Kompozycja i dynamika chromatyny w obszarach terminacji transkrypcji u człowieka

Magda Kopczyńska

Praca doktorska realizowana pod kierunkiem
dr hab. Kingi Kamieniarz-Gduli, Prof. UAM

Uniwersytet Adama Mickiewicza w Poznaniu
Wydział Biologii i Centrum Zaawansowanych Technologii
Szkoła Doktorska Nauk Przyrodniczych
Instytut Biologii Molekularnej i Biotechnologii
Zakład Biologii Molekularnej i Komórkowej

Poznań, 2025

Genome-wide analysis of chromatin composition and dynamics at human transcription termination regions

Magda Kopczyńska

Doctoral dissertation conducted under the supervision of

dr hab. Kinga Kamieniarz-Gdula, Prof. UAM

Adam Mickiewicz University in Poznań

Faculty of Biology and Center for Advanced Technologies

Doctoral School of Natural Sciences

Institute of Molecular Biology and Biotechnology

Department of Molecular and Cellular Biology

Poznań, 2025

Acknowledgements

I would like to express my gratitude to my supervisor **Kinga Kamieniarz-Gdula**. When I started the PhD, I never thought that I would gain such knowledge and skills. Thank you for your trust and support. Under your mentorship I could truly develop myself and learn how science is done.

I am also very grateful to the whole **STOPlab team**. I feel extremely privileged that I could work in such a professional but also fun environment. I would especially like to thank my assistant supervisor **Agata Stępień** – your astonishing patience and willingness to help and listen to me saved me countless times. I was also very lucky to find the best lab buddy I could have imagined – **Martyna Plens-Gałąska**, thank you for everything.

I would not have started the PhD if I had not been encouraged by my husband **Adam**. Thank you for being my rock and for motivating me to always move at least one step forward. You deserve this PhD as much as I do!

Dziękuję też mojej rodzinie. **Mamo, Tato, Bartku i Olu** – dziękuję, że we mnie wierzyliście i zawsze byliście ze mnie dumni.

To everyone who might find this work helpful

Table of contents

Streszczenie

Transkrypcja genów kodujących białka dzieli się na trzy etapy: inicjację, elongację i terminację, a enzymem odpowiedzialnym za syntezę RNA podczas tego procesu jest polimeraza RNA II (RNAPII). Na przebieg transkrypcji wpływa zarówno architektura chromatyny, jak i czynniki epigenetyczne. Nie jest to jednak relacja jednostronna – transkrypcja i chromatyna oddziałują na siebie wzajemnie. Ze względu na błędne przekonanie, zakładające, że tylko inicjacja i elongacja stanowią kluczowe etapy transkrypcji, terminacja przez stosunkowo długi czas była marginalizowana. Jednak piętrzące się dowody naukowe wskazują, że terminacja jest kluczowym etapem regulacyjnym, niezbędnym dla precyzyjnej kontroli ekspresji genów.

Niniejsza praca skupia się na wpływie, jaki organizacja i dynamika chromatyny wywiera na terminację transkrypcji. W pierwszej części pokazuję, że regiony terminacyjne można zdefiniować w genomie jako „okna" charakteryzujące się wzbogaceniem fosforylacji treoniny 4 (T4ph) w C-terminalnej domenie (CTD) polimerazy RNAPII. T4ph pełni rolę uniwersalnego markera pauzowania RNAPII, który jest obecny zarówno na końcach genów jak i w miejscach przedwczesnej terminacji transkrypcji (PTT). Dalsza analiza „okien terminacyjnych" ujawniła, że ich występowanie koreluje ze wzbogaceniem czynników odpowiedzialnych za tworzenie pętli chromatynowych. Co więcej, w „oknach terminacyjncyh" stopniowo zanika modyfikacja H3K36me3, charakterystyczna dla regionów, w których zachodzi aktywna transkrypcja.

W drugiej części analizuję rolę metylotransferazy SETD2, odpowiedzialnej za metylację H3K36me3. W większości przypadków, enzym ten jest konieczny do prawidłowej inicjacji transkrypcji, jednak dla 15–25% genów jest niezbędny do właściwej terminacji. Dla tych właśnie genów aktywność katalityczna SETD2 jednocześnie stymuluje efektywne cięcie 3' końca pre-mRNA oraz zapobiega kryptycznej inicjacji transkrypcji wewnątrz genów.

W ostatniej części, zwracam uwagę na kliniczne znaczenie prawidłowej terminacji. Skupiam się na roli PTT w utrzymaniu różnorodności transkryptomu na przykładzie rzadkiego zaburzenia neurorozwojowego – hipoplazji mostowo-móżdżkowej typu 10. PTT reguluje równowagę między różnymi izoformami mRNA i obniża poziom ekspresji danego genu. Z tego względu, niezwykle istotne jest dalsze badanie roli przedwczesnej terminacji w kontekście klinicznym.

Podsumowując, niniejsza praca kładzie nacisk na rolę, jaką stanowią końce genów. Wpływając na stabilność mRNA, jego lokalizację oraz ostateczną funkcję kodowanych białek, pełnią one istotną regulacyjną funkcję. Definicja „okien terminacyjnych" w genomie może mieć realny wpływ na ulepszanie adnotacji genomowych i identyfikację mutacji związanych z chorobami, które znajdują się w rejonach niekodujących. Ponadto, odkrycie nowej roli SETD2 w obróbce 3' końca pre-mRNA podkreśla złożoność powiązań między czynnikami epigenetycznymi a transkrypcją. Tym samym, terminacja pełni istotną funkcję regulacji całego procesu transkrypcji zarówno w warunkach fizjologicznych jak i w przypadku wielu jednostek chorobowych.

Abstract

Transcription termination of protein-coding genes represents the last step of nascent RNA synthesis by RNA polymerase II (RNAPII), following initiation and elongation. Because transcription occurs in a dynamic nuclear environment, chromatin architecture and epigenetic factors strongly influence its progression, and transcription and chromatin organization mutually shape each other. To date, most studies have focused on initiation and elongation, leaving termination comparatively overlooked due to the assumption that RNAPII had already completed its most important tasks. Yet accumulating evidence demonstrates that termination is a crucial regulatory step, essential for safeguarding transcriptional integrity and fine-tuning gene expression.

This work investigates how chromatin organization and dynamics affect transcription termination. In the first part, I show that termination regions can be successfully defined genome-wide as windows marked by threonine 4 phosphorylation (T4ph) of the RNAPII C-terminal domain (CTD). T4ph acts as a universal marker of RNAPII terminal pausing, present at both annotated gene ends and premature transcription termination (PTT) sites within gene bodies. Genome-wide definition of termination windows further reveals their association with the occupancy of chromatin looping factors and their association with a gradual decline of H3K36me3, a histone modification characteristic of active transcription.

In the second part, I explore the role of the H3K36me3-depositing methyltransferase SETD2. While SETD2 is generally required for correct transcription initiation, I demonstrate that in ~15–25% of genes it is additionally essential for proper termination. At these loci, SETD2's catalytic activity promotes efficient pre-mRNA 3' end cleavage and prevents cryptic transcription initiation within gene bodies, particularly near gene ends. Importantly, this function of SETD2 in termination occurs independently of alternative polyadenylation (APA).

Finally, I address the clinical relevance of accurate termination by showing how PTT contributes to transcriptome diversity in pontocerebellar hypoplasia type 10, a rare neurodevelopmental disorder. Since PTT regulates isoform balance and attenuates gene expression, its role in disease pathogenesis merits further investigation.

Taken together, this work emphasizes that gene ends represent important regulatory regions influencing transcript stability, localization and the eventual function of encoded proteins. Genome-wide definition of termination windows offers a valuable resource for improving genomic annotations and for identifying pathogenic variants in noncoding regions. Moreover, the discovery of a novel role for the tumor suppressor SETD2 in 3' end processing underscores the intricate interplay between epigenetic regulation and transcription, and highlights termination as a one of central safeguards of transcriptional stability in both normal physiology and disease.

List of abbreviations

AD – Alzheimer's Disease
APA – Alternative Polyadenylation
ARE – AU-rich Element
ccRCC – Clear Cell Renal Cell Carcinoma
CFIm – Cleavage Factor Im
CFIIm – Cleavage Factor IIm
CPA – Cleavage and Polyadenylation
CPSF – Cleavage and Polyadenylation Specificity Factor
CSTF – Cleavage Stimulation Factor
CTD – C-terminal Domain (of RNA Polymerase II)
ELAVL1 – Embryonic Lethal Abnormal Vision-Like Protein 1
GRE – Glucocorticoid Response Element
GTF – General Transcription Factor
IPEX – Immune Dysregulation, Polyendocrinopathy, Enteropathy, X-linked syndrome
iPSCs – Induced Pluripotent Stem Cells
mESCs – Mouse Embryonic Stem Cells
NDR – Nucleosome-Depleted Region
PAP – Poly(A) Polymerase
PAS – Polyadenylation Site
PCH10 – Pontocerebellar Hypoplasia Type 10
PD – Parkinson's Disease
PIC – Pre-initiation Complex
PST-condensates – Phase-Separated Transcriptional-Condensates
PTT – Premature Transcription Termination
RBBP6 – Retinoblastoma Binding Protein 6
RBP – RNA-Binding Protein
RNAPII – RNA Polymerase II
SEC – Super Elongation Complex
SNCA – Synuclein Alpha
TADs – Topologically Associating Domains
TSS – Transcription Start Site
UTR – Untranslated Region

Introduction

Transcription termination: mechanisms and challenges

Transcription of protein-coding genes and non-coding transcription units by RNA Polymerase II (RNAPII) controls every step of cellular development and growth, as well as metabolic and signaling pathways. Transcription is a multi-step process that can be divided into three main phases: initiation, elongation, and termination[1]. All stages are precisely regulated by specific phosphorylations of the C-terminal domain (CTD) of RNAPII, in which distinct residues within the heptapeptide repeats are modified at different steps of transcription[2–4]. These modifications facilitate binding of unique proteins such as capping enzymes[5] and 3' end processing factors[6].

Transcription initiates when RNAPII is recruited to the gene promoter by the preinitiation complex (PIC)[7]. Approximately 20-40 nucleotides downstream of the transcription start site (TSS) RNAPII pauses[8]; it can then either be released into productive elongation[9], or displaced from the DNA by promoter-proximal premature termination, resulting in attenuated gene expression[10]. After RNAPII transcribes the polyadenylation signal at the 3' end of the gene, pre-mRNA is cleaved and polyadenylated to produce a mature transcript[11–13]. In mammals, RNAPII continues synthesizing RNA for several hundred to thousands of nucleotides downstream of the polyadenylation site (PAS)[14,15]. Transcription ultimately terminates when the polymerase detaches from the DNA template (Figure 1A). It is important to terminate transcription properly, as disrupted termination leads to transcriptional interference, which can result in the downregulation of expression of a downstream gene (Figure 1B)[16–18]. Such delayed termination, known as readthrough, has been shown to arise in response to various cellular stresses[19,20], viral infection[21,22], and cancer[23].

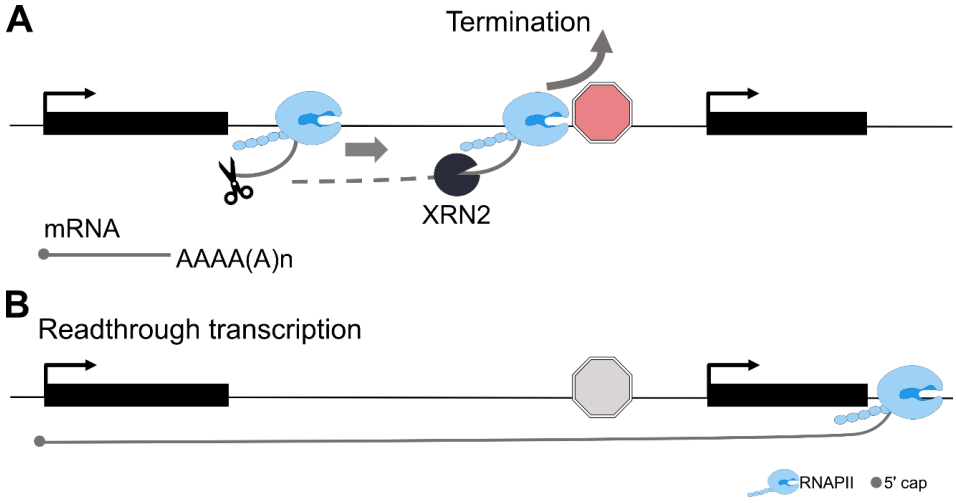


Figure 1. Model of transcription termination and readthrough. A) In mammals, transcription termination typically occurs ~2 kb downstream of the PAS. After RNAPII transcribes the polyadenylation signal, it undergoes allosteric changes, and the 5′ end of the cleaved RNA provides an entry site for the XRN2 exonuclease. Termination is achieved when XRN2 catches up with RNAPII and displaces it from the DNA. B) Disruption of termination can lead to transcriptional readthrough. This may cause transcriptional interference if a neighboring downstream gene is located nearby, potentially resulting in aberrant gene expression.

Mechanistically, there are two models describing how termination is triggered. They were initially described separately, but are now regarded as unified processes[24,25]. In the first model – the allosteric model – transcription of the polyadenylation signal induces allosteric changes in RNAPII. These changes occur through the exchange of associated factors or alterations in protein modifications (including RNAPII phosphorylation), ultimately resulting in termination[26,27]. The second model, known as the torpedo model, involves the 5'-to-3' exonuclease XRN2, which degrades RNAPII-associated RNA exposed following 3' end cleavage (Figure 1A)[28,29]. When XRN2 catches up with transcription machinery, RNAPII is displaced from the DNA template, leading to termination[30,31] However, recent evidence from yeast has revised the current torpedo model. According to the new data, Xrn2 forms a stable complex with elongating RNAPII, implying that a chasing mechanism may not be involved[32].

In addition to gene-end termination, protein-coding genes in metazoans frequently undergo premature transcription termination (PTT) – also known as transcription attenuation – which plays a critical regulatory role. This process can be triggered by factors such as PCF11 at intronic polyadenylation sites[15], the Integrator complex near promoters[10,33,34], in response to heat shock[35] or disruption of RNA processing factos[36,37]. In addition, premature termination occurs also on non-coding transcripts stimulated by the Restrictor complex[38]. Although premature termination may appear energetically wasteful, its evolutionary conservation across prokaryotes, yeast, and metazoans, as well as its role in regulating both coding and non-coding transcripts, demonstrates that it has essential regulatory functions[39].

The study of transcription termination was historically challenging because appropriate methodological tools were lacking. However, the introduction of nascent transcriptomics technologies – such as chromatin-bound RNA-seq, GRO-seq, PRO-seq, mNET-seq, TT-seq, and POINT-seq – has markedly advanced the field[14,37,40–42]. These methods rely either on immunoprecipitating RNAPII bound to newly synthesized RNA (mNET-seq and POINT-seq), on incorporation of biotinylated NTPs in nascent RNA (PRO-seq, GRO-seq) or metabolic labelling with 4-thiouridine (TT-seq). Nonetheless, these approaches still present difficulties in pinpointing the exact genomic windows where RNAPII dissociates from the DNA template (Figure 2). This is due to the need for defining a threshold that distinguishes diminishing signal downstream of the gene end from the intergenic signal originating from pervasive transcription of non-coding regions[43]. This challenge is even more pronounced when assessing premature termination events within gene bodies.

To overcome these limitations, scientists have taken advantage of long-read sequencing platforms such as PacBio or Nanopore, combining them with nascent RNA isolation to develop new methods like POINT-nano[37] and FLEP-seq[44,45]. These approaches enable the simultaneous investigation of 3' end cleavage of nascent RNA and termination on a single RNA molecule. However, they remain more expensive than short-read sequencing methods, are biased toward specific transcript lengths[46], and require substantially greater analysis time and computational resources[47].

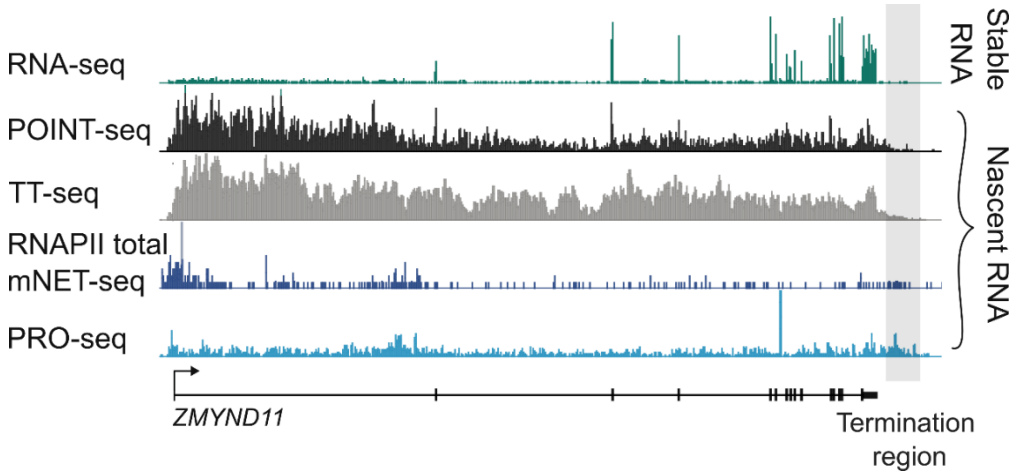


Figure 2. Comparison of nascent RNA techniques illustrating the challenge of defining transcription termination (indicated in grey) based on signal decrease. The first track shows stable RNA detected by RNA-seq, which captures mainly exon-derived sequences. Data from: GSM7179805, GSM4826609, GSM7119959, GSE81662, GSM2692352 (from the top, respectively).

Chromatin and transcription: reciprocal relationship

Transcription occurs within a highly dynamic chromatin environment and is significantly influenced by changes in chromatin structure. DNA wrapped around the nucleosome forms one of the most stable protein-DNA complexes under physiological conditions[48]. Nevertheless, nucleosomes are not static entities; rather, they are dynamic structures that undergo various repositioning events and are subject to multiple chemical modifications.

Histone proteins can be posttranslationally modified at both their tail regions and globular domains. Lysines (K) may be acetylated, methylated, ubiquitinated, ADP-ribosylated, or sumoylated; arginines (R) primarily undergo methylation; and serines (S) and threonines (T) can be phosphorylated[49]. Specific histone modifications are correlated with both active and inactive genes; however, they are not necessarily the cause of particular transcriptional events and may instead be a consequence of transcription itself[50–52]. Active promoters are typically marked by acetylation of histones H3 and H4[53], as well as H3K4me3[54]; active enhancers are associated with H3K27ac and H3K4me1[55]. Additionally, gene bodies of actively transcribed genes are decorated with H3K36me3[56], H3K79me3[57] and H2BK120u1[58].

In addition to histone modifications, chromatin is subject to numerous remodeling events mediated by protein complexes that use ATP hydrolysis to alter its structure. The consequences of chromatin remodeling include nucleosome assembly, changes in chromatin accessibility, and the installation or removal of histone variants[59]. From the perspective of transcription, the most prominent role is played by remodelers of the SWI/SNF subfamily. These remodelers reposition nucleosomes by ejecting or sliding them, thereby exposing binding sites for transcription factors at gene promoters or enhancers[60]. SWI/SNF complexes can facilitate gene expression by increasing chromatin accessibility for RNAPII and its associated factors[61], or repress transcription by inserting nucleosomes into the nucleosome-depleted region (NDR) near the gene promoter[62].

However, the most prominent mechanism of gene silencing depends on DNA methylation within CpG islands located in gene promoters[63,64]. Methylation of these CpG-rich regions leads to stable repression by preventing transcription factor binding and recruiting repressive complexes, including Polycomb proteins. During embryonic development, methyl groups are deposited de novo by two DNA methyltransferases, DNMT3A and DNMT3B[65]. While both enzymes contribute to establishing new methylation patterns, DNMT3A is specifically required for methylating promoters and gene bodies of Polycomb group target developmental genes, whereas DNMT3B preferentially methylates the X chromosome and repetitive elements[66]. To maintain correct gene expression in somatic cells, established DNA methylation patterns are preserved during cell division by the DNMT1 methyltransferase[67].

Transcription and chromatin dynamics are closely intertwined, with evidence supporting both models of interaction: chromatin structure influences transcription, while transcriptional activity also drives changes in chromatin. Depending on the context, chromatin remodelers can be recruited in advance to facilitate transcription initiation, or the transcription machinery itself can promote the recruitment of remodeling complexes. For example, the SWI/SNF complex is typically recruited to gene promoters prior to the full assembly of the transcription initiation machinery[68]. However, at certain promoters, pioneer factors first bind to chromatin and subsequently recruit SWI/SNF, thereby facilitating the recruitment of coactivator complexes[69].

Gene promoter activation involves both the action of nearby enhancers and the removal of nucleosomes. As chromatin becomes more open and accessible, the PIC composed of RNAPII and six general transcription factors (GTFs) – assembles[70]. Two histone marks, H3K4me3 and H3K27ac, are strongly correlated with active promoters; however, their direct causal involvement in initiation remains debated[70]. Notably, some studies indicate that H3K4me3 can enhance PIC formation and stimulate the expression of p53-dependent genes[71], as well as promote transcription *in vitro*[72]. On the other hand, methylation of H3K4 is not essential for transcription in *Drosophila*[73,74], and acetylation of H3K27ac at canonical H3 appears likewise dispensable, as point mutations at H3K27 are sufficient to induce transcription of genes previously repressed by Polycomb[75]. These findings suggest that the causal relationship between histone marks and initiation is highly context dependent.

Following promoter-proximal pausing, RNAPII transitions to productive elongation. During this phase, progressive nucleosome disassembly is facilitated by the histone chaperone FACT, which removes H2A-H2B dimers, disrupts the histone octamer, and then reassembles nucleosomes behind RNAPII[76]. In addition to FACT, transcription through nucleosomes is promoted by the chromatin remodeler Chd1, which, together with FACT, functions in the presence of elongation factors Spt4/5 and TFIIS[77]. Additionally, the N-terminal tails of histones on actively transcribed genes are marked by H3K36me3, a modification that progressively accumulates toward the 3' end of the gene[78].

In contrast to initiation and elongation, the role of chromatin in transcription termination – especially in mammals – remains poorly understood. Evidence from yeast indicates that

nucleosomes can promote termination by serving as barriers to the transcription machinery[79]. Similarly, transcription factors such as Reb1 and Rap1 can trigger termination simply by binding to DNA[80]. Moreover, increased nucleosome density at 3' gene ends in mammals correlates with RNAPII accumulation, suggesting that nucleosome positioning after the PAS may induce polymerase pausing and facilitate termination[81]. Nevertheless, comprehensive data for these chromatin-related mechanisms in mammalian systems are still limited.

SETD2 methyltransferase: role in transcription

One of the most prevalent histone marks decorating the gene bodies of actively transcribed genes is H3K36me3. This modification is deposited co-transcriptionally by the SETD2 methyltransferase, which is recruited through interaction with the RNAPII CTD phosphorylated at serine 2 (S2ph)[82]. As transcription progresses, DNA is unwrapped from nucleosomes, rendering them substrates for SETD2-mediated methylation. SETD2 preferentially deposits H3K36me3 on the upstream histone octamer (Figure 3A), in coordination with the histone chaperone SPT6[78]. H3K36me3 mark gradually increases within the gene body and peaks at PAS (Figure 3B).

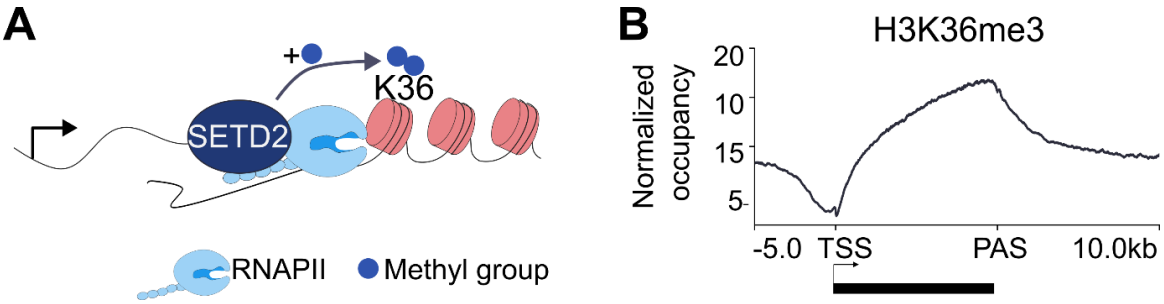


Figure 3. SETD2 and H3K36me3. A) SETD2 deposits H3K36me3 on the upstream histone octamer. B) H3K36me3 distribution on protein-coding genes (n = 4370), data from GSM1177767.

In budding yeast, H3K36me2/3 promotes chromatin restoration during transcription through the recruitment of the histone deacetylase Rpd3S, which, in turn, prevents cryptic transcription initiation from within gene bodies[83]. In mammals, SETD2-mediated H3K36me3 is also thought to suppress spurious initiation events, but this process operates independently of histone deacetylation. In mouse embryonic stem cells (mESCs), H3K36me3 recruits Dnmt3b, which methylates CpG islands within actively transcribed genes, thereby suppressing cryptic initiation[84]. By contrast, direct mutation of H3K36 in *Drosophila* does not result in spurious initiation events – likely because flies possess low levels of DNA cytosine methylation and lack DNMT3 and DNMT1 methyltransferases[85,86]. In human cells, RNA-seq revealed that SETD2 downregulation is associated with cryptic initiation events within gene bodies. However, since this technique captures only mature RNA, the evidence provided is indirect[87].

SETD2 has also been reported to influence splicing. *SETD2*-deficient clear cell renal cell carcinoma (ccRCC) tumors display pervasive alternative splicing events and widespread splicing deficiencies[88]. Additionally, the specific histone variant H3.3K36me3 recruits ZMYND11, a chromatin reader that associates with the spliceosome and regulates intron

retention[89]. As emphasized in the earlier chapter, chromatin and transcription mutually influence each other; similarly, RNA processing events such as splicing can regulate SETD2 activity. In mammals, H3K36me3 is primarily deposited on genes with introns, in proportion to transcriptional activity, whereas intronless genes consistently display significantly lower H3K36me3 levels. Furthermore, splicing inhibition impedes SETD2 recruitment and substantially reduces H3K36me3 levels[90]. Similarly to cryptic initiation, H3K36 is not required for proper splicing in flies. However, *Drosophila* cells carrying point mutation at H3K36 display defective mRNA polyadenylation, implicating this histone mark in mRNA maturation[91].

SETD2 has been also implicated in transcription termination. Analysis of total RNA-seq data from ccRCC tumors shows that mutations causing catalytic inactivation of SETD2 result in the generation of chimeric transcripts that span adjacent transcriptional units. This finding suggests that SETD2 activity – and, by extension, H3K36me3 – serves to prevent transcriptional readthrough and intergenic splicing events[23]. However, these conclusions were not corroborated by nascent RNA sequencing methods, leaving the mechanisms underlying SETD2/H3K36me3-driven transcription termination unresolved.

Alternative polyadenylation

Transcription termination is closely linked to 3' end processing, but it is important to emphasize that these are distinct molecular events. The site of nascent RNA cleavage – cleavage and polyadenylation site (PAS) – determines the end of the 3' untranslated region (3' UTR). Together with alternative promoter usage, which defines various 5' UTRs, selection of alternative cleavage sites, accounts for extensive variation in transcript isoforms[92]. Messenger RNAs that share identical coding sequences but differ in their UTRs can exhibit significant differences in stability and cellular localization. Most human genes possess multiple polyadenylation sites, and their utilization is highly context- and tissue-dependent[93].

The primary machinery responsible for 3' end processing of pre-mRNA is the cleavage and polyadenylation complex (CPA), which comprises four subcomplexes: cleavage stimulation factor (CSTF), cleavage and polyadenylation specificity factor (CPSF), cleavage factor I (CFIm), and cleavage factor II (CFIIm), alongside other proteins such as poly(A) polymerase (PAP) and symplekin[94]. Many proteins in these subcomplexes function as RNA-binding proteins. Within CPSF, CPSF30 and WDR33 recognize and bind the canonical polyadenylation signal, the hexameric sequence AAUAAA[95]. In humans, this consensus motif is typically located approximately 21 nucleotides upstream of the cleavage site[96]. The endonucleolytic cleavage of pre-mRNA is catalyzed by CPSF73, after which PAP adds the poly(A) tail to the newly generated 3' end of the transcript[97].

Alternative polyadenylation (APA) predominantly occurs in the 3' UTRs of genes, producing various mRNA isoforms with differing 3' UTR lengths (Figure 4)[94]. These 3' UTRs contain specific sequence motifs and structural elements that are recognized by regulatory factors such as microRNAs (miRNAs) and RNA-binding proteins (RBPs),

which collectively influence mRNA stability, localization, and translation efficiency[98]. In mammals, more than 50% of conserved miRNA target sites are located within alternative UTRs[99]. Early studies demonstrated that both cancer cells and activated T cells exhibit global 3' UTR shortening relative to non-cancerous or non-activated counterparts, and therefore escape repressive effects of miRNAs[99,100]. Over the past decade, research interest in the roles of 3' UTRs and APA have grown substantially, underscoring the critical impact of non-coding elements on cellular physiology and metabolic regulation.

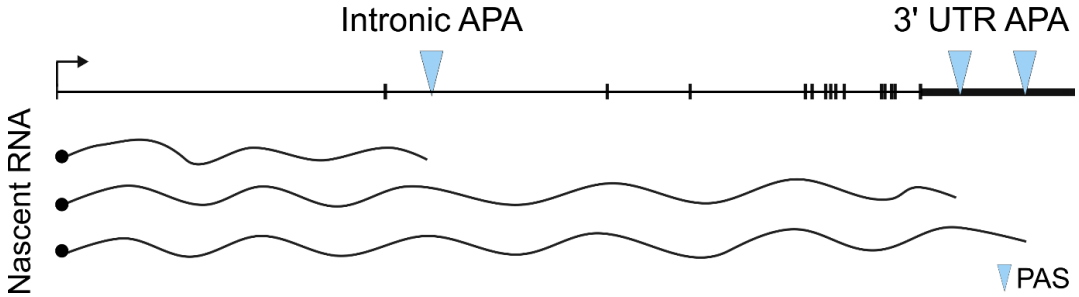


Figure 4. Alternative polyadenylation (APA) generates transcript diversity. Alternative PASs can be located within 3' UTRs or introns.

3' UTRs are rich in sequence motifs that function as destabilization elements: AU-rich elements (AREs), GU-rich elements (GREs) and PUF protein-binding elements[101]. These motifs are recognized by specific RBPs, which modulate mRNA stability. For instance, in human IFN-regulatory factor 5 (IRF5), a genetic polymorphism produces two APA isoforms with differing decay rates due to the presence of an ARE in the alternative UTR; the less stable transcripts are associated with increased risk for systemic lupus erythematosus[102]. Additionally, APA can impact mRNA nuclear export and localization – isoforms with longer 3' UTRs tend to be more prevalent in the nucleus compared to the cytoplasm[103]. Beyond mRNA localization, APA events can also influence the localization of the encoded protein: in the cytoplasm, embryonic lethal abnormal vision-like protein 1 (ELAVL1) binds to AREs in the extended 3' UTR of the *CD47* gene, facilitating the translocation of CD47 protein from the endoplasmic reticulum to the plasma membrane[104].

In addition to altering the 3' UTR sequence composition, alternative polyadenylation can also result in changes to the coding sequence, thereby influencing protein diversification[105]. A classic example of APA affecting regions upstream of the 3' UTR is the immunoglobulin M (IgM) heavy chain gene, first described in 1980. During B cell activation, PAS usage switches from a distal PAS in the most downstream exon to a proximal PAS in a terminal exon. This switch causes a shift from production of membrane-bound IgM protein to its secreted antibody form[106]. APA upstream of the 3' UTR may also yield truncated proteins. For instance, the polyadenylation factor retinoblastoma-binding protein 6 (RBBP6) undergoes differential RNA processing, resulting in several transcript isoforms. One isoform, produced by the usage of an intronic PAS, encodes a truncated protein (Iso3) that competes with full-length RBBP6 for binding to CPSF/CSTF complexes. This competitive interaction inhibits 3' end processing and modulates APA[107,108].

Alternative polyadenylation and transcription termination in disease

Choice of PAS is recognized as a key mechanism responsible for transcription isoform diversity and for modulating the function of the encoded protein. Accordingly, PAS usage is often disrupted in human diseases, including various cancers, as well as immunological and neurological disorders[98]. Remarkably, even the loss or gain of a single polyadenylation signal can have profound consequences. For example, immunodysregulation polyendocrinopathy enteropathy X-linked (IPEX) syndrome results from a mutation in the polyadenylation signal sequence of the *FOXP3* gene[109]. FOXP3's proper polyadenylation is thus critical, as its disruption directly leads to regulatory T cell deficiency and the onset of autoimmunity in IPEX syndrome[110].

Importantly, APA plays a pivotal regulatory role in central nervous system fundamentally affecting neuronal transcript isoform diversity. Neurons are generally characterized by longer 3' untranslated regions (3'UTRs)[111], which emphasizes the complex regulation of gene expression in this system. APA has been reported to play a significant role in the pathogenesis of neurodegenerative disorders by altering the 3' UTR lengths of disease-specific genes[112]. For example, in the brains of patients with Alzheimer's disease (AD), the *BIN1* gene expresses isoforms with longer 3' UTRs, potentially leading to increased BIN1 protein abundance[112,113]. BIN1 is involved in synaptic vesicle endocytosis and has already been identified as a risk factor for AD. In Parkinson's disease, the formation of Lewy bodies – intraneuronal aggregates rich in misfolded α-synuclein (SNCA) – disrupts normal cellular processes, contributing directly to neurodegeneration and the characteristic motor symptoms of the disorder[114]. It has been demonstrated that in PD patients, the *SNCA* gene undergoes distal APA, resulting in protein isoforms that localize to mitochondria and exhibit an increased propensity for aggregation[115].

Global changes in APA are frequently driven by alterations of CPA factors. For example, reduced levels of CPSF5 (also known as NUDT21) in glioblastoma are strongly associated with global 3' UTR shortening, increased cell proliferation, and poor patients' survival outcomes[116]. Similar downregulation of CPSF5 is observed in bladder cancer and hepatocellular carcinoma, contributing to shortened 3' UTRs length and adverse clinical prognosis[117,118]. In neuroblastoma cells, depletion of CPA complex component PCF11 triggers widespread transcript lengthening and is linked to spontaneous tumor regression and favorable outcomes[119]. Collectively, mounting evidence demonstrates that various cancers are characterized by a predominance of shorter 3' UTRs, a change that enhances proliferative capacity and alters cellular behavior[98,120].

Since transcription termination is often viewed as a linear outcome of nascent RNA cleavage and polyadenylation, direct evidence for how correct or defective termination impacts cell physiology and metabolism remains limited. Most insights come from studies employing nascent RNA sequencing to examine the effects of cellular stress on transcriptional dynamics. Notably, the induction of osmotic or oxidative stress in cultured human cells leads to widespread defects in transcription termination, resulting in the

generation of downstream-of-gene transcripts (DoGs)[19,20]. These readthrough transcripts have been proposed to play a protective role in maintaining nuclear integrity during stress[121], counterintuitive to the potentially disruptive effects of transcriptional interference with neighboring genes[122].

Additionally, termination is perturbed during viral infection. The NS1 protein of influenza A virus binds to CPSF30 – a subunit of the CPSF complex – thus blocking CPSF binding to the RNA substrate and inhibiting 3' end cleavage and polyadenylation of host pre-mRNAs. This results in extensive transcriptional readthrough, as host mRNA processing is disrupted[123]. Similarly, infection with herpes simplex virus 1 triggers widespread readthrough transcription: the viral protein ICP27 directly interacts with CPSF subunits and interferes with proper assembly of the CPSF complex, thereby impeding mRNA 3' end formation and termination[124]. As a consequence, translation of hundreds of host genes can be halted[21].

Transcriptional readthrough events have been also observed in cancer. As described in the chapter "SETD2 methyltransferase: role in transcription", in ccRCC, readthrough transcription leads to the production of chimeric transcripts and correlates with poor patients' prognosis[23]. Fusion transcripts have been also reported in breast cancer[125] and prostate cancer[126], and may give rise to circular transcript isoforms[127]. These chimeric transcripts are likely a consequence of transcription termination defects and may even result in the production of novel proteins[25].

To summarize, APA shapes mRNA isoform diversity and gene regulation, and its disruption is broadly linked to immune, neurological, and cancer pathologies. 3' end processing and termination are frequently perturbed by stress, viral infection, and oncogenic states, leading to readthrough transcription and aberrant chimeric RNAs. For these reasons, investigating termination mechanisms is crucial for understanding disease and identifying novel therapeutic opportunities.

Aims and research questions

The chromatin context of transcription termination remains insufficiently explored. While there is some evidence supporting roadblock-induced termination in yeast, as well as a role for nucleosome positioning in this process, the precise pathways and mechanisms by which chromatin influences transcription termination – particularly in mammals – are still unclear.

To address this knowledge gap, this study aims to:

1. Define transcription termination regions genome-wide using RNAPII C-terminal domain phosphorylation at threonine 4 as a specific termination marker.
2. Examine correlations between chromatin marks/chromatin-associated factors and established termination regions to elucidate their involvement in termination.
3. Investigate the role of the chromatin modifier SETD2 in transcription termination.

Results

Defining gene ends: RNA polymerase II CTD threonine 4 phosphorylation marks transcription termination regions genome-wide (published)

Magda Kopczyńska, Upasana Saha, Anastasiia Romanenko, Takayuki Nojima, Michał R Gdula, Kinga Kamieniarz-Gdula

Nucleic Acids Research, Volume 53, Issue 2, 27 January 2025, gkae1240

doi: https://doi.org/10.1093/nar/gkae1240

This publication is included as an appendix.

Investigating transcription termination is especially challenging because, unlike other regulatory processes where factors remain associated with chromatin, termination involves the detachment of RNAPII from the DNA template. This absence of a stable, traceable intermediate makes it exceptionally difficult to capture and study the molecular events that define transcription termination (described in more detail in the chapter "Transcription termination: mechanisms and challenges"). In contrast, assessing the appearance of a termination mark is more straightforward. In this article, we argue that the C-terminal domain of RNAPII phosphorylated on threonine 4 (CTD-T4ph) can serve as a useful marker for RNAPII termination, specifically marking terminal pausing regions.

In this work, we comprehensively review current literature on RNAPII CTD-T4ph, with a particular emphasis on its role in transcription termination in metazoans. We systematically analyze all available genomic datasets concerning T4ph, revealing that, across species, T4ph is consistently enriched at the ends of active protein-coding genes. Using T4ph enrichment, we define termination windows, which coincide with a significant decrease in nascent RNA signal. Notably, methods sensitive to RNAPII pausing detect a pronounced signal enrichment within these windows. Taken together, our findings establish CTD-T4ph as a marker for transcription termination at protein-coding genes in metazoans. This study has the potential to enhance genome annotations by providing detailed information about the intervals where transcription terminates. Such insights can improve the identification of mutations within non-coding regions that may be associated with various diseases, thereby facilitating more precise mapping of regulatory elements.

My contribution to this work included performing bioinformatic analyses of all T4ph and nascent RNA datasets, visualizing the results through figure preparation and creating UCSC Genome Browser sessions. Additionally, I participated in manuscript writing and editing.

Enrichment of chromatin factors in transcription termination regions (unpublished)

Defining termination windows genome-wide

To define genomic regions where transcription terminates – here referred to as termination windows – I employed a bioinformatic pipeline detailed in Publication 1. Briefly, I used T4ph mNET-seq datasets from HeLa cells published in two separate studies as input[15,128]. I then identified genomic intervals with significant enrichment of T4ph mNET-seq signal and determined their overlap with coordinates of pre-selected protein-coding genes (see "Materials and methods" section). This approach allowed me to distinguish normal termination windows associated with gene ends from premature terminators located within gene bodies (Figure 5). As a result, I defined 3,981 normal terminators and 1,502 premature termination windows.

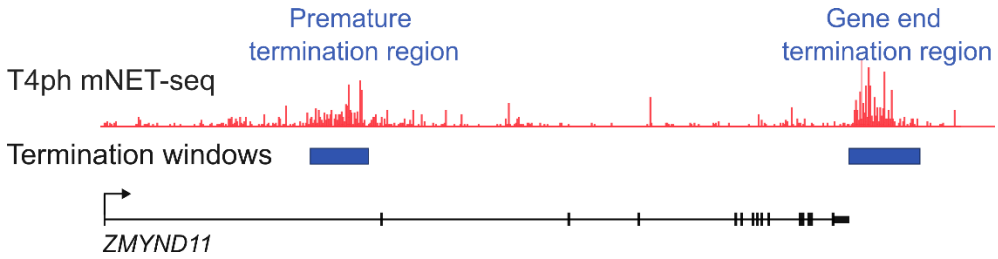


Figure 5. Termination windows were defined based on the five biological replicates of T4ph mNET-seq published in two different studies. The regions of significant enrichment were established using macs2 tool.

After defining these termination windows, I used them as input for downstream analyses. I selected 17 ENCODE ChIP-seq datasets performed on HeLa cells, covering various histone marks and chromatin-associated factors (Table 1). To assess enrichment of these features in termination windows, I calculated the average signal of each factor within the termination windows and compared it to flanking control regions (Figure 6). Here, I describe the four factors that were most significantly enriched/changed in termination windows relative to control regions.

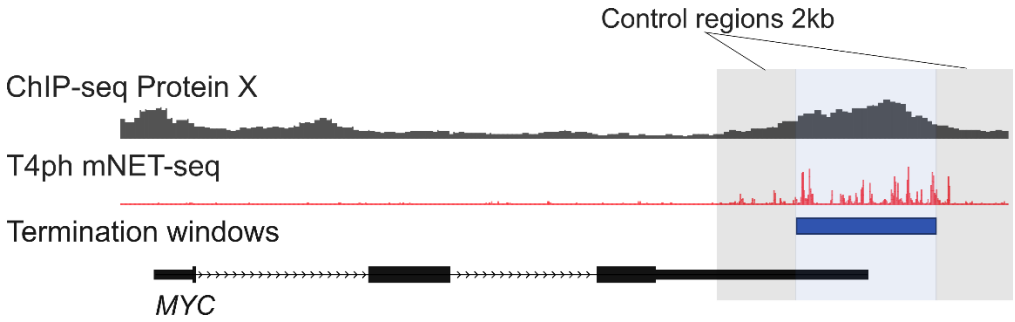


Figure 6. To assess the association of chromatin factors with termination regions, the average signal of the protein/histone mark of interest was calculated in termination windows and flanking control regions.

Termination windows are enriched in factors involved in chromatin looping

In silico data mining revealed significant enrichment of three key players in chromatin loop formation – CTCF, SMC3, and MAZ – within both normal and premature termination windows. Genomic DNA is organized into loops and topologically associating domains

(TADs) through the coordinated action of CTCF and cohesin complexes. Cohesin loads onto DNA and extrudes loops by sliding until it encounters CTCF, which functions as a loop anchor[129]. SMC3 is a cohesin subunit[130], while MAZ acts to stabilize CTCF binding[131].

Table 1. Histone marks and chromatin modifiers analyzed in the *in silico* data mining.

Factor	Data GEO accession number	Control region localization	p-val termination window vs control region - normal terminators	p-val termination window vs control region - premature terminators
BRCA1	GSM935552	upstream	5.71E-28	3.44E-10
		downstream	4.62E-92	7.16E-18
CTCF	GSM733785	upstream	2.09E-83	2.44E-33
		downstream	7.92E-87	7.88E-45
E2F1	GSM935484	upstream	9.00E-57	1.59E-11
		downstream	5.69E-21	2.48E-22
H3K36me3	GSM733711	upstream	2.74E-268	4.35E-01
		downstream	5.31E-216	6.34E-11
EP300	GSM935500	upstream	2.34E-05	0.058
		downstream	4.00E-40	0.003
LSD1	GSM1104352	upstream	2.25E-01	7.03E-05
		downstream	4.98E-65	4.10E-11
MAZ	GSM935272	upstream	4.49E-62	1.47E-39
		downstream	6.28E-99	5.94E-47
MYC	GSM822286	upstream	7.47E-37	4.40E-11
		downstream	1.81E-50	2.18E-20
SMARCA4	GSM935511	upstream	1.21E-44	1.14E-05
		downstream	6.72E-59	5.52E-13
SMC3	GSM935384	upstream	2.44E-68	3.74E-33
		downstream	9.60E-107	3.58E-41
TBP	GSM935606	upstream	1.97E-13	0.15
		downstream	4.14E-06	0.385
H3K4me3	GSM733682	upstream	1.12E-32	1.20E-21
		downstream	9.98E-57	9.62E-42
H3K27me3	GSM733696	upstream	1.12E-30	1.68E-19
		downstream	2.28E-89	1.08E-35
H3K79me2	GSM733669	upstream	4.19E-06	8.30E-02
		downstream	2.56E-54	2.04E-10
H3K27ac	GSM733684	upstream	1.12E-30	1.68E-19
		downstream	2.28E-89	1.08E-35
H3K4me1	GSM798322	upstream	9.25E-21	1.70E-02
		downstream	1.99E-35	3.22E-07
R-loops	GSM2891661	upstream	6.06E-10	4.00E-03
		downstream	3.70E-31	2.42E-07

The enrichment of these looping factors is apparent from average ChIP-seq signal analyses within termination windows compared to flanking control regions and is visually evident in the corresponding heatmaps (Figure 7A,B,C). Notably, this enrichment is more pronounced in premature termination windows, particularly for SMC3 and MAZ.

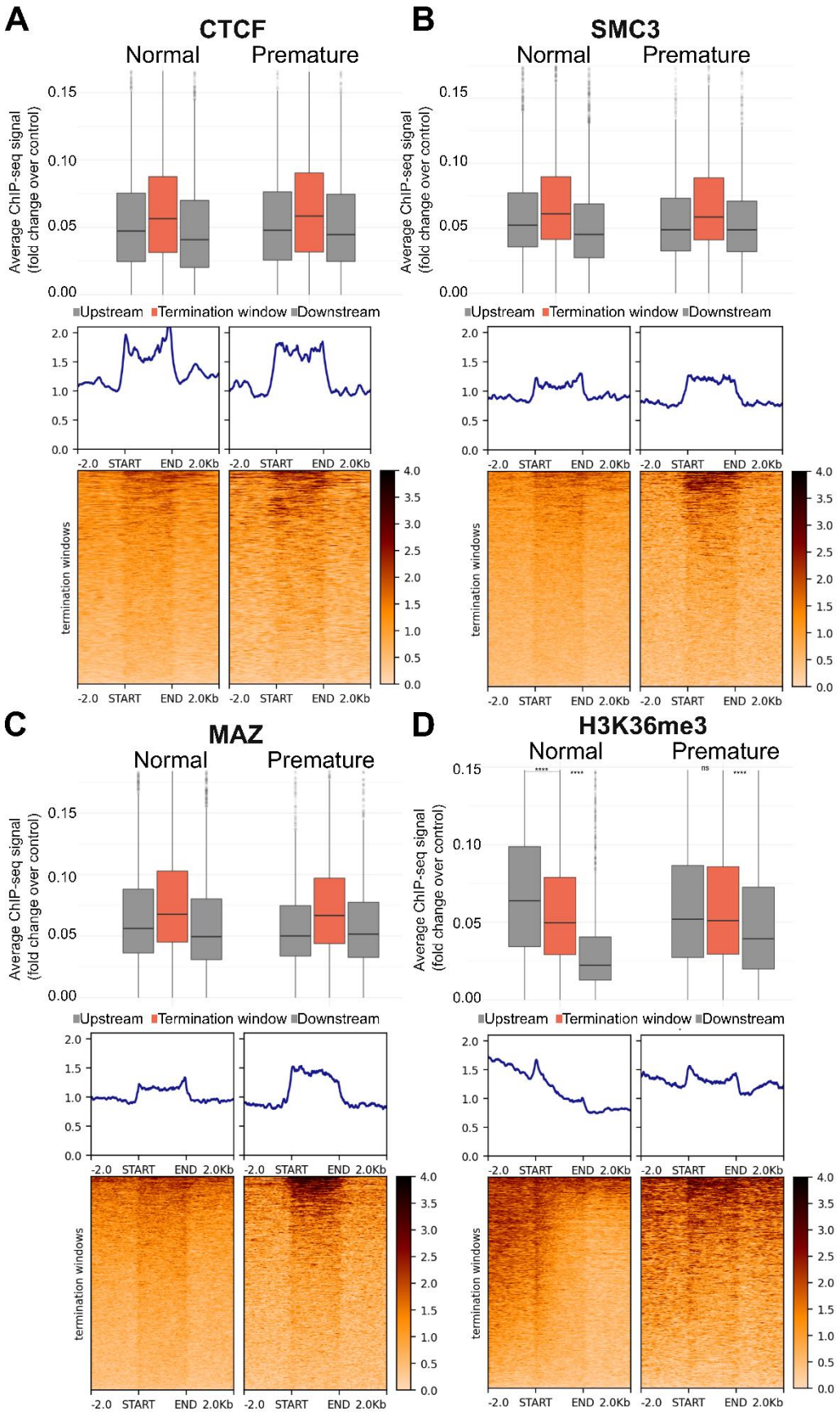


Figure 7. Enrichment of selected factors calculated as shown in Figure 6 (top panel) and represented as a heatmap (bottom panel) for: A) CTCF, B) SMC3, C) MAZ and D) H3K36me3.

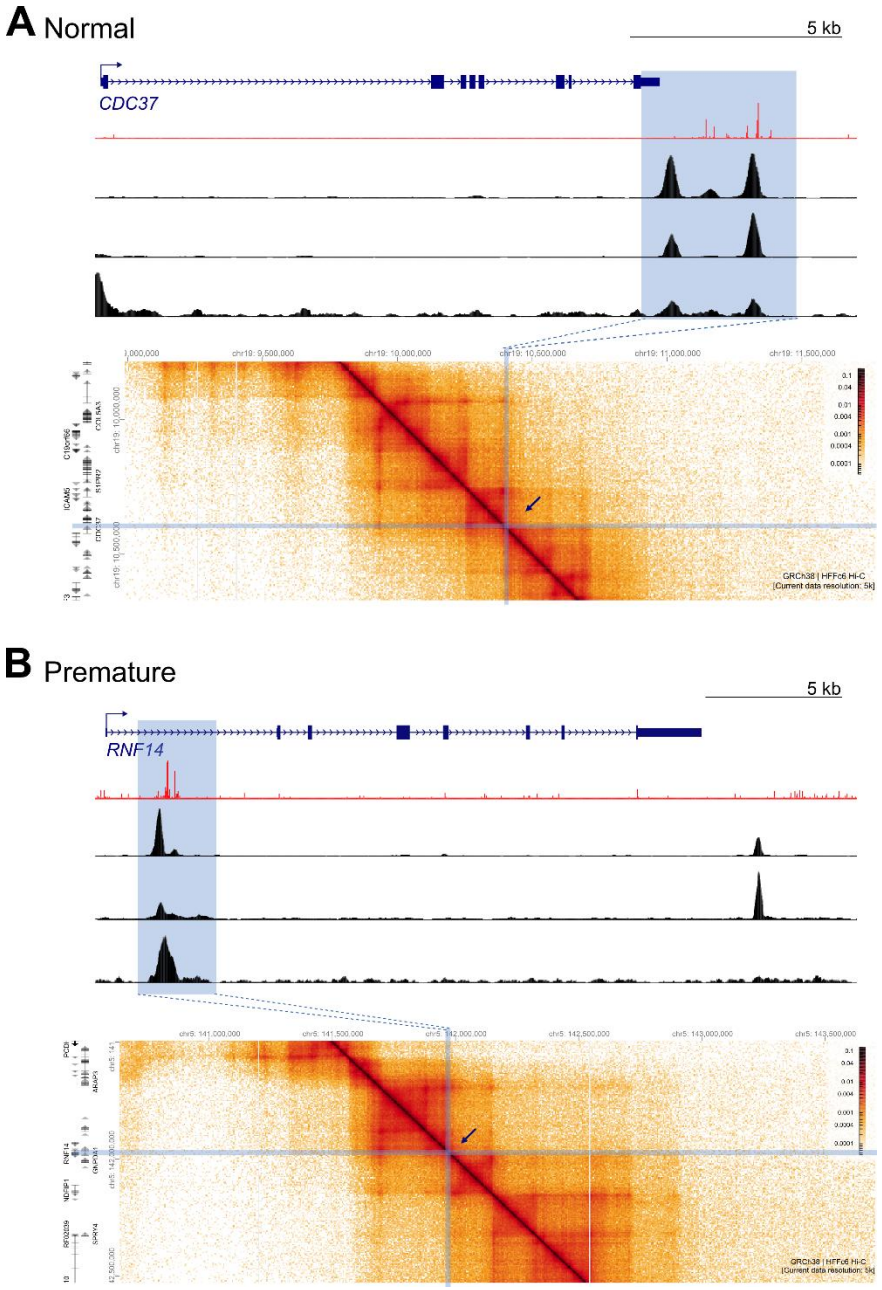


Figure 8. Colocalization of termination windows enriched in CTCF, SMC3 and MAZ (top panel) with Micro-C domain borders (bottom panel) for A) normal terminator of *CDC37* gene and B) premature terminator of *RNF14* gene.

In addition, I inspected representative genes with high looping factors enrichments and compared these findings with publicly available Micro-C data obtained via the 4D Nucleome Data Portal. Micro-C is a high-resolution chromatin conformation capture method that provides detailed information about chromatin interactions and chromosome folding. At a conceptual level, this method represents chromosomes as contact maps, where areas with a higher probability of interaction – indicating features such as loops or TADs – appear darker. Loop anchors can sometimes be seen as individual dots. Analysis of Micro-C dataset revealed multiple instances where both normal and premature termination windows coincide with Micro-C domain boundaries (Figure 8). However, this dataset is derived from a different cell line – HFF – so the evidence is not definitive. Furthermore, 23% of termination windows overlapped with computationally predicted

loop anchors (Figure 9)[132]. These data suggest that a proportion of termination windows may be located at chromatin loop boundaries. Nonetheless, since the described approach relies on correlation, further investigation is needed to clarify the relationship between loop formation and transcription termination events.

H3K36me3 significantly drops in termination windows

Another factor that stands out from the analysis is H3K36me3. Unlike the looping factors, this histone mark is not enriched within termination windows; instead, the H3K36me3 ChIP-seq signal significantly decreases at both normal and premature terminators (Figure 7D). This reduction mirrors the decrease in nascent RNA signal reported in Publication 1, which is expected because H3K36me3 is deposited co-transcriptionally. Nevertheless, given prior report suggesting a role for this histone mark in transcription termination, we chose to investigate its involvement in this process more thoroughly.

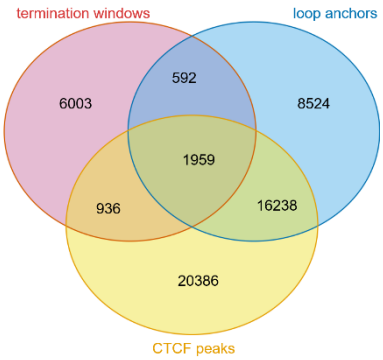


Figure 9. Overlap of termination windows, loop anchors and CTCF peaks (obtained through ENCODE).

Materials and methods

For the reference genome, I applied hg38/GRCh38, and gene annotations were taken from GENCODE release 40 (basic set). From this, I selected 10,387 protein-coding genes that fulfilled two criteria: (i) they did not overlap with any other annotated gene on the same strand, and (ii) their 3' ends were separated from the next downstream gene on the same strand by at least 6 kb.

To obtain termination windows genome-wide, I analyzed five biological replicates of T4ph mNET-seq derived from two independent studies: three replicates published by Kamieniarz-Gdula et al.[15] and two biological replicates from Schlackow et al[128]. For each dataset, I first generated an average T4ph mNET-seq signal track using wiggletools mean[133]. The resulting single-nucleotide resolution files were then processed in R: sequencing reads were resized to 150 bp, and bigWig signal files were converted into bedGraph format.

To detect regions enriched for T4ph signal, I used the bdgbroadcall function of macs2[134] with default parameters. Only regions that overlapped between the averaged files from at least two datasets were retained, after which overlapping intervals were merged (Figure 10). To reduce noise, I further restricted the dataset to windows that were positioned ≥2 kb away from their neighboring downstream and upstream windows.

Finally, termination windows were classified by overlap with the preselected protein-coding gene set. Two classes of windows were defined:

1. Normal termination windows – overlapping the annotated polyadenylation site (PAS) plus a downstream extension of 6 kb, n = 3,981.

2. Premature termination windows – located entirely within gene bodies of the selected protein-coding genes, n = 1,502.

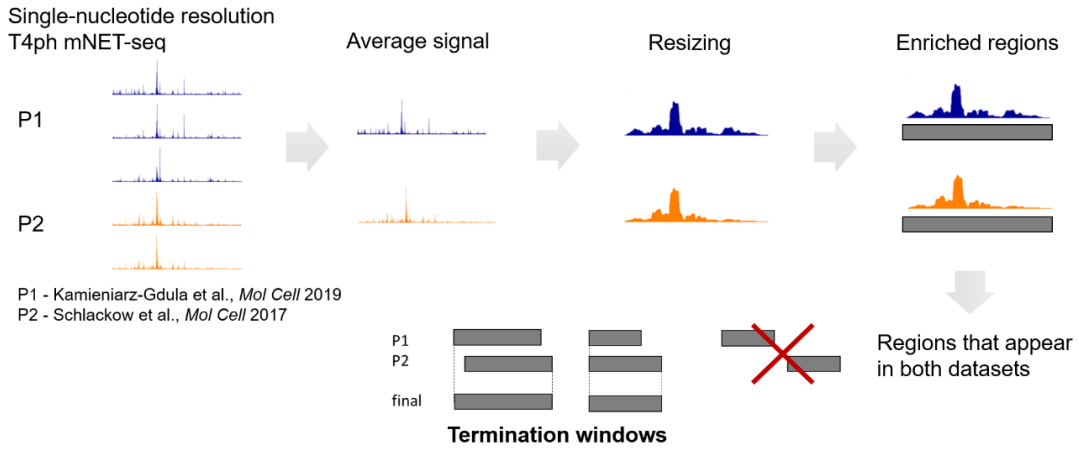


Figure 10. Termination windows definition. First, T4ph mNET-seq signal coming from distinct biological replicates was averaged for each dataset. Then, it was resized to 150bp and macs2 tool was used to extract T4ph enriched regions. Only the regions that appeared in both datasets were selected and subsequently merged.

To check the enrichment of chromatin factors in termination windows, I first defined 2 kb long flanking control regions. Then, using the getCountsByRegions() function from the BRGenomics library, I calculated the average signal within three regions (control upstream, termination window, and control downstream). This average was defined as the sum of the scores within each region divided by the number of ranges in that region. The calculated signals were then visualized in a boxplot, distinguishing between normal and premature terminators.

Termination windows and flanking control regions were used as input to create heatmaps with the deeptools[135] plotHeatmap tool, following computeMatrix. Parameters were set as --beforeRegionStartLength, --regionBodyLength, and --afterRegionStartLength to 2000 bp, and the --skipZeros option was applied.

SETD2 methyltransferase activity aids gene definition by promoting correct transcription initiation and termination (published)

Chihiro Nakayama*, **Magda Kopczyńska***, Agata Stępień, Shoko Ito, Koshi Imami, Michal R. Gdula, Takayuki Nojima, Kinga Kamieniarz-Gdula

bioRxiv 2025.07.14.664659, doi: https://doi.org/10.1101/2025.07.14.664659

***Co-first authors**

This publication is included as an appendix.

As described in the chapter "SETD2 methyltransferase: role in transcription", histone methyltransferase SETD2 has been implicated in transcription termination[23]. However, the study relied on indirect measures of mature RNA and did not investigate potential mechanisms underlying SETD2-mediated termination. Furthermore, correlation analysis described in the chapter "H3K36me3 histone mark significantly drops in termination regions" showed that the product of SETD2 activity, H3K36me3, significantly decreases within transcription termination regions, mirroring the drop in nascent RNA signal reported in Publication 1. Additionally, readthrough transcription has been identified as a prognostic marker in ccRCC, with higher levels correlating with poorer patient survival[23].

In this publication, we employ nascent transcriptomics on SETD2 knockout (KO) and RCC patient-derived cells. We demonstrate that SETD2 loss or inactivity does not affect all genes uniformly. Following SETD2 loss or inactivation, most genes – designated as class I – terminate transcription at the expected regions but fail to initiate transcription at the levels observed in wild-type conditions. Conversely, 15-25% of genes, classified as class II, exhibit readthrough transcription, while their initiation remains unchanged in the absence of SETD2 activity. Our data indicate that termination defects due to SETD2 loss or mutation likely result from increased cryptic initiation combined with impaired 3' pre-mRNA cleavage. We further show that readthrough is independent of alternative polyadenylation, highlighting the importance of context-dependent distinction between termination and 3' end processing. Surprisingly, the effect of SETD2 loss on termination appears indirect: treatment with the specific inhibitor EPZ-719 induces readthrough only after 9 days, despite earlier loss of H3K36me3. In summary, our findings demonstrate that SETD2 regulates transcription by enabling proper initiation, preventing cryptic initiation within gene bodies, and promoting efficient 3' pre-mRNA cleavage.

I contributed to this work by analyzing all generated genomic datasets (POINT-seq, POINT-5, 3' mRNA-seq, and T4ph mNET-seq). I performed downstream analyses such as class I and class II genes separation, analysis of cryptic initiation, APA, and development of a bioinformatic method for calculating cleavage index. Additionally, I carried out 3' mRNA-seq and T4ph mNET-seq experiments on U2OS WT/SETD2 KO cells. I visualized the data, prepared all manuscript figures, and took part in writing and editing the manuscript.

CLP1-dependent premature transcription termination opposes neurodegeneration (published)

Michał R. Gdula, **Magda Kopczyńska**, Upasana Saha, Kinga Kamieniarz-Gdula

Neuron. 2022 Apr 20;110(8):1277-1280

doi: https://doi.org/10.1016/j.neuron.2022.03.012

This publication is included as an appendix.

In this preview, we summarize the work by LaForce et al.[136] placing it in the broader context of how premature transcription termination shapes transcriptional regulation and cell physiology. The authors investigate the molecular mechanism underlying pontocerebellar hypoplasia type 10 (PCH10), a neurodegenerative disorder characterized by cerebellar and brainstem atrophy, microcephaly, and motor neuron disease. PCH10 is caused by a missense mutation (R140H) in CLP1, a component of the pre-mRNA 3' end processing machinery.

The authors utilize induced pluripotent stem cells (iPSCs) derived from fibroblasts of a healthy parent and a child with PCH10, which are then differentiated into motor neurons. In cell culture, the PCH10 model exhibits reduced cell density, likely due to enhanced differentiation of motor neuron progenitor cells into post-mitotic neurons. At the molecular level, motor neurons show a strikingly high level of PAS usage within gene bodies, including introns and coding regions, indicating widespread premature transcription termination (PTT). The authors report a substantial reduction in mRNA 3' end isoform diversity in the PCH10 model, with a relative increase in longer transcript isoforms. They also examine a second model – CLP1 knockout (KO) in motor neurons – in which complete loss of CLP1 shifts usage toward proximal sites, the opposite pattern to PCH10. In the preview, we argue that the claim that CLP1 promotes distal PAS usage may be overstated; instead, the proximal shift observed in CLP1 KO could reflect an indirect consequence of CFIm68 downregulation, a factor known to drive major APA changes[119,137].

In summary, LaForce et al. conclude that impaired pre-mRNA 3' end processing contributes to PCH10 pathogenesis by reducing transcript isoform diversity. These findings underscore the pivotal role of 3' end formation in cell differentiation and encourage further studies investigating 3' pre-mRNA processing, particularly in neurodegenerative disorders.

My contribution to this work included preparing the figure and participating in the writing and editing of the manuscript.

Discussion

Because transcription is often viewed in one dimension, its epigenetic context and spatial organization remain insufficiently explored, particularly regarding the final phase – termination. Chromatin architecture is highly dynamic, and the nucleus is a crowded environment where numerous protein complexes interact simultaneously. This complexity adds another layer of challenge in studying transcription, a process that requires hundreds of proteins to function in a highly organized manner.

On the other hand, it has been now demonstrated that transcription occurs pervasively throughout the genome, which may initially appear random and wasteful but is thought to play a regulatory role in maintaining chromatin accessibility and shaping 3D chromosome structure[43]. Since termination enables RNAPII turnover, severe disruption of this process could have catastrophic effects on nuclear and cellular physiology. Termination is essential for proper cell function, and multiple fail-safe termination pathways exist, likely including mechanisms that have yet to be discovered[138].

This work aims to uncover how chromatin influences transcription termination, integrating multiple layers of regulation into the seemingly linear, one-dimensional process. First, it focuses on defining a universal marker for transcription termination regions: RNAPII CTD phosphorylation at threonine 4. Next, using this marker, genome-wide termination windows were identified, and various histone marks and chromatin-associated factors were screened for possible enrichment within these regions. Finally, the role of the chromatin modifier SETD2 methyltransferase in transcription termination was assessed, and a novel function of this factor in 3' end processing was uncovered.

Through a comprehensive literature review and analysis of all available T4ph genomic datasets presented in the publication "Defining gene ends: RNA polymerase II CTD threonine 4 phosphorylation marks transcription termination regions genome-wide", we propose that RNAPII CTD phosphorylation at threonine 4 marks genomic regions where terminal polymerase pausing occurs. This approach is particularly advantageous for investigating premature termination events within gene bodies, where only a fraction of RNAPII molecules terminates while others continue transcribing. Nascent RNA sequencing techniques, which capture complete RNA molecules, have certain limitations in this context because quantifying signal drop within gene bodies is challenging and less straightforward than assessing the presence of a specific marker. Admittedly, researchers from the Cramer lab have developed mathematical models to estimate RNAPII drop-off probability using TT-seq experiments[139], demonstrating that nascent RNA sequencing can be successfully applied to study these events.

Notably, premature transcription termination has been increasingly recognized as an important regulatory mechanism in gene expression. As briefly outlined in the preview "CLP1-dependent premature transcription termination opposes neurodegeneration", PTT may contribute to maintaining transcript diversity, thereby preventing the onset of neurological disorders such as pontocerebellar hypoplasia type 10. Conversely, widespread

PTT observed in patients with chronic lymphocytic leukemia gives rise to truncated proteins, which frequently lack the tumor-suppressive functions exhibited by their full-length counterparts[140].

However, not all PTT events result in stable, polyadenylated transcripts; in many cases, they produce RNA molecules that are rapidly degraded[39]. This scenario is particularly evident near TSS, where it has been shown that ~99% of initiation events terminate prematurely at the gene promoter[141]. Such promoter-proximal termination is mediated by the Integrator complex and is considered a crucial mechanism that prevents RNAPII release into the gene body, thereby attenuating gene expression[142,143]. In line with these findings, we also observed T4ph enrichment at TSS prior to the appearance of productive elongation signals detected by nascent RNA sequencing.

It is important to emphasize that we propose T4ph as a marker for termination regions without demonstrating its mechanistic involvement in the process. While there is a clear correlation between the presence of this phospho-mark and decreasing nascent RNA signal, there is no direct evidence that T4ph is required for termination. Nevertheless, because T4ph specifically occurs in termination regions and the CTD threonine 4 residue is essential for metazoan cell viability[144,145], we speculate that it may play a functional role in the final step of transcription. Similar to other phospho-marks, T4ph could serve as a binding platform for termination factors or block elongation factor binding. However, the causal relationship between T4ph and transcription termination has yet to be elucidated.

Mapping termination regions across the genome offers an opportunity to enhance current genomic annotations. Pinpointing where transcription ends may help enable more detailed studies of mutation hotspots that contribute to disease. There is growing evidence that mutations in non-coding regions – which comprise 98% of nuclear DNA – play a substantial role in cancer development[146–148]. DNA sequence, particularly T-rich regions (T-tracks), has been shown to trigger transcription termination on the coding strand in human cells[149]. Thus, mutations in termination regions may directly impact the termination process. Changes in DNA sequence can also alter RNA structure, which is essential for termination in prokaryotes and for processing histone genes, recently increasingly recognized as important for efficient 3' processing of eukaryotic protein-coding genes[150]. Consequently, mutations in these regulatory regions may directly affect RNA processing itself.

Genome-wide definition of transcription termination regions enabled *in silico* data mining to identify chromatin features associated with terminating RNAPII. This analysis revealed that the defined termination windows are enriched for three key chromatin-looping factors: CTCF, SMC3, and MAZ. Evidence from multiple model organisms shows that actively transcribed genes form promoter–terminator loops, first in yeast[151–155], and later in higher eukaryotes[99,156–158]. Gene loops are thought to be dynamic structures formed by physical interactions between initiation and termination factors[158], and have been implicated in RNAPII recycling from terminators back to promoters[152] and in preserving

transcriptional memory[154]. Notably, these loops are specific to active transcription, and their formation is abrogated in the absence of transcriptional activators[153].

However, in recent years, with the development of elegant high-resolution techniques such as Micro-C, the model of functional gene looping as a driver of transcription regulation has been challenged, and a more complex picture has emerged. Recent analyses indicate that cohesin-mediated loop extrusion does not determine the organization of transcribed genes. Nevertheless, active genes form central subcompartments of the 3D genome driven by RNAPII binding and elongation, without requiring direct promoter–terminator loops[159]. Moreover, acute depletion of CTCF and cohesin only modestly affects gene expression and does not disrupt enhancer–promoter interactions, although cohesin may facilitate target searching and binding by transcription factors[160].

The enrichment of looping factors at termination windows may be purely correlational and does not imply causation. However, their enrichment is also observed at premature termination windows, which is unexpected and could indicate a potential role for loop formation in facilitating early termination. Just as transcriptional activation can be driven by diverse stimuli, termination may likewise be governed by distinct factors depending on genomic context. In some cases, loop formation might act as a roadblock to transcribing RNAPII, as reported for specific proteins in yeast[80]. Although a subset of termination windows coincides with predicted loop anchors, these observations are entirely correlational and require further experimental validation to determine whether chromatin loops can trigger termination at specific loci.

It is important to note that recent advancements, including 3D genome capture assays and sophisticated imaging approaches, have revealed that active transcription occurs within distinct nuclear compartments – historically termed transcription factories and more recently unified under the concept of phase-separated transcriptional condensates (PST-condensates)[161]. Structurally, these condensates are organized around a protein-rich core containing transcription factors and co-activators, with RNAPII molecules positioned at the periphery. Such assemblies are considered highly dynamic, yet capable of maintaining a metastable structure[162].

However, comprehensive studies evaluating the functional relevance of these higher-order assemblies, and their impact on multiple stages of transcription, are still lacking. The integration of methodologies from molecular biology, 3D genomics, and chromatin imaging represents a particularly promising direction to fully elucidate the dynamics of transcription termination within its nuclear environment.

In this thesis, one histone modification was investigated in detail. *In silico* screening revealed that H3K36me3 levels drop significantly within both normal and premature termination windows, parallelling the decrease in nascent RNA signal. This observation was somewhat expected, as H3K36me3 is deposited co-transcriptionally. However, it also points to a potential functional role of this histone mark in promoting RNAPII dissociation from the DNA template. Furthermore, H3K36me3 has been functionally linked to APA in

mESCs[163], and mutations in its depositing methyltransferase, SETD2, have been implicated in readthrough transcription in ccRCC[23].

In the publication "SETD2 methyltransferase activity aids gene definition by promoting correct transcription initiation and termination", we demonstrate the dichotomy in SETD2 activity depending on gene context. Our initial aim was to investigate the role of this methyltransferase in transcription termination; however, we found that most protein-coding genes (class I) rely on SETD2 activity primarily for proper transcription initiation, while the regions where termination occurs remain largely unaffected. In contrast, only 15–25% of genes (class II) exhibit transcriptional readthrough upon SETD2 knockout or catalytic inactivation, while their initiation levels remain unchanged. We further show that disrupted termination is associated with increased cryptic initiation, particularly near gene ends. Moreover, our analysis of the 3' cleavage index revealed a significant reduction only in class II genes, both under SETD2 knockout and catalytic inactivation conditions. We therefore propose a model in which SETD2 activity is essential at both gene ends: at the 5' ends of class I genes, it promotes accurate transcription initiation, whereas at the 3' ends of class II genes, it enhances 3' cleavage efficiency and prevents spurious initiation, thereby facilitating proper termination.

It remains unclear why different genes respond differently to SETD2 loss or inactivation. We found that class II genes are significantly more transcriptionally active, longer, and contain longer introns. This is in contrast to previous evidence showing a negative correlation between gene length and expression level[164,165]. However, it is important to emphasize that our study employed cancer cell lines, which inherently display disrupted expression patterns across multiple physiological pathways. Because transcription proceeds more rapidly across introns[166], longer genes are transcribed at a higher elongation rate. As a result, they are also more vulnerable to perturbations affecting elongation factors[167,168]. In mESCs, it has been demonstrated that H3K36me3 binds the Npac protein to facilitate efficient elongation of *Nanog* and *Rif1*[169]. Therefore, SETD2-dependent H3K36me3 deposition may enable or stabilize the recruitment of 3' processing and termination factors in long, fast-transcribed genes, whereas in the absence of this modification, RNA processing may become inefficient. However, as our experiments provide no direct proof for this hypothesis, additional investigation will be necessary to validate it.

Our study reveals, for the first time, a role for SETD2 in promoting 3' end cleavage efficiency. Previous studies based on reconstitution of the CPA machinery have shown that 3' cleavage is relatively inefficient *in vitro* (30–40%)[170,171], highlighting the importance of the nuclear and epigenetic context in promoting its effectiveness. The mechanism underlying SETD2 involvement in 3' end processing remains unknown; however, it may facilitate the recruitment of processing factors, as has been shown for heterogeneous ribonucleoproteins (hnRNPs) in the regulation of alternative splicing events[172,173]. In that case, the SHI domain of SETD2 is crucial for hnRNP binding, and its deletion leads to reduced H3K36me3 deposition[172].

In our work, we assessed the 3' cleavage index in two models: U2OS cells lacking SETD2 entirely, and A498 ccRCC cells harboring a mutation (c.6098_6099) that drastically reduces its catalytic activity. Since this mutation lies in the interdomain region between the catalytic SET/Post-SET domain and the regulatory SHI/SRI domains[174], it may affect both: the SHI domain, which is essential for hnRNP binding, and the SRI domain, which interacts with the RNAPII CTD[86]. Deleting the SRI and SHI domains causes a global reduction in H3K36me3 deposition, and a SETD2 mutant lacking both domains is nearly inactive[172]. Based on our findings, we propose that the catalytic activity of SETD2 is important for enhancing 3' cleavage, although contributions from the SHI and SRI domains to 3' processing cannot be excluded. To fully resolve this question, further structural and biochemical studies will be required.

Notably, we also demonstrate an uncoupling between APA events and transcriptional readthrough upon loss of SETD2 activity. 3' pre-mRNA cleavage and termination are normally tightly linked, as rapid depletion or inhibition of the cleavage factor CPSF73 endonuclease leads to widespread termination defects[24,175]. However, the chemical inhibition of the CPSF73 by JTE-607 primarily resulted in 3' UTR lengthening[175] and revealed that this endonuclease acts in a sequence-specific manner[176]. Our findings show that, although transcriptional readthrough occurs in both experimental models – whether through complete SETD2 loss or catalytic inactivation – APA is predominantly observed in cells lacking SETD2. Unexpectedly, SETD2-deficient cells preferentially utilize proximal PASs, a phenomenon evident in both class I and class II genes. In the RCC model, however, APA is not prominent, suggesting that PAS choice depends more on the cellular context than on transcription termination defects or SETD2 activity itself. Collectively, this evidence underscores a role for SETD2 in stimulating efficient termination and 3' cleavage, while indicating that these processes may occur independently of PAS selection. This perspective is relatively new, as 3' pre-mRNA processing and termination have traditionally been viewed as tightly coupled, also 3' end cleavage efficiency and APA are often treated synonymous.

Surprisingly, we found that the effects of SETD2 on transcription termination are indirect. Using the specific SETD2 inhibitor EPZ-719[177], we showed that although H3K36me3 levels were markedly reduced after 3 days of treatment, termination remained largely unaffected at this time point. Transcriptional readthrough appeared only after 9 days of inhibition, but even then, it remained weaker than that observed upon SETD2 knockout. This delayed effect may arise from cumulative changes in chromatin architecture and dynamics during prolonged inhibition, as the cells undergo multiple divisions.

Mass spectrometry analysis of nuclear fractions from SETD2-knockout cells revealed a general trend toward protein downregulation, including factors involved in all stages of transcription as well as RNAPII itself. Thus, prolonged absence of SETD2-mediated H3K36me3 deposition may gradually alter the expression of genes encoding RNA processing factors, eventually reaching a threshold at which transcriptional defects become unavoidable. These changes could be mediated not only by chromatin remodeling but also

by accumulation of DNA methylation at specific gene promoters, as we observed upregulation of DNMT3A and DNMT3B. To date, the effects of SETD2 depletion (knockout or knockdown) have primarily been attributed directly to SETD2 loss or H3K36me3 reduction. Our findings underscore the importance of distinguishing between direct and indirect effects – not only in the case of SETD2, but more broadly for chromatin modifiers and transcription-related factors, by using inhibitors and degron approaches.

Since *SETD2* is frequently mutated in several cancers, including ccRCC, multiple types of leukemia, melanoma, gliomas and colon adenocarcinoma[178,179] it is regarded as a tumor-suppressor gene. Moreover, immunohistochemical analyses of ccRCC samples shows a gradual decrease in H3K36me3 levels, progressing from primary tumors to distant metastases[180]. Therefore, assessing the molecular consequences of SETD2 loss or inactivation is crucial for understanding cancer biology, even if these alterations are not the primary drivers of tumor progression. This is particularly relevant as SETD2 also methylates additional substrates, such as α-tubulin[181], actin[182] and STAT1[183]. Readthrough transcription, on the one hand, promotes the formation of potentially harmful chimeric transcripts[23]. On the other hand, the induction of widespread termination defects may act as a mechanism to suppress cell proliferation and induce DNA damage, as has been demonstrated for the CPSF73 inhibitor JTE-607[175].

In summary, this thesis shows that transcription termination plays a crucial role in maintaining proper gene expression, and its function extends beyond merely displacing RNAPII from the DNA template. Gene ends are increasingly recognized as essential regulatory regions that dictate mRNA stability, localization, and ultimately the function of the encoded protein. Understanding the mechanisms and chromatin contexts of transcription termination is therefore important not only to satisfy scientific curiosity but also to address critical gaps in our knowledge of the transcriptional cycle – an essential process for the survival of every cell, and hence for human health.

This work provides insights into the chromatin context and epigenetic regulation of transcription termination, demonstrating correlations between termination windows and chromatin loop borders and exploring SETD2-dependent termination events. By defining termination windows genome-wide, it highlights the importance of investigating – often overlooked – regulatory regions, particularly at mutation hotspots. It also points to additional mechanisms that may govern termination, such as SETD2-driven enhancement of 3'-end cleavage efficiency and a potential role of chromatin loop formation in influencing termination.

References

1. Osman, S. & Cramer, P. Structural Biology of RNA Polymerase II Transcription: 20 Years On. *Annu Rev Cell Dev Biol* **36**, 1–34 (2020).
2. Buratowski, S. The CTD code. *Nat Struct Biol* **10**, 679–680 (2003).
3. Heidemann, M., Hintermair, C., Voß, K. & Eick, D. Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochim Biophys Acta* **1829**, 55–62 (2013).
4. Zaborowska, J., Egloff, S. & Murphy, S. The pol II CTD: new twists in the tail. *Nat Struct Mol Biol* **23**, 771–777 (2016).
5. Fabrega, C., Shen, V., Shuman, S. & Lima, C. D. Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Mol Cell* **11**, 1549–1561 (2003).
6. Fong, N. & Bentley, D. L. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. *Genes Dev* **15**, 1783–1795 (2001).
7. Petrenko, N., Jin, Y., Dong, L., Wong, K. H. & Struhl, K. Requirements for RNA polymerase II preinitiation complex formation in vivo. *Elife* **8**, e43654 (2019).
8. Fong, N., Sheridan, R. M., Ramachandran, S. & Bentley, D. L. The pausing zone and control of RNA polymerase II elongation by Spt5: Implications for the pause-release model. *Molecular Cell* **82**, 3632-3645.e4 (2022).
9. Core, L. & Adelman, K. Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes Dev* **33**, 960–982 (2019).
10. Wagner, E. J., Tong, L. & Adelman, K. Integrator is a global promoter-proximal termination complex. *Mol Cell* **83**, 416–427 (2023).
11. Takagaki, Y. & Manley, J. L. RNA recognition by the human polyadenylation factor CstF. *Mol Cell Biol* **17**, 3907–3914 (1997).
12. Meinhart, A. & Cramer, P. Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. *Nature* **430**, 223–226 (2004).
13. Kumar, A., Clerici, M., Muckenfuss, L. M., Passmore, L. A. & Jinek, M. Mechanistic insights into mRNA 3'-end processing. *Curr Opin Struct Biol* **59**, 143–150 (2019).
14. Nojima, T. *et al.* Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. *Cell* **161**, 526–540 (2015).
15. Kamieniarz-Gdula, K. *et al.* Selective Roles of Vertebrate PCF11 in Premature and Full-Length Transcript Termination. *Molecular Cell* **74**, 158-172.e9 (2019).
16. Proudfoot, N. J. Transcriptional interference and termination between duplicated alpha-globin gene constructs suggests a novel mechanism for gene regulation. *Nature* **322**, 562–565 (1986).
17. Greger, I. H., Demarchi, F., Giacca, M. & Proudfoot, N. J. Transcriptional interference perturbs the binding of Sp1 to the HIV-1 promoter. *Nucleic Acids Res* **26**, 1294–1301 (1998).
18. Eszterhas, S. K., Bouhassira, E. E., Martin, D. I. K. & Fiering, S. Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position. *Mol Cell Biol* **22**, 469–479 (2002).
19. Vilborg, A., Passarelli, M. C., Yario, T. A., Tycowski, K. T. & Steitz, J. A. Widespread Inducible Transcription Downstream of Human Genes. *Mol Cell* **59**, 449–461 (2015).
20. Vilborg, A. *et al.* Comparative analysis reveals genomic features of stress-induced transcriptional readthrough. *Proc Natl Acad Sci U S A* **114**, E8362–E8371 (2017).

21. Rutkowski, A. J. *et al.* Widespread disruption of host transcription termination in HSV-1 infection. *Nat Commun* **6**, 7126 (2015).
22. Hennig, T. *et al.* HSV-1-induced disruption of transcription termination resembles a cellular stress response but selectively increases chromatin accessibility downstream of genes. *PLoS Pathog* **14**, e1006954 (2018).
23. Grosso, A. R. *et al.* Pervasive transcription read-through promotes aberrant expression of oncogenes and RNA chimeras in renal carcinoma. *eLife* **4**, e09214 (2015).
24. Eaton, J. D., Francis, L., Davidson, L. & West, S. A unified allosteric/torpedo mechanism for transcriptional termination on human protein-coding genes. *Genes Dev* **34**, 132–145 (2020).
25. Rodríguez-Molina, J. B., West, S. & Passmore, L. A. Knowing when to stop: Transcription termination on protein-coding genes by eukaryotic RNAPII. *Mol Cell* **83**, 404–415 (2023).
26. Ahn, S. H., Kim, M. & Buratowski, S. Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Mol Cell* **13**, 67–76 (2004).
27. Kim, M., Ahn, S.-H., Krogan, N. J., Greenblatt, J. F. & Buratowski, S. Transitions in RNA polymerase II elongation complexes at the 3' ends of genes. *EMBO J* **23**, 354–364 (2004).
28. Connelly, S. & Manley, J. L. A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes Dev* **2**, 440–452 (1988).
29. Proudfoot, N. J. Ending the message: poly(A) signals then and now. *Genes Dev* **25**, 1770–1782 (2011).
30. Eaton, J. D. *et al.* Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CPSF73 activity. *Genes Dev.* **32**, 127–139 (2018).
31. West, S., Gromak, N. & Proudfoot, N. J. Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature* **432**, 522–525 (2004).
32. Kuś, K. *et al.* DSIF factor Spt5 coordinates transcription, maturation and exoribonucleolysis of RNA polymerase II transcripts. *Nat Commun* **16**, 10 (2025).
33. Skaar, J. R. *et al.* The Integrator complex controls the termination of transcription at diverse classes of gene targets. *Cell Res* **25**, 288–305 (2015).
34. Elrod, N. D. *et al.* The Integrator Complex Attenuates Promoter-Proximal Transcription at Protein-Coding Genes. *Mol Cell* **76**, 738-752.e7 (2019).
35. Cugusi, S. *et al.* Heat shock induces premature transcript termination and reconfigures the human transcriptome. *Mol Cell* **82**, 1573-1588.e10 (2022).
36. Kaida, D. *et al.* U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* **468**, 664–668 (2010).
37. Sousa-Luís, R. *et al.* POINT technology illuminates the processing of polymerase-associated intact nascent transcripts. *Molecular Cell* **81**, 1935-1950.e6 (2021).
38. Estell, C., Davidson, L., Steketee, P. C., Monier, A. & West, S. ZC3H4 restricts non-coding transcription in human cells. *Elife* **10**, e67305 (2021).
39. Kamieniarz-Gdula, K. & Proudfoot, N. J. Transcriptional Control by Premature Termination: A Forgotten Mechanism. *Trends Genet* **35**, 553–564 (2019).
40. Core, L. J., Waterfall, J. J. & Lis, J. T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845–1848 (2008).
41. Kwak, H., Fuda, N. J., Core, L. J. & Lis, J. T. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* **339**, 950–953 (2013).

42. Schwalb, B. *et al.* TT-seq maps the human transient transcriptome. *Science* **352**, 1225–1228 (2016).
43. Jensen, T. H., Jacquier, A. & Libri, D. Dealing with pervasive transcription. *Mol Cell* **52**, 473–484 (2013).
44. Jia, J. *et al.* An atlas of plant full-length RNA reveals tissue-specific and monocots-dicots conserved regulation of poly(A) tail length. *Nat Plants* **8**, 1118–1126 (2022).
45. Jin, X. *et al.* Pre-mRNA processing factors differentially impact coordination between co-transcriptional cleavage and transcription termination. *Nat Commun* **16**, 7086 (2025).
46. Mo, W. *et al.* Landscape of transcription termination in Arabidopsis revealed by single-molecule nascent RNA sequencing. *Genome Biol* **22**, 322 (2021).
47. Szakállas, N. *et al.* Can long-read sequencing tackle the barriers, which the next-generation could not? A review. *Pathol Oncol Res* **30**, 1611676 (2024).
48. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251–260 (1997).
49. Li, B., Carey, M. & Workman, J. L. The Role of Chromatin during Transcription. *Cell* **128**, 707–719 (2007).
50. Henikoff, S. & Shilatifard, A. Histone modification: cause or cog? *Trends Genet* **27**, 389–396 (2011).
51. Zhang, H., Gao, L., Anandhakumar, J. & Gross, D. S. Uncoupling Transcription from Covalent Histone Modification. *PLOS Genetics* **10**, e1004202 (2014).
52. Zhang, T., Cooper, S. & Brockdorff, N. The interplay of histone modifications – writers that read. *EMBO reports* **16**, 1467–1481 (2015).
53. Deckert, J. & Struhl, K. Histone acetylation at promoters is differentially affected by specific activators and repressors. *Mol Cell Biol* **21**, 2726–2735 (2001).
54. Barrera, L. O. *et al.* Genome-wide mapping and analysis of active promoters in mouse embryonic stem cells and adult organs. *Genome Res* **18**, 46–59 (2008).
55. Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A* **107**, 21931–21936 (2010).
56. Edmunds, J. W., Mahadevan, L. C. & Clayton, A. L. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *The EMBO Journal* **27**, 406–420 (2008).
57. Cao, K. *et al.* DOT1L-controlled cell-fate determination and transcription elongation are independent of H3K79 methylation. *Proceedings of the National Academy of Sciences* **117**, 27365–27373 (2020).
58. Batta, K., Zhang, Z., Yen, K., Goffman, D. B. & Pugh, B. F. Genome-wide function of H2B ubiquitylation in promoter and genic regions. *Genes Dev* **25**, 2254–2265 (2011).
59. Clapier, C. R., Iwasa, J., Cairns, B. R. & Peterson, C. L. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* **18**, 407–422 (2017).
60. Boeger, H., Griesenbeck, J., Strattan, J. S. & Kornberg, R. D. Removal of promoter nucleosomes by disassembly rather than sliding in vivo. *Mol Cell* **14**, 667–673 (2004).
61. Maassen, A. *et al.* SWI/SNF-type complexes–transcription factor interplay: a key regulatory interaction. *Cellular & Molecular Biology Letters* **30**, 30 (2025).
62. Morse, K., Bishop, A. L., Swerdlow, S., Leslie, J. M. & Ünal, E. Swi/Snf chromatin remodeling regulates transcriptional interference and gene repression. *Molecular Cell* **84**, 3080-3097.e9 (2024).

63. Weber, M. *et al.* Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* **39**, 457–466 (2007).
64. Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
65. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–257 (1999).
66. Yagi, M. *et al.* Identification of distinct loci for de novo DNA methylation by DNMT3A and DNMT3B during mammalian development. *Nat Commun* **11**, 3199 (2020).
67. Jones, P. A. & Liang, G. Rethinking how DNA Methylation Patterns are Maintained. *Nat Rev Genet* **10**, 805–811 (2009).
68. Schwabish, M. A. & Struhl, K. The Swi/Snf Complex Is Important for Histone Eviction during Transcriptional Activation and RNA Polymerase II Elongation In Vivo. *Mol Cell Biol* **27**, 6987–6995 (2007).
69. Hossain, I. *et al.* ZIC2 and ZIC3 promote SWI/SNF recruitment to safeguard progression towards human primed pluripotency. *Nat Commun* **15**, 8539 (2024).
70. Haberle, V. & Stark, A. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat Rev Mol Cell Biol* **19**, 621–637 (2018).
71. Lauberth, S. M. *et al.* H3K4me3 Interactions with TAF3 Regulate Preinitiation Complex Assembly and Selective Gene Activation. *Cell* **152**, 1021–1036 (2013).
72. Lin, J. J. *et al.* Mediator coordinates PIC assembly with recruitment of CHD1. *Genes Dev.* **25**, 2198–2209 (2011).
73. Hödl, M. & Basler, K. Transcription in the absence of histone H3.3. *Curr Biol* **19**, 1221–1226 (2009).
74. Hödl, M. & Basler, K. Transcription in the absence of histone H3.2 and H3K4 methylation. *Curr Biol* **22**, 2253–2257 (2012).
75. Pengelly, A. R., Copur, Ö., Jäckle, H., Herzig, A. & Müller, J. A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. *Science* **339**, 698–699 (2013).
76. Ehara, H., Kujirai, T., Shirouzu, M., Kurumizaka, H. & Sekine, S. Structural basis of nucleosome disassembly and reassembly by RNAPII elongation complex with FACT. *Science* **377**, eabp9466 (2022).
77. Farnung, L., Ochmann, M., Engeholm, M. & Cramer, P. Structural basis of nucleosome transcription mediated by Chd1 and FACT. *Nat Struct Mol Biol* **28**, 382–387 (2021).
78. Markert, J. W., Soffers, J. H. & Farnung, L. Structural basis of H3K36 trimethylation by SETD2 during chromatin transcription. *Science* **387**, 528–533 (2025).
79. Hildreth, A. E. *et al.* The nucleosome DNA entry-exit site is important for transcription termination and prevention of pervasive transcription. *eLife* **9**, e57757 (2020).
80. Candelli, T. *et al.* High-resolution transcription maps reveal the widespread impact of roadblock termination in yeast. *EMBO J* **37**, e97490 (2018).
81. Grosso, A. R., de Almeida, S. F., Braga, J. & Carmo-Fonseca, M. Dynamic transitions in RNA polymerase II density profiles during transcription termination. *Genome Res* **22**, 1447–1456 (2012).
82. Fuchs, S. M., Kizer, K. O., Braberg, H., Krogan, N. J. & Strahl, B. D. RNA polymerase II carboxyl-terminal domain phosphorylation regulates protein stability

of the Set2 methyltransferase and histone H3 di- and trimethylation at lysine 36. *J Biol Chem* **287**, 3249–3256 (2012).
83. Carrozza, M. J. *et al.* Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**, 581–592 (2005).
84. Neri, F. *et al.* Intragenic DNA methylation prevents spurious transcription initiation. *Nature* **543**, 72–77 (2017).
85. Bewick, A. J., Vogel, K. J., Moore, A. J. & Schmitz, R. J. Evolution of DNA Methylation across Insects. *Mol Biol Evol* **34**, 654–665 (2017).
86. Molenaar, T. M. & van Leeuwen, F. SETD2: from chromatin modifier to multipronged regulator of the genome and beyond. *Cell Mol Life Sci* **79**, 346 (2022).
87. Carvalho, S. *et al.* Histone methyltransferase SETD2 coordinates FACT recruitment with nucleosome dynamics during transcription. *Nucleic Acids Res* **41**, 2881–2893 (2013).
88. Simon, J. M. *et al.* Variation in chromatin accessibility in human kidney cancer links H3K36 methyltransferase loss with widespread RNA processing defects. *Genome Res* **24**, 241–250 (2014).
89. Guo, R. *et al.* BS69/ZMYND11 reads and connects histone H3.3 lysine 36 trimethylation-decorated chromatin to regulated pre-mRNA processing. *Mol Cell* **56**, 298–310 (2014).
90. de Almeida, S. F. *et al.* Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. *Nat Struct Mol Biol* **18**, 977–983 (2011).
91. Meers, M. P. *et al.* Histone gene replacement reveals a post-transcriptional role for H3K36 in maintaining metazoan transcriptome fidelity. *eLife* **6**, e23249.
92. Reyes, A. & Huber, W. Alternative start and termination sites of transcription drive most transcript isoform differences across human tissues. *Nucleic Acids Res* **46**, 582–592 (2018).
93. Tian, B., Hu, J., Zhang, H. & Lutz, C. S. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res* **33**, 201–212 (2005).
94. Tian, B. & Manley, J. L. Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol* **18**, 18–30 (2017).
95. Chan, S. L. *et al.* CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3' processing. *Genes Dev* **28**, 2370–2380 (2014).
96. Gruber, A. J. *et al.* A comprehensive analysis of 3' end sequencing data sets reveals novel polyadenylation signals and the repressive role of heterogeneous ribonucleoprotein C on cleavage and polyadenylation. *Genome Res* **26**, 1145–1159 (2016).
97. Mandel, C. R. *et al.* Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature* **444**, 953–956 (2006).
98. Gruber, A. J. & Zavolan, M. Alternative cleavage and polyadenylation in health and disease. *Nat Rev Genet* **20**, 599–614 (2019).
99. Sandberg, R., Neilson, J. R., Sarma, A., Sharp, P. A. & Burge, C. B. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* **320**, 1643–1647 (2008).
100. Mayr, C. & Bartel, D. P. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**, 673–684 (2009).
101. Garneau, N. L., Wilusz, J. & Wilusz, C. J. The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol* **8**, 113–126 (2007).

102. Graham, R. R. *et al.* Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. *Proc Natl Acad Sci U S A* **104**, 6758–6763 (2007).
103. Djebali, S. *et al.* Landscape of transcription in human cells. *Nature* **489**, 101–108 (2012).
104. Berkovits, B. D. & Mayr, C. Alternative 3'UTRs act as scaffolds to regulate membrane protein localization. *Nature* **522**, 363–367 (2015).
105. Yao, P. *et al.* Coding region polyadenylation generates a truncated tRNA synthetase that counters translation repression. *Cell* **149**, 88–100 (2012).
106. Alt, F. W. *et al.* Synthesis of secreted and membrane-bound immunoglobulin mu heavy chains is directed by mRNAs that differ at their 3' ends. *Cell* **20**, 293–301 (1980).
107. Mbita, Z. *et al.* De-regulation of the RBBP6 isoform 3/DWNN in human cancers. *Mol Cell Biochem* **362**, 249–262 (2012).
108. Di Giammartino, D. C. *et al.* RBBP6 isoforms regulate the human polyadenylation machinery and modulate expression of mRNAs with AU-rich 3' UTRs. *Genes Dev* **28**, 2248–2260 (2014).
109. Bennett, C. L. *et al.* A rare polyadenylation signal mutation of the FOXP3 gene (AAUAAA-->AAUGAA) leads to the IPEX syndrome. *Immunogenetics* **53**, 435–439 (2001).
110. Bacchetta, R., Barzaghi, F. & Roncarolo, M.-G. From IPEX syndrome to FOXP3 mutation: a lesson on immune dysregulation. *Ann N Y Acad Sci* **1417**, 5–22 (2018).
111. Miura, P., Shenker, S., Andreu-Agullo, C., Westholm, J. O. & Lai, E. C. Widespread and extensive lengthening of 3' UTRs in the mammalian brain. *Genome Res* **23**, 812–825 (2013).
112. Patel, R., Brophy, C., Hickling, M., Neve, J. & Furger, A. Alternative cleavage and polyadenylation of genes associated with protein turnover and mitochondrial function are deregulated in Parkinson's, Alzheimer's and ALS disease. *BMC Med Genomics* **12**, 60 (2019).
113. Kaur, I. *et al.* BIN1 in the Pursuit of Ousting the Alzheimer's Reign: Impact on Amyloid and Tau Neuropathology. *Neurotox Res* **41**, 698–707 (2023).
114. Chinta, S. J. *et al.* Environmental stress, ageing and glial cell senescence: a novel mechanistic link to Parkinson's disease? *J Intern Med* **273**, 429–436 (2013).
115. Rhinn, H. *et al.* Alternative α-synuclein transcript usage as a convergent mechanism in Parkinson's disease pathology. *Nat Commun* **3**, 1084 (2012).
116. Masamha, C. P. *et al.* CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. *Nature* **510**, 412–416 (2014).
117. Xiong, M. *et al.* NUDT21 inhibits bladder cancer progression through ANXA2 and LIMK2 by alternative polyadenylation. *Theranostics* **9**, 7156–7167 (2019).
118. Tan, S. *et al.* NUDT21 negatively regulates PSMB2 and CXXC5 by alternative polyadenylation and contributes to hepatocellular carcinoma suppression. *Oncogene* **37**, 4887–4900 (2018).
119. Ogorodnikov, A. *et al.* Transcriptome 3'end organization by PCF11 links alternative polyadenylation to formation and neuronal differentiation of neuroblastoma. *Nat Commun* **9**, 5331 (2018).
120. Nourse, J., Spada, S. & Danckwardt, S. Emerging Roles of RNA 3'-end Cleavage and Polyadenylation in Pathogenesis, Diagnosis and Therapy of Human Disorders. *Biomolecules* **10**, 915 (2020).

121. Rosa-Mercado, N. A. & Steitz, J. A. Who let the DoGs out?: Biogenesis of stress-induced readthrough transcripts. *Trends Biochem Sci* **47**, 206–217 (2022).
122. Proudfoot, N. J. Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science* **352**, aad9926 (2016).
123. Nemeroff, M. E., Barabino, S. M., Li, Y., Keller, W. & Krug, R. M. Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. *Mol Cell* **1**, 991–1000 (1998).
124. Wang, X. *et al.* Herpes simplex virus blocks host transcription termination via the bimodal activities of ICP27. *Nat Commun* **11**, 293 (2020).
125. Varley, K. E. *et al.* Recurrent read-through fusion transcripts in breast cancer. *Breast Cancer Res Treat* **146**, 287–297 (2014).
126. Kannan, K. *et al.* Recurrent chimeric RNAs enriched in human prostate cancer identified by deep sequencing. *Proc Natl Acad Sci U S A* **108**, 9172–9177 (2011).
127. Vo, J. N. *et al.* The Landscape of Circular RNA in Cancer. *Cell* **176**, 869-881.e13 (2019).
128. Schlackow, M. *et al.* Distinctive Patterns of Transcription and RNA Processing for Human lincRNAs. *Mol Cell* **65**, 25–38 (2017).
129. Ganji, M. *et al.* Real-time imaging of DNA loop extrusion by condensin. *Science* **360**, 102–105 (2018).
130. Sun, M., Nishino, T. & Marko, J. F. The SMC1-SMC3 cohesin heterodimer structures DNA through supercoiling-dependent loop formation. *Nucleic Acids Res* **41**, 6149–6160 (2013).
131. Xiao, T., Li, X. & Felsenfeld, G. The Myc-associated zinc finger protein (MAZ) works together with CTCF to control cohesin positioning and genome organization. *Proc Natl Acad Sci U S A* **118**, e2023127118 (2021).
132. Akgol Oksuz, B. *et al.* Systematic evaluation of chromosome conformation capture assays. *Nat Methods* **18**, 1046–1055 (2021).
133. Zerbino, D. R., Johnson, N., Juettemann, T., Wilder, S. P. & Flicek, P. WiggleTools: parallel processing of large collections of genome-wide datasets for visualization and statistical analysis. *Bioinformatics* **30**, 1008–1009 (2014).
134. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* **9**, R137 (2008).
135. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research* **44**, W160–W165 (2016).
136. LaForce, G. R. *et al.* Suppression of premature transcription termination leads to reduced mRNA isoform diversity and neurodegeneration. *Neuron* **110**, 1340-1357.e7 (2022).
137. Li, W. *et al.* Systematic profiling of poly(A)+ transcripts modulated by core 3' end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation. *PLoS Genet* **11**, e1005166 (2015).
138. Lemay, J.-F. & Bachand, F. Fail-safe transcription termination: Because one is never enough. *RNA Biol* **12**, 927–932 (2015).
139. Žumer, K. *et al.* FACT maintains chromatin architecture and thereby stimulates RNA polymerase II pausing during transcription *in vivo*. *Molecular Cell* **84**, 2053-2069.e9 (2024).
140. Lee, S.-H. *et al.* Widespread intronic polyadenylation inactivates tumour suppressor genes in leukaemia. *Nature* **561**, 127–131 (2018).

141. Steurer, B. *et al.* Live-cell analysis of endogenous GFP-RPB1 uncovers rapid turnover of initiating and promoter-paused RNA Polymerase II. *Proc Natl Acad Sci U S A* **115**, E4368–E4376 (2018).
142. Huang, K.-L. *et al.* Integrator Recruits Protein Phosphatase 2A to Prevent Pause Release and Facilitate Transcription Termination. *Mol Cell* **80**, 345-358.e9 (2020).
143. Stein, C. B. *et al.* Integrator endonuclease drives promoter-proximal termination at all RNA polymerase II transcribed loci. *Mol Cell* **82**, 4232-4245.e11 (2022).
144. Hsin, J.-P., Sheth, A. & Manley, J. L. RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3' end processing. *Science* **334**, 683–686 (2011).
145. Hintermair, C. *et al.* Threonine-4 of mammalian RNA polymerase II CTD is targeted by Polo-like kinase 3 and required for transcriptional elongation. *EMBO J* **31**, 2784–2797 (2012).
146. Mosquera Orgueira, A. *et al.* Novel Mutation Hotspots within Non-Coding Regulatory Regions of the Chronic Lymphocytic Leukemia Genome. *Sci Rep* **10**, 2407 (2020).
147. Weinhold, N., Jacobsen, A., Schultz, N., Sander, C. & Lee, W. Genome-wide analysis of noncoding regulatory mutations in cancer. *Nat Genet* **46**, 1160–1165 (2014).
148. Diederichs, S. *et al.* The dark matter of the cancer genome: aberrations in regulatory elements, untranslated regions, splice sites, non-coding RNA and synonymous mutations. *EMBO Molecular Medicine* **8**, 442–457 (2016).
149. Davidson, L. *et al.* DNA-directed termination of mammalian RNA polymerase II. *Genes Dev* **38**, 998–1019 (2024).
150. Xu, J., Duncan, S. & Ding, Y. The role of RNA structure in 3' end processing in eukaryotes. *Current Opinion in Structural Biology* **89**, 102933 (2024).
151. Lainé, J.-P., Singh, B. N., Krishnamurthy, S. & Hampsey, M. A physiological role for gene loops in yeast. *Genes Dev* **23**, 2604–2609 (2009).
152. Allepuz-Fuster, P. *et al.* RNA polymerase II plays an active role in the formation of gene loops through the Rpb4 subunit. *Nucleic Acids Res* **47**, 8975–8987 (2019).
153. Kaderi, B. E., Medler, S., Raghunayakula, S. & Ansari, A. Gene Looping Is Conferred by Activator-dependent Interaction of Transcription Initiation and Termination Machineries *. *Journal of Biological Chemistry* **284**, 25015–25025 (2009).
154. Tan-Wong, S. M., Wijayatilake, H. D. & Proudfoot, N. J. Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex. *Genes Dev* **23**, 2610–2624 (2009).
155. Moabbi, A. M., Agarwal, N., El Kaderi, B. & Ansari, A. Role for gene looping in intron-mediated enhancement of transcription. *Proc Natl Acad Sci U S A* **109**, 8505–8510 (2012).
156. Henriques, T. *et al.* Transcription termination between polo and snap, two closely spaced tandem genes of D. melanogaster. *Transcription* **3**, 198–212 (2012).
157. Sun, L. *et al.* Mapping nucleosome-resolution chromatin organization and enhancer-promoter loops in plants using Micro-C-XL. *Nat Commun* **15**, 35 (2024).
158. Goel, S., Krishnamurthy, S. & Hampsey, M. Mechanism of start site selection by RNA polymerase II: interplay between TFIIB and Ssl2/XPB helicase subunit of TFIIH. *J Biol Chem* **287**, 557–567 (2012).
159. Salari, H., Fourel, G. & Jost, D. Transcription regulates the spatio-temporal dynamics of genes through micro-compartmentalization. *Nat Commun* **15**, 5393 (2024).
160. Hsieh, T.-H. S. *et al.* Enhancer–promoter interactions and transcription are largely maintained upon acute loss of CTCF, cohesin, WAPL or YY1. *Nat Genet* **54**, 1919–1932 (2022).

161. Rippe, K. & Papantonis, A. RNA polymerase II transcription compartments — from factories to condensates. *Nat Rev Genet* 1–14 (2025) doi:10.1038/s41576-025-00859-6.
162. Cho, W.-K. *et al.* Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412–415 (2018).
163. Jia, J. *et al.* FUS reads histone H3K36me3 to regulate alternative polyadenylation. *Nucleic Acids Research* **52**, 5549–5571 (2024).
164. Urrutia, A. O. & Hurst, L. D. The signature of selection mediated by expression on human genes. *Genome Res* **13**, 2260–2264 (2003).
165. Castillo-Davis, C. I., Mekhedov, S. L., Hartl, D. L., Koonin, E. V. & Kondrashov, F. A. Selection for short introns in highly expressed genes. *Nat Genet* **31**, 415–418 (2002).
166. Mimoso, C. A. & Adelman, K. U1 snRNP increases RNA Pol II elongation rate to enable synthesis of long genes. *Mol Cell* **83**, 1264-1279.e10 (2023).
167. Krajewska, M. *et al.* CDK12 loss in cancer cells affects DNA damage response genes through premature cleavage and polyadenylation. *Nat Commun* **10**, 1757 (2019).
168. Oh, J.-M. *et al.* U1 snRNP telescripting regulates a size-function-stratified human genome. *Nat Struct Mol Biol* **24**, 993–999 (2017).
169. Yu, S. *et al.* Npac Is a Co-Factor of Histone H3K36me3 and Regulates Transcriptional Elongation in Mouse Embryonic Stem Cells. *genom. proteom. bioinform.* **20**, 110–128 (2022).
170. Boreikaite, V., Elliott, T. S., Chin, J. W. & Passmore, L. A. RBBP6 activates the pre-mRNA 3' end processing machinery in humans. *Genes Dev* **36**, 210–224 (2022).
171. Schmidt, M. *et al.* Reconstitution of 3' end processing of mammalian pre-mRNA reveals a central role of RBBP6. *Genes & Development* **36**, 195 (2022).
172. Bhattacharya, S. *et al.* The methyltransferase SETD2 couples transcription and splicing by engaging mRNA processing factors through its SHI domain. *Nat Commun* **12**, 1443 (2021).
173. Bhattacharya, S. *et al.* Structural basis of the interaction between SETD2 methyltransferase and hnRNP L paralogs for governing co-transcriptional splicing. *Nat Commun* **12**, 6452 (2021).
174. Bhattacharya, S. *et al.* The disordered regions of the methyltransferase SETD2 govern its function by regulating its proteolysis and phase separation. *J Biol Chem* **297**, 101075 (2021).
175. Cui, Y. *et al.* Elevated pre-mRNA 3' end processing activity in cancer cells renders vulnerability to inhibition of cleavage and polyadenylation. *Nature Communications* **14**, 4480 (2023).
176. Liu, L. *et al.* The anticancer compound JTE-607 reveals hidden sequence specificity of the mRNA 3' processing machinery. *Nat Struct Mol Biol* **30**, 1947–1957 (2023).
177. Lampe, J. W. *et al.* Discovery of a First-in-Class Inhibitor of the Histone Methyltransferase SETD2 Suitable for Preclinical Studies. *ACS Med Chem Lett* **12**, 1539–1545 (2021).
178. Skucha, A., Ebner, J. & Grebien, F. Roles of SETD2 in Leukemia-Transcription, DNA-Damage, and Beyond. *Int J Mol Sci* **20**, 1029 (2019).
179. Li, J. *et al.* SETD2: an epigenetic modifier with tumor suppressor functionality. *Oncotarget* **7**, 50719–50734 (2016).
180. Ho, T. H. *et al.* High-resolution profiling of histone h3 lysine 36 trimethylation in metastatic renal cell carcinoma. *Oncogene* **35**, 1565–1574 (2016).
181. Park, I. Y. *et al.* Dual Chromatin and Cytoskeletal Remodeling by SETD2. *Cell* **166**, 950–962 (2016).

182. Seervai, R. N. H. *et al.* The Huntingtin-interacting protein SETD2/HYPB is an actin lysine methyltransferase. *Sci Adv* **6**, eabb7854 (2020).
183. Chen, K. *et al.* Methyltransferase SETD2-Mediated Methylation of STAT1 Is Critical for Interferon Antiviral Activity. *Cell* **170**, 492-506.e14 (2017).

Funding

This work was supported by:

1. National Science Centre Sonata BIS 2018/30/E/NZ1/00073
2. European Molecular Biology Organization EMBO IG 4728-2020
3. European Union - Horizon Europe - European Research Council ERC 'AlternativeEnds' [101042642]

Authors contributions

Poznań, 4 sierpnia 2025

Magda Kopczyńska
Zakład Biologii Molekularnej i Komórkowej
Instytut Biologii Molekularnej i Biotechnologii
Wydział Biologii UAM
Centrum Zaawansowanych Technologii

Oświadczenie współautora artykułu

Oświadczam, że mój udział w przygotowanie artykułu "Defining gene ends: RNA polymerase II CTD threonine 4 phosphorylation marks transcription termination regions genome-wide" (autorzy: Magda Kopczyńska, Upasana Saha, Anastasiia Romanenko, Takayuki Nojima, Michał Gdula oraz Kinga Kamieniarz-Gdula) opublikowanego w czasopiśmie *Nucleic Acids Research,* Volume 53, Issue 2, 27 January 2025, gkae1240, doi: 10.1093/nar/gkae1240, który jest częścią mojej rozprawy doktorskiej polegał na wykonaniu wszystkich analiz bioinformatycznych, przygotowaniu figur oraz pisaniu i edycji manuskryptu.

.....
Podpis współautora

.....
Podpis promotora

Poznań, September 2, 2025

Upasana Saha, PhD
Department of Molecular and Cellular Biology
Institute of Molecular Biology and Biotechnology
Faculty of Biology, AMU
Center for Advanced Technologies

Co-author declaration

I hereby declare that my contribution to the publication titled "Defining gene ends: RNA polymerase II CTD threonine 4 phosphorylation marks transcription termination regions genome-wide" (authors: Magda Kopczyńska, Upasana Saha, Anastasiia Romanenko, Takayuki Nojima, Michał Gdula, and Kinga Kamieniarz-Gdula), published in Nucleic Acids Research, Volume 53, Issue 2, January 27, 2025, gkae1240, doi: 10.1093/nar/gkae1240, which forms part of Magda Kopczyńska's doctoral dissertation, consisted of writing and editing the manuscript.

Upasana Saha 4/9/2025
Co-author's signature

Kyiv, August 4, 2025

Anastasiia Romanenko
Junior biotechnologist at YURiA-PHARM LLC

Co-author declaration

I hereby declare that my contribution to the publication titled "Defining gene ends: RNA polymerase II CTD threonine 4 phosphorylation marks transcription termination regions genome-wide" (authors: Magda Kopczyńska, Upasana Saha, Anastasiia Romanenko, Takayuki Nojima, Michał Gdula, and Kinga Kamieniarz-Gdula), published in Nucleic Acids Research, Volume 53, Issue 2, January 27, 2025, gkae1240, doi: 10.1093/nar/gkae1240, which forms part of Magda Kopczyńska's doctoral dissertation, consisted of writing and editing the manuscript.

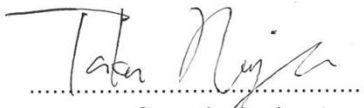
.....
Co-author's signature

Fukuoka, August 4, 2025

Prof. Takayuki Nojima
Medical Institute of Bioregulation
Kyushu University

Co-author declaration

I hereby declare that my contribution to the publication titled "Defining gene ends: RNA polymerase II CTD threonine 4 phosphorylation marks transcription termination regions genome-wide" (authors: Magda Kopczyńska, Upasana Saha, Anastasiia Romanenko, Takayuki Nojima, Michał Gdula, and Kinga Kamieniarz-Gdula), published in Nucleic Acids Research, Volume 53, Issue 2, January 27, 2025, gkae1240, doi: 10.1093/nar/gkae1240, which forms part of Magda Kopczyńska's doctoral dissertation, consisted of editing the manuscript.



Co-author's signature

Poznań, 4 sierpnia 2025

Dr hab. Michał Gdula
Zakład Ekspresji Genów
Instytut Biologii Molekularnej i Biotechnologii
Wydział Biologii UAM
Centrum Zaawansowanych Technologii

Oświadczenie współautora artykułu

Oświadczam, że mój udział w przygotowanie artykułu "Defining gene ends: RNA polymerase II CTD threonine 4 phosphorylation marks transcription termination regions genome-wide" (autorzy: Magda Kopczyńska, Upasana Saha, Anastasiia Romanenko, Takayuki Nojima, Michał Gdula oraz Kinga Kamieniarz-Gdula) opublikowanego w czasopiśmie *Nucleic Acids Research,* Volume 53, Issue 2, 27 January 2025, gkae1240, doi: 10.1093/nar/gkae1240, który jest częścią mojej rozprawy doktorskiej polegał na edycji manuskryptu.

.....
Podpis współautora

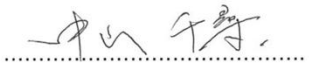
Fukuoka, August 4, 2025

Chihiro Nakayama
Medical Institute of Bioregulation
Kyushu University

Co-author declaration

I hereby declare that my contribution to the preparation of the article "SETD2 methyltransferase activity aids gene definition by promoting correct transcription initiation and termination" (authors: Chihiro Nakayama*, Magda Kopczyńska*, Agata Stępień, Shoko Ito, Koshi Imami, Michał Gdula, Takayuki Nojima#, and Kinga Kamieniarz-Gdula#), published as a preprint on the bioRxiv platform, doi: 10.1101/2025.07.14.664659, which forms part of Magda Kopczyńska's doctoral dissertation, consisted of performing POINT-seq and POINT-5 experiments, assessing the effect of EPZ-719 on cells, Western Blot and RT-qPCR experiments, and writing and editing the manuscript.

* These authors contributed equally
# Co-corresponding authors



Co-author's signature

Poznań, 4 sierpnia 2025

Magda Kopczyńska
Zakład Biologii Molekularnej i Komórkowej
Instytut Biologii Molekularnej i Biotechnologii
Wydział Biologii UAM
Centrum Zaawansowanych Technologii

Oświadczenie współautora artykułu

Oświadczam, że mój udział w przygotowanie artykułu "SETD2 methyltransferase activity aids gene definition by promoting correct transcription initiation and termination" (autorzy: Chihiro Nakayama*, Magda Kopczyńska*, Agata Stępień, Shoko Ito, Koshi Imami, Michał Gdula, Takayuki Nojima# oraz Kinga Kamieniarz-Gdula#) opublikowanego jako preprint na platformie bioRxiv, doi: 10.1101/2025.07.14.664659, który jest częścią mojej rozprawy doktorskiej polegał na wykonaniu wszystkich analiz bioinformatycznych, przeprowadzeniu eksperymentów mNET-seq i 3'mRNA-seq, przygotowaniu figur oraz pisaniu i edycji manuskryptu.

* Autorzy wnieśli równy wkład pracy
# Współautorzy korespondecyjni

.....
Podpis współautora

.....
Podpis promotora

Poznań, 4 sierpnia 2025

Dr inż. Agata Stępień
Zakład Biologii Molekularnej i Komórkowej
Instytut Biologii Molekularnej i Biotechnologii
Wydział Biologii UAM
Centrum Zaawansowanych Technologii

Oświadczenie współautora artykułu

Oświadczam, że mój udział w przygotowanie artykułu "SETD2 methyltransferase activity aids gene definition by promoting correct transcription initiation and termination" (autorzy: Chihiro Nakayama*, Magda Kopczyńska*, Agata Stępień, Shoko Ito, Koshi Imami, Michał Gdula, Takayuki Nojima# oraz Kinga Kamieniarz-Gdula#) opublikowanego jako preprint na platformie bioRxiv, doi: 10.1101/2025.07.14.664659, który jest częścią rozprawy doktorskiej Magdy Kopczyńskiej polegał na przeprowadzeniu eksperymentów 3'mRNA-seq oraz edycji manuskryptu.

* Autorzy wnieśli równy wkład pracy
# Współautorzy korespondecyjni

Stępień
.....
Podpis współautora

Fukuoka, August 4, 2025

Shoko Ito
Medical institute of Bioregulation
Kyushu University

Co-author declaration

I hereby declare that my contribution to the preparation of the article "SETD2 methyltransferase activity aids gene definition by promoting correct transcription initiation and termination" (authors: Chihiro Nakayama*, Magda Kopczyńska*, Agata Stępień, Shoko Ito, Koshi Imami, Michał Gdula, Takayuki Nojima#, and Kinga Kamieniarz-Gdula#), published as a preprint on the bioRxiv platform, doi: 10.1101/2025.07.14.664659, which forms part of Magda Kopczyńska's doctoral dissertation, consisted of cell culture maintenance, Western Blot experiments and RT-qPCR experiments.

* These authors contributed equally
# Co-corresponding authors

伊藤 祥子

Co-author's signature

Yokohama, August 4, 2025

Prof. Koshi Imami
Proteome Homeostasis Research Unit
RIKEN Center for Integrative Medical Sciences

Co-author declaration

I hereby declare that my contribution to the preparation of the article "SETD2 methyltransferase activity aids gene definition by promoting correct transcription initiation and termination" (authors: Chihiro Nakayama*, Magda Kopczyńska*, Agata Stępień, Shoko Ito, Koshi Imami, Michał Gdula, Takayuki Nojima#, and Kinga Kamieniarz-Gdula#), published as a preprint on the bioRxiv platform, doi: 10.1101/2025.07.14.664659, which forms part of Magda Kopczyńska's doctoral dissertation, consisted of conducting the mass spectrometry experiment and raw signal annotation.

* These authors contributed equally
# Co-corresponding authors

.....

Co-author's signature

Poznań, 4 sierpnia 2025

Dr hab. Michał Gdula
Zakład Ekspresji Genów
Instytut Biologii Molekularnej i Biotechnologii
Wydział Biologii UAM
Centrum Zaawansowanych Technologii

Oświadczenie współautora artykułu

Oświadczam, że mój udział w przygotowanie artykułu "SETD2 methyltransferase activity aids gene definition by promoting correct transcription initiation and termination" (autorzy: Chihiro Nakayama*, Magda Kopczyńska*, Agata Stępień, Shoko Ito, Koshi Imami, Michał Gdula, Takayuki Nojima# oraz Kinga Kamieniarz-Gdula#) opublikowanego jako preprint na platformie bioRxiv, doi: 10.1101/2025.07.14.664659, który jest częścią rozprawy doktorskiej Magdy Kopczyńskiej polegał na ustawieniu i zoptymalizowaniu analizy danych 3'mRNA-seq i alternatywnej poliadenylacji oraz POINT-seq.

* Autorzy wnieśli równy wkład pracy
# Współautorzy korespondecyjni

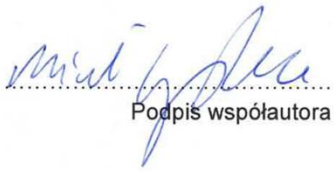
Podpis współautora

Poznań, 4 sierpnia 2025

Dr hab. Michał Gdula
Zakład Ekspresji Genów
Instytut Biologii Molekularnej i Biotechnologii
Wydział Biologii UAM
Centrum Zaawansowanych Technologii

Oświadczenie współautora artykułu

Oświadczam, że mój udział w przygotowanie artykułu "CLP1-dependent premature transcription termination opposes neurodegeneration" (autorzy: Michał Gdula, Magda Kopczyńska, Upasana Saha and Kinga Kamieniarz-Gdula), opublikowanego w czasopiśmie Neuron, 2022 Apr 20;110(8):1277-1280, doi: 10.1016/j.neuron.2022.03.012, który jest częścią rozprawy doktorskiej Magdy Kopczyńskiej polegał na pisaniu i edycji manuskryptu.



Podpis współautora

Poznań, 4 sierpnia 2025

Magda Kopczyńska
Zakład Biologii Molekularnej i Komórkowej
Instytut Biologii Molekularnej i Biotechnologii
Wydział Biologii UAM
Centrum Zaawansowanych Technologii

Oświadczenie współautora artykułu

Oświadczam, że mój udział w przygotowanie artykułu "CLP1-dependent premature transcription termination opposes neurodegeneration" (autorzy: Michał Gdula, Magda Kopczyńska, Upasana Saha and Kinga Kamieniarz-Gdula), opublikowanego w czasopiśmie Neuron, 2022 Apr 20;110(8):1277-1280, doi: 10.1016/j.neuron.2022.03.012, który jest częścią mojej rozprawy doktorskiej polegał na pisaniu i edycji manuskryptu.

.....
Podpis współautora

.....
Podpis promotora

Poznań, September 4, 2025

Upasana Saha, PhD
Department of Molecular and Cellular Biology
Institute of Molecular Biology and Biotechnology
Faculty of Biology, AMU
Center for Advanced Technologies

Co-author declaration

I hereby declare that my contribution to the publication titled "CLP1-dependent premature transcription termination opposes neurodegeneration" (authors: Michał Gdula, Magda Kopczyńska, Upasana Saha and Kinga Kamieniarz-Gdula), published in Neuron. 2022 Apr 20;110(8):1277-1280, doi: 10.1016/j.neuron.2022.03.012, which forms part of Magda Kopczyńska's doctoral dissertation, consisted of writing and editing the manuscript.

Upasana Saha 4/9/2025
Co-author's signature

Poznań, 4 sierpnia 2025

Dr hab. Kinga Kamieniarz-Gdula, prof. UAM
Zakład Biologii Molekularnej i Komórkowej
Instytut Biologii Molekularnej i Biotechnologii
Wydział Biologii UAM
Centrum Zaawansowanych Technologii

Oświadczenie współautora artykułu

Oświadczam, że mój udział w przygotowanie artykułu "CLP1-dependent premature transcription termination opposes neurodegeneration" (autorzy: Michał Gdula, Magda Kopczyńska, Upasana Saha and Kinga Kamieniarz-Gdula), opublikowanego w czasopiśmie Neuron, 2022 Apr 20;110(8):1277-1280, doi: 10.1016/j.neuron.2022.03.012, który jest częścią rozprawy doktorskiej Magdy Kopczyńskiej polegał na pisaniu i edycji manuskryptu.

Podpis współautora

Appendix

Appendix includes three publications:

1. Magda Kopczyńska, Upasana Saha, Anastasiia Romanenko, Takayuki Nojima, Michał R Gdula, Kinga Kamieniarz-Gdula, Defining gene ends: RNA polymerase II CTD threonine 4 phosphorylation marks transcription termination regions genome-wide, Nucleic Acids Research, Volume 53, Issue 2, 27 January 2025, gkae1240, doi: https://doi.org/10.1093/nar/gkae1240
2. Chihiro Nakayama*, Magda Kopczyńska*, Agata Stępień, Shoko Ito, Koshi Imami, Michał R. Gdula, Takayuki Nojima, Kinga Kamieniarz-Gdula, SETD2 methyltransferase activity aids gene definition by promoting correct transcription initiation and termination, bioRxiv 2025.07.14.664659
   doi: https://doi.org/10.1101/2025.07.14.664659
   *Co-first authors
3. Michał R. Gdula, Magda Kopczyńska, Upasana Saha, Kinga Kamieniarz-Gdula, CLP1-dependent premature transcription termination opposes neurodegeneration, Neuron. 2022 Apr 20;110(8):1277-1280
   doi: https://doi.org/10.1016/j.neuron.2022.03.012