Biochimica et Biophysica Acta xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

Review

Transcription elongation factors DSIF and NELF: Promoter-proximal pausing and beyond ☆

Yuki Yamaguchi [a,b], Hirotaka Shibata [a], Hiroshi Handa [a,*]

[a] Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama 226-8501, Japan
[b] PRESTO, Japan Science and Technology Agency, Kawaguchi 332-0012, Japan

ARTICLE INFO

Article history:
Received 17 July 2012
Received in revised form 15 November 2012
Accepted 17 November 2012
Available online xxxx

Keywords:
Transcription elongation
Promoter-proximal pausing
DSIF
NELF
P-TEFb

ABSTRACT

DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) were originally identified as factors responsible for transcriptional inhibition by 5,6-dichloro-1-beta-D-ribofuranosyl-benzimidazole (DRB) and were later found to control transcription elongation, together with P-TEFb, at the promoter-proximal region. Although there is ample evidence that these factors play roles throughout the genome, other data also suggest gene- or tissue-specific roles for these factors. In this review, we discuss how these apparently conflicting data can be reconciled. In light of recent findings, we also discuss the detailed mechanism by which these factors control the elongation process at the molecular level. This article is part of a Special Issue entitled: RNA polymerase II Transcript Elongation.

1. Introduction

Promoter-proximal pausing is a key regulatory step of the post-initiation process, first described by John Lis and coworkers in the 1980s [1,2]. In uninduced *Drosophila* heat-shock genes, RNA polymerase II (Pol II) is associated with the promoter-proximal region (around +20 to +50), leading to the idea that pausing is an important rate-limiting step in heat shock induction and contributes to rapid induction of these genes [3,4]. Subsequently, a growing number of genes, such as mammalian *c-myc* and *c-fos*, have been shown to be regulated by promoter-proximal pausing [5,6]. Furthermore, recent genome-wide studies in *Drosophila* and mammalian cell culture have revealed that promoter-proximal pausing is a widespread phenomenon that is observed in the vast majority of genes [7–10].

Several possibilities have been proposed to explain how promoter-proximal pausing is induced. First, local DNA and/or RNA sequence may contribute to site-specific pausing, as has been observed in many bacterial genes [11]. However, ubiquitous *cis* elements downstream of transcriptional start sites (TSS) that might be associated with promoter-proximal pausing have not been discovered in higher eukaryotes. Second, because nucleosome structure clearly serves as a barrier to transcription elongation [12,13], nucleosomes, particularly the first one downstream of the TSS, have been implicated in promoter-proximal pausing [7,14,15]. Third, in vitro biochemical studies using the small-molecule transcriptional inhibitor DRB led to the discovery of a new set of positive and negative transcription elongation factors in the 1990s, and these factors have been implicated in promoter-proximal pausing [16–20]. It is very plausible that the second and third mechanisms might work together to control promoter-proximal pausing and pause release in vivo.

Small molecule compounds have proven useful in elucidating complex biological processes. DRB is one such compound that has been used to study transcription for decades. First developed as a potential antiviral compound active against several viruses including adenovirus and influenza virus [21,22], DRB was later found to inhibit the elongation step of Pol II transcription [23]. Since DRB is a nucleoside analog, the most likely explanation for its action would be that DRB is incorporated into the Pol II active site to inhibit polymerase activity directly. This must not be the case, however, because DRB-mediated transcription inhibition cannot be reproduced in vitro using purified Pol II and general transcription factors (GTFs) [19,23,24]. Because DRB affects transcription of most, if not all, of the protein-coding genes in cultured cells [25], putative targets of DRB were thought to play important roles in vivo. These observations prompted the search for the targets of DRB, and subsequently three novel elongation factors (DSIF, NELF, and P-TEFb) were identified as involved in DRB-mediated transcription inhibition.

DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) were isolated from nuclear extracts of HeLa cells and identified as factors that mediate the inhibitory effect of DRB [19,20]. DSIF is a heterodimer composed of Spt4 and Spt5, and NELF is a protein complex composed of four subunits (A, B, C/D, and E). These factors cooperatively induce transcriptional pausing by binding to

☆ This article is part of a Special Issue entitled: RNA polymerase II Transcript Elongation.

* Corresponding author at: Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259 Nagatsuta, Yokohama 226-8501, Japan. Tel.: +81 45 924 5872; fax: +81 45 924 5145.
E-mail address: handa.h.aa@m.titech.ac.jp (H. Handa).

1874-9399/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbagrm.2012.11.007

initiated Pol II and possibly also to newly synthesized transcript [20,26,27]. Meanwhile, P-TEFb was purified and identified in *Drosophila* Kc cell extract as a factor that stimulates transcription elongation in a DRB-sensitive manner [28–30]. P-TEFb is a protein kinase composed of Cdk9 and Cyclin T, and it phosphorylates the C-terminal domain (CTD) of the largest Pol II subunit in a DRB-sensitive manner. [16–18] One of the consequences of CTD phosphorylation is the reversal of transcription inhibition by DSIF and NELF.

2. The roles of DSIF, NELF, and P-TEFb: general or specific?

The next important questions are how, when, and which genes are controlled by these elongation factors. Initial studies revealed that many of the immediate-early genes that were previously shown to associate with paused polymerase, such as *Drosophila hsp70* and *hsp26* and mammalian *c-fos* and *junb*, are indeed controlled by these elongation factors [31–36].

There is ample evidence that DSIF, NELF, and P-TEFb play general roles in elongation, as GTFs do in initiation. First, because these elongation factors exert their effects even in promoter-independent transcription assays in vitro [20], they may act as promiscuous regulators of transcription regardless of the types of promoters involved. Second, these elongation factors are ubiquitously expressed and evolutionarily conserved across eukaryotes. The only likely exception is NELF, which is absent in some species of nematodes, *S. cerevisiae*, and plants [37]. Coincidentally, chromatin immunoprecipitation (ChIP) studies in *C. elegans* and *S. cerevisiae* have not provided sufficient evidence for widespread promoter-proximal pausing [38,39]. Third, DSIF, NELF, and P-TEFb are essential for development in higher eukaryotes, and depletion of these factors from *Drosophila*, zebrafish, or mice results in various lethal phenotypes [40–43]. For example, *Nelf-b*-deficient mouse embryos die at the time of implantation [41]. Fourth and most importantly, recent genome-wide studies demonstrate that promoter-proximal pausing is more widespread than initially thought. According to recent estimates from ChIP-seq and global run-on (GRO)-seq analyses in *Drosophila* and mammalian cells, 30–80% of the total genes exhibit promoter-proximal enrichment of Pol II [7–10]. (The estimation of the number of affected genes varies widely because these studies use different definitions associated with paused Pol II and therefore count different sets of genes.) Moreover, there is a strong correlation between the level of DSIF, NELF, and Pol II in the promoter-proximal region, and depletion of DSIF or NELF from *Drosophila* or mammalian cells results in a substantial reduction, but not total abolition, of promoter-proximal Pol II occupancy throughout the genome [7–10,44]. Collectively, these findings support the idea that DSIF- and NELF-induced promoter-proximal pausing is a general phenomenon.

However, other studies have suggested that DSIF, NELF, and P-TEFb may function in a tissue- or gene-specific manner. One of the earliest examples comes from the zebrafish mutant called *foggy*^{m806}, which carries a single point mutation in the conserved C-terminal region of Spt5 (the large subunit of DSIF) and exhibits highly specific developmental defects, such as a reduction of dopaminergic neurons and a corresponding surplus of serotonergic neurons in the brain at 2–3 days post fertilization [45]. Because null mutation of the zebrafish *Spt5* gene causes more severe, pleiotropic developmental defects [40], the highly specific phenotypes of *foggy*^{m806} are likely due to a partial loss of function. Nonetheless, the specificity of the hypomorphic phenotypes suggests that not all protein-coding genes may be equally regulated by DSIF. In addition, microarray analyses of DSIF- or NELF-depleted *Drosophila*, zebrafish, mouse, and human cells have revealed that surprisingly few genes are affected by depletion of DSIF or NELF [7,44,46–52]. For example, Narita et al. reported that 204 genes were affected by shRNA-mediated knockdown of NELF-E in human HeLa cells [46]. Comparative analysis of NELF-A, -C/D, and -E depletion in human breast cancer T47D cells identified a common set of 543 genes that were affected by depletion of NELF subunits [49]. In *Drosophila*, NELF depletion from S2 cells significantly altered the levels of 241 transcripts [7,51]. As for DSIF, only 65 of ~8500 genes studied were differentially expressed in Spt5-depleted HeLa cells 1 day before the cells started to die [47]. In zebrafish, expression profiling revealed that 455 genes were differentially expressed in *Spt5*-deficient embryos 24 h after fertilization [48]. Ontology analysis of the aforementioned gene expression data indicates that the set of NELF-affected genes exhibits a significant enrichment in replication-dependent histone genes [47], cell cycle-associated genes [49], and genes responding to environmental or intracellular stimuli 7,51], whereas DSIF depletion upregulates genes induced by environmental or cellular stress and genes encoding certain transcription factors [47,48]. Thus, regardless of the underlying mechanism, expression of some genes appears to be more sensitive than that of others to the depletion of NELF or DSIF.

Although DSIF, NELF, and Pol II co-occupy the TSS in a large fraction of protein-coding genes, depletion of DSIF or NELF seems to affect only a small number of those genes. How can these apparently conflicting observations be reconciled? One might argue that the limited effect on gene expression is due to the presence of residual DSIF or NELF. Indeed, in many of the aforementioned studies, RNAi was used to deplete DSIF or NELF, allowing for the possibility of incomplete knockdown. In addition, because DSIF is essential for cell growth and survival at the single-cell level [47], expression profiling of DSIF-depleted cells must be performed prior to the onset of cell death. Consequently, the number of affected genes may have been underestimated. However, at least in mammals, NELF is not absolutely necessary at the single-cell level. Genetic ablation of *Nelf-b* in mouse embryonic fibroblasts results in slow cell growth and increased apoptosis, but the cells are nonetheless viable and continue to grow [50]. Similar observations were made in *Nelf-a*-deficient mouse embryonic stem cells [our unpublished data], and the slow-growth phenotype is similar to what has been observed after shRNA-mediated NELF-E depletion from HeLa cells [46]. Even under these circumstances, only a small fraction of NELF-associated genes were affected ([50] and our unpublished data). Thus, the limited effect of NELF depletion on gene expression is not simply due to residual protein expression, but NELF intrinsically has a limited effect on gene expression.

An alternative explanation that could account for the discrepancy is that the reversal of promoter-proximal pausing may not always lead to the overexpression of target genes, because mRNA synthesis is a complex process involving multiple rate-limiting steps. For example, nucleosomes may serve as another barrier to elongation. In support of this idea, promoter-proximally paused Pol II is not completely eliminated following NELF depletion [9,44], and the +1 nucleosome has been implicated in promoter-proximal pausing [14,15]. Furthermore, promoter-proximal pausing may not have the same effect on all target genes, even if it is a general phenomenon. The efficiency of natural pause release may vary among genes because of differences in availability of factors that counteract promoter-proximal pausing, such as P-TEFb, its regulators, or other elongation factors. Perhaps, such differences cause the severity or duration of intrinsic pausing to vary among genes, as does the effect of DSIF or NELF depletion. This point will be further discussed in the following section.

Based on biochemical evidence, the likely function of DSIF and NELF is to repress transcription by inducing transcriptional pausing. This model predicts that inhibition of these factors would result in derepression of target genes. In reality, however, microarray analyses have demonstrated that DSIF or NELF depletion results in both upregulation and downregulation of target genes. Some of these changes may be secondary effects, rather than direct consequences, of the depletion. Nonetheless, it is conceivable that transcriptional pausing also contributes to the upregulation of mRNA synthesis, possibly through the formation of a “poised” or “preactivated” state. In this regard, it has been suggested that DSIF- and NELF-induced formation

Please cite this article as: Y. Yamaguchi, et al., Transcription elongation factors DSIF and NELF: Promoter-proximal pausing and beyond, Biochim. Biophys. Acta (2012), http://dx.doi.org/10.1016/j.bbagrm.2012.11.007