feedback Regulation of Ribosome Assembly. *Curr. Genet.* 64 (2), 393–404. doi:10.1007/s00294-017-0764-x

Deng, W., and Roberts, S. G. E. (2007). TFIIB and the Regulation of Transcription by RNA Polymerase II. *Chromosoma.* 116 (5), 417–429. doi:10.1007/s00412-007-0113-9

Gallegos, J. E., and Rose, A. B. (2015). The Enduring Mystery of Intron-Mediated Enhancement. *Plant Sci.* 237, 8–15. doi:10.1016/j.plantsci.2015.04.017

Gómez-Herreros, F., Rodríguez-Galán, O., Morillo-Huesca, M., Maya, D., Arista-Romero, M., de la Cruz, J., et al. (2013). Balanced Production of Ribosome Components Is Required for Proper G1/S Transition in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 288 (44), 31689–31700. doi:10.1074/jbc.M113.500488

Lane, W. J., and Darst, S. A. (2010a). Molecular Evolution of Multisubunit RNA Polymerases: Sequence Analysis. *J. Mol. Biol.* 395, 671–685. doi:10.1016/j.jmb.2009.10.062

Lane, W. J., and Darst, S. A. (2010b). Molecular Evolution of Multisubunit RNA Polymerases: Structural Analysis. *J. Mol. Biol.* 395, 686–704. doi:10.1016/j.jmb.2009.10.063

Lanz, M. C., Yugandhar, K., Gupta, S., Sanford, E. J., Faça, V. M., Vega, S., et al. (2021). In-Depth and 3-dimensional Exploration of the Budding Yeast Phosphoproteome. *EMBO Rep.* 22 (2), e51121. doi:10.15252/embr.202051121

Ream, T., Haag, J., and Heidelberg, K. S. M. a. M. A. Trakselis. (2014). *Plant Multisubunit RNA Polymerases IV and V. Nucleic Acid Polymerases*. 30, Heidelberg: Springer-Verlag.

Ream, T., Haag, J., and Pikaard, C. (2013). "Plant Multisubunit RNA Polymerases IV and V," in *Nucleic Acid Polymerases*. Editor K. S. M. a. M. A. Trakselis (Heidelberg: Springer-Verlag). doi:10.1007/978-3-642-39796-7_13

Roy, A. L. (2012). Biochemistry and Biology of the Inducible Multifunctional Transcription Factor TFII-I: 10years Later. *Gene.* 492, 32–41. doi:10.1016/j.gene.2011.10.030

Selby, C. P., and Sancar, A. (1993). Molecular Mechanism of Transcription-Repair Coupling. *Science.* 260, 53–58. doi:10.1126/science.8465200

Shaul, O. (2017). How Introns Enhance Gene Expression. *Int. J. Biochem. Cell Biol.* 91, 145–155. doi:10.1016/j.biocel.2017.06.016

Šoštarić, N., O'Reilly, F. J., Giansanti, P., Heck, A. J. R., Gavin, A.-C., and van Noort, V. (2018). Effects of Acetylation and Phosphorylation on Subunit Interactions in Three Large Eukaryotic Complexes. *Mol. Cel Proteomics.* 17 (12), 2387–2401. doi:10.1074/mcp.RA118.000892

Werner, F., and Grohmann, D. (2011). Evolution of Multisubunit RNA Polymerases in the Three Domains of Life. *Nat. Rev. Microbiol.* 9 (2), 85–98. doi:10.1038/nrmicro2507

Wild, T., and Cramer, P. (2012). Biogenesis of Multisubunit RNA Polymerases. *Trends Biochem. Sci.* 37, 99–105. doi:10.1016/j.tibs.2011.12.001

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

MINI REVIEW
published: 21 April 2021
doi: 10.3389/fmolb.2021.669004

Gene Architecture Facilitates Intron-Mediated Enhancement of Transcription

*Katherine Dwyer, Neha Agarwal, Lori Pile and Athar Ansari**

Department of Biological Science, Wayne State University, Detroit, MI, United States

Introns impact several vital aspects of eukaryotic organisms like proteomic plasticity, genomic stability, stress response and gene expression. A role for introns in the regulation of gene expression at the level of transcription has been known for more than thirty years. The molecular basis underlying the phenomenon, however, is still not entirely clear. An important clue came from studies performed in budding yeast that indicate that the presence of an intron within a gene results in formation of a multi-looped gene architecture. When looping is defective, these interactions are abolished, and there is no enhancement of transcription despite normal splicing. In this review, we highlight several potential mechanisms through which looping interactions may enhance transcription. The promoter-5′ splice site interaction can facilitate initiation of transcription, the terminator-3′ splice site interaction can enable efficient termination of transcription, while the promoter-terminator interaction can enhance promoter directionality and expedite reinitiation of transcription. Like yeast, mammalian genes also exhibit an intragenic interaction of the promoter with the gene body, especially exons. Such promoter-exon interactions may be responsible for splicing-dependent transcriptional regulation. Thus, the splicing-facilitated changes in gene architecture may play a critical role in regulation of transcription in yeast as well as in higher eukaryotes.

Keywords: transcription, splicing, intron, exon, gene architecture, gene regulation, gene looping

OPEN ACCESS

Edited by:
Andrea Cerase,
Queen Mary University of London,
United Kingdom

Reviewed by:
Karla M. Neugebauer,
Yale University, United States
Sherif Abou Elela,
Université de Sherbrooke, Canada
Alan Rose,
University of California, Davis,
United States

***Correspondence:**
Athar Ansari
bb2749@wayne.edu

Specialty section:
This article was submitted to Protein and RNA Networks, a section of the journal Frontiers in Molecular Biosciences

Received: *17 February 2021*
Accepted: *31 March 2021*
Published: *21 April 2021*

Citation:
Dwyer K, Agarwal N, Pile L and Ansari A (2021) Gene Architecture Facilitates Intron-Mediated Enhancement of Transcription. Front. Mol. Biosci. 8:669004. doi: 10.3389/fmolb.2021.669004

INTRODUCTION

Introns are intervening non-coding sequences in eukaryotic genes that are removed from the primary transcripts by the process of splicing (**Figure 1**). An elaborate splicing machinery is needed to remove introns to form the mature transcript (Padgett et al., 1986). The energy, time consumption, and complex nature of the spliceosome indicate that introns impose a large burden on eukaryotic organisms (Jo and Choi, 2015). However, in spite of all of these drawbacks, introns have been evolutionarily conserved, indicative of their having a fundamental and significant role in the cell (Carmel and Chorev, 2012; Rogozin et al., 2012). In fact, nearly half of all common genetic disorders in humans may be attributed to a disruption in the splicing process (Faustino and Cooper, 2003; Wang and Cooper, 2007; Padgett, 2012). An important question therefore is what is the physiological significance of introns in eukaryotes?

Research conducted during the last few decades has revealed novel biological functions of introns. An apparent advantage conferred by introns is their ability to increase proteomic complexity through the process of alternative splicing, meaning that a single gene can produce multiple isoforms of a protein dependent on cell type and environment (Nilsen and Graveley, 2010). This offers immense proteomic plasticity since the cell can modulate what protein isoform is produced depending on the developmental context, cell type and environmental cues (Marquez et al., 2015; Chaudhary et al., 2019). The presence of introns has also been shown to protect genomic integrity. The process of transcription is accompanied by formation of genotoxic R-loops, which are three stranded nucleic acid structures harboring an RNA-DNA hybrid (Niehrs and Luke, 2020). Intron-containing genes alleviate R-loop formation and in doing so confer genomic stability and protect cells from harmful stress responses (Bonnet et al., 2017). Recent studies have led to the identification of yet another novel function of introns. Some excised introns provide protection against environmental stress (Morgan et al., 2019; Parenteau et al., 2019). Introns are also the source of snoRNA, microRNA, and lncRNA (Carmel and Chorev, 2012). These non-coding RNA species regulate gene expression at the level of transcription and RNA stability. Outside of acting as a non-coding regulatory RNA molecule, introns possess the unique function of acting as mutational buffers that can protect coding regions from incurring deleterious mutations (Jo and Choi, 2015). A harmful mutation in the coding region may affect the function of the protein. In contrast, a mutation in introns, which are non-coding regions, has minimal chances of affecting the function of the protein. In human genes, introns make up the bulk of a gene and therefore are able to absorb detrimental mutations without affecting the protein function. Although introns have been implicated in a variety of functions in eukaryotes, not all introns are associated with every function described above.

Of all the known functions of introns, one of the best known and evolutionarily conserved, is their ability to regulate the expression of genes that harbor them. Introns affect gene expression at multiple levels. They have been implicated in altering nucleosome positioning, transcription, RNA stability, nucleo-cytoplasmic export of mRNA, and translation efficiency (reviewed in Le Hir et al., 2003; Rose, 2008; Gallegos and Rose, 2015; Laxa, 2017; Shaul, 2017). The details regarding posttranscriptional regulation of gene expression by an intron have been covered in many recently published reviews (Gallegos and Rose, 2015; Laxa, 2017; Shaul, 2017). Here, we highlight the function of introns in regulating gene expression at the level of transcription. The transcription-enhancing potential of introns was first observed in cultured maize cells and transgenic mice (Callis et al., 1987; Brinster et al., 1988). Furthermore, cDNA of a number of human genes are not transcribed to the wild type level unless a promoter-proximal intron is included (Palmiter et al., 1991; Charron et al., 2007). Soon thereafter it was realized that introns play a general role in activating transcription in a variety of eukaryotes. Intron-mediated regulation of transcription has been observed in simple eukaryotes like yeast and Chlamydomonas as well as in higher eukaryotes like flies, worms, plants, and mammals including humans (Callis et al., 1987; Brinster et al., 1988; Bieberstein et al., 2012; Moabbi et al., 2012; Agarwal and Ansari, 2016; Baier et al., 2020). This enhancement property could be a very crucial function of introns in eukaryotes as a number of genes are dependent on introns for their normal transcription (Rose, 2019). In general, intron-containing genes in eukaryotes exhibit higher expression than their non-intronic counterparts (Ares et al., 1999; Juneau et al., 2006; Bieberstein et al., 2012; Gallegos and Rose, 2015; Ding and Elowitz, 2019; Baier et al., 2020). An intron may enhance transcription by a meager 2–3-fold, or it may augment mRNA output in the range of 10–100-fold or higher depending on the gene. Not all introns, however, can stimulate transcription. Some naturally occurring genes do not contain introns but are expressed efficiently, while some transgenes fail to express even in the presence of an intron (Pasleau et al., 1987; Malim et al., 1988; Rose, 2008). Introns are dispensable for enhancing transcription from a strong promoter (Huang and Liang, 1993). The transcription activation potential of introns though, is crucial for high expression of genes with a weak promoter. Some of the highly expressed genes like H2A and hepatitis B virus genes have *cis*-acting elements that appear to function like introns (Huang and Liang, 1993; Huang and Yen, 1995; Liu and Mertz, 1995). Despite some introns lacking the transcription enhancement potential, introns in general are emerging as an important component of the transcription regulatory machinery in eukaryotes (Rose, 2019).

Intron-mediated transcriptional regulation can be broadly divided into two categories: (1) splicing-independent, and (2) splicing-dependent regulation. Splicing-independent regulation is due to the presence of an enhancer or a promoter element within the intron (Kaneda et al., 1992; Bianchi et al., 2009; Beaulieu et al., 2011). Such introns can influence transcription even if their splicing function is compromised. In contrast, the splicing-dependent regulation requires a functional, splicing-competent intron within the transcribed region of the gene. Such introns cannot affect transcription if their splicing is inhibited by a mutation in the conserved sequences at the 5′ splice site, 3′ splice site and branchpoint, or if they are inserted in an anti-sense orientation (Furger et al., 2002; Moabbi et al., 2012; Agarwal and Ansari, 2016). This direct effect of a splicing-competent intron on transcription of a gene represents splicing-dependent regulation. It is often referred to as "intron-mediated enhancement" (IME). The focus of this review is to highlight key findings related to the possible mechanism of this enhancement with an emphasis on data indicating that genome architecture plays a critical role in the process.

INTRON-MEDIATED ENHANCEMENT EFFECT MAY TARGET INITIATION, ELONGATION OR TERMINATION STEPS

To enhance transcription, the intron must be located within the transcribed region of the gene and should be spliced (**Figures 1A**, **2A**). The activation is maximum when an intron

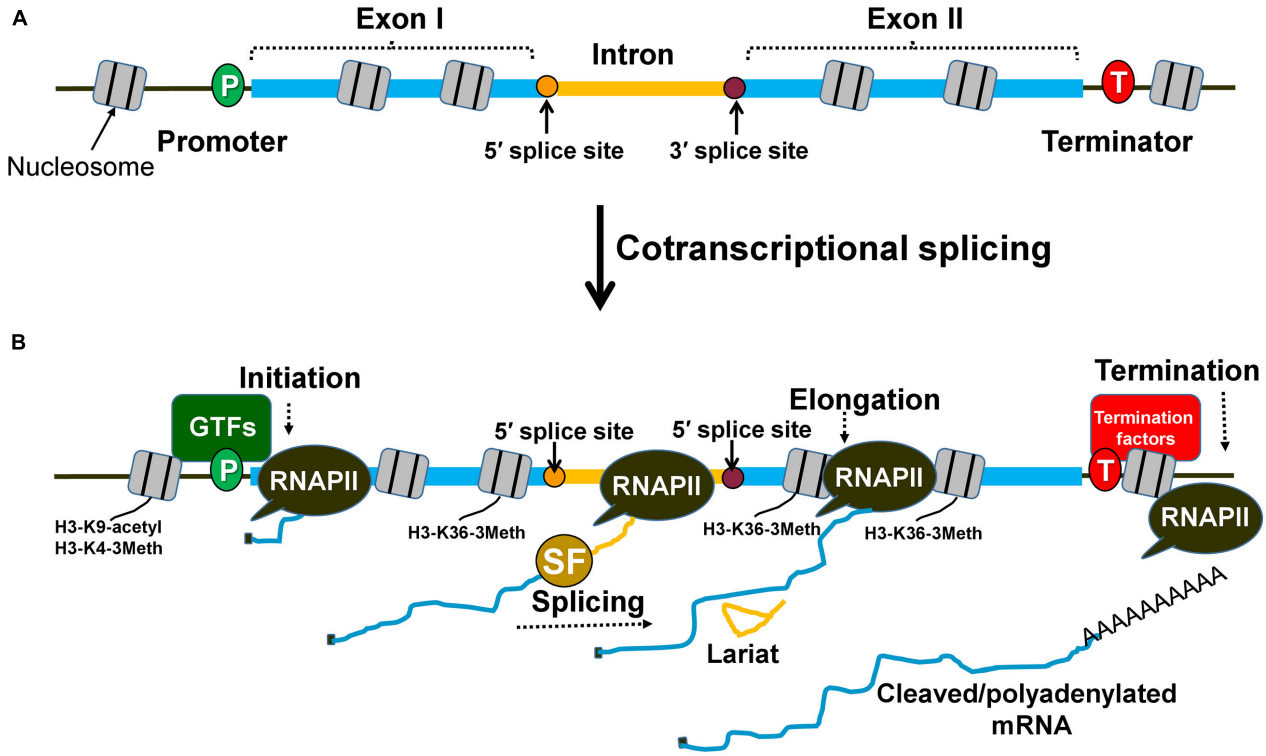


FIGURE 1 | Splicing occurs cotranscriptionally and affects different steps of transcription. **(A)** An intron within the transcribed region is flanked by exon I and exon II. The splice sites are designated as 5′ splice site and 3′ splice site and are the sites for spliceosome assembly during transcription. **(B)** Splicing factors are recruited cotranscriptionally to the intron with the help of the RNAPII carboxy-terminal domain. Spliceosomal assembly on the splice sites can facilitate the stabilization of general transcription factors (GTFs) at the promoter region of the gene and prime nucleosomes with activation marks (H3-K9 acetylation and H3-K4 trimethylation) for initiation. The splicing factors can also interact with transcription elongation factors and influence nucleosome modifications (H3-K36 trimethylation) to promote elongation. Similarly, splicing factors can contribute to enhanced termination of transcription by facilitating the recruitment of termination factors and removal of elongation marks that block effective termination.

is present close to the promoter, and gradually decreases with increasing distance from the promoter (Callis et al., 1987; Rose, 2004; Bieberstein et al., 2012; Gallegos and Rose, 2017). There are, however, reports that in some genes, the activation potential of the intron is partially restored with increasing proximity to the terminator (Callis et al., 1987; Kim et al., 2011). Thus, the position of an intron within a gene is an important determinant of its transcription activation potential.

Although introns can impact any step of the transcription cycle to achieve higher mRNA output, the initiation step is the most frequent target. The transcription regulating ability of an intron is dependent on the cotranscriptional nature of splicing (**Figure 1B**). The carboxy-terminal domain (CTD) of RNA Polymerase II (RNAPII) is used as a recruitment and docking site for several splicing factors that help in coupling of transcription with splicing (Millhouse and Manley, 2005; Nojima et al., 2018). During cotranscriptional splicing, the spliceosome assembled on the elongating mRNA facilitates initiation/reinitiation by stabilizing assembly of the preinitiation complex (PIC) on the promoter (**Figure 2B**; Tian, 2001; Kwek et al., 2002; Das et al., 2007; Damgaard et al., 2008; Jobert et al., 2009). The 5′ splice site of nascent transcripts plays a crucial role in these processes. In fact, a 5′ splice site alone, without being a part of an intron, can enhance transcription to some extent (Damgaard et al., 2008). The 5′ splice site enhances recruitment of TFIID, TFIIB, and TFIIH on the promoter (Das et al., 2007; Damgaard et al., 2008). Splicing factors like U1-snRNP and HNRNPU, which bind to the 5′ splice site, facilitate the recruitment of these general transcription factors (Tian, 2001; Das et al., 2007). U1-snRNP physically interacts with RNAPII during cotranscriptional splicing (Nojima et al., 2018). U1-snRNP also contacts the cyclin H subunit of TFIIH (Kwek et al., 2002). U1-snRNP may facilitate initiation and reinitiation of transcription by enabling recruitment of TFIIH, and probably TFIID and TFIIB as well (Kwek et al., 2002; Damgaard et al., 2008). HNRNPU similarly promotes initiation by enabling the recruitment of TFIIF to the pre-initiation complex (PIC) (Fiszbein et al., 2019). Furthermore, the promoter-proximal introns also affect histone H3-K9 acetylation and H3-K4 trimethylation near the 5′ end of genes (**Figure 1B**; Bieberstein et al., 2012). These two chromatin modifications facilitate the recruitment of the PIC on the promoter leading to the activation of genes, both in yeast and higher eukaryotes. Thus, an intron may affect initiation directly by facilitating the recruitment of general transcription factors on the promoter or indirectly by affecting the chromatin structure in the promoter region.

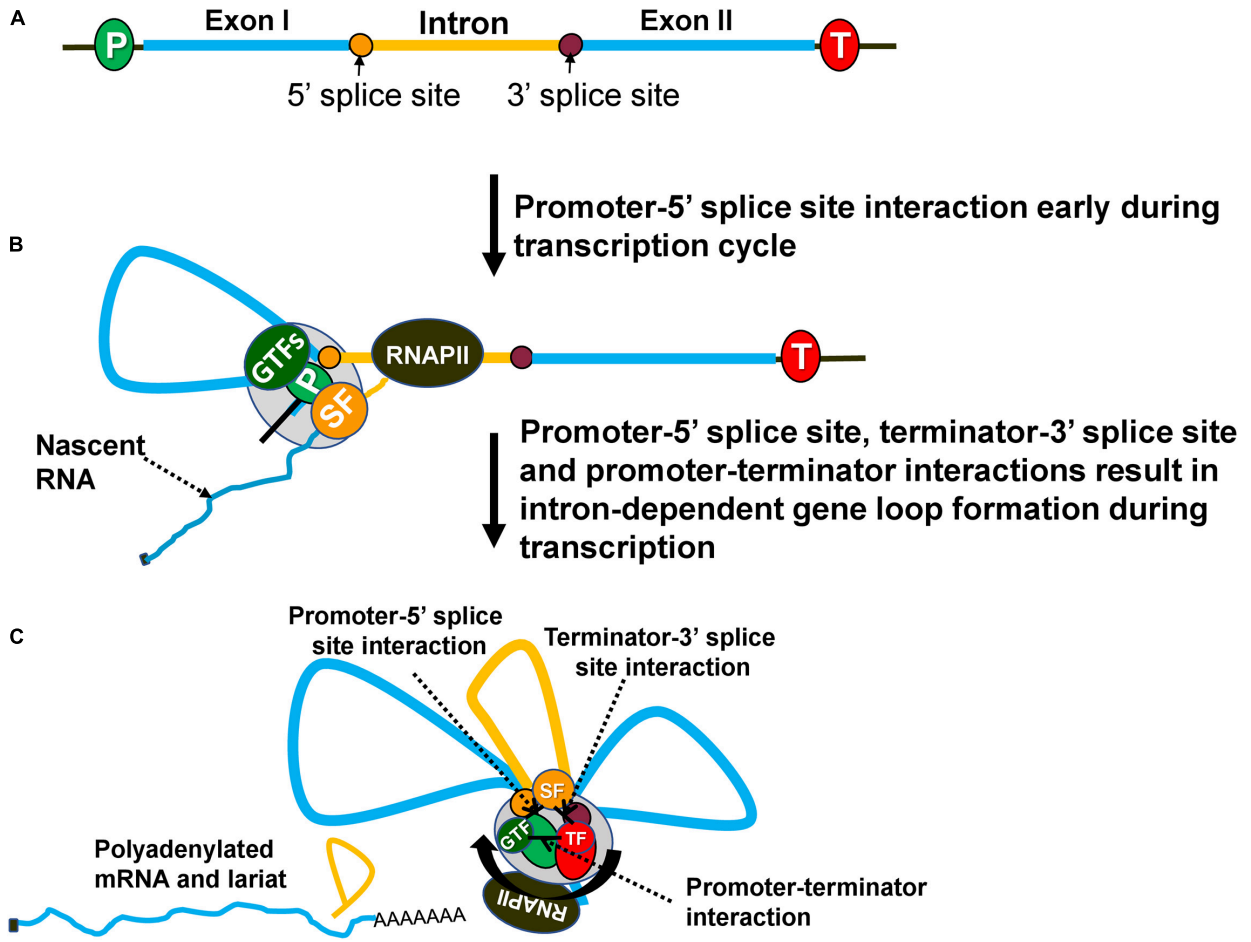


FIGURE 2 | Intron-dependent gene looping enhances transcription by facilitating initiation, reinitiation and promoter directionality. **(A)** A gene with an intron (yellow) and exons (blue). **(B)** During transcription, an elongating RNAPII transcribes the gene and recruits splicing factors (SF) to splice out the intervening intronic sequence. Splicing components (SF) at the 5′ splice site interact with the general transcription factors (GTFs) at the promoter forming a loop between the promoter and 5′ splice site. **(C)** Once RNAPII has transcribed the intron, splicing components at the 3′ splice site associate with the termination factors near the 3′ end of the gene forming a loop between the 3′ splice site and terminator. Finally, the gene forms an overall three-looped conformation where the promoter and terminator physically interact with one another to assist in reinitiation. The three unique interactions that take place are between the promoter-terminator, promoter-5′ splice site and terminator-3′ splice site. GTFs, general transcription factors; SFs, splicing factors; TF, termination factors.

Introns have also been found to stimulate transcription at the elongation as well as termination steps of transcription. Many independent studies have found U1-snRNP, SKIP and SC35, which are splicing factors, stimulating elongation by interacting with various transcription elongation factors (Fong and Zhou, 2001; Brès et al., 2005; Lin et al., 2008). Introns also influence elongation indirectly by affecting chromatin structure. The H3-K36 trimethylation mark, which is specifically associated with elongation of transcription, is enriched in intron-containing genes compared to the intron-less genes (**Figure 1B**; De Almeida et al., 2011). There are also a few reports of terminator-proximal introns influencing transcription by affecting termination of transcription either directly by helping in the recruitment of termination factors or indirectly by affecting the chromatin structure near the 3′ end of genes (Lutz et al., 1996; McCracken et al., 2002). The 3′ splice site of a terminator-proximal intron has been shown to enhance utilization of a downstream poly(A) site by facilitating recruitment of the F 3′ end processing complex and poly(A) polymerase in mammalian cell lines (**Figure 1B**; Lutz et al., 1996; McCracken et al., 2002). The 3′ splice site also has an adverse effect on H3-K36 trimethylation, which is a transcription elongation mark that needs to be removed to facilitate termination of transcription (Kim et al., 2011). Although introns may target the elongation and termination steps to enhance transcription, the initiation step is emerging as the most frequent target.

INTRON-DEPENDENT LOOPED GENE ARCHITECTURE FACILITATES ENHANCEMENT OF TRANSCRIPTION

The classical view of transcriptional regulation by *cis*-acting regulatory sequences and *trans*-acting protein factors has

undergone radical changes due to research carried out during the last few decades (Papantonis and Cook, 2010; Ansari, 2019; Al-Husini et al., 2020). Genome topology or chromatin conformations formed by enhancer-promoter interactions and promoter-terminator interactions have been found to play a crucial role in regulation of transcription (Misteli, 2007). The physical interaction of the promoter and terminator regions of a gene during transcription results in the formation of a looped gene architecture (Ansari and Hampsey, 2005). Such gene loops are formed in an activator-dependent manner and have been observed in yeast as well as in higher eukaryotes (reviewed in Al-Husini et al., 2020). Activator-dependent gene looping has been shown to enhance transcription by facilitating direct transfer of polymerase from the terminator to the promoter for reinitiation, and by enhancing promoter directionality (Tan-Wong et al., 2012; Al Husini et al., 2013).

In budding yeast, the presence of an intron in a gene also results in a looped gene architecture (**Figure 2C**; Moabbi et al., 2012; Tan-Wong et al., 2012; Agarwal and Ansari, 2016). Intron-dependent gene looping, however, is mechanistically different from the activator-dependent looped conformation. The "Chromosome Conformation Capture" (3C) approach revealed that an intron-dependent loop is characterized by additional contacts of the promoter with the 5′ splice site and of the terminator with 3′ splice site (**Figures 2B,C**; Moabbi et al., 2012). The promoter-5′ splice site loop may not be observed in all genes as the first intron often is located very close to the promoter and can make direct contact with the promoter without loop formation. How the presence of an intron facilitates gene loop formation, however, is not yet clear. It has been proposed that the interaction of a 5′ splice site with the promoter and of a 3′ splice site with terminator bring the two ends of a gene in close physical proximity and facilitate the promoter-terminator contact (Al-Husini et al., 2020). Only a splicing-competent intron facilitates gene loop formation (Moabbi et al., 2012; Agarwal and Ansari, 2016). Mutation of either the 5′ or the 3′ splice sites abolishes intron-facilitated looped gene structure. Since the mutation of splice sites also adversely affects IME of transcription, it was proposed that intron-mediated gene looping may also enhance transcription in a manner similar to activator-dependent enhancement. Examination of the IME effect in the looping-defective mutants of yeast revealed that although splicing was normal, there was no enhancement of transcription (Moabbi et al., 2012; Agarwal and Ansari, 2016). The IME effect therefore is not due to splicing *per se* but due to the formation of a splicing-dependent looped gene architecture in budding yeast.

Evidence suggests that the three contact points in an intron-dependent looped gene structure; promoter-terminator contact, promoter-5′ splice site contact and terminator-3′ splice site (**Figure 2C**), are established by a protein-protein interaction of factors occupying promoter, terminator and intronic sites. A combination of ChIP and 3C approaches identified the crucial role of the general transcription factor TFIIB and CF1A termination complex in the promoter-terminator interaction (Medler et al., 2011). Similarly, the interaction of another general transcription factor TFIIH with U1-snRNP could lead to promoter-5′ splice site contact (Kwek et al., 2002). The interactions contributing to terminator-3′ splice site contact are yet to be established. Splicing factors do not contact DNA directly, but proximity of splice sites on RNA with the corresponding DNA region during cotranscriptional splicing results in splicing factors getting crosslinked to the splice sites on DNA as well (Kotovic et al., 2003; Lin et al., 2008; Oesterreich et al., 2016; Minocha et al., 2018; Nojima et al., 2018). This evidence suggests that the intron-mediated gene loop formation is due to an interaction of the initiation and termination factors occupying distal ends of a gene with splicing factors bound to the intronic regions.

A critical issue is how intron-dependent looped gene architecture brings about enhancement of transcription. All three of the physical interactions in an intron-facilitated gene loop have the potential to enhance transcription. The promoter-5′ splice site interaction can facilitate initiation/reinitiation by stabilizing the assembly of the PIC (**Figure 2B**). U1-snRNP, which binds the 5′ splice site and exhibits an interaction with TFIIH, may play a crucial role in this regard (Kwek et al., 2002). U1-snRNP may directly help in the recruitment of TFIIH, and possibly TFIIB and TFIID on the promoter if the 5′ splice site is located in close proximity to the promoter (Damgaard et al., 2008). The 5′ splice sites, however, may be located several hundred nucleotides away from the promoter. In such a scenario, a loop formed by promoter-5′ splice site interaction may play a critical role in assembly or stabilization of the PIC on the promoter (**Figure 2B**). The net result will be enhanced initiation or reinitiation of transcription. This may explain why a functional 5′ splice site alone could bring about an increase in transcription of HIV-1 and β-globin genes (Damgaard et al., 2008). The enhancement of transcription elicited by a 5′ splice site alone, however, was much lower (75% less) compared to that brought about by a full-length intron. The interaction of the 5′ splice site with the promoter is therefore not sufficient to achieve enhancement of transcription fully. Contacts of the promoter with the terminator and of the 3′ splice site with the terminator may contribute significantly to the IME effect (**Figure 2C**). Thus, all three interactions in an intron-dependent gene loop may play a role in the enhancement of transcription.

The promoter-terminator interaction is especially crucial as it can enhance transcription by facilitating reinitiation and by conferring promoter directionality (Tan-Wong et al., 2012; Al Husini et al., 2013). It is conceivable that proximity of promoter and terminator in the intron-dependent gene loop may similarly contribute to the enhancement of transcription by a similar mechanism (**Figure 2C**). Juxtaposition of the terminator and promoter facilitates release of polymerase from the terminator region near the promoter. The polymerase is then recycled back to the juxtaposed promoter for reinitiation of transcription (Al Husini et al., 2013). Such a coupling of termination to reinitiation, with a concomitant increase in the transcriptional activity, has been observed for RNAPIII, RNAPI, mitochondrial polymerase and archaeal polymerase (Dieci and Sentenac, 1996; Jansa et al., 2001; Martin et al., 2005; Spitalny and Thomm, 2008). The promoter-terminator interaction is emerging as a critical player in regulation of transcription in eukaryotes.

In budding yeast, juxtaposition of the promoter and terminator also confers promoter directionality, which is the enhancement of promoter-initiated transcription of mRNA while keeping the upstream antisense transcription in check (Tan-Wong et al., 2012; Al Husini et al., 2013). The presence of an intron also enhances promoter directionality of yeast genes (Agarwal and Ansari, 2016). In the absence of a splicing-competent intron, mRNA synthesis exhibits a decline, while uaRNA (upstream antisense RNA) transcription increases. In a looping-defective strain, despite the presence of a splicing-competent intron and normal splicing, promoter directionality is adversely affected. The proximity of the promoter and terminator in the intron-mediated gene loop allows the termination factors bound to the 3′ end to contact the 5′ end of a gene and bring about termination of uaRNA transcription (Agarwal and Ansari, 2016). The net result is higher transcription of the gene by enhancing promoter directionality.

Although, the terminator-3′ splice site interaction can promote cotranscriptional recruitment of termination factors, leading to efficient termination of transcription, its role in IME effect needs further exploration. It is, however, clear from the studies described above that intron-dependent looped gene architecture likely plays a crucial role in enhancement of transcription in budding yeast.

SPLICING-DEPENDENT GENE LOOPING IN HIGHER EUKARYOTES

In budding yeast, merely 4% of genes contain introns, but these few genes produce more than 25% of total cellular mRNA (Ares et al., 1999). On average, yeast intron-containing genes produce 3.7 times more mRNA than their non-intronic counterparts (Juneau et al., 2006). In higher eukaryotes, a far higher proportion of genes harbor introns, and splicing-dependent activation of transcription is prevalent in higher eukaryotes as well (reviewed in Rose, 2008; Laxa, 2017; Shaul, 2017; Rose, 2019). This raises the question if splicing-associated changes in gene architecture also occur in higher eukaryotes, and if they contribute to enhancement of transcription.

Chromatin interaction analyses have revealed the presence of physical interactions of promoters with their gene body in mammalian systems. 3C analysis of the human BRCA1 gene revealed an interaction of the promoter and terminator regions of the gene with the intronic regions in a transcription-dependent manner (Tan-Wong et al., 2008). Intragenic gene loops formed by the interaction of exons with cognate promoters have been identified in human cell lines on a genomewide scale using the ChIA-PET (Chromatin Interaction Analysis with Paired-End Tag) approach (Mercer et al., 2013). Hi-C analysis corroborated the presence of intragenic chromatin loops formed by the interaction of promoters with exons in the human genome (Ruiz-Velasco et al., 2017). This genomewide study also demonstrated intragenic gene loops between exons and the 3′ ends of genes. This is reminiscent of the intron-terminator interaction observed in budding yeast and the human BRCA1 gene (Tan-Wong et al., 2008; Moabbi et al., 2012). ChIA-PET identified CTCF (CCCTC-binding factor) as the protein that facilitates the interaction of promoter with exons (Mercer et al., 2013). It was further shown that the CTCF-mediated promoter-exon loops are prevalent in genes coding for proteins involved in cell signaling and response to stimuli. The promoter-exon loop formation is accompanied by trimethylation of histone H3-K4, acetylation of histone H3-K27 and trimethylation of histone H3-K36 in the coding region of the gene (Ruiz-Velasco et al., 2017). These three histone modifications are associated with transcriptionally active chromatin (Bannister and Kouzarides, 2011). Modification of histone H3-K4 and H3K9 by methylation and acetylation have been implicated in enhancement of transcription by the promoter-proximal intron in mammalian cells thereby suggesting that the promoter-exon interaction may enhance transcription by affecting chromatin structure (Bieberstein et al., 2012). The promoter-exon interaction may be the mammalian structural equivalent of yeast intron-mediated gene loops.

The promoter-exon interaction may be responsible for recently reported internal exon-mediated activation of transcription in mammalian cells (Fiszbein et al., 2019). Internal exons are small exons generally less than 300 nucleotides in length and are flanked by at least one exon on the 5′ side and one exon on the 3′ side. They are responsible for transcriptional regulation of thousands of mammalian genes. The phenomenon is called EMAT (exon-mediated activation of transcription). Activation of transcription by internal exons is not dependent on the sequence of the exon, but on the splice sites flanking the intron. Mutation of either the 5′ or 3′ splice site completely abrogated transcription activation potential of an internal exon (Fiszbein et al., 2019). Furthermore, exon-mediated activation involved recruitment of the general transcription factor TFIIF on the promoter by the splicing factor HNRNPU (heterogeneous nuclear ribonucleoprotein U), which is known to interact with TFIIF through its N-terminal domain (Kim and Nikodem, 1999). In addition, the exon-promoter interaction is accompanied by trimethylation and acetylation of histone H3, which may also result in enhanced transcription by affecting both transcription initiation and elongation (Bieberstein et al., 2012; Ruiz-Velasco et al., 2017).

The exon-mediated transcriptional regulation in mammalian cells exhibits striking similarities to the intron-mediated regulation in budding yeast: (1) internal exon or intron must be located within one kbp of the promoter to activate transcription; (2) transcription activation by both occurs in a splicing-dependent manner; and (3) transcriptional regulation in both cases involves splicing-dependent recruitment of general transcription factors on the promoter. The possibility of exon-mediated transcriptional activation through intragenic gene loops formed by the interaction of a promoter with the gene body therefore cannot be ruled out. Both intron-mediated and exon-mediated enhancement of transcription are in fact splicing-mediated regulations. Gene architecture playing a general role in splicing-mediated regulation of transcription is an attractive possibility. The studies from yeast strongly suggest the involvement of gene looping in the splicing-mediated regulation, but evidence from higher eukaryotes are still preliminary and needs further investigation. Nevertheless, it is clear that introns

are not merely junk coding sequence but are an important regulator of cellular functions.

AUTHOR CONTRIBUTIONS

AA conceptualized. KD and NA prepared the original draft. KD made figures and helped with editing. AA and LP edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by NSF grant MCB-1936030.

ACKNOWLEDGMENTS

We would like to thank Michael O'Brien and Zulifa Khan for critical reading of the manuscript.

REFERENCES

Agarwal, N., and Ansari, A. (2016). Enhancement of transcription by a splicing-competent intron is dependent on promoter directionality. *PLoS Genet.* 12:e1006047. doi: 10.1371/journal.pgen.1006047

Al Husini, N., Kudla, P., and Ansari, A. (2013). A role for CF1A 3′ end processing complex in promoter-associated transcription. *PLoS Genet.* 9:e1003722. doi: 10.1371/journal.pgen.1003722

Al-Husini, N., Medler, S., and Ansari, A. (2020). Crosstalk of promoter and terminator during RNA polymerase II transcription cycle. *Biochim. Biophys. Acta Gene Regulat. Mechanis.* 1863:194657. doi: 10.1016/j.bbagrm.2020.194657

Ansari, A. (2019). Recent Trends in Eukaryotic Transciption: Crucial Role of Gene Architecture in Transcriptional Regulation. *J. Cytol. Mol. Biol.* 4:2.

Ansari, A., and Hampsey, M. (2005). A role for the CPF 3′-end processing machinery in RNAP II-dependent gene looping. *Genes Dev.* 19, 2969–2978. doi: 10.1101/gad.1362305

Ares, M., Grate, L., and Pauling, M. H. (1999). A handful of intron-containing genes produces the lion's share of yeast mRNA. *RNA* 5, 1138–1139. doi: 10.1017/s1355838299991379

Baier, T., Jacobebbinghaus, N., Einhaus, A., Lauersen, K. J., and Kruse, O. (2020). Introns mediate post-transcriptional enhancement of nuclear gene expression in the green microalga Chlamydomonas reinhardtii. *PLoS Genet.* 16:e1008944. doi: 10.1371/journal.pgen.1008944

Bannister, A. J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395. doi: 10.1038/cr.2011.22

Beaulieu, E., Green, L., Elsby, L., Alourfi, Z., Morand, E. F., Ray, D. W., et al. (2011). Identification of a novel cell type-specific intronic enhancer of macrophage migration inhibitory factor (MIF) and its regulation by mithramycin. *Clin. Exp. Immunol.* 163, 178–188. doi: 10.1111/j.1365-2249.2010.04289.x

Bianchi, M., Crinelli, R., Giacomini, E., Carloni, E., and Magnani, M. (2009). A potent enhancer element in the 5′-UTR intron is crucial for transcriptional regulation of the human ubiquitin C gene. *Gene* 448, 88–101. doi: 10.1016/j.gene.2009.08.013

Bieberstein, N. I., Oesterreich, F. C., Straube, K., and Neugebauer, K. M. (2012). First exon length controls active chromatin signatures and transcription. *Cell Rep.* 2, 62–68. doi: 10.1016/j.celrep.2012.05.019

Bonnet, A., Grosso, A. R., Elkaoutari, A., Coleno, E., Presle, A., Sridhara, S. C., et al. (2017). Introns protect eukaryotic genomes from transcription-associated genetic instability. *Mol. Cell* 67, 608–621. doi: 10.1016/j.molcel.2017.07.002

Brès, V., Gomes, N., Pickle, L., and Jones, K. A. (2005). A human splicing factor, SKIP, associates with P-TEFb and enhances transcription elongation by HIV-1 Tat. *Genes Dev.* 19, 1211–1226. doi: 10.1101/gad.1291705

Brinster, R. L., Allen, J. M., Behringer, R. R., Gelinas, R. E., and Palmiter, R. D. (1988). Introns increase transcriptional efficiency in transgenic mice. *Proc. Natl. Acad. Sci.* 85, 836–840. doi: 10.1073/pnas.85.3.836

Callis, J., Fromm, M., and Walbot, V. (1987). Introns increase gene expression in cultured maize cells. *Genes Dev.* 1, 1183–1200. doi: 10.1101/gad.1.10.1183

Carmel, L., and Chorev, M. (2012). The function of introns. *Front. Genet.* 3:55. doi: 10.3389/fgene.2012.00055

Charron, M., Chern, J. Y., and Wright, W. W. (2007). The cathepsin L first intron stimulates gene expression in rat sertoli cells. *Biol. Reproduct.* 76, 813–824. doi: 10.1095/biolreprod.106.057851

Chaudhary, S., Khokhar, W., Jabre, I., Reddy, A. S., Byrne, L. J., Wilson, C. M., et al. (2019). Alternative splicing and protein diversity: plants versus animals. *Front. Plant Sci.* 10:708. doi: 10.3389/fpls.2019.00708

Damgaard, C. K., Kahns, S., Lykke-Andersen, S., Nielsen, A. L., Jensen, T. H., and Kjems, J. (2008). A 5′ splice site enhances the recruitment of basal transcription initiation factors in vivo. *Mol. Cell* 29, 271–278. doi: 10.1016/j.molcel.2007.11.035

Das, R., Yu, J., Zhang, Z., Gygi, M. P., Krainer, A. R., Gygi, S. P., et al. (2007). SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol. Cell* 26, 867–881. doi: 10.1016/j.molcel.2007.05.036

De Almeida, S. F., Grosso, A. R., Koch, F., Fenouil, R., Carvalho, S., Andrade, J., et al. (2011). Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. *Nat. Struct. Mol. Biol.* 18:977. doi: 10.1038/nsmb.2123

Dieci, G., and Sentenac, A. (1996). Facilitated recycling pathway for RNA polymerase III. *Cell* 84, 245–252. doi: 10.1016/s0092-8674(00)80979-4

Ding, F., and Elowitz, M. B. (2019). Constitutive splicing and economies of scale in gene expression. *Nat. Struct. Mol. Biol.* 26, 424–432. doi: 10.1038/s41594-019-0226-x

Faustino, N. A., and Cooper, T. A. (2003). Pre-mRNA splicing and human disease. *Genes Dev.* 17, 419–437. doi: 10.1101/gad.1048803

Fiszbein, A., Krick, K. S., Begg, B. E., and Burge, C. B. (2019). Exon-mediated activation of transcription starts. *Cell* 179, 1551–1565. doi: 10.1016/j.cell.2019.11.002

Fong, Y. W., and Zhou, Q. (2001). Stimulatory effect of splicing factors on transcriptional elongation. *Nature* 414, 929–933. doi: 10.1038/414929a

Furger, A., O'Sullivan, J. M., Binnie, A., Lee, B. A., and Proudfoot, N. J. (2002). Promoter proximal splice sites enhance transcription. *Genes Dev.* 16, 2792–2799. doi: 10.1101/gad.983602

Gallegos, J. E., and Rose, A. B. (2015). The enduring mystery of intron-mediated enhancement. *Plant Sci.* 237, 8–15. doi: 10.1016/j.plantsci.2015.04.017

Gallegos, J. E., and Rose, A. B. (2017). Intron DNA sequences can be more important than the proximal promoter in determining the site of transcript initiation. *Plant Cell* 29, 843–853. doi: 10.1105/tpc.17.00020

Huang, J., and Liang, T. J. (1993). A novel hepatitis B virus (HBV) genetic element with Rev response element-like properties that is essential for expression of HBV gene products. *Mol. Cell. Biol.* 13, 7476–7486. doi: 10.1128/mcb.13.12.7476

Huang, Z. M., and Yen, T. S. (1995). Role of the hepatitis B virus posttranscriptional regulatory element in export of intronless transcripts. *Mol. Cell. Biol.* 15, 3864–3869. doi: 10.1128/mcb.15.7.3864

Jansa, P., Burek, C., Sander, E. E., and Grummt, I. (2001). The transcript release factor PTRF augments ribosomal gene transcription by facilitating reinitiation of RNA polymerase I. *Nucleic Acids Res.* 29, 423–429. doi: 10.1093/nar/29.2.423

Jo, B. S., and Choi, S. S. (2015). Introns: the functional benefits of introns in genomes. *Genomics Informat.* 13:112. doi: 10.5808/gi.2015.13.4.112

Jobert, L., Pinzon, N., Van Herreweghe, E., Jady, B. E., Guialis, A., Kiss, T., et al. (2009). Human U1 snRNA forms a new chromatin-associated snRNP with TAF15. *EMBO Rep.* 10, 494–500. doi: 10.1038/embor.2009.24

Juneau, K., Miranda, M., Hillenmeyer, M. E., Nislow, C., and Davis, R. W. (2006). Introns regulate RNA and protein abundance in yeast. *Genetics* 174, 511–518. doi: 10.1534/genetics.106.058560

Kaneda, S., Horie, N., Takeishi, K., Takayanagi, A., Seno, T., and Ayusawa, D. (1992). Regulatory sequences clustered at the 5′ end of the first intron of the human thymidylate synthase gene function in cooperation with the promoter region. *Somat. Cell Mol. Genet.* 18, 409–415. doi: 10.1007/bf01233080

Kim, M. K., and Nikodem, V. M. (1999). hnRNP U inhibits carboxy-terminal domain phosphorylation by TFIIH and represses RNA polymerase II elongation. *Mol. Cell. Biol.* 19, 6833–6844. doi: 10.1128/mcb.19.10.6833

Kim, S., Kim, H., Fong, N., Erickson, B., and Bentley, D. L. (2011). Pre-mRNA splicing is a determinant of histone H3K36 methylation. *Proc. Natl. Acad. Sci.* 108, 13564–13569. doi: 10.1073/pnas.1109475108

Kotovic, K. M., Lockshon, D., Boric, L., and Neugebauer, K. M. (2003). Cotranscriptional recruitment of the U1 snRNP to intron-containing genes in yeast. *Mol. Cell. Biol.* 23, 5768–5779. doi: 10.1128/mcb.23.16.5768-5779.2003

Kwek, K. Y., Murphy, S., Furger, A., Thomas, B., O'Gorman, W., Kimura, H., et al. (2002). U1 snRNA associates with TFIIH and regulates transcriptional initiation. *Nat. Struct. Biol.* 9, 800–805.

Laxa, M. (2017). Intron-mediated enhancement: a tool for heterologous gene expression in plants? *Front. Plant Sci.* 7:1977. doi: 10.3389/fpls.2016.01977

Le Hir, H., Nott, A., and Moore, M. J. (2003). How introns influence and enhance eukaryotic gene expression. *Trends Biochem. Sci.* 28, 215–220. doi: 10.1016/s0968-0004(03)00052-5

Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S., and Fu, X. D. (2008). The splicing factor SC35 has an active role in transcriptional elongation. *Nat. Struct. Mol. Biol.* 15:819. doi: 10.1038/nsmb.1461

Liu, X., and Mertz, J. E. (1995). HnRNP L binds a cis-acting RNA sequence element that enables intron-dependent gene expression. *Genes Dev.* 9, 1766–1780. doi: 10.1101/gad.9.14.1766

Lutz, C. S., Murthy, K. G., Schek, N., O'Connor, J. P., Manley, J. L., and Alwine, J. C. (1996). Interaction between the U1 snRNP-A protein and the 160-kD subunit of cleavage-polyadenylation specificity factor increases polyadenylation efficiency in vitro. *Genes Dev.* 10, 325–337. doi: 10.1101/gad.10.3.325

Malim, M. H., Hauber, J., Fenrick, R., and Cullen, B. R. (1988). Immunodeficiency virus rev trans-activator modulates the expression of the viral regulatory genes. *Nature* 335, 181–183. doi: 10.1038/335181a0

Marquez, Y., Höpfler, M., Ayatollahi, Z., Barta, A., and Kalyna, M. (2015). Unmasking alternative splicing inside protein-coding exons defines exitrons and their role in proteome plasticity. *Genome Res.* 25, 995–1007. doi: 10.1101/gr.186585.114

Martin, M., Cho, J., Cesare, A. J., Griffith, J. D., and Attardi, G. (2005). Termination factor-mediated DNA loop between termination and initiation sites drives mitochondrial rRNA synthesis. *Cell* 123, 1227–1240. doi: 10.1016/j.cell.2005.09.040

McCracken, S., Lambermon, M., and Blencowe, B. J. (2002). SRm160 splicing coactivator promotes transcript 3′-end cleavage. *Mol. Cell. Biol.* 22, 148–160. doi: 10.1128/mcb.22.1.148-160.2002

Medler, S., Al Husini, N., Raghunayakula, S., Mukundan, B., Aldea, A., and Ansari, A. (2011). Evidence for a complex of transcription factor IIB (TFIIB) with Poly (A) polymerase and cleavage factor 1 subunits required for gene looping. *J. Biol. Chem.* 286, 33709–33718. doi: 10.1074/jbc.m110.193870

Mercer, T. R., Edwards, S. L., Clark, M. B., Neph, S. J., Wang, H., Stergachis, A. B., et al. (2013). DNase I–hypersensitive exons colocalize with promoters and distal regulatory elements. *Nat. Genet.* 45, 852–859. doi: 10.1038/ng.2677

Millhouse, S., and Manley, J. L. (2005). The C-terminal domain of RNA polymerase II functions as a phosphorylation-dependent splicing activator in a heterologous protein. *Mol. Cell. Biol.* 25, 533–544. doi: 10.1128/mcb.25.2.533-544.2005

Minocha, R., Popova, V., Kopytova, D., Misiak, D., Hüttelmaier, S., Georgieva, S., et al. (2018). Mud2 functions in transcription by recruiting the Prp19 and TREX complexes to transcribed genes. *Nucleic Acids Res.* 46, 9749–9763. doi: 10.1093/nar/gky640

Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. *Cell* 128, 787–800. doi: 10.1016/j.cell.2007.01.028

Moabbi, A. M., Agarwal, N., El Kaderi, B., and Ansari, A. (2012). Role for gene looping in intron-mediated enhancement of transcription. *Proc. Natl. Acad. Sci.* 109, 8505–8510. doi: 10.1073/pnas.1112400109

Morgan, J. T., Fink, G. R., and Bartel, D. P. (2019). Excised linear introns regulate growth in yeast. *Nature* 565, 606–611. doi: 10.1038/s41586-018-0828-1

Niehrs, C., and Luke, B. (2020). Regulatory R-loops as facilitators of gene expression and genome stability. *Nat. Rev. Mol. Cell Biol.* 21, 167–178. doi: 10.1038/s41580-019-0206-3

Nilsen, T. W., and Graveley, B. R. (2010). Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463, 457–463. doi: 10.1038/nature08909

Nojima, T., Rebelo, K., Gomes, T., Grosso, A. R., Proudfoot, N. J., and Carmo-Fonseca, M. (2018). RNA polymerase II phosphorylated on CTD serine 5 interacts with the spliceosome during co-transcriptional splicing. *Mol. Cell* 72, 369–379. doi: 10.1016/j.molcel.2018.09.004

Oesterreich, F. C., Herzel, L., Straube, K., Hujer, K., Howard, J., and Neugebauer, K. M. (2016). Splicing of nascent RNA coincides with intron exit from RNA polymerase II. *Cell* 165, 372–381. doi: 10.1016/j.cell.2016.02.045

Padgett, R. A. (2012). New connections between splicing and human disease. *Trends Genet.* 28, 147–154. doi: 10.1016/j.tig.2012.01.001

Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S., and Sharp, P. A. (1986). Splicing of messenger RNA precursors. *Annu. Rev. Biochem.* 55, 1119–1150.

Palmiter, R. D., Sandgren, E. P., Avarbock, M. R., Allen, D. D., and Brinster, R. L. (1991). Heterologous introns can enhance expression of transgenes in mice. *Proc. Natl. Acad. Sci.* 88, 478–482. doi: 10.1073/pnas.88.2.478

Papantonis, A., and Cook, P. R. (2010). Genome architecture and the role of transcription. *Curr. Opin. Cell Biol.* 22, 271–276. doi: 10.1016/j.ceb.2010.03.004

Parenteau, J., Maignon, L., Berthoumieux, M., Catala, M., Gagnon, V., and Abou Elela, S. (2019). Introns are mediators of cell response to starvation. *Nature* 565, 612–617. doi: 10.1038/s41586-018-0859-7

Pasleau, F., Leung, F., and Kopchick, J. J. (1987). A comparison of bovine growth hormone expression directed by bGH genomic or intronless DNA in transiently transfected eukaryotic cells. *Gene* 57, 47–52. doi: 10.1016/0378-1119(87)90175-2

Rogozin, I. B., Carmel, L., Csuros, M., and Koonin, E. V. (2012). Origin and evolution of spliceosomal introns. *Biol. Direct* 7, 1–28.

Rose, A. B. (2004). The effect of intron location on intron-mediated enhancement of gene expression in Arabidopsis. *Plant J.* 40, 744–751. doi: 10.1111/j.1365-313x.2004.02247.x

Rose, A. B. (2008). Intron-mediated regulation of gene expression. *Nuclear Pre mRNA Process. Plants* 326, 277–290. doi: 10.1007/978-3-540-76776-3_15

Rose, A. B. (2019). Introns as gene regulators: a brick on the accelerator. *Front. Genet.* 9:672. doi: 10.3389/fgene.2018.00672

Ruiz-Velasco, M., Kumar, M., Lai, M. C., Bhat, P., Solis-Pinson, A. B., Reyes, A., et al. (2017). CTCF-mediated chromatin loops between promoter and gene body regulate alternative splicing across individuals. *Cell Syst.* 5, 628–637. doi: 10.1016/j.cels.2017.10.018

Shaul, O. (2017). How introns enhance gene expression. *Int. J. Biochem. Cell Biol.* 91, 145–155. doi: 10.1016/j.biocel.2017.06.016

Spitalny, P., and Thomm, M. (2008). A polymerase III-like reinitiation mechanism is operating in regulation of histone expression in archaea. *Mol. Microbiol.* 67, 958–970. doi: 10.1111/j.1365-2958.2007.06084.x

Tan-Wong, S. M., French, J. D., Proudfoot, N. J., and Brown, M. A. (2008). Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene. *Proc. Natl. Acad. Sci.* 105, 5160–5165. doi: 10.1073/pnas.0801048105

Tan-Wong, S. M., Zaugg, J. B., Camblong, J., Xu, Z., Zhang, D. W., Mischo, H. E., et al. (2012). Gene loops enhance transcriptional directionality. *Science* 338, 671–675. doi: 10.1126/science.1224350

Tian, H. (2001). RNA ligands generated against complex nuclear targets indicate a role for U1 snRNP in co-ordinating transcription and RNA splicing. *FEBS Lett.* 509, 282–286. doi: 10.1016/s0014-5793(01)03188-x

Wang, G. S., and Cooper, T. A. (2007). Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat. Rev. Genet.* 8, 749–761. doi: 10.1038/nrg2164

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Molecular Biosciences

MINI REVIEW
published: 21 April 2021
doi: 10.3389/fmolb.2021.663209

Eukaryotic RNA Polymerases: The Many Ways to Transcribe a Gene

*Marina Barba-Aliaga[1,2], Paula Alepuz[1,2] and José E. Pérez-Ortín[1,2]**

[1] Instituto de Biotecnología y Biomedicina (Biotecmed), Universitat de València, València, Spain, [2] Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias Biológicas, Universitat de València, València, Spain

In eukaryotic cells, three nuclear RNA polymerases (RNA pols) carry out the transcription from DNA to RNA, and they all seem to have evolved from a single enzyme present in the common ancestor with archaea. The multiplicity of eukaryotic RNA pols allows each one to remain specialized in the synthesis of a subset of transcripts, which are different in the function, length, cell abundance, diversity, and promoter organization of the corresponding genes. We hypothesize that this specialization of RNA pols has conditioned the evolution of the regulatory mechanisms used to transcribe each gene subset to cope with environmental changes. We herein present the example of the homeostatic regulation of transcript levels *versus* changes in cell volume. We propose that the diversity and instability of messenger RNAs, transcribed by RNA polymerase II, have conditioned the appearance of regulatory mechanisms based on different gene promoter strength and mRNA stability. However, for the regulation of ribosomal RNA levels, which are very stable and transcribed mainly by RNA polymerase I from only one promoter, different mechanisms act based on gene copy variation, and a much simpler regulation of the synthesis rate.

Keywords: RNA pol I, RNA pol II, transcription, nucleus, evolution, RNA pol III

OPEN ACCESS

Edited by:
Athar Ansari,
Wayne State University, United States

Reviewed by:
Giorgio Dieci,
University of Parma, Italy
Christoph Müller,
European Molecular Biology Laboratory Heidelberg, Germany

***Correspondence:**
José E. Pérez-Ortín
jose.e.perez@uv.es

Specialty section:
This article was submitted to Protein and RNA Networks, a section of the journal Frontiers in Molecular Biosciences

Received: *02 February 2021*
Accepted: *09 March 2021*
Published: *21 April 2021*

Citation:
Barba-Aliaga M, Alepuz P and Pérez-Ortín JE (2021) Eukaryotic RNA Polymerases: The Many Ways to Transcribe a Gene. Front. Mol. Biosci. 8:663209. doi: 10.3389/fmolb.2021.663209

INTRODUCTION

A key step in the central dogma of molecular biology is the transcription of pieces of DNA information into RNA molecules, which will, in some cases, be translated into proteins but will remain, in other cases, as functional non-coding RNAs (ncRNAs). In all living systems, the transcription of cellular genomes is carried out by cellular multisubunit DNA-dependent RNA polymerases (RNA pols). Eubacteria and archaea possess a single such enzyme, while eukaryotes carry out nuclear transcription with at least three RNA pols with functional specialization by each one transcribing different non-overlapping subsets of genes. Although all these enzymes have originated from a common ancestral enzyme, the increasing complexity of genomes, cells, and organisms has imposed the evolution of transcription machineries to more sophisticated systems in terms of composition, interactions, selection of target genes, and regulation. In this mini review, we summarize the presumed evolutionary origin and functional reasons that have led to the multiplicity of nuclear RNA pols in eukaryotes, and its consequences for their regulation and the homeostasis of their different RNA products. We finally focus on the different adaptation of transcription regulation by eukaryotic RNA pols to changes in cellular volume. Other eukaryotic RNA pols aspects have been extensively reviewed elsewhere (Hahn, 2004; Dieci et al., 2007; Cramer et al., 2008; Werner and Grohmann, 2011; Engel et al., 2013; Cramer, 2019).

THE EVOLUTIONARY SCHEME OF RNA POL FROM EUBACTERIA TO EUKARYOTES

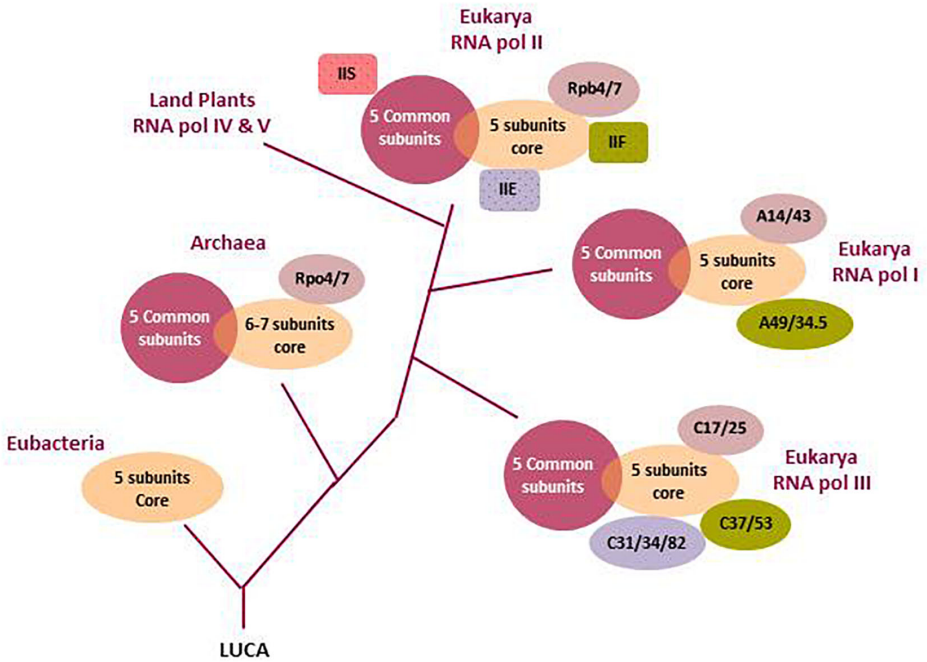
All RNA pols, from eubacteria to higher eukaryotes, share basic mechanistic functioning: use of a DNA template, processive translocation on the template during RNA synthesis, utilization of ribonucleoside triphosphate as substrates, Watson–Crick base pairing of the new added nucleotide with the complementary one in the template DNA, and formation of a new phosphodiester bound by a metal-dependent mechanism. To perform these basic functions, all RNA pols contain two largest subunits (**Figure 1**) with double-ψ β-barrel motifs that create an active site at the interface of the subunits with three key aspartic residues conserved across all domains of life. Additionally, multisubunit RNA pols contain a variable number of additional smaller subunits (**Figure 1**). The two largest catalytic subunits of RNA pols are thought to have evolved from the duplication and diversification of a gene that encoded a protein cofactor of a common ancestral ribozyme, which performed RNA polymerase activity in the primal RNA world (Iyer et al., 2003). At some point of evolution, the new protein heterodimer would have gained polymerase activity and acquired different subunits with specialized assembly and auxiliary functions. Thus, all multisubunit RNA pols share a common structural core and similar basic molecular mechanisms and must derive from the RNA pol of the last universal common ancestor (LUCA) of archaea, eubacteria, and eukaryotes, assumed to have existed 3.5–3.8 billion years ago (Burton, 2014). This ancestral multisubunit RNA pol was probably similar to the simple RNA pol found today in eubacteria, which is formed (see **Figure 1**) by two large β and β' catalytic subunits, two assembly subunits (2α), and one auxiliary subunit (ω), as all these five subunits are highly conserved in the structure/function of all organisms (Werner, 2007; Werner and Grohmann, 2011).

RNA pol gained greater complexity in terms of acquiring new subunits following the split of the eubacterial and archaeal–eukaryotic branches from the universal tree of life (Werner, 2007; Spang et al., 2015). Archaeal RNA pol has three or four catalytic polypeptides and three assembly and auxiliary subunits, which are closely related to bacterial ones (**Figure 1**). However, archaeal RNA pol has gained five additional periphery subunits with no homologs in eubacteria but resembling eukaryotic subunits, which stabilize the interactions of polymerase with template DNA, newly synthesized RNA, and different transcription factors to ensure efficient functioning in the transcription cycle (Werner, 2007; Werner and Grohmann, 2011; Fouqueau et al., 2017). The more complex transcription machineries of archaea and eukaryotes are linked with the fact that their genomes, which differ from the eubacterial genome, are stabilized and compacted by histone or histone-like proteins that impose more restrictive access to DNA and the need for additional basal transcription factors (Reeve, 2003; Geiduschek and Ouhammouch, 2005; Kwapisz et al., 2008; Jun et al., 2011; Werner and Grohmann, 2011; Koster et al., 2015).

Archaeal and eukaryotic lineages diverged more than 2 billion years ago, with eukaryotes originating from an archaeal linage with already diverse eukaryotic signature proteins (Spang et al., 2015). Other important differences include that eukaryotes have an extended system of intracellular membranes that compartmentalizes the intracellular space, and the cellular volume is three to four orders of magnitude larger than that of archaea and bacteria (Lane and Martin, 2010; Koonin, 2015). They also contain organelles (mitochondria and chloroplasts) that derive from two kinds of eubacteria and have their own RNA pol (De Duve, 2007). The most prominent difference for nuclear transcription that arises with eukaryotes is diversification into three different nuclear RNA pols with specialized functions: RNA pol I is responsible for the synthesis of a single transcript, namely, precursor ribosomal RNA, which is processed into 28S, 5.8S, and 18S rRNAs; RNA pol II synthesizes a wide diversity of transcripts, including protein-coding messenger RNA (mRNA) and many ncRNAs, such as microRNAs (mi), small nuclear (sn), and small nucleolar (sno) RNAs; RNA pol III synthesizes diverse transfer RNA (tRNA) and 5S rRNA, and also U6 small nuclear RNA and other non-coding small RNAs (Dieci et al., 2007). There are two additional nuclear RNA pols in plants (IV and V), involved in the transcription of ncRNAs that are required for transcriptional gene silencing via the RNA-directed DNA methylation (Zhou and Law, 2015). In this review, we will focus on the structure and function of RNA pols I, II, and III.

The most well-studied eukaryotic RNA pols are those of the budding yeast *Saccharomyces cerevisiae*, and it is thought that they are good models for other eukaryotic RNA pols. For this reason, we use the names of yeast RNA pols genes and subunits throughout this review (**Figure 1**). Yeast RNA pols I, II, and III have a structurally conserved horseshoe-shaped core formed by 10 subunits (**Figure 1**) homologous to archaeal RNA pol subunits and a different number of additional periphery eukaryote specific subunits (Darst, 2001; Werner, 2007; Cramer et al., 2008). The 10 subunit cores include the two largest catalytic subunits (the two upper rows in **Figure 1B**), five additional subunits (Rpb5, 6, 8, 10, and 12) common to the three nuclear RNA pol, the A12/Rpb9/C11 subunit involved in proofreading (see below) and the AC40–AC19 heterodimer, shared between RNA pols I and III and homologous to Rpb3–Rpb11 in RNA pol II (Fernández-Tornero et al., 2013). The additional periphery yeast RNA pol subunits are mostly essential for cell viability but are not strictly required for RNA polymerization. Instead, they increase the regulatory potential and allow the specialization of each RNA pol in the transcription of a non-redundant subset of genes (Werner, 2007; Cramer et al., 2008; Koster et al., 2015). RNA pol II has a dissociable dimer (Rpb4/7) that plays important roles during the multifaceted transcription elongation of this RNA pol. This dimer has a homology with the Rpo4/7 dimer of archaeal RNA pol and has a counterpart (with low homology) in the A14/A43 and C17/C25 dimers of RNA pols I and III, respectively (**Figure 1**). RNA pol I has a further dimer (A49/A34) that has an equivalent in RNA pol III (C37/C53) but is not a constitutive part of RNA pol II where its function is conducted by the independent TFIIF factor (α/β dimer; Vannini and Cramer, 2012). This dimer plays a specific role in the particular mode of

A



B

Eubacteria	Archaea	Eukarya RNA pol I	Eukarya RNA pol II	Eukarya RNA pol III
β′	Rpo1	A190	Rpb1	C160
β	Rpo2	A135	Rpb2	C128
α	Rpo3	AC40	Rpb3	AC40
α	Rpo11	AC19	Rpb11	AC19
	TFS	A12	Rpb9+ TFIIS	C11
	Rpo5	Rpb5	Rpb5	Rpb5
ω	Rpo6	Rpb6	Rpb6	Rpb6
	Rpo8	Rpb8	Rpb8	Rpb8
	Rpo10	Rpb10	Rpb10	Rpb10
	Rpo12	Rpb12	Rpb12	Rpb12
	Rpo4	A14	Rpb4	C17
	Rpo7	A43	Rpb7	C25
		A49	TFIIFα	C37
		A34.5	TFIIFβ	C53
	Rpo13		TFIIE(α/β)/TFIIF	C31/34/81

FIGURE 1 | Evolutionary history and subunit organization of nuclear eukaryotic RNA polymerases. **(A)** The last universal common ancestor (LUCA) of all organisms is assumed to have a multisubunit DNA-dependent RNA polymerase. Nowadays, all living beings have RNA pols with a core of five to seven subunits. After *Eubacteria* separation, the common ancestor of *Archaea* and *Eukarya* added additional peripheral subunits. Finally, after eukaryote emergence, the *Archaea*-derived nucleus started to develop specialized RNA polymerases. Specialized RNA pols I and III integrated some transcription factors as permanent subunits which, in RNA pol II, remain independent (TFIIS, TFIIF, TFIIE). RNA pol IV and V are not fully described. Only the branching after RNA pol I separation is indicated. See the main text for further descriptions. **(B)** The table shows a comparative scheme of the RNA pol subunits aligned according to sequence and/or functional homology. Colors correspond to the structural scheme of part **(A)**. Note that the Rpb5 and 6 subunits are part of both the core and the five unit sets of common subunits to all three eukaryotic RNA pols. Archaeal Rpo13 has no equivalent in eukaryotes, and the TFS from *Archaea* is an independent homologous factor to eukaryotic TFIIS. See Werner and Grohmann (2011); Vannini and Cramer (2012), and Huang et al. (2015) for more details on RNA pol subunit structure and evolution.

initiation of all three RNA pols (Abascal-Palacios et al., 2018) and in RNA pol III termination (Hoffmann et al., 2015; Arimbasseri and Maraia, 2016) that very much differs from the other two RNA pols in this stage (Proshkina et al., 2006; Werner and Grohmann, 2011). RNA pol III has an additional and totally specific trimer (C31/C34/C82) that is homologous to RNA pol II TFIIE and is proposed to be involved in the mechanism of RNA pol III initiation (Hoffmann et al., 2015). This trimer has been proposed to be TFIIF–TFIIE hybrid rather than simply a TFIIE-like subcomplex (Abascal-Palacios et al., 2018).

The coexistence of the conserved, but different, largest core subunits of the three RNA pols (A190/A135, Rpb1/Rpb2, and C160/C128 in RNA pols I, II, and III, respectively) in all eukaryotes is remarkable and suggests their early evolutionary divergence. At the same time, the substantial conservation of the central RNA pol core since LUCA indicates that it performs essential processes required for gene expression that allows very little innovation. Therefore, in order to generate complex eukaryotes, most evolutionary innovation is expected to occur in periphery subunits, especially in RNA pol II, which specifies the cellular proteome that confers unique characteristics to different cell types through mRNA synthesis. Additionally, the unique *C*-terminal domain (CTD) of the largest catalytic subunit (Rpb1) of RNA pol II is also one source for innovation in mRNA transcription regulation and a mark of the eukaryotic lineage (Burton, 2014). CTD consists of a repeating structure that is rich in serine and other phosphorylable amino acids, which increases in number of repetitions with greater evolutionary complexity. Another consequence of eukaryotic innovation is the complex structure of RNA pol III with 17 subunits, which are all conserved to a certain degree in eukaryotes from yeast to humans. This supports the notion of the early divergence of RNA pol III from RNA pols I and II (Proshkina et al., 2006; **Figure 1**). Of all these considerations, it can be suggested that the last eukaryote common ancestor is likely to have already had distinct RNA pols I, II, and III, as well as the repetitive structure at the CTD of RNA pol II (Proshkina et al., 2006; Yang and Stiller, 2014). It can be concluded that the existence and evolution of the three specialized RNA pols in eukaryotic cells would have allowed the division of labor and enabled intricate gene regulation in multicellular complex organisms that requires the cell cycle, tissue-specific, environmental, and developmental regulation of gene expression (Dieci et al., 2007; Cramer, 2019). RNA pols IV and V are thought to have evolved more recently from RNA pol II through subfunctionalization of silencing activities performed by RNA pol II in fungi and metazoans in the earliest land plants (Huang et al., 2015).

DIFFERENCES IN THE THREE RNA POL STRUCTURE LINKED TO DIFFERENCES IN FUNCTION

Although the transcription cycle (initiation, elongation, and termination) has similar principles in all three nuclear RNA pols, the specific features of their transcription modes are reflected in their subunit structures. RNA pol II targets a large set of differently regulated genes, which requires the capacity to interact with a bigger set of transcription initiation and elongation factors than the other two RNA pols. Perhaps this was accomplished by having less permanent subunits than the other two RNA pols, but by also having dissociable subunits (Rpb4/7) and independent initiation and elongation factors (TFIIF, TFIIS, and TFIIE), while the equivalent factors in other polymerases form an intrinsic part (subunits) of the RNA pol complex. For example, RNA pol I has a single promoter to recognize but requires high-speed, efficient elongation to avoid collisions between polymerases in its highly crowded genes (Goodfellow and Zomerdijk, 2013). This is perhaps the reason why RNA pol I possesses important intrinsic RNA cleavage activity for proofreading and a rapid resumption of elongation after pausing (Fernández-Tornero et al., 2013). This activity resides in its A12 subunit with homology to both the RNA pol II Rpb9 subunit and the TFIIS elongation factor. Thus, A12 seems to be a fusion protein that comprises the amino-terminal domain of the RNA pol II Rbp9 subunit and the carboxy-terminal domain of TFIIS (Hoffmann et al., 2015). A similar reasoning can be done for RNA pol III where the C11 subunit has homology to Rpb9 and TFIIS (Chédin et al., 1998). The more complicated process of resuming elongation after pausing in RNA pol II suggests the need for specific regulation, which is not required for simpler and faster RNA pol I/III elongation (Engel et al., 2013). Another example of functions that fall in RNA pol III intrinsic subunits but in external transcription factors in RNA pol II is related to transcription termination. RNA pol III specific dimer (C53/C37) together with C11 subunit are particularly required for the very fast efficient termination and coupled re-initiation needed by this RNA pol due to the highly transcribed and very short genes that it targets (Dieci et al., 2013; Arimbasseri and Maraia, 2016). In fact, RNA pol III termination is distinct from that of the other two nuclear RNA pols because its genes present a tract of oligo-T at the 3′ end, which induces termination. On the contrary, RNA pols I and II require additional *cis*-acting elements and ancillary factors for termination (Arimbasseri and Maraia, 2016). In short, both RNA pols I and III seem to have integrated some transcription factor-like subunits into the core enzyme during evolution to prioritize rapid efficient transcription *versus* regulation (Carter and Drouin, 2010).

Chromatin imposes a major limitation to transcription by three eukaryotic RNA pols preventing their direct targeting to gene promoters, which probably explains why all nuclear RNA pols are first engaged in pre-initiation complexes before starting transcription. Pre-initiation complexes minimally consist of the TATA box-binding protein (TBP), which is common to all three transcription systems, initiation factors TFIIB (RNA pol II) and Brf1 (RNA pol III), and the RNA pol II-specific TFIIE factor (Hahn, 2004; Naidu et al., 2011). Moreover, during elongation, chromatin imposes clearly different conditions to each RNA pol. Active rRNA genes are totally covered by transcribing RNA pol I complexes to form characteristic Christmas trees with no nucleosomes (Albert et al., 2012; Goodfellow and Zomerdijk, 2013). Most RNA pol III genes (tRNAs and 5S, especially) are so short that the whole transcribing unit lies in a short track free of nucleosomes (Shukla and Bhargava, 2018), unlike RNA pol II that transcribes longer genes and deals with nucleosomes

during elongation. The arrest and backtracking of RNA pol II occur at nucleosome barriers, and elongation is resumed by the stimulation of weak intrinsic RNA pol II cleavage activity by TFIIS to form a new RNA 3′ end in its active site (Cramer, 2019). This more complicated way of solving backtracking could serve to refine the elongation regulation process (Bradsher et al., 1993; Shilatifard et al., 1996).

THE NEED FOR COORDINATION OF THE THREE RNA POL ACTIVITIES: THE CASE OF TRANSLATION MACHINERY

Translation machinery (ribosomes and tRNAs) synthesis requires the tight coordination among all nuclear RNA pols because rRNAs are synthesized by RNA pols I and III and ribosomal proteins are made from mRNAs transcribed by RNA pol II. Hence their coordination at all times and in all growth regimes is clearly necessary. The existence of five common subunits and one universal initiation factor, TBP, in all three RNA pols may be used to establish common regulatory mechanisms for nuclear transcription. RP mRNAs are some of the most abundant mRNAs in actively growing cells (Pelechano et al., 2010), and, thus, their synthesis forms a significant part of the total RNA pol II effort (Warner, 1999). Moreover, many other RNA pol II genes encode proteins involved in ribosome biogenesis but are not part of ribosomes. These include the RNA pol I and III subunits and the proteins involved in rRNA and tRNA maturation, and transport and translation factors, which are coordinately regulated (RiBi regulon in yeast) and also share some regulatory mechanisms with RP genes (Martin et al., 2006; Bosio et al., 2017). Therefore, the coordination of ribosome biogenesis and its regulation by growth must require the existence of regulatory mechanisms that coordinate their output. Candidates for this role are mammalian c-Myc and the yeast Sfp1 transcription factors (Lempiäinen and Shore, 2009). Regulation of RNA pols by growth is dependent on the target of rapamycin and Ras–PKA pathways that link ribosome production to nutrient availability (Warner, 1999; Martin et al., 2006; Mayer and Grummt, 2006; Lempiäinen and Shore, 2009). These pathways act by regulating the activity of several transcription activators, such as Rap1, Abf1, or Sfp1, in yeast (see Bosio et al., 2017, for further details).

AN EXAMPLE OF DIFFERENT REGULATION OF RNA POLS RELATED TO THEIR DIFFERENT FUNCTION: RIBOSTATIC CONTROL DURING CELL VOLUME VARIATIONS

The different properties of eukaryotic RNA pols and their RNA products predict that the regulatory mechanisms used by each one to cope with changes will be different. We discuss here an example that we have recently studied in yeast *S. cerevisiae*: the regulation of global RNA pol I and II activities with respect to changes in cell volume.

Homeostasis is defined as the state of steady internal conditions maintained by living beings and includes the control of concentrations of cell molecules. The terms ribostasis and proteostasis refer to the modulation of RNA and protein levels, respectively, in response to changes in the environment. Proteins are mostly the final goal of gene expression and are in charge of catalytic and structural functions. For this reason, their homeostasis is very strictly controlled, and the total protein concentration remains quite constant (Liebermeister et al., 2014; Milo and Phillips, 2015; Benet et al., 2017). Nonetheless, gene expression regulation occurs chiefly at the mRNA level. For this purpose, mRNAs are mostly unstable, and the overall mRNA concentration is controlled within a certain range (Pérez-Ortín et al., 2013; Benet et al., 2017). On the contrary, rRNAs and tRNAs remain very stable during active growth and only degrade under stress conditions or when defects in the molecule occur (Deutscher, 2006; Pérez-Ortín et al., 2019).

Homeostasis deals with the molecular concentration, and not with the number of molecules. Therefore, changes in the cell volume are expected to provoke adaptation mechanisms to maintain homeostasis. In yeast cells, and probably in other organisms, volume varies depending on the genotype, the cell cycle phase (Jorgensen et al., 2002; Ferrezuelo et al., 2010), aging (Egilmez et al., 1990), ploidy (Cook and Tyers, 2007; Lee et al., 2009), and the growth rate (Aldea et al., 2017). To maintain ribostasis and proteostasis, increases in cell volume must be compensated by the coordinated increase in the amounts of RNA molecules and proteins (Bustamante et al., 2014; Walters and Parker, 2015).

Studies carried out in different model organisms have established differences between transcription regulatory responses to cell volume depending on the organism and the RNA pol studied and suggest the existence of a size-sensing mechanism that produces alterations in transcription (Wu et al., 2010). Changes in the RNA pol II transcription rate (TR) with volume increase have been widely studied to show that it is differentially regulated in cells with different cell division types. Thus, for symmetrically dividing cells, such as mammalian fibroblasts (Padovan-Merhar et al., 2015), or *Schizosaccharomyces pombe* (Zhurinsky et al., 2010), RNA pol II nascent TR increases in parallel with volume due to a bigger and faster recruitment of polymerase onto chromatin (Sun et al., 2020). Thus, for symmetrically dividieng cells, such as *S. cerevisiae*, nascent TR remains constant by controlling the expression of RNA pol II coding genes, while mRNA stability increases to maintain mRNA ribostasis (Mena et al., 2017). This difference is explained by asymmetric cell division in *S. cerevisiae* resulting in two cells with different volumes: a small daughter cell and a large mother cell. In this scenario, the strategy adopted by eukaryotes with symmetric cell division, such as *S. pombe* or fibroblasts, is not applicable, as it would result in a higher mRNA net synthesis rate in small daughter cells (Mena et al., 2017). However, the strategy adopted to adapt ribostasis to increased cell volume is very different for RNA pol I. In this case, nascent TR increases with cell volume by increasing the number of copies of the rDNA gene. The higher gene copy number can occur by increasing cell ploidy or by expanding the number of rDNA repeats

(Mena et al., 2017; Pérez-Ortín et al., 2021). This mode of regulation is a slow form of TR regulation because changes in the genome can occur only during replication (Kobayashi, 2006, 2011; Nelson et al., 2019).

Why is there a different solution for an identical problem in RNA pols I and II? We hypothesize that the differences in their targets, the 35S gene, and protein-encoding genes conditioned the evolution of different regulatory mechanisms for RNA pol I and RNA pol II. As the rRNA TR needs to reach much higher levels than that of any of the RNA pol II genes, eukaryotic cells evolved a specialized faster polymerase with a single gene template with many repeated copies. RNA pol is able to form extremely dense head-to-tail "camel caravans" in which the A49 subunit from one molecule contacts directly with A43 from the neighboring molecule. Thus, the specialized dimer A49/A43 allows a higher RNA pol loading rate than in RNA pol II (Albert et al., 2011). On the other hand, the repeated nature of the rDNA locus is prone to cause homologous recombination (Iida and Kobayashi, 2019) and offers the opportunity to alter the rDNA copy number and, thus, total TR without changing nascent TR per gene copy. In this way, RNA pol I can be controlled in the short term at the transcription initiation and elongation levels, as with other RNA pols, but also in the long term by changing its copy number during genome replication (Kobayashi et al., 1998; Pérez-Ortín et al., 2021). An interesting question arises here: what happens to the RNA pol III that transcribes tRNAs and 5S genes, whose gene number is also a few hundred copies (Turowski and Tollervey, 2016)? Interestingly, 5S genes are localized within the rDNA repeats in the genome of *Saccharomycotina* clade (Bergeron and Drouin, 2008), which comprises mostly asymmetrically dividing yeasts, which could imply a common TR regulation strategy for RNA pols I and III in rRNA synthesis. To support this idea, in other yeasts and most of other eukaryotes with symmetric cell division, 5S genes are usually dispersed along the genome (Drouin and De Sa, 1995).

To summarize, eukaryotes differentiate from prokaryotes not only because of a more complex intracellular organization with nuclear and organelle evolutionarily independent genomes but also because the unprecedented job division occurs between several distinct nuclear RNA pols. The specialization of each one in the synthesis of a specific subset of transcripts with different abundance, stability, and function has forced differences in transcription initiation, elongation, termination, and regulation strategies but has provided, at the same time, the versatility to make phenotypically different cells from the same genome as a requisite for multicellular organisms.

AUTHOR CONTRIBUTIONS

JP-O conceived the manuscript. PA, MB-A, and JP-O wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

JP-O and PA were supported by grants from the Spanish MiNECO (BFU2016-77728-C3-3-P) and from the Regional Valencian Government (Generalitat Valenciana AICO/2019/088 to JP-O and AICO/2020/086 to PA). These projects received support from the European Union funds (FEDER). MB-A is a recipient of a predoctoral fellowship (FPU2017/03542) from the Spanish Ministry of Science, Innovation and Universities.

ACKNOWLEDGMENTS

The authors are especially grateful to F. Navarro for critically reviewing the manuscript.

REFERENCES

Abascal-Palacios, G., Ramsay, E. P., Beuron, F., Morris, E., and Vannini, A. (2018). Structural basis of RNA polymerase III transcription initiation. *Nature* 553, 301–306. doi: 10.1038/nature25441

Albert, B., Léger-Silvestre, I., Normand, C., Ostermaier, M. K., Pérez-Fernández, J., Panov, K. I., et al. (2011). RNA polymerase I-specific subunits promote polymerase clustering to enhance the rRNA gene transcription cycle. *J. Cell Biol.* 192, 277–293. doi: 10.1083/jcb.201006040

Albert, B., Pérez-Fernández, J., Léger-Silvestre, I., and Gadal, O. (2012). Regulation of ribosomal RNA production by RNA Polymerase I: does elongation come first? *Genet. Res. Int.* 2012:276948. doi: 10.1155/2012/276948

Aldea, M., Jenkins, K., and Csikász-Nagy, A. (2017). Growth rate as a direct regulator of the start network to set cell size. *Front. Cell Dev. Biol.* 5:57. doi: 10.3389/fcell.2017.00057

Arimbasseri, A. G., and Maraia, R. J. (2016). RNA polymerase III advances: structural and tRNA functional views. *Trends Biochem. Sci.* 41, 546–559. doi: 10.1016/j.tibs.2016.03.003

Benet, M., Miguel, A., Carrasco, F., Li, T., Planells, J., Alepuz, P., et al. (2017). Modulation of protein synthesis and degradation maintains proteostasis during yeast growth at different temperatures. *Biochim. Biophys. Acta Gene Regul. Mech.* 1860, 794–802. doi: 10.1016/j.bbagrm.2017.04.003

Bergeron, J., and Drouin, G. (2008). The evolution of 5S ribosomal RNA genes linked to the rDNA units of fungal species. *Curr. Genet.* 54, 123–131. doi: 10.1007/s00294-008-0201-2

Bosio, M. C., Fermi, B., and Dieci, G. (2017). Transcriptional control of yeast ribosome biogenesis: a multifaceted role for general regulatory factors. *Transcription* 8, 254–260. doi: 10.1080/21541264.2017.1317378

Bradsher, J. N., Jackson, K. W., Conaway, R. C., and Conaway, J. W. (1993). RNA polymerase II transcription factor SIII. I. Identification, purification, and properties. *J. Biol. Chem.* 268, 25587–25593. doi: 10.1016/s0021-9258(19)74431-7

Burton, Z. F. (2014). The old and new testaments of gene regulation: evolution of multi-subunit RNA polymerases and co-evolution of eukaryote complexity with the RNAP II CTD. *Transcription* 5:e28674. doi: 10.4161/trns.28674

Bustamante, C. J., Kaiser, C. M., Maillard, R. A., Goldman, D. H., and Wilson, C. A. M. (2014). Mechanisms of cellular proteostasis: insights from single-molecule approaches. *Annu. Rev. Biophys.* 43, 119–140. doi: 10.1146/annurev-biophys-051013-022811

Carter, R., and Drouin, G. (2010). The increase in the number of subunits in eukaryotic RNA polymerase III relative to RNA polymerase II Is due to the permanent recruitment of general transcription factors. *Mol. Biol. Evol.* 27, 1035–1043. doi: 10.1093/molbev/msp316

Chédin, S., Riva, M., Schultz, P., Sentenac, A., and Caries, C. (1998). The RNA cleavage activity of RNA polymerase III is mediated an essential TFIIS-like

subunit and is important for transcription termination. *Genes Dev.* 12, 3857–3871. doi: 10.1101/gad.12.24.3857

Cook, M., and Tyers, M. (2007). Size control goes global. *Curr. Opin. Biotechnol.* 18, 341–350. doi: 10.1016/j.copbio.2007.07.006

Cramer, P. (2019). Eukaryotic transcription turns 50. *Cell* 179, 808–812. doi: 10.1016/j.cell.2019.09.018

Cramer, P., Armache, K. J., Baumli, S., Benkert, S., Brueckner, F., Buchen, C., et al. (2008). Structure of eukaryotic RNA polymerases. *Annu. Rev. Biophys.* 37, 337–352. doi: 10.1146/annurev.biophys.37.032807.130008

Darst, S. A. (2001). Bacterial RNA polymerase. *Curr. Opin. Struct. Biol.* 11, 155–162. doi: 10.1016/S0959-440X(00)00185-8

De Duve, C. (2007). The origin of eukaryotes: a reappraisal. *Nat. Rev. Genet.* 8, 395–403. doi: 10.1038/nrg2071

Deutscher, M. P. (2006). Degradation of RNA in bacteria: comparison of mRNA and stable RNA. *Nucleic Acids Res.* 34, 659–666. doi: 10.1093/nar/gkj472

Dieci, G., Bosio, M. C., Fermi, B., and Ferrari, R. (2013). Transcription reinitiation by RNA polymerase III. *Biochim. Biophys. Acta Gene Regul. Mech.* 1829, 331–341. doi: 10.1016/j.bbagrm.2012.10.009

Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M., and Pagano, A. (2007). The expanding RNA polymerase III transcriptome. *Trends Genet.* 23, 614–622. doi: 10.1016/j.tig.2007.09.001

Drouin, G., and De Sa, M. M. (1995). The concerted evolution of 5S ribosomal genes linked to the repeat units of other multigene families. *Mol. Biol. Evol.* 12, 481–493. doi: 10.1093/oxfordjournals.molbev.a040223

Egilmez, N. K., Chen, J. B., and Jazwinski, S. M. (1990). Preparation and partial characterization of old yeast cells. *J. Gerontol.* 45, 9–17. doi: 10.1093/geronj/45.1.B9

Engel, C., Sainsbury, S., Cheung, A. C., Kostrewa, D., and Cramer, P. (2013). RNA polymerase i structure and transcription regulation. *Nature* 502, 650–655. doi: 10.1038/nature12712

Fernández-Tornero, C., Moreno-Morcillo, M., Rashid, U. J., Taylor, N. M. I., Ruiz, F. M., Gruene, T., et al. (2013). Crystal structure of the 14-subunit RNA polymerase i. *Nature* 502, 644–649. doi: 10.1038/nature12636

Ferrezuelo, F., Colomina, N., Futcher, B., and Aldea, M. (2010). The transcriptional network activated by Cln3 cyclin at the G1-to-S transition of the yeast cell cycle. *Genome Biol.* 11:R67. doi: 10.1186/gb-2010-11-6-r67

Fouqueau, T., Blombach, F., and Werner, F. (2017). Evolutionary origins of two-barrel RNA polymerases and site-specific transcription initiation. *Annu. Rev. Microbiol.* 71, 331–348. doi: 10.1146/annurev-micro-091014-104145

Geiduschek, E. P., and Ouhammouch, M. (2005). Archaeal transcription and its regulators. *Mol. Microbiol.* 56, 1397–1407. doi: 10.1111/j.1365-2958.2005.04627.x

Goodfellow, S. J., and Zomerdijk, J. C. B. M. (2013). Basic mechanisms in RNA polymerase I transcription of the ribosomal RNA genes. *Subcell. Biochem.* 61, 211–236. doi: 10.1007/978-94-007-4525-4_10

Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* 11, 394–403. doi: 10.1038/nsmb763

Hoffmann, N. A., Jakobi, A. J., Moreno-Morcillo, M., Glatt, S., Kosinski, J., Hagen, W. J. H., et al. (2015). Molecular structures of unbound and transcribing RNA polymerase III. *Nature* 528, 231–236. doi: 10.1038/nature16143

Huang, Y., Kendall, T., Forsythe, E. S., Dorantes-Acosta, A., Li, S., Caballero-Pérez, J., et al. (2015). Ancient origin and recent innovations of RNA polymerase IV and V. *Mol. Biol. Evol.* 32, 1788–1799. doi: 10.1093/molbev/msv060

Iida, T., and Kobayashi, T. (2019). RNA Polymerase I activators count and adjust ribosomal RNA gene copy number. *Mol. Cell* 73, 545.e13–654.e13. doi: 10.1016/j.molcel.2018.11.029

Iyer, L. M., Koonin, E. V., and Aravind, L. (2003). Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases. *BMC Struct. Biol.* 3:1. doi: 10.1186/1472-6807-3-1

Jorgensen, P., Nishikawa, J. L., Breitkreutz, B. J., and Tyers, M. (2002). Systematic identification of pathways that couple cell growth and division in yeast. *Science* 297, 395–400. doi: 10.1126/science.1070850

Jun, S. H., Reichlen, M. J., Tajiri, M., and Murakami, K. S. (2011). Archaeal RNA polymerase and transcription regulation. *Crit. Rev. Biochem. Mol. Biol.* 46, 27–40. doi: 10.3109/10409238.2010.538662

Kobayashi, T. (2006). Strategies to maintain the stability of the ribosomal RNA gene repeats: collaboration of recombination, cohesion, and condensation. *Genes Genet. Syst.* 81, 155–161. doi: 10.1266/ggs.81.155

Kobayashi, T. (2011). Regulation of ribosomal RNA gene copy number and its role in modulating genome integrity and evolutionary adaptability in yeast. *Cell. Mol. Life Sci.* 68, 1395–1403. doi: 10.1007/s00018-010-0613-2

Kobayashi, T., Heck, D. J., Nomura, M., and Horiuchi, T. (1998). Expansion and contraction of ribosomal DNA repeats in Saccharomyces cerevisiae: requirement of replication fork blocking (Fob1) protein and the role of RNA polymerase I. *Genes Dev.* 12, 3821–3830. doi: 10.1101/gad.12.24.3821

Koonin, E. V. (2015). Origin of eukaryotes from within archaea, archaeal eukaryome and bursts of gene gain: eukaryogenesis just made easier? *Philos. Trans. R. Soc. B Biol. Sci.* 370:20140333. doi: 10.1098/rstb.2014.0333

Koster, M. J. E., Snel, B., and Timmers, H. T. M. (2015). Genesis of chromatin and transcription dynamics in the origin of species. *Cell* 161, 724–736. doi: 10.1016/j.cell.2015.04.033

Kwapisz, M., Beckouët, F., and Thuriaux, P. (2008). Early evolution of eukaryotic DNA-dependent RNA polymerases. *Trends Genet.* 24, 211–215. doi: 10.1016/j.tig.2008.02.002

Lane, N., and Martin, W. (2010). The energetics of genome complexity. *Nature* 467, 929–934. doi: 10.1038/nature09486

Lee, H. O., Davidson, J. M., and Duronio, R. J. (2009). Endoreplication: polyploidy with purpose. *Genes Dev.* 23, 2461–2477. doi: 10.1101/gad.1829209

Lempiäinen, H., and Shore, D. (2009). Growth control and ribosome biogenesis. *Curr. Opin. Cell Biol.* 21, 855–863. doi: 10.1016/j.ceb.2009.09.002

Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual account of protein investment in cellular functions. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8488–8493. doi: 10.1073/pnas.1314810111

Martin, D. E., Powers, T., and Hall, M. N. (2006). Regulation of ribosome biogenesis: where is TOR? *Cell Metab.* 4, 259–260. doi: 10.1016/j.cmet.2006.09.002

Mayer, C., and Grummt, I. (2006). Ribosome biogenesis and cell growth: mTOR coordinates transcription by all three classes of nuclear RNA polymerases. *Oncogene* 25, 6384–6391. doi: 10.1038/sj.onc.1209883

Mena, A., Medina, D. A., Garciá-Martínez, J., Begley, V., Singh, A., Chávez, S., et al. (2017). Asymmetric cell division requires specific mechanisms for adjusting global transcription. *Nucleic Acids Res.* 45, 12401–12412. doi: 10.1093/nar/gkx974

Milo, R., and Phillips, R. (2015). *Cell Biology by the Numbers*. New York, NY: Taylor & Francis Group, doi: 10.1201/9780429258770

Naidu, S., Friedrich, J. K., Russell, J., and Zomerdijk, J. C. B. M. (2011). TAF1B is a TFIIB-like component of the basal transcription machinery for RNA polymerase I. *Science* 333, 1640–1642. doi: 10.1126/science.1207656

Nelson, J. O., Watase, G. J., Warsinger-Pepe, N., and Yamashita, Y. M. (2019). Mechanisms of rDNA copy number maintenance. *Trends Genet.* 5, 734–742. doi: 10.1016/j.tig.2019.07.006

Padovan-Merhar, O., Nair, G. P., Biaesch, A. G., Mayer, A., Scarfone, S., Foley, S. W., et al. (2015). Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. *Mol. Cell* 58, 339–352. doi: 10.1016/j.molcel.2015.03.005

Pelechano, V., Chávez, S., and Pérez-Ortín, J. E. (2010). A complete set of nascent transcription rates for yeast genes. *PLoS One* 5:e15442. doi: 10.1371/journal.pone.0015442

Pérez-Ortín, J. E., Alepuz, P., Chávez, S., and Choder, M. (2013). Eukaryotic mRNA decay: methodologies, pathways, and links to other stages of gene expression. *J. Mol. Biol.* 425, 3750–3775. doi: 10.1016/j.jmb.2013.02.029

Pérez-Ortín, J. E., Mena, A., Barba-Aliaga, M., Singh, A., Chávez, S., and García-Martínez, J. (2021). Cell volume homeostatically controls the rDNA repeat copy number and rRNA synthesis rate in yeast. *PLoS Genet.* 17(4):e1009520. doi: 10.1371/journal.pgen.1009520

Pérez-Ortín, J. E., Tordera, V., and Chávez, S. (2019). Homeostasis in the central dogma of molecular biology: the importance of mRNA instability. *RNA Biol.* 16, 1659–1666. doi: 10.1080/15476286.2019.1655352

Proshkina, G. M., Shematorova, E. K., Proshkin, S. A., Zaros, C., Thuriaux, P., and Shpakovski, G. V. (2006). Ancient origin, functional conservation and

fast evolution of DNA-dependent RNA polymerase III. *Nucleic Acids Res.* 34, 3615–3624. doi: 10.1093/nar/gkl421

Reeve, J. N. (2003). Archaeal chromatin and transcription. *Mol. Microbiol.* 48, 587–598. doi: 10.1046/j.1365-2958.2003.03439.x

Shilatifard, A., Lane, W. S., Jackson, K. W., Conaway, R. C., and Conaway, J. W. (1996). An RNA polymerase II elongation factor encoded by the human ELL gene. *Science* 271, 1873–1876. doi: 10.1126/science.271.5257.1873

Shukla, A., and Bhargava, P. (2018). Regulation of tRNA gene transcription by the chromatin structure and nucleosome dynamics. *Biochim. Biophys. Acta Gene Regul. Mech.* 1861, 295–309. doi: 10.1016/j.bbagrm.2017.11.008

Spang, A., Saw, J. H., Jørgensen, S. L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A. E., et al. (2015). Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* 521, 173–179. doi: 10.1038/nature14447

Sun, X. M., Bowman, A., Priestman, M., Bertaux, F., Martinez-Segura, A., Tang, W., et al. (2020). Size-dependent increase in RNA Polymerase II initiation rates mediates gene expression scaling with cell size. *Curr. Biol.* 30, 1217.e7–1230.e7. doi: 10.1016/j.cub.2020.01.053

Turowski, T. W., and Tollervey, D. (2016). Transcription by RNA polymerase III: insights into mechanism and regulation. *Biochem. Soc. Trans.* 44, 1367–1375. doi: 10.1042/BST20160062

Vannini, A., and Cramer, P. (2012). Conservation between the RNA Polymerase I, II, and III transcription initiation machineries. *Mol. Cell* 45, 439–446. doi: 10.1016/j.molcel.2012.01.023

Walters, R. W., and Parker, R. (2015). Coupling of ribostasis and proteostasis: Hsp70 proteins in mRNA metabolism. *Trends Biochem. Sci.* 40, 552–559. doi: 10.1016/j.tibs.2015.08.004

Warner, J. R. (1999). The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.* 24, 437–440. doi: 10.1016/S0968-0004(99)01460-7

Werner, F. (2007). Structure and function of archaeal RNA polymerases. *Mol. Microbiol.* 65, 1395–1404. doi: 10.1111/j.1365-2958.2007.05876.x

Werner, F., and Grohmann, D. (2011). Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* 9, 85–98. doi: 10.1038/nrmicro2507

Wu, C. Y., Alexander Rolfe, P., Gifford, D. K., and Fink, G. R. (2010). Control of transcription by cell size. *PLoS Biol.* 8:e1000523. doi: 10.1371/journal.pbio.1000523

Yang, C., and Stiller, J. W. (2014). Evolutionary diversity and taxon-specific modifications of the RNA polymerase II C-terminal domain. *Proc. Natl. Acad. Sci. U.S.A.* 111, 5920–5925. doi: 10.1073/pnas.1323616111

Zhou, M., and Law, J. A. (2015). RNA Pol IV and V in gene silencing: rebel polymerases evolving away from Pol II's rules. *Curr. Opin. Plant Biol.* 27, 154–164. doi: 10.1016/j.pbi.2015.07.005

Zhurinsky, J., Leonhard, K., Watt, S., Marguerat, S., Bähler, J., and Nurse, P. (2010). A coordinated global control over cellular transcription. *Curr. Biol.* 20, 2010–2015. doi: 10.1016/j.cub.2010.10.002

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Molecular Biosciences

MINI REVIEW
published: 28 April 2021
doi: 10.3389/fmolb.2021.669300

Biogenesis of RNA Polymerases in Yeast

*Ana I. Garrido-Godino[1], Francisco Gutiérrez-Santiago[1] and Francisco Navarro[1,2]**

[1] Departamento de Biología Experimental-Genética, Universidad de Jaén, Jaén, Spain, [2] Centro de Estudios Avanzados en Aceite de Oliva y Olivar, Universidad de Jaén, Jaén, Spain

Eukaryotic RNA polymerases (RNA pols) transcriptional processes have been extensively investigated, and the structural analysis of eukaryotic RNA pols has been explored. However, the global assembly and biogenesis of these heteromultimeric complexes have been narrowly studied. Despite nuclear transcription being carried out by three RNA polymerases in eukaryotes (five in plants) with specificity in the synthesis of different RNA types, the biogenesis process has been proposed to be similar, at least for RNA pol II, to that of bacteria, which contains only one RNA pol. The formation of three different interacting subassembly complexes to conform the complete enzyme in the cytoplasm, prior to its nuclear import, has been assumed. In *Saccharomyces cerevisiae*, recent studies have examined in depth the biogenesis of RNA polymerases by characterizing some elements involved in the assembly of these multisubunit complexes, some of which are conserved in humans. This study reviews the latest studies governing the mechanisms and proteins described as being involved in the biogenesis of RNA polymerases in yeast.

Keywords: RNA polymerases, biogenesis, assembly, transcription, yeast

OPEN ACCESS

Edited by:
Teng Ma,
Capital Medical University, China

Reviewed by:
Magdalena Boguta,
Institute of Biochemistry and Biophysics (PAN), Poland
Mehdi Kabani,
Centre National de la Recherche Scientifique (CNRS), France

***Correspondence:**
Francisco Navarro
fngomez@ujaen.es

Specialty section:
This article was submitted to Protein and RNA Networks, a section of the journal Frontiers in Molecular Biosciences

Received: *18 February 2021*
Accepted: *31 March 2021*
Published: *28 April 2021*

Citation:
Garrido-Godino AI, Gutiérrez-Santiago F and Navarro F (2021) Biogenesis of RNA Polymerases in Yeast. Front. Mol. Biosci. 8:669300. doi: 10.3389/fmolb.2021.669300

INTRODUCTION

Transcription is the most studied step of the gene expression catalyzed by RNA polymerases (RNA pols). Eukaryotes contain at least three RNA pols (RNA pols I, II, III), while archaea and bacteria consist of a single enzyme (Werner and Grohmann, 2011). In addition, two additional RNA pols have been described to be present in plants (RNA pols IV and V) (Wierzbicki et al., 2008; Haag and Pikaard, 2011; Haag et al., 2014). Although single-subunit RNA pols exist (as in bacteriophage T7) (Kwapisz et al., 2008), the bacterial, archaeal, and eukaryotic RNA pols are heteromultimeric complexes (Werner and Grohmann, 2011; Cramer, 2019). In eukaryotes, RNA pol I comprises 14 subunits and synthesizes a precursor of the three largest rRNAs (Werner et al., 2009; Lane et al., 2011; Werner and Grohmann, 2011; Moreno-Morcillo et al., 2014). RNA pol II contains 12 subunits and is responsible for the transcription of mRNAs and some non-coding RNAs (Armache et al., 2003, 2005; Werner and Grohmann, 2011). RNA pol III is composed of 17 subunits and catalyzes the synthesis of tRNAs and 5S rRNA, as well as other non-coding RNAs (Werner et al., 2009; Fernández-Tornero et al., 2011; Lane et al., 2011; Werner and Grohmann, 2011; Dieci et al., 2012; Khatter et al., 2017).

Despite the fact that the transcription process and regulation have been extensively studied along with the structure of RNA pols (Ishihama, 1981; Briand et al., 2001; Cramer, 2002, 2019; Armache et al., 2003, 2005; Werner, 2007; Werner et al., 2009; Fernández-Tornero et al., 2011, 2013; Lane et al., 2011; Werner and Grohmann, 2011; Moreno-Morcillo et al., 2014), very little is known about the biogenesis of multisubunit RNA pols and how these processes occur in yeast. Several studies have identified many factors involved in the assembly and/or nuclear transport of RNA polymerases in both yeast and human cells, with most of them operating for the biogenesis of RNA pol II. In light of this, this review focuses on pre-existing knowledge of the assembly of RNA polymerases in yeast.

GENERAL OVERVIEW OF THE ASSEMBLY PROCESSES OF RNA POLYMERASES

A model for bacterial RNA pol (αααββ'ω subunits) assembly has been proposed based on *in vitro* experiments. Assembly would start with the formation of the αα dimer, which would interact with the β subunit. Later, the ααβ module would associate with the β' subunit, which probably forms a complex with the ω subunit (Ishihama, 1981). Interestingly, the ω subunit, which is not essential, seems to stabilize the β' subunit (Minakhin et al., 2001).

In yeast, a model for the biogenesis of RNA pol II based on bacterial RNA pol formation (Ishihama, 1981) has been suggested (Wild and Cramer, 2012). A similar model has been proposed in human cells, which suggests the conservation of these processes. Yeast RNA pol II possesses the bacterial homolog αααββ'ω core composed of subunits Rpb1, Rpb2, Rpb3, Rpb11, and Rpb6 (Zhang et al., 1999; Werner and Grohmann, 2011; Wild and Cramer, 2012; Cramer, 2019). These subunits are conserved in yeast RNA pol I and III (Werner and Grohmann, 2011; Cramer, 2019). Accordingly, the Rpb3 subassembly complex (corresponding to αα dimer: Rpb3, Rpb10, Rpb11, and Rpb12) would form and interact with the Rpb2 subassembly complex (similar β subunit: Rpb2 and Rpb9) prior to the association with the Rpb1 subassembly module (β'ω subunits: composed of Rpb1, Rpb5, Rpb6, and Rpb8). Rpb6 (ω, subunit) would act by stabilizing the largest subunit of RNA pol II for the assembly of RNA pol II (Nouraini et al., 1996; Minakhin et al., 2001; Garrido-Godino et al., 2013) in line with the role proposed for the ω subunit (Minakhin et al., 2001). Finally, the stalk subcomplex (Rpb4/7) would associate with the preassembled core enzyme (Wild and Cramer, 2012 and our unpublished data), although it can dissociate (Armache et al., 2005).

While the proposed model accounts for RNA pol II, the question of what occurs for RNA pol I and III arises. By taking into account the conservation of the different subunits αααββ'ω in RNA pol I and III, similar pathways may act for the assembly of all eukaryotic RNA pols, and this process is likely coordinated by the existence of five RNA pols common subunits: Rpb5, Rpb6, Rpb8, Rpb10, and Rpb12 (Wild and Cramer, 2012). It is worth noting that the yeast homologous ω subunit Rpb6 has also been described to stabilize RNA pol I in addition to RNA pol II (Nouraini et al., 1996; Minakhin et al., 2001; Garrido-Godino et al., 2013). Furthermore, the availability of the shared subunit Rpb12 must be a limiting step in the assembly of the three RNA pols in yeast (Rubbi et al., 1999; Wild and Cramer, 2012). In addition, homologous Rpb4/7 complexes exist for RNA pol I and III (Rpa14/Rpa43 and Rpc17/Rpc25, respectively) (Werner and Grohmann, 2011). Nevertheless, the mechanisms governing the assembly of RNA pol I and III are described in another article published in the same issue by Boguta and Turowski (in press).

In line with previously proposed mechanisms for yeast, quantitative proteomic analyses in human cells have demonstrated the existence of a cytoplasmic RNA pol II subcomplex formed by subunits RPB2, RPB3, RPB10, RPB11, and RPB12 (Boulon et al., 2010), which suggests that the interaction of the RPB2 and RPB3 subassembly complexes may occur prior to the association with the RPB1 subassembly complex.

Another important question relates to where the assembly of the RNA polymerases occurs and how they enter the nucleus. RNA pol II assembly has been proposed to occur in the cytoplasm before its nuclear import in yeast (Boulon et al., 2010; Corden, 2011; Wild and Cramer, 2012; Mirón-García et al., 2013; Gómez-Navarro and Estruch, 2015), as it similarly occurs in human cells (Boulon et al., 2010; Corden, 2011; Wild and Cramer, 2012; Mirón-García et al., 2013; Gómez-Navarro and Estruch, 2015), and as it is suggestes by both the cytoplasmic accumulation of RNA pol II subunits after blocking biogenesis and by the identification of RNA pol II transport factors (i.e., Iwr1 and Rtp1) suggest (Czeko et al., 2011; Gómez-Navarro et al., 2013). However, additional Iwr1-independent mechanisms have been proposed to allow some RNA pol II subunits to passively diffuse into the nucleus in both yeast and human (Boulon et al., 2010; Gómez-Navarro and Estruch, 2015). Notably, cytoplasmic biogenesis has also been proposed for yeast RNA pol I and III based on the cytoplasmic accumulation of the largest subunits of their RNA pols under impairing assembly (Mirón-García et al., 2013). Although the mechanisms governing the assembly of RNA pol I and III are not as clear, mass spectrometry approaches and dissociation studies of the elongation complexes in yeast have identified the disassembly of RNA pol I and III and probably assembly subcomplexes (Schneider and Nomura, 2004; Lane et al., 2011). In yeast, the RNA pol III core may be assembled in the cytoplasm (Hardeland and Hurt, 2006; Mirón-García et al., 2013), whereas additional subcomplexes or free subunits must bind the core in the nucleus (Hardeland and Hurt, 2006). In fact, there are reports informing that the cytoplasmic accumulation of the second largest subunit Rpc128 also leads to the accumulation of other subunits, such as Rpc160, Rpc53, and Rpc11, whereas others remain nuclear (Hardeland and Hurt, 2006). Interestingly, dissociation analyses by mass spectrometry have evidenced that several subcomplexes appear after disturbing the RNA pol III structure *in vitro*, including stable trimer Rpc31/82/34 and the two heterodimers Rpc82/31 and Rpc17/25, as well as some free subunits like Rpb10, Rpc11, Rpc82, and Rpc34 (Lane et al., 2011). Notably,

human RNA pol I assembly has been proposed to be sequential, but highly inefficient, even *in vivo*, with individual subunits entering the nucleolus rather than the preassembled holoenzyme (Dundr et al., 2002).

ASSEMBLY FACTORS ARE REQUIRED FOR THE ASSEMBLY PROCESSES OF RNA POLYMERASES

In yeast, the assembly of RNA pols and their transport to the nucleus require the action of assembly and/or transport factors, most of which are conserved in human cells. The pre-existing knowledge about their role in the assembly of yeast RNA pols and their comparison with that in humans are summarized in **Table 1**. How these factors act to mediate the sequential assembly of RNA pols and the nuclear transport is shown in **Figure 1**.

R2TP/Prefoldin-like

R2TP was initially identified in yeast as an Hsp90- associated multiprotein complex (R2TP-Hsp90 complex) (Zhao et al., 2005). This complex is well conserved from yeast to humans (Zhao et al., 2005; Boulon et al., 2008). Yeast R2TP components Rvb1/Rvb2 associate independently with whole RNA pol II and the Hsp90 complex (Lakshminarasimhan et al., 2016). R2TP subunits have also been detected in the polysomes interacting with newly synthesized Rpb1 subunits (Villanyi et al., 2014). Furthermore, the co-translational assembly between R2TP, Hsp90, and Rpb1 is mediated by Not5 (Villanyi et al., 2014). Similarly, in human cells, the co-chaperone R2TP complex works with HSP90 in the activation and assembly of several macromolecular complexes, including RNA pol II (Boulon et al., 2010). Proteomic analyses in human cells have evidenced the presence of RPB1-RPB8 dimer (RPB8 also called RPABC3) that interacts with the full R2TP/PFDL complex (Boulon et al., 2010). Human R2TP subunit RPAP3 delivers unassembled RPB1 to HSP90 and also associates with the largest subunits RPA190 (also called RPA1) (Boulon et al., 2010) and RPC160 (also called RPC1) of the free RNA pol I and RNA pol III (Jeronimo et al., 2007), respectively. These data suggest that R2TP subunit RPAP3 may be involved in the assembly of all three RNA pols (Boulon et al., 2010).

The human R2TP complex interacts with components of the prefoldin (PFD) complex (PFDN2 and PFDN6) and the prefoldin-like complex (URI, UXT, and PDRG1) to form the R2TP/prefoldin-like complex (R2TP/PFDL) (Cloutier et al., 2009; Cloutier and Coulombe, 2010; Martínez-Fernández et al., 2018). URI yeast ortholog Bud27 interacts with prefoldin subunits 2 and 6 (Pfd2 and Pfd6) and with the RNA pols common subunit Rpb5 and plays a role in the cytoplasmic assembly of RNA pol I, II, and III (Cloutier et al., 2009; Cloutier and Coulombe, 2010; Mirón-García et al., 2013; Martínez-Fernández et al., 2018).

These data suggest that R2TP, in both yeast and human, participates in the assembly of subassembly complex Rpb1 to the rest of the RNA pol II enzyme and likely occurs for RNA pol I and III.

HSP90

Yeast Hsp90 and human HSP90 are well-conserved molecular chaperones that participate in protein folding and avoid the non-specific aggregation of non-native proteins (Pearl and Prodromou, 2006; Wandinger et al., 2008; Taipale et al., 2010; Makhnevych and Houry, 2012; Schopf et al., 2017).

As indicated above, yeast Hsp90 associates with RNA pol II (Lakshminarasimhan et al., 2016) and human HSP90 and R2TP/PFDL mediate the assembly of RNA pol II through the interaction with the RPB1 subcomplex (Boulon et al., 2010, 2012; Makhnevych and Houry, 2012). Interestingly, yeast Rpa135 (RNA pol I) and Rpc40 (RNA pol III) feature among Hsp90 clients (McClellan et al., 2007). These data suggest that Hsp90 could mediate the assembly of RNA pol I, II, and III. HSP90 is required for RPB1 stabilization through most of the assembly pathway, particularly for the RPB1 subunit association with RPB8 (also called RPABC3) and RPB5 (also called RPABC1) and also with the RPB2-RPB3-RPB10-RPB11-RPB12 subcomplex (also called RPB2-RPB3-RPABC5-RPB11-a and RPABC4), by facilitating the assembly of the complete enzyme (Boulon et al., 2010, 2012). Therefore, HSP90/R2TP could mediate a quality control mechanism for RNA pol II formation by ensuring its correct assembly before its nuclear import (Boulon et al., 2012). Interestingly, and as previously demonstrated for yeast (Lakshminarasimhan et al., 2016), proteomic analyses have also revealed not only RPB1 but also RPA190 (also called RPA1) and RPC160 (also called RPC1) to be HSP90/R2TP interactors, which also suggests their role in the assembly of RNA pol I and III (Boulon et al., 2010).

Bud27

Bud27, and its human ortholog URI, are members of the PFD family of the ATP-independent molecular chaperones considered to function as scaffold proteins capable of assembling additional members of the PFD family in both human and yeast (Gstaiger et al., 2003; Martínez-Fernández and Navarro, 2018). Bud27 contacts the Pfd6 and Pfd2 components of the PFD/GimC complex (Gstaiger et al., 2003; Mirón-García et al., 2013), whereas URI contacts the PFD complex (Gstaiger et al., 2003).

Both Bud27 and URI interact with Rpb5, a common subunit of eukaryotic RNA pols (Dorjsuren et al., 1998; Gstaiger et al., 2003; Mirón-García et al., 2013; Martínez-Fernández and Navarro, 2018). Bud27 mediates the cytoplasmic assembly of the three RNA pols in *Saccharomyces cerevisiae* in an Rpb5-dependent manner before nuclear translocation, which probably occurs similarly for human URI, at least for RNA pol II (Mirón-García et al., 2013). Furthermore, proteomic analyses reveal URI to be an R2TP/PFDL component involved in, at least, RNA pol II assembly in human (Cloutier et al., 2009; Cloutier and Coulombe, 2010; Martínez-Fernández et al., 2018).

Bud27 shuttles between the nucleus and the cytoplasm (Mirón-García et al., 2013) and participates in the transcription mediated by the three RNA pols (Mirón-García et al., 2014; Vernekar and Bhargava, 2015; Martínez-Fernández et al., 2020). These results suggest that Bud27 could be imported to the nucleus in association with RNA pols and then remain associated with

TABLE 1 | Assembly factors.

Assembly factor		Eukaryotic RNA pol	References
S. cerevisiae	**Human**		
R2TP complex	R2TP/prefoldin-like complex	RNA pol II. Also suggested for RNA pol I and III	Boulon et al., 2010
Hsp90	HSP90	RNA pol II. Also suggested for RNA pol I and III	Boulon et al., 2010
Npa3	GPN1/RPAP4/XAB1/MBD*in*	RNA pol II	Forget et al., 2010; Staresincic et al., 2011; Niesser et al., 2015
Gpn2	GPN2	RNA pol II and III	Staresincic et al., 2011; Minaker et al., 2013; Zeng et al., 2018
Gpn3	GPN3/Parcs	RNA pol II and III	Calera et al., 2011; Minaker et al., 2013; Liu et al., 2020
Rba50	RPAP1	RNA pol II	Jeronimo et al., 2004; Zeng et al., 2018; Liu et al., 2020
Rtr1	RPAP2	RNA pol II	Forget et al., 2013; Gómez-Navarro and Estruch, 2015
Bud27	URI	RNA pol I, II, III	Mirón-García et al., 2013; Vernekar and Bhargava, 2015
Rtp1		RNA pol II	Gómez-Navarro et al., 2013
Rbs1		RNA pol III	Cieśla et al., 2015
Iwr1		RNA pol II	Czeko et al., 2011

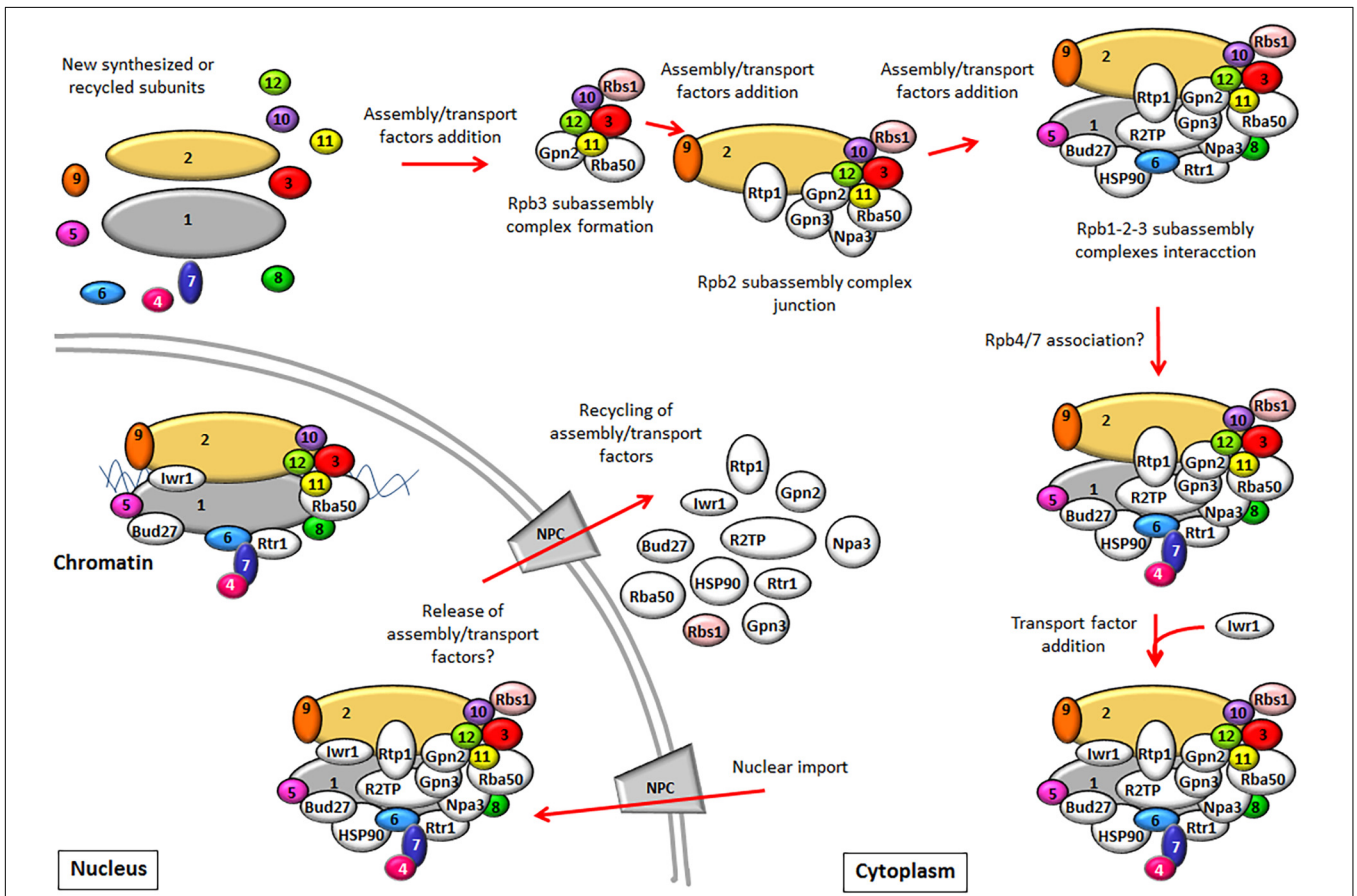


FIGURE 1 | The biogenesis model of RNA pol II (RNA pol I and III assembly has been proposed to be similar). RNA pol II is composed of several subassembly complexes that sequentially interact to form the whole enzyme. The Rpb3 subassembly complex is formed prior to its interaction with the Rpb2 subassembly complex. Later, the Rpb1 subassembly complex junction leads to the core formation of RNA pol II. The association of the Rpb4/7 dimer with the rest of the complex probably occurs at a later step in the cytoplasm, although nuclear association must not be ruled out. RNA pol assembly requires the participation of assembly factors. After the core formation of RNA pol II, some assembly factors could be released, while others could participate in RNA pol II import, such as Npa3 or Rtr1. Import factor Iwr1 binds the active center of the full enzyme prior to its nuclear import. In the nucleus, assembly and import factors release RNA pol II and could be recycled to the cytoplasm, while others like Rtr1, Bud27, and Iwr1 could continue to be associated with RNA pol to mediate a role in transcription. The assembly factors described in yeast (some with human counterparts that play similar roles) are shown in white. Rbs1 is depicted in pink because it has been shown to participate only in the assembly of RNA pol III.

the transcriptional complexes, probably by a tripartite interaction with Rpb5 and remodeler complexes like RSC (Mirón-García et al., 2014; Vernekar and Bhargava, 2015).

The GPN-Loop GTPase Family

GPN-loop GTPase proteins, highly conserved from archaea to humans, contain a highly conserved GPN-loop motif of Gly-Pro-Asn inserted into the GTPase core-fold that functions in GTP hydrolysis (Gras et al., 2007). In yeast, three GPN-loop GTPase, Npa3, Gpn2, and Gpn3 (GPN1, GPN2, and GPN3 in human, respectively), have been described, which evolved from a single archaeal counterpart (Gras et al., 2007; Forget et al., 2010; Carre and Shiekhattar, 2011; Staresincic et al., 2011; Minaker et al., 2013).

The three yeast and human small GTPases have been described to participate in RNA pol II assembly and/or transport to the nucleus (Forget et al., 2010; Carre and Shiekhattar, 2011; Staresincic et al., 2011; Minaker et al., 2013; Niesser et al., 2015; Li et al., 2018; Zeng et al., 2018; Liu et al., 2020). Furthermore, interactions among the three yeast and human members of the GPN-loop GTPase family have been detected (Uetz et al., 2000; Boulon et al., 2010; Forget et al., 2010; Carre and Shiekhattar, 2011; Staresincic et al., 2011; Minaker et al., 2013; Liu et al., 2020). A more detailed overview of these proteins is shown below.

Npa3

Npa3 interacts with RNA pol II and the R2TP complex in yeast (Forget et al., 2010; Niesser et al., 2015). Similarly, the human Npa3 ortholog, GPN1, has been described as an RNA pol II-associated protein that interacts with not only the complex R2TP/PFDL, and the cytosolic chaperonin, CCT, but also with other proteins involved in protein assembly and/or folding (Jeronimo et al., 2007; Forget et al., 2010). Npa3 depletion leads to the cytoplasmic accumulation of Rpb1 and Rpb3 (Staresincic et al., 2011). Cytoplasmic Rpb1 accumulation is also observed in mutants of the Npa3 GTP-binding domain or GPN motifs (similarly for human GPN1) (Forget et al., 2010; Carre and Shiekhattar, 2011; Staresincic et al., 2011). Nevertheless, Rpb3 immunoprecipitation evidences that Npa3 coordinates not only with Gpn3 but also with Rba50 (described in detail in the next paragraph), for the correct association of Rpb1 and Rpb2 with the Rpb3 subcomplex, likely in the cytoplasm (Liu et al., 2020). These data point to a major role of this protein in the cytoplasmic assembly of RNA pol II as has been previously proposed (Niesser et al., 2015). In line with this, two-hybrid assays have shown that the Rpb2 subunit contacts Npa3 and Rba50, which suggests that both proteins may coordinate the Rpb2 subcomplex-dependent assembly of RNA pol II (Liu et al., 2020). In fact, Npa3 is found mainly in the cytoplasm (Huh et al., 2003), but it contains a nuclear export sequence (NES) (Staresincic et al., 2011) that is also conserved in its human ortholog GPN1 (Reyes-Pardo et al., 2012). Npa3 and, similarly human GPN1, translocates to the cytoplasm by the action of the Xpo1/Crm1 pathway (Forget et al., 2010; Carre and Shiekhattar, 2011; Staresincic et al., 2011) and has been proposed to participate mainly in the nuclear import of RNA pol II (Forget et al., 2010; Carre and Shiekhattar, 2011; Staresincic et al., 2011). In line with data in yeast, two-hybrid experiments in human cells have demonstrated the interaction between GPN1 and Rba50 human ortholog RPAP1 and also between GPN1 and RPB2 (Liu et al., 2020). These findings suggest that the role proposed for Npa3 in the assembly of RNA pol II is also conserved in human cells.

Gpn2

Gpn2 is another GPN-loop GTPase family member (Alonso et al., 2013). Similar to Gpn1 and Gpn3, Gpn2 loss-of-function leads to altered cytoplasmic RNA pol II localization, which suggests a role for Gpn2 in the transport of RNA pol II to the nucleus (Staresincic et al., 2011; Minaker et al., 2013; Zeng et al., 2018; Liu et al., 2020). Gpn2 physically interacts with Rba50 and Rpb12 by cooperating to assemble the Rpb3 subcomplex prior to its association with Rpb1 and Rpb2 (Zeng et al., 2018). Yeast *gpn2* mutants show genetic interactions with RNA pol I and III mutants (Minaker et al., 2013). Although *gpn2* mutants affect the localization of both RNA pol II and III, they do not mislocalize RNA pol I subunits (Minaker et al., 2013). These results suggest that Gpn2 acts not only in RNA pol II assembly but also in the RNA pol I and III assembly process.

Gpn3

Gpn3 has been proposed to mediate RNA pols biogenesis, assembly, and transport to the nucleus (Minaker et al., 2013; Liu et al., 2020) and has been demonstrated to form a stable complex with Gpn1 in both yeast and human (Carre and Shiekhattar, 2011; Cristóbal-Mondragón et al., 2019; Liu et al., 2020). In addition, yeast *gpn3* mutants mislocalize RNA pol II and III subunits, which suggests a role for Gpn3 (and Gpn2) in not only RNA pol II but also in RNA pol III assembly and/or transport (Minaker et al., 2013). Gpn3 has been proposed to act upstream of import factor Iwr1 during RNA pol II biogenesis (Minaker et al., 2013). Proteomic analyses have evidenced the human GPN3, as well as GPN1 and GPN2, to be an interactor of the RNA pol II subunits (Forget et al., 2010; Calera et al., 2011). The human GPN1/GPN3 complex associates with RNA pol II in both the nucleus and the cytoplasm (Carre and Shiekhattar, 2011) and specifically binds RPB7 (likely the RPB7/RPB4 dimer) and the C-terminal domain (CTD) of RPB1 *in vitro* (Carre and Shiekhattar, 2011). Furthermore, the depletion of human *GPN3* or *GPN1* by small interfering RNAs (siRNAs) leads to RPB1 cytoplasmic accumulation (Calera et al., 2011; Carre and Shiekhattar, 2011). In line with a role for GPN3 in RNA pol II transport to the nucleus, the Q279* mutation of GPN3, related to cancer, has been described to lead to GPN3 entering the cell nucleus and inhibiting GPN1 nuclear export (Barbosa-Camacho et al., 2017).

By taking these data collectively, the GPN-loop GTPase family would act by favoring the assembly and/or transport of the three RNA pols in yeast, and these roles could be conserved in human cells.

Rba50/RPAP1

Rba50 has been described as a cytoplasmic protein that interacts not only with the Rpb10 subunit in two-hybrid screening (Ito et al., 2001; Huh et al., 2003) but also with Rpb2, Rpb3, and Rpb11 in TAP-tagging analyses (Hazbun et al., 2003).

In yeast, physical and functional interactions have been demonstrated between Rba50 and the small GTPases Gpn2 and Npa3, and these interactions are conserved in human and *Arabidopsis* (Muñoz et al., 2017; Li et al., 2018; Zeng et al., 2018; Liu et al., 2020).

Recent studies propose the coordinated action between Rba50 and Gpn2 in the Rpb3 subcomplex assembly prior to its association with the Rpb2 subassembly complex (Zeng et al., 2018). The association among Rba50, Rpb3, Rpb10, and Rpb11 and between Gpn2 and Rpb12 would allow the Rpb3 subassembly complex formation (Zeng et al., 2018). Although yeast *rba50-3* mutant cells affect Rpb1 distribution, no interactions between Rpb1 and Rba50 or Gpn2 have been identified, which suggests that Rba50 and Gpn2 only transiently associate with the Rpb3 subcomplex and dissociate once Rpb1 is associated during the assembly of RNA pol II (Zeng et al., 2018). The recent observation that Rba50 and Npa3 not only interact but also target Rpb2 during the biogenesis of RNA pol II suggests that Rba50 also associates with Npa3 by increasing its affinity to Rpb2 to facilitate Rpb2 assembly to the previously formed Rpb3 subassembly complex (Liu et al., 2020). As previously indicated, human Rba50 ortholog RPAP1 interacts with GPN1, which associates with RPB2 (Liu et al., 2020). RPAP1 enters the nucleus and has been proposed to be required for the transcription of cell identity genes (those genes regulating the developmental process and fibroblastic/mesenchymal identity) by operating at the interface between the Mediator and RNA pol II (Lynch et al., 2018).

Based on yeast and human data, some authors propose that Rba50 (and probably its human ortholog RPAP1) functions by favoring a platform for other assembly factors like Gpn2 and Npa3 to sequentially mediate the association of RNA pol II subassembly complexes (Liu et al., 2020).

Rtr1/RPAP2

Rtr1 ("regulator of transcription" 1) has been described as a phosphorylated RNA pol II interactor by acting as an S5-P CTD phosphatase during the transition from the initiation of the transcription to elongation *in vivo* (Gibney et al., 2008; Mosley et al., 2009, 2013; Hsu et al., 2014; Smith-Kinnaman et al., 2014; Hunter et al., 2016). Additional roles have been proposed for Rtr1 in transcription and mRNA stability (Mosley et al., 2013; Hsu et al., 2014; Hodko et al., 2016; Victorino et al., 2020) (our unpublished data).

It has been proposed that Rtr1 acts as a nuclear RNA pol II import factor as *RTR1* deletion causes cytoplasmic accumulation of Rpb1 and Rpb2 (Gómez-Navarro and Estruch, 2015). Rtr1 shuttles between the nucleus and cytoplasm in a Crm1-dependent manner (Gibney et al., 2008) and interacts with the nucleocytoplasmic transport protein Ran (a small Ras-like GTPase) (Braunwarth et al., 2003). The deletion of *RTR1* paralog *RTR2* does not cause mislocalization of Rpb1, but it increases in the double *rtr1 rtr2* mutant. These findings suggest that both proteins may play a redundant role in RNA pol II import (Gómez-Navarro and Estruch, 2015). Interestingly, Rtr1 also copurifies with the GTPases Gpn3 and Npa3 (Mosley et al., 2013; Smith-Kinnaman et al., 2014), which are proteins that have also been implicated in the nuclear import of RNA pol II (Staresincic et al., 2011).

Human Rtr1 ortholog RPAP2 has been proposed to act as an RNA pol II import factor, given its silencing results in RPB1 cytoplasmic accumulation (Forget et al., 2013). RPAP2 interacts with RNA pol II through its nuclear retention domain *in vitro* (Forget et al., 2013) and directly binds the RPB6 subunit of the enzyme (Wani et al., 2014). In addition, proteomic analyses of human subassembly complexes have identified that RPAP2 is preferentially associated with the free RPB3 and RPB1 subunits (Boulon et al., 2010), which suggests a role for RPAP2 in the biogenesis of RNA pol II. Furthermore, RPAP2 shuttles to the cytoplasm in association with GPN1 (Forget et al., 2013; Guerrero-Serrano et al., 2017). As in yeast, this evidences the relationship between RPAP2 and small GTPases. Interestingly, *Arabidopsis* Rtr1 ortholog RIMA interacts with MINIYO (the Rba50 yeast ortholog) (Muñoz et al., 2017).

By considering the role of Rtr1 in only the transcription of RNA pol II (Gibney et al., 2008; Mosley et al., 2009; Hsu et al., 2014; Victorino et al., 2020) and the above data, we speculate that this protein is specific to the assembly and/or transport of RNA pol II in yeast and also in other organisms.

Rtp1

Rtp1 has been proposed to be an important factor for nuclear RNA pol II localization *via* Iwr1-independent pathways (Gómez-Navarro et al., 2013). In fact, it has been suggested to participate in transporting RNA pol II through the nuclear pore complex (Gómez-Navarro et al., 2013; Gómez-Navarro and Estruch, 2015). Rtp1 physically interacts with R2TP complex components and also with several RNA pol II subunits (Gómez-Navarro et al., 2013). Mass spectrometry data suggest that Rtp1 can facilitate the interaction between subassembly complexes Rpb2 and Rpb3, and their later interaction with the subassembly complex Rpb1 for the assembly of RNA pol II (Gómez-Navarro et al., 2013). *RTP1* gene depletion in yeast leads to the cytoplasmic accumulation of Rpb1 and Rpb2 (Gómez-Navarro et al., 2013). Interestingly, the fact that the *rtp1* mutant shows no clear cytoplasmic accumulation of small RNA pol II subunits, such as Rpb3 and Rpb11, suggests the existence of passive nuclear diffusion of small RNA pol II subunits (Gómez-Navarro and Estruch, 2015).

Iwr1

Iwr1 was initially reported as a protein that interacts with almost every RNA pol II subunit to regulate the transcription of some genes (Gavin et al., 2002; Peiro-Chova and Estruch, 2009) and was later reported as being important for preinitiation complex formation by all three nuclear RNA pols in *S. cerevisiae* (Esberg et al., 2011). It has been proposed that the main pathway for the nuclear RNA pol II import involves yeast Iwr1 (Czeko et al., 2011). Iwr1 binds the cleft of the active center of RNA pol II once the enzyme is fully assembled and uses its NLS signal to direct the nuclear import of RNA pol II (Czeko et al., 2011). Iwr1 is displaced from active RNA pol II in the nucleus, which facilitates its export and recycling (Czeko et al., 2011; Wild and Cramer, 2012). Nevertheless, *iwr1*Δ mutant strains show cytoplasmic

accumulation of both Rpb1 and Rpb3 (Czeko et al., 2011), which suggests interferences with the assembly of RNA pol II.

In addition to Iwr1-dependent nuclear RNA pol II import, Iwr1-independent import pathways have been proposed as mechanisms that maintain cell viability when the main pathways are blocked, which involves the nuclear import of individual or partially assembled subunits, even by diffusion (Gómez-Navarro and Estruch, 2015).

Rbs1

Rbs1 was originally identified as a PAS kinase suppressor in a genetic high-copy suppressor study (Rutter et al., 2002). More recently, a role for Rbs1 in RNA pol III assembly and transport to the nucleus has been proposed (Cieśla et al., 2015). This role for Rbs1 has also been described in another article published in the same issue by Boguta and Turowski (in press). Rbs1 physically interacts not only with several RNA pol III subunits, such as Rpc19 and Rpc40 but also with the RNA pol common subunit Rpb5 (Cieśla et al., 2015). Rbs1 has been described as a Crm1 interactor, and it shuttles between the nucleus and the cytoplasm. Accordingly, it has been suggested to likely interact with the RNA pol III complex to mediate its nuclear translocation (Cieśla et al., 2015). Rbs1 has also been proposed to mediate the biogenesis of RNA pol III by controlling the steady-state levels of *RPB10* mRNA by interacting with its 3′ UTR region (Cieśla et al., 2020).

CONCLUDING REMARKS

Despite the pre-existing knowledge about RNA pols assembly in yeast and their transport to the nucleus, mainly focused on RNA pol II and likely well-conserved for human RNA pol II, many interesting questions still need answers: although the assembly of RNA pols in eukaryotes seems to be similar to that in bacteria, why do RNA pols-specific processes exist? Why do RNA pol-specific assembly factors exist, while others seem to be general? Can different and additional mechanisms act for the assembly of RNA pols in yeast and other eukaryotes? Which mechanisms account for the nuclear transport of different RNA pols?

Resolving these important questions and others must be the goal of future studies, to help us better understand the mechanisms governing the assembly and nuclear transport of different RNA pols.

AUTHOR CONTRIBUTIONS

All authors wrote and reviewed this manuscript.

FUNDING

This work was supported by grants from the Spanish Ministry of Economy and Competitiveness (MINECO) and ERDF (BFU2016-77728-C3-2-P), the Spanish Ministry of Science and Innovation (MICINN) and ERDF (RED2018-102467-T), the Junta de Andalucía-Universidad de Jaén (FEDER-UJA 1260360), and the Junta de Andalucía (BIO258) to FN. AG-G was financed by the Universidad de Jaen, MINECO, and ERDF funds (BFU2016-77728-C3-2-P to FN). FG-S is a recipient of a predoctoral fellowship from the Universidad de Jaen.

REFERENCES

Alonso, B., Beraud, C., Meguellati, S., Chen, S. W., Pellequer, J. L., Armengaud, J., et al. (2013). Eukaryotic GPN-loop GTPases paralogs use a dimeric assembly reminiscent of archeal GPN. *Cell Cycle* 12, 463–472. doi: 10.4161/cc.23367

Armache, K. J., Kettenberger, H., and Cramer, P. (2003). Architecture of initiation-competent 12-subunit RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6964–6968. doi: 10.1073/pnas.1030608100

Armache, K. J., Mitterweger, S., Meinhart, A., and Cramer, P. (2005). Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *J. Biol. Chem.* 280, 7131–7134. doi: 10.1074/jbc.m413038200

Barbosa-Camacho, A. A., Méndez-Hernández, L. E., Lara-Chacón, B., Peña-Gómez, S. G., Romero, V., González-González, R., et al. (2017). The Gpn3 Q279* cancer−associated mutant inhibits Gpn1 nuclear export and is deficient in RNA polymerase II nuclear targeting. *FEBS Lett.* 591, 3555–3566. doi: 10.1002/1873-3468.12856

Boulon, S., Bertrand, E., and Pradet-Balade, B. (2012). HSP90 and the R2TP co-chaperone complex: building multi-protein machineries essential for cell growth and gene expression. *RNA Biol.* 9, 148–154. doi: 10.4161/rna.18494

Boulon, S., Marmier-Gourrier, N., Pradet-Balade, B., Wurth, L., Verheggen, C., Jady, B. E., et al. (2008). The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. *J. Cell Biol.* 180, 579–595. doi: 10.1083/jcb.200708110

Boulon, S., Pradet-Balade, B., Verheggen, C., Molle, D., Boireau, S., Georgieva, M., et al. (2010). HSP90 and its R2TP/Prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. *Mol. Cell* 39, 912–924. doi: 10.1016/j.molcel.2010.08.023

Braunwarth, A., Fromont-Racine, M., Legrain, P., Bischoff, F. R., Gerstberger, T., Hurt, E., et al. (2003). Identification and characterization of a novel RanGTP-binding protein in the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.* 278, 15397–15405. doi: 10.1074/jbc.m210630200

Briand, J. F., Navarro, F., Rematier, P., Boschiero, C., Labarre, S., Werner, M., et al. (2001). Partners of Rpb8p, a small subunit shared by yeast RNA polymerases I. II and III. *Mol. Cell Biol.* 21, 6056–6065. doi: 10.1128/mcb.21.17.6056-6065.2001

Calera, M. R., Zamora-Ramos, C., Araiza-Villanueva, M. G., Moreno-Aguilar, C. A., Peña-Gómez, S. G., Castellanos-Terán, F., et al. (2011). Parcs/Gpn3 is required for the nuclear accumulation of RNA polymerase II. *Biochim. Biophys. Acta Mol. Cell Res.* 1813, 1708–1716. doi: 10.1016/j.bbamcr.2011.07.005

Carre, C., and Shiekhattar, R. (2011). Human GTPases associate with RNA polymerase II To mediate its nuclear import. *Mol. Cell Biol.* 31, 3953–3962. doi: 10.1128/mcb.05442-11

Cieśla, M., Makała, E., Płonka, M., Bazan, R., Gewartowski, K., Dziembowski, A., et al. (2015). Rbs1, a new protein implicated in RNA polymerase III biogenesis in yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 35, 1169–1181. doi: 10.1128/MCB.01230-14

Cieśla, M., Turowski, T. W., Nowotny, M., Tollervey, D., and Boguta, M. (2020). The expression of Rpb10, a small subunit common to RNA polymerases, is modulated by the R3H domain-containing Rbs1 protein and the Upf1 helicase. *Nucleic Acids Res.* 48, 12252–12268. doi: 10.1093/nar/gkaa1069

Cloutier, P., Al-Khoury, R., Lavallee-Adam, M., Faubert, D., Jiang, H., Poitras, C., et al. (2009). High-resolution mapping of the protein interaction network for the human transcription machinery and affinity purification of RNA polymerase II-associated complexes. *Methods* 48, 381–386. doi: 10.1016/j.ymeth.2009.05.005

Cloutier, P., and Coulombe, B. (2010). New insights into the biogenesis of nuclear RNA polymerases? *Biochem. Cell Biol.* 88, 211–221.

Corden, J. (2011). Going nuclear: transcribers in transit. *Mol. Cell* 42, 143–145. doi: 10.1016/j.molcel.2011.04.001

Cramer, P. (2002). Multisubunit RNA polymerases. *Curr. Opin. Struct. Biol.* 12, 89–97. doi: 10.1016/s0959-440x(02)00294-4

Cramer, P. (2019). Eukaryotic transcription turns 50. *Cell* 179, 808–812. doi: 10.1016/j.cell.2019.09.018

Cristóbal-Mondragón, G. R., Lara-Chacón, B., Santiago, Á, De—la—Rosa, V., González-González, R., Muñiz-Luna, J. A., et al. (2019). FRET—based analysis and molecular modeling of the human GPN—loop GTP ases 1 and 3 heterodimer unveils a dominant—negative protein complex. *FEBS J.* 286, 4797–4818. doi: 10.1111/febs.14996

Czeko, E., Seizl, M., Augsberger, C., Mielke, T., and Cramer, P. (2011). Iwr1 directs RNA polymerase II nuclear import. *Mol. Cell* 42, 261–266. doi: 10.1016/j.molcel.2011.02.033

Dieci, G., Conti, A., Pagano, A., and Carnevali, D. (2012). Identification of RNA polymerase III-transcribed genes in eukaryotic genomes. *Biochim. Biophys. Acta* 1829, 296–305. doi: 10.1016/j.bbagrm.2012.09.010

Dorjsuren, D., Lin, Y., Wei, W., Yamashita, T., Nomura, T., Hayashi, N., et al. (1998). RMP, a novel RNA polymerase II subunit 5-interacting protein, counteracts transactivation by hepatitis B virus X protein. *Mol. Cell Biol.* 18, 7546–7555. doi: 10.1128/mcb.18.12.7546

Dundr, M., Hoffmann-Rohrer, U., Hu, Q., Grummt, I., Rothblum, L. I., Phair, R. D., et al. (2002). A kinetic framework for a mammalian RNA polymerase in vivo. *Science* 298, 1623–1626. doi: 10.1126/science.1076164

Esberg, A., Moqtaderi, Z., Fan, X., Lu, J., Struhl, K., and Bystrom, A. (2011). Iwr1 protein is important for preinitiation complex formation by all three nuclear RNA polymerases in Saccharomyces cerevisiae. *PLoS One* 6:e20829. doi: 10.1371/journal.pone.0020829

Fernández-Tornero, C., Bottcher, B., Rashid, U. J., and Muller, C. W. (2011). Analyzing RNA polymerase III by electron cryomicroscopy. *RNA Biol.* 8, 760–765. doi: 10.4161/rna.8.5.16021

Fernández-Tornero, C., Moreno-Morcillo, M., Rashid, U. J., Taylor, N. M., Ruiz, F. M., Gruene, T., et al. (2013). Crystal structure of the 14-subunit RNA polymerase I. *Nature* 502, 644–649.

Forget, D., Lacombe, A. A., Cloutier, P., Al-Khoury, R., Bouchard, A., Lavallee-Adam, M., et al. (2010). The protein interaction network of the human transcription machinery reveals a role for the conserved GTPase RPAP4/GPN1 and microtubule assembly in nuclear import and biogenesis of RNA polymerase II. *Mol. Cell Proteomics* 9, 2827–2839. doi: 10.1074/mcp.m110.003616

Forget, D., Lacombe, A. A., Cloutier, P., Lavallee-Adam, M., Blanchette, M., and Coulombe, B. (2013). Nuclear import of RNA polymerase II is coupled with nucleocytoplasmic shuttling of the RNA polymerase II-associated protein 2. *Nucleic Acids Res.* 41, 6881–6891. doi: 10.1093/nar/gkt455

Garrido-Godino, A. I., Garcíaa-López, M. C., and Navarro, F. (2013). Correct assembly of RNA polymerase II depends on the foot domain and is required for multiple steps of transcription in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* 33, 3611–3626. doi: 10.1128/mcb.00262-13

Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., et al. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415, 141–147.

Gibney, P. A., Fries, T., Bailer, S. M., and Morano, K. A. (2008). Rtr1 is the *Saccharomyces cerevisiae* homolog of a novel family of RNA polymerase II-binding proteins. *Eukaryot. Cell* 7, 938–948. doi: 10.1128/ec.00042-08

Gómez-Navarro, N., and Estruch, F. (2015). Different pathways for the nuclear import of yeast RNA polymerase II. *Biochim. Biophys. Acta Gene Regul. Mech.* 1849, 1354–1362. doi: 10.1016/j.bbagrm.2015.10.003

Gómez-Navarro, N., Peiro-Chova, L., Rodríguez-Navarro, S., Polaina, J., and Estruch, F. (2013). Rtp1p is a karyopherin-like protein required for RNA polymerase II biogenesis. *Mol. Cell Biol.* 33, 1756–1767. doi: 10.1128/MCB.01449-12

Gras, S., Chaumont, V., Fernandez, B., Carpentier, P., Charrier-Savournin, F., Schmitt, S., et al. (2007). Structural insights into a new homodimeric self—activated GTPase family. *EMBO Rep.* 8, 569–575. doi: 10.1038/sj.embor.7400958

Gstaiger, M., Luke, B., Hess, D., Oakeley, E. J., Wirbelauer, C., Blondel, M., et al. (2003). Control of nutrient-sensitive transcription programs by the unconventional prefoldin URI. *Science* 302, 1208–1212. doi: 10.1126/science.1088401

Guerrero-Serrano, G., Castanedo, L., Cristóbal-Mondragón, G. R., Montalvo-Arredondo, J., Riego-Ruiz, L., DeLuna, A., et al. (2017). Npa3/ScGpn1 carboxy-terminal tail is dispensable for cell viability and RNA polymerase II nuclear targeting but critical for microtubule stability and function. *Biochim. Biophys. Acta Mol. Cell Res.* 1864, 451–462. doi: 10.1016/j.bbamcr.2016.12.010

Haag, J. R., Brower-Toland, B., Krieger, E. K., Sidorenko, L., Nicora, C. D., Norbeck, A. D., et al. (2014). Functional diversification of maize RNA polymerase IV and V subtypes via alternative catalytic subunits. *Cell Rep.* 9, 378–390. doi: 10.1016/j.celrep.2014.08.067

Haag, J. R., and Pikaard, C. S. (2011). Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. *Nat. Rev. Mol. Cell Biol.* 12, 483–492. doi: 10.1038/nrm3152

Hardeland, U., and Hurt, E. (2006). Coordinated nuclear import of RNA polymerase III subunits. *Traffic* 7, 465–473. doi: 10.1111/j.1600-0854.2006.00399.x

Hazbun, T. R., Malmstrom, L., Anderson, S., Graczyk, B. J., Fox, B., Riffle, M., et al. (2003). Assigning function to yeast proteins by integration of technologies. *Mol. Cell* 12, 1353–1365. doi: 10.1016/s1097-2765(03)00476-3

Hodko, D., Ward, T., and Chanfreau, G. (2016). The Rtr1p CTD phosphatase autoregulates its mRNA through a degradation pathway involving the REX exonucleases. *RNA* 22, 559–570. doi: 10.1261/rna.055723.115

Hsu, P. L., Yang, F., Smith-Kinnaman, W., Yang, W., Song, J.-E., Mosley, A. L., et al. (2014). Rtr1 is a dual specificity phosphatase that dephosphorylates Tyr1 and Ser5 on the RNA polymerase II CTD. *J. Mol. Biol.* 426, 2970–2981. doi: 10.1016/j.jmb.2014.06.010

Huh, W. K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S., et al. (2003). Global analysis of protein localization in budding yeast. *Nature* 425, 686–691.

Hunter, G. O., Fox, M. J., Smith-Kinnaman, W. R., Gogol, M., Fleharty, B., and Mosley, A. L. (2016). Phosphatase Rtr1 regulates global levels of serine 5 RNA polymerase II C-terminal domain phosphorylation and cotranscriptional histone methylation. *Mol. Cell. Biol.* 36, 2236–2245. doi: 10.1128/mcb.00870-15

Ishihama, A. (1981). Subunit of assembly of *Escherichia coli* RNA polymerase. *Adv. Biophys.* 14, 1–35. doi: 10.1016/0022-2836(91)90865-4

Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001). A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. U.S.A.* 98, 4569–4574. doi: 10.1073/pnas.061034498

Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., et al. (2007). Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol. Cell* 27, 262–274. doi: 10.1016/j.molcel.2007.06.027

Jeronimo, C., Langelier, M. F., Zeghouf, M., Cojocaru, M., Bergeron, D., Baali, D., et al. (2004). RPAP1, a novel human RNA polymerase II-associated protein affinity purified with recombinant wild-type and mutated polymerase subunits. *Mol. Cell Biol.* 24, 7043–7058. doi: 10.1128/MCB.24.16.7043-7058.2004

Khatter, H., Vorlaender, M. K., and Mueller, C. W. (2017). RNA polymerase I and III: similar yet unique. *Curr. Opin. Struct. Biol.* 47, 88–94. doi: 10.1016/j.sbi.2017.05.008

Kwapisz, M., Beckouet, F., and Thuriaux, P. (2008). Early evolution of eukaryotic DNA-dependent RNA polymerases. *Trends Genet.* 24, 211–215. doi: 10.1016/j.tig.2008.02.002

Lakshminarasimhan, M., Boanca, G., Banks, C. A., Hattem, G. L., Gabriel, A. E., Groppe, B. D., et al. (2016). Proteomic and genomic analyses of the Rvb1 and Rvb2 interaction network upon deletion of R2TP complex components. *Mol. Cell. Proteomics* 15, 960–974. doi: 10.1074/mcp.m115.053165

Lane, L. A., Fernandez-Tornero, C., Zhou, M., Morgner, N., Ptchelkine, D., Steuerwald, U., et al. (2011). Mass spectrometry reveals stable modules in holo and apo RNA polymerases I and III. *Structure* 19, 90–100. doi: 10.1016/j.str.2010.11.009

Li, Y., Yuan, Y., Fang, X., Lu, X., Lian, B., Zhao, G., et al. (2018). A role for MINIYO and QUATRE-QUART2 in the assembly of RNA polymerases II, IV, and V in *Arabidopsis*. *Plant Cell* 30, 466–480. doi: 10.1105/tpc.17.00380

Liu, X., Xie, D., Hua, Y., Zeng, P., Ma, L., and Zeng, F. (2020). Npa3 interacts with Gpn3 and assembly factor Rba50 for RNA polymerase II biogenesis. *FASEB J.* 34, 15547–15558. doi: 10.1096/fj.202001523r

Lynch, C. J., Bernad, R., Calvo, I., Nóbrega-Pereira, S., Ruiz, S., Ibarz, N., et al. (2018). The RNA polymerase II factor RPAP1 is critical for mediator-driven transcription and cell identity. *Cell Rep.* 22, 396–410. doi: 10.1016/j.celrep.2017.12.062

Makhnevych, T., and Houry, W. A. (2012). The role of Hsp90 in protein complex assembly. *Biochim. Biophys. Acta Mol. Cell Res.* 1823, 674–682. doi: 10.1016/j.bbamcr.2011.09.001

Martínez-Fernández, V., Cuevas-Bermúdez, A., Gutiérrez-Santiago, F., Garrido-Godino, A. I., Rodríguez-Galán, O., Jordán-Pla, A., et al. (2020). Prefoldin-like Bud27 influences the transcription of ribosomal components and ribosome biogenesis in *Saccharomyces cerevisiae*. *RNA* 27, 1360–1379. doi: 10.1261/rna.075507.120

Martínez-Fernández, V., Garrido-Godino, A. I., Cuevas-Bermúdez, A., and Navarro, F. (2018). "The yeast Prefoldin Bud27," in *Prefoldins: the New Chaperones*, ed. N. Djouder (Cham: Springer International Publishing), 109–118. doi: 10.1007/978-3-030-00737-9_8

Martínez-Fernández, V., and Navarro, F. (2018). Rpb5, a subunit shared by eukaryotic RNA polymerases, cooperates with prefoldin-like Bud27/URI. *AIMS Genet.* 5, 63–74. doi: 10.3934/genet.2018.1.63

McClellan, A. J., Xia, Y., Deutschbauer, A. M., Davis, R. W., Gerstein, M., and Frydman, J. (2007). Diverse cellular functions of the Hsp90 molecular chaperone uncovered using systems approaches. *Cell* 131, 121–135. doi: 10.1016/j.cell.2007.07.036

Minaker, S. W., Filiatrault, M. C., Ben-Aroya, S., Hieter, P., and Stirling, P. C. (2013). Biogenesis of RNA polymerases II and III requires the conserved GPN small GTPases in *Saccharomyces cerevisiae*. *Genetics* 193, 853–864. doi: 10.1534/genetics.112.148726

Minakhin, L., Bhagat, S., Brunning, A., Campbell, E. A., Darst, S. A., Ebright, R. H., et al. (2001). Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proc. Natl. Acad. Sci. U.S.A.* 98, 892–897. doi: 10.1073/pnas.98.3.892

Mirón-García, M. C., Garrido-Godino, A. I., García-Molinero, V., Hernández-Torres, F., Rodríguez-Navarro, S., and Navarro, F. (2013). The prefoldin bud27 mediates the assembly of the eukaryotic RNA polymerases in an rpb5-dependent manner. *PLoS Genet.* 9:e1003297. doi: 10.1371/journal.pgen.1003297

Mirón-García, M. C., Garrido-Godino, A. I., Martínez-Fernández, V., Fernández-Pévida, A., Cuevas-Bermúdez, A., Martín-Expósito, M., et al. (2014). The yeast prefoldin-like URI-orthologue Bud27 associates with the RSC nucleosome remodeler and modulates transcription. *Nucleic Acids Res.* 42, 9666–9676. doi: 10.1093/nar/gku685

Moreno-Morcillo, M., Taylor, N. M., Gruene, T., Legrand, P., Rashid, U. J., Ruiz, F. M., et al. (2014). Solving the RNA polymerase I structural puzzle. *Acta Crystallogr. Section D Biol. Crystallogr.* 70, 2570–2582.

Mosley, A. L., Hunter, G. O., Sardiu, M. E., Smolle, M., Workman, J. L., Florens, L., et al. (2013). Quantitative proteomics demonstrates that the RNA polymerase II subunits Rpb4 and Rpb7 dissociate during transcriptional elongation. *Mol. Cell. Proteomics* 12, 1530–1538. doi: 10.1074/mcp.m112.024034

Mosley, A. L., Pattenden, S. G., Carey, M., Venkatesh, S., Gilmore, J. M., Florens, L., et al. (2009). Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. *Mol. Cell* 34, 168–178. doi: 10.1016/j.molcel.2009.02.025

Muñoz, A., Mangano, S., González-García, M. P., Contreras, R., Sauer, M., De Rybel, B., et al. (2017). RIMA-dependent nuclear accumulation of IYO triggers auxin-irreversible cell differentiation in *Arabidopsis*. *Plant Cell* 29, 575–588. doi: 10.1105/tpc.16.00791

Niesser, J., Wagner, F. R., Kostrewa, D., Muhlbacher, W., and Cramer, P. (2015). Structure of GPN-Loop GTPase Npa3 and implications for RNA polymerase II assembly. *Mol. Cell Biol.* 36, 820–831. doi: 10.1128/MCB.01009-15

Nouraini, S., Archambault, J., and Friesen, J. D. (1996). Rpo26p, a subunit common to yeast RNA polymerases, is essential for the assembly of RNA polymerases I and II and for the stability of the largest subunits of these enzymes. *Mol. Cell Biol.* 16, 5985–5996. doi: 10.1128/mcb.16.11.5985

Pearl, L. H., and Prodromou, C. (2006). Structure and mechanism of the Hsp90 molecular chaperone machinery. *Annu. Rev. Biochem.* 75, 271–294. doi: 10.1146/annurev.biochem.75.103004.142738

Peiro-Chova, L., and Estruch, F. (2009). The yeast RNA polymerase II-associated factor Iwr1p is involved in the basal and regulated transcription of specific genes. *J. Biol. Chem.* 284, 28958–28967. doi: 10.1074/jbc.m109.012153

Reyes-Pardo, H., Barbosa-Camacho, A. A., Pérez-Mejía, A. E., Lara-Chacón, B., Salas-Estrada, L. A., Robledo-Rivera, A. Y., et al. (2012). A nuclear export sequence in GPN-loop GTPase 1, an essential protein for nuclear targeting of RNA polymerase II, is necessary and sufficient for nuclear export. *Biochim. Biophys. Acta Mol. Cell Res.* 1823, 1756–1766. doi: 10.1016/j.bbamcr.2012.07.001

Rubbi, L., Labarre-Mariotte, S., Chedin, S., and Thuriaux, P. (1999). Functional characterization of ABC10alpha, an essential polypeptide shared by all three forms of eukaryotic DNA-dependent RNA polymerases. *J. Biol. Chem.* 274, 31485–31492. doi: 10.1074/jbc.274.44.31485

Rutter, J., Probst, B. L., and McKnight, S. L. (2002). Coordinate regulation of sugar flux and translation by PAS kinase. *Cell* 111, 17–28. doi: 10.1016/s0092-8674(02)00974-1

Schneider, D. A., and Nomura, M. (2004). RNA polymerase I remains intact without subunit exchange through multiple rounds of transcription in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15112–15117. doi: 10.1073/pnas.0406746101

Schopf, F. H., Biebl, M. M., and Buchner, J. (2017). The HSP90 chaperone machinery. *Nat. Rev. Mol. Cell Biol.* 18, 345–360. doi: 10.1038/nrm.2017.20

Smith-Kinnaman, W. R., Berna, M. J., Hunter, G. O., True, J. D., Hsu, P., Cabello, G. I., et al. (2014). The interactome of the atypical phosphatase Rtr1 in *Saccharomyces cerevisiae*. *Mol. Biosyst.* 10, 1730–1741. doi: 10.1039/c4mb00109e

Staresincic, L., Walker, J., Dirac-Svejstrup, A. B., Mitter, R., and Svejstrup, J. Q. (2011). GTP-dependent binding and nuclear transport of RNA polymerase II by Npa3 protein. *J. Biol. Chem.* 286, 35553–35561. doi: 10.1074/jbc.m111.286161

Taipale, M., Jarosz, D. F., and Lindquist, S. (2010). HSP90 at the hub of protein homeostasis: emerging mechanistic insights. *Nat. Rev. Mol. Cell Biol.* 11, 515–528. doi: 10.1038/nrm2918

Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., et al. (2000). A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* 403, 623–627.

Vernekar, D. V., and Bhargava, P. (2015). Yeast Bud27 modulates the biogenesis of Rpc128 and Rpc160 subunits and the assembly of RNA polymerase III. *Biochim. Biophys. Acta Gene Regul. Mech.* 1849, 1340–1353. doi: 10.1016/j.bbagrm.2015.09.010

Victorino, J. F., Fox, M. J., Smith-Kinnaman, W. R., Justice, S. A. P., Burriss, K. H., Boyd, A. K., et al. (2020). RNA Polymerase II CTD phosphatase Rtr1 fine-tunes transcription termination. *PLoS Genet.* 16:e1008317. doi: 10.1371/journal.pgen.1008317

Villanyi, Z., Ribaud, V., Kassem, S., Panasenko, O. O., Pahi, Z., Gupta, I., et al. (2014). The Not5 subunit of the ccr4-not complex connects transcription and translation. *PLoS Genet.* 10:e1004569. doi: 10.1371/journal.pgen.1004569

Wandinger, S. K., Richter, K., and Buchner, J. (2008). The Hsp90 chaperone machinery. *J. Biol. Chem.* 283, 18473–18477. doi: 10.1074/jbc.r800007200

Wani, S., Hirose, Y., and Ohkuma, Y. (2014). Human RNA polymerase II-associated protein 2 (RPAP2) interacts directly with the RNA polymerase II subunit Rpb6 and participates in pre-mRNA 3′-end formation. *Drug Discover. Ther.* 8, 255–261. doi: 10.5582/ddt.2014.01044

Werner, F. (2007). Structure and function of archaeal RNA polymerases. *Mol. Microbiol.* 65, 1395–1404. doi: 10.1111/j.1365-2958.2007.05876.x

Werner, F., and Grohmann, D. (2011). Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* 9, 85–98. doi: 10.1038/nrmicro2507

Werner, M., Thuriaux, P., and Soutourina, J. (2009). Structure-function analysis of RNA polymerases I and III. *Curr. Opin. Struct. Biol.* 19, 740–745. doi: 10.1016/j.sbi.2009.10.005

Wierzbicki, A. T., Haag, J. R., and Pikaard, C. S. (2008). Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell* 135, 635–648. doi: 10.1016/j.cell.2008.09.035

Wild, T., and Cramer, P. (2012). Biogenesis of multisubunit RNA polymerases. *Trends Biochem. Sci.* 37, 99–105. doi: 10.1016/j.tibs.2011.12.001

Zeng, F., Hua, Y., Liu, X., Liu, S., Lao, K., Zhang, Z., et al. (2018). Gpn2 and Rba50 directly participate in the assembly of the Rpb3 subcomplex in the biogenesis of RNA polymerase II. *Mol. Cell. Biol.* 38:e00091-18.

Zhang, G., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K., and Darst, S. A. (1999). Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 A resolution [see comments]. *Cell* 98, 811–824. doi: 10.1016/s0092-8674(00)81515-9

Zhao, R., Davey, M., Hsu, Y. C., Kaplanek, P., Tong, A., Parsons, A. B., et al. (2005). Navigating the chaperone network: an integrative map of physical and genetic interactions mediated by the hsp90 chaperone. *Cell* 120, 715–727. doi: 10.1016/j.cell.2004.12.024

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Molecular Biosciences

MINI REVIEW
published: 30 April 2021
doi: 10.3389/fmolb.2021.669044

Critical Involvement of TFIIB in Viral Pathogenesis

*Michael J. O'Brien and Athar Ansari**

Department of Biological Science, Wayne State University, Detroit, MI, United States

Viral infections and the harm they cause to their host are a perpetual threat to living organisms. Pathogenesis and subsequent spread of infection requires replication of the viral genome and expression of structural and non-structural proteins of the virus. Generally, viruses use transcription and translation machinery of the host cell to achieve this objective. The viral genome encodes transcriptional regulators that alter the expression of viral and host genes by manipulating initiation and termination steps of transcription. The regulation of the initiation step is often through interactions of viral factors with gene specific factors as well as general transcription factors (GTFs). Among the GTFs, TFIIB (Transcription Factor IIB) is a frequent target during viral pathogenesis. TFIIB is utilized by a plethora of viruses including human immunodeficiency virus, herpes simplex virus, vaccinia virus, Thogoto virus, hepatitis virus, Epstein-Barr virus and gammaherpesviruses to alter gene expression. A number of viral transcriptional regulators exhibit a direct interaction with host TFIIB in order to accomplish expression of their genes and to repress host transcription. Some viruses have evolved proteins with a three-dimensional structure very similar to TFIIB, demonstrating the importance of TFIIB for viral persistence. Upon viral infection, host transcription is selectively altered with viral transcription benefitting. The nature of viral utilization of TFIIB for expression of its own genes, along with selective repression of host antiviral genes and downregulation of general host transcription, makes TFIIB a potential candidate for antiviral therapies.

Keywords: virus, TFIIB, transcription, RNA polymerase II, pathogenesis, gene expression

OPEN ACCESS

Edited by:
Silvia M. L. Barabino,
University of Milano-Bicocca, Italy

Reviewed by:
Alessandro Vannini,
Institute of Cancer Research (ICR), United Kingdom
Linda F. Van Dyk,
University of Colorado Denver, United States

***Correspondence:**
Athar Ansari
bb2749@wayne.edu

Specialty section:
This article was submitted to Protein and RNA Networks, a section of the journal Frontiers in Molecular Biosciences

Received: *17 February 2021*
Accepted: *08 April 2021*
Published: *30 April 2021*

Citation:
O'Brien MJ and Ansari A (2021) Critical Involvement of TFIIB in Viral Pathogenesis. Front. Mol. Biosci. 8:669044. doi: 10.3389/fmolb.2021.669044

INTRODUCTION

Viruses have always been a threat to living creatures. Humans alone are the target of more than 200 viral species (Woolhouse et al., 2012; Knipe et al., 2013). The known viruses and the newly evolving strains, which are being discovered on a regular basis, have the potential to pose a global threat to humanity. They have caused pandemics in the past and the current COVID-19 pandemic is due to a recently evolved strain of coronavirus (Xie and Chen, 2020). It is necessary to identify, understand, and block replication of viruses to combat the hazards they pose. Viruses take advantage of their host to persist, replicate, and ultimately spread to a new host.

Necessary for viral infection is replication of the viral genome and production of viral proteins. Transcription of viral genes is the first step toward production of viral proteins (An et al., 2019; Liu et al., 2020). A number of DNA viruses and retroviruses use host transcription machinery to achieve this objective (Agostini et al., 1996; Gelev et al., 2014; Liu et al., 2020). Transcription is an essential biological process that results in production of RNA from the DNA template, a necessary step before the eventual production of proteins. At the center of transcription is the RNA

polymerase (RNAP), the enzyme responsible for catalyzing RNA synthesis. In eukaryotes, RNA polymerase II (RNAPII) is responsible for synthesizing mRNAs, which subsequently are translated into proteins. The first step of transcription, known as initiation, involves recruitment of RNAPII by gene-specific transcription factors and a suite of general transcription factors (GTFs) on the promoter to form a preinitiation complex (PIC) (Krishnamurthy and Hampsey, 2009). The PIC consists of TFIID, TFIIA, TFIIB, TFIIF, TFIIE, and TFIIH together with RNAPII and Mediator complex (Woychik and Hampsey, 2002; Luse, 2014). Viruses target both gene-specific and GTFs to alter gene expression during pathogenesis. Among gene-specific factors IRF3, pro-inflammatory cytokines, NFκB and STATs are the most common viral targets (Lyles, 2000; Haas et al., 2018). A number of viruses also target GTFs to repress transcription of host anti-viral genes and to transcribe genes coding for viral proteins. The TATA-binding protein (TBP), which is a subunit of TFIID, is the target of HPV16 E7 protein, adenovirus E1A protein, and poliovirus 3C protein; TFIIE is the target of varicella virus IE63; while rift valley fever virus (RVFV) targets the TFIIH complex during viral pathogenesis (**Figure 1**; Maldonado et al., 2002; Dasgupta and Scovell, 2003; di Valentin et al., 2005; Kundu et al., 2005; Kalveram et al., 2011). TFIIB, however, is emerging as a critical viral target.

TFIIB is canonically involved in initiation of transcription by RNAPII and is an essential component of the PIC (Deng and Roberts, 2005; Luse, 2014). It is crucially important for recruitment of RNAPII on the promoter for initiation of transcription and is therefore taken advantage of by a number of viruses to transcribe their essential genes. Viral genomes encode transcriptional regulators, which alter viral and host transcription. A number of these viral regulatory proteins have been shown to interact with host TFIIB (Sundseth and Hansen, 1992; Smith et al., 1993; Yu et al., 1995; Agostini et al., 1996; Haviv et al., 1998; Jang et al., 2001; Vogt et al., 2008). The binding of TFIIB to a template is seemingly the critical, often rate limiting step in viral transcription where initiation will not occur if TFIIB is not present. The interaction of viral transcriptional regulators with TFIIB is therefore critical for transcription of protein-coding genes in a number of viruses. TFIIB has proven to be an important protein in viral transcription to such an extent that some viruses have evolved to possess proteins with structures similar to TFIIB in order to effectively transcribe their genes (Grimm et al., 2019; Cackett et al., 2020). Evidence suggests that TFIIB is not only intricately linked to viral pathogenesis, but is also selectively targeted by viruses to sustain viability. The widespread utilization of host TFIIB and TFIIB-like viral proteins by a number of viruses makes it an important, often essential, aspect of viral pathogenesis and a potential antiviral target.

VIRUSES TARGET TFIIB TO ENHANCE EXPRESSION OF THEIR GENES AND REPRESS HOST ANTIVIRAL GENES

Viruses employ a two-pronged attack on host transcription machinery during pathogenesis. They exploit host transcription machinery to achieve expression of their own genes, and simultaneously downregulate transcription of host genes or gene products especially those linked to antiviral immune response (Gelev et al., 2014; Haas et al., 2018; Yang and You, 2020; Wang et al., 2020). A number of viruses use host transcription machinery to transcribe their protein-coding genes, and a number of them including herpes simplex virus-1 (HSV-1) target the general transcription factor TFIIB to achieve this goal. With RNAi knockdown of TFIIB, HeLa cells transfected with HSV-1 formed fewer plaques compared to the cells containing TFIIB (Gelev et al., 2014). Viral replication and plaque formation were reduced more than sevenfold upon knockdown of TFIIB. In the absence of TFIIB, transcription of most of the herpes genes examined was reduced, while expression of only a subset of host genes was adversely affected (Gelev et al., 2014). Likewise, human immunodeficiency virus (HIV), hepatitis B virus (HBV), human cytomegalovirus (HCMV), equine herpesvirus type 1 (EHV), and Epstein-Barr viruses (EBV) use TFIIB to enhance transcription of their genes (**Figure 1**; Caswell et al., 1993; Tong et al., 1995; Agostini et al., 1996; Haviv et al., 1998; Jang et al., 2001; Simmen et al., 2001).

Viral infection downregulates expression of a number of host genes related to anti-viral immune response. One such host gene normally induced upon viral infection is interferon regulatory factor 3 (IRF-3) (Lyles, 2000). IRF-3 along with IRF-7 are gene-specific transcription activators of type-I IFNs (IFN-α/β), which bind type-I IFN receptor present on the surface of the cell. Upon binding of IFN-α/β, the signal is transduced inside the cell by the signal transducer and activator of transcription STAT-1 and STAT-2. STATs translocate to the nucleus where they induce transcription of several hundred genes encoding proteins that protect the cell from viral infection. Viruses have evolved a variety of strategies to combat activity of such antiviral genes or gene products during pathogenesis. Polio and vesicular stomatitis viruses target TBP, while RVFV destabilizes the general transcription factor TFIIH to induce shut-off of host cell transcription (Yuan et al., 2001; Kundu et al., 2005; Kalveram et al., 2011). Gammaherpesviruses, including Kaposi's sarcoma-associated herpesvirus (KSHV) and murine gammaherpesvirus 68 (MHV68), as well as orthomyxovirus Thogoto virus downregulate TFIIB to repress host gene expression (Hartenian et al., 2020). Gammaherpesvirus adversely affects general host transcription (Hartenian et al., 2020), while Thogoto virus selectively represses transcription of antiviral genes without affecting general host transcription. Thogoto virus specifically inhibits transcription of type-I IFNs (IFN-α/β), inflammatory cytokines, and antiviral effector genes (Jennings et al., 2005; Vogt et al., 2008; Haas et al., 2018). Thus, TFIIB is targeted in multiple ways during viral pathogenesis.

VIRAL EXPRESSION IS INTERTWINED WITH TFIIB

An important issue is how viruses use TFIIB to alter transcription of their own genes and inhibit transcription of selected host

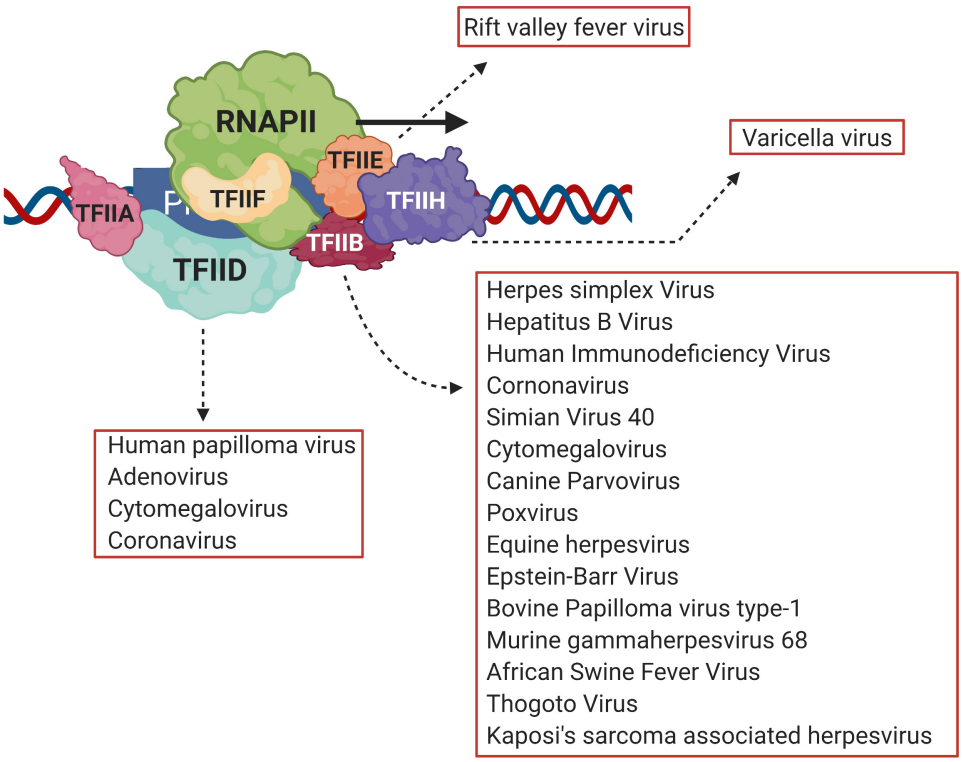


FIGURE 1 | Viral transcriptional regulators target different general transcription factors, but TFIID and TFIIB are the prime targets. Viruses target general transcription factors such as TFIIB, TFIID, TFIIE and TFIIH in order to transcribe viral genes. TFIIB, however, is emerging as the most common target of many different viruses such as HIV, HSV, HCMV and HBV among others. Created with BioRender.com.

genes. Evidence demonstrates that viruses accomplish this objective by physically interacting with TFIIB and recruiting it to the promoter of viral specific genes while preventing it from binding to selected host gene promoters. A number of viral transcriptional regulators like Vpr of HIV, HBx of Hepatitis B, VP16 and ICP4 of HSV, ML of Thogoto virus, IE2 of cytomegalovirus, E2TA of bovine papillomavirus type I, and IE of equine herpes virus exhibit a direct physical interaction with TFIIB (**Table 1**; Lin et al., 1991; Caswell et al., 1993; Rank and Lambert, 1995; Agostini et al., 1996; Haviv et al., 1998; Haas et al., 2018; Li et al., 2020). These viral transcription factors direct transcription of a subset of viral genes that are often vital to viability of the virus. Transcription of all viral genes, however, is not dependent on interaction of viral transcriptional factors with host TFIIB. The mechanism of regulation of transcription of viral and host genes by viral transcriptional regulators through their interaction with TFIIB is described in detail below.

Vpr and Tat: Vpr (Viral Protein R) is a highly conserved HIV-1 encoded protein required for replication, transcription and proliferation of the virus and is critical for viral pathogenicity. Vpr stimulates viral transcription, possibly due to its ability to interact with TFIIB (Agostini et al., 1996; Kino et al., 1999). *In vitro* studies have demonstrated a direct physical interaction of Vpr with TFIIB. Vpr binding brings about a change in the three-dimensional structure of TFIIB from "closed" to "open," which is the transcriptionally active conformation (Agostini et al., 1996). Another HIV-1 protein Tat (Trans-Activator of Transcription), which drastically enhances viral transcription, also targets TFIIB to achieve its transactivation function (Yu et al., 1995). It forms a ternary complex with TFIIB and the human TAP protein. TAP protein binds strongly to a carboxy-terminal region of TFIIB as well as to the conserved activation domain of Tat. TAP acts as a bridge, facilitating interaction of the Tat transactivator with the host general transcription machinery through TFIIB (Yu et al., 1995). A direct protein-protein interaction has been observed between TFIIB and Tat (Veschambre et al., 1997).

VP16: Human herpes simplex virus 1 and 2 (HSV-1 and HSV-2) cause cold sores and genital herpes in humans. Both these viral species are common and contagious. Approximately 67% of the world population under the age of 50 has HSV-1 (James et al., 2020). Viral spread is dependent on the onset of the lytic cycle and reactivation. Viral tegument protein (VP16), which is the transcriptional activator of the immediate-early (IE) gene products (alpha genes), is the key activator of lytic infection (Mossman et al., 2000). VP16-induced transcription of IE genes acts as a regulatory switch; when it is on, it promotes lytic infection, and when it is off latent infection is favored. TFIIB is among several molecular targets of VP16 during transcription of IE genes. VP16 exhibits a physical interaction with both native and recombinant human TFIIB, but not with yeast or fly TFIIB (Lin et al., 1991; Gupta et al., 1996). A mutation in the VP16

TABLE 1 | Viruses, their respective viral proteins, and the mechanism by which they regulate transcription.

Virus	Viral protein	Mechanism of transcriptional regulation
Herpes Simplex Virus I (HSV-I)	VP16	Direct interaction with TFIIB; TFIIB knockdown decreases HSVI gene transcription, viral replication, and plaque formation
Equine Herpesvirus 1 (EHV-1)	IE	Direct interaction with TFIIB; IE utilizes TFIIB to activate EHV-1 promoters
Thogoto Virus (THOV)	ML	Direction interaction with TFIIB; Downregulates TFIIB, represses antiviral gene transcription, and relocalizes TFIIB from nucleus to cytoplasm
Hepatitis B Virus (HBV)	HBx	Interacts with TFIIB to recruit RNAPII on viral promoters.
African Swine Fever Virus (ASFV)	ASFV-RNAP	Encodes a viral homolog similar to TFIIB
Human Immunodeficiency Virus Type 1 (HIV-1)	Vpr/Tat	Vpr: Direct interaction with TFIIB and induces transcriptionally active form of TFIIB. Tat: Forms ternary complex with TFIIB and human TAP protein facilitating interaction with transcriptional machinery
Epstein-Barr virus (EBV)	EBNA-2	Specific interaction with TFIIB crucial for transcription activation potential of EBNA-2
Murine gammaherpesvirus 68 (MHV68)	N/A	Reduced promoter occupancy of RNAPII due to viral induced degradation of RNAPII subunits and TFIIB
Kaposi's sarcoma-associated herpesvirus (KSHV)	N/A	Reduced promoter occupancy of RNAPII due to viral induced degradation of RNAPII subunits and TFIIB
Poxviruses	vRNAP	Viral RNAP contains a subunit with homology to TFIIB.
Canine parvovirus (CBPV)	N/A	Sequesters TFIIB to viral induced nuclear compartment which is site of viral mRNA transcription
Simian virus 40 (SV40)	LSF	Increases rate of association of TFIIB to viral promoters leading to efficient PIC assembly
Cytomegalovirus	IE2	Direct interaction with TFIIB and likely crucial for transcriptional regulation of viral early promoters.
Coronavirus	N/A	TFIIB is a high confidence transcriptional target

activation domain, which adversely affects its transactivation function, reduced its binding to native TFIIB. Structural studies suggest that binding of VP16 brings about a conformational change in TFIIB that primes it for binding to the promoter-bound TBP leading to enhanced transcription of IE genes (Hayashi et al., 1998; Dion and Coulombe, 2003). The infected-cell polypeptide 4 (ICP4) protein of HSV acts as a transcriptional regulator, also affecting VP16 (Gu et al., 1995). The ICP4 protein has been observed to form a complex involving TFIIB in order to alter viral transcription (Smith et al., 1993; Gu et al., 1995).

HBx: Hepatitis B virus (HBV), which causes inflammation of the liver, also manipulates TFIIB during pathogenesis. HBV causes fatal liver infection and chronically infects more than 250 million people worldwide. The HBV minigenome persists in the nucleus of infected cells and is the template for production of four viral proteins. One of these proteins, HBx, is crucial for viral pathogenesis as it interacts with a number of host factors to facilitate viral replication and prevent antiviral response. One of the host proteins that HBx interacts with to enable viral transcription is TFIIB (Turton et al., 2020). HBx directly interacts with TFIIB through its B-finger motif (Haviv et al., 1998; Zhou et al., 2015). This interaction is critical for the coactivator function of HBx. TFIIB-HBx interaction facilitates recruitment of RNAPII on viral promoters leading to upregulation of viral RNA transcription.

IE1 and IE2: Human cytomegalovirus (HCMV) is another virus that targets TFIIB during pathogenesis. It is a common virus that infects people of all ages. Normally the infected people show only mild symptoms, but occasionally it causes serious disorders like mononucleosis and hepatitis. During viral pathogenesis, IE1 and IE2 proteins, which are the product of immediate early genes, transactivate a number of homologous (HCMV) and heterologous (non-HCMV) promoters. Transcription activation potential of IE2 is dependent on its interaction with two GTFs, TBP and TFIIB. IE2 exhibits a direct physical interaction with TFIIB (Caswell et al., 1993). The region of IE2 that mediates binding to TFIIB overlaps with that required for TBP binding. This is also the region linked to the transcriptional regulatory function of the protein. The IE2 gene produces three IE2 protein isoforms, IE2-86, IE2-60, IE2-40, late in infection. IE2-86 is essential for viral replication. An independent study demonstrated binding of IE2-86 to TFIIB under *in vitro* conditions (Xu and Ye, 2002). IE2-86-TFIIB interaction could be crucial for transcriptional regulation of viral early promoters, by facilitating downregulation of its own promoter and activating expression of many host cellular genes necessary for progression of viral infection.

ML: Thogoto virus is an orthomyxovirus that is transmitted to vertebrates through ticks. Thogoto virus targets TFIIB to repress transcription of host antiviral genes linked to innate immunity by a unique mechanism (Vogt et al., 2008; Haas et al., 2018). The ML (Matrix Long) protein of the virus physically interacts with TFIIB and relocalizes it to the cytoplasm (Haas et al., 2018). ML-mediated nuclear depletion of TFIIB represses transcription of genes that require *de novo* recruitment of RNAPII, while transcription of genes with paused polymerases continues unabated (Haas et al., 2018). Among host genes that require *de novo* recruitment of polymerases are antiviral immune

response genes like pro-inflammatory cytokines, IRF3 and its regulated genes, as well as NFκB regulated genes. Sequestration of TFIIB by the ML protein therefore facilitates viral infection by selectively inhibiting host antiviral genes without affecting the bulk of general host transcription.

IE: In addition, equine herpesvirus type 1 (EHV-1) also targets TFIIB during pathogenesis (Jang et al., 2001). At least six viral transcriptional regulators including the IE (Immediate Early) protein regulates the coordinated expression of EHV-1 genes. The transcription activation potential of IE is dependent on its ability to interact with TFIIB. The IE interaction domain of TFIIB spans residues 125–174 in the first direct repeat of the protein. Transient transfection assays demonstrated that exogenous native TFIIB did not perturb transcription activation potential of the IE protein, but a TFIIB mutant that lacked the IE interactive domain adversely affected the ability of the IE protein to activate EHV-1 promoters (Albrecht et al., 2003). These results demonstrate that direct interaction of IE with TFIIB is essential for its ability to activate EHV-1 promoters.

VIRUSES TARGET TFIIB OVER TFIID AND OTHER GENERAL TRANSCRIPTION FACTORS

Since transcription by RNAPII requires at least six GTFs, it is likely that GTFs other than TFIIB are also targeted during viral pathogenesis. There are reports of viruses interacting with the TFIID subunit TBP, TFIIE and TFIIH to achieve transcription of viral genes and turn off host transcription. TBP and TFIIB, however, have emerged as the preferred target of viral transcriptional regulators (Sundseth and Hansen, 1992; Tong et al., 1995; Ihalainen et al., 2012). Transcription regulators of some viruses like SV40, Epstein-Barr virus and canine parvovirus (CPV) interact with both TFIIB and TBP. TFIIB, however, is more critical for viral pathogenesis. In the case of gammaherpesvirus, despite both TFIIB and TFIIA being the viral targets, TFIIB is the key factor for viral transcription (Hartenian et al., 2020).

LSF (Late Simian virus 40 transcription Factor) is a cellular transcriptional activator of SV40 that dramatically increases transcription from viral major late promoters. LSF enhances transcription by facilitating assembly of the PIC on major late promoters. LSF, however, does not affect binding of TFIID to the promoter. Instead, it increases the rate of association of TFIIB, which leads to efficient PIC assembly and increased transcription of SV40 genes (Sundseth and Hansen, 1992). Similarly, Epstein-Barr virus nuclear antigen 2 (EBNA-2) has an acidic domain, which is essential for the transcription activation potential of EBNA-2 (Tong et al., 1995). EBNA-2 exhibits specific interaction with TFIIB, while it's binding affinity for TBP is much less (Tong et al., 1995). The EBNA-2-TFIIB interaction is crucial for transcription activation potential of EBNA-2 during the viral life cycle.

Canine parvovirus (CPV) infection leads to formation of a proteinaceous sub-compartment within the nucleus, which is the site of transcription of viral mRNA. Fluorescence recovery after photobleaching (FRAP) revealed accumulation of both TBP and TFIIB in the nuclear sub-compartment during viral infection. TBP and TFIIB, however, exhibited different kinetics of diffusion and binding affinities to the nuclear sub-compartment (Ihalainen et al., 2012). The binding affinity of TBP to the nuclear body area decreased upon viral infection while that of TFIIB slightly increased. The measured binding time of TFIIB reflected the time scale of TFIIB association with the PIC and recruitment of RNAPII to the promoter, indicating selective usefulness of TFIIB over TBP for transcription of CPV genes in the nuclear body area.

Kaposi's sarcoma-associated herpesvirus (KSHV) and murine gammaherpesvirus 68 (MHV68) infection leads to mRNA-decay induced repression of transcription of antiviral genes and overall general transcription of host cells. Viral induced downregulation of host transcription was due to the genomewide decrease in RNAPII promoter occupancy (Hartenian et al., 2020). Reduced promoter occupancy was attributed to viral induced degradation of subunits of RNAPII and TFIIB. The amount of TFIIA also registered a slight decline upon viral infection. There was, however, absolutely no decrease in the level of TBP in the cell. Clearly, SV40, Epstein-Barr virus, canine parvovirus and gammaherpesviruses target TFIIB over TBP or TFIIA during viral pathogenesis.

VIRUSES HAVE EVOLVED PROTEINS SIMILAR TO TFIIB

Some viruses have evolved proteins with a structure and function similar to that of TFIIB, resulting in less reliance upon the host. Vaccinia virus, which is a prototype poxvirus, encodes its own RNAP (vRNAP). The vRNAP is a multi-subunit protein capable of carrying out transcription in the cytoplasm, not reliant upon host RNAPII to transcribe viral genes (Moss, 2013; Grimm et al., 2019). Replication in the cytoplasm not only rules out the utilization of host RNAPII, but of natively localized host TFIIB as well. The vRNAP complex consists of eight subunits with varying degrees of homology to RNAPII and one subunit, Rap94, displays homology to TFIIB (Grimm et al., 2019; Hillen et al., 2019; Liu et al., 2020). A central region of Rap94 possesses a "B-homology region" containing elements homologous to eukaryotic TFIIB B-ribbon, B-cyclin, and B-reader domains. These regions allow Rap94 to interact with RNAPII, bind to DNA, and determine start site selection in a manner similar to host TFIIB (Nikolov and Burley, 1997; Bushnell et al., 2004; Weinzierl and Wiesler, 2011; Sainsbury et al., 2013; Grimm et al., 2019; Hillen et al., 2019). These similarities suggest that Rap94 likely functions in transcription initiation, bypassing the need to utilize host TFIIB (Grimm et al., 2019). African Swine Fever Virus (ASFV) also encodes a vRNAP with subunits exhibiting homology to RNAPII subunits and TBP. One of the subunits also exhibits remarkable structural and functional similarities to TFIIB (Cackett et al., 2020). For such viruses like vaccinia virus and ASFV, which assemble a viral PIC using a vRNAP complex, the critical usage of host TFIIB is bypassed, as they have evolved TFIIB-like proteins. Concerning the gammaherpesvirus, host TFIIB is used in conjunction with the TFIIB-like viral protein for transcribing

viral genes. Host TFIIB is required for early transcription of viral genes that produces components of the vPIC. Once vPIC proteins are synthesized, the requirement of host TFIIB for transcription of late viral genes is bypassed (Nandakumar and Glaunsinger, 2019). Thus, early transcription during viral life cycle is vPIC-independent, while late transcription is vPIC-dependent. The presence of these TFIIB-like proteins in viral transcription complexes simply reflects the vital role of TFIIB in viral gene expression.

DISCUSSION

TFIIB is a general transcription factor that is essential for initiation of transcription from a majority of RNAPII-transcribed genes in eukaryotes. Viruses can target any general transcription factor to inhibit transcription of antiviral genes for successful pathogenesis. The general target of most viruses, however, are TFIID and TFIIB. The preferential targeting of these GTFs prevents assembly of PICs at an early stage prior to the recruitment of RNAPII. Recent studies have uncovered rather unexpected novel roles of TFIIB in the transcription cycle. The studies have revealed that TFIIB is not merely an initiation factor but plays pleiotropic roles in the transcription cycle (Wang et al., 2010; Medler et al., 2011; Tan-Wong et al., 2012). TFIIB affects gene architecture by facilitating interaction of the terminator with the promoter of the cognate gene during transcription (Medler et al., 2011). The promoter-terminator interaction results in the formation of a looped gene architecture (Ansari and Hampsey, 2005). Gene looping affects termination, reinitiation and promoter directionality (Grzechnik et al., 2014; Al-Husini et al., 2020). Thus, viruses have the potential to affect multiple aspects of the transcription cycle by targeting TFIIB. This could be one of the reasons why viral transcriptional regulators prefer TFIIB over TBP and more so than other GTFs during viral pathogenesis. Future research must focus on which of the TFIIB-dependent processes described above are the target of viral transcriptional regulators during pathogenesis.

The importance of TFIIB in completion of the viral life cycle is corroborated by the fact that multiple viruses have evolved proteins with structural and functional similarity to host TFIIB. Bypassing the need for host TFIIB, these viruses are now self-sufficient in terms of their TFIIB requirement, thus demonstrating the critical role of TFIIB in viral transcription. It is not surprising that TFIIB was also identified along with a number of other proteins as a high confidence transcriptional target (HCT) during infection by the current coronavirus (Ochsner et al., 2020). Taken together, selective downregulation of viral transcription without compromising host gene expression, specific viral targeting of TFIIB over other GTFs, and evolution of viral TFIIB-like proteins are compelling evidence for the critical role of TFIIB in viral pathogenesis. Involvement of TFIIB in pathogenicity of multiple human viruses by altering viral and host gene expression makes TFIIB a potential target of antiviral therapies. Future research concerning the extent of abrogation or inhibition of TFIIB necessary to invoke negative viral responses while maintaining normal host function would further foretell how TFIIB may be targeted to control viral infection. This is likely to be accomplishable as organisms or cells with mutant or inhibited TFIIB have been shown to be viable for study (Gelev et al., 2014). A three-dimensional structure of viral transcription regulators with TFIIB will elucidate the region of TFIIB essential for viral pathogenesis. The regions of TFIIB targeted by virus but not critical for host transcription may be the ideal drug target for future antiviral therapies. The viral TFIIB-like proteins of gammaherpesvirus and ASFV are also the potential drug targets for stopping the lifecycle of these viruses as selective inhibition of vPIC-dependent transcription may stop viral replication without adversely affecting host cell functions.

AUTHOR CONTRIBUTIONS

AA conceptualized and edited the manuscript. MJO wrote the original draft. Both authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by NSF grant MCB-1936030.

ACKNOWLEDGMENTS

We would like to thank Dr. Haidong Gu for critical reading of the manuscript.

REFERENCES

Agostini, I., Navarro, J. M., Rey, F., Bouhamdan, M., Spire, B., Vigne, R., et al. (1996). The human immunodeficiency virus type 1 Vpr transactivator: cooperation with promoter-bound activator domains and binding to TFIIB. *J. Mol. Biol.* 261, 599–606. doi: 10.1006/jmbi.1996.0485

Albrecht, R. A., Jang, H. K., Kim, S. K., and O'Callaghan, D. J. (2003). Direct interaction of TFIIB and the IE protein of equine herpesvirus 1 is required for maximal trans-activation function. *Virology* 316, 302–312. doi: 10.1016/j.virol.2003.08.017

Al-Husini, N., Medler, S., and Ansari, A. (2020). Crosstalk of promoter and terminator during RNA polymerase II transcription cycle. *Biochim. Biophys. Acta Gene Regul. Mech.* 1863:194657. doi: 10.1016/j.bbagrm.2020.194657

An, P., Guo, J. T., and Winkler, C. A. (2019). Host genetics in viral pathogenesis and control. *Front. Genet.* 10:1038. doi: 10.3389/fgene.2019.01038

Ansari, A., and Hampsey, M. (2005). A role for the CPF 3′-end processing machinery in RNAP II-dependent gene looping. *Genes Dev.* 19, 2969–2978. doi: 10.1101/gad.1362305

Bushnell, D. A., Westover, K. D., Davis, R. E., and Kornberg, R. D. (2004). Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. *Science* 303, 983–988. doi: 10.1126/science.1090838

Cackett, G., Sýkora, M., and Werner, F. (2020). Transcriptome view of a killer: African swine fever virus. *Biochem. Soc. Trans.* 48, 1569–1581. doi: 10.1042/bst20191108

Caswell, R., Hagemeier, C., Chiou, C. J., Hayward, G., Kouzarides, T., and Sinclair, J. (1993). The human cytomegalovirus 86K immediate early (IE) 2 protein requires the basic region of the TATA-box binding protein (TBP) for binding, and interacts with TBP and transcription factor TFIIB via regions of IE2 required for transcriptional regulation. *J. Gen. Virol.* 74, 2691–2698. doi: 10.1099/0022-1317-74-12-2691

Dasgupta, A., and Scovell, W. M. (2003). TFIIA abrogates the effects of inhibition by HMGB1 but not E1A during the early stages of assembly of the transcriptional preinitiation complex. *Biochim. Biophys. Acta Gene Struct. Expr.* 1627, 101–110. doi: 10.1016/s0167-4781(03)00080-0

Deng, W., and Roberts, S. G. (2005). A core promoter element downstream of the TATA box that is recognized by TFIIB. *Genes Dev.* 19, 2418–2423. doi: 10.1101/gad.342405

di Valentin, E., Bontems, S., Habran, L., Jolois, O., Markine-Goriaynoff, N., Vanderplasschen, A., et al. (2005). Varicella-zoster virus IE63 protein represses the basal transcription machinery by disorganizing the pre-initiation complex. *Biol. Chem.* 386, 255–267. doi: 10.1515/bc.2005.031

Dion, V., and Coulombe, B. (2003). Interactions of a DNA-bound transcriptional activator with the TBP-TFIIA-TFIIB-promoter quaternary complex. *J. Biol. Chem.* 278, 11495–11501. doi: 10.1074/jbc.m211938200

Gelev, V., Zabolotny, J. M., Lange, M., Hiromura, M., Yoo, S. W., Orlando, J. S., et al. (2014). A new paradigm for transcription factor TFIIB functionality. *Sci. Rep.* 4:3664.

Grimm, C., Hillen, H. S., Bedenk, K., Bartuli, J., Neyer, S., Zhang, Q., et al. (2019). Structural basis of poxvirus transcription: vaccinia RNA polymerase complexes. *Cell* 179, 1537–1550. doi: 10.1016/j.cell.2019.11.024

Grzechnik, P., Tan-Wong, S. M., and Proudfoot, N. J. (2014). Terminate and make a loop: regulation of transcriptional directionality. *Trends Biochem. Sci.* 39, 319–327. doi: 10.1016/j.tibs.2014.05.001

Gu, B., Kuddus, R., and DeLuca, N. A. (1995). Repression of activator-mediated transcription by herpes simplex virus ICP4 via a mechanism involving interactions with the basal transcription factors TATA-binding protein and TFIIB. *Mol. Cell. Biol.* 15, 3618–3626. doi: 10.1128/mcb.15.7.3618

Gupta, R., Emili, A., Pan, G., Xiao, H., Shales, M., Greenblatt, J., et al. (1996). Characterization of the interaction between the acidic activation domain of VP16 and the RNA polymerase II initiation factor TFIIB. *Nucleic Acids Res.* 24, 2324–2330. doi: 10.1093/nar/24.12.2324

Haas, D. A., Meiler, A., Geiger, K., Vogt, C., Preuss, E., Kochs, G., et al. (2018). Viral targeting of TFIIB impairs de novo polymerase II recruitment and affects antiviral immunity. *PLoS Pathog.* 14:e1006980. doi: 10.1371/journal.ppat.1006980

Hartenian, E., Gilbertson, S., Federspiel, J. D., Cristea, I. M., and Glaunsinger, B. A. (2020). RNA decay during gammaherpesvirus infection reduces RNA polymerase II occupancy of host promoters but spares viral promoters. *PLoS Pathog.* 16:e1008269. doi: 10.1371/journal.ppat.1008269

Haviv, I., Shamay, M., Doitsh, G., and Shaul, Y. (1998). Hepatitis B virus pX targets TFIIB in transcription coactivation. *Mol. Cell. Biol.* 18, 1562–1569. doi: 10.1128/mcb.18.3.1562

Hayashi, F., Ishima, R., Liu, D., Tong, K. I, Kim, S., Reinberg, D., et al. (1998). Human general transcription factor TFIIB: conformational variability and interaction with VP16 activation domain. *Biochemistry* 37, 7941–7951. doi: 10.1021/bi9801098

Hillen, H. S., Bartuli, J., Grimm, C., Dienemann, C., Bedenk, K., Szalay, A. A., et al. (2019). Structural basis of poxvirus transcription: transcribing and capping vaccinia complexes. *Cell* 179, 1525–1536. doi: 10.1016/j.cell.2019.11.023

Ihalainen, T. O., Willman, S. F., Niskanen, E. A., Paloheimo, O., Smolander, H., Laurila, J. P., et al. (2012). Distribution and dynamics of transcription-associated proteins during parvovirus infection. *J. Virol.* 86, 13779–13784. doi: 10.1128/jvi.01625-12

James, C., Harfouche, M., Welton, N. J., Turner, K. M., Abu-Raddad, L. J., Gottlieb, S. L., et al. (2020). Herpes simplex virus: global infection prevalence and incidence estimates, 2016. *Bull. World Health Organ.* 98, 315–329. doi: 10.2471/blt.19.237149

Jang, H. K., Albrecht, R. A., Buczynski, K. A., Kim, S. K., Derbigny, W. A., and O'Callaghan, D. J. (2001). Mapping the sequences that mediate interaction of the equine herpesvirus 1 immediate-early protein and human TFIIB. *J. Virol.* 75, 10219–10230. doi: 10.1128/jvi.75.21.10219-10230.2001

Jennings, S., Martínez-Sobrido, L., García-Sastre, A., Weber, F., and Kochs, G. (2005). Thogoto virus ML protein suppresses IRF3 function. *Virology* 331, 63–72. doi: 10.1016/j.virol.2004.10.015

Kalveram, B., Lihoradova, O., and Ikegami, T. (2011). NSs protein of rift valley fever virus promotes posttranslational downregulation of the TFIIH subunit p62. *J. Virol.* 85, 6234–6243. doi: 10.1128/jvi.02255-10

Kino, T., Gragerov, A., Kopp, J. B., Stauber, R. H., Pavlakis, G. N., and Chrousos, G. P. (1999). The HIV-1 virion-associated protein vpr is a coactivator of the human glucocorticoid receptor. *J. Exp. Med.* 189, 51–62. doi: 10.1084/jem.189.1.51

Knipe, D. M., Howley, P. M., Griffin, D. E., Lamb, R. A., Martin, M. A., Roizman, B., et al. (2013). *Fields Virology*. Philadelphia, PA: Lippincott Williams & Wilkins.

Krishnamurthy, S., and Hampsey, M. (2009). Eukaryotic transcription initiation. *Curr. Biol.* 19, R153–R156.

Kundu, P., Raychaudhuri, S., Tsai, W., and Dasgupta, A. (2005). Shutoff of RNA polymerase II transcription by poliovirus involves 3C protease-mediated cleavage of the TATA-binding protein at an alternative site: incomplete shutoff of transcription interferes with efficient viral replication. *J. Virol.* 79, 9702–9713. doi: 10.1128/jvi.79.15.9702-9713.2005

Li, M., Ball, C. B., Collins, G., Hu, Q., Luse, D. S., Price, D. H., et al. (2020). Human cytomegalovirus IE2 drives transcription initiation from a select subset of late infection viral promoters by host RNA polymerase II. *PLoS Pathog.* 16:e1008402. doi: 10.1371/journal.ppat.1008402

Lin, Y. S., Ha, I., Maldonado, E., Reinberg, D., and Green, M. R. (1991). Binding of general transcription factor TFIIB to an acidic activating region. *Nature* 353, 569–571. doi: 10.1038/353569a0

Liu, X., Hong, T., Parameswaran, S., Ernst, K., Marazzi, I., Weirauch, M. T., et al. (2020). Human virus transcriptional regulators. *Cell* 182, 24–37. doi: 10.1016/j.cell.2020.06.023

Luse, D. S. (2014). The RNA polymerase II preinitiation complex: through what pathway is the complex assembled? *Transcription* 5:e27050. doi: 10.4161/trns.27050

Lyles, D. S. (2000). Cytopathogenesis and inhibition of host gene expression by RNA viruses. *Microbiol. Mol. Biol. Rev.* 64, 709–724. doi: 10.1128/mmbr.64.4.709-724.2000

Maldonado, E., Cabrejos, M. E., Banks, L., and Allende, J. E. (2002). Human papillomavirus-16 E7 protein inhibits the DNA interaction of the TATA binding transcription factor. *J. Cell. Biochem.* 85, 663–669. doi: 10.1002/jcb.10172

Medler, S., Al Husini, N., Raghunayakula, S., Mukundan, B., Aldea, A., and Ansari, A. (2011). Evidence for a complex of transcription factor IIB (TFIIB) with poly (A) polymerase and cleavage factor 1 subunits required for gene looping. *J. Biol. Chem.* 286, 33709–33718. doi: 10.1074/jbc.m110.193870

Moss, B. (2013). "Poxviridae," in *Fields Virology*, Vol. 2, eds D. M. Knipe, and P. M. Howley (Philadelphia, PA: Lippincott Williams & Wilkins), 2129–2159.

Mossman, K. L., Sherburne, R., Lavery, C., Duncan, J., and Smiley, J. R. (2000). Evidence that herpes simplex virus VP16 is required for viral egress downstream of the initial envelopment event. *J. Virol.* 74, 6287–6299. doi: 10.1128/jvi.74.14.6287-6299.2000

Nandakumar, D., and Glaunsinger, B. (2019). An integrative approach identifies direct targets of the late viral transcription complex and an expanded promoter recognition motif in Kaposi's sarcoma-associated herpesvirus. *PLoS Pathog.* 15:e1007774. doi: 10.1371/journal.ppat.1007774

Nikolov, D. B., and Burley, S. K. (1997). RNA polymerase II transcription initiation: a structural view. *Proc. Natl. Acad. Sci. U.S.A.* 94, 15–22. doi: 10.1073/pnas.94.1.15

Ochsner, S. A., Pillich, R. T., and McKenna, N. J. (2020). Consensus transcriptional regulatory networks of coronavirus-infected human cells. *Sci. Data* 7:314.

Rank, N. M., and Lambert, P. F. (1995). Bovine papillomavirus type 1 E2 transcriptional regulators directly bind two cellular transcription factors, TFIID and TFIIB. *J. Virol.* 69, 6323–6334. doi: 10.1128/jvi.69.10.6323-6334.1995

Sainsbury, S., Niesser, J., and Cramer, P. (2013). Structure and function of the initially transcribing RNA polymerase II–TFIIB complex. *Nature* 493, 437–440. doi: 10.1038/nature11715

Simmen, K. A., Singh, J., Luukkonen, B. M., Lopper, M., Bittner, A., Miller, N. E., et al. (2001). Global modulation of cellular transcription by human cytomegalovirus is initiated by viral glycoprotein B. *Proc. Natl. Acad. Sci. U.S.A.* 98, 7140–7145. doi: 10.1073/pnas.121177598

Smith, C. A., Bates, P., Rivera-Gonzalez, R., Gu, B., and DeLuca, N. A. (1993). ICP4, the major transcriptional regulatory protein of herpes simplex virus type 1, forms a tripartite complex with TATA-binding protein and TFIIB. *J. Virol.* 67, 4676–4687. doi: 10.1128/jvi.67.8.4676-4687.1993

Sundseth, R., and Hansen, U. (1992). Activation of RNA polymerase II transcription by the specific DNA-binding protein LSF. Increased rate of binding of the basal promoter factor TFIIB. *J. Biol. Chem.* 267, 7845–7855. doi: 10.1016/s0021-9258(18)42591-4

Tan-Wong, S. M., Zaugg, J. B., Camblong, J., Xu, Z., Zhang, D. W., Mischo, H. E., et al. (2012). Gene loops enhance transcriptional directionality. *Science* 338, 671–675. doi: 10.1126/science.1224350

Tong, X., Wang, F., Thut, C. J., and Kieff, E. (1995). The Epstein-Barr virus nuclear protein 2 acidic domain can interact with TFIIB, TAF40, and RPA70 but not with TATA-binding protein. *J. Virol.* 69, 585–588. doi: 10.1128/jvi.69.1.585-588.1995

Turton, K. L., Meier-Stephenson, V., Badmalia, M. D., Coffin, C. S., and Patel, T. R. (2020). Host transcription factors in hepatitis B virus RNA synthesis. *Viruses* 12:160. doi: 10.3390/v12020160

Veschambre, P., Roisin, A., and Jalinot, P. (1997). Biochemical and functional interaction of the human immunodeficiency virus type 1 Tat transactivator with the general transcription factor TFIIB. *J. Gen. Virol.* 78, 2235–2245. doi: 10.1099/0022-1317-78-9-2235

Vogt, C., Preuss, E., Mayer, D., Weber, F., Schwemmle, M., and Kochs, G. (2008). The interferon antagonist ML protein of Thogoto virus targets general transcription factor IIB. *J. Virol.* 82, 11446–11453. doi: 10.1128/jvi.01284-08

Wang, X., Hennig, T., Whisnant, A. W., Erhard, F., Prusty, B. K., Friedel, C. C., et al. (2020). Herpes simplex virus blocks host transcription termination via the bimodal activities of ICP27. *Nat. Commun.* 11:293.

Wang, Y., Fairley, J. A., and Roberts, S. G. (2010). Phosphorylation of TFIIB links transcription initiation and termination. *Curr. Biol.* 20, 548–553. doi: 10.1016/j.cub.2010.01.052

Weinzierl, R. O., and Wiesler, S. C. (2011). Revealing the functions of TFIIB. *Transcription* 2, 254–257. doi: 10.4161/trns.2.6.18076

Woolhouse, M., Scott, F., Hudson, Z., Howey, R., and Chase-Topping, M. (2012). Human viruses: discovery and emergence. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 2864–2871. doi: 10.1098/rstb.2011.0354

Woychik, N. A., and Hampsey, M. (2002). The RNA polymerase II machinery: structure illuminates function. *Cell* 108, 453–463. doi: 10.1016/s0092-8674(02)00646-3

Xie, M., and Chen, Q. (2020). Insight into 2019 novel coronavirus—an updated interim review and lessons from SARS-CoV and MERS-CoV. *Int. J. Infect. Dis.* 94, 119–124. doi: 10.1016/j.ijid.2020.03.071

Xu, J., and Ye, L. (2002). Human cytomegalovirus IE2 protein interacts with transcription activating factors. *Sci. China Ser. C Life Sci.* 45, 604–612. doi: 10.1360/02yc9067

Yang, J. F., and You, J. (2020). Regulation of polyomavirus transcription by viral and cellular factors. *Viruses* 12:1072. doi: 10.3390/v12101072

Yu, L., Loewenstein, P. M., Zhang, Z., and Green, M. (1995). In vitro interaction of the human immunodeficiency virus type 1 Tat transactivator and the general transcription factor TFIIB with the cellular protein TAP. *J. Virol.* 69, 3017–3023. doi: 10.1128/jvi.69.5.3017-3023.1995

Yuan, H., Puckett, S., and Lyles, D. S. (2001). Inhibition of host transcription by vesicular stomatitis virus involves a novel mechanism that is independent of phosphorylation of TATA-binding protein (TBP) or association of TBP with TBP-associated factor subunits. *J. Virol.* 75, 4453–4458. doi: 10.1128/jvi.75.9.4453-4458.2001

Zhou, Q., Huang, F., Chen, L., Chen, E., Bai, L., Cheng, X., et al. (2015). RPB5-mediating protein suppresses hepatitis B virus (HBV) transcription and replication by counteracting the transcriptional activation of hepatitis B virus X protein in HBV replication mouse model. *Jundishapur J. Microbiol.* 8:e21936.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Molecular Biosciences

REVIEW
published: 05 May 2021
doi: 10.3389/fmolb.2021.651134

Early Evolution of Transcription Systems and Divergence of Archaea and Bacteria

Lei Lei[1] *and Zachary F. Burton*[2]*

[1] Department of Biology, University of New England, Biddeford, ME, United States, [2] Department of Biochemistry and Molecular Biology, Michigan State University, E. Lansing, MI, United States

DNA template-dependent multi-subunit RNA polymerases (RNAPs) found in all three domains of life and some viruses are of the two-double-Ψ-β-barrel (DPBB) type. The 2-DPBB protein format is also found in some RNA template-dependent RNAPs and a major replicative DNA template-dependent DNA polymerase (DNAP) from Archaea (PolD). The 2−DPBB family of RNAPs and DNAPs probably evolved prior to the last universal common cellular ancestor (LUCA). Archaeal Transcription Factor B (TFB) and bacterial σ factors include homologous strings of helix-turn-helix units. The consequences of TFB-σ homology are discussed in terms of the evolution of archaeal and bacterial core promoters. Domain-specific DPBB loop inserts functionally connect general transcription factors to the RNAP active site. Archaea appear to be more similar to LUCA than Bacteria. Evolution of bacterial σ factors from TFB appears to have driven divergence of Bacteria from Archaea, splitting the prokaryotic domains.

Keywords: archaea, bacteria, double-Ψ-β-barrel, general transcription factor evolution, promoter evolution, transcription, transcription factor B, sigma factor

OPEN ACCESS

Edited by:
Athar Ansari,
Wayne State University, United States

Reviewed by:
Ludovic Sauguet,
Institut Pasteur, France
Richard Burgess,
University of Wisconsin-Madison, United States

***Correspondence:**
Zachary F. Burton
burton@msu.edu

Specialty section:
This article was submitted to Protein and RNA Networks, a section of the journal Frontiers in Molecular Biosciences

Received: *08 January 2021*
Accepted: *06 April 2021*
Published: *05 May 2021*

Citation:
Lei L and Burton ZF (2021) Early Evolution of Transcription Systems and Divergence of Archaea and Bacteria. Front. Mol. Biosci. 8:651134. doi: 10.3389/fmolb.2021.651134

INTRODUCTION

The purpose of this review is to provide a conceptual overview of transcription systems in the early phase of their evolution, in order to explain how RNA polymerases (RNAPs), general transcription factors and promoters may have evolved. The review also touches on the divergence of Archaea and Bacteria that appears to have partly been driven by the divergence of transcription systems. The proper way to view structures is using molecular graphics such as UCSF ChimeraX (Goddard et al., 2018; Pettersen et al., 2021). Viewing structures in 2−dimensions is challenging to the human eyes and mind. We recommend downloading ChimeraX, running tutorials and using it to follow along with this manuscript. For instance, some figures in this paper are difficult to fully appreciate without a more 3-dimensional representation.

Our opinion is that analyzing the structure-function-dynamics of any protein requires a combination of approaches: i.e., (1) structure analysis; (2) evolution; (3) functional studies; and (4) dynamics. To appreciate structural analysis and dynamics, Cryo-electron microscopy

Abbreviations: BH, bridge helix; BRE, TFB-recognition element; CLR, cyclin-like repeat (TFB HTH domains); DNAP, DNA polymerase; DPBB, double-Ψ - β-barrel; HTH, helix-turn-helix; InR, initiator element; LUCA, last universal (cellular) common ancestor; Pfu, *Pyrococcus furiosis*; Pol, DNA polymerases (i.e., PolA, PolB, PolC, and PolD); PPE, promoter-proximal element; RNAP, RNA polymerase; RRM, RNA-recognition motif; SBHM, sandwich barrel hybrid motif; Sso, *Sulfolobus* solfataricus; TBP, TATA-box binding protein; TFB, transcription factor B; TFE, transcription factor E; TIM, triose phosphate isomerase; TL, trigger loop.

becomes an ever more powerful tool. Cryo-EM provides ensembles of structures often indicating a dynamic progression through a reaction mechanism. Evolutionary studies have the potential to dissect a protein into its component parts to better appreciate how the protein came to have its eventual form and function. In some cases, structural studies have not been combined fully with evolutionary studies, and the historic naming of protein domains can be confusing. Also, very large structures are difficult to analyze unless they can be broken into component parts. We see two potential problems. Without an evolutionary view, structures may be difficult to understand and analyze. Also, the evolution literature can be complex and challenging to read unless one is reasonably expert or determined. In this paper, we attempt to apply a combination of structural and evolutionary principles to the analysis and description of multi-subunit RNAPs, general transcription factors and promoters.

EVOLUTION OF 2-DPBB RNAPs AND DNAPs

2-Double-Ψ-β-Barrel Type RNAPs

Near the dawn of evolution of life on Earth, RNAPs of the 2-DPBB type evolved (Iyer et al., 2003; Lane and Darst, 2010a,b; Werner and Grohmann, 2011; Iyer and Aravind, 2012; Fouqueau et al., 2017; Sauguet, 2019; Madru et al., 2020; Zatopek et al., 2020). These enzymes are found in all domains of life and some viruses. 2-DPBB RNAPs can be either RNA template-dependent or DNA template-dependent, indicating that this important class of enzyme may have arisen in an RNA world before DNA genomes became prominent. The DPBB is a particular fold of cradle-loop barrel (**Figure 1**; Coles et al., 2005, 2006; Alva et al., 2008). The crossing chains make a Ψ pattern, hence the barrel name. 2−DPBB type RNAPs have 2-DPBBs at their active sites (**Figure 2**). Loops from the barrels hold the two Mg^{2+} that retain the phosphates of the NTP substrate and activate the RNA 3′-O to catalyze NMP addition. In addition to the 2-DPBBs, both RNA and DNA template-dependent RNAPs have a bridge helix and trigger loop, indicating that these elements are ancient (Salgado et al., 2006; Iyer and Aravind, 2012; Qian et al., 2016). In DNA template-dependent RNAPs, the β-subunit DPBB1 has a sandwich-barrel hybrid motif (SBHM) inserted into one of the barrel loops (Lane and Darst, 2010a,b; Iyer and Aravind, 2012; Fouqueau et al., 2017). The SBHM loop extension forms the historically-named "flap" or "wall" motif in multi-subunit RNAPs.

Barrels are frequent motifs in ancient evolution. In earliest evolution, barrels were selected to form compact, structured units with reasonable solubility and structural closure (Burton et al., 2016). For instance, 8−β-sheet barrels [$(\beta-\alpha)_8$; i.e., TIM barrels (TIM for triose phosphate isomerase)] are found in most glycolytic enzymes. Rossmann folds appear to be sheets that are rearranged from $(\beta-\alpha)_8$ barrels. Most of the citric acid cycle is made up of Rossmann fold proteins. So, much of core metabolism was generated from barrels and, also, from refolded barrels rendered into more linear sheets. Cradle-loop barrels are a similar ancient evolution story (Alva et al., 2008). If early evolution was partly a race to form stable and soluble scaffolds, formation of barrels helped to build these and, among other possible advantages, helped to avoid generation of β−sheet amyloids and liquid-liquid phase separated compartments that resisted ordered protein folding. Clearly, barrels were a successful evolutionary innovation that, once formed, persisted throughout evolution. From this point of view, an important evolutionary event can be viewed as the race to form stable and soluble protein structures with a degree of structural closure. Barrels were typically formed in evolution by repeated motif duplications, so barrels often won races to higher order structure, solubility and closure. After generation of barrels, primitive catalytic sites could be modified to generate many new, more efficient and more specific enzyme functions. So, for instance, in metabolism, an enzyme with broad specificity built around an 8−β-sheet barrel was duplicated genetically many times and then refined, generating specialist enzymes that formed a more sophisticated and integrated pathway (i.e., glycolysis).

Similarly, the DPBB evolved by duplication of a $\beta-\beta-\alpha-\beta$ unit followed by refolding into a barrel (Alva et al., 2008; Burton et al., 2016). In **Figure 1**, a $\beta-\beta-\alpha-\beta--\beta-\beta-\alpha-\beta$ DPBB enzyme domain is shown in which the basic DPBB form is preserved without much modification (Coles et al., 1999). The β−sheets are numbered 1−6, so that the chain can be traced. The α−helices are numbered 1 and 2. The Ψ patterns of the crossing chains are indicated. The ability to identify a DPBB helps with understanding the 2−DPBB enzyme patterns when analyzing more complex structures. Because of modifications of the pattern during evolution or disorder in structures, DPBBs can be a challenge to identify and, in a complex structure, can be potentially difficult to locate.

2−DPBB type enzymes include RNA template−dependent RNAPs (found in some Eukaryotes), multi−subunit RNAPs (found in all domains and some viruses) and DNA template−dependent DNAPs (PolD in most Archaea) (**Figures 2**, **3**; Iyer et al., 2003; Lane and Darst, 2010a,b; Werner and Grohmann, 2011; Iyer and Aravind, 2012; Fouqueau et al., 2017; Koonin et al., 2020). In 2−DPBB type enzymes, the basic $\beta-\beta-\alpha-\beta--\beta-\beta-\alpha-\beta$ form can be modified by insertions into barrel loops. In RNA template-dependent 2−DPBB RNAPs, neither DPBB1 (corresponding to the β−subunit DPBB1 in 2−DPBB bacterial RNAPs) nor DPBB2 (corresponding to the β′-subunit DPBB2 in 2−DPBB bacterial RNAPs) includes very large inserts or modifications in the basic DPBB pattern (Salgado et al., 2006; Iyer and Aravind, 2012; Qian et al., 2016).

In DNA template-dependent 2−DPBB type RNAPs, by contrast, there are large identifying inserts (Iyer and Aravind, 2012). Significantly, the β−subunit (referring to bacterial RNAPs) DPBB1, includes a sandwich-barrel hybrid motif (SBHM) inserted between β2 and β3 after α1. The SBHM can be recognized because it includes long β−sheets. The SBHM forms the "flap" or "wall" domain of the RNAP that contacts σ (Bacteria) and TFB (Archaea) general transcription factors. The SBHM also contacts the general elongation factors NusG (Bacteria) and Spt5/Spt4 (Archaea). Because the SBHM is missing in RNA template−dependent RNAPs of the 2−DPBB type, the SBHM

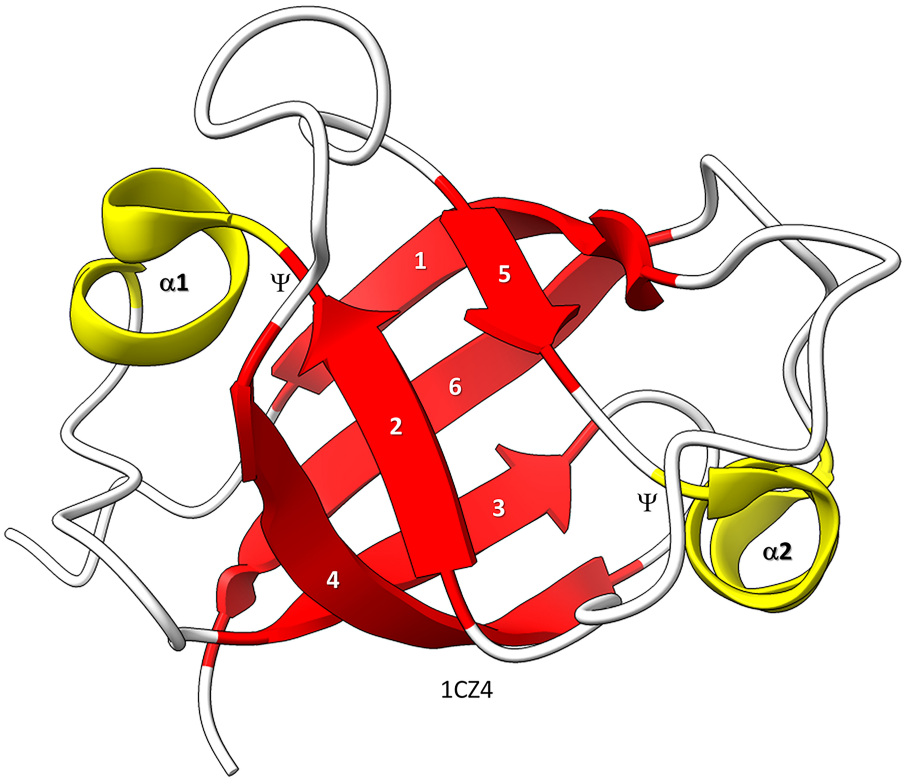


FIGURE 1 | Bacterial VAT (VCP−like ATPase) includes a simple DPBB. ChimeraX was used for molecular graphics (Goddard et al., 2018; Pettersen et al., 2021). The structure is PDB 1CZ4 (Coles et al., 1999). β−sheets are red; α−helices are yellow. Ψ indicates the Ψ pattern of crossing peptide chains.

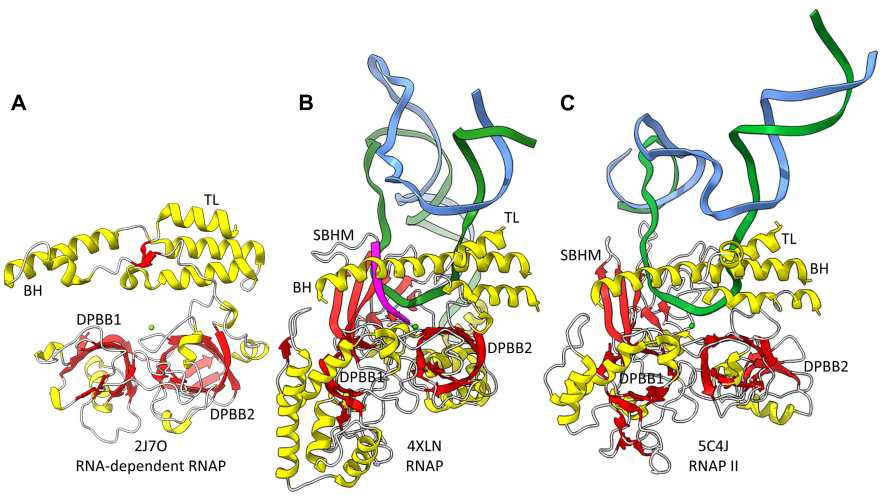


FIGURE 2 | The catalytic core of 2-DPBB type RNAPs. **(A)** A RNA template-dependent RNAP from *Neurospora crassa* (PDB 2J7O) (Salgado et al., 2006). **(B)** A bacterial multi-subunit RNAP (PDB 4XLN) (Bae et al., 2015). **(C)** A human multi-subunit RNAP (PDB 5C4J) (Barnes et al., 2015). α-helices are yellow; β-sheets are red; Mg is green; RNA is magenta; template DNA is green; non-template DNA is blue. BH indicates the bridge helix. TL indicates the trigger loop. The active site is identified by the Mg (Mg1) and the 3′-end of the RNA **(B,C)**.

is considered to be a feature for the transcription of DNA templates (Iyer and Aravind, 2012). Because the SBHM interacts with initiation factors, the SBHM is considered to be evolved to facilitate initiation from DNA templates. A large mostly α−helical insert is found between DPBB1 β5 and β6, after α2. This insert is only partially homologous in archaeal and bacterial RNAPs and appears to make domain-specific contacts to RNAP rather than contacts to transcription factors. In some structures,

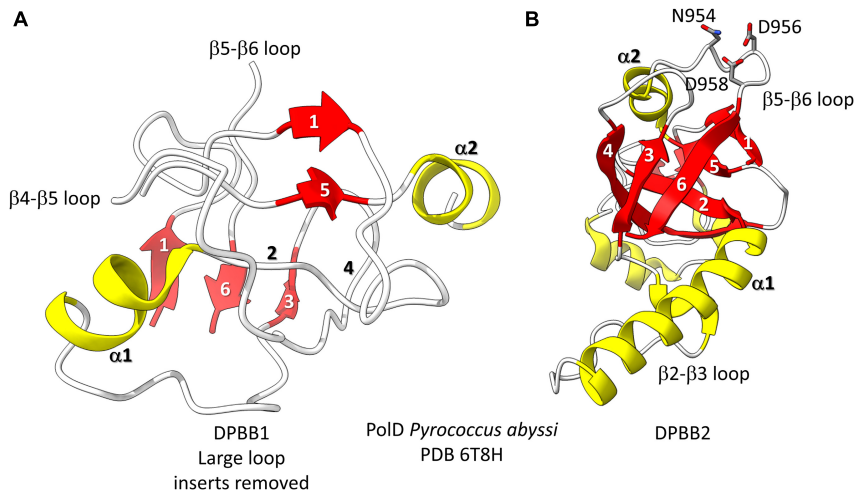


FIGURE 3 | The two DPBBs of a DNA template-dependent DNAP (archaeal PolD) (PDB 6T8H) (Madru et al., 2020). Colors are as in **Figures 1**, **2**. **(A)** DPBB1 is somewhat disordered in the structure, so not all β-sheets were scored as such by ChimeraX. In **(B)** DPBB2, N954, D956 and D958 may hold the active site Mg (missing in the structure) (Zatopek et al., 2020).

DPBB1 is somewhat disordered in 2−DPBB DNA template-dependent RNAPs, making some of the β−sheets difficult to discern. The β′-subunit DPBB2 (referring to bacterial RNAPs) has a largely α−helical insert between β2 and β3 (distinct from the SBHM that includes long β−sheets). In Archaea, the insert between DPBB2 β2 and β3 is referred to as a RAGNYA domain that includes β−sheets and α−helices (Balaji and Aravind, 2007; Iyer and Aravind, 2012). The archaeal and bacterial DPBB2 β2−β3 inserts are very different in sequence and make domain-specific contacts to TFB and σ for initiation.

Found in many Archaea, PolD are DNA template-dependent DNAPs of the 2−DPBB form engaged in genomic replication (Raia et al., 2019; Sauguet, 2019; Koonin et al., 2020; Madru et al., 2020). In these enzymes DPBB1 includes two large inserts, one between β4 and β5 and one between β5 and β6. In available structures, PolD DPBB1 appears to be somewhat disordered, similarly to DPBB1 (β−subunit of bacterial RNAPs) in some structures of DNA template-dependent RNAPs. The significance of this possible similarity in some structures is not known to us. One idea is that DPBB1 is somewhat more dynamic because it accommodates to the presence and absence of substrate to a larger extent than DPBB2, which holds active site Mg1 more tightly than DPBB1 holds Mg2. We would be interested to know whether dNTP binding tightens the PolD DPBB1 and whether similar changes might occur in multi-subunit RNAPs with NTP binding. In PolD, DPBB2 includes an insert between β1 and β2. The inserts in the DNA template-dependent DNAPs (PolD) discriminate PolD enzymes from multi-subunit RNAPs and RNA template-dependent RNAPs and indicate how these more complex enzymes diverged from RNA template-dependent RNAPs of the 2-DPBB form (Koonin et al., 2020).

The story of evolution of these ancient 2-DPBB-type enzymes cannot now be told with certainty, but we construct a possible narrative. We posit that RNA template-dependent RNAPs may have evolved in an RNA-dominated world prior to LUCA (Iyer and Aravind, 2012; Koonin et al., 2020). These enzymes include no large inserts in their DPBBs, indicating that RNA template-dependent RNAPs probably comprise the most ancient 2-DPBB enzyme form. DNA template-dependent RNAPs (multi-subunit RNAPs) and DNAPs (PolD) appear to have radiated mostly independently from the primitive form, although, multi-subunit RNAPs and PolD may share one or two Zn motifs that are missing from 2-DPBB RNA template-dependent RNAPs (see below). Multi-subunit RNAPs and Pol D, however, have distinct DPBB loop inserts. To our knowledge, comparative sequence analyses of these enzymes provides limited insight into details of their divergence, because sequences among enzyme classes are only weakly conserved (Sauguet, 2019; Madru et al., 2020; Zatopek et al., 2020). Because PolD is ancient, this 2-DPBB type enzyme may be the initial evolved DNA template-dependent DNAP for genomic replication (i.e., at LUCA), and other DNAPs, i.e., PolA, PolB and PolC, may have evolved later (Koonin et al., 2020).

RNA template-dependent RNAPs and multi-subunit RNAPs have a recognizable bridge helix and trigger loop (**Figure 2**), and these features are altered and rearranged in DNA template-dependent DNAPs (PolD) of the 2-DPBB type (see below) (Madru et al., 2020). It appears that 2-DPBB multi-subunit RNAPs from Archaea and Eukaryotes and PolD from Archaea may share a Zn-finger motif that is missing from RNA template-dependent RNAPs and bacterial multi-subunit RNAPs. We posit that Archaea are older than Bacteria and closer to LUCA (Battistuzzi et al., 2004; Lei and Burton, 2020; Long et al., 2020), but also see Forterre (2015), Da Cunha et al. (2017, 2018), Castelle and Banfield (2018), Eme et al. (2018). We, therefore, posit that this Zn-finger was lost in bacterial multi-subunit RNAPs, which appear to be a simplified form compared to archaeal multi-subunit RNAPs. We posit that bacterial RNAPs were driven to diverge from archaeal RNAPs primarily because bacterial RNAPs co-evolved with bacterial σ factors.

RNAP Catalytic Subunits (A Guided Tour)

Our view is that Archaea are older than Bacteria, and, therefore, Archaea are closer to LUCA (Battistuzzi et al., 2004; Marin et al., 2017; Lei and Burton, 2020; Long et al., 2020). For other views, see Forterre (2015), Da Cunha et al. (2017, 2018), Castelle and Banfield (2018), Eme et al. (2018). Because of horizontal gene transfer, some phylogenetic analyses may be misleading in determining the deep branching of prokaryotic domains. We believe Bacteria were derived from Archaea. Our opinions are based on ancient evolution studies of transcription systems, tRNA, aminoacyl-tRNA synthetases, ribosomes and the genetic code. In every comparison we have made, Archaea appear to be the more ancient lineage, and Bacteria appear to be more innovated and more derived evolutionarily from root sequences. Therefore, to describe the multi-subunit RNAP catalytic subunits, we use an archaeal RNAP as the example. The RNAP we selected is from *Saccharolobus shibatae* (PDB 2WB1) (Korkhin et al., 2009). The catalytic subunits include 2WB1_A and 2WB1_C (_A and _C indicates the chain designation), which correspond to the β′ subunit of bacterial RNAP, a subunit that is split in some Archaea. 2WB1_B corresponds to the β subunit of bacterial RNAPs. We compare similar motifs in DNAP PolD to emphasize early evolution of RNAPs.

Figure 4 shows the Rpo1N (2WB1_A; A′) and Rpo1C (2WB1_C; A″) chains. We describe some recognizable protein motifs, reading from the N-terminus of the 2WB1_A chain through the 2WB1_C chain. Zn1 is very close to the 2WB1_A N-terminus. Evolutionarily-related motifs in PolD are indicated below the blue bar. Zn1 in 2WB1_A may correspond to archaeal DNAP PolD Zn2, based on its position in the structure and its distance from a Zn motif in chain 2WB1_B (Madru et al., 2020). The N-terminal β-sheet of the β-hairpin is next, followed by 2WB1_A Zn2, which is missing in bacterial RNAP. Next is the C-terminal β-hairpin. From D234 to L302 is a helix-loop-helix motif that connects the AT-like hooks (Iyer and Aravind, 2012). The AT-like hook loop contacts single-stranded DNA in the RNAP open complex and elongation complex. Next is the DPBB2 barrel. Between DPBB2 β2 and β3 after α1 is the RAGNYA insert. In Bacteria, a DPBB2 β2-β3 insert after α1 shows no detectable homology and is primarily α-helical (see below). DPBB2 holds Mg1 within the loop between DPBB2 β5 and β6 (NADFDGD). The "funnel" is located in the primary sequence between the DPBB2 and bridge helix. In the open transcription complex or elongation complex, the DNA template bends by about 90° and DNA strands separate over the bridge helix. DNA PolD has a similar DPBB2 and, also, modified structures that are probably genetically related to the bridge helix and trigger loop, although these features in PolD appear to be rearranged and repurposed (see below).

The *Saccharolobus shibatae* RNAP is separated into two genes relative to the bacterial RNAP β′ subunit, and the subunit separation is between the bridge helix and the trigger loop. The trigger loop is near the archaeal Rpo1C subunit (2WB1_C) N-terminus. The RNAP trigger loop appears to correlate with

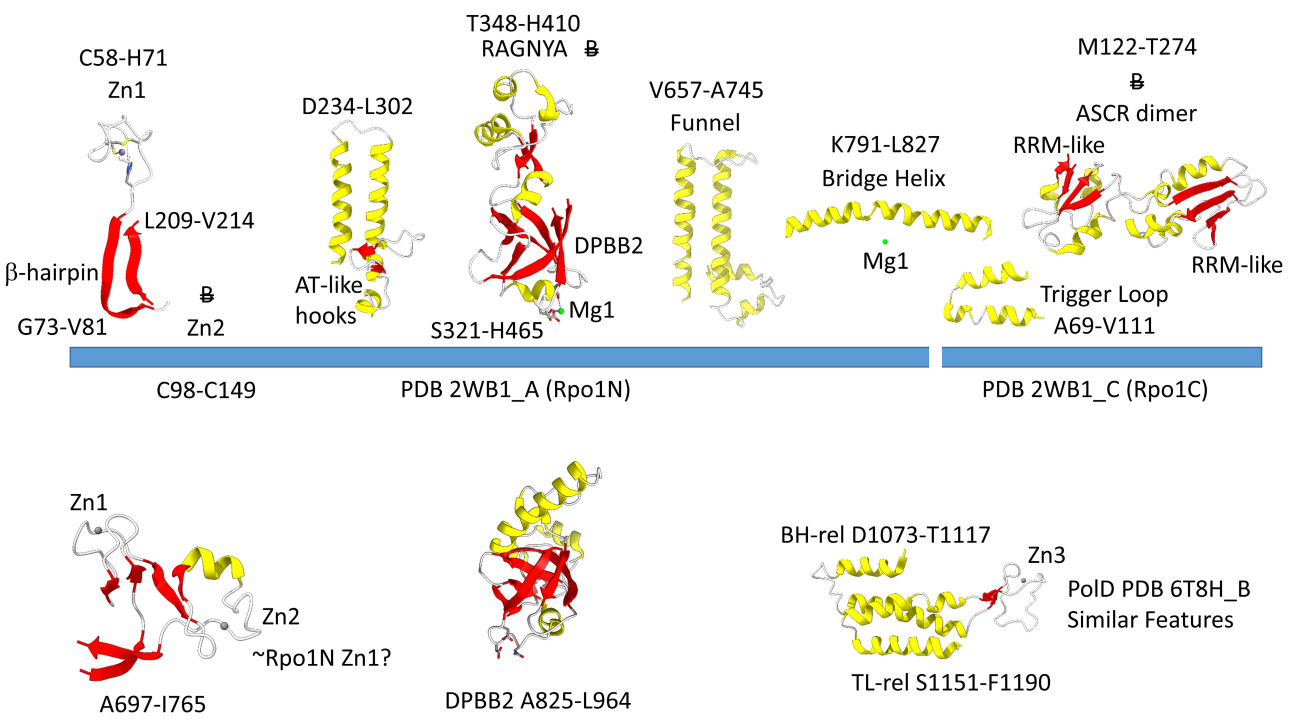


FIGURE 4 | Some recognizable motifs that characterize the RpoA′ and RpoA″ subunits of archaeal RNAP, corresponding to the β′ subunit of bacterial RNAP (Korkhin et al., 2009). "B" with a double strike through indicates a motif in archaeal RNAP that is not identified in bacterial RNAP. Similar motifs in DNAP PolD are shown below the blue bar.

the PolD "clamp" structure (PDB 6T8H_B; S1151-F1190) (Madru et al., 2020). Near the C-terminus of archaeal RNAP Rpo1C, the ASCR dimer is located, with two RRM-like features (RRM for RNA-recognition motif) (Iyer and Aravind, 2012). The ASCR dimer motif is missing in bacterial RNAP and may have been lost by deletion.

In **Figure 5**, a comparison is shown of bacterial RNAP DPBB2, the bridge helix and the trigger loop (**Figure 5A**) and related features in DNAP PolD (**Figure 5B**). In **Figure 5A**, an α-helical domain separates DPBB2 β2 and β3. The α-helical loop insert corresponds to and may have replaced the RAGNYA region in archaeal RNAP (**Figure 4**). The bacterial RNAP β′ subunit includes a Zn motif separating the bridge helix and the trigger loop that is missing in Archaea (**Figure 5A**). PolD also has a Zn motif (Zn3) separating its bridge helix-related and trigger loop-related features, although we do not think these Zn motifs in bacterial RNAP and PolD are related by homology. Rather these Zn motifs may be the result of convergent evolution. In bacterial RNAP, the trigger loop is closer to the active site than the bridge helix and closes over the NTP substrate to expel water from the active site and tighten the substrate for addition to the RNA chain (Vassylyev et al., 2007b). In the image in **Figure 5A**, the trigger loop is in the closed and catalytic conformation. In PolD, the trigger loop-related feature is further from the active site than the bridge helix-related feature. In PolD, the bundle of C-terminal α-helices (bridge helix-related and trigger loop-related features) bind DNA and, also, the proofreading PolD subunit (DP1; the 2-DPBBs are part of the DP2 subunit) (**Figure 6**). The DP1 subunit includes an exonuclease domain. Loops from the bridge helix-related and trigger loop-related PolD features also contact the sliding clamp that maintains PolD processivity (Madru et al., 2020). It appears, therefore, that, although bridge helix- and trigger loop-related features in PolD and RNAPs may be related by evolution, they fulfill different roles.

The archaeal RNAP Rpo2 subunit corresponds to the β-subunit in bacterial RNAP. Features of the Rpo2 RNAP subunit (PDB 2WB1_B; B) are shown in **Figure 7**. There is a 2-lobed N-terminal domain extending from position 1–722. The DPBB1 extends from G723 to K995. There are two notable inserts in DPBB1 loops. Between β2 and β3, just after α1, a SBHM is inserted (Iyer and Aravind, 2012). The SBHM is characterized by long β-sheets. In archaeal RNAP, the SBHM is referred to as the "wall" domain, which interacts with the general transcription factor TFB. In bacterial RNAP, the SBHM has been referred to as the "flap" domain, which interacts with the bacterial σ factor. Between β5 and β6, just after α2, an α-helical segment is inserted (~N914-R985). At the C-terminus of the Rpo2 chain, a Zn finger is located in archaeal RNAPs but missing in bacterial RNAPs. Although the sequences are different, this Zn finger may correspond to Zn1 in archaeal DNAP PolD (Madru et al., 2020). As in PolD, the Rpo2 Zn finger and the Rpo1N Zn1 are close in space in archaeal RNAP, similar to PolD Zn1 and Zn2.

The description of the catalytic subunits of multi-subunit RNAPs here is incomplete. The intention is to provide some visible and conceptual guide posts for researchers as they begin to probe and familiarize themselves with RNAP structures. Also, we emphasize features that appear most important for interactions between general transcription factors and the RNAP catalytic center (see below). A more detailed description of RNAP evolution and domains is provided by Iyer and Aravind (2012). Reviews of the subunit structures of multi-subunit RNAPs are also published elsewhere (Jun et al., 2011; Osman and Cramer, 2020).

2-Mg Mechanism of Transcription by Multi-Subunit RNAPs

We have described the basic catalytic core of multi-subunit RNAPs: 2-DPBBs, a bridge helix and a trigger loop (**Figures 2B,C**). These enzymes utilize a 2-Mg mechanism for transcription (**Figure 8**; Vassylyev et al., 2007b). The 2-Mg (Mg1 and Mg2) are held by acidic groups (E and D) on loops of the 2-DPBBs. DPBB1 includes 685-ED-686 (*Thermus thermophilus*

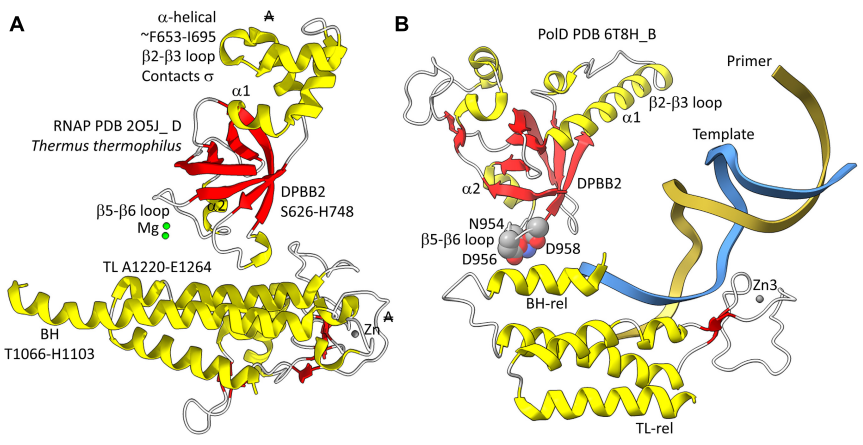


FIGURE 5 | Similarities between the DPBB2, bridge helix and trigger loop of bacterial RNAP and related motifs in DNAP PolD. **(A)** Bacterial RNAP features. **(B)** Related PolD features. The similarly placed Zn motifs are not thought to be homologous. "A" with a double strike through indicates that a feature of bacterial RNAP is not present in archaeal RNAP.

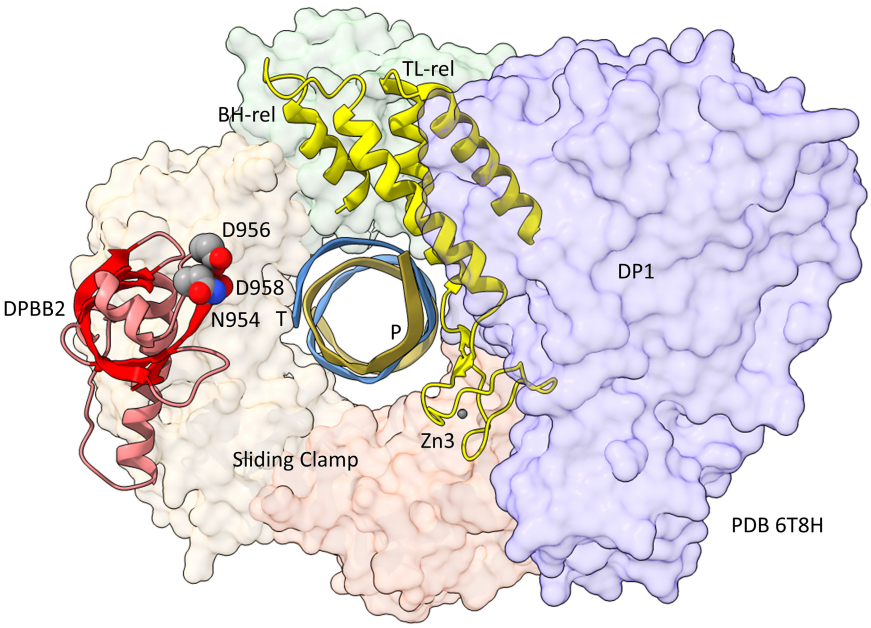


FIGURE 6 | Repurposing of the bridge helix-related (BH-rel) and trigger loop-related (TL-rel) motifs in PolD. The DPBB2 (light red with red β-sheets) and BH-rel, Zn3 and TL-rel region (yellow) is shown for the DP2 2-DPBB subunit. (T) template DNA (blue); (P) primer DNA (gold). The sliding clamp trimer is shown (green, beige and orange). The DP1 subunit is blue. Active site residues that hold Mg1 are indicated in space-filling representation.

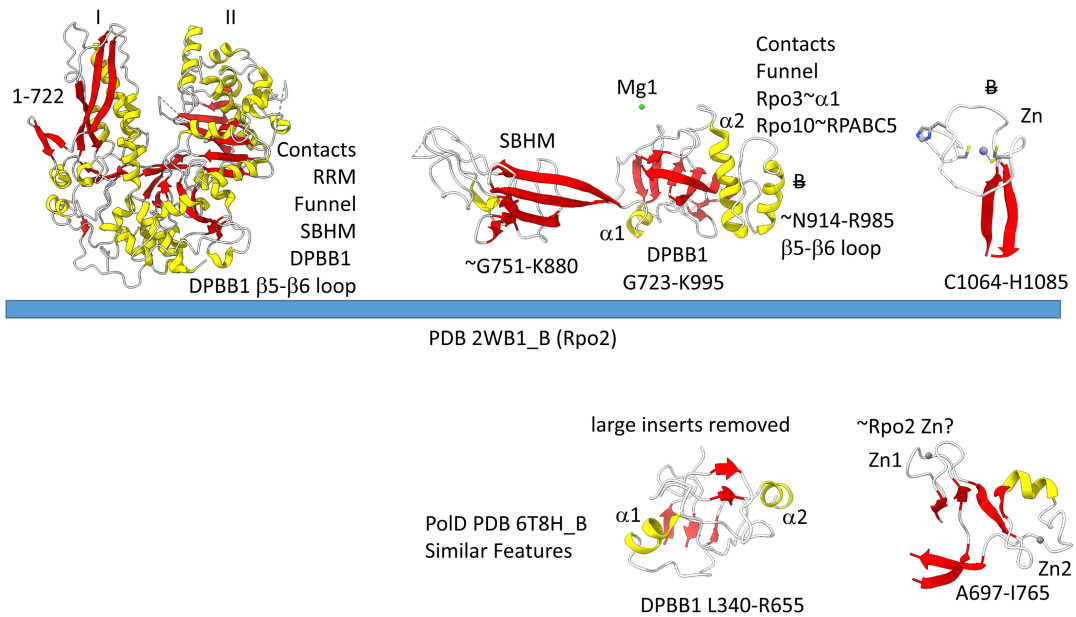


FIGURE 7 | Some recognizable motifs in the Rpo2 subunit of archaeal RNAP (corresponding to the β subunit of bacterial RNAP) (Korkhin et al., 2009). Colors and abbreviations are as in **Figure 4**. Related motifs in DNAP PolD are indicated beneath the blue bar. "B" with a double strike through indicates a feature in Archaea that is missing or very different in Bacteria.

RNAP numbering) located on the DPBB1 loop between β4 and β5. D686 appears to interact with Mg2 during phosphodiester bond formation. Mg2 is loosely held in the RNAP structure. DPBB2 includes the highly conserved sequence 737-NADFDGD-743 within the loop between β5 and β6. D739, D741 and D743 strongly hold Mg1. It is thought that Mg1 remains bound to RNAP, but Mg2 may exchange with each NTP addition. Mg2 normally enters the RNAP bound to the NTP as NTP-Mg. The NADFDGD motif in multi-subunit RNAPs corresponds to 954-NCDGDED-961 in archaeal *Pyrococcus abyssi* DNAP

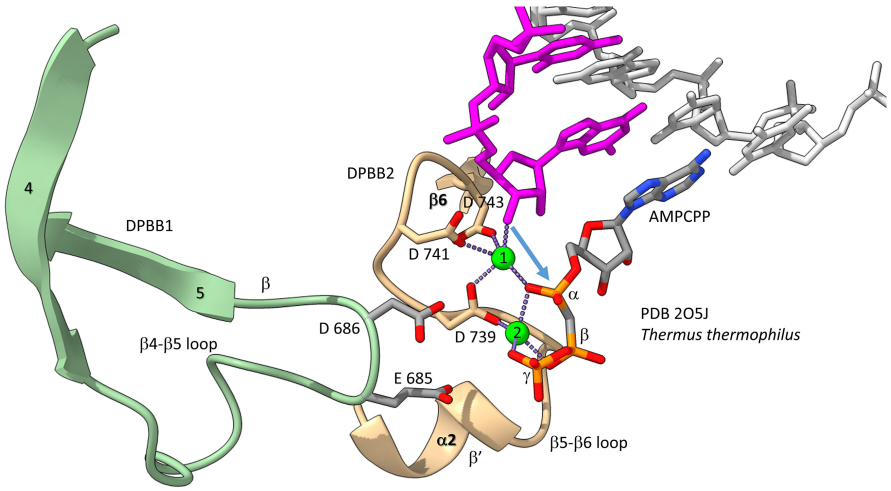


FIGURE 8 | The two Mg mechanism for transcription by RNAP. The structure (PDB 205J) is from *Thermus thermophilus* (Vassylyev et al., 2007b). Mg1 and Mg2 (green spheres) are labeled. The RNA chain is magenta. The β′ subunit is beige. The β subunit is lime. Some active site residues are labeled. AMPCPP (a non-hydrolyzable substrate) is in the substrate site.

PolD (Madru et al., 2020), although, in PolD, the active site Mg1 is held by N954, D956 and D958, so the Mg1-contacting residues are slightly shifted in PolD (Zatopek et al., 2020). In *Neurospora crassa* RNA template-dependent RNAP, Mg1 is held by 1005-GGDYDGD-1011 (Salgado et al., 2006; Qian et al., 2016). Acidic groups retaining Mg1 at the active enzyme site are highly conserved in 2-DPBB type enzymes, although PolD has slightly shifted the set of interacting residues. In the simplest cradle loop barrel enzymes, similar acidic groups can be identified in the same DPBB location (just before β3 and β6), indicating that the initial evolution of DPBBs may have been to chelate Mg (Coles et al., 1999).

Figure 8 shows the 2-Mg mechanism for RNA polymerization. The 3′-O of the RNA chain attacks the α-phosphate of the incoming NTP substrate to add a single NMP unit to the chain and to release pyrophosphate (Vassylyev et al., 2007a,b). Mg1 is held tightly by D739, D741 and D743 within the NADFDGD loop between β5 and β6 of the DPBB2 (β′subunit). Mg2 enters with the NTP substrate and probably interacts with D686 of the DPBB1 (β subunit). Mg2 probably leaves with pyrophosphate.

EVOLUTION OF ARCHAEAL AND BACTERIAL GTFs

Because we posit that Archaea are older than Bacteria, we first consider general transcription factors (GTFs) in Archaea (Jun et al., 2011; Blombach et al., 2015). To recognize a core promoter, Archaea utilize TBP (TATA-box binding protein), TFB (transcription factor B) and TFE (transcription factor E). It appears that Bacteria evolved σ factors from TFB and lost TBP and TFE in evolution. **Figure 9** shows a promoter-TBP-TFB complex from Archaea (Littlefield et al., 1999). **Figure 9A** is a detail of the image in **Figure 9B** to indicate the helix-turn-helix (HTH) motif of the most C-terminal HTH domain. TBP contacts the 8-nt TATA-box. TBP includes a C-terminal repeat sequence that forms a pseudo-dimer of β-sheet folds to align with pseudo-dimeric DNA. TBP occupies the minor groove of the DNA. TFB includes two cyclin-like repeats (CLR) formed as 5-α-helix bundles that bind DNA upstream and downstream of TATA (Lagrange et al., 1998; Renfrow et al., 2004). The last 3-helices of each CLR comprise a typical HTH DNA-binding motif (**Figure 9A**). HTH motifs are comprised of H1-T1-H2-T2-H3 (H for helix; T for turn). Characteristically, H1 braces H2 and H3. H2 is generally a short helix. The N-terminus of H3 penetrates the major groove of DNA and makes most sequence-specific contacts. **Figure 9A** emphasizes the typical DNA contacts of HTH2 of TFB to the BREup (TFB-recognition element upstream of TATA) of the archaeal promoter. **Figure 9B** is a more complex image that includes TBP and CLR1 and CLR2 of TFB. H3 of CLR1 and CLR2 interacts with the major groove of DNA at BREdown and BREup. TFE is another GTF in Archaea that does not make extensive sequence-specific contacts to DNA (Blombach et al., 2015). In Bacteria, TBP and TFE appear to have been lost in evolution. The TFB C-terminal CLR/HTH repeats appear to have been duplicated and modified in evolution to generate bacterial σ factors.

Bacterial σ factors are homologs of TFB (Iyer and Aravind, 2012; Burton, 2014; Burton and Burton, 2014; Burton et al., 2016; **Figure 10**). This idea was first postulated by Aravind and co-workers, based on the similarities of HTH units. Similarly to TFB, σ factors were initially strings of HTH units. For instance, σA appears to be derived from 4-HTH units (HTH1-4). We posit that σA was derived from duplication of the TFB C-terminus CLR/HTH units. σ54, by contrast, might be derived from 6–7 (or possibly 8) HTH units. σ54 might have resulted from early duplication of σA. The more N-terminal HTH units in both σA and σ54 are more degenerate, and, therefore, less recognizable. Here, we consider the four most C-terminal HTH units, which are in common comparing σA and σ54, and number them 1→4, from the N-terminal end, so HTH4 is the most C-terminal σ HTH unit. TFB, by contrast, includes two HTH units, numbered HTH1

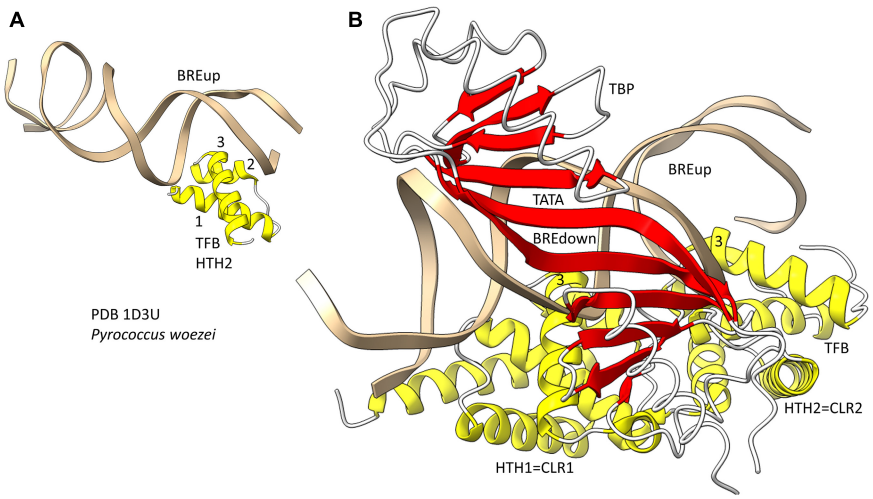


FIGURE 9 | The promoter-TBP-TFB complex in Archaea. **(A)** A detail of the image in panel **(B)**, showing that TFB HTH units are typical and make typical contacts to the major groove of DNA. **(B)** The promoter-TBP-TFB complex. HTH1 and HTH2 are the last 3 helices of 5-helix cyclin-like repeats (CLR1 and CLR2).

and HTH2, C-terminal to an N-terminal Zn finger domain. So, HTH4 in σA and σ54 corresponds to HTH2 in TFB. HTH3 in σA and σ54 corresponds to HTH1 in TFB. The concept of σ and TFB homology is necessary to consider archaeal and bacterial divergence and the evolution and divergence of promoters.

To further support the homology of σ factors and TFB, we prepared overlays of initiation complexes from bacterial and human systems (**Figure 10**). Human TFIIB is a close homolog of archaeal TFB. RNAP and other GTFs were removed from the image to attempt simplification. **Figure 10** is an overlay of three structures: (1) a human preinitiation complex (PDB 5IY7) (He et al., 2016), (2) a bacterial σA early initiation complex, with a short RNA (PDB 5I2D) (Feng et al., 2016), and (3) a bacterial σ54 holoenzyme (PDB 5BYH) (Yang et al., 2015). Because the image is somewhat busy, two views and a detail view are shown. TFIIB HTH1, σA HTH3 and σ54 HTH3 co-localize at the upstream end of the transcription bubble. TFIIB HTH2 and σ54 HTH4 partly overlay in the upstream DNA region. By contrast, σA HTH4 follows the diverging trajectory of the upstream DNA to which HTH4 binds at the −35 promoter region (detail image). Notice that σA HTH4 makes typical HTH contacts to the −35 region of the bacterial promoter (**Figure 10**; detail image), just as TFB makes typical HTH contacts to BREup and BREdown (**Figure 9**). We conclude from the overlay of these structures that HTH4 and HTH3 of bacterial σ factors correspond to HTH2 and HTH1 of human TFIIB (Iyer and Aravind, 2012; Burton, 2014; Burton and Burton, 2014; Burton et al., 2016).

Promoter-Specific Regulatory HTH Factors

We speculate that GTFs TBP and TFB may have been present at LUCA as part of the earliest mechanisms for opening and managing DNA templates. In Archaea and Bacteria, many promoter-specific transcription factors are dimeric HTH or winged-HTH (HTH factors with β-sheet "wings") factors (Aravind et al., 2005; Iyer and Aravind, 2012). These promoter-specific HTH factors may somehow have been derived by simplification of the CLR domains of TFB (5-α-helix bundles), followed generally by homodimerization. We note that bacterial σ factor HTH units are simplified from the TFB 5-helix CLR formats, from which σ factors appear to be derived (Iyer and Aravind, 2012; Burton and Burton, 2014). The HTH motif was, therefore, a core founding feature in Archaea and Bacteria of early evolution of both transcriptional GTFs (TFB and σ) and regulatory (HTH and winged-HTH factors) mechanisms.

Evolution of Archaeal and Bacterial Promoters

A model for the divergence of archaeal and bacterial promoters is described (**Figure 11**). Because of the long passage of time, we are not certain that all aspects of a core promoter model can precisely be stated. The model is presented in order to provide a simple possible narrative that may stimulate more sophisticated bioinformatics approaches to this problem than we were able to do. Also, the model is based partly on our opinion that Archaea is most similar to LUCA, that Bacteria are more derived and that Bacteria evolved from Archaea (Battistuzzi et al., 2004; Marin et al., 2017; Lei and Burton, 2020; Long et al., 2020). There are reasons to consider this idea. A recent paper indicated that LUCA was most similar to Archaea, and that Bacteria were derived from Archaea. tRNAs and tRNAomes (all the tRNAs for an organism) are simpler and more similar to the primordial tRNA sequence in Archaea (Pak et al., 2018; Kim et al., 2019; Lei and Burton, 2020). Also, aminoacyl-tRNA synthetases and the genetic code are simpler to model in Archaea than in Bacteria, indicating that Archaea are more similar to LUCA than Bacteria.

Figure 11 compares a bacterial σA promoter and its GTF contacts and an archaeal promoter and its GTF contacts. The bacterial promoter shows sequences characteristic of a strong promoter with multiple contacts to different regions of σA. Bacteria lack TBP and TFE, which we posit may have been

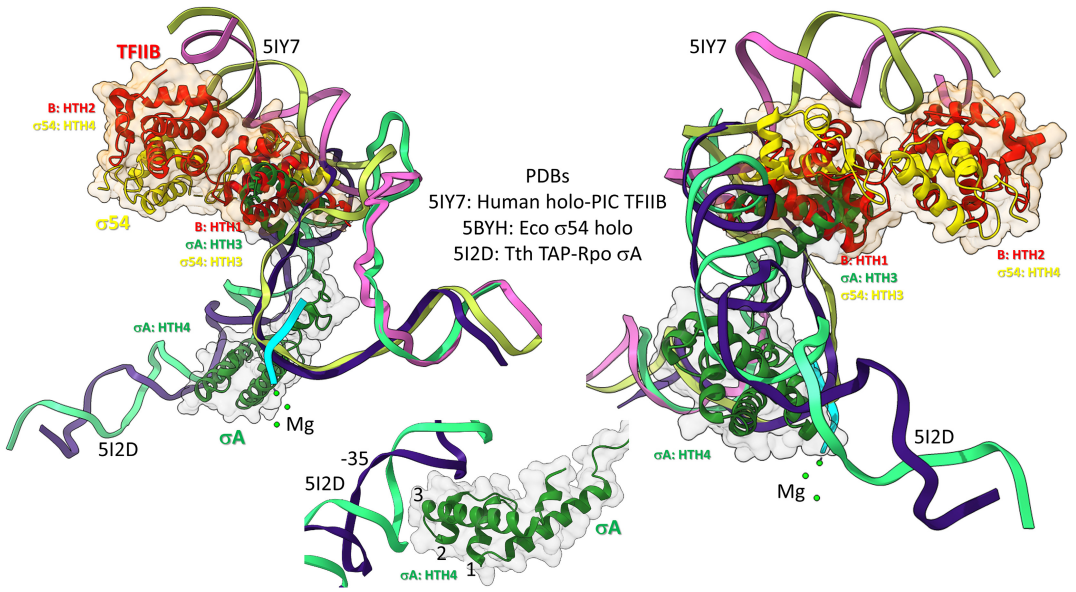


FIGURE 10 | Bacterial σ factors and human TFIIB are homologs. Two views and one detail are shown. Two initiation complexes (human and *Thermus thermophilus*) and a σ54 holoenzyme structure (*Escherichia coli*) were overlaid. σA HTH3, σ54 HTH3, and TFIIB (B) HTH1 overlay at the upstream edge of the transcription bubble. σ54 HTH4 and TFIIB HTH2 partly overlay upstream (i.e., BREup). The detail is of σA HTH4 showing characteristic HTH contacts to the promoter –35 region. RNA is cyan. Mg is green. Upstream DNA strands are labeled: 5IY7: (pink) non-template; (yellow) template; and 5I2D: (green) non-template; (blue) template.

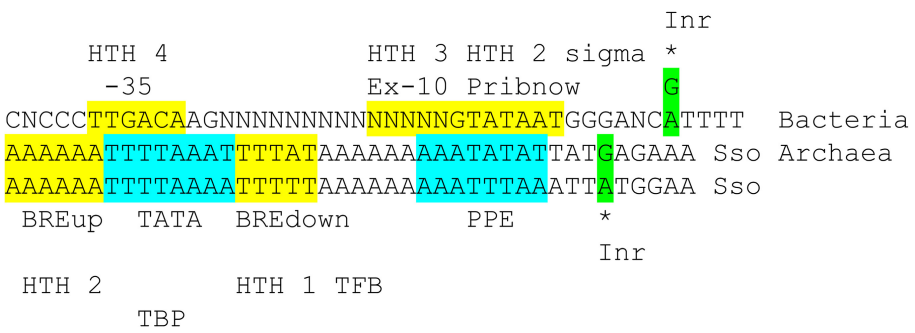


FIGURE 11 | Comparison of bacterial σA promoters and archaeal promoters from *Sulfolobus solfataricus* (Sso; an ancient Archaea). See the text for details. Inr for initiator element.

lost during bacterial divergence. Bacteria include RNase HIII that includes a TBP fold (Brindefalk et al., 2013), however, possibly indicating that Bacteria had TBP as a transcription factor from Archaea and then lost TBP in evolution, as we propose. According to the structural overlay (**Figure 10**), bacterial σA HTH4 and HTH3 correspond to archaeal TFB HTH2 and HTH1 (Iyer and Aravind, 2012; Burton, 2014; Burton and Burton, 2014; Burton et al., 2016). Bacterial σA HTH4 contacts the −35 region of promoters [i.e., (-34)-TTGACA-(-29)]. Archaeal TFB HTH2 contacts the BREup (TFB-recognition element upstream of the TATA-box). TBP binds the 8-nt TATA-box [i.e., (-30)-TTTTAAAA-(-23) in *Sulfolobus solfataricus*] (Ao et al., 2013), but TBP is missing in Bacteria. Bacterial σA HTH3 partly contacts the Extended −10 sequence in double-stranded DNA, found in some promoters, and then resides on double-stranded DNA at the upstream edge of the transcription bubble, as the promoter opens (**Figure 10**). Archaeal TFB HTH1 contacts the BREdown (TFB-recognition element downstream of the TATA-box) (an A/T-rich sequence downstream from TATA in *Sulfolobus solfataricus*) (**Figure 9B**). After promoter opening, TFB HTH1 occupies double-stranded DNA just upstream of the transcription bubble (**Figure 10**).

The Promoter-Proximal Element (PPE) is an A/T-rich sequence in *Sulfolobus solfataricus* promoters upstream of the transcription start [i.e., ∼(−11)-AATATTAA-(−4)] (Ao et al., 2013). To us, the PPE resembles a TATA-box and may be derived from one. The PPE appears to be positioned similarly to the bacterial Pribnow box [i.e., (−12)-TATAAT-(−7)] and is similar in sequence. We, therefore, posit that the Pribnow box of bacterial promoters may be derived from an archaeal PPE sequence. Notably, the Pribnow box is recognized by σA HTH2, which is a modified HTH with interesting characteristics.

The σA HTH2 opens the bacterial promoter by flipping bases. A(−11) is first flipped out followed by T(−7), leading to promoter opening (Feklistov and Darst, 2011; Feklistov et al., 2014; Boyaci et al., 2019).

Archaeal promoters typically have an initiator sequence surrounding + 1, the transcription start (Ao et al., 2013). Many promoters have (−1)-TATG-(+3). In this case, no 5′-untranslated sequence may be present in the mRNA, which may initiate translation at (+1)-AUG-(+3). (−1)-TGAG-(+3) is also common. In this case, translation generally initiates at a downstream AUG. The initiator element is thought to be recognized directly by RNAP. Bacteria also have an initiator

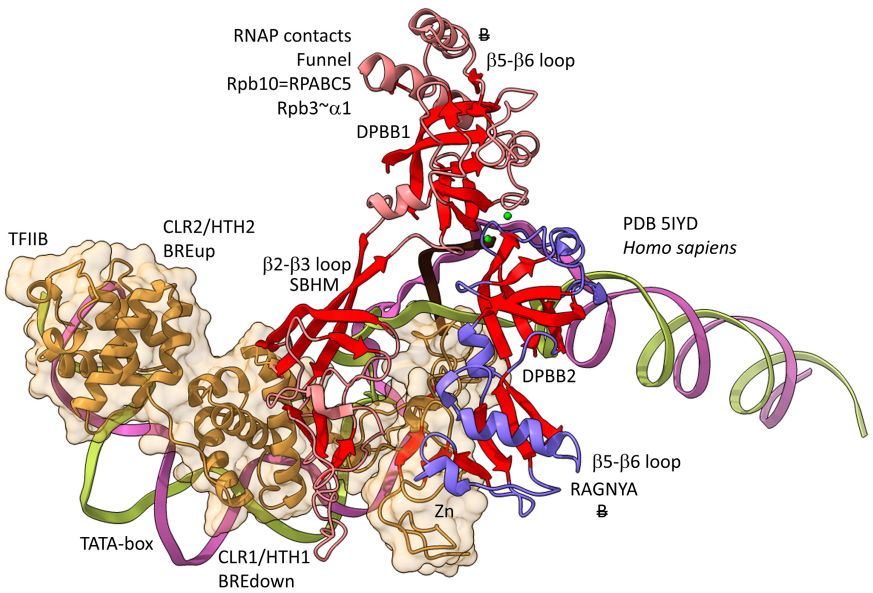


FIGURE 12 | Archaea/Eukaryote-specific contacts of TFB/TFIIB with DPBB insert loops. β-sheets are red. Other features of Rpb1 are blue and Rpb2 are light red. TFIIB is orange with transparent space-filling representation. "B" with double strike through indicates a contact specific to Archaea and not found or very different in Bacteria.

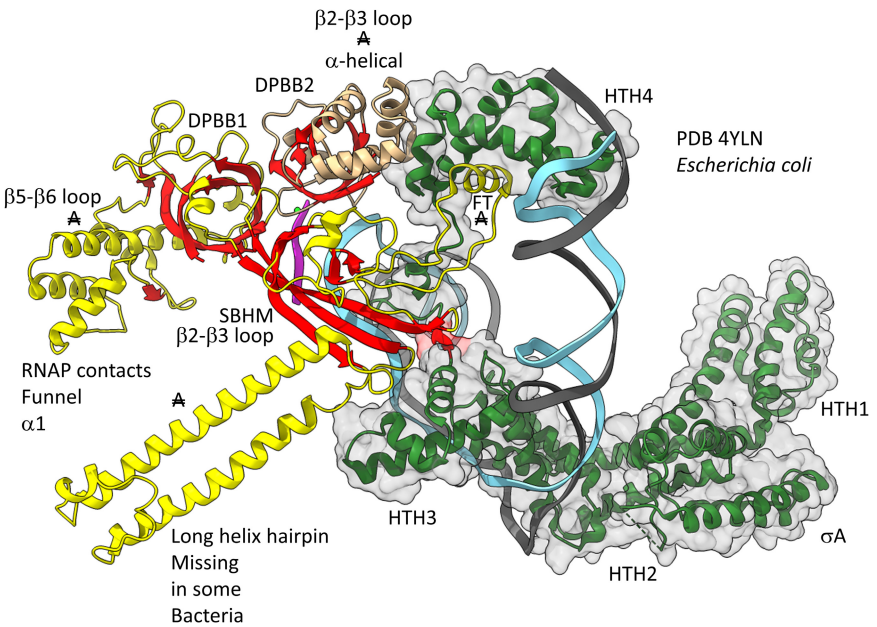


FIGURE 13 | Bacteria-specific contacts of σA with DPBB insert loops. β-sheets are red. Other β′ features are beige, and β features are yellow. σA is green with transparent space-filling representation. FT for flap tip helix. "A" with double strike through indicates a feature found in Bacteria but very different or not identified in Archaea.

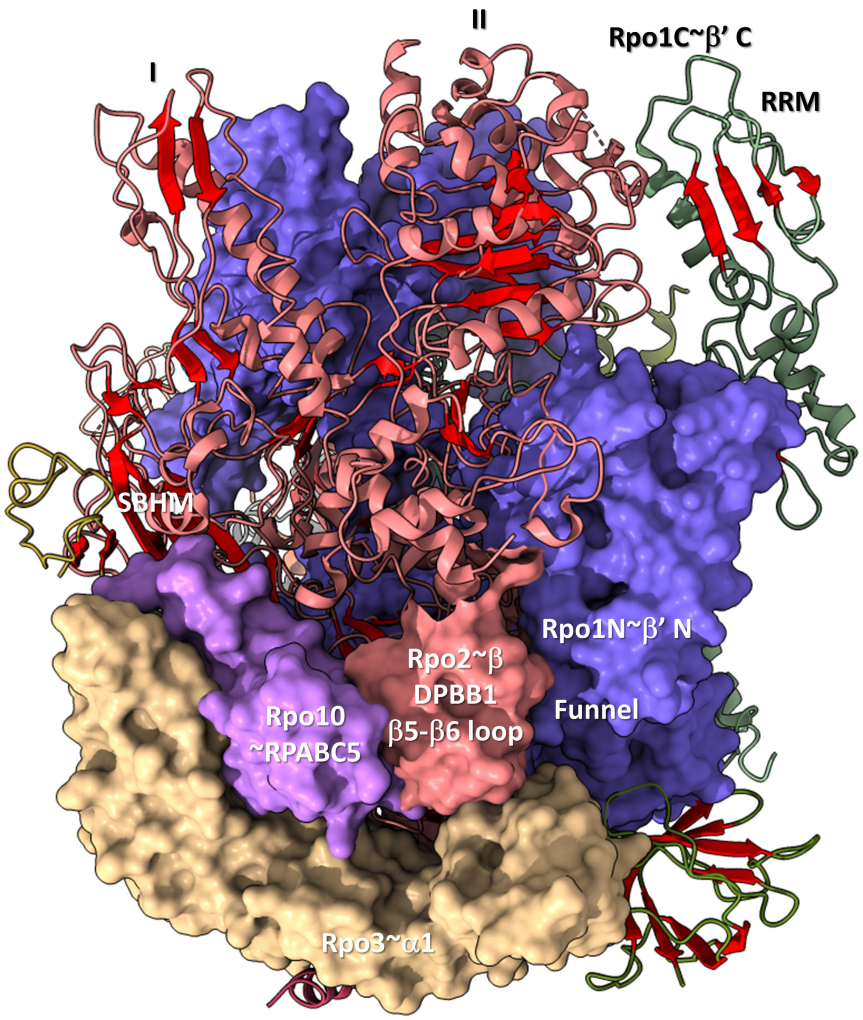


FIGURE 14 | The DPBB1 β5-β6 loop (space-filling representation) contacts RNAP. In Archaea, the DPBB1 β5-β6 loop contacts the Rpo1N (homolog of β′ in Bacteria) funnel, the Rpo2 (homolog of β in Bacteria) N-terminal domain (lobe II), Rpo3 (homolog of α1 in Bacteria) and Rpo10 (homolog of RPABC5 in Eukarya). The SBHM contacts lobe I of the N-terminal Rpo2 domain and Rpo3.

sequence (Cassiano and Silva-Rocha, 2020). Both Archaea and Bacteria utilize ribosome attachment sequences (i.e., AGGA) on some mRNAs with a corresponding interaction sequence near the 3′-end of 16S rRNA (i.e., UCCU).

Interactions of DPBB Loops With GTFs

One hypothesis might be that multi-subunit RNAP DPBB loops that include inserts contact GTFs in a domain-specific fashion. The idea underlying this hypothesis is that DPBBs form the catalytic center and hold the active site Mg1 and Mg2. The RNAP active site is deeply sequestered within the RNAP core, limiting access to the catalytic center. Inserts in the DPBB loops might allow GTFs binding closer to the RNAP periphery to communicate with catalytic functions. Because archaeal GTFs and TFB are so different from bacterial σ factors, TFB and σ might be expected to interact with DPBB loops with distinct, domain-specific inserts.

Figures 12, **13** show domain-specific functional contacts of DPBB loops with GTFs. **Figure 12** shows a simplified view of a human preinitiation complex (PDB 5IYD) (He et al., 2016). Most of the factors in the structure have been removed to simplify the image. The human DPBB1 SBHM (β2-β3 insert) contacts TFIIB HTH1/CLR1 located at the upstream edge of the transcription bubble. Interestingly, the human DPBB2 RAGNYA β2-β3 insert, specific for Archaea and Eukaryotes, contacts the N-terminal Zn finger of TFIIB. In **Figure 13**, a detail of the *Escherichia coli* RNAP initiation complex is shown (PDB 4YLN) (Zuo and Steitz, 2015). Bacterial σA HTH3, at the upstream end of the transcription bubble, contacts the SBHM. Thus, homologous

GTFs in Archaea (TFB) and Bacteria (σA) make domain-specific contacts to their domain-specific SBHMs. In Bacteria, the flap tip helix is an extension of the SBHM that contacts the σA HTH4, bound to the −35 promoter region. Interestingly, the *Escherichia coli* RNAP SBHM includes a long helix hairpin motif as an insert, missing in Archaea and many Bacteria (i.e., missing in *Thermus thermophilus*, an ancient Bacteria). The long helix hairpin insert contacts σA HTH3 in the initiating complex. The DPBB2 β2-β3 insert in *Escherichia coli* RNAP is an α-helical motif that substitutes for the very different RAGNYA insert in Archaea, which contacts the N-terminal Zn motif in TFIIB (**Figure 12**). The corresponding DPBB2 β2-β3 α-helical insert in Bacteria makes domain-specific contacts to αA HTH4, bound at the −35 promoter region (**Figure 13**).

The DPBB1 β5-β6 insert shows some homology in Archaea and Bacteria but, also, significant domain-specific character, so we attempted to identify a GTF that might contact this region. We were unsuccessful. So far as we can discern, the β5-β6 DPBB1 inserts in Archaea and Bacteria make domain-specific contacts to other regions of RNAP (**Figure 14**). In Archaea, the β5-β6 DPBB1 insert contacts: (1) the Rpo1N funnel (A′; homolog of β′ in Bacteria); (2) Rpo10 (N; homolog of RPABC5 in Eukarya); and (3) Rpo3 (C; homolog of α1 in Bacteria). In Bacteria, the β5-β6 DPBB1 insert makes similar domain-specific contacts to RNAP (not shown).

During transcription elongation, TFB and σ factors cycle off RNAP and are replaced by the elongation factor homologs Spt5/Spt4 in Archaea and NusG in Bacteria (Werner, 2012; Blombach et al., 2013; Hartzog and Fu, 2013; Tomar and Artsimovitch, 2013; Yakhnin and Babitzke, 2014; Wang and Artsimovitch, 2020). These elongation factors occupy approximately the same positions on RNAP as HTH2 and HTH3 of bacterial σA (not shown). These elongation factors, therefore, make domain-specific contacts to the SBHM of their DPBB1 (i.e., see PDB 5TBZ) (Liu and Steitz, 2017). Contacts to GTFs are also specific to the initiation and elongation phases of the transcription cycle. For instance, in Bacteria, the flap tip helix contacts σA during initiation (**Figure 13**) but does not contact NusG during elongation.

DIVERGENCE OF ARCHAEA AND BACTERIA

Evolution of life on Earth appears to be a simple outline with overwhelming detail. According to our view, pre-life evolved to LUCA, which we interpret as an ancient Archaea. Archaea diverged to generate Bacteria, which became a more flexible and, in many ways, more successful prokaryotic domain, restricting Archaea somewhat to the margins (i.e., to extremophile environments). Multiple Archaea and Bacteria fused to form Eukaryotes, which have occupied many new niches on Earth (Forterre, 2015; Castelle and Banfield, 2018; Eme et al., 2018). Ancient Archaea, therefore, are very similar to LUCA. Bacteria are more innovated than Archaea and more derived evolutionarily. Because of their mitochondria and complex genomes and development, Eukaryotes have many new capacities lacking in Archaea and Bacteria. We refer to the splitting of the archaeal and bacterial domains as "the great divergence," and we consider this event to be one of the most important advances in evolution of life as we know it on Earth.

There are several defining differences comparing Archaea and Bacteria: i.e., (1) evolution of TFB (Archaea) versus σ factors (Bacteria); (2) utilization of DNAPs PolD and PolB (Archaea) versus PolC (Bacteria) (Koonin et al., 2020), and (3) archaeal versus bacterial membranes (Lane and Martin, 2012; Lane, 2020). Above, we have discussed the divergence of archaeal and bacterial GTFs and promoters in some detail. We consider modifications of bacterial transcription systems to be fundamental and possibly the founding difference in the great divergence of Bacteria from Archaea. For instance, evolution of bacterial σ factors appears to have driven the simplification and divergence of bacterial RNAPs from archaeal ancestors.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

Alva, V., Koretke, K. K., Coles, M., and Lupas, A. N. (2008). Cradle-loop barrels and the concept of metafolds in protein classification by natural descent. *Curr. Opin. Struct Biol.* 18, 358–365. doi: 10.1016/j.sbi.2008.02.006

Ao, X., Li, Y., Wang, F., Feng, M., Lin, Y., Zhao, S., et al. (2013). The Sulfolobus initiator element is an important contributor to promoter strength. *J. Bacteriol.* 195, 5216–5222. doi: 10.1128/jb.00768-13

Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M., and Iyer, L. M. (2005). The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol. Rev.* 29, 231–262. doi: 10.1016/j.femsre.2004.12.008

Bae, B., Feklistov, A., Lass-Napiorkowska, A., Landick, R., and Darst, S. A. (2015). Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *eLife* 4:e08504.

Balaji, S., and Aravind, L. (2007). The RAGNYA fold: a novel fold with multiple topological variants found in functionally diverse nucleic acid, nucleotide and peptide-binding proteins. *Nucleic Acids Res.* 35, 5658–5671. doi: 10.1093/nar/gkm558

Barnes, C. O., Calero, M., Malik, I., Graham, B. W., Spahr, H., Lin, G., et al. (2015). Crystal structure of a transcribing RNA Polymerase II complex reveals a complete transcription bubble. *Mol. Cell* 59, 258–269. doi: 10.1016/j.molcel.2015.06.034

Battistuzzi, F. U., Feijao, A., and Hedges, S. B. (2004). A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol. Biol.* 4:44. doi: 10.1186/1471-2148-4-44

Blombach, F., Daviter, T., Fielden, D., Grohmann, D., Smollett, K., and Werner, F. (2013). Archaeology of RNA polymerase: factor swapping during the transcription cycle. *Biochem. Soc. Trans.* 41, 362–367. doi: 10.1042/bst20120274

Blombach, F., Salvadori, E., Fouqueau, T., Yan, J., Reimann, J., Sheppard, C., et al. (2015). Archaeal TFEalpha/beta is a hybrid of TFIIE and the RNA polymerase III subcomplex hRPC62/39. *eLife* 4:e08378.

Boyaci, H., Chen, J., Jansen, R., Darst, S. A., and Campbell, E. A. (2019). Structures of an RNA polymerase promoter melting intermediate elucidate DNA unwinding. *Nature* 565, 382–385. doi: 10.1038/s41586-018-0840-5

Brindefalk, B., Dessailly, B. H., Yeats, C., Orengo, C., Werner, F., and Poole, A. M. (2013). Evolutionary history of the TBP-domain superfamily. *Nucleic Acids Res.* 41, 2832–2845. doi: 10.1093/nar/gkt045

Burton, S. P., and Burton, Z. F. (2014). The sigma enigma: bacterial sigma factors, archaeal TFB and eukaryotic TFIIB are homologs. *Transcription* 5:e967599. doi: 10.4161/21541264.2014.967599

Burton, Z. F. (2014). The old and new testaments of gene regulation. evolution of multi-subunit rna polymerases and co-evolution of eukaryote complexity with the RNAP II CTD. *Transcription* 5:e28674. doi: 10.4161/trns.28674

Burton, Z. F., Opron, K., Wei, G., and Geiger, J. H. (2016). A model for genesis of transcription systems. *Transcription* 7, 1–13. doi: 10.1080/21541264.2015.1128518

Cassiano, M. H. A., and Silva-Rocha, R. (2020). Benchmarking available bacterial promoter prediction tools: potentialities and limitations. *bioRxiv* [Preprint]. doi: 10.1128/msystems.00439-20

Castelle, C. J., and Banfield, J. F. (2018). Major new microbial groups expand diversity and alter our understanding of the tree of life. *Cell* 172, 1181–1197. doi: 10.1016/j.cell.2018.02.016

Coles, M., Diercks, T., Liermann, J., Groger, A., Rockel, B., Baumeister, W., et al. (1999). The solution structure of VAT-N reveals a 'missing link' in the evolution of complex enzymes from a simple Betaalphabetabeta element. *Curr. Biol.* 9, 1158–1168. doi: 10.1016/s0960-9822(00)80017-2

Coles, M., Djuranovic, S., Soding, J., Frickey, T., Koretke, K., Truffault, V., et al. (2005). AbrB-like transcription factors assume a swapped hairpin fold that is evolutionarily related to double-psi beta barrels. *Structure* 13, 919–928. doi: 10.1016/j.str.2005.03.017

Coles, M., Hulko, M., Djuranovic, S., Truffault, V., Koretke, K., Martin, J., et al. (2006). Common evolutionary origin of swapped-hairpin and double-psi beta barrels. *Structure* 14, 1489–1498. doi: 10.1016/j.str.2006.08.005

Da Cunha, V., Gaia, M., Gadelle, D., Nasir, A., and Forterre, P. (2017). Lokiarchaea are close relatives of *Euryarchaeota*, not bridging the gap between prokaryotes and eukaryotes. *PLoS Genet.* 13:e1006810. doi: 10.1371/journal.pgen.1006810

Da Cunha, V., Gaia, M., Nasir, A., and Forterre, P. (2018). Asgard archaea do not close the debate about the universal tree of life topology. *PLoS Genet.* 14:e1007215. doi: 10.1371/journal.pgen.1007215

Eme, L., Spang, A., Lombard, J., Stairs, C. W., and Ettema, T. J. G. (2018). Archaea and the origin of eukaryotes. *Nat. Rev. Microbiol.* 16:120. doi: 10.1038/nrmicro.2017.154

Feklistov, A., and Darst, S. A. (2011). Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell* 147, 1257–1269. doi: 10.1016/j.cell.2011.10.041

Feklistov, A., Sharon, B. D., Darst, S. A., and Gross, C. A. (2014). Bacterial sigma factors: a historical, structural, and genomic perspective. *Annu. Rev. Microbiol.* 68, 357–376. doi: 10.1146/annurev-micro-092412-155737

Feng, Y., Zhang, Y., and Ebright, R. H. (2016). Structural basis of transcription activation. *Science* 352, 1330–1333. doi: 10.1126/science.aaf4417

Forterre, P. (2015). The universal tree of life: an update. *Front. Microbiol.* 6:717. doi: 10.3389/fmicb.2015.00717

Fouqueau, T., Blombach, F., and Werner, F. (2017). Evolutionary origins of two-barrel RNA polymerases and site-specific transcription initiation. *Annu. Rev. Microbiol.* 71, 331–348. doi: 10.1146/annurev-micro-091014-104145

Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H., et al. (2018). Meeting modern challenges in visualization and analysis. *Protein Sci.* 27, 14–25. doi: 10.1002/pro.3235

Hartzog, G. A., and Fu, J. (2013). The Spt4-Spt5 complex: a multi-faceted regulator of transcription elongation. *Biochim. Biophys. Acta* 1829, 105–115. doi: 10.1016/j.bbagrm.2012.08.007

He, Y., Yan, C., Fang, J., Inouye, C., Tjian, R., Ivanov, I., et al. (2016). Near-atomic resolution visualization of human transcription promoter opening. *Nature* 533, 359–365. doi: 10.1038/nature17970

Iyer, L. M., and Aravind, L. (2012). Insights from the architecture of the bacterial transcription apparatus. *J. Struct. Biol.* 179, 299–319. doi: 10.1016/j.jsb.2011.12.013

Iyer, L. M., Koonin, E. V., and Aravind, L. (2003). Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases. *BMC Struct. Biol.* 3:1. doi: 10.1186/1472-6807-3-1

Jun, S. H., Reichlen, M. J., Tajiri, M., and Murakami, K. S. (2011). Archaeal RNA polymerase and transcription regulation. *Crit. Rev. Biochem. Mol. Biol.* 46, 27–40.

Kim, Y., Opron, K., and Burton, Z. F. (2019). A tRNA- and anticodon-centric view of the evolution of Aminoacyl-tRNA Synthetases, tRNAomes, and the genetic code. *Life* 9:37. doi: 10.3390/life9020037

Koonin, E. V., Krupovic, M., Ishino, S., and Ishino, Y. (2020). The replication machinery of LUCA: common origin of DNA replication and transcription. *BMC Biol.* 18:61. doi: 10.1186/s12915-020-00800-9

Korkhin, Y., Unligil, U. M., Littlefield, O., Nelson, P. J., Stuart, D. I., Sigler, P. B., et al. (2009). Evolution of complex RNA polymerases: the complete archaeal RNA polymerase structure. *PLoS Biol.* 7:e1000102. doi: 10.1371/journal.pbio.1000102

Lagrange, T., Kapanidis, A. N., Tang, H., Reinberg, D., and Ebright, R. H. (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev.* 12, 34–44. doi: 10.1101/gad.12.1.34

Lane, N. (2020). How energy flow shapes cell evolution. *Curr. Biol.* 30, R471–R476.

Lane, N., and Martin, W. F. (2012). The origin of membrane bioenergetics. *Cell* 151, 1406–1416. doi: 10.1016/j.cell.2012.11.050

Lane, W. J., and Darst, S. A. (2010a). Molecular evolution of multisubunit RNA polymerases: sequence analysis. *J. Mol. Biol.* 395, 671–685. doi: 10.1016/j.jmb.2009.10.062

Lane, W. J., and Darst, S. A. (2010b). Molecular evolution of multisubunit RNA polymerases: structural analysis. *J. Mol. Biol.* 395, 686–704. doi: 10.1016/j.jmb.2009.10.063

Lei, L., and Burton, Z. F. (2020). Evolution of life on earth: tRNA, Aminoacyl-tRNA synthetases and the genetic code. *Life* 10:21. doi: 10.3390/life10030021

Littlefield, O., Korkhin, Y., and Sigler, P. B. (1999). The structural basis for the oriented assembly of a TBP/TFB/promoter complex. *Proc. Natl. Acad. Sci. U.S.A.* 96, 13668–13673. doi: 10.1073/pnas.96.24.13668

Liu, B., and Steitz, T. A. (2017). Structural insights into NusG regulating transcription elongation. *Nucleic Acids Res.* 45, 968–974. doi: 10.1093/nar/gkw1159

Long, X., Xue, H., and Wong, J. T. (2020). Descent of bacteria and Eukarya from an archaeal root of life. *Evol. Bioinform. Online* 16, 1176934320908267.

Madru, C., Henneke, G., Raia, P., Hugonneau-Beaufet, I., Pehau-Arnaudet, G., England, P., et al. (2020). Structural basis for the increased processivity of D-family DNA polymerases in complex with PCNA. *Nat. Commun.* 11:1591.

Marin, J., Battistuzzi, F. U., Brown, A. C., and Hedges, S. B. (2017). The timetree of prokaryotes: new insights into their evolution and speciation. *Mol. Biol. Evol.* 34, 437–446.

Osman, S., and Cramer, P. (2020). Structural biology of RNA Polymerase II transcription: 20 years on. *Annu. Rev. Cell Dev. Biol.* 36, 1–34. doi: 10.1146/annurev-cellbio-042020-021954

Pak, D., Du, N., Kim, Y., Sun, Y., and Burton, Z. F. (2018). Rooted tRNAomes and evolution of the genetic code. *Transcription* 9, 137–151. doi: 10.1080/21541264.2018.1429837

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., et al. (2021). Structure visualization for researchers, educators, and developers. *Protein Sci.* 30, 70–82. doi: 10.1002/pro.3943

Qian, X., Hamid, F. M., El Sahili, A., Darwis, D. A., Wong, Y. H., Bhushan, S., et al. (2016). Functional evolution in orthologous cell-encoded RNA-dependent RNA polymerases. *J. Biol. Chem.* 291, 9295–9309. doi: 10.1074/jbc.m115.685933

Raia, P., Carroni, M., Henry, E., Pehau-Arnaudet, G., Brule, S., Beguin, P., et al. (2019). Structure of the DP1-DP2 PolD complex bound with DNA and its implications for the evolutionary history of DNA and RNA polymerases. *PLoS Biol.* 17:e3000122. doi: 10.1371/journal.pbio.3000122

Renfrow, M. B., Naryshkin, N., Lewis, L. M., Chen, H. T., Ebright, R. H., and Scott, R. A. (2004). Transcription factor B contacts promoter DNA near the transcription start site of the archaeal transcription initiation complex. *J. Biol. Chem.* 279, 2825–2831. doi: 10.1074/jbc.m311433200

Salgado, P. S., Koivunen, M. R., Makeyev, E. V., Bamford, D. H., Stuart, D. I., and Grimes, J. M. (2006). The structure of an RNAi polymerase links RNA silencing and transcription. *PLoS Biol.* 4:e434. doi: 10.1371/journal.pbio.0040434

Sauguet, L. (2019). The extended "Two-Barrel" polymerases superfamily: structure, function and evolution. *J. Mol. Biol.* 431, 4167–4183. doi: 10.1016/j.jmb.2019.05.017

Tomar, S. K., and Artsimovitch, I. (2013). NusG-Spt5 proteins-Universal tools for transcription modification and communication. *Chem. Rev.* 113, 8604–8619. doi: 10.1021/cr400064k

Vassylyev, D. G., Vassylyeva, M. N., Perederina, A., Tahirov, T. H., and Artsimovitch, I. (2007a). Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* 448, 157–162. doi: 10.1038/nature05932

Vassylyev, D. G., Vassylyeva, M. N., Zhang, J., Palangat, M., Artsimovitch, I., and Landick, R. (2007b). Structural basis for substrate loading in bacterial RNA polymerase. *Nature* 448, 163–168. doi: 10.1038/nature05931

Wang, B., and Artsimovitch, I. (2020). NusG, an Ancient Yet rapidly evolving transcription factor. *Front. Microbiol.* 11:619618. doi: 10.3389/fmicb.2020.619618

Werner, F. (2012). A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. *J. Mol. Biol.* 417, 13–27. doi: 10.1016/j.jmb.2012.01.031

Werner, F., and Grohmann, D. (2011). Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* 9, 85–98. doi: 10.1038/nrmicro2507

Yakhnin, A. V., and Babitzke, P. (2014). NusG/Spt5: are there common functions of this ubiquitous transcription elongation factor? *Curr. Opin. Microbiol.* 18, 68–71. doi: 10.1016/j.mib.2014.02.005

Yang, Y., Darbari, V. C., Zhang, N., Lu, D., Glyde, R., Wang, Y. P., et al. (2015). Structures of the RNA polymerase-sigma54 reveal new and conserved regulatory strategies. *Science* 349, 882–885. doi: 10.1126/science.aab1478

Zatopek, K. M., Alpaslan, E., Evans, T. C., Sauguet, L., and Gardner, A. F. (2020). Novel ribonucleotide discrimination in the RNA polymerase-like two-barrel catalytic core of Family D DNA polymerases. *Nucleic Acids Res.* 48, 12204–12218. doi: 10.1093/nar/gkaa986

Zuo, Y., and Steitz, T. A. (2015). Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble. *Mol. Cell* 58, 534–540. doi: 10.1016/j.molcel.2015.03.010

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 11 May 2021
doi: 10.3389/fmolb.2021.669314

Transcribing Genes the Hard Way: *In Vitro* Reconstitution of Nanoarchaeal RNA Polymerase Reveals Unusual Active Site Properties

Sven Nottebaum[1,2] *and Robert O. J. Weinzierl*[1]*

[1]Department of Life Sciences, Imperial College London, London, United Kingdom, [2]Orthomol Pharmazeutische Vertriebs GmbH, Langenfeld, Germany

Nanoarchaea represent a highly diverged archaeal phylum that displays many unusual biological features. The *Nanoarchaeum equitans* genome encodes a complete set of RNA polymerase (RNAP) subunits and basal factors. Several of the standard motifs in the active center contain radical substitutions that are normally expected to render the polymerase catalytically inactive. Here we show that, despite these unusual features, a RNAP reconstituted from recombinant *Nanoarchaeum* subunits is transcriptionally active. Using a sparse-matrix high-throughput screening method we identified an atypical stringent requirement for fluoride ions to maximize its activity under *in vitro* transcription conditions.

Keywords: archaea, nanoarchaea, RNA polymerase, catalytic center, active site, high-throughput assay, sparse matrix sampling, fluoride

OPEN ACCESS

Edited by:
Francisco Navarro, University of Jaén, Spain

Reviewed by:
Thomas J. Santangelo, Colorado State University, United States
Zachary Burton, Michigan State University, United States

***Correspondence:**
Robert O. J. Weinzierl
r.weinzierl@imperial.ac.uk

Specialty section:
This article was submitted to Protein and RNA Networks, a section of the journal Frontiers in Molecular Biosciences

Received: *18 February 2021*
Accepted: *26 April 2021*
Published: *11 May 2021*

Citation:
Nottebaum S and Weinzierl ROJ (2021) Transcribing Genes the Hard Way: In Vitro *Reconstitution of Nanoarchaeal RNA Polymerase Reveals Unusual Active Site Properties. Front. Mol. Biosci. 8:669314. doi: 10.3389/fmolb.2021.669314*

INTRODUCTION

The basal transcriptional machineries of Archaea are intriguingly similar to the core components of the eukaryotic RNA polymerase II (RNAPII) transcriptional machinery (Cramer et al., 2001). This close similarity to eukaryotic systems, combined with the greater experimental accessibility, has established archaeal systems as key model systems for in-depth structure/function analyses of the transcriptional machinery (Werner and Weinzierl 2002; Ouhammouch et al., 2004; Werner et al., 2006; Naji et al., 2007; Hirata et al., 2008; Tan et al., 2008; Thomm et al., 2009; Weinzierl 2013; Fouqueau et al., 2018; Blombach et al., 2019; Wenck and Santangelo, 2020). Apart from serving as model systems for eukaryotic systems, archaea also include numerous examples of extremophiles that do not fit the general pattern (Adam et al., 2017). Such species often provide unusual examples of molecular organization that have the capacity of enlarging our understanding of fundamental molecular mechanisms by illustrating the degree of flexibility that is possible, or by providing examples for achieving the same goal in a variety of alternative ways (Coker, 2019). Some of the best-known examples include the adaption of enzymes to operate in high-salt environments (halophiles), over a wide range of temperatures (psychrophile, mesophiles, thermophiles, hyperthermophiles), or at low or high pH (acidophiles and alkaliphiles, respectively). Another interesting class of archaea are the evolutionary "outliers", such as *Methanopyrus kandleri*, *Cenarchaeum symbiosum,* and *Nanoarchaeum equitans*. The phylogenetic classification of these species is contentious, and their protein sequences frequently contain unique and unusual substitutions that are not shared by other archaea. Such unorthodox features raise many, yet unanswered, questions regarding the evolutionary origin of such species (deep-branching evolutionary ancestry or recent degeneracy?) and often

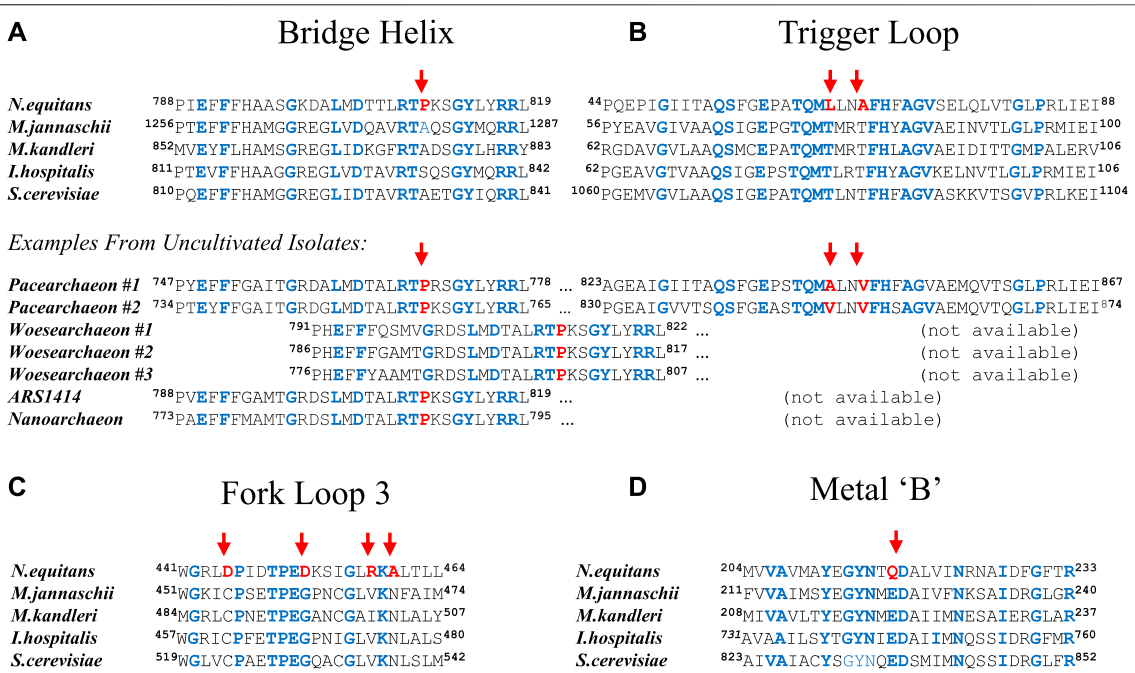


FIGURE 1 | Substitutions in key regions and domains of the nanoarchaeal RNAP catalytic site. Sequences from four archaeal (*Nanoarchaeum equitans*, *Methanocaldococcus jannaschii*, *Methanopyrus kandleri*, and *Ignicoccus hospitalis*) and one eukaryotic (*Saccharomyces cerevisiae*; RNA polymerase II) species are shown in all panels. Unusual substitutions in the nanoarchaeal sequence are highlighted with a red arrow. Residues identical in all sequences shown are highlighted in blue. The beginning and end positions of the sequences shown relative to the full length protein sequence are indicated as superscripts **(A)** Alignment of Bridge Helix sequences. *N. equitans* (SeqID: AAR39345.1), *M. jannaschii* (SeqID: WP_064,496,945.1), *M. kandleri* (SeqID: AAM01900.1), *I. hospitalis* (SeqID WP_011,998,279.1) and *S. cerevisiae* (SeqID: NP_010141.1). Below, additional examples from uncultivated and yet unnamed species are shown (*Candidatus Pacearchaeota archaeon #1* [SeqID: MAG61561.1; RNAP subunit combines A′ and A″ as continuous polypeptide]; *Candidatus Pacearchaeota archaeon #2* [isolate CG_2015–01t_39_43; SeqID: NCO11196.1; RNAP subunit combines A′ and A″ as continuous polypeptide]; *Candidatus Woesearchaeota archaeon* #1 [CG1_02_47_18; SeqID: OIO63522.1; A′ only) (Probst et al., 2017)]; *Candidatus Woesearchaeota archaeon#2* [isolate SpSt-512; SeqID: HGS79070.1; A′ only) (Zhou et al., 2020)]; *Candidatus Woesearchaeota archaeon#3* [isolate SpSt-1178; SeqID: HDP74066.1) (Zhou et al., 2020)]; Archaeal isolate ARS1414 [SeqID: MAG50098.1; A′ only) (Tully et al., 2018)], Nanoarchaeota archaeon [SeqID: NTV23449.1; Breister et al.]) **(B)** Alignment of Trigger Loop sequences. *N. equitans* (SeqID: AAR39272.1), *M. jannaschii* (SeqID: WP_010,870,556.1), *M. kandleri* (SeqID: WP_0,11,019,054.1), *I. hospitalis* (SeqID WP_052,570,437.1), and *S. cerevisiae* (SeqID: NP_010141.1) **(C)** Sequence alignment of the Fork-Loop 3 motif. *N. equitans* (SeqID: AAR39027.1), *M. jannaschii* (SeqID: Q58444.1), *M. kandleri* (SeqID: WP_088,335,828.1), *I. hospitalis* (SeqID: WP_052,570,488.1), and *S. cerevisiae* (SeqID: AAA68096.1) **(D)** Sequence alignment of the Metal-B motif. *N. equitans* (SeqID: AAR39011.1), *M. jannaschii* (SeqID: Q60181.1), *M. kandleri* (SeqID: WP_193,333,232.1), *I. hospitalis* (SeqID: WP_052,570,488.1), and *S. cerevisiae* (SeqID: AAA68096.1).

challenge fundamental concepts of apparently well-understood enzymatic pathways and mechanisms (Randau et al., 2005; Randau et al., 2008).

Here we focus on the molecular organization and properties of the RNAP from the hyperthermophile *Nanoarchaeum equitans* (from here on abbreviated as *N. equitans*, or "*n*" as a prefix). *N. equitans* is a highly unusual archaeon because of its diminutive size (0.35–0.5 μm cell diameter), drastically reduced genome and parasitic lifestyle (Huber et al., 2002; Waters et al., 2003; Forterre et al., 2009; Rawle et al., 2017). The genome lacks most genes that are required to produce cellular precursors, such as amino acids, nucleotides, cofactors, and lipids. These are most likely imported directly from the host cell, the crenarchaeote *Ignicoccus hospitalis*. Depending on the criteria chosen, *N. equitans* has been plausibly classified as a new and early diverging archaeal phylum (the "Nanoarchaeota" (Huber et al., 2003)), a sister branch of the Crenarchaea (Ciccarelli et al., 2006), or as a fast-evolving Euryarchaeon (Brochier et al., 2005). Recent studies have demonstrated that Nanoarchaea are widespread and occur in a variety of locations, including mesophilic and halophilic environments (Hohn et al., 2002; McCliment et al., 2006; Probst et al., 2017; Tully et al., 2018; Zhou et al., 2020).

The *N. equitans* genome encodes a full complement of all RNA polymerase (RNAP) subunits and basal factors TBP, TFB, TFE, and TF-S (Huber et al., 2002; Waters et al., 2003). Considering the minimal size of the genome, the presence of a set of genes encoding a complete basal transcriptional machinery strongly suggests that *N. equitans* is fully capable of transcribing its own genome. We observe, however, a distinct set of substitutions in several key positions of the *ne*RNAP catalytic center that are of a unique and radical nature and raise the question whether such an enzyme could display a substantial amount of catalytic activity. The Bridge Helix (BH), Trigger Loop (TL), Fork Loop-3 (FL-3), as well as the "Metal B" binding domain (Me-B; responsible for positioning one of the two catalytically active Mg^{2+} ions) display substitutions in positions that are typically absolutely or highly conserved in all other archaeal and eukaryotic RNAPs (**Figure 1**). Some of these substitutions (such as the presence of a proline in the Bridge Helix (BH); **Figure 1A**) are predicted to have highly disruptive, non-local effects by destabilizing the *a*-helical integrity

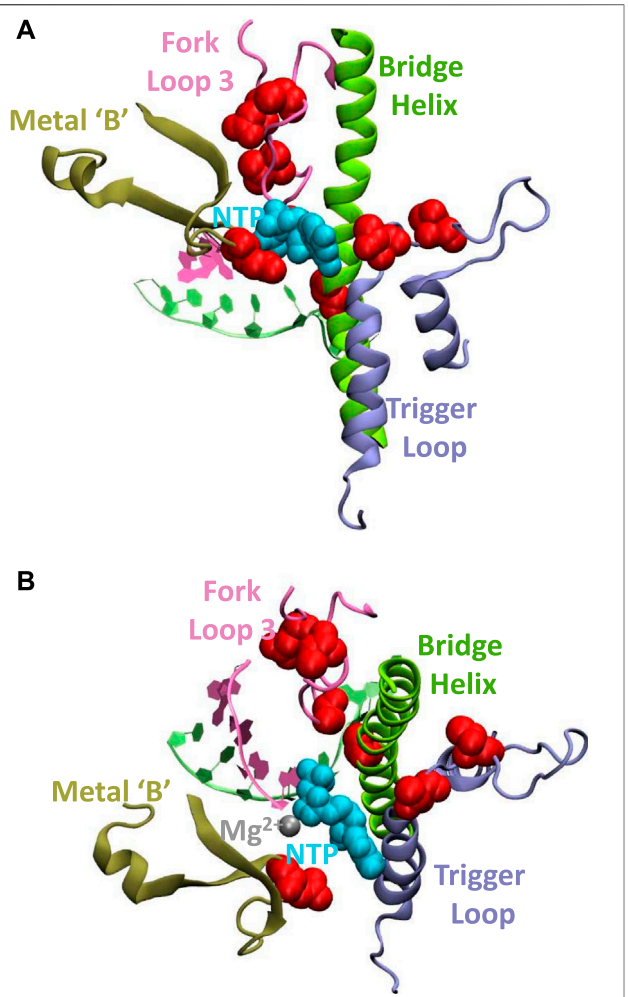


FIGURE 2 | Spatial arrangement of nanoarchaeal-specific substitutions within the catalytic site of RNAP. The Bridge Helix (BH) is shown in green, the Trigger Loop (TL) in blue and Fork Loop 3 (FL-3) in purple. The nucleotide triphosphate is represented as a cyan space-filling model. Positions substituted in *N. equitans* are shown as red space-filling van der Waals representations.

of such a key element in a particularly critical position (Tan et al., 2008; Weinzierl, 2010b, Weinzierl, 2010a, Weinzierl, 2011). Although proline substitutions in particular places of BH results in a substantial increase the specific activity of the structurally closely related euryarchaeal RNAP from *Methanocaldococcus jannaschii* (*mj*RNAP) (Tan et al., 2008; Weinzierl, 2010b), a proline located in the position characteristic for *ne*A′ causes a substantial drop in activity in *mj*RNAP (Tan et al., 2008). Several other unusual substitutions in other key elements of the catalytic site (Cramer et al., 2001) are evident, including the Trigger Loop (TL; **Figure 1B**), Fork Loop-3 (FL3; **Figure 1C**) and the Metal-B motif required to coordinate the Mg^{2+} ions facilitating the various types of catalytic chemistries (Sosunov et al., 2003); Me-B; **Figure 1D**). All these nanoarchaeal substitutions are spatially in close vicinity within the catalytic site of RNAP (**Figure 2**). Based on our current understanding of the structural basis of the nucleotide addition cycle, such substitutions would be predicted to have a substantially deleterious effect on the catalytic function of the *ne*RNAP active site. In comparison, the RNAP of the archaeon *I. hospitalis* - the host to *N. equitans* - does not encode any of these unusual substitutions found in the *ne*RNAP (**Figure 1**), thus essentially ruling out that the substitutions are required to survive in a particular environment.

Recent large-scale sequencing efforts have demonstrated that similar unusual substitution patterns can been found in hundreds of sequence samples derived from fresh- and marine water sources (Probst et al., 2017; Tully et al., 2018; Zhou et al., 2020). Several data base entries, labeled as yet unnamed representatives of Woesearchaea or Pacearchaea, show the same types of substitutions as originally found in *N. equitans* (**Figure 1A**). Although the Bridge Helix and Trigger Loop are usually encoded by separate subunits of archaeal RNAPs (A′ and A″, respectively) - and can therefore usually not be allocated to the same species in environmental sequencing samples - there are two pacearchaeal sequences where A' and A" appear to be fused into a single subunit (**Figure 1A**; directly comparable to the eukaryotic large RNAP subunits). We can therefore see from these examples, that - like in *N. equitans* - the specific substitution pattern in both Bridge Helix and Trigger Loop are encoded within an RNAP subunit from the same species. This suggests that, although for a long time considered unusual, *N. equitans* is actually a fairly typical representative of a larger group of archaea (including Pacearchaea, Woesearchaea etc.) that display comparable, but structurally diverged RNAP active site architectures.

The goal of this study was to determine whether the RNAP encoded by the *N. equitans* genome was 1) enzymatically active and 2) to what extent the substitution pattern resulted in altered catalytic properties.

MATERIALS AND METHODS

Identification of neRNAP Subunits and Basal Transcription Factors

*ne*RNAP subunit-encoding open reading frames were identified using existing data base annotations and tblastn searches of the *Nanoarchaeum equitans* genome sequence (SeqID: AE017199.1; see **Supplementary Table S1** for more details).

Markov chain Monte Carlo Simulations

Markov Chain Monte Carlo MCMC simulations were carried out as described previously (Sullivan and Weinzierl, 2020). Briefly, the simulations employed the PROFASI forcefield in the PHAISTOS package (Boomsma et al., 2013). Due to the origin of the proteins from hyperthermophilic organisms, the simulation temperature was set to 355 K (81.85°C). The resulting trajectory data (based on 50,000 calculated states per simulation) was analyzed for secondary structure elements using cpptraj (Roe and Cheatham, 2013) and processed/visualized with custom scripts on *Python* Jupyter notebooks.

In Vitro Reconstitution *ne*RNAP

The protein-coding portions of RNAP subunits A′, A″, B′, B″, D, H, L, N and *p* were PCR amplified from purified *N. equitans*

genomic DNA (a gift from Prof. M. Thomm, University of Regensburg) as full-length, non-tagged sequences and cloned as NdeI-BamHI (*ne*A″, *ne*B′, *ne*B″, *ne*F, *ne*K), or NdeI-EcoRI (*ne*A′, *ne*D, *ne*E, *ne*H, *ne*L, *ne*N, *ne*P) fragments into the bacterial expression vector pET21a. Recombinant proteins were expressed with IPTG-induction in *E. coli* BL21-DE3 Rosetta 2 (Merck) under standard conditions (Werner and Weinzierl, 2002). Subunits *ne*A′, *ne*A″, *ne*B' and *ne*B″ were purified as insoluble inclusion bodies. Briefly, bacterial cells expressing these recombinant subunits were resuspended in T/G_0 (25 mM Tris-base, 200 mM glycine, 10 mM magnesium acetate, 100 µM zinc acetate, 14 mM ß-mercaptoethanol and 10% glycerol at pH7.5) and sonicated. The inclusion bodies were washed extensively with 1 x deoxycholate buffer (1 mg/ml deoxycholate, 15 mM ß-mercaptoethanol) and water/15 mM ß-mercaptoethanol before solubilizing them in T/G_0 in the presence of saturating urea or 6 M guanidine-hydrochloride. Subunits *ne*D, *ne*L, *ne*H, *ne*N and *ne*P were expressed similarly as soluble recombinant proteins. Bacterial cells expressing these recombinant subunits were resuspended in P300 Buffer (300 mM potassium acetate, 20 mM Tris-acetate pH 7.9, 10 mM magnesium acetate, 100 µM zinc acetate, 14 mM ß-mercaptoethanol and 10% glycerol) and sonicated. The supernatant containing the solubilized proteins were heat-inactivated of at 70°C for 10 min to precipitate the bacterial proteins present in the extract (the hyperthermophilic *ne* subunits remain completely soluble during this treatment).

The urea-solubilized inclusion bodies, or the soluble subunits, were passed over ~5 ml SP- or Q-Sepharose (Fast flow, Amersham) in chromatography columns. Proteins were eluted in a salt gradient from T/G0 to T/G1000 using a DuoFlow BioLogic FPLC system (BioRad). The purified subunits were assembled by mixing them in the presence of 8 M urea in a dialysis cell (Slide-A-Lyzer 3500MCOW frames [Pierce], or 96-well microdialyser (SpectraPor) on a Theonix robotic platform (Aviso) for high-throughput assembly (Nottebaum et al., 2008; Weinzierl, 2013)), followed by lowering the urea concentration by gradual dilution in the dialysis buffer (Werner and Weinzierl, 2002; Naji et al., 2007; Nottebaum et al., 2008; Weinzierl, 2013). Equimolar amounts of the large subunits were mixed with small subunits, which were in at least four-fold excess to the large subunits, under denaturing conditions. The highest yield of enzymatically active *ne*RNAP (due to increased folding efficiency) was obtained in the presence of 500 mM salt (either sodium chloride, potassium- or sodium acetate) in the refolding buffer (**Supplementary Figure S1**). The assembly of large complexes was monitored by analytical size exclusion chromatography on Superose-6 and Superose-12 columns (Amersham) at a flow-rate of 0.5 ml/min. When required, soluble protein complexes were concentrated further using centrifugal YM-50 Centricon (Millipore) units according to the manufacturers instructions.

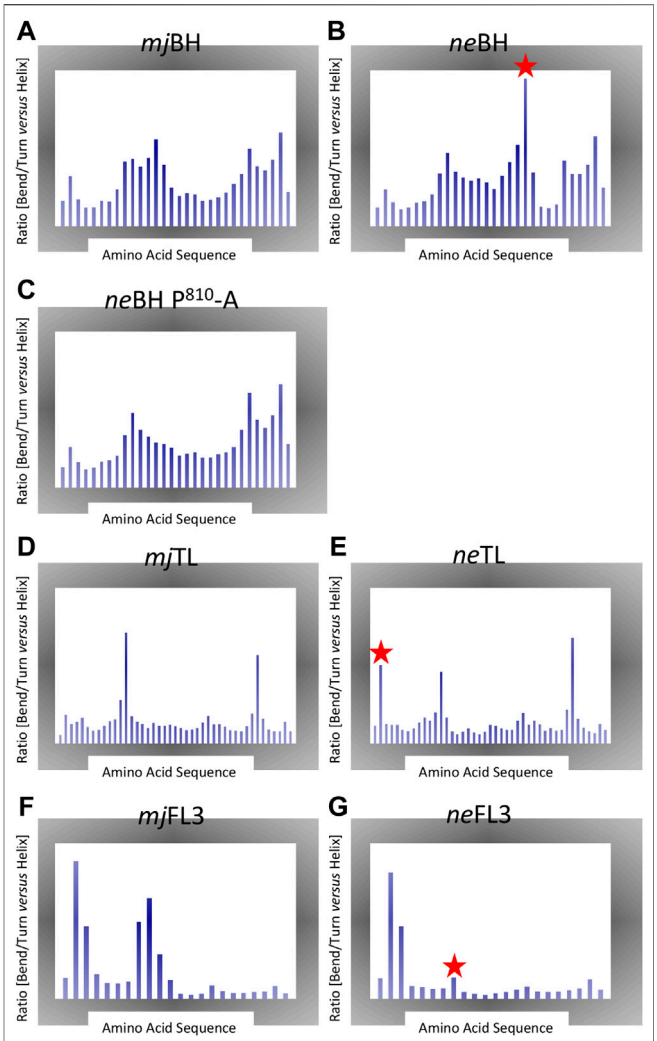


FIGURE 3 | Markov Chain Monte Carlo (MCMC). Each plot shows the ratio of percentage of disordered secondary structure elements (bends, turns) vs. percentage of helical structures (including *a*-, 3–10 and *p* helix data) created during the simulations. Large peaks therefore highlight regions displaying local structural instability. The *x*-axis shows the amino acid positions for each element. Note that the range of *y*-axes is variable between different plot groups. Significant local structural variations present in the *ne* motifs are highlighted with a red star **(A)** Bridge Helix from *M. jannaschii* **(B)** Bridge Helix from *N. equitans* **(C)** Bridge Helix from *M. jannaschii* simulated with *in silico* mutated *ne*A′ P^{810}-A **(D)** Trigger Loop from *M. jannaschii* **(E)** Trigger Loop from *N. equitans* (E) Fork Loop **3** from *M. jannaschii* (F) Fork Loop 3 from *N. equitans*.

In Vitro Transcription Assay

Refolded RNAPs were assayed for transcriptional activity by measuring incorporation of *a*-^{32}P-UTP into RNA. Refolded RNAP was added to 1 x transcription buffer (1 x TB) containing 500 µM ATP, CTP, GTP, 1 µM UTP, 27 nM *a*-^{32}P-UTP (6000 Ci/mmol, Amersham), 1.5 µg nuclease-activated calf thymus DNA (Fluka), 120 mM potassium acetate, 10 mM magnesium acetate, 10 mM Tris-acetate pH 7.5 and 10 mM DTT, which was incubated at 37–65°C for 45 min. The final reaction volume was 50 µL. The reactions were stopped by addition of 15% (w/v) trichloroacetic acid followed by 30 min incubation on ice. The precipitate was collected on 96-well GF/F glass fiber filter plates (Whatman), washed twice with excess 10% TCA, once with 95% ethanol, and quantitated in a scintillation counter in presence of scintillant fluid (Opti-fluor, Packard Bioscience). These steps were fully automated on a Theonyx liquid

handling platform (Aviso) (Nottebaum et al., 2008; Weinzierl, 2013). Independent repetitions ("biological replicates") of the same transcription reaction are reproducible within a 12% error margin. Transcripts originating solely from abortive initiation are not precipitated using this method. Therefore, only transcripts from elongation-competent RNAPs (longer than ~20 nucleotides) give rise to a signal in this assay.

High-Throughput "Sparse-Matrix" Sampling

Crystallization buffer sets ICL-1, -3, -4, and -5 (Hampton Research, Aliso Viejo, United States; **Supplementary Figures S2A–F**) were used as 10 x concentrates for high-throughput transcription assays based on nicked DNA templates as described previously (Werner and Weinzierl, 2002; Nottebaum et al., 2008; Tan et al., 2008; Boomsma et al., 2013). Briefly, the assay measures the incorporation of *a*-^{32}P-UTP into acid-insoluble RNA by liquid scintillation counting in a robotically implemented workflow. Similar to the strategy used when employing such buffer sets for crystallization screens, the initial screen was only carried out with one assay per buffer set. Buffer sets that gave high levels of activity were subsequently tested in triplicate to confirm the result.

RESULTS

Computational Simulations

The structural consequences of several of the substitutions were determined by comparing secondary structure propensities of sequences from *N. equitans* to equivalent domains from *M. jannaschii* (*mj*). Both species are hyperthermophiles thus containing similar sequence-encoded features that stabilize their protein structures at elevated temperatures. Markov Chain Monte Carlo (MCMC) simulations is the method of choice for a systematic and comprehensive exploration of conformational space (Boomsma et al., 2013) (**Figure 3**). As expected, the presence of a proline in the *ne*BH causes a substantial disruption of a region that displays high *a*-helical propensity in the *mj*BH (Tan et al., 2008). On a structural level, the presence of proline in *ne*A′ in position 810 (*ne*A′ P^{810}) is predicted to cause a substantial destabilization of the *a*-helical conformation of the Bridge Helix in a slightly more N-terminal location (mostly affecting *ne*A′ R^{808}; **Figure 3B**). Simulating a *ne* Bridge Helix with a "corrected" *in silico* point mutation (*ne*A′ P^{810}-A) restores the predicted conformational population to one that is very close to the *mj* Bridge Helix (**Figure 3C**). This proves that the unusual conformational properties are predominantly due to *ne*A′ P^{810} position, rather than any of the other differences in the primary amino acid sequence. The region in the *ne* Bridge Helix most distorted corresponds to the orthologous region in *M. jannaschii* (*mj*A′-R^{820}) which is a structure with one of the highest *a*-helical propensities of the entire domain (**Figure 3A**). High-throughput mutagenesis studies of *mj*A′-R^{820} in *mj*RNAP have shown it to be highly sensitive to point mutations, with only phenylalanine and tryptophane substitutions not resulting in substantial loss of catalytic activity (Tan et al., 2008; Weinzierl, 2013).

Similar comparisons of the *ne* and *mj* Trigger Loop conformations yield a less clear-cut result (**Figure 3D,E**), although the unusual position of a proline in the *ne* motif near the edges of the domain (*ne*A"-P^{47}) again is likely to contribute a destabilizing influence (**Figure 3E**).

For *ne*FL3, the structural consequences of replacing highly conserved residues that are virtually invariant in other archaeal and eukaryotic polymerases in a non-conservative manner (for example, in FL3: C-D^{445}, G-D^{452}, V/I-R^{458}, and N-A^{460}; **Figure 1C**) suggest that this would cause distinct changes in the functional contributions of these residues to catalysis. Especially one of these substitutions *ne*B" G^{452}-D is predicted to reduce the flexibility of the central region of *ne*FL3 considerably in comparison to the orthologous sequence of *mj*FL3 (**Figures 3F,G**). Similarly, Metal B contains two highly conserved acidic residues that coordinate of binding of the Mg^{2+} ion brought along by the incoming NTP, but in *N. equitans* one of them is converted to glutamine (*ne*A' Q217) and thus is predicted bind the metal less strongly (**Figure 1D**).

Overall, based on previous insights from a range of structural and functional studies from archaeal and eukaryotic RNAPs representative of the majority of such organisms, a picture of a structurally diverged catalytic site in nanoarchaeal RNAPs emerges that suggests that the catalytic site may be more flexible in some areas (prolines in the *ne*Bridge Helix and *ne*Trigger Loop domains, stiffer in the diverged *ne*FL-3 domain and potentially compromised electrostatically by a diminished *ne*Metal-B motif).

In Vitro Assembly of *ne*RNAP and High-Throughput "Sparse-Matrix" Sampling of *ne*RNAP Assay Conditions

The conformational distortions caused by potentially disruptive radical substitution suggest that the *Nanoarchaeum* RNAP may display only very low - or even no - catalytic activity. On the other hand, the presence of all known RNAP subunits in an otherwise minimal genome implies selective pressure responsible for maintaining an active transcriptional machinery. Technical problems with obtaining *N. equitans* in quantities sufficient for biochemical analysis preclude a direct purification of native enzymes from cells. We therefore decided to investigate this question by adopting the *in vitro* assembly approach that has been applied successfully for the assembly of RNAPs from other hyperthermophilic archaea (Werner and Weinzierl, 2002; Naji et al., 2007). The *in vitro* assembly of *ne*RNAP followed essentially the same procedure that we employed successfully in the past for *mj*RNAP (Werner and Weinzierl, 2002). Each of the subunits essential for catalytic activity was expressed as a recombinant protein in *E. coli*, followed by chromatographic purification and *in vitro* assembly by controlled dialysis from denaturing conditions (**Figure 4A**). Under these conditions, a portion of the *ne*RNAP subunits assembled into a complex that - comparable to *mj*RNAP (**Figure 4B**) - eluted as a distinct peak of activity during size exclusion chromatography (**Figure 4C**). As expected from its hyperthermophilic origin, the temperature optimum for catalytic activity was around 76°C (**Supplementary Figures S3**). Initial transcription

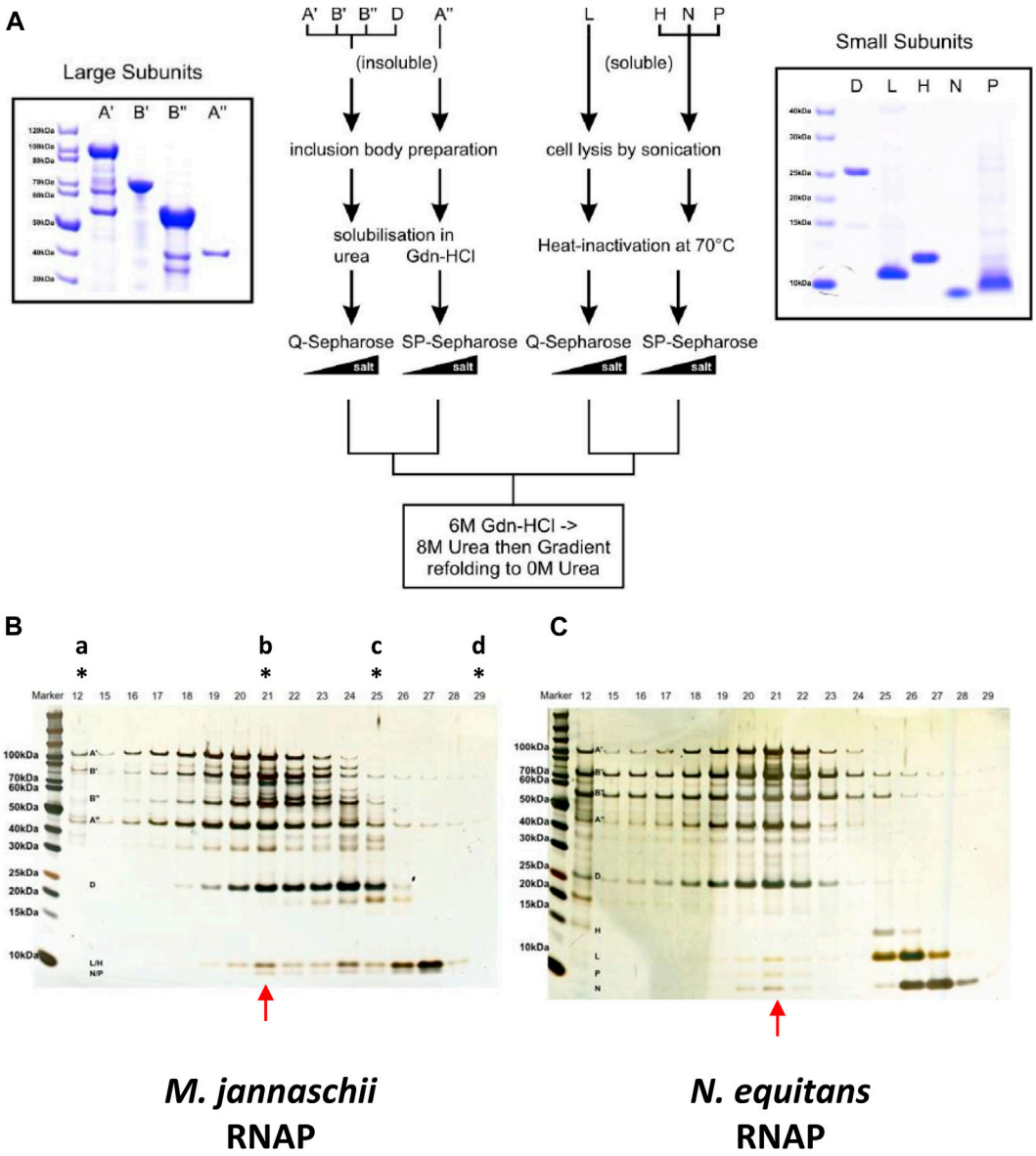


FIGURE 4 | (A) Overview of the purification of *ne*RNAP subunits. The central scheme outlining the purification procedure for each subunit is flanked by Coomassie-stained gels of the purified subunits (left: *ne*A′, *ne*A", *ne*B′ and *ne*B"; right: *ne*D, *ne*H, *ne*L, *ne*N and *ne*P) **(B)** Elution profile of the *mj*RNAP *in vitro* assembly reaction from a Superose-6 size-exclusion column (similar to (Werner and Weinzierl, 2002)) shown on a silver-stained Bis/Tris 4–12% gradient SDS-protein gel. Fraction 21 (indicated with red arrow) contains the fully assembled enzyme (and peak transcriptional activity; data not shown) as revealed by the presence of all subunits within a single fraction. The letters with stars on top show the fractions where the size exclusion markers ("a", Blue Dextran 2,000 kDa; "b", *β*-amylase 200 kDa; "c", carbonic anhydrase 25 kDa; "d", cytochrome c 12.4 kDa) eluted **(C)** Similar to (B), but for the *ne*RNAP *in vitro* assembly.

experiments with *ne*RNAP suggested that the standard buffer conditions (120 mM potassium acetate, 10 mM magnesium acetate, 20 mM Tris-acetate, pH 8.6) were probably suboptimal because we observed a ~ 7-fold lower specific activity for *ne*RNAP as compared to *mj*RNAP when assembled in parallel. We therefore attempted to optimize the assay conditions over a wider range of pH values, salt concentrations and in the presence of various additives. The concept of "sparse-matrix" sampling is well established in the macromolecular crystallization community where the method is used to identify the optimal (yet initially unknown) conditions to obtain macromolecular crystals for structural studies (Jancarik and Kim, 1991). Such approaches have also been employed usefully to identify optimal renaturation conditions (Hofmann et al., 1995), or for stabilizing macromolecular complexes (Chari et al., 2015). Here, we employed such a strategy to identify the best assay conditions for *ne*RNAP that included a wide range of different concentrations of various cations and anions, buffers at different pHs, and the presence

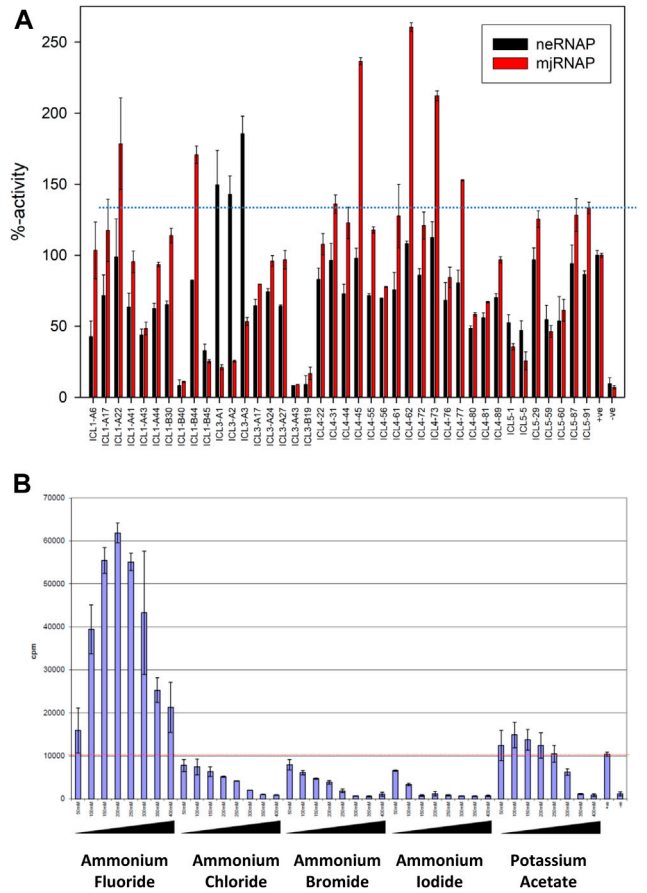


FIGURE 5 | Effect of different assay buffer compositions on the catalytic activities of *ne*RNAP and *mj*RNAP **(A)** The transcriptional activity under "standard conditions" is defined as 100%. The activities of *ne*RNAP in this set of buffers is shown in black, and the performance of *mj*RNAP under the same conditions is shown in red (B) Fluoride-specific effect among halogen salts. The catalytic activity of *ne*RNAP at various salt concentrations ((50–400 mM) is shown, including ammonium fluoride, ammonium chloride, ammonium bromide, ammonium iodide and potassium acetate. Fluoride has the most distinct effect.

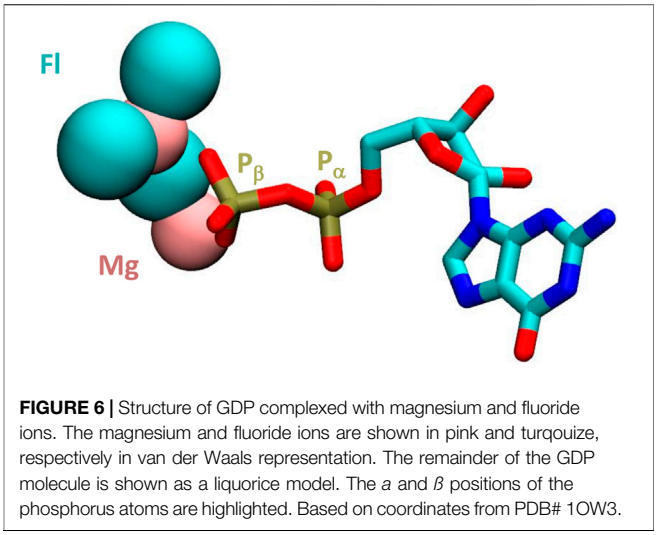


FIGURE 6 | Structure of GDP complexed with magnesium and fluoride ions. The magnesium and fluoride ions are shown in pink and turqouize, respectively in van der Waals representation. The remainder of the GDP molecule is shown as a liquorice model. The *a* and *ß* positions of the phosphorus atoms are highlighted. Based on coordinates from PDB# 1OW3.

of a variety of detergents and stabilizing reagents. A series of buffer sets (ICL-1, ICL-3, ICL-4, and ICL-5; Hampton Research), comprised of 386 different cocktails (see **Supplementary Figures S2A–F** for composition), were used as 10 x stock solutions after supplementing them with Mg^{2+} and Zn^{2+} in automated high-throughput transcription assays. Most of the mixtures include an inorganic or organic salt, a buffering compound (with pH ranges from 4.5 to 9.5) and a "precipitant", such as polyethylene glycol. In our assays, the precipitant may display stabilizing effects on protein structure - especially quaternary structures - under hyperthermophilic assay conditions.

A summary of the results (see **Supplementary Figure S4** for the complete data set) shows that *ne*RNAP had a clear preference for a group of three buffers (ICL-3 #A1, A2, A3) that contained 20 mM sodium fluoride, potassium fluoride and ammonium fluoride, respectively (**Figure 5A**). This apparent preference for fluoride is unique to *ne*RNAP because *mj*RNAP only performed moderately (30–50% in comparison to standard conditions) in these buffers (**Figure 5A**). To test this potential requirement for fluoride further, *ne*RNAP activity was assayed in transcription buffers containing varying amounts of fluoride salts. Optimal *ne*RNAP stimulation was achieved with 200–300 mM potassium fluoride or ammonium fluoride (**Figure 5B** and **Supplementary Figure S5**). The stimulation of *ne*RNAP activity by fluoride ions raised the question of whether other halogen ions (chloride, bromide, or iodide) would have a similar effect on *ne*RNAP. This, however, was not the case, suggesting that the stimulating effect on the catalytic activity of *ne*RNAP is indeed highly specific for fluoride.

DISCUSSION

Nanoarchaea are, in many ways, puzzling organisms. Their unique parasitic lifestyle has substantial effects on their cell- and genome size, which are both greatly minimized (Huber et al., 2002; Huber et al., 2003; Waters et al., 2003). Therefore, the cells depend on their host, *I. hospitalis* for many metabolites and precursors (Rawle et al., 2017). Analysis of the *N. equitans* genome has, however, revealed the presence of orthologs of all RNAP subunits and other components of the basal transcriptional machinery (TBP, TFB, and TFS; (Huber et al., 2002; Waters et al., 2003). It therefore looks as if *N. equitans* is capable of transcribing its own genome without help from its host cell in terms of imported basal transcription factors. Nevertheless, a number of key domains and motifs that constitute the active site of RNAP contains a distinct set of highly unusual and radical substitutions that appear to be deleterious to its catalytic activity.

Here we show, by *in vitro* assembly of nanoarchaeal RNAP from recombinant subunits expressed in and purified individually from *E. coli*, that the resulting enzyme displays catalytic activity. The temperature, pH optimum and specific activity are within the expected range of a hyperthermophilic organism and comparable to a similar enzyme assembled from *M. jannaschii* (Werner and Weinzierl 2002). We therefore conclude that the changes in

sequence, unusual as they may be, do not preclude catalytic activity. In a search for optimal assay conditions involving a sparse matrix approach, we discovered, however, an unexpected property: *ne*RNAP responded favourably to the presence of a high concentration of fluoride ions in the reaction buffer (optimal fluoride concentration for neRNAP ~200–300 mM). Reports in the research literature from the 1970s describe a similar stimulatory effect of fluoride on adenylate cyclase (Drummond et al., 1971; Stalmans and Hers, 1975). These biochemical analyses showed that the reaction velocity (V_{max}) of adenylate cyclase increased in the presence of fluoride but had no effect on the affinity (K_m) for substrate molecules. It later became apparent that it was a regulatory subunit that was the target of the fluoride stimulation, and not adenylate cyclase itself (Hebdon et al., 1978; Sahyoun et al., 1981). The identity of the regulatory protein turned out to be a subunit of a membrane bound, heterotrimeric G-protein complex. This G-protein is a gtpase and upon binding of GTP activates adenylate cyclase activity. The stimulatory effect of fluoride is believed to be the result of the ability of fluoride to form multi-fluorinated complexes with metal ions, such as Mg^{2+} (Antonny et al., 1993). Such "MgFx" complexes are capable of mimicking the γ-phosphate of a GTP molecule (Higashijima et al., 1991) and are thus able to occupy the phosphate binding pocket of the nucleotide-binding site of the G-protein. Several other G-protein dependent regulatory enzymes (such as Erk, Rho, Ras) have been shown to respond to fluoride in such a way (Bogatcheva et al., 2006). Fluoride has also been shown to bind to pyrophosphate (Baykov et al., 2000). We therefore hypothesize that the stimulation of transcription by high levels of fluoride ions may have a comparable cause in nanoarchaal RNAP. The presence of mono- or multi-fluorinated NTP complexes (see **Figure 6** for a GDP-based example) may assist with binding of NTPs to a structurally more flexible active site in *ne*RNAP and/or help to stabilize some transition complexes in the nucleotide addition cycle. It is possible that especially the binding of Mg^{2+} ions to the divergent Metal B motif could be influenced in such a manner. According to such a model, the observed lack of effect of fluoride on the catalytic activity *mj*RNAP would reflect the fact that "conventional" RNAPs do not require this kind of assistance for their catalytic sites to operate.

Future studies will focus on the potential interplay between fluoride, Mg^{2+} and NTPs, as well as defining in more detail which of the diverged motifs is most susceptible to this effect. By replacing some of the substitutions - either individually or in groups - with residues that are normally found in their position in other RNAPs, we will be able to study which of them are most likely to be responsible for this unusual behavior of *ne*RNAP in presence of fluoride.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

RW designed the project. SN carried out the experimental work and analyzed the data. The manuscript is a joint effort of SN and RW.

FUNDING

The project was funded by the Wellcome Trust (Grant 078043/Z/05/Z).

ACKNOWLEDGMENTS

We would like to thank Prof. Michael Thomm for providing us a sample of purified *N. equitans* genomic DNA.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.669314/full#supplementary-material

REFERENCES

Adam, P. S., Borrel, G., Brochier-Armanet, C., and Gribaldo, S. (2017). The Growing Tree of Archaea: New Perspectives on Their Diversity, Evolution and Ecology. *Isme J.* 11, 2407–2425. doi:10.1038/ismej.2017.122

Antonny, B., Sukumar, M., Bigay, J., Chabre, M., and Higashijima, T. (1993). The Mechanism of Aluminum-independent G-Protein Activation by Fluoride and Magnesium. 31P NMR Spectroscopy and Fluorescence Kinetic Studies. *J. Biol. Chem.* 268, 2393–2402. doi:10.1016/s0021-9258(18)53789-3

Baykov, A. A., Fabrichniy, I. P., Pohjanjoki, P., Zyryanov, A. B., and Lahti, R. (2000). Fluoride Effects along the Reaction Pathway of Pyrophosphatase: Evidence for a Second Enzyme·Pyrophosphate Intermediate†. *Biochemistry* 39, 11939–11947. doi:10.1021/bi000627u

Blombach, F., Matelska, D., Fouqueau, T., Cackett, G., and Werner, F. (2019). Key Concepts and Challenges in Archaeal Transcription. *J. Mol. Biol.* 431, 4184–4201. doi:10.1016/j.jmb.2019.06.020

Bogatcheva, N. V., Wang, P., Birukova, A. A., Verin, A. D., and Garcia, J. G. N. (2006). Mechanism of Fluoride-Induced MAP Kinase Activation in Pulmonary Artery Endothelial Cells. *Am. J. Physiol.-Lung Cell Mol. Physiol.* 290, L1139–L1145. doi:10.1152/ajplung.00161.2005

Boomsma, W., Frellsen, J., Harder, T., Bottaro, S., Johansson, K. E., Tian, P., et al. (2013). PHAISTOS: A Framework for Markov Chain Monte Carlo Simulation and Inference of Protein Structure. *J. Comput. Chem.* 34, 1697–1705. doi:10.1002/jcc.23292

Brochier, C., Gribaldo, S., Zivanovic, Y., Confalonieri, F., and Forterre, P. (2005). 'Nanoarchaea: Representatives of a Novel Archaeal Phylum or a Fast-Evolving Euryarchaeal Lineage Related to Thermococcales? *Genome Biol.* 6. doi:10.1186/gb-2005-6-5-r42

Chari, A., Haselbach, D., Kirves, J.-M., Ohmer, J., Paknia, E., Fischer, N., et al. (2015). ProteoPlex: Stability Optimization of Macromolecular Complexes by Sparse-Matrix Screening of Chemical Space. *Nat. Methods* 12, 859–865. doi:10.1038/nmeth.3493

Ciccarelli, F. D., Doerks, T., von Mering, C., Creevey, C. J., Snel, B., and Bork, P. (2006). Toward Automatic Reconstruction of a Highly Resolved Tree of Life. *Science* 311, 1283–1287. doi:10.1126/science.1123061

Coker, J. A. (2019). Recent Advances in Understanding Extremophiles. *F1000Res* 8, F1000 Faculty Rev-1917. doi:10.12688/f1000research.20765.1

Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001). Structural Basis of Transcription: RNA Polymerase II at 2.8 Angstrom Resolution. *Science* 292, 1863–1876. doi:10.1126/science.1059493

Drummond, G. I., Severson, D. L., and Duncan, L. (1971). Adenyl Cyclase. Kinetic Properties and Nature of Fluoride and Hormone Stimulation. *J. Biol. Chem.* 246, 4166–4173. doi:10.1016/s0021-9258(18)62068-x

Forterre, P., Gribaldo, S., and Brochier-Armanet, C. (2009). Happy Together: Genomic Insights into the Unique Nanoarchaeum/Ignicoccus Association. *J. Biol.* 8, 7. doi:10.1186/jbiol110

Fouqueau, T., Blombach, F., Cackett, G., Carty, A. E., Matelska, D. M., Ofer, S., et al. (2018). The Cutting Edge of Archaeal Transcription. *Emerg. Top. Life Sci.* 2, 517–533. doi:10.1042/etls20180014

Hebdon, M., Le Vine, H., 3rd, Sahyoun, N., Schmitges, C. J., and Cuatrecasas, P. (1978). Properties of the Interaction of Fluoride- and Guanylyl-5'-Imidodiphosphate-Regulatory Proteins with Adenylate Cyclase. *Proc. Natl. Acad. Sci.* 75, 3693–3697. doi:10.1073/pnas.75.8.3693

Higashijima, T., Graziano, M. P., Suga, H., Kainosho, M., and Gilman, A. G. (1991). 19F and 31P NMR Spectroscopy of G Protein Alpha Subunits. Mechanism of Activation by Al3+ and F-. *J. Biol. Chem.* 266, 3396–3401. doi:10.1016/s0021-9258(19)67806-3

Hirata, A., Klein, B. J., and Murakami, K. S. (2008). Erratum: The X-Ray Crystal Structure of RNA Polymerase from Archaea. *Nature* 452, 248. doi:10.1038/nature06844

Hofmann, A., Tai, M., Wong, W., and Glabe, C. G. (1995). A Sparse Matrix Screen to Establish Initial Conditions for Protein Renaturation. *Anal. Biochem.* 230, 8–15. doi:10.1006/abio.1995.1429

Hohn, M. J., Hedlund, B. P., and Huber, H. (2002). Detection of 16S rDNA Sequences Representing the Novel Phylum "Nanoarchaeota": Indication for a Wide Distribution in High Temperature Biotopes. *Syst. Appl. Microbiol.* 25, 551–554. doi:10.1078/07232020260517698

Huber, H., Hohn, M. J., Rachel, R., Fuchs, T., Wimmer, V. C., and Stetter, K. O. (2002). A New Phylum of Archaea Represented by a Nanosized Hyperthermophilic Symbiont. *Nature* 417, 63–67. doi:10.1038/417063a

Huber, H., Hohn, M. J., Stetter, K. O., and Rachel, R. (2003). The Phylum Nanoarchaeota: Present Knowledge and Future Perspectives of a Unique Form of Life. *Res. Microbiol.* 154, 165–171. doi:10.1016/s0923-2508(03)00035-4

Jancarik, J., and Kim, S. H. (1991). Sparse Matrix Sampling: a Screening Method for Crystallization of Proteins. *J. Appl. Cryst.* 24, 409–411. doi:10.1107/s0021889891004430

McCliment, E. A., Voglesonger, K. M., O'Day, P. A., Dunn, E. E., Holloway, J. R., and Cary, S. C. (2006). Colonization of Nascent, Deep-Sea Hydrothermal Vents by a Novel Archaeal and Nanoarchaeal Assemblage. *Environ. Microbiol.* 8, 114–125. doi:10.1111/j.1462-2920.2005.00874.x

Naji, S., Grünberg, S., and Thomm, M. (2007). The RPB7 Orthologue E′ Is Required for Transcriptional Activity of a Reconstituted Archaeal Core Enzyme at Low Temperatures and Stimulates Open Complex Formation. *J. Biol. Chem.* 282, 11047–11057. doi:10.1074/jbc.m611674200

Nottebaum, S., Tan, L., Trzaska, D., Carney, H. C., and Weinzierl, R. O. J. (2008). The RNA Polymerase Factory: a Robotic In Vitro Assembly Platform for High-Throughput Production of Recombinant Protein Complexes. *Nucleic Acids Res.* 36, 245–252. doi:10.1093/nar/gkm1044

Ouhammouch, M., Werner, F., Weinzierl, R. O. J., and Geiduschek, E. P. (2004). A Fully Recombinant System for Activator-dependent Archaeal Transcription. *J. Biol. Chem.* 279, 51719–51721. doi:10.1074/jbc.c400446200

Probst, A. J., Castelle, C. J., Singh, A., Brown, C. T., Anantharaman, K., Sharon, I., et al. (2017). Genomic Resolution of a Cold Subsurface Aquifer Community Provides Metabolic Insights for Novel Microbes Adapted to High CO2concentrations. *Environ. Microbiol.* 19, 459–474. doi:10.1111/1462-2920.13362

Randau, L., Münch, R., Hohn, M. J., Jahn, D., and Söll, D. (2005). Nanoarchaeum Equitans Creates Functional tRNAs from Separate Genes for Their 5′- and 3′-halves. *Nature* 433, 537–541. doi:10.1038/nature03233

Randau, L., Schröder, I., and Söll, D. (2008). Life without RNase P. *Nature* 453, 120–123. doi:10.1038/nature06833

Rawle, R. A., Hamerly, T., Tripet, B. P., Giannone, R. J., Wurch, L., Hettich, R. L., et al. (2017). Multi-omics Analysis Provides Insight to the Ignicoccus Hospitalis-Nanoarchaeum Equitans Association. *Biochim. Biophys. Acta (Bba) - Gen. Subjects* 1861, 2218–2227. doi:10.1016/j.bbagen.2017.06.001

Roe, D. R., and Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theor. Comput.* 9, 3084–3095. doi:10.1021/ct400341p

Sahyoun, N. E., LeVine, H., 3rd, Davis, J., Hebdon, G. M., and Cuatrecasas, P. (1981). Molecular Complexes Involved in the Regulation of Adenylate Cyclase. *Proc. Natl. Acad. Sci.* 78, 6158–6162. doi:10.1073/pnas.78.10.6158

Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V., and Goldfarb, A. (2003). Unified Two-Metal Mechanism of RNA Synthesis and Degradation by RNA Polymerase. *EMBO J.* 22, 2234–2244. doi:10.1093/emboj/cdg193

Stalmans, W., and Hers, H.-G. (1975). The Stimulation of Liver Phosphorylase B by AMP, Fluoride and Sulfate. A Technical Note on the Specific Determination of the a and B Forms of Liver Glycogen Phosphorylase. *Eur. J. Biochem.* 54, 341–350. doi:10.1111/j.1432-1033.1975.tb04144.x

Sullivan, S. S., and Weinzierl, R. O. J. (2020). Optimization of Molecular Dynamics Simulations of C-Myc(1-88)-An Intrinsically Disordered System. *Life (Basel)* 10 (7), 109. doi:10.3390/life10070109

Tan, L., Wiesler, S., Trzaska, D., Carney, H. C., and Weinzierl, R. O. (2008). Bridge Helix and Trigger Loop Perturbations Generate Superactive RNA Polymerases. *J. Biol.* 7, 40. doi:10.1186/jbiol98

Thomm, M., Reich, C., Grünberg, S., and Naji, S. (2009). Mutational Studies of Archaeal RNA Polymerase and Analysis of Hybrid RNA Polymerases. *Biochem. Soc. Trans.* 37, 18–22. doi:10.1042/bst0370018

Tully, B. J., Graham, E. D., and Heidelberg, J. F. (2018). The Reconstruction of 2,631 Draft Metagenome-Assembled Genomes from the Global Oceans. *Sci. Data* 5, 170203. doi:10.1038/sdata.2017.203

Waters, E., Hohn, M. J., Ahel, I., Graham, D. E., Adams, M. D., Barnstead, M., et al. (2003). The Genome of Nanoarchaeum Equitans: Insights into Early Archaeal Evolution and Derived Parasitism. *Proc. Natl. Acad. Sci.* 100, 12984–12988. doi:10.1073/pnas.1735403100

Weinzierl, R. O. J. (2010a). Nanomechanical Constraints Acting on the Catalytic Site of Cellular RNA Polymerases. *Biochem. Soc. Trans.* 38, 428–432. doi:10.1042/bst0380428

Weinzierl, R. O. J. (2010b). 'The Nucleotide Addition Cycle of RNA Polymerase Is Controlled by Two Molecular Hinges in the Bridge Helix Domain'. *BMC Biol.* 8, 134. doi:10.1186/1741-7007-8-134

Weinzierl, R. O. J. (2011). The Bridge Helix of RNA Polymerase Acts as a Central Nanomechanical Switchboard for Coordinating Catalysis and Substrate Movement. *Archaea* 2011, 608385. doi:10.1155/2011/608385

Weinzierl, R. O. J. (2013). The RNA Polymerase Factory and Archaeal Transcription. *Chem. Rev.* 113, 8350–8376. doi:10.1021/cr400148k

Wenck, B. R., and Santangelo, T. J. (2020). Archaeal Transcription. *Transcription* 11, 199–210. doi:10.1080/21541264.2020.1838865

Werner, F., and Weinzierl, R. O. J. (2002). A Recombinant RNA Polymerase II-like Enzyme Capable of Promoter-specific Transcription. *Mol. Cel* 10, 635–646. doi:10.1016/s1097-2765(02)00629-9

Werner, F., Wiesler, S., Nottebaum, S., and Weinzierl, R. O. J. (2006). Modulation of RNA Polymerase Core Functions by Basal Transcription Factor TFB/TFIIB. *Biochem. Soc. Symp.* 73, 49–58. doi:10.1042/bss0730049

Zhou, Z., Liu, Y., Xu, W., Pan, J., Luo, Z. H., and Li, M. (2020). Genome- and Community-Level Interaction Insights into Carbon Utilization and Element Cycling Functions of Hydrothermarchaeota in Hydrothermal Sediment. *mSystems* 5, e00795–19. doi:10.1128/msystems.00795-19

Conflict of Interest: SN is currently employed by Orthomol Pharmazeutische Vertriebs GmbH.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Molecular Biosciences

MINI REVIEW
published: 13 May 2021
doi: 10.3389/fmolb.2021.681550

Regulation of RNA Polymerase II Transcription Initiation and Elongation by Transcription Factor TFII-I

Niko Linzer[1†], Alexis Trumbull[1†], Rukiye Nar[1], Matthew D. Gibbons[1], David T. Yu[1], John Strouboulis[2] and Jörg Bungert[1]*

[1] Department of Biochemistry and Molecular Biology, College of Medicine, UF Health Cancer Center, Genetics Institute, Powell Gene Therapy Center, University of Florida, Gainesville, FL, United States, [2] Comprehensive Cancer Center, School of Cancer and Pharmaceutical Sciences, King's College London, United Kingdom

OPEN ACCESS

Edited by:
Athar Ansari, Wayne State University, United States

Reviewed by:
Ananda L. Roy, National Institutes of Health (NIH), United States
Mehdi Kabani, Centre National de la Recherche Scientifique, Center for the National Scientific Research (CNRS), France

***Correspondence:**
Jörg Bungert
jbungert@ufl.edu

[†] *These authors have contributed equally to this work*

Specialty section:
This article was submitted to Protein and RNA Networks, a section of the journal Frontiers in Molecular Biosciences

Received: *16 March 2021*
Accepted: *20 April 2021*
Published: *13 May 2021*

Citation:
Linzer N, Trumbull A, Nar R, Gibbons MD, Yu DT, Strouboulis J and Bungert J (2021) Regulation of RNA Polymerase II Transcription Initiation and Elongation by Transcription Factor TFII-I. Front. Mol. Biosci. 8:681550. doi: 10.3389/fmolb.2021.681550

Transcription by RNA polymerase II (Pol II) is regulated by different processes, including alterations in chromatin structure, interactions between distal regulatory elements and promoters, formation of transcription domains enriched for Pol II and co-regulators, and mechanisms involved in the initiation, elongation, and termination steps of transcription. Transcription factor TFII-I, originally identified as an initiator (INR)-binding protein, contains multiple protein–protein interaction domains and plays diverse roles in the regulation of transcription. Genome-wide analysis revealed that TFII-I associates with expressed as well as repressed genes. Consistently, TFII-I interacts with co-regulators that either positively or negatively regulate the transcription. Furthermore, TFII-I has been shown to regulate transcription pausing by interacting with proteins that promote or inhibit the elongation step of transcription. Changes in TFII-I expression in humans are associated with neurological and immunological diseases as well as cancer. Furthermore, TFII-I is essential for the development of mice and represents a barrier for the induction of pluripotency. Here, we review the known functions of TFII-I related to the regulation of Pol II transcription at the stages of initiation and elongation.

Keywords: RNA polymerase II, transcription regulation, TFII-I, GTF2I, transcription elongation

DISCOVERY OF TFII-I AS AN INITIATOR-BINDING PROTEIN

The discovery of the three eukaryotic RNA polymerases and the development of powerful *in vitro* techniques for the analysis of the transcription process initiated a large body of work that led to the identification of basal promoter elements and trans-acting proteins involved in initiating the transcription of protein-coding genes by RNA polymerase II (Pol II; Roeder and Rutter, 1969; Weil et al., 1979; Roeder, 2019; Schier and Taatjes, 2020). Earlier, most of the studies were performed using viral genes containing strong promoter elements that recruit the transcription machinery with high efficiency. A critical component in the initiation step of the transcription process by Pol II is the TFIID (transcription factor II D) complex, which is composed of the TATA-binding protein (TBP) and TBP-associated factors (TAFs; Patel et al., 2020). TATA-box-containing promoters are usually found

at developmentally regulated genes and characterized by the presence of a focused transcription start site (TSS), while TATA-less promoters are often found at housekeeping genes and exhibit transcription initiation over broad regions (Bhuiyan and Timmers, 2019). At TATA-box-containing promoter regions, TBP is sufficient for the reconstitution of basal transcription *in vitro*. Among the pioneering work on Pol II transcription was the discovery of the initiator element by Smale and Baltimore (1989). The initiator is a pyrimidine-rich DNA sequence that overlaps with the sequence of the TSS and was shown to be able to direct accurate transcription in the absence of a TATA box. Transcription factor TFII-I was one of the early proteins identified to interact with the initiator and to recruit transcription complexes to TATA-less promoters (Roy et al., 1993b). Subsequent studies have shown that components of the TFIID complex, including TAF1 and TAF2, interact with the initiator as well as with downstream promoter elements (DPEs), which were discovered by the Kadonaga laboratory (Burke and Kadonaga, 1996; Patel et al., 2018; Vo Ngoc et al., 2020). However, efficient transcription of TATA-less promoters cannot be reconstituted with TFIID and the other basal transcription factors alone, suggesting that additional components are essential for the initial recruitment of TFIID or stabilization of TFIID at TATA-less promoters. Furthermore, TFII-I may act to regulate the transcription of a specific set of genes *via* the initiator and/or in response to specific environmental signals *in vivo* (Roy, 2012).

STRUCTURE AND FUNCTION OF TFII-I

TFII-I is an unusual transcription factor consisting of a basic region (BR) DNA-binding domain, a nuclear localization sequence (NLS), and multiple protein–protein interaction domains, including a leucine zipper (LZ) and six helix-loop-helix (HLH)-like domains, also referred to as I-repeats (R1–R6, **Figure 1A**; Doi-Katayama et al., 2007; Roy, 2012). TFII-I has been shown to interact not only with initiator sequences and E-boxes (HLH-binding motif, CANNTG) but also with other sequences, including the serum response element (SRE) *in vitro* (Grueneberg et al., 1997; Roy, 2012). Some of these interactions are likely mediated by other transcription factors that are associated with TFII-I. For example, TFII-I interacts with the HLH- and E-box-binding proteins USF (upstream stimulatory factor) and Myc (myelocytomatosis; Roy et al., 1991, 1993a). While cooperative interactions between TFII-I and USF activate the transcription process, the interactions between TFII-I and Myc repress transcription at the adenovirus 2 major late promoter (Ad2MLP) region. These studies on the aforementioned interactions have already provided evidence that TFII-I can function as a transcription activator or repressor depending on the interacting partner protein(s).

Since these initial studies, subsequent work has demonstrated that TFII-I is a multifunctional protein that exerts activities in both cytoplasm and nucleus (Roy, 2012). In the cytoplasm, TFII-I inhibits agonist-induced calcium entry by the transient receptor potential cation channel subfamily C member 3 (TRPC-3; Caraveo et al., 2006). In the nucleus, TFII-I functions as an activator or repressor of gene expression and also plays a role in translesion DNA repair (Roy, 2012; Fattah et al., 2014). Furthermore, TFII-I shuttles between the cytoplasm and the nucleus in response to specific signals, which is regulated by tyrosine phosphorylation, and several studies have shown that TFII-I facilitates the nuclear import of transcription factors, including members of the nuclear factor NF-κb family (Ashworth and Roy, 2007; Roy, 2012).

Alternative splicing generates four isoforms of TFII-I (α, β, Δ, and γ), which are expressed in a ubiquitous or cell type-specific manner (Roy, 2012). The Δ-isoform is ubiquitously expressed and shuttles between the cytoplasm and the nucleus. The α- and β-isoforms contain additional but exclusive short exons upstream of the NLS. The γ-isoform contains both alternative exons found in the α- and β-isoforms, and is predominantly expressed in neuronal tissues. In addition to these isoforms, there are also TFII-I-related genes expressed in humans and mice (Roy, 2012). One of these genes, *GTF2IRD1*, is located in close proximity to the *GTF2I* gene and encodes BEN (binding factor of early enhancer) (Bayarsaihan and Ruddle, 2000). Williams–Beuren syndrome (WBS) is characterized by haploinsufficiency of a relatively large genomic region on human chromosome 7 encompassing *GTF2I* and *GTF2IRD1* genes (Pober, 2010). Individuals afflicted by this large genomic deletion exhibit craniofacial, cardiovascular, and neurologic defects. Some aspects of this compound genetic disease are recapitulated in mice lacking TFII-I or BEN (Tassabehji et al., 2005; Enkhmandakh et al., 2009).

TFII-I FUNCTION IN PROLIFERATING CELLS AND CANCER

TFII-I has been implicated in a variety of diseases, including neurological abnormalities. This has recently been reviewed by **Roy (2017)** and will not be repeated here. The role of TFII-I in the proliferation of cells and cancer will be briefly reviewed here as it relates to the multiple functions associated with this complex transcription factor. TFII-I regulates the proliferation of cells in response to serum and mitogenic signals (Roy, 2012). Phosphorylation of TFII-I by tyrosine kinases like Src (Sarcoma) and Bruton's tyrosine kinase (BTK) leads to nuclear translocation and activation of serum response genes, including the c-fos gene (Cheriyath et al., 2002). It has also been shown that TFII-I is sequestered in the cytoplasm by the p190 Rho GTPase-activating protein through the FF domain, which is characterized by the presence of two conserved phenylalanine (FF) residues (Jiang et al., 2005). Phosphorylation of the FF domain by Src in response to serum releases TFII-I and leads to the translocation of TFII-I to the nucleus and activation of serum response genes. One of the target genes of TFII-I in the nucleus is glucose-regulated protein 78 (GRP78), a protein chaperone involved in endoplasmic reticulum (ER) stress (Hong et al., 2005). GRP78 plays an essential role in the prosurvival machinery, and high-level expression of GRP78 is associated with drug resistance, carcinogenesis, and metastasis (Ibrahim et al., 2019). TFII-I also regulates genes involved in DNA repair and is directly involved

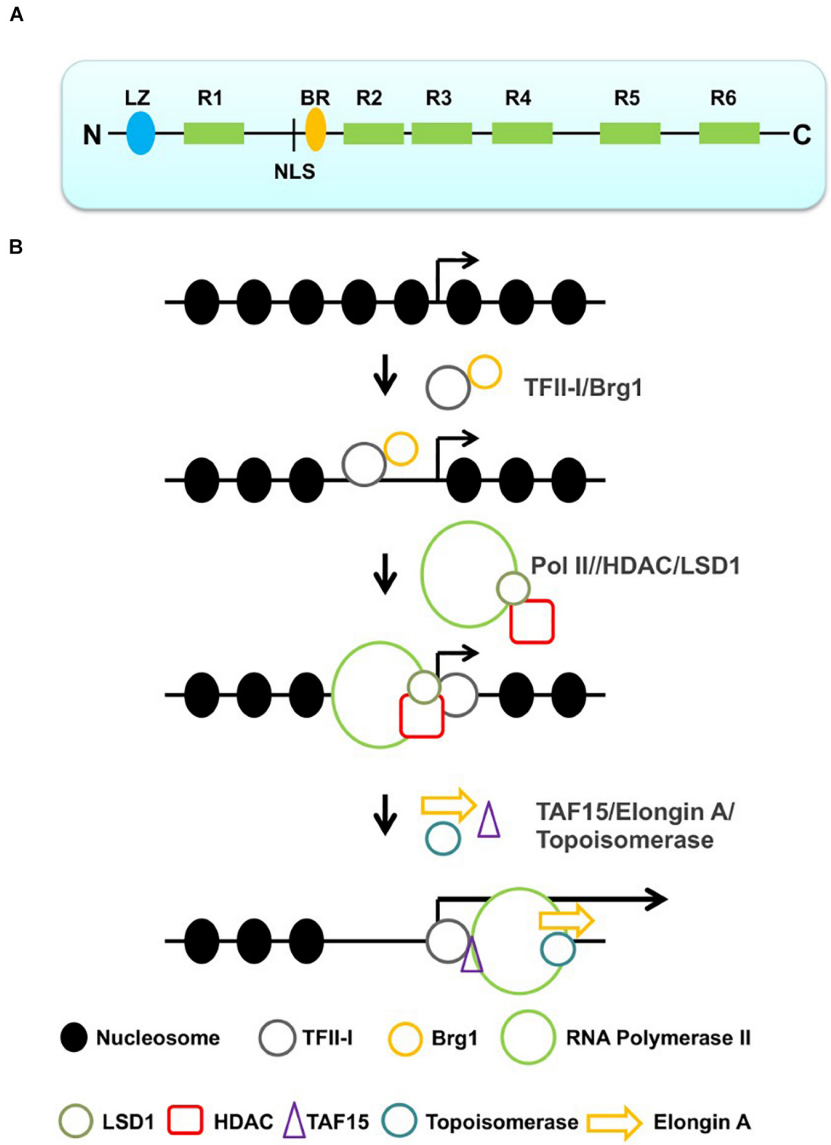


FIGURE 1 | Structure of TFII-I and sequential action of TFII-I leading to the recruitment of a productive Pol II transcription complex. **(A)** Structure of TFII-I (LZ, leucine zipper; R1–R6, I-repeats; NLS, nuclear localization sequence; BR, basic region). **(B)** TFII-I interacts with the chromatin remodeler Brg1 and establishes an accessible chromatin configuration at a specific promoter. Interactions with negative co-regulators (HDAC and LSD1) keep the promoter in an accessible but inactive configuration. Dissociation of the negative co-regulators and association with positive transcription elongation factors (TAF15, topoisomerase and Elongin A) converts Pol II into a productive elongation complex.

in DNA translesion repair (Roy, 2012; Fattah et al., 2014), thus leading to genome stability, which may be an important function during the proliferation of cells.

Mutations in TFII-I are associated with a number of different tumors, including T-cell lymphoma and thymus epithelial tumors (TETs; Radovich et al., 2018; Nathany et al., 2021). Point mutations in the TFII-I-coding region were found in about 6% of patients with angioimmunoblastic T-cell lymphomas (Vallois et al., 2016). A missense mutation (Leu404His) in TFII-I was found in a large number of type A and type AB thymomas (Petrini et al., 2014). This mutation was found to increase the expression of TFII-I, which may be due to the disruption of a potential destruction box (Oberndorfer and Müllauer, 2020). Previous studies have shown that TFII-I is subjected to ubiquitylation and proteasomal degradation in response to genotoxic stress (Desgranges et al., 2005). Increased expression of TFII-I in thymomas is consistent with its role in activating genes involved in proliferation (Roy, 2017). However, TFII-I was not found to be overexpressed in the most aggressive forms of thymomas (Petrini et al., 2014; Oberndorfer and Müllauer, 2020). This is interesting in light of the fact that previous studies have shown that TFII-I represents a roadblock in the generation of induced pluripotent

stem cells (IPSCs; Yang et al., 2014). This suggests that although TFII-I contributes to the proliferation of cells by modulating the expression of cell cycle genes, it prevents dedifferentiation of cells. This could be due to the fact that TFII-I also regulates genes that constitute cell identity.

TFII-I AND ESTABLISHMENT OF CHROMATIN DOMAINS

Transcription factors impact gene expression in many different ways. They can bind to and act in close proximity to the genes they regulate, or they act at a distance by binding to enhancer or insulator sequences. Recent advances in our understanding of the organization of genes within the nucleus demonstrate that the genome is organized in defined topologically associating domains (TADs) that are often multiple Mb-long and are characterized by frequent chromosomal interactions within TADs and limited contacts between TADs (Sun et al., 2019). TADs are separated from each other by boundary elements that interact with CCCTC-binding protein (CTCF) and cohesin and/or condensin complexes. Within TADs, genes are organized in insulated neighborhoods in which one or several, sometimes co-regulated, genes and corresponding enhancer elements are localized. At least a subset of insulated neighborhoods is established by dimerization of CTCF proteins that interact with insulator sequences flanking the neighborhoods (Luo et al., 2020).

The genome-wide analysis of TFII-I-chromatin interactions revealed that TFII-I associates with active and with repressed genes (Makeyev et al., 2012; Fan et al., 2014). While it is evident that TFII-I regulates genes directly *via* interactions with specific DNA elements, only 8% of genomic binding sites for TFII-I in IPSCs correspond to nearby genes that change expression in response to TFII-I depletion (Makeyev et al., 2012). Furthermore, a vast majority of genes that changed expression upon TFII-I depletion did not contain binding sites for TFII-I, suggesting they are regulated indirectly. TFII-I interacts with CTCF and often associates with genomic sites occupied by CTCF (Peña-Hernández et al., 2015). Furthermore, TFII-I peaks also overlap with peaks for Rad21, a component of the cohesin complex, and pull-down experiments identified subunits of cohesin and condensin [structural maintenance of chromosomes (SMC) 2, 3, and 6] as interaction partners of TFII-I in human erythroleukemia K562 cells (Fan et al., 2014; Kim et al., 2015). These data point to the possibility that TFII-I exerts part of its function by assisting CTCF and cohesins in the establishment of TADs and/or insulated neighborhoods.

TFII-I REPRESSES AND ACTIVATES TRANSCRIPTION BY POL II

Pull-down experiments identified chromatin remodeling complexes, particularly Brg1, histone deacetylases (HDACs), lysine-specific demethylase 1 (LSD1), topoisomerases, and transcription elongation factors as TFII-I-interacting proteins, suggesting diverse functions of TFII-I during the regulation of transcription (Fan et al., 2014; Adamo et al., 2015). LSD1 removes methyl groups from H3K4, and H3K4 methylation is associated with transcriptionally active or permissive chromatin (Meier and Brehm, 2014). Thus, TFII-I likely inhibits transcription through interactions with HDACs and LSD1. Genome-wide TFII-I peaks are often associated with binding sites for related (e.g., USF) or unrelated transcription factors (e.g., E2F and CTCF) (Makeyev et al., 2012; Fan et al., 2014; Peña-Hernández et al., 2015). Consistent with these findings, TFII-I interacts with E-box-binding proteins (e.g., USF and cMyc) as well as with E2F transcription factors and CTCF (Roy et al., 1991, 1993a; Fan et al., 2014; Peña-Hernández et al., 2015; Shen et al., 2018). Thus, a large fraction of TFII-I-binding events in the context of chromatin may be mediated by other DNA-binding transcription factors rather than direct interactions of TFII-I with DNA.

As mentioned before, TFII-I interacts with negative and positive co-regulators. At some gene loci, perhaps at those involved in stress response or in cell cycle control, TFII-I may play negative and positive roles at different stages of induction (**Figure 1B**). It is conceivable that TFII-I recruits the Brg1 chromatin remodeling complex to these genomic loci and establishes short regions of accessibility. These regions may further associate with other repressor or co-repressor proteins, e.g., repressor E2Fs, HDACs, and LSD1. Binding of these components will keep regulatory regions in an accessible but inactive configuration. Upon specific signals, e.g., growth factors, stress, or cell cycle progression, the inhibitory proteins leave the promoter from the DNA and TFII-I recruits positive factors that mediate the recruitment of Pol II or stimulate the elongation step of transcription. The poised state may also involve a paused RNA polymerase, which is outlined in the next section.

TFII-I REGULATES THE TRANSITION FROM TRANSCRIPTION PAUSING TO ELONGATION

At a subset of genes, Pol II pauses near the TSS and several activities have been identified to mediate the transition from pausing to productive elongation (Gonzales et al., 2021). At mRNA genes, Pol II pauses to allow capping of the 5′end of the RNA. This pausing is mediated by negative elongation factor (NELF) and 5,6-dichloro-1-β-d-ribofuranosylbenzimidazole (DRB) sensitivity-inducing factor (DSIF; Yamaguchi et al., 2013; Schier and Taatjes, 2020). DRB is a nucleoside homolog that inhibits the elongation step of transcription. DSIF interacts with the initially transcribed RNA and recruits NELF. Pol II consists of a relatively unstructured C-terminal domain (CTD) that contains a repeated heptapeptide sequence (Harlen and Churchman, 2017). Within the repeated heptapeptide sequence, there are several serine (S) residues that are subject to phosphorylation during transcription (Harlen and Churchman, 2017; Schier and Taatjes, 2020). Upon the initiation of transcription, the basal transcription factor TFII-H phosphorylates S5 (Schier and Taatjes, 2020). This phosphorylation event disrupts interactions with basal

transcription factors and the mediator–co-activator complex and promotes interactions with DSIF, NELF, and the capping complex. Interactions of NELF with Pol II prevent the association with positive elongation factor TFIIS and with the RNA polymerase-associated factor (PAF) complex (Vos et al., 2018). The binding of NELF also leads to an inactive conformation of Pol II that prevents translocations and base pairing of nucleotides in the active site. After capping, positive transcription elongation factor b (pTEFb) phosphorylates DSIF, NELF, and the CTD residue S2 (Yamaguchi et al., 2013; Schier and Taatjes, 2020). S2P assists in recruiting the PAF complex as well as RNA-processing factors. Phosphorylation of DSIF converts it from a negative to a positive elongation factor. Phosphorylation of NELF causes its dissociation from the transcription complex allowing interactions of Pol II with PAF and TFIIS, and transitioning from the paused to the elongation-competent form (Yamaguchi et al., 2013; Vos et al., 2018; Schier and Taatjes, 2020).

The Elongin complex has been shown to regulate Pol II transcription elongation activity (Conaway and Conaway, 1999). Studies by the Conaway laboratory demonstrated that Elongin A associates with genes at regions occupied by S5P-modified Pol II (Kawauchi et al., 2013). Furthermore, *in vitro* studies demonstrated that Elongin A stimulates the elongation step of transcription (Conaway and Conaway, 1999). Recent genome-wide analysis of Elongin A-deficient cells did not reveal strong defects in overall Pol II transcription elongation rates but showed increased accumulation of Pol II at TSSs, suggesting that Elongin A regulates the transition from pause to transcription elongation (Ardehali et al., 2020; Wang et al., 2020). Interestingly, the genome-wide occupancy data suggest that Elongin A is preferentially recruited to sites upstream of the TSS and to enhancer elements. This could indicate that Elongin A is recruited by sequence-specific transcription activators that bind promoters and/or enhancers. Furthermore, RNA-seq data show that Elongin A deficiency only affects the expression of a small set of genes (Ardehali et al., 2020; Wang et al., 2020).

TFII-I has been shown to interact with NELF as well as Elongin A (Fan et al., 2014; McCleary-Wheeler et al., 2020). It appears that these interactions play a role in the inducible expression of specific genes. For example, both Elongin A and TFII-I are important for maximal stress-dependent induction of the activating transcription factor 3 (ATF3) gene (Fan et al., 2014). ATF3 is induced in response to a variety of cellular stress signals, including ER stress (Ku and Cheng, 2020). TFII-I has previously been implicated in gene regulation following ER stress (Parker et al., 2001). Elongin A was shown to interact with the transcribed region of ATF3, and this interaction is increased upon the induction of ER stress (Fan et al., 2014). Pol II peaks are associated with a putative enhancer element located far upstream of the gene and with the promoter. TFII-I interacts immediately downstream of the Pol II peak at the enhancer element. Upon stress, increased transcription is not only observed at the promoter but also downstream of the enhancer (Fan et al., 2014). Importantly, increased transcription is associated with enhanced recruitment of Elongin A to the ATF3 promoter region. TFII-I also interacts with topoisomerases that remove torsional stress of the DNA during the elongation step of transcription. These data suggest that TFII-plays a positive role in recruiting and/or modulating the activity of positive transcription elongation factors.

As mentioned before, TFII-I was shown to interact with the insulator protein CTCF and to regulate genes in response to metabolic stress (Peña-Hernández et al., 2015). This is interesting in light of previous studies showing that CTCF, in addition to serving as an insulator-binding protein (IBP), regulates transcriptional pausing (Shukla et al., 2011; Herrera Paredes et al., 2013). Binding of CTCF to the proximal promoter has been shown to increase the pausing index. However, CTCF has also been implicated in the positive regulation of the elongation step of transcription by mediating the recruitment of pTEFb (Laitem et al., 2015). Ablating TFII-I expression led to a reduction of CTCF binding at specific promoters concomitant with the reduced expression of these genes (Marques et al., 2014; Peña-Hernández et al., 2015). Moreover, TFII-I deficiency was not only associated with reduced CTCF binding but also with an impaired recruitment of CDK8 and a reduction of Pol II S5P at CTCF target genes (Marques et al., 2014; Peña-Hernández et al., 2015). Thus, at certain genes, TFII-I cooperates with CTCF in mediating transcription likely by modulating early Pol II transcription initiation events. In *Drosophila*, other IBPs have also been shown to modulate Pol II pausing at distantly located genes. This is mediated by their common cofactor CP190 (Liang et al., 2014).

The above-discussed studies implicate TFII-I in the positive regulation of the elongation step of transcription. A recent study implicates TFII-I in the negative regulation of genes induced by TGF-β (McCleary-Wheeler et al., 2020). At a subset of TGF-β-induced genes, Pol II is paused downstream of the TSS (**Figure 2**). The paused Pol II is associated with NELF and DSIF. TFII-I was shown to bind at the TSS of these genes and to interact with NELF and DSIF. The authors propose that small mothers against decapentaplegic 3 (SMAD3), induced by transforming growth factor β (TGF-β), displaces TFII-I from the TSS, thus dissociating NELF and converting Pol II into an elongation-competent form. It is an intriguing idea that TFII-I may interact with TSSs after Pol II initiates transcription to regulate the pausing step. This is consistent with the fact that TFIID is the major protein complex recruiting Pol II to basal promoter elements, and reinforces the idea that TFII-I may regulate transcription at a step post Pol II recruitment.

Recently, it was shown that NELF forms nuclear condensates in response to stress (Rawat et al., 2021). Formation of nuclear condensates is dependent on the presence of an intrinsically disordered region (IDR) and is driven by dephosphorylation and sumoylation of NELF. This process increases the recruitment of NELF to promoters and causes transcription repression. TFII-I has previously been shown to interact with TAF15, one of the FET (Fus/EWS/TAF15) proteins (Fan et al., 2014). FET proteins contain IDRs that drive phase separation (Wang et al., 2018). Loci-specific phase separation mediated by Fet proteins recruits RNA Pol II to promoters and activates transcription (Wei et al.,

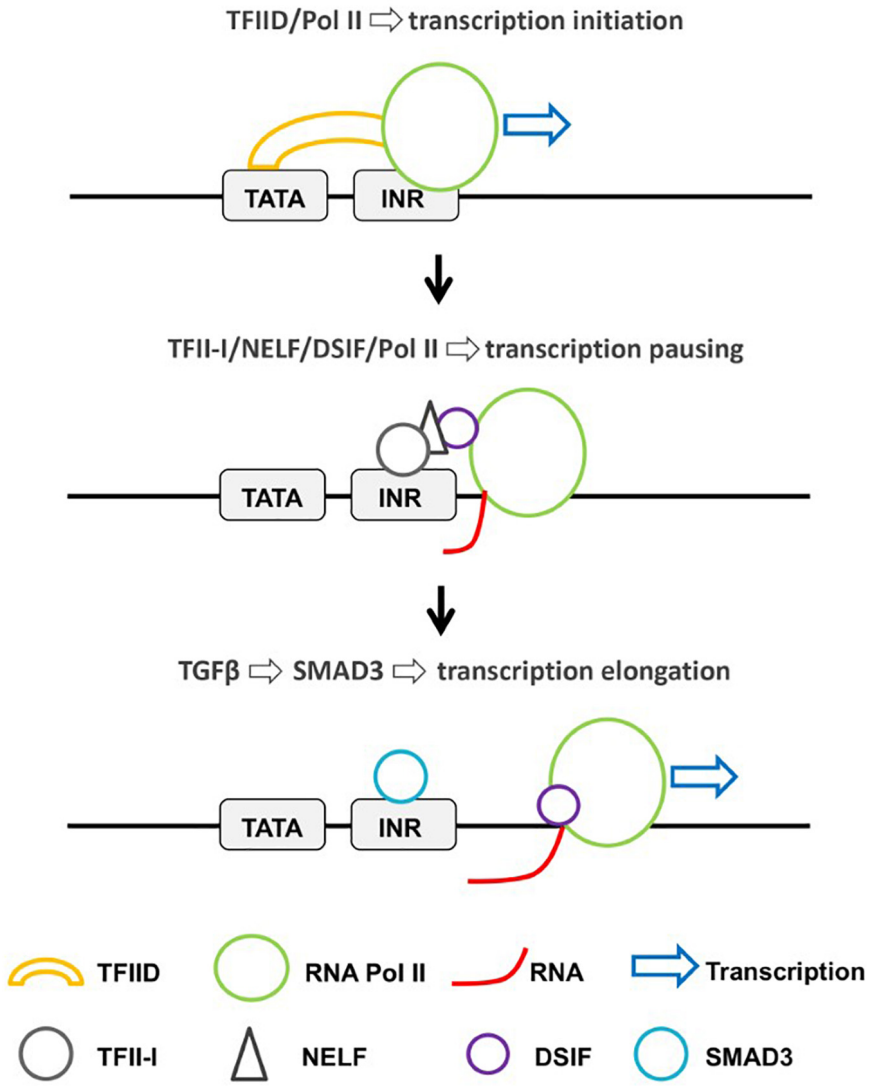


FIGURE 2 | Regulation of Pol II recruitment and transcriptional pausing by TFFI-I. At a subset of TGF-β-inducible genes, TFII-I interacts with NELF and DSIF at a post-initiation step and prevents the release of Pol II from the paused state (McCleary-Wheeler et al., 2020). TGF-β signaling increases the nuclear localization of SMAD3, which displaces TFII-I and converts Pol II into a productive elongation complex.

2020; Zuo et al., 2021). Thus, TFII-I may regulate transcription negatively through NELF-driven phase separation, or positively through TAF15-driven phase separation.

CONCLUSION AND OUTLOOK

Because of its unique structure, its multiple functions in the nucleus and cytoplasm, and its implication in cancer as well as neurological and immunological disorders, there is interest and a significant requirement to elucidate the mechanisms by which TFII-I affects gene expression patterns and other cellular functions during development and differentiation. It is clear that TFII-I is a DNA-binding protein that interacts with co-regulators to positively or negatively affect the transcription of specific target genes, but many aspects of TFII-I function remain enigmatic. It seems that a fraction of TFII-I chromatin associations is mediated by interactions with other DNA-binding proteins, including associations with E-box sequences together with HLH proteins, associations with E2F sites together with E2F transcription factors, and associations with insulator sequences together with CTCF and perhaps components of cohesin. In addition, TFII-I interacts with a variety of proteins that regulate different steps in the process of transcription, including histone-modifying enzymes, topoisomerases, and transcription elongation factors. Due to its relatively large size and the presence of multiple protein–protein interaction domains, it is possible that TFII-I functions as a hub to regulate the coordinated recruitment of activities involved in gene regulation. Its putative involvement in

the creation of chromatin domains and insulated chromatin neighborhoods will be an exciting focus of future research. Furthermore, the recent findings implicating TFII-I in regulating the elongation step of transcription suggest that it will be important to determine its function in response to stress or other signals. As mentioned before, TFII-I has been shown to interact with TAF15, which is capable of forming phase-separated domains (Fan et al., 2014; Wang et al., 2018). Recent evidence suggests that the transcription of highly expressed genes is regulated by phase-separated super-enhancers (Ishov et al., 2020). The elongation step of transcription occurs away from these domains and toward RNA processing domains. It will be interesting to investigate if TFII-I is involved in the formation of phase-separated transcription initiation domains and/or in associations of specific genes with RNA processing compartments.

AUTHOR CONTRIBUTIONS

JB, JS, RN, NL, and AT discussed the contents and outline of the review. All authors contributed to the writing and generation of figures.

FUNDING

This work was funded by a grant from the American Society of Hematology and by matching funds from the UF College of Medicine (JB).

ACKNOWLEDGMENTS

We thank the past and present colleagues in the Bungert and Strouboulis laboratory for stimulating discussions.

REFERENCES

Adamo, A., Atashpaz, S., Germain, P. L., Zanella, M., D'Agostino, G., Albertin, V., et al. (2015). 7q11.23 dosage-dependent dysregulation in human pluripotent stem cells affects transcriptional programs in disease-relevant lineages. *Nat. Genet.* 47, 132–141. doi: 10.1038/ng.3169

Ardehali, M. B., Damle, M., Perea-Resa, C., Blower, M. D., and Kingston, R. E. (2020). Elongin A associates with actively transcribed genes and modulates enhancer RNA levels with limited impact on transcription elongation rate *in vivo*. *J. Biol. Chem.* 296:100202. doi: 10.1074/jbc.RA120.015877

Ashworth, T., and Roy, A. L. (2007). Cutting Edge: TFII-I controls B cell proliferation via regulating NF-kappaB. *J. Immunol.* 178, 2631–2635. doi: 10.4049/jimmunol.178.5.2631

Bayarsaihan, D., and Ruddle, F. H. (2000). Isolation and characterization of BEN, a member of the TFII-I family of DNA-binding proteins containing distinct helix-loop-helix domains. *Proc. Natl. Acad. Sci. U.S.A.* 97, 7342–7347. doi: 10.1073/pnas.97.13.7342

Bhuiyan, T., and Timmers, H. T. M. (2019). Promoter recognition: putting TFIID on the spot. *Trends Cell Biol.* 29, 752–763. doi: 10.1016/j.tcb.2019.06.004

Burke, T. W., and Kadonaga, J. T. (1996). Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes Dev.* 10, 711–724. doi: 10.1101/gad.10.6.711

Caraveo, G., Van Rossum, D. B., Patterson, R. L., Snyder, S. H., and Desiderio, S. (2006). Action of TFII-I outside the nucleus as an inhibitor of agonist-induced calcium entry. *Science* 314, 122–125. doi: 10.1126/science.1127815

Cheriyath, V., Desgranges, Z. P., and Roy, A. L. (2002). c-Src-dependent transcriptional activation of TFII-I. *J. Biol. Chem.* 277, 22798–22805. doi: 10.1074/jbc.M202956200

Conaway, J. W., and Conaway, R. C. (1999). Transcription elongation and human disease. *Annu. Rev. Biochem.* 68, 301–319. doi: 10.1146/annurev.biochem.68.1.301

Desgranges, Z. P., Ahn, J., Lazebnik, M. B., Ashworth, T., Lee, C., Pestell, R. C., et al. (2005). Inhibition of TFII-I-dependent cell cycle regulation by p53. *Mol. Cell. Biol.* 25, 10940–10952. doi: 10.1128/MCB.25.24.10940-10952.2005

Doi-Katayama, Y., Hayashi, F., Inoue, M., Yabuki, T., Aoki, M., Seki, E., et al. (2007). Solution structure of the general transcription factor 2I domain in mouse TFII-I protein. *Protein Sci.* 16, 1788–1792. doi: 10.1110/ps.072792007

Enkhmandakh, B., Makeyev, A. V., Erdenechimeg, L., Ruddle, F. H., Chimge, N. O., Tussie-Luna, M. I., et al. (2009). Essential functions of the Williams-Beuren syndrome-associated TFII-I genes in embryonic development. *Proc. Natl. Acad. Sci. U.S.A.* 106, 181–186. doi: 10.1073/pnas.0811531106

Fan, A. X., Papadopoulos, G. L., Hossain, M. A., Lin, I. J., Hu, J., Tang, T. M., et al. (2014). Genomic and proteomic analysis of transcription factor TFII-I reveals insight into the response to cellular stress. *Nucl. Acids Res.* 42, 7625–7641. doi: 10.1093/nar/gku467

Fattah, F. J., Hara, K., Fattah, K. R., Yang, C., Wu, N., Warrington, R., et al. (2014). The transcription factor TFII-I promotes DNA translesion synthesis and genomic stability. *PLoS Genet.* 10:e1004419. doi: 10.1371/journal.pgen.1004419

Gonzales, N. M., Blears, D., and Svejstrup, J. Q. (2021). Causes and consequences of RNA polymerase II stalling during transcript elongation. *Nat. Rev. Mol. Cell Biol.* 22, 3–21. doi: 10.1038/s41580-020-00308-8

Grueneberg, D. A., Henry, R. W., Brauer, A., Novina, C. D., Cheriyath, V., Roy, A. L., et al. (1997). A multifunctional DNA-binding protein that promotes the formation of serum response factor/homeodomain complexes: identity to TFII-I. *Genes Dev.* 11, 2482–2493. doi: 10.1101/gad.11.19.2482

Harlen, K. M., and Churchman, L. S. (2017). The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nat. Rev. Mol. Cell Biol.* 18, 263–273. doi: 10.1038/nrm.2017.10

Herrera Paredes, S., Melgar, M. F., and Sethupathy, P. (2013). Promoter-proximal CCCTC-factor binding is associated with an increase in the transcriptional pausing index. *Bioinformatics* 29, 1485–1487. doi: 10.1093/bioinformatics/bts596

Hong, M., Lin, M. Y., Huang, J. M., Baumeister, P., Hakre, S., Roy, A. L., et al. (2005). Transcriptional regulation of the Grp78 promoter by endoplasmic reticulum stress: role of TFII-I and its tyrosine phosphorylation. *J. Biol. Chem.* 280, 16821–16828. doi: 10.1074/jbc.M413753200

Ibrahim, I. M., Abdelmalek, D. H., and Elfiky, A. A. (2019). GRP78: a cell's response to stress. *Life Sci.* 226, 156–163. doi: 10.1016/j.lfs.2019.04.022

Ishov, A. M., Gurumurthy, A., and Bungert, J. (2020). Coordination of transcription, processing, and export of highly expressed RNAs by distinct biomolecular condensates. *Emerg. Top. Life Sci.* 4, 281–291.

Jiang, W., Sordella, R., Chen, G. C., Hakre, S., Roy, A. L., and Settleman, J. (2005). An FF domain-dependent protein interaction mediates a signaling pathway for growth factor-induced gene expression. *Mol. Cell* 17, 23–35. doi: 10.1016/j.molcel.2004.11.024

Kawauchi, J., Inoue, M., Fukuda, M., Uchida, Y., Yasukawa, T., Conaway, R. C., et al. (2013). Transcriptional properties of mammalian elongin A and its role in stress response. *J. Biol. Chem.* 288, 24302–24315. doi: 10.1074/jbc.M113.496703

Kim, R., Smith, O. K., Wong, W. C., Ryan, A. M., Ryan, M. C., and Aladjem, M. I. (2015). ColoWeb: a resource for analysis of colocalization of genomic features. *BMC Genomics* 16:142. doi: 10.1186/s12864-015-1345-3

Ku, H. C., and Cheng, C. F. (2020). Master regulator activating transcription factor 3 (ATF3) in metabolic homeostasis and cancer. *Front. Endocrinol.* 11:556. doi: 10.3389/fendo.2020.00556

Laitem, C., Zaborowska, J., Tellier, M., Yamaguchi, Y., Cao, Q., Egloff, S., et al. (2015). CTCF regulates NELF, DSIF, and p-TEFb recruitment during transcription. *Transcription* 6, 79–90. doi: 10.1080/21541264.2015.1095269

Liang, J., Lacroix, L., Gamot, A., Cuddapah, S., Queille, S., Lhoumaud, P., et al. (2014). Chromatin immunoprecipitation indirect peaks highlight long-range interactions of insulator proteins and Pol II pausing. *Mol. Cell* 53, 672–681. doi: 10.1016/j.molcel.2013.12.029

Luo, H., Yu, Q., Liu, Y., Tang, M., Liang, M., Zhang, D., et al. (2020). LATS kinase-mediated CTCF phosphorylation and selective loss of genomic binding. *Sci. Adv.* 6:eaaw4651. doi: 10.1126/sciadv.aaw4651

Makeyev, A. V., Enkhmandakh, B., Hong, S. H., Joshi, P., Shin, D. G., and Bayarsaihan, D. (2012). Diversity and complexity in chromatin recognition by TFII-I transcription factors in pluripotent embryonic stem cells and embryonic tissues. *PLoS One* 7:e44443. doi: 10.1371/journal.pone.0044443

Marques, M., Hernandez, R. P., and Witcher, M. (2014). Analysis of changes to mRNA levels and CTCF occupancy upon TFII-I knockdown. *Genom. Data* 4, 17–21. doi: 10.1016/j.gdata.2014.09.012

McCleary-Wheeler, A. L., Paradise, B. D., Almada, L. L., Carlson, A. J., Marks, D. L., Vrabel, A., et al. (2020). TFII-I-mediated polymerase pausing antagonizes GLI2 induction by TGFβ. *Nucl. Acids Res.* 48, 7169–7181. doi: 10.1093/nar/gkaa476

Meier, K., and Brehm, A. (2014). Chromatin regulation: how complex does it get? *Epigenetics* 9, 1485–1495. doi: 10.4161/15592294.2014.971580

Nathany, S., Tripathi, R., and Mehta, A. (2021). Gene of the month: *GTF2I*. *J. Clin. Pathol.* 74, 1–4. doi: 10.1136/jclinpath-2020-207013

Oberndorfer, F., and Müllauer, L. (2020). Genomic alterations in thymoma-molecular pathogenesis? *J. Thorac. Dis.* 12, 7536–7544. doi: 10.21037/jtd.2019.12.52

Parker, R., Phan, T., Baumeister, P., Roy, B., Cheriyath, V., Roy, A. L., et al. (2001). Identification of TFII-I as the endoplasmic reticulum stress response element binding factor ERSF: its autoregulation by stress and interaction with ATF6. *Mol. Cell. Biol.* 21, 3220–3233. doi: 10.1128/MCB.21.9.3220-3233.2001

Patel, A. B., Greber, B. J., and Nogales, E. (2020). Recent insights into the structure of TFIID, its assembly, and its binding to core promoter. *Curr. Opin. Struct. Biol.* 61, 17–24. doi: 10.1016/j.sbi.2019.10.001

Patel, A. B., Louder, R. K., Greber, B. J., Grünberg, S., Luo, J., Fang, J., et al. (2018). Structure of human TFIID and mechanism of TBP loading onto promoter DNA. *Science* 362:eaau8872. doi: 10.1126/science.aau8872

Peña-Hernández, R., Marques, M., Hilmi, K., Zhao, T., Saad, A., Alaoui-Jamali, M. A., et al. (2015). Genome-wide targeting of the epigenetic regulatory protein CTCF to gene promoters by the transcription factor TFII-I. *Proc. Natl. Acad. Sci. U.S.A.* 112, E677–E686. doi: 10.1073/pnas.1416674112

Petrini, I., Meltzer, P. S., Kim, I. K., Lucchi, M., Park, K. S., Fontanini, G., et al. (2014). A specific missense mutation in GTF2I occurs at high frequency in thymic epithelial tumors. *Nat. Genet.* 46, 844–849.

Pober, B. R. (2010). Williams-Beuren syndrome. *N. Engl. J. Med.* 362, 239–252. doi: 10.1056/NEJMra0903074

Radovich, M., Pickering, C. R., Felau, I., Ha, G., Zhang, H., Jo, H., et al. (2018). The integrated genomic landscape of thymic epithelial tumors. *Cancer Cell* 33, 244–258. doi: 10.1016/j.ccell.2018.01.003

Rawat, P., Boehning, M., Hummel, B., Aprile-Garcia, F., Pandit, A. S., Eisenhardt, N., et al. (2021). Stress-induced nuclear condensation of NELF drives transcriptional downregulation. *Mol. Cell* 81, 1013–1026. doi: 10.1016/j.molcel.2021.01.016

Roeder, R. G. (2019). 50+ years of eukaryotic transcription: an expanding universe of factors and mechanisms. *Nat. Struct. Mol. Biol.* 26, 783–791. doi: 10.1038/s41594-019-0287-x

Roeder, R. G., and Rutter, W. J. (1969). Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* 224, 234–237. doi: 10.1038/224234a0

Roy, A. L. (2012). Biochemistry and biology of the inducible multifunctional transcription factor TFII-I: 10 years later. *Gene* 492, 32–41. doi: 10.1016/j.gene.2011.10.030

Roy, A. L. (2017). Pathophysiology of TFII-I: old guard wearing new hats. *Trends Mol. Med.* 23, 501–511. doi: 10.1016/j.molmed.2017.04.002

Roy, A. L., Carruthers, C., Gutjahr, T., and Roeder, R. G. (1993a). Direct role for Myc in transcription initiation mediated by interactions with TFII-I. *Nature* 365, 359–361. doi: 10.1038/365359a0

Roy, A. L., Malik, S., Meisterernst, M., and Roeder, R. G. (1993b). An alternative pathway for transcription initiation involving TFII-I. *Nature* 365, 355–359. doi: 10.1038/365355a0

Roy, A. L., Meisterernst, M., Pognonec, P., and Roeder, R. G. (1991). Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF. *Nature* 354, 245–248. doi: 10.1038/354245a0

Schier, A. C., and Taatjes, D. J. (2020). Structure and mechanism of the RNA polymerase II transcription machinery. *Genes Dev.* 34, 465–488. doi: 10.1101/gad.335679.119

Shen, Y., Nar, R., Fan, A. X., Aryan, M., Hossain, M. A., Gurumurthy, A., et al. (2018). Functional interrelationship between TFII-I and E2F transcription factors at specific cell cycle gene loci. *J. Cell. Biochem.* 119, 712–722. doi: 10.1002/jcb.26235

Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., et al. (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* 479, 74–79. doi: 10.1038/nature10442

Smale, S. T., and Baltimore, D. (1989). The "initiator" as a transcription control element. *Cell* 57, 103–113. doi: 10.1016/0092-8674(89)90176-1

Sun, F., Chronis, C., Kronenberg, M., Chen, X. F., Su, T., Lay, F. D., et al. (2019). Promoter-enhancer communication occurs primarily within insulated neighborhoods. *Mol. Cell* 73, 250–263. doi: 10.1016/j.molcel.2018.10.039

Tassabehji, M., Hammond, P., Karmiloff-Smith, A., Thompson, P., Thorgeirsson, S. S., Durkin, M. E., et al. (2005). GTFIRD1 in craniofacial development of humans and mice. *Science* 310, 1184–1187. doi: 10.1126/science.1116142

Vallois, D., Dobay, M. P., Morin, R. D., Lemonnier, F., Missiaglia, E., Juilland, M., et al. (2016). Activating mutations in genes related to TCR signaling in angioimmunoblastic and other follicular helper T-cell-derived lymphomas. *Blood* 128, 1490–1502. doi: 10.1182/blood-2016-02-698977

Vo Ngoc, L., Huang, C. V., Cassidy, C. J., Medrano, C., and Kadonaga, J. T. (2020). Identification of the human DPR core promoter element using machine learning. *Nature* 585, 459–463. doi: 10.1038/s41586-020-2689-7

Vos, S. M., Farnung, L., Urlaub, H., and Cramer, P. (2018). Structure of paused transcription complex Pol II-DSIF-NELF. *Nature* 560, 601–606. doi: 10.1038/s41586-018-0442-2

Wang, J., Choi, J. M., Holehouse, A. S., Lee, H. O., Zhang, X., Jahnel, M., et al. (2018). A molecular grammar governing the driving forces for phase separation of prion-like RNA binding protein. *Cell* 174, 688–699. doi: 10.1016/j.cell.2018.06.006

Wang, Y., Hou, L., Ardehali, M. B., Kingston, R. E., and Dynlacht, B. D. (2020). Elongin A regulates transcription *in vivo* through enhanced RNA polymerase processivity. *J. Biol. Chem.* 296:100170. doi: 10.1074/jbc.RA120.015876

Wei, M. T., Chang, Y. C., Shimobayashi, S. F., Shin, Y., Strom, A. R., and Brangwynne, C. P. (2020). Nucleated transcriptional condensates amplify gene expression. *Nat. Cell Biol.* 22, 1187–1196. doi: 10.1038/s41556-020-00578-6

Weil, P. A., Luse, D. S., Segall, J., and Roeder, R. G. (1979). Selective and accurate transcription at the Ad2 major late promoter in a soluble system dependent on purified RNA polymerase II and DNA. *Cell* 18, 469–484. doi: 10.1016/0092-8674(79)90065-5

Yamaguchi, Y., Shibata, H., and Handa, H. (2013). Transcription elongation factors DSIF and NELF: promoter proximal pausing and beyond. *Biochim. Biophys. Acta* 1829, 98–104. doi: 10.1016/j.bbagrm.2012.11.007

Yang, C. S., Chang, K. Y., and Rana, T. M. (2014). Genome-wide functional analysis reveals factors needed at the transition steps of induced reprogramming. *Cell Rep.* 8, 327–337. doi: 10.1016/j.celrep.2014.07.002

Zuo, L., Zhang, G., Massett, M., Cheng, J., Guo, Z., Wang, L., et al. (2021). Loci-specific phase separation of FET fusion oncoproteins promotes gene transcription. *Nat. Commun.* 12:1491. doi: 10.1038/s41467-021-21690-7

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Molecular Biosciences

MINI REVIEW
published: 14 May 2021
doi: 10.3389/fmolb.2021.680090

Specific Features of RNA Polymerases I and III: Structure and Assembly

Tomasz W. Turowski[1,2†] *and Magdalena Boguta*[2*†]

[1] *Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh, United Kingdom,* [2] *Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland*

RNA polymerase I (RNAPI) and RNAPIII are multi-heterogenic protein complexes that specialize in the transcription of highly abundant non-coding RNAs, such as ribosomal RNA (rRNA) and transfer RNA (tRNA). In terms of subunit number and structure, RNAPI and RNAPIII are more complex than RNAPII that synthesizes thousands of different mRNAs. Specific subunits of the yeast RNAPI and RNAPIII form associated subcomplexes that are related to parts of the RNAPII initiation factors. Prior to their delivery to the nucleus where they function, RNAP complexes are assembled at least partially in the cytoplasm. Yeast RNAPI and RNAPIII share heterodimer Rpc40-Rpc19, a functional equivalent to the αα homodimer which initiates assembly of prokaryotic RNAP. In the process of yeast RNAPI and RNAPIII biogenesis, Rpc40 and Rpc19 form the assembly platform together with two small, bona fide eukaryotic subunits, Rpb10 and Rpb12. We propose that this assembly platform is co-translationally seeded while the Rpb10 subunit is synthesized by cytoplasmic ribosome machinery. The translation of Rpb10 is stimulated by Rbs1 protein, which binds to the 3′-untranslated region of *RPB10* mRNA and hypothetically brings together Rpc19 and Rpc40 subunits to form the αα-like heterodimer. We suggest that such a co-translational mechanism is involved in the assembly of RNAPI and RNAPIII complexes.

Keywords: RNA polymerase I, RNA polymerase III, complex assembly, transcription factors, tRNA, rRNA

OPEN ACCESS

Edited by:
Francisco Navarro, University of Jaén, Spain

Reviewed by:
Giorgio Dieci, University of Parma, Italy
Carlos Fernandez-Tornero, Margarita Salas Center for Biological Research, Spanish National Research Council, Spain
Olivier Gadal, FR3743 Centre de Biologie Intégrative (CBI), France

***Correspondence:**
Magdalena Boguta
magda@ibb.waw.pl

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Protein and RNA Networks, a section of the journal Frontiers in Molecular Biosciences

Received: *13 March 2021*
Accepted: *16 April 2021*
Published: *14 May 2021*

Citation:
Turowski TW and Boguta M (2021) Specific Features of RNA Polymerases I and III: Structure and Assembly. Front. Mol. Biosci. 8:680090. doi: 10.3389/fmolb.2021.680090

INTRODUCTION

Gene expression is one of the most fundamental processes in all domains of life. DNA is transcribed to RNA by complex machinery, the core component of which is RNA polymerase (RNAP). Both bacteria and archaea have single RNAPs, multiprotein complexes that originated from two-barrel RNA polymerase enzymes and present a high degree of similarity, including other core subunits and various auxiliary factors (**Figure 1**; Werner and Grohmann, 2011; Fouqueau et al., 2017). Eukaryotes have at least three RNAPs that transcribe nuclear genes. RNAPII, which transcribes messenger RNAs (mRNAs), is most similar to archaeal RNAP (Werner and Weinzierl, 2002). RNAPI and RNAPIII specialize in transcribing highly abundant non-coding RNAs, including ribosomal RNA (rRNA) and transfer RNA (tRNA).

The mechanisms that allowed for the evolution of RNAPI and RNAPIII remain unknown. Recent findings suggest that eukaryotic cells evolved from Asgard archaea, which are able to form a stable interface with bacteria (Zaremba-Niedzwiedzka et al., 2017; Imachi et al., 2020).

This evolutionary step may be associated with the establishment of the compact nucleoprotein organization which formed pre-nucleus and thus reflect a physical limitation that is available for transcription.

OVERVIEW OF TRANSCRIPTION SYSTEMS

Yeast RNAPII transcribes various different transcripts, mainly mRNAs, the abundance of which spans slightly more than two orders of magnitude (Lahtvee et al., 2017). Transcripts undergo various co-transcriptional modifications, including 5′ capping, splicing, cleavage, and polyadenylation. RNAPI transcribes only one 7-kb-long pre-rRNA, a polycistronic transcript from ∼150 rDNA repeats in yeast. RNAPI undergoes general regulation, and its transcriptional output is regulated by the availability of rDNA repeats (Wittner et al., 2011; Turowski, 2013). RNAPIII transcribes short, abundant non-coding RNA, including tRNA and 5S rRNA (Leśniewska and Boguta, 2017).

Transcription initiation by RNAPII depends on multiple transcription factors (TFs), including TATA-binding protein (TBP), TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH (Schier and Taatjes, 2020). RNAPI and RNAPIII initiate transcription *in vivo* by utilizing dedicated TFs. RNAPI utilizes Rrn3, TBP, core factor (CF), and upstream-associated factor (UAF) (**Figures 1C,D**; Albert et al., 2012). The RNAPIII preinitiation complex includes binding of internal promoters by multisubunit TFIIIC followed by recruitment of TFIIIB (consisted of TBP, Brf1/Brf2, and Bdp1) to the transcription start site (**Figures 1E,F**). TBP is involved in transcription initiation by all three RNAPs, and is recruited to TATA-containing as well to TATA-less promoters, while Brf1 is functionally related to the TFIIB (Turowski and Tollervey, 2016; Ciesla et al., 2018; Ramsay and Vannini, 2018). TFs play a key role in transcription initiation which requires opening of the DNA double helix and directing initial RNA synthesis. When formed, the DNA-RNA-RNAP ternary complex has extraordinary stability (Cai and Luse, 1987; Churchman and Weissman, 2011). Biochemical data clearly indicate that all RNAPs have high affinity for an RNA-DNA hybrid (Greive and von Hippel, 2005), confirming that opening of the DNA double helix is a key step in transcription initiation for all eukaryotic RNAPs whereas additional RNAP-specific factors account for differences in promoter recognition and gene-class specific regulation.

Research during the last decade revealed new mechanisms that are important for the regulation of eukaryotic transcription. RNAPII was shown to transcribe nearly the entire genome at a low level, a process referred to as pervasive transcription (Churchman and Weissman, 2011; Milligan et al., 2016). Many RNAPII promoters are bidirectional, and antisense transcription is common. This is in marked contrast to RNAPI and RNAPIII transcription, which uses very specific promoters and remains unidirectional (Turowski et al., 2016, 2020; Clarke et al., 2018). Finally, transcription is regulated by the local concentration of TFs and three-dimensional chromatin organization (Hnisz et al., 2017). A high number of very weak, multivalent interactions within transcription preinitiation complexes may lead to liquid-liquid phase separation. This phenomenon was previously reported for yeast RNAPI and pre-rRNA transcription and processing (Lafontaine, 2019). Recently, phase separation was demonstrated to drive chromatin function in the human genome (Cook and Marenduzzo, 2018; Sabari et al., 2018; Frottin et al., 2019; Brackey et al., 2020).

RNA POLYMERASE STRUCTURE: SIMILARITIES AND DIFFERENCES

There is remarkable structural and functional conservation among RNAP enzymes in all eukaryotes, from yeast to man. RNAPI and RNAPIII are homologous to RNAPII, but their structures incorporated additional subunit homologs to RNAPII TFs (**Figures 1A,B**). The majority of subunits are encoded by independent, RNAP-specific genes. Two subunits, Rpc40 and Rpc19, are homologous to bacterial α and shared between RNAPI and RNAPIII (Wild and Cramer, 2012). Moreover, all three eukaryotic RNAPs share five relatively small subunits: Rpb5, Rpb6, Rpb8, Rpb10, and Rpb12. Four subunits common for RNAPI and RNAPIII, Rpc40, Rpc19, Rpb10, and Rpb12, form a subcomplex called the assembly platform corresponding to the assembly platform that was defined for archaeal RNAP (Werner et al., 2000; Werner and Weinzierl, 2002). All RNAPs contain the two largest subunits that are homologous to bacterial β and β′ and slightly vary in size. For RNAPI, these are Rpa190 and Rpa135. For RNAPIII, these are Rpc160 and Rpc128. Both RNAPI and RNAPIII lack the long unstructured C-terminal domain (CTD) that is present in Rpb1, the largest subunit of RNAPII. The CTD is responsible for binding and orchestrating many RNA processing factors, such as capping enzymes or the spliceosome, and its role is tightly coupled to phosphorylation status of the CTD (Hsin and Manley, 2012). Moreover, the CTD was shown to regulate RNAPII clustering *via* a phase separation mechanism (Boehning et al., 2018).

The RNAPII Rpb4/7 (stalk) subcomplex interacts with Rpb1 directly and *via* an Rpb6 interaction (Armache et al., 2005). Interestingly, Rpb6, a subunit that is common to all three RNAPs and homologous to a small ω subunit of bacterial RNAP, participates in anchoring stalk homologs in RNAPI and RNAPIII (i.e., the heterodimers Rpa14/43 and Rpc17/25, respectively (Minakhin et al., 2001; Jasiak et al., 2006; Engel et al., 2013; Fernández-Tornero et al., 2013).

In contrast to RNAPII, specialized RNAPs incorporated TFIIF-like heterodimers as stable Rpa49/34 subunits for RNAPI and Rpc37/53 subunits for RNAPIII. Additionally, the C-terminal region of Rpa49 forms a "tandem winged helix" domain that is predicted in TFIIE (Geiger et al., 2010). The Rpa49/34 heterodimer plays a role in transcription initiation and interactions with the TF Rrn3 (Beckouet et al., 2008; Albert et al., 2011). Furthermore, RNAPIII contains a heterotrimeric subcomplex, Rpc82/34/31, that is similar to TFIIE and crucial for transcription initiation (Fernández-Tornero et al., 2007).

Another interesting feature of specialized RNAPs is incorporation of the TFIIS zinc-finger domain into polymerase

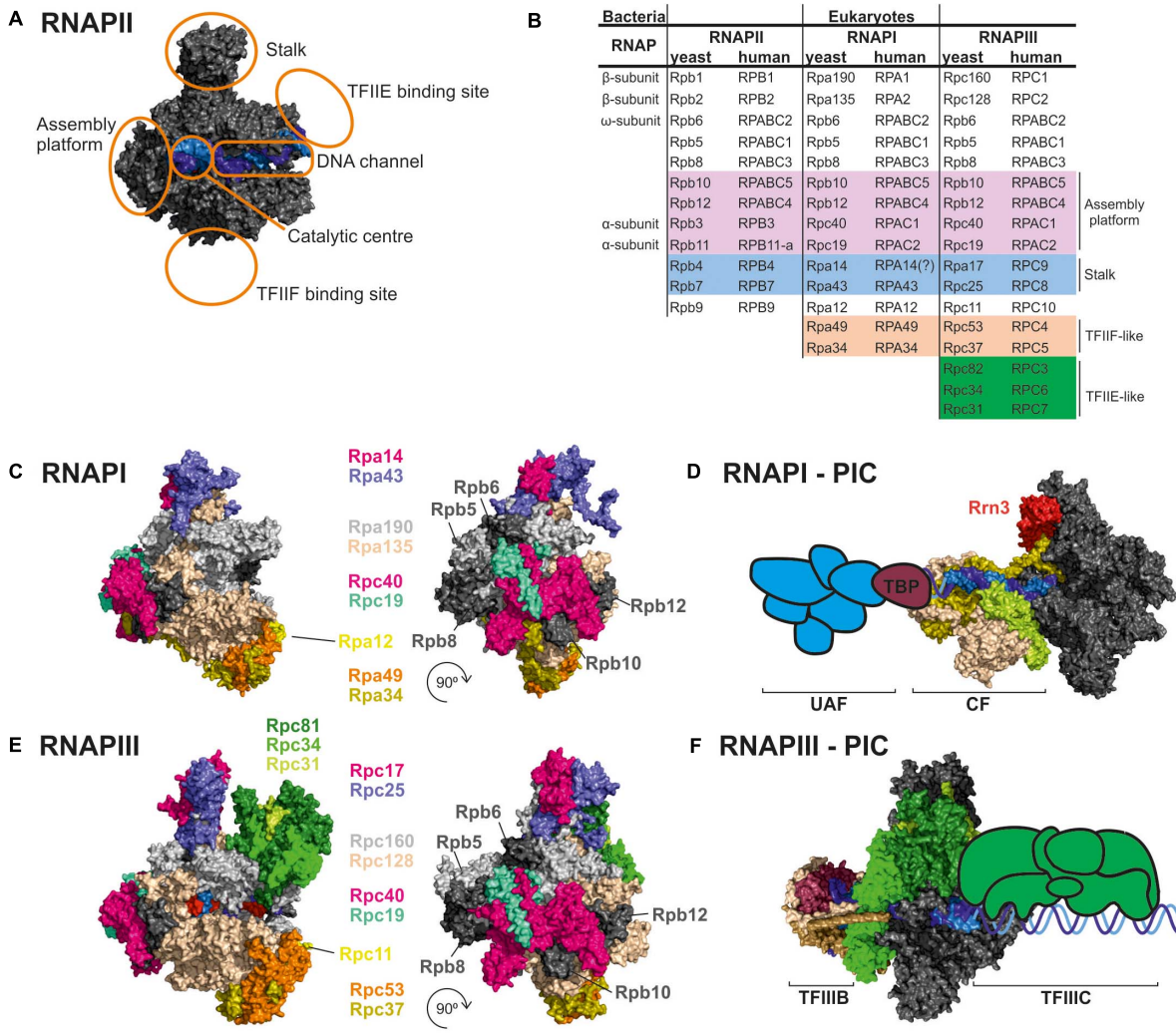


FIGURE 1 | Comparison of RNAPI and RNAPIII structures and transcription factors. **(A)** General architecture of RNAPII, consisting of the catalytic core and stalk. RNAPII core consists of a DNA binding channel, catalytic center, and assembly platform. RNAPII binds multiple transcription factors (TFs). Some TFs are homologous to additional subunits of specialized RNAPs (i.e., TFIIF). **(B)** Subunit composition of eukaryotic RNAPs. Human nomenclature is shown for comparision. Please note that C-terminal region of Rpa49 subunit harbors a "tandem winged helix" which is predicted in TFIIE and that human RNAPIII RPC7 subunit is coded by two isoforms α and β. The question mark indicates name unconfirmed. **(C)** Subunit composition of yeast RNAPI. **(D)** Model of the RNAPI pre-initiation complex, showing an early intermediate with visible Rrn3 and core factor (CF). TATA-binding protein (TBP) and upstream-associated factor (UAF) are added schematically. **(E)** Subunit composition of yeast RNAPIII. **(F)** Atomic model of RNAPIII pre-initiation complex with TFIIIB. The Rpc82/34/31 heterotrimer is involved in initiation and marked in green as in E. TFIIIC is added schematically. PDB: 5C4X, 5FJ8, 4C3J, 6EU0, and 6TPS (Fernández-Tornero et al., 2013; Barnes et al., 2015; Hoffmann et al., 2015; Abascal-Palacios et al., 2018; Pilsl and Engel, 2020).

subunits (Ruan et al., 2011; Khatter et al., 2017). This domain is responsible for the endonucleolytic cleavage of the nascent RNA 3′ end. In RNAPI and RNAPIII this domain fuses with Rpa12 and Rpc11 subunits, respectively. Therefore, specialized RNAPs are predicted to more effectively release from polymerase backtracking. In summary, the permanent recruitment of TFs might contribute to the efficiency of RNAPI and RNAPIII that is fundamental for optimization of the cell growth rate.

Finally, RNAPI incorporated unique features that allow complex dimerization. The dimerization of RNAPI has been shown for *S. cerevisiae* and *S. pombe*, suggesting that this is a conserved phenomenon. A homodimer of RNAPI is assembled in response to environmental stress, such as nutrient deprivation. This mechanism is reversible and can also be induced by perturbations in the ribosome biogenesis pathway, suggesting that homodimer assembly may be a storage mechanism of RNAPI (Torreira et al., 2017; Heiss et al., 2021).

The specialization of RNAP machinery appears to be a driver upon the archaea-to-eukaryote transition. Nevertheless, the incorporation of TFs may suggest an additional mechanism. We speculate that limited space within a crowded environment of the pre-nucleus transformed transient interactions into the stable incorporation of TFs into structures of RNAPI and RNAPIII. In fact, archaeal general TFB binds upstream protein

coding genes but is depleted upstream the rRNA, indicating that differences between the occupancy of TFs between rRNA and mRNA transcription units are also present in archaea (Smollett et al., 2017). Additionally, ribosomal components loop together in archaeal chromatin, suggesting the spatial organization of ribosome biogenesis (Takemata and Bell, 2021). Therefore, we suggest that spatial organization of the eukaryotic genome promoted the evolution of RNAP-specific and co-evolution of specific TFs. Ultimately, the evolution of specialized transcription machinery allowed the optimal use of limited space in the nucleus organized by chromatin.

ASSEMBLY OF RNAPI AND RNAPIII

Detailed knowledge of the structures of yeast RNA polymerases contrasts with the incomplete information on the control of their assembly. A hypothetical model of RNAPI and RNAPIII assembly is based on the relatively well-recognized assembly pathway of bacterial RNAP (Ghosh et al., 2001; Kannan et al., 2001; Patel et al., 2020). The initial complex is formed by two α-like subunits, Rpc40 and Rpc19 (Wild and Cramer, 2012). As supported by genetic data, formation of the Rpc19/40 heterodimer additionally involves a small common subunit, Rpb10, which has no equivalent in the prokaryotic enzyme. Mutations of the conserved motif of Rpb10 lead to a complete depletion of the largest RNAPI subunit (Rpa190) suggesting that the mutant enzyme is not properly assembled (Gadal et al., 1999).

Rbp10 overexpression suppresses conditional *rpc40* and *rpc19* mutations that prevent RNAPIII assembly (Lalo et al., 1993) as well as a conditional *rpc128-1007* mutant that is located in the Rpc128 subunit near contact points for the association between Rpc128 and Rpc40 contact points (Cieśla et al., 2015). Rpb10 may function in the RNAP assembly platform by acting as structural adaptor between the α-like dimer Rpc40-Rpc19 and catalytic β-like subunit Rpc128. Such a role was suggested for the archaeal subunit N, which is homologous to yeast Rpb10 (Werner et al., 2000). Essential function in the formation of assembly platform of all RNAPs, by bridging between the Rpc40-Rpc19-Rpb10 subcomplex (or Rpb3-Rpb11-Rpb10 in RNAP II) and the β-like subunit, was postulated for Rpb12 (Cramer et al., 2000). A role of Rpb12 in RNAPIII assembly was also supported by earlier genetic data (Rubbi et al., 1999).

The existence of intermediate complexes in the process of yeast RNAP assembly was suggested by the mass spectrometry analysis of RNAPIII disassembly (Lorenzen et al., 2007; Lane et al., 2011). These analyses revealed two stable subcomplexes, Rpc128-Rpc40-Rpc19-Rpb12 and Rpc160-Rpb8-Rpb5. In addition to Rpb10, other small subunits also contribute to the association of these macromolecular assembles (Minakhin et al., 2001; Mirón-García et al., 2013). Although common to all RNAPs, the small subunits may have distinct functions in the assembly of each RNAP, thereby providing an interaction platform for other molecules (Voutsina et al., 1999).

According to an existing model (Wild and Cramer, 2012), eukaryotic RNAP enzymes are at least partially assembled in the cytoplasm and then imported to the nucleus as a complex with specific adaptor proteins. A set of RNAPIII subunits exhibit coordinated nuclear import, indicating that the RNAPIII core is assembled in the cytoplasm, with additional components that bind in the nucleus (Hardeland and Hurt, 2006). This suggests that the specific subcomplexes, particularly Rpc82-Rpc34-Rpc31, would only bind the core in the nucleus (Hardeland and Hurt, 2006). Interestingly, efficient RNAPIII assembly requires sumoylation of the Rpc82 subunit, which is RNAPIII-specific (Chymkowitch et al., 2017).

Several auxiliary factors, originally implicated in RNAPII assembly and nuclear import and subsequently shown to be common to RNAPI and RNAPIII were described in another article published in the same issue by Navarro and colleagues. Here we focus on the Rbs1 protein, a candidate RNAPIII assembly/import factor, which was identified in a genetic screen for suppressors of the RNAPIII assembly mutant *rpc128-1007* (Cieśla et al., 2015). Genetic suppression correlated with an increase in the stability of RNAPIII subunits and an increase in their interaction. Additionally, Rbs1 physically interacts with a subset of RNAPIII subunits (i.e., Rpc19, Rpc40, and Rpb5) and the exportin Crm1. We postulated that Rbs1 binds to the RNAPIII complex or subcomplex and facilitates its translocation to the nucleus. Following dissociation from RNAPIII in the nucleus, Rbs1 is exported back to the cytoplasm in complex with Crm1 (Cieśla et al., 2015).

It is reasonable that the Rbs1 function in RNAP assembly is not limited to RNAPIII. Rbs1 interacts with Rpc19 and Rpc40 subunits common to RNAP I and RNAPIII and Rpb5, a component of all three RNAPs (Cieśla et al., 2015). Moreover, Rpb5 participates in the assembly of all three polymerases mediated by Bud27 (Mirón-García et al., 2013).

Genetic and functional suppression of the RNAPIII assembly defect by Rbs1 correlated with higher levels of *RPB10* mRNA and Rpb10 protein. This regulatory mechanism, however, relies on the control of steady-state levels of *RPB10* mRNA by Rbs1 protein, which interacts with the 3′-untranslated region (UTR) of this transcript (Cieśla et al., 2020).

By exploring specific features of the Rbs1 protein sequence, we identified two regions: a highly ordered N-terminal region that comprises two RNA-interacting domains (R3H and SUZ) and a mostly disordered C-terminal region with a prionogenic (aggregation-promoting) sequence. Investigations of possible roles of these regions in *RBS1* led to the conclusion that the R3H domain was essential for suppressing both genetic and molecular phenotypes of the *rpc128-1007* mutation and function of Rbs1 protein in RNAPIII assembly, whereas the role of the prionogenic domain remains unknown (Cieśla et al., 2020).

By applying ultraviolet crosslinking, we identified the transcriptome-wide binding of Rbs1, which predominately targets 3′-UTRs of mRNAs. The list of high-confidence Rbs1 targets included *RPB10* mRNA and *RPC19* mRNA, which encodes Rpc19, another subunit involved in formation of the assembly platform for RNAPIII (Cieśla et al., 2020).

Notably, homologs of Rbs1 have been identified in other eukaryotes, including the human proteins R3H domain protein 2 (R3HDM2) and cyclic adenosine monophosphate-regulated

phosphoprotein 21 (ARPP21), which are also known to interact with mRNA (Castello et al., 2012; Rehfeld et al., 2018).

MODEL OF CO-TRANSLATIONAL FORMATION OF THE RPC40-RPC19-RPB12-RPB10 ASSEMBLY PLATFORM

Based on established interactions between Rbs1 and the subunits of RNAPIII (Cieśla et al., 2015, 2020), we propose a co-translational mechanism of formation of the early-stage assembly intermediate of the RNAPIII complex and potentially also RNAPI (**Figure 2**). According to our hypothesis, RNAPIII assembly might be seeded while the Rpb10 subunit of the enzyme core is being synthesized by cytoplasmic ribosome machinery. This assembly pathway would be proceeded by the co-translational association of other subunits, including Rpc19 and Rpc40 (**Figure 2**), to build an initial assembly subcomplex that is common to RNAPI and RNAPIII. Currently unknown, however, is how the Rpc40-Rpc19-Rpb10 complex discriminates among Rpc128 and Rpa135 proteins to proceed with formation of the RNAPIII and RNAPI assembly platform.

Co-translational assembly has been reported for several multisubunit complexes (e.g., TFIID, TREX-2, SAGA, and fatty acid synthase; Kamenova et al., 2019; Schwarz and Beck, 2019; Shiber et al., 2018) but has not yet been considered for RNA polymerases. Our hypothesis is in line with the idea that that co-translational subunit association is likely to be a general principle in yeast and mammalian cells as an efficient assembly pathway in eukaryotes (Shiber et al., 2018).

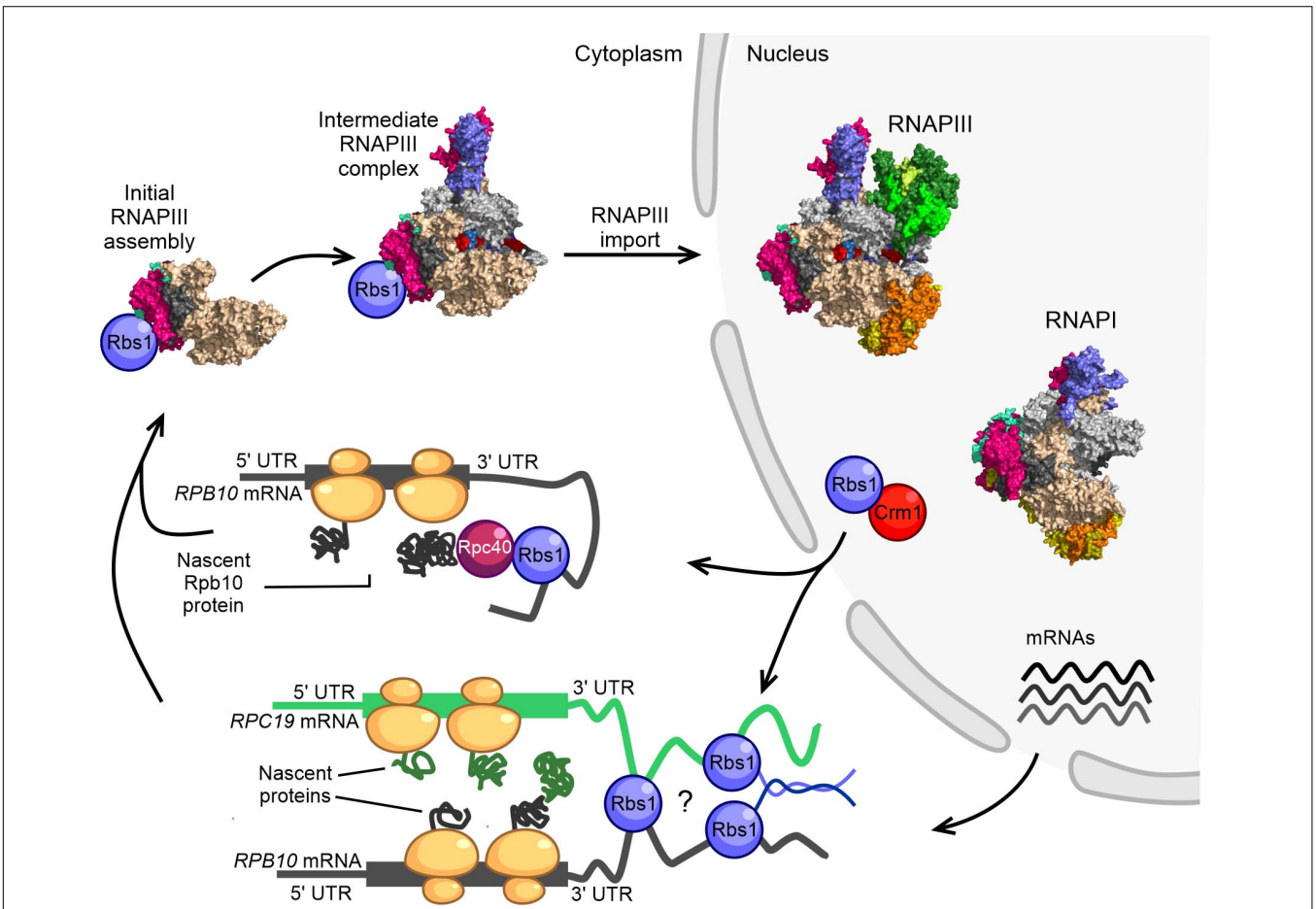


FIGURE 2 | Model of RNAPIII biogenesis in the yeast *Saccharomyces cerevisiae*. The control of Rpb10 expression and role of Rpb10 in assembly of the RNAP III complex are connected *via* a regulatory loop that involves Rbs1 protein. Possible ways in which subunits of the RNAPIII intermediate complex are brought together for co-translational assembly are shown. The initial step of RNAPIII assembly in *Saccharomyces cerevisiae* occurs in the cytoplasm. Formation of the intermediate Rpc128-Rpc40-Rpc19-Rpb12-Rpb10 subcomplex is seeded co-translationally while the Rpb10 subunit is being synthesized by cytoplasmic ribosomes. Rbs1 is an RNA binding protein that stimulates the translation of Rpb10 protein through an interaction of the R3H domain with the 3′-UTR in *RPB10* mRNA. Rpb10 brings together the Rpc19 and Rpc40 subunits to form the α-like heterodimer. One possibility is that Rbs1 binds and recruits the mature Rpc40 subunit to the 3′-UTR of *RPB10* mRNA, which undergoes translation. The Rpc40-Rbs1 interaction has been previously demonstrated by co-immunoprecipitation. Alternatively, Rbs1 protein directly bridges *RPB10* mRNA and *RPC19* mRNA. A fully folded subunit that formed on one mRNA was recently shown to detach from its ribosome and interact with a nascent protein on another mRNA (Cieśla et al., 2015, 2020).

For RNAPIII, we propose two plausible models that are not necessarily mutually exclusive and could be applicable to RNAPI (**Figure 2**). In the first model, the long 3′-UTR of *RPB10* acts as a scaffold to recruit Rbs1 that is associated with another RNAPIII subunit (e.g., Rpc40) that interacts with Rbs1 through co-immunoprecipitation (Cieśla et al., 2015) to the site of Rpb10 translation. This facilitates the association of this subunit with the newly translated Rpb10 to form the RNAP assembly platform subcomplex. Such a scenario corresponds to a sequential assembly model, in which RNA-binding protein recruits a fully folded subunit to the 3′-UTR of mRNA that encodes the second subunit that undergoes translation. The 3′-UTR regions can act as scaffolds for RNA binding proteins that serve as adaptors to deliver preferred proteins to the site of translation (Berkovits and Mayr, 2015). The sequential assembly pathway has been proposed for the co-translational assembly of TAF8-TAF10 subunits of TFIID and TAF1-TBP assembly. TAF10 binds the nascent TAF8 subunit, and TAF10 protein co-localizes with *TAF8* mRNA in cytoplasmic foci (Kamenova et al., 2019).

In the second model, *RPB10* and *RPC19* mRNAs are bridged together by Rbs1, which interacts with 3′-regulatory regions of both transcripts (Cieśla et al., 2020). Additionally, unstructured parts of Rbs1 may facilitate interaction among Rbs1 molecules allowing the Rpb10 and Rpc19 subunits to be translated in proximity to each other, thereby enabling their co-translational interaction (**Figure 2**). A simultaneous model has been proposed for the co-translational assembly of TAF6 and TAF9 subunits of the transcription factor TFIID (Kamenova et al., 2019). Physical linkage of the two mRNAs could also be accomplished by their co-localization in phase-separated compartments that allow translation at defined subcellular locations (Mayr, 2018).

Rbs1 exhibits all characteristics of the postulated protein that bridges mRNA. The two RNA-interacting domains, R3H and SUZ, have been identified in the sequence of Rbs1 protein, and this sequence also contains a prionogenic, disordered region. The specific mRNA motifs and potential effect of Rbs1 binding on the translation of these targets need to be determined. R3H likely cooperates with the SUZ domain in the recognition of specific mRNA targets and bridging them into proximity with each other. A disordered region of Rbs1 may be involved in multivalent interactions that bring Rbs1-associated mRNAs together. Such an Rbs1-mediated co-localization of mRNAs would allow them to be translated at defined subcellular locations.

STOICHIOMETRY OF SUBUNITS OF SPECIALIZED RNAPS

The assembly platform Rpc40-Rpc19-Rpb10-Rpb12 is shared between yeast RNAPI and III what arises question about the stoichiometry of RNAPs subunits during the assembly pathway. The absolute quantification of yeast proteins indicated that RNAPI and RNAPII are present in 5,000 copies per cell, whereas RNAPIII is present in 2,500 copies (Turowski et al., 2020). Consequently, common subunits are shared between RNAPI, RNAPII, and RNAPIII in a 2:2:1 ratio. RNAPI and RNAPIII share an assembly platform that contains the Rpc19 and Rpc40 subunits and two additional subunits (Rpb10 and Rpb12) among the five common subunits. Both specialized RNAPs utilize the assembly platform, sharing RNAPI:RNAPIII in a 2:1 ratio. The platform is attached *via* the second largest subunit Rpa135 to RNAPI and *via* Rpc128 to RNAPIII. Limited data suggest a difference in binding strength at this stage. A biochemical disassembly approach demonstrated that RNAPI disassembles the platform from the dimer of the two largest subunits, Rpa135 and Rpa190, whereas RNAPIII disassembles the interface between the two largest subunits before detachment of the assembly platform (Lane et al., 2011). This suggests that the Rpc128-platform interaction might be stronger than the interaction between the two largest subunits. This would be in contrast to RNAPI, in which the interaction with the two largest subunits would be stronger than the interaction with the platform. In the consequence, a common assembly platform could be preferentially incorporated by less abundant RNAPIII.

DISCUSSION

Despite recent progress, the RNAP assembly process remains poorly described. Knowledge about its basic mechanism is necessary to ask more detailed questions about disease and developmental biology. The structure of human RNAPI awaits to be determined. Recently published structures of human RNAPIII revealed a high level of conservation (Ramsay et al., 2020; Li et al., 2021; Wang et al., 2021). Moreover, mutations of specialized RNAPs lead to genetic disorders, such as Treacher-Collins syndrome and hypomyelinating leukodystrophy (Ramsay et al., 2020; Girbig et al., 2021), demonstrating the requirement for precise coordination among all three RNAPs and their assembly. Research on RNAPIII assembly in yeast focused on *rpc128-1007* mutations that disturbed the interface between the Rpc128 and Rpc40 subunits. Interestingly, multiple disease-associated mutations of human RNAPIII subunits tend to cluster within the region of the RNAPIII assembly platform, suggesting that defects in RNAPIII biogenesis may have severe health consequences (Ramsay et al., 2020; Girbig et al., 2021). Further studies of RNAP assembly should reveal additional factors that are involved in this process and improve our understanding of this vital pathway.

AUTHOR CONTRIBUTIONS

Both authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was supported by National Science Centre (UMO-2017/25/B/NZ1/01889) to MB and Polish National Agency for Academic Exchange (PPN/PPO/2020/2/00004/U/00001) to TWT.

REFERENCES

Abascal-Palacios, G., Ramsay, E. P., Beuron, F., Morris, E., and Vannini, A. (2018). Structural basis of RNA polymerase III transcription initiation. *Nature* 553, 301–306. doi: 10.1038/nature25441

Albert, B., Léger-Silvestre, I., Normand, C., Ostermaier, M. K., Pérez-Fernández, J., Panov, K. I., et al. (2011). RNA polymerase I–specific subunits promote polymerase clustering to enhance the rRNA gene transcription cycle. *J. Cell Biol.* 192, 277–293. doi: 10.1083/jcb.201006040

Albert, B., Perez-Fernandez, J., Léger-Silvestre, I., and Gadal, O. (2012). Regulation of ribosomal RNA production by RNA polymerase I: does elongation come first? *Genet. Res. Int.* 2012:276948.

Armache, K.-J., Mitterweger, S., Meinhart, A., and Cramer, P. (2005). Structures of complete RNA polymerase II and Its subcomplex, Rpb4/7*. *J. Biol. Chem.* 280, 7131–7134. doi: 10.1074/jbc.m413038200

Barnes, C. O., Calero, M., Malik, I., Graham, B. W., Spahr, H., Lin, G., et al. (2015). Crystal structure of a transcribing RNA polymerase II complex reveals a complete transcription bubble. *Mol. Cell* 59, 258–269. doi: 10.1016/j.molcel.2015.06.034

Beckouet, F., Labarre-Mariotte, S., Albert, B., Imazawa, Y., Werner, M., Gadal, O., et al. (2008). Two RNA polymerase I subunits control the binding and release of Rrn3 during transcription. *Mol. Cell. Biol.* 28, 1596–1605. doi: 10.1128/mcb.01464-07

Berkovits, B. D., and Mayr, C. (2015). Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. *Nature* 522, 363–367. doi: 10.1038/nature14321

Boehning, M., Dugast-Darzacq, C., Rankovic, M., Hansen, A. S., Yu, T., Marie-Nelly, H., et al. (2018). RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat. Struct. Mol. Biol.* 25, 833–840. doi: 10.1038/s41594-018-0112-y

Brackey, C. A., Marenduzzo, D., and Gilbert, N. (2020). Mechanistic modeling of chromatin folding to understand function. *Nat. Methods* 17, 767–775. doi: 10.1038/s41592-020-0852-6

Cai, H., and Luse, D. S. (1987). Transcription initiation by RNA polymerase II in vitro. Properties of preinitiation, initiation, and elongation complexes. *J. Biol. Chem.* 262, 298–304. doi: 10.1016/s0021-9258(19)75926-2

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B. M., Strein, C., et al. (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 149, 1393–1406. doi: 10.1016/j.cell.2012.04.031

Churchman, L. S., and Weissman, J. S. (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* 469, 368–373. doi: 10.1038/nature09652

Chymkowitch, P., Nguéa, P. A., Aanes, H., Robertson, J., Klungland, A., and Enserink, J. M. (2017). TORC1-dependent sumoylation of Rpc82 promotes RNA polymerase III assembly and activity. *Proc. Natl. Acad. Sci. U.S.A.* 114, 1039–1044. doi: 10.1073/pnas.1615093114

Cieśla, M., Makała, E., Płonka, M., Bazan, R., Gewartowski, K., Dziembowski, A., et al. (2015). Rbs1, a new protein implicated in RNA polymerase III biogenesis in yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 35, 1169–1181. doi: 10.1128/mcb.01230-14

Ciesla, M., Skowronek, E., and Boguta, M. (2018). Function of TFIIIC, RNA polymerase III initiation factor, in activation and repression of tRNA gene transcription. *Nucleic Acids Res.* 46, 9444–9455. doi: 10.1093/nar/gky656

Cieśla, M., Turowski, T. W., Nowotny, M., Tollervey, D., and Boguta, M. (2020). The expression of Rpb10, a small subunit common to RNA polymerases, is modulated by the R3H domain-containing Rbs1 protein and the Upf1 helicase. *Nucleic Acids Res.* 48, 12252–12268. doi: 10.1093/nar/gkaa1069

Clarke, A. M., Engel, K. L., Giles, K. E., Petit, C. M., and Schneider, D. A. (2018). NETSeq reveals heterogeneous nucleotide incorporation by RNA polymerase I. *Proc. Natl. Acad. Sci. U.S.A.* 115, E11633–E11641.

Cook, P. R., and Marenduzzo, D. (2018). Transcription-driven genome organization: a model for chromosome structure and the regulation of gene expression tested through simulations. *Nucleic Acids Res.* 46, 9895–9906. doi: 10.1093/nar/gky763

Cramer, P., Bushnell, D. A., Fu, J., Gnatt, A. L., Maier-Davis, B., Thompson, N. E., et al. (2000). Architecture of RNA Polymerase II and implications for the transcription mechanism. *Science* 288, 640–649. doi: 10.1126/science.288.5466.640

Engel, C., Sainsbury, S., Cheung, A. C., Kostrewa, D., and Cramer, P. (2013). RNA polymerase I structure and transcription regulation. *Nature* 502, 650–655.

Fernández-Tornero, C., Böttcher, B., Riva, M., Carles, C., Steuerwald, U., Ruigrok, R. W. H., et al. (2007). Insights into transcription initiation and termination from the electron microscopy structure of yeast RNA polymerase III. *Mol. Cell* 25, 813–823. doi: 10.1016/j.molcel.2007.02.016

Fernández-Tornero, C., Moreno-Morcillo, M., Rashid, U. J., Taylor, N. M. I., Ruiz, F. M., Gruene, T., et al. (2013). Crystal structure of the 14-subunit RNA polymerase I. *Nature* 502, 644–649.

Fouqueau, T., Blombach, F., and Werner, F. (2017). Evolutionary origins of two-barrel RNA polymerases and site-specific transcription initiation. *Annu. Rev. Microbiol.* 71, 331–348. doi: 10.1146/annurev-micro-091014-104145

Frottin, F., Schueder, F., Tiwary, S., Gupta, R., Körner, R., Schlichthaerle, T., et al. (2019). The nucleolus functions as a phase-separated protein quality control compartment. *Science* 365, 342–347. doi: 10.1126/science.aaw9157

Gadal, O., Shpakovski, G. V., and Thuriaux, P. (1999). Mutants in ABC10beta, a conserved subunit shared by all three yeast RNA polymerases, specifically affect RNA polymerase I assembly. *J. Biol. Chem.* 274, 8421–8427. doi: 10.1074/jbc.274.13.8421

Geiger, S. R., Lorenzen, K., Schreieck, A., Hanecker, P., Kostrewa, D., Heck, A. J. R., et al. (2010). RNA polymerase I contains a TFIIF-related DNA-binding subcomplex. *Mol. Cell* 39, 583–594. doi: 10.1016/j.molcel.2010.07.028

Ghosh, P., Ishihama, A., and Chatterji, D. (2001). *Escherichia coli* RNA polymerase subunit omega and its N-terminal domain bind full-length beta' to facilitate incorporation into the alpha2beta subassembly. *Eur. J. Biochem.* 268, 4621–4627. doi: 10.1046/j.1432-1327.2001.02381.x

Girbig, M., Misiaszek, A. D., Vorländer, M. K., Lafita, A., Grötsch, H., Baudin, F., et al. (2021). Cryo-EM structures of human RNA polymerase III in its unbound and transcribing states. *Nat. Struct. Mol. Biol.* 28, 210–219. doi: 10.1038/s41594-020-00555-5

Greive, S. J., and von Hippel, P. H. (2005). Thinking quantitatively about transcriptional regulation. *Nat. Rev. Mol. Cell Biol.* 6, 221–232. doi: 10.1038/nrm1588

Hardeland, U., and Hurt, E. (2006). Coordinated nuclear import of RNA Polymerase III subunits. *Traffic* 7, 465–473. doi: 10.1111/j.1600-0854.2006.00399.x

Heiss, F. B., Daiß, J. L., Becker, P., and Engel, C. (2021). Conserved strategies of RNA polymerase I hibernation and activation. *Nat. Commun.* 12:758.

Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K., and Sharp, P. A. (2017). A phase separation model for transcriptional control. *Cell* 169, 13–23. doi: 10.1016/j.cell.2017.02.007

Hoffmann, N. A., Jakobi, A. J., Moreno-Morcillo, M., Glatt, S., Kosinski, J., Hagen, W. J. H., et al. (2015). Molecular structures of unbound and transcribing RNA polymerase III. *Nature* 528, 231–236. doi: 10.1038/nature16143

Hsin, J.-P., and Manley, J. L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev.* 26, 2119–2137. doi: 10.1101/gad.200303.112

Imachi, H., Nobu, M. K., Nakahara, N., Morono, Y., Ogawara, M., Takaki, Y., et al. (2020). Isolation of an archaeon at the prokaryote–eukaryote interface. *Nature* 577, 519–525.

Jasiak, A. J., Armache, K.-J., Martens, B., Jansen, R.-P., and Cramer, P. (2006). Structural biology of RNA polymerase III: subcomplex C17/25 X-ray structure and 11 subunit enzyme model. *Mol. Cell* 23, 71–81. doi: 10.1016/j.molcel.2006.05.013

Kamenova, I., Mukherjee, P., Conic, S., Mueller, F., El-Saafin, F., Bardot, P., et al. (2019). Co-translational assembly of mammalian nuclear multisubunit complexes. *Nat. Commun.* 10:1740.

Kannan, N., Chander, P., Ghosh, P., Vishveshwara, S., and Chatterji, D. (2001). Stabilizing interactions in the dimer interface of α-subunit in *Escherichia coli* RNA polymerase: a graph spectral and point mutation study. *Protein Sci.* 10, 46–54. doi: 10.1110/ps.26201

Khatter, H., Vorländer, M. K., and Müller, C. W. (2017). RNA polymerase I and III: similar yet unique. *Curr. Opin. Struct. Biol.* 47, 88–94. doi: 10.1016/j.sbi.2017.05.008

Lafontaine, D. L. J. (2019). Birth of nucleolar compartments: phase separation-driven ribosomal RNA sorting and processing. *Mol. Cell* 76, 694–696. doi: 10.1016/j.molcel.2019.11.015

Lahtvee, P.-J., Sánchez, B. J., Smialowska, A., Kasvandik, S., Elsemman, I. E., Gatto, F., et al. (2017). Absolute quantification of Protein and mRNA abundances demonstrate variability in gene-specific translation efficiency in yeast. *Cell Syst.* 4, 495–504.e5.

Lalo, D., Carles, C., Sentenac, A., and Thuriaux, P. (1993). Interactions between three common subunits of yeast RNA polymerases I and III. *Proc. Natl. Acad. Sci. U.S.A.* 90, 5524–5528. doi: 10.1073/pnas.90.12.5524

Lane, L. A., Fernández-Tornero, C., Zhou, M., Morgner, N., Ptchelkine, D., Steuerwald, U., et al. (2011). Mass spectrometry reveals stable modules in holo and apo RNA polymerases I and III. *Structure* 19, 90–100. doi: 10.1016/j.str.2010.11.009

Leśniewska, E., and Boguta, M. (2017). Novel layers of RNA polymerase III control affecting tRNA gene transcription in eukaryotes. *Open Biol.* 7:170001. doi: 10.1098/rsob.170001

Li, L., Yu, Z., Zhao, D., Ren, Y., Hou, H., and Xu, Y. (2021). Structure of human RNA polymerase III elongation complex. *Cell Res.* doi: 10.1038/s41422-021-00472-2 [Epub ahead of print].

Lorenzen, K., Vannini, A., Cramer, P., and Heck, A. J. R. (2007). Structural biology of RNA polymerase III: mass spectrometry elucidates subcomplex architecture. *Structure* 15, 1237–1245. doi: 10.1016/j.str.2007.07.016

Mayr, C. (2018). Protein complexes assemble as they are being made. *Nature* 561, 186–187. doi: 10.1038/d41586-018-05905-4

Milligan, L., Huynh-Thu, V. A., Delan-Forino, C., Tuck, A., Petfalski, E., Lombraña, R., et al. (2016). Strand-specific, high-resolution mapping of modified RNA polymerase II. *Mol. Syst. Biol.* 12:874. doi: 10.15252/msb.20166869

Minakhin, L., Bhagat, S., Brunning, A., Campbell, E. A., Darst, S. A., Ebright, R. H., et al. (2001). Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proc. Natl. Acad. Sci. U.S.A.* 98, 892–897. doi: 10.1073/pnas.98.3.892

Mirón-García, M. C., Garrido-Godino, A. I., García-Molinero, V., Hernández-Torres, F., Rodríguez-Navarro, S., and Navarro, F. (2013). The prefoldin bud27 mediates the assembly of the eukaryotic RNA polymerases in an rpb5-dependent manner. *PLoS Genet.* 9:e1003297. doi: 10.1371/journal.pgen.1003297

Patel, U. R., Gautam, S., and Chatterji, D. (2020). Validation of omega subunit of RNA polymerase as a functional entity. *Biomolecules* 10:1588. doi: 10.3390/biom10111588

Pilsl, M., and Engel, C. (2020). Structural basis of RNA polymerase I pre-initiation complex formation and promoter melting. *Nat. Commun.* 11:1206.

Ramsay, E. P., Abascal-Palacios, G., Daiß, J. L., King, H., Gouge, J., Pilsl, M., et al. (2020). Structure of human RNA polymerase III. *Nat. Commun.* 11:6409.

Ramsay, E. P., and Vannini, A. (2018). Structural rearrangements of the RNA polymerase III machinery during tRNA transcription initiation. *Biochim. Biophys. Acta* 1861, 285–294. doi: 10.1016/j.bbagrm.2017.11.005

Rehfeld, F., Maticzka, D., Grosser, S., Knauff, P., Eravci, M., Vida, I., et al. (2018). The RNA-binding protein ARPP21 controls dendritic branching by functionally opposing the miRNA it hosts. *Nat. Commun.* 9:1235.

Ruan, W., Lehmann, E., Thomm, M., Kostrewa, D., and Cramer, P. (2011). Evolution of two modes of intrinsic RNA polymerase transcript cleavage. *J. Biol. Chem.* 286, 18701–18707. doi: 10.1074/jbc.m111.222273

Rubbi, L., Labarre-Mariotte, S., Chédin, S., and Thuriaux, P. (1999). Functional characterization of ABC10α, an essential polypeptide shared by all three forms of eukaryotic DNA-dependent RNA polymerases. *J. Biol. Chem.* 274, 31485–31492. doi: 10.1074/jbc.274.44.31485

Sabari, B. R., Dall'Agnese, A., Boija, A., Klein, I. A., Coffey, E. L., Shrinivas, K., et al. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. *Science* 361:eaar3958. doi: 10.1126/science.aar3958

Schier, A. C., and Taatjes, D. J. (2020). Structure and mechanism of the RNA polymerase II transcription machinery. *Genes Dev.* 34, 465–488. doi: 10.1101/gad.335679.119

Schwarz, A., and Beck, M. (2019). The benefits of cotranslational assembly: a structural perspective. *Trends Cell Biol.* 29, 791–803. doi: 10.1016/j.tcb.2019.07.006

Shiber, A., Döring, K., Friedrich, U., Klann, K., Merker, D., Zedan, M., et al. (2018). Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. *Nature* 561, 268–272. doi: 10.1038/s41586-018-0462-y

Smollett, K., Blombach, F., Reichelt, R., Thomm, M., and Werner, F. (2017). A global analysis of transcription reveals two modes of Spt4/5 recruitment to archaeal RNA polymerase. *Nat. Microbiol.* 2:17021.

Takemata, N., and Bell, S. D. (2021). Multi-scale architecture of archaeal chromosomes. *Mol. Cell* 81, 473–487.e6.

Torreira, E., Louro, J. A., Pazos, I., González-Polo, N., Gil-Carton, D., Duran, A. G., et al. (2017). The dynamic assembly of distinct RNA polymerase I complexes modulates rDNA transcription. *eLife* 6:e20832.

Turowski, T. W. (2013). The impact of transcription on posttranscriptional processes in yeast. *Gene* 526, 23–29. doi: 10.1016/j.gene.2013.04.021

Turowski, T. W., Leśniewska, E., Delan-Forino, C., Sayou, C., Boguta, M., and Tollervey, D. (2016). Global analysis of transcriptionally engaged yeast RNA polymerase III reveals extended tRNA transcripts. *Genome Res.* 26, 933–944. doi: 10.1101/gr.205492.116

Turowski, T. W., Petfalski, E., Goddard, B. D., French, S. L., Helwak, A., and Tollervey, D. (2020). Nascent transcript folding plays a major role in determining RNA polymerase elongation rates. *Mol. Cell* 79, 488–503.e11.

Turowski, T. W., and Tollervey, D. (2016). Transcription by RNA polymerase III: insights into mechanism and regulation. *Biochem. Soc. Trans.* 44, 1367–1375. doi: 10.1042/bst20160062

Voutsina, A., Riva, M., Carles, C., and Alexandraki, D. (1999). Sequence divergence of the RNA polymerase shared subunit ABC14.5 (Rpb8) selectively affects RNA polymerase III assembly in Saccharomyces cerevisiae. *Nucleic Acids Res.* 27, 1047–1055. doi: 10.1093/nar/27.4.1047

Wang, Q., Li, S., Wan, F., Xu, Y., Wu, Z., Cao, M., et al. (2021). Structural insights into transcriptional regulation of human RNA polymerase III. *Nat. Struct. Mol. Biol.* 28, 220–227. doi: 10.1038/s41594-021-00557-x

Werner, F., Eloranta, J. J., and Weinzierl, R. O. (2000). Archaeal RNA polymerase subunits F and P are bona fide homologs of eukaryotic RPB4 and RPB12. *Nucleic Acids Res.* 28, 4299–4305. doi: 10.1093/nar/28.21.4299

Werner, F., and Grohmann, D. (2011). Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* 9, 85–98. doi: 10.1038/nrmicro2507

Werner, F., and Weinzierl, R. O. J. (2002). A recombinant RNA polymerase II-like enzyme capable of promoter-specific transcription. *Mol. Cell* 10, 635–646. doi: 10.1016/s1097-2765(02)00629-9

Wild, T., and Cramer, P. (2012). Biogenesis of multisubunit RNA polymerases. *Trends Biochem. Sci.* 37, 99–105. doi: 10.1016/j.tibs.2011.12.001

Wittner, M., Hamperl, S., Stöckl, U., Seufert, W., Tschochner, H., Milkereit, P., et al. (2011). Establishment and maintenance of alternative chromatin states at a multicopy gene locus. *Cell* 145, 543–554. doi: 10.1016/j.cell.2011.03.051

Zaremba-Niedzwiedzka, K., Caceres, E. F., Saw, J. H., Bäckström, D., Juzokaite, L., Vancaester, E., et al. (2017). Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* 541, 353–358. doi: 10.1038/nature21031

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Molecular Biosciences

REVIEW
published: 20 May 2021
doi: 10.3389/fmolb.2021.668290

The Transcription-Repair Coupling Factor Mfd Prevents and Promotes Mutagenesis in a Context-Dependent Manner

Laura A. Lindsey-Boltz and Aziz Sancar**

Department of Biochemistry and Biophysics, University of North Carolina School of Medicine, Chapel Hill, NC, United States

The *mfd* (mutation frequency decline) gene was identified by screening an auxotrophic *Escherichia coli* strain exposed to UV and held in a minimal medium before plating onto rich or minimal agar plates. It was found that, under these conditions, holding cells in minimal (nongrowth) conditions resulted in mutations that enabled cells to grow on minimal media. Using this observation as a starting point, a mutant was isolated that failed to mutate to auxotrophy under the prescribed conditions, and the gene responsible for this phenomenon (mutation frequency decline) was named *mfd*. Later work revealed that *mfd* encoded a translocase that recognizes a stalled RNA polymerase (RNAP) at damage sites and binds to the stalled RNAP, recruits the nucleotide excision repair damage recognition complex $UvrA_2UvrB$ to the site, and facilitates damage recognition and repair while dissociating the stalled RNAP from the DNA along with the truncated RNA. Recent single-molecule and genome-wide repair studies have revealed time-resolved features and structural aspects of this transcription-coupled repair (TCR) phenomenon. Interestingly, recent work has shown that in certain bacterial species, *mfd* also plays roles in recombination, bacterial virulence, and the development of drug resistance.

Keywords: mutation frequency decline (MFD), nucleotide excision repair (NER), excision repair-sequencing (XR-seq), transcription-coupled repair (TCR), uvrABC excinuclease, UvrD

OPEN ACCESS

Edited by:
Athar Ansari,
Wayne State University, United States

Reviewed by:
Nigel Savery,
University of Bristol, United Kingdom
Harshad Ghodke,
University of Wollongong, Australia

****Correspondence:***
Laura A. Lindsey-Boltz
llindsey@med.unc.edu
Aziz Sancar
asancar@med.unc.edu

Specialty section:
This article was submitted to Protein and RNA Networks, a section of the journal Frontiers in Molecular Biosciences

Received: *15 February 2021*
Accepted: *14 April 2021*
Published: *20 May 2021*

Citation:
Lindsey-Boltz LA and Sancar A (2021) The Transcription-Repair Coupling Factor Mfd Prevents and Promotes Mutagenesis in a Context-Dependent Manner. Front. Mol. Biosci. 8:668290. doi: 10.3389/fmolb.2021.668290

MUTATION FREQUENCY DECLINE

The "mutation frequency decline" (MFD) phenomenon was discovered by Evelyn Witkin 65 years ago (Witkin, 1956). Notably, this was 4 years before the discovery of *Escherichia coli* RNA polymerase (RNAP; Hurwitz et al., 1960), and several years before it was even known that thymine dimers were the major UV lesions in *E. coli* DNA (Wacker et al., 1962) and that such dimers are repaired in *E. coli* either by a visible light–dependent photoreactivating enzyme (Rupert et al., 1958),

Abbreviations: Mfd, mutation frequency decline; RNAP, RNA polymerase; NER, nucleotide excision repair; GGR, global genome repair; TCR, transcription-coupled repair; TRCF, transcription-repair coupling factor; CPD, cyclobutane pyrimidine dimer; XR-seq, excision repair-sequencing; ChIP-seq, chromatin immunoprecipitation followed by high-throughput sequencing; RID, RNA polymerase interaction domain; EM, electron microscopy.

later named photolyase (Sancar, 2008), or by another mechanism called nucleotide excision repair (NER; Setlow and Carrier, 1964). The MFD phenomenon describes the observation that the yield of UV-induced mutations in specific auxotrophic *E. coli* strains is dependent on the number of nutrients present during the first cell division after irradiation. In other words, when Witkin briefly held the UV-irradiated auxotrophic strain for a few minutes under a condition where protein synthesis was inhibited before plating them onto rich agar plates, a decrease in the frequency of mutations was observed. She found that protein synthesis-inhibiting posttreatments that caused MFD, such as incubation in low nutritional media or the addition of the protein synthesis inhibitor chloramphenicol, did not affect the overall survival or change the yields of other kinds of mutations. Witkin went on to isolate a mutant *E. coli* strain, *mfd*-, that failed to mutate to auxotrophy under the prescribed conditions (Witkin, 1966), and then, 25 years later, she sent this strain to the Sancar Lab where Christopher Selby determined that the strain lacks the transcription-repair coupling activity (Selby et al., 1991) that he had been characterizing (Selby and Sancar, 1990, 1991). Thus, Mfd was the long sought-after *E. coli* transcription-repair coupling factor (TRCF) (reviewed in Selby, 2017).

TRANSCRIPTION-COUPLED REPAIR

Nucleotide excision repair is a versatile DNA repair pathway that removes all types of DNA-helix-distorting lesions (Sancar et al., 2004). NER occurs *via* two pathways: the predominant pathway, global genome repair (GGR), functions throughout the whole genome and a sub-pathway, transcription-coupled repair (TCR), specifically repairs the template strand of genes by acting upon lesions that block transcription by RNAP. NER plays a critical role in genome integrity, and both NER pathways can be found in all domains of life; however, the excision repair proteins in prokaryotes are not evolutionarily related to those in eukaryotes.

Using defined *in vitro* systems, the Sancar Lab has elucidated the reaction mechanisms of both global and TCRs in *E. coli* (reviewed in Sancar, 2016). Except for differences in initial damage recognition, both NER pathways are essentially the same in that they utilize the excision repair proteins UvrA, UvrB, and UvrC to perform dual incisions that remove the DNA damage in the form of 12- to 13-nucleotide oligomers and then utilize the proteins UvrD (helicase II), DNA polymerase I, and DNA ligase to release the excised oligomer, resynthesize the resulting gap, and ligate, respectively (**Figure 1**).

Global NER is initiated when the damage recognition factor, UvrA, which exists as a dimer together in a complex with UvrB (denoted $UvrA_2UvrB$), facilitates the formation of a stable UvrB–DNA complex in an ATP hydrolysis–dependent reaction (Hu et al., 2017). UvrA then disassociates from the complex, and UvrB recruits the UvrC endonuclease to the damage site. UvrC is a multidomain nuclease, which first incises the DNA at the 3rd or 4th phosphodiester bond 3′ to the lesion *via* its GIY-YIG catalytic domain and then at the 7th phosphodiester bond 5′ to the lesion using its C-terminal RNase H-like catalytic domain. The UvrD helicase then displaces the excised damaged strand.

Damage recognition is the rate-limiting step in NER, and some damage, such as the UV light–induced cyclobutane pyrimidine dimer (CPD), causes only minimal distortion to the DNA double helix and is thus poorly recognized by GGR (Hu et al., 2017). As a result, repair of such damage is greatly facilitated by RNAP scanning the DNA to initiate the repair of these lesions. When RNAP encounters DNA damage, it forms a stable complex at the damage site that inhibits repair by interfering with the access of $UvrA_2UvrB$ to the damage (Selby and Sancar, 1990). Mfd recognizes stalled RNAP and displaces it from the damage site while concomitantly recruiting $UvrA_2UvrB$ (Selby and Sancar, 1993a). Even though the Mfd protein has been extensively studied for nearly three decades since it was cloned (Selby and Sancar, 1993a) and characterized (Selby and Sancar, 1993b, 1994, 1995a,b) by Selby, several recent reports have significantly advanced our understanding of Mfd, including discoveries from whole-genome analyses (Adebali et al., 2017a,b; Ragheb et al., 2021), as well as from structural (Brugger et al., 2020; Kang et al., 2021) and single-molecule (Fan et al., 2016; Ho et al., 2018, 2020; Ghodke et al., 2020) studies which will be reviewed here.

RECENT ADVANCES: WHOLE-GENOME STUDIES

Transcription-coupled repair in *E. coli* was first described by the Hanawalt Lab when they reported 10-fold faster repair of the transcribed strand of the *lac* operon (Mellon and Hanawalt, 1989). Although this was subsequently confirmed with the analysis of several other *E. coli* genes in the 30 years since the original report, a significant advance in the field occurred when the Sancar Lab recently developed a method named eXcision Repair-sequencing (XR-seq) to map NER events throughout the whole genome at single nucleotide resolution (Hu et al., 2015) and employed this method to map CPD repair in *E. coli* (Adebali et al., 2017a,b). Briefly, the XR-seq method consists of purifying excised damaged oligos by immunoprecipitation with CPD-specific antibodies, ligating the isolated DNA to adapters, repairing the CPDs by photoreactivation, amplifying the DNA by PCR, and then next-generation sequencing and mapping the reads to the genome (Hu et al., 2017).

The Sancar Lab generated XR-seq CPD repair maps from several different *E. coli* strains, including *mfd*- and *uvrD*-, to assess the roles of these proteins in TCR (Adebali et al., 2017a,b). The maps revealed a rather complex genome-wide pattern of repair in the regions of annotated genes because of the widespread antisense transcription throughout most of the *E. coli* genome. Nevertheless, it was clear that Mfd is required for TCR. In fact, they found that the nontemplate strand is preferentially repaired in the *mfd*- strain, likely due to the interference of damage recognition by $UvrA_2UvrB$ when RNAP is stalled at the damage sites in the template strand. In contrast, TCR slightly increased in the absence of UvrD, consistent with the role of UvrD in the catalytic turnover of the Uvr(A)BC excision nuclease.

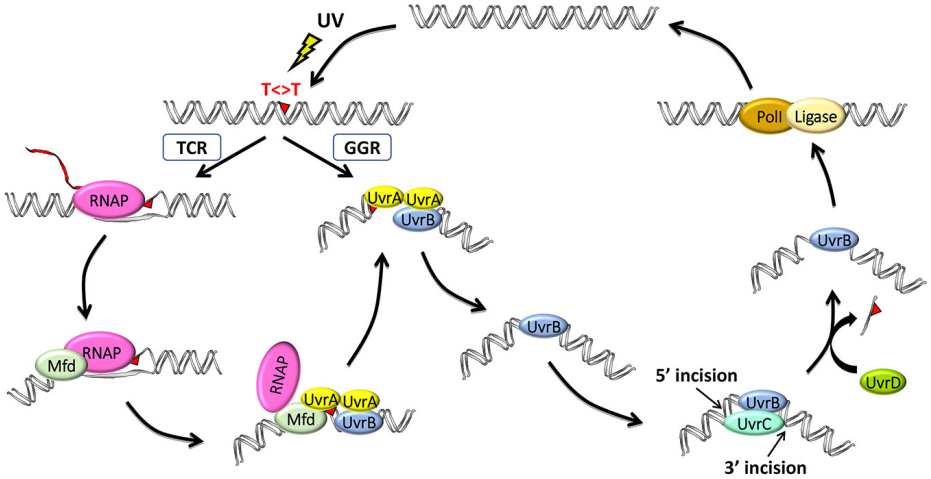


FIGURE 1 | Model for the two nucleotide excision repair pathways in *E. coli*: general global repair (GGR) and transcription-coupled repair (TCR). UV light induces thymine dimers in DNA which are either directly recognized by $UvrA_2B$ in the GGR pathway or indirectly recognized by RNA polymerase (RNAP) in the TCR pathway. Elongating RNAP stalls when it encounters a dimer in the template strand and recruits the mutation frequency decline (Mfd) translocase, which, in turn, removes RNAP while recruiting $UvrA_2B$. The two pathways then converge after these initial damage-recognition steps, and $UvrA_2$ dissociates, leaving a stable preincision complex consisting of UvrB bound to damaged DNA, which now has an altered structure. UvrC is recruited to generate the coupled dual incisions, and UvrD removes UvrC and the damaged oligonucleotide. Repair is completed by synthesis and ligation of the repair patch by DNA polymerase I (PolI) and DNA ligase, respectively.

Following the dual incisions, the UvrB-UvrC-excised oligomer complex remains bound to the duplex, and this complex is displaced by the UvrD helicase to release UvrC, the limiting repair factor, for new rounds of repair (**Figure 1**). In the *uvrD*-strain, there is a higher yield of recovered excised oligos due to their protection from nucleases when complexed with UvrB–UvrC, yet there is lower overall repair, and the slight increase in TCR seen in the *uvrD*- strain is due to Mfd facilitating the first and the only round of repair in the template strand. These same findings were also observed when XR-seq was used to analyze the *lac* operon under conditions where *lacZ* is either not expressed, in glucose-containing medium, or expressed, by the addition of isopropyl β-d-1-thiogalactopyranoside (IPTG; Adebali et al., 2017b). They found that the transcribed strand of *lacZ* is repaired ∼5-fold faster in the wild-type strain. In the *mfd*- strain, TCR was abolished and the nontranscribed strand was repaired in a higher level, whereas in the *uvrD*- strain, TCR was slightly enhanced although the overall repair was reduced.

As previously mentioned, NER is evolutionarily conserved, and indeed, the generation of XR-seq repair maps of parental and *uvrD*- strains of another prokaryote, *Mycobacterium smegmatis*, which is a very close relative of the human pathogen *Mycobacterium tuberculosis*, demonstrated that the TCR repair mechanism in Mycobacteria is the same as in *E. coli* (Selby et al., 2020). In conclusion, genome-wide studies provide detailed repair maps that complement the curated transcription maps of *E. coli* and *M. smegmatis* and confirm a central role of the Mfd protein in coupling transcription to repair in prokaryotes.

A very recent *E. coli* whole-genome analysis from Houra Merrikh's lab mapped Mfd-associated genomic loci using chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) and found a very high correlation ($r = 0.98$) between sites bound by Mfd and those bound by RNAP (Ragheb et al., 2021). Interestingly, this study was performed in the absence of exogenous DNA damage, and the Mfd-bound sites correlated ($r = 0.6$) with the sites of the *E. coli* RNA secondary structure as determined by parallel analysis of RNA structure (PARS-seq). They compared RNAP chromatin association in the presence and absence of Mfd and found that 40% of the genes that had at least a twofold increase in RNAP association in the *mfd*- strain contained a regulatory RNA or structural element. This, together with their other results from experiments in *Bacillus subtilis*, led them to conclude that Mfd regulates RNAP in hard-to-transcribe regions such as those with structured RNAs. These and other discoveries have direct implications on the role of Mfd in bacterial virulence and the development of drug resistance and will be discussed further below.

RECENT ADVANCES: Mfd STRUCTURAL STUDIES

Recent progress on the structural biology of Mfd has provided significant details and resolution to our understanding of how this enzyme functions in TCR. Mfd is a multidomain protein composed of eight domains, namely, D1a, D1b, and D2–D7 (**Figure 2**). The x-ray crystallography structure of full-length Mfd (Deaconescu et al., 2006) showed that it normally exists in a repressed conformation with its C-terminal D7 domain interacting with its N-terminal D2 domain, which is homologous to the UvrA-interacting domain of UvrB. Multiple large conformational changes occur in Mfd when it engages with and displaces RNAP and while recruiting $UvrA_2UvrB$ to the DNA damage, and recent cryo-electron microscopy (cryo-EM)

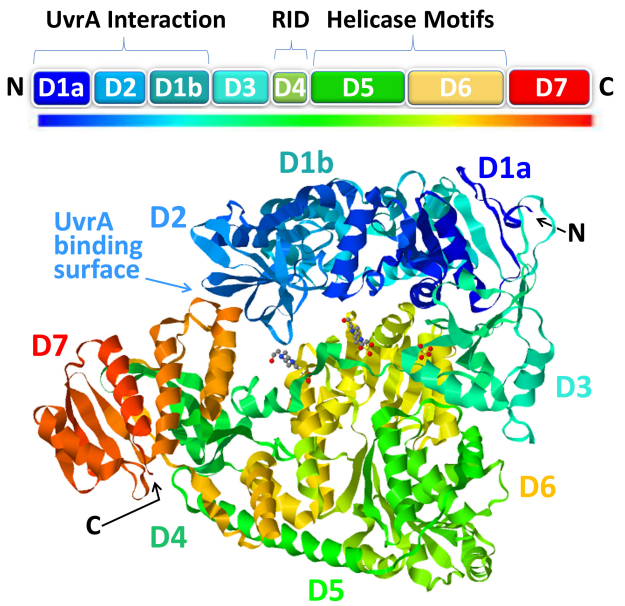


FIGURE 2 | The structure of mutation frequency decline (Mfd). The modular Mfd protein consists of eight domains indicated as boxes in the cartoon (top) and as a rainbow ribbon representation of the crystal structure [PDB ID: 2EYQ (Deaconescu et al., 2006)] viewed with JSmol (bottom). The N-terminus (N) contains the UvrA interaction region (D1a, D2, D1b) which is structurally homologous to the region of UvrB that binds to UvrA. This region of the protein is sequestered in a locked state *via* interactions with the D7 autoinhibitory domain in the C-terminus (C). Upon interaction with stalled RNAP, the D4 domain, containing the RNA polymerase (RNAP)-interacting domain (RID), binds to the β' subunit of RNAP which triggers ATP hydrolysis by the helicase motifs in D5-D6 and subsequent DNA translocation and rearrangements ultimately resulting in release of the nascent RNA, removal of RNAP, and recruitment of $UvrA_2B$ repair factors.

studies have provided enough high-resolution images of the intermediates to provide a clear understanding of this cycle (Brugger et al., 2020; Kang et al., 2021). The cryo-EM structures beautifully illustrate how Mfd is remodeled from its repressed conformation to expose the UvrA-interacting surface in D2, which is hidden during most of the remodeling process to prevent premature interactions with $UvrA_2UvrB$. Domain D4 of Mfd, which contains the RNAP interaction domain (RID), interacts with the β-subunit of RNAP, and the EM images demonstrate how Mfd engages with the RNAP bound to DNA damage (Kang et al., 2021). Although initial binding of the RID to the RNAP does not require conformational changes, at least one round of ATP hydrolysis is required for Mfd to form a stable complex with RNAP, and this allows tethering of the Mfd translocation module (domains D5 and D6) to the upstream duplex DNA (Kang et al., 2021). Domains D5 and D6 of the Mfd are homologous to the RecG bacterial motor protein that couples ATP hydrolysis to double-stranded DNA translocation; however, unlike RecG and other helicases, Mfd cannot separate the DNA strands (Selby and Sancar, 1995b). After the initial interaction of Mfd with stalled RNAP, a series of stepwise dynamic conformational changes is triggered resulting in Mfd completely encircling the upstream duplex DNA and culminating in the ATP-hydrolysis-powered disruption of the RNAP (Kang et al., 2021).

RECENT ADVANCES: Mfd SINGLE-MOLECULE STUDIES

Single-molecule approaches can be very useful in studying multicomponent, multistep reactions such as TCR. *In vitro* single-molecule experiments allow one to answer questions such as which proteins are present, what are their stoichiometries, and how quickly do they come and go (Strick and Portman, 2019). Many observations from *in vitro* Mfd single-molecule studies have confirmed and added a more detailed understanding to aspects of the mechanism that was determined by Selby's original population-averaging *in vitro* biochemistry experiments (Selby and Sancar, 1993a,b, 1994, 1995a,b), such as the observation that Mfd binds stalled RNAP and uses the energy from ATP hydrolysis to displace the stalled RNAP from DNA (Howan et al., 2012) and that the displacement of RNAP from DNA is accompanied by the loss of the nascent RNA (Graves et al., 2015; Kang et al., 2021). One surprising result is that the displaced RNAP remains in a long-lived complex with Mfd on the DNA, and in the absence of DNA damage, this Mfd–RNAP complex is capable of translocating thousands of base pairs in the same direction as the initial transcription (Howan et al., 2012; Graves et al., 2015). The interaction of Mfd with DNA induced bending or wrapping of the DNA, and it has been proposed that the high processivity of Mfd translocation is due to this topological wrapping. In a different single-molecule study, Mfd was added to DNA alone and was found to translocate for a few hundred base pairs (Le et al., 2018). However, as previously discussed, Mfd is thought to exist in a repressed state when not bound to stalled RNAP (Deaconescu et al., 2006, 2012) and has been shown to have only weak DNA-binding activity on its own (Selby and Sancar, 1995a), and thus, the physiological relevance is unclear. Single-molecule studies also showed that Mfd can rescue RNAP at pause sites, but more severe obstacles to RNAP movement such as DNA damage lead to eventual transcription termination (Le et al., 2018). The addition of either $UvrA_2$ or $UvrA_2B$ to the single-molecule system arrested the translocating Mfd–RNAP complex, and then, both Mfd and RNAP were released from the DNA (Fan et al., 2016). Then, with the further addition of UvrC, incision was observed in the damaged DNA, and the kinetics was in agreement with previous estimates of ∼3-fold faster repair by TCR than by GGR (Fan et al., 2016).

In vivo single-molecule experiments are very useful for analyzing the diffusion of proteins inside cells as they search for and bind to their targets (Strick and Portman, 2019). The van Oijen Lab fluorescently labeled Mfd in live *E. coli* and found that it interacts with RNAP even in the absence of exogenous DNA damage (Ho et al., 2018). The authors proposed that the interactions involved naturally stalled RNAP because they were enriched in the presence of a drug that stalls RNAP, they were absent in cells treated with a transcription inhibitor, and the presence of UvrA shortened the lifetime of the Mfd-RNAP-DNA complexes. In back-to-back follow-up reports, they analyzed

fluorescently labeled Mfd in UV-irradiated cells and also analyzed fluorescently labeled UvrA (Ghodke et al., 2020; Ho et al., 2020). They reported that the lifetime of the Mfd-RNAP-DNA complex decreased from ∼18 s, in the absence of exogenous DNA damage, to 12 s in UV-irradiated cells and that this mirrored what was seen with fluorescently labeled UvrA, and the lifetime of UvrA was dependent on the presence of Mfd indicating that the proteins function together (Ghodke et al., 2020). In the companion report, they employed ATPase mutants of UvrA and damage-recognition mutants of UvrB to analyze $UvrA_2B$ recruitment and Mfd dissociation *in vivo*. As predicted from earlier genetic and biochemistry studies, they found that Mfd is stably arrested on DNA in both mutant backgrounds relative to wild-type cells and concluded that Mfd dissociation is coupled with successful loading of UvrB (Ho et al., 2020). In conclusion, the *in vitro* and *in vivo* single-molecule studies on *E. coli* TCR provide detailed resolution that advances our understanding of this complex, multicomponent, and multistep reaction.

ROLES OF Mfd IN RECOMBINATION, BACTERIAL VIRULENCE, AND DEVELOPMENT OF DRUG RESISTANCE

As discussed above, the role of Mfd in *E. coli* TCR has been extensively characterized; however, many of the other reported cellular functions of Mfd are less well understood (Strick and Portman, 2019). There are reports suggesting that Mfd can enhance prokaryotic virulence and survival *via* the promotion of mutations in various genes involved in cell wall biosynthesis, translation, and transcription and has led to Mfd being called a "proevolutionary factor" (Strick and Portman, 2019) or "evolvability factor" (Ragheb et al., 2019; Brugger et al., 2020); however, there is no clear consensus as to the underlying mechanism. This has become a topic of wide interest due to the clinical implications of possibly targeting Mfd for antimicrobial drug resistance prevention, and recent results that shed some light on the subject will be discussed below.

When Witkin isolated the *mfd*-mutant strain, she reported that it produced ∼5-fold more UV-induced mutations than the parent strain even though it was no more sensitive to UV irradiation than its parent (Witkin, 1966). Witkin's assay specifically selected mutations in tRNA suppressor genes (Witkin, 1994), but later studies of the *lacI* gene also showed that the UV-induced mutation frequency was reduced ∼5-fold by Mfd and, at one particular site, by more than 300-fold (Oller et al., 1992). Thus, for DNA damage–induced mutagenesis, Mfd clearly functions as an antimutator; however for "spontaneous mutagenesis," it appears to function as a mutator. For example, the antimicrobial drugs to which resistance develops in an Mfd-dependent manner are not known as DNA-damaging agents, and thus, the function of Mfd was not clear in this phenomenon. Mfd plays a role in recombination (Ayora et al., 1996) and facilitates the generation of R-loops (Portman et al., 2021) which initiate DNA breakage and genome instability (Wimberly et al., 2013), consistent with the notion that Mfd facilitates transcription-associated mutagenesis (Ayora et al., 1996; Jinks-Robertson and Bhagwat, 2014; Gomez-Marroquin et al., 2016; Ragheb et al., 2019; Portman et al., 2021). This has recently gained support from the recent study discussed above (from the Merrikh Lab), showing that Mfd regulates RNAP in hard-to-transcribe regions such as those with structured RNAs (Ragheb et al., 2021). The authors analyzed the Mfd-bound genes and found that they were involved in a variety of cellular functions including toxin–antitoxin systems. Indeed, they went on to show that cell viability is compromised by the overexpression of toxin genes in the absence of Mfd and that the mutation rate of one particular toxin gene is lower by ∼7-fold in *mfd*- cells compared to wild-type cells. In conclusion, it is evident that Mfd plays a role in RNAP transcriptional control at regions of frequent RNAP pausing, and specific regions of the genome may be prone to transcription-associated mutagenesis due to inherent RNA structure. Thus, Mfd prevents and promotes mutagenesis in a context-dependent manner.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was supported by National Institutes of Health grant GM118102 to AS.

REFERENCES

Adebali, O., Chiou, Y. Y., Hu, J., Sancar, A., and Selby, C. P. (2017a). Genome-wide transcription-coupled repair in *Escherichia coli* is mediated by the Mfd translocase. *Proc. Natl. Acad. Sci. U.S.A.* 114, E2116–E2125.

Adebali, O., Sancar, A., and Selby, C. P. (2017b). Mfd translocase is necessary and sufficient for transcription-coupled repair in *Escherichia coli*. *J. Biol. Chem.* 292, 18386–18391. doi: 10.1074/jbc.c117.818807

Ayora, S., Rojo, F., Ogasawara, N., Nakai, S., and Alonso, J. C. (1996). The Mfd protein of *Bacillus subtilis* 168 is involved in both transcription-coupled DNA repair and DNA recombination. *J. Mol. Biol.* 256, 301–318. doi: 10.1006/jmbi.1996.0087

Brugger, C., Zhang, C., Suhanovsky, M. M., Kim, D. D., Sinclair, A. N., Lyumkis, D., et al. (2020). Molecular determinants for dsDNA translocation by the transcription-repair coupling and evolvability factor Mfd. *Nat. Commun.* 11:3740.

Deaconescu, A. M., Chambers, A. L., Smith, A. J., Nickels, B. E., Hochschild, A., Savery, N. J., et al. (2006). Structural basis for bacterial transcription-coupled DNA repair. *Cell* 124, 507–520. doi: 10.1016/j.cell.2005.11.045

Deaconescu, A. M., Sevostyanova, A., Artsimovitch, I., and Grigorieff, N. (2012). Nucleotide excision repair (NER) machinery recruitment by the transcription-repair coupling factor involves unmasking of a conserved intramolecular interface. *Proc. Natl. Acad. Sci. U.S.A.* 109, 3353–3358.

Fan, J., Leroux-Coyau, M., Savery, N. J., and Strick, T. R. (2016). Reconstruction of bacterial transcription-coupled repair at single-molecule resolution. *Nature* 536, 234–237. doi: 10.1038/nature19080

Ghodke, H., Ho, H. N., and van Oijen, A. M. (2020). Single-molecule live-cell imaging visualizes parallel pathways of prokaryotic nucleotide excision repair. *Nat. Commun.* 11:1477.

Gomez-Marroquin, M., Martin, H. A., Pepper, A., Girard, M. E., Kidman, A. A., Vallin, C., et al. (2016). Stationary-phase mutagenesis in stressed *Bacillus subtilis* cells operates by Mfd-dependent mutagenic pathways. *Genes* (*Basel*) 7:33. doi: 10.3390/genes7070033

Graves, E. T., Duboc, C., Fan, J., Stransky, F., Leroux-Coyau, M., and Strick, T. R. (2015). A dynamic DNA-repair complex observed by correlative single-molecule nanomanipulation and fluorescence. *Nat. Struct. Mol. Biol.* 22, 452–457. doi: 10.1038/nsmb.3019

Ho, H. N., van Oijen, A. M., and Ghodke, H. (2018). The transcription-repair coupling factor Mfd associates with RNA polymerase in the absence of exogenous damage. *Nat. Commun.* 9:1570.

Ho, H. N., van Oijen, A. M., and Ghodke, H. (2020). Single-molecule imaging reveals molecular coupling between transcription and DNA repair machinery in live cells. *Nat. Commun.* 11:1478.

Howan, K., Smith, A. J., Westblade, L. F., Joly, N., Grange, W., Zorman, S., et al. (2012). Initiation of transcription-coupled repair characterized at single-molecule resolution. *Nature* 490, 431–434. doi: 10.1038/nature11430

Hu, J., Adar, S., Selby, C. P., Lieb, J. D., and Sancar, A. (2015). Genome-wide analysis of human global and transcription-coupled excision repair of UV damage at single-nucleotide resolution. *Genes Dev.* 29, 948–960. doi: 10.1101/gad.261271.115

Hu, J., Selby, C. P., Adar, S., Adebali, O., and Sancar, A. (2017). Molecular mechanisms and genomic maps of DNA excision repair in *Escherichia coli* and humans. *J. Biol. Chem.* 292, 15588–15597. doi: 10.1074/jbc.r117.807453

Hurwitz, J., Bresler, A., and Diringer, R. (1960). The enzymic incorporation of Ribonucleotides into Polyribonucleotides and the effect of DNA. *Biochem. Biophys. Res. Commun.* 3, 15–19. doi: 10.1016/0006-291x(60)90094-2

Jinks-Robertson, S., and Bhagwat, A. S. (2014). Transcription-associated mutagenesis. *Annu. Rev. Genet.* 48, 341–359. doi: 10.1146/annurev-genet-120213-092015

Kang, J. Y., Llewellyn, E., Chen, J., Olinares, P. D. B., Brewer, J., Chait, B. T., et al. (2021). Structural basis for transcription complex disruption by the Mfd translocase. *eLife* 10:e62117.

Le, T. T., Yang, Y., Tan, C., Suhanovsky, M. M., Fulbright, R. M. Jr., Inman, J. T., et al. (2018). Mfd dynamically regulates transcription via a release and catch-up mechanism. *Cell* 173:1823. doi: 10.1016/j.cell.2018.06.002

Mellon, I., and Hanawalt, P. C. (1989). Induction of the *Escherichia coli* lactose operon selectively increases repair of its transcribed DNA strand. *Nature* 342, 95–98. doi: 10.1038/342095a0

Oller, A. R., Fijalkowska, I. J., Dunn, R. L., and Schaaper, R. M. (1992). Transcription-repair coupling determines the strandedness of ultraviolet mutagenesis in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 89, 11036–11040. doi: 10.1073/pnas.89.22.11036

Portman, J. R., Brouwer, G. M., Bollins, J., Savery, N. J., and Strick, T. R. (2021). Co-transcriptional R-loop formation by Mfd involves topological partitioning of DNA. *Proc. Natl. Acad. Sci. U.S.A.* 118:e2019630118. doi: 10.1073/pnas.2019630118

Ragheb, M. N., Merrikh, C., Browning, K., and Merrikh, H. (2021). Mfd regulates RNA polymerase association with hard-to-transcribe regions in vivo, especially those with structured RNAs. *Proc. Natl. Acad. Sci. U.S.A.* 118:e2008498118. doi: 10.1073/pnas.2008498118

Ragheb, M. N., Thomason, M. K., Hsu, C., Nugent, P., Gage, J., Samadpour, A. N., et al. (2019). Inhibiting the evolution of antibiotic resistance. *Mol. Cell* 73, 157–165.e5.

Rupert, C. S., Goodgal, S. H., and Herriott, R. M. (1958). Photoreactivation in vitro of ultraviolet-inactivated *Hemophilus influenzae* transforming factor. *J. Gen. Physiol.* 41, 451–471. doi: 10.1085/jgp.41.3.451

Sancar, A. (2008). Structure and function of photolyase and in vivo enzymology: 50th anniversary. *J. Biol. Chem.* 283, 32153–32157. doi: 10.1074/jbc.r800052200

Sancar, A. (2016). Mechanisms of DNA repair by Photolyase and excision nuclease (Nobel Lecture). *Angew. Chem. Int. Ed. Engl.* 55, 8502–8527. doi: 10.1002/anie.201601524

Sancar, A., Lindsey-Boltz, L. A., Unsal-Kacmaz, K., and Linn, S. (2004). Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu. Rev. Biochem.* 73, 39–85. doi: 10.1146/annurev.biochem.73.011303.073723

Selby, C. P. (2017). Mfd protein and transcription-repair coupling in *Escherichia coli*. *Photochem. Photobiol.* 93, 280–295. doi: 10.1111/php.12675

Selby, C. P., Lindsey-Boltz, L. A., Yang, Y., and Sancar, A. (2020). Mycobacteria excise DNA damage in 12- or 13-nucleotide-long oligomers by prokaryotic-type dual incisions and performs transcription-coupled repair. *J. Biol. Chem.* 295, 17374–17380. doi: 10.1074/jbc.ac120.016325

Selby, C. P., and Sancar, A. (1990). Transcription preferentially inhibits nucleotide excision repair of the template DNA strand in vitro. *J. Biol. Chem.* 265, 21330–21336. doi: 10.1016/s0021-9258(17)45364-6

Selby, C. P., and Sancar, A. (1991). Gene- and strand-specific repair in vitro: partial purification of a transcription-repair coupling factor. *Proc. Natl. Acad. Sci. U.S.A.* 88, 8232–8236. doi: 10.1073/pnas.88.18.8232

Selby, C. P., and Sancar, A. (1993b). Transcription-repair coupling and mutation frequency decline. *J. Bacteriol.* 175, 7509–7514. doi: 10.1128/jb.175.23.7509-7514.1993

Selby, C. P., and Sancar, A. (1993a). Molecular mechanism of transcription-repair coupling. *Science* 260, 53–58. doi: 10.1126/science.8465200

Selby, C. P., and Sancar, A. (1994). Mechanisms of transcription-repair coupling and mutation frequency decline. *Microbiol. Rev.* 58, 317–329. doi: 10.1128/mr.58.3.317-329.1994

Selby, C. P., and Sancar, A. (1995a). Structure and function of transcription-repair coupling factor. I. Structural domains and binding properties. *J. Biol. Chem.* 270, 4882–4889. doi: 10.1074/jbc.270.9.4882

Selby, C. P., and Sancar, A. (1995b). Structure and function of transcription-repair coupling factor. II. Catalytic properties. *J. Biol. Chem.* 270, 4890–4895. doi: 10.1074/jbc.270.9.4890

Selby, C. P., Witkin, E. M., and Sancar, A. (1991). *Escherichia coli* mfd mutant deficient in "mutation frequency decline" lacks strand-specific repair: in vitro complementation with purified coupling factor. *Proc. Natl. Acad. Sci. U.S.A.* 88, 11574–11578. doi: 10.1073/pnas.88.24.11574

Setlow, R. B., and Carrier, W. L. (1964). The disappearance of thymine dimers from DNA: an error-correcting mechanism. *Proc. Natl. Acad. Sci. U.S.A.* 51, 226–231. doi: 10.1073/pnas.51.2.226

Strick, T. R., and Portman, J. R. (2019). Transcription-coupled repair: from cells to single molecules and back again. *J. Mol. Biol.* 431, 4093–4102. doi: 10.1016/j.jmb.2019.05.040

Wacker, A., Dellweg, H., and Jacherts, D. (1962). Thymine dimerization and survival of bacteria. *J. Mol. Biol.* 4, 410–412.

Wimberly, H., Shee, C., Thornton, P. C., Sivaramakrishnan, P., Rosenberg, S. M., and Hastings, P. J. (2013). R-loops and nicks initiate DNA breakage and genome instability in non-growing *Escherichia coli*. *Nat. Commun.* 4:2115.

Witkin, E. M. (1956). Time, temperature, and protein synthesis: a study of ultraviolet-induced mutation in bacteria. *Cold Spring Harb. Symp. Quant. Biol.* 21, 123–140. doi: 10.1101/sqb.1956.021.01.011

Witkin, E. M. (1966). Radiation-induced mutations and their repair. *Science* 152, 1345–1353. doi: 10.1126/science.152.3727.1345

Witkin, E. M. (1994). Mutation frequency decline revisited. *Bioessays* 16, 437–444. doi: 10.1002/bies.950160613

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Molecular Biosciences

MINI REVIEW
published: 25 June 2021
doi: 10.3389/fmolb.2021.681865

Regulation of Eukaryotic RNAPs Activities by Phosphorylation

Araceli González-Jiménez[1], *Adrián Campos*[1], *Francisco Navarro*[2,3], *Andrés Clemente-Blanco*[1] *and Olga Calvo*[1]*

[1]Instituto de Biología Funcional y Genómica, Consejo Superior de Investigaciones Científicas, Universidad de Salamanca, Salamanca, Spain, [2]Departamento de Biología Experimental-Genética, Universidad de Jaén, Jaén, Spain, [3]Centro de Estudios Avanzados en Aceite de Oliva y Olivar, Universidad de Jaén, Jaén, Spain

Evolutionarily conserved kinases and phosphatases regulate RNA polymerase II (RNAPII) transcript synthesis by modifying the phosphorylation status of the carboxyl-terminal domain (CTD) of Rpb1, the largest subunit of RNAPII. Proper levels of Rpb1-CTD phosphorylation are required for RNA co-transcriptional processing and to coordinate transcription with other nuclear processes, such as chromatin remodeling and histone modification. Whether other RNAPII subunits are phosphorylated and influences their role in gene expression is still an unanswered question. Much less is known about RNAPI and RNAPIII phosphorylation, whose subunits do not contain functional CTDs. However, diverse studies have reported that several RNAPI and RNAPIII subunits are susceptible to phosphorylation. Some of these phosphorylation sites are distributed within subunits common to all three RNAPs whereas others are only shared between RNAPI and RNAPIII. This suggests that the activities of all RNAPs might be finely modulated by phosphorylation events and raises the idea of a tight coordination between the three RNAPs. Supporting this view, the transcription by all RNAPs is regulated by signaling pathways that sense different environmental cues to adapt a global RNA transcriptional response. This review focuses on how the phosphorylation of RNAPs might regulate their function and we comment on the regulation by phosphorylation of some key transcription factors in the case of RNAPI and RNAPIII. Finally, we discuss the existence of possible common mechanisms that could coordinate their activities.

Keywords: phosphorylation, transcription regulation, gene expression, RNA polymerase I, RNA polymerase II, RNA polymerase III

OPEN ACCESS

Edited by:
Silvia M. L. Barabino, University of Milano-Bicocca, Italy

Reviewed by:
Jose C. Reyes, Andalusian Center of Molecular Biology and Regenerative Medicine (CABIMER), Spain
Joe Reese, Pennsylvania State University (PSU), United States

***Correspondence:**
Olga Calvo
ocalvo@usal.es

Specialty section:
This article was submitted to Protein and RNA Networks, a section of the journal Frontiers in Molecular Biosciences

Received: *17 March 2021*
Accepted: *07 June 2021*
Published: *25 June 2021*

Citation:
González-Jiménez A, Campos A, Navarro F, Clemente-Blanco A and Calvo O (2021) Regulation of Eukaryotic RNAPs Activities by Phosphorylation. Front. Mol. Biosci. 8:681865. doi: 10.3389/fmolb.2021.681865

INTRODUCTION

The transcription of cellular RNAs is carried out by DNA-dependent RNA polymerases (RNAPs). In bacteria and archaea, only one RNAP transcribes all RNAs. In Eukarya, three RNAPs (RNAPI, -II and -III) are required for RNA transcription, except plants containing two other RNAPs (RNAPIV and -V). RNAPI synthesizes the precursor ribosomal RNA (rRNA 35S in yeast, 47S in mammals), RNAPIII produces 5S rRNA and transfer RNAs (tRNAs) and RNAPII transcribes all the protein-coding genes synthesizing messenger RNAs (mRNAs). Additionally, RNAPII and RNAPIII can synthesize other types of transcripts, such as small non-coding RNAs (ncRNAs), whose specific synthesis may differ depending on the species (Huet et al., 1985; Dieci et al., 2007). Finally, RNAPIV and RNAPV produce small interfering (siRNAs) and ncRNAs in plants (Onodera et al., 2005; Zhang et al., 2007; Haag and Pikaard, 2011; Lopez et al., 2011; Wang and Ma, 2015). All RNAPs are related

at the evolutionary level, displaying common structures and functions. The minimum preserved structure of RNAPs is that of bacteria, consisting of five subunits. Archaeal RNAP has 12 subunits and eukaryotic RNAPs are complexes of 12 (RNAPII), 14 (RNAPI) and 17 (RNAPIII) subunits (Cramer et al., 2008; Werner and Grohmann, 2011; Wang and Ma, 2015; Cramer, 2019b). They all have a structurally conserved core formed by 10 subunits, with additional factors located on the polymerase complex periphery. Moreover, they all share five subunits (Rpb5, Rpb6, Rpb8, Rpb10, and Rpb12) with common functions but also with specific roles in their corresponding RNAPs (Cramer et al., 2008; Cuevas-Bermúdez et al., 2017). The structures of the three eukaryotic RNAPs, first solved in *Saccharomyces cerevisiae,* are highly conserved and their resolution has tremendously helped to understand the mechanism of transcription (Cramer et al., 2000; Armache et al., 2005; Engel et al., 2013; Fernandez-Tornero et al., 2013; Hoffmann et al., 2015; Sainsbury et al., 2015; Ramsay et al., 2020; Schier and Taatjes, 2020). The correct regulation of gene transcription depends on mechanisms that regulate the formation of large multiprotein complexes (RNAPs and their cognate factors) and their dynamics through all the transcription process. One of the most prominent mechanisms is post-translational modification (PTM) of proteins (Deribe et al., 2010), phosphorylation being the most frequent (Beltrao et al., 2013). A clear example is the dynamic phosphorylation of the carboxyl-terminal domain (CTD) of Rpb1, key for gene transcription (Buratowski, 2009; Calvo and García, 2012; Hsin and Manley, 2012; Eick and Geyer, 2013; Harlen and Churchman, 2017). Unfortunately, while most of the available data refer mainly to the phospho-regulation of transcription factors implicated in the modulation of all RNAP activities, little is known about the phosphorylation of other RNAP subunits and their implications in RNA biogenesis. Here, we have compiled all the phospho-sites identified to date for *S. cerevisiae* and human RNAPs (**Supplementary Tables S1, S2**). We discuss the localization and possible roles of the three RNAP subunit phosphorylations in budding yeast, as the structures of the different transcription complexes are better known in this organism. Finally, we review the possible conservation of RNAP phospho-regulation with evolution.

RNAPII PHOSPHORYLATION

RNAPII is the best known of the eukaryotic RNA polymerases. Transcription by RNAPII is a very complex, dynamic and finely regulated process. A sophisticated network of protein–protein and protein–nucleic acid interactions is established, producing conformational and activity changes in RNAPII through the transcription cycle. Thus, a pre-initiation complex (PIC), composed basically of general transcription factors (GTFs: TFIIA, B, D, E, F, and H), Mediator and RNAPII, is assembled at the gene promoters, opening the DNA to initiate transcription (Greber and Nogales, 2019; Schier and Taatjes, 2020). Other factors acting as activators/co-activators and repressors/co-repressors can modulate the transcription activity (Ho and Shuman, 1999; Thomas and Chiang, 2006; Hahn and Young, 2011; Roeder, 2019). Subsequently, RNAPII activity is regulated by elongation and termination factors (Kwak and Lis, 2013). Because pre-mRNA maturation (capping, splicing and polyadenylation) occurs co-transcriptionally, a set of processing factors also interacts with the transcription machinery. Moreover, chromatin and histone modifiers act to facilitate and regulate the passage of RNAPII through the genes being transcribed in concert with the transcription complex. It is well known that the correct orchestration of all these processes involved in mRNA biogenesis is coordinated and fine-tuned by the phosphorylation status of the Rpb1-CTD (Perales and Bentley, 2009; Calvo and García, 2012; Hsin and Manley, 2012; Harlen and Churchman, 2017).

Functional and structural studies with *S. cerevisiae* have provided the majority of the existing knowledge about RNAPII transcription mechanisms, regulation and coordination with other cellular processes (Cramer, 2019a; Cramer, 2019b; Roeder, 2019). Recent structural data combined with functional studies have advanced our understanding of RNAPII transcription in general and that of PIC function, structure and dynamics in particular (Greber and Nogales, 2019; Schier and Taatjes, 2020). Resolution of the RNAPII structure by X-ray crystallography about 20 years ago showed that its twelve subunits are folded and assembled into four mobile modules: the *core* module, formed by the active center (Rpb1 and Rpb2) and assembly platform (Rpb3, Rpb10, Rpb11, and Rpb12); the *jaw-lobe* module, made up of Rpb1 and Rpb9; the *shelf* module containing the foot and cleft domains of Rpb1 and the lower jaw and assembly domains of Rpb5; and the *stalk* module, formed by Rpb4 and Rpb7, which in the case of *S. cerevisiae* can be dissociated from the 10-subunit core polymerase (Cramer et al., 2000; Cramer et al., 2001; Gnatt et al., 2001; Armache et al., 2003; Bushnell and Kornberg, 2003). Within these modules there are some key structural domains with basic roles in transcription, such as the active site, cleft, clamp, wall, protrusion, funnel and RNA exit channel. Movement of these regions is accompanied by binding of the GTFs with essential roles in transcription initiation. Subsequent binding of elongation factors replaces the GTFs, thus regulating further steps of the transcription cycle. How all these events take place is not fully understood, although some are explained by conformational changes of the transcription complex and/or phosphorylation of specific factors and the Rpb1-CTD (Wang et al., 2010; Larochelle et al., 2012; He et al., 2013; Sainsbury et al., 2015; He et al., 2016; Harlen and Churchman, 2017; Nogales et al., 2017; Greber and Nogales, 2019; Nogales and Greber, 2019; Patel et al., 2019; Schier and Taatjes, 2020).

The Rpb1-CTD is an unstructured and flexible domain that is crucial for the regulation of RNAPII transcription. It consists of multiple repeats of the heptapeptide sequence Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7, which is not present in other RNAPs. It is evolutionarily well conserved from protozoa to metazoa. The number of repeats ranges from 26 repetitions in *S. cerevisiae* to 52 in mammals (Corden et al., 1985; Chapman et al., 2008). Five of the seven residues are susceptible to phosphorylation: Tyr1; Ser2, -5, and -7; and Thr4 (Buratowski, 2003; Hsin et al., 2011; Hsin

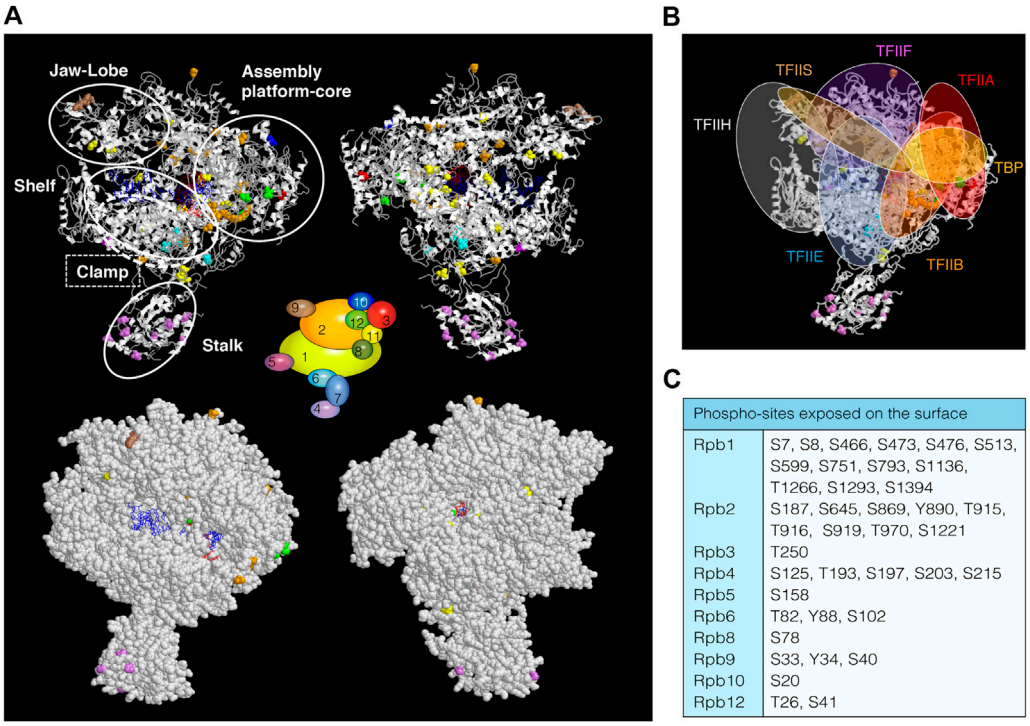


FIGURE 1 | RNAPII phospho-sites. **(A)** Upper, schematic views (ribbon representation) of *Saccharomyces cerevisiae* RNAPII (PDB: 1y1w), displaying phospho-sites that have been labelled in different colours according to the 12 subunits diagram shown in the middle of the figures. RNAPII mobile modules are indicated with white open circles. DNA is represented in blue and RNA in red. Surface ***(bottom)*** views showing exposed phospho-sites. **(B)** Schematic representation of GTFs localizations according to published works (i.e., (Sainsbury et al., 2015; Schier and Taatjes, 2020)). **(C)** Table with phospho-sites exposed on the surface of RNAPII whose phosphorylation status could be important for the association/dissociation of transcription regulators.

and Manley, 2012; Allepuz-Fuster et al., 2014; Yurko and Manley, 2018). Prior to transcription initiation, the Rpb1-CTD likely interacts with Mediator and helps to recruit it to the promoter (Kim et al., 1994; Naar et al., 2002; Robinson et al., 2016). During initiation, the CTD becomes phosphorylated by transcription-associated kinases, generating phospho-marks required for the binding of elongation and RNA processing factors, among others, and to proceed to a productive elongation phase. As this subject has been extensively reviewed (Perales and Bentley, 2009; Calvo and García, 2012; Hsin and Manley, 2012; Harlen and Churchman, 2017), we will focus on this section on the phosphorylations of other RNAPII subunits (**Supplementary Table S1**), although in most cases a role in transcription and/or RNA processing has not been yet stated.

Several phospho-proteomic studies have identified at least 75 new phospho-sites in 10 of the 12 subunits of *S. cerevisiae* RNAPII (**Supplementary Table S1**), 55 of them distributed along specific RNAPII subunits (Rpb1, Rpb2, Rpb3 Rpb4, and Rpb9) and 20 in shared subunits with RNAPI and RNAPIII (Rpb5, Rpb6, Rpb8, Rpb10, and Rpb12) (Albuquerque et al., 2008; Pultz et al., 2012; Swaney et al., 2013; Sostaric et al., 2018; MacGilvray et al., 2020; Lanz et al., 2021; Richard et al., 2021). However, it is unknown whether all these residues are phosphorylated *in vivo* and if they form part of a regulatory mechanism to control the biogenesis of RNAPII transcripts. Localization of these residues on the RNAPII structure (**Figure 1A**, upper panel) shows a broad distribution, with 50 of the 75 phospho-sites localized in structured regions. Notably, many of them correspond to defined RNAPII regions, suggesting that post-translational modifications by phosphorylation may influence how these specific regions act during the transcription steps (i.e., DNA contact, NTP addition, clamp movement, etc.). It is worth noting that 39 phospho-sites are exposed on the surface of the enzyme (**Figure 1A**, bottom panel, and **Figure 1C**). Thus, it is tempting to speculate that the phosphorylation of these residues might be important for protein–protein interaction between RNAPII and different transcriptional regulators. In fact, these exposed residues localize in regions that are described to contact the GTFs (TFIIB, TBP, TFIIA, TFIIE, TFIIF, TFIIH or TFIIS; **Figure 1B**), Mediator and elongation factor Spt5/4 (Cai et al., 2010; Martinez-Rucobo et al., 2011; Nogales et al., 2017; Greber and Nogales, 2019; Schier and Taatjes, 2020). Interestingly, there are three phospho-sites (S1793, T1471, and Y1473) within the Rpb1 *linker*, an unstructured region that connects the CTD to the rest of the protein, whose phosphorylation is required for Spt6 interaction with RNAPII and the re-assembly of repressive chromatin during transcription (Sdano et al., 2017).

Rpb2 phosphorylation sites lie in the external 1, protrusion, fork, wall, hybrid binding and anchor domains. In the wall, near

the active site, the RNA:DNA hybrid separates and upstream DNA makes a 90° turn to exit RNAPII (Cramer et al., 2001). The protrusion is an external, positively charged domain, placed above the wall where the DNA exits from the cleft. Re-annealing of transcribed DNA occurs as it exits the enzyme, and the protrusion may participate in this process. Therefore, phosphorylation of residues lying in these domains could be involved in the separation of the RNA:DNA hybrid, the re-annealing of the transcribed DNA as it exits the enzyme and/or in the association/dissociation of transcription and processing factors (Pappas and Hampsey, 2000). Another example is the association and function of TFIIB and TFIIF during transcription initiation, factors important to position the DNA over the RNAPII active center cleft (Sainsbury et al., 2015; Greber and Nogales, 2019; Schier and Taatjes, 2020). Indeed, TFIIB interacts with the clamp, with the dock and cleft (Rpb1), and with the wall and protrusion domains (Rpb2) (Kostrewa et al., 2009; Liu et al., 2010; Sainsbury et al., 2015). TFIIF binds upstream and downstream DNA and RNAPII near the Rpb2 lobe and protrusion domains (Muhlbacher et al., 2014; Plaschka et al., 2015). Again, phosphorylation of residues in TFIIB and TFIIF binding regions could be involved in their association with the RNAPII.

Rpb4 and Rpb7 form a heterodimer known as the stalk domain, and only Rpb4 contains phosphorylation sites (Richard et al., 2021). The stalk extends from the foot domain at the base of the RNAPII enzyme and its movement helps to coordinate opening and closing of the clamp (Armache et al., 2003; Bushnell and Kornberg, 2003). It is contacted by initiation and elongation factors (Cai et al., 2010; Martinez-Rucobo et al., 2011; Li et al., 2014; Plaschka et al., 2015; Greber and Nogales, 2019; Schier and Taatjes, 2020). Rpb4 contains several phospho-sites whose phosphorylation may be important for interaction with Rpb7 and/or the 10-subunit polymerase. Accordingly, the Rpb4 S125 residue resides within a region exclusively present in *S. cerevisiae* that could regulate specific functions in this organism, such as dissociation of Rpb4/7 from the core polymerase (Sharma and Kumari, 2013; Duek et al., 2018). Moreover, exposed residues could mediate the association of different factors (Babbarwal et al., 2014; Garavis et al., 2017; Allepuz-Fuster et al., 2019; Calvo, 2020), depending on their phosphorylation status, such as TFIIE, TFIIF, Mediator (Cai et al., 2010; Sainsbury et al., 2015), Spt5/4 (Martinez-Rucobo et al., 2011; Li et al., 2014) some CTD phosphatases (Kimura et al., 2002; Allepuz-Fuster et al., 2014), and termination factors (Mitsuzawa et al., 2003; Runner et al., 2008), and thus regulate the function of Rpb4/7. Similarly to Rpb4/7, Rpb3 forms a heterodimer with Rpb11 (Cramer et al., 2001). Moreover, some phospho-sites fall in a region comprising the heterodimerization domain of Rpb3. This suggests that phosphorylation of this region might be important for the formation of the heterodimer. Rpb9 phosphorylated residues localized in the jaw and linker domains (Cramer et al., 2001) and, because TFIIH interacts with RNAPII near the Rpb9 jaw, we could speculate that modification of these residues could be functionally linked to this factor (Muhlbacher et al., 2014; Plaschka et al., 2015).

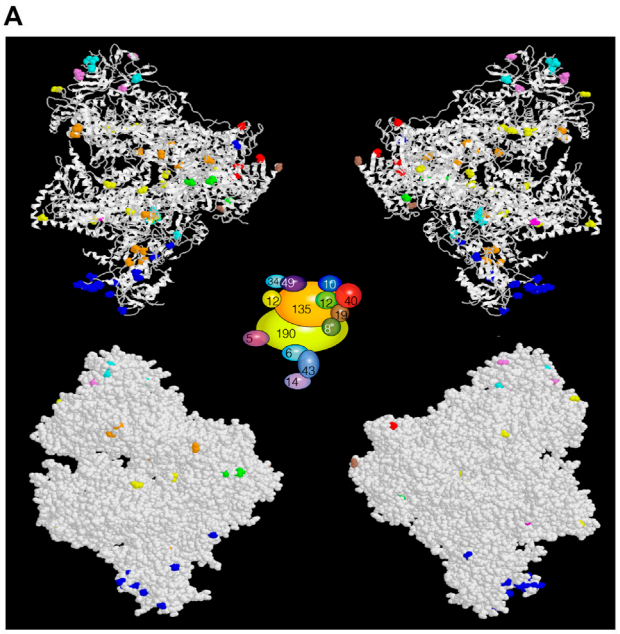
The subunits shared by the three RNAPs (Rpb5, Rpb6, Rpb8, Rpb10, and Rpb12) are also phospho-proteins (**Supplementary Table S1**) (Albuquerque et al., 2008; Swaney et al., 2013; Sostaric et al., 2018; MacGilvray et al., 2020; Lanz et al., 2021). For instance, Rpb5 and Rpb6 contain phospho-sites (S158, and Y88 and T82, respectively) localized in regions important for Rpb5 and Rpb6 assembly to RNAPII (Cramer et al., 2001; Tan et al., 2003; Zaros et al., 2007).

Modification by phosphorylation of some residues of RNAPII subunits could be important not only for the association of different factors along the transcription cycle but also for exchange of factors occupying the same or close surfaces on RNAPII. This is the case of initiation and elongation factors that compete during the transcription cycle for binding to the polymerase complex, for instance TFIIE and Spt5 (Li et al., 2014). How these mutually exclusive interactions of the transcription factors with RNAPII are regulated without affecting the efficiency of all the transcription steps (initiation, pausing and elongation) remains to be understood. Recently, it has been shown that the Rpb1-CTD undergoes liquid-phase separation, which could explain the association of initiation and elongation factors (Boehning et al., 2018; Cramer, 2019b). First, a dynamic condensate is formed near the promoter during initiation that contains a non-phosphorylated RNAPII and initiation factors. This condensate facilitates transcription initiation, RNA synthesis and Rpb1-CTD phosphorylation. Second, a transient condensate containing phosphorylated RNAPII and elongation factors is produced and maintained until RNAPII reaches the end of the genes, where RNAPII is dephosphorylated, recycled and transferred to the first condensate. As the transfer of RNAPII from one condensate to another is controlled by CTD phosphorylation, it is possible that this mechanism might be crucial for optimal transcriptional regulation (Boehning et al., 2018; Guo et al., 2019; Peng et al., 2020). However, we cannot rule out that the exchange of factors during the initiation/elongation transition could be regulated by the phosphorylation of other RNAPII subunits and/or even that the transfer of RNAPII between both condensates could require the post-translational modification of additional subunits.

Finally, the high sequence conservation within the RNAPII core between yeast and humans suggests similar mechanisms of RNA synthesis. However, sequences are more divergent toward the exterior/surface residues, suggesting that biochemically distinct interfaces interact with different factors (Cramer et al., 2001; He et al., 2013; He et al., 2016; Nogales et al., 2017; Schier and Taatjes, 2020). Accordingly, phospho-sites localized in the surface of RNAPII may contribute to the association/dissociation of species-specific factors.

RNAPI PHOSPHORYLATION

Initially, 15 phospho-sites were identified in *S. cerevisiae* distributed to five of the 14 subunits. Mutation of 13 of these phospho-sites indicated that most are non-essential PTMs,



Phospho-sites exposed on the surface	
Rpa190	S174, S354, T447, S664, S685 S1374, S1383
Rpa135	S81, S86, S88, S115, S507, S737, S1054, S1124, S1156
Rpc40	S46
Rpc19	S141
Rpb5	S158
Rpb6	T82, Y88, S102
Rpb8	S78
Rpa12	S40, S59, S70
Rpb10	S20
Rpb12	T26, S41
Rpa43	T25, T100, S220, S244, S251, Y260, S262, S263, S265, S269, S285, S306
Rpa49	S8, S34, T36, S70
Rpa34	S60, S113, S115, S118

Phospho-sites exposed on the surface	
Rpc128	T40
Rpc40	S46, S226, S243
Rpc19	S141
Rpb5	S158
Rpb6	T82, Y88, S102
Rpb8	S78
Rpb10	S20
Rpb12	T26, S41
Rpc25	S162
Rpc82	S230, S306, S323, S344
Rpc37	T192

FIGURE 2 | RNAPI and RNAPIII phospho-sites. **(A)** Ribbon ***(upper)*** and surface ***(bottom)*** schematic views of RNAPI (PDB: 4c3h) from *Saccharomyces cerevisiae*, displaying phospho-sites labelled in different colours according to the subunit diagram shown in the middle of the figures. **(B)** Schematic representation of Rrn3 localization (Torreira et al., 2017). **(C)** Table with phospho-sites exposed on the surface of RNAPI. **(D)** Ribbon ***(upper)*** and surface ***(bottom)*** representations of RNAPIII (PDB: 5fj9) displaying coloured phospho-sites. **(E)** Schematic representation of Maf1 and Brf1 associations with RNAPIII (Vorlander et al., 2020a; Vorlander et al., 2020b). **(F)** Table with RNAPIII phospho-sites exposed on the surface. As in the case of RNAPII, these residues could be important for the interaction with transcription regulators.

suggesting that they might contribute to non-essential RNAPI functions. Only one residue, Rpa190-S685, was suggested to play a role in rRNA cleavage/elongation or termination (Gerber et al., 2008). To date, 115 site-specific phosphorylations have been identified, mostly in phospho-proteomic studies, distributed along all the 14 RNAPI subunits (Ficarro et al., 2002; Albuquerque et al., 2008; Holt et al., 2009; Soulard et al., 2010; Pultz et al., 2012; Swaney et al., 2013; Sostaric et al., 2018; MacGilvray et al., 2020; Lanz et al., 2021). We have compiled all these sites in **Supplementary Table S1**. In summary, 81 sites reside in specific subunits and 34 are shared: 20 with RNAPII and RNAPIII and 14 with RNAPIII. Among these 81 phospho-sites, 63

are localized in regions of solved structure (**Figure 2A**, upper panel). Remarkably, 49 sites are exposed on the surface of the enzyme (**Figure 2A**, bottom panel, and **Figure 2C**), which again suggests a role for the association of RNAPI with transcription regulators, for instance Rrn3 (**Figure 2B**) (Torreira et al., 2017).

One of the first observations implicating protein phosphorylation in regulating RNAPI activity was the discovery that Fcp1, a Rpb1-CTD phosphatase, interacted with the RNAPI transcription machinery and was essential for rDNA efficient transcription, probably by facilitating RNAPI dephosphorylation and chain elongation during rRNA synthesis (Fath et al., 2004). Later, it was documented that the Rpa43 subunit was phosphorylated in several specific residues (S208, S220, S262, S263, S285) (Gerber et al., 2008). This subunit, together with Rpa14, forms the stalk domain and creates a platform for binding initiation factors and newly synthesized RNA (Fernandez-Tornero et al., 2013; Torreira et al., 2017). This stalk domain is required for RNAPI homodimerization and transcription inactivation (Torreira et al., 2017). However, it is unknown if Rpa43 phosphorylation levels play a role in this process. Nonetheless, it was reported that Cdc14 dephosphorylates Rpa43 in mitosis to exclude it from the nucleolus, thereby restraining rDNA transcription and facilitating condensin loading, an essential step for correct segregation of the nucleolus (Clemente-Blanco et al., 2009).

Rpa43 interacts with Rrn3 (Milkereit and Tschochner, 1998; Moorefield et al., 2000), a crucial RNAPI factor whose phosphorylation has been implicated in the regulation of the holoenzyme, after activation of growth factor signaling pathways that connect nutrient availability and rDNA production. Rrn3 is the yeast homologue of the mammalian growth-dependent rRNA synthesis factor TIF-IA (Grummt and Voit, 2010). This interaction depends on the phosphorylation of RNAPI and on Rrn3-P/TIF-IA association, and is essential to establish a competent transcriptional initiation complex (Fath et al., 2001; Cavanaugh et al., 2002; Torreira et al., 2017). Interestingly, in mice, casein kinase 2 (C1K2) has been implicated in Rrn3/TIF-IA phosphorylation at S170/172 to trigger its release from the RNAPI complex after transcription initiation, a prerequisite for transcription elongation (Bierhoff et al., 2008). This suggests that Rrn3/TIF-IA is subjected to a complex phospho-code that regulates its interaction with the RNAPI holoenzyme during ribosome biogenesis. Importantly, human RNAPI activity is also controlled in response to different types of environmental stresses throughout the phosphorylation of Rrn3/TIF-IA. Under glucose restriction, Rrn3/TIF-IA phosphorylation by the AMPK kinase prevents the assembly of a functional PIC (Hoppe et al., 2009). On the other hand, Rrn3/TIF-IA phosphorylation by the JNK kinase in mice restrains its interaction with RNAPI in response to oxidative stress, thus abrogating the formation of new PICs (Mayer et al., 2005).

RNAPIII PHOSPHORYLATION

In terms of structural composition, RNAPIII is the largest eukaryotic RNA polymerase complex in mass and molecular conformation (Vannini and Cramer, 2012). It is formed by 17 subunits, 10 of which are unique to RNAPIII. Novel phospho-proteomic studies have shed light on the post-transcriptional phospho-mapping of multiple RNAPIII subunits (Albuquerque et al., 2008; Holt et al., 2009; Swaney et al., 2013; MacGilvray et al., 2020; Lanz et al., 2021). Fifteen of the 17 subunits are phosphorylated in both yeast and humans (**Supplementary Table S1, S2**). In the case of *S. cerevisiae*, there are 76 phospho-sites, 42 of them localized in specific subunits. Only 28 residues are localized in regions of known structure (**Figure 2D**) and 19 are exposed on the surface of the polymerase complex (**Figure 2D**, bottom, and **Figure 2F**). Three specific residues of Rpc53 (S224, T228 and T232) are of known function (see below) (Lee et al., 2012). Therefore, it is intuitive to think that RNAPIII activity might also be highly regulated by phosphorylation events, as described for RNAPI and RNAPII complexes. Interestingly, in *S. cerevisiae* only two phospho-sites in the two largest subunits have been shown to be phosphorylated, whereas 32 phospho-sites distributed along these subunits have been identified in humans (**Supplementary Table S2**). This observation suggests that regulation of RNAPIII activity by phosphorylation could be species specific.

Probably the best-known regulator of RNAPIII is the repressor Maf1 (Willis and Moir, 2018; Vorlander et al., 2020a), whose activity is controlled by its phosphorylation at multiple sites by protein kinase A, the rapamycin-sensitive TOR kinase (TORC1) and the TORC1-regulated kinase Sch9. Phosphorylation of Maf1 by these kinases leads to changes in its subcellular localization, a mechanism that ensures the accurate activation/repression of RNAPIII (Moir et al., 2006; Lee et al., 2009; Wei and Zheng, 2009; Willis and Moir, 2018). Additionally, casein kinase 2 (CK2) phosphorylation of Maf1 in favorable growth conditions releases this protein from the RNAPIII complex bound to genes for tRNAs, thus activating their transcription (Graczyk et al., 2011). Maf1 regulation also depends on protein phosphatases. It has been postulated that in response to nutrient starvation, poor carbon sources or several cellular stresses, Maf1 is dephosphorylated in a PP4/PP2A-dependent manner and translocated to the nucleus, thus repressing RNAPIII activity (Oler and Cairns, 2012; Ahn et al., 2019). Interestingly, nuclear localization of Maf1 is not enough to completely inhibit RNAPIII activity, suggesting the existence of alternative mechanisms that co-regulate RNAPIII transcription under these conditions (Huber et al., 2009). In agreement with this observation, recent studies have demonstrated that the RNAPIII subunit Rpc53 is also subjected to a phosphorylation switch in response to nutrient limitation and other types of cellular stress. Rpc53 phosphorylation by the two conserved kinases Kns1 and Mck1 modifies the ability of RNAPIII to interact with the DNA molecule, thus avoiding recycling rounds of transcription and allowing dephosphorylated Maf1 to join and inhibit RNAPIII activity (Lee et al., 2012). Another component of RNAPIII controlled by phosphorylation is its Rpc82 subunit, whose concomitant phosphorylation with the

TFIIIB subunit Bdp1 by the Sch9 and CK2 kinases opposes Maf1-mediated transcriptional repression (Lee et al., 2015). Finally, it has been reported that the TATA-binding protein (TBP) is also a preferred substrate of CK2 *in vitro*, which suggests a new mechanism to regulate RNAPIII transcription by phosphorylation *in vivo* (Ghavidel and Schultz, 1997).

It is important to remark that RNAPIII transcription is regulated in response to environmental cues and during the different stages encompassed in the cell cycle. It has been reported that tRNA levels fluctuate during the cell cycle in a process controlled by the Cdk1/Clb5 kinase complex, boosting tRNA expression during the S phase. This is attained by the cycling phosphorylation of Bdp1, an event that triggers the recruitment of TFIIIC to the genes for tRNAs, stimulates interaction between TFIIIB and TFIIIC and enhances RNAPIII activity (Herrera et al., 2018). However, the physiological significance of cell cycle regulation of RNAPIII transcription remains to be elucidated and undoubtedly will be a fascinating question for the future.

COORDINATION OF RNAP ACTIVITIES: ADAPTING GENE EXPRESSION TO ENVIRONMENTAL CONDITIONS

RNAP activities are essential for cellular viability and a limiting step in regulating gene expression. All RNAPs respond to growth cell conditions and nutrient availability. In actively growing cells, the majority of the transcriptional output is due to RNAPI and RNAPIII activities, which are required for the synthesis of ribosomes. The activity of RNAPII is also essential because it transcribes all ribosomal protein genes and genes encoding factors required for ribosome assembly (Ribi regulon) (Warner, 1999; de la Cruz et al., 2018). How eukaryotic RNAPs are regulated has been extensively studied and is still a field of great interest. However, less is known about the mechanisms coordinating and communicating with the three RNAP machineries to adapt cell growth to environmental conditions. Factors that coordinate at least the function of two RNAPs have been identified. For instance, Spt4/5, Paf1C and Ccr4 regulate both RNAPI and RNAPII transcription (Zhang et al., 2010; Anderson et al., 2011; Hartzog and Fu, 2013; Laribee et al., 2015). Similarly, TFIIS and Sub1 influence RNAPII and RNAPIII (Guglielmi et al., 2007; Ghavi-Helm et al., 2008; Carriere et al., 2012; Garcia et al., 2012; Garavis et al., 2017; Calvo, 2018). Recently, the yeast prefoldin-like Bud27 has been shown to be a regulator of the three RNAPs, most likely via its association with a common subunit, Rpb5 (Martinez-Fernandez et al., 2020). One possibility is that RNAPs might be coordinated through regulation of the phosphorylation state of their shared subunits. In support of this hypothesis, Rpb5, Rpb6, Rpb8, Rpb10, and Rpb12 contain phospho-sites (Albuquerque et al., 2008; Swaney et al., 2013; Sostaric et al., 2018; MacGilvray et al., 2020) (**Supplementary Table S1**).

TOR serine/threonine kinases play an essential role in controlling many aspects of living cells, such as growth, proliferation and survival in response to nutrients (Loewith and Hall, 2011; Kim and Guan, 2019; Laribee and Weisman, 2020). Initially, it was reported that TOR proteins only localized in the cytoplasm, with a crucial role in regulating protein synthesis (Barbet et al., 1996; Gingras et al., 2004). We currently know that TOR and its associated proteins also localize in the nucleus, where they regulate gene expression to guarantee the appropriate ribogenesis (Tsang and Zheng, 2007; Laribee, 2018; Laribee and Weisman, 2020). When *S. cerevisiae* is grown under nutrient-replete conditions, Tor1 localizes in both the cytoplasm and the nucleus. In the nucleus, Tor1 and Kog1 (the Raptor subunit in *S. cerevisiae*) bind to the 35S (RNAPI) and 5S (RNAPIII) promoters. However, after starvation or rapamycin treatment, they are removed from these regions, thus inhibiting transcription (Li et al., 2006). In mammals, mTOR and Raptor also interact with the RNAPIII factor TFIIIC to induce 5S and tRNA transcription (Kantidakis et al., 2010). This and other evidence suggest that TORC1 complexes are RNAPI and RNAPIII regulators and likely coordinators of these two RNAP activities. Whether any of the RNAPI or RNAPIII subunits are phosphorylated by TORC1 is unknown. Nevertheless, RNAPIII transcription is also activated by TORC1 via phosphorylation of Maf1 in yeasts (Huber et al., 2009; Lee et al., 2009) and mammals (Kantidakis et al., 2010; Michels et al., 2010). Whereas mTOR and Raptor contribute to RNAPII transcription regulation of a number of genes in mammals (Cunningham et al., 2007; Chaveroux et al., 2013; Laribee, 2018), in budding yeast only the *HMO1* gene is known to be directly activated by Tor1 (Panday et al., 2017). Hmo1 activates the transcription of genes regulated by TORC1, including RP, 5S and 35S genes (Gadal et al., 2002; Hall et al., 2006). Both Tor1 and Hmo1 bind to the *HMO1* promoter, facilitating its transcription. After rapamycin treatment or DNA damage, Tor1 and Hmo1 are released, thus inhibiting transcription (Panday et al., 2017). Interestingly, promoter binding by the Tor1 kinase is a prerequisite for transcription inhibition, which suggests that Tor1 may phosphorylate a specific target to repress transcription in response to stress conditions. One of these targets might be Paf1C, whose activity is needed to attenuate RNAPI transcription after TORC1 inhibition (Zhang et al., 2010). Similarly, Ccr4 couples nutrient signaling through TORC1 with Rrn3-RNAPI transcription inhibition (Laribee et al., 2015). It would be reasonable to think that Tor1 kinases could also phosphorylate RNAPs, maybe common subunits, to coordinate and modulate their activities in response to environmental conditions.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

Funding was provided by the Spanish Ministry of Economy and Competitiveness (MINECO) and ERDF with Grants conceded

to OC (BFU2017-84694-P), AC-B (PGC2018-097963-B-100), FN (BFU2016-77728-C3-2-P), and by Spanish Ministry of Science and Innovation (MICINN) and ERDF to OC and FN (RED2018-102467-T); and by the Junta de Andalucía-Universidad de Jaén (FEDER-UJA 1260360), and the Junta de Andalucía (BIO258) to FN, AG-J, and AC were supported by pre-doctoral contracts from "Junta de Castilla y León". The IBFG is supported in part by an institutional grant from the "Junta y Castilla y León" (Programa "Escalera de Excelencia" de la Junta de Castilla y León, CLU-2017-03), cofunded by ERDF OP de Castilla y León 14-20.

ACKNOWLEDGMENTS

We thank members of our laboratories for helpful discussions on the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.681865/full#supplementary-material

REFERENCES

Ahn, C. S., Lee, D.-H., and Pai, H.-S. (2019). Characterization of Maf1 in Arabidopsis: Function under Stress Conditions and Regulation by the TOR Signaling Pathway. *Planta* 249 (2), 527–542. doi:10.1007/s00425-018-3024-5

Albuquerque, C. P., Smolka, M. B., Payne, S. H., Bafna, V., Eng, J., and Zhou, H. (2008). A Multidimensional Chromatography Technology for In-Depth Phosphoproteome Analysis. *Mol. Cell Proteomics* 7 (7), 1389–1396. doi:10.1074/mcp.M700468-MCP200

Allepuz-Fuster, P., Martínez-Fernández, V., Garrido-Godino, A. I., Alonso-Aguado, S., Hanes, S. D., Navarro, F., et al. (2014). Rpb4/7 Facilitates RNA Polymerase II CTD Dephosphorylation. *Nucleic Acids Res.* 42 (22), 13674–13688. doi:10.1093/nar/gku1227

Allepuz-Fuster, P., O'Brien, M. J., González-Polo, N., Pereira, B., Dhoondia, Z., Ansari, A., et al. (2019). RNA Polymerase II Plays an Active Role in the Formation of Gene Loops through the Rpb4 Subunit. *Nucleic Acids Res.* 47 (17), 8975–8987. doi:10.1093/nar/gkz597

Anderson, S. J., Sikes, M. L., Zhang, Y., French, S. L., Salgia, S., Beyer, A. L., et al. (2011). The Transcription Elongation Factor Spt5 Influences Transcription by RNA Polymerase I Positively and Negatively. *J. Biol. Chem.* 286 (21), 18816–18824. doi:10.1074/jbc.M110.202101

Armache, K.-J., Kettenberger, H., and Cramer, P. (2003). Architecture of Initiation-Competent 12-subunit RNA Polymerase II. *Proc. Natl. Acad. Sci.* 100 (12), 6964–6968. doi:10.1073/pnas.1030608100

Armache, K.-J., Mitterweger, S., Meinhart, A., and Cramer, P. (2005). Structures of Complete RNA Polymerase II and its Subcomplex, Rpb4/7. *J. Biol. Chem.* 280 (8), 7131–7134. doi:10.1074/jbc.M413038200

Babbarwal, V., Fu, J., and Reese, J. C. (2014). The Rpb4/7 Module of RNA Polymerase II Is Required for Carbon Catabolite Repressor Protein 4-negative on TATA (Ccr4-Not) Complex to Promote Elongation. *J. Biol. Chem.* 289 (48), 33125–33130. doi:10.1074/jbc.C114.601088

Barbet, N. C., Schneider, U., Helliwell, S. B., Stansfield, I., Tuite, M. F., and Hall, M. N. (1996). TOR Controls Translation Initiation and Early G1 Progression in Yeast. *MBoC* 7 (1), 25–42. doi:10.1091/mbc.7.1.25

Beltrao, P., Bork, P., Krogan, N. J., and Noort, V. (2013). Evolution and Functional Cross-talk of Protein post-translational Modifications. *Mol. Syst. Biol.* 9, 714. doi:10.1002/msb.201304521

Bierhoff, H., Dundr, M., Michels, A. A., and Grummt, I. (2008). Phosphorylation by Casein Kinase 2 Facilitates rRNA Gene Transcription by Promoting Dissociation of TIF-IA from Elongating RNA Polymerase I. *Mcb* 28 (16), 4988–4998. doi:10.1128/MCB.00492-08

Boehning, M., Dugast-Darzacq, C., Rankovic, M., Hansen, A. S., Yu, T., Marie-Nelly, H., et al. (2018). RNA Polymerase II Clustering through Carboxy-Terminal Domain Phase Separation. *Nat. Struct. Mol. Biol.* 25 (9), 833–840. doi:10.1038/s41594-018-0112-y

Buratowski, S. (2009). Progression through the RNA Polymerase II CTD Cycle. *Mol. Cel* 36 (4), 541–546. doi:10.1016/j.molcel.2009.10.019

Buratowski, S. (2003). The CTD Code. *Nat. Struct. Mol. Biol.* 10 (9), 679–680. doi:10.1038/nsb0903-679

Bushnell, D. A., and Kornberg, R. D. (2003). Complete, 12-subunit RNA Polymerase II at 4.1-A Resolution: Implications for the Initiation of Transcription. *Proc. Natl. Acad. Sci.* 100 (12), 6969–6973. doi:10.1073/pnas.1130601100

Cai, G., Imasaki, T., Yamada, K., Cardelli, F., Takagi, Y., and Asturias, F. J. (2010). Mediator Head Module Structure and Functional Interactions. *Nat. Struct. Mol. Biol.* 17 (3), 273–279. doi:10.1038/nsmb.1757

Calvo, O., and García, A. (2012). "RNA Polymerase II Phosphorylation and Gene Expression Regulation," in *Protein Phosphorylation in Human Cells* (Rijeka, Croatia: Intech Open), 151–194.

Calvo, O. (2020). RNA Polymerase II Phosphorylation and Gene Looping: New Roles for the Rpb4/7 Heterodimer in Regulating Gene Expression. *Curr. Genet.* 66 (5), 927–937. doi:10.1007/s00294-020-01084-w

Calvo, O. (2018). Sub1 and RNAPII, until Termination Does Them Part. *Transcription* 9 (1), 52–60. doi:10.1080/21541264.2017.1333558

Carrière, L., Graziani, S., Alibert, O., Ghavi-Helm, Y., Boussouar, F., Humbertclaude, H., et al. (2012). Genomic Binding of Pol III Transcription Machinery and Relationship with TFIIS Transcription Factor Distribution in Mouse Embryonic Stem Cells. *Nucleic Acids Res.* 40 (1), 270–283. doi:10.1093/nar/gkr737

Cavanaugh, A. H., Hirschler-Laszkiewicz, I., Hu, Q., Dundr, M., Smink, T., Misteli, T., et al. (2002). Rrn3 Phosphorylation Is a Regulatory Checkpoint for Ribosome Biogenesis. *J. Biol. Chem.* 277 (30), 27423–27432. doi:10.1074/jbc.M201232200

Chapman, R. D., Heidemann, M., Hintermair, C., and Eick, D. (2008). Molecular Evolution of the RNA Polymerase II CTD. *Trends Genet.* 24 (6), 289–296. doi:10.1016/j.tig.2008.03.010

Chaveroux, C., Eichner, L. J., Dufour, C. R., Shatnawi, A., Khoutorsky, A., Bourque, G., et al. (2013). Molecular and Genetic Crosstalks between mTOR and ERRα Are Key Determinants of Rapamycin-Induced Nonalcoholic Fatty Liver. *Cel Metab.* 17 (4), 586–598. doi:10.1016/j.cmet.2013.03.003

Clemente-Blanco, A., Mayán-Santos, M., Schneider, D. A., Machín, F., Jarmuz, A., Tschochner, H., et al. (2009). Cdc14 Inhibits Transcription by RNA Polymerase I during Anaphase. *Nature* 458 (7235), 219–222. doi:10.1038/nature07652

Corden, J. L., Cadena, D. L., Ahearn, J. M., Jr., and Dahmus, M. E. (1985). A Unique Structure at the Carboxyl Terminus of the Largest Subunit of Eukaryotic RNA Polymerase II. *Proc. Natl. Acad. Sci.* 82 (23), 7934–7938. doi:10.1073/pnas.82.23.7934

Cramer, P., Armache, K.-J., Baumli, S., Benkert, S., Brueckner, F., Buchen, C., et al. (2008). Structure of Eukaryotic RNA Polymerases. *Annu. Rev. Biophys.* 37, 337–352. doi:10.1146/annurev.biophys.37.032807.130008

Cramer, P., Bushnell, D. A., Fu, J., Gnatt, A. L., Maier-Davis, B., Thompson, N. E., et al. (2000). Architecture of RNA Polymerase II and Implications for the Transcription Mechanism. *Science* 288 (5466), 640–649. doi:10.1126/science.288.5466.640

Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001). Structural Basis of Transcription: RNA Polymerase II at 2.8 Angstrom Resolution. *Science* 292 (5523), 1863–1876. doi:10.1126/science.1059493

Cramer, P. (2019a). Eukaryotic Transcription Turns 50. *Cell* 179 (4), 808–812. doi:10.1016/j.cell.2019.09.018

Cramer, P. (2019b). Organization and Regulation of Gene Transcription. *Nature* 573 (7772), 45–54. doi:10.1038/s41586-019-1517-4

Cuevas-Bermúdez, A., Martinez-Fernandez, V., Garrido-Godino, A. I., and Navarro, F. (2017). "Subunits Common to RNA Polymerases," in *The Yeast Role in Medical Applications* (Rijeka, Croatia: Intech Open), 151–165.)

Cunningham, J. T., Rodgers, J. T., Arlow, D. H., Vazquez, F., Mootha, V. K., and Puigserver, P. (2007). mTOR Controls Mitochondrial Oxidative Function through a YY1-PGC-1α Transcriptional Complex. *Nature* 450 (7170), 736–740. doi:10.1038/nature06322

de la Cruz, J., Gómez-Herreros, F., Rodríguez-Galán, O., Begley, V., de la Cruz Muñoz-Centeno, M., and Chávez, S. (2018). Feedback Regulation of Ribosome Assembly. *Curr. Genet.* 64 (2), 393–404. doi:10.1007/s00294-017-0764-x

Deribe, Y. L., Pawson, T., and Dikic, I. (2010). Post-translational Modifications in Signal Integration. *Nat. Struct. Mol. Biol.* 17 (6), 666–672. doi:10.1038/nsmb.1842

Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M., and Pagano, A. (2007). The Expanding RNA Polymerase III Transcriptome. *Trends Genet.* 23 (12), 614–622. doi:10.1016/j.tig.2007.09.001

Duek, L., Barkai, O., Elran, R., Adawi, I., and Choder, M. (2018). Dissociation of Rpb4 from RNA Polymerase II Is Important for Yeast Functionality. *PLoS One* 13 (10), e0206161. doi:10.1371/journal.pone.0206161

Eick, D., and Geyer, M. (2013). The RNA Polymerase II Carboxy-Terminal Domain (CTD) Code. *Chem. Rev.* 113 (11), 8456–8490. doi:10.1021/cr400071f

Engel, C., Sainsbury, S., Cheung, A. C., Kostrewa, D., and Cramer, P. (2013). RNA Polymerase I Structure and Transcription Regulation. *Nature* 502 (7473), 650–655. doi:10.1038/nature12712

Fath, S., Kobor, M. S., Philippi, A., Greenblatt, J., and Tschochner, H. (2004). Dephosphorylation of RNA Polymerase I by Fcp1p Is Required for Efficient rRNA Synthesis. *J. Biol. Chem.* 279 (24), 25251–25259. doi:10.1074/jbc.M401867200

Fath, S., Milkereit, P., Peyroche, G., Riva, M., Carles, C., and Tschochner, H. (2001). Differential Roles of Phosphorylation in the Formation of Transcriptional Active RNA Polymerase I. *Proc. Natl. Acad. Sci.* 98 (25), 14334–14339. doi:10.1073/pnas.231181398

Fernández-Tornero, C., Moreno-Morcillo, M., Rashid, U. J., Taylor, N. M. I., Ruiz, F. M., Gruene, T., et al. (2013). Crystal Structure of the 14-subunit RNA Polymerase I. *Nature* 502 (7473), 644–649. doi:10.1038/nature12636

Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., et al. (2002). Phosphoproteome Analysis by Mass Spectrometry and its Application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* 20 (3), 301–305. doi:10.1038/nbt0302-301

Gadal, O., Labarre, S., Boschiero, C., and Thuriaux, P. (2002). Hmo1, an HMG-Box Protein, Belongs to the Yeast Ribosomal DNA Transcription System. *EMBO J.* 21 (20), 5498–5507. doi:10.1093/emboj/cdf539

Garavís, M., González-Polo, N., Allepuz-Fuster, P., Louro, J. A., Fernández-Tornero, C., and Calvo, O. (2017). Sub1 Contacts the RNA Polymerase II Stalk to Modulate mRNA Synthesis. *Nucleic Acids Res.* 45 (5), 2458–2471. doi:10.1093/nar/gkw1206

García, A., Collin, A., and Calvo, O. (2012). Sub1 Associates with Spt5 and Influences RNA Polymerase II Transcription Elongation Rate. *MBoC* 23 (21), 4297–4312. doi:10.1091/mbc.E12-04-0331

Gerber, J., Reiter, A., Steinbauer, R., Jakob, S., Kuhn, C.-D., Cramer, P., et al. (2008). Site Specific Phosphorylation of Yeast RNA Polymerase I. *Nucleic Acids Res.* 36 (3), 793–802. doi:10.1093/nar/gkm1093

Ghavi-Helm, Y., Michaut, M., Acker, J., Aude, J.-C., Thuriaux, P., Werner, M., et al. (2008). Genome-wide Location Analysis Reveals a Role of TFIIS in RNA Polymerase III Transcription. *Genes Develop.* 22 (14), 1934–1947. doi:10.1101/gad.471908

Ghavidel, A., and Schultz, M. C. (1997). Casein Kinase II Regulation of Yeast TFIIIB Is Mediated by the TATA-Binding Protein. *Genes Develop.* 11 (21), 2780–2789. doi:10.1101/gad.11.21.2780

Gingras, A.-C., Raught, B., and Sonenberg, N. (2004). mTOR Signaling to Translation. *Curr. Top. Microbiol. Immunol.* 279, 169–197. doi:10.1007/978-3-642-18930-2_11

Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A., and Kornberg, R. D. (2001). Structural Basis of Transcription: an RNA Polymerase II Elongation Complex at 3.3 A Resolution. *Science* 292 (5523), 1876–1882. doi:10.1126/science.1059495

Graczyk, D., Debski, J., Muszynska, G., Bretner, M., Lefebvre, O., and Boguta, M. (2011). Casein Kinase II-Mediated Phosphorylation of General Repressor Maf1 Triggers RNA Polymerase III Activation. *Proc. Natl. Acad. Sci.* 108 (12), 4926–4931. doi:10.1073/pnas.1010010108

Greber, B. J., and Nogales, E. (2019). The Structures of Eukaryotic Transcription Pre-initiation Complexes and Their Functional Implications. *Subcell Biochem.* 93, 143–192. doi:10.1007/978-3-030-28151-9_5

Grummt, I., and Voit, R. (2010). Linking rDNA Transcription to the Cellular Energy Supply. *Cell Cycle* 9 (2), 225–226. doi:10.4161/cc.9.2.10614

Guglielmi, B., Soutourina, J., Esnault, C., and Werner, M. (2007). TFIIS Elongation Factor and Mediator Act in Conjunction during Transcription Initiation *In Vivo*. *Proc. Natl. Acad. Sci.* 104 (41), 16062–16067. doi:10.1073/pnas.0704534104

Guo, Y. E., Manteiga, J. C., Henninger, J. E., Sabari, B. R., Dall'Agnese, A., Hannett, N. M., et al. (2019). Pol II Phosphorylation Regulates a Switch between Transcriptional and Splicing Condensates. *Nature* 572 (7770), 543–548. doi:10.1038/s41586-019-1464-0

Haag, J. R., and Pikaard, C. S. (2011). Multisubunit RNA Polymerases IV and V: Purveyors of Non-coding RNA for Plant Gene Silencing. *Nat. Rev. Mol. Cel Biol* 12 (8), 483–492. doi:10.1038/nrm3152

Hahn, S., and Young, E. T. (2011). Transcriptional Regulation in *Saccharomyces cerevisiae*: Transcription Factor Regulation and Function, Mechanisms of Initiation, and Roles of Activators and Coactivators. *Genetics* 189 (3), 705–736. doi:10.1534/genetics.111.127019

Hall, D. B., Wade, J. T., and Struhl, K. (2006). An HMG Protein, Hmo1, Associates with Promoters of many Ribosomal Protein Genes and throughout the rRNA Gene Locus in *Saccharomyces cerevisiae*. *Mol. Cel Biol* 26 (9), 3672–3679. doi:10.1128/MCB.26.9.3672-3679.2006

Harlen, K. M., and Churchman, L. S. (2017). The Code and beyond: Transcription Regulation by the RNA Polymerase II Carboxy-Terminal Domain. *Nat. Rev. Mol. Cel Biol* 18 (4), 263–273. doi:10.1038/nrm.2017.10

Hartzog, G. A., and Fu, J. (2013). The Spt4-Spt5 Complex: a Multi-Faceted Regulator of Transcription Elongation. *Biochim. Biophys. Acta (Bba) - Gene Regul. Mech.* 1829 (1), 105–115. doi:10.1016/j.bbagrm.2012.08.007

He, Y., Fang, J., Taatjes, D. J., and Nogales, E. (2013). Structural Visualization of Key Steps in Human Transcription Initiation. *Nature* 495 (7442), 481–486. doi:10.1038/nature11991

He, Y., Yan, C., Fang, J., Inouye, C., Tjian, R., Ivanov, I., et al. (2016). Near-atomic Resolution Visualization of Human Transcription Promoter Opening. *Nature* 533 (7603), 359–365. doi:10.1038/nature17970

Herrera, M. C., Chymkowitch, P., Robertson, J. M., Eriksson, J., Bøe, S. O., Alseth, I., et al. (2018). Cdk1 gates Cell Cycle-dependent tRNA Synthesis by Regulating RNA Polymerase III Activity. *Nucleic Acids Res.* 46 (22), 11698–11711. doi:10.1093/nar/gky846

Ho, C. K., and Shuman, S. (1999). Distinct Roles for CTD Ser-2 and Ser-5 Phosphorylation in the Recruitment and Allosteric Activation of Mammalian mRNA Capping Enzyme. *Mol. Cel* 3 (3), 405–411. doi:10.1016/s1097-2765(00)80468-2

Hoffmann, N. A., Jakobi, A. J., Moreno-Morcillo, M., Glatt, S., Kosinski, J., Hagen, W. J. H., et al. (2015). Molecular Structures of Unbound and Transcribing RNA Polymerase III. *Nature* 528 (7581), 231–236. doi:10.1038/nature16143

Holt, L. J., Tuch, B. B., Villen, J., Johnson, A. D., Gygi, S. P., and Morgan, D. O. (2009). Global Analysis of Cdk1 Substrate Phosphorylation Sites Provides Insights into Evolution. *Science* 325 (5948), 1682–1686. doi:10.1126/science.1172867

Hoppe, S., Bierhoff, H., Cado, I., Weber, A., Tiebe, M., Grummt, I., et al. (2009). AMP-activated Protein Kinase Adapts rRNA Synthesis to Cellular Energy Supply. *Proc. Natl. Acad. Sci.* 106 (42), 17781–17786. doi:10.1073/pnas.0909873106

Hsin, J.-P., and Manley, J. L. (2012). The RNA Polymerase II CTD Coordinates Transcription and RNA Processing. *Genes Develop.* 26 (19), 2119–2137. doi:10.1101/gad.200303.112

Hsin, J.-P., Sheth, A., and Manley, J. L. (2011). RNAP II CTD Phosphorylated on Threonine-4 Is Required for Histone mRNA 3' End Processing. *Science* 334 (6056), 683–686. doi:10.1126/science.1206034

Huber, A., Bodenmiller, B., Uotila, A., Stahl, M., Wanka, S., Gerrits, B., et al. (2009). Characterization of the Rapamycin-Sensitive Phosphoproteome Reveals that Sch9 Is a central Coordinator of Protein Synthesis. *Genes Develop.* 23 (16), 1929–1943. doi:10.1101/gad.532109

Huet, J., Riva, M., Sentenac, A., and Fromageot, P. (1985). Yeast RNA Polymerase C and its Subunits. Specific Antibodies as Structural and Functional Probes. *J. Biol. Chem.* 260 (28), 15304–15310. doi:10.1016/s0021-9258(18)95736-4

Kantidakis, T., Ramsbottom, B. A., Birch, J. L., Dowding, S. N., and White, R. J. (2010). mTOR Associates with TFIIIC, Is Found at tRNA and 5S rRNA Genes, and Targets Their Repressor Maf1. *Proc. Natl. Acad. Sci.* 107 (26), 11823–11828. doi:10.1073/pnas.1005188107

Kim, J., and Guan, K.-L. (2019). mTOR as a central Hub of Nutrient Signalling and Cell Growth. *Nat. Cel Biol* 21 (1), 63–71. doi:10.1038/s41556-018-0205-1

Kim, Y.-J., Björklund, S., Li, Y., Sayre, M. H., and Kornberg, R. D. (1994). A Multiprotein Mediator of Transcriptional Activation and its Interaction with the C-Terminal Repeat Domain of RNA Polymerase II. *Cell* 77 (4), 599–608. doi:10.1016/0092-8674(94)90221-6

Kimura, M., Suzuki, H., and Ishihama, A. (2002). Formation of a Carboxy-Terminal Domain Phosphatase (Fcp1)/TFIIF/RNA Polymerase II (Pol II) Complex in *Schizosaccharomyces pombe* Involves Direct Interaction between Fcp1 and the Rpb4 Subunit of Pol II. *Mol. Cel Biol* 22 (5), 1577–1588. doi:10.1128/mcb.22.5.1577-1588.2002

Kostrewa, D., Zeller, M. E., Armache, K.-J., Seizl, M., Leike, K., Thomm, M., et al. (2009). RNA Polymerase II-TFIIB Structure and Mechanism of Transcription Initiation. *Nature* 462 (7271), 323–330. doi:10.1038/nature08548

Kwak, H., and Lis, J. T. (2013). Control of Transcriptional Elongation. *Annu. Rev. Genet.* 47, 483–508. doi:10.1146/annurev-genet-110711-155440

Lanz, M. C., Yugandhar, K., Gupta, S., Sanford, E. J., Faça, V. M., Vega, S., et al. (2021). In-depth and 3-dimensional Exploration of the Budding Yeast Phosphoproteome. *EMBO Rep.* 22 (2), e51121. doi:10.15252/embr.202051121

Laribee, R. N., Hosni-Ahmed, A., Workman, J. J., and Chen, H. (2015). Ccr4-Not Regulates RNA Polymerase I Transcription and Couples Nutrient Signaling to the Control of Ribosomal RNA Biogenesis. *Plos Genet.* 11 (3), e1005113. doi:10.1371/journal.pgen.1005113

Laribee, R. N. (2018). Transcriptional and Epigenetic Regulation by the Mechanistic Target of Rapamycin Complex 1 Pathway. *J. Mol. Biol.* 430 (24), 4874–4890. doi:10.1016/j.jmb.2018.10.008

Laribee, R. N., and Weisman, R. (2020). Nuclear Functions of TOR: Impact on Transcription and the Epigenome. *Genes* 11 (6), 641. doi:10.3390/genes11060641

Larochelle, S., Amat, R., Glover-Cutter, K., Sansó, M., Zhang, C., Allen, J. J., et al. (2012). Cyclin-dependent Kinase Control of the Initiation-To-Elongation Switch of RNA Polymerase II. *Nat. Struct. Mol. Biol.* 19 (11), 1108–1115. doi:10.1038/nsmb.2399

Lee, J., Moir, R. D., McIntosh, K. B., and Willis, I. M. (2012). TOR Signaling Regulates Ribosome and tRNA Synthesis via LAMMER/Clk and GSK-3 Family Kinases. *Mol. Cel* 45 (6), 836–843. doi:10.1016/j.molcel.2012.01.018

Lee, J., Moir, R. D., and Willis, I. M. (2015). Differential Phosphorylation of RNA Polymerase III and the Initiation Factor TFIIIB in *Saccharomyces cerevisiae*. *PLoS One* 10 (5), e0127225. doi:10.1371/journal.pone.0127225

Lee, J., Moir, R. D., and Willis, I. M. (2009). Regulation of RNA Polymerase III Transcription Involves SCH9-dependent and SCH9-independent Branches of the Target of Rapamycin (TOR) Pathway. *J. Biol. Chem.* 284 (19), 12604–12608. doi:10.1074/jbc.C900020200

Li, H., Tsang, C. K., Watkins, M., Bertram, P. G., and Zheng, X. F. S. (2006). Nutrient Regulates Tor1 Nuclear Localization and Association with rDNA Promoter. *Nature* 442 (7106), 1058–1061. doi:10.1038/nature05020

Li, W., Giles, C., and Li, S. (2014). Insights into How Spt5 Functions in Transcription Elongation and Repressing Transcription Coupled DNA Repair. *Nucleic Acids Res.* 42 (11), 7069–7083. doi:10.1093/nar/gku333

Liu, X., Bushnell, D. A., Wang, D., Calero, G., and Kornberg, R. D. (2010). Structure of an RNA Polymerase II-TFIIB Complex and the Transcription Initiation Mechanism. *Science* 327 (5962), 206–209. doi:10.1126/science.1182015

Loewith, R., and Hall, M. N. (2011). Target of Rapamycin (TOR) in Nutrient Signaling and Growth Control. *Genetics* 189 (4), 1177–1201. doi:10.1534/genetics.111.133363

López, A., Ramírez, V., García-Andrade, J., Flors, V., and Vera, P. (2011). The RNA Silencing Enzyme RNA Polymerase V Is Required for Plant Immunity. *Plos Genet.* 7 (12), e1002434. doi:10.1371/journal.pgen.1002434

MacGilvray, M. E., Shishkova, E., Place, M., Wagner, E. R., Coon, J. J., and Gasch, A. P. (2020). Phosphoproteome Response to Dithiothreitol Reveals Unique versus Shared Features of *Saccharomyces cerevisiae* Stress Responses. *J. Proteome Res.* 19 (8), 3405–3417. doi:10.1021/acs.jproteome.0c00253

Martínez-Fernández, V., Cuevas-Bermúdez, A., Gutiérrez-Santiago, F., Garrido-Godino, A. I., Rodríguez-Galán, O., Jordán-Pla, A., et al. (2020). Prefoldin-like Bud27 Influences the Transcription of Ribosomal Components and Ribosome Biogenesis in *Saccharomyces cerevisiae*. *RNA* 26 (10), 1360–1379. doi:10.1261/rna.075507.120

Martinez-Rucobo, F. W., Sainsbury, S., Cheung, A. C., and Cramer, P. (2011). Architecture of the RNA Polymerase-Spt4/5 Complex and Basis of Universal Transcription Processivity. *EMBO J.* 30 (7), 1302–1310. doi:10.1038/emboj.2011.64

Mayer, C., Bierhoff, H., and Grummt, I. (2005). The Nucleolus as a Stress Sensor: JNK2 Inactivates the Transcription Factor TIF-IA and Down-Regulates rRNA Synthesis. *Genes Develop.* 19 (8), 933–941. doi:10.1101/gad.333205

Michels, A. A., Robitaille, A. M., Buczynski-Ruchonnet, D., Hodroj, W., Reina, J. H., Hall, M. N., et al. (2010). mTORC1 Directly Phosphorylates and Regulates Human MAF1. *Mol. Cel Biol* 30 (15), 3749–3757. doi:10.1128/MCB.00319-10

Milkereit, P., and Tschochner, H. (1998). A Specialized Form of RNA Polymerase I, Essential for Initiation and Growth-dependent Regulation of rRNA Synthesis, Is Disrupted during Transcription. *EMBO J.* 17 (13), 3692–3703. doi:10.1093/emboj/17.13.3692

Mitsuzawa, H., Kanda, E., and Ishihama, A. (2003). Rpb7 Subunit of RNA Polymerase II Interacts with an RNA-Binding Protein Involved in Processing of Transcripts. *Nucleic Acids Res.* 31 (16), 4696–4701. doi:10.1093/nar/gkg688

Moir, R. D., Lee, J., Haeusler, R. A., Desai, N., Engelke, D. R., and Willis, I. M. (2006). Protein Kinase A Regulates RNA Polymerase III Transcription through the Nuclear Localization of Maf1. *Proc. Natl. Acad. Sci.* 103 (41), 15044–15049. doi:10.1073/pnas.0607129103

Moorefield, B., Greene, E. A., and Reeder, R. H. (2000). RNA Polymerase I Transcription Factor Rrn3 Is Functionally Conserved between Yeast and Human. *Proc. Natl. Acad. Sci.* 97 (9), 4724–4729. doi:10.1073/pnas.080063997

Mühlbacher, W., Sainsbury, S., Hemann, M., Hantsche, M., Neyer, S., Herzog, F., et al. (2014). Conserved Architecture of the Core RNA Polymerase II Initiation Complex. *Nat. Commun.* 5, 4310. doi:10.1038/ncomms5310

Naar, A. M., Taatjes, D. J., Zhai, W., Nogales, E., and Tjian, R. (2002). Human CRSP Interacts with RNA Polymerase II CTD and Adopts a Specific CTD-Bound Conformation. *Genes Dev.* 16 (11), 1339–1344. doi:10.1101/gad.987602

Nogales, E., and Greber, B. J. (2019). High-resolution Cryo-EM Structures of TFIIH and Their Functional Implications. *Curr. Opin. Struct. Biol.* 59, 188–194. doi:10.1016/j.sbi.2019.08.002

Nogales, E., Louder, R. K., and He, Y. (2017). Structural Insights into the Eukaryotic Transcription Initiation Machinery. *Annu. Rev. Biophys.* 46, 59–83. doi:10.1146/annurev-biophys-070816-033751

Oler, A. J., and Cairns, B. R. (2012). PP4 Dephosphorylates Maf1 to Couple Multiple Stress Conditions to RNA Polymerase III Repression. *EMBO J.* 31 (6), 1440–1452. doi:10.1038/emboj.2011.501

Onodera, Y., Haag, J. R., Ream, T., Nunes, P. C., Pontes, O., and Pikaard, C. S. (2005). Plant Nuclear RNA Polymerase IV Mediates siRNA and DNA Methylation-dependent Heterochromatin Formation. *Cell* 120 (5), 613–622. doi:10.1016/j.cell.2005.02.007

Panday, A., Gupta, A., Srinivasa, K., Xiao, L., Smith, M. D., and Grove, A. (2017). DNA Damage Regulates Direct Association of TOR Kinase with the RNA Polymerase II-Transcribed HMO1 Gene. *MBoC* 28 (18), 2449–2459. doi:10.1091/mbc.E17-01-0024

Pappas, D. L., Jr., and Hampsey, M. (2000). Functional Interaction between Ssu72 and the Rpb2 Subunit of RNA Polymerase II in *Saccharomyces cerevisiae*. *Mol. Cel Biol* 20 (22), 8343–8351. doi:10.1128/mcb.20.22.8343-8351.2000

Patel, A. B., Greber, B. J., and Nogales, E. (2020). Recent Insights into the Structure of TFIID, its Assembly, and its Binding to Core Promoter. *Curr. Opin. Struct. Biol.* 61, 17–24. doi:10.1016/j.sbi.2019.10.001

Peng, L., Li, E.-M., and Xu, L.-Y. (2020). From Start to End: Phase Separation and Transcriptional Regulation. *Biochim. Biophys. Acta (Bba) - Gene Regul. Mech.* 1863 (12), 194641. doi:10.1016/j.bbagrm.2020.194641

Perales, R., and Bentley, D. (2009). "Cotranscriptionality": The Transcription Elongation Complex as a Nexus for Nuclear Transactions. *Mol. Cel* 36 (2), 178–191. doi:10.1016/j.molcel.2009.09.018

Plaschka, C., Larivière, L., Wenzeck, L., Seizl, M., Hemann, M., Tegunov, D., et al. (2015). Architecture of the RNA Polymerase II-Mediator Core Initiation Complex. *Nature* 518 (7539), 376–380. doi:10.1038/nature14229

Pultz, D., Bennetzen, M. V., Rødkær, S. V., Zimmermann, C., Enserink, J. M., Andersen, J. S., et al. (2012). Global Mapping of Protein Phosphorylation

Events Identifies Ste20, Sch9 and the Cell-Cycle Regulatory Kinases Cdc28/Pho85 as Mediators of Fatty Acid Starvation Responses in *Saccharomyces cerevisiae*. *Mol. Biosyst.* 8 (3), 796–803. doi:10.1039/c2mb05356j

Ramsay, E. P., Abascal-Palacios, G., Daiss, J. L., King, H., Gouge, J., Pilsl, M., et al. (2020). Structure of Human RNA Polymerase III. *Nat. Commun.* 11 (1), 6409. doi:10.1038/s41467-020-20262-5

Richard, S., Gross, L., Fischer, J., Bendalak, K., Ziv, T., Urim, S., et al. (2021). Numerous Post-translational Modifications of RNA Polymerase II Subunit Rpb4/7 Link Transcription to Post-transcriptional Mechanisms. *Cel Rep.* 34 (2), 108578. doi:10.1016/j.celrep.2020.108578

Robinson, P. J., Trnka, M. J., Bushnell, D. A., Davis, R. E., Mattei, P.-J., Burlingame, A. L., et al. (2016). Structure of a Complete Mediator-RNA Polymerase II Pre-initiation Complex. *Cell* 166 (6), 1411–1422 e1416. doi:10.1016/j.cell.2016.08.050

Roeder, R. G. (2019). 50+ Years of Eukaryotic Transcription: an Expanding Universe of Factors and Mechanisms. *Nat. Struct. Mol. Biol.* 26 (9), 783–791. doi:10.1038/s41594-019-0287-x

Runner, V. M., Podolny, V., and Buratowski, S. (2008). The Rpb4 Subunit of RNA Polymerase II Contributes to Cotranscriptional Recruitment of 3′ Processing Factors. *Mol. Cel Biol* 28 (6), 1883–1891. doi:10.1128/MCB.01714-07

Sainsbury, S., Bernecky, C., and Cramer, P. (2015). Structural Basis of Transcription Initiation by RNA Polymerase II. *Nat. Rev. Mol. Cel Biol* 16 (3), 129–143. doi:10.1038/nrm3952

Schier, A. C., and Taatjes, D. J. (2020). Structure and Mechanism of the RNA Polymerase II Transcription Machinery. *Genes Dev.* 34 (7-8), 465–488. doi:10.1101/gad.335679.119

Sdano, M. A., Fulcher, J. M., Palani, S., Chandrasekharan, M. B., Parnell, T. J., Whitby, F. G., et al. (2017). A Novel SH2 Recognition Mechanism Recruits Spt6 to the Doubly Phosphorylated RNA Polymerase II Linker at Sites of Transcription. *Elife* 6. doi:10.7554/eLife.28723

Sharma, N., and Kumari, R. (2013). Rpb4 and Rpb7: Multifunctional Subunits of RNA Polymerase II. *Crit. Rev. Microbiol.* 39 (4), 362–372. doi:10.3109/1040841X.2012.711742

Šoštarić, N., O'Reilly, F. J., Giansanti, P., Heck, A. J. R., Gavin, A. C., and van Noort, V. (2018). Effects of Acetylation and Phosphorylation on Subunit Interactions in Three Large Eukaryotic Complexes. *Mol. Cel Proteomics* 17 (12), 2387–2401. doi:10.1074/mcp.RA118.000892

Soulard, A., Cremonesi, A., Moes, S., Schütz, F., Jenö, P., and Hall, M. N. (2010). The Rapamycin-Sensitive Phosphoproteome Reveals that TOR Controls Protein Kinase A toward Some but Not All Substrates. *MBoC* 21 (19), 3475–3486. doi:10.1091/mbc.E10-03-0182

Swaney, D. L., Beltrao, P., Starita, L., Guo, A., Rush, J., Fields, S., et al. (2013). Global Analysis of Phosphorylation and Ubiquitylation Cross-Talk in Protein Degradation. *Nat. Methods* 10 (7), 676–682. doi:10.1038/nmeth.2519

Tan, Q., Prysak, M. H., and Woychik, N. A. (2003). Loss of the Rpb4/Rpb7 Subcomplex in a Mutant Form of the Rpb6 Subunit Shared by RNA Polymerases I, II, and III. *Mol. Cel Biol* 23 (9), 3329–3338. doi:10.1128/mcb.23.9.3329-3338.2003

Thomas, M. C., and Chiang, C.-M. (2006). The General Transcription Machinery and General Cofactors. *Crit. Rev. Biochem. Mol. Biol.* 41 (3), 105–178. doi:10.1080/10409230600648736

Torreira, E., Louro, J. A., Pazos, I., González-Polo, N., Gil-Carton, D., Duran, A. G., et al. (2017). The Dynamic Assembly of Distinct RNA Polymerase I Complexes Modulates rDNA Transcription. *Elife* 6, e20832. doi:10.7554/eLife.20832

Tsang, C. K., and Zheng, X. F. S. (2007). TOR-in(g) the Nucleus. *Cell Cycle* 6 (1), 25–29. doi:10.4161/cc.6.1.3675

Vannini, A., and Cramer, P. (2012). Conservation between the RNA Polymerase I, II, and III Transcription Initiation Machineries. *Mol. Cel* 45 (4), 439–446. doi:10.1016/j.molcel.2012.01.023

Vorländer, M. K., Baudin, F., Moir, R. D., Wetzel, R., Hagen, W. J. H., Willis, I. M., et al. (2020a). Structural Basis for RNA Polymerase III Transcription Repression by Maf1. *Nat. Struct. Mol. Biol.* 27 (3), 229–232. doi:10.1038/s41594-020-0383-y

Vorländer, M. K., Jungblut, A., Karius, K., Baudin, F., Grötsch, H., Kosinski, J., et al. (2020b). Structure of the TFIIIC Subcomplex τA Provides Insights into RNA Polymerase III Pre-initiation Complex Formation. *Nat. Commun.* 11 (1), 4905. doi:10.1038/s41467-020-18707-y

Wang, Y., Fairley, J. A., and Roberts, S. G. E. (2010). Phosphorylation of TFIIB Links Transcription Initiation and Termination. *Curr. Biol.* 20 (6), 548–553. doi:10.1016/j.cub.2010.01.052

Wang, Y., and Ma, H. (2015). Step-wise and Lineage-specific Diversification of Plant RNA Polymerase Genes and Origin of the Largest Plant-specific Subunits. *New Phytol.* 207 (4), 1198–1212. doi:10.1111/nph.13432

Warner, J. R. (1999). The Economics of Ribosome Biosynthesis in Yeast. *Trends Biochem. Sci.* 24 (11), 437–440. doi:10.1016/s0968-0004(99)01460-7

Wei, Y., and Zheng, X. F. S. (2009). Sch9 Partially Mediates TORC1 Signaling to Control Ribosomal RNA Synthesis. *Cell Cycle* 8 (24), 4085–4090. doi:10.4161/cc.8.24.10170

Werner, F., and Grohmann, D. (2011). Evolution of Multisubunit RNA Polymerases in the Three Domains of Life. *Nat. Rev. Microbiol.* 9 (2), 85–98. doi:10.1038/nrmicro2507

Willis, I. M., and Moir, R. D. (2018). Signaling to and from the RNA Polymerase III Transcription and Processing Machinery. *Annu. Rev. Biochem.* 87, 75–100. doi:10.1146/annurev-biochem-062917-012624

Yurko, N. M., and Manley, J. L. (2018). The RNA Polymerase II CTD "orphan" Residues: Emerging Insights into the Functions of Tyr-1, Thr-4, and Ser-7. *Transcription* 9 (1), 30–40. doi:10.1080/21541264.2017.1338176

Zaros, C., Briand, J.-F., Boulard, Y., Labarre-Mariotte, S., Garcia-Lopez, M. C., Thuriaux, P., et al. (2007). Functional Organization of the Rpb5 Subunit Shared by the Three Yeast RNA Polymerases. *Nucleic Acids Res.* 35 (2), 634–647. doi:10.1093/nar/gkl686

Zhang, X., Henderson, I. R., Lu, C., Green, P. J., and Jacobsen, S. E. (2007). Role of RNA Polymerase IV in Plant Small RNA Metabolism. *Proc. Natl. Acad. Sci.* 104 (11), 4536–4541. doi:10.1073/pnas.0611456104

Zhang, Y., Smith, A. D., Renfrow, M. B., and Schneider, D. A. (2010). The RNA Polymerase-Associated Factor 1 Complex (Paf1C) Directly Increases the Elongation Rate of RNA Polymerase I and Is Required for Efficient Regulation of rRNA Synthesis. *J. Biol. Chem.* 285 (19), 14152–14159. doi:10.1074/jbc.M110.115220

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Molecular Biosciences

REVIEW
published: 30 July 2021
doi: 10.3389/fmolb.2021.696438

RNA Polymerase III Subunit Mutations in Genetic Diseases

*Elisabeth Lata[1†], Karine Choquet[2†], Francis Sagliocco[1], Bernard Brais[3], Geneviève Bernard[4,5,6] and Martin Teichmann[1]**

[1]Bordeaux University, Inserm U 1212, CNRS UMR 5320, ARNA laboratory, Bordeaux, France, [2]Department of Genetics, Harvard Medical School, Boston, MA, United States, [3]Montreal Neurological Institute, McGill University, Montreal, QC, Canada, [4]Departments of Neurology and Neurosurgery, Pediatrics and Human Genetics, McGill University, Montreal, QC, Canada, [5]Department of Specialized Medicine, Division of Medical Genetics, McGill University Health Center, Montreal, QC, Canada, [6]Child Health and Human Development Program, Research Institute of the McGill University Health Center, Montreal, QC, Canada

RNA polymerase (Pol) III transcribes small untranslated RNAs such as 5S ribosomal RNA, transfer RNAs, and U6 small nuclear RNA. Because of the functions of these RNAs, Pol III transcription is best known for its essential contribution to RNA maturation and translation. Surprisingly, it was discovered in the last decade that various inherited mutations in genes encoding nine distinct subunits of Pol III cause tissue-specific diseases rather than a general failure of all vital functions. Mutations in the POLR3A, POLR3C, POLR3E and POLR3F subunits are associated with susceptibility to varicella zoster virus-induced encephalitis and pneumonitis. In addition, an ever-increasing number of distinct mutations in the POLR3A, POLR3B, POLR1C and POLR3K subunits cause a spectrum of neurodegenerative diseases, which includes most notably hypomyelinating leukodystrophy. Furthermore, other rare diseases are also associated with mutations in genes encoding subunits of Pol III (POLR3H, POLR3GL) and the BRF1 component of the TFIIIB transcription initiation factor. Although the causal relationship between these mutations and disease development is widely accepted, the exact molecular mechanisms underlying disease pathogenesis remain enigmatic. Here, we review the current knowledge on the functional impact of specific mutations, possible Pol III-related disease-causing mechanisms, and animal models that may help to better understand the links between Pol III mutations and disease.

Keywords: RNA polymerase III (Pol III), Pol III-related hypomyelinating leukodystrophy (POLR3-HLD), innate immunity, neurodegenerative disease, Pol III subunits (POLR3A, POLR3B, POLR3C, POLR3E, POLR3F, POLR3GL, POLR3H, POLR3K, POLR1C)

OPEN ACCESS

Edited by:
Olga Calvo,
Instituto de Biología Funcional y Genómica, Spain

Reviewed by:
Jian Wu,
Shanghai Jiao Tong University, China
Francisco Navarro,
University of Jaén, Spain

****Correspondence:***
Martin Teichmann
Martin.Teichmann@inserm.fr

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Protein and RNA Networks, a section of the journal Frontiers in Molecular Biosciences

Received: *16 April 2021*
Accepted: *21 July 2021*
Published: *30 July 2021*

Citation:
Lata E, Choquet K, Sagliocco F, Brais B, Bernard G and Teichmann M (2021) RNA Polymerase III Subunit Mutations in Genetic Diseases. Front. Mol. Biosci. 8:696438. doi: 10.3389/fmolb.2021.696438

INTRODUCTION

Transcription is essential to make genome-encoded information accessible, which is a basic condition for the creation of all life forms. It represents the first step in gene expression and is coordinated by regulatory mechanisms allowing cells to respond not only according to their own needs, but also, if necessary, to demands from neighboring cells or to differentiation programs.

Nuclear RNA polymerases are responsible for the transcription of genomic DNA into RNA. In eukaryotes, up to five different nuclear DNA-dependent RNA polymerases (Pol I-V) have been described, each of which transcribes specific groups of genes. RNA polymerases I to III are expressed by all eukaryotes. RNA polymerase (Pol) I transcribes the large ribosomal gene, which is present in up to several hundred copies within eukaryotic genomes. The resulting ribosomal (r)RNAs represent

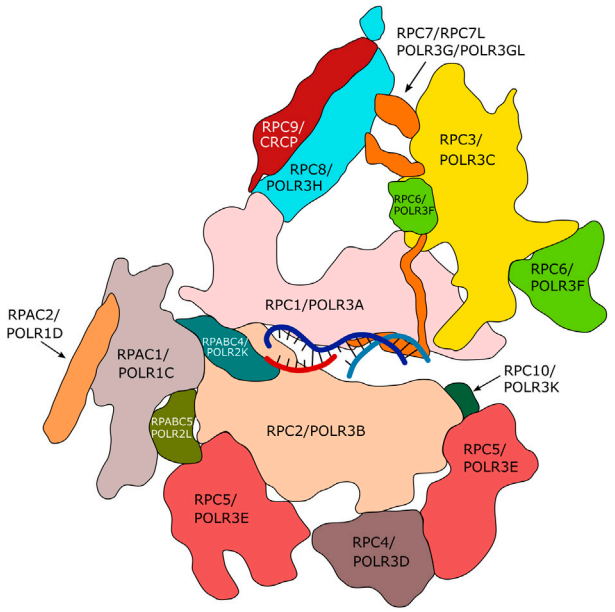


FIGURE 1 | Subunits of human RNA polymerase III. Fourteen of the seventeen subunits of human Pol III are appropriately assigned. Subunits POLR2E, POLR2F and POLR2H are not displayed. The Figure is inspired by the structure of human Pol III published by Girbig et al. (2021) (PDB7AE3). The RPC4/RPC5 hetero-dimeric complex is required by termination/reinitiation. The RPC3/RPC6/RPC7 hetero-trimeric subcomplex is required for transcription initiation.

the major constituents of ribosomes (Khatter et al., 2017). Pol II is responsible for transcription of all protein-coding genes and is also involved in the expression of several non-coding RNAs (Roeder, 2019 and references therein). Pol III synthesizes a variety of small (<350 nt) and highly expressed RNAs (e.g. 5S rRNA, transfer RNA (tRNA), U6 RNA) that do not code for proteins (Dieci et al., 2007). RNA polymerases IV and V, which have been described exclusively in plants, are involved in RNA-dependent gene silencing (Zhou and Law, 2015). In terms of protein composition, Pol III is the most complex enzyme performing DNA-dependent transcription in eukaryotic cells. It is composed of 17 subunits (in contrast to 14 subunits in Pol I and 12 subunits in Pol II; **Figure 1**).

Here, we will review recent discoveries connecting Pol III (also referred to as POLR3) transcription to diseases. We will focus on mutations in genes encoding subunits of the Pol III transcription system that have been associated with microbial infections or with neurodegenerative diseases, including Pol III-related hypomyelinating leukodystrophy (POLR3-HLD). These mutations are also referred to as pathogenic variants in medical genetics. First, we will give an overview of the regulation of Pol III expression. Subsequently, we will review mutations in genes encoding Pol III subunits that were shown to be altered in disease and discuss potential underlying pathophysiological mechanisms that may depend on altered expression of Pol III transcripts. Finally, we will discuss the role of Pol III in innate immunity and related diseases.

TRANSCRIPTION BY RNA POLYMERASE III

The Pol III Promoter Types

Three main promoter types are employed by Pol III: types 1 and 2 have gene internal elements, while type 3 possesses regulatory elements in the 5′ region upstream of the transcriptional start site (TSS) (**Figures 2A–C**; reviewed in Dumay-Odelot et al. (2014)). The type 1 promoter, consisting of an A- and C-Box, is exclusively used by the 5S rRNA genes (**Figure 2A**). Expression of tRNA and the adenoviral VA1 and VA2 genes depends on type 2 promoters, which are comprised of A- and B-Boxes. Furthermore, type 2 promoters are encountered in short interspersed nuclear elements (SINEs) (**Figure 2B**). Type 3 promoters regulate transcription of the U6 small nuclear (sn) RNA, the H1 RNA component of RNase P, the RNA component of RNase MRP, Y RNAs and the 7SK RNA (Dieci et al., 2007; **Figure 2C**). The type 3 gene regulatory elements include a TATA-box, a proximal sequence element (PSE) and a distal sequence element (DSE), which are respectively located approximately 30, 50 and 200 nt upstream of the TSS. This promoter type emerged during evolution from single cell to multicellular eukaryotes and has been accompanied by the appearance of new transcription factors (Teichmann et al., 2010; Girbig et al., 2021). In addition, there are promoter variations, which are composed of combinations of regulatory elements from type 2 and 3 promoters, as well as of enhancer elements that are known from Pol II transcription. Such hybrid promoter-dependent genes include the selenocysteine tRNA gene ($tRNA^{Sec}$), the Epstein Barr virus EBER gene, the 7SL RNA gene, vault RNA genes (Howe and Shu, 1989; Bredow et al., 1990a,b; Carbon and Krol, 1991; Kickhoefer et al., 2003) and the BC200 RNA gene (Khanam et al., 2007) (**Figure 2D**).

The Pol III Transcription Factors

Expression of genes regulated by intragenic promoters requires the six subunit transcription factor TFIIIC (type 1 and 2 promoters) and the transcription factor TFIIIA (type 1 promoter only) to recruit the transcription initiation factor TFIIIB-β (**Figures 2A,B**). The regulatory elements upstream of the TSS in type 3 and the promoter of the selenocysteine tRNA ($tRNA^{Sec}$) gene are recognized by STAF/ZNF143 and OCT1 (DSE), as well as by SNAPc/PTF (PSE), which stimulate the recruitment of TFIIIB-α to the TSS, whereupon Pol III is recruited (reviewed in Schramm and Hernandez (2002), Dumay-Odelot et al. (2010); **Figure 2C**). TFIIIB-α is composed of the TATA-binding protein (TBP), the B double prime 1 (BDP1) component and the TFIIB-related factor 2 (BRF2), whereas TFIIIB-β contains the TFIIB-related factor 1 (BRF1) instead of BRF2 (**Figure 2**) (Teichmann and Seifart, 1995; Teichmann et al., 2000; Schramm et al., 2000; reviewed in Schramm and Hernandez (2002), Dumay-Odelot et al. (2010)). Hybrid promoters display gene-specific transcription factor requirements. While transcription of the 7SL and EBER genes is stimulated by binding of the Pol II transcriptional activator ATF upstream of the TATA-like box, activation of the $tRNA^{Sec}$ gene is dependent on transcription factors that recognize the PSE and DSE (SNAPc/PTF, STAF/ZNF143; Schaub et al., 2000; **Figure 2D**), which also regulate the transcription rate of type

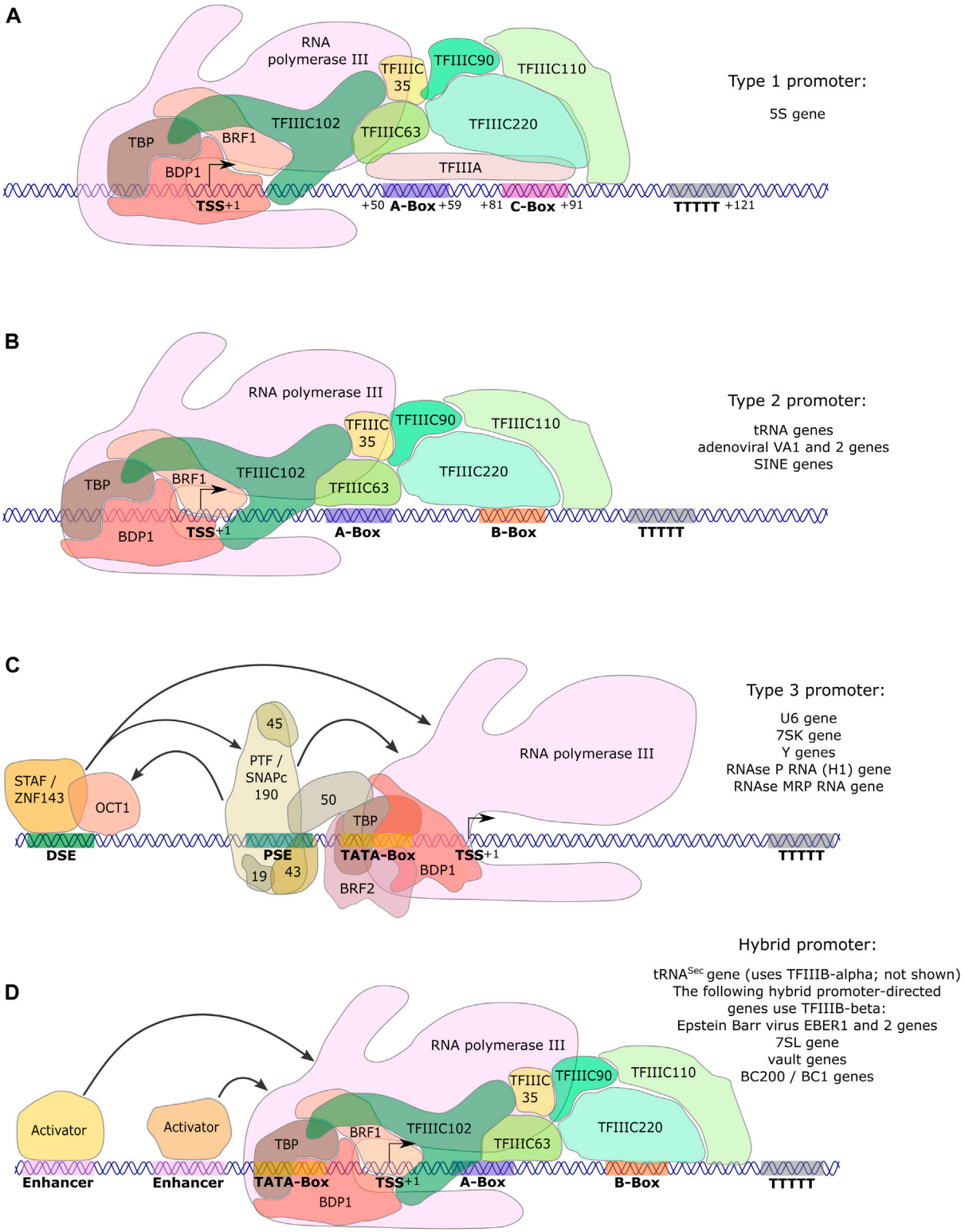


FIGURE 2 | Promoters directing human RNA polymerase III transcription. Pol III type 1 and type 2 genes contain gene-internal promoter elements. **(A)** Type 1 gene transcription of 5S ribosomal (r)RNA is directed by A- and C-boxes that are located relative to the transcription start site (TSS) as indicated. These promoter elements are bound by TFIIIA, permitting the recruitment of TFIIIC and subsequently of TFIIIB-β (composed of TBP, BDP1 and BRF1), which altogether recruit Pol III. **(B)** Type 2 genes (tRNA genes, VA1, VA2, SINEs) contain A- and B-boxes as promoter elements at varying positions relative to the TSS. They are bound by TFIIIC, subsequently

(*Continued*)

FIGURE 2 | allowing recruitment of TFIIIB-β and in turn of Pol III. **(C)** Type 3 gene regulatory elements are entirely located upstream (5′) of the TSS. They are composed of a TATA-box at -30, as well as a proximal sequence element (PSE) and a distal sequence element (DSE) at variable distances with respect to the TSS, depending on the gene. Transcriptional activators (STAF/ZNF143; SNAPc/PTF) bind to the DSE and PSE, respectively, and regulate transcriptional activity. The TATA-box is the only promoter element required for directing Pol III to the TSS (Teichmann et al., 1997). It is bound by TFIIIB-α (composed of TBP, BDP1 and BRF2), which in turn recruits Pol III. **(D)** Hybrid promoter-directed transcription is regulated by gene-internal elements of type 2 promoters (A- and B-boxes) and additionally by gene regulatory elements upstream of the TSS. All these elements vary in their distance to the TSS depending on the gene. In the presence of the PSE ($tRNA^{Sec}$ gene), TFIIIB-α is recruited (not shown in the Figure). In the case of all other enhancer-activator combinations with gene-internal A- and B-boxes (EBER 1 and 2; 7SL; vault; BC1 and BC200), TFIIIB-β is recruited, allowing the subsequent recruitment of RNA polymerase III. For all promoter types, the stretch of T's represents the transcription termination site. Arrows in panels C. and D. symbolize protein-protein-interactions that contribute to activation of Pol III transcription from these promoters. Promoter types were reviewed in Dumay-Odelot et al. (2014).

3 promoters in multicellular organisms (Carbon and Krol, 1991; Meissner et al., 1994; reviewed in Dieci et al. (2007), Dumay-Odelot et al. (2010)). Furthermore, only the promoters that depend on a PSE and SNAPc/PTF transcription factors recruit the BRF2-containing TFIIIB-α transcription initiation factor, whereas other enhancer-activator combinations with gene-internal A- and B-Boxes result in the recruitment of the BRF1-containing TFIIIB-β.

Pol III Transcription: High Efficiency Through Compact Gene Organization

Human genes transcribed by Pol III are composed of maximally ~300–350 nucleotides, with the longest RNAs generated by transcription of SINEs as well as of 7SK and 7SL genes. Functional elements required for regulating gene expression rate are all found within less than 500 base pairs relative to the TSS. In contrast, enhancer elements are often distributed over tens or hundreds of kilobases in the case of Pol II genes (reviewed in Dieci et al. (2007), Dumay-Odelot et al. (2014)). The Pol III gene regulatory elements include DNA sequences showing enhancer-like features to regulate transcription levels ([PSE]; [DSE]; B-box) and promoter elements (TATA-like box; A-box) that are required for positioning Pol III at the TSS. At the protein level, the functional entities for DNA recognition and polymerase recruitment described in the Pol II transcription system are also found in Pol III transcription. However, some of the functions performed by general transcription factors in the Pol II system appear to be fully integrated into the Pol III enzyme. Indeed, structural and/or functional similarities with Pol II transcription factors were identified in four of the 17 Pol III subunits. TFIIE-comparable structure-function modules were described in the POLR3C (RPC3) and POLR3F (RPC6) subunits (Lefevre et al., 2011; Blombach et al., 2015; Ramsay et al., 2020; Girbig et al., 2021; Li G. et al., 2021) and similarities to TFIIF were found in the POLR3D (RPC4) and POLR3E (RPC5) subunits (Ramsay et al., 2020; Li L. et al., 2021; Wang et al., 2021; Girbig et al., 2021 and references therein; **Figure 1**). The complex composition of Pol III by 17 subunits can probably be explained by the structure of the genes that are transcribed by this enzyme. These genes are short and often possess gene internal or hybrid promoters (type 1, type 2 without or with regulatory elements upstream the TSS; **Figure 2**), which are bound by TFIIIC. As a consequence, TFIIIC needs to be removed to allow for Pol III to progress through the gene during transcription. Additional transcription factors might complicate this task. Therefore, stable integration of TFIIE- and TFIIF-like activities into polymerase subunits may contribute to a highly efficient transcription mode deemed "facilitated reinitiation" (Dieci and Sentenac, 1996; Dieci et al., 2002; Ferrari et al., 2004).

Furthermore, activities attributed to either general transcription factors or transcriptional activators in the Pol II system were found in the same protein or protein complex in the Pol III transcription system. On the one hand, *Xenopus laevis* TFIIIA is indispensable to recognize the 5S type 1 promoter (A- and C-box; **Figure 2A**), which corresponds to a function attributed to a general transcription factor. On the other hand, it possesses a transcriptional activation domain that is not needed for promoter recognition but is essential for transcriptional activation. Without this 14-amino acid domain at the C-terminus of *Xenopus laevis* TFIIIA, 5S rRNA gene transcription is undetectable *in vitro* (Mao and Darby, 1993). Furthermore, it has been reported that the three most C-terminal of the nine TFIIIA zinc fingers also exert a higher influence on transcription rate than on promoter recognition (Del Rio and Setzer, 1993).

The functions of TFIIIC in promoter recognition and transcriptional activation can also be separated. In the yeast *Saccharomyces cerevisiae*, the type 2 promoter of the U6 gene is localized partly within the transcribed region and partly downstream of the RNA coding sequence. The intragenic A-box is involved in start site selection along with the TATA box, whereas the B-box is required downstream of the transcription termination site for transcriptional activation *in vivo*. Importantly, the orientation of the B-box is irrelevant for transcriptional activation, demonstrating characteristics of a typical enhancer element (Gabrielsen and Sentenac, 1991; Burnol et al., 1993). Since *S. cerevisiae* TFIIIC is composed of two submodules, τA and τB, which bind to A-Box and B-Box, respectively, general transcription factor activity can be assigned to τA and activator-like functions to τB (Baker et al., 1987; Vorländer et al., 2020). It has not been determined whether this separation of transcriptional activities also holds true for TFIIIC in higher eukaryotes.

In summary, it should be noted that the compact Pol III transcription system combines within the same proteins or protein complexes the functions that are either attributed to transcriptional activators or to general transcription factors in the Pol II system. In addition, functions of some general Pol II transcription factors have been intrinsically integrated into Pol III.

The compact organization of Pol III genes and their requirement for a small number of regulatory DNA elements, the limited number of Pol III transcription factors as well as the major functions that were described for the most prominent Pol III-transcribed RNAs (tRNAs and 5S rRNA in translation; U6

snRNA in mRNA splicing) led to the suggestion that Pol III transcription fulfills primarily housekeeping functions in cells. These housekeeping functions supporting RNAs are thought to be essential for cell survival, but it was long assumed that their expression did not require any regulation since they are thought to be provided in excess compared to the needs of cells (reviewed in Dieci et al. (2007)). The identification of mutations in genes encoding Pol III subunits that lead to the development of hypomorphic diseases, including neurodegenerative disorders, could be considered to result from a failure of these housekeeping functions. Alternatively, it could indicate that Pol III transcription or its RNA products require cell type-specific regulation, which could explain why cells of the central nervous system are more vulnerable than other cells in the body to a loss of homeostasis upon Pol III mutations.

In addition to housekeeping functions, several discoveries unraveled central roles of Pol III in regulatory rather than simply supportive cellular functions. It has become clear that Pol III transcription cannot be separated from the regulation of hypermorphic processes such as tumorigenesis (not discussed here but exemplified or reviewed in White (2008), Pavon-Eternod et al. (2009), Dumay-Odelot et al. (2010), Goodarzi et al. (2016), Durrieu-Gaillard et al. (2018), Gouge and Vannini (2018), Petrie et al. (2019), Yang et al. (2020), Yeganeh and Hernandez (2020)). Moreover, Pol III transcription is an integral part of innate immune defense mechanisms (Carter-Timofte et al., 2018b).

RNA POLYMERASE III AND DISEASES

Hypomyelinating Leukodystrophy and Related Disorders

Biallelic pathogenic variants in genes encoding Pol III subunits cause a wide spectrum of neurodegenerative disorders.

Within the past decade, it was discovered that biallelic pathogenic variants in six genes encoding subunits of Pol III cause a spectrum of rare inherited disorders (Bernard et al., 2011; Saitsu et al., 2011; Tetreault et al., 2011; Thiffault et al., 2015; Dorboz et al., 2018; Franca et al., 2019; Beauregard-Lacroix et al., 2020; Terhal et al., 2020). The hypomyelinating leukodystrophy (HLD) called 4H leukodystrophy was the first and most commonly identified disease associated with Pol III dysfunction. Since then, the phenotypic spectrum has continued to widen to include both milder and more severe neurodegenerative diseases, as well as rare forms of premature aging or impaired puberty, and gave rise to the name POLR3-related disorders. In this section, we will first describe the major clinical and genetic characteristics of each disease entity within this spectrum. Next, we will review the current state of knowledge on the possible pathogenic mechanisms underlying these diseases. It is important to note that the phenotypic heterogeneity of POLR3-related disorders suggests that several distinct disease mechanisms are likely responsible for different clinical manifestations, perhaps by affecting different functional domains of the enzyme and/or in a cell-type specific manner.

Leukodystrophies are a group of genetically determined diseases of the cerebral white matter (Vanderver et al., 2015; van der Knaap and Bugiani, 2017). They are divided according to their Magnetic Resonance Imaging (MRI) characteristics and whether the pathophysiological mechanism is thought to be a lack of myelin deposition during development (hypomyelinating) or alteration of myelin homeostasis (i.e. demyelination or other mechanisms) (Schiffmann and van der Knaap, 2009; Steenweg et al., 2010; Parikh et al., 2015). POLR3-HLD is now recognized as one of the most common hypomyelinating leukodystrophies (Schmidt et al., 2020). It is also referred to as 4H leukodystrophy, where the 4Hs represent the cardinal clinical features: Hypomyelination, Hypodontia and Hypogonadotropic Hypogonadism (Bernard and Vanderver, 1993; Vanderver et al., 2015). Clinical manifestations and anatomical structures involved in POLR3-HLD are shown and described in **Table 1** and **Figure 3**. From 2003 until 2011, before the discovery of the first causal genes, five distinct disorders were described that are now recognized as phenotypes of POLR3-HLD: leukodystrophy with oligodontia (Atrouni et al., 2003), 4H syndrome (Timmons et al., 2006), ataxia, delayed dentition and hypomyelination (Wolf et al., 2007), hypomyelination with cerebellar atrophy and hypoplasia of the corpus callosum (Sasaki et al., 2009), and tremor ataxia with central hypomyelination (Bernard et al., 2010; Tétreault et al., 2012).

Our group and others identified the first and most commonly mutated genes in POLR3-HLD, *POLR3A* and *POLR3B* (Bernard et al., 2011; Saitsu et al., 2011; Tétreault et al., 2011; Daoud et al., 2013). We later described a third, less commonly mutated gene, *POLR1C* (Thiffault et al., 2015), and also *POLR3K* as a fourth and rare causal gene (Dorboz et al., 2018). Patients with POLR3-HLD typically present in early childhood with motor delay or regression (Vanderver et al., 2013). POLR3-HLD primarily affects the central nervous system (CNS). The predominant neurological features are cerebellar (i.e. gait ataxia, dysmetria, dysarthria), followed by pyramidal (i.e. spasticity, brisk reflexes, etc., often affecting predominantly the lower extremities), extrapyramidal (especially dystonia) (Osterman et al., 2012; Al Yazidi et al., 2019) and cognitive (i.e. intellectual disability and/or cognitive regression) features (Wolf et al., 2014a; Gauquelin et al., 2019). The disease is progressive or neurodegenerative, resulting in progressive motor impairment leading to loss of ambulation, progressive dysarthria leading to loss of speech, progressive dysphagia leading to tube feeding dependency and eventually to premature death. Non-neurological features are typically but not universally present (Wolf et al., 2014a; Gauquelin et al., 2019) and include myopia, typically progressive over several years, dental abnormalities (e.g. hypodontia, oligodontia, delayed or abnormal pattern of tooth eruption, natal tooth/teeth, etc.) (Wolff et al., 2010) and endocrine abnormalities, typically but not exclusively, hypogonadotropic hypogonadism leading to arrested or absence of puberty (Potic et al., 2012, 2015; Pelletier et al., 2021). The MRI of patients with a typical POLR3-HLD is characterized by a specific and recognizable pattern of hypomyelination (Schiffmann and van der Knaap, 2009; Steenweg et al., 2010), with relative preservation of the myelination of certain structures (i.e. dentate nucleus, optic radiations, anterolateral nucleus of the thalamus, globus pallidus, and in

TABLE 1 | Description of the main clinical manifestations observed in POLR3-related disorders and anatomical structures involved.

Clinical manifestation	Description	Anatomical structure(s) involved
Neurological manifestations		
Cerebellar		Cerebellum and/or cerebellar tracts
Gait ataxia	Incoordination or clumsiness of gait	
Dysmetria	Incoordination in limb movements	
Dysarthria	Slurred and dysrhythmic speech	
Dysphagia	Difficulty swallowing	
Pyramidal tract signs		Corticospinal tracts
Spasticity	Velocity dependent increased muscle tone	
Brisk reflexes	Abnormally brisk stretch reflexes	
Extrapyramidal signs		Basal ganglia (striatum) connections
Dystonia	Movement disorder characterized by involuntary contractions of muscles leading to abnormal postures, twisting movements and/or tremor	
Non-neurological manifestations		
Hypodontia	Developmental absence of tooth/teeth	Teeth
Hypogonadotropic hypogonadism	Delayed/absent/arrested puberty, growth hormone deficiency	Pituitary gland
Endosteal sclerosis	Sclerosis of the endosteum (layer of vascular connective tissue lining the medullary cavities of bone)	Bone
Progeroid appearance, progeria	Aged appearance, premature aging	N/A

some cases, of the corticospinal tracts at the level of the posterior limb of the internal capsule), as shown in **Figure 3**. Atrophy of the cerebellum and thinning of the corpus callosum are commonly seen (Steenweg et al., 2010; La Piana et al., 2014; Wolf et al., 2014a) (**Figure 3**). Patients with POLR3-HLD require multidisciplinary care for their complex medical needs (Adang et al., 2017).

Although POLR3-HLD is the most common form of POLR3-related disorders, there is a spectrum of several disease entities caused by mutations in genes encoding Pol III subunits. At the most severe end of the spectrum are patients with a specific combination of *POLR3A* variants, leading to the severe striatal variant, which is clinically and radiologically distinct from the typical POLR3-HLD, with prominent involvement of the basal ganglia (**Figure 3**). These patients present at 2–3 months of life with developmental delay and regression and severe dysphagia (Perrier et al., 2020a; Harting et al., 2020; Hiraide et al., 2020). They develop respiratory failure and a significant proportion of them become bedridden and/or die during early childhood.

Another form of POLR3-related disorder is the Wiedemann-Rautenstrauch syndrome (WRS), caused by specific combinations of *POLR3A* mutations (Jay et al., 2016; Paolacci et al., 2018; Wambach et al., 2018). These patients present intrauterine growth retardation and post-natal failure to thrive, together with a progeroid appearance. They also typically have a triangular face, convex or pinched nose, a small mouth, sparse hair and lipodystrophy. Their fontanelles may be enlarged and pseudohydrocephalus with prominent scalp veins may be observed. Dental abnormalities reminiscent of POLR3-HLD can be seen, including natal tooth/teeth. Some of these patients have both WRS and POLR3-HLD.

At the other end of the spectrum are mild presentations. This category includes patients homozygous for the common *POLR3B* mutation c.1568T>A (p.Val523Glu), who may remain asymptomatic or paucisymptomatic until adulthood and even late adulthood (DeGasperis et al., 2020; Perrier et al., 2020a; Verberne et al., 2020; Wolf et al., 2014a). Also in this category are the patients with the mild striatal variant, without hypomyelination but with basal ganglia involvement on the brain MRI (**Figure 3**), who carry a very specific combination of *POLR3A* splice site variants (Azmanov et al., 2016). Another group presenting a milder presentation include patients with spastic ataxia and spastic paraparesis without hypomyelination (La Piana et al., 2016; Minnerop et al., 2017; Gauquelin et al., 2018; Rydning et al., 2019). Some patients with biallelic pathogenic variants in *POLR3B* can present mainly or uniquely with endocrine manifestations (Richards et al., 2017). Patients with mutations in *POLR3B* can present with cerebellar involvement and the bone manifestation of endosteal sclerosis (Ghoumid et al., 2017). Most recently, specific *de novo* pathogenic variants in *POLR3B* have been associated with ataxia, spasticity and demyelinating neuropathy without CNS hypomyelination (Djordjevic et al., 2021). Patients with biallelic variants in *POLR3GL* can present with endosteal hyperostosis and oligodontia (Terhal et al., 2020) or WRS (Beauregard-Lacroix et al., 2020). Finally, a homozygous variant in *POLR3H* has been associated with primary ovarian insufficiency (Franca et al., 2019).

Interestingly, although not technically a part of POLR3-related disorders, mutations in *BRF1*, encoding a subunit of the Pol III transcription factor TFIIIB-β, cause a cerebellar-facial-dental syndrome with clinical overlap with POLR3-related disorders (Borck et al., 2015; Jee et al., 2017), emphasizing the vulnerability of these tissues to Pol III dysfunction.

Although POLR3-related disorders, and more specifically POLR3-HLD, have been extensively characterized at the clinical and genetic levels, the functional consequences of the various mutations in genes encoding Pol III subunits are not well understood. To this date, no curative treatment is available and

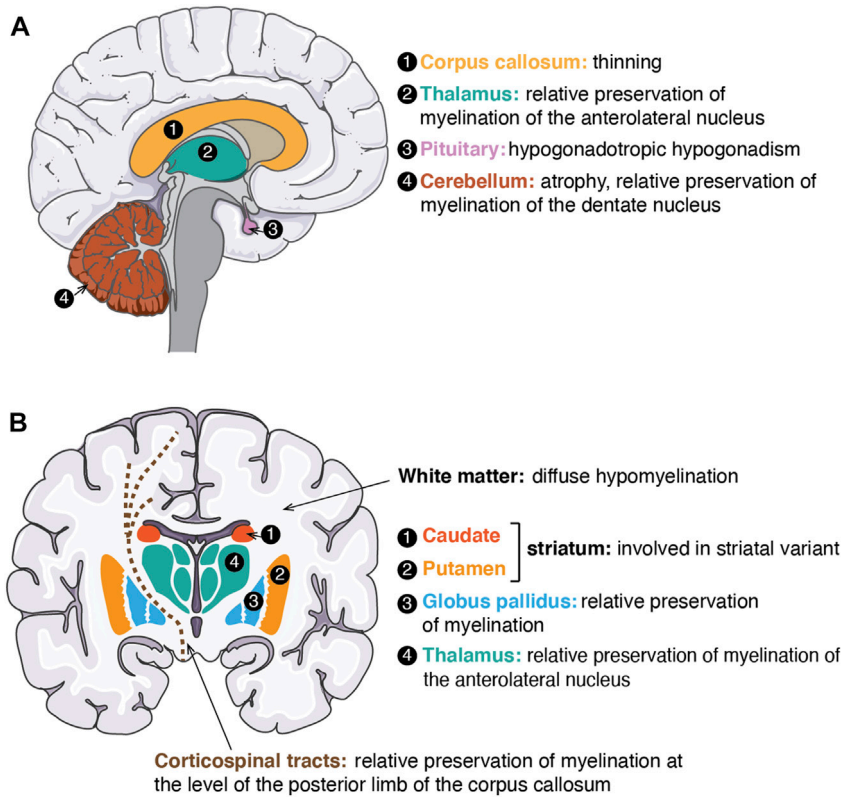


FIGURE 3 | Neuro-anatomical structures affected or for which myelination is preserved in POLR3-related disorders. **(A)** Schematic of a sagittal view of the human brain. Structures involved/preserved in POLR3-HLD are depicted in distinct colours and labeled with a number. On the right side, the names of anatomical structures corresponding to each number are shown in the same colour as the structure, followed by a description of how the structure is affected/preserved in POLR3-HLD. **(B)** Schematic of a coronal view of the human brain. Structures involved/preserved in the striatal variant of POLR3-related disorders (caudate and putamen) or in POLR3-HLD (other structures) are shown in distinct colours and labeled with a number. The legend on the right side follows the same description as in **(A)**. White matter (in white on the brain schematic) is indicated by an arrow. Corticospinal tracts are displayed as brown dashed lines. The figure was adapted from images available on https://smart.servier.com.

supportive care is the standard. Understanding the pathophysiology of these diseases will be key in order to develop therapies that can be tested in the pre-clinical setting and eventually translated to the clinic (Perrier et al., 2020b). Specifically, it remains enigmatic how mutations in a ubiquitously expressed and essential enzyme such as Pol III lead to disorders with clinical features that are largely restricted to the CNS and a few other tissues, all of which originate from neural crest cells. The pathophysiological mechanisms underlying such a wide spectrum of phenotypes are also unclear. Importantly, depending on the phenotype, different CNS cell types are affected, including oligodendrocytes, the cells that produce myelin, several populations of neurons, and/or their respective progenitor cells (**Figure 3** and **Table 1**). Hypomyelination in POLR3-HLD is thought to result from oligodendrocyte dysfunction, but cerebellar atrophy indicative of cerebellar neuron involvement is also observed (Vanderver et al., 2013; Wolf et al., 2014b). The other neurodegenerative phenotypes are postulated to result from abnormalities of cerebellar neurons (spastic ataxia) or of the basal ganglia, or brain atrophy (striatal variants) (Minnerop et al., 2017; Perrier et al., 2020a). Thus, identifying one unified disease mechanism for all POLR3-related disorders is not expected. Instead, distinct cell types may be differently affected by Pol III dysfunction, leading to a mechanistic diversity that would reflect the genetic and phenotypic heterogeneity of POLR3-related disorders.

There are two main pathophysiological hypotheses in the field, which are not mutually exclusive (**Figure 4**). Specifically for hypomyelination in POLR3-HLD, the first hypothesis states that hypofunctional Pol III, secondary to mutations in genes encoding Pol III subunits, leads to reduced levels of tRNA (either globally or of specific anticodons or isodecoders) and/or other small non-coding RNA (ncRNA) important for translation in a critical developmental period such as myelination. Since most of the myelination process occurs in a relatively short period of time, i.e. in the first 2 years of life in humans, it is thought that oligodendrocytes, the cells responsible for myelin production in the CNS, are more susceptible to a hypofunctional Pol III or reduced translation capacity due to the high metabolic requirements of producing myelin. Indeed, oligodendrocytes must produce a large amount of lipids and myelin-specific proteins to deposit on axons during myelination (Pfeiffer

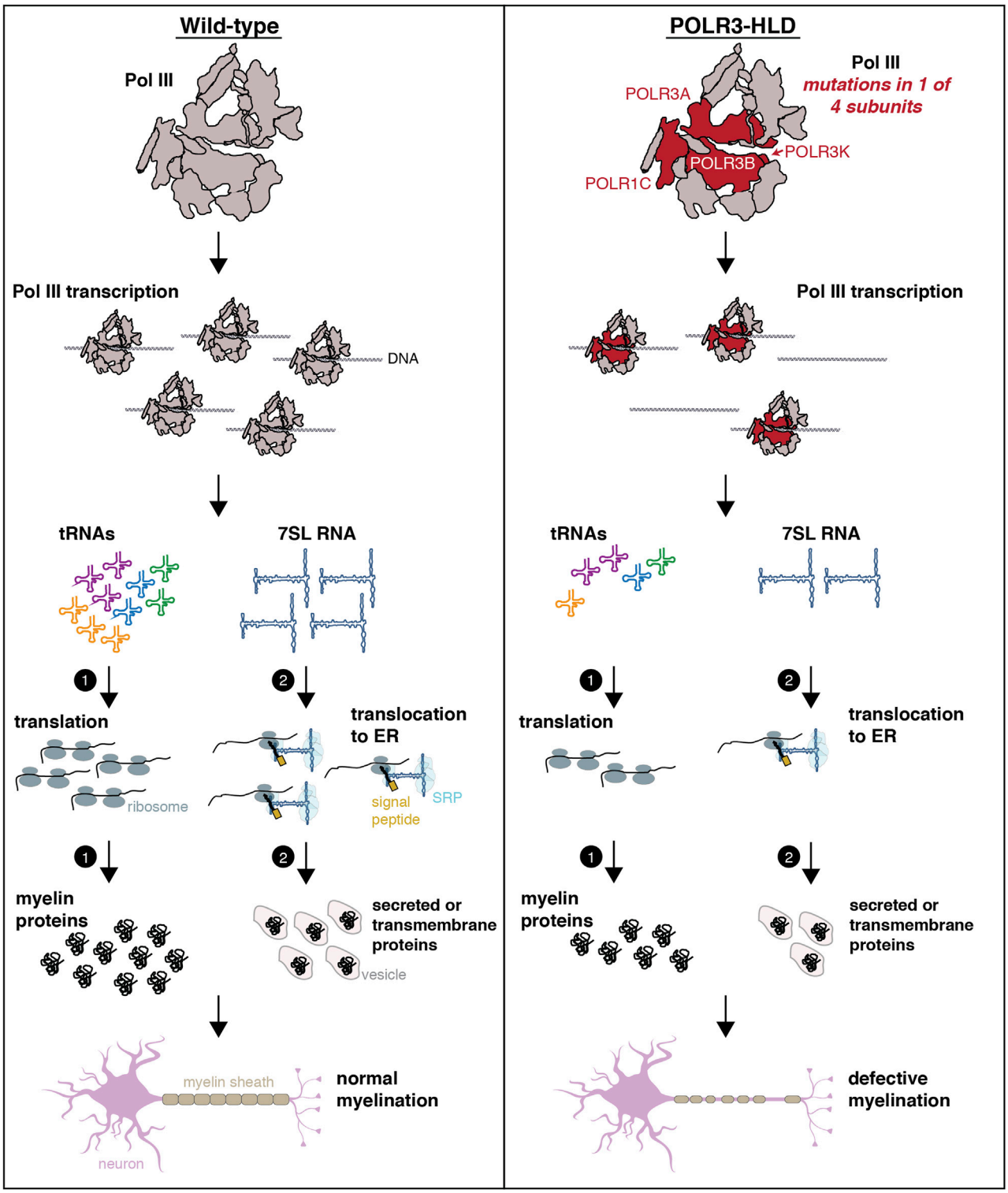


FIGURE 4 | Schematic representing possible mechanisms underlying POLR3-HLD. In wild-type conditions (healthy individuals), Pol III synthesizes small ncRNAs that play essential roles in housekeeping processes such as translation and co-translational targeting of nascent peptides, which are necessary for the production of myelin. In individuals with POLR3-HLD, it is hypothesized that mutations in Pol III subunits (POLR3A, POLR3B, POLR1C or POLR3K) result in reduced Pol III transcription and decreased levels of Pol III transcripts. The "tRNA-centric" hypothesis postulates that lower levels of tRNAs (either globally, for specific anticodons or for specific isodecoders) will impact translation and synthesis of proteins that are essential for myelination. Alternatively or in addition, reduced levels of other Pol III transcripts may contribute to POLR3-HLD pathogenesis through suboptimal performance of their respective functions that will particularly affect oligodendrocytes and/or neurons. An example is shown for 7SL RNA, where reduced levels of this ncRNA could impair translocation of secreted or transmembrane proteins to the ER, which could impact production of myelin. The schematic of the neuron was adapted from images available on https://smart.servier.com.

et al., 1993; Anitei and Pfeiffer, 2006). A hypomorphic Pol III would therefore impair global protein production during this critical developmental window leading to improper formation of myelin, ultimately causing the hypomyelination phenotype (Lin and Popko, 2009; Fröhlich et al., 2018; Torrent et al., 2018). This hypothesis is supported by the recent description of several hypomyelinating disorders caused by mutations in genes important for protein translation such as those encoding for tRNA-aminoacyl synthetases, including *DARS1, EPRS1* and *RARS1,* amongst others (Taft et al., 2013; Wolf et al., 2014a; Mendes et al., 2018, 2020). This raises the possibility that certain codons are particularly important for proper CNS function, and that reduced availability of the corresponding aminoacyl-tRNA through Pol III or tRNA-synthetase mutations is particularly detrimental to the CNS. Moreover, the CNS may have a lower threshold than other tissues for tolerating hypofunction of these enzymes. Another supportive element is that the brain MRI of patients with hypomyelination that carry mutations in genes encoding Pol III subunits show an arrested myelination, with myelination of the early myelinating structures, which are the smallest in size, but not the rest of the brain, suggesting that myelination began properly but could not be completed, perhaps due to impaired protein synthesis.

The second hypothesis, which can be generalized to all POLR3-related disorders, states that Pol III hypofunction leads to decreased levels of specific Pol III transcripts involved in transcription, RNA processing and/or translation, which preferentially perturbs the expression and/or translation of mRNAs that are essential for the development, survival and function of oligodendrocytes and/or neurons (Tétreault et al., 2011; Thiffault et al., 2015; Azmanov et al., 2016; Minnerop et al., 2017; Choquet et al., 2019a). An example of this hypothesis is shown in **Figure 4** for 7SL RNA, but it can be extended to any Pol III transcript and their specific function. These two non-mutually exclusive hypotheses may both contribute to the distinct phenotypes observed in POLR3-related disorders, with perturbation of different Pol III transcripts and their downstream functions having cell type- or temporal-specific effects.

Recent efforts to better understand the pathophysiological mechanisms of POLR3-related disorders have focused on three main areas: the impact of Pol III subunit mutations on biogenesis of the Pol III complex; the downstream consequences of mutations on the Pol III transcriptome; and the development of animal models of the disease.

Impact of Pol III Subunit Mutations on Enzyme Biogenesis

The recessive mode of inheritance and the nature of most disease-causing mutations (missense, splice site, truncating) in genes encoding Pol III subunits suggests a hypomorphic disease mechanism, either through decreased protein abundance or because of abnormal interactions of the mutated subunit with other subunits, with DNA or with RNA (Bernard et al., 2011). Given the genetic and phenotypic heterogeneity of POLR3-related disorders, distinct mutations may have different effects on mRNA or protein stability or on Pol III function itself, leading to different phenotypes or modulating disease severity. Decreased levels of mRNA or protein encoded by the mutated gene have been observed in fibroblasts, blood, white matter or cortex of individuals with *POLR3A* or *POLR3GL* mutations (Bernard et al., 2011; Azmanov et al., 2016; Minnerop et al., 2017; Perrier et al., 2020a; Báez-Becerra et al., 2020; Beauregard-Lacroix et al., 2020), the majority of which carried a truncating mutation on one allele. While mRNA or protein levels have not been examined extensively in individuals with missense mutations, two reports suggest that they are not always altered. First, mice homozygous for the *Polr3a* c.2015G>A (p.Gly672Glu) mutation had normal POLR3A protein levels (Choquet et al., 2017). Second, *POLR3K* mRNA levels were unchanged in individuals carrying missense mutations in this gene (Dorboz et al., 2018).

Missense mutations in *POLR3A*, *POLR3B* and *POLR1C* causing POLR3-HLD are located throughout the three genes without clear hotspots (Wolf et al., 2014a; Gauquelin et al., 2019; Ramsay et al., 2020; Li G. et al., 2021; Girbig et al., 2021) and affect most major structural regions (Arimbasseri and Maraia, 2016). Prior to the publication of the first Pol III yeast structures (Abascal-Palacios et al., 2018; Vorländer et al., 2018), the potential impact of *POLR3A* and *POLR3B* HLD mutations was predicted *in silico* by extrapolating them onto the yeast Pol II structure. This suggested that most of these amino acid changes would impair the interaction with other Pol III subunits or with the DNA template (Bernard et al., 2011; Saitsu et al., 2011; Tétreault et al., 2011; Wolf et al., 2014a). Similarly, the only reported *POLR3K* mutation was predicted using the yeast Pol III structure to decrease protein stability and to impair the interaction between POLR3K and POLR3B (Dorboz et al., 2018). Recently, the tridimensional structure of the human Pol III was resolved by cryogenic electron microscopy (Ramsay et al., 2020; Li L. et al., 2021; Girbig et al., 2021; Wang et al., 2021). First, Ramsay et al., mapped 47 POLR3-HLD mutations in *POLR3A, POLR3B* and *POLR1C* onto the Pol III structure, which revealed that they cluster in regions at the interface of several subunits and are predicted to disrupt these interfaces, consistent with the earlier predictions made using the yeast Pol II structure. Second, Girbig et al., investigated 110 point mutations found in patients with POLR3-related disorders and classified them into four types, showing that the majority of POLR3-HLD mutations are predicted to disturb the core of a given subunit (Type I) or are located at the interface between subunits and have the potential to impair complex assembly (Type III), while a smaller number affect functional elements such as the bridge helix or the trigger loop (Type II). Li et al., mapped several mutations in Treacher Collins syndrome (TCS), WRS and POLR3-HLD. They suggest that these mutations may impair complex integrity or enzymatic activity.

In addition to these *in silico* predictions, the effect of some mutations on Pol III complex assembly has been assessed experimentally (Thiffault et al., 2015; Choquet et al., 2017, 2019a, 2019b; Djordjevic et al., 2021). In this series of experiments, the wild-type or mutated Pol III subunit of interest was exogenously expressed with a FLAG tag, allowing

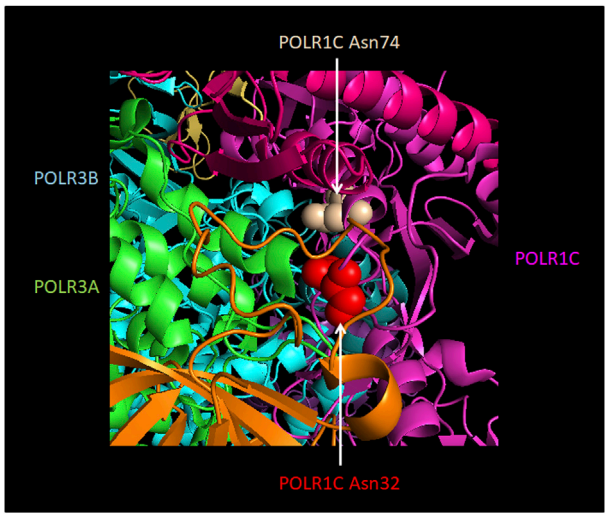


FIGURE 5 | Localization of POLR1C Asn32Ile and Asn74Ser mutations that are associated with POLR3-HLD. Amino acids in POLR1C which cause an assembly defect in Pol III (Thiffault et al., 2015) upon mutation are shown in red (POLR1C Asn32) and in ochre (POLR1C Asn74). They are localized at the interface with Pol III subunits POLR3A (green) and POLR3B (turquoise). The remainder of POLR1C is colored in pink. The Figure was modified from PDB 7AE3 by employing Pymol.

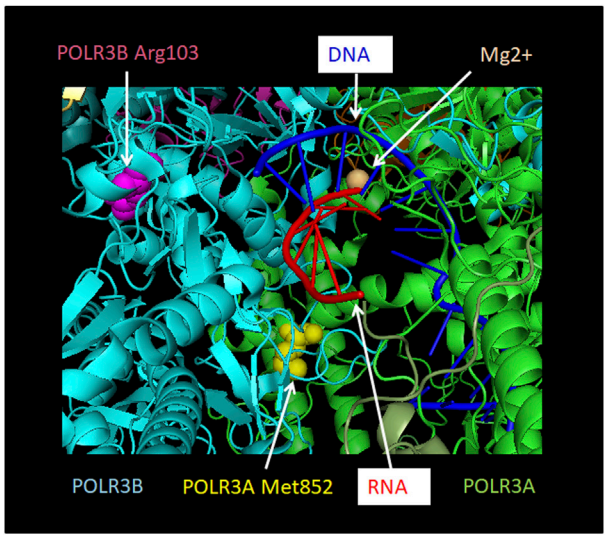


FIGURE 6 | Localization of POLR3A Met852Val and POLR3B Arg103His mutations relative to the active site. POLR3A is shown in green and POLR3B in turquoise. Pol III mutations described in the text that are found close to the active site and which may thus affect catalytic activity are depicted. Methionine 852 of POLR3A as part of the bridge helix is highlighted as a sphere in yellow. Arginine 103 of POLR3B is highlighted as a sphere in magenta. RNA is shown in red and DNA in blue. The Mg2+ ion of the active site is shown as a sphere in light orange. The Figure was modified from PDB 7AE3 by employing Pymol.

subsequent affinity purification and shotgun proteomics to identify interacting partners. The first such study focused on POLR1C and demonstrated that two HLD-causing mutant versions of this subunit (Asn32Ile and Asn74Ser) pulled down significantly lower levels of other Pol III subunits compared to the wild-type subunit, indicating a defect in Pol III complex assembly (Thiffault et al., 2015). This was supported by immunofluorescence data showing that while wild-type POLR1C was predominantly present in the nucleus, mutated POLR1C variants tended to accumulate in the cytoplasm, where Pol III biogenesis takes place. Consistent with these results, mapping of these residues onto the human Pol III structures suggested a function in mediating interactions with POLR3A and POLR3B (Ramsay et al., 2020; **Figure 5**) and postulated that they would impair complex assembly (Girbig et al., 2021). Only one POLR3-HLD-causing *POLR3B* mutation (Arg103His; **Figure 6**) was assayed in a similar manner and was also found to severely impair Pol III complex assembly (Choquet et al., 2019b), while it was predicted to disrupt the core of the subunit in structural studies (Girbig et al., 2021). In contrast, two *POLR3A* mutations, Gly672Glu and Met852Val (**Figure 6**), had no impact on Pol III biogenesis using the same assay (Choquet et al., 2017; 2019a), although they were predicted to impact assembly with POLR2H and to destabilize the POLR3A/POLR3B interface, respectively, in one Pol III structural study (Ramsay et al., 2020), while they were classified as disrupting the core of the subunit and impacting functional elements, respectively, in the second structural study (Girbig et al., 2021). Indeed, POLR3A Met852Val is localized in the vicinity of the bridge helix (**Figure 6**), so it could impair interaction with DNA or transcription itself rather than enzyme assembly (Bernard et al., 2011). Moreover, it is worth noting that POLR3A Gly672Glu can cause a relatively mild phenotype in human individuals (Bernard et al., 2011; Wolf et al., 2014a) and does not lead to neurological abnormalities in mice (Choquet et al., 2017) (see below), which may be due in part to the correct biogenesis and nuclear import of Pol III when this mutation is present.

Using the same experimental system, recently described *POLR3B de novo* heterozygous mutations, which cause a distinct phenotype and are thought to act through a dominant negative mechanism, were found to disrupt the interaction of POLR3B with only one or two Pol III subunits instead of causing an assembly defect of the entire complex (Djordjevic et al., 2021), as was seen for the *POLR1C* and *POLR3B* mutations above. Mapping of these mutations onto the yeast Pol III structure suggests that they are involved in DNA melting or transcription itself (Djordjevic et al., 2021). This indicates that the structural and mechanistic impact of various Pol III pathogenic variants may underlie some of the phenotypic differences observed in patients.

Impact of Pol III Subunit Mutations on the Pol III Transcriptome

Decreased protein abundance, defective Pol III biogenesis and nuclear import or impaired interaction with DNA are all hypothesized to lead to the common outcome of reduced Pol III transcriptional output (**Figure 4**), resulting in some shared clinical symptoms despite differences in the structural and mechanistic consequences of the mutations. Nonetheless, Pol

III transcript deficiencies may be different across cell types or as a consequence of different mutations, thus underlying some of the observed phenotypic heterogeneity. Indeed, expression profiling of the Pol III transcriptome in patient cells or disease models by several groups has revealed a complex picture.

As quantification of most Pol III transcripts is challenging due to their small size, post-transcriptional modifications and repetitive nature, measurement of Pol III occupancy on DNA has often been used as a proxy for Pol III transcription levels (Kutter et al., 2011; Canella et al., 2012). ChIP-seq of FLAG-tagged mutated versions of POLR1C showed a global decreased occupancy at all types of Pol III target promoters, consistent with the low nuclear levels of these POLR1C variants (Thiffault et al., 2015). In contrast, no significant differences in Pol III occupancy were observed by ChIP-qPCR of three Pol III-transcribed loci with exogenous POLR3A-Gly672Glu (Choquet et al., 2017) or by ChIP-seq in cell lines carrying an endogenous *POLR3A* Met852Val mutation (Choquet et al., 2019a), suggesting that these POLR3-HLD-causing mutations may directly impact transcription itself rather than Pol III binding to DNA.

While several recent studies have reported decreased levels of some Pol III transcripts as a result of disease-causing mutations, the identity of these transcripts varies from one study to another (Azmanov et al., 2016; Dorboz et al., 2018; Choquet et al., 2019a). Azmanov et al. (2016) were the first to perform a transcriptome-wide characterization of blood cells from patients with the mild striatal variant of POLR3-related disorders and a specific homozygous splice site mutation (c.1771-6C>G) in *POLR3A*. They observed a global but mild decrease in mature tRNA levels, with only seven tRNAs reaching statistical significance, a reduction in 7SL RNA levels and an increase in the levels of 5S rRNA, RNase P RNA (H1), 7SK RNA and RNase MRP RNA (Azmanov et al., 2016). In a second study, targeted analysis of fibroblasts from two HLD patients carrying *POLR3K* mutations found decreased levels of initiator $tRNA^{Met}$ but no change for three other tRNAs, as well as reduced expression of 7SK RNA and a more severe decrease of 7SL and 5S rRNA levels (Dorboz et al., 2018). Third, CRISPR-Cas9 was used to introduce an endogenous HLD *POLR3A* Met852Val mutation in HEK293 cells in compound heterozygosity with a null allele. Transcriptome-wide analysis uncovered a global decrease in precursor tRNA levels, 7SL RNA and the primate-specific neural BC200 RNA, while other transcripts were not affected (Choquet et al., 2019a). BC200 RNA was also downregulated in the oligodendroglial cell line MO3.13 edited with the same genotype and in two small cohorts of HLD patient-derived fibroblasts carrying *POLR3A* mutations (individual mutations listed in Table S5 of Choquet et al., 2019a). The levels of this RNA were not assessed in the two aforementioned studies (Azmanov et al., 2016; Dorboz et al., 2018). Lastly, qRT-PCR analysis of fibroblasts from a WRS patient showed an increase of tRNA-Leu-CAA, a decrease in 7SK RNA and a virtual absence of 5S rRNA (Báez-Becerra et al., 2020).

These four datasets present important differences in terms of pathogenic variants, nature of the variants, associated phenotypes, cell types assayed and experimental approaches, thus it is not surprising that the affected transcripts vary. Nonetheless, some common trends are starting to emerge. Among disorders with predominant CNS manifestations (mild striatal phenotype or HLD) (Azmanov et al., 2016; Dorboz et al., 2018; Choquet et al., 2019a), 7SL RNA stands out as possibly the most commonly affected transcript. However, an earlier report observed that this transcript had comparable levels in fibroblasts from one POLR3-HLD patient and a healthy control (Shimojima et al., 2014). A subset of tRNAs were also downregulated in each of the three studies, including a common decrease in the initiator $tRNA^{Met}$ levels, which reached statistical significance in two out of three datasets. In the study using POLR3A-edited cell lines and two small patient cohorts, BC200 RNA also emerged as a downregulated Pol III transcript.

7SL and BC200 are both transcribed through a hybrid Pol III promoter (**Figure 2D**), while tRNA genes use a standard type 2 promoter (**Figure 2B**) (Choquet et al., 2019a). In contrast, the two downregulated transcripts in the WRS patient, 5S rRNA and 7SK RNA, are transcribed through type 1 and 3 promoters, respectively (Báez-Becerra et al., 2020; **Figures 2A,C**). It is tempting to hypothesize that alterations in the levels of distinct Pol III transcripts or promoter types may be responsible for different phenotypes. However, in fibroblasts from two patients with *POLR3K* HLD-causing mutations (Dorboz et al., 2018), 5S rRNA levels were decreased in both patients and 7SK levels were diminished in one patient, suggesting a more complex picture. Indeed, mutations in *BRF1*, encoding a subunit of the transcription factor TFIIIB-β specific to Pol III type 1 and type 2, as well as some hybrid promoters, cause a cerebellar-facial-dental syndrome. Analysis of the corresponding mutations in yeast showed impaired Pol III transcription of a tRNA gene *in vitro* (Borck et al., 2015). This disorder overlaps phenotypically with POLR3-related disorders but does not include hypomyelination (Borck et al., 2015; Jee et al., 2017). Together with alterations of tRNAs and 7SL RNA in the mild striatal variant (Azmanov et al., 2016), this argues against the hypothesis that perturbation of the transcription of type 2 or hybrid Pol III target genes specifically leads to myelination defects. Furthermore, as type 2 target genes, especially tRNA genes, far outnumber those with type 1 or 3 target genes, it is not unexpected that the most affected genes in these studies would belong to the former group.

Nonetheless, these gene expression studies emphasize that Pol III transcript levels are remarkably resistant to genetic perturbations in the enzyme, since only a proportion of Pol III transcripts are affected, while many show no change. Importantly, the majority of these datasets were obtained from cell types that are not affected in POLR3-related disorders. Pol III mutations may have a much stronger impact on the transcriptome of affected cell types. Consistent with this idea, the c.1909+22G>A mutation that is common in POLR3-related spastic ataxia results in an aberrant POLR3A splice isoform that is present at higher levels in neuroepithelial cells compared to induced pluripotent stem cells (iPSCs) (Minnerop et al., 2017). Although the Pol III transcriptome was not profiled in these cells, it would be interesting to determine if a higher ratio of aberrantly spliced to wild-type isoform results in stronger alterations of Pol III transcript levels.

The top down-regulated transcripts, 7SL RNA, tRNAs and BC200 RNA, are all involved in mRNA translation and protein homeostasis (Dieci et al., 2007). Quantitative proteomics in POLR3A-edited MO3.13 cells uncovered only a small number of deregulated proteins compared to normal cells (Choquet et al., 2019a). However, since this oligodendroglial cell line was established from a tumor (McLaurin et al., 1995), similar experiments in oligodendrocyte precursor cells (OPCs) derived from human iPSCs or in mouse OPCs, along with ribosome profiling or analysis of nascent proteins, would allow to better determine how translation is impacted upon *POLR3A* mutations in oligodendrocytes. Nevertheless, these cells showed decreased expression of Myelin Basic Protein (MBP) mRNA upon differentiation into more mature oligodendroglial cells, indicating that the mild Pol III transcriptome alterations may be sufficient to alter oligodendrocyte differentiation and/or MBP expression (Choquet et al., 2019a). The observation of nucleolar disruption, activation of p53 and premature senescence in WRS fibroblasts (Báez-Becerra et al., 2020) suggests an alternative mechanism for the pathophysiology of this progeroid syndrome that could be associated with the near absence of rRNAs.

Future research will require larger cohorts from each disease entity within POLR3-related disorders to determine which specific Pol III transcripts are affected and to pinpoint phenotype- and cell-type specific disease mechanisms. Moreover, in order to better understand the pathophysiology of POLR3-related disorders in the relevant cell types, animal models of the diseases are required.

Development of Animal Models of POLR3-Related Disorders

Initial efforts to generate an animal model of POLR3-HLD were not successful. Homozygous knockout of *Polr3a* in mice is embryonic lethal (Choquet et al., 2017), but whole body knock-in (KI) of the French-Canadian founder mutation *Polr3a* c.2015G>A (p.Gly672Glu) did not lead to any neurological or developmental abnormalities in homozygous animals (Choquet et al., 2017). Pol III transcript levels were also normal in the brain of these KI mice. In contrast, homozygosity for the *Polr3b* c.308G>A (p.Arg103His) mutation, which has only been reported in compound heterozygosity with another missense mutation in humans, is embryonic lethal in mice (Choquet et al., 2019b). Interestingly, the drastically different impacts of these two mutations in mice are consistent with the severity of their effect on Pol III biogenesis in human cells (see above) (Choquet et al., 2017; 2019b).

As the *POLR3A* Gly672Glu mutation leads to disease in humans but not in mice, this could suggest that the latter species is less vulnerable to Pol III mutations or that primate-specific transcripts, such as *BC200* RNA, are involved in the pathogenesis of the disease (Choquet et al., 2017). This is also consistent with the observation that mouse models for leukodystrophies tend to have a milder phenotype (Lu et al., 1997; Pujol et al., 2002; Odermatt et al., 2003; Geva et al., 2010; Tress et al., 2011; Fröhlich et al., 2020), which may be due to the lower amount of myelin in mouse brains compared to humans (Fields, 2008; Jakovcevski et al., 2009; Ornelas et al., 2016; Choquet et al., 2017). Thus, two strategies have been attempted to increase the Pol III mutational burden in the hopes that it would lead to a phenotype in mice. First, $Polr3a^{G672/G672E}$ and $Polr3b^{+/R103H}$ were interbred to generate mice with a homozygous mutation in *Polr3a* and a heterozygous mutation in *Polr3b*. However, these mice did not display neurological abnormalities or alterations in Pol III transcript levels (Choquet et al., 2019b).

The second approach was more successful and generated the first mouse model demonstrating hypomyelination as seen in POLR3-HLD, but with a very mild phenotype and absent motor features (pre-print on https://www.biorxiv.org/content/10.1101/2020.12.09.418657v2, currently under peer review at the time this manuscript is written) (Merheb et al., 2021). To achieve this, the Willis laboratory first screened a panel of *POLR3A* HLD mutations by introducing them in the *S. cerevisiae* orthologous gene, *Rpc160*, focusing on a cluster of mutations in the pore region of Pol III, which included Gly672Glu (Moir et al., 2021). Double mutants were also generated by combining Gly686Glu (corresponding to Gly672Glu in humans) with every other mutation in the pore region. Individually, none of these mutations impaired growth, Pol III transcription or mature Pol III transcript levels in *S. cerevisiae*. However, the double mutants displayed phenotypes ranging from wild type to lethal as well as various sensitivity levels to high or cold temperatures. The authors focused on one double mutant carrying the adjacent Tyr685Lys and Gly686Glu mutations, which had an intermediate growth defect and displayed temperature sensitivity. They observed decreased levels of a subset of Pol III transcripts in this mutant (RNAse P RNA [*RPR1*] and small nucleolar RNA 52 [*SNR52*]), while other RNAs, notably those encoding 7SL RNA and 5S rRNAs, were not affected. *In vitro* transcription experiments demonstrated a defect in both factor-independent and factor-dependent transcription for genes representative of the yeast Pol III transcriptome in this double mutant (Moir et al., 2021). Next, the authors generated mice with the corresponding human double allele Trp671Arg/Gly672Glu (Merheb et al., 2021). Since homozygosity for the whole-body KI of this allele was embryonic lethal, a conditional KI mouse was engineered using an Olig2-Cre driver, directing expression of the mutant allele throughout the oligodendrocyte lineage and in a subset of other CNS cells (Merheb et al., 2021). Homozygous conditional KI mice displayed growth defects, neurobehavioral deficits and impaired myelination, myelin integrity and oligodendrogliogenesis. Although the mouse did not display a motor phenotype compatible with POLR3-HLD, it did show mild neurobehavioral features, and myelin defects reminiscent of HLD. Thus, the Trp671Arg/Gly672Glu KI mouse is the first animal model of POLR3-HLD that recapitulates some of the pathological features of the disease. This model can now be used to better understand the relationship between impaired Pol III function and myelin deficits.

A handful of other animal models with mutations in Pol III subunits have also been engineered. A mutation in *POLR3H* was recently found to cause primary ovarian failure (Franca et al.,

2019) and this phenotype was well-recapitulated in whole-body knock-in mice homozygous for the *POLR3H* Asp50Gly mutation (Franca et al., 2019). In zebrafish, a splice site mutation causing the deletion of 41 amino acids in the Polr3b protein led to defects in the development of the intestine, intestinal epithelium and exocrine pancreas (Yee et al., 2007). This mutation impacted the interaction of Polr3b with Polr3k in yeast, and overexpression of *Polr3k* cDNA in zebrafish partially rescued the exocrine pancreas defects. Moreover, conditional deletion of *Polr3b* exon 10 in the mouse intestinal epithelium also led to reduced survival and growth, defective crypt development and increased apoptosis (Kieckhaefer et al., 2016). Interestingly, HLD patients with *POLR3K* mutations present severe digestive dysfunctions that are not typically observed in individuals with mutations in other Pol III subunits (Dorboz et al., 2018). Both patients with biallelic pathogenic variants in *POLR3K* as well as the zebrafish model displayed decreased levels of 7SL RNA, suggesting that the interaction between *POLR3B* and *POLR3K* may be particularly important for transcription of the 7SL RNA gene. Without normal levels of 7SL RNA, its function in protein secretion may be impaired, which could be especially detrimental to normal gut function. The future generation of animal models with a range of mutation types in different Pol III subunits will hopefully help to delineate genotype-phenotype correlations and provide a better understanding of the tissue- and cell type-specific manifestations of POLR3-related disorders. When possible, direct modulation of candidate Pol III transcript (e.g. 7SL RNA) levels in animal models would also help understand the developmental and tissue-specific consequences of their depletion.

As described above, possible reasons for tissue-specific differences may reside in particular dependencies of individual cell types on Pol III transcription products. Below, we will focus on the major Pol III-transcribed RNAs that have shown altered expression in cells carrying mutations in genes associated with POLR3-related disorders. We will describe characteristics of tRNAs, 7SL RNA and BC200 RNA.

FUNCTION OF POL III TRANSCRIPTS AND ROLE IN POLR3-RELATED DISORDERS

Expression and Functions of tRNAs

tRNAs are short (76–90 nucleotides) non-coding RNAs that act as essential adapters during mRNA translation. Each tRNA is loaded at their 3' end with the amino acid corresponding to its anticodon by cytoplasmic aminoacyl tRNA synthetases (aaRS1). tRNAs allow decoding of the genetic code by recognizing cognate codons in translating mRNA and providing the corresponding amino acid for addition to the nascent peptide (reviewed in Lant et al. (2019)). Of the >600 putative tRNA genes in human, approximately 300–400 are expressed in a given human cell (Canella et al., 2010; Oler et al., 2010; Gogakos et al., 2017), resulting in multiple expressed genes with minor sequence differences encoding tRNAs with the same anticodon (isodecoders; Pan, 2018). Sequence changes or imbalanced expression of tRNAs can lead to deregulated translation (reviewed in Lant et al. (2019); Kapur et al. (2020)).

Several studies have shown that pools of expressed tRNA isodecoders vary by cell type and cell state, suggesting that certain isodecoders are more important in specific contexts and that their dysregulation could impair cellular homeostasis. Indeed, distinct pools of tRNAs are expressed between proliferating and differentiating cells (Gingold et al., 2014) and the corresponding anticodons match the codon usage of mRNAs expressed in each state. Thus, a specific pool of tRNAs may be required to match the codon usage of genes important for oligodendrocyte differentiation and/or myelination or neuronal development or function, and reduced levels due to mutations in Pol III subunits or aaRS1 may contribute to the pathogenesis of HLD. Moreover, a recent study optimized next-generation sequencing of mature tRNAs to demonstrate a distinct expression profile of tRNA isodecoders in mouse CNS tissues compared to non-CNS tissues, with several isodecoders varying more than 4-fold, while total isoacceptor pools were relatively stable across these tissues (Pinkard et al., 2020). In an earlier study done by microarray, tRNA levels were found to vary across tissues, with the brain having among the highest levels of nuclear-encoded tRNAs (Dittmar et al., 2006). Together, these data suggest that CNS cell types could be particularly vulnerable to reduced tRNA levels, particularly for certain isodecoders that are more abundant in the CNS.

Consistent with this hypothesis, the Ackerman group identified the first instance of a tissue-specific mammalian tRNA gene, *n-Tr20*, which is exclusively expressed in the mouse CNS (Ishimura et al., 2014). *n-Tr20* encodes a tRNA-Arg-UCU isodecoder and contains a single nucleotide polymorphism (SNP) in the T stem loop in the C57BL/6J strain compared to other mouse strains. This results in accumulation of a precursor form of *n-Tr20* and decreased levels of the mature form and leads to increased ribosome pausing on AGA codons. On its own, the *n-Tr20* polymorphism was found to modulate seizure susceptibility and synaptic transmission (Kapur et al., 2020). Together with loss-of-function mutations in the recently characterized ribosome rescue factor genes *Gtpbp1* and *Gtpbp2*, the *n-Tr20* SNP leads to widespread neurodegeneration (Ishimura et al., 2014; Terrey et al., 2020), suggesting that these factors are essential to resolve ribosome pausing defects induced by decreased tRNA levels. Deletion of *n-Tr20* led to increased pausing at AGA codons genome-wide and reprogramming of the translatome and induced the integrated stress response (ISR) (Kapur et al., 2020), an important component of regulated translation (reviewed in Tahmasebi et al. (2018)). Moreover, a deletion in one of four expressed tRNA-Ile-UAU isodecoders (*n-Ti17*) decreased total tRNA-Ile-UAU levels and similarly increased the ISR in mouse brains, indicating that this is not specific to n-*Tr20* but rather a common response to deficient tRNA levels. Thus, specific tRNA isodecoders play essential roles in maintaining normal translation in mouse brains. Although isodecoders with CNS-specific expression have not yet been identified in humans, these results suggest that deficient expression of any single tRNA important for brain function could lead to translation deregulation. In the context of POLR3-HLD, reduced levels of specific isodecoder(s)

important in certain spatio-temporal contexts could induce ribosome stalling at the corresponding codons and impair translation of proteins important for normal oligodendrocyte and/or neuronal function and underlie disease pathogenesis.

As described above, a subset of tRNAs were found to be downregulated in patient cells or in cell lines carrying POLR3-HLD mutations. The *POLR3A* Met852Val mutation (**Figure 6**), causing POLR3-HLD, significantly reduced pre-tRNA levels, but not those of selected mature tRNAs (Choquet et al., 2019a) in a cellular model of POLR3-HLD, whereas the *POLR3K* Arg41Trp mutation mildly decreased levels of mature $tRNA^{Met}$ but not those of three other mature tRNAs (Dorboz et al., 2018). It should be noted that due to their extensive post-transcriptional modifications, tRNA expression levels are more difficult to determine by RT-qPCR or by RNA-sequencing, the primary methods used in these studies. DM-tRNA-seq (Zheng et al., 2015), ARM-seq (Cozen et al., 2015) and Hydro-tRNA-seq (Arimbasseri et al., 2015) were developed to overcome this obstacle and improved sequencing results. Recently, the mim-tRNA-seq (Behrens et al., 2021) and QuantM-tRNAseq approaches were published (Pinkard et al., 2020), which may help to further improve the quantification of mature tRNA expression levels, thereby allowing to determine whether POLR3-HLD can be the consequence of a modest reduction of pre-tRNA levels due to higher demand for translation in oligodendrocytes, or whether other mechanisms may also account for the development of this disease. Ribosome profiling in relevant cell types could also determine if ribosome stalling occurs at certain codons, as was observed with the *n-Tr20* polymorphism in mice.

Since HLDs are not only caused by mutations in genes encoding Pol III subunits, but also by alterations in several aaRS1 genes (e.g. *DARS1*; *EPRS1*; *RARS1*; Taft et al., 2013; Wolf et al., 2014a; Mendes et al., 2018, 2020), it has been speculated that changes in tRNA abundance (POLR3-HLD) or dysfunctions in the attachment of amino acids to tRNAs (HLD caused by aaRS1 mutations) could represent a unified disease-causing mechanism, in which reduced availability of specific aminoacylated tRNA(s) would lead to altered or insufficient translation by stalling ribosomes on the corresponding codons. A defect in aminoacylation was reported for disease-causing mutations in *KARS1, EPRS1* and *AARS1 in vitro* (Simons et al., 2015; Nakayama et al., 2017; Mendes et al., 2018; Itoh et al., 2019), while the aminoacylation activity of ArgRS1 was impaired upon some *RARS1* mutations but not with the most common mutation (Li G. et al., 2021). However, the potential impact of these mutations on translation was not investigated. Dominant mutations in several genes encoding tRNA aminoacyl synthetases are associated with Charcot-Marie-Tooth (CMT) disease and characterization of the corresponding mutants has demonstrated that aminoacylation activity is frequently not impaired. Instead, the mutations induce an alternative open conformation of the enzyme, which exposes a surface for new protein interactions (He et al., 2011; Blocquel et al., 2017; Bervoets et al., 2019; Blocquel et al., 2019; Sun et al., 2021), indicating a gain-of-function mechanism. In the case of leukodystrophies caused by bi-allelic mutations in genes encoding aaRS1, caused by hypomorphic mutations, further studies are required to determine if there is an underlying mechanism that involves translation deregulation and/or shares features with POLR3-HLD. Alternatively, aaRS1 possess numerous non-canonical functions (Wagasugi and Yokosawa, 2020; Yao and Fox, 2020) that could contribute to disease pathogenesis, although those differ between different aaRS1.

Upstream from their role in translation, misexpression of tRNA genes could affect their transcription, post-transcriptional processing and/or modifications. First, tRNAs are expressed from type 2 gene internal promoters, requiring TFIIIC and TFIIIB-β transcription factor complexes in order to recruit Pol III to the TSS (Dumay-Odelot et al., 2010; **Figure 2B**). High transcriptional efficiency at tRNA genes is at least in part enabled by facilitated recycling (Dieci et al., 2014). It is conceivable that Pol III mutations exert a negative effect on facilitated recycling, which could result in the decreased expression of tRNA genes observed in POLR3-HLD studies (Azmanov et al., 2016; Dorboz et al., 2018; Choquet et al., 2019a).

Second, upon transcription termination, tRNAs undergo extensive post-transcriptional modifications including 1) the removal of the 5′ leader sequence by RNAse P (Jarrous, 2017), 2) processing of the 3′ end by ELAC2, the human orthologue of RNAse Z (Takaku et al., 2003; Siira et al., 2018), 3) the addition of CCA nucleotides to the 3’ terminus of tRNAs by the tRNA nucleotidyl transferase 1 (TRNT1) (Xiong and Steitz, 2006) and 4) removal of possible introns by the tRNA splicing endonuclease (TSEN) complex and CLP1 (Hayne et al., 2020). Subsequently, an average of 13 post-transcriptional modifications is brought upon individual tRNA molecules (reviewed in Pan (2018); tRNA transcription and maturation is summarized in **Figure 7**), many of which are important for normal brain function (reviewed in Ramos and Fu (2019)). Post-transcriptional modifications alter local and overall tRNA folding, affecting their stability (reviewed in Ramos and Fu (2019)). The half-life of precursor (pre)-tRNAs was estimated to be 15 to 30 min whereas it is about 100 h for mature tRNAs (Choe and Taylor, 1972). Most of these processing steps have been associated with neurological disorders: changes in modifications of nucleotides in the anticodon stem loop or at transitions from stem to D-loop or T-loop structures were shown to be related to the development of neurodevelopmental disorders, including intellectual disabilities or amyotrophic lateral sclerosis (Freude et al., 2004; Bento-Abreu et al., 2018; Sharkia et al., 2019). Several enzymes involved in pre-tRNA processing and tRNA post-transcriptional modification are associated with inherited neurodegenerative disorders (reviewed in Schaffer et al. (2019)). Thus, the CNS appears to be particularly vulnerable to any defect in tRNA metabolism, further indicating that reduced tRNA expression in POLR3-HLD is a likely mechanism underlying dysfunction of neurons and/or oligodendrocytes.

Of particular relevance to POLR3-HLD, accumulation of a tRNA processing intermediate retaining the 5’ leader was observed in a *S. cerevisiae* mutant strain carrying two POLR3-HLD mutations at the homologous positions in *Rpc160* (yeast homolog of *POLR3A*) (Moir et al., 2021). Although this was not

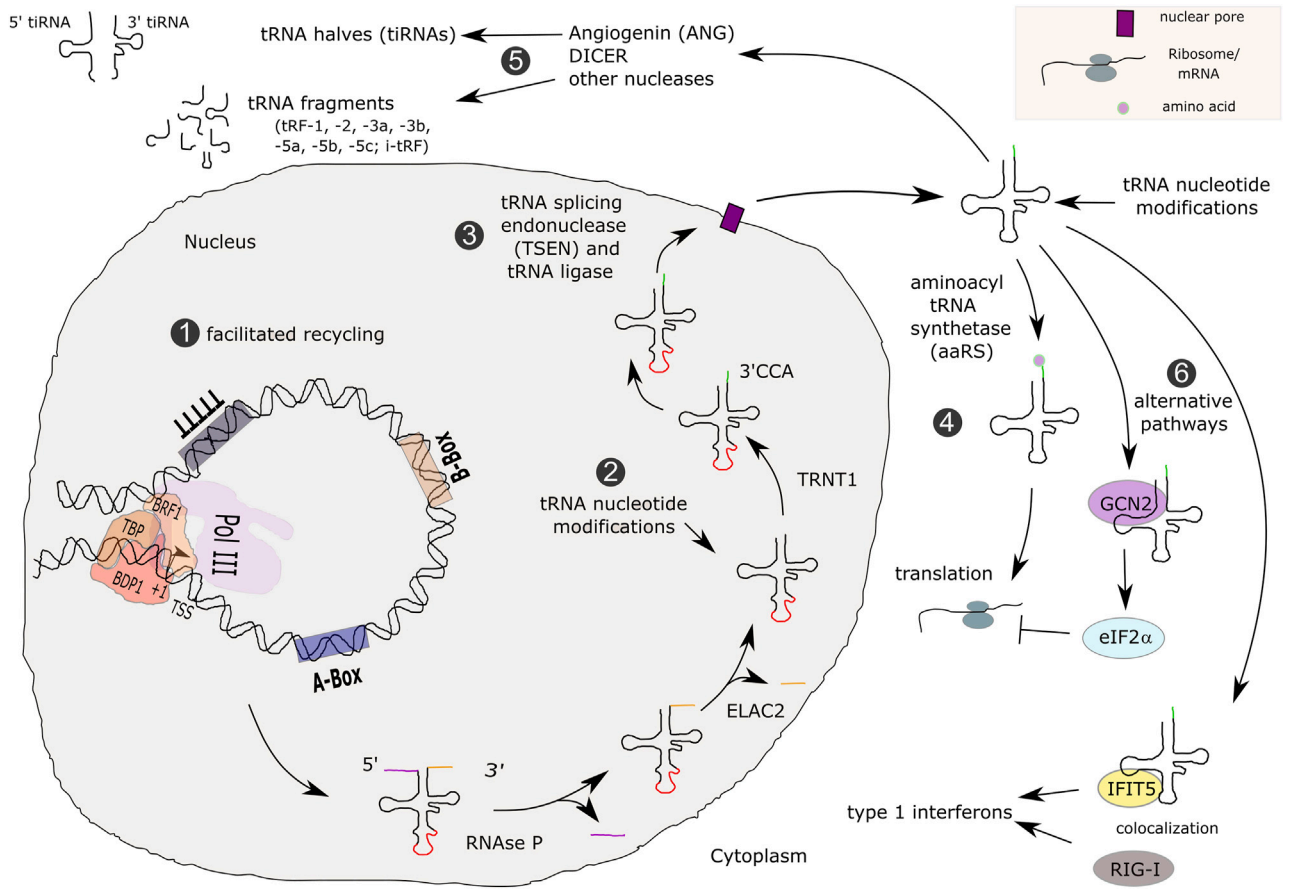


FIGURE 7 | Transcription, maturation and functions of tRNAs. Type 2 promoter tRNA genes are transcribed by Pol III, possibly involving the facilitated recycling pathway, which allows transcription reinitiation in the absence of TFIIIC (1). Primary tRNAs contain 5′ leader and 3′ trailer sequences (colored lines) and a few human tRNAs contain introns (red line). tRNAs are processed by RNAse P, ELAC2 and TRNT1 (2), as well as spliced in the nucleus (3). Mature tRNAs are exported to the cytoplasm and loaded with amino acids for participating in translation (4). Other functions for tRNAs that were described involve enzymatic processing for yielding tRNA fragments (tRFs) or tRNA halves (tiRNAs) (5). In addition, they can enter alternative functional pathways (6).

sufficient to impair mature tRNA levels (Moir et al., 2021), it indicates that tRNA processing could also be affected by POLR3-HLD, which may have particular importance in certain tissues. In *S. pombe*, introduction of two homologous POLR3-HLD mutations led to decreased transcription of three tRNA genes, but also to increased tRNA N2,N2-dimethyl G26 (m^2_2G26) modification efficiency (Arimbasseri et al., 2015). Global repression of Pol III transcription through rapamycin treatment yielded a similar effect, both in *S. pombe* and in human HEK293 cells, suggesting that this response is conserved. Thus, reduced transcription of tRNA genes may lead to increased modification efficiency of tRNAs, which could be detrimental for brain function. Furthermore, stress-correlated modification of tRNAs was reported to occur (Gu et al., 2014), which might be affected by lower tRNA transcription rates and subsequently affect aminoacylation and translation.

Aspects of mRNA and tRNA modification have also been shown to be coordinated (Ontiveros et al., 2020; Levi and Arava, 2021). For example, the enzyme TRMT10A, which is known for depositing m^1G on tRNAs, also influences m^6A deposition on mRNAs by interacting with FTO (Ontiveros et al., 2020). Depletion of TRMT10A decreased m^1G levels on tRNAs but increased m^6A levels on mRNA. Some pseudouridine synthases were also shown to modify both tRNAs and mRNAs (Borchardt et al., 2020). This opens the door to the possibility that alterations in transcription of tRNAs in POLR3-HLD could influence post-transcriptional modification of mRNAs, with reduced levels of tRNAs liberating more enzymes for acting on mRNAs, thus modulating processing, stability and/or translation of these mRNAs.

Finally, tRNAs have non-canonical functions outside of translation that could be misregulated in POLR3-HLD. tRNAs were shown to bind to and modify the activities of proteins that are not directly involved in translation control. For instance, it was demonstrated that tRNA-GCN2 kinase interactions regulated the phosphorylation of eIF2α, thereby reprogramming translation towards general repression whilst activating translation of selected mRNAs (Dong et al., 2000; Castilho et al., 2014; **Figure 7**). In addition, the interferon-induced tetratricopeptide repeat 5 (IFIT5) protein binds

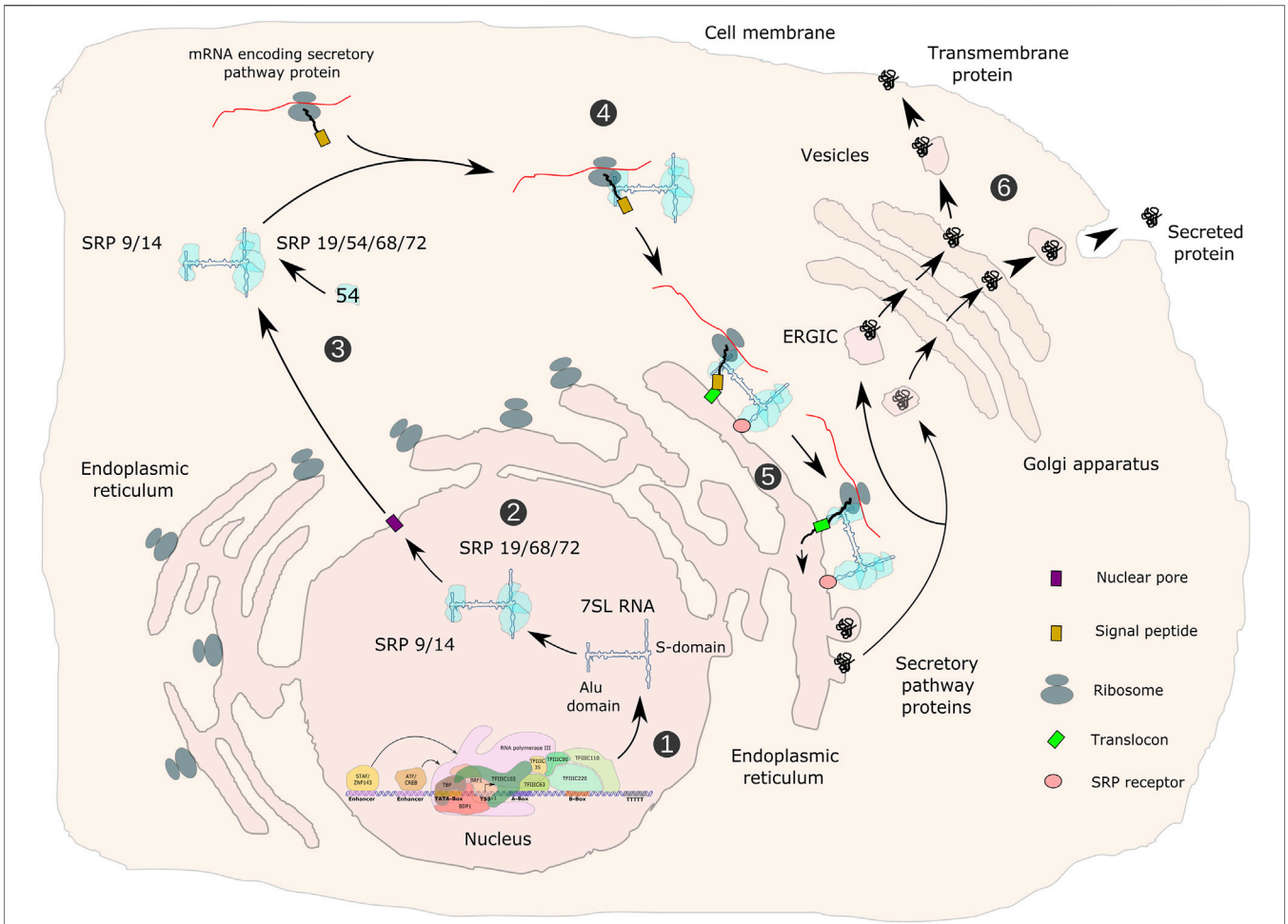


FIGURE 8 | Transcription of the 7SL RNA gene, assembly of the signal recognition particle and functions in the secretory pathway. 7SL RNA gene transcription is directed by gene-internal A- and B-boxes as well as by an ATF/CREB-response element and a distal sequence element (DSE) upstream of the TSS. Transcription is initiated by TFIIIC- and TFIIIB-β-dependent recruitment of Pol III (1). The SRP 9/14/19/68/72 pre-signal recognition particle (pre-SRP) is assembled in the nucleus (2), exported to the cytoplasm and completed with SRP54 (3). SRP recognizes hydrophobic signal peptides in nascent polypeptide chains of secretory pathway proteins (4), arrests translation, translocates arrested ribosomes to the endoplasmic reticulum (ER), where translation resumes and proteins are translocated either into the lumen or into the membrane of the ER (5). Proteins of the secretory pathway are either incorporated into membranes or exported from cells (6). ERGIC refers to the ER-Golgi-intermediate compartment.

tRNAs, thereby modulating double-stranded DNA sensing receptor RIG-I and being involved in regulating type I interferon response (Katibah et al., 2013; **Figure 7**). Such mechanisms may also be sensitive to changes in tRNA transcription and could therefore be of importance for the development of POLR3-HLD.

7SL RNA – Transcription, Structure and Functions

7SL RNA is a major component of the signal recognition particle (SRP). In addition to 7SL RNA, the SRP is composed of the SRP9, 14, 54, 68 and 72 proteins. It is responsible for co-translational targeting of nascent secretory and transmembrane peptides to the endoplasmic reticulum (ER) through interaction with its SRP receptor (Lakkaraju et al., 2008; Akopian et al., 2013). Pol III transcription of the 7SL gene is directed by promoter elements that are located within the transcribed region (A- and B-boxes), as well as by TATA-like, ATF/CRE and STAF-binding sequences upstream of the TSS (Ullu and Weiner, 1985; Bredow et al., 1990a; Englert et al., 2004; Dumay-Odelot et al., 2014) (hybrid promoter; **Figure 2D**).

Genomic occupancy of POLR3G (RPC32α; Haurie et al., 2010), BRF1/TFIIIB-β (Teichmann and Seifart, 1995; Wang and Roeder, 1995; Mital et al., 1996), GTF3C4/TFIIIC63 (Hsieh et al., 1999) and BDP1 (Schramm et al., 2000; Teichmann et al., 2000) was analyzed, suggesting that transcription of the 7SL gene is carried out *in vivo* by TFIIIC, TFIIIB-β and Pol III (Canella et al., 2010; Oler et al., 2010). *In vitro*, 7SL transcription was shown to be stimulated by ATF (Bredow et al., 1990b). *Ex vivo*, 7SL was identified as the most abundant non-rRNA transcript in two cell lines (Boivin et al., 2018).

The 300 nt human 7SL RNA contains two domains that were identified by micrococcal nuclease digestion. Base pairing of the 5′ and 3' parts of 7SL RNA forms the Alu domain, whereas the central part folds into the S domain (Gundelfinger et al., 1983; Zwieb, 1985; **Figure 8**). The Alu domain represents the binding site for SRP14/SRP9, whilst the S-domain is recognized by SRP19, SRP54 and the SRP68/SRP72 heterodimer, altogether composing the SRP (Gundelfinger et al., 1983). A pre-SRP, consisting of 7SL RNA and SRP proteins 9, 14, 19, 68 and 72 is assembled in the nucleus. Upon export to the cytoplasm, it is completed by the addition of SRP54 (Massenet, 2019; **Figure 8**). SRP54 within the S-domain-associated protein complex recognizes a N-terminal hydrophobic signal sequence in nascent peptide chains of proteins. The SRP14/SRP9-containing Alu domain in turn interacts with translation elongation factor binding sites within cytoplasmic ribosomes, thereby inducing an elongation arrest (Halic et al., 2004; Voorhees and Hegde, 2015). Proteins containing a signal peptide are either secreted or are an integral part of the cell membrane. Stalled ribosomes are then targeted to the Sec61 core component of the translocon within the ER via a GTP- and SRP54-dependent process, resulting in proteins being synthesized and translocated either into the lumen or into the membrane of the ER and secreted or delivered to the cellular membrane (**Figure 8**). GTP hydrolysis triggers the release of SRP from the Sec61 translocon, allowing translation to resume (Fulga et al., 2001; Pool, 2005). Depletion of SRP14, SRP54 or SRP72 in HEK293 or HeLa cells leads to decreased 7SL RNA levels, inefficient ER targeting and impaired post-ER membrane trafficking (Lakkaraju et al., 2007). Thus, decreased 7SL RNA levels in POLR3-HLD may impair translocation of secreted proteins to the ER, which could contribute to the pathophysiology of POLR3-related disorders in several different cell types. 7SL RNA is at its highest level of expression in the hypothalamus compared to other non-CNS tissues (Castle et al., 2010). The expression of 7SL RNA was demonstrated to be positively regulated during differentiation of mouse embryonic stem cells into a differentiated heterogeneous population of neurons and glial cells (Skreka et al., 2012), suggesting that it may be of particular importance in these cell types, both of which are affected in POLR3-related disorders.

In oligodendrocytes, proteolipid protein (PLP) is a major myelin protein that is targeted to the ER and follows the secretory pathway to reach the site of myelination (Woodward, 2008). Reduced SRP function could potentially impair PLP trafficking and contribute to POLR3-HLD pathogenesis (**Figure 4**). In depth analysis of such mechanism may require the establishment of appropriate experimental systems, including the study of primary oligodendrocytes derived from iPSCs and/or in co-culture with other CNS cell types. For example, iPSCs from individuals carrying mutations in the gene encoding PLP, which cause Pelizaeus-Merzbacher HLD, were differentiated into oligodendrocytes to show mislocalization of PLP to the ER and to identify modulators of ER stress (Numasawa-Kuroiwa et al., 2014; Nevin et al., 2017). Similar experiments could be undertaken with iPSCs from POLR3-HLD patients to determine if PLP or other myelin proteins are mislocalized. In addition, three-dimensional growth of human iPSC-derived oligodendrocytes in organoid cultures (Marton et al., 2019) may allow reproducing a cellular environment that better reflects the *in vivo* situation. Cells grown under these conditions may show higher dependency on optimal Pol III transcription and may therefore be more vulnerable to protein mislocalization in the case of reduced 7SL RNA expression.

In neurons, protein trafficking through the ER is crucial for both dendritic and axonal function, including for synaptic plasticity and neurotransmitter trafficking (reviewed in Ramirez and Couve (2011), Kennedy and Hanus (2019)). In mouse motor neurons, 7SL RNA was found to be more abundant in axons than in the somatodendritic compartment, indicating an important role in axons (Briese et al., 2016). Moreover, the importance of ER function in neurons is exemplified by the fact that around half of hereditary spastic paraplegias (HSPs) are caused by mutations in genes encoding ER-shaping proteins (Ozturk et al., 2020). At least one HSP-associated mutation causes a kinetic delay in ER protein secretion (Slosarek et al., 2018). Spasticity is observed in patients with POLR3-related spastic ataxia, suggesting that pyramidal neurons, the primary affected cell type in HSPs, are also involved in POLR3-related disorders. A possible hypothesis is that reduced 7SL RNA levels could affect ER targeting in these neurons or others, leading to the observed neurodegeneration.

Another connection between 7SL RNA and neurodegeneration is the fact that *in vivo* assembly of the SRP complex depends on the survival of motoneuron (SMN) complex (Piazzon et al., 2013; Massenet, 2019), which is responsible for assembly of ribonucleoprotein (RNP) complexes, most notably the spliceosomal snRNPs (Li et al., 2014), and is composed by the SMN protein and Gemin2-8 proteins. Mutations in *SMN1*, encoding the SMN protein, cause spinal muscular atrophy (SMA) (Lefebvre et al., 1995; Wirth, 2000; Wirth et al., 2020). Reduced levels of 7SL RNA were detected in the spinal cord but not in the brain and heart of an SMA mouse model (Piazzon et al., 2013), suggesting that 7SL RNA levels are regulated by SMN function in a cell-type specific manner. Recently, loss-of-function mutations in *GEMIN5* were found to cause a neurodevelopmental syndrome that includes cerebellar ataxia (Kour et al., 2021). *GEMIN5* mutations decreased levels of snRNP complexes *in vivo* and disrupted SMN complex assembly *in vitro*. Thus, dysfunction of the SMN complex can specifically affect cells of the CNS. Since 7SL RNA was found to compete with U1 and U2 snRNPs for binding to SMN complexes *in vitro* (Piazzon et al., 2013), it is possible that similar competition occurs *in vivo* when SMN function is compromised, resulting in impaired protein secretion that could contribute to the disease phenotypes. Future investigation of protein secretion in SMA and other SMN-related disorders will help clarify the potential role of 7SL RNA in neuronal dysfunction.

Reduced 7SL RNA levels may also contribute to other disease phenotypes, such as the digestive dysfunction observed in patients with *POLR3K* mutations (Dorboz et al., 2018) and in *polr3b* mutant zebrafish (Yee et al., 2007). Indeed, intestinal epithelial cells are highly secretory and are sensitive to ER stress and unfolded protein response (reviewed in Coleman and Haller 2019). Future experiments are required to

determine the role of the POLR3K-POLR3B interface and of 7SL RNA in intestinal cell homeostasis.

BC200 - Transcription, Structure and Functions

BC200 (BCYRN1 – brain cytoplasmic RNA 1) is a monomeric Alu RNA that is predominantly expressed in the brain of primates. It was discovered by northern blot analyses (Watson and Sutcliffe, 1987) employing an identifier (ID) sequence as a probe. This ID element, the rodent-specific BC1 RNA of 154 nt, was previously shown to be specifically expressed in rat brain (Sutcliffe et al., 1982, 1984). BC1 RNA is derived from a $tRNA^{Ala}$ retrotransposition event (Daniels and Deininger, 1985), whereas the 200 nt primate BC200 RNA is a 7SL RNA-derived exapted monomeric Alu element (Watson and Sutcliffe, 1987; Brosius, 1999). As a consequence, BC1 and BC200 RNA share little sequence homology and are thought to be functional analogs rather than true homologs. The fact that two independent retrotransposition events resulted in the generation of distinct but related brain-specific RNAs indicated that these RNAs fulfil important roles within the brain. However, KO of the BC1 gene in mice resulted in healthy animals which showed at the first sight only discrete neurological abnormalities such as neuronal hyperexcitability (Skryabin et al., 2003; Zhong et al., 2009). Closer inspection in later behavioral trials demonstrated that BC1 KO mice had impaired cognitive abilities (Chung et al., 2017; Iacoangeli et al., 2017). In addition to the neuron-specific expression in healthy animals, BC1 and BC200 expression was also detected in tumor samples and tumor cell lines (for review see Samson et al. (2018)). A recent study using qRT-PCR found that BC200 RNA levels were comparable in primary cell lines and tumor cell lines from the same tissue. Expression of BC200 RNA in three primary or non-tumorigenic cell lines was also surprisingly similar to GAPDH mRNA levels (Booy et al., 2017).

The BC200 gene contains classical type 2 intragenic promoter elements (A-box and B-box) and in addition a TATA-like sequence upstream of the TSS (hybrid promoter). Sequences up to 100 nucleotides upstream of the TSS were suggested to be important for BC200 transcription efficiency in transient transfection experiments. Stepwise deletion of these sequences led to a gradual decline in transcription rate without changing the ability of Pol III to correctly recognize the TSS. In addition, mutation of gene internal A- or B-boxes abolished transcription of the BC200 gene (Kim et al., 2017). These results indicate that gene internal promoter elements are the crucial determinants of BC200 expression and TSS selection. The importance of gene internal control elements and of sequences upstream of the TSS were also demonstrated by *in vitro* transcription of rodent BC1 gene (Martignetti and Brosius, 1995). In addition, BRF1, TFIIIC and Pol III subunits, but not BRF2 were detected by ChIP-seq at the BC1 gene promoter in mice, underscoring that it is regulated by gene internal promoter elements and stimulated by regulatory elements upstream of the TSS (Carrière et al., 2012).

As the only Pol III transcript with brain-specific expression, BC200 RNA represents an attractive candidate for a role in POLR3-related disorders. In the tumor cell line MO3.13, which has characteristics of oligodendrocyte progenitor cells (OPC) (McLaurin et al., 1995), KO of BC200 led to significant gene expression changes (Choquet et al., 2019a), suggesting a function for BC200 RNA in OPCs, although these findings must be confirmed in primary cells to draw definite conclusions. It should be noted that early *in situ* hybridization experiments did not detect BC200 RNA expression in adult brain white matter (Tiedge et al., 1993), but this does not exclude the possibility that BC200 RNA is expressed in OPCs or in oligodendrocytes earlier in development, such as when myelination occurs, especially given that recent studies have detected BC200 RNA expression in non-neuronal primary cell lines (Booy et al., 2017; Choquet et al., 2019a), albeit at lower levels than in the brain. Additional functional studies will be required to determine if BC200 RNA is important for other cell types, and expression profiling in different CNS cell types from fetal and adult tissues will help establish how BC200 RNA is modulated spatially and temporally.

Alternatively or in addition, impaired expression of BC200 RNA may contribute to some of the neuronal phenotypes (e.g. cerebellar, striatal) observed in POLR3-related disorders. Functional studies on BC200 RNA have mostly been performed *in vitro*, in tumor cell lines or by analogy with BC1 RNA. Nonetheless, the many identified interacting partners and potential functions for this non-coding RNA provide hypotheses as to how it may contribute to the pathogenesis of POLR3-related disorders.

According to a structural model of BC200 RNA, the first 120 nucleotides at the 5′-end, together with nucleotides 175–200 of the C-terminal unique region, fold into an Alu-domain. The 5′ part of the Alu-domain and the unique C-rich domain at the 3′-end of BC200 RNA are separated by a loop-forming A-rich domain. The Alu-domain of BC200 is highly similar to that of 7SL RNA (Sosińska-Zawierucha et al., 2018; **Figure 9**). Consequently, the 7SL-interacting SRP9/14 heterodimer was also shown to interact with BC200 RNA (Bovia et al., 1997; Kremerskothen et al., 1998) and possible consequences on translation inhibition were discussed. Other proteins interacting with BC1 and/or BC200 RNAs were described, including Pur α (Kobayashi et al., 2000; Johnson et al., 2006), Fragile X Mental Retardation Protein (FMRP; Zalfa et al., 2003), Poly(A)-binding Protein (PABP; Muddashetty et al., 2002), Synaptotagmin-binding cytoplasmic RNA interacting Protein (SYNCRIP/hnRNP Q1; Duning et al., 2008), RNA helicase associated with AU-rich element (RHAU/DHX36; Booy et al., 2016), eukaryotic translation initiation factor 4A (eIF4A; Lin et al., 2008) and heterogeneous nuclear ribonucleoproteins E1 and E2 (hnRNP E1 and E2; Jang et al., 2017).

These proteins were proposed to influence the stability and/or the export of BC200 from the nucleus to the cytoplasm (SRP9/14), mRNA transport in neuronal dendrites (Pur α, FMRP, SYNCRIP) and/or interfere with BC1/BC200 effects on translation. Whilst eIF4A and PABP are targets of translational inhibition by BC200/BC1 RNAs (Lin et al., 2008), hnRNP E1 and E2 were proposed to counteract BC200-mediated translation inhibition (Jang et al., 2017; **Figure 9**). The helicase RHAU/DHX36 was shown to mediate the binding of unwound G-quadruplexes to BC200 (Booy et al., 2016), thereby possibly indirectly intervening with translation. The question of whether BC1/BC200 stimulates or inhibits translation

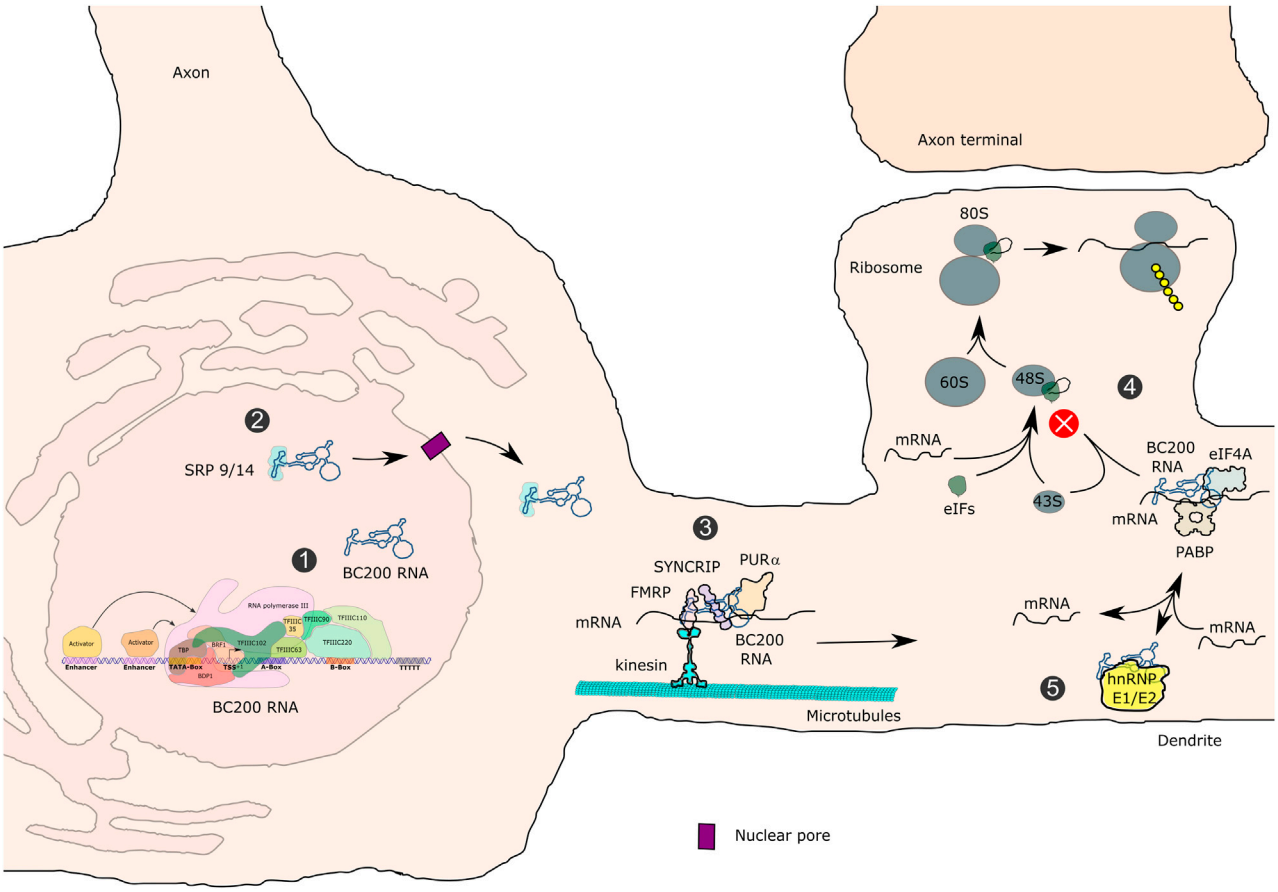


FIGURE 9 | Transcription of the BC200 RNA and its functions in mRNA transport and regulation of translation. BC200 RNA gene transcription depends on TFIIIC bound to gene-internal A- and B-boxes, which allows the subsequent recruitment of TFIIIB-β and Pol III. Sequences upstream of the TSS including a TATA-like box at -30 regulate transcription activity (1). BC200 RNA assembles with SRP 9/14 in the nucleus (2) and is exported to the cytoplasm where it was shown to interact with a variety of proteins. BC200 was shown to interact with FMRP, SYNCRIP and Pur-alpha, suggesting a role in mRNA transport similar to what was demonstrated for BC1 in complex with these proteins (3). In association with eukaryotic initiation factor 4A (eIF4A) and poly A binding protein (PABP), BC200 RNA prevents the assembly of the 48S ribosomal subunit (4). This repression of translation can be counteracted by hnRNP E1/E2 binding to BC200 RNA *in vitro* (5).

has not been definitively solved. Most reports indicate that these RNAs contribute to repression of translation in postsynaptic dendrites through interactions with eukaryotic initiation factors (eIFs) 4A and 4B (reviewed in Iacoangeli and Tiedge (2013)), with most of these experiments performed *in vitro* or only for BC1 RNA.

Recent studies performed in human cell lines indicate that the role of BC200 RNA goes beyond what was learned from *in vitro* experiments or by analogy with BC1 RNA. Indeed, a recent report shows that depletion of BC200 in MCF-7 breast cancer cells resulted in the reduction of translation (Booy et al., 2020). Thus, the impact of BC1/BC200 RNAs on translation may be context-dependent and vary in neurons where these RNAs act as translational repressors inhibiting eIF4A helicase activity and in tumor cells where translation is executed at sites within the cytoplasm that differ largely from specialized compartments such as synapses. Moreover, the BC200 interactome was analyzed in three transformed cell lines (MCF-7; MDA-MB231; HEK293T) by exogenous expression of a 3′-end labeled BC200 RNA (Booy et al., 2018). This confirmed previous interactors of BC200 RNA (e.g. SRP9/14, PABPC1, DDX36) but also identified new interactors (e.g. TRIM24, HNRNPK, CSDE1), several of which are involved in regulating RNA stability. Some binding partners may influence the stability of BC200 RNA itself, while others may be functional partners. A reciprocal interaction was shown for at least one binding partner, in which CSDE1 regulates BC200 RNA levels, while BC200 RNA influences CSDE1 post-transcriptional regulation, likely by affecting translation rate or protein stability. In addition, BC200 RNA was also found to regulate alternative splicing (Singh et al., 2016) and mRNA stability (Shin et al., 2017a; Shin et al., 2017b) of specific transcripts in cell lines, suggesting that its role does go beyond the analogies drawn from studying BC1 RNA and that BC200 RNA may modulate RNA processing or stability of the same or other target mRNAs in CNS cells. Establishing whether BC200 RNA accomplishes these functions or interacts with the same protein partners in normal cells such as neurons in addition to tumor cells represents an important avenue of future research. Furthermore, characterizing the mRNA interactome of BC200 RNA in normal

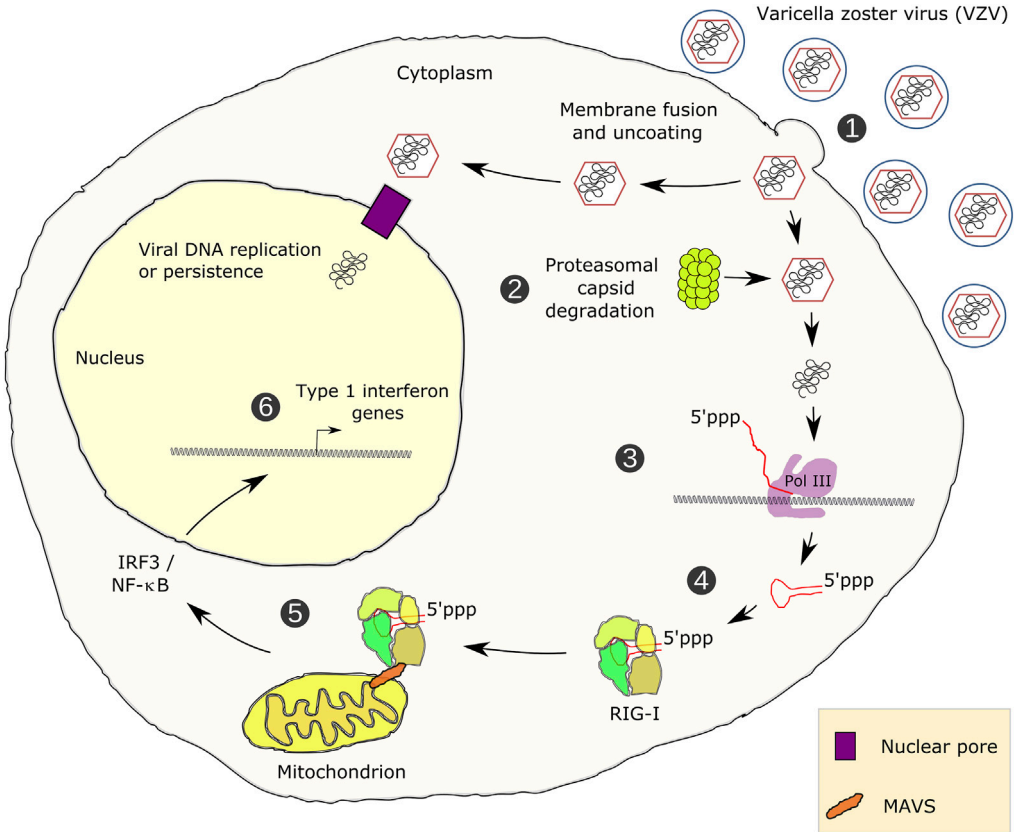


FIGURE 10 | Cytoplasmic transcription by RNA polymerase III in innate immune defense. Upon infection (1), the varicella zoster virus (VZV) DNA is found in the nucleus and the cytoplasm (Mahalingam et al., 1999; Zerboni et al., 2014). Although not specifically reported for VZV, cytoplasmic liberation of DNA from capsids may involve proteasomal degradation (2), as was described for herpes simplex virus 1 (HSV-1; Horan et al., 2013). Cytoplasmic AT-rich VZV DNA is transcribed by Pol III (3), the 5′ tri-phosphorylated RNA folds into stem-loop structures, interacts with the retinoic acid-inducible gene I (RIG-I) (4), which, via mitochondrial antiviral signaling (MAVS) protein and NF-kB/IRF3 transcriptional activators (5) increases transcription of type I interferon genes (6).

and tumor cells from different tissues may help to clarify its function(s).

Interestingly, repression of myelin basic protein (MBP) translation during transport is mediated in part by hnRNP-E1 (Torvund-Jensen et al., 2014), which was also shown to regulate BC200 RNA function *in vitro* (Jang et al., 2017), providing a tenuous but potential link between BC200 RNA and myelination.

The data summarized here indicate that BC200 RNA plays important roles in post-transcriptional mRNA regulation, with most studies so far having focused on its function in postsynaptic translation. The Alu domain in BC200 RNA associated with SRP9/14 proteins resembles the translation arrest domain of the 7SL RNA-containing SRP, suggesting that it is likewise involved in translation inhibition. As a consequence, reduced BC200 RNA levels may lead to imbalanced postsynaptic translation. Although oligodendrocytes are the primary cell type affected in POLR3-HLD, a direct neuronal dysfunction is thought to be responsible for the neuronal loss observed in POLR3-HLD (e.g. cerebellum) and other POLR3-related disorders (e.g. cerebellum, striatum) (Wolf et al., 2014a; Azmanov et al., 2016; La Piana et al., 2016; Minnerop et al., 2017; Perrier et al., 2020a). The role of BC200 RNA in dendrites could affect the function and integrity of neurons in these brain regions. Furthermore, as mentioned above, a potential role for BC200 RNA in oligodendrocytes and/or their progenitor cells could contribute to the hypomyelination phenotype.

LINKING MUTATIONS IN GENES ENCODING SUBUNITS OF POL III TO THE INNATE IMMUNE SYSTEM

Although Pol III is best known for the nuclear transcription of small non-coding RNA genes, its function in the immune response is becoming increasingly clear. The discovery that Pol III does not only participate in nuclear transcription of small RNAs, including viral RNAs, but is also involved in the detection of invading DNA, expanded cellular activities of this enzyme to innate immunity. Pol III recognizes double-stranded transfected linear (ds)AT-rich DNA in the cytoplasm and transcribes it into 5' triphosphorylated RNA, which triggers the activation of retinoic acid-inducible gene I (RIG-I). RIG-I activation is dependent on the AT-content of the produced RNA since it can be abolished by

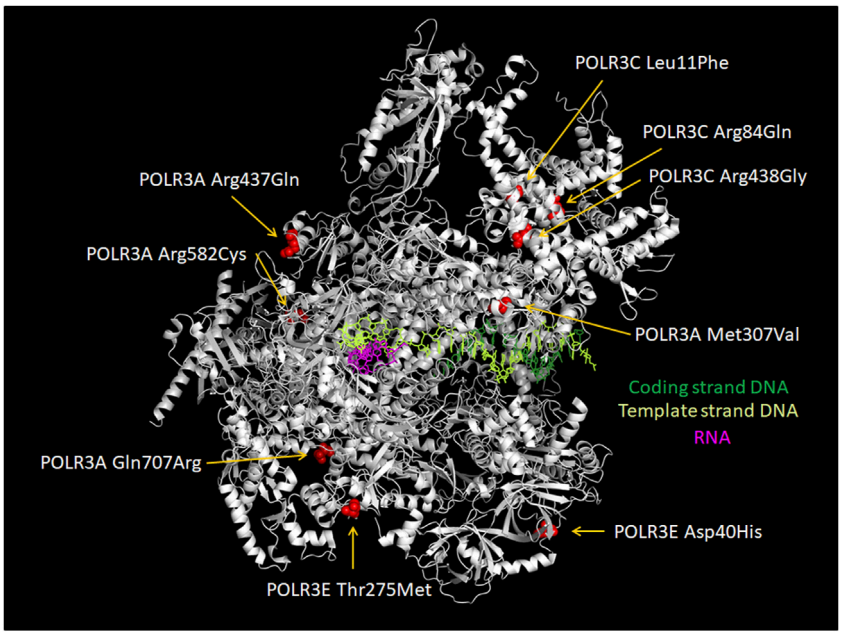


FIGURE 11 | Distribution of mutations in RNA polymerase III subunits associated with impaired innate immune defense. Spatial distribution of amino acid mutations that are associated with severe immune deficiency (POLR3E Asp40His) or with varicella zoster virus (VZV) encephalitis and/or pneumonitis (other displayed mutations). Mutated amino acids are depicted as spheres and highlighted in red. Individual mutations in subunits POLR3A, 3C, 3E are appropriately designated. The mutation POLR3F Arg50Trp cannot be displayed since the corresponding sequence was not resolved in the cryo-EM structure (PDB:7ae1). The Figure was created using Pymol.

insertion of GC sequences (Chiu et al., 2009). Activated RIG-I signals to the mitochondrial antiviral signaling protein (MAVS), resulting in the induction of the production of type 1 interferons (Ablasser et al., 2009; Chiu et al., 2009; **Figure 10**). The Pol III-dependent pathway of inducing innate immune response by production of RNAs from AT-rich DNA is complementary to the Cyclic GMP–AMP synthase (cGAS) pathway, which is activated by binding to DNA from invading microbes and production of cGAMP (Sun et al., 2013; Luecke et al., 2017; reviewed in; Tan et al. (2018)). Cytoplasmic DNA recognized by Pol III is derived from infections with Gram-negative bacteria (*Shigella flexneri* or *Legionella pneumophilia*), Gram-positive bacteria (*Listeria* monocytogenes) or from viral infection (herpes simplex virus 1), suggesting that both bacterial and viral sources can trigger the Pol III-dependent innate immune system (Chiu et al., 2009; Pollpeter et al., 2011; Jehl et al., 2012; Crill et al., 2015). In addition to defending cells against acute infectious threats, Pol III nuclear transcription also contributes to induction of interferon production in cells having been transfected with adenoviral DNA (Minamitani et al., 2011) or latently infected by the Epstein-Barr virus. Thus, nuclear Pol III transcribing adenoviral VA RNAs or Epstein-Barr viral EBER1 and EBER2 genes is also able to trigger RIG-I-dependent type 1 interferon production (Samanta et al., 2006; Minamitani et al., 2011).

Consistent with this role of Pol III in innate immunity, rare heterozygous genetic variants in the genes encoding Pol III subunits POLR3A (RPC1), POLR3C (RPC3), POLR3F (RPC6) and POLR3E (RPC5) were shown to strongly impair immune response to varicella-zoster virus (VZV) infections in humans. This reduced immune response resulted in the development of VZV pneumonitis (mutations in *POLR3A*, *POLR3C*) or of VZV encephalitis (mutations in *POLR3A*, *POLR3C*, *POLR3E* and *POLR3F*) (Ogunjimi et al., 2017; Carter-Timofte et al., 2018a, 2019; **Figure 11**). Importantly, Pol III mutations associated with susceptibility towards VZV infections have not been linked to POLR3-HLD or other neurodegenerative diseases (Ogunjimi et al., 2017). Together with the different mode of inheritance and the fact that these patients are healthy until VZV infections, this indicates that pathogenic mechanisms are likely fundamentally different.

Heterozygous missense variants affecting VZV immune response were first described in *POLR3A* and *POLR3C*. These variants led to reduced interferon production in peripheral blood mononuclear cells (PBMCs) from patients, which could be rescued *in vitro* by introduction of wild type alleles of the mutated genes into these cells, suggesting defects in Pol III cytoplasmic function. Nuclear 5S rRNA gene transcription was not affected in PBMCs carrying *POLR3A*- and *POLR3C* mutations (Ogunjimi et al., 2017). Although it remains possible that expression of other nuclear Pol III transcripts is impaired in these cells, these data suggest that cytoplasmic DNA transcription by Pol III has unique requirements in terms of polymerase-DNA- or polymerase-associated protein-interactions compared to the nuclear gene expression by the Pol III enzyme. In line with this hypothesis, structural studies of human Pol III found that mutations associated with severe VZV infections mapped to the periphery/surface of Pol III (Ramsay et al., 2020; Girbig et al., 2021) and made few contacts with other residues or subunits (Type IV in Girbig et al. (2021)), in contrast

to mutations associated with recessive POLR3-related disorders. One possibility, which is at present favored, would be that cytoplasmic AT-rich DNA does not require the presence of Pol III promoter elements (A-, B- or TATA-boxes) and transcription is therefore independent of Pol III transcription factors. However, it cannot be excluded that TATA-like elements, which may be present in AT-rich DNA and may result in the recruitment of TATA-binding protein (TBP)-containing TFIIIB-α or -β transcription initiation factors, would in turn enable the recruitment of Pol III to participate in cytoplasmic immune response. Consistent with this hypothesis, it should be noted that all Pol III transcription factors have been detected in the cytoplasm and would therefore be available for transcription of cytoplasmic DNA (for example, TBP: Hardivillé et al., 2020; BRF1: Mital et al., 1996; BDP1: Weser et al., 2004; TFIIIC: Dumay-Odelot et al., 2007; Schneider et al., 1989; Pol III: Jones et al., 2000; Haurie et al., 2010). Furthermore, it should be considered that viral genomes are not present in the cytoplasm as short dsDNA with possibly 3′ overhanging ssDNA elements comparable to AT-rich DNA fragments, but may be present as hundreds of kilobases long DNA elements without structural elements that would be required for factor-independent Pol III transcription initiation. This means that these viral DNAs cannot be transcribed highly efficiently like the short AT-rich dsDNAs without Pol III transcription factors but may have to rely on transcription factor-dependent mechanisms to elicit an efficient RIG-I-dependent immune response. Furthermore, it should be considered that the VZV genome in human cells is circular (Cohen, 2010) and thus lacks free 3′ overhangs which would be indispensable for factor-independent Pol III transcription initiation. As a consequence, it seems quite possible that the cellular, Pol III transcription-dependent immune response could also rely on Pol III transcription factors in the cytoplasm. Future research will clarify whether RIG-I activation by Pol III-transcribed RNAs occurs independently from transcription factor or with the involvement of TFIIIB and TFIIIC. Regardless of whether Pol III transcription factors are involved in these processes or not, the fact remains that sequences transcribed by Pol III must be AT-rich to elicit a RIG-I response.

Mechanistically, an observation may also link cytoplasmic immune response to a helicase activity intrinsic to Pol III. Arginine 84 in POLR3C (RPC3) was shown to be replaced by glutamine (Arg84Gln) in a patient who developed VZV-induced encephalitis (Ogunjimi et al., 2017; **Figure 11**). Interestingly, a mutation of the same residue of POLR3C, Arg84Ala, was shown to result in a defect in its intrinsic helicase activity *in vitro* (Ayoubi et al., 2019). This finding could be in line with a model in which Pol III helicase activity is required for dsDNA unwinding of viral cytoplasmic DNA.

Interestingly, a rare homozygous variant in *POLR3E* (Asp40His) was recently identified in a child with recurrent and systemic viral infections and Langerhans cell histiocytosis, indicating a more severe immune deficiency than in individuals carrying heterozygous mutations conferring specific VZV susceptibility who are otherwise healthy. Induction of interferon expression was triggered by the CG-rich (57,5%; Marti-Carreras and Maes, 2019) genome of human cytomegaly virus (HCMV) in control cells and abolished in patient cells homozygous for the POLR3E Asp40His mutation (Ramanathan et al., 2020), suggesting a different mechanism than the reported Pol III immune response to AT-rich DNA. The authors found that HCMV and sindbis virus infection induced POLR3E expression in control cells. Transfection of plasmid DNA also induced POLR3E expression and led to increased expression of 5S rRNA and a tRNA gene. In contrast, ectopic expression of wild-type POLR3E had a lower effect on expression of these Pol III target genes compared to the empty vector, and this response was absent or much lower with ectopic expression of mutant POLR3E. Finally, expression of the mutant POLR3E impaired formation of Pol III initiation complexes. These results suggest a role for Pol III nuclear transcription in the response to foreign viral and non-viral nucleic acids. Although the identification of additional patients with a similar phenotype is necessary to confirm that the POLR3E mutation is indeed causal, these data showing impaired Pol III transcription in response to foreign DNA, combined with the recessive mode of inheritance, suggest a different disease mechanism than in patients with VZV susceptibility. Despite immune dysfunction being the main feature, this phenotype could be related to the role of Pol III in nuclear transcription. As with other POLR3-related disorders, further investigations will help delineate the contribution of cytoplasmic and nuclear functions of Pol III to immune phenotypes.

CONCLUSION

The molecular-phenotypic relationships that may explain the development of POLR3-related disorders are as of yet only fragmentarily understood. The phenotypic heterogeneity, vast distribution of the mutations known so far across six subunits of Pol III and the resulting mutation-specific structural changes, as well as effects on transcription, argue against a single unifying disease-causing mechanism.

Although formal evidence is still lacking that reduction of Pol III transcript levels is the triggering factor for pathogenesis of POLR3-related disorders, it seems clear that affected neuroanatomical structures and their cellular components have particular vulnerabilities to impaired Pol III transcription or to altered amino acid loading of tRNAs. Mutations in genes encoding the TFIIIB-β component BRF1, Pol III subunits and aminoacyl-tRNA synthetases suggest that perturbations of protein synthesis and proper delivery of protein products to membranes, and possibly ectopic translation in neurons and oligodendrocytes, may play special roles in the development of POLR3-related disorders. Therefore, a dedicated analysis of the transcription of the tRNA, 7SL, and BC200 RNA genes is also necessary to obtain an integral picture of these diseases and their causes, as well as to generate therapeutic strategies for the future. Furthermore, small interspersed nuclear elements (SINEs) and especially the Alu gene subgroup might also play a role in the development of POLR3-HLD, since their promoters and RNA products show sequence and structural similarities to 7SL and BC200 RNAs. However, to the best of our knowledge, their involvement in POLR3-related disorders has not yet been investigated.

Surprisingly, and independently of POLR3-related neurological disorders, mutations in four Pol III subunits

affecting innate immune defense have also been described. Apparently, the key factor in the development of these diseases is the activity of the Pol III enzyme in the cytoplasm rather than the transcription of specific Pol III target genes. It will also be of interest for this spectrum of diseases to determine the exact underlying mechanism in order to develop potential therapies to reduce the risk of life-threatening complications from viral infections.

In summary, Pol III transcription has emerged as a key factor in the pathogenesis of several rare debilitating diseases. The establishment of molecular-pathological correlations will facilitate the development of rational therapies in the future.

AUTHOR CONTRIBUTIONS

MT: Conception and writing of the review article. EL: Conception and writing of the review article. KC: Conception and writing of the review article. FS: Writing of the review article. BB: Writing of the review article. GB: Writing of the review article.

FUNDING

MT's work received funding from the INSERM, the CNRS and from the Ligue Contre le Cancer Comités Dordogne and Landes. GB's work on POLR3-related disorders is supported by grants from the Canadian Institute of Health Research (377869, 426534). GB has received a Research Scholar Junior 1 award from the Fonds de Recherche du Quebec – Santé (FRQS) (2012–2016), the New Investigator Salary Award from the Canadian Institutes of Health Research (2017–2022) and a Research Scholar Senior award from the FRQS (2022–2025). KC received post-doctoral fellowships from the Canadian Institutes of Health Research (CIHR) and the Fonds de recherche du Québec-Santé (FRQS).

REFERENCES

Abascal-Palacios, G., Ramsay, E. P., Beuron, F., Morris, E., and Vannini, A. (2018). Structural Basis of RNA Polymerase III Transcription Initiation. *Nature* 553, 301–306. doi:10.1038/nature25441

Ablasser, A., Bauernfeind, F., Hartmann, G., Latz, E., Fitzgerald, K. A., and Hornung, V. (2009). RIG-I-dependent Sensing of poly(dA:dT) through the Induction of an RNA Polymerase III-Transcribed RNA Intermediate. *Nat. Immunol.* 10, 1065–1072. doi:10.1038/ni.1779

Adang, L. A., Sherbini, O., Ball, L., Bloom, M., Darbari, A., Amartino, H., et al. (2017). Revised Consensus Statement on the Preventive and Symptomatic Care of Patients with Leukodystrophies. *Mol. Genet. Metab.* 122, 18–32. doi:10.1016/j.ymgme.2017.08.006

Akopian, D., Shen, K., Zhang, X., and Shan, S.-o. (2013). Signal Recognition Particle: an Essential Protein-Targeting Machine. *Annu. Rev. Biochem.* 82, 693–721. doi:10.1146/annurev-biochem-072711-164732

Al Yazidi, G., Tran, L. T., Guerrero, K., Vanderver, A., Schiffmann, R., Wolf, N. I., et al. (2019). Dystonia in RNA Polymerase III-Related Leukodystrophy. *Mov Disord. Clin. Pract.* 6, 155–159. doi:10.1002/mdc3.12715

Anitei, M., and Pfeiffer, S. E. (2006). Myelin Biogenesis: Sorting Out Protein Trafficking. *Curr. Biol.* 16, R418–R421. doi:10.1016/j.cub.2006.05.010

Arimbasseri, A. G., Blewett, N. H., Iben, J. R., Lamichhane, T. N., Cherkasova, V., Hafner, M., et al. (2015). RNA Polymerase III Output Is Functionally Linked to tRNA Dimethyl-G26 Modification. *Plos Genet.* 11, e1005671. doi:10.1371/journal.pgen.1005671

Arimbasseri, A. G., and Maraia, R. J. (2016). RNA Polymerase III Advances: Structural and tRNA Functional Views. *Trends Biochem. Sci.* 41, 546–559. doi:10.1016/j.tibs.2016.03.003

Atrouni, S., Darazé, A., Tamraz, J., Cassia, A., Caillaud, C., and Mégarbané, A. (2003). Leukodystrophy Associated with Oligodontia in a Large Inbred Family: Fortuitous Association or New Entity? *Am. J. Med. Genet.* 118A, 76–81. doi:10.1002/ajmg.a.10019

Ayoubi, L. E., Dumay-Odelot, H., Chernev, A., Boissier, F., Minvielle-Sébastia, L., Urlaub, H., et al. (2019). The hRPC62 Subunit of Human RNA Polymerase III Displays Helicase Activity. *Nucleic Acids Res.* 47, 10313–10326. doi:10.1093/nar/gkz788

Azmanov, D. N., Siira, S. J., Chamova, T., Kaprelyan, A., Guergueltcheva, V., Shearwood, A.-M. J., et al. (2016). Transcriptome-wide Effects of aPOLR3Agene Mutation in Patients with an Unusual Phenotype of Striatal Involvement. *Hum. Mol. Genet.* 25, 4302–4314. doi:10.1093/hmg/ddw263

Báez-Becerra, C. T., Valencia-Rincón, E., Velásquez-Méndez, K., Ramírez-Suárez, N. J., Guevara, C., Sandoval-Hernandez, A., et al. (2020). Nucleolar Disruption, Activation of P53 and Premature Senescence in POLR3A-Mutated Wiedemann-Rautenstrauch Syndrome Fibroblasts. *Mech. Ageing Development* 192, 111360. doi:10.1016/j.mad.2020.111360

Baker, R. E., Camier, S., Sentenac, A., and Hall, B. D. (1987). Gene Size Differentially Affects the Binding of Yeast Transcription Factor Tau to Two Intragenic Regions. *Proc. Natl. Acad. Sci.* 84, 8768–8772. doi:10.1073/pnas.84.24.8768

Beauregard-Lacroix, E., Salian, S., Kim, H., Ehresmann, S., D'Amours, G., Gauthier, J., et al. (2020). A Variant of Neonatal Progeroid Syndrome, or Wiedemann-Rautenstrauch Syndrome, Is Associated with a Nonsense Variant in POLR3GL. *Eur. J. Hum. Genet.* 28, 461–468. doi:10.1038/s41431-019-0539-6

Behrens, A., Rodschinka, G., and Nedialkova, D. D. (2021). High-resolution Quantitative Profiling of tRNA Abundance and Modification Status in Eukaryotes by mim-tRNAseq. *Mol. Cel* 81, 1802–1815. doi:10.1016/j.molcel.2021.01.028

Bento-Abreu, A., Jager, G., Swinnen, B., Rué, L., Hendrickx, S., Jones, A., et al. (2018). Elongator Subunit 3 (ELP3) Modifies ALS through tRNA Modification. *Hum. Mol. Genet.* 27, 1276–1289. doi:10.1093/hmg/ddy043

Bernard, G., Chouery, E., Putorti, M. L., Tétreault, M., Takanohashi, A., Carosso, G., et al. (2011). Mutations of POLR3A Encoding a Catalytic Subunit of RNA Polymerase Pol III Cause a Recessive Hypomyelinating Leukodystrophy. *Am. J. Hum. Genet.* 89, 415–423. doi:10.1016/j.ajhg.2011.07.014

Bernard, G., Thiffault, I., Tetreault, M., Putorti, M. L., Bouchard, I., Sylvain, M., et al. (2010). Tremor-ataxia with central Hypomyelination (TACH) Leukodystrophy Maps to Chromosome 10q22.3-10q23.31. *Neurogenetics* 11, 457–464. doi:10.1007/s10048-010-0251-8

Bernard, G., and Vanderver, A. (1993). POLR3-Related Leukodystrophy, GeneReviews. (Seattle (WA): University of Washington, Seattle).

Bervoets, S., Wei, N., Erfurth, M. L., Yusein-Myashkova, S., Ermanoska, B., Mateiu, L., et al. (2019). Transcriptional Dysregulation by a Nucleus-Localized Aminoacyl-tRNA Synthetase Associated with Charcot-Marie-Tooth Neuropathy. *Nat. Commun.* 10, 5045. doi:10.1038/s41467-019-12909-9

Blocquel, D., Li, S., Wei, N., Daub, H., Sajish, M., Erfurth, M.-L., et al. (2017). Alternative Stable Conformation Capable of Protein Misinteraction Links tRNA Synthetase to Peripheral Neuropathy. *Nucleic Acids Res.* 45, 8091–8104. doi:10.1093/nar/gkx455

Blocquel, D., Sun, L., Matuszek, Z., Li, S., Weber, T., Kuhle, B., et al. (2019). CMT Disease Severity Correlates with Mutation-Induced Open Conformation of Histidyl-tRNA Synthetase, Not Aminoacylation Loss, in Patient Cells. *Proc. Natl. Acad. Sci. USA* 116, 19440–19448. doi:10.1073/pnas.1908288116

Blombach, F., Salvadori, E., Fouqueau, T., Yan, J., Reimann, J., Sheppard, C., et al. (2015). Archaeal TFEα/β Is a Hybrid of TFIIE and the RNA Polymerase III Subcomplex hRPC62/39. *Elife* 4, e08378. doi:10.7554/elife.08378

Boivin, V., Deschamps-Francoeur, G., Couture, S., Nottingham, R. M., Bouchard-Bourelle, P., Lambowitz, A. M., et al. (2018). Simultaneous Sequencing of Coding and Noncoding RNA Reveals a Human Transcriptome Dominated by a

Small Number of Highly Expressed Noncoding Genes. *RNA* 24, 950–965. doi:10.1261/rna.064493.117

Booy, E. P., Gussakovsky, D., Choi, T., and McKenna, S. A. (2020). The Non-coding RNA BC200 Associates with Polysomes to Positively Regulate mRNA Translation in Tumour Cells. *J. Biol. Chem.* 296, 100036. doi:10.1074/jbc.RA120.015775

Booy, E. P., McRae, E. K., Ezzati, P., Choi, T., Gussakovsky, D., and McKenna, S. A. (2018). Comprehensive Analysis of the BC200 Ribonucleoprotein Reveals a Reciprocal Regulatory Function with CSDE1/UNR. *Nucleic Acids Res.* 46, 11575–11591. doi:10.1093/nar/gky860

Booy, E. P., McRae, E. K., Koul, A., Lin, F., and McKenna, S. A. (2017). The Long Non-coding RNA BC200 (BCYRN1) Is Critical for Cancer Cell Survival and Proliferation. *Mol. Cancer* 16, 109. doi:10.1186/s12943-017-0679-7

Booy, E. P., McRae, E. K. S., Howard, R., Deo, S. R., Ariyo, E. O., Dzananovic, E., et al. (2016). RNA Helicase Associated with AU-Rich Element (RHAU/DHX36) Interacts with the 3′-Tail of the Long Non-coding RNA BC200 (BCYRN1). *J. Biol. Chem.* 291, 5355–5372. doi:10.1074/jbc.m115.711499

Borchardt, E. K., Martinez, N. M., and Gilbert, W. V. (2020). Regulation and Function of RNA Pseudouridylation in Human Cells. *Annu. Rev. Genet.* 54, 309–336. doi:10.1146/annurev-genet-112618-043830

Borck, G., Hög, F., Dentici, M. L., Tan, P. L., Sowada, N., Medeira, A., et al. (2015). BRF1 Mutations Alter RNA Polymerase III-dependent Transcription and Cause Neurodevelopmental Anomalies. *Genome Res.* 25, 609. doi:10.1101/gr.176925.114

Bovia, F., Wolff, N., Ryser, S., and Strub, K. (1997). The SRP9/14 Subunit of the Human Signal Recognition Particle Binds to a Variety of Alu-like RNAs and with Higher Affinity Than its Mouse Homolog. *Nucleic Acids Res.* 25, 318–326. doi:10.1093/nar/25.2.318

Bredow, S., Kleinert, H., and Benecke, B.-J. (1990a). Sequence and Factor Requirements for Faithful *In Vitro* Transcription of Human 7SL DNA. *Gene* 86, 217–225. doi:10.1016/0378-1119(90)90282-v

Bredow, S., Sürig, D., Müller, J., Kleinert, H., and Benecke, B.-J. (1990b). Activating-transcription-factor (ATF) Regulates Human 7S L RNA Transcription by RNA Polymerase III *in Vivo* and *in Vitro*. *Nucl. Acids Res.* 18, 6779–6784. doi:10.1093/nar/18.23.6779

Briese, M., Saal, L., Appenzeller, S., Moradi, M., Baluapuri, A., and Sendtner, M. (2016). Whole Transcriptome Profiling Reveals the RNA Content of Motor Axons. *Nucleic Acids Res.* 44, e33. doi:10.1093/nar/gkv1027

Brosius, J. (1999). RNAs from All Categories Generate Retrosequences that May Be Exapted as Novel Genes or Regulatory Elements. *Gene* 238, 115–134. doi:10.1016/s0378-1119(99)00227-9

Burnol, A.-F., Margottin, F., Schultz, P., Marsolier, M.-C., Oudet, P., and Sentenac, A. (1993). Basal Promoter and Enhancer Element of Yeast U6 snRNA Gene. *J. Mol. Biol.* 233, 644–658. doi:10.1006/jmbi.1993.1542

Canella, D., Bernasconi, D., Gilardi, F., LeMartelot, G., Migliavacca, E., Praz, V., et al. (2012). A Multiplicity of Factors Contributes to Selective RNA Polymerase III Occupancy of a Subset of RNA Polymerase III Genes in Mouse Liver. *Genome Res.* 22, 666–680. doi:10.1101/gr.130286.111

Canella, D., Praz, V., Reina, J. H., Cousin, P., and Hernandez, N. (2010). Defining the RNA Polymerase III Transcriptome: Genome-wide Localization of the RNA Polymerase III Transcription Machinery in Human Cells. *Genome Res.* 20, 710–721. doi:10.1101/gr.101337.109

Carbon, P., and Krol, A. (1991). Transcription of the *Xenopus laevis* Selenocysteine tRNA(Ser)Sec Gene: a System that Combines an Internal B Box and Upstream Elements Also Found in U6 snRNA Genes. *EMBO J.* 10, 599–606. doi:10.1002/j.1460-2075.1991.tb07987.x

Carrière, L., Graziani, S., Alibert, O., Ghavi-Helm, Y., Boussouar, F., Humbertclaude, H., et al. (2012). Genomic Binding of Pol III Transcription Machinery and Relationship with TFIIS Transcription Factor Distribution in Mouse Embryonic Stem Cells. *Nucleic Acids Res.* 40, 270–283. doi:10.1093/nar/gkr737

Carter-Timofte, M. E., Hansen, A. F., Mardahl, M., Fribourg, S., Rapaport, F., Zhang, S. Y., et al. (2018a). Varicella-zoster Virus CNS Vasculitis and RNA Polymerase III Gene Mutation in Identical Twins. *Neurol. Neuroimmunol Neuroinflamm* 5, e500. doi:10.1212/nxi.0000000000000500

Carter-Timofte, M. E., Hansen, A. F., Christiansen, M., Paludan, S. R., and Mogensen, T. H. (2019). Mutations in RNA Polymerase III Genes and Defective DNA Sensing in Adults with Varicella-Zoster Virus CNS Infection. *Genes Immun.* 20, 214–223. doi:10.1038/s41435-018-0027-y

Carter-Timofte, M. E., Paludan, S. R., and Mogensen, T. H. (2018b). RNA Polymerase III as a Gatekeeper to Prevent Severe VZV Infections. *Trends Mol. Med.* 24, 904–915. doi:10.1016/j.molmed.2018.07.009

Castilho, B. A., Shanmugam, R., Silva, R. C., Ramesh, R., Himme, B. M., and Sattlegger, E. (2014). Keeping the eIF2 Alpha Kinase Gcn2 in Check. *Biochim. Biophys. Acta (Bba) - Mol. Cel Res.* 1843, 1948–1968. doi:10.1016/j.bbamcr.2014.04.006

Castle, J. C., Armour, C. D., Löwer, M., Haynor, D., Biery, M., Bouzek, H., et al. (2010). Digital Genome-wide ncRNA Expression, Including SnoRNAs, across 11 Human Tissues Using polyA-Neutral Amplification. *PLoS One* 5, e11779. doi:10.1371/journal.pone.0011779

Chiu, Y.-H., Macmillan, J. B., and Chen, Z. J. (2009). RNA Polymerase III Detects Cytosolic DNA and Induces Type I Interferons through the RIG-I Pathway. *Cell* 138, 576–591. doi:10.1016/j.cell.2009.06.015

Choe, B. K., and Taylor, M. W. (1972). Kinetics of Synthesis and Characterization of Transfer RNA Precursors in Mammalian Cells. *Biochim. Biophys. Acta (Bba) - Nucleic Acids Protein Synth.* 272, 275–287. doi:10.1016/0005-2787(72)90251-1

Choquet, K., Forget, D., Meloche, E., Dicaire, M.-J., Bernard, G., Vanderver, A., et al. (2019a). Leukodystrophy-associated POLR3A Mutations Down-Regulate the RNA Polymerase III Transcript and Important Regulatory RNA BC200. *J. Biol. Chem.* 294, 7445–7459. doi:10.1074/jbc.ra118.006271

Choquet, K., Pinard, M., Yang, S., Poitras, C., Sgarioto, N., Lanviere, R., et al. (2019b). The Leukodystrophy Mutation Polr3b R103H Causes Homozygote Mouse Embryonic Lethality and Impairs RNA Polymerase III Biogenesis. *Mol. Brain* 12, 59. doi:10.1186/s13041-019-0479-7

Choquet, K., Yang, S., Moir, R. D., Forget, D., Larivière, R., Bouchard, A., et al. (2017). Absence of Neurological Abnormalities in Mice Homozygous for the Polr3a G672E Hypomyelinating Leukodystrophy Mutation. *Mol. Brain* 10, 13. doi:10.1186/s13041-017-0294-y

Chung, A., Dahan, N., Alarcon, J. M., and Fenton, A. A. (2017). Effects of Regulatory BC1 RNA Deletion on Synaptic Plasticity, Learning, and Memory. *Learn. Mem.* 24, 646–649. doi:10.1101/lm.045617.117

Cohen, J. I. (2010). The Varicella-Zoster Virus Genome. *Curr. Top. Microbiol. Immunol.* 342, 1–14. doi:10.1007/82_2010_10

Cozen, A. E., Quartley, E., Holmes, A. D., Hrabeta-Robinson, E., Phizicky, E. M., and Lowe, T. M. (2015). ARM-seq: AlkB-Facilitated RNA Methylation Sequencing Reveals a Complex Landscape of Modified tRNA Fragments. *Nat. Methods* 12, 879–884. doi:10.1038/nmeth.3508

Crill, E. K., Furr-Rogers, S. R., and Marriott, I. (2015). RIG-I Is Required for VSV-Induced Cytokine Production by Murine Glia and Acts in Combination with DAI to Initiate Responses to HSV-1. *Glia* 63, 2168–2180. doi:10.1002/glia.22883

Coleman, O. I., and Haller, D. (2019). ER Stress and the UPR in Shaping Intestinal Tissue Homeostasis and Immunity. *Front. Immunol.* 10, 2825. doi:10.3389/fimmu.2019.02825

Daniels, G. R., and Deininger, P. L. (1985). Repeat Sequence Families Derived from Mammalian tRNA Genes. *Nature* 317, 819–822. doi:10.1038/317819a0

Daoud, H., Tétreault, M., Gibson, W., Guerrero, K., Cohen, A., Gburek-Augustat, J., et al. (2013). Mutations inPOLR3AandPOLR3Bare a Major Cause of Hypomyelinating Leukodystrophies with or without Dental Abnormalities And/or Hypogonadotropic Hypogonadism. *J. Med. Genet.* 50, 194–197. doi:10.1136/jmedgenet-2012-101357

DeGasperis, S. M., Bernard, G., Wolf, N. I., Miller, E., and Pohl, D. (2020). 4H Leukodystrophy. *Neurol. Genet.* 6, e409. doi:10.1212/nxg.0000000000000409

Del Rio, S., and Setzer, D. R. (1993). The Role of Zinc Fingers in Transcriptional Activation by Transcription Factor IIIA. *Proc. Natl. Acad. Sci.* 90, 168–172. doi:10.1073/pnas.90.1.168

Dieci, G., Fermi, B., and Bosio, M. C. (2014). Investigating Transcription Reinitiation through *In Vitro* Approaches. *Transcription* 5, e27704. doi:10.4161/trns.27704

Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M., and Pagano, A. (2007). The Expanding RNA Polymerase III Transcriptome. *Trends Genet.* 23, 614–622. doi:10.1016/j.tig.2007.09.001

Dieci, G., Giuliodori, S., Catellani, M., Percudani, R., and Ottonello, S. (2002). Intragenic Promoter Adaptation and Facilitated RNA Polymerase III Recycling in the Transcription of SCR1, the 7SL RNA Gene ofSaccharomyces Cerevisiae. *J. Biol. Chem.* 277, 6903–6914. doi:10.1074/jbc.m105036200

Dieci, G., and Sentenac, A. (1996). Facilitated Recycling Pathway for RNA Polymerase III. *Cell* 84, 245–252. doi:10.1016/s0092-8674(00)80979-4

Dittmar, K. A., Goodenbour, J. M., and Pan, T. (2006). Tissue-specific Differences in Human Transfer RNA Expression. *Plos Genet.* 2, e221. doi:10.1371/journal.pgen.0020221

Djordjevic, D., Pinard, M., Gauthier, M.-S., Smith-Hicks, C., Hoffman, T. L., Wolf, N. I., et al. (2021). De Novo variants in POLR3B Cause Ataxia, Spasticity, and Demyelinating Neuropathy. *Am. J. Hum. Genet.* 108, 186–193. doi:10.1016/j.ajhg.2020.12.002

Dong, J., Qiu, H., Garcia-Barrio, M., Anderson, J., and Hinnebusch, A. G. (2000). Uncharged tRNA Activates GCN2 by Displacing the Protein Kinase Moiety from a Bipartite tRNA-binding Domain. *Mol. Cel* 6, 269–279. doi:10.1016/s1097-2765(00)00028-9

Dorboz, I., Dumay-Odelot, H., Boussaid, K., Bouyacoub, Y., Barreau, P., Samaan, S., et al. (2018). Mutation in POLR3K Causes Hypomyelinating Leukodystrophy and Abnormal Ribosomal RNA Regulation. *Neurol. Genet.* 4, e289. doi:10.1212/nxg.0000000000000289

Dumay-Odelot, H., Durrieu-Gaillard, S., Da Silva, D., Roeder, R. G., and Teichmann, M. (2010). Cell Growth- and Differentiation-dependent Regulation of RNA Polymerase III Transcription. *Cell Cycle* 9, 3687–3699. doi:10.4161/cc.9.18.13203

Dumay-Odelot, H., Durrieu-Gaillard, S., El Ayoubi, L., Parrot, C., and Teichmann, M. (2014). Contributions of *In Vitro* Transcription to the Understanding of Human RNA Polymerase III Transcription. *Transcription* 5, e27526. doi:10.4161/trns.27526

Dumay-Odelot, H., Marck, C., Durrieu-Gaillard, S., Lefebvre, O., Jourdain, S., Prochazkova, M., et al. (2007). Identification, Molecular Cloning, and Characterization of the Sixth Subunit of Human Transcription Factor TFIIIC. *J. Biol. Chem.* 282, 17179–17189. doi:10.1074/jbc.m611542200

Duning, K., Buck, F., Barnekow, A., and Kremerskothen, J. (2008). SYNCRIP, a Component of Dendritically Localized mRNPs, Binds to the Translation Regulator BC200 RNA. *J. Neurochem.* 105, 351–359. doi:10.1111/j.1471-4159.2007.05138.x

Durrieu-Gaillard, S., Dumay-Odelot, H., Boldina, G., Tourasse, N. J., Allard, D., André, F., et al. (2018). Regulation of RNA Polymerase III Transcription during Transformation of Human IMR90 Fibroblasts with Defined Genetic Elements. *Cell Cycle* 17, 605–615. doi:10.1080/15384101.2017.1405881

Englert, M., Felis, M., Junker, V., and Beier, H. (2004). Novel Upstream and Intragenic Control Elements for the RNA Polymerase III-dependent Transcription of Human 7SL RNA Genes. *Biochimie* 86, 867–874. doi:10.1016/j.biochi.2004.10.012

Ferrari, R., Rivetti, C., Acker, J., and Dieci, G. (2004). Distinct Roles of Transcription Factors TFIIIB and TFIIIC in RNA Polymerase III Transcription Reinitiation. *Proc. Natl. Acad. Sci.* 101, 13442–13447. doi:10.1073/pnas.0403851101

Fields, R. D. (2008). White Matter Matters. *Sci. Am.* 298, 42–49. doi:10.1038/scientificamerican0308-54

Franca, M. M., Han, X., Funari, M. F. A., Lerario, A. M., Nishi, M. Y., Fontenele, E. G. P., et al. (2019). Exome Sequencing Reveals the POLR3H Gene as a Novel Cause of Primary Ovarian Insufficiency. *J. Clin. Endocrinol. Metab.* 104, 2827–2841. doi:10.1210/jc.2018-02485

Freude, K., Hoffmann, K., Jensen, L.-R., Delatycki, M. B., des Portes, V., Moser, B., et al. (2004). Mutations in the FTSJ1 Gene Coding for a Novel S-Adenosylmethionine-Binding Protein Cause Nonsyndromic X-Linked Mental Retardation. *Am. J. Hum. Genet.* 75, 305–309. doi:10.1086/422507

Fröhlich, D., Mendes, M., Kuen, A. J., Bongers, A., Herold, M., Salomons, G. S., et al. (2020). A Hypomorphic Dars1D367Y Model Recapitulates Key Aspects of the Leukodystrophy HBSL. *Front Cel Neurosci.* 14, 625879.

Fröhlich, D., Suchowerska, A. K., Voss, C., He, R., Wolvetang, E., von Jonquieres, G., et al. (2018). Expression Pattern of the Aspartyl-tRNA Synthetase DARS in the Human Brain. *Front. Mol. Neurosci.* 11, 81. doi:10.3389/fnmol.2018.00081

Fulga, T. A., Sinning, I., Dobberstein, B., and Pool, M. R. (2001). SRbeta Coordinates Signal Sequence Release from SRP with Ribosome Binding to the Translocon. *EMBO J.* 20, 2338–2347. doi:10.1093/emboj/20.9.2338

Gabrielsen, O. S., and Sentenac, A. (1991). RNA Polymerase III (C) and its Transcription Factors. *Trends Biochem. Sci.* 16, 412–416. doi:10.1016/0968-0004(91)90166-s

Gauquelin, L., Cayami, F. K., Sztriha, L., Yoon, G., Tran, L. T., Guerrero, K., et al. (2019). Clinical Spectrum of POLR3-Related Leukodystrophy Caused by Biallelic POLR1C Pathogenic Variants. *Neurol. Genet.* 5, e369. doi:10.1212/nxg.0000000000000369

Gauquelin, L., Tétreault, M., Thiffault, I., Farrow, E., Miller, N., Yoo, B., et al. (2018). POLR3A Variants in Hereditary Spastic Paraplegia and Ataxia. *Brain* 141, e1. doi:10.1093/brain/awx290

Geva, M., Cabilly, Y., Assaf, Y., Mindroul, N., Marom, L., Raini, G., et al. (2010). A Mouse Model for Eukaryotic Translation Initiation Factor 2B-Leucodystrophy Reveals Abnormal Development of Brain white Matter. *Brain* 133, 2448–2461. doi:10.1093/brain/awq180

Ghoumid, J., Petit, F., Boute-Benejean, O., Frenois, F., Cartigny, M., Vanlerberghe, C., et al. (2017). Cerebellar Hypoplasia with Endosteal Sclerosis Is a POLR3-Related Disorder. *Eur. J. Hum. Genet.* 25, 1011–1014. doi:10.1038/ejhg.2017.73

Gingold, H., Tehler, D., Christoffersen, N. R., Nielsen, M. M., Asmar, F., Kooistra, S. M., et al. (2014). A Dual Program for Translation Regulation in Cellular Proliferation and Differentiation. *Cell* 158, 1281–1292. doi:10.1016/j.cell.2014.08.011

Girbig, M., Misiaszek, A. D., Vorländer, M. K., Lafita, A., Grötsch, H., Baudin, F., et al. (2021). Cryo-EM Structures of Human RNA Polymerase III in its Unbound and Transcribing States. *Nat. Struct. Mol. Biol.* 28, 210–219. doi:10.1038/s41594-020-00555-5

Gogakos, T., Brown, M., Garzia, A., Meyer, C., Hafner, M., and Tuschl, T. (2017). Characterizing Expression and Processing of Precursor and Mature Human tRNAs by Hydro-tRNAseq and PAR-CLIP. *Cel Rep.* 20, 1463–1475. doi:10.1016/j.celrep.2017.07.029

Goodarzi, H., Nguyen, H. C. B., Zhang, S., Dill, B. D., Molina, H., and Tavazoie, S. F. (2016). Modulated Expression of Specific tRNAs Drives Gene Expression and Cancer Progression. *Cell* 165, 1416–1427. doi:10.1016/j.cell.2016.05.046

Gouge, J., and Vannini, A. (2018). New Tricks for an Old Dog: Brf2-dependent RNA Polymerase III Transcription in Oxidative Stress and Cancer. *Transcription* 9, 61–66. doi:10.1080/21541264.2017.1335269

Gu, C., Begley, T. J., and Dedon, P. C. (2014). tRNA Modifications Regulate Translation during Cellular Stress. *FEBS Lett.* 588, 4287–4296. doi:10.1016/j.febslet.2014.09.038

Gundelfinger, E. D., Krause, E., Melli, M., and Dobberstein, B. (1983). The Organization of the 7SL RNA in the Signal Recognition Particle. *Nucl. Acids Res.* 11, 7363–7374. doi:10.1093/nar/11.21.7363

Halic, M., Becker, T., Pool, M. R., Spahn, C. M. T., Grassucci, R. A., Frank, J., et al. (2004). Structure of the Signal Recognition Particle Interacting with the Elongation-Arrested Ribosome. *Nature* 427, 808–814. doi:10.1038/nature02342

Hardivillé, S., Banerjee, P. S., Selen Alpergin, E. S., Smith, D. M., Han, G., Ma, J., et al. (2020). TATA-box Binding Protein O-GlcNAcylation at T114 Regulates Formation of the B-TFIID Complex and Is Critical for Metabolic Gene Regulation. *Mol. Cel* 77, 1143–1152. doi:10.1016/j.molcel.2019.11.022

Harting, I., Al-Saady, M., Krägeloh-Mann, I., Bley, A., Hempel, M., Bierhals, T., et al. (2020). POLR3A Variants with Striatal Involvement and Extrapyramidal Movement Disorder. *Neurogenetics* 21, 121–133. doi:10.1007/s10048-019-00602-4

Haurie, V., Durrieu-Gaillard, S., Dumay-Odelot, H., Da Silva, D., Rey, C., Prochazkova, M., et al. (2010). Two Isoforms of Human RNA Polymerase III with Specific Functions in Cell Growth and Transformation. *Proc. Natl. Acad. Sci.* 107, 4176–4181. doi:10.1073/pnas.0914980107

Hayne, C. K., Schmidt, C. A., Haque, M. I., Matera, A. G., and Stanley, R. E. (2020). Reconstitution of the Human tRNA Splicing Endonuclease Complex: Insight into the Regulation of Pre-tRNA Cleavage. *Nucleic Acids Res.* 48, 7609–7622. doi:10.1093/nar/gkaa438

He, W., Zhang, H.-M., Chong, Y. E., Guo, M., Marshall, A. G., and Yang, X.-L. (2011). Dispersed Disease-Causing Neomorphic Mutations on a Single Protein Promote the Same Localized Conformational Opening. *Proc. Natl. Acad. Sci.* 108, 12307–12312. doi:10.1073/pnas.1104293108

Hiraide, T., Kubota, K., Kono, Y., Watanabe, S., Matsubayashi, T., Nakashima, M., et al. (2020). POLR3A Variants in Striatal Involvement without Diffuse Hypomyelination. *Brain Development* 42, 363–368. doi:10.1016/j.braindev.2019.12.012

Horan, K. A., Hansen, K., Jakobsen, M. R., Holm, C. K., Søby, S., Unterholzner, L., et al. (2013). Proteasomal Degradation of Herpes Simplex Virus Capsids in Macrophages Releases DNA to the Cytosol for Recognition by DNA Sensors. *J.I.* 190, 2311–2319. doi:10.4049/jimmunol.1202749

Howe, J. G., and Shu, M.-D. (1989). Epstein-Barr Virus Small RNA (EBER) Genes: Unique Transcription Units that Combine RNA Polymerase II and III Promoter Elements. *Cell* 57, 825–834. doi:10.1016/0092-8674(89)90797-6

Hsieh, Y.-J., Wang, Z., Kovelman, R., and Roeder, R. G. (1999). Cloning and Characterization of Two Evolutionarily Conserved Subunits (TFIIIC102 and TFIIIC63) of Human TFIIIC and Their Involvement in Functional Interactions with TFIIIB and RNA Polymerase III. *Mol. Cel Biol.* 19, 4944–4952. doi:10.1128/mcb.19.7.4944

Iacoangeli, A., Dosunmu, A., Eom, T., Stefanov, D. G., and Tiedge, H. (2017). Regulatory BC1 RNA in Cognitive Control. *Learn. Mem.* 24, 267–277. doi:10.1101/lm.045427.117

Iacoangeli, A., and Tiedge, H. (2013). Translational Control at the Synapse: Role of RNA Regulators. *Trends Biochem. Sci.* 38, 47–55. doi:10.1016/j.tibs.2012.11.001

Ishimura, R., Nagy, G., Dotu, I., Zhou, H., Yang, X.-L., Schimmel, P., et al. (2014). Ribosome Stalling Induced by Mutation of a CNS-specific tRNA Causes Neurodegeneration. *Science* 345, 455–459. doi:10.1126/science.1249749

Itoh, M., Dai, H., Horike, S.-i., Gonzalez, J., Kitami, Y., Meguro-Horike, M., et al. (2019). Biallelic KARS Pathogenic Variants Cause an Early-Onset Progressive Leukodystrophy. *Brain* 142, 560–573. doi:10.1093/brain/awz001

Jakovcevski, I., Filipovic, R., Mo, Z., Rakic, S., and Zecevic, N. (2009). Oligodendrocyte Development and the Onset of Myelination in the Human Fetal Brain. *Front. Neuroanat.* 3, 5. doi:10.3389/neuro.05.005.2009

Jang, S., Shin, H., Lee, J., Kim, Y., Bak, G., and Lee, Y. (2017). Regulation of BC200 RNA-Mediated Translation Inhibition by hnRNP E1 and E2. *FEBS Lett.* 591, 393–405. doi:10.1002/1873-3468.12544

Jarrous, N. (2017). Roles of RNase P and its Subunits. *Trends Genet.* 33, 594–603. doi:10.1016/j.tig.2017.06.006

Jay, A. M., Conway, R. L., Thiffault, I., Saunders, C., Farrow, E., Adams, J., et al. (2016). Neonatal Progeriod Syndrome Associated with Biallelic Truncating Variants inPOLR3A. *Am. J. Med. Genet.* 170, 3343–3346. doi:10.1002/ajmg.a.37960

Jee, Y. H., Sowada, N., Markello, T. C., Rezvani, I., Borck, G., and Baron, J. (2017). BRF1mutations in a Family with Growth Failure, Markedly Delayed Bone Age, and central Nervous System Anomalies. *Clin. Genet.* 91, 739–747. doi:10.1111/cge.12887

Jehl, S. P., Nogueira, C. V., Zhang, X., and Starnbach, M. N. (2012). IFNγ Inhibits the Cytosolic Replication of Shigella Flexneri via the Cytoplasmic RNA Sensor RIG-I. *Plos Pathog.* 8, e1002809. doi:10.1371/journal.ppat.1002809

Johnson, E. M., Kinoshita, Y., Weinreb, D. B., Wortman, M. J., Simon, R., Khalili, K., et al. (2006). Role of Purα in Targeting mRNA to Sites of Translation in Hippocampal Neuronal Dendrites. *J. Neurosci. Res.* 83, 929–943. doi:10.1002/jnr.20806

Jones, E., Kimura, H., Vigneron, M., Wang, Z., Roeder, R. G., and Cook, P. R. (2000). Isolation and Characterization of Monoclonal Antibodies Directed against Subunits of Human RNA Polymerases I, II, and III. *Exp. Cel Res.* 254, 163–172. doi:10.1006/excr.1999.4739

Kapur, M., Ganguly, A., Nagy, G., Adamson, S. I., Chuang, J. H., Frankel, W. N., et al. (2020). Expression of the Neuronal tRNA N-Tr20 Regulates Synaptic Transmission and Seizure Susceptibility. *Neuron* 108, 193–208. doi:10.1016/j.neuron.2020.07.023

Katibah, G. E., Lee, H. J., Huizar, J. P., Vogan, J. M., Alber, T., and Collins, K. (2013). tRNA Binding, Structure, and Localization of the Human Interferon-Induced Protein IFIT5. *Mol. Cel* 49, 743–750. doi:10.1016/j.molcel.2012.12.015

Kennedy, M. J., and Hanus, C. (2019). Architecture and Dynamics of the Neuronal Secretory Network. *Annu. Rev. Cel Dev. Biol.* 35, 543–566. doi:10.1146/annurev-cellbio-100818-125418

Khanam, T., Rozhdestvensky, T. S., Bundman, M., Galiveti, C. R., Handel, S., Sukonina, V., et al. (2007). Two Primate-specific Small Non-protein-coding RNAs in Transgenic Mice: Neuronal Expression, Subcellular Localization and Binding Partners. *Nucleic Acids Res.* 35, 529–539. doi:10.1093/nar/gkl1082

Khatter, H., Vorländer, M. K., and Müller, C. W. (2017). RNA Polymerase I and III: Similar yet Unique. *Curr. Opin. Struct. Biol.* 47, 88–94. doi:10.1016/j.sbi.2017.05.008

Kickhoefer, V. A., Emre, N., Stephen, A. G., Poderycki, M. J., and Rome, L. H. (2003). Identification of Conserved Vault RNA Expression Elements and a Non-expressed Mouse Vault RNA Gene. *Gene* 309, 65–70. doi:10.1016/s0378-1119(03)00507-9

Kieckhaefer, J. E., Lukovac, S., Ye, D. Z., Lee, D., Beetler, D. J., Pack, M., et al. (2016). The RNA Polymerase III Subunit Polr3b Is Required for the Maintenance of Small Intestinal Crypts in Mice. *Cell Mol. Gastroenterol. Hepatol.* 2, 783–795. doi:10.1016/j.jcmgh.2016.08.003

Kim, Y., Lee, J., Shin, H., Jang, S., Kim, S. C., Lee, Y., et al. (2017). Biosynthesis of Brain Cytoplasmic 200 RNA. *Sci. Rep.* 7, 6884. doi:10.1038/s41598-017-05097-3

Kobayashi, S., Agui, K., Kamo, S., Li, Y., and Anzai, K. (2000). Neural BC1 RNA Associates with Pur α, a Single-Stranded DNA and RNA Binding Protein, Which Is Involved in the Transcription of the BC1 RNA Gene. *Biochem. Biophysical Res. Commun.* 277, 341–347. doi:10.1006/bbrc.2000.3683

Kour, S., Rajan, D. S., Fortuna, T. R., Anderson, E. N., Ward, C., Lee, Y., et al. (2021). Loss of Function Mutations in GEMIN5 Cause a Neurodevelopmental Disorder. *Nat. Commun.* 12, 2558. doi:10.1038/s41467-021-22627-w

Kremerskothen, J., Zopf, D., Walter, P., Cheng, J.-G., Nettermann, M., Niewerth, U., et al. (1998). Heterodimer SRP9/14 Is an Integral Part of the Neural BC200 RNP in Primate Brain. *Neurosci. Lett.* 245, 123–126. doi:10.1016/s0304-3940(98)00215-8

Kutter, C., Brown, G. D., Gonçalves, Â., Wilson, M. D., Watt, S., Brazma, A., et al. (2011). Pol III Binding in Six Mammals Shows Conservation Among Amino Acid Isotypes Despite Divergence Among tRNA Genes. *Nat. Genet.* 43, 948–955. doi:10.1038/ng.906

La Piana, R., Cayami, F. K., Tran, L. T., Guerrero, K., van Spaendonk, R., Õunap, K., et al. (2016). Diffuse Hypomyelination Is Not Obligate for POLR3-Related Disorders. *Neurology* 86, 1622–1626. doi:10.1212/wnl.0000000000002612

Lakkaraju, A. K. K., Luyet, P.-P., Parone, P., Falguières, T., and Strub, K. (2007). Inefficient Targeting to the Endoplasmic Reticulum by the Signal Recognition Particle Elicits Selective Defects in post-ER Membrane Trafficking. *Exp. Cel Res.* 313, 834–847. doi:10.1016/j.yexcr.2006.12.003

Lakkaraju, A. K. K., Mary, C., Scherrer, A., Johnson, A. E., and Strub, K. (2008). SRP Keeps Polypeptides Translocation-Competent by Slowing Translation to Match Limiting ER-Targeting Sites. *Cell* 133, 440–451. doi:10.1016/j.cell.2008.02.049

Lant, J. T., Berg, M. D., Heinemann, I. U., Brandl, C. J., and O'Donoghue, P. (2019). Pathways to Disease from Natural Variations in Human Cytoplasmic tRNAs. *J. Biol. Chem.* 294, 5294–5308. doi:10.1074/jbc.rev118.002982

Lefebvre, S., Bürglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., et al. (1995). Identification and Characterization of a Spinal Muscular Atrophy-Determining Gene. *Cell* 80, 155–165. doi:10.1016/0092-8674(95)90460-3

Lefèvre, S., Dumay-Odelot, H., El-Ayoubi, L., Budd, A., Legrand, P., Pinaud, N., et al. (2011). Structure-function Analysis of hRPC62 Provides Insights into RNA Polymerase III Transcription Initiation. *Nat. Struct. Mol. Biol.* 18, 352–358. doi:10.1038/nsmb.1996

Levi, O., and Arava, Y. S. (2021). RNA Modifications as a Common Denominator between tRNA and mRNA. *Curr. Genet.* 67, 545–551. doi:10.1007/s00294-021-01168-1

Li, D. K., Tisdale, S., Lotti, F., and Pellizzoni, L. (2014). SMN Control of RNP Assembly: from post-transcriptional Gene Regulation to Motor Neuron Disease. *Semin. Cel Developmental Biol.* 32, 22–29. doi:10.1016/j.semcdb.2014.04.026

Li, G., Eriani, G., Wang, E.-D., and Zhou, X.-L. (2021a). Distinct Pathogenic Mechanisms of Various RARS1 Mutations in Pelizaeus-merzbacher-like Disease. *Sci. China Life Sci.* doi:10.1007/s11427-020-1838-2

Li, L., Yu, Z., Zhao, D., Ren, Y., Hou, H., and Xu, Y. (2021b). Structure of Human RNA Polymerase III Elongation Complex. *Cell Res.* 31, 791–800. doi:10.1038/s41422-021-00472-2

Lin, D., Pestova, T. V., Hellen, C. U. T., and Tiedge, H. (2008). Translational Control by a Small RNA: Dendritic BC1 RNA Targets the Eukaryotic Initiation Factor 4A Helicase Mechanism. *Mol. Cel Biol.* 28, 3008–3019. doi:10.1128/mcb.01800-07

Lin, W., and Popko, B. (2009). Endoplasmic Reticulum Stress in Disorders of Myelinating Cells. *Nat. Neurosci.* 12, 379–385. doi:10.1038/nn.2273

Lu, J.-F., Lawler, A. M., Watkins, P. A., Powers, J. M., Moser, A. B., Moser, H. W., et al. (1997). A Mouse Model for X-Linked Adrenoleukodystrophy. *Proc. Natl. Acad. Sci.* 94, 9366–9371. doi:10.1073/pnas.94.17.9366

Luecke, S., Holleufer, A., Christensen, M. H., Jønsson, K. L., Boni, G. A., Sørensen, L. K., et al. (2017). cGAS Is Activated by DNA in a Length-dependent Manner. *EMBO Rep.* 18, 1707–1715. doi:10.15252/embr.201744017

Mahalingam, R., Kennedy, P. G., and Gilden, D. H. (1999). The Problems of Latent Varicella Zoster Virus in Human Ganglia: Precise Cell Location and Viral Content. *J. Neurovirol.* 5, 445–448. doi:10.3109/13550289909045372

Mao, X., and Darby, M. K. (1993). A Position-dependent Transcription-Activating Domain in TFIIIA. *Mol. Cel Biol* .13, 7496–7506. doi:10.1128/mcb.13.12.7496-7506.1993

Martí-Carreras, J., and Maes, P. (2019). Human Cytomegalovirus Genomics and Transcriptomics through the Lens of Next-Generation Sequencing: Revision and Future Challenges. *Virus Genes* 55, 138–164. doi:10.1007/s11262-018-1627-3

Martignetti, J. A., and Brosius, J. (1995). BC1 RNA: Transcriptional Analysis of a Neural Cell-specific RNA Polymerase III Transcript. *Mol. Cel Biol.* 15, 1642–1650. doi:10.1128/mcb.15.3.1642

Marton, R. M., Miura, Y., Sloan, S. A., Li, Q., Revah, O., Levy, R. J., et al. (2019). Differentiation and Maturation of Oligodendrocytes in Human Three-Dimensional Neural Cultures. *Nat. Neurosci.* 22, 484–491. doi:10.1038/s41593-018-0316-9

Massenet, S. (2019). *In Vivo* assembly of Eukaryotic Signal Recognition Particle: A Still Enigmatic Process Involving the SMN Complex. *Biochimie* 164, 99–104. doi:10.1016/j.biochi.2019.04.007

McLaurin, J., Trudel, G. C., Shaw, I. T., Antel, J. P., and Cashman, N. R. (1995). A Human Glial Hybrid Cell Line Differentially Expressing Genes Subserving Oligodendrocyte and Astrocyte Phenotype. *J. Neurobiol.* 26, 283–293. doi:10.1002/neu.480260212

Meissner, W., Wanandi, I., Carbon, P., Krol, A., and Seifart, K. H. (1994). Transcription Factors Required for the Expression of *Xenopus laevis* Selenocysteine tRNA *In Vitro*. *Nucleic Acids Res.* 22, 553–559. doi:10.1093/nar/22.4.553

Mendes, M. I., Green, L. M. C., Bertini, E., Tonduti, D., Aiello, C., Smith, D., et al. (2020). RARS1 -related Hypomyelinating Leukodystrophy: Expanding the Spectrum. *Ann. Clin. Transl Neurol.* 7, 83–93. doi:10.1002/acn3.50960

Mendes, M. I., Gutierrez Salazar, M., Guerrero, K., Thiffault, I., Salomons, G. S., Gauquelin, L., et al. (2018). Bi-allelic Mutations in EPRS, Encoding the Glutamyl-Prolyl-Aminoacyl-tRNA Synthetase, Cause a Hypomyelinating Leukodystrophy. *Am. J. Hum. Genet.* 102, 676–684. doi:10.1016/j.ajhg.2018.02.011

Merheb, E., Cui, M.-H., DuBois, J. C., Branch, C. A., Gulinello, M., and Shafit-Zagardo, B. (2021). Defective Myelination In An RNA Polymerase III Mutant Leukodystrophic Mouse.. *bioRxiv*. doi:10.1101/2020.12.09.418657

Minamitani, T., Iwakiri, D., and Takada, K. (2011). Adenovirus Virus-Associated RNAs Induce Type I Interferon Expression through a RIG-I-Mediated Pathway. *J. Virol.* 85, 4035–4040. doi:10.1128/jvi.02160-10

Minnerop, M., Kurzwelly, D., Wagner, H., Soehn, A. S., Reichbauer, J., Tao, F., et al. (2017). Hypomorphic Mutations in POLR3A Are a Frequent Cause of Sporadic and Recessive Spastic Ataxia. *Brain* 140, 1561–1578. doi:10.1093/brain/awx095

Mital, R., Kobayashi, R., and Hernandez, N. (1996). RNA Polymerase III Transcription from the Human U6 and Adenovirus Type 2 VAI Promoters Has Different Requirements for Human BRF, a Subunit of Human TFIIIB. *Mol. Cel Biol.* 16, 7031–7042. doi:10.1128/mcb.16.12.7031

Moir, R. D., Lavados, C., Lee, J., and Willis, I. M. (2021). Functional Characterization of Polr3a Hypomyelinating Leukodystrophy Mutations in the *S. cerevisiae* Homolog, RPC160. *Gene* 768, 145259. doi:10.1016/j.gene.2020.145259

Muddashetty, R. S., Khanam, T., Kondrashov, A., Bundman, M., Iacoangeli, A., Kremerskothen, J., et al. (2002). Poly(A)-binding Protein Is Associated with Neuronal BC1 and BC200 Ribonucleoprotein Particles. *J. Mol. Biol.* 321, 433–445. doi:10.1016/s0022-2836(02)00655-1

Nakayama, T., Wu, J., Galvin-Parton, P., Weiss, J., Andriola, M. R., Hill, R. S., et al. (2017). Deficient Activity of Alanyl-tRNA Synthetase Underlies an Autosomal Recessive Syndrome of Progressive Microcephaly, Hypomyelination, and Epileptic Encephalopathy. *Hum. Mutat.* 38, 1348–1354. doi:10.1002/humu.23250

Nevin, Z. S., Factor, D. C., Karl, R. T., Douvaras, P., Laukka, J., Windrem, M. S., et al. (2017). Modeling the Mutational and Phenotypic Landscapes of Pelizaeus-Merzbacher Disease with Human iPSC-Derived Oligodendrocytes. *Am. J. Hum. Genet.* 100, 617–634. doi:10.1016/j.ajhg.2017.03.005

Numasawa-Kuroiwa, Y., Okada, Y., Shibata, S., Kishi, N., Akamatsu, W., Shoji, M., et al. (2014). Involvement of ER Stress in Dysmyelination of Pelizaeus-Merzbacher Disease with PLP1 Missense Mutations Shown by iPSC-Derived Oligodendrocytes. *Stem Cel Rep.* 2, 648–661. doi:10.1016/j.stemcr.2014.03.007

Odermatt, B., Wellershaus, K., Wallraff, A., Seifert, G., Degen, J., Euwens, C., et al. (2003). Connexin 47 (Cx47)-Deficient Mice with Enhanced green Fluorescent Protein Reporter Gene Reveal Predominant Oligodendrocytic Expression of Cx47 and Display Vacuolized Myelin in the CNS. *J. Neurosci.* 23, 4549–4559. doi:10.1523/jneurosci.23-11-04549.2003

Ogunjimi, B., Zhang, S.-Y., Sørensen, K. B., Skipper, K. A., Carter-Timofte, M., Kerner, G., et al. (2017). Inborn Errors in RNA Polymerase III Underlie Severe Varicella Zoster Virus Infections. *J. Clin. Invest.* 127, 3543–3556. doi:10.1172/jci92280

Oler, A. J., Alla, R. K., Roberts, D. N., Wong, A., Hollenhorst, P. C., Chandler, K. J., et al. (2010). Human RNA Polymerase III Transcriptomes and Relationships to Pol II Promoter Chromatin and Enhancer-Binding Factors. *Nat. Struct. Mol. Biol.* 17, 620–628. doi:10.1038/nsmb.1801

Ontiveros, R. J., Shen, H., Stoute, J., Yanas, A., Cui, Y., Zhang, Y., et al. (2020). Coordination of mRNA and tRNA Methylations by TRMT10A. *Proc. Natl. Acad. Sci. USA* 117, 7782–7791. doi:10.1073/pnas.1913448117

Ornelas, I. M., McLane, L. E., Saliu, A., Evangelou, A. V., Khandker, L., and Wood, T. L. (2016). Heterogeneity in Oligodendroglia: Is it Relevant to Mouse Models and Human Disease? *J. Neurosci. Res.* 94, 1421–1433. doi:10.1002/jnr.23900

Osterman, B., Sylvain, M., Chouinard, S., and Bernard, G. (2012). Tremor-ataxia with central Hypomyelination (TACH): Dystonia as a New Clinical Feature. *Mov Disord.* 27, 1829–1830. doi:10.1002/mds.25270

Öztürk, Z., O'Kane, C. J., and Pérez-Moreno, J. J. (2020). Axonal Endoplasmic Reticulum Dynamics and its Roles in Neurodegeneration. *Front. Neurosci.* 14, 48. doi:10.3389/fnins.2020.00048

Pan, T. (2018). Modifications and Functional Genomics of Human Transfer RNA. *Cel Res.* 28, 395–404. doi:10.1038/s41422-018-0013-y

Paolacci, S., Li, Y., Agolini, E., Bellacchio, E., Arboleda-Bustos, C. E., Carrero, D., et al. (2018). Specific Combinations of Biallelic POLR3A Variants Cause Wiedemann-Rautenstrauch Syndrome. *J. Med. Genet.* 55, 837–846. doi:10.1136/jmedgenet-2018-105528

Parikh, S., Bernard, G., Leventer, R. J., van der Knaap, M. S., van Hove, J., Pizzino, A., et al. (2015). A Clinical Approach to the Diagnosis of Patients with Leukodystrophies and Genetic Leukoencephelopathies. *Mol. Genet. Metab.* 114, 501–515. doi:10.1016/j.ymgme.2014.12.434

Pavon-Eternod, M., Gomes, S., Geslain, R., Dai, Q., Rosner, M. R., and Pan, T. (2009). tRNA Over-expression in Breast Cancer and Functional Consequences. *Nucleic Acids Res.* 37, 7268–7280. doi:10.1093/nar/gkp787

Pelletier, F., Perrier, S., Cayami, F. K., Mirchi, A., Saikali, S., Tran, L. T., et al. (2021). Endocrine and Growth Abnormalities in 4H Leukodystrophy Caused by Variants in POLR3A, POLR3B, and POLR1C. *J. Clin. Endocrinol. Metab.* 106, e660–e674.

Perrier, S., Michell-Robinson, M. A., and Bernard, G. (2020b). POLR3-Related Leukodystrophy: Exploring Potential Therapeutic Approaches. *Front. Cel Neurosci.* 14, 631802. doi:10.3389/fncel.2020.631802

Perrier, S., Gauquelin, L., Fallet-Bianco, C., Dishop, M. K., Michell-Robinson, M. A., Tran, L. T., et al. (2020a). Expanding the Phenotypic and Molecular Spectrum of RNA Polymerase III-Related Leukodystrophy. *Neurol. Genet.* 6, e425. doi:10.1212/nxg.0000000000000425

Petrie, J. L., Swan, C., Ingram, R. M., Frame, F. M., Collins, A. T., Dumay-Odelot, H., et al. (2019). Effects on Prostate Cancer Cells of Targeting RNA Polymerase III. *Nucleic Acids Res.* 47, 3937–3956. doi:10.1093/nar/gkz128

Pfeiffer, S., Warrington, A., and Bansal, R. (1993). The Oligodendrocyte and its many Cellular Processes. *Trends Cel Biol.* 3, 191–197. doi:10.1016/0962-8924(93)90213-k

Piana, R. L., Tonduti, D., Dressman, H. G., Schmidt, J. L., Murnick, J., Brais, B., et al. (2014). Brain Magnetic Resonance Imaging (MRI) Pattern Recognition in Pol III-Related Leukodystrophies. *J. Child. Neurol.* 29, 214–220. doi:10.1177/0883073813503902

Piazzon, N., Schlotter, F., Lefebvre, S., Dodre, M., Mereau, A., Soret, J., et al. (2013). Implication of the SMN Complex in the Biogenesis and Steady State Level of the Signal Recognition Particle. *Nucleic Acids Res.* 41, 1255–1272. doi:10.1093/nar/gks1224

Pinkard, O., McFarland, S., Sweet, T., and Coller, J. (2020). Quantitative tRNA-sequencing Uncovers Metazoan Tissue-specific tRNA Regulation. *Nat. Commun.* 11, 4104. doi:10.1038/s41467-020-17879-x

Pollpeter, D., Komuro, A., Barber, G. N., and Horvath, C. M. (2011). Impaired Cellular Responses to Cytosolic DNA or Infection with Listeria Monocytogenes and Vaccinia Virus in the Absence of the Murine LGP2 Protein. *PLoS One* 6, e18842. doi:10.1371/journal.pone.0018842

Pool, M. R. (2005). Signal Recognition Particles in Chloroplasts, Bacteria, Yeast and Mammals (Review). *Mol. Membr. Biol.* 22, 3–15. doi:10.1080/09687860400026348

Potic, A., Brais, B., Choquet, K., Schiffmann, R., and Bernard, G. (2012). 4H Syndrome with Late-Onset Growth Hormone Deficiency Caused by POLR3A Mutations. *Arch. Neurol.* 69, 920–923. doi:10.1001/archneurol.2011.1963

Potic, A., Popovic, V., Ostojic, J., Pekic, S., Kozic, D., Guerrero, K., et al. (2015). Neurogenic Bladder and Neuroendocrine Abnormalities in Pol III-Related Leukodystrophy. *BMC Neurol.* 15, 22. doi:10.1186/s12883-015-0283-7

Pujol, A., Hindelang, C., Callizot, N., Bartsch, U., Schachner, M., Mandel, J. L., et al. (2002). Late Onset Neurological Phenotype of the X-ALD Gene Inactivation in Mice: a Mouse Model for Adrenomyeloneuropathy. *Hum. Mol. Genet.* 11, 499–505. doi:10.1093/hmg/11.5.499

Ramanathan, A., Weintraub, M., Orlovetskie, N., Serruya, R., Mani, D., Marcu, O., et al. (2020). A Mutation in POLR3E Impairs Antiviral Immune Response and RNA Polymerase III. *Proc. Natl. Acad. Sci. USA* 117, 22113–22121. doi:10.1073/pnas.2009947117

Ramírez, O. A., and Couve, A. (2011). The Endoplasmic Reticulum and Protein Trafficking in Dendrites and Axons. *Trends Cel Biol.* 21, 219–227. doi:10.1016/j.tcb.2010.12.003

Ramos, J., and Fu, D. (2019). The Emerging Impact of tRNA Modifications in the Brain and Nervous System. *Biochim. Biophys. Acta (Bba) - Gene Regul. Mech.* 1862, 412–428. doi:10.1016/j.bbagrm.2018.11.007

Ramsay, E. P., Abascal-Palacios, G., Daiß, J. L., King, H., Gouge, J., Pilsl, M., et al. (2020). Structure of Human RNA Polymerase III. *Nat. Commun.* 11, 6409. doi:10.1038/s41467-020-20262-5

Richards, M. R., Plummer, L., Chan, Y.-M., Lippincott, M. F., Quinton, R., Kumanov, P., et al. (2017). Phenotypic Spectrum ofPOLR3Bmutations: Isolated Hypogonadotropic Hypogonadism without Neurological or Dental Anomalies. *J. Med. Genet.* 54, 19–25. doi:10.1136/jmedgenet-2016-104064

Roeder, R. G. (2019). 50+ Years of Eukaryotic Transcription: an Expanding Universe of Factors and Mechanisms. *Nat. Struct. Mol. Biol.* 26, 783–791. doi:10.1038/s41594-019-0287-x

Rydning, S. L., Koht, J., Sheng, Y., Sowa, P., Hjorthaug, H. S., Wedding, I. M., et al. (2019). Biallelic POLR3A Variants Confirmed as a Frequent Cause of Hereditary Ataxia and Spastic Paraparesis. *Brain* 142, e12. doi:10.1093/brain/awz041

Saitsu, H., Osaka, H., Sasaki, M., Takanashi, J.-i., Hamada, K., Yamashita, A., et al. (2011). Mutations in POLR3A and POLR3B Encoding RNA Polymerase III Subunits Cause an Autosomal-Recessive Hypomyelinating Leukoencephalopathy. *Am. J. Hum. Genet.* 89, 644–651. doi:10.1016/j.ajhg.2011.10.003

Samanta, M., Iwakiri, D., Kanda, T., Imaizumi, T., and Takada, K. (2006). EB Virus-Encoded RNAs Are Recognized by RIG-I and Activate Signaling to Induce Type I IFN. *EMBO J.* 25, 4207–4214. doi:10.1038/sj.emboj.7601314

Samson, J., Cronin, S., and Dean, K. (2018). BC200 (BCYRN1) - the Shortest, Long, Non-coding RNA Associated with Cancer. *Non-coding RNA Res.* 3, 131–143. doi:10.1016/j.ncrna.2018.05.003

Sasaki, M., Takanashi, J.-i., Tada, H., Sakuma, H., Furushima, W., and Sato, N. (2009). Diffuse Cerebral Hypomyelination with Cerebellar Atrophy and Hypoplasia of the Corpus Callosum. *Brain Development* 31, 582–587. doi:10.1016/j.braindev.2008.09.003

Schaffer, A. E., Pinkard, O., and Coller, J. M. (2019). tRNA Metabolism and Neurodevelopmental Disorders. *Annu. Rev. Genom. Hum. Genet.* 20, 359–387. doi:10.1146/annurev-genom-083118-015334

Schaub, M., Krol, A., and Carbon, P. (2000). Structural Organization of Staf-DNA Complexes. *Nucleic Acids Res.* 28, 2114–2121. doi:10.1093/nar/28.10.2114

Schiffmann, R., and van der Knaap, M. S. (2009). Invited Article: an MRI-Based Approach to the Diagnosis of white Matter Disorders. *Neurology* 72, 750–759. doi:10.1212/01.wnl.0000343049.00540.c8

Schmidt, J. L., Pizzino, A., Nicholl, J., Foley, A., Wang, Y., Rosenfeld, J. A., et al. (2020). Estimating the Relative Frequency of Leukodystrophies and Recommendations for Carrier Screening in the Era of Next-generation Sequencing. *Am. J. Med. Genet.* 182, 1906–1912. doi:10.1002/ajmg.a.61641

Schneider, H. R., Waldschmidt, R., Jahn, D., and Seifart, K. H. (1989). Purification of Human Transcription Factor IIIC and its Binding to the Gene for Ribosomal 5S RNA. *Nucl. Acids Res.* 17, 5003–5016. doi:10.1093/nar/17.13.5003

Schramm, L., and Hernandez, N. (2002). Recruitment of RNA Polymerase III to its Target Promoters. *Genes Dev.* 16, 2593–2620. doi:10.1101/gad.1018902

Schramm, L., Pendergrast, P. S., Sun, Y., and Hernandez, N. (2000). Different Human TFIIIB Activities Direct RNA Polymerase III Transcription from TATA-Containing and TATA-Less Promoters. *Genes Dev.* 14, 2650–2663. doi:10.1101/gad.836400

Sharkia, R., Zalan, A., Jabareen-Masri, A., Zahalka, H., and Mahajnah, M. (2019). A New Case Confirming and Expanding the Phenotype Spectrum of ADAT3-Related Intellectual Disability Syndrome. *Eur. J. Med. Genet.* 62, 103549. doi:10.1016/j.ejmg.2018.10.001

Shimojima, K., Shimada, S., Tamasaki, A., Akaboshi, S., Komoike, Y., Saito, A., et al. (2014). Novel Compound Heterozygous Mutations of POLR3A Revealed by Whole-Exome Sequencing in a Patient with Hypomyelination. *Brain Development* 36, 315–321. doi:10.1016/j.braindev.2013.04.011

Shin, H., Lee, J., Kim, Y., Jang, S., Lee, Y., Kim, S., et al. (2017b). Knockdown of BC200 RNA Expression Reduces Cell Migration and Invasion by Destabilizing mRNA for Calcium-Binding Protein S100A11. *RNA Biol.* 14, 1418–1430. doi:10.1080/15476286.2017.1297913

Shin, H., Lee, J., Kim, Y., Jang, S., Ohn, T., and Lee, Y. (2017a). Identifying the Cellular Location of Brain Cytoplasmic 200 RNA Using an RNA-Recognizing Antibody. *BMB Rep.* 50, 318–322. doi:10.5483/bmbrep.2017.50.6.217

Siira, S. J., Rossetti, G., Richman, T. R., Perks, K., Ermer, J. A., Kuznetsova, I., et al. (2018). Concerted Regulation of Mitochondrial and Nuclear Non-coding RNAs by a Dual-Targeted RNase Z. *EMBO Rep.* 19. doi:10.15252/embr.201846198

Simons, C., Griffin, L. B., Helman, G., Golas, G., Pizzino, A., Bloom, M., et al. (2015). Loss-of-function Alanyl-tRNA Synthetase Mutations Cause an Autosomal-Recessive Early-Onset Epileptic Encephalopathy with Persistent Myelination Defect. *Am. J. Hum. Genet.* 96, 675–681. doi:10.1016/j.ajhg.2015.02.012

Singh, R., Gupta, S. C., Peng, W.-X., Zhou, N., Pochampally, R., Atfi, A., et al. (2016). Regulation of Alternative Splicing of Bcl-X by BC200 Contributes to Breast Cancer Pathogenesis. *Cell Death Dis.* 7, e2262. doi:10.1038/cddis.2016.168

Skreka, K., Schafferer, S., Nat, I.-R., Zywicki, M., Salti, A., Apostolova, G., et al. (2012). Identification of Differentially Expressed Non-coding RNAs in Embryonic Stem Cell Neural Differentiation. *Nucleic Acids Res.* 40, 6001–6015. doi:10.1093/nar/gks311

Skryabin, B. V., Sukonina, V., Jordan, U., Lewejohann, L., Sachser, N., Muslimov, I., et al. (2003). Neuronal Untranslated BC1 RNA: Targeted Gene Elimination in Mice. *Mol. Cel Biol.* 23, 6435–6441. doi:10.1128/mcb.23.18.6435-6441.2003

Slosarek, E. L., Schuh, A. L., Pustova, I., Johnson, A., Bird, J., Johnson, M., et al. (2018). Pathogenic TFG Mutations Underlying Hereditary Spastic Paraplegia Impair Secretory Protein Trafficking and Axon Fasciculation. *Cel Rep.* 24, 2248–2260. doi:10.1016/j.celrep.2018.07.081

Sosińska-Zawierucha, P., Zawierucha, P., Bręborowicz, A., and Barciszewski, J. (2018). Prediction of Secondary and Tertiary Structures of Human BC200 RNA (BCYRN1) Based on Experimental and Bioinformatic Cross-Validation. *Biochem. J.* 475, 2727–2748. doi:10.1042/BCJ20180239

Steenweg, M. E., Vanderver, A., Blaser, S., Bizzi, A., de Koning, T. J., Mancini, G. M. S., et al. (2010). Magnetic Resonance Imaging Pattern Recognition in Hypomyelinating Disorders. *Brain* 133, 2971–2982. doi:10.1093/brain/awq257

Sun, L., Wei, N., Kuhle, B., Blocquel, D., Novick, S., Zhou, H., et al. (2021). CMT2N-causing Aminoacylation Domain Mutants Enable Nrp1 Interaction with AlaRS. *Proc. Natl. Acad. Sci. U S A.* 118, e2012898118. doi:10.1073/pnas.2012898118

Sun, L., Wu, J., Du, F., Chen, X., and Chen, Z. J. (2013). Cyclic GMP-AMP Synthase Is a Cytosolic DNA Sensor that Activates the Type I Interferon Pathway. *Science* 339, 786–791. doi:10.1126/science.1232458

Sutcliffe, J. G., Milner, R. J., Bloom, F. E., and Lerner, R. A. (1982). Common 82-nucleotide Sequence Unique to Brain RNA. *Proc. Natl. Acad. Sci.* 79, 4942–4946. doi:10.1073/pnas.79.16.4942

Sutcliffe, J. G., Milner, R. J., Gottesfeld, J. M., and Lerner, R. A. (1984). Identifier Sequences Are Transcribed Specifically in Brain. *Nature* 308, 237–241. doi:10.1038/308237a0

Taft, R. J., Vanderver, A., Leventer, R. J., Damiani, S. A., Simons, C., Grimmond, S. M., et al. (2013). Mutations in DARS Cause Hypomyelination with Brain Stem and Spinal Cord Involvement and Leg Spasticity. *Am. J. Hum. Genet.* 92, 774–780. doi:10.1016/j.ajhg.2013.04.006

Tahmasebi, S., Khoutorsky, A., Mathews, M. B., and Sonenberg, N. (2018). Translation Deregulation in Human Disease. *Nat. Rev. Mol. Cel Biol.* 19, 791–807. doi:10.1038/s41580-018-0034-x

Takaku, H., Minagawa, A., Takagi, M., and Nashimoto, M. (2003). A Candidate Prostate Cancer Susceptibility Gene Encodes tRNA 3' Processing Endoribonuclease. *Nucleic Acids Res.* 31, 2272–2278. doi:10.1093/nar/gkg337

Tan, X., Sun, L., Chen, J., and Chen, Z. J. (2018). Detection of Microbial Infections through Innate Immune Sensing of Nucleic Acids. *Annu. Rev. Microbiol.* 72, 447–478. doi:10.1146/annurev-micro-102215-095605

Teichmann, M., Dieci, G., Huet, J., Rüth, J., Sentenac, A., Seifart, K. H., et al. (1997). Functional Interchangeability of TFIIIB Components from Yeast and Human Cells Invitro. *EMBO J.* 16, 4708–4716. doi:10.1093/emboj/16.15.4708

Teichmann, M., Dieci, G., Pascali, C., and Boldina, G. (2010). General Transcription Factors and Subunits of RNA Polymerase III. *Transcription* 1, 130–135. doi:10.4161/trns.1.3.13192

Teichmann, M., and Seifart, K. H. (1995). Physical Separation of Two Different Forms of Human TFIIIB Active in the Transcription of the U6 or the VAI Gene *In Vitro*. *EMBO J.* 14, 5974–5983. doi:10.1002/j.1460-2075.1995.tb00286.x

Teichmann, M., Wang, Z., and Roeder, R. G. (2000). A Stable Complex of a Novel Transcription Factor IIB- Related Factor, Human TFIIIB50, and Associated Proteins Mediate Selective Transcription by RNA Polymerase III of Genes with Upstream Promoter Elements. *Proc. Natl. Acad. Sci.* 97, 14200–14205. doi:10.1073/pnas.97.26.14200

Terhal, P. A., Vlaar, J. M., Middelkamp, S., Nievelstein, R. A. J., Nikkels, P. G. J., Ross, J., et al. (2020). Biallelic Variants in POLR3GL Cause Endosteal Hyperostosis and Oligodontia. *Eur. J. Hum. Genet.* 28, 31–39. doi:10.1038/s41431-019-0427-0

Terrey, M., Adamson, S. I., Gibson, A. L., Deng, T., Ishimura, R., Chuang, J. H., et al. (2020). GTPBP1 Resolves Paused Ribosomes to Maintain Neuronal Homeostasis. *Elife* 9, e62731. doi:10.7554/elife.62731

Tétreault, M., Choquet, K., Orcesi, S., Tonduti, D., Balottin, U., Teichmann, M., et al. (2011). Recessive Mutations in POLR3B, Encoding the Second Largest Subunit of Pol III, Cause a Rare Hypomyelinating Leukodystrophy. *Am. J. Hum. Genet.* 89, 652–655. doi:10.1016/j.ajhg.2011.10.006

Tétreault, M., Putorti, M. L., Thiffault, I., Sylvain, M., Venderver, A., Schiffmann, R., et al. (2012). TACH Leukodystrophy: Locus Refinement to Chromosome 10q22.3-23.1. *Can. J. Neurol. Sci.* 39, 122–123. doi:10.1017/s0317167100022174

Thiffault, I., Wolf, N. I., Forget, D., Guerrero, K., Tran, L. T., Choquet, K., et al. (2015). Recessive Mutations in POLR1C Cause a Leukodystrophy by Impairing Biogenesis of RNA Polymerase III. *Nat. Commun.* 6, 7623. doi:10.1038/ncomms8623

Tiedge, H., Chen, W., and Brosius, J. (1993). Primary Structure, Neural-specific Expression, and Dendritic Location of Human BC200 RNA. *J. Neurosci.* 13, 2382–2390. doi:10.1523/jneurosci.13-06-02382.1993

Timmons, M., Tsokos, M., Asab, M. A., Seminara, S. B., Zirzow, G. C., Kaneski, C. R., et al. (2006). Peripheral and central Hypomyelination with Hypogonadotropic Hypogonadism and Hypodontia. *Neurology* 67, 2066–2069. doi:10.1212/01.wnl.0000247666.28904.35

Torrent, M., Chalancon, G., de Groot, N. S., Wuster, A., and Madan Babu, M. (2018). Cells Alter Their tRNA Abundance to Selectively Regulate Protein Synthesis during Stress Conditions. *Sci. Signal.* 11. doi:10.1126/scisignal.aat6409

Torvund-Jensen, J., Steengaard, J., Reimer, L., Fihl, L. B., and Laursen, L. S. (2014). Transport and Translation of MBP mRNA Is Regulated Differently by Distinct hnRNP Proteins. *J. Cel Sci.* 127, 1550–1564. doi:10.1242/jcs.140855

Tress, O., Maglione, M., Zlomuzica, A., May, D., Dicke, N., Degen, J., et al. (2011). Pathologic and Phenotypic Alterations in a Mouse Expressing a Connexin47 Missense Mutation that Causes Pelizaeus-merzbacher-like Disease in Humans. *Plos Genet.* 7, e1002146. doi:10.1371/journal.pgen.1002146

Ullu, E., and Weiner, A. M. (1985). Upstream Sequences Modulate the Internal Promoter of the Human 7SL RNA Gene. *Nature* 318, 371–374. doi:10.1038/318371a0

van der Knaap, M. S., and Bugiani, M. (2017). Leukodystrophies: a Proposed Classification System Based on Pathological Changes and Pathogenetic Mechanisms. *Acta Neuropathol.* 134, 351–382. doi:10.1007/s00401-017-1739-1

Vanderver, A., Prust, M., Tonduti, D., Mochel, F., Hussey, H. M., Helman, G., et al. (2015). Case Definition and Classification of Leukodystrophies and Leukoencephalopathies. *Mol. Genet. Metab.* 114, 494–500. doi:10.1016/j.ymgme.2015.01.006

Vanderver, A., Tonduti, D., Bernard, G., Lai, J., Rossi, C., Carosso, G., et al. (2013). More Than Hypomyelination in Pol-III Disorder. *J. Neuropathol. Exp. Neurol.* 72, 67–75. doi:10.1097/nen.0b013e31827c99d2

Verberne, E. A., Dalen Meurs, L., Wolf, N. I., and Haelst, M. M. (2020). 4H Leukodystrophy Caused by a Homozygous POLR3B Mutation: Further Delineation of the Phenotype. *Am. J. Med. Genet.* 182, 1776–1779. doi:10.1002/ajmg.a.61600

Voorhees, R. M., and Hegde, R. S. (2015). Structures of the Scanning and Engaged States of the Mammalian SRP-Ribosome Complex. *Elife* 4. doi:10.7554/elife.07975

Vorländer, M. K., Jungblut, A., Karius, K., Baudin, F., Grötsch, H., Kosinski, J., et al. (2020). Structure of the TFIIIC Subcomplex τA Provides Insights into RNA Polymerase III Pre-initiation Complex Formation. *Nat. Commun.* 11, 4905. doi:10.1038/s41467-020-18707-y

Vorländer, M. K., Khatter, H., Wetzel, R., Hagen, W. J. H., and Müller, C. W. (2018). Molecular Mechanism of Promoter Opening by RNA Polymerase III. *Nature* 553, 295–300. doi:10.1038/nature25440

Wakasugi, K., and Yokosawa, T. (2020). Non-canonical Functions of Human Cytoplasmic Tyrosyl-, Tryptophanyl- and Other Aminoacyl-tRNA Synthetases. *Enzymes* 48, 207–242. doi:10.1016/bs.enz.2020.04.001

Wambach, J. A., Wegner, D. J., Patni, N., Kircher, M., Willing, M. C., Baldridge, D., et al. (2018). Bi-allelic POLR3A Loss-Of-Function Variants Cause Autosomal-Recessive Wiedemann-Rautenstrauch Syndrome. *Am. J. Hum. Genet.* 103, 968–975. doi:10.1016/j.ajhg.2018.10.010

Wang, Q., Li, S., Wan, F., Xu, Y., Wu, Z., Cao, M., et al. (2021). Structural Insights into Transcriptional Regulation of Human RNA Polymerase III. *Nat. Struct. Mol. Biol.* 28, 220–227. doi:10.1038/s41594-021-00557-x

Wang, Z., and Roeder, R. G. (1995). Structure and Function of a Human Transcription Factor TFIIIB Subunit that Is Evolutionarily Conserved and Contains Both TFIIB- and High-Mobility-Group Protein 2-related Domains. *Proc. Natl. Acad. Sci.* 92, 7026–7030. doi:10.1073/pnas.92.15.7026

Watson, J. B., and Sutcliffe, J. G. (1987). Primate Brain-specific Cytoplasmic Transcript of the Alu Repeat Family. *Mol. Cel Biol.* 7, 3324–3327. doi:10.1128/mcb.7.9.3324-3327.1987

Weser, S., Gruber, C., Hafner, H. M., Teichmann, M., Roeder, R. G., Seifart, K. H., et al. (2004). Transcription Factor (TF)-like Nuclear Regulator, the 250-kDa Form of *Homo sapiens* TFIIIB″, Is an Essential Component of Human TFIIIC1 Activity. *J. Biol. Chem.* 279, 27022–27029. doi:10.1074/jbc.m312790200

White, R. J. (2008). RNA Polymerases I and III, Non-coding RNAs and Cancer. *Trends Genet.* 24, 622–629. doi:10.1016/j.tig.2008.10.003

Wirth, B. (2000). An Update of the Mutation Spectrum of the Survival Motor Neuron Gene (SMN1) in Autosomal Recessive Spinal Muscular Atrophy (SMA). *Hum. Mutat.* 15, 228–237. doi:10.1002/(sici)1098-1004(200003)15:3<228::aid-humu3>3.0.co;2-9

Wirth, B., Karakaya, M., Kye, M. J., and Mendoza-Ferreira, N. (2020). Twenty-Five Years of Spinal Muscular Atrophy Research: From Phenotype to Genotype to Therapy, and what Comes Next. *Annu. Rev. Genom. Hum. Genet.* 21, 231–261. doi:10.1146/annurev-genom-102319-103602

Wolf, N., Harting, I., Innes, A., Patzer, S., Zeitler, P., Schneider, A., et al. (2007). Ataxia, Delayed Dentition and Hypomyelination: a Novel Leukoencephalopathy. *Neuropediatrics* 38, 64–70. doi:10.1055/s-2007-985137

Wolf, N. I., Salomons, G. S., Rodenburg, R. J., Pouwels, P. J. W., Schieving, J. H., Derks, T. G. J., et al. (2014a). Mutations inRARScause Hypomyelination. *Ann. Neurol.* 76, 134–139. doi:10.1002/ana.24167

Wolf, N. I., Vanderver, A., van Spaendonk, R. M. L., Schiffmann, R., Brais, B., Bugiani, M., et al. (2014b). Clinical Spectrum of 4H Leukodystrophy Caused by POLR3A and POLR3B Mutations. *Neurology* 83, 1898–1905. doi:10.1212/wnl.0000000000001002

Wolff, A., Koch, M. J., Benzinger, S., van Waes, H., Wolf, N. I., Boltshauser, E., et al. (2010). Rare Dental Peculiarities Associated with the Hypomyelinating Leukoencephalopathy 4H Syndrome/ADDH. *Pediatr. Dent* 32, 386–392.

Woodward, K. J. (2008). The Molecular and Cellular Defects Underlying Pelizaeus-Merzbacher Disease. *Expert Rev. Mol. Med.* 10, e14. doi:10.1017/s1462399408000677

Xiong, Y., and Steitz, T. (2006). A story with a Good Ending: tRNA 3′-end Maturation by CCA-Adding Enzymes. *Curr. Opin. Struct. Biol.* 16, 12–17. doi:10.1016/j.sbi.2005.12.001

Yang, J., Smith, D. K., Ni, H., Wu, K., Huang, D., Pan, S., et al. (2020). SOX4-mediated Repression of Specific tRNAs Inhibits Proliferation of Human Glioblastoma Cells. *Proc. Natl. Acad. Sci. USA* 117, 5782–5790. doi:10.1073/pnas.1920200117

Yao, P., and Fox, P. L. (2020). Aminoacyl-tRNA Synthetases in Cell Signaling. *Enzymes* 48, 243–275. doi:10.1016/bs.enz.2020.04.002

Yee, N. S., Gong, W., Huang, Y., Lorent, K., Dolan, A. C., Maraia, R. J., et al. (2007). Mutation of RNA Pol III Subunit Rpc2/polr3b Leads to Deficiency of Subunit Rpc11 and Disrupts Zebrafish Digestive Development. *Plos Biol.* 5, e312. doi:10.1371/journal.pbio.0050312

Yeganeh, M., and Hernandez, N. (2020). RNA Polymerase III Transcription as a Disease Factor. *Genes Dev.* 34, 865–882. doi:10.1101/gad.333989.119

Zalfa, F., Giorgi, M., Primerano, B., Moro, A., Di Penta, A., Reis, S., et al. (2003). The Fragile X Syndrome Protein FMRP Associates with BC1 RNA and Regulates the Translation of Specific mRNAs at Synapses. *Cell* 112, 317–327. doi:10.1016/s0092-8674(03)00079-5

Zerboni, L., Sen, N., Oliver, S. L., and Arvin, A. M. (2014). Molecular Mechanisms of Varicella Zoster Virus Pathogenesis. *Nat. Rev. Microbiol.* 12, 197–210. doi:10.1038/nrmicro3215

Zheng, G., Qin, Y., Clark, W. C., Dai, Q., Yi, C., He, C., et al. (2015). Efficient and Quantitative High-Throughput tRNA Sequencing. *Nat. Methods* 12, 835–837. doi:10.1038/nmeth.3478

Zhong, J., Chuang, S.-C., Bianchi, R., Zhao, W., Lee, H., Fenton, A. A., et al. (2009). BC1 Regulation of Metabotropic Glutamate Receptor-Mediated Neuronal Excitability. *J. Neurosci.* 29, 9977–9986. doi:10.1523/jneurosci.3893-08.2009

Zhou, M., and Law, J. A. (2015). RNA Pol IV and V in Gene Silencing: Rebel Polymerases Evolving Away from Pol II's Rules. *Curr. Opin. Plant Biol.* 27, 154–164. doi:10.1016/j.pbi.2015.07.005

Zwieb, C. (1985). The Secondary Structure of the 7SL RNA in the Signal Recognition Particle: Functional Implications. *Nucl. Acids Res.* 13, 6105–6124. doi:10.1093/nar/13.17.6105

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

MINI REVIEW
published: 02 August 2021
doi: 10.3389/fmolb.2021.691636

Coupling Between Cell Cycle Progression and the Nuclear RNA Polymerases System

Irene Delgado-Román[1,2] *and Mari Cruz Muñoz-Centeno*[1,2]*

[1]*Instituto de Biomedicina de Sevilla, Universidad de Sevilla-CSIC-Hospital Universitario V. Del Rocío, Seville, Spain,*
[2]*Departamento de Genética, Facultad de Biología, Universidad de Sevilla, Seville, Spain*

Eukaryotic life is possible due to the multitude of complex and precise phenomena that take place in the cell. Essential processes like gene transcription, mRNA translation, cell growth, and proliferation, or membrane traffic, among many others, are strictly regulated to ensure functional success. Such systems or vital processes do not work and adjusts independently of each other. It is required to ensure coordination among them which requires communication, or crosstalk, between their different elements through the establishment of complex regulatory networks. Distortion of this coordination affects, not only the specific processes involved, but also the whole cell fate. However, the connection between some systems and cell fate, is not yet very well understood and opens lots of interesting questions. In this review, we focus on the coordination between the function of the three nuclear RNA polymerases and cell cycle progression. Although we mainly focus on the model organism *Saccharomyces cerevisiae*, different aspects and similarities in higher eukaryotes are also addressed. We will first focus on how the different phases of the cell cycle affect the RNA polymerases activity and then how RNA polymerases status impacts on cell cycle. A good example of how RNA polymerases functions impact on cell cycle is the ribosome biogenesis process, which needs the coordinated and balanced production of mRNAs and rRNAs synthesized by the three eukaryotic RNA polymerases. Distortions of this balance generates ribosome biogenesis alterations that can impact cell cycle progression. We also pay attention to those cases where specific cell cycle defects generate in response to repressed synthesis of ribosomal proteins or RNA polymerases assembly defects.

Keywords: RNA polymerases I, II and III, RNA polymerases assembly, cell cycle progression, regulatory networks, *Sacharomyces cerevisiae*

OPEN ACCESS

Edited by:
Olga Calvo,
Universidad de Salamanca, Spain

Reviewed by:
Małgorzata Cieśla,
Institute of Biochemistry and Biophysics (PAN), Poland
Isabel Chillón,
European Molecular Biology Laboratory Grenoble, France

***Correspondence:**
Mari Cruz Muñoz-Centeno
mcmunoz@us.es

Specialty section:
This article was submitted to Protein and RNA Networks, a section of the journal Frontiers in Molecular Biosciences

Received: *06 April 2021*
Accepted: *28 June 2021*
Published: *02 August 2021*

Citation:
Delgado-Román I and Muñoz-Centeno MC (2021) Coupling Between Cell Cycle Progression and the Nuclear RNA Polymerases System. Front. Mol. Biosci. 8:691636. doi: 10.3389/fmolb.2021.691636

INTRODUCTION

The eukaryotic cell cycle is controlled by a regulatory network, whose general features are conserved from yeast to humans (Lubischer, 2007). It proceeds through firmly regulated transitions to ensure that specific events take place in a correct and organized manner. This, in turn, ensures viability and the correct transmission of genetic information (Haase and Wittenberg, 2014). A fundamental element of cell cycle regulation consists of arrests at particular steps to guarantee the completion of a previous cell cycle event, to repair cellular or DNA damage, or to resolve a challenging situation. Accordingly, eukaryotic cell cycle regulation integrates a huge multitude of internal and external signals to optimize survival. Failures in these processes reduce cell survival and, in higher metazoans,

lead to cancer, and other diseases (Moriel-Carretero et al., 2019; She et al., 2019; Klemm et al., 2020; Lai et al., 2020; Matellán and Monje-Casas, 2020; Niwa, 2020).

RNA synthesis in the eukaryotic nucleus is carried out by three multisubunit complexes. RNA polymerase II (RNA pol II) transcribes the vast majority of genes, including all protein coding and many other non-coding RNAs (ncRNAs) such as snRNAs, miRNAs, and snoRNAs. RNA polymerase I (RNA pol I) transcribes ribosomal RNAs (rRNA) as a single polycistronic gene: rRNA 35-47S, which is processed into 3 mature rRNAs: 28S (25S in yeast), 18S and 5.8S. This gene appears repeatedly in all eukaryotes with hundreds of copies arranged in tandem. RNA polymerase III (RNA pol III) transcribes an intermediate number of small, non-coding genes (150–400 different), including 5S rRNA and tRNAs (Chan and Lowe, 2016). RNA pol I transcription accounts for almost 60% of global transcription and RNA pol III for around 25%. Of the latter, the 5S rRNA constitutes between 10–15%; and the rest, mostly corresponds to tRNAs. Finally, RNA pol II transcription corresponds to approximately 15% of the total. An important part of this corresponds to RNAs that encode ribosomal proteins (Warner, 1999; Pelechano et al., 2010).

The connection between this transcriptional network and cell cycle progression, can be divided into two different aspects with different levels of knowledge. Regarding what we can call "better known word of the RNA polymerases and the cell cycle," lot of information has been generated describing the dramatic reorganization of gene expression that takes place through the cell cycle. Nearly 20% of *S. cerevisiae* yeast genome is transcribed periodically during each cell division cycle. Abundant information is available on the waves of genes expression associated to the different phases (G1, S, G2/M, M/G1), the complex regulatory connection between them, and on the technological approaches to study this phenomenon (Haase and Wittenberg, 2014). Obviously, in this better-known world, we can understand how transcription impairment of specific genes can disturb the normal cell cycle progression. At this level, RNA pol II has a relevant and direct role on cell cycle regulation (Hartwell et al., 1973; Bähler, 2005; Nurse, 2020).

In this review, we focus in the "lesser known world of the RNA polymerases and the cell cycle." During years, there has been an increase in the knowledge of connections between complex regulatory networks as the transcriptional machinery and cell cycle progression. General changes on transcription levels depending on the cell cycle phases has been known for over decades (Gottesfeld and Forbes, 1997). Different biochemical events underlying this coupled regulation have been elucidated. Here we will focus on mechanisms affecting the three nuclear RNA polymerases. Here we also address this crosstalk between cell cycle and RNA polymerases in the opposite sense, that is, how the status and function of nuclear RNA polymerases can affect cell cycle progression, a much lesser known aspect. It is important to highlight that this interplay coordinates different aspects of the overall status of the three polymerases system with cell cycle progression. In this sense, we review how cell cycle regulation is affected by the balance between the three RNA polymerases products and, secondly, by RNA polymerases assembly. Finally, we also analyze the parallelism between these regulatory interplays in yeast and metazoan, suggesting that it could exist a general control strategy extended throughout eukaryotes.

CELL CYCLE PHASES IMPACTS ON RNA POLYMERASES FUNCTION

Since several years, it is well known that transcription activity in eukaryotes is affected by cell cycle phases. Thus, transcription is repressed during mitosis and highly active in interphase (G1, S, and G2). This mitotic repression has been observed *in vivo* for genes transcribed by all three nuclear RNA polymerases. Different mechanisms contribute to mitotic repression as global transcriptional silencing, including dissociation of transcription factors and cofactors from target genes and profound reorganization of chromatin structure (Gottesfeld and Forbes, 1997, and references therein; Taylor, 1960; Marsden and Laemmli, 1979; Martínez-Balbás et al., 1995). We will focus on how cell cycle phases modulate transcription affecting the basal transcription machinery (RNA pol I, II, and II) in *S. cerevisiae* although some aspects in higher eukaryotes will also be addressed.

RNA Pol II Transcribing Through the Cell Phases

Early works, interestingly described a cell cycle arrest for some RNA pol II mutants. Thus, mutations in the largest RNA pol II subunit, Rpb1, impaired cell cycle progression in budding yeast *S. cerevisiae* (Drebot et al., 1993), fission yeast *Schizosaccharomyces pombe* (Sugaya et al., 1998) and mammalian cells (Sugaya et al., 2001). RNA pol II activity is regulated during the cell cycle by changes in the phosphorylation status of the carboxyl-terminal domain (CTD) of its largest subunit Rpb1 both in yeast and mammal cells (Bregman et al., 2000; Oelgeschläger, 2002; Chymkowitch and Enserink, 2013). The Rpb1 CTD contains 26 heptapeptide repeats in yeast (Allison et al., 1988) and 52 in mammals (Corden et al., 1985). The direct regulation of CTD phosphorylation serves as a switch to regulate transcription machinery during the cell cycle. In the budding yeast *S. cerevisiae*, early in the transcription cycle, Kin28 phosphorylates the CTD which serves as a mark for recruitment of the mRNA capping system (Rodriguez et al., 2000). Interestingly, and coupling cell cycle to RNA pol II activity, it has been demonstrated that Cdc28 (also called Cdk1, and the main CDK cell cycle regulator in budding yeast) is a CTD kinase sharing a partially redundant role with Kin28 (Chymkowitch et al., 2012; Chymkowitch and Enserink, 2013).

RNA Pol III Transcribing Through the Cell Phases

A tRNA synthesis fluctuation during cell cycle has been described both in mammals and yeast (Scott et al., 2001; Frenkel-Morgenstern et al.,

2012; Chen and Gartenberg, 2014; Herrera et al., 2018). Previous results had proposed a tRNA peak in M phase (Chen and Gartenberg, 2014). However, a more recent research has demonstrated that *tDNA* transcription peaked in S phase. The authors, interestingly, propose that this apparent discrepancy can be explained by the overlapping between the S phase and metaphase in *S. cerevisiae*, concluding that the cell cycle-dependent increase in tDNA transcription occurs in the overlapping time span of late S phase/early metaphase. The same authors demonstrate the regulatory mechanism coupling cell cycle to RNA pol III activity: the S phase cyclin Clb5 recruits Cdc28 (Cdk1) to *tDNA* genes; Cdc28 promotes the recruitment of TFIIIC and stimulates the interaction TFIIIC/TFIIIB which directly increases the dynamics of RNA pol III *in vivo*. Bdp1, a component of the TFIIIB complex, has been proposed as the direct target for Cdc28 (Herrera et al., 2018). Recently, new post-translational modifications of RNA pol III, as sumoylation, has been proposed to be involved in stress response in yeast (Nguéa P et al., 2019). The role of this modifications in cell cycle would also be a very interesting open question.

RNA Pol I Transcribing Through the Cell Phases

Transcription by RNA pol I oscillates during the cell cycle, being repressed during mitosis, recovered during G1 and maximal in S/G2 phases. In mammals, repression during M phase is caused by inactivation of a RNA pol I specific factor (TIF-IB/SL1) by an inhibitory cdc2 mediated phosphorylation (Heix et al., 1998). Then, transcription recovery during G1 is mediated by reactivation of another specific factor, UBF (Klein and Grummt, 1999). In the budding yeast, the locus containing *rDNA* genes, segregate after the rest of the genome, in late anaphase. Only in anaphase, yeast repress RNA pol I transcription by the Cdc14 phosphatase acting on Rpa43 subunit, inducing the dissociation of RNA pol I from the 35S *rDNA* (Clemente-Blanco et al., 2009). More recently in *S. cerevisiae*, it has been demonstrated that Rio1 downregulates RNA pol I in a cell cycle dependent manner through Rpa43 subunit as a target. Moreover, Rio1 promotes rDNA stability to ensure rDNA segregation during anaphase (Iacovella et al., 2015).

IMBALANCE OF RNA POL I, II, AND III PRODUCTS PROVOKES G1 ARREST

Balanced Production of Ribosomal Components Prevents G1 Arrest in Budding Yeast

NTP-depleting drugs, as 6-Azauracil (6AU) and mycophenolic acid (MPA) interfere with transcription elongation *in vivo* by strongly inhibiting inosine monophosfate (IMP) dehydrogenase, a rate-limiting enzyme in the *novo* synthesis in guanine nucleotides (Shaw and Reines, 2000; Shaw et al., 2001). Our studies revealed that *S. cerevisiae* cells accumulate at G1 after NTP-depleting drug treatment. As NTP are substrates for three RNA polymerases, we could clearly establish that NTP depletion differentially impacts the RNA products of the three RNA polymerases: products from RNA pol I and III presented a strong and early reduction after treatment but mRNAs showed a very slight reduction at the same conditions. Thus, NTP-depletion drugs generate a clear imbalance between pre-rRNAs, tRNAs and mRNAs (Gómez-Herreros et al., 2013). Using conditional mutants affecting essential subunits of RNA pol I (Rpa43) or III (Rpc17), where their normal transcripts production (rRNAs or 5S rRNA respectively) decreased but not mRNAs generated from wild type RNA pol II, cells also arrested at G1, indicating that any imbalance in RNA polymerases products negatively impacts G1/S transition (Gómez-Herreros et al., 2013).

Ribosome biogenesis is a highly resource-consuming process and, therefore, involves the tight regulation and balanced synthesis of all its components. This complicated pathway requires the coordinated assembly of rRNAs, synthesized by RNA pol I and III, and ribosomal proteins (r-proteins), whose mRNAs are transcribed by RNA pol II. This coordination is critical for an effective utilization of cell resources and requires a balanced function of the RNA pol I, II, and III transcription activities. Thus, the synthesis of rRNAs and r-proteins are two coordinated pathways that lead to efficient ribosome biogenesis [(Warner, 1999; de la Cruz et al., 2018) and references therein]. Data from mammalian cells also showed a G1 arrest after disturbances in ribosome biogenesis, moreover, a key role of mammalian r-proteins L5 and L11 for this essential response has been demonstrated very well (Sun et al., 2008). L11 and L5 r-proteins assembly to 5S rRNA on pre-60S ribosomal particles in a process mediated by Rrs1 (Miyoshi et al., 2004). These mammalian r-proteins L5 and L11 have been reported to accumulate as free proteins and to induce p53 stabilization and G1 arrest after ribosomal biogenesis stress (Sun et al., 2008; Bursać et al., 2012). Therefore, we proposed that in yeast, the imbalance in the three RNA polymerases transcripts provoked defects in ribosomal biogenesis and generated the accumulation of free r-proteins due to the drop in rRNAs. This ribosomal assembly defect could induce a G1 arrest through the accumulation of free r-proteins. Thus, we demonstrated the accumulation of free L5 r-protein in these conditions, as was the case for mammalian cells. **Figure 1** summarizes the model that has been proposed (Gómez-Herreros et al., 2013). In this model, the balanced activity of the three eukaryotic RNA polymerases (I, II, and III) is a prerequisite for an equimolar production of the different ribosomal components. When this balance is disturbed, the accumulation of free L5 occurs and acts as a signal to arrest cell cycle at G1 (**Figure 1**).

Specific cell cycle defects have been described in response to repressed synthesis of r-proteins. After several hours of repression of r-proteins, systematic analyses of cell cycle progression, cell morphology, and bud site selection were performed after repression of 54 individual r-proteins genes in *S. cerevisiae*. In this study, most of the repressed genes involved a G1 arrest (nine encoding 60S subunit components and twenty-two encoding r-proteins of the 40S subunit) and only nine repressed genes encoding components of the 60S subunit resulted in a G2/M delay (Thapa et al., 2013). A later work from the same laboratory, explore cell cycle changes during the transition from normal cell cycle to arrest after inhibition of ribosome formation or translation capacity. Both inhibitions are sensed after a short time and the G1 stage was reached. No spindles or mitotic actin rings were visible,

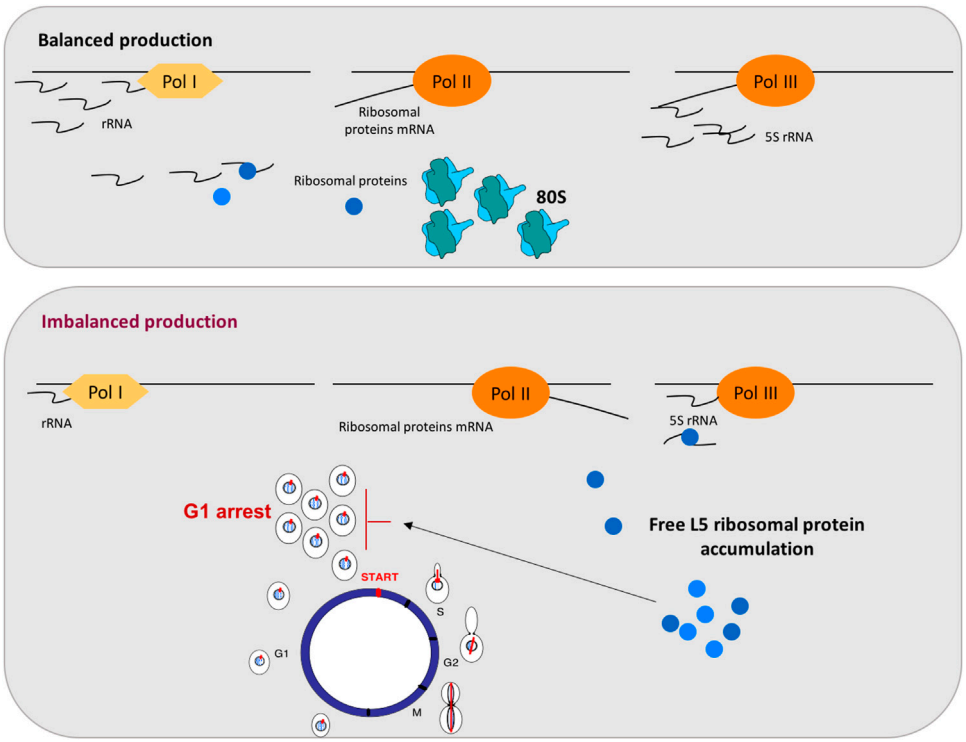


FIGURE 1 | Coupling RNA polymerases production to cell cycle through the free accumulation of the r-protein L5 in yeast. The top panel represents balanced production of ribosomal components: rRNA, r-proteins mRNA and 5S rRNA, transcribed by RNA pol I, II, and III, respectively, are synthesized in the balanced proportion required for correct ribosomal particles assembly. The bottom panel represents situations where this balance is disturbed by a decrease in rRNAs levels but not in r-proteins mRNAs, generating free L5 accumulation and a G1 arrest. As indicated in the figure, rRNA is represented by waves and mRNA by lines.

but membrane ingression was completed in most cells and Ace2, a transcription factor with asymmetric localization to daughter cell nuclei after cell division (Herrero et al., 2020), was localized to daughter cell nuclei demonstrating that, even in the budded arrested cells, G1 phase was reached (Shamsuzzaman et al., 2017). Finally, and very recently, it has been shown that disruption of the assembly of the 40S subunit affected the assembly of the 60S subunit (Rahman et al., 2020). As the r-proteins in each ribosomal subunit are essential only for the assembly of the cognate subunit (Gregory et al., 2019), it was unexpected that disruption of the 40S subunit assembly affected the kinetics of assembly of the 60S subunit, causing accumulation of free/extra-ribosomal 60S L5 (also named uL8) (Rahman et al., 2020). These results indicate that an interaction between the assembly of ribosomal subunits 40S and 60S exists, and that free L5 is a good marker of this generated ribosomal stress.

Nucleolar Stress Induces a G1 Arrest in Mammalian Cells

Nucleolar stress is the term used to described failures in ribosome biogenesis or function that ultimately leads to disruption in cell homeostasis (James et al., 2014). In human cells, mycophenolic acid (MPA) acts as an NTP-depleting drug, as in yeast. Thus, in mammalian, MPA treatment results in both a drastic reduction of pre-rRNA synthesis and the disruption of the nucleolus, causing p53 activation and the subsequent G1 arrest. This treatment provokes the accumulation of free human r-proteins L5 and L11 that bind and inhibit MDM2, the p53 E3 ubiquitin ligase. Therefore, ribosomal imbalance causes MDM2 inhibition, which induces p53 stabilization (Sun et al., 2008; Bursać et al., 2012; Fumagalli et al., 2012; Lo et al., 2012).

Cell responses to the imbalance between RNA polymerases activities, described in yeast and human cells, show very strong analogies: i) in both systems the outcome is a G1 arrest; ii) in both organisms, the G1 arrest responses are mediated by a ribosomal stress; iii) in both scenarios the accumulation of free r-proteins (as L5) is essential for coupling to cell cycle. This strong parallelism between the mechanisms responding to nucleolar stress in yeast and metazoan suggests that it reflects a general control strategy extended throughout eukaryotes. However, a major difference between the two systems exists: yeast does not contain p53 or MDM2. The interpretation of these differences has been extensively discussed and other systems exhibiting nucleolar stress without p53 have been described (James et al., 2014).

DEFFECTS IN RNA POLYMERASE ASSEMBLY PROVOKES ARREST AT G1

As we have just described, the ribosome biogenesis process has been extensively studied [(de la Cruz et al., 2018) and references therein] and its relevant role in interplaying complex networks, as

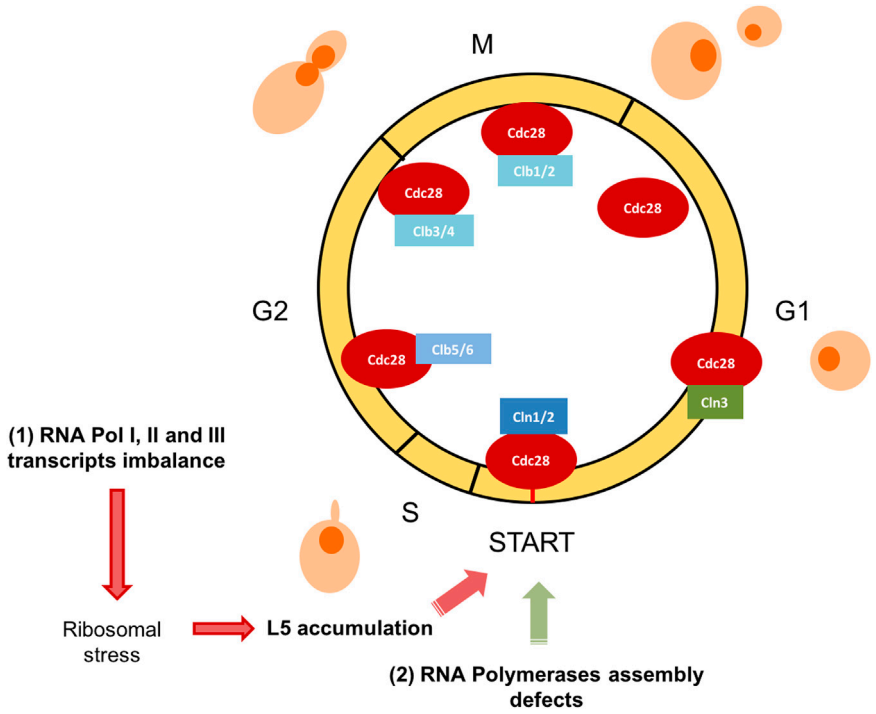


FIGURE 2 | *S. cerevisiae* yeast cell cycle scheme. The RNA polymerases system status impacts on G1/S transition. Graphical representation of the cell cycle phases and main regulators in *S. cerevisiae* are designed. The cyclin-dependent kinase Cdc28 (also known as Cdk1) is sufficient and necessary for cell cycle regulation. Different substrates are phosphorylated based on their association with G1 phase cyclins (Cln1, 2, and 3), S phase cyclins (Clb 5 and 6) or mitotic cyclins (Clb 1, 2, 3, and 4). In this figure, we represent how two abnormal situations involving RNA polymerases, affects specifically G1/S transition: (1) RNA pol I, II, and III transcripts imbalance and (2) RNA polymerases assembly defects.

cell cycle regulation, has been revealed. The assembly of eukaryotic RNA polymerases (RNA pol I, II, and III), is not completely understood although some elements involved in that process has been recently identified. We focus on yeast RNA pol III assembly, as coupling between this assembly process and cell cycle progression has been described (Płonka et al., 2019). The authors had previously isolated and characterized conditional mutants affecting the Rpc128, the second largest RNA pol III subunit. The mutant allele *rpc128-1007* presents a severe defect in RNA pol III assembly as well as an expected reduction in tRNA levels (Cieśla et al., 2007; Cieśla et al., 2015). This conditional mutant, at the restrictive temperature, shows a G1 arrest phenotype which is partially suppressed by overexpression of *RBS1*, the gene encoding a protein involved in RNA pol III assembly (Cieśla et al., 2015). Also, cells lacking Rbs1 showed moderated delay in G1/S transition, indicating that impaired RNA pol III assembly is connected to the cell cycle default. Moreover, the G1 arrest phenotype is not suppressed after inactivation of Maf1, conditions in which elevated levels of tRNAs are produced (Pluta et al., 2001). Thus, they conclude that impairment of RNA pol III complex assembly, and not decreased tRNA transcription levels, is the primary reason for the G1 arrest observed in the rpc128 mutant (Płonka et al., 2019). Very interestingly, Rbs1 was identified as a substrate of cyclin-dependent kinase Cdc28, the main cell cycle regulator in *S. cerevisiae*, in a global proteomic approach (Ubersax et al., 2003).

However, there is evidence that RNA pol III defects can affect cell cycle progression regardless of assembly defects. Thus, mutants affecting the Rpc53 RNA pol III subunit, which has not been described as involved in assembly, leads to a G1 arrest both in yeast (Mann et al., 1992) and mammals (Ittmann et al., 1993). Moreover, depletion of *RPC17* (encoding another RNA pol III subunit), also led to a delay in the G1 phase of the cell cycle (Gómez-Herreros et al., 2013) but, interestingly, *RBS1* overexpression did not overcome G1 arrest (Płonka et al., 2019). These results indicate that G1 arrest coupled to defects in RNA Pol III can be mediated by different regulatory inputs.

CONCLUSIONS, APPLICATIONS AND OPEN QUESTIONS

In this work, we have revisited some aspects of the crosstalk between cell cycle progression and RNA polymerases function. We have focused on those situations where the cell cycle defect is not mediated by the limiting transcription of a specific gene, but those situations where the signal for the cell cycle regulation is the consequence of impaired activity of RNA polymerases or this activity is modulated by the cell cycle phase. First, we have revisited how the three RNA polymerases modulates their

transcription capacity by cell cycle. Then, we have discussed two models in yeast. The first one, when cell cycle arrest is generated by an imbalanced production of RNA pol I, II, and III, which induces an imbalance in ribosomal components and the accumulation of the free r-protein L5 (**Figure 1**). Secondly, when a defect in RNA polymerases assembly is sensed and cell cycle arrested. In both cases, cells arrest at G1, indicating that yeast cells are able to detect internal signals, derived from the activity of the transcriptional machinery. These signals can impact the dynamics of START, the main regulatory event that takes place towards the end of G1 and involves an extensive transcriptional program (Costanzo et al., 2004; de Bruin et al., 2004; Haase and Wittenberg, 2014). It is a very attractive concept that complex processes like gene transcription and ribosomal biogenesis are coupled and sensed to take decisions at START (**Figure 2**).

We have also highlighted that the surveillance mechanism that couples balanced production of yeast ribosomal components and cell cycle, resembles the p53-dependent nucleolar stress checkpoint described in human cells, which indicates that this is a general control strategy extended throughout eukaryotes. In human cells, the molecular components of the regulatory pathway are well known. Clinicians uses the induction of nucleolar stress in cancer cells as an anti-cancer therapy. Moreover, selective inhibition of ribosomal gene transcription in the nucleolus has been shown to be an effective therapeutic strategy to promote cancer-specific activation of p53 (Bywater et al., 2012; Hein et al., 2013; James et al., 2014; Woods et al., 2015; Carotenuto et al., 2019).

Relevant questions remain to be answered in the yeast regulatory systems presented in this work. First, it would be interesting to figure out if all G1 arrest phenomena induced by different defects in RNA polymerases are mediated by the ribosomal stress. Finally, it would be extremely challenging to elucidate the molecular elements that connect the signals (imbalanced production of ribosomal components or defects in assembly) to the G1 arrest. The different elements that participate in the G1/S transition regulatory network, are good candidates. This knowledge would have a relevant translational potential as more than 50% of human cancers lack functional p53. Identification of new p53-independent response pathways could potentially reveal new therapy strategies for p53-defective cancer.

In summary, only understanding both regulatory aspects of this crosstalk, how cell cycle modulates transcription and *viceversa*, a precise knowledge of this complex regulatory interplay will be achieved with a huge translational potential that it has already begun promisingly.

AUTHOR CONTRIBUTIONS

ID contributed to discuss, write and review the manuscript, and drew the figures. MM designed and contributed to write the draft and the final version of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work has been supported by grants from the Ministerio de Ciencia e Innovación-Agencia Estatal de Investigación (BFU2016-77728-C3-1-P to M.C.M.C.), European Union funds (FEDER) and the Andalusian Govermment (BIO271 to S.C. and US-1256285 to M.C.M.C).

ACKNOWLEDGMENTS

We thank all the people of the IBiS (Instituto de Biomedicina de Sevilla) Gene Expression lab for helpful discussion. We also thank the Genetics Department from Universidad de Sevilla for support to the authors.

REFERENCES

Allison, L. A., Wong, J. K., Fitzpatrick, V. D., Moyle, M., and Ingles, C. J. (1988). The C-Terminal Domain of the Largest Subunit of RNA Polymerase II of *Saccharomyces cerevisiae*, *Drosophila melanogaster*, and Mammals: a Conserved Structure with an Essential Function. *Mol. Cel Biol.* 8 (1), 321–329. doi:10.1128/mcb.8.1.321-329.1988

Bähler, J. (2005). Cell-cycle Control of Gene Expression in Budding and Fission Yeast. *Annu. Rev. Genet.* 39, 69–94. doi:10.1146/annurev.genet.39.110304.095808

Bregman, D. B., Pestell, R. G., and Kidd, V. J. (2000). Cell Cycle Regulation and RNA Polymerase II. *Front. Biosci.* 5, D244–D257. doi:10.2741/bregman10.2741/a508

Bursac, S., Brdovcak, M. C., Pfannkuchen, M., Orsolic, I., Golomb, L., Zhu, Y., et al. (2012). Mutual protection of Ribosomal Proteins L5 and L11 from Degradation Is Essential for P53 Activation upon Ribosomal Biogenesis Stress. *Proc. Natl. Acad. Sci.* 109 (50), 20467–20472. doi:10.1073/pnas.1218535109

Bywater, M. J., Poortinga, G., Sanij, E., Hein, N., Peck, A., Cullinane, C., et al. (2012). Inhibition of RNA Polymerase I as a Therapeutic Strategy to Promote Cancer-specific Activation of P53. *Cancer Cell* 22 (1), 51–65. doi:10.1016/j.ccr.2012.05.019

Carotenuto, P., Pecoraro, A., Palma, G., Russo, G., and Russo, A. (2019). Therapeutic Approaches Targeting Nucleolus in Cancer. *Cells* 8 (9), 1090. doi:10.3390/cells8091090

Chen, M., and Gartenberg, M. R. (2014). Coordination of tRNA Transcription with export at Nuclear Pore Complexes in Budding Yeast. *Genes Develop.* 28 (9), 959–970. doi:10.1101/gad.236729.113

Chymkowitch, P., Eldholm, V., Lorenz, S., Zimmermann, C., Lindvall, J. M., Bjoras, M., et al. (2012). Cdc28 Kinase Activity Regulates the Basal Transcription Machinery at a Subset of Genes. *Proc. Natl. Acad. Sci.* 109 (26), 10450–10455. doi:10.1073/pnas.1200067109

Chymkowitch, P., and Enserink, J. M. (2013). The Cell Cycle Rallies the Transcription Cycle. *Transcription* 4 (1), 3–6. doi:10.4161/trns.22456

Cieśla, M., Makała, E., Płonka, M., Bazan, R., Gewartowski, K., Dziembowski, A., et al. (2015). Rbs1, a New Protein Implicated in RNA Polymerase III Biogenesis in Yeast *Saccharomyces cerevisiae*. *Mol. Cel Biol.* 35 (7), 1169–1181. doi:10.1128/MCB.01230-14

Cieśla, M., Towpik, J., Graczyk, D., Oficjalska-Pham, D., Harismendy, O., Suleau, A., et al. (2007). Maf1 Is Involved in Coupling Carbon Metabolism to RNA Polymerase III Transcription. *Mol. Cel Biol.* 27 (21), 7693–7702. doi:10.1128/MCB.01051-07

Clemente-Blanco, A., Mayán-Santos, M., Schneider, D. A., Machín, F., Jarmuz, A., Tschochner, H., et al. (2009). Cdc14 Inhibits Transcription by RNA Polymerase I during Anaphase. *Nature* 458 (7235), 219–222. doi:10.1038/nature07652

Corden, J. L., Cadena, D. L., Ahearn, J. M., and Dahmus, M. E. (1985). A Unique Structure at the Carboxyl Terminus of the Largest Subunit of Eukaryotic RNA Polymerase II. *Proc. Natl. Acad. Sci.* 82 (23), 7934–7938. doi:10.1073/pnas.82.23.7934

Costanzo, M., Nishikawa, J. L., Tang, X., Millman, J. S., Schub, O., Breitkreuz, K., et al. (2004). CDK Activity Antagonizes Whi5, an Inhibitor of G1/S Transcription in Yeast. *Cell* 117 (7), 899–913. doi:10.1016/j.cell.2004.05.024

de Bruin, R. A. M., McDonald, W. H., Kalashnikova, T. I., Yates, J., and Wittenberg, C. (2004). Cln3 Activates G1-specific Transcription via Phosphorylation of the SBF Bound Repressor Whi5. *Cell* 117 (7), 887–898. doi:10.1016/j.cell.2004.05.025

de la Cruz, J., Gómez-Herreros, F., Rodríguez-Galán, O., Begley, V., de la Cruz Muñoz-Centeno, M., and Chávez, S. (2018). Feedback Regulation of Ribosome Assembly. *Curr. Genet.* 64 (2), 393–404. doi:10.1007/s00294-017-0764-x

Drebot, M. A., Johnston, G. C., Friesen, J. D., and Singer, R. A. (1993). An Impaired RNA Polymerase II Activity in *Saccharomyces cerevisiae* Causes Cell-Cycle Inhibition at START. *Mol. Gen. Genet.* 241 (3-4), 327–334. doi:10.1007/BF00284685

Frenkel-Morgenstern, M., Danon, T., Christian, T., Igarashi, T., Cohen, L., Hou, Y. M., et al. (2012). Genes Adopt Non-optimal Codon Usage to Generate Cell Cycle-dependent Oscillations in Protein Levels. *Mol. Syst. Biol.* 8, 572. doi:10.1038/msb.2012.3

Fumagalli, S., Ivanenkov, V. V., Teng, T., and Thomas, G. (2012). Suprainduction of P53 by Disruption of 40S and 60S Ribosome Biogenesis Leads to the Activation of a Novel G2/M Checkpoint. *Genes Develop.* 26 (10), 1028–1040. doi:10.1101/gad.189951.112

Gómez-Herreros, F., Rodríguez-Galán, O., Morillo-Huesca, M., Maya, D., Arista-Romero, M., de la Cruz, J., et al. (2013). Balanced Production of Ribosome Components Is Required for Proper G1/S Transition in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 288 (44), 31689–31700. doi:10.1074/jbc.M113.500488

Gottesfeld, J., and Forbes, D. J. (1997). Mitotic Repression of the Transcriptional Machinery. *Trends Biochem. Sci.* 22 (6), 197–202. doi:10.1016/s0968-0004(97)01045-1

Gregory, B., Rahman, N., Bommakanti, A., Shamsuzzaman, M., Thapa, M., Lescure, A., et al. (2019). The Small and Large Ribosomal Subunits Depend on Each Other for Stability and Accumulation. *Life Sci. Alliance* 2 (2), e201800150. doi:10.26508/lsa.201800150

Haase, S. B., and Wittenberg, C. (2014). Topology and Control of the Cell-Cycle-Regulated Transcriptional Circuitry. *Genetics* 196 (1), 65–90. doi:10.1534/genetics.113.152595

Hartwell, L. H., Mortimer, R. K., Culotti, J., and Culotti, M. (1973). Genetic Control of the Cell Division Cycle in Yeast: V. Genetic Analysis of Cdc Mutants. *Genetics* 74 (2), 267–286. doi:10.1093/genetics/74.2.267

Hein, N., Hannan, K. M., George, A. J., Sanij, E., and Hannan, R. D. (2013). The Nucleolus: an Emerging Target for Cancer Therapy. *Trends Mol. Med.* 19 (11), 643–654. doi:10.1016/j.molmed.2013.07.005

Heix, J., Vente, A., Voit, R., Budde, A., Michaelidis, T. M., and Grummt, I. (1998). Mitotic Silencing of Human rRNA Synthesis: Inactivation of the Promoter Selectivity Factor SL1 by Cdc2/cyclin B-Mediated Phosphorylation. *EMBO J.* 17 (24), 7373–7381. doi:10.1093/emboj/17.24.7373

Herrera, M. C., Chymkowitch, P., Robertson, J. M., Eriksson, J., Bøe, S. O., Alseth, I., et al. (2018). Cdk1 gates Cell Cycle-dependent tRNA Synthesis by Regulating RNA Polymerase III Activity. *Nucleic Acids Res.* 46 (22), 11698–11711. doi:10.1093/nar/gky846

Herrero, E., Stinus, S., Bellows, E., Berry, L. K., Wood, H., and Thorpe, P. H. (2020). Asymmetric Transcription Factor Partitioning during Yeast Cell Division Requires the FACT Chromatin Remodeler and Cell Cycle Progression. *Genetics* 216 (3), 701–716. doi:10.1534/genetics.120.303439

Iacovella, M. G., Golfieri, C., Massari, L. F., Busnelli, S., Pagliuca, C., Dal Maschio, M., et al. (2015). Rio1 Promotes rDNA Stability and Downregulates RNA Polymerase I to Ensure rDNA Segregation. *Nat. Commun.* 6, 6643. doi:10.1038/ncomms7643

Ittmann, M., Ali, J., Greco, A., and Basilico, C. (1993). The Gene Complementing a Temperature-Sensitive Cell Cycle Mutant of BHK Cells Is the Human Homologue of the Yeast RPC53 Gene, Which Encodes a Subunit of RNA Polymerase C (III). *Cell Growth Differ.* 4 (6), 503–511.

James, A., Wang, Y., Raje, H., Rosby, R., and DiMario, P. (2014). Nucleolar Stress with and without P53. *Nucleus* 5 (5), 402–426. doi:10.4161/nucl.32235

Klein, J., and Grummt, I. (1999). Cell Cycle-dependent Regulation of RNA Polymerase I Transcription: the Nucleolar Transcription Factor UBF Is Inactive in Mitosis and Early G1. *Proc. Natl. Acad. Sci.* 96 (11), 6096–6101. doi:10.1073/pnas.96.11.6096

Klemm, C., Thorpe, P. H., and Ólafsson, G. (2020). Cell-cycle Phospho-Regulation of the Kinetochore. *Curr. Genet.* 67, 177–193. doi:10.1007/s00294-020-01127-2

Lai, L., Shin, G. Y., and Qiu, H. (2020). The Role of Cell Cycle Regulators in Cell Survival-Dual Functions of Cyclin-dependent Kinase 20 and p21Cip1/Waf1. *Int. J. Mol. Sci.* 21 (22), 8504. doi:10.3390/ijms21228504

Lo, D., Dai, M.-S., Sun, X.-X., Zeng, S. X., and Lu, H. (2012). Ubiquitin- and MDM2 E3 Ligase-independent Proteasomal Turnover of Nucleostemin in Response to GTP Depletion. *J. Biol. Chem.* 287 (13), 10013–10020. doi:10.1074/jbc.M111.335141

Lubischer, J. L. (2007). The Cell Cycle, Principles of Control. David O. Morgan. *Integr. Comp. Biol.* 47 (5), 794–795. doi:10.1093/icb/icm066

Mann, C., Micouin, J. Y., Chiannilkulchai, N., Treich, I., Buhler, J. M., and Sentenac, A. (1992). RPC53 Encodes a Subunit of *Saccharomyces cerevisiae* RNA Polymerase C (III) Whose Inactivation Leads to a Predominantly G1 Arrest. *Mol. Cel. Biol.* 12 (10), 4314–4326. doi:10.1128/mcb.12.10.4314

Marsden, M. P. F., and Laemmli, U. K. (1979). Metaphase Chromosome Structure: Evidence for a Radial Loop Model. *Cell* 17 (4), 849–858. doi:10.1016/0092-8674(79)90325-8

Martínez-Balbás, M. A., Dey, A., Rabindran, S. K., Ozato, K., and Wu, C. (1995). Displacement of Sequence-specific Transcription Factors from Mitotic Chromatin. *Cell* 83 (1), 29–38. doi:10.1016/0092-8674(95)90231-7

Matellán, L., and Monje-Casas, F. (2020). Regulation of Mitotic Exit by Cell Cycle Checkpoints: Lessons from *Saccharomyces cerevisiae*. *Genes* 11 (2), 195. doi:10.3390/genes11020195

Miyoshi, K., Shirai, C., Horigome, C., Takenami, K., Kawasaki, J., and Mizuta, K. (2004). Rrs1p, a Ribosomal Protein L11-Binding Protein, Is Required for Nuclear export of the 60S Pre-ribosomal Subunit inSaccharomyces Cerevisiae. *FEBS Lett.* 565 (1-3), 106–110. doi:10.1016/j.febslet.2004.03.087

Moriel-Carretero, M., Pasero, P., and Pardo, B. (2019). DDR Inc., One Business, Two Associates. *Curr. Genet.* 65 (2), 445–451. doi:10.1007/s00294-018-0908-7

Nguéa P, A., Robertson, J., Herrera, M. C., Chymkowitch, P., and Enserink, J. M. (2019). Desumoylation of RNA Polymerase III Lies at the Core of the Sumo Stress Response in Yeast. *J. Biol. Chem.* 294 (49), 18784–18795. doi:10.1074/jbc.RA119.009721

Niwa, M. (2020). A Cell Cycle Checkpoint for the Endoplasmic Reticulum. *Biochim. Biophys. Acta (Bba) - Mol. Cel Res.* 1867 (12), 118825. doi:10.1016/j.bbamcr.2020.118825

Nurse, P. (2020). Fission Yeast Cell Cycle Mutants and the Logic of Eukaryotic Cell Cycle Control. *Mol. Biol. Cell* 31 (26), 2871–2873. doi:10.1091/mbc.E20-10-0623

Oelgeschläger, T. (2002). Regulation of RNA Polymerase II Activity by CTD Phosphorylation and Cell Cycle Control. *J. Cel. Physiol.* 190 (2), 160–169. doi:10.1002/jcp.10058

Pelechano, V., Chávez, S., and Pérez-Ortín, J. E. (2010). A Complete Set of Nascent Transcription Rates for Yeast Genes. *PLoS One* 5 (11), e15442. doi:10.1371/journal.pone.0015442

Pluta, K., Lefebvre, O., Martin, N. C., Smagowicz, W. J., Stanford, D. R., Ellis, S. R., et al. (2001). Maf1p, a Negative Effector of RNA Polymerase III in *Saccharomyces cerevisiae*. *Mol. Cel Biol.* 21 (15), 5031–5040. doi:10.1128/MCB.21.15.5031-5040.2001

Płonka, M., Wawrzycka, D., Wysocki, R., Boguta, M., and Cieśla, M. (2019). Coupling of RNA Polymerase III Assembly to Cell Cycle Progression in *Saccharomyces cerevisiae*. *Cell Cycle* 18 (4), 500–510. doi:10.1080/15384101.2019.1578134

Rahman, N., Shamsuzzaman, M., and Lindahl, L. (2020). Interaction between the Assembly of the Ribosomal Subunits: Disruption of 40S Ribosomal Assembly Causes Accumulation of Extra-ribosomal 60S Ribosomal Protein uL18/L5. *PLoS One* 15 (1), e0222479. doi:10.1371/journal.pone.0222479

Rodriguez, C. R., Cho, E.-J., Keogh, M.-C., Moore, C. L., Greenleaf, A. L., and Buratowski, S. (2000). Kin28, the TFIIH-Associated Carboxy-Terminal Domain Kinase, Facilitates the Recruitment of mRNA Processing Machinery to RNA Polymerase II. *Mol. Cel Biol.* 20 (1), 104–112. doi:10.1128/MCB.20.1.104-112.2000

Scott, P. H., Cairns, C. A., Sutcliffe, J. E., Alzuherri, H. M., McLees, A., Winter, A. G., et al. (2001). Regulation of RNA Polymerase III Transcription during Cell Cycle Entry. *J. Biol. Chem.* 276 (2), 1005–1014. doi:10.1074/jbc.M005417200

Shamsuzzaman, M., Bommakanti, A., Zapinsky, A., Rahman, N., Pascual, C., and Lindahl, L. (2017). Analysis of Cell Cycle Parameters during the Transition from Unhindered Growth to Ribosomal and Translational Stress Conditions. *PLoS One* 12 (10), e0186494. doi:10.1371/journal.pone.0186494

Shaw, R. J., and Reines, D. (2000). *Saccharomyces cerevisiae* Transcription Elongation Mutants Are Defective in PUR5 Induction in Response to Nucleotide Depletion. *Mol. Cel Biol.* 20 (20), 7427–7437. doi:10.1128/mcb.20.20.7427-7437.2000

Shaw, R. J., Wilson, J. L., Smith, K. T., and Reines, D. (2001). Regulation of an IMP Dehydrogenase Gene and its Overexpression in Drug-Sensitive Transcription Elongation Mutants of Yeast. *J. Biol. Chem.* 276 (35), 32905–32916. doi:10.1074/jbc.M105075200

She, Z. Y., Wei, Y. L., Lin, Y., Li, Y. L., and Lu, M. H. (2019). Mechanisms of the Ase1/PRC1/MAP65 Family in central Spindle Assembly. *Biol. Rev.* 94 (6), 2033–2048. doi:10.1111/brv.12547

Sugaya, K., Ajimura, M., Tsuji, H., Morimyo, M., and Mita, K. (1998). Alteration of the Largest Subunit of RNA Polymerase II and its Effect on Chromosome Stability in *Schizosaccharomyces pombe*. *Mol. Gen. Genet.* 258 (3), 279–287. doi:10.1007/s004380050732

Sugaya, K., Sasanuma, S., Cook, P. R., and Mita, K. (2001). A Mutation in the Largest (Catalytic) Subunit of RNA Polymerase II and its Relation to the Arrest of the Cell Cycle in G(1) Phase. *Gene* 274 (1-2), 77–81. doi:10.1016/s0378-1119(01)00615-1

Sun, X.-X., Dai, M.-S., and Lu, H. (2008). Mycophenolic Acid Activation of P53 Requires Ribosomal Proteins L5 and L11. *J. Biol. Chem.* 283 (18), 12387–12392. doi:10.1074/jbc.M801387200

Taylor, J. H. (1960). Nucleic Acid Synthesis in Relation to the Cell Division Cycle. *Ann. N. Y Acad. Sci.* 90, 409–421. doi:10.1111/j.1749-6632.1960.tb23259.x

Thapa, M., Bommakanti, A., Shamsuzzaman, M., Gregory, B., Samsel, L., Zengel, J. M., et al. (2013). Repressed Synthesis of Ribosomal Proteins Generates Protein-specific Cell Cycle and Morphological Phenotypes. *Mol. Biol. Cell* 24 (23), 3620–3633. doi:10.1091/mbc.E13-02-0097

Ubersax, J. A., Woodbury, E. L., Quang, P. N., Paraz, M., Blethrow, J. D., Shah, K., et al. (2003). Targets of the Cyclin-dependent Kinase Cdk1. *Nature* 425 (6960), 859–864. doi:10.1038/nature02062

Warner, J. R. (1999). The Economics of Ribosome Biosynthesis in Yeast. *Trends Biochem. Sci.* 24 (11), 437–440. doi:10.1016/s0968-0004(99)01460-7

Woods, S. J., Hannan, K. M., Pearson, R. B., and Hannan, R. D. (2015). The Nucleolus as a Fundamental Regulator of the P53 Response and a New Target for Cancer Therapy. *Biochim. Biophys. Acta (Bba) Gene Regul. Mech.* 1849 (7), 821–829. doi:10.1016/j.bbagrm.2014.10.007

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

MINI REVIEW
published: 26 October 2021
doi: 10.3389/fmolb.2021.778778

Coupling Between Production of Ribosomal RNA and Maturation: Just at the Beginning

Chaima Azouzi[1], Mariam Jaafar[1†], Christophe Dez[1], Raghida Abou Merhi[2], Annick Lesne[3,4], Anthony K. Henras[1] and Olivier Gadal[1]**

[1]Laboratoire de Biologie Moléculaire, Cellulaire et du Développement (MCD), Centre de Biologie Intégrative (CBI), CNRS, UPS, Université de Toulouse, Toulouse, France, [2]Genomic Stability and Biotherapy (GSBT) Laboratory, Faculty of Sciences, Rafik Hariri Campus, Lebanese University, Beirut, Lebanon, [3]CNRS, Laboratoire de Physique Théorique de la Matière Condensée, LPTMC, Sorbonne Université, Paris, France, [4]Institut de Génétique Moléculaire de Montpellier, IGMM, CNRS, Université Montpellier, Montpellier, France

OPEN ACCESS

Edited by:
Francisco Navarro,
University of Jaén, Spain

Reviewed by:
Kostantin Panov,
Queen's University Belfast, United Kingdom
Tomasz Wojciech Turowski,
Institute of Biochemistry and Biophysics (PAN), Poland

***Correspondence:**
Anthony K. Henras
anthony.henras@univ-tlse3.fr
Olivier Gadal
olivier.gadal@univ-tlse3.fr

†Present address:
Inserm U1052, CNRS UMR5286, Centre de Recherche en Cancérologie de Lyon, Université de Lyon, Université Claude Bernard Lyon 1, Centre Léon Bérard, Lyon, France

Specialty section:
This article was submitted to Protein and RNA Networks, a section of the journal Frontiers in Molecular Biosciences

Received: 17 September 2021
Accepted: 12 October 2021
Published: 26 October 2021

Citation:
Azouzi C, Jaafar M, Dez C, Abou Merhi R, Lesne A, Henras AK and Gadal O (2021) Coupling Between Production of Ribosomal RNA and Maturation: Just at the Beginning. Front. Mol. Biosci. 8:778778. doi: 10.3389/fmolb.2021.778778

Ribosomal RNA (rRNA) production represents the most active transcription in the cell. Synthesis of the large rRNA precursors (35S/47S in yeast/human) is achieved by up to hundreds of RNA polymerase I (Pol I) enzymes simultaneously transcribing a single rRNA gene. In this review, we present recent advances in understanding the coupling between rRNA production and nascent rRNA folding. Mapping of the distribution of Pol I along ribosomal DNA at nucleotide resolution, using either native elongating transcript sequencing (NET-Seq) or crosslinking and analysis of cDNAs (CRAC), revealed frequent Pol I pausing, and CRAC results revealed a direct coupling between pausing and nascent RNA folding. High density of Pol I per gene imposes topological constraints that establish a defined pattern of polymerase distribution along the gene, with a persistent spacing between transcribing enzymes. RNA folding during transcription directly acts as an anti-pausing mechanism, implying that proper folding of the nascent rRNA favors elongation *in vivo*. Defects in co-transcriptional folding of rRNA are likely to induce Pol I pausing. We propose that premature termination of transcription, at defined positions, can control rRNA production *in vivo*.

Keywords: RNA polymerase I (Pol I), ribosomal RNA (rRNA) processing, transcription, termination of transcription, ribosomal RNA (rRNA) genes, RNA folding, premature termination of transcription

SYNTHESIS OF THE 35S PRIMARY TRANSCRIPT BY POL I

Yeast haploid cells contain between 150 and 200 copies of tandemly repeated rRNA genes while the diploid human genome contains around 400 copies. Although present at a high copy number in the genomes, not all rRNA genes are actively transcribed. In budding yeast, only about 50% of the genes on average are transcribed in exponentially growing cells. Each ribosomal gene unit spreads over 9.1 kb of DNA and contains two transcribed regions encoding the 35S pre-rRNA, transcribed by RNA Polymerase I (Pol I), and the 5S rRNA, transcribed by Pol III (**Figure 1A**). These transcribed regions are separated by intergenic spacers (IGSs): IGS1 starts at the transcription termination site of the 35S gene and ends at the 5S rRNA gene terminator and IGS2 corresponds to the region between the 5S rRNA gene promoter and the promoter of the next 35S gene (Nomura, 2001). Pol I transcription accounts for almost 60% of total transcriptional activity in yeast cells (Warner, 1999). This process occurs in the nucleolus

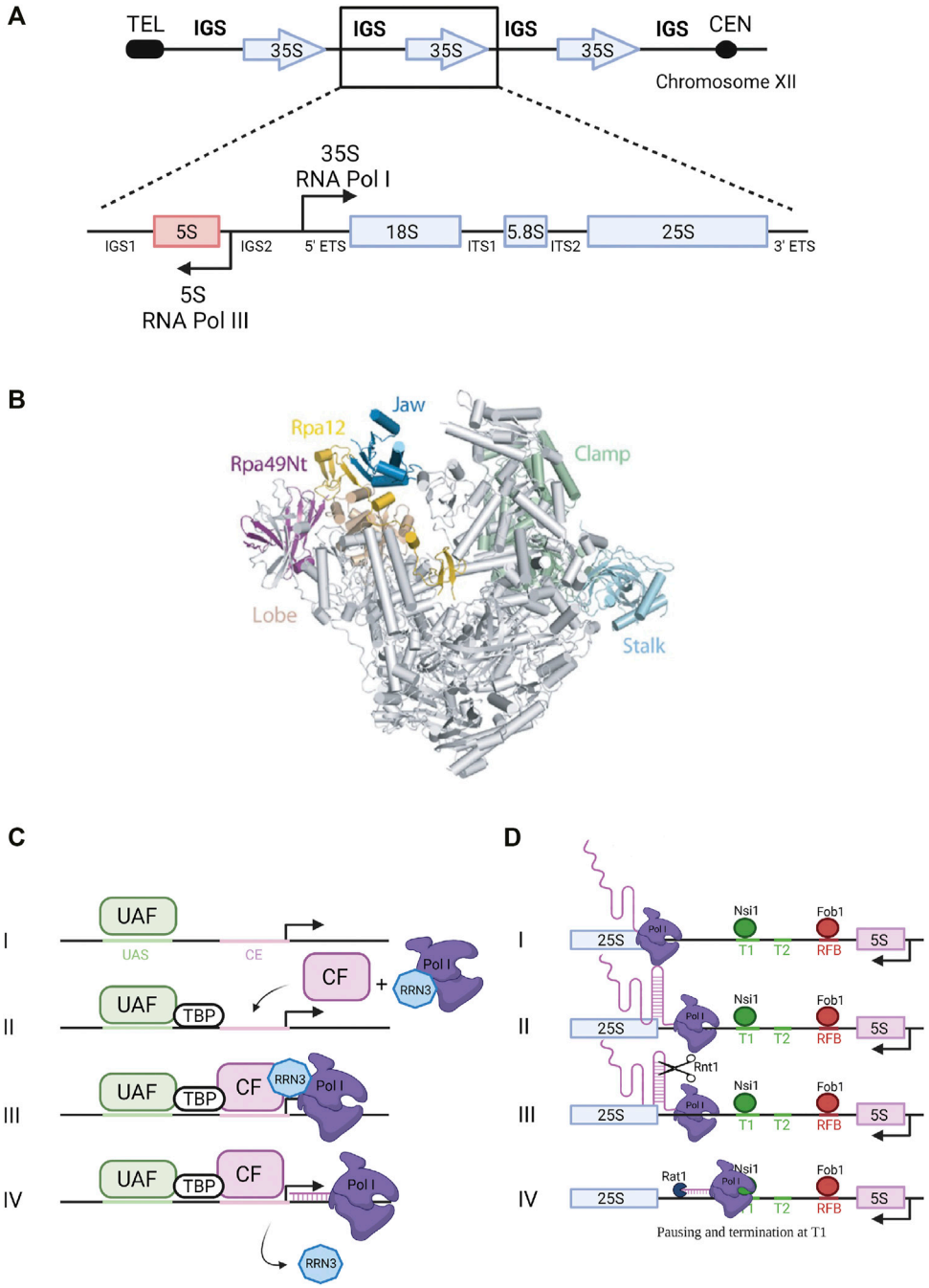


FIGURE 1 | Ribosomal DNA transcription by RNA Pol I. **(A)** Ribosomal DNA. The rDNA repeats (150–200 copies) are located on chromosome XII. A single repeated unit is transcribed by RNA polymerase I (Pol I) to synthesize the 35S primary pre-rRNA transcript, which is then processed to produce the mature 18S, 5.8S and 25S rRNAs (arrow pointing to the right). RNA Polymerase III synthesizes the 5S rRNA (arrow pointing to the left). IGS, intergenic sequence; ETS, external transcribed spacer; ITS, internal transcribed spacer. **(B)** RNA Polymerase I. Pol I 3D structure (Darrière et al., 2019). View of the initially transcribing complex model and its four different subunits - PDB 5W66 (Han et al., 2017). Catalytic amino acids are located in the center of the central cleft. The two main modules are mobile and allow cleft opening and closure, depending of the transcription step. **(C)** Transcription initiation. Composition of Pol I pre-initiation complex (*see* text for details). UAF, Upstream Activating Factor; TBP, TATA-binding protein; CF, Core Factor. **(D)** Transcription termination. Pol I termination mechanisms (*see* text for details).

and results in the synthesis of the 35S pre-rRNA containing the sequences of three of the four rRNAs composing the mature ribosome, the 18S, 5.8S and 25S rRNAs. These sequences are flanked and separated by sequences that are not retained in the mature ribosomes: respectively the 5′ and 3′ external transcribed spacers (5′ ETS and 3′ ETS) and the internal transcribed spacers 1 and 2 (ITS1 and ITS2) (**Figure 1A**). This 35S precursor will be co-transcriptionally packaged into

pre-ribosomal particles that will undergo a complex maturation pathway to generate the mature ribosomal subunits.

TRANSCRIPTION INITIATION AND TERMINATION

Pol I enzyme in yeast is composed of 14 subunits (global molecular weight of 590 kDa) including two large subunits, Rpa190 and Rpa135, jointly forming the active site of the enzyme (Riva et al., 1987) (**Figure 1B**). Crystal structure of yeast *Saccharomyces cerevisiae* Pol I revealed the interactions occurring between its 14 subunits: the two large subunits Rpa190 and Rpa135 organize the enzyme in two modules of similar mass (Engel et al., 2013; Fernández-Tornero et al., 2013). The Pol I-specific subunits whose role during transcription has been partially characterized include Rpa43 and Rpa14 subunits in the stalk, and Rpa34, Rpa49 and Rpa12 subunits associated with the jaw/lobe module (**Figure 1B**).

Formation of preinitiation complex (PIC) is presented in **Figure 1C**. Pol I promoter contains two sequences required for efficient transcription initiation: the upstream activating sequence (UAS) and the core element (CE) (Nomura, 2001; Boukhgalter et al., 2002). Recruitment of the polymerase to the promoter to form the PIC relies on four transcription factors: upstream activating factor (UAF), core factor (CF), TATA-binding protein (TBP) and the Rrn3 transcription factor (Keener et al., 1998). UAF is the first complex to associate with the UAS of the rDNA promoter to initiate PIC assembly (Steffan et al., 1996). TBP was shown to bind to both CF and UAF, thus serving as a bridge to position CF downstream of the UAS. Binding of CF to the CE allows further recruitment of Pol I stably associated with Rrn3 (Aprikian et al., 2001). Rrn3 is a highly conserved transcription factor that associates with the Rpa43-Rpa14 heterodimer of Pol I and interacts with the Rrn6 subunit of the CF. It is therefore a crucial element required for transcription initiation (Peyroche et al., 2000; Aprikian et al., 2001). Transcription begins at the transcription start site (TSS) and Pol I and Rrn3 are released from the PIC upon transcription initiation. Several structural studies gave new insights into Pol I promoter recognition and melting, and more broadly into transcription initiation by yeast Pol I (Blattner et al., 2011; Engel et al., 2013, 2016; Moreno-Morcillo et al., 2014; Neyer et al., 2016; Tafur et al., 2016; Han et al., 2017; Sadian et al., 2017; Smith et al., 2018; Sadian et al., 2019; Tafur et al., 2019; Knutson et al., 2020). These studies will not be detailed here.

Pol I transcription termination involves pausing induced by a terminator protein, leading to dissociation of the polymerase and release of the primary transcript. Paradoxically, termination is not required for rRNA production since nascent transcript is released through the endonucleolytic cleavage by Rnt1 (**Figure 1D**) (Henras et al., 2005). In fission yeast, Reb1 protein interacts with the Rpa12 subunit of Pol I to stimulate termination (Jaiswal et al., 2016). In budding yeast, 90% of Pol I transcription termination occurs at a well-defined primary terminator element (T1) downstream of the 25S rRNA sequence (**Figure 1D**). Transcription termination at this site implicates the DNA-binding factor Nsi1, a Reb1 paralog, which promotes termination upstream of T1 at a T-rich element that likely operates as a polymerase release element (Lang and Reeder, 1993; Merkl et al., 2014; Reiter et al., 2012). In 10% of the cases, Pol I reads through this first terminator and stops at a downstream, "fail-safe" terminator (T2) located around position +250 from the 3′ end of the 25S rRNA sequence (Reeder et al., 1999). Transcription termination on Pol II-transcribed genes was shown to involve the 5′-3′ exoribonuclease Rat1 through a mechanism called "torpedo" (West et al., 2004; Luo et al., 2006; Kim et al., 2004). According to this model, Rat1 binds and degrades the transcript emerging from the polymerase following cleavage and release of the pre-mRNA, and given its high processivity, Rat1 catches up and dissociates Pol II from the DNA template. In the context of Pol I transcription, Rat1 was shown to interact with terminator sequences T1 and T2 and to be required for efficient termination. Its catalytic activity is required for this function since expression of a catalytically inactive mutant of Rat1 ($Rat1_{D235A}$) could not suppress the Pol I termination defect observed in absence of Rat1. The absence of both Rat1 and Fob1, bound to the replication fork barrier (RFB) site (**Figure 1D**), increases polymerase read-through of T2 and the RFB site, indicating that Fob1 is also partly involved in termination (El Hage et al., 2008).

POL I SUBUNITS AND TRANS-ACTING FACTORS INVOLVED IN ELONGATION DYNAMICS

Transcription elongation properties involve in particular three Pol I subunits present on the lobe (**Figure 1B**): Rpa12 and the heterodimer Rpa34/Rpa49 (Liljelund et al., 1992; Nogi et al., 1993; Gadal et al., 1997). In absence of Rpa34/Rpa49, Pol I activity is altered (Huet et al., 1975; Liljelund et al., 1992). Pol I lacking the Rpa34/Rpa49 subunits does not produce RNA to the same extent as a wild-type enzyme (Kuhn et al., 2007; Beckouet et al., 2008; Albert et al., 2011). Furthermore, this heterodimer plays an important role in transcription by improving the recruitment of the Rrn3-Pol I complex to the rDNA and by triggering the release of Rrn3 from elongating Pol I. Indeed, in an *rpa49* deletion strain, Rrn3 is recruited less efficiently at the promoter and fails to dissociate from elongating polymerases following transcription initiation (Beckouet et al., 2008). Interestingly, Rpa49 and Rpa34 are important for nucleolar assembly and formation of a property of actively transcribed rRNA genes called "Pol I caravans" or "Pol I convoys," reflecting a spatial proximity between adjacent polymerases (Albert et al., 2011; Neyer et al., 2016). Rpa12 subunit stabilizes the Rpa49/Rpa34 heterodimer on the polymerase (Van Mullem et al., 2002; Tafur et al., 2019). In the absence of Rpa12, Pol I catalytic properties are affected (Appling et al., 2018; Scull et al., 2021). Furthermore, Pol I transcription through a linear mono-nucleosomal template was shown to be defective in the absence of the lobe-binding subunits (Merkl et al., 2020). Mutations affecting the Rpa135 subunit were also shown to affect transcription elongation. In particular,

mutation of the amino acid at position 784 (rpa135-D784G), suspected to play a role in loading NTP substrates, caused reduced transcription compared to a wild-type Pol I. Calculation of Pol I elongation rate *in vitro* showed that this Rpa135 mutant is ten times slower than the wild-type polymerase (Schneider et al., 2007).

In addition to the role of Pol I subunits in transcription elongation, transcription factor Spt5 in complex with Spt4, was also shown to be required for efficient Pol I transcription (Schneider et al., 2006). Immunoprecipitation and mass spectrometry experiments showed that this complex interacts directly with multiple Pol I subunits (Rpa49, Rpa34, Rpa135 and Rpa190), through the NGN and KOW domains of Spt5 (Schneider et al., 2006). Moreover, Spt5 also associates with the transcription factor Rrn3 and with the 35S rRNA gene (coding region and promoter) (Viktorovskaya et al., 2011). Depletion of Spt4 in yeast results in a temperature-sensitive slow growth phenotype associated with a decreased rRNA synthesis rate as well as a reduced Pol I elongation efficiency, also impacting pre-rRNA processing and ribosome assembly (Schneider et al., 2006). Furthermore, Spt5 mutations suppress the cold-sensitive phenotype of an *rpa49Δ* strain. All these data support a function of the Spt4-Spt5 complex in Pol I transcription elongation, which remains to be understood at the molecular level. Another related protein, Spt6, interacts with the Spt4/Spt5 complex and was also proposed to play a role in Pol I transcription (Swanson and Winston, 1992). Spt6 interacts with Pol I subunit Rpa43 (Beckouët et al., 2011). It was shown that Spt6 associates with rDNA and is required for Pol I transcription since a strain carrying an in-frame deletion allele of *SPT6 (Spt6-1004)* showed reduced Pol I occupancy on the rDNA (Engel et al., 2015). Other factors including Hmo1 also modulate Pol I elongation properties, but the underlying mechanisms remain elusive (Albert et al., 2013; Higashino et al., 2015).

MAPPING POL I POSITION AT NUCLEOTIDE RESOLUTION TO INVESTIGATE POL I ELONGATION *IN VIVO*

In addition to the implication of Pol I subunits and trans-acting factors, Pol I elongation is also regulated by mechanisms intrinsic to the transcription process. Elongation is fundamentally discontinuous, with events of pausing, backtracking and possible premature termination, which remain to be explored. Pol I elongation was studied using the native elongating transcript sequencing (NET-seq) method, based on deep sequencing of the 3′ ends of nascent transcripts associated with the polymerase (Churchman and Weissman, 2011). This study revealed hundreds of positions within rDNA that reproducibly induce pausing (Clarke et al., 2018). Unfortunately, fragments of mature rRNAs co-purifying with Pol I in the NET-seq procedure could introduce bias in the analysis. Turowski and co-workers used the crosslinking and analysis of cDNAs (CRAC) technique to map the position of Pol I on rDNA during elongation. CRAC consists in crosslinking Pol I to its associated nascent rRNAs during elongation *in vivo*, followed by complex purification, reverse transcription of associated rRNAs and sequencing of cDNAs (Turowski et al., 2020). Applied to a population of cells, this method provides a statistical snapshot of the position of transcribing Pol I all along the rDNA unit and allows the determination of areas of the gene in which Pol I is accumulated (**Figure 2A**). It is noteworthy that a high polymerase occupancy reflects a low elongation rate. This CRAC analysis revealed a massive Pol I enrichment in the 5′ end of rRNA genes. Enrichment of polymerases at the 5′ end of rRNA genes was previously observed, but to a much lower extent, using the chromatin spread method developed by Oskar Miller, allowing a direct observation of Pol I in complex along rDNA (Miller and Beatty, 1969; Osheim et al., 2009; French et al., 2003). It was speculated that the high density of polymerases in the 5′ETS region, called "Low Entrainment Region" (LER), results in polymerases moving more slowly (decreasing Pol I elongation rate <20%) and being more closely over the initial 2 kb. As an underlying mechanism, Turowski and collaborators proposed that in the LER, where Pol I is associated with only short nascent transcripts, Pol I molecules are able to rotate freely along DNA grooves during elongation, while they become progressively unable to do so due to viscous drag 2 kb after initiation (**Figure 2B**). Accordingly, polymerase activity in the LER would not generate torsion in DNA, which allows changes in the relative positions of adjacent polymerases. This results in increased freedom for movement, likely increasing the probability of backtracking events, which would explain the accumulation of Pol I in the 5′ region of the genes. The high density of polymerases in the 5′ETS region is also correlated with the fact that major early pre-rRNA assembly events take place on the 5′ region of nascent rRNA (Chaker-Margot et al., 2017).

It is important to note that Pol I translocation is based on Brownian ratchet motion making elongation prone to frequent backtracking and potentially sensitive to quite modest forces (Dangkulwanich et al., 2013). Co-transcriptional folding of the nascent rRNA has direct consequences on elongation by preventing backtracking, thereby favoring productive elongation (Turowski et al., 2020). Any co-transcriptional association with the nascent transcript of trans-acting factors (UTPs, snoRNPs) should have the same stimulatory effect on transcription. With up to 200 transcribing Pol I per rRNA gene, each enzyme is influenced by its neighbors along the template directly through steric constraints. Indirectly, away of the LER predicted to occur only in the first 2 kb, torsional constraints on DNA plays a major role (torsional coupling):

- When the rotation around DNA of the transcribing polymerases is prevented by viscous drag due to the size and structure of nascent rRNA, elongation can be described within the twin-supercoiled domain model: DNA screws into the polymerase and experiences positive supercoiling downstream and negative supercoiling upstream (Liu and Wang, 1987).

- When all polymerases transcribe at the same rate, the negative DNA supercoiling created in the wake of one translocating polymerase is rapidly cancelled out by the positive DNA supercoiling created in front of the following one. The torsional stress between polymerases is alleviated and

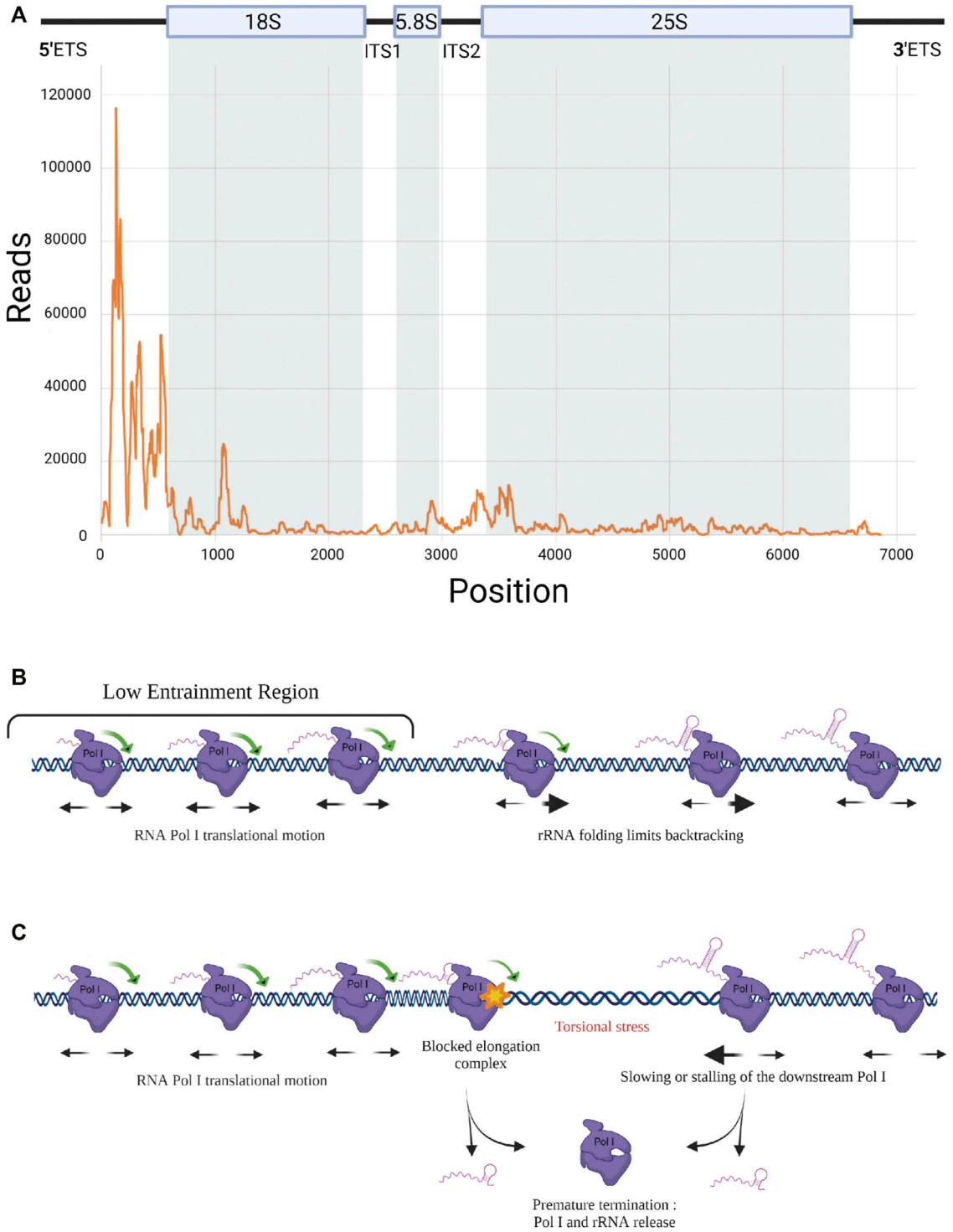


FIGURE 2 | RNA Pol I elongation dynamics. **(A)** RNA Pol I distribution along rDNA template. Rpa135-CRAC results showing strong Pol I accumulation at 5' end of the rRNA gene (Turowski et al., 2020). **(B)** Elongation dynamics in the Low Entrainment Region. Schematic representation of Pol I elongation dynamics in the LER (Turowski et al., 2020). Associated with short nascent transcripts, Pol I can easily rotate around rDNA in the 5' region, leading to free translocation and a higher rate of backtracking. Beyond the LER, viscous drag limits the rotation of the pre-RNA/Pol I complex around DNA. **(C)** Premature termination. Model including the propensity of the elongation complex to dissociate and release rRNA, leading to premature termination. When a Pol I is stalled, a torsional stress occurs that could be resolved by a premature transcription termination (PTT) event.

a fast and processive collective translocation is allowed, leading to polymerase convoys (Lesne et al., 2018; Kim et al., 2019). Therefore, all polymerases in convoys translocate at the same rate, their spacing remains constant (**Figure 2C**). Any change in the relative positions of transcribing polymerases generates torsional stress, which will quickly exceed the low stalling force of the polymerases (Ma et al., 2013; Heberling et al., 2016; Tantale et al., 2016). Any local modification of Pol I spacing within rDNA modifies DNA supercoiling, and the associated increase of local torsional energy generates an apparent force sufficiently strong to restore the initial distance between the polymerases and ensures the cohesion of the convoy (Lesne et al., 2018). Deletion or rapid depletion of topoisomerase I, results in defective rRNA synthesis (El Hage et al., 2010; Albert et al., 2019), highlighting the importance of resolving DNA supercoiling (downstream and upstream of each convoy) for efficient Pol I transcription elongation.

However, this cooperative long-distance group behavior may also induce antagonist effects on elongation. It was observed that promoter shut-off reduces the apparent elongation rate of the engaged polymerases, which is associated with a significant increase in premature termination (Kim et al., 2019). It is rational to suppose that the same effect occurs when elongating Pol I gets stalled on rDNA, thus leading to accumulation of negative torsional stress in the wake of the downstream Pol I (*i.e.* the nearby Pol I farthest from the promoter). Pol I stalling is known to increase premature termination of the paused Pol I and possibly also of the downstream polymerases (**Figure 2C**). Such a phenomenon was previously described as premature termination of transcription (PTT) for Pol II (Kamieniarz-Gdula and Proudfoot, 2019).

These premature termination events could also potentially explain the 5′ bias observed in the Pol I CRAC profile. Pol II is known to undergo a transition from initiation to elongation states that is associated with changes of the phosphorylation status of the C-terminal domain (CTD) of the largest Pol II subunit (Milligan et al., 2016). It is possible that Pol I undergoes a similar transition, the 5′ accumulation bias reflecting a region in which the polymerase has an elevated probability to terminate prematurely. However, consideration of premature termination in the model of Turowski and collaborators, even though it recapitulated the overall profile, reduced by 30% the total number of polymerases per transcription unit, which falls below the number of Pol I molecules per rDNA observed using Miller spreads (Turowski et al., 2020). Nevertheless, premature termination of Pol I cannot be excluded and could, at least partially, play a role in establishing the 5′ bias.

WORKING HYPOTHESIS: POL I PROCESSIVITY AND PREMATURE TERMINATION

In order to better understand transcription regulation, Pol I mutants are of particular interest. We have recently identified in a genetic screen a super-active Pol I mutant, bearing a single substitution on the second largest subunit: Rpa135-F301S allele, hereafter named SuperPol I. This mutant induces an increase of rRNA production in yeast (Darrière et al., 2019). The mechanism leading to this increased rRNA production is not well understood. We proposed that this mutation alleviates an intrinsic repressive element of the polymerase, leading to increased processivity during elongation, *i.e.* the ability of Pol I to carry out continuous RNA synthesis on the DNA template without premature termination. This hypothesis is based on several experimental evidences. First, Miller spreads showed that the amounts of Pol I engaged in transcription are comparable in wild-type (WT) and mutant cells, meaning that the increased production of rRNA is not due to a major enhancement of Pol I initiation rate (Darrière et al., 2019). Moreover, *in vitro* promoter-dependent transcription assays confirmed that transcription initiation rate is similar between WT and SuperPol I. On the other hand, a tailed template assay, measuring elongation rate *in vitro*, revealed an increased rRNA production by the SuperPol I, likely due to a higher processivity (Darrière et al., 2019). Taken together, these elements suggest that the Rpa135-F301S mutation induces modifications in the elongation process, and more precisely on processivity. Premature termination directly affects processivity and likely influences Pol I distribution along the DNA template. Importantly, premature termination can not be measured by CRAC, which relies on detection of rRNA still bound to Pol I. To demonstrate the occurrence of premature termination events, defined as a dissociation of the elongation complex and release of the nascent rRNA, it will be necessary to correlate Pol I complex stalling with the production of abortive rRNAs. This could be achieved by combining Pol I CRAC data, highlighting precise pause sites, with a mapping of the corresponding abortive transcripts. Detection of rRNA species resulting from abortive transcription in differential amounts in cells expressing the SuperPol I or WT polymerase should allow to better understand what features of elongating Pol I lead to premature termination. The increased processivity of the SuperPol I mutant could likely be the consequence of a lower occurrence of premature termination, *i.e.* a lower production of abortive rRNAs.

CONCLUSION AND PERSPECTIVES

Methods allowing to map at nucleotide resolution Pol I pausing sites during elongation revealed a key interplay between RNA folding and elongation rate: formation of rRNA secondary structures prevents backtracking, hence enhances elongation rate. With a large amount of co-transcriptional folding of rRNA, we are now able to study how processing events are affecting Pol I elongation rate. So far limited to budding yeast, there is no doubt that some Pol I regulatory mechanisms are evolutionary conserved, as Pol I elongation rate is limiting for rRNA synthesis in metazoan cells (Hung et al., 2017). The understanding of the precise mechanisms of Pol I

transcription and the implication of each inherent elongation feature opens wide prospects on health-related areas of research, particularly to understand a large number of genetic diseases collectively called ribosomopathies. Pol I inhibition used in cancer therapy these recent years will also benefit from such mechanistic breakthroughs (Sulima et al., 2019; Ferreira et al., 2020; Kampen et al., 2020).

AUTHOR CONTRIBUTIONS

OG and AH conceived the manuscript. CA and MJ wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

OG has been supported by the "Fondation ARC pour la Recherche sur le Cancer," by the "Ligue contre le cancer," and by the "Centre de Biologie Integrative (CBI)" grant. CA benefited from PhD fellowship from "Ligue contre le cancer." AH has been supported by ANR grant "CHROMRIB" (ANR-16-CE12-0027-01).

ACKNOWLEDGMENTS

We would like to thank all of members of Gadal and Henry-Henras groups for helpful discussions.

REFERENCES

Albert, B., Colleran, C., Léger-Silvestre, I., Berger, A. B., Dez, C., Normand, C., et al. (2013). Structure-function Analysis of Hmo1 Unveils an Ancestral Organization of HMG-Box Factors Involved in Ribosomal DNA Transcription from Yeast to Human. *Nucleic Acids Res.* 41, 10135–10149. doi:10.1093/nar/gkt770

Albert, B., Kos-Braun, I. C., Henras, A., Dez, C., Rueda, M. P., Zhang, X., et al. (2019). A Ribosome Assembly Stress Response Regulates Transcription to Maintain Proteome Homeostasis. bioRxiv [Preprint]. Available at: https://www.biorxiv.org/content/10.1101/512665v1.full. (Accessed January 6, 2019) doi:10.1101/512665

Albert, B., Léger-Silvestre, I., Normand, C., Ostermaier, M. K., Pérez-Fernández, J., Panov, K. I., et al. (2011). RNA Polymerase I-specific Subunits Promote Polymerase Clustering to Enhance the rRNA Gene Transcription Cycle. *J. Cel Biol.* 192, 277–293. doi:10.1083/jcb.201006040

Appling, F. D., Scull, C. E., Lucius, A. L., and Schneider, D. A. (2018). The A12.2 Subunit Is an Intrinsic Destabilizer of the RNA Polymerase I Elongation Complex. *Biophysical J.* 114, 2507–2515. doi:10.1016/j.bpj.2018.04.015

Aprikian, P., Moorefield, B., and Reeder, R. H. (2001). New Model for the Yeast RNA Polymerase I Transcription Cycle. *Mol. Cel. Biol.* 21, 4847–4855. doi:10.1128/MCB.21.15.4847-4855.2001

Beckouet, F., Labarre-Mariotte, S., Albert, B., Imazawa, Y., Werner, M., Gadal, O., et al. (2008). Two RNA Polymerase I Subunits Control the Binding and Release of Rrn3 during Transcription. *Mol. Cel. Biol.* 28, 1596–1605. doi:10.1128/MCB.01464-07

Beckouët, F., Mariotte-Labarre, S., Peyroche, G., Nogi, Y., and Thuriaux, P. (2011). Rpa43 and its Partners in the Yeast RNA Polymerase I Transcription Complex. *FEBS Lett.* 585, 3355–3359. doi:10.1016/j.febslet.2011.09.011

Blattner, C., Jennebach, S., Herzog, F., Mayer, A., Cheung, A. C. M., Witte, G., et al. (2011). Molecular Basis of Rrn3-Regulated RNA Polymerase I Initiation and Cell Growth. *Genes Develop.* 25, 2093–2105. doi:10.1101/gad.17363311

Boukhgalter, B., Liu, M., Guo, A., Tripp, M., Tran, K., Huynh, C., et al. (2002). Characterization of a Fission Yeast Subunit of an RNA Polymerase I Essential Transcription Initiation Factor, SpRrn7h/TAF I 68, that Bridges Yeast and Mammals: Association with SpRrn11h and the Core Ribosomal RNA Gene Promoter. *Gene* 291, 187–201. doi:10.1016/s0378-1119(02)00597-8

Chaker-Margot, M., Barandun, J., Hunziker, M., and Klinge, S. (2017). Architecture of the Yeast Small Subunit Processome. *Science* 355, eaal1880. doi:10.1126/science.aal1880

Churchman, L. S., and Weissman, J. S. (2011). Nascent Transcript Sequencing Visualizes Transcription at Nucleotide Resolution. *Nature* 469, 368–373. doi:10.1038/nature09652

Clarke, A. M., Engel, K. L., Giles, K. E., Petit, C. M., and Schneider, D. A. (2018). NETSeq Reveals Heterogeneous Nucleotide Incorporation by RNA Polymerase I. *Proc. Natl. Acad. Sci. USA* 115, E11633–E11641. doi:10.1073/pnas.1809421115

Dangkulwanich, M., Ishibashi, T., Liu, S., Kireeva, M. L., Lubkowska, L., Kashlev, M., et al. (2013). Complete Dissection of Transcription Elongation Reveals Slow Translocation of RNA Polymerase II in a Linear Ratchet Mechanism. *Elife* 2, e00971. doi:10.7554/eLife.00971

Darrière, T., Pilsl, M., Sarthou, M.-K., Chauvier, A., Genty, T., Audibert, S., et al. (2019). Genetic Analyses Led to the Discovery of a Super-active Mutant of the RNA Polymerase I. *Plos Genet.* 15, e1008157. doi:10.1371/journal.pgen.1008157

El Hage, A., French, S. L., Beyer, A. L., and Tollervey, D. (2010). Loss of Topoisomerase I Leads to R-Loop-Mediated Transcriptional Blocks During Ribosomal RNA Synthesis. *Genes Develop.* 24, 1546–1558. doi:10.1101/gad.573310

El Hage, A., Koper, M., Kufel, J., and Tollervey, D. (2008). Efficient Termination of Transcription by RNA Polymerase I Requires the 5' Exonuclease Rat1 in Yeast. *Genes Develop.* 22, 1069–1081. doi:10.1101/gad.463708

Engel, C., Plitzko, J., and Cramer, P. (2016). RNA Polymerase I-Rrn3 Complex at 4.8 Å Resolution. *Nat. Commun.* 7, 12129. doi:10.1038/ncomms12129

Engel, C., Sainsbury, S., Cheung, A. C., Kostrewa, D., and Cramer, P. (2013). RNA Polymerase I Structure and Transcription Regulation. *Nature* 502, 650–655. doi:10.1038/nature12712

Engel, K. L., French, S. L., Viktorovskaya, O. V., Beyer, A. L., and Schneider, D. A. (2015). Spt6 Is Essential for rRNA Synthesis by RNA Polymerase I. *Mol. Cel. Biol.* 35, 2321–2331. doi:10.1128/MCB.01499-14

Fernández-Tornero, C., Moreno-Morcillo, M., Rashid, U. J., Taylor, N. M. I., Ruiz, F. M., Gruene, T., et al. (2013). Crystal Structure of the 14-subunit RNA Polymerase I. *Nature* 502, 644–649. doi:10.1038/nature12636

Ferreira, R., Schneekloth, J. S., Panov, K. I., Hannan, K. M., and Hannan, R. D. (2020). Targeting the RNA Polymerase I Transcription for Cancer Therapy Comes of Age. *Cells* 9, 266. doi:10.3390/cells9020266

French, S. L., Osheim, Y. N., Cioci, F., Nomura, M., and Beyer, A. L. (2003). In Exponentially Growing *Saccharomyces cerevisiae* Cells, rRNA Synthesis Is Determined by the Summed RNA Polymerase I Loading Rate Rather Than by the Number of Active Genes. *Mol. Cel. Biol.* 23, 1558–1568. doi:10.1128/mcb.23.5.1558-1568.2003

Gadal, O., Mariotte-Labarre, S., Chedin, S., Quemeneur, E., Carles, C., Sentenac, A., et al. (1997). A34.5, a Nonessential Component of Yeast RNA Polymerase I, Cooperates with Subunit A14 and DNA Topoisomerase I to Produce a Functional rRNA Synthesis Machine. *Mol. Cel. Biol.* 17, 1787–1795. doi:10.1128/mcb.17.4.1787

Han, Y., Yan, C., Nguyen, T. H. D., Jackobel, A. J., Ivanov, I., Knutson, B. A., et al. (2017). Structural Mechanism of ATP-independent Transcription Initiation by RNA Polymerase I. *Elife* 6, e27414. doi:10.7554/eLife.27414

Heberling, T., Davis, L., Gedeon, J., Morgan, C., and Gedeon, T. (2016). A Mechanistic Model for Cooperative Behavior of Co-transcribing RNA Polymerases. *Plos Comput. Biol.* 12, e1005069. doi:10.1371/journal.pcbi.1005069

Henras, A. K., Sam, M., Hiley, S. L., Wu, H., Hughes, T. R., Feigon, J., et al. (2005). Biochemical and Genomic Analysis of Substrate Recognition by the Double-Stranded RNA Binding Domain of Yeast RNase III. *RNA* 11, 1225–1237. doi:10.1261/rna.2760705

Higashino, A., Shiwa, Y., Yoshikawa, H., Kokubo, T., and Kasahara, K. (2015). Both HMG Boxes in Hmo1 Are Essential for DNA Binding In Vitro and In Vivo. *Biosci. Biotechnol. Biochem.* 79, 384–393. doi:10.1080/09168451.2014.978258

Huet, J., Buhler, J. M., Sentenac, A., and Fromageot, P. (1975). Dissociation of Two Polypeptide Chains from Yeast RNA Polymerase A. *Proc. Natl. Acad. Sci.* 72, 3034–3038. doi:10.1073/pnas.72.8.3034

Hung, S. S., Lesmana, A., Peck, A., Lee, R., Tchoubrieva, E., Hannan, K. M., et al. (2017). Cell Cycle and Growth Stimuli Regulate Different Steps of RNA Polymerase I Transcription. *Gene* 612, 36–48. doi:10.1016/j.gene.2016.12.015

Jaiswal, R., Choudhury, M., Zaman, S., Singh, S., Santosh, V., Bastia, D., et al. (2016). Functional Architecture of the Reb1-Ter Complex of *Schizosaccharomyces pombe*. *Proc. Natl. Acad. Sci. USA* 113, E2267–E2276. doi:10.1073/pnas.1525465113

Kamieniarz-Gdula, K., and Proudfoot, N. J. (2019). Transcriptional Control by Premature Termination: A Forgotten Mechanism. *Trends Genet.* 35, 553–564. doi:10.1016/j.tig.2019.05.005

Kampen, K. R., Sulima, S. O., Vereecke, S., and De Keersmaecker, K. (2020). Hallmarks of Ribosomopathies. *Nucleic Acids Res.* 48, 1013–1028. doi:10.1093/nar/gkz637

Keener, J., Josaitis, C. A., Dodd, J. A., and Nomura, M. (1998). Reconstitution of Yeast RNA Polymerase I Transcription In Vitro from Purified Components. *J. Biol. Chem.* 273, 33795–33802. doi:10.1074/jbc.273.50.33795

Kim, M., Ahn, S.-H., Krogan, N. J., Greenblatt, J. F., and Buratowski, S. (2004). Transitions in RNA Polymerase II Elongation Complexes at the 3′ Ends of Genes. *EMBO J.* 23, 354–364. doi:10.1038/sj.emboj.7600053

Kim, S., Beltran, B., Irnov, I., and Jacobs-Wagner, C. (2019). Long-Distance Cooperative and Antagonistic RNA Polymerase Dynamics via DNA Supercoiling. *Cell* 179, 106–119. doi:10.1016/j.cell.2019.08.033

Knutson, B. A., Smith, M. L., Belkevich, A. E., and Fakhouri, A. M. (2020). Molecular Topology of RNA Polymerase I Upstream Activation Factor. *Mol. Cel Biol* 40. doi:10.1128/MCB.00056-20

Kuhn, C.-D., Geiger, S. R., Baumli, S., Gartmann, M., Gerber, J., Jennebach, S., et al. (2007). Functional Architecture of RNA Polymerase I. *Cell* 131, 1260–1272. doi:10.1016/j.cell.2007.10.051

Lang, W. H., and Reeder, R. H. (1993). The REB1 Site Is an Essential Component of a Terminator for RNA Polymerase I in *Saccharomyces cerevisiae*. *Mol. Cel. Biol.* 13, 649–658. doi:10.1128/mcb.13.1.649

Lesne, A., Victor, J.-M., Bertrand, E., Basyuk, E., and Barbi, M. (2018). The Role of Supercoiling in the Motor Activity of RNA Polymerases. *Methods Mol. Biol.* 1805, 215–232. doi:10.1007/978-1-4939-8556-2_11

Liljelund, P., Mariotte, S., Buhler, J. M., and Sentenac, A. (1992). Characterization and Mutagenesis of the Gene Encoding the A49 Subunit of RNA Polymerase A in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* 89, 9302–9305. doi:10.1073/pnas.89.19.9302

Liu, L. F., and Wang, J. C. (1987). Supercoiling of the DNA Template During Transcription. *Proc. Natl. Acad. Sci.* 84, 7024–7027. doi:10.1073/pnas.84.20.7024

Luo, W., Johnson, A. W., and Bentley, D. L. (2006). The Role of Rat1 in Coupling mRNA 3'-end Processing to Transcription Termination: Implications for a Unified Allosteric-torpedo Model. *Genes Develop.* 20, 954–965. doi:10.1101/gad.1409106

Ma, J., Bai, L., and Wang, M. D. (2013). Transcription Under Torsion. *Science* 340, 1580–1583. doi:10.1126/science.1235441

Merkl, P. E., Pilsl, M., Fremter, T., Schwank, K., Engel, C., Längst, G., et al. (2020). RNA Polymerase I (Pol I) Passage Through Nucleosomes Depends on Pol I Subunits Binding its Lobe Structure. *J. Biol. Chem.* 295, 4782–4795. doi:10.1074/jbc.RA119.011827

Merkl, P., Perez-Fernandez, J., Pilsl, M., Reiter, A., Williams, L., Gerber, J., et al. (2014). Binding of the Termination Factor Nsi1 to its Cognate DNA Site Is Sufficient to Terminate RNA Polymerase I Transcription In Vitro and to Induce Termination In Vivo. *Mol. Cell Biol.* 34, 3817–3827. doi:10.1128/MCB.00395-14

Miller, O. L., and Beatty, B. R. (1969). Visualization of Nucleolar Genes. *Science* 164, 955–957. doi:10.1126/science.164.3882.955

Milligan, L., Huynh-Thu, V. A., Delan-Forino, C., Tuck, A., Petfalski, E., Lombraña, R., et al. (2016). Strand-specific, High-resolution Mapping of Modified RNA Polymerase II. *Mol. Syst. Biol.* 12, 874. doi:10.15252/msb.20166869

Moreno-Morcillo, M., Taylor, N. M. I., Gruene, T., Legrand, P., Rashid, U. J., Ruiz, F. M., et al. (2014). Solving the RNA Polymerase I Structural Puzzle. *Acta Cryst. D Biol. Crystallogr.* 70, 2570–2582. doi:10.1107/S1399004714015788

Mullem, V. V., Landrieux, E., Vandenhaute, J., and Thuriaux, P. (2002). Rpa12p, a Conserved RNA Polymerase I Subunit with Two Functional Domains. *Mol. Microbiol.* 43, 1105–1113. doi:10.1046/j.1365-2958.2002.02824.x

Neyer, S., Kunz, M., Geiss, C., Hantsche, M., Hodirnau, V.-V., Seybert, A., et al. (2016). Structure of RNA Polymerase I Transcribing Ribosomal DNA Genes. *Nature* 540, 607–610. doi:10.1038/nature20561

Nogi, Y., Yano, R., Dodd, J., Carles, C., and Nomura, M. (1993). Gene RRN4 in *Saccharomyces cerevisiae* Encodes the A12.2 Subunit of RNA Polymerase I and Is Essential Only at High Temperatures. *Mol. Cel. Biol.* 13, 114–122. doi:10.1128/mcb.13.1.114-122.1993

Nomura, M. (2001). Ribosomal RNA Genes, RNA Polymerases, Nucleolar Structures, and Synthesis of rRNA in the Yeast *Saccharomyces cerevisiae*. *Cold Spring Harbor Symposia Quantitative Biol.* 66, 555–566. doi:10.1101/sqb.2001.66.555

Osheim, Y. N., French, S. L., Sikes, M. L., and Beyer, A. L. (2009). Electron Microscope Visualization of RNA Transcription and Processing in *Saccharomyces cerevisiae* by Miller Chromatin Spreading. *Methods Mol. Biol.* 464, 55–69. doi:10.1007/978-1-60327-461-6_4

Peyroche, G., Milkereit, P., Bischler, N., Tschochner, H., Schultz, P., Sentenac, A., et al. (2000). The Recruitment of RNA Polymerase I on rDNA Is Mediated by the Interaction of the A43 Subunit with Rrn3. *EMBO J.* 19, 5473–5482. doi:10.1093/emboj/19.20.5473

Reeder, R. H., Guevara, P., and Roan, J. G. (1999). *Saccharomyces cerevisiae* RNA Polymerase I Terminates Transcription at the Reb1 Terminator In Vivo. *Mol. Cel. Biol.* 19, 7369–7376. doi:10.1128/mcb.19.11.7369

Reiter, A., Hamperl, S., Seitz, H., Merkl, P., Perez-Fernandez, J., Williams, L., et al. (2012). The Reb1-Homologue Ydr026c/Nsi1 Is Required for Efficient RNA Polymerase I Termination in Yeast. *EMBO J.* 31, 3480–3493. doi:10.1038/emboj.2012.185

Riva, M., Schäffner, A. R., Sentenac, A., Hartmann, G. R., Mustaev, A. A., Zaychikov, E. F., et al. (1987). Active Site Labeling of the RNA Polymerases A, B, and C from Yeast. *J. Biol. Chem.* 262, 14377–14380. doi:10.1016/s0021-9258(18)47803-9

Sadian, Y., Baudin, F., Tafur, L., Murciano, B., Wetzel, R., Weis, F., et al. (2019). Molecular Insight into RNA Polymerase I Promoter Recognition and Promoter Melting. *Nat. Commun.* 10, 5543. doi:10.1038/s41467-019-13510-w

Sadian, Y., Tafur, L., Kosinski, J., Jakobi, A. J., Wetzel, R., Buczak, K., et al. (2017). Structural Insights into Transcription Initiation by Yeast RNA Polymerase I. *EMBO J.* 36, 2698–2709. doi:10.15252/embj.201796958

Schneider, D. A., French, S. L., Osheim, Y. N., Bailey, A. O., Vu, L., Dodd, J., et al. (2006). RNA Polymerase II Elongation Factors Spt4p and Spt5p Play Roles in Transcription Elongation by RNA Polymerase I and rRNA Processing. *Proc. Natl. Acad. Sci.* 103, 12707–12712. doi:10.1073/pnas.0605686103

Schneider, D. A., Michel, A., Sikes, M. L., Vu, L., Dodd, J. A., Salgia, S., et al. (2007). Transcription Elongation by RNA Polymerase I Is Linked to Efficient rRNA Processing and Ribosome Assembly. *Mol. Cel* 26, 217–229. doi:10.1016/j.molcel.2007.04.007

Scull, C. E., Lucius, A. L., and Schneider, D. A. (2021). The N-Terminal Domain of the A12.2 Subunit Stimulates RNA Polymerase I Transcription Elongation. *Biophys. J.* 120 (10), 1883–1893. doi:10.1016/j.bpj.2021.03.007

Smith, M. L., Cui, W., Jackobel, A. J., Walker-Kopp, N., and Knutson, B. A. (2018). Reconstitution of RNA Polymerase I Upstream Activating Factor and the Roles of Histones H3 and H4 in Complex Assembly. *J. Mol. Biol.* 430, 641–654. doi:10.1016/j.jmb.2018.01.003

Steffan, J. S., Keys, D. A., Dodd, J. A., and Nomura, M. (1996). The Role of TBP in rDNA Transcription by RNA Polymerase I in *Saccharomyces cerevisiae*: TBP Is Required for Upstream Activation Factor-dependent Recruitment of Core Factor. *Genes Develop.* 10, 2551–2563. doi:10.1101/gad.10.20.2551

Sulima, S., Kampen, K., and De Keersmaecker, K. (2019). Cancer Biogenesis in Ribosomopathies. *Cells* 8, 229. doi:10.3390/cells8030229

Swanson, M. S., and Winston, F. (1992). SPT4, SPT5 and SPT6 Interactions: Effects on Transcription and Viability in *Saccharomyces cerevisiae*. *Genetics* 132, 325–336. doi:10.1093/genetics/132.2.325

Tafur, L., Sadian, Y., Hanske, J., Wetzel, R., Weis, F., and Müller, C. W. (2019). The Cryo-EM Structure of a 12-subunit Variant of RNA Polymerase I Reveals Dissociation of the A49-A34.5 Heterodimer and Rearrangement of Subunit A12.2. *Elife* 8, e43204. doi:10.7554/eLife.43204

Tafur, L., Sadian, Y., Hoffmann, N. A., Jakobi, A. J., Wetzel, R., Hagen, W. J. H., et al. (2016). Molecular Structures of Transcribing RNA Polymerase I. *Mol. Cel* 64, 1135–1143. doi:10.1016/j.molcel.2016.11.013

Tantale, K., Mueller, F., Kozulic-Pirher, A., Lesne, A., Victor, J.-M., Robert, M.-C., et al. (2016). A Single-Molecule View of Transcription Reveals Convoys of RNA Polymerases and Multi-Scale Bursting. *Nat. Commun.* 7, 12248. doi:10.1038/ncomms12248

Turowski, T. W., Petfalski, E., Goddard, B. D., French, S. L., Helwak, A., and Tollervey, D. (2020). Nascent Transcript Folding Plays a Major Role in Determining RNA Polymerase Elongation Rates. *Mol. Cel* 79, 488–503. doi:10.1016/j.molcel.2020.06.002

Viktorovskaya, O. V., Appling, F. D., and Schneider, D. A. (2011). Yeast Transcription Elongation Factor Spt5 Associates with RNA Polymerase I and RNA Polymerase II Directly. *J. Biol. Chem.* 286, 18825–18833. doi:10.1074/jbc.M110.202119

Warner, J. R. (1999). The Economics of Ribosome Biosynthesis in Yeast. *Trends Biochem. Sci.* 24, 437–440. doi:10.1016/s0968-0004(99)01460-7

West, S., Gromak, N., and Proudfoot, N. J. (2004). Human 5′ → 3′ Exonuclease Xrn2 Promotes Transcription Termination at Co-transcriptional Cleavage Sites. *Nature* 432, 522–525. doi:10.1038/nature03035

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 20 December 2021
doi: 10.3389/fmolb.2021.679292

Several Isoforms for Each Subunit Shared by RNA Polymerases are Differentially Expressed in the Cultivated Olive Tree (*Olea europaea* L.)

Isabel Fernández-Parras[1], *Jorge Antolín Ramírez-Tejero*[1], *Francisco Luque*[1,2]* *and Francisco Navarro*[1,2]*

[1]Departamento de Biología Experimental-Genética, Jaén, Spain, [2]Centro de Estudios Avanzados en Aceite de Oliva y Olivar, Universidad de Jaén, Jaén, Spain

Plants contain five nuclear RNA polymerases, with RNA pols IV and V in addition to conserved eukaryotic RNA pols I, II, and III. These transcriptional complexes share five common subunits, which have been extensively analyzed only in yeasts. By taking advantage of the recently published olive tree cultivar (*Olea europaea* L. cv. Picual) genome, we performed a genome-wide analysis of the genomic composition corresponding to subunits common to RNA pols. The cultivated olive tree genome is quite complex and contains many genes with several copies. We also investigated, for the first time, gene expression patterns for subunits common to RNA pols using RNA-Seq under different economically and biologically relevant conditions for the cultivar "Picual": tissues/organs, biotic and abiotic stresses, and early development from seeds. Our results demonstrated the existence of a multigene family of subunits common to RNA pols, and a variable number of paralogs for each subunit in the olive cultivar "Picual." Furthermore, these isoforms display specific and differentiated expression profiles depending on the isoform and growth conditions, which may be relevant for their role in olive tree biology.

Keywords: RNA polymerases, plants, olive, expression conditions, common subunits

OPEN ACCESS

Edited by:
Manuela Helmer-Citterich, University of Rome Tor Vergata, Italy

Reviewed by:
Angelo Gismondi, University of Rome Tor Vergata, Italy
Rosanna Petrella, University of Milan, Italy

***Correspondence:**
Francisco Luque
fjluque@ujaen.es
Francisco Navarro
fngomez@ujaen.es

Specialty section:
This article was submitted to Protein and RNA Networks, a section of the journal Frontiers in Molecular Biosciences

Received: *11 March 2021*
Accepted: *07 October 2021*
Published: *20 December 2021*

Citation:
*Fernández-Parras I, Ramírez-Tejero JA, Luque F and Navarro F (2021) Several Isoforms for Each Subunit Shared by RNA Polymerases are Differentially Expressed in the Cultivated Olive Tree (*Olea europaea *L.). Front. Mol. Biosci. 8:679292. doi: 10.3389/fmolb.2021.679292*

INTRODUCTION

Gene expression is a highly regulated process that comprises coordinated steps to ensure appropriate RNA levels and to allow cells to correctly respond and adapt to any situation. Transcription is the most widely studied step in gene expression that is carried out by RNA polymerases (RNA pols). In bacteria, archaea and eukarya RNA pols are heteromultimeric complexes responsible for the specific synthesis of different RNA types (Werner and Grohmann, 2011). Most eukaryotes possess three heteromultimeric RNA polymerases, namely, RNA pol I, RNA pol II, and RNA pol III (also known as RNA pols A, B, and C in plants). While RNA pol I synthesizes the precursor of the three largest rRNAs, RNA pol III synthesizes tRNAs, 5S rRNA, and several short non-translated RNAs. RNA pol II produces all mRNAs and many non-coding RNAs, including miRNA (Kwapisz et al., 2008; Werner et al., 2009; Werner and Grohmann, 2011; Cramer, 2019). Furthermore, plants contain two additional RNA pols—RNA pols IV and V (or RNA pols D and E)—that play roles in epigenetic regulation. They synthesize siRNAs that play roles in transcriptional silencing *via* RNA-directed

DNA methylation (RdDM) and also non-coding RNAs with a role in the development and response to environmental changes (Wierzbicki et al., 2008; Tucker et al., 2010; Haag and Pikaard, 2011; Lopez et al., 2011; Ream et al., 2013). Both RNA pols IV and V have evolved as specialized forms of RNA pol II, as demonstrated by mass spectrometry and phylogenetic analyses (Huang et al., 2009; Ream et al., 2009; Tucker et al., 2010; Wang and Ma, 2015).

RNA pols I and III are composed of 14 and 17 subunits, respectively, while RNA pol II contains 12 (Kwapisz et al., 2008; Werner et al., 2009; Werner and Grohmann, 2011; Cramer, 2019). Plant RNA pols IV and V, which have evolved from RNA pol II, are also composed of 12 subunits, some of which are shared with RNA pol II (Wierzbicki et al., 2008; Tucker et al., 2010; Haag and Pikaard, 2011; Lopez et al., 2011; Ream et al., 2013; Wang and Ma, 2015). It has been described that the NRP4 subunit, shared by RNA pols II, IV, and V, is missing in cauliflower RNA pol V. However, this enzyme maintains its role in RNA silencing (Huang et al., 2009).

Eukaryotic RNA pols I, II, and III share five common subunits (Rpb5, Rpb6, Rpb8, Rpb10, and Rpb12) with archaeal homologs (Woychik et al., 1990; Shpakovski et al., 1995; Cuevas-Bermúdez et al., 2017). In plants, several paralogs of subunits common to RNA pols have been identified. Some are shared by several of, or all five, RNA pols, while others are RNA pol–specific. This is the case for subunits NRP5 and NRP6, which are specific or shared only by RNA pol IV and/or V, or even shared by RNA pols II, IV, and V (Wierzbicki et al., 2008; Huang et al., 2009; Tucker et al., 2010; Haag and Pikaard, 2011; Ream et al., 2013; Wang and Ma, 2015). Similarly, Rpb5 and Rpb6 paralogs have been described in trypanosomes (Kelly et al., 2005; Devaux et al., 2006). Subunits common to RNA pols have been extensively analyzed in yeast (Cuevas-Bermúdez et al., 2017), while their role and contribution to transcription are still unclear in plants. Notably, although subunits common to RNA pols must perform similar functions in the corresponding RNA pols, some of these subunits have been described to also play specific roles in transcription (Woychik et al., 1990; Zaros et al., 2007; Cuevas-Bermúdez et al., 2017; Martínez-Fernández et al., 2017). This must also account for plant NRPD/E5 subunits (RNA pol IV/V) with roles in gene silencing and RNA-directed DNA methylation, RdDM (Huang et al., 2009; Pikaard and Tucker, 2009; Lopez et al., 2011; Zhou and Law, 2015). Notably, subunit NRP6, one of the subunits common to RNA pols having a bacterial homolog (the ω subunit), is thought to be important for RNA pols assembly and stability and to play specific roles in transcription (Werner and Grohmann, 2011; Garrido-Godino et al., 2013; Nouraini et al., 1996; Lanzendorfer et al., 1997; Minakhin et al., 2001). It is worth noting that the Rpb8 common subunit is described as being eukaryote-specific, although an Rpb8 archaeal ortholog, called G or Rpo8, has been identified in Crenarchaeota, and it is thought to be a protein that appears at an early step in eukaryotic evolution (Koonin et al., 2007; Kwapisz et al., 2008).

The olive tree (*Olea europaea* L.) is one of the most important fruit trees in the Mediterranean Basin. Olive cultivation is important in economic, agronomic, and agro-ecological terms in many countries. Given its health benefits and economic impact in Mediterranean countries, olive oil is probably the most important vegetable oil in the world (Conde et al., 2008). Extra virgin olive oil is appreciated worldwide, thanks to its benefits for human health (Donaire et al., 2011). Its global demand continuously rises, and "Picual" is one of the most extensively cultivated olive varieties, thanks to the organoleptic properties of its extra virgin olive oil and excellent oxidative stability (Gutiérrez et al., 2001; Talhaoui et al., 2016). The cultivar "Picual," of Andalusian origin, is the leading cultivar in Spain that accounts for 50% of national oil production and 20% of world oil production. The *O. europaea* genome is diploid, with 46 chromosomes (2n), whose size ranges from 1.48 to 2.2 Gb depending on the sequenced variety (Rugini et al., 1996; Loureiro et al., 2007). Recently, the "Picual" genome has been reported (Jiménez-Ruiz et al., 2020). "Picual" is a diploid organism (2n = 2x = 46) whose genome is larger than that of *O. europaea* var. *sylvestris*, with an estimated size of 1.68 Gb and 79667 genes, more than wild genomes (Jiménez-Ruiz et al., 2020). The cultivated olive genome results from two independent whole-genome duplications (WGDs) from around 62 and 25 million years ago, but with very recent partial genome duplications (Unver et al., 2017; Jiménez-Ruiz et al., 2020). "Picual" displays an excellent capacity to adapt to a wide variety of growth conditions, soils, stress, or pathogenic agent infections. "Picual" is sensitive to *Verticillium dahliae* infection, as are most olive cultivars, which has an important impact on economy or ecology. Adaptation to all these situations results from complex transcriptional responses, among other regulatory events (Leyva-Pérez et al., 2015; Jiménez-Ruiz et al., 2017; Jiménez-Ruiz et al., 2018).

Although extensive global transcriptomic studies have been performed on "Picual" (tissues and organs, abiotic cold stress, biotic stress by *V. dahliae* infection) and in early development from seeds, very little is known about transcriptional machinery regulation, and specifically about regulation of RNA pols. Taking advantage of this cultivar's recently reported genome (Jiménez-Ruiz et al., 2020) and its economic, agronomic, and ecological importance, at least in the Mediterranean Basin, we analyzed the RNA pol gene composition and gene expression regulation in different relevant situations that impact growth or are of interest for genetic improvement. To do so, we focused on subunits common to RNA pols by understanding that they are shared by different RNA pols and that plants contain several specific paralogs of some RNA pol subunits. Based on both "Picual" genome and RNA-Seq datasets from tissues and organs, abiotic cold stress, biotic stress by *V. dahliae* infection, and early development transcriptomic studies, we investigated and described the existence of multigene families coding for subunits common to RNA pols and elucidated their differential transcriptional responses under these conditions.

MATERIALS AND METHODS

Genome-Wide Identification of Genes and Proteins for Common Subunits of RNA Polymerases in the Olive "Picual" Genome

Arabidopsis RNA pol common subunit genes were identified in The Arabidopsis Information Resource (TAIR) (www.arabidopsis.org). Protein sequence queries were used to search for homologs by BlastP with an E value of $<1 \times 10^{-5}$ to identify common subunit proteins of plant RNA polymerases.

The common subunits of the *Arabidopsis thaliana* RNA pols used as queries were NRPA/D5, At3g22320; NRPE5, At3g57080; NRPE5-Like, At2g41340; NRPA/E6a, At5g51940; NRPA/E6b, At2g04630; NRPA/E8a, At1g54250; NRPA/E8b, At3g59600; NRPA/E10, At1g11475; NRPB10-like, At1g61700; NRPA/E12a, At5g41010; NRPB12-like, At1g53690. The identified sequences of the common subunits of Arabidopsis RNA pols were subsequently employed as queries to recover their homologs from the "Picual" genome using BlastP searches, available at the OliveTreeDB website (https://genomaolivar.dipujaen.es/db/index.php). Genomic, cDNA, CDS, and protein sequences were obtained for each common subunit of the RNA pols.

The retrieved "Picual" NRP5 sequences were aligned to other plant RNA pol sequences for further analyses. The other common subunits of plant RNA polymerases herein used were NRPB5a_*Zea mays*, NP_001141164; NRPB5b_*Zea mays*, NP_001132429.1; NRPE_*Zea mays*, ACG37268; NRPE5_*Pinus canariensis*, AJA90785.1; NRPE5_*Ginkgo biloba*, AJA90777.1; NRPE5_*Ephedra trifurca*, AJA90766.1; NRPE5_*Cycas revoluta*, AJA90761.1; NRP5A-like.a (*O. europapea sylvestris*), XP_022875925; NRP5A-like.c (*O. europapea sylvestris*), XP_022871082.1; NRPE5 (*O. europapea sylvestris*) XP_022872077.1.

RNA-Seq Analysis

All RNA-Seq datasets used for the different studies have been previously described and are indicated later. The transcriptional steady-state levels of the olive cultivar "Picual" genes for subunit common to the RNA pols in organs/tissues were obtained from previously described datasets (Ramírez-Tejero et al., 2020). In brief, samples were collected from the roots, stems, meristems, leaves, flowers, and fruit of three healthy 10-year-old "Picual" olive trees under field conditions at the World Olive Germplasm Collection (WOGC) of the Andalusian Institute of Agricultural and Fisheries Research and Training (IFAPA), Córdoba, Spain. Two biological replicates (consisting of an equilibrated pool of three plant RNAs per sample) were sequenced.

The datasets described later were used for early development plant samples (Jiménez-Ruiz et al., 2018). For plant material preparation purposes, the seeds from the open pollinated cultivar Arbequina were induced to geminate at the Agrarian Research and Training Center (IFAPA) in Churriana, Spain. Seedlings were grown *in vitro* under chamber conditions with a 16-h photoperiod of fluorescent light at a constant temperature of 25°C until they were 2 months old. Then they were potted and grown in a conditioned greenhouse (25°C). The aerial parts of 10 plants were collected 1, 2, 3, 4, 5, and 6 months after seed activation. Two biological replicates of 10 pooled plants per sample were sequenced.

The *V. dahliae*–infected plants were obtained at the Department of Crop Protection, Institute for Sustainable Agriculture, Córdoba, Spain, and data from previously reported datasets were used (Jiménez-Ruiz et al., 2017). 40 plants were infected by root-dip inoculation in a conidial suspension (10^7 conidia ml^{-1}) of defoliating *V. dahliae* isolate V937I. As a control group, 40 non-inoculated plants were handled in the same way to be used in the absence of the pathogen cited before. For each biological replicate, roots from three plants were pooled from the control or after 2 and 7 days postinfection, and the cDNA from the samples was sequenced.

The cold stress and cold acclimation data were obtained from the previously described datasets (Leyva-Pérez et al., 2015) and were obtained as the previous *V. dahliae*–infected plants at the Department of Crop Protection, Institute for Sustainable Agriculture, Córdoba, Spain. For this purpose, thirty-five 4-month-old potted olive "Picual" cultivar plants were used, acclimated at 24°C, and then incubated with a 14-h photoperiod of fluorescent light at 65 μmol m^2 s (10°C day/4°C night) for 10 days and constant 76–78% relative humidity. Another group of 15 plants was used as the control treatment. Aerial tissues were harvested at 0, 10, and 24 days, and three plants were pooled for each biological replicate for RNA extraction and sequencing purposes.

For RNA sequencing, samples were immediately frozen in liquid nitrogen and stored at −80°C. Total RNA was extracted with the Spectrum Plant Total RNA Kit (Sigma-Aldrich, St. Louis, MO, United States) according to the manufacturer's instructions. Two technical replicates of each sample were sequenced by paired-end sequencing (101 × 2) in an Illumina® HiSeq sequencer (Illumina, San Diego, CA, United States) at Sistemas Genómicos company (Valencia, Spain).

The expression analysis was performed with DNAstar (ArrayStar 17, Rockville, MD, United States) for the RNA-seq analyses (www.dnastar.com). Reads were mapped to the "Picual" genome as reference Oleur061 (Jiménez-Ruiz et al., 2020). Mapping was performed with high-stringency parameters to differentiate between highly similar paralogs, k-mer = 63 and 95% matches. Data were normalized using parameter reads per kilobase of transcript, per million mapped reads (RPKM). A basal expression level of $\log_2$ RPKM = −2 was considered. Therefore, the genes with expression values above this threshold level were considered expressed, whereas those with values that equaled or were below the threshold level were considered not expressed.

Data Availability Statement

The RNAseq data are available at the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO). The organs/tissues data are available with accession numbers GSE140648, GSM4176229, GSM4176230, GSM4176231, GSM4176232, GSM4176233, GSM4176234, GSM4176235, GSM4176236, GSM4176237, GSM4176238, GSM4176239, and GSM4176240 for Project PRJNA556567 (Ramírez-Tejero et al., 2020).

The early development data are available with accession Numbers (NCBI: SAMN07603885, SAMN07603886, SAMN07603887, SAMN07603888, SAMN07603889, SAMN07603890, SAMN07603891,

SAMN07603892, SAMN07603893, SAMN07603894, SAMN07603895, and SAMN07603896) for Project PRJNA401310 (Jiménez-Ruiz et al., 2018). The data corresponding to the response to cold stress and to *V. dahliae* infection are available with accession numbers SRR1525051, SRR1525052, SRR1524949, SRR1524950, SRR1524951, SRR1524952, SRR1525086, SRR1525087, SRR1525113, SRR1525114) SRR1525231, SRR1525237, SRR1524947, SRR1524948, SRR1525213, SRR1525114, SRR1525224, SRR1525226, SRR1525284, SRR1525285, SRR1525286, SRR1525287, SRR1525415, SRR1525416, SRR1525436, and SRR1525437 (Leyva-Pérez et al., 2015; Jiménez-Ruiz et al., 2017).

Phylogenetic Analysis

For phylogenetic analysis, the Phylogeny.fr interface was used (www.phylogeny.fr) (Dereeper et al., 2008). To do so, amino acid sequences were aligned with MUSCLE, and Gblocks was used for alignment curation. The phylogenetic tree was reconstructed using the maximum parsimony method with software PhyML (Guindon et al., 2010). Finally TreeDyn for tree drawing was used.

The amino acid sequences used were: Arabidopsis subunits common to RNA polymerases NRPA/D5, At3g22320, NRPE5, At3g57080, and NRPE5-Like, At2g41340, as well as NRPB5a _*Zea mays*, NP_001141164; NRPB5b _*Zea mays*, NP_001132429.1; NRPE_*Zea mays*, ACG37268; NRPE5_*Pinus canariensis*, AJA90785.1; NRPE5_*Ginkgo biloba*, AJA90777.1; NRPE5_*Ephedra trifurca*, AJA90766.1; NRPE5_*Cycas revoluta*, AJA90761.1; NRP5A-like.a (*Olea europapea sylvestris*), XP_022875925; NRP5A-like.c (*Olea europapea sylvestris*), XP_022871082.1; NRPE5 (*Olea europapea sylvestris*) XP_022872077.1.

RESULTS

Identification of Genes for Common Subunits of RNA Polymerases

In order to identify the subunits common to all the RNA pols from olive (NRP5, NRP6, NRP8, NRP10, NRP12), we used the recently reported "Picual" olive genome (Jiménez-Ruiz et al., 2020). We searched for loci containing the ORFs putatively coding for subunits common to RNA pols using BLAST search and the corresponding subunits common to RNA pols from *A. thaliana* as queries.

Arabidopsis has six genes that putatively encode NRP5 subunits (Larkin et al., 1999; Ream et al., 2009). However, only one subunit shared by RNA pols I-IV, NRPA/D5, and a second one specific to RNA pol V, NRPE5, have been identified in proteomic and functional analyses (Larkin et al., 1999; Ream et al., 2009; Law et al., 2011; Zhang et al., 2013). More recently, an NRPE5-like subunit was identified in proteomic analyses as being a component of RNA pol V, while this is not the case for two other putative NRPE-like subunits (Law et al., 2011). By using Arabidopsis NRPA/D5 as a query, we identified three putative genes coding for the "Picual" olive homolog subunits with identities falling within the 74–79% range. Concomitantly, we named them NRPA/D5a/b/c (**Table 1** and **Supplemental Figure S1**). The NRPE5 homolog search permitted the identification of only one putative gene coding for the NRPE5 subunit with about 62% identity (**Table 1** and **Supplemental Figure S1**). However, and unlike Arabidopsis, no NRPE5-like subunits were identified (58% and 68% identities between the Arabidopsis NRPE5-like and olive and the Arabidopsis NRPE5 subunits, respectively; **Table 1** and **Supplemental Figure S1**). Furthermore, the existence of two classes of NRP5 subunits in olive was corroborated by the phylogenetic analysis (**Figure 1**). In addition, the NRPE5 subunit maintained the short N-terminal extension described for the plant NRPE5, as compared to NRPA/D5 (**Supplemental Figure S2**), which was suggested to be important for protein stability *in vivo* in Arabidopsis (Ream et al., 2009).

Two NRP6 subunits, NRPA/E6a and NRPA/E6b, have been described in Arabidopsis (Ream et al., 2009; Law et al., 2011; Ream et al., 2015). By using them as queries, we identified five putative NRP6 coding genes in olive (**Table 1** and **Supplemental Figure S1**). Three of the corresponding subunits (named NRPA/E6a/b/c) showed high identity among them (88–96%) and ranged from 69 to 79% identity in relation to the Arabidopsis NRPE6 subunits. Strikingly, the other two subunits (named NRPA/E6d/e) were small in size, with 113 and 119 amino acids. NRPA/E6d showed about 59% identity compared to the Arabidopsis subunits, while NRPA/E6e displayed the least identity of about 49%.

Two NRP8 subunits in Arabidopsis have been shown to form part of all five RNA pols: NRPA/E8a and NRPA/E8b (Ream et al., 2009; Law et al., 2011; Ream et al., 2015). In olive, three putative coding genes for the NRPA/E8a/b/c subunits were identified. The NRPA/E8a identity range was 53–55% with the Arabidopsis subunits, while NRPA/E8b/c range was 45% (**Table 1** and **Supplemental Figure S1**).

In Arabidopsis, two NRP10 subunits (NRPA/E10 and NRPB10-like) and two NRP12 subunits (NRPA/E12 and NRPB12-like) have been detected (Ream et al., 2009; Law et al., 2011; Ream et al., 2015), although the association of NRPB10-like and NRPB12-like with RNA pols remains controversial (Ream et al., 2015). By using them as queries, two putative coding genes for NRPA/E10a/b have been identified in the olive "Picual" with identities within the 90–95% range with their Arabidopsis homologs (**Table 1** and **Supplemental Figure S1**). Conversely, four putative coding genes for the NRPA/E12a/b/c/d subunits were found in "Picual," whose identities were greater with Arabidopsis NRPA/E12 (67–90%) than with NRPB12-like (52–71%) (**Table 1** and **Supplemental Figure S1**). Interestingly, NRPA/E12d showed the fewest identities with the other NRPA/E12 olive subunits when they were all compared.

The Genes for Subunits Shared by RNA Polymerases are Spatially and Temporally Regulated

In line with the aforementioned data, we have speculated that several members of each distinct subunit common to RNA pols existed in the olive cultivar. To explore whether the corresponding putative coding genes were functional and expressed, we analyzed their spatial and temporal expression patterns.

We first investigated the expression of genes putatively coding for the different NRP5, NRP6, NRP8, NRP10, and NRP12 subunits from the "Picual" olive cultivar by analyzing their

TABLE 1 | Identified common subunits of RNA polymerase from olive.

Subunit		Gene accession	Protein size (amino acids)	mRNA expression			
				Organs and tissues	Biotic stress (*V. dahliae* infection)	Abiotic stress (Cold acclimation)	Development
NRP5	NRPA/D5a	Oleur061Scf3785g07006.1	203	+	+	+	+
	NRPA/D5b	Oleur061Scf0084g07007.1	217	+	+	+	+
	NRPA/D5c	Oleur061Scf3324g05023.1	206	+	+	+	+
	NRPE5	Oleur061Scf4420g01012.1	228	+	+	+	+
NRP6	NRPA/E6a	Oleur061Scf2238g07030.1	141	+	+	+	+
	NRPA/E6b	Oleur061Scf0173g01014.1	143	+	+	+	+
	NRPA/E6c	Oleur061Scf0677g02024.1	131	+	+	+	+
	NRPA/E6d	Oleur061Scf5121g00005.1	113	+	+	+	+
	NRPA/E6e	Oleur061Scf0350g03011.1	119	–	–	–	–
NRP8	NRPA/E8a	Oleur061Scf0592g02022.1	165	+	+	+	+
	NRPA/E8b	Oleur061Scf0022g02019.1	148	+	+	+	±
	NRPA/E8c	Oleur061Scf5855g00014.1	148	+	+	+	+
NRP10	NRPA/E10a	Oleur061Scf0656g01027.1	71	+	+	+	+
	NRPA/E10b	Oleur061Scf3000g03019.1	63	+	+	+	+
NRP12	NRPA/E12a	Oleur061Scf2481g09006.1	127	+	+	+	+
	NRPA/E12b	Oleur061Scf2607g00039.1	97	+	+	+	+
	NRPA/E12c	Oleur061Scf1394g00001.1	72	±	+	±	±
	NRPA/E12d	Oleur061Scf1987g03032.1	94	–	–	–	–

+ Means expressed in all samples. ± Means expressed in some samples. – Means not expressed in any sample.
Based on Blast analysis using A. thaliana *common subunits of RNA polymerases as queries:* ***NRPA/D5,*** *At3g22320****; NRPE5,*** *At3g57080****; NRPE5-Like,*** *At2g41340****; NRPA/E6a,*** *At5g51940* ***; NRPA/E6b,*** *At2g04630;* ***NRPA/E8a****, At1g54250;* ***NRPA/E8b****, At3g59600;* ***NRPA/E10****, At1g11475;* ***NRPB10-like****, At1g61700 ;* ***NRPA/E12a****, At5g41010 ;* ***NRPB12-like****, At1g53690.*

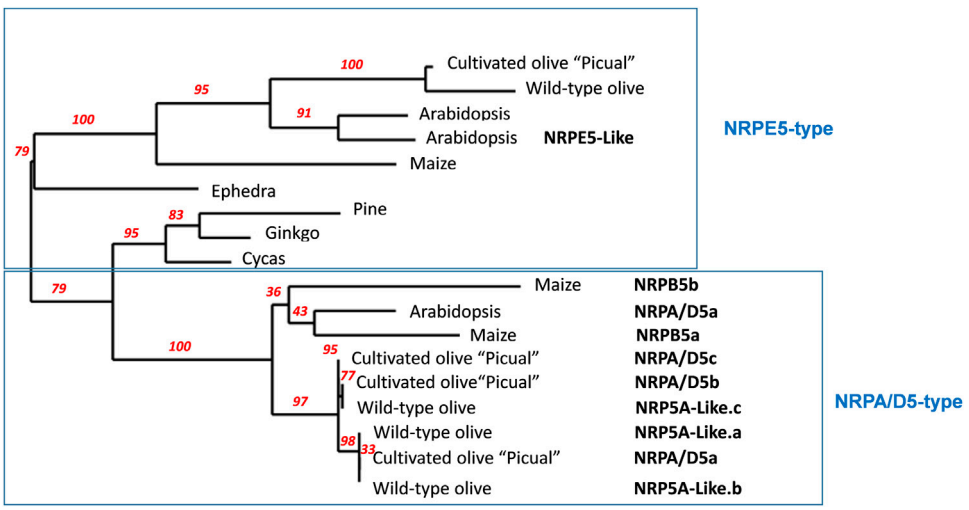


FIGURE 1 | Schematic phylogenetic diagram of NRP5 genes. NRP5 sequences were aligned with MUSCLE, and the unrooted phylogenetic tree was reconstructed using the maximum likelihood method with the PhyML algorithm. The numbers at the nodes represent the percentage bootstrap values (only those higher than 50% were represented). The reliability for the internal branch was assessed using 100 bootstrap replicates. The NRPE5-type denoted the RNA pol V–specific subunit, including the corresponding NRPE5-like described in Arabidopsis. Cultivated olive corresponds to *Olea europaea sylvestris* (Unver et al., 2017). Arabidopsis RNA polymerase common subunits NRPA/D5, At3g22320, NRPE5, At3g57080 and NRPE5-Like, At2g41340 were used and NRPB5a *_Zea mays*, NP_001141164; NRPB5b *_Zea mays*, NP_001132429.1; NRPE_*Zea mays*, ACG37268; NRPE5_*Pinus canariensis*, AJA90785.1; NRPE5_*Ginkgo biloba*, AJA90777.1; NRPE5_*Ephedra trifurca*, AJA90766.1; NRPE5_*Cycas revoluta*, AJA90761.1; NRP5A-like.a (*Olea europapea sylvestris*), XP_022875925.1; NRP5A-like.b (*Olea europapea sylvestris*), XP_022875924.1; NRP5A-like.c (*Olea europapea sylvestris*), XP_022871082.1; NRPE5 (*Olea europapea sylvestris*) XP_022872077.1.

mRNA levels in several organs (fruits, flowers, leaves, roots, stems) and tissues (meristems) with the RNA-Seq data from a previously detailed transcriptomic analysis (Ramírez-Tejero et al., 2020).

As shown in **Figure 2**, all the identified NRP genes from olive were expressed in all the analyzed organs and tissues, except Oleur061Scf0350g03011.1 for NRPA/E6e and Oleur061Scf 1987g03032.1 for NRPA/E12d, which were not expressed for

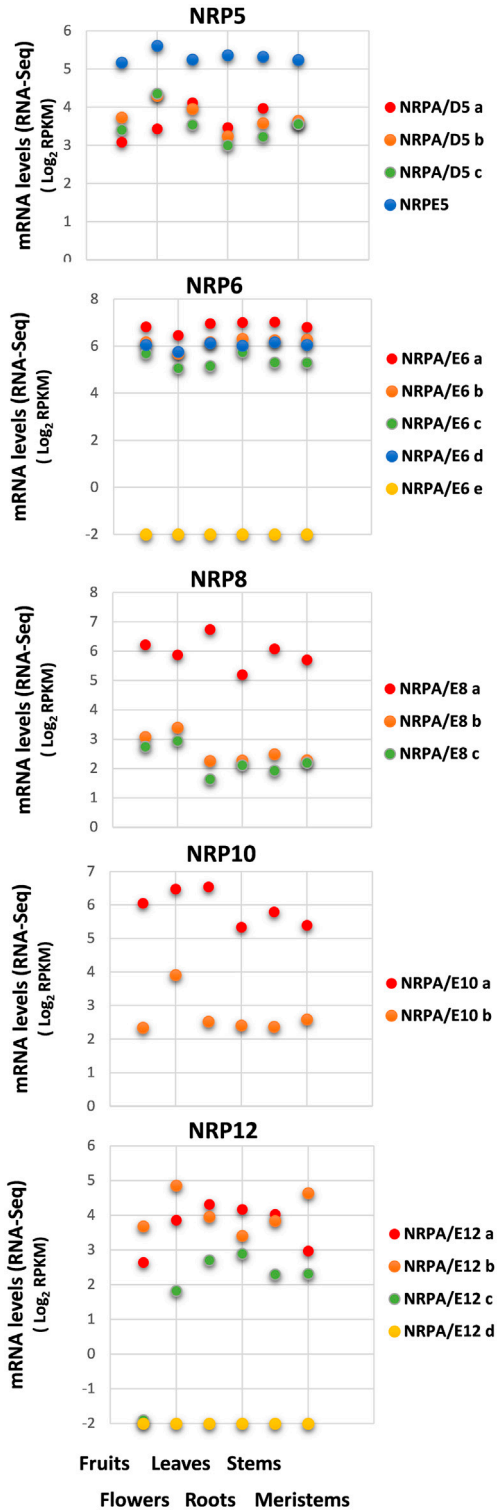


FIGURE 2 | Gene expression pattern of subunits common to RNA polymerases in different organs and tissues. Data from previously published RNA-Seq datasets (Ramírez-Tejero et al., 2020). Expression profile of the different NRP common subunit genes from olive cultivar "Picual" in fruit, flowers, leaves, roots, stems, and meristems. Data correspond to reads per kilobase of transcript per million mapped reads (RPKM). Data are represented as $\log_2$ (RPKM). Values ≤ −2 are considered no expression and are represented by a value of −2.

either condition. A complex expression pattern was also observed when we compared the genes of the different NRP subunits. However, when independently comparing each subset of NRP genes corresponding to each NRP subunit, our results showed that the different paralog genes tended to maintain similar expression levels in the different organs and tissues, except for the NRP12 genes and NRPA/E10b.

Regarding expression levels, NRPE5 showed the highest mRNA levels of all the NRP5 genes, while the other three NRPA/D5 genes were similarly expressed in the different organs and tissues. Moreover, the gene for the NRPA/E8a subunit was more highly expressed than the other two identified NRPA/E8 genes. This was also the case for the NRPA/E10a gene *versus* NRPA/E10b. Notably, no major differences in expression levels were observed for the four expressed NRP6 coding genes. Finally, the NRP12 genes presented the biggest differences when comparing their expression levels in the different analyzed organs and tissues. The NRPA/E12c gene was not expressed in fruits and was similarly expressed in the other analyzed organs and tissues. These differences in expression levels did not seem to maintain a relation with the total mRNA amount detected in any analyzed organ and tissue (**Supplemental Figure S3**).

These results collectively suggested that most NRP coding genes were expressed, with no major differences between different organs and tissues for each gene, except the NRP12c gene that was not expressed in fruits, and NRPA/E10b that was overexpressed in flowers. On the contrary, a clear spatial regulation with evident differences in the expression levels between paralog genes was observed, thus implying that some of these subunits have major contribution. We further investigated whether NRP common subunit genes from the olive cultivar could be temporally regulated by exploring their expression patterns using previously reported and corroborated datasets during the early juvenile development period from 1-month to 6-month seedlings (from germinated embryos to juvenile trees), the end of which corresponds to the juvenile development stage (Jiménez-Ruiz et al., 2018). Note that a complete transcriptomic study of olive development during the early juvenile period demonstrated that after 3–4 months of development, all plant structure and cell and organ differentiation have occurred, and thus, the juvenile tree development from seed is complete (Jiménez-Ruiz et al., 2018).

As shown in **Figure 3**, a general decrease in NRP gene expression was observed until development at 4 months, although differences were found in mRNA levels among distinct gene paralogs. We were unable to exclude some gene expressions not being mainly altered, as with NRPA/E6a. It is worth noting that 4 months corresponded to completed juvenile tree development from seed (Jiménez-Ruiz et al., 2018). Interestingly, some differences were evident: NRPA/D5a, NRPA/D6c, and NRPA/E12a gene expressions peaked at 2-month development and then lowered, while NRPA/E8b and NRPA/E12c expressions drastically decreased. These results could account for major transcriptional activity during early development from seed before later reaching the levels maintained mainly in juvenile and/or adult trees.

Notably, after juvenile tree formation (between 4 and 6 development months), the expression of most NRP genes

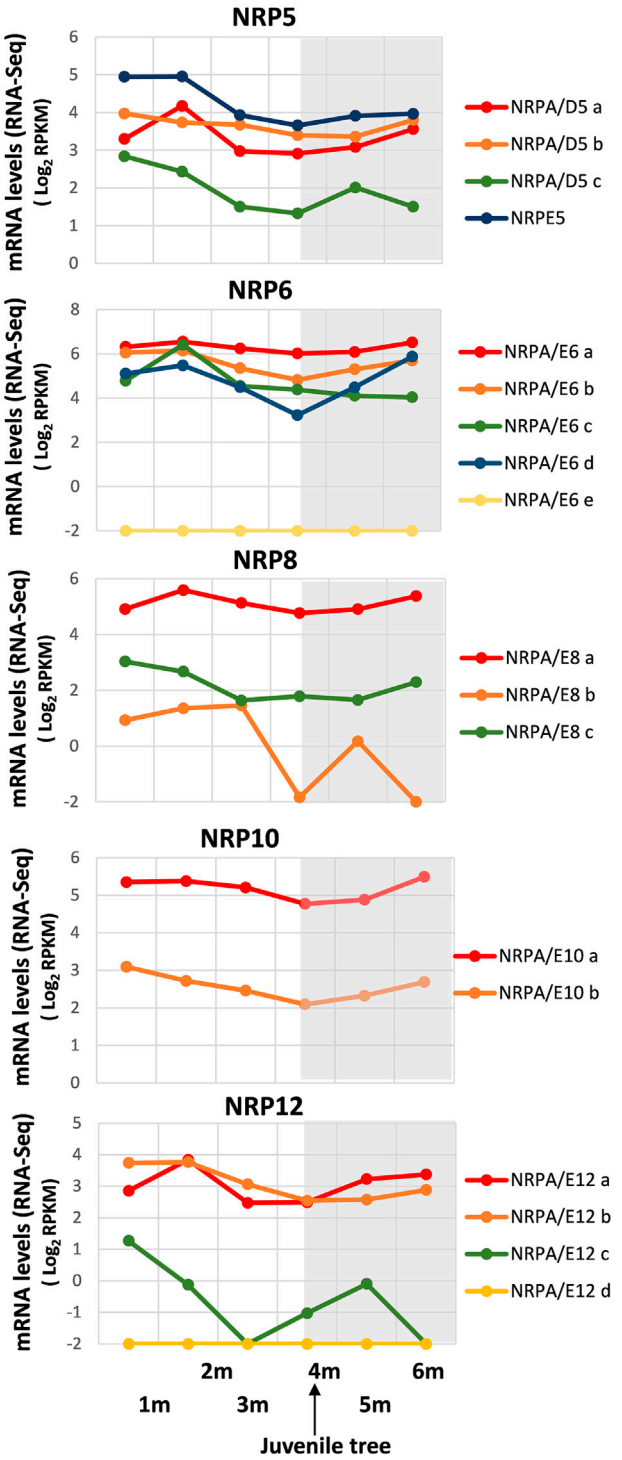


FIGURE 3 | Gene expression pattern of subunits common to RNA polymerases during early development from germinated embryos to juvenile trees. Data from the RNA-Seq datasets (Jiménez-Ruiz et al., 2018). Expression profile of the different NRP genes from olive cultivar "Picual" during early development 1, 2, 3, 4, 5, and 6 months after seed activation. Data correspond to reads per kilobase of transcript per million mapped reads (RPKM). 4 months corresponds to the time considered for plant to be a juvenile tree. Data are represented as $\log_2$ (RPKM). Values ≤ −2 are considered no expression and represented by a value of −2. The gray square represents the late developmental period once plants were juvenile trees.

(*Continued*)

increased or remained unaltered, except for NRPA/E6c, whose slow decrease in gene expression continued from 2 development months. In addition, the NRPA/D5c expression pattern differed by lowering between 5 and 6 development months. This also occurred for NRPA/E8b and NRPA/E12c, although their expression drastically dropped, or even disappeared, after 3 and 1 development months, respectively, before increasing to 5 development months.

Furthermore, we corroborated that genes NRPA/E6e and NRPA/E12c were not expressed during development from germinated embryos to juvenile trees, and NRPE5, NRPA/E8a, NRPA/E10a, and NRPA/E12a and b were the most expressed gene paralogs, as observed under other conditions analyzed herein. Altogether, these data demonstrated that most NRP common subunit genes from olive were expressed and were spatially and temporally regulated.

The Genes for Subunits Shared by RNA Polymerases are Regulated by Stress Conditions

Biotic and abiotic stresses impact olive tree cultivars, leading to vast economic loss and agronomic damage (López-Escudero and Mercado-Blanco, 2011; Trapero et al., 2013). Accordingly, we investigated whether NRP common subunit genes were expressed under biotic and abiotic stresses.

A wide variety of biotic constraints affects olive cultivation, including *Verticillium* wilt of olive caused by the pathogenic fungus *V. dahliae*, which is detected in almost all the regions where olive culture exists, and is one of the most harmful diseases that affect this woody crop, leading to vast economic loss and agronomic damage, particularly in the Mediterranean Basin (López-Escudero and Mercado-Blanco, 2011). Most olive tree cultivars are susceptible to this disease, including "Picual" (Trapero et al., 2013).

We collected data from recently published genome-wide transcriptomic studies conducted during infection and the plant–*V. dahliae* interaction (Jiménez-Ruiz et al., 2017; Leyva-Pérez et al., 2018), we investigated the expression pattern of genes for subunits shared by RNA polymerases from olive cultivars associated with biotic stress during *V. dahliae* root infection. We also used the RNA-Seq data from the total RNA extracted from the roots of three groups of three randomly selected plants after 48 h and 7 days of infection, as well as from the control plants, taken as time 0 before infection (Jiménez-Ruiz et al., 2017). As shown in **Figure 4**, the analysis of the control plants (time 0 before infection) corroborated our previous results observed in roots (**Figure 2**). Notably, in the susceptible "Picual" cultivar, the gene expression for most NRP subunits decreased during infection and plant–*V. dahliae* interaction (**Figure 4**). On the contrary, in the resistant cultivar "Frantoio," none of the NRP common subunit genes decreased their expression during

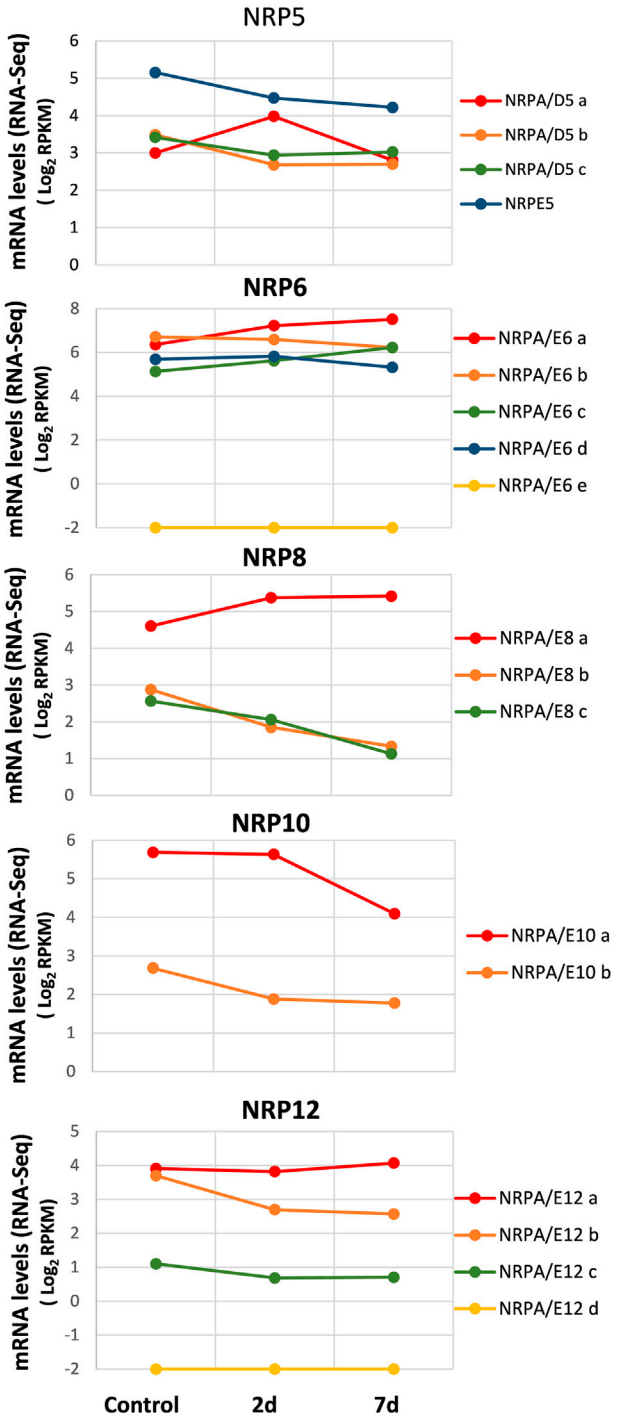


FIGURE 4 | Gene expression pattern of subunits common to RNA polymerases during the *V. dahliae* early infection process. Data from previously published RNA-Seq datasets (Jiménez-Ruiz et al., 2017). Expression profile of the different NRP genes from olive cultivar "Picual" 1 and 7 days after *V. dahliae* infection. The control corresponds to the control group of non-inoculated plants, handled in the same way as in the absence of the pathogen. Data correspond to reads per kilobase of transcript per million mapped reads (RPKM). Data are represented as Log_2(RPKM). Values ≤ −2 are considered no expression and represented by a value of −2.

V. dahliae infection, suggesting that this differential expression may contribute to the resistance or sensitivity to *V. dahliae* infection (**Supplemental Figure S4**; compare with **Figure 4**). However, some exceptions were observed in the "Picual" NRP expression pattern (**Figure 4**). This was the case of the NRPA/D5a gene, which was upregulated 2 days after infection, and the levels lowered to those noted in the control (time 0) at seven days. The expression of genes NRPA/E6a, NRPA/E6c, and NRPA/E8a increased during infection, which suggests a specific response to *V. dahliae* infection in olive cultivars. Furthermore, NRPA/E12a, which was the most expressed of the four NRPA/E12 genes, showed no significant wide variation in gene expression during *V. dahliae* infection. This behavior differed from that of the other three NRP12 genes, whose expression decreased or was absent (NRPA/E12d). Notably, as observed before for the different organs and tissues, NRPE5, NRPA/E8a, and NRPA/E10a were still the most expressed genes among their paralogs, while NRPA/E6e was not expressed at all, which was also the case for the NRPA/E12d gene, as indicated before.

These results demonstrated the expression of most NRP common subunit genes during biotic stress by *V. dahliae* infection and suggest some NRPs' major contribution to this response.

To further investigate the expression of NRP common subunit genes under abiotic stress, we have paid special attention to the olive cultivar response to abiotic cold stress as olive is sensitive to winter chilling temperatures, with severe leaf damage occurring at −7°C (D'Angeli et al., 2003). In this situation, an adaptive response, called cold acclimation (Chinnusamy et al., 2006), has evolved to overcome damage related to this abiotic stress.

In order to explore the expression pattern of genes for common subunits of RNA polymerases from the olive cultivar in response to abiotic cold stress, we used the RNA-Seq data from the whole-transcriptome analysis of cold acclimation in "Picual" plant leaves (Leyva-Pérez et al., 2015). To do so, the acclimated plants were subjected to cold stress, and aerial tissues (leaves) were harvested at 0 h (control), 24 h, and 10 days after cold stress (Leyva-Pérez et al., 2015). It is worth noting that cold stress symptoms were detected after 24 h of treatment, and plants completely recovered after 5 days (Leyva-Pérez et al., 2015).

As observed in **Figure 5**, our results at time 0 (control) corroborated mainly the aforementioned data in leaves (**Figure 2**). Furthermore, as observed in previously mentioned analyses, some NRP genes appeared to be the most expressed of their paralogs: NRPE5, NRPA/E8a, NRPA/E10a, and NRPA/E12a and b. In addition, as in previous analyses, NRPA/E6e and NRPA/E12d were not expressed. It is worth noting that these two features were also observed for all conditions analyzed herein. Interestingly, most expressed NRP genes showed a similar general response and were induced during cold acclimation (**Figure 5**). On the contrary, NRPA/E10b gene expression suggested not responding to this abiotic stress. This feature did not seem to be the result of NRPA/E10b gene constitutive expression because this gene responded to *V. dahliae* infection (**Figure 4**).

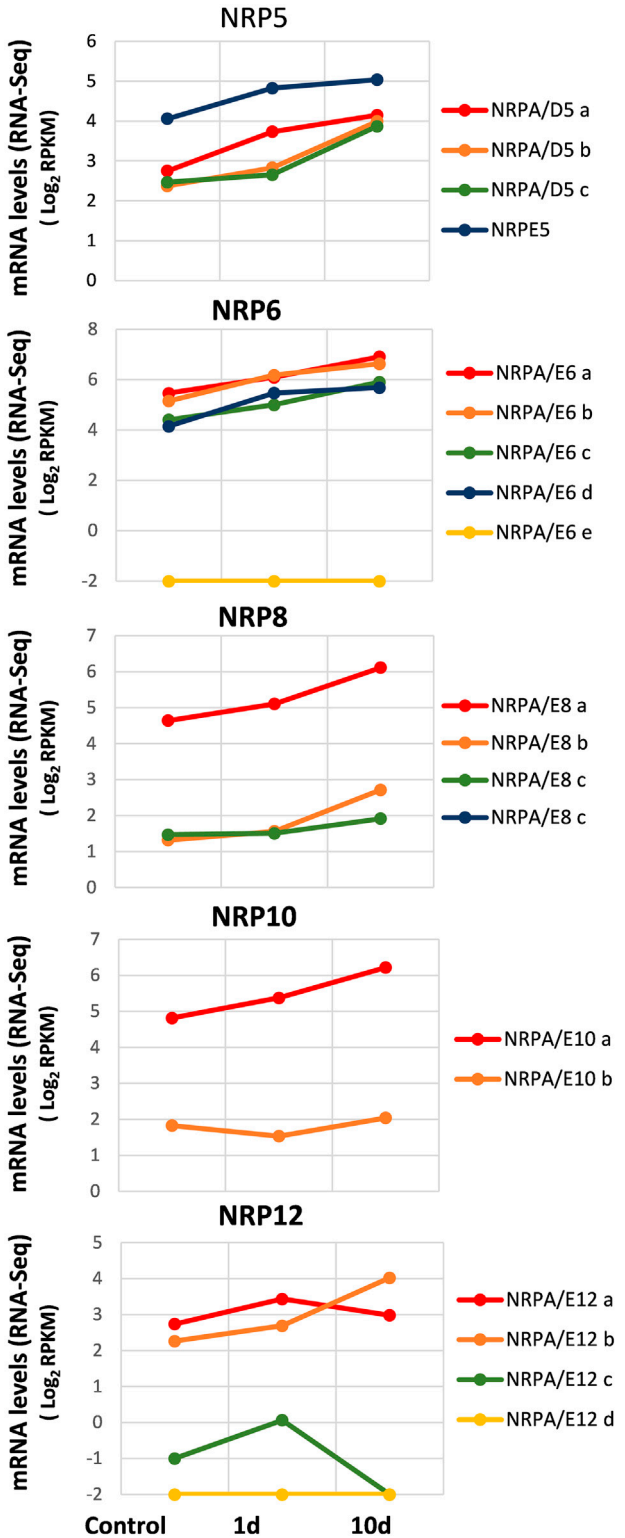


FIGURE 5 | Gene expression pattern of subunits common to RNA polymerases under abiotic stress during cold acclimation. Data from previously published RNA-Seq datasets (Leyva-Pérez et al., 2015). Expression profile of the different NRP genes from olive cultivar "Picual" 1 and 10 days after cold acclimation. Control corresponds to the control group of acclimated plants at time 0. Data correspond to reads per kilobase of transcript per million mapped reads (RPKM). Data are represented as $\log_2$ (RPKM). Values ≤ –2 are considered no expression and represented by a value of –2.

The NRP12 gene paralogs showed the most complex differentiated expression pattern (**Figure 5**), similar to that observed during *V. dahliae* infection (**Figure 4**). As shown, while NRPA/E12b gene expression was induced by cold acclimation, NRPA/E12a gene expression did not significantly alter. Strikingly, NRPA/E12c gene expression disappeared after 10 cold acclimation days.

Our results indicated a major global response to cold acclimation that resulted in NRP gene expression increasing mainly with time. Globally, our results demonstrated a major global transcriptional regulatory response to biotic and abiotic stresses.

DISCUSSION

Plants present a high genomic plasticity, and many species are polyploids or have been polyploids during some evolutionary events, such as olive trees. This fact is why unique genes in other eukaryote groups frequently have some paralog genes in plants. Duplicated genes may evolve and be silenced, or be specialized in a specific condition. Plants have two additional RNA pols (IV and V) to the general three-eukaryote ones as specialized enzymes that have evolved from RNA pol II. All eukaryotic RNA pols share five common subunits, which are mostly coded by unique genes. However, this is not true for plants, containing several paralog genes for these NRPs. Based on the marked agronomic, economic, and ecological interest of olive trees, we searched for genes of subunits shared by RNA polymerases (RNA pols are major elements in gene expression regulation) and studied if the different genes coding for each subunit were regulated and differentially expressed. For this reason and based on the recently reported cultivar "Picual" genome (Jiménez-Ruiz et al., 2020), we analyzed the composition of genes for common subunits of RNA polymerases and their expression patterns in several situations of interest, such as early development, organ/tissue profile, and biotic or abiotic stresses. Furthermore, globally analyzing their expression can help elucidate not only their contribution but also that of RNA pols to global transcriptional responses of interest to the cultivar "Picual."

We identified distinct genes for all five subunits shared by RNA pols (**Table 1**). These results fall in line with those previously described for Arabidopsis and maize, or other angiosperms and gymnosperm plants (Haag et al., 2014; Wang and Ma, 2015; Ream et al., 2009), which suggests high divergence and large differences in evolutionary gene patterns for these different gene subunits (Wang and Ma, 2015). This feature must apply not only to RNA pols common subunits but also to any RNA pol subunits (Ream et al., 2009; Tucker et al., 2010; Ream et al., 2013; Haag et al., 2014; Ream et al., 2015; Wang and Ma, 2015). However, except for subunits five and ten, "Picual" possesses more genes than other plants for subunits 6 (five), 8 (four), and 12 (four) (Ream et al., 2009; Haag et al., 2014; Wang and Ma, 2015). These results agree with the olive cultivar genome resulting from two independent whole-genome duplication (WGD) events during domestication dating back some 62 and 25 million years ago, in addition to very

recent partial genome duplications (Unver et al., 2017; Jiménez-Ruiz et al., 2020). On the contrary, angiosperms like Arabidopsis and maize genomes have resulted from one WGD and partial genome duplications (Jiao et al., 2011; Wang and Ma, 2015; Jiao, 2018; Ren et al., 2018). Only two genes for subunit ten were found in "Picual." This suggest gene loss during evolution after gene duplications, which is described to be a general evolutionary mechanism (Wang and Ma, 2015; Julca et al., 2018). In angiosperms, subunit NRP5 gene duplication seemed to lead to the appearance of a gene coding for subunits NRPA/C5 and/or NRPA/D5 (RNA pols I-III or I-IV, respectively) and a second one for the NRPD/E5 subunit (RNA pols IV/V), whose duplication gave rise to the specific gene for the NRPE5 subunit (RNA pols V) in Arabidopsis (Wang and Ma, 2015). Similarly, in cauliflower, a specific NRPB5b subunit has evolved from an RNA pol II precursor into a functionally different subunit in RNA pol V (Huang et al., 2009). Notably, these occurrences are essential features for RNA pol IV and V specialization from RNA pol II (Tucker et al., 2010; Haag et al., 2014; Huang et al., 2015). The olive cultivar "Picual" only contains one gene that putatively codes for the NRPE5 subunit, according to the amino acid identity with Arabidopsis NRP5 proteins (**Table 1** and **Supplemental Figure S1**), while three genes putatively coding for almost identical subunits (89–95%) showed a closer identity to the NRPA/D5 subunit. Accordingly, these data agree with a "Picual" NRPD/E5 coding gene being duplicated during the ancient WGD in gymnosperms (Jiao et al., 2011; Wang and Ma, 2015; Jiao, 2018; Ren et al., 2018), with the loss of at least one resulting duplicated gene after the second WGD occurring during *Olea–Fraxinus* ancestor speciation (Unver et al., 2017; Jiménez-Ruiz et al., 2020). Furthermore, this ancient WGD has been hypothesized to come from an ancestral allotetraploid produced by the hybridization of an ancestral *Fontanesia*-related species and an ancestral *Jasminum–Forsythia* species (Taylor, 1945; Unver et al., 2017; Jiménez-Ruiz et al., 2020). Conversely, the three "Picual" NRPA/D5 copies account for the very recent partial genome duplication, which may have occurred during domestication.

WGDs seem to account for the reciprocal loss or subfunctionalization of duplicated genes in different species, which enhances the adaptation of organisms to environmental challenge (Hegarty and Hiscock, 2008; Julca et al., 2018; Ren et al., 2018; Qiao et al., 2019). Gene inactivation mechanisms could occur for the "Picual" genes that putatively code for NRPA/E6e and NRPA/E12d as we were unable to detect mRNA expression under any analyzed condition (different organs and tissues, biotic and abiotic stress, plant development). Furthermore, we cannot rule out gene loss for other RNA pol subunits in "Picual" during the domestication period.

Except for the putative coding genes for NRPA/E6e and NRPA/E12d that are indicated earlier, all the other identified genes for the common subunits of RNA pols in "Picual" were expressed, according to our analyses of the different RNA-Seq datasets (Leyva-Pérez et al., 2015; Jiménez-Ruiz et al., 2017; Jiménez-Ruiz et al., 2018; Ramírez-Tejero et al., 2020). However, we cannot rule out that NRPA/E6e and NRPA/E12d can be expressed under other conditions, although it seems unlikely. Nevertheless, other genes for some Arabidopsis or maize RNA pol subunits have been found to be expressed according to mRNA analyses, but no corresponding proteins have been identified in proteomic and/or biochemical studies (Wang and Ma, 2015).

A holistic expression analysis reveals some interesting findings: *1*) most genes for subunits shared by RNA polymerases show similar expression patterns for most analyzed conditions, which suggests coordinated responses; *2*) global differences in the gene expression levels between the distinct paralogs are chiefly maintained under any of the analyzed conditions, albeit with specific differences for some common subunits of RNA polymerases, which indicate specific expression regulation for those NRP genes; *3*) one gene or two per subunit show the highest expression for any analyzed condition, save the NRP6 paralog genes with similar expression levels. This finding implies that major contribution to general gene expression depends on some gene paralogs; *4*) the NRP common subunit genes show spatial and temporal transcriptional regulation and respond to biotic and abiotic stress. Furthermore, certain specificities exist for each analyzed condition.

In terms of the spatial transcriptional regulation, gene expression of NRP common subunits was variable for different organs and tissues (**Figure 2**), and not only for different subunit genes, but also among diverse paralogs of the same NRP common subunit genes. Similarly, differences in NRP subunit expression in organs and tissues have also been observed for Arabidopsis, maize, and other plants (Ream et al., 2009; Ream et al., 2015; Wang and Ma, 2015). These results suggest that specific expression regulation for those NRP genes may be physiologically relevant in different organs and tissues and similarly in other growth conditions (*see* below).

Given cultivated olive trees' agronomic importance, knowledge of early tree development gene regulation steps is relevant to manipulate and shorten the unproductive juvenile period (Moreno-Alías et al., 2010). Notably, transcriptomic analyses of seedlings during early development show a major alteration of gene expression in the first 3–4 months, and gene expression subsequently grows more stable once juvenile tree development from seed is complete (Jiménez-Ruiz et al., 2018). This also seems to be the case for most genes of subunits shared by RNA polymerases, with a slight general trend for a decrease in gene expression decrease during the 1- to 4-month periods (from seed to juvenile tree) and mostly maintained later. These results could imply greater transcriptional activity during early development from seed to juvenile tree being completed (4–6 months development), after which transcriptional activity lowered and remained the same in juvenile and/or adult trees. However, differences were observed for some genes (**Figure 3**). Our results also suggest a minor contribution of NRPA/E8b and NRPA/E12c to seedling development, which was mostly constrained to very early development steps as they were underexpressed 2 or 3 months after inducing germination. A specific transcriptional response in early development has been described for the genes involved in DNA methylation, which were upregulated during the 6-month follow-up (Jiménez-Ruiz et al., 2018). Interestingly, the role of RNA pols IV and V in RNA-directed DNA methylation has been clearly demonstrated (Huang et al., 2009; Ream et al., 2009; Haag and Pikaard, 2011; Haag et al., 2014; Zhou and Law, 2015) by acting during development (Pikaard and Tucker, 2009; Moo et al., 2012; Haag et al., 2014). However, the NRPE5 gene for RNA pol V (and/or IV) did not specifically and differently modulate its expression. This suggests that RNA pol V transcriptional regulation did not make any major contribution to olive

development or that RNA pols are globally regulated at activity levels by the protein–protein interactions of transcriptional complexes, or even by posttranslational modification, rather than by the gene expression of RNA pol subunits.

Susceptible olive cultivar "Picual" responds to biotic stress provoked by *V. dahliae* infection by initiating a specific transcriptional stress response, similarly to that observed in other plants (Jiménez-Ruiz et al., 2017; Yuan et al., 2020). This complex transcriptional response may involve the regulation of some or all transcriptional machineries. Along this line, most NRP genes transcriptionally respond to this biotic stress as some decrease their expression in the plant–fungi interaction, while others specifically and distinctly respond by showing rapid upregulation at 2-days postinfection, or even remain generally unaltered (**Figure 4**). Notably, plants acquire immunity to pathogen infections, a response that involves the participation of different transcription factors and, at least, RNA pol V (Lopez et al., 2011; Amorim et al., 2017) *via* RdDM mechanism (Huang et al., 2009; Ream et al., 2009; Haag and Pikaard, 2011; Haag et al., 2014; Zhou and Law, 2015). However, subunit gene NRPE5, specific to RNA pol V (and/or IV), did not respond by specifically modulating its expression after *V. dahliae* infection in relation to other NRP5 genes (**Figure 4**) as could be expected. This result poses several considerations, as discussed above, during olive development. First is the notion that RNA pols are regulated at activity levels, although this did not seem to be a general feature of gene regulation due the differential transcriptional response of the NRP genes (**Figure 4**). Another possibility of NRPE5 expression not being specifically altered may suggest that RNA pol V transcriptional regulation did not make any major contribution to "Picual" infection by *V. dahliae* to mediate plant immunity, or could be the consequence of this olive tree being sensitive to *V. dahliae* (Leyva-Pérez et al., 2018). In line with this, some genes implicated in *V. dahliae* infection have been seen to clearly upregulate in olive cultivar "Frantoio" which resists this fungus, while their expression remains unaltered or even decreases in "Picual" (Leyva-Pérez et al., 2018). Notably, none of the "Frantoio" NRP common subunit genes decreased their expression during *V. dahliae* infection, in contrast to the transcriptional response observed for "Picual," suggesting that this differential expression may contribute to the resistance or sensitivity to *V. dahliae* infection (**Supplemental Figure S4**; compare with **Figure 4**).

The NRP genes for common subunits respond to abiotic cold stress (**Figure 5**) in line with general transcriptomic responses of this cultivar to cold (Leyva-Pérez et al., 2015) and with that of many other plants (Matsui et al., 2008; Qi and Zhang, 2019). Olive trees achieve a cold acclimation response that provokes metabolic, physiological, and developmental changes that are genetically controlled (Chinnusamy et al., 2006). In olive leaves, cold acclimation leads to a rapid cold stress response during the first 24-h exposure and a long-term expression response during 10-day cold exposure (Leyva-Pérez et al., 2015). However, while a general gene expression downregulation tendency is observed, the expression of most NRP common subunit genes increased during cold acclimation, with only clear distinct responses occurring for some NRP12 paralogs (**Figure 5**). These results suggest increased transcriptional activity that allowed olive plants to acclimatize and physiologically recover after initial cold stress exposure. We could speculate about this response being accompanied by increased cell cycle progression after cell cycle arrest by cold stress (Qi and Zhang, 2019), although the gene expression analyses performed with "Picual" during cold acclimation have revealed the downregulation of some cell cycle genes (Leyva-Pérez et al., 2015), and we found no alteration or minor increase in some CDK gene expression (not shown). It is worth noting that CDKs have been suggested to be relevant, at least *via* posttranscriptional modulation, during biotic and abiotic stress in plants (Kitsios and Doonan, 2011).

Although our data suggest the scenario of genes for subunits shared by RNA polymerases showing coordinated regulation to mediate the global transcriptional responses observed under different growth conditions or abiotic and biotic stress (Leyva-Pérez et al., 2015; Jiménez-Ruiz et al., 2017; Jiménez-Ruiz et al., 2018; Ramírez-Tejero et al., 2020), we found specific responses of some NRP common subunit paralog genes. These data suggest the contribution of some NRP common subunit genes to the transcriptional regulation mediated by RNA pols for olive "Picual" biology to adapt to different growth situations. Finally, based on our data, we cannot rule out that some NRP common subunit genes code for subunits with RNA pol specificity, which will be the goal of future studies.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

FN and FL contributed to conceptualization. IF-P, JR-T, FL, and FN assisted with Genomic and RNA-seq analyses. FN and FL acquired funding. FN and FL contributed to writing. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work has been supported by grants from the Spanish Ministry of Science and Innovation (MCIN) and ERDF (PID2020-112853GB-C33 to FN and PID2020-115853RR-C33 to FL), the Junta de Andalucía-Universidad de Jaén (FEDER-UJA 1260360) to FN, and the Junta de Andalucía (BIO258) to FN and FL.

ACKNOWLEDGMENTS

We thank the "Servicios Centrales de Apoyo a la Investigación (SCAI)" of the University of Jaen for technical support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.679292/full#supplementary-material

REFERENCES

Amorim, L., Santos, R., Neto, J., Guida-Santos, M., Crovella, S., and Benko-Iseppon, A. (2017). Transcription Factors Involved in Plant Resistance to Pathogens. *Cpps* 18 (4), 335–351. doi:10.2174/1389203717666160619185308

Chinnusamy, V., Zhu, J., and Zhu, J.-K. (2006). Gene Regulation During Cold Acclimation in Plants. *Physiol. Plant* 126 (1), 52–61. doi:10.1111/j.1399-3054.2006.00596.x

Conde, C., Delrot, S., and Gerós, H. (2008). Physiological, Biochemical and Molecular Changes Occurring During Olive Development and Ripening. *J. Plant Physiol.* 165 (15), 1545–1562. doi:10.1016/j.jplph.2008.04.018

Cramer, P. (2019). Eukaryotic Transcription Turns 50. *Cell* 179 (4), 808–812. doi:10.1016/j.cell.2019.09.018

Cuevas-Bermúdez, A., Martínez-Fernández, V., Garrido-Godino, A. I., and Navarro, F. (2017). "Subunits Common to RNA Polymerases," in *The Yeast Role in Medical Applications* (London: W.M.H. Abdulkhair. IntechOpen), 151–165.

D'Angeli, S., Malhó, R., and Altamura, M. M. (2003). Low-temperature Sensing in Olive Tree: Calcium Signalling and Cold Acclimation. *Plant Sci.* 165 (6), 1303–1313. doi:10.1016/S0168-9452(03)00342-X

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., et al. (2008). Phylogeny.fr: Robust Phylogenetic Analysis for the Non-specialist. *Nucleic Acids Res.* 36, W465–W469. doi:10.1093/nar/gkn180

Devaux, S., Lecordier, L., Uzureau, P., Walgraffe, D., Dierick, J. F., Poelvoorde, P., et al. (2006). Characterization of RNA Polymerase II Subunits of *Trypanosoma Brucei*. *Mol. Biochem. Parasitol.* 148 (1), 60–68. doi:10.1016/j.molbiopara.2006.02.020

Donaire, L., Pedrola, L., de la Rosa, R., and Llave, C. (2011). High-throughput Sequencing of RNA Silencing-Associated Small RNAs in Olive (Olea Europaea L.). *PLoS One* 6 (11), e27916. doi:10.1371/journal.pone.0027916

Garrido-Godino, A. I., García-López, M. C., and Navarro, F. (2013). Correct Assembly of RNA Polymerase II Depends on the Foot Domain and Is Required for Multiple Steps of Transcription in *Saccharomyces cerevisiae*. *Mol. Cel. Biol.* 33 (18), 3611–3626. doi:10.1128/MCB.00262-13

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* 59 (3), 307–321. doi:10.1093/sysbio/syq010

Gutiérrez, F., Arnaud, T., and Garrido, A. (2001). Contribution of Polyphenols to the Oxidative Stability of Virgin Olive Oil. *J. Sci. Food Agric.* 81 (15), 1463–1470. doi:10.1002/jsfa.958

Haag, J. R., Brower-Toland, B., Krieger, E. K., Sidorenko, L., Nicora, C. D., Norbeck, A. D., et al. (2014). Functional Diversification of maize RNA Polymerase IV and V Subtypes via Alternative Catalytic Subunits. *Cel. Rep.* 9, 378–390. doi:10.1016/j.celrep.2014.08.067

Haag, J. R., and Pikaard, C. S. (2011). Multisubunit RNA Polymerases IV and V: Purveyors of Non-coding RNA for Plant Gene Silencing. *Nat. Rev. Mol. Cel. Biol.* 12 (8), 483–492. doi:10.1038/nrm3152

Hegarty, M. J., and Hiscock, S. J. (2008). Genomic Clues to the Evolutionary Success of Polyploid Plants. *Curr. Biol.* 18 (10), R435–R444. doi:10.1016/j.cub.2008.03.043

Huang, L., Jones, A. M. E., Searle, I., Patel, K., Vogler, H., Hubner, N. C., et al. (2009). An Atypical RNA Polymerase Involved in RNA Silencing Shares Small Subunits with RNA Polymerase II. *Nat. Struct. Mol. Biol.* 16 (1), 91–93. doi:10.1038/nsmb.1539

Huang, Y., Kendall, T., Forsythe, E. S., Dorantes-Acosta, A., Li, S., Caballero-Pérez, J., et al. (2015). Ancient Origin and Recent Innovations of RNA Polymerase IV and V. *Mol. Biol. Evol.* 32 (7), 1788–1799. doi:10.1093/molbev/msv060

Jiao, Y. (2018). Double the Genome, Double the Fun: Genome Duplications in Angiosperms. *Mol. Plant* 11 (3), 357–358. doi:10.1016/j.molp.2018.02.009

Jiao, Y., Wickett, N. J., Ayyampalayam, S., Chanderbali, A. S., Landherr, L., Ralph, P. E., et al. (2011). Ancestral Polyploidy in Seed Plants and Angiosperms. *Nature* 473 (7345), 97–100. doi:10.1038/nature09916

Jiménez-Ruiz, J., Leyva-Pérez, M. d. l. O., Schilirò, E., Barroso, J. B., Bombarely, A., Mueller, L., et al. (2017). Transcriptomic Analysis of Olea Europaea L. Roots During the *Verticillium dahliae* Early Infection Process. *Plant Genome* 10 (1), 1–15. doi:10.3835/plantgenome2016.07.0060

Jiménez-Ruiz, J., Ramírez-Tejero, J. A., Fernández-Pozo, N., Leyva-Pérez, M. d. l. O., Yan, H., Rosa, R. d. l., et al. (2020). Transposon Activation Is a Major Driver in the Genome Evolution of Cultivated Olive Trees (Olea Europaea L.). *Plant Genome* 13, e20010. doi:10.1002/tpg2.20010

Jiménez-Ruiz, J., de la O Leyva-Pérez, M., Vidoy-Mercado, I., Barceló, A., and Luque, F. (2018). Transcriptomic Time-Series Analysis of Early Development in Olive from Germinated Embryos to Juvenile Tree. *BMC Genomics* 19 (1), 824. doi:10.1186/s12864-018-5232-6

Julca, I., Marcet-Houben, M., Vargas, P., and Gabaldón, T. (2018). Phylogenomics of the Olive Tree (Olea Europaea) Reveals the Relative Contribution of Ancient Allo- and Autopolyploidization Events. *BMC Biol.* 16 (1), 15. doi:10.1186/s12915-018-0482-y

Kelly, S., Wickstead, B., and Gull, K. (2005). An In Silico Analysis of Trypanosomatid RNA Polymerases: Insights into Their Unusual Transcription. *Biochem. Soc. Trans.* 33 (Pt 6), 1435–1437. doi:10.1042/BST20051435

Kitsios, G., and Doonan, J. H. (2011). Cyclin Dependent Protein Kinases and Stress Responses in Plants. *Plant Signaling Behav.* 6 (2), 204–209. doi:10.4161/psb.6.2.14835

Koonin, E. V., Makarova, K. S., and Elkins, J. G. (2007). Orthologs of the Small RPB8 Subunit of the Eukaryotic RNA Polymerases Are Conserved in Hyperthermophilic Crenarchaeota and "Korarchaeota". *Biol. Direct* 2, 38. doi:10.1186/1745-6150-2-38

Kwapisz, M., Beckouët, F., and Thuriaux, P. (2008). Early Evolution of Eukaryotic DNA-dependent RNA Polymerases. *Trends Genet.* 24 (5), 211–215. doi:10.1016/j.tig.2008.02.002

Lanzendorfer, M., Smid, A., Klinger, C., Schultz, P., Sentenac, A., Carles, C., et al. (1997). A Shared Subunit Belongs to the Eukaryotic Core RNA Polymerase. *Genes Dev.* 11 (8), 1037–1047. doi:10.1101/gad.11.8.1037

Larkin, R. M., Hagen, G., and Guilfoyle, T. J. (1999). *Arabidopsis thaliana* RNA Polymerase II Subunits Related to Yeast and Human RPB5. *Gene* 231 (1-2), 41–47. doi:10.1016/s0378-1119(99)00090-6

Law, J. A., Vashisht, A. A., Wohlschlegel, J. A., and Jacobsen, S. E. (2011). SHH1, a Homeodomain Protein Required for DNA Methylation, as Well as RDR2, RDM4, and Chromatin Remodeling Factors, Associate with RNA Polymerase IV. *Plos Genet.* 7 (7), e1002195. doi:10.1371/journal.pgen.1002195

Leyva-Pérez, M. d. l. O., Jiménez-Ruiz, J., Gómez-Lama Cabanás, C., Valverde-Corredor, A., Barroso, J. B., Luque, F., et al. (2018). Tolerance of Olive (Olea Europaea) Cv Frantoio to Verticillium Dahliaerelies on Both Basal and Pathogen-Induced Differential Transcriptomic Responses. *New Phytol.* 217 (2), 671–686. doi:10.1111/nph.14833

Leyva-Pérez, M. d. l. O., Valverde-Corredor, A., Valderrama, R., Jiménez-Ruiz, J., Muñoz-Merida, A., Trelles, O., et al. (2015). Early and Delayed Long-Term Transcriptional Changes and Short-Term Transient Responses During Cold Acclimation in Olive Leaves. *DNA Res.* 22 (1), 1–11. doi:10.1093/dnares/dsu033

López, A., Ramírez, V., García-Andrade, J., Flors, V., and Vera, P. (2011). The RNA Silencing Enzyme RNA Polymerase V Is Required for Plant Immunity. *Plos Genet.* 7 (12), e1002434. doi:10.1371/journal.pgen.1002434

López-Escudero, F. J., and Mercado-Blanco, J. (2011). Verticillium Wilt of Olive: A Case Study to Implement an Integrated Strategy to Control a Soil-Borne Pathogen. *Plant Soil* 344 (1), 1–50. doi:10.1007/s11104-010-0629-2

Loureiro, J., Rodriguez, E., Costa, A., and Santos, C. (2007). Nuclear DNA Content Estimations in Wild Olive (*Olea Europaea L.* Ssp. Europaea Var. Sylvestris Brot.) and Portuguese Cultivars of O. Europaea Using Flow Cytometry. *Genet. Resour. Crop Evol.* 54 (1), 21–25. doi:10.1007/s10722-006-9115-3

Martínez-Fernández, V., Garrido-Godino, A. I., Mirón-García, M. C., Begley, V., Fernández-Pévida, A., de la Cruz, J., et al. (2018). Rpb5 Modulates the RNA Polymerase II Transition from Initiation to Elongation by Influencing Spt5 Association and Backtracking. *Biochim. Biophys. Acta (Bba) - Gene Regul. Mech.* 1861 (1), 1–13. doi:10.1016/j.bbagrm.2017.11.002

Matsui, A., Ishida, J., Morosawa, T., Mochizuki, Y., Kaminuma, E., Endo, T. A., et al. (2008). Arabidopsis Transcriptome Analysis under Drought, Cold,

High-Salinity and ABA Treatment Conditions Using a Tiling Array. *Plant Cel. Physiol.* 49 (8), 1135–1149. doi:10.1093/pcp/pcn101

Minakhin, L., Bhagat, S., Brunning, A., Campbell, E. A., Darst, S. A., Ebright, R. H., et al. (2001). Bacterial RNA Polymerase Subunit and Eukaryotic RNA Polymerase Subunit RPB6 Are Sequence, Structural, and Functional Homologs and Promote RNA Polymerase Assembly. *Proc. Natl. Acad. Sci.* 98 (3), 892–897. doi:10.1073/pnas.98.3.892

Moo, L. d. R. C., González, A. K., Rodríguez-Zapata, L. C., Suarez, V., and Castaño, E. (2012). Expression of RNA Polymerase IV and V in *Oryza Sativa*. *Electron. J. Biotechnol.* 15, 9. doi:10.2225/vol15-issue2-fulltext-8

Moreno-Alías, I., Rapoport, H. F., López, R., León, L., and de la Rosa, R. (2010). Optimizing Early Flowering and Pre-selection for Short Juvenile Period in Olive Seedlings. *Horts* 45, 519–522. doi:10.21273/hortsci.45.4.519

Nouraini, S., Archambault, J., and Friesen, J. D. (1996). Rpo26p, a Subunit Common to Yeast RNA Polymerases, Is Essential for the Assembly of RNA Polymerases I and II and for the Stability of the Largest Subunits of These Enzymes. *Mol. Cel. Biol.* 16 (11), 5985–5996. doi:10.1128/mcb.16.11.5985

Pikaard, C. S., and Tucker, S. (2009). RNA-silencing Enzymes Pol IV and Pol V in maize: More Than One Flavor. *Plos Genet.* 5 (11), e1000736. doi:10.1371/journal.pgen.1000736

Qi, F., and Zhang, F. (2019). Cell Cycle Regulation in the Plant Response to Stress. *Front. Plant Sci.* 10, 1765. doi:10.3389/fpls.2019.01765

Qiao, X., Li, Q., Yin, H., Qi, K., Li, L., Wang, R., et al. (2019). Gene Duplication and Evolution in Recurring Polyploidization-Diploidization Cycles in Plants. *Genome Biol.* 20 (1), 38. doi:10.1186/s13059-019-1650-2

Ramírez-Tejero, J. A., Jiménez-Ruiz, J., Leyva-Pérez, M. d. l. O., Barroso, J. B., and Luque, F. (2020). Gene Expression Pattern in Olive Tree Organs (Olea Europaea L.). *Genes* 11 (5), 544. doi:10.3390/genes11050544

Ream, T., Haag, J., and Pikaard, C. (2013). "Plant Multisubunit RNA Polymerases IV and V," in *Nucleic Acid Polymerases*. Editor K.S.M.A. Trakselis (Heidelberg: Springer-Verlag). doi:10.1007/978-3-642-39796-7_13

Ream, T. S., Haag, J. R., Pontvianne, F., Nicora, C. D., Norbeck, A. D., Paša-Tolić, L., et al. (2015). Subunit Compositions of Arabidopsis RNA Polymerases I and III Reveal Pol I- and Pol III-specific Forms of the AC40 Subunit and Alternative Forms of the C53 Subunit. *Nucleic Acids Res.* 43 (8), 4163–4178. doi:10.1093/nar/gkv247

Ream, T. S., Haag, J. R., Wierzbicki, A. T., Nicora, C. D., Norbeck, A. D., Zhu, J.-K., et al. (2009). Subunit Compositions of the RNA-Silencing Enzymes Pol IV and Pol V Reveal Their Origins as Specialized Forms of RNA Polymerase II. *Mol. Cel.* 33 (2), 192–203. doi:10.1016/j.molcel.2008.12.015

Ren, R., Wang, H., Guo, C., Zhang, N., Zeng, L., Chen, Y., et al. (2018). Widespread Whole Genome Duplications Contribute to Genome Complexity and Species Diversity in Angiosperms. *Mol. Plant* 11 (3), 414–428. doi:10.1016/j.molp.2018.01.002

Rugini, E., Pannelli, G., Ceccarelli, M., and Muganu, M. (1996). Isolation of Triploid and Tetraploid Olive (Olea Europaea L.) Plants from Mixoploid Cv. 'Frantoio' and 'Leccino' Mutants by In Vivo and In Vitro Selection. *Plant Breed.* 115 (1), 23–27. doi:10.1111/j.1439-0523.1996.tb00865.x

Shpakovski, G. V., Acker, J., Wintzerith, M., Lacroix, J. F., Thuriaux, P., and Vigneron, M. (1995). Four Subunits that Are Shared by the Three Classes of RNA Polymerase Are Functionally Interchangeable Between *Homo sapiens* and *Saccharomyces cerevisiae*. *Mol. Cel. Biol.* 15 (9), 4702–4710. doi:10.1128/mcb.15.9.4702

Talhaoui, N., Gómez-Caravaca, A., León, L., De la Rosa, R., Fernández-Gutiérrez, A., and Segura-Carretero, A. (2016). From Olive Fruits to Olive Oil: Phenolic Compound Transfer in Six Different Olive Cultivars Grown under the Same Agronomical Conditions. *Ijms* 17 (3), 337. doi:10.3390/ijms17030337

Taylor, H. (1945). Cyto-taxonomy and Phylogeny of the Oleaceae. *Brittonia* 5 (4), 337–367. doi:10.2307/2804889

Trapero, C., Serrano, N., Arquero, O., Del Río, C., Trapero, A., and López-Escudero, F. J. (2013). Field Resistance to Verticillium Wilt in Selected Olive Cultivars Grown in Two Naturally Infested Soils. *Plant Dis.* 97 (5), 668–674. doi:10.1094/PDIS-07-12-0654-RE

Tucker, S. L., Reece, J., Ream, T. S., and Pikaard, C. S. (2010). Evolutionary History of Plant Multisubunit RNA Polymerases IV and V: Subunit Origins via Genome-wide and Segmental Gene Duplications, Retrotransposition, and Lineage-specific Subfunctionalization. *Cold Spring Harbor Symposia Quantitative Biol.* 75, 285–297. doi:10.1101/sqb.2010.75.037

Unver, T., Wu, Z., Sterck, L., Turktas, M., Lohaus, R., Li, Z., et al. (2017). Genome of Wild Olive and the Evolution of Oil Biosynthesis. *Proc. Natl. Acad. Sci. USA* 114 (44), E9413–E9422. doi:10.1073/pnas.1708621114

Wang, Y., and Ma, H. (2015). Step-wise and Lineage-specific Diversification of Plant RNA Polymerase Genes and Origin of the Largest Plant-specific Subunits. *New Phytol.* 207 (4), 1198–1212. doi:10.1111/nph.13432

Werner, F., and Grohmann, D. (2011). Evolution of Multisubunit RNA Polymerases in the Three Domains of Life. *Nat. Rev. Microbiol.* 9 (2), 85–98. doi:10.1038/nrmicro2507

Werner, M., Thuriaux, P., and Soutourina, J. (2009). Structure-function Analysis of RNA Polymerases I and III. *Curr. Opin. Struct. Biol.* 19 (6), 740–745. doi:10.1016/j.sbi.2009.10.005

Wierzbicki, A. T., Haag, J. R., and Pikaard, C. S. (2008). Noncoding Transcription by RNA Polymerase Pol IVb/Pol V Mediates Transcriptional Silencing of Overlapping and Adjacent Genes. *Cell* 135 (4), 635–648. doi:10.1016/j.cell.2008.09.035

Woychik, N. A., Liao, S. M., Kolodziej, P. A., and Young, R. A. (1990). Subunits Shared by Eukaryotic Nuclear RNA Polymerases. *Genes Dev.* 4 (3), 313–323. doi:10.1101/gad.4.3.313

Yuan, G., He, X., Li, H., Xiang, K., Liu, L., Zou, C., et al. (2020). Transcriptomic Responses in Resistant and Susceptible Maize Infected with Fusarium Graminearum. *Crop J.* 8 (1), 153–163. doi:10.1016/j.cj.2019.05.008

Zaros, C., Briand, J.-F., Boulard, Y., Labarre-Mariotte, S., García-López, M. C., Thuriaux, P., et al. (2007). Functional Organization of the Rpb5 Subunit Shared by the Three Yeast RNA Polymerases. *Nucleic Acids Res.* 35 (2), 634–647. doi:10.1093/nar/gkl686

Zhang, H., Ma, Z.-Y., Zeng, L., Tanaka, K., Zhang, C.-J., Ma, J., et al. (2013). DTF1 Is a Core Component of RNA-Directed DNA Methylation and May Assist in the Recruitment of Pol IV. *Proc. Natl. Acad. Sci.* 110 (20), 8290–8295. doi:10.1073/pnas.1300585110

Zhou, M., and Law, J. A. (2015). RNA Pol IV and V in Gene Silencing: Rebel Polymerases Evolving Away from Pol II's Rules. *Curr. Opin. Plant Biol.* 27, 154–164. doi:10.1016/j.pbi.2015.07.005

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership