Transcription termination and the control of the transcriptome: why, where and how to stop

Odil Porrua and Domenico Libri

Abstract | Transcription termination occurs when the polymerase is released after a transcription event, thus delimitating transcription units; however, the functional importance of termination extends beyond the mere definition of gene borders. By determining the cellular fate of the generated transcripts, transcription termination pathways shape the transcriptome. Recent reports have underscored the crucial role of these pathways in limiting the extent of pervasive transcription, which has attracted interest in post-initiation events in gene expression control. Transcription termination pathways involved in the production of non-coding RNAs — such as the Nrd1–Nab3–Sen1 (NNS) pathway in yeast and the cap-binding complex (CBC)–ARS2 pathway in humans — are key determinants of transcription quality control. Understanding the mechanisms leading to the timely and efficient dismantling of elongation complexes remains a major unmet challenge, but new insights into the molecular basis of termination at mRNA-coding and non-coding RNA gene targets have been gained in eukaryotes.

Pervasive transcription
Widespread transcription, almost always by RNA polymerase II, that is not associated with annotated features such as protein-coding genes, small nuclear RNAs and small nucleolar RNAs.

Elongation complexes
Ternary complexes composed of the transcribing RNA polymerase, the DNA template and the nascent RNA.

Institut Jacques Monod, Centre Nationale pour la Recherche Scientifique (CNRS), UMR 7592, Université Paris Diderot, Sorbonne Paris Cité, F-75205 Paris, France.
Correspondence to D.L.
e-mail: Libri.Domenico@ijm.univ-paris-diderot.fr
doi:10.1038/nrm3943
Published online
4 February 2015

Transcription termination is needed to partition the genetic information encoded in DNA by defining the boundaries of transcription units. However, recent studies have uncovered unexpected regulatory roles of this final step of transcription in gene expression. As many termination factors interact with RNA processing and degradation enzymes, they are crucial in defining both the cellular fate and the half-life of the transcript. As a consequence, the choice of the termination pathway has a large influence on the ability of the newly synthesized RNAs to carry a genetic message: transcription termination of mRNA-coding genes generally leads to the production of stable transcripts that are directed to the cytoplasm for translation, whereas termination of some classes of non-coding RNAs (ncRNAs) results in their nuclear restriction and degradation.

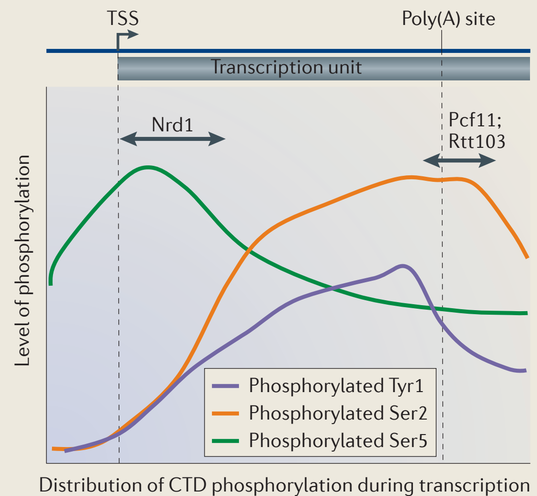
Many recent studies have also bolstered regulatory roles for termination. In some model systems, transcription is initiated, often in a constitutive manner, but gene expression is modulated by early termination (known as transcription attenuation). However, arguably one of the most prominent roles of transcription termination is in the control of pervasive transcription. In eukaryotes, transcription initiation by RNA polymerase II (Pol II) occurs promiscuously and is largely irrespective of annotated coding regions, which adds an additional layer of complexity to the transcriptome. A wealth of ncRNAs are produced, the function of which is subject to intense speculation. Such a crowded transcriptional landscape demands sophisticated control mechanisms to prevent disruptive overlapping sense or antisense transcription events and the accumulation of non-functional RNAs that might interfere with the metabolism of functional molecules[1–3]. Recent reports in many eukaryotic and bacterial model systems have underscored the crucial role of transcription termination pathways in limiting the extent of pervasively initiated transcription and in eliminating its products. Such a posteriori transcriptional quality control might be considered uneconomical, but it is likely to have a major evolutionary role, for instance, in the generation of new genes.

Transcription termination occurs when the polymerase and the nascent RNA are released from the DNA template; however, the molecular mechanisms that lead to the timely and efficient dismantling of elongation complexes remain poorly understood. The mechanistic basis of transcription termination is the subject of many studies, and new concepts on the mechanism of termination are emerging in the light of the extended roles of this process in gene expression.

© 2015 Macmillan Publishers Limited. All rights reserved

Box 1 | **The carboxy-terminal domain of RNA polymerase II**

The carboxy-terminal domain (CTD) of the largest subunit of RNA polymerase II (Pol II) is exclusively found in eukaryotic Pol II[121] and consists of tandem repeats of the heptapeptide Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7 (26 repeats in yeast and 52 in humans). The CTD functions as a binding platform for many factors that are involved in all stages of transcription and transcription-coupled processes, such as capping, splicing and RNA processing (reviewed in REFS 122–124). The CTD undergoes post-translational modifications throughout the transcription cycle — in particular, phosphorylation at Tyr1, Ser2, Ser5, Ser7 and Thr4 (REFS 44–46,124–127). The most relevant modifications for transcription termination are phosphorylation at Tyr1, Ser2, Ser5 and Ser7, which alter the binding specificity of termination factors (see the figure). For instance, in yeast the cleavage and polyadenylation factor (CPF)–cleavage factor (CF) complex component Pcf11 and the Rat1-interacting protein Rtt103 interact preferentially with the Ser2-phosphorylated (Ser2P) form of the CTD[29]. These proteins interact with the Ser2P CTD only when Tyr1 is dephosphorylated[127]. Phosphorylation of Ser2 and Tyr1 accumulate during transcription over the body of the gene, but Tyr1 phoshorylation level sharply decreases just before the 3′ end of the gene, whereas Ser2 phosphorylation level remains high, enabling the interaction of Pcf11 and Rtt103 with the CTD. Conversely, the Nrd1–Nab3–Sen1 (NNS) component Nrd1 recognizes the Ser5P form of the CTD[41,43], which predominates early in transcription, but only before the phosphorylation of Tyr1, which possibly helps to restrict the recruitment of the NNS complex to the early stages of transcription[127]. Aside from phosphorylation, isomerization of the peptidyl-prolyl bond of the CTD from the *cis* to the *trans* conformation by the prolyl isomerase Ess1 (REF. 128) can also affect the recruitment of termination factors; for instance, Nrd1 binding to the CTD requires the *cis* conformation[41].



Distribution of CTD phosphorylation during transcription

TSS, transcription start site.

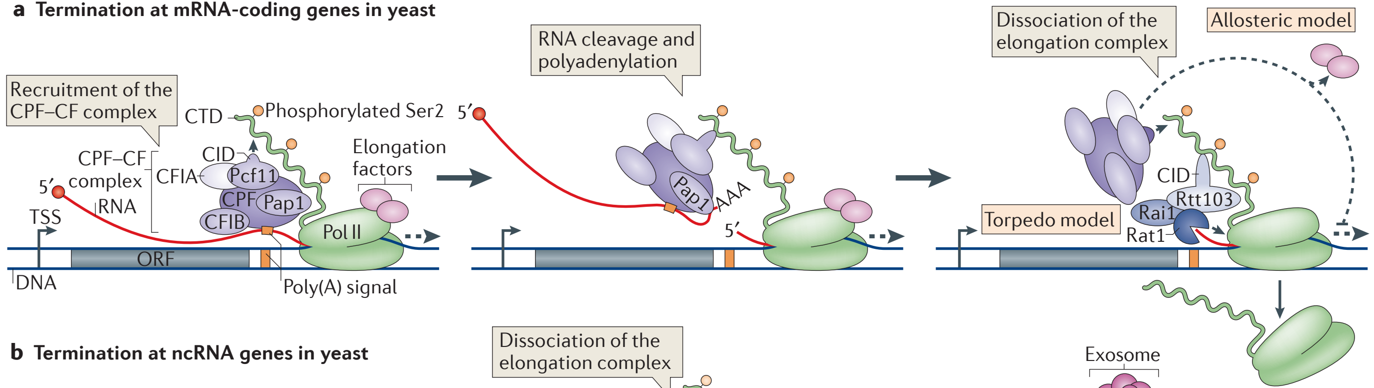
In this Review, we describe the various termination pathways for Pol II-dependent transcription, with a main focus on the most well-studied system, yeast, including the cleavage and polyadenylation factor (CPF)–cleavage factor (CF)- and Nrd1–Nab3–Sen1 (NNS)-dependent pathways. We discuss how the different pathways are specifically assigned to mRNA-coding and ncRNA gene targets, and how they cooperate to enforce transcription boundaries. We describe the role of termination pathways in regulating gene expression, in maintaining the stability of the transcriptome and in controlling pervasive transcription. Metazoan systems are addressed with a similar focus; in particular, we review transcription termination at regions encoding mRNAs, small nuclear RNAs (snRNAs) and replication-dependent histones. We refrain from a detailed description of the factors involved and their function because this subject has recently been covered by some excellent reviews[4–6].

Transcription termination in yeast

Termination of mRNA-coding genes: the CPF–CF pathway. Transcription termination of protein-coding genes is mainly dependent on three complexes: CPF, CFIA and CFIB. These termination factors are generally conserved in eukaryotes (reviewed in REFS 4,5,7; see below).

Several components of the CPF–CF complex (Rna15, Cft1, Cft2, Yth1, Mpe1 and Hrp1 (also known as Nab4)) recognize termination and processing signals in the 3′ untranslated region (UTR) of the nascent RNA. The Pcf11 subunit recognizes the Ser2-phosphorylated (Ser2P) form of the carboxy-terminal domain (CTD) of the largest Pol II subunit with its CTD-interacting domain (CID)[4–8] (BOX 1; FIG. 1a). Pcf11, in complex with another CPF–CF subunit, Clp1, has also been shown to bind to the body of the polymerase by interacting with the flap loop domain of the second largest subunit of Pol II, Rpb2 (REF. 9). It is commonly accepted that both sequences on the nascent RNA and the interaction with the polymerase contribute redundantly to the recruitment of the CPF–CF complex at the 3′ end of genes, although the relative contribution of each remains unclear. Subsequently, the RNA is cleaved by the CPF endonuclease Ysh1 at the poly(A) site, and adenosine nucleotides are added to the free hydroxyl group on the 3′ end (3′OH) by the CPF-associated poly(A) polymerase Pap1. The newly formed poly(A) tail is bound by a poly(A)-binding protein, which is thought to affect the length of the tail and to be required both for protecting the RNA from 3′ end degradation and for promoting nuclear export. However, there is some controversy concerning the identity of the poly(A)-binding protein, and two shuttling factors — Pab1 and Nab2 — have been proposed to bind to early to poly(A) tails and escort the RNAs to the cytoplasm[10,11]. Pab1, which has a mainly cytoplasmic localization, is important for translation, whereas Nab2 is mainly nuclear and has recently been reported to have a role in targeting RNAs for degradation in the nucleus in both *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*[12,13].

© 2015 Macmillan Publishers Limited. All rights reserved



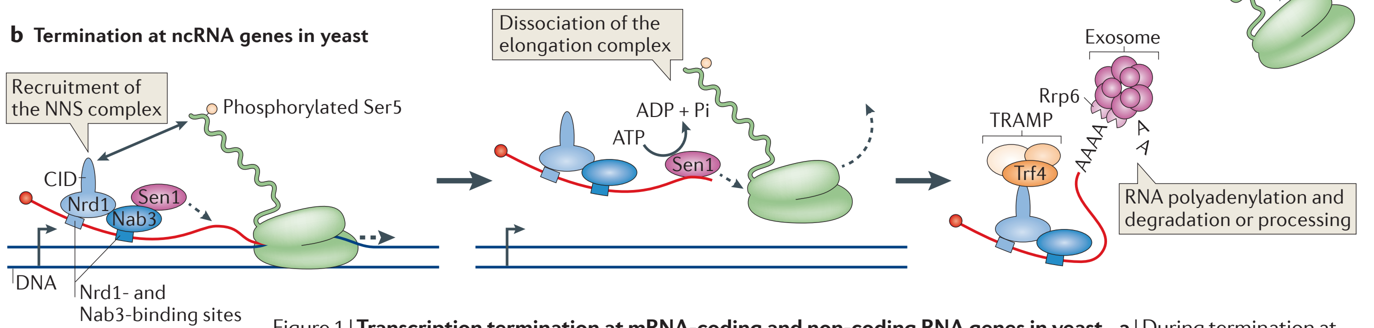


Figure 1 | **Transcription termination at mRNA-coding and non-coding RNA genes in yeast.** **a** | During termination at mRNA-coding genes, components of the cleavage and polyadenylation factor (CPF) and cleavage factor (CF) complexes — CPF, CFIA and CFIB — recognize specific sequences in the 3′ untranslated region (UTR) of the transcript. The CFIA component Pcf11 interacts with the Ser2-phosphorylated form of the carboxy-terminal domain (CTD) of RNA polymerase II (Pol II) through its CTD-interaction domain (CID). Upon endonucleolytic cleavage of the transcript at the poly(A) site, poly(A) tails are added by the CPF-associated poly(A) polymerase Pap1. The 5′ end of the downstream portion of the transcript is then targeted by the Rat1 5′–3′ exonuclease. Two alternative models are proposed for the mechanism of termination after transcript cleavage. The allosteric model posits that loss of elongation factors and/or conformational changes in the polymerase after transcription of the poly(A) signal destabilizes the elongation complex. The torpedo model postulates that the Rat1 exonuclease (alone or in complex with its cofactor Rai1) is recruited by the CTD-interacting protein Rtt103 and degrades the nascent RNA after cleavage. The interaction of Rat1 with the polymerase leads to the dissociation of the elongation complex. **b** | During termination at non-coding RNA (ncRNA) genes, the Nrd1–Nab3–Sen1 (NNS) complex is recruited to the elongation complex through the recognition of specific motifs on the nascent RNA by Nrd1 and Nab3, and the interaction of the Nrd1 CID with the Ser5-phosphorylated form of the CTD. The RNA and DNA helicase Sen1 is then loaded onto the RNA, where it uses the energy of ATP hydrolysis to 'catch up' with Pol II and elicit termination. In a subsequent phase, the RNA-bound Nrd1–Nab3 heterodimer interacts with the TRAMP (Trf4–Air2–Mtr4) complex through the Nrd1 CID, which promotes polyadenylation of the transcript and its degradation or processing by the exosome and Rrp6. ORF, open reading frame; Pi, inorganic phosphate; TSS, transcription start site.

RNAs that are terminated by the CPF–CF pathway are rapidly exported to the cytoplasm, and their half-lives are generally determined by the cytoplasmic turnover pathways. Export to the cytoplasm is generally fast and not limiting[14,15], which has been suggested to subtract mRNAs from competing nuclear degradation activities[3,16]. However, the mechanisms underlying the coupling between CPF–CF termination and nuclear export remain unclear. Export competence might be acquired early during transcription by the recruitment of specific factors[17] or be conferred when 3′ end processing occurs.

After the mRNA is cleaved from the nascent RNA, the polymerase transcribes the downstream DNA for a length that is unlikely to exceed 150 nucleotides[18]. The polymerase is then released from the DNA by mechanisms that are under debate. Although it has generally been possible to address the function of individual CPF–CF components in the RNA cleavage and polyadenylation steps using *in vitro* systems, reconstitution of the termination reaction using purified components has proved to be more challenging, and most studies rely on *in vivo* data. Mutation of many components of the CPF–CF complex induces both a processing and an Pol II release defect *in vivo,* which suggests that this complex is required for both steps. Although cleavage can occur *in vitro* in Pcf11 mutants that are defective for polymerase release[19], cleavage-defective mutants are generally also impaired for termination[19]. These findings support the notion that cleavage is important for polymerase release. Consistent with (but not proving) this model, a recent genome-wide analysis of Pol II distribution after depletion of the CPF Ysh1 endonuclease revealed termination defects at protein-coding genes[20].

Two models have been proposed for CPF–CF-dependent termination (FIG. 1a). The allosteric model posits that after transcription of the poly(A) site, binding of the termination complex results in a conformational change of the elongation complex owing to the loss of elongation or anti-termination factors, which decreases processivity and ultimately leads to termination (reviewed in REF. 6). In support of this model, it has been shown that the polymerase loses associated elongation factors before being released[21,22]. It has also been shown that Pcf11 alone

Anti-termination factors
Protein or protein complexes that protect the transcribing RNA polymerase from transcription termination.

© 2015 Macmillan Publishers Limited. All rights reserved

Table 1 | **Transcripts and associated termination, processing and degradation pathways**

Transcript	Termination pathway	Stability	Degradation factors	Refs
Yeast				
mRNA	CPF–CF and possibly Sen1	Stable	None	7
snRNA and snoRNA	NNS	Stable (3′ end processed)	TRAMP, Rrp6, exosome, Rex1 (3′ end processing)	16,129
CUT	NNS	Unstable	TRAMP, Rrp6, exosome	16,33
SUT	CPF–CF and possibly NNS	Partially unstable	Rrp6, exosome, Xrn1 (NMD)	16,58
XUT	CPF–CF	Unstable	Xrn1 (NMD)	59,60
RUT	Reb1 roadblock	Unstable	TRAMP, Rrp6, exosome	54
Metazoan				
mRNA	CPSF–CF and SETX	Stable	None	4
snRNA	Integrator complex, CBC–ARS2, PCF11 and NELF	Stable (3′ end processed)	Exosome (3′ end processing)	89,90,92, 93,130
mRNAs encoding replication-dependent histones	CBC–ARS2	Stable	None	93–96
PROMPT	CPSF–CF and CBC–ARS2	Unstable	NEXT and exosome	93,109, 116–118

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.

can destabilize an elongation complex *in vitro* by simultaneous binding to the Pol II CTD and to the nascent RNA[23]. However, whether this actually occurs *in vivo* in the presence of the whole CPF–CF complex remains to be determined.

The torpedo model proposes that 3′ end cleavage of the precursor mRNA by the CPF–CF complex provides an entry point for a 5′–3′ exonuclease (Rat1 in yeast and XRN2 in humans), which degrades the nascent RNA up to the transcribing Pol II, leading to the dissociation of the elongation complex[24,25]. This would explain the coupling between cleavage and termination, and the occurrence of readthrough transcription when the 5′–3′ exonuclease is defective[24,25]. However, how degradation of the nascent RNA prompts destabilization of the elongation complex remains unclear. In fact, it has been shown that neither Rat1 (alone or in complex with its cofactor Rai1) nor degradation of the nascent RNA by itself is sufficient to elicit termination in a highly purified *in vitro* assay[26]. However, in another recent study[27] using a more complex system in which elongation complexes are only partially purified from whole-cell extracts, it was shown that Rat1–Rai1 can dismantle a stalled elongation complex, suggesting that additional factors besides Rat1–Rai1 are essential for this mode of termination. Importantly, an exonucleolytically inactive mutant of Rat1 that does not support termination *in vivo*[24] can dismantle the elongation complex *in vitro*, but only on addition of extra Rtt103, a CTD-interacting factor that recruits Rat1–Rai1 to the 3′ end of genes[24,28,29]. This suggests that degradation of the nascent RNA is not the primary event that determines termination but a means for Rat1 to 'catch up' with the elongating polymerase and bind to it to trigger termination, possibly by allosterically altering the conformation or the composition of the elongation complex. Direct recruitment of exonucleolytically defective Rat1 by Rtt103 to the CTD would enable bypass of this 'RNA' route to the elongation complex. Consistent with this, redirecting the cytoplasmic Rat1 paralogue Xrn1 to the nucleus results in the degradation of the nascent RNA but does not rescue the termination defect of a Rat1 mutant[28].

A unified allosteric–torpedo model has also been proposed on the basis of the mutual recruitment of Rat1 and Pcf11 at sites of cleavage and termination. In this scenario, a complex containing both Rat1 and CPF–CF would assemble at poly(A) sites, where it would mediate cleavage, nascent RNA degradation and termination through an allosteric change in the elongation complex[28].

Termination of ncRNAs: the NNS-dependent pathway. In addition to protein-coding genes, Pol II transcribes a plethora of ncRNAs. Some of these exert important cellular functions, such as snRNAs and small nucleolar RNAs (snoRNAs), but most are by-products of transcription that occurs in the wrong place or in the wrong direction as a result of 'leaky' transcription initiation[3].

In *S. cerevisiae*, the NNS complex is responsible for transcription termination at genes encoding snRNAs and snoRNAs[30] and at cryptic unstable transcripts (CUTs), which are an important class of transcription units that code for ncRNAs[31–34] (TABLE 1). These RNAs are degraded rapidly after transcription by the nuclear exosome and can generally only be detected when the nuclear exosome is not fully functional. As CUTs are widespread and overlap with functional genes in both the sense direction and the antisense direction, timely termination of their transcription is of paramount importance for preventing interference with the coding transcriptome (see below).

Cryptic unstable transcripts (CUTs). Yeast non-coding transcripts that become detectable only on inactivation of nuclear RNA decay pathways.

© 2015 Macmillan Publishers Limited. All rights reserved

The essential NNS complex contains two RNA-binding proteins, Nrd1 and Nab3, and the conserved superfamily I RNA and DNA helicase Sen1 (REFS 35,36) (FIG. 1b). 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[37–39]. These motifs are often clustered and associated with AU-rich sequences, which contribute to the efficiency of termination[38]. The NNS complex interacts with Ser5P CTD of Pol II through the CID of Nrd1 (REFS 40–43). The Ser5 phosphorylation mark is predominant early during transcription[44–46] (BOX 1), and this interaction pre-recruits the NNS complex to the elongation complex, favouring subsequent recognition of the specific motifs as they emerge from the transcribing polymerase[40,42]. 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 (see below).

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 two exosome nucleases, Rrp6 and Dis3, contribute to trimming the 3′ end of snRNA and snoRNA precursors to convert them into the mature species[16,47,48], and they completely degrade CUTs[16,31–33]. Efficient degradation and processing of these transcripts requires the multimeric TRAMP (Trf4–Air2–Mtr4) complex, which catalyses polyadenylation of the transcript and facilitates degradation by the exosome. The coupling between NNS-dependent termination and RNA degradation has been proposed to depend on physical interactions between NNS, TRAMP and the exosome[36]. However, molecular details of the mechanism by which termination and degradation are coordinated have been elucidated only recently. 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[42].

Although it is well established how the NNS complex identifies its targets, the actual mechanisms of termination and the precise role of each complex component are not completely understood. An important step forward in the comprehension of the mechanism was provided by a recent study showing that the Sen1 helicase alone can dissociate an elongation complex in a highly purified *in vitro* transcription termination system. Termination *in vitro* occurs preferentially at pause sites and requires both the interaction of Sen1 with the nascent RNA and ATP hydrolysis[49]. Together with *in vivo* experiments showing that the speed of transcription affects the position of NNS-dependent termination[50], these data support a model whereby Sen1 translocates on the nascent RNA to catch up with the transcribing polymerase to provoke transcription termination (FIG. 1b), akin to bacterial Rho-dependent termination (see Supplementary information S1 (box)).

Reb1-dependent unstable transcripts
(RUTs). Yeast unstable transcripts with 3′ ends that are determined by the collision of RNA polymerase II with DNA-bound Reb1.

Nucleosome-free regions
(NFRs). Chromatin regions that are devoid of nucleosomes. They are generally located at promoter and intergenic regions.

Nrd1 and Nab3 probably act upstream in the pathway by ensuring efficient and specific recruitment of Sen1, which is present at low levels in the cell (125 molecules per cell[51]) and recognizes RNA in a seemingly non-specific manner[18,49]. However, the possibility that Nrd1 and Nab3 have a more direct role in the release of the polymerase from its template cannot be excluded.

Rnt1- and Reb1-dependent termination. Although the CPF–CF and NNS pathways account for the generation of most coding RNAs and ncRNAs in yeast, two other mechanisms of termination have been described. The first depends on the Rnt1 endonuclease, the eukaryotic homologue of bacterial RNaseIII, which functions in the generation of mature snRNAs and snoRNAs. It was suggested that Rnt1 provides an entry point for the exonuclease Rat1 by cleaving nascent transcripts at defined sites, which triggers termination by the torpedo mechanism as described above[52,53]. This pathway is therefore a variation of CPF–CF-dependent termination, whereby recognition of the termination motif and RNA cleavage occur independently of CPF–CF. Rnt1-dependent termination was proposed to function as a fail-safe mechanism when termination by the CPF–CF pathway is leaky.

All known mechanisms of termination for Pol II-dependent transcription rely on the recognition of signals on the nascent RNA. A novel study demonstrates that Pol II termination can also occur via a roadblock mechanism in yeast, which is akin to the mechanism by which transcription of ribosomal DNA by Pol I is terminated[54]. It has been shown that Pol II pauses at sites occupied by the DNA-binding protein Reb1 and is released by a mechanism that involves its ubiquitylation and, most likely, its degradation by the proteasome, akin to the removal of polymerases stalled upstream of DNA lesions[55]. ncRNAs generated by this pathway are polyadenylated by TRAMP and degraded in the nucleus by the exosome; on the basis of their similarity to CUTs, they have been named Reb1-dependent unstable transcripts (RUTs) (TABLE 1). Termination by Reb1 was also shown to function as a fail-safe mechanism at CPF–CF terminators, in a similar way to Rnt1-dependent termination. Termination by Reb1 occurs at intergenic regions where Reb1 is required for the positioning of nucleosome-free regions (NFRs)[131]. This mode of termination is mechanistically different from the other known pathways in that it only requires the binding of a single protein to DNA to prevent progression of the enzyme. This pathway is 'disruptive' in the sense that every termination event probably leads to the destruction of a Pol II molecule.

A novel, provocative, double-torpedo model of termination has been proposed in fission yeast[56]. It was shown that depletion of the exosome catalytic subunit Dis3 or the core exosome component Rrp41 (but not Rrp6) induces a termination defect in approximately 30% of Pol II transcripts. When Pol II pauses, it backtracks and the 3′ end of the nascent RNA is displaced from the catalytic centre. It has been proposed that the exosome promotes termination by degrading the 3′ extension in backtracked Pol II, possibly

© 2015 Macmillan Publishers Limited. All rights reserved

complementing the concurrent 5′–3′ torpedo degradation by Rat1. Although it is debatable whether the 3′ RNA extension handle in backtracked Pol II is long enough to be targeted by the exosome, this model has the merit of explaining analogous observations made on the role of the exosome in termination in the *S. cerevisiae* system[16,57].

Termination of other ncRNAs. In addition to CUTs and RUTs, numerous, generally longer, ncRNAs have been described in yeast, such as stable unannotated transcripts (SUTs)[58] and Xrn1-dependent unstable transcripts (XUTs)[59]. These species are distinguished mainly by their stability and their turnover pathways (TABLE 1). Transcription termination by the CPF–CF pathway has been demonstrated for a few model SUTs[60], but it is likely that this pathway is also responsible for the termination of XUTs. SUTs can be detected in wild-type cells, but these RNAs are also subject to exosome degradation in the nucleus for reasons that remain unclear[16] (TABLE 1). To some extent, this contradicts the notion that CPF–CF-terminated transcripts are not targeted by nuclear degradation and might parallel the degradation of unstable promoter-proximal transcripts (PROMPTs) in humans (see below); alternatively, SUTs might be longer forms of CUTs that are not terminated efficiently[61]. In addition to nuclear degradation, a substantial fraction of SUTs are exported to the cytoplasm and degraded by the nonsense-mediated decay (NMD) pathway (A. Jacquier, personal communication), suggesting that the steady-state levels of these RNAs are determined by incomplete nuclear and cytoplasmic RNA quality control. XUTs are efficiently degraded in the cytoplasm by the Xrn1-dependent[59] and NMD pathways (A. Jacquier, personal communication) (TABLE 1), suggesting that both of these classes of ncRNAs take the mRNA route to escape nuclear degradation.

Functional relationships between the CPF–CF and the NNS pathways. The coexistence of many transcription termination pathways in a compact genome such as that of *S. cerevisiae*, as well as the different fate that each imposes on the transcripts produced, raises the question of their mutual relationships.

Specificity for the CPF–CF and the NNS pathways is determined by both the recognition of signals on the nascent RNA and the interaction of the CID modules of Pcf11 and Nrd1 with specific phosphorylated forms of the CTD. The phosphorylation pattern of the CTD (BOX 1) contributes to specify a positional parameter: NNS termination occurs preferentially early in transcription (within 1 kb of the transcription start site (TSS)), whereas most CPF–CF target genes are longer[38,62–64].

However, despite having generally well-differentiated functions and targets, the CPF–CF and the NNS pathways are much more interconnected than one would expect. Indeed, NNS terminators can efficiently be used by the CPF–CF pathway when located beyond the 1-kb positional threshold[38,62], which implies that the two protein complexes recognize substantially overlapping sequences. Consistent with this notion, termination at snoRNA genes is also dependent on CPF–CF components[9,65–68], and the NNS complex has been shown to be involved in the termination of short protein-coding genes[42,69]. Furthermore, Sen1 and Rat1, and their respective human homologues senataxin (SETX) and XRN2 (see below), have been proposed to cooperate for fully efficient termination at model protein-coding genes[53,70,71]. However, it should be noted that a recent high-resolution genome-wide analysis of Pol II distribution[20] did not detect termination defects at mRNA-coding genes on depletion of Sen1.

Further underscoring these functional connections, Sen1 was shown to interact with the Glc7 phosphatase, a component of the CPF–CF machinery that dephosphorylates the Tyr1P CTD and that is required for efficient termination at genes encoding both snoRNAs and mRNAs[35,72].

Transcription termination in metazoans

So far, three pathways of Pol II transcription termination have been described in metazoans, generating mRNAs, snRNAs and transcripts encoding replication-dependent histones, respectively.

General termination of mRNA-coding genes. Many components of the yeast CPF–CF complex are conserved in mammals and form similar complexes called cleavage and polyadenylation specificity factor (CPSF), cleavage stimulatory factor (CstF), cleavage factor I (CFI) and CFII (reviewed in REFS 4,5). As for yeast, the mechanism of transcription termination is not fully understood, and analogous models have been proposed (FIG. 2a).

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; however, it remains a matter of debate whether this occurs following cleavage and exonucleolytic degradation of the nascent RNA by the Rat1 homologue XRN2 (REF. 25) or by an allosteric mechanism that is possibly dependent on the CFII component PCF11 (REF. 73). A variation of the torpedo model for termination has been described for the human gene encoding β-globin and for roughly 80 additional genes[74], whereby the cleavage event triggering XRN2 entry occurs at a downstream co-transcriptional cleavage (CoTC) sequence. However, the molecular mechanism of CoTC cleavage and the factor or factors involved are unknown.

Electron microscopy visualization of Miller's chromatin spreading in *Xenopus laevis* and *Drosophila melanogaster*[75,76] revealed that, in most cases, transcribing polymerases are associated with full-length, uncleaved nascent transcripts up to the site of termination. Polymerases associated with cleaved nascent transcripts could also be visualized, but these were not systematically present in the termination region, as would be expected if cleavage were an absolute requirement for polymerase release. These findings would argue, at least in these experimental systems, against the torpedo model, although it cannot be excluded that after cleavage the elongation complex is released too quickly to be detected[76].

Stable unannotated transcripts
(SUTs). Yeast non-coding transcripts that are detected by transcriptome analyses in a wild-type background.

Xrn1-dependent unstable transcripts
(XUTs). Yeast transcripts that are stabilized on inactivation of the Xrn1 5′–3′ exonuclease.

Promoter-proximal transcripts
(PROMPTs). Human non-coding RNAs that are transcribed divergently to protein-coding genes. Generally, they can be detected only on exosome inactivation.

Nonsense-mediated decay
(NMD). A cytoplasmic RNA decay pathway that typically targets transcripts containing premature stop codons.

RNA quality control
Cellular mechanisms that identify and discard aberrant RNA molecules or ribonucleoprotein particles.

Miller's chromatin spreading
A technique for depositing dispersed chromatin that enables the visualization of RNA transcription and processing *in vivo* using electron microscopy.

© 2015 Macmillan Publishers Limited. All rights reserved

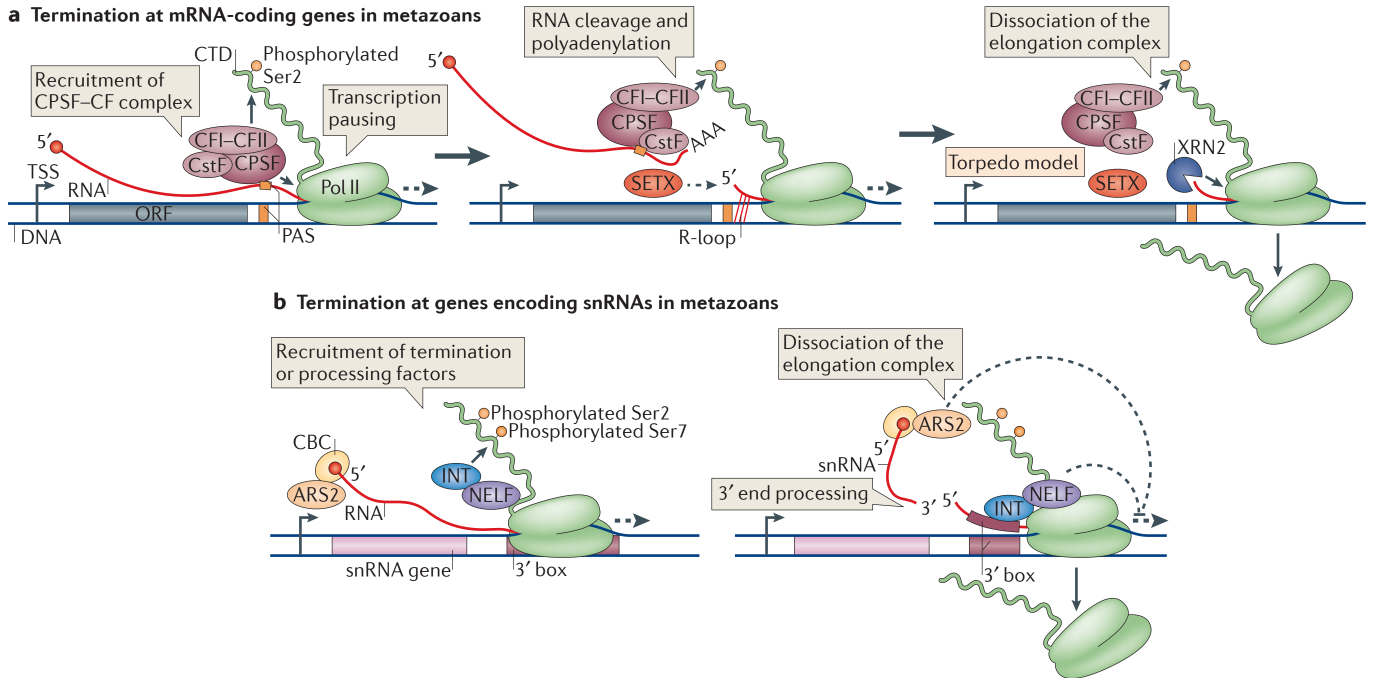


Figure 2 | **Transcription termination at mRNA-coding genes and small nuclear RNA genes in metazoans.** **a** | Transcription termination at protein-coding genes is triggered by cleavage and polyadenylation specificity factor (CPSF), cleavage stimulatory factor (CstF), cleavage factor I (CFI) and CFII, which contain homologues of components of the yeast cleavage and polyadenylation factor (CPF)–CF complex. Multiple interactions underlie the recruitment of the termination complex and the triggering of termination: CPSF directly binds to the body of the polymerase; CstF and CFI–CFII bind to the Ser2-phosphorylated form of the RNA polymerase II (Pol II) carboxy-terminal domain (CTD); and specific motifs (including the polyadenylation signal (PAS)) are recognized in the 3′ untranslated region of the nascent RNA by CPSF and CstF. The RNA is cleaved and polyadenylated at the 3′ end, which favours subsequent export to the cytoplasm and translation. Transcription pausing is thought to be required for termination. The homologue of yeast Sen1, senataxin (SETX), has been suggested to participate in termination of some mRNA genes, possibly by resolving R-loops to allow the entry of the 5′–3′ exoribonuclease XRN2, the homologue of Rat1. Degradation of the 3′ end fragment of the nascent transcript is thought to subsequently elicit transcription termination (the torpedo model). An allosteric model similar to the one shown in FIG. 1a has also been proposed for termination in metazoans, but for simplicity only the torpedo model is depicted. **b** | Termination at small nuclear RNA (snRNA) genes involves the integrator (INT), a large protein complex that contains homologues of CPSF subunits. This complex recognizes the Ser7-phosphorylated form of the Pol II CTD and a sequence at the 3′ end of the snRNA (3′ box). The negative elongation factor (NELF) is also recruited to the elongation complex, possibly as a result of interaction with INT. Concomitantly or subsequently to RNA 3′end cleavage by INT, transcription termination is elicited by unknown mechanisms that require INT, NELF, the cap-binding complex (CBC) and its associated factor ARS2. ORF, open reading frame; TSS, transcription start site.

Using an *in vitro* transcription termination approach, a recent study shows that the presence of a PAS is sufficient to trigger termination even under conditions in which cleavage does not occur. However, release of the polymerase in these conditions is delayed, suggesting that cleavage (presumably providing an entry site for XRN2) affects the efficiency and the timely occurrence of termination (H. Zhang and H. G. Martinson, personal communication).

Transcription termination in metazoans is thought to be associated with Pol II pausing[76–80], which might favour catching up of the elongating enzyme by XRN2 or might ensue from (or promote) allosteric changes that lead to termination. However, whether pausing is a general requirement for termination remains a subject of controversy in both mammals and yeast[18,79,81–83]. Mechanistically, it has been suggested that pausing is provoked by the simultaneous interaction of the CPSF complex with the PAS and the body of the polymerase (mediated by the CPSF30 (also known as CPSF4) subunit). This would precede entry of the CstF complex, which would bind the RNA and CPSF, displacing the latter from the body of the polymerase, and would also interact with the CTD, presumably in complex with PCF11 (REF. 80). Possibly consistent with these findings, it has been reported that deletion of the Ctk1 kinase in yeast, which is responsible for Ser2 CTD phosphorylation, leads to accumulation of Pol II signal at the 3′ end of several genes[46], possibly because Pol II pausing cannot be resolved by the interaction of the CPF–CF complex with the CTD. However, it should be noted that the lack of Ser2 CTD phosphorylation does not substantially affect termination of mRNAs[46,84]. Pausing at termination sites has also been correlated to the formation of RNA–DNA hybrids when the nascent transcript anneals to the template strand in the wake of transcribing (or paused)

© 2015 Macmillan Publishers Limited. All rights reserved

polymerase[71]. R-loops have been recently proposed to induce antisense transcription and the ensuing formation of repressive histone H3 Lys9 dimethylation (H3K9me2) marks by the RNAi pathway. Recognition of this mark by heterochromatin protein 1γ (HP1γ; also known as CBX3) would induce or stabilize pausing over termination sites[85]. Release of the polymerase at R-loop sites has been proposed to depend on the human Sen1 homologue SETX, which would unwind R-loops and grant access to XRN2 after nascent RNA cleavage[71]. A role for SETX in transcription termination of mRNAs has been proposed in several reports[71,86–88].

Genes encoding snRNAs and replication-dependent histones. Although most snoRNAs are encoded within introns in mammals, snRNAs are independently transcribed and undergo 3′ end processing and transcription termination by a specific pathway (FIG. 2b). No clear homologues or analogues of the Nrd1 and Nab3 components of the NNS complex have been described in metazoans. SETX is unlikely to be involved in transcription termination of genes encoding snRNAs because termination defects have not been seen in SETX-knockdown experiments[86,89]. Genes encoding snRNAs contain a conserved 13–16-nucleotide sequence element (termed the 3′ box) that is required both for 3′ end processing and for transcription termination. Processing and termination choices are made early for snRNA genes because recognition of the 3′ box only occurs *in vivo* when transcription is driven from an snRNA promoter. The 3′ box is recognized by a large complex, known as the integrator (INT) complex, and cleaved by its catalytic subunits INT9 and INT11, which are, respectively, homologues of the CPSF73 and CPSF100 subunits of the CPSF complex[90,91]. As the INT complex is recruited through its interaction with the Ser7P CTD and recognizes the nascent RNA, the pattern of concurrent recognition of signals on the nascent RNA and the CTD is conserved for snRNA 3′ processing and termination. However, termination is unlikely to be triggered by cleavage of the nascent transcript through an XRN2-dependent torpedo mechanism[89]. Rather, release of the polymerase has been linked to the particular structure of these genes, with a nucleosome-depleted region that spans the whole transcription unit, and to the action of negative elongation factor (NELF), which is involved in promoter-proximal pausing for mRNA transcription[89,92]. A role for the cap-binding complex (CBC) and the associated factor ARS2 (also known as SRRT) has also been described in the termination of genes encoding snRNAs, which possibly involves the CFII factors CLP1 and PCF11 (REFS 89,93). It has been proposed that the CBC–ARS2 complex specifically functions in the recognition of 3′ signals early in transcription, perhaps by mediating a link between the cap and the elongating polymerase. Although the details of its function remain unclear, this complex would be functionally analogous to the NNS complex, which also promotes early termination by sensing the distance from the TSS, presumably by recognizing the phosphorylation status of Pol II CTD[62].

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 processing[94–97]. However, the mechanism of termination remains obscure, and only depletion of ARS2 and CBC has been shown to induce transcriptional readthrough[93].

Regulating and protecting the transcriptome

Several reports over the past decade have highlighted the notion that transcription termination and RNA degradation are widely used for controlling gene expression and pervasive transcription. This has shifted the focus of gene expression regulation from the activation and initiation step to post-initiation events.

Termination and the regulation of gene expression. 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 — or to a mere constitutive attenuation of gene expression. Seemingly constitutive attenuation, or at least attenuation that has not been found to be regulated, has been shown for many genes, such as *HRP1*, *PCF11* (for which it depends on the NNS pathway[69]) and *HIS5* (for which it depends on the Reb1 roadblock pathway[54]). 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[31] (FIG. 3a). This is similar to the autoregulatory mechanism that has been reported in bacteria, whereby the transcription of the terminator protein Rho is controlled by Rho-dependent early termination[98]. Regulated attenuation has also been suggested to occur at the *FKS2* (also known as *GSC2*) locus, which encodes 1,3-β-D-glucan synthase, a protein involved in the synthesis of a major structural component of the cell wall. Recruitment of the NNS complex at this gene and the consequential early transcription termination can be prevented by the MAPK Mpk1 (also known as Slt2), leading to activation of *FKS2* under cell wall stress conditions[99] (FIG. 3b). In mammals, early transcription termination was reported to control HIV-1 provirus expression by the transactivator response (TAR) element. Transcription of TAR leads to the recruitment of Drosha, which cleaves the nascent RNA, thus triggering XRN2-dependent release of Pol II[87]. SETX and RRP6 are also involved in early termination, although their precise roles remain to be defined[87].

In some cases, early transcription termination is involved in the control of gene expression, but modulation of termination efficiency is not the primary regulatory event. Paradigmatic examples are genes involved in nucleotide biosynthesis in yeast[100,101]. For instance, in the

© 2015 Macmillan Publishers Limited. All rights reserved

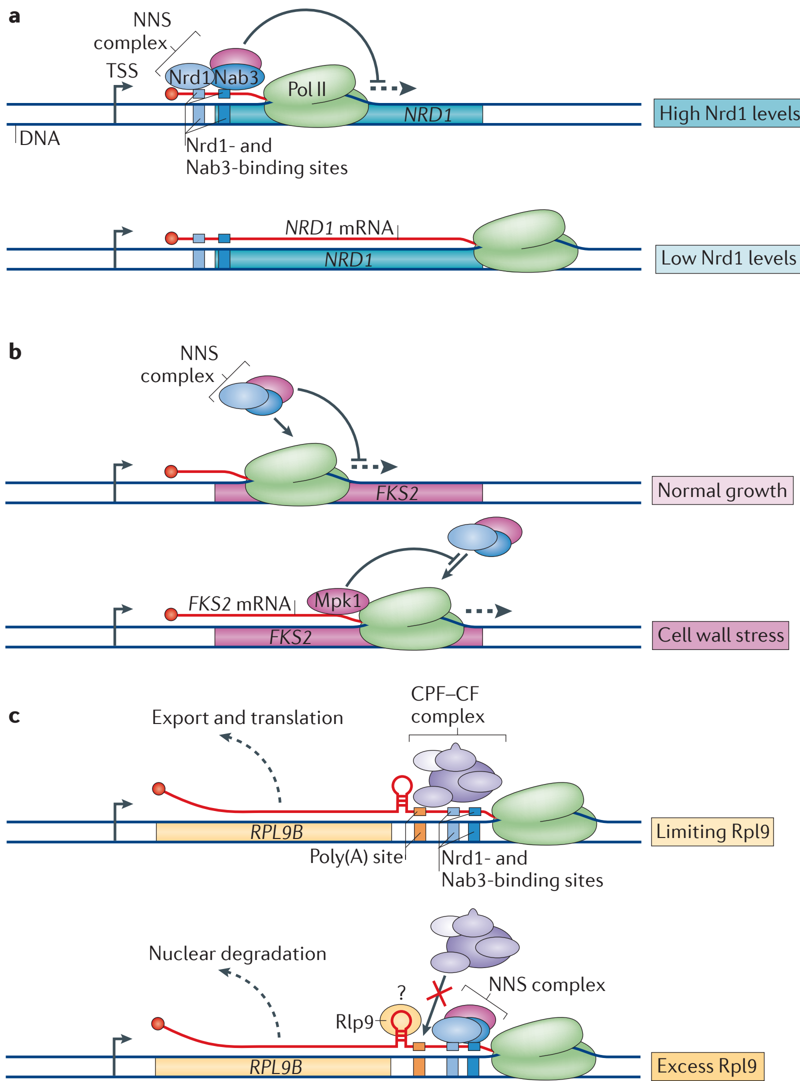


Figure 3 | **Regulation of gene expression by transcription termination in yeast.** **a** | Autoregulation of *NRD1* expression by premature termination is shown. Nrd1- and Nab3-binding sites are present in the 5′ untranslated region (UTR) and 5′ end of the *NRD1* gene. In a negative feedback loop, high Nrd1 levels promote early termination and RNA degradation by the Nrd1–Nab3–Sen1 (NNS) pathway, whereas low Nrd1 levels lead to the production of a full-length and functional *NRD1* transcript. **b** | Repression of *FKS2* expression is mediated by regulated attenuation. Under normal growth conditions, the NNS complex elicits early termination of *FKS2* transcription, which is constitutively activated. During cell wall stress, the MAPK Mpk1 prevents the recruitment of the NNS complex to the 5′ end of *FKS2* by an unknown mechanism, therefore enabling full induction. **c** | Feedback regulation of *RPL9B* expression is achieved by switching between alternative termination pathways. 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. TSS, transcription start site.

case of *URA2* and *IMD2*, transcription can initiate within two regions separated by a short sequence that induces termination by the NNS pathway. When transcription initiates at sites in the upstream region, this sequence is transcribed and early termination occurs, shutting off expression. Conversely, initiation downstream of the NNS terminator leads to full transcription and expression of the gene. In these examples, the bona fide regulatory event occurs at the level of TSS selection. However, NNS-dependent termination is required to prevent the production of a full-length RNA from the upstream initiation sites, which could be inappropriately translated.

Finally, an interesting example of autoregulation by termination has been described for the *RPL9B* gene in yeast, which encodes a ribosomal protein[102]. Transcription of this short gene can be terminated by either the CPF–CF pathway or the NNS pathway, but only the CPF–CF pathway yields a stable and functional RNA. Shifting between the two pathways is controlled by Rpl9 itself, which is proposed to bind to a conserved structure in the 3′ end of the gene, thus masking CPF–CF termination signals (FIG. 3c).

Shielding regulatory regions. Gene promoters are located in NFRs or nucleosome-depleted regions (NDRs), where transcription preferentially initiates. The position and extension of the NFR or NDR, as well as the characteristics of the flanking nucleosomes, are essential for efficient initiation. Transcription through these regions strongly inhibits the function of the local promoters, which is actually the basis for regulation at several loci[57,103–106]. It has been shown that transcription through NFRs or NDRs alters the epigenetic state of these regions and favours the assembly of nucleosomes[104], thus effectively preventing the access of transcriptional activators and the assembly of pre-initiation complexes (reviewed in REFS 3,107). As even low levels of transcription through these regions can induce transcriptional interference, it is important that NFRs or NDRs are shielded from RNA polymerases that might read through neighbouring termination signals. This is the most likely reason why enforcing regulatory region boundaries is essential, a function for which the various termination pathways often act redundantly. Yeast snoRNA genes are paradigmatic in this respect because these strongly transcribed units generally contain bipartite terminators, whereby NNS termination signals are followed by CPF–CF signals[66,64,84]. It has also been shown that cryptic CPF–CF termination is frequently revealed at genes encoding CUTs when NNS termination is impaired or is only partially effective[32–34]. Similarly, CPF–CF termination can be backed up by alternative pathways[53,52,54], which generally direct the RNAs that escape 'canonical' termination to nuclear degradation. Thus, in addition to preventing interference by readthrough transcription, these fail-safe pathways provide a nuclear 'filter' to eliminate readthrough transcripts. The extent to which these 'secondary' termination events might also contribute to the production of functional RNAs remains unclear[53,64,66].

Taming pervasive transcription. Genome-wide studies have unveiled the intrinsic bidirectional nature of many (if not all) promoters in yeast and humans, and presumably in many other species[58,79,108–110], with the notable exception of *D. melanogaster*, in which divergent transcription is rare[111]. Transcription of canonical protein-coding or functional RNA genes is often associated with the production

© 2015 Macmillan Publishers Limited. All rights reserved

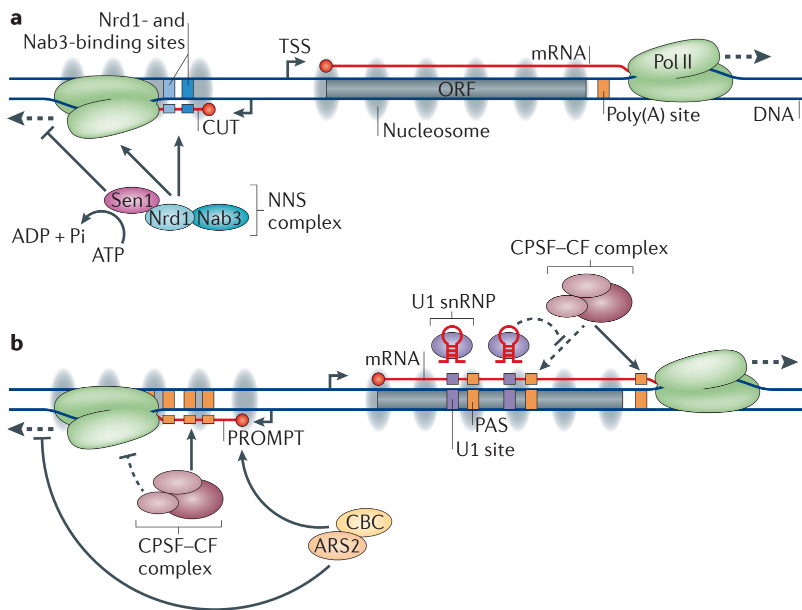


Figure 4 | **Transcription termination in the control of pervasive transcription throughout evolution.** **a** | Post-initiation control of bidirectional and pervasive transcription in yeast is mediated by the Nrd1–Nab3–Sen1 (NNS) pathway. Transcription events initiated spuriously from bidirectional promoters are preferentially terminated early by the NNS complex, which results in the elimination of non-functional transcripts at a post-initiation stage. The preferential presence of Nrd1- and Nab3-binding sites in the non-coding cryptic unstable transcript (CUT) induces the recruitment of the Nrd1–Nab3 complex and the Sen1 ATPase for transcription termination. **b** | Proposed mechanisms for limiting the intrinsic bidirectionality of promoters by early termination in metazoans are shown. Polyadenylation signals (PASs) are depleted in the direction of the coding gene (mRNA) and enriched in the divergent orientation (non-coding promoter-proximal transcript (PROMPT)) from a same promoter. In the coding orientation, the presence of U1 small nuclear ribonucleoprotein (snRNP)-binding sites near the rare, intragenic PAS prevents their recognition by the cleavage and polyadenylation specificity factor (CPSF)–cleavage factor (CF) complex. In the divergent orientation, CPSF–CF, together with cap-binding complex (CBC)–ARS2, promotes early transcription termination and subsequent degradation of the RNA. ORF, open reading frame; Pi, inorganic phosphate; TSS, transcription start site.

of a divergently initiated transcript, which may or may not have a functional role. 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[81]). 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 NFRs and NDRs, 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[81,112–115].

However, bidirectionality still substantially contributes to pervasive transcription and is a threat for the transcriptome. For instance, in yeast, CUT-producing transcription initiated divergently from the downstream gene of a tandem pair generally overlaps the 3′ region of the upstream gene and might potentially silence it if left uncontrolled[34,58,108]. Transcription termination is essential to suppress bidirectionality a posteriori: the crucial role of the NNS complex in this respect had earlier been postulated on the basis of the study of model genes[31,32] and was more recently extended to a genome-wide perspective[34]. Spurious transcription events are discriminated from functional ones after bidirectional initiation because NNS termination signals are more frequently present in the 'wrong' direction and are generally depleted from coding regions (FIG. 4a).

A similar functional pattern, although with different factors and mechanisms, has recently been described in human cells, which represents a major development in understanding the dynamics of the pervasive transcriptome in metazoans. PROMPTs are the equivalent of CUTs in humans. Generally originating divergently from canonical transcription units, these RNAs are shorter than normal transcripts[109] and are present at low steady-state levels because they are degraded by the human exosome[116] (FIG. 4b). As in yeast, transcription termination has a major role in the a posteriori selection of the correct one of two divergent transcripts: the CPSF–CF pathway recognizes the PASs that are present more frequently in the non-functional transcript and induces promoter proximal termination[109]. Interestingly, when present in the mRNA-coding direction, these termination signals are suppressed by the presence of antagonistic U1 snRNP-binding sites that have previously been shown to inhibit polyadenylation and termination[117]. The mechanistic reason that PROMPTs are unstable, despite their dependence on the CPSF–CF pathway, is unclear but might relate to the early position of the PASs relative to the TSS[109], possibly paralleling the position factor described above for NNS termination[62]. Interestingly, the mechanism of PROMPT termination is dependent on a set of factors that are similar to the ones directing transcription termination at genes encoding snRNAs and replication-dependent histones[93,118], which are short transcription units. This suggests the existence of a specific pathway for short transcription units that also operates to neutralize spurious transcription by early termination. It has been shown that CBC, together with ARS2, stimulates transcription termination at proximal sites, possibly by recruiting the CFII factor CLP1, while having poor effects at distally positioned PASs. Importantly, CBC–ARS2 is found to associate with the nuclear exosome targeting (NEXT) complex, which promotes PROMPT degradation. In light of these features, CBC–ARS2–NEXT can be considered as the functional analogue of the NNS–TRAMP complex in yeast. Finally, it has been suggested that promoter bidirectionality in mammals is also suppressed by de-capping and XRN2-promoted torpedo termination[119]. How this coordinates with the role of the CBC–ARS2 in promoting termination requires further study.

Conclusions and perspectives

Transcription termination has recently attracted much interest because of its implication in the control of pervasive transcription and gene expression. The notion that, in many cases, termination is crucial for removing non-functional RNAs underscores the fundamental importance of this step in gene expression.

© 2015 Macmillan Publishers Limited. All rights reserved

However, in spite of intense investigation, several questions remain open, and a detailed mechanistic understanding of how the elongation complex is destabilized is still missing. The action of the two main yeast termination pathways might converge towards a single (or analogous) mechanism for the polymerase release step. Chasing the polymerase with a helicase that translocates on the nascent RNA (Sen1 in the NNS pathway) or with a processive exonuclease degrading the Pol II-associated transcript (Rat1 in the CPF–CF pathway) might be functionally analogous events aimed at targeting the same 'button' on the polymerase (for example, an allosteric change) to shut off transcription. Alternatively, the NNS and the CPF–CF pathways might inactivate the elongation complex by completely different mechanisms, possibly linked to the specific catalytic activities of the two complexes.

Particularly exciting is the perspective of understanding the so far cryptic role of the CBC–ARS2 complex in suppressing bidirectionality of RNA production in mammals. Although probing the position of the mRNA cap could be an easy way to sense the length of the transcribed region (and therefore to recognize a prospective non-functional transcript), how this translates into early termination and degradation remains unexplained to a large extent. Addressing the mechanism or mechanisms of termination is a major and important future challenge that will require the implementation of appropriate *in vitro* component-controlled systems.

The use of termination to control gene expression is universally conserved from bacteria to humans. In these cases, transcription is started but attenuated or diverted to non-productive termination pathways. At first sight this might seem uneconomical — why initiate a process to interrupt it shortly after initiation? Perhaps controlling initiation is generally more expensive or simply slower than modulating a limiting post-initiation mechanism under constitutive activation.

If the efficiency of termination depends, to some extent, on transcription rates, attenuation might turn out to be an economical mechanism to fine-tune expression levels. For instance, it is possible that the levels of nascent transcripts might influence, positively or negatively, the availability of termination factors. Different scenarios can be imagined for a positive or negative dependency. For instance, high termination efficiency at low transcription rates (negative dependency) would contribute to a reduction of basal expression levels under non-activating conditions but would have little effect on expression under strong activation. Conversely, more efficient termination at higher transcription rates (positive dependency) could effectively limit maximal expression levels but would have less impact when the gene is expressed at lower levels. The prediction is that the occurrence of attenuation is not randomly distributed and instead reflects its mechanistic impact on expression.

Another conserved pattern is the genome-wide implication of transcription termination in the control of pervasive transcription. This might reflect a universal difficulty in establishing a robust control on initiation. The transcription machinery might have adapted to a default state that is inherently prone to initiation, and this could antagonistically limit the efficiency of preventing mis-initiation. Using post-initiation strategies to limit unwanted transcription might therefore be less demanding, particularly when exploiting pathways that have additional and essential functions (for example, Reb1 as a transcriptional activator). Importantly, it has been speculated[1–3,32] that the production of even short-lived ncRNAs as a result of pervasive transcription has an important evolutionary interest because it exposes additional information to the driving forces of natural selection, with the potential to generate new functional genes. This idea has been supported by a recent comparative genomics study showing that several protein-coding genes in humans correspond to loci expressing ncRNAs in primates[120]. The findings that the yeast NNS and CPF–CF pathways recognize largely overlapping termination signals and that termination of human PROMPTs is also CPSF–CF-dependent suggest that such a strong interconnection between termination pathways might facilitate the drift towards the evolution of new functions. Unstable ncRNAs, at least in yeast and mammals, are only a few evolutionary steps away from acquiring stability by shifting the termination pathway when their expression translates into a selective advantage for the cell.

Transcription is clearly not only a means to generate RNA but also a powerful tool to regulate gene expression, two functions that do not necessarily need to be coupled. Transcription termination associated with the disposal of the RNA produced might provide an evolutionary advantage by allowing regulatory transcription without the production of potentially harmful RNA by-products.

1. Jacquier, A. The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. *Nature Rev. Genet.* **10**, 833–844 (2009).
2. Colin, J., Libri, D. & Porrua, O. Cryptic transcription and early termination in the control of gene expression. *Genet. Res. Int.* **2011**, 653494 (2011).
3. Jensen, T. H., Jacquier, A. & Libri, D. Dealing with pervasive transcription. *Mol. Cell* **52**, 473–484 (2013).
A recent review on the generation and control of pervasive transcription in yeast and mammals.
4. Kuehner, J. N., Pearson, E. L. & Moore, C. Unravelling the means to an end: RNA polymerase II transcription termination. *Nature Rev. Mol. Cell. Biol.* **12**, 283–294 (2011).
5. Xiang, K., Tong, L. & Manley, J. L. Delineating the structural blueprint of the pre-mRNA 3′-end processing machinery. *Mol. Cell. Biol.* **34**, 1894–1910 (2014).
6. Richard, P. & Manley, J. L. Transcription termination by nuclear RNA polymerases. *Genes Dev.* **23**, 1247–1269 (2009).
7. Mischo, H. E. & Proudfoot, N. J. Disengaging polymerase: terminating RNA polymerase II transcription in budding yeast. *Biochim. Biophys. Acta* **1829**, 174–185 (2013).
8. Baejen, C. *et al.* Transcriptome maps of mRNP biogenesis factors define pre-mRNA recognition. *Mol. Cell* **55**, 745–757 (2014).
9. Pearson, E. & Moore, C. The evolutionarily conserved Pol II flap loop contributes to proper transcription termination on short yeast genes. *Cell Rep.* **9**, 821–828 (2014).
10. Hector, R. E. *et al.* Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. *EMBO J.* **21**, 1800–1810 (2002).
11. Dunn, E. F., Hammell, C. M., Hodge, C. A. & Cole, C. N. Yeast poly(A)-binding protein, Pab1, and PAN, a poly(A) nuclease complex recruited by Pab1, connect mRNA biogenesis to export. *Genes Dev.* **19**, 90–103 (2005).
12. Grenier St-Sauveur, V., Soucek, S., Corbett, A. H. & Bachand, F. Poly(A) tail-mediated gene regulation by opposing roles of Nab2 and Pab2 nuclear poly(A)-binding proteins in pre-mRNA decay. *Mol. Cell. Biol.* **33**, 4718–4731 (2013).

© 2015 Macmillan Publishers Limited. All rights reserved

13. Schmid, M. *et al.* Rrp6p controls mRNA poly(A) tail length and its decoration with poly(A) binding proteins. *Mol. Cell* **47**, 267–280 (2012).
14. Mouaikel, J. *et al.* High-frequency promoter firing links THO complex function to heavy chromatin formation. *Cell Rep.* **5**, 1082–1094 (2013).
15. Zenklusen, D., Larson, D. R. & Singer, R. H. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nature Struct. Mol. Biol.* **15**, 1263–1271 (2008).
16. Gudipati, R. K. *et al.* Extensive degradation of RNA precursors by the exosome in wild-type cells. *Mol. Cell* **48**, 409–421 (2012).
17. Tuck, A. C. & Tollervey, D. A transcriptome-wide atlas of RNP composition reveals diverse classes of mRNAs and lncRNAs. *Cell* **154**, 996–1009 (2013).
18. Creamer, T. J. *et al.* Transcriptome-wide binding sites for components of the *Saccharomyces cerevisiae* non-poly(A) termination pathway: Nrd1, Nab3, and Sen1. *PLoS Genet.* **7**, e1002329 (2011).
A genome-wide analysis of NNS components binding to the RNA by photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR–CLIP).
19. Sadowski, M., Dichtl, B., Hübner, W. & Keller, W. Independent functions of yeast Pcf11p in pre-mRNA 3′ end processing and in transcription termination. *EMBO J.* **22**, 2167–2177 (2003).
20. Schaughency, P., Merran, J. & Corden, J. L. Genome-wide mapping of yeast RNA polymerase II termination. *PLoS Genet.* **10**, e1004632 (2014).
Describes the high-resolution PAR-CLIP analysis of Pol II distribution in wild-type cells and in mutants of the NNS and CPF–CF pathways.
21. 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).
22. 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).
23. Zhang, Z., Fu, J. & Gilmour, D. S. CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3′-end processing factor, Pcf11. *Genes Dev.* **19**, 1572–1580 (2005).
Uses *in vitro* assays to show the dismantling of the elongation complex by purified Pcf11, arguing for an allosteric model for termination.
24. Kim, M. *et al.* The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* **432**, 517–522 (2004).
25. 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).
References 24 and 25 describe the role of Rat1 and XRN2 in termination (torpedo model).
26. Dengl, S. & Cramer, P. Torpedo nuclease Rat1 is insufficient to terminate RNA polymerase II *in vitro*. *J. Biol. Chem.* **284**, 21270–21279 (2009).
27. Pearson, E. L. & Moore, C. L. Dismantling promoter-driven RNA polymerase II transcription complexes *in vitro* by the termination factor Rat1. *J. Biol. Chem.* **288**, 19750–19759 (2013).
28. Luo, W., Johnson, A. W. & Bentley, D. L. The role of Rat1 in coupling mRNA 3′-end processing to transcription termination: implications for a unified allosteric-torpedo model. *Genes Dev.* **20**, 954–965 (2006).
Proposes a combined allosteric torpedo model for CPF–CF termination in yeast.
29. Lunde, B. M. *et al.* Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nature Struct. Mol. Biol.* **17**, 1195–1201 (2010).
30. Steinmetz, E. J., Conrad, N. K., Brow, D. A. & Corden, J. L. RNA-binding protein Nrd1 directs poly(A)-independent 3′-end formation of RNA polymerase II transcripts. *Nature* **413**, 327–331 (2001).
The first demonstration of the role of Nrd1 in termination of snoRNAs and snRNAs in yeast.
31. Arigo, J. T., Eyler, D. E., Carroll, K. L. & Corden, J. L. Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Mol. Cell* **23**, 841–851 (2006).
32. Thiebaut, M., Kisseleva-Romanova, E., Rougemaille, M., Boulay, J. & Libri, D. Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the Nrd1–Nab3 pathway in genome surveillance. *Mol. Cell* **23**, 853–864 (2006).
References 31 and 32 show that Nrd1 and Nab3 function in termination of CUTs in yeast.
33. Wyers, F. *et al.* Cryptic Pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121**, 725–737 (2005).
The first study describing CUTs and demonstrating the role of TRAMP and the exosome in their degradation.
34. Schulz, D. *et al.* Transcriptome surveillance by selective termination of noncoding RNA synthesis. *Cell* **155**, 1075–1087 (2013).
Extends to a genome-wide perspective the role of Nrd1 in termination of CUTs and the control of pervasive transcription.
35. Nedea, E. *et al.* The Glc7 phosphatase subunit of the cleavage and polyadenylation factor is essential for transcription termination on snoRNA genes. *Mol. Cell* **29**, 577–587 (2008).
36. Vasiljeva, L. & Buratowski, S. Nrd1 interacts with the nuclear exosome for 3′ processing of RNA polymerase II transcripts. *Mol. Cell* **21**, 239–248 (2006).
37. Carroll, K. L., Ghirlando, R., Ames, J. M. & Corden, J. L. Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements. *RNA* **13**, 361–373 (2007).
38. Porrua, O. *et al. In vivo* SELEX reveals novel sequence and structural determinants of Nrd1–Nab3–Sen1-dependent transcription termination. *EMBO J.* **31**, 3935–3948 (2012).
39. Wlotzka, W., Kudla, G., Granneman, S. & Tollervey, D. The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. *EMBO J.* **30**, 1790–1803 (2011).
40. Heo, D. *et al.* The RNA polymerase II C-terminal domain-interacting domain of yeast Nrd1 contributes to the choice of termination pathway and couples to RNA processing by the nuclear exosome. *J. Biol. Chem.* **288**, 36676–36690 (2013).
41. Kubicek, K. *et al.* Serine phosphorylation and proline isomerization in RNAP II CTD control recruitment of Nrd1. *Genes Dev.* **26**, 1891–1896 (2012).
42. Tudek, A. *et al.* Molecular basis for coordinating transcription termination with noncoding RNA degradation. *Mol. Cell* **55**, 467–481 (2014).
References 36 and 42 demonstrate the recruitment of TRAMP and the exosome by the NNS complex and the dual role of the CID of Nrd1.
43. Vasiljeva, L., Kim, M., Mutschler, H., Buratowski, S. & Meinhart, A. The Nrd1–Nab3–Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nature Struct. Mol. Biol.* **15**, 795–804 (2008).
44. Tietjen, J. R. *et al.* Chemical-genomic dissection of the CTD code. *Nature Struct. Mol. Biol.* **17**, 1154–1161 (2010).
45. Mayer, A. *et al.* Uniform transitions of the general RNA polymerase II transcription complex. *Nature Struct. Mol. Biol.* **17**, 1272–1278 (2011).
46. Kim, H. *et al.* Gene-specific RNA polymerase II phosphorylation and the CTD code. *Nature Struct. Mol. Biol.* **17**, 1279–1286 (2011).
47. Allmang, C. *et al.* Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J.* **18**, 5399–5410 (1999).
48. Van Hoof, A., Lennertz, P. & Parker, R. Yeast exosome mutants accumulate 3′-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. *Mol. Cell. Biol.* **20**, 441–452 (2000).
49. Porrua, O. & Libri, D. A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast. *Nature Struct. Mol. Biol.* **20**, 884–891 (2013).
Shows that purified Sen1 is sufficient to terminate transcription *in vitro*, for which it requires ATP hydrolysis and recognition of the nascent RNA.
50. Hazelbaker, D. Z., Marquardt, S., Wlotzka, W. & Buratowski, S. Kinetic competition between RNA polymerase II and Sen1-dependent transcription termination. *Mol. Cell* **49**, 55–66 (2012).
51. Ghaemmaghami, S. *et al.* Global analysis of protein expression in yeast. *Nature* **425**, 737–741 (2003).
52. Ghazal, G. *et al.* Yeast RNase III triggers polyadenylation-independent transcription termination. *Mol. Cell* **36**, 99–109 (2009).
53. Rondón, A. G., Mischo, H. E., Kawauchi, J. & Proudfoot, N. J. Fail-safe transcriptional termination for protein-coding genes in *S. cerevisiae*. *Mol. Cell* **36**, 88–98 (2009).
54. Colin, J. *et al.* Roadblock termination by Reb1p restricts cryptic and readthrough transcription. *Mol. Cell* **56**, 667–680 (2014).
Demonstrates the role of Reb1-dependent roadblock termination as a fail-safe pathway for termination and for controlling pervasive transcription.
55. Wilson, M. D., Harreman, M. & Svejstrup, J. Q. Ubiquitylation and degradation of elongating RNA polymerase II: the last resort. *Biochim. Biophys. Acta* **1829**, 151–157 (2013).
56. Lemay, J.-F. *et al.* The RNA exosome promotes transcription termination of backtracked RNA polymerase II. *Nature Struct. Mol. Biol.* **21**, 919–926 (2014).
57. Castelnuovo, M. *et al.* Bimodal expression of PHO84 is modulated by early termination of antisense transcription. *Nature Struct. Mol. Biol.* **20**, 851–858 (2013).
58. Xu, Z. *et al.* Bidirectional promoters generate pervasive transcription in yeast. *Nature* **457**, 1033–1037 (2009).
59. Van Dijk, E. L. *et al.* XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature* **475**, 114–117 (2011).
The first study showing the existence of XUTs, which are suggested to have a role in gene regulation.
60. Marquardt, S., Hazelbaker, D. Z. & Buratowski, S. Distinct RNA degradation pathways and 3′ extensions of yeast non-coding RNA species. *Transcription* **2**, 145–154 (2011).
61. Castelnuovo, M. *et al.* Role of histone modifications and early termination in pervasive transcription and antisense-mediated gene silencing in yeast. *Nucleic Acids Res.* **42**, 4348–4362 (2014).
62. Gudipati, R. K., Villa, T., Boulay, J. & Libri, D. Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice. *Nature Struct. Mol. Biol.* **15**, 786–794 (2008).
63. Jenks, M. H., O'Rourke, T. W. & Reines, D. Properties of an intergenic terminator and start site switch that regulate *IMD2* transcription in yeast. *Mol. Cell. Biol.* **28**, 3883–3893 (2008).
64. Steinmetz, E. J., Ng, S. B. H., Cloute, J. P. & Brow, D. A. *Cis-* and *trans*-acting determinants of transcription termination by yeast RNA polymerase II. *Mol. Cell. Biol.* **26**, 2688–2696 (2006).
65. Dheur, S. *et al.* Pti1p and Ref2p found in association with the mRNA 3′ end formation complex direct snoRNA maturation. *EMBO J.* **22**, 2831–2840 (2003).
66. Kim, M. *et al.* Distinct pathways for snoRNA and mRNA termination. *Mol. Cell* **24**, 723–734 (2006).
67. Morlando, M. *et al.* Coupling between snoRNP assembly and 3′ processing controls box C/D snoRNA biosynthesis in yeast. *EMBO J.* **23**, 2392–2401 (2004).
68. Fatica, A., Morlando, M. & Bozzoni, I. Yeast snoRNA accumulation relies on a cleavage-dependent/polyadenylation-independent 3′-processing apparatus. *EMBO J.* **19**, 6218–6229 (2000).
69. Steinmetz, E. J. *et al.* Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol. Cell* **24**, 735–746 (2006).
70. Kawauchi, J., Mischo, H., Braglia, P., Rondon, A. & Proudfoot, N. J. Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. *Genes Dev.* **22**, 1082–1092 (2008).
71. Skourti-Stathaki, K., Proudfoot, N. J. & Gromak, N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol. Cell* **42**, 794–805 (2011).
Proposes a role for SETX in the termination of transcription in mammals.
72. Schreieck, A. *et al.* RNA polymerase II termination involves C-terminal-domain tyrosine dephosphorylation by CPF subunit Glc7. *Nature Struct. Mol. Biol.* **21**, 175–179 (2014).
73. Zhang, Z. & Gilmour, D. S. Pcf11 is a termination factor in *Drosophila* that dismantles the elongation complex by bridging the CTD of RNA polymerase II to the nascent transcript. *Mol. Cell* **21**, 65–74 (2006).
74. Nojima, T., Dienstbier, M., Murphy, S., Proudfoot, N. J. & Dye, M. J. Definition of RNA polymerase II CoTC terminator elements in the human genome. *Cell Rep.* **3**, 1080–1092 (2013).
75. Osheim, Y. N., Proudfoot, N. J. & Beyer, A. L. EM visualization of transcription by RNA polymerase II: downstream termination requires a poly(A) signal but not transcript cleavage. *Mol. Cell* **3**, 379–387 (1999).

© 2015 Macmillan Publishers Limited. All rights reserved

76. Osheim, Y. N., Sikes, M. L. & Beyer, A. L. EM visualization of Pol II genes in *Drosophila*: most genes terminate without prior 3′ end cleavage of nascent transcripts. *Chromosoma* **111**, 1–12 (2002).
In references 75 and 76, individual transcription units are visualized by electron microscopy. Few polymerases were found to transcribe at the 3′ end of genes associated with short nascent RNAs, which provides evidence against the occurrence of cleavage before (and as a requirement for) termination.
77. Gromak, N., West, S. & Proudfoot, N. J. Pause sites promote transcriptional termination of mammalian RNA polymerase II. *Mol. Cell. Biol.* **26**, 3986–3996 (2006).
78. Plant, K. E., Dye, M. J., Lafaille, C. & Proudfoot, N. J. Strong polyadenylation and weak pausing combine to cause efficient termination of transcription in the human $^{G}\gamma$-globin gene. *Mol. Cell. Biol.* **25**, 3276–3285 (2005).
79. 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).
80. Nag, A., Narsinh, K. & Martinson, H. G. The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase. *Nature Struct. Mol. Biol.* **14**, 662–669 (2007).
81. Churchman, L. S. & Weissman, J. S. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* **469**, 368–373 (2011).
82. Glover-Cutter, K., Kim, S., Espinosa, J. & Bentley, D. L. RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. *Nature Struct. Mol. Biol.* **15**, 71–78 (2008).
83. 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).
84. Lenstra, T. L. *et al.* The role of Ctk1 kinase in termination of small non-coding RNAs. *PLoS ONE* **8**, e80495 (2013).
85. Skourti-Stathaki, K., Kamieniarz-Gdula, K. & Proudfoot, N. J. R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature* **516**, 436–439 (2014).
86. Suraweera, A. *et al.* Functional role for senataxin, defective in ataxia oculomotor apraxia type 2, in transcriptional regulation. *Hum. Mol. Genet.* **18**, 3384–3396 (2009).
87. Wagschal, A. *et al.* Microprocessor, Setx, Xrn2, and Rrp6 co-operate to induce premature termination of transcription by RNAPII. *Cell* **150**, 1147–1157 (2012).
88. Banerjee, A., Sammarco, M. C., Ditch, S., Wang, J. & Grabczyk, E. A. Novel tandem reporter quantifies RNA polymerase II termination in mammalian cells. *PLoS ONE* **4**, e6193 (2009).
89. O'Reilly, D. *et al.* Human snRNA genes use polyadenylation factors to promote efficient transcription termination. *Nucleic Acids Res.* **42**, 264–275 (2014).
90. Baillat, D. *et al.* Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* **123**, 265–276 (2005).
Describes the INT complex and its role in snRNA processing in mammals.
91. Ezzeddine, N. *et al.* A subset of *Drosophila* integrator proteins is essential for efficient U7 snRNA and spliceosomal snRNA 3′-end formation. *Mol. Cell. Biol.* **31**, 328–341 (2011).
92. Yamamoto, J. *et al.* DSIF and NELF interact with Integrator to specify the correct post-transcriptional fate of snRNA genes. *Nature Commun.* **5**, 4263 (2014).
93. Hallais, M. *et al.* CBC–ARS2 stimulates 3′-end maturation of multiple RNA families and favors cap-proximal processing. *Nature Struct. Mol. Biol.* **20**, 1358–1366 (2013).
Describes the role of CBC–ARS2 in the termination of PROMPTs.
94. Narita, T. *et al.* NELF interacts with CBC and participates in 3′ end processing of replication-dependent histone mRNAs. *Mol. Cell* **26**, 349–365 (2007).
95. Gruber, J. J. *et al.* Ars2 promotes proper replication-dependent histone mRNA 3′ end formation. *Mol. Cell* **45**, 87–98 (2012).
96. Yang, X.-C. *et al.* A complex containing the CPSF73 endonuclease and other polyadenylation factors associates with U7 snRNP and is recruited to histone pre-mRNA for 3′-end processing. *Mol. Cell. Biol.* **33**, 28–37 (2013).
97. Dominski, Z., Yang, X. & Marzluff, W. F. The polyadenylation factor CPSF-73 is involved in histone-pre-mRNA processing. *Cell* **123**, 37–48 (2005).
98. Matsumoto, Y., Shigesada, K., Hirano, M. & Imai, M. Autogenous regulation of the gene for transcription termination factor rho in *Escherichia coli*: localization and function of its attenuators. *J. Bacteriol.* **166**, 945–958 (1986).
99. Kim, K.-Y. & Levin, D. E. Mpk1 MAPK association with the Paf1 complex blocks Sen1-mediated premature transcription termination. *Cell* **144**, 745–756 (2011).
100. Kuehner, J. N. & Brow, D. A. Regulation of a eukaryotic gene by GTP-dependent start site selection and transcription attenuation. *Mol. Cell* **31**, 201–211 (2008).
101. Thiebaut, M. *et al.* Futile cycle of transcription initiation and termination modulates the response to nucleotide shortage in *S. cerevisiae*. *Mol. Cell* **31**, 671–682 (2008).
References 100 and 101 demonstrate the role of the NNS termination pathway in the regulation of genes involved in nucleotide metabolism.
102. Gudipati, R. K., Neil, H., Feuerbach, F., Malabat, C. & Jacquier, A. The yeast *Rpl9b* gene is regulated by modulation between two modes of transcription termination. *EMBO J.* **31**, 2427–2437 (2012).
103. Bumgarner, S. L. *et al.* Single-cell analysis reveals that noncoding RNAs contribute to clonal heterogeneity by modulating transcription factor recruitment. *Mol. Cell* **45**, 470–482 (2012).
104. Hainer, S. J., Pruneski, J. A., Mitchell, R. D., Monteverde, R. M. & Martens, J. A. Intergenic transcription causes repression by directing nucleosome assembly. *Genes Dev.* **25**, 29–40 (2011).
105. Martens, J. A., Wu, P.-Y. J. & Winston, F. Regulation of an intergenic transcript controls adjacent gene transcription in *Saccharomyces cerevisiae*. *Genes Dev.* **19**, 2695–2704 (2005).
106. Van Werven, F. J. *et al.* Transcription of two long noncoding RNAs mediates mating-type control of gametogenesis in budding yeast. *Cell* **150**, 1170–1181 (2012).
107. Smolle, M. & Workman, J. L. Transcription-associated histone modifications and cryptic transcription. *Biochim. Biophys. Acta* **1829**, 84–97 (2013).
108. Neil, H. *et al.* Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* **457**, 1038–1042 (2009).
References 58 and 108 describe the first high-resolution analyses of cryptic transcription in yeast.
109. Ntini, E. *et al.* Polyadenylation site-induced decay of upstream transcripts enforces promoter directionality. *Nature Struct. Mol. Biol.* **20**, 923–928 (2013).
110. Seila, A. C. *et al.* Divergent transcription from active promoters. *Science* **322**, 1849–1851 (2008).
111. Core, L. J. *et al.* Defining the status of RNA polymerase at promoters. *Cell Rep.* **2**, 1025–1035 (2012).
112. Tan-Wong, S. M. *et al.* Gene loops enhance transcriptional directionality. *Science* **338**, 671–675 (2012).
113. Whitehouse, I., Rando, O. J., Delrow, J. & Tsukiyama, T. Chromatin remodelling at promoters suppresses antisense transcription. *Nature* **450**, 1031–1035 (2007).
114. Yadon, A. N. *et al.* Chromatin remodeling around nucleosome-free regions leads to repression of noncoding RNA transcription. *Mol. Cell. Biol.* **30**, 5110–5122 (2010).
115. Marquardt, S. *et al.* A chromatin-based mechanism for limiting divergent noncoding transcription. *Cell* **157**, 1712–1723 (2014).
116. Preker, P. *et al.* RNA exosome depletion reveals transcription upstream of active human promoters. *Science* **322**, 1851–1854 (2008).
Provides the first description of PROMPTs. Together with references 79 and 110, it describes the occurrence of cryptic transcription in mammals.
117. Almada, A. E., Wu, X., Kriz, A. J., Burge, C. B. & Sharp, P. A. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature* **499**, 360–363 (2013).
118. Andersen, P. R. *et al.* The human cap-binding complex is functionally connected to the nuclear RNA exosome. *Nature Struct. Mol. Biol.* **20**, 1367–1376 (2013).
Describes the role of the CBC complex in terminating transcription of PROMPTs. Together with references 117, it describes the role of PAS signals and U1 snRNP in controlling promoter directionality.
119. Brannan, K. *et al.* mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. *Mol. Cell* **46**, 311–324 (2012).
120. Wu, D.-D., Irwin, D. M. & Zhang, Y.-P. *De novo* origin of human protein-coding genes. *PLoS Genet.* **7**, e1002379 (2011).
121. Cramer, P. *et al.* Structure of eukaryotic RNA polymerases. *Annu. Rev. Biophys.* **37**, 337–352 (2008).
122. Buratowski, S. Progression through the RNA polymerase II CTD cycle. *Mol. Cell* **36**, 541–546 (2009).
123. Hsin, J.-P. & Manley, J. L. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev.* **26**, 2119–2137 (2012).
124. Eick, D. & Geyer, M. The RNA polymerase II carboxy-terminal domain (CTD) code. *Chem. Rev.* **113**, 8456–8490 (2013).
A comprehensive review of the modifications and role of the CTD.
125. Bataille, A. R. *et al.* A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. *Mol. Cell* **45**, 158–170 (2012).
126. Hsin, J.-P., Sheth, A. & Manley, J. L. RNAPII CTD phosphorylated on threonine-4 is required for histone mRNA 3′ end processing. *Science* **334**, 683–686 (2011).
127. Mayer, A. *et al.* CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II. *Science* **336**, 1723–1725 (2012).
Describes the role of CTD Tyr1 phosphorylation in transcription termination.
128. Singh, N. *et al.* The Ess1 prolyl isomerase is required for transcription termination of small noncoding RNAs via the Nrd1 pathway. *Mol. Cell* **36**, 255–266 (2009).
129. Garland, W., Feigenbutz, M., Turner, M. & Mitchell, P. Rrp47 functions in RNA surveillance and stable RNA processing when divorced from the exoribonuclease and exosome-binding domains of Rrp6. *RNA* **19**, 1659–1668 (2013).
130. Egloff, S., O'Reilly, D. & Murphy, S. Expression of human snRNA genes from beginning to end. *Biochem. Soc. Trans.* **36**, 590–594 (2008).
131. Hartley, P. D.& Madhani, H.D. Mechanisms that specify promoter nucleosome location and identity. *Cell* **137**, 445–458 (2009).

Acknowledgements
The authors thank F. Stutz, N. Proudfoot and T. H. Jensen for critical reading of this manuscript, and F. Bachand, A. Jacquier, H. G. Martinson, N. Proudfoot and H. Zhang for providing access to unpublished data. They apologize to the many colleagues whose work could not be cited owing to a lack of space. Work in the authors' laboratory is supported by the Centre National de la Recherche Scientifique, the Danish National Research Foundation, the Agence Nationale pour la Recherche (ANR-08-Blan-0038-01 and ANR-12-BSV8-0014-01), the Fondation pour la Recherche Médicale (FRM; programme Equipes 2013) and the Fondation Bettencourt-Schueller (prix coup d'elan 2009 to D.L.). O.P. was supported by fellowships from the European Molecular Biology Organization and the FRM.

Competing interests statement
The authors declare no competing interests.

SUPPLEMENTARY INFORMATION
See online article: S1 (box)
ALL LINKS ARE ACTIVE IN THE ONLINE PDF

© 2015 Macmillan Publishers Limited. All rights reserved