Title
A global function for transcription factors in assisting RNA polymerase II termination

Journal
Transcription, 9(1)

ISSN
2154-1264

Authors
Roy, Kevin
Chanfreau, Guillaume F

Publication Date
2018

DOI
10.1080/21541264.2017.1300121

Peer reviewed

TRANSCRIPTION
2018, VOL. 9, NO. 1, 41–46
http://dx.doi.org/10.1080/21541264.2017.1300121

Taylor & Francis
Taylor & Francis Group

POINT-OF-VIEW

A global function for transcription factors in assisting RNA polymerase II termination

Kevin Roy [†] and Guillaume F. Chanfreau

Department of Chemistry and Biochemistry and the Molecular Biology Institute, University of California Los Angeles, CA, USA

ABSTRACT
The role of transcription factors (TFs) on nucleosome positioning, RNA polymerase recruitment, and transcription initiation has been extensively characterized. Here, we propose that a subset of TFs such as Reb1, Abf1, Rap1, and TFIIIB also serve a major function in partitioning transcription units by assisting the Nrd1p-Nab3p-Sen1p Pol II termination pathway.

ARTICLE HISTORY
Received 2 February 2017
Revised 21 February 2017
Accepted 22 February 2017

KEYWORDS
exosome; GRFs; Nab3; Nrd1; Sen1; Reb1; TFIIIB; Transcription termination

Introduction: Pervasive transcription and the role of the NNS termination pathway

A large fraction of eukaryotic genomes is transcribed by RNA polymerase II, yet the large majority of these transcripts do not encode any known functional proteins and likely have no functional role. The pervasive nature of Pol II transcription appears to be a consequence of the intrinsic ability of Pol II to initiate transcription at nucleosome-depleted regions (NDRs). As NDRs are found throughout the genome, eukaryotes have evolved means of limiting the impact of pervasive transcription. For instance, specific mechanisms have emerged for early termination of transcription or to ensure the rapid degradation of these apparently non-functional RNAs. Most of the mechanisms identified to date involve the recognition of signals in the nascent RNA by *trans*-acting RNA-binding proteins. In metazoan genes, cryptic polyadenylation signals are frequently found upstream of protein-coding genes and couple Pol II transcription termination to degradation of nascent transcripts termed PROMoter uPstream Transcripts (PROMPTs).[1] In fungi, the Nrd1p-Nab3p-Sen1p (NNS) termination pathway[2] is responsible for taming the majority of pervasive transcription, and also is responsible for termination of stable ncRNAs.[3] Following termination by the NNS pathway, the RNA 3′-ends are oligoadenylated by the TRAMP complex, and trimmed or fully degraded by the nuclear RNA exosome. Because the signals recognized by the NNS machinery involve tetrameric motifs with some degree of degeneracy,[4] virtually all protein-coding transcripts contain multiple NNS motifs, and therefore it has not been completely clear how the specificity of this pathway is achieved *in vivo*.

Known functions of general regulatory factors (GRFs)

What specifies the arrangement of nucleosomes on DNA? While some degree of nucleosome positioning can be explained by the intrinsic DNA sequence preference for nucleosomes,[5] a group of DNA-binding proteins termed general regulatory factors (GRFs) exert a broad influence on chromatin structure.[6] Comparison of nucleosome occupancy *in vivo* to genomic DNA packaged with reconstituted histones *in vitro* (in the absence of GRFs) revealed significant differences around GRF binding sites genome-wide, with depletion of nucleosomes at GRF sites *in vivo* and a phasing of the adjacent nucleosomes into well-positioned arrays.[5] GRFs, defined by their ability to outcompete nucleosomes for DNA binding, their essentiality, and by an abundance of their binding sites in genomes, are

CONTACT Guillaume F. Chanfreau guillom@chem.ucla.edu Chemistry and Biochemistry, Box 951569, University of California Los Angeles, CA 90095-1569, USA.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/ktrn.

[†]Present address: Department of Genetics, Stanford University.

also responsible for partitioning chromosomes into distinct transcriptional units by acting as insulators.[6] In budding yeast, three major GRFs (Abf1p, Reb1p, and Rap1p) provide key functions in promoting chromatin structure by establishing nucleosome free regions.[7,8,9] GRFs also play a major role in defining the functional architecture of specific promoters, and their role in structuring chromatin can be recapitulated *in vitro* in a purified system.[9] Interestingly, GRFs exhibit little intrinsic ability to activate Pol II transcription on their own, but instead function to amplify the activating effect of proximally-bound activator proteins, explaining how the same GRF can activate some genes while repressing others.[10] In addition to promoting the formation of nucleosome-free regions for transcriptional activation, these proteins moonlight in other DNA-based transactions. For example, Abf1p and Rap1p both function in DNA repair, replication, and transcriptional silencing.[11,12,13] Originally characterized for its role in RNA polymerase I transcription of the ribosomal RNAs,[14] Reb1p was also recently shown to promote roadblock termination of Pol II in a mechanism that was suggested to use ubiquitylation of the polymerase.[15]

The NNS termination pathway uses GRFs and TFIIIB to roadblock pol II and promote pol II release from the chromatin

Sequencing of the 3′-ends of RNAs that accumulate in strains inactivated for components of the nuclear exosome allowed us to identify at the nucleotide level sites of NNS termination genome-wide in *Saccharomyces cerevisiae*.[16] Using this strategy, we found that sites of NNS termination can be divided into two classes: those showing a diffuse pattern, with 3′-ends scattered throughout a region spanning dozens to hundreds of nucleotides, and those exhibiting punctuated peaks, with 3′-ends accumulating within a very narrow window of several nucleotides, suggesting that Pol II terminates at well-defined locations for these loci (Fig. 1). Analysis of genomic elements located near these sites showed that they corresponded precisely to the distance between the active site of Pol II and the boundaries of binding sites for the GRFs Reb1p, Rap1p, and Abf1p, as well as the Pol III transcription factor TFIIIB.[16] These observations led to the idea that these proteins can physically block Pol II progression and assist the NNS pathway in promoting Pol II termination, extending the roadblocking model initially proposed for Reb1p.[15] The observation that TFIIIB can physically roadblock Pol II also provided a mechanistic model to support previous observations that inactivating TFIIIB or Pol III transcription led to an increase of intergenic transcription by Pol II.[17,18]

High-resolution analysis of the sites of NNS termination relied on the sequencing of polyadenylated RNA 3′-ends in strains inactivated for exosome components. Thus, it could be argued that RNA 3′-ends accumulating in exosome mutants may not correspond to *bona fide* sites of NNS termination, especially considering that some of the exosome components may be implicated in termination.[19] To confirm that the 3′-ends stabilized in exosome mutants correspond to sites of Pol II roadblock termination, we analyzed nascent elongating transcript (NET)-seq 3′-ends obtained after inactivation of the TFIIS factor Dst1, which marks the site of pausing of Pol II.[20] Upon encountering obstacles in the template, Pol II pauses and can backtrack several nucleotides, leaving a protruded 3′ end and repositioning the active site of Pol II over an upstream phosphodiester bond. TFIIS then stimulates the intrinsic hydrolytic activity of Pol II on this phosphodiester bond and thus inactivating TFIIS reveals the location of the initial pause site.[21] These pause sites mapped precisely to the 3′-ends stabilized in exosome mutants, indicating that paused rather than backtracked Pol II is the preferred state of NNS-mediated release. Finally, the sites of NNS termination mapped by 3′-end sequencing globally matched with those identified in previous studies using Photoactivatable Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP) and Pol II ChIP.[22,23] Thus, inactivating the nuclear exosome followed by 3′-end sequencing offers a valuable strategy to identify sites of NNS termination genome-wide and provides a higher-resolution view of the actual termination sites.

Interestingly, the NNS pathway and transcription factors (TFs) that roadblock Pol II seemed to play a synergistic function to promote Pol II termination. For instance, removing the binding site for proteins that roadblock Pol II led to a switch of the termination pattern from punctuate to diffuse; reciprocally, inactivating the NNS pathway resulted in Pol II reading through the sites of GRF binding and in the use of downstream polyA signals to promote termination. This functional synergy between the NNS pathway

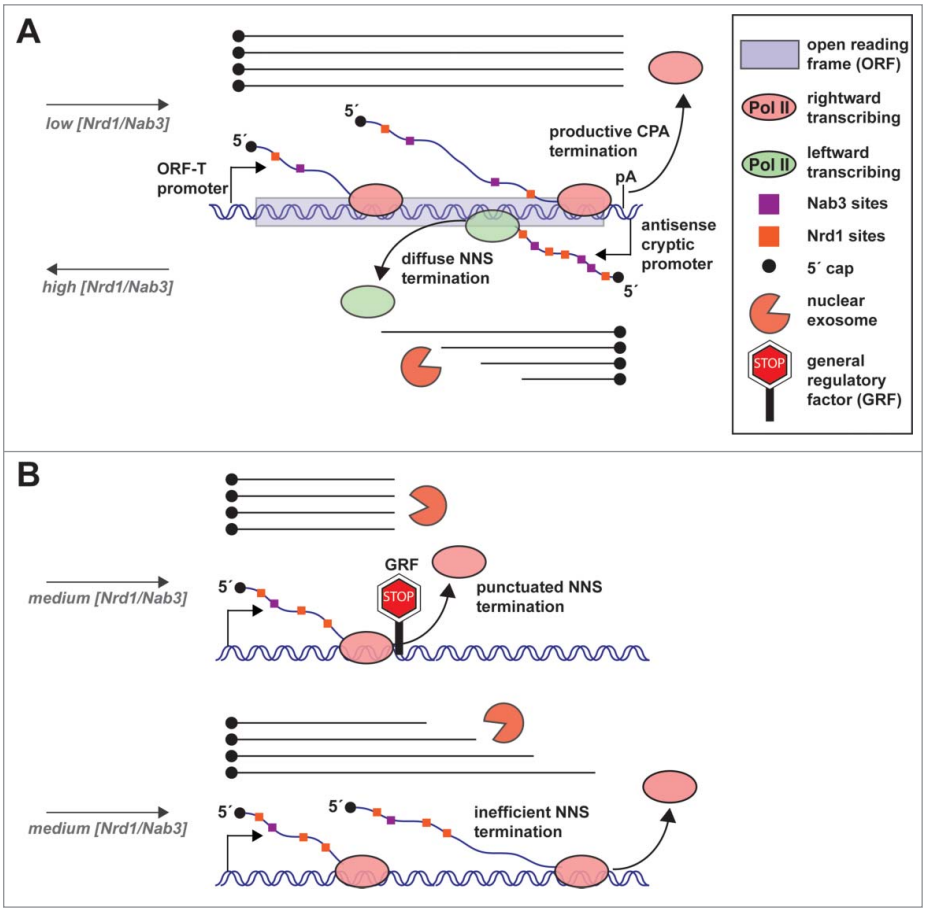

Figure 1. The role of the Nrd1-Nab3-Sen1 (NNS) pathway in regulating the transcriptome. (A) Overall, ORF transcripts (ORF-Ts) in yeast are statistically under-enriched for NNS motifs, in particular the extended Nab3 motif UCUUG. The presence of NNS motifs on the antisense strand promotes early termination of antisense cryptic transcription (green Pol II) to minimize transcriptional interference on the ORF-T promoter. NNS-terminated transcripts are degraded by the combined action of TRAMP and the nuclear exosome (orange Pac-Man). (B) Medium and high levels of NNS motifs promote the release of Pol II paused at general regulatory factor (GRF) binding sites by roadblock-dependent, punctuated NNS termination (top panel). In the absence of GRFs, medium levels of NNS signals result in inefficient and diffuse termination by the NNS pathway (bottom panel).

and roadblock-promoting TFs was also revealed by a genome-wide analysis of Pol II behavior at GRF binding sites, as GRF sites that function as roadblocks were found to be enriched for upstream NNS signals relative to those that do not function as roadblocks. Thus, the NNS pathway dictates the fate of Pol II collisions with GRFs, and a lack of NNS signals upstream GRF binding sites enables Pol II to read through roadblocks (Fig. 2). Furthermore, we found that in the context of upstream convergent transcription, the function of the Reb1p binding site in maintaining transcriptional activity of a downstream gene depended on an intact NNS pathway terminating transcription at the Reb1p site. In this manner, one can view the NNS pathway as modulating GRF function by controlling the ability of Pol II to evict GRFs and thereby influencing GRF chromatin modifying behavior. These findings show that the GRFs Reb1, Rap1, and Abf1 and TFIIIB play major roles in limiting intergenic transcription and readthrough by Pol II when it initiates from ncRNA transcription units or intergenic regions. Thus, these TFs not only define promoters by establishing nucleosome-free regions, but they also establish boundaries for upstream transcription units by promoting the pausing and release of Pol II through the NNS pathway.

Studies of pervasive Pol II termination in yeast are aided by the fact that the NNS pathway releases the terminally transcribed nucleotide, and this nucleotide can be sequenced by stabilizing the RNAs normally degraded by the nuclear exosome. In contrast, 3′-end processing of pervasive transcripts in human cells appears to rely exclusively on the cleavage and polyadenylation apparatus (e.g. for PROMPTs), such that

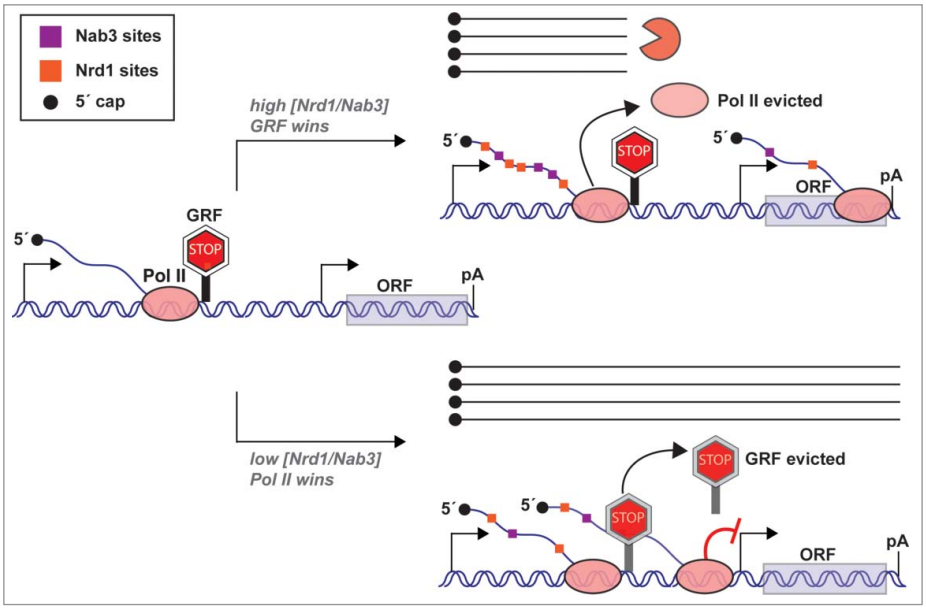

Figure 2. The fate of collisions between Pol II and GRFs is dictated by the level of NNS signals in the nascent RNA. High levels of NNS signals result in sensitization of Pol II to chromatin roadblocks, and an inability to evict GRFs (top panel, legend as in Fig. 1). Low levels of NNS endow Pol II with increased stability and a lower tendency for termination (bottom panel). This results in the eventual eviction of GRFs by Pol II, and resumed transcriptional elongation that can interfere with transcription initiation at the downstream promoter.

the terminally transcribed nucleotide is part of the 3′-cleavage product, which also harbors a destabilizing 5′-phosphate. Therefore, dissecting how DNA-bound proteins assist termination in human cells may require other techniques to afford high-resolution analysis of transcription termination sites. For example, data from PAR-CLIP or NET-seq of Pol II could be analyzed for enrichment of Pol II-associated transcript 3′-ends proximally upstream of protein binding sites. Artificial removal of these binding sites could then reveal the effect of these roadblocks on 3′-end formation and Pol II transcriptional interference into downstream units.

Why do GRFs and TFIIIB promote roadblocking of RNA and DNA polymerases?

It is not completely clear why Reb1p, Rap1p, Abf1p and TFIIIB are able to promote Pol II release by the NNS pathway, while other TFs do not seem to be able to roadblock Pol II. Interestingly, the exact same set of proteins was shown to promote DNA polymerase delta (Pol δ) termination during lagging strand synthesis.[24] The three GRFs Reb1p, Abf1p, and Rap1p rank in the top 10% of proteins by abundance, with Reb1p being almost twice as abundant as Rap1p and Abf1p and having reported abundances ranging from 2533 to 7510 protein molecules per cell.[25,26] Not all GRFs function as roadblocks, as substitution of a Reb1 roadblock at a model NNS target locus by several other well-characterized GRF binding sites (Tbf1p, Mcm1p, or Cbf1p) revealed that none of these could roadblock Pol II, while substituting a Rap1p site restored termination.[16] This difference cannot be explained by protein abundance alone, as Mcm1p and Cbf1p are similar in abundance to Rap1p.[27] We also note that artificial removal of TFIIIB from Pol III promoters revealed the ability of TFIIIC to roadblock Pol II[16]; however, it appears that this is rare *in vivo* and that nearly all tRNA promoters roadblock Pol II using TFIIIB alone. Thus, a common, yet highly specific set of roadblocking proteins act on both RNA and DNA polymerases to promote pausing and termination. Why only these proteins? One possibility is that the rate of binding and dissociation of these factors is such that they are so stably bound to the DNA (i.e. a fast on-rate and slow off-rate) that they can provide a physical block to Pol II or Pol δ progression, while other factors cannot. In the case of Pol δ, rapid binding of GRFs would be essential for their roadblocking function as the Okazaki fragment synthesis occurs shortly after replication fork progression. In support of this hypothesis was the observation that the dynamics of Rap1p binding were strongly correlated to its

roadblocking behavior; for instance, sites where Rap1p promotes the roadblocking of Pol II were associated with longer residence time, as measured in a previous study.[28] Although we do not know whether this model also applies to other GRFs and to TFIIIB, it is tempting to speculate that the residence time for other GRFs is a strong indicator of their ability to promote roadblock. It is also possible that the mode of binding, and not just the residence time may have an impact on the behavior of the polymerase. Binding of TFs that promote roadblocks could be associated with a particular distortion of the geometry of the DNA, leading to topological constraints that may impede the progress of both DNA and RNA polymerases. Aside from these hypotheses, it is possible that other TFs can also physically block Pol II, but that the binding of these proteins was not significant or stable enough in the conditions used in our experiments to influence the behavior of RNA polymerase. Analysis of nuclear exosome targeted 3′-ends in non-standard growth culture conditions may lead to the identification of other TFs that can promote Pol II pausing, leading to termination by the NNS pathway.

The roadblock model for promoting pausing and termination of Pol II in *S. cerevisiae* may potentially extend to other RNA polymerases and organisms. For instance, it was shown that Nsi1p, a Reb1p-homolog is required for efficient termination of transcription by RNA polymerase I at the rDNA loci.[29] Although the mechanisms of transcription termination of Pol I and that of Pol II termination by the NNS pathway differ substantially, the similarities between Reb1p and Nsi1p make it tempting to speculate that the binding of Nsi1p may provide a roadblock mechanism to promote the pausing of Pol I and its release from the rDNA terminator region. The results we obtained in *S. cerevisiae* might also be relevant to transcriptional control in higher eukaryotes. With regards to roadblock termination in human cells, the MAZ protein was shown to promote termination between closely spaced genes by binding to its G_5AG_4 consensus sequence.[30] However, it is not clear how widespread MAZ-assisted termination is across the human transcriptome and this function might be limited to closely spaced genes. Interestingly, CTCF is the closest analog of a GRF in human cells, and it is tempting to speculate that it may provide assistance to Pol II termination at some loci in human cells.[6] Overall, our study underscores the notion that Pol II dynamics are governed by cooperation between *cis*-acting RNA signals and chromatin obstacles, and it is likely that further examples of this cooperation in other model systems will emerge in future studies.

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

Funding

This work was funded by Grant GM61518 (GC). KR was supported by a Philip Whitcome Graduate Fellowship and the National Institutes of Health training program T32GM007185.

ORCID

Kevin Roy http://orcid.org/0000-0001-8799-6207
Guillaume F. Chanfreau http://orcid.org/0000-0002-2505-1950

References

[1] Ntini E, Jarvelin AI, Bornholdt J, Chen Y, Boyd M, Jorgensen M, Andersson R, Hoof I, Schein A, Andersen PR et al. Polyadenylation site-induced decay of upstream transcripts enforces promoter directionality. Nat Struct Mol Biol 2013; 20:923-928; PMID:23851456; http://dx.doi.org/10.1038/nsmb.2640

[2] Steinmetz EJ, Conrad NK, Brow DA, Corden JL. RNA-binding protein Nrd1 directs poly(A)-independent 3′-end formation of RNA polymerase II transcripts. Nature 2001; 413:327-331; PMID:11565036; http://dx.doi.org/10.1038/35095090

[3] Arndt KM, Reines D. Termination of transcription of short noncoding RNAs by RNA polymerase II. Annu Rev Biochem 2014; 84:381-404; PMID:25747400; http://dx.doi.org/10.1146/annurev-biochem-060614-034457

[4] Porrua O, Hobor F, Boulay J, Kubicek K, D'Aubenton-Carafa Y, Gudipati RK, Stefl R, Libri D. In vivo SELEX reveals novel sequence and structural determinants of Nrd1-Nab3-Sen1-dependent transcription termination. EMBO J 2012; 31:3935-3948; PMID:23032188; http://dx.doi.org/10.1038/emboj.2012.237

[5] Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, Leproust EM, Hughes TR, Lieb JD, Widom J et al. The DNA-encoded nucleosome organization of a eukaryotic genome. Nature 2009; 458:362-366; PMID:19092803; http://dx.doi.org/10.1038/nature07667

[6] Fourel G, Miyake T, Defossez PA, Li R, Gilson É. General regulatory factors (GRFs) as genome partitioners. J Biol Chem 2002; 277:41736-41743; PMID:12200417; http://dx.doi.org/10.1074/jbc.M202578200

[7] Hartley PD, Madhani HD. Mechanisms that specify promoter nucleosome location and identity. Cell 2009;

137:445-458; PMID:19410542; http://dx.doi.org/10.1016/j.cell.2009.02.043

[8] Kubik S, Bruzzone MJ, Jacquet P, Falcone JL, Rougemont J, Shore D. Nucleosome stability distinguishes two different promoter types at all protein-coding genes in yeast. Mol Cell 2015; 60:422-434; PMID:26545077; http://dx.doi.org/10.1016/j.molcel.2015.10.002

[9] Krietenstein N, Wal M, Watanabe S, Park B, Peterson CL, Pugh BF, Korber P. Genomic nucleosome organization reconstituted with pure proteins. Cell 2016; 167:709-721. e12; PMID:27768892; http://dx.doi.org/10.1016/j.cell.2016.09.045

[10] Chasman DI, Lue NF, Buchman AR, LaPointe JW, Lorch Y, Kornberg RD. A yeast protein that influences the chromatin structure of UASG and functions as a powerful auxiliary gene activator. Genes Dev 1990; 4:503-514; PMID:2361590; http://dx.doi.org/10.1101/gad.4.4.503

[11] Morse RH. RAP, RAP, open up! New wrinkles for RAP1 in yeast. Trends Genet 2000; 16:51-53; PMID:10652526; http://dx.doi.org/10.1016/S0168-9525(99)01936-8

[12] Yarragudi A, Parfrey LW, Morse RH. Genome-wide analysis of transcriptional dependence and probable target sites for Abf1 and Rap1 in *Saccharomyces cerevisiae*. Nucleic Acids Res 2007; 35:193-202; PMID:17158163; http://dx.doi.org/10.1093/nar/gkl1059

[13] Schlecht U, Erb I, Demougin P, Robine N, Borde V, van Nimwegen E, Nicolas A, Primig M. Genome-wide expression profiling, in vivo DNA binding analysis, and probabilistic motif prediction reveal novel Abf1 target genes during fermentation, respiration, and sporulation in yeast. Mol Biol Cell 2008; 19:308-317; PMID:18003976; http://dx.doi.org/10.1091/mbc.E07-12-1242

[14] Morrow BE, Johnson SP, Warner JR. Proteins that bind to the yeast rDNA enhancer. J Biol Chem 1989; 264:9061-9068; PMID:2656694.

[15] Colin J, Candelli T, Porrua O, Boulay J, Zhu C, Lacroute F, Steinmetz LM, Libri D. Roadblock termination by reb1p restricts cryptic and readthrough transcription. Mol Cell 2014; 56:667-680; PMID:25479637; http://dx.doi.org/10.1016/j.molcel.2014.10.026

[16] Roy K, Gabunilas J, Gillespie A, Ngo D, Chanfreau GF. Common genomic elements promote transcriptional and DNA replication roadblocks. Genome Res 2016; 26:1363-1375; PMID:27540088; http://dx.doi.org/10.1101/gr.204776.116

[17] Korde A, Rosselot JM, Donze D. Intergenic transcriptional interference is blocked by RNA polymerase III transcription factor TFIIIB in *Saccharomyces cerevisiae*. Genetics 2014; 196:427-438; PMID:24336746; http://dx.doi.org/10.1534/genetics.113.160093

[18] Wang Q, Nowak CM, Korde A, Oh DH, Dassanayake M, Donze D. Compromised RNA polymerase III complex assembly leads to local alterations of intergenic RNA polymerase II transcription in *Saccharomyces cerevisiae*. BMC Biol 2014; 12:89; PMID:25348158; http://dx.doi.org/10.1186/s12915-014-0089-x

[19] Fox MJ, Gao H, Smith-Kinnaman WR, Liu Y, Mosley AL. The exosome component Rrp6 is required for RNA polymerase II termination at specific targets of the Nrd1-Nab3 pathway. PLoS Genet 2015; 11:e1004999; PMID:25680078; http://dx.doi.org/10.1371/journal.pgen.1004999

[20] Churchman LS, Weissman JS. Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature 2011; 469:368-373; PMID:21248844; http://dx.doi.org/10.1038/nature09652

[21] Reines D, Ghanouni P, Li QQ, Mote J. The RNA polymerase II elongation complex. Factor-dependent transcription elongation involves nascent RNA cleavage. J Biol Chem 1992; 267:15516-15522; PMID:1379232

[22] Schaughency P, Merran J, Corden JL. Genome-wide mapping of yeast RNA polymerase II termination. PLoS Genet 2014; 10:e1004632; PMID:25299594; http://dx.doi.org/10.1371/journal.pgen.1004632

[23] Schulz D, Schwalb B, Kiesel A, Baejen C, Torkler P, Gagneur J, Soeding J, Cramer P. Transcriptome surveillance by selective termination of noncoding RNA synthesis. Cell 2013; 155:1075-1087; PMID:24210918; http://dx.doi.org/10.1016/j.cell.2013.10.024

[24] Smith DJ, Whitehouse I. Intrinsic coupling of lagging-strand synthesis to chromatin assembly. Nature 2012; 483:434-438; PMID:22419157; http://dx.doi.org/10.1038/nature10895

[25] Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS. Global analysis of protein expression in yeast. Nature 2003; 425:737-741; PMID:14562106; http://dx.doi.org/10.1038/nature02046

[26] Kulak NA, Pichler G, Paron I, Nagaraj N, Mann M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nat Methods 2014; 11:319-324; PMID:24487582; http://dx.doi.org/10.1038/nmeth.2834

[27] Lawless C, Holman SW, Brownridge P, Lanthaler K, Harman VM, Watkins R, Hammond DE, Miller RL, Sims PF, Grant CM et al. Direct and absolute quantification of over 1800 yeast proteins via selected reaction monitoring. Mol Cell Proteom 2016; 15:1309-1322; PMID:26750110; http://dx.doi.org/10.1074/mcp.M115.054288

[28] Lickwar CR, Mueller F, Hanlon SE, McNally JG, Lieb JD. Genome-wide protein-DNA binding dynamics suggest a molecular clutch for transcription factor function. Nature 2012; 484:251-255; PMID:22498630; http://dx.doi.org/10.1038/nature10985

[29] Reiter A, Hamperl S, Seitz H, Merkl P, Perez-Fernandez J, Williams L, Gerber J, Németh A, Léger I, Gadal O et al. The Reb1-homologue Ydr026c/Nsi1 is required for efficient RNA polymerase I termination in yeast. EMBO J 2012; 31:3480-3493; PMID:22805593; http://dx.doi.org/10.1038/emboj.2012.185

[30] Ashfield R, Patel AJ, Bossone SA, Brown H, Campbell RD, Marcu KB, Proudfoot NJ. MAZ-dependent termination between closely spaced human complement genes. EMBO J 1994; 13:5656-5667; PMID:7988563