

Transcription termination

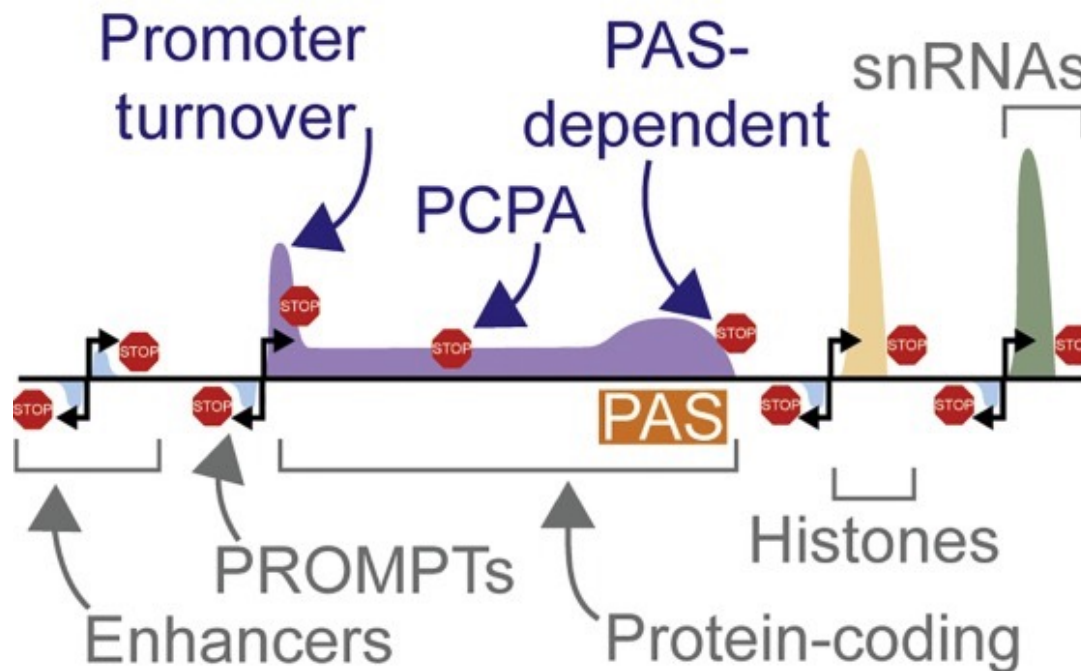
Transcription termination

- Transcription initiates pervasively in all organisms.
- Termination of transcription is essential for sorting out the functional RNAs from a plethora of transcriptional products that seemingly have no use in the cell.
- Terminating transcription is not that easy, given the high robustness of the elongation process.

Termination Occurs at All Points in the Transcription Cycle

Because Pol II transcribes multiple gene types, its termination occurs in a variety of ways

Where to terminate?



Trends in Genetics

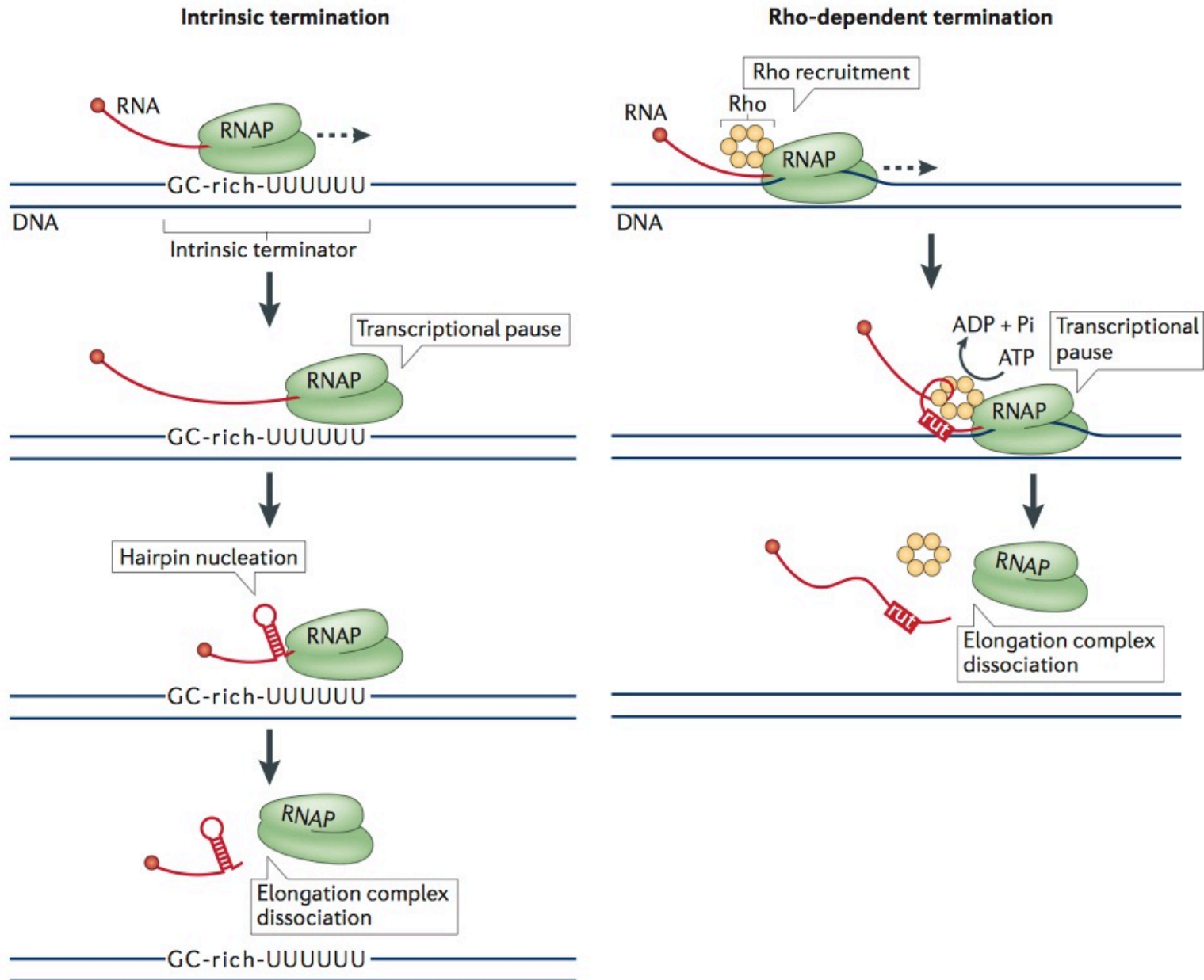
Abbreviations: PAS, polyadenylation signal; PCPA, premature cleavage and polyadenylation; PROMPT, promoter upstream transcript; snRNA, small nuclear RNA.

Transcripts and associated termination, processing and degradation pathways

Transcript	Termination pathway	Stability	Degradation factors
Yeast			
mRNA	CPF–CF and possibly Sen1	Stable	None
snRNA and snoRNA	NNS	Stable (3' end processed)	TRAMP, Rrp6, exosome, Rex1 (3' end processing)
CUT	NNS	Unstable	TRAMP, Rrp6, exosome
SUT	CPF–CF and possibly NNS	Partially unstable	Rrp6, exosome, Xrn1 (NMD)
XUT	CPF–CF	Unstable	Xrn1 (NMD)
RUT	Reb1 roadblock	Unstable	TRAMP, Rrp6, exosome
Metazoan			
mRNA	CPSF–CF and SETX	Stable	None
snRNA	Integrator complex, CBC–ARS2, PCF11 and NELF	Stable (3' end processed)	Exosome (3' end processing)
Non productive transcription	Integrator complex	Unstable	Exosome (3' end processing)
mRNAs encoding replication-dependent histones	CBC–ARS2	Stable	None
PROMPT	CPSF–CF and CBC–ARS2	Unstable	NEXT and exosome

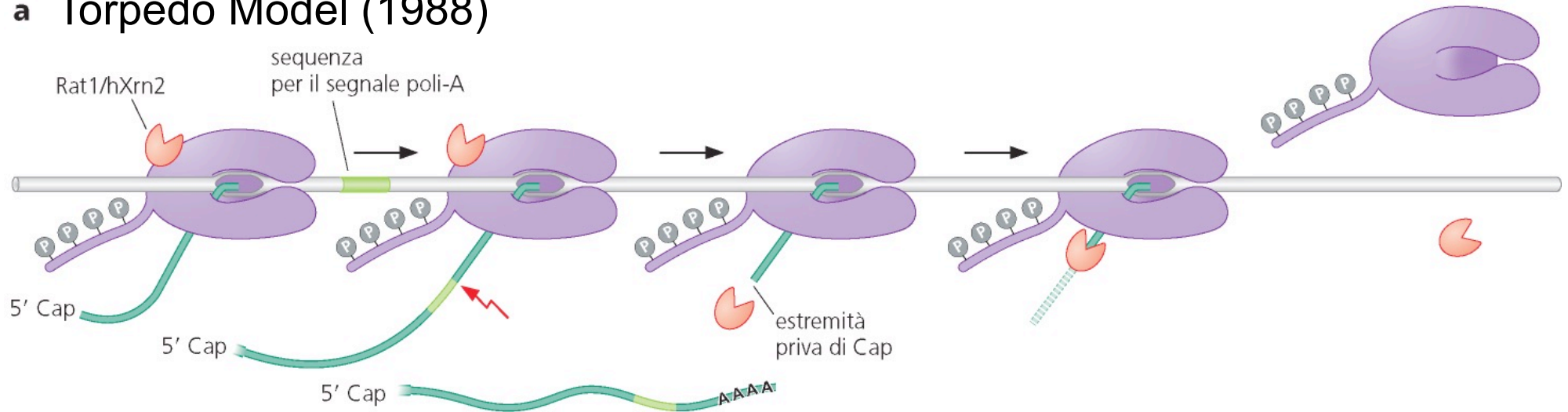
CBC, cap-binding complex; CF, cleavage factor; CPSF, cleavage and polyadenylation specificity factor; CUT, cryptic unstable transcript; NELF, negative elongation factor; NEXT, nuclear exosome targeting; NMD, nonsense-mediated decay; NNS, Nrd1–Nab3–Sen1; PROMPT, promoter-proximal transcript; RUT, Reb1-dependent unstable transcript; SETX, senataxin; snoRNA, small nucleolar RNA; snRNA, small nuclear RNA; SUT, stable unannotated transcript; TRAMP, Trf4–Air2–Mtr4; XUT, Xrn1-dependent unstable transcript.

Termination of transcription in bacteria

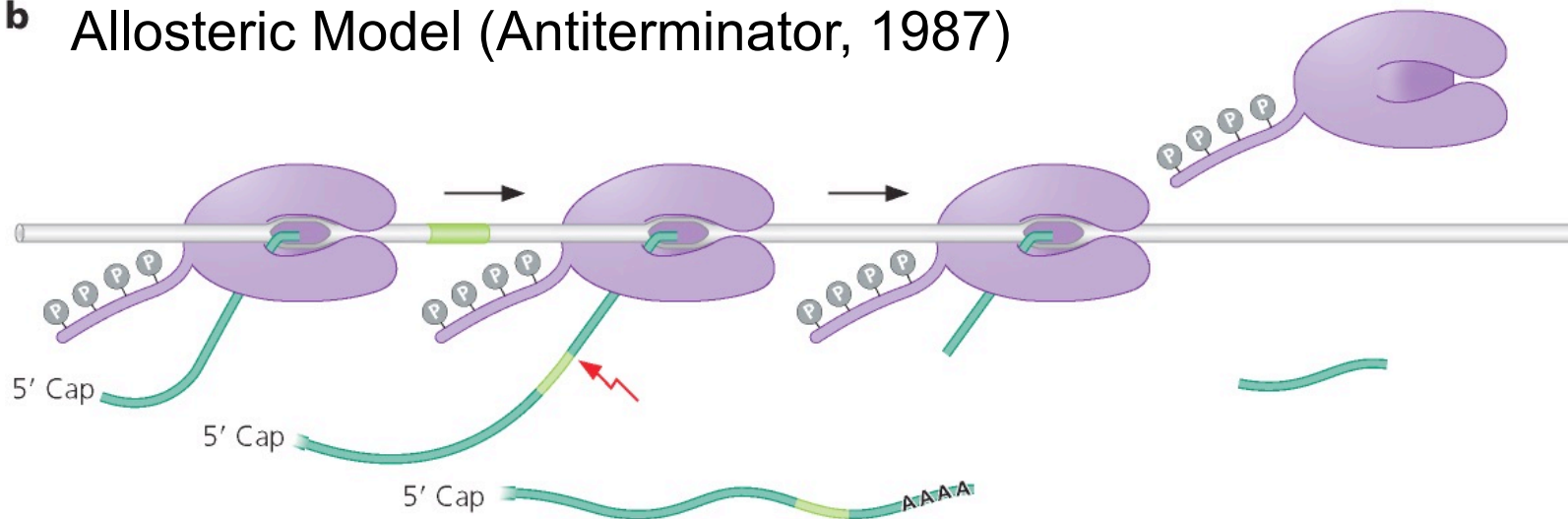


Models for PAS-Dependent Pol II Termination

a Torpedo Model (1988)

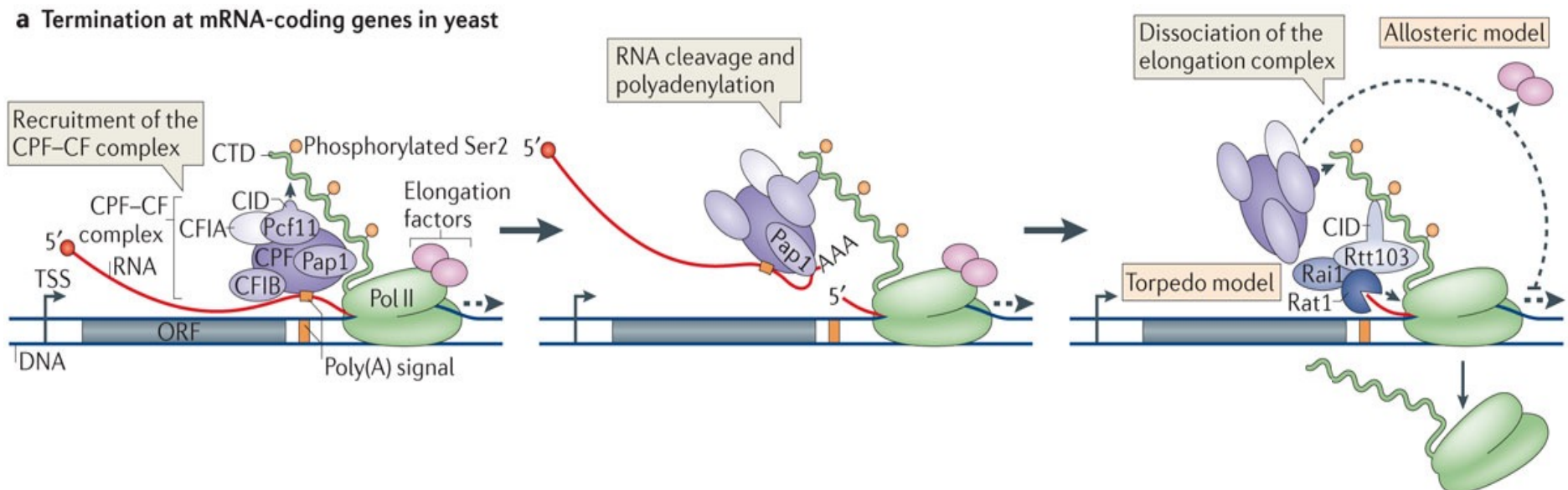


b Allosteric Model (Antiterminator, 1987)



Termination of mRNA-coding genes: the CPF–CF pathway

3'-end processing and termination of mRNAs are triggered by multipartite signals on the nascent RNA, which are recognized by components of the CPF complex that also directly interacts with the polymerase. It is generally accepted that cleavage of the nascent transcripts occurs before release of the enzyme from the DNA template, which occurs further downstream. The biochemical details of the termination reaction are still unclear.

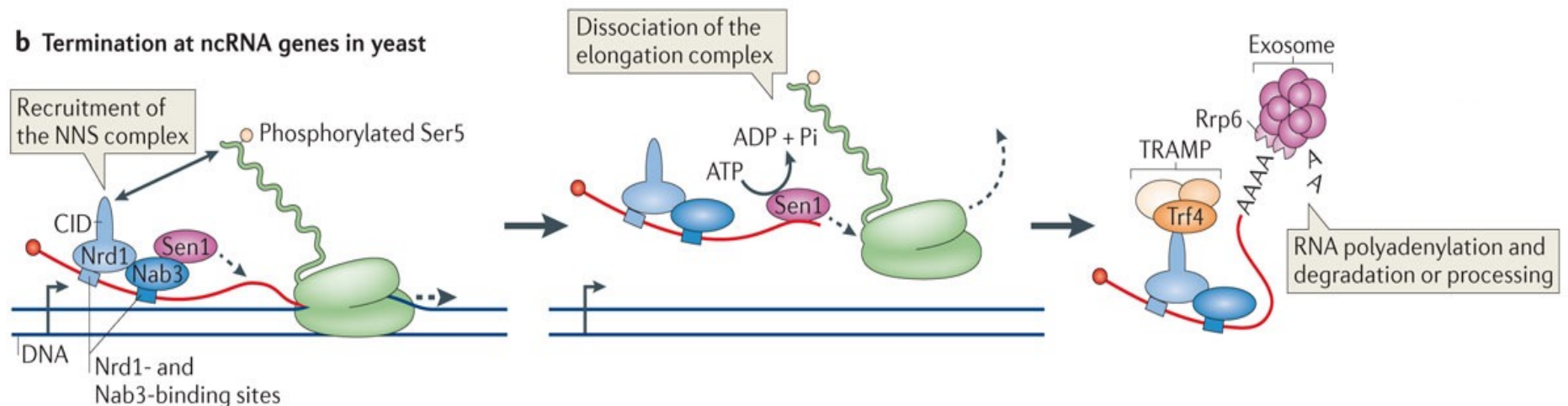


Termination of ncRNAs: the NNS-dependent pathway

In *S. cerevisiae*, the **Nrd1–Nab3–Sen1 (NNS) complex** is responsible for transcription termination at genes encoding **snRNAs** and **snoRNAs** and at **cryptic unstable transcripts (CUTs)**. The essential NNS complex contains two RNA-binding proteins, Nrd1 and Nab3, and the conserved superfamily I RNA and DNA helicase Sen1. Cleavage of the primary transcript has never been demonstrated for this termination pathway, and release of the polymerase occurs by a mechanism that strictly requires the action of the helicase Sen1 (similar to bacterial Rho-dependent termination).

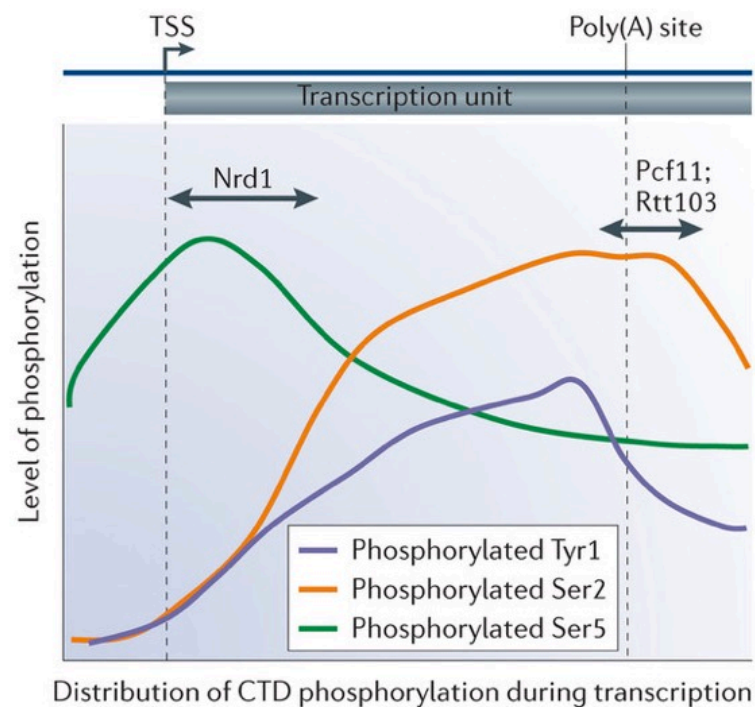
An important and distinctive trait of the transcripts produced by NNS-dependent termination is that they are targeted by the nuclear exosome rapidly after their release.

The presence of short sequence motifs on the nascent RNA that are recognized by Nrd1 and Nab3 (GUAA/G and UCUUG, respectively) has been shown to be a crucial specificity determinant of NNS-dependent termination. These motifs are often clustered and associated with AU-rich sequences, which contribute to the efficiency of termination.



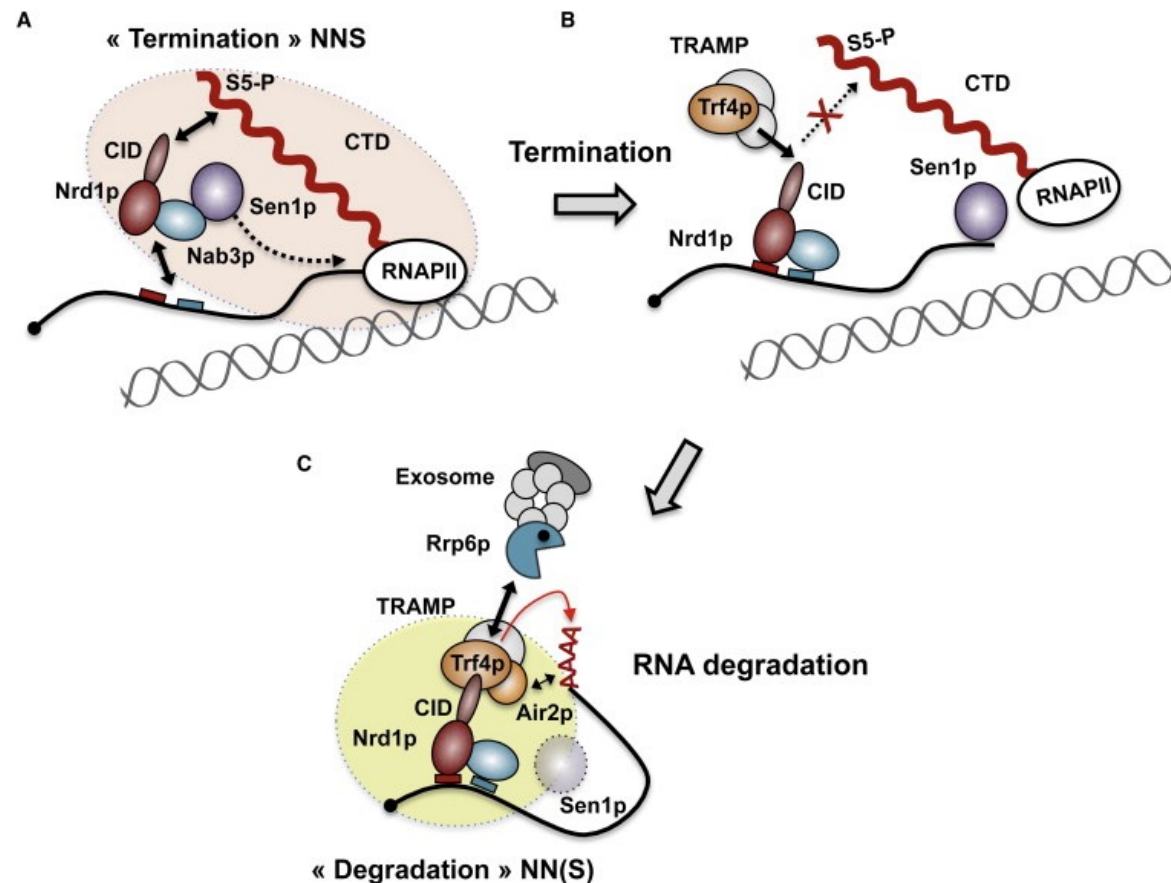
The CTD of RNA Pol II and termination

The most relevant modifications for transcription termination are phosphorylation at Tyr1, Ser2, Ser5 and Ser7, which alter the binding specificity of termination factors. For instance, in yeast the CPF complex component Pcf11 interact preferentially with the Ser2P form of the CTD only when Tyr1 is dephosphorylated. Conversely, Nrd1 recognizes the Ser5P form of the CTD, which predominates early in transcription, but only before the phosphorylation of Tyr1, which helps to restrict the recruitment of the NNS complex to the early stages of transcription.



Model for the Coordination of Transcription Termination with RNA Degradation at NNS

After interacting with the Ser5P CTD for the termination step, Nrd1 recruits TRAMP through the direct recognition of a CTD mimic — known as the **Nrd1-interacting motif (NIM)** — in the TRAMP component Trf4. The sequential (and mutually exclusive) interaction of Nrd1 with the CTD and Trf4 contributes to the temporal coordination of termination with degradation.

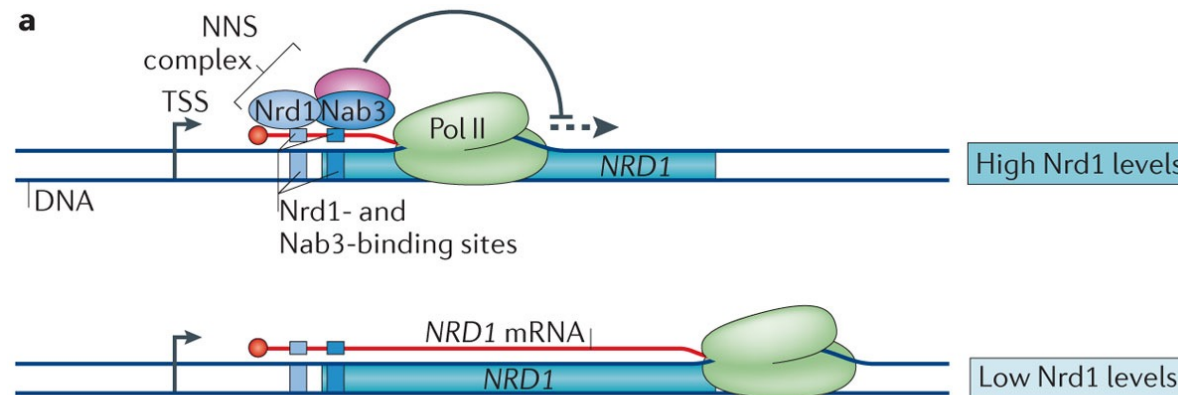


Regulation of gene expression by transcription termination

The occurrence of premature termination or termination that is associated with degradation of the transcript effectively prevents or limits gene expression. This can lead to bona fide regulation of gene expression, for example, when the occurrence of premature termination is modulated in response to an external stimulus or a physiological condition.

Nrd1:

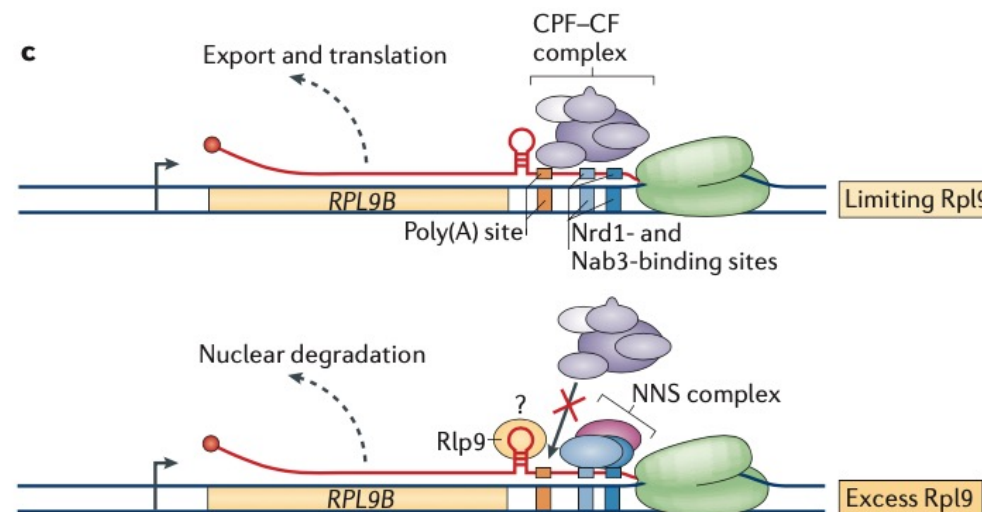
In yeast, the *NRD1* locus contains Nrd1- and Nab3-binding sites in the 5' region of the gene, and NNS-dependent termination occurs with a suboptimal efficiency that depends on the levels of the Nrd1 protein, thus establishing a negative feedback loop



Regulation of gene expression by transcription termination

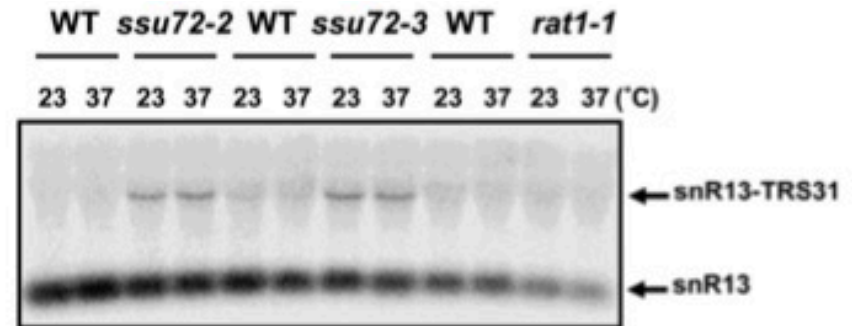
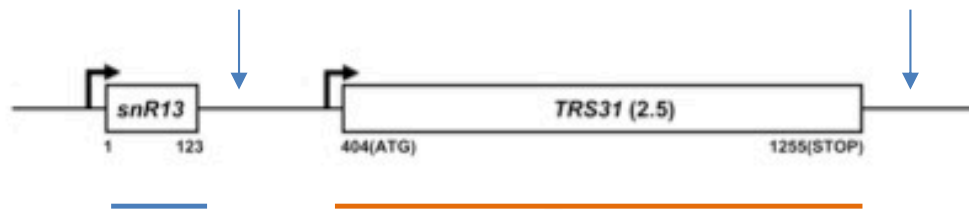
RPL9b

When the ribosomal protein Rpl9 is limiting, transcription termination of the *RPL9B* locus is driven by the cleavage and polyadenylation factor (CPF)–cleavage factor (CF) complex, which generates functional transcripts that are exported to the cytoplasm for translation. When Rpl9 is in excess, it is thought to bind to an RNA stem–loop in the vicinity of the poly(A) site, thus masking CPF–CF termination signals and preventing CPF–CF-dependent termination. This enables the interaction of the NNS complex with downstream sites, which induces transcription termination and the generation of transcripts that are rapidly degraded by the exosome.



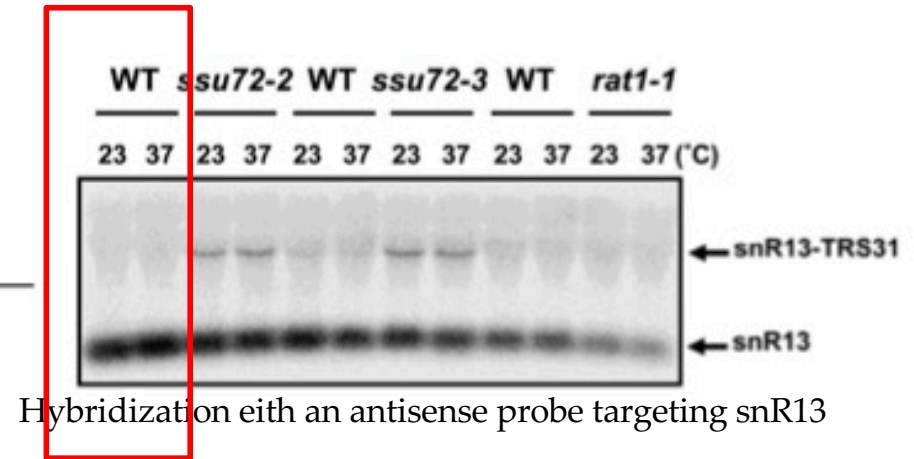
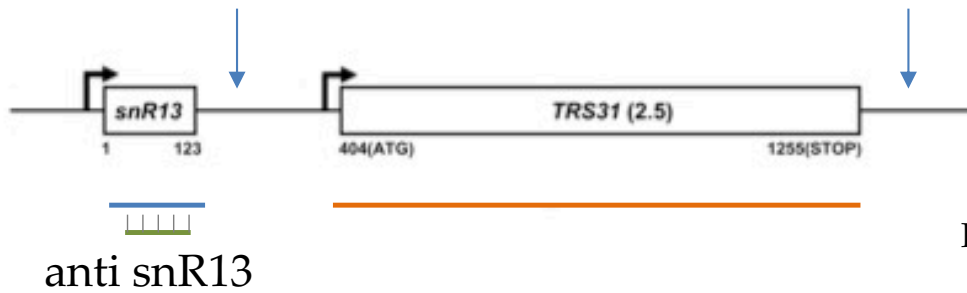
Experimental methodologies for transcription termination

Northern blot



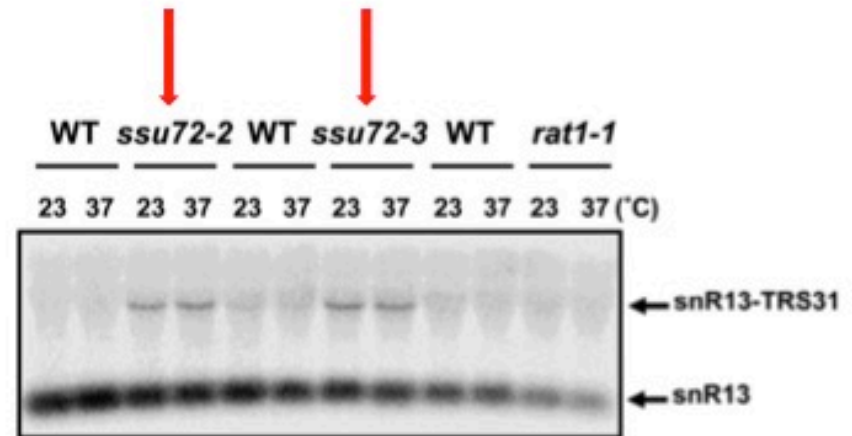
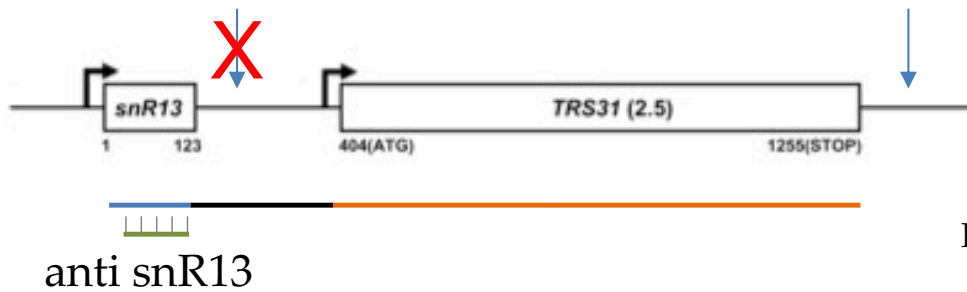
Experimental methodologies for transcription termination

Northern blot



Experimental methodologies for transcription termination

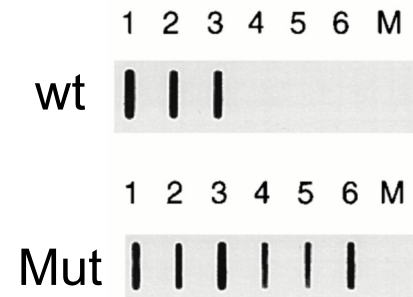
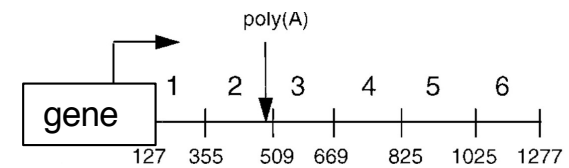
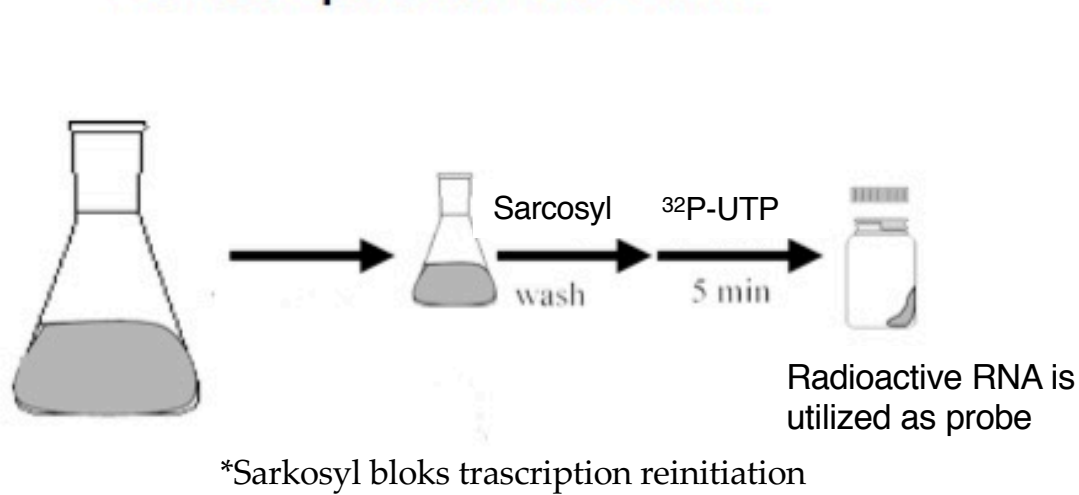
Northern blot



Hybridization with an antisense probe targeting nbR13

Experimental methodologies for transcription termination

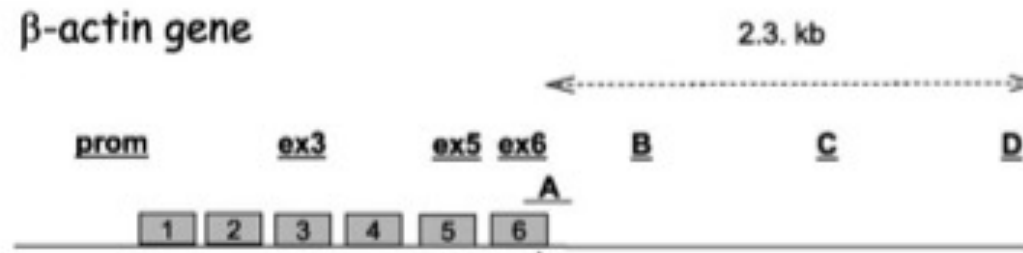
Transcriptional Run ON



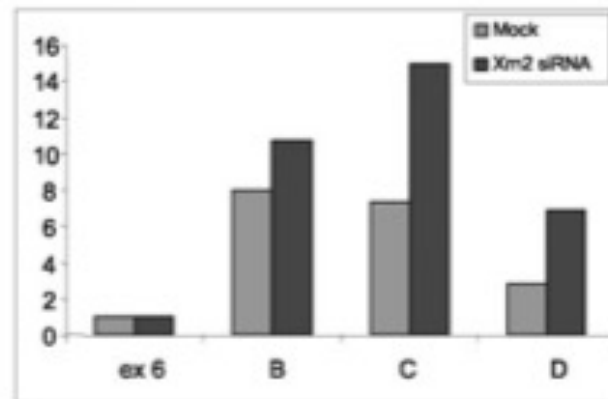
Antisense DNA oligos (1-6) are spotted on a membrane

Experimental methodologies for termination

RNA Pol II ChIP

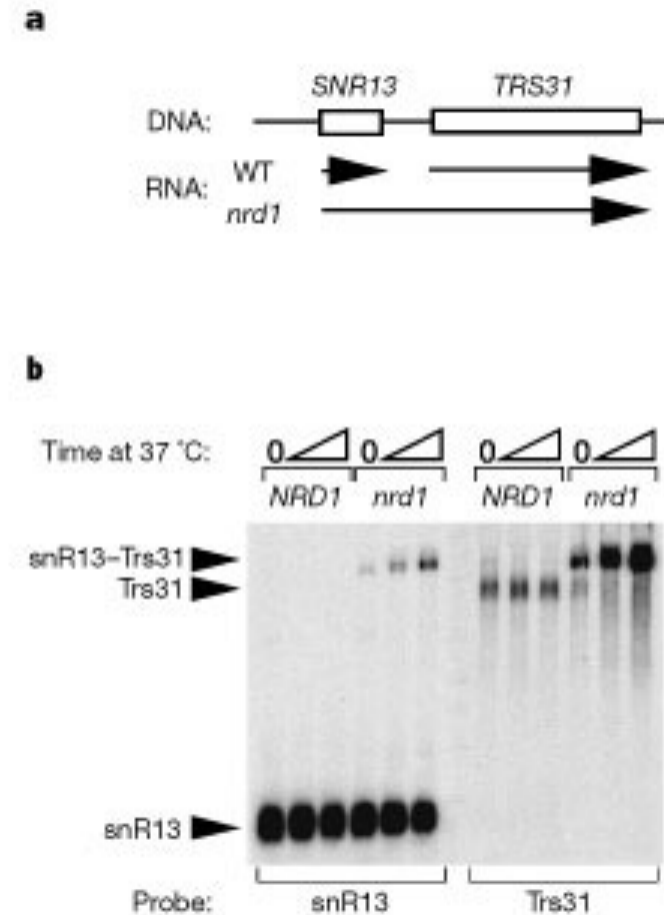


C Pol II ChIP



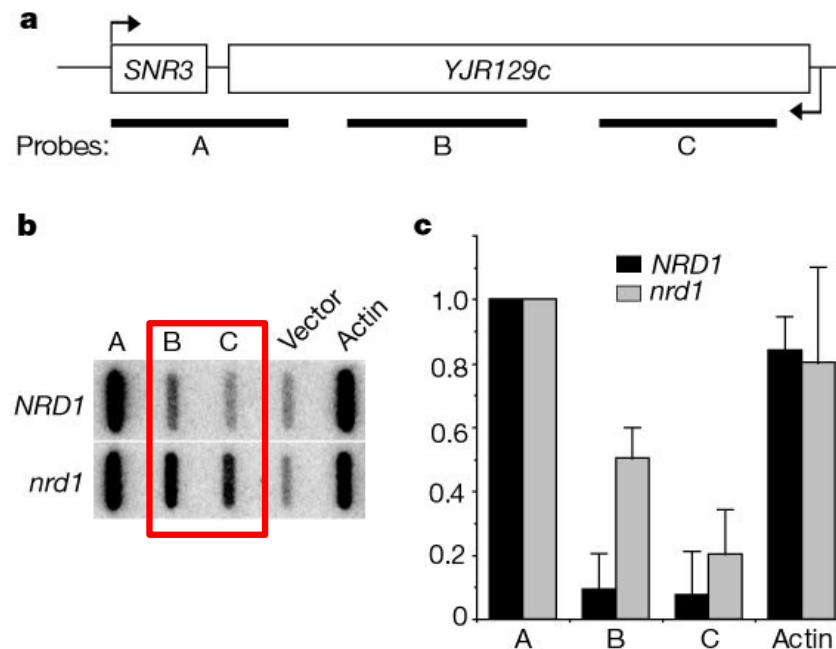
RNA-binding protein Nrd1 directs poly(A)-independent 3-end formation of RNA Pol II transcripts

- Direct evidence that ncRNAs might be natural targets for **Nrd1** was first obtained from expression profiling of poly(A)⁺ RNA (microarray) derived from temperature-sensitive *nrd1* yeast strains as compared with wild type. Several open reading frames (ORFs) exhibiting increased expression in the *nrd1* mutant strain are located downstream of snoRNA genes in the yeast genome.



RNA-binding protein Nrd1 directs poly(A)-independent 3-end formation of RNA Pol II transcripts

A significant proportion of polymerase molecules that terminate transcription downstream of the *snoRNA* coding region in the wild-type strain fail to do so in the *nrd1* mutant



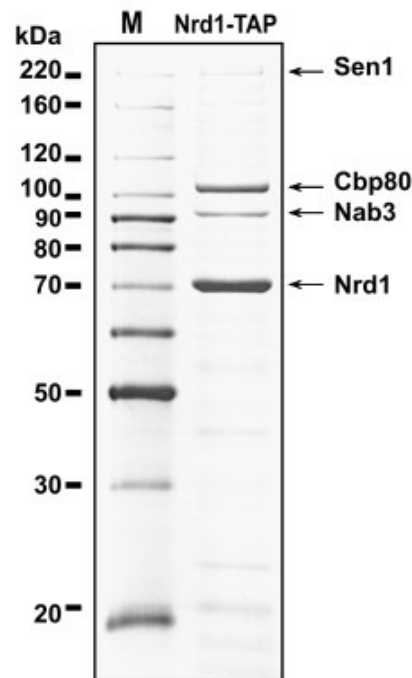
In contrast to *SNR13*, the downstream ORF is in a convergent orientation, therefore single-stranded probes for *SNR3* read-through transcripts will not hybridize with mRNA from the downstream ORF

transcription run-on analysis

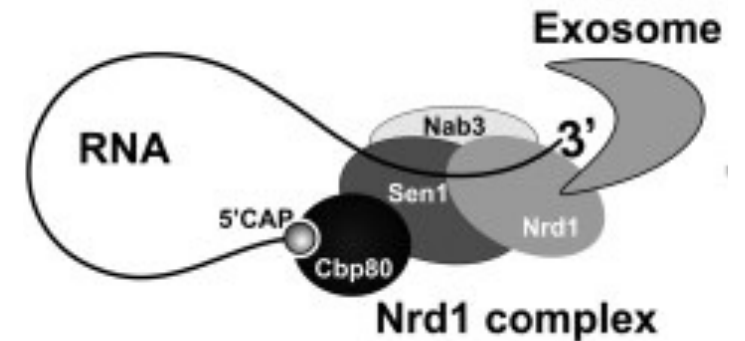
The Sen1-Nab3-Nrd1 complex

The alternative Pol II termination pathway for non-coding RNAs in yeast is composed by the RNA-binding proteins **Nrd1** and **Nab3**, and the RNA and DNA helicase **Sen1**.

TAP-purification

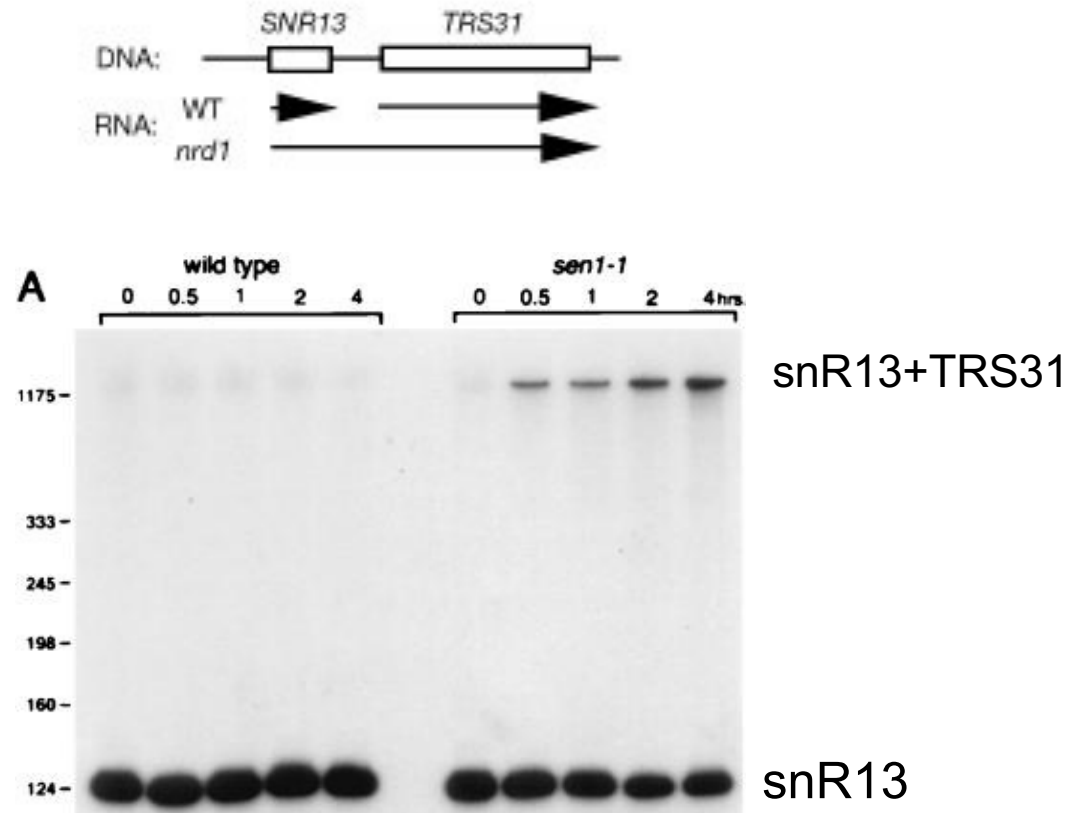


	Protein	Number of peptides	Protein function
Nrd1 complex	Sen1	57	RNA helicase, sn/snoRNA processing
	Nab3	42	sn/snoRNA termination
	Cbp80	32	Cap-binding
	Cbp20	5	Cap-binding
	Nrd1	27	sn/snoRNA termination
RNA pol II complex	Spt5	24	transcription factor
	Rpb1	24	RNApol II subunit
	Rpb2	19	RNApol II subunit
	Rpb3	9	RNApol II subunit
	Rpb5	3	RNApol II subunit
	Rpb4	2	RNApol II subunit
RNA processing complex	Rrp6	15	nuclear exosome
	Rrp44	5	exosome core
	Rrp4, Csl4, Rrp45, Rrp43	8	exosome core
	Air2	9	Trf4 cofactor
	Trf4	7	poly(A) polymerase
	Pab1	5	poly(A) binding
	Rnt1	3	sno/snRNA processing, endonuclease



The Sen1-Nab3-Nrd1 complex

The alternative Pol II termination pathway for non-coding RNAs in yeast is composed by the RNA-binding proteins **Nrd1** and **Nab3**, and the RNA and DNA helicase **Sen1**



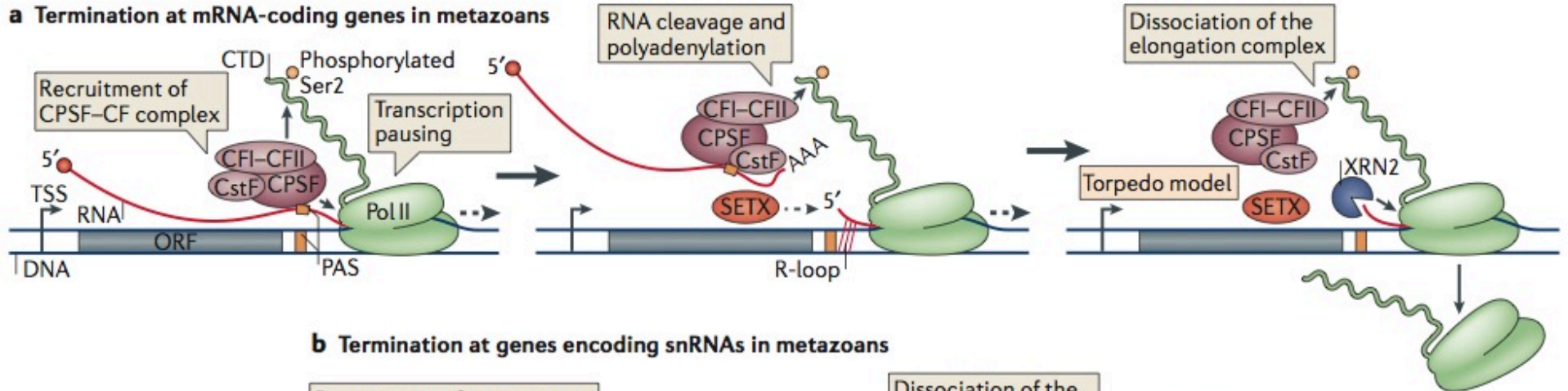
Transcription termination in metazoans

Three pathways of Pol II transcription termination have been described in metazoans, generating:

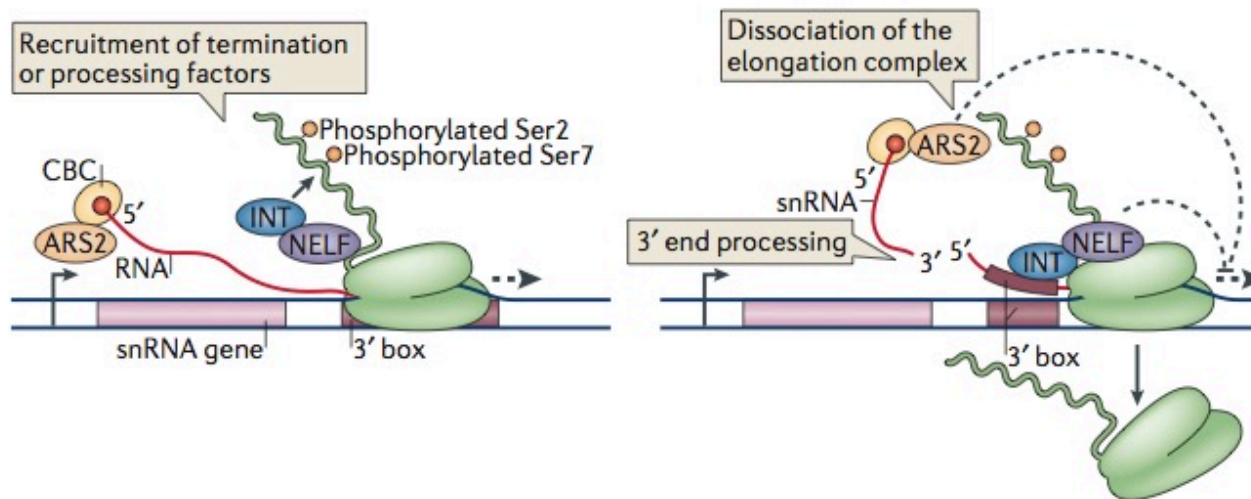
1. mRNAs,
2. snRNAs
3. transcripts encoding replication-dependent histones.

Transcription termination in metazoans

a Termination at mRNA-coding genes in metazoans

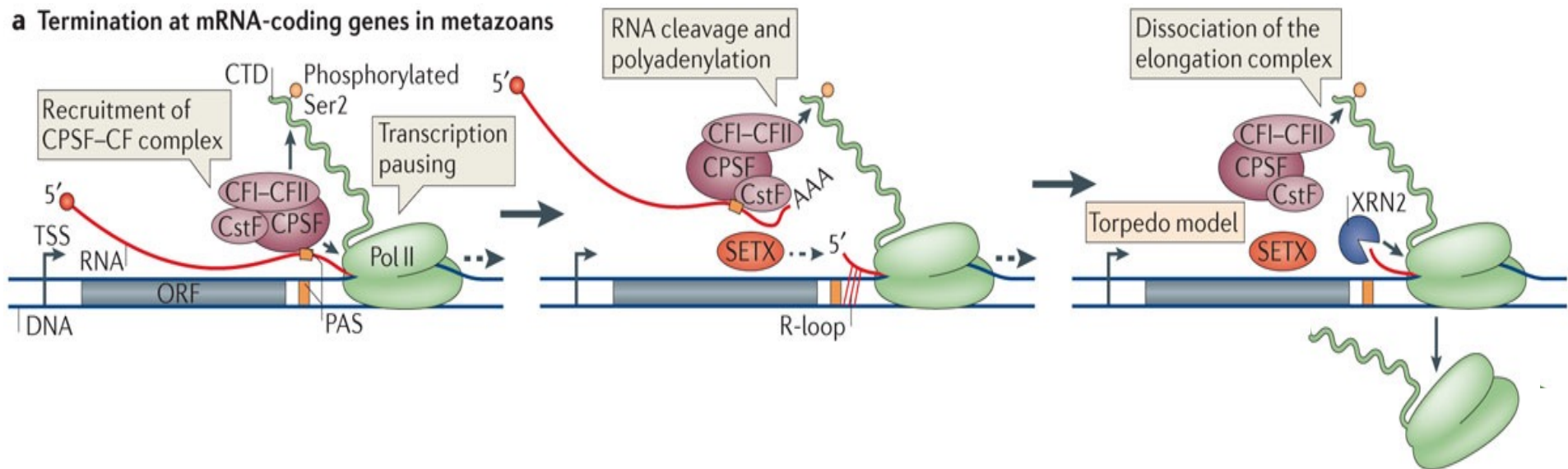


b Termination at genes encoding snRNAs in metazoans



General termination of mRNAs in metazoans

Cleavage of the nascent transcripts by the CPSF component CPSF73 (also known as CPSF3) occurs 18–30 nucleotides downstream of a polyadenylation signal (PAS; AAUAAA). It is commonly accepted that the PAS is required to trigger termination; A role for **senataxin (SETX)** in transcription termination of mRNAs has been proposed in several reports.

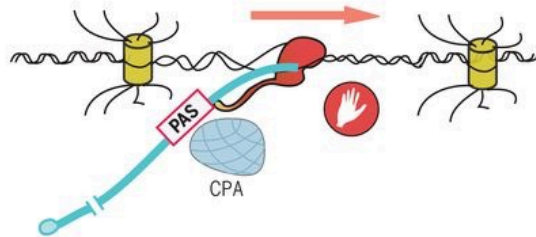


General termination of mRNAs in metazoans

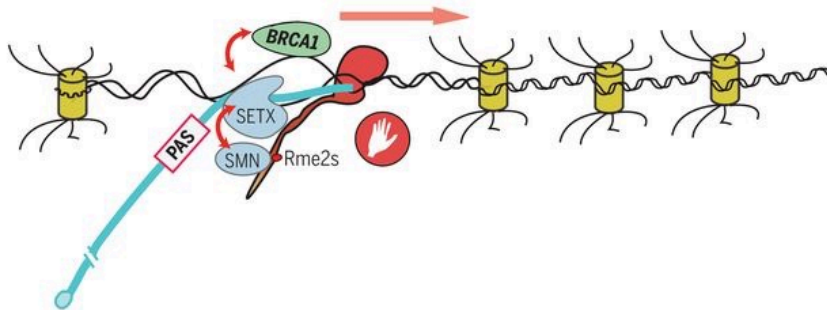
Transcription termination in metazoans is thought to be associated with Pol II pausing.

A Pausing

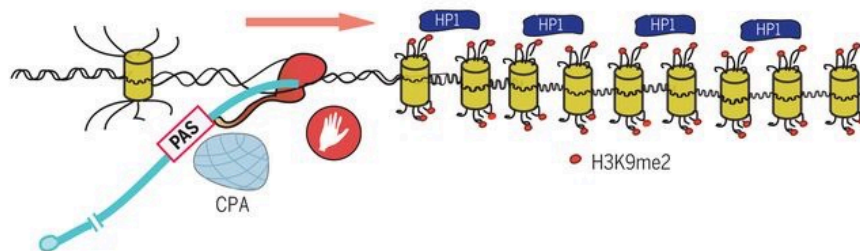
1 PAS-dependent pausing



2 R-loop-dependent pausing



3 Heterochromatin-dependent pausing



Three different types of **Pol II pausing**:

1. induced by CPA recognition of the PAS
2. R-loop formation
3. heterochromatin patches

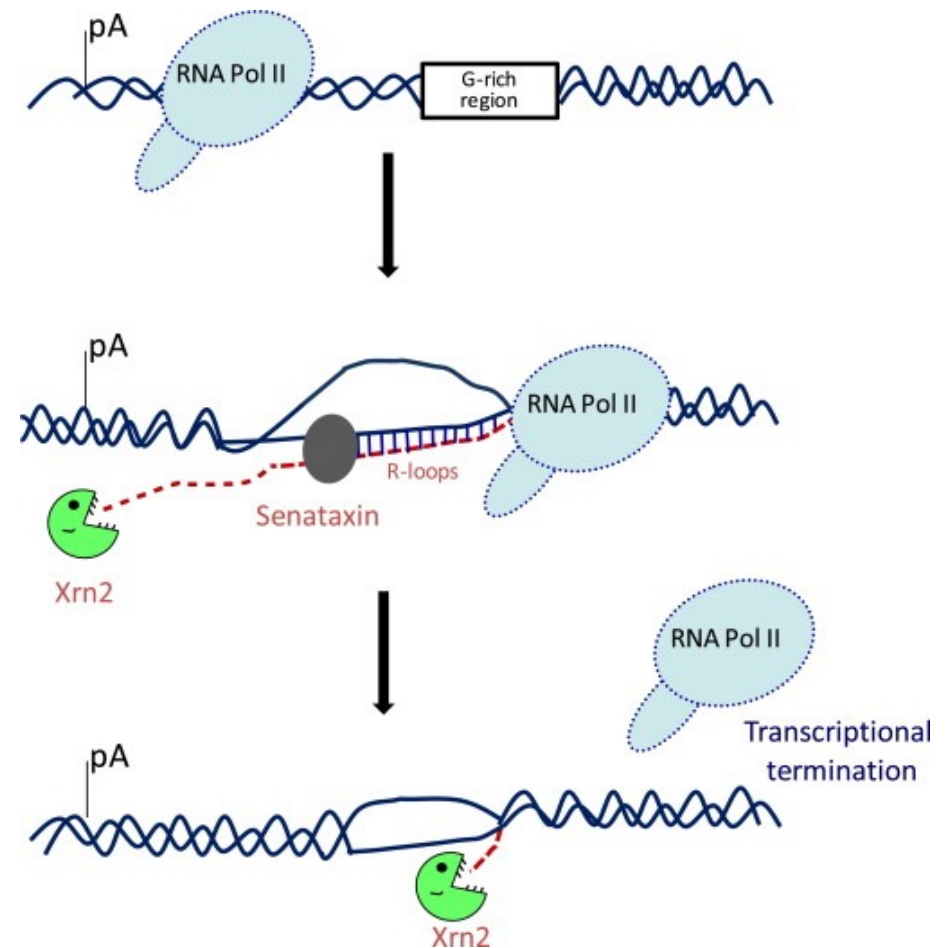
Elongating Pol II (red) is shown transcribing the DNA template, with extruded, capped RNA transcript (blue) indicated. Nucleosomes are depicted by yellow barrels, with histone N-terminal tails indicated. Pol II CTD is shown as an extended tail. Red dots on the CTD and histone tails denote methylation. The hand denotes Pol II pausing.



Senataxin is involved in RNA-Pol II transcription termination

Nascent transcripts form **RNA/DNA hybrid structures (R-loops)** behind elongating Pol II and are especially prevalent over G-rich pause sites positioned downstream of gene poly(A) signals. **Senataxin**, a helicase protein associated with neurodegenerative disorders (*Ataxia-Oculomotor Apraxia 2* and *amyotrophic lateral sclerosis 4*) acts to resolve these R-loop structures and by so doing allows access of the 5'-3' exonuclease Xrn2 at 3' cleavage poly(A) sites. This affords 3' transcript degradation and consequent Pol II termination.

R-loops formed over G-rich pause sites, followed by their resolution by senataxin, are key steps in the termination process.



R loops in human disease

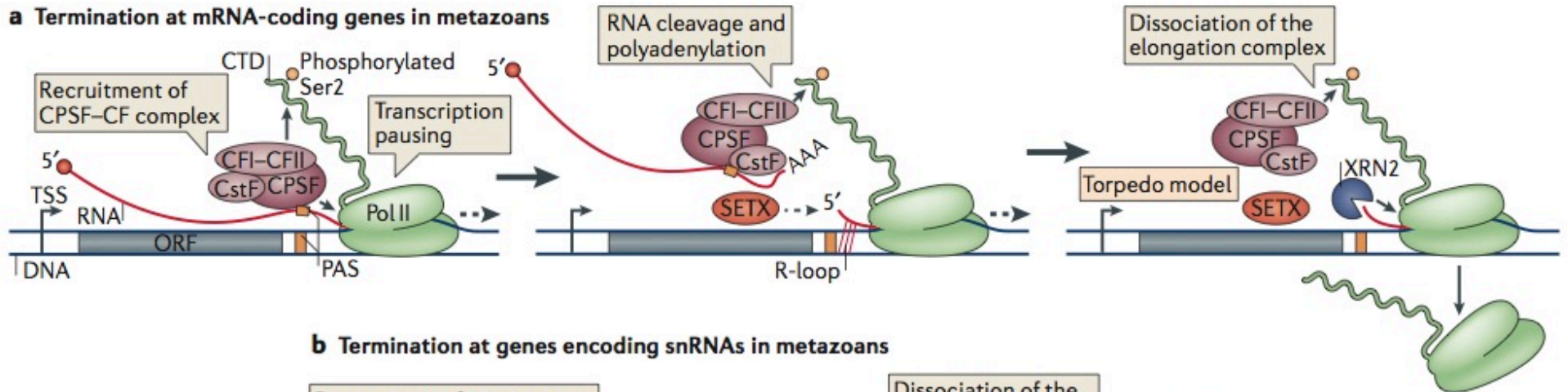
Table 2 | **Genes related to R-loop metabolism that can cause human diseases if dysfunctional**

Gene	Disease	Cause	Refs
<i>SETX</i>	Ataxia-ocular apraxia type 2 (AOA2) and amyotrophic lateral sclerosis type 4 (ALS4)	Mutations in the RNA–DNA helicase <i>SETX</i>	102,103
<i>FXN</i>	Friedreich ataxia (FRDA)	Expansion of GAA repeats in <i>FXN</i> gene promotes R-loop formation, H3K9me2 and decreased <i>FXN</i> expression	93,98
<i>FMR1</i>	Fragile X syndrome (FXS) and fragile X-associated tremor/ataxia syndrome (FXTAS)	Expansion of CGG repeats in <i>FMR1</i> gene promotes R-loop formation, H3K9me2 and decreased <i>FMR1</i> expression	93,99, 100
<i>C9orf72</i>	Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD)	Expansion of GGGGCC repeats causes R-loop formation and accumulation of aborted transcripts	101
<i>BRCA1</i>	Cancer	Genome instability caused by R-loop accumulation in <i>BRCA1</i> -deficient cells	80,95, 108
<i>BRCA2</i>	Cancer and Fanconi anaemia (FA)	Genome instability caused by R-loop accumulation in <i>BRCA2</i> -deficient cells	80
<i>FIP1L1</i>	Cancer	Genome instability caused by R-loop accumulation in <i>FIP1L1</i> -deficient cells inferred by the yeast mutant <i>fip1Δ</i>	21
<i>BRE1</i>	Cancer	Genome instability caused by R-loop accumulation in <i>BRE1</i> -deficient cells	111
<i>SRSF1</i>	Cancer	Deregulation of cancer-associated genes due to <i>SRSF1</i> overexpression	112
<i>ORF57</i>	Kaposi sarcoma-associated herpesvirus (KSHV)	Sequestration of human TREX complex by <i>ORF57</i> causes R-loop formation and DNA damage	113

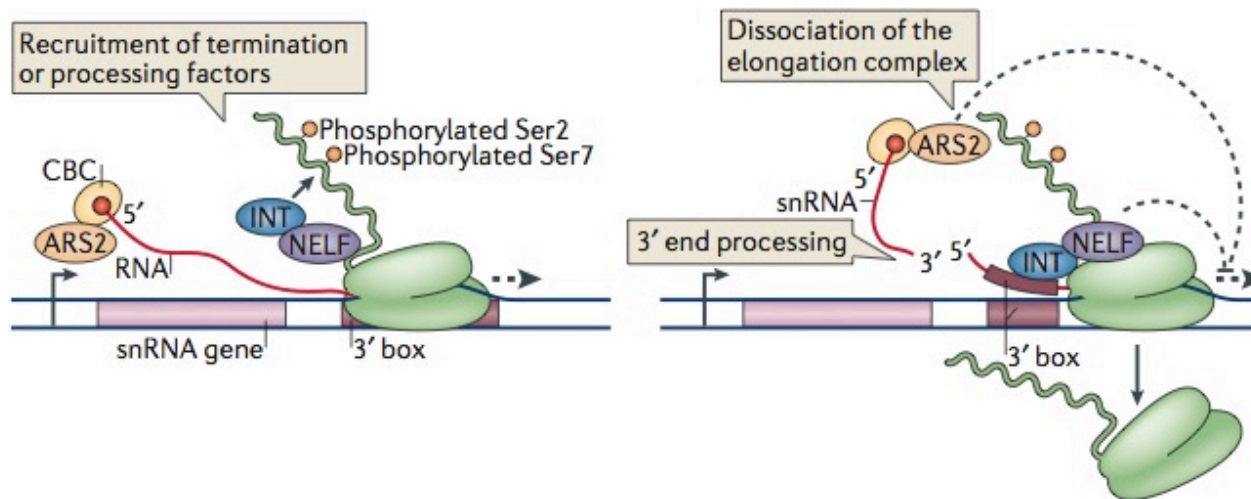
C9orf72, chromosome 9 open reading frame 72; *FIP1L1*, factor interacting with PAPOLA and CPSF1; *FMR1*, fragile X mental retardation 1; *FXN*, frataxin; H3K9me2, histone H3 lysine 9 dimethylation; *SETX*, senataxin; *SRSF1*, serine/arginine-rich splicing factor 1.

Transcription termination in metazoans

a Termination at mRNA-coding genes in metazoans

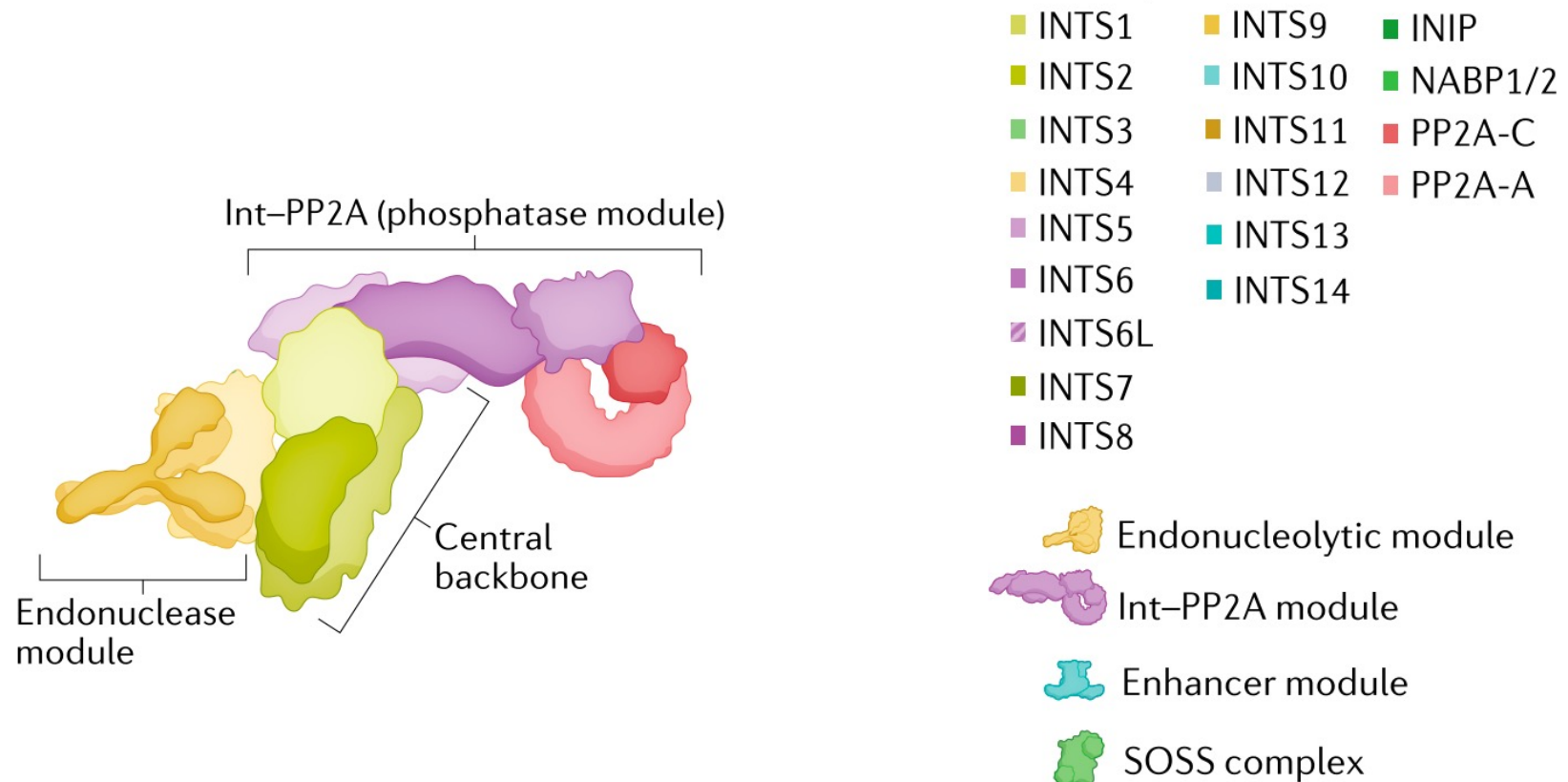


b Termination at genes encoding snRNAs in metazoans



The integrator (Int) complex

Integrator was discovered in 2005 as a new multisubunit complex in human cells capable of binding the Pol II CTD. **Orthologues of all subunits were identified throughout metazoans, but not in yeast, suggesting that the complex is unique to multicellular eukaryotes.** Sequence homology revealed that Integrator complex subunit 9 (INTS9) and INTS11 are highly homologous to cleavage and polyadenylation specificity factor subunit 73 (CPSF73) and CPSF100, providing the first hint that these INT subunits may be endowed with RNA endonuclease activity. Depletion of either the largest subunit (INTS1) or the putative catalytic core (INTS11) of Integrator resulted in specific accumulation of unprocessed, precursor snRNAs.



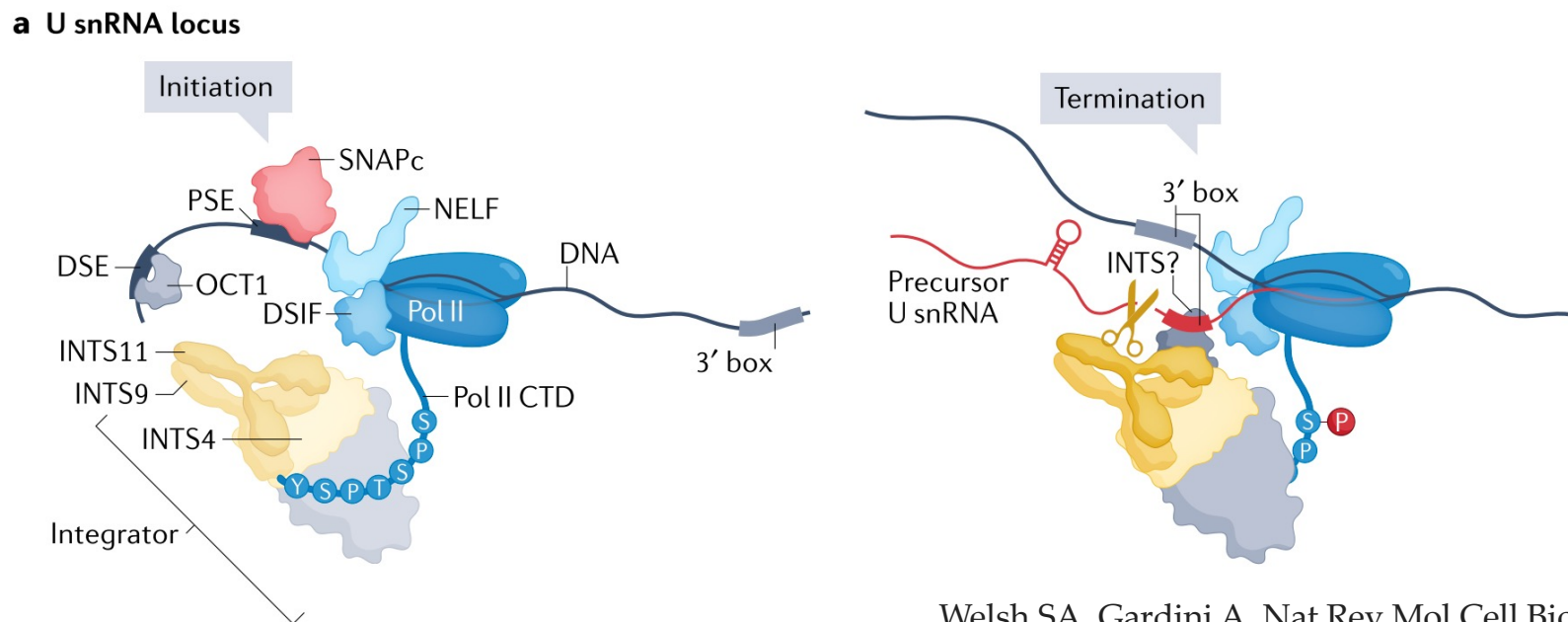
The functions of integrator at different non-coding RNAs

RNA species	Function of Integrator	Cell type, organism	Refs.
<i>Short RNAs (<0.2 kb)</i>			
U snRNAs	3' box-mediated cleavage	All cell types, various metazoans	21,23, 25,59,63
Viral miRNAs	Transcript release and maturation	Infected lymphocytes, marmosets	83,84
piRNAs	3' cleavage (unknown motif)	Germ cells, <i>Caenorhabditis elegans</i>	87,88
<i>Long RNAs (>0.2 kb)</i>			
eRNAs	3' cleavage (at unknown motif)	Multiple cell types, humans	58,81,82
TERC	3' cleavage (at unknown motif)	Human cell lines	92
NEAT1	Support of early transcription termination	Human cell lines	94
Other lncRNAs	Support of early transcription termination	Human cell lines	82

eRNA, enhancer RNA; lncRNA, long non-coding RNA; miRNA, microRNA; NEAT1, nuclear paraspeckle assembly transcript 1; piRNA, PIWI-interacting RNA; snRNA, small nuclear RNA; TERC, telomerase RNA template component.

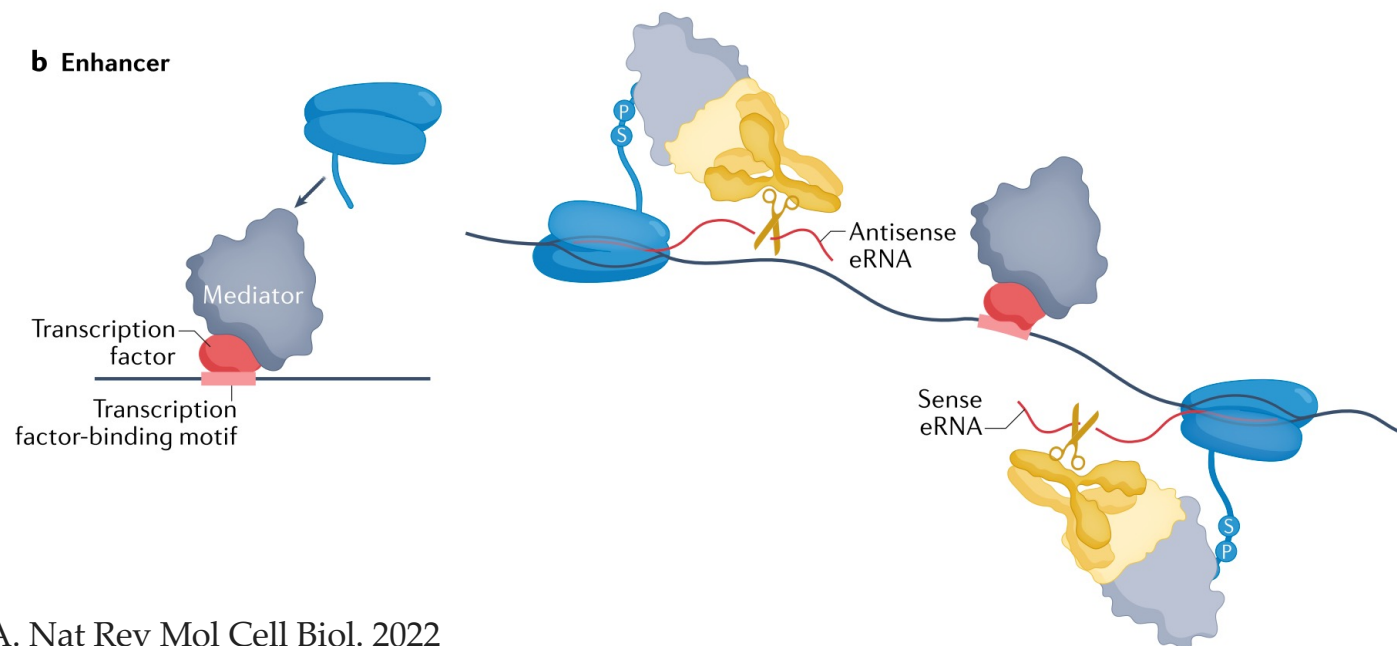
Integrator terminates transcription of snRNAs

snRNA loci have distinctive promoter elements that recruit the transcription initiation-competent RNA polymerase II (Pol II) holoenzyme including DSIF, NELF and the **Integrator complex**. Shortly after transcribing through the 3' box (13-16 nts), which is a highly conserved motif at the termination site of all U snRNAs, Integrator cleaves the nascent small RNA, triggered by phosphorylation (P) of Ser7 of the carboxy-terminal domain (CTD) of Pol II's largest subunit (RBP1). A 3' stem-loop in the precursor U snRNA and recognition of the ensuing 3' box RNA sequence by a set of Integrator accessory subunits (INTS?) may support an efficient cleavage process



Integrator terminates transcription of eRNAs

Enhancer loci are activated by sequence-specific transcription factors that recruit the co-activator Mediator complex. Upon Mediator recruitment and assembly of the transcription pre-initiation complex, bidirectional transcription of the enhancer locus occurs, producing long (>200-bp) sense and antisense transcripts called 'enhancer RNAs' (eRNAs). Both sense- transcribing and antisense-transcribing Pol II holoenzymes recruit the Integrator complex to terminate transcription and release eRNAs without eliciting their polyadenylation.

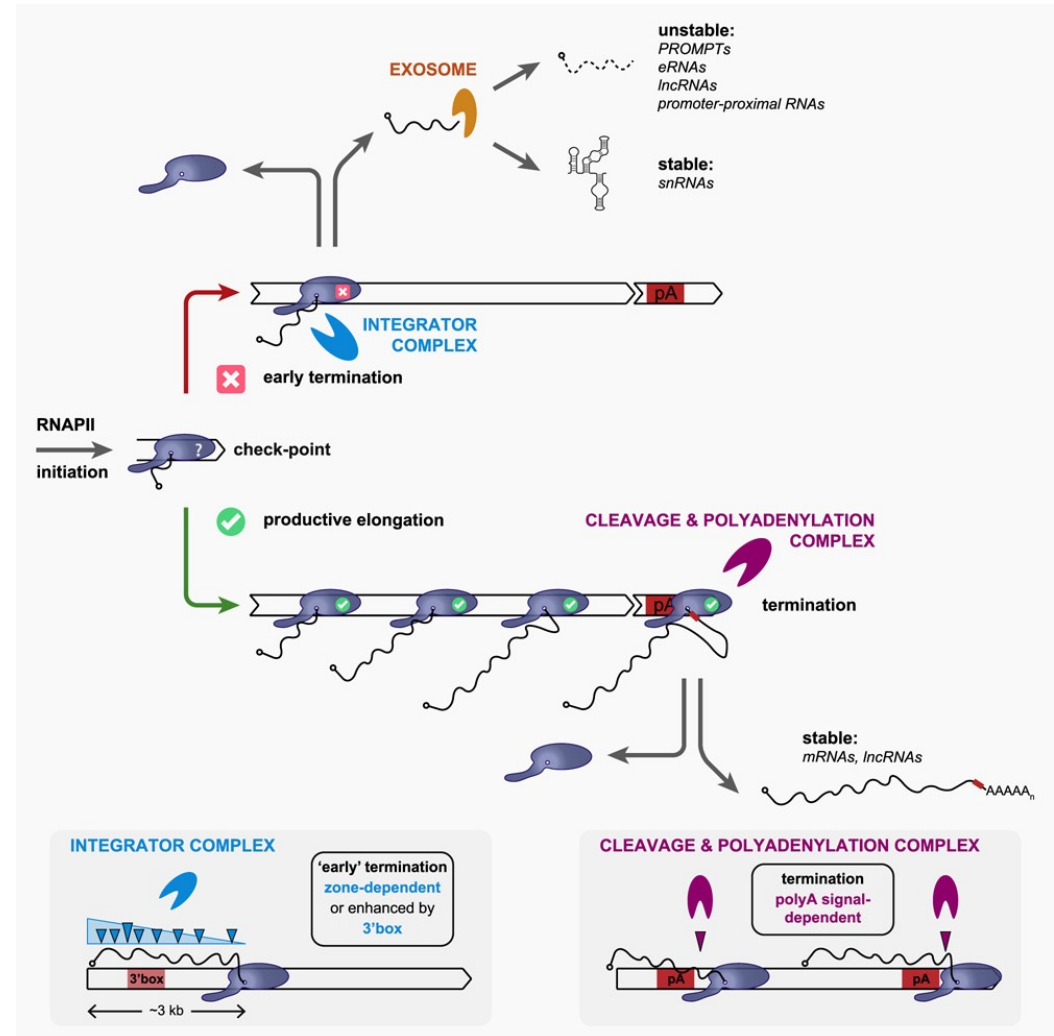


Integrator (INT) is a genome-wide attenuator of non-productive transcription

Two functions of genome-wide INT activity:

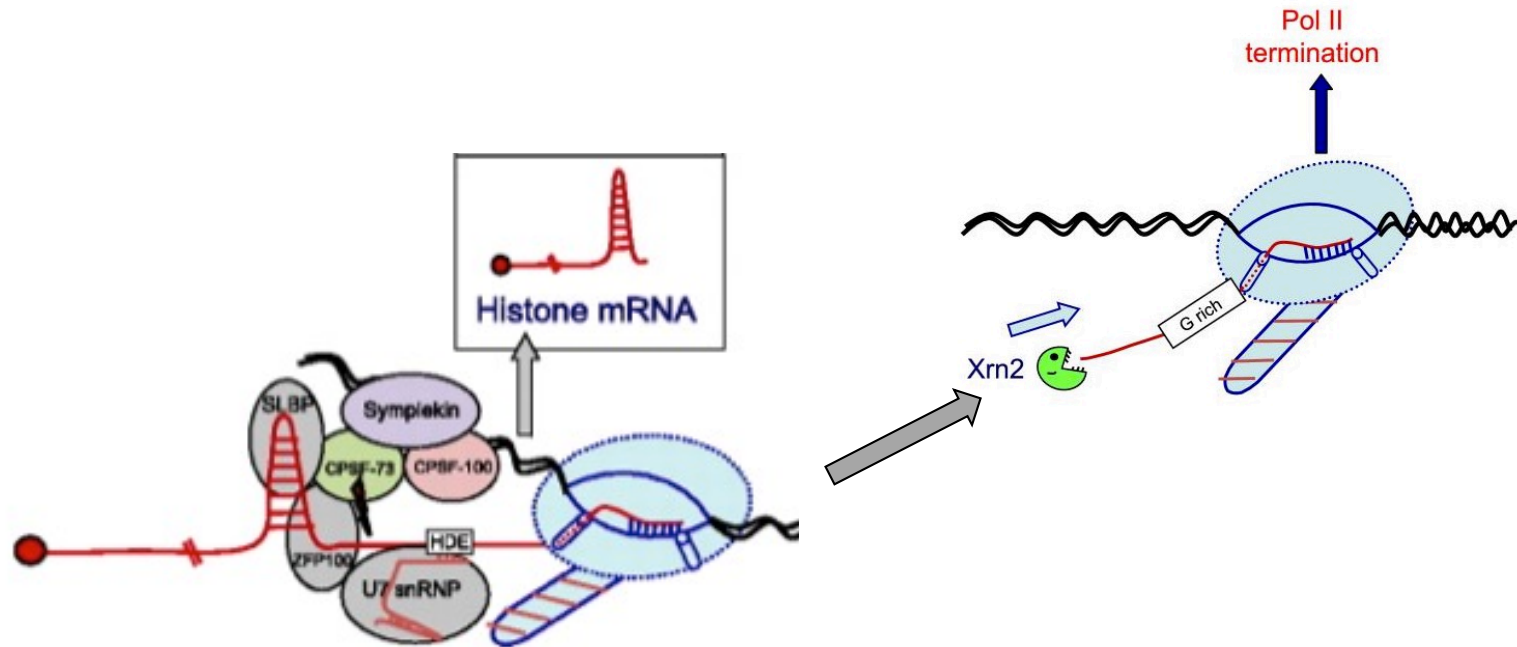
1. it dampens transcriptional output from weak promoters,
2. it provides quality control of RNAPII complexes that are unfavorably configured for transcriptional elongation.

The function of INT in stable snRNA production is an exception from its general cellular role, the attenuation of non-productive transcription.



Transcription termination at replication-dependent histones mRNAs

In contrast to many mRNAs, transcripts encoding replication-dependent histones are not polyadenylated but rather undergo cleavage at a particular stem-loop structure. U7 small nuclear ribonucleoprotein (snRNP), CBC, NELF, ARS2 and CPSF factors, including the CPSF73 endonuclease, have been shown to be involved in termination. A torpedo model has been proposed in which the Exonuclease (Xrn2) entry site is produced by CPSF-73 recruited by the U7/SLBP/ZFP100 complex.

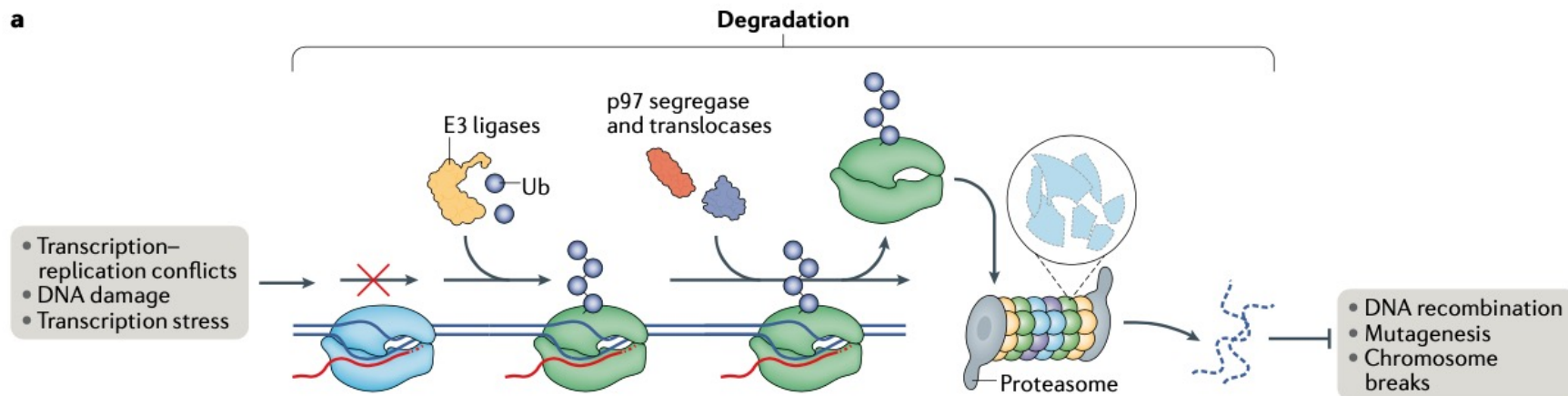


Fates of stalled or arrested RNA Pol II

Conditions that induce high levels of transcription stress can produce RNA Pol II stalling and arrest. Several mechanisms may contribute to its removal from the DNA template.

1. Pol II degradation

A variety of transcription stresses, including transcription–replication conflicts and DNA damage, lead to the proteasomal degradation of Pol II, suggesting that degradation may serve a general function in removal of Pol II from DNA. E3 ligases ubiquitylate Pol II and target it for degradation, whereas removal of Pol II from the DNA template for proteasomal degradation is aided by ATP-dependent segregases and translocases.

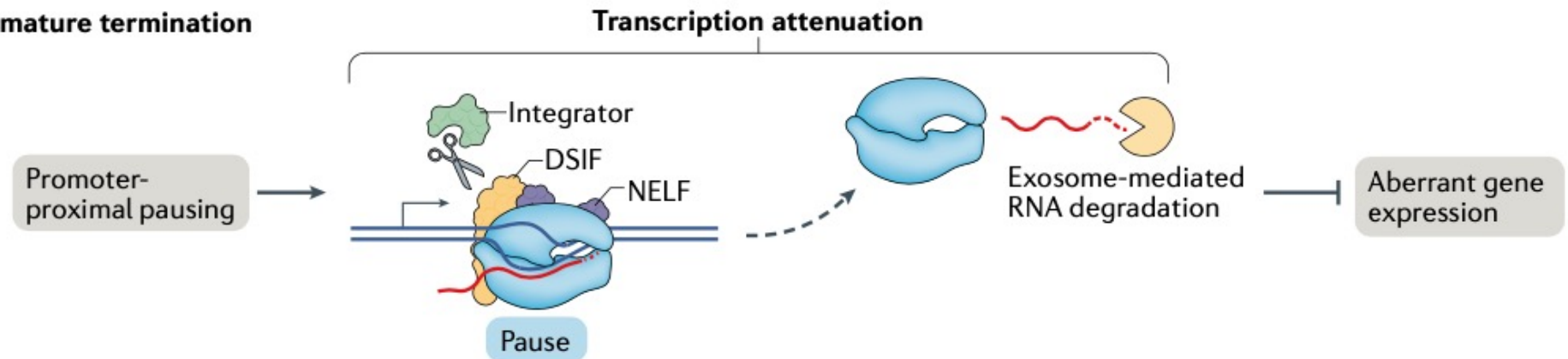


Fates of stalled or arrested RNA Pol II

2. Premature termination

During promoter-proximal pausing, Pol II may be removed from a gene by RNA cleavage coupled with Pol II dissociation, leading to premature transcription termination. This is likely an important aspect of gene regulation by promoter-proximal pausing. Paused Pol II may be terminated by the RNA cleavage complex Integrator

b Premature termination



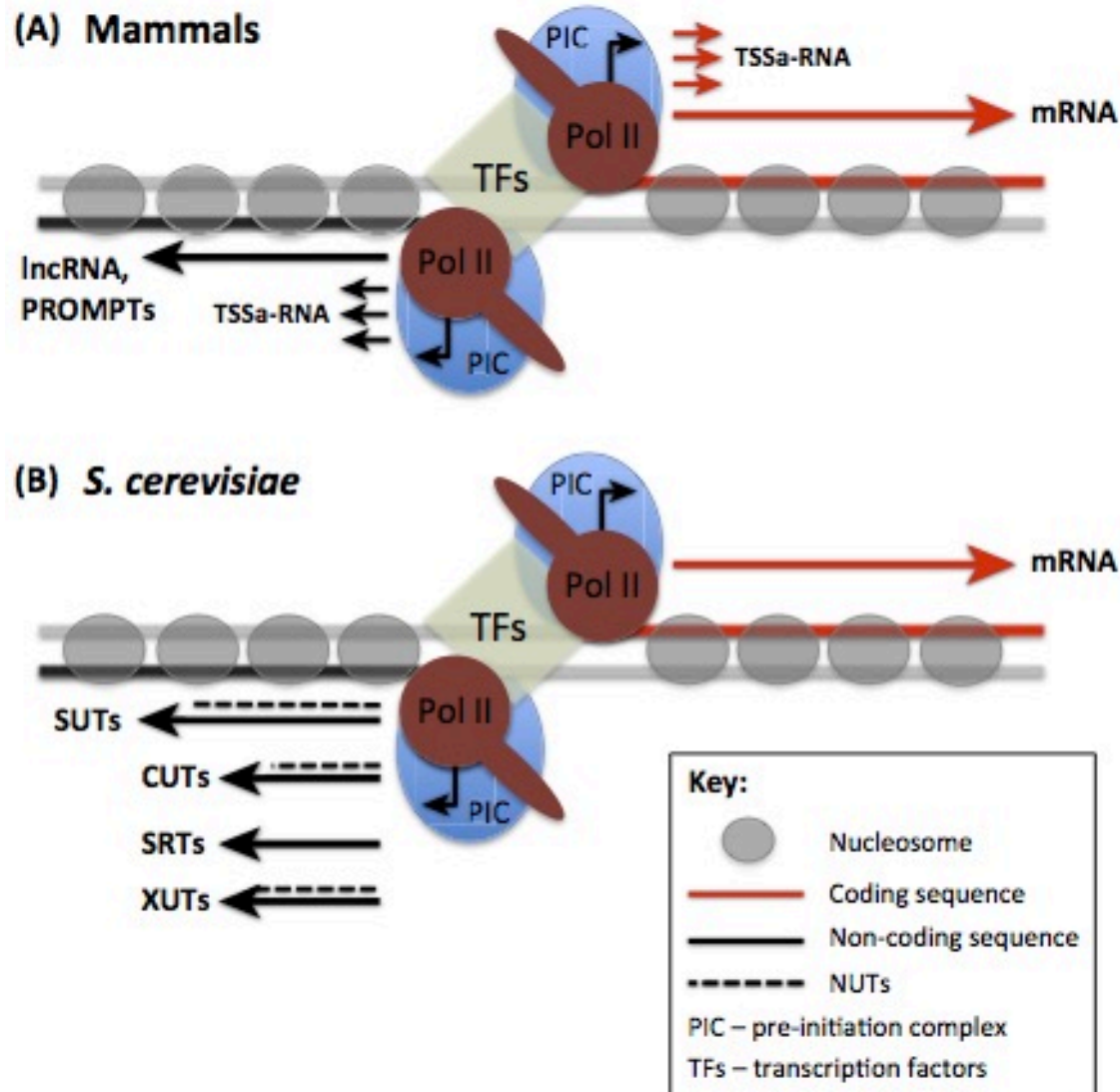
Pervasive transcription

- Only 10% of the elongating Pol II molecules in the yeast *Saccharomyces cerevisiae* are engaged in transcription that initiates from conventional promoters. The remaining 90% of the elongating Pol II molecules represent transcriptional noise. Of these, 60% are hyperphosphorylated on the C-terminal domain and associated with chromatin in a salt-stable manner, indicating that they are in the act of transcriptional elongation.
- Most RNAs initiated at inappropriate positions are unstable and rapidly degraded (exosome and nonsense-mediated decay) but a high proportions of eukaryotic genomes produce numerous stable noncoding and antisense RNAs.
- The relative proportions of biologically significant noncoding RNAs and transcriptional noise are unknown.
- Pervasive transcription provides the opportunity for the evolution of new genes.

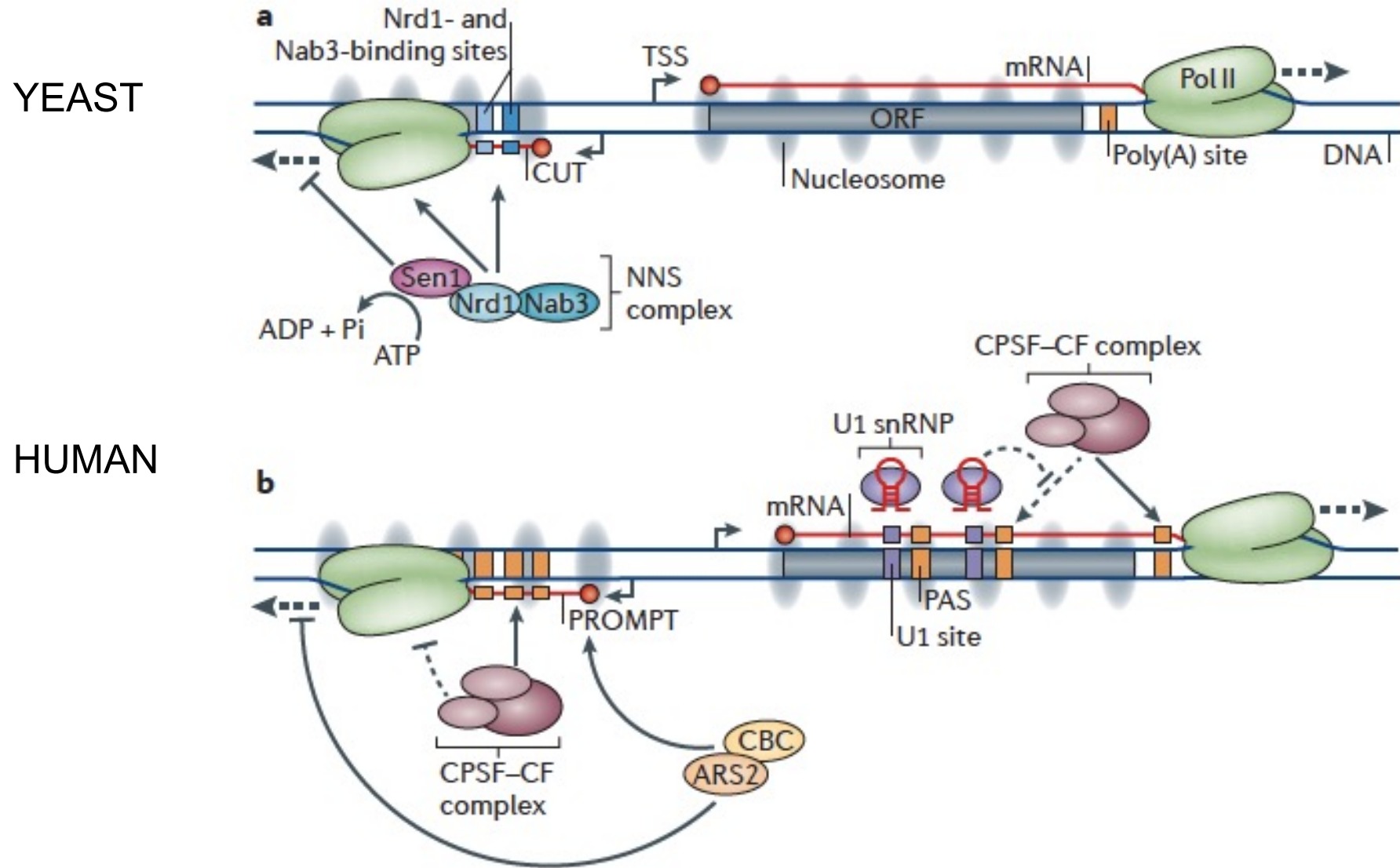
Promoter-associated noncoding RNAs

Genome-wide studies have unveiled the intrinsic bidirectional nature of many (if not all) promoters in yeast and humans,

The efficiency of bidirectional initiation after (generally) common transcription activation events is not symmetric because ‘meaningful’ transcription (that is, mRNA production) is generally preferred over non-functional transcription (that is, CUT production in yeast). The reason for this is that directional specificity is strongly influenced by the chromatin structure of the region of initiation in terms of the position of the Nucleosome Free Regions as well as the asymmetry in the chromatin marks of the flanking nucleosomes. Although the exact mechanisms have not been fully elucidated, many factors involved in chromatin remodelling or modification have been shown to suppress the bidirectionality of initiation and to control pervasive initiation



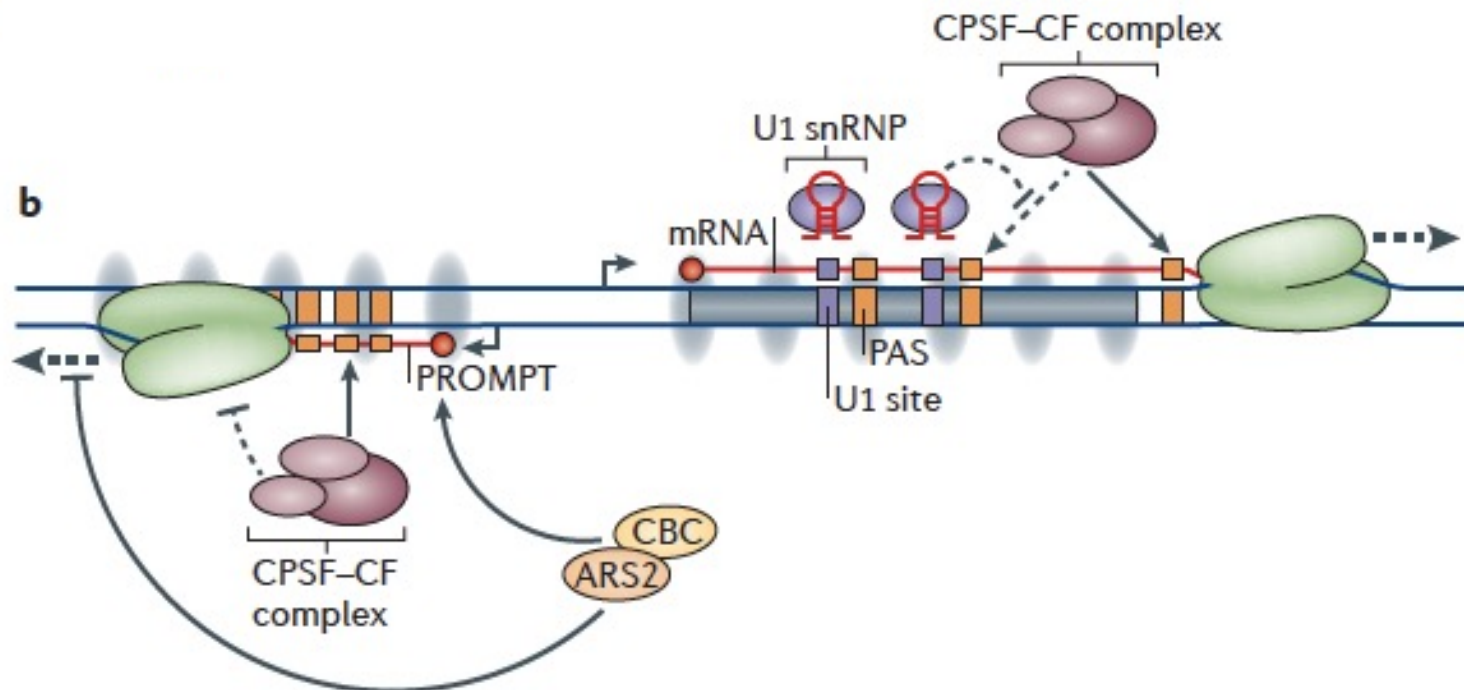
Transcription termination in the control of pervasive transcription throughout evolution



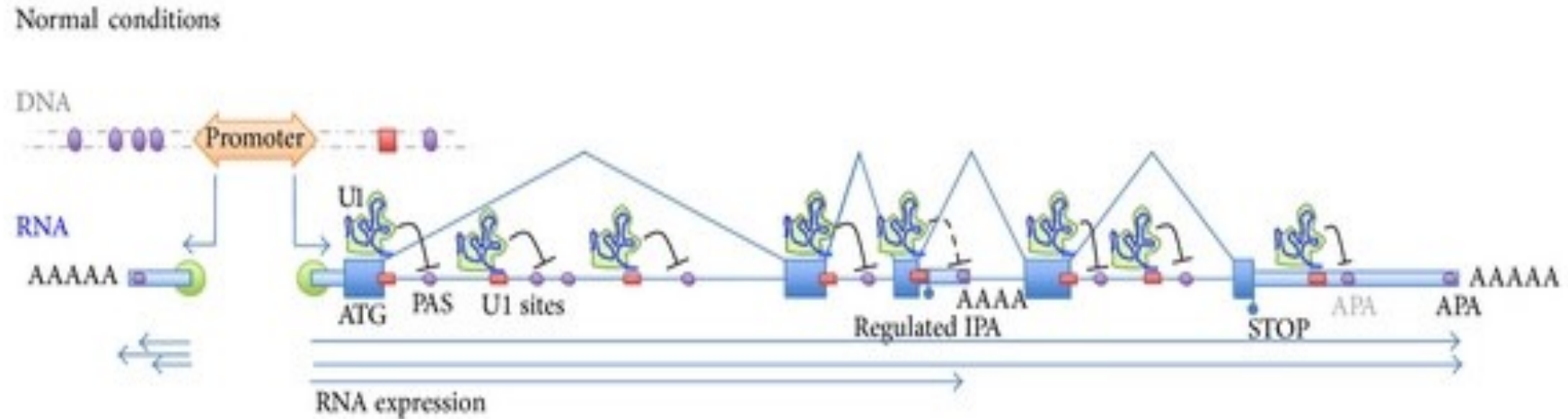
Directing transcription to the right way

Mammals

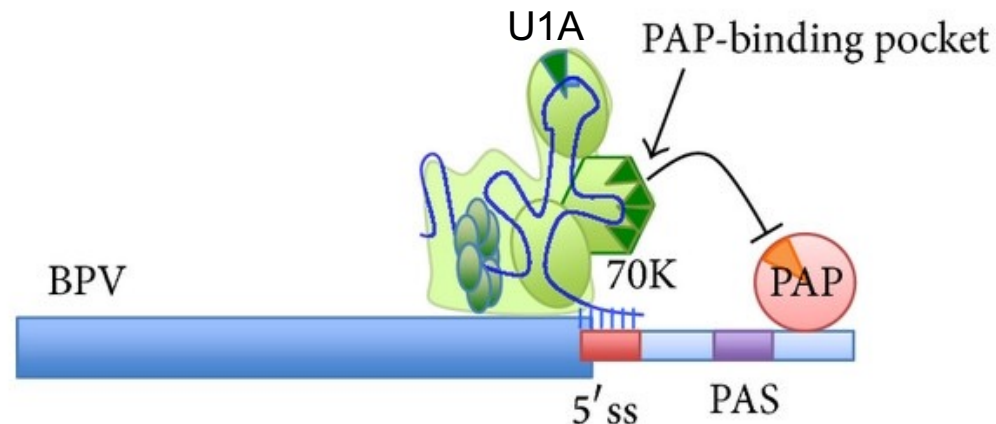
RNAPIIs are initiated in two opposite orientations from an active divergent promoter, and paused at the -1 and +1 nucleosomes, respectively. The paused RNAPIIs reinforce the boundaries of Nucleosome Free Regions and maintain the accessibility of promoter to transcription factors. **The CPSF-CF pathway** recognizes the PASs that are present more frequently in the non-functional transcript and induces promoter proximal termination. Interestingly, when present in the mRNA-coding direction, these termination signals are suppressed by the presence of antagonistic U1 snRNP-binding sites that have been shown to inhibit polyadenylation and termination.



U1 snRNP suppression of cleavage and polyadenylation safeguards transcriptome integrity

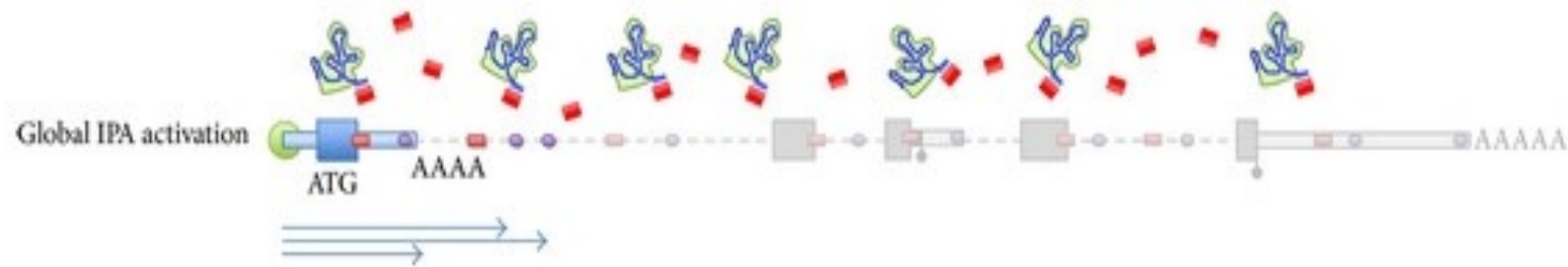


The U1 interacting protein U1-70K and in some context also U1A inhibit the PAS usage through direct PAP inhibition

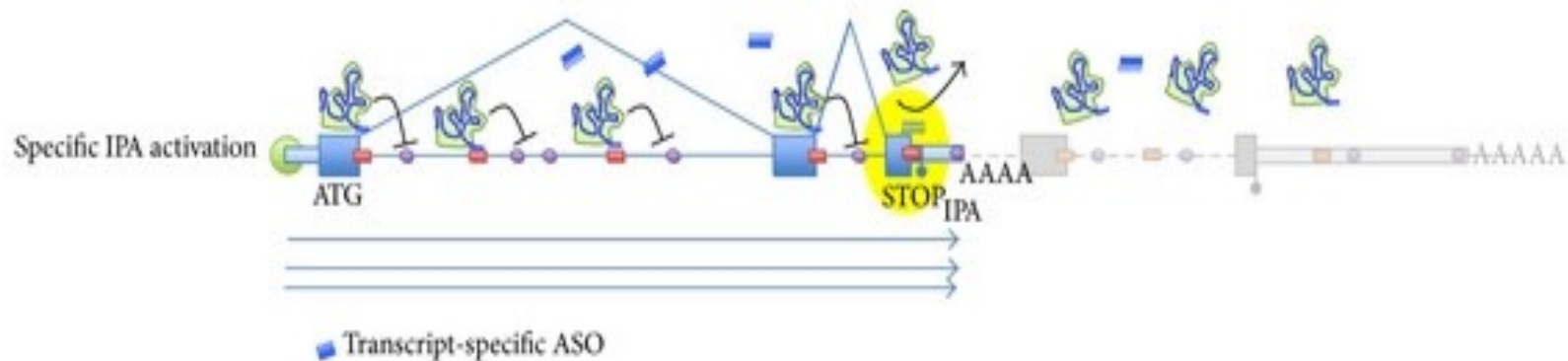


U1 snRNP-Dependent Suppression of Polyadenylation: Physiological Role and Therapeutic Opportunities in Cancer

Complete disruption of U1 activity by sequestering ASOs leads to loss of splicing and release of global IPA activation



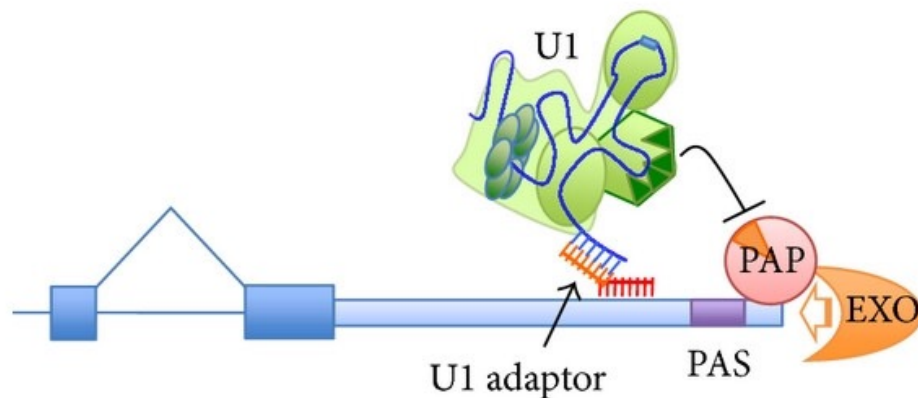
When ASOs targeted to a specific 5' ss are used, U1 binding is disrupted in that particular location but still functions normally elsewhere. The result is the selective activation of the targeted IPA site (highlighted), with expression of a truncated variant



Alternative 3' End Processing: Physiological Regulation and Deregulation in Cancer

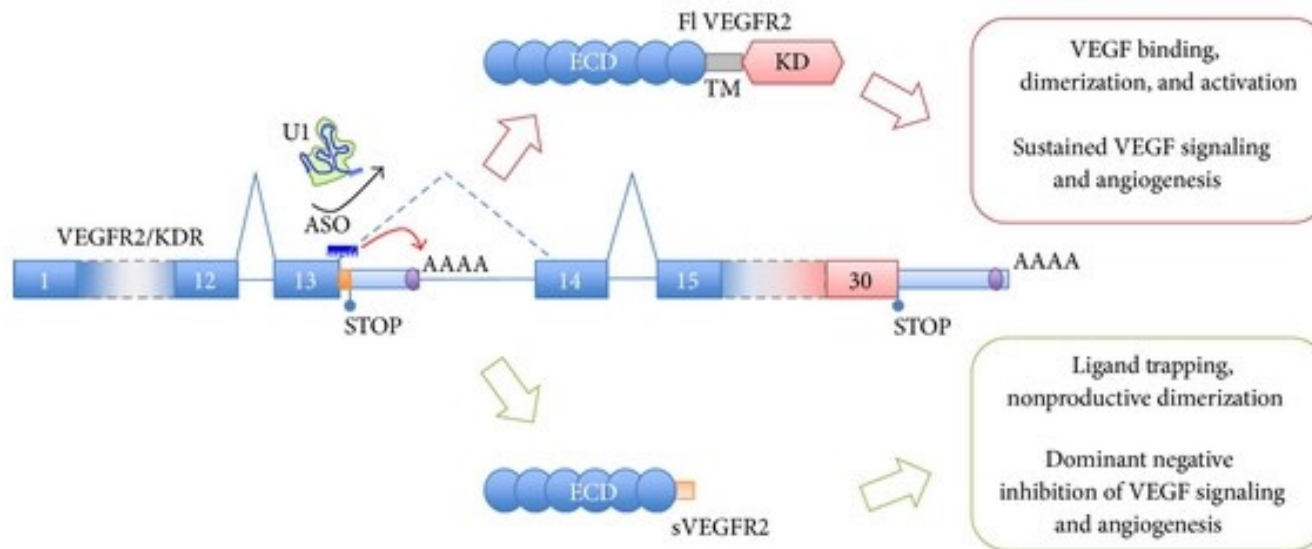
Specific and global modulation of APA are important aspects of many physiological processes, and their deregulation can contribute to the etiology of numerous diseases, including cancer. APA is associated with increased cellular proliferation and potentially with oncogenic transformation, **with the switch toward isoforms with a shorter 3' UTR in proliferating/cancer cells.**

In addition to global 3' UTR shortening, **activation of specific IPA sites can also be modulated in cancer cells,** with the generation of truncated variants possessing oncogenic properties or the suppression in tumors of antitumorigenic variants.



Thetering U1 to a specific site in order to inhibit the usage of PAS o IPA restoring to the production of full lenght mRNA

soluble VEGFR2 is a powerful natural inhibitor of angiogenesis and is underrepresented in tumors. Its induction by activation of a PAS in intron 13 of the VEGFR2/KDR pre-mRNA resulted in the generation of a soluble protein isoform that potently inhibited angiogenesis in a paracrine and autocrine fashion and also showed activity *in vivo*



Therapeutic potential of IPA activation: induction of secreted decoy VEGFR2. An IPA site in intron 13 of VEGFR2 can be specifically and effectively activated using ASOs targeted to the 5' ss immediately upstream, preventing U1 from binding and thus releasing suppression

Escaping nuclear decay

RNA survival in the degradative environment of the cell nucleus requires end-protective features.. To achieve this, more specialized RNA structures, such as the **triple-helical 3'-ends** of the lncRNAs NEAT1 and MALAT1 (part b) can be established. Long-term RNA survival is further provided by its **export to the cytoplasm** (part c) or its **sequestration in ribonucleoprotein (RNP) granules** such as paraspeckles (part d), as exemplified by the NEAT1 lncRNA, or **on chromatin** (part e), as exemplified by the accumulation of XIST on the inactive X chromosome. RNAs that lack such features get degraded by XRN2 or the nuclear RNA exosome, assisted by their respective decapping and poly(A) RNA exosome targeting (PAXT) cofactors (part f), composed by Trf4 and Zfc3h1.

