



HAL
open science

New insights into the mechanisms of transcription termination by RNA polymerase III

Juanjuan Xie

► **To cite this version:**

Juanjuan Xie. New insights into the mechanisms of transcription termination by RNA polymerase III. Genetics. Université Paris Cité, 2021. English. NNT : 2021UNIP7321 . tel-04100280

HAL Id: tel-04100280

<https://theses.hal.science/tel-04100280>

Submitted on 17 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Paris

École doctorale [ED 562 Bio Sorbonne Paris Cité](#)
Laboratoire [Institut Jacques Monod](#)

New Insights into the Mechanisms of Transcription Termination by RNA Polymerase III

Par

XIE Juanjuan

Thèse de doctorat de [GENETIQUE](#)

Présentée et soutenue publiquement le [19 Octobre 2021](#)

Composition du Jury:

Directeurs de thèse

| | |
|--------------------|-------------------------------------|
| PORRUA FUERTE Odil | HDR, Institut Jacques Monod, France |
| LIBRI Domenico | DR, Institut Jacques Monod, France |

Rapporteurs

| | |
|------------------|----------------------------------|
| HERMAND Damien | DR, University of Namur, Belgium |
| SOUTOURINA Julie | DR, CEA Paris-Saclay, France |

Examineurs

| | |
|----------------|---------------------------------|
| JACQUIER Alain | DR, Institut Pasteur, France |
| LESAGE Pascale | DR, Université de Paris, France |

Membre invité du jury

| | |
|----------------|-------------------------------|
| SENTENAC André | HDR, CEA Paris-Saclay, France |
|----------------|-------------------------------|

Acknowledgements

Firstly, I would like to thank my supervisors, Dr. Odil Porrua and Dr. Domenico Libri, for mentoring me in every stage of my PhD. Dom, thank you for having accepted me to join your laboratory. Odil, thanks for your time and help to my experiments and postdoc application. Thank both of you for your patience, support, encouragement, as well as your help and advice to not only the research project, but also my life in Paris and my personal development. I have learned a lot from all of our discussions and I'm very grateful for my time working with you.

I would like to thank people of the Lab, Tommaso Villa, Umberto Aiello, Yves Clement, Griselda Wentzinger, and former members, Zhong Han, Drice Challal, Nouhou Haidara and Julien Gros, for your help with my experiments, your inspiring discussions and advice for my project, and thank you for always being very kind to me.

Thank you to Dr. Michel Werner and Dr. Benoît Palancade of our institute, for the insightful comments and suggestions to my work during our joint lab meetings. And thanks to Dr. Christoph Müller and Mathias Girbig of the EMBL Heidelberg, for the delighted collaboration and help to my project.

I am very grateful to my thesis committee members, Dr. Vincent Vanoosthuyse of the ENS Lyon, Dr. Olivier Lefebvre of the CEA Paris, and Dr. Sebastien Leon of our institute, thank you for giving me valuable feedback and suggestions during our three meetings.

I would like to express my gratitude to the PhD jury members, Dr. Damien Hermand, Dr. Julie Soutourina, Dr. Alain Jacquier, Dr. Pascale Lesage and Dr. André Sentenac, for taking the time to read, comment and evaluate my work. I am so honored to have all of you in the committee.

Thank you to my parents and brother, for your endless support and your belief in me. I'm sorry I wasn't there when you needed me the most during the past four years, thank you for all your understanding. Thanks to my best friend, Yi, for your trust and making me feel important, thank you for having been with me for 16 years.

I would also like to give special thanks to my country, for the support from China Scholarship Council, for the timely help and consideration from the Educational Office of Chinese Embassy in Paris.

Abstract

Transcription is the process whereby RNA molecules are synthesized using DNA as the template. In eukaryotic cells, this process is carried out by three different kinds of RNA polymerases (RNAP). RNAPI is mainly responsible for the production of rRNA, RNAPII transcribes protein-coding genes and several classes of non-coding genes and RNAPIII is mostly dedicated to the synthesis of tRNAs, 5S rRNA as well as a few short non-coding RNAs. The genomic distribution of different polymerases needs to be tightly controlled to avoid disruptive interferences between adjacent transcription events, which largely depends on the process of transcription termination. The current model posits that RNAPIII termination relies solely on a stretch of thymines (T-tract) in the non-template DNA strand that are presumably sufficient for both RNAPIII pausing and its release from the DNA. However, previous results from our group identified an interaction between RNAPIII and a well-characterized RNAPII transcription termination factor, the helicase Sen1, which prompted us to investigate a possible role for Sen1 in the termination of RNAPIII transcription in budding yeast.

In this study, to address specifically the function of Sen1 in RNAPIII transcription termination I have employed a Sen1 mutant (*sen1-3*) containing three point mutations in a conserved region of the Sen1 N-terminal domain, which are sufficient to abolish the interaction with RNAPIII without affecting RNAPII transcription termination. By generating high-resolution maps of transcribing RNAPIII, I have observed that a significant fraction of RNAPIIIs normally read through the primary terminator (i.e. the first T-tract downstream of the 3' end of genes). Importantly, I have shown that the mutations in *sen1-3* induce transcription termination defects at most RNAPIII-dependent genes, indicating that Sen1 is globally required for efficient termination of RNAPIII transcription *in vivo*, and that this function relies on the interaction of its N-terminal domain with RNAPIII. In addition, I have shown that Sen1 acts mainly at read-through regions as a fail-safe mechanism to promote termination of RNAPIIIs that override the primary terminator.

In order to explore whether Sen1 can directly induce the release of RNAPIII from the DNA, I have performed *in vitro* transcription termination assays with purified proteins. First, I have tested the termination efficiency at T-tracts of different lengths and I have shown that six consecutive thymines are sufficient for RNAPIII transcription termination *in vitro*. Furthermore, I have demonstrated that Sen1 can promote termination at weak terminators (i.e. containing 4 or 5 Ts), as well as at other kinds of pausing sequences. By analysing the function of several Sen1 variants, I have shown that the helicase domain of Sen1 alone can induce transcription termination of RNAPIII *in vitro* as the full-length protein. Moreover, I

have found that Sen1-mediated termination requires Sen1 binding to the RNA and Sen1 ATP hydrolysis activity. Our results indicate that Sen1 employs a similar mechanism to promote transcription termination of both RNAPII and RNAPIII.

Finally, I have found that hairpin-like structures that typically form at RNAPIII transcripts near the primary terminator can also stimulate RNAPIII release at weak terminators. Moreover, I have obtained evidence that hairpin-like structures can form inside the polymerase. I have also shown that the presence of these structures hinders the loading of Sen1 onto the RNA, indicating that Sen1 and RNA hairpins work in a mutually exclusive manner. While Sen1 can promote release of RNAPIII at pausing sites other than T-tracts, I have found that RNA structures can only work together with a canonical termination sequence.

Taken together, our data offer a detailed model for RNAPIII transcription termination, which in budding yeast involves the cooperation of T-tracts, RNA secondary structures and the helicase Sen1.

Keywords: RNA polymerase III, transcription termination, T-tract, helicase Sen1, RNA secondary structure

Résumé

La transcription est le processus par lequel les molécules d'ARN sont synthétisées en utilisant l'ADN comme matrice. Chez les eucaryotes, ce processus est réalisé par trois types d'ARN polymérases (ARNpol). L'ARNpol I est responsable de la production d'ARNr, l'ARNpol II transcrit les gènes codant pour de protéines et certains gènes non-codants et l'ARNpol III est surtout dédiée à la synthèse des ARNt, de l'ARNr 5S ainsi que d'autres ARN non-codants courts. La distribution des différentes polymérases dans le génome doit être finement contrôlée pour éviter les interférences entre les gènes adjacents, ce qui dépend largement de la terminaison de la transcription. Selon le modèle actuel, la terminaison de l'ARNpol III repose uniquement sur une suite de thymines (T) dans le brin d'ADN non-matrice qui serait suffisant pour induire une pause de l'ARNpol III et sa dissociation de l'ADN. Cependant, mon groupe a précédemment trouvé une interaction entre l'ARNpol III et l'hélicase Sen1, un facteur de terminaison de la transcription de l'ARNpol II bien caractérisé, ce qui nous a incité à étudier un rôle possible de Sen1 dans la terminaison de la transcription de l'ARNpol III chez la levure bourgeonnante.

Pour élucider la fonction spécifique de Sen1 dans la terminaison de l'ARNpol III, j'ai utilisé un mutant (*sen1-3*) contenant trois mutations ponctuelles dans le domaine N-terminal de Sen1, qui sont suffisantes pour abolir son interaction avec l'ARNpol III sans affecter la terminaison de la transcription de l'ARNpol II. En générant des cartes à haute résolution de la transcription par l'ARNpol III, j'ai observé qu'une fraction significative d'ARNpol III lit normalement à travers le terminateur primaire (c'est-à-dire la première suite de Ts en aval de l'extrémité 3' des gènes). J'ai montré que les mutations dans *sen1-3* induisent des défauts de terminaison dans la plupart des gènes dépendants de l'ARNpol III, ce qui indique que l'interaction de Sen1 avec l'ARNpol III est globalement requise pour une terminaison efficace de la transcription de l'ARNpol III *in vivo*. De plus, j'ai montré que Sen1 agit principalement comme un mécanisme de "fail-safe" pour promouvoir la terminaison des ARNpol IIIs qui dépassent le terminateur primaire.

Afin d'explorer si Sen1 peut induire directement la dissociation de l'ARNpol III de l'ADN, j'ai réalisé des essais de terminaison de la transcription *in vitro* avec des protéines purifiées. Tout d'abord, j'ai montré que six Ts consécutives sont nécessaires pour une terminaison efficace de la transcription par l'ARNpol III *in vitro*. De plus, j'ai démontré que Sen1 peut promouvoir la terminaison au niveau des terminateurs faibles contenant 4 ou 5 Ts, ainsi que d'autres types de séquences de pause. Ensuite, j'ai montré que le domaine hélicase de Sen1 peut induire la terminaison de l'ARNpol III de façon similaire à la protéine entière *in vitro*. De

plus, j'ai découvert que la terminaison médiée par Sen1 nécessite la liaison de Sen1 à l'ARN et l'activité ATPasique de Sen1, comme montré précédemment pour la terminaison de la transcription de l'ARNpol II.

Enfin, j'ai découvert que les structures secondaires qui se trouvent généralement sur les ARN transcrits par l'ARNpol III peuvent aussi compléter la fonction des terminateurs faibles. J'ai également montré que la présence de ces structures empêche l'interaction de Sen1 avec l'ARN, ce qui indique que Sen1 et les structures de l'ARN fonctionnent d'une manière mutuellement exclusive. Alors que Sen1 peut favoriser le relâchement de l'ARNpol III de différentes types de séquences, les structures d'ARN ne peuvent fonctionner qu'avec une séquence de terminaison canonique.

Ensemble, nos données offrent un modèle détaillé pour la terminaison de la transcription de ARNpol III qui implique la coopération des suites de Ts, des structures secondaires d'ARN et de l'hélicase Sen1.

Mots clés: ARNpol III, terminaison de la transcription, suite de T, l'hélicase Sen1, Structures secondaires d'ARN

Table of Contents

| | |
|---|------------|
| Acknowledgements | i |
| Abstract | iii |
| Résumé | v |
| Table of Contents | vii |
| List of Figures | x |
| List of Tables | xi |
| List of Boxes | xii |
| INTRODUCTION | 1 |
| Chapter 1 General introduction of RNA polymerase and transcription | 2 |
| 1.1 Discovery of the RNA polymerase | 3 |
| 1.2 Bacterial RNA polymerase | 3 |
| 1.3 Eukaryotic RNA polymerases..... | 4 |
| 1.4 Archaeal RNA polymerase | 5 |
| 1.5 RNA polymerase IV and V..... | 6 |
| 1.6 Transcription cycle | 7 |
| 1.7 Yeast as a model organism | 8 |
| Chapter 2 Transcription by RNA polymerase II | 9 |
| 2.1 Transcription initiation | 9 |
| 2.1.1 The RNAPII core promoter..... | 9 |
| 2.1.2 PIC assembly | 10 |
| 2.2 Transcription elongation | 15 |
| 2.2.1 Promoter clearance | 15 |
| 2.2.2 Elongation in the body of genes | 15 |
| 2.3 The phosphorylation cycle of the RNAPII CTD..... | 16 |
| 2.4 Chromatin dynamics during transcription elongation..... | 18 |
| 2.4.1 Histone modifiers | 18 |
| 2.4.2 Histone chaperones..... | 20 |
| 2.4.3 ATP-dependent chromatin remodelers..... | 20 |

| | | |
|------------------|--|-----------|
| 2.5 | Transcription termination | 21 |
| 2.5.1 | The CPF-CF pathway | 21 |
| 2.5.2 | The NNS-dependent pathway..... | 26 |
| 2.5.3 | Alternative pathways for RNAPII termination | 32 |
| Chapter 3 | Transcription by RNA polymerase III | 37 |
| 3.1 | The RNAPIII structure | 37 |
| 3.2 | The RNAPIII transcriptome | 39 |
| 3.3 | Transcription initiation | 44 |
| 3.3.1 | Basal RNAPIII transcription factors | 44 |
| 3.3.2 | Promoters of class III genes | 46 |
| 3.3.3 | Recruitment of TFIIC to the promoter | 48 |
| 3.3.4 | Recruitment of TFIIB by TFIIC | 49 |
| 3.3.5 | Recruitment of RNAPIII and assembly of the PIC..... | 51 |
| 3.3.6 | Promoter opening..... | 52 |
| 3.4 | Transcription elongation | 56 |
| 3.5 | Transcription termination | 58 |
| 3.5.1 | T-tract-dependent termination | 58 |
| 3.5.2 | The role of RNA secondary structures in termination | 63 |
| 3.5.3 | Extrinsic termination factors | 64 |
| 3.6 | Facilitated reinitiation by RNAPIII..... | 65 |
| | Research outline and main objectives | 68 |
| | MANUSCRIPT | 70 |
| | An integrated model for termination of RNA polymerase III transcription | 71 |
| | Abstract | 72 |
| | Introduction | 73 |
| | Results | 76 |
| | The N-terminal domain of Sen1 interacts with RNAPIII..... | 76 |
| | Sen1 is required for efficient termination of RNAPIII transcription <i>in vivo</i> | 80 |
| | Sen1 functions in RNAPIII transcription independently of the NNS-complex | 82 |
| | The function of Sen1 in RNAPIII transcription termination is not mediated by the replisome | 82 |
| | Sen1 operates in a fail-safe transcription termination pathway | 84 |
| | Sen1 can promote termination of RNAPIII transcription <i>in vitro</i> | 88 |
| | Sen1 employs a similar mechanism to terminate transcription of RNAPII and RNAPIII | 92 |
| | RNA structures upstream of T-tracts can promote the release of paused RNAPIIs | 94 |

| | |
|---|------------|
| RNA secondary structures can form within RNAPIII | 97 |
| Discussion | 99 |
| <i>S. cerevisiae</i> Sen1 is a fail-safe transcription termination factor for RNAPIII | 99 |
| The mechanism of Sen1-dependent RNAPIII transcription termination | 100 |
| RNA structures are enhancers of canonical termination signals | 101 |
| Multiple mechanisms partake in RNAPIII transcription termination | 103 |
| Supplementary material | 105 |
| Methods | 119 |
| Construction of yeast strains and plasmids | 119 |
| Co-immunoprecipitation (Co-IP)..... | 119 |
| Mass spectrometry analysis and label-free quantification | 119 |
| UV crosslinking and analysis of cDNA (CRAC) | 120 |
| Synchronization of cells in G1 and analysis by flow cytometry | 121 |
| Dataset processing..... | 122 |
| Bioinformatic analyses..... | 122 |
| RNA analyses | 123 |
| Protein purification..... | 124 |
| <i>In vitro</i> transcription termination assays..... | 125 |
| Data availability | 127 |
| Acknowledgements | 128 |
| Funding | 129 |
| Authors contributions | 130 |
| References | 131 |
| DISCUSSION & PERSPECTIVES | 135 |
| What is the Sen1-RNAPIII interaction surface? | 136 |
| How and when is Sen1 recruited to RNAPIII? | 138 |
| Structural differences between <i>S. cerevisiae</i> and <i>S. pombe</i> Sen1 proteins: an implication in functional divergence? | 139 |
| Uneven distribution of RNAPIII along tRNAs: bias in data analysis? | 145 |
| BIBLIOGRAPHY | 147 |
| RESUME DE THESE | 168 |
| Étude sur les différents mécanismes de termination de la transcription par l'ARNpol III | 169 |

List of Figures

| | |
|---|----|
| Figure 1-1: Overview of the main classes of RNA in eukaryotes and their role in gene expression. | 2 |
| Figure 1-2: Scheme of the RNA polymerase subunits composition in the three domains of life..... | 4 |
| Figure 2-1: The core promoter elements for transcription by RNAPII. | 10 |
| Figure 2-2: <i>In vitro</i> model of RNAPII transcription initiation..... | 11 |
| Figure 2-3: The composition and conservation of the CTD. | 17 |
| Figure 2-4: The CTD phosphorylation patterns across protein-coding genes in humans and yeast. | 18 |
| Figure 2-5: Cleavage and polyadenylation complex and its binding sequences. | 22 |
| Figure 2-6: Major pathways for RNAPII transcription termination. | 24 |
| Figure 2-7: The domain structure of the Nab3 and Nrd1 proteins from <i>S. cerevisiae</i> | 26 |
| Figure 2-8: The domain organization and structural features of Sen1..... | 30 |
| Figure 2-9: Alternative pathways of RNAPII transcription termination. | 33 |
| Figure 2-10: Transcription termination in bacteria. | 34 |
| Figure 2-11: Transcription termination by RNAPI in yeast. | 36 |
| Figure 3-1: Structures of eukaryotic RNA polymerases..... | 38 |
| Figure 3-2: tRNA structure, isotypes and isoacceptors. | 40 |
| Figure 3-3: Promoter architecture of class III genes. | 47 |
| Figure 3-4: Architecture of TFIIC subunits..... | 49 |
| Figure 3-5: Model of TFIIB assembly by τ A/TFIIC. | 50 |
| Figure 3-6: Architecture of TFIIB subunits..... | 51 |
| Figure 3-7: Domain architecture of RNAPIII key subunits involved in PIC assembly..... | 53 |
| Figure 3-8: Schematic of the mechanism of promoter opening and DNA melting by RNAPIII..... | 54 |
| Figure 3-9: Uneven distribution of RNAPIII on transcription units..... | 56 |
| Figure 3-10: Models of Transcription termination by RNAPIII. | 62 |
| Figure 3-11: The RNAPIII transcription cycle. | 66 |

List of Tables

| | |
|--|----|
| Table 1-1: Subunit composition of RNA polymerases in the three domains of life..... | 6 |
| Table 2-1: Subunit composition of RNAPII and its general transcription factors. | 13 |
| Table 2-2: Factors involved in RNAPII transcription termination. | 25 |
| Table 3-1: The RNAPIII transcriptome. | 43 |
| Table 3-2: Subunit composition of RNAPIII and its transcription factors. | 45 |

List of Boxes

| | |
|--|----|
| Box 2-1: The SAGA and the Mediator complex | 14 |
| Box 2-2: The exosome and the TRAMP complex..... | 28 |
| Box 2-3: Pervasive transcription..... | 31 |
| Box 2-4: Transcription termination by bacterial RNAP | 34 |
| Box 2-5: Transcription termination by RNAPI | 36 |
| Box 3-1: The TPR, WH, WD40 and SANT protein domains..... | 55 |
| Box 3-2: UV crosslinking and analysis of the cDNA (CRAC) | 57 |

INTRODUCTION

Chapter 1 General introduction of RNA polymerase and transcription

It has been more than 50 years since the Central Dogma of molecular biology was enunciated by Francis Crick (Crick, 1958). The general idea of the Central Dogma is that the genetic information is copied from DNA into RNA, and then used to make a functional protein. RNA molecules that can encode proteins are known as messenger RNAs (mRNAs). The process by which the genetic information is converted into functional products is called gene expression, which typically contains two key stages: transcription and translation. However, this notion does not apply to a large group of RNA molecules which are transcribed but not destined to become proteins, and thus are termed non-coding RNAs (ncRNAs). Among them, transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs) are the most abundant ones and function as either carriers of amino acids, in the case of tRNAs, or as structural components of the ribosome, in the case of rRNAs, which are both indispensable for protein synthesis. Other classes of non-coding RNAs (such as small nuclear RNAs, snRNAs) also perform many diverse and important functions in the regulation of gene expression. Moreover, prokaryotic and eukaryotic genomes are pervasively transcribed to generate a large ensemble of different RNA molecules, most of them without any known function (Jensen et al., 2013). Overall, RNA is a central

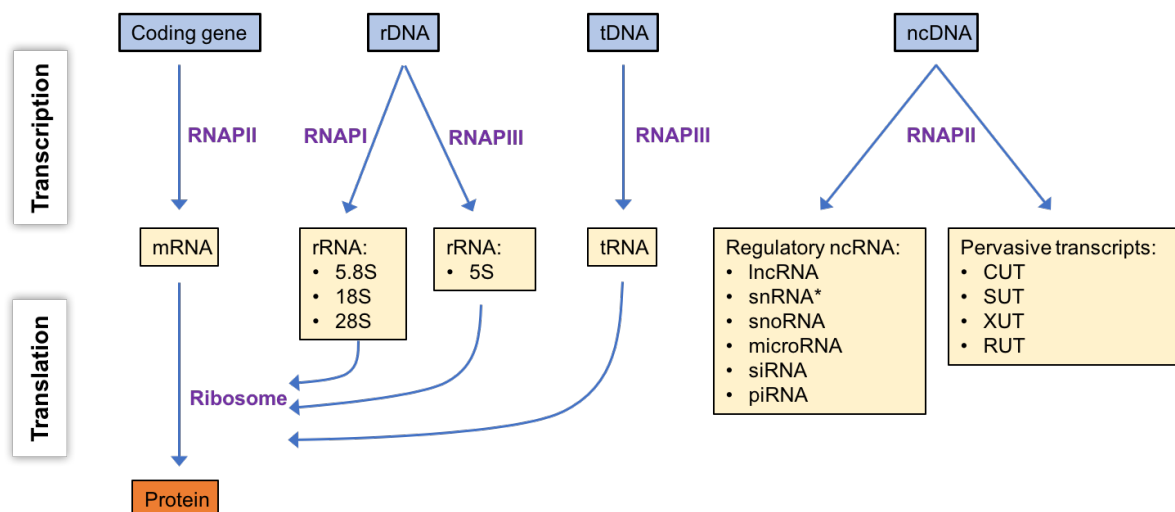


Figure 1-1: Overview of the main classes of RNA in eukaryotes and their role in gene expression.

Abbreviations: rDNA, ribosomal DNA; tDNA or transfer RNA gene; ncDNA, non-protein coding gene; RNAPI, RNA polymerase I; RNAPII, RNA polymerase II; RNAPIII, RNA polymerase III; mRNA, messenger RNA; rRNA, ribosomal RNA; tRNA, transfer RNA; ncRNA, non-coding RNA; lncRNA, long non-coding RNA; snoRNA, small nucleolar RNA; snRNA, small nuclear RNA; miRNA, microRNA; siRNA, small interfering RNA; piRNA, Piwi-interacting RNA; CUT, cryptic unstable transcript; SUT, stable unannotated transcript; XUT, Xrn1-sensitive unstable transcript; RUT, Reb1-dependent unstable transcript. For pervasive transcripts, see **Box 2-3** in section 2.5.2. *U6 snRNA is transcribed by RNAPIII.

component in the flow of genetic information, and therefore plays a fundamental role in all biological processes (Figure 1-1).

1.1 Discovery of the RNA polymerase

The biosynthesis of RNA is carried out by a complex molecular machine, the DNA-dependent RNA polymerase (RNAP), which exists in all living organisms. The RNA polymerase was first discovered in 1959 by Samuel Weiss and Leonard Gladstone when they found that rat liver nuclei support the synthesis of RNA in reaction mixtures containing the four types of nucleotides (Weiss & Gladstone, 1959). One year after, the RNA polymerase was simultaneously discovered in *Escherichia coli* extracts by several independent studies (Huang et al., 1960; Hurwitz et al., 1960; Stevens, 1960). Bacteria RNAP was then purified to homogeneity and found to require an additional subunit, the sigma factor, in order to perform specific transcription initiation (Burgess et al., 1969). Whereas bacteria contain a single type of RNA polymerase, multiple forms of DNA-dependent RNA polymerases were found in eukaryotic cells by Robert Roeder and William Rutter in 1969. In developing sea urchin embryos, they identified three chromatographically distinct RNA polymerases that later proved to contain distinct subunit composition and to produce distinct subsets of RNAs in nuclei (Reeder & Roeder, 1972; Roeder & Rutter, 1969; Sklar et al., 1975; Weinmann et al., 1974; Weinmann & Roeder, 1974; reviewed in Roeder, 2019). Only in the early 1990s the polypeptide sequences of subunits of the three RNAPs were initially revealed for yeast (Sentenac et al., 1992).

Owing to the substantial progress of biochemistry, molecular biology, genomics and structural biology, all forms of nuclear RNAPs in bacteria, archaea and eukaryotes have so far been very well described at the structural and functional level. Even though the subunit composition of these RNAPs vary among the three domains of life, they have a common structural framework and operate by closely related molecular mechanisms, indicating that the last universal common ancestor (LUCA) of all life forms on earth had an RNAP very similar to the simplest form of contemporary RNAPs found in bacteria (Werner & Grohmann, 2011). The composition of the different RNAPs is summarized in Figure 1-2 and **Table 1-1**, and their detailed mechanisms during transcription will be discussed in the following chapters.

1.2 Bacterial RNA polymerase

Bacteria have the simplest form of RNA polymerase comprising five subunits, which are two copies of α , β , β' , and ω , encoded by four genes. These five subunits ($\alpha_2\beta\beta'\omega$) constitute the bacterial RNAP core enzyme, with a total molecular mass of around 400 kDa (Figure 1-2). All

RNAPs in archaea and eukaryotes contain homologues of the bacterial core RNAP subunits (**Table 1-1**). The bacterial core RNAP can bind DNA in a non-sequence-specific manner and initiate transcription from DNA ends or nicks. For initiating transcription from promoter DNA, the core polymerase must bind to a single regulatory subunit known as sigma (σ) factor, which confers specificity and partakes in promoter DNA opening. The core enzyme associated with σ factor is referred to as the holoenzyme. Bacteria express various kinds of σ factors which direct the polymerase to specific promoters in response to environmental cues (reviewed in Sutherland & Murakami, 2018). The simplicity of subunit composition and evolutionary conservation among all organisms make bacterial RNAP an ideal model for studying the mechanisms of transcription.

1.3 Eukaryotic RNA polymerases

In eukaryotic cells transcription of nuclear DNA is carried out by three kinds of multi-subunit RNA polymerases, namely RNA polymerase I (RNAPI), RNA polymerase II (RNAPII) and RNA

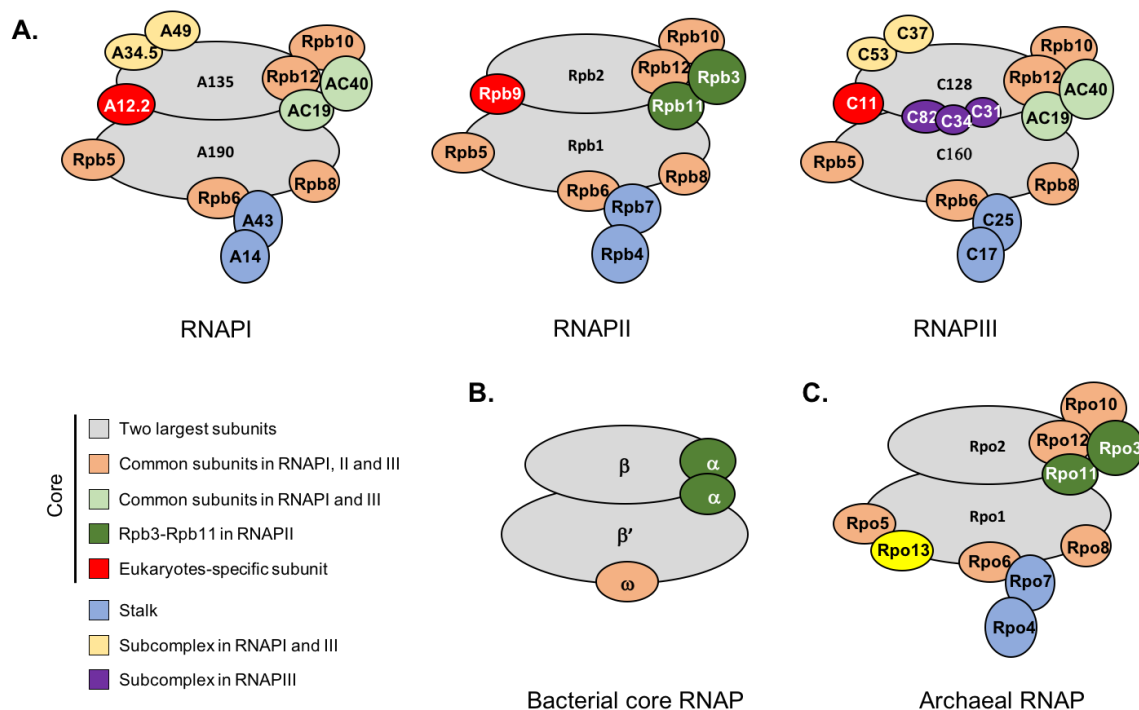


Figure 1-2: Scheme of the RNA polymerase subunits composition in the three domains of life.

(A) Subunit composition of eukaryotic RNAPI (left), RNAPII (middle) and RNAPIII (right). Subunits are labelled according to **Table 1-1**. **(B)** Subunit composition of the bacterial core RNA polymerase. **(A)** and **(B)** are adapted from (Wild & Cramer, 2012). **(C)** Subunit composition of the archaeal RNA polymerase. Scheme adapted from (Hirata & Murakami, 2009). Subunits in **(B)** and **(C)** are coloured according to their eukaryotic homologs in **(A)**, except for Rpo13 which is only found in the domain of Archaea.

polymerase III (RNAPIII). Specifically, RNAPI mainly produces rRNAs, RNAPII synthesises mRNAs and several classes of ncRNAs, and RNAPIII is mostly responsible for the production of tRNAs, the 5S rRNA and a few additional short and abundant ncRNAs. Although RNAPII transcribes the largest number of genes, its transcripts only constitute 5-10% of all cellular RNAs due to their low expression level and high turn-over rates. The vast majority of RNA molecules are rRNAs transcribed by RNAPI which take up to ~75% of total RNAs. RNAPIII transcripts (tRNAs and 5S rRNA) are the second most abundant RNA species and constitute around 15% of all RNAs in a cell (Khatter et al., 2017).

RNAPI, RNAPII and RNAPIII contain 14, 12 and 17 subunits, respectively (Figure 1-2; **Table 1-1**). These polymerases share a 10-subunit catalytic core that consists of the two largest subunits (A190-A135 in RNAPI, Rpb1-Rpb2 in RNAPII, and C160-C128 in RNAPIII) related to bacterial RNAP β and β' subunits; five subunits shared by the three RNAPs (Rpb5-Rpb6-Rpb8-Rpb10-Rpb12); two subunits common to RNAPI and RNAPIII (AC40-AC19) and equivalent to RNAPII subunits Rpb3-Rpb11; and one eukaryote-specific subunit (A12.2, Rpb9 and C11 in RNAPI, RNAPII and RNAPIII, respectively). They also share a heterodimeric stalk composed of A43-A14 in RNAPI, Rpb4-Rpb7 in RNAPII and C25-C17 in RNAPIII. This subdomain is conserved in archaea but not in bacteria, it mediates the interaction with exiting RNA, and has multiple roles in transcription initiation, elongation and termination. RNAPI and RNAPIII share an additional peripheral heterodimeric subcomplex (A49-A34.5 in RNAPI and C53-C37 in RNAPIII) that is related to RNAPII transcription factor TFIIF (see section 2.1) and is involved in transcription initiation and termination. Moreover, RNAPIII contains a specific peripheral trimeric subcomplex (C82-C34-C31) partially resembling TFIIE (see section 2.1) and contributing to transcription initiation (Reviewed in Vannini and Cramer, 2012; Wild and Cramer, 2012).

1.4 Archaeal RNA polymerase

Archaea are prokaryotic microorganisms that were once considered to be closely related to bacteria as their cellular organization resembles bacteria. However, they are now believed to be more closely related to the eukaryotic cells at the molecular level than bacteria. Structural evidence shows that the archaeal transcription machinery including RNAP and some general transcription factors (such as TBP and TFB) is similar to that of eukaryotes, while archaeal transcription regulatory factors, such as activators and repressors, are found in close relationship with bacterial factors. Thus, transcription in archaeal appears to be a combination of eukaryotic-type transcription apparatus with bacterial-like regulatory mechanisms (Jun et al., 2011). Like bacteria, archaea possess only one kind of RNAP to

transcribe all genes and its subunit composition and architecture are very similar to the eukaryotic RNAPII, including a catalytic core and a heterodimeric stalk (Figure 1-2; **Table 1-1**). Archaeal RNAP consists of 10-12 subunits depending on the species, most of which are conserved in eukaryotes except for Rpo13, a subunit with unknown function that is only present in some groups of archaea (reviewed in Jun et al., 2011, Fouqueau et al., 2018).

1.5 RNA polymerase IV and V

Plants have evolved two additional DNA-dependent RNA polymerases, RNAPIV and RNAPV. These two odd polymerases are thought to evolve from eukaryotic RNAPII because of the striking similarities with RNAPII regarding their subunit composition and architecture. RNAPIV and RNAPV are composed of 12 subunits, as RNAPII, with half of them identical to those of RNAPII and the remaining subunits being encoded by paralogues of genes encoding RNAPII subunits (**Table 1-1**). RNAPIV and RNAPV are not essential for viability in plants but they both cooperate to play an important role in RNA-mediated gene silencing and heterochromatin formation. These processes are involved in development, transposon control, genome defence against viruses and allelic crosstalk (reviewed in Haag and Pikaard, 2011).

Table 1-1: Subunit composition of RNA polymerases in the three domains of life.

| | Bacteria | Archaea | Eukaryotes | | | | |
|---------------------------------------|-------------------|--------------|--------------|-------------------------|---------|---------------|---------------|
| | | | RNAPI | RNAPII | RNAPIII | Plant RNAPIV | Plant RNAPV |
| Core | β' | Rpo1 (RpoA) | Rpa190 | Rpb1 (B220) | Rpc160 | NRPD1 | NRPE1 |
| | β | Rpo2 (RpoB) | Rpa135 | Rpb2 (B150) | Rpc128 | NRPD2 (NRPE2) | NRPD2 (NRPE2) |
| | α | Rpo3 (RpoD) | Rpc40 | Rpb3 (B44) | Rpc40 | Rpb3 | Rpb3 |
| | α | Rpo11 (RpoL) | Rpc19 | Rpb11 (B12.5) | Rpc19 | Rpb11 | Rpb11 |
| | ω | Rpo6 (RpoK) | Rpb6 | Rpb6 (ABC23) | Rpb6 | Rpb6 | Rpb6 |
| | | Rpo5 (RpoH) | Rpb5 | Rpb5 (ABC27) | Rpb5 | Rpb5 | NRPE5 |
| | | Rpo8* (RpoG) | Rpb8 | Rpb8 (ABC14.5) | Rpb8 | Rpb8 | Rpb8 |
| | | Rpo10 (RpoN) | Rpb10 | Rpb10 (ABC10 β) | Rpb10 | Rpb10 | Rpb10 |
| | | Rpo12 (RpoP) | Rpb12 | Rpb12 (ABC10 α) | Rpb12 | Rpb12 | Rpb12 |
| | | Rpa12 | Rpb9 (B12.6) | Rpc11 | NRPD9b | Rpb9 | |
| Stalk | | Rpo4 (ProF) | Rpa14 | Rpb4 (B32) | Rpc17 | NRPD4 (NRPE4) | NRPD4 (NRPE4) |
| | | Rpo7 (RpoE) | Rpa43 | Rpb7 (B16) | Rpc25 | NRPD7 | NRPE7 |
| Peripheral subcomplex (TFIIF-like) | | | Rpa49 | | Rpc37 | | |
| | | | Rpa34 | | Rpc53 | | |
| Peripheral subcomplex (TFIIE-related) | | | | | Rpc82 | | |
| | | | | | Rpc34 | | |
| | | | | | Rpc31 | | |
| Bacteria-specific | σ -factors | | | | | | |
| Archaea-specific | | Rpo13* | | | | | |

*Found in some but not all archaeal species.

Alternative names for subunits of RNAPs are indicated in brackets

Adapted from Werner and Grohmann, 2011, Wild and Cramer, 2012, Porrua et al., 2016

1.6 Transcription cycle

As mentioned above, transcription is the first step in gene expression and leads to the production of RNA molecules from DNA templates. The transcription cycle includes the following events:

(1) Initiation, which involves the recruitment of transcription initiation factors and RNAP to gene promoter regions and the melting of promoter sequences to permit RNAP to launch RNA synthesis.

(2) Elongation, which consists in the escape of RNAP from the promoter to start processive addition of nucleotides to the growing RNA chain.

(3) Termination or the cessation of RNA synthesis and the disassembly of the transcription elongation complex (EC) composed of the RNA polymerase, the nascent RNA transcript and the DNA template.

Every step in the transcription cycle is highly controlled by a multitude of proteins called transcription factors (TFs). Some transcription factors bind to the promoter sequence to help form the transcription initiation complex, thus known as basal transcription factors or general transcription factors (GTFs). Other transcription factors bind to DNA regulatory sequences, for instance, the enhancers in metazoans or upstream activation sequences (UAS) in yeast, to regulate the transcription of the related gene. These factors are thus termed regulators, including transcription activators (TAs) and repressors (TRs). Activators and repressors, collectively known as specific transcription factors, act, in part, by recruiting the transcription machinery or repressive complexes to gene regulatory regions. Their activities are further modulated by the transcription co-regulators (co-activators or co-repressors), which, in general, function by transmitting the signals from the specific transcription factors to the transcription preinitiation complex (PIC) (Hahn and Young, 2011; Soutourina, 2018). In contrast to the high degree of homology between RNAPs from different organisms, the evolutionary conservation of TFs is very limited. Furthermore, even though the overall process of transcription is similar in the three domains of life, some significant divergences and variabilities can be found between different lineages and different kinds of polymerases (Werner and Grohmann, 2011).

In the following chapters, I will introduce the mechanisms of transcription by eukaryotic RNAPII and RNAPIII, respectively, including the process of transcription initiation, elongation and termination. For RNAPI and bacterial RNAP, I will only briefly describe the current models governing transcription termination, for comparison. The activities of archaeal RNAP and plant-specific RNAPs will not be covered by this manuscript.

1.7 Yeast as a model organism

Our laboratory works on budding yeast *Saccharomyces cerevisiae*, the simplest and best-characterized eukaryotic cellular model. *S. cerevisiae* has been cultured and extensively studied in laboratory for many decades and has been exploited for understanding all kinds of basic biological processes. The main advantages of this model are that: it is a unicellular organism with a short generation time (~100 min); it is easy to grow and to manipulate genetically; it is the first eukaryotic organism having its entire genome sequenced (Goffeau et al., 1996); and many of the essential cellular processes, among which the mechanisms of transcription, are highly conserved from yeast to humans.

In this manuscript, I will focus on the mechanisms governing transcription in *S. cerevisiae* but comparisons with other organisms will also be provided when necessary.

Chapter 2 Transcription by RNA polymerase II

Eukaryotic RNA polymerase II (RNAPII) is a 12-subunit protein complex dedicated to the transcription of all protein-coding genes and many non-coding regions in eukaryotic genomes. The structure of the 10-subunit core RNAPII from *S. cerevisiae* lacking the subcomplex Rpb4-Rpb7 was first solved by crystallography (Cramer et al., 2000). 20 years afterwards, atomic structures of RNAPII holo-enzyme from different species and under different transcriptional states are available (reviewed in Osman and Cramer, 2020), which contribute greatly to our understanding of the mechanisms of eukaryotic transcription. Transcription by RNAPII is so far the best characterized. RNAPII has the simplest subunit composition among the three eukaryotic RNAPs (Figure 1-2). Nevertheless, RNAPII transcription is clearly the most highly organized and tightly controlled process, which requires regulation at multiple steps by a large number of transcription factors (Woychik and Hampsey, 2002).

2.1 Transcription initiation

Transcription initiation of RNAPII-dependent genes (class II genes) is mediated by the basal transcription machinery, which consists of RNAPII and several GTFs including TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH. GTFs cooperate with RNAPII to recognize and open the promoter DNA, to nucleate the synthesis of RNA and stimulate the escape of RNAPII from the promoter region (Sainsbury et al., 2015; Schier and Taatjes, 2020).

2.1.1 The RNAPII core promoter

The RNAPII core promoter is generally defined to be the minimal stretch of DNA that is sufficient to direct the accurate initiation of RNAPII transcription. A core promoter comprises several specific DNA sequence motifs, which include the TATA box, the initiator (Inr), the TFIIB recognition element (BRE), the motif ten element (MTE), and downstream core promoter element (DPE) (Figure 2-1). RNAPII core promoters are found to be structurally and functionally diverse as there are no universal elements that are found in all promoters (reviewed in Smale and Kadonaga, 2003; Juven-Gershon and Kadonaga, 2010; Kadonaga, 2012).

The TATA box or TATA-related sequence was the first core promoter motif discovered (Goldberg, 1979; Breathnach and Chambon, 1981) and is the best known core promoter element so far. It is a consensus sequence of ~8 nt localized ~30 bp and 40-120 bp upstream of the transcription start site (TSS) in metazoans and yeast, respectively, and is recognized

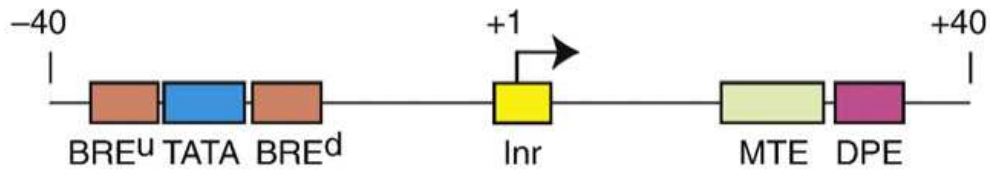


Figure 2-1: The core promoter elements for transcription by RNAPII.

This diagram is roughly to scale. -40 and +40 represent the distance in base pair from the transcription start site (black arrow). Abbreviations: BRE^u, upstream TFIIB recognition element; BRE^d, downstream TFIIB recognition element; TATA, TATA box; Inr, initiator; MTE, motif ten element; DPE, downstream core promoter element. From Juven-Gershon and Kadonage, 2010.

and bound by the TATA-binding protein (TBP). The TATA box and TBP are both conserved from archaea to humans (Reeve, 2003), however, only 10-20% of yeast and human core promoters are TATA-containing ones (Basehoar et al., 2004; Yang et al., 2007). There are two BRE motifs, located upstream (BRE^u) or downstream (BRE^d) of a TATA box, both of which function in conjunction with a TATA box in transcriptional regulation (Juven-Gershon and Kadonage, 2010). The sequence encompassing the TSS is called the initiator (Inr), which is possibly the most commonly occurring core promoter motif. The Inr element is recognized by TFIID, it is functionally similar to the TATA box and can function independently of it (Smale and Kadonage, 2003). The DPE is a recognition site for TFIID downstream of the Inr. The MTE element is localized immediately upstream of the DPE and stimulates the binding of TFIID. The MTE functions in cooperation with the Inr but can act independently of the DPE and the TATA box. BRE, MTE and DPE elements are only found in metazoan (Juven-Gershon and Kadonage, 2010).

2.1.2 PIC assembly

According to the *in vitro* experiments, transcription initiation commences with the assembly of GTFs and RNAPII on the core promoter to form a large complex known as the transcription preinitiation complex (PIC). To nucleate PIC assembly, TBP, a subunit of TFIID, binds first to the TATA box. TFIID is a multisubunit complex composed of TBP and, in yeast, 14 TBP-associated factors (TAFs) that also involved in promoter recognition (Huisinga and Pugh, 2004). TBP binding to the TATA box induces a ~90-degree bent in the DNA (Geiger et al., 1996; Tan et al., 1996), which is subsequently stabilized by the joining of TFIIB and TFIIA. TFIIA is a two-subunit auxiliary factor that is not strictly required for basal transcription but can stabilize the TBP-DNA complex (Imbalzano et al., 1994). TFIIB is composed of a single polypeptide chain that is required for the recruitment of RNAPII to the promoter and facilitates TBP binding to DNA and DNA bending (Malik et al., 1991; Zhao and Herr, 2002). The complex containing these

GTFs bound to the upstream promoter recruits the RNAPII-TFIIF complex, leading to the formation of a stable complex called core PIC (cPIC) (Buratowski et al., 1989). Mammalian TFIIF is constituted of the TFIIF α/β heterodimer, corresponding to Tfg1/2 in yeast. However, yeast TFIIF additionally contains a third subunit, Tfg3 that is not essential for transcription (Henry et al., 1992). TFIIF plays an important role in the PIC stabilization, TSS selection and early RNA synthesis (Sainsbury et al., 2015). Finally, the heterodimeric factor TFIIE is recruited to the cPIC to assemble an intermediate PIC (mPIC), followed by the 10-subunit TFIIH complex to form a complete holo PIC (hPIC) (Zawel and Reinberg, 1993). The TFIIH is an essential and multifunctional factor, which includes the ATPase/helicase activity that is necessary for promoter opening, and the kinase activity that promotes the phosphorylation of the RNAPII to initiate transcription (Svejstrup et al., 1996). TFIIE binds to RNAPII, facilitates the recruitment of TFIIH and stimulates the ATPase and kinase activity of TFIIH (Figure 2-2; Table 2-1; Sainsbury et al., 2015).

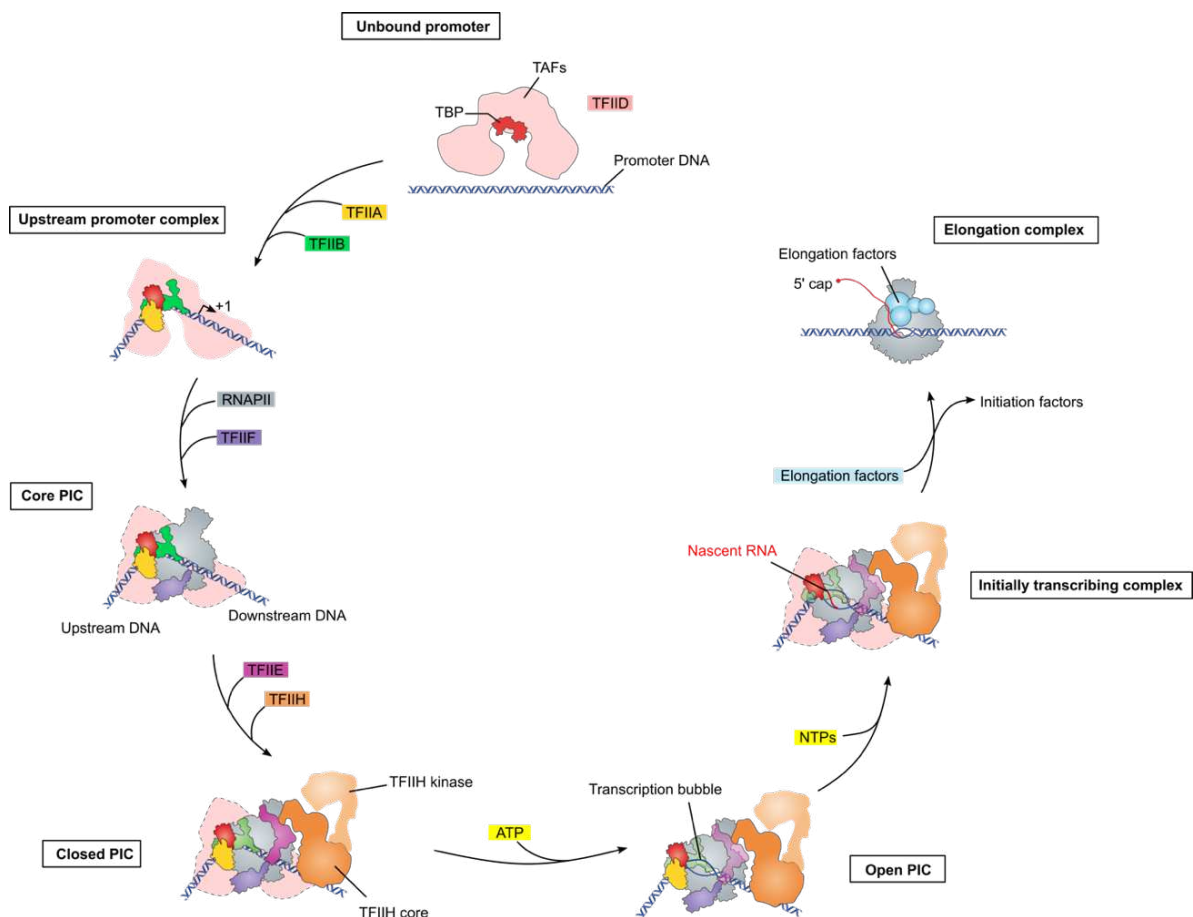


Figure 2-2: *In vitro* model of RNAPII transcription initiation.

Representation of the canonical model for stepwise pre-initiation complex (PIC) assembly from general transcription factors and RNAPII on promoter DNA. The names of the intermediate PIC complexes are provided in the boxes. Details described in the text. Adapted from Sainsbury et al., 2015.

The aforementioned stepwise PIC assembly is dependent on the TBP-TATA system. However, a vast majority (up to 85%) of core promoters lack a consensus TATA box motif or TATA-like sequence, which are commonly called TATA-less promoters (Basehoar et al., 2004; Yang et al., 2007). Nevertheless, TFIID is required for the transcription of almost all class II genes (Warfield et al., 2017) and it was not clear until recently how TFIID can recognize highly diversified RNAPII promoters. A newly structural study on human TFIID-containing PIC revealed that TATA box and TATA-less promoters employ a shared TFIID-binding pattern and loading of TBP, and TBP similarly bends TATA box and TATA-less promoters, providing structural insights into how TFIID can support PIC assembly on TATA-less promoters (Chen et al., 2021).

In general, a complete PIC consists of RNAPII and six basal transcription factors (TFIIA, TFIIB, TFIID, TFIIIE, TFIIF and TFIIH) as well as the closed, double-stranded DNA (**Table 2-1**). It is also worth mentioning that PIC assembly is normally preceded and activated by the binding of activators to DNA regions located at various distances from the core promoter, and often require co-activators, such as the SAGA complex that is necessary for all regulated transcription and is conserved among eukaryotes, and the Mediator complex, which plays an important role in the assembly and stabilization of the PIC by interacting with GTFs and now is widely considered as a part of the PIC (Nguyen et al., 2021) (**Box 2-1**).

Table 2-1: Subunit composition of RNAPII and its general transcription factors.

| Human | | | Yeast | | | |
|-----------------------|---------------|----------|--------------|--------------|----------|--|
| Subunit | Gene | MW (kDa) | Subunit | Gene | MW (kDa) | |
| RNAPII | | | | | | |
| RPB1 | <i>POLR2A</i> | 217.2 | RPB1 | <i>RPO21</i> | 191.6 | |
| RPB2 | <i>POLR2B</i> | 133.9 | RPB2 | <i>RPB2</i> | 138.8 | |
| RPB3 | <i>POLR2C</i> | 31.4 | RPB3 | <i>RPB3</i> | 35.3 | |
| RPB4 | <i>POLR2D</i> | 16.3 | RPB4 | <i>RPB4</i> | 25.4 | |
| RPB5 | <i>POLR2E</i> | 24.6 | RPB5 | <i>RPB5</i> | 25.1 | |
| RPB6 | <i>POLR2F</i> | 14.5 | RPB6 | <i>RPO26</i> | 17.9 | |
| RPB7 | <i>POLR2G</i> | 19.3 | RPB7 | <i>RPB7</i> | 19.1 | |
| RPB8 | <i>POLR2H</i> | 17.1 | RPB8 | <i>RPB8</i> | 16.5 | |
| RPB9 | <i>POLR2I</i> | 14.5 | RPB9 | <i>RPB9</i> | 14.3 | |
| RPB10 | <i>POLR2L</i> | 7.6 | RPB10 | <i>RPB10</i> | 8.3 | |
| RPB11 | <i>POLR2J</i> | 13.3 | RPB11 | <i>RPB11</i> | 13.6 | |
| RPB12 | <i>POLR2K</i> | 7 | RPB12 | <i>RPB12</i> | 7.7 | |
| TFIID | | | | | | |
| TBP | <i>TBP</i> | 37.7 | TBP | <i>SPT15</i> | 27 | |
| TAF1 | <i>TAF1</i> | 212.7 | TAF1 | <i>TAF1</i> | 120.7 | |
| TAF2 | <i>TAF2</i> | 137 | TAF2 | <i>TAF2</i> | 161.5 | |
| TAF3 | <i>TAF3</i> | 103.6 | TAF3 | <i>TAF3</i> | 40.3 | |
| TAF4 | <i>TAF4</i> | 110.1 | TAF4 | <i>TAF4</i> | 42.3 | |
| TAF5 | <i>TAF5</i> | 86.8 | TAF5 | <i>TAF5</i> | 89 | |
| TAF6 | <i>TAF6</i> | 72.7 | TAF6 | <i>TAF6</i> | 57.9 | |
| TAF7 | <i>TAF7</i> | 40.3 | TAF7 | <i>TAF7</i> | 67.6 | |
| TAF8 | <i>TAF8</i> | 34.3 | TAF8 | <i>TAF8</i> | 58 | |
| TAF9 | <i>TAF9</i> | 29 | TAF9 | <i>TAF9</i> | 17.3 | |
| TAF10 | <i>TAF10</i> | 21.7 | TAF10 | <i>TAF10</i> | 23 | |
| TAF11 | <i>TAF11</i> | 23.3 | TAF11 | <i>TAF11</i> | 40.6 | |
| TAF12 | <i>TAF12</i> | 17.9 | TAF12 | <i>TAF12</i> | 61.1 | |
| TAF13 | <i>TAF13</i> | 14.3 | TAF13 | <i>TAF13</i> | 19.1 | |
| n.a. | n.a. | n.a. | TAF14 | <i>TAF14</i> | 27.4 | |
| TFIIA | | | | | | |
| Large subunit | <i>GTF2A1</i> | 41.5 | Toa1 | <i>TOA1</i> | 32.2 | |
| Small subunit | <i>GTF2A2</i> | 12.5 | Toa2 | <i>TOA2</i> | 13.5 | |
| TFIIB | | | | | | |
| TFIIB (TFB*) | <i>GTF2B</i> | 34.8 | TFIIB | <i>SUA7</i> | 38.2 | |
| TFIIF | | | | | | |
| TFIIF α | <i>GTF2F1</i> | 58.2 | Tfg1 | <i>TFG1</i> | 82.2 | |
| TFIIF β | <i>GTF2F2</i> | 28.4 | Tfg2 | <i>TFG2</i> | 46.6 | |
| n.a. | n.a. | n.a. | Taf14 | <i>TAF14</i> | 27.4 | |
| TFIIE | | | | | | |
| TFIIE α (TFE*) | <i>GTF2E1</i> | 49.5 | Tfa1 | <i>TFA1</i> | 54.7 | |
| TFIIE β | <i>GTF2E2</i> | 33 | Tfa2 | <i>TFA2</i> | 37 | |
| TFIIH | | | | | | |
| p62 | <i>GTF2H1</i> | 62 | Tfb1 | <i>TFB1</i> | 72.9 | |
| p44 | <i>GTF2H2</i> | 44.4 | Ssl1 | <i>SSL1</i> | 52.3 | |
| p34 | <i>GTF2H3</i> | 34.4 | Tfb4 | <i>TFB4</i> | 37.5 | |
| p52 | <i>GTF2H4</i> | 52.2 | Tfb2 | <i>TFB2</i> | 58.5 | |
| p8 | <i>GTF2H5</i> | 8.1 | Tfb5 | <i>TFB5</i> | 8.2 | |
| XPD | <i>ERCC2</i> | 86.9 | Rad3 | <i>RAD3</i> | 89.8 | |
| XPB | <i>ERCC3</i> | 89.3 | Ssl2 (Rad25) | <i>SSL2</i> | 95.3 | |
| Cyclin H | <i>CCNH</i> | 37.6 | Ccl1 | <i>CCL1</i> | 45.2 | |
| CDK7 | <i>CDK7</i> | 39 | Kin28 | <i>KIN28</i> | 35.2 | |
| MAT1 | <i>MNAT1</i> | 35.8 | Tfb3 | <i>TFB3</i> | 38.1 | |

Adapted from Sainsbury et al., 2015
n.a., not available

Box 2-1: The SAGA and the Mediator complex

SAGA (Spt-Ada-Gcn5 acetyltransferase) is a multi-subunit transcriptional co-activator complex conserved between yeast and humans that controls transcription by modifying histones. Yeast SAGA contains 19 subunits, with a total molecular mass of 1.8 MDa, which are organized into four modules with distinct functions, namely:

- 1) the histone acetyltransferase (HAT) module, composed of Gcn5, Ada2, Ada3 and Sgf29;
- 2) the histone deubiquitinase (DUB) module, composed of Ubp8, Sgf11, Sgf73 and Sus1;
- 3) the Tra1 module, containing only the large Tra1 protein that serves as a docking platform for transcription factors binding;
- 4) the 10-subunit core module, composed of TBP-associated factor (TAF) Taf5/6/9/10/12 as well as Spt3/7/8/20 and Ada1.

Although not considered as a PIC factor, SAGA can be recruited to promoters by gene-specific transcription factors, can bind TBP, and contains activities to acetylate and to deubiquitylate histones. Reviewed in Schier and Taatjes, 2020; Osman and Cramer, 2020; Wang et al., 2020; Papai et al., 2020.

The mediator of RNAPII transcription (Mediator) complex is a transcription co-activator conserved from yeast to metazoans. Budding yeast Mediator comprises 25 subunits that are organized into four distinct modules: the head module, the middle module, and the tail module, and the CDK8 kinase module, which is transiently associated with the complex. The main function of Mediator is to transduce signals from the activators to the preinitiation complex to assist in PIC assembly on core promoters. The Mediator complex can also stimulate the phosphorylation of the carboxy-terminal domain (CTD) of the largest RNAPII subunit, which in turn triggers RNAPII release from the promoter. Reviewed in Poss et al., 2013; Soutourina, 2018.

2.2 Transcription elongation

2.2.1 Promoter clearance

The transition from transcription initiation to productive elongation must go through a stage known as promoter clearance, during which the contact with initiation factors is lost and stable association with the nascent transcript and elongation factors is established (Luse, 2013). Once the PIC assembles, the closed promoter DNA needs to be opened to initiate transcription. Unlike other DNA-dependent RNAPs, RNAPII cannot open the promoter DNA by itself. Promoter opening requires ATP and the Ssl2 (XPB in human) subunit of TFIIH, which possesses ATP-dependent DNA-translocase activity. Following formation of the open complex, RNA synthesis commences at the TSS in the presence of NTPs (Sainsbury et al., 2015). Nevertheless, in the very early stage of RNA synthesis, the RNA:DNA hybrid within the transcription bubble is too short to be stable and the nascent transcript can be released from the elongation complex (EC), resulting in abortive initiation. Abortive initiation is considerably reduced when the length of RNA:DNA hybrid reaches 8-9 nt (Sims et al., 2004). As elongation continues, the 5' end of the nascent RNA is released from the template DNA and enters the RNA exit channel of the polymerase. After synthesis of ~30 nt of RNA, RNAPII is thought to lose contact with the core promoter and the rest of the transcription machinery, and promoter clearance is complete (Luse, 2013). At this stage, a subset of GTFs remain at the core promoter, serving as a scaffold for the assembly of the next transcription initiation complex, which is believed to be much faster relative to the initial round. Of all the GTFs, only TFIIF and TFIIB need to be re-assembled for a new cycle of transcription event (Hahn, 2004).

2.2.2 Elongation in the body of genes

RNAPII transcription elongation is not a smooth, continuous process. During transcription elongation, RNA polymerase may encounter obstacles that can slow down or stall the transcribing polymerase. These obstacles can be caused by the positioning of nucleosomes, DNA-binding factors, DNA damages/mismatches, depletion of NTPs, and DNA sequences that are intrinsically difficult to transcribe. Many factors are required for normal transcription elongation in the body of genes, such as TFIIF, Spt4 and Spt5 (DSIF in humans), and the PAF complex (Selth et al., 2010). During elongation RNAPII might stall and backtrack, i.e. move in the reverse direction along the DNA template. Polymerase backtracking results in the displacement of the neo-synthesized RNA 3' end from the catalytic sites, which renders transcriptionally inactive the enzyme (Wilson et al., 2013). Arrested RNAPII can be reactivated by the recruitment of TFIIS (Dst1 in yeast) that acts by stimulating the intrinsic RNA endonuclease activity of RNAPII to cleave the displaced portion of the transcript, so that the

RNA 3' end is located again in the polymerase active center (Fish and Kane, 2002; Cheung and Cramer, 2011). However, transcribing RNAPII can become permanently stalled or arrested under a wide variety of conditions. If the arrested RNAPII cannot be restarted, it becomes poly-ubiquitylated by ubiquitin ligases, and is then disassembled and degraded by the proteasome (Wilson et al., 2013). Because transcription occurs on a chromatin template, factors that affect chromatin dynamics are important for elongation. These will be briefly discussed in section 2.4.

2.3 The phosphorylation cycle of the RNAPII CTD

Phosphorylation of the carboxy-terminal domain (CTD) of Rpb1, the largest subunit of RNAPII is an important event in the transcription cycle. The RNAPII CTD is a long unstructured domain composed of multiple heptapeptide repeats with the consensus sequence Tyr₁-Ser₂-Pro₃-Thr₄-Ser₅-Pro₆-Ser₇ (YSPTSPS). The number of these repeats varies between species: 26 in *S. cerevisiae*, 29 in *S. pombe*, 32 in *C. elegans*, 37 in *D. melanogaster* and *A. thaliana*, and 52 in *H. sapiens*. Most of the repeats in yeast are identical to the consensus, while in human, only the first 26 are highly conserved (Figure 2-3). A full-length CTD is not required for the function of RNAPII or viability, as genetic studies showed that the minimal viable CTD in yeast contains eight repeats (West and Corden, 1995). The CTD is subject to extensive post-translational modifications (PTMs), most notably phosphorylation, which define a “CTD code” based on the combination of modified residues. CTD post-translational modifications are crucial for the regulation of the transcription process, from initiation to termination, as well as for co-transcriptional processes, such as pre-mRNA capping, splicing and chromatin modification (reviewed in Harlen and Churchman, 2017).

Of the seven residues in a CTD repeat, Tyr1, Ser2, Thr4, and Ser5 and Ser7 residues are dynamically phosphorylated and dephosphorylated by several CTD kinases and phosphatases throughout the transcription cycle (Figure 2-4, Harlen and Churchman, 2017). Among them, Ser5 phosphorylation (Ser5P) and Ser2P patterns are the best studied and characterized. Initially, RNAPII assembled in the PIC has an unmodified CTD, which has high affinity for the Mediator complex (Lu et al., 1991; Myers et al., 1998). Cyclin-dependent kinase 7 (CDK7; known as Kin28 in *S. cerevisiae*), the kinase subunit of TFIIH, phosphorylates Ser5 and Ser7 early in the transcription cycle and these phosphorylations are believed to favour promoter escape by decreasing the affinity of RNAPII for the Mediator complex (Wong et al., 2014; Jeronimo and Robert, 2014). Ser5P also promotes the recruitment of capping and splicing factors (Perales and Bentley, 2009), the COMPASS (Complex Protein Associated with Set1, see section 2.4.1) as well as the NNS complex (Nrd1-Nab3-Sen1, see section 2.5.2).

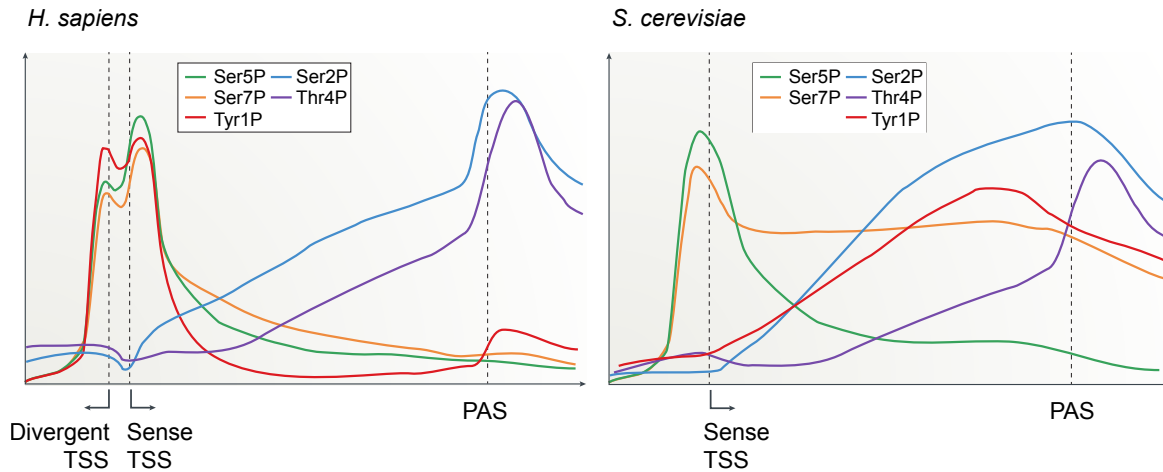


Figure 2-4: The CTD phosphorylation patterns across protein-coding genes in humans and yeast.

The average chromatin immunoprecipitation (ChIP) profiles of phosphorylated residues of the carboxy-terminal domain (CTD) of RNAPII at protein-coding genes during the transcription cycle is represented. The profiles in budding yeast and humans are highly similar, except for Tyr1 phosphorylation. TSS indicates the position of the transcription start site, and PAS indicates the polyadenylation site. From Harlen and Churchman, 2017.

2.4 Chromatin dynamics during transcription elongation

Transcription by RNAPII occurs on a DNA template that is organized in chromatin fibers, the elementary unit of which is the nucleosome. A nucleosome consists of ~147 nt of DNA wrapped 1.65 turns around a set of eight proteins called histones, which are known as a histone octamer. A histone octamer is composed of two copies each of the histone proteins H2A, H2B, H3 and H4. There are multiple interactions between the histones and the DNA, making the nucleosome one of the most stable protein-DNA complexes that thus serves as a strong physical barrier to RNAPII movement (Luger et al., 1997). However, the nucleosome is not a static but rather a dynamic unit, which is under control of various protein complexes that favour the passage of RNAPII during transcription. The best characterized of these factors include the histone modifiers, the histone chaperones and the ATP-dependent chromatin remodelers (Sims et al., 2004; Li et al., 2007; Selth et al., 2010).

2.4.1 Histone modifiers

During transcription, histones are subject to a vast array of post-translational modifications including methylation of arginine residues; methylation, acetylation, ubiquitination, ADP-ribosylation, and sumolation of lysines; and phosphorylation of serines and threonines (Li et al., 2007; Kouzarides, 2007). These modifications are found primarily in the unstructured, amino terminal segments of histones that protrude from the nucleosome (known as histone

tails) (Luger et al., 1997). Histone modifications play an important role in transcription elongation, working by affecting internucleosomal contacts or changing electrostatic charges to alter the packaging of chromatin (Selth et al., 2010). Histone modifications can also serve as a binding surface for elongation-associated effector complexes (Dechassa et al., 2010; Hassan et al., 2001; Spain et al., 2014).

Histone acetylation is mostly associated with activation of transcription. It occurs at multiple lysine residues and is carried out by histone acetyltransferases (HAT). Acetylation decreases the positive charge on the histone tails, thereby affecting the interactions of the histone octamer with the negatively charged phosphate group of the DNA. As a consequence, the chromatin is more relaxed and the DNA more accessible to transcription factors. This effect can be reversed by deacetylation catalysed by histone deacetylases (HDAC), which normally correlates with transcriptional repression (Selth et al., 2010; Zentner and Henikoff, 2013). As an example, the SAGA complex (**Box 2-1**) containing the HAT Gcn5 protein can stimulate acetylation of histones and thus can cause the eviction of nucleosomes in transcribed coding sequences and promote RNAPII elongation (Govind et al., 2007). Also, NuA4, the major H4 lysine acetyltransferase (KAT) complex in *S. cerevisiae*, is recruited through the interaction with phosphorylated RNAPII CTD to acetylate H4 and promote histone eviction (Ginsburg et al., 2009).

Unlike acetylation, histone methylation does not change the net charge of nucleosomes, but rather acts as a tag for effector proteins containing methyl-binding domains. Lysine (K) residues of H3 and H4 can be modified by one, two, or three methyl groups (me), and these different methylation sites can have distinct functions. Methylation on histone 3 lysine 4 (H3K4me), H3K36me and H3K79me are implicated in activation of transcription and commonly referred to as euchromatin modifications, while H3K9me, H3K20me and H3K27me are localized to inactive genes or regions and are often termed heterochromatin modifications (Kouzarides, 2007). In yeast, H3K4me and H3K36me are carried out by Set1, which is part of the COMPASS complex (see below), and Set2, respectively. The recruitment of Set1 and Set2 depends on the phosphorylation state of the RNAPII CTD (Sims et al., 2004; Selth et al., 2010).

The COMPASS (Complex Protein Associated with Set1) was the first H3K4 methylase identified in *S. cerevisiae* and is capable of catalysing the mono-, di-, and trimethylation of histone H3K4 (Miller et al., 2001). This complex consists of eight subunits, including Set1, Cps60, Cps50, Cps40, Cps35, Cps30, Cps25, and Cps15. Among them, Cps35 (also known as SWD2) is the only essential subunit of the complex in yeast, which is shared with other complexes, such as the cleavage and polyadenylation factor complex. Set1 is the catalytic subunit that possess histone (H) or lysine (K) methyltransferase (HMTase or KMTase) activity.

However, Set1 alone is not active as a KMTase, as Set1 within COMPASS is the active form of the enzyme (Shilatifard, 2012).

2.4.2 Histone chaperones

Histone chaperones are histone-binding proteins involved in intracellular histone dynamics, as well as histone storage and replication-associated chromatin assembly (Selth et al., 2010). The FACT (facilitates chromatin transcription) complex and Spt6 are two histone chaperones that affect the chromatin structure during transcription elongation. The FACT complex is a heterodimer composed of Spt16 and Pob3 in yeast (SPT16 and SSRP1 in human). This complex is highly conserved among eukaryotes, and functions to destabilize the nucleosome by selectively displacing the histone dimer H2A/H2B and reassemble nucleosomes after RNAPII passage (Orphanides et al., 1999; Hsieh et al., 2013). Spt6 (suppressor of Ty6) is also well conserved throughout eukaryotes and is involved in the maintenance of chromatin structure during elongation (Selth et al., 2010).

2.4.3 ATP-dependent chromatin remodelers

ATP-dependent chromatin remodeling complexes (remodelers) utilize the energy of ATP hydrolysis to modify the structure of chromatin. The outcomes of these factors include transient unwrapping the end DNA from histone octamers (histone ejection), or moving nucleosomes to different positions (histone sliding), all of which change the accessibility of nucleosomal DNA to transcription factors (Li et al., 2007). Chromatin remodelers can be classified into four subfamilies: SWI-SNF (switch/sucrose non-fermentable), ISWI (imitation switch), CHD (chromodomain helicase DNA-binding), and INO80, on the basis of the similarities in their catalytic ATPases. Each subfamily employs unrelated enzymatic mechanisms to alter the nucleosome organization of chromatin. However, they all use ATP-dependent DNA translocation to break histone-DNA contacts and to propel DNA along the histone surface (for review, see Clapier et al., 2017).

2.5 Transcription termination

The elongation complex is removed from the DNA template at the end of genes, to allow the release of the transcript and the recycling of the polymerase and to prevent the interference with neighbouring transcription units. In budding yeast, there are two main pathways for termination of RNAPII transcription (Porrua et al., 2016). The first one depends on the Cleavage and Polyadenylation Factor-Cleavage Factor (CPF-CF) complex and is responsible for the termination of protein-coding genes; the second one depends on the Nrd1-Nab3-Sen1 (NNS) complex and is dedicated to termination of genes coding for snRNAs, snoRNAs, and cryptic unstable transcripts (CUTs), which constitute the major product of pervasive transcription (for pervasive transcription, see **Box 2-3**).

2.5.1 The CPF-CF pathway

RNAPII termination of most mRNA genes is functionally connected with the 3' end processing of the nascent transcript, which requires a set of different proteins that bind to the polymerases as well as to specific sequences in the nascent RNA. This event in yeast is dependent on a large, multisubunit complex composed of three subcomplexes: CPF (cleavage and polyadenylation factor), CFIA (cleavage factor IA) and CFIB (cleavage factor IB), hereafter CPF-CF. CFIA consists of Rna15, Rna14, Clp1 and Pcf11, whereas CFIB includes only one component, the RNA-binding protein Hrp1. The enzymatic activities for cleavage and polyadenylation are present in the 14 subunit CPF complex that are further organized into three prominent modules: the nuclease module that is comprised of three subunits (Ysh1, Cft2 and Mpe1); the phosphatase module that consists of six subunits (Pta1, Ref2, Pti1, Swd2, Glc7, Ssu72); and the poly(A) polymerase module that contains five subunits (Cft1, Pfs2, Pap1, Fip1 and Yth1) (Casanal et al., 2017). The six subunits of the phosphatase module together with an additional factor, Syc1, form a distinct complex called APT (associated with Pta1), with partially distinct functions in termination of sn/snoRNA genes (Nedea et al., 2003; Lidschreiber et al., 2018). The subunit composition of CPF-CF and their related roles are presented in **Table 2-2**.

Termination by the CPF-CF pathway involves three successive steps: 1) recruitment of CPF-CF through interaction with RNAPII and the recognition of *cis*-regulatory elements on the nascent RNA; 2) cleavage of the nascent RNA at the poly(A) site, followed by the polyadenylation of the upstream cleavage product and degradation of the downstream cleavage fragment; 3) dismantling of the EC (Figure 2-6A).

For most *S. cerevisiae* protein coding genes, a poly(A) signal (PAS) serves as a 3' end processing site and a transcription termination signal (TTS). The PAS contains five different

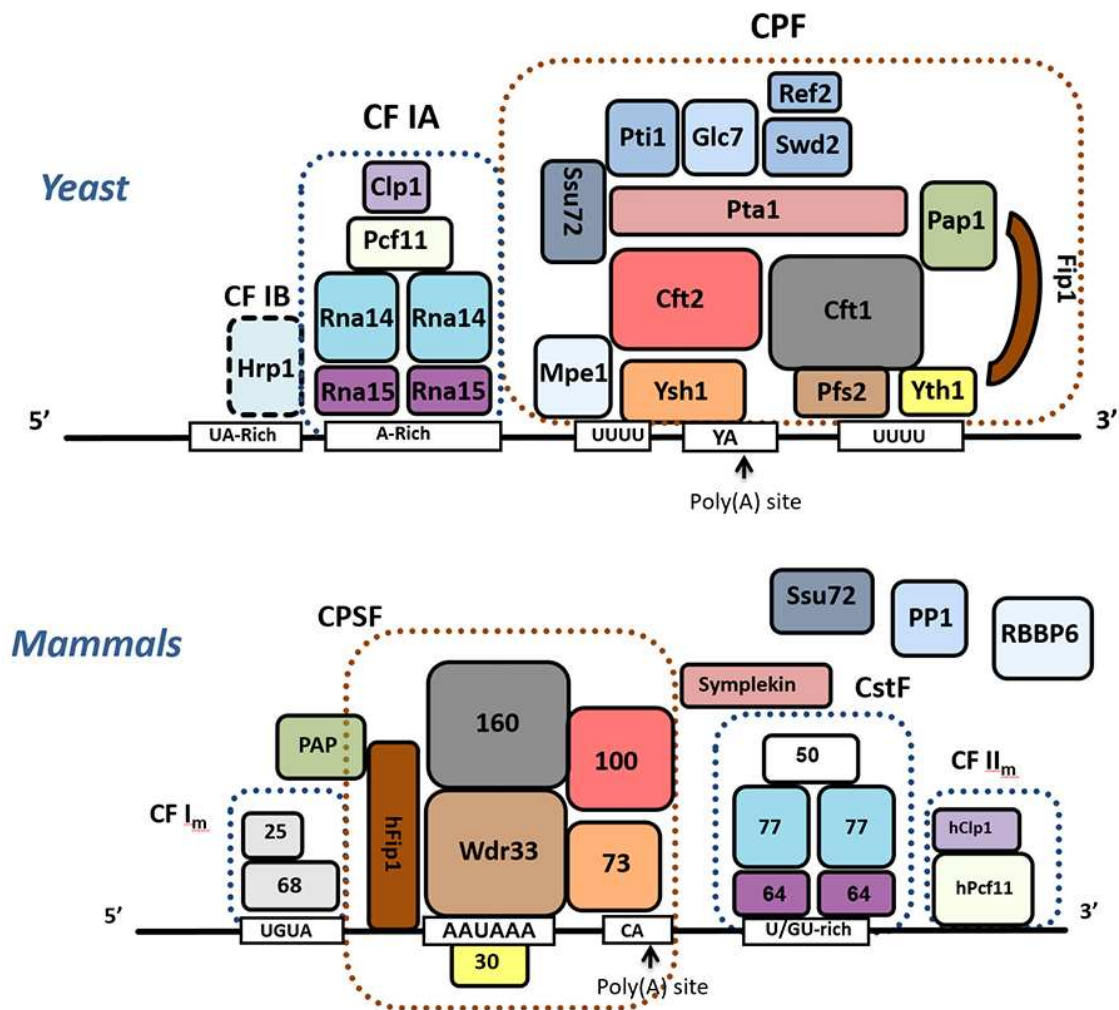


Figure 2-5: Cleavage and polyadenylation complex and its binding sequences.

Depicted is the subunit composition of the CPF-CF complex in yeast (upper panel). Mammals homologues are also presented here for comparison (lower panel). Conserved subunits are indicated in similar colors. Details in the text. Adapted from <https://gsbs.tufts.edu/facultyResearch/faculty/moore-claire/research>.

elements: the AU-rich efficiency element (EE) responsible for polyadenylation efficiency; the A-rich positioning element (PE) critical for precise 3' end processing; the poly(A) cleavage site; and the upstream and downstream U-rich enhancer elements (UUE and DUE) (Mischo and Proudfoot, 2013; Figure 2-5). Once RNAPII transcribes the PAS, CTD phosphorylation at Ser2 promotes the recruitment of the CPF-CF complex to these elements in the nascent pre-mRNA. Among the various CPF-CF proteins that interact with phosphorylated CTD, a key component of CFIA called Pcf11 contains a CTD interacting domain (CID) and can specifically interact with Ser2P CTD, thus enhancing its recruitment at the 3' end of genes (Komarnitsky et al., 2000; Lunde et al., 2010; Mayer et al., 2012). In addition, the EE element is loosely associated by the Hrp1 subunit of the CFIB subcomplex, and the PE element is bound by the Rna15

component of the CFIA subcomplex, which together define the RNA cleavage site (Valentini et al., 1999; Gross and Moore, 2001). Then, the endonuclease Ysh1, one of the catalytic subunits of CPF, cleaves the RNA at the polyadenylation site (Jenny et al., 1996). The co-transcriptional cleavage of the nascent transcript results in the formation of an uncapped 5' end in the downstream cleavage fragment attached to the EC, which is subsequently degraded by the nuclear 5'-3' exonuclease Rat1 in complex with Rai1 (Rat1 interacting protein) and Rtt103 that also contains a CID domain (Kim et al., 2004). Rtt103 recognizes the Ser2P and the Thr4P forms of the CTD (Harlen et al., 2016). On the other hand, the resulting upstream cleavage product containing a 3' OH is recognized by Pap1, the CPF poly(A)-polymerase subunit that catalyses the addition of a poly(A) tail (Lingner et al., 1991). The activity of Pap1 is regulated by Nab2, a protein that interacts with the CPF-CF complex. This protein binds to poly(A) tails and limits the addition of adenine residues to the 3' end of the RNA (Hector, 2002).

It is generally accepted that the cleavage of the nascent RNA by Ysh1 precedes the release of the EC from the DNA template, which occurs about ~200 bp downstream of the poly(A) site (Baejen et al., 2017; Schaughency et al., 2014). With regard to the mechanisms provoking RNAPII dissociation, two non-mutually exclusive models have been proposed: the allosteric model and the *torpedo* model (Figure 2-6A; reviewed in Porrua and Libri, 2015b; Richard and Manley, 2009). The allosteric model posits that the transcription through the poly(A) site causes a conformational change in the EC due to the loss of elongation factors and/or association of termination factors, which is followed by dissociation of RNAPII from the DNA template (Logan et al., 1987). The *torpedo* model proposes that progressive degradation by Rat1 of the downstream RNA product after cleavage leads to destabilization of the EC and promotes RNAPII release upon “catching up” with RNAPII (Kim et al., 2004; West et al., 2004; Park et al., 2015; Pearson and Moore, 2013; Baejen et al., 2017). However, several evidences suggest that these two mechanisms may act in concert to efficiently terminate transcription at protein-coding genes (Luo, 2006; Kaneko et al., 2007; West et al., 2008).

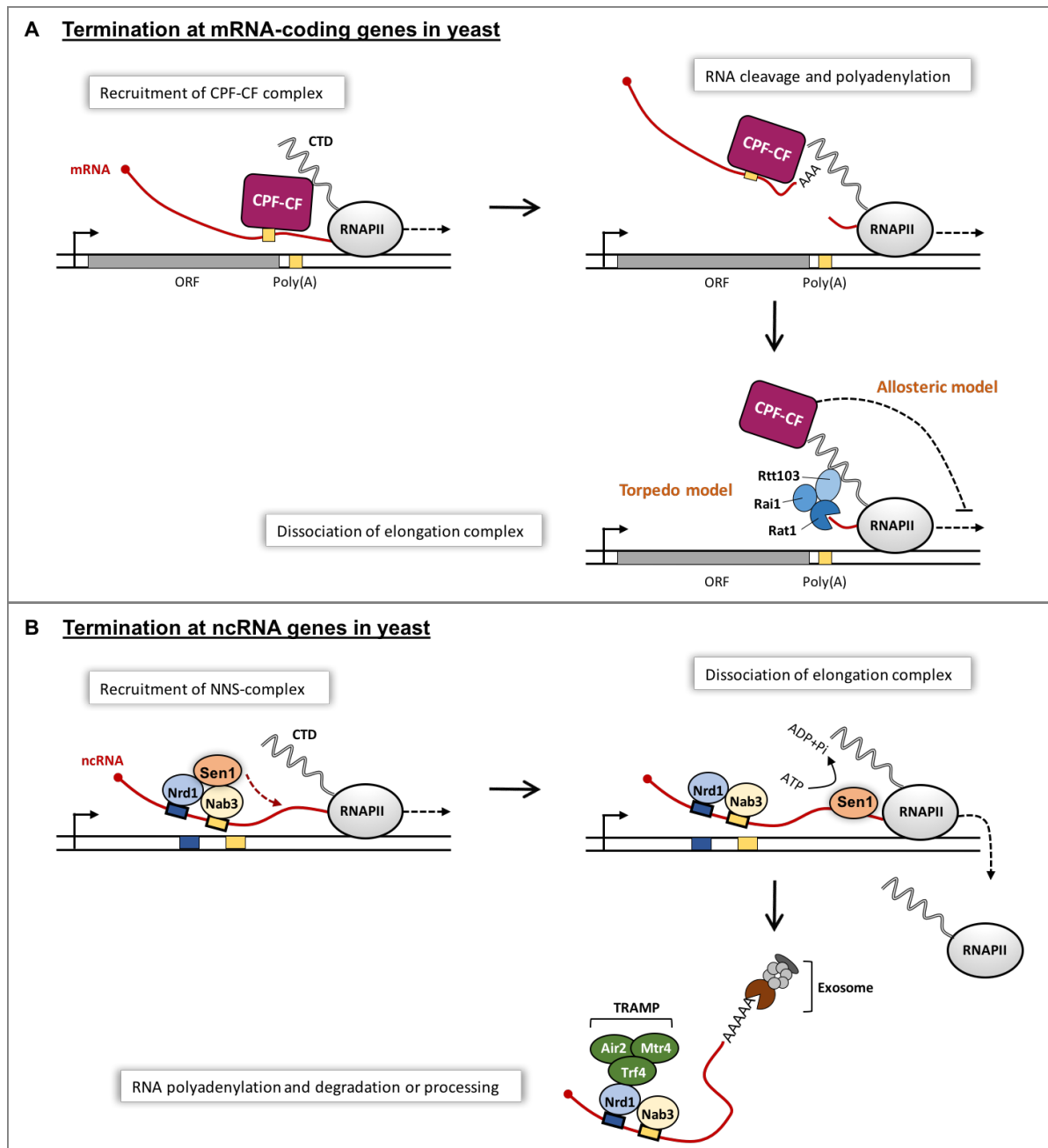


Figure 2-6: Major pathways for RNAPII transcription termination.

Transcription termination by the CPF-CF pathway. The CPF-CF is recruited via the interaction with the CTD of RNAPII and with sequence signals present on the nascent transcript (yellow box). The Ysh1 subunit of the CPF-CF complex cleaves the RNA, generating an uncapped 5' end that serves as an entry point for the Rat1 exonuclease. Two models are proposed for the dissociation of the elongation complex: the torpedo mechanism and the allosteric mechanism (details in the text). **(B)** Transcription termination by the NNS pathway. Nrd1 and Nab3 are recruited by recognition of Ser5P CTD and specific motifs on the nascent transcript. Sen1 is then recruited, translocates along the RNA and releases RNAPII. The resulting RNA transcript is subsequently degraded by the nuclear exosome and the TRAMP complex. Adapted from Porrua et al, 2016.

Table 2-2: Factors involved in RNAPII transcription termination.

| Yeast Complex | Yeast subunit | Human subunit | Proposed role | |
|---------------|---------------|---------------|--|---|
| CF | Pcf11 | PCF11 | Binds RNAPII Ser2P CTD; scaffolding protein; promotes RNA cleavage and Rat1 recruitment | |
| | CFIA | Rna14 | CSTF77 | Scaffolding protein, promotes RNA cleavage |
| | | Rna15 | CSTF64 | Binds poly(A) site RNA; promotes RNA cleavage and Rat1 recruitment |
| | | Clp1 | hClp1 | RNA kinase in human |
| | CFIB | Hrp1 | n.d. | Binds poly(A) site RNA; promotes recruitment of Pcf11 to RNAPII |
| Core CPF | Ysh1 | CPSF73 | Endonuclease, cleaves poly(A) site RNA, provides entry point for Rat1 | |
| | nuclease | Cft2 | CPSF100 | Binds RNAPII Ser2P and Ser5P CTD; binds poly(A) site RNA; promotes RNAPII pausing and RNA cleavage for Rat1 entry |
| | | Mpe1 | RBBP6 | RNA binding |
| | Polymerase | Pap1 | PAP | Poly(A) polymerase |
| | | Cft1 | CPSF160 | Binds RNAPII Ser5P and Ser2P CTD, binds poly(A) site RNA; promote RNAPII pausing and RNA cleavage for Rat1 entry |
| | | Pfs2 | WDR33 | Scaffold, RNA binding |
| | | Yth1 | CPSF30 | Binds poly(A) site RNA and RNAPII; promotes RNAPII pausing |
| | | Fip1 | hFip1 | Binds Pap1 |
| | APT | Pta1 | Symplekin | Scaffolding protein, bridges CFIA and CPF |
| | | Ref2 | n.d. | RNA-binding protein, scaffolding protein, bridges Pti1 and Glc7 |
| Pti1 | | n.d. | Scaffold protein, bridges Ref2 and Pta1 | |
| Phosphatase | | Swd2 | WDR82 | Scaffolding protein, subunit of histone methylation complex Set1 |
| | | Glc7 | PP1A | Ser/Thr phosphatase of RNAPII CTD; dephosphorylates Sen1 |
| Ssu72 | | SSU72 | RNAPII Ser5P CTD phosphatase; promotes recruitment of Pcf11 to RNAPII | |
| Syc1 | | n.d. | Paralog of Ysh1 | |
| NNS | Nrd1 | n.d. | RNA binding; binds RNAPII Ser5P CTD; | |
| | Nab3 | n.d. | RNA binding | |
| | Sen1 | Senataxin | ATP-dependent 5'-3' DNA/RNA helicase; binds RNAPII CTD; promotes Rat1 activity by exposing RNA | |
| | Rat1 | XRN2 | 5'-3' exonuclease, degrades Ysh1 (or CPSF73)-generated downstream cleavage product; promotes Pcf11 and Rna15 recruitment | |
| | Rai1 | DOM3Z | De-capping endonuclease, pyrophosphohydrolase; promote Rat1 stability and activity | |
| | Rtt103 | n.d. | Binds RNAPII Ser2P CTD; recruits Rat1 and Pcf11 to RNAPII | |

n.d., not detected;

Adapted from Kuehner et al., 2011 and Kumar et al., 2019; Lidschreiber et al., 2018

2.5.2 The NNS-dependent pathway

Transcription termination at most non-coding genes in *S. cerevisiae* is dependent on the NNS complex, which consists of two RNA-binding proteins, Nrd1 and Nab3, and a DNA/RNA helicase called Sen1 (reviewed in Arndt and Reines, 2015). All three NNS subunits are essential for growth in *S. cerevisiae* (Steinmetz and Brow, 1996; Wilson et al., 1994; Winey and Culbertson, 1988).

The NNS components

Nrd1 (Nuclear pre-mRNA Down-regulation) is a 63 kDa protein containing an N-terminal RNAPII CTD interaction domain (CID), a central RNA recognition motif (RRM), a Nab3 interaction domain (NabID) and a C-terminal proline and glutamine (Q/P) rich region (Figure 2-7). Nrd1 interacts genetically and physically with RNAPII (Conrad et al., 2000) and preferentially binds the Ser5P CTD, which is a mark of early elongation (Vasiljeva et al., 2008; Kubicek et al., 2012). Nrd1 CID also interacts with Trf4, a component of the TRAMP complex (see Box 2-2) involved in RNA degradation and processing (Tudek et al., 2014). Nrd1 is recruited to the nascent transcript by recognizing the GUAA/G motif (Carroll et al., 2004; Wlotzka et al., 2011; Porrua et al., 2012).

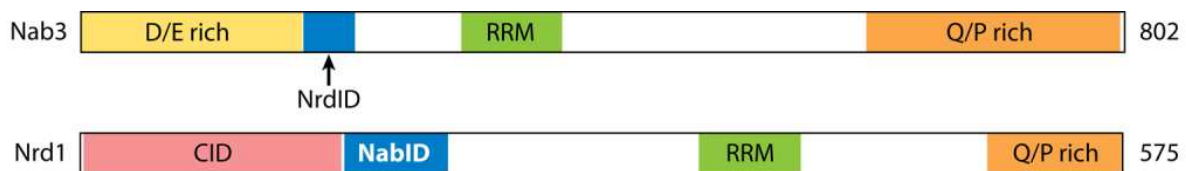


Figure 2-7: The domain structure of the Nab3 and Nrd1 proteins from *S. cerevisiae*.

The lengths of the proteins are indicated on the right (in amino acids). The Nrd1 interaction domain of Nab3 (NrdID; amino acids 204-248), the RNAPII CTD interaction domain of Nrd1 (CID; amino acids 6-151); the Nab3 interaction domain of Nrd1 (NabID; amino acids 151-214), the RNA recognition motifs (RRMs), and the Q/P rich and D/E rich domains are indicated in boxes with different colors, respectively. From Arndt and Reines, 2015.

Nab3 (Nuclear polyAdenylated RNA-Binding) is a 90 kDa protein that contains an N-terminal glutamate/aspartate (D/E) rich domain, a central RRM, a Nrd1 interaction domain, as well as a Q/P rich domain at its C-terminus (Figure 2-7). Nab3 interacts directly with Nrd1 forming a stable heterodimer, and specifically recognizes UCUUG motifs present on the nascent RNA (Carroll et al., 2004; Wlotzka et al., 2011; Porrua et al., 2012;). Moreover, AU-rich sequences frequently present close to Nab3 binding sites are found to be important for efficient termination (Porrua et al., 2012).

Sen1 (Splicing ENdonuclease) belongs to the superfamily 1B (SF1B) Upf1-like family of

helicases. It is a large (molecular mass of 252 kDa), and low-abundance nuclear protein (63-498 molecules/cell according to Ghaemmaghami et al., 2003; Newman et al., 2006; Kulak et al., 2014; Chong et al., 2015) that consists of a central helicase domain, an N-terminal domain (NTD) and a C-terminal unstructured region (Figure 2-8A). Sen1 interacts with the RNA without any known sequence specificity (Creamer et al., 2011; Porrua and Libri, 2013) and both the N- and C-terminal domains are involved in protein-protein interactions. Specifically, the N-terminal domain mediates the interaction with the CTD of RNAPII (Chinchilla et al., 2012; Han et al., 2020), while the C-terminal domain contains sequences that are important for its nuclear localisation and for the interaction with the phosphatase Glc7 and Nrd1 (Nedeia et al., 2008; Ursic et al., 2004; Chen et al., 2014; Han et al., 2020, see Figure 2-8). Sen1 is the only subunit of the NNS complex that possesses catalytic activity. It is an ATP-dependent helicase, which can unwind both DNA and RNA substrates (Han et al., 2017; Martin-Tumasz and Brow, 2015). Deletion of its N-terminal domain or mutation of the helicase domain provoke transcription termination defects *in vivo* (Chen et al., 2014; Finkel et al., 2010; Steinmetz et al., 2006; Han et al., 2020). Besides its role in transcription termination, Sen1 was also shown to function in the resolution of R-loops, which are structures that typically form during transcription when the nascent RNA invades and anneals with the DNA template (Mischo et al., 2011). Sen1 was also proposed to play an important role in the resolution of transcription-replication conflicts (Alzu et al., 2012), and in the repair of DNA damage (Li et al., 2016).

Mechanisms of NNS-dependent termination

Unlike the CPF-CF termination pathway, NNS-mediated termination is not linked to the endonucleolytic cleavage of the nascent RNA, but rather relies on the translocase activity of Sen1 helicase and is coupled to the nuclear RNA degradation pathway. The current model proposes that the NNS complex is recruited by the interaction of Nrd1 CID with RNAPII CTD, as well as by the binding of Nrd1 and Nab3 to the nascent RNA, which facilitates the recruitment of Sen1, albeit is not a strict requirement for Sen1 binding to its targets (Han et al., 2020). Following recruitment, Sen1 can translocate along the nascent RNA and dismantle the EC upon “catching up” with RNAPII (Figure 2-6B). Once the EC is dissociated, Nrd1 and Nab3 remain bound to the nascent RNA and promote the recruitment of the TRAMP complex and the exosome (**Box 2-2**) to trim or degrade the nascent transcript (Tudek et al., 2014). RNA degradation allows the release of Nrd1 and Nab3 for further cycles of termination (Villa et al., 2020).

Our team has long-standing interest in the mechanism of transcription termination by the NNS pathway. During the past years, the team has characterized in detail the function of the NNS complex. By using a highly purified *in vitro* system, it was shown that Sen1 alone is

sufficient to terminate RNAPII transcription, which requires its interaction with the nascent RNA and the hydrolysis of ATP to dismantle the paused EC. Importantly, the CTD of RNAPII was shown not to be necessary for Sen1-dependent termination *in vitro*. The ability to terminate transcription is specific to Sen1, as Upf1, a related helicase involved in the Nonsense-Mediated mRNA Decay (NMD) pathway for RNA quality control, cannot terminate RNAPII transcription *in vitro* (Porrúa and Libri, 2013a).

A subsequent study from our team established that the helicase domain of Sen1 is sufficient for transcription termination by RNAPII *in vitro*, indicating that this domain retains the essential features and activities involved in termination (Han et al., 2017). These study showed that Sen1 is a low-processivity helicase that can translocate along both single-stranded (ss) RNA and ssDNA in the 5' to 3' direction. Sen1 translocation along the RNA allows dismantling of the EC in a reaction that depends on the activity of its helicase domain. Furthermore, Sen1 can also promote forward translocation of stalled RNAPII, possibly by acting on the nascent RNA to exert a mechanical force on the EC (Han et al., 2017). The observation that Sen1 can alternatively promote either forward translocation or dissociation of paused RNAPII suggests that termination might require a particular state of the polymerase, possibly a persistent pause. The mechanism employed by Sen1 closely resembles that of the bacterial termination factor Rho, which is the key actor of one of the major pathways for

Box 2-2: The exosome and the TRAMP complex

The exosome is a highly conserved RNA-processing protein complex that plays a key role in RNA surveillance. It is localized in the nucleus and the cytoplasm and degrades aberrant non-coding and coding RNAs and catalyse the 3' end maturation of rRNAs, snRNAs and snoRNAs (Butler, 2002; Houseley and Tollervy, 2006). *In S. cerevisiae*, the exosome is composed of a nine-subunit core (Rrp4, Rrp40, Rrp41, Rrp42, Rrp43, Rrp45, Rrp46, Mtr3, and Csl4) and two catalytic subunits, the 3'-5' exonucleases Dis3 and Rrp6. However, Rrp6 is only present in the nuclear form of the exosome, thus distinguishing the nuclear exosome from the cytoplasmic exosome (Synowsky et al., 2009; Zhang et al., 2019).

The TRAMP complex is one of the best-characterized nuclear exosome cofactors. In *S. cerevisiae*, TRAMP is a heterotrimeric complex composed of a poly(A) polymerase (Trf4 or Trf5); a zinc-knuckle RNA-binding protein (Air 1 or Air2), and the RNA helicase Mtr4 (LaCava et al., 2005; Wyer et al., 2005; Houseley and Tollervy, 2006; Vanacova et al., 2005). TRAMP interacts with the exosome in the nucleus, and polyadenylates RNAs destined for Rrp6 and the core exosome, assisting in transcript recognition and exosome activation (Schmidt and Butler, 2013).

transcription termination in bacteria (see Box 2-4).

A single-molecule study performed in collaboration with Terence Strick's laboratory (IJM, Paris) (Wang et al, 2019) provided further insights into the process of Sen1-mediated transcription termination. By using a DNA scaffold construct compatible with magnetic-trapping assays, we could study in detail the kinetics of dissociation of a stalled RNAPII by Sen1. Interestingly, this study detected an intermediate RNAPII state during Sen1 action, in which the transcription bubble is apparently partially rewound. Partially rewinding of the transcription bubble might lead to a short and thus unstable RNA:DNA hybrid in the active center, which likely commits the EC to dissociation.

To explore the mechanisms of action of Sen1, the team has performed structural and biochemical studies in collaboration with Elena Conti's group (Max Planck, Munich) (Leonaitè et al., 2017). The crystal structure of the ~90 kDa helicase core (Sen1_{Hel}, aa 1095-1904) was obtained at 1.8 Å resolution. Based on the structure obtained, the overall domain organization of Sen1_{Hel} is similar to that of the helicase domain of other Upf1-like helicases, which includes two tandem RecA domains (RecA1 and RecA2) and additional subdomains. As Upf1 and other similar helicases, Sen1_{Hel} also contains two SF1B-specific accessory domains that extend on the surface of RecA1: subdomain 1B (the “barrel”) and subdomain 1C (the “prong”). Subdomain 1B is flanked by two antiparallel helices that pack against each other, forming the so-called “stalk” (Figure 2-8B). Common to Upf1-like helicases is that they bind nucleic acids in the same orientation, with the 3' end at RecA1 and the 5' end at RecA2, and unwind duplexes processively in the 5'-3' direction. However, Sen1_{Hel} shows several distinct features in the accessory subdomains when compared with the helicase domain of other Upf1-like helicases: the “barrel” has a more elaborate topology; the ordered portion of the “prong” is shorter; and most importantly, Sen1 has a distinct and evolutionarily conserved domain localized at the N-terminal end of Sen1_{Hel} called “brace” (aa 1097-1149). The “brace” connects the “stalk” and the “barrel” and appears to stabilize the overall fold of the protein. The structural and biochemical data suggested that the “brace” helps pulling the “barrel” toward the “prong”, thus shaping a favourable conformation for RNA binding and unwinding. More importantly, it was found that the “prong” is an essential element for 5' to 3' unwinding and for Sen1-mediated transcription termination both *in vivo* and *in vitro* (Leonaitè et al, 2017). Based on these observations, it was speculated that once Sen1 encounters RNAPII, the “prong” will insert into the RNA exit channel, which would lead to conformational changes and destabilization of the EC (discussed in Han and Porrua, 2018). In conclusion, the particular conformation determined by the “brace” and the distinctive characteristics of the “prong” are likely important for the specific function for Sen1 in termination of non-coding transcription.

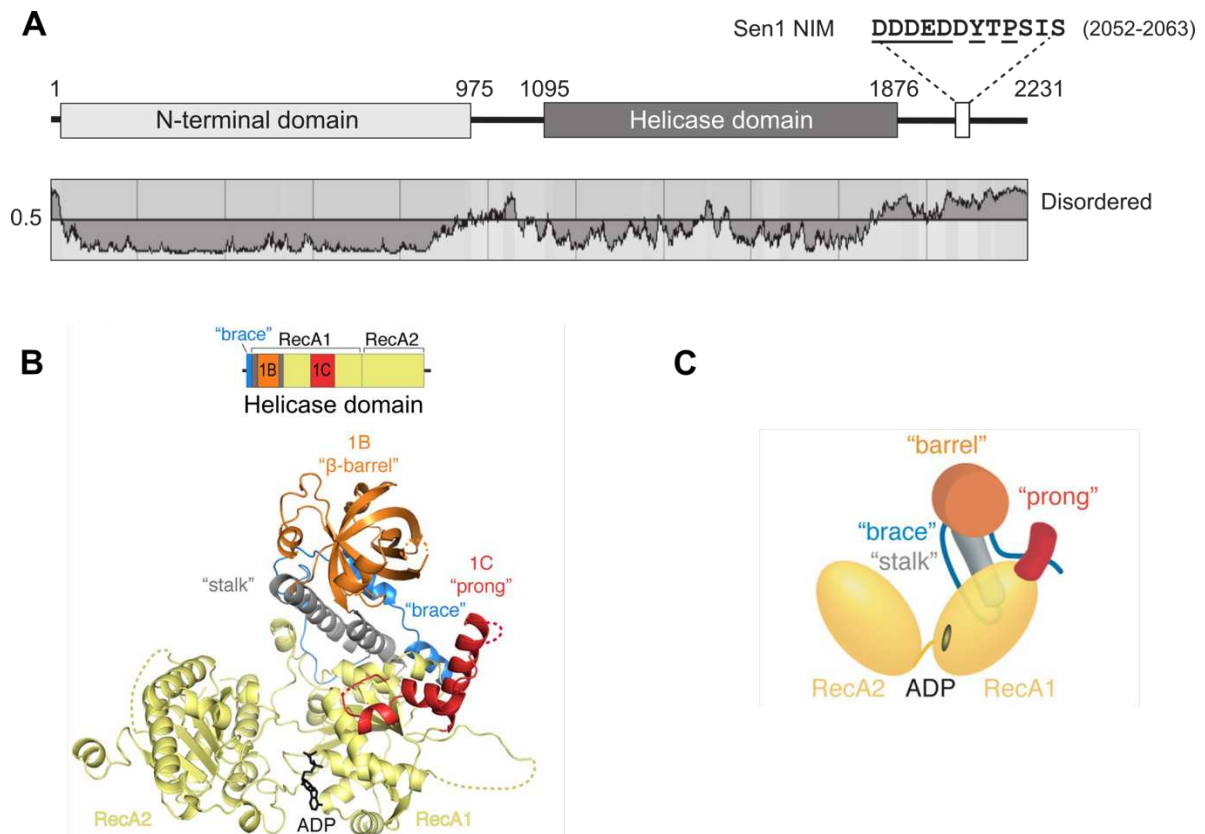


Figure 2-8: The domain organization and structural features of Sen1.

(A) Scheme of Sen1 protein. Globular domains are denoted by solid bars, and intrinsically disordered regions are indicated by a line. The disorder prediction was obtained by IUPred (Dosztányi et al., 2005). The sequence of the Nrd1-nteraction motif (NIM) is shown on top. From Han et al., 2020. (B) Structure of the Sen1 helicase domain determined in the absence of RNA. Dashed lines indicate disordered loops not resolved in the structure. A scheme with the domain organization is shown on top. (C) Simplified diagram of the main structural features according to (B). (B) and (C) from Leonaité et al., 2017.

The structure of the full-length Sen1 protein is currently missing, as Sen1 has a large N-terminal domain (aa 1-975) and a C-terminal intrinsically disordered region (aa 1930-2231). Although the N-terminal of Sen1 is not required for termination *in vitro* (Han et al., 2017), it is essential for cell growth as well as for RNAPII transcription termination *in vivo* (Han et al., 2020). However, the detailed function of these domains of Sen1 is still not completely understood.

The precise role of Nrd1 and Nab3 in RNAPII termination remains also unclear. As mentioned before, Nrd1 and Nab3 can recognize specific motifs on the nascent RNA, which is required for termination and is thought to provide the necessary specificity to the NNS complex (Wlotzka et al., 2011; Porrua et al., 2012; Schulz et al., 2013). Nrd1-Nab3 also recruits the Trf4 subunit of TRAMP, thus coupling termination by the NNS complex to RNA degradation (Tudek et al., 2014). It was proposed that Nrd1 and Nab3 also function as

adaptors to position Sen1 for timely and specific termination. However, the regions responsible for the interactions between Sen1 and Nrd1-Nab3 were identified by our team and removal of these regions did not significantly affect the efficiency of termination and only partially reduced the association of Sen1 with non-coding RNAs, indicating that the main role of Nrd1 and Nab3 in termination is not to promote Sen1 recruitment (Han et al., 2020).

Box 2-3: Pervasive transcription

The notion of “pervasive” or “hidden” transcription refers to the generation of a large ensemble of different RNAs distinct from those encode protein and those with established functions like tRNAs, rRNAs, snRNAs and snoRNAs (Neil et al., 2009; Jensen et al., 2013). Because these transcripts are rapidly degraded, they can be generally revealed only in cells defective for defined RNA degradation pathways. For instance, the absence of Rrp6, a catalytic subunit of the nuclear exosome, exposes a layer of RNAs known as “cryptic unstable transcripts” (CUTs) (Davis and Ares, 2006; Houalla et al., 2006; Wyers et al., 2005). Pervasive transcripts that are less sensitive to Rrp6 are therefore named “stable unannotated transcripts” (SUTs) (Xu et al., 2009). The absence of the cytoplasmic 5’-3’ exonuclease Xrn1 revealed another group of RNAs referred to as “Xrn1-sensitive unstable transcripts” (XUTs) (van Dijk et al., 2011). In addition, RNAs that are generated from Reb1-dependent termination events (see Figure 2-9B) are unstable and called “Reb1-dependent unstable transcripts” (RUTs) (Colin et al., 2014). CUTs, SUTs, XUTs and RUTs are all transcribed by RNAPII. CUTs and SUTs almost exclusively originate from nucleosome-depleted regions (NDRs) at the 5’ and 3’ ends of genes, and often appear to result from divergent transcription from gene promoters (reviewed in Jensen et al., 2013). Pervasive transcription is a common phenomenon conserved in prokaryotes and eukaryotes. It is potentially harmful for cell homeostasis, as it can interfere with transcription of canonical genes and generate toxic non-coding RNA molecules. Therefore, pervasive transcription needs to be controlled, which often depends on transcription termination and RNA degradation (Jensen et al., 2013). Some SUTs and XUTs probably rely on CPF-CF for termination (Marquardt et al., 2011; van Dijk et al., 2011), while for others, such as CUTs and a fraction of SUTs, termination and RNA degradation depends on the NNS pathway (Arigo et al., 2006; Thiebaut et al., 2006, Schulz et al., 2013).

2.5.3 Alternative pathways for RNAPII termination

Apart from the CPF-CF and NNS termination pathways, RNAPII can also be terminated through other mechanisms. In *S. cerevisiae*, two alternative pathways have been revealed: one is dependent on the Rnt1 endonuclease, the other one relies on various DNA-binding proteins that function as a roadblock for RNAPII (Figure 2-9).

Rnt1-dependent termination

Rnt1 is a nuclear dsRNA-specific ribonuclease, a homolog of bacterial RNase III, that plays a role in rDNA transcription, rRNA processing and snRNA 3' end maturation (Abou Elela et al., 1996; Abou Elela and Ares, 1998; Catala et al., 2008). Rnt1 is also involved in transcription termination in the absence of polyadenylation signals and mainly serves as a fail-safe termination pathway for protein-coding genes (Ghazal et al., 2009; Rondón et al., 2009). The model posits that Rnt1 recognizes and cleaves a stem-loop structure in the nascent transcript, generating a free 5'-OH at the downstream cleavage product that is subsequently targeted by Rat1 for degradation (Figure 2-9A). RNAPII would be destabilized and released from the DNA by Rat1 as the *torpedo* model proposed for the CPF-CF termination pathway. The upstream cleavage transcript is subjected to trimming or degradation by the TRAMP and exosome complexes, similarly to what occurs for transcripts produced by NNS-dependent termination (Figure 2-6B). Overall, Rnt1-dependent termination of RNAPII transcription is similar to that of RNAPI (see **Box 2-5**).

The roadblock termination pathway

A roadblock mechanism of termination was first studied for the transcription factor Reb1 by our laboratory (Colin et al., 2014). Reb1 is a sequence-specific DNA binding protein localized in the nucleus that was originally described as an activator for RNAPII and RNAPI transcription (Brandl and Struhl, 1990; Kulkens et al., 1992). It also plays an important role in the positioning and protection of nucleosome-free regions (NFRs) (Hartley and Madhani, 2009). Our team has shown that in yeast Reb1 can promote termination of RNAPII transcription. Reb1 bound to DNA is a roadblock for RNAPII as it induces pausing of the polymerase. The stalled polymerase is then ubiquitylated and degraded by the proteasome, thus provoking termination (Figure 2-9B). Reb1-dependent termination generates a class of unstable transcripts that are degraded in the nucleus by the TRAMP and exosome complexes, which were dubbed Reb1-dependent unstable transcripts, or RUTs (Colin et al., 2014). It was further demonstrated that roadblock termination can extend to various DNA-binding proteins including general regulatory factors (GRFs) such as Rap1, centromere-binding proteins, and

RNAPIII transcription factors like TFIIB. Roadblock termination occurs genomewide and functions as a fail-safe mechanism to neutralize transcriptional leakage from canonical termination pathways, which is a significant source of pervasive transcription (Candelli et al., 2018).

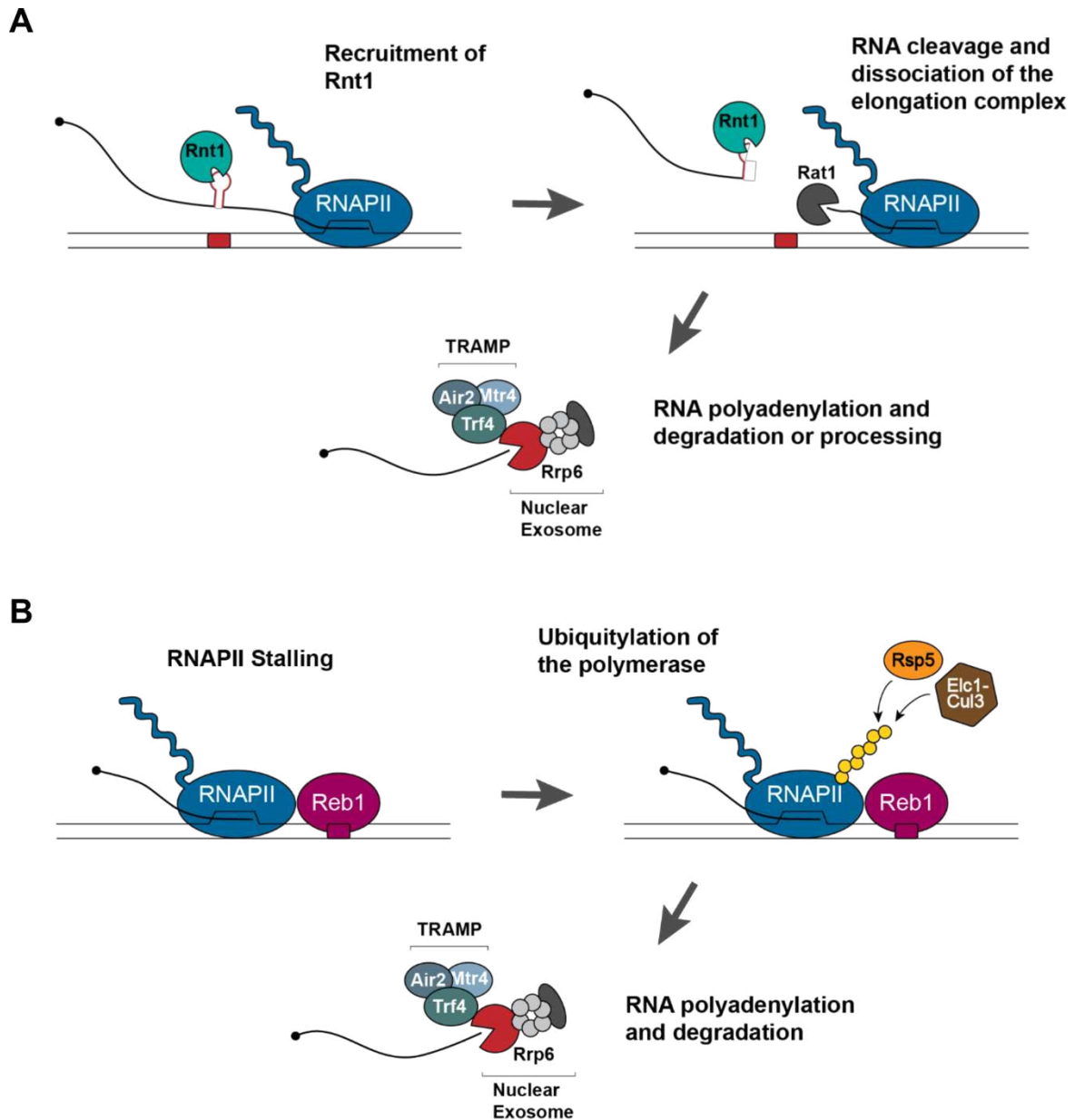


Figure 2-9: Alternative pathways of RNAPII transcription termination.

(A) Rnt1-dependent termination pathway. (B) Reb1-mediated roadblock termination pathway. Details see in the text. Rsp5, ubiquitin ligase; Elc1, elongin C that forms a complex with Cul3 that polyubiquitylates monoubiquitylated RNAPII to trigger its proteolysis. From Challal thesis, 2019.

Box 2-4: Transcription termination by bacterial RNAP

Bacterial transcription termination occurs via two major pathways: intrinsic termination and Rho-dependent termination (Figure 2-10; Ray-Soni et al., 2016; Porrua et al., 2016; Roberts, 2019). Intrinsic termination relies solely on a DNA sequence that, when transcribed, forms a GC-rich hairpin followed by a U-rich region. Upon pausing of RNAP at the U-tract, the hairpin forms into the RNA exit channel of RNAP, which induces dissociation of the EC. Three alternative mechanistic models have been proposed. The hyper-translocation model (Yarnell and Roberts, 1999; Santangelo and Roberts, 2004) posits that a steric clash of the hairpin with the RNA exit channel pushes RNAP forward without addition of nucleotides to the RNA 3' end, resulting in shortening of the RNA:DNA hybrid in RNAP main channel and destabilization of the EC. The hybrid shearing model (Larson et al., 2008; Molodtsov et al., 2014), proposes that formation of the hairpin rather generates a shearing force that pulls the transcript out of the complex, resulting in hybrid shortening. Finally, the allosteric model (Gusarov and Nudler, 1999; Touloukhonov et al., 2001) proposes that formation of the hairpin triggers a conformational change within the RNAP which destabilizes the EC (Figure 2-10).

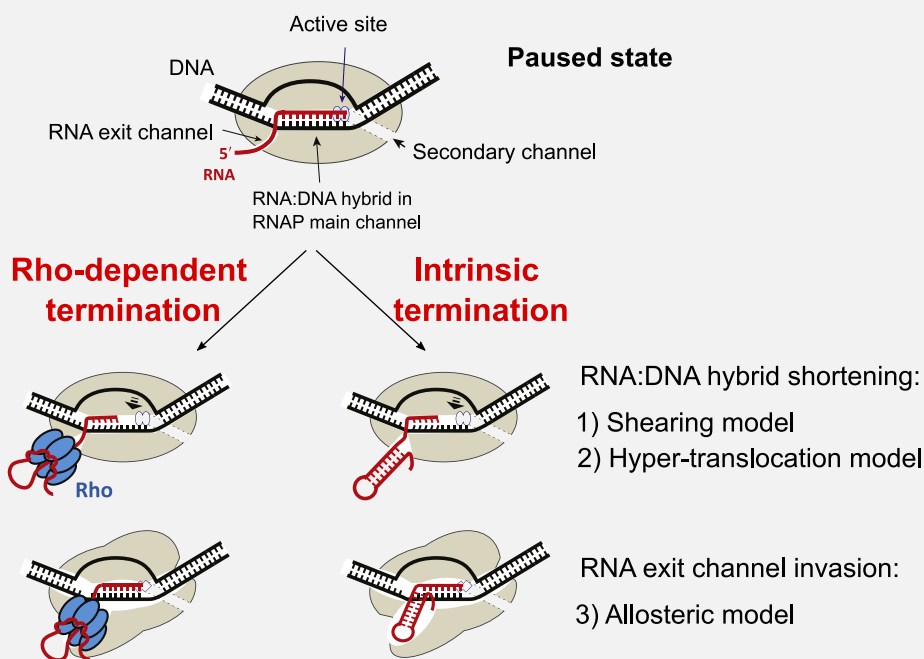


Figure 2-10: Transcription termination in bacteria.

A paused intermediate is common to all termination models. In the shearing model, RNA is pulled (by the action of Rho or hairpin folding) from an otherwise immobile EC, whereas, in the hyper-translocation model, hybrid shortening results from forward movement of RNAP. In the allosteric model, the hairpin or Rho invades the RNA exit channel triggering catalytic inactivation and conformational destabilization of the EC. Adapted from Porrua et al., 2016.

Box 2-4: Transcription termination by bacterial RNAP (continued)

The other main pathway for termination depends on the protein Rho, a highly conserved homohexameric RNA helicase (Brennan et al., 1987). Rho interacts with the nascent RNA at the so-called Rho utilization (rut) sites, which are 85-100 nt long regions, rich in cytosine and poor in guanine with little propensity to form secondary structures (Ciampi, 2006). Once bound, Rho translocates along the nascent RNA in the 5' to 3' direction to catch up with the RNAP and promote its release from the DNA. The same models as for intrinsic termination have been proposed to explain the mechanisms of EC dissociation by Rho (Figure 2-10; reviewed in Ray-Soni et al., 2016 and Porrua et al., 2016).

Box 2-5: Transcription termination by RNAPI

In yeast, ribosomal DNA (rDNA) is located in chromosome XII and consists in an approximately 1-2 Mb region composed of 150-200 tandem copies of a 9.1 kb region. Each repeat contains the genes encoding for the 18S, 5.8S and 25S rRNAs, which are transcribed by RNAPI, and the 5S rRNA, which is transcribed by RNAPIII. The different genes are separated by various spacer regions (Figure 2-11) and the tandem repeats are separated by an intergenic sequence (IGS), where termination of RNAPI transcription occurs (Venema and Tollervey, 1999). Termination has been proposed to involve a roadblock mechanism, similar to the one previously mentioned in Figure 2-9B. The IGS contains one or several recognition sequences for NTS1 silencing protein (Nsi1), a DNA-binding protein of the Myb-family (Reiter et al., 2012; Merkl et al., 2014) preceded by an Rnt1 cleavage site. RNAPI is roadblocked by Nsi1 bound to the IGS DNA and Rnt1 cleaves a stem-loop in the pre-rRNA generating an entry point for the exonuclease Rat1. Subsequently, Rat1 progressively degrades the RNAPI-bound transcript and promotes the release of RNAPI (Kawauchi et al., 2008; Hage et al., 2008). It was proposed that Sen1 aids Rat1 function by removing possible RNA secondary structures that might impair Rat1 progression (Kawauchi et al., 2008).

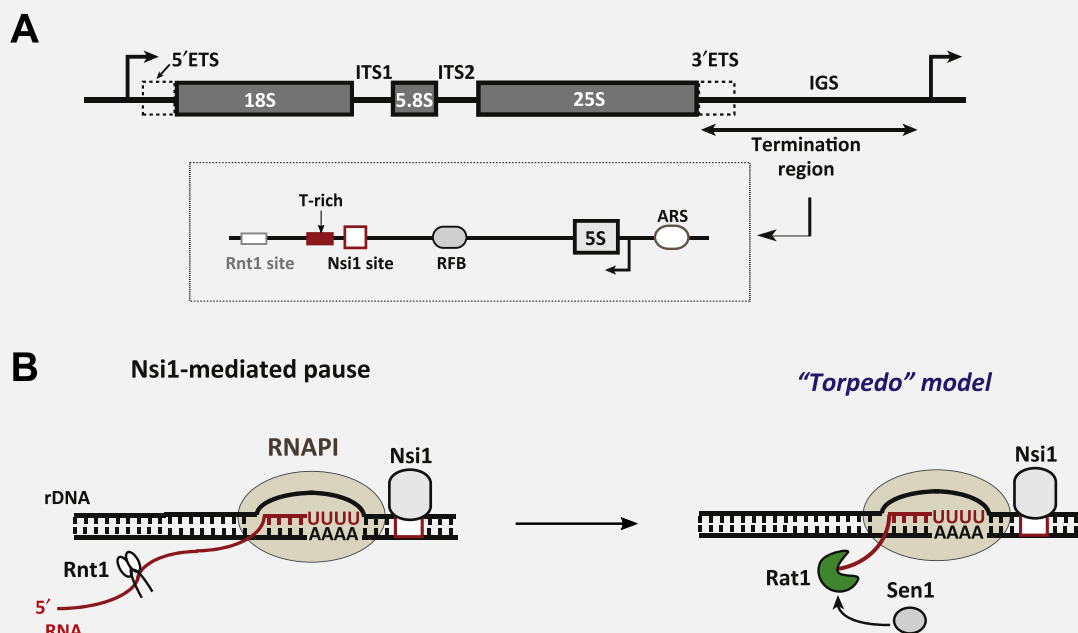


Figure 2-11: Transcription termination by RNAPI in yeast.

(A) Structure of the rDNA locus. (B) Model for RNAPI transcription termination. More details in the text. Abbreviations: ITS, internal transcribed spacer; ETS, external transcribed spacer; IGS, intergenic sequence; RFB, replication fork barrier; ARS: autonomously replicating sequence. Adapted from Porrua et al., 2016.

Chapter 3 Transcription by RNA polymerase III

RNA polymerase III (RNAPIII) is the largest of the three classical eukaryotic RNA polymerases (Figure 1-2; **Table 1-1**). It is specialized in the synthesis of short, abundant and structured non-coding RNAs (ncRNAs), such as nuclear tRNAs, the 5S rRNA and the spliceosomal U6 snRNA, many of which are involved in protein biosynthesis. Studies of the genomic occupancy of RNAPIII and its associated transcription factors have expanded the transcriptome of RNAPIII to include a number of additional non-coding RNAs, which have been found to actively regulate various essential processes in the cell (**Table 3-1**). RNAPIII transcriptional activity is highly regulated in response to multiple environmental cues, thus making it a key target for the regulation of cell growth, proliferation and differentiation (Hoffmann et al., 2016a, Willis and Moir, 2018).

3.1 The RNAPIII structure

RNAPIII is a multi-protein complex composed of 17 subunits with an overall molecular mass of approximately 700 kDa (Figure 1-2; **Table 1-1**; and Figure 3-1A). The subunit composition and architecture of RNAPIII have been summarized in a previous review (Vannini and Cramer, 2012). Former insight into the RNAPIII topology was based on several low-resolution electron cryo-microscopy (cryo-EM) studies (Fernandez-Tornero et al., 2007, Fernandez-Tornero et al., 2010, Vannini et al., 2010). In 2015 the first atomic structure of yeast elongating RNAPIII (holo RNAPIII, with a DNA:RNA hybrid in the active centre and downstream double-stranded DNA in the cleft) was solved at 3.9 Å resolution, together with two different conformations of the unbound RNAPIII (apo RNAPIII) at 4.6 and 4.7 Å resolution (Hoffmann et al., 2015). Three years later, atomic structures of the yeast RNAPIII preinitiation complex (PIC) comprising promoter-bound RNAPIII and TFIIIB have been published by two independent groups (Abascal-Palacios et al., 2018, Vorländer et al., 2018). Finally, very recently, cryo-EM structures of human RNAPIII have also been solved by the same laboratories (Ramsay et al., 2020, Girbig et al., 2021). These structural studies allow a better understanding of RNAPIII subunit positioning and its transcriptional mechanisms.

All multi-subunit RNAPs resemble a “crab claw” including several functional domains (Figure 3-1D). The two largest subunits of RNAPIII, C160 and C128, form the active site and the DNA-binding cleft. The cleft and ABC27, ABC23, ABC14.5, ABC10 α , ABC10 β , AC40 and AC19, together with C11 constitute the core enzyme. C11 is involved in transcription termination and RNA cleavage, with its C-terminal domain being structurally and functionally related to the C-terminal zinc-ribbon domain of the RNAPII elongation factor TFIIIS

(Arimbasseri and Maraia, 2015, Chédin et al., 1998).

RNAPIII also contains three distinct subcomplexes on the periphery of the core enzyme (C25-C17, C53-C37, and C82-C34-C31). The C25-C17 subcomplex forms the RNAPIII stalk that protrudes from the polymerase core on the C160 side and is involved in transcription initiation (Vannini and Cramer, 2012) and in the binding of single-stranded exiting RNA (Jasiak et al., 2006). The ten-subunit core and the heterodimeric stalk are structurally conserved among the three RNAPs in eukaryotes. The second subcomplex, the C53-C37 heterodimer, is situated on the C128 lobe and extends into the DNA-binding channel. This subcomplex is involved in RNAPIII transcription initiation and termination, and

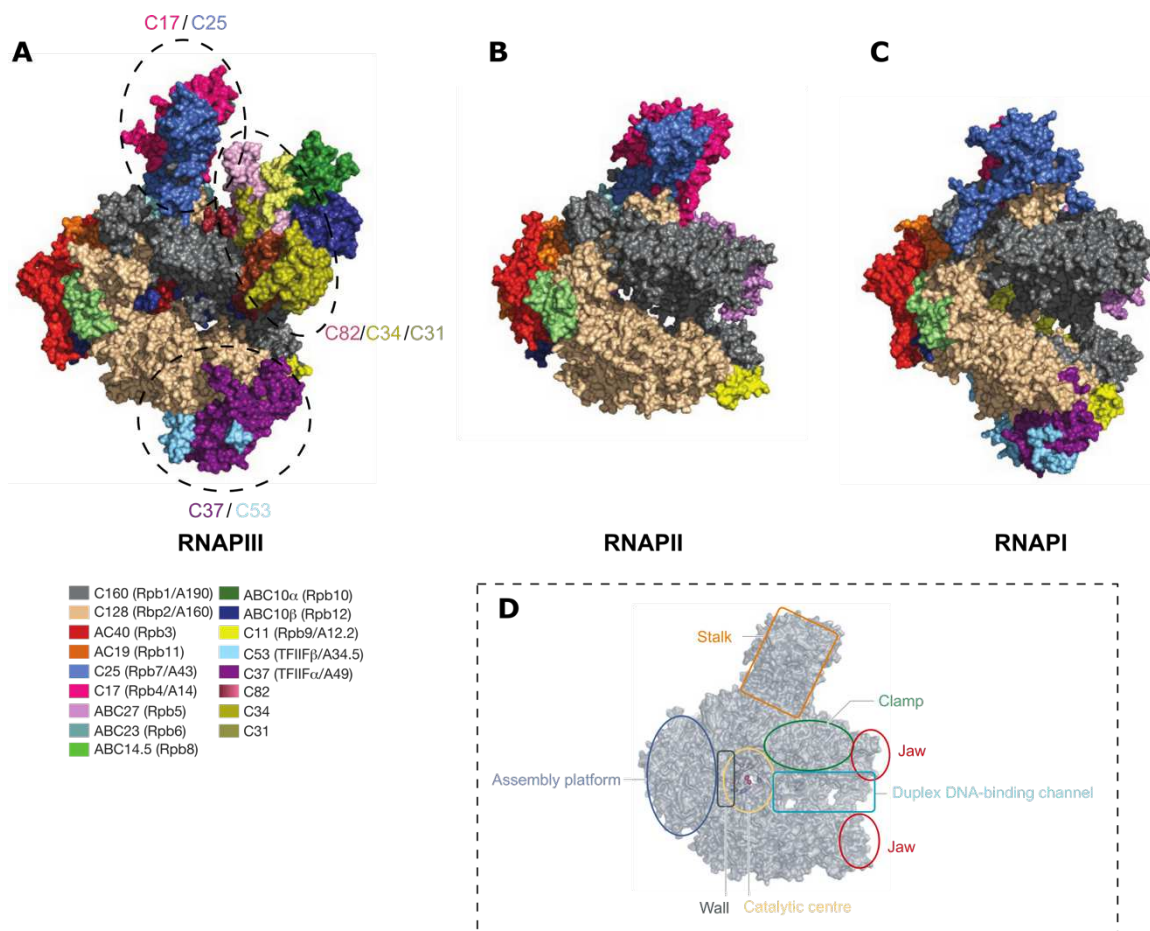


Figure 3-1: Structures of eukaryotic RNA polymerases.

(A) Surface view of the elongating RNAPIII compared to (B) RNAPII and (C) RNAPI. Homologous subunits in RNAPII and RNAPI are coloured based on RNAPIII as indicated in the legend. Three peripheral subcomplexes of RNAPIII (C25-C17, C53-C37, and C82-C34-C31) are indicated by black circles. (A-C) Adapted from Hoffmann et al., 2015. (D) Overall architecture of RNA polymerase based on the 12-subunit RNAPII (PDB 1Y1W). This simplified diagram of RNA polymerase shows important structural and functional features including the assembly platform, the active site, the DNA-binding channel, the jaws and the wall, clamp and stalk domains. The two active-site Mg^{2+} ions are indicated as magenta spheres. Adapted from Werner and Grohmann, 2011.

participates with C11 in the process of facilitated reinitiation (Arimbasseri and Maraia, 2015, Kassavetis et al., 2010, Landrieux et al., 2006; see below). C37 and C53 are distantly related to RNAPI A49 and A34.5, respectively, and to the RNAPII transcription factor TFIIF α and TFIIF β , respectively (Vannini and Cramer, 2012, Carter and Drouin, 2010, Geiger et al., 2010). The last subcomplex, the C82-C34-C31 heterotrimer, is placed on C160 clamp head in close proximity to the stalk and is involved in transcription initiation (Vannini and Cramer, 2012, Thuillier et al., 1995) and in the recruitment of RNAPIII to TFIIB (Brun et al., 1997, Khoo et al., 2014). The C82-C34-C31 subcomplex is RNAPIII-specific, but C82 and C34 are distantly related to TFIIE α and TFIIE β , respectively (Carter and Drouin, 2010).

3.2 The RNAPIII transcriptome

In eukaryotes, the genes transcribed by RNAPIII are normally referred to as class III genes. RNAPIII transcripts were previously thought to be restricted to only a small set of infrastructural non-coding RNAs. However, extensive work on RNAPIII transcriptomes in yeast and metazoans has discovered many new RNAPIII-dependent genes. The known set of yeast RNAPIII-synthesized RNAs is provided in **Table 3-1**. RNAPIII in mammals produces additional RNAs such as the vault RNA, Y RNAs and many small interspersed elements (SINEs). The biological functions of some of these RNAs have been well characterized, for instance those of tRNAs, the 5S rRNA, U6, the RNA component of RNase P, 7SL and snR52. However, the cellular functions of many others remain unclear.

tRNAs

Transfer RNAs are RNAPIII transcripts that serve as adaptor molecules in the biosynthesis of proteins. Mature tRNAs have a cloverleaf-shaped secondary structure including a D-loop (also known as DHU loop because it contains the dihydrouridine base), a T-loop (also named T ψ C loop due to the presence of thymidine, pseudouridine and cytidine bases), an anticodon loop (that specifies an amino acid), a variable loop and an acceptor stem. The cloverleaf structure further folds into a tertiary structure, a “L-shape” structure, which is maintained by hydrogen bonds (Figure 3-2A; Kirchner and Ignatova, 2015).

In *S. cerevisiae*, there are 275 nuclear tRNA genes including *tX(XXX)D*, with unknown specificity but very similar to a serine tRNA gene. The length of tRNA genes is 71-133 nt and 61 of them have an intron (Lesniewska and Boguta, 2017). tRNAs that bear the same anticodon belong to the same isoacceptor family, and isoacceptors that translate to the same amino acid are grouped into the same isotype. There are 64 isoacceptors in theory according to the genetic code, however, most eukaryotic genomes encode only ~42 isoacceptors and

some isoacceptors are universally missing (Chan and Lowe, 2016). On the other hand, tRNA genes show an important redundancy both in prokaryotes and in eukaryotes so that most isoacceptors are produced by more than one gene in the genome. A summary of isoacceptor tRNAs in *S. cerevisiae* is provided in Figure 3-2B.

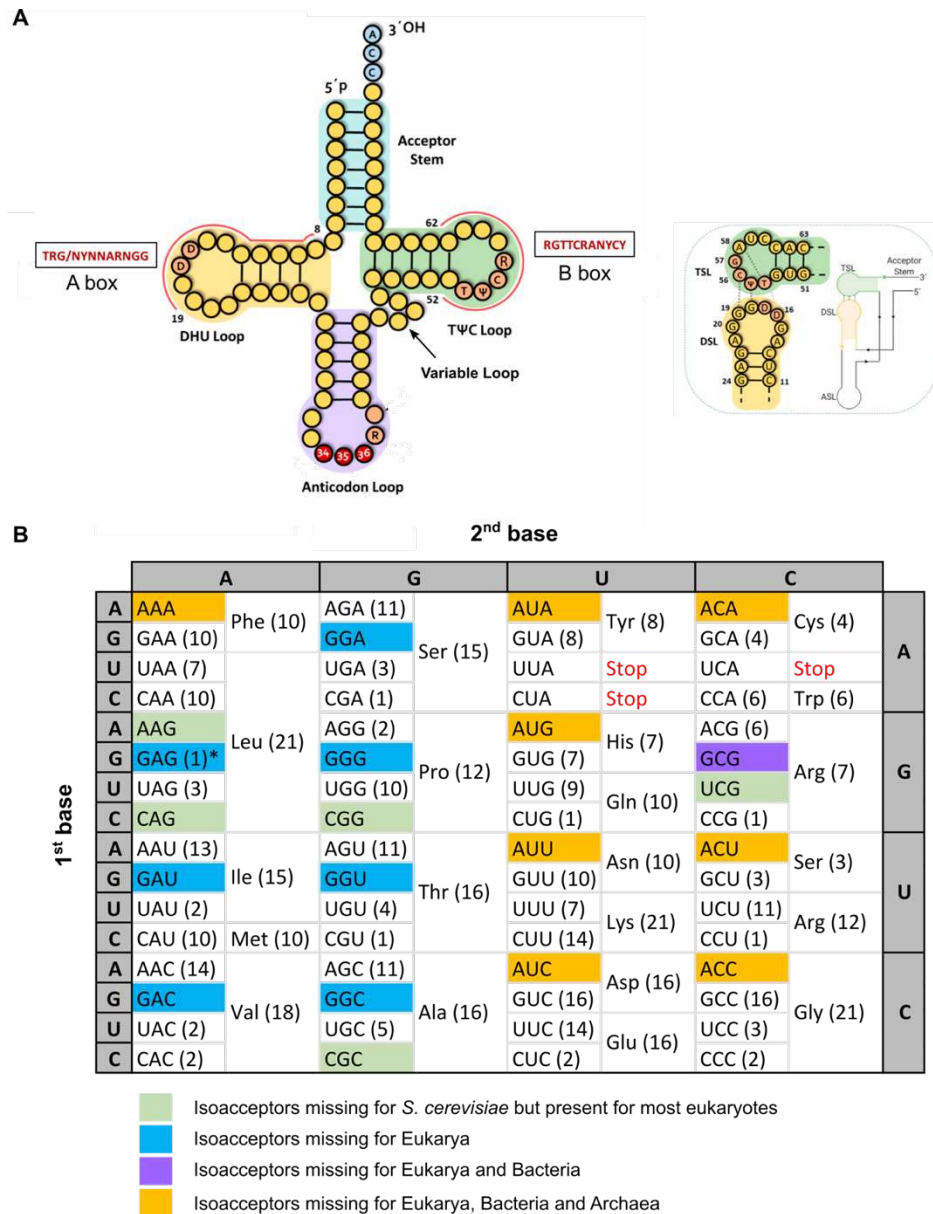


Figure 3-2: tRNA structure, isotypes and isoacceptors.

(A) Classical cloverleaf structure of tRNA. The conventional IUB/IUPAC degenerate DNA alphabet is used in this figure: **R** (purine), A or G; **N** (any), A, C, G or U; **Y** (pyrimidine), C or T. Main interactions supporting the L-shape of tRNAs are shown in the right dashed box. Adapted from Ehrlich et al., 2021. (B) Anticodon chart showing gene copy number of tRNA isotypes and isoacceptors in *S. cerevisiae* (genome version R64-2-1). 274 of nuclear tRNA genes are analysed (the 'tX(XXX)D' tRNA gene with unknown function is excluded). Number of genes for each isotype or isoacceptor is indicated within brackets. The coloured boxes correspond to missing tRNA isoacceptors in the different kingdoms, according to Ehrlich et al., 2021. *The 'GAG' isoacceptor tRNA is missing for most eukaryotes but there is one copy in *S. cerevisiae*.

5S rRNA

5S ribosomal RNA is a component of the large (60S) ribosomal subunit that in *S. cerevisiae* is encoded by 6 repeated transcription units known as *RDN5* genes (*RDN5-1* to *RDN5-6*). *RDN5-1* and *RDN5-2* are placed within the *RDN1* locus, while *RDN5-3* through *RDN5-6* are positioned at sites distal to *RDN1*, in a 3.6 kb repeated region (McMahon et al., 1984). The scheme of one *RDN1* locus has been shown in Figure 2-11A. These RNAs, together with 5.8S, 18S and 25S rRNAs (28S in mammals) that are processed from 35S (45S in mammals) rRNA precursor produced by RNAPI, are fundamental parts of the protein synthesis machinery.

U6 snRNA

SNR6-encoded U6 spliceosomal RNA (U6 snRNA) is the RNA component of U6 small nuclear ribonucleoprotein (snRNP). snRNP is an RNA-protein complex that combine with various other proteins to form a large ribonucleoprotein complex, namely the spliceosome that removes introns from precursor mRNA (pre-mRNA) before its translation into protein. U6 snRNA is the most highly conserved spliceosomal snRNA across species (Brow and Guthrie, 1988) which can directly mediate the catalysis of pre-mRNA splicing by the spliceosome (Fica et al., 2013).

RNase P RNA and human RNase MRP RNA

The yeast *RPR1* gene (*RPPH1* in humans, for ribonuclease P RNA component H1) encodes the RNA component of the nuclear ribonuclease P (RNase P), an enzyme mainly involved in the maturation of the 5' ends of tRNA precursors (Baer et al., 1989, Lee et al., 1991). Another related gene is the human *RMRP*, which encodes the RNA subunit of the RNase MRP complex (MRP stands for mitochondrial RNA processing). RNase MRP is a ribonucleoprotein complex evolutionarily linked to RNase P (Zhu et al., 2006) that plays a role in the initiation of mitochondrial DNA replication and in the processing of rRNA precursors in the nucleus (Chang and Clayton, 1989, Schmitt and Clayton, 1993).

7SL RNA

The 7SL RNA, also called SPR RNA, encoded by *SCR1* is the longest RNAPIII transcript in *S. cerevisiae* with a length of 522 nt. It forms the scaffold of the signal recognition particle (SRP), a ribonucleoprotein involved in targeting proteins to the endoplasmic reticulum membrane (Hann and Walter, 1991). The *SCR1* gene is in general well conserved among eukaryotes but the type of promoter is different in yeast and metazoans (Dieci et al., 2002, Dieci et al., 2007).

snR52 snoRNA

The small nucleolar RNA snR52 is the only snoRNA transcribed by RNAPIII (Harismendy et al., 2003, Roberts et al., 2003, Marck et al., 2006). snoRNAs are also often referred to as guide RNAs because they direct chemical modifications in other RNAs. In particular, snR52 belongs to the box C/D class of snoRNAs, which contain the conserved C box (UGAUGA) and the D box (CUGA) and function in directing site-specific 2'-O-methylation of other RNAs (Galardi et al., 2002).

Other RNAPIII-transcribed ncRNAs

Yeast RNAPIII synthesizes several additional short ncRNAs whose function is yet to be determined, such as RNA170 and ZOD1 RNAs, as shown in **Table 3-1**. Human RNAPIII is also responsible for producing SINEs, 7SK RNAs, Vault RNAs, Y RNAs, *etc* (Dieci et al., 2007).

Short interspersed nuclear elements (SINEs) are non-autonomous retrotransposons evolutionally derived from other RNAPIII-transcribed genes. Eukaryotic genomes can harbour more than a million of SINE copies, the bulk of which (66%) have a length from 150 to 300 bp (Kramerov and Vassetzky, et al., 2011).

7SK RNA is an abundant non-coding RNA found in a small nuclear ribonucleoprotein complex (snRNP), that has been involved in the regulation of RNAPII transcription by controlling the positive transcription elongation factor P-TEFb (Diribarne and Bensaude, 2009, Peterlin et al., 2012).

Vault RNAs, or vtRNAs, are small RNAs (~100 nt) present in a very large cytoplasmic ribonucleoprotein particle (RNP) known as vault (Stadler et al., 2009). They have been implicated in a broad range of cellular functions including multidrug resistance of human tumors and are thought to partake in intracellular and nucleocytoplasmic transport (Berger et al., 2009; van Zon et al., 2006).

Y RNAs are a family of small RNAs (~100 nt) that show high conservation in metazoans, and appear to be required for mediating the initiation of chromosomal DNA replication, regulating the autoimmune protein Ro60 and generating smaller RNA fragments following cellular stress (Hall et al., 2013).

Table 3-1: The RNAPIII transcriptome.

| Name | Gene | Length (bp) | Description | References |
|-------------|------------------|-------------|--|---|
| tRNAs | <i>tDNAs</i> | 71-133 | 275 nuclear tRNA genes found in <i>S. cerevisiae</i> , involved in protein synthesis | Goffeau et al., 1996 |
| 5S rRNA | <i>RND5</i> | 121 | Component of large (60S) ribosomal subunit | - |
| U6 snRNA | <i>SNR6</i> | 112 | Involved in pre-mRNA splicing, promoter organization is evolutionarily variable | Brow and Guthrie, 1988 |
| RNase P RNA | <i>RPR1</i> | 369 | RNA component of nuclear endoribonuclease RNase P, involved in pre-tRNAs 5' leader processing | Lee et al., 1991 |
| RNA170 | <i>RNA170</i> | 169 | Unknown function | Olivas et al., 1997 |
| 7SL RNA | <i>SCR1</i> | 522 | RNA subunit of signal recognition particle (SRP) ribonucleoprotein complex. | Briand et al., 2001 |
| snR52 | <i>SNR52</i> | 92 | C/D box small nucleolar RNA, served as a methylation guide for rRNA | Harismendy et al., 2003, Roberts et al., 2003 |
| ETC1-ETC8 | <i>ETC1-ETC8</i> | 23 | Abbreviation for "extra TFIIC", eight loci occupied only by TFIIC (<i>ETC5</i> is located within <i>RNA170</i>) | Moqtaderi and Struhl, 2004 |
| ZOD1 | <i>ZOD1</i> | 58 | Abbreviation for "zone of disparity", an ancient gene for tRNA-Ile | Guffanti et al., 2006 |
| ETC10 | <i>ETC10</i> | - | Loci occupied by TFIIC only | Nagarajavel et al., 2013 |
| TLT1-TLT6 | <i>TLT1-TLT6</i> | - | Potential loci revealed by RNAPIII CRAC, named for "tRNA-like transcripts", of unknown function, expression of TLT1 and TLT6 RNA were confirmed by northern blot | Turowski et al., 2016 |

Summarized according to Leśniewska and Boguta, 2017

3.3 Transcription initiation

RNAPIII transcription involves three general transcription factors (GTFs): TFIIIA, TFIIIB and TFIIIC. Metazoans additionally require SNAPc to transcribe a specific group of class III genes (**Table 3-2**). Transcription initiation follows several key conserved steps. First, GTFs are recruited sequentially to the promoter, followed by the association of RNAPIII around the transcription start site (TSS) to form the pre-initiation complex (PIC). Afterwards, the DNA double strand is melted and the transcription bubble is formed, followed by the initiation of RNA synthesis (Hoffmann et al., 2016b).

3.3.1 Basal RNAPIII transcription factors

TFIIIA

Transcription Factor III A (TFIIIA) is a single protein that plays a major role in the synthesis of the 5S rRNA by binding to the internal control regions (ICR) of the 5S rRNA genes and then serving as a platform for TFIIIC recruitment. It contains nine conserved zinc fingers at its N-terminal domain, which carry out sequence-specific DNA and RNA binding activity. The C-terminal domain of TFIIIA is involved in the transactivation process possibly by interacting with other general factors. In addition to the recognition of promoter sequence in DNA, TFIIIA can also bind to the 5S RNA to form the 7S ribonucleoprotein particle (RNP) and the 42S RNP complex (Layat et al., 2013).

TFIIIB

Transcription Factor III B (TFIIIB) is composed of three subunits: the TATA binding protein (TBP), TFIIIB-related factor 1 (Brf1) and B'' (Bdp1, for B double prime). Early work showed that Brf1 interacts with TBP forming the B' domain, which is able to bind the Bdp1 subunit (Kassavetis et al., 1992a). In vertebrates, Brf1 is replaced by the TFIIIB-related factor 2 (Brf2) at type 3 promoters of class III genes. TBP interacts tightly with Brf1 to form the so called B' fraction, which binds a specific sequence upstream of the transcription start site. Bdp1 is only weakly associated with other components of the TFIIIB complex in the absence of DNA but it is required for the formation of a stable TFIIIB-DNA complex. As a transcription initiation factor, TFIIIB is involved in the recruitment of RNAPIII to the promoter and promotes the transition from a closed to an open RNAPIII preinitiation complex. Moreover, TFIIIB has been shown to act as a genomic roadblock to induce the termination of neighboring transcribing RNAPIIs and the dissociation of the replisome (Gouge et al., 2017, Roy et al., 2016; Candelli et al., 2018).

Table 3-2: Subunit composition of RNAPIII and its transcription factors.

| | Human Subunit | Gene | MW (kDa) | Yeast Subunit | Gene | MW (kDa) |
|--|--------------------|---------------|----------|------------------|--------------|----------|
| RNAPIII: transcribing enzyme | | | | | | |
| | RPC1 (RPC155) | <i>POLR3A</i> | 155.6 | RPC1 (C160) | <i>RPO31</i> | 162.3 |
| | RPC2 | <i>POLR3B</i> | 127.8 | RPC2 (C128) | <i>RET1</i> | 129.5 |
| | RPC3 (RPC62) | <i>POLR3C</i> | 60.6 | RPC3 (C82) | <i>RPC82</i> | 74 |
| | RPC4 (RPC53) | <i>POLR3D</i> | 44.4 | RPC4 (C53) | <i>RPC53</i> | 46.7 |
| | RPC5 (C37) | <i>POLR3E</i> | 79.9 | RPC5 (C37) | <i>RPC37</i> | 32.1 |
| | RPC6 (RPC39) | <i>POLR3F</i> | 35.7 | RPC6 (C34) | <i>RPC34</i> | 36.1 |
| | RPC7 (RPC32) | <i>POLR3G</i> | 25.9 | RPC7 (C31) | <i>RPC31</i> | 27.7 |
| | RPC8 | <i>POLR3H</i> | 22.9 | RPC8 (C25) | <i>RPC25</i> | 24.3 |
| | RPC9 (CGRP-RC) | <i>CRCP</i> | 16.9 | RPC9 (C17) | <i>RPC17</i> | 18.6 |
| | RPC10 (RPC11) | <i>POLR3K</i> | 12.3 | RPC10 (C11) | <i>RPC11</i> | 12.5 |
| | RPAC1 (RPA5) | <i>POLR1C</i> | 39.3 | RPAC1 (AC40) | <i>RPC40</i> | 37.7 |
| | RPAC2 (RPA9) | <i>POLR1D</i> | 15.2 | RPAC2 (AC19) | <i>RPC19</i> | 16.1 |
| | RPABC1 (RPB5) | <i>POLR2E</i> | 24.6 | RPABC1 (ABC27) | <i>RPB5</i> | 25.1 |
| | RPABC2 (RPB6) | <i>POLR2F</i> | 14.5 | RPABC2 (ABC23) | <i>RPO26</i> | 17.9 |
| | RPABC3 (RPB8) | <i>POLR2H</i> | 17.1 | RPABC3 (ABC14.5) | <i>RPB8</i> | 16.5 |
| | RPABC4 (RPB7) | <i>POLR2K</i> | 7 | RPABC4 (ABC10a) | <i>RPC10</i> | 7.7 |
| | RPABC5 (RPB10) | <i>POLR2L</i> | 7.6 | RPABC5 (ABC10b) | <i>RPB10</i> | 8.3 |
| TFIIIA: GTF for type 1 promoter, binds to C box | | | | | | |
| | TFIIIA | <i>GTF3A</i> | 41.5 | TFIIIA (Tfc2) | <i>PZF1</i> | 50 |
| TFIIIC: GTF for type 1 and 2 promoters, binds to TFIIIA (on type 1 promoter), Box A and Box B | | | | | | |
| tA module | TFIIIC102 | <i>GTF3C3</i> | 101.3 | Tfc4 (t131) | <i>TFC4</i> | 120.2 |
| | TFIIIC63 | <i>GTF3C5</i> | 59.6 | Tfc1 (t95) | <i>TFC1</i> | 73.5 |
| | TFIIIC35 | <i>GTF3C6</i> | 24.1 | Tfc7 (t55) | <i>TFC7</i> | 49.2 |
| tB module | TFIIIC220 | <i>GTF3C1</i> | 238.9 | Tfc3 (t138) | <i>TFC3</i> | 132.1 |
| | TFIIIC110 | <i>GTF3C2</i> | 100.7 | Tfc6 (t91) | <i>TFC6</i> | 74.7 |
| | TFIIIC90 | <i>GTF3C4</i> | 92 | Tfc8 (t60) | <i>TFC8</i> | 67.7 |
| TFIIB: GTF for all types of promoter, binds to TATA-like sequence, recruit RNAPIII, facilitates re-initiation | | | | | | |
| B' fraction | TBP | <i>TBP</i> | 37.7 | TBP | <i>SPT15</i> | 27 |
| | Brf1 (TFIIB90) | <i>BRF1</i> | 73.8 | Brf1 | <i>BRF1</i> | 66.9 |
| | or Brf2 (TFIIB50)* | <i>BRF2</i> | 46.5 | | | |
| B'' fraction | Bdp1 | <i>BDP1</i> | 293.9 | Bdp1 (Tfc5) | <i>BDP1</i> | 67.7 |
| SNAPc: snRNA-activating protein complex in higher eukaryotes, GTF for type 3 promoter, binds to PSE | | | | | | |
| | SNAPC43 | <i>SNAPC1</i> | 43 | | | |
| | SNAPC45 | <i>SNAPC2</i> | 35.6 | | | |
| | SNAPC50 | <i>SNAPC3</i> | 46.8 | | | |
| | SNAPC190 | <i>SNAPC4</i> | 159.4 | | | |
| | SNAPC19 | <i>SNAPC5</i> | 11.3 | | | |
| Activators: stimulate transcription from external promoters, binds to DSEs | | | | | | |
| | Staf (SBF) | <i>ZNF143</i> | 68.9 | | | |
| | Oct1 | <i>POU2F1</i> | 76.5 | | | |

*Brf2, TFIIB-related factor 2, replacing Brf1 for transcription at type 3 promoter

Abbreviations: GTF, general transcription factor; PSE, proximal sequence element; DSE, distal sequence element.

Alternative names of *Saccharomyces cerevisiae* and *Homo sapiens* specific subunits are given in brackets.

Adapted from Hoffmann et al., 2016, Nikitina et al., 2011.

TFIIIC

Transcription Factor III C (TFIIIC) is a large protein complex composed of 6 subunits, which are organized in two DNA-binding subcomplexes called τ A and τ B. In yeast τ A is assembled from three subunits, namely τ 131 (Tfc4), τ 95 (Tfc1) and τ 55 (Tfc7), while τ B is composed of τ 138 (Tfc3), τ 91 (Tfc6) and τ 60 (Tfc8), they together form a complex with a molecular mass of about 500 kDa. TFIIIC is required for the transcription of class III genes that are controlled by type 1 and 2 promoters (see section 0). It recognizes the highly conserved promoter elements, thus allowing the recruitment of TFIIIB upstream of the transcription start site, which subsequently leads to the recruitment of RNAPIII and formation of the preinitiation complex (Male et al., 2015). It has also been shown in yeast that TFIIIC binds some chromosomal locations called ETC loci (for extra-TFIIIC) without the rest of the RNAPIII transcription apparatus, which appears to be able to mediate some additional functions, for example, to create nucleosome-free genomic landmarks, and as boundary elements in separating chromatin domains (reviewed in Donze, 2011).

SNAPc

Small nuclear RNA (snRNA)-activating protein complex (SNAPc), also called proximal sequence element (PSE)-binding transcription factor (PTF), is a sequence-specific DNA binding complex composed of 5 subunits: SNAP190, SNAP50, SNAP45, SNAP43 and SNAP19 (Mittal et al., 1999). SNAPc is required for transcription of RNAPII-dependent snRNA genes and RNAPIII-dependent genes with type 3 promoters (see section 0) in metazoans. SNAPc is involved in a number of well-defined functions, for instance, specific binding to the PSE, nucleation of the assembly of RNAPII and RNAPIII transcription initiation complexes, cooperative binding with TBP and with its corresponding activators, etc.

3.3.2 Promoters of class III genes

Based on the organization of transcriptional control elements and transcription factor dependence, the promoters of known RNAPIII-transcribed genes are divided into three categories, as shown in Figure 3-3. **Type 1 promoters** contain the A box and the internal control region (ICR) recognized by transcription factors TFIIIC and TFIIIA, respectively. **Type 2 promoters** are composed of an A box and a B box bound by TFIIIC. The distance between the A box and the B box is highly variable because TFIIIC exhibits a naturally elastic structure so that it does not require a specific spacing between its recognition sites (Nagarajavel et al., 2013). Sometimes an upstream TATA box is also present in type 2 promoters, which can help TFIIIB recruitment and influence transcription start site selection. **Type 3 promoters**, present

in metazoans, lack control elements within the transcribed region and, instead, harbor an upstream proximal sequence element (PSE) which is bound by SNAPc. The majority of SNAPc-dependent promoters include an external enhancer region, the distal sequence element (DSE), occupied by activators Staf and/or Oct1. Moreover, in type 3 RNAPIII promoters, there is also a TATA box bound by TFIIB, which determines their RNAPIII specificity (Figure 3-3).

5S rRNA genes are the only set of class III genes that rely on type 1 promoters. All tRNA genes (with the exception of the selenocysteine tRNA gene) employ a type 2 promoter and the two conserved A box and B box correspond to the universally conserved D-loop and T-loop of mature tRNAs. In the structure of mature tRNAs, the A box normally starts from the end of the acceptor stem and extends to the D-loop (i.e. N8 to N19) and the B-box is positioned from N52 to N62 (Figure 3-2A). The synthesis of U6 and other RNAPIII-dependent transcripts from metazoans like the 7SK and Y RNAs is dependent on type 3 upstream promoters (Figure 3-3).

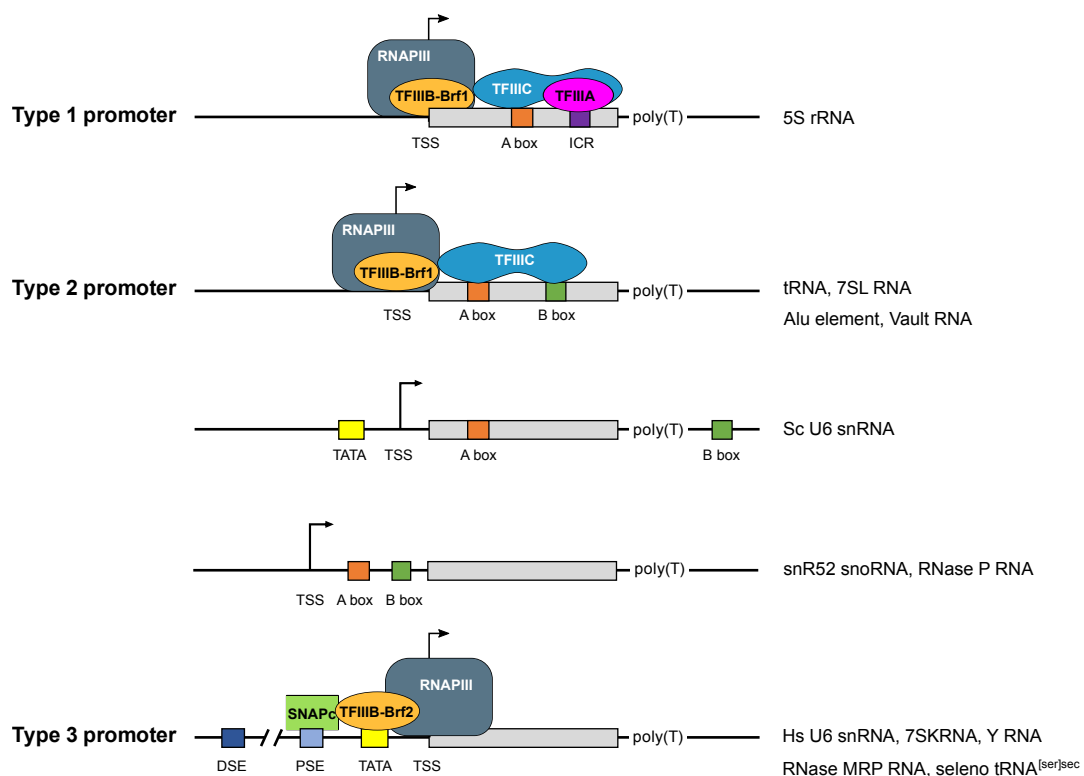


Figure 3-3: Promoter architecture of class III genes.

Three types of class III gene promoters are shown. Detail described in the text. RNA species relying on distinct promoter are indicated on the right site. Primary transcripts originating from type 1 and type 3 promoters do not undergo removal of a 5'-leader and, thus, the 5'-terminal nucleotide of the mature RNA coincides with the transcription start site, as indicated by the bent arrows. In contrast, primary transcripts originated from type 2 promoters generally carry a 5'-leader portion that must be processed. Abbreviations: TSS, transcription start site; Sc, *S. cerevisiae*; Hs, *Homo sapiens*. Adapted from Dieci et al., 2013.

3.3.3 Recruitment of TFIIC to the promoter

The first step for transcription of tRNA genes is the binding of TFIIC to the internal promoter sequences, the A box and B box, which does not require any additional transcription factor (Lassar et al., 1983; Ruet et al., 1984). For 5S rRNA genes, which lack the B box, the recruitment of TFIIC requires the preceding binding of TFIIA to the ICR (Orioli et al., 2012). Photochemical crosslinking and antibody-based interference experiments revealed that τ 95 is responsible for A box binding and τ 138 for B box binding, with both interactions being essential for RNAPIII transcription (Gabrielsen et al., 1989; Bartholomew et al., 1990).

It was previously shown that τ B binding to B box was predominant over the τ A binding to the A box (Stillman and Geiduschek, 1984; Schultz et al., 1989). The low affinity of τ A for the A box was supported by a recent structural study (Vorländer et al., 2020) in which the authors observed that τ 95 binding to the A box is auto-inhibited by its C-terminal tail (Figure 3-4). The negatively charged acidic tail transiently associates with the positively charged DNA binding domain (DBD) of τ 95, thereby competing with DNA and reducing the affinity to promoter. The auto-inhibition by the acidic tail of τ 95 is thought to increase the specificity of the interaction with the A box by outcompeting suboptimal DNA sequences (Vorländer et al., 2020).

tRNA genes share a similar organization but they differ in length either because of the variable arm of the tRNA or because of the presence of an intron. Thus, the distance between the A and the B box can also be very different. However, the τ A and τ B modules of TFIIC can bind the A and B boxes regardless of the distance between them, which was believed to largely owe to the remarkable structural elasticity of TFIIC (Schultz et al., 1989; Nagarajavel et al., 2013). But how TFIIC achieves this elasticity was not clear.

Based on earlier low-resolution scanning transmission electron microscopy (STEM) analyses, TFIIC was observed as a “dumb-bell” shaped molecule with a flexible linker connecting two subdomains. Thus, this linker was proposed to underlie the structural elasticity of TFIIC (Marzouki et al., 1986, Schultz et al., 1989). Chemical cross-linking coupled to mass spectrometry (XL-MS) and crystal structure analyses further mapped the region of this linker (Male et al., 2015). The authors identified that τ 131 contains tetra-trico peptide repeat (TPR) arrays at both its N-terminal and C-terminal halves (Figure 3-4). Importantly, they found that the N-terminal TPR arrays of τ 131 establish contacts with an unstructured, central region of τ 138 (in the τ B module) termed τ -Interacting Region (τ IR). It was therefore suggested that τ IR establishes the main link between τ A and τ B, as its disordered nature may provide flexibility to recognize variously spaced A and B boxes (Male et al., 2015).

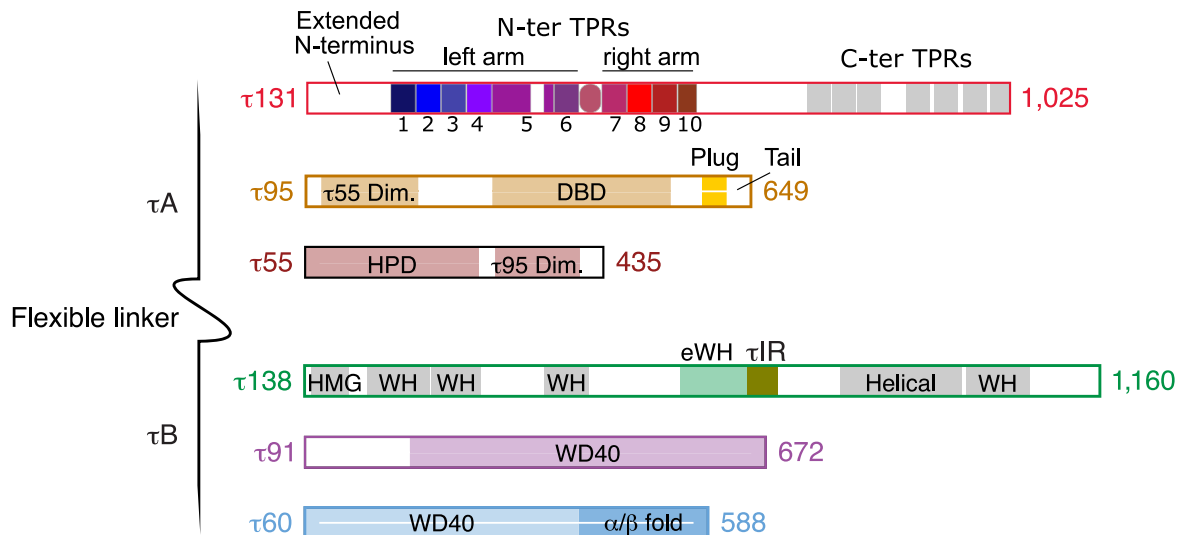


Figure 3-4: Architecture of TFIIC subunits.

Schematic representation of the six subunits of *S. cerevisiae* TFIIC. The amino acid length of each subunit is indicated at the C terminus. Domains available in the crystal structure obtained by Male et al., 2015 are colored. Abbreviations: $\tau55$ and $\tau95$ Dim, $\tau55$ and $\tau95$ dimerization domains; DBD, DNA-binding domain; HPD, histidine phosphatase domain; HMG, high mobility group DNA-binding domain; TPR, tetra-trico peptide repeats; eWH, extended winged helix. For more details about TPR, WH and WD40 see Box 3-1.

3.3.4 Recruitment of TFIIB by TFIIC

At type 1 and type 2 promoters, TFIIB is recruited by TFIIC to a region upstream of the transcription unit (Kassavetis et al., 1990). Previous studies suggested that $\tau131$ binds to TFIIB subunits Brf1 and Bdp1 in a stepwise manner to help TFIIB assembly, using overlapping sites on its N-terminal TPR arrays (Dumay-Odelot et al., 2002; Liao et al., 2003; Moir et al., 1997), and that conformational changes occur within $\tau131$ upon binding of Brf1 and Bdp1 (Moir et al., 2000; Kassavetis et al., 1992b). However, biochemical and structural data obtained by Male et al. (2015) showed that Brf1 and Bdp1 binds to distinct sites on $\tau131$, but the Bdp1- $\tau131$ interaction sites overlap with the $\tau138$ - $\tau131$ interaction region. Thus, it was proposed that the interaction of $\tau131$ with Bdp1 could cause a conformational change in TFIIC leading to the displacement of the τB module, which would be a regulatory mechanism essential for the initial round of RNAPIII transcription (Male et al., 2015). These results support the notion that $\tau131$ TPR arrays play a major role in linking τA , τB and TFIIB during PIC assembly.

A couple of evidence also suggested that the binding of τA to the A box mediates TFIIB assembly, and is important for transcription activation and TSS selection (Baker et al., 1987; Gerlach et al., 1995; Joazeiro et al., 1996). However, it is not clear how TSS selection is achieved. It has been shown that TFIIB is placed preferentially ~ 30 bp upstream of the TSS,

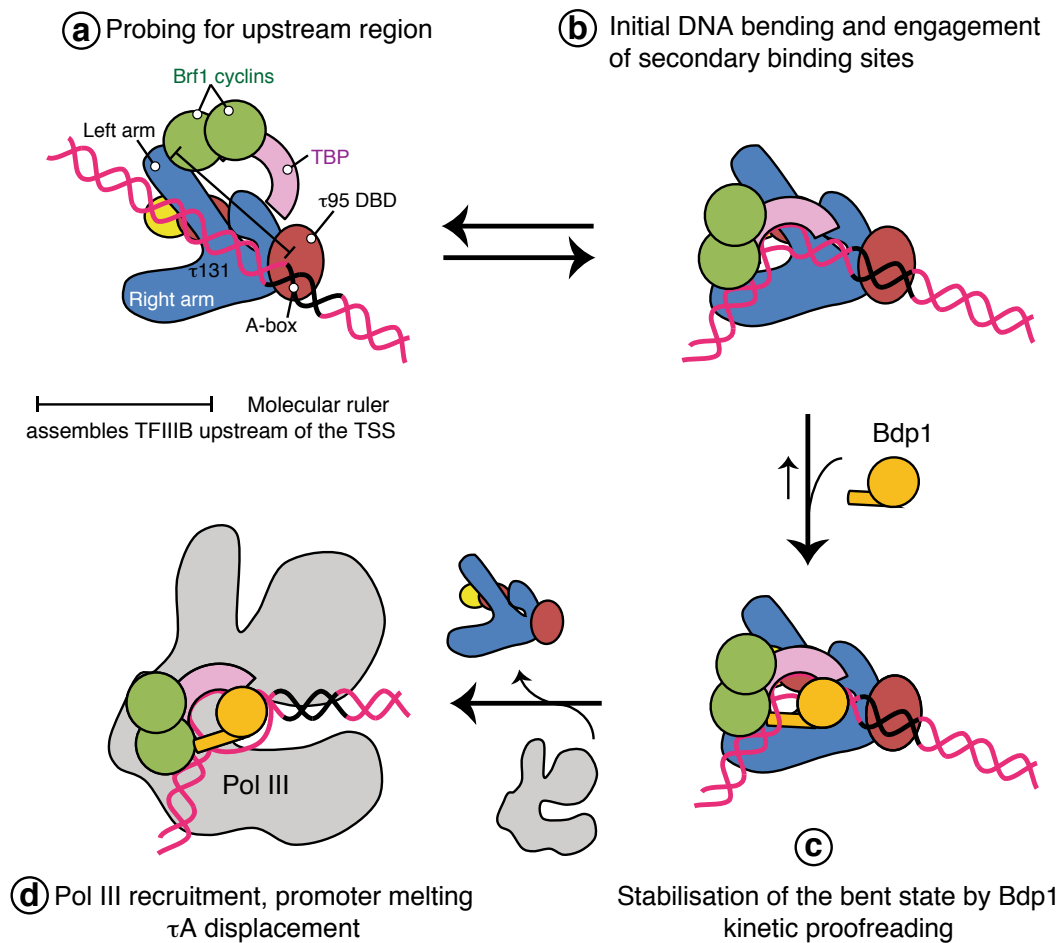


Figure 3-5: Model of TFIIB assembly by τ A/TFIIIC.

(a) Initially, Brf1 and TBP are recruited to the N-terminal TPR array of τ 131. The distance between the τ 95 DBD and the τ 131 TPR array serves as a molecular ruler that places TFIIB within a certain distance upstream of the TSS. (b) TBP binds and bends the upstream DNA sequence. The lifetime of this complex depends on the upstream DNA sequence engaged. (c) Bdp1 then enters the complex and stabilizes the bent state. (d) Recruitment of RNAPIII and promoter opening displaces τ A, freeing the transcribed region. From Vorländer et al., 2020.

and its positioning is determined by TFIIC and by the direct interaction of TBP with the upstream DNA sequence (Kassavetis et al., 1989; Joazeiro et al., 1996). A recent study has redefined the model for TFIIB sequential assembly (Vorländer et al., 2020). Initially Brf1 is bound to the N-terminal region of τ 131. TBP is then recruited via interactions with the C-terminus of Brf1 and with the τ B subunit τ 60 (Mylona et al., 2006; Deprez et al., 1999). TBP subsequently binds and bends the upstream DNA sequence, and is further stabilized through the incorporation of Bdp1. This process would allow TFIIB to assemble on a suitable DNA sequence by using a proofreading mechanism, in which the lifetime of the initial TBP-DNA complex helps selecting the correct sequence around which TFIIB assembles, and therefore the correct TSS. Importantly, the position of the DBD of τ 95 and the Brf1- τ 131 TPR array is

such that the distance between them might serve as a molecular ruler to place TFIIB at a relatively constant position upstream of the TSS (Vorländer et al., 2020; Figure 3-5). Previous cryo-EM structure revealed that Brf1, TBP and Bdp1 form a positively charged ring around the TATA box, which may explain why the binding of TFIIB to DNA is so unusually stable and why they can serve as a “roadblock” for the transcription and replication machineries (Vorländer et al., 2018).

3.3.5 Recruitment of RNAPIII and assembly of the PIC

Once bound to the promoter region, TFIIC and TFIIB recruit RNAPIII to form the pre-initiation complex (PIC). Assembly of RNAPIII PIC at tRNA genes in yeast involves numerous interactions between the polymerase and the GTFs. For instance, yeast two hybrid experiments showed that C34 interacts with Brf1 (Werner et al., 1993), this interaction was subsequently shown to be essential for RNAPIII recruitment and open complex formation based on a mutagenic analysis of C34 (Brun et al., 1997). Photo-cross-linking studies found that C34 is located immediately downstream from Brf1 in the assembled RNAPIII transcription complex, supporting its role in TFIIB recognition (Bartholomew et al., 1993). Additional cross-links between Brf1 and the RNAPIII subunits C160 and C128 were also detected (Khoo et al., 2014). Furthermore, the stalk subunit C17 was found to interact with the N-terminal cyclin repeats of Brf1 (Figure 3-6), suggesting a role for C17 in the recruitment of RNAPIII (Ferri et al., 2000) and Bdp1 was found to interact with C37 (Wu et al., 2011). Besides TFIIB interactions, C53 and ABC10 α were also found to contact the TFIIC subunit τ 131 (Wu et al., 2011; Dumay et al., 1999). A role for C31 in the preinitiation complex recognition was also characterized since a small deletion of the C-terminal end of C31 impaired RNA chain initiation (Thuillier et al., 1995).

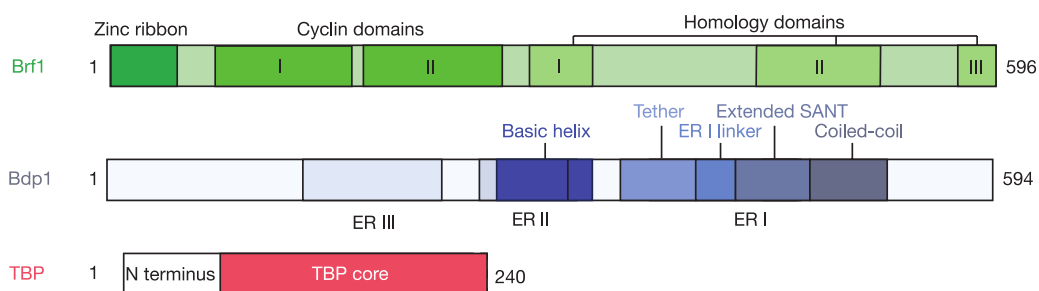


Figure 3-6: Architecture of TFIIB subunits.

Domain architecture of TFIIB involved in PIC assembly. Protein regions are depicted according to their presence (solid color boxes) or absence (empty boxes) in the open complex of the pre-initiation complex (OC-PIC) structure built by Abascal-Palacios et al., 2018. Abbreviations: ER, essential region. For SANT domain see **Box 3-1**.

Several studies revealed a structurally and functionally conserved core transcription initiation complex in all eukaryotic RNAPs which contains promoter DNA, RNAP, TBP, a TFIIB-like factor, a TFIIIF-like factor and a TFIIE-related factor. In the RNAPIII transcription system, the C53-C37 heterodimer is considered to be homologous to TFIIIF, while the C82-C34-C31 heterotrimer is regarded as the homolog of TFIIE. Among the three subunits of TFIIB, TBP is common to the three eukaryotic RNAPs; the N-terminal half of Brf1 shares a high degree of sequence similarity with TFIIB; but Bdp1 has no counterpart in RNAPI and RNAPII transcription machineries (Vannini and Cramer, 2012). It was also suggested that TFIIC acts as an assembly factor, but they are not *bona fide* components of the PIC. Recruitment of RNAPIII and promoter opening may displace τ A from the A box to free the transcribing unit, as the τ A module was found to elute separately from the TFIIB-RNAPIII-DNA complex during size-exclusion chromatography (Vorländer et al., 2020). This idea is in agreement with previous *in vitro* data showing that TFIIC is only required for TFIIB assembly but is dispensable for RNAPIII transcription (Bardeleben et al., 1994). However, other *in vitro* studies showed that TFIIC is not released from the DNA template once it is bound to it (Ruet et al., 1984), and, moreover, TFIIC is required to support re-initiation at genes as long as 300 bp (Ferrari et al., 2004). Thus, it still remains unclear whether TFIIC is indeed disassembled during transcription initiation *in vivo*.

3.3.6 Promoter opening

Upon RNAPIII PIC assembly on promoter, the transcription machinery forms a closed complex (CC) where the duplex DNA that is located at the DNA-binding cleft of the polymerase remains double stranded. The promoter DNA is then melted into single strands and the template strand is engaged by the polymerase active site, leading to the formation of an open complex (OC). After synthesis of a short piece of RNA transcript, the initially transcribing complex (ITC) forms and the polymerase successively escapes the promoter and transits to the phase of transcription elongation (Ramsay and Vannini, 2018).

An early study showed that RNAPIII does not open its promoter uniformly, with the upstream segment (bp -9 to -5, relative to the TSS as +1) opening at a lower temperature than the downstream segment (bp -3 to +7), suggesting that promoter opening by RNAPIII may nucleate at the upstream end of the transcription bubble (Kassavetis et al., 1992b). TFIIB was found to participate in promoter opening in addition to its role in RNAPIII recruitment (Kassavetis et al., 1998). Further analyses proposed that TFIIB participates in two steps of promoter opening: the N-terminal domain of Bdp1 is involved in initiating duplex DNA separation at the upstream edge of the transcription bubble, and then, the zinc-ribbon domain in the N-terminus of Brf1 functions by extending the initial transcription bubble

towards and beyond the TSS (Figure 3-6; Kassavetis et al., 2001). It was suggested that both Bdp1 and Brf1 drive promoter opening by inducing conformational changes in RNAPIII, particularly by altering the arrangement of the C82-C34-C31 heterotrimer, which has been long implicated in promoter opening (Ramsay and Vannini, 2018).

Atomic models of yeast RNAPIII PIC in different functional states have been built simultaneously by two laboratories, providing structural insights into the mechanisms of promoter opening (Abascal-Palacios et al., 2018; Vorländer et al., 2018). The cryo-EM structures of open RNAPIII PIC (OC-PIC) revealed an intricate interaction network between TFIIIB and several RNAPIII subunits. The architecture of the subunits involved in promoter opening are shown in Figure 3-7. TFIIIB was observed to completely enclose the DNA around the TATA box. The overall topology of the PICs is highly similar in RNAPIII and RNAPII. As mentioned before, the C82-C34-C31 heterotrimer is regarded as a TFIIIE-like subcomplex. Indeed, the unwound non-template DNA strand is stabilized by the C82 cleft loop (Figure 3-7) resembling the mechanism of action of TFIIIE. In addition, C31 physically bridges the stalk and the clamp, which is functionally conserved in TFIIIE. However, based on the structures obtained, the upstream region of the transcription bubble is stabilized by the WH domains of C34 (Figure 3-7) as by the TFIIIF subunit Tfg2 in the case of the RNAPII PIC. Thus, the heterotrimer combines the roles of both TFIIIF and TFIIIE.

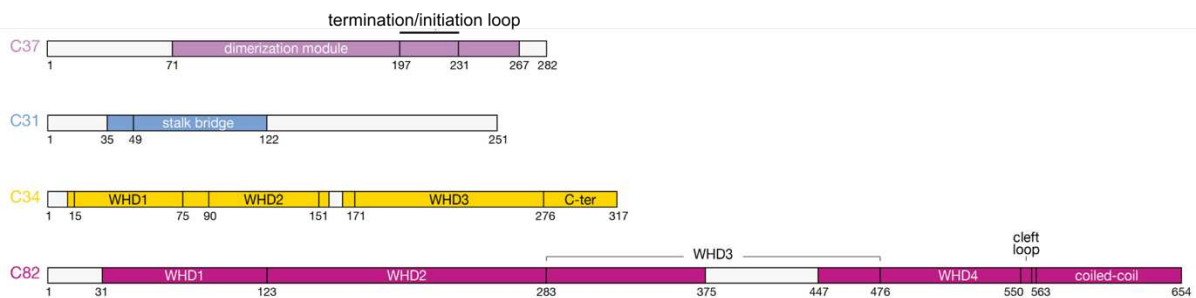


Figure 3-7: Domain architecture of RNAPIII key subunits involved in PIC assembly.

Protein regions are depicted according to their presence (solid color boxes) or absence (empty boxes) in the open complex of the pre-initiation complex (OC-PIC) structure built by Abascal-Palacios et al., 2018. Abbreviations: WHD, winged helix domain. More details see in the text.

Through the comparison between the RNAPIII CC and OC structures, Vorländer et al. proposed a model for the structural rearrangements occurring during promoter opening (Figure 3-8). In the early closed DNA complex, the C34 WH domains and the Bdp1 tether are disordered, the upstream DNA is kinked away by the clamp head and C82 cleft loops, leading to a 30-degree bend introduced around position -15. The transition from the closed to the open promoter complex proceeds via an open-clamp intermediate. The open and closed

clamp conformations of unbound RNAPIII has also been reported before in yeast, where it was suggested that a moving stalk comprising subunits C25 and C17 can mediate this conformational change (Hoffmann et al., 2015). The open-clamp conformation, where the clamp and the C82 cleft loop move upwards and lie on top of the closed DNA, allows a better association of the polymerase with the target DNA. Subsequently, C34 WH domains become ordered followed by clamp closing. The transition to the closed clamp position enforces DNA melting by a steric clash of the C82 cleft loop and the DNA duplex, and then the template strand is loaded into the active site (Figure 3-8). This model is similar to that of RNAPII, since both employ a movement of the clamp, an extended loop (C82 or TFIIIE), and WH domains (heterotrimer/TFIIIE-TFIIF) for promoter opening and transcription bubble stabilization (Vorländer et al., 2018). But unlike RNAPII, for which DNA opening generally requires ATP hydrolysis, in the case of RNAPIII (as well as RNAPI) the DNA is open only with the aid of the binding energy generated by interactions newly established during the assembly of PIC (Cramer, 2019).

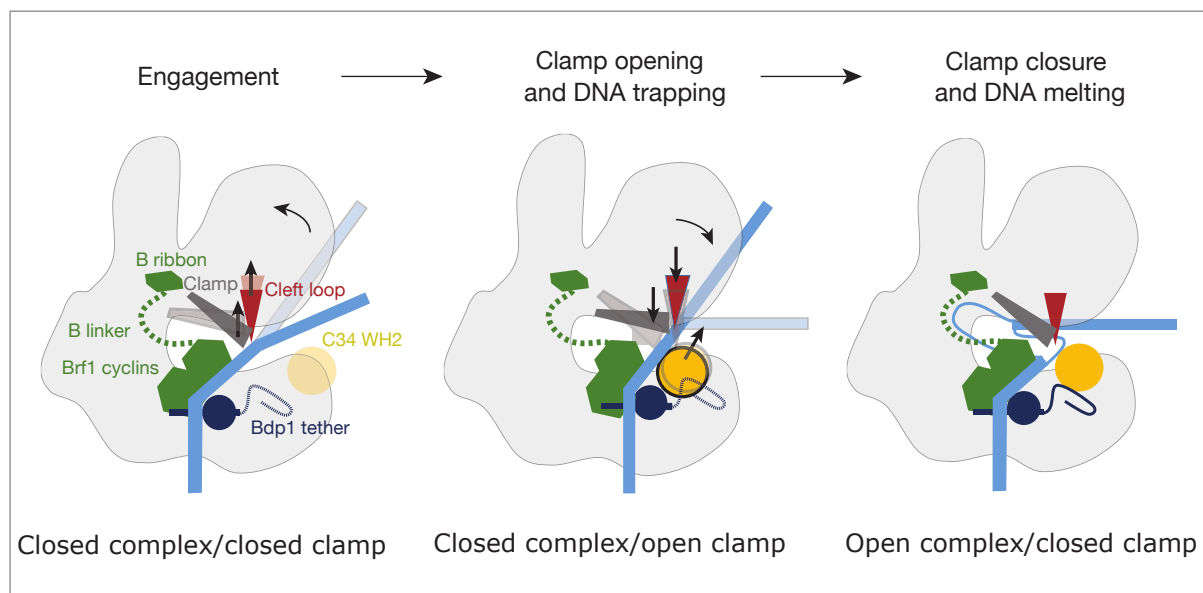


Figure 3-8: Schematic of the mechanism of promoter opening and DNA melting by RNAPIII.

Initially, the closed DNA is bent away by RNAPIII subunit C82 cleft loop and the clamp, while the TFIIIB subunit Bdp1 tether and WH1 (winged-helix 1) and WH2 (winged-helix 2) domains of RNAPIII subunit C34 are disordered. Opening of the clamp enables the closed DNA to slide between clamp and lobe. Subsequently, C34 WH1 and WH2 become ordered and enclosed double-stranded DNA. Closing of the clamp leads to the downwards movement of the cleft loop and DNA melting, while establishing stabilizing interactions between the Bdp1 tether and C34 WH1 and WH2. The non-template strand is stabilized by C34 and C38; the template strand is presumably stabilized by the Brf1 linker. Adapted from Vorländer et al., 2018.

Box 3-1: The TPR, WH, WD40 and SANT protein domains

Tetra-trico peptide repeat (TPR) is a degenerate 34 amino acid sequence arranged into two antiparallel α -helices, present in tandem arrays of 3-16 motifs, which can serve as a platform for protein-protein interactions and for the assembly of multiprotein complexes (Das et al., 1998). (Presented in TFIIIC subunit τ 131)

Winged helix (WH) DNA-binding domain is a structural motif belonging to the helix-turn-helix (HTH) family. The classical WH fold consists of two wings (W1 and W2), three α -helices (H1, H2 and H3) and three β -strands (S1, S2 and S3) arranged in the canonical order H1-S1-H2-H3-S2-W1-S3-W2. It is found in core components of transcription systems in eukaryotes and prokaryotes, participating in the establishment of protein-DNA and protein-protein-interactions (Gajiwala and Burley, 2000; Teichmann et al., 2012). "Extended" winged helix (eWH) is a WH domain extended by specific α -helices at the N- and C- termini (Meinhart et al., 2003). (Present in TFIIIC subunit τ 138, C34 and C82)

WD40 was named by the conserved WD dipeptide and the length of approximately 40 amino acid residues in a single repeat. Each WD40 repeat comprises a four-stranded antiparallel β -sheet. The repeats in a protein fold into a β -propeller architecture, often comprising seven blades. Proteins containing WD40 domain are very abundant in eukaryotic organisms, and are rarely present in prokaryotes. This domain is among the top ten most abundant domains in the eukaryotic genomes (Stirnemann et al., 2010). WD40 domain proteins are involved in a large variety of cellular processes, in which WD40 domains function as a protein-protein or protein-DNA interaction platform. No enzymatic activity has been detected so far for this domain (Xu and Min, 2011). (Present in TFIIIC subunit τ 91 and τ 60)

SANT domain, for "switching-defective protein 3 (Swi3), adaptor 2 (Ada2), nuclear receptor co-repressor (N-CoR), transcription factor (TF)IIIB", is a protein domain found in many chromatin-remodelling proteins that function as a unique histone-interaction module coupling histone binding to enzyme catalysis. It has high sequence similarity with the DNA binding domain of Myb-related proteins (Boyer et al., 2004). (Present in TFIIIB subunit Bdp1)

3.4 Transcription elongation

Transcription elongation is the process following initiation by which an RNA chain is processively synthesized as the polymerase moves along the template DNA. This transcription step is the less well characterized for RNAPIII. Genome-wide mapping of RNAPIII transcription in yeast by the CRAC method (UV crosslinking and analysis of the cDNA, details see **Box 3-2**) revealed unequal distribution of RNAPIII along tRNA genes (Turowski et al., 2016). According to this study, RNAPIII was enriched in two regions close to the 5' and the 3' ends of tRNA genes, respectively, with the 5' end peak being much higher than the 3' end peak. These peaks could represent regions with decreased elongation rate and/or RNAPIII pausing. Interestingly, the 5' and 3' peaks coincided with the beginning of the A and B boxes bound by TFIIIC, respectively. Thus, Turowski et al. proposed that TFIIIC binding to the A and B boxes forms a physical barrier that interferes with RNAPIII elongation, therefore supporting the idea that TFIIIC remains associated with the DNA during transcription elongation. RNAPIII would

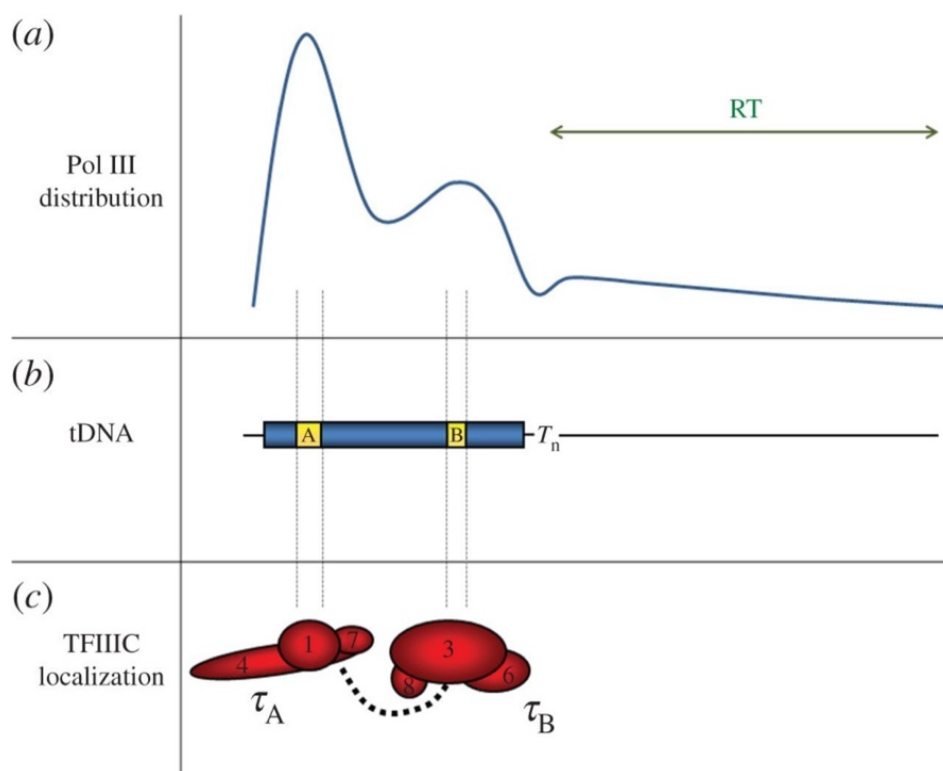


Figure 3-9: Uneven distribution of RNAPIII on transcription units.

(a) RNAPIII distribution detected by the CRAC method, across most genes. The profile presents a high peak of nascent transcript density over the 5' end of the transcription unit and a weaker peak before the 3' end of mature tRNA (intron-less tRNA gene is shown). Read-through (RT) of termination signal is observed on many tRNAs, typically extending 50-200 nt beyond the expected canonical termination site. (b) Localization of A- and B-boxes of the bipartite internal promoter, and termination site (T_n) in a tRNA-encoding gene (tDNA). (c) The τ_A and τ_B modules of TFIIIC factor binding the A and B boxes. Regions of postulated transient pausing of RNAPIII correspond to the TFIIIC binding sites. From Lesniewska and Boguta, 2017.

displace the τ A module from the A box and then τ B module from the B box sequentially, and thus, one TFIIC module would always remain in contact with the transcribed gene. With regards to the prominent 5' peak, they proposed that it could be due to slow clearance from the initiation site, perhaps a delay in the dissociation of the polymerase from the transcription factors (Turowski et al., 2016, Turowski and Tollervey, 2016), which, however, needs to be further investigated.

Box 3-2: UV crosslinking and analysis of the cDNA (CRAC)

The UV crosslinking and analysis of cDNA (CRAC) method was originally developed by the Tollervey group and used for the identification of binding sites of RNA-interacting proteins (Granneman et al., 2009; Bohnsack et al., 2012). Briefly, proteins of interest containing a bipartite tag are crosslinked with their target RNA and isolated under highly denaturing conditions, ensuring that only direct interactions are detected. Then, RNA fragments are recovered and deep-sequenced after linker ligation and cDNA synthesis. This method allows a genome-wide analysis of the interactome of RNA-binding proteins with nucleotide resolution.

3.5 Transcription termination

The final step of a transcription cycle is the process of transcription termination, which allows the release of the transcript and the recycling of the polymerase for the next round of transcription. Proper termination is critical for maintaining genome stability by avoiding the interference between neighboring transcription units as well as the conflicts between transcribing RNAPs and other DNA-associated machineries such as the replisome (reviewed in Porrua and Libri, 2015b). The three eukaryotic RNAPs employ different strategies to terminate transcription and among them, RNAPIII seems to adopt the simplest mechanism.

3.5.1 T-tract-dependent termination

Unlike RNAPI and RNAPII, which require both *cis*-acting sequences and *trans*-acting factors for transcription termination, the most widely accepted model posits that RNAPIII terminates autonomously and efficiently at a stretch of thymidines (T-tract) of variable length on the non-template strand, without the need for ancillary factors (Geiduschek and Tocchini-Valentini, 1988; Arimbasseri et al., 2013; Porrua et al., 2016). Early studies revealed that a cluster of four or more consecutive T residues on the non-coding strand could terminate the synthesis of *Xenopus* 5S RNA and tRNA^{Lys} *in vitro*, in the absence of additional factors (Bogenhagen and Brown, 1981; Mazabraud et al., 1987; Cozzarelli et al., 1983). Introducing a T-tract within the *Xenopus* tRNA^{Tyr} coding region caused premature termination (Koski et al., 1980), while deletions of natural oligo(dT) signals resulted in transcription readthrough in yeast and human tRNA genes (Allison and Hall, 1985; Adeniyi-Jones et al., 1984). Further analyses showed that the minimum length of T-tract required for RNAPIII termination is species-specific. For *Xenopus* and mammals, as few as 4 Ts can support efficient termination (Bogenhagen and Brown, 1981; Cozzarelli et al., 1983; Hamada et al., 2000) and, indeed, most tRNA genes harbour a stretch of 4 Ts (i.e. a T4 terminator) in these organisms (Allison and Hall, 1985; Braglia et al., 2005). The fission yeast *Schizosaccharomyces pombe* requires at least 5 Ts for efficient termination (Hamada et al., 2000), while in *S. cerevisiae*, 5 Ts support only very moderate levels of termination and at least 6 Ts are necessary for relatively efficient termination (Allison and Hall, 1985; Arimbasseri and Maraia, 2015; Mishra and Maraia, 2019). These ideas are in agreement with the fact that T5/T6 and T6/T7 are the most frequent terminators in *S. pombe* and *S. cerevisiae*, respectively (Braglia et al., 2005). In the case of human RNAPIII, *in vitro* and *in vivo* data have demonstrated that termination can occur at non-canonical terminators consisting in interrupted long T-tracts, generally composed of several portions of 1, 2 or 3 consecutive Ts separated by another nucleotide (Orioli et al., 2011).

Effect of the T-tract sequence context on termination

Several pieces of evidence indicate that the nucleotide sequence surrounding the T-tract could also influence the efficiency of transcription termination, especially for T-tracts of short length. For the *Xenopus* 5S RNA, efficient transcription termination was observed whenever GC-rich sequences surrounded the T4 terminator. In contrast, the presence of two or more consecutive A nucleotides within the three nucleotides preceding or following the T4 terminator significantly reduced the efficiency of termination (Bogenhagen and Brown, 1981). Similarly, human RNAPIII normally readthrough *in vitro* a *B1-Alu* gene that contains a T4 terminator flanked by AA, but replacing AA with GC increased the termination efficiency dramatically. Moreover, 4 Ts flanked by GC can be as efficient as a T5 terminator (Goodier and Maraia, 1998). In the case of *S. cerevisiae*, for a model tRNA gene, a CT dinucleotide placed immediately downstream of a T5 can significantly weaken the T5 termination potential. In contrast, termination was highly efficient when T5 was followed by an A or G residue (Braglia et al., 2005). Although an impact of the sequence context on termination at short T-tracts has been observed both in yeast and in metazoans, there is so far no universal rule with regard to the effect of the flanking sequence.

Molecular mechanisms of termination at T-tracts

It has been shown that the 8-9 nt RNA:DNA hybrid in the polymerase catalytic center is a major determinant of the stability of the elongation complex (Korzheva et al., 1998; Sidorenkov et al., 1998; Kireeva et al., 2000). Stability analyses have revealed that oligo (rU:dA) sequences are exceptionally unstable, as a DNA:RNA duplex containing a (rU:dA)₅ is at least 200 times less stable at room temperature than the corresponding duplex containing an (rA:dT)₅ (Martin and Tinoco, 1980), providing a strong rationale for the biological selection of a short run of U residues at transcription termination sites.

An early biochemical study showed that a *S. cerevisiae* RNAPIII variant lacking the three subunits C53-C37-C11, dubbed RNAPIII Δ (core enzyme) failed to terminate transcription efficiently. C53-C37 forms a stable heterodimer related to TFIIIF α/β (Vannini and Cramer, 2012). The efficient association of C53-C37 with RNAPIII is dependent on C11, a 11 kDa protein containing two zinc ribbon domains, with its NTD similar to Rpb9 and CTD similar to TFIIIS exhibiting RNA 3' cleavage activity (Chédin et al., 1998). Adding back recombinant C53-C37 to RNAPIII Δ restored the recognition of the termination signal, while adding C11 only restored the RNA cleavage activity of RNAPIII but not termination (Landrieux et al., 2006). So it is possible that the addition of high amounts of recombinant C53-C37 could be enough to support some interactions with the rest of the RNAPIII even in the absence of the C11 subunit. Kinetic analyses showed that RNAPIII Δ elongated faster than the wild-type enzyme,

suggesting that C53-C37 promotes termination by reducing the elongation rate, therefore increasing the dwelling time of the polymerase at the terminator (Landrieux et al., 2006). Arimbasseri and Maraia (2013) further compared the properties of RNAPIII Δ and the wild-type RNAPIII and proposed relatively distinct termination mechanisms for each form of RNAPIII. They observed that the 17-subunit holoenzyme can terminate on T-tracts composed of as few as 5 or 6 Ts, or on the proximal part of a 9 Ts tract, whereas termination by the 14-subunit RNAPIII Δ requires longer T-tracts (≥ 9 Ts) and termination indeed occurs mainly at the distal part of the terminator. The requirement for a longer T-tract by RNAPIII Δ could reflect its lower sensitivity to the instability of the oligo(rU:dA) hybrid, as it would require longer, and therefore less stable, rU:dA hybrids. Accordingly, the C53-C37-C11 subcomplex might play a role in increasing the sensibility to the instability of oligo(rU:dA) hybrid (Arimbasseri and Maraia 2013). These hypotheses were supported by a following study, in which the authors showed that RNAPIII Δ and holo enzyme differ in their sensitivity to the rU:dA hybrid and that weakening the short rU:dA hybrid by using 4-thio-UTP, a uridine-triphosphate analog with decreased base pairing strength, can compensate for the lack of C53-C37-C11 (Mishra and Maraia, 2019).

Former structural data provided a possible explanation for the higher sensibility of RNAPIII to the stability of the RNA:DNA duplex at the active site, compared to the other eukaryotic RNAPs (Hoffmann et al., 2015). Indeed, RNAPIII appears to bind the RNA:DNA hybrid less tightly than the other polymerases, although more extensive interactions were observed with the downstream DNA duplex, which was suggested to compensate the loose grip of the DNA:RNA hybrid during elongation.

Using purified *S. cerevisiae* RNAPIII and various nucleic acid scaffolds, Arimbasseri and Maraia (2015) revealed an unexpected role of the non-template DNA strand in RNAPIII termination. The authors observed that a majority of ECs paused upon transcribing the first four Ts of a terminator and entered into a metastable, yet catalytically active, intermediate that they called the pre-termination complex (PTC). The data suggested that both the first four Ts in the non-template strand and the C53-C37-C11 subunits are required for the formation of a PTC. Substitution of the first four Ts abolished polymerase pausing and RNA release, whereas a mutation of the fifth T retained pausing but abrogated RNA release, indicating that a stretch of 4 Ts is only sufficient for polymerase pausing but transcript release requires the presence of a fifth T in the non-template strand. On the other hand, they found that the A-tract in the template strand is a strict requirement for termination, further supporting the idea that the weak oligo(rU:dA) hybrid is an essential determinant for termination. Based on these and former data, Arimbasseri and Maraia proposed that the template strand promotes destabilization of the EC through an unstable oligo(rU:dA) hybrid,

and the non-template strand provides distinct signals for RNAPIII pausing and release.

With regard to the recognition of the non-template strand signals, Arimbasseri and Maraia observed that five amino acids (226-230) in the C-terminal domain of C37 were important for eliciting RNAPIII release, thus they proposed that this C37 region would mediate specific interactions with the fifth T nucleotide of the terminator (Arimbasseri and Maraia, 2015). This region of C37 was previously identified by cross-linking experiments as binding to C128 and localizing to the active center (Wu et al., 2011), and by cryo-EM study as positioned in close proximity to the non-template strand (Hoffmann et al., 2015). However, very recent structural analyses of the yeast RNAPIII PTC have shown that the recognition of the Ts in the non-template strand does not involve C37, but is actually mediated by several residues of the second largest subunit of RNAPIII, C128 (Mathias Girbig and Christoph W. Muller, personal communication). Although direct engagement between C53-C37 and the non-template strand could not be observed, the RNAPIII Δ construct fails to stabilize the non-template strand as shown by cryo-EM (Mathias Girbig and Christoph W. Muller, personal communication). Taken all together, the precise role of C53-C37-C11 in termination is more complicated than expected, probably by aiding C128 to sense the termination signals, which remains to be further elucidated.

In conclusion, termination by RNAPIII is driven by an interplay between the DNA template, the RNA transcript and the polymerase itself, mainly mediated by the subunits C53-C37, C11 and C128. The T-tract is a bipartite termination signal as both its template and non-template strands carry distinct information to direct different stages of RNAPIII termination. Specifically, the oligo(dT) sequence in the non-template strand promotes both RNAPIII pausing and release, while the oligo(dA) in the template strand leads to formation of a weak rU:dA hybrid in the polymerase active center that acts as a destabilizing signal (Figure 3-10; Arimbasseri and Maraia, 2016; Porrua et al., 2016; Mishra and Maraia, 2019).

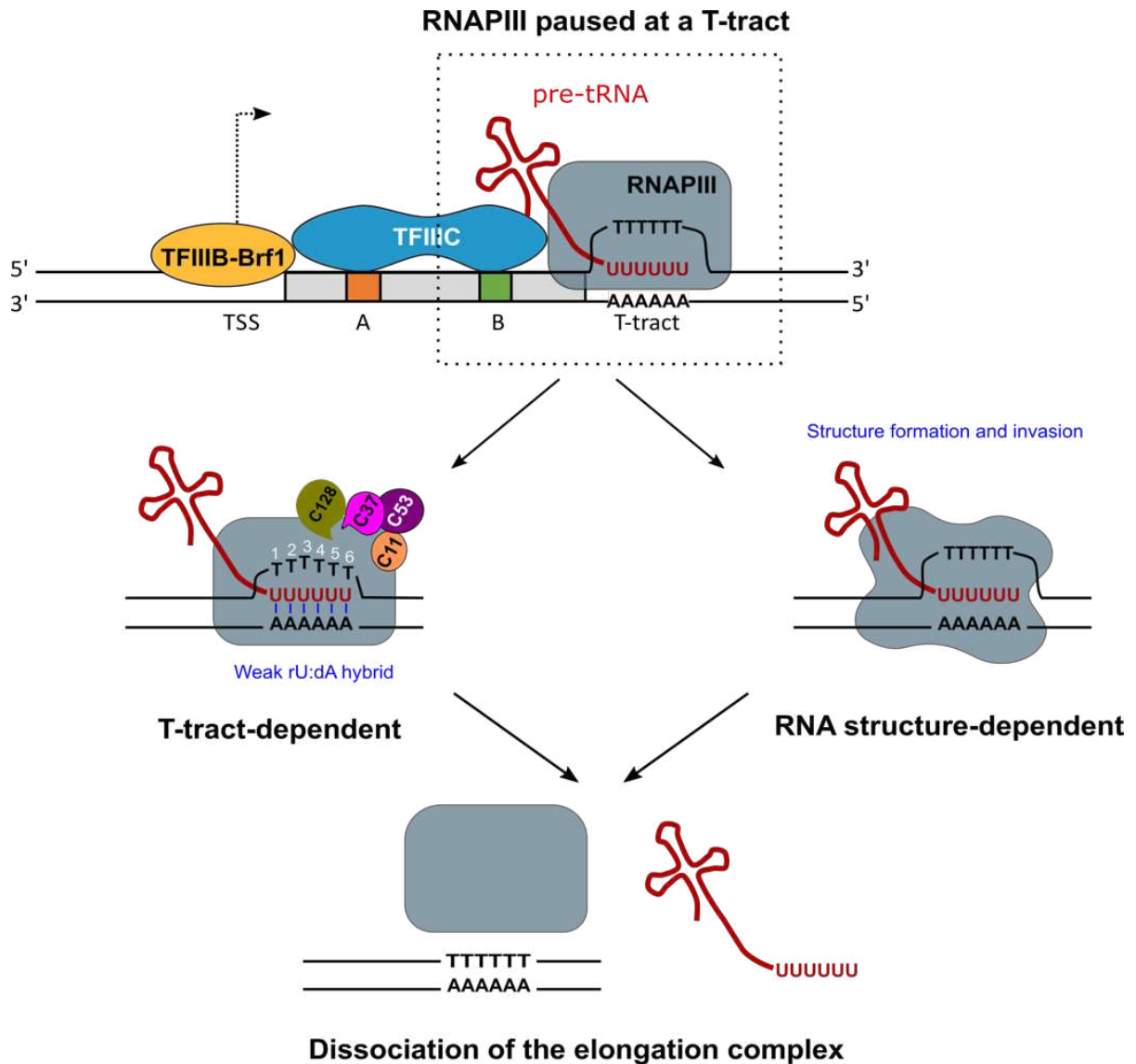


Figure 3-10: Models of Transcription termination by RNAPIII.

Transcription termination occurs after pausing upon synthesis of a U-tract. Two mechanisms are proposed for the dissociation of the elongation complex. The T-tract dependent model posits that RNAPIII is released by the weakness of the rU:dA hybrid in the catalytic sites and some specific interactions between several subunits of RNAPIII and the T-tract on the non-template strand. While the RNA-structure-dependent model suggests that invasion of the RNA structure induces RNAPIII conformational change, which destabilizes and releases the elongation complex.

3.5.2 The role of RNA secondary structures in termination

A new model for RNAPIII transcription termination was proposed by Nielsen et al. (2013) according to which a hairpin structure formed by the nascent RNA is an absolute requirement for RNAPIII termination, somewhat similar to the intrinsic termination pathway for bacterial RNAP (see **Box 2-4**). Using purified yeast RNAPIII and *in vitro* transcription termination assays, the authors of that study observed that these polymerases paused on stretches of 7, 8 or 12 Ts but failed to dissociate from the DNA. However, introducing a hairpin-like structure immediately upstream of the poly-U sequence significantly facilitated RNAPIII release. They further found that a distance as long as ~12 nt between the poly-U stretch and the upstream RNA structure still allowed efficient termination, while longer spacers prevented the RNA hairpin from promoting termination. The fact that hairpin-mediated termination could occur even when placed at a certain distance from the poly-U sequence suggested that, upon pausing within the T-tract, RNAPIII can move backward to some extent to approach the hairpin. Importantly, their data challenged the idea that RNAPIII requires a weak RNA:DNA hybrid in the active center, as efficient hairpin-dependent termination could also be achieved when RNAPIII paused at a G-tract. Thus, Nielsen et al. suggested that the function of the T-tract would be to induce the pausing and catalytic inactivation of RNAPIII by switching off the RNA cleavage activity mediated by C11. RNAPIII termination would then be triggered by the formation of RNA secondary structures within the RNA exit channel or by the invasion of this channel by the RNA hairpin upon RNAPIII backtracking, which would destabilize the EC via an allosteric mechanism (Nielsen et al. 2013).

However, this model remains very controversial as it is in contradiction with data from several groups. Indeed, prior *in vitro* transcription assays showed that a T-tract alone, without the presence of an upstream hairpin, is sufficient for efficient termination by RNAPIII (Bogenhagen and Brown, 1981; Wang and Folk, 1994; Arimbasseri et al., 2013). In addition, it was shown both *in vivo* and *in vitro* that RNA 3' end cleavage is active during RNAPIII termination (Huang et al., 2005; Bobkova et al., 1999; Rijal and Maraia, 2013). A follow-up study performed by Arimbasseri et al. (2014) implicated that the discrepant results obtained by Nielsen et al. could be due to the specific polymerase preparation used rather than variables of the different assays. By comparing the RNAPIII preparation assessed by Nielsen and co-authors with two other RNAPIII preparations from distinct groups, Arimbasseri et al. found that Nielsen's protein was far more active than the other two proteins and, indeed, it was more prone to read through the terminator. Arimbasseri et al. also showed that, unlike Nielsen's protein, the other two proteins can terminate efficiently at a T-tract in the absence of an upstream hairpin. They further showed that RNA structures are not required for RNAPIII termination *in vivo* in a particular yeast reporter system (Arimbasseri et al. 2014).

A very recent study using a reporter system in human cells showed that an RNA hairpin placed just upstream of a run of 4 Ts can enhance transcription termination of RNAPIII initiating from a type 3 promoter. However, the RNA structure by itself could not induce termination in the absence of an oligo(dT) sequence. These findings indicate that RNA structures can also partake in RNAPIII termination (Verosloff et al., 2021). Nonetheless, whether these structures are functional only in particular sequence contexts or in specific organisms is still unclear.

3.5.3 Extrinsic termination factors

Several extrinsic factors have shown to promote RNAPIII termination *in vitro*. These include the La protein (Lhp1 in budding yeast), TFIIIC, topoisomerase-1 (topo-1) and PC4 (reviewed in Arimbasseri et al., 2013). TFIIIC were found to be cross-linked to the terminators of RNAPIII-transcribed genes, suggesting a role in termination (Bartholomew et al., 1990; Bartholomew et al., 1991). Both topo-1 and PC4 were found to copurify with TFIIIC and can promote both termination and reinitiation by RNAPIII (Wang and Roeder, 1998). The La protein is an RNA binding protein that associates with all newly synthesized RNAPIII transcripts and protects the 3' end of these RNAs from exonuclease digestion. La association is required for the maturation of pre-tRNAs and the assembly of RNP, and contributes to nuclear retention of certain small RNAs (reviewed in Wolin and Cedervall, 2002). A couple of evidence obtained from *in vitro* transcription assays showed that La protein is required for proper termination and reinitiation of RNAPIII, as well as for transcript release (Goodier and Maraia, 1998; Maraia et al., 1994; Gottlieb and Steitz, 1989; Fan et al., 1997; Goodier et al., 1997), which, however, has been contested by other studies (Lin-Marq and Clarkson, 1998; Hu et al., 2003; Schramm and Hernandez, 2002). Whether or not these factors are directly involved in RNAPIII transcription termination needs to be verified.

During the course of my PhD, a recent study showed that one of the homologs of the *S. cerevisiae* helicase Sen1 in *S. pombe* is required for robust transcription termination by RNAPIII *in vivo* (Rivosecchi et al., 2019). Fission yeast *S. pombe* contains two Sen1 homologs: Sen1 (hereafter *Sp* Sen1) and Dbl8. Unlike *S. cerevisiae* Sen1, none of the two *S. pombe* Sen1 homologs is essential for viability or is involved in RNAPII transcription termination (Larochelle et al., 2018). A previous study indicated that *Sp* Sen1 interacts with RNAPIII but not with RNAPI or RNAPII (Legros et al., 2014). Rivosecchi et al. further showed *Sp* Sen1 but not Dbl8 is involved in RNAPIII termination in fission yeast. Specifically, ChIP-seq analyses performed in a $\Delta sen1$ strain showed that RNAPIII accumulated downstream of class III genes, indicating transcription termination defects. Previous data had revealed that unstable R-loops form at tRNA loci in fission yeast and that *Sp* Sen1 was recruited to these loci, raising the question

whether *Sp* Sen1 recruitment and action on RNAPIII transcription would involve the resolution of R-loops, which was proposed to be an important function of *S. cerevisiae* Sen1 (Legros et al., 2014). However, Rivosecchi et al. showed that the function of Sen1 in RNAPIII transcription termination was independent on the presence of R-loops at tRNA regions. The precise mechanisms by which *Sp* Sen1 would stimulate termination have not been explored, but based on its homology with *S. cerevisiae* Sen1, it was proposed that it would employ the same mode of action as *S. cerevisiae* Sen1 on RNAPII-dependent transcription units. Whether the function of *Sp* Sen1 in RNAPIII transcription termination is conserved in other organisms remains an open question.

3.6 Facilitated reinitiation by RNAPIII

After transcription termination, polymerases are available to initiate a new transcription cycle. In the case of RNAPIII, transcription reinitiation occurs preferentially on the same gene, a phenomenon that is referred to as facilitated recycling. This process was discovered by Dieci and Sentenac (1996), when performing *in vitro* transcription assays with yeast RNAPIII. They found that RNAPIIIs completed each new cycle 5-10 times faster than the first round of transcription and were quite resistant to the presence of heparin. The polyanion heparin mimics the structure of nucleic acids and is used as preinitiation inhibitor because it can sequester RNAPIIIs not engaged in elongation, for instance unassembled or terminating RNAPIIIs (Kassavetis et al., 1989, Dieci et al., 2014). From these observations they concluded that RNAPIIIs were somehow retained on the same template for multiple rounds of transcription, implying that RNAPIIIs are not fully released from the DNA after termination but instead are committed to reinitiation on the same gene. In addition, they noticed that the RNAPIII canonical termination signal was required for efficient reinitiation, as RNAPIIIs that read-through the termination site exhibited reduced recycling ability, probably due to the loss of physical proximity between terminating RNAPIII and transcription factors, indicating an important link between the termination and the reinitiation steps (Figure 3-11A). Studies in both yeast and humans revealed that a substantial amount of RNAPIIIs escape the canonical terminators (Canella et al., 2010; Orioli et al., 2011; Turowski et al., 2016), and then preferentially terminate at U-rich tracts downstream of the canonical terminators (Orioli et al., 2011; Turowski et al., 2016), with readthrough transcripts targeted for degradation by the exosome (Figure 3-11C; Turowski et al., 2016; Rivosecchi et al., 2019).

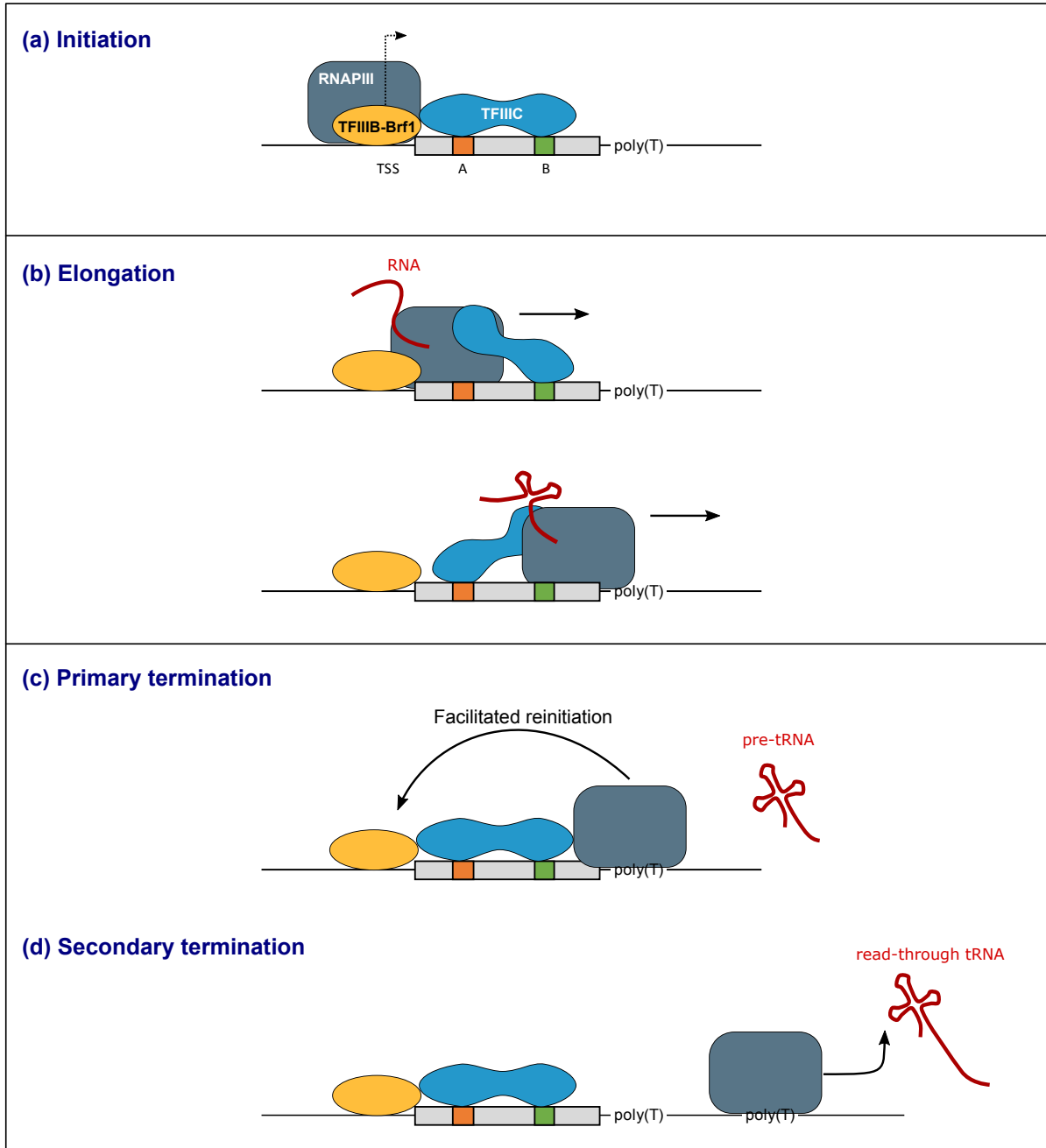
This ability of RNAPIII to be reloaded on the same transcription unit to achieve high transcription rates (Dieci and Sentenac 1996) was later found to be conserved from yeast to humans (Cabart et al., 2008). However, the molecular mechanism of termination-coupled

facilitated reinitiation by RNAPIII remains poorly understood. An *in vitro* study on RNAPIII reinitiation properties suggested that facilitated recycling is not a stochastic process, but relies on a specific polymerase recapture pathway involving the promoter-bound transcription factors TFIIB and TFIIC. Particularly, TFIIB was capable of directing recycling on a short template (~100 bp) in the absence of TFIIC. However, on long genes (> 300 bp) such as *SCR1*, TFIIC was further required to support a high reinitiation rate (Ferrari et al., 2004). Biochemical evidence showed that the termination-linked reinitiation also involved the action of the RNAPIII subunits C53, C37 and C11 (Landrieux et al., 2006). Among them, C11 was critical for transcription reinitiation independently of its RNA cleavage stimulation activity. The interaction of C11 with the C53-C37 heterodimer would induce a conformational change in the RNAPIII needed for the recapture pathway. The C53-C37 heterodimer was proposed to slow down the elongation rate and increase the duration of pausing either during elongation or at the terminator, thus giving time for the correct recognition of termination element and, thus, also favoring facilitated reinitiation (Landrieux et al., 2006). However, how the interactions between the terminating RNAPIII and initiation factors assembled on the promoter are established remains unclear, and structural evidence for the “reinitiation complex” is greatly needed.

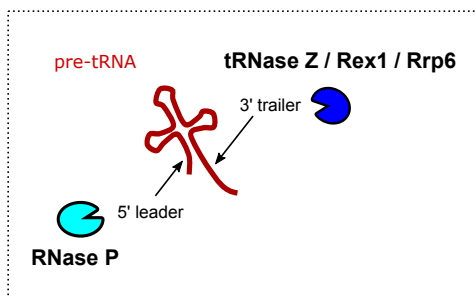
Figure 3-11: The RNAPIII transcription cycle (next page).

(A) RNAPIII Transcription cycle on tRNA genes. **(a)** Transcription initiation: the internal promoters (A box and B box) of tRNA genes are bound by the transcription factor TFIIC, which recruits TFIIB to the DNA. RNAPIII is then recruited and transcription initiates at the transcription start site (TSS). **(b)** Transcription elongation: during elongation, TFIIC remains associated with the promoter elements, possibly because RNAPIII displaces TFIIC from A box and subsequently from B box during elongation, but not from both sites simultaneously. **(c)** Primary (canonical) termination: RNAPIII terminates at a canonical terminator sequence and releases the pre-tRNA transcript. The released RNAPIII can rapidly re-associate with the same transcription unit with the help of promoter-bound TFIIC and TFIIB and start a new transcription cycle through a process known as “facilitated reinitiation” (Dieci and Sentenac, 1996). **(d)** Secondary termination: a substantial fraction of RNAPIII overrides the canonical terminator and then terminates at regions downstream of 3'-end of tRNA, thus producing aberrant readthrough transcripts. **(B)** Pre-tRNA processing. The primary transcripts of tRNA genes must undergo maturation at both ends to generate mature tRNAs. The 5'-end leader of the pre-tRNA is generally removed by the RNase P endonuclease, and the 3'-end trailer is thought to be cleaved by the tRNase Z endonuclease or trimmed by the exonucleases Rex1 and Rrp6 (Skowronek et al., 2014). **(C)** Readthrough transcripts processing. Readthrough tRNAs can be degraded by the nuclear surveillance machineries such as the exosome, or can be processed through other mechanisms. Adapted from Turowski and Tollervey, 2016.

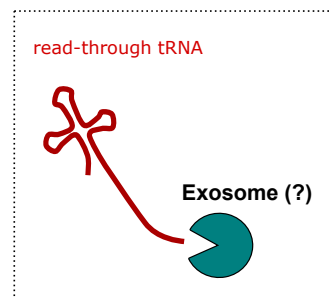
A



B



C



Research outline and main objectives

The Nrd1-Nab3-Sen1 (NNS) complex in budding yeast is dedicated to terminate non-coding RNAs transcribed by RNAPII. Previously, studies in our group identified a major role for the helicase Sen1 in dismantling the elongation complex, which requires its RNA-binding and ATP-hydrolysis activities (Porrua and Libri, 2013b; Han et al., 2017). The team showed that the helicase domain of Sen1 alone is sufficient to induce RNAPII transcription termination *in vitro* (Han et al., 2017). However, the N-terminal domain (NTD) of Sen1 is essential for viability as well as for RNAPII transcription termination *in vivo*, likely by recognizing the CTD of RNAPII (Han et al., 2020). In order to investigate other functional interactions mediated by the Sen1 NTD, we studied the protein interactomes of Sen1 and several Sen1 variants. Surprisingly, we found that deletion of the NTD of Sen1 leads to the loss of its association with replication factors and RNAPIII, whereas the NTD alone could interact with the replisome and RNAPIII, suggesting a direct role for Sen1 NTD in binding the replisome and RNAPIII (data obtained by Nouhou Haidara). But little was known about the biological function of these interactions.

In a collaborative study with the De Piccoli group (Coventry, UK), it was found that budding yeast Sen1 binds to the replisome via the replication factors Ctf4 and Mrc1. Our collaborators identified a region of Sen1 NTD that is involved in these interactions and generated a mutant version of Sen1 containing substitutions at three conserved residues (W773, E774, W777), referred to as *sen1-3*, that lost the capacity to bind the replisome but was fully competent for cell growth and RNAPII transcription termination (Appanah et al., 2020). Further analyses of the interactome of the Sen1-3 protein demonstrated that it was impaired in the interaction with RNAPIII as well (data obtained by Umberto Aiello). Therefore, the *sen1-3* allele is a separation-of-function mutant that allows us to study the role of Sen1 in RNAPIII transcription and replication without affecting the efficiency of RNAPII transcription termination.

This was the moment when I joined the team and started my PhD project. My study aimed at characterizing in more detail the interaction of Sen1 with RNAPIII and explore the function of this interaction in RNAPIII transcription in *S. cerevisiae*. To this end:

- 1) I analysed RNAPIII interacting partners by co-immunoprecipitation followed by mass-spectrometry analysis to understand whether the interaction of Sen1 with RNAPIII could be mediated by the replisome.

- 2) I generated high-resolution maps of transcribing RNAPIII using the CRAC technique (**Box 3-2**) in a wild-type, *sen1-3*, or a Sen1-depleted strain, to determine the role of Sen1 and the Sen1-RNAPIII interaction on RNAPIII transcription. I also performed RNAPIII CRAC under additional conditions to assess whether the role of Sen1 in RNAPIII transcription requires other components of the NNS complex or the presence of replisome.
- 3) I performed *in vitro* transcription termination assays using purified RNAPIII and Sen1, to characterize the mechanisms of Sen1-mediated RNAPIII termination.
- 4) I also tested the role of RNA secondary structures in RNAPIII transcription termination using similar *in vitro* systems.

By these approaches, we expected to unveil a new function for the highly conserved helicase Sen1 and revisit the current model for RNAPIII transcription termination.

MANUSCRIPT

An integrated model for termination of RNA polymerase III transcription

Juanjuan Xie¹, Umberto Aiello^{1°}, Yves Clement^{1°}, Nouhou Haidara^{1°}, Mathias Girbig^{2,3}, Jana Schmitzova^{4,5}, Vlad Pena^{4,6}, Christoph W. Müller², Domenico Libri^{1*} and Odil Porrua^{1*}

°Equal contribution

*Correspondence should be addressed to domenico.libri@ijm.fr and odil.porrua@ijm.fr

Affiliations:

¹ Université de Paris, CNRS, Institut Jacques Monod, F-75006 Paris, France.

² European Molecular Biology Laboratory (EMBL), Structural and Computational Biology Unit, 69117 Heidelberg, Germany.

³ (Candidate for) Joint PhD degree from EMBL and Heidelberg University, Faculty of Biosciences.

⁴ Max Planck Institute for Biophysical Chemistry, Macromolecular Crystallography, Am Fassberg 11, 37077, Goettingen, Germany

⁵ Present address: Max Planck Institute for Biophysical Chemistry, Department of Molecular Biology Am Fassberg 11, 37077 Göttingen, Germany.

⁶ The Institute of Cancer Research, Structural Biology Division, Fulham Road, SW7 3RP, London, UK.

Abstract

RNA polymerase III (RNAPIII) synthesizes essential and abundant non-coding RNAs such as tRNAs. Controlling RNAPIII span of activity by accurate and efficient termination is a challenging necessity to ensure robust gene expression and to prevent conflicts with other DNA-associated machineries. The mechanism of RNAPIII termination is believed to be simpler than that of other eukaryotic RNA polymerases, solely relying on the recognition of a T-tract in the non-template strand. Here we combine high-resolution genome-wide analyses and *in vitro* transcription termination assays to revisit the mechanism of RNAPIII transcription termination in budding yeast. We show that T-tracts are necessary but not always sufficient for termination and that secondary structures of the nascent RNAs are important auxiliary cis-acting elements. Moreover, we show that the helicase Sen1 plays a key role in a fail-safe termination pathway. Our results provide a comprehensive model illustrating how multiple mechanisms cooperate to ensure efficient RNAPIII transcription termination.

Introduction

Transcription termination is an essential process that sets the borders between genes, therefore avoiding the interference between neighboring transcription units. Furthermore, transcription termination plays an important role in the maintenance of genome integrity by limiting the possible conflicts between transcribing RNA polymerases (RNAPs) and other cellular machineries involved in DNA replication or repair (reviewed in Porrua and Libri, 2015a).

Transcription termination can be envisioned as a multi-step process consisting in the recruitment of termination factors, the recognition of sequence motifs, RNAP pausing, and finally the release of the RNAP and the transcript from the DNA. This last step involves a remodeling of an intricate network of interactions between the RNAP, the nascent RNA and the DNA template (reviewed in Porrua et al., 2016). Within this network, the interactions between the polymerase and the RNA:DNA hybrid are considered as the main determinant of the stability of the elongation complex (EC) (Kireeva et al., 2000). Most eukaryotic organisms possess three different RNAPs that are specialized in producing different classes of transcripts and seem to adopt different strategies to efficiently terminate transcription. RNAPI is responsible for the synthesis of ribosomal RNAs; RNAPII transcribes all protein-coding genes and several classes of non-coding genes and RNAPIII synthesizes short and abundant transcripts among which all tRNAs, the 5S rRNA, and several additional non-coding RNAs.

The mechanisms of transcription termination of the three polymerases have been extensively characterized in the eukaryotic model *Saccharomyces cerevisiae* and many of the principles uncovered in this organism seem to be highly conserved from yeast to humans (reviewed in Porrua et al., 2016). RNAPI and RNAPII require extrinsic protein factors to terminate transcription. RNAPI pauses when it encounters a Myb-like factor bound to the DNA downstream of each rRNA gene (Merkl et al., 2014; Reiter et al., 2012). The release of the paused RNAPI is then mediated by additional proteins, specifically the Rat1 exonuclease and the helicase Sen1 (El Hage et al., 2008; Kawauchi et al., 2008), which are also major termination factors for RNAPII (see below).

The mechanism of RNAPII transcription termination is more complex and involves the action of a larger number of proteins. There are two major termination pathways for RNAPII (reviewed in Porrua and Libri, 2015a). Transcription termination at protein-coding genes relies on a multi-subunit complex that is responsible for the co-transcriptional cleavage of the pre-mRNA at the poly(A) site and the addition of a poly(A) tail. The downstream portion of

the nascent transcript is then targeted by Rat1 (XRN2 in humans), which degrades the RNA molecule until it encounters RNAPII and promotes its release from the DNA (Baejen et al., 2017; Kim et al., 2004; Park et al., 2015; Pearson and Moore, 2013; West et al., 2004).

The second pathway is devoted to termination of non-coding transcription and plays an essential role in the control of pervasive transcription as well as in the biogenesis of snoRNAs (Arndt and Reines, 2015; Porrua and Libri, 2015a). This pathway depends on a complex composed of two RNA-binding proteins, Nrd1 and Nab3, and the aforementioned helicase Sen1 (i.e. the NNS complex). Whereas Nrd1 and Nab3 recognize specific sequence motifs that are enriched in the target non-coding RNAs, the helicase Sen1 induces the dissociation of the EC (Porrua and Libri, 2013b; Porrua et al., 2012; Schulz et al., 2013; Steinmetz et al., 2006; Wlotzka et al., 2011). The mechanisms of action of Sen1 in RNAPII transcription have been extensively characterized at the molecular level by our group and others (Han et al., 2017; Hazelbaker et al., 2013; Leonaitė et al., 2017; Porrua and Libri, 2013b; Wang et al., 2019). Briefly, Sen1 uses the energy of ATP hydrolysis to translocate along the nascent RNA towards the transcribing RNAPII and, upon transcriptional pausing, it collides with the polymerase and induces its dissociation from the DNA.

A large body of evidence supports the notion that, in contrast to the other RNAPs, RNAPIII can terminate precisely and efficiently at a particular DNA sequence without the need for accessory proteins (reviewed in Arimbasseri et al., 2013 and Porrua et al., 2016). A typical RNAPIII terminator consists in a stretch of thymidines (T) of variable length in the non-template DNA strand that, according to the current model, is sufficient to promote both pausing and release of RNAPIII. Upon transcription of a T-tract, the weakness of the resulting rU:dA hybrid is thought to be central to the destabilization of the RNAPIII EC (Mishra and Maraia, 2019). The particular sensitivity of RNAPIII to weak rU:dA hybrids compared to other RNAPs that do not sense T-tracts as terminators is believed to depend on the less-extensive interactions between RNAPIII and the RNA:DNA hybrid (Hoffmann et al., 2015). The Ts in the non-template strand play an additional critical role in transcription termination (Arimbasseri and Maraia, 2015), as they have been proposed to be recognized by the C37 and C53 subunits of RNAPIII that also contribute to termination (Landrieux et al., 2006; Rijal and Maraia, 2013). An alternative model proposed by Nielsen and coauthors (Nielsen et al., 2013) posits that T-tracts are required for RNAPIII pausing but are not sufficient for its release from the DNA. These authors have proposed that the folding of the nascent RNA into a hairpin-like structure in the vicinity of the paused RNAPIII is an absolute requirement for termination. The hairpin would invade the RNA exit channel of the polymerase, thus provoking its dissociation from the DNA. The proposed mechanism is reminiscent of the so-called intrinsic termination pathway described for bacterial RNAP. This hairpin-dependent model remains, however,

highly controversial since it is seemingly in disagreement with a large body of former experimental evidence (Arimbasseri et al., 2014).

The model according to which sequence signals are the sole determinant of RNAPIII termination has also been challenged in the fission yeast *Schizosaccharomyces pombe* by a recent report showing that one of the homologues of the *S. cerevisiae* Sen1 (hereafter designated *Sp* Sen1) is involved in RNAPIII termination *in vivo* (Rivosecchi et al., 2019). Deletion of this gene that in *S. pombe* is non-essential leads to a global shift of RNAPIII occupancy downstream of tRNA genes, consistent with the notion that *Sp* Sen1, in addition to T-tracts, is required for RNAPIII termination in this organism. The precise role of *Sp* Sen1 in termination as well as its mechanism of action were, however, not addressed in this study. Thus, much uncertainty remains about the relative contribution of sequence elements, RNA structures and *trans*-acting factors to the efficiency of RNAPIII transcription termination. Also, to what extent the different termination mechanisms are evolutionary conserved remains an open question.

In the present study we combine high-resolution genome-wide approaches with *in vitro* transcription termination assays using highly-purified components to dissect the mechanism of RNAPIII transcription termination in *S. cerevisiae*. We observe that termination at the primary terminator of RNAPIII-dependent genes (i.e. the first T-tract after the gene), is only partially efficient and, thus, a considerable fraction of polymerases terminate in the downstream region. We provide *in vivo* and *in vitro* evidence that the helicase Sen1 plays a global role in RNAPIII transcription termination and that this function relies on the interaction of its N-terminal domain with RNAPIII. However, we find that Sen1 contributes very little to the efficiency of primary termination and that it mainly functions as a fail-safe mechanism to promote termination of RNAPIIIs that override the first termination signal. Our data indicate that only T-tracts within a particular length range are sufficient to promote autonomous termination by RNAPIII. Nevertheless, we show that tRNA genes often contain suboptimal termination signals and that their capacity to induce termination can be complemented by Sen1 as well as by secondary structures of the nascent RNA. These two factors act in a mutually exclusive manner since the presence of RNA structures prevent the loading of Sen1 onto the transcript, which is strictly required for Sen1-mediated termination. While Sen1 can also promote the release of RNAPIII at pausing sites other than T-tracts, we find that RNA structures can only function in association with canonical termination signals.

Together, our data allow revisiting former models for RNAPIII transcription termination and offer a novel and detailed view of how intrinsic components of the EC (i.e. T-tracts and RNA structures) and the extrinsic factor Sen1 concur to promote efficient termination of RNAPIII transcription.

Results

The N-terminal domain of Sen1 interacts with RNAPIII

S. cerevisiae Sen1 is a modular protein composed of a large N-terminal domain (aa 1-975), a central helicase domain (aa 1095-1867) and a C-terminal disordered region (aa 1930-2231, see figure 1A). We have recently shown that the N-terminal domain (NTD) is essential for viability and for termination of RNAPII transcription and that it recognizes the CTD of RNAPII, although it is not the only RNAPII-interacting region in Sen1 (Han et al., 2020). In a quest for other functional interactions mediated by the Sen1 Nter, we performed co-immunoprecipitation (co-IP) experiments followed by mass spectrometry (MS) analyses using either a full-length or a Δ NTD version of Sen1 as a bait (table 1 and S1). We expressed both *SEN1* variants from the *GAL1* promoter (p*GAL1*) because only overexpression of the *sen1* Δ NTD allele supports viability (Han et al., 2020). In agreement with previous reports (Appanah et al., 2020; Yüce and West, 2013), we detected an RNase-resistant interaction of Sen1 with its partners within the NNS-complex Nrd1 and Nab3, several replication and transcription-related factors, as well as with the three RNAPs. Strikingly, deletion of the NTD abolished the association of Sen1 with RNAPIII and most replication factors without markedly affecting other interactions. Additional co-IP/MS experiments using the isolated NTD as a bait confirmed the interaction with replication factors (e.g. Ctf4) and RNAPIII subunits, strongly suggesting direct protein-protein interactions between the NTD and these factors (table 1 and S2).

The interaction of the N-terminal domain of Sen1 with the replisome was found to depend on the replication factors Ctf4 and Mrc1 in a parallel, collaborative study (Appanah et al., 2020). In that work, we found that combination of three point mutations in a conserved region of the Sen1 NTD (W773A, E774A, W777A; defining the Sen1-3 variant) abolishes the interaction with these proteins. Importantly, we showed that Sen1-3 is expressed at similar levels as WT Sen1 and is fully proficient for terminating transcription of NNS target genes (Appanah et al., 2020). To assess whether these mutations also affect the association with RNAPIII, we analysed the protein interactome of Sen1-3 by co-IP/MS (figure 1A, tables 1 and S3). The interaction with RNAPII was not significantly altered in this mutant, in agreement with its proficiency in RNAPII transcription termination (Appanah et al., 2020). Interestingly, we observed that the mutations introduced in Sen1-3 strongly affect the interaction with RNAPIII subunits.

These results are compatible with the notion that the same surface of Sen1 mediates mutually exclusive interactions with the replisome and RNAPIII. Alternatively, the interaction

between Sen1 NTD and RNAPIII could be mediated by the replisome. To distinguish between these possibilities, we conducted quantitative MS and western blot analyses on RNAPIII coimmunoprecipitates from WT and *sen1-3* cells (figure 1B-D and table S4). We observed a clear association of RNAPIII with protein components of the Ty1 transposon, which was previously reported and validates our experimental conditions (figure 1C, Bridier-Nahmias et al., 2015). Importantly, while Sen1 was among the most enriched RNAPIII interactors, we did not detect the two replisome anchoring factors, Ctf4 and Mrc1, indicating that Sen1 interacts in a mutually exclusive manner with RNAPIII and the replisome. RNase A treatment induced a ~2-fold decrease in the level of RNAPIII-bound Sen1, indicating that this interaction is also partially mediated or stabilized by the RNA (figure 1E). As expected, the association of Sen1-3 with RNAPIII was strongly reduced compared to WT Sen1 (figure 1D), even in the absence of RNase treatment, suggesting that the protein-protein interaction mediated by Sen1 NTD is a major pre-requisite for the association of Sen1 with RNAPIII transcripts. Strikingly, the Sen1 NNS partners Nrd1 and Nab3 were very poorly enriched in RNAPIII coimmunoprecipitates (figure 1C-E), strongly suggesting that Sen1 plays a role in RNAPIII transcription termination independently from its function within the NNS-complex.

Taken together, our results support the notion that Sen1 associates with RNAPIII and the replisome within two alternative complexes that are also distinct from the NNS-complex and likely exert different functions.

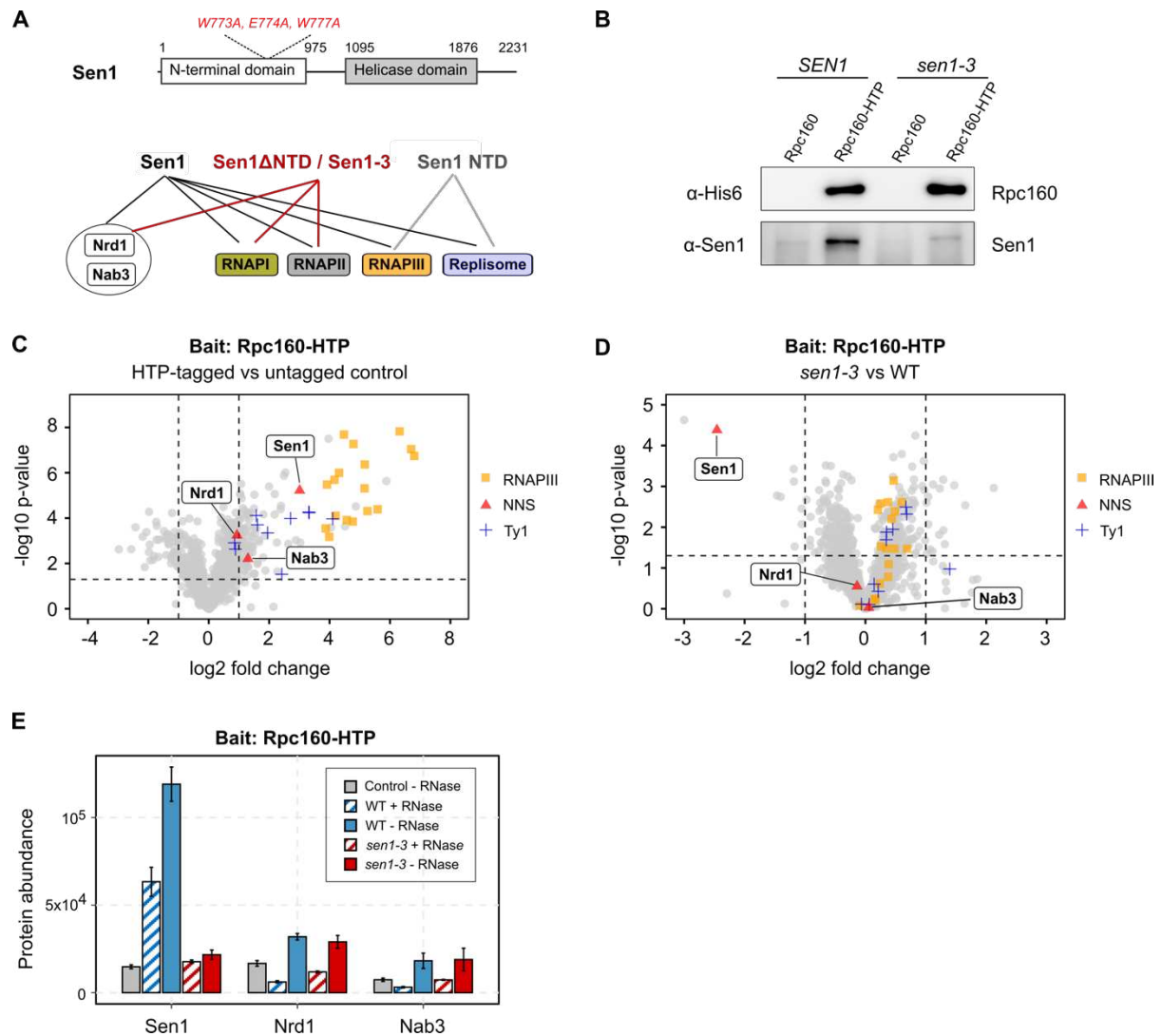


Figure 1: The N-terminal domain of Sen1 interacts with RNAPIII.

A) Summary of the results of coimmunoprecipitation-MS experiments using different versions of TAP-tagged Sen1 as baits that are included in table 1. A scheme of Sen1 protein indicating the different functional domains as well as the position of the mutations introduced in the *sen1-3* strain is shown on the top. Globular domains are denoted by solid bars while protein regions predicted to be intrinsically disordered are indicated by a line.

B) Western blot analysis of a representative coimmunoprecipitation experiment using a C-terminally His₆-TEV-Protein A (HTP)-tagged version of the largest subunit of RNAPIII (Rpc160) as the bait.

C) and D) Label-free quantitative MS analysis of coimmunoprecipitation assays using Rpc160-HTP as the bait. Data correspond to experiments performed in the absence of RNase A treatment. C) Volcano plot representing the enrichment of the detected proteins in the HTP-tagged strain relative to the untagged control in a WT (*SEN1*) background. D) Quantitative comparison of the proteins that are associated with tagged RNAPIII in a *sen1-3* mutant relative to the WT. Only proteins with a fold change ≥ 2 relative to the control and p-value < 0.05 are considered as significantly changed among the compared conditions.

E) Comparison of the abundance (arbitrary units) of the different NNS-components in Rpc160-HTP coimmunoprecipitates with or without treatment with RNase A.

Table 1: mass spectrometry analyses of coimmunoprecipitation experiments using different versions of Sen1 as bait.

Values correspond to the mascot score. ND, not detected. Note that the mascot score depends on the size of the protein and therefore, truncated versions of Sen1 have lower values. ND, not detected. The full datasets and the results of additional replicates are included in tables S1, S2 and S3.

| Protein | Complex | TAP-Sen1 vs TAP-Sen1 Δ NTD | | | TAP-Sen1 NTD | | Sen1-TAP vs Sen1-3-TAP | | |
|---------|-----------|-----------------------------------|-------|-------------------|--------------|----------|------------------------|------|--------|
| | | Ctrl | Sen1 | Sen1 Δ NTD | Ctrl | Sen1 NTD | Ctrl | Sen1 | Sen1-3 |
| Sen1 | NNS | 50 | 24819 | 14182 | 153 | 4749 | 0 | 8280 | 7299 |
| Nrd1 | NNS | 0 | 439 | 665 | ND | ND | 0 | 315 | 240 |
| Nab3 | NNS | 0 | 417 | 575 | ND | ND | 0 | 185 | 89 |
| Ctf4 | Replisome | 0 | 2465 | 0 | 19 | 1343 | 0 | 181 | 0 |
| Mrc1 | Replisome | 0 | 40 | 0 | ND | ND | 0 | 63 | 21 |
| Rpa190 | RNAPI | 126 | 463 | 1326 | 31 | 0 | 492 | 2041 | 1812 |
| Rpa135 | RNAPI | 87 | 122 | 664 | 29 | 0 | 400 | 1505 | 1060 |
| Rpb1 | RNAPII | 0 | 2770 | 3003 | 81 | 102 | 113 | 2313 | 2494 |
| Rpb2 | RNAPII | 0 | 2302 | 2328 | 86 | 55 | 34 | 1482 | 1684 |
| Rpc160 | RNAPIII | 0 | 7479 | 0 | 0 | 358 | 25 | 2216 | 59 |
| Rpc128 | RNAPIII | 0 | 4731 | 0 | 0 | 125 | 42 | 938 | 204 |
| Rpc82 | RNAPIII | 0 | 2735 | 0 | 0 | 233 | 0 | 1135 | 34 |
| Rpc53 | RNAPIII | 68 | 1255 | 183 | 0 | 56 | 0 | 500 | 34 |
| Rpc37 | RNAPIII | 0 | 1982 | 0 | 0 | 132 | 0 | 197 | 50 |
| Rpc34 | RNAPIII | 0 | 2022 | 0 | 0 | 28 | 0 | 433 | 174 |
| Rpc31 | RNAPIII | 0 | 1212 | 0 | 0 | 19 | 0 | 426 | 0 |
| Rpc25 | RNAPIII | 0 | 410 | 0 | 0 | 0 | 0 | 79 | 0 |
| Rpc17 | RNAPIII | 0 | 422 | 0 | 0 | 91 | 0 | 128 | 0 |
| Rpc11 | RNAPIII | 0 | 191 | 0 | 15 | 37 | 33 | 138 | 62 |

Sen1 is required for efficient termination of RNAPIII transcription *in vivo*.

The most widely accepted model for RNAPIII transcription termination posits that the polymerases recognize a *cis*-acting element composed of a stretch of thymidines on the non-template DNA and is released without the need for additional *trans*- or *cis*-acting factors (reviewed in Arimbasseri et al., 2013 and Porrua et al., 2016). However, the evidence supporting a direct interaction between RNAPIII and Sen1 prompted us to investigate a possible role for the latter in terminating RNAPIII transcription. To this end, we generated high-resolution maps of transcribing RNAPIII by CRAC (crosslinking analysis of cDNAs) (Candelli et al., 2018; Granneman et al., 2009). Briefly, the nascent RNAs are UV-crosslinked to RNAPIII *in vivo* and the RNAPIII-RNA complexes are purified under stringent conditions. The extracted RNAs are then used to generate cDNAs that are deep-sequenced, providing the position of RNAPIII with nucleotide resolution. We performed these experiments in WT or *sen1-3* cells as well as in a Sen1-AID (auxin-inducible degron) strain, which allowed assessing the effect of Sen1 depletion (figure 2).

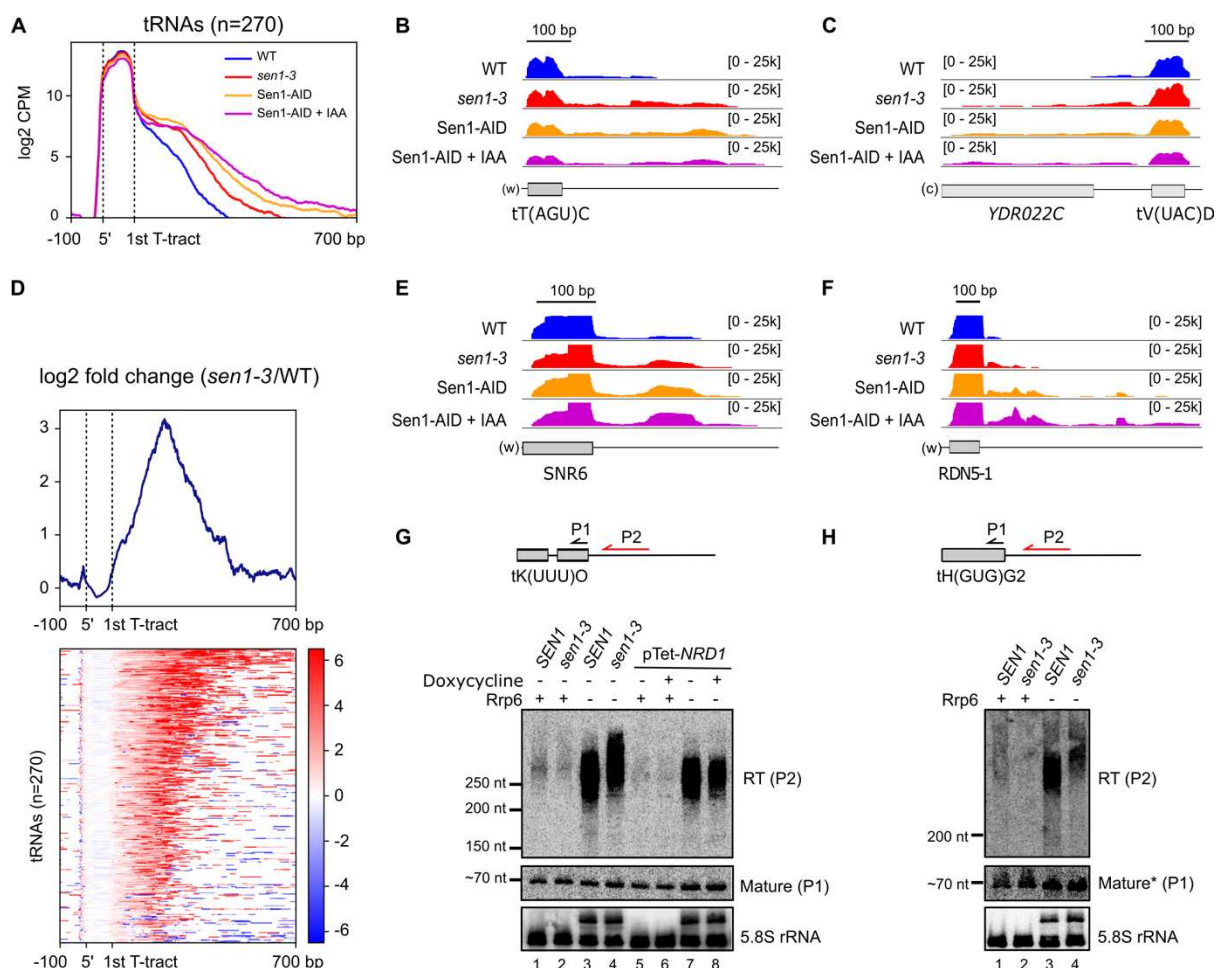


Figure 2: The interaction of Sen1 with RNAPIII is globally required for efficient transcription termination at RNAPIII-dependent genes.

A) Metagene analysis of the RNAPIII distribution around tRNA genes. The signal covering the region between the 5' and the primary terminator (i.e. the 1st T-tract after the 3' end of the mature tRNA) is scaled to 100 bp. Values on the y-axis correspond to the mean coverage expressed as counts per million (CPM) multiplied by 10. Sen1-AID denotes the strain expression an Auxin Inducible Degron version of Sen1. IAA: indole-3-acetic acid, an analogue of auxin.

B) and C) Integrative Genomics Viewer (IGV) screenshots of examples of tRNA genes displaying termination defects upon mutation or depletion of Sen1. “w” and “c” denote the Watson and the Crick strands, respectively. The values under brackets correspond to the scale of the RNAPIII signal expressed in 10xCPM.

D) Heatmap analysis representing the log₂ of the fold change (FC) of the RNAPIII signal around tRNA genes in the *sen1-3* mutant relative to the WT. The summary plot on the top was calculated using the average values for each position.

E) and F) Examples of RNAPIII-dependent genes other than tRNA genes that present termination defects upon mutation or depletion of Sen1.

G) and H) Northern blot analysis of transcripts derived from two different tRNA genes in the indicated backgrounds. Schemes on the top indicate the approximate position of the probes (P1 and P2) used for the detection of the different RNA species (RT, for read-through, and mature tRNA). The RNA probe is indicated in red, while DNA oligonucleotide probes are indicated in black (more details in table S5). pTet-*NRD1* denotes strains expressing *NRD1* from a Tet-Off promoter. Depletion of Nrd1 in those strains was achieved by incubation with the tetracycline analogue doxycycline for 10.5h. This system was employed instead of the Nrd1-AID system because RT species are only detectable in a $\Delta rrp6$ background and the Nrd1-AID, $\Delta rrp6$ strain is not viable even in the absence of IAA. The 5.8S rRNA is used as a loading control. Note that Rrp6 is responsible for the processing of the 5.8S rRNA and thus, 5.8S precursors are detected in the $\Delta rrp6$ background. The asterisk in panel H indicates that the signal of mature tH(GUG)G2 corresponds to the same samples loaded in the blot in panel G.

We obtained very specific and reproducible RNAPIII occupancy signals in the crosslinked samples relative to un-crosslinked controls, with most reads mapping at RNAPIII-dependent genes (figure S1A-C). Consistent with a former genome-wide study (Turowski et al., 2016), our metagene analyses revealed significant RNAPIII signals downstream of the first T-tract after the 3' end of tRNA genes, (hereafter referred to as the primary terminator), indicating that termination at this sequence element is only partially efficient *in vivo* (figure 2A-C). Importantly, we observed a clear increase in the RNAPIII signal downstream of the primary terminator in the *sen1-3* mutant, indicating that the interaction with Sen1 promotes termination of RNAPIII, either at the primary terminator or downstream of it. Read-through (RT) transcription was also increased in the Sen1-AID strain even under non-depletion conditions, most likely because the presence of the tag affects the amount or the function of Sen1 even in the absence of auxin as observed for other proteins. Transcriptional read-through was further exacerbated when Sen1 was depleted by the addition of the auxin analogue 3-indoleacetic acid (IAA). The stronger effect of Sen1 depletion relative to the *sen1-3* mutation might imply either that Sen1-3 can still interact weakly with RNAPIII *in vivo*, or

that Sen1 functions in RNAPIII termination to some extent in the absence of interaction with the polymerase. Nevertheless, because full depletion of Sen1 also affects termination of many RNAPII non-coding RNA genes, we focused on the more specific *sen1-3* mutant for the rest of our study.

Heatmap analyses of the RNAPIII differential signal (\log_2 ratio) in the *sen1-3* mutant relative to the WT showed that an increase in the signal downstream of the primary terminator could be observed for the vast majority of tRNA genes (figure 2D). Furthermore, inspection of other RNAPIII-dependent genes such as the 5S and U6 genes revealed similar transcription termination defects, indicating that the role of Sen1 in favouring RNAPIII transcription termination is not restricted to tRNA genes (figure 2E-F).

Taken together, our results indicate that Sen1 is globally required for fully efficient termination of RNAPIII transcription *in vivo* and that this Sen1 function relies to a large extent on its interaction with RNAPIII.

Sen1 functions in RNAPIII transcription independently of the NNS-complex

Nrd1 and Nab3 have been found to bind the precursors of several tRNAs *in vivo* (Wlotzka et al., 2011), and it remain possible that these proteins also partake in RNAPIII termination although they did not appear significantly associated with RNAPIII in our MS analyses (Figure 1E). To address this possibility we conducted RNAPIII CRAC experiments in a Nrd1-AID strain. Depletion of Nrd1 upon treatment with IAA for 1h was sufficiently efficient to provoke clear termination defects at two well-characterized non-coding genes transcribed by RNAPII (i.e. *NEL025c* and *SNR13*, see figure S2C). However, neither the metagene analyses of RNAPIII distribution around tRNAs (figure 3A) nor the inspection of individual RNAPIII-dependent genes (figure 3B-C) revealed any significant effect on RNAPIII transcription termination efficiency. We conclude that, unlike Sen1, Nrd1 is not required for efficient termination of RNAPIII transcription. Because Nab3 is not known to function separately from Nrd1, our results indicate that Sen1 plays a role in RNAPIII transcription independently from the NNS-complex.

The function of Sen1 in RNAPIII transcription termination is not mediated by the replisome

Our analyses of Sen1 and RNAPIII protein interaction network support a model whereby Sen1 interacts with RNAPIII and the replisome in a mutually exclusive manner. However, they do not exclude the possibility that the replisome mediates the loading of Sen1 onto RNAPIII, for instance when a collision between these complexes occurs (e.g. Sen1 could interact

sequentially with the replisome and RNAPIII). RNAPIII transcription units are indeed hot spots of conflicts between the transcription and the replication machineries (Osmundson et al., 2017). Therefore, we considered the possibility that Sen1 might only function in RNAPIII transcription termination in the presence of ongoing replication. To explore this possibility, we performed parallel RNAPIII CRAC experiments in asynchronous cells and in cells arrested in the G1 phase by treatment with α -factor, in a WT and a *sen1-3* background (G1-arrest was verified by FACS analysis, figure S1D-E). Importantly, we observed a very similar RNAPIII pattern in G1-arrested and asynchronously growing cells (figure 2A-C and 3D-E), namely prominent RNAPIII termination defects in *sen1-3*.

The finding that abolishing the interaction between Sen1 and RNAPIII reduces the efficiency of termination even in the absence of the replisome (i.e. G1-arrested cells) indicates that Sen1 plays a role in termination of RNAPIII transcription independently of its association with the replisome.

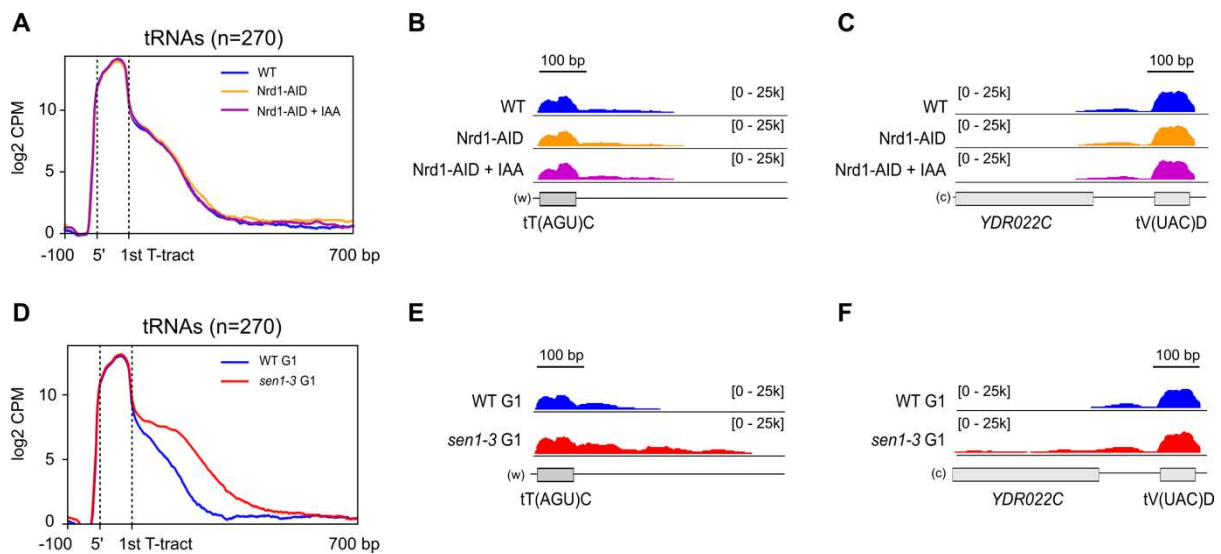


Figure 3: The function of Sen1 in RNAPIII transcription termination does not rely on Sen1 interaction with its partners Nrd1 and Nab3 or with the replisome.

A) Metagenome analysis of the RNAPIII distribution around tRNA genes as in figure 2 in a WT or in a Nrd1 Auxin-Inducible Degron (AID) strain in the absence or in the presence of IAA. Additional experiments validating the efficiency of Nrd1 depletion on well-characterized NNS target RNAs are included in figure S2.

B) and C) Individual examples of tRNA genes that exhibit clear termination defects in a *sen1-3* mutant (see figure 2B-C) but not in Nrd1-depleted conditions.

D) Metagenome analysis as in figure 2A but in cells blocked in the G1 phase of the cell cycle.

E) and F) Individual examples of tRNA genes that display termination defects in the *sen1-3* mutant in cells blocked in G1 as well as in asynchronous cells (compare with figure 2B-C).

Sen1 operates in a fail-safe transcription termination pathway

Our genome-wide data indicate that the association of Sen1 with RNAPIII globally increases the efficiency of transcription termination. However, these results are consistent with both a function for Sen1 in promoting termination at the primary terminator and/or a role in removing polymerases that constitutively escape primary termination.

To distinguish between these possibilities, we first analysed the distribution of the RNAPIII CRAC signal in WT and *sen1-3* cells. The total transcription levels, inferred from the RNAPIII signal within the gene body, were virtually identical in WT and *sen1-3* cells, indicating that the mutations in Sen1-3 do not impact transcription initiation or elongation (figure S3A-B).

We then computed for each tRNA gene both the RT index (i.e. the ratio of the RNAPIII signal downstream versus upstream of the primary terminator) and the RT length (i.e. the distance between the primary terminator and the 3' end of the RT signal) in the WT and in *sen1-3* (figure 4A). For most genes, we observed an increase in the RT index in *sen1-3* cells compared to WT cells (figure 4B-E), which is compatible with Sen1 functioning in primary or in secondary termination, since failure in either one of these processes alone would result in the accumulation of RNAPIII within RT regions. However, the heatmap analyses shown in figure 2D revealed that for most tRNA genes, very little or no RNAPIII accumulation could be observed immediately after the primary terminator in the mutant, with the largest increase of RNAPIII signal occurring further downstream, arguing against a major role for Sen1 at the primary termination site. Consistent with this notion, we observed a clear increase in the RT length in the mutant (figure 4B-E), indicating that polymerases that have escaped primary termination transcribe for longer because downstream termination is defective in the Sen1 mutant.

Because termination defects would lead to the production of different RNA species from tRNA genes depending on whether they occur at the primary terminator or at read-through regions, we set out to analyse these RNAs by northern blot (figure 2G-H). Mature tRNAs are generated by termination at the primary terminator and eventually by the processing of short 5' and 3'-extensions. Therefore, defects in primary termination are expected to result in lower amounts of mature tRNAs with a concomitant increase in the amount of RT transcripts. We could only detect RT RNAs for the tRNA genes tK(UUU)O and tH(GUG)G2 in the absence of the exosome-associated exonuclease *RRP6* (figure 2G-H), consistent with former data indicating that RT species are degraded by the RNA exosome (Turowski et al., 2016). In the case of tG(GCC)F2, simultaneous deletion of *RRP6* and depletion of the tRNase Z endonuclease Trz1, involved in the processing of tRNA precursors (Skowronek et al., 2014), was required for the strong detection of RT transcripts (figure S2A), indicating

that RT transcripts can also be targeted by Trz1.

Importantly, in all these cases we did not observe a significant decrease in the abundance of mature tRNAs in *sen1-3*, not even upon depletion of Trz1, excluding the possibility that RT transcripts are recognized as tRNA precursors by this endonuclease and cleaved to generate mature tRNAs (figure S2A and data not shown). Accordingly, the overall abundance of RT RNAs was similar in the WT and in *sen1-3*, but these species were globally longer in *sen1-3* cells, confirming CRAC data suggesting that they result from defective Sen1-dependent termination occurring downstream of tRNA primary terminators (figures 2G-H and S2A). This increase in size was not observed, as expected, when the NNS subunit Nrd1 was depleted, consistent with the Nrd1-AID RNAPIII CRAC data (figures 2G and 3A-C).

To further support the notion that Sen1 functions mainly on RNAPIIIIs that have escaped the primary termination site, we performed more detailed analyses of our CRAC data. If Sen1 does not function in primary termination, its failure to interact with RNAPIII should affect similarly genes with weak or strong primary terminators. Based on *in vitro* data, the minimal length for a functional terminator is 5 Ts (Arimbasseri and Maraia, 2015; Mishra and Maraia, 2019) but 6 Ts are required for relatively efficient termination and it is generally assumed that the termination efficiency is higher as the T-tract length increases. In partial agreement with these notions, we observed that i) the first T-tract rarely contains 4 Ts, ii) 6 Ts and 7 Ts are the most frequent terminators at this position and iii) tracts longer than 8 Ts are rarely found as primary terminators (figure 4F). We analysed the RT index of tRNAs clustered according to the length of their primary terminator and, as expected, we found that the RT index in these clusters tends to decrease as the T-tract length increases (figure 4G) in inverse correlation with the termination efficiency. Importantly, in *sen1-3* cells the RT index increases similarly for all clusters suggesting that having an inefficient primary terminator does not make termination more sensitive to Sen1, arguing against a role of Sen1 at these sites.

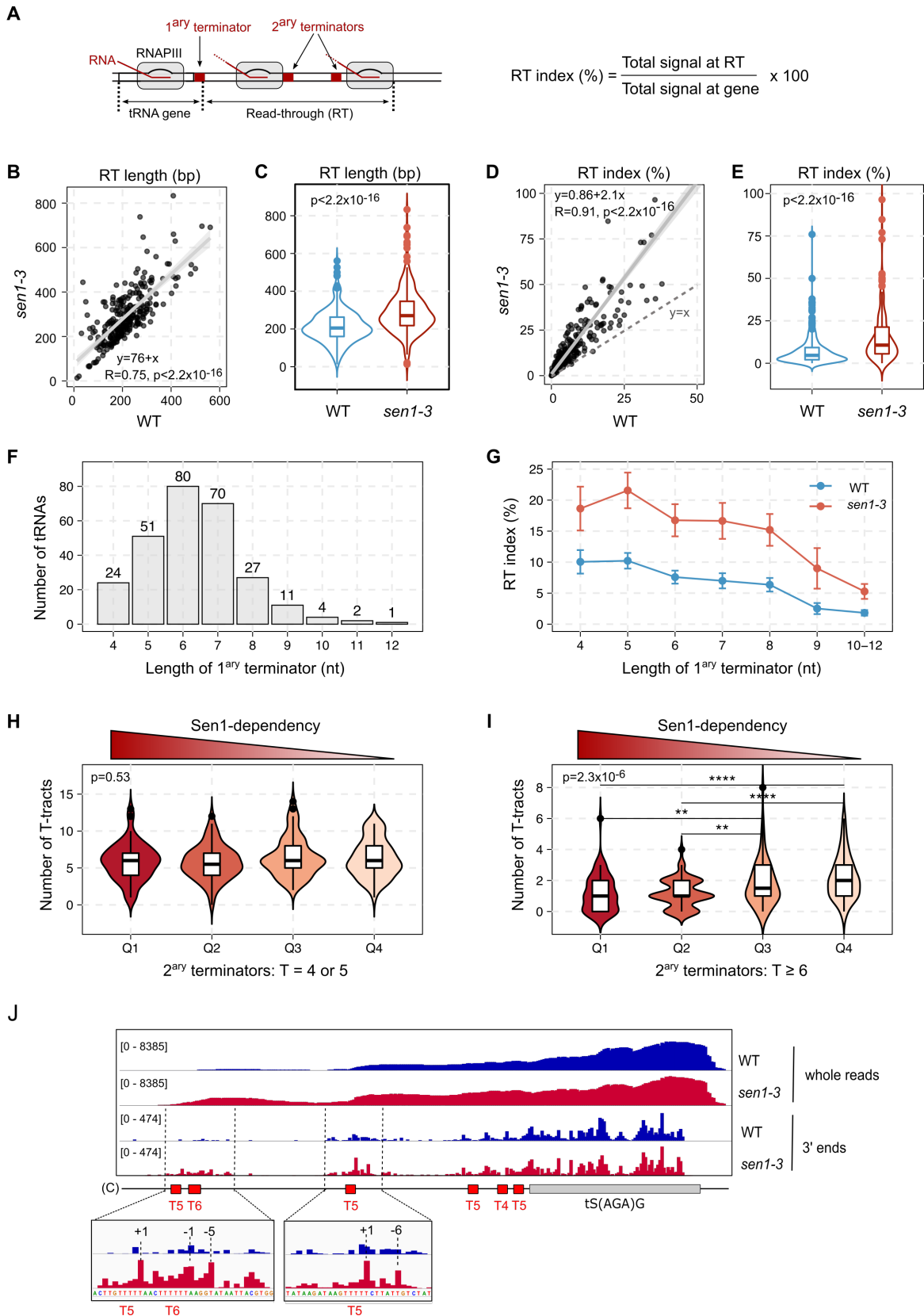


Figure 4: Sen1 functions mainly on secondary termination.

A) Scheme of tRNA transcription units indicating the relevant elements and parameters used for the assessment of the transcription termination efficiency in the WT and the *sen1-3* mutant.

B) and C) Comparison the RT length for the different tRNA genes in the mutant relative to the WT. B) Correlation plot. The grey zone corresponds to the confidence interval whereas R is the Pearson's correlation coefficient. p is the p-value associated with Pearson's correlation.

C) Violin plot showing the distribution of RT lengths in the WT and in *sen1-3*. The p-value (p) was calculated with the Wilcoxon test.

D) and E) Comparison of the RT index measured in the indicated strains for each tRNA gene. D) Correlation plot generated as in B). Three outliers in *sen1-3* are not shown. E) Violin plots as in C) but with RT index values. Three outliers in *sen1-3* are excluded. Note that both RT length and the RT index are inversely proportional to the termination efficiency (e.g. higher RT index indicates lower termination efficiency).

F) Histogram representing the number of tRNA genes that possess a primary terminator of each indicated length. Only consecutive thymidines are considered when computing the length of the primary terminator.

G) Analysis of the RT index of tRNA genes grouped according to the length of their primary terminator in either the WT or the *sen1-3* mutant. Data points correspond to the average value whereas error bars denote the standard error.

H) and I) Analysis of the number of either “weak” (H) or “strong” (I) terminators located at the 700 bp region downstream of the primary terminator for tRNA genes grouped according to the extent of termination defects in the *sen1-3* mutant (i.e. dependency on Sen1 for efficient transcription termination). Groups correspond to quartiles (Q) defined by the tRNA gene ranking obtained in the heatmap analyses in figure 2D, where Q1 includes the 25% of genes with the highest impairment in transcription termination in the *sen1-3* mutant. p corresponds to the p-value for the global comparison of the four groups according to the Kruskal-Wallis test. Asterisks denote the p-values of pairwise comparisons (*: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$; ****: $p \leq 0.0001$).

J) IGV screenshot of an individual tRNA showing the distribution of RNAPIII CRAC signal in the WT and the *sen1-3* mutant. The 3' ends datasets provide the position of individual RNAPIII with single-nucleotide resolution. Insets are zoom in views of the main regions where RNAPIII accumulates in the mutant. Coordinates in insets correspond to the position relative to the beginning of the nearest downstream T-tract.

The region downstream of tRNA genes contains T-stretches that were previously proposed to play a role as secondary termination sites (Turowski et al., 2016). We considered the possibility that Sen1 might be preferentially required for tRNA genes having a lower number of secondary termination sites or less efficient ones. To address this possibility, we ranked the different tRNAs according to the extent of the RNAPIII accumulation in *sen1-3* relative to WT cells, thus defining a hierarchy of Sen1-dependency. For each tRNA, we computed the number of weak (4 or 5 Ts) or strong (≥ 6 Ts) terminators in regions of secondary termination and we compared the average number of terminators of each kind in the different quartiles. Interestingly, we found that the tRNA genes that are more dependent on Sen1 for termination (i.e. Q1) tend to have a lower number of efficient terminators compared to those that are less dependent (i.e. Q3 and Q4). In contrast, the number of weak terminators, which have a lower impact on RNAPIII progression, was similar in all groups of tRNAs (figures 4H-I and S3).

These results strongly suggest that Sen1 compensates for the lack of efficient terminators in regions of secondary termination. This could imply that Sen1 improves termination at weak terminators or that it promotes termination at other sequences. A careful analysis of the RNAPIII CRAC signal at individual tRNA genes provided evidence supporting both possibilities (figure 4J and S4). Mapping only the 3' end of the nascent RNA allows obtaining a precise readout of RNAPIII position with single-nucleotide resolution. We observed very little if any effect of the *sen1-3* mutation at positions around the primary terminator, while RNAPIII was clearly found to accumulate preferentially around T-tracts, but also at other sequences, in the downstream regions. Together our results support the notion that Sen1 does not play a prominent role in primary termination and rather promotes the release of RNAPIII that pause within regions of secondary termination.

Sen1 can promote termination of RNAPIII transcription *in vitro*

We have previously demonstrated that Sen1 can directly promote termination of RNAPII transcription in a sequence-independent manner (Porrua and Libri, 2013b). To assess whether Sen1 can also directly induce RNAPIII transcription termination and whether it requires the presence of canonical termination signals, we employed an *in vitro* transcription termination system containing purified proteins (i.e. RNAPIII and full-length Sen1), transcription templates and nascent RNA (figures 5A and B).

We first analysed the capacity of canonical terminator sequences to induce RNAPIII transcription termination by comparing the behaviour of RNAPIII on transcription templates containing T-tracts of variable lengths (i.e. from 4 to 12 Ts, see figures 5C-D and S5). Consistent with former data (Arimbasseri and Maraia, 2015; Mishra and Maraia, 2019), we

observed only very weak polymerase pausing at the T4 terminator and no detectable RNAPIII release. The T5 terminator induced stronger pausing and intermediate levels of RNAPIII release, while the T6 terminator promoted very efficient release. Stretches of 9 or 12 Ts induced very strong RNAPIII pausing as virtually no transcription signal could be detected downstream of these terminators but a substantial proportion of RNAPIIIs remained associated with the proximal part of these long T-tracts (~ 50% for the T9 and ~ 80% for the T12). This might be due to the recognition of the distal portion of the T-tract in the downstream DNA by RNAPIII, which might induce strong pausing and disfavour release (see Discussion). Indeed, for shorter T-tracts we also observed several prominent pausing sites a few nt upstream of these sequences, both *in vitro* (figure 5C) and *in vivo* (figures 4J and S5) supporting the idea that RNAPIII can sense downstream untranscribed T-tracts.

Importantly, the presence of Sen1 in the reaction provoked a substantial increase in the levels of transcription termination at the T4 terminator, and, to a lesser extent, at the T5 terminator, while no significant effect was observed for the more efficient T6 terminator. This result indicates that Sen1 can enhance RNAPIII release at weak terminators.

Interestingly, we found that Sen1 could also promote the release of RNAPIIIs that are paused at the proximal part of long T-tracts, especially in the case of the T12 terminator, for which roughly 50% of paused RNAPIIIs were released by Sen1. Finally, and importantly, we also observed Sen1-dependent release of RNAPIIIs that are paused at sequences other than T-tracts, for instance, at pausing sites upstream of the canonical terminators, which corroborates our *in vivo* RNAPIII CRAC analyses (figure 4J and S4). Taken together, these results indicate that Sen1 can both enhance termination at inefficient terminators and promote termination at unrelated sequences.

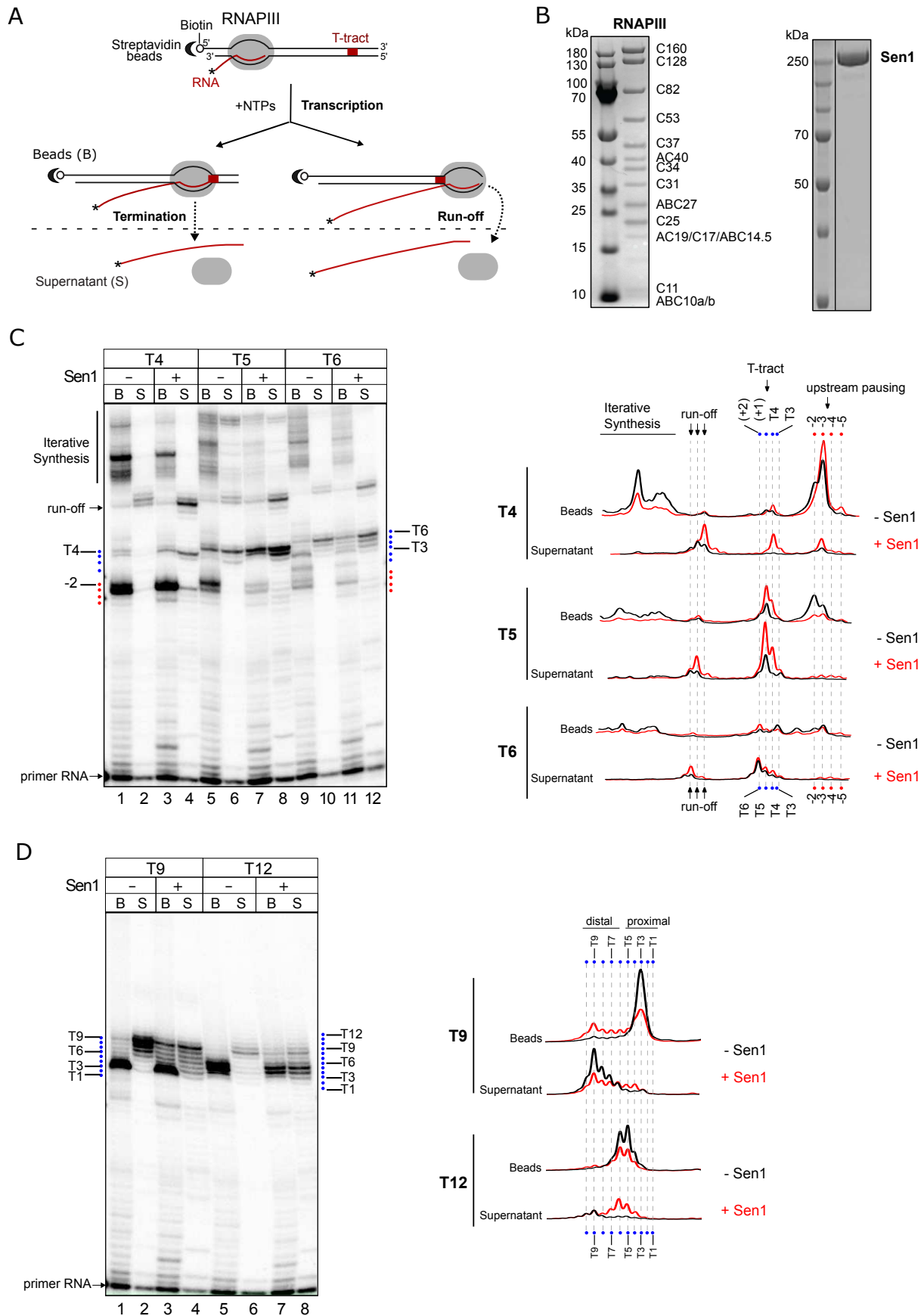


Figure 5: Sen1 can induce termination of RNAPIII transcription *in vitro*.

A) Scheme of an *in vitro* transcription termination (IVTT) assay. Ternary elongation complexes (ECs) composed of purified RNAPIII, nascent RNA, and DNA transcription templates are assembled by step-wise addition of the different components (see methods) and associated with streptavidin beads via the 5' biotin of the non-template strand to allow subsequent separation of beads-associated (B) and supernatant (S) fractions. The RNA is radioactively labeled at the 5' end to allow its detection (indicated by an asterisk). Each transcription template contains a T-tract of a particular length on the non-template strand. After addition of nucleotides, RNAPIII transcribes and can pause at different positions, including the T-tract. RNAPIIIs that pause at a T-tract can either dissociate from the DNA to the supernatant (i.e. undergo transcription termination) or remain paused, and thus associated with the beads, or resume transcription. Polymerases that read-through the T-tract and reach the end of the template can either run-off, with concomitant release of full-length transcripts into the supernatant, or perform iterative synthesis. For transcripts associated with paused RNAPIIIs, the comparison of the fraction that is retained in the beads with the fraction that is released provides an estimate of the efficiency of termination at each site.

B) SDS-PAGE analyses of the protein preparations used in IVTT assays.

C) Analyses performed on templates containing T-tracts composed of 4 (T4), 5 (T5) or 6 (T6) consecutive Ts. Left: Denaturing PAGE analysis of transcripts from an IVTT assay testing the capacity of Sen1 and T-tracts of different lengths to induce RNAPIII transcription termination. "B" corresponds to the beads fraction while "S" denotes the supernatant. Representative gel of one out of three independent experiments. Right: Profile of the signal over the region of interest for each gel lane.

D) Analysis of IVTT reactions performed on templates containing stretches of 9 (T9) or 12 (T12) Ts. These reactions were performed in parallel with those in panel C) but migrated on different gels. Left: Representative gel of one out of three independent experiments. Right: Profile of the signal over the region of interest for each gel lane.

The position of the nucleotides of interest was determined by migrating in parallel a radioactively-labelled ladder (not shown).

Sen1 employs a similar mechanism to terminate transcription of RNAPII and RNAPIII

According to previous studies, canonical terminators contain signals that induce both RNAPIII pausing and release from the DNA (reviewed in Arimbasseri et al., 2013 and Porrua et al., 2016). The above results indicate that Sen1 requires polymerase pausing but not necessarily the presence of a T-tract for terminating RNAPIII. To further explore this idea, we performed *in vitro* transcription assays with modified templates containing a G-less cassette followed by a run of Gs to force the stalling of RNAPIII at the G-stretch in the absence of GTP (figure 6A). In these conditions, and similarly to what was observed for RNAPII, Sen1 could induce

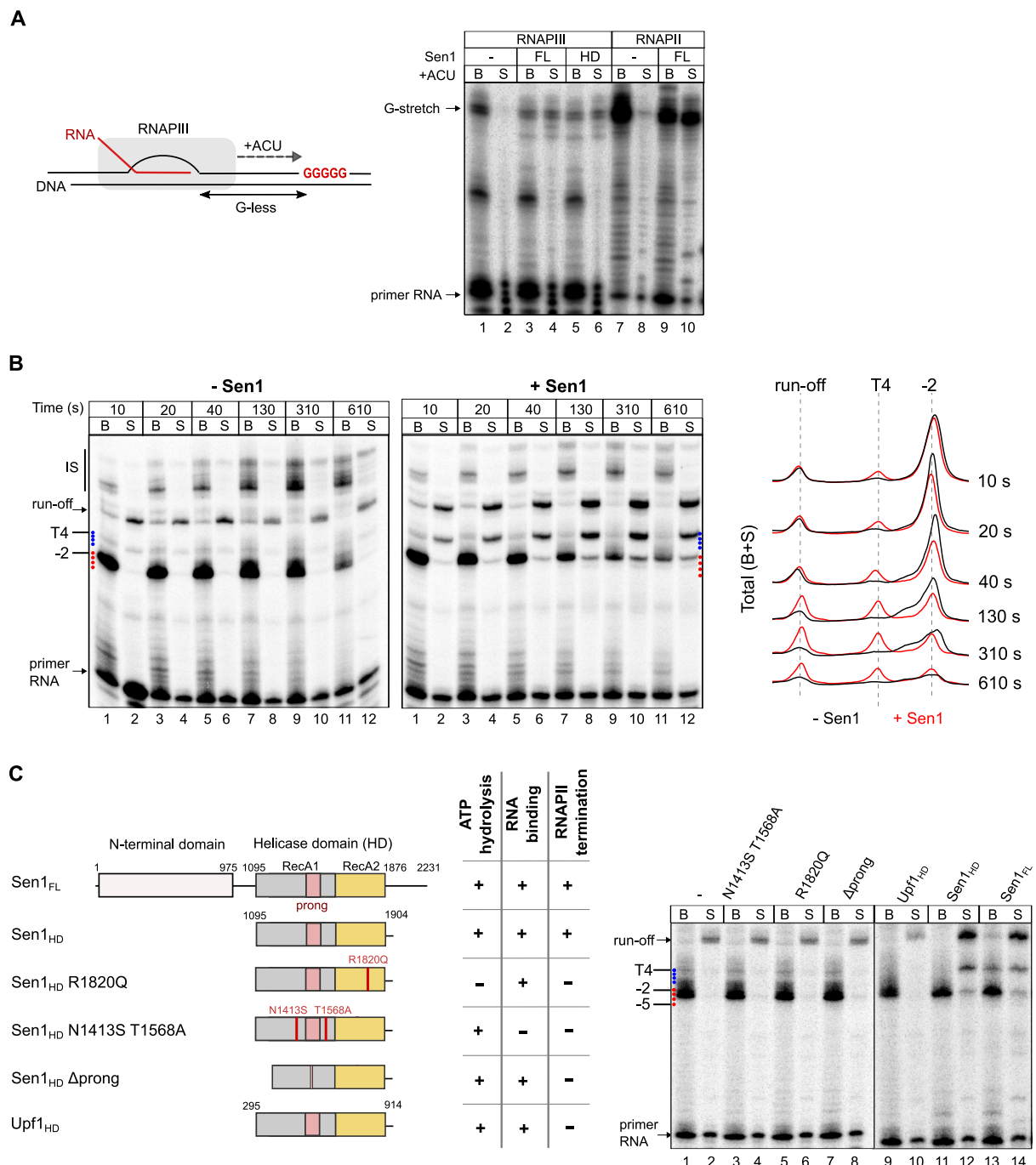


Figure 6: Analysis of the mechanisms of Sen1-mediated termination of RNAPIII transcription.

A) Analysis of the capacity of Sen1 to promote the release of RNAPIII paused at a sequence other than a T-tract. The transcription templates contain a G-less cassette followed by a stretch of Gs so that in the absence of G in the nucleotide mixture RNAPs are stalled at the first G. Experiments were performed in parallel with purified RNAPII as a positive control for Sen1 activity. The efficiency of transcription termination observed for RNAPII is similar as in our former studies (Han et al., 2017; Leonaitė et al., 2017).

B) Time-course IVTT assay performed on transcription templates containing a T4 terminator. All the reactions were performed in parallel but were migrated on different gels. Left: Representative gel of one out of two independent experiments. Right: Profile of the total signal (beads and supernatant) over the region of interest.

C) Analysis of the role of different protein regions and activities of Sen1 in RNAPIII transcription termination. Left: Scheme of the different proteins used in IVTT assays and summary of the relevant phenotypes. The different variants of Sen1 helicase domain (HD) were purified and characterized within the frame of a previous study (Leonaitė et al., 2017). A constitutively active version of the Upf1 HD that possess helicase activity but cannot induce RNAPII transcription termination *in vitro* is used as a negative control (see Porrua and Libri, 2013). The Δ prong version of Sen1 HD contains a deletion of amino acids 1461-1554, which corresponds to most of this subdomain. Right: Representative gel of one out of two independent experiments. All the reactions were performed in parallel but were migrated on different gels.

the release of roughly 50% of paused RNAPIII, demonstrating that it can terminate transcription at pausing sites other than T-tracts.

We next set out to investigate the mechanism by which Sen1 induces RNAPIII transcription termination. We have previously shown that in order to dissociate the RNAPII elongation complex, Sen1 needs to load on the nascent RNA and translocate towards the RNAPII using the energy of ATP hydrolysis. Upon colliding with RNAPII, Sen1 can induce forward translocation of the polymerase, which, in the appropriate context, results in its release from the DNA (Han et al., 2017). Our *in vitro* assays support a similar mechanism for RNAPIII release (figure 5C), with evidence for “pushing” the elongation complex at sites of pausing. This is for instance manifest at the T5 terminator (figure 5C, compare lanes 5-6 to 7-8) where a decrease in the pausing signal at position -2 is not due to release at this position but rather by forward translocation and release at a downstream site. This is best illustrated in a time-course experiment performed with the T4 template, in which we quantified the total signal in the presence and absence of Sen1 (figure 6B, right). If a decrease in the signal from a paused polymerase is due to its release, the total signal (i.e. beads + supernatant) at that position should not change. On the contrary, if polymerases are “pushed” by Sen1 and eventually released at a later stage the signal distribution should be shifted downward. Indeed, upon addition of Sen1 we observe such a signal shift as well as the accumulation of RNA signal over time at positions where Sen1 induces its release. These findings support the notion that Sen1 promotes both RNAPIII translocation and its release from the DNA, similarly to what we previously showed for RNAPII.

To further explore the mechanisms of RNAPIII termination by Sen1, we first assessed whether the interaction of Sen1 with RNAPIII, mediated by its N-terminal domain, is required for the actual step of polymerase release. To this end we first analysed the capacity of the helicase domain of Sen1 alone to induce termination *in vitro*. We have previously shown that this domain is sufficient for inducing termination of RNAPII transcription (Han et al, 2017, Leonaite et al, 2017). Strikingly, we found that the helicase domain of Sen1 could induce termination of RNAPIII transcription *in vitro* as efficiently as the full-length protein (figure 6C), suggesting that the association of Sen1 with RNAPIII via its N-terminal domain is not a strict requirement for termination but might rather play a role in the recruitment of Sen1 to RNAPIII *in vivo*. As a negative control, we assessed a catalytically active version of the closely-related helicase Upf1 (Chakrabarti et al., 2011), which could not provoke termination of RNAPIII transcription, indicating that termination is not induced unspecifically by any active helicase but rather requires specific Sen1 activities or features (figure 6C). Finally, we analysed several mutant variants of the Sen1 helicase domain that are deficient for RNA binding (N1413S T1568A) or ATP hydrolysis (R1820Q), or a mutant that retains the catalytic activity but lacks the “prong” domain (i.e. Δ 1461-1554), which is essential for viability and for RNAPII transcription termination (Leonaitė et al., 2017). Importantly, none of these mutants could promote RNAPIII transcription termination *in vitro*, indicating that Sen1 employs the same structural features and activities to induce transcription termination of RNAPII and RNAPIII.

RNA structures upstream of T-tracts can promote the release of paused RNAPIII

The above results indicate that, akin to the RNAPII system, Sen1-mediated termination of RNAPIII transcription involves Sen1 translocation along the nascent transcript, and our former structural and biochemical data showed that Sen1 can only interact with single-stranded RNA (Porrua and Libri, 2013b, Leonaite et al, 2017). tRNAs are highly structured RNA molecules and for a vast majority of them (i.e. 251 out of 270 tRNAs) the spacer between the 3' end of the mature tRNA and the primary terminator is at most 7 nt. We envisioned that a possible reason for which Sen1 does not function at sites of primary termination is that its binding to the nascent RNA is hindered by the co-transcriptional formation of stable structures in the vicinity of the primary terminator. Conversely, less structured RNAs in the read-through region would allow Sen1 loading and function.

To explore these possibilities we performed *in vitro* transcription assays with modified transcription templates containing a natural hairpin from the 5S RNA, an RNAPIII-dependent transcript, upstream of T-tracts of different lengths. As a control we used a mutated hairpin with substitutions in the left arm preventing stem formation (figure 7A-C and S5). Surprisingly, the presence of a hairpin in the transcribed RNA could significantly stimulate transcription

termination at a T4 terminator, similarly to the addition of Sen1 to the unstructured version of the same RNA (figure 7A).

This result indicates that not only Sen1 but also RNA secondary structures can improve the function of weak terminators. In agreement with this idea, the presence of the hairpin did not enhance termination at the T6 terminator, since this sequence already supports full release of paused RNAPIII (figure 7B). However, the RNA structure could induce the release of polymerases paused at the proximal part of the T12 terminator, even more efficiently than Sen1 (figure 7C). These observations support the notion that, similarly to Sen1, RNA hairpins have the capacity to promote RNAPIII release.

Because RNA structures naturally form close to the primary terminator of RNAPIII-dependent genes, we next assessed to what extent secondary structures need to be in proximity to T-stretches to function in termination. To this end, we compared the efficiency of termination on templates containing a T4 or a T12 terminator when the hairpin was located immediately upstream (figure 7A-C), 7 nt or 18 nt upstream of the corresponding T-tract (figure 7D-F). We observed a clear enhancement of RNAPIII release at both T4- and T12-containing templates when the hairpin was located immediately upstream or 7 nt away from the T-tract but not in the presence of a 18 nt spacer. These results indicate that RNA structures can enhance transcription termination only when they are in close proximity to T-tracts. The results with the 18 nt spacer provided a tool to address the impact of secondary structures of the RNA on the function of Sen1 in termination. Indeed, in this case, formation of a hairpin 18 nt upstream of the T-tract only allows exposing a segment of roughly 7 nt of single strand RNA outside of the RNAPIII, which is not sufficient for the loading of Sen1 (Leonaite et al, 2017). Importantly, Sen1 could not release RNAPIII at the T4 terminator in this construct, likely because of the insufficient single stranded RNA span between the structure and the polymerase. We observed, however, an increase in the amount of released run-off transcripts in the presence of Sen1, indicating that at further downstream positions Sen1 can load on the nascent RNA and promote the release of polymerases at the end of the template (figure 7E, lanes 5-8).

Taken together, our results strongly suggest that RNA secondary structures forming in the vicinity of weak primary terminators can markedly improve their function. However, they can also hamper the recruitment of Sen1 to the nascent RNA and, thus, would likely prevent Sen1 from functioning at primary terminators, regardless of their strength.

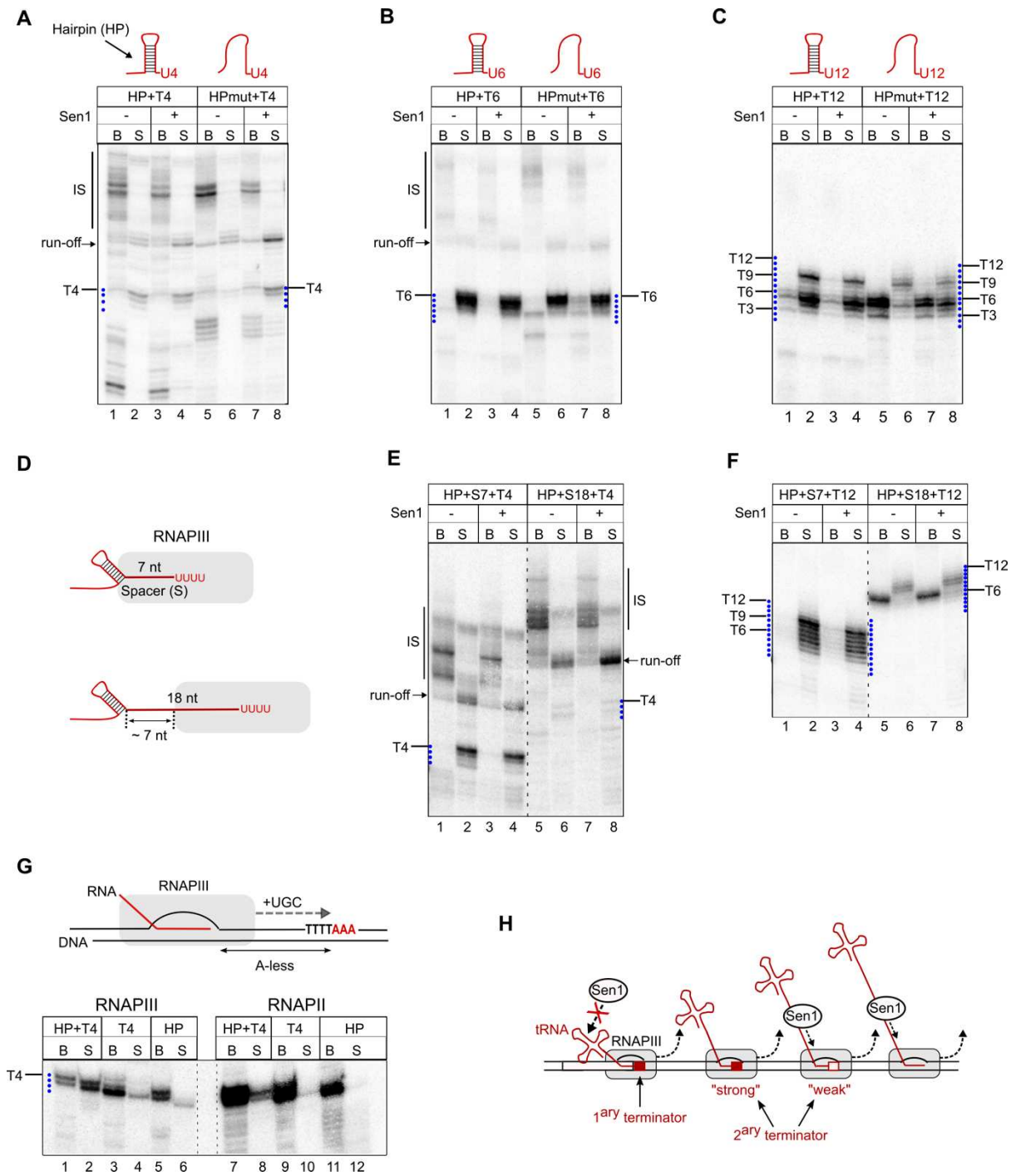


Figure 7: Hairpin-like structures forming in the nascent RNA can complement the function of canonical termination signals.

A-C) Analysis of the role of RNA structures in transcription termination at a T4 (A), T6 (B) or a T12 (C) terminator. IVTT assays as in figure 5 but with modified templates to introduce a hairpin (HP) in the transcribed RNA immediately upstream of the T-tract. The control template (HPmut) harbors several mutations at the left arm that disrupt hairpin formation (see figure S4 for sequence details and structure prediction of the resulting RNAs).

D-F) Analysis of the impact of the positioning of the hairpin relative to the T-tract on its capacity to stimulate RNAPIII release. Experiments performed as in A) and C) but with modified templates to introduce a spacer (S) of the indicated lengths between the hairpin and the T-tract (see figure S4 for details). D) Scheme showing the predicted position of the hairpin relative to RNAPIII in the presence of the indicated spacers. IVTT assays with templates containing either a T4 terminator (E) or a T12 terminator (F).

G) Functional dissection of a transcription terminator composed of an RNA hairpin and a stretch of 4 Ts (T4). IVTT assays performed as in A) but with modified templates (see figure S4 for details) to include an A-less cassette followed by a stretch of As so that, in the absence of A in the nucleotide mixture, RNAPs are stalled at the first A. In the HP template the T-tract is mutated to CTCT. Experiments were performed in parallel with purified RNAPII to compare the sensitivity of both RNAPs to termination signals.

H) Model for the role of canonical termination signals, RNA structures and Sen1 in termination of RNAPIII transcription (shown for tRNA genes). At the primary terminator, termination typically involves the action of a T-tract and the secondary structure of the nascent tRNA. RNA structures are required only for T-tracts of sub-optimal length. In downstream regions transcription by RNAPIII is typically terminated either by “strong” secondary terminators, without the aid of Sen1, or by “weak” termination signals if Sen1 can access and load onto the nascent RNA. Sen1 can in principle also promote termination at sites of pausing other than T-tracts.

RNA secondary structures can form within RNAPIII

A previously proposed model (Nielsen et al., 2013) posits that T-tracts are sufficient for RNAPIII pausing but not for its release from the DNA, for which RNA secondary structures would be strictly required. This model opposes the most widely accepted one, which points to an exclusive role for the T-tract in termination. Our data indicate that secondary structures can promote RNAPIII release but only at defective terminators. We decided to further investigate the mechanism of action of RNA structures in termination. The model of Nielsen et al. postulates that one of the main functions of T-tracts would be to promote RNAPIII backtracking to bring the polymerase in contact with the nearest upstream structure, which would invade the RNA exit channel of RNAPIII, thus destabilizing the elongation complex. Nonetheless, we have observed that the RNA hairpin is functional when located immediately upstream or very close to a T4 or a T12 sequence, which implies either that i) RNAPIII transcribes beyond the T-tract to allow formation of the hairpin, pauses at downstream sequences and undergoes subsequent backtracking; or that ii) the RNA folds at least partially

within the polymerase to induce its release. The former possibility appears to be hardly compatible with the case of the T12 terminator, for which we did not find any evidence of RNAPIII transcribing through the terminator (figure 5D).

To address these possibilities, we conducted *in vitro* transcription assays with modified templates where the RNA hairpin is encoded in an A-less cassette and is followed by a T4 sequence and three As (figure 7G and S5). By performing reactions with these templates in the absence of ATP, polymerases cannot transcribe through the terminator and stall at the fourth T of the T4. In these conditions, the hairpin cannot form outside of the polymerase, its downstream arm still being within the polymerase.

Interestingly, we observed that stalled RNAPIIIIs were released in the presence of the hairpin but not the corresponding mutant version, indicating that transcription through the terminator is not required for the folding of the hairpin, which must occur in the RNA exit channel of RNAPIII. Importantly, very little, if any, polymerase release was observed when the T4 sequence was mutated, even in the presence of the hairpin, indicating that the hairpin is not an autonomous termination signal and can only function together with a canonical termination sequence. Finally, the concomitant presence of an RNA hairpin and a T-tract induced poor release of RNAPIIs, indicating that these nucleic acid elements function specifically as RNAPIII termination signals.

These findings strongly support the notion that the RNA can fold within the RNAPIII and that backtracking is not required to promote RNAPIII termination. Together, our data comfort the notion that RNA secondary structures are not absolutely required for RNAPIII termination, but can nevertheless function as auxiliary elements that work in concert with weak or defective termination signals.

Discussion

RNAPIII synthesises short ncRNAs like tRNAs and the 5S rRNA that are absolutely essential for mRNA translation and, therefore, for cell growth and survival. Timely termination of RNAPIII transcription is critical not only for the correct synthesis of these RNAs but also for preventing RNAPIII from invading neighbouring genomic regions and interfering with the function of other machineries that operate in these regions. This is even more relevant considering the high expression levels of RNAPIII-dependent genes.

The traditional model for termination of RNAPIII transcription posits that termination exclusively relies on the presence of T-tracts at the 3' end of RNAPIII-dependent genes that are specifically recognized by the polymerase as termination signals. The implication of additional *cis*-acting factors, such as RNA secondary structures has been previously proposed (Nielsen et al., 2013) but has remained hitherto controversial (Arimbasseri et al., 2014).

Here we show that the mechanisms governing RNAPIII transcription termination *in vivo* are considerably more complex than those represented in former models, involving the interplay between distinct *cis*-acting elements and the extrinsic termination factor Sen1. We propose an integrated model whereby T-tracts and RNA secondary structures function in concert at primary terminators (and possibly other sites) while Sen1 concurs to release polymerases that have escaped “intrinsic” termination.

S. cerevisiae Sen1 is a fail-safe transcription termination factor for RNAPIII

One of the important findings of this study is the demonstration that Sen1 can directly promote termination of RNAPIII transcription, a conclusion that is supported both by genome-wide data and by *in vitro* transcription termination assays with purified components. However, multiple lines of evidence support the notion that Sen1 functions to remove polymerases that have escaped primary termination at the very 3'-end of RNAPIII-dependent genes. These include the high-resolution detection of RNAPIII occupancy by CRAC as well as the analysis of the different RNA species produced from model tRNA genes (figures 2, 4, S2 and S4). A mechanistic explanation for our observation that Sen1 cannot operate at primary terminators is provided by our finding that *in vitro* Sen1 function in termination is hindered by RNA secondary structures, which are typically present in RNAPIII transcripts close to the first terminator.

A recent study has reported that one of the two *Schizosaccharomyces pombe* homologues of Sen1, *Sp* Sen1, also interacts with RNAPIII and its deletion leads to global defects in RNAPIII transcription termination (Rivosecchi et al., 2019). However, some

important conclusions of this study are substantially different from the ones supported by our results. It was shown that deletion of *Sp* Sen1 leads to a global downstream shift of the RNAPIII occupancy peak at tRNA genes, as determined by CHIP-seq, and a reduction in the levels of mature tRNAs, which we did not observe in *S. cerevisiae*. These findings have been interpreted in support of a model whereby efficient primary termination in *S. pombe* relies on *Sp* Sen1 and would be only partially dependent on intrinsic termination signals. In contrast, in *S. cerevisiae* primary termination mainly depends on *cis*-signals (T-tracts and secondary structures) and Sen1 operates in downstream regions to remove read-through polymerases. Therefore, in *S. cerevisiae* Sen1 rather plays an important genome safeguarding role in preventing inappropriate extension of RNAPIII transcription.

The divergency between these models might be due to the different resolution of the techniques employed in the two studies (e.g. CRAC and CHIP-seq) but might also reflect mechanistic differences between two organisms. For instance, differences in the biochemical properties of the two Sen1 proteins or in the mode they are loaded onto the nascent transcript might be at stake. Indeed, substantial sequence homology between the two proteins can be found only in their helicase domains and, contrary to *S. cerevisiae* Sen1, none of the *S. pombe* Sen1 homologues is essential for viability, interacts with RNAPII or partakes in RNAPII transcription termination (Larochelle et al., 2018; Rivosecchi et al., 2019).

Importantly, despite the possible mechanistic differences, the fact that the two Sen1 orthologues play a role on RNAPIII transcription opens up the possibility that this function is conserved in other organisms.

The mechanism of Sen1-dependent RNAPIII transcription termination

Although its best-characterized function is the termination of non-coding transcription by RNAPII within the NNS-complex, *S. cerevisiae* Sen1 is also implicated in other processes like the control of R-loop formation, the resolution of transcription-replication conflicts and DNA-repair (Alzu et al., 2012; Appanah et al., 2020; Li et al., 2016; Mischo et al., 2011). The N-terminal domain of Sen1 is an important hub for protein-protein interactions that might modulate these different functions of Sen1 (Appanah et al., 2020; Han et al., 2020). The function of Sen1 in RNAPIII termination *in vivo* strongly depends on the interaction with RNAPIII, which is mediated by a region in the Sen1 N-terminal domain containing the amino acids mutated in Sen1-3 (W773, E774 and W777). This region is not required for termination *in vitro* indicating that it is not a critical molecular determinant of the process of RNAPIII release, and therefore we suggest it drives the recruitment of Sen1 to RNAPIII, which might be a limiting step for termination *in vivo*.

Mutation of the same amino acids in Sen1-3 also prevents the interaction with the replisome components Ctf4 and Mrc1 (Appanah et al., 2020), yet we show that the interactions of Sen1 with the replisome and RNAPIII are not interdependent but rather mutually exclusive. This suggests either that the same surface mediates the interaction with RNAPIII and the replication fork, or that these mutations alter the conformation of two distinct regions of interaction. The observation that Sen1 promotes RNAPIII transcription termination even in the absence of replication (i.e. in G1) indicates that the action of Sen1 is not restricted to situations of transcription-replication conflicts. However, we cannot exclude that such conflicts might trigger Sen1-dependent termination in some circumstances.

Our *in vitro* data strongly support the notion that Sen1 terminates RNAPIII transcription essentially by the same mechanism employed to induce RNAPII release, for which we have previously provided a detailed molecular model (Han et al., 2017; Leonaitė et al., 2017; Porrua and Libri, 2013b). We have shown that Sen1 translocates along the nascent RNA and induces a forward motion of paused RNAPII that results in its release from the DNA. The helicase domain of Sen1 retains all the properties that are necessary for transcription termination and we proposed that a particular subdomain protruding from the helicase core (the "prong") enters the RNA exit channel provoking destabilizing conformational changes in the elongation complex (Leonaitė et al., 2017; discussed in Han and Porrua, 2017). We show here that the helicase domain is also sufficient for RNAPIII transcription termination and that the essential activities involved in translocation (RNA binding and ATP hydrolysis) as well as the "prong" are required. Akin to what was shown for RNAPII, Sen1 "pushes" paused RNAPIII, which either promotes elongation resumption or results in its release from the DNA (figure 6). Whether the outcome of "pushing" (elongation or termination) is determined by alternative, pre-existing conformations of paused RNAPIII or it is stochastic remains to be determined

A former study reported transcription termination defects at RNAPI-dependent genes upon inactivation of Sen1 *in vivo* (Kawauchi et al., 2008). The interpretation of these data is blurred by the fact that Sen1 inhibition can have multiple indirect effects due to its widespread role in termination of RNAPII transcription. However, we have found that Sen1 associates with RNAPI *in vivo* (table 1) and can also promote the release of paused RNAPI *in vitro* (figure S6). Therefore, altogether these data could point at a common mechanism of transcription termination operating at the three eukaryotic RNAPs and relying on the helicase Sen1.

RNA structures are enhancers of canonical termination signals

The two fundamental steps in RNAPIII transcription termination are RNAPIII pausing and

release from the DNA. The most widely accepted model (Arimbasseri et al., 2013), posits that a stretch of Ts in the non-template DNA strand is sufficient for both pausing and release of RNAPIII. An alternative model was proposed by Nielsen and co-authors (Nielsen et al., 2013), according to which, while T-tracts can promote RNAPIII pausing, an RNA hairpin in the vicinity of the paused RNAPIII is the main determinant for the dissociation of the polymerase from the DNA. These fundamental disparities were attributed to differences in the purity of the RNAPIII preparations employed in the studies supporting these models (Arimbasseri et al., 2014; Nielsen and Zenkin, 2014). Here we use a high-purity preparation of the RNAPIII holoenzyme validated in structural and functional analyses (Hoffmann et al., 2015; Vorländer et al., 2018) to investigate the different mechanisms involved in RNAPIII transcription termination.

We find that the capacity of T-tracts to promote RNAPIII pausing is directly linked to the T-tract length, with T4 terminators supporting very little pausing and $T \geq 9$ terminators inducing a complete block of RNAPIII elongation (figure 5C-D). Our results show that T6 terminators, which are the most frequently found *in vivo* (figure 4F) are not fully efficient in supporting pausing but can induce RNAPIII release in the absence of any adjacent RNA structure (figure 5C and 7B), indicating that RNA secondary structures are not always required for termination. In contrast, in the case of T4 terminators, which are essentially non-functional *in vitro*, we find that an adjacent RNA secondary structure can convert these sequences into moderately efficient terminators (figure 7A). This behaviour likely phenocopies the situation *in vivo* where the tRNA acceptor stem typically folds very close to the primary terminator and might explain why T4 terminators can be found as primary terminators (figure 4F). Although some tracts of 4 Ts are separated by only 1-2 nt from downstream T-tracts and, thus, could be part of longer interrupted termination signals, more isolated T4 terminators appear to function independently (figure S7) and likely in concert with native RNA secondary structures.

Strikingly, in our assays, very long T-tracts ($T \geq 9$) are defective in promoting RNAPIII release, and these defects are more pronounced as the length of the T-tract increases. More precisely, we observe that a fraction of RNAPIII stall at the proximal portion of these long T-tracts after "reading" only the first 3-6 nt of the T-tract and fail to dissociate from the DNA. Our interpretation of these observations is that RNAPIII can recognize the T-tract in the downstream duplex region, either because of its sequence or because of the particular structure T-tracts impose to the DNA helix (Stefl et al., 2004). Such interactions would stabilize the EC, thus compensating for the destabilizing effect of the weak rU-dA hybrid and the interaction with the unpaired thymidines in the non-template strand within the transcription bubble. Interestingly, we find that an RNA hairpin forming in the vicinity of these long T-tracts

can promote full release of stalled RNAPIII (figure 7C), suggesting that *in vivo* long T-tracts might require the concomitant presence of an adjacent secondary structure to be fully proficient in transcription termination.

We provide mechanistic evidence that RNA hairpins can form within the RNA exit channel of RNAPIII (figure 7G) to promote termination. Consistent with this finding, a recent structural study has provided evidence that an RNA hairpin can fold within the RNA exit channel of a bacterial polymerase, leading to a rearrangement of the EC (Kang et al., 2018). Interestingly, structural comparisons indicate that eukaryotic polymerases can also accommodate such RNA secondary structures within their RNA exit channels (Kang et al., 2018). Indeed, a very recent structural study on human RNAPII has provided evidence for the presence of double-stranded RNA in the RNA exit channel (Misiaszek et al., <https://doi.org/10.1101/2021.05.31.446457>). Therefore, we propose that, in the case of RNAPIII, the formation of an RNA hairpin can induce destabilizing conformational changes in the RNAPIII that would contribute to the dissociation of the EC. Importantly, unlike Nielsen and co-authors (Nielsen et al., 2013), we find that an RNA hairpin can only promote efficient release of paused RNAPIII when a T-tract resides in the polymerase main channel (figure 7G). While this work was in progress, a study using a reporter system in human cells provided evidence that an RNA hairpin located close to a short T-tract (T4) can enhance RNAPIII transcription termination *in cellulo* (Verosloff et al., 2021), pointing to an evolutionarily conserved role for RNA structures in termination.

Taken together, our results allow proposing a revisited model for autonomous RNAPIII transcription termination that can partially reconcile former contradictory findings. According to our model, T-tracts are strictly required for termination, but adjacent RNA structures are important auxiliary elements when the length of the T-tract falls outside of the optimal range. Thus, the protein-independent mechanism of termination of RNAPIII transcription has more commonalities with the so-called intrinsic termination pathway for bacterial RNAP than previously appreciated.

Multiple mechanisms partake in RNAPIII transcription termination

We and others have observed that RNAPIII read through the primary terminator quite frequently and termination at downstream regions was proposed to rely on secondary canonical terminators (Turowski et al., 2016). Indeed, tRNA read-through regions contain T-tracts that are more frequent in the sense orientation than in the antisense orientation (figure S8A-B), suggesting they are under positive selection. However, long T-tracts (T>5) are scarce in these regions (figure 4I), suggesting that alternative evolutionary routes have been

undertaken for ensuring efficient termination.

We have shown that both RNA structures and the helicase Sen1 can complement the function of short termination signals. These two factors act in a mutually exclusive manner because: i) both employ a similar mechanism likely involving a conformational change initiated at the level of the RNA exit channel of RNAPIII; and ii) the presence of secondary structures in the nascent transcript prevents Sen1 loading. Our data support the idea that Sen1 would play a more prominent role in termination at read-through regions than RNA structures. A possible explanation is that Sen1 can function both at weak terminators and at other pausing sites, while RNA structures can only work when located sufficiently close to a T-tract.

We have observed that the transcripts encoded in the ~250 bp region immediately downstream of the primary terminator have a lower propensity to fold into secondary structures than the genomic average (figure S8C-E). While this could be partially due to the higher frequency of T-tracts in this region, which lowers the GC content, it might also be a consequence of Sen1 involvement in fail-safe termination. We suggest that "repurposing" the RNAPII transcription termination factor Sen1 for terminating RNAPIII might have a lower evolutionary cost than generating the appropriate arrangements of T-tracts and RNA structures in tRNA read-through regions.

These considerations do not exclude the possibility that more than one mechanism operate in secondary termination for the same gene. This is for instance illustrated by the tH(GUG)G2 gene (figure S2B), where a secondary T8 terminator is present 60 bp downstream of the primary terminator. Termination at this site is independent of Sen1 most likely because a strong secondary structure forms immediately before T8. The fraction of RNAPIII that escape termination at this site terminates at downstream sites, in the apparent absence of strong secondary structures, in a Sen1-dependent manner.

Taken together, our findings reveal the existence of multiple mechanisms cooperating to promote termination of RNAPIII transcription. We propose that RNA structures contribute to the efficiency of primary termination in some instances (i.e. genes with suboptimal terminators), thanks to the natural proximity of the tRNA acceptor stem to the first T-tract, whereas Sen1 would preferentially function at downstream regions (figure 7H). Efficient termination is important for the rapid recycling of RNAPIII for new cycles of transcription and, thus, for maintaining robust expression of tRNAs and other RNAPIII-dependent transcripts that are essential to sustain cell proliferation. Also, it is crucial to prevent or to minimize the conflicts with other transcribing polymerases as well as with other DNA-associated machineries.

Supplementary material

List of supplementary material:

- Figures S1: Complementary analyses validating CRAC experiments in figures 2 and 3.
- Figures S2: Experiments related to figures 2 and 3.
- Figures S3: Complementary analyses related to figure 4.
- Figures S4: Examples of tRNA genes illustrating the role of Sen1 in enhancing secondary termination.
- Figures S5: Sequence of transcription templates and predicted structure of the different transcribed RNAs in *in vitro* transcription termination assays.
- Figures S6: Sen1 can promote the release of paused RNAPs *in vitro*.
- Figures S7: Examples of tRNA genes harbouring a T4 primary terminator.
- Figures S8: Analysis of T-tracts and RNA structures at regions of secondary termination.
- Table S6: Yeast strains used in this study.

Material provided as separate xls files and accessible by the following link:

<https://www.dropbox.com/sh/th60zut0jz3bvj8/AAAUV1Df5ub6Hli7qqdMten2a?dl=0>

- Table S1: mass spectrometry analyses of TAP-Sen1 and TAP-Sen1 Δ NTD coimmunoprecipitates.
- Table S2: mass spectrometry analyses of TAP-Sen1 NTD coimmunoprecipitates.
- Table S3: mass spectrometry analyses of Sen1-TAP and Sen1-3-TAP coimmunoprecipitates.
- Table S4: label-free quantitative mass spectrometry analyses of Rpc160-HTP coimmunoprecipitates in a WT and a *sen1-3* background.
- Table S5: list of oligonucleotides used in this study.
- Table S7: annotations of tRNA genes from the 5' end to the mature tRNA to the primary terminator in bed format.
- Table S8: annotations of potential secondary terminators of tRNA genes in bed format.
- Table S9: annotations of tRNA genes readthrough regions in a WT and a *sen1-3* mutant in bed format.

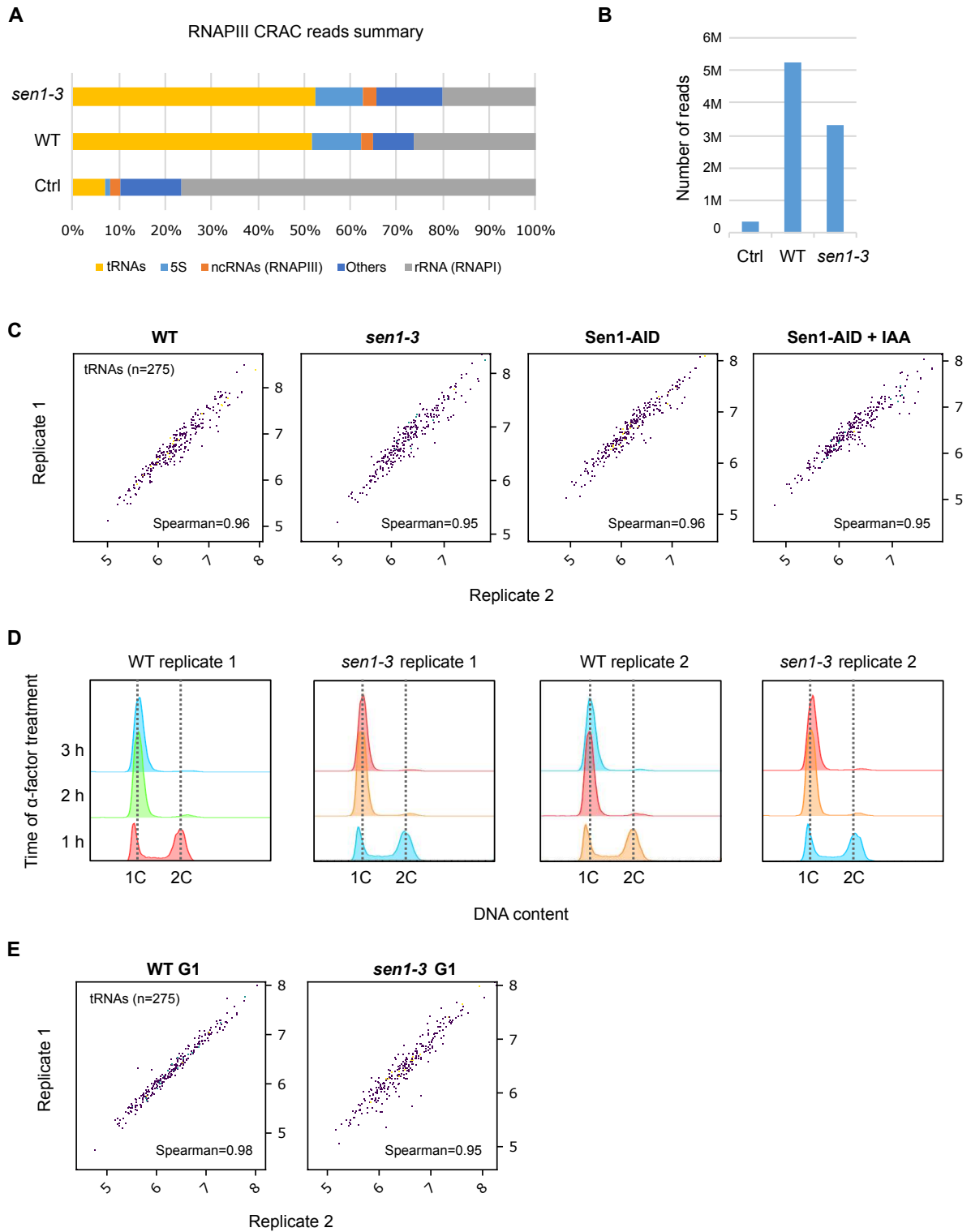


Figure S1: Complementary analyses validating CRAC experiments in figures 2 and 3.

A) Comparison of the reads distribution among different genomic regions in crosslinked samples (WT and *sen1-3*) relative to the un-crosslinked control (Ctrl) in a typical RNAPIII CRAC experiment. The "others" category corresponds to RNAPII genes and intergenic regions. Note that tRNA read-through regions are included in this category and the larger proportion of reads in this group in the *sen1-3* mutant could be due to the observed increased RNAPIII presence at those regions.

B) Plot representing the number of mapped reads obtained in a typical CRAC experiment in the different samples. Note that the number of reads in cross-linked samples is at least one order of magnitude higher than in the un-crosslinked control (Ctrl). "M" denotes millions.

C) Scatter plots showing the high correlation between the two biological replicates of each condition/strain for the CRAC experiments showed in figure 2.

D) Analysis of DNA copy-number for samples in figure 3D-F by flow cytometry. 1C and 2C corresponds to 1 and 2 copies of the genome, respectively. Cultures were used for CRAC analyses after 3h of treatment with α -factor (see methods).

E) Correlation plots for the two biological replicates of samples used in experiments in figure 3D-F.

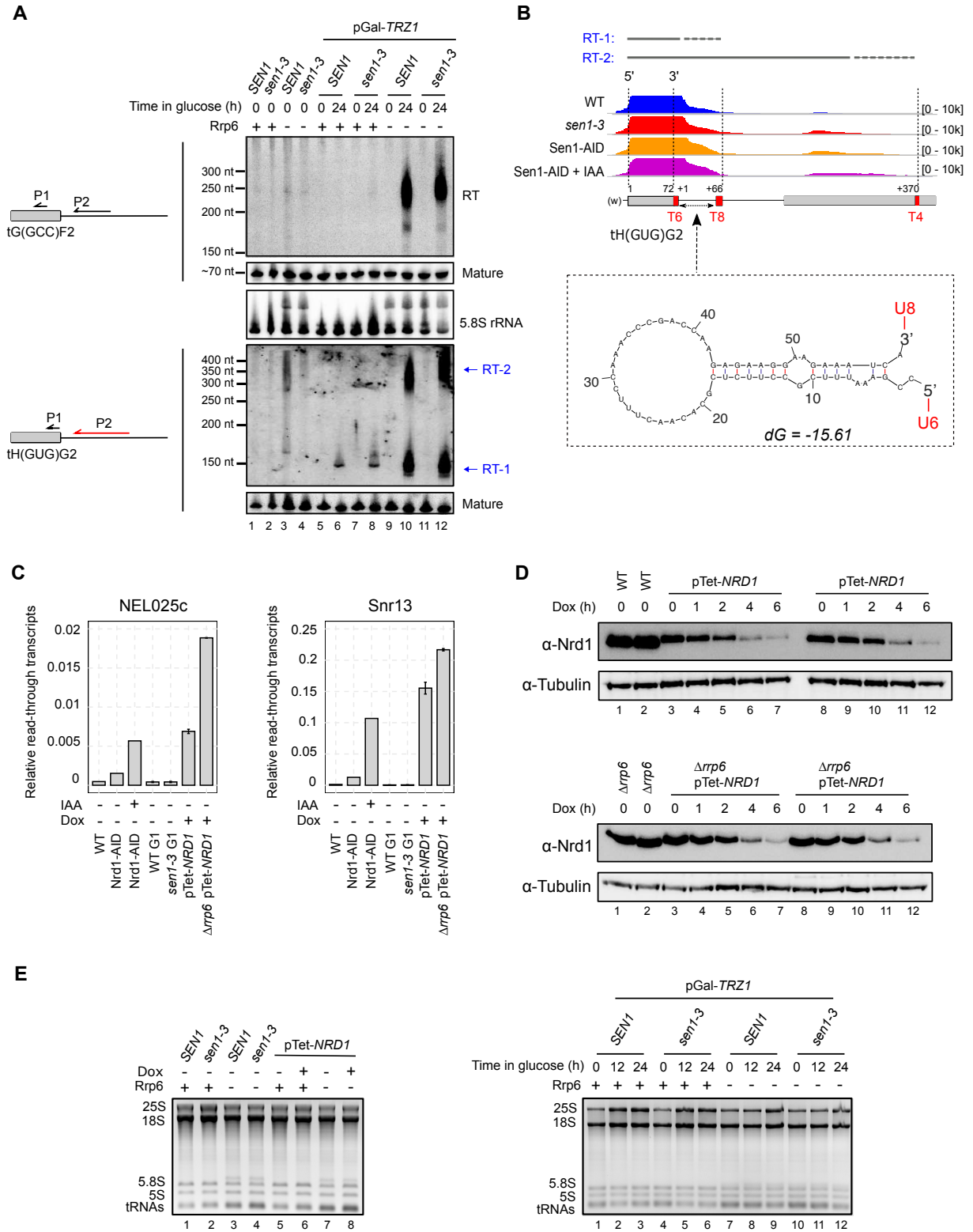


Figure S2: Experiments related to figures 2 and 3.

A) Northern blot analysis of transcripts derived from two different tRNA genes in the indicated backgrounds. In pGal-*TRZ1* strains, the essential gene *TRZ1* is expressed from the *GAL1* promoter and the different strains are either grown on galactose for the whole experiment (t=0) or shifted to medium containing glucose for 24h to repress *TRZ1*. Experiments performed with 12h-incubation in glucose-containing medium provided similar results (data not shown). Schemes on the left indicate the approximate position of the probes (P1 and P2) used for the detection of the different RNA species (RT, for read-through, and mature tRNA). The RNA probe is indicated in red, while DNA oligonucleotide probes are indicated in black (more details in table S5). The 5.8S rRNA is used as a loading control.

B) IGV screenshot of the region around the tH(GUG)G2 gene indicating the position of different T-tracts found and the two major groups of RT transcripts detected by northern blot (**A**, bottom blot). Note that the fact that we detect multiple bands for each termination region could be due to the existence of several termination sites and/or the presence of heterogeneous poly(A) tails. The structure of the RNA between the T6 and the T8 T-tract predicted by the mFold software of the UNAFold package (<http://www.unafold.org/>) is shown on the bottom.

C) Analysis of transcription termination defects at two well-characterized NNS-targets (the CUT NEL025c and the snoRNA gene *SNR13*) in the indicated strains. RNAs were prepared from the same cultures used for CRAC experiments and northern blot analyses in figures 2 and 3. Typical read-through transcripts resulting from inefficient termination by the NNS-complex were detected by RT-qPCR with oligonucleotides listed in table S5. Values are normalized relative to the levels of the *ACT1* mRNA.

D) Western blot analysis of Nrd1 depletion by incubation of pTet-*NRD1* strains (2 biological replicates) with doxycycline (Dox) for the indicated times. Tubulin detection was used as a loading control.

E) Native agarose gels showing total RNA levels in the indicated samples stained with ethidium bromide.

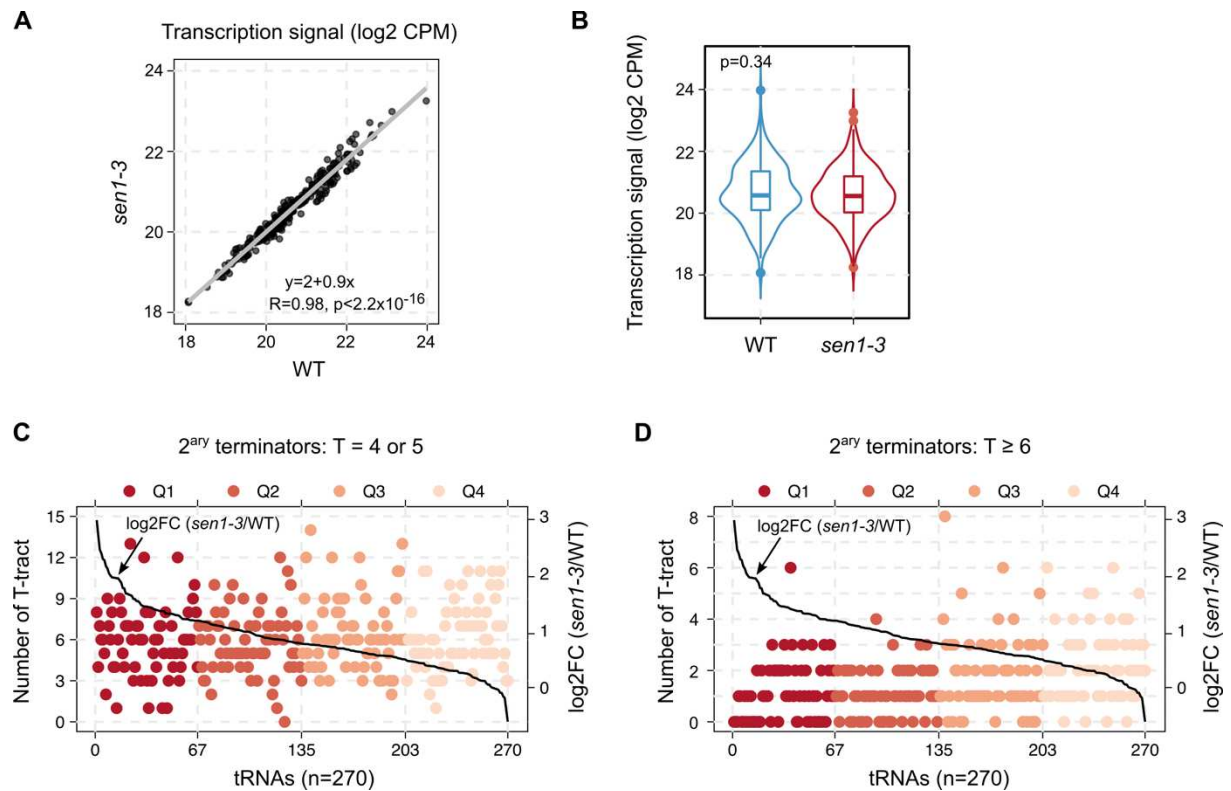


Figure S3: Complementary analyses related to figure 4.

A) and **B)** Comparison of the transcription signal at tRNA genes in the WT and in the *sen1-3* mutant measured by the total RNAPIII CRAC signal at the region from the 5' end of the mature tRNA to the first T-tract. **A)** Dispersion plot where “R” corresponds to Pearson’s correlation coefficient and p is the associated p-value. **B)** Violin plot where p corresponds to the p-value calculated by the Wilcoxon test.

C) and **D)** Representation of the number of T-tracts of indicated lengths located in the 700 bp region downstream of the primary terminator of each tRNA gene. Data points are coloured according to the quartile (Q) they belong to. Quartiles are defined according to the log₂ FC of the RNAPIII signal in the *sen1-3* relative to the WT at the same region, which provides an estimation of the dependency on Sen1 for termination. Thus, Q1 includes the most Sen1-dependent tRNA genes.

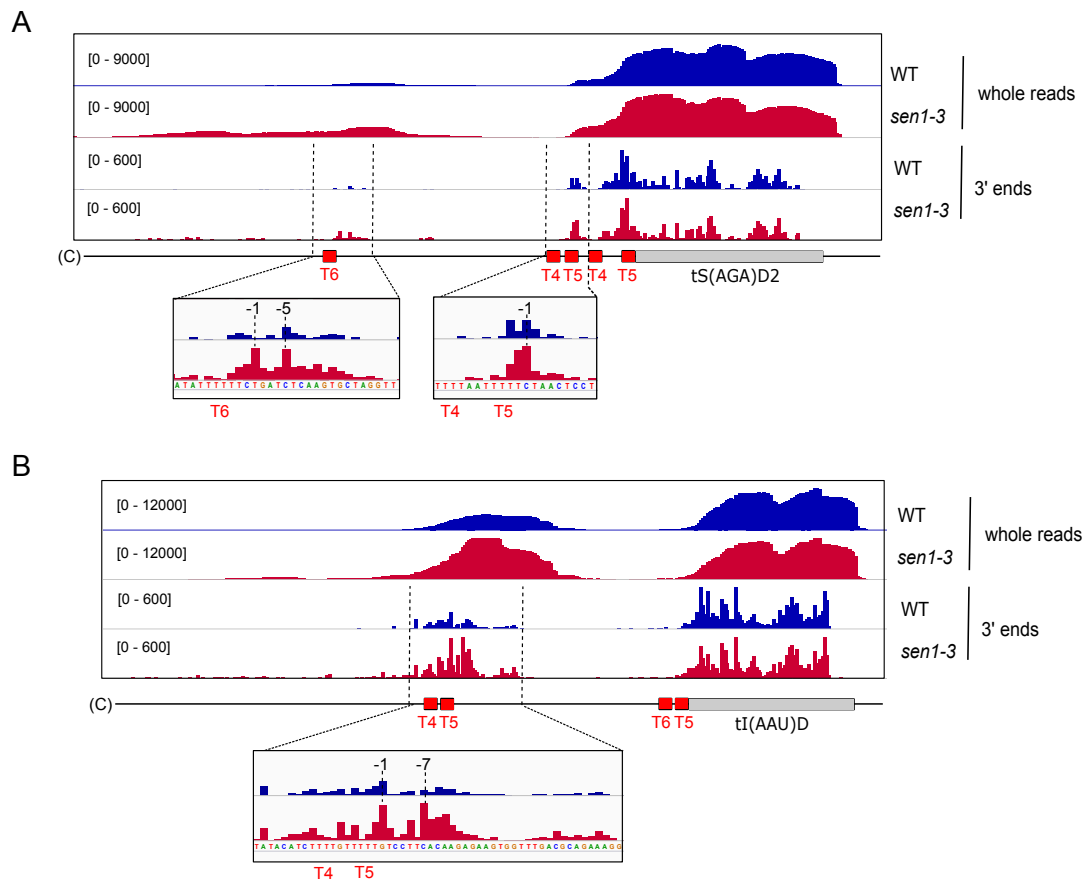
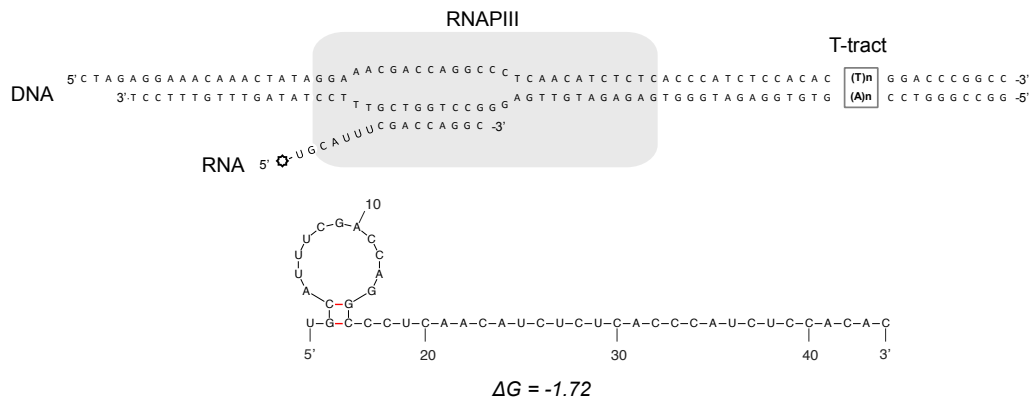
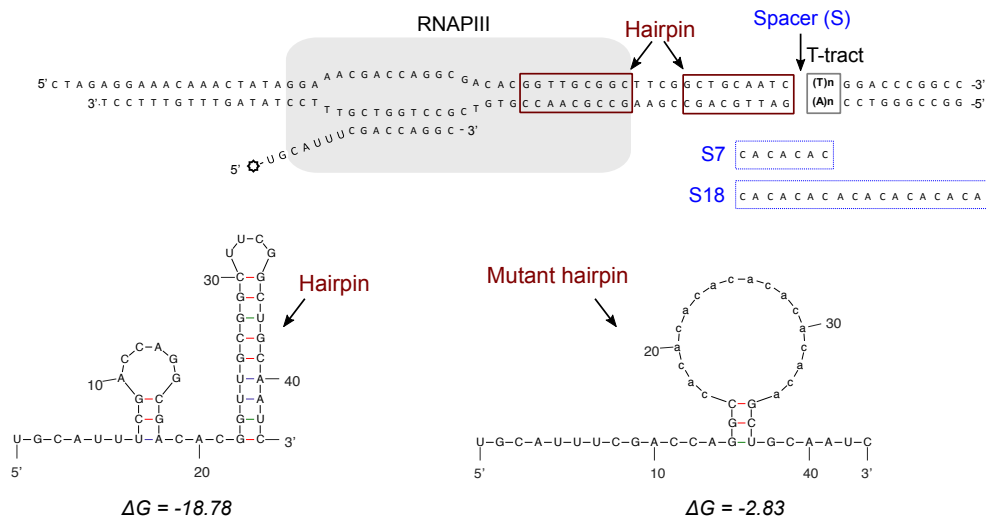


Figure S4: Examples of tRNA genes illustrating the role of Sen1 in enhancing secondary termination. IGV screenshots showing the distribution of RNAPIII CRAC signal in the WT and the *sen1-3* mutant with zoom in views of the main regions where RNAPIII accumulates in the mutant. The 3' ends datasets provide the position of individual RNAPIIIIs with single-nucleotide resolution. Most accumulation is observed just upstream of or at the first Ts of secondary weak terminators, suggesting impaired RNAPIII release by Sen1 at these sites. The indicated coordinates correspond to the position relative to the beginning of the nearest downstream T-tract.

1) No hairpin and T4, T5, T6, T9 or T12 terminators (figures 5, and 6)



2) Hairpin (HP) and T4, T6 or T12 terminators (figure 7E-F)



3) Hairpin (HP) - A-less cassette - T4 terminator - A-tract (figure 7E)

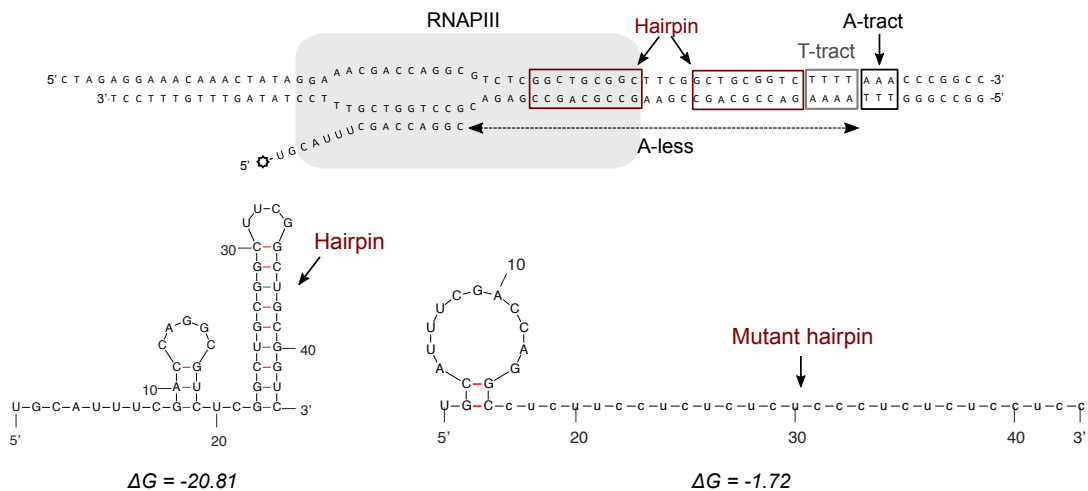


Figure S5: Sequence of transcription templates and predicted structure of the different transcribed RNAs in *in vitro* transcription termination assays.

The sequence of the wild-type version of each template are indicated in the schemes. The mutant version of the transcribed RNAs is shown together with the wild-type version under the corresponding scheme. The sequence of the spacers (S) correspond to the non-template strand. RNA structure predictions and ΔG calculation for each structure were performed with the mFold software.

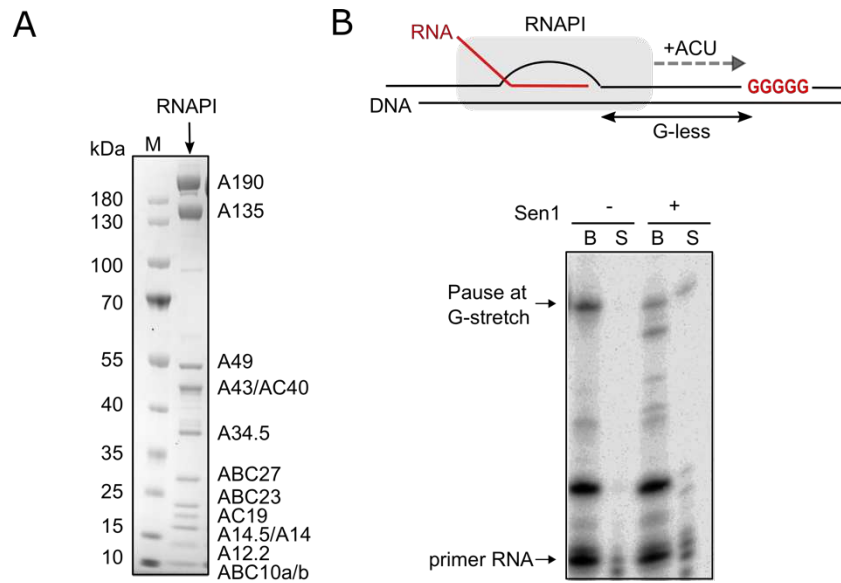


Figure S6: Sen1 can promote the release of paused RNAPIs *in vitro*.

A) SDS-PAGE analysis of the RNAPI preparation used in these assays.

B) IVTT assay performed on templates containing a G-less cassette followed by a run of Gs to promote stalling of RNAPI at the first G in the absence of guanine in the reaction. Top: scheme of the transcription templates. Bottom: Denaturing PAGE analysis of transcripts from one out of two independent IVTT assays, which produced very similar results.

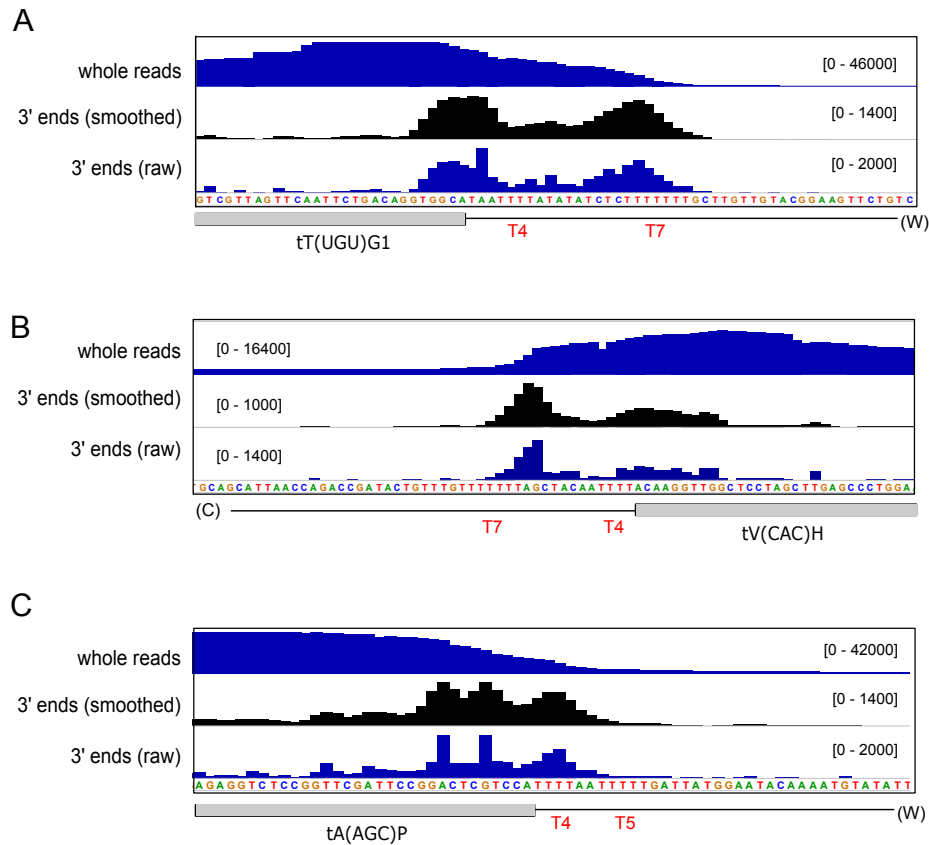


Figure S7: Examples of tRNA genes harbouring a T4 primary terminator.

A) and **B)** Cases where the T4 terminator seems to function autonomously to promote moderate levels of termination. The 3' end datasets provide accurately the position of individual RNAPIII, and therefore the location of pausing sites. The decrease in the whole reads RNAPIII signal within the T4 sequence despite the presence of very strong pausing at close downstream T-tracts supports the idea that a fraction of RNAPIII terminate at the T4 terminator.

C) Example where the T4 sequence seems to function in combination with the downstream T5 T-tract, as suggested by the strong termination observed at these sequences compared to A) and B), where the T4 sequence is further from other T-tracts. The pausing pattern, with accumulation of RNAPIII at the first 3 thymidines of the T4 sequence rather resembles the pattern observed for long T-tracts (e.g. T9 terminators) *in vitro*.

Datasets correspond to the WT strain.

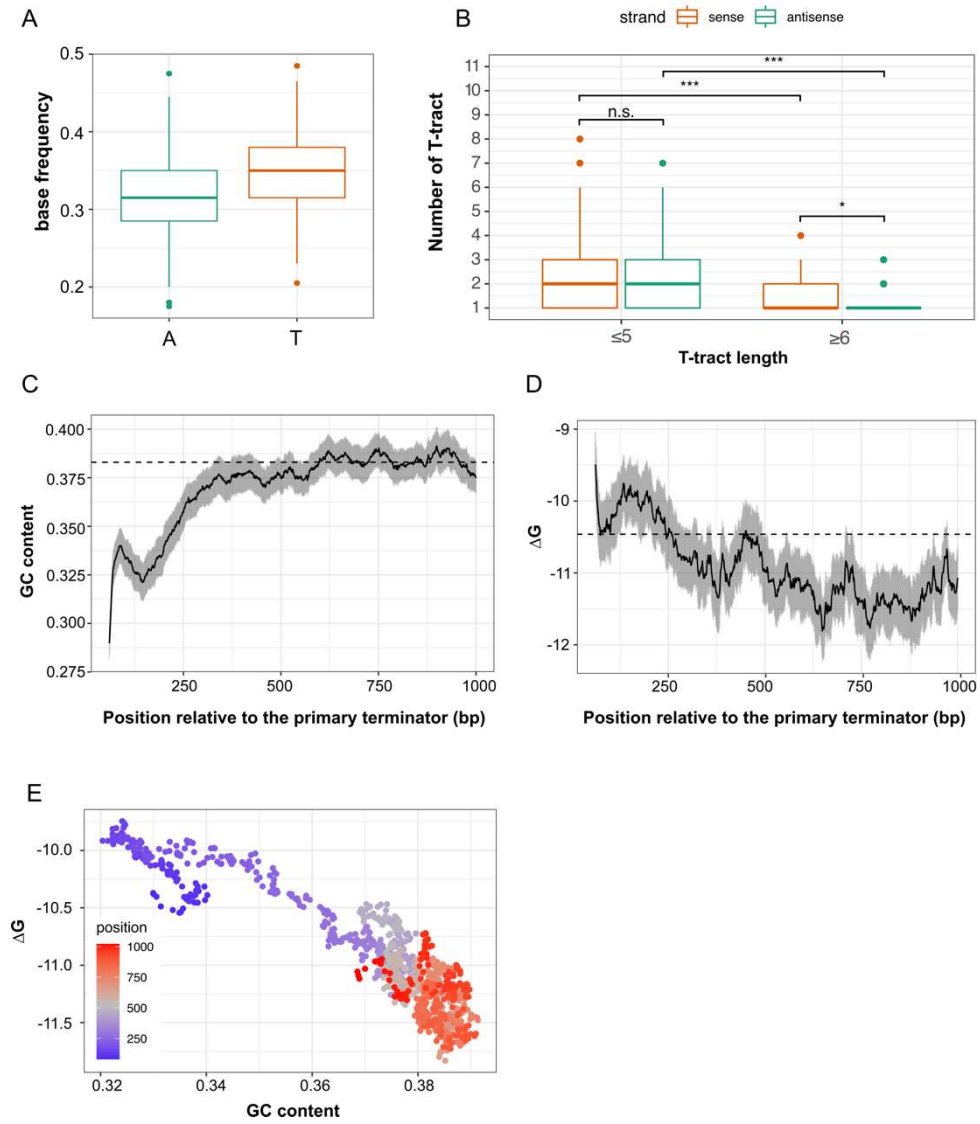


Figure S8: Analysis of T-tracts and RNA structures at regions of secondary termination.

A) Analysis of the frequency of A and T nucleotides in the 200 bp region downstream of the primary terminator of tRNA genes.

B) Comparison of the number of weak ($T \leq 5$) or strong ($T \geq 6$) terminators at the 200 bp region downstream of the primary terminator of tRNA genes in the sense orientation versus to the antisense orientation relative to transcription. Statistical significance was calculated using a Wilcoxon rank sum test. n.s indicates no significant difference between the compared groups whereas * denotes a p-value ≤ 0.05 and *** a p-value ≤ 0.001 .

C) Analysis of the GC content (fraction of G and C nucleotides) of regions downstream of the primary terminator of tRNA genes. Values were calculated for 60 bp sliding windows. The black line corresponds to the average value whereas the grey zone represents the 95% confidence interval of the average value. A dashed line indicates the average value for the whole genome.

D) Analysis of the Gibbs free energy (ΔG) as a proxy for the propensity of the transcribed regions to form secondary structures. We plotted values calculated for 65 bp sliding windows published in Turowski et al. 2020. A dashed line indicates the average ΔG value for the whole genome.

E) Combined representation of the GC content and the ΔG of the different regions coloured according to the distance from the primary terminator. Closer regions (blue) tend to be less GC-rich and less structured while further regions (red) tend to be more GC-rich and structured.

Table S6: Yeast strains used in this study.

| Number | Name | Genotype | Source |
|---------|-------------------------------------|---|---------------------|
| DLY671 | BMA | as W303, $\Delta trp1$ | F. Lacroute |
| DLY1152 | $\Delta rrp6$ | as W303, $rrp6::URA3$ | F. Lacroute |
| DLY1605 | $pTet-NRD1$ | as BMA, $pTet::FLAG::NRD1$ | J. Colin |
| DLY1626 | $pTet-NRD1, \Delta rrp6$ | as BMA, $pTet::FLAG::NRD1, rrp6::KAN$ | This work |
| DLY1656 | $P_{GAL1}-TAP-SEN1$ | as BMA, $TRP1::Pgal::TAP::SEN1$ | Porrua et al, 2012 |
| DLY2692 | $P_{GAL1}-TAP-sen1\Delta Nter$ | as BMA, $TRP1::Pgal::TAP::sen1\Delta Nter$ ($\Delta 1-975$) | Han et al, 2020 |
| DLY3171 | Sen1-TAP | as W303, $SEN1::TAP::KAN$ | Appanah et al, 2020 |
| DLY3173 | $sen1-3$ -TAP | as W303, $sen1W773A, E774A, W777A::TAP::KAN$ | Appanah et al, 2020 |
| DLY3197 | $sen1-3$ | as W303, $sen1W773A, E774A, W777A$ | This work |
| DLY3246 | $sen1-3, \Delta rrp6$ | as W303, $sen1W773A, E774A, W777A rrp6::URA3$ | This work |
| DLY3262 | Rpc160-HTP | as BMA, $RPC160::HTP::TRP1$ | This work |
| DLY3265 | Rpc160-HTP, $sen1-3$ | as W303, $RPC160::HTP::TRP1, sen1W773A, E774A, W777A$ | This work |
| DLY3343 | Rpc160-HTP, Sen1-AID | as BMA, $RPC160::HTP::TRP1, SEN1-AID::KAN::OsTIR1$ | This work |
| DLY3377 | Rpc160-HTP, Nrd1-AID | as BMA, $RPC160::HTP::TRP1, NRD1-3Flag-AID::KAN::OsTIR1$ | This work |
| DLY3462 | $GAL::HA-Trz1$ | $GAL-HA-Trz1::KAN$ | This work |
| DLY3463 | $GAL::HA-Trz1, sen1-3$ | $GAL-HA-Trz1::KAN, sen1W773A, E774A, W777A$ | This work |
| DLY3464 | $GAL::HA-Trz1, \Delta rrp6$ | $GAL-HA-Trz1::KAN, rrp6::URA3$ | This work |
| DLY3465 | $GAL::HA-Trz1, \Delta rrp6, sen1-3$ | $GAL-HA-Trz1::KAN, rrp6::URA3, sen1W773A, E774A, W777A$ | This work |

Methods

Construction of yeast strains and plasmids

All the strains used in this paper are listed in table S6. Tagging of *RPC160* with the HTP-tag was performed with standard procedures (Longtine et al., 1998; Rigaut et al., 1999) using plasmid pDL599. Plasmid pDL995 for expression of recombinant Sen1 in insect cells was constructed using the SLIC (sequence and ligation-independent cloning) method (Li and Elledge, 2012).

Co-immunoprecipitation (Co-IP)

For immunoprecipitation of proteins expressed under their own promoter cells were grown on YPD medium. For proteins expressed under the control of the *GAL1* promoter (i.e. full-length Sen1, Sen1 Δ NTD and the N-terminal domain of Sen1), cells were grown on rich medium containing 20 g/L of galactose instead of glucose as the carbon source. Cultures (typically 250 mL) were grown to $OD_{600} \sim 1$ and then cells were collected by centrifugation and resuspended in 1.5 mL of lysis buffer (10 mM sodium phosphate pH 7.5, 200 mM sodium acetate, 0.25% NP-40, 2 mM EDTA, 1 mM EGTA, 5% glycerol) containing protease inhibitors. Suspensions were frozen in liquid nitrogen and lysed using a Retsch MM301 Ball Mill (5 cycles of 3 minutes at 15 Hz). Lysates were clarified by centrifugation at 13 krpm for 30 min at 4°C and, unless otherwise indicated, treated with 20 μ g/mL RNase A for 20 min at 25°C prior to immunoprecipitation. For HTP-tagged proteins, the extracts were then incubated with 2.5 mg of IgG-coupled M-280 tosylactivated dynabeads (Thermo Fisher) for 2 h at 4°C with rotation. After incubation, beads were washed three times with lysis buffer and once with H₂O and used directly in mass spectrometry analyses.

For proteins overexpressed from the *GAL1* promoter IgG sepharose (GE HEathcare) was used instead. After washes with lysis buffer, beads were washed with TEV cleavage buffer (10 mM Tris pH 8, 150 mM NaCl, 0.1 % NP-40, 0.5 mM EDTA, 1 mM DTT, 5% glycerol) and proteins were then eluted by cleaving the protein A moiety with the TEV protease in TEV cleavage overnight at 4°C.

Mass spectrometry analysis and label-free quantification

Analysis of Sen1 and RNAPIII coimmunoprecipitates by mass spectrometry was conducted by the proteomics core facility of the Institut Jacques Monod. Proteins were digested by adding 0.2 μ g of trypsin (Promega, Madison, WI, USA) per sample followed by incubation in 25 mM NH₄HCO₃ at 37°C overnight. The resulting peptides were desalted using ZipTip μ -C18 Pipette

Tips (Pierce Biotechnology, Rockford, IL, USA) and analyzed using an Orbitrap Fusion equipped with an easy spray ion source and coupled to a nano-LC Proxeon 1200 (Thermo Scientific, Waltham, MA, USA). Peptides were loaded with an online pre-concentration method and separated by chromatography using a Pepmap-RSLC C18 column (0.75 x 750 mm, 2 μ m, 100 Å) from Thermo Scientific, equilibrated at 50°C and operating at a flow rate of 300 nl/min. Peptides were eluted by a gradient of solvent A (H₂O, 0.1 % FA) and solvent B (ACN/H₂O 80/20, 0.1% FA), the column was first equilibrated 5 min with 95 % of A, then B was raised to 28 % in 105 min and to 40% in 15 min. Finally, the column was washed with 95% of B during 20 min and re-equilibrated with 95% of A during 10 min. Peptide masses were analyzed in the Orbitrap cell in full ion scan mode, at a resolution of 120,000, a mass range of m/z 350-1550 and an AGC target of $4 \cdot 10^5$. MS/MS were performed in the top speed 3 s mode. Peptides were selected for fragmentation by Higher-energy C-trap Dissociation (HCD) with a Normalized Collisional Energy of 27% and a dynamic exclusion of 60 seconds. Fragment masses were measured in an Ion trap in the rapid mode, with and an AGC target of $1 \cdot 10^4$. Monocharged peptides and unassigned charge states were excluded from the MS/MS acquisition. The maximum ion accumulation times were set to 100 ms for MS and 35 ms for MS/MS acquisitions respectively.

Label-free quantification was done on Progenesis Q1 for Proteomics (Waters, Milford, MA, USA) in Hi-3 mode for protein abundance calculation. MGF peak files from Progenesis were processed by Proteome Discoverer 2.4 with the Mascot search engine. The Swissprot protein database was typically used for interrogation. A maximum of 2 missed cleavages was authorized. Precursor and fragment mass tolerances were set to 7 ppm and 0.5 Da, respectively. The following post-translational modifications were included as variable: Oxidation (M), Phosphorylation (STY). Spectra were filtered using a 1% FDR using the percolator node.

UV crosslinking and analysis of cDNA (CRAC)

The CRAC protocol used in this study is derived from Granneman et al. (Granneman et al., 2009) with several modifications as previously described (Candelli et al., 2018). Briefly, 2 L of cells expressing an HTP-tagged version of Rpc160 (the largest subunit of RNAPIII) at the endogenous locus were grown at 30°C to $OD_{600} \sim 0.6$ in CSM-TRP medium. Cells were crosslinked for 50 seconds using a W5 UV crosslinking unit (UVO3 Ltd) and harvested by centrifugation. Cell pellets were washed once with ice-cold 1x PBS and resuspended in 2.4 mL of TN150 buffer (50 mM Tris pH 7.8, 150 mM NaCl, 0.1% NP-40 and 5 mM β -mercaptoethanol) per gram of cells in the presence of protease inhibitors (Complete™ EDTA-free Protease Inhibitor Cocktail, Roche). Suspensions were flash frozen in droplets and cells subjected to

cryogenic grinding using a Ball Mill MM 400 (5 cycles of 3 minutes at 20 Hz). The resulting frozen lysates were thawed on ice and digested with DNase I (165 units per gram of cells) at 25°C for 1h to solubilize chromatin and then clarified by centrifugation at 16 krpm for 30 min at 4°C.

RNA-protein complexes were immobilized on M-280 tosylactivated dynabeads coupled with rabbit IgGs (10 mg of beads per sample), washed with TN1000 buffer (50 mM Tris pH 7.8, 1 M NaCl, 0.1% NP-40 and 5 mM β -mercaptoethanol), and eluted by digestion with the TEV protease. RNAs were subjected to partial degradation to reduce their size by adding with 0.2 U of RNase cocktail (RNase-IT, Agilent) and the reaction was stopped by the addition of guanidine-HCl to a final concentration of 6 M. RNA-protein complexes were then incubated with Ni-NTA sepharose (Qiagen, 100 μ l of slurry per sample) overnight at 4°C and extensively washed. Sequencing adaptors were ligated to the RNA molecules as described in the original procedure. RNA-protein complexes were eluted with elution buffer containing 50 mM Tris pH 7.8, 50 mM NaCl, 150 mM imidazole, 0.1% NP-40 and 5 mM β -mercaptoethanol fractionated using a Gel Elution Liquid Fraction Entrapment Electrophoresis (GelFree) system (Expedeon) following manufacturer's specifications. The fractions containing Rpc160 were treated with 100 μ g of proteinase K, and RNAs were purified and reverse-transcribed using reverse transcriptase Superscript IV (Invitrogen).

The cDNAs were amplified by PCR using LA Taq polymerase (Takara), and then, the PCR reactions were treated with 200 U/mL of Exonuclease I (NEB) for 1 h at 37°C. Finally, the DNA was purified using NucleoSpin columns (Macherey-Nagel) and sequenced on a NextSeq 500 Illumina sequencer.

Synchronization of cells in G1 and analysis by flow cytometry

2L of cells were synchronized in the G1 phase of the cell cycle by adding 4 mg of α -factor. To maintain cells in G1, 8 mg and 4 mg of α -factor were subsequently added to the culture after 1h and 2h of incubation at 30°C, respectively. Cells were collected and processed 1h after the last addition of α -factor.

To analyse the DNA content of synchronized cells, 2 mL of culture were collected at different time points and cells were harvested by centrifugation. Cell pellets were resuspended in 50 mM sodium citrate buffer and treated with RNase A (QIAGEN) for 2 hours at 50°C, followed by proteinase K (Sigma) treatment for 2 hours at 50°C. Cell aggregates were then dissociated by sonication and 40 μ L of cell suspension were incubated with 170 μ L of 50 mM sodium citrate buffer containing 0.5 μ M Sytox Green (Invitrogen). Data were acquired on a MASQuant Analyzer (Miltenyi Biotec) and analyzed with FlowJo Software.

Dataset processing

CRAC reads were demultiplexed using the `pyBarcodeFilter` script from the `pyCRACUtility` suite (Webb et al., 2014). Next, the 5' adaptor was clipped with `Cutadapt` and the resulting insert quality-trimmed from the 3' end using `Trimmomatic` rolling mean clipping (Bolger et al., 2014). We used the `pyCRAC` script `pyFastqDuplicateRemover` to collapse PCR duplicates using a 6-nucleotide random tag included in the 3' adaptor. The resulting sequences were reverse complemented with the `Fastx` reverse complement that is part of the `fastx` toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) and mapped to the R64 genome with `bowtie2` using “-N 1” option. Reads shorter than 20 nt were filtered out after mapping and coverage files were generated and normalized to counts per million (CPM) using the `bamCoverage` tool from the `deepTools` package (Ramírez et al., 2016) using a bin size of 1.

Bioinformatic analyses

All sequence files and annotations were obtained from *Saccharomyces* Genome Database (*S. cerevisiae* genome version R64-2-1). T-tracts were annotated by first searching for sequences containing at least 4 consecutive thymines (for the plus strand) or adenines (for the minus strand) using the unix command line tool `grep` and then generating coordinate files by the `awk` command. The resulting files were then combined into a single BED file (table S7) using `BEDOPS` suite (Neph et al., 2012) with the “*everything*” option. For each tRNA gene, the primary terminator was defined as the 1st T-tract after the 3' end of the mature tRNA. Such primary terminators were identified by comparing the mentioned T-tract annotations and the tRNAs annotations with the `closest` tool from `BEDTools` (Quinlan and Hall, 2010). T-tracts falling within the 700 bp region immediately downstream of the primary terminator of each tRNA gene were identified with the `BEDTools intersect` tool and defined as secondary terminators (table S8).

Reads mapped to different classes of RNAs were summarized by `BEDTools coverage`. Metagene analyses of RNAPIII occupancy were performed with `deepTools` suite (Ramírez et al., 2016). Strand-specific coverage bigwig files and modified tRNA coordinates (from the 5' end to the end of 1st T-tract) were used as inputs for the `computeMatrix` tool using a bin size of 1 and the `scale-regions` mode. The matrices generated for each strand were subsequently combined by the `computeMatrixOperations` tool with the `rbind` option and used as inputs for the `plotProfile` tool to create a summary plot. For heatmap analyses the log₂ ratio of the RNAPIII signal in the *sen1-3* mutant relative to the WT was calculated by the `bigwigCompare` tool using a bin size of 1 and the corresponding bigwig coverage files as inputs. Matrices were generated and combined as for metagene analyses and the final matrix was used as the input for the `plotHeatmap` tool. To analyze the correlation between two replicates, the average

RNAPIII signal over regions comprising tRNA genes and 500 bp upstream and downstream regions was computed using the *multiBigwigSummary* tool. The resulting tables were used as inputs for the *plotCorrelation* tool to generate scatter plots and calculate the correlation coefficients using the Spearman method.

To annotate tRNA genes read-through regions in the WT and the *sen1-3* mutant, we first determined a threshold below which CRAC signal was considered as background signal. To do so, genomic regions corresponding to protein-coding genes, which are transcribed by RNAPII, were divided into 20 bp non-overlapping windows and the total signal was computed for each of them using normalized coverage files. The value corresponding to the 95% quantile (value below which 95% of windows values fall) was set as threshold. For each tRNA gene, the 1 kb region immediately downstream of the primary terminator was then divided into 20 bp windows with 1 bp overlap and the RNAPIII CRAC signal was computed for all of them. Contiguous windows with values above the threshold were merged and, most often, this resulted in a single read-through region for each gene. When this was not the case, we manually merged the fragmented regions that were separated by small gaps. Final annotations are provided as BED files in table S9.

The efficiency of transcription termination in the WT and the *sen1-3* mutant was estimated by calculating the read-through index defined as the percentage of RNAPIII signal over the read-through regions relative to the signal over tRNA gene regions. The total RNAPIII signal at each region was computed with the UCSC *bigWigAverageOverBed* package (<http://genome.ucsc.edu>) using the aforementioned annotations.

Data representation and statistical analyses were performed with R using the *ggplot2* and *plyr* (<https://cran.r-project.org/web/packages/plyr/index.html>) packages.

RNA analyses

Yeast cells were grown on 30 mL of YPD medium containing the appropriate additives, depending on the experiment, at 30°C to OD₆₀₀ 0.3 to 0.6. Cells were harvested by centrifugation and RNAs were prepared using standard methods. 4 µg of total RNA were reverse-transcribed by the M-MLV reverse transcriptase (New England BioLabs) following the manufacturer specifications and using oligo d(T) and a mixture of random hexamers at 37°C for 45 min. The resulting cDNAs were analysed by quantitative PCR using the LightCycler 480 SYBR Green I Master reagent (Roche) and LightCycler 480 instrument (Roche) using primers specific for the regions to detect (table S5).

For northern blot assays, typically 10 µg of total RNA were loaded onto a 10% denaturing polyacrylamide gel and separated by electrophoresis at 20 W for 2 h. RNAs were

then transferred to a nitrocellulose membrane (GE Amersham HybondTM-N⁺) using a wet transfer system (Trans-Blot cell, Bio-Rad) at 100 V for 2 h at 4°C. Membranes were UV cross-linked and hybridized with the corresponding radioactively labeled probe in commercial buffer (Ultrascreen, Amersham) overnight. For abundant RNA species we employed 5' end labeled DNA oligonucleotides as probes and hybridizations and subsequent washes were performed at 42°C. For RNA species that were very poorly detected using DNA oligonucleotide probes, we employed RNA probes generated by *in vitro* transcription in the presence of $\alpha^{32}\text{P}$ -UTP using the MAXIscript Kit (Ambion). Hybridization was then performed at 68°C overnight, and the membrane was washed two times for 15 min at 42°C with buffer 2x SSC (30 mM sodium citrate pH 7.0, 300 mM NaCl) containing 0.5% SDS and two times for 15 min at 60°C with buffer 0.1x SSC containing 0.1% SDS. After washes, blots were exposed on a phosphorimager screen and finally scanned using a Typhoon scanner (GE healthcare). Images were analysed using the ImageQuant software (GE healthcare).

Protein purification

RNAPIII and RNAPI were purified from *Saccharomyces cerevisiae* by heparin chromatography, followed by IgG-affinity chromatography and finally anion-exchange using a previously described procedure (Moreno-Morcillo et al., 2014) with the following modifications: for cells lysis and equilibration of the heparin column (GE Healthcare), we used instead a buffer containing 250 mM Tris-HCl pH 8, 250 mM ammonium sulfate, 20% glycerol, 1 mM EDTA, 10 mM MgCl₂, 10 μM ZnCl₂, 12 mM β -mercaptoethanol and a protease inhibitor cocktail (0.3 $\mu\text{g}/\text{mL}$ leupeptin, 1.4 $\mu\text{g}/\text{mL}$ pepstatin A, 170 $\mu\text{g}/\text{mL}$ PMSF and 330 $\mu\text{g}/\text{mL}$ benzamidin). Purified RNAPIII and RNAPI were buffer-exchanged to storage buffer (15 mM HEPES pH 7.5, 150 mM ammonium sulfate, 10 mM DTT), concentrated to 14.9 mg/mL (RNAPIII) and 10.4 mg/mL (RNAPI), snap-frozen in liquid nitrogen and stored at -80°C.

Full-length Sen1 was overexpressed from pFL vector in insect cells (*Trichoplusia ni*) using an optimized synthetic gene (GeneArt, Life Technologies) and the baculovirus expression system (Berger et al., 2004). Briefly, Hi5 cells (Thermo Fisher Scientific) expressing a C-terminally His₆-tagged version of Sen1 were harvested by centrifugation and lysed by sonication at 4°C in ice-cold buffer A1 (50 mM HEPES-NaOH, pH 7.75, 600 mM NaCl, 15% (v/v) glycerol, 5 mM β -mercaptoethanol, 10 mM imidazole, 2 mM MgCl₂) supplemented with a cocktail of protease inhibitors EDTA free (Roche). The lysate was clarified by centrifugation at 15000 rpm for 1h at 4°C, filtered through a 0.45 μm filter and loaded on a 5 mL Protino Ni-NTA agarose prepacked column (Machery Nagel), equilibrated with buffer A1. To get rid of nucleic acids bound to Sen1, the column was washed with high salt buffer B2 (50 mM HEPES-

NaOH, pH 7.75, 1M NaCl, 15% (v/v) glycerol, 5 mM β -mercaptoethanol, 2 mM MgCl_2) and equilibrated back to buffer A1. Proteins were eluted using a linear gradient from buffer A1 to B (50 mM HEPES-NaOH, pH 7.75, 200 mM NaCl, 10 % (v/v) glycerol, 5 mM β -mercaptoethanol, 400 mM imidazole, 2 mM MgCl_2), diluted with 2 volumes of buffer D1 (50 mM HEPES-NaOH, pH 7.75, 50 mM NaCl, 15 % (v/v) glycerol, 2 mM β -mercaptoethanol, 2 mM MgCl_2) and then loaded onto two tandem 5 mL heparin HP pre-packed columns (GE Healthcare). Sen1 was eluted using a linear gradient from 20% to 100% of buffer B2 containing 50 mM HEPES-NaOH, pH 7.75, 1 M NaCl, 5% (v/v) glycerol, 2 mM MgCl_2 and 2 mM DTT. Peak fractions were pooled and subjected to size exclusion chromatography using a Superdex 200 16/600 column (GE healthcare) equilibrated in buffer A3 (50 mM HEPES-NaOH, pH 7.75, 300 mM NaCl, 5% (v/v) glycerol, 2 mM Mg acetate, 2 mM DTT). Finally, the fractions of interest were concentrated using an Amicon Ultra-100 centrifugal filter (Millipore), aliquoted, flash frozen in liquid nitrogen and stored at -80°C .

***In vitro* transcription termination assays**

RNAPIII transcription termination assays were performed using essentially the previously described method for RNAPII (Porrua and Libri, 2015b) with some modifications. For each reaction, elongation complexes (ECs) were assembled by annealing 2.5 pmol of 5'-end radioactively labeled RNA primer with 2.5 pmol of template DNA oligo in hybridization buffer (HB) buffer (20 mM Hepes pH 7.6, 100 mM NaCl, 12 mM MgCl_2 , 10 mM DTT). Subsequently, the RNA:DNA hybrids were incubated with 2 pmol of highly purified RNAPIII in transcription reaction buffer (TRB) buffer (20 mM Hepes pH 7.6, 60 mM $(\text{NH}_4)_2\text{SO}_4$, 10 mM MgCl_2 , 10% glycerol, 10 mM DTT) at 20°C for 10 min at 550 rpm. Next, 5 pmol of 5'-end biotinylated non-template DNA were added to the mixture and incubated at 20°C for 10 min with shaking. The resulting ternary ECs were mixed with streptavidin beads (Dynabeads MyOne Streptavidin T1 from Invitrogen, 10 μL of slurry per reaction) pre-washed 4 times with TRB buffer containing 0.1% triton X-100 and then incubated at 20°C for 30 min with gentle shaking. After binding, the beads were washed with 1 volume of TRB containing 0.1% triton X-100, then with 1 volume of TRB containing 250 mM $(\text{NH}_4)_2\text{SO}_4$, and finally with 1 volume of TRB. After washes, beads were resuspended in 13 μL of TRB buffer. The reaction was started by adding 7 μL of nucleotides mixture (1 mM each in TRB buffer) and incubating at 28°C for 10 min, and then stopped by the addition of 1 μL of 0.5 M EDTA. Beads and supernatant fractions were then collected separately. RNAs in the supernatant were ethanol-precipitated and resuspended in 10 μL of loading buffer containing 1x TBE and 8 M urea and incubated at 95°C for 3 min before loading onto a 10% denaturing polyacrylamide gel. To isolate RNAs from beads, 10 μL of

loading buffer was added to the beads and boiled at 95°C for 3 min and then recovered supernatants as “bead fractions”. Finally, sample were subjected to 10% denaturing PAGE, running for 1 h at 40 W in 1x TBE buffer. Gels were exposed on a phosphorimager screen overnight at -80°C and screens were scanned using a Typhoon scanner (GE healthcare). Images were analysed using the ImageQuant software (GE healthcare).

Data availability

The RNAPIII CRAC data have been deposited in NCBI's Gene Expression Omnibus (GEO) and are accessible through GEO Series accession number GSE174738.

Acknowledgements

We thank G. Wentzinger for technical assistance and other members of the Libri lab for fruitful discussions. We thank F. Fiorini and Hervé Le Hir for sharing the Upf1 protein. We thank the Roscoff Bioinformatics platform ABiMS (<http://abims.sb-roscoff.fr>) for providing computational resources and support. This work has benefited from the facilities and expertise of the high throughput sequencing core facility of I2BC (<http://www.i2bc.paris-saclay.fr/>). We thank the proteomics facility of the Institut Jacques Monod, supported by the Region Ile-de-France (SESAME), Université de Paris and the CNRS, for their technical assistance.

Funding

This study was supported by the Centre National de la Recherche Scientifique (CNRS), the Agence National pour la Recherche (ANR-16-CE12-0001-01 to O.P. and ANR-16-CE12-0022-01 to D.L) and the Fondation pour la Recherche Medical (F.R.M., programme Equipes 2019). J.X. was supported by the China Scholarship Council, by the FRM (FDT202012010433) and the LabEx “Who Am I?” (ANR-11-LABX-0071 and the Université de Paris IdEx ANR-18-IDEX-0001) funded by the French Government through its “Investments for the Future”. U.A. was supported by the French Ministry of Education and Research and by the Fondation ARC pour la recherche sur le cancer. M.G. was supported by a Boehringer Ingelheim Fonds PhD fellowship and the EMBL International PhD program. C.W.M. was supported by the EMBL. J.S. and V.P. were supported by the DFG grant PE 2079/2-2.

Authors contributions

J.X. conducted all experiments except for those in table 1, performed bioinformatic analyses together with Y.C. and contributed to manuscript writing. U.A. and N.H. performed experiments in table 1. M.G. and C.W.M. purified RNAPIII and RNAPI and provide experimental advice. J.S. and V.P. purified Sen1 protein. O.P. conceived the project. O.P. and D.L. designed experiments, supervised research and wrote the manuscript with inputs from all coauthors.

References

- Alzu, A., Bermejo, R., Begnis, M., Lucca, C., Piccini, D., Carotenuto, W., Saponaro, M., Brambati, A., Cocito, A., Foiani, M., et al. (2012). Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. *Cell* *151*, 835–846.
- Appanah, R., Lones, E.C., Aiello, U., Libri, D., and De Piccoli, G. (2020). Sen1 Is Recruited to Replication Forks via Ctf4 and Mrc1 and Promotes Genome Stability. *Cell Rep* *30*, 2094-2105.e9.
- Arimbasseri, A.G., and Maraia, R.J. (2015). Mechanism of Transcription Termination by RNA Polymerase III Utilizes a Non-template Strand Sequence-Specific Signal Element. *Mol. Cell* *58*, 1124–1132.
- Arimbasseri, A.G., Rijal, K., and Maraia, R.J. (2013). Transcription termination by the eukaryotic RNA polymerase III. *Biochim. Biophys. Acta* *1829*, 318–330.
- Arimbasseri, A.G., Kassavetis, G.A., and Maraia, R.J. (2014). Transcription. Comment on “Mechanism of eukaryotic RNA polymerase III transcription termination.” *Science* *345*, 524.
- Arndt, K.M., and Reines, D. (2015). Termination of Transcription of Short Noncoding RNAs by RNA Polymerase II. *Annual Review of Biochemistry* *84*, 381–404.
- Baejen, C., Andreani, J., Torkler, P., Battaglia, S., Schwalb, B., Lidschreiber, M., Maier, K.C., Boltendahl, A., Rus, P., Esslinger, S., et al. (2017). Genome-wide Analysis of RNA Polymerase II Termination at Protein-Coding Genes. *Molecular Cell* *66*, 38-49.e6.
- Berger, I., Fitzgerald, D.J., and Richmond, T.J. (2004). Baculovirus expression system for heterologous multiprotein complexes. *Nat Biotechnol* *22*, 1583–1587.
- Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* *30*, 2114–2120.
- Bridier-Nahmias, A., Tchalikian-Cosson, A., Baller, J.A., Menouni, R., Fayol, H., Flores, A., Saïb, A., Werner, M., Voytas, D.F., and Lesage, P. (2015). Retrotransposons. An RNA polymerase III subunit determines sites of retrotransposon integration. *Science* *348*, 585–588.
- Candelli, T., Challal, D., Briand, J.-B., Boulay, J., Porrua, O., Colin, J., and Libri, D. (2018). High-resolution transcription maps reveal the widespread impact of roadblock termination in yeast. *EMBO J.* *37*.
- Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H., and Conti, E. (2011). Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. *Mol. Cell* *41*, 693–703.
- 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 Dev.* *22*, 1069–1081.
- Granneman, S., Kudla, G., Petfalski, E., and Tollervey, D. (2009). Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc. Natl. Acad. Sci. U.S.A.* *106*, 9613–9618.
- Han, Z., and Porrua, O. (2017). Helicases as transcription termination factors: Different solutions for a common problem. *Transcription* *1–7*.
- Han, Z., Libri, D., and Porrua, O. (2017). Biochemical characterization of the helicase Sen1 provides new insights into the mechanisms of non-coding transcription termination. *Nucleic Acids Res.* *45*, 1355–1370.
- Han, Z., Jasnovidova, O., Haidara, N., Tudek, A., Kubicek, K., Libri, D., Stefl, R., and Porrua, O. (2020). Termination of non-coding transcription in yeast relies on both an RNA Pol II CTD interaction domain and a CTD-mimicking region in Sen1. *EMBO J.* e101548.
- Hazelbaker, D.Z., Marquardt, S., Wlotzka, W., and Buratowski, S. (2013). Kinetic competition between RNA Polymerase II and Sen1-dependent transcription termination. *Mol. Cell* *49*, 55–66.
- Hoffmann, N.A., Jakobi, A.J., Moreno-Morcillo, M., Glatt, S., Kosinski, J., Hagen, W.J.H., Sachse, C., and Müller, C.W. (2015). Molecular structures of unbound and transcribing RNA polymerase III. *Nature* *528*, 231–236.
- Kang, J.Y., Mishanina, T.V., Bellecourt, M.J., Mooney, R.A., Darst, S.A., and Landick, R. (2018). RNA Polymerase Accommodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing. *Molecular Cell* *69*, 802-815.e1.

- Kawauchi, J., Mischo, H., Braglia, P., Rondon, A., and Proudfoot, N.J. (2008). Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. *Genes Dev.* 22, 1082–1092.
- Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedeá, E., Greenblatt, J.F., and Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* 432, 517–522.
- Kireeva, M.L., Komissarova, N., Waugh, D.S., and Kashlev, M. (2000). The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex. *J. Biol. Chem.* 275, 6530–6536.
- Landrieux, E., Alic, N., Ducrot, C., Acker, J., Riva, M., and Carles, C. (2006). A subcomplex of RNA polymerase III subunits involved in transcription termination and reinitiation. *EMBO J* 25, 118–128.
- Larochelle, M., Robert, M.-A., Hébert, J.-N., Liu, X., Matteau, D., Rodrigue, S., Tian, B., Jacques, P.-É., and Bachand, F. (2018). Common mechanism of transcription termination at coding and noncoding RNA genes in fission yeast. *Nat Commun* 9.
- Leonaitė, B., Han, Z., Basquin, J., Bonneau, F., Libri, D., Porrua, O., and Conti, E. (2017). Sen1 has unique structural features grafted on the architecture of the Upf1-like helicase family. *EMBO J.* 36, 1590–1604.
- Li, M.Z., and Elledge, S.J. (2012). SLIC: a method for sequence- and ligation-independent cloning. *Methods Mol Biol* 852, 51–59.
- Li, W., Selvam, K., Rahman, S.A., and Li, S. (2016). Sen1, the yeast homolog of human senataxin, plays a more direct role than Rad26 in transcription coupled DNA repair. *Nucleic Acids Res.* 44, 6794–6802.
- Longtine, M.S., McKenzie, A., 3rd, Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J.R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* 14, 953–961.
- Merkel, P., Perez-Fernandez, J., Pilsl, M., Reiter, A., Williams, L., Gerber, J., Böhm, M., Deutzmann, R., Griesenbeck, J., Milkereit, P., 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.
- Mischo, H.E., Gomez-Gonzalez, B., Grzechnik, P., Rondon, A.G., Wei, W., Steinmetz, L., Aguilera, A., and Proudfoot, N.J. (2011). Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol Cell* 41, 21–32.
- Mishra, S., and Maraia, R.J. (2019). RNA polymerase III subunits C37/53 modulate rU:dA hybrid 3' end dynamics during transcription termination. *Nucleic Acids Res* 47, 310–327.
- Moreno-Morcillo, M., Taylor, N.M.I., Gruene, T., Legrand, P., Rashid, U.J., Ruiz, F.M., Steuerwald, U., Müller, C.W., and Fernández-Tornero, C. (2014). Solving the RNA polymerase I structural puzzle. *Acta Crystallogr D Biol Crystallogr* 70, 2570–2582.
- Neph, S., Kuehn, M.S., Reynolds, A.P., Haugen, E., Thurman, R.E., Johnson, A.K., Rynes, E., Maurano, M.T., Vierstra, J., Thomas, S., et al. (2012). BEDOPS: high-performance genomic feature operations. *Bioinformatics* 28, 1919–1920.
- Nielsen, S., and Zenkin, N. (2014). Transcription. Response to Comment on “Mechanism of eukaryotic RNA polymerase III transcription termination.” *Science* 345, 524.
- Nielsen, S., Yuzenkova, Y., and Zenkin, N. (2013). Mechanism of eukaryotic RNA polymerase III transcription termination. *Science* 340, 1577–1580.
- Osmundson, J.S., Kumar, J., Yeung, R., and Smith, D.J. (2017). Pif1-family helicases cooperatively suppress widespread replication-fork arrest at tRNA genes. *Nat Struct Mol Biol* 24, 162–170.
- Park, J., Kang, M., and Kim, M. (2015). Unraveling the mechanistic features of RNA polymerase II termination by the 5'-3' exoribonuclease Rat1. *Nucleic Acids Res* 43, 2625–2637.
- Pearson, E.L., and Moore, C.L. (2013). Dismantling Promoter-driven RNA Polymerase II Transcription Complexes in Vitro by the Termination Factor Rat1. *J. Biol. Chem.* 288, 19750–19759.
- Porrua, O., and Libri, D. (2013). A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast. *Nat. Struct. Mol. Biol.* 20, 884–891.
- Porrua, O., and Libri, D. (2015a). Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat. Rev. Mol. Cell Biol.* 16, 190–202.
- Porrua, O., and Libri, D. (2015b). Characterization of the mechanisms of transcription termination by the helicase

- Sen1. *Methods Mol. Biol.* **1259**, 313–331.
- Porrúa, O., Hobor, F., Boulay, J., Kubicek, K., D'Aubenton-Carafa, Y., Gudipati, R.K., Stefl, R., and Libri, D. (2012). In vivo SELEX reveals novel sequence and structural determinants of Nrd1-Nab3-Sen1-dependent transcription termination. *EMBO J.* **31**, 3935–3948.
- Porrúa, O., Boudvillain, M., and Libri, D. (2016). Transcription Termination: Variations on Common Themes. *Trends Genet.* **32**, 508–522.
- Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842.
- Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–165.
- Reiter, A., Hamperl, S., Seitz, H., Merkl, P., Perez-Fernandez, J., Williams, L., Gerber, J., Németh, A., Léger, I., Gadal, O., et al. (2012). The Reb1-homologue Ydr026c/Nsi1 is required for efficient RNA polymerase I termination in yeast. *EMBO J.* **31**, 3480–3493.
- Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Séraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **17**, 1030–1032.
- Rijal, K., and Maraia, R.J. (2013). RNA polymerase III mutants in TFIIIF-like C37 that cause terminator readthrough with no decrease in transcription output. *Nucleic Acids Res* **41**, 139–155.
- Rivosecchi, J., Laroche, M., Teste, C., Grenier, F., Malapert, A., Ricci, E.P., Bernard, P., Bachand, F., and Vanoosthuyse, V. (2019). Senataxin homologue Sen1 is required for efficient termination of RNA polymerase III transcription. *EMBO J* **38**, e101955.
- Schulz, D., Schwalb, B., Kiesel, A., Baejen, C., Torkler, P., Gagneur, J., Soeding, J., and Cramer, P. (2013). Transcriptome Surveillance by Selective Termination of Noncoding RNA Synthesis. *Cell*.
- Skowronek, E., Grzechnik, P., Späth, B., Marchfelder, A., and Kufel, J. (2014). tRNA 3' processing in yeast involves tRNase Z, Rex1, and Rrp6. *RNA* **20**, 115–130.
- Stefl, R., Wu, H., Ravindranathan, S., Sklenár, V., and Feigon, J. (2004). DNA A-tract bending in three dimensions: solving the dA4T4 vs. dT4A4 conundrum. *Proc Natl Acad Sci U S A* **101**, 1177–1182.
- Steinmetz, E.J., Warren, C.L., Kuehner, J.N., Panbehi, B., Ansari, A.Z., and Brow, D.A. (2006). Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol. Cell* **24**, 735–746.
- 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.
- Verosloff, M.S., Corcoran, W.K., Dolberg, T.B., Bushhouse, D.Z., Leonard, J.N., and Lucks, J.B. (2021). RNA Sequence and Structure Determinants of Pol III Transcriptional Termination in Human Cells. *J Mol Biol* **433**, 166978.
- Vorländer, M.K., Khatter, H., Wetzels, R., Hagen, W.J.H., and Müller, C.W. (2018). Molecular mechanism of promoter opening by RNA polymerase III. *Nature* **553**, 295–300.
- Wang, S., Han, Z., Libri, D., Porrúa, O., and Strick, T.R. (2019). Single-molecule characterization of extrinsic transcription termination by Sen1 helicase. *Nat Commun* **10**, 1545.
- Webb, S., Hector, R.D., Kudla, G., and Granneman, S. (2014). PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. *Genome Biol* **15**, R8.
- 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.
- Wlotzka, W., Kudla, G., Granneman, S., and Tollervey, D. (2011). The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. *EMBO J.* **30**, 1790–1803.
- Yüce, Ö., and West, S.C. (2013). Senataxin, defective in the neurodegenerative disorder ataxia with oculomotor apraxia 2, lies at the interface of transcription and the DNA damage response. *Mol. Cell. Biol.* **33**, 406–417.

DISCUSSION & PERSPECTIVES

In this study, we provided evidence that Sen1 interacts directly with RNAPIII and this interaction is essential for efficient RNAPIII termination *in vivo*. Our *in vitro* results showed that Sen1 can directly induce the release of RNAPIII from the DNA by the same mechanism employed in termination of RNAPII transcription. We also showed that the presence of RNA secondary structures can promote RNAPIII termination in the appropriate context *in vitro*. The function of Sen1 and RNA structures in terminating RNAPIII have been discussed in the manuscript. Here I will mention some additional data and discuss some questions that are not included in the manuscript.

What is the Sen1-RNAPIII interaction surface?

Although the W773, E774 and W777 residues of Sen1 are essential for the interaction of Sen1 with RNAPIII, the specific surface of RNAPIII that interacts with Sen1 is unknown. Identifying this surface could provide important insights into the mechanisms of Sen1 recruitment and its function in termination of RNAPIII transcription. Furthermore, this would allow us to generate RNAPIII mutants specifically affected in the interaction of Sen1 with RNAPIII to further validate the results obtained with the *sen1-3* mutant.

Therefore, we tried to map the Sen1-RNAPIII interaction regions by yeast two hybrid assays. Briefly, components of the RNAPIII transcription machinery were fused to the DNA binding domain (BD) of the Gal4 transcription factor (Flores et al., 1999) and transformed into a yeast strain containing a *lacZ* reporter gene, while Sen1 protein or its NTD were fused with the activation domain (AD) of Gal4 and transformed into a strain carrying a *HIS3* reporter gene. These two strains were mated and then tested for β -galactosidase activity as well as for their resistance to the 3-AT drug (inhibitor of the *HIS3* reporter gene). As shown in Figure S9, the interaction of C82 with C34, C31 with C34 and TFIIB70 with C34 were clearly observed by both assays, consistent with former results obtained by a similar two hybrid system (Flores et al., 1999). Note that one of the two replicates for interactions with C34 did not give the expected result, possibly because of a problem with this particular clone. If the experiment were to be pursued other replicates would be obviously needed. However, we didn't detect the interaction of Sen1 or the NTD with any of the RNAPIII subunits used, or with the replisome component Ctf4 (Figure S9). The ABC23 subunit appeared to interact with Sen1 NTD only in the assay scoring for *HIS3* expression but not for β -galactosidase activity, which casts a doubt on this result. Also, we know that only the interaction with RNAPIII is lost in a *sen1-3* mutant, which suggests that the mediator of this interaction is unlikely to be a common RNAP subunit.

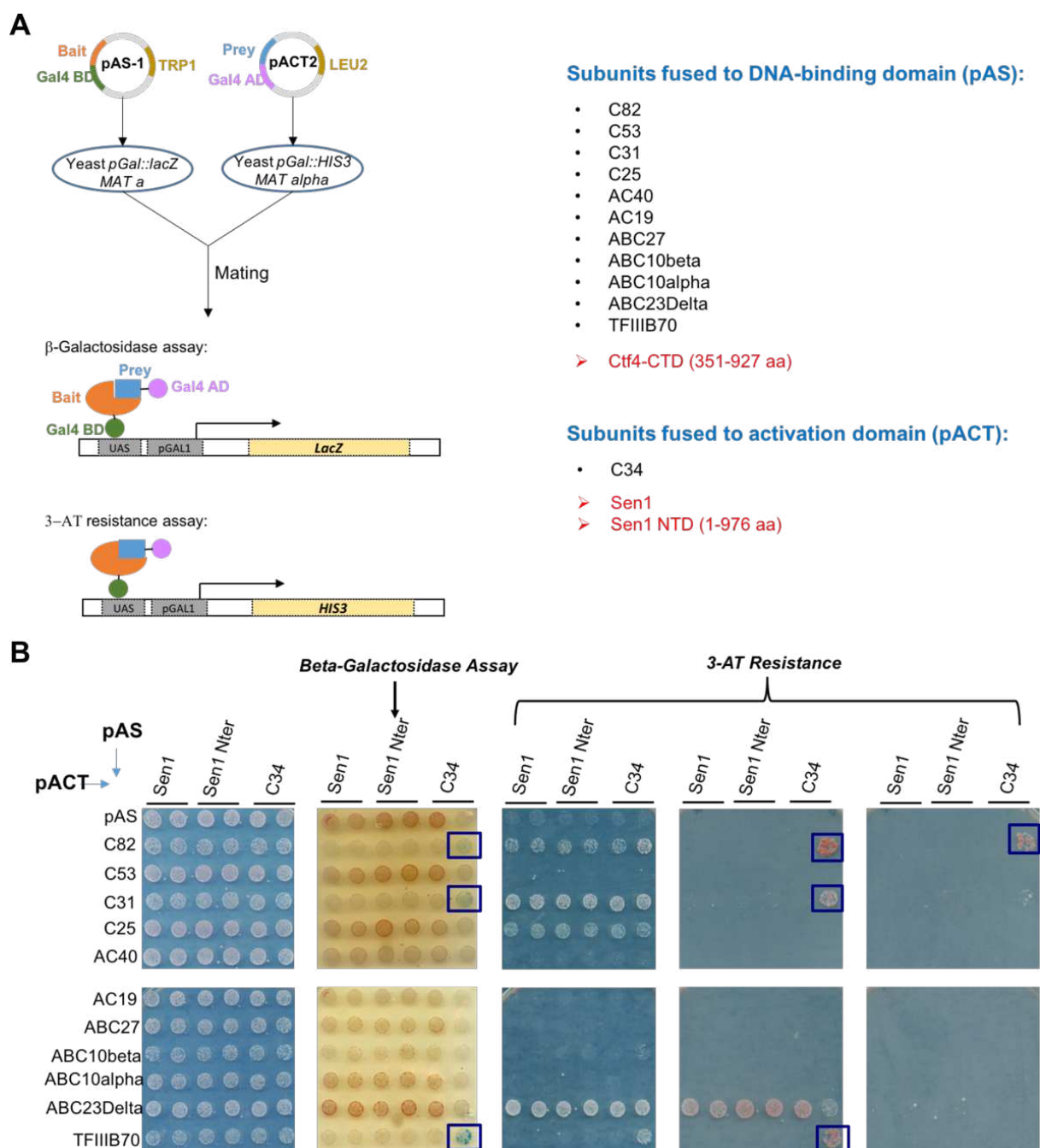


Figure S9: Yeast two hybrid assays to identify RNAPIII components interacting with Sen1.

(A) Scheme showing the strategies of two hybrid screening. Plasmids containing RNAPIII subunits were from Flores et al., 1999. Plasmids containing Sen1 and Ctf4 (indicated in red color) were constructed in this study. Only the CTD of Ctf4, which contains the Sen1-interaction region was used as putative positive control (Appanah et al., 2020) (B) Assays of the two hybrid screens. Positive results from both methods are highlighted in boxes.

While these experiments were ongoing, we learned that our collaborator, G. De Piccoli had committed a large-scale two-hybrid screen to a company that is a leader in this field, Hybrigenix, which failed to identify any interaction of Sen1 with any subunit of RNAPIII or with

the replisome, which might suggest that the Sen1-Gal4 fusion protein does not fold properly or is not stable enough for these assays. Alternatively, Sen1 might associate with a composite surface of RNAPIII involving several polymerase subunits and, therefore, the interaction with separate subunits could not be efficiently detected. It is also possible that the presence of RNA might be important to stabilize this interaction as we observed a decreased level of RNAPIII-bound Sen1 upon RNase A treatment in colP experiments (Figure 1E).

Crosslinking of reconstituted Sen1-RNAPIII complexes coupled to mass spectrometry analysis might be an alternative approach to identify the precise interacting peptides of RNAPIII and Sen1. However, those experiments require large amounts of purified proteins, and the protein preparations we possessed for this study were only sufficient for the experiments shown in the research manuscript, which we decided to prioritize. Our collaborators from the group of C. Muller will perform in the future a structural analysis of RNAPIII-Sen1 complex, which, if successful, will inform about the interaction surfaces.

How and when is Sen1 recruited to RNAPIII?

As introduced before, Sen1 functions in the termination of RNAPII transcription in the context of the NNS pathway. In short, Nrd1 and Nab3 within the NNS complex are initially recruited via both the interaction with the Ser5P CTD of RNAPII and specific sequence motifs on the target RNA, which would enhance the recruitment of Sen1 (see details in section 2.5.2). Sen1 itself can also recognize the Ser5P CTD, which is important for efficient termination (Han et al., 2020). Unlike in RNAPII termination, our data showed that the action of Sen1 in termination of RNAPIII transcription is independent of the NNS complex. Moreover, RNAPIII does not have a CTD that is dynamically phosphorylated and recognized by various of factors. So far, it is not clear what kind of signal triggers the recruitment of Sen1 to the RNAPIII transcription machinery.

Our genome-wide studies of RNAPIII transcription together with northern blot analyses of different tRNA species suggest that the moment when Sen1 acts in RNAPIII termination is when RNAPIII have read through the primary terminator, and Sen1 acts as a fail-safe transcription termination factor. Our data suggest that the activity of Sen1 in primary termination is hindered by the presence of tRNA secondary structures. However, our data do not distinguish at which moment Sen1 associates with RNAPIII. It is possible that Sen1 binds to RNAPIII that have escaped primary termination and paused at downstream regions. Alternatively, Sen1 could associate with RNAPIII at the very beginning of the transcription, travel along with RNAPIII and load onto the nascent RNA only when there is a sufficiently long portion of unstructured RNA available, which tends to happen at readthrough regions.

In our Sen1 CRAC data (Han et al., 2020), we observe significant Sen1 signal downstream of the 3' end of RNAPIII-dependent genes but also within the body of RNAPIII-dependent genes. The latter signal is not fully reliable because highly abundant RNAs are common contaminants in this kind of experiments, though the signal in the bodies remained much higher than in the uncrosslinked controls, indicating that part of the signal observed could correspond to bona fide RNAPIII signal. Thus, a different method would be needed to map more reliably the position of Sen1 within and around class III genes, which would help us to understand if the association of Sen1 with RNAPIII occurs at early stages of transcription.

Structural differences between *S. cerevisiae* and *S. pombe* Sen1 proteins: an implication in functional divergence?

One of the main discoveries of our work is that *S. cerevisiae* Sen1 (ScSen1) can directly induce RNAPIII termination. During the course of my PhD, another study has shown that *S. pombe* Sen1 (SpSen1) also interacts with RNAPIII and is required for efficient termination of RNAPIII transcription (Rivosecchi et al., 2019). There are several features shared by ScSen1 and SpSen1. For instance, they are both constituted by a large NTD and a helicase domain, though significant sequence homology can be found only in their helicase domains (Figure S10). Both proteins were shown to translocate along single-stranded nucleic acids in the 5' to 3' direction and to be able to unwind DNA:DNA and DNA:RNA duplexes *in vitro* (Han et al., 2017; Kim et al., 1999). In addition, the ATP hydrolysis activity of ScSen1 and SpSen1 were found to be essential for their function in RNAPIII termination (this thesis and Rivosecchi thesis, 2019).

However, as discussed in the research manuscript, significant functional differences do exist between these two proteins. Contrary to ScSen1, neither SpSen1 nor the other Sen1 homolog in fission yeast (Dlb8) are essential for viability or for RNAPII transcription termination (Larochelle et al., 2018). Furthermore, none of the *S. pombe* Sen1 homologues form a stable complex with the homologues of *S. cerevisiae* Nrd1 (Seb1) and Nab3 (Larochelle et al., 2018; Legros et al., 2014; Lemay et al., 2016). With regards to their function in RNAPIII transcription termination, SpSen1 was implicated in primary termination (Rivosecchi et al., 2019), whereas our data strongly support the notion that ScSen1 functions mainly on secondary termination.

What different properties of the two Sen1 proteins underlie the functional divergences between the two organisms? As mentioned before, very little if any sequence similarity can be found in the NTDs of ScSen1 and SpSen1, which appear to be protein-protein interaction hubs. Moreover, ScSen1 has an additional disordered CTD, which was shown to interact with Nrd1 and Glc7 (Figure S10; Han et al., 2020; Nedeja et al., 2008). Thus, does the

To approach these questions, it could be useful to compare the shape of the two proteins, which is closely linked with their function but, so far, only the structure of the helicase domain of ScSen1 is available (Leonaité et al., 2017). However, very recently, an artificial intelligence (AI) program AlphaFold developed by Google's DeepMind has been released, which is able to predict quite accurately a protein's structure from its amino acid sequence (Jumper et al., 2021). An AlphaFold database has been simultaneously built containing predicted structures for the human proteome and 20 other key organisms (<https://alphafold.ebi.ac.uk>). Thanks to this scientific advance, the 3D models of the structure of full-length Sen1 from *S. cerevisiae* and *S. pombe* are now available, as shown in Figure S11. The structure of the human homologue of Sen1, Senataxin (HsSETX), is also included for comparison. The experimentally-generated structure of ScSen1 helicase domain (HD) (Leonaité et al., 2017) has already been shown in Figure 2-8 (herein Sen1_{HD}).

When comparing the predicted structures (Figure S11A) with Sen1_{HD} (Figure 2-8B), they all have a similar organization as Sen1_{HD} in the helicase core including two RecA domains that are positioned side by side. One accessory subdomain, the "barrel", is extended on the surface of RecA1. The "stalk", containing two antiparallel helices, is also well recognized. Another accessory subdomain, the "prong", is formed by two long antiparallel α -helices protruding from the helicase core. Importantly, structural comparisons of ScSen1 with SpSen1 reveal distinct features in the orientations of the "stalk" and "prong", especially for the latter. In the structure of ScSen1, the "prong" is extending towards the NTD. However, in SpSen1, the "prong" is rotated about 90° toward the RecA2, which is opposite to the face encountering the polymerase. This different positioning of the "prong" in SpSen1 relative to ScSen1 should be confirmed by *bona fide* structural analyses, since the model provided by AlphaFold remains a prediction. However, it opens up the possibility that the different orientation of this subdomain provides different properties with respect to termination. Indeed, in *S. cerevisiae* the "prong" is essential for transcription termination by RNAPII (Leonaité et al., 2017) and RNAPIII (this study), and its position suggests that, upon Sen1 collision with RNAPIII, the "prong" could enter the RNA exit channel of the polymerase thus inducing conformational changes and termination. Leonaité et al. have also identified that another subdomain called "brace" (Figure 2-8 and Figure S10) is important for shaping a favourable conformation for RNA binding and unwinding by pulling the "barrel" towards the "prong".

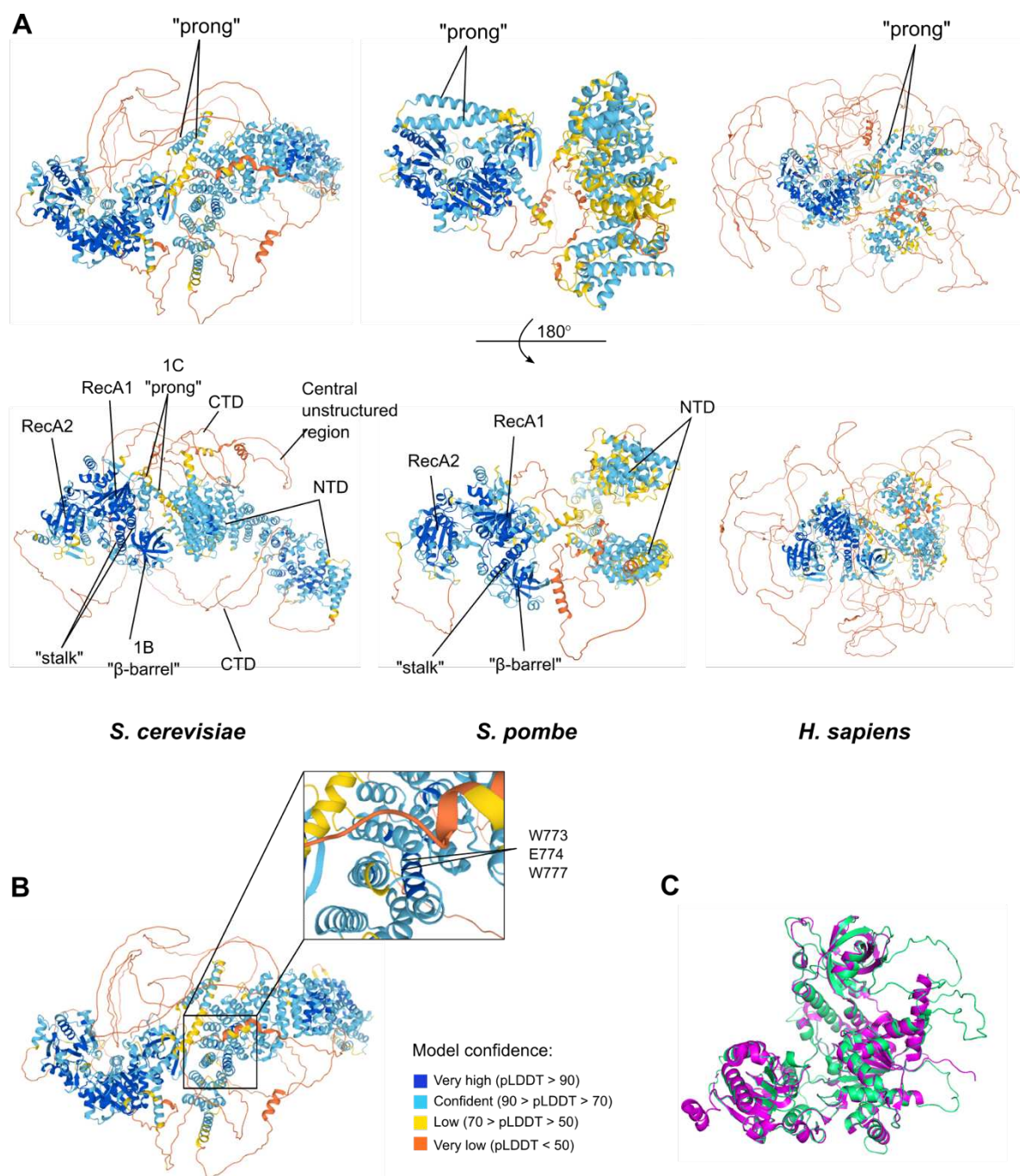


Figure S11: Structure of various Sen1 proteins predicted with AlphaFold .

(A) AlphaFold-predicted structures of Sen1 in *S. cerevisiae* (left, ID: AF-Q00416-F1), *S. pombe* (middle, ID: AF-Q92355-F1) and *H. sapiens* (right, ID: AF-Q7Z333-F1). (B) Zoom-in view (with the corresponding overall view) showing the three amino acids mutated in *S. cerevisiae* *sen1-3* variant. (Jumper et al., 2021). AlphaFold produces a per-residue confidence score (pLDDT) between 0 and 100. Some regions below 50 pLDDT may be unstructured in isolation. (C) A model of the structure of HsSTEX helicase domain (HD) (green) based on ScSen1 HD (Leonaité et al., 2017) (magenta). Generated by Richard Stefl and Marek Sebesta (CEITEC, Czech Republic) with HHPRED suite.

Another pronounced difference is the conformation of the NTD. As shown in Figure S11A, the NTD of ScSen1 is more extended and might be pushed by the “prong” as well. As introduced before, the NTD and CTD of ScSen1 mediate various protein-protein interactions (see section 2.5.2). On the contrary, in SpSen1, the NTD is well organized into a “clamp” shape. The difference in the NTD might result in a different way of recruitment by the RNAPIII and different mode of loading onto the nascent RNA.

The information we get from these predicted structures needs, of course, to be verified. In the future, it would be interesting to test if the “prong” and NTD of SpSen1 are required for its activity in RNAPIII termination. It will also be interesting to understand in more detail the function of SpSen1 unstructured regions.

The structure of HsSETX is poorly modeled, but the overall fold of the helicase domain is more similar to ScSen1, especially for the orientation of the “prong”, which might indicate a conserved function between the two proteins. This result is in agreement with another model made by our collaborators Marek Sebesta and Richard Stefl (CEITEC, Czech Republic), in which the two structures fit very well when overlaid, as shown in Figure S11C.

Senataxin was shown to play a role in transcription termination of at least a subset of RNAPII-transcribed genes (Suraweera et al., 2009; Skourti-Stathaki et al., 2011; Wagschal et al., 2012; Zhao et al., 2016). It has been proposed that, as Sen1, Senataxin is involved in the resolution of R-loops and that this activity is important for genome stability (Becherel et al., 2013; Skourti-Stathaki et al., 2011; Yüce and West, 2013). Senataxin has also been implicated in DNA repair (Andrews et al., 2018; Cohen et al., 2018) and in the resolution of transcription–replication conflicts (Richard et al., 2013; Yüce and West, 2013). Importantly, mutations in the most conserved regions of Senataxin, specifically the N-terminal and the helicase domains, are linked to two neurodegenerative disorders: amyotrophic lateral sclerosis type 4 (ALS4) and ocular ataxia-apraxia type 2 (AOA2) (Chen et al., 2004; Moreira et al., 2004; Bennett and La Spada, 2015; Groh et al., 2017). How distinct Senataxin mutations cause these diseases remains unclear. The introduction of AOA2-associated mutations in the equivalent residues of the budding yeast Sen1 provokes transcription termination defects both *in vivo* (Chen et al., 2014) and *in vitro* (Leonaitė et al., 2017), suggesting that the development of this disorder could be linked to some extent to termination defects.

In the future, it would be interesting to understand whether Senataxin does also play an important role in termination of RNAPIII transcription, and eventually whether a dysfunction in this process is linked to Senataxin-associated neurodegenerative disorders.

The positions of the mutations in Sen1-3 are also indicated in Figure S11B. The three mutated amino acids reside in a helix that is modelled with very high confidence. More importantly, this helix is placed on the base of the NTD, which could be important to sustain

the conformation of the NTD. Thus, mutating these residues might provoke conformational changes in the NTD, which in turn could affect RNAPIII and replisome binding. However, it is also possible that these residues mediate directly contacts with RNAPIII and the replisome. More work is needed in the future to distinguish between these possibilities.

Uneven distribution of RNAPIII along tRNAs: bias in data analysis?

The last question I would like to discuss is the distribution of RNAPIII during transcription. Turowski et al. (2016) observed that RNAPIIIIs were enriched at two regions corresponding to the beginning of the A box and the B box, respectively, which they proposed were caused by the binding of TFIIIC at promoters. They also detected much higher RNAPIII signals at the first peak (i.e. at the A box) and they suggested it could be due to a delay in the dissociation of RNAPIII from the transcription initiation factors (Figure 3-9). However, we do not observe the same pattern with our CRAC data (Figure 2A; Figure S12). Metagene analysis using the whole reads did not show two well-separated peaks and we rather observed an increased signal towards the 3' end of tRNA genes (Figure 2A). We also analysed the RNAPIII occupancy by mapping only the 3' end of each read, which can more precisely reflect the position of individual RNAPIIIIs with single-nucleotide resolution. Although the RNAPIII distribution seems to vary according to the tRNA isotype, in general, we observe an enrichment of RNAPIIIIs around the 3' end of the mature tRNAs, in other words, near the primary terminator of tRNA genes (Figure S12). This would be consistent with RNAPIII undergoing strong and relatively long-lived pausing at the primary terminator, as is expected. Indeed, correct recognition of termination signal and reinitiation by RNAPIII requires slowing down or/and pausing of the polymerase mediated by the C53-C37 subunits.

With regards to the 5' end bias observed by Turowski et al. (2016), it could actually be due to the way they processed their data. In their study, their sequencing procedure generated 50-bp single-end reads and the cDNA fragments were sequenced from the 5' end. Only reads including both the 5' and 3' linkers, which flank the cDNA, were analysed to ensure that bona fide 3' ends were mapped, which means a fraction of long reads containing real 3' end signals were discarded. Furthermore, reads corresponding to RNAPIIIIs transcribing the first 50 bp would be enriched as they would meet the criteria to be selected and used in their analyses. Therefore, the different RNAPIII distribution patterns originated by two different studies might be due to the different ways of data processing, which still needs to be verified.

CRAC of RNAPIII

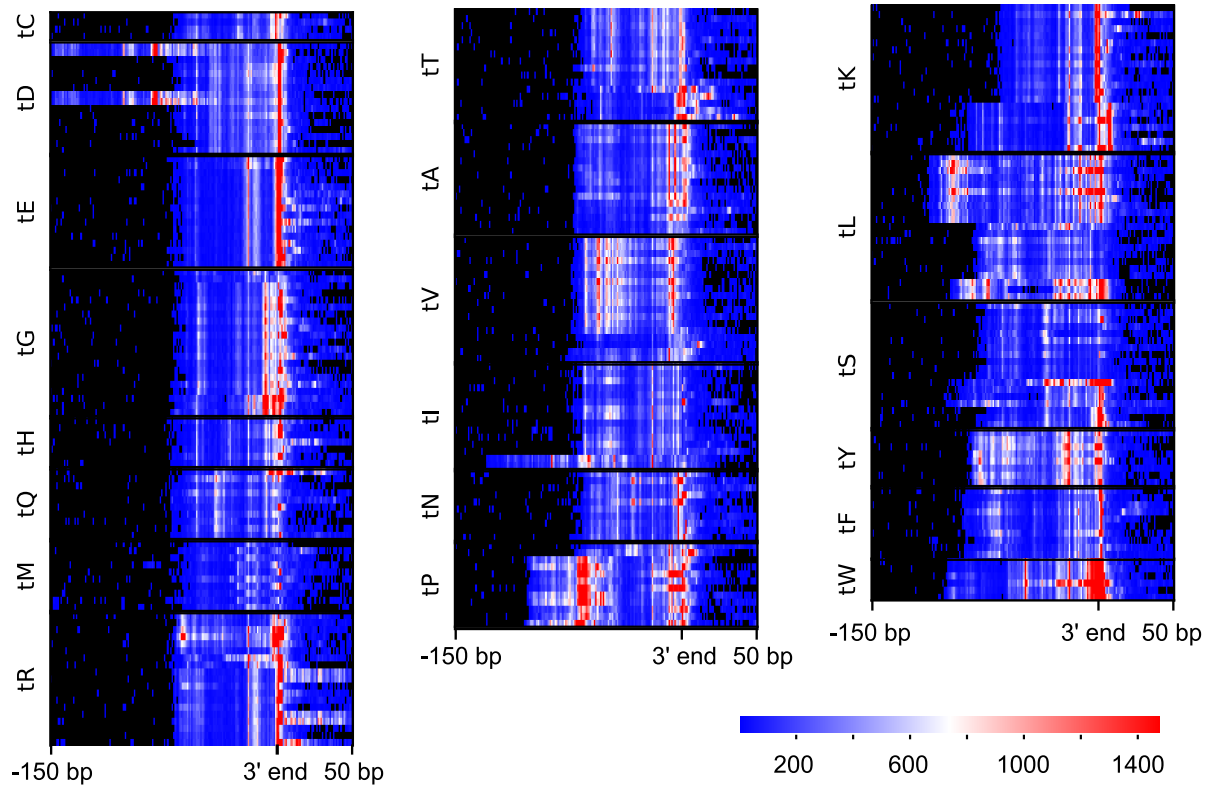


Figure S12: Heatmap analysis of the distribution of RNAPIII along tRNA genes.

Only 3' ends of the reads are analysed in the heatmap. All nuclear tRNA genes are aligned to the 3' end of the mature tRNAs, with 150 nt upstream and 50 nt downstream, and clustered by the amino acid isotype.

BIBLIOGRAPHY

- Abascal-Palacios, G., Ramsay, E. P., Beuron, F., Morris, E., & Vannini, A. (2018). Structural basis of RNA polymerase III transcription initiation. *Nature*, 553(7688), 301–306. <https://doi.org/10.1038/nature25441>
- Abou Elela, S., & Ares, M. (1998). Depletion of yeast RNase III blocks correct U2 3' end formation and results in polyadenylated but functional U2 snRNA. *The EMBO Journal*, 17(13), 3738–3746. <https://doi.org/10.1093/emboj/17.13.3738>
- Adeniyi-Jones, S., Romeo, P. H., & Zasloff, M. (1984). Generation of long read-through transcripts in vivo and in vitro by deletion of 3' termination and processing sequences in the human tRNA imet gene. *Nucleic Acids Research*, 12(2), 1101–1115. <https://doi.org/10.1093/nar/12.2.1101>
- Allison, D. S., & Hall, B. D. (1985). Effects of alterations in the 3' flanking sequence on in vivo and in vitro expression of the yeast SUP4-o tRNATyr gene. *The EMBO Journal*, 4(10), 2657–2664.
- Alzu, A., Bermejo, R., Begnis, M., Lucca, C., Piccini, D., Carotenuto, W., Saponaro, M., Brambati, A., Cocito, A., Foiani, M., & Liberi, G. (2012). Senataxin Associates with Replication Forks to Protect Fork Integrity across RNA-Polymerase-II-Transcribed Genes. *Cell*, 151(4), 835–846. <https://doi.org/10.1016/j.cell.2012.09.041>
- Andrews, A. M., McCartney, H. J., Errington, T. M., D'Andrea, A. D., & Macara, I. G. (2018). A senataxin-associated exonuclease SAN1 is required for resistance to DNA interstrand cross-links. *Nature Communications*, 9, 2592. <https://doi.org/10.1038/s41467-018-05008-8>
- Appanah, R., Lones, E. C., Aiello, U., Libri, D., & De Piccoli, G. (2020). Sen1 Is Recruited to Replication Forks via Ctf4 and Mrc1 and Promotes Genome Stability. *Cell Reports*, 30(7), 2094-2105.e9. <https://doi.org/10.1016/j.celrep.2020.01.087>
- Arigo, J. T., Carroll, K. L., Ames, J. M., & Corden, J. L. (2006). Regulation of yeast NRD1 expression by premature transcription termination. *Molecular Cell*, 21(5), 641–651. <https://doi.org/10.1016/j.molcel.2006.02.005>
- Arimbasseri, A. G., Kassavetis, G. A., & Maraia, R. J. (2014). Comment on “Mechanism of eukaryotic RNA polymerase III transcription termination.” *Science*, 345(6196), 524–524. <https://doi.org/10.1126/science.1253783>
- Arimbasseri, A. G., & Maraia, R. J. (2013). Distinguishing Core and Holoenzyme Mechanisms of Transcription Termination by RNA Polymerase III. *Molecular and Cellular Biology*, 33(8), 1571–1581. <https://doi.org/10.1128/MCB.01733-12>
- Arimbasseri, A. G., & Maraia, R. J. (2015). Mechanism of Transcription Termination by RNA Polymerase III Utilizes a Non-template Strand Sequence-Specific Signal Element. *Molecular Cell*, 58(6), 1124–1132. <https://doi.org/10.1016/j.molcel.2015.04.002>
- Arimbasseri, A. G., & Maraia, R. J. (2016). A high density of cis-information terminates RNA Polymerase III on a 2-rail track. *RNA Biology*, 13(2), 166–171. <https://doi.org/10.1080/15476286.2015.1116677>
- Arimbasseri, A. G., Rijal, K., & Maraia, R. J. (2013). Transcription termination by the eukaryotic RNA polymerase III. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1829(3), 318–330. <https://doi.org/10.1016/j.bbagr.2012.10.006>
- Arndt, K. M., & Reines, D. (2015). Termination of Transcription of Short Noncoding RNAs by RNA Polymerase II. *Annual Review of Biochemistry*, 84(1), 381–404. <https://doi.org/10.1146/annurev-biochem-060614-034457>
- Baejen, C., Andreani, J., Torkler, P., Battaglia, S., Schwalb, B., Lidschreiber, M., Maier, K. C., Boltendahl, A., Rus, P., Esslinger, S., Söding, J., & Cramer, P. (2017). Genome-wide Analysis of RNA Polymerase II Termination at Protein-Coding Genes. *Molecular Cell*, 66(1), 38-49.e6. <https://doi.org/10.1016/j.molcel.2017.02.009>
- Baer, M. F., Wesolowski, D., & Altman, S. (1989). Characterization in vitro of the defect in a temperature-sensitive mutant of the protein subunit of RNase P from Escherichia coli. *Journal of Bacteriology*, 171(12), 6862–6866. <https://doi.org/10.1128/jb.171.12.6862-6866.1989>
- Baker, R. E., Camier, S., Sentenac, A., & Hall, B. D. (1987). Gene size differentially affects the binding of yeast transcription factor tau to two intragenic regions. *Proceedings of the National Academy of Sciences of the United States of America*, 84(24), 8768–8772. <https://doi.org/10.1073/pnas.84.24.8768>
- Bardeleben, C., Kassavetis, G. A., & Geiduschek, E. P. (1994). Encounters of Saccharomyces cerevisiae RNA Polymerase III with its Transcription Factors during RNA Chain Elongation. *Journal of Molecular Biology*, 235(4), 1193–1205. <https://doi.org/10.1006/jmbi.1994.1073>
- Bartholomew, B., Durkovich, D., Kassavetis, G. A., & Geiduschek, E. P. (1993). Orientation and topography of RNA polymerase III in transcription complexes. *Molecular and Cellular Biology*, 13(2), 942–952.

- Bartholomew, B., Kassavetis, G. A., Braun, B. R., & Geiduschek, E. P. (1990a). The subunit structure of *Saccharomyces cerevisiae* transcription factor IIIC probed with a novel photocrosslinking reagent. *The EMBO Journal*, *9*(7), 2197–2205.
- Bartholomew, B., Kassavetis, G. A., Braun, B. R., & Geiduschek, E. P. (1990b). The subunit structure of *Saccharomyces cerevisiae* transcription factor IIIC probed with a novel photocrosslinking reagent. *The EMBO Journal*, *9*(7), 2197–2205.
- Bartholomew, B., Kassavetis, G. A., & Geiduschek, E. P. (1991). Two components of *Saccharomyces cerevisiae* transcription factor IIIB (TFIIIB) are stereospecifically located upstream of a tRNA gene and interact with the second-largest subunit of TFIIIC. *Molecular and Cellular Biology*, *11*(10), 5181–5189. <https://doi.org/10.1128/mcb.11.10.5181-5189.1991>
- Basehoar, A. D., Zanton, S. J., & Pugh, B. F. (2004). Identification and distinct regulation of yeast TATA box-containing genes. *Cell*, *116*(5), 699–709. [https://doi.org/10.1016/s0092-8674\(04\)00205-3](https://doi.org/10.1016/s0092-8674(04)00205-3)
- Bataille, A. R., Jeronimo, C., Jacques, P.-É., Laramée, L., Fortin, M.-È., Forest, A., Bergeron, M., Hanes, S. D., & Robert, F. (2012). A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. *Molecular Cell*, *45*(2), 158–170. <https://doi.org/10.1016/j.molcel.2011.11.024>
- Becherel, O. J., Yeo, A. J., Stellati, A., Heng, E. Y. H., Luff, J., Suraweera, A. M., Woods, R., Fleming, J., Carrie, D., McKinney, K., Xu, X., Deng, C., & Lavin, M. F. (2013). Senataxin plays an essential role with DNA damage response proteins in meiotic recombination and gene silencing. *PLoS Genetics*, *9*(4), e1003435. <https://doi.org/10.1371/journal.pgen.1003435>
- Bennett, C. L., & La Spada, A. R. (2015). Unwinding the role of senataxin in neurodegeneration. *Discovery Medicine*, *19*(103), 127–136.
- Berger, W., Steiner, E., Grusch, M., Elbling, L., & Micksche, M. (2009). Vaults and the major vault protein: Novel roles in signal pathway regulation and immunity. *Cellular and Molecular Life Sciences: CMLS*, *66*(1), 43–61. <https://doi.org/10.1007/s00018-008-8364-z>
- Bobkova, E. V., Habib, N., Alexander, G., & Hall, B. D. (1999). Mutational analysis of the hydrolytic activity of yeast RNA polymerase III. *The Journal of Biological Chemistry*, *274*(30), 21342–21348. <https://doi.org/10.1074/jbc.274.30.21342>
- Bogenhagen, D. F., & Brown, D. D. (1981). Nucleotide sequences in *Xenopus* 5S DNA required for transcription termination. *Cell*, *24*(1), 261–270. [https://doi.org/10.1016/0092-8674\(81\)90522-5](https://doi.org/10.1016/0092-8674(81)90522-5)
- Bohnsack, M. T., Tollervey, D., & Granneman, S. (2012). Identification of RNA helicase target sites by UV cross-linking and analysis of cDNA. *Methods in Enzymology*, *511*, 275–288. <https://doi.org/10.1016/B978-0-12-396546-2.00013-9>
- Boyer, L. A., Latek, R. R., & Peterson, C. L. (2004). The SANT domain: A unique histone-tail-binding module? *Nature Reviews Molecular Cell Biology*, *5*(2), 158–163. <https://doi.org/10.1038/nrm1314>
- Braglia, P., Percudani, R., & Dieci, G. (2005). Sequence context effects on oligo(dT) termination signal recognition by *Saccharomyces cerevisiae* RNA polymerase III. *The Journal of Biological Chemistry*. <https://doi.org/10.1074/jbc.M412238200>
- Brandl, C. J., & Struhl, K. (1990). A nucleosome-positioning sequence is required for GCN4 to activate transcription in the absence of a TATA element. *Molecular and Cellular Biology*, *10*(8), 4256–4265. <https://doi.org/10.1128/mcb.10.8.4256-4265.1990>
- Breathnach, R., & Chambon, P. (1981). Organization and Expression of Eucaryotic Split Genes Coding for Proteins. *Annual Review of Biochemistry*, *50*(1), 349–383. <https://doi.org/10.1146/annurev.bi.50.070181.002025>
- Brennan, C. A., Dombroski, A. J., & Platt, T. (1987). Transcription termination factor rho is an RNA-DNA helicase. *Cell*, *48*(6), 945–952. [https://doi.org/10.1016/0092-8674\(87\)90703-3](https://doi.org/10.1016/0092-8674(87)90703-3)
- Briand, J.-F., Navarro, F., Gadal, O., & Thuriaux, P. (2001). Cross Talk between tRNA and rRNA Synthesis in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, *21*(1), 189–195. <https://doi.org/10.1128/MCB.21.1.189-195.2001>
- Brow, D. A., & Guthrie, C. (1988). Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. *Nature*, *334*(6179), 213–218. <https://doi.org/10.1038/334213a0>

- Brun, I., Sentenac, A., & Werner, M. (1997). Dual role of the C34 subunit of RNA polymerase III in transcription initiation. *The EMBO Journal*, 16(18), 5730–5741. <https://doi.org/10.1093/emboj/16.18.5730>
- Buratowski, S., Hahn, S., Guarente, L., & Sharp, P. A. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell*, 56(4), 549–561. [https://doi.org/10.1016/0092-8674\(89\)90578-3](https://doi.org/10.1016/0092-8674(89)90578-3)
- Burgess, R. R., Travers, A. A., Dunn, J. J., & Bautz, E. K. F. (1969). Factor Stimulating Transcription by RNA Polymerase. *Nature*, 221(5175), 43–46. <https://doi.org/10.1038/221043a0>
- Butler, J. S. (2002). The yin and yang of the exosome. *Trends in Cell Biology*, 12(2), 90–96. [https://doi.org/10.1016/S0962-8924\(01\)02225-5](https://doi.org/10.1016/S0962-8924(01)02225-5)
- Cabart, P., Lee, J., & Willis, I. M. (2008). Facilitated recycling protects human RNA polymerase III from repression by Maf1 in vitro. *The Journal of Biological Chemistry*, 283(52), 36108–36117. <https://doi.org/10.1074/jbc.M807538200>
- Candelli, T., Challal, D., Briand, J.-B., Boulay, J., Porrua, O., Colin, J., & Libri, D. (2018). High-resolution transcription maps reveal the widespread impact of roadblock termination in yeast. *The EMBO Journal*, 37(4), e97490. <https://doi.org/10.15252/embj.201797490>
- Canella, D., Praz, V., Reina, J. H., Cousin, P., & Hernandez, N. (2010). Defining the RNA polymerase III transcriptome: Genome-wide localization of the RNA polymerase III transcription machinery in human cells. *Genome Research*, 20(6), 710–721. <https://doi.org/10.1101/gr.101337.109>
- Carroll, K. L., Pradhan, D. A., Granek, J. A., Clarke, N. D., & Corden, J. L. (2004). Identification of cis Elements Directing Termination of Yeast Nonpolyadenylated snoRNA Transcripts. *Molecular and Cellular Biology*, 24(14), 6241–6252. <https://doi.org/10.1128/MCB.24.14.6241-6252.2004>
- Carter, R., & 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. *Molecular Biology and Evolution*, 27(5), 1035–1043. <https://doi.org/10.1093/molbev/msp316>
- Casañal, A., Kumar, A., Hill, C. H., Easter, A. D., Emsley, P., Degliesposti, G., Gordiyenko, Y., Santhanam, B., Wolf, J., Wiederhold, K., Dornan, G. L., Skehel, M., Robinson, C. V., & Passmore, L. A. (2017). Architecture of eukaryotic mRNA 3'-end processing machinery. *Science*, 358(6366), 1056–1059. <https://doi.org/10.1126/science.aao6535>
- Catala, M., Tremblay, M., Samson, E., Conconi, A., & Abou Elela, S. (2008). Deletion of Rnt1p alters the proportion of open versus closed rRNA gene repeats in yeast. *Molecular and Cellular Biology*, 28(2), 619–629. <https://doi.org/10.1128/MCB.01805-07>
- Challal, D. (2019). *Role of general regulatory factors in the control of gene expression and transcription fidelity* [These de doctorat, Université Paris-Saclay (ComUE)]. <http://www.theses.fr/2019SACL5176>
- Chan, P. P., & Lowe, T. M. (2016). GtRNADB 2.0: An expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Research*, 44(D1), D184–189. <https://doi.org/10.1093/nar/gkv1309>
- Chang, D. D., & Clayton, D. A. (1989). Mouse RNAase MRP RNA is encoded by a nuclear gene and contains a decamer sequence complementary to a conserved region of mitochondrial RNA substrate. *Cell*, 56(1), 131–139. [https://doi.org/10.1016/0092-8674\(89\)90991-4](https://doi.org/10.1016/0092-8674(89)90991-4)
- Chédin, S., Riva, M., Schultz, P., Sentenac, A., & Carles, C. (1998). The RNA cleavage activity of RNA polymerase III is mediated by an essential TFIIIS-like subunit and is important for transcription termination. *Genes & Development*, 12(24), 3857–3871. <https://doi.org/10.1101/gad.12.24.3857>
- Chen, X., Müller, U., Sundling, K. E., & Brow, D. A. (2014). *Saccharomyces cerevisiae* Sen1 as a Model for the Study of Mutations in Human Senataxin That Elicit Cerebellar Ataxia. *Genetics*, 198(2), 577–590. <https://doi.org/10.1534/genetics.114.167585>
- Chen, X., Qi, Y., Wu, Z., Wang, X., Li, J., Zhao, D., Hou, H., Li, Y., Yu, Z., Liu, W., Wang, M., Ren, Y., Li, Z., Yang, H., & Xu, Y. (2021). Structural insights into preinitiation complex assembly on core promoters. *Science*, 372(6541). <https://doi.org/10.1126/science.aba8490>
- Chen, Y.-Z., Bennett, C. L., Huynh, H. M., Blair, I. P., Puls, I., Irobi, J., Dierick, I., Abel, A., Kennerson, M. L., Rabin, B. A., Nicholson, G. A., Auer-Grumbach, M., Wagner, K., De Jonghe, P., Griffin, J. W., Fischbeck, K. H., Timmerman, V., Cornblath, D. R., & Chance, P. F. (2004). DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *American Journal of Human Genetics*, 74(6), 1128–1135.

- <https://doi.org/10.1086/421054>
- Cheung, A. C. M., & Cramer, P. (2011). Structural basis of RNA polymerase II backtracking, arrest and reactivation. *Nature*, 471(7337), 249–253. <https://doi.org/10.1038/nature09785>
- Chinchilla, K., Rodriguez-Molina, J. B., Ursic, D., Finkel, J. S., Ansari, A. Z., & Culbertson, M. R. (2012). Interactions of Sen1, Nrd1, and Nab3 with Multiple Phosphorylated Forms of the Rpb1 C-Terminal Domain in *Saccharomyces cerevisiae*. *Eukaryotic Cell*, 11(4), 417–429. <https://doi.org/10.1128/EC.05320-11>
- Chong, Y. T., Koh, J. L. Y., Friesen, H., Kaluarachchi Duffy, S., Cox, M. J., Moses, A., Moffat, J., Boone, C., & Andrews, B. J. (2015). Yeast Proteome Dynamics from Single Cell Imaging and Automated Analysis. *Cell*, 161(6), 1413–1424. <https://doi.org/10.1016/j.cell.2015.04.051>
- Chu, W.-M., Ballard, R. E., & Schmid, C. W. (1997). Palindromic sequences preceding the terminator increase polymerase III template activity. *Nucleic Acids Research*, 25(11), 2077–2082. <https://doi.org/10.1093/nar/25.11.2077>
- Ciampi, M. S. (2006). Rho-dependent terminators and transcription termination. *Microbiology (Reading, England)*, 152(Pt 9), 2515–2528. <https://doi.org/10.1099/mic.0.28982-0>
- Clapier, C. R., Iwasa, J., Cairns, B. R., & Peterson, C. L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nature Reviews Molecular Cell Biology*, 18(7), 407–422. <https://doi.org/10.1038/nrm.2017.26>
- Cohen, S., Puget, N., Lin, Y.-L., Clouaire, T., Aguirrebengoa, M., Rocher, V., Pasero, P., Canitrot, Y., & Legube, G. (2018). Senataxin resolves RNA:DNA hybrids forming at DNA double-strand breaks to prevent translocations. *Nature Communications*, 9, 533. <https://doi.org/10.1038/s41467-018-02894-w>
- Colin, J., Candelli, T., Porrúa, O., Boulay, J., Zhu, C., Lacroute, F., Steinmetz, L. M., & Libri, D. (2014). Roadblock Termination by Reb1p Restricts Cryptic and Readthrough Transcription. *Molecular Cell*, 56(5), 667–680. <https://doi.org/10.1016/j.molcel.2014.10.026>
- Conrad, N. K., Wilson, S. M., Steinmetz, E. J., Patturajan, M., Brow, D. A., Swanson, M. S., & Corden, J. L. (2000). A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II. *Genetics*, 154(2), 557–571.
- Cozzarelli, N. R., Gerrard, S. P., Schlissel, M., Brown, D. D., & Bogenhagen, D. F. (1983). Purified RNA polymerase III accurately and efficiently terminates transcription of 5s RNA genes. *Cell*, 34(3), 829–835. [https://doi.org/10.1016/0092-8674\(83\)90540-8](https://doi.org/10.1016/0092-8674(83)90540-8)
- Cramer, P. (2019). Organization and regulation of gene transcription. *Nature*, 573(7772), 45–54. <https://doi.org/10.1038/s41586-019-1517-4>
- Cramer, P., Bushnell, D. A., Fu, J., Gnatt, A. L., Maier-Davis, B., Thompson, N. E., Burgess, R. R., Edwards, A. M., David, P. R., & Kornberg, R. D. (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. *Science (New York, N.Y.)*, 288(5466), 640–649. <https://doi.org/10.1126/science.288.5466.640>
- Creamer, T. J., Darby, M. M., Jamonnak, N., Schaugency, P., Hao, H., Wheelan, S. J., & Corden, J. L. (2011). Transcriptome-Wide Binding Sites for Components of the *Saccharomyces cerevisiae* Non-Poly(A) Termination Pathway: Nrd1, Nab3, and Sen1. *PLOS Genetics*, 7(10), e1002329. <https://doi.org/10.1371/journal.pgen.1002329>
- Crick, F. H. (1958). On protein synthesis. *Symposia of the Society for Experimental Biology*, 12, 138–163.
- Das, A. K., Cohen, P. T. W., & Barford, D. (1998). The structure of the tetratricopeptide repeats of protein phosphatase 5: Implications for TPR-mediated protein–protein interactions. *The EMBO Journal*, 17(5), 1192–1199. <https://doi.org/10.1093/emboj/17.5.1192>
- Davis, C. A., & Ares, M. (2006). Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences*, 103(9), 3262–3267. <https://doi.org/10.1073/pnas.0507783103>
- Dechassa, M. L., Sabri, A., Pondugula, S., Kassabov, S. R., Chatterjee, N., Kladde, M. P., & Bartholomew, B. (2010). SWI/SNF has intrinsic nucleosome disassembly activity that is dependent on adjacent nucleosomes. *Molecular Cell*, 38(4), 590–602. <https://doi.org/10.1016/j.molcel.2010.02.040>
- Deprez, E., Arrebola, R., Conesa, C., & Sentenac, A. (1999). A subunit of yeast TFIIC participates in the recruitment of TATA-binding protein. *Molecular and Cellular Biology*, 19(12), 8042–8051.

- <https://doi.org/10.1128/MCB.19.12.8042>
- Dieci, G., Conti, A., Pagano, A., & Carnevali, D. (2013). Identification of RNA polymerase III-transcribed genes in eukaryotic genomes. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1829(3), 296–305. <https://doi.org/10.1016/j.bbagr.2012.09.010>
- Dieci, G., Fermi, B., & Bosio, M. C. (2014). Investigating transcription reinitiation through in vitro approaches. *Transcription*, 5(1), e27704. <https://doi.org/10.4161/trns.27704>
- Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M., & Pagano, A. (2007). The expanding RNA polymerase III transcriptome. *Trends in Genetics*, 23(12), 614–622. <https://doi.org/10.1016/j.tig.2007.09.001>
- Dieci, G., Giuliodori, S., Catellani, M., Percudani, R., & Ottonello, S. (2002). Intragenic promoter adaptation and facilitated RNA polymerase III recycling in the transcription of SCR1, the 7SL RNA gene of *Saccharomyces cerevisiae*. *The Journal of Biological Chemistry*, 277(9), 6903–6914. <https://doi.org/10.1074/jbc.M105036200>
- Dieci, G., & Sentenac, A. (1996). Facilitated Recycling Pathway for RNA Polymerase III. *Cell*, 84(2), 245–252. [https://doi.org/10.1016/S0092-8674\(00\)80979-4](https://doi.org/10.1016/S0092-8674(00)80979-4)
- Diribarne, G., & Bensaude, O. (2009). 7SK RNA, a non-coding RNA regulating P-TEFb, a general transcription factor. *Rna Biology*, 6(2), 122–128. <https://doi.org/10.4161/rna.6.2.8115>
- Donze, D. (2012). Extra-transcriptional functions of RNA Polymerase III complexes: TFIIC as a potential global chromatin bookmark. *Gene*, 493(2), 169–175. <https://doi.org/10.1016/j.gene.2011.09.018>
- Dosztányi, Z., Csizmok, V., Tompa, P., & Simon, I. (2005). IUPred: Web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics (Oxford, England)*, 21(16), 3433–3434. <https://doi.org/10.1093/bioinformatics/bti541>
- Dumay, H., Rubbi, L., Sentenac, A., & Marck, C. (1999). Interaction between Yeast RNA Polymerase III and Transcription Factor TFIIC via ABC10 α and τ 131 Subunits *. *Journal of Biological Chemistry*, 274(47), 33462–33468. <https://doi.org/10.1074/jbc.274.47.33462>
- Dumay-Odelot, H., Acker, J., Arrebola, R., Sentenac, A., & Marck, C. (2002). Multiple roles of the tau131 subunit of yeast transcription factor IIC (TFIIC) in TFIIB assembly. *Molecular and Cellular Biology*, 22(1), 298–308. <https://doi.org/10.1128/MCB.22.1.298-308.2002>
- Egloff, S., Dienstbier, M., & Murphy, S. (2012). Updating the RNA polymerase CTD code: Adding gene-specific layers. *Trends in Genetics: TIG*, 28(7), 333–341. <https://doi.org/10.1016/j.tig.2012.03.007>
- Ehrlich, R., Davyt, M., López, I., Chalar, C., & Marín, M. (2021). On the Track of the Missing tRNA Genes: A Source of Non-Canonical Functions? *Frontiers in Molecular Biosciences*, 8. <https://doi.org/10.3389/fmolb.2021.643701>
- Elela, S. A., Igel, H., & Ares, M. (1996). RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. *Cell*, 85(1), 115–124. [https://doi.org/10.1016/S0092-8674\(00\)81087-9](https://doi.org/10.1016/S0092-8674(00)81087-9)
- Fan, H., Sakulich, A. L., Goodier, J. L., Zhang, X., Qin, J., & Maraia, R. J. (1997). Phosphorylation of the human La antigen on serine 366 can regulate recycling of RNA polymerase III transcription complexes. *Cell*, 88(5), 707–715. [https://doi.org/10.1016/S0092-8674\(00\)81913-3](https://doi.org/10.1016/S0092-8674(00)81913-3)
- Fernández-Tornero, C., Böttcher, B., Rashid, U. J., Steuerwald, U., Flörchinger, B., Devos, D. P., Lindner, D., & Müller, C. W. (2010). Conformational flexibility of RNA polymerase III during transcriptional elongation. *The EMBO Journal*, 29(22), 3762–3772. <https://doi.org/10.1038/emboj.2010.266>
- Fernández-Tornero, C., Böttcher, B., Riva, M., Carles, C., Steuerwald, U., Ruigrok, R. W. H., Sentenac, A., Müller, C. W., & Schoehn, G. (2007). Insights into transcription initiation and termination from the electron microscopy structure of yeast RNA polymerase III. *Molecular Cell*, 25(6), 813–823. <https://doi.org/10.1016/j.molcel.2007.02.016>
- Ferrari, R., Rivetti, C., Acker, J., & Dieci, G. (2004). Distinct roles of transcription factors TFIIB and TFIIC in RNA polymerase III transcription reinitiation. *Proceedings of the National Academy of Sciences*, 101(37), 13442–13447. <https://doi.org/10.1073/pnas.0403851101>
- Ferri Maria-Laura, Peyroche Gérald, Siaut Magali, Lefebvre Olivier, Carles Christophe, Conesa Christine, & Sentenac André. (2000). A Novel Subunit of Yeast RNA Polymerase III Interacts with the TFIIB-Related Domain of TFIIB70. *Molecular and Cellular Biology*, 20(2), 488–495. <https://doi.org/10.1128/MCB.20.2.488-495.2000>

- Fica, S. M., Tuttle, N., Novak, T., Li, N.-S., Lu, J., Koodathingal, P., Dai, Q., Staley, J. P., & Piccirilli, J. A. (2013). RNA catalyses nuclear pre-mRNA splicing. *Nature*, *503*(7475), 229–234. <https://doi.org/10.1038/nature12734>
- Finkel, J. S., Chinchilla, K., Ursic, D., & Culbertson, M. R. (2010). Sen1p Performs Two Genetically Separable Functions in Transcription and Processing of U5 Small Nuclear RNA in *Saccharomyces cerevisiae*. *Genetics*, *184*(1), 107–118. <https://doi.org/10.1534/genetics.109.110031>
- Fish, R. N., & Kane, C. M. (2002). Promoting elongation with transcript cleavage stimulatory factors. *Biochimica Et Biophysica Acta*, *1577*(2), 287–307. [https://doi.org/10.1016/s0167-4781\(02\)00459-1](https://doi.org/10.1016/s0167-4781(02)00459-1)
- Flores, A., Briand, J. F., Gadal, O., Andrau, J. C., Rubbi, L., Van Mullem, V., Boschiero, C., Goussot, M., Marck, C., Carles, C., Thuriaux, P., Sentenac, A., & Werner, M. (1999). A protein-protein interaction map of yeast RNA polymerase III. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(14), 7815–7820. <https://doi.org/10.1073/pnas.96.14.7815>
- Fouqueau, T., Blombach, F., Cackett, G., Carty, A. E., Matelska, D. M., Ofer, S., Pilotto, S., Phung, D. K., & Werner, F. (2018). The cutting edge of archaeal transcription. *Emerging Topics in Life Sciences*, *2*(4), 517–533. <https://doi.org/10.1042/ETLS20180014>
- Gabrielsen, O. S., Marzouki, N., Ruet, A., Sentenac, A., & Fromageot, P. (1989). Two polypeptide chains in yeast transcription factor tau interact with DNA. *The Journal of Biological Chemistry*, *264*(13), 7505–7511.
- Gajiwala, K. S., & Burley, S. K. (2000). Winged helix proteins. *Current Opinion in Structural Biology*, *10*(1), 110–116. [https://doi.org/10.1016/S0959-440X\(99\)00057-3](https://doi.org/10.1016/S0959-440X(99)00057-3)
- Galardi, S., Fatica, A., Bachi, A., Scaloni, A., Presutti, C., & Bozzoni, I. (2002). Purified Box C/D snoRNPs Are Able To Reproduce Site-Specific 2'-O-Methylation of Target RNA In Vitro. *Molecular and Cellular Biology*, *22*(19), 6663–6668. <https://doi.org/10.1128/MCB.22.19.6663-6668.2002>
- Geiduschek, E., & Tocchinvalentini, G. (1988). Transcription by Rna Polymerase-III. *Annual Review of Biochemistry*, *57*, 873–914. <https://doi.org/10.1146/annurev.bi.57.070188.004301>
- Geiger, J. H., Hahn, S., Lee, S., & Sigler, P. B. (1996). Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science (New York, N.Y.)*, *272*(5263), 830–836. <https://doi.org/10.1126/science.272.5263.830>
- Geiger, S. R., Lorenzen, K., Schrieck, A., Hanecker, P., Kostrewa, D., Heck, A. J. R., & Cramer, P. (2010). RNA polymerase I contains a TFIIF-related DNA-binding subcomplex. *Molecular Cell*, *39*(4), 583–594. <https://doi.org/10.1016/j.molcel.2010.07.028>
- Gerlach, V. L., Whitehall, S. K., Geiduschek, E. P., & Brow, D. A. (1995). TFIIB placement on a yeast U6 RNA gene in vivo is directed primarily by TFIIC rather than by sequence-specific DNA contacts. *Molecular and Cellular Biology*, *15*(3), 1455–1466.
- Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'Shea, E. K., & Weissman, J. S. (2003). Global analysis of protein expression in yeast. *Nature*, *425*(6959), 737–741. <https://doi.org/10.1038/nature02046>
- Ghazal, G., Gagnon, J., Jacques, P.-É., Landry, J.-R., Robert, F., & Elela, S. A. (2009). Yeast RNase III Triggers Polyadenylation-Independent Transcription Termination. *Molecular Cell*, *36*(1), 99–109. <https://doi.org/10.1016/j.molcel.2009.07.029>
- Ginsburg, D. S., Govind, C. K., & Hinnebusch, A. G. (2009). NuA4 Lysine Acetyltransferase Esa1 Is Targeted to Coding Regions and Stimulates Transcription Elongation with Gcn5. *Molecular and Cellular Biology*, *29*(24), 6473–6487. <https://doi.org/10.1128/MCB.01033-09>
- Girbig, M., Misiaszek, A. D., Vorländer, M. K., Lafita, A., Grötsch, H., Baudin, F., Bateman, A., & Müller, C. W. (2021). Cryo-EM structures of human RNA polymerase III in its unbound and transcribing states. *Nature Structural & Molecular Biology*, *28*(2), 210–219. <https://doi.org/10.1038/s41594-020-00555-5>
- Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H., & Oliver, S. G. (1996). Life with 6000 Genes. *Science*, *274*(5287), 546–567. <https://doi.org/10.1126/science.274.5287.546>
- Goldberg, M. L. (1979). *Sequence analysis of Drosophila histone genes*. Stanford University.
- Goodier, J. L., Fan, H., & Maraia, R. J. (1997). A carboxy-terminal basic region controls RNA polymerase III transcription factor activity of human La protein. *Molecular and Cellular Biology*, *17*(10), 5823–5832.
- Goodier, J. L., & Maraia, R. J. (1998a). Terminator-specific recycling of a B1-Alu transcription complex by RNA

- polymerase III is mediated by the RNA terminus-binding protein La. *Journal of Biological Chemistry*, 273(40), 26110–26116. <https://doi.org/10.1074/jbc.273.40.26110>
- Goodier, J. L., & Maraia, R. J. (1998b). Terminator-specific recycling of a B1-Alu transcription complex by RNA polymerase III is mediated by the RNA terminus-binding protein La. *The Journal of Biological Chemistry*, 273(40), 26110–26116. <https://doi.org/10.1074/jbc.273.40.26110>
- Gottlieb, E., & Steitz, J. A. (1989). The RNA binding protein La influences both the accuracy and the efficiency of RNA polymerase III transcription in vitro. *The EMBO Journal*, 8(3), 841–850.
- Gouge, J., Guthertz, N., Kramm, K., Dergai, O., Abascal-Palacios, G., Satia, K., Cousin, P., Hernandez, N., Grohmann, D., & Vannini, A. (2017). Molecular mechanisms of Bdp1 in TFIIIB assembly and RNA polymerase III transcription initiation. *Nature Communications*, 8(1), 130. <https://doi.org/10.1038/s41467-017-00126-1>
- Govind, C. K., Zhang, F., Qiu, H., Hofmeyer, K., & Hinnebusch, A. G. (2007). Gcn5 Promotes Acetylation, Eviction, and Methylation of Nucleosomes in Transcribed Coding Regions. *Molecular Cell*, 25(1), 31–42. <https://doi.org/10.1016/j.molcel.2006.11.020>
- Granneman, S., Kudla, G., Petfalski, E., & Tollervey, D. (2009). Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proceedings of the National Academy of Sciences*, 106(24), 9613–9618.
- Groh, M., Albulescu, L. O., Cristini, A., & Gromak, N. (2017). Senataxin: Genome Guardian at the Interface of Transcription and Neurodegeneration. *Journal of Molecular Biology*, 429(21), 3181–3195. <https://doi.org/10.1016/j.jmb.2016.10.021>
- Gross, S., & Moore, C. (2001). Five subunits are required for reconstitution of the cleavage and polyadenylation activities of *Saccharomyces cerevisiae* cleavage factor I. *Proceedings of the National Academy of Sciences*, 98(11), 6080–6085. <https://doi.org/10.1073/pnas.101046598>
- Guffanti, E., Percudani, R., Harismendy, O., Soutourina, J., Werner, M., Iacovella, M. G., Negri, R., & Dieci, G. (2006). Nucleosome depletion activates poised RNA polymerase III at unconventional transcription sites in *Saccharomyces cerevisiae*. *The Journal of Biological Chemistry*, 281(39), 29155–29164. <https://doi.org/10.1074/jbc.M600387200>
- Gusarov, I., & Nudler, E. (1999). The Mechanism of Intrinsic Transcription Termination. *Molecular Cell*, 3(4), 495–504. [https://doi.org/10.1016/S1097-2765\(00\)80477-3](https://doi.org/10.1016/S1097-2765(00)80477-3)
- Haag, J. R., & Pikaard, C. S. (2011). Multisubunit RNA polymerases IV and V: Purveyors of non-coding RNA for plant gene silencing. *Nature Reviews Molecular Cell Biology*, 12(8), 483–492. <https://doi.org/10.1038/nrm3152>
- Hage, A. E., Koper, M., Kufel, J., & Tollervey, D. (2008). Efficient termination of transcription by RNA polymerase I requires the 5' exonuclease Rat1 in yeast. *Genes & Development*, 22(8), 1069–1081. <https://doi.org/10.1101/gad.463708>
- Hahn, S. (2004). Structure and mechanism of the RNA Polymerase II transcription machinery. *Nature Structural & Molecular Biology*, 11(5), 394–403. <https://doi.org/10.1038/nsmb763>
- Hahn, S., & 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. <https://doi.org/10.1534/genetics.111.127019>
- Hall, A. E., Turnbull, C., & Dalmay, T. (2013). Y RNAs: Recent developments. *BioMolecular Concepts*, 4(2), 103–110. <https://doi.org/10.1515/bmc-2012-0050>
- Hamada, M., Sakulich, A. L., Koduru, S. B., & Maraia, R. J. (2000). Transcription termination by RNA polymerase III in fission yeast—A genetic and biochemically tractable model system. *Journal of Biological Chemistry*, 275(37), 29076–29081. <https://doi.org/10.1074/jbc.M003980200>
- Han, Z., Jasnovidova, O., Haidara, N., Tudek, A., Kubicek, K., Libri, D., Stefl, R., & Porrua, O. (2020). Termination of non-coding transcription in yeast relies on both an RNA Pol II CTD interaction domain and a CTD-mimicking region in Sen1. *The EMBO Journal*, 39(7). <https://doi.org/10.15252/embj.2019101548>
- Han, Z., Libri, D., & Porrua, O. (2017). Biochemical characterization of the helicase Sen1 provides new insights into the mechanisms of non-coding transcription termination. *Nucleic Acids Research*, 45(3), 1355–1370. <https://doi.org/10.1093/nar/gkw1230>
- Han, Z., & Porrua, O. (2018). Helicases as transcription termination factors: Different solutions for a common

- problem. *Transcription*, 9(3), 152–158. <https://doi.org/10.1080/21541264.2017.1361503>
- Hann, B. C., & Walter, P. (1991). The signal recognition particle in *S. cerevisiae*. *Cell*, 67(1), 131–144. [https://doi.org/10.1016/0092-8674\(91\)90577-1](https://doi.org/10.1016/0092-8674(91)90577-1)
- Harismendy, O., Gendrel, C.-G., Soularue, P., Gidrol, X., Sentenac, A., Werner, M., & Lefebvre, O. (2003). Genome-wide location of yeast RNA polymerase III transcription machinery. *The EMBO Journal*, 22(18), 4738–4747. <https://doi.org/10.1093/emboj/cdg466>
- Harlen, K. M., & Churchman, L. S. (2017). The code and beyond: Transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nature Reviews Molecular Cell Biology*, 18(4), 263–273. <https://doi.org/10.1038/nrm.2017.10>
- Harlen, K. M., Trotta, K. L., Smith, E. E., Mosaheb, M. M., Fuchs, S. M., & Churchman, L. S. (2016). Comprehensive RNA Polymerase II Interactomes Reveal Distinct and Varied Roles for Each Phospho-CTD Residue. *Cell Reports*, 15(10), 2147–2158. <https://doi.org/10.1016/j.celrep.2016.05.010>
- Hartley, P. D., & Madhani, H. D. (2009). Mechanisms that specify promoter nucleosome location and identity. *Cell*, 137(3), 445–458. <https://doi.org/10.1016/j.cell.2009.02.043>
- Hassan, A. H., Neely, K. E., & Workman, J. L. (2001). Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. *Cell*, 104(6), 817–827. [https://doi.org/10.1016/s0092-8674\(01\)00279-3](https://doi.org/10.1016/s0092-8674(01)00279-3)
- Hector, R. E., Nykamp, K. R., Dheur, S., Anderson, J. T., Non, P. J., Urbinati, C. R., Wilson, S. M., Minvielle-Sebastia, L., & Swanson, M. S. (2002). Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. *The EMBO Journal*, 21(7), 1800–1810. <https://doi.org/10.1093/emboj/21.7.1800>
- Henry, N. L., Sayre, M. H., & Kornberg, R. D. (1992). Purification and characterization of yeast RNA polymerase II general initiation factor g. *The Journal of Biological Chemistry*, 267(32), 23388–23392.
- Hirata, A., & Murakami, K. S. (2009). Archaeal RNA polymerase. *Current Opinion in Structural Biology*, 19(6), 724–731. <https://doi.org/10.1016/j.sbi.2009.10.006>
- Hoffmann, N. A., Jakobi, A. J., Moreno-Morcillo, M., Glatt, S., Kosinski, J., Hagen, W. J. H., Sachse, C., & Müller, C. W. (2015). Molecular structures of unbound and transcribing RNA polymerase III. *Nature*, 528(7581), 231–236. <https://doi.org/10.1038/nature16143>
- Hoffmann, N. A., Jakobi, A. J., Vorländer, M. K., Sachse, C., & Müller, C. W. (2016a). Transcribing RNA polymerase III observed by electron cryomicroscopy. *The FEBS Journal*, 283(15), 2811–2819. <https://doi.org/10.1111/febs.13732>
- Hoffmann, N. A., Sadian, Y., Tafur, L., Kosinski, J., & Mueller, C. W. (2016b). Specialization versus conservation: How Pol I and Pol III use the conserved architecture of the pre-initiation complex for specialized transcription. *Transcription-Austin*, 7(4), 127–132. <https://doi.org/10.1080/21541264.2016.1203628>
- Houalla, R., Devaux, F., Fatica, A., Kufel, J., Barrass, D., Torchet, C., & Tollervey, D. (2006). Microarray detection of novel nuclear RNA substrates for the exosome. *Yeast*, 23(6), 439–454. <https://doi.org/10.1002/yea.1369>
- Houseley, J., & Tollervey, D. (2006). Yeast Trf5p is a nuclear poly(A) polymerase. *EMBO Reports*, 7(2), 205–211. <https://doi.org/10.1038/sj.embor.7400612>
- Hsieh, F.-K., Kulaeva, O. I., Patel, S. S., Dyer, P. N., Luger, K., Reinberg, D., & Studitsky, V. M. (2013). Histone chaperone FACT action during transcription through chromatin by RNA polymerase II. *Proceedings of the National Academy of Sciences*, 110(19), 7654–7659. <https://doi.org/10.1073/pnas.1222198110>
- Hu, P., Wu, S., & Hernandez, N. (2003). A minimal RNA polymerase III transcription system from human cells reveals positive and negative regulatory roles for CK2. *Molecular Cell*, 12(3), 699–709. <https://doi.org/10.1016/j.molcel.2003.08.011>
- Huang, R. C., Maheshwari, N., & Bonner, J. (1960). Enzymatic synthesis of RNA. *Biochemical and Biophysical Research Communications*, 3, 689–694. [https://doi.org/10.1016/0006-291x\(60\)90088-7](https://doi.org/10.1016/0006-291x(60)90088-7)
- Huang, Y., Intine, R. V., Mozlin, A., Hasson, S., & Maraia, R. J. (2005). Mutations in the RNA Polymerase III Subunit Rpc11p That Decrease RNA 3' Cleavage Activity Increase 3'-Terminal Oligo(U) Length and La-Dependent tRNA Processing. *Molecular and Cellular Biology*, 25(2), 621–636. <https://doi.org/10.1128/MCB.25.2.621-636.2005>
- Huisinga, K. L., & Pugh, B. F. (2004). A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Molecular Cell*, 13(4), 573–585.

- [https://doi.org/10.1016/s1097-2765\(04\)00087-5](https://doi.org/10.1016/s1097-2765(04)00087-5)
- Hurwitz, J., Bresler, A., & Dinger, R. (1960). The enzymic incorporation of ribonucleotides into polyribonucleotides and the effect of DNA. *Biochemical and Biophysical Research Communications*, 3(1), 15–19. [https://doi.org/10.1016/0006-291X\(60\)90094-2](https://doi.org/10.1016/0006-291X(60)90094-2)
- Imbalzano, A. N., Kwon, H., Green, M. R., & Kingston, R. E. (1994). Facilitated binding of TATA-binding protein to nucleosomal DNA. *Nature*, 370(6489), 481–485. <https://doi.org/10.1038/370481a0>
- Jasiak, A. J., Armache, K.-J., Martens, B., Jansen, R.-P., & Cramer, P. (2006). Structural Biology of RNA Polymerase III: Subcomplex C17/25 X-Ray Structure and 11 Subunit Enzyme Model. *Molecular Cell*, 23(1), 71–81. <https://doi.org/10.1016/j.molcel.2006.05.013>
- Jenny, A., Minvielle-Sebastia, L., Preker, P. J., & Keller, W. (1996). Sequence Similarity Between the 73-Kilodalton Protein of Mammalian CPSF and a Subunit of Yeast Polyadenylation Factor I. *Science*, 274(5292), 1514–1517. <https://doi.org/10.1126/science.274.5292.1514>
- Jensen, T. H., Jacquier, A., & Libri, D. (2013). Dealing with Pervasive Transcription. *Molecular Cell*, 52(4), 473–484. <https://doi.org/10.1016/j.molcel.2013.10.032>
- Jeronimo, C., & Robert, F. (2014). Kin28 regulates the transient association of Mediator with core promoters. *Nature Structural & Molecular Biology*, 21(5), 449–455. <https://doi.org/10.1038/nsmb.2810>
- Joazeiro, C. A., Kassavetis, G. A., & Geiduschek, E. P. (1996). Alternative outcomes in assembly of promoter complexes: The roles of TBP and a flexible linker in placing TFIIB on tRNA genes. *Genes & Development*, 10(6), 725–739. <https://doi.org/10.1101/gad.10.6.725>
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 1–11. <https://doi.org/10.1038/s41586-021-03819-2>
- Jun, S.-H., Reichlen, M. J., Tajiri, M., & Murakami, K. S. (2011). Archaeal RNA polymerase and transcription regulation. *Critical Reviews in Biochemistry and Molecular Biology*, 46(1), 27–40. <https://doi.org/10.3109/10409238.2010.538662>
- Juven-Gershon, T., & Kadonaga, J. T. (2010). Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Developmental Biology*, 339(2), 225–229. <https://doi.org/10.1016/j.ydbio.2009.08.009>
- Kadonaga, J. T. (2012). Perspectives on the RNA polymerase II core promoter. *WIREs Developmental Biology*, 1(1), 40–51. <https://doi.org/10.1002/wdev.21>
- Kaneko, S., Rozenblatt-Rosen, O., Meyerson, M., & Manley, J. L. (2007). The multifunctional protein p54nrb/PSF recruits the exonuclease XRN2 to facilitate pre-mRNA 3' processing and transcription termination. *Genes & Development*, 21(14), 1779–1789. <https://doi.org/10.1101/gad.1565207>
- Kassavetis, G. A., Blanco, J. A., Johnson, T. E., & Geiduschek, E. P. (1992b). Formation of open and elongating transcription complexes by RNA polymerase III. *Journal of Molecular Biology*, 226(1), 47–58. [https://doi.org/10.1016/0022-2836\(92\)90123-2](https://doi.org/10.1016/0022-2836(92)90123-2)
- Kassavetis, G. A., Braun, B. R., Nguyen, L. H., & Peter Geiduschek, E. (1990). S. cerevisiae TFIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIA and TFIIC are assembly factors. *Cell*, 60(2), 235–245. [https://doi.org/10.1016/0092-8674\(90\)90739-2](https://doi.org/10.1016/0092-8674(90)90739-2)
- Kassavetis, G. A., Joazeiro, C. A. P., Pisano, M., Geiduschek, E. P., Colbert, T., Hahn, S., & Blanco, J. A. (1992a). The role of the TATA-binding protein in the assembly and function of the multisubunit yeast RNA polymerase III transcription factor, TFIIB. *Cell*, 71(6), 1055–1064. [https://doi.org/10.1016/0092-8674\(92\)90399-W](https://doi.org/10.1016/0092-8674(92)90399-W)
- Kassavetis, G. A., Kumar, A., Letts, G. A., & Geiduschek, E. P. (1998). A post-recruitment function for the RNA polymerase III transcription–initiation factor IIB. *Proceedings of the National Academy of Sciences*, 95(16), 9196–9201.
- Kassavetis, G. A., Letts, G. A., & Geiduschek, E. P. (2001). The RNA polymerase III transcription initiation factor TFIIB participates in two steps of promoter opening. *The EMBO Journal*, 20(11), 2823–2834. <https://doi.org/10.1093/emboj/20.11.2823>
- Kassavetis, G. A., Prakash, P., & Shim, E. (2010). The C53/C37 Subcomplex of RNA Polymerase III Lies Near the Active Site and Participates in Promoter Opening *. *Journal of Biological Chemistry*, 285(4), 2695–2706.

- <https://doi.org/10.1074/jbc.M109.074013>
- Kassavetis, G. A., Riggs, D. L., Negri, R., Nguyen, L. H., & Geiduschek, E. P. (1989). Transcription factor IIIB generates extended DNA interactions in RNA polymerase III transcription complexes on tRNA genes. *Molecular and Cellular Biology*, 9(6), 2551–2566. <https://doi.org/10.1128/mcb.9.6.2551-2566.1989>
- Kawauchi, J., Mischo, H., Braglia, P., Rondon, A., & Proudfoot, N. J. (2008). Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. *Genes & Development*, 22(8), 1082–1092. <https://doi.org/10.1101/gad.463408>
- Khatter, H., Vorländer, M. K., & Müller, C. W. (2017). RNA polymerase I and III: Similar yet unique. *Current Opinion in Structural Biology*, 47, 88–94. <https://doi.org/10.1016/j.sbi.2017.05.008>
- Khoo, S.-K., Wu, C.-C., Lin, Y.-C., Lee, J.-C., & Chen, H.-T. (2014). Mapping the Protein Interaction Network for TFIIB-Related Factor Brf1 in the RNA Polymerase III Preinitiation Complex. *Molecular and Cellular Biology*, 34(3), 551–559. <https://doi.org/10.1128/MCB.00910-13>
- Kim, H.-D., Choe, J., & Seo, Y.-S. (1999). The sen1+ Gene of Schizosaccharomyces pombe, a Homologue of Budding Yeast SEN1, Encodes an RNA and DNA Helicase. *Biochemistry*, 38(44), 14697–14710. <https://doi.org/10.1021/bi991470c>
- Kim, M., Krogan, N. J., Vasiljeva, L., Rando, O. J., Nedeja, E., Greenblatt, J. F., & Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature*, 432(7016), 517–522. <https://doi.org/10.1038/nature03041>
- Kirchner, S., & Ignatova, Z. (2015). Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. *Nature Reviews. Genetics*, 16(2), 98–112. <https://doi.org/10.1038/nrg3861>
- Kireeva, M. L., Komissarova, N., Waugh, D. S., & Kashlev, M. (2000). The 8-Nucleotide-long RNA:DNA Hybrid Is a Primary Stability Determinant of the RNA Polymerase II Elongation Complex *. *Journal of Biological Chemistry*, 275(9), 6530–6536. <https://doi.org/10.1074/jbc.275.9.6530>
- Komarnitsky, P., Cho, E. J., & Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes & Development*, 14(19), 2452–2460. <https://doi.org/10.1101/gad.824700>
- Korzheva, N., Mustaev, A., Nudler, E., Nikiforov, V., & Goldfarb, A. (1998). Mechanistic model of the elongation complex of Escherichia coli RNA polymerase. *Cold Spring Harbor Symposia on Quantitative Biology*, 63, 337–345. <https://doi.org/10.1101/sqb.1998.63.337>
- Koski, R. A., Clarkson, S. G., Kurjan, J., Hall, B. D., & Smith, M. (1980). Mutations of the yeast SUP4 tRNATyr Locus: Transcription of the mutant genes in vitro. *Cell*, 22(2, Part 2), 415–425. [https://doi.org/10.1016/0092-8674\(80\)90352-9](https://doi.org/10.1016/0092-8674(80)90352-9)
- Kouzarides, T. (2007). Chromatin Modifications and Their Function. *Cell*, 128(4), 693–705. <https://doi.org/10.1016/j.cell.2007.02.005>
- Kramerov, D. A., & Vassetzky, N. S. (2011). Origin and evolution of SINEs in eukaryotic genomes. *Heredity*, 107(6), 487–495. <https://doi.org/10.1038/hdy.2011.43>
- Kubicek, K., Cerna, H., Holub, P., Pasulka, J., Hrossova, D., Loehr, F., Hofr, C., Vanacova, S., & Stefl, R. (2012). Serine phosphorylation and proline isomerization in RNAP II CTD control recruitment of Nrd1. *Genes & Development*, 26(17), 1891–1896. <https://doi.org/10.1101/gad.192781.112>
- Kuehner, J. N., Pearson, E. L., & Moore, C. (2011). Unravelling the means to an end: RNA polymerase II transcription termination. *Nature Reviews Molecular Cell Biology*, 12(5), 283–294. <https://doi.org/10.1038/nrm3098>
- Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N., & Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nature Methods*, 11(3), 319–324. <https://doi.org/10.1038/nmeth.2834>
- Kulkens, T., van der Sande, C. A., Dekker, A. F., van Heerikhuizen, H., & Planta, R. J. (1992). A system to study transcription by yeast RNA polymerase I within the chromosomal context: Functional analysis of the ribosomal DNA enhancer and the RBP1/REB1 binding sites. *The EMBO Journal*, 11(12), 4665–4674.
- Kumar, A., Clerici, M., Muckenfuss, L. M., Passmore, L. A., & Jinek, M. (2019). Mechanistic insights into mRNA 3'-end processing. *Current Opinion in Structural Biology*, 59, 143–150. <https://doi.org/10.1016/j.sbi.2019.08.001>

- LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A., & Tollervey, D. (2005). RNA Degradation by the Exosome Is Promoted by a Nuclear Polyadenylation Complex. *Cell*, 121(5), 713–724. <https://doi.org/10.1016/j.cell.2005.04.029>
- Landrieux, E., Alic, N., Ducrot, C., Acker, J., Riva, M., & Carles, C. (2006). A subcomplex of RNA polymerase III subunits involved in transcription termination and reinitiation. *The EMBO Journal*, 25(1), 118–128. <https://doi.org/10.1038/sj.emboj.7600915>
- Larochelle, M., Robert, M.-A., Hébert, J.-N., Liu, X., Matteau, D., Rodrigue, S., Tian, B., Jacques, P.-É., & Bachand, F. (2018). Common mechanism of transcription termination at coding and noncoding RNA genes in fission yeast. *Nature Communications*, 9(1), 4364. <https://doi.org/10.1038/s41467-018-06546-x>
- Larson, M. H., Greenleaf, W. J., Landick, R., & Block, S. M. (2008). Applied force reveals mechanistic and energetic details of transcription termination. *Cell*, 132(6), 971–982. <https://doi.org/10.1016/j.cell.2008.01.027>
- Lassar, A. B., Martin, P. L., & Roeder, R. G. (1983). Transcription of class III genes: Formation of preinitiation complexes. *Science (New York, N.Y.)*, 222(4625), 740–748. <https://doi.org/10.1126/science.6356356>
- Layat, E., Probst, A. V., & Tourmente, S. (2013). Structure, function and regulation of Transcription Factor IIIA: From *Xenopus* to *Arabidopsis*. *Biochimica Et Biophysica Acta- Gene Regulatory Mechanisms*, 1829(3–4), 274–282. <https://doi.org/10.1016/j.bbagr.2012.10.013>
- Lee, J. Y., Rohlman, C. E., Molony, L. A., & Engelke, D. R. (1991). Characterization of RPR1, an essential gene encoding the RNA component of *Saccharomyces cerevisiae* nuclear RNase P. *Molecular and Cellular Biology*, 11(2), 721–730. <https://doi.org/10.1128/mcb.11.2.721-730.1991>
- Legros, P., Malapert, A., Niinuma, S., Bernard, P., & Vanoosthuyse, V. (2014). RNA Processing Factors Swd2.2 and Sen1 Antagonize RNA Pol III-Dependent Transcription and the Localization of Condensin at Pol III Genes. *PLOS Genetics*, 10(11), e1004794. <https://doi.org/10.1371/journal.pgen.1004794>
- Lemay, J.-F., Marguerat, S., Larochelle, M., Liu, X., van Nues, R., Hunyadkürti, J., Hoque, M., Tian, B., Granneman, S., Bähler, J., & Bachand, F. (2016). The Nrd1-like protein Seb1 coordinates cotranscriptional 3' end processing and polyadenylation site selection. *Genes & Development*, 30(13), 1558–1572. <https://doi.org/10.1101/gad.280222.116>
- Leonaitė, B., Han, Z., Basquin, J., Bonneau, F., Libri, D., Porrua, O., & Conti, E. (2017). Sen1 has unique structural features grafted on the architecture of the Upf1-like helicase family. *The EMBO Journal*, 36(11), 1590–1604. <https://doi.org/10.15252/emboj.201696174>
- Leśniowska, E., & Boguta, M. (2017). Novel layers of RNA polymerase III control affecting tRNA gene transcription in eukaryotes. *Open Biology*, 7(2). <https://doi.org/10.1098/rsob.170001>
- Li, B., Carey, M., & Workman, J. L. (2007). The Role of Chromatin during Transcription. *Cell*, 128(4), 707–719. <https://doi.org/10.1016/j.cell.2007.01.015>
- Li, W., Selvam, K., Rahman, S. A., & Li, S. (2016). Sen1, the yeast homolog of human senataxin, plays a more direct role than Rad26 in transcription coupled DNA repair. *Nucleic Acids Research*, 44(14), 6794–6802. <https://doi.org/10.1093/nar/gkw428>
- Liao, Y., Willis, I. M., & Moir, R. D. (2003). The Brf1 and Bdp1 Subunits of Transcription Factor TFIIB Bind to Overlapping Sites in the Tetratricopeptide Repeats of Tfc4*. *Journal of Biological Chemistry*, 278(45), 44467–44474. <https://doi.org/10.1074/jbc.M308354200>
- Lidschreiber, M., Easter, A. D., Battaglia, S., Rodríguez-Molina, J. B., Casañal, A., Carminati, M., Baejen, C., Grzechnik, P., Maier, K. C., Cramer, P., & Passmore, L. A. (2018). The APT complex is involved in non-coding RNA transcription and is distinct from CPF. *Nucleic Acids Research*, 46(21), 11528–11538. <https://doi.org/10.1093/nar/gky845>
- Lingner, J., Kellermann, J., & Keller, W. (1991). Cloning and expression of the essential gene for poly(A) polymerase from *S. cerevisiae*. *Nature*, 354(6353), 496–498. <https://doi.org/10.1038/354496a0>
- Lin-Marq, N., & Clarkson, S. G. (1998). Efficient synthesis, termination and release of RNA polymerase III transcripts in *Xenopus* extracts depleted of La protein. *The EMBO Journal*, 17(7), 2033–2041. <https://doi.org/10.1093/emboj/17.7.2033>
- Logan, J., Falck-Pedersen, E., Darnell, J. E., & Shenk, T. (1987). A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. *Proceedings of the National Academy of Sciences of the United States of America*,

- 84(23), 8306–8310. <https://doi.org/10.1073/pnas.84.23.8306>
- Lu, H., Flores, O., Weinmann, R., & Reinberg, D. (1991). The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. *Proceedings of the National Academy of Sciences of the United States of America*, 88(22), 10004–10008. <https://doi.org/10.1073/pnas.88.22.10004>
- Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., & Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, 389(6648), 251–260. <https://doi.org/10.1038/38444>
- Lunde, B. M., Reichow, S. L., Kim, M., Suh, H., Leeper, T. C., Yang, F., Mutschler, H., Buratowski, S., Meinhart, A., & Varani, G. (2010). Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nature Structural & Molecular Biology*, 17(10), 1195–1201. <https://doi.org/10.1038/nsmb.1893>
- Luo, W., Johnson, A. W., & 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 & Development*, 20(8), 954–965. <https://doi.org/10.1101/gad.1409106>
- Luse, D. S. (2013). Promoter clearance by RNA polymerase II. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1829(1), 63–68. <https://doi.org/10.1016/j.bbagrm.2012.08.010>
- Male, G., von Appen, A., Glatt, S., Taylor, N. M. I., Cristovao, M., Groetsch, H., Beck, M., & Müller, C. W. (2015). Architecture of TFIIIC and its role in RNA polymerase III pre-initiation complex assembly. *Nature Communications*, 6(1), 7387. <https://doi.org/10.1038/ncomms8387>
- Malik, S., Hisatake, K., Sumimoto, H., Horikoshi, M., & Roeder, R. G. (1991). Sequence of general transcription factor TFIIIB and relationships to other initiation factors. *Proceedings of the National Academy of Sciences*, 88(21), 9553–9557. <https://doi.org/10.1073/pnas.88.21.9553>
- Maraia, R. J., Kenan, D. J., & Keene, J. D. (1994). Eukaryotic transcription termination factor La mediates transcript release and facilitates reinitiation by RNA polymerase III. *Molecular and Cellular Biology*, 14(3), 2147–2158. <https://doi.org/10.1128/mcb.14.3.2147-2158.1994>
- Marck, C., Kachouri-Lafond, R., Lafontaine, I., Westhof, E., Dujon, B., & Grosjean, H. (2006). The RNA polymerase III-dependent family of genes in hemiascomycetes: Comparative RNomics, decoding strategies, transcription and evolutionary implications. *Nucleic Acids Research*, 34(6), 1816–1835. <https://doi.org/10.1093/nar/gkl085>
- Marquardt, S., Hazelbaker, D. Z., & Buratowski, S. (2011). Distinct RNA degradation pathways and 3' extensions of yeast non-coding RNA species. *Transcription*, 2(3), 145–154. <https://doi.org/10.4161/trns.2.3.16298>
- Martin, F. H., & Tinoco, I., Jr. (1980). DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. *Nucleic Acids Research*, 8(10), 2295–2300. <https://doi.org/10.1093/nar/8.10.2295>
- Martin-Tumasch, S., & Brow, D. A. (2015). *Saccharomyces cerevisiae* Sen1 Helicase Domain Exhibits 5'- to 3'-Helicase Activity with a Preference for Translocation on DNA Rather than RNA. *The Journal of Biological Chemistry*, 290(38), 22880–22889. <https://doi.org/10.1074/jbc.M115.674002>
- Marzouki, N., Camier, S., Ruet, A., Moenne, A., & Sentenac, A. (1986). Selective proteolysis defines two DNA binding domains in yeast transcription factor τ . *Nature*, 323(6084), 176–178. <https://doi.org/10.1038/323176a0>
- Mayer, A., Heidemann, M., Lidschreiber, M., Schreieck, A., Sun, M., Hintermair, C., Kremmer, E., Eick, D., & Cramer, P. (2012). CTD Tyrosine Phosphorylation Impairs Termination Factor Recruitment to RNA Polymerase II. *Science*, 336(6089), 1723–1725. <https://doi.org/10.1126/science.1219651>
- Mazabraud, A., Scherly, D., Müller, F., Rungger, D., & Clarkson, S. G. (1987). Structure and transcription termination of a lysine tRNA gene from *Xenopus laevis*. *Journal of Molecular Biology*, 195(4), 835–845. [https://doi.org/10.1016/0022-2836\(87\)90488-8](https://doi.org/10.1016/0022-2836(87)90488-8)
- McMahon, M. E., Stamenkovich, D., & Petes, T. D. (1984). Tandemly arranged variant 5S ribosomal RNA genes in the yeast *Saccharomyces cerevisiae*. *Nucleic Acids Research*, 12(21), 8001–8016. <https://doi.org/10.1093/nar/12.21.8001>
- Meinhart, A., Blobel, J., & Cramer, P. (2003). An Extended Winged Helix Domain in General Transcription Factor E/II α *. *Journal of Biological Chemistry*, 278(48), 48267–48274. <https://doi.org/10.1074/jbc.M307874200>
- Merkel, P., Perez-Fernandez, J., Pilsl, M., Reiter, A., Williams, L., Gerber, J., Böhm, M., Deutzmann, R., Griesenbeck,

- J., Milkereit, P., & Tschochner, H. (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. *Molecular and Cellular Biology*, 34(20), 3817–3827. <https://doi.org/10.1128/MCB.00395-14>
- Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F., & Shilatifard, A. (2001). COMPASS: A complex of proteins associated with a trithorax-related SET domain protein. *Proceedings of the National Academy of Sciences of the United States of America*, 98(23), 12902–12907. <https://doi.org/10.1073/pnas.231473398>
- Mischo, H. E., Gómez-González, B., Grzechnik, P., Rondón, A. G., Wei, W., Steinmetz, L., Aguilera, A., & Proudfoot, N. J. (2011). Yeast Sen1 Helicase Protects the Genome from Transcription-Associated Instability. *Molecular Cell*, 41(1), 21–32. <https://doi.org/10.1016/j.molcel.2010.12.007>
- Mischo, H. E., & Proudfoot, N. J. (2013). Disengaging polymerase: Terminating RNA polymerase II transcription in budding yeast. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1829(1), 174–185. <https://doi.org/10.1016/j.bbagr.2012.10.003>
- Mishra, S., & Maraia, R. J. (2019). RNA polymerase III subunits C37/53 modulate rU:dA hybrid 3' end dynamics during transcription termination. *Nucleic Acids Research*, 47(1), 310–327. <https://doi.org/10.1093/nar/gky1109>
- Mittal, V., Ma, B., & Hernandez, N. (1999). SNAPc: A core promoter factor with a built-in DNA-binding damper that is deactivated by the Oct-1 POU domain. *Genes & Development*, 13(14), 1807–1821.
- Moir, R. D., Puglia, K. V., & Willis, I. M. (2000). Interactions between the tetratricopeptide repeat-containing transcription factor TFIIC131 and its ligand, TFIIB70. Evidence for a conformational change in the complex. *The Journal of Biological Chemistry*, 275(34), 26591–26598. <https://doi.org/10.1074/jbc.M003991200>
- Moir, R. D., Sethy-Coraci, I., Puglia, K., Librizzi, M. D., & Willis, I. M. (1997). A tetratricopeptide repeat mutation in yeast transcription factor IIC131 (TFIIC131) facilitates recruitment of TFIIB-related factor TFIIB70. *Molecular and Cellular Biology*, 17(12), 7119–7125.
- Molodtsov, V., Anikin, M., & McAllister, W. T. (2014). The Presence of an RNA:DNA Hybrid That Is Prone to Slippage Promotes Termination by T7 RNA Polymerase. *Journal of Molecular Biology*, 426(18), 3095–3107. <https://doi.org/10.1016/j.jmb.2014.06.012>
- Moqtaderi, Z., & Struhl, K. (2004). Genome-Wide Occupancy Profile of the RNA Polymerase III Machinery in *Saccharomyces cerevisiae* Reveals Loci with Incomplete Transcription Complexes. *Molecular and Cellular Biology*, 24(10), 4118–4127. <https://doi.org/10.1128/MCB.24.10.4118-4127.2004>
- Moreira, M.-C., Klur, S., Watanabe, M., Németh, A. H., Le Ber, I., Moniz, J.-C., Tranchant, C., Aubourg, P., Tazir, M., Schöls, L., Pandolfo, M., Schulz, J. B., Pouget, J., Calvas, P., Shizuka-Ikeda, M., Shoji, M., Tanaka, M., Izatt, L., Shaw, C. E., ... Koenig, M. (2004). Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. *Nature Genetics*, 36(3), 225–227. <https://doi.org/10.1038/ng1303>
- Mosley, A. L., Pattenden, S. G., Carey, M., Venkatesh, S., Gilmore, J. M., Florens, L., Workman, J. L., & Washburn, M. P. (2009). Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. *Molecular Cell*, 34(2), 168–178. <https://doi.org/10.1016/j.molcel.2009.02.025>
- Myers, L. C., Gustafsson, C. M., Bushnell, D. A., Lui, M., Erdjument-Bromage, H., Tempst, P., & Kornberg, R. D. (1998). The Med proteins of yeast and their function through the RNA polymerase II carboxy-terminal domain. *Genes & Development*, 12(1), 45–54. <https://doi.org/10.1101/gad.12.1.45>
- Mylona, A., Fernández-Tornero, C., Legrand, P., Haupt, M., Sentenac, A., Acker, J., & Müller, C. W. (2006). Structure of the $\tau 60/\Delta\tau 91$ Subcomplex of Yeast Transcription Factor IIIC: Insights into Preinitiation Complex Assembly. *Molecular Cell*, 24(2), 221–232. <https://doi.org/10.1016/j.molcel.2006.08.013>
- Nagarajavel, V., Iben, J. R., Howard, B. H., Maraia, R. J., & Clark, D. J. (2013). Global 'bootprinting' reveals the elastic architecture of the yeast TFIIB–TFIIC transcription complex in vivo. *Nucleic Acids Research*, 41(17), 8135–8143. <https://doi.org/10.1093/nar/gkt611>
- Nedea, E., He, X., Kim, M., Pootoolal, J., Zhong, G., Canadien, V., Hughes, T., Buratowski, S., Moore, C. L., & Greenblatt, J. (2003). Organization and Function of APT, a Subcomplex of the Yeast Cleavage and Polyadenylation Factor Involved in the Formation of mRNA and Small Nucleolar RNA 3'-Ends *. *Journal of Biological Chemistry*, 278(35), 33000–33010. <https://doi.org/10.1074/jbc.M304454200>

- Nedea, E., Nalbant, D., Xia, D., Theoharis, N. T., Suter, B., Richardson, C. J., Tatchell, K., Kislinger, T., Greenblatt, J. F., & Nagy, P. L. (2008). The Glc7 Phosphatase Subunit of the Cleavage and Polyadenylation Factor Is Essential for Transcription Termination on snoRNA Genes. *Molecular Cell*, 29(5), 577–587. <https://doi.org/10.1016/j.molcel.2007.12.031>
- Neil, H., Malabat, C., d'Aubenton-Carafa, Y., Xu, Z., Steinmetz, L. M., & Jacquier, A. (2009). Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature*, 457(7232), 1038–1042. <https://doi.org/10.1038/nature07747>
- Newman, J. R. S., Ghaemmighami, S., Ihmels, J., Breslow, D. K., Noble, M., DeRisi, J. L., & Weissman, J. S. (2006). Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature*, 441(7095), 840–846. <https://doi.org/10.1038/nature04785>
- Nguyen, V. Q., Ranjan, A., Liu, S., Tang, X., Ling, Y. H., Wisniewski, J., Mizuguchi, G., Li, K. Y., Jou, V., Zheng, Q., Lavis, L. D., Lionnet, T., & Wu, C. (2021). Spatiotemporal coordination of transcription preinitiation complex assembly in live cells. *Molecular Cell*, 81(17), 3560-3575.e6. <https://doi.org/10.1016/j.molcel.2021.07.022>
- Nielsen, S., Yuzenkova, Y., & Zenkin, N. (2013). Mechanism of Eukaryotic RNA Polymerase III Transcription Termination. *Science*, 340(6140), 1577–1580. <https://doi.org/10.1126/science.1237934>
- Nikitina, T. V., Tischenko, L. I., & Schulz, W. A. (2011). *Recent insights into regulation of transcription by RNA polymerase III and the cellular functions of its transcripts*. 392(5), 395–404. <https://doi.org/10.1515/bc.2011.049>
- Olivas, W. M., Muhlrad, D., & Parker, R. (1997). Analysis of the yeast genome: Identification of new non-coding and small ORF-containing RNAs. *Nucleic Acids Research*, 25(22), 4619–4625. <https://doi.org/10.1093/nar/25.22.4619>
- Orioli, A., Pascali, C., Pagano, A., Teichmann, M., & Dieci, G. (2012). RNA polymerase III transcription control elements: Themes and variations. *Gene*, 493(2), 185–194. <https://doi.org/10.1016/j.gene.2011.06.015>
- Orioli, A., Pascali, C., Quartararo, J., Diebel, K. W., Praz, V., Romascano, D., Percudani, R., van Dyk, L. F., Hernandez, N., Teichmann, M., & Dieci, G. (2011). Widespread occurrence of non-canonical transcription termination by human RNA polymerase III. *Nucleic Acids Research*, 39(13), 5499–5512. <https://doi.org/10.1093/nar/gkr074>
- Orphanides, G., Wu, W. H., Lane, W. S., Hampsey, M., & Reinberg, D. (1999). The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. *Nature*, 400(6741), 284–288. <https://doi.org/10.1038/22350>
- Osman, S., & Cramer, P. (2020). Structural Biology of RNA Polymerase II Transcription: 20 Years On. *Annual Review of Cell and Developmental Biology*. <https://doi.org/10.1146/annurev-cellbio-042020-021954>
- Pannone, B. K., Xue, D., & Wolin, S. L. (1998). A role for the yeast La protein in U6 snRNP assembly: Evidence that the La protein is a molecular chaperone for RNA polymerase III transcripts. *The EMBO Journal*, 17(24), 7442–7453. <https://doi.org/10.1093/emboj/17.24.7442>
- Papai, G., Frechard, A., Kolesnikova, O., Crucifix, C., Schultz, P., & Ben-Shem, A. (2020). Structure of SAGA and mechanism of TBP deposition on gene promoters. *Nature*, 577(7792), 711–716. <https://doi.org/10.1038/s41586-020-1944-2>
- Park, J., Kang, M., & Kim, M. (2015). Unraveling the mechanistic features of RNA polymerase II termination by the 5'-3' exoribonuclease Rat1. *Nucleic Acids Research*, 43(5), 2625–2637. <https://doi.org/10.1093/nar/gkv133>
- Pearson, E. L., & Moore, C. L. (2013). Dismantling promoter-driven RNA polymerase II transcription complexes in vitro by the termination factor Rat1. *The Journal of Biological Chemistry*, 288(27), 19750–19759. <https://doi.org/10.1074/jbc.M112.434985>
- Perales, R., & Bentley, D. (2009). “Cotranscriptionality”: The transcription elongation complex as a nexus for nuclear transactions. *Molecular Cell*, 36(2), 178–191. <https://doi.org/10.1016/j.molcel.2009.09.018>
- Peterlin, B. M., Brogie, J. E., & Price, D. H. (2012). 7SK snRNA: A noncoding RNA that plays a major role in regulating eukaryotic transcription. *Wiley Interdisciplinary Reviews-Rna*, 3(1), 92–103. <https://doi.org/10.1002/wrna.106>
- Porrua, O., Boudvillain, M., & Libri, D. (2016). Transcription Termination: Variations on Common Themes. *Trends in Genetics*, 32(8), 508–522. <https://doi.org/10.1016/j.tig.2016.05.007>

- Porrúa, O., Hobor, F., Boulay, J., Kubicek, K., D'Aubenton-Carafa, Y., Gudipati, R. K., Stefl, R., & Libri, D. (2012). In vivo SELEX reveals novel sequence and structural determinants of Nrd1-Nab3-Sen1-dependent transcription termination. *The EMBO Journal*, *31*(19), 3935–3948. <https://doi.org/10.1038/emboj.2012.237>
- Porrúa, O., & Libri, D. (2013a). RNA quality control in the nucleus: The Angels' share of RNA. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1829*(6–7), 604–611. <https://doi.org/10.1016/j.bbagr.2013.02.012>
- Porrúa, O., & Libri, D. (2013b). A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast. *Nature Structural & Molecular Biology*, *20*(7), 884–891. <https://doi.org/10.1038/nsmb.2592>
- Porrúa, O., & Libri, D. (2015a). Characterization of the mechanisms of transcription termination by the helicase Sen1. *Methods in Molecular Biology (Clifton, N.J.)*, *1259*, 313–331. https://doi.org/10.1007/978-1-4939-2214-7_19
- Porrúa, O., & Libri, D. (2015b). Transcription termination and the control of the transcriptome: Why, where and how to stop. *Nature Reviews Molecular Cell Biology*, *16*(3), 190–202. <https://doi.org/10.1038/nrm3943>
- Poss, Z. C., Ebmeier, C. C., & Taatjes, D. J. (2013). The Mediator complex and transcription regulation. *Critical Reviews in Biochemistry and Molecular Biology*, *48*(6), 575–608. <https://doi.org/10.3109/10409238.2013.840259>
- Qiu, H., Hu, C., & Hinnebusch, A. G. (2009). Phosphorylation of the Pol II CTD by KIN28 enhances BUR1/BUR2 recruitment and Ser2 CTD phosphorylation near promoters. *Molecular Cell*, *33*(6), 752–762. <https://doi.org/10.1016/j.molcel.2009.02.018>
- Ramsay, E. P., Abascal-Palacios, G., Daiß, J. L., King, H., Gouge, J., Pils, M., Beuron, F., Morris, E., Gunkel, P., Engel, C., & Vannini, A. (2020). Structure of human RNA polymerase III. *Nature Communications*, *11*(1), 6409. <https://doi.org/10.1038/s41467-020-20262-5>
- Ramsay, E. P., & Vannini, A. (2018). Structural rearrangements of the RNA polymerase III machinery during tRNA transcription initiation. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1861*(4), 285–294. <https://doi.org/10.1016/j.bbagr.2017.11.005>
- Ray-Soni, A., Bellecourt, M. J., & Landick, R. (2016). Mechanisms of Bacterial Transcription Termination: All Good Things Must End. *Annual Review of Biochemistry*, *85*(1), 319–347. <https://doi.org/10.1146/annurev-biochem-060815-014844>
- Reeder, R. H., & Roeder, R. G. (1972). Ribosomal RNA synthesis in isolated nuclei. *Journal of Molecular Biology*, *67*(3), 433–441. [https://doi.org/10.1016/0022-2836\(72\)90461-5](https://doi.org/10.1016/0022-2836(72)90461-5)
- Reeve, J. N. (2003). Archaeal chromatin and transcription. *Molecular Microbiology*, *48*(3), 587–598. <https://doi.org/10.1046/j.1365-2958.2003.03439.x>
- Reiter, A., Hamperl, S., Seitz, H., Merkl, P., Perez-Fernandez, J., Williams, L., Gerber, J., Németh, A., Léger, I., Gadal, O., Milkereit, P., Griesenbeck, J., & Tschochner, H. (2012). The Reb1-homologue Ydr026c/Nsi1 is required for efficient RNA polymerase I termination in yeast. *The EMBO Journal*, *31*(16), 3480–3493. <https://doi.org/10.1038/emboj.2012.185>
- Richard, P., Feng, S., & Manley, J. L. (2013). A SUMO-dependent interaction between Senataxin and the exosome, disrupted in the neurodegenerative disease AOA2, targets the exosome to sites of transcription-induced DNA damage. *Genes & Development*, *27*(20), 2227–2232. <https://doi.org/10.1101/gad.224923.113>
- Richard, P., & Manley, J. L. (2009). Transcription termination by nuclear RNA polymerases. *Genes & Development*, *23*(11), 1247–1269. <https://doi.org/10.1101/gad.1792809>
- Rijal, K., & Maraia, R. J. (2013). RNA polymerase III mutants in TFIIIF-like C37 that cause terminator readthrough with no decrease in transcription output. *Nucleic Acids Research*, *41*(1), 139–155. <https://doi.org/10.1093/nar/gks985>
- Rivosecchi, J. (2019). *Sen1-mediated RNAPIII transcription termination controls the positioning of condensin on mitotic chromosomes* [These de doctorat, Lyon]. <http://www.theses.fr/2019LYSEN041>
- Rivosecchi, J., Laroche, M., Teste, C., Grenier, F., Malapert, A., Ricci, E. P., Bernard, P., Bachand, F., & Vanoosthuyse, V. (2019). Senataxin homologue Sen1 is required for efficient termination of RNA polymerase III transcription. *The EMBO Journal*, *38*(16). <https://doi.org/10.15252/emj.2019101955>
- Roberts, D. N., Stewart, A. J., Huff, J. T., & Cairns, B. R. (2003). The RNA polymerase III transcriptome revealed by genome-wide localization and activity–occupancy relationships. *Proceedings of the National Academy of Sciences*, *100*(12), 7107–7112. <https://doi.org/10.1073/pnas.0305071100>

- Sciences*, 100(25), 14695–14700. <https://doi.org/10.1073/pnas.2435566100>
- Roberts, J. W. (2019). Mechanisms of Bacterial Transcription Termination. *Journal of Molecular Biology*, 431(20), 4030–4039. <https://doi.org/10.1016/j.jmb.2019.04.003>
- Roeder, R. G. (2019). 50+ years of eukaryotic transcription: An expanding universe of factors and mechanisms. *Nature Structural & Molecular Biology*, 26(9), 783–791. <https://doi.org/10.1038/s41594-019-0287-x>
- Roeder, R. G., & Rutter, W. J. (1969). Multiple Forms of DNA-dependent RNA Polymerase in Eukaryotic Organisms. *Nature*, 224(5216), 234–237. <https://doi.org/10.1038/224234a0>
- Rondón, A. G., Mischo, H. E., Kawauchi, J., & Proudfoot, N. J. (2009). Fail-safe transcriptional termination for protein-coding genes in *S. cerevisiae*. *Molecular Cell*, 36(1), 88–98. <https://doi.org/10.1016/j.molcel.2009.07.028>
- Roy, K., Gabunilas, J., Gillespie, A., Ngo, D., & Chanfreau, G. F. (2016). Common genomic elements promote transcriptional and DNA replication roadblocks. *Genome Research*, 26(10), 1363–1375. <https://doi.org/10.1101/gr.204776.116>
- Ruet, A., Camier, S., Smagowicz, W., Sentenac, A., & Fromageot, P. (1984). Isolation of a class C transcription factor which forms a stable complex with tRNA genes. *The EMBO Journal*, 3(2), 343–350. <https://doi.org/10.1002/j.1460-2075.1984.tb01809.x>
- Sainsbury, S., Bernecky, C., & Cramer, P. (2015). Structural basis of transcription initiation by RNA polymerase II. *Nature Reviews Molecular Cell Biology*, 16(3), 129–143. <https://doi.org/10.1038/nrm3952>
- Santangelo, T. J., & Roberts, J. W. (2004). Forward translocation is the natural pathway of RNA release at an intrinsic terminator. *Molecular Cell*, 14(1), 117–126. [https://doi.org/10.1016/s1097-2765\(04\)00154-6](https://doi.org/10.1016/s1097-2765(04)00154-6)
- Schaughency, P., Merran, J., & Corden, J. L. (2014). Genome-wide mapping of yeast RNA polymerase II termination. *PLoS Genetics*, 10(10), e1004632. <https://doi.org/10.1371/journal.pgen.1004632>
- Schier, A. C., & Taatjes, D. J. (2020). Structure and mechanism of the RNA polymerase II transcription machinery. *Genes & Development*, 34(7–8), 465–488. <https://doi.org/10.1101/gad.335679.119>
- Schmidt, K., & Butler, J. S. (2013). Nuclear RNA Surveillance: Role of TRAMP in Controlling Exosome Specificity. *Wiley Interdisciplinary Reviews. RNA*, 4(2), 217–231. <https://doi.org/10.1002/wrna.1155>
- Schmitt, M. E., & Clayton, D. A. (1993). Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, 13(12), 7935–7941. <https://doi.org/10.1128/mcb.13.12.7935-7941.1993>
- Schultz, P., Marzouki, N., Marck, C., Ruet, A., Oudet, P., & Sentenac, A. (1989). The two DNA-binding domains of yeast transcription factor tau as observed by scanning transmission electron microscopy. *The EMBO Journal*, 8(12), 3815–3824. <https://doi.org/10.1002/j.1460-2075.1989.tb08559.x>
- Schulz, D., Schwalb, B., Kiesel, A., Baejen, C., Torkler, P., Gagneur, J., Soeding, J., & Cramer, P. (2013). Transcriptome Surveillance by Selective Termination of Noncoding RNA Synthesis. *Cell*, 155(5), 1075–1087. <https://doi.org/10.1016/j.cell.2013.10.024>
- Selth, L. A., Sigurdsson, S., & Svejstrup, J. Q. (2010). Transcript Elongation by RNA Polymerase II. *Annual Review of Biochemistry*, 79(1), 271–293. <https://doi.org/10.1146/annurev.biochem.78.062807.091425>
- Sentenac, A., Riva, M., Thuriaux, P., Buhler, J., Treich, I., Carles, C., Werner, M., Ruet, A., Huet, J., & Mann, C. (1992). Yeast RNA polymerase subunits and genes. In *Transcriptional Regulation* (pp. 27–54). Cold Spring Harbor Laboratory Press. <https://ci.nii.ac.jp/naid/10005183319/en/>
- Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: Mechanisms of regulation in development and disease pathogenesis. *Annual Review of Biochemistry*, 81, 65–95. <https://doi.org/10.1146/annurev-biochem-051710-134100>
- Sidorenkov, I., Komissarova, N., & Kashlev, M. (1998). Crucial role of the RNA:DNA hybrid in the processivity of transcription. *Molecular Cell*, 2(1), 55–64. [https://doi.org/10.1016/s1097-2765\(00\)80113-6](https://doi.org/10.1016/s1097-2765(00)80113-6)
- Sims, R. J., Belotserkovskaya, R., & Reinberg, D. (2004). Elongation by RNA polymerase II: The short and long of it. *Genes & Development*, 18(20), 2437–2468. <https://doi.org/10.1101/gad.1235904>
- Sklar, V. E., Schwartz, L. B., & Roeder, R. G. (1975). Distinct molecular structures of nuclear class I, II, and III DNA-dependent RNA polymerases. *Proceedings of the National Academy of Sciences*, 72(1), 348–352. <https://doi.org/10.1073/pnas.72.1.348>

- Skourti-Stathaki, K., Proudfoot, N. J., & Gromak, N. (2011). Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination. *Molecular Cell*, 42(6), 794–805. <https://doi.org/10.1016/j.molcel.2011.04.026>
- Smale, S. T., & Kadonaga, J. T. (2003). The RNA polymerase II core promoter. *Annual Review of Biochemistry*, 72, 449–479. <https://doi.org/10.1146/annurev.biochem.72.121801.161520>
- Soutourina, J. (2018). Transcription regulation by the Mediator complex. *Nature Reviews Molecular Cell Biology*, 19(4), 262–274. <https://doi.org/10.1038/nrm.2017.115>
- Spain, M. M., Ansari, S. A., Pathak, R., Palumbo, M. J., Morse, R. H., & Govind, C. K. (2014). The RSC Complex localizes to coding sequences to regulate Pol II and histone occupancy. *Molecular Cell*, 56(5), 653–666. <https://doi.org/10.1016/j.molcel.2014.10.002>
- Stadler, P. F., Chen, J. J.-L., Hackermüller, J., Hoffmann, S., Horn, F., Khaitovich, P., Kretschmar, A. K., Mosig, A., Prohaska, S. J., Qi, X., Schutt, K., & Ullmann, K. (2009). Evolution of Vault RNAs. *Molecular Biology and Evolution*, 26(9), 1975–1991. <https://doi.org/10.1093/molbev/msp112>
- Steinmetz, E. J., & Brow, D. A. (1996). Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1. *Molecular and Cellular Biology*, 16(12), 6993–7003. <https://doi.org/10.1128/MCB.16.12.6993>
- Steinmetz, E. J., Warren, C. L., Kuehner, J. N., Panbehi, B., Ansari, A. Z., & Brow, D. A. (2006). Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Molecular Cell*, 24(5), 735–746. <https://doi.org/10.1016/j.molcel.2006.10.023>
- Stevens, A. (1960). Incorporation of the adenine ribonucleotide into RNA by cell fractions from *E. coli* B. *Biochemical and Biophysical Research Communications*, 3(1), 92–96. [https://doi.org/10.1016/0006-291X\(60\)90110-8](https://doi.org/10.1016/0006-291X(60)90110-8)
- Stillman, D. J. (2010). Nhp6: A small but powerful effector of chromatin structure in *Saccharomyces cerevisiae*. *Biochimica Et Biophysica Acta-Genetic Regulatory Mechanisms*, 1799(1–2), 175–180. <https://doi.org/10.1016/j.bbagr.2009.11.010>
- Stirnemann, C. U., Petsalaki, E., Russell, R. B., & Müller, C. W. (2010). WD40 proteins propel cellular networks. *Trends in Biochemical Sciences*, 35(10), 565–574. <https://doi.org/10.1016/j.tibs.2010.04.003>
- Suraweera, A., Lim, Y., Woods, R., Birrell, G. W., Nasim, T., Becherel, O. J., & Lavin, M. F. (2009). Functional role for senataxin, defective in ataxia oculomotor apraxia type 2, in transcriptional regulation. *Human Molecular Genetics*, 18(18), 3384–3396. <https://doi.org/10.1093/hmg/ddp278>
- Sutherland, C., & Murakami, K. S. (2018). An Introduction to the Structure and Function of the Catalytic Core Enzyme of *Escherichia coli* RNA Polymerase. *EcoSal Plus*, 8(1). <https://doi.org/10.1128/ecosalplus.ESP-0004-2018>
- Svejstrup, J. Q., Vichi, P., & Egly, J.-M. (1996). The multiple roles of transcription/repair factor TFIIH. *Trends in Biochemical Sciences*, 21(9), 346–350. [https://doi.org/10.1016/S0968-0004\(96\)10046-3](https://doi.org/10.1016/S0968-0004(96)10046-3)
- Synowsky, S. A., van Wijk, M., Raijmakers, R., & Heck, A. J. R. (2009). Comparative multiplexed mass spectrometric analyses of endogenously expressed yeast nuclear and cytoplasmic exosomes. *Journal of Molecular Biology*, 385(4), 1300–1313. <https://doi.org/10.1016/j.jmb.2008.11.011>
- Tan, S., Hunziker, Y., Sargent, D. F., & Richmond, T. J. (1996). Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature*, 381(6578), 127–151. <https://doi.org/10.1038/381127a0>
- Teichmann, M., Dumay-Odelot, H., & Fribourg, S. (2012). Structural and functional aspects of winged-helix domains at the core of transcription initiation complexes. *Transcription*, 3(1), 2–7. <https://doi.org/10.4161/trns.3.1.18917>
- Thiebaut, M., Kisseleva-Romanova, E., Rougemaille, M., Boulay, J., & Libri, D. (2006). Transcription termination and nuclear degradation of cryptic unstable transcripts: A role for the nrd1-nab3 pathway in genome surveillance. *Molecular Cell*, 23(6), 853–864. <https://doi.org/10.1016/j.molcel.2006.07.029>
- Thuillier, V., Stettler, S., Sentenac, A., Thuriaux, P., & Werner, M. (1995). A mutation in the C31 subunit of *Saccharomyces cerevisiae* RNA polymerase III affects transcription initiation. *The EMBO Journal*, 14(2), 351–359.
- Touloukhonov, I., Artsimovitch, I., & Landick, R. (2001). Allosteric Control of RNA Polymerase by a Site That Contacts Nascent RNA Hairpins. *Science*, 292(5517), 730–733. <https://doi.org/10.1126/science.1057738>

- Tudek, A., Porrua, O., Kabzinski, T., Lidschreiber, M., Kubicek, K., Fortova, A., Lacroute, F., Vanacova, S., Cramer, P., Stefl, R., & Libri, D. (2014). Molecular Basis for Coordinating Transcription Termination with Noncoding RNA Degradation. *Molecular Cell*, 55(3), 467–481. <https://doi.org/10.1016/j.molcel.2014.05.031>
- Turowski, T. W., Leśniewska, E., Delan-Forino, C., Sayou, C., Boguta, M., & Tollervey, D. (2016). Global analysis of transcriptionally engaged yeast RNA polymerase III reveals extended tRNA transcripts. *Genome Research*, 26(7), 933–944. <https://doi.org/10.1101/gr.205492.116>
- Turowski, T. W., & Tollervey, D. (2016). Transcription by RNA polymerase III: Insights into mechanism and regulation. *Biochemical Society Transactions*, 44(5), 1367–1375. <https://doi.org/10.1042/BST20160062>
- Ursic, D. (2004). Multiple protein/protein and protein/RNA interactions suggest roles for yeast DNA/RNA helicase Sen1p in transcription, transcription-coupled DNA repair and RNA processing. *Nucleic Acids Research*, 32(8), 2441–2452. <https://doi.org/10.1093/nar/gkh561>
- Valentini, S. R., Weiss, V. H., & Silver, P. A. (1999). Arginine methylation and binding of Hrp1p to the efficiency element for mRNA 3'-end formation. *RNA*, 5(2), 272–280. <https://doi.org/10.1017/S1355838299981633>
- van Dijk, E. L., Chen, C. L., d'Aubenton-Carafa, Y., Gourvenec, S., Kwapisz, M., Roche, V., Bertrand, C., Silvain, M., Legoix-Né, P., Loeillet, S., Nicolas, A., Thermes, C., & Morillon, A. (2011). XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature*, 475(7354), 114–117. <https://doi.org/10.1038/nature10118>
- van Zon, A., Mossink, M. H., Houtsmuller, A. B., Schoester, M., Scheffer, G. L., Scheper, R. J., Sonneveld, P., & Wiemer, E. A. C. (2006). Vault mobility depends in part on microtubules and vaults can be recruited to the nuclear envelope. *Experimental Cell Research*, 312(3), 245–255. <https://doi.org/10.1016/j.yexcr.2005.10.016>
- Vanáčová, S., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G., & Keller, W. (2005). A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS Biology*, 3(6), e189. <https://doi.org/10.1371/journal.pbio.0030189>
- Vannini, A., & Cramer, P. (2012). Conservation between the RNA Polymerase I, II, and III Transcription Initiation Machineries. *Molecular Cell*, 45(4), 439–446. <https://doi.org/10.1016/j.molcel.2012.01.023>
- Vannini, A., Ringel, R., Kusser, A. G., Berninghausen, O., Kassavetis, G. A., & Cramer, P. (2010). Molecular Basis of RNA Polymerase III Transcription Repression by Maf1. *Cell*, 143(1), 59–70. <https://doi.org/10.1016/j.cell.2010.09.002>
- Vasiljeva, L., Kim, M., Mutschler, H., Buratowski, S., & Meinhart, A. (2008). The Nrd1–Nab3–Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nature Structural & Molecular Biology*, 15(8), 795–804. <https://doi.org/10.1038/nsmb.1468>
- Venema, J., & Tollervey, D. (1999). Ribosome Synthesis in *Saccharomyces cerevisiae*. *Annual Review of Genetics*, 33(1), 261–311. <https://doi.org/10.1146/annurev.genet.33.1.261>
- Verosloff, M. S., Corcoran, W. K., Dolberg, T. B., Bushhouse, D. Z., Leonard, J. N., & Lucks, J. B. (2021). RNA Sequence and Structure Determinants of Pol III Transcriptional Termination in Human Cells. *Journal of Molecular Biology*, 433(13), 166978. <https://doi.org/10.1016/j.jmb.2021.166978>
- Villa, T., Barucco, M., Martin-Niclos, M.-J., Jacquier, A., & Libri, D. (2020). Degradation of Non-coding RNAs Promotes Recycling of Termination Factors at Sites of Transcription. *Cell Reports*, 32(3). <https://doi.org/10.1016/j.celrep.2020.107942>
- Vorländer, M. K., Jungblut, A., Karius, K., Baudin, F., Grötsch, H., Kosinski, J., & Müller, C. W. (2020). Structure of the TFIIC subcomplex τ A provides insights into RNA polymerase III pre-initiation complex formation. *Nature Communications*, 11(1), 4905. <https://doi.org/10.1038/s41467-020-18707-y>
- Vorländer, M. K., Khatler, H., Wetzels, R., Hagen, W. J. H., & Müller, C. W. (2018). Molecular mechanism of promoter opening by RNA polymerase III. *Nature*, 553(7688), 295–300. <https://doi.org/10.1038/nature25440>
- Wagschal, A., Rousset, E., Basavarajaiah, P., Contreras, X., Harwig, A., Laurent-Chabalier, S., Nakamura, M., Chen, X., Zhang, K., Meziane, O., Boyer, F., Parrinello, H., Berkhout, B., Terzian, C., Benkirane, M., & Kiernan, R. (2012). Microprocessor, Setx, Xrn2, and Rrp6 Co-operate to Induce Premature Termination of Transcription by RNAPII. *Cell*, 150(6), 1147–1157. <https://doi.org/10.1016/j.cell.2012.08.004>
- Wang, H., Dienemann, C., Stützer, A., Urlaub, H., Cheung, A. C. M., & Cramer, P. (2020). Structure of the

- transcription coactivator SAGA. *Nature*, 577(7792), 717–720. <https://doi.org/10.1038/s41586-020-1933-5>
- Wang, S., Han, Z., Libri, D., Porrua, O., & Strick, T. R. (2019). Single-molecule characterization of extrinsic transcription termination by Sen1 helicase. *Nature Communications*, 10(1), 1545. <https://doi.org/10.1038/s41467-019-09560-9>
- Wang, X., & Folk, W. (1994). Termination of transcription by RNA polymerase III from wheat germ. *The Journal of Biological Chemistry*. [https://doi.org/10.1016/s0021-9258\(17\)37644-5](https://doi.org/10.1016/s0021-9258(17)37644-5)
- Wang, Z., & Roeder, R. G. (1998). DNA topoisomerase I and PC4 can interact with human TFIIIC to promote both accurate termination and transcription reinitiation by RNA polymerase III. *Molecular Cell*, 1(5), 749–757. [https://doi.org/10.1016/s1097-2765\(00\)80074-x](https://doi.org/10.1016/s1097-2765(00)80074-x)
- Warfield, L., Ramachandran, S., Baptista, T., Devys, D., Tora, L., & Hahn, S. (2017). Transcription of Nearly All Yeast RNA Polymerase II-Transcribed Genes Is Dependent on Transcription Factor TFIIID. *Molecular Cell*, 68(1), 118–129.e5. <https://doi.org/10.1016/j.molcel.2017.08.014>
- Weinmann, R., Raskas, H. J., & Roeder, R. G. (1974). Role of DNA-Dependent RNA Polymerases II and III in Transcription of the Adenovirus Genome Late in Productive Infection. *Proceedings of the National Academy of Sciences*, 71(9), 3426–3430. <https://doi.org/10.1073/pnas.71.9.3426>
- Weinmann, R., & Roeder, R. G. (1974). Role of DNA-Dependent RNA Polymerase III in the Transcription of the tRNA and 5S RNA Genes. *Proceedings of the National Academy of Sciences of the United States of America*, 71(5), 1790–1794.
- Weiss, S. B., & Gladstone, L. (1959). A mammalian system for the incorporation of cytidine triphosphate into ribonucleic Acid1. *Journal of the American Chemical Society*, 81(15), 4118–4119. <https://doi.org/10.1021/ja01524a087>
- Werner, F., & Grohmann, D. (2011). Evolution of multisubunit RNA polymerases in the three domains of life. *Nature Reviews Microbiology*, 9(2), 85–98. <https://doi.org/10.1038/nrmicro2507>
- Werner, M., Chaussivert, N., Willis, I. M., & Sentenac, A. (1993). Interaction between a complex of RNA polymerase III subunits and the 70-kDa component of transcription factor IIIB. *Journal of Biological Chemistry*, 268(28), 20721–20724. [https://doi.org/10.1016/S0021-9258\(19\)36839-5](https://doi.org/10.1016/S0021-9258(19)36839-5)
- West, M. L., & Corden, J. L. (1995). Construction and Analysis of Yeast RNA Polymerase II Ctd Deletion and Substitution Mutations. *Genetics*, 140(4), 1223–1233.
- West, S., Gromak, N., & Proudfoot, N. J. (2004). Human 5'→3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature*, 432(7016), 522–525. <https://doi.org/10.1038/nature03035>
- West, S., Proudfoot, N. J., & Dye, M. J. (2008). Molecular dissection of mammalian RNA polymerase II transcriptional termination. *Molecular Cell*, 29(5), 600–610. <https://doi.org/10.1016/j.molcel.2007.12.019>
- Wild, T., & Cramer, P. (2012). Biogenesis of multisubunit RNA polymerases. *Trends in Biochemical Sciences*, 37(3), 99–105. <https://doi.org/10.1016/j.tibs.2011.12.001>
- Willis, I. M., & Moir, R. D. (2018). Signaling to and from the RNA Polymerase III Transcription and Processing Machinery. *Annual Review of Biochemistry*, 87(1), 75–100. <https://doi.org/10.1146/annurev-biochem-062917-012624>
- Wilson, M. D., Harreman, M., & Svejstrup, J. Q. (2013). Ubiquitylation and degradation of elongating RNA polymerase II: The last resort. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1829(1), 151–157. <https://doi.org/10.1016/j.bbagr.2012.08.002>
- Wilson, S. M., Datar, K. V., Paddy, M. R., Swedlow, J. R., & Swanson, M. S. (1994). Characterization of nuclear polyadenylated RNA-binding proteins in *Saccharomyces cerevisiae*. *Journal of Cell Biology*, 127(5), 1173–1184. <https://doi.org/10.1083/jcb.127.5.1173>
- Winey, M., & Culbertson, M. R. (1988). Mutations affecting the tRNA-splicing endonuclease activity of *Saccharomyces cerevisiae*. *Genetics*, 118(4), 609–617.
- Wlotzka, W., Kudla, G., Granneman, S., & Tollervey, D. (2011). The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. *The EMBO Journal*, 30(9), 1790–1803. <https://doi.org/10.1038/emboj.2011.97>
- Wolin, S. L., & Cedervall, T. (2002). The La Protein. *Annual Review of Biochemistry*, 71(1), 375–403.

- <https://doi.org/10.1146/annurev.biochem.71.090501.150003>
- Wong, K. H., Jin, Y., & Struhl, K. (2014). TFIIH phosphorylation of the Pol II CTD stimulates mediator dissociation from the preinitiation complex and promoter escape. *Molecular Cell*, 54(4), 601–612. <https://doi.org/10.1016/j.molcel.2014.03.024>
- Woychik, N. A., & Hampsey, M. (2002). The RNA Polymerase II Machinery: Structure Illuminates Function. *Cell*, 108(4), 453–463. [https://doi.org/10.1016/S0092-8674\(02\)00646-3](https://doi.org/10.1016/S0092-8674(02)00646-3)
- Wu, C.-C., Lin, Y.-C., & Chen, H.-T. (2011). The TFIIIF-like Rpc37/53 dimer lies at the center of a protein network to connect TFIIIC, Bdp1, and the RNA polymerase III active center. *Molecular and Cellular Biology*. <https://doi.org/10.1128/MCB.05151-11>
- Wyers, F., Rougemaille, M., Badis, G., Rousselle, J.-C., Dufour, M.-E., Boulay, J., Régault, B., Devaux, F., Namane, A., Séraphin, B., Libri, D., & Jacquier, A. (2005). Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell*, 121(5), 725–737. <https://doi.org/10.1016/j.cell.2005.04.030>
- Xu, C., & Min, J. (2011). Structure and function of WD40 domain proteins. *Protein & Cell*, 2(3), 202–214. <https://doi.org/10.1007/s13238-011-1018-1>
- Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Münster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W., & Steinmetz, L. M. (2009). Bidirectional promoters generate pervasive transcription in yeast. *Nature*, 457(7232), 1033–1037. <https://doi.org/10.1038/nature07728>
- Yang, C., Bolotin, E., Jiang, T., Sladek, F. M., & Martinez, E. (2007). Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene*, 389(1), 52–65. <https://doi.org/10.1016/j.gene.2006.09.029>
- Yarnell, W. S., & Roberts, J. W. (1999). Mechanism of intrinsic transcription termination and antitermination. *Science (New York, N.Y.)*, 284(5414), 611–615. <https://doi.org/10.1126/science.284.5414.611>
- Yoo, C. J., & Wolin, S. L. (1997). The yeast La protein is required for the 3' endonucleolytic cleavage that matures tRNA precursors. *Cell*, 89(3), 393–402. [https://doi.org/10.1016/s0092-8674\(00\)80220-2](https://doi.org/10.1016/s0092-8674(00)80220-2)
- Yüce, Ö., & West, S. C. (2013). Senataxin, defective in the neurodegenerative disorder ataxia with oculomotor apraxia 2, lies at the interface of transcription and the DNA damage response. *Molecular and Cellular Biology*, 33(2), 406–417. <https://doi.org/10.1128/MCB.01195-12>
- Zawel, L., & Reinberg, D. (1993). Initiation of transcription by RNA polymerase II: A multi-step process. *Progress in Nucleic Acid Research and Molecular Biology*, 44, 67–108. [https://doi.org/10.1016/s0079-6603\(08\)60217-2](https://doi.org/10.1016/s0079-6603(08)60217-2)
- Zentner, G. E., & Henikoff, S. (2013). Regulation of nucleosome dynamics by histone modifications. *Nature Structural & Molecular Biology*, 20(3), 259–266. <https://doi.org/10.1038/nsmb.2470>
- Zhao, D. Y., Gish, G., Braunschweig, U., Li, Y., Ni, Z., Schmitges, F. W., Zhong, G., Liu, K., Li, W., Moffat, J., Vedadi, M., Min, J., Pawson, T. J., Blencowe, B. J., & Greenblatt, J. F. (2016). SMN and symmetric arginine dimethylation of RNA polymerase II C-terminal domain control termination. *Nature*, 529(7584), 48–53. <https://doi.org/10.1038/nature16469>
- Zhao, X., & Herr, W. (2002). A regulated two-step mechanism of TBP binding to DNA: A solvent-exposed surface of TBP inhibits TATA box recognition. *Cell*, 108(5), 615–627. [https://doi.org/10.1016/s0092-8674\(02\)00648-7](https://doi.org/10.1016/s0092-8674(02)00648-7)
- Zhang, Y., Liu, Y., Liu, H., & Tang, W. H. (2019). Exosomes: Biogenesis, biologic function and clinical potential. *Cell & Bioscience*, 9(1), 19. <https://doi.org/10.1186/s13578-019-0282-2>
- Zhu, Y., Stribinskis, V., Ramos, K. S., & Li, Y. (2006). Sequence analysis of RNase MRP RNA reveals its origination from eukaryotic RNase P RNA. *RNA*, 12(5), 699–706. <https://doi.org/10.1261/rna.22849>

RESUME DE THESE

Étude sur les différents mécanismes de terminaison de la transcription par l'ARNpol III

Résumé

L'ARN polymérase III (ARNpol III) synthétise des ARN non codants essentiels et abondants tels que les ARNt. Contrôler l'étendue de l'activité de la ARNpol III par une terminaison précise et efficace est un défi nécessaire pour assurer une expression génétique robuste et prévenir les conflits avec d'autres machineries associées à l'ADN. On pense que le mécanisme de terminaison de l'ARNpol III est plus simple que celui des autres ARN polymérases eucaryotes, car il repose uniquement sur la reconnaissance d'une suite de T dans le brin hors matrice. Nous combinons ici des analyses à haute résolution à l'échelle du génome et des essais de terminaison de la transcription *in vitro* pour réexaminer le mécanisme de terminaison de la transcription par l'ARNpol III chez la levure bourgeonnante. Nous montrons que les suites de Ts sont nécessaires mais pas toujours suffisantes pour la terminaison et que les structures secondaires dans les ARNs naissants sont des éléments auxiliaires importants agissant en *cis*. De plus, nous montrons que l'hélicase Sen1 joue un rôle clé dans une voie de terminaison de secours. Nos résultats fournissent un modèle intégré illustrant comment de multiples mécanismes coopèrent pour assurer une terminaison efficace de la transcription de l'ARNpol III.

Introduction

La terminaison de la transcription est un processus essentiel qui fixe les frontières entre les gènes, évitant ainsi les interférences entre unités de transcription voisines. De plus, la terminaison de la transcription joue un rôle important dans le maintien de l'intégrité du génome en limitant les conflits possibles entre les ARN polymérases transcriptrices (RNAP) et les autres machineries cellulaires impliquées dans la réplication ou la réparation de l'ADN (revue dans Porrua et Libri, 2015a).

La terminaison de la transcription peut être envisagée comme un processus en plusieurs étapes consistant au recrutement de facteurs de terminaison, à la reconnaissance de motifs de séquence, à la mise en pause de l'ARNpol, et enfin à la libération de l'ARNpol et du transcrit de l'ADN. Cette dernière étape implique un remodelage d'un réseau complexe d'interactions entre l'ARNpol, l'ARN naissant et la matrice d'ADN (revue dans Porrua et al., 2016). Au sein de ce réseau, les interactions entre la polymérase et l'hybride ARN:ADN sont

considérées comme le principal déterminant de la stabilité du complexe d'élongation (CE) (Kireeva et al., 2000). La plupart des organismes eucaryotes possèdent trois ARNpol différents qui sont spécialisées dans la production de différentes classes de transcrits et semblent adopter des stratégies différentes pour terminer efficacement la transcription. L'ARNpol I est responsable de la synthèse des ARN ribosomiques ; l'ARNpol II transcrit tous les gènes codant pour les protéines et plusieurs classes de gènes non codants et l'ARNpol III synthétise des transcrits courts et abondants parmi lesquels tous les ARNr, l'ARNr 5S et plusieurs ARN non codants supplémentaires.

Les mécanismes de terminaison de la transcription des trois polymérases ont été largement caractérisés dans le modèle eucaryote *Saccharomyces cerevisiae* et beaucoup des principes découverts dans cet organisme semblent être hautement conservés de la levure à l'homme (revu dans Porrua et al., 2016). ARNpol I et ARNpol II ont besoin de facteurs protéiques accessoires pour mettre fin à la transcription. L'ARNpol I se met en pause lorsqu'elle rencontre un facteur de type Myb lié à l'ADN en aval de chaque gène d'ARNr (Merkl et al., 2014 ; Reiter et al., 2012). La libération de l'ARNpol I en pause est ensuite médiée par d'autres protéines, notamment l'exonucléase Rat1 et l'hélicase Sen1 (El Hage et al., 2008 ; Kawauchi et al., 2008), qui sont également des facteurs de terminaison majeurs pour l'ARNpol II (voir ci-dessous).

Le mécanisme de terminaison de la transcription de l'ARNpol II est plus complexe et implique l'action d'un plus grand nombre de protéines. Il existe deux voies principales de terminaison pour l'ARNpol II (revue dans Porrua et Libri, 2015a). La terminaison de la transcription au niveau des gènes codant pour des protéines repose sur un complexe multi-protéique qui est responsable du clivage co-transcriptionnel du pré-ARNm au niveau du site poly(A) et de l'ajout d'une queue poly(A). La partie du transcrit naissant en aval de ce site est ensuite ciblée par Rat1 (XRN2 chez l'homme), qui dégrade la molécule d'ARN jusqu'à ce qu'elle rencontre l'ARNpol II et favorise sa libération de l'ADN (Baejen et al., 2017 ; Kim et al., 2004 ; Park et al., 2015 ; Pearson et Moore, 2013 ; West et al., 2004).

La deuxième voie est consacrée à la terminaison de la transcription non codante et joue un rôle essentiel dans le contrôle de la transcription pervasive ainsi que dans la biogenèse des snoRNAs (Arndt et Reines, 2015 ; Porrua et Libri, 2015a). Cette voie dépend d'un complexe composé de deux protéines de liaison à l'ARN, Nrd1 et Nab3, et de l'hélicase Sen1 mentionnée plus haut (c'est-à-dire le complexe NNS). Alors que Nrd1 et Nab3 reconnaissent des motifs de séquence spécifiques qui sont enrichis dans les ARN non codants cibles, l'hélicase Sen1 induit la dissociation des CEs (Porrua et Libri, 2013b ; Porrua et al., 2012 ; Schulz et al., 2013 ; Steinmetz et al., 2006 ; Wlotzka et al., 2011). Les mécanismes d'action de Sen1 dans la transcription par l'ARN pol II ont été largement caractérisés au niveau

moléculaire par notre groupe ainsi que par d'autres groupes (Han et al., 2017 ; Hazelbaker et al., 2013 ; Leonaitė et al., 2017 ; Porrua et Libri, 2013b ; Wang et al., 2019). En bref, Sen1 utilise l'énergie de l'hydrolyse de l'ATP pour se déplacer le long de l'ARN naissant vers l'ARNpol II en train de transcrire et, lors de la pause transcriptionnelle, il entre en collision avec la polymérase et induit sa dissociation de l'ADN.

De nombreuses données soutiennent l'idée que, contrairement aux autres ARNpols, l'ARNpol III peut terminer la transcription de manière précise et efficace au niveau d'une séquence d'ADN particulière sans avoir besoin de protéines accessoires (voir Arimbasseri et al., 2013 et Porrua et al., 2016). Un terminateur typique de l'ARNpol III consiste en une suite de thymidines (T) de longueur variable dans le brin d'ADN non-matrice qui, selon le modèle actuel, est suffisant pour favoriser à la fois la pause et la libération de l'ARNpol III. Lors de la transcription d'une suite de Ts, on pense que la faiblesse de l'hybride rU:dA qui en résulte joue un rôle central dans la déstabilisation du CE de l'ARNpol III (Mishra et Maraia, 2019). La sensibilité particulière de l'ARNpol III aux hybrides rU:dA faibles par rapport aux autres ARNpols qui ne détectent pas les suites de T comme terminateurs dépendrait des interactions moins étendues entre l'ARNpol III et l'hybride ARN:ADN (Hoffmann et al., 2015). Les Ts dans le brin non-matrice jouent un rôle critique supplémentaire dans la terminaison de la transcription (Arimbasseri et Maraia, 2015), car il a été proposé qu'ils sont reconnus par les sous-unités C37 et C53 de l'ARNpol III qui contribuent également à la terminaison (Landrieux et al., 2006 ; Rijal et Maraia, 2013). Un modèle alternatif proposé par Nielsen et ses coauteurs (Nielsen et al., 2013) postule que les suites de T sont nécessaires à la pause de l'ARNpol III mais ne sont pas suffisantes pour sa libération de l'ADN. Ces auteurs ont proposé que le repliement de l'ARN naissant dans une structure en épingle à cheveux à proximité de l'ARNpol III en pause est une condition absolue pour la terminaison. Cette structure envahirait le canal de sortie de l'ARN de la polymérase, provoquant ainsi sa dissociation de l'ADN. Le mécanisme proposé rappelle la voie de terminaison dite intrinsèque décrite pour l'ARNpol bactérienne. Ce modèle dépendant de l'épingle à cheveux reste cependant très controversé car il semble en désaccord avec un grand nombre de données expérimentales antérieures (Arimbasseri et al., 2014).

Le modèle selon lequel les éléments de séquence au niveau de l'ARN et de l'ADN sont les seuls déterminants de la terminaison de l'ARNpol III a également été remis en question chez la levure de fission *Schizosaccharomyces pombe*. En effet, un rapport récent a montré que l'un des homologues de Sen1 de *S. cerevisiae* (ci-après désigné Sp Sen1) est impliqué dans la terminaison de l'ARNpol III *in vivo* (Rivosecchi et al., 2019). La délétion de ce gène qui chez *S. pombe* est non-essentiel, entraîne un déplacement global de la distribution de l'ARNpol III en aval des gènes d'ARNt, ce qui est cohérent avec l'idée que Sp Sen1, en plus des suites de

T, serait nécessaire à la terminaison de ARNpol III dans cet organisme. Le rôle précis de Sp Sen1 dans la terminaison ainsi que son mécanisme d'action n'ont cependant pas été abordés dans cette étude.

Ainsi, une grande incertitude demeure quant à la contribution relative des éléments de séquence, des structures d'ARN et des facteurs de transcription à l'efficacité de la terminaison de la transcription par l'ARNpol III. En outre, la question de savoir dans quelle mesure les différents mécanismes de terminaison sont conservés au cours de l'évolution reste ouverte.

Résultats

1) Le domaine N-terminal de Sen1 interagit avec ARNpol III

S. cerevisiae Sen1 est une protéine modulaire composée d'un grand domaine N-terminal (aa 1-975), d'un domaine hélicase central (aa 1095-1867) et d'une région désordonnée C-terminale (aa 1930-2231). Il a été démontré que le domaine N-terminal (NTD) est essentiel pour la viabilité et pour la terminaison de la transcription de l'ARNpol II (Han et al., 2020). Nous avons réalisé des expériences de co-immunoprécipitation (co-IP) suivies d'analyses de spectrométrie de masse (MS) en utilisant soit une version complète soit une version Δ NTD de Sen1 comme appât. Nos données ont montré que Sen1 interagit avec l'ARNpol III principalement via son domaine N-terminal. Un rapport récent a montré que Sen1 interagit avec le replisome via ce même domaine. Nous avons obtenu d'un collaborateur un mutant de Sen1 contenant trois substitutions au domaine N-terminal (W773A, E774A, W777A ; définissant le variant Sen1-3) qui empêche l'interaction avec le replisome sans affecter la fonction générale de Sen1 dans la transcription par l'ARNpol II. Des expériences supplémentaires de co-IP/MS utilisant Sen1-3 comme appât ont montré que ces mutations empêchent également l'interaction de Sen1 avec l'ARNpol III. De plus, nous avons montré que ARNpol III interagit avec Sen1 indépendamment du replisome. Par conséquent, Sen1 utilise la même surface protéique pour interagir avec l'ARNpol III et le replisome d'une manière mutuellement exclusive. Enfin, nous n'avons pas observé d'association claire entre l'ARNpol III et les partenaires de Sen1, Nrd1 et Nab3, suggérant une fonction indépendante de Sen1 sur la transcription de l'ARNpol III.

Dans l'ensemble, nos résultats soutiennent l'idée que Sen1 s'associe à ARNpol III et au replisome au sein de deux complexes alternatifs qui sont également distincts du complexe NNS et qui exercent probablement des fonctions différentes.

2) Sen1 est nécessaire pour une terminaison efficace de la transcription de ARNpol III *in vivo*

Le modèle le plus largement accepté pour la terminaison de la transcription de l'ARNpol III postule que les polymérases reconnaissent un élément agissant en cis composé d'une suite de thymidines sur l'ADN non-matrice et qu'il est libéré sans qu'il soit nécessaire d'avoir recours à des facteurs supplémentaires (examiné dans Arimbasseri et al., 2013 et Porrua et al., 2016). Cependant, nos résultats montrant une interaction directe entre l'ARNpol III et Sen1 nous ont incités à étudier un rôle possible de ce dernier dans la terminaison de la transcription de l'ARNpol III. À cette fin, nous avons généré des cartes à haute résolution de l'ARNpol III en train de transcrire par CRAC (crosslinking analysis of cDNAs) (Candelli et al., 2018 ; Granneman et al., 2009). Cette technique permet de déterminer précisément la position de l'ARNpol III à partir du séquençage de l'ARN naissant. Nous avons réalisé ces expériences dans des cellules WT ou *sen1-3* ainsi que dans une souche Sen1-AID (auxin-inducible degron), ce qui nous a permis d'évaluer l'effet de la déplétion de Sen1. De manière intéressante, nous avons observé une accumulation substantielle d'ARNpol III en aval du terminateur primaire (c'est-à-dire la première suite de T en aval du gène) de la plupart des gènes transcrits par l'ARNpol III chez le mutant. Ceci indique que Sen1 est globalement requis pour une terminaison pleinement efficace de la transcription par l'ARNpol III *in vivo* et que cette fonction de Sen1 repose en grande partie sur son interaction avec l'ARNpol III.

Nos données à l'échelle du génome indiquent que l'association de Sen1 avec l'ARNpol III augmente globalement l'efficacité de la terminaison de la transcription. Cependant, ces résultats sont cohérents avec une fonction de Sen1 dans la stimulation de la terminaison au niveau du terminateur primaire et/ou avec un rôle dans l'élimination des polymérases qui échappent constitutivement à la terminaison primaire. Nos résultats suggèrent que Sen1 ne joue pas un rôle prépondérant dans la terminaison primaire et favorise plutôt la libération des ARNpol IIIs qui font une pause dans les régions de terminaison secondaire. Nos données suggèrent également que Sen1 compense le manque de terminateurs efficaces dans les régions de terminaison secondaire.

Nrd1 et Nab3 sont essentiels pour la terminaison au niveau des gènes non codants transcrits par l'ARNpol II. Cependant, l'absence d'une interaction claire entre ARNpol III et les partenaires de Sen1 suggère que ces protéines pourraient ne pas participer à la terminaison de l'ARNpol III. Afin de corroborer cette notion, nous avons réalisé une expérience de CRAC d'ARNpol III dans des cellules de type sauvage et des cellules dépourvues de Nrd1. Nous avons constaté que l'absence de Nrd1 n'affecte pas la terminaison des gènes transcrits par l'ARNpol III. Puisque Nrd1 et Nab3 fonctionnent toujours comme un hétérodimère, par extension cette expérience suggère qu'aucune de ces protéines n'est impliquée dans la terminaison de la transcription par l'ARNpol III.

Puisque la même région de Sen1 interagit à la fois avec ARNpol III et le replisome, nous

avons également effectué une expérience de CRAC d'ARNpol III sur des cellules de type sauvage et *sen1-3* arrêtées en phase G1, lorsque le replisome est absent des cellules. De manière importante, nous avons observé des défauts de terminaison similaires chez le mutant dans les cellules arrêtées en G1 et dans les cellules asynchrones, ce qui indique que la fonction de Sen1 dans la transcription par l'ARNpol III est indépendante de la présence du replisome.

3) Sen1 peut induire la terminaison de la transcription de l'ARNpol III *in vitro*

Nous avons précédemment démontré que Sen1 peut promouvoir directement la terminaison de la transcription de l'ARN pol II d'une manière indépendante de la séquence de l'ARN (Porrúa et Libri, 2013b). Pour tester si Sen1 peut également induire directement la terminaison de la transcription de l'ARNpol III et si cela nécessite la présence de signaux de terminaison canoniques, nous avons utilisé un système de terminaison de la transcription *in vitro* contenant des protéines purifiées (c'est-à-dire ARNpol III et Sen1), des matrices de transcription et de l'ARN naissant. Nous avons constaté que Sen1 peut à la fois renforcer la terminaison au niveau des terminateurs inefficaces et favoriser la terminaison au niveau des séquences non liées. Nous proposons que Sen1 assiste ARNpol III pour la terminaison au niveau des terminateurs sous-optimaux et fonctionne donc comme une voie de terminaison de secours.

Nous avons également étudié le mécanisme par lequel Sen1 induit la terminaison de la transcription par l'ARNpol III. Plus précisément, nous avons montré que le domaine hélicase de Sen1 peut déclencher la libération de l'ARNpol III *in vitro* avec une efficacité similaire à celle de la version complète de Sen1. Ceci suggère fortement que l'interaction entre Sen1 et l'ARNpol III via le domaine N-terminal est importante pour le recrutement de Sen1 aux transcrits ARNpol III *in vivo*. De plus, la terminaison de ARNpol III médiée par Sen1 est ATP-dépendante et nécessite l'interaction de Sen1 avec l'ARN naissant. Ces résultats indiquent que Sen1 utilise un mécanisme similaire pour induire la terminaison de la transcription de l'ARNpol II et l'ARNpol III.

4) Les structures d'ARN peuvent favoriser la terminaison de la transcription de l'ARNpol III *in vitro*

Les résultats ci-dessus indiquent que, comme pour l'ARNpol II, la terminaison de la transcription de l'ARNpol III médiée par Sen1 implique la translocation de Sen1 le long du transcrit naissant, et nos données structurales et biochimiques antérieures ont montré que Sen1 ne peut interagir qu'avec un ARN simple brin (Porrúa et Libri, 2013b, Leonaite et al,

2017). Les ARNt sont des molécules d'ARN hautement structurées et pour une grande majorité d'entre elles, la région entre l'extrémité 3' de l'ARNt mature et le terminateur primaire est d'au plus 7 nt. Nous avons imaginé qu'une raison possible pour laquelle Sen1 ne fonctionne pas aux sites de terminaison primaire est que sa liaison à l'ARN naissant est entravée par la formation co-transcriptionnelle de structures stables à proximité du terminateur primaire. À l'inverse, des ARN moins structurés permettraient le chargement et la fonction de Sen1.

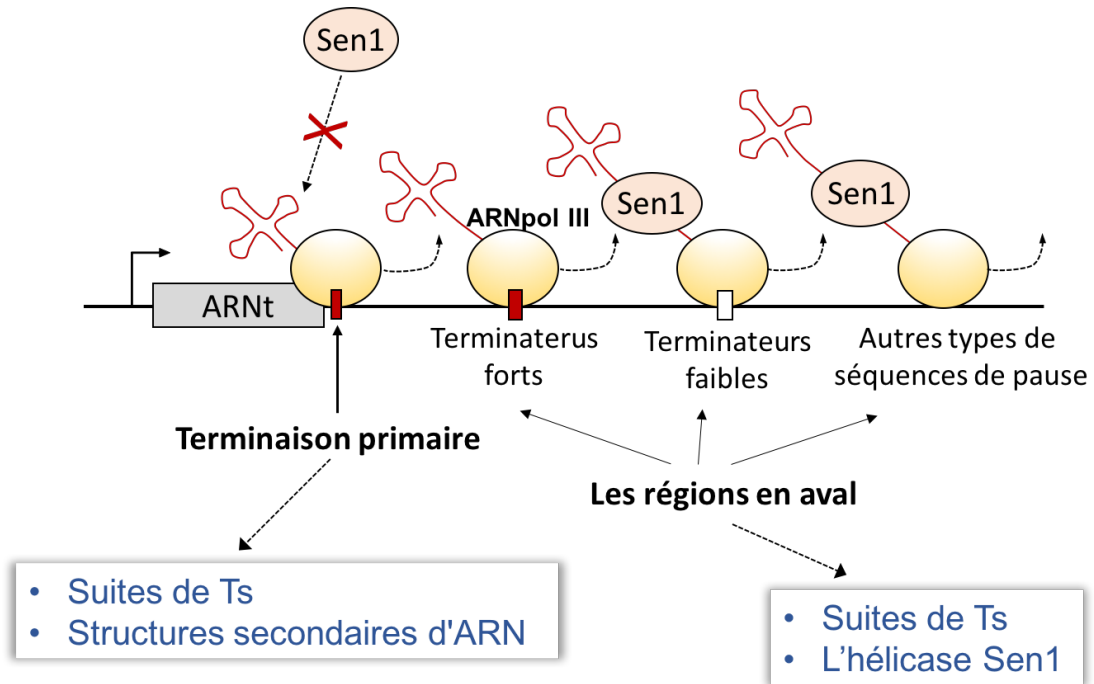
Pour explorer ces possibilités, nous avons réalisé des essais de transcription *in vitro* avec des matrices de transcription modifiées pour former une structure du type épingle à cheveux qui se trouve naturellement dans l'ARN 5S, un transcrit dépendant de ARNpol III, en amont de suites de T de différentes longueurs. De façon surprenante, nous avons constaté que la présence de cette structure dans l'ARN transcrit pouvait augmenter de façon significative l'efficacité de terminaison de la transcription au niveau d'un terminateur sous-optimal. L'effet observé est similaire au résultat de l'addition de Sen1 à la version non structurée du même ARN. Ceci indique que non seulement Sen1 mais aussi les structures secondaires de l'ARN peuvent améliorer la fonction des terminateurs inefficaces. Nous avons également montré que les structures d'ARN peuvent améliorer la terminaison de la transcription uniquement lorsqu'elles se trouvent à proximité des segments T. Nous avons également montré, avec des approches similaires, que les structures d'ARN entravent le recrutement de Sen1 à l'ARN naissant et, par conséquent, empêcheraient Sen1 de fonctionner au niveau des terminateurs primaires.

Conclusion

Dans la présente étude, nous combinons des approches à haute résolution à l'échelle du génome avec des essais de terminaison de transcription *in vitro* utilisant des composants hautement purifiés afin de disséquer le mécanisme de terminaison de transcription de l'ARNpol III chez *S. cerevisiae*. Nous observons que la terminaison au niveau du terminateur primaire des gènes dépendant de l'ARNpol III (c'est-à-dire la première suite de T après le gène), n'est que partiellement efficace et, ainsi, une fraction considérable de polymérase se terminent dans la région en aval. Nous fournissons des données *in vivo* et *in vitro* prouvant que l'hélicase Sen1 joue un rôle global dans la terminaison de la transcription par l'ARNpol III et que cette fonction repose sur l'interaction de son domaine N-terminal avec l'ARNpol III. Cependant, nous constatons que Sen1 contribue très peu à l'efficacité de la terminaison primaire et qu'il fonctionne principalement comme un mécanisme de sécurité pour promouvoir la terminaison des ARNpol III qui outrepassent le premier signal de terminaison. Nos données indiquent que seuls les suites de T dans une gamme de longueur particulière

sont suffisants pour promouvoir la terminaison autonome par l'ARNpol III. Néanmoins, nous montrons que les gènes ARNt contiennent souvent des signaux de terminaison sous-optimaux et que leur capacité à induire la terminaison peut être complétée par Sen1 ainsi que par les structures secondaires de l'ARN naissant. Ces deux facteurs agissent de manière mutuellement exclusive puisque la présence de structures d'ARN empêche l'association de Sen1 avec le transcrit, ce qui est strictement requis pour que Sen1 puisse induire la terminaison de la transcription. Bien que Sen1 puisse également promouvoir la terminaison des ARNpol IIIs qui sont arrêtées dans des sites autres que les suites de T, nous constatons que les structures d'ARN ne peuvent fonctionner qu'en association avec des signaux de terminaison canoniques, c'est à dire des suites de T.

Ensemble, nos données permettent de revoir les anciens modèles de terminaison de la transcription par l'ARNpol III et offrent une vision nouvelle et détaillée de la façon dont les composants intrinsèques du CE (c'est-à-dire les suites de T et les structures d'ARN) et le facteur extrinsèque Sen1 coopèrent pour promouvoir une terminaison efficace de la transcription par l'ARNpol III. Nous proposons que les structures de l'ARN contribuent à l'efficacité de la terminaison primaire dans certains cas (c'est-à-dire les gènes avec des terminateurs sous-optimaux), grâce à la proximité naturelle de la tige acceptrice de l'ARNt au premier T-tract, alors que Sen1 fonctionnerait préférentiellement dans les régions en aval. Une terminaison efficace est importante pour le recyclage rapide des polymerases pour de nouveaux cycles de transcription et, ainsi, pour le maintien d'une expression robuste des ARNt et d'autres transcrits dépendants de l'ARNpol III qui sont essentiels pour soutenir la prolifération cellulaire. De plus, il est crucial de prévenir ou de minimiser les conflits avec les autres polymérase en train de transcrire ainsi qu'avec les autres machineries associées à l'ADN.



Modèle pour le rôle des signaux de terminaison canoniques, des structures d'ARN et de Sen1 dans la terminaison de la transcription par l'ARNpol III (montré pour les gènes d'ARNt).

Les mécanismes régissant la terminaison de la transcription du ARNpol III *in vivo* sont considérablement plus complexes que ceux représentés dans les modèles précédents et impliquent l'interaction entre des éléments distincts agissant en cis et le facteur de terminaison extrinsèque Sen1. Au niveau du terminateur primaire, la terminaison implique généralement l'action d'une suite de T et la structure secondaire de l'ARNt naissant. Les structures d'ARN ne sont requises que pour les terminateurs de longueur non optimale. Dans les régions en aval, la transcription par l'ARNpol III est typiquement terminée soit par des terminateurs secondaires "forts", sans l'aide de Sen1, soit par des signaux de terminaison "faibles" si Sen1 peut accéder et se charger sur l'ARN naissant. D'après nos données, Sen1 peut également favoriser la terminaison au niveau des sites de pause autres que les suites de T.