

# Human Transcription Elongation Factor NELF: Identification of Novel Subunits and Reconstitution of the Functionally Active Complex

Takashi Narita,<sup>1</sup> Yuki Yamaguchi,<sup>1,2</sup> Keiichi Yano,<sup>3</sup> Seiji Sugimoto,<sup>3</sup> Sittinan Chanarat,<sup>1</sup>  
Tadashi Wada,<sup>1</sup> Dong-ki Kim,<sup>1</sup> Jun Hasegawa,<sup>1</sup> Masashi Omori,<sup>1</sup> Naoto Inukai,<sup>1</sup>  
Masaki Endoh,<sup>1</sup> Tomoko Yamada,<sup>1</sup> and Hiroshi Handa<sup>4\*</sup>

Graduate School of Bioscience and Biotechnology<sup>1</sup> and Frontier Collaborative Research Center,<sup>4</sup> Tokyo Institute of Technology, and PRESTO, Japan Science and Technology Corporation,<sup>2</sup> Yokohama 226-8503, and Tokyo Research Laboratories, Kyowa Hakko Kogyo Company, Tokyo 194-8533,<sup>3</sup> Japan

Received 8 October 2002/Returned for modification 25 November 2002/Accepted 23 December 2002

**The multisubunit transcription elongation factor NELF (for negative elongation factor) acts together with DRB (5,6-dichloro-1- $\beta$ -D-ribofuranosylbenzimidazole) sensitivity-inducing factor (DSIF)/human Spt4-Spt5 to cause transcriptional pausing of RNA polymerase II (RNAPII). NELF activity is associated with five polypeptides, A to E. NELF-A has sequence similarity to hepatitis delta antigen (HDAG), the viral protein that binds to and activates RNAPII, whereas NELF-E is an RNA-binding protein whose RNA-binding activity is critical for NELF function. To understand the interactions of DSIF, NELF, and RNAPII at a molecular level, we identified the B, C, and D proteins of human NELF. NELF-B is identical to COBRA1, recently reported to associate with the product of breast cancer susceptibility gene BRCA1. NELF-C and NELF-D are highly related or identical to the protein called TH1, of unknown function. NELF-B and NELF-C or NELF-D are integral subunits that bring NELF-A and NELF-E together, and coexpression of these four proteins in insect cells resulted in the reconstitution of functionally active NELF. Detailed analyses using mutated recombinant complexes indicated that the small region of NELF-A with similarity to HDAG is critical for RNAPII binding and for transcriptional pausing. This study defines several important protein-protein interactions and opens the way for understanding the mechanism of DSIF- and NELF-induced transcriptional pausing.**

The elongation step of RNA polymerase II (RNAPII) transcription is emerging as a critical control point for the expression of various genes and for diverse biological processes. Examples include neuronal fate determination during embryonic development (6, 44), gene expression of human immunodeficiency virus (5, 11, 13, 19, 43), replication and transcription of hepatitis delta virus (38), and transcriptional regulation of heat shock genes (1, 10, 18). In all these cases, the involvement of three transcription elongation factors, namely, DRB (5,6-dichloro-1- $\beta$ -D-ribofuranosylbenzimidazole) sensitivity-inducing factor (DSIF), NELF (negative elongation factor), and positive transcription elongation factor b (P-TEFb), has been demonstrated or implicated.

Shortly after the initiation of transcription, RNAPII comes under the negative and positive control of DSIF, NELF, and P-TEFb. DSIF and NELF cause transcriptional pausing through physical association with RNAPII. DSIF binds to RNAPII directly and stably (33, 36). However, this appears to have little effect on the catalytic activity of RNAPII (37). A previous study has pointed out that NELF does not bind substantially to DSIF or RNAPII alone but does bind to the complex of DSIF and RNAPII (40). This association is the likely trigger of transcriptional pausing. Conversely, P-TEFb allows RNAPII to enter the productive elongation phase by preventing the action of DSIF and NELF (27, 37). P-TEFb is

the protein kinase whose primary target is thought to be the C-terminal domain (CTD) of RNAPII (26). Most, but not all, evidence suggests that P-TEFb-dependent phosphorylation of the CTD facilitates the release of DSIF and NELF from RNAPII, thereby reversing the inhibition (3, 24, 37). In theory, such regulation at the elongation step allows for rapid change in mRNA levels and for highly sophisticated control over gene expression when combined with regulation at the (pre)initiation step.

The structures and functions of DSIF and P-TEFb have been extensively characterized. Human DSIF is a heterodimer composed of p14 (14 kDa) and p160 (160 kDa), whose *Saccharomyces cerevisiae* counterparts are Spt4 and Spt5 (7, 33). In addition to its role in transcriptional pausing, DSIF has a potential to activate RNAPII elongation. The activation mechanism is not well understood: interaction partners of DSIF other than NELF may be involved (13, 14, 20, 23, 28). Spt5 has a highly acidic N-terminal region, multiple copies of the KOW motifs, and a repetitive C-terminal region analogous to the RNAPII CTD (9, 25, 36). RNAPII interacts with Spt5 through a region encompassing the KOW motifs. KOW motifs are also found in the bacterial transcription elongation factor NusG, which binds to prokaryotic RNA polymerase and controls termination and antitermination (15, 17, 29). In addition, the extreme C terminus of Spt5 is specifically involved in the transcriptional repression pathway (6). Human P-TEFb is a heterodimer composed of Cdk9 (41 kDa) and one of multiple cyclin subunits T1, T2a, T2b, and K (50 to 90 kDa) (26). The kinase activity of P-TEFb is essential for its ability to stimulate transcriptional elongation and to counteract DSIF and NELF

\* Corresponding author. Mailing address: Frontier Collaborative Research Center, Tokyo Institute of Technology, 4259 Nagatsuta, Yokohama 226-8503, Japan. Phone: 81-45-924-5872. Fax: 81-45-924-5145. E-mail: hhanda@bio.titech.ac.jp.

inhibition. Recent studies have shown that the P-TEFb kinase is negatively regulated by a cellular RNA called 7SK (22, 41). The P-TEFb kinase is also strongly inhibited by synthetic compounds such as DRB, *N*-[(2-methylamino)ethyl]-5-isoquinolinesulfonamide (H-8), and flavopiridol (4, 19, 43). In the past and the present, DRB has played important roles in the identification of DSIF, NELF, and P-TEFb and in the elucidation of their actions. In *in vitro* transcription assays using HeLa nuclear extracts, the P-TEFb kinase is often prevalent and obscures the actions of DSIF and NELF. Inhibition of P-TEFb by DRB uncovers the DSIF- and NELF-dependent transcriptional pausing (33, 34, 37).

The activity of human NELF is associated with five polypeptides, A (66 kDa), B (62 kDa), C (60 kDa), D (59 kDa), and E (46 kDa), of which only NELF-A and NELF-E have been identified (37, 38). The structure of NELF-E is characterized by an N-terminal leucine zipper motif, a central domain rich in Arg-Asp dipeptide repeats (the RD motif), and a C-terminal RNA recognition motif (RRM). A previous study has shown that the RRM binds to a set of RNA sequences and is critical for NELF activity (40). NELF-A is encoded by *WHSC2*, a candidate gene for Wolf-Hirschhorn syndrome (35). Interestingly, its N-terminal segment shows sequence similarity to the hepatitis delta antigen (HDAG), the hepatitis delta virus protein that binds to RNAPII and activates transcriptional elongation (38).

The goals of this study are to define the structure of NELF and, at a molecular level, to understand the interactions of DSIF, NELF, and RNAPII, which are likely relevant to transcriptional pausing. We report here the identification of the B, C, and D proteins of NELF, the reconstitution of functionally active NELF, and detailed analyses with mutated recombinant complexes. Our data define several important protein-protein interactions within the NELF complex as well as between NELF and RNAPII.

## MATERIALS AND METHODS

**Peptide sequencing and molecular cloning.** Preparation of protein samples and microsequencing analysis were performed as described previously (37). With regard to NELF-B, three peptide sequences were obtained. Translated BLAST searches identified several expressed sequence tags (ESTs) that matched. Primers were synthesized based on available information, and partial cDNAs were obtained by reverse transcription PCR. With the cDNAs as probes, a HeLa cDNA library was screened for a full-length cDNA coding for NELF-B. Among three clones thus obtained, the longest one (GenBank accession number E36720) was 2,471 bp long and predicted to encode a 580-amino-acid (aa) protein. With regard to NELF-C and NELF-D, elution profiles of Lys-C-digested materials obtained by reverse-phase high-performance liquid chromatography were quite similar. Three (NELF-C) and four (NELF-D) peptide sequences were determined by microsequencing analysis. Some of the sequences obtained independently from NELF-C and NELF-D were identical, and all of the sequences matched parts of the single protein called human TH1 (GenBank accession number NM016397).

**Preparation of antibodies and bacterial recombinant proteins.** Antibodies N12 and C14 were generated in rabbits against synthetic peptides AGAVP-GAIMDED and EGEHDPVTEFIAHC, respectively, conjugated to keyhole limpet hemocyanin. These peptides correspond to amino acid regions 2 to 13 and 568 to 582 of the 590-aa TH1 protein. Epitope-specific antibodies were affinity purified over a peptide-conjugated column. cDNA fragments encoding the 581- and 590-aa TH1 proteins were cloned into pET-14b (Novagen). The recombinant proteins were expressed in *Escherichia coli* strain BL21(DE3) and purified by nickel affinity chromatography.

**Baculoviral expression of recombinant NELF subunits.** Oligonucleotides encoding YPYDVPDYA (hemagglutinin [HA]), EQKLISEEDL (c-Myc), YTDI-

EMNRLGK (vesicular stomatitis virus glycoprotein [VSV-G]), and DYKD-DDDK (Flag) were attached to the 5' ends of open reading frames for NELF-A, NELF-B, NELF-D, and NELF-E by PCR. The tagged sequences were cloned into pFastBac-HT (Invitrogen), thus generating pFB-HA-NELF-A, pFB-myc-NELF-B, pFB-VSV-NELF-D, and pFB-Flag-NELF-E. *E. coli* DH10<sub>BAC</sub> cells were transformed with these plasmids and grown on Luria-Bertani plates containing 50 µg of kanamycin/ml, 7 µg of gentamicin/ml, 10 µg of tetracycline/ml, 40 µg of isopropyl-β-D-thiogalactopyranoside/ml, and 0.1 mg of Bluo-gal (Invitrogen)/ml to allow for DNA recombination. Recombinant baculoviral DNAs were recovered from white colonies by the alkaline lysis procedure and transfected into Sf9 cells by using Lipofectin reagent (Invitrogen). Four days posttransfection, culture supernatants were collected as the primary virus solutions (~2 × 10<sup>7</sup> PFU/ml). For virus amplification, Sf9 cells were infected at a multiplicity of infection of 0.1 and culture supernatants were collected 96 h postinfection as the secondary virus solutions (~10<sup>8</sup> PFU/ml), which were used in the subsequent experiments. For protein expression, Sf9 cells (7 × 10<sup>6</sup> to 10 × 10<sup>6</sup> cells) that were seeded onto 100-mm-diameter dishes 1 h prior to infection were infected with various combinations of the recombinant viruses at a multiplicity of infection of 2. Two days postinfection, the cells were washed twice with phosphate-buffered saline and lysed in 1 ml of Nonidet P-40 (NP-40) lysis buffer (50 mM Tris-HCl [pH 7.9], 150 mM NaCl, 1% NP-40, 1 mM EDTA [pH 8.0]) or 500 µl of high-salt buffer (50 mM Tris-HCl [pH 7.9], 500 mM NaCl, 1% NP-40, 2 mM EDTA [pH 8.0]) with brief sonication. The cell lysates were cleared by centrifugation at 20,000 × *g* for 15 min at 4°C.

**Immunoprecipitation and immunoblotting.** To purify recombinant NELF complexes, Sf9 lysates (total, 1 ml), prepared by using high-salt buffer and subsequently diluted with the same volume of deionized water to reduce salt concentrations, were incubated with 20 µl of anti-Flag M2 agarose (Sigma) equilibrated with 0.5× high-salt buffer for 1 h at 4°C. After the matrices were washed three times with 180 µl of 0.5× high-salt buffer and twice with 180 µl of NE buffer (20 mM HEPES-NaOH [pH 7.9], 20% glycerol, 100 mM KCl, 0.2 mM EDTA [pH 8.0], 0.1% NP-40), bound materials were eluted four times with 20 µl of NE buffer containing 0.1 mg of Flag peptide (Sigma)/ml.

For immunoprecipitation with anti-HA antibody, Sf9 lysates (200 µl) prepared by using NP-40 lysis buffer were incubated for 1 h at 4°C with 20 µl of protein G Sepharose (Amersham) coupled with 3 µg of anti-HA monoclonal antibody (3F10; Roche) and equilibrated with WB buffer (20 mM HEPES-NaOH [pH 7.9], 20% glycerol, 100 mM KCl, 0.2 mM EDTA [pH 8.0], 0.5% NP-40, 0.5% skim milk). After the matrices were washed four times with 180 µl of WB buffer and four times with 180 µl of WH buffer (20 mM HEPES-NaOH [pH 7.9], 20% glycerol, 100 mM KCl, 0.2 mM EDTA [pH 8.0], 0.5% NP-40), bound materials were eluted with 60 µl of Laemmli buffer. For immunoprecipitation with anti-VSV-G antibody, Sf9 lysates (total, 800 µl), prepared by using high-salt buffer and subsequently diluted with the same volume of deionized water, were incubated for 1 h at 4°C with 20 µl of protein G Sepharose coupled with 5 µg of anti-VSV-G monoclonal antibody (P5D4; Roche) and equilibrated with 0.5× high-salt buffer. After the matrices were washed five times with 180 µl of 0.5× high-salt buffer, bound materials were eluted with 60 µl of Laemmli buffer.

NELF affinity purified from a HeLa cell line constitutively expressing Flag-E (containing 200 ng of NELF-B protein) was incubated for 2 h at 4°C with 20 µl of protein G Sepharose coupled with 15 µg of anti-NELF-C (N12) or anti-c-Myc (9E10; Roche) antibody. After the matrices were washed three times with 500 µl of high-salt buffer and once with 500 µl of NE buffer, bound materials were eluted with 30 µl of Laemmli buffer (see Fig. 2C).

The HeLa nuclear extract-derived fraction (P.3/D.3), rich in DSIF and RNAPII but devoid of NELF, was used for immunoprecipitation (see Fig. 5C and 7D). The 0.3 M KCl fraction of the phosphocellulose P11 column was subjected to DEAE Sepharose chromatography, and the 0.225 to 0.3 M KCl fraction was collected as described previously (37). This fraction was dialyzed against NE buffer and used the same way as P.3/D.3. P.3/D.3 (500 µl) and Sf9 lysates (200 µl) prepared by using NP-40 lysis buffer were incubated for 2 h on ice. After the removal of precipitated materials by centrifugation, a 1/10 volume of 5% skim milk was added to the supernatants, which were incubated further for 2 h at 4°C with 20 µl of anti-Flag agarose or protein G Sepharose coupled with anti-HA antibody. After extensive washing, bound materials were eluted with Flag peptide or Laemmli buffer as described above.

For immunoblotting, primary antibodies were used as follows: anti-HA antibody, 1 µg/ml; anti-c-Myc antibody, 2 µg/ml; anti-VSV-G antibody, 1 µg/ml; anti-Flag M2 monoclonal antibody (Sigma), 4.4 µg/ml; anti-NELF-C (N12) sera, 1:2,000 dilution; anti-NELF-C and -NELF-D (C14) sera, 1:2,000 dilution; anti-DSIF p160 rat monoclonal antibody (33), 3 µg/ml; and anti-CTD monoclonal antibody (8WG16; Babco), 0.45 µg/ml. Additional materials used were horseradish peroxidase-conjugated anti-rat immunoglobulin (Dako), 1:1,000 dilution;

biotinylated anti-mouse immunoglobulin (Amersham), 1:1,000 dilution; streptavidin-horseradish peroxidase (Amersham), 1:3,000 dilution; and the ECL detection kit (Amersham).

**GST pull-down assays.** To prepare NELF-D protein, pBluescript-NELF-D was digested with *EcoRV* and transcribed with T3 RNA polymerase, followed by translation with the rabbit reticulocyte lysate system (Promega) in the presence of [<sup>35</sup>S]methionine. Various fragments of NELF-A were prepared by PCR or by restriction endonuclease digestion and cloned into pGEX vectors (Amersham). Glutathione *S*-transferase (GST)-NELF-A mutant proteins (~1 μg) expressed in *E. coli* strain DH5α or BL21(DE3) were bound to 20 μl of glutathione Sepharose (Amersham). After being washed three times with 180 μl of NETN (20 mM Tris-HCl [pH 7.9], 100 mM NaCl, 1 mM EDTA [pH 8.0], 0.5% NP-40) and twice with 180 μl of NE buffer, the beads were incubated with <sup>35</sup>S-labeled NELF-D (3 μl) in 100 μl of NE buffer for 1 h at 4°C. After the beads were washed five times with 400 μl of NE buffer, bound materials were eluted with 30 μl of Laemmli buffer, resolved by sodium dodecyl sulfate-10% polyacrylamide gel electrophoresis (SDS-PAGE), and visualized by Coomassie blue staining and autoradiography. Twenty microliters of glutathione Sepharose coupled with GST-NELF-A proteins was incubated for 2 h at 4°C with 500 μl of P.3/D.3 that was treated with RNase A (15 μg). This nuclease treatment was done to reduce background due to nonspecific interactions mediated by endogenous RNA molecules. After the mixture was washed five times with 400 μl of NE buffer, bound materials were eluted with 45 μl of Laemmli buffer. The eluates were analyzed by SDS-PAGE and Coomassie blue staining or immunoblotting (see Fig. 8A).

**Transcription assays.** Transcription assays were performed as described previously (33, 37).

RESULTS

**Identification of B, C, and D proteins of NELF.** To further understand the role of NELF in elongational control, we identified three proteins associated with NELF activity: B, C, and D (Fig. 1A).

With regard to NELF-B (62 kDa), three peptide sequences were determined by microsequencing analysis. Translated BLAST searches identified several ESTs that matched those sequences. Based on the ESTs, cDNAs encoding NELF-B were identified from a HeLa cDNA library. The longest clone (2,471 bp; GenBank accession number E36720) encodes a 580-aa protein (Fig. 1B). Recently, the same protein was independently identified by a two-hybrid screen with the carboxy-terminal domain of BRCA1, the breast cancer susceptibility gene product, as bait (42). NELF-B or COBRA1 (for cofactor of BRCA1) was shown to coimmunoprecipitate with BRCA1 when overexpressed and to affect large-scale chromatin structure when bound to a subchromosomal DNA region at a high density (42). These findings are potentially interesting; however, it should be noted that these results were based on transient overexpression of NELF-B in cultured cells. We suspect that, under such conditions, most of the exogenously expressed NELF-B would not form a functional NELF complex. Computer analysis identified two regions predicted to form the coiled-coil structure in NELF-B. These regions may be involved in interaction with NELF-E: NELF-E has a leucine zipper motif, through which it binds to NELF-B (data not shown; see below).

With regard to NELF-C (60 kDa) and NELF-D (59 kDa), three and four peptide sequences, respectively, were determined by microsequencing analysis (Fig. 1C). Unexpectedly, all these peptides were found to match parts of the human TH1 protein (GenBank accession number NM016397), indicating that NELF-C and NELF-D are highly related proteins. TH1 was previously identified by a search for genes located on chromosome 20q13 (2), and no function has been assigned to

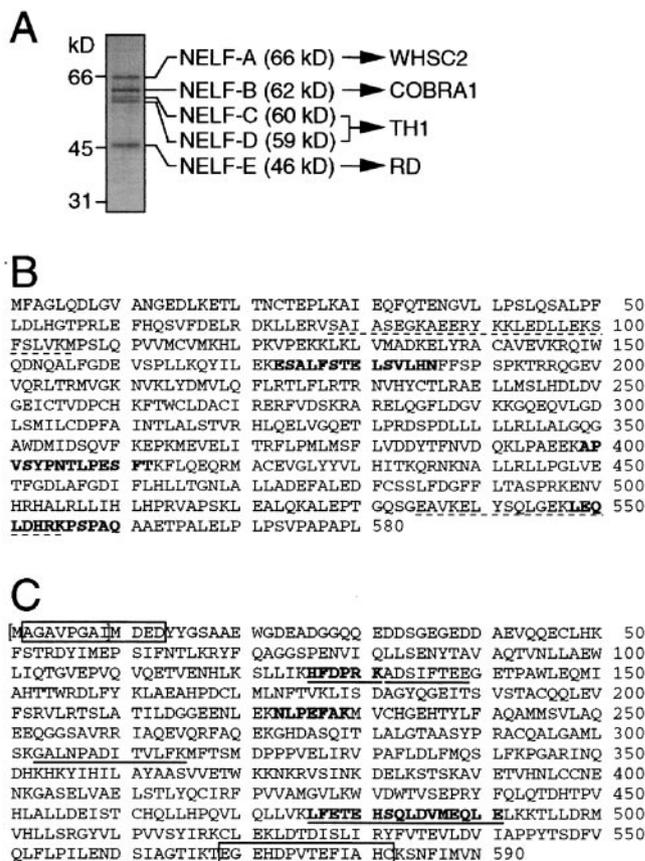


FIG. 1. Cloning of NELF-B, NELF-C, and NELF-D. (A) NELF was purified from human HeLa cell nuclear extracts as described previously (37). NELF from the final purification step (mono S) was visualized by silver staining. (B) Amino acid sequence of human NELF-B. The three obtained peptide sequences are shown in boldface. Dotted lines indicate two regions predicted (probability, >0.5) by a 21-residue scan with the program COILS to form the coiled-coil structure. (C) Amino acid sequence of human NELF-C and NELF-D. The three peptide sequences obtained from NELF-C are shown in boldface. The four peptide sequences obtained from NELF-D are underlined. Boxed are the sequences used for the generation of antipeptide antibodies N12 and C14. Amino acids in brackets are probably unique to NELF-C.

it. Besides the coiled-coil structure of NELF-B, NELF-B, NELF-C, and NELF-D had no sequence similarities or motifs indicative of their functions. Putative orthologs of NELF-B and of NELF-C and NELF-D were found in mouse and *Drosophila melanogaster* but not in *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, and *Arabidopsis thaliana*.

Recently, we found longer TH1 cDNAs in GenBank that are capable of encoding 9 extra amino acids at the N terminus of the previously described 581-aa TH1 protein (Fig. 2A). To test the possibility that NELF-C and NELF-D differ at their N termini, we cloned a longer cDNA for TH1 under the T3 promoter and performed transcription and translation with rabbit reticulocyte lysates. As a result, two products with sizes similar to those of NELF-C and NELF-D were obtained (Fig. 2A), and the smaller one was indistinguishable in size from a unique product that was obtained from a parallel reaction with TH1 cDNA lacking the open reading frame that encodes the

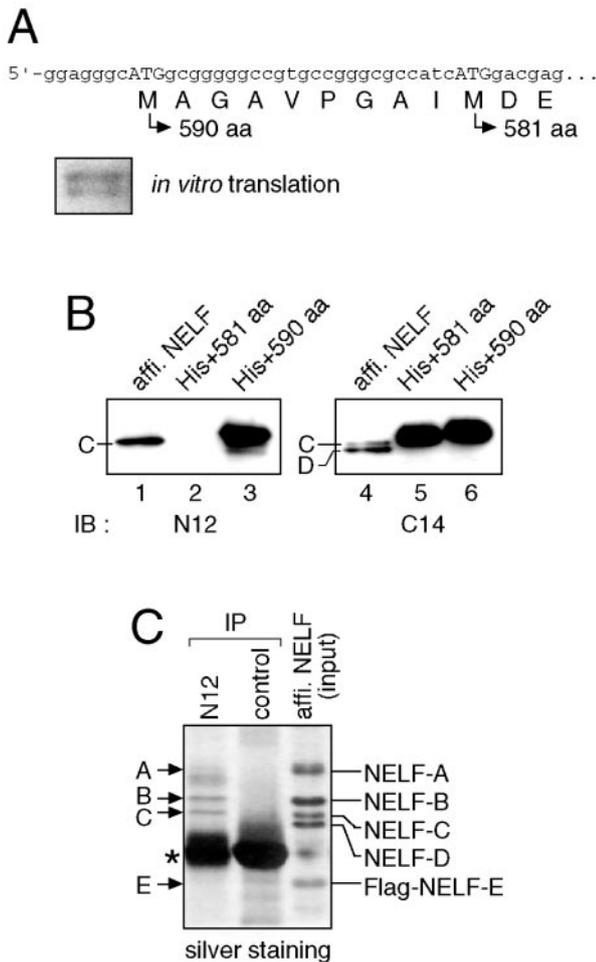


FIG. 2. Distinction between NELF-C and NELF-D. (A) A cDNA encoding the 590-aa form of TH1 was cloned under the T3 promoter and used as a template for *in vitro* transcription and translation. Translation was performed with rabbit reticulocyte lysates in the presence of [<sup>35</sup>S]cysteine, and reaction products were analyzed by SDS-PAGE and autoradiography. The 5'-terminal sequence of the cDNA and putative translation initiation sites are shown. (B) Affinity-purified NELF (affi. NELF) and the 581- and 590-aa forms of recombinant TH1 proteins with histidine tags (His) were examined by immunoblotting (IB) with the antibodies N12 and C14. (C) NELF affinity purified from a HeLa cell line constitutively expressing Flag-E was immunoprecipitated (IP) with the antibodies N12 and c-Myc (control). Bound materials and 10% of the input were analyzed by SDS-PAGE and silver staining. The asterisk denotes the position of immunoglobulin heavy chains. Though bands of NELF-A and NELF-E are not clearly visible in lane 1 on this gel, their presence has been confirmed by immunoblot analysis (data not shown).

extra 9 aa (data not shown). These results suggest that NELF-C and NELF-D arise from a common mRNA species by the alternative usage of translation initiation codons. To distinguish these proteins, we made a polyclonal antibody against the N-terminal peptide unique to the 590-aa protein (N12) and an antibody against the C-terminal peptide common to both proteins (C14) as a control (Fig. 1C). Immunoblot analysis using bacterially expressed TH1 proteins showed that N12 specifically reacts with the 590-aa form but that C14 reacts with both forms of TH1 (Fig. 2B). When affinity-purified NELF was

used instead, N12 detected only NELF-C, whereas C14 detected both NELF-C and NELF-D (Fig. 2B). These results strongly indicate that NELF-C contains the 9-aa extension at the N terminus but that NELF-D does not.

Given the similarity of NELF-C to NELF-D, NELF might be four, rather than five, subunit complexes composed of A, B, either C or D, and E. To test this possibility, we performed immunoprecipitation of NELF-C from affinity-purified NELF complexes by using the NELF-C-specific N12 antibody. Silver staining (Fig. 2C) and immunoblot analysis (data not shown) of the precipitated materials showed the presence of A, B, C, and E polypeptides but not D. It is therefore likely that NELF-C and NELF-D are contained in distinct NELF complexes.

Northern blot analysis indicated that all the NELF subunits were expressed in all human tissues examined (Fig. 3). NELF-A, NELF-B, and NELF-E had major 2.4-, 2.7-, and 1.6-kb mRNAs, respectively. NELF-C and NELF-D had a major 2.4-kb mRNA with an additional 4.4-kb species in heart and skeletal muscle. The relative abundances of mRNA in these tissues were similar among different subunits (e.g., less in lung and liver). These results are consistent with the proposed general roles of NELF in transcriptional elongation.

**Reconstitution of functionally active NELF with recombinant proteins.** The identification of the five proteins associated with NELF activity prompted us to reconstitute the NELF complex with recombinant proteins. As a first step, the 581-aa form of TH1, which likely corresponds to NELF-D, was used on behalf of NELF-C and NELF-D. To ease characterization, four different epitope tags, HA, c-Myc, VSV-G, and Flag, were attached to the N termini of NELF-A, NELF-B, NELF-D, and NELF-E, respectively (Fig. 4A). These fusion proteins were overexpressed together in insect Sf9 cells, and the cell lysates were subjected to anti-Flag affinity chromatography. Purification under stringent conditions (250 mM NaCl and 0.5% NP-40) yielded several proteins ranging from 70 kDa to 50 kDa (Fig. 4B). Immunoblot analysis identified them as HA-A, VSV-D, Myc-B, and Flag-E in accordance with their decreasing molecular weights on SDS-PAGE (Fig. 4C, lane 5). The stoichiometry of recombinant NELF-A was relatively low: this subunit seemed to be partially dissociated during purification under the stringent conditions. Recombinant NELF-B showed a mobility slightly faster (indicating a difference of ~3 kDa) than that of native NELF-B (data not shown), migrating faster than NELF-D on SDS-PAGE; we do not currently know the reason for this discrepancy. In addition, recombinant NELF-E appeared as multiple bands, probably resulting from proteolytic degradation at the C terminus. Despite these facts, the above results demonstrated that coexpression of the four proteins in Sf9 cells allows for the formation of a NELF-like complex, which can be isolated by one-step affinity chromatography.

We then studied the molecular basis for NELF complex formation. The four subunits were overexpressed in various combinations in Sf9 cells, and NELF subassemblies were analyzed by using a set of anti-epitope tag antibodies. Briefly, as shown in Fig. 4C, (i) NELF-E interacts directly with NELF-B (lane 3) but not with NELF-A (lanes 2 and 3) or NELF-D (lane 2), (ii) NELF-D interacts directly with NELF-A (lane 8) and NELF-B (lane 11), and (iii) NELF-B does not interact directly with NELF-A (lane 3). All the available data are con-

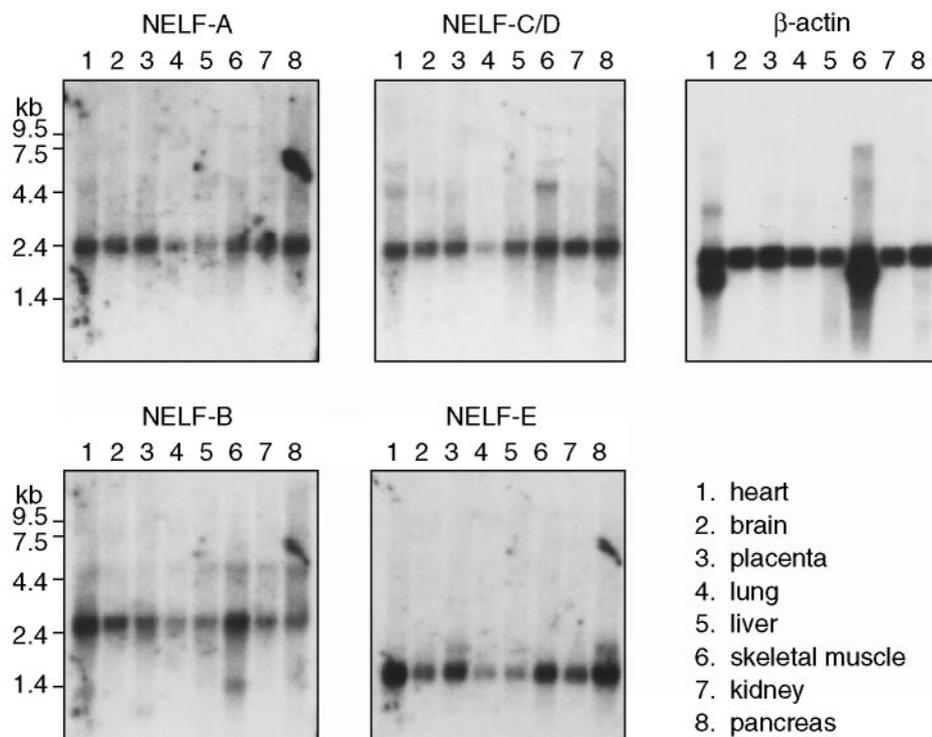


FIG. 3. NELF subunits are ubiquitously expressed in various human tissues. Multiple-tissue Northern blot (Clontech) containing 2  $\mu$ g of poly(A)-containing mRNA per lane was used to determine expression levels of NELF subunit mRNAs in a variety of human tissues. Appropriate restriction fragments of NELF subunit cDNAs were gel purified and  $^{32}$ P labeled by using Klenow DNA polymerase and random hexamers. Human  $\beta$ -actin probe (Clontech) was also used to confirm equivalent loadings of intact RNA. The positions of markers in kilobases are indicated to the left.

sistent with the model that NELF-B and NELF-D (or NELF-C) bring NELF-A and NELF-E together via three protein-protein interactions (Fig. 4D).

We asked whether recombinant NELF is functionally equivalent to native NELF. In HeLa nuclear extracts, NELF represses transcription when P-TEFb is inhibited by DRB. As reported previously (37, 40), immunodepletion of NELF from the extracts increased the level of transcription in the presence of DRB, whereas adding affinity-purified NELF back repressed transcription in the presence of DRB (Fig. 5A, lanes 2, 4, 6, and 8). Recombinant NELF (A-D-B-E) also strongly repressed transcription (Fig. 5A, lanes 10, 12, 18, and 19). In contrast, neither affinity-purified NELF nor recombinant NELF (A-D-B-E) affected transcription in the absence of DRB (Fig. 5A, lanes 5, 7, 9, and 11). We performed another test using a deoxycytidine (dC)-tailed template. This template is efficiently transcribed by pure RNAPII without additional factors and is useful for assaying the activity of transcription elongation factors. This reaction proceeds in a manner sensitive to DSIF and NELF (37, 40). While recombinant DSIF and NELF (A-D-B-E) had little effect individually on transcription by RNAPII, the addition of both proteins strongly repressed transcription (Fig. 5B). As has been shown previously, NELF binds to RNA of various sequences through the RRM of NELF-E (40). This finding was again recapitulated by the results of experiments with recombinant NELF (A-D-B-E) (data not shown). From these results, we concluded that recombinant NELF is functionally fully active *in vitro*.

**All the subunits of NELF are required for its function.** To elucidate the functional significance of each NELF subunit, we wished to examine the activities of NELF subassemblies. It has already been shown that NELF-E is critical for NELF activity (40). Considering the linear configuration of NELF subunits (Fig. 4D), we decided to delete NELF subunits serially from the terminus carrying NELF-A. Two NELF subcomplexes, D-B-E and B-E, were prepared as described above and used for transcription assays. In depletion-add back experiments, these subcomplexes had no detectable effect on transcription in the presence and absence of DRB (Fig. 5A, lanes 14 to 17, and data not shown). In addition, the D-B-E complex did not repress RNAPII transcription on the dC-tailed template regardless of the presence of DSIF (Fig. 5B). Given the importance of NELF-A, we assayed the effect of NELF-A alone on transcriptional elongation (Fig. 5A). Bacterially expressed NELF-A protein had no detectable effect (lanes 21 to 23), while further addition of the D-B-E complex strongly repressed transcription (lanes 24 to 26). Taken together, these results demonstrated that NELF-A is essential, but not sufficient, for NELF's ability to cooperate with DSIF to repress transcriptional elongation.

It has been shown previously that NELF binds to the complex of DSIF and RNAPII (40). To know if NELF subcomplexes have a compromised ability to associate with DSIF/RNAPII, Sf9 cell lysates containing various combinations of NELF subunits were incubated with a HeLa nuclear extract-derived fraction rich in DSIF and RNAPII but devoid of NELF

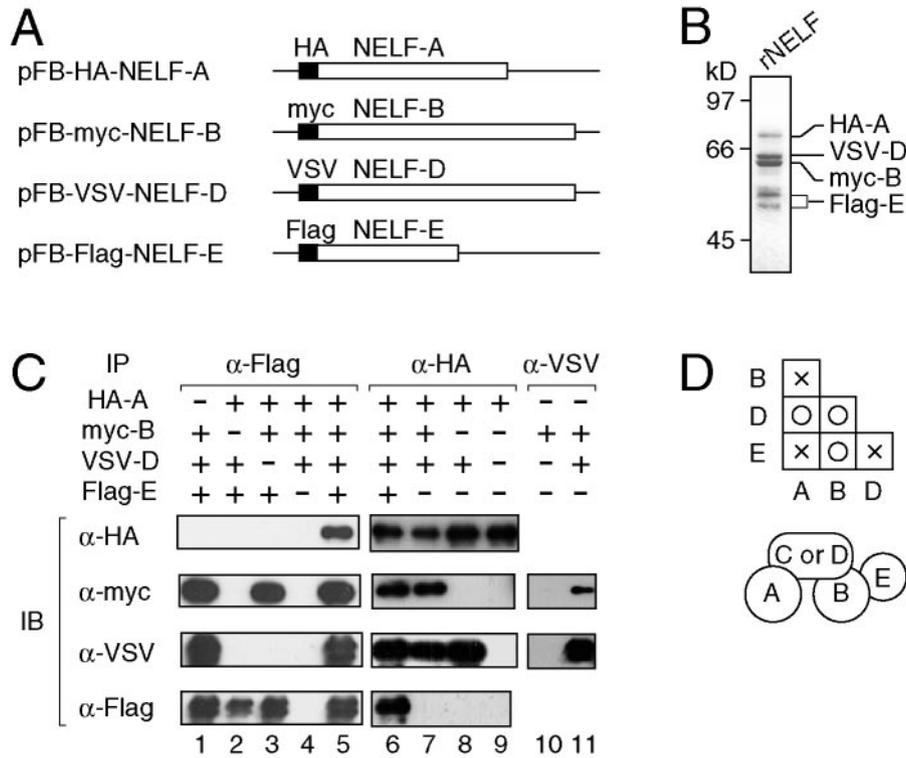


FIG. 4. Reconstitution of NELF complex with recombinant proteins. (A) pFastBac (pFB)-based DNAs constructed for expression of recombinant NELF subunits. Each subunit was fused to one of four epitope tags, as shown. (B) Coomassie blue staining of purified recombinant NELF (rNELF). Sf9 cells that were multiply infected by recombinant baculoviruses for HA-A, Myc-B, VSV-D, and Flag-E were lysed and subjected to anti-Flag affinity chromatography as described in Materials and Methods. Purified proteins were resolved by SDS-7.5% PAGE and stained by Coomassie blue. The identities of these bands were determined by using anti-epitope tag antibodies and are indicated to the right. The positions of molecular mass markers in kilodaltons (kD) are indicated to the left. (C) Identification of multiple interactions among NELF subunits. Sf9 cells were infected by the recombinant baculoviruses in various combinations. These cells were lysed, immunoprecipitated (IP) with anti-Flag ( $\alpha$ -Flag) (lanes 1 to 5), anti-HA ( $\alpha$ -HA) (lanes 6 to 9), and anti-VSV ( $\alpha$ -VSV) (lanes 10 and 11) antibodies, and analyzed by immunoblotting (IB) with the indicated antibodies. (D) Summary of the results. (Top) O's and X's indicate the presence and absence of interactions, respectively. (Bottom) Schematic structure of NELF complex.

(P.3/D.3). The mixtures were immunoprecipitated, and NELF-associated proteins were analyzed. Under these conditions, DSIF and RNAPII bound to recombinant holo-NELF but not to the subcomplex lacking NELF-A or NELF-E (Fig. 5C). Thus, NELF-A and NELF-E may be directly involved in interaction with DSIF/RNAPII or, alternatively, may play structural roles to maintain the NELF complex in an active form. Regardless of the roles of these subunits, these results demonstrated that four types of the NELF subunits are all required for its function: NELF-B and either NELF-C or NELF-D serve as structural components, while NELF-A and NELF-E play structural and/or functional roles.

**The NELF-A regions with similarity to HDag are docking sites for NELF-C or NELF-D and RNAPII.** In light of the above data, we focused on the molecular structure of NELF-A in the subsequent study. The N-terminal segment of NELF-A shows weak homology to HDag, the recently identified viral transcription elongation factor, while the C-terminal region is rich in proline (P), serine (S), and threonine (T) residues (Fig. 6). PST-rich regions are often found in transactivation domains of transcription factors (30, 31). To perform structure-function analysis of NELF-A in the context of the NELF complex, we needed prior knowledge about the NELF-A region that is

responsible for NELF complex formation. For this purpose, various fragments of NELF-A were expressed as GST fusion proteins, immobilized to glutathione Sepharose beads, and incubated with NELF-D or the 581-aa form of TH1 that was prepared by *in vitro* transcription and translation. As shown in Fig. 6, NELF-D bound to the short fragment of NELF-A encompassing aa 125 to 188 as strongly as to full-length NELF-A (lanes 2 and 10). Further deletion from either end of this fragment significantly reduced the A-D interaction (lanes 11 and 12). These results demonstrated that the short sequence of NELF-A (aa 125 to 188) within the HDag homology region is required and sufficient for interaction between NELF-A and NELF-D (or NELF-C). In this experiment, two bands of NELF-D were generated by *in vitro* transcription and translation, but this observation is not reproducible.

A series of C-terminally truncated mutants of NELF-A were constructed so as not to disturb the NELF-C- or NELF-D-binding region (Fig. 7A). These mutant subunits were expressed together with the other subunits in Sf9 cells and purified by anti-Flag affinity chromatography. As expected, Coomassie blue staining (Fig. 7B) and immunoblotting with anti-HA antibody (data not shown) showed that purified complexes contained these mutants with similar stoichiometries.

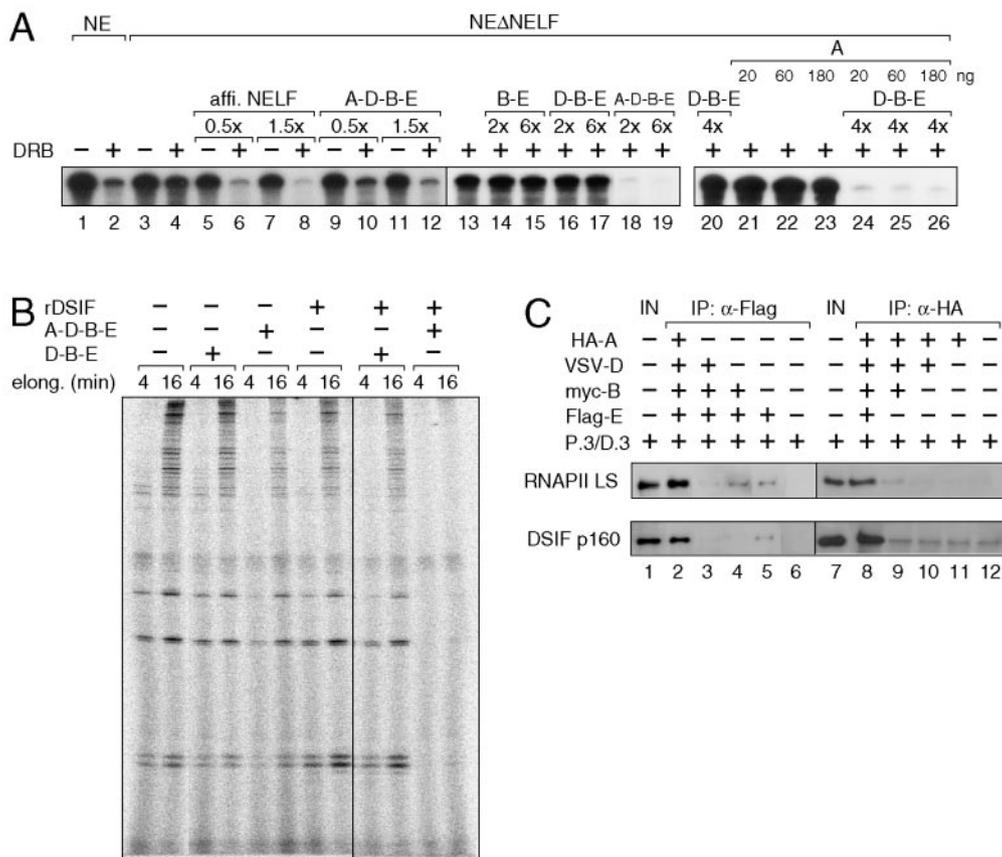


FIG. 5. Functional analyses of recombinant NELF and its subassemblies. (A) Promoter- and factor-dependent transcription with HeLa nuclear extracts (NE). Supercoiled plasmid pTF3-6C<sub>2</sub>AT (25 ng) and either untreated (lanes 1 and 2) or NELF-immunodepleted (NEΔNELF) (lanes 3 to 26) nuclear extracts (2 μl) were incubated in the presence or absence of DRB (50 μM). NELF affinity purified (affi. NELF) from a HeLa cell line constitutively expressing Flag-E, recombinant NELF complexes containing various numbers of subunits, and bacterially expressed NELF-A were added to the reaction mixtures where indicated. The notations 0.5x, 1.5x, 2x, 4x, and 6x indicate the amounts of exogenously added NELF relative to that of endogenous NELF contained in 2-μl untreated nuclear extracts (i.e., 10 ng of NELF-B protein). Full-length (380-nucleotide [nt]) G-free transcripts are shown. (B) Promoter- and factor-independent transcription with RNAPII. dC-tailed template (25 fmol) and purified RNAPII (25 fmol) were incubated. Recombinant DSIF (rDSIF) (p160, 30 ng; p14, 3 ng) and recombinant NELF complexes containing 12 ng of NELF-B were included where indicated. Nucleoside triphosphates were added, and elongation (elong.) was allowed to proceed for 4 and 16 min. Note that short transcripts of <20 nt flowed through the 6% polyacrylamide gel containing 8 M urea. (C) Binding of recombinant NELF to DSIF and RNAPII. Sf9 cells infected with various combinations of the recombinant baculoviruses were lysed and incubated with P.3/D.3, a HeLa nuclear extract-derived fraction rich in DSIF and RNAPII but devoid of NELF. The mixtures were immunoprecipitated (IP) with anti-Flag (α-Flag) (lanes 2 to 6) and anti-HA (α-HA) (lanes 8 to 12) antibodies. The presence of DSIF and RNAPII was determined by immunoblotting with anti-DSIF p160 and anti-RNAPII large subunit (LS) antibodies. Input (IN) lanes represent 2% of P.3/D.3 used for IP.

Thus, the C-terminal deletions of NELF-A did not affect the assembly of the NELF complex. In depletion-add back experiments, NELF complexes containing regions from aa 1 to 427, 1 to 316, and 1 to 248 of NELF-A more or less repressed transcription to the same extent as wild-type NELF, while a complex containing aa 1 to 188 of NELF-A did not (Fig. 7C). Thus, the C-terminal half of NELF-A, including the PST-rich region, could be deleted without a significant loss of NELF activity, whereas further deletion into the HDAG homology region strongly impaired transcriptional repression. Abilities of these mutant NELF complexes to associate with DSIF/RNAPII were then assayed as shown in Fig. 5C. As shown in Fig. 7D, the complexes containing aa 1 to 248 or longer regions of NELF-A bound to DSIF and RNAPII, whereas the complex containing aa 1 to 188 of NELF-A did not. Taken together, these results demonstrated that the short segment of NELF-A

(aa 189 to 248) within the HDAG homology region is important both for transcriptional repression and for interaction with DSIF and RNAPII. The more C-terminal region (aa 249 to 316) of NELF-A may also contribute to the stability of the DSIF/NELF/RNAPII complex.

A previous study reported evidence that NELF binds to the preformed DSIF/RNAPII complex but not to either DSIF or RNAPII alone (40). A possible explanation is that NELF indeed interacts with both DSIF and RNAPII, these binary interactions being very weak individually but stabilized together. Consistent with the previous findings, DSIF and RNAPII behaved similarly with respect to interactions with various NELF mutants (Fig. 5C and 7D). It thus appears that only the stable ternary complex formation, but not the weak binary interactions, could be detected under the experimental conditions. To dissect the multiple roles of NELF, we sought to look at the

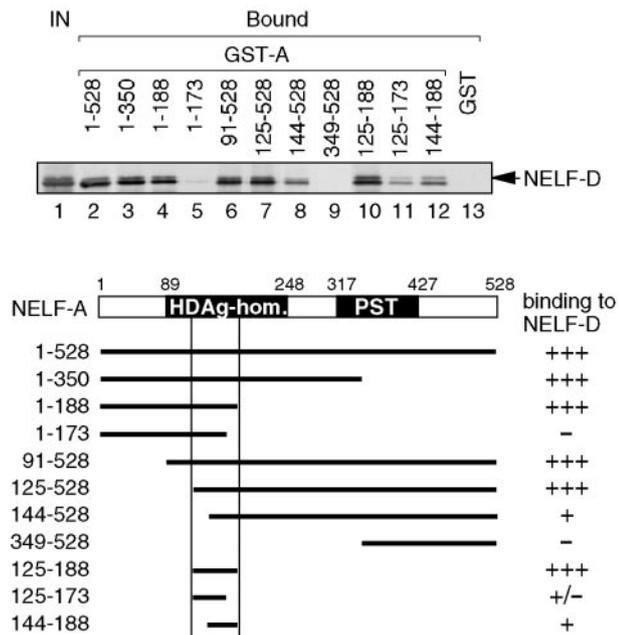


FIG. 6. NELF-A interacts with NELF-C or NELF-D through a short segment in the middle of the HDag homology region. Various fragments of NELF-A were expressed as GST fusion proteins (GST-A) in *E. coli*. These proteins ( $\sim 1 \mu\text{g}$ ) were coupled to glutathione Sepharose and incubated with NELF-D protein that was synthesized and  $^{35}\text{S}$  labeled in vitro. Bound materials as well as 50% of the input (IN) were analyzed by SDS-PAGE and autoradiography. +++, strong (>25%) binding; +, weak (5 to 10%) binding; +/-, very weak (2 to 5%) binding; -, no (<2%) binding. HDag-hom., the region of NELF-A with similarity to HDag; PST, the region rich in proline, serine, and threonine residues.

individual interactions by increasing the sensitivity of binding assays. Each subunit of NELF was expressed as a GST fusion protein, immobilized to glutathione Sepharose beads, and then incubated with P.3/D.3. Analysis of bound materials revealed that RNAPII selectively bound to GST-NELF-A while DSIF bound to none of these proteins (Fig. 8A and data not shown). These results suggested that NELF-A may be responsible for RNAPII binding. Importantly, assays using the C-terminally truncated mutants of NELF-A revealed that the region from aa 189 to 248 is essential for interaction with RNAPII. These results coincide with those shown in Fig. 7 and collectively suggest that the region from aa 189 to 248 of NELF-A binds to RNAPII directly in the context of NELF.

Note that the results shown in Fig. 8A do not imply that isolated NELF-A selectively associates with DSIF-free RNAPII. Usually only  $\sim 10\%$  of RNAPII is in complex with DSIF. If NELF-A associates with RNAPII regardless of the presence of DSIF, the amounts of DSIF coprecipitated with RNAPII would approach background.

## DISCUSSION

**Difference between NELF-C and NELF-D.** We showed here that NELF-C and NELF-D are highly related proteins. Three lines of evidence strongly suggest that NELF is a four- rather than five- subunit complex composed of A, B, either C or D,

and E. First, NELF-C and NELF-D proteins contained in a purified NELF fraction were stained somewhat weakly by silver (Fig. 1A) and by Coomassie blue (data not shown) compared to the other polypeptides. Second, a previous study noted a subtle difference in the elution profiles of NELF-C and NELF-D obtained by mono Q column chromatography of NELF (40). Third and most importantly, Fig. 2C shows that NELF-C and NELF-D are contained in distinct NELF complexes. The reconstitution experiments presented here used only the 581-aa form of TH1, which likely corresponds to NELF-D. Since the reconstituted complex had all the activities associated with native NELF, the structural differences between NELF-C and NELF-D at their N termini may not be functionally important.

The data presented here strongly suggest that NELF-C and NELF-D arise from a common mRNA species by the alternative usage of translation initiation codons. In this regard, it is interesting that the sequence around the second ATG codon has a typical Kozak consensus sequence while the sequence around the first ATG codon does not. NELF-C and NELF-D may also arise from different mRNA species: in GenBank, there are multiple cDNAs and ESTs for TH1 whose 5' ends span around the initiation codon for the 590-aa protein. At present, however, it is not clear whether multiple transcription initiation sites do exist or whether the variability is merely a reflection of incomplete sequence data.

**Possible model for DSIF- and NELF-induced transcriptional pausing.** Based on the available data, we propose a possible model for the DSIF- and NELF-induced transcriptional pausing (Fig. 8B). This model postulates that NELF binds through the different surfaces to DSIF and RNAPII, these binary interactions being very weak individually but stabilized together. Figure 4C demonstrates the three interactions among NELF subunits. Comparison of Fig. 7D and 8A suggests that the small region of NELF-A binds to RNAPII directly. Though we could not see predicted direct interactions between DSIF and individual NELF subunits (Fig. 8A and data not shown), it is possible to assume that two or more subunits of NELF are required for this interaction. With regard to NELF-E, its RNA-binding domain has been shown to be important for transcriptional repression but not for binding to DSIF/RNAPII (40). In addition, we showed here that the deletion of the entire subunit strongly impaired binding to DSIF/RNAPII, suggesting that a region of NELF-E other than the RRM plays an important role in the ternary complex formation.

The sequence alignment of NELF-A and HDag shows that the 60-aa region of NELF-A that is important for RNAPII binding corresponds to the C-terminal 66-aa region of HDag. Surprisingly, this region of HDag exactly matches the region previously determined to be critical for RNAPII binding (38). It is therefore likely that these sequences of NELF-A and HDag form the RNAPII-binding motif conserved in humans and virus. The NELF-C- or NELF-D-binding region of NELF-A was also mapped within the HDag homology region (Fig. 6). However, several lines of evidence suggest that HDag does not directly interact with NELF. Under various conditions with GST pull-down assays (38), immunoprecipitations (unpublished data), and yeast two-hybrid interaction assays (unpublished data), we were unable to detect direct interaction

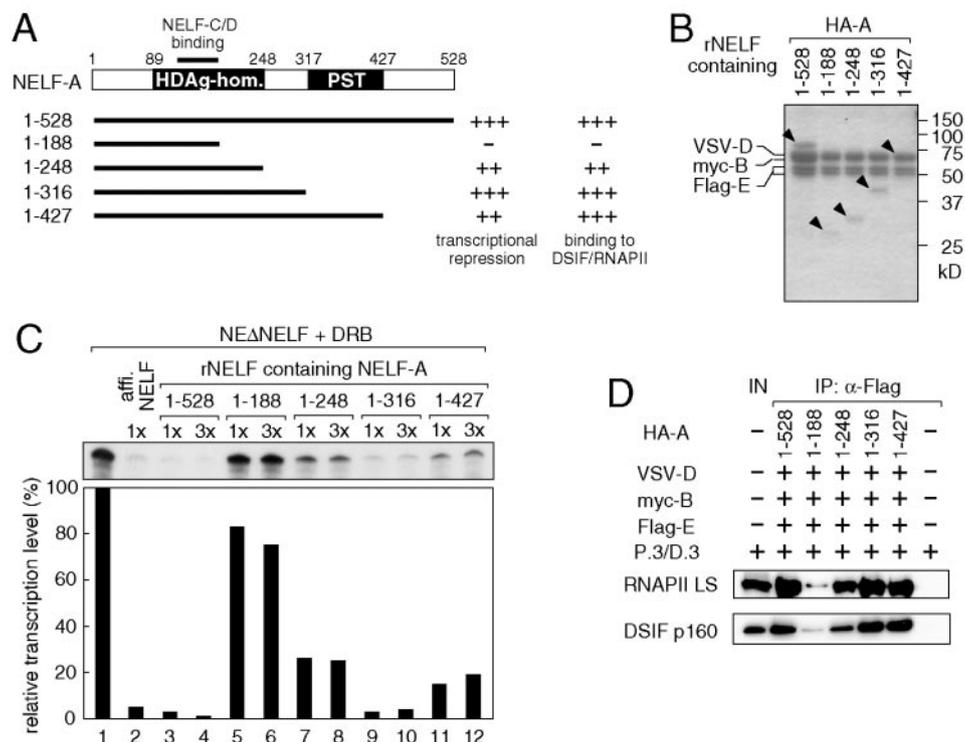


FIG. 7. A part of the NELF-A region with similarity to HDAG (HDAG-hom.) is important for RNAPII binding and for transcriptional repression. (A) C-terminally truncated mutants of NELF-A used in this study. The region responsible for NELF-C or NELF-D binding is indicated at the top. A summary of the results presented in panels C and D is shown to the right. +++, comparable to wild-type activity; ++, slightly weaker than wild-type activity; -, significantly weaker than wild-type activity. PST, the region rich in proline, serine, and threonine residues. (B) Recombinant NELF (rNELF) complexes containing wild-type and mutant NELF-As were purified by anti-Flag affinity chromatography and analyzed by SDS-12.5% PAGE and Coomassie blue staining. Arrowheads indicate the positions of wild-type and mutant NELF-As. (C) Transcriptional repression by the rNELF mutants. pTF3-6C<sub>2</sub>AT (25 ng) and NELF-depleted nuclear extracts (NEΔNELF; 2 μl) were incubated in the presence of DRB (50 μM). NELF affinity purified (affi. NELF) from a HeLa cell line constitutively expressing Flag-E (lane 2) and rNELF complexes containing wild-type and mutant NELF-As (lanes 3 to 12) were included in the reaction mixtures where indicated. The notations 1× and 3× indicate the amounts of exogenously added NELF relative to that of endogenous NELF contained in 2-μl untreated nuclear extracts (i.e., 10 ng of NELF-B protein). The counts of 380-nt G-free transcripts were quantified by using the image analyzer STORM (Molecular Dynamics). Transcription levels relative to that of lane 1 (the reaction without NELF) are shown at the bottom. (D) Binding of the rNELF mutants to DSIF and RNAPII. Binding assays were carried out as described in the legend for Fig. 5C. IN, input; IP, immunoprecipitate; α-Flag, anti-Flag antibody; LS, large subunit.

between these proteins. Instead, the region of HDAG corresponding to the NELF-C- or NELF-D-binding region of NELF-A has an RNA-binding function (16). It is interesting that the arginine-rich motif-type RNA-binding domain of HDAG and the RRM-type RNA-binding domain of NELF-E may have a related function during transcriptional elongation (Fig. 8B).

**Evolutionary conservation of NELF.** Another important finding of this study is that NELF appears to be conserved only in a subset of metazoans including vertebrates and insects. Using human NELF subunits as query sequences, we identified putative orthologs of A, B, C, D, and E in flies. However, extensive searches against yeast, nematode, and plant genomes failed to identify any entry with even partial (but significant) similarity to the NELF subunits both at nucleotide and protein levels. The only exceptions were those belonging to the large RRM family, which showed limited similarity to the NELF-E RRM. Thus, NELF may have evolved to control the biological processes unique to vertebrates and insects. In this regard, two recent findings are intriguing. First, human NELF-A is en-

coded by *WHSC2*, a candidate gene for Wolf-Hirschhorn syndrome (35). This genetic disease causes a wide range of developmental defects, including some in the brain that lead to mental retardation. Second, mutations of the *Spt5* gene in zebra fish cause multiple developmental defects, including discrete problems with neuronal and cardiac differentiation (6, 12).

We were surprised not to find orthologs of NELF in yeast since potential orthologs of DSIF and P-TEFb are found in yeast. DSIF, a heterodimer composed of Spt4 and Spt5 subunits, is widely conserved across species (1, 6, 7, 10, 33). Spt4/Spt5, originally identified in a genetic screen of yeast, has been implicated in transcriptional control in various species, including yeast, flies, fish, and humans (39). P-TEFb has been studied with fly and human cells. Since P-TEFb belongs to the large Cdk-cyclin family, it is difficult to find orthologs of P-TEFb based solely on the sequence information. However, recent genetic and biochemical evidence from yeast suggests two Cdk-containing complexes, Ctk1/Ctk2/Ctk3 and Bur1/Bur2, as possible counterparts of P-TEFb (8, 21). If there is no NELF,

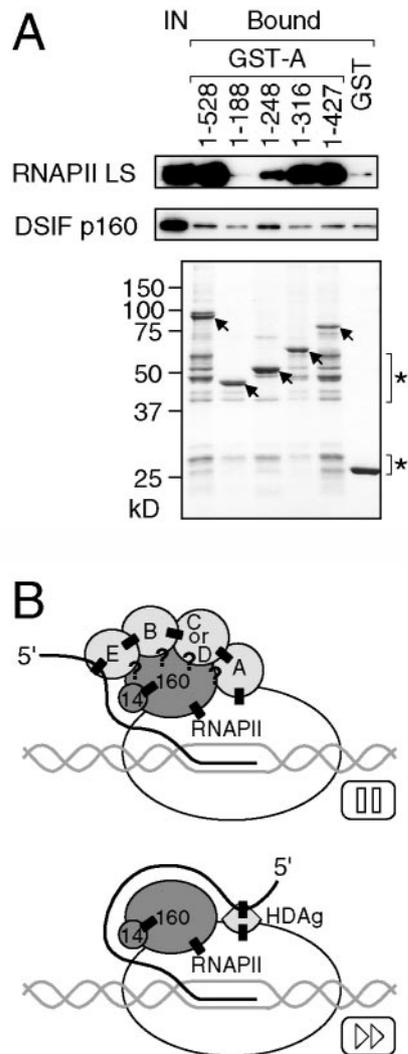


FIG. 8. Evidence that NELF-A binds directly to RNAPII. (A) The same NELF-A mutant proteins used in the analyses described in Fig. 7 were expressed as GST fusion proteins (GST-As). These proteins (~1  $\mu$ g) were coupled to glutathione Sepharose and incubated with P.3/D.3 as described in Materials and Methods. Bound materials as well as 2% of the input (IN) were analyzed by SDS-PAGE and either immunoblotting (top) or Coomassie blue staining (bottom). Arrows indicate the expected sizes of GST-As. Most of the smaller polypeptides visible on the gel (indicated by asterisks) originated from bacterial lysates and probably represent the C-terminally truncated forms of GST-As. LS, large subunit. (B) Possible model for DSIF- and NELF-induced transcriptional pausing (top). As a comparison, a model for the HDAG-activated elongation is shown at the bottom. See the text for details. Subunits of NELF and DSIF are shown in light gray and dark gray, respectively. Filled boxes indicate experimentally proven interactions, while question marks indicate hypothetical interactions. Interactions of DSIF p160/Spt5 with DSIF p14/Spt4 and RNAPII were demonstrated previously (9, 36). RNA bindings by NELF-E and HDAG were shown previously (16, 40). RNAPII binding by HDAG was also shown previously (38).

DSIF may act solely as an activator of transcriptional elongation (3, 33). A previous study with yeast showed that cold-sensitive *spt5* mutants show 6-azauracil (6AU)-dependent cell growth at a nonpermissive temperature while strains containing other *spt5* alleles grow in a 6AU-sensitive manner (7). 6AU

is known to reduce the efficiency of transcriptional elongation by decreasing the intracellular levels of nucleotides, and indeed, mutations of genes encoding TFIIS and RNAPII often cause a 6AU-sensitive phenotype (32). Thus, these observations suggest that yeast Spt5 may also have dual roles in transcriptional elongation. In this scenario, Spt4/Spt5 may be capable of inducing transcriptional pausing without NELF in yeast (and perhaps in nematodes and plants). Alternatively, these organisms may have a protein factor that is functionally similar to, but structurally different from, NELF.

#### ACKNOWLEDGMENTS

We thank David Gilmour for helpful discussions and comments on the manuscript. We also thank Sophie Delehouzee for manuscript preparation and Keiko Watanabe and Junko Kato for technical assistance.

This work was supported in part by a Grant for Research and Development Projects in Cooperation with Academic Institutions from the New Energy and Industrial Technology Development Organization to H.H. T.N. is a JSPS research fellow.

T.N. and Y.Y. contributed equally to this work.

#### REFERENCES

- Andrulis, E. D., E. Guzman, P. Doring, J. Werner, and J. T. Lis. 2000. High-resolution localization of *Drosophila* Spt5 and Spt6 at heat shock genes in vivo: roles in promoter proximal pausing and transcription elongation. *Genes Dev.* **14**:2635–2649.
- Bonthron, D. T., B. E. Hayward, V. Moran, and L. Strain. 2000. Characterization of TH1 and CTSZ, two non-imprinted genes downstream of GNAS1 in chromosome 20q13. *Hum. Genet.* **107**:165–175.
- Bourgeois, C. F., Y. K. Kim, M. J. Churcher, M. J. West, and J. Karn. 2002. Spt5 cooperates with human immunodeficiency virus type 1 Tat by preventing premature RNA release at terminator sequences. *Mol. Cell. Biol.* **22**:1079–1093.
- Chao, S. H., K. Fujinaga, J. E. Marion, R. Taube, E. A. Sausville, A. M. Senderowicz, B. M. Peterlin, and D. H. Price. 2000. Flavopiridol inhibits P-TEFb and blocks HIV-1 replication. *J. Biol. Chem.* **275**:28345–28348.
- Garber, M. E., and K. A. Jones. 1999. HIV-1 Tat: coping with negative elongation factors. *Curr. Opin. Immunol.* **11**:460–465.
- Guo, S., Y. Yamaguchi, S. Schilbach, T. Wada, J. Lee, A. Goddard, D. French, H. Handa, and A. Rosenthal. 2000. A regulator of transcriptional elongation controls vertebrate neuronal development. *Nature* **408**:366–369.
- Hartzog, G. A., T. Wada, H. Handa, and F. Winston. 1998. Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in *Saccharomyces cerevisiae*. *Genes Dev.* **12**:357–369.
- Hautbergue, G., and V. Goguel. 2001. Activation of the cyclin-dependent kinase CTDK-I requires the heterodimerization of two unstable subunits. *J. Biol. Chem.* **276**:8005–8013.
- Ivanov, D., Y. T. Kwak, J. Guo, and R. B. Gaynor. 2000. Domains in the SPT5 protein that modulate its transcriptional regulatory properties. *Mol. Cell. Biol.* **20**:2970–2983.
- Kaplan, C. D., J. R. Morris, C. Wu, and F. Winston. 2000. Spt5 and Spt6 are associated with active transcription and have characteristics of general elongation factors in *D. melanogaster*. *Genes Dev.* **14**:2623–2634.
- Karn, J. 1999. Tackling Tat. *J. Mol. Biol.* **293**:235–254.
- Keegan, B. R., J. L. Feldman, D. H. Lee, D. S. Koos, R. K. Ho, D. Y. Stainier, and D. Yelon. 2002. The elongation factors Pandora/Spt6 and Foggy/Spt5 promote transcription in the zebrafish embryo. *Development* **129**:1623–1632.
- Kim, J. B., Y. Yamaguchi, T. Wada, H. Handa, and P. A. Sharp. 1999. Tat-SF1 protein associates with RAP30 and human SPT5 proteins. *Mol. Cell. Biol.* **19**:5960–5968.
- Krogan, N. J., M. Kim, S. H. Ahn, G. Zhong, M. S. Kobar, G. Cagney, A. Emili, A. Shilatifard, S. Buratowski, and J. F. Greenblatt. 2002. RNA polymerase II elongation factors of *Saccharomyces cerevisiae*: a target proteomics approach. *Mol. Cell. Biol.* **22**:6979–6992.
- Kyrpides, N. C., C. R. Woese, and C. A. Ouzounis. 1996. KOW: a novel motif linking a bacterial transcription factor with ribosomal proteins. *Trends Biochem. Sci.* **21**:425–426.
- Lai, M. M. 1995. The molecular biology of hepatitis delta virus. *Annu. Rev. Biochem.* **64**:259–286.
- Li, J., R. Horwitz, S. McCracken, and J. Greenblatt. 1992. NusG, a new *Escherichia coli* elongation factor involved in transcriptional antitermination by the N protein of phage lambda. *J. Biol. Chem.* **267**:6012–6019.
- Lis, J. T., P. Mason, J. Peng, D. H. Price, and J. Werner. 2000. P-TEFb kinase recruitment and function at heat shock loci. *Genes Dev.* **14**:792–803.

19. Mancebo, H. S., G. Lee, J. Flygare, J. Tomassini, P. Luu, Y. Zhu, J. Peng, C. Blau, D. Hazuda, D. Price, and O. Flores. 1997. P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro. *Genes Dev.* **11**:2633–2644.
20. Mueller, C. L., and J. A. Jaehning. 2002. Ctr9, Rtf1, and Leo1 are components of the Paf1/RNA polymerase II complex. *Mol. Cell. Biol.* **22**:1971–1980.
21. Murray, S., R. Udupa, S. Yao, G. Hartzog, and G. Prelich. 2001. Phosphorylation of the RNA polymerase II carboxy-terminal domain by the Bur1 cyclin-dependent kinase. *Mol. Cell. Biol.* **21**:4089–4096.
22. Nguyen, V. T., T. Kiss, A. A. Michels, and O. Bensaude. 2001. 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature* **414**:322–325.
23. Parada, C. A., and R. G. Roeder. 1999. A novel RNA polymerase II-containing complex potentiates Tat-enhanced HIV-1 transcription. *EMBO J.* **18**:3688–3701.
24. Ping, Y. H., and T. M. Rana. 2001. DSIF and NELF interact with RNA polymerase II elongation complex and HIV-1 Tat stimulates P-TEFb-mediated phosphorylation of RNA polymerase II and DSIF during transcription elongation. *J. Biol. Chem.* **276**:12951–12958.
25. Ponting, C. P. 2002. Novel domains and orthologues of eukaryotic transcription elongation factors. *Nucleic Acids Res.* **30**:3643–3652.
26. Price, D. H. 2000. P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol. Cell. Biol.* **20**:2629–2634.
27. Renner, D. B., Y. Yamaguchi, T. Wada, H. Handa, and D. H. Price. 2001. A highly purified RNA polymerase II elongation control system. *J. Biol. Chem.* **276**:42601–42609.
28. Squazzo, S. L., P. J. Costa, D. L. Lindstorm, K. E. Kumer, R. Simic, J. L. Jennings, A. J. Link, K. M. Arndt, and G. A. Hartzog. 2002. The Paf1 complex physically and functionally associates with transcription elongation factors in vivo. *EMBO J.* **21**:1764–1774.
29. Sullivan, S. L., and M. E. Gottesman. 1992. Requirement for E. coli NusG protein in factor-dependent transcription termination. *Cell* **68**:989–994.
30. Tang, H. K., S. Singh, and G. F. Saunders. 1998. Dissection of the transactivation function of the transcription factor encoded by the eye developmental gene PAX6. *J. Biol. Chem.* **273**:7210–7221.
31. Thirunavukkarasu, K., M. Mahajan, K. W. McLarren, S. Stifani, and G. Karsenty. 1998. Two domains unique to osteoblast-specific transcription factor Osf2/Cbfa1 contribute to its transactivation function and its inability to heterodimerize with Cbfb. *Mol. Cell. Biol.* **18**:4197–4208.
32. Uptain, S. M., C. M. Kane, and M. J. Chamberlin. 1997. Basic mechanisms of transcript elongation and its regulation. *Annu. Rev. Biochem.* **66**:117–172.
33. Wada, T., T. Takagi, Y. Yamaguchi, A. Ferdous, T. Imai, S. Hirose, S. Sugimoto, K. Yano, G. A. Hartzog, F. Winston, S. Buratowski, and H. Handa. 1998. DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev.* **12**:343–356.
34. Wada, T., T. Takagi, Y. Yamaguchi, D. Watanabe, and H. Handa. 1998. Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription *in vitro*. *EMBO J.* **17**:7395–7403.
35. Wright, T. J., J. L. Costa, C. Naranjo, P. Francis-West, and M. R. Altherr. 1999. Comparative analysis of a novel gene from the Wolf-Hirschhorn/Pitt-Rogers-Danks syndrome critical region. *Genomics* **59**:203–212.
36. Yamaguchi, Y., T. Wada, D. Watanabe, T. Takagi, J. Hasegawa, and H. Handa. 1999. Structure and function of the human transcription elongation factor DSIF. *J. Biol. Chem.* **274**:8085–8092.
37. Yamaguchi, Y., T. Takagi, T. Wada, K. Yano, A. Furuya, S. Sugimoto, J. Hasegawa, and H. Handa. 1999. NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* **97**:41–51.
38. Yamaguchi, Y., J. Filipovska, K. Yano, A. Furuya, N. Inukai, T. Narita, T. Wada, S. Sugimoto, M. M. Konarska, and H. Handa. 2001. Stimulation of RNA polymerase II elongation by hepatitis delta antigen. *Science* **293**:124–127.
39. Yamaguchi, Y., T. Narita, N. Inukai, T. Wada, and H. Handa. 2001. SPT genes: key players in the regulation of transcription, chromatin structure and other cellular processes. *J. Biochem.* **129**:185–191.
40. Yamaguchi, Y., N. Inukai, T. Narita, T. Wada, and H. Handa. 2002. Evidence that negative elongation factor represses transcription elongation through binding to a DRB sensitivity-inducing factor/RNA polymerase II complex and RNA. *Mol. Cell. Biol.* **22**:2918–2927.
41. Yang, Z., Q. Zhu, K. Luo, and Q. Zhou. 2001. The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* **414**:317–322.
42. Ye, Q., Y.-F. Hu, H. Zhong, A. C. Nye, A. S. Belmont, and R. Li. 2001. BRCA1-induced large-scale chromatin unfolding and allele-specific effects of cancer-predisposing mutations. *J. Cell Biol.* **155**:911–921.
43. Zhu, Y., T. Pe'ery, J. Peng, Y. Ramanathan, N. Marshall, T. Marshall, B. Amendt, M. B. Mathews, and D. H. Price. 1997. Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro. *Genes Dev.* **11**:2622–2632.
44. Zorio, D. A. R., and D. L. Bentley. 2001. Transcription elongation: the 'Foggy' is lifting. *Curr. Biol.* **11**:R144–R146.